Investigation of the clinical utility of two potential pro-oncogenic genes in prostate cancer and breast cancer

LUISA BARBATO

A Thesis submitted in Partial Fulfilment of the Requirements of Nottingham Trent University for the degree of Doctor of Philosophy.

Acknowledgements

This thesis is a result of four years of experiments and experiences. It has been a long and beautiful journey with a lot of obstacles, but I have been packed with the best company I could ever ask for and they all deserved to be thanked. I want to start by thanking my first director of the studies Dr Tarik Regad that selected me for the PhD program because he let me know the best friends and the most expert academics at John van Geest.

I have never been so inspired, helped and guided as I have been during the most difficult moments of my PhD. Dr David Boocock took the strain as my new director of the studies and his knowledge, and sincere help was vital during this period. He is the best supervisor that every PhD could ask for. There are not enough words to describe how I am grateful and honoured to be his student.

I have been granted the most wonderful supervisor team: To Dr Jayakumar Vadakekolathu a safe harbour for all PhD students an incredible and knowledgeable scientist, thank you for your patience and tremendous support during this adventure. Professor Graham Pockley my temporary director of the studies and then second supervisor is a fantastic scientist and thoughtful person. I still have in my mind his words: “Science is funny” and I was reminding myself during the hard moments when things were not going as I expected and pushed me to give my best. Thank you for being so helpful and always giving the most accurate feedback on everything.

My time at John van Geest has been priceless thanks to the people working in there. Thanks to Anne a wonderful colleague and person, even if the time we spent together was limited, I felt so close to her with her words and a big help, I have learned the importance to stand up for myself. Thank you to Steve for the best time during my lab work talking about science and the Beatles surrounded by great music from Smooth Radio station (I am still listening to that) and my mind goes to those moments. Thank you both for making me feel appreciated and for all the support you gave to me and all students during our laboratory time and the patience you had, is fair to say you are the angels of Van Geest. Thanks to Murrium to be always there for help in the most difficult moments and always there for a chat with her incredible kindness and warm words. Thanks to all the academics working at Van Geest with their incredible expertise provided to all of us, students.

Thanks to Gemma Foulds her name couldn’t be more appropriate as in Italian means “precious stone”, the goodness of flowcytometry but also a wonderful and the most courageous person I have met. Thanks for your time spent together during the flow running. I am so happy to know you and I hope to see you one day in Italy. Thanks to Clare Convey to help me during my last experiments and always and be so nice with all of us, I am going to miss you. Thanks to all the PhD colleagues and friends Nikki, Elena, Marco, Anna, Josh, and Harish for sharing those years I will miss you all. To Arif Surani a passionate scientist, to be present during my lab time with his pure friendship, help, kindness and with the most contagious laugh to make me smile even in the bad time, I know you are already doing great things and I am sure you will keep do it. I would like to also thank my colleagues at the MTIF Dr Elena Hunt and Dr Philippe Wilson for allowing me to work for them and feel welcome in their group. I will miss them a lot. I hope to still meet them one day.

A big thanks go to what I consider my stepfamily, we have been through together to the hardest time. To Tania, Jenny and Melissa I think I wouldn’t be able to cope with the last months without you. No words are enough to describe how lucky I felt to meet you all. To be patient with my Italian accent but mostly to be there when I was falling you gave me your hands to stand up. Thank you so much it can’t be forgotten.

I have been lucky to see different places in the UK and in all of them I have left part of my heart and memories which are in every single street I have walked; Nottingham is also one of those. I have met wonderful people and I have learned from them something. To Elisabetta, Giulia, Marco, and Davide for their friendship and laughs thanks for the beautiful evening spent together and thanks for the support showed me.

To my best friends Federica and Cristina our long friendship is still strong and despite being far with the body we are even closer with the heart year after year. You have been next to me always no matter where I was, and I hope that this friendship will remain forever. Thank you to those gorgeous ladies to be in my life.

I have never thought that in this long travel I would have been also so lucky to meet the most intelligent, funny, and kind person, Carmine. You easily came into my life and just changed for the better with your smile and laughter. You helped me during the most difficult times, and you made me feel protected but at the same time strong and independent, qualities I thought lost for a while. We supported each other and started a beautiful walk together I hope will never stop. I am not sure if you know how important you are to me. Furthermore, you made me the honour to meet your wonderful family who strongly deserves to be mentioned in my acknowledgement as they have been part of this experience too. Thank you to Mrs Luisa Salvati and Ms Vincenzo Napoli to make me part of your daily video calls and making me feel less alone. Thank you for the care showed me.

Last but not least, my family without them I wouldn’t be here in the first place. Thanks to be next to me in every step of my life with unconditional love and support. Thanks to have encouraged the dream of a daughter to travel to the UK, to study in a different country for so much time far from a loved house. You accepted it even if it was hard, believe me, it was hard for me too, but somehow you knew I had to do it. Thank you to be with me on the phone all night long until I was ready to sleep and being next to me when I felt sad. Finally, thank you for teaching me to be a respectful and responsible person and showing me that hardworking always pays back. I would not be the person I am proud of today without you, dearest dad, mum, and sister thank you I love you the most.

To all thank you, each of you taught me something precious and will be forever in me.

Grazie di cuore

“Fix your course on a star and you will navigate any storm”
“Traccia la tua rotta verso una stella e supererai qualsiasi tempesta”
Leonardo da Vinci

Table of Contents

List of Figures

List of Tables

Abstract

The Identification of novel and specific biomarkers is crucial to diagnosis, and prognosis, in patients with prostate and breast cancer. Because cancer therapies have side effects in patients, discovering and potentially targeting specific biomarkers could promote the use of personalised approach for a more effective treatment.

Firstly, we have focused on the development of a monoclonal antibody-drug-based therapy, targeting prostate cancer stem cells (PCSCs), using a monoclonal antibody (mAb) previously generated in our laboratory against human endothelial protein C receptor (EPCR). PCSCs were isolated using lentivirus expressing the enhanced green fluorescent protein (EGFP) under NANOG-promoter generating two populations *NANOG-EGFP+* and *NANOG-EGFP-* and analysed for EPCR expression.

No significant difference was observed in the expression of EPCR between *NANOG-EGFP+* and *NANOG-EGFP-* cell populations. A lack of conclusive correlation was observed between EPCR deficient cells with epithelial-mesenchymal transition (EMT) markers, cancer stem cells (CSCs), and stem cell markers. Finally, Gene Expression Profiling Interactive Analysis (GEPIA) was used to look at the tissue expression in normal and tumour tissue, showing high expression of EPCR in endothelial cells. Finally, based tissue expression profiling, EPCR is not a suitable candidate for antibody targeting as it would lead to off-target effects in multiple tissues, therefore no further experiments were designed using EPCR as a target biomarker.

Following this, a feasible study on the effect of Sperm-Associated Antigen 5 (SPAG5) chemoresistance and cancer progression in prostate and breast cancer was performed. The transcriptome and proteome of SPAG5 deficient were investigated in triple-negative breast cancer (TBC) MDA-MB-231 and androgen-independent prostate cancer DU145 cell lines, by RNA-sequencing and mass spectrometry (MS) analysis. Transcriptome was performed and a total of 2,201 differentially expressed genes (DEGs) in MDA-MB-231 SPAG5 deficient cells, while 907 DEGs DU145 SPAG5 deficient cells, versus control empty vector pLKO.1 cells, were identified. No significant differences in the cell cycle were observed in Doxorubicin and Epirubicin treatment DU145 and MDA-MB-231 SPAG5 deficient cells versus controls.

A list of the most statistically significant genes upregulated and downregulated was taken forward for verification for common and unique pathways, through free available online resources such as METASCAPE, and Kyoto Encyclopaedia of Genes and Genomes (KEGG) pathway and Gene Ontology (GO). Using *StatsPro* free online sources proteomics analysis generated 230 differentially expressed proteins (DEPs) in MDA-MB-231 SPAG5 deficient cells and 65 DEPs DU145 SPAG5 deficient versus control cells. Protein-protein interaction (PPI) network using Cytoscape has been conducted for enrichment KEGG analysis.

Cross-over data from MS and RNAseq upregulated and downregulated genes in MDA-MB-231 and DU145 SPAG5 deficient were compared to *in silico* data from cBioPortal tool. Interestingly, positive correlation was observed in genes involved in cell cycle, but also in genes involved in catalyse and biosynthesis of cholesterol.

Collectively those data offer a wider insight into the association of SPAG5 in cancer progression and its potential role not only in pathways involved in cell cycle but also how in lipid metabolism in cancer.

1. Introduction

1.1 Cancer

The term neoplasia indicates an abnormal growth of the tissue. These abnormalities can be benign (noncancerous) or malignant. In benign neoplasia, the tumour grows slowly but doesn't spread while in malignant neoplasia it grows rapidly and spreads around the body. Cancer is a disease caused by uncontrolled cell growth that can happen in almost all the tissues and organs of the body. Particularly, the World Health Organisation sponsored a symposium in 1950 where the dramatic variations different types of cancer in a different part of the body (Kesharwani et al., 2019) was discussed. It was interesting to see that people who migrated to other countries developed a type of cancer, mainly present in that country rather than develop types of cancer present in the homeland. That led to considering the environment was also a possible cause of most cancers (Shimkin, 1977). Molecular modifications are involved in the origin of cancer, leading to the cell on uncontrolled cell division (Adjiri, 2016).

Genetic modification contributes to cancer growth altering normal cell function. Modification can occur during cell division or environmental exposure that can damage the DNA. Tobacco smoke, radiation including the ultraviolet B rays from the sun are responsible for DNA damages (Jacob et al., 2018) (Li, Y. & Ma, 2020).

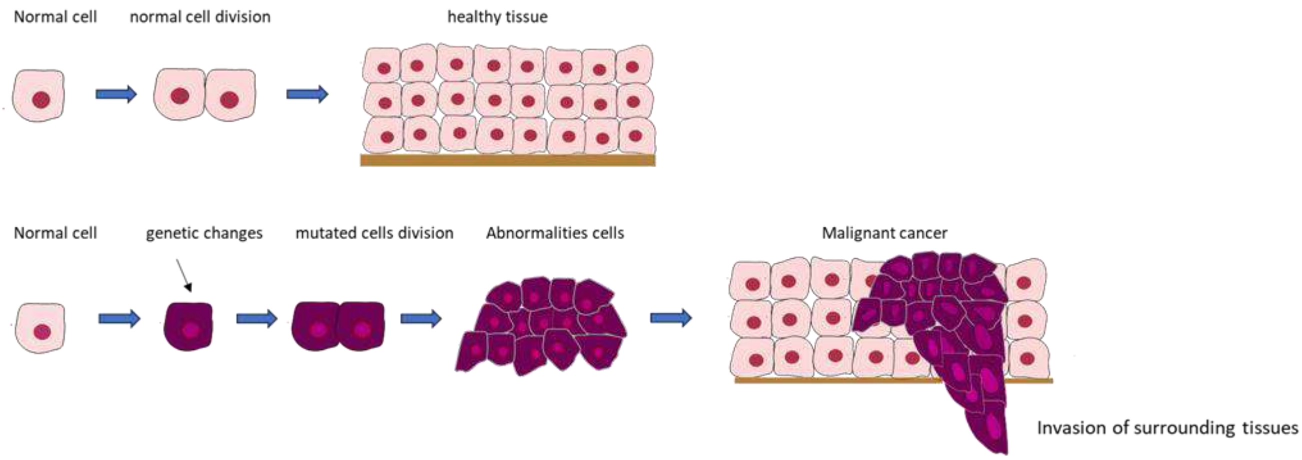


Figure 1.1 – Development normal cells to cancerous cells. Schematic representation of normal cell versus development of tumour: genetic mutations can lead to progressive cells abnormalities. Malignant abnormalities may become cancerous, and cells can rapidly grow and invade surrounding tissues.

1.2.1 Cancer incidence

In 2020, there were nearly 10 million of deaths recorded worldwide, with breast and lung cancer the most diagnosed and lung cancer the leading cause of death (1.8 million) (Piñeros et al., 2021a). Incidence and the risks related to cancer increase with age (Piñeros et al., 2021b).

1.2.2 Cancer worldwide

Cancer represents the leading cause of death in people aged in their 70s in 112 countries out of 183 while it is the third and fourth cause of death in 23 countries (Fig.1.2). The population increasing, also in the distribution of the risk factors, associated with a different socio-economic development are responsible of cancer growth and mortality (Sung et al., 2021a).

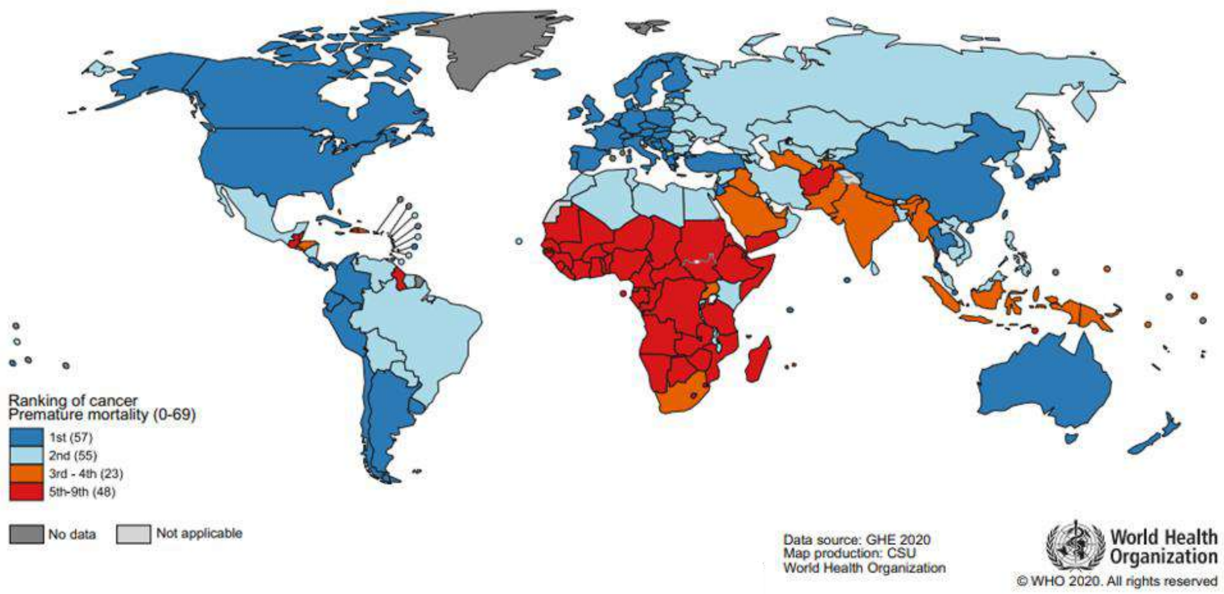


Figure 1.2 - Age associated mortality of cancer worldwide. The map is an estimation of age cancer related death in 2019 in people <70y old. (Data source: GHE2020, map production: CSU; *(Sung et al., 2021b)*.

1.2.2 Cancer incidence in the United Kingdom

In 2020 data from GLOBOCAN has registered 457,960 new cancer cases in the UK with 53% incidence in men and 46% in women. Excluding the non-melanoma skin cancer in males, the most common is prostate (23.1%) while in females breast cancer became the most diagnosed (25.5%) in the UK and 11.7 % worldwide surpassing lung cancer (Fig 1.3) (WHO, 2021b).

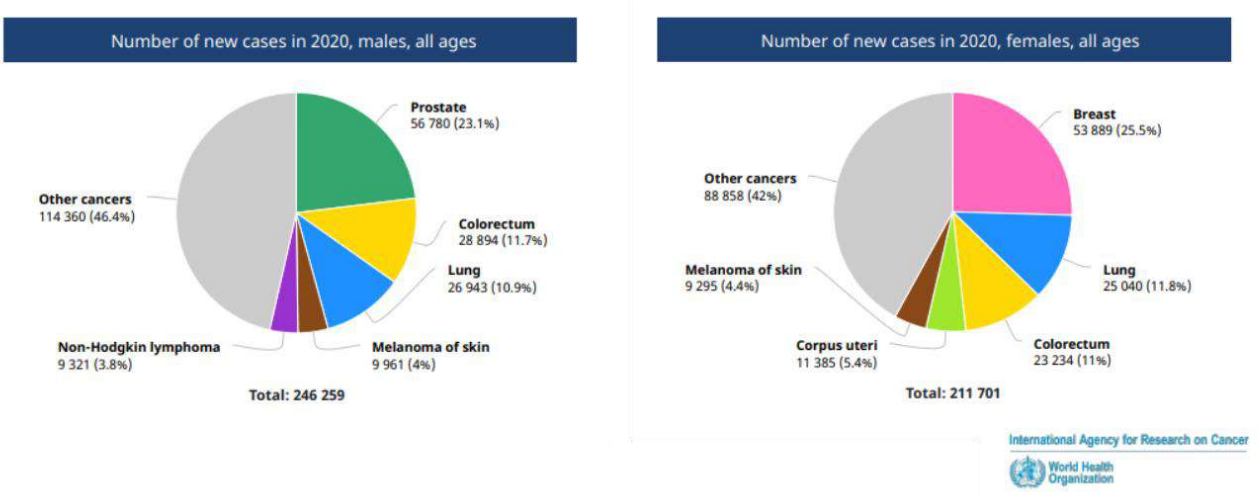


Figure 1.3 - Incidence of different types of cancer in males and female in UK 2020. Pie graphs represent the percentage of different types of cancer based on the most recent data analysis in males and females in UK (Data source: GLOBOCAN 2020, pie map) *(WHO, 2021a)* .

1.3 Carcinogenesis

Among the risk factors, virus, which has been known since the 1980s, were discovered as the cause of cancers. The International Agency for Research on Cancer (IARC) reported that the human papilloma virus, Epstein-Barr virus, and human immunodeficiency virus (HIV) and human herpesvirus are involved in carcinogenesis (International Agency for Research on Cancer, 2012) .

A study conducted on pharmaceutical compounds were identified as responsible for carcinogenesis. In the 1960s study conducted on analgesics demonstrated that high concentrations were responsible for renal cancers (Bengtsson et al., 1968). Researchers focused also on the use of some warfare agents. Particularly their evidence suggested that the use of chemical compounds such as the mustard sulfide that share analogies with a series of nitrogen analogues bis, and tris (β-chloroethyl)-amines, were responsible for the cytotoxic activity leading to the rapidly proliferating tissue in lymph nodes, involving bone marrow and epithelium in the gastrointestinal tract (Gilman & Philips, 1946). The dangers due to radioactivity were already documented. Marie Currie-Skłodowska, despite the enormous contribution in science and medicine on the study of radioactive with her husband Pierre died from aplastic anaemia due to her long exposure to radiation. Even the radioisotope used in medical practice is responsible for cancers (Anna Gasinska, 2016). Many patients that were treated with the radioactive phosphorus (phosphate (3-) ion) used to treat polycythaemia (high production of red blood cells from the bone marrow), also showed a high risk of leukaemia (Osgood, 1964).

1.3. Hallmarks of cancer

In 2001, Hanahan and Weinberg published an interesting review describing the complexity of neoplastic disease as six biological capabilities that cancer acquires during the development of the human tumours (Hanahan & Weinberg, 2000). In 2011, a second review published from Hanahan and Weinberg update the existent featured adding two more to the six already described (Hanahan & Weinberg, 2011a). The eight hallmarks identified include the capabilities for sustaining proliferative signalling, evading growth suppressors, resisting cell death, enabling replicative immortality, inducing/accessing vasculature, activating invasion and metastasis, reprogramming cellular metabolism, and avoiding immune destruction. However, already in 2010 Lazebenik argued that all the hallmarks except for the invasion/metastasis are involved in both benign and malignant neoplasia even if cancer is already referred to a malignant disease (Lazebnik, 2010). Finally, in 2017 Foaud and Aanai suggested that cancer hallmarks could be considered as specific features that cells acquired along with their life leading to cell transformation, and a malignant progression using surrounding tissues for their growth (Fouad & Aanei, 2017).

In 2022, Hanahan published an interesting follow-up review in which they stated that the cellular metabolism and the avoiding immune destruction were much like the rest of the six hallmarks and therefore should be considered as core hallmarks of cancers. The review also suggests the incorporation with four new hallmarks (i) unlocking phenotypic plasticity, (ii) Nonmutational epigenetic reprogramming, (iii) polymorphic microbiomes, and (iv) senescent cells considered as core components of the hallmark cancer characterisation (Hanahan, 2022a).

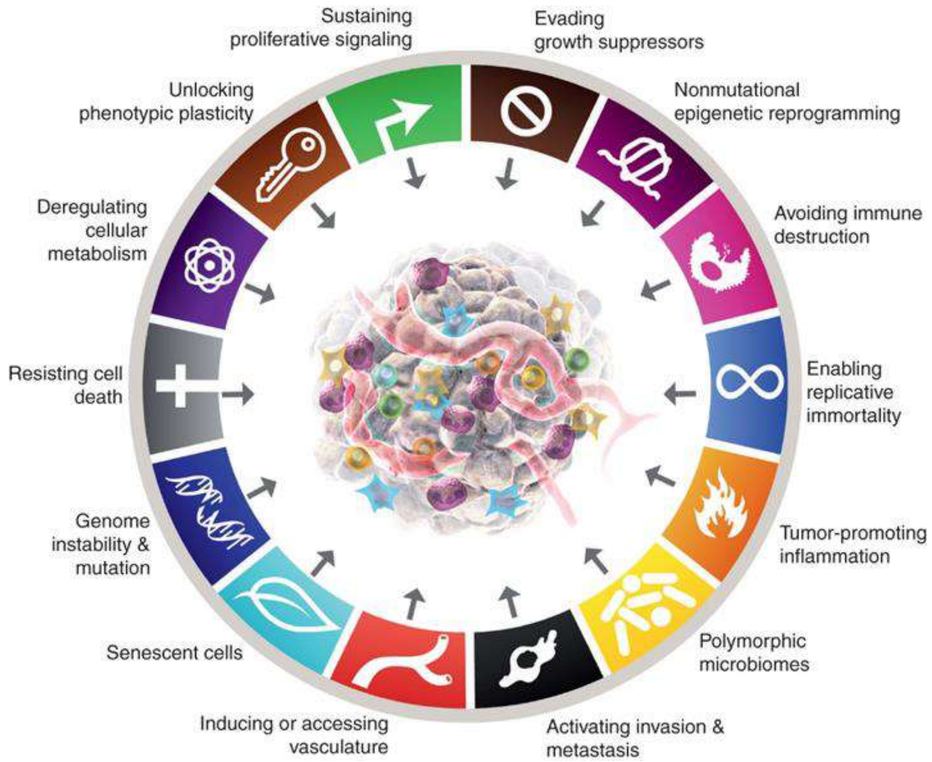


Figure 1.4 - 'Hallmarks of cancer'. The Hallmarks of cancer currently proposing eight hallmarks' capabilities and two characteristics. Those were added to the previously explained in Hallmarks of cancer 2000. Picture adapted from *(Hanahan, 2022b)*.

1.3.1 Sustaining proliferative signalling

Normal tissue controls the production and the release of specific growth-promoting signals that ensure the cell progress and cell number homeostasis, particularly in cancer those signals represent an important event in its development. This is due to an alteration in the mechanism of activation of the cell cycle in specific proteins. The normal barriers to proliferation are responsible to regulate the low oxygen tension, low glucose level and an acidic extracellular pH level that can create a hostile environment (Feitelson et al., 2015). In several solid tumours, hypoxia is a hallmark for tumour microenvironment, due to imbalance of oxygen because of an increase of oxygen request and inadequate supply as a consequence of the rapid cells proliferation (McKeown, 2014). To deal with the imbalance oxygen, cancer cells adapted to hypoxia through changes in gene expressions and proteomics that affect the metabolism of cancerous cells (Paredes et al., 2021). Cell proliferation is reduced under hypoxia condition avoiding further consume of O_2. In cancer cells, mutations in oncogene and in tumour-suppressor genes together with highly metabolism changes, allow those cells to survive even in low oxygen condition (Luo et al., 2011). Hypoxia-inducible factor HIF-1 is a transcriptional factor involved in cellular adaptive metabolism when a reduction of oxygen level occurs. Particularly, many oxygen-related genes contain specific hypoxia response elements (HREs) that bind to HIF and allow to adapt in response to hypoxia in favouring tumour proliferation and metastasis (Yfantis et al., 2023). Metabolic reprogramming in hypoxia induced by HIF-1 involves the glycolysis. In hypoxic condition, different genes, such as glucose transport genes (GLUT-1 and GLUT-3) as also glycolytic enzymes such as the pyruvate kinases 2 (*PKM2*) and lactate dehydrogenase (LDHA) are under control of the transcriptional factor HIF-1 (Chen, C. et al., 2001) (Doherty & Cleveland, 2013). In cancer cells PKM2 promotes proliferation as a coactivator of HIF-1 stimulation chromatin binding, and transcriptional activator leading to a metabolism reprogramming of cancer cells promoting cancer progression and angiogenesis (Luo et al., 2011). In non-small lung cancer cells, studies *in vitro* and *in vivo* documented that transforming growth factor TGF-β inhibits the glycolysis in normal oxygen condition, however, promotes tumour cells under hypoxia condition through the binding with HIF-1 to MH2 domain of phosphorylated Smad3 (Huang, Y. et al., 2021).

1.3.2 Evading growth suppressors

Cells are capable to control tightly cell division. Tumour protein p53 (TP53) is considered “the guardian of the genome” because of its essential role in the regulation of DNA repair and cell division. Localised inside the nucleus, protein p53 binds to the DNA when damage occurs, p53 can activate repair mechanisms or alternatively if cannot be repaired the protein stops the cell from dividing and trigger apoptosis (Ozaki & Nakagawara, 2011).

Another important tumour suppressor gene is retinoblastoma (*RB*). RB is known to be involved in cell cycle regulation. It exerts its effect by interaction with the transcription factor E2F (E2F) and with chromatin remodelers and modifiers by repressing genes involved in the cell cycle. Once the RB is bound to the E2F protein the link can either block the expression of transcriptional co-activator or recruiting a transcriptional co-repressor, thus repressing the passage from stage G1/S of the cell cycle (Vélez-Cruz & Johnson, 2017).

The hyperphosphorylation of the RB by a group of cyclin-dependent kinases (cyclin D-cdk4/ cdk6 and cyclin E-cdk2) dissociates the bound with the E2F activation (Burke et al., 2010). In this way E2F will allow the transcription of co-activators that will allow the progression of the cell cycle by releasing the transcriptional repressor. During the cell cycle progression, the activity of CDKs decreases and the activity of protein phosphatase 1 (PP1) dephosphorylates the RB that will form again a complex with E2F proteins repressing the transcription of the cell cycle progression genes. This negative cell cycle control seems to be the basis of the tumour suppressor develop of the RB (Kolupaeva & Janssens, 2013).

1.3.3 Resisting cell death

The term apoptosis (in Ancient Greek apóptōsis, "falling off") describes a programmed cell death that happens in a multicellular organism (Alberts B, Johnson A, Lewis J, et al., 2002). The apoptosis mechanism involves an upstream regulator and a downstream effector component. The regulators are then divided into two pathways (extrinsic and intrinsic pathways). The extrinsic pathway, which involved in transmembrane receptor-mediated interactions, includes the tumour necrosis factor (TNF) receptor responsible for the transmission of the death signals to the intracellular signalling pathway, and the FAS ligand /FAD receptor (Locksley et al., 2001). The intrinsic pathway involves a non-receptor- mediated stimuli responsible to produce intracellular signals that directly affect the target cells, and these are mitochondria-initiated events (Kim, Ryungsa, 2005). However, both pathways culminate with the activation of the protease enzymes caspase8 and caspase9, which initiate the proteolysis cascade that ends with the apoptosis of the cells. TP53 gene can induce apoptosis by the upregulation of PMAIP1 (phorbol-12-myristate-13-acetate-induced protein 1), also known as NOXA and p53 upregulated modulator of apoptosis BH3 (Puma BH3) only because of substantial DNA damages or other chromosomal abnormalities (Oda et al., 2000).

In cancer, disruption of apoptosis is responsible for its development and progression. The resistance to apoptosis is due to whether the expression of anti-apoptotic protein as Bcl-2 and low expression of the pro-apoptotic proteins such as BCL2 Associated X (Bax). Both proteins are regulated by the p53 tumour suppressor (Basu & Haldar, 1998).

1.3.4 Enabling replicative immortality

Cells are subjective to a limited number of divisions after that they can enter in a senescence state where they are unable to divide or die. This is because at the end of DNA molecules there is a region of repetitive nucleotide sequence associated with specialised protein at the end of the chromatin, called telomere. At every cell division, the telomeric DNA gets shorter until it becomes too short and activate the quiescence state and so it stops the cell division. Cancer cells can bypass senescence by manipulating the enzyme telomerase to increase the length of telomere and keep replicating (Blasco, 2005).

1.3.6 Activating invasion & metastasis

Metastasis is a process by which cancer cells can spread to different tissues and organs of the body, from the primary tumour site and will generate a new tumour. The metastasis process is followed by a cascade of events: a local invasion of the cancer cells and this is due to a loss of cell-cell adhesion capacity, which allows the malignant cell to leave the primary tumour allowing the cell to invade the stroma (Martin et al., 2013). The alteration of the shape and the attachment to other cells and the extracellular matrix (ECM) involve loss of specific molecules involved in cell-cell adhesion as E-Cadherin. In contrast, it is well known that in different carcinoma cells N-cadherin is upregulated, even though this molecule is normally expressed in mesenchymal cells and in migrating neurons during the organogenesis (Cavallaro & Christofori, 2004).

The important progress in cancer research brought the two authors to revisit their original publication by adding two more hallmarks to the six previously discussed (Hanahan & Weinberg, 2011a).

1.3.7 Deregulating cellular metabolism

The metabolism of glucose is an important process for the sustaining cells in human body for nourishment and energy. This mechanism allows the energy to be used in the easiest form of adenosine triphosphate (ATP) through the oxidation of the glucose molecule carbon bond. Lactate can be the final product of the metabolism or CO_2 whether the glucose if fully oxidised in the respiration process inside the mitochondria (Liberti & Locasale, 2016). In normal cells, the metabolic activities rely on the mitochondrial oxidative phosphorylation (OXPHOS) for ATP generation for energy (Koppenol et al., 2011). In contrary, glucose degradation to pyruvate (or glycolysis), is a non-oxidative dependent process and is relatively inefficient for ATP generation compared to OXPHOS (Wang, Zixi et al., 2020). However, in tumour cells, there is an increase in glucose uptake associated with enhanced glycolytic flux leading to accumulation of metabolic by-product such as lactate. This process is called “the Warburg effect” and occurs even in presence of normal oxygen level and with fully mitochondria.

The observation that tumour cells use an enormous amount of glucose comparing with normal nearby tissues, was studied by Otto Warburg and his colleagues in 1920. Particularly they noticed that cancer cells were able to reprogram their energy metabolism to increase glycolysis, even in presence of oxygen. To do so cancer cells upregulate the glucose transported GLUT-1 that increase the amount of glucose transport in the cytoplasm. In addition to that, the glycolytic increase is due to a particular form of hexokinase linked to mitochondria which enable the upregulation of the glycolytic activity without the use of oxygen (Vander Heiden et al., 2009).

Finally, a study conducted on diabetic patients has shown that there is an increase in the chance of developing tumours because the production of ketones together with lactate are responsible for metastasis and tumour cells. High presence of glucose in different types of cancer can be visualised through the positron emission tomography (PET) using a radiolabelled analogue of the glucose (18 F-fluorodeoxyglucose, FDG) as a reporter (Croteau et al., 2016).

1.3.8 Avoiding immune destruction

The ability of the tumour to resist the immune system is still understudied. The existing theory is that cells and tissues are under the immune system surveillance that can recognise and eliminate most early cancer cells and nascent tumours. However, it seems that cancer cells can evade the immune system mainly in immunocompromised individuals. Studies conducted on mice, genetically engineered to be deficient in specific immune system components, revealed that tumours arise more frequently. Mainly it was found that the lack of immune system cells as CD8+ cytotoxic T lymphocyte (CTLs), CD4+ Th 1 helper T cells or natural killers (NK) were responsible for tumour development. Mice with both deficiency of T cells and NK, would even more susceptible to tumour development (Teng et al., 2008).

Finally, two more emerging characteristics were taken into consideration: the genetic impact and inflammation.

1.3.8.1 Genome instability & mutation

Accumulation of extra copies of DNA, chromosome deletion as also double or single strands break, are common in cancer. Even epigenetic lesions are responsible for increasing of the mutations that can lead to the acquisition of more mutation ending with tumour developing (Corcos, 2012).

1.3.5 Inducing or accessing vasculature

The word angiogenesis identifies a physiological process by which new blood vessels are grown from a pre-existent vessel. Sprouting angiogenesis, a process where endothelial cells and their growth depends on angiogenic factors such as the vascular endothelial grow factor A (VEGF-A) (Schmidt et al., 2007). By this process, the blood vessel can be established trough out the tissue. The splitting process also called intussusceptive angiogenesis involve vessels formation by splitting the existing blood vessel in two (Ribatti & Crivellato, 2012). During tumour progression, there is a continuation of vascularisation, an "angiogenic switch" to an activated state that ensure the expansion and tumour growth. VEGF-A is an angiogenic stimulator and is involved in the growth of blood vessels during embryo and postnatal development. Angiogenic inhibitors such as thrombospondin-1 (TSP-1) inhibitor acts to suppress the proangiogenic stimuli by binding to the transmembrane receptor localised on the endothelial cells (Rohrs et al., 2016). Angiogenesis and inflammation are involved in different pathological condition including cancer as documented in several studies (Carmeliet & Jain, 2000) Beta fibroblast growth factor (βFGF) is involved in the angiogenesis promoted by the inflammation (Sajib et al., 2018). Inflammatory cells expressing βFGF as also inflammatory mediator synthesise and release βFGF, though the endothelium, which in turn stimulates angiogenesis in autocrine manner (Andrés et al., 2009).

1.3.8.2 Tumour-promoting Inflammation

Rudolf Virchow in 1863 sustained that cancer was localised in the chronic inflammation site too. It is now known that inflammation is a key component in developing human cancer (Multhoff et al., 2012). Inflammation mediators are responsible for cancer develop such as prostaglandins as also the cytokinesis, TNF-α, IL-1β, IL-6 and IL-5. Reactive oxygen species (ROS) and nitrogen species (RNS) are stimulated transcriptional factor nuclear factor kappa-light-chain-enhancer of activated B cells (NF-KB) and Signal Transducer and Activator of Transcription 3 (STAT-3) that lead to cellular proliferation, genomic instability, invasion, and metastasis (Shrihari, 2017). Chemokines show an important role in the cancer-related inflammation and the components of chemokines system affect the stages of tumour progression including leukocyte recruitment, neo-angiogenesis survival, invasion, and metastasis of tumour cells (Fernandes et al., 2015). It is also demonstrated in pre-clinical and clinical trial that acting on chemokines system could be the target for development of future therapeutic strategy against cancer (Yeung & Jeang, 2011).

Recently, new prospective hallmarks were incorporated as core components of the hallmarks of cancer conceptualisation: "unlocking phenotypic plasticity," "nonmutational epigenetic reprogramming," "polymorphic microbiomes," and "senescent cells" (Hanahan, 2022c).

1.3.8.3 Unlocking phenotypic plasticity

Cells undergoing to development, determination, and organisation into tissue, and to reach homeostatic functions, are associated with terminal differentiation in which progenitor cells stop growing, sometimes permanently, upon culmination of these processes. However, there are different evidences suggesting that unlocking the restricted ability for the phenotypic plasticity to evade from terminal differentiation state, represent an important component for cancer progression (Yuan, S. et al., 2019). There are several ways by which this plasticity can operate: (i) Dedifferentiation in which nascent cancer cells originated from normal cells that are approaching to a fully differentiated state could reverse back to the progenitor-like cell state (Yao & Wang, 2020); (ii) Blocked differentiation where incomplete differentiation can maintain the expanding cancer cells in a partially differentiated progenitor-like state (He, L. et al., 2000); (iii) Transdifferentiation manifestation, in which cells that were initially committed into one differentiation pathway, switch to different development program from the preordained normal cell-of-origin (Tosh & Slack, 2002).

1.3.8.4 Nonmutational epigenetic reprogramming

Nonmutational epigenetic regulation of gene expression is already explained in the mechanism of embryonic development, differentiation, and organogenesis (Zeng, Y. & Chen, 2019). Examples in adults involve the long-term memory that is due to changes in gene and histone modification, chromatin structure and the gene expression switches that are maintained by positive and negative feedback (Hegde & Smith, 2019). However, several evidence suggest that epigenetic alteration can be responsible for the acquisition of hallmark for cancer progression to malignant state. Examples are: (i) Microenvironment Mechanism of Epigenetic Reprogramming in which microenvironment can cause broad changes in the epigenome for example hypoxia, a common characteristic of tumours, that lead to insufficient vascularisation, reduces the activity of ten-eleven translocation (TET) demethylase, that causes changes in the methylome causing hypermethylation (Thienpont et al., 2016); (ii) Epigenetic Regulatory Heterogeneity which is revealed by increasingly powerful technologies for profiling genome wide DNA methylation (Heyn et al., 2016), histone modification (Audia & Campbell, 2016), chromatin accessibility and posttranscriptional modification and translation of RNA (Janin et al., 2020); (iii) Epigenetic Regulation of the Stromal Cell Types Populating the Tumour Microenvironment where a recent study as suggested reprogramming can effect modification in epigenome in addition to exchanges of cytokines, chemokines growth factors that can alter intracellular signalling networks in all of those cell types (Lu et al., 2020).

1.3.8.5 Polymorphic microbiomes

There is increase evidence suggesting that the polymorphic variability in the microbiomes can have a profound impact on cancer phenotypes (Kim, Donghyun et al., 2017). Study in humans, but also mouse model manipulation in cancer, revealed that specific microorganisms not only bacteria can have protective or deleterious effect on cancer development, malignant progression, and response to therapy (Helmink et al., 2019).

1.3.8.6 Senescent cells

Cells senescence represent an irreversible form of proliferative arrest and evolved as a protective mechanism for maintaining tissue homeostasis, as also a complementary way to the apoptosis that inactivate or remove diseased, dysfunctional, or unnecessary cells (Gorgoulis et al., 2019). Senescence program also determine a change in the morphology and metabolism in cells through the activation of a senescence-associated secretory phenotype (SASP) that releases different proteins including chemokines, cytokines, and protease in which identity depends on a senescent cell arises (Birch & Gil, 2020). Despite the protective benefits of senescence in limiting malignant progression, increasing evidence suggest that senescent cells promote tumour development and malignant progression (Wang, Boshi et al., 2020). The SASP mechanism is thought to be responsible for promoting tumour phenotypes in paracrine way to viable cells in proximity, as well as to other cells in the tumour microenvironment (TME), including single molecules (proteases that activate and/or desequester them) to deliver hallmark capabilities (Faget et al., 2019). An example is the activation of B-Raf Proto-Oncogene, Serine/Threonine Kinase (*BRAF*) in primary fibroblast that led to the secretion of Insulin Like Growth Factor Binding Protein 7 (*IGFBP-7*) , that acts through autocrine/paracrine way to induce senescence and apoptosis in neighboured cells (Wajapeyee et al., 2008).

1.4 Biomarkers

The World Health Organisation suggests that a biomarker is any substance, structure, or process able to be measured and can predict disease state. From a clinical point of view, a cancer biomarker can give information on the risk of developing the disease in a specific tissue and cancer progression and the potential response to the therapy (Henry & Hayes, 2012).

1.4.1 Cancer biomarkers

Genetic, epigenetic, proteomic and glycomic biomarkers can be useful for cancer identification, progression and improving health outcomes of the disease in a population. Classification of cancer biomarkers can be separated into: (i) Predictive, (ii) Prognostic and (iii) Diagnostic (Goossens et al., 2015).

-***Predictive:*** it is related to how good a patient's response to specific drug therapy is. An example was described in a study conducted on breast cancer patients in whom there is an overexpression of the Human Epidermal Growth Factor Receptor (*HER2*). The data showed the clinical benefit of using the trastuzumab in metastatic breast cancer that overexpressed HER2 (Slamon et al., 2001). As also in colorectal cancer, the use of drug therapy such as cetuximab, an epidermal growth factor (*EGF*) inhibitor, has reported a good response in patients having wile-type-KRAS mutation, but to be resistant in patients with KRAS-activating mutation (Van Cutsem et al., 2009).

-***Prognostic*** biomarker is more related to the status of cancer and progression. It also aims to inform the patients on the risk of the future outcomes in case of cancer recurrence after primary treatment. Although markers can be a simple measure, for example stage of diseases or size of the tumour, they are more complex such as protein abnormal level or genetic mutations (Riley, R. D. et al., 2009). For instance, MYCN Proto-Oncogene (*MYCN*) in paediatric oncology, amplification of this gene is related to poor outcome in patients with neuroblastoma (Riley, Richard D. et al., 2004). With the availability and improvement in the treatments, cancer patients survival is increasing in the last decades (Oldenhuis et al., 2008). Although, there are patients receiving treatments, which they do not benefit instead they experience cytotoxicity. The search of new therapies, arise some elucidation between prognostic and predictive biomarkers to offer a better patient selection of treatments. The difference between prognostic and predictive is the first provides information about the patient overall cancer outcome, while the predictive biomarkers give information regarding the effect of the therapeutic treatment (Colburn et al., 2001) Also, predictive biomarkers can be used a target for therapy. An example in breast cancer the estrogen receptor (ER), the progesterone receptor (PR), and the human epidermal growth factor receptor 2 *Neu* (HER2/neu) (Payne et al., 2008).

-***Diagnostic*** biomarkers help to early diagnose a disease in non-invasive way. Recently, in colorectal cancer, there was an improvement in diagnosis by looking at the stool DNA (Imperiale et al., 2014). An example of the diagnostic marker in the treatment of colorectal cancer is the Cologuard which combine faecal immunochemical test (FIT) with a multigene DNA stool test in people with high risk to develop this type of cancer. The studies conducted on this test and the encouraging results obtained led to be approved by the Food and Drug Administration (FDA) in August 2014 (Goossens et al., 2015).

EGFR is now accepted to be the target for the treatment in patients suffering from non-small-cell lung cancer (NSCLC). Particularly, individuals with a mutation in EGFR gene can be treated with EGFR-inhibitor (Schulze et al., 2019). Anaplastic lymphoma kinase (*ALK*) is responsible for the emerging of NSCLC and anti-ALK is now approved for the treatment of a patient with this type of cancer, together with c-ros oncogene 1 (ROS1) (Ulivi, 2020). In 50% of patients with melanoma showed a mutation in the gene BRAF a serin/threonine-protein kinases B-Raf (Ulivi, 2020). In prostate cancer mutation in the gene *BRAC1/BRAC2* increase the risk to develop pathology and is now suggested as a biomarker (Castro & Eeles, 2012).

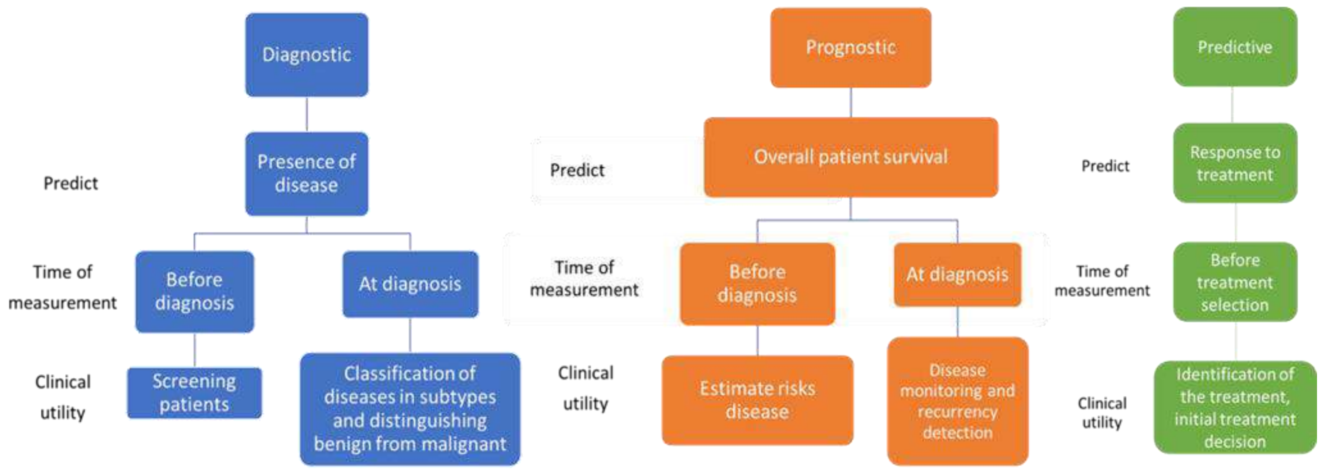


Figure 1.5 – Clinical biomarkers for diagnosis. Biomarkers are presented along with each predict and the clinical utility in which they can be sued.

1.5 Cancer stem cells

Cancer stem cells (CSCs) are a small population of cells within the tumour that are responsible for cancer recurrence in patients following chemotherapy and radiotherapy. Like normal stem cells, CSCs possess stemness properties such as self-renewal and the capacity to generate differentiated cells that contribute to the heterogeneity of the tumour. Those cells also called tumour initiating cells (TICs) are responsible for tumour establishment and growth and seem to be associated with aggressiveness, tumour relapse and therapy resistance. CSC appear to be generated after mutations occur affecting adult normal stem cells that represent the source of the organogenesis and tissue homeostasis. Sajiro Makino in 1959 published an article in which for the first time introduce the concept of cancer stem cells but calling them "tumour stem cells", identifying those groups of cells that are resistant to the chemotherapy and showing a chromosomal difference with the original cell bulk. Makino's theory was confirmed through colony forming assays stating the tumour could have originated from cells that share characteristics of the stem cells. More confirmation arrived studies on acute myeloid leukaemia (AML) (Lapidot et al., 1994). Particularly, it was described that a population of $CD34^{+}CD38^{-}$ AML stem cells were associated with AML in severe immune-deficient (SCID) mice (Lapidot et al., 1994) reviewed in (Barbato et al., 2019).

1.5.1 Origin of cancer stem cells

To explain the origin of CSCs two hypotheses were proposed: the stochastic model also called clonal evolution (CE) and the hierarchy model also called the CSC model. The stochastic model describes that every single cell within the tumour could be the cell-of-origin and facilitate tumour initiation and progression malignant cells as biologically homologues whereby, the functionalities depend on extrinsic factors originated from the tumour microenvironment or intrinsic factors such as signalling pathways and transcription factors (Plaks et al., 2015). The stochastic model depends on the fact that tumorigenesis occurs in normal differentiated somatic cells, that can acquire genetic alterations affecting cell cycle genes which contribute the aberrant proliferation and expansion (Easwaran et al., 2014). Example of tumour types presenting stochastic models are colorectal cancer (Odoux et al., 2008) and B cell lymphatic leukaemia (Williams et al., 2007). The acquisition of those new characteristics is associated with cell heterogeneity (Gerdes et al., 2014). The hierarchy model instead states that within the tumour there is a minority of subpopulation of cells that have the stem cells properties of self-renewal and the ability to differentiate their abilities and phenotypes. Those cells are also able to recreate the heterogeneous tumour (Rich, 2016).Examples of hierarchy model is in some solid tumour such as breast cancer and in non-solid tumour such as acute myeloid leukaemia (AML) (Al-Hajj et al., 2003a) (Bonnet & Dick, 1997) (Fig 1.5).

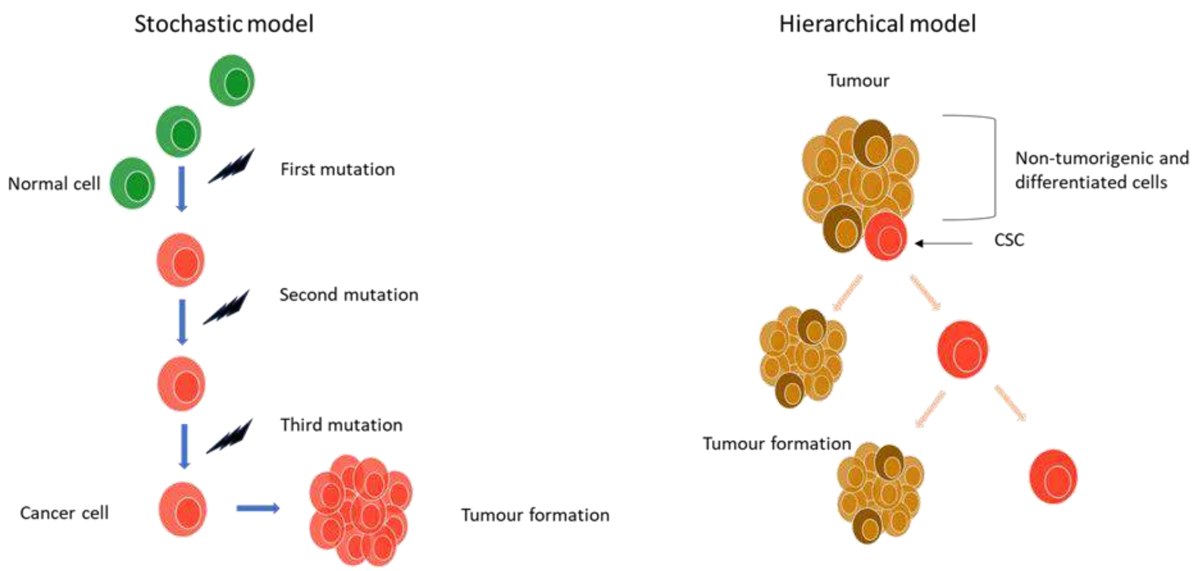


Figure 1.6 ¬- Stochastic and Hierarchical model for CSCs. In stochastic model hypothesis that the normal cells can undergo to a serial of mutations which culminate to a bulk of tumour. The hierarchical model proposes that within the tumour there are cancer cells with the same characteristic of pluripotency and self-renewing of the stem cells defined CSCs. Those cells are highly tumorigenic and form new tumour. CSCs dived asymmetrically and are able to develop new CSCs and progenitor cells that leading to differentiated cancer cell and the tumour bulk.

1.5.2 CSCs Plasticity model

Cancer cells can exist within the tumour in various phenotypes such as CSC and non-stem cancer cells (NSCCs) and these states affect the functional properties of the cell (Visvader & Lindeman, 2008). Tumour microenvironment (TME) is the environment around the tumour that comprehends various cell types that ensure the tumour development, angiogenesis, inflammation, progression, and metastasis (Visvader & Lindeman, 2008). Healthy fibroblasts can transform into cancer-associated fibroblasts (CAFs) in response to signals from TME (Madar et al., 2013). Tumour progression and stemness is maintained by CAFs which express growth factors including the insulin-growth factor-II (IGF-II), hepatocyte growth factor (HGF), VEGF in Wingless-related integration site (Wnt) and Neurogenic locus notch homolog protein (Notch)- dependent manner (Kalluri & Zeisberg, 2006). Contrary to the normal homeostasis in which Wnt Notch is regulated and limited to the stem cell niche, TME-directed signalling can affect both CSC and NSCCs (Junttila & De Sauvage, 2013a). Because these signalling pathways regulate stemness, they can also enable dedifferentiation from NSCCs into CSCs (Chen, W. et al., 2014).

Non-tumorigenic, immortalised human mammary epithelial cells (HMLEs) were induced to EMT by ectopic expression of transcription factors Twist or Snail to determine whether adult cells that go through EMT and adult stem cells have similar characteristics and are capable of inducing EMTs in epithelial cells (Cano et al., 2000) (Yang, J. et al., 2004). Results revealed that cells acquired fibroblasts like mesenchymal form, downregulated mRNAs encoding for epithelial markers such as E-cadherin, and upregulated mRNAs encoding mesenchymal markers as N-cadherin, vimentin, and fibronectin (Mani et al., 2008). Andriani et all in 2016 demonstrate a link between plasticity and stemness with the regulation of the epithelial-mesenchymal transition (EMT) in lung cancer. Principally, in human lung cancer, the treatment with the tumour growth factor-β1 (TGF-β1) is responsible for the switching of the cell in stem cell-like when other remained unresponsive to stimulation (Andriani et al., 2016). In breast cancer, it has been shown that the modulation of transcription factor ZEB1 and other EMT transcription factors are involved in the switch of non-stem cells to CSC- like in immortalised human mammary epithelial cells (Chaffer et al., 2013).

1.5.3 Cancer stem cells in the tumour microenvironment

The concept of the stem cells niche was hypothesised by Schofield in 1978 that proposed that the stem cell niche is essential for the determination of the stem cell fate and their behaviour was able to influence other cells inside the niche (Schofield, 1978) reviewed in (Barbato et al., 2019). CSC microenvironments show a high variety of cells and heterogenous complex belonging to stromal cells, immune cells and epithelial cells as also extracellular macromolecules involved in the support of the extracellular matrix (EC) (Chaffer et al., 2013)(Figure 1.4).

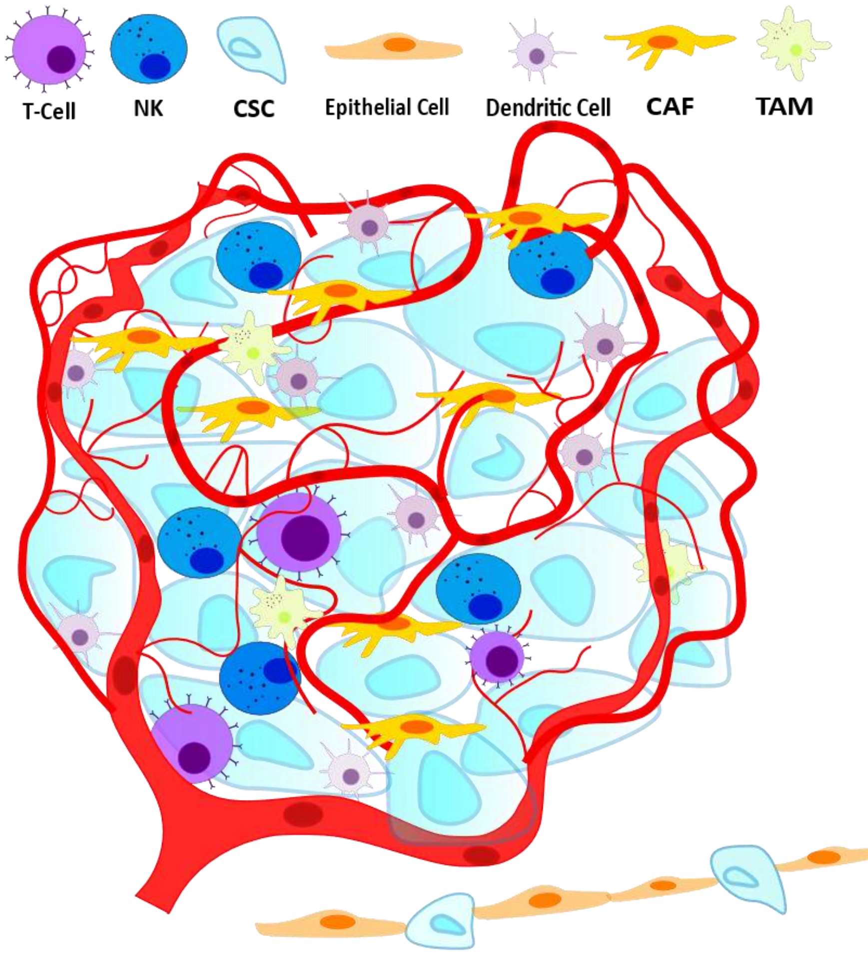


Figure 1.7 - Schematic representation of the cancer stem cell microenvironment. The CSC niche is composed from a complex network involving inflammatory and immune cells including tumour associate macrophages (TAMs), immune T-cells, endothelial cells, cancer associate fibroblasts (CAFs), dendritic cells supporting tumour community.

1.5.3.1 Immune cells

TME is a complex system in which co-exist non-immune cells (endothelial and stromal cells) and immune cells [including macrophages, polymorphonuclear cells, mast cells, natural killer (NK), dendritic cells (DC)] together creating an interaction ensuring tumour survival (Giraldo et al., 2014). Immune cells can infiltrate tumour, and their composition and organisation with TME are strongly associated with the clinical outcome of cancer patients (Giraldo et al., 2019). In several types of solid tumours and haematologic malignancies cancer patients have been treated with monoclonal antibody therapy targeting inhibitory receptor of immune cells, has reported remarkable response rates (Giraldo et al., 2018). An important factor that is involved in invasion and metastasis and contributes to EMT is TGFβ1. This molecule is secreted from tumour-associated macrophage (TAM) and myeloid-derived suppressor cells (MDSCs) (Buczek, Miles, et al. 2016).

1.5.3.2 Cancer associate fibroblasts (CAFs)

CAFs are responsible for the increase of proliferation as also of the enhancing of the extracellular matrix production and cytokines secretion. Examples are the stromal cell-derived factor-1, (SDF-1); VEGF and plated-derived growth factor (PDGF) and HGF (Junttila & De Sauvage, 2013b). The progression of cancer is due also to mesenchymal cells such as the adipose cells (Dirat et al., 2011). CSC mechanical support is ensured by CAFs through the production of fibrillar collagen, as also the secretion of C-X-C motif chemokine 12 (CXCL-12) (Najafi et al., 2019). Finally, is also involved in the EMT by the secretion of TGFβ in the early step of the invasive and metastatic process (Dianat-Moghadam et al., 2018).

1.5.3.3 Mesenchymal Stem cells (MSC)

MSCs are multipotent stromal cells, involved in chronic inflammatory sites and allow EMT through the secretion of TGFβ which enhances the metastatic process (Ridge et al., 2017). Those cells are also able to differentiate in various cells such as adipocytes and osteocytes. In gastric cancer, those cells are responsible for angiogenesis and cancer progression through the secretion of VEGF and other inflammatory molecules and macrophage inflammatory protein-2 (MIP2), TGFβ-1 and pro-inflammatory cytokines (IL-6 and IL-9) (Li, W. et al., 2015).

1.5.3.4 Endothelial cells

In the CSC microenvironment, angiogenesis is very important as it delivers the nutrients for the CSC metabolism necessary for the self-renewal, invasion, and metastatic process (Fessler et al., 2015). Together with perivascular cells, endothelial cells create blocks of vessels that, with angiogenic factors such as VEGF enables the tumour vasculature which permits the growth and the proliferation of cancers. Endothelial cells promote self-renewal and CSC progression by the secretion of cytokines as IL-3, Granulocyte-Macrophage Colony-Stimulating Factor (GM-CSF) as also IL-1, IL-6, VEGF-A and bFGF (Pirtskhalaishvili & Nelson, 2000).

1.5.4 Epithelial-mesenchymal transition (EMT) and cancer stem cells (CSC)

An important property of cancer stem cells is their invasiveness and metastatic potential that involves a process called epithelial-mesenchymal transition (EMT). This process, by which cells lose their cell-cell adhesion and transdifferentiate in a more mesenchymal phenotype, gain migratory and more invasive properties (Grünert et al., 2003). This mechanism is essential in embryonic development as it occurs also in the process of healing wounds and in organ fibrosis. Epithelial cells differ from the mesenchymal cells. Epithelial cells are connected by the tight junction and adhesion junction including desmosome. Mesenchymal cells show a more spindle-shaped morphology and lack of polarisation. Also, epithelial cells express high-level of markers such as E-cadherin while in mesenchymal cells N-cadherin, fibronectin and vimentin are largely expressed. In cancer, epithelial cells acquire a mesenchymal property that allows them to invade local tissue and to disseminate to distant tissues (Kalluri & Weinberg, 2009)

As already mentioned, the role of TGFβ secreted in the CSC microenvironment, also plays an important role in EMT due to the binding to its receptor which determines the activation of "small and mother against the decapentaplegic protein" (SMAD) (Thiery, 2002). Once activated, it translates to the nucleus allowing the activation of target genes involved in EMT such as *SNAIL, bHLH* and *ZEB* transcription factors (Massagué, 2012). Besides TGFβ signalling, other signalling pathways are involved in the EMT. The WNT signalling pathway promotes EMT by stabilising β-catenin through the inhibition of glycogen synthetase kinases-3β (GSK3β). Once β-catenin is translocated in the nucleus this will activate the transcription factors lymphoid enhancer-binding factor (LEF), and T cell factor (TCF) promote EMT (Niehrs, 2012). The Notch signalling pathway also is involved in EMT via activation of SNAIL2 expression directly by intracellular Notch (Xie et al., 2012).

1.5.5 CSC biomarkers in solid tumour

Although development in therapies targeting CSCs has been made, specificity in the targeted antigen remains a challenge. Different CSCs biomarkers have been identified. Cell surface biomarkers of CSCs identified include cluster of differentiation 133 or human prominin-1 (CD133), CD44 a cell adhesion receptor expressed in different cancers, the epithelial cell adhesion molecule (EpCAM) which is expressed in most human carcinomas (Macdonald et al., 2018). However, the challenge in the CSCs selection is related to specificity for antigens that are also found in healthy tissue increasing the chances for off-target effects in treated patients.

1.5.5.1 CD133

CD133 belongs to the transmembrane glycoproteins, and despite the function still not being fully clear, it is supposed to be involved in the organisation of membrane topology (Yin, A. H. et al., 1997). This biomarker is expressed in human embryonic stem cells (hESC) as also in the normal tissue (Kim, Won-Tae & Ryu, 2017).

In cancer, this marker has been shown to characterise cells with high tumorigenicity as also the ability to form spheroids. This marker is present in different types of cancer such as breast, liver, stomach, and colon (Brugnoli et al., 2019). Interestingly, CD133 expression are affected by Metastasis Associated Lung Adenocarcinoma Transcript 1 (MALAT1) a long coding RNA involved in cancer progression and ELAV Like RNA Binding Protein 1 (HuR), by regulating EMT features, suggesting a potential control of Hur/MALAT1 on CD133-related tumour progression (Latorre et al., 2016). CD133 showed a relationship with EMT markers in tumour cells, particularly patients with metastatic breast cancer, present a high level of N-cadherin (a marker for mesenchymal cells) (Armstrong et al., 2011). While in patients with lung cancer was shown to promote EMT in combination with B cell specific Moloney murine leukaemia virus integration site 1 (BMI1), another cancer stem cell marker (Koren et al., 2016). Gene expression analysis of the mRNA level of CD133 has shown to be a good marker for predication however, CD133 protein is correlated to poor prognosis in patients with breast cancer (Joseph et al., 2019).

1.5.5.2 CD44

The CD44 antigen is a cell-surface protein that in humans is encoded by the gene CD44. Is involved in cell-cell interactions and cell adhesion and migration (Su et al., 2016). Alternative splicing during the transcription process leads to the formation of two isoforms, the standard isoform (CD44s) and variant isoform (CD44v) (Xu, H. et al., 2015). CD44 is expressed also in haematological cancer. It was shown that expression of CD44 is linked to an increase in proliferation, self-renewal, and metastasis (Su et al., 2016). In breast cancer was shown that the percentage of $CD44^+/CD24^-$/cytokeratin positive (CK^+)/$CD45^-$ is related to high tumorgenicity (Paula et al., 2017). It has been shown that in gastric cancer knocking down CD44 resulted to a reduction of spheroids formation and a decrease tumour production in severe combined immunodeficiency mice (Takaishi et al., 2009). In CSCs CD44 is used as a biomarker for diagnosis, therapeutic and prognosis. As shown in gastric cancer the presence of circulating $CD44^+$ is correlated to poor prognosis (Zhang, H. et al., 2019). Therapeutical approaches are now made to target $CD44^+$ using the adenoviral techniques of siRNA *in vitro* (Nam et al., 2015).

1.5.6 Regulators of Stemness in solid cancer

CSCs share with stem cells unlimited growth and are characterised by undifferentiated states. Hadjimichael et all described pluripotency in embryonic stem cells (ESCs) as a process that involves a group of core-network of transcription factors (Hadjimichael et al., 2015). Those consisted of Sex determining region Y SYR Box-2 (Sox2), homeobox protein NANOG, octamer-binding transcription factor 4 (Oct3/4), Kruppel-like factor 4 (Klf4) also proto-oncogene c-MYC together with signalling pathways WNT/β-catenin, Hedgehog/Notch and TGFβ (Lundberg et al., 2016a). Particularly, the genes *NANOG1, Sox2* and *Oct3/4* are responsible for the activation of self-renewal and suppress the genes involved in differentiation. (Rosner et al., 1990) Those factors are also potential CSC biomarkers.

1.5.6.1 NANOG

Homeobox protein NANOG coded by the *NANOG1* gene, is a transcription factor involved in embryonic stem cell development by maintaining pluripotency (Jeter et al., 2015). It is highly expressed in pluripotent cells such as ESCs and embryonic germ (EG) and embryonal carcinoma (EC) cells during the differentiation its expression is downregulated (Chambers et al., 2003a). NANOG affects the fate of pluripotent cells as ectopic expression can benefit the reprogramming in a cell-division-rate-independent manner however, a downregulation of NANOG in mouse and embryonic ESC induce differentiation (Darr et al., 2006) (Hanna et al., 2009). It is expressed in different types of cancer, in breast cancer, the expression of both Nanog and Oct3/4 is correlated to a poor prognosis and involved in EMT (Gawlik-Rzemieniewska & Bednarek, 2016) (Wang, Dan et al., 2014). A study conducted on colorectal cancer documented that CD133+ cells displayed a high-level expression of NANOG (Xu, F. et al., 2012). In non-small-cell lung cancer (NSLCLC), A549 it was shown that SNAIL2 can activate NANOG through SMAD1/Akt/Gsk3β pathway (Liu, Chen-Wei et al., 2014). Double knockdown of NANOG and OCT4 lead to reduction in the expression of SNAIL2 and increasing expression of E-cadherin protein levels in A549 cells (Chiou et al., 2010). It was shown that PCa cell express *NANOG* mRNA primarily from the *NANOGP8* locus on chromosome 15q14 and using a lentivirus promoter reporter construct *NANOGP8-EGFP* isolated PCa cells NANOGP8-GFP+ showed CSC characteristics such as enhanced clonal growth and tumour regenerative capacity (Jeter et al., 2011).

1.5.6.2 SOX2

Sox2 is a transcription factor and is involved in the maintenance of self-renewal and pluripotency in undifferentiated stem cells (Lundberg et al., 2016b). In gastric cancer, its expression is correlated to a poor prognosis as also with lymphoid metastasis in patients with cardiac gastric cancer (Yang, L. et al., 2017) . Analysis *in vitro* of tumour spheroids revealed a high expression level of CD44 as well CD133 and transcriptional factor Sox2, Nanog and Oct3/4. Expression of Sox2 was found also in high Gleason grade prostate cancer, and in glioblastoma Sox2 represents the stem transcription factor responsible for the maintenance of stemness properties in glioblastoma cells (Zhang, X. et al., 2016).

1.5.6.3 Oct3/4

OCT4 is a protein that in humans is encoded by the *POU5F1* gene and is involved in the self-renewal and the undifferentiated embryonic stem cells. It is also used as a marker for undifferentiated cells (Niwa et al., 2000). High levels of Oct4 is found in prostate cancer and also in breast cancer (Wang, Ying-Jie & Herlyn, 2015a). Conventional treatment using cisplatin, etoposide and doxorubicin and gamma radiation are inefficient in lung cancer patients expressing a high level of Oct4 (Prabavathy et al., 2018). Resistant to the treatment for mesothelioma was shown in cells with increased level of Oct4 and Sox2 (Mohiuddin et al., 2020).

1.6 Prostate cancer

The International Agency for Research on Cancer has estimated that 7.1% of men died worldwide from prostate cancer in 2018 and it is the second most frequent cause of death from cancer in men after lung cancer. The incidence of, and mortality from prostate cancer increases with age, with the highest rate in men over 65 years of age. Current approaches are ineffective at detecting disease relapse and metastatic prostate cancer remains incurable (Bray et al., 2018).

The events that lead the prostate to change into prostate cancer are due to a multistep process that starts as prostatic intraepithelial neoplasia (PIN) (Brawer, 2005). Later follows the more localised form of prostate cancer and then reaches the more invasive state called prostate adenocarcinoma that concludes with the metastatic prostate cancer. Metastasis represents the main cause of death for people affected by prostate cancer. The metastasis spreads to the liver and then the lung and bones (Ziaee et al., 2015). Different studies have been conducted to better understand bone metastasis to propose a possible treatment. It is suggested that the EMT process is involved in the metastasis in different cancer types including prostate cancer because where cells circulate as circulating tumour cells (CTCs), during the EMT (Ruscetti et al., 2015). Those cells can be resistant to physical barriers during the process of bone metastasis, till reaching the bone marrow stroma, the internal site of the bone via sinusoids localised inside the bone marrow cavity, and finally proceed to metastasise. Once prostate cancer cells reached the bone marrow prostate cancer start to secrete growth factors including endothelin-1, adrenomedullin and fibroblast growth factors that allow maintaining the tumour growth and cancer cell survival (Hiratsuka et al., 2006).

Symptoms of prostate cancer are still not fully understood. That can be confused with benign prostatic hyperplasia discomfort during urination (urinating frequently mainly during the night), blood in the urine are common symptoms. However, prostate cancer can also cause urination dysfunction. Metastatic prostate cancer can show other symptoms such as pain in the bones and pelvis (Merriel et al., 2018).

Gleason score represents a system to grade prostate cancer and it is fundamental for diagnosis of prostate cancer disease since it indicates the disease aggressiveness (Tagai et al., 2019). Identified for the first time by Donald Gleason in 1974 these methods help the doctor to understand how the cancer cell behaves and suggest the optimal therapy for patients (Epstein et al., 2016). The mortality of prostate cancer depends on geographic area as also other factors. Survival of patients is correlated on the stages of cancer for example the survival at 5 years for patients affected by metastatic prostate cancer is around 30% when compared with patients in which cancer is localised the 5 years survival is 100% (American Cancer Society, 2021).

1.6.1 Prostate Cancer diagnosis

Diagnosis of prostate cancer represents an important tool for clinical and patient care. Prostate cancer diagnosis relays on the digital rectal examination (DRE) and the blood test screening the presence of prostate-specific antigen (PSA) as also the transrectal ultrasound (TRUS) guided biopsy (Loeb & Catalona, 2009). Still today the TRUS guided biopsy is used for prostate cancer diagnosis. Nonetheless, limitation in the procedure is linked to sepsis that historically was defined as “infection plus systemic inflammatory response syndrome” (Rhodes et al., 2017). The risk related to this procedure does not surprise as the needle needs to pass through the rectal wall to enter in the prostate, inoculating rectal flora. Antibiotic used for TRUS biopsy is ciprofloxacin, however, several studies were performed to reduce the risk of infections and in treatment of antibiotic resistance (Moe & Hayne, 2020).

The digital rectal examination (DRE) is a valuable diagnostic technique for valuation of dysfunction in males however, it does show some limitations for patients who were affected by prostate cancer because the presence of the prostate specific antigen (PSA) appeared in patients with a normal DRE (Loeb & Catalona, 2009). The total PSA represent a good biological test for prostate cancer diagnosis. Therefore, patients with a high level in the blood of this antigen are at high risk to develop the pathology. However, the test is not able to discriminate prostate cancer (PCa) from benign prostatitis (BHP) (Oesterling et al., 1993).

For this reason, different methods were used and optimised to better improve the screening. One optimisation was optimised based on age. It is recommended a threshold between 2.5ng/ml and 6.5ng/mL for men at the age of 40s,50s,60s and 70s (Oesterling et al., 1995). That would have enhanced the specificity of the test. The results that Lobi et all obtained, were in favour to consider a baseline when they analysed 13,943 men younger than 60 that showed a total prostatic-specific antigen (tPSA) higher than 2.5 ng/mL and with a DRE which gave doubtful results (Loeb et al., 2006). Another method aiming to improve the screening test was to consider the ratio of free PSA (Fpsa) with PSA. Particularly this method was used when patients showed a good DRE and a value of PSA between 4 ng/mL and 10ng/mL. In 15% of cases, a value below this range was sufficient to diagnose a patient without PCa (Lee, R. et al., 2006). Diagnostic accuracy of PSA testing was performed in 2,620 men 40 years and older showing that on 930 cancers detected the PSA cut point of 4 ng/mL had a sensitivity of 86% and a specificity of 33%, concluding, that despite the fact that PSA has fair discriminating power for prostate cancer detection, and sensitivity is relatively non-specific test (Hoffman et al., 2002).

Another way for PCa diagnosis is PCa gene 3 (*PCA3*) analysis in the urine stating a threshold of 35 score. Particularly, this test evaluates the mRNA level in the urine. This method was found very useful mostly in patients who shown persistent level of PSA > 4 ng/mL with previously negative biopsy (Wei, J. T. et al., 2014).

1.6.2 PCa diagnosis by imaging

Ultrasound and magnetic resonance imaging (MRI) are commonly used for PCa diagnosis (Galfano, 2009). The improvement in the MRI technology enhanced the ability to detect *in vivo* PCa within the gland. Another use of MRI is for surgical prostatectomy. MRI was also shown to be used as a type of treatment for focal therapy and radiotherapy. Multiparametric magnetic resonance imaging (mpMRI) is an interesting method for PCa diagnosis, also in patients with localised PCa it was revealed to be useful for the observation of cancer progression (Descotes, 2019).

Positron emission tomography (PET) is a PCa diagnostic method offered in the hospital. It is often combined with a CT- scan (PET-CT) that increase the accuracy of the analysis. Patients generally are injected with a small amount of dye fluorocholine (FCH), and lay down for 30 minutes the time to take pictures of the body. The scan will take the picture based on the radiation given from the dye (Beheshti et al., 2010).

1.6.3 PCa Histopathological diagnosis

The Gleason score is a prognosis method the prostate cancer. The score is based on the examination under the microscope. Pathological scores go from 6 lowest risk to 10 the highest risk of mortality. The scoring takes into consideration adding the two most common grade of the cells in a tissue and calculating the overall grade (Moch et al., 2016) (Lawson et al., 2019).

A

B

Grade groups	Cells arrangement	Gleason score	Risk group
Grade group 1	Small uniform gland	< 6 or less	Low grade cancer
Grade group 2	More space (stroma) between gland	3+4=7	Medium-grade cancer
Grade group 3	Distinct infiltration of cells from gland to margins	4+3=7	Medium-grade cancer with more cells abnormalities
Grade group 4	Irregular masses of neoplastic cells with few glands	4+4=8;5+3=8 ;3+5=8	High-grade cancer
Grade group 5	Lack of or occasional glands cells sheets	4+5=9;5+4=9 ;5+5=10	High-grade cancer

Figure 1.8 - Gleason scoring system. A Schematic presentation of histological pattern in prostate biopsy. B Prostate cancer grade group, cells arraignment, Gleason grade. Figure adapted from Veterans Administration Cooperative Urological Research Group.

The TNM system is a wide method used for cancer staging. In the TNM the (T) is referred to the size and the extension of the tumour. The (N) is referred to the number of the lymph node that has cancer and (M) is referred if cancer is metastasised that is mean cancer has spread from the primary tumour (Borley and Feneley 2009) (Rosario & Rosario, 2020).

Table 1.1 - TNM staging of prostate cancer (PCa). The TNM (tumour, nodes, metastases) table describing the size of cancer and how far it is growth according to the American Joint Cancer Committee prostate cancer staging guidelines (Adapted from Localized Prostate Cancer StatPearls Publishing)

Primary Tumour (pT)	
T0	No evidence of primary tumour
T1	No detectable on the digital rectal exam (DRE)/imaging (could be find during surgery for reason other cancer)
T2	Tumour can be detected on DRE
T2a	Tumour is half or less than one side (lobe) of the prostate
T2b	Tumour is more than half of one prostate, but it hasn’t invaded the other lobe yet
T2c	Tumour is localised in both lobe
T3	Tumour is spread outside the prostate
T3a	Tumour is spread to the prostate capsule
T3b	Tumour is spread to seminal vesicles (the gland of each side the bladder)
T4	Tumour is spread to other tissue near the prostate other than vesicles
Nearby lymph nodes	
N0	No cancer cells have been found in the nodes
N1	Cancer cells have been found in the nodes
Distant Metastasis	
M0	Cancer has not metastases yet beyond the prostate
M1	Cancer has spread beyond the prostate
M1a	Cancer has spread to distant lymph nodes
M1b	Cancer has spread to the bones
M1c	Cancer has spread to other organs with or without bones disease

1.6.4 Prostate cancer treatment

Continued monitoring of the prostate condition is a good practice for the patient. Monitoring of prostate condition requirs taking the test for the PSA, digital rectal exam, transrectal exam and needle biopsy (Screening & Board, 2002).

In the case of the growth of PCa, there are different treatments available for patients. For a patient who is affected by the PCa, but the tumour is localised inside the gland, surgical proceed to remove the prostate gland and tissue close to the gland represents the best choice to remove all the portions of the tumour. Still radiation is a technique used for the treatment of cancer. These procedures request high energy x-ray that kill the cancer cells. These methodologies depend on the stage of the tumour as also can determine side effect that could be difficult on the urination as also impotency that can increase with the age (Gay & Michalski, 2018).

Another non-surgical treatment is hormone therapies. This type of treatment consists of blocking the hormone actions which will stop cancer cells growth or hormone removal. Androgen deprivation therapy (ADT) is commonly used as first-line in the treatment symptomatic metastatic and in men who received radical prostatectomy (Perlmutter & Lepor, 2007). Example of drugs therapy include the abiraterone acetate that affects the production of androgen from the prostate cancer cells and is generally used for patients at a high stage of PCa (De Bono et al., 2011).

Different immunotherapy approaches were identified PCa treatment. Sipuleucel-T (Provenge) represents an immunotherapy strategy, that targets the prostatic phosphatase (PAP) a tissue antigen expressed in prostate cancer. This therapy, that is approved by FDA, acts by inducing a T cell-mediated response through the stimulation of the antigen-presenting cell (APC) in combination with the recombinant PAP to co-stimulate the granulocyte-macrophage colony-stimulate factor (GM-CSF). This type of therapy, which concluded phase III clinical trial, showed an improvement of the overall survival as also a reduction of the 22% of mortality in patients affected by metastatic castration-resistant prostate cancer (mCRPC) (Hammerstrom et al., 2011). Chemotherapy has shown low benefit in patients with mCRPC (Tannock et al., 2004). Docetaxel-based treatment has a survival advantage however, patients were reporting disturbing side effect including fluid retention, with peripheral edema and weight gain leading the physicians to be reluctant to use chemotherapy in men presenting no or few symptoms of mCRPC, preferring sipuleucel-T as only therapy specifically approved for asymptomatic mCRPC due to low toxicity when compared with docetaxel (Cheever & Higano, 2011).

1.7 Breast cancer

Breast cancer is the most common cancer in women with 1 new case in 10 (Alkabban FM, 2022) It represents the second principal cause of death in women and the evolution of the pathology is very silent, so women don't know unless by a routine screening. Physically, both women and men have the breast. The breast is generally composed from adiposity tissue conversely to the man it presents more glandular tissue. Female breasts are formed with 12-20 lobes that are still divided in other smaller lobules (Moore et al., 2010). A network of nerves, blood vessels, lymph vessels and lymph nodes are localised in the adipose tissue. The female breast also is composed of fibrous connective tissue (Thomsen & Tatman, 1998). As glandular gland breasts are highly sensitive the hormonal changes as they change under the menstrual cycle like the genital female system (Jagannathan & Sharma, 2017).

Invasive breast cancer occurs in 1 in 8 women in the United States and is thought that in 2018 around 266,120 will have experienced invasive breast cancer and just 63,960 localised breast cancer. In the UK between 2015-2018, 55,900 new cases were reported with an increase of 15% (2016-2018)(Cancer Research, 2021). Risk factors in the development of breast cancer are related to age. In the UK British females develop cancer between 45-50 years of age (Akram et al., 2017a). It was also shown that women in their 20s that had a child have the lowest risk to develop breast cancer compared to women in their 30s that they didn't have any children (Franca-Botelho et al., 2012).

Another important factor that contributes to the development of breast cancer is the diet. The combination of a diet low in fibre and rich in fat can trigger breast cancer (Lee, H. P. et al., 1991). Ultimately there is also a genetic predisposition to develop breast cancer (Eberl et al., 2005). It was demonstrated that first-degree parent increases the chance of heredity breast cancer 2-fold to 3-fold. Genetic risk is responsible for 10% probability to develop breast cancer and in women younger than 30 it is a 25% risk to develop breast cancer. As with other cancers, one of the causes of the development of breast cancer is a disruption at DNA level and defects in DNA damage repairs or cell-cycle checkpoints. Defects in DNA damage response (DDR) gene expression represent a common breast cancer phenotype is DNA damage as gene mutations in genes such as *TP53, BRCA1* and *BRAC2,* phosphatase and tensin homolog *(PTEN)* (Alkabban & Ferguson, 2018). Those tumour suppressors are mostly involved in in maintenance of DNA fidelity, in particular *BRCA1* (DNA damage repair), *TP53* (cell cycle checkpoint) and *PTEN* (blockage of cell cycle progression in G1 and involved in DNA damage repair) (Guler et al., 2011). Breast cancer can be also classified based upon the hormone receptor (estrogen and progesterone receptor) and HER-2 positivity. A well-established classification scheme estrogen receptor (ER) and progesterone receptor and ErbB2/HER2 groups is divided into three categories. (i) hormone receptor sensitive (ii) hormone receptor negative with HER2 over-expression; (iii) “triple negative” (this represents a breast cancer which doesn’t express any of the three receptors) (Foulkes et al., 2010). Triple-negative breast cancers show inactivation of DNA repair genes *BRCA1* (Adamo & Anders, 2011). Hormone receptor positive cancer, not refractory to anti-estrogen, can be treated with selective estrogen receptor modulators (SERMs) or selective estrogen receptor down-regulators (SERDs) to slow down cancer cells growth (Lumachi et al., 2015). A combination of radiotherapy (instigating DNA damages) with surgical and hormonal therapies could be successful. However, if the DNA damage response do not affect cancer cells, inside the tumour, the apoptosis will fail (Davis & Lin, 2011). Another challenge that could occur in hormone-sensitive breast cancer, could be developing hormone insensitivity and therefore becoming resistant to SERMs and SERDs drugs (Rao et al., 2011). For hormone receptor negative breast cancer, HER-2 overexpressing, treatment include Trastuzumab or other HER-2 antagonists (Duffy et al., 2011). For triple-negative breast cancers (TNBC) are sensitive to traditional chemotherapies, however, the treatment outcome is poor when cancer relaps or metastases aggressively (Gonzalez-Angulo et al., 2011). Breast cancers presenting aberrant DNA damages could be more sensitive to drugs which are poly (ADP) ribose-1 (PARP) inhibitors, causing chromosomal abnormalities in those cancer cells that defect in genes involved in DNA damage repair such as *BRCA1* (McCabe et al., 2006). That could be potential effective for TNBC cancer cells (Carey et al., 2010).

Breast cancer can be divided depending on invasive or non-invasive state. Ductal carcinoma in situ (DCIS) non-invasive breast cancer, is a cancer not spreading through the walls of the ducts into the nearby breast tissue. DCIS happens when abnormal cells occurred in the milk ductal, but they are limited to that area and do not invade adjacent tissues (Kitamura et al., 2019). Invasive ductal carcinoma (IDC) is invasive breast cancer where atypical cells can spread from outside the milk duct and invade the closest tissues and is the most common type of breast cancer. An example of that is triple-negative breast cancer (Liedtke et al., 2008). This type of cancer is highly present in women in the premenopausal stage with an increase in incidence. A mutation in genes and changes in the environmental factors can cause breast cancer. Also, dysregulation of a gene involved specific pathways such as RAS/MEK/ERK pathways and PI3K/AKT pathways fail in apoptosis and develop cancer (Cavalieri et al., 2006). Another type of invasive breast cancer is the invasive lobular carcinoma (ILC) and represents 1 in 10 of invasive breast cancer and starts from the lobules where the breast glands make milk and as IDC can spread to other parts of the body. IDC is difficult to detect with normal exam such as physical and mammography exams. Women affected from IDC are likely to show invasive carcinoma in both breasts (Chen, H. et al., 2019).

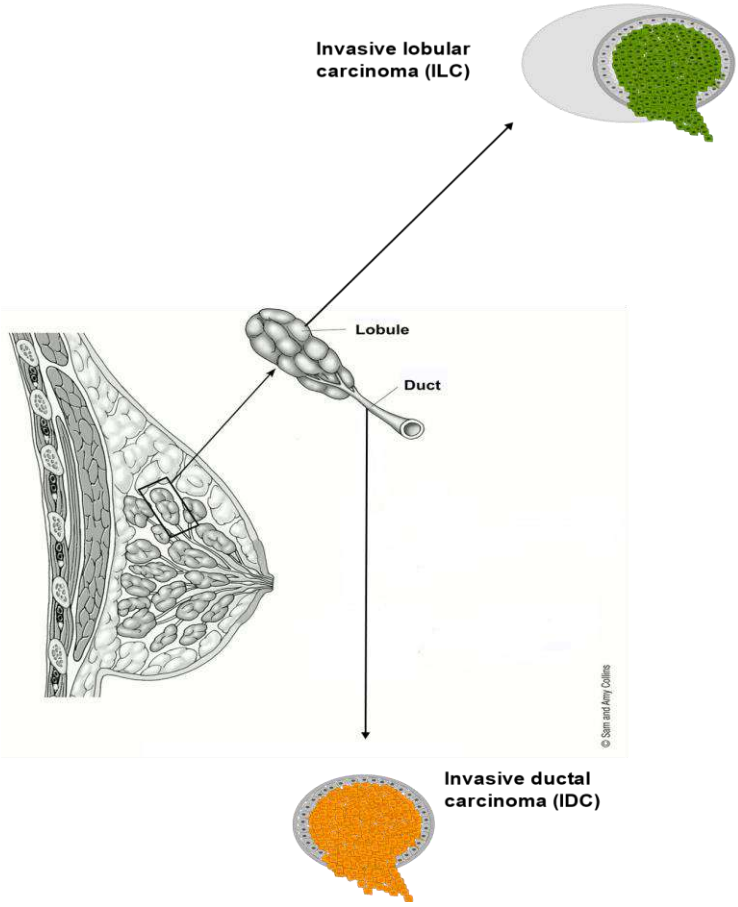


Figure 1.9 - Schematic representation of breast cancer invasive progression. Cross section of the lobe and duct that compares different types of cancer. The epithelial cells will misplace, and abnormal cells will take place following several stages of the cancer progression. Invasive carcinoma is divided in invasive ductal carcinoma (IDC) and invasive lobular carcinoma (ILC). Picture adapted from https://www.cancer.org/cancer/breast-cancer/about/types-of-breast-cancer/dcis.html

1.7.1 Breast cancer diagnosis

Because a primary prevention for breast cancer is not available at the moment unless in the most extreme measure as in mastectomy is important to continue to promote early detection (Caplan, 2014). It is now accepted that an early diagnosis correlates with a reduction in morbidity and mortality (Oeffinger et al., 2015).

1.7.1.1 Self- examination

This technique can be taught by the physician and is a valuable way to check breast condition (Kosters & Gotzsche, 2003). Women in this way can find abnormalities in the size and shape of the breast. In 2010 Ozkan et all conducted a study on 113 midwives and nurses to understand the level of the knowledge of the self-examination. Interestingly, the studies demonstrated that a continued education program was having a good impact on the female community regarding the importance of these pathologies (Ozkan et al., 2010).

1.7.1.2 Position emission tomography (PET-CT)

As already described for prostate cancer, in breast cancer the PET- CT is also used for the diagnosis. Using a specific dye that is injected or drank from the patients this technology can give more accurate information on the different organs which might be affected by cancer. Is also able to detect possible lymph node localised or area in which there is metastasis (Antoch et al., 2004).

1.7.1.3 Digital mammography

This diagnosis is used for the identification of lumps in dense tissue. It is also considered and the gold standard for the early detection of breast cancer (American College of Radiology, 2018). Digital mammography (DG) use computer-based electronic conductors to display a picture of the interior of the breast for most accurate images leading to a correct diagnosis (American cancer society, 2022). However, this type of diagnosis can also generate false negatives and false positives mainly in patients with dense breast tissue (Sprague et al., 2014). DG has shown a lower sensitivity and higher false positive among younger women (Lehman et al., 1999). Digital breast tomosynthesis (DBT) is an imaging technique that allows to generate pseudo three-dimensional (3D) images focusing on a single section of the anatomy (Østerås et al., 2019). Contrary to DM the screening performed by DBT reduce the superposition of breast tissue and improve the cancer detection for both fatty and dense breast in the age groups that are relevant for mammography screening (Zackrisson et al., 2018).

1.7.2 Breast cancer staging

As for prostate cancer, the TNM method is used also for the breast cancer stage. Also, here is considered the size of the tumour (T) and the presence of lymph nodes close to the armpit (N) and if the tumour metastasised (M) (Teichgraeber et al., 2021). Stages are described from 0 in which tumour is not spread and the cancerous cells did not invade adjacent tissues while stage 4 describes the worst scenario where cancerous cells invaded the close tissue and tumour metastasised in another organ such as brain, lung, bones, and liver (Cserni et al., 2018).

Table 1.2 - TNM staging breast cancer (BCa). The most recent edition calls the pathologic stage the "AJCC Anatomic Stage Group." The stage listed in the pathology report is the Anatomic Stage Group. Table adapted from "Breast Cancer Staging: Updates in the AJCC Cancer Staging Manual, 8th Edition, and Current Challenges for Radiologists, From the AJR Special Series on Cancer Staging" American Journal of Roentgenology 2021.

Primary Tumour (pT)	
pT0	No evidence of primary tumour
pTis	Ductal carcinoma in situ, or Paget's disease
pT1	Invasive tumour ≤2.0 cm in greatest dimension
pT1mi: Tumour ≤0.1 cm in greatest dimension (microinvasion)	
pT1a: Tumour >0.1 cm but ≤0.5 cm in greatest dimension	
pT1b: Tumour >0.5 cm but ≤1.0 cm in greatest dimension	
pT1c: Tumour >1.0 cm but ≤2.0 cm in greatest dimension	
pT2	Invasive tumour >2.0 cm but ≤5.0 cm in greatest dimension
pT3	Invasive tumour>5.0 cm in greatest dimension
pT4	Invasive tumour of any size with direct extension to the chest wall and/or to the skin, or inflammatory breast carcinoma
pT4a: Tumour extension to chest wall (does not include tumours limited to pectoralis muscle)	
pT4b: Tumour extension into the skin with ulceration	
pT4c: Tumour invasion into the chest wall and the skin	
Inflammatory breast carcinoma (requires clinical symptoms)	
Regional Lymph Nodes (pN)	
pNX	regional lymph nodes cannot be assessed (eg, previously removed, or not removed)
pN0	No regional lymph node metastasis identified histologically, or isolated tumour cells only
pN0 (i+): Isolated tumour cells (ITC): malignant cells <0.2 mm and no more than 200 cells	
pN1mi	Micrometastases (malignant cells measuring 0.2 mm - 2.0 mm, or more than 200 cells)
pN1a	Metastases in 1 to 3 axillary lymph nodes (at least one measuring > 2.0 mm)
pN2a	Metastases in 4 to 9 axillary lymph nodes
pN3a	Metastases in 10 or more axillary lymph nodes
Distant Metastasis (M) (required only if confirmed pathologically in this case)	
pM0	No metastases
pM1	Histologically proven metastasis larger than 0.2 mm

1.7.3 Breast cancer treatment

Anti-cancer drugs vinblastine and vincristine for breast cancer were introduced the first time in 1961, together with surgery, radiotherapy and chemotherapies represent the main form of treatment against breast cancer (Debnath et al., 2006). Surgery represents one of the main treatments against breast cancer. It represents the strategy used when breast cancer is not spread in the other part of the body, and it correlated with cancer characteristics, however, is important to take into consideration patients feel. An example of a surgical procedure is the lumpectomy this type of surgery involved just a part of the breast that is affected by the tumour and a small portion of the breast (Fisher et al., 1995). This type of surgery is more accepted by women as part of the breast is still intact. Those proceedings are more frequent in women affected at the first stages of the disease but is important that the treatment keep going with the use of radiation and, chemotherapy and hormone replacement therapies (Dorval et al., 1998).

Mastectomy represents is the entire breast remove through surgery (Rebbeck et al., 2004). This type of procedure is the most effective and decrease the risk of breast cancer relapse. However, it can affect women psychology by feeling less attractive with the consequence of depression (Keskin & Gumus, 2011).

Cancer affected by hormones can be treated with the anti-estrogen receptor (Clarke et al., 2003). Anti-estrogen drugs that are now used in breast cancer include tamoxifen as also the raloxifene and toremifene. Tamoxifen is an anti-estrogen mainly used in women who are estrogen receptor positive. It acts by inhibiting the hormone oestrogen to enter breast cancer (Mehta et al., 2003). In 1998 the drugs were accepted by the Food and Drug Administration (FDA) (Jordan, 2007). A study conducted on the drugs showed the efficacy of tamoxifen decreased breast cancer recurrence, some doubts were raised mainly because there was an increase in the risk of breast cancer for 5 years and of the specificity of the drugs. Mostly, because the effect of tamoxifen is on patients who were estrogen receptor-positive but with no evidence of estrogen receptor-negative breast cancer (Cuzick et al., 2003). It was also demonstrated that side effects such as venous thrombosis, cataracts, and menstrual disorders were frequent in women under tamoxifen treatment. For this reason, the FDA in 2007 accepted to use of raloxifene as an anti-cancer drug in women in the postmenopausal stage because decrease cancer progression with less risk of side effects (Akram et al., 2017b).

1.8 EPCR is a potential biomarker for prostate cancer stem cells (PCSCs)

The first part of the project focused on the potential involvement of EPCR in prostate cancer stem cells. Study conducted on the nasopharyngeal carcinoma demonstrated that EPCR is involved in the maintenance of tumour cells stemness through the regulation of lipid metabolism and mitochondrial fission (Zhang, P. et al., 2022). Additionally, in breast cancer cell line MDA-MB-231 mammary fatpad (mfp) EPCR+ generated abundant mammospheres and displayed aldehyde dehydrogenase (ALDH) activity, another marker for cancer stem cells is linked to poor prognosis (Schaffner et al., 2013). EPCR is also expressed in prostate cancer however, its role in carcinogenesis is still unclear (Menschikowski et al., 2011a). Endothelial protein C receptor (EPCR) also known as CD201, is a 26.6 KDa type 1 transmembrane glycoprotein mainly expressed by arterial and venous endothelial cells in the lung and heart (Laszik et al., 1997)(Fig 1.10). It shows homology for a sequence and 3D structure with the major histocompatibility complex class 1/CD1 family protein, particularly with CD1d (Oganesyan et al., 2002). Human EPCR is encoded by the PROCR gene (OMIM 600646). EPCR was isolated from endothelial cell-specific and cloned and it has a high affinity with protein C and activated PC (APC). The binding of ECPR with the thrombin-thrombomodulin complex activate protein C (Mohan Rao et al., 2014). The activation of protein C generates the anticoagulant enzyme activated protein C (APC) which determines inhibits coagulation in combination with protein S, through the inactivation of two coagulation factors Va and factor VIIIa (Stearns-Kurosawa et al., 1996). It has been demonstrated that deficiency of protein C can lead to thrombotic complications (de Wouwer et al., 2004)

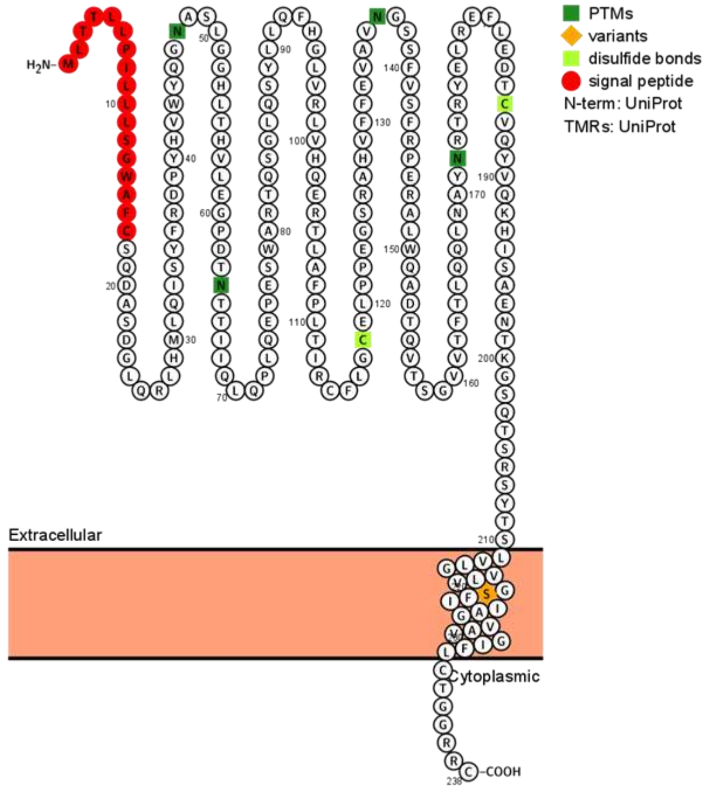


Figure 1.10 - Protter illustration of human protein (hEPCR) (Uniprot Q9UNN8) with annotation of the various UniProt features.

The coagulation pathway is a cascade of events leading to the homeostasis (Chaudhry et al., 2018). It involved two pathways the intrinsic pathway, consisting of transmembrane receptor tissue factor (TF) that under physiological condition surrounds blood vessel and initiate the clotting (Smith, S. A. et al., 2015). The extrinsic pathway of blood coagulation is required for the thrombosis. In thrombotic complications and cancer at advanced stages, the activation of the coagulation can affect mortality in cancer patients (Mackman et al., 2007). The activation of anticoagulant and anti-inflammatory responses is due to the binding of the proteinase-activated protein C (aPC) to the endothelial protein C receptor (EPCR) thereby triggering the protease-activated receptor-1 (PAR1) signalling in endothelial cells and initiating an anti-inflammatory and apoptotic response to protect the integrity of the endothelial barrier (Pendurthi & Rao, 2018). Analysing the influence of aPC/EPCR on tumour migration and metastasis has suggested that an endothelial overexpression of EPCR decreases tumour adhesion and transmigration in B16-F10 melanoma cells (Bezuhly et al., 2009) . The role of EPCR-dependent PAR1 activation by aPC in the migration of breast cancer cells and the inhibition of lung cancer cell apoptosis which enables metastasis has been demonstrated. Also, different studies have demonstrated that EPCR could be used as a marker for poor clinical outcome in different cancers such as lung and breast (Stearns-Kurosawa et al., 1996) hEPCR has also been shown to be a marker for identifying cancer stem cells, including CD44+/CD24- and the expression of ALDH (Al-Hajj et al., 2003b). Isolated prostate cancer stem cells (PCSCs) were previously obtained in our laboratory using a second generation of lentivirus expressing the enhanced green fluorescent protein (EGFP) under NANOG-promoter (Buczek, Reeder and Regad 2018). Generated monoclonal antibody against EPCR in our laboratory (Di Biase 2020) was tested on isolated prostate cancer stem cell to determine whether EPCR could be a potential marker for prostate cancer stem cells (PCSCs) and potentially develop a monoclonal drug-based therapy.

1.9 SPAG5 as a novel marker involved in drug resistance and cancer progression in aggressive breast and prostate cancer

The second part of the project is focused on the investigation of pro-oncogenic marker previously studied on breast cancer. Sperm Associated Antigen 5 (SPAG5) is a protein that in humans is encoded by the *SPAG5* gene, it is an important component of the mitotic spindle for normal segregation and progression during the anaphase of the mitotic cycle. It is required for chromosome integrity as also participates in the chromosomal alignment and in the sister chromatid segregation (Cheng, T. et al., 2008). The gene is located on chromosome17 (17q11.2) with 24 exons and nine reported mRNA spliced variants. Amplification of the 17q11 region seems common in different types of cancer (Mack & Compton, 2001a).

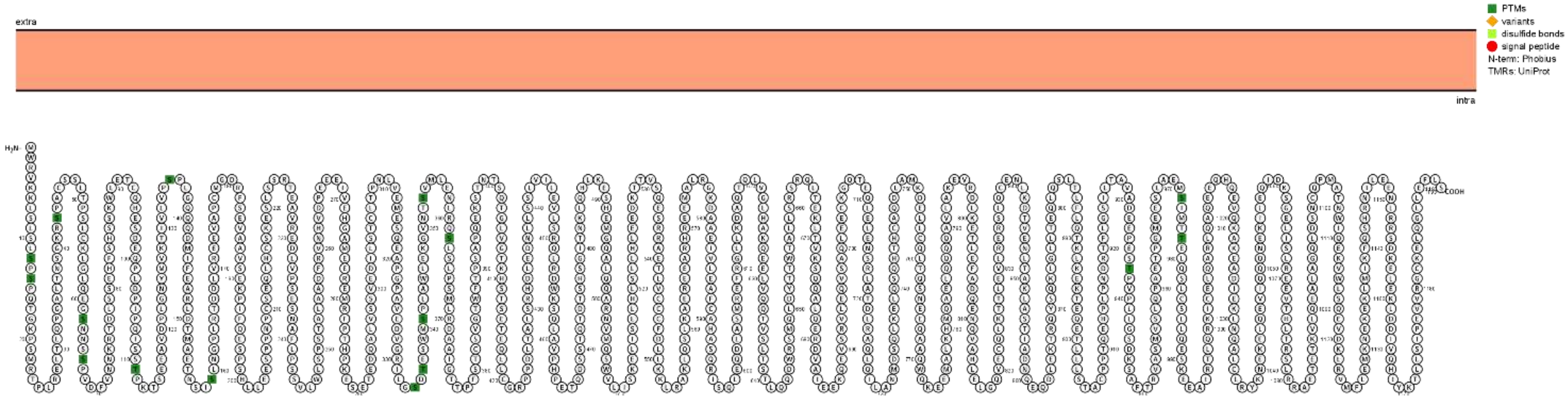


Figure 1.11 - Protter illustration of SPAG5 (Uniprot Q96R06) with annotation of the various UniProt features.

In 2016 collaboration between Nottingham Trent University and Nottingham University Hospitals NHS Trust identified SPAG5 as factor that drives proliferation in breast cancer using an artificial neural network (ANN) based integrative data-mining methodology that was applied to three cohorts [(Nottingham-discovery (ND), Uppsala and METABRIC (Molecular Taxonomy of Breast Cancer International Consortium)] (Lancashire et al., 2010) (Abdel-Fatah et al., 2016a). Following this study reported in 2016, they analysed the association of SPAG5 gene and SPAG5 protein expression in estrogen receptor-positive breast cancer with treatment response, resulting in a prolonged 5-years distal relapse-free survival in estrogen-positive breast cancer with or without lymph node (Abdel-Fatah et al., 2020a). Amplification in the expression of SPAG5 was seen in different types of cancer including lung cancer, breast cancer, prostate, and bladder cancer (Mohamadalizadeh-Hanjani et al., 2020). A study conducted on 112 patients showed that upregulation of SPAG5 was responsible for the poor survival, in primary bladder urothelial carcinoma (BUC). The finding also demonstrated that SPAG5 enable cancer progression and apoptosis suppression by upregulation of the Wnt3 through AKT/mammalian target of rapamycin (mTOR) pathways (Liu, J. Y. et al., 2018) .The role of SPAG5 in the progression and promotion of cancer was documented by Jiang et all in which it was reported that knockdown of SPAG5 can modulate the expression through a reduction of Wnt3 expression in breast cancer cell lines, contrary there was a high expression of the Wnt3 when SPAG5 was overexpressed (Jiang et al., 2019). The study conducted on prostate cancer suggested a role of SPAG5 in cancer progression. Using a small non-coding RNA (miR-529) they demonstrated that was able to decrease PCs metastasis *in vivo* while they observed a reduction in the progression, migration, and invasion *in vitro* when SPAG5 was downregulated (Zhang, Hongtuan, et al 2016a).

1.10 EPCR studying on prostate cancer stem cell (PCSCs): Aims and objectives

- Hypothesis: EPCR is a marker for cancer stem cells in prostate cancer.

- Aim: Test and validate this for potential development of a combination of antibody-drug therapies against prostate cancer stem cells (PCSCs) and studying the effect on the epithelial-mesenchymal transition in isolated PCSCs.

- Objective

- Transduction of human DU145 and PC3 prostate cancer cells using lentiviral constructs carrying the promoter of NANOG controlling the expression of EGFP, and isolate EGFP positive cells using a Beckman Coulter MoFlo™ XDP cell.
- Confirm the presence of hEPCR expression in sorted cells population by flow cytometer assay.
- Investigate the expression of cancer stem cell markers (CSCs), embryonic stem cells markers (ESCs) and Epithelial-mesenchymal transition markers (EMT).
- Investigate the capacity of hEPCR antibodies to trigger antibody-dependent cytotoxicity (ADCC) *in vitro*.

1.11 SPAG5 studying on breast cancer and prostate cancer: Aims and objective

Hypothesis: SPAG5 pathways are involved in the drug resistance to chemotherapies and cancer progression.

Aims: Develop a stable cell lines SPAG5 silenced and evaluate the effects in genes involved in pathways implicated in cancer progression and drug resistance. Assess the consequences in cell cycle when treated with chemotherapy drugs Doxorubicin in prostate cancer and Epirubicin in breast cancer cell SPAG5 deficient.

Objective:

- Generation of stable prostate and breast cancer cell line SPAG5 deficient.
- Investigate the transcriptome and proteome of the stable cell lines obtained using RNA- sequencing and mass spectrometry technology.
- Using free gene analysis resources to study pathways and protein networks involved in SPAG5 regulation.
- *In vitro* functional study on proliferation and drug resistance analysis on cell cycling in DU145 and MDA-MB-231 SPAG5 deficient cells

2. EPCR is a potential therapeutic target for prostate cancer (PCa)

2.1 Introduction

Prostate cancer is the most common cancer affecting men. Complications involving coagulation events that can led to the haemorrhage are frequent in patients with metastatic hormone-refractory prostate carcinoma (De la Fouchardiere et al., 2003). Other pathologies linked to the coagulation are thrombocytopenic thrombotic purpura, thrombosis and Trousseau's syndrome (Desai et al., 2015) and the development of factor VIII inhibitor, an important blood-clotting protein involved in the blood coagulation (Toole et al., 1984). Patients affected by cancer exhibit an upregulation of thrombin which could led to pro-carcinogenic events that could be stopped/controlled by anticoagulant or anti-inflammatory protein C/thrombomodulin-mediated mechanisms (Castellino & Ploplis, 2009). These facts and observation justify studying the activity and the mechanisms of action of the anticoagulant protein C (PC) pathway in patients with prostate cancer (Esmon, 2009). An important co-factor involved in the PC-pathway is the endothelial protein C receptor (EPCR) (Menschikowski et al., 2011b).

Previous work in our laboratory has shown that the cytoplasmic promyelocytic leukaemia (cPML) isoform I is involved in the Epithelial to Mesenchymal Transition (EMT) of human prostate cancer cells (PCCs) (Buczek., Miles, et al. 2016). Recently, our group has generated a model for studying *in vitro* induction of EMT in prostate cancer cell lines using a lentiviral-driven expression of PML I in prostate cancer cells. Three different constructs of PML have been generated in the human DU145 and PC3 cell lines, but only one induced a mesenchymal and invasive phenotype (Di Biase et al., 2018). Cells with a mesenchymal behaviour showed an increased migration and invasion and, proteomic mass spectrometry analysis enabled the proteome of both cell lines to be compared and specific proteins that were overexpressed in the invasive cells to be identified.

Using criteria such as expression in prostate cancer, and low or absent expression in normal tissue, this approach identified the endothelial protein C receptor (EPCR, CD201) as a potential therapeutic target as it is not expressed in normal prostate but is notably expressed in some PCa (Menschikowski et al., 2011c). Using hybridoma technology our team has generated monoclonal antibodies (mAb) against human EPCR (hEPCR) that have the potential to be used for antibody-based therapeutics. The first part of the current project was to assess if EPCR expressed on the cell membrane of prostate cancer stem cells (PCSCs) triggers antibody-dependent cellular cytotoxicity (ADCC) using our generated monoclonal antibodies (JvGCRC-H61.3 and JvGCRC-H599.5) using a surrogate *in vitro* assay for ADCC. The ADCC is a mechanism by which antibody that bounds to structures on target cells recruits effector cells via the binding of its Fc domain to the gamma receptor III a (FcγRIIIa, CD16) expressed on effector cells (typically NK cells) and thereby triggers the killing of the target cell by the effector cell (Román et al., 2014).

2.1.2 EPCR is expressed in prostate cancer stem cells (PCSCs)

Previous studies have shown that in addition to its role in the anticoagulant pathway in vascular cells, EPCR is involved in the tumour progression. Principally, the activation of the EPCR-PAR1 by aPC is responsible of the cell migration in breast cancer and downregulates apoptosis in lung cancer (Antón et al., 2012). EPCR has also been identified as a potential cancer stem cell biomarker (Hwang-Verslues et al., 2009).In aggressive basal-like breast cancer, EPCR is highly expressed (Park et al., 2010). Furthermore, in breast cancer tissue enriched for stem cell-like subpopulations expressing high levels of cancer stem cell biomarkers such as $CD44^{high}/CD24^{-}$ surface subtype and ALDH express EPCR (Al-Hajj et al., 2003c); (Ginestier et al., 2007) and EPCR defines a specific subpopulation $CD44^{high}/CD24^{-}$ cells in triple negative breast cancer, the most aggressive form of the disease (Ruf & Schaffner, 2014). Due to the role of EPCR in cancer progression in different types of cancer, including prostate cancer (Wojtukiewicz et al., 2019) a potential involvement in prostate cancer stem cells (PCSC) has been proposed. Previous work in our laboratory has established a technique for isolating a population of cells which exhibit cancer stem cell characteristics such as self-renewal and differentiation. This uses a lentiviral construct carrying the promoter of a stem cell marker, NANOG, which drives the expression of an enhanced green fluorescent protein (EGFP) reporter (Buczek et al., 2018a). The characterisation of the stemness has been assessed and confirmed using sphere formation and differentiation assay.

2.2 Material and Methods

2.2.1 Cell culture

The human DU145 (prostate carcinoma derived from metastatic site: brain) and PC3 (prostate carcinoma derived from metastatic site: bone) prostate cancer cells were purchased from the American Type Culture Collection (ATCC). Cell culture was performed using sterile class II laminar flow cabinet. Cells were incubated in a 5% v/v CO_2 humidified atmosphere at 37° C and passaged/harvested when they reached 80%-90% confluence. The supernatants were removed, cells washed and then incubated with 0.25% (v/w) Trypsin-EDTA for 5 minutes. DU145 cells were maintained in EMEM medium (Corning catalog #10-010-CVR) supplemented with 10% v/v foetal calf serum (FCS), 1% L-glutamine, 1% non-essential amino acids

(NEEA,Sigma-Aldrich), 1% sodium pyruvate (Sigma-Aldrich). PC3 cells were maintained in F-12K medium (Corning #10-025-CV) supplemented with 10% v/v foetal calf serum (FCS). Cell counting was performed by re-suspending the harvest cell pellet in 1-3 mL of cell dedicated medium and re-suspending the cell pellet in a trypan blue solution 1:10. A haemocytometer was used to count the total number of living cells from the death cells (blue stained).

2.2.2 Evaluation of generated monoclonal antibody EPCR expression in prostate cancer using flow cytometry

Wild type DU145 and PC3 human prostate cancer cells were incubated with JvGCRC-H61.3 mAb (patented application submitted: application no. US17/428,816) (Di Biase, 2020) to evaluate the expression of EPCR by both cells line. For this, cell were grown, harvested and counted as described in section 2.1, and $1x10^5$ cell for each cells type were aliquoted in 12x75 mm polycarbonate flow cytometry tubes. Cells were centrifuged for 5 min at 259 xg and potential non-specific binding blocked by incubation with 50µL of FcR blocking solution (MACS Miltenyil Biotech 130-059-901) for 15 min at room temperature. Cells were incubated at 4° C with fixable LIVE/DEAD™ Violet (Life Technologies L34955) viability stain (1µL for 1 mL of PBS) and with unconjugated JvGCRC-H61.3 EPCR mouse mAb (application no.1901640.1), monoclonal antibody generated in our laboratory, in PBS (100 µL, 1µg/mL). Cells were washed three times with 2 mL of PBS and centrifuged for 5 min at 300 xg. Cells were incubated with 5 µg/mL of Goat anti-mouse/IgG polyclonal secondary antibody Alexa Flour™633 (Life Technologies # A21052) for 30 min at room temperature protected from the light. Cells were washed three times with 2mL of PBS and finally resuspended with 300µL of Isoton II™sheath fluid (Beckman Coulter). Samples were analysed using unstained cells and cells incubated with secondary antibody alone as control. Data were acquired using a Beckman Coulter Gallios™ flow cytometry and analysed using Beckman Coulter Kaluza™ software.

2.2.3 Bacterial transformation and plasmid production

Lentivirus vector *PL-SIN-NANOG-EGFP, PL-SIN-EF1α-EGFP, PL-SIN-pLKO.1-puro*, *psPAX* and *pMD2.G* in LB bacterial stab culture were purchased from Addgene. XL-1 Blue super-competent bacteria used for transformation and plasmid amplification were gently thawed on ice. For this, 100 µL of cells were added to a pre-chilled round bottom for each of the clone to be transformed. For *NANOG-EGFP*, *PL-SIN-EF1α-EGFP* and *pLKO.1-puro* plasmids, 1.5 µL was transferred, while for *psPAX* 0.9 µL and *pMD2.G* 2.5 µL were transferred to each of the reaction tube for collecting 5 µg of plasmid mix. Each reaction was gently mixed without pipetting and then incubated for 30 min. Then the transformation was heat-shocked for 3 min at 42° C before being place on ice for 5 min. 250 µl LB broth was added to each reaction and left on to shake horizontally at 1,200 rpm, for 1h at 37°C. 100 µL of each transformation reaction was then spread on pre-warmed LB-agar selective plate containing 50 µg/mL of Ampicillin and left in the incubator at 37°C overnight. After incubation, single colonies, from each plasmid of interest, were inoculated in 50 mL of LB-broth supplemented with Ampicillin 50 µg/mL (stock 100 µg/µL) and incubated overnight on the shaker at 37°C. Cell culture with visibly cloud bacteria were used for plasmid isolation using QIAGEN Plasmid Midi Kit 25 (12143). Isolated plasmids were then quantify using NanoDrop™ 8000 Spectrophotometer and stored in tris EDTA buffer (TE) at -20°C.

2.2.4 Transfection of HEK-293T human embryonic kidney epithelial cells and lentivirus production

Human HEK-293 embryonic kidney cells carrying the SV40 T-antigen (HEK-293T) were grown at 90% confluence in 4 mL of dedicated media DMEM 10% FCS and 1%L-Glu in T75 flasks. On the day of the transfection, 20 µL of Lipofectamine™ 3000 (Invitrogen) were incubated with 500 µL of OPTIMEM™ 1x (Gibco) for 30 min at room temperature. For the lentivirus production, 6 µg of the packing plasmid psPAX2 and 2 µg of the envelope pMD2.G were added. Finally, 8µg of the plasmid of interest (*NANOG-EGFP*, *EF1-α EGFP*) was added to the solution containing 500 µL of OPTIMEM™, the packaging and envelop, in which HEK-293T cells were incubated at 37°C in viral incubator for 12-15 hours in T25 flasks. On the second day medium was changed to a dedicated medium HEK-293T. The third day, the first viral fraction surnatant, was collected, filtered in 40 µm nylon, and 1 mL aliquots 1mL stored at -20°C. On the fourth day, the second viral fraction was collected at the same way and aliquots stored at -20°C.

2.2.5 Infection of prostate cancer cells

DU145 and PC3 cell were cultured to 60% confluency in 6 well plates. For cell infection, 1mL of cell dedicated medium combined with 1mL of viral fraction and 16 µL of hexadimethrine bromide solution (1 mg/mL in 0.9% w/v NaCl) was added to DU145 and PC3 cell and cells incubated at 37°C in the viral incubator. Cells transduced with lentivirus construct *NANOG-EGFP*, *EF1-α EGFP* were further grown in cell-line dedicated medium. The efficiency of the infection was assessed by measuring the expression of the Enhanced Green Fluorescent Protein (EGF) at 488 nm wavelength using a Carl Zeiss MicroBeam fluorescent microscope.

2.2.6 Isolation of DU145 *NANOG-EGFP* positive and DU145 *NANOG-EGFP* negative cell population using MoFlo™ flow cytometry-based cell sorting

Cells transfected with *NANOG-EGFP* lentivirus were enriched for EGFP+ and EGFP- populations by flow cytometry cell sorting. For this, cells were harvested as described above after which they were washed with Dulbecco's Phosphate-Buffered Saline (DPBS). Following detachment, equal amounts of medium was added, cells harvested and centrifuged for 5 min at 260 xg. Cells were resuspended in 1-2 mL of filter-sterilised sorting medium (EMEM supplemented with EDTA 3mM (Ambion), HEPES 25mM (Sigma-Aldrich), 99% Benzonase 1:5000 (Millipore) and 2% Pen/Strep (Corning)) and filtered through 40nm nylon to remove the clumps. EGFP+ and EGFP- were sorted based on green fluorescent using a Beckman Coulter MoFlo™ ADP cell sorter (488 nm emission wavelength band-pass filter 530/40).

2.2.7 Flow cytometric analysis of isolated *NANOG+* and *NANOG-* DU145 cells

Sorted EGFP+ and EGF- cells were centrifuged for 5 min at 300 xg and potential non-specific binding blocked by incubating cells with 50 µL of FcR blocking solution (MACS Miltenyil Biotech 130-059-901) for 15 min at room temperature. Cells were then incubated for 1h at room temperature with fixable LIVE/ DEAD™ Violet (Life Technologies L34955) viability stain (1 µL for 1mL of PBS) and with unconjugated JvGCRC-H61.3, EPCR mAb (Di Biase, 2020), in PBS (100µL, 1 µg/mL). Cells were washed three times with 2mL of PBS and centrifuged for 5 min at 300 xg. Cells were then incubated with an Alexa Fluor™ 633 conjugated polyclonal goat anti-mouse secondary antibody for 30 min protected from the light 5 µg/mL, (Invitrogen Life Technologies A21052). Cells were washed three times with 2mL of PBS and finally suspended in 300 µL of Isoton II™ sheath fluid (Beckman Coulter). Samples were analysed using unstained cells and cells incubated with secondary antibody alone as control. Data were acquired using a Beckman Coulter Gallios™ flow cytometer and analysed using Beckman Coulter Kaluza™ software.

2.2.8 Susceptibility of *NANOG-EGFP+* DU145 and NANOG-*EGFP+* PC3 cells to antibody-dependent cellular cytotoxicity (ADCC)

The mFcγRIV ADCC Report Bioassay is a biological relevant assay that can be used to measure the activity of mouse antibodies that specifically bind and it actives FcγRIV. For the commercial assay (mFcγRIV ADCC Reporter Bioassay, Core Kit, Promega, UK) used, Jurkat (T cell leukaemia) cells have been engineered to stably express the mouse Fc Gamma Receptor IV (FcγRIV) linked to the Nuclear Factor of Activated T-cells (NFAT) response element which drives the expression of luciferase activity on ligation of the receptor with antibody. The day before the experiment, *NANOG-EGFP+* DU145 and *NANOG-EGFP+* PC3 cells were seeded in 96 well-plate (20,000 cells/well and 25,000 cells/well respectively) and cultured overnight with dedicated culture medium at 37° C. The assay was performed following the manufacturer's protocol (Promega

M1211). Target cells were incubated with 15 µg/mL of JvGCRC-H61.3 (IgG2b) and JvGCRC-H599.5 (IgG2b) and mFcyRIV 'effector 'cells (Promega M115A) for 6h at 37°C 5% v/v CO_2. The resultant luciferase activity was detected by incubating cells with Bio- Glo™ reagent for 15 min and measuring the luminescence using a luminometer plate reader.

2.2.9 Evaluation of mRNA level of *PROCR* silencing in DU145

To confirm the efficiency of EPCR knockdown in prostate cancer cell line, total RNA was isolated from cells expressing pLKO.1-puro, shRNA1 target sequence GAATCACCTGAGGCGTTCAAA, shRNA4 target sequence GCAGCAGCTCAATGCCTACAA in 1 ug/mL of Puromycin in DU145 previously obtained in our laboratories (Di Biase, 2020) (harvested as described in paragraph 2.2.2) using the RNeasy® Mini Kit (cat. #74104 Qiagen) according to the manufacturer's protocol and stored at -80° C. RNA quantification and purity was obtained using NanoDrop™ Spectrophotometry. 1.5 µg was reverse transcribed into cDNA by incubating with 1µL of Promega Oligo-$(dT)_{15}$ primers in nuclease-free water (NFW) (Ambion #AM9937). Depending on the RNA concentration of each sample, each reaction was transferred in 0.5 mL Eppendorf tubes for a final volume of 10 µL. Samples were gently mixed and incubated for 5 min at 70°C in a thermal block, after which samples were incubated for 5 min on ice. In mean time, the second mix was prepared according to the table 2.1.

Table 2.1 - cDNA Master Mix. Reagents and volume required for the preparation of cDNA for single tube purchased form Promega company.

Reagents	Volume for single reaction
RT 5x Buffer	5 µL
Reverse Transcriptase	1 µL
RNasin	0.7 µL
dNTPs	1 µL
NFW	7.3 µL
Total	15 µL

Master Mix (15 µL) was added to each reaction tube mixed thoroughly by pipetting up and down gently. Tubes were incubated in a thermal block at 40°C for 60 min followed by inactivation of the reaction by incubating the tubes at 95°C for 5 min. Finally, cDNA samples were stored at -20°C until use.

2.2.10 Quantitative Real-Time PCR (qRT-PCR) amplification

Level of targeted mRNA (*PROCR*) were normalised with that of control housekeeping gene (*GUSB*). *GUSB* is a suitable housekeeping gene because its annealing temperature and the Ct are comparable with *PROCR* allowing to be analysed together. Forward and reverse primers for PROCR and GUSB were purchased from Merck (Table 2.2) and resuspended according to the manufacturers' recommendations by adding the relevant amount of nuclease-free water to achieve 100 pmol followed by vortex and mixing. The vials were left for 30 min to dissolve completely. A working concentration 10 pmol was obtained by diluting the stock concentration in nuclease-free water and stored at -20°C. Each cDNA sample was analysed in triplicate.

Table 2.2 - Primer specifications. Forward and reverse primers for PROCR Sigma-Aldrich and the housekeeping gene GUSB Eurofins used for all the qPCR experiments in this thesis.

Oligo Name	Oligo#	Tm°	GC%	Sequence (5'-3')
FH_*PROCR*	8815837094-60/0	57.7	45	TTCTCTTTTCCCTAGACTGC
RH_*PROCR*	8815837094-60/1	58.8	45	CATATGAAGTCTTTGGAGGC

Oligo Name	Oligo#	Tm°	GC%	Sequence (5'-3')
FH_*GUSB*	8812878372-80/0	53.7	52.9	ACTGAACAGTCACCGAC
RH_*GUSB*	8812878372-80/1	56.7	40	AAACATTGTGACTTGGCTAC

Table 2.3 – qPCR Master Mix. Reagents and the volume used for a single reaction.

Reagents	Volume for single reaction
iTA™ Universal SYBR® Green Supermix BioRad #172-5124	6.75 µL
Forward Primer (10 pmol)	0.5 µL
Reverse Primer (10 pmol)	0.5 µL
NFW	2.75 µL
Total	**10.5 µL**

1 µL of cDNA was added to 10.5 µL of master mix PCR tube and left on ice. Tubes were tightly closed and placed into QIAGEN's real time PCR cycler, Rotor-Gene Q. The reaction was then carried out using the primers temperature according to the primers condition. Acquisition on the green channel was set during the elongation step as shown in table 2.4.

Table 2.4 - PCR condition.

PCR condition	
Initial denaturation	5 minutes at 95°C
Denaturation	10 sec at 95°C
Annealing	15 sec at 58°C
Elongation	20 sec at 72°C
40 cycles	

2.2.11 Effect of EPCR knockdown on the expression 'stemness' genes

The differential effect of EPCR knockdown (shRNA1 and shRNA4) in DU145 cells on the expression of embryonic and cancer stem cell genes was assessed using primers purchased from Eurofins (Table 2.5) and Merck (table 2.6). Gene expression was compared to that of the pLKO.1 control (empty vector) (Di Biase, 2020).

Table 2.5 - Primers used for the analysis of ESCs marker.

Stem Cell makers	
Oligo Name	**Sequence (5'-3')**
FH_*NANOG*	CCAGAACCAGAGAATGAAATC
RH_*NANOG*	TGGTGGTAGGAAGAGTAAAG
FH_*OCT4*	GATCACCCTGGGATATACAC
RH_*OCT4*	GCTTTGCATATCTCCTGAAG
FH_*SOX2*	ATAATAACAATCATCGGCGG
RH_*SOX2*	AAAAAGAGAGAGGCAAACTG

Table 2.6 - Primers for used for the analysis of CSCs marker.

Cancer stem cell markers				
Oligo Name	**Oligo#**	**Tm°**	**GC%**	**Sequence (5'-3')**
FH_*PROM1*	8815665218-10/0	59.3	40	AAGCATTGGCATCTTCTATG
RH_*PROM1*	8815665218-10/1	59.4	40	TTTGCTCTGGAGTTTCATTC
FH_*CD44*	8815665218-20/0	56.6	45	TTATCAGGAGACCAAGACAC
RH_*CD44*	8815665218-20/1	61.5	40	ATCAGCCATTCTGGAATTTG
FH_*CD24*	8815665218-30/0	56.4	42.8	CAGTAGTCTTGATGACCAAAG
RH_*CD24*	8815665218-30/1	58.4	40	ACAGCATTCTGGAATAAAGC

2.2.12 Analysis of EPCR silencing for EMT

The influence of EPCR knockdown (shRNA1 and shRNA4) in DU145 cells on the expression of EMT-related genes was assessed using primers from Eurofin (Table 2.7) and the EMT transcription factors (EMT-TFs) purchased from Merck. Gene expression was compared to that of pLKO.1 control (empty vector) (Di Biase, 2020).

Table 2.7 - Primers used for the analysis of EMT-related gene expression after EPCR silencing in DU145 cells.

EMT-markers	
Oligo Name	**Sequence (5'-3')**
FH_*VIM*	GAGAACTTTGCCGTTGAAGC
RH_*VIM*	GCTTCCTGTAGGTGGCAATC
FH_*CDH1*	TGCCCAGAAAATGAAAAAGG
RH_*CDH1*	GTGTATGTGGCAATGCGTTC
FH_*FN1*	CAGTGGGAGACCTCGAGAAG
RH_*FN1*	TCCCTCGGAACATCAGAAAC

EMT-TFs				
Oligo Name	**Oligo#**	**Tm°**	**GC%**	**Sequence (5'-3')**
FH_*SNAI2*	8815837094-10/0	59.7	42.8	CAGTGATTATTTCCCCGTATC
RH_*SNAI2*	8815837094-10/01	59.9	50	CCCCAAAGATGAGGAGTATC
FH_*SNAI1*	8815837094-20/0	55.4	40.9	CTCTAATCCAGAGTTTACCTTC
RH_*SNAI1*	8815837094-20/01	55.2	55.5	GACAGAGTCCCAGATGAG
FH_*ZEB1*	8815837094-20/0	59.3	40	AAAGATGATGAATGCGAGTC
RH_*ZEB1*	8815837094-20/01	60.8	40	TCCATTTTCATCATGACCAC
FH_*TWIST1*	8815837094-50/0	57.9	40.9	CTAGATGTCATTGTTTCCAGAG
RH_*TWIST1*	8815837094-50/01	60.9	40	CCCTGTTTCTTTGAATTTGG

2.2.13 *In silico* analysis of publicly available RNA-seq datasets

The Gene Expression Profiling Interactive Analysis (GEPIA) http://gepia.cancer-pku.cn/ webserver was used to study the expression of EPCR across different types of cancer (versus normal tissue). For this, two data sets from The Cancer Genome Atlas (TCGA) and one from the international Cancer Genome Consortium (ICGC) were available from cBioPortal platform (Cerami et al., 2012a) (Gao, J. et al., 2013). Data were transformed in $\log_2$ scale and sorted for EPCR (*PROCR*) low expression and high expression. The expression of different genes analysed in correlation with EPCR low and high level are shown in Table 2.8. Genes of interest sorted for EPCR low and high were then analysed as heat map using Morpheus matrix software https://software.broadinstitute.org/morpheus/ .

Table 2.8 - Gene for correlation analysis for EPCR.

Epithelial-mesenchymal transition (EMT) markers	• CDH1: E-cadherin • CDH2:N-cadherin • VIM: Vimentin • FN1:Fibronectin • FGF: Fibroblastic growth factor • EGF: Epidermal growth factor • TGFB: Transforming growth factor • YAP1:Yes-association protein
Epithelial-mesenchymal transition transcription factor (EMT-TFs) markers	• SNAI1:Snail Zinc finger protein 1 • SNAI2:Slug Zinc finger protein 2 • ZEB1: Zinc Finger E-box-binding homoebox1 • ZEB2: Zinc Finger E-box-binding homeobox 2 • TWIST: *Twist-related* protein
Embryonic stem cell markers	• NANOG1:Homeobox protein NANOG • SOX: (Sex determining region Y) box2 • PUO5F: Octamer-binding transcription factor 4 OCT4 • KLF5:Kruppel-like factor 5 • LIN28: Lin-28 homolog A
Cancer stem cell markers	• ITGA2: Integrin α2 β1 • PROM1:CD133 • CD44 • CD24 • ALDHA1: Aldehyde dehydrogenase 1 family • TP63: p63
Prognostic prostate markers	• KLK3: Prostate-Antigen-Specific (PSA) • AR: Androgen receptor • PSAP: Prostatic-acid phosphates • KRT18:Keratin18 • KRT8:Keratin8 • KRT14:Keratin14 • MUC1:cell surface associated • LY6E: Stem cell Antigen Sca-1 • TACSTD2: Trop-2 • SYP: Synaptophysin • CHGA: Chromogranin A • ENO2: Enolase2

2.2.14 *In silico* analysis of publicly clinical data

Clinical data from the same three consortia (2.13 paragraph) were analysed based on EPCR low and high expression for age at diagnosis, Neoplasm Disease Lymph Node Stage (from American Joint Committee on Cancer Codes) and the Cancer Tumour stage code.

2.2.15 Statistical Analysis

All statistical analysis was performed using GraphPad Prism 8.4.2 software (GraphPad, Inc, USA). For experiments where two groups were compared, a two-tailed Student's t-test was performed. For comparison of three or more groups, a one-way ANOVA was performed followed by post-hoc Sidak's multiple comparison tests. For non-parametric data, Kruskal-Wallis followed by uncorrected Dunn's test was used. Unless otherwise stated, histogram columns represent the mean and error bars indicate the standard deviation (SD). The standard error of mean (SEM) was used to displayed variance in the polysome results. The number of replicate (n) for each experiment is stated in the figure legend. The data is considered to be statistically significant if $P < 0.05$ and are indicated in the figure legends by asterisks. For *in silico* data statistical analysis, comparing two variables was performed, and using simple liner regression for model the relationship between two variables by using the line-of-best-fit was created. Pearson correlation coefficient, r, was then calculated. Finally, for comparison of two or more group contingency table was used and chi square with Yates' correction was performed.

2.3 Results

2.3.1 EPCR is differentially expressed in DU145 and PC3 human prostate cancer cell lines

To assess whether EPCR is a good candidate target for monoclonal antibody-based therapy against prostate cancer, EPCR expression was analysed in DU145 and PC3 cells line by flow cytometry. Cells were initially identified based on the side (SSC) and forward light scatter (FSC) which provide information on granularity and cell size respectively. Viable single cells were gated based on staining with LIVE/DEAD™ fixable Violet Dead Cell reagent and EPCR expression identified by staining using the JvGCRC-H61.3 mAb. Results demonstrated that EPCR is differentially expressed in DU145 and PC3, with higher expression by DU145 cells (Fig 2.1).

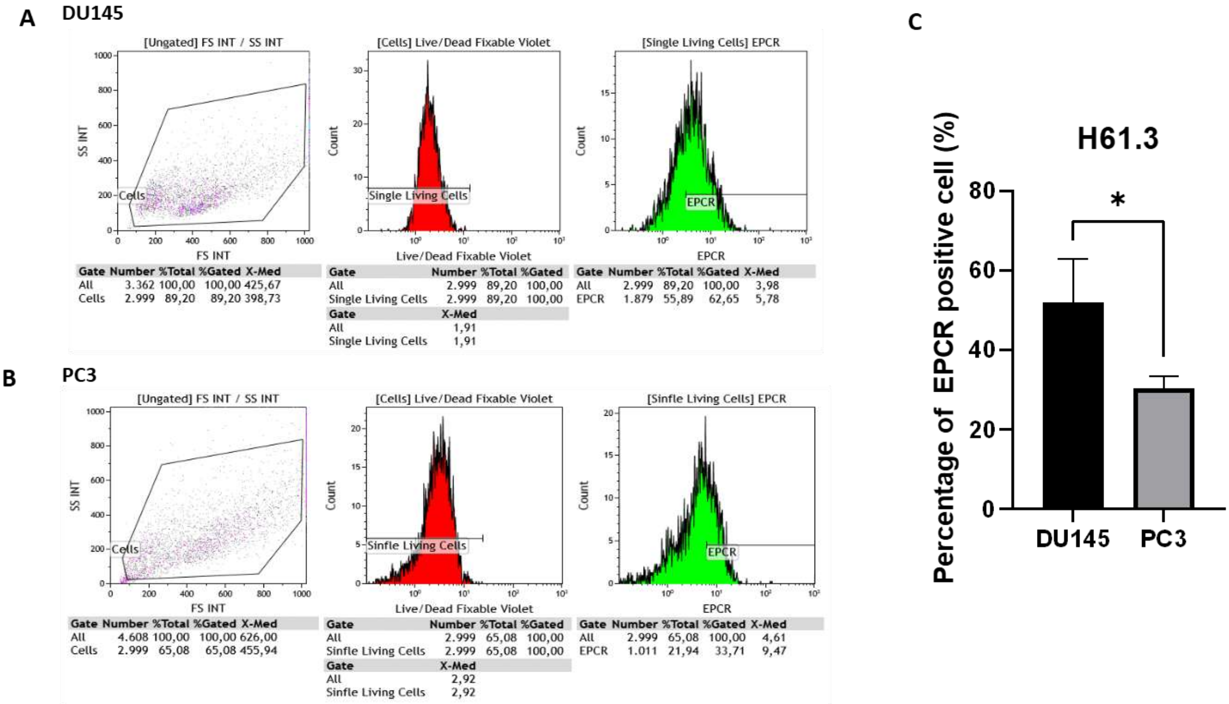


Figure 2.1 - Expression of EPCR by DU145 and PC3 human prostate cancer. **A-B** Panels from left to right represent the gating of DU145 and PC3 cells. Single, viable cells were stained for the expression of EPCR using the unconjugated mouse JvGCRC-H61.3 mAb, (1µg / mL in PBS). **C** Percentage of DU145 and PC3 cells expressing ECPR positive cells to EPCR. Data presented as mean ± SD of three independent experiment, (*p=0.0485; unpaired t test; N=3). Data analysed using Kaluza™ 3.1 software.

2.3.2 Analysis of EPCR expression in *NANOG+* and *NANOG*- DU145 prostate cancer cells

To confirm the expression of hEPCR on prostate cancer stem cells (PCSCs), DU145 were transfected with a lentivirus construct containing the NANOG promoter that controls the expression of a fluorescent protein (EGFP). For this experiment 600,000 cell /samples EGFP+ and EGFP- were isolated using the Beckman Coulter MoFlo™ cell sorter and hEPCR expression isolated cells was examined using flow cytometry. Cells were initially identified based on the side (SSC) and forward light scatter (FSC) which provide information on granularity and cell size respectively. Viable, single cells were gated based on staining with LIVE/DEAD™ fixable Violet Dead Cell reagent and EPCR expression identified by staining unfixed cells using the JcGCRC-H61.3 mAb.

Analysis was conducted considering the expression of EPCR in terms of median fluorescence intensity (MFI) and percentage of cells positive for EPCR. Data obtained from three independent experiments demonstrated that there is no statistically significant difference in the intensity of EPCR expression from isolated DU145 *NANOG-EGFP+* and *NANOG-EGFP*- DU145 cells. In contrast, a higher percentage of NANOG-EGFP- DU145 cells express EPCR. Although there is no statistical significant difference in the EPCR expression between cell population *NANOG-EGFP-* and *NANOG-EGFP+* in DU145 the results obtained show a trend in the expression (Fig. 2.2). Possibly, the number of the experiments (n=3) is not sufficient to see a difference between *NANOG-EGFP+* and *NANOG-EGFP-.*

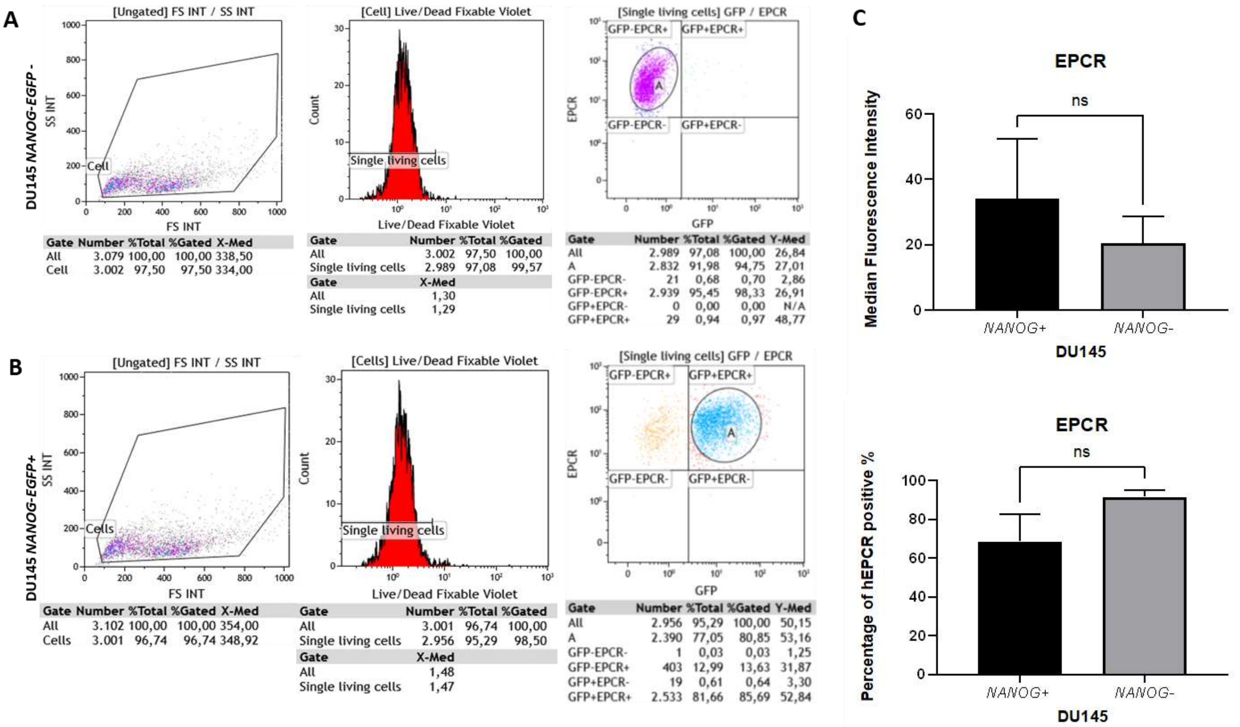


Figure 2.2 - Expression of hEPCR on *NANOG-EGFP-* and *NANOG-EGFP+* DU145. Panels from left bottom to right represent the of gating of *NANOG-EGFP+ NANOG-EGFP–* in DU145 cell population. Single, viable cells were stained for EPCR using unconjugated JvGCRC-H61.3 mAb, (1µg / mL in PBS).**C** Median fluorescent intensity and percentage of DU145 and PC3 cells expressing EPCR. Data presented as mean ± SD of three independent experiment (ns=non-significant; *p-value=0.0485; unpaired t test; n=3). Data analysed using Kaluza™ 3.1 software.

2.3.3 Capacity of EPCR monoclonal antibodies to trigger antibody dependent cytotoxicity (ADCC) *in vitro*

The commercial (mFcγRVI ADCC Reporter Bioassay (Core Kit, Promega, UK) was used to study the potential capacity of the JvGCRC-H61.3 and JvGCRC-599.5 mAbs to mediate antibody-dependent cell-mediated cytotoxicity (ADCC) against *NANOG-EGFP+* DU145 and PC3 cells.

Serial dilutions of the JvGCRC-H61.3 and JvGCRC-H599.5 IGg2b mAbs identified 3X (15 µg/mL final concentration) which mediate the activation of mFcyRIV with luciferase activity being detected for both cell lines. Particularly, for the *NANOG-EGFP+* PC3 cells, JvGCRC-H61.3 shows a higher fold induction with low EC_{50} when compared with the EC_{50} for the clone JvGCRC-599.5. In contrast, there was no difference in the activity of the two mAbs when using *NANOG-EGFP+* DU145 cells (Figure 2.3, Table 2.1).

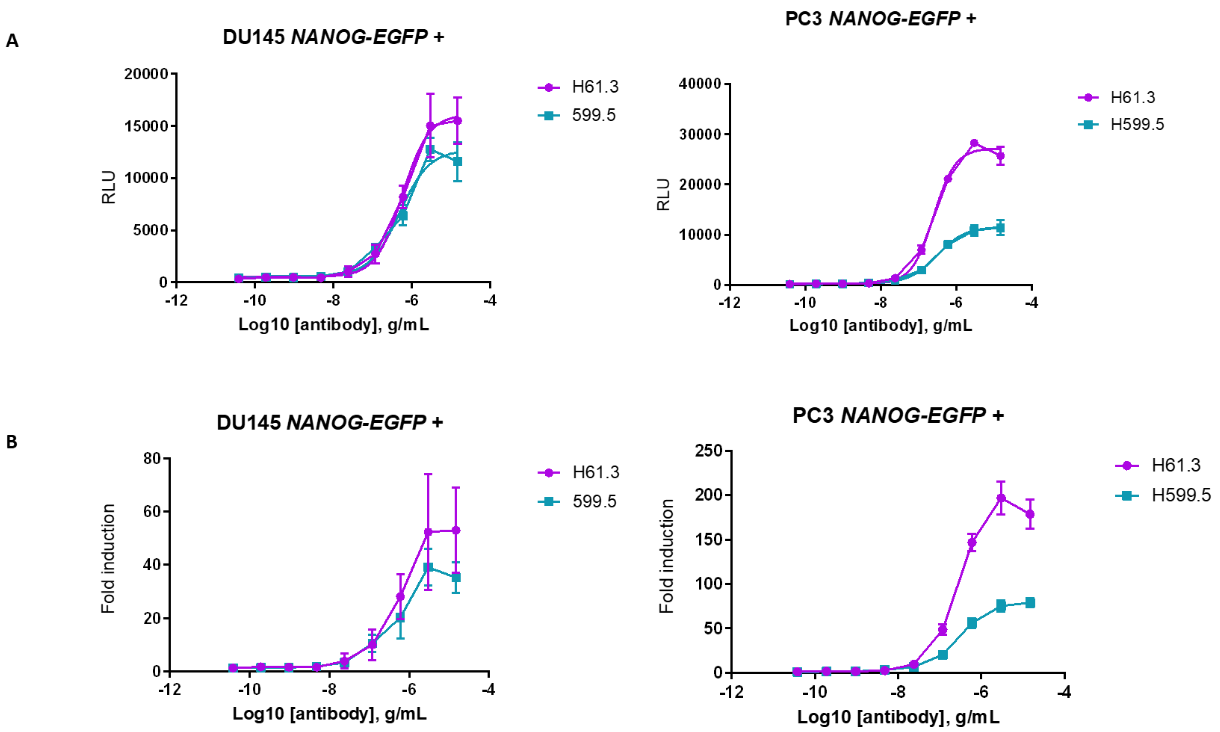


Figure 2.3 - Triggering of mFcγRIV reporter cells by JvGCRC-H61.3 and JvGCRC-599.5 mAbs. Effects of increasing concentrations of the JvGCRC-H61.3 (IgG2b) and JvGCRC-599.5 (IgG2b) mAbs on the luminescent readout of the mFcγRIV effector cell reporter assay using *NANOG-EGFP+* DU145 and PC3 target cells. Each experiment was run in triplicates in 10-point dilution series. BioGlo™ reagent was added and samples read using a luminometer. Luminescence data were fitted using log(agonist) vs. response – variable slope (four parameters) curve using GraphPad Prism 8.4.2 software n=1. Result is expressed in fluorescence activity RLU (A) or as fold induction (B).

Table 2.9 - Half maximal effective concentration (EC_{50}) and maximum fold change. Table showing the antibody against EPCR tested (JvGRC-H61.3 and JvGRC-H599.5), the concentration of the antibody used to induce a response halfway between the baseline (EC_{50} µg / mL) and the maximum effect generated (Maximum Fold Induction) in NANOG-*EGFP+* DU145 and PC3 cells.

Antibody tested	EC_{50} (µg/mL)		Maximum Fold Induction	
	DU145	PC3	DU145	PC3
JvGRC-H61.3	0.583	0.253	53.0	178.9
JvGRC-H599.5	0.489	0.31	35.2	79.0

2.3.4 EPCR knockdown in DU145 cells.

RT-qPCR was used to assess the efficiency of silencing by analysing the endogenous level of *PROCR* mRNA level in total RNA samples. Significant reduction of EPCR expression was seen in two knockdowns with the efficiency of knockdown for shRNA1 being approximately 93% and that for shRNA4 approximately 87% (Figure 2.4).

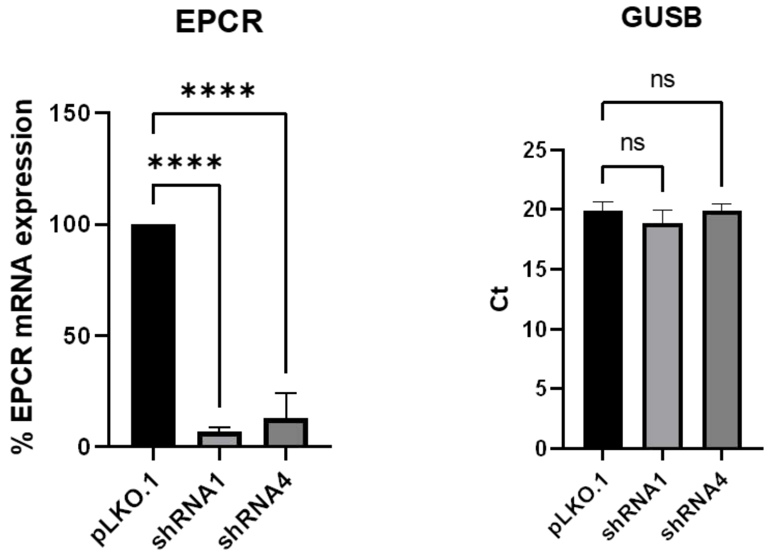


Figure 2.4 - Quantitative PCR on mRNA samples in DU145 of control pLKO.1 and EPCR shRNAs. Values were normalised on GUSB mRNA expression in all cell lines. Statistical analysis was carried out using one-way ANOVA with post-hoc Sidack's multiple comparison test against the empty vector (pLKO.1), P value was shown. All the data shown are the mean ± SD of three independent experiment (n=3) error bars represent S.D. (****= p-value <0.0001, ns=non-significant).

2.3.5 Effect of EPCR knockdown on the expression of 'stemness' genes

Aberrant expression of specific embryonic stem cell markers such as NANOG, OCT4 and SOX2 have been reported in many types of human cancer including breast, brain, cervix, head and neck, prostate, oral, ovary and others (Wang, Gang et al., 2018) (Gong et al., 2015). In order to determine if EPCR could be used as a potential marker to target cancer stem cells, DU145 cells EPCR silenced were analysed for the expression of embryonic stem cell markers NANOG, OCT4 (POU5F) and SOX2 using quantitative RT-qPCR. Results shown that NANOG expression is decreased in both knockdowns whereas OCT4 expression is decreased in one knockdown compared with the control. There is no statistical difference in expression of SOX2 when compared with the empty vector (Fig 2.5).

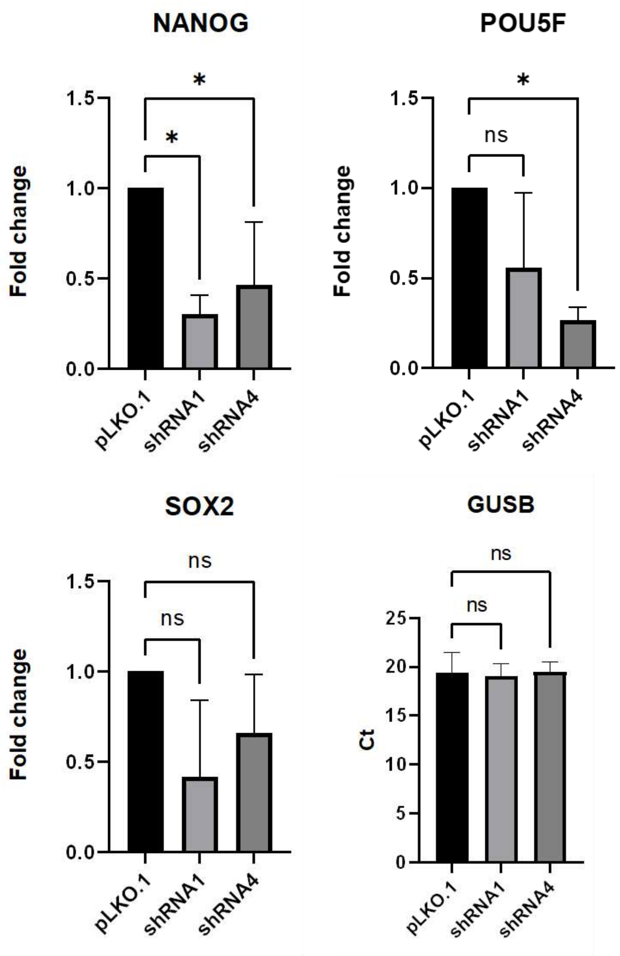


Figure 2.5 - Effect of EPCR knockdown on the expression of embryonic stem cell markers by DU145 cells. Values were normalised on GUSB mRNA expression in all cell lines. Statistical analysis was carried out using one-way ANOVA with post-hoc Sidack's multiple comparison test, P value was shown. All the data shown are the mean ± SD of three independent experiment (n=3) error bars represent S.D. (*=p-value=0.0128; *=p-value=0.0199; *=p-value=0.01830; ns=non-significant).

2.3.6 Effect of EPCR knockdown on the expression of cancer stem cell genes

To determine whether there is a correlation between the expression of EPCR and cancer stem cell markers, we analysed the influence of EPCR knockdown on the expression of cancer stem cells markers which are shown to be responsible of cancer progression (*PROM1*(CD133), CD24 and CD44). The analysis of mRNA expression lever in three independent experiments show that EPCR knockdown has no significant effect on CD44 expression, but that it decreases *PROM1* (CD133) expression and increases CD24 expression (Figure 2.6)

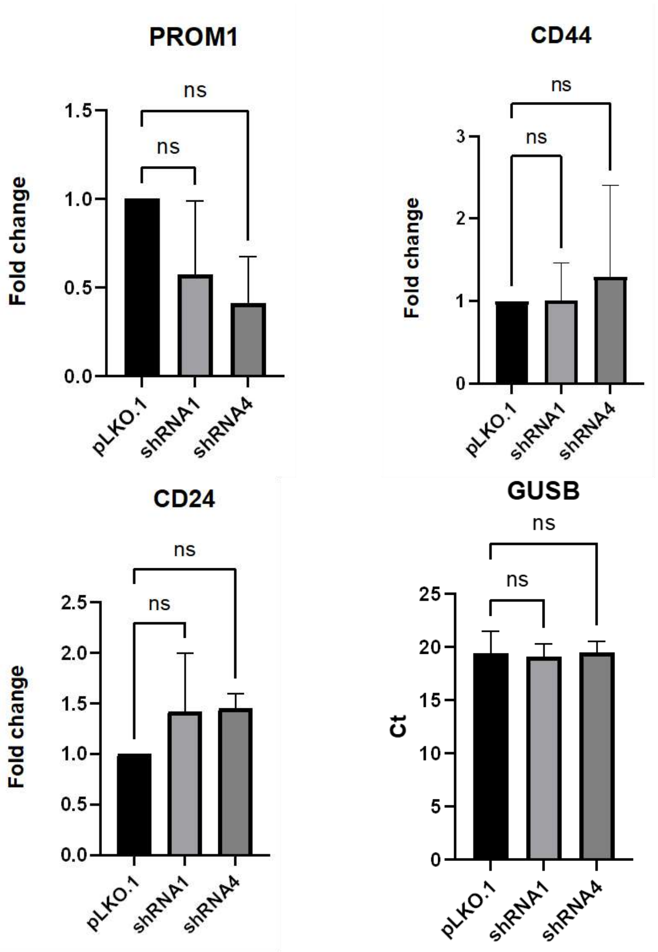


Figure 2.6 - Effect of EPCR knockdown on the expression of cancer stem cell genes. Values were normalised on GUSB mRNA expression in all cell lines. Statistical analysis was carried out using one-way ANOVA with post-hoc Sidack's multiple comparison test, P value was shown. All the data shown are the mean ± SD of three independent experiment (n=3) error bars represent S.D (ns=non-significant).

2.3.7 Effect of EPCR knockdown on the expression of EMT genes

The effect of EPCR knockdown on the expression of EMT markers in DU145 was also evaluated. Results demonstrated that EPCR knockdown had no effect on the expression of epithelial marker E-cadherin or the fibroblast marker fibronectin. With regards to vimentin, knockdown of EPCR with shRNA1 had no effect on expression, whereas knockdown with shRNA4 increased the expression (Figure 2.7A). The effect of EPCR knockdown in the expression of specific EMT transcription factors (EMT-TFs) was also analysed. EPCR knockdown had no significant effect on SLUG or ZEB1 expression but did reduce the expression of TWIST1. Knockdown of EPCR with shRNA1 had no effect on the expression of SNAI1, whereas knockdown with shRNA4 increased the expression (Figure 2.7B).

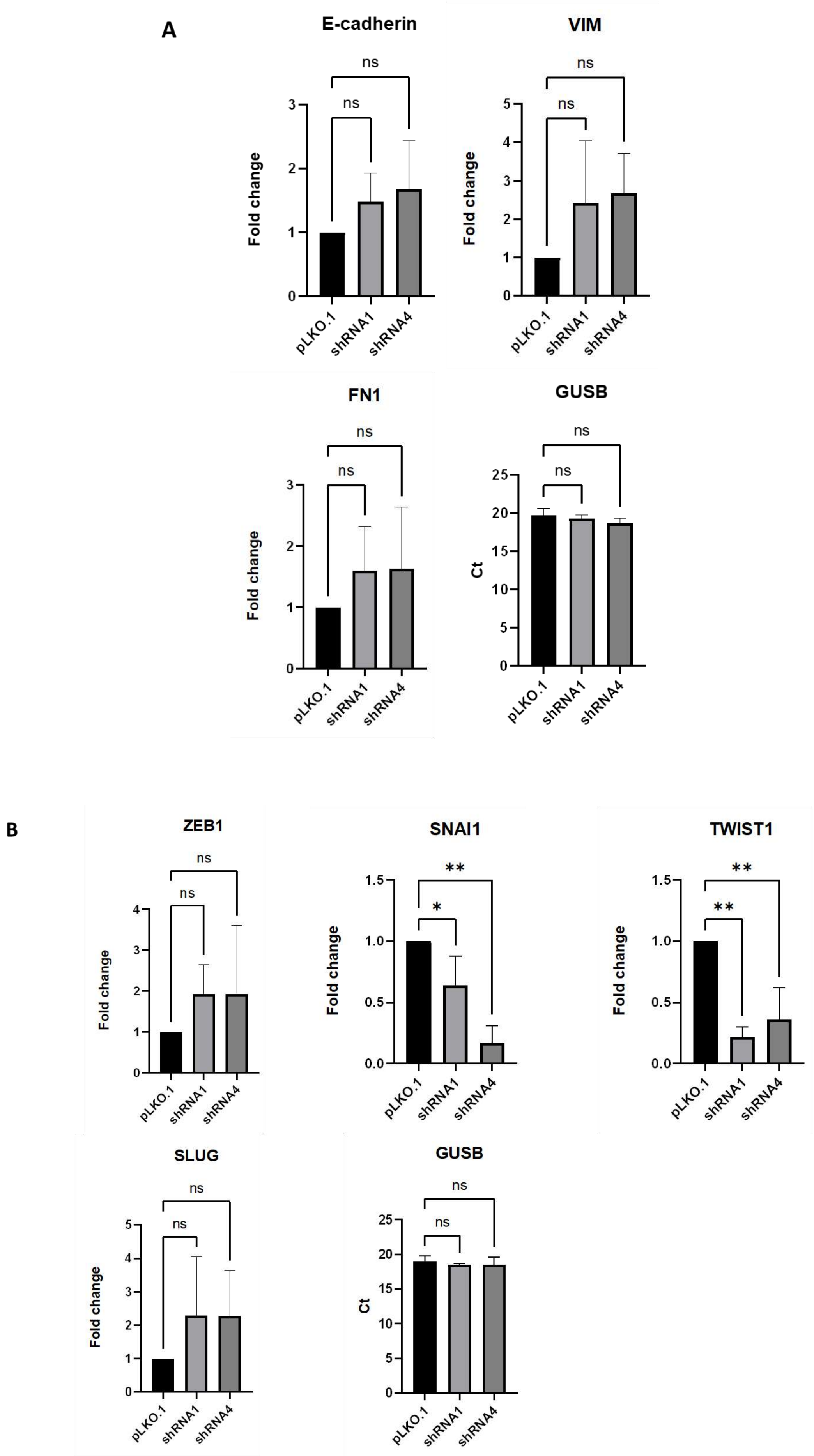

A
E-cadherin
VIM
FN1
GUSB
Fold change
Ct
ns
pLKO.1
shRNA1
shRNA4
B
ZEB1
SNAI1
TWIST1
SLUG
GUSB
Fold change
Ct
ns
*
**

Figure 2.7 - Effect of EPCR knockdown on the expression of EMT markers. A EMT markers. **B** EMT transcription factors (EMT-TFs). Values were normalised on GUSB mRNA expression. Statistical analysis was carried out using one-way ANOVA with post-hoc Sidack's multiple comparison test, P value was shown. All the data shown are the mean ± SD of three independent experiment (n=3) error bars represent S.D. (*=p-value=0.0319; **=p-value=0.0014; **=p-value=0.0015; **=p-value=0.0042 ns=non-significant).

2.3.8 Correlation between EPCR (PROCR) expression and the embryonic stem cell (ESC) markers in patients with prostate cancer

Datasets from three consortia (TCGA2018, TCGA2015, ICGC) were used for the analysis (Table 2.11). Data of the sample type were provided for each consortium therefore, TCGA 2018 and ICGC the samples type is primary tumour while the data TCGA 2015 samples type is metastatic prostate adenocarcinoma. *In silico* analysis using datasets from three different consortia show a positive correlation for Oct4 (*POU5F*) and *SOX2* in TCGA 2018 whereas a negative correlation is seen in the other two datasets (TCGA2018 and ICGC). Different results are seen for *NANOG,* whereas a positive correlation is seen the TCGA 2015 datasets while negative correlation is seen in the ICGC and TCGA 2018 datasets (Figure 2.8).

Table 2.10 - Data links from three prostate cancer consortia

Consortia	**Reference data from cBioPortal**
TCGA2018	https://www.cbioportal.org/study/summary?id=prad_tcga_pan_can_atlas_2018
TCGA2015	https://www.cbioportal.org/study/summary?id=prad_tcga_pub
ICGC	https://www.cbioportal.org/study/summary?id=prostate_dkfz_2018

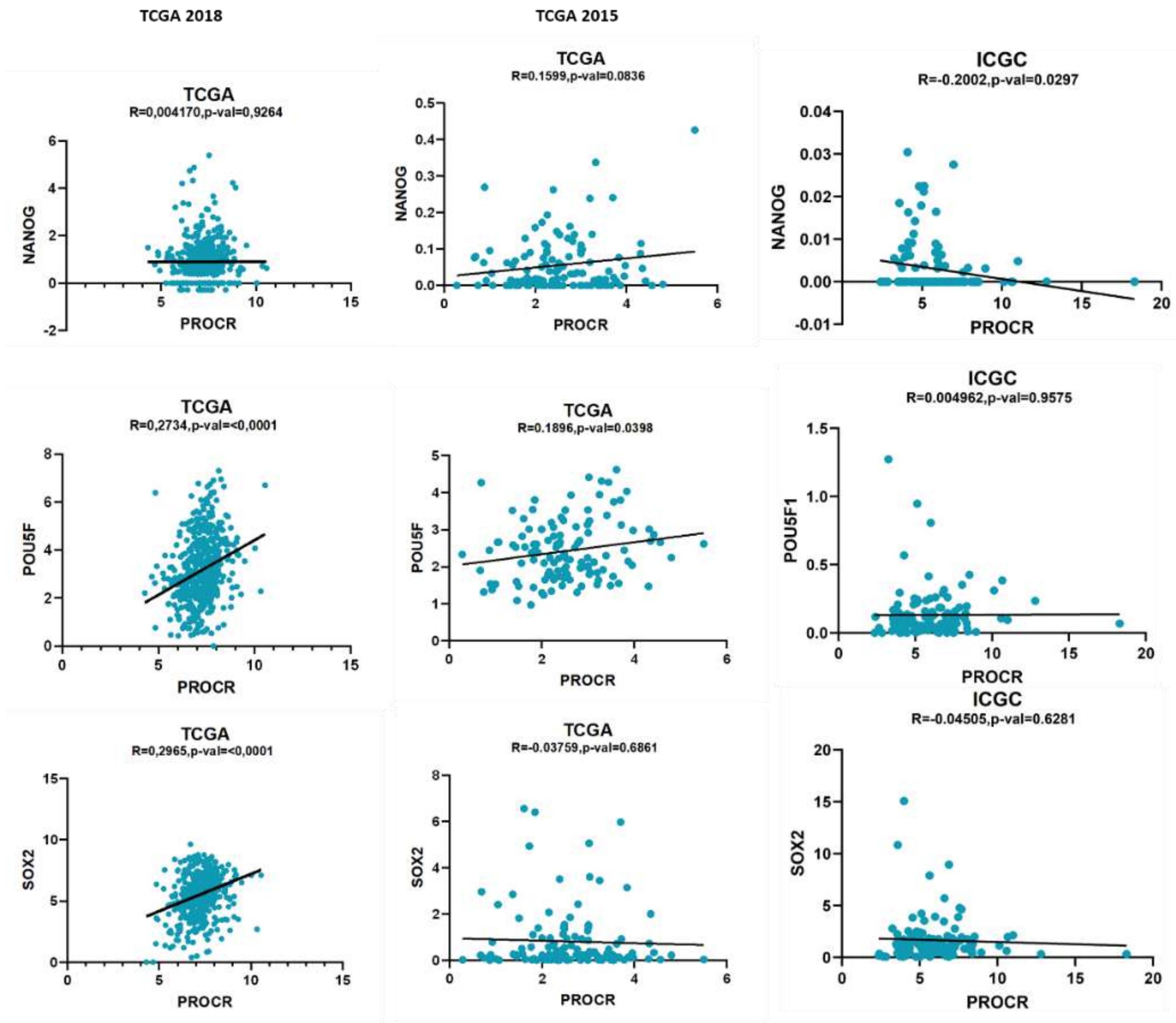


Figure 2.8 - Correlation between EPCR (*PROCR*) expression and the expression of embryonic stem cell (ESC) markers in patients with prostate cancer. Analysis was obtained by sorting the gene expression based on EPCR level low and high after data had been $\log_2$ transformed. Correlation analysis using GraphPad prism and simple linear regression to find the best line that predicts Y (ESCs marker) with X EPCR (*PROCR*). Correlation is calculated with Pearson's coefficient, r values range between -1 to 1.

2.3.9 Correlation between EPCR (*PROCR*) expression and the expression of cancer stem cell (CSC) markers in patients with prostate cancer.

In silico experiments using datasets from three different consortia were used to identify potential correlations between the expression of EPCR and cancer stem cell markers (Table 2.8). The expression of CD133 *(PROM1)* shows a positive correlation with EPCR expression in TCGA2018 dataset, but not in the TCGA 2015 and ICGC datasets, For CD44, there appears to be a positive correlation in the TCGA 2018 and TCGA 2015, but not in the ICGC dataset. CD24 shows a negative correlation in the TCGA 2018 and a positive correlation in the TCGA2015 datasets, with no data being available in the IGCG dataset (Figure 2.9).

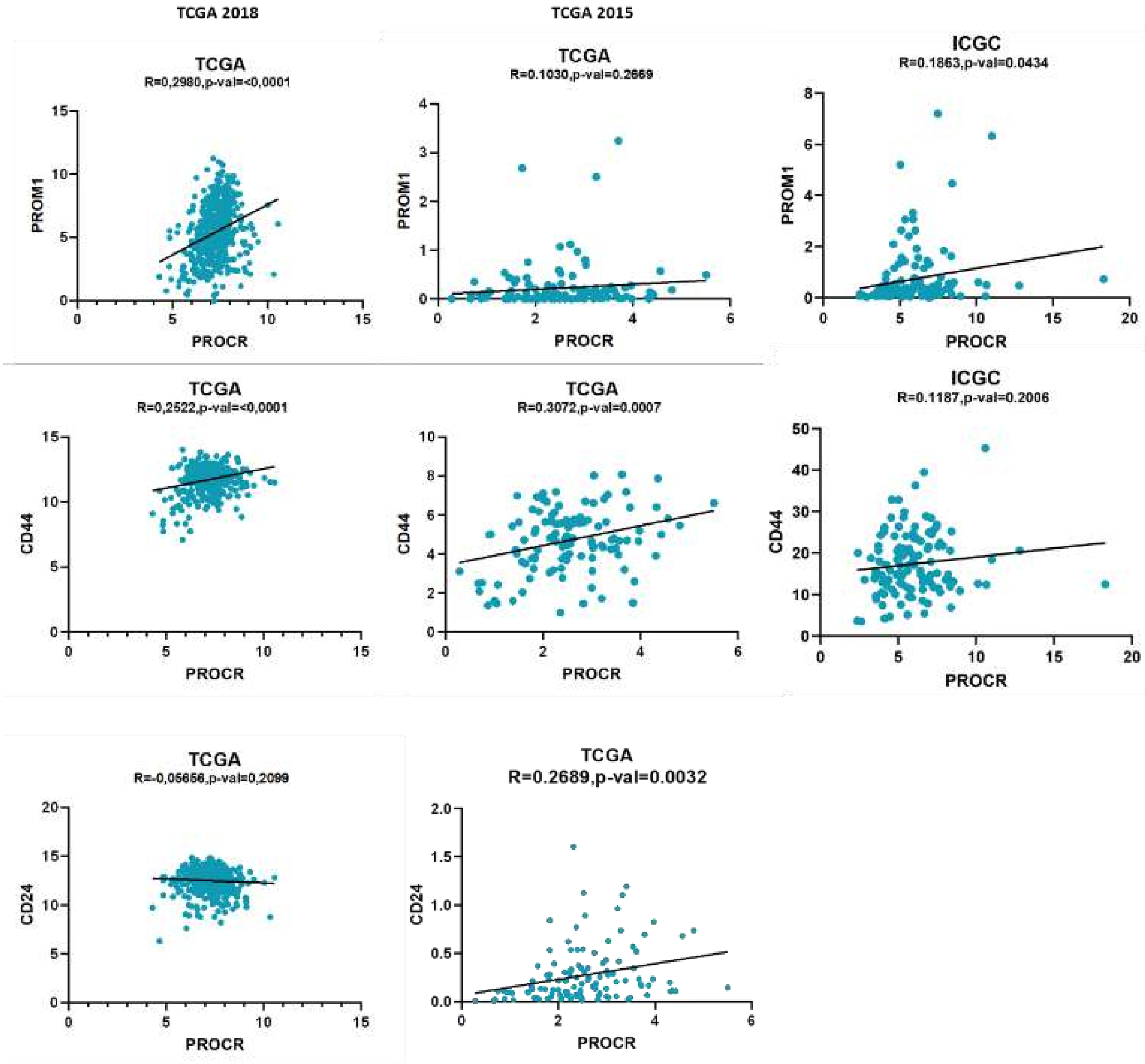


Figure 2.9 - Correlation between EPCR (*PROCR*) expression and the expression of cancer stem cell (CSC) markers in patients with prostate cancer. Analysis was obtained by sorting the gene expression based on EPCR level low and high after transformed in log_2. Correlation analysis using GraphPad prism and simple linear regression to find the best line that predicts Y (ESCs marker) with X EPCR (*PROCR*). Correlation is calculated with Pearson's coefficient, r values range between -1 to 1.

2.3.10 Correlation between EPCR (*PROCR*) expression and the expression of epithelial mesenchymal transition marker (EMT) markers in patients with prostate cancer.

In silico analysis using datasets from three different consortia were used to identify potential correlations between the expression of EPCR and the expression of factors that are commonly used for studying the mechanism of EMT in cancer cells and the expression of EMT transcription factors. The analysis demonstrated that the expression of vimentin and fibronectin (*FN-1*) is positively correlated with EPCR in the three datasets, whereas there is no correlation with E-cadherin (*CDH1*) in the TCGA2018 and TCGA 2015 datasets, but a negative correlation in the ICGC dataset (Figure.2.10).

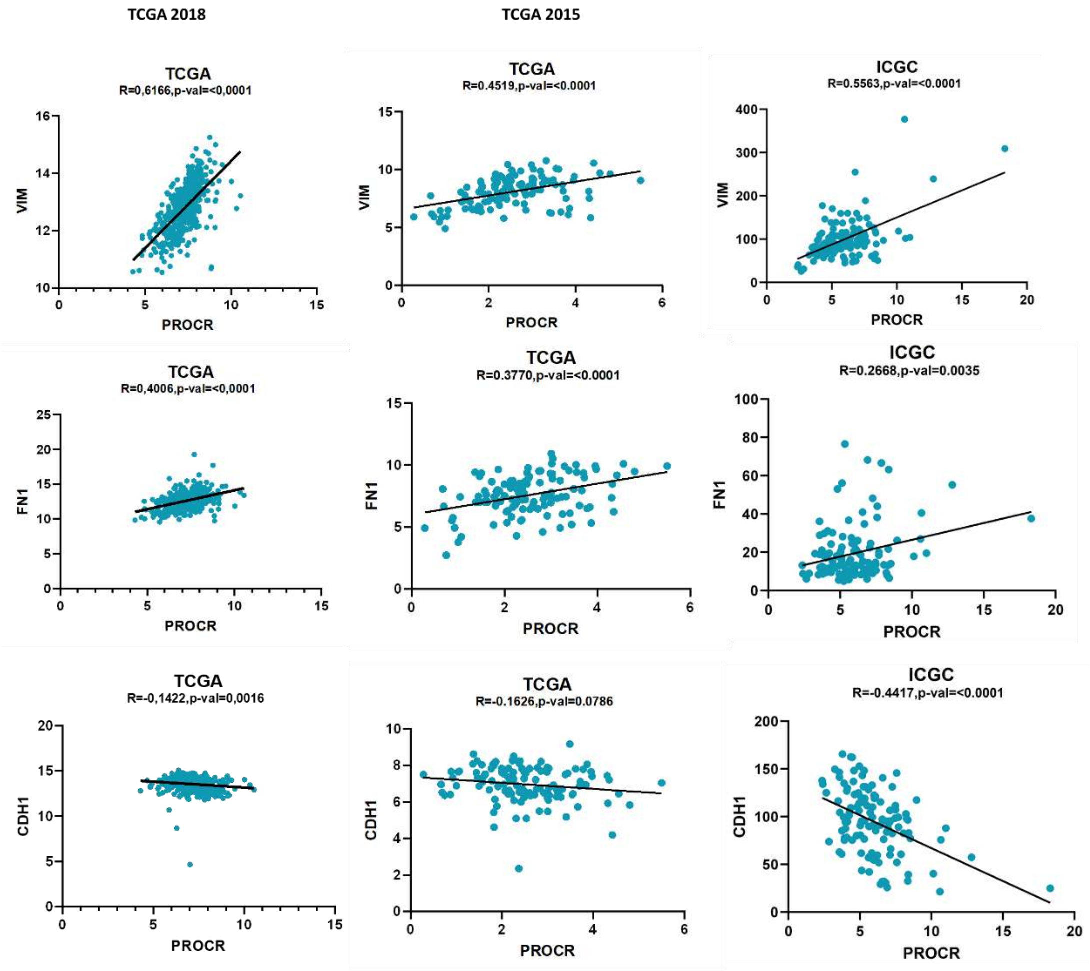


Figure 2.10 - Correlation between EPCR (*PROCR*) expression and the expression of epithelial mesenchymal transition (EMT) markers in patients with prostate cancer. Analysis was obtained by sorting the gene expression based on EPCR level low and high after transformed in $\log_2$. Correlation analysis using GraphPad prism and simple linear regression to find the best line that predicts Y (ESCs marker) with X EPCR (*PROCR*). Correlation is calculated with Pearson's coefficient, r values range between -1 to 1.

The analysis for EMT-TFs show that TWIST1 is positively correlates with EPCR in TCGA 2018, but not correlated in the TCGA2015 and ICGC dataset. *SNAI1, SNAI2, ZEB1* and *ZEB2* expression is positively correlated with EPCR expression in all three datasets. Morpheus matrix visualisation software was used to generate and display the all the pairwise correlations for the set of variables analysed in the TCGA 2018 dataset (Figure 2.11 and Table 2.8).

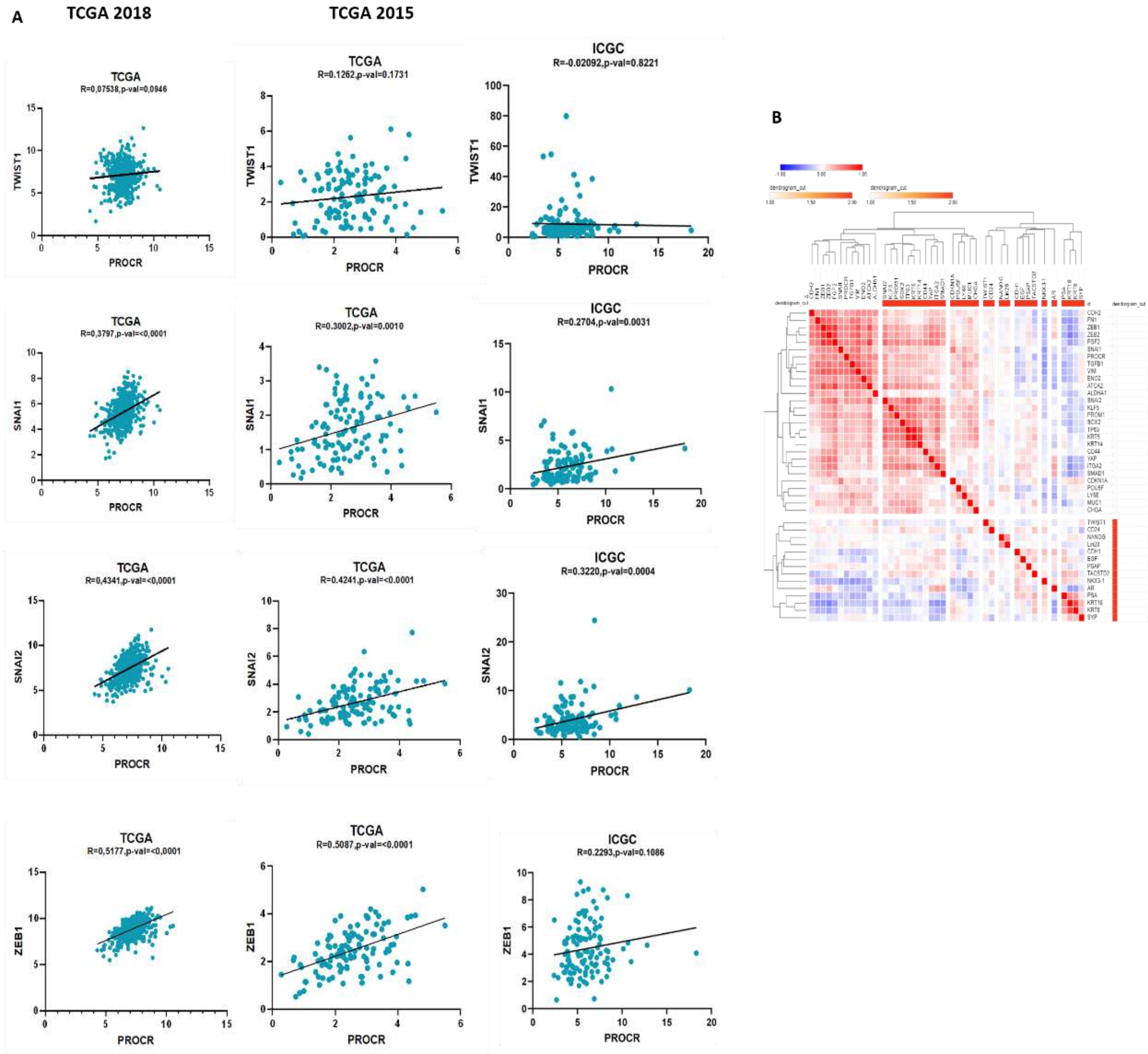


Figure 2.11 - Correlation between EPCR (*PROCR*) expression and the expression of epithelial to mesenchymal transition (EMT) transcription factors (EMT-TFs) in patients with prostate cancer. A. Analysis was obtained by sorting the gene expression based on EPCR level low and high after transformed in $\log_2$. Correlation analysis using GraphPad prism and simple linear regression to find the best line that predicts Y (ESCs marker) with X EPCR (*PROCR*). Correlation is calculated with Pearson's coefficient, r values range between -1 to 1. **B.** Correlation matrix using Morpheus-Broad Institute open-source software. The red boxes represent variables that have positive correlation and blue boxes negative correlation. The darker the box is, the closer the correlation is to negative or positive to 1. Dendrogram was obtained to show the hierarchical relationship between the variables. (https://software.broadinstitute.org/morpheus).

2.3.11 Relationship(s) between EPCR (*PROCR*) expression and clinical features in patients with prostate cancer.

In order to investigate a potential role for EPCR in cancer therapy, patients included in the TCGA2018, TCGA2015 and ICGC datasets obtained from cBioPortal were sorted for high and low expression of EPCR and analysed based on the age of diagnosis and tumour stages (see Table 2.8). Data presented below are from TCGA2018 the data from other consortia are included in the Appendix chapter 2.

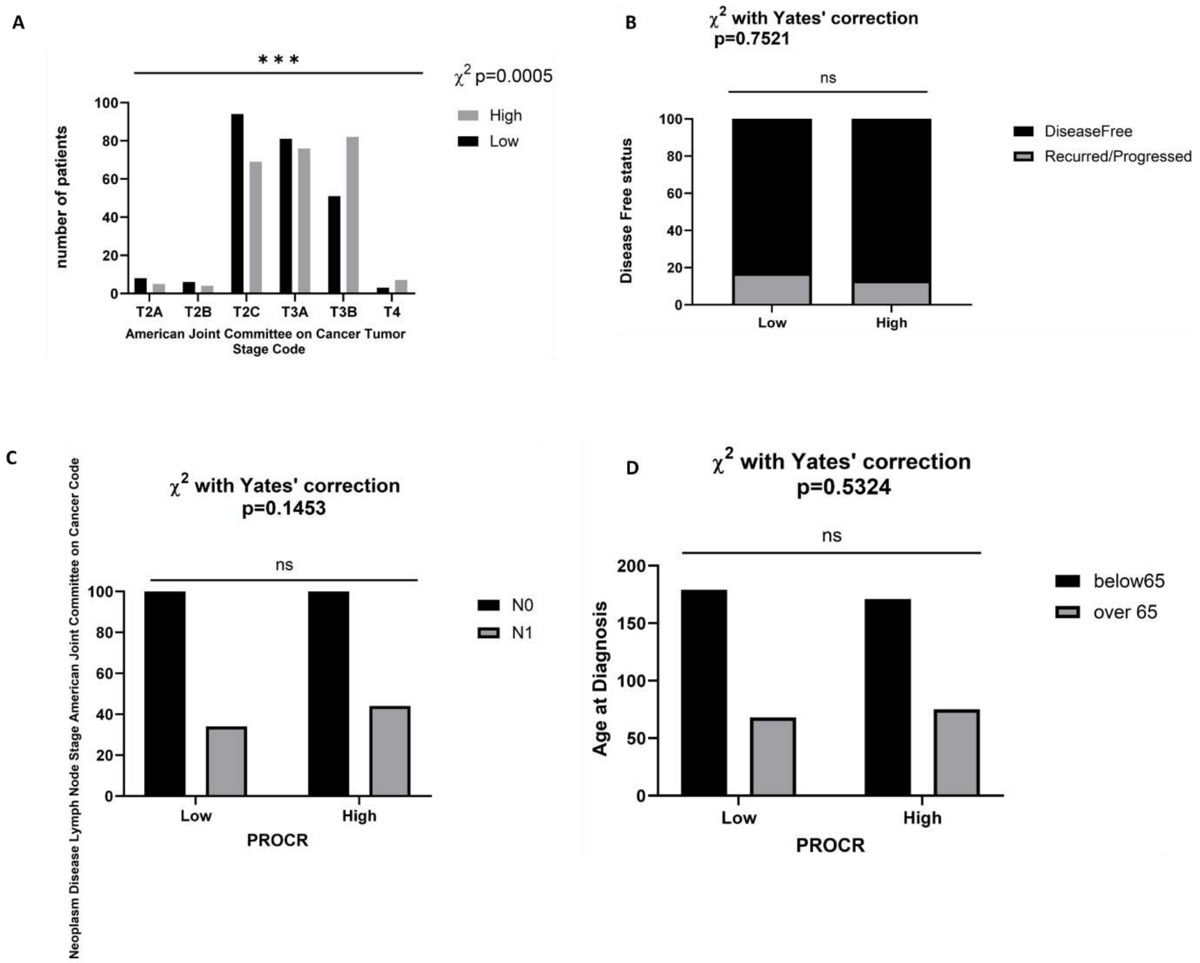


Figure 2.12 - Clinical characterises of EPCR patients expressing high and low EPCR. Analysis was undertaken after log_2. transformation. Data were used to analyse age of disease and tumour stages. Contingency analysis, and Chi-square analysed Yate's correction was used (***p=0.0005, ns=not significant).

Patients were analysed by tumour stage using the American Joint Committee stage code based on the high and low expression of EPCR (Fig. 2.12 A). Statistical difference across all the stage. Disease-free status (Fig. 2.12 B), the Neoplasm Disease Lymph Node stages and age (Fig. 2.12 C) of diagnosis were also analysed (Fig. 2.12 D). No statistical differences in the clinical characteristics of patients exhibiting low and high EPCR expression were observed (Figure 2.12).

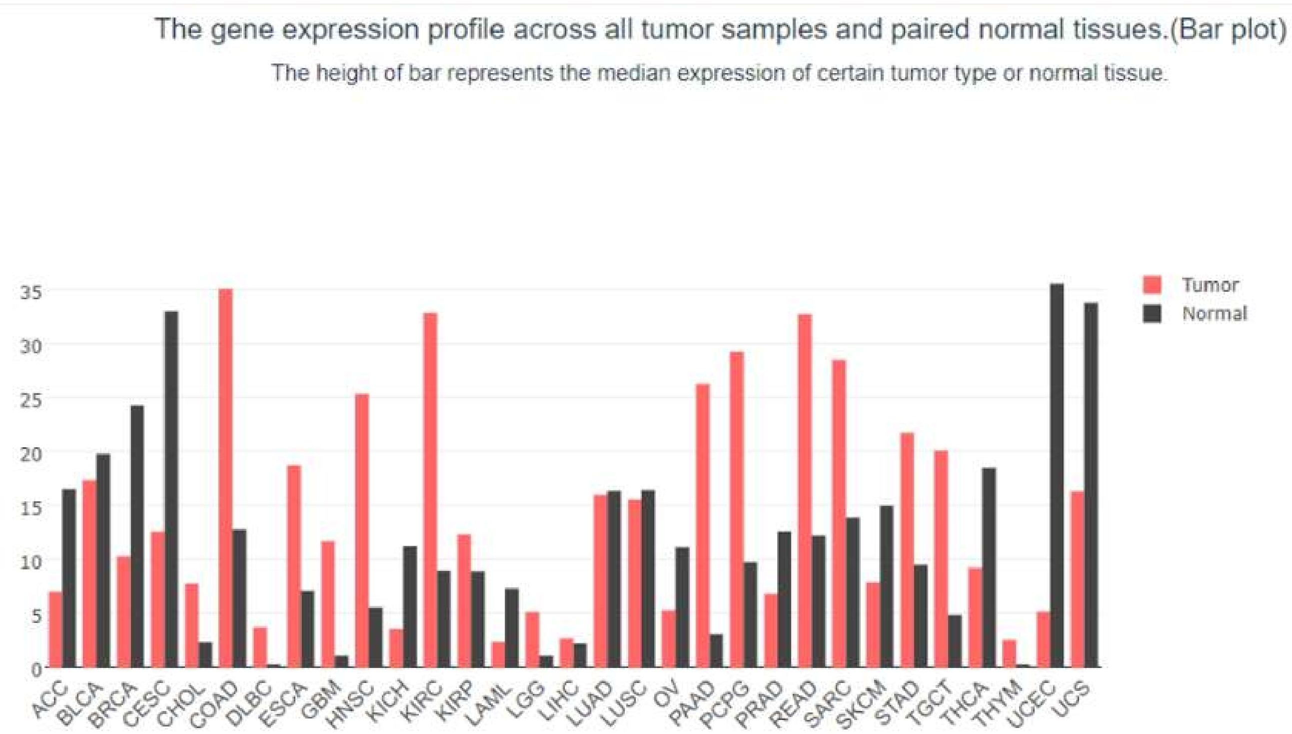


Figure 2.13 - EPCR (*PROCR*) gene expression profile in tumour and normal tissues. The panel shows the gene expression profile of EPCR (*PROCR*) in different type of tumours and normal tissues. Gene expression level is represented with bar graph as median expression between the normal and tumour tissues. (http://gepia.cancer-pku.cn/detail.php?gene=PROCR).

According to Gene Expression Profiling Interactive Analysis (GEPIA), EPCR gene expression is detected in different tumour samples paired with normal tissue. Gene expression in prostate adenocarcinoma (PRAD) was higher and almost double the expression in normal tissue compared with normal tissue (12.62 versus 8.86) (Figure 2.13).

2.4 Discussion

The aim of this element of the project was to determine whether EPCR was a suitable marker for a potential monoclonal antibody-based drug therapy targeting prostate cancer stem cell (PCSCs). NANOG represents an important marker for the maintenance of renewing the embryonic stem cells (ESCs) as also represents a key regulator in the clonogenic growth, tumorigenesis and therapy resistance (Jeter et al., 2016). Studies on hepatocellular carcinoma (HCC) have demonstrated that using NANOG promoter as a marker it is possible to successfully isolate a small sub-population of cells NANOG-positive cells with high ability of self-renewal, clonogenicity and initiation of tumour important hallmarks in the definition of cancer stem cells (CSC) (Shan et al., 2012).

Previously work in our laboratory identified prostate cancer stem cells using a 2nd generation of lentivirus NANOG promoter which controlled the expression of the enhanced green fluorescent protein (EGFP) when transduced into DU145 and LNCaP prostate cancer cells and isolated populations of *NANOG-EGFP* positive and *NANOG-EGFP* negative cells using the MoFlow cell sorter (Buczek et al., 2018a). For the first part of the project, lentivirus containing the NANOG promoter controlling EGFP expression was transduced into DU145 cells, after which NANOG-*EGFP+* and *NANOGEGFP-* cells were sorted. Isolated cells were later stained with the EPCR monoclonal antibody that was generated in our laboratory. However, results shown no statistical difference in EPCR expression between NANOG positive and NANOG negative (figure 2.2). Therefore, the experiment suggests that EPCR is not a suitable target for prostate cancer stem cells (PCSCs).

To assess whether EPCR mAb that had been generated had the potential to trigger ADCC against prostate cancer stem cells expressing EPCR, an *in vitro* ADCC bioluminescent assay was used to measure the activation of reporter effector cells following the incubation of clone JvGCRC-H61.3 (IgG2b) and JvGCRC-599.5 (IgG2b) mAbs with NANOG-EGFP+ DU145 and PC3. The kit provides an engineered Jukart effector cell, derived from immortalised line of human T lymphocyte cells that express the mouse mFcyRIV and a luciferase reported driven by an NFAT-response element (NFAT-RE). The ligation of the mFcyRIV of the effector cells with the Fc of the antibody and the NFAT-response element drives the expression of luciferase reporter. Although this is not a functional cytotoxic assay, using this approach reduces the variability that commonly happens using other assays that rely on the isolation of primary peripheral blood mononuclear cells (PBMCs).

Although results show that both monoclonal antibodies can activate the effector cells, it is important to clarify that the data come from a single experiment, and that there were a few issues with the experimental approach (Figure 2.3). Particularly, NANOG+ PC3 cells were not isolated on the same day as

NANOG+ DU145 cells because *NANOG+* PC3 cells were in very low numbers and required culture to generate sufficient number for the assay. This could affect the fold induction of PC3 NANOG+ observed compared with DU145 *NANOG+* (table 2.10). These experiments had to be suspended due to COVID-19 lockdown in March2020 and it was not possible to repeat them afterwards.

The correlation of between EPCR expression and 'stemness' was evaluated using NANOG+ and NANOG- populations and the silencing of EPCR. These experiments used the EPCR silencing construct of previously generated in our laboratory (Di Biase, 2020). Three stemness marker were analysed NANOG, OCT4 and SOX2. Both NANOG and OCT4 are responsible of the maintenance of pluripotency and self-renewal, fundamental properties for the embryonic stem cells (ESCs) (Boiani & Schöler, 2005). In breast cancer OCT4 promotes the tumorigenesis and self-renewal by the activation of its downstream genes as NANOG and SOX2 (Ponti et al., 2005). NANOG is responsible for the cell fate determination of the pluripotency of the inner cell mass during embryonic development (Chambers et al., 2003b) . Based on their role in cancers, the influence of EPCR knockdown in DU145 cells on the expression of these markers was determined. Results interestingly suggested that there is a reduction of the expression of OCT4 in shRNA4 and NANOG in shRNA1 knockdown when compared with the control pLKO.1 empty vector, while no statistically significant difference was found between the knockdown and the empty vector pLKO1. Decrease of OCT4 and NANOG expression after EPCR knockdown could suggest a potential correlation between them.

Knockdown studies assessing whether there is any correlation between EPCR expression and cancer stem cell markers CD44, CD133/PROM1 and CD24. CD133 is not very highly expressed in prostate cancer confirming the low expression even in EPCR knockdown (Xiang et al., 2016). It is demonstrated that in PC cells intracellular CD24, which encode for a glycosylphosphatidylinositol (GPI)-anchored cell surface protein, promotes cell proliferation and inhibition of apoptosis leading tumour progression and metastasis in both xenogenic and transgenic tumour models (Zhang, W. et al., 2016). Results shown that there is an increased expression of CD24 in one of the two knockdowns (shRNA4) when compared with the control empty vector pLKO.1 suggesting that a EPCR silencing could affect positively in cancer progression. No statistical difference in the expression of CD44 was observed.

Recent review stated that EMT involves changing in the cell behaviour which determine the loss of some epithelial features and gain of more mesenchymal ones. All those changes that occur in the cells might require the cooperation of a large number of molecules and cannot rely on just the epithelial marker E-cadherin and vimentin (Yang, J. et al., 2020). In order to analyse the involvement of EPCR in the mechanism of EMT in prostate cancer, different EMT markers together with EMT-TFs were analysed. Results showed that there is no statistical difference in the expression of E-cadherin versus the control pLKO.1 in both knockdown construct shRNA1 and shRNA4. Same for the vimentin and fibronectin, mesenchymal markers, when compared to control. A high level of vimentin expression is correlated with tumour progression, metastasis and poor clinical outcome in different cancers (Satelli & Li, 2011). Results suggested that silencing EPCR expression does not reduce vimentin or fibronectin expression. Therefore, targeting EPCR would not negatively impact EMT. We then analysed the effect of EPCR knockdown on the expression of E-cadherin transcriptional repressors SNAI1, SLUG/SNAI2, twist-related protein 1 (Twist 1) and zinc finger E-box-binding homeobox 1 (ZEB1), all of which are highly expressed during EMT (Lo et al., 2017). ZEB1 induces transendothelial migration and mediates migration repressing the E-cadherin expression, while loss of ZEB1 in prostate cancer undergoing to EMT showed a moderate re-expression of E-cadherin allowing the acquisition of epithelial characteristics (Drake et al., 2009). A more mesenchymal phenotype is also confirmed by the high expression of vimentin and fibronectin. Findings showed that there is no correlation between and the reduction of Twist-1 and mesenchymal factor vimentin and fibronectin after EPCR knockdown. Statistically significant low expression of TWIST1 and SNAI1 is observed in EPCR knockdown cells, when compared with control pLKO.1 (empty vector), this could suggest a reverse from the epithelial-mesenchymal transition (EMT) towards a mesenchymal to epithelial transition (MET) (Dongre & Weinberg, 2019). MET could promote a secondary metastasis formation of cell that have already spread into bloodstream (Pitsidianaki et al., 2021).

The cBio Cancer Genomic Portal (cBioPortal) is an open-access source used for the study of multidimensional cancer genomic datasets (Cerami et al., 2012a). The *In vitro* analysis performed on DU145 were included with *in silico* data from cancer genomic data set. Datasets from three consortia (TCGA 2018, TCGA 2015, ICGC) were used for the analysis. Data of the sample type were provided for each consortium, in TCGA 2018 and ICGC the samples type is primary tumour while the data TCGA 2015 samples type is metastatic prostate adenocarcinoma. Results show that there is a different correlation between EPCR and stemness factor across the different consortia. Particularly, NANOG, while there is no correlation with EPCR in TCGA 2018 in the ICGC data there is a negative correlation so increasing levels of EPCR correspond to decreasing levels of NANOG expression. NANOG is expressed in different primary tumours including brain, breast esophagus, ovary and prostate (Zhang, S. et al., 2008) (Jeter et al., 2009) . In HCC, expression of NANOG in the primary tumour is correlated with poorer clinical outcome (Yin, X. et al., 2012) . The fact that these data are derived from primary tumour could suggest an involvement of EPCR and cancer progression. However, in the TCGA 2015 where the sample types are metastatic there is a positive correlation that determine an increasing of NANOG when EPCR increasing (Figure2.8). These findings could suggest a different mechanism in cancer development and progression. The octamer-binding transcription factor (OCT4), *POU5F* gene name, is an important factor involved in the control of self-renewal and pluripotency (Wang, Ying-Jie & Herlyn, 2015b). Analysing the correlation between EPCR and OCT4 in primary tumour data set TCGA 2018 and in the metastatic tumour types shows a positive correlation between the two markers (figure 2.8). A recent genomic study suggested that OCT-4 is expressed in about 11% of primary tumours and 25% of metastatic prostate cancers (Taylor, B. S. et al., 2010). Results could suggest a potential involvement of EPCR in cancer progression and its effect on pluripotency and self-renewal control factor OCT4 (Figure 2.8). In order to study a correlation with cancer stem cell markers commonly used the data were sorted for EPCR high and low level. Results shown that all three genes analysed are positively correlated with EPCR. The increasing level of EPCR corresponds to an increased of CD133(*PROM1)*, CD44 and CD24 across the three datasets in primary tumour and metastatic samples type (Figure 2.8).

Finally, EMT factors were analysed in correlation with EPCR in the three data set. Results shown that the mesenchymal factor involved in the process of EMT were positively correlated with EPCR in all the three tumour types. Vimentin is a filamental protein in mesenchymal cells and is often correlated to cell invasion via EMT (Wang, Ying-Jie & Herlyn, 2015b). This finding could suggest a potential involvement of EPCR in the progression of cell invasion via Vimentin (Figure 2.10). EMT-TFs expression also correlated with EPCR. Results reported that there was a positive correlation in all the three data set for the genes (*ZEB1*, *SNAI1, SNAI2*) (Figure 2.11). These genes are involved in the EMT by altering the expression of specific surface-proteins. Even the E-cadherin action is negatively affected by EMT-TFs that is responsible for cancer progression (Acs et al., 2001). This finding could propose a potential role of EPCR in cancer progression.

Lastly using the clinical data from the cBioPortal, the age of diagnosis in patients expressing high and low level of the gene was examined. Results show statistical differences for TCGA 2018. No statistical difference in the disease-free status and in the neoplasm lymph nodes stage was observed in patients expressing high and low levels of EPCR gene. A statistical difference was observed in the American Joint Committee Tumour stage (Figure 2.12).

In conclusion, the use of EPCR as a potential biomarker in cancer has already been observed in colorectal cancer (Lal et al., 2017). However, our aim was to demonstrate that EPCR could be a potential marker for prostate cancer stem cells (PCSCs). Our data shown that EPCR expression in prostate cancer stem cells and normal cancer cells is the same. We also demonstrated that EPCR is not a suitable marker for studying EMT as knockdown of EPCR induces an increase in the expression of mesenchymal markers. This suggests a role in cancer progression but could also suggest a potential prognostic biomarker. We finally confirmed our findings using *in silico* data from three genomic datasets. Ultimately, we analysed the gene profile of EPCR in different types of tumours tissue compared with normal tissue using GEPIA BioSource. EPCR is expressed in different types of cancer in normal tissue and tumour tissue, which suggest significant off-target effect for an EPCR-target therapy. It is therefore concluded that EPCR is not a suitable marker for a monoclonal-antibody drug therapy in prostate cancer. The study of EPCR was therefore concluded and no other experiments were designed.

3. Gene expression studies of MDA-MB-231 and DU145 SPAG5 deficient

3.1 Introduction

Sperm-Associate Antigen 5 (SPAG5) is involved in the mitotic spindle formation essential for cell to enter in the anaphase (Mack & Compton, 2001b). In the metaphase SPAG5 is localised in the centromere (Chu et al., 2016) and interacts with other proteins to form a molecular switch to the centromere to regulate the centromere-microtubule dynamics and allowing the mitotic process (Thein et al., 2007). Several studies also demonstrated that the SPAG5 is involved in the progression in different types of cancer including cervical cancer, breast cancer and prostate cancer (Yuan, L. et al., 2014) (Zhou, X. et al., 2019). High expression of SPAG5 was associated with a poor prognosis in lung adenocarcinoma (LUAD) (Huang, R. & Li, 2020). The role of SPAG5 in cancer progression and proliferation was conducted in different types of cancers, in hepatocellular carcinoma (HCC) SPAG5 highly expressed in HCC tissue when compared to normal renal tissue. Interestingly the downregulation of SPAG5 (negatively or positively) affected the proliferation of cancer cells by the upregulation of Scavenger receptor class A member 5 (SCAR-5), a member of scavenger family involved in many human cancers (Liu, Hongliang et al., 2018). Study conducted in malignant melanoma (MM) demonstrated that a downregulation of SPAG5 inhibits the progression through the inhibition of the forkhead box protein M1 (FOXM1) (Dang et al., 2022). In a study conducted in 2016, it has been shown that interaction of the binding site of microRNA(mir539) at 3'-UTR of SPAG5, resulted in proliferation, migration, and invasion *in vitro* (Zhang, Hongtuan, et al. 2016). In triple-negative breast cancer high SPAG5 expression is correlated with a poor prognosis however, a downregulation of its expression strongly affects the cancer cell cycling, progression, proliferation, and migration (Canu et al., 2021).

Role of SPAG5 in the progression of breast cancer is known in the literature, however, its role in prostate cancer has not been investigated in detail. Moreover, the effect of the gene expression changes after SPAG5 silencing has not been studied well in both prostate and breast cancer. This chapter investigates the effect of SPAG5 silencing at transcriptomic level using high-throughput RNA sequencing (RNA-seq).

3.2 Materials and Methods

3.2.1 Cell line and reagents

Breast cancer cell line MDA-MB-231, MDA-MB-468 and MCF-7 (breast adenocarcinoma derived from metastatic site: pleural effusion) and MDA-MB453 (metastatic carcinoma derived from metastatic site: pleural effusion), representing five molecular breast cancer subtypes (Claudin-low, Basal-like, Luminal A) and prostate cancer cell line DU145 (Human prostate cancer) and PC3 (Human prostate adenocarcinoma) were used from Jon Vann Geest Cancer Research Centre / Nottingham Trent University where they were purchased from American Type Culture Collection (ATCC). All cell lines were cultured in their dedicated media. For MCF-7 Eagle's Minimum Essential Medium (EMEM) (Lonza) supplemented with 10% FBS and 0.01 mg/ml insulin solution. Leibovitz's (L-15) with 1% w/v L-Glutamine (Lonza) was used for MDA-MB-231, MDA-MB-468 and MDA-MB-453. For prostate cancer cell line Eagle's Minimum Essential Medium (EMEM) with 1% w/v of L-Glutamine (Lonza) was used for DU145 and Kaighn's Modification of Ham's F-12 Medium (F-12K) was used for PC3 cell line supplemented with 10% of foetal calf serum (FCS) according to ATCC culture methods. No antibiotics were used. MCF-7, DU145 and PC3 cell line were incubated at 37°C, in a humidified atmosphere with 5% v/v CO_2 and humidified air. MDA-MB-231, MDA-MB-468 and MDA-MB-453 were incubated at 37°C in humidified atmosphere without CO_2. Cells were routinely passaged at 70-80% of confluency. During passaging, cells were washed twice with Dulbecco's Phosphate Saline (DPBS) and detached by incubating with 0.25% w/v of Trypsin – 0.53 mM EDTA (Lonza) solution for 5-15 min at 37°C. Equal amount of specific medium were added immediately upon cell detachment and cell were centrifuged at 260 xg for 5 min. Cell counting was carried out, by resuspending a harvested cell pellet in 1-3 mL of cell dedicated medium and resuspending cell solution in Trypan blue 1:10. The haemocytometer was used to count the total number of living cell excluding the number of dead cells (blue stained) from the count. The cell pellet was re-suspended in fresh medium and cells re-cultured in culture flasks (Sarstedt, UK) by passaging. Stock of cell line was prepared at approximately $1x10^6$ cell number in 1 mL FCS + 10% v/v Dimethyl sulfoxide (DMSO) (Insight Biotechnology) (freezing medium) and stored at -80°C (Revco/Sanyo). When required, cells were thawed, gently re-suspended in 10 mL of relevant culture media, and centrifuged at 150 xg for 5 min. Cell pellets were then gently re-suspended in a fresh batch of cell dedicated medium and plated in suitable flasks. Medium changes ensured removal of DMSO frozen cells samples and increased the viability of thawed samples.

3.2.2. Gene expression profile of SPAG5 in breast and prostate cancer: qPCR

3.2.2.1 RNA extraction and cDNA synthesis

Total RNA was isolated from MDA-MB-231, MDA-MB-468, MDA-MB-453, MCF-7 and prostate cells line DU145 and PC3 and processed for cDNA as described in the paragraph 2.2.9 in chapter 2.

3.2.2.2 Primer reconstitution

SPAG5 forward and reverse primers were purchased from Merck (Table 3.1) and resuspended according to the manufacturer's recommendations 100 pmol/ µL and vortexed. The vials were kept for 30 min to dissolve completely. Aliquots with 10 pmol working solutions were prepared by adding 10 µL from the primer stock to 90 µL of nuclease free water (NFW) (1:10) dilution, and stored at -20°C.

Table 3.1 - SPAG5 primers specifications. Table provides the technical characteristics relative to Forward Human 1_SPAG5 (FH1_SPAG5) and Reverse Human_SPAG5 (RH1_SPAG5)

Oligo Name	Oligo#	Tm°	GC%	Sequence (5'-3')
FH_*SPAG5*	8815837094 70/0	58.4	45	GGATTATACAACATGGACAGC
RH_*SPAG5*	8815837094 70/1	59.2	40	TTTTCCAACTCCAGTTGTTC

3.2.2.3 Antibiotic titration assay

DU145 and MDA-231 were grown to 50 % of confluence in 1 mL of cell-dedicated medium in 6-well culture plates and incubated overnight at 37°C. Once ready cells were replaced with a fresh media containing varying concentrations of puromycin (0.5 µg/mL, 1 µg/mL ,1.5 µg/mL, 2 µg/mL, 2.5 µg/mL, 3 µg/mL) in each well and one well without antibiotic as control. Titrations was conducted in triplicate in the 6-well and the medium was replaced every two days. Cells were cultured for seven days, cells were checked the viability using a cell counter (Thermofisher; Countess 3 Automated Cell Counter). The concentration of the antibiotic used for both cell lines were determined as 1 µg/mL.

3.2.2.4 Plasmid purification

Lentivirus vectors Mission® #SHC001 pLKO.1-puro empty vector, Mission® #SHCLNG-NM_006461 SPAG5 – shRNA 1, shRNA2, -shRNA3, - shRNA4 in bacterial glycerol stock were purchased from Sigma (Table 3.2). For viral particles formation lentivirus vector psPAX2 (Addegene; #PL-12260) and packaging pMD2.G (Addegene; #PL-12259) and transduction positive control PL-SIN-EF1α-EGFP (Addegene; #21320) were obtained from Addgene in bacteria stab format. Each bacterial stock tubes were gently spin down before opening. Using a sterile loop ice splinters were removed and placed in 0.5 mL Bijou tubes containing 1 mL of S.O.C Medium (Invitrogen #15544-034) without antibiotics and incubated at 37°C with shaking for 30 min approximately at 300 rpm. Nutritional media Lauria Broth agar was prepared (LB)(Sigma-Aldrich) was prepared according to the manufacturer's instruction. LB powder was dissolved in DD H_2O and autoclaved for sterilisation. Using a sterile loop 50 µL from each incubated culture were streaked onto plates containing a pre-warmed LB-agar selective plates containing 100 µg/mL of ampicillin, and incubated overnight at 37°C. After incubation, single colonies for each plasmid were then inoculated in 5 mL of LB agar containing appropriate selective antibiotic (ampicillin 100 µg/mL) and incubated for 8h at 37°C with vigorous shaking (approximately 300 rpm). Using a conical flask at least 4 times the volume of the culture, 100 µL of each lentivirus starter culture were inoculated in 100 mL of medium containing 100 µg / mL of ampicillin antibiotic and incubated for 16h at 37 °C with shaking approximately 300 rpm. After overnight culturing, cultures with visible cloudy bacterial growth were selected for plasmid purification using QIAGEN QIAfilter Plasmid Midi according to manufacturer's protocol. Isolated plasmids were quantified using Nanodrop 8000 Spectrophotometer and stored in TE buffer at -20°C.

3.2.2.5 Cell transfection using HEK-293T cell line

HEK-293 (human embryonic kidney cells containing SV40 T-antigen) cells were grown to 90% confluency in 5 mL of Dulbecco's Modified Eagle medium (DMEM) supplemented with 10% of FCS and 1% w/v of L-Glutamine in T25 flasks. On the first day, 20 µL of Lipofectamine™ 3000 Transfection Reagent (Invitrogen # L3000008) and 500 µL of Gibco OPTIMEM medium were incubated for 30 min at RT (L+O mixture). For viral particles formation 6 µg of the packaging plasmid psPAX2 and 2 µg of the envelope expressing plasmid pMD2.G were mixed with 500 µL OPTIMEM medium (psPAX2 + pMD2.G + O). For each plasmid of interest 8 µg (pLKO.1-puro empty vector, SPAG5-shRNA2 and SPAG5-shRNA4 and PL-SIN-EF1α-EGFP) were added to each mix containing psPAX2 + pMD2.G + O, mixed and pre-incubated with L+O mixture for 30 min and finally added to each HEK-293T T25 flask with 5mL of cell dedicated medium, and incubated for 16h at 37°C. On the second day, medium was changed to all the T25 flask and replaced with 5 mL of fresh HEK-293T dedicated medium. On the third day, fractions of medium (F1) from transfected cells were collected and filtered through 0.45 nm of nylon strainers (Sartorius) to remove any residual particles. Filtered fractions were aliquoted and stored at -20°C for short term storage. Flasks were replaced with 5 mL of fresh HEK-293T cell medium. On the fourth days a second fraction (F2) was collected at the same way and stored at -20°C.

3.2.2.6 Infection of breast and prostate cancer cell lines

Triple-negative MDA-MB-231 breast cancer cell line and DU145 prostate cancer cell line were grown to 60% confluence in 6-well culture plates. Infection of target cells was obtained combining 1 mL of F1 containing pLKO.1-puro empty vector, SPAG5-shRNA2 and SPAG5-shRNA4 (Table 3.2) and PL-SIN-EF1α-EGFP of collected lentiviral supernatant with 1 mL of cell-line dedicated medium supplemented with 8 µg/mL of Hexadimethrine bromide solution (10 mg/mL in 0.9% NaCl filtered #H9268-10G), added to the target cells. Cells were incubated at 37°C and for MDA-MB-231 without CO_2 overnight. The efficiency of infections between fractions (F1-F2) in the positive control PL-SIN-EF1α-EGFP was assessed through the observation of the expression of the Enhanced Green Fluorescent Protein (EGFP) at 488 nm using the fully integrated digital inverted benchtop microscope (EVOS™ M5000 Imaging System, Invitrogen # AMF5000) and by flow cytometry for quantification difference. After 72h after the infections MDA-MB-231 and DU145 were treated using cell-dedicated media supplemented with 1 µg/mL of puromycin for both cell lines to select the cells that efficiently integrated the plasmid. A control with antibiotic was prepared by adding 1 µg/mL of puromycin for wild-type MDA-MB-231 and DU145.

Table 3.2 - Specification of Mission® shRNA clone SPAG5 silencing. The table shows the region on SPAG5 gene were target by each shRNA.

	Species	Clone ID	Target Seq	Match position	Match Region
shRNA 1	**Human**	**TRCN0000156202**	CCAGAATCTGCTTCACCTCTT	10615	3UTR
shRNA 2	**Human**	**TRCN0000153765**	CAGAATCTGCTTCACCTCTTT	10615	3UTR
shRNA 3	**Human**	**TRCN0000322618**	CAGAATCTGCTTCACCTCTTT	10615	3UTR
shRNA 4	**Human**	**TRCN0000154768**	GCAGCAGATTTCCGTGTCAAT	10615	CDS
shRNA 5	**Human**	**TRCN0000322546**	CCAAATTAGCTCTACTCCTAA	10615	CDS

3.2.2.7 Testing knockdown efficiency of SPAG5 in MDA-MB-231 and DU145 by q-PCR.

Total RNA was isolated from cell expressing pLKO.1, SPAG5 shRNA2 and shRNA4 in both MDA-MB-231 and prostate cells line DU145 and processed for cDNA as described in the paragraph 2.2.9 in chapter 2.

3.2.2.8 Western Blot

3.2.2.8.1 Reagents used for Western Blot

10% Resolving Gel	20mL
ddH_2O	7.9 mL
30% Acrylamide Geneflow #A2-0084	6.7 mL
1.5 M Tris pH 8.8	5.0 mL
10% SDS Sigma #151-21-3	0.2 mL
10% APS Geneflow #A2-0200	0.2 mL
TEMED Geneflow #A2-0104	0.008 mL

5% Stacking Gel	8 mL
ddH_2O	5.5 mL
30% Acrylamide GeneFlow #A2-0084	1.3 mL
1 M Tris pH 6.8	1mL
10 % SDS Sigma #151-21-3	0.08 mL
10% APS GeneFlow #A2-0200	0.08 mL
TEMED GeneFlow #A2-0104	0.008 mL

10X Running Buffer	For 1L
SDS	10 g
Trizma Base Sigma #77-86-1	30.3 g
Glycine Sigma #56-40-6	144 g
ddH_2O	Up to 1 L
10X Transfer Buffer	**For 2 L**
Glycine Sigma #56-40-6	5.8 g
Trizma Base Sigma #77-86-1	11.6 g
10% SDS Sigma #151-21-3	0.75 g
Methanol Fisher scientific #10499560	400 mL
ddH_2O	Up to 2 L
10X Tris-Buffered saline (10X TBS)	**For 1 L**
Sodium Chloride (NaCl) Calbiochem	80 g
Trizma Base Sigma #77-86-1	24.2 g
ddH_2O	Up to 1 L
pH adjusted to 7.6 with concentrated HCl	
Tris-Buffered saline with Tween (TBST)	**For 1 L**
10X TBS	100 mL
ddH_2O	900 mL
Tween20 Sigma #9005-64-5	1 mL

3.2.2.8.2 Total protein extract preparation

For this experiment, total protein lysate from cells expressing pLKO.1, SPAG5-shRNA2, -shRNA4 in both cell lines MDA-MB-231 and DU145 were prepared. Cells were grown and harvested as described in paragraph 3.2.1 of this chapter and snap frozed on ice for 5 min. Cells were lysed by adding 100 µL of IP Lysis buffer (pH 7.4, 25mM Tris, 150mM NaCl, 1mM EDTA, 1% NP40, 5% glycerol; Pierce™ ThermoFisher Scientific; # 88805) supplemented with 1% v/v of protease inhibitor (Halt™ Protease and Phosphatase Inhibitor Single-Use Cocktail, EDTA-Free (100X); #78442) to prevent protein degradation. Tubes were transferred into ice bath for sonication at the max power for 10 min then stored on ice. Cell lysate was passed 10 times through 29G 12.7mm needle 0.5 mL syringe (BD Micro-Fine™+ insulin syringe; # 324824) for complete lysis. Those steps were repeated three times, and then samples were centrifuged at 14,000 xg for 15 min at 4°C. Finally, each supernatant, corresponding to the protein lysate, was carefully removed, and stored in new individual Eppendorf tubes at -80°C. The amount of protein was assessed using protein assay kit (Pierce™ BCA Protein Assay Kit; #23227) according to the manufacture's protocol. Bicinchoninic acid (BCA) protein assay working reagent (WR) was prepared by mixing 50 parts of BCA Reagent A with 1 part of Reagent B (50:1, Reagent A: B). For each sample is required 200 µL of working reagent (WR) in flat bottom 96 well (Sarstedt®; #83.3924). Protein standards were prepared using different bovine serum albumin (BSA) protein concentrations (20–2000 µg/mL) in distilled water. Samples were diluted 1:10 in distilled water and incubated for 30 min at 37°C covered with foil to protect from the light.

Table 3.3 - BSA standard curve. Each standard was prepared in duplicates for a final volume of 20 µL**.**

Final concentration (mg/mL)	2 mg/mL BSA (µL)	Buffer (µl)
0	0	20
0.2	2	18
0.4	4	16
0.6	6	14
0.8	8	12
1.2	12	8
1.6	16	4
2.0	20	0

Standards and samples absorbance were read at 562 nm using a plate reader (iMark™ Microplate Absorbance Reader; Biorad#1681130). The average of each absorbance value was taken from the control and subtracted from the averages of samples to remove the background signal. Protein concentration from each sample was calculated using a standard curve.

3.2.2.8.3 Gel preparation and electrophoresis

SDS-PAGE was performed for protein separation and prepared as shown in paragraph 3.2.2.8.1 of this chapter. Gels were mounted on a Mini-PROTEAN® Electrophoresis System containing 1x SDS running buffer. Each gel-lane were loaded with 30 µg of specific protein lysate sample containing 4x of Protein loading buffer containing 10 µL / mL of β-mercaptoethanol (Sigma) at ratio 3:1 (4x Laemmli Sample Buffer, *Bio-Rad* #1610747), vortex and incubated at 95°C for 10 min to allow complete protein denaturation. After denaturation each sample were carefully loaded on well-set gel alongside with 5 µL of protein standard (Precision Plus Protein™ WesternC™ Blotting Standards, *Bio-Rad* #1610376). Samples were run at 60V for 10 min to let the proteins to pass through the stacking gel then increased at 100V until the tracking dye ran off the bottom of the gel.

3.2.2.8.4 Protein transfer

Once protein run through the gel (approximately 2h), gel was removed from the tank and carefully transferred onto 0.2 µm pore size nitrocellulose blotting membrane (Amersham™ Protran® Western blotting membranes, # GE10600001) for a ''wet transfer''. The nitrocellulose membranes were cut into 8.5 cm x 6 cm pieces to fit the size of the gel. The ''sandwich'' was assembled using the assembly cassette in the order: sponge-filter paper- gel-membrane-filter- paper-sponge. The sandwich was prepared taking care to remove the air bubble during the assembly because could affect the transfer. Once the sandwiches are prepared and immediately inserted into the transfer tank filled up with ice-cold transfer buffer. Proteins were transferred for 90 min at 100V at 4°C.

3.2.2.8.5 Membrane probing

After the transfer, the membranes were carefully removed from the sandwich and blocked in a 5% w/v of Marvel Skimmed milk in TBST for 1h at RT on a rocking shaker. Membranes were incubated with the specific primary antibody at 4°C on a rocking platform overnight. In the following day membranes were washed six times for 5 min each with TBST solution and then incubated with the corresponding secondary antibody and antiladder Precision Strep Tactin 5x (Biorad # 161-0380) for detection of the of Precision Plus Protein unstained standards for 1h on rocking platform at RT. All the antibodies specification used are shown in table 3.4. After the incubation, membranes were washed six times for 5 min on a rocking platform and incubated using Clarity Western ECL Substrate (Bio-Rad, # 1705060) for imaging using a CCD camera (Syngen, Western Blot imager).

Table 3.4 - Antibodies and dilution used

Primary Antibody	Dilution	catalogue Number	Class	Supplier
Rabbit anti-SPAG5	1:2000	#14726-1-AP	polyclonal	proteintech®
Mouse anti-β-actin	1:5000	#SAB1305567	monoclonal	Sigma

Secondary Antibody	Dilution	catalogue Number	Class	Supplier
Anti-Rabbit IgG HRP-linked	1:1000	#7074	monoclonal	Cell Signalling
Anti- MouseIgG HRP-linked	1:1000	#7076	monoclonal	Cell Signalling

3.2.2.9 RNA sequencing in MDA-MB-231 and DU145 SPAG5 knockdown

RNA sequencing was performed to investigate differentially expressed genes in MDA-MB-231 and DU145 SPAG5 deficient cell lines versus control empty vector. Library generation and next sequencing of pLKO.1-puro (Empty vector) and SPAG5-shRNA4 in both MDA-MB-231 and DU145 cell lines generation was performed at the Novogene (UK) in Cambridge. Total RNA for this experiment was isolated from pLKO.1-puro (Empty vector) and SPAG5-shRNA4 MDA-MB-231 and DU145 cells using Qiagen RNAeasy columns. Cells were cultured in three independent T75 flasks for each cell line for a total of 12 samples. Cells were cultured and harvested as described in the paragraph 3.2.1. Samples were processed and analysed according to paragraph 2.2.9 in chapter 2. RNA purity and concentration was measured using Nanodrop 8000 and the best quality of RNAs was selected based on the 280/260 ratio (~ 2.0) and 260/230 ratio (~ 2.0-2.2) ratio. For this study, 500ng total RNA (100 ng/µL) from each sample were sent for sequencing. Sequencing was performed via illumina NovaSeq 6000 PE150 platform, a technology based on the mechanism of SBS (sequencing by synthesis). The raw data originated from high-throughput sequencing platform is transformed to sequenced reads by CASAVA base recognition. Raw data are stored in FASTQ (fq) format files. Gene expression quantitation was conducted for all the samples using the Fragments Per Kilobase of Transcript sequence per Millions base pairs sequenced (FPKM) method. Finally, a differential gene expression was assessed after the gene expression quantitation using DESeq2 software (Love et al., 2014)

3.2.2.10 Statistical analysis

All statistical analysis was performed using GraphPad Prism 8.4.2 software (GraphPad, Inc, USA). For experiments where two groups were compared, and for comparison of three or more groups, statistical analysis was performed as described in chapter 2 paragraph 2.2.15.

3.2.2.11 Bioinformatics analysis on DEGs in SPAG5 deficient cell lines MDA-MB-231 and DU145

For matrix visualization of the DEGs a heatmap was generated using the opensource software Morpheus (https://software.broadinstitute.org/morpheus/). To provide a comprehensive gene list annotation and analysis, METASCAPE open source was used (https://metascape.org/gp/index.html#/main/step1). Gene functions were divided into three parts including cellular components (CCs), biological processes (BPs), and molecular functions (MFs). The Kyoto Encyclopaedia of Genes and Genome (KEGG) was used for enrichment analysis. Functional protein-protein interaction (PPI) showed physical contact between two or more proteins and was performed using Cytoscape (https://cytoscape.org/) for PPI network.

3.3 Results

3.3.1 Gene expression profile of SPAG5 in breast and prostate cancer cell lines.

To assess the expression of SPAG5, a gene expression profile was generated using breast and prostate cancer cell lines. Four breast cancer cell lines MDA-MB-453, MDA-MB-468, MCF-7 and MDA-MB-231 and two prostate cancer cell lines DU145 and PC3 were used in this study. qPCR was used to determine the relative endogenous expression levels of *SPAG5* mRNA in the total RNA samples extracted from those cell lines. Total RNA extraction and retrotranscription in cDNA from 1.5 µg of RNA was obtained as described in the chapter 2 paragraph 2.2.9. Values were normalised using housekeeping *GUSB* mRNA expression level. Data were analysed using the $2^{\wedge -\Delta Ct,}$ a variation of Livak method (Livak & Schmittgen, 2001) where ΔCt = Ct (reference gene)- Ct (target gene).

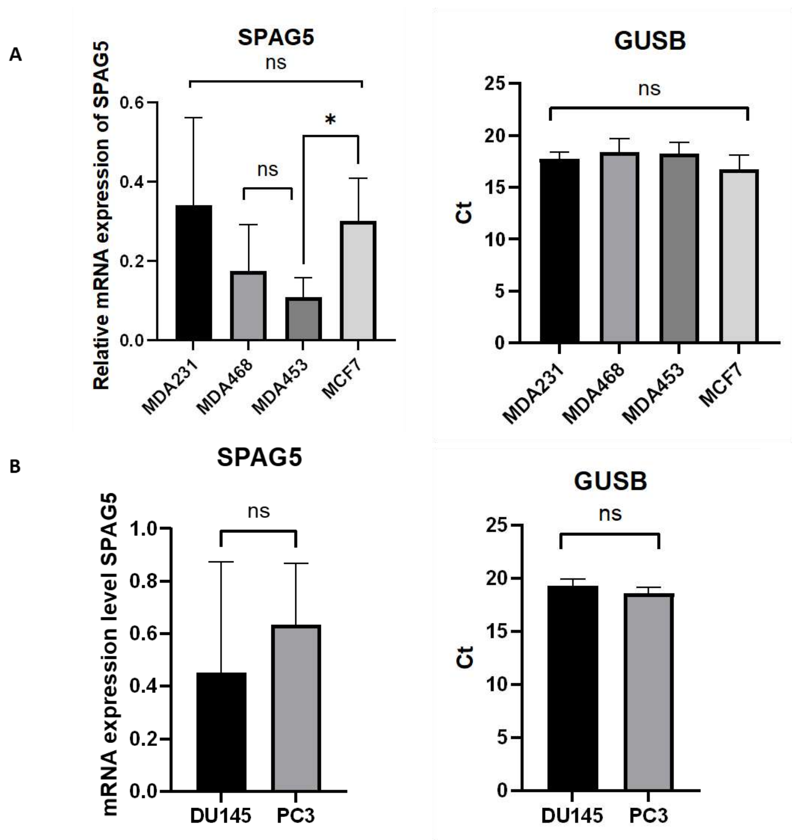


Figure 3.1 - Quantitative PCR (qPCR) gene expression studies of SPAG5 showing varying levels in breast cancer and prostate cancer cell lines. **A** Quantitative PCR on mRNA SPAG5 expression level in breast cancer cell lines MDA-MB-231, MDA-MB-468, MDA-MB-453 and MCF-7. **B** Quantitative PCR on mRNA SPAG5 expression level in prostate cancer

cell lines DU145 and PC3.Data represented as mean ±SD and are representative of three independent experiments. Values are normalised for GUSB mRNA expression in all the cell lines. Statistical significance was calculated using one way ANOVA multiple comparison and unpaired t-test indicated when significant with asterisk (*p-value =0.0480, ns, n=3).

Results showed that all breast cancer cell lines express SPAG5 however, the expression levels do not significantly (one way ANOVA) vary between the cell lines but between MDA-MB453 and MCF-7 cell line (unpaired t test p-value=0.0480). Same for prostate cancer cell lines DU145 and PC3 cell line. Results suggested that in both cancer cell lines SPAG5 is expressed with no statistical difference between the cells. Therefore, we decided to consider highly aggressive, invasive, and poorly differentiated triple-negative breast cancer (TNBC) MDA-MB-231 breast cancer cell line (Parekh et al., 2018) and the hormone-refractory and aggressive DU145 prostate cancer cell line (Belochitski et al., 2007) for further studies.

3.3.2 SPAG5 silencing in DU145 and MDA-MB-231 cancer cells

To investigate the role of SPAG5 in breast cancer and prostate cancer, both MDA-MB-231 and DU145 cell lines were transduced with lentivirus particles containing SPAG5-targeting shRNAs as described in the material and methods paragraph of this chapter 3.2. As a negative control an empty vector, without the shRNA insert, consisting of pLKO.1-puro backbone was used. All the lentivirus used for this study contain a gene with resistance to puromycin therefore a titration to select the minimum concentration of antibiotic able to kill all the cell in specific period was assessed. Cells were treated at different puromycin concentrations. Old media was replaced with fresh media containing the antibiotic concentration and checked under the light microscope every two days to assess the cell death. After seven days cells viability was checked using the cell counter and a concentration of 1 µg/mL of puromycin antibiotic that killed all the cells was selected for both cell lines (Figure 3.2).

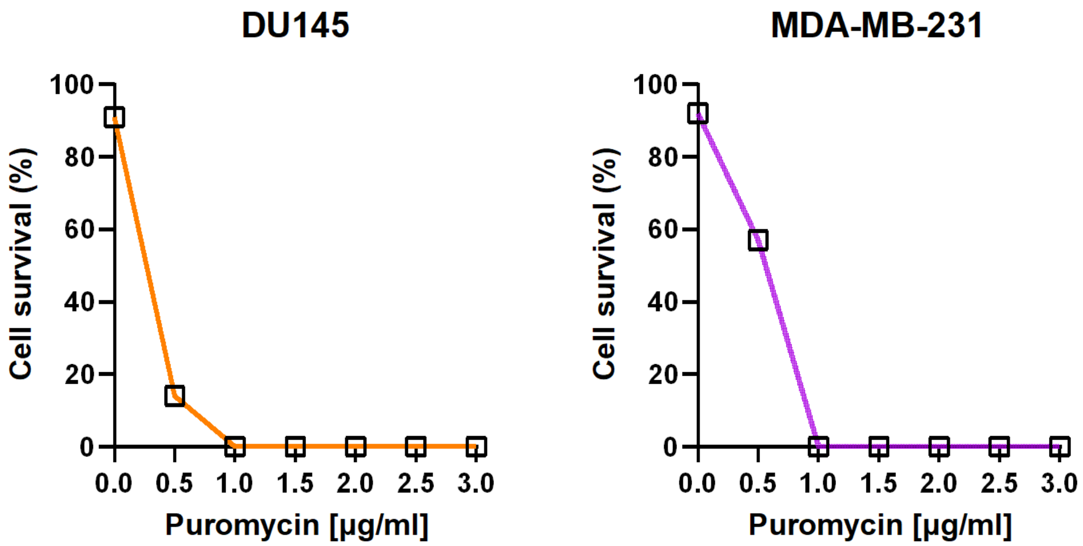


Figure 3.2 - Antibiotic killing curve. MDA-MB-231 and DU145 cell lines were incubated at different concentrations of puromycin in triplicate wells. After seven days cells viability was assessed using the cell counter and plotted. The Y axis shows the cell viability and puromycin concentration represented on X axis in both cell lines.

The efficiency of silencing was assessed initially at transcriptome level. Therefore, after three passages with antibiotic selection, total RNA was extracted from pLKO.1-puro and SPAG5-shRNA2 and -shRNA4 transduced cells and a RT- qPCR was performed on the cDNA obtained from 1.5 µg of RNA. In both cell line MDA-MB-231 and DU145, SPAG5-shRNA4 is statistically significant reduced when compared with the empty vector pLKO.1 (around 85% for MDA-MB-231 and 78 % for DU145).

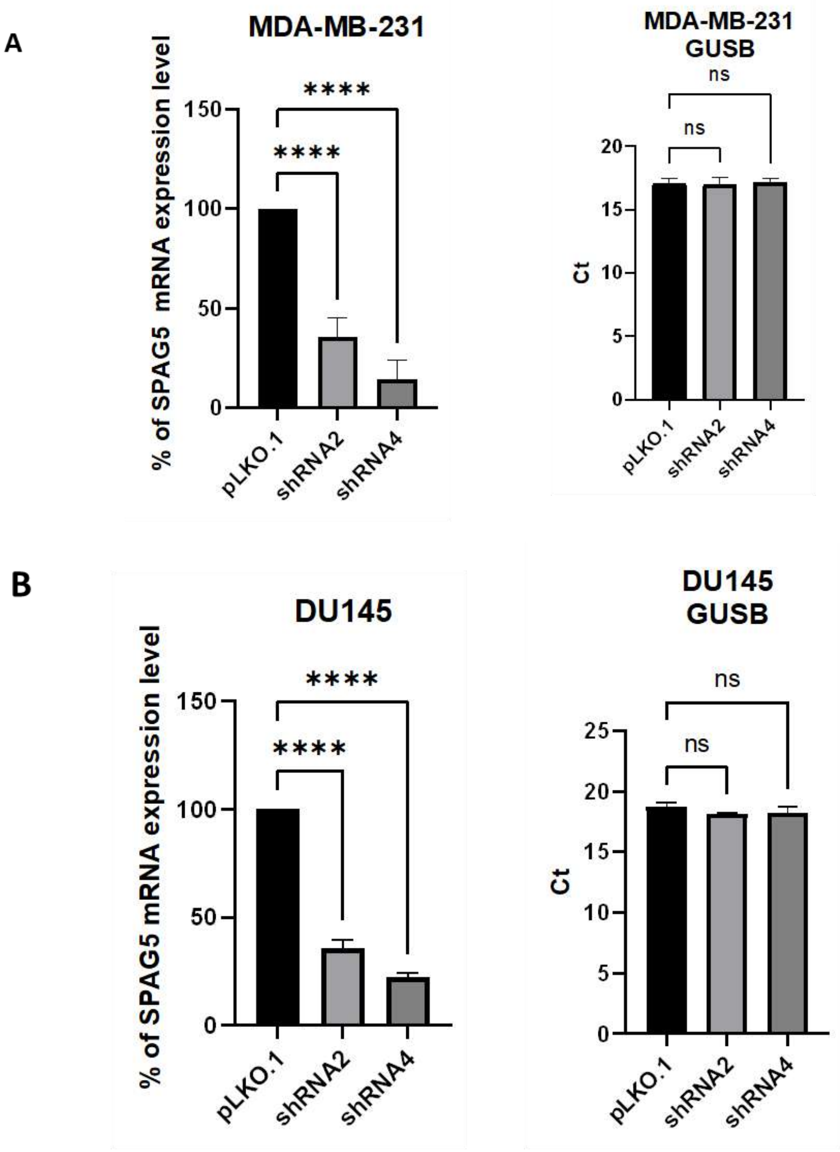


Figure 3.3 - mRNA expression levels of SPAG5 deficient in MDA-MB-231 and DU145. **A** Quantitative PCR on mRNA sample from MDA-MB-231 and DU145 in control (pLKO.1) and SPAG5 shRNAs cells Statistical analysis was carried out using one-way ANOVA with post-hoc Sidack's multiple comparison test, P value was shown. All the data shown are the mean ± SD of three independent experiment (n=3) error bars represent S.D. **B** Values are normalised on GUSB mRNA expression. Statistical significance was calculated between the control, and each using unpaired t-test and indicated when significant with asterisk. (****p=value < 0.0001, ns= non-significant; n=3).

Because the efficiency of SPAG5 silencing in shRNA4 cells was statistically higher than in shRNA2 cells, in both cells MDA-MB-231 and DU145, SPAG5 construct 4 (shRNA4) was selected for further investigation and pLKO.1 (empty vector) as a negative control.

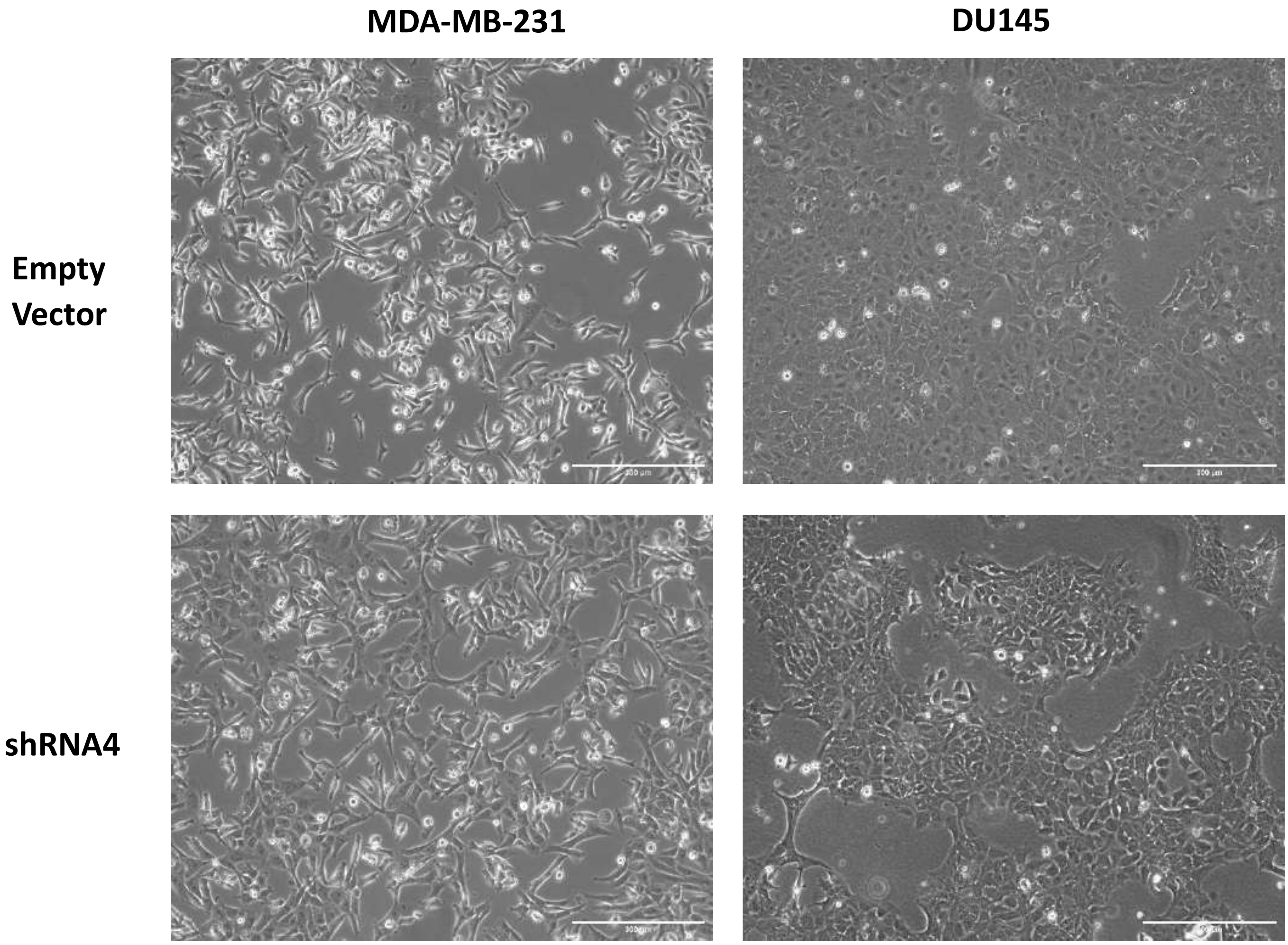


Figure 3.4 - Micrographs showing empty vector and SPAG5 deficient population in DU145 and MDA-MB-231. Representative images showing different morphologies in cell populations SPAG5 knockdown. Pictures were taken at 10x magnification. Scale bar 300µm.

To validate the efficiency of knockdown at protein level of SPAG5 silencing, in both MDA-MB-231 and DU145 were assessed by western blot analysis.

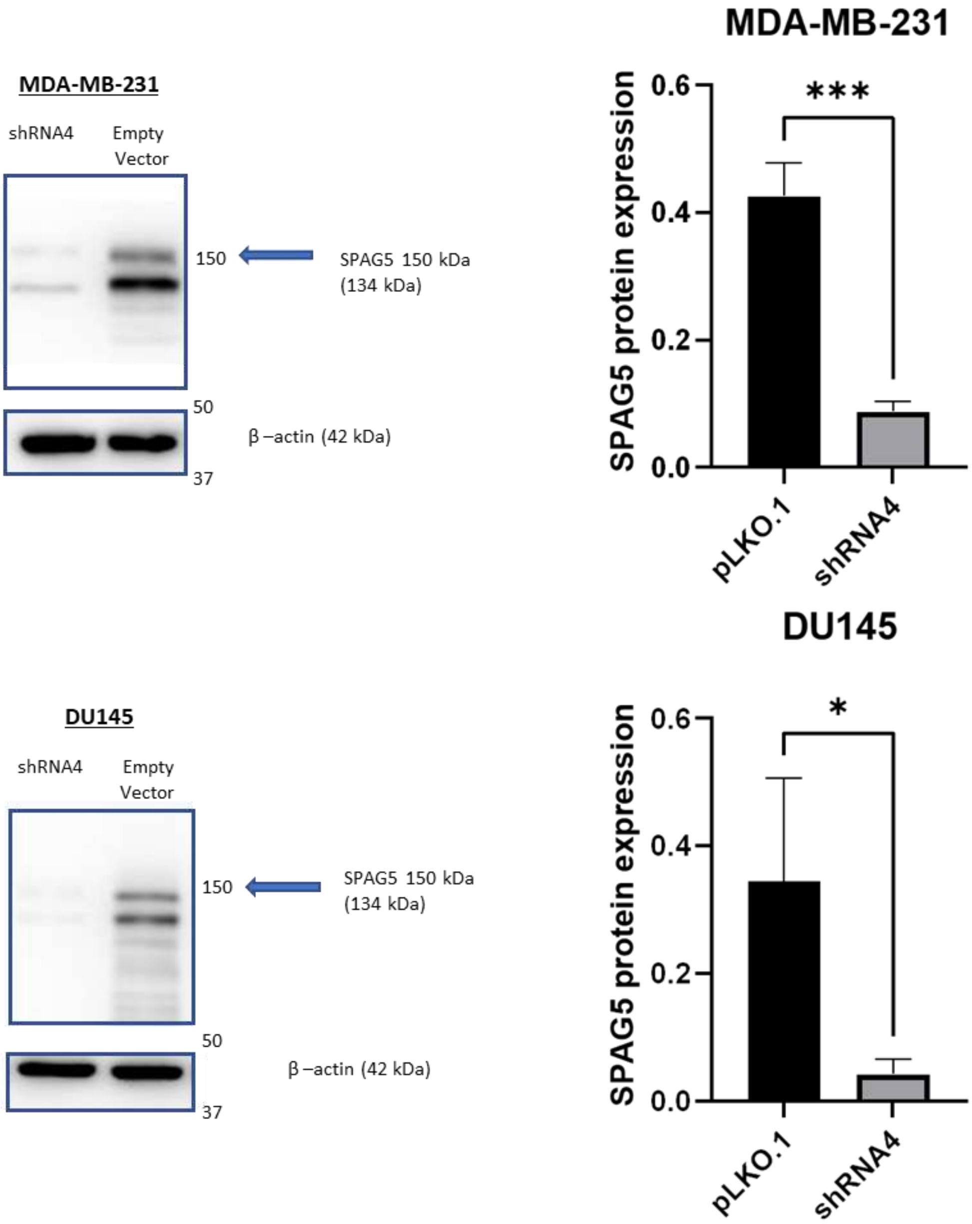


Figure 3.5 - SPAG5 protein expression in SPAG5 deficient MDA-MB-231 and DU145. Protein expression change in SPAG5 knock down using anti-SPAG5 rabbit monoclonal antibody (1:2000) in 5%Milk prepared in 1x TBS-T incubated at 4°C, overnight. ECL expose for 2 second for MDA-MB-231 (top) and 6 second for DU145 (bottom). Densitometry of the bands calculated using image-J software and the values were normalised for β-actin, the results are shown on bar graphs Calculated molecular weight 134 kDa. Observed molecular weight 134 kDa - 150 kDa Statistical significance was calculated between the control (empty vector) and each test group using the asterisk when significant (*=p-value = 0.0329, ***=p-value = 0.004; n=3).

3.3.3 Assessment of gene expression changes in MDA-MB-231 and DU145 SPAG5 after knockdown using RNA sequencing (RNA-seq) technology.

3.3.3.1 Relative gene expression level in SPAG5 deficient MDA-MB-231 and DU145 cells

Total transcriptome of MDA-MB-231 and DU145 SPAG5 knockdown was analysed through RNA-sequencing. Aim of this study was to examine differentially expressed genes and when SPAG5 is downregulated in both MDA-MB-231 (Breast Cancer) and DU145 (Prostate Cancer) cell line. For this experiment, RNAs was extracted as shown in paragraph 2.2.9 chapter 2. RNAs was quantified using Nanodrop 8000 Spectrophotometer and the RNAs purity choose based on the 280/260 ratio (~1.8-2.2) and 260/230 ratio (~ 1.8-2.2). Gene expression quantitation was conducted in all 4 samples (MDA-MB-231 pLKO.1, MDA-MB-231 shRNA4, DU145 pLKO.1 and DU145 shRNA4 each in triplicate) comparing the knockdown cell population SPAG5-shRNA4 with the control pLKO.1 (empty vector) in MDA-MB-231 and same approach was applied for DU145 cell line.

For MDA231, 51 million clean reads were generated in the two samples pPLKO.1 and SPAG5-shRNA4 (each with three biological replicates) and 49 million in the two samples pLKO.1 and SPAG-shRNA4 (each with three biological replicates) for DU145. After passed the quality control (QC) samples were processed for the alignment to the reference the complete genome annotation (GRCh38/hg38). Alignment of RNA-seq data to the reference genome was achieved using HISAT2, a faster and more sensitive graph-based alignment program for mapping next-generation sequencing reads (oligonucleotides sequenced).

HISAT2 the first run discovers splice sites supported by the reads with long anchor. The second run align reads with short anchor (1-7bp) by using the list of the splicing collected from the first run. Making more accurate and increasing the number of the alignment even if these proceed might take the double of the time to run. In contrast with the previous alignment HISAT2 includes, besides the global Ferragina-Manzini (GFM) representing all the genome, also a small GFM (local index) that cover the genome allowing an effective alignment of RNA-seq reads (Fig 3.6) (Kim, Daehwan et al., 2014).

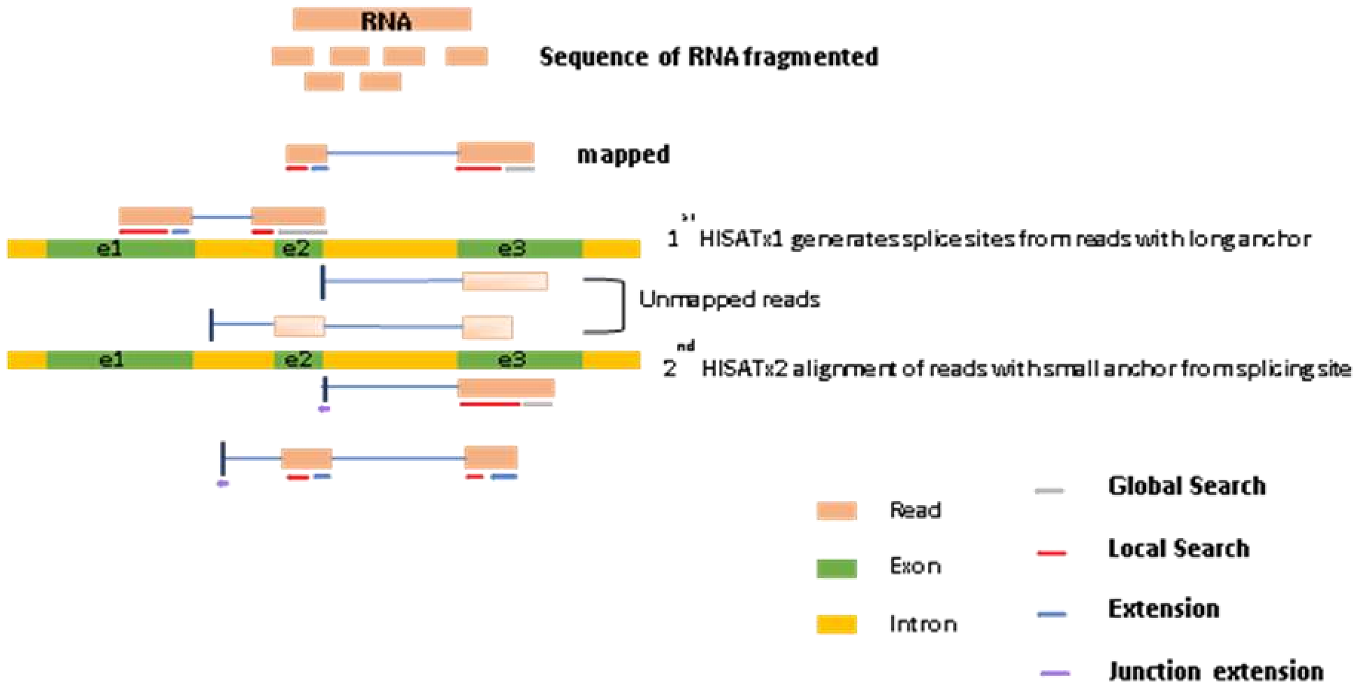


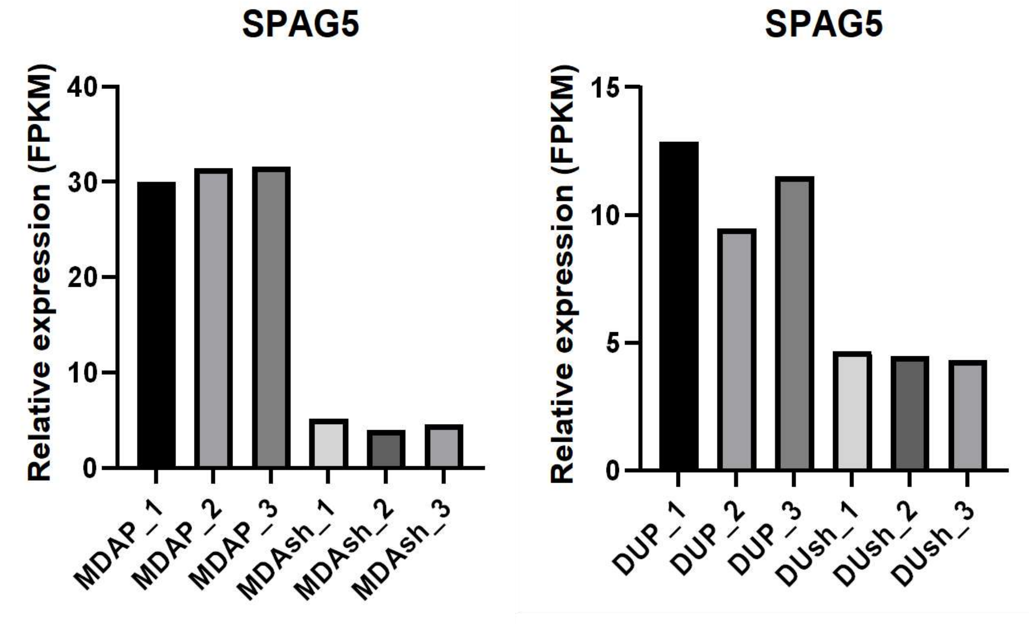
Figure 3.6 - HISAT2 algorithm for the alignment reads. Picture represents the HISAT2 alignment program for mapping the NGS reads that present short anchors. (Adapted from Kim, Langmead and Salzberg 2014).

The human annotation (GRCh38/hg38) were used for the sample comparisons and all samples shown a unique map ratio (number of total reads aligned to the unique position of the reference genome) greater than 93% which guarantees a good library generation (>90%) (Appendix chapter 3 Table 3.1). A low mapping ratio <50% would have suggested problems during the samples preparation or during data processing.

In total 56,709 genes were detected from the RNA-seq. The effective knockdown of SPAG5 in MDA-MB-231 and DU145 was confirmed by the RNA seq (Fig.3.7 A) which is consistent with result obtained from the qPCR (Fig.3.3) and by western blot (Fig.3.5).

A correlation analysis was performed using the normalised data as the first step of analysis to assess the degree of variation between the replicates and to understand the extend of biological variability. For the correlation analysis, the Fragments Per Kilobase of transcript sequence per Millions base pairs sequenced (FPKM) normalised reads were used for each sample to calculate Pearson correlation. Pearson correlation coefficient was calculated using R language package and plotted as a heatmap with a pairwise correlation coefficient indicated in each square (Fig.3.7). Samples for gene expression quantification were assessed using Feature Counts software. The heatmap generated shows that the three replicates for each sample in both cell lines MDA-MB-231 and DU145 are highly correlated (Fig.3.7 B).

After quantification of the genes, differential expression genes (DEGs) were calculated using DESeq2 software and the adjustment method used Benjamini-Hochberg. From the analysis a total of 1,121 genes were upregulated and 1080 downregulated in MDA-MB-231 (Fig.3.8 C) while for DU145, 472 genes were identified as upregulated and 435 were downregulated (log2(Fold Change) >= 0.58 and padj <= 0.05). The expression heatmap of DEGs for MDA-MB-231 SPAG5 shRNA4 are shown (total of 2201 genes are shown in the appendix section chapter 3 Table 3.1) (Fig.3.8 A). The thirty most upregulated and downregulated DEGs associated with SPAG5 shRNA4 in MDA-MB-231 are presented in volcano plot using the ggVolcanoR (https://ggvolcanor.erc.monash.edu/) (Fig.3.8 B). A list of gene identities sorted for the most significant upregulated and downregulated was obtained from the HGNC (HUGO gene nomenclature) and NCBI (National Center of Biotechnology Information) for MDA-MB-231 SPAG5 shRNA4 and is represented in the Table 3.5 (Full list of the genes identified is showed in Appendix of this chapter 3 Table 3.5)



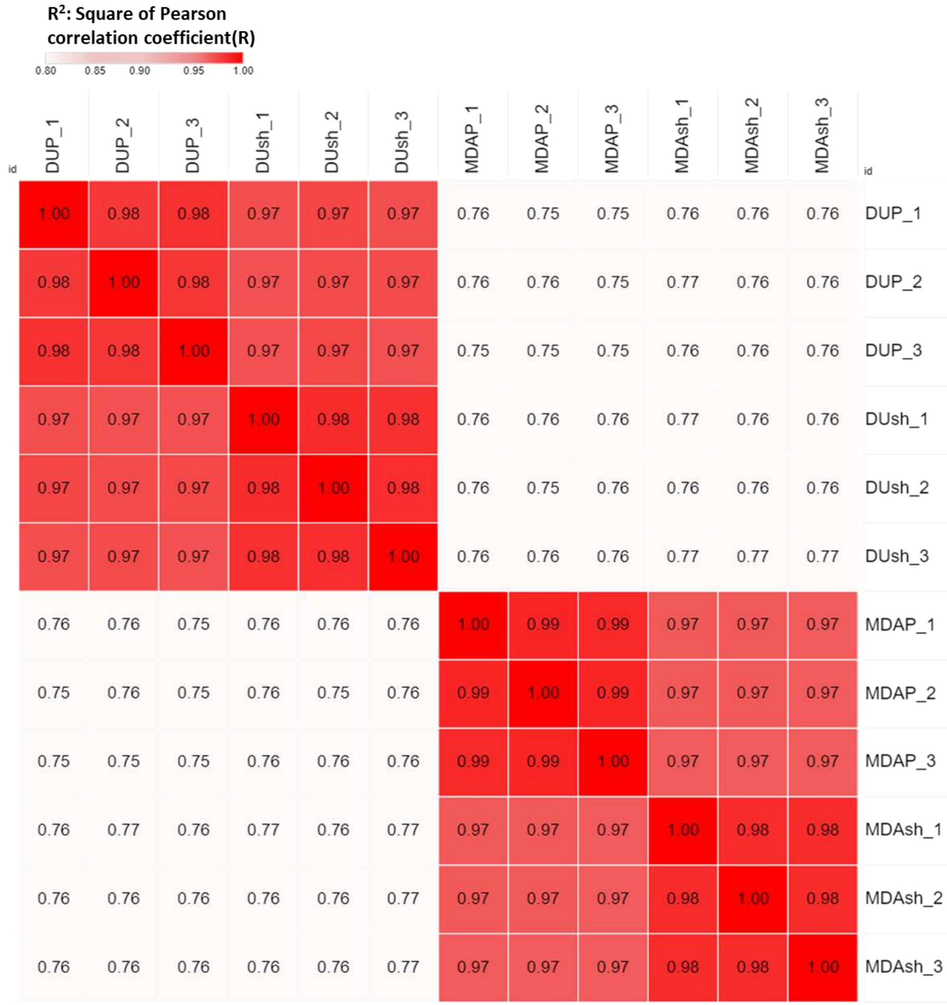


Figure 3.7 - Gene expression quantitation and correlation in MDA-MB-231 and DU145 defient versus Control. (A) Gene expression quantitation on MDA-MB-231 SPAG5-shRNA4 (MDAsh) versus MDA-MB-231 control pLKO.1 (MDAP) and DU145 SPAG5- shRNA4 (DUsh) versus control pLKO.1 (DUP). Analysis was conducted considering Fragment Per Kilobase of transcript per Million base pairs sequenced (FPKM) using FeatureCounts software (n=3). A higher FPKM corresponds to a higher expression of the gene. (B) Pearson's correlation analysis of gene expression in MDA-MB-231 and DU145 SPAG5-shRNA4 versus the control pLKO.1 (empty vector). Correlation coefficients were indicated in each square, the closer the correlation is to 1 the higher is the similarity between the samples.

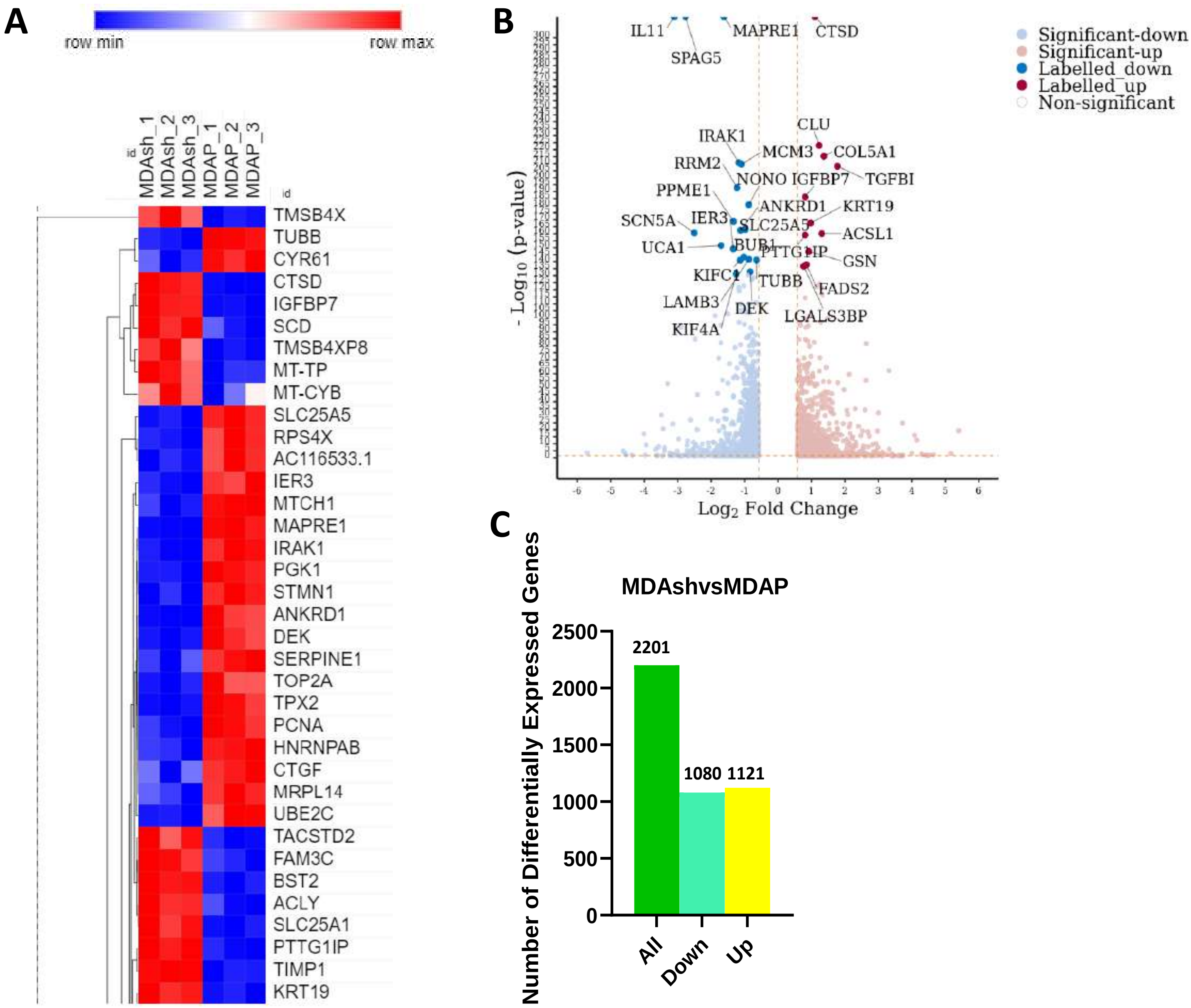


Figure 3.8 - RNA-seq analysis and differentially expressed genes (DEGs) in MDA-MB-231 SPAG5 deficient vs control (empty vector). (A) Expression heatmap showing 36 out of 2201 most significant DEGs expressed in SPAG5-shRNA4 compared with control pLKO.1 in MDA-MB-231. (B) Volcano plot showing the fold change (Log2FC) and padj value <=0.05 of significantly upregulated (red dots) and downregulated (blue dots) differentially gene expression in triple negative MDA-MB-231 shRNA4-SPAG5 using ggVolcanoR. Transcriptome data were obtained following RNA sequencing via Illumina platform from Novogene genomic service (padj <0.05 and log2FC >= 0.58; n=3). (C) Bar graph showing the DEGs expression analysis of RNA-seq in MDA-MB-231 SPAG5 knockdown.

Table 3.5 - Most significant differentially expressed genes (DEGs) in MDA-MB-231 SPAG5 deficient cells. The tables present the list of the thirty most upregulated genes out of 2201 identified, obtained applying Log2FC and a cut-off of 0.58. Gene listed in blue (right side) are downregulated in knockdown vs control cell populations, whereas the orange table (left side) are genes upregulated in knockdown vs control cell populations. (A complete table for both population is shown in the supplementary section of this

Gene ID	Log2FC	Pvalue	Gene Description
CTSD	1.10	0.00	cathepsin D [HGNC:2529]
CLU	1.23	0.00	clusterin [HGNC:2095]
COL5A1	1.37	0.00	collagen type V alpha 1 chain [HGNC:2209]
TGFBI	1.78	0.00	transforming growth factor beta induced [HGNC:11771]
IGFBP7	0.81	0.00	insulin like growth factor binding protein 7 [HGNC:5476]
KRT19	0.98	0.00	keratin 19 [HGNC:6436]
ACSL1	1.31	0.00	acyl-CoA synthetase long chain family member 1 [HGNC:3569]
PTTG1IP	0.81	0.00	PTTG1 interacting protein [HGNC:13524]
GSN	0.92	0.00	gelsolin [HGNC:4620]
FADS2	0.85	0.00	fatty acid desaturase 2 [HGNC:3575]
LGALS3BP	0.77	0.00	galectin 3 binding protein [HGNC:6564]
TACSTD2	1.24	0.00	tumor associated calcium signal transducer 2 [HGNC:11530]
HTRA1	1.13	0.00	HtrA serine peptidase 1 [HGNC:9476]
CERCAM	1.29	0.00	cerebral endothelial cell adhesion molecule [HGNC:23723]
MXRA8	1.33	0.00	matrix remodeling associated 8 [HGNC:7542]
TMSB4X	0.82	0.00	thymosin beta 4 X-linked [HGNC:11881]
LSS	0.95	0.00	lanosterol synthase [HGNC:6708]
CLIC3	1.50	0.00	chloride intracellular channel 3 [HGNC:2064]
FAM3C	0.85	0.00	family with sequence similarity 3 member C [HGNC:18664]
LOX	1.28	0.00	lysyl oxidase [HGNC:6664]
ATP1A1	0.65	0.00	ATPase Na+/K+ transporting subunit alpha 1 [HGNC:799]
PRSS23	0.67	0.00	serine protease 23 [HGNC:14370]
DHCR7	0.91	0.00	7-dehydrocholesterol reductase [HGNC:2860]
CYBRD1	0.84	0.00	cytochrome b reductase 1 [HGNC:20797]
TIMP1	0.62	0.00	TIMP metallopeptidase inhibitor 1 [HGNC:11820]
CTSA	0.82	0.00	cathepsin A [HGNC:9251]
MELTF	0.89	0.00	melanotransferrin [HGNC:7037]
TUBA1A	1.08	0.00	tubulin alpha 1a [HGNC:20766]
CPNE7	2.64	0.00	copine 7 [HGNC:2320]
LPIN1	0.94	0.00	lipin 1 [HGNC:13345]

Gene ID	log2FC	Pvalue	Gene Description
IL11	-3.10	0.00	interleukin 11 [:HGNC:5966]
SPAG5	-2.77	0.00	sperm associated antigen 5 [:HGNC:13452]
MAPRE1	-1.62	0.00	microtubule associated protein RP/EB family member 1 [:HGNC:6890]
IRAK1	-1.17	0.00	interleukin 1 receptor associated kinase 1 [:HGNC:6112]
MCM3	-1.09	0.00	minichromosome maintenance complex component 3 [:HGNC:6945]
RRM2	-1.22	0.00	ribonucleotide reductase regulatory subunit M2 [:HGNC:10452]
NONO	-0.88	0.00	non-POU domain containing octamer binding [:HGNC:7871]
PPME1	-1.33	0.00	protein phosphatase methylesterase 1 [:HGNC:30178]
SLC25A5	-0.98	0.00	solute carrier family 25 member 5 [:HGNC:10991]
ANKRD1	-1.12	0.00	ankyrin repeat domain 1 [:HGNC:15819]
SCN5A	-2.50	0.00	sodium voltage-gated channel alpha subunit 5 [:HGNC:10593]
UCA1	-1.70	0.00	urothelial cancer associated 1 [:HGNC:37126]
IER3	-1.34	0.00	immediate early response 3 [:HGNC:5392]
BUB1	-1.02	0.00	BUB1 mitotic checkpoint serine/threonine kinase [:HGNC:1148]
LAMB3	-0.87	0.00	laminin subunit beta 3 [:HGNC:6490]
TUBB	-0.64	0.00	tubulin beta class I [:HGNC:20778]
KIFC1	-1.13	0.00	kinesin family member C1 [:HGNC:6389]
DEK	-0.83	0.00	DEK proto-oncogene [:HGNC:2768]
KIF4A	-1.25	0.00	kinesin family member 4A [:HGNC:13339]
NCAPG2	-0.89	0.00	non-SMC condensin II complex subunit G2 [:HGNC:21904]
PGK1	-0.73	0.00	phosphoglycerate kinase 1 [:HGNC:8896]
RPL7L1	-0.78	0.00	ribosomal protein L7 like 1 [:HGNC:21370]
BRPF3	-1.02	0.00	bromodomain and PHD finger containing 3 [:HGNC:14256]
IDS	-1.18	0.00	iduronate 2-sulfatase [:HGNC:5389]
RANBP9	-1.16	0.00	RAN binding protein 9 [:HGNC:13727]
PRRC2A	-0.65	0.00	proline rich coiled-coil 2A [:HGNC:13918]
TOP2A	-0.87	0.00	DNA topoisomerase II alpha [:HGNC:11989]
AL161431.1	-1.09	0.00	novel transcript
TCF19	-1.08	0.00	transcription factor 19 [:HGNC:11629]
NUP153	-0.98	0.00	nucleoporin 153 [:HGNC:8062]

In DU145 SPAG5 knockdown a total of 907 differentially expressed genes were identified (Fig.3.9 C). The analysis was conducted as same as MDA-MB-231 with log2FC cut off 0.58 and padj value <= 0.05. In DU145 SPAG5 knockdown the heatmap is showing the DEGs (Fig.3.9 A). The thirty most upregulated and downregulated DEGs associated with SPAG5 shRNA4 in DU145 are presented in volcano plot using the ggVolcanoR (https://ggvolcanor.erc.monash.edu/))(Fig.3.9 B). A list of gene identities sorted for the most significant upregulated and downregulated was obtained from the HGNC (HUGO gene nomenclature) for DU145 SPAG5 shRNA4 and is represented in the Table 3.6. (Complete table is shown in the appendix section of this chapter 3 Table 3.6).

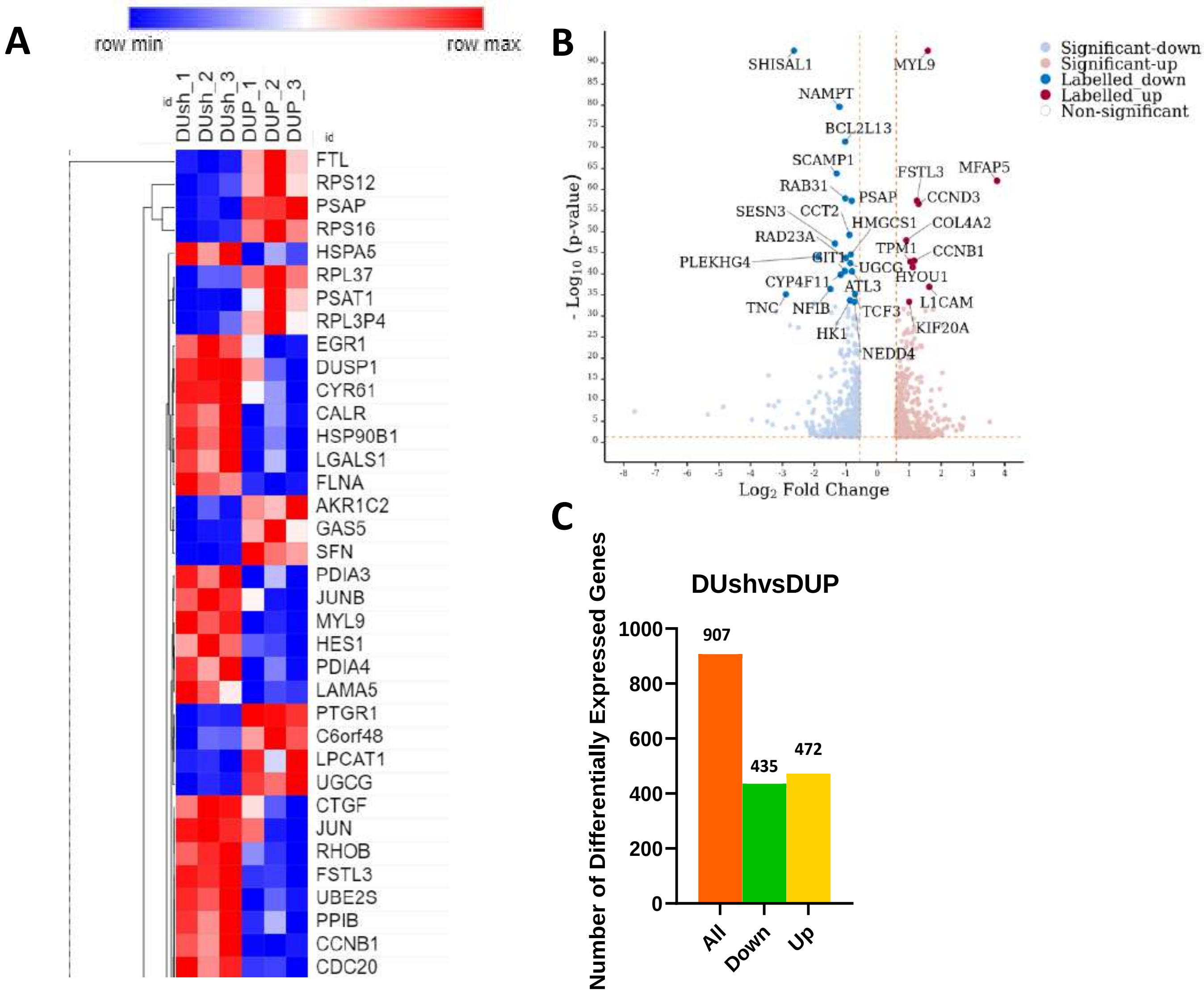


Figure 3.9 - RNA-seq analysis and differentially expressed genes (DEGs) in DU145 SPAG5 deficient vs control (pLKO.1). (A) Expression heatmap showing 36 out of 907 most significant DEGs expressed in SPAG5-shRNA4 compared with control pLKO.1 in DU145. (B) Volcano plot showing the fold change (Log2FC) and padj value <=0.05 of significantly upregulated (red dots) and downregulated (blue dots) differentially gene expression in triple negative MDA-MB-231 shRNA4-SPAG5 using ggVolcanoR. Transcriptome data were obtained following RNA sequencing via Illumina platform from Novogene genomic service (padj <=0.05 and log2FC >= 0.58; n=3). (C) Bar graph showing the DEGs expression analysis of RNA-seq in DU145 SPAG5 knockdown.

Table 3.6 - Most significant Differentially expressed genes (DEGs) in DU145 deficient cells. The tables present the list of the thirty most upregulated genes out of 907 identified, obtained applying Log2FC 0.58 and padj <= 0.05. Gene listed in blue (right side) are downregulated in knockdown vs control cell populations, whereas the orange table (left side) are genes upregulated in knockdown vs control cell populations.

Gene ID	log2FC	Pvalue	Gene Description
MYL9	1.58	0.00	myosin light chain 9 [HGNC:15754]
MFAP5	3.76	0.00	microfibril associated protein 5 [HGNC:29673]
FSTL3	1.24	0.00	follistatin like 3 [HGNC:3973]
CCND3	1.30	0.00	cyclin D3 [HGNC:1585]
COL4A2	0.90	0.00	collagen type IV alpha 2 chain [HGNC:2203]
CCNB1	1.15	0.00	cyclin B1 [HGNC:1579]
TPM1	1.03	0.00	tropomyosin 1 [HGNC:12010]
HYOU1	1.11	0.00	hypoxia up-regulated 1 [HGNC:16931]
L1CAM	1.62	0.00	L1 cell adhesion molecule [HGNC:6470]
KIF20A	0.99	0.00	kinesin family member 20A [HGNC:9787]
SDC2	0.82	0.00	syndecan 2 [HGNC:10659]
TAGLN	1.74	0.00	transgelin [HGNC:11553]
MANF	1.50	0.00	mesencephalic astrocyte derived neurotrophic factor [HGNC:15461]
TGFB2	0.97	0.00	transforming growth factor beta 2 [HGNC:11768]
KLF10	1.13	0.00	Kruppel like factor 10 [HGNC:11810]
FLNA	0.67	0.00	filamin A [HGNC:3754]
ANXA8	2.04	0.00	annexin A8 [HGNC:546]
AURKA	1.16	0.00	aurora kinase A [HGNC:11393]
CPT1A	0.95	0.00	carnitine palmitoyltransferase 1A [HGNC:2328]
ADAMTSL4	1.16	0.00	ADAMTS like 4 [HGNC:19706]
DNAJB11	0.87	0.00	DnaJ heat shock protein family (Hsp40) member B11 [HGNC:14889]
PLK1	1.09	0.00	polo like kinase 1 [HGNC:9077]
CDC20	0.97	0.00	cell division cycle 20 [HGNC:1723]
BMI1	0.83	0.00	BMI1 proto-oncogene, polycomb ring finger [HGNC:1066]
PTTG1	1.16	0.00	pituitary tumor-transforming 1 [HGNC:9690]
ANXA6	0.76	0.00	annexin A6 [HGNC:544]
IQGAP3	0.92	0.00	IQ motif containing GTPase activating protein 3 [HGNC:20669]
HES1	1.08	0.00	hes family bHLH transcription factor 1 [HGNC:5192]
SLC1A1	1.09	0.00	solute carrier family 1 member 1 [HGNC:10939]
CDHR1	1.35	0.00	cadherin related family member 1 [HGNC:14550]

Gene ID	log2FC	Pvalue	Gene Description
SHISAL1	-2.63	0.00	shisa like 1 [HGNC:29335]
NAMPT	-1.20	0.00	nicotinamide phosphoribosyltransferase [HGNC:30092]
BCL2L13	-1.03	0.00	BCL2 like 13 [HGNC:17164]
SCAMP1	-1.29	0.00	secretory carrier membrane protein 1 [HGNC:10563]
RAB31	-1.02	0.00	RAB31, member RAS oncogene family [HGNC:9771]
PSAP	-0.82	0.00	prosaposin [HGNC:9498]
CCT2	-0.89	0.00	chaperonin containing TCP1 subunit 2 [HGNC:1615]
SESN3	-1.34	0.00	sestrin 3 [HGNC:23060]
HMGCS1	-0.85	0.00	3-hydroxy-3-methylglutaryl-CoA synthase 1 [HGNC:5007]
PLEKHG4	-1.88	0.00	pleckstrin homology and RhoGEF domain containing G4 [HGNC:24501]
RAD23A	-1.01	0.00	RAD23 homolog A, nucleotide excision repair protein [HGNC:9812]
UGCG	-0.87	0.00	UDP-glucose ceramide glucosyltransferase [HGNC:12524]
GIT1	-1.04	0.00	GIT ArfGAP 1 [HGNC:4272]
ATL3	-0.81	0.00	atlastin GTPase 3 [HGNC:24526]
CYP4F11	-1.16	0.00	cytochrome P450 family 4 subfamily F member 11 [HGNC:13265]
NFIB	-1.48	0.00	nuclear factor I B [HGNC:7785]
TNC	-2.89	0.00	tenascin C [HGNC:5318]
TCF3	-0.72	0.00	transcription factor 3 [HGNC:11633]
HK1	-0.88	0.00	hexokinase 1 [HGNC:4922]
NEDD4	-0.74	0.00	E3 ubiquitin protein ligase [HGNC:7727]
BRDT	-1.99	0.00	bromodomain testis associated [HGNC:1105]
SPAG5	-1.33	0.00	sperm associated antigen 5 [HGNC:13452]
MLXIP	-0.93	0.00	MLX interacting protein [HGNC:17055]
SLC4A4	-3.21	0.00	solute carrier family 4 member 4 [HGNC:11030]
NT5E	-1.46	0.00	5'-nucleotidase ecto [HGNC:8021]
LYPLA2	-0.92	0.00	lysophospholipase II [HGNC:6738]
RPS16	-0.75	0.00	ribosomal protein S16 [HGNC:10396]
FTL	-0.94	0.00	ferritin light chain [HGNC:3999]
FLNC	-0.77	0.00	filamin C [HGNC:3756]
PTGR1	-0.61	0.00	prostaglandin reductase 1 [HGNC:18429]

In order to investigate a potential biological role of SPAG5 deficient in MDA-MB-231 and DU145 cell line, a GO and KEGG enrichment analysis was performed. (GO Biological process, GO Cellular components, GO Molecular functions are presented in the appendix Table 3.10).

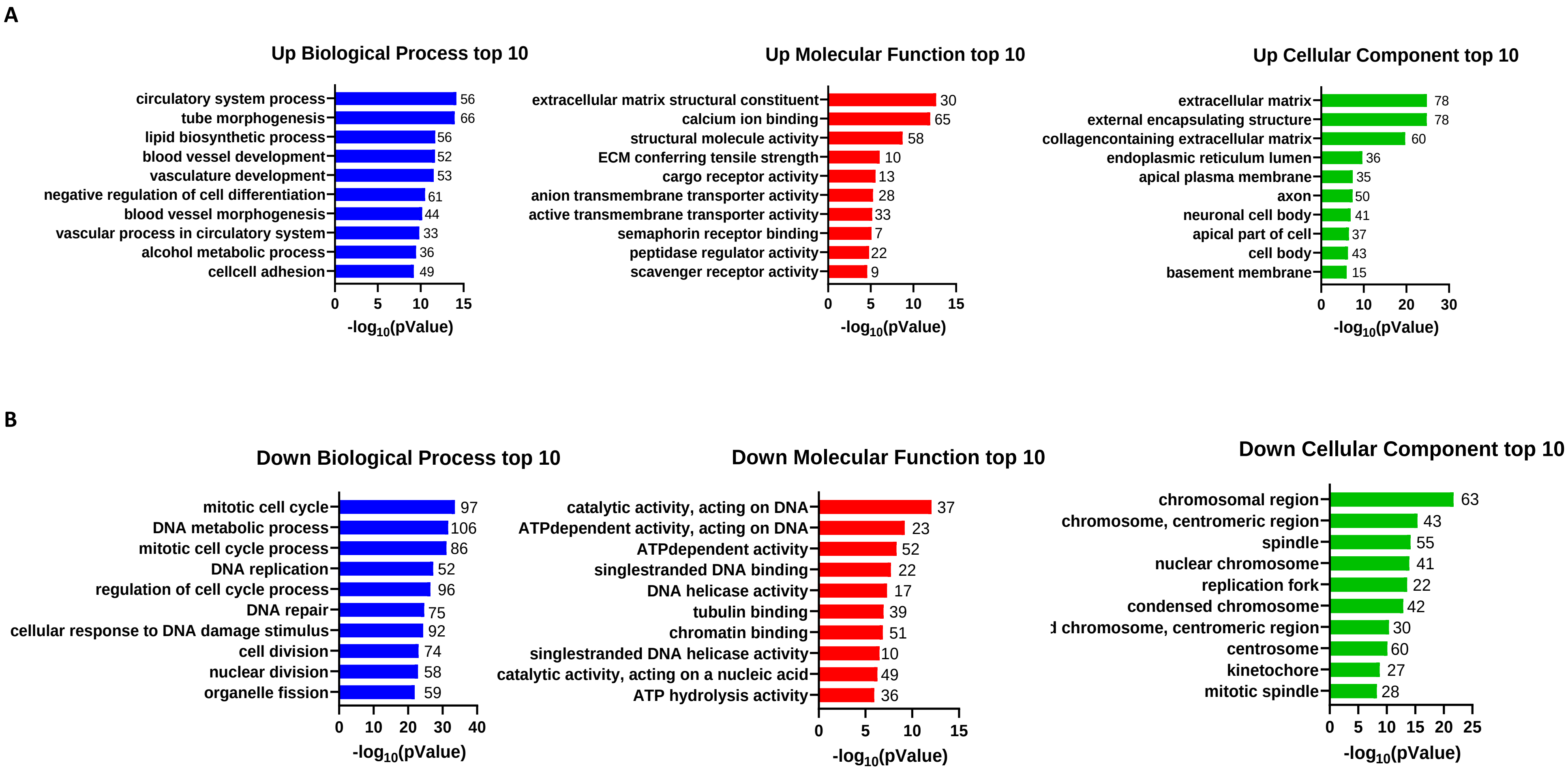


Figure 3.10 - GO analysis of DEGs in MDA-MB-231 shRNA4 SPAG5 deficient. Bar graphs show the most 10 most enriched Gene Ontology (GO) terms (Biological process, molecular function, cellular component) upregulated (A) and downregulated in MDA-MB-231 SPAG5 silencing (B). Data were obtained using Metascape software and sorted for p-value <= 0.05.

From the GO analysis is demonstrated that the shRNA4- SPAG5 upregulated genes (Fig.3.9 A) in the biological process, were significantly enriched in the 'circulatory system process', 'tube morphogenesis' and 'lipid biosynthesis process' while for the Molecular functions the most significant genes were enriched in the 'Extracellular matrix structural consistent', 'calcium ion binding' and 'structural molecule activity'. Finally, SPAG5 upregulated genes in the cellular components were found significantly enriched in 'extracellular matrix', 'external encapsulated structure' and 'collagencontaining extracellular matrix'. The downregulated genes (Fig.3.9 B) were enriched in biological process in 'mitotic cell cycle', 'DNA metabolic process' and 'mitotic cell process', while for molecular function the most significant enriched genes were in 'catalytic activity, acting on DNA', 'ATP dependent activity, acting on DNA' and 'ATP dependent activity'. Finally, in the cellular component the most significant genes were enriched in 'chromosomal activity,' 'chromosome centromeric region' and 'spindle.' Kyoto Encyclopaedia of Genes and Genomes (KEGG) pathways enriched was performed on MDA-MB-231 SPAG5 deficient. Results shown that shRNA4 upregulated genes were significantly enriched in 'ECM receptor interaction', 'protein digestion and absorption' and' PI3K Akt signalling pathway'. 'DNA replication,' 'cell cycle' and 'Fanconi anaemia pathway' were significantly enriched in downregulated genes set.'

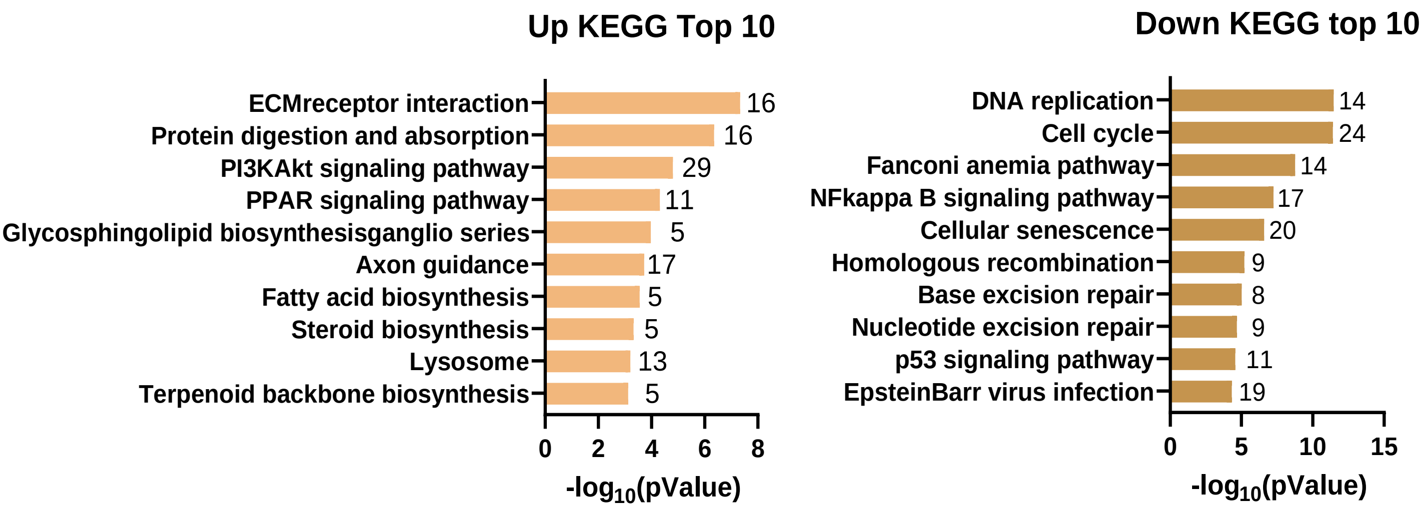


Figure 3.11 - KEGG analysis of differentially gene expression in MDA-MB-231 SPAG5 deficient. Bar graphs show the most 10 most enriched KEGG pathways upregulated and downregulated in MDA-MB-231 SPAG5 silencing. Data were obtained using Metascape software and sorted for p-value <= 0.05. X-Axis represented $-\log_{10}$ p-Value and Y axis represent KEGG pathway classes. Number of genes in each pathway category enriched in the present data set given at the end of each bar. (Complete list KEGG enrichment in the appendix chapter 3 Table 3.11).

Same analysis was applied for DU145 shRNA4 SPAG5 silencing. Differential expressed gene (DEGs) upregulated and downregulated were processed through Metascape free gene annotation sources for GO (Fig.3.11) and KEGG pathways enrichment. (Full list with GO and KEGG in the appendix chapter 3 Table 3.12-Table 3.13).

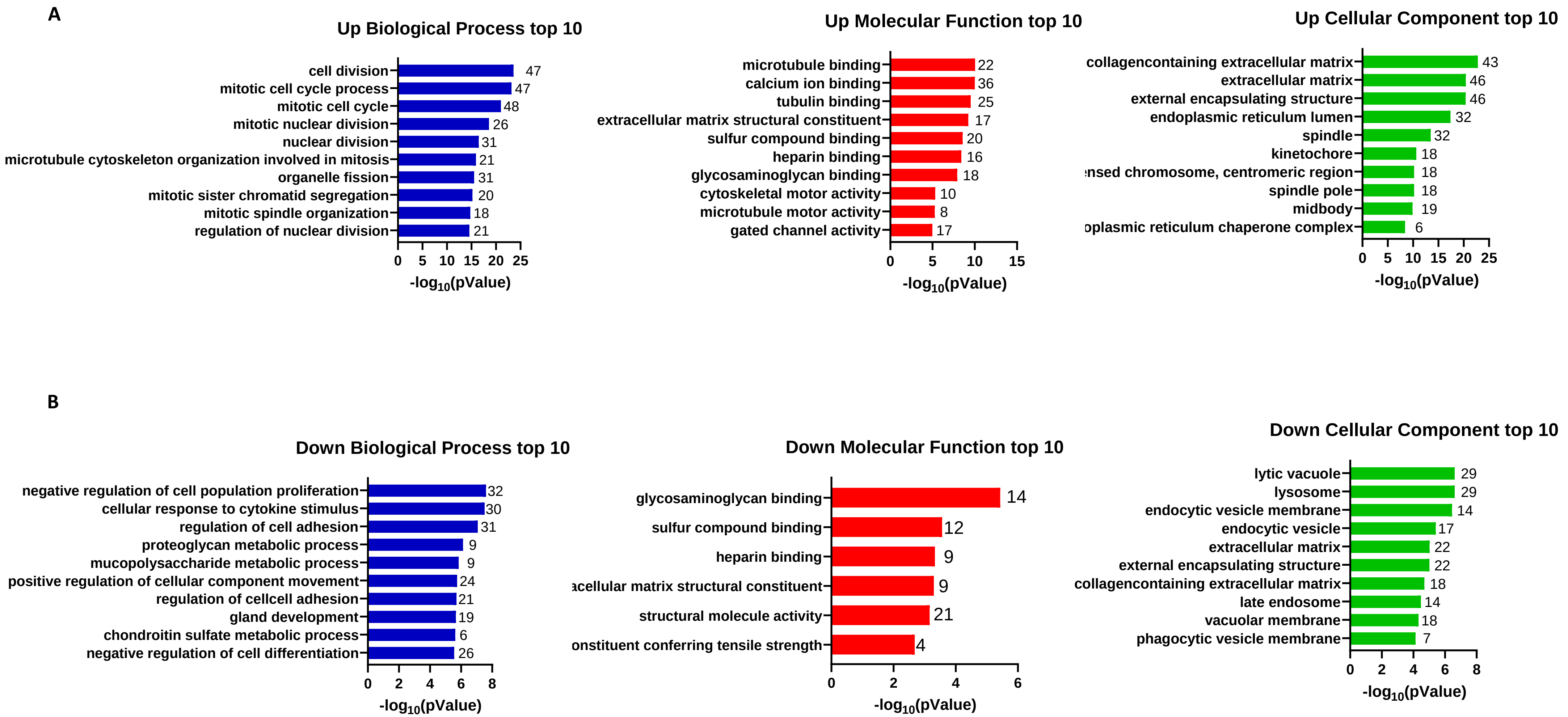


Figure 3.12 - GO analysis of DEGs in DU145 shRNA4 SPAG5 deficient. Bar graphs show the most 10 most enriched Gene Ontology (GO) terms (biological process, molecular function, cellular component) upregulated (A) and downregulated (B) in DU145 SPAG5 silencing. Data were obtained using Metascape software and sorted for p-value <= 0.05. Molecular function enrichment just six terms were found statistically significant.

For top 10 biological process, the GO enrichment analysis showed that shRNA4 SPAG5-upregulated gene (Fig.3.11 A) were statistically significant in ‘cell division’, ‘mitotic cell cycle process’ and ‘mitotic cell division’ while for molecular function the top 10 GO for upregulated gene were shown in ‘microtubule binding’, ‘calcium binding’ and ‘tubule binding’. For downregulated genes (Fig.3.11 B) the GO enrichment showed that the most statistical significant gene were present in ‘negative regulation of cell population proliferation’, ‘cellular response to cytokines stimulus’ and ‘regulation of cell adhesion’. In the molecular function enrichment, analysis with Metascape software revealed that the in SPAG5 deficient cells, downregulated genes were mostly enriched in ‘glycosaminoglycan binding’, ‘sulphur compound binding’ and ‘heparin biding’. ‘Lytic vacuole’, ‘lysosome’ and ‘endocytic vesicle membrane’.

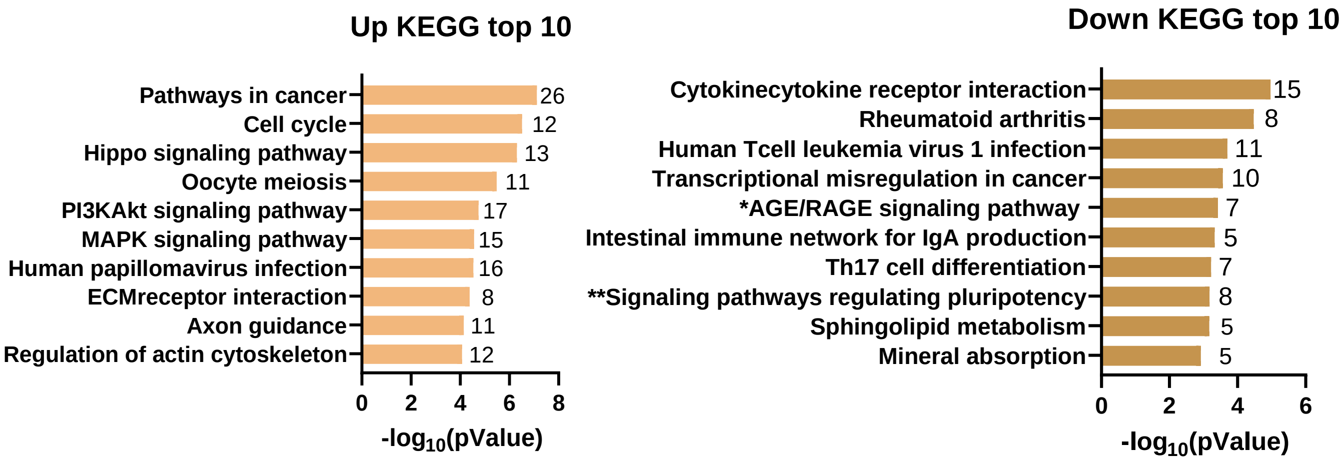


Figure 3.13 - KEGG analysis of differentially gene expression in DU145 SPAG5 deficient. Bar graphs show the most 10 most enriched KEGG pathways upregulated and downregulated in DU145 SPAG5 silencing. Data were obtained using Metascape software and sorted for p-value <= 0.05. *AGE/RAGE signalling pathway in diabetic complication. ** Signalling pathways regulating pluripotency of stem cells.

From the KEGG pathway analysis it was shown that the upregulated genes were mainly associated with ‘pathways in cancer’, ‘cell cycle’ and ‘Hippo signalling pathway’. ‘Cytokines receptor interaction’, ‘Rheumatoid arthritis’ and ‘human T cell leukaemia virus 1 infection’ pathways were mainly enriched in the downregulated gene set.

Next the DEGs were uploaded in Cytoscape (v3.9.1) to construct the protein-protein interaction (PPI network) using STRING protein query for MDA-MB-231 and DU145 shRNA4 SPAG5. For the analysis, 500 genes which were significantly modulated were uploaded in the software using a high confidence (0.700). GO analysis was performed using the coloured donut cart for clustering genes in the specific GO biological process based on the p-value (Fig. 3.11 and 3.12). Upregulated genes were mainly enriched in the ‘Regulation of multicellular organismal process’ and ‘Biological adhesion’ in MDA-MB-231 shRNA4 cells (Fig.3.11). Downregulated genes were enriched in ‘Cell cycle’ and ‘Cell migration’ (Fig.3.12). Generation of PPI network was obtained for DU145 SPAG5 shRNA4 using Cytoscape (v3.9.1). For DU145 shRNA4 SPAG5

200 upregulated genes were analysed, and GO Biological process showed that most of the upregulated genes were enriched in the 'Multicellular organism development' and 'regulation of cell communication' (Fig.3.13). Finally, downregulated genes network was analysed considering all the 435 genes for GO Biological process and the most of DEGs were enriched in the 'Regulation of metabolic process' and 'Regulation of response to stimulus' (Fig.3.14).

A

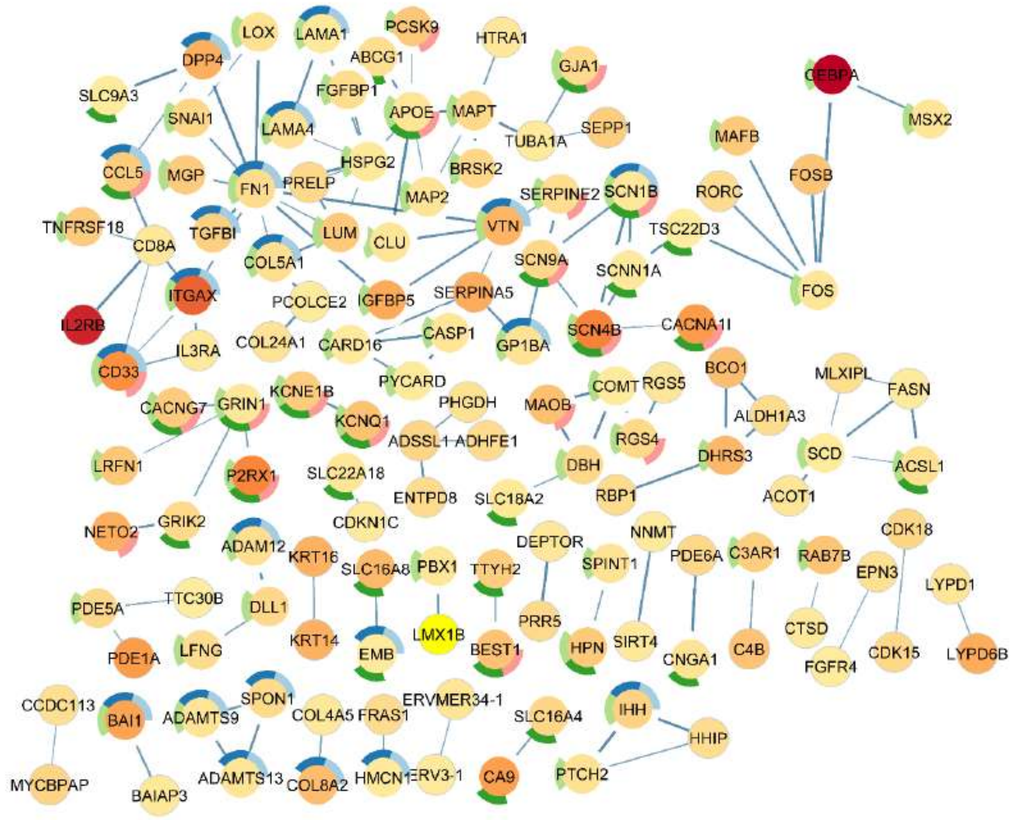


Figure 3.14 - Protein-protein interaction networks on DEGs upregulated identified in RNA-seq analysis in MDA-MB-231 shRNA4 SPAG5 deficient cells. A The PPI of the DEGs upregulated genes were visualised in Cytoscape (v3.9.1). Each node (proteins) was coloured based on the Log2FC with the darkest red indicating the highest expression. The thickness of the node connecting represents the size of the comparisons. Split donuts chart around the nodes, represents five non-redundant enrichment cluster of the upregulated genes, and are reported in table (**B**). The entire set of genes exhibiting a Log2FC cut off 0.58 in gene expression included 1121 upregulated. To allow a proper visualisation network, it was chosen arbitrarily 500 genes that exhibited the largest modulation with setting at high confidence (0.700). **C** Table showing the analysis of the network generated from Cytoscape where nodes representing proteins and the edges the interaction between two proteins. Significancy in the number of edges and nodes are represented from the PPI enrichment p-value.

B

Category	**Chart colour**	**Description**	**pValue**
GO Biological process		Biological adhesion	0.0000954
GO Biological process		Cell adhesion	0.0001
GO Biological process		Regulation of multicellular organismal process	0.0028
GO Biological process		Ion transport	0.0076
GO Biological process		Regulation of ion transport	0.0081

C

Network Stats
number of nodes: 409
number of edges: 125
PPI enrichment p-value: 3.71e-12

A

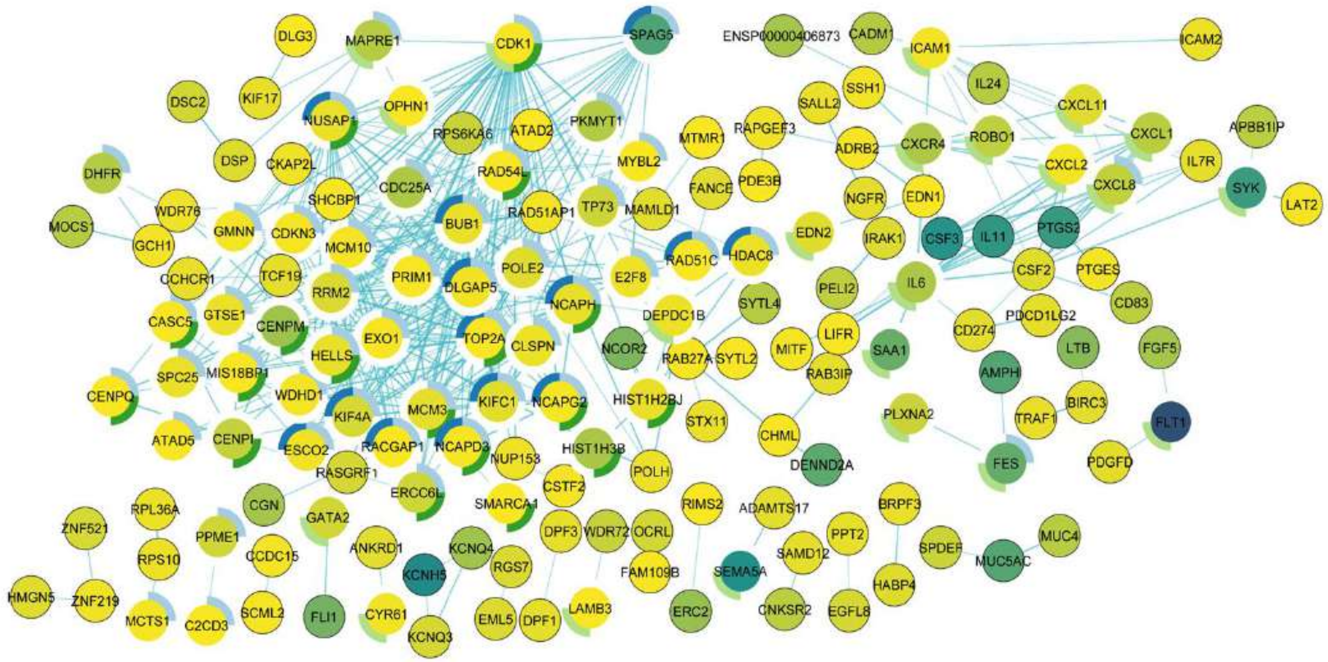


Figure 3.15 - Protein-protein interaction networks on DEGs downregulated identified in RNA-seq analysis in MDA-MB-231 shRNA4 SPAG5 deficient cells. (A) The PPI of the DEGs downregulated genes were visualised in Cytoscape (v3.9.1). Each node (proteins) was coloured based on the Log2FC with the darkest red indicating the highest expression. The thickness of the node connecting represents the size of the comparisons. Split donuts chart around the nodes represents five non-redundant enrichment clusters of the downregulated genes as reported in table **(B)**. The entire set of genes exhibiting a Log2 FC cut-off 0.58 in gene expression included 1080 downregulated. To permit a proper visualisation it was chosen arbitrarily selected the 500 genes that exhibited the largest modulation with a setting at high confidence (0.700). **(C)** Table showing the analysis of the network generating from Cytoscape where nodes representing proteins and the edges the interaction between two proteins. The significancy of the number of edges and nodes are represented by the PPI enrichment p-value.

B

Category	Chart colour	Description	pValue
GO Biological process		Cell cycle	0.00037
GO Biological process		Sister chromatid segregation	0.0046
GO Biological process		Cell migration	0.0332
GO Biological process		DNA conformation change	0.0332

C

Network Stats
number of nodes: 396
number of edges: 466
PPI enrichment p-value: < 1.0e-16

A

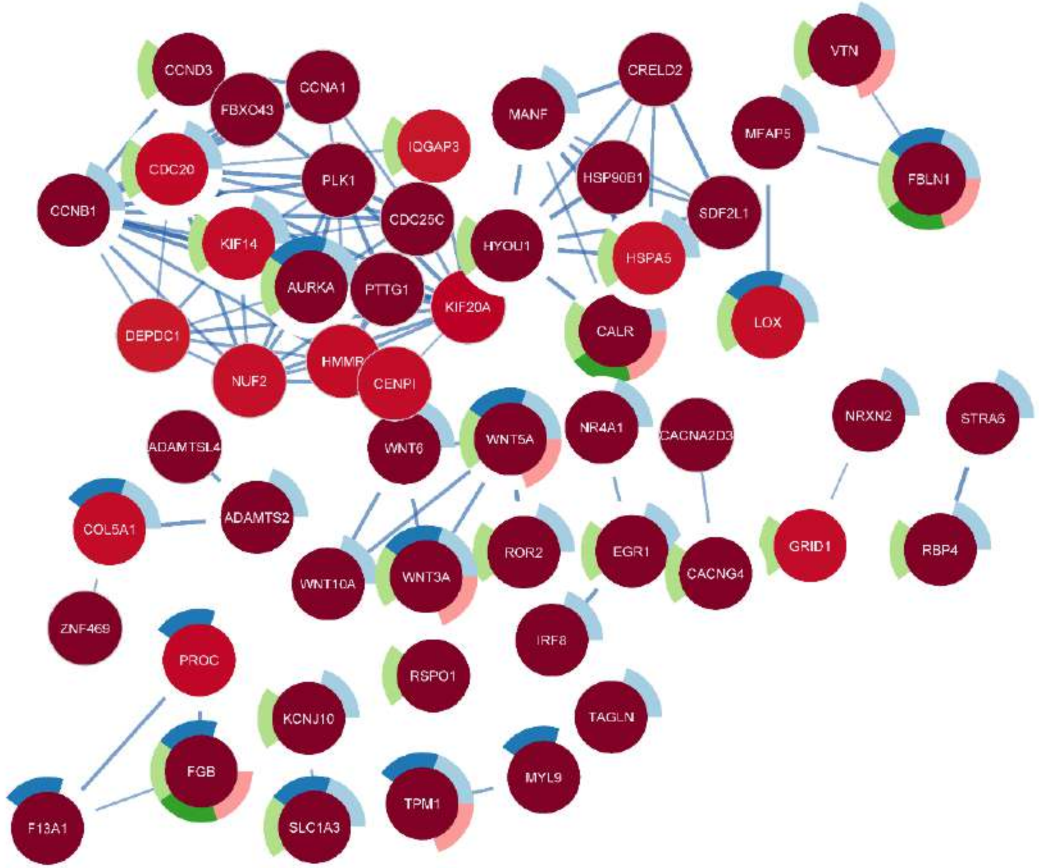


Figure 3.16 - Protein-protein interaction networks on DEGs upregulated identified in RNA-seq analysis in DU145 shRNA4 SPAG5 deficient cells. (A) The PPI of the DEGs upregulated genes were visualised in Cytoscape (v3.9.1). Each node (proteins) was coloured based on the Log2FC with the darkest red indicating the highest expression. The thickness of the node connecting represents the size of the comparisons. Split donuts chart around the nodes represents five non-redundant enrichment cluster of the upregulated genes and reported in table **(B)**. The entire set of gene exhibiting a Log2 FC cut off 0.58 in gene expression included 472 upregulated. To permit a proper visualisation it was chosen arbitrarily selected the 200 genes that exhibited the largest modulation with setting at high confidence (0.700). (**C)** Table showing the analysis of the network generating from CYTOSCAPE where nodes representing proteins and the edges the interaction between two proteins. Significancy in the number of edges and nodes are represented from the PPI enrichment p-value.

B

Category	Chart colour	Description	pValue
GO Biological process	(light blue)	Multicellular organism development	0.00068
GO Biological process	(dark blue)	Response to wounding	0.00068
GO Biological process	(light green)	Regulation of cell communication	0.00068
GO Biological process	(dark green)	Positive of substrate adhesion-dependent cell spreading	0.00088
GO Biological process	(pink)	Positive regulation of cell adhesion	0.0012

C

Network Stats
number of nodes: 168
number of edges: 110
PPI enrichment p-value: < 1.0e-16

A

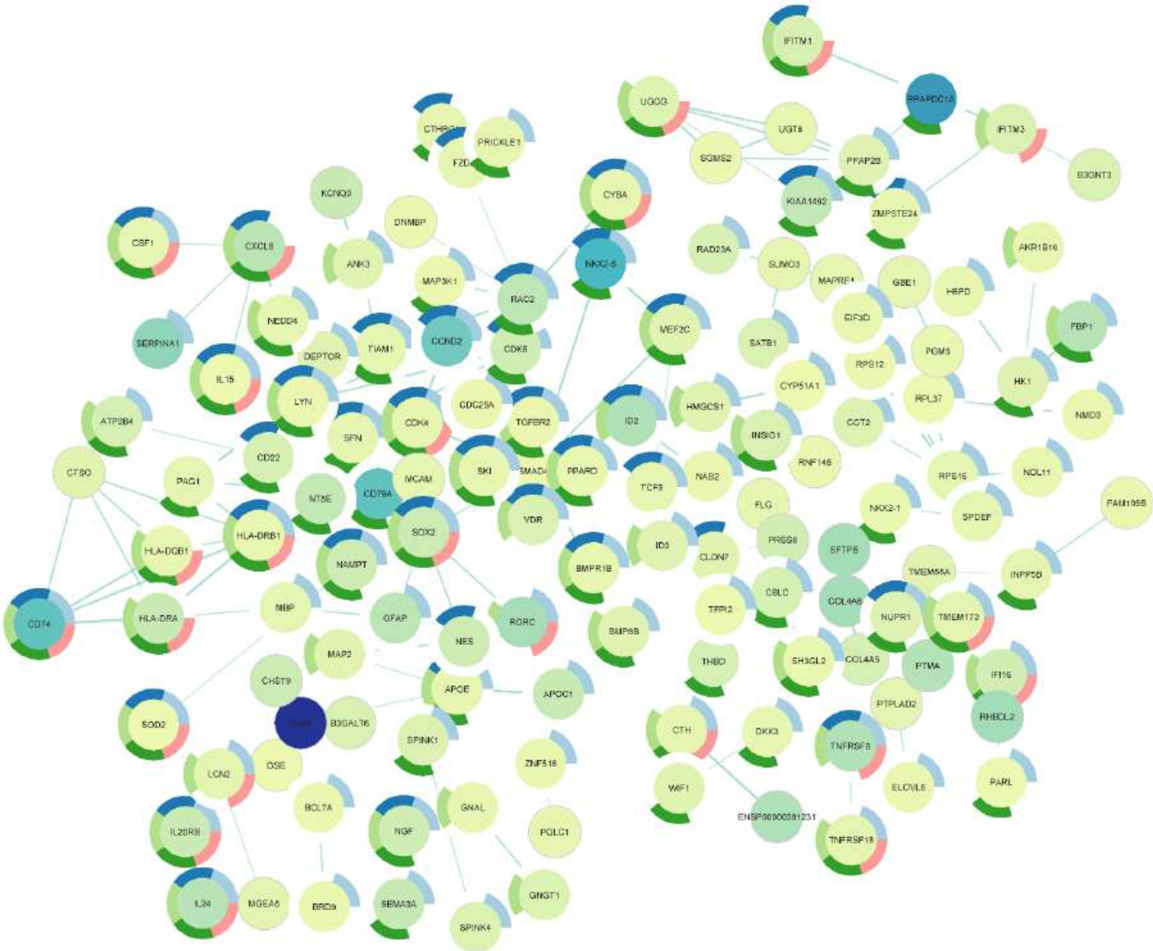


Figure 3.17 - Protein-protein interaction networks on DEGs downregulated identified in RNA-seq analysis in DU145 shRNA4 SPAG5 deficient cells. (A) The PPI of the DEGs downregulated genes were visualised in Cytoscape (v3.9.1). Each node (proteins) was coloured based on the Log2FC with the darkest red indicating the highest expression. The thickness of the node connecting represents the size of the comparisons. Split donuts chart around the nodes represents five non-redundant enrichment cluster of the downregulated genes and reported in table **(B)**. The entire set of gene exhibiting a Log2 FC cut off 0.58 in gene expression included 435 downregulated. Network was generated at high confidence (0.700). **(C)** Table showing the analysis of the network generating from CYTOSCAPE where nodes representing proteins and the edges the interaction between two proteins. Significancy in the number of edges and nodes are represented from the PPI enrichment p-value.

B

Category	Chart colour	Description	pValue
GO Biological process		Regulation of metabolic process	0.00048
GO Biological process		Regulation of cell population proliferation	0.00048
GO Biological process		Cellular response to chemical stimulus	0.0014
GO Biological process		Regulation of response to stimulus	0.0035
GO Biological process		Cellular response to cytokine stimulus	0.0035

C

Network Stats
number of nodes: 375
number of edges: 138
PPI enrichment p-value: 3.22e-05

In order to investigate commonly upregulated and downregulated genes between MDA-MB-231 and DU145 SPAG5 deficient, DEGs from both cell lines were analysed using an interacting tool using Venny software (v 2.0.2) (Oliveros, 2007).

Table 3.7 - Top 20 the most upregulated genes commonly expressed in MDA-MB-231 and DU145 SPAG5 deficient cells. Commonly upregulated genes are presented for Log2 FC showing the difference of expression between the two cell lines and sorted for the most significant p-Value <0.05.

MDA-MB-231

Gene ID	log2FC	Pvalue
COL5A1	1.3712204	4.29E-216
LOX	1.2770694	1.46E-97
UACA	0.6492913	3.79E-67
ABCG1	1.3546954	3.8E-55
ACSS1	0.8555519	1.64E-36
PPIB	0.6283404	2.7E-35
MLPH	0.6696059	4.24E-32
ABCA1	0.6121098	1.46E-24
SLC1A1	1.7651492	1.1E-21
MAP2K6	1.6420816	1.01E-17
IFITM10	1.8268234	1.92E-17
VTN	2.4822614	1.35E-16
TBX1	1.1605432	2.55E-16
COL4A1	0.646916	9.97E-16
PSMG3-AS1	1.387122	1.1E-14
GXYLT2	0.7763276	2.03E-11
ADRA2C	2.5515702	3.18E-09
RNF224	1.2135133	3.68E-08
PLA2R1	1.0386759	1.34E-07
GPRC5B	1.135041	0.00000017

DU145

Gene ID	Log2FC	Pvalue
SLC1A1	1.087195	4.71E-24
COL5A1	0.957377	3.11E-18
VTN	2.036796	2.81E-17
COL4A1	0.658129	4.02E-17
RHOB	0.644178	3.08E-14
MLPH	0.712739	2.53E-13
C1QTNF6	0.661745	8.49E-13
PCSK1N	0.717357	1.21E-11
PPIB	0.719939	9.43E-11
PCYOX1L	0.664453	2.91E-10
FBXL16	0.662561	8.64E-10
GPRC5B	0.745429	2.37E-09
IFITM10	0.818397	1.19E-08
PIGP	0.587777	3.82E-08
UACA	0.595434	4.54E-08
IQCA1	1.508737	3.21E-07
DISP2	1.466851	3.8E-07
ABCG1	0.910617	8.73E-07
SLC9A3-AS1	0.655729	1.39E-06
VASH1	0.589448	4.66E-06

Table 3.8 - Top 20 the most downregulated genes commonly expressed in MDA-MB-231 and DU145 SPAG5 deficient cells. Commonly downregulated genes are presented for Log2 FC showing the difference of expression between the two cell lines and sorted for the most significant p-Value <0.05.

MDA-MB-231

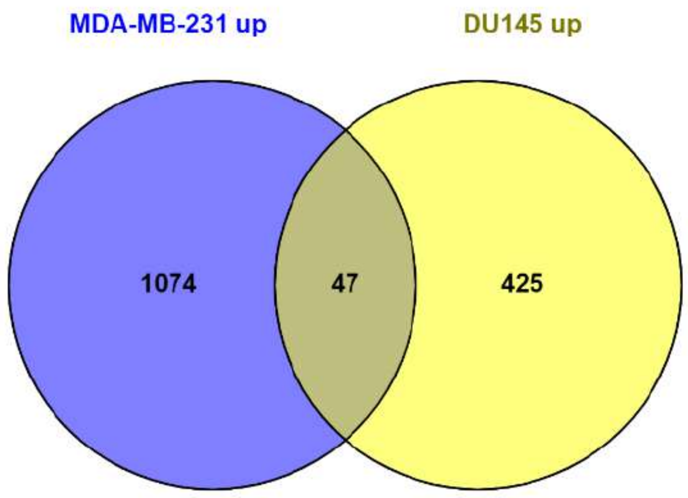
Gene ID	log2FC	Pvalue
SPAG5	-2.76653	0
MAPRE1	-1.61809	0
SMIM13	-1.49776	2.1E-103
CENPM	-1.87795	4.1E-102
AIDA	-1.70707	1.87E-95
SSH1	-0.94024	1.33E-86
CDC25A	-1.70765	8.15E-83
SLC16A3	-0.73274	9.25E-81
GIT1	-1.10858	5.19E-73
ZMPSTE24	-1.00217	7.85E-72
CCT2	-0.73496	1.18E-69
RAD23A	-0.94445	1.16E-65
HK1	-0.63923	1.52E-61
SLC2A4RG	-0.81297	9.44E-61
SCAMP1	-1.27594	5.25E-59
ATL3	-0.649	1.41E-58
TRIM37	-1.04507	4.6E-58
TMEM87B	-1.20655	6.36E-58
CUL2	-0.99373	6.05E-49
AC138392.1	-1.80321	1.07E-46

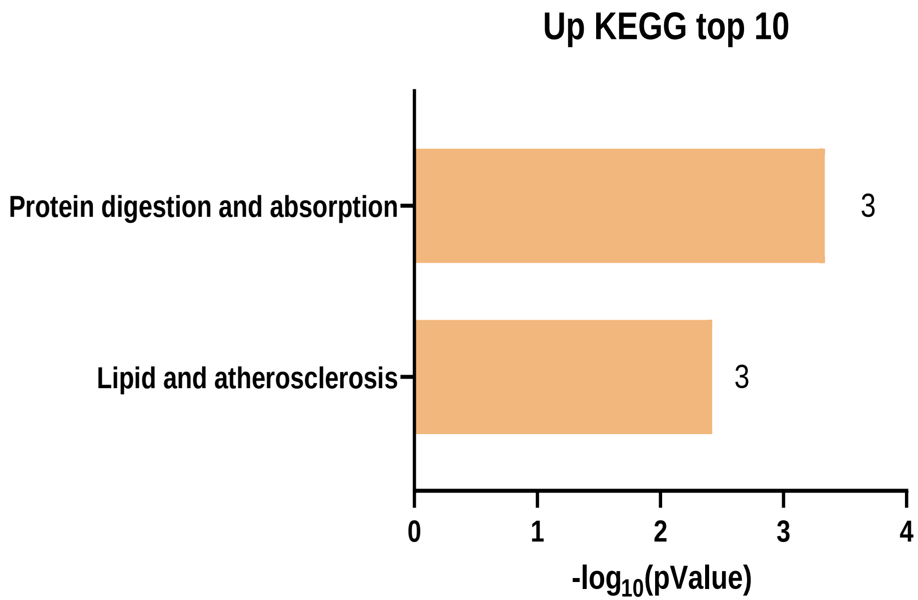
DU145

Gene ID	Log2FC	Pvalue
BCL2L13	-1.02969	4.19E-72
SCAMP1	-1.28882	1.48E-64
CCT2	-0.89256	5.39E-50
RAD23A	-1.00647	1.59E-44
GIT1	-1.03589	2.15E-41
ATL3	-0.81326	2.62E-41
HK1	-0.87562	1.88E-34
NEDD4	-0.73802	4.14E-34
SPAG5	-1.33389	6.32E-33
SLC4A4	-3.20788	1.05E-31
HIPK3	-0.76363	1.35E-28
B3GALT6	-0.94366	7.69E-28
VPS37B	-0.97745	2.96E-26
TRIM37	-0.82082	4.42E-26
SLC16A3	-0.60681	7.5E-24
LIFR	-0.90427	1.02E-22
PPP1R37	-1.14861	1.09E-22
MAPRE1	-0.70997	1.46E-18
SSH1	-0.59976	1.18E-17
KCNQ3	-1.32988	3.26E-17

The common 47 upregulated genes and 70 downregulated genes were then analysed for enrichment using Metascape tools. For each given list of genes pathway and process enrichment analysis was generated KEGG pathways for both downregulated and upregulated genes were shown while complete ontology sources (Fig.3.12) (GO Biological process, GO Cellular components, GO Molecular functions are presented in the appendix Table 3.7).

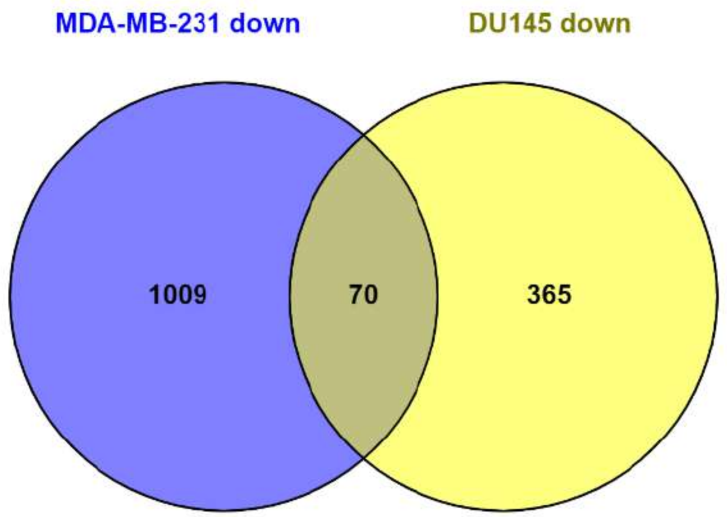
Terms with a p-value < 0.05, a minimum of count of 3, and an enrichment factor > 1.5 were collected and grouped in cluster based on their membership similarities. Graph for each term was created for top 10 with the best p-value was shown. Tables presenting the most significant upregulated and downregulated genes (complete table is shown in the appendix Table 3.7- Table 3.8 chapter 3). KEGG pathway showed that genes commonly upregulated in the two cell lines MDA-MB-231 and DU145 SPAG5 deficient, were enriched in just two terms: 'Protein digestion and absorption and 'lipid atherosclerosis' for upregulated genes and for the downregulated, genes are most enriched in 'cytokine-cytokine receptor interaction' pathway (Fig.3.15).

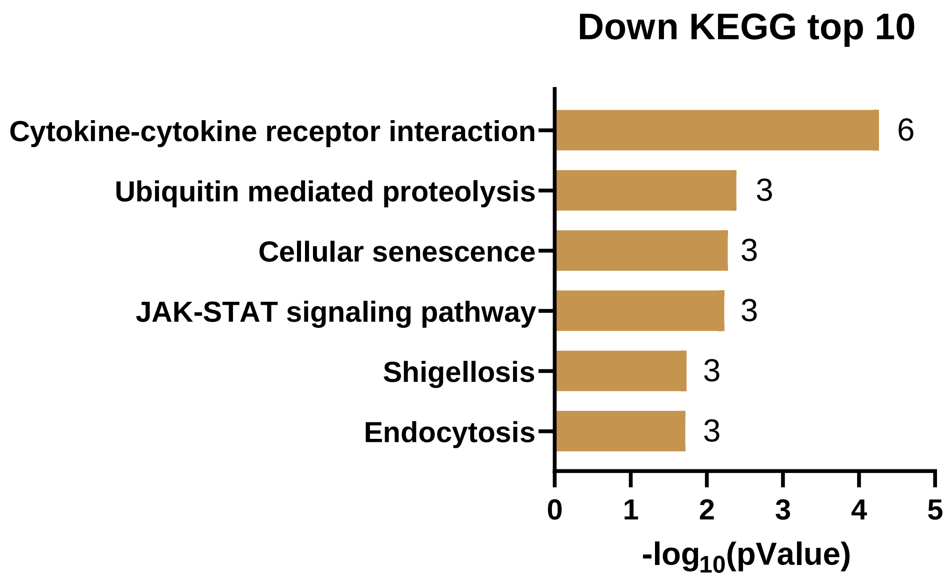




Term	Description	Log pValue	Gene list	Symbols
hsa04974	Protein digestion and absorption	-3.33627	3/103	COL4A1, COL5A1, SLC1A1
hsa05417	Lipid and atherosclerosis	-2.42002	3/215	ABCA1, MAP2K6, ABCG1

Figure 3.18 - Enrichment analysis on common upregulated in MDA-MB-231 and DU145 SPAG5 deficient cells. Commonly DEGs upregulated and downregulated genes between MDA-MB-231 and DU145 shRNA4 SPAG5 were analysed using Venny diagram tool (v 2.0.2). Enrichment analysis was obtained considering the most significant 47 upregulated and 70 downregulated genes commonly expressed between the two-cell lines and analysed by Metascape. KEGG pathways were generated, and the top 10 terms were shown based on the best p-value <0.05. The numbers next to each bar represent the upregulated and downregulated genes enriched in each term represented in the tables.





Term	Description	Log pValue	Gene list	Symbols
hsa04060	Cytokine-cytokine receptor interaction	-4.26453	6/295	BMP8B, CXCL8, TNFRSF9, LIFR, IL32, IL24
hsa04120	Ubiquitin mediated proteolysis	-2.38926	3/142	TRIM37, NEDD4, CUL2
hsa04218	Cellular senescence	-2.27562	3/156	CDC25A, CXCL8, HIPK3
hsa04630	JAK-STAT signalling pathway	-2.23029	3/162	GFAP, LIFR, IL24
hsa05131	Shigellosis	-1.73656	3/247	HK1, CXCL8, RPS6KA5
hsa04144	Endocytosis	-1.71829	3/251	NEDD4, GIT1, VPS37B

Figure 3.19 - Enrichment analysis on common downregulated genes in MDA-MB-231 and DU145 SPAG5 deficient cells. Commonly DEGs upregulated and downregulated genes between MDA-MB-231 and DU145 shRNA4 SPAG5 were analysed using Venny diagram tool (v 2.0.2). Enrichment analysis was obtained considering the most significant 47 upregulated and 70 downregulated genes commonly expressed between the two-cell line and analysed by Metascape. KEGG pathways were generated, and the top 10 terms were shown based on the best p-value <0.05. The numbers next to each bar represent the upregulated and downregulated genes enriched in each term represented in the tables.

3.4 Discussion

The present part of the study aimed to perform the analysis of the entire transcriptome in two cancer cell line, the triple-negative breast cancer cell line MDA-MB-231 and the androgen-receptor positive prostate cancer cell line DU145 after the stable knockdown of SPAG5. The first part of the study was focused on the gene expression profiling of SPAG5 to determine which cell line in both breast and prostate cancer would be suitable for the RNA silencing. Four breast cancer cell line and two for prostate cancer were used. Through the RT-qPCR analysis there was shown no statistical difference between the cell line therefore it was decided to transduce with the shRNAs SPAG5 for breast cancer the triple negative MDA-MB-231 cell line (Fig.3.1A) as considered the most aggressive and poor differentiated form, while for prostate cancer the most aggressive form and easier cell to culture DU145 cell line (Fig.3.1B). Stable SPAG5 knockdown was generated for MDA-MB-231 and DU145 cell line and each construct efficiency was assessed at transcript level (Fig.3.3) using qPCR. The construct demonstrated the best knockdown efficiency was also assessed at protein level (Fig.3.5) in both cell line. The best efficiency knockdown shRNA4 -SPAG5 was selected for the RNA-seq analysis in both breast and prostate cancer cell line using a negative control pLKO.1 (empty vector) construct for its off-target effect. Stable cell lines generated that integrated the shRNA-SPAG5 silencing were then sent for RNA-seq analysis. In the first part of the analysis, it was assessed the differential expressed genes in MDA-MB-231 shRNA4 SPAG5 versus Control pLKO.1 empty vector and DU145 shRNA4 SPAG5 versus control pLKO.1. Normalisation of the count by FPKM methods confirmed the silencing of SPAG5 in both cell line (Fig.3.7 A) as obtained from qRT-PCR and western blot. The transcriptome analysis identified for MDA-MB-231 2,201 DEGs in which 1,121 were upregulated and 1,080 downregulated (Fig.3.8 C). Collagens are essential components the ECM composition. In the vertebrated were identified 46 different types of collage that can be grouped in 28 different collagens (Shoulders & Raines, 2009). In breast cancer modification of collagen composition are observed with increasing of the collagen types I, III and V. Here it was shown that DEGs that upregulated genes were enriched in extracellular matrix (ECM), including *COL1A2, COL4A5, COL5A1, COL6A2, COL8A2*. It was also demonstrated that downregulation of SPAG5 is correlated with enrichment in the negative regulation of cell differentiation term. Here was shown that the downregulation is related to an increase of MMP11. The matrix-metalloproteinase (MMP11) was demonstrated to be highly expressed in breast cancer intratumoral mononuclear inflammatory cells (MICs) (Eiro et al., 2019).Upregulated genes *CTSD* and *CLU* were seen in the DEGs MDA-MB-231 SPAG5 knockdown. Cathepsin D deficiency (CTSD) is a major endopeptidase with primary location in the endosomal/lysosomal compartments. In breast cancer is associated with poor prognosis. In a recent study CTSD was linked to estrogen receptor (Kang, J. et al., 2020).Clustering gene (*CLU*) is highly conserved gene during the evolution, and it appears to be expressed in different cell compartments and widely distributed in different species (Koltai, 2014). CLU is implicated in different biological function including DNA repair, cell adhesion, tumorigenesis (Rohne et al., 2016). Study conducted on breast cancer demonstrated that CLU is correlated with tumour grade and progression from primary carcinoma to invasive breast cancer (Ming et al., 2018). Secreted protein clusterin (sCLU) is also stress-activated survival factors functionally associated with treatment resistance (Redondo et al., 2000).

DEGs upregulated in SPAG5 knockdown were analysed for PPI trough Cytoscape and a map was created (Fig.3.14). Protein-protein association of *FN-1* with *TGF-β* was observed. DEGs for SPAG5 knockdown shown an upregulation of genes involved in cancer progression as the beta transforming growth factor (*TGF-β*). In breast cancer *TGF-β* is hyperactivated and is responsible of cancer progression and metastases (Tang et al., 2018).The *TGF-β* pathways has been used as a target in different solid cancer, in particular small molecules inhibitors of *TGF-β* receptor kinases activity were designed to bind the ATP-binding domain in *TGF-β* receptor kinases to inhibit the ATP kinases activity (Huynh et al., 2019). Vactosertib an orally available inhibitor of *TGF-β* receptor has shown antitumour effect in various xenograft models with hepatocellular carcinoma, B16/F1 melanoma and 4T1 breast cancer cell line (Jung, Su Young et al., 2020). Based on the protein network generated *FN-1* is also associated with EMT marker the zinc finger protein *SNAI1* which is involved in drug resistance, tumour recurrence and metastasis. In human breast carcinoma, MDA-MB-231 *SNAI1* has been shown to be required for lymph node metastases (Olmeda et al., 2007). A study conducted on breast cancer showed that *SNAI1* influences the sensitivity to the chemotherapy drug tamoxifen through its silencing which affect the estrogen receptor (ER) transcriptional activity and contrary to its activation decreases the ER activation maintaining the cell resistance to tamoxifen (Scherbakov et al., 2012). Lysis oxidase (*LOX*) are family of proteins involved in cell adhesion, motility, and invasion. Downregulation of SPAG5 determines an increase of *LOX* gene expression and PPI network shown an association with *FN-1*(Fig.3.15). Studies conducted in pre-clinical breast cancer models showed the capacity of *LOX* to generate a pre-metastatic niche (Erler et al., 2009). In lung cancer, inhibition of its expression is correlated with a suppression of metastasis (Rachman-Tzemah et al., 2017). *LOX* expression is also correlated with resistance to chemotherapy. In triple-negative breast cancer (TNBC) it was demonstrated that inhibition of *LOX* decreases the expression of *FN-1* and the integrin alpha 5 (*ITGA5*) which affect the extracellular collagen cross-linking and decrease fibronectin/assembly facilitating drug penetration (Saatci et al., 2020). Downregulation of SPAG5 showed an upregulation of the insulin-like growth factor binding protein 5 (*IGFBP5*) which it showed to be co-expressed with *FN-1* protein in the protein-protein interaction network. In breast cancer, *IGFBP5* is associated with metastatic tumour phenotype and poor prognosis that could potentially use as a target for therapeutic development (Li, X. et al., 2007) (Wang, Huamin et al., 2008).Interestingly, a study conducted in breast cancer cell line MCF-7 indicated that low expression of *IGFBP5* was associated with a low survival rate after being treated with tamoxifen, indicating a role of *IGFBP5* as a reverse of tamoxifen resistance (Ahn et al., 2010). A precedent was observed in mammary carcinoma xenograft microarray data in which relative expression of *IGFBP5* affects tamoxifen responsiveness (Becker et al., 2005).Downregulation of SPAG5 showed an upregulation of Fos Proto-Oncogene, AP-1 Transcription Factor Subunit (*FOS*). In study conducted using The Cancer Genome Atlas dataset and bioinformatic tools it was shown that high expression of *FOS* is correlated with higher survival rate and using microarray data it was demonstrated a high correlation between *FOS* and pro-apoptotic gene expression (Fisler et al., 2018).A co-expression between *FOS* and TSC22 Domain Family Member 3 (*TSC22D3*), an anti-inflammatory protein glucocorticoid (GC)- induced leucine zipper. A study conducted by Yang and co-workers showed that high levels of *TSC22D3* were seen in sera in patients affected by colorectal cancer and non-small-cell lung cancer (NSCLC) when compared with healthy volunteers (Yang, H. et al., 2019). In mice was noticed that psychology and metabolic stress, via glucocorticoids, affect cancer progression and treatments. In breast cancer a normal level of glucocorticoids, and stress linked are correlated to a reduced survival rate (Sephton et al., 2000). SPAG5 plays an important role in the cell cycle as it is involved in the chromosome alignment, sister chromatid segregation and spindle pole formation. The transcriptome analysis showed that the downregulation of SPAG5 affects the mitotic cell cycle and cell division including the microtubule-associated protein RP/EB Family Member 1 (*MAPRE1*) which is significantly downregulated. A study conducted on breast cancer demonstrated the role of *MAPRE1* as a prognostic and diagnostic biomarker, which also correlated with a poor clinical outcome (Rodrigues-Ferreira et al., 2019). Protein-protein Interaction network generated showed a co-expression between MAPRE1 and the Cyclin-dependent kinase 1 (*CDK1*) (Fig 3.12). CDK1 drives cells from the G2 phase to the mitosis process. MAPRE1 is involved in the mitosis process as it is associated with centrosome and spindle microtubule orientation as also spindle positioning in response to cell shape and adhesion (Sun, L. et al., 2008). Results showed co-expression between MAPRE1 and CDK1, this could suggest a potential involvement of MAPRE1 in the progression of cells through the mitosis process. One of the cancer hallmarks is aberrant cell proliferation due to cell cycle dysregulation, therefore hypothesised to target the cell division process for cancer treatment (Hanahan & Weinberg, 2011b). Cyclin-kinase dependent 4/6 (CDK4/6) inhibitors are already approved by FDA for the treatment of metastatic hormone receptor-positive breast cancer (Fry et al., 2004). However, most of the first-generation inhibitors are not approved for clinical application due to their non-selectivity and toxicity (Whittaker et al., 2017). Flavopiridol an Inhibitor of CDK1, CDK2, CDK4, CDK6, CDK7 and CDK9 shown to be effective for the treatment of different solid cancer, however, poor efficacy was observed in some solid cancer and a new combination of therapies is promoting (Asghar et al., 2015). Downregulation of MAPRE1 in SPAG5 knockdown breast cancer cell line and co-expression with CDK1 could be a potential target for dysregulation of cell cycle division. The previous study in patients with estrogen receptor-positive breast cancer has shown SPAG5 transcript as a potential target for novel therapeutic strategy in endocrine therapy resistance (Abdel-Fatah et al., 2016b). DEGs also identified that the knockdown of SPAG5 affected the downregulation of genes involved in the

DNA damage response stimuli including the mini-chromosome maintenance component complex 3 (MCM3) and the Non-POU Domain-Containing Octamer-Binding Protein (NONO). MCM3 ensures that the DNA is replicated once per cell cycle in eukaryotes, and it also drives the elongation during replication acting as a helicase and is overexpressed in different tumours including breast cancer (Ha et al., 2004). NONO is an RNA-binding protein involved in transcriptional regulation and RNA splicing (Emili et al., 2002). In a study conducted on breast cancer NONO has been shown to be responsible for cancer proliferation through regulation of the E3 ubiquitin ligase *SKP2* and the transcription factor 8 (*E2F8)* (Iino et al., 2020). Interestingly *E2F8*, which belongs to the E2F family a group of the protein involved in the control of DNA damage, is also downregulated in MDA-MB-231 SPAG5 knockdown associated with the biological process of mitotic cell cycle and regulation of the cell cycle process. Furthermore, KEGG enrichment analysis demonstrated that upregulated genes were enriched in the 'PI3K-Akt signalling pathway' and in 'ECM receptor interaction' while downregulated genes were enriched in 'Cell cycle', 'DNA replication' and 'NF kappa B signalling pathway' all together those can let to cancer progression (Fig.3.11).

Transcriptome analysis was generated for DU145 shRNA4 SPAG5 compared with pLKO.1 empty vector as a control. The analysis showed 907 DEGs of which 435 were downregulated and 472 upregulated. Nicotinamide phosphoribosyltransferase (NAMPT) is an enzyme that in human is encoded by *NAMPT* gene. This enzyme catalises the conversion of nicotinamide mononucleotide (NMN) and is essential for NAD biosynthesis. Inhibition of NAMPT determine inhibition of ATP through the depletion of NAD+ (Wei, Y. et al., 2022). Upregulation of NAMPT is accountable of malignancy including breast, colon, and prostate cancer (Gallí et al., 2010). Particularly, high expression of NAMPT is related to invasion and increased metastasis as also chemoresistance (Wang, B. et al., 2011). In colon cancer it was shown that NAMPT is a suitable oncogene that induce the stem cell pathways through NAMPT downstream effector SIRT1 and PARP1 (Lucena-Cacace et al., 2018). Inhibition of NAMPT is responsible of attenuation of glycolysis leading to in a further perturbation of metabolism in cancer cell (Busso et al., 2008). Result obtained from RNA-seq showed a significant downregulation of NAMPT in prostate cancer cell DU145 silenced for SPAG5 this could suggest a potential target using SPAG5 as guide for therapies. BCL2-like 13 (apoptotic facilitator) is a protein encoded by the gene *BCL2L13* and exhibits capacity of apoptosis-mediating in different cells line. In clear cell renal cell carcinoma (ccRCC) and papillary renal cell carcinoma (pRCC) showed reduced mRNA level and correlated with a lower survival probability (Meng et al., 2021). However, in glioblastoma (GBM) and childhood lymphoblastic leukaemia (ALL) BCL3L13 level were found elevated (Jensen et al., 2014) (Holleman et al., 2006).Secretory carrier-associated membrane protein 1 (SCAMP1), a molecule involved in the post-Golgi recycling pathway and in endosome cell membrane cycling is statistically downregulated in DU145 SPAG5 silencing (Castle & Castle, 2005) (Hubbard et al., 2000). Loss of SCAMP1 together with MTSS1 was already demonstrated to be correlated with reduced disease-specific free survival and in more aggressive cancer cell phenotype in HER2+/ER/PR- breast cancer (Vadakekolathu et al., 2018). As till now the role of SCAMP1 is still not clear our result could suggest that even in prostate cancer the downregulation of SCAMP1 could be related to increase of aggressiveness that could lead to resistance to chemotherapy. Sestrin3 (SESN3) belong to a small protein family implicated in biological process including anti-oxidative stress, anti-aging, cell signalling and metabolic homeostasis (Ho et al., 2016). In hepatocellular carcinoma (HCC) deficiency of SESN3 promotes carcinogenesis and accumulation of ECM and impairment of tissue repair in liver (Liu, Yunjian et al., 2019). In prostate cancer patients, to study the effect following external beam radiation therapy (EBRT) it was evaluate the change in the expression of sestrin genes family to see if there was association with changes during EBRT, and SESN3 was related to an intensification of EBRT and dysregulation of mTOR-AMPK pathway as also mitochondria impairment and oxidative stress (Gonzalez et al., 2018). Result obtained in DU145 SPAG5 silencing could suggest a potential role of SPAG5 in the SESN3 biological process. Interestingly downregulation of SPAG5 in DU145 is also led to the upregulation gene involved in the 'extracellular matrix' and 'collagen extracellular matrix' term as MDA-MB-231 shRNA4 SPAG5.

Genes upregulated in this term include *MYL9*, *MFAP5, ANXA6, ANXA8*.Recently studies on myosin light chain 9 (MYL9) have shown to be associated with tumorigenesis, invasion, and metastasis, however a prognostic and immunological role of MYL9 is not yet reported (Feng et al., 2022). Study using public dataset for TCGA showed that in different cancer MTL9 was lowly expressed including squamous carcinoma, stomach but highly expressed in head and neck carcinoma lever hepatocellular carcinoma (Lv et al., 2022). Result obtained from RNAseq in DU145 SPAG5 silencing has reported a significant upregulation of MYL9 in agreement with the finding obtained from TCGA dataset analysis. Microfibril-associated protein 5 (*MFAP5*) is a glycoprotein highly secreted from cancer-associated fibroblast (CAF) (Principe et al., 2018). In CAF multiple cancer MFAP5 is upregulated including prostate cancer and its over-expression led to consider MFAP5 as a marker for early detection (Jia et al., 2011a). Annexin 6 (*ANXA6*) and annexin 8 (*ANXA8*) are members of a family of Ca^{2+}-dependent membrane-binding annexing proteins involved in cell development particularly in the membrane cytoskeleton organisation also including cholesterol homeostasis and cell adhesion and signal transduction (Rescher & Gerke, 2004). In prostate cancer was documented that a lower expression of this marker *ANXA6* is correlated with a progression of cancer (Xin et al., 2003a). Differently from *ANXA8* which study conducted on a different type of cancer including breast, ovarian and prostate cancer showed that this marker is significantly overexpressed (Labrecque et al., 2019a). Results showed that both markers were upregulated in SPAG5 knockdown, however, *ANXA8* with a log2FC of 2.04 while *ANXA6* with a log2FC of 0.76 confirmed the different expression pattern between the two annexins in prostate cancer progression. KEGG pathways analysis showed that the downregulation of SPAG5 genes resulted in the most enriched pathways including the 'cancer pathways', 'PI3KAkt signalling pathway' and the 'Hippo signalling pathway' while downregulated genes were the most enriched in the 'Cytokine-cytokine receptor interaction'(Fig.3.13). GO Enrichment analysis showed that downregulation of SPAG5 led to an upregulation gene 'mitotic cell process' and 'cell division' including genes like BUB1 which in MDA-MB-231 are downregulated in DU145 are significantly upregulated (Fig.3.12). Aurora A Kinase (AURKA) an enzyme involved in the cell division process, together with BUB1 can positively regulate the cell process in DU145 shRNA4 SPAG5 populations. A positive correlation between SPAG5 with BUB1 and AURKA was also documented in breast cancer progression, therefore this could suggest a correlation in prostate cancer too (Zhu et al., 2019).

Finally, data sets from MDA-MB-231 and DU145 SPAG5 knockdown were combined to investigate whether the two transcriptomes could show common upregulated and downregulated genes and which pathways are commonly upregulated and downregulated. Venn diagram sorted genes for the most upregulated and downregulated genes in both cells line MDA-MB-231 and DU145 SPAG5 knockdown and enrichment analysis was generated. KEGG pathways analysis revealed that the most upregulated DEGs in both cell lines were mainly involved in the 'protein digestion absorption' and 'lipid atherosclerosis' (Fig.3.18). The results showed that the collagens gene (*COL4A1, COL5A1*) and glutamate transporter (*SLC1A1*) gene were directly associated with protein digestion absorption. Besides those two pathways, COL4A1 is also enriched in PI3K-Akt signalling pathway (Fig.3.11-3.13) which is proven to be important in the cell cycle as also proliferation and gene mutation and its activation is also seen in different cancer (Hennessy et al., 2005a)r. Overexpression of COL4A1 affects the progression and migration of breast cancer cells (Jin, R. et al., 2017). Downregulation of SPAG5 was already demonstrated in MCF-7 cell line through PI3K inhibitor and mTOR inhibitor and trastuzumab combined taxol therapy (Thedieck et al., 2013). In this study, triple negative MDA-MB-231 and DU145 genetically silenced with shRNA-SPAG5 determined the increased expression of genes involved in PI3K-Atk pathways.

Commonly downregulated genes between MDA-MB-231 and DU145 SPAG5 knockdown were then analysed, and results revealed that 70 genes were commonly expressed and presented in table 3.5 as the top 20 most significant downregulated. The most significant genes were sorted and presented in table 3.8. Among this MAPE1, SCAMP1, CUL2, RAD23A, LIFR shown. Interesting SCAMP1 where its loss is documented with increasing aggressiveness in HER2+/ER-/PR- shown in both cell line almost similar downregulation. This could be suggesting a potential involvement of SPAG5 with SCAMP1. KEGG pathways analysis showed that the most downregulated commonly expressed in both cell lines were enriched in the 'cytokine-cytokine receptor interaction pathways', 'Ubiquitin mediated proteolysis', 'Cellular senescence', 'JAK-STAT signalling pathway', 'Shigellosis' and 'Endocytosis' pathways (Fig.3.19). Cytokine-cytokine interaction pathways are involved in cancer development as the releasing of cytokine in response to inflammation or immunity could affect cancer progression (Mantovani et al., 2008). Results showed that the downregulation of SPAG5 interleukin -32 (IL-32) which is highly expressed (Sloot et al., 2018) in different cancer, is downregulated. An important role in cancers is related to the leukaemia inhibitor factor receptor (LIFR) including breast cancer (Chen, D. et al., 2012). It is documented that this receptor together with the leukaemia inhibitor factor (LIF) is commonly overexpressed in different solid cancer and is also responsible for the activation of the oncogenic pathway such as mTOR, and MAPK including JAK/STAT pathways (Liu, Shu-Chen et al., 2013). In particular, the LIF/LIFR axis is involved in tumour growth, progression, metastasis and stemness in solid cancers including therapy resistance (Morton et al., 2015). A study conducted on breast cancer revealed that targeting LIFR could inhibit tumour growth and demonstrated a more efficient effect by immunotherapy LIFR compared to immunotherapy LIF (Ghanei et al., 2020). Inhibition of LIF also suppresses tumour growth in prostate cancer. Because the LIF/LIFR axis is important in cancer progression this could suggest that a potential therapy targeting this axis could be effective also in prostate cancer. A recent study identified small molecules that specifically target LIF and LIFR under development (Viswanadhapalli et al., 2021). Overexpression of SPAG5 combined with tetracycline therapy showed an increase of 5 years of free survival in women with estrogen-receptor-positive breast cancer (Abdel-Fatah et al., 2020a). Results obtained showed that LIFR is commonly downregulated in both MDA-MB-231 and DU145 SPAG5 knockdown.

This experiment was designed to investigate what gene were affected by SPAG5 in MDA-MB-231 and DU145 and what pathways were involved. Results from this study has demonstrated that SPAG5 silencing could upregulated, and downregulated genes involved not only in cell cycle but also in pathways involved in drug resistance and in pathways highly involved in cancer progression. The combination of the prostate and breast data set showed commonly upregulated and downregulate gene giving new inside in triple-negative breast cancer as also In prostate cancer.

4.Quantitative proteomic mass spectrometry: identification of differentially expressed proteins in prostate DU145 and triple-negative breast cancer MDA-231 SPAG5 deficient cells

4.1 Introduction

Despite advanced cancer treatment, for many patients cancer can still be fatal (Brouckaert et al., 2013). As such it is of great benefit to use biomarkers to monitor or predict disease state, treatment outcome and prognosis. Specific biomarkers can be divided into five subtypes (Parise and Caggiano 2014). This classification is essential to guide on specific therapy; however, cancer recurrence can cause therapy resistance and metastasis (Van Nguyen et al., 2021). In prostate cancer (PCa), most patients show an indolent disease that can be effortlessly managed without immediate treatment, ensuring a good quality of life. However, in some cases, the disease can metastasise reaching other parts of the body at which point the prognosis worsens for patients. Only 28% of patients with metastatic PCa survive beyond 5-6 years (Nandana and Chung 2014). Prostate-Specific Antigen (PSA) represents the gold standard and the most used biomarker for diagnosing men with PCa, contributing to an over-treatment of the patients and negatively affecting the quality of life (Wachtel et al. 2013). Therefore, a new way to classify tumours could help to easily predict which patient group responds better to therapy and discovering improved biomarkers.

Mass spectrometry (MS) is an important and widely used technique in the area of small molecules and proteomics.

This technique is based on the ionisation and fragmentation of molecules in a gas phase. The technology of MS is generally linked with gas chromatography (GC) and or liquid chromatography (LC) for separation with the MS used as a detector (Pitt J.J. 2009) (Emwas et al. 2015). Therefore, samples are separated by chromatography in compounds and then enter into the mass spectrometry for sequentially ionisation, separation and detection of the ions generated. A mass spectrometry is formed from five crucial components: a vacuum system, an ion source, a mass analyser, an ion detector, and a data recording system (Cañas et al. 2006). Mainly, ions in the samples are produced in the ion source and then accelerated based on their *m/z,* which represents the ratio between the mass (*m*) and the charge number of the ion (*z*) (Bull, Lee and Vallance 2013). Then the ion signal is amplified and stored in a data recording system; this will generate a mass spectrum in which the x-axis represents the mass to charge (*m/z*) ratio and the y-axis the relative intensity (Sarvin et al., 2020). During the mass analysis inside the mass analyser sector, ions are separated under magnetic or electric fields travelling towards the detector. The presence of an electrostatic analyser before the magnetic sector can enhance the focus of the ions with the same (*m/z*) but, slightly different kinetic that improves the resolution (Smith 2013).

The time-of-flight analyser (TOF) is a high-resolution analyser that use an electric field to accelerate generated ions through the same electrical potential and measure the time by which ions reach the detector. The Ion's velocity depends on their *m/z* ratio, with the same kinetic energy. That means ions with lower-mass travel faster through the flight tube and can be separated before higher-mass ions (Boesl 2017).

In past years MS emerged as a fundamental technology for quantifying a multitude of proteins as well as their localisation, protein modification and interactions (Kragstrup et al., 2013). Tandem mass spectrometry, also called MS/MS, is used for this analysis. This technique uses two or more analysers that increase the ability to analyse samples. Once samples are ionised, the first spectrometer (MS1) separates the ions based on their *m/z* ratio and splits them into smaller ion fragments by collision-induced dissociation. Subsequently, they are introduced in another mass spectrometer (MS2) which separates the fragments by their *m/z* ratio and detects them (Gillet et al. 2012).

Two strategies for proteomic analysis /typically from digested proteins-a "bottom up" approach) are: full-scan, data-dependent acquisition (DDA) and data-independent acquisition (DIA). In the full-scan one, MS is used for ions generation of the molecules tested without fragmentation. This acquisition method provides a lower level of spectral information and metabolites, or in the case of proteomics, peptide identification. However, it is still used for some metabolic studies because of its easy acquisition, data processing and differentiation of biological samples (Clancy et al. 2018). In DDA mode, the acquisition-specific ions already selected from the full-scan spectrum are processed through MS full-scan and selected, isolated in MS1 and fragmented to generate MS2 or MSMS spectra (mainly different methods of fragmentation exist, commonly collisional induced dissociation or CID is used) usually in order of intensity e.g., the top 10,30 or 45 ions from the full scan per cycle. This mode of acquisition obtains some quantitative and structural information from the analyte. A limitation in DDA mode is due to high variety abundance in the analities eluting; therefore, if some of those analytes are low in abundance, there might be the risk that some of them are not detected from the MS spectrum or are not in the "Top X" ions detected in the full scan. This leads to DDA mode not being able to capture the MS/MS spectra for everything detected from the MS mode, leading to a loss in data acquisition, and "missing value" (Guo and Huan 2020).

Recent development of MS equipment improved not only the processing speed of the samples but also increased the sensitivity. This development led to a new strategy in proteomic analysis data-independent acquisition (DIA), known as sequential window acquisition of all theoretical fragment ion spectra (SWATH), the most reliable and accurate for studying proteome compared to the typically used DDA-MS (Arnhard et al. 2015). In DIA-MS, all the precursor ions identified in the first survey scan (MS1) within a selected mass-to-charge ratio (*m/z*) range are fragmented and analysed by MS. In DDA-MS, the MS spectra are obtained from a broad (*m/z*) range. Peaks are detected and sorted by descending intensity. In DIA-MS, the whole mass charge is considered during the LC time frame, allowing complete MS and MS/MS pictures of all the peaks present in the samples (Krasny and Huang, 2021) above the limit of detection.

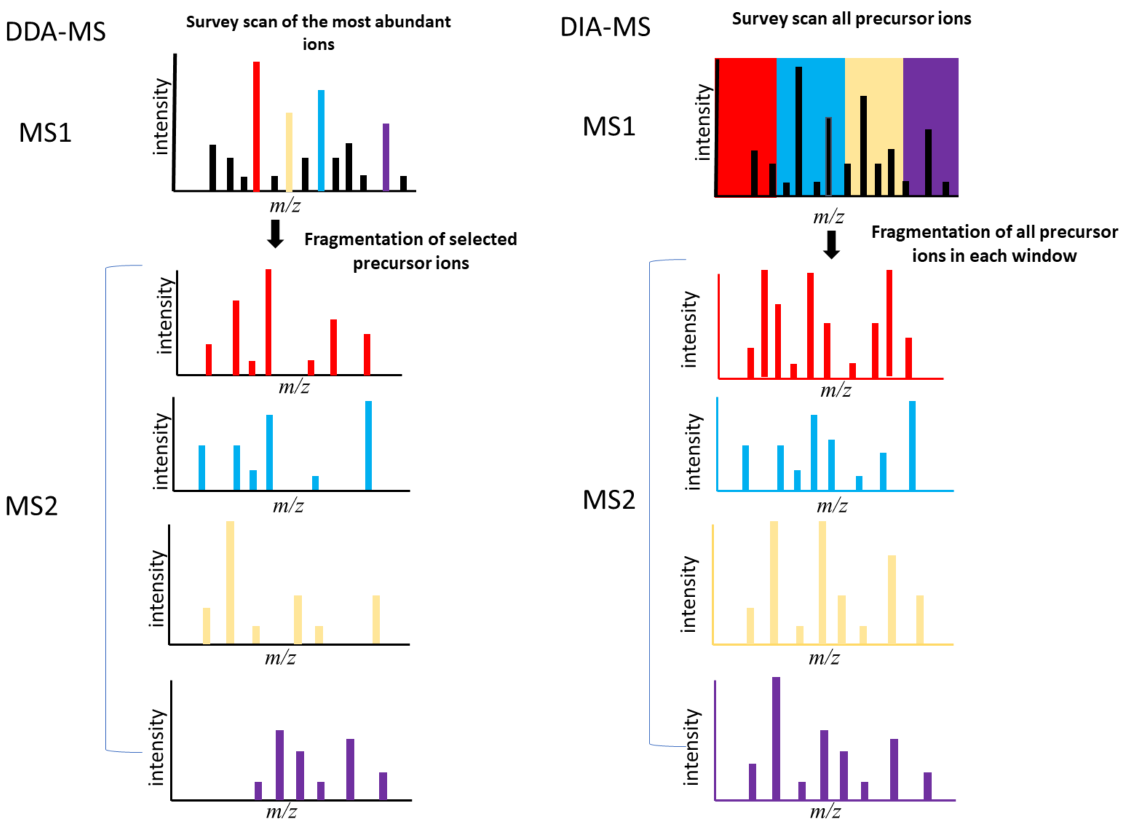


Figure 4.1 - Schematic overview of mass spectrometry acquisition methods. In DDA-MS the mass spectrometry selects precursor ions from the MS1 the most intense precursor ions and just selected peptides are fragmented for the MS analysis. In DIA-MS all the precursor ions within the m/z range window are selected for fragmentation and are analysed. Adapted from *(Krasny & Huang, 2021)* with permission from the Royal Society of Chemistry.

Aim of this chapter was to identify potential proteins in breast and prostate cancer SPAG5 deficient cells, using quantitative proteome approach based on sequential windowed acquisition of all theoretical fragment ion spectra (SWATH) mass spectrometry. This study was conducted in all 4 samples (MDA-MB-231 pLKO.1, MDA-MB-231 shRNA4, DU145 pLKO.1 and DU145 shRNA4 each in sextuplicate) comparing SPAG5 deficient cells population with the control pLKO.1 (empty vector) in MDA-MB-231 and same approach was applied for DU145 cell line. The obtained SWATH-MS data produced differentially expressed proteins (DEPs) and analysis for enrichment pathways and protein-protein interaction (PPI) was performed. In addition, SWATH-MS data were used to compare protein and transcript levels of DEGs identified in chapter 3 by RNA-sequencing. This study describes the application of the SWATH-MS technique in generation of large scale of quantitative proteomics profile in both breast and prostate cancer cell line SPAG5 deficient.

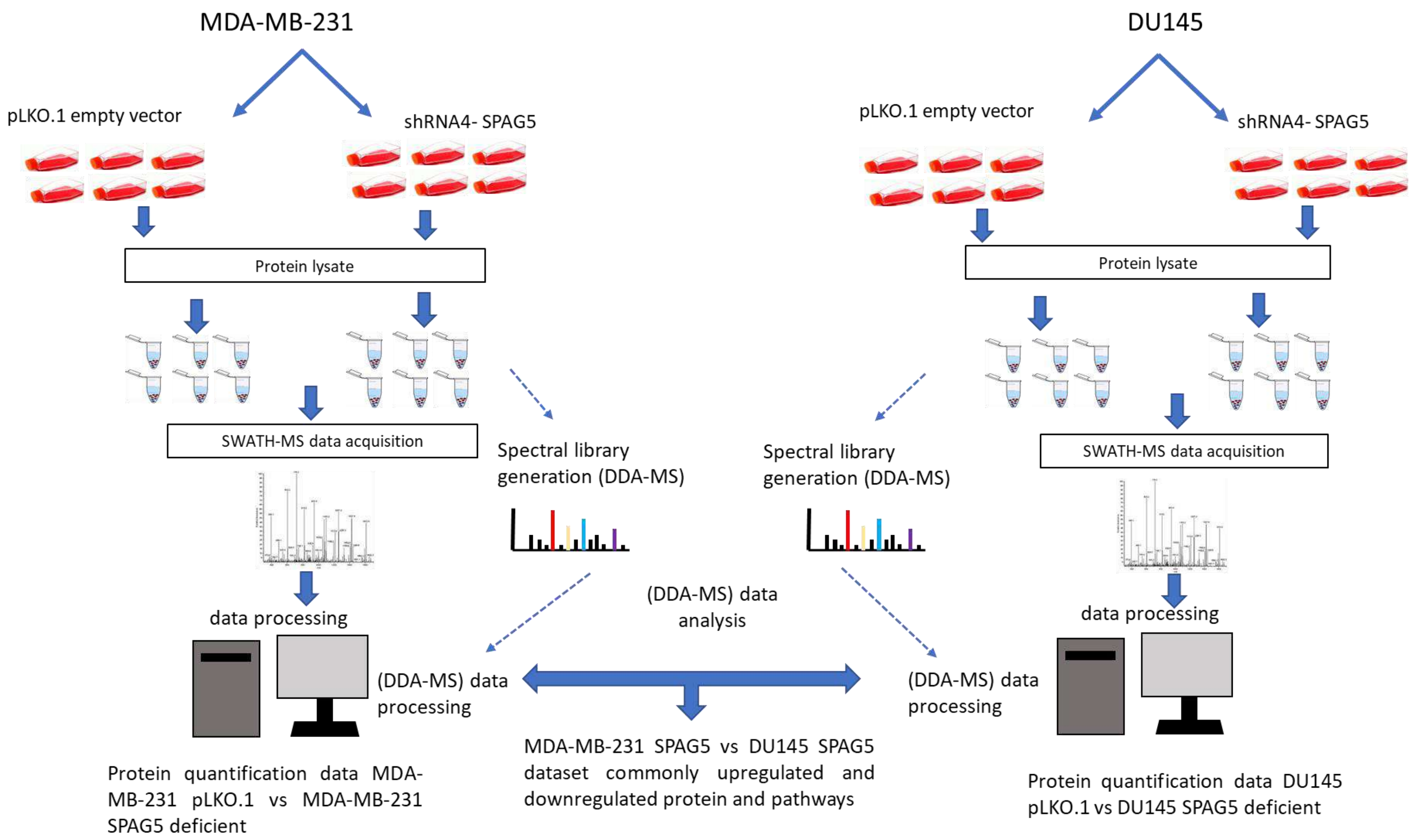


Fig. 4.2- Experimental workflow for the detection. The study is divided into protein samples preparation for processing through SMAWTH-mass spectrometry acquisition for protein quantitation in MDA-MB-231 and DU145 SPAG5 deficient cell lines.

4.2 Materials and Methods

4.2.1 Mass spectrometry on MDA-MB-231 and DU145 SPAG5 deficient cells

4.2.1.1 Whole lysate preparation

Triple-negative breast cancer MDA-MB-231 and prostate cancer DU145 cell lines were growing in six T75 flasks replicates for pLKO.1 (empty vector) and SPAG5 deficient cell shRNA4 as described in chapter 3 paragraph 3.2.1 in 1 µg/mL of puromycin antibiotic selection. The medium was removed, and cells washed three times with DPBS. Cells were detached by incubating with 0.25% w/v Trypsin-0.53 mM EDTA solution for 5-15 min at 37°C. An equal volume of cell-specific media was added to the detached cells to stop the reaction and centrifuged at 260 xg for 5 min. Cells were counted by re-suspending the cells in 1-3 mL of cell dedicated media and resuspending the cells in Trypan blue solution at 1:10 dilution. Cells were counted using hemocytometry counting the total number of cells living and excluding the number of dead cells (blue stained) from the count. Approximately $5x10^6$ cells were counted for mass spectrometry analysis. Cells were centrifuged at 260 xg for 5 min and gently washed with DPBS twice. Cell pellets were transferred in Eppendorf tubes. Whole protein lysate was obtained by adding 100 µL of PI Lysis buffer (pH 7.4, 25mM Tris, 150mM NaCl, 1mM EDTA, 1% NP40, 5% glycerol; Pierce™ ThermoFisher Scientific; #88805) supplemented with 1% v/v of protease inhibitor (Halt™ Protease and Phosphatase Inhibitor Single-Use Cocktail, EDTA-Free (100X); #78442) to prevent protein degradation. Tubes were transferred into an ice bath for sonication at the max power for 10 min then stored on ice then passed 10 times through a 29G 12.7 mm needle 0.5 mL syringe (BD Micro-Fine™+ insulin syringe; #324824). Those steps were repeated three times, and then samples were centrifuged at 14,000 xg for 15 min at 4°C. Finally, each supernatant, corresponding to the protein lysate, was carefully removed, and stored in new individual Eppendorf tubes at -80°C.

4.2.1.2 Protein extract quantification

The amount of protein was assessed using a protein assay kit (Pierce™ BCA Protein Assay Kit; #23227) according to the manufacturer's protocol. Bicinchoninic acid (BCA) protein assay working reagent (WR) was prepared by mixing 50 parts of BCA Reagent A with 1 part of Reagent B (50:1, Reagent A: B). For each sample 200 µL of WR in flat bottom 96 well is required (SARSTEDT; #83.3924). Protein standards were prepared using different bovine serum albumin (BSA) protein concentrations (20–2000 µg/mL) in distilled water. Samples were diluted 1:10 in distilled water and incubated for 30 min at 37°C covered with foil to protect from the light. Standards and samples absorbance were read at 562 nm using a plate reader (iMark™ Microplate Absorbance Reader; #1681130). The average of each absorbance value was taken from the control (blank) and subtracted from samples to remove the background signal. Protein concentration from each sample was obtained as a reference from the standard curve.

4.2.1.3 Samples preparation MDA-MB-231 and DU145 SPAG5 deficient cells for mass spectrometry

Equal amounts of protein extract (50 µg) were processed using the S-trap™ Micro spin column digestion methodology. Samples were diluted in 50 mM of tri-ethyl ammonium bicarbonate (TEAB) then reduced by adding 1 µL of 0.5M of dithiothreitol (DTT) to the protein solution in SDS and incubated for 20 min at 56°C in a shaking thermomixer. Protein solutions were cooled down at room temperature and alkylated by adding 2 µL of 0.5M of Iodoacetamide (IAA) and incubated for 15 min at room temperature in the dark. Protein lysates were incubated with 12% of aqueous phosphoric acid in dilution 1:10 and vortexed allowing the binding at this pH level. Protein lysates were incubated with 185 µL of S-Trap buffer (90% aqueous methanol + 100 mM TEAB, pH 7.1) and transferred into a 1.7 mL tube for flow through with S-Trap microcolumn. Micro columns for each sample were centrifuged at 4,000xg until all the acidified lysate/S-Trap buffer mix solution was passed through the S-Trap column. Proteins were washed three times with S-Trap buffer. Proteins solutions were digested using Digestion buffer prepared by adding 50 mM of TEAB pH 7.5 to 20 µg of trypsin vial and mixed. Proteins were digested for 1.5 h at 47°C with 25 µL of digestion buffer containing protease 1:10 wt:wt (ProteaseMAX™ Surfactant; Promega# V2072). Elution of the peptides was obtained by centrifuges samples firstly with 50 mL of TEAB and then with 0.2% of aqueous formic acid at 4,000xg. Hydrophobic peptides recovery was achieved with elution of 35 µL 50% acetonitrile containing 0.2 % formic acid. Samples were dried through a vacuum concentrator (Concentrator plus/Vacufuge® plus – Eppendorf) at 60°C for 1h and stored at -20°C before resuspension in 5% acetonitrile and 0.1% of formic acid for subsequent analysis.

4.2.1.4 Processing mass spectrometry generated data

DU145 and MDA-MB-231 SPAG5 knockdown samples were analysed on a SCIEX TripleTOF® 6600 mass spectrometer linked to an Eksigent nanoLC 425 HPLC system. The LC system was operating in microflow (5 µl/min) and 3 µl of each sample was directly injected on a YMC 25 cm x 0.3 mm Triart-C18 column (12 nm, 3 µm particle size). Chromatographic separation was achieved over a 60-minute time frame for sequential window acquisition of all theoretical fragment ion spectra (SWATH) analysis and 87 min for Information Dependent Acquisition (IDA). The chromatography separation consisted of the following mobile phase gradients; 2% v/v mobile phase B (2% v/v acetonitrile, 5% v/v DMSO in 0.1% v/v FA) to 40% v/v over 50 min; to 80% v/v B at 55 min, held for 2 min, then returned to 2% v/v over 1 min. MS analysis was performed using two acquisition methods, IDA for spectral library generation and MS-SWATH data acquisition. The library data for both stable cell lines were generated through DIANN- SWATH (https://github.com/vdemichev/DiaNN/releases/tag/1.8.1) using FASTA format and adding the sequence database for each protein from Swiss-Prot (June 2021) database containing human species at a 1% False Discovery Rate (FDR) cut-off.

4.2.1.5 Quantitative analysis of mass spectrometry data

Detection of the differential expressed protein was performed using different statistical methods with the *StatsPro* tool (https://www.omicsolution.com/wukong/StatsPro). For this analysis proteins with a P-value of 0.05 and a Log2 fold change (FC) absolute 0.58 were considered differentially expressed proteins (DEPs). Subsequently, the DEPs were visualised using Morpheus online software for heatmap generation (https://software.broadinstitute.org/morpheus/). Using the statistical *t*test option proteins were sorted for the most upregulated and downregulated in both MDA-MB-231 and DU145 SPAG5 deficient cell line. Using an R-based Shiny application, called ggVolcanoR (https://ggvolcanor.erc.monash.edu/), DEPs were visualised in a volcano plot and customized for the 30 most significantly upregulated and downregulated proteins. Both *StatsPro* and ggVolcanoR are provided with the option to download the customised list of filtered dysregulation expression data, that were been used for downstream pathway analysis.

4.2.1.6 Functional enrichment analysis for pathway generation and protein-protein interaction (PPI) reconstruction

The results obtained from mass spectrometry were used to perform pathway enrichment analysis and gene network reconstruction. Data were analysed separately for MDA-MB-231 SPAG5 knockdown and DU145 SPAG5 knockdown using the online software Metascape tool (https://metascape.org) (Zhou, Y. et al., 2019) with default parameters. Once obtained proteins data set from *StatsPro* analysis, pathways and enrichment analysed were performed by selecting genomic sources: KEGG pathway, GO Biological Process, GO Cellular components, and GO Molecules function. Data for input species *Homo sapiens* were analysed considering p-value cut off < 0.01, minimum count of 3, and enrichment factor of 1.5 defined as the ratio between the observed count and the count expected by chance. Data were collected and grouped into clusters based on their membership similarities. Particularly, p-values and enrichment factors were calculated based on accumulative hypergeometric distribution, while q value with the Benjamin-Hochberg procedure was generated to correct for multiple testing. The remaining significant terms were hierarchically clustered into a tree based on Kappa-statistical similarities among their gene membership. Finally, the 0.3 Kappa score was applied as the threshold to cast the tree into the term cluster. Terms with the best p-value from each of the 20 clusters were considered. Cytoscape software was performed for Protein-protein interaction (PPI) to identify potential interactions of the selected proteins based on their gene IDs. PPI analysis for MDA-MB-231 SPAG5 deficient cells, for active interaction source the minimum required confidence used was (0.400) . However, to make the analysis more stringent the level of confidence was increased at the highest (0.700) threshold setting.

4.2.1.7 Global proteome analysis of MDA-MB-231 and DU145 SPAG5 deficient cells

To investigate whether MDA-MB-231 and DU145 SPAG5 deficient cells showed common genes a Venn diagram was generated using the Venny tool containing shared or specific proteins database. The common proteins dataset processed from MS were submitted to *StatsPro* for statistical analysis *t*test. Data were provided based on the padj<0.05 and Log2 FC 0.58.

4.2.1.8 Cross-over gene and proteome data of MDA-MB-231 and DU145 SPAG5 deficient

Cross-over gene and proteome data were combined to identify potential common markers. A Venn diagram was generated with Venny tool. DEGs identified in chapter 3 for MDA-MB-231 and DU145 SPAG5 were combined and sorted for the most significant based on the padj<0.05.

4.3 Results

4.3.1 Identification of differentially expressed proteins in MDA-MB-231 and DU145 SPAG5 deficient cells

To investigate the effect of SPAG5 silencing in both prostate cancer and breast cancer cell lines, quantitative proteomic profiling was assessed by mass spectrometry (MS) analysis from the same cell population used for RNA sequencing technology. Mass spectrometry analysis was performed to identify potential proteins whose endogenous expression could be altered from SPAG5 silencing. Therefore, the experiment was performed on both the control pLKO.1 (empty vector) and silenced cells line with shRNA4 knockdown lentivirus in both DU145 and MDA-MB-231 cell lines. Sample preparation was conducted as described in section 3.3.3 and analysed as explained in section 3.3.4 of this chapter. Data were analysed separately by applying 1% of the false discovery rate (FDR). From the analysis conducted, 5007 proteins from breast cancer cell line MDA-MB-231 and 5563 proteins for prostate cancer cell line DU145 were identified and quantified. Information on protein identities was obtained using UniProt Knowledgebase (UniProtKB). Using *StatsPro* tool 230 proteins for breast cancer (Table 4.1) and 65 proteins for prostate cancer, significantly modulated were identified (Table 4.2).

Figure 4.3 shows the heatmap created using the online software “Morpheus” representing 126 upregulated proteins and 104 proteins downregulated in MDA-MB-231 SPAG5 deficient cells versus the control (empty vector), and 22 upregulated protein and 43 downregulated proteins for DU145 in SPAG5 deficient cells versus the control (empty vector) (Fig. 4.4 A). The proteins significantly upregulated and downregulated in shRNA4 are indicated with a coloured square red for upregulated and blue square for downregulated proteins in the heatmap.

The volcano plot (Fig. 4.2 **B**) (Fig. 4.3 **B**) showed the most statistically significant upregulated and downregulated proteins, generated by ggVolcanoR after processing with *StatsPro* online software sorted for Log2FC and padj value < 0.05 in MD-MB-231 and DU145 SPAG5 deficient cells.

Table 4.1 - List of proteins differentially expressed in whole cell lysate from MDA-MB-231 SPAG5 deficient vs MDA-MB-231 control (empty vector pLKO.1) cell population. The tables present 44 protein (22 in each downregulated and upregulated gene) form the list of 230 proteins out of 5007 identified, using an absolute Log2 FC ≥ 0.58. Proteins listed in blue are downregulated in knockdown vs control cell populations, whereas in the orange table are proteins upregulated in knockdown vs control cell populations. Full proteins list is shown in the appendix of this chapter.

Protein ID	log2FC	Pvalue	Protein names
MYL6	0.8056733	0.000045	myosin light chain 6
FASN	0.7727183	0.000045	fatty acid synthase
CALB2	0.673465	0.000294174	calbindin 2
NDUFA4	0.8102433	0.000451162	NDUFA4, mitochondrial complex associated
ACACA	0.8507667	0.000551243	acetyl-CoA carboxylase alpha
CKAP4	0.735005	0.000557335	cytoskeleton associated protein 4
THBS1	1.38988	0.000921797	thrombospondin 1
PGLS	0.706205	0.000921797	6-phosphogluconolactonase
GSTK1	0.6611033	0.000921797	glutathione S-transferase kappa 1
AHNAK2	1.7353433	0.001132068	AHNAK nucleoprotein 2
LSS	0.7825617	0.001135313	lanosterol synthase
SNX9	0.8484367	0.001268741	sorting nexin 9
CASK	0.8451483	0.001268741	calcium/calmodulin dependent serine protein kinase
DHCR7	0.7306267	0.001268741	7-dehydrocholesterol reductase
CPOX	1.1915783	0.001351428	coproporphyrinogen oxidase
MYO18A	0.6174167	0.001374104	myosin XVIIIA
NIPSNAP1	1.442075	0.001377712	nipsnap homolog 1
SCRIB	0.8392683	0.001514356	scribbled planar cell polarity protein
TACSTD2	1.318375	0.001565657	tumor associated calcium signal transducer 2
MRRF	0.9911617	0.001574805	mitochondrial ribosome recycling factor
PTTG1IP	1.1009183	0.001749585	PTTG1 interacting protein

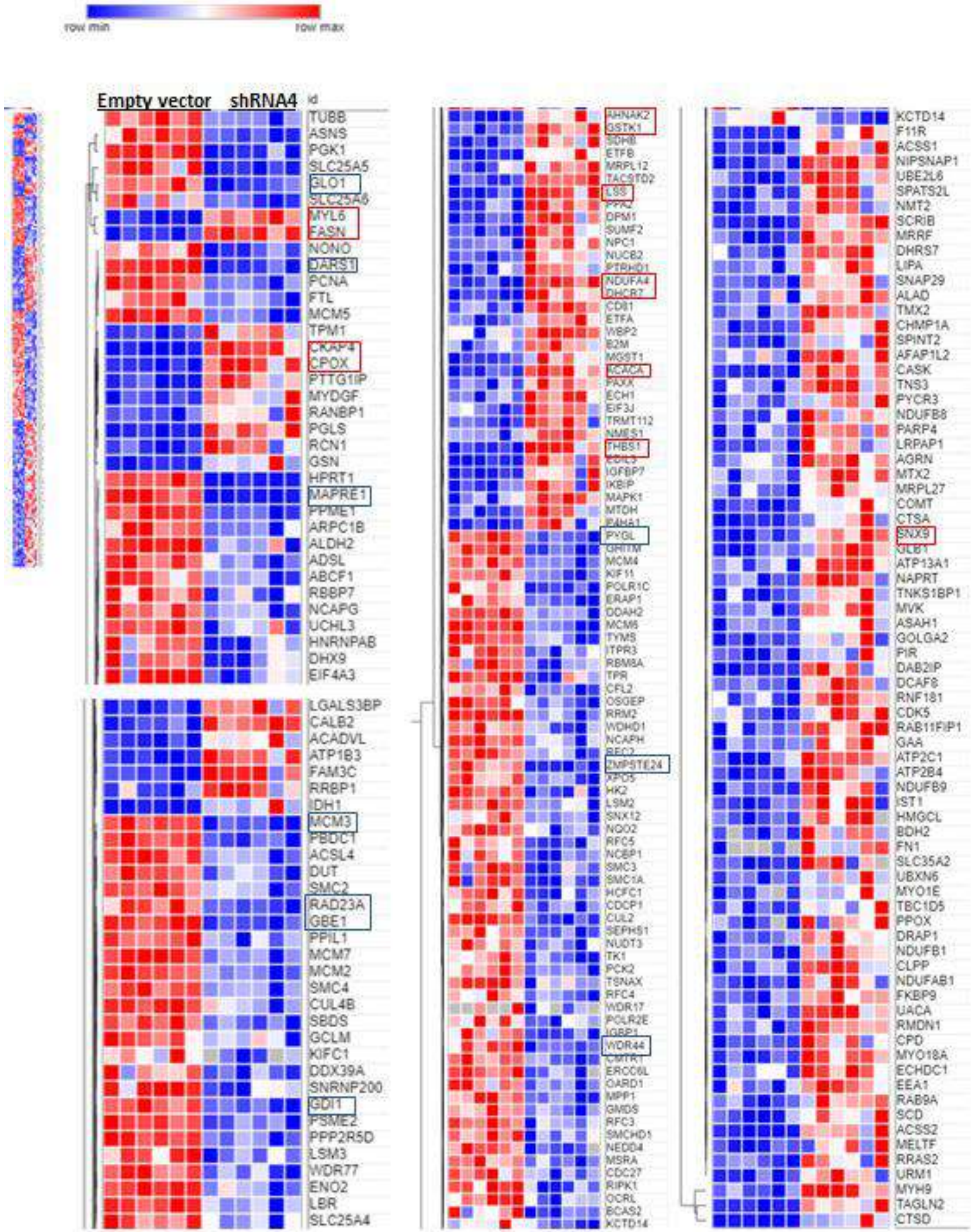
Entry name	log2FC	Pvalue	Protein names
HPRT1	-1.252545	5.92E-08	hypoxanthine phosphoribosyltransferase 1
MAPRE1	-1.3980017	1.68E-06	microtubule associated protein RP/EB family member 1
ZMPSTE24	-0.9384217	0.000106	zinc metallopeptidase STE24
GHITM	-0.8055733	0.000279	growth hormone inducible transmembrane protein
GBE1	-0.9666867	0.00029	1,4-alpha-glucan branching enzyme 1
GLO1	-0.6487933	0.000352	glyoxalase I
DARS1	-0.59797	0.000449	Aspartyl-TRNA Synthetase 1
RAD23A	-0.99648	0.000556	RAD23 homolog A, nucleotide excision repair protein
ASNS	-1.0340183	0.000557	asparagine synthetase domain containing 1
PYGL	-0.7982167	0.000893	glycogen phosphorylase L
CUL2	-0.9719667	0.00092	cullin 2
MCM3	-1.013495	0.001071	minichromosome maintenance complex component 3 associated protein
GDI1	-0.7964817	0.001103	GDP dissociation inhibitor 1
WDR44	-0.7441833	0.001269	WD repeat domain 44
PCNA	-0.67534	0.001378	proliferating cell nuclear antigen
POLR1C	-0.72079	0.001514	RNA polymerase I and III subunit C
PBDC1	-1.04886	0.001575	polysaccharide biosynthesis domain containing 1
SEPHS1	-0.791175	0.001575	selenophosphate synthetase 1
SBDS	-0.603655	0.001575	SBDS, ribosome maturation factor
KIF11	-0.623325	0.00166	kinesin family member 11
NONO	-0.8818133	0.001776	non-POU domain containing octamer binding
TUBB	-0.5979	0.00198	tubulin beta class I

Table 4.2 - List of proteins differentially expressed in whole cell lysate from DU145 SPAG5 deficient vs DU145 control (empty vector pLKO.1) cell population. The tables present 44 protein (22 in each downregulated and upregulated gene) form the list of 65 proteins out of 5563 identified, obtained from the absolute Log2FC ≥ 0.58. Proteins listed in blue are downregulated in knockdown vs control cell populations, whereas the orange table are protein upregulated in knockdown vs control cell populations. Full proteins list is shown in the appendix of this chapter.

Protein ID	log2FC	Pvalue	Protein names
PLOD2	0.9102233	0.0000018	procollagen-lysine,2-oxoglutarate 5-dioxygenase 2
TAGLN	1.9978000	0.0000281	transgelin 2
ASS1	0.9117067	0.0002215	argininosuccinate synthase 1
SLC7A5	0.6795283	0.0002215	solute carrier family 7 member 5
P4HA2	0.6859083	0.0010722	prolyl 4-hydroxylase subunit alpha 2
SPINT1	0.7130467	0.0014671	serine peptidase inhibitor, Kunitz type 1
EHHADH	0.7213000	0.0020397	enoyl-CoA hydratase and 3-hydroxyacyl CoA dehydrogenase
IL18	0.6753500	0.0021343	interleukin 18
CNTN1	0.8171367	0.0024670	contactin 1
CDH1	0.6199767	0.0024670	cadherin 1
CPT1A	0.7492967	0.0028184	carnitine palmitoyltransferase 1A
LDHAL6A	0.8232723	0.0064972	lactate dehydrogenase A like 6A
EDIL3	0.9038333	0.0099024	EGF like repeats and discoidin domains 3
CNN2	0.7463550	0.0137128	calponin 2
DSP	0.8526917	0.0140276	desmoplakin
CBR3	0.6576333	0.0146019	carbonyl reductase 3
TUBB3	0.6077567	0.0149972	tubulin beta 3 class III
ALDH3B1	0.8438483	0.0220629	aldehyde dehydrogenase 3 family member B1
POMP	0.8479359	0.0255891	proteasome maturation protein
FKBP9	0.5843483	0.0394885	FK506 binding protein 9
CAAP1	0.9661603	0.0409898	caspase activity and apoptosis inhibitor 1
FNDC3B	0.5897333	0.0469626	fibronectin type III domain containing 3B

protein ID	log2FC	Pvalue	Protein names
CCT2	-0.6710517	0.00000350	chaperonin containing TCP1 subunit 2
HK1	-0.9666550	0.0000136	hexokinase 1
NAMPT	-0.6151333	0.0000136	nicotinamide phosphoribosyltransferase
FLNC	-1.3137983	0.0000344	filamin C
ATL3	-0.8932283	0.0000846	atlastin GTPase 3
AKR1C2	-0.8908967	0.0000846	aldo-keto reductase family 1 member C2
MAPRE1	-0.7842400	0.0001412	microtubule associated protein RP/EB family member 1
SCAMP1	-2.0134300	0.0001433	secretory carrier membrane protein 1
BCL2L13	-1.5891683	0.0001446	BCL2 like 13
SLC16A3	-0.7439383	0.0001650	solute carrier family 16 member 3
CUL2	-1.5723500	0.0003184	cullin 2
SLC38A2	-1.1839450	0.0003301	solute carrier family 38 member 2
GBE1	-0.6666900	0.0004077	1,4-alpha-glucan branching enzyme 1
DBNL	-0.6274200	0.0004871	drebrin like
PADI2	-1.7476717	0.0006640	peptidyl arginine deiminase 2
NEDD4	-1.2189133	0.0009678	NEDD4 E3 ubiquitin protein ligase
NMD3	-0.6576550	0.0022002	NMD3 ribosome export adaptor
GLUL	-0.9610100	0.0027392	glutamate-ammonia ligase
NT5E	-1.3649800	0.0028184	5'-nucleotidase ecto
SPAG5	-0.8722488	0.0036863	sperm associated antigen 5
CAPG	-1.3539633	0.0052776	capping actin protein, gelsolin like
ASF1B	-0.5925217	0.0059416	anti-silencing function 1B histone chaperone

A



B

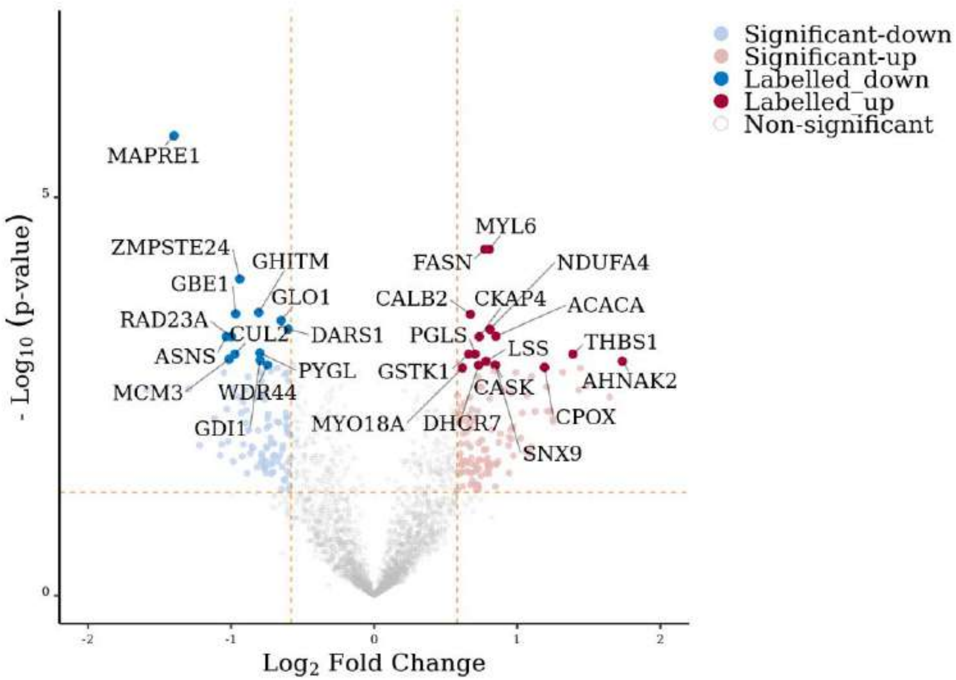


Figure 4.3 - Heatmap and volcano plot of the top 230 significantly modulated proteins in MDA-MB-231 SPAG5 deficient vs pLKO.1. (A) Heatmap shows 126 proteins upregulated and 104 proteins downregulated in SPAG5 knockdown (shRNA4) vs. control (Empty vector) MDA-MB-231 cell population after normalisation. Box in read represents the proteins showed in volcano plot. **(B)** Volcano plot showing the fold change (Log2FC) and padj value of significantly upregulated (red dots) and downregulated (blue dots) protein abundance in triple-negative MDA-MB-231 shRNA4-SPAG5 using ggVolcanoR. Protein change of no significance is showed in grey dots. Data were obtained following protein quantitation on the whole lysate and processed using StatsPro web tool using ttest methods (padj <0.05 and log2FC absolute ≥ 0.58; n=6)

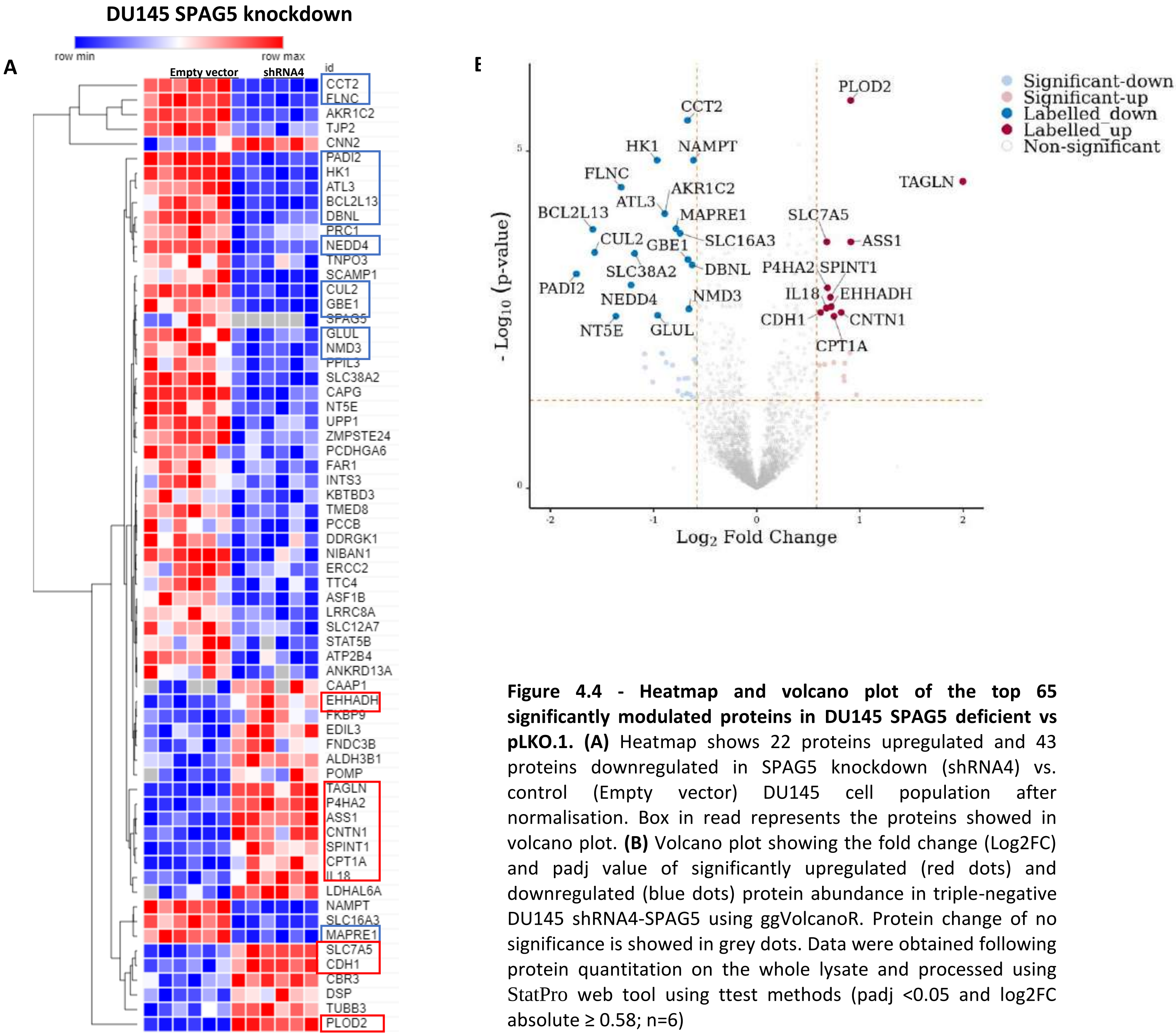


Figure 4.4 - Heatmap and volcano plot of the top 65 significantly modulated proteins in DU145 SPAG5 deficient vs pLKO.1. (A) Heatmap shows 22 proteins upregulated and 43 proteins downregulated in SPAG5 knockdown (shRNA4) vs. control (Empty vector) DU145 cell population after normalisation. Box in read represents the proteins showed in volcano plot. **(B)** Volcano plot showing the fold change (Log2FC) and padj value of significantly upregulated (red dots) and downregulated (blue dots) protein abundance in triple-negative DU145 shRNA4-SPAG5 using ggVolcanoR. Protein change of no significance is showed in grey dots. Data were obtained following protein quantitation on the whole lysate and processed using StatPro web tool using ttest methods (padj <0.05 and log2FC absolute ≥ 0.58; n=6)

The significant DEPs obtained from mass spectrometry were later used to perform pathway enrichment analysis and gene network reconstruction. Data were analysed separately for MDA-MB-231 SPAG5 knockdown (Fig.4.5) using the online software METASCAPE.

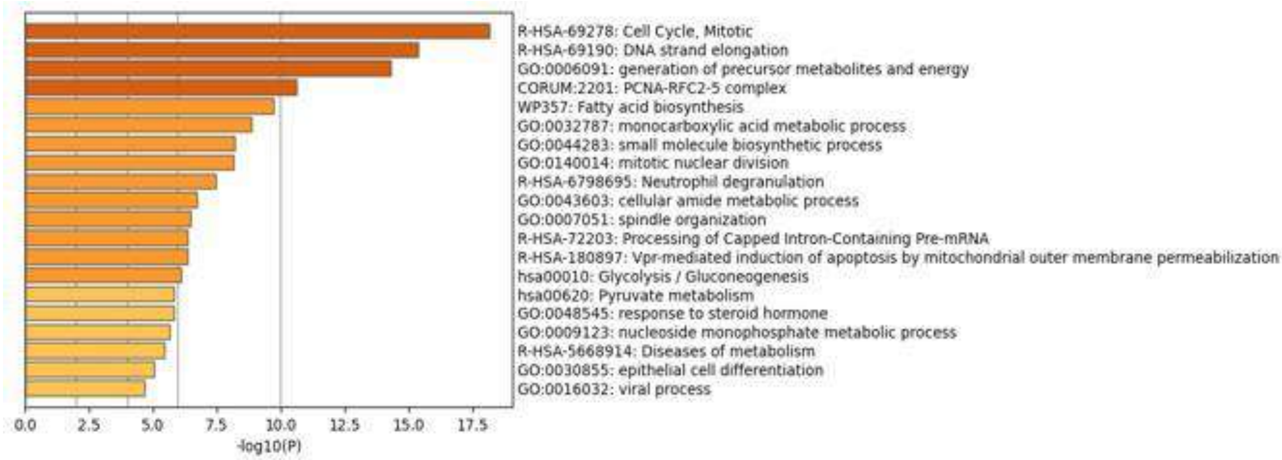


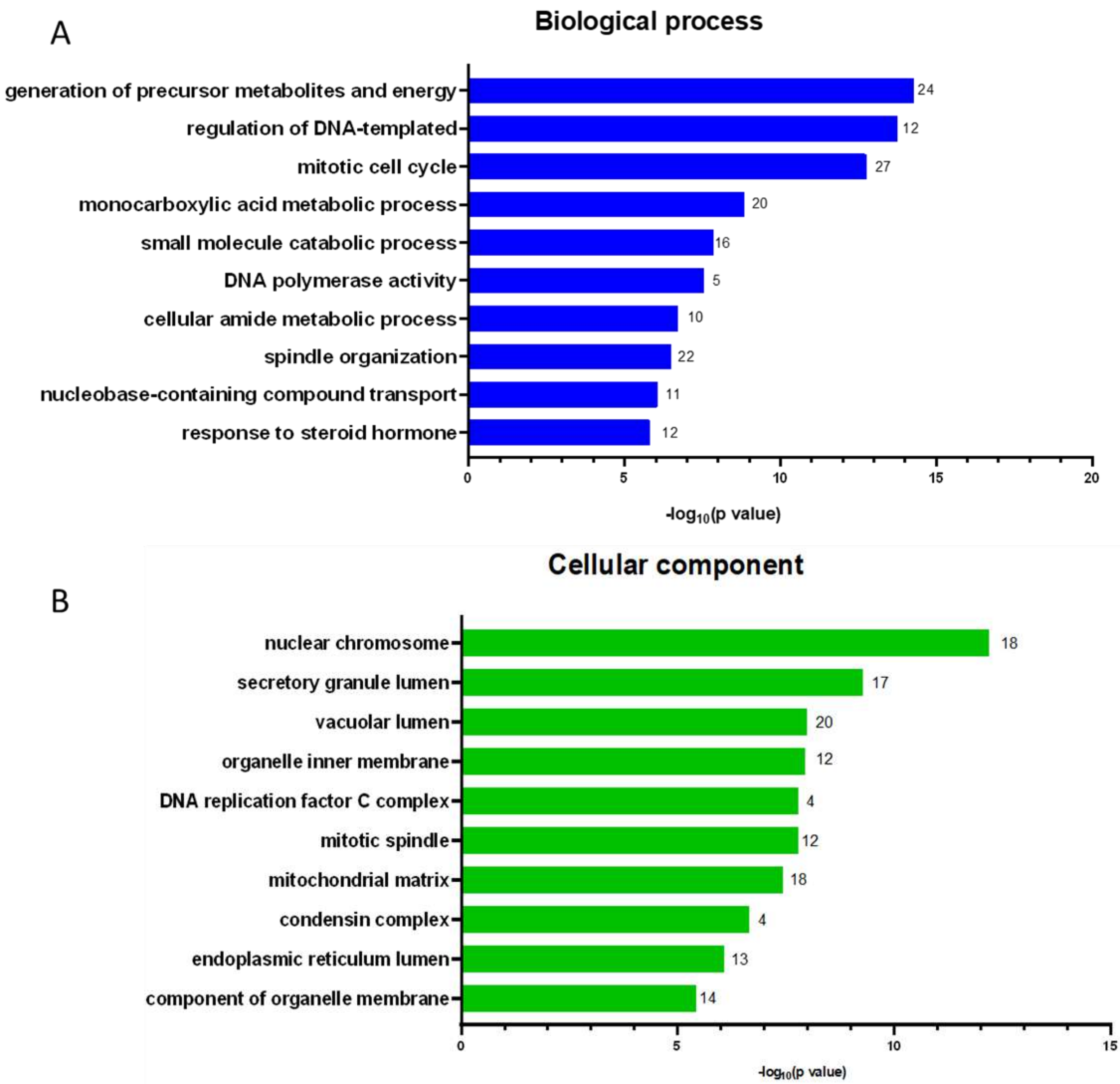
Figure 4.5 - Bar graph of the enriched terms from DEPs in MDA-MB-231 SPAG5 deficient cells. Bar graph showing the enrich term across the input gene list coloured by p-value <0.01, via Metascape software in MDA-MB-231 SPAG5 knockdown.

The genes associated with each pathway are described in table 4.3 for MDA-MB-231 and in table 4.4 for DU145. Each gene is differentially coloured whether upregulated (red) or downregulated (blue). Among the most significantly 4 pathways ‘DNA strand elongation’, ‘Processing of Capped Intron-Containing Pre-mRNA’, ‘Vpr-mediated induction of apoptosis by mitochondrial outer membrane permeabilization and ‘nucleoside monophosphate metabolic process’ are affected by SPAG5 deficient cells population.

Table 4.3 - Tables show the list of proteins differentially expressed in different pathways from MDA-MB-231 SPAG5 deficient cells. The table shows the list of the pathways sorted according to the p-value <0.01. The 'entities found' represents the number of genes from the list provided with membership with the specific term (pathway) and specifically defined as upregulated (red) and downregulated (blue).

Pathway name	Protein ID upregulated	Protein ID downregulated	Entities found	p-value
Cell Cycle, Mitotic	GOLGA2, MAPK1, IST1	CDC27, LBR, MCM2, MCM3, MCM4, MCM5, MCM6, MCM7, PCNA, PPP2R5D, PSME2, RFC2, RFC3, RFC4, RFC5, RRM2, TK1, TPR, TYMS, SMC1A, SMC3, SMC4, SMC2, MAPRE1, NCAPH, PPME1, ERCC6L, NCAPG, TUBB	32/561	7.27E-19
DNA strand elongation		MCM2, MCM3, MCM4, MCM5, MCM6, MCM7, PCNA, RFC2, RFC3, RFC4, RFC5	11/32	4.20E-16
generation of precursor metabolites and energy	ACADVL, ETFA, ETFB, GAA, HMGCL, IDH1, NDUFA4, NDUFAB1, NDUFB1, NDUFB8, NDUFB9, SDHB, PGLS, ACSS2, ACSS1	ADSL, ALDH2, SLC25A4, ENO2, GBE1, HK2, NQO2, PGK1, PYGL,	24/394	5.15E-15
ribose phosphate metabolic process	MVK, NDUFAB1, NDUFB1, NDUFB8, NDUFB9, SDHB, CASK, NUDT3, ACSS2, ACSS1, ACACA	ADSL, ENO2, ACSL4, HK2, HPRT1, MPP1, PGK1, PYGL,	19/382	1.07E-07
PCNA-CHL12-RFC2-5 complex		PCNA, RFC2, RFC3, RFC4, RFC5	5/6	2.40E-11
Fatty acid biosynthesis	ACACA, ECH1 FASN, SCD, ECHDC1, ACSS2	ACSL4	7/22	2.04E-10
monocarboxylic acid metabolic process	ACACA, ACADVL, ASAH1, ECH1 ETFA, ETFB, FASN, IDH1, NDUFAB1, NPC1, SCD, ECHDC1, ACSS2, BDH2, ACSS1	ENO2, ACSL4, HK2, PCK2, PGK1	20/492	1.41E-09
small molecule biosynthetic process	ACACA, ASAH1, DHCR7, FASN, HMGCL, LSS, MVK, NDUFAB1, SCD, PYCR3, ACSS1	ASNS, ENO2, HPRT1, LBR, PCK2, PGK1, SEPHS1	18/430	6.16E-09
mitotic nuclear division	GOLGA2, RMDN1, CHMP1A,	KIF11, KIFC1, SMC1A, SMC3, SMC4, SMC2, MAPRE1, NCAPH, NCAPG	12/169	7.23E-09
Neutrophil degranulation	ALAD, ASAH1, B2M, CTSD, GAA, GLB1, GSN, IDH1, MGST1, CTSA, MAPK1, SNAP29, IST1, CKAP4, NAPRT	FTL, PYGL, TUBB	18/482	3.53E-08
cellular amide metabolic process	ACACA, ASAH1, CPD, IDH1, MVK, MRPL12, RRBP1, EIF3J, MRPL27, ACSS2, BDH2, ACSS1, MRRF, GSTK1	ABCF1, ASNS, DARS1, ACSL4, GCLM, GLO1, NCBP1, ERAP1,	22/792	1.93E-07
spindle organization	GOLGA2, MYH9, RMDN1	KIF11, KIFC1, SMC1A, SMC3, MAPRE1, SBDS, TUBB	10/155	3.25E-07
Processing of Capped Intron-Containing Pre-mRNA		DHX9, NCBP1, POLR2E, TPR, EIF4A3, RBM8A, DDX39A, BCAS2, SNRNP200, LSM3, PPIL1, LSM2	12/245	4.27E-07
Vpr-mediated induction of apoptosis by mitochondrial outer membrane permeabilization		SLC25A4, SLC25A5, SLC25A6	3/3	4.30E-07
Glycolysis / Gluconeogenesis	ACSS2, ACSS1	ALDH2, ENO2, HK2, PCK2, PGK1	7/67	7.77E-07
Pyruvate metabolism	ACSS2, ACSS1	ACACA, ALDH2, GLO1, PCK2	6/47	1.47E-06
response to steroid hormone	ALAD, GLB1, IDH1, IGFBP7, NPC1, THBS1, WBP2	NEDD4, PCK2, PCNA, RBBP7, TYMS	12/276	1.50E-06
nucleoside monophosphate metabolic process		ADSL, DUT, HPRT1, MPP1, TK1, TYMS, CASK	7/78	2.20E-06
Diseases of metabolism	ACACA, GAA, GLB1, IDH1, CTSA, THBS1, DPM1, AGRN	GBE1, GCLM, HPRT1	11/249	4.83E-04
epithelial cell differentiation	ACADVL, ASAH1, FASN, MYO1E, TAGLN2, SPINT2, SCRIB, F11R, BDH2, GSTK1	PCK2, PCNA, PGK1, TYMS, WDR77, TUBB	16/576	9.52E-06
Viral Process	CD81, NPC1, CHMP1A, EEA1, NMT2, IST1, F11R	DHX9, HCFC1, NEDD4	10/248	2.17E-05

The list of the most upregulated and downregulated genes was also sorted for GO enrichment analysis. Particularly the analysis was performed for Biological Process (BP), Molecular Function (MF) and Cellular Components (CC) in MDA-MB-231 SPAG5 knockdown. Tables with all gene names in the specific pathways are shown in the Appendix for this chapter (Chapter 4 Table 4.6).



C

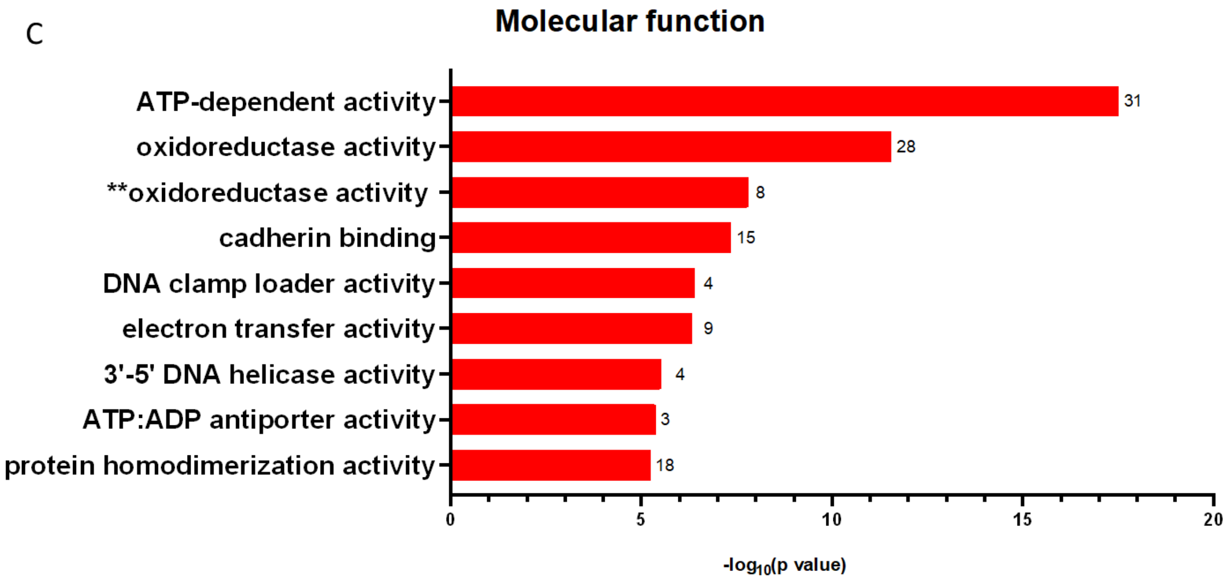


Figure 4.6 - Enrichment analysis for the biological function of the 230 genes obtained from mass spectrometry analysis through Metascape online tool. Analysis was performed on Biological Process (BP) (A), Molecular Function (MF) (B), and Cellular Component (CC) (C) in MDA-MB-231 SPAG5 knockdown. Pathways are shown in descending order based on -log_{10} P-value. The number of genes associated is shown above each bar. Only pathways with a p-value <0.01 are shown. ** oxidoreductase activity acting on the CH-CH group of donors.

The functional enrichment analysis was performed also from the MS proteins list. For DU145 SPAG5 knockdown, 65 genes were found with 22 upregulated and 43 downregulated when compared with the control (Empty vector pLKO.1). Protein list was processed using Metascape online tool and a bar graph was generated based on the pathways with statistically significance across the input gene list provided as shown in figure 4.7.

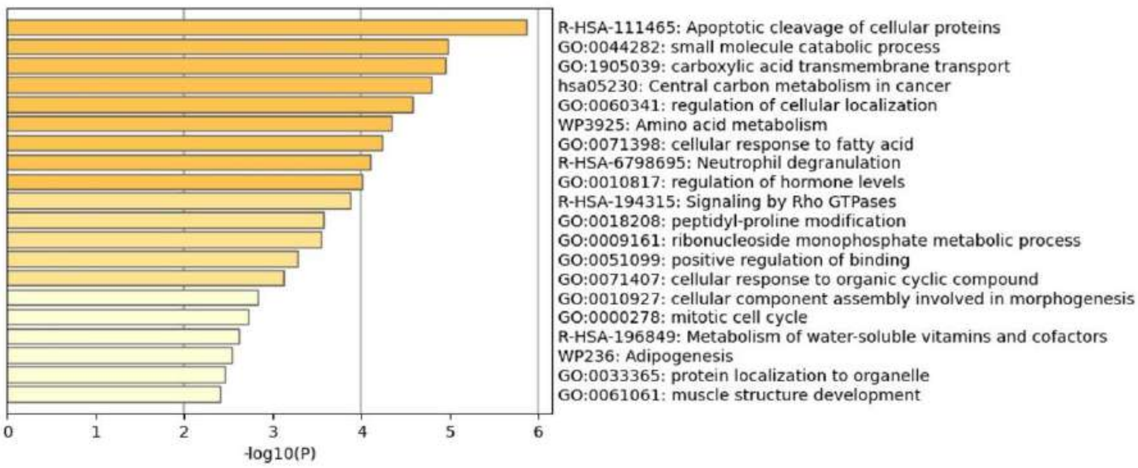


Figure 4.7 - Bar graph of the enriched terms from DEPs in DU145 SPAG5 deficient cells. Bar graph showing the enrich term across the input gene list coloured by p-value <0.01, via Metascape software in DU145 SPAG5 knockdown.

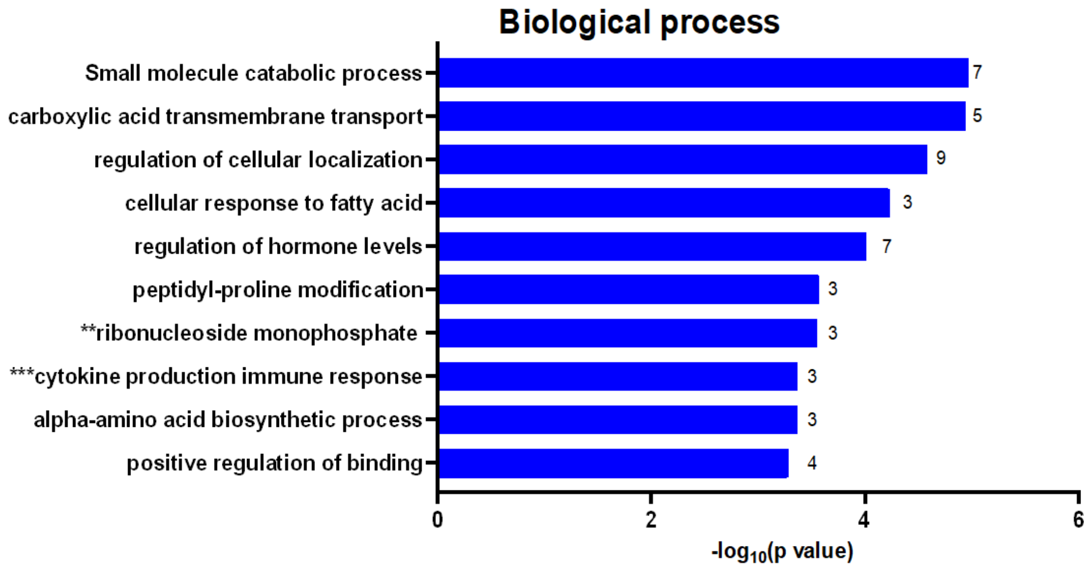
Proteins associated with specific pathways is coloured for upregulated (red) and downregulated (blue) specifically. Particularly, among the most significant pathways for prostate cancer 'ribonucleoside monophosphate metabolic process', 'Metabolism of water-soluble vitamins and cofactors' and 'Adipogenesis' are shown to be affected by the SPAG5 deficient cell population.

Table 4.4 - Tables show the list of proteins differentially expressed in different pathways from DU145 SPAG5 deficient cells . The table shows the list of the pathways sorted according to the p-value <0.01. The 'entities found' represents the number of genes from the list provided with membership with the specific term (pathway) and specifically defined as upregulated (red) and downregulated (blue).

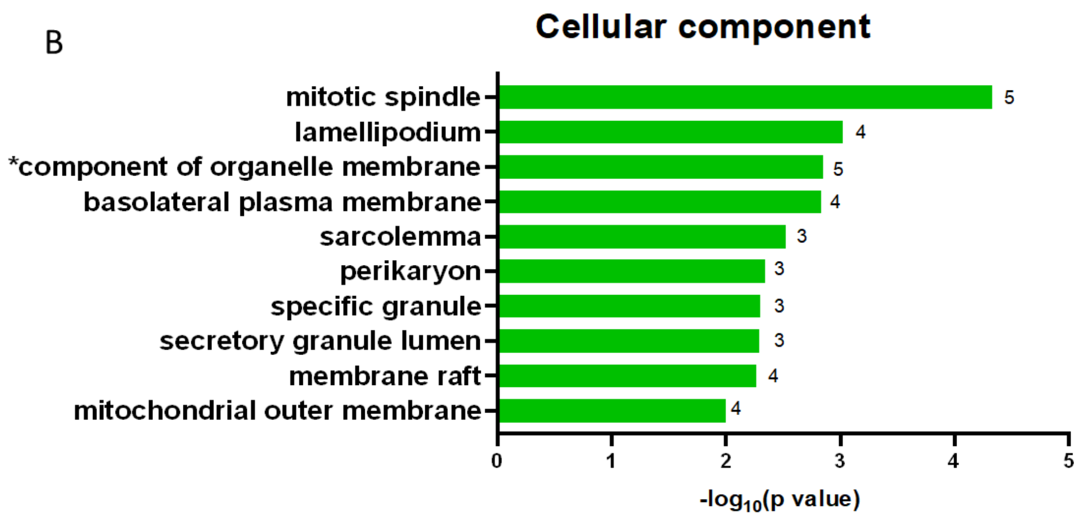
Pathway name	Protein ID upregulated	Protein ID downregulated	Entities found	p-value
Apoptotic cleavage of cellular proteins	CDH1, DSP,	TJP2, DBNL	4/38	1.37E-06
small molecule catabolic process	ALDH3B1, CPT1A, EHHADH	GLUL, HK1, PCCB, UPP1	7/352	1.07E-05
carboxylic acid transmembrane transport	CPT1A, SLC7A5	SLC16A3, SLC38A2, LRRC8A	5/136	1.14E-05
Central carbon metabolism in cancer	SLC7A5, LDHAL6A	HK1, SLC16A3	4/70	1.61E-05
regulation of cytokine production involved in immune response	IL18, SLC7A5	HK1	3/98	1.25E-03
Amino acid metabolism	ASS1, EHHADH, P4HA2	GLUL	4/91	4.54E-05
Neutrophil degranulation	ALDH3B1, CNN2, DSP	SCAMP1, CCT2, PADI2, DBNL	7/482	7.91E-05
regulation of hormone levels	CPT1A, SLC7A5	AKR1C2, GLUL, STAT5B, ZMPSTE24, LRRC8A	7/498	9.70E-05
Signalling by Rho GTPases	CDH1, DSP, TUBB3, DDRGK1	PRC1, TJP2, CCT2, MAPRE1,	8/707	1.35E-04
peptidyl-proline modification	P4HA2, FKBP9	PPIL3	3/58	2.70E-04
ribonucleoside monophosphate metabolic process		NT5E, UPP1, TJP2	3/59	2.84E-04
positive regulation of binding	DDRGK1	ERCC2, MAPRE1, NMD3,	4/172	5.25E-04
cellular response to an organic cyclic compound	ASS1, CDH1, IL18	ATP2B4, NEDD4, PADI2	6/505	7.61E-04
cellular component assembly involved in morphogenesis	CNTN1	ERCC2, FLNC	3/104	1.48E-03
mitotic cell cycle	TUBB3	CUL2, PRC1, SPAG5, MAPRE1, INTS3	6/605	1.91E-03
Metabolism of water-soluble vitamins and cofactors		NT5E, PCCB, NAMPT	3/124	2.45E-03
Adipogenesis		STAT5B, NAMPT, ZMPSTE24	3/131	2.86E-03
protein localization to organelle	DDRGK1	HK1, NEDD4, SPAG5, MAPRE1, TNPO3	6/684	3.52E-03
muscle structure development	ASS1, DSP, TAGLN	FLNC, ZMPSTE24	5/484	3.89E-03

The list of the most upregulated and downregulated genes was also sorted for GO enrichment analysis for DU145 SPAG5 knockdown. Particularly the analysis was performed for Biological Process (BP)(**A**), Molecular Function (MF)(**B**) and Cellular Components (CC)(**C**) in DU145 SPAG5 knockdown (Fig.4.8). Tables with all gene names in the specific pathways are shown in the Appendix for this chapter (Chapter 4 Table 4.8).

A


Biological process
Small molecule catabolic process
carboxylic acid transmembrane transport
regulation of cellular localization
cellular response to fatty acid
regulation of hormone levels
peptidyl-proline modification
**ribonucleoside monophosphate
***cytokine production immune response
alpha-amino acid biosynthetic process
positive regulation of binding
7
5
9
3
7
3
3
3
3
4
0
2
4
6
-log10(p value)

B


Cellular component
mitotic spindle
lamellipodium
*component of organelle membrane
basolateral plasma membrane
sarcolemma
perikaryon
specific granule
secretory granule lumen
membrane raft
mitochondrial outer membrane
5
4
5
4
3
3
3
3
4
4
0
1
2
3
4
5
-log10(p value)

C

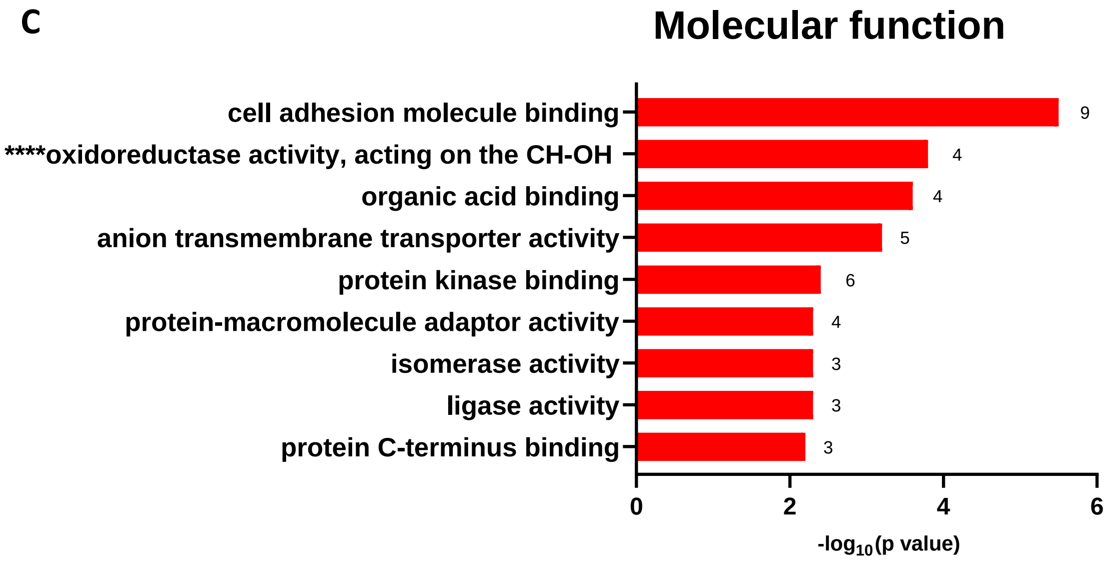
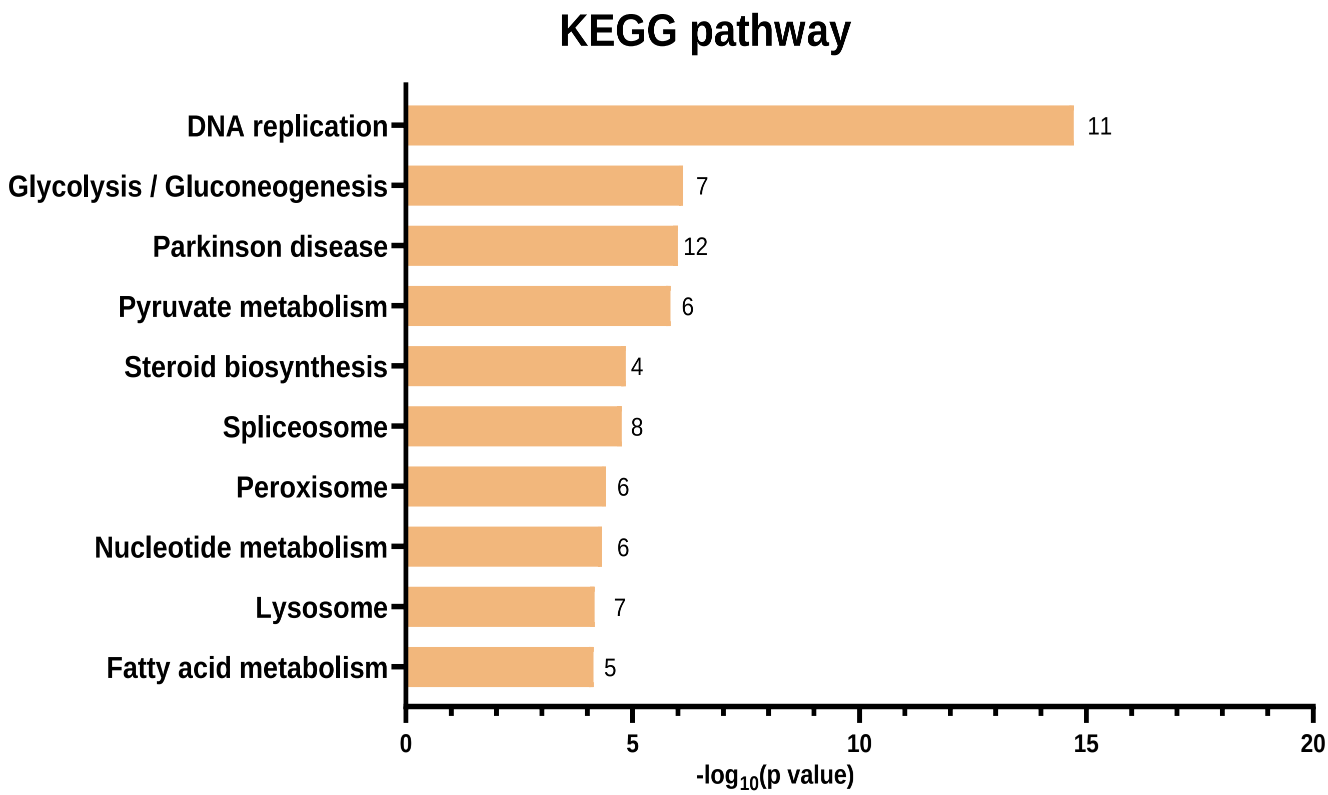


Figure 4.8 - Enrichment analysis for the biological function of the 65 genes obtained from mass spectrometry analysis through Metascape online tool. Analysis was performed on Biological Process (BP) (**A**), Molecular Function (MF) (**B**), and Cellular Component (CC) (**C**) in DU145 SPAG5 knockdown. Pathways are shown in descending order based on -$\log_{10}$ P-value. The number of genes associated is shown above each bar. Only pathways with a p-value <0.01 are shown.

KEGG analysis was performed on 230 and 65 gene lists (proteins) provided from mass spectrometry data. Protein lists were processed separately using Metascape and the analysis was customised just for KEGG pathways. For MDA-MB-231 20 pathways identified with a p-Value of 0.01 were associated with 230 proteins (Fig. 4.9A). For DU145 SPAG5 deficient using the same customised analysis 3 pathways were associated with 65 genes (Fig. 4.9 B). Tables with all the gene IDs for each pathway for the cell line are present in the Appendix for this chapter (Chapter 4 Table 4.9).

A



B

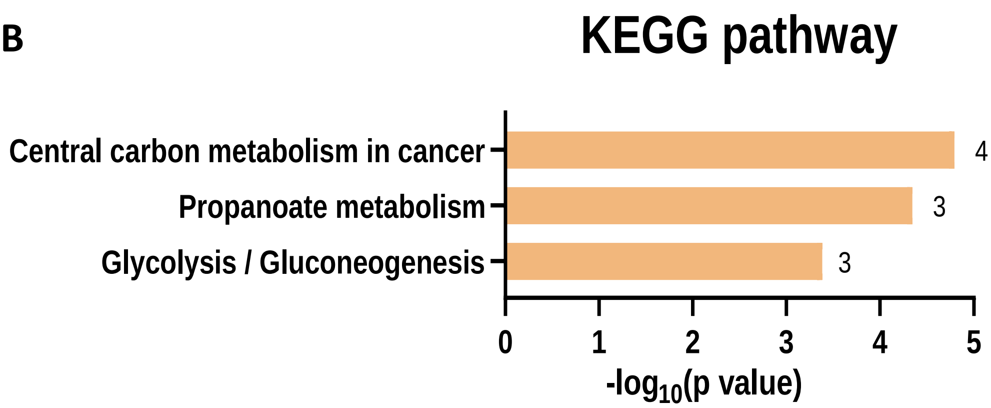
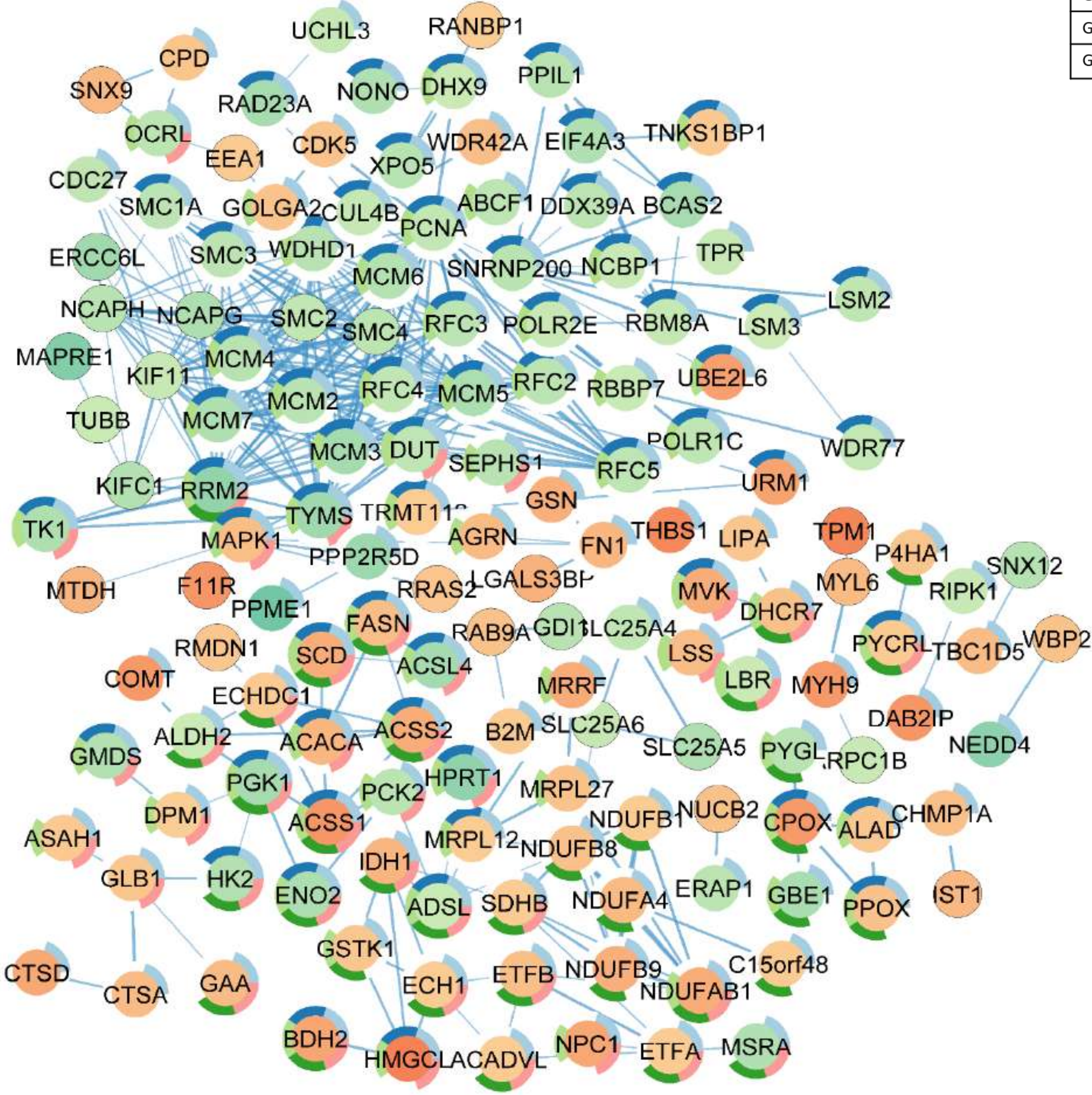


Figure 4.9 - KEGG pathways enrichment in MDA-MB-231 and DU145 SPAG5 deficient cells. KEGG pathways were generated from 239 proteins for MDA-231 (A) and 65 proteins DU145 (B) SPAG5 silencing found from mass spectrometry data using *StatPro*.Pathways in both bar graphs are shown in descending order based on the $\log_{10}$ P-value. The number of genes associated with the pathways is shown above each bar. Only pathways with a p-value <0.01 are shown.

To investigate the interaction of the DEPs a protein-protein interaction (PPI) was performed using Cytoscape online software. The analysis was performed on the 230 proteins of MDA-MB-231 and 65 genes in DU145 SPAG5 deficient obtained from the MS dataset. For the analysis, text mining, experiments and database for DU145 SPAG5 knockdown the medium confidence of 0.400 (Fig.4.11) was used while for MDA-MB-231 at confidence 0.700 (Fig.4.10 A).

A



B

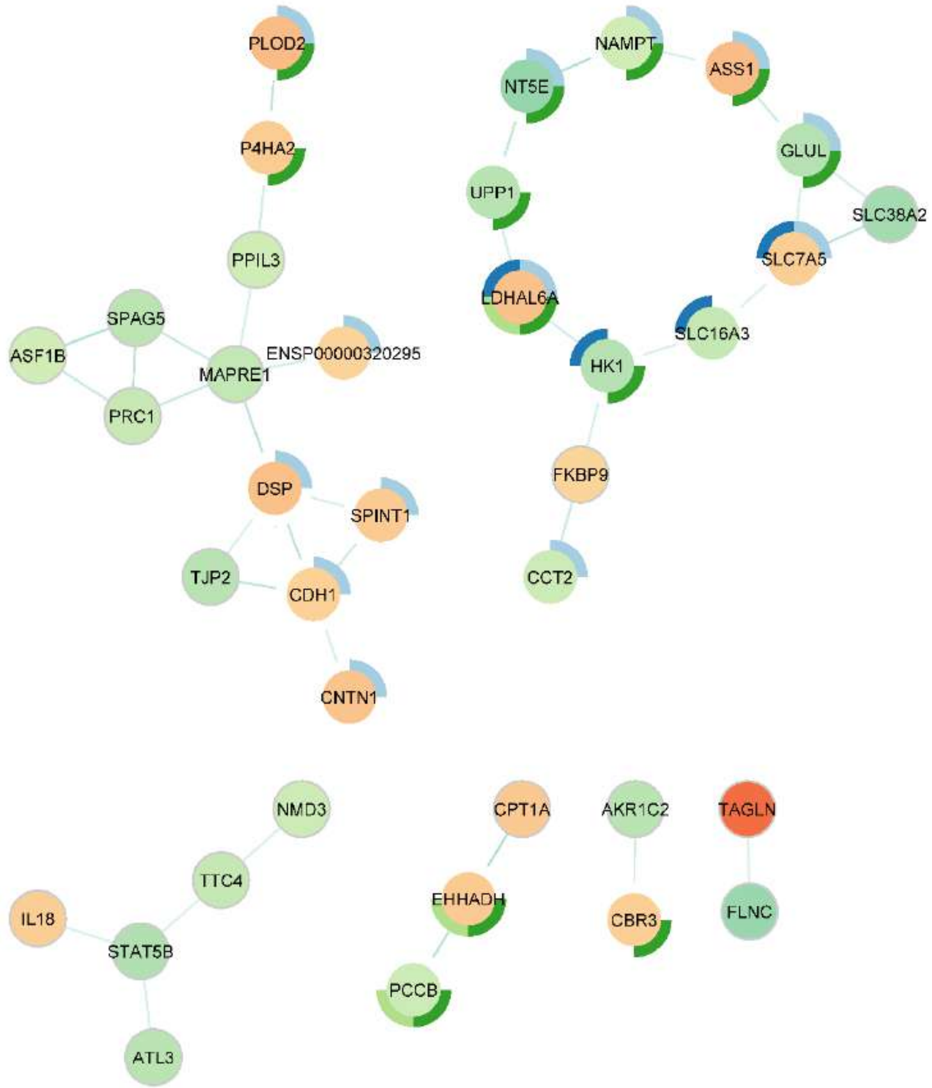
Category	Chart colour	Description	P-value
GO Biological process		Metabolic process	4.54E-15
GO Biological process		Heterocycle metabolic process	1.04E-09
GO Biological process		Biosynthetic process	1.08E-09
GO Biological process		Oxidation-reduction process	1.08E-09
GO Biological process		Small molecule metabolic process	1.22E-08

C

Network Stats
number of nodes: 239
number of edges: 218
PPI enrichment p-value: < 1.0e-16

Figure 4.10 - Protein-protein interaction networks on DEPs upregulated and downregulated identified in MS analysis in MDA-MB-231 SPAG5 deficient cells. **(A)** The PPI of the DEPs upregulated genes was visualised in Cytoscape (v3.9.1). Each node (proteins) was coloured based on the Log2FC with the darkest red indicating the highest expression. The thickness of the node connecting represents the size of the comparisons. The split donuts chart around the nodes represents five non-redundant enrichment clusters of the upregulated genes as reported in table **(B)**. The entire set of genes exhibiting an absolute Log2 FC ≥ 0.58 in gene expression included 126 upregulated and 104 downregulated. Network was generated exhibiting the largest modulation with setting at high confidence (0.700). **(C)** Table showing the analysis of the network generating from CYTOSCAPE where nodes representing proteins and the edges the interaction between two proteins. Significance in the number of edges and nodes is represented by the PPI enrichment p-value.

A



B

Category	Chart colour	Description	P-value
GO Cellular component		Extracellular exosome	0.001
KEGG pathway		Central carbon metabolism in cancer	0.0337
KEGG pathway		Propanoate metabolism	0.0417
KEGG pathway		Metabolic pathway	0.0417

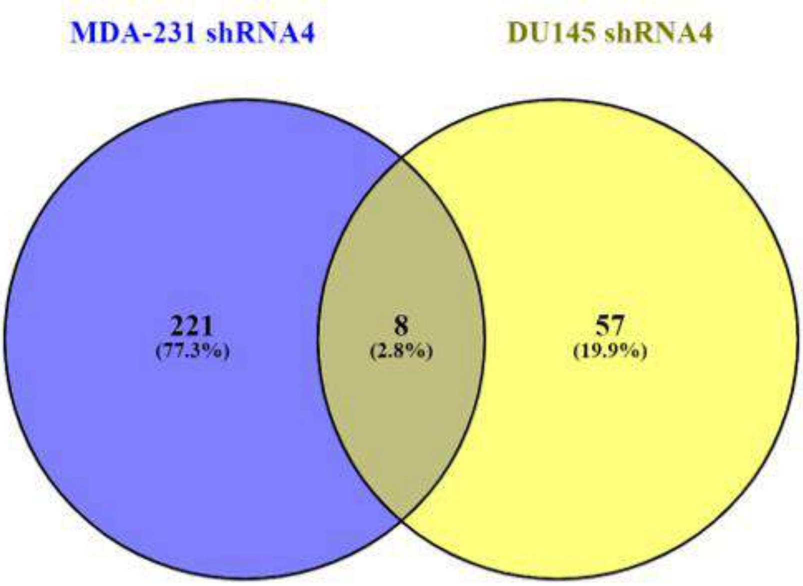
C

Network Stats
number of nodes: 65
number of edges: 37
PPI enrichment p-value: 0.0053

Figure 4.11 - Protein-protein interaction networks on DEPs upregulated and downregulated identified in MS analysis in DU145 SPAG5 deficient cells. (A) The PPI of the DEPs upregulated genes was visualised in Cytoscape (v3.9.1). Each node (proteins) was coloured based on the Log2FC with the darkest red indicating the highest expression. The thickness of the node connecting represents the size of the comparisons. The split donuts chart around the nodes represents five non-redundant enrichment clusters of the upregulated genes as reported in table **(B)**. The entire set of genes exhibiting an absolute Log2 FC ≥0.58 in gene expression included 22 upregulated and 43 downregulated. The network was generated exhibiting with the setting of confidence (0.400). **(C)** Table showing the analysis of the network generating from CYTOSCAPE where nodes representing proteins and the edges the interaction between two proteins. Significance in the number of edges and nodes is represented by the PPI enrichment p-value. ENSP00000320295 (TUBB3).

Finally, to investigate whether MDA-MB-231 and DU145 SPAG5 knockdown present common genes a Venn diagram was generated using an online tool Venny 2.1.0 (Fig. 4.12).

A



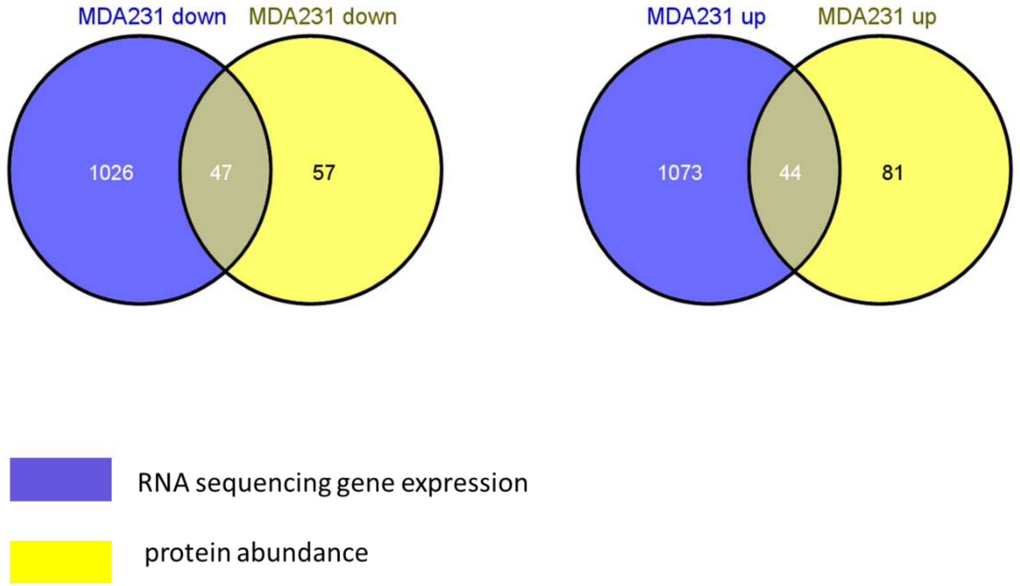
B

IDs	MDA-MB-231 shRNA4	DU145 shRNA4
MAPRE1	dowregulated	dowregulated
ZMPSTE24	dowregulated	dowregulated
GBE1	dowregulated	dowregulated
CUL2	dowregulated	dowregulated
NEDD4	dowregulated	dowregulated
EDIL3	upregulated	upregulated
FKBP9	upregulated	upregulated
ATP2B4	upregulated	downregulated

Figure 4.12 - Venn diagram analysis on MDA-MB231 and DU145 SPAG5 deficient. (**A)** Venn diagram was created by using the Venny tool (https://bioinfogp.cnb.csic.es/tools/venny/) considering the significant protein abundance downregulated and upregulate in MDA-MB-231 shRNA4 and DU145 shRNA4 SPAG5 silencing. padj <0.05 cut off and Flog2FC 0.58. **(B)** Table on the right shows the list of the common upregulated and downregulated genes from the Venn diagram. Eight common genes are represented.

4.3.2 Combined gene and proteome data set for common upregulated and downregulated markers in MDA-MB-231 and DU145 SPAG5 silencing.

Cross-over gene and proteome analysis was performed using DEGs identified in chapter 3 with DEPs from MDA-MB-231 SPAG5 using Venny diagram. Among 47 common downregulated and 44 common upregulated proteins 28 are showed in the table sorted on significant. Same analysis was performed for DU145 SPAG5 deficient cells. Analysis showed that 21 common downregulated proteins and 3 upregulated proteins were identified when combined gene and proteome data set.

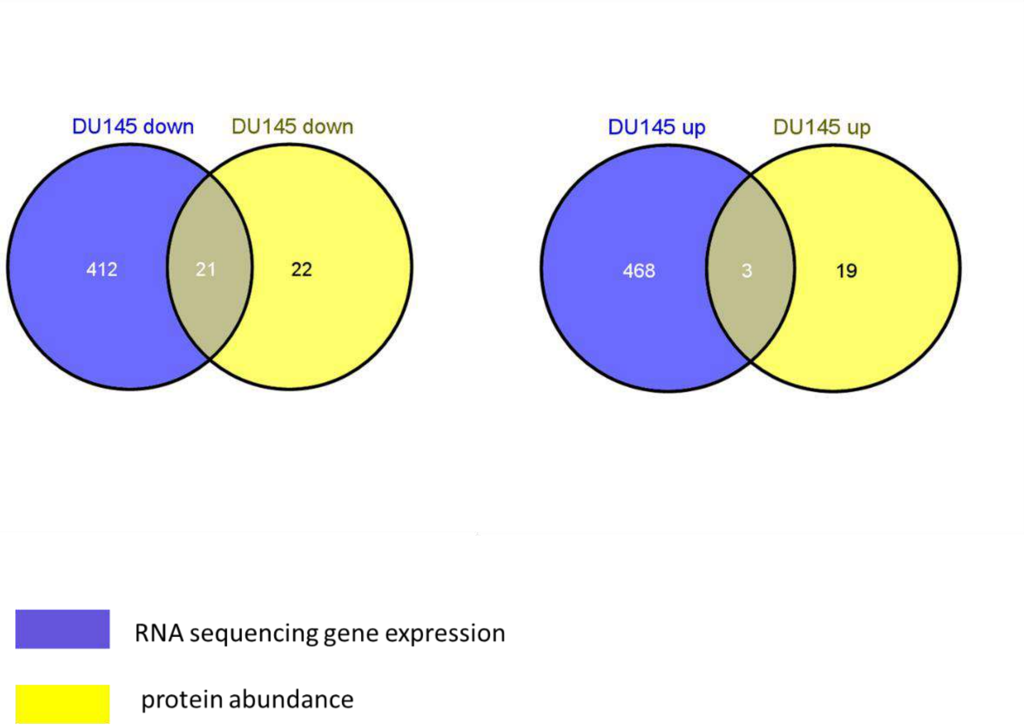


Cross-over between gene data and proteome data set in breast cancer MDA-MB-231 SPAG5 silencing

Downregulated	Upregulated
HPRT1	CTSD
MAPRE1	IGFBP7
MCM3	PTTG1IP
RRM2	GSN
NONO	LGALS3BP
PPME1	TACSTD2
SLC25A5	LSS
TUBB	FAM3C
KIFC1	DHCR7
PGK1	CTSA
OCRL	MELTF
ZMPSTE24	ACSS2
NCAPH	CPOX
ABCF1	MYH9
RAD23A	UACA
NCAPG	MGST1
HNRNPAB	EDIL3
ERCC6L	MYDGF
PCNA	IDH1
KIF11	MYO18A
CUL2	LIPA
PPP2R5D	ACACA
MCM5	COMT
MCM6	DAB2IP
ENO2	ACSS1
GDI1	PGLS
XPO5	FN1
WDHD1	AHNAK2

Figure 4.13 - Cross-over genes and proteome data set in MDA-MB-231 SPAG5 deficient. Venn diagram was created by using Venny tool considering the significant gene and protein abundance downregulated and upregulate in MDA-MB-231 SPAG5 silencing. padj <0.05 cut off and Flog2FC 0.58. Tables represent the common upregulate (orange) and downregulated (blue) proteins/gens commonly expressed in two data sets. The table on the left shows 28 out of 47 downregulated proteins and 28 out of 44 upregulated

proteins/genes identified in the cross-over gene and proteome data in MDA-MB-231 SPAG5 silencing (full list of common proteins is showed in the appendix of this chapter Table 4.13).



Cross-over between gene data and proteome data set in prostate cancer DU145 SPAG5 silencing

Downregulated	Upregulated
NAMPT	TUBB3
BCL2L13	TAGLN
SCAMP1	CPT1A
CCT2	
ATL3	
HK1	
NEDD4	
SPAG5	
NT5E	
FLNC	
SLC16A3	
LRRC8A	
ATP2B4	
MAPRE1	
CAPG	
AKR1C2	
ZMPSTE24	
CUL2	
GBE1	
NMD3	
PADI2	

Figure 4.14 - Cross-over genes and proteome data set in DU145 SPAG5 deficient. Venn diagram was created by using Venny tool considering the significant gene and protein abundance downregulated and upregulate in DU145 SPAG5 silencing. padj <0.05 cut off and Flog2FC 0.58. The table on the right shows 21 markers downregulated and 3 upregulated proteins/genes identified in the cross-over gene and proteome data in DU145 SPAG5 silencing.

4.4 Discussion

Sperm-Associated Antigen 5 (SPAG5) is a microtubule-associated protein, an essential component of the mitotic spindle required for chromosome segregation and progression through the anaphase of the mitotic process. It is also demonstrated to be responsible for progression in different types of cancer, including lung, and cervical cancer, as also Triple-negative breast cancer (Huang and Li 2020) (Yuan et al. 2014). In this chapter, the stable total proteome of SPAG5 knockdown generated in Triple-negative breast cancer cell line MDA-MB-231 and prostate cancer DU145 cell line was processed for mass spectrometry analysis (MS). Through mass spectrometry is possible to identify and profile a huge amount of protein from complex biological samples. This technology aims to detect differentially expressed proteins (DEPs) under different experimental conditions.

The whole lysate was processed as described in paragraph 4.1.1 of this chapter in the two-cell line SPAG5 knockdown cell population with their control (cell population expressing pLKO.1 empty vector) in both MDA-MB-231 and DU145 cell lines. Bioinformatic analysis was performed to investigate whether there were common and unique upregulated and downregulated genes between the two cells when comparing the knockdown with the ctrl. Using Morpheus online tool for heatmap generation, the DEPs of MDA-231 SPAG5 knockdown were presented and using an R-base Shiny application, the 30 most upregulated and downregulated were shown by volcano plot user-friendly web tool, called *StatsPro* was used for statistical approach (Fig.4.3 A-B). Data obtained from the *StatsPro* showed that for MDA-MB-231 SPAG5 knockdown, 126 upregulated genes and 104 genes were generated (Table 4.1). For DU145 SPAG5 knockdown, applying the same process, 22 upregulated and 43 downregulated genes were generated (Table 4.2).

***-Breast cancer* -** The analysis conducted on MS results in MDA-MB-231 SPAG5 knockdown highlighted those proteins involved in significant modulated pathways are the minichromosomal maintenance (MCM2-7) as also the replication factor C (RFC2-5) complex (Table 4.3). The MCM complex represents a family protein highly conserved hexameric complex of DNA binding protein composed of six subtypes, MCM2, MCM3, MCM4, MCM5, MCM6, and MCM7, involved in the replication of DNA. Dysregulation of this complex is responsible for cancer development and progression (Forsburg 2004). Studies conducted on different types of cancer, including lung, cervical and medulloblastoma, demonstrated a significant role of the subtype protein MCM2 in cancer progression (Amaro Filho et al., 2014). Each subtype was demonstrated to have a role in the progression of cancer and be potentially considered as a marker for proliferation or prognosis. Notably, subtypes MCM4 and MCM7 demonstrated a role in the proliferation is squamous cell carcinoma and oesophagal adenocarcinoma as also in lesions in which abnormal cells are associated with a high risk of developing cancer (Choy et al. 2016). Patients affected with glioma, hepatocellular carcinoma and endometrioid endometrial adenocarcinoma showing overexpression in the MCM6 subtype was also diagnosed with a poor survival rate (Hotton et al., 2018). In their studies, Issac and colleagues showed that a low transcription of the three subtypes MCM2, MCM4, and MCM6, together with the MKI67 gene, an essential gene for cancer proliferation, are associated with an increase in the probability of free-survival relapse from cancer (Issac et al. 2019). Another important complex that is highly modulated in SPAG5 knockdown is the human replication factor C (RFC 2-5) complex (Bowman, O'Donnell and Kuriyan 2004). This complex is involved in DNA replication as also in DNA damage repair. The RFC complex explicates its role in DNA replication as a clamp loader for the proliferating cell nuclear antigen (PCNA) to link into the DNA, allowing the DNA elongation (Sakato, O'Donnell and Hingorani 2012). It also works in DNA damage binding to the cell cycle protein checkpoints (Majka, Chung and Burgers 2004). It was demonstrated that upregulation of RFC is common in different types of cancer, including breast cancer, prostate cancer, colorectal cancer, and hepatocellular carcinoma (Arai et al. 2009) (Ji, Li and Wang 2021). In Triple-negative breast cancer and ER-positive and ER-negative RFC2 subunit acts as a marker for progression and metastasis in those forms of breast cancers (Ji, Li and Wang, 2021). In a study conducted using five datasets for prostate cancer shown, 33 genes were identified, including the RFC4 and RFC5 subunits that are overexpressed and are helpful as prognostic cell cycle progression markers (Barfeld et al. 2014). In this study, specifically, results show that those complexes are downregulated in MDA-MB-231 SPAG5 knockdown cell populations; this could suggest a correlation between those downregulation complexes with SPAG5 activity in cell mitosis.

The dataset generated from mass spectrometry analysis was then used for the enrichment analysis by online Metascape software. The data were customised for GO Biological process (BP), Molecular Function (MF) and Cellular Component (CC), shown on a bar graph with the number of a gene associated above each bar (Fig 4.6). The GO enrichment analysis conformed analysis identified in the most significant pathways (BP, MF, CC) present the complex MCM 2-7 and RFC2, RFC3, RFC4 and RFC5 subunit. Furthermore, KEGG pathways were performed on the 230 genes from mass spectrometry data, and a bar graph was generated with the most significant pathways (Fig.4.9 A) with 'DNA replication' were the MCM2-7 complex and RFC2, RFC3, RFC4, RFC5 are associated with it, followed by 'glycolysis /gluconeogenesis. Interesting, the genes associated with the glycolysis ENO2 and ACSS1 and ACSS2 are associated with MF, BP, and CC pathways. Finally, PPI analysis was performed using Cytoscape software. Split coloured donuts chart customize proteins in different GO terms, 6 proteins (MCM2, MCM3, MCM6, MCM5, MCM7) are enriched in the metabolic process, 2 proteins that are upregulated in MDA-MB-231 SPAG5 knockdown (ACSS1 and ACSS2) are enriched is the biosynthetic and oxidation-reduction process and (Fig 4.10 A). Acetyl-CoA synthetases ACSS1 and ACSS2 promote the conversion of acetyl-CoA in lipid synthesis and are also involved in energy production and protein acetylation (Miller and Schug, 2021). It was also demonstrated that in some cancer, this enzyme is present at a high level (Comerford et al., 2014). In melanoma cells, ACSS1 is involved in the survival and tumour growth, while ACSS2 is responsible for the regulation of carcinogenesis in different types of cancer, including glioblastoma, prostate, and breast cancer (Mashimo et al., 2014) (Schug et al. 2015). In this study, those two ACSS1 and ACSS2 proteins are upregulated in the SPAG5 knockdown cell population in the MDA-MB-231 cells line, which could suggest a link between SPAG5 and metabolic stress. In a study conducted on melanoma cells, it was reported that the deprivation of glucose depends on acetate metabolism to maintain ATP production. They also demonstrated that a reduction in glucose increased the expression of ACSS1 and ACSS2 (Lakhter et al., 2016). Glucose is essential for cancer cells to survive, and glycolysis is the pathway by which a glucose molecule is separated into two molecules of pyruvate to produce energy (Lin et al., 2020). Neuron-specific enolase 2 is an important enzyme involved in glycolysis. It is demonstrated that its overexpression of ENO2 is responsible for malignancies in breast cancer as also small cell lung cancer (SCLC) and neuroendocrine tumours (Miremadi et al. 2002) (Schofield, Lincz and Skelding 2020). In this study, the enolase 2 (ENO2), an enzyme involved in the glycolysis process, is downregulated in the SPAG5 knockdown cell population. ENO2, together with ACSS1 and ACSS2, is associated with 'Glycolysis, gluconeogenesis' in KEGG enrichment pathways that could suggest a potential correlation between SPAG5 downregulation and the metabolism process in breast cancer progression.

-Prostate cancer- The same bioinformatic approach used for MDA-MB-231 SPAG5 knockdown was applied for the DU145 SPAG5 knockdown cell population. From mass spectrometry analysis, 65 DEPs were identified with 22 upregulated and 43 downregulated. With the same online tool for heatmap generation, Morpheus, the DEPs 65 gene were presented and using an R-base Shiny application, the 30 most upregulated and downregulated were shown by volcano plot (Fig 4.4 A-B). Amon the 30 most upregulated gene procollagen-lysin,2-oxoglutarate 5-dioxygenase (PLOD2), is a membrane-bound homodimeric enzyme involved in the extracellular matrix formation and various pathological process. High expression of PLOD2 has been found in different malignancy included: breast cancer, sarcoma, and hepatocellular carcinoma (Gilkes et al., 2013) (Noda et al., 2012) (Eisinger-Mathason et al., 2013). The function of PLOD2 involving the cross-link switch and tumour cell invasion and migration has been validated also in relation with metastasis (Du et al., 2017). Cadherin 1 (CDH1) is a tumour suppressor gene and germinal mutation has been associated with hereditary diffuse gastric cancer (HDGC) and lobular breast cancer. Dysregulation of CDH1 is responsible of tumour proliferation, invasion, migration, and metastasis (Shenoy, 2019). Study conducted RNA-binding protein Apobec1 complementation factor (AICF) on hepatocellular carcinoma showed a positive correlation between SPAG5 and CDH1 and both markers showed an altered expression (Blanc et al., 2021). Results has showed that downregulation of SPAG5 showed an increased expression of CDH1 leading to the questions wheatear SPAG5 downregulation could affect progression in prostate cancer. High enrichment pathways were generated with the Metascape tool, and a table with the associated genes is shown in table 4.4. Among the most significantly affected three pathways are 'Ribonucleoside Monophosphate Metabolic process', 'metabolism of water-soluble vitamins and cofactor' and 'Adipogenesis' are affected by SPAG5 knockdown. Particularly, the gene involved in the adipogenesis and ribonucleoside monophosphate metabolic pathways, ecto-5'-nucleotidase (NT5E) is demonstrated to be dysregulated in different types of cancer, including prostate cancer (Yang, Q., Du, and Zu 2013). Interestingly, even in DEPs MAPRE1 is significant downregulated as showed in chapter 3 results, DEGs for DU145 showed a downregulation of protein. Downregulation also led to a decrease in the expression of hexokinase 1 (HK1) in DU145 SPAG5 silencing but also in MDA-MB-231 SPAG5 silencing. Hexokinase 1 is a key regulator in glycolysis and oncogenic marker (Gao, Y. et al., 2015). Study conducted on ovarian cancer has showed that deletion of this marker suppressed the growth in xenotransplant ovarian cancer and nearly abolished the tumour growth in mice fed with glucose-free diet (Šimčíková et al., 2021). Enrichment analysis showed that HK1 is involved in glycolysis/gluconeogenesis and central carbon metabolism pathways, extending a potential involvement of SPAG5. Enrichment analysis was obtained using Metascape and differentially expressed proteins were represented on a bar graph and customised for GO biological process, molecular function, and cellular components (Fig.4.8). Data were analysed for protein-protein interaction (PPI) using Cytoscape online tool, with less protein interaction than seen in MDA-MB-231 (Fig.4.10). The genes associated with cellular components and KEGG pathways were generated as shown (Fig. 4.).

-Breast and prostate combined dataset- Finally, a Venn diagram was generated to investigate whether the two-cell line could present any common gene together (Oliveros 2007) (Fig.4.12). The two data sets were input into the online tool, and results show that eight genes are common between the two-cell line while no statistically significant pathways were commonly present. Among them, five are commonly downregulated (MAPRE1, ZMPSTE24, GBE1, CUL2 and NEDD4) and two are commonly upregulated (EDIL3 and FKBP9) between the two cell lines downregulated for SPAG5. Microtubule-associated protein RB/BE family member 1 (MAPRE1) is associated with the mitotic cell cycle and cell adhesion pathways in both cell lines. A different study demonstrated that deregulation in MAPRE1 is associated with different types of cancer, including colorectal cancer, gastric cancer, and acute lymphoblastic leukaemia (Ladd et al., 2012) (Kim et al., 2011). In breast cancer MDA-MB-231 SPAG5 knockdown, downregulated MAPRE1 is enriched in the mitotic cell cycle and spindle organisation and mitotic cell cycle and signalling by Rho GTPase for DU145 SPAG5 knockdown. In hepatocellular carcinoma (HCC), it was demonstrated that high levels of MAPRE1 are associated with cell proliferation enhancing tumorigenesis and poor prognosis in patients with HCC (Liang et al., 2020).In Breast cancer, MDA-MB-231 and DU145 SPAG5 knockdown of the E3 ubiquitin-protein ligase (NEDD4) is commonly downregulated and this protein is enriched in adipogenesis and cellular component assembly involved in morphogenesis terms for DU145 SPAG5 knockdown and in MDA-MB-231 SPAG5 is significantly enriched in respond to steroid stimuli and viral process term. Although the role of NEDD4 as an oncogene is known in different cancers (Amodio et al., 2010) (Jung, Samil et al., 2013), recently been demonstrated to suppress some cancer targeting Myc and RAS oncoprotein for ubiquitination and degradation (Zeng, T. et al., 2014) (Liu, P. Y. et al., 2013). Finally, evidence of the regulation of NEDD4 in cancer progression was obtained in breast cancer in which silencing of NEDD4 decreases cell proliferation in all breast cancer cell lines studied (Wan et al., 2019). Commonly upregulated extracellular matrix protein (EDIL3) is enriched in integrin binding and structural molecule activity terms for MDA-MB-231 SPAG5 knockdown and cell adhesion molecules terms in DU145 SPAG5 knockdown. EDIL3 protein has been shown to promote the epithelial-mesenchymal transition through the interaction of integrin $\alpha_v\beta_3$ (Gasca et al., 2020).

DEPs data set and DEGs from MDA-MB-231 SPAG5 silencing showing 47 commonly expressed markers downregulated and 44 commonly upregulated markers, showing to the similar expression in both data sets RNA-seq and quantitative MS. Among those commonly markers MAPRE1, NONO, NEDD4, ZAMPSTE24, GBE1 and CUL2 are downregulated while common upregulated include ACSS1, ACSS2. In prostate cancer DU145 silencing for SPAG5 21 are the proteins-genes commonly downregulated in both data sets and just 3 are the proteins-gene commonly upregulated in both data set. Among 21 commonly downregulated markers MAPRE1, NEDD4 are present while TUBB3, TAGLN, and CPT1A are the only commonly upregulated markers in both data set. Interestingly the actin-associated protein Transgelin (TAGLN) is strongly upregulated in prostate cancer DU145 SPAG5 silencing in both data set. Study conducted in prostate cancer cell line demonstrated that downregulation of TAGLN is promoting proliferation and suppressed the migration of prostate cells (Wen, F. et al., 2021). Our result showed that SPAG5 downregulation induces and upregulation of this gene this could suggest an involvement of SPAG5 in the regulation of TAGLN in prostate cancer cells proliferation. Finally, downregulated data set from RNA-seq and quantitative MS of MDA-MB-231 SPAG5 silencing combined with DU145 SPAG5 silencing has NEDD4, MAPRE1, ZMPSTE24, CUL2, and GBE1 while no common upregulated markers were observed between RNA-seq and quantitative MS data sets.

In conclusion, this study described the application of the SWATH-MS technique to generate large-scale quantitative proteomics profiles in two cancer cell line triple-negative MDA-MB-231 and DU145 cell line SPAG5 knockdown. Data showed that downregulating SPAG5 affects different proteins involved in the cancer progression at the cell cycle level and in cellular supports components such as the extracellular matrix. Breast and prostate cancer combined data set has shown that five gene are commonly downregulated and two are commonly upregulated in both cell line. In addition to allowing to classify the DEPs obtained in pathways involved in DNA replication, cancer metabolism and glycolysis pathways, the obtained SWATH-MS data allowed to compare the protein and transcript level of the DEGs identified in the previous chapter 3. Although, for some the expression at protein and transcript level was difference, it was strong for some of them. Further studies on these proteins significantly downregulated and upregulated in MDA-MB-231 and DU14 after being silenced for SPAG5 could lead to a broader knowledge of SPAG5 function in cancer progression.

5. Effect in cell cycle mechanism in treated MDA-MB-231 and DU145 SPAG5 deficient cell population with chemotherapeutic drugs and correlation analysis of commonly genes/proteins identified *in silico* TCGA datasets.

5.1 Introduction

Despite the inevitable side effects such as nausea and hair loss, chemotherapies still represent an essential treatment modality in fighting the cancer (Behranvand et al., 2021). The high sensitivity of tumours to chemotherapeutic drugs, partially due to altered gene expression or variation in the kinetic growth. This could partially explain the difference in sensitivity of response to the chemotherapies drugs between the patients. In chemotherapy, multidrug resistance (MDR) also plays an essential role as it can negatively affect the treatment (Hamed et al., 2019). Therefore, a single chemotherapy drug is not sufficient for the treatment, combining chemotherapies or therapies, including radiotherapy, is necessary to avoid cancer cell proliferation, invasion, and metastasis (Randrian et al., 2020).

Cell cycle regulation during cancer progression influences tumour cell proliferation, metastasis, and recurrence (Rozengurt, 1999). The regulation of the cell cycle, a primary purpose of cancer treatment, is to control the expression of specific genes involved in cancer progression and the activity of enzymes and proteins or signal factors (Bonacci & Emanuele, 2020). The cell cycle represents a series of events that leads to cell growth and division. It has divided into a stage called G1 representing cell growth characterised by increasing RNA and protein synthesis, preparing for the DNA synthesis In the S stage. In this stage , DNA replication and the synthesis of chromosomal proteins such as histone and non-histone proteins are expected (Vermeulen et al., 2003) .

In the G2 phase, after the DNA replication, condensation of genetic material will prepare for the mitosis process, the M (mitosis) phase, which will lead to the division of chromosomes into two daughter cells with partitioning of genetic material (DNA) and proteins (Norbury & Nurse, 1992). Genetic information is then transmitted from the DNA replication to the cell daughter in the M phase, which is essential for the stability of genetic behaviours (Sun, Y. et al., 2021). For this reason, the S phase represents a crucial and critical stage in the cell cycle and, therefore, an important target of the chemotherapy drugs (Yano et al., 2020). In treating different types of cancer, anthracyclines are the most effective chemotherapies available targeting not only topoisomerase II (topo II), but also topoisomerase I (topo I), or the cytoplasmic targeting protein kinase C (PKC) (Lothstein et al., 2001) (Martins-Teixeira & Carvalho, 2020).

Approved for the first time in France in 1982, Epirubicin is the most used chemotherapy drug marketed in 80 countries, and 200 publications document the efficiency and safety for treating breast cancer and other malignancies (McGowan et al., 2017). Epirubicin explicates its mechanism of action as DNA intercalate through Topoisomerase II activity inhibition, causing the DNA strands to break. At the cell cycle level, Epirubicin is active mainly in the S and G2 phases of cell cycle (Martins-Teixeira & Carvalho, 2020).

Doxorubicin, another standard chemotherapy drug, acts in the same way described for Epirubicin; it also inhibits Topoisomerase I; ultimately, it is responsible for inducing programmed cell death (Tacar et al., 2013). A study on prostate cancer has shown that Doxorubicin induces apoptosis in LnCaP and DU145 cell lines and, at different level doses, is responsible for the fragmentation of DNA in > 40% of LnCap cells (Collins, L. et al., 2006).

SPAG5 is involved in the mitosis process of the cell cycle and is involved in maintenance of the sister chromatids' cohesion and centrosome integrity. It is involved in tumorigenesis and already been documented in different types of cancer, including cervical and bladder cancer (Abdel-Fatah et al., 2016b).

In breast cancer, SPAG5 is associated with cancer cell proliferation. A study on Triple-negative breast cancer (TNBC), the most lethal subtype of breast cancer, showed that inhibition of SPAG influences the transition from the S/G1 inducing S to G1 arrest (Li, M. et al., 2019) .

In this chapter, SPAG5 deficient MDA-MB-231 and DU145 were treated with two doses of chemotherapy drug and processed for cell cycle analysis versus control pLKO.1 (empty vector). Therefore, cell populations MDA-MB-231 SPAG5 deficient were treated with Epirubicin, while DU145 SPAG5 knockdown was treated with Doxorubicin. Aim of this chapter is to investigate whether the treatment of MDA-MB-231 and DU145 SPAG5 deficient affects the cell cycle and at what concentrations those effects are potentially observed. Finally, commonly upregulated, and downregulated genes identified in chapter 3 and 4 were validated using *in silico* data sets from cBioPortal portal for correlation analysis with SPAG5. The cBioPortal for Cancer Genomic provides a Web source for the analysis of multidimensional cancer genomic data. This portal was specifically designed to easily access to the complex data sets and improve the translation of genomic data into new biological insight, therapies, and clinical trials (Cerami et al., 2012b). Large scale cancer genomic projects such as The Cancer Genome Atlas (TCGA) are collected in cBioPortal (Cerami et al., 2012b) (Collins, F. S. & Barker, 2007).

5.2 Materials and Methods

5.2.1 Effect of SPAG5 silencing on cell proliferation using IncuCyte®

For this assay pLKO.1 and shRNA4 in both MDA-MB-231 and DU145 were used, and the protocol provided by the manufacturer applied. Cells were cultured and harvested, according to the protocol described in the section 3.2.1, at 80% confluence. DU145 and MDA-MB-231 cells were then resuspended at 1.0 x 10^5 cell/mL cell density, 100 µL of cell suspension were transferred (1.0 x10^4 cell/mL) in 96 -well flat bottom plate in six replicates. Then, the plate was transferred to the IncuCyte ZOOM® system incubator, allowed to warm to 37°C for 30 min, and scanned in Phase Contrast every hour for a total of one week for MDA-MB-231 and 72h for DU145 and time points collected cells growth. For MDA-MB-231 cells were grown in the IncuCyte ZOOM® system incubator without CO_2 per the suggested culture conditions of ATCC.

5.2.2 Drug titration Doxorubicin and Epirubicin

Cell viability was assessed using ready-to-use, non-toxic, resazurin-based solutions AlamarBlue™ cell viability reagent (ThermoFisher; DAL1100). To test cytotoxicity MDA-MB-231 SPAG5 deficient cell and MDA-MB-231 pLKO.1 empty vector were treated with Epirubicin (Epi), and for DU145 SPAG5 deficient cell pLKO.1 empty vector with Doxorubicin (DOX). For this assay cells were cultured and harvested as described in the section 3.2.1 at below 80% of confluence. MDA-MB-231 shRNA4 and MDA-MB-231 pLKO.1 cells were then resuspended at 1.0 x 10^5 cell/mL density, 100 µL of cell suspension was transferred (1.0 x10^4 cell/mL) in 96 -well flat bottom plate in six replicates. Epi (Selleckchem.com; S1223) was reconstituted by adding DMSO (100 mg/mL) to a working concentration of 10 mM. Cells were grown in their specific media supplemented with Epi at concentrations (0; 0.2; 0.4; 0.8; 1.6; 3.12; 6.25; 12.5; 50; 100; 150 µM) for 24h. DU145 shRNA4 and DU145 pLKO.1 were growth in their dedicated media with DOX (Sigma; PHR1789) at concentration (0; 0.2; 0.4; 1.6; 3.12; 12.5; 50; 100; 150 µM) for 48h. At the end of the incubation both cell line MDA-MB-231 and DU145 cell viability was detected using 1/10th of alamarBlue® reagent and incubated for 1h at 37°C protected from the light. The result was recorded using a Tecan i-control infinite200Pro microplate reader at 570 nm. Fluorescence data were fitted using IC_{50} [Inhibitor] vs. response (three parameters), and the half maximal inhibitory concentration (IC_{50}) of drugs response was determined using GraphPad Prism 9.0.0 software.

5.2.3 DNA content in MDA-MB-231 and DU145 SPAG5 treated with Epi and DOX

To investigate the cytotoxic effect of the two-chemotherapy drug in MDA-MB-231 and DU145 SPAG5 silencing, cells were incubated in their dedicated media supplemented with specific concentration of Epi for MDA-MB-231 and DOX for DU145. Based on the IC_{50}, there were two concentrations used; first, two times higher than calculated IC_{50} and second two time lower than the IC_{50} value obtained from the titration curve for both Epi and DOX. Cell population shRNA4 – SPAG5 and pLKO.1 empty vector MDA-MB-231 and DU145 growth in T25 flask with their dedicated media supplemented with specific high and low dose of Epi for 48 h and DOX for 24 has shown in table 1.

Table 5.1 - Chemotherapy drugs concentration used for MDA-MB-231 and DU145 SPAG5 deficient and pLKO.1 empty vector. The table shows the different amount of Epirubicin (Epi) used for MDA-MB-231 shRNA4 SPAG5 and pLKO.1 versus the amount of Doxorubicin (DOX) used for DU145 shRNA4 SPAG5 and pLKO.1 empty vector. Concentrations was selected two times high (2 high) and two times low (2 low) based on the IC_{50} of Epirubicin (Epi) for MDA-MB-231 and doxorubicin (DOX) for DU145.

	pLKO.1		shRNA4	
Epi	2x high	2x low	2x high	2x low
MDA-MB-231	7.5(μM)	1.9(μM)	7.3(μM)	1.8(μM)
DOX	2x high	2x low	2x high	2x low
DU145	1.1(μM)	0.3(μM)	0.6(μM)	0.3(μM)

Samples of untreated and treated SPAG5 deficient and pLKO.1 cell population were analysed for DNA content/cell cycle in both in MDA-MB-231 and DU145. For this analysis, cells were incubated with specific drugs and harvested in accordance with procedure as described in chapter 3 section 3.2.1. Cells were counted as described in section 3.2.1 and $1x10^6$ were transferred in 12 labelled centrifuge tubes for each cell line. After centrifuge for 5 min at 300xg cells were fixed with 500 μL of ice cold 70% ethanol in dH_2O added gradually whilst vortex to avoid precipitation and incubated on ice for 30 min. Ethanol was removed carefully, after centrifuge for 5 min at higher centrifugal force (850xg) and washed twice with 2mL of PBS and centrifuged for 5 min at 300xg. Each sample tube was incubated for 10 min with 100 μg / mL of Ribonuclease A from Bovine pancreas Type I (RNase I) (Sigma # R4875) at room temperature (RT). Finally, 50 μg / mL of propidium iodide (PI) (Sigma # 81845) were added to each tube and incubated at RT for 10 min. After incubation with PI 200 μL of isoton solution was added to each tube and analysed directly on Beckman Coulter flow cytometry. Assay was carried out three times and results are representative of three independent experiments. Cell cycle analysis was performed after appropriate gating of the population in FL-2 vs the FL-2 region plots with PI fluorescence (λex 482 nm: λem 608 nm). Acquisition rate was set at no more than 500 cell / s. Based on PI fluorescence three region should expected to be observed. The most-left peak corresponding to G0/G1 phase, the right-most peak represents cells in G2/M phase while the centre region are cells in S phase. An incubation of cells with drug but without PI was performed to ensure no interference in the PI-region was observed with the drug during the reading.

5.2.4 *In silico* correlation analysis of publicly available RNA-seq datasets

Data sets from The Cancer Genome Atlas (TCGA) for breast and prostate cancer were available from cBioPortal platform (Cerami et al., 2012b) . Data were transformed in Log_2 scale and sorted for *SPAG5.* The expression of commonly upregulated and downregulated genes identified in chapter 4 (Fig. 4.13-4.14), was analysed in correlation with *SPAG5* and represented in Table 5.2. The Pearson correlation coefficient (*r*) was used for measuring a linear correlation and scaled to the range from -1 to +1, where 0 indicates no correlation association (Schober et al., 2018). Correlation coefficient can be described as low, moderate, or strong relationship (Mukaka, 2012).

Table 5.2. Cross-over genes and proteome data set in MDA-MB-231 and DU145 SPAG5 deficient. Tables represent the common upregulate (orange) and downregulated (blue) markers commonly expressed in two data sets. The table on the left shows 28 out of 47 downregulated markers and 28 out of 44 upregulated markers identified in the cross-over gene and proteome data in MDA-MB-231 SPAG5 silencing (full list of common markers is showed in the appendix of this chapter). The table on the right shows 21 markers downregulated and 3 upregulated markers identified in the cross-over gene and proteome data in DU145 SPAG5 silencing.

Cross-over between gene data and proteome data set in breast cancer MDA-MB-231 SPAG5 silencing

Downregulated	Upregulated
HPRT1	CTSD
MAPRE1	IGFBP7
MCM3	PTTG1IP
RRM2	GSN
NONO	LGALS3BP
PPME1	TACSTD2
SLC25A5	LSS
TUBB	FAM3C
KIFC1	DHCR7
PGK1	CTSA
OCRL	MELTF
ZMPSTE24	ACSS2
NCAPH	CPOX
ABCF1	MYH9
RAD23A	UACA
NCAPG	MGST1
HNRNPAB	EDIL3
ERCC6L	MYDGF
PCNA	IDH1
KIF11	MYO18A
CUL2	LIPA
PPP2R5D	ACACA
MCM5	COMT
MCM6	DAB2IP
ENO2	ACSS1
GDI1	PGLS
XPO5	FN1
WDHD1	AHNAK2

Cross-over between gene data and proteome data set in prostate cancer DU145 SPAG5 silencing

Downregulated	Upregulated
NAMPT	TUBB3
BCL2L13	TAGLN
SCAMP1	CPT1A
CCT2	
ATL3	
HK1	
NEDD4	
SPAG5	
NT5E	
FLNC	
SLC16A3	
LRRC8A	
ATP2B4	
MAPRE1	
CAPG	
AKR1C2	
ZMPSTE24	
CUL2	
GBE1	
NMD3	
PADI2	

5.2.5 Statistical analysis

All statistical analysis was performed using GraphPad Prism 8.4.2 software (GraphPad, Inc, USA). For experiments where two groups were compared, a two-tailed Student's t-test was performed. For doses-response analysis for drug titration Fluorescence data were fitted using for IC50 [Inhibitor] vs. response (three parameters). For the analysis of how the response of drug is effective in different phases of the cell, cycle 2-way ANOVA with unpaired t-test for post-test for the single phase between the two concentrations compared with the untreated was performed. Unless otherwise stated, histogram columns represent the mean and error bars indicate the standard deviation (SD). The number of replicate (n) for each experiment is stated in the figure legend. The data is statistically significant if $P < 0.05$ and are indicated in the figure legends by asterisks.

5.3 Results

5.3.1 Effect of SPAG5 silencing on cell proliferation

In order to determine whether SPAG5 silencing affected cell proliferation in both MDA-MB-231 and DU145 a proliferation assay was performed using the Incucyte technology according to the manufacturer's protocol and shown in section 5.2.1. Incucyte is a real-time quantitative live imaging and analysis perform that efficiently analyse cell proliferation rate in a time dependent manner by analysing the imaging data. Measurement of cell proliferation over the time, percentage of cell confluence was obtained through IncuCyte (Essen BioScience, Ann Arbor, MI, USA) software by phase-contrast images. Data are presented as mean of phase area confluence for time point and normalised to time 0.

Figure 5.1 shows the results for the proliferation assay in both MDA-MB-231 and DU145 shRNA4-SPAG5 vs control pLKO.1 empty vector. The experiment was performed as described in Material and methods section 5.2.1 of this chapter. Data points were plotted using GraphPad Prism 8.4.2 software.

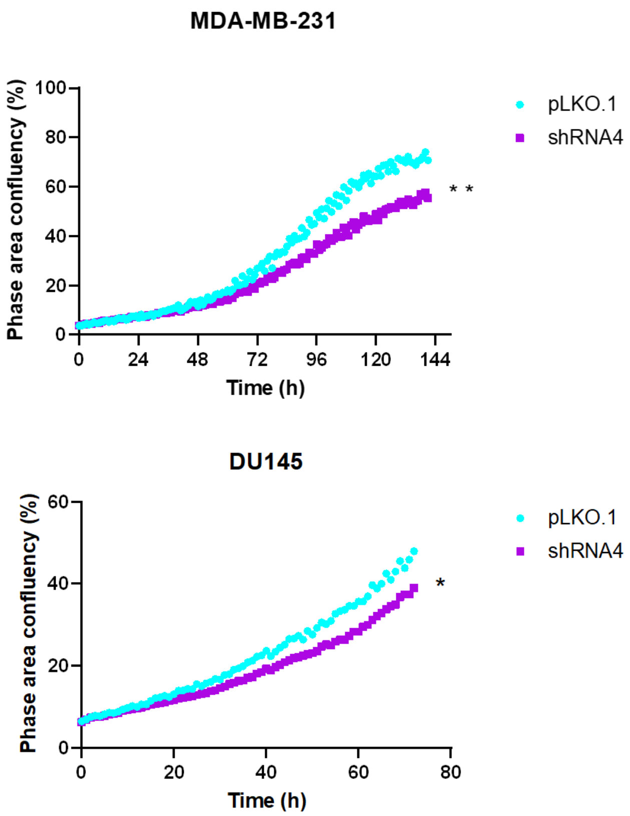
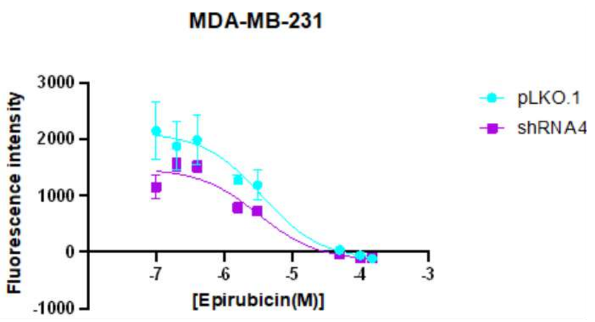
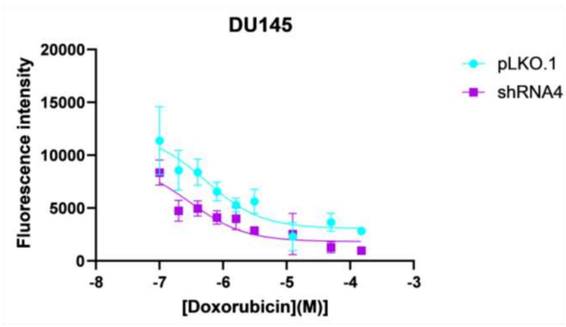


Figure 5.1 - Proliferation growth curve in MDA-MB-231 andDU145 cells SPAG5 deficient vs control pLKO.1. IncuCyte cell proliferation assay was assessed on MDA-MB-231 and DU145. Negative control, shRNA4-SPAG5 cell population were seeded in 96 Multiwell plate and let growing for 1 week for MDA-MB-231 and 72h for DU145. Each data point represents 6 biological replicates within the same experiment. Value represents the mean percentage phase area confluency ± SD for each point and normalised to the initial value at time 0.

5.3.2 Drug response to MDA-MB-231 and DU145 SPAG5 deficient

To study the effect of the chemotherapy drug Epi and DOX on MDA-MB-231 and DU145 SPAG5 silencing a titration was performed to determine the half minimum concentration of drugs able to trigger a response in both cell population and in their control. The experiment was performed as described in the section 5.2.2 and data read using Tecan i-control infinite200Pro microplate reader at 570 nm and shown in figure 5.3.





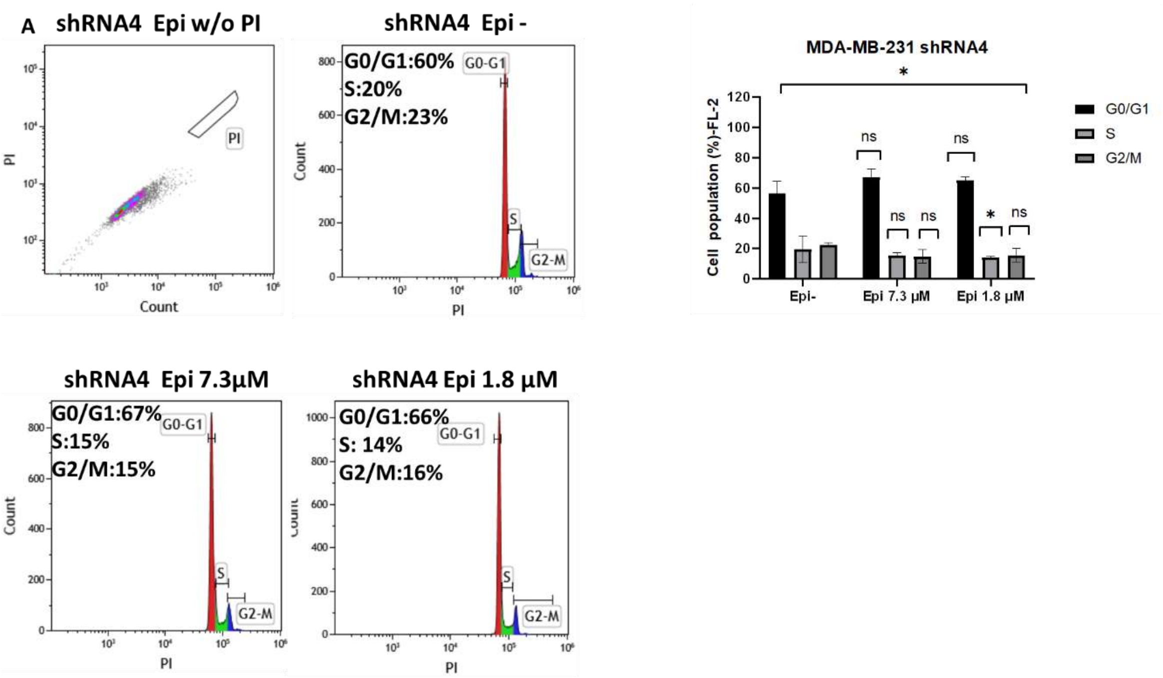
Epirubicin (M)	pLKO.1	shRNA4
IC_{50} M	3.624e-006	3.212e-006
R squared	0.9044	0.9298

Doxorubicin (M)	pLKO.1	shRNA4
IC_{50} M	5.422e-007	3.040e-007
R squared	0.75.18	0.7233

Figure 5.2 - Drug titration Epi and DOX in MDA-MB-231 and DU145 vs control pLKO.1. Cells were seeded in 96-flat Microplate and treated with Epi at different concentrations for 24h for MDA-MB-231 and with DOX 48h DU145. Cell viability was detected using 1/10th of alamarBlue® reagent and incubated for 1h at 37°C protected from the light. Result was recorded using Tecan i-control infinite200Pro microplate reader at 570 nm. Fluorescence data were fitted using for IC50 [Inhibitor] vs. response (three parameters) using GraphPad Prism 8.4.2 software. Biological replicate=6.Tables shown the concentration in Molar for both cell population and the goodness of the fit R^2 .

5.3.3 Analysis of cell cycle in MDA-MB-231 and DU145 SPAG5 treated with Epi and DOX

Flow cytometry analysis was performed on both cell lines after been treated with specific volume of chemotherapy drug. For this experiment cells were differently incubated with the specific drug at specific time. To test the effect of the drugs for both cell lines for each cell population it was chosen to use two doses of drug, two times higher and two time lower than the starting concentration for both Epi and DOX. The exact amount in µM is shown in table 5.1. Cells were prepared as described in the section 5.2.3. Quantitation of DNA content was obtained using the DNA- binding fluorochrome, propidium iodide for cell cycle analysis. To avoid interference between PI red-fluorescent dye as also Epi and DOX treated cells without PI was performed. Three regions were supposed to be observed that identify the DNA content in each of the three cell cycle stages as peaks. Meanly a left-most peak corresponding to the G0/G1 phase a right-most peak corresponding to a G2/M phase and finally a S phase in the central region. Therefore, each cell population SPAG5 deficient and pLKO.1 were prepared also with no drug as a control. The experiment was performed three times. After treated for 24h with two different concentrations cells showed a typical DNA-pattern representing by the G0/G1 phase, S phase and the G2/M phase of the cell cycle. Treated cells MDA-MB-231 shRNA4-SPAG5 deficient with 7.3 µM and 1.8 µM, showed an accumulation of DNA in G0/G1 phase of 66% and 67% when compared with the untreated (60%) (Fig 5.3 A). The treatment caused a reduction of the DNA accumulation to 15% and 14% respectively in the treated cells in G2/M phase while in untreated cells showed a DNA accumulation of 20%. Comparing means of the phases with the untreated no statistical difference were observed except for the G2/M phase. Same result is shown in the MDA-MB-231 pLKO.1 empty vector. An increasing of DNA is localised in G0/G1 phase in treated cell (62%) with 7.5 µM and 1.9 µM (61%) when compared with the untreated (55%) (Fig.5.3 B). Also, for pLKO.1 the treatment caused a decrease in the G2/M phase in both cell populations compared with the untreated. Statistical difference was seen between MDA-MB-231 empty vector pLKO.1 and SPAG5- shRNA4 MDA-MB-231 cells treated with Epirubicin with two higher dose in S phase of cell cycle (Fig.5.3 C) and not in the other cell cycle phase at different concentrations. No statistical differences were seen between the untreated and the treated cell line across the cell cycle phases in both MDA-MB-231 SPAG5 deficient and pLKO.1. Nevertheless, the decreasing of DNA in G2/M could suggest that low concentration could induce an arrest in the G0/G1 phase of the cell cycle however, these results should be taken carefully and in order to provide a stronger and clearer results, repetition of this experiment should be performed in the future.



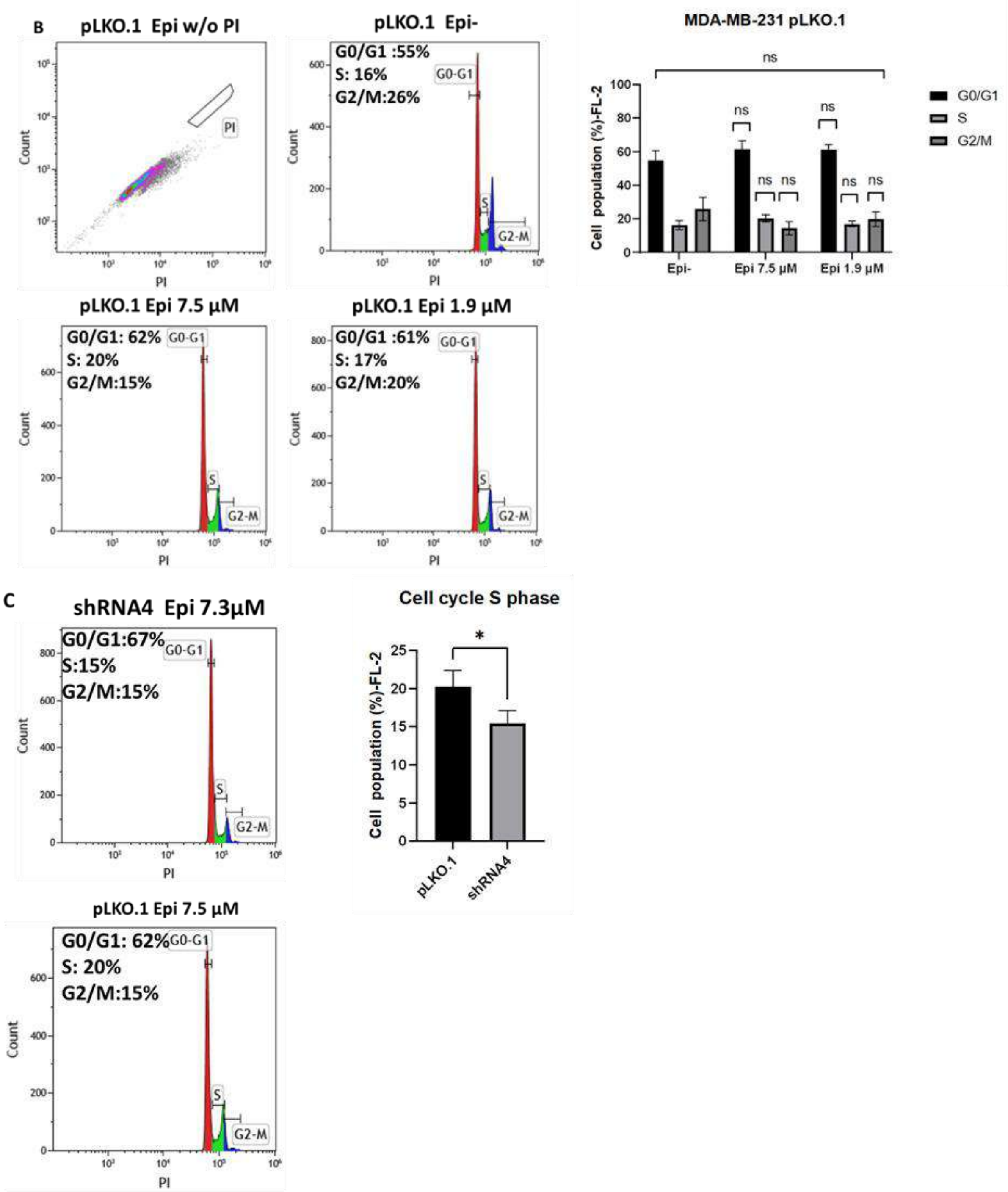


Figure 5.3 - Epirubicin effect on cell cycle distribution in MDA-MB-231 SPAG5 and MDA-MB-231 pLKO.1. Treated MDA-MB-231 SPAG5 shRNA4 (A) and MDA-MB-231 pLKO.1 (B) with two concentrations of Epi for 24h were stained with PI to analyse cell cycle distribution of each cell type by flow cytometry. Analysis of cell number % of each cell phase relative to total phase. Untreated cell was indicated with Epi- and treated with the concentration expressed in µM. One panel (top left) represent the cell stained with the drug without PI to avoid potential interference between the PI and drug wavelength. Each data point represents the mean of three independent experiments, were calculated using 2-way ANOVA with unpaired t-test for post-test for the single phase between the two concentrations compared with the untreated and indicated with asterisks where significant (p-Value <0.05, ns). (C) Bar graph showing the comparison of pLKO.1 and SPAG5-shRNA4 MDA-MB-231 treated with two times high EC_{50} Epi concentration. Each data point represents the mean of three independent experiments calculated using GraphPad Prism 8.4.2 software for unpaired *t-test* and indicated with asterisks were significant (p-Value=*0.0353).

DU145 SPAG5 deficient and pLKO.1 were treated with DOX using two different concentration and the analysis of flow cytometry was the same as processed MDA-MB-231. Cells were incubated with DOX for 48h results are shown in figure 5.4.

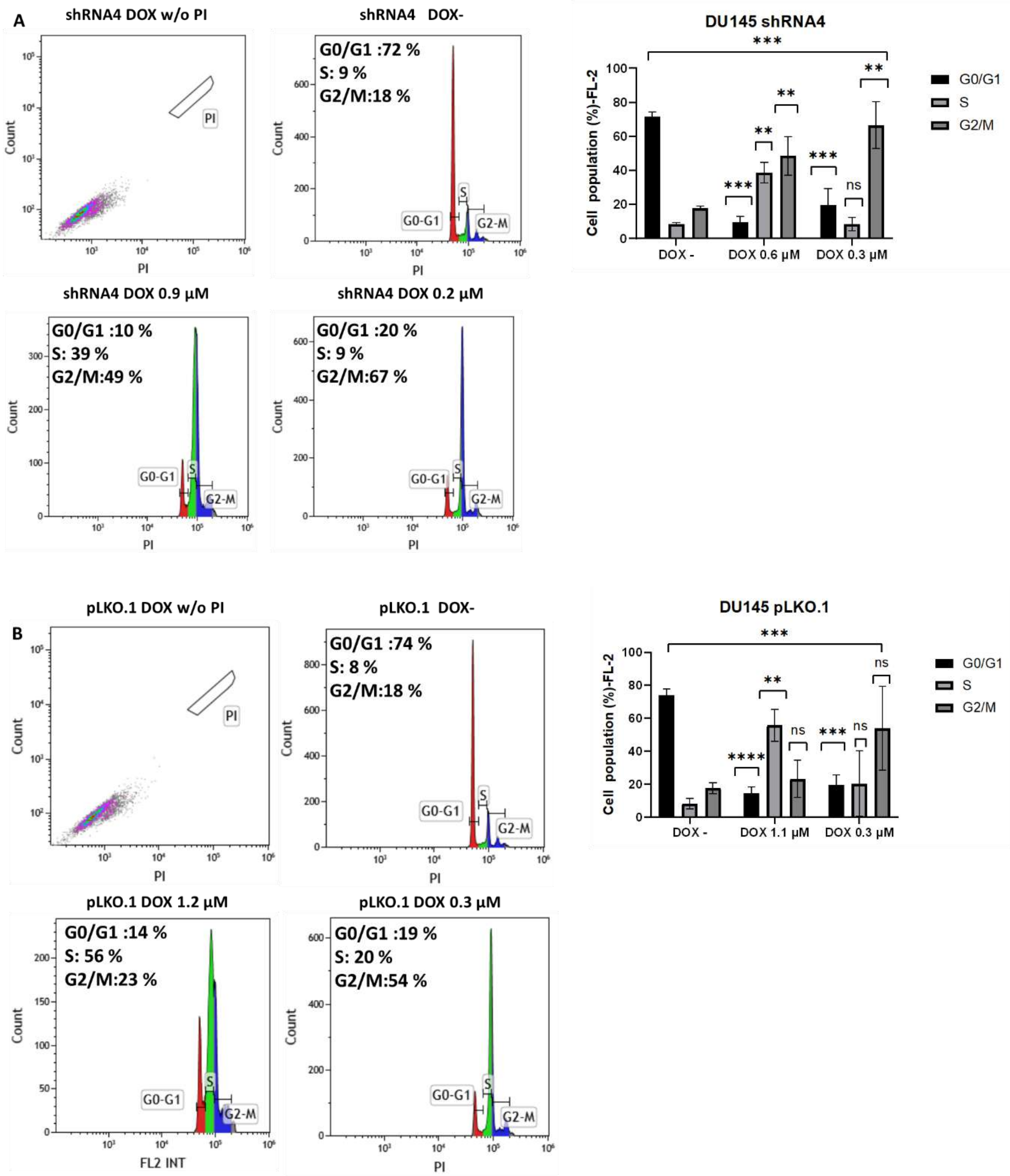


Figure 5.4 - Doxorubicin effect on cell cycle distribution in DU145 SPAG5 deficient and DU145 pLKO.1. Treated DU145 SPAG5 shRNA4 (A) and DU145 pLKO.1(B) with two concentrations of DOX for 48h were stained with PI to analyse cell cycle distribution of each cell type by flow cytometry. Analysis of cell number % of each cell phase relative to total phase. Untreated cell was indicated with DOX- and treated with the concentration expressed in µM. One

panel (top left) represent the cell stained with the drug without PI to avoid potential interference between the PI and drug wavelength. Each data point represents the mean of three independent experiments, were calculated using 2-way ANOVA with unpaired t-test for post-test for the single phase between the two concentrations compared with the untreated, and indicated with asterisks where significant (p-Value <0.05, ns)

Results showed that DU145 SPAG5 deficient treated with 0.9 µM and 0.2 µM were significant decreased in DNA content in the G0/G1 phase when compared with the untreated (72%). Particularly, a statistically significant increase of DNA content in the treated cells DU145 SPAG5 deficient (0.9 µM) was observed when compared with the untreated cells but not in at lower concentration of DOX. This could suggest that 0.3 µM is not able to trigger the arrest in S phase (Fig.5.4 A). In pLKO.1 cells showed a statistical difference in the DNA content when treated with DOX compared with the untreated. The DNA content in S phase increased in pLKO.1 cell population treated with 1.2 µM when compared with the untreated while no statistical difference was shown in DNA content in the S phase between the cell DU145 treated with 0.3 µM with the untreated (Fig5.4 B).

5.3.4 Correlation between SPAG5 expression with commonly upregulated and downregulated genes identified in MDA-MB-231 and DU145 SPAG5 deficient cells with prostate and breast cancer patient

In silico experiments using two data sets from TCGA consortia one from breast cancer and one for prostate cancer, were used to identify potential correlation between the expression of SPAG5 and commonly markers identified in MDA-MB-231 and DU145 SPAG5 silenced cell lines and showed in figure 5.5 and 5.6. The markers used for correlation analysis were sorted based on the statistical differential expression in both data sets. For breast cancer cell MDA-MB-231 SPAG5 silencing among the 47 commonly downregulated markers between the gene and proteome data set four markers (HAPRT1, MAPRE1, MCM3 and NONO) were correlated with SPAG5 expression in patients from TCGA data set. Two markers, HPRT1 and MAPRE1, showed a significant low correlation (Pearson correlation = between ±0.30 and ± 0.49), and one marker, NONO, significant low correlation (Pearson correlation = between ± 0.30 and ± 0.49) with SPAG5 expression in breast cancer patients. Significant moderate positive correlation is observed in MCM3 (Pearson correlation= ± 0.50 and ±0.7) (Fig. 5.5 A). Cross-over upregulated markers between gene and proteome data set identified 44 markers. Results showed four significant upregulated markers (DHCR7, IGFBP7, LSS, and GSN), correlated with SPAG5 expression level in TCGA breast cancer data set in which two markers (IGFBP7 and GSN) show a low negative correlation (Pearson correlation = between ± -0.30 and ± -0.49). Significant moderate positive correlation is observed in DHCR7 while a small positive correlation is seen in LSS markers (Fig 5.5 B). Combined gene and proteome data set from DU145 SPAG5 silenced showed 21 commonly downregulated marker and 3 commonly upregulated markers. Correlation analysis between SPAG5 expression from prostate cancer TCGA data set and 4 most significant commonly downregulated markers were analysed. Results indicated that one marker (FLNC) is low negative correlated (Pearson correlation = between ± -0.30 and ± -0.49) while SCAMP1 showed a small negative correlation with SPAG5 expression in prostate cancer data set. Moderate positive correlation is observed in CCT2 and ATL3 (Pearson correlation = between ±0.30 and ± 0.49) (Fig.5.6 A). Finally, only 3 markers were identified in gene and proteome cross-over in DU145 SPAG5 silencing (TAGLN, CPTA1, and TUBB3). Low positive correlation is observed in TUBB3 (Pearson correlation = between ±0.30 and ± 0.49) and a small positive correlation in CPTA1 (Pearson correlation = ≤ 0.29). Low negative correlation is showed in TAGLN markers (Pearson correlation = between ± -0.30 and ± -0.49) when correlated with SPAG5 prostate data set from TCGA consortia (Fig.5.6 B).

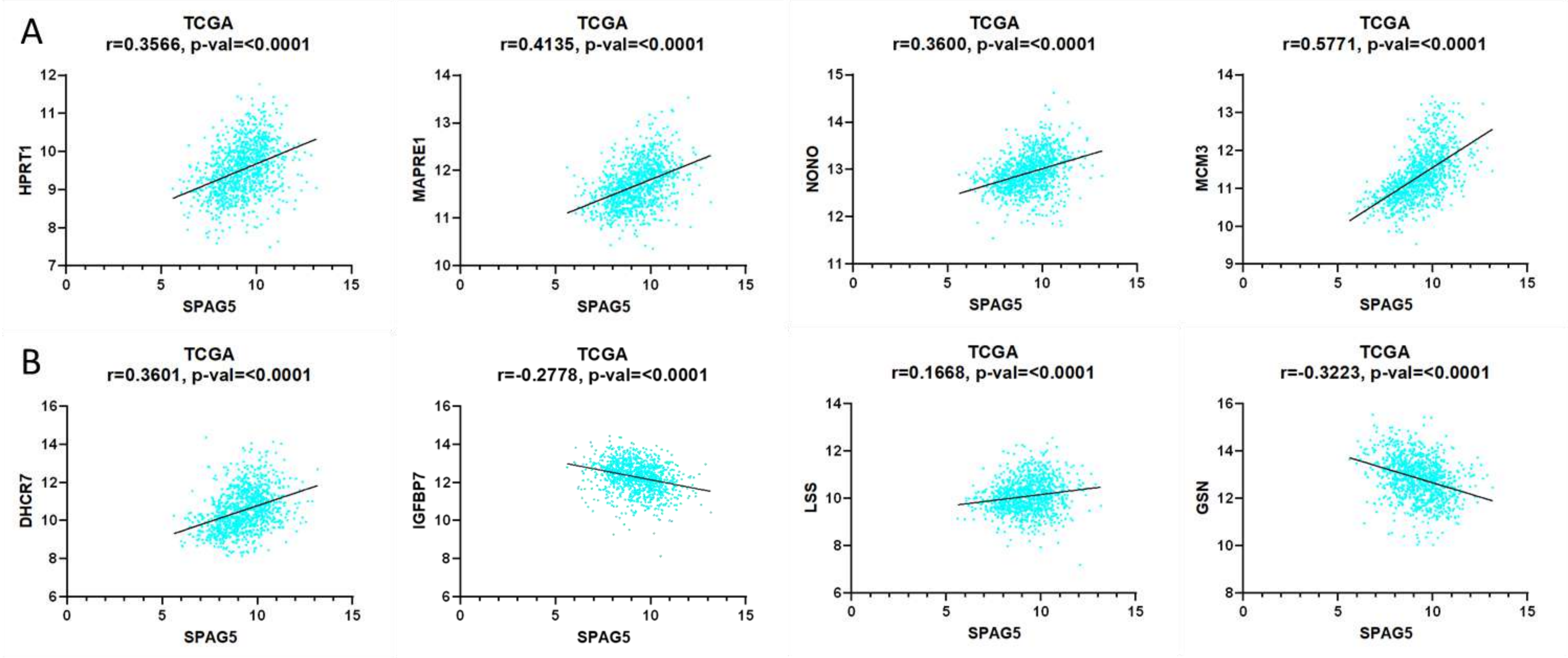


Figure 5.5 – Correlation between SPAG5 expression of commonly markers upregulated and downregulated identified in MDA-MB-231 SPAG5 deficient with *in silico* data. Analysis was obtained by sorting the gene expression SPAG5 from the TCGA breast cancer data set after transformed in Log_2 and correlated with commonly markers identified in MDA-MB-231 SPAG5. (A) Panel shows the correlation between SPAG5 expression from TCGA data set and four of the most statistical downregulated common markers identified in MDA-MB-231. (B) Panel shows the correlation between SPAG5 expression from TCGA data set and four of the most statistical upregulated common markers identified in MDA-MB-231. Correlation analysis using GraphPad Prism 8.4.2 and simple linear regression to find the best line that predict Y (commonly makers identified) with X SPAG5. Correlation is calculated with Pearson's coefficient, r values range between -1 to 1.

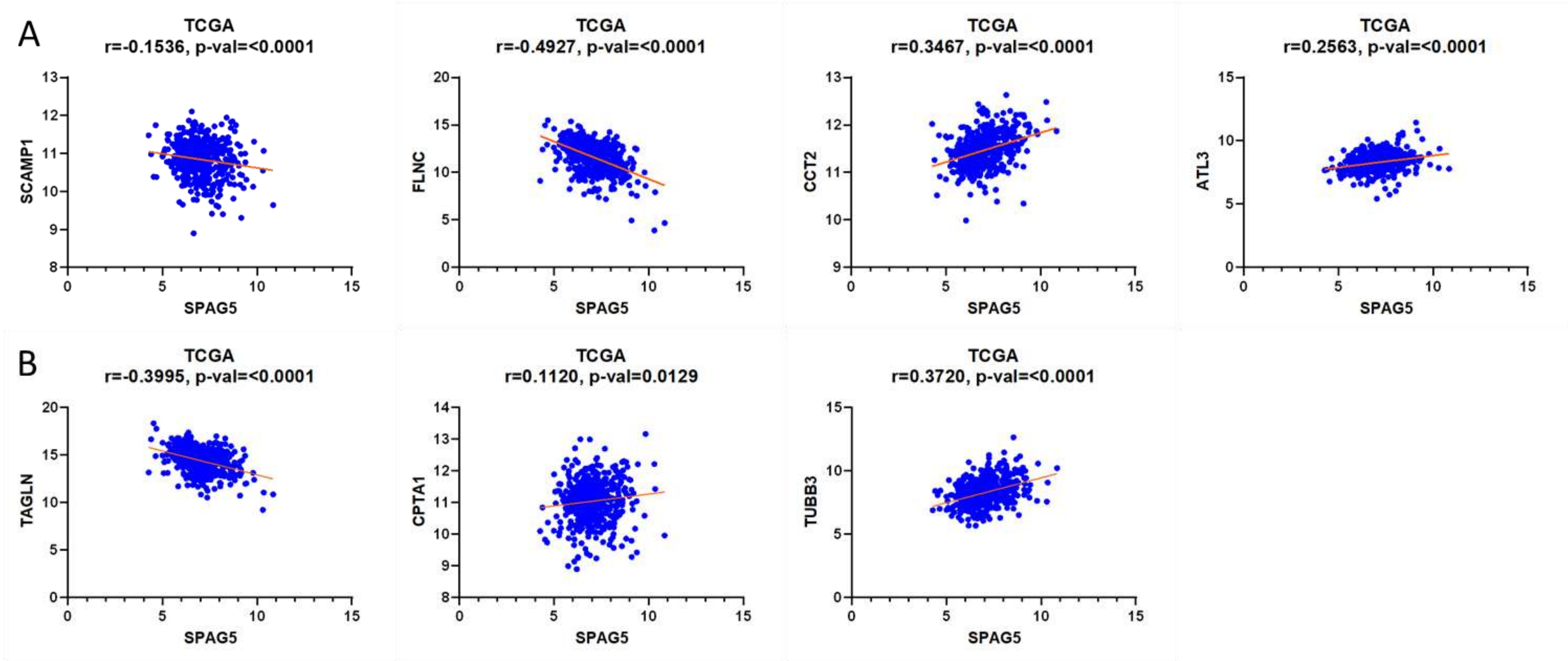


Figure 5.6 – Correlation between SPAG5 expression of commonly markers upregulated and downregulated identified in DU145 SPAG5 deficient with *in silico* data. Analysis was obtained by sorting the gene expression SPAG5 from the TCGA prostate cancer data set after transformed in Log2 and correlated with commonly markers identified in DU145. (A) Panel shows the correlation between SPAG5 expression from TCGA data set and four of the most statistical downregulated common markers identified in DU145. (B) Panel shows the correlation between SPAG5 expression from TCGA data set and four of the most statistical upregulated common markers identified in DU145. Correlation analysis using GraphPad Prism 8.4.2 and simple linear regression to find the best line that predict Y (commonly makers identified) with X SPAG5. Correlation is calculated with Pearson's coefficient, r values range between -1 to 1.

5.4 Discussion

Sperm-associated antigen (SPAG5) attracted the attention of researchers due to its involvement in the cell cycle progression. It is involved in the maintenance of sister chromatid's cohesion and centromere integrity during anaphase of the mitosis process. It also demonstrated the role of SPAG5 in cancer progression in different types of cancers, including breast, prostate cancer. In ovarian cancer is correlated with poor survival and promoting proliferation (Zhang, M. et al., 2020).

Due to its role in the cell cycle, it was hypothesised that its inhibition could affect the stability of chromosomes and make them more sensitive to chemotherapy drugs. Previous studies on breast cancer demonstrated that amplification of SPAG5 locus (Ch17q11.2) represents 10-20% of all breast cancers and that the high expression of mRNA transcript and protein represents the independent predictors of chemotherapy. This characteristic is due to the position of the chromosome 17 (CEP17) amplification in the centromeric region, which represents a marker of chromosomal instability and, therefore, is susceptible to the anthracyclines treatment.

A follow-up study published in 2020 aimed to link the gene expression level of SPAG5 and protein with treatment response in estrogen receptor-positive breast cancer (Abdel-Fatah et al., 2016b). These studies demonstrated that a combination of low levels of SPAG5 and anthracycline chemotherapy, was associated with a prolonged 5-year, relapse-free survival in patients with no lymph node involvement (Abdel-Fatah et al., 2020b). Although the role of SPAG5 chemosensitivity has been demonstrated in the breast cancer, its role in prostate cancer remains unclear.

In prostate cancer, knocking down SPAG5 inhibits proliferation, invasion and metastasis and is confirmed in vivo using a small non-coding RNA miR-539 (Zhang, Hongtuan, et al. 2016). Given these preliminary observations in patients and in few cells line models, this chapter aimed to understand the role of SPAG5 in detail using two genetically modified cell lines. This study also investigated the effect of two chemotherapeutic agents Epirubicin and Doxorubicin in breast cancer cell lines, MDA-MB-231 prostate cancer cell line DU145 respectively. Both chemotherapy drugs are anthracycline , their mechanism of action is mainly through the inhibition of the Topoisomerase II an enzyme involved in DNA replication. Doxorubicin is also used for the treatment of breast cancer. However, the rationale for the choice of Epirubicin over Doxorubicin came from studies conducted on advanced metastatic breast cancer cardiotoxicity effects, that were statistically decreased compared with doxorubicin (Smith, L. A. et al., 2010). The studies demonstrated that with an equal dose ratio of Doxorubicin and Epirubicin, fewer side effects, related to nausea and vomiting, were observed in patients treated with Epirubicin compared to

Doxorubicin (Coukell & Faulds, 1997) (Burnell et al., 2010). In prostate cancer, Doxorubicin represents the most used chemotherapy targeting highly dividing cells, acting on the inhibition of the synthesis of RNA and DNA thus inducing the apoptosis of cells (He, H. et al., 2016).

Despite being the primary drug used for the treatment of prostate cancer, Doxorubicin reported a resistance, including different side effects, cardiac arrest due to its cardiotoxicity; to overcome to those consequences several drug conjugates have been developed such as bovine lactoferrin (bLf) as a carrier protein for delivering Doxorubicin where in DU145 cells, this complex enhanced nuclear Dox retention up to 24h and was proved to be significantly effective (Shankaranarayanan et al., 2016). Previous studies suggested anthracyclines worked better in tumours with higher proliferation and chromosomal instability, while endocrine therapy and taxane (chemotherapeutic agent) work better in chromosomally stable low proliferative breast cancer (Munro et al., 2012) (Miller & Larionov, 2010). In this study, both prostate and breast cancer cell lines were treated with maximum and minimum doses of Epirubicin and doxorubicin and the effects on the cell cycle using a flow cytometry. Results were obtained, and a comparison was made between MDA-MB-231 SPAG5 silencing treated with Epirubicin and without treatment (Fig.5.3 A). Despite no statistical differences, provisional results between the untreated MDA-MB-231 cell population and the treated across the cell cycle phases showed a trend in which Epirubicin suppressed cell proliferation in G0/G1 of the cell cycle. That could suggest that cells check for and repair DNA before progress in the S phase of cell cycle. The arrest of cell cycle in G1 phase was also seen in 4T1 cells used as a control to investigate the effect of Epi and cyclophosphamide (CTX) on microglia (de la Hoz-Camacho et al., 2022). Interesting a similar trend is observed between pLKO.1 empty vector and SPAG5-shRNA4 treated at maximum dose of Epi where a lower DNA content in the S phase of cell cycle is observed (Fig 5.3 C).

Cyclin-Dependent Kinases (CDKs) are involved in cell cycle progression, particularly the link with the cyclin proteins and activated CDKs during a specific phase of the cell cycle (Lim & Kaldis, 2013). From the results obtained, the next step would be to investigate the types of cyclin involved and compare them with data obtained from the RNAseq results described in chapter 3 of this thesis. The MDA-MB-231 SPAG5 cell population transcriptome was mainly analysed for differentially upregulated and downregulated genes (chapter 3, section 3.3.31). Among the most DEGs, CDK18 is statistically upregulated (log2FC 1.58) and is involved in regulating transcription in the G1/S phase. It was demonstrated that a high level of CDK18 was related to a more efficient response to replication stress (Barone et al., 2018), in contradiction with the previous finding from G. Barone, which observed that high levels of CDK18 were associated with reduced survival in ER- but not in ER+ (Barone et al., 2016). Therefore, linked with this finding, it would be interesting to study if a high expression of CDK18 combined with Epirubicin treatment could predict a better response to chemotherapy. For pLKO.1 empty vector, results showed no statistically difference in DNA content across the three-phase untreated and treated samples.

Inhibition of the cell cycle affects mitotic entry; therefore, a mechanism that arrests the cell cycle is an essential strategy in cancer treatment. A study on the effect of PectaSol and Doxorubicin on cell cycle arrest in prostate cancer demonstrated that DOX alone caused an arrest in the G2/M phase after 48 h of treatment in prostate cancer cell line DU145 (20% in treated compared with 8.7 % of the control) (Tehranian et al., 2012). Results showed that DU145 SPAG5 knockdown showed a statistically singificance DNA content between the cell cycle phases. In specific, high amount of DNA content is observed in the G2/M phase of the cell cycle in the treated cell population (49%) with 0.6 µM and 0.3 µM (67%) compared with the untreated (18%) (Fig 5.4 A). Accumulation of cell population in late S (39%) was seen in treated DU145 SPAG5 with 0.6 µM of DOX compared with the untreated (9%). In cell population treated with 0.3 µM of DOX showed the same results. DOX induces G2/M, and another study has shown that DOX can induce arrest in G2/M (White et al., 2007). In agreement with this finding, cell population treated with DOX 0.3 µM were already in progress, this could suggest that DOX-induced a damage to DNA in G1 potentially during the DNA replication phase and recognised at G2 checkpoint, leading to a prolonged block. Similar results were observed in the pLKO.1 empty vector (Fig. 5.4B).

G2/M phase transition represents an important checkpoint for the cell cycle progression, the activation and deactivation of CDC-family proteins and cyclins are responsible for regulating this transition (Taylor, W. R. & Stark, 2001). M-phase inducer phosphatase (CDC25A) protein that works as a mitotic activator and is involved in the progression of a cell from the S phase to the M phase (Graves et al., 2000). A study conducted in prostate cancer demonstrated that the CDC25A is upregulated in several cancers, including prostate cancer (Chiu et al., 2009). Results obtained from RNA sequencing showed that the downregulation of SPAG5 induces a statistical downregulation of CDC25A. Together with the results obtained from the flow cytometry analysis for cell cycle, it would be interesting to study the association of CDC25A and SPAG5 in Doxorubicin treatment and which could predict a better response to chemotherapy.

The cBio Cancer Genomic Portal (cBioPortal) is an effective resource for the analysis of genomic data, containing all TCGA projects and datasets curated from literature (Cerami et al., 2012b). Using cBioPortal features, researchers can expand the analysis and interactively explore genetic alterations across samples, gene pathways and if available also correlate these to clinical outcomes (Wu et al., 2019). A section of this chapter was dedicated to the incorporation of the *in vitro* analysis with *in silico* data from the cancer genomic data sets. Data set from the consortia TCGA of breast and prostate cancer, were correlated with commonly downregulated and upregulated genes/proteins from cross-over between gene and proteome data from MDA-MB-231 and DU145 SPAG5 cell lines. Commonly upregulated and downregulated markers for MDA-MB-231 SPAG5 knockdown were correlated with SPAG5 expression in breast cancer TCGA dataset. Analysing the four most significantly downregulated markers, hypoxanthine Phosphoribosyltransferase 1 (HPRT1) and Microtubule Associated Protein RP/EB Family Member 1 (MAPRE1) a statistically significant positive correlation was observed when analysed with TCGA breast cancer data set therefore, increasing level of SPAG5 corresponds an increasing of HPRT1 levels. Study conducted on breast cancer showed that upregulation of HPRT1 was correlated with poor clinical outcomes and elevated HPRT1 RNA level was found in malignant tissue when compared with normal tissue (J. Sedano et al., 2020). Our results suggest a potential involvement of SPAG5 in HPRT1 expression in triple-negative breast cancer. The minichromosomal maintenance 3 protein (MCM3) is commonly downregulated in gene and proteome data set. Correlation analysis with *in silico* from breast TCGA data set showed a strong positive correlation between MCM3 and SPAG5 increasing level of SPAG5 relates increasing level of MCM3. Our *in vitro* results confirm a positive correlation therefore, decreasing level of SPAG5 corresponds a decreasing level of MCM3. The mechanism by which MCM3 affect the endocrine resistance and its predictive and prognostic function was documented in ER^+ breast cancer, showing that lowering MCM3 expression restored sensitivity to tamoxifen (Løkkegaard et al., 2021a). Our data could suggest that downregulation of SPAG5 not only, could affect key molecules involved in the cell cycle and DNA replication as MCM3 but also be involved in chemotherapy resistance mechanism in triple-negative breast cancer. Positive correlation was observed between SPAG5 expression level and non-POU Domain containing Octamer Binding (NONO) marker indicating that at increasing level of SPAG5, corresponds increased level of NONO. A positive correlation is presented *in vitro* data in which SPAG5 downregulation led to a decreasing expression level of NONO in both data set. Non-POU Domain containing Octamer Binding (NONO) is an important component of ribonucleoprotein bodies, and it was demonstrated to promote chemoresistance in hepatocellular carcinoma (HCC) and is increased in tumour tissues (Kessler et al., 2019). Combined significant upregulated gene data and proteome in MDA-MB-231 SPAG5 silencing identified 44 common markers. Among those markers the four most significant were analysed in correlation with SPAG5 expression from TCGA breast cancer data set. Results showed that two markers DHCR7 and LSS significant negative correlation, therefore the decreasing level of SPAG5 corresponds increased level of DHCR7 and LSS. Our data showed the same trend. The enzyme 7-dehydrocholesterol reductase (DHCR7) is showed to be highly expressed in different types of cancer and identified as potential marker for prognosis (Dai et al., 2022). Several evidence demonstrated that dysregulation of cholesterol is involved in carcinogenesis and cancer development (Mok & Lee, 2020). Common upregulated insulin Like Growth Factor Binding Protein 7 (IGFBP7) and gelsolin (GSN) genes were analysed in correlation with SPAG5 breast cancer TCGA data set. Correlation analysis showed that there is a negative correlation between SPAG5 expression and IGFBP7 and GSN indicating that decreasing level of SPAG5 correlated increasing level of IGFBP7 and GSN. *In silico* data showed the same outcome, SPAG5 downregulation led to an increased level of those markers. In follicular thyroid cancer (FTC) showed that overexpression of IGFBP7 is correlated with significant inhibition of cell proliferation acting as cell cycle repressor (Zhang, L. et al., 2019). Our results could suggest a potential role of SPAG5 in inhibition of IGBP7 in triple-negative breast cancer. Gelsolin (GNS) is an active-binding protein playing an important role in cell motility and it is also able to regulate cell morphology, proliferation, and apoptosis (Chen, Z. et al., 2019). Several publication data demonstrated that GSN was downregulated in different types of cancer including colon carcinoma, gastric cancer, cervical cancer, and ovarian cancer (Noske et al., 2005). Cross-over gene and proteome data identified 21 common downregulated and 3 common upregulated markers in DU145 SPAG5 silencing. Among 21 common downregulated markers the four most significant were correlate with SPAG5 expression from prostate cancer TCGA data set. Correlation analysis on four markers was generated (SCAMP1, FLNC, CCT2, and ATLE). Result showed that secretory Carrier Membrane Protein1 (SCAMP1) and Filamin C (FLNC) showed a negative correlation with SPAG5, therefore decreasing expression of SPAG5 led to increased level of SCAMP1 and FLNC. Our result indicating that downregulation of SPAG5 lead to a downregulation of SCAMP1 and FLNC expression. Study conducted in pancreatic adenocarcinoma (PAAD) showed that SCAMP1 is upregulated compared to normal tissue and the loss of SCAMP1 is linked to an improve overall survival, whilst the downregulation led to a decreasing (Mao et al., 2021). Quantitative proteomic analysis in hepatocellular carcinoma (HCC) showed that filamin C is upregulated and potentially affects invasion and metastasis (Qi, Y. et al., 2016). Positive correlation was observed for Chaperonin Containing TCP1 Subunit 2 (CCT2) and Atlastin GTPase 3 (ATL3) with SPAG5 expression from TCGA data set, indicating that increasing of SPAG5 corresponds an increasing of CCT2 and ATL3. Our data indicating that when SPAG5 is downregulated in prostate cancer DU145 there is a decrease of CCT2 and ATL3. In breast cancer study CCT2 showed to be significant upregulated in HER2-positive group, and it showed that the expression was enriched in cell cycle (Liu, Qiang et al., 2021). Our results showed that CCT is enriched in Rho GTPase pathway which is involved in the different type of biological process included cell adhesion, cytokines and their effector are involved in cell cycle (David et al., 2012). Combined upregulated gene and proteome data set identified 3 markers (TAGLN, CPTA1, and TUBB3). Negative correlation was observed with TAGLN when correlated with SPAG5 expression from prostate TCGA dataset, indicating that an increase of SPAG5 level corresponds to a decrease of TAGLN expression. Our results showed that downregulation of SPAG5 corresponds and increase of TGLN confirming the correlation with *in silico* data. Trangelin (TAGLN) expression is significant low in the prostate tumour tissue when compared to normal tissue and TAGLN suppressor during prostate cancer progression seems to be an important element in the dysregulation of the actin in cytoskeleton (Prasad et al., 2010). Positive correlation was observed in Carnitine palmitoyltransferase 1A (CPTA1) and Tubulin Beta 3 class III (TUBB3) showing that the increasing of SPAG5 induces an increase of CPTA1 and TUBB3. Carnitine palmitoyltransferase 1A (CPTA1) showed to be upregulated in prostate cancer study and responsible of PCa progression (Rios-Colon et al., 2021). Our data showed that downregulation of SPAG5 led an increase of those markers therefore, there is no correlation between *in vitro* and *in silico.*

In conclusion, this chapter investigated the effect of common drug used in breast cancer and prostate cancer. Results obtained for MDA-MB-231 suggested that treatment with Epirubicin affect cell proliferation and activate cell checkpoint to DNA repair before to enter in phase S and M of cell cycle, however, because not statistical differences were observed between MDA-MB-231 and pPLKO.1 except in S phase, it would advised to repeat the experiment perhaps increasing the time of the incubation to 48h and to examine the cell cycle phases. Dox treatment of DU145 SPAG5 silencing cells showed that the high amount of DNA was found in the phase G2/M of cell cycle indicating the activation of checkpoint to allow progression. Taking together those experiments could show a trend of the effect of SPAG5 downregulation combined to drug treatment, however, is necessary to repeat this experiment. Correlation analysis from TCGA data set for breast and prostate cancer patients validate the *In vitro* results obtained from SPAG5 silencing in triple-negative MDA-MB-231 and DU145. Those results can expand our knowledge on SPAG5 role in cancer progression pathways and its involvement in metabolism pathways as responsible in cancer progression.

6. Summary of discussion

Prostate cancer (PCa) represents the second leading cause of death with more than 1.41 million new case and 375,000 deaths are expected worldwide every year (Gandaglia et al., 2021). The radiation therapy represents potentially cure for patients with local recurrence at the earliest sign of rise in serum of prostate-specific antigen (APS) level, after local radical prostatectomy (RP) (Tendulkar et al., 2016). However, in case of metastasis androgen deprivation therapy (ADT) and chemotherapy are the treatment options (Yossepowitch et al., 2014). Cancer stem cells (CSC) represent the model to explain heterogeneity, tumour-initiating capability chemotherapy and metastasis (Eun et al., 2017). CSCs express markers that are strictly tumour or tissue of origin; therefore, there is no universal CSCs marker (Sterlacci et al., 2014). CSC are therapy-resistant and increase the expression of multi-drug resistance (MDR) transported and decrease apoptosis through increasing DNA-repair mechanism (Phi et al., 2018). Beyond the drug resistance, CSCs can also enter the dormant state, leading to tumour recurrence and metastasis (Patel & Chen, 2012). However, because of the multiple heterogenous phenotypes and the no specific markers, there is an urgency to identify potential molecular target for identify CSCs to a given tumour to develop a tailored therapy. Aim of this study was to develop an antibody-drug based therapy against prostate cancer stem cells (PCSCs) using a monoclonal antibody previously generated in our laboratory the human endothelia protein C receptor (hEPCR). Subsequently, silencing of EPCR was sued to investigate the correlation with specific EMT, stemness and cancer stem factors. Finally, the capacity of EPCR to trigger antibody-dependent cellular cytotoxicity (ADCC) was tested in isolated PCSCs using the monoclonal antibody generated in our laboratory.

As presented in chapter 2, the idea of developing a monoclonal antibody drug targeting PCSCs was conceived following a study by Dr Tarik Regad and his team. In previous work, PCSCs were identified using a second-generation lentivirus NANOG promoter, which controlled the expression of the enhanced green fluorescent protein (EGFP) when transducing in specific prostate cancer in this study DU145 (Buczek et al., 2018b). NANOG is an important marker for the maintenance of renewal in embryonic stem cells. As a monoclonal antibody, the endothelial protein C receptor (EPCR) was selected by Dr Tarik Regad and generated by Dr Anna Di Biase as a suitable marker of invasive cells and mAbs-drug therapy. EPCR is a cell surface transmembrane protein with an expected molecular weight of ~26/36-46 KDa, depending on its degree of glycosylation. In a study conducted on breast cancer, EPCR was used as a suitable marker for CSCs. The hypothesis was that EPCR is a convenient marker for a potential mAbs-drug therapy against PCSCs. DU145 PC cells were isolated using NANOG-EGFP lentivirus, generating two populations of DU145 *NANOG positive*, potentially CSCs and DU145 NANOG negative, potentially non-CSCs. The two populations were then stained for EPCR and analysed by flow cytometry. The relative results shown are presented in chapter 2. Data reported not statically differences between cell populations *NANOG positive* and *NANOG negative*, suggesting that EPCR is not a suitable therapeutic target for prostate cancer cells. However, the results shown in the first part of chapter 2, the generated JvGCRC-H61.3 (IgG2b) and JvGCRC-599.5 (IgG2b) mAb against EPCR were used to assess the potential to trigger antibody antibody-dependent cellular cytotoxicity (ADCC) against prostate cancer stem cell expressing EPCR through an *in vitro* ADCC bioluminescent assay which represents a measure of potential killing. Results from one experiment showed that the incubation with the generated monoclonal antibody could activate the effector cells; however, no information on killing efficiency is provided from this assay or other statistical information as the experiment was performed once, and issues were found in the experimental approach. In the specific cell line, PC3 transduce with lentivirus NANOG-EGFP was sorted twice because of the low number of PC3 NANOG positive population needed for the ADCC assay; this could affect the assay and potentially the activation of the effector cell when incubated with the mAbs.

Du145 cells deficient for EPCR previously generated from Dr Anna Di Biase, were used to investigate the expression of stemness markers such as NANOG, SOX2 and OCT4. Besides their essential role in maintaining pluripotency in embryonic stem cells and the role of OCT4 in tumorigenesis is documented (Hatefi et al., 2012). Gene expression analysis showed a decreasing level of OCT4 and NANOG in the DU145 EPCR deficient compared with the control pLKO.1. Conversely, no statistical difference was seen with SOX2. CSCs markers were analysed for CD133, CD44 and CD24. No statistical difference was observed when compared with pLKO.1 empty vector cells. Finally, gene expression analysis on EMT markers such as vimentin and fibronectin mesenchymal markers shown and high expression in the cell population downregulated for EPCR in contrast with the control pLKO. 1 and low expression of the epithelial marker as E-cadherin.

In conclusion, the project aimed to determine whether EPCR was a suitable candidate for a monoclonal antibody therapy drug based on targeting PCSCs results confirmed that EPCR is not differentially expressed in cell populations isolated for NANOG positive and NANOG negative. However, limitation in this experiment should be mentioned. The cell sorter isolated cell population which are *NANOG* - lack of a selectable marker would confuse whether, the cells are *NANOG* - because the plasmid is not integrated in the cells or expressing low level of NANOG. Ultimately, using GEPIA online software mRNA, EPCR was expressed at a high level across the normal and tumour tissues, but meanly EPCR is expressed in endothelial cells suggesting that a potential treatment targeting EPCR would give off target effects. Therefore, no further experiments were conducted on EPCR as it has been deemed as no specific therapeutic target at this stage of the project.

The second part of my PhD course was to investigate another potential pro-oncogenic gene sperm-associate antigen 5 (SPAG5) that has been associated with cell cycle progression and chemosensitivity in prostate cancer and breast cancer.

As mentioned previously for prostate cancer incidence in male, breast cancer worldwide ranks first for incidence and mortality, accounting for 24.5% of all cancer cases and 14.5% of cancer deaths in women (Sung et al., 2021c). Gene inherited in mutant form confer high risk to develop breast cancer and other cancers. It is supposed that approximately 5 % -10 % of the all-breast cancer are believed to be hereditary, BRCA genes are identified to be the mostly linked germ line mutation (Larsen et al., 2014). Mutation in genes *BRCA1* and *BRCA2* are responsible of the high-rate risk of developing breast cancer by up to 85% as also ovarian cancer by up to 54% (Wooster & Weber, 2003) (Wooster et al., 1995).

The aim of this study was to investigate whether SPAG5 pathways were involved in chemotherapy resistance and cancer progression. The rationale for that is chemoresistance represents the limiting factor for achieving a cure in cancer patients. Therefore, early diagnosis and discovery of new potential biomarkers for predicting chemoresistance essential is a valuable tool for cancer treatment. A study conducted in 2016 by Dr Abdel-Fatah in collaboration with Nottingham Trent University and Nottingham University Hospital NHS Trust identified a new prognostic marker which predict chemotherapy sensitivity in HER2 positive breast cancer, the gene sperm-associated antigen 5 (SPAG5). Also, a new recent paper published in 2020 by Dr Fatah in collaboration with Nottingham Trent University demonstrated an association between SPAG5 and treatment response in estrogen receptor-positive breast cancer; where women with tumour, that showed expression of SPAG5 transcript and protein expression level showed a better survival and prolonged 5-years distal relapse-free survival when treated with anthracycline chemotherapy in adjuvant endocrine therapy (Abdel-Fatah et al., 2020c). Little is known about SPAG5 involvement in prostate cancer, with one publication in 2016 showing its role in cancer progression (Zhang, et al. 2016).

To validate the expression of SPAG5 in breast cancer and prostate cancer a gene expression profile was generated on four breast cancer MDA-MB-231, MDA-MB-453, MDA-MB-468 and MCF-7 and two prostate cancer cell lines PC3 and DU145. The results are presented in chapter 3. The data reported that SPAG5 was expressed in all the four-breast cancer cell lines and two prostate cancer cell lines and no statistical difference in the expression was seen between them, similar results were observed in the prostate cancer cell lines. Based on the tumour cell aggressive characteristic, a highly aggressive, poorly differentiated triple negative MDA-MB-231 cell line and metastatic prostate cancer cell line DU145 were chosen for this study. The role of SPAG5 in chemotherapy resistance and cancer progression in breast and prostate cancer a functional study was performed on MDA-MB-231 and DU145 cell lines in which SPAG5 expression was silenced and results presented in chapter 3. SPAG5 silencing in MDA-MB-231 and DU145 cell lines were later sent for RNA sequencing providing an insight into the transcriptome. RNA-sequencing workflow goes from sample preparation to sequencing platform to bioinformatic data analysis offering deep profiling of the transcriptome allowing to elucidate of different physiological and pathological conditions (Wang, Zhong et al., 2009).

Differentially expressed gene for MDA-MB-231 and DU145 was sorted for upregulated and downregulated genes. DEGs in MDA-MB-231 SPAG5 silencing showed upregulating genes involved in the cancer progression and chemoresistance such as the transforming growth factor-1 (TGF-β) in different solid tumours including breast cancer (Palomeras et al., 2019). Despite different inhibitors being developed to target TGF-β signalling results were so far inconsistent suggesting that the role of TGF-β is not fully clear. MDA-MB-231 cells downregulated for SPAG5 showed an upregulation of cathepsin D (CTSD) gene encoding for a lysosomal protein protease involved in the degradation of various substrate including disease-associated proteins like α-synuclein (α-syn) amyloid precursor protein (APP) and tau. Variation in this gene is responsible of neurodegeneration disease such as Alzheimer and Parkinson disease. In breast cancer CTSD is a marker of poor prognosis, and a deficiency of this maker is mammary epithelium blocked the tumour development in CTSD knockout mouse crossed to transgenic *MMTV-PyMT* breast cancer model (Ketterer et al., 2020). Clusterin (CLU) a secreted chaperone, implicated in several pathological state including Alzheimer's disease; however, CLU is also involved in pathways such as cell death and survival and oxidative stress (Foster et al., 2019). In triple-negative breast cancer cell line secreted CLU (sCLU) were positive as also in spontaneous breast cancer mouse strain with triple-negative genotype suggesting a role of CLU in the initiation of triple-negative breast cancer. Our results showed upregulation of CLU in MDA-MB-231 SPAG5 silencing and could therefore, confirm the involvement of CLU in the initiation of triple negative breast cancer (Zhang, D. et al., 2012). Downregulation of SPAG5 also determined upregulation of insulin Growth factor Binding Protein 7 (IGFBP7) gene. This gene (IGFBP7) possesses an IGF-independent activity, and it has shown to be upregulated in cells treated with TGF-β1 and retinoic acid (Swisshelm et al., 1995). IGFBP7 is also involved in a variety of cancer including breast cancer, prostate cancer, and colon cancer (Jin, L. et al., 2020). Interestingly several studies have reported that while IGFBP7 is upregulated in some types of malignancy it is downregulated in others suggesting a potential role as an oncogene or suppressor gene in distinct types of cancer (Chen, Y. et al., 2007). In the top statistically upregulate genes in MDA-MB-231 SPAG5 silencing *KRT19* gene was present. This gene encodes for protein member of the keratin family responsible for the structural integrity of epithelial cells (Coulombe & Wong, 2004). In *KRT19*-knock out mice is responsible for skeletal myopathy and in breast cancer *KRT19* regulates breast cancer properties through the activation of protein kinase B (AKT) pathway (Chen, Y. et al., 2007) (Ju et al., 2013). Study conducted on triple-negative breast cancer cell line showed that KRT19 knockdown led to an increase in the proliferation, migration, invasion, drug resistance through upregulation of NOTCH signalling pathway (Saha et al., 2017). In the BC cell line, MDA-MD-231 silenced for SPAG5 transcriptome analyses DEGs upregulated were enriched in the Phosphatidylinositol 3-kinase (PI3K-Akt) signalling pathways. This pathway is well known to be involved in various cancers and the control of the hallmarks of cancer, such as survival, metastasis, and metabolism (Lawrence et al., 2014). PI3K-Akt pathways is also involved in chemoresistance. Aberrant activation of this pathway through gene mutations such as AKT, TSC1 and also phosphate tensin homolog (PTEN) and mammalian target rapamycin (mTOR) (Hennessy et al., 2005b).

Transcriptome analysis showed that downregulation of SPAG5 affect genes involved in the mitotic cell cycle and cell division such as the microtubule-associated protein RP/EB Family Member 1 encoded by *MAPRE1* gene (Wen, Y. et al., 2004). This gene is commonly mutated in colorectal cancer and increased level of circulating of MAPRE1 was validated in plasma samples and used as diagnosis and prediagnostic marker for this type of cancer (Taguchi et al., 2015). Chemotherapy is a primary treatment in cancer patients together with surgery however, the molecular mechanism that induces sensitivity and resistance is still unclear. Chemotherapies, generally cause DNA damage resulting in activating the cell cycle which will determine cell cycle arrest/apoptosis. However, cancer progression affects the DNA damage response (DDR) itself leading to mutations that can cause chemoresistance (Bartkova et al., 2006). Results presented in chapter 3 revealed that DEGs showed the statistical significance of downregulated genes enriched in the DNA replication and cell cycle including genes associated with the mitosis process and spindle formation and orientation such as MAPRE1 and cell driver to mitosis process gene CDK1. Interaction between MAPRE1 and CDK1 with SPAG5 was seen in the PPI network. Taken together these results could suggest a potential target for therapeutic strategy. Downregulation of SPAG5 in MDA-MB-231 led to downregulation of interleukin-1 receptor kinase 1 (*IRAK1*). *IRAK1* comprise of a class of serine-threonine kinases recently described as involved in the inflammatory regulation, innate immunity, and metabolic disease (Rhyasen & Starczynowski, 2015). Particularly, IRAK1 as active kinase, play a critical role in the IL-1/TLR signalling pathways implicated in the inflammation and innate immune response (Vollmer et al., 2017). In breast cancer it was reported that IRAK1 is overexpressed, and the inhibition reduces the proliferation and metastasis and is also involved in the chemoresistance in nasopharyngeal carcinoma through the IRAK1-S100A9 axis (Wee et al., 2015) (Liu, Lizhen et al., 2021). Finally, its overexpression of IRAK1 is observed hepatocellular carcinoma, increasing cancer proliferation, stemness and drug resistance (Cheng, B. Y. et al., 2018). KEGG pathways showed that IRAK1 is enriched in nuclear factor of κ-light chain of enhancer-activated N cell (NFκB) pathways and plays an important role in the innate immunity via IL-1β. IRAK1 regulates level of another important cytokine as interleukin 6 (IL-6) (Kang, S. et al., 2015). Study conducted on MCF7 cells showed that IL-1β induced IL-6 production in transglutaminase 2 - expressing MCF7 (TG2-expressing MCF7) cells through NFκB dependent mechanism, and cancer aggressiveness was attenuated by either anti-IL6 or anti-IL2β antibody treatment (Oh et al., 2016). Our results showed IRAK1 and IL-6 in MDA-MB-231 SPAG5 deficient and suggesting beneficial effect not only in inflammatory disease but also cancer. Downregulation of SPAG5 in MDA-MB-231 cell line led to downregulation of genes involved in DNA damage response stimuli as mini-chromosome maintenance complex3 (MCM3) and the Non-POU Domain-Containing Octamer-Binding Protein (NONO). In hepatocellular carcinoma (HCC) study, MCM3 upregulation is correlated with poor prognosis and resistance to radiotherapy through the activation of NF-κB pathway (Yang, Q. et al., 2019). Inhibition of NF-κB in cells overexpressing MCM3, showed a reduction of chemoresistance, however, in estrogen-receptor positive breast cancer, MCM3 causes endocrine resistance affecting the tamoxifen action (Løkkegaard et al., 2021b). Clinicopathological study has correlated high expression of NONO with poor prognosis and contribution of progression of breast cancer and cell cycle promotion (Kim, Seong-Jin et al., 2020). In triple negative breast cancer NONO is highly expressed directly interact with STAT protein affecting the stability and transcription factor thus contributing the oncogenic function and chemoresistance correlated with DNA repair pathway (Kim, Seong-Jin et al., 2020). Our results determined that SPAG5 downregulation reduce the expression of NONO and MCM3 and could potentially suggest a potential extension of the role of SPAG5 as guide for the optimal therapies in both estrogen receptor-positive and triple-negative breast cancer.

Prostate cancer cell line DU145 SPAG5 deficient were analysed for RNA-seq and mass spectrometry analysis. Results of RNA-seq provided in chapter 3 showed that DEGs upregulated genes were enriched in the cell cycle and cancer pathways. Genes such as BUB1 mitotic checkpoint serine/threonine kinase (BUB1*)* involved in the cell cycle checkpoint are upregulated in DU145 SPAG5 knockdown, and Aurora kinase A (AURKA*)* is a critical gene involved in the mitotic process. This data could suggest that the downregulation of SPAG5 can aberrantly affect AURKA, leading to the dysregulation of genes and cancer progression. Significantly upregulation of genes in DU145 SPAG5 such as Myosin Light Chain 9 (MYL9) is described to be a key regulator in tumour progression as also in metastasis as described in the pancreatic ductal adenocarcinoma a very aggressive malignancy (Matsushita et al., 2022). In a study conducted on different types of cancers, including prostate cancer, using public dataset from TCGA, investigating a potential role of MYL9 as oncogenesis, it demonstrated that the expression of MYL9 was significantly associated with the infiltration of cancer-associated fibroblasts (CAFs) (Lv et al., 2022). Interestingly, downregulation of SPAG5 in DU145 led to upregulation of Microfibril Associated Protein 5 (MFAP5) and it has been shown recently to be up regulated in CAF of multiple tumours including prostate cancer and be associated with poor prognosis and used as diagnostic marker in prostate cancer (Leung et al., 2014) (Jia et al., 2011b). Results obtained could potentially indicate that SPAG5 downregulation affect not only genes involved in the cell cycle control but also genes which are up regulated in CAFs, highly involved in cancer progression and increased chemoresistance (Ham et al., 2021), therefore, could be suggested as potential guide for optimal therapy in prostate cancer. Finally, downregulation of SPAG5 in DU145 showed upregulation of genes involved in the development of membrane cytoskeleton organisation such as ANXA6 and ANXA8 (Qi, H. et al., 2015). In a large study it was reported that ANXA8 significantly overexpressed in metastatic castration-resistant prostate squamous cancer (Labrecque et al., 2019b). In contradiction with ANXA6 in which it was demonstrated that reduction of this marker was related to progression of from benign to malignant state in previous model of PCa (Xin et al., 2003b). Our data confirm the overexpression in ANXA8 and the lower expression of ANXA6 in DU145 SPAG5 knockdown.

Transcriptome analysis in DU145 SPAG5 deficient revealed a downregulation of NAMPT protein, in which upregulation is associated with several human malignancies including breast and prostate cancer (Sawicka-Gutaj et al., 2015) (Lucena-Cacace et al., 2017). It was shown that inhibition of NAMPT attenuates the glycolysis at glyceraldehyde 3-phospate dehydrogenase step, leading to ATP depletion, metabolic perturbation and consequently the inhibition of tumour growth (Tan et al., 2013) . Overall, this data could suggest that downregulation of SPAG5 can affect genes involved in glycolysis pathway. Glucose metabolism in cancer cells is characterised from the increased intake of glucose and aerobic glycolysis that guarantee cancer cell survival and therefore could be a potential therapy pathway to target (Ganapathy-Kanniappan & Geschwind, 2013). Downregulation of SPAG5 also led to downregulation of genes involved in apoptosis mechanism such as BCL2L13 but downregulation is also responsible of lower survival in different cancers such as clear cell renal carcinoma (ccRCC) and papillary renal carcinoma (pRCC) and is strongly correlated with *SLC25A4* one of the hub genes involved in the physiological function BCL2L13 in kidney cancer (Meng et al., 2021) (Kim, Jee-Youn et al., 2012). Interestingly, our data showed that SPAG5 knockdown downregulated both genes BCL2L13 and SLC25A4 this could suggest that even in prostate cancer the correlation between these two genes is also present. Downregulation of SPAG5 in DU145 showed a downregulation of SCAMP1 which is resulted in aggressive breast cancer (Vadakekolathu et al., 2018). Finally, DU145 SPAG5 silencing led to downregulation of SESN3 that in prostate cancer change in the SESN3 expression was related to a change in the fatigue during the external beam radiation therapy (Gonzalez et al., 2018).

Finally, downregulated and upregulated genes from the two cell lines were combined for identification of common genes (Fig. 3.18-Fig.3.19). Results showed that 47 genes were found commonly upregulated including, based on the most significant upregulated, COL4A1, MAP2K6, COL5A1, VTN, UACA. Vitronectin an adhesive glycoprotein is strongly upregulated in both cell lines (Reuning, 2011). Patients with amplified vitronectin has shown lower rate survival compared to ones without amplified vitronectin (Bera et al., 2020). In breast cancer study it was shown that serum vitronectin level could be used as early marker for breast cancer survival and the PI3K/AKT axis is responsible for vitronectin expression (Bera et al., 2020). In prostate cancer also VTN was used as a potential marker for the diagnosis and assessment of disease progression and metastasis by detecting it expression in the tissue and blood serum of PCa patients (Niu et al., 2016). Common 70 genes downregulated were expressed in both MDA-MB-231 and DU145 SPAG5 deficient cells. The main significant downregulated included: SCAMP1, MAPRE1, CUL2, LIFR, RAD23A. SCAMP1 is almost equally downregulated in both cell lines that could indicate that in both cell lines low expression of this gene is responsible of increasing in the aggressiveness of cancer. Because not much is present in literature, it would interesting to investigate the expression of this marker in prostate cancer tissue.

Proteomic analysis presented in chapter 4 showed that in KEGG pathways enrichment, most of the DEPs were enriched in the glycolysis/gluconeogenesis pathway in MDA-MB-231 SPAG5 silencing. Alterations of metabolism are involved in chemotherapy resistance, particularly ENO2 involved in the glycolysis process, its overexpression is responsible for cancer proliferation in Nod/Scid mice through the upregulation of GLUT-1, LDH and PMK2 which is increasing glycolysis and resistance to glucocorticoids (Liu, Cheng-cheng et al., 2018). Downregulation of ENO2 is related to a decrease in proliferation and sensitivity restored to glucocorticoids. Results showed in chapter 4 showed downregulation of ENO2 when SPAG5 is silenced suggesting a potential involvement of SPAG5 in the pathway. In DU145 SPAG5 silencing proteomic data the most significant downregulated genes were enriched in the adipogenesis and metabolic process. Interestingly, MS data showed that also in breast cancer glycolysis pathways and metabolism pathways are involved in the downregulation of SPAG5, this could extend the role of SPAG5 not only in the cell cycle process but also in other pathways which are involved in cancer progression and chemoresistance. As done for DEGs obtained from RNAseq data set from MDA-MB-231 and DU145 SPAG5 knockdown were combined for commonly upregulated and downregulated proteins. Results presented 8 common proteins divided in five downregulated (MAPRE1, ZMPSTE24, GBE1, CUL2, NEDD4) and three upregulated (EDIL3, FKBP9, ATP2B4). Interestingly two genes are also commonly present in RNAseq data sets MAPRE1 and CUL2 and it would be interesting to study the expression of this marker using clinical patient tissue. ATPase Plasma Membrane CA^{2+} Transporting 4 is downregulated in DU145 SPAG5 knockdown in both transcriptomic and proteomic data set while in breast cancer MDA-MB-231 SPAG5 knockdown is upregulate in proteomic data result. Based only on the results obtained, in prostate cancer DU145 SPAG5 deficient, ATP2BA is enriched in "cellular response to an organic cyclic compound" terms and therefore related to any process involving a change in state or activity of cells. Study conducted on invasive breast cancer using three data set showed a low mRNA expression level when compared with normal breast tissue (Varga et al., 2018). Our results proposed instead that downregulation of SPAG5 in triple-negative breast cancer cell line MDA-MB-231 ATP2BA is upregulated. Crossing-over the gene and proteome data generated a subset of potential candidates to be used for further investigation (Fig. 4.13-Fig.4.14). Result showed that combined data set from RNA-seq and MS analysis identified 47 common significant downregulated and 44 upregulated genes. In prostate cancer cell line DU145 SPAG5 deficient combination of gene and proteome data sets identified 21 common downregulated and 3 common upregulated markers. Those markers were then combined to investigate whether there were any common markers in both cell line. Only 5 markers are common downregulated when combined gene and proteome data set, in both cell MDA-MB-231 and DU145 SPAG5 silencing (MAPRE1, NEDD4, ZMPSTE24, CUL2, GBE1). Those genes are involved in the cell cycle control as described before, DNA repair mechanism but also in metabolism control, confirming the role of SPAG5 in cell cycle and its downregulation decreases genes involved, but also could suggest a potential role also in the metabolism in cancer progression. Finally, the identification of markers that are expressed in gene and proteome data set in two cancers could be also used for future investigations.

As presented in chapter 5 functional study was performed in MDA-MB-231 SPAG5 silencing incubating the cell with Epirubicin, the most common chemotherapy used for BC, and analysed how cells responded in the cells cycle phases G0/1G1-S-G2/M. After 24h, flow cytometry analysis showed no statistical differences between the three cell cycle phases in the cell population. This could be due to a reduced time of incubation with the drug and potentially increasing to 48h of incubation to observe an effect in the cell cycle. The functional study presented in chapter 5 showed that the DU145 SPAG5 cell population treated with different doses of Doxorubicin drug were analysed for cell cycle. The results showed that cells treated with a low concentration of Doxorubicin affect the G2/M phase of the cell cycle, and increasing the dose also affect the S phase of the cell cycle.

Using the cBioPortal bioinformatic software cross-over gene and proteome markers were identified in commonly upregulated and downregulated in MDA-MB-231 and DU145 SPAG5 silencing. Correlation analysis was generated using TCGA data set for breast cancer and four of the most downregulated and upregulated markers were analysed. Data showed that there is a correlation between the *in vitro* analysis with *in silico* data. The same analysis was obtained for DU145 SPAG5 silencing, and the most downregulated and upregulated markers were analysed. Correlation analysis showed that the in vitro experiment and in silico analysis could be correlated.

In conclusion, the second part of the thesis identified a hub of genes and key pathways associated with the downregulation of SPAG5 by RNA-sequencing and mass spectrometry technology. However, is important to point out some limitations in this study, particularly RNAseq is a simpler more comprehensive analysis but on its own doesn't give all the insight needed, that's why it was interesting to look at protein level which is the product of genes. Nonetheless, the interpretation of MS results is different because in this study we only looked at a small subset of unmodified proteins with a small subset of post translational modification (PTMs) that should be quantified but we didn't perform. Even so, this is a discovery-based experiment we only looked initially what unmodified proteins are changed and if we could correlate to RNAseq. Bioinformatic tools allow to analyse the large amount of information in specific cellular pathways. The use of bioinformatic portal such as cBioPortal allowed to combine the *in vitro* clinical data and potential correlation with the data obtained from RNA-sequencing and mass spectrometry analysis. Breast cancer and prostate cancer are treated with anthracycline in combination with surgery. Because the cancer characteristics are different from individual, also the response to therapy can be different. The anthracycline is aggressive treatment for patients, and they can develop different side effects therefore, it would be important to spare people who do not benefit from this treatment. Our results showed a panel of common upregulated and downregulated genes involved in cancer progression and chemoresistance, so it will be interesting as future work to integrate this with few more experiments perhaps, by validating on at gene expression level and on tissues from clinical patients. This could lead to a develop to a prognostic or predict biomarker in response to chemotherapy that will give information to guide the choice of treatment and achieve a better patients outcome.

References

ABDEL-FATAH, T.M., AGARWAL, D., LIU, D., RUSSELL, R., RUEDA, O.M., LIU, K., XU, B., MOSELEY, P.M., GREEN, A.R. AND POCKLEY, A.G., 2016a. SPAG5 as a prognostic biomarker and chemotherapy sensitivity predictor in breast cancer: a retrospective, integrated genomic, transcriptomic, and protein analysis. The Lancet Oncology, **17** (7), 1004-1018.

ABDEL-FATAH, T.M., BALL, G.R., THANGAVELU, P.U., REID, L.E., REED, A.E.M., SAUNUS, J.M., DUIJF, P.H., SIMPSON, P.T., LAKHANI, S.R. AND PONGOR, L., 2020a. Association of Sperm-Associated Antigen 5 and Treatment Response in Patients With Estrogen Receptor–Positive Breast Cancer. JAMA Network Open, **3** (7), e209486.

ABDULKHALEQ, L.A., ASSI, M.A., ABDULLAH, R., ZAMRI-SAAD, M., TAUFIQ-YAP, Y.H. and HEZMEE, M., 2018. The crucial roles of inflammatory mediators in inflammation: A review. Veterinary World, **11** (5), 627.

ACS, G., LAWTON, T.J., REBBECK, T.R., LIVOLSI, V.A. AND ZHANG, P.J., 2001. Differential expression of E-cadherin in lobular and ductal neoplasms of the breast and its biologic and diagnostic implications. American Journal of Clinical Pathology, **115** (1), 85-98.

ADAMO, B., AND ANDERS, C.K., 2011. Stratifying triple-negative breast cancer: which definition (s) to use? Breast Cancer Research, **13** (2), 1-3.

ADJIRI, A., 2016. Identifying and targeting the cause of cancer is needed to cure cancer. Oncology and Therapy, **4** (1), 17-33.

AHN, B.Y., ELWI, A.N., LEE, B., TRINH, D.L., KLIMOWICZ, A.C., YAU, A., CHAN, J.A., MAGLIOCCO, A. AND KIM, S., 2010. Genetic Screen Identifies Insulin-like Growth Factor Binding Protein 5 as a Modulator of Tamoxifen Resistance in Breast CancerIGFBP5 Reverses Tamoxifen Resistance. Cancer Research, **70** (8), 3013-3019.

AKRAM, M., IQBAL, M., DANIYAL, M. and KHAN, A.U., 2017b. Awareness and current knowledge of breast cancer. Biological Research, **50** (1), 1-23.

ALBERTS B, JOHNSON A, LEWIS J, ET AL., 2002. Molecular Biology of the Cell. 4th edition. New York: Garland Science.

AL-HAJJ, M., WICHA, M.S., BENITO-HERNANDEZ, A., MORRISON, S.J. AND CLARKE, M.F., 2003A. Prospective identification of tumorigenic breast cancer cells. Proceedings of the National Academy of Sciences, **100** (7), 3983-3988.

ALKABBAN FM, F.T., 2022. Breast Cancer [online]. Available at: https://www.ncbi.nlm.nih.gov/books/NBK482286/ Accessed: [November, 2022].

AMARO FILHO, S.M., NUOVO, G.J., CUNHA, C.B., DE OLIVEIRA RAMOS PEREIRA, LUIZA, OLIVEIRA-SILVA, M., RUSSOMANO, F., PIRES, A. AND NICOL, A.F., 2014. Correlation of MCM2 detection with stage and virology of cervical cancer. The International Journal of Biological Markers, **29** (4), 363-371.

AMERICAN CANCER SOCIETY, 2021. SURVIVAL RATES FOR PROSTATE CANCER [online]. Available at: https://www.cancer.org/cancer/prostate-cancer/detection-diagnosis-staging/survival-rates.Accessed: [August, 2022]

AMERICAN CANCER SOCIETY, 2022. MAMMOGRAMS BASICS [online].Available at https://www.cancer.org/cancer/breast-cancer/screening-tests-and-earlydetection/mammograms/mammogram-basics. Accessed: [November, 2022].

AMERICAN CANCER SOCIETY., 2021. Survival Rates for Prostate Cancer [online]. Available at: https://www.cancer.org/cancer/prostate-cancer/detection-diagnosis-staging/survival-rates.html. Accessed: [December,2021]

AMERICAN COLLEGE OF RADIOLOGY, 2018.APPROPRIATENESS CRITERIA [online]. Available at: https://acsearch.acr.org/list [November, 2022].

AMODIO, N., SCRIMA, M., PALAIA, L., SALMAN, A.N., QUINTIERO, A., FRANCO, R., BOTTI, G., PIROZZI, P., ROCCO, G. AND DE ROSA, N., 2010. Oncogenic role of the E3 ubiquitin ligase NEDD4-1, a PTEN negative regulator, in non-small-cell lung carcinomas. The American Journal of Pathology, **177** (5), 2622-2634.

ANDRIANI, F., BERTOLINI, G., FACCHINETTI, F., BALDOLI, E., MORO, M., CASALINI, P., CASERINI, R., MILIONE, M., LEONE, G. AND PELOSI, G., 2016. Conversion to stem-cell state in response to microenvironmental cues is regulated by balance between epithelial and mesenchymal features in lung cancer cells. Molecular Oncology, **10** (2), 253-271.

ANNA GASINSKA, 2016. THE CONTRIBUTION OF WOMEN TO RADIOBIOLOGY: MARIE CURIE AND BEYOND. Rep Pract Oncol Radiother, , 250-258.

ANTOCH, G., SAOUDI, N., KUEHL, H., DAHMEN, G., MUELLER, S.P., BEYER, T., BOCKISCH, A., DEBATIN, J.F. and FREUDENBERG, L.S., 2004. Accuracy of whole-body dual-modality fluorine-18–2-fluoro-2-deoxy-D-glucose positron emission tomography and computed tomography (FDG-PET/CT) for tumor staging in solid tumors: comparison with CT and PET. Journal of Clinical Oncology, **22** (21), 4357-4368.

ANTÓN, I., MOLINA, E., LUIS-RAVELO, D., ZANDUETA, C., VALENCIA, K., ORMAZABAL, C., MARTÍNEZ-CANARIAS, S., PERURENA, N., PAJARES, M.J. AND AGORRETA, J., 2012. Receptor of activated protein C promotes metastasis and correlates with clinical outcome in lung adenocarcinoma. American Journal of Respiratory and Critical Care Medicine, **186** (1), 96-105.

ARAI, M., KONDOH, N., IMAZEKI, N., HADA, A., HATSUSE, K., MATSUBARA, O. AND YAMAMOTO, M., 2009. The knockdown of endogenous replication factor C4 decreases the growth and enhances the chemosensitivity of hepatocellular carcinoma cells. Liver International, **29** (1), 55-62.

ARMSTRONG, A.J., MARENGO, M.S., OLTEAN, S., KEMENY, G., BITTING, R.L., TURNBULL, J.D., HEROLD, C.I., MARCOM, P.K., GEORGE, D.J. AND GARCIA-BLANCO, M.A., 2011. Circulating Tumor Cells from Patients with Advanced Prostate and Breast Cancer Display Both Epithelial and Mesenchymal MarkersEpithelial/Mesenchymal Markers on Circulating Tumor Cells. Molecular Cancer Research, **9** (8), 997-1007.

ARNHARD, K., GOTTSCHALL, A., PITTERL, F. AND OBERACHER, H., 2015. Applying 'Sequential Windowed Acquisition of All Theoretical Fragment Ion Mass Spectra'(SWATH) for systematic toxicological analysis with liquid chromatography-high-resolution tandem mass spectrometry. Analytical and Bioanalytical Chemistry, **407** (2), 405-414.

ASGHAR, U., WITKIEWICZ, A.K., TURNER, N.C. AND KNUDSEN, E.S., 2015. The history and future of targeting cyclin-dependent kinases in cancer therapy. Nature Reviews Drug Discovery, **14** (2), 130-146.

AUDIA, J.E., AND CAMPBELL, R.M., 2016. Histone modifications and cancer. Cold Spring Harbor Perspectives in Biology, **8** (4), a019521.

BARBATO, L., BOCCHETTI, M., DI BIASE, A. AND REGAD, T., 2019. Cancer stem cells and targeting strategies. Cells, **8** (8), 926.

BARFELD, S.J., EAST, P., ZUBER, V. AND MILLS, I.G., 2014. Meta-analysis of prostate cancer gene expression data identifies a novel discriminatory signature enriched for glycosylating enzymes. BMC Medical Genomics, **7** (1), 1-26.

BARONE, G., ARORA, A., GANESH, A., ABDEL-FATAH, T., MOSELEY, P., ALI, R., CHAN, S.Y., SAVVA, C., SCHIAVONE, K. AND CARMELL, N., 2018. The relationship of CDK18 expression in breast cancer to clinicopathological parameters and therapeutic response. Oncotarget, **9** (50), 29508.

BARTKOVA, J., REZAEI, N., LIONTOS, M., KARAKAIDOS, P., KLETSAS, D., ISSAEVA, N., VASSILIOU, L.F., KOLETTAS, E., NIFOROU, K. AND ZOUMPOURLIS, V.C., 2006. Oncogene-induced senescence is part of the tumorigenesis barrier imposed by DNA damage checkpoints. Nature, **444** (7119), 633-637.

BASU, A., AND HALDAR, S., 1998. The relationship between BcI2, Bax and p53: consequences for cell cycle progression and cell death. Molecular Human Reproduction, **4** (12), 1099-1109.

BEAULIEU, L.M., AND CHURCH, F.C., 2007. Activated protein C promotes breast cancer cell migration through interactions with EPCR and PAR-1. Experimental Cell Research, **313** (4), 677-687.

BECKER, M., SOMMER, A., KRÄTZSCHMAR, J.R., SEIDEL, H., POHLENZ, H. AND FICHTNER, I., 2005. Distinct gene expression patterns in a tamoxifen-sensitive human mammary carcinoma xenograft and its tamoxifen-resistant subline MaCa 3366/TAM. Molecular Cancer Therapeutics, **4** (1), 151-170.

BEHESHTI, M., IMAMOVIC, L., BROINGER, G., VALI, R., WALDENBERGER, P., STOIBER, F., NADER, M., GRUY, B., JANETSCHEK, G. and LANGSTEGER, W., 2010. 18F choline PET/CT in the preoperative staging of prostate cancer in patients with intermediate or high risk of extracapsular disease: a prospective study of 130 patients. Radiology, **254** (3), 925-933.

BEHRANVAND, N., NASRI, F., ZOLFAGHARI EMAMEH, R., KHANI, P., HOSSEINI, A., GARSSEN, J. AND FALAK, R., 2021. Chemotherapy: a double-edged sword in cancer treatment. Cancer Immunology, Immunotherapy, , 1-20.

BELOCHITSKI, O., ARIAD, S., SHANY, S., FRIDMAN, V. AND GAVRILOV, V., 2007. Efficient dual treatment of the hormone-refractory prostate cancer cell line DU145 with cetuximab and 1, 25-dihydroxyvitamin D3. In Vivo, **21** (2), 371-376.

BENGTSSON, U., ANGERVALL, L., EKMAN, H. and LEHMANN, L., 1968. Transitional cell tumors of the renal pelvis in analgesic abusers. Scandinavian Journal of Urology and Nephrology, **2** (3), 145-150.

BEZUHLY, M., CULLEN, R., ESMON, C.T., MORRIS, S.F., WEST, K.A., JOHNSTON, B. and LIWSKI, R.S., 2009. Role of activated protein C and its receptor in inhibition of tumor metastasis. Blood, the Journal of the American Society of Hematology, **113** (14), 3371-3374.

BIRCH, J., AND GIL, J., 2020. Senescence and the SASP: many therapeutic avenues. Genes & Development, **34** (23-24), 1565-1576.

BLASCO, M.A., 2005. Telomeres and human disease: ageing, cancer and beyond. Nature Reviews Genetics, **6** (8), 611-622.

BOESL, U., 2017. Time-of-flight mass spectrometry: introduction to the basics. Mass Spectrometry Reviews, **36** (1), 86-109.

BOIANI, M., AND SCHÖLER, H.R., 2005. Regulatory networks in embryo-derived pluripotent stem cells. Nature Reviews Molecular Cell Biology**, 6** (11), 872-881.

BONACCI, T., AND EMANUELE, M.J., 2020. Dissenting degradation: Deubiquitinases in cell cycle and cancer. In: Seminars in cancer biology, Elsevier, pp. 145-158.

BOWMAN, G.D., O'DONNELL, M. AND KURIYAN, J., 2004. Structural analysis of a eukaryotic sliding DNA clamp–clamp loader complex. Nature, **429** (6993), 724-730.

BRAWER, M.K., 2005. Prostatic intraepithelial neoplasia: an overview. Reviews in Urology, **7** (Suppl 3), S11.

BRAWER, M.K., 2006. Hormonal therapy for prostate cancer. Reviews in Urology, **8** (Suppl 2), S35.

BRAY, F., FERLAY, J., SOERJOMATARAM, I., SIEGEL, R.L., TORRE, L.A. and JEMAL, A., 2018. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: A Cancer Journal for Clinicians, **68** (6), 394-424.

BROUCKAERT, O., SCHONEVELD, A., TRUYERS, C., KELLEN, E., VAN ONGEVAL, C., VERGOTE, I., MOERMAN, P., FLORIS, G., WILDIERS, H. AND CHRISTIAENS, M.R., 2013. Breast cancer phenotype, nodal status and palpability may be useful in the detection of overdiagnosed screening-detected breast cancers. Annals of Oncology, **24** (7), 1847-1852.

BRUGNOLI, F., GRASSILLI, S., AL-QASSAB, Y., CAPITANI, S. and BERTAGNOLO, V., 2019. CD133 in breast cancer cells: more than a stem cell marker. Journal of Oncology, 2019.

BUCZEK, M.E., MILES, A.K., GREEN, W., JOHNSON, C., BOOCOCK, D.J., POCKLEY, A.G., REES, R.C., HULMAN, G., VAN SCHALKWYK, G. and PARKINSON, R., 2016. Cytoplasmic PML promotes TGF-β-associated epithelial–mesenchymal transition and invasion in prostate cancer. Oncogene**, 35** (26), 3465-3475.

BUCZEK, M.E., REEDER, S.P. AND REGAD, T., 2018, Identification and isolation of cancer stem cells using NANOG-EGFP reporter system. In: Identification and isolation of cancer stem cells using NANOG-EGFP reporter system. Cancer Stem Cells. Springer, 2018, pp. 139-148.

BULL, J.N., LEE, J.W. AND VALLANCE, C., 2013. Quantification of ions with identical mass-to-charge (m/z) ratios by velocity-map imaging mass spectrometry. Physical Chemistry Chemical Physics, **15** (33), 13796-13800.

BURKE, J.R., DESHONG, A.J., PELTON, J.G. and RUBIN, S.M., 2010. Phosphorylation-induced conformational changes in the retinoblastoma protein inhibit E2F transactivation domain binding. Journal of Biological Chemistry, **285** (21), 16286-16293.

BURNELL, M., LEVINE, M.N., CHAPMAN, J.W., BRAMWELL, V., GELMON, K., WALLEY, B., VANDENBERG, T., CHALCHAL, H., ALBAIN, K.S. AND PEREZ, E.A., 2010. Cyclophosphamide, epirubicin, and fluorouracil versus dose-dense epirubicin and cyclophosphamide followed by paclitaxel versus doxorubicin and cyclophosphamide followed by paclitaxel in node-positive or high-risk node-negative breast cancer. Journal of Clinical Oncology, **28** (1), 77.

BUSSO, N., KARABABA, M., NOBILE, M., ROLAZ, A., VAN GOOL, F., GALLI, M., LEO, O., SO, A. AND DE SMEDT, T., 2008. Pharmacological inhibition of nicotinamide phosphoribosyltransferase/visfatin enzymatic activity identifies a new inflammatory pathway linked to NAD. PloS One, **3** (5), e2267.

CABARCAS, S.M., MATHEWS, L.A. and FARRAR, W.L., 2011. The cancer stem cell niche—there goes the neighborhood? International Journal of Cancer, **129** (10), 2315-2327.

CAÑAS, B., LÓPEZ-FERRER, D., RAMOS-FERNÁNDEZ, A., CAMAFEITA, E. AND CALVO, E., 2006. Mass spectrometry technologies for proteomics. Briefings in Functional Genomics, **4** (4), 295-320.

CANCER RESEARCH, U.K., 2021. BREAST CANCER STATISTICS [ONLINE]. Available at: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer#heading-Zero.

CANCER RESEARCH, U.K., 2021. Breast cancer statistics [online]. Available at: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer#heading-Zero. Accessed [December,2021]

CANCER, WORLD HEALTH ORGANIZATION (WHO), 2022. Cancer [online] Available at: https://www.who.int/news-room/fact-sheets/detail/cancer. [Accessed: March 2022].

CANO, A., PÉREZ-MORENO, M.A., RODRIGO, I., LOCASCIO, A., BLANCO, M.J., DEL BARRIO, M.G., PORTILLO, F. AND NIETO, M.A., 2000. The transcription factor snail controls epithelial–mesenchymal transitions by repressing E-cadherin expression. Nature Cell Biology, **2** (2), 76-83.

CANU, V., DONZELLI, S., SACCONI, A., LO SARDO, F., PULITO, C., BOSSEL, N., DI BENEDETTO, A., MUTI, P., BOTTI, C. AND DOMANY, E., 2021. Aberrant transcriptional and post-transcriptional regulation of SPAG5, a YAP-TAZ-TEAD downstream effector, fuels breast cancer cell proliferation. Cell Death & Differentiation, **28** (5), 1493-1511.

CAPLAN, L., 2014. Delay in breast cancer: implications for stage at diagnosis and survival. Frontiers in Public Health, **2**, 87.

CAREY, L., WINER, E., VIALE, G., CAMERON, D. AND GIANNI, L., 2010. Triple-negative breast cancer: disease entity or title of convenience? Nature Reviews Clinical Oncology, **7** (12), 683-692.

CASTELLINO, F.J., AND PLOPLIS, V.A., 2009. The protein C pathway and pathologic processes. Journal of Thrombosis and Haemostasis, **7**, 140-145.

CASTLE, A., AND CASTLE, D., 2005. Ubiquitously expressed secretory carrier membrane proteins (SCAMPs) 1-4 mark different pathways and exhibit limited constitutive trafficking to and from the cell surface. Journal of Cell Science, **118** (16), 3769-3780.

CASTRO, E., and EELES, R., 2012. The role of BRCA1 and BRCA2 in prostate cancer. Asian Journal of Andrology, 14 (3), 409.

CAVALIERI, E., CHAKRAVARTI, D., GUTTENPLAN, J., HART, E., INGLE, J., JANKOWIAK, R., MUTI, P., ROGAN, E., RUSSO, J. and SANTEN, R., 2006. Catechol estrogen quinones as initiators of breast and other human cancers: implications for biomarkers of susceptibility and cancer prevention. Biochimica Et Biophysica Acta (BBA)-Reviews on Cancer, **1766** (1), 63-78.

CAVALLARO, U., and CHRISTOFORI, G., 2004. Cell adhesion and signalling by cadherins and Ig-CAMs in cancer. Nature Reviews Cancer, **4** (2), 118-132.

CERAMI, E., GAO, J., DOGRUSOZ, U., GROSS, B.E., SUMER, S.O., AKSOY, B.A., JACOBSEN, A., BYRNE, C.J., HEUER, M.L. AND LARSSON, E., 2012. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discovery, **2** (5), 401-404.

CHAFFER, C.L., MARJANOVIC, N.D., LEE, T., BELL, G., KLEER, C.G., REINHARDT, F., D'ALESSIO, A.C., YOUNG, R.A. and WEINBERG, R.A., 2013. Poised chromatin at the ZEB1 promoter enables breast cancer cell plasticity and enhances tumorigenicity. Cell, **154** (1), 61-74.

CHAMBERS, I., COLBY, D., ROBERTSON, M., NICHOLS, J., LEE, S., TWEEDIE, S. AND SMITH, A., 2003. Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. Cell, **113** (5), 643-655.

CHEEVER, M.A., AND HIGANO, C.S., 2011. PROVENGE (Sipuleucel-T) in Prostate Cancer: The First FDA-Approved Therapeutic Cancer VaccineTherapeutic Prostate Cancer Vaccine. Clinical Cancer Research, **17** (11), 3520-3526.

CHEN, D., SUN, Y., WEI, Y., ZHANG, P., REZAEIAN, A.H., TERUYA-FELDSTEIN, J., GUPTA, S., LIANG, H., LIN, H. AND HUNG, M., 2012. LIFR is a breast cancer metastasis suppressor upstream of the Hippo-YAP pathway and a prognostic marker. Nature Medicine, **18** (10), 1511-1517.

CHEN, H., BAI, F., WANG, M., ZHANG, M., ZHANG, P. and WU, K., 2019. The prognostic significance of co-existence ductal carcinoma in situ in invasive ductal breast cancer: a large population-based study and a matched case-control analysis. Annals of Translational Medicine, **7** (18).

CHEN, W., HO, C., CHANG, Y., CHEN, H., LIN, C., LING, T., YU, S., YUAN, S., LOUISA CHEN, Y. AND LIN, C., 2014. Cancer-associated fibroblasts regulate the plasticity of lung cancer stemness via paracrine signalling. Nature Communications**, 5** (1), 1-17.

CHEN, Y., PACYNA-GENGELBACH, M., YE, F., KNÖSEL, T., LUND, P., DEUTSCHMANN, N., SCHLÜNS, K., KOTB, W., SERS, C. AND YASUMOTO, H., 2007. Insulin-like growth factor binding protein-related protein 1 (IGFBP-rP1) has potential tumour-suppressive activity in human lung cancer. The Journal of Pathology: A Journal of the Pathological Society of Great Britain and Ireland, **211** (4), 431-438.

CHENG, B.Y., LAU, E.Y., LEUNG, H., LEUNG, C.O., HO, N.P., GURUNG, S., CHENG, L.K., LIN, C.H., LO, R.C. AND MA, S., 2018. IRAK1 Augments Cancer Stemness and Drug Resistance via the AP-1/AKR1B10 Signaling Cascade in Hepatocellular CarcinomaIRAK1 Regulates Cancer Stemness and Drug Resistance in HCC. Cancer Research, **78** (9), 2332-2342.

CHENG, T., HSIAO, Y., LIN, C., YU, C.R., HSU, C., CHANG, M., LEE, C., HUANG, C.F., HOWNG, S. and HONG, Y., 2008. Glycogen synthase kinase 3β interacts with and phosphorylates the spindle-associated protein astrin. Journal of Biological Chemistry, **283** (4), 2454-2464.

CHIOU, S., WANG, M., CHOU, Y., CHEN, C., HONG, C., HSIEH, W., CHANG, H., CHEN, Y., LIN, T. AND HSU, H., 2010. Coexpression of Oct4 and Nanog enhances malignancy in lung adenocarcinoma by inducing cancer stem cell–like properties and epithelial–mesenchymal transdifferentiation. Cancer Research, **70** (24), 10433-10444.

CHIU, Y., HAN, H., LEUNG, S.C., YUEN, H., CHAU, C., GUO, Z., QIU, Y., CHAN, K., WANG, X. AND WONG, Y., 2009. CDC25A functions as a novel Ar corepressor in prostate cancer cells. Journal of Molecular Biology, **385** (2), 446-456.

CHOY, B., LALONDE, A., QUE, J., WU, T. AND ZHOU, Z., 2016. MCM4 and MCM7, potential novel proliferation markers, significantly correlated with Ki-67, Bmi1, and cyclin E expression in esophageal adenocarcinoma, squamous cell carcinoma, and precancerous lesions. Human Pathology, **57**, 126-135.

CHU, X., CHEN, X., WAN, Q., ZHENG, Z. AND DU, Q., 2016. Nuclear mitotic apparatus (NuMA) interacts with and regulates astrin at the mitotic spindle. Journal of Biological Chemistry, **291** (38), 20055-20067.

CLANCY, M.V., ZYTYNSKA, S.E., MORITZ, F., WITTING, M., SCHMITT-KOPPLIN, P., WEISSER, W.W. AND SCHNITZLER, J., 2018. Metabotype variation in a field population of tansy plants influences aphid host selection. Plant, Cell & Environment, **41** (12), 2791-2805.

CLARKE, R., LIU, M.C., BOUKER, K.B., GU, Z., LEE, R.Y., ZHU, Y., SKAAR, T.C., GOMEZ, B., O'BRIEN, K. and WANG, Y., 2003. Antiestrogen resistance in breast cancer and the role of estrogen receptor signaling. Oncogene, **22** (47), 7316-7339.

COLLINS, L., ZHU, T., GUO, J., XIAO, Z.J. AND CHEN, C.Y., 2006. Phellinus linteus sensitises apoptosis induced by doxorubicin in prostate cancer. British Journal of Cancer, **95** (3), 282-288.

COMERFORD, S.A., HUANG, Z., DU, X., WANG, Y., CAI, L., WITKIEWICZ, A.K., WALTERS, H., TANTAWY, M.N., FU, A. AND MANNING, H.C., 2014. Acetate dependence of tumors. Cell, **159** (7), 1591-1602.

CORCOS, D., 2012. Unbalanced replication as a major source of genetic instability in cancer cells. American Journal of Blood Research**,** **2** (3), 160.

COUKELL, A.J., AND FAULDS, D., 1997. Epirubicin. Drugs, **53** (3), 453-482.

COULOMBE, P.A., AND WONG, P., 2004. Cytoplasmic intermediate filaments revealed as dynamic and multipurpose scaffolds. Nature Cell Biology, **6** (8), 699-706.

CROTEAU, E., RENAUD, J.M., RICHARD, M.A., RUDDY, T.D., BÉNARD, F. and DEKEMP, R.A., 2016. PET Metabolic Biomarkers for Cancer: Supplementary Issue: Biomarkers and their Essential Role in the Development of Personalised Therapies (A). Biomarkers in Cancer, 8, BIC. S27483.

CSERNI, G., CHMIELIK, E., CSERNI, B. and TOT, T., 2018. The new TNM-based staging of breast cancer. Virchows Archiv, **472** (5), 697-703.

CUZICK, J., POWLES, T., VERONESI, U., FORBES, J., EDWARDS, R., ASHLEY, S. and BOYLE, P., 2003. Overview of the main outcomes in breast-cancer prevention trials. The Lancet, **361** (9354), 296-300.

DANG, L., SHI, C., ZHANG, Q., LIAO, P. AND WANG, Y., 2022. Downregulation of sperm-associated antigen 5 inhibits melanoma progression by regulating forkhead box protein M1/A disintegrin and metalloproteinase 17/NOTCH1 signaling. Bioengineered, **13** (3), 4744-4756.

DARR, H., MAYSHAR, Y. AND BENVENISTY, N., 2006. Overexpression of NANOG in human ES cells enables feeder-free growth while inducing primitive ectoderm features.

DAVIS, J.D., AND LIN, S., 2011. DNA damage and breast cancer. World Journal of Clinical Oncology, **2** (9), 329.

DE BONO, J.S., LOGOTHETIS, C.J., MOLINA, A., FIZAZI, K., NORTH, S., CHU, L., CHI, K.N., JONES, R.J., GOODMAN JR, O.B. and SAAD, F., 2011. Abiraterone and increased survival in metastatic prostate cancer. New England Journal of Medicine, **364** (21), 1995-2005.

DE LA FOUCHARDIERE, C., FLECHON, A. AND DROZ, J.P., 2003. Coagulopathy in prostate cancer. Neth J Med, **61** (11), 347-354.

DE LA HOZ-CAMACHO, R., RIVERA-LAZARÍN, A.L., VÁZQUEZ-GUILLEN, J.M., CABALLERO-HERNÁNDEZ, D., MENDOZA-GAMBOA, E., MARTÍNEZ-TORRES, A.C. AND RODRÍGUEZ-PADILLA, C., 2022. Cyclophosphamide and epirubicin induce high apoptosis in microglia cells while epirubicin provokes DNA damage and microglial activation at sub-lethal concentrations. EXCLI Journal, **21**, 197.

DE WOUWER, M.V., COLLEN, D. AND CONWAY, E.M., 2004. Thrombomodulin-protein C-EPCR system: integrated to regulate coagulation and inflammation. Arteriosclerosis, Thrombosis, and Vascular Biology, **24** (8), 1374-1383.

DE WOUWER, M.V., COLLEN, D. and CONWAY, E.M., 2004. Thrombomodulin-protein C-EPCR system: integrated to regulate coagulation and inflammation. Arteriosclerosis, Thrombosis, and Vascular Biology, **24** (8), 1374-1383.

DEBNATH, M., MALIK, C.P. and BISEN, P.S., 2006. Micropropagation: a tool for the production of high quality plant-based medicines. Current Pharmaceutical Biotechnology, **7** (1), 33-49.

DESAI, M., JOHN, B., EVANS, G. AND EDDY, B., 2015. Prostate cancer: beware of disseminated intravascular coagulation. Case Reports, 2015, bcr2014206814.

DESCOTES, J., 2019. Diagnosis of prostate cancer. Asian Journal of Urology, **6** (2), 129-136.

DI BIASE, A., 2020. Generation of α-Epcr Monoclonal Antibodies and Their Use in the Development of an Antibody-Based Targeting Therapy of Invasive Prostate Cancer. Nottingham Trent University.

DI BIASE, A., MILES, A.K. AND REGAD, T., 2018,. Generation of in vitro model of epithelial mesenchymal transition (EMT) via the expression of a cytoplasmic mutant form of promylocytic leukemia protein (PML). In: Generation of in vitro model of epithelial mesenchymal transition (EMT) via the expression of a cytoplasmic mutant form of promylocytic leukemia protein (PML). Cancer Stem Cells. Springer, 2018, pp. 129-138.

DIANAT-MOGHADAM, H., HEIDARIFARD, M., JAHANBAN-ESFAHLAN, R., PANAHI, Y., HAMISHEHKAR, H., POUREMAMALI, F., RAHBARGHAZI, R. and NOURI, M., 2018. Cancer stem cells-emanated therapy resistance: implications for liposomal drug delivery systems. Journal of Controlled Release, **288,** 62-83.

DIRAT, B., BOCHET, L., DABEK, M., DAVIAUD, D., DAUVILLIER, S., MAJED, B., WANG, Y.Y., MEULLE, A., SALLES, B. and LE GONIDEC, S., 2011. Cancer-associated adipocytes exhibit an activated phenotype and contribute to breast cancer invasion. Cancer Research, **71** (7), 2455-2465.

DORVAL, M., MAUNSELL, E., DESCHÊNES, L. and BRISSON, J., 1998. Type of mastectomy and quality of life for long term breast carcinoma survivors. Cancer: Interdisciplinary International Journal of the American Cancer Society, **83** (10), 2130-2138.

DOWNARD, K.M., 2012. 1912: a titanic year for mass spectrometry. Journal of Mass Spectrometry, **47** (8), 1034-1039.

DUFFY, M.J., O'DONOVAN, N. AND CROWN, J., 2011. Use of molecular markers for predicting therapy response in cancer patients. Cancer Treatment Reviews, **37** (2), 151-159.

EASWARAN, H., TSAI, H. and BAYLIN, S.B., 2014. Cancer epigenetics: tumor heterogeneity, plasticity of stem-like states, and drug resistance. Molecular Cell, **54** (5), 716-727.

EBERL, M.M., SUNGA, A.Y., FARRELL, C.D. and MAHONEY, M.C., 2005. Patients with a family history of cancer: identification and management. The Journal of the American Board of Family Practice, **18** (3), 211-217.

EIRO, N., CID, S., FERNÁNDEZ, B., FRAILE, M., CERNEA, A., SÁNCHEZ, R., ANDICOECHEA, A., DEANDRES GALIANA, E.J., GONZÁLEZ, L.O. AND FERNÁNDEZ-MUÑIZ, Z., 2019. MMP11 expression in intratumoral inflammatory cells in breast cancer. Histopathology, **75** (6), 916-930.

ELDER, E.E., BRANDBERG, Y., BJÖRKLUND, T., RYLANDER, R., LAGERGREN, J., JURELL, G., WICKMAN, M. and SANDELIN, K., 2005. Quality of life and patient satisfaction in breast cancer patients after immediate breast reconstruction: a prospective study. The Breast, **14** (3), 201-208.

ELMORE, S., 2007. Apoptosis: a review of programmed cell death. Toxicologic Pathology, **35** (4), 495-516.

EMILI, A., SHALES, M., MCCRACKEN, S., XIE, W., TUCKER, P.W., KOBAYASHI, R., BLENCOWE, B.J. AND INGLES, C.J., 2002. Splicing and transcription-associated proteins PSF and p54nrb/nonO bind to the RNA polymerase II CTD. Rna, **8** (9), 1102-1111.

EMWAS, A.M., AL-TALLA, Z.A., YANG, Y. AND KHARBATIA, N.M., 2015, GAS CHROMATOGRAPHY–MASS SPECTROMETRY OF BIOFLUIDS AND EXTRACTS. In: Gas chromatography–mass spectrometry of biofluids and extracts. Metabonomics. Springer, 2015, pp. 91-112.

EPSTEIN, J.I., EGEVAD, L., AMIN, M.B., DELAHUNT, B., SRIGLEY, J.R. and HUMPHREY, P.A., 2016. The 2014 International Society of Urological Pathology (ISUP) consensus conference on Gleason grading of prostatic carcinoma. The American Journal of Surgical Pathology, **40** (2), 244-252.

ERLER, J.T., BENNEWITH, K.L., COX, T.R., LANG, G., BIRD, D., KOONG, A., LE, Q. AND GIACCIA, A.J., 2009. Hypoxia-induced lysyl oxidase is a critical mediator of bone marrow cell recruitment to form the premetastatic niche. Cancer Cell, **15** (1), 35-44.

ESMON, C.T., 2009. Basic mechanisms and pathogenesis of venous thrombosis. Blood Reviews, **23** (5), 225-229.

EUN, K., HAM, S.W. AND KIM, H., 2017. Cancer stem cell heterogeneity: origin and new perspectives on CSC targeting. BMB Reports, **50** (3), 117.

FAGET, D.V., REN, Q. AND STEWART, S.A., 2019. Unmasking senescence: context-dependent effects of SASP in cancer. Nature Reviews Cancer, **19** (8), 439-453.

FEITELSON, M.A., ARZUMANYAN, A., KULATHINAL, R.J., BLAIN, S.W., HOLCOMBE, R.F., MAHAJNA, J., MARINO, M., MARTINEZ-CHANTAR, M.L., NAWROTH, R. and SANCHEZ-GARCIA, I., 2015. Sustained proliferation in cancer: Mechanisms and novel therapeutic targets. In: Seminars in cancer biology, Elsevier, pp. S25-S54.

FENG, M., DONG, N., ZHOU, X., MA, L. AND XIANG, R., 2022. Myosin light chain 9 promotes the proliferation, invasion, migration and angiogenesis of colorectal cancer cells by binding to Yes-associated protein 1 and regulating Hippo signaling. Bioengineered, **13** (1), 96-106.

FERNANDES, J.V., COBUCCI, R.N.O., JATOBÁ, C.A.N., DE MEDEIROS FERNANDES, T.A.A., DE AZEVEDO, J.W.V. AND DE ARAÚJO, J.M.G., 2015. The role of the mediators of inflammation in cancer development. Pathology & Oncology Research, **21** (3), 527-534.

FESSLER, E., BOROVSKI, T. and MEDEMA, J.P., 2015. Endothelial cells induce cancer stem cell features in differentiated glioblastoma cells via bFGF. Molecular Cancer, **14** (1), 1-12.

FISHER, B., ANDERSON, S., REDMOND, C.K., WOLMARK, N., WICKERHAM, D.L. and CRONIN, W.M., 1995. Reanalysis and results after 12 years of follow-up in a randomized clinical trial comparing total mastectomy with lumpectomy with or without irradiation in the treatment of breast cancer. New England Journal of Medicine, **333** (22), 1456-1461.

FISLER, D.A., SIKARIA, D., YAVORSKI, J.M., TU, Y.N. AND BLANCK, G., 2018. Elucidating feed forward apoptosis signatures in breast cancer datasets: Higher FOS expression associated with a better outcome. Oncology Letters, **16** (2), 2757-2763.

FORSBURG, S.L., 2004. Eukaryotic MCM proteins: beyond replication initiation. Microbiology and Molecular Biology Reviews, **68** (1), 109-131.

FOSTER, E.M., DANGLA-VALLS, A., LOVESTONE, S., RIBE, E.M. AND BUCKLEY, N.J., 2019. Clusterin in Alzheimer's disease: mechanisms, genetics, and lessons from other pathologies. Frontiers in Neuroscience, **13**, 164.

FOUAD, Y.A., and AANEI, C., 2017. Revisiting the hallmarks of cancer. American Journal of Cancer Research, **7** (5), 1016.

FOULKES, W.D., SMITH, I.E. AND REIS-FILHO, J.S., 2010. Triple-negative breast cancer. New England Journal of Medicine, **363** (20), 1938-1948.

FRANCA-BOTELHO, A.D.C., FERREIRA, M.C., FRANCA, J.L., FRANCA, E.L. and HONORIO-FRANCA, A.C., 2012. Breastfeeding and its relationship with reduction of breast cancer: a review. Asian Pacific Journal of Cancer Prevention, **13** (11), 5327-5332.

FRY, D.W., HARVEY, P.J., KELLER, P.R., ELLIOTT, W.L., MEADE, M., TRACHET, E., ALBASSAM, M., ZHENG, X., LEOPOLD, W.R. AND PRYER, N.K., 2004. Specific inhibition of cyclin-dependent kinase 4/6 by PD 0332991 and associated antitumor activity in human tumor xenografts. Molecular Cancer Therapeutics, **3** (11), 1427-1438.

GALFANO, A., 2009. Editorial Comment on: The European Association of Urology (EAU) Guidelines Methodology: A Critical Evaluation. European Urology, **5** (56), 864.

GALLÍ, M., VAN GOOL, F., RONGVAUX, A., ANDRIS, F. AND LEO, O., 2010. The nicotinamide phosphoribosyltransferase: a molecular link between metabolism, inflammation, and cancer. Cancer Research, **70** (1), 8-11.

GANDAGLIA, G., LENI, R., BRAY, F., FLESHNER, N., FREEDLAND, S.J., KIBEL, A., STATTIN, P., VAN POPPEL, H. AND LA VECCHIA, C., 2021. Epidemiology and prevention of prostate cancer. European Urology Oncology, .

GAO, J., AKSOY, B.A., DOGRUSOZ, U., DRESDNER, G., GROSS, B., SUMER, S.O., SUN, Y., JACOBSEN, A., SINHA, R. AND LARSSON, E., 2013. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Science Signaling**, 6** (269), pl1.

GASCA, J., FLORES, M.L., JIMÉNEZ-GUERRERO, R., SÁEZ, M.E., BARRAGÁN, I., RUÍZ-BORREGO, M., TORTOLERO, M., ROMERO, F., SÁEZ, C. AND JAPÓN, M.A., 2020. EDIL3 promotes epithelial–mesenchymal transition and paclitaxel resistance through its interaction with integrin αVβ3 in cancer cells. Cell Death Discovery, **6** (1), 1-14.

GAWLIK-RZEMIENIEWSKA, N., AND BEDNAREK, I., 2016. The role of NANOG transcriptional factor in the development of malignant phenotype of cancer cells. Cancer Biology & Therapy, **17** (1), 1-10.

GAY, H.A., and MICHALSKI, J.M., 2018. Radiation therapy for prostate cancer. Missouri Medicine, **115** (2), 146.

GERDES, M.J., SOOD, A., SEVINSKY, C., PRIS, A.D., ZAVODSZKY, M.I. and GINTY, F., 2014. Emerging understanding of multiscale tumor heterogeneity. Frontiers in Oncology**, 4,** 366.

GHANEI, Z., MEHRI, N., JAMSHIDIZAD, A., JOUPARI, M.D. AND SHAMSARA, M., 2020. Immunization against leukemia inhibitory factor and its receptor suppresses tumor formation of breast cancer initiating cells in BALB/c mouse. Scientific Reports, **10** (1), 1-11.

GILLET, L.C., NAVARRO, P., TATE, S., RÖST, H., SELEVSEK, N., REITER, L., BONNER, R. AND AEBERSOLD, R., 2012. Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. Molecular & Cellular Proteomics, **11** (6).

GILMAN, A., and PHILIPS, F.S., 1946. The biological actions and therapeutic applications of the B-chloroethyl amines and sulfides. Science, **103** (2675), 409-436.

GINESTIER, C., HUR, M.H., CHARAFE-JAUFFRET, E., MONVILLE, F., DUTCHER, J., BROWN, M., JACQUEMIER, J., VIENS, P., KLEER, C.G. AND LIU, S., 2007. ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. Cell Stem Cell, **1** (5), 555-567.

GIRALDO, N.A., BECHT, E., REMARK, R., DAMOTTE, D., SAUTÈS-FRIDMAN, C. AND FRIDMAN, W.H., 2014. The immune contexture of primary and metastatic human tumours. Current Opinion in Immunology, **27**, 8-15.

GIRALDO, N.A., NGUYEN, P., ENGLE, E.L., KAUNITZ, G.J., COTTRELL, T.R., BERRY, S., GREEN, B., SONI, A., CUDA, J.D. AND STEIN, J.E., 2018. Multidimensional, quantitative assessment of PD-1/PD-L1 expression in patients with Merkel cell carcinoma and association with response to pembrolizumab. Journal for Immunotherapy of Cancer, **6** (1), 1-11.

GIRALDO, N.A., SANCHEZ-SALAS, R., PESKE, J.D., VANO, Y., BECHT, E., PETITPREZ, F., VALIDIRE, P., INGELS, A., CATHELINEAU, X. AND FRIDMAN, W.H., 2019. The clinical role of the TME in solid cancer. British Journal of Cancer, **120** (1), 45-53.

GONG, S., LI, Q., JETER, C.R., FAN, Q., TANG, D.G. AND LIU, B., 2015. Regulation of NANOG in cancer cells. Molecular Carcinogenesis, **54** (9), 679-687.

GONZALEZ, V.J., ABBAS-AGHABABAZADEH, F., FRIDLEY, B.L., GHANSAH, T. AND SALIGAN, L.N., 2018. Expression of sestrin genes in radiotherapy for prostate cancer and its association with fatigue: A proof-of-concept study. Biological Research for Nursing, **20** (2), 218-226.

GONZALEZ-ANGULO, A.M., TIMMS, K.M., LIU, S., CHEN, H., LITTON, J.K., POTTER, J., LANCHBURY, J.S., STEMKE-HALE, K., HENNESSY, B.T. AND ARUN, B.K., 2011. Incidence and Outcome of BRCA Mutations in Unselected Patients with Triple Receptor-Negative Breast CancerTriple-Negative Breast Cancer and BRCA Mutations. Clinical Cancer Research, **17** (5), 1082-1089.

GOOSSENS, N., NAKAGAWA, S., SUN, X. and HOSHIDA, Y., 2015. Cancer biomarker discovery and validation. Translational Cancer Research, **4** (3), 256.

GORGOULIS, V., ADAMS, P.D., ALIMONTI, A., BENNETT, D.C., BISCHOF, O., BISHOP, C., CAMPISI, J., COLLADO, M., EVANGELOU, K. AND FERBEYRE, G., 2019. Cellular senescence: defining a path forward. Cell, **179** (4), 813-827.

GRAVES, P.R., YU, L., SCHWARZ, J.K., GALES, J., SAUSVILLE, E.A., O'CONNOR, P.M. AND PIWNICA-WORMS, H., 2000. The Chk1 protein kinase and the Cdc25C regulatory pathways are targets of the anticancer agent UCN-01. Journal of Biological Chemistry, **275** (8), 5600-5605.

GRÜNERT, S., JECHLINGER, M. and BEUG, H., 2003. Diverse cellular and molecular mechanisms contribute to epithelial plasticity and metastasis. Nature Reviews Molecular Cell Biology, **4** (8), 657-665.

GULER, G., HIMMETOGLU, C., JIMENEZ, R.E., GEYER, S.M., WANG, W.P., COSTINEAN, S., PILARSKI, R.T., MORRISON, C., SUREN, D. AND LIU, J., 2011. Aberrant expression of DNA damage response proteins is associated with breast cancer subtype and clinical features. Breast Cancer Research and Treatment, **129** (2), 421-432.

GUO, J., AND HUAN, T., 2020. Comparison of full-scan, data-dependent, and data-independent acquisition modes in liquid chromatography–mass spectrometry based untargeted metabolomics. Analytical Chemistry, **92** (12), 8072-8080.

HA, S., SHIN, S.M., NAMKOONG, H., LEE, H., CHO, G.W., HUR, S.Y., KIM, T.E. AND KIM, J.W., 2004. Cancer-associated expression of minichromosome maintenance 3 gene in several human cancers and its involvement in tumorigenesis. Clinical Cancer Research, **10** (24), 8386-8395.

HADJIMICHAEL, C., CHANOUMIDOU, K., PAPADOPOULOU, N., ARAMPATZI, P., PAPAMATHEAKIS, J. and KRETSOVALI, A., 2015. Common stemness regulators of embryonic and cancer stem cells. World Journal of Stem Cells, **7** (9), 1150.

HAM, I., LEE, D. AND HUR, H., 2021. Cancer-associated fibroblast-induced resistance to chemotherapy and radiotherapy in gastrointestinal cancers. Cancers, **13** (5), 1172.

HAMED, A.R., ABDEL-AZIM, N.S., SHAMS, K.A. AND HAMMOUDA, F.M., 2019. Targeting multidrug resistance in cancer by natural chemosensitizers. Bulletin of the National Research Centre, **43** (1), 1-14.

HAMMERSTROM, A.E., CAULEY, D.H., ATKINSON, B.J. and SHARMA, P., 2011. Cancer immunotherapy: sipuleucel-T and beyond. Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy, **31** (8), 813-828.

HANAHAN, D., 2022. Hallmarks of cancer: new dimensions. Cancer Discovery, **12** (1), 31-46.

HANAHAN, D., and WEINBERG, R.A., 2000. The hallmarks of cancer. Cell, **100 (**1), 57-70.

HANAHAN, D., and WEINBERG, R.A., 2011. Hallmarks of cancer: the next generation. Cell, **144** (5), 646-674.

HANNA, J., SAHA, K., PANDO, B., VAN ZON, J., LENGNER, C.J., CREYGHTON, M.P., VAN OUDENAARDEN, A. AND JAENISCH, R., 2009. Direct cell reprogramming is a stochastic process amenable to acceleration. Nature, **462** (7273), 595-601.

HE, H., TIAN, W., CHEN, H. AND DENG, Y., 2016. MicroRNA-101 sensitizes hepatocellular carcinoma cells to doxorubicin-induced apoptosis via targeting Mcl-1. Molecular Medicine Reports, **13** (2), 1923-1929.

HE, L., BHAUMIK, M., TRIBIOLI, C., REGO, E.M., IVINS, S., ZELENT, A. AND PANDOLFI, P.P., 2000. Two critical hits for promyelocytic leukemia. Molecular Cell, **6** (5), 1131-1141.

HEGDE, A.N., AND SMITH, S.G., 2019. Recent developments in transcriptional and translational regulation underlying long-term synaptic plasticity and memory. Learning & Memory, **26** (9), 307-317.

HELMINK, B.A., KHAN, M.A., HERMANN, A., GOPALAKRISHNAN, V. AND WARGO, J.A., 2019. The microbiome, cancer, and cancer therapy. Nature Medicine, **25** (3), 377-388.

HENNESSY, B.T., SMITH, D.L., RAM, P.T., LU, Y. AND MILLS, G.B., 2005. Exploiting the PI3K/AKT pathway for cancer drug discovery. Nature Reviews Drug Discovery, **4** (12), 988-1004.

HENRY, N.L., and HAYES, D.F., 2012. Cancer biomarkers. Molecular Oncology, **6** (2), 140-146.

HEYN, H., VIDAL, E., FERREIRA, H.J., VIZOSO, M., SAYOLS, S., GOMEZ, A., MORAN, S., BOQUE-SASTRE, R., GUIL, S. AND MARTINEZ-CARDUS, A., 2016. Epigenomic analysis detects aberrant super-enhancer DNA methylation in human cancer. Genome Biology, **17** (1), 1-16.

HIRATSUKA, S., WATANABE, A., ABURATANI, H. and MARU, Y., 2006. Tumour-mediated upregulation of chemoattractants and recruitment of myeloid cells predetermines lung metastasis. Nature Cell Biology, **8** (12), 1369-1375.

HO, A., CHO, C., NAMKOONG, S., CHO, U. AND LEE, J.H., 2016. Biochemical basis of sestrin physiological activities. Trends in Biochemical Sciences, **41** (7), 621-632.

HOFFMAN, R.M., GILLILAND, F.D., ADAMS-CAMERON, M., HUNT, W.C. AND KEY, C.R., 2002. Prostate-specific antigen testing accuracy in community practice. BMC Family Practice, **3** (1), 1-8.

HOLLEMAN, A., DEN BOER, M.L., DE MENEZES, R.X., CHEOK, M.H., CHENG, C., KAZEMIER, K.M., JANKA-SCHAUB, G.E., GÖBEL, U., GRAUBNER, U.B. AND EVANS, W.E., 2006. The expression of 70 apoptosis genes in relation to lineage, genetic subtype, cellular drug resistance, and outcome in childhood acute lymphoblastic leukemia. Blood, **107** (2), 769-776.

HOTTON, J., AGOPIANTZ, M., LEROUX, A., CHARRA-BRUNAUD, C., MARIE, B., BUSBY-VENNER, H., MOREL, O., GUÉANT, J., VIGNAUD, J. AND BATTAGLIA-HSU, S., 2018. Minichromosome maintenance complex component 6 (MCM6) expression correlates with histological grade and survival in endometrioid endometrial adenocarcinoma. Virchows Archiv, **472** (4), 623-633.

HUANG, R., AND LI, A., 2020. SPAG5 is associated with unfavorable prognosis in patients with lung adenocarcinoma and promotes proliferation, motility and autophagy in A549 cells. Experimental and Therapeutic Medicine, **20** (5), 1.

HUBBARD, C., SINGLETON, D., RAUCH, M., JAYASINGHE, S., CAFISO, D. AND CASTLE, D., 2000. The secretory carrier membrane protein family: structure and membrane topology. Molecular Biology of the Cell, **11** (9), 2933-2947.

HUYNH, L.K., HIPOLITO, C.J. AND TEN DIJKE, P., 2019. A Perspective on the Development of TGF-β Inhibitors for Cancer Treatment. Biomolecules, **9** (11), 743.

HWANG-VERSLUES, W.W., KUO, W., CHANG, P., PAN, C., WANG, H., TSAI, S., JENG, Y., SHEW, J., KUNG, J.T. AND CHEN, C., 2009. Multiple lineages of human breast cancer stem/progenitor cells identified by profiling with stem cell markers. PloS One, **4** (12), e8377.

IINO, K., MITOBE, Y., IKEDA, K., TAKAYAMA, K., SUZUKI, T., KAWABATA, H., SUZUKI, Y., HORIE-INOUE, K. AND INOUE, S., 2020. RNA-binding protein NONO promotes breast cancer proliferation by post-transcriptional regulation of SKP2 and E2F8. Cancer Science, **111** (1), 148-159.

IMPERIALE, T.F., RANSOHOFF, D.F., ITZKOWITZ, S.H., LEVIN, T.R., LAVIN, P., LIDGARD, G.P., AHLQUIST, D.A. and BERGER, B.M., 2014. Multitarget stool DNA testing for colorectal-cancer screening. New England Journal of Medicine, **370** (14), 1287-1297.

INTERNATIONAL AGENCY FOR RESEARCH ON CANCER, 2012. Biological agents. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, .

ISSAC, M.S.M., YOUSEF, E., TAHIR, M.R. AND GABOURY, L.A., 2019. MCM2, MCM4, and MCM6 in breast cancer: clinical utility in diagnosis and prognosis. Neoplasia, **21** (10), 1015-1035.

JACOB, L., FREYN, M., KALDER, M., DINAS, K. and KOSTEV, K., 2018. Impact of tobacco smoking on the risk of developing 25 different cancers in the UK: a retrospective study of 422,010 patients followed for up to 30 years. Oncotarget, **9** (25), 17420.

JAGANNATHAN, N.R., and SHARMA, U., 2017. Breast tissue metabolism by magnetic resonance spectroscopy. Metabolites, **7** (2), 25.

JANIN, M., COLL-SANMARTIN, L. AND ESTELLER, M., 2020. Disruption of the RNA modifications that target the ribosome translation machinery in human cancer. Molecular Cancer, **19** (1), 1-13.

JENSEN, S.A., CALVERT, A.E., VOLPERT, G., KOURI, F.M., HURLEY, L.A., LUCIANO, J.P., WU, Y., CHALASTANIS, A., FUTERMAN, A.H. AND STEGH, A.H., 2014. Bcl2L13 is a ceramide synthase inhibitor in glioblastoma. Proceedings of the National Academy of Sciences, **111** (15), 5682-5687.

JETER, C.R., BADEAUX, M., CHOY, G., CHANDRA, D., PATRAWALA, L., LIU, C., CALHOUN-DAVIS, T., ZAEHRES, H., DALEY, G.Q. AND TANG, D.G., 2009. Functional evidence that the self-renewal gene NANOG regulates human tumor development. Stem Cells, **27** (5), 993-1005.

JETER, C.R., LIU, B., LIU, X., CHEN, X., LIU, C., CALHOUN-DAVIS, T., REPASS, J., ZAEHRES, H., SHEN, J.J. AND TANG, D.G., 2011. NANOG promotes cancer stem cell characteristics and prostate cancer resistance to androgen deprivation. Oncogene, **30** (36), 3833-3845.

JETER, C.R., LIU, B., LU, Y., CHAO, H., ZHANG, D., LIU, X., CHEN, X., LI, Q., RYCAJ, K. AND CALHOUN-DAVIS, T., 2016. NANOG reprograms prostate cancer cells to castration resistance via dynamically repressing and engaging the AR/FOXA1 signaling axis. Cell Discovery, **2** (1), 1-19.

JETER, C.R., YANG, T., WANG, J., CHAO, H. AND TANG, D.G., 2015. Concise review: NANOG in cancer stem cells and tumor development: an update and outstanding questions. Stem Cells, **33** (8), 2381-2390.

JI, Z., LI, J. AND WANG, J., 2021. Up-regulated RFC2 predicts unfavorable progression in hepatocellular carcinoma. Hereditas, **158** (1), 1-10.

JIA, Z., WANG, Y., SAWYERS, A., YAO, H., RAHMATPANAH, F., XIA, X., XU, Q., PIO, R., TURAN, T. AND KOZIOL, J.A., 2011. Diagnosis of prostate cancer using differentially expressed genes in stroma. Cancer Research, **71** (7), 2476-2487.

JIANG, J., WANG, J., HE, X., MA, W., SUN, L., ZHOU, Q., LI, M. and YU, S., 2019. High expression of SPAG 5 sustains the malignant growth and invasion of breast cancer cells through the activation of Wnt/β-catenin signalling. Clinical and Experimental Pharmacology and Physiology, **46** (6), 597-606.

JIN, L., SHEN, F., WEINFELD, M. AND SERGI, C., 2020. Insulin growth factor binding protein 7 (IGFBP7)-related cancer and IGFBP3 and IGFBP7 crosstalk. Frontiers in Oncology, **10**, 727.

JIN, R., SHEN, J., ZHANG, T., LIU, Q., LIAO, C., MA, H., LI, S. AND YU, Z., 2017. The highly expressed COL4A1 genes contributes to the proliferation and migration of the invasive ductal carcinomas. Oncotarget, **8** (35), 58172.

JORDAN, V.C., 2007. New insights into the metabolism of tamoxifen and its role in the treatment and prevention of breast cancer. Steroids, **72** (13), 829-842.

JOSEPH, C., ARSHAD, M., KUROZOMI, S., ALTHOBITI, M., MILIGY, I.M., AL-IZZI, S., TOSS, M.S., GOH, F.Q., JOHNSTON, S.J. and MARTIN, S.G., 2019. Overexpression of the cancer stem cell marker CD133 confers a poor prognosis in invasive breast cancer. Breast Cancer Research and Treatment, **174** (2), 387-399.

JU, J., YANG, W., LEE, K., OH, S., NAM, K., SHIM, S., SHIN, S.Y., GYE, M.C., CHU, I. AND SHIN, I., 2013. Regulation of Cell Proliferation and Migration by Keratin19-Induced Nuclear Import of Early Growth Response-1 in Breast Cancer CellsKRT19 Regulates Cell Proliferation and Migration. Clinical Cancer Research, **19** (16), 4335-4346.

JUNG, S., LI, C., JEONG, D., LEE, S., OHK, J., PARK, M., HAN, S., DUAN, J., KIM, C. AND YANG, Y., 2013. Oncogenic function of p34SEI-1 via NEDD4 1 mediated PTEN ubiquitination/degradation and activation of the PI3K/AKT pathway. International Journal of Oncology, **43** (5), 1587-1595.

JUNG, S.Y., HWANG, S., CLARKE, J.M., BAUER, T.M., KEEDY, V.L., LEE, H., PARK, N., KIM, S. AND LEE, J.I., 2020. Pharmacokinetic characteristics of vactosertib, a new activin receptor-like kinase 5 inhibitor, in patients with advanced solid tumors in a first-in-human phase 1 study. Investigational New Drugs, **38** (3), 812-820.

JUNTTILA, M.R., and DE SAUVAGE, F.J., 2013. Influence of tumour micro-environment heterogeneity on therapeutic response. Nature, **501** (7467), 346-354.

KALLURI, R., and WEINBERG, R.A., 2009. The basics of epithelial-mesenchymal transition. The Journal of Clinical Investigation**, 119** (6), 1420-1428.

KALLURI, R., AND ZEISBERG, M., 2006. Fibroblasts in cancer. Nature Reviews Cancer, **6** (5), 392-401.

KANG, J., YU, Y., JEONG, S., LEE, H., HEO, H.J., PARK, J.J., NA, H.S., KO, D.S. AND KIM, Y.H., 2020. Prognostic role of high cathepsin D expression in breast cancer: a systematic review and meta-analysis. Therapeutic Advances in Medical Oncology, **12**, 1758835920927838.

KANG, S., TANAKA, T. AND KISHIMOTO, T., 2015. Therapeutic uses of anti-interleukin-6 receptor antibody. International Immunology, **27** (1), 21-29.

KESHARWANI, V., KUMARI, S., KUSHWAHA, N. and SINGH, K., 2019. A review on dangerous neoplasm cancer.

KESKIN, G., and GUMUS, A.B., 2011. Turkish hysterectomy and mastectomy patients-depression, body image, sexual problems and spouse relationships. Asian Pac J Cancer Prev, **12** (2), 425-432.

KETTERER, S., MITSCHKE, J., KETSCHER, A., SCHLIMPERT, M., REICHARDT, W., BAEUERLE, N., HESS, M.E., METZGER, P., BOERRIES, M. AND PETERS, C., 2020. Cathepsin D deficiency in mammary epithelium transiently stalls breast cancer by interference with mTORC1 signaling. Nature Communications, **11** (1), 1-18.

KIM, D., LANGMEAD, B. AND SALZBERG, S.L., 2014. HISAT: hierarchical indexing for spliced alignment of transcripts. BioRxiv, , 012591.

KIM, D., ZENG, M.Y. AND NÚÑEZ, G., 2017. The interplay between host immune cells and gut microbiota in chronic inflammatory diseases. Experimental & Molecular Medicine, **49** (5), e339.

KIM, K., LEE, H., PARK, J., KIM, M., KIM, S., NOH, S., SONG, K., KIM, J.C. AND KIM, Y.S., 2011. Epigenetic regulation of microRNA-10b and targeting of oncogenic MAPRE1 in gastric cancer. Epigenetics, **6** (6), 740-751.

KIM, R., 2005. Recent advances in understanding the cell death pathways activated by anticancer therapy. Cancer: Interdisciplinary International Journal of the American Cancer Society, **103** (8), 1551-1560.

KIM, S., JU, J., KANG, M., WON, J.E., KIM, Y.H., RANINGA, P.V., KHANNA, K.K., GYŐRFFY, B., PACK, C. AND HAN, H., 2020. RNA-binding protein NONO contributes to cancer cell growth and confers drug resistance as a theranostic target in TNBC. Theranostics, **10** (18), 7974.

KIM, W., AND RYU, C.J., 2017. Cancer stem cell surface markers on normal stem cells. BMB Reports, 50 (6), 285.

KITAMURA, M., NAKAYAMA, T., MUKAISHO, K., MORI, T., UMEDA, T., MORITANI, S., KUSHIMA, R., TANI, M. and SUGIHARA, H., 2019. Progression potential of ductal carcinoma in situ assessed by genomic copy number profiling. Pathobiology, **86** (2-3), 92-101.

KOLTAI, T., 2014. Clusterin: a key player in cancer chemoresistance and its inhibition. OncoTargets and Therapy, **7**, 447.

KOLUPAEVA, V., and JANSSENS, V., 2013. PP1 and PP2A phosphatases–cooperating partners in modulating retinoblastoma protein activation. The FEBS Journal, **280** (2), 627-643.

KOPPENOL, W.H., BOUNDS, P.L. AND DANG, C.V., 2011. Otto Warburg's contributions to current concepts of cancer metabolism. Nature Reviews Cancer, **11** (5), 325-337.

KOREN, A., RIJAVEC, M., KERN, I., SODJA, E., KOROSEC, P. and CUFER, T., 2016. BMI1, ALDH1A1, and CD133 transcripts connect epithelial-mesenchymal transition to cancer stem cells in lung carcinoma. Stem Cells International, 2016.

KOSTERS, J.P., and GOTZSCHE, P.C., 2003. Regular self-examination or clinical examination for early detection of breast cancer. Cochrane Database Syst Rev**, 2** (2).

KRAGSTRUP, T.W., VORUP-JENSEN, T., DELEURAN, B. AND HVID, M., 2013. A simple set of validation steps identifies and removes false results in a sandwich enzyme-linked immunosorbent assay caused by anti-animal IgG antibodies in plasma from arthritis patients. Springerplus, **2** (1), 1-10.

KRASNY, L., AND HUANG, P.H., 2021. Data-independent acquisition mass spectrometry (DIA-MS) for proteomic applications in oncology. Molecular Omics, **17** (1), 29-42.

KVÅLE, R., MØLLER, B., WAHLQVIST, R., FOSSÅ, S.D., BERNER, A., BUSCH, C., KYRDALEN, A.E., SVINDLAND, A., VISET, T. and HALVORSEN, O.J., 2009. Concordance between Gleason scores of needle biopsies and radical prostatectomy specimens: a population-based study. BJU International, **103** (12), 1647-1654.

LABRECQUE, M.P., COLEMAN, I.M., BROWN, L.G., TRUE, L.D., KOLLATH, L., LAKELY, B., NGUYEN, H.M., YANG, Y.C., DA COSTA, R.M.G. AND KAIPAINEN, A., 2019. Molecular profiling stratifies diverse phenotypes of treatment-refractory metastatic castration-resistant prostate cancer. The Journal of Clinical Investigation, **129** (10), 4492-4505.

LADD, J.J., BUSALD, T., JOHNSON, M.M., ZHANG, Q., PITTERI, S.J., WANG, H., BRENNER, D.E., LAMPE, P.D., KUCHERLAPATI, R. AND FENG, Z., 2012. Increased plasma levels of the APC-interacting protein MAPRE1, LRG1, and IGFBP2 preceding a diagnosis of colorectal cancer in women. Cancer Prevention Research, **5** (4), 655-664.

LAKHTER, A.J., HAMILTON, J., KONGER, R.L., BRUSTOVETSKY, N., BROXMEYER, H.E. AND NAIDU, S.R., 2016. Glucose-independent acetate metabolism promotes melanoma cell survival and tumor growth. Journal of Biological Chemistry, **291** (42), 21869-21879.

LAL, N., WILLCOX, C.R., BEGGS, A., TANIERE, P., SHIKOTRA, A., BRADDING, P., ADAMS, R., FISHER, D., MIDDLETON, G. AND TSELEPIS, C., 2017. Endothelial protein C receptor is overexpressed in colorectal cancer as a result of amplification and hypomethylation of chromosome 20q. The Journal of Pathology: Clinical Research**, 3** (3), 155-170.

LANCASHIRE, L.J., POWE, D.G., REIS-FILHO, J.S., RAKHA, E., LEMETRE, C., WEIGELT, B., ABDEL-FATAH, T.M., GREEN, A.R., MUKTA, R. AND BLAMEY, R., 2010. A validated gene expression profile for detecting clinical outcome in breast cancer using artificial neural networks. Breast Cancer Research and Treatment, **120** (1), 83-93.

LAPIDOT, T., SIRARD, C., VORMOOR, J., MURDOCH, B., HOANG, T., CACERES-CORTES, J., MINDEN, M., PATERSON, B., CALIGIURI, M.A. and DICK, J.E., 1994. A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. Nature, **367** (6464), 645-648.

LARSEN, M.J., THOMASSEN, M., GERDES, A. AND KRUSE, T.A., 2014. Hereditary breast cancer: clinical, pathological and molecular characteristics. Breast Cancer: Basic and Clinical Research, **8**, BCBCR. S18715.

LASZIK, Z., MITRO, A., TAYLOR JR, F.B., FERRELL, G. and ESMON, C.T., 1997. Human protein C receptor is present primarily on endothelium of large blood vessels: implications for the control of the protein C pathway. Circulation, **96** (10), 3633-3640.

LATORRE, E., CARELLI, S., RAIMONDI, I., D'AGOSTINO, V., CASTIGLIONI, I., ZUCAL, C., MORO, G., LUCIANI, A., GHILARDI, G. AND MONTI, E., 2016. The Ribonucleic Complex HuR-MALAT1 Represses CD133 Expression and Suppresses Epithelial–Mesenchymal Transition in Breast CancerEMT Can Be Modulated by Fine Tuning MALAT1 and HuR. Cancer Research, **76** (9), 2626-2636.

LAWRENCE, M.S., STOJANOV, P., MERMEL, C.H., ROBINSON, J.T., GARRAWAY, L.A., GOLUB, T.R., MEYERSON, M., GABRIEL, S.B., LANDER, E.S. AND GETZ, G., 2014. Discovery and saturation analysis of cancer genes across 21 tumour types. Nature, **505** (7484), 495-501.

LAWSON, P., SHOLL, A.B., BROWN, J., FASY, B.T. and WENK, C., 2019. Persistent homology for the quantitative evaluation of architectural features in prostate cancer histology. Scientific Reports, **9** (1), 1-15.

LAZEBNIK, Y., 2010. What are the hallmarks of cancer? Nature Reviews Cancer, **10** (4), 232-233.

LEE, H.P., LEE, J., GOURLEY, L., DUFFY, S.W., DAY, N.E. and ESTÈVE, J., 1991. Dietary effects on breast-cancer risk in Singapore. The Lancet, **337** (8751), 1197-1200.

LEE, R., LOCALIO, A.R., ARMSTRONG, K., MALKOWICZ, S.B., SCHWARTZ, J.S. and FREE PSA STUDY GROUP, 2006. A meta-analysis of the performance characteristics of the free prostate-specific antigen test. Urology, **67** (4), 762-768.

LEHMAN, C.D., WHITE, E., PEACOCK, S., DRUCKER, M.J. AND URBAN, N., 1999. Effect of age and breast density on screening mammograms with false-positive findings. AJR.American Journal of Roentgenology, **173** (6), 1651-1655.

LEUNG, C.S., YEUNG, T., YIP, K., PRADEEP, S., BALASUBRAMANIAN, L., LIU, J., WONG, K., MANGALA, L.S., ARMAIZ-PENA, G.N. AND LOPEZ-BERESTEIN, G., 2014. Calcium-dependent FAK/CREB/TNNC1 signalling mediates the effect of stromal MFAP5 on ovarian cancer metastatic potential. Nature Communications, **5** (1), 1-15.

LI, M., LI, A., ZHOU, S., LV, H. AND YANG, W., 2019. SPAG5 upregulation contributes to enhanced c-MYC transcriptional activity via interaction with c-MYC binding protein in triple-negative breast cancer. Journal of Hematology & Oncology, **12** (1), 1-18.

LI, W., ZHOU, Y., YANG, J., ZHANG, X., ZHANG, H., ZHANG, T., ZHAO, S., ZHENG, P., HUO, J. and WU, H., 2015. Gastric cancer-derived mesenchymal stem cells prompt gastric cancer progression through secretion of interleukin-8. Journal of Experimental & Clinical Cancer Research, **34** (1), 1-15.

LI, X., CAO, X., LI, X., ZHANG, W. AND FENG, Y., 2007. Expression level of insulin-like growth factor binding protein 5 mRNA is a prognostic factor for breast cancer. Cancer Science, **98** (10), 1592-1596.

LI, Y., AND MA, L., 2020. Exposure to solar ultraviolet radiation and breast cancer risk: A dose-response meta-analysis. Medicine, **99** (45).

LIANG, X., FENG, Z., LIU, F., YAN, R., YIN, L., SHEN, H. AND LU, H., 2020. MAPRE1 promotes cell cycle progression of hepatocellular carcinoma cells by interacting with CDK2. Cell Biology International, **44** (11), 2326-2333.

LIBERTI, M.V., AND LOCASALE, J.W., 2016. The Warburg effect: how does it benefit cancer cells? Trends in Biochemical Sciences, **41** (3), 211-218.

LIEDTKE, C., MAZOUNI, C., HESS, K.R., ANDRÉ, F., TORDAI, A., MEJIA, J.A., SYMMANS, W.F., GONZALEZ-ANGULO, A.M., HENNESSY, B. and GREEN, M., 2008. Response to neoadjuvant therapy and long-term survival in patients with triple-negative breast cancer. Journal of Clinical Oncology, **26** (8), 1275-1281.

LIM, S., AND KALDIS, P., 2013. Cdks, cyclins and CKIs: roles beyond cell cycle regulation. Development, **140** (15), 3079-3093.

LIN, X., XIAO, Z., CHEN, T., LIANG, S.H. AND GUO, H., 2020. Glucose metabolism on tumor plasticity, diagnosis, and treatment. Frontiers in Oncology, 10, 317.

LIU, C., LI, C., YI-JEN, P., CHENG, Y., CHEN, H., LIAO, P., KANG, J. AND YENG, M., 2014. Snail regulates Nanog status during the epithelial–mesenchymal transition via the Smad1/Akt/GSK3β signaling pathway in non-small-cell lung cancer. Oncotarget, **5** (11), 3880.

LIU, C., WANG, H., WANG, L., LIU, W., WANG, J., GENG, Q. AND LU, Y., 2018. ENO2 promotes cell proliferation, glycolysis, and glucocorticoid-resistance in acute lymphoblastic leukemia. Cellular Physiology and Biochemistry, **46** (4), 1525-1535.

LIU, H., HU, J., WEI, R., ZHOU, L., PAN, H., ZHU, H., HUANG, M., LUO, J. AND XU, W., 2018. SPAG5 promotes hepatocellular carcinoma progression by downregulating SCARA5 through modifying β-catenin degradation. Journal of Experimental & Clinical Cancer Research, **37 (**1), 1-14.

LIU, J.Y., ZENG, Q.H., CAO, P.G., XIE, D., YANG, F., HE, L.Y., DAI, Y.B., LI, J.J., LIU, X.M. and ZENG, H.L., 2018. SPAG5 promotes proliferation and suppresses apoptosis in bladder urothelial carcinoma by upregulating Wnt3 via activating the AKT/mTOR pathway and predicts poorer survival. Oncogene, **37** (29), 3937-3952.

LIU, L., LIU, S., DENG, P., LIANG, Y., XIAO, R., TANG, L., CHEN, J., CHEN, Q., GUAN, P. AND YAN, S., 2021. Targeting the IRAK1–S100A9 Axis Overcomes Resistance to Paclitaxel in Nasopharyngeal Carcinoma. Cancer Research, **81** (5), 1413-1425.

LIU, P.Y., XU, N., MALYUKOVA, A., SCARLETT, C.J., SUN, Y.T., ZHANG, X.D., LING, D., SU, S.P., NELSON, C. AND CHANG, D.K., 2013. The histone deacetylase SIRT2 stabilizes Myc oncoproteins. Cell Death & Differentiation, **20** (3), 503-514.

LIU, S., TSANG, N., CHIANG, W., CHANG, K., HSUEH, C., LIANG, Y., JUANG, J., CHOW, K.N. AND CHANG, Y., 2013. Leukemia inhibitory factor promotes nasopharyngeal carcinoma progression and radioresistance. The Journal of Clinical Investigation, **123** (12), 5269-5283.

LIU, Y., KIM, H.G., DONG, E., DONG, C., HUANG, M., LIU, Y., LIANGPUNSAKUL, S. AND DONG, X.C., 2019. Sesn3 deficiency promotes carcinogen-induced hepatocellular carcinoma via regulation of the hedgehog pathway. Biochimica Et Biophysica Acta (BBA)-Molecular Basis of Disease, **1865** (10), 2685-2693.

LIVAK, K.J., AND SCHMITTGEN, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2− ΔΔCT method. Methods, **25** (4), 402-408.

LO, U., LEE, C., LEE, M. AND HSIEH, J., 2017. The role and mechanism of epithelial-to-mesenchymal transition in prostate cancer progression. International Journal of Molecular Sciences, **18** (10), 2079.

LOCKSLEY, R.M., KILLEEN, N. and LENARDO, M.J., 2001. The TNF and TNF receptor superfamilies: integrating mammalian biology. Cell, **104** (4), 487-501.

LOEB, S., and CATALONA, W.J., 2009. What is the role of digital rectal examination in men undergoing serial screening of serum PSA levels? Nature Clinical Practice Urology, **6** (2), 68-69.

LOEB, S., ROEHL, K.A., ANTENOR, J.A.V., CATALONA, W.J., SUAREZ, B.K. and NADLER, R.B., 2006. Baseline prostate-specific antigen compared with median prostate-specific antigen for age group as predictor of prostate cancer risk in men younger than 60 years old. Urology, **67** (2), 316-320.

LØKKEGAARD, S., ELIAS, D., ALVES, C.L., BENNETZEN, M.V., LÆNKHOLM, A., BAK, M., GJERSTORFF, M.F., JOHANSEN, L.E., VEVER, H. AND BJERRE, C., 2021. MCM3 upregulation confers endocrine resistance in breast cancer and is a predictive marker of diminished tamoxifen benefit. NPJ Breast Cancer, **7** (1), 1-15.

LOTHSTEIN, L., ISRAEL, M. AND SWEATMAN, T.W., 2001. Anthracycline drug targeting: cytoplasmic versus nuclear–a fork in the road. Drug Resistance Updates, **4** (3), 169-177.

LOVE, M.I., HUBER, W. AND ANDERS, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology, **15** (12), 1-21.

LU, Z., ZOU, J., LI, S., TOPPER, M.J., TAO, Y., ZHANG, H., JIAO, X., XIE, W., KONG, X. AND VAZ, M., 2020. Epigenetic therapy inhibits metastases by disrupting premetastatic niches. Nature, **579** (7798), 284-290.

LUCENA-CACACE, A., OTERO-ALBIOL, D., JIMÉNEZ-GARCÍA, M.P., MUÑOZ-GALVAN, S. AND CARNERO, A., 2018. NAMPT Is a Potent Oncogene in Colon Cancer Progression that Modulates Cancer Stem Cell Properties and Resistance to Therapy through Sirt1 and PARPNAMPT Overexpression Induces CSCs in Colon Cancer. Clinical Cancer Research, **24** (5), 1202-1215.

LUCENA-CACACE, A., OTERO-ALBIOL, D., JIMÉNEZ-GARCÍA, M.P., PEINADO-SERRANO, J. AND CARNERO, A., 2017. NAMPT overexpression induces cancer stemness and defines a novel tumor signature for glioma prognosis. Oncotarget, **8** (59), 99514.

LUMACHI, F., SANTEUFEMIA, D.A. AND BASSO, S.M., 2015. Current medical treatment of estrogen receptor-positive breast cancer. World Journal of Biological Chemistry, **6** (3), 231.

LUNDBERG, I.V., EDIN, S., EKLÖF, V., ÖBERG, Å, PALMQVIST, R. and WIKBERG, M.L., 2016a. SOX2 expression is associated with a cancer stem cell state and down-regulation of CDX2 in colorectal cancer. BMC Cancer, **16** (1), 1-11.

LV, M., LUO, L. AND CHEN, X., 2022. The landscape of prognostic and immunological role of myosin light chain 9 (MYL9) in human tumors. Immunity, Inflammation and Disease, **10** (2), 241-254.

MACDONALD, J., HENRI, J., ROY, K., HAYS, E., BAUER, M., VEEDU, R.N., POULIOT, N. and SHIGDAR, S., 2018. EpCAM immunotherapy versus specific targeted delivery of drugs. Cancers, **10** (1), 19.

MACK, G.J., and COMPTON, D.A., 2001. Analysis of mitotic microtubule-associated proteins using mass spectrometry identifies astrin, a spindle-associated protein. Proceedings of the National Academy of Sciences, **98** (25), 14434-14439.

MADAR, S., GOLDSTEIN, I. AND ROTTER, V., 2013. 'Cancer associated fibroblasts'–more than meets the eye. Trends in Molecular Medicine, **19** (8), 447-453.

MAJKA, J., CHUNG, B.Y. AND BURGERS, P.M., 2004. Requirement for ATP by the DNA damage checkpoint clamp loader. Journal of Biological Chemistry, **279** (20), 20921-20926.

MANI, S.A., GUO, W., LIAO, M., EATON, E.N., AYYANAN, A., ZHOU, A.Y., BROOKS, M., REINHARD, F., ZHANG, C.C. AND SHIPITSIN, M., 2008. The epithelial-mesenchymal transition generates cells with properties of stem cells. Cell, **133** (4), 704-715.

MANTOVANI, A., ALLAVENA, P., SICA, A. AND BALKWILL, F., 2008. Cancer-related inflammation. Nature, **454** (7203), 436-444.

MARTIN, T.A., YE, L., SANDERS, A.J., LANE, J. and JIANG, W.G., 2013, Cancer invasion and metastasis: molecular and cellular perspective. In: Cancer invasion and metastasis: molecular and cellular perspective. Madame Curie Bioscience Database [Internet]. Landes Bioscience,2000- 2013.

MARTINS-TEIXEIRA, M.B., AND CARVALHO, I., 2020. Antitumour anthracyclines: progress and perspectives. ChemMedChem, **15** (11), 933-948.

MASHIMO, T., PICHUMANI, K., VEMIREDDY, V., HATANPAA, K.J., SINGH, D.K., SIRASANAGANDLA, S., NANNEPAGA, S., PICCIRILLO, S.G., KOVACS, Z. AND FOONG, C., 2014. Acetate is a bioenergetic substrate for human glioblastoma and brain metastases. Cell, **159** (7), 1603-1614.

MASSAGUÉ, J., 2012. TGFβ signalling in context. Nature Reviews Molecular Cell Biology, **13** (10), 616-630.

MATSUSHITA, K., KOBAYASHI, S., AKITA, H., KONNO, M., ASAI, A., NODA, T., IWAGAMI, Y., ASAOKA, T., GOTOH, K. AND MORI, M., 2022. Clinicopathological significance of MYL9 expression in pancreatic ductal adenocarcinoma. Cancer Reports, **5** (10), e1582.

MCCABE, N., TURNER, N.C., LORD, C.J., KLUZEK, K., BIAŁKOWSKA, A., SWIFT, S., GIAVARA, S., O'CONNOR, M.J., TUTT, A.N. AND ZDZIENICKA, M.Z., 2006. Deficiency in the repair of DNA damage by homologous recombination and sensitivity to poly (ADP-ribose) polymerase inhibition. Cancer Research, **66** (16), 8109-8115.

MCGOWAN, J.V., CHUNG, R., MAULIK, A., PIOTROWSKA, I., WALKER, J.M. AND YELLON, D.M., 2017. Anthracycline chemotherapy and cardiotoxicity. Cardiovascular Drugs and Therapy, **31** (1), 63-75.

MEHTA, S.H., DHANDAPANI, K.M., DE SEVILLA, L.M., WEBB, R.C., MAHESH, V.B. and BRANN, D.W., 2003. Tamoxifen, a selective estrogen receptor modulator, reduces ischemic damage caused by middle cerebral artery occlusion in the ovariectomized female rat. Neuroendocrinology, **77** (1), 44-50.

MENG, F., ZHANG, L., ZHANG, M., YE, K., GUO, W., LIU, Y., YANG, W., ZHAI, Z., WANG, H. AND XIAO, J., 2021. Down-regulation of BCL2L13 renders poor prognosis in clear cell and papillary renal cell carcinoma. Cancer Cell International, **21** (1), 1-12.

MENSCHIKOWSKI, M., HAGELGANS, A., TIEBEL, O., KLINSMANN, L., EISENHOFER, G. and SIEGERT, G., 2011. Expression and shedding of endothelial protein C receptor in prostate cancer cells. Cancer Cell International, **11** (1), 1-10.

MERRIEL, S.W., FUNSTON, G. and HAMILTON, W., 2018. Prostate cancer in primary care. Advances in Therapy, **35** (9), 1285-1294.

MILLER, K.D., AND SCHUG, Z.T., 2021. Targeting acetate metabolism: Achilles' nightmare. British Journal of Cancer, **124** (12), 1900-1901.

MILLER, W.R., AND LARIONOV, A., 2010. Changes in expression of oestrogen regulated and proliferation genes with neoadjuvant treatment highlight heterogeneity of clinical resistance to the aromatase inhibitor, letrozole. Breast Cancer Research, **12** (4), 1-9.

MING, X., BAO, C., HONG, T., YANG, Y., CHEN, X., JUNG, Y. AND QIAN, Y., 2018. Clusterin, a Novel DEC1 Target, Modulates DNA Damage–Mediated Cell DeathDEC1 Regulates sCLU-Mediated Cytoprotective Activity. Molecular Cancer Research, **16** (11), 1641-1651.

MIREMADI, A., PINDER, S.E., LEE, A., BELL, J.A., PAISH, E.C., WENCYK, P., ELSTON, C.W., NICHOLSON, R.I., BLAMEY, R.W. AND ROBERTSON, J.F., 2002. Neuroendocrine differentiation and prognosis in breast adenocarcinoma. Histopathology, **40** (3), 215-222.

MOCH, H., CUBILLA, A.L., HUMPHREY, P.A., REUTER, V.E. and ULBRIGHT, T.M., 2016. The 2016 WHO classification of tumours of the urinary system and male genital organs—part A: renal, penile, and testicular tumours. European Urology, **70** (1), 93-105.

Moe, A., and HAYNE, D., 2020. Transrectal ultrasound biopsy of the prostate: does it still have a role in prostate cancer diagnosis? Translational Andrology and Urology, **9** (6), 3018.

MOHAMADALIZADEH-HANJANI, Z., SHAHBAZI, S. and GERANPAYEH, L., 2020. Investigation of the SPAG5 gene expression and amplification related to the NuMA mRNA levels in breast ductal carcinoma. World Journal of Surgical Oncology, **18** (1), 1-10.

MOHAN RAO, L.V., ESMON, C.T. and PENDURTHI, U.R., 2014. Endothelial cell protein C receptor: a multiliganded and multifunctional receptor. Blood, the Journal of the American Society of Hematology, **124** (10), 1553-1562.

MOHIUDDIN, I.S., WEI, S. and KANG, M.H., 2020. Role of OCT4 in cancer stem-like cells and chemotherapy resistance. Biochimica Et Biophysica Acta (BBA)-Molecular Basis of Disease, **1866** (4), 165432.

MOORE, M.A., ARIYARATNE, Y., BADAR, F., BHURGRI, Y., DATTA, K., MATHEW, A., GANGADHARAN, P., NANDAKUMAR, A., PRADHANANGA, K.K. and TALUKDER, M.H., 2010. Cancer epidemiology in South Asia-past, present and future. Asian Pac J Cancer Prev, **11** (Suppl 2), 49-66.

MORTON, S.D., CADAMURO, M., BRIVIO, S., VISMARA, M., STECCA, T., MASSANI, M., BASSI, N., FURLANETTO, A., JOPLIN, R.E. AND FLOREANI, A., 2015. Leukemia inhibitory factor protects cholangiocarcinoma cells from drug-induced apoptosis via a PI3K/AKT-dependent Mcl-1 activation. Oncotarget, **6** (28), 26052.

MULTHOFF, G., MOLLS, M. and RADONS, J., 2012. Chronic inflammation in cancer development. Frontiers in Immunology, **2**, 98.

MUNRO, A.F., TWELVES, C., THOMAS, J.S., CAMERON, D.A. AND BARTLETT, J., 2012. Chromosome instability and benefit from adjuvant anthracyclines in breast cancer. British Journal of Cancer, **107** (1), 71-74.

NAJAFI, M., FARHOOD, B. and MORTEZAEE, K., 2019. Cancer stem cells (CSCs) in cancer progression and therapy. Journal of Cellular Physiology, **234** (6), 8381-8395.

NAM, K., OH, S., LEE, K., YOO, S. and SHIN, I., 2015. CD44 regulates cell proliferation, migration, and invasion via modulation of c-Src transcription in human breast cancer cells. Cellular Signalling, **27** (9), 1882-1894.

NANDANA, S., AND CHUNG, L.W., 2014. Prostate cancer progression and metastasis: potential regulatory pathways for therapeutic targeting. American Journal of Clinical and Experimental Urology, **2** (2), 92.

NIEHRS, C., 2012. The complex world of WNT receptor signalling. Nature Reviews Molecular Cell Biology, **13** (12), 767-779.

NIWA, H., MIYAZAKI, J. and SMITH, A.G., 2000. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. Nature Genetics, **24** (4), 372-376.

NORBURY, C., AND NURSE, P., 1992. Animal cell cycles and their control. Annual Review of Biochemistry, **61** (1), 441-468.

ODA, E., OHKI, R., MURASAWA, H., NEMOTO, J., SHIBUE, T., YAMASHITA, T., TOKINO, T., TANIGUCHI, T. and TANAKA, N., 2000. Noxa, a BH3-only member of the Bcl-2 family and candidate mediator of p53-induced apoptosis. Science, **288** (5468), 1053-1058.

ODOUX, C., FOHRER, H., HOPPO, T., GUZIK, L., STOLZ, D.B., LEWIS, D.W., GOLLIN, S.M., GAMBLIN, T.C., GELLER, D.A. AND LAGASSE, E., 2008. A stochastic model for cancer stem cell origin in metastatic colon cancer. Cancer Research, **68** (17), 6932-6941.

OEFFINGER, K.C., FONTHAM, E.T., ETZIONI, R., HERZIG, A., MICHAELSON, J.S., SHIH, Y.T., WALTER, L.C., CHURCH, T.R., FLOWERS, C.R. and LAMONTE, S.J., 2015. Breast cancer screening for women at average risk: 2015 guideline update from the American Cancer Society. Jama, **314** (15), 1599-1614.

OESTERLING, J.E., JACOBSEN, S.J. and COONER, W.H., 1995. The Use of Age-Specific Reference Ranges for Serum Prostate Specific Anitgen in Men 60 years Old or Older. The Journal of Urology, **153** (4), 1160-1163.

OESTERLING, J.E., RICE, D.C., GLENSKI, W.J. and BERGSTRALH, E.J., 1993. Effect of cystoscopy, prostate biopsy, and transurethral resection of prostate on serum prostate-specific antigen concentration. Urology, **42** (3), 276-282.

OGANESYAN, V., OGANESYAN, N., TERZYAN, S., QU, D., DAUTER, Z., ESMON, N.L. and ESMON, C.T., 2002. The crystal structure of the endothelial protein C receptor and a bound phospholipid. Journal of Biological Chemistry, **277** (28), 24851-24854.

OH, K., LEE, O., PARK, Y., SEO, M.W. AND LEE, D., 2016. IL-1β induces IL-6 production and increases invasiveness and estrogen-independent growth in a TG2-dependent manner in human breast cancer cells. BMC Cancer, **16** (1), 1-11.

OLIVEROS, J.C., 2007. VENNY. An interactive tool for comparing lists with Venn Diagrams. Http://Bioinfogp.Cnb.Csic.Es/Tools/Venny/Index.Html, .

OLMEDA, D., MORENO-BUENO, G., FLORES, J.M., FABRA, A., PORTILLO, F. AND CANO, A., 2007. SNAI1 is required for tumor growth and lymph node metastasis of human breast carcinoma MDA-MB-231 cells. Cancer Research, **67** (24), 11721-11731.

OSGOOD, E.E., 1964. Contrasting incidence of acute monocytic and granulocytic leukemias in P32-treated patients with polycythemia vera and chronic lymphocytic leukemia. The Journal of Laboratory and Clinical Medicine, **64** (4), 560-573.

OSHIMA, S., KISA, K., TERASHITA, T., HABARA, M., KAWABATA, H. and MAEZAWA, M., 2011. A qualitative study of Japanese patients' perspectives on post-treatment care for gynecological cancer. Asian Pac J Cancer Prev, **12** (9), 2255-2261.

ØSTERÅS, B.H., MARTINSEN, A.C.T., GULLIEN, R. AND SKAANE, P., 2019. Digital mammography versus breast tomosynthesis: impact of breast density on diagnostic performance in population-based screening. Radiology, **293** (1), 60-68.

OZAKI, T., and NAKAGAWARA, A., 2011. Role of p53 in cell death and human cancers. Cancers, **3** (1), 994-1013.

OZKAN, A., MALAK, A.T., GURKAN, A. and TURGAY, A.S., 2010. Do Turkish nursing and midwifery students teach breast self-examination to their relatives. Asian Pac J Cancer Prev, **11** (6), 1569-1573.

PAIK, S., SHAK, S., TANG, G., KIM, C., BAKER, J., CRONIN, M., BAEHNER, F.L., WALKER, M.G., WATSON, D. and PARK, T., 2004. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. New England Journal of Medicine, **351** (27), 2817-2826.

PALOMERAS, S., DIAZ-LAGARES, Á, VIÑAS, G., SETIEN, F., FERREIRA, H.J., OLIVERAS, G., CRUJEIRAS, A.B., HERNÁNDEZ, A., LUM, D.H. AND WELM, A.L., 2019. Epigenetic silencing of TGFBI confers resistance to trastuzumab in human breast cancer. Breast Cancer Research, **21** (1), 1-16.

PAREKH, A., DAS, D., DAS, S., DHARA, S., BISWAS, K., MANDAL, M. AND DAS, S., 2018. Bioimpedimetric analysis in conjunction with growth dynamics to differentiate aggressiveness of cancer cells. Scientific Reports, **8** (1), 1-10.

PARISE, C.A., AND CAGGIANO, V., 2014. Breast cancer survival defined by the ER/PR/HER2 subtypes and a surrogate classification according to tumor grade and immunohistochemical biomarkers. Journal of Cancer Epidemiology, 2014.

PARK, S.Y., LEE, H.E., LI, H., SHIPITSIN, M., GELMAN, R. AND POLYAK, K., 2010. Heterogeneity for stem cell–related markers according to tumor subtype and histologic stage in breast cancer. Clinical Cancer Research, **16** (3), 876-887.

PATEL, P., AND CHEN, E.I., 2012. Cancer stem cells, tumor dormancy, and metastasis. Frontiers in Endocrinology, **3**, 125.

PAULA, A.D.C., LEITÃO, C., MARQUES, O., ROSA, A.M., SANTOS, A.H., RÊMA, A., DE FÁTIMA FARIA, M., ROCHA, A., COSTA, J.L. and LIMA, M., 2017. Molecular characterization of CD44 /CD24−/Ck /CD45− cells in benign and malignant breast lesions. Virchows Archiv, **470** (3), 311-322.

PHI, L.T.H., SARI, I.N., YANG, Y., LEE, S., JUN, N., KIM, K.S., LEE, Y.K. AND KWON, H.Y., 2018. Cancer stem cells (CSCs) in drug resistance and their therapeutic implications in cancer treatment. Stem Cells International, 2018.

PIÑEROS, M., MERY, L., SOERJOMATARAM, I., BRAY, F. and STELIAROVA-FOUCHER, E., 2021. Scaling up the surveillance of childhood cancer: a global roadmap. JNCI: Journal of the National Cancer Institute, **113** (1), 9-15.

PIRTSKHALAISHVILI, G., and NELSON, J.B., 2000. Endothelium-derived factors as paracrine mediators of prostate cancer progression. The Prostate, **44** (1), 77-87.

PITT, J.J., 2009. Principles and applications of liquid chromatography-mass spectrometry in clinical biochemistry. The Clinical Biochemist Reviews, **30** (1), 19.

PLAKS, V., KONG, N. and WERB, Z., 2015. The cancer stem cell niche: how essential is the niche in regulating stemness of tumor cells? Cell Stem Cell, **16** (3), 225-238.

PONTI, D., COSTA, A., ZAFFARONI, N., PRATESI, G., PETRANGOLINI, G., CORADINI, D., PILOTTI, S., PIEROTTI, M.A. AND DAIDONE, M.G., 2005. Isolation and in vitro propagation of tumorigenic breast cancer cells with stem/progenitor cell properties. Cancer Research, **65** (13), 5506-5511.

PRABAVATHY, D., SWARNALATHA, Y. and RAMADOSS, N., 2018. Lung cancer stem cells—origin, characteristics and therapy. Stem Cell Investigation, **5**.

PRINCIPE, S., MEJIA-GUERRERO, S., IGNATCHENKO, V., SINHA, A., IGNATCHENKO, A., SHI, W., PEREIRA, K., SU, S., HUANG, S.H. AND O'SULLIVAN, B., 2018. Proteomic analysis of cancer-associated fibroblasts reveals a paracrine role for MFAP5 in human oral tongue squamous cell carcinoma. Journal of Proteome Research**,** **17** (6), 2045-2059.

Qi, H., Liu, S., Guo, C., Wang, J., Greenaway, F.T. and Sun, M., 2015. Role of annexin A6 in cancer. Oncology Letters**,** **10** (4), 1947-1952.

QING, G., AND SIMON, M.C., 2009. Hypoxia inducible factor-2α: a critical mediator of aggressive tumor phenotypes. Current Opinion in Genetics & Development, **19** (1), 60-66.

RACHMAN-TZEMAH, C., ZAFFRYAR-EILOT, S., GROSSMAN, M., RIBERO, D., TIMANER, M., MÄKI, J.M., MYLLYHARJU, J., BERTOLINI, F., HERSHKOVITZ, D. AND SAGI, I., 2017. Blocking surgically induced lysyl oxidase activity reduces the risk of lung metastases. Cell Reports, **19 (**4), 774-784.

RANDRIAN, V., BIAU, J., BENOIT, C., PEZET, D., LAPEYRE, M. AND MOREAU, J., 2020. Preoperative intensity-modulated radiotherapy of rectal cancers: Relevance and modalities. Cancer Radiotherapie: Journal De La Societe Francaise De Radiotherapie Oncologique, **24** (4), 345-353.

RAO, X., DI LEVA, G., LI, M., FANG, F., DEVLIN, C., HARTMAN-FREY, C., BUROW, M.E., IVAN, M., CROCE, C.M. AND NEPHEW, K.P., 2011. MicroRNA-221/222 confers breast cancer fulvestrant resistance by regulating multiple signaling pathways. Oncogene, **30** (9), 1082-1097.

REBBECK, T.R., FRIEBEL, T., LYNCH, H.T., NEUHAUSEN, S.L., VAN'T VEER, L., GARBER, J.E., EVANS, G.R., NAROD, S.A., ISAACS, C. and MATLOFF, E., 2004. Bilateral prophylactic mastectomy reduces breast cancer risk in BRCA1 and BRCA2 mutation carriers: the PROSE Study Group. Journal of Clinical Oncology, **22** (6), 1055-1062.

REDONDO, M., VILLAR, E., TORRES-MUNOZ, J., TELLEZ, T., MORELL, M. AND PETITO, C.K., 2000. Overexpression of clusterin in human breast carcinoma. The American Journal of Pathology, **157** (2), 393-399.

RESCHER, U., AND GERKE, V., 2004. Annexins–unique membrane binding proteins with diverse functions. Journal of Cell Science, **117** (13), 2631-2639.

RHYASEN, G.W., AND STARCZYNOWSKI, D.T., 2015. IRAK signalling in cancer. British Journal of Cancer, **112** (2), 232-237.

RIBATTI, D., and CRIVELLATO, E., 2012. "Sprouting angiogenesis", a reappraisal. Developmental Biology, **372** (2), 157-165.

RICH, J.N., 2016. Cancer stem cells: understanding tumor hierarchy and heterogeneity. Medicine, **95** (Suppl 1).

RIDGE, S.M., SULLIVAN, F.J. and GLYNN, S.A., 2017. Mesenchymal stem cells: key players in cancer progression. Molecular Cancer, **16** (1), 1-10.

RODRIGUES-FERREIRA, S., NEHLIG, A., MONCHECOURT, C., NASR, S., FUHRMANN, L., LACROIX-TRIKI, M., GARBERIS, I., SCOTT, V., DELALOGE, S. AND PISTILLI, B., 2019. Combinatorial expression of microtubule-associated EB1 and ATIP3 biomarkers improves breast cancer prognosis. Breast Cancer Research and Treatment, **173** (3), 573-583.

ROHNE, P., PROCHNOW, H. AND KOCH-BRANDT, C., 2016. The CLU-files: disentanglement of a mystery. Biomolecular Concepts, **7** (1), 1-15.

ROHRS, J.A., SULISTIO, C.D. AND FINLEY, S.D., 2016. Predictive model of thrombospondin-1 and vascular endothelial growth factor in breast tumor tissue. NPJ Systems Biology and Applications, **2** (1), 1-11.

ROMÁN, V.R.G., MURRAY, J.C. AND WEINER, L.M., 2014, Antibody-dependent cellular cytotoxicity (ADCC). In: Antibody-dependent cellular cytotoxicity (ADCC). Antibody Fc. Elsevier, 2014, pp. 1-27.

ROSARIO, E., and ROSARIO, D.J., 2020. Localized Prostate Cancer.

ROSNER, M.H., VIGANO, M.A., OZATO, K., TIMMONS, P.M., POIRIE, F., RIGBY, P.W. and STAUDT, L.M., 1990. A POU-domain transcription factor in early stem cells and germ cells of the mammalian embryo. Nature, **345** (6277), 686-692.

ROZENGURT, E., 1999. Autocrine loops, signal transduction, and cell cycle abnormalities in the molecular biology of lung cancer. Current Opinion in Oncology, **11** (2), 116.

RUF, W., AND SCHAFFNER, F., 2014. Role of the protein C receptor in cancer progression. Thrombosis Research, **133**, S85-S89.

RUSCETTI, M., QUACH, B., DADASHIAN, E.L., MULHOLLAND, D.J. and WU, H., 2015. Tracking and functional characterization of epithelial–mesenchymal transition and mesenchymal tumor cells during prostate cancer metastasis. Cancer Research, **75** (13), 2749-2759.

SAATCI, O., KAYMAK, A., RAZA, U., ERSAN, P.G., AKBULUT, O., BANISTER, C.E., SIKIRZHYTSKI, V., TOKAT, U.M., AYKUT, G. AND ANSARI, S.A., 2020. Targeting lysyl oxidase (LOX) overcomes chemotherapy resistance in triple negative breast cancer. Nature Communications, **11** (1), 1-17.

SAHA, S.K., CHOI, H.Y., KIM, B.W., DAYEM, A.A., YANG, G.M., KIM, K.S., YIN, Y.F. AND CHO, S.G., 2017. KRT19 directly interacts with β-catenin/RAC1 complex to regulate NUMB-dependent NOTCH signaling pathway and breast cancer properties. Oncogene, **36** (3), 332-349.

SAKATO, M., O'DONNELL, M. AND HINGORANI, M.M., 2012. A central swivel point in the RFC clamp loader controls PCNA opening and loading on DNA. Journal of Molecular Biology, **416** (2), 163-175.

SARVIN, B., LAGZIEL, S., SARVIN, N., MUKHA, D., KUMAR, P., AIZENSHTEIN, E. AND SHLOMI, T., 2020. Fast and sensitive flow-injection mass spectrometry metabolomics by analyzing sample-specific ion distributions. Nature Communications, **11** (1), 1-11.

SATELLI, A., AND LI, S., 2011. Vimentin in cancer and its potential as a molecular target for cancer therapy. Cellular and Molecular Life Sciences, **68** (18), 3033-3046.

SAWICKA-GUTAJ, N., WALIGÓRSKA-STACHURA, J., ANDRUSIEWICZ, M., BICZYSKO, M., SOWIŃSKI, J., SKROBISZ, J. AND RUCHAŁA, M., 2015. Nicotinamide phosphorybosiltransferase overexpression in thyroid malignancies and its correlation with tumor stage and with survivin/survivin DEx3 expression. Tumor Biology, **36** (10), 7859-7863.

SCHAFFNER, F., YOKOTA, N., CARNEIRO-LOBO, T., KITANO, M., SCHAFFER, M., ANDERSON, G.M., MUELLER, B.M., ESMON, C.T. and RUF, W., 2013. Endothelial protein C receptor function in murine and human breast cancer development. PloS One, **8** (4), e61071.

SCHERBAKOV, A.M., ANDREEVA, O.E., SHATSKAYA, V.A. AND KRASIL'NIKOV, M.A., 2012. THE relationships between snail1 and estrogen receptor signaling in breast cancer cells. Journal of Cellular Biochemistry, **113** (6), 2147-2155.

SCHMIDT, A., BRIXIUS, K. and BLOCH, W., 2007. Endothelial precursor cell migration during vasculogenesis. Circulation Research, **101** (2), 125-136.

SCHOFIELD, R., 1978. The relationship between the spleen colony-forming cell and the haemopoietic stem cell. Blood Cells, **4** (1-2), 7-25.

SCHUG, Z.T., PECK, B., JONES, D.T., ZHANG, Q., GROSSKURTH, S., ALAM, I.S., GOODWIN, L.M., SMETHURST, E., MASON, S. AND BLYTH, K., 2015. Acetyl-CoA synthetase 2 promotes acetate utilization and maintains cancer cell growth under metabolic stress. Cancer Cell, **27** (1), 57-71.

SCHULZE, A.B., EVERS, G., KERKHOFF, A., MOHR, M., SCHLIEMANN, C., BERDEL, W.E. and SCHMIDT, L.H., 2019. Future options of molecular-targeted therapy in small cell lung cancer. Cancers, **11** (5), 690.

SCREENING, P., and BOARD, P.E., 2002. Prostate Cancer Screening (PDQ®): Patient Version. PDQ Cancer Information Summaries [Internet].

SEGHEZZI, G., PATEL, S., REN, C.J., GUALANDRIS, A., PINTUCCI, G., ROBBINS, E.S., SHAPIRO, R.L., GALLOWAY, A.C., RIFKIN, D.B. and MIGNATTI, P., 1998. Fibroblast growth factor-2 (FGF-2) induces vascular endothelial growth factor (VEGF) expression in the endothelial cells of forming capillaries: an autocrine mechanism contributing to angiogenesis. The Journal of Cell Biology, **141** (7), 1659-1673.

SEPHTON, S.E., SAPOLSKY, R.M., KRAEMER, H.C. AND SPIEGEL, D., 2000. Diurnal cortisol rhythm as a predictor of breast cancer survival. Journal of the National Cancer Institute, **92** (12), 994-1000.

SHAN, J., SHEN, J., LIU, L., XIA, F., XU, C., DUAN, G., XU, Y., MA, Q., YANG, Z. AND ZHANG, Q., 2012. Nanog regulates self-renewal of cancer stem cells through the insulin-like growth factor pathway in human hepatocellular carcinoma. Hepatology, **56** (3), 1004-1014.

SHANKARANARAYANAN, J.S., KANWAR, J.R., AL-JUHAISHI, A.J.A. AND KANWAR, R.K., 2016. Doxorubicin conjugated to immunomodulatory anticancer lactoferrin displays improved cytotoxicity overcoming prostate cancer chemo resistance and inhibits tumour development in TRAMP mice. Scientific Reports, **6** (1), 1-16.

SHIMKIN, M.B., 1977. Contrary to Nature: Being an Illustrated Commentary on Some Persons and Events of Historical Importance in the Development of Knowledge Concerning... Cancer. US Department of Health, Education, and Welfare, Public Health Service ….

SHOULDERS, M.D., AND RAINES, R.T., 2009. Collagen structure and stability. Annual Review of Biochemistry, **78**, 929.

SHRIHARI, T.G., 2017. Dual role of inflammatory mediators in cancer. Ecancermedicalscience, **11**.

SLAMON, D.J., LEYLAND-JONES, B., SHAK, S., FUCHS, H., PATON, V., BAJAMONDE, A., FLEMING, T., EIERMANN, W., WOLTER, J. and PEGRAM, M., 2001. Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. New England Journal of Medicine, **344** (11), 783-792.

SLOOT, Y.J., SMIT, J.W., JOOSTEN, L.A. AND NETEA-MAIER, R.T., 2018. Insights into the role of IL-32 in cancer. In: Seminars in immunology, Elsevier, pp. 24-32.

SMITH, L.A., CORNELIUS, V.R., PLUMMER, C.J., LEVITT, G., VERRILL, M., CANNEY, P. AND JONES, A., 2010. Cardiotoxicity of anthracycline agents for the treatment of cancer: systematic review and meta-analysis of randomised controlled trials. BMC Cancer, **10** (1), 1-14.

SMITH, R.W., 2013, Mass Spectrometry. In: J.A. Siegel, P.J. Saukko and M.M. Houck, eds., Encyclopedia of Forensic Sciences (Second Edition). Waltham: Academic Press, 2013, pp. 603-608.

SPERANDIO, S., DE BELLE, I. and BREDESEN, D.E., 2000. An alternative, nonapoptotic form of programmed cell death. Proceedings of the National Academy of Sciences**, 97** (26), 14376-14381.

SPRAGUE, B.L., GANGNON, R.E., BURT, V., TRENTHAM-DIETZ, A., HAMPTON, J.M., WELLMAN, R.D., KERLIKOWSKE, K. AND MIGLIORETTI, D.L., 2014. Prevalence of mammographically dense breasts in the United States. JNCI: Journal of the National Cancer Institute, **106** (10).

STEARNS-KUROSAWA, D.J., KUROSAWA, S., MOLLICA, J.S., FERRELL, G.L. and ESMON, C.T., 1996. The endothelial cell protein C receptor augments protein C activation by the thrombin-thrombomodulin complex. Proceedings of the National Academy of Sciences, **93** (19), 10212-10216.

STERLACCI, W., SAVIC, S., FIEGL, M., OBERMANN, E. AND TZANKOV, A., 2014. Putative stem cell markers in non–small-cell lung cancer: A Clinicopathologic characterization. Journal of Thoracic Oncology, **9** (1), 41-49.

SU, J., WU, S., WU, H., LI, L. and GUO, T., 2016. CD44 is functionally crucial for driving lung cancer stem cells metastasis through Wnt/β-catenin-FoxM1-Twist signaling. Molecular Carcinogenesis, **55** (12), 1962-1973.

SUN, L., GAO, J., DONG, X., LIU, M., LI, D., SHI, X., DONG, J., LU, X., LIU, C. AND ZHOU, J., 2008. EB1 promotes Aurora-B kinase activity through blocking its inactivation by protein phosphatase 2A. Proceedings of the National Academy of Sciences, **105** (20), 7153-7158.

SUN, Y., LIU, Y., MA, X. AND HU, H., 2021. The influence of cell cycle regulation on chemotherapy. International Journal of Molecular Sciences, **22** (13), 6923.

SUNG, H., FERLAY, J., SIEGEL, R.L., LAVERSANNE, M., SOERJOMATARAM, I., JEMAL, A. and BRAY, F., 2021. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: A Cancer Journal for Clinicians, **71** (3), 209-249.

SWISSHELM, K., RYAN, K., TSUCHIYA, K. AND SAGER, R., 1995. Enhanced expression of an insulin growth factor-like binding protein (mac25) in senescent human mammary epithelial cells and induced expression with retinoic acid. Proceedings of the National Academy of Sciences, **92** (10), 4472-4476.

TACAR, O., SRIAMORNSAK, P. AND DASS, C.R., 2013. Doxorubicin: an update on anticancer molecular action, toxicity and novel drug delivery systems. Journal of Pharmacy and Pharmacology, **65** (2), 157-170.

TAGUCHI, A., RHO, J., YAN, Q., ZHANG, Y., ZHAO, Y., XU, H., TRIPATHI, S.C., WANG, H., BRENNER, D.E. AND KUCHERLAPATI, M., 2015. MAPRE1 as a Plasma Biomarker for Early-Stage Colorectal Cancer and AdenomasMAPRE1 for Early Detection of Colorectal Cancer. Cancer Prevention Research, **8** (11), 1112-1119.

TAKAISHI, S., OKUMURA, T., TU, S., WANG, S.S., SHIBATA, W., VIGNESHWARAN, R., GORDON, S.A., SHIMADA, Y. and WANG, T.C., 2009. Identification of gastric cancer stem cells using the cell surface marker CD44. Stem Cells, **27** (5), 1006-1020.

TANG, J., GIFFORD, C.C., SAMARAKOON, R. AND HIGGINS, P.J., 2018. Deregulation of negative controls on TGF-β1 signaling in tumor progression. Cancers, **10** (6), 159.

TANNOCK, I.F., DE WIT, R., BERRY, W.R., HORTI, J., PLUZANSKA, A., CHI, K.N., OUDARD, S., THÉODORE, C., JAMES, N.D. AND TURESSON, I., 2004. Docetaxel plus prednisone or mitoxantrone plus prednisone for advanced prostate cancer. New England Journal of Medicine, **351** (15), 1502-1512.

TAYLOR, B.S., SCHULTZ, N., HIERONYMUS, H., GOPALAN, A., XIAO, Y., CARVER, B.S., ARORA, V.K., KAUSHIK, P., CERAMI, E. AND REVA, B., 2010. Integrative genomic profiling of human prostate cancer. Cancer Cell, **18** (1), 11-22.

TAYLOR, W.R., AND STARK, G.R., 2001. Regulation of the G2/M transition by p53. Oncogene, **20** (15), 1803-1815.

TEHRANIAN, N., SEPEHRI, H., MEHDIPOUR, P., BIRAMIJAMAL, F., HOSSEIN-NEZHAD, A., SARRAFNEJAD, A. AND HAJIZADEH, E., 2012. Combination effect of PectaSol and Doxorubicin on viability, cell cycle arrest and apoptosis in DU-145 and LNCaP prostate cancer cell lines. Cell Biology International, **36** (7), 601-610.

TEICHGRAEBER, D.C., GUIRGUIS, M.S. AND WHITMAN, G.J., 2021. Breast Cancer Staging: Updates in the AJCC Cancer Staging Manual, and Current Challenges for Radiologists, From the AJR Special Series on Cancer Staging. American Journal of Roentgenology, **217** (2), 278-290.

TENDULKAR, R.D., AGRAWAL, S., GAO, T., EFSTATHIOU, J.A., PISANSKY, T.M., MICHALSKI, J.M., KOONTZ, B.F., HAMSTRA, D.A., FENG, F.Y. AND LIAUW, S.L., 2016. Contemporary update of a multi-institutional predictive nomogram for salvage radiotherapy after radical prostatectomy. Journal of Clinical Oncology, **34** (30), 3648-3654.

TENG, M.W., SWANN, J.B., KOEBEL, C.M., SCHREIBER, R.D. and SMYTH, M.J., 2008. Immune-mediated dormancy: an equilibrium with cancer. Journal of Leukocyte Biology, **84** (4), 988-993.

THEDIECK, K., HOLZWARTH, B., PRENTZELL, M.T., BOEHLKE, C., KLÄSENER, K., RUF, S., SONNTAG, A.G., MAERZ, L., GRELLSCHEID, S. AND KREMMER, E., 2013. Inhibition of mTORC1 by astrin and stress granules prevents apoptosis in cancer cells. Cell**, 154** (4), 859-874.

THEIN, K.H., KLEYLEIN-SOHN, J., NIGG, E.A. AND GRUNEBERG, U., 2007. Astrin is required for the maintenance of sister chromatid cohesion and centrosome integrity. The Journal of Cell Biology, **178** (3), 345-354.

THIENPONT, B., VAN DYCK, L. AND LAMBRECHTS, D., 2016. Tumors smother their epigenome. Molecular & Cellular Oncology, **3** (6), e1240549.

THIERY, J.P., 2002. Epithelial–mesenchymal transitions in tumour progression. Nature Reviews Cancer, **2** (6), 442-454.

THOMSEN, S., and TATMAN, D., 1998. Physiological and Pathological Factors of Human Breast Disease That Can Influence Optical Diagnosis a. Annals of the New York Academy of Sciences, **838** (1), 171-193.

TOOLE, J.J., KNOPF, J.L., WOZNEY, J.M., SULTZMAN, L.A., BUECKER, J.L., PITTMAN, D.D., KAUFMAN, R.J., BROWN, E., SHOEMAKER, C. AND ORR, E.C., 1984. Molecular cloning of a cDNA encoding human antihaemophilic factor. Nature, **312** (5992), 342-347.

TOSH, D., AND SLACK, J.M., 2002. How cells change their phenotype. Nature Reviews Molecular Cell Biology, **3** (3), 187-194.

ULIVI, P., 2020. Predictive biomarkers in clinical practice: State of the art and perspectives in solid tumors. The International Journal of Biological Markers, **35** (1_suppl), 16-19.

VADAKEKOLATHU, J., AL-JUBOORI, S.I.K., JOHNSON, C., SCHNEIDER, A., BUCZEK, M.E., DI BIASE, A., POCKLEY, A.G., BALL, G.R., POWE, D.G. AND REGAD, T., 2018. MTSS1 and SCAMP1 cooperate to prevent invasion in breast cancer. Cell Death & Disease, **9** (3), 1-12.

VAN CUTSEM, E., KÖHNE, C., HITRE, E., ZALUSKI, J., CHANG CHIEN, C., MAKHSON, A., D'HAENS, G., PINTÉR, T., LIM, R. and BODOKY, G., 2009. Cetuximab and chemotherapy as initial treatment for metastatic colorectal cancer. New England Journal of Medicine, **360** (14), 1408-1417.

VAN NGUYEN, C., NGUYEN, Q.T., VU, H.T.N., PHAM, K.H. AND PHUNG, H.T., 2021. Molecular classification predicts survival for breast cancer patients in Vietnam: a single institutional retrospective analysis. International Journal of Clinical and Experimental Pathology, **14** (3), 322.

VANDER HEIDEN, M.G., 2011. Targeting cancer metabolism: a therapeutic window opens. Nature Reviews Drug Discovery, **10** (9), 671-684.

VANDER HEIDEN, M.G., CANTLEY, L.C. AND THOMPSON, C.B., 2009. Understanding the Warburg effect: the metabolic requirements of cell proliferation. Science, **324** (5930), 1029-1033.

VÉLEZ-CRUZ, R., and JOHNSON, D.G., 2017. The retinoblastoma (RB) tumor suppressor: pushing back against genome instability on multiple fronts. International Journal of Molecular Sciences, **18** (8), 1776.

VERMEULEN, K., VAN BOCKSTAELE, D.R. AND BERNEMAN, Z.N., 2003. The cell cycle: a review of regulation, deregulation and therapeutic targets in cancer. Cell Proliferation**, 36** (3), 131-149.

VISVADER, J.E., AND LINDEMAN, G.J., 2008. Cancer stem cells in solid tumours: accumulating evidence and unresolved questions. Nature Reviews Cancer, **8** (10), 755-768.

VISWANADHAPALLI, S., DILEEP, K.V., ZHANG, K.Y., NAIR, H.B. AND VADLAMUDI, R.K., 2021. Targeting LIF/LIFR signaling in cancer. Genes & Diseases, .

VOLLMER, S., STRICKSON, S., ZHANG, T., GRAY, N., LEE, K.L., RAO, V.R. AND COHEN, P., 2017. The mechanism of activation of IRAK1 and IRAK4 by interleukin-1 and Toll-like receptor agonists. Biochemical Journal, **474** (12), 2027-2038.

WACHTEL, M.S., NELIUS, T., HAYNES, A.L., DAHLBECK, S. AND DE RIESE, W., 2013. PSA screening and deaths from prostate cancer after diagnosis—a population based analysis. The Prostate, **73** (12), 1365-1369.

WAN, L., LIU, T., HONG, Z., PAN, Y., SIZEMORE, S.T., ZHANG, J. AND MA, Z., 2019. NEDD4 expression is associated with breast cancer progression and is predictive of a poor prognosis. Breast Cancer Research, **21** (1), 1-16.

WANG, B., HASAN, M.K., ALVARADO, E., YUAN, H., WU, H. AND CHEN, W.Y., 2011. NAMPT overexpression in prostate cancer and its contribution to tumor cell survival and stress response. Oncogene, **30** (8), 907-921.

WANG, B., KOHLI, J. AND DEMARIA, M., 2020. Senescent cells in cancer therapy: friends or foes? Trends in Cancer, **6** (10), 838-857.

WANG, D., LU, P., ZHANG, H., LUO, M., ZHANG, X., WEI, X., GAO, J., ZHAO, Z. AND LIU, C., 2014. Oct-4 and Nanog promote the epithelial-mesenchymal transition of breast cancer stem cells and are associated with poor prognosis in breast cancer patients. Oncotarget, **5** (21), 10803.

WANG, G., ZHOU, H., GU, Z., GAO, Q. AND SHEN, G., 2018. Oct4 promotes cancer cell proliferation and migration and leads to poor prognosis associated with the survivin/STAT3 pathway in hepatocellular carcinoma. Oncology Reports, **40** (2), 979-987.

WANG, H., ARUN, B.K., WANG, H., FULLER, G.N., ZHANG, W., MIDDLETON, L.P. AND SAHIN, A.A., 2008. IGFBP2 and IGFBP5 overexpression correlates with the lymph node metastasis in T1 breast carcinomas. The Breast Journal, **14** (3), 261-267.

WANG, Y., AND HERLYN, M., 2015. The emerging roles of Oct4 in tumor-initiating cells. American Journal of Physiology-Cell Physiology, **309** (11), C709-C718.

WANG, Z., GERSTEIN, M. AND SNYDER, M., 2009. RNA-Seq: a revolutionary tool for transcriptomics. Nature Reviews Genetics, **10** (1), 57-63.

WANG, Z., GUAN, D., WANG, S., CHAI, L.Y.A., XU, S. AND LAM, K., 2020. Glycolysis and oxidative phosphorylation play critical roles in natural killer cell receptor-mediated natural killer cell functions. Frontiers in Immunology, **11**, 202.

WEE, Z.N., YATIM, S.M.J., KOHLBAUER, V.K., FENG, M., GOH, J.Y., BAO, Y., LEE, P.L., ZHANG, S., WANG, P.P. AND LIM, E., 2015. IRAK1 is a therapeutic target that drives breast cancer metastasis and resistance to paclitaxel. Nature Communications, **6** (1), 1-16.

WEI, J.T., FENG, Z., PARTIN, A.W., BROWN, E., THOMPSON, I., SOKOLL, L., CHAN, D.W., LOTAN, Y., KIBEL, A.S. AND BUSBY, J.E., 2014. Can urinary PCA3 supplement PSA in the early detection of prostate cancer? Journal of Clinical Oncology, **32** (36), 4066.

WEI, J.T., FENG, Z., PARTIN, A.W., BROWN, E., THOMPSON, I., SOKOLL, L., CHAN, D.W., LOTAN, Y., KIBEL, A.S. and BUSBY, J.E., 2014. Can urinary PCA3 supplement PSA in the early detection of prostate cancer? Journal of Clinical Oncology, **32** (36), 4066.

WEI, Y., XIANG, H. AND ZHANG, W., 2022. Review of various NAMPT inhibitors for the treatment of cancer. Frontiers in Pharmacology, **13**.

WEN, Y., ENG, C.H., SCHMORANZER, J., CABRERA-POCH, N., MORRIS, E.J., CHEN, M., WALLAR, B.J., ALBERTS, A.S. AND GUNDERSEN, G.G., 2004. EB1 and APC bind to mDia to stabilize microtubules downstream of Rho and promote cell migration. Nature Cell Biology, **6** (9), 820-830.

WHITE, S.J., KASMAN, L.M., KELLY, M.M., LU, P., SPRUILL, L., MCDERMOTT, P.J. AND VOELKEL-JOHNSON, C., 2007. Doxorubicin generates a proapoptotic phenotype by phosphorylation of elongation factor 2. Free Radical Biology and Medicine, **43** (9), 1313-1321.

WHITTAKER, S.R., MALLINGER, A., WORKMAN, P. AND CLARKE, P.A., 2017. Inhibitors of cyclin-dependent kinases as cancer therapeutics. Pharmacology & Therapeutics, **173**, 83-105.

WHO, 2021. GLOBOCAN 2020 [online].Available at: https://gco.iarc.fr/today/data/factsheets/populations/826-united-kingdom-fact-sheets.pdf.Accessed [December,2021]

WILLIAMS, R.T., DEN BESTEN, W. AND SHERR, C.J., 2007. Cytokine-dependent imatinib resistance in mouse BCR-ABL , Arf-null lymphoblastic leukemia. Genes & Development, **21** (18), 2283-2287.

WOJTUKIEWICZ, M.Z., HEMPEL, D., SIERKO, E., TUCKER, S.C. AND HONN, K.V., 2019. Endothelial protein C receptor (EPCR), protease activated receptor-1 (PAR-1) and their interplay in cancer growth and metastatic dissemination. Cancers, **11** (1), 51.

WOOSTER, R., AND WEBER, B.L., 2003. Breast and ovarian cancer. New England Journal of Medicine, 348 (23), 2339-2347.

WOOSTER, R., BIGNELL, G., LANCASTER, J., SWIFT, S., SEAL, S., MANGION, J., COLLINS, N., GREGORY, S., GUMBS, C. AND MICKLEM, G., 1995. Identification of the breast cancer susceptibility gene BRCA2. Nature, **378** (6559), 789-792.

XIANG, Y., YANG, T., PANG, B., ZHU, Y. AND LIU, Y., 2016. The progress and prospects of putative biomarkers for liver cancer stem cells in hepatocellular carcinoma. Stem Cells International, 2016.

XIE, M., ZHANG, L., HE, C., XU, F., LIU, J., HU, Z., ZHAO, L. AND TIAN, Y., 2012. Activation of notch-1 enhances epithelial–mesenchymal transition in gefitinib-acquired resistant lung cancer cells. Journal of Cellular Biochemistry, **113** (5), 1501-1513.

XIN, W., RHODES, D.R., INGOLD, C., CHINNAIYAN, A.M. AND RUBIN, M.A., 2003. Dysregulation of the annexin family protein family is associated with prostate cancer progression. The American Journal of Pathology, **162** (1), 255-261.

XU, F., DAI, C., ZHANG, R., ZHAO, Y., PENG, S. AND JIA, C., 2012. Nanog: a potential biomarker for liver metastasis of colorectal cancer. Digestive Diseases and Sciences, **57** (9), 2340-2346.

XU, H., TIAN, Y., YUAN, X., WU, H., LIU, Q., PESTELL, R.G. AND WU, K., 2015. The role of CD44 in epithelial–mesenchymal transition and cancer development. OncoTargets and Therapy, **8,** 3783.

YANG, H., XIA, L., CHEN, J., ZHANG, S., MARTIN, V., LI, Q., LIN, S., CHEN, J., CALMETTE, J. AND LU, M., 2019. Stress–glucocorticoid–TSC22D3 axis compromises therapy-induced antitumor immunity. Nature Medicine, **25** (9), 1428-1441.

YANG, J., ANTIN, P., BERX, G., BLANPAIN, C., BRABLETZ, T., BRONNER, M., CAMPBELL, K., CANO, A., CASANOVA, J. AND CHRISTOFORI, G., 2020. Guidelines and definitions for research on epithelial–mesenchymal transition. Nature Reviews Molecular Cell Biology, **21** (6), 341-352.

YANG, J., MANI, S.A., DONAHER, J.L., RAMASWAMY, S., ITZYKSON, R.A., COME, C., SAVAGNER, P., GITELMAN, I., RICHARDSON, A. AND WEINBERG, R.A., 2004. Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. Cell, **117** (7), 927-939.

YANG, L., XU, J., KANG, Q., LI, A., JIN, P., WANG, X., HE, Y., LI, N., CHENG, T. AND SHENG, J., 2017. Predictive value of stemness factor Sox2 in gastric cancer is associated with tumor location and stage. PLoS One, **12** (1), e0169124.

YANG, Q., DU, J. AND ZU, L., 2013. Overexpression of CD73 in prostate cancer is associated with lymph node metastasis. Pathology & Oncology Research, **19** (4), 811-814.

YANG, Q., XIE, B., TANG, H., MENG, W., JIA, C., ZHANG, X., ZHANG, Y., ZHANG, J., LI, H. AND FU, B., 2019. Minichromosome maintenance 3 promotes hepatocellular carcinoma radioresistance by activating the NF-κB pathway. Journal of Experimental & Clinical Cancer Research, **38** (1), 1-12.

YANG, Y., CHENG, J., WANG, S. AND YANG, H., 2022. StatsPro: Systematic integration and evaluation of statistical approaches for detecting differential expression in label-free quantitative proteomics. Journal of Proteomics, **250**, 104386.

YANO, S., TAZAWA, H., KAGAWA, S., FUJIWARA, T. AND HOFFMAN, R.M., 2020. FUCCI real-time cell-cycle imaging as a guide for designing improved cancer therapy: A review of innovative strategies to target quiescent chemo-resistant cancer cells. Cancers, **12** (9), 2655.

YAO, Y., AND WANG, C., 2020. Dedifferentiation: Inspiration for devising engineering strategies for regenerative medicine. NPJ Regenerative Medicine, **5** (1), 1-11.

YEUNG, M.L., AND JEANG, K., 2011. MicroRNAs and cancer therapeutics. Pharmaceutical Research, 28 (12), 3043-3049.

YIN, A.H., MIRAGLIA, S., ZANJANI, E.D., ALMEIDA-PORADA, G., OGAWA, M., LEARY, A.G., OLWEUS, J., KEARNEY, J. and BUCK, D.W., 1997. AC133, a novel marker for human hematopoietic stem and progenitor cells. Blood, the Journal of the American Society of Hematology, **90** (12), 5002-5012.

YIN, X., LI, Y., ZHANG, B., REN, Z., QIU, S., YI, Y. AND FAN, J., 2012. Coexpression of stemness factors Oct4 and Nanog predict liver resection. Annals of Surgical Oncology, **19** (9), 2877-2887.

YOSSEPOWITCH, O., BRIGANTI, A., EASTHAM, J.A., EPSTEIN, J., GRAEFEN, M., MONTIRONI, R. AND TOUIJER, K., 2014. Positive surgical margins after radical prostatectomy: a systematic review and contemporary update. European Urology, **65** (2), 303-313.

YUAN, L., LI, J.D., ZHANG, L., WANG, J.H., WAN, T., ZHOU, Y., TU, H., YUN, J.P., LUO, R.Z. AND JIA, W.H., 2014. SPAG5 upregulation predicts poor prognosis in cervical cancer patients and alters sensitivity to taxol treatment via the mTOR signaling pathway. Cell Death & Disease, **5** (5), e1247.

YUAN, S., NORGARD, R.J. AND STANGER, B.Z., 2019. Cellular Plasticity in CancerCancer Cells Change Identity during Tumor Progression. Cancer Discovery, **9** (7), 837-851.

ZACKRISSON, S., LÅNG, K., ROSSO, A., JOHNSON, K., DUSTLER, M., FÖRNVIK, D., FÖRNVIK, H., SARTOR, H., TIMBERG, P. AND TINGBERG, A., 2018. One-view breast tomosynthesis versus two-view mammography in the Malmö Breast Tomosynthesis Screening Trial (MBTST): a prospective, population-based, diagnostic accuracy study. The Lancet Oncology, **19** (11), 1493-1503.

ZENG, T., WANG, Q., FU, J., LIN, Q., BI, J., DING, W., QIAO, Y., ZHANG, S., ZHAO, W. AND LIN, H., 2014. Impeded Nedd4-1-mediated Ras degradation underlies Ras-driven tumorigenesis. Cell Reports, **7** (3), 871-882.

ZENG, Y., AND CHEN, T., 2019. DNA methylation reprogramming during mammalian development. Genes, **10** (4), 257.

ZHANG, D., SUN, B., ZHAO, X., CUI, Y., XU, S., DONG, X., ZHAO, J., MENG, J., JIA, X. AND CHI, J., 2012. Secreted CLU is associated with the initiation of triple-negative breast cancer. Cancer Biology & Therapy, **13** (5), 321-329.

ZHANG, H., BROWN, R.L., WEI, Y., ZHAO, P., LIU, S., LIU, X., DENG, Y., HU, X., ZHANG, J. and GAO, X.D., 2019. CD44 splice isoform switching determines breast cancer stem cell state. Genes & Development, **33** (3-4), 166-179.

ZHANG, H., LI, S., YANG, X., QIAO, B., ZHANG, Z. AND XU, Y., 2016. miR-539 inhibits prostate cancer progression by directly targeting SPAG5. Journal of Experimental & Clinical Cancer Research, **35** (1), 1-9.

ZHANG, M., SHA, L., HOU, N., SHI, C. AND TAN, L., 2020. High expression of sperm-associated antigen 5 correlates with poor survival in ovarian cancer. Bioscience Reports, **40** (2).

ZHANG, P., HE, Q., WANG, Y., ZHOU, G., CHEN, Y., TANG, L., ZHANG, Y., HONG, X., MAO, Y. AND HE, Q., 2022. Protein C receptor maintains cancer stem cell properties via activating lipid synthesis in nasopharyngeal carcinoma. Signal Transduction and Targeted Therapy, **7** (1), 1-11.

ZHANG, S., BALCH, C., CHAN, M.W., LAI, H., MATEI, D., SCHILDER, J.M., YAN, P.S., HUANG, T.H. AND NEPHEW, K.P., 2008. Identification and characterization of ovarian cancer-initiating cells from primary human tumors. Cancer Research, **68** (11), 4311-4320.

ZHANG, W., YI, B., WANG, C., CHEN, D., BAE, S., WEI, S., GUO, R., LU, C., NGUYEN, L.L. AND YANG, W., 2016. Silencing of CD24 Enhances the PRIMA-1–Induced Restoration of Mutant p53 in Prostate Cancer Cells. Clinical Cancer Research, **22** (10), 2545-2554.

ZHANG, X., HUA, R., WANG, X., HUANG, M., GAN, L., WU, Z., ZHANG, J., WANG, H., CHENG, Y. AND LI, J., 2016. Identification of stem-like cells and clinical significance of candidate stem cell markers in gastric cancer. Oncotarget, **7** (9), 9815.

ZHOU, X., JIA, L., SUN, Y., XU, L., WANG, X. AND TANG, Q., 2019. Sperm associated antigen 5 is a potential biomarker for poor prognosis in breast cancer. Oncology Letters, **17** (1), 1146-1152.

ZHOU, Y., ZHOU, B., PACHE, L., CHANG, M., KHODABAKHSHI, A.H., TANASEICHUK, O., BENNER, C. AND CHANDA, S.K., 2019. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nature Communications, **10** (1), 1-10.

ZHU, C., MENYHART, O., GYŐRFFY, B. AND HE, X., 2019. The prognostic association of SPAG5 gene expression in breast cancer patients with systematic therapy. BMC Cancer, **19** (1), 1-12.

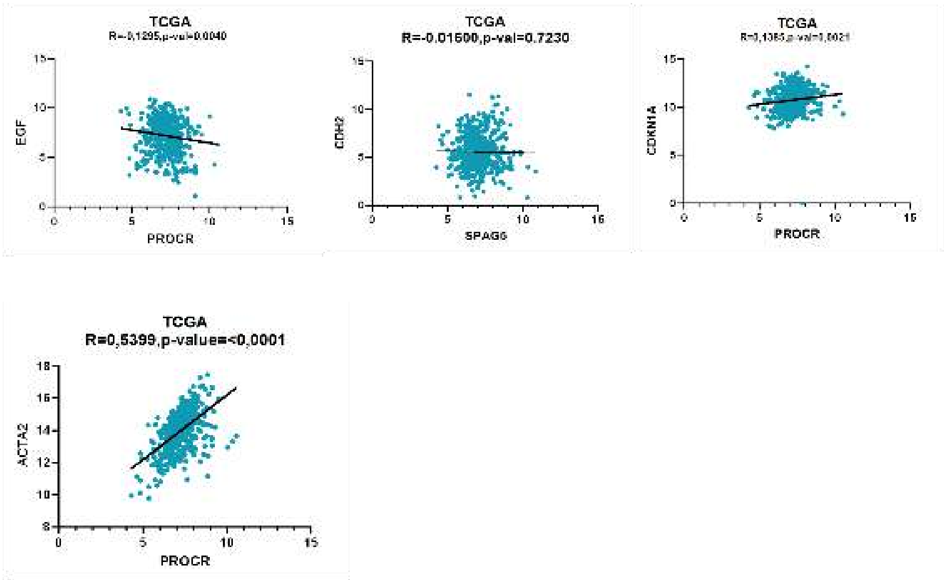
ZIAEE, S., CHU, G.C., HUANG, J., SIEH, S. and CHUNG, L.W., 2015. Prostate cancer metastasis: roles of recruitment and reprogramming, cell signal network and three-dimensional growth characteristics. Translational Andrology and Urology, **4** (4), 438.

Appendix

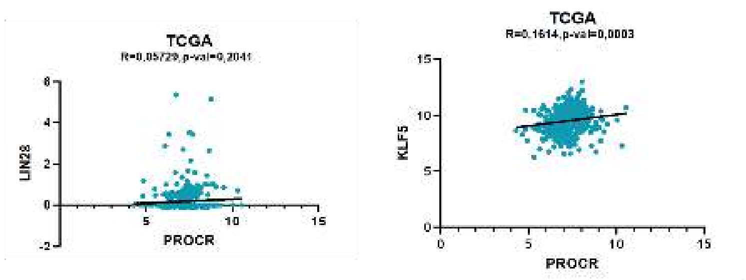
Appendix 2

Correlation analysis for EPCR from three independent consortia. RNA-sequencing data was obtained form cBioportal

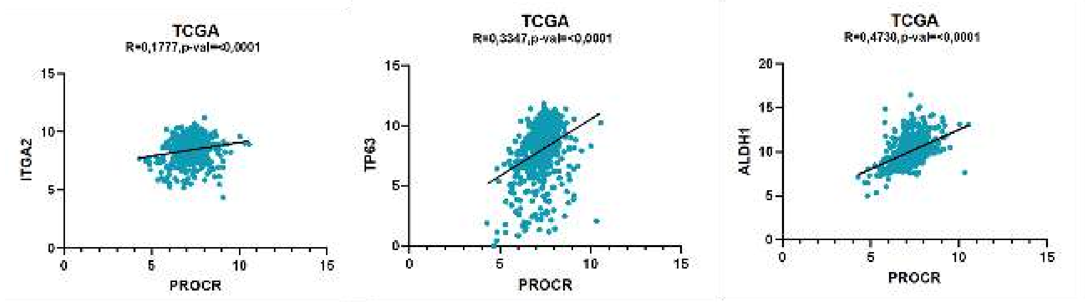
EMT marker



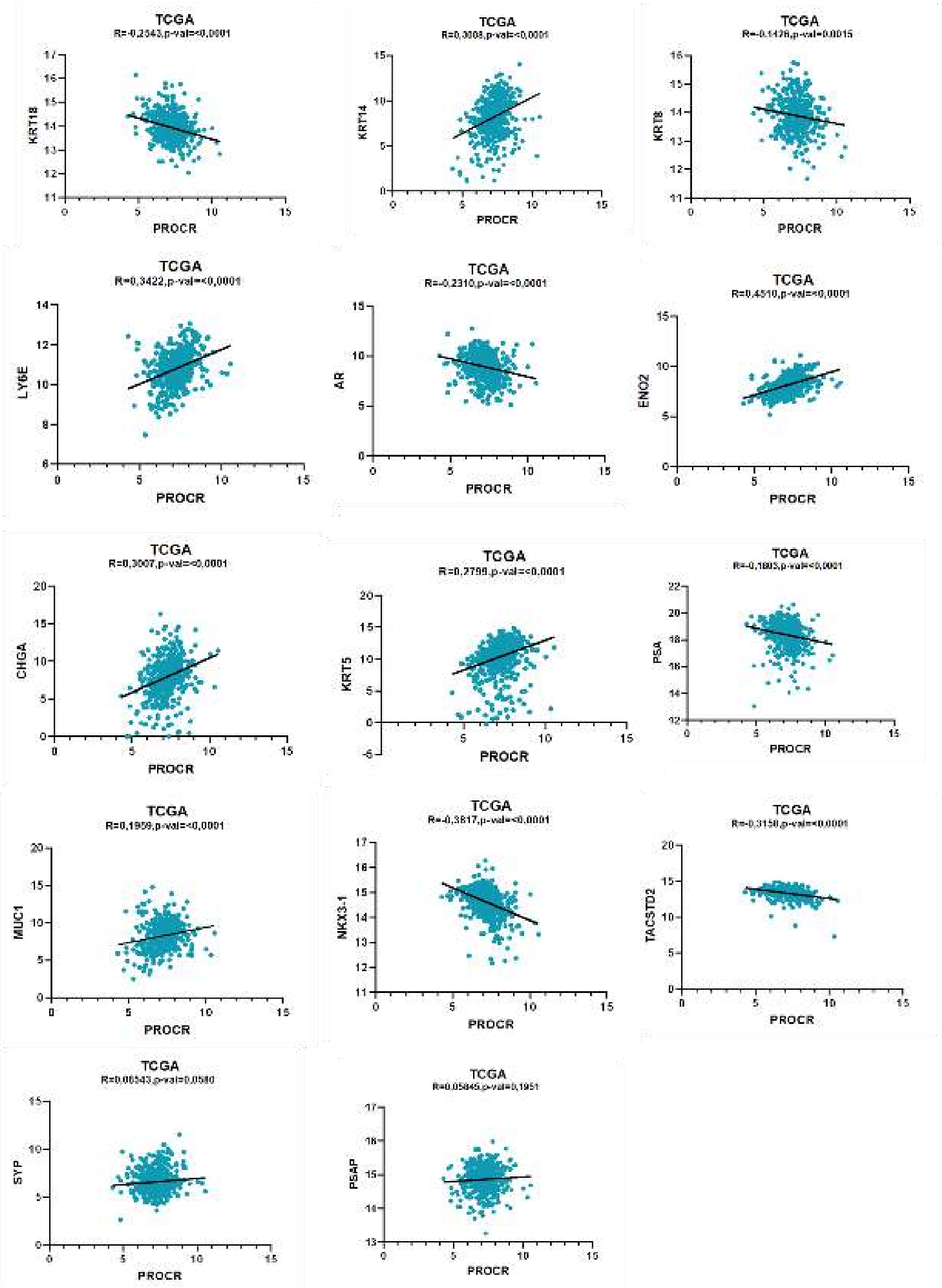
Embryonic stem cell marker



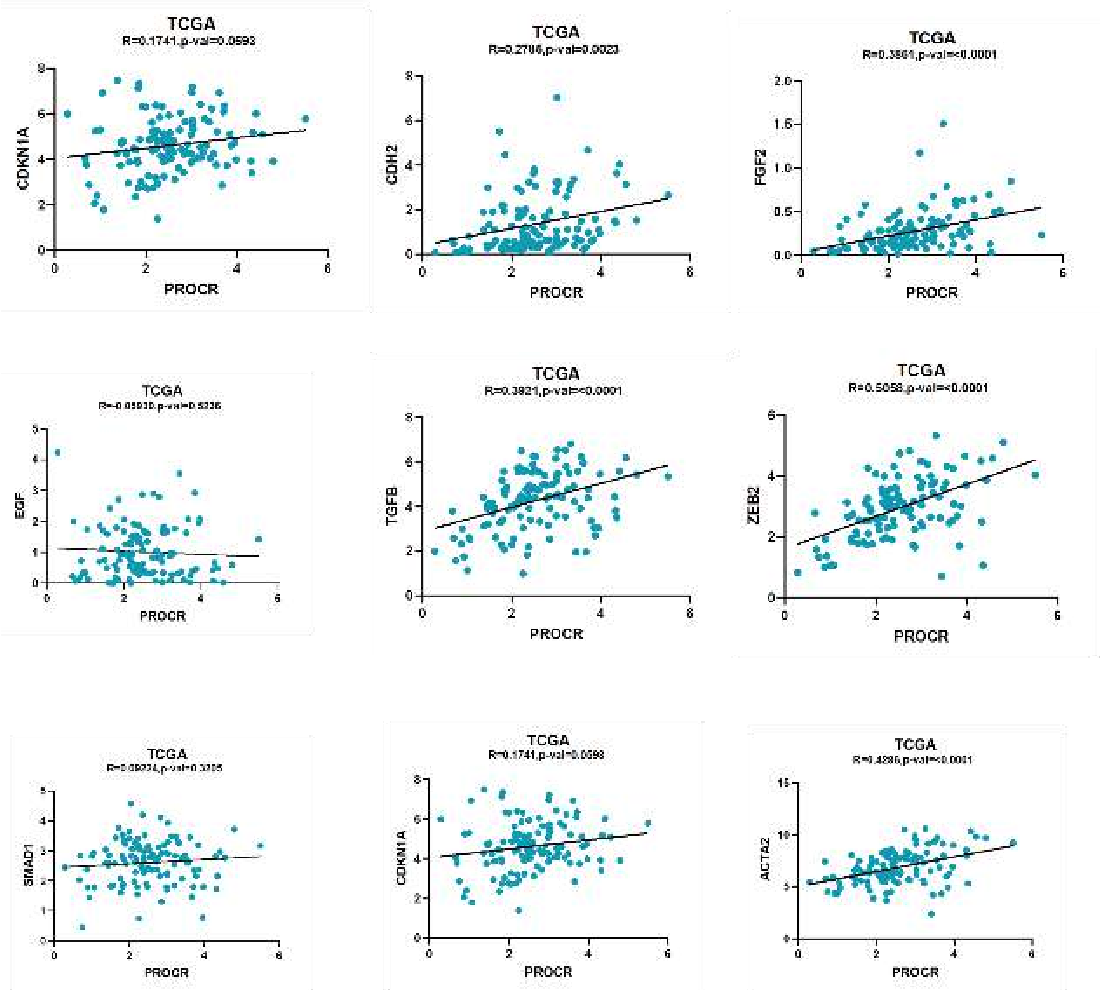
Cancer stem cell



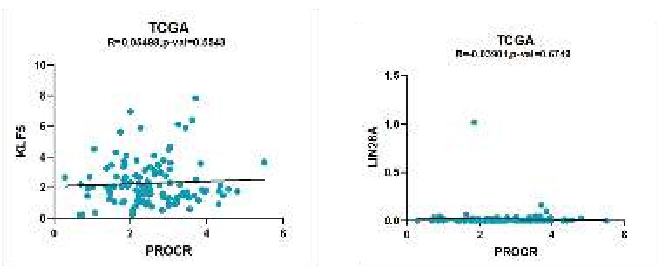
Prognostic Markers



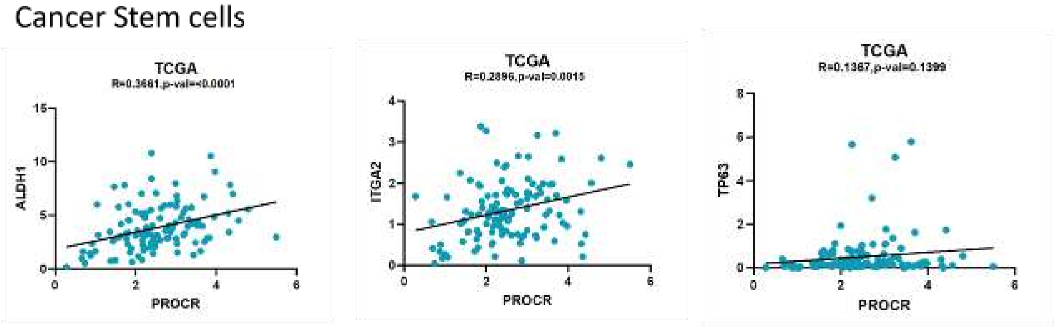
EMT markers TCGA 2015



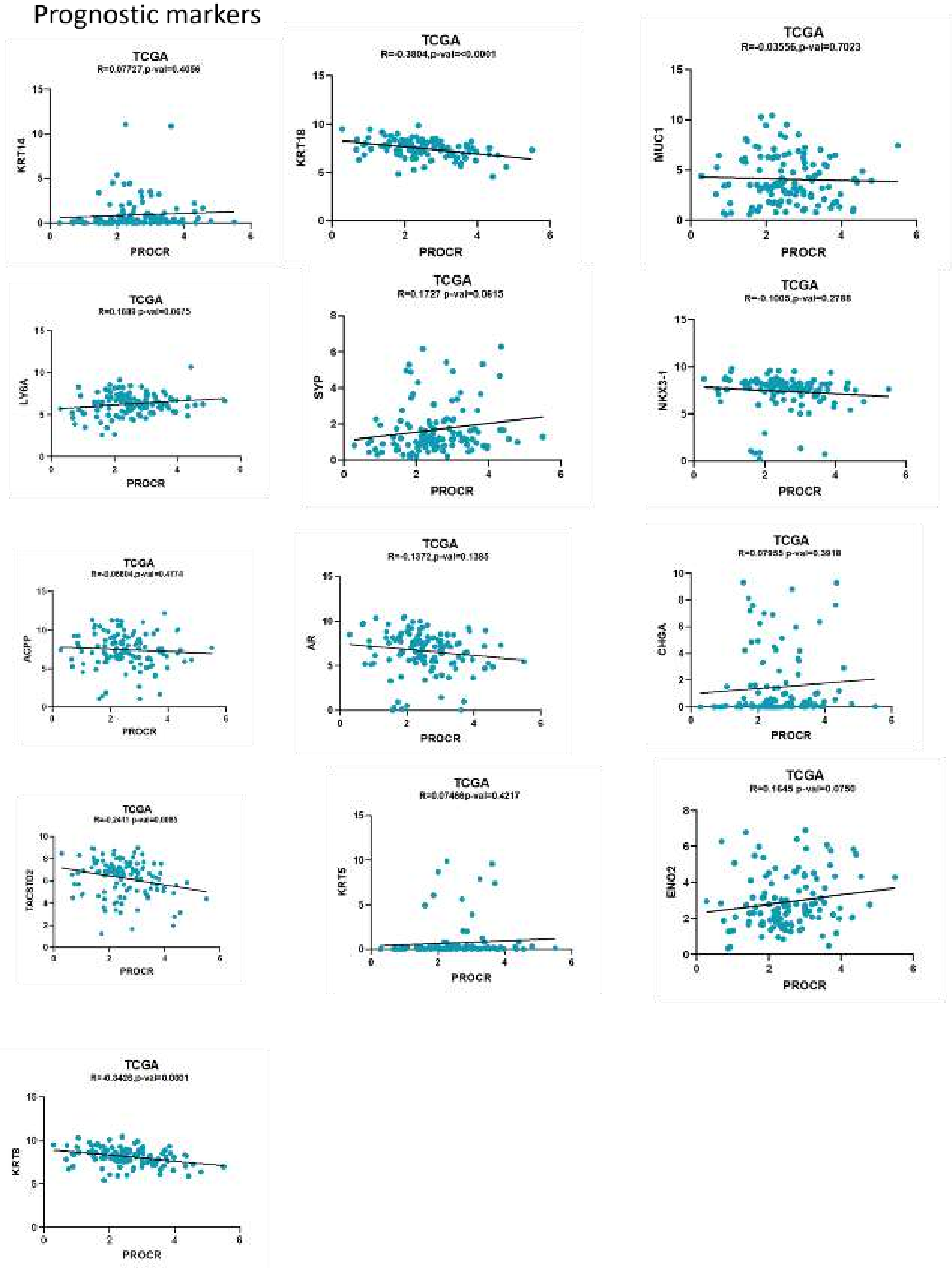
Embryonic stem cells markers



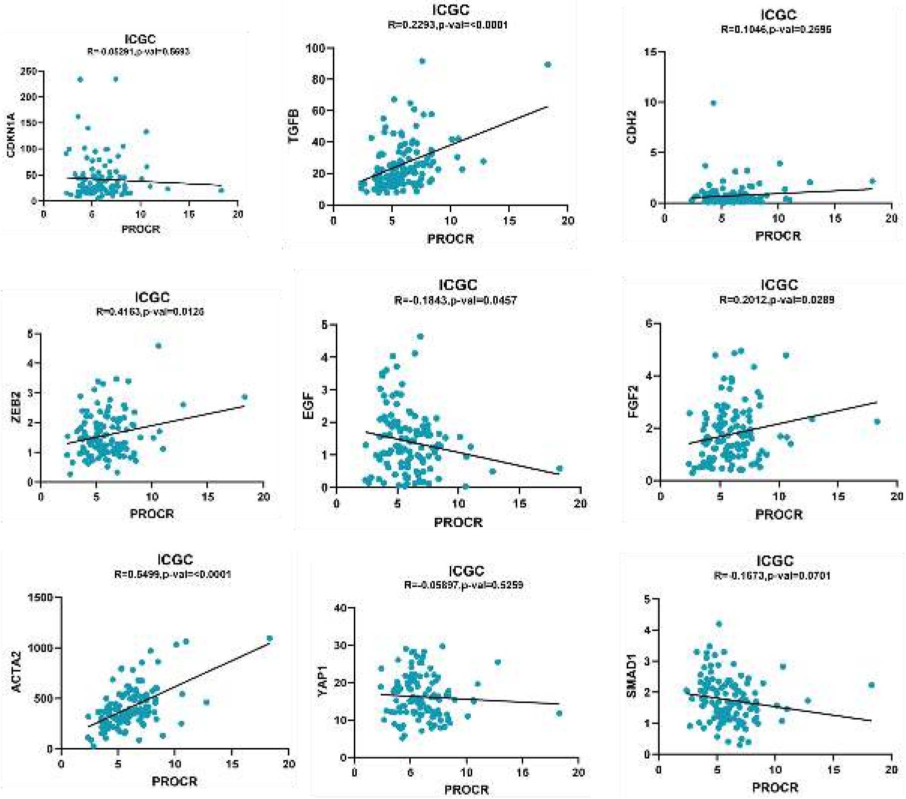
Cancer Stem cells



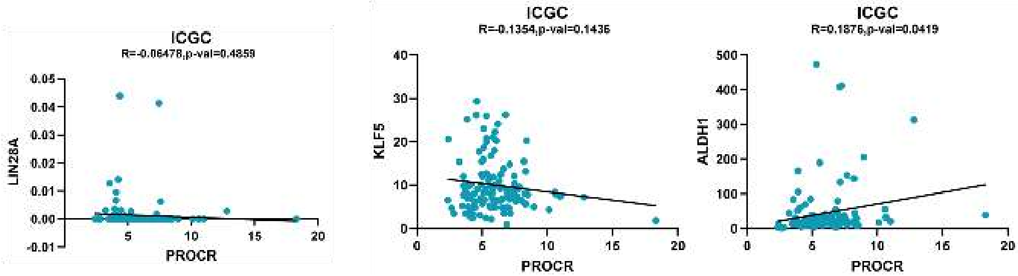
Prognostic markers



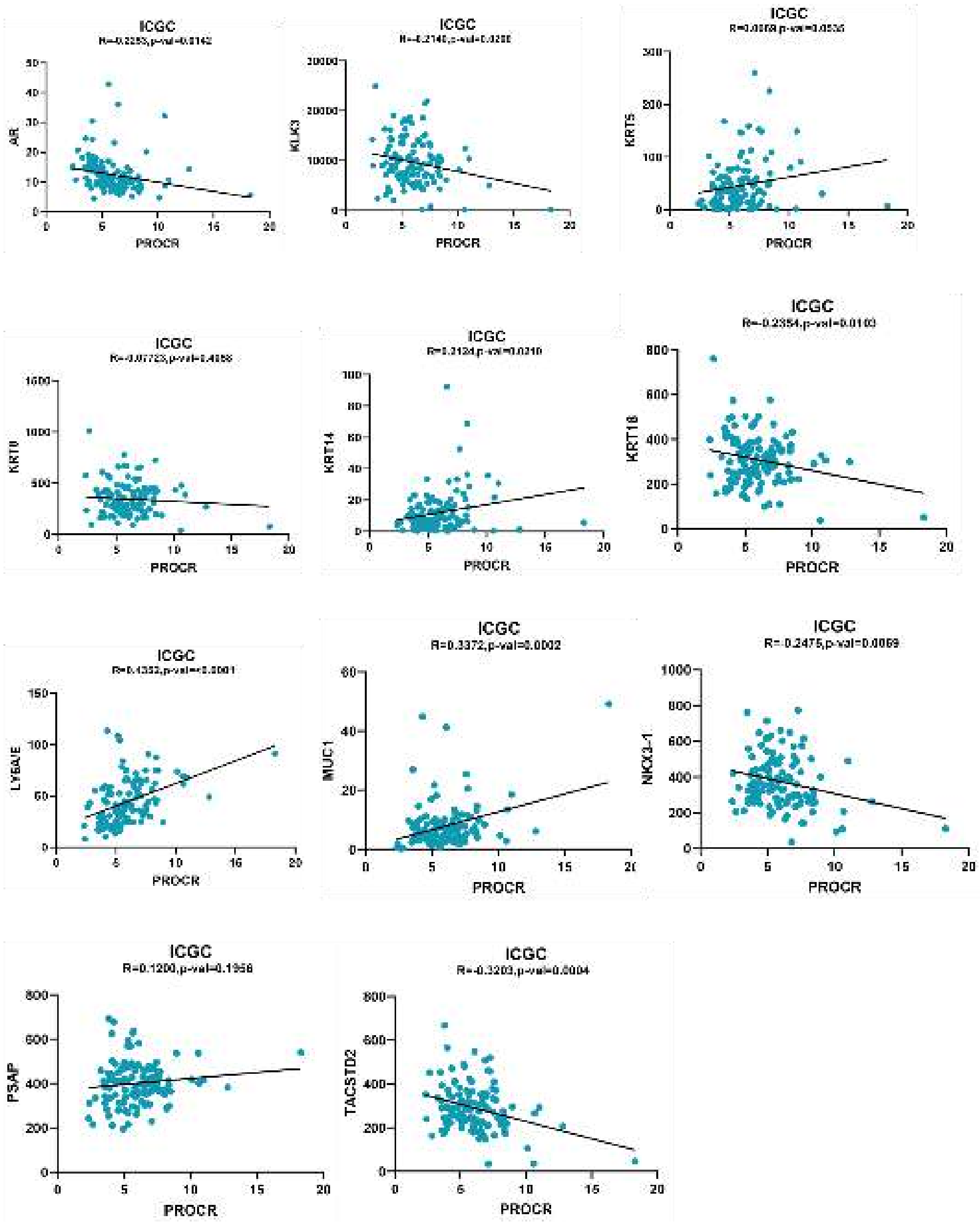
EMT markers ICGC



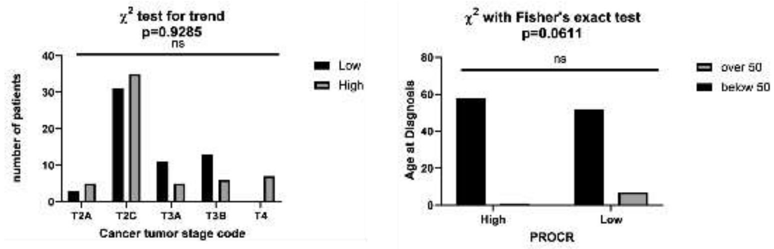
Embryonic stem cell markers ICGC



Prognostic correlation PROCR



Clinical analysis data PROCR ICGC



Appendix 3

Sample comparisons with genome reference.

sample	total reads	total map	unique map	Multimap	read1_map	read2_map	positive map	negative map	splice map	Unsplice map	proper map
DUP_1	48588504	46750361(96.22%)	45479546(93.6%)	1270815(2.62%)	22712703(46.75%)	22766843(46.86%)	22717644(46.76%)	22761902(46.85%)	19933898(41.03%)	25545648(52.58%)	44374210(91.33%)
DUP_2	42386266	40764391(96.17%)	39657046(93.56%)	1107345(2.61%)	19801164(46.72%)	19855882(46.85%)	19806067(46.73%)	19850979(46.83%)	17459996(41.19%)	22197050(52.37%)	38641764(91.17%)
DUP_3	51383676	49450256(96.24%)	48125106(93.66%)	1325150(2.58%)	24051652(46.81%)	24073454(46.85%)	24038118(46.78%)	24086988(46.88%)	21255215(41.37%)	26869891(52.29%)	46956704(91.38%)
DUsh_1	47226134	45519903(96.39%)	44309778(93.82%)	1210125(2.56%)	22121653(46.84%)	22188125(46.98%)	22131181(46.86%)	22178597(46.96%)	19100932(40.45%)	25208846(53.38%)	43197226(91.47%)
DUsh_2	60288472	58067469(96.32%)	56498736(93.71%)	1568733(2.6%)	28192249(46.76%)	28306487(46.95%)	28219569(46.81%)	28279167(46.91%)	24495397(40.63%)	32003339(53.08%)	55046016(91.3%)
DUsh_3	44675280	42981667(96.21%)	41829949(93.63%)	1151718(2.58%)	20883097(46.74%)	20946852(46.89%)	20891101(46.76%)	20938848(46.87%)	17985859(40.26%)	23844090(53.37%)	40761650(91.24%)
MDAP_1	51714506	49423853(95.57%)	48169111(93.14%)	1254742(2.43%)	24061833(46.53%)	24107278(46.62%)	24060402(46.53%)	24108709(46.62%)	20934180(40.48%)	27234931(52.66%)	46981360(90.85%)
MDAP_2	59654294	57144725(95.79%)	55692104(93.36%)	1452621(2.44%)	27830501(46.65%)	27861603(46.71%)	27817193(46.63%)	27874911(46.73%)	23959099(40.16%)	31733005(53.19%)	54313440(91.05%)
MDAP_3	49347446	47398884(96.05%)	46228043(93.68%)	1170841(2.37%)	23097176(46.81%)	23130867(46.87%)	23090757(46.79%)	23137286(46.89%)	19879260(40.28%)	26348783(53.39%)	45073866(91.34%)
MDAsh_1	49873432	47940715(96.12%)	46735025(93.71%)	1205690(2.42%)	23343056(46.8%)	23391969(46.9%)	23343715(46.81%)	23391310(46.9%)	20271597(40.65%)	26463428(53.06%)	45589784(91.41%)
MDAsh_2	55115772	52977842(96.12%)	51616376(93.65%)	1361466(2.47%)	25790715(46.79%)	25825661(46.86%)	25781893(46.78%)	25834483(46.87%)	21962312(39.85%)	29654064(53.8%)	50364158(91.38%)
MDAsh_3	47342646	45751050(96.64%)	44656178(94.33%)	1094872(2.31%)	22306381(47.12%)	22349797(47.21%)	22305416(47.11%)	22350762(47.21%)	19242447(40.65%)	25413731(53.68%)	43552368(91.99%)

Table 3.1 – Differentially expressed genes (DEGs) in MDA-MB-231 SPAG5 knockdown. The tables present the list of the 1,121 the most upregulated and 1,080 downregulated genes out of 2201 identified, obtained applying Log2FC and a cut-off of 0.58. Gene listed in orange are upregulated in knockdown vs control cell populations, whereas the blue table are genes downregulated in knockdown vs control cell populations.

Gene ID	Log2FC	Pvalue	Gene Description
CTSD	1.100999	0	cathepsin D [HGNC:2529]
CLU	1.227572	1.06E-223	clusterin [HGNC:2095]
COL5A1	1.371220	4.29E-216	collagen type V alpha 1 chain [HGNC:2209]
TGFBI	1.776142	6.86E-209	transforming growth factor beta induced [HGNC:11771]
IGFBP7	0.811291	6.05E-187	insulin like growth factor binding protein 7 [HGNC:5476]
KRT19	0.976821	2.36E-168	keratin 19 [HGNC:6436]

ACSL1	1.309492	8.39E-161	acyl-CoA synthetase long chain family member 1 [HGNC:3569]
PTTG1IP	0.809386	9.22E-160	PTTG1 interacting protein [HGNC:13524]
GSN	0.920503	5.25E-148	gelsolin [HGNC:4620]
FADS2	0.848034	1.54E-138	fatty acid desaturase 2 [HGNC:3575]
LGALS3BP	0.770571	1.85E-137	galectin 3 binding protein [HGNC:6564]
TACSTD2	1.236221	8.88E-125	tumor associated calcium signal transducer 2 [HGNC:11530]
HTRA1	1.132950	1.09E-122	HtrA serine peptidase 1 [HGNC:9476]
CERCAM	1.290603	4.58E-122	cerebral endothelial cell adhesion molecule [HGNC:23723]
MXRA8	1.329760	4.72E-117	matrix remodeling associated 8 [HGNC:7542]
TMSB4X	0.815413	6.10E-115	thymosin beta 4 X-linked [HGNC:11881]
LSS	0.949232	5.77E-102	lanosterol synthase [HGNC:6708]
CLIC3	1.500519	1.59E-100	chloride intracellular channel 3 [HGNC:2064]
FAM3C	0.853764	2.58E-99	family with sequence similarity 3 member C [HGNC:18664]
LOX	1.277069	1.46E-97	lysyl oxidase [HGNC:6664]
ATP1A1	0.651910	4.07E-97	ATPase Na+/K+ transporting subunit alpha 1 [HGNC:799]
PRSS23	0.671656	2.21E-95	serine protease 23 [HGNC:14370]
DHCR7	0.911368	6.41E-95	7-dehydrocholesterol reductase [HGNC:2860]
CYBRD1	0.837705	7.00E-94	cytochrome b reductase 1 [HGNC:20797]
TIMP1	0.624589	3.07E-85	TIMP metallopeptidase inhibitor 1 [HGNC:11820]
CTSA	0.822385	3.30E-85	cathepsin A [HGNC:9251]
MELTF	0.892454	4.20E-84	melanotransferrin [HGNC:7037]
TUBA1A	1.079727	3.61E-83	tubulin alpha 1a [HGNC:20766]
CPNE7	2.637410	2.74E-82	copine 7 [HGNC:2320]
LPIN1	0.939809	1.94E-80	lipin 1 [HGNC:13345]
ACSS2	0.987903	3.52E-73	acyl-CoA synthetase short chain family member 2 [HGNC:15814]
SARDH	1.251055	4.68E-72	sarcosine dehydrogenase [HGNC:10536]
CPOX	0.775519	6.35E-72	coproporphyrinogen oxidase [HGNC:2321]
MYH9	0.610003	4.68E-70	myosin heavy chain 9 [HGNC:7579]
PLEKHG4	0.928493	4.06E-68	pleckstrin homology and RhoGEF domain containing G4 [HGNC:24501]
EFEMP1	0.920176	1.60E-67	EGF containing fibulin extracellular matrix protein 1 [HGNC:3218]
SLC25A23	0.875694	2.33E-67	solute carrier family 25 member 23 [HGNC:19375]
UACA	0.649291	3.79E-67	uveal autoantigen with coiled-coil domains and ankyrin repeats [HGNC:15947]
BST2	0.769066	8.16E-67	bone marrow stromal cell antigen 2 [HGNC:1119]
FRAS1	1.440813	3.34E-65	Fraser extracellular matrix complex subunit 1 [HGNC:19185]
MGST1	0.877180	5.19E-65	microsomal glutathione S-transferase 1 [HGNC:7061]
DPP7	0.795920	1.06E-64	dipeptidyl peptidase 7 [HGNC:14892]
ACLY	0.636034	2.00E-64	ATP citrate lyase [HGNC:115]
ID2	3.319291	5.35E-63	inhibitor of DNA binding 2 [HGNC:5361]
LLGL2	1.232299	1.41E-62	LLGL2, scribble cell polarity complex component [HGNC:6629]
CALD1	0.644994	1.58E-62	caldesmon 1 [HGNC:1441]
CDH11	0.865862	1.04E-61	cadherin 11 [HGNC:1750]
SLC2A6	1.044254	5.41E-61	solute carrier family 2 member 6 [HGNC:11011]
CALCR	2.125815	2.72E-60	calcitonin receptor [HGNC:1440]
SLC25A1	0.633322	3.55E-60	solute carrier family 25 member 1 [HGNC:10979]
EGFL7	1.048705	9.23E-59	EGF like domain multiple 7 [HGNC:20594]

ABCC3	1.052087	1.11E-58	ATP binding cassette subfamily C member 3 [HGNC:54]
EDIL3	0.968263	1.49E-58	EGF like repeats and discoidin domains 3 [HGNC:3173]
MYDGF	0.657372	2.26E-58	myeloid derived growth factor [HGNC:16948]
IDH1	0.865755	4.05E-58	isocitrate dehydrogenase (NADP(+)) 1, cytosolic [HGNC:5382]
MYO18A	0.635633	1.07E-57	myosin XVIIIA [HGNC:31104]
LIPA	0.811867	2.40E-57	lipase A, lysosomal acid type [HGNC:6617]
CMTM7	0.875621	2.54E-57	CKLF like MARVEL transmembrane domain containing 7 [HGNC:19178]
TMEM250	0.770166	4.99E-57	transmembrane protein 250 [HGNC:31009]
DHCR24	0.580105	5.48E-57	24-dehydrocholesterol reductase [HGNC:2859]
PCYT2	0.718028	8.00E-57	phosphate cytidylyltransferase 2, ethanolamine [HGNC:8756]
PEG10	0.721566	1.01E-56	paternally expressed 10 [HGNC:14005]
NUPR1	2.625645	4.42E-56	nuclear protein 1, transcriptional regulator [HGNC:29990]
NSMF	0.794077	5.53E-56	NMDA receptor synaptonuclear signaling and neuronal migration factor [HGNC:29843]
ACSL5	0.667706	7.37E-56	acyl-CoA synthetase long chain family member 5 [HGNC:16526]
GPX4	0.586173	2.06E-55	glutathione peroxidase 4 [HGNC:4556]
SLC2A3	0.968973	2.72E-55	solute carrier family 2 member 3 [HGNC:11007]
ABCG1	1.354695	3.80E-55	ATP binding cassette subfamily G member 1 [HGNC:73]
ABCG2	1.527271	6.01E-55	ATP binding cassette subfamily G member 2 (Junior blood group) [HGNC:74]
ALDH3B1	0.970965	1.24E-54	aldehyde dehydrogenase 3 family member B1 [HGNC:410]
RXRA	0.722606	1.71E-54	retinoid X receptor alpha [HGNC:10477]
B4GALNT4	1.788899	1.74E-54	beta-1,4-N-acetyl-galactosaminyltransferase 4 [HGNC:26315]
VSTM2L	0.910748	9.17E-54	V-set and transmembrane domain containing 2 like [HGNC:16096]
PBX1	1.162290	2.17E-53	PBX homeobox 1 [HGNC:8632]
ACACA	0.733132	2.27E-53	acetyl-CoA carboxylase alpha [HGNC:84]
COMT	1.069211	2.35E-53	catechol-O-methyltransferase [HGNC:2228]
SCNN1A	1.103496	5.43E-52	sodium channel epithelial 1 alpha subunit [HGNC:10599]
SDC3	0.691289	6.30E-52	syndecan 3 [HGNC:10660]
CSF1	0.677079	1.12E-51	colony stimulating factor 1 [HGNC:2432]
PCSK9	1.977877	2.20E-49	proprotein convertase subtilisin/kexin type 9 [HGNC:20001]
MAN1B1	0.733309	5.86E-49	mannosidase alpha class 1B member 1 [HGNC:6823]
EEF1A2	1.409279	7.65E-48	eukaryotic translation elongation factor 1 alpha 2 [HGNC:3192]
MSMO1	0.854485	7.95E-48	methylsterol monooxygenase 1 [HGNC:10545]
UNC93B1	0.747414	8.82E-48	unc-93 homolog B1, TLR signaling regulator [HGNC:13481]
DBH-AS1	1.498074	1.36E-47	DBH antisense RNA 1 [HGNC:24155]
SPATA20	0.735402	3.51E-47	spermatogenesis associated 20 [HGNC:26125]
YARS	0.654360	4.07E-47	tyrosyl-tRNA synthetase [HGNC:12840]
GPRC5C	1.602989	1.67E-46	G protein-coupled receptor class C group 5 member C [HGNC:13309]
SLFN5	1.313099	3.63E-46	schlafen family member 5 [HGNC:28286]
RHOBTB3	0.584153	1.06E-45	Rho related BTB domain containing 3 [HGNC:18757]
SCARA3	1.237639	6.72E-45	scavenger receptor class A member 3 [HGNC:19000]
LFNG	1.418102	7.86E-45	LFNG O-fucosylpeptide 3-beta-N-acetylglucosaminyltransferase [HGNC:6560]
ST3GAL5	1.341488	2.70E-44	ST3 beta-galactoside alpha-2,3-sialyltransferase 5 [HGNC:10872]
ITGB4	0.596424	3.56E-44	integrin subunit beta 4 [HGNC:6158]
VASN	1.103910	1.21E-43	vasorin [HGNC:18517]
DAB2IP	0.583608	8.57E-43	DAB2 interacting protein [HGNC:17294]

RGS4	1.504903	1.42E-42	regulator of G protein signaling 4 [HGNC:10000]
DDIT4	0.646327	2.18E-42	DNA damage inducible transcript 4 [HGNC:24944]
SEMA3C	0.605523	3.25E-42	semaphorin 3C [HGNC:10725]
PLPP2	0.979671	4.34E-42	phospholipid phosphatase 2 [HGNC:9230]
SNTB1	1.563468	5.60E-42	syntrophin beta 1 [HGNC:11168]
UAP1L1	0.869605	2.25E-41	UDP-N-acetylglucosamine pyrophosphorylase 1 like 1 [HGNC:28082]
ARMC9	1.294311	2.30E-41	armadillo repeat containing 9 [HGNC:20730]
SERPINE2	1.395945	4.27E-41	serpin family E member 2 [HGNC:8951]
ELFN2	1.047513	2.15E-40	extracellular leucine rich repeat and fibronectin type III domain containing 2 [HGNC:29396]
NFIX	0.655545	2.74E-40	nuclear factor I X [HGNC:7788]
IER5L	1.624762	4.58E-40	immediate early response 5 like [HGNC:23679]
ADGRL2	1.029790	3.40E-39	adhesion G protein-coupled receptor L2 [HGNC:18582]
CBLB	1.172149	3.62E-39	Cbl proto-oncogene B [HGNC:1542]
NOXA1	1.042298	7.20E-39	NADPH oxidase activator 1 [HGNC:10668]
AK4	0.655154	9.98E-39	adenylate kinase 4 [HGNC:363]
TNS1	1.464813	1.73E-38	tensin 1 [HGNC:11973]
COL6A2	0.666541	3.28E-38	collagen type VI alpha 2 chain [HGNC:2212]
PTPRK	0.699271	4.11E-38	protein tyrosine phosphatase, receptor type K [HGNC:9674]
IRX5	1.345591	4.98E-38	iroquois homeobox 5 [HGNC:14361]
CA12	0.953194	7.10E-38	carbonic anhydrase 12 [HGNC:1371]
SRPX	0.725885	1.11E-37	sushi repeat containing protein X-linked [HGNC:11309]
LGR4	0.738369	1.32E-37	leucine rich repeat containing G protein-coupled receptor 4 [HGNC:13299]
CXXC5	0.914751	1.57E-37	CXXC finger protein 5 [HGNC:26943]
CAMK1D	1.297249	1.93E-37	calcium/calmodulin dependent protein kinase ID [HGNC:19341]
SLC37A2	0.636161	7.08E-37	solute carrier family 37 member 2 [HGNC:20644]
CDK15	1.548925	1.33E-36	cyclin dependent kinase 15 [HGNC:14434]
ACSS1	0.855552	1.64E-36	acyl-CoA synthetase short chain family member 1 [HGNC:16091]
PGLS	0.772347	2.20E-36	6-phosphogluconolactonase [HGNC:8903]
FN1	1.421041	2.26E-36	fibronectin 1 [HGNC:3778]
TSC1	0.733577	4.20E-36	TSC complex subunit 1 [HGNC:12362]
TNIK	1.318586	5.45E-36	TRAF2 and NCK interacting kinase [HGNC:30765]
RCAN2	2.309374	9.57E-36	regulator of calcineurin 2 [HGNC:3041]
EVI2A	0.796510	1.09E-35	ecotropic viral integration site 2A [HGNC:3499]
NCAM2	1.604565	1.26E-35	neural cell adhesion molecule 2 [HGNC:7657]
SOD3	0.778684	1.69E-35	superoxide dismutase 3 [HGNC:11181]
CREB3L1	0.636482	1.72E-35	cAMP responsive element binding protein 3 like 1 [HGNC:18856]
PBXIP1	0.738962	2.11E-35	PBX homeobox interacting protein 1 [HGNC:21199]
PTPRM	0.613919	2.23E-35	protein tyrosine phosphatase, receptor type M [HGNC:9675]
PPIB	0.628340	2.70E-35	peptidylprolyl isomerase B [HGNC:9255]
MBNL2	0.612351	3.31E-35	muscleblind like splicing regulator 2 [HGNC:16746]
KITLG	1.022944	4.20E-35	KIT ligand [HGNC:6343]
DPP4	2.601998	6.92E-35	dipeptidyl peptidase 4 [HGNC:3009]
TMEM37	2.064087	1.59E-34	transmembrane protein 37 [HGNC:18216]
LRP3	1.564714	5.44E-34	LDL receptor related protein 3 [HGNC:6695]
TMSB4XP8	0.790723	6.02E-34	TMSB4X pseudogene 8 [HGNC:11885]

AHNAK2	1.200027	1.54E-33	AHNAK nucleoprotein 2 [HGNC:20125]
ME1	0.800331	4.52E-33	malic enzyme 1 [HGNC:6983]
PNPLA3	1.606168	4.90E-33	patatin like phospholipase domain containing 3 [HGNC:18590]
ASAP3	0.593355	7.59E-33	ArfGAP with SH3 domain, ankyrin repeat and PH domain 3 [HGNC:14987]
CA9	3.005570	8.62E-33	carbonic anhydrase 9 [HGNC:1383]
ST3GAL1	0.757961	2.05E-32	ST3 beta-galactoside alpha-2,3-sialyltransferase 1 [HGNC:10862]
MLPH	0.669606	4.24E-32	melanophilin [HGNC:29643]
TMEM229B	1.657266	4.33E-32	transmembrane protein 229B [HGNC:20130]
RTN4R	2.087614	1.19E-31	reticulon 4 receptor [HGNC:18601]
GSTK1	0.598273	1.75E-31	glutathione S-transferase kappa 1 [HGNC:16906]
SERPINA1	0.743319	2.45E-31	serpin family A member 1 [HGNC:8941]
GRINA	0.583553	2.69E-31	glutamate ionotropic receptor NMDA type subunit associated protein 1 [HGNC:4589]
CYBA	0.604754	2.85E-31	cytochrome b-245 alpha chain [HGNC:2577]
FAM86DP	0.927558	3.74E-31	family with sequence similarity 86 member D, pseudogene [HGNC:32659]
MEGF6	1.178841	3.97E-31	multiple EGF like domains 6 [HGNC:3232]
MVD	0.894648	6.02E-31	mevalonate diphosphate decarboxylase [HGNC:7529]
LRRC15	1.409491	9.02E-31	leucine rich repeat containing 15 [HGNC:20818]
HMGCS1	0.645563	1.17E-30	3-hydroxy-3-methylglutaryl-CoA synthase 1 [HGNC:5007]
SMIM14	0.978847	1.80E-30	small integral membrane protein 14 [HGNC:27321]
FASN	1.054994	6.85E-30	fatty acid synthase [HGNC:3594]
CALB2	0.674351	6.85E-30	calbindin 2 [HGNC:1435]
TNFSF9	0.821212	7.95E-30	TNF superfamily member 9 [HGNC:11939]
NIPSNAP1	1.141721	3.25E-29	nipsnap homolog 1 [HGNC:7827]
LRP1	0.821611	6.05E-29	LDL receptor related protein 1 [HGNC:6692]
TGFB1	0.690765	7.45E-29	transforming growth factor beta 1 [HGNC:11766]
GAA	0.653796	8.64E-29	glucosidase alpha, acid [HGNC:4065]
CD276	0.787604	1.79E-28	CD276 molecule [HGNC:19137]
SEMA3B	0.615549	2.22E-28	semaphorin 3B [HGNC:10724]
LOXL4	0.637178	2.40E-28	lysyl oxidase like 4 [HGNC:17171]
PPL	0.701199	6.87E-28	periplakin [HGNC:9273]
CSPG4	0.861278	9.53E-28	chondroitin sulfate proteoglycan 4 [HGNC:2466]
TRIB3	0.928296	2.98E-27	tribbles pseudokinase 3 [HGNC:16228]
GLB1	0.611373	3.66E-27	galactosidase beta 1 [HGNC:4298]
C9orf3	0.630017	3.67E-27	chromosome 9 open reading frame 3 [HGNC:1361]
KCNN4	0.639487	5.57E-27	potassium calcium-activated channel subfamily N member 4 [HGNC:6293]
PPIC	1.035063	7.48E-27	peptidylprolyl isomerase C [HGNC:9256]
DAAM1	0.707261	1.01E-26	dishevelled associated activator of morphogenesis 1 [HGNC:18142]
NPTXR	0.937213	1.80E-26	neuronal pentraxin receptor [HGNC:7954]
ANXA4	0.647679	3.88E-26	annexin A4 [HGNC:542]
CABLES1	0.781222	4.66E-26	Cdk5 and Abl enzyme substrate 1 [HGNC:25097]
PAPLN	1.751331	8.85E-26	papilin, proteoglycan like sulfated glycoprotein [HGNC:19262]
GCNT1	0.758469	9.56E-26	glucosaminyl (N-acetyl) transferase 1, core 2 [HGNC:4203]
LINC00963	0.939441	9.97E-26	long intergenic non-protein coding RNA 963 [HGNC:48716]
GPR108	0.662519	1.09E-25	G protein-coupled receptor 108 [HGNC:17829]
IRX3	0.890436	1.10E-25	iroquois homeobox 3 [HGNC:14360]

CST7	1.205534	1.10E-25	cystatin F [HGNC:2479]
HFE	0.789457	1.20E-25	homeostatic iron regulator [HGNC:4886]
TP53I11	0.905609	1.29E-25	tumor protein p53 inducible protein 11 [HGNC:16842]
ZFAS1	0.827988	1.44E-25	ZNFX1 antisense RNA 1 [HGNC:33101]
SEC14L6	3.724857	1.68E-25	SEC14 like lipid binding 6 [HGNC:40047]
OAF	0.624782	1.72E-25	out at first homolog [HGNC:28752]
LINC01444	1.228187	1.78E-25	long intergenic non-protein coding RNA 1444 [HGNC:50769]
DBH	1.529824	2.67E-25	dopamine beta-hydroxylase [HGNC:2689]
SURF1	0.849161	3.43E-25	SURF1, cytochrome c oxidase assembly factor [HGNC:11474]
ZNF467	1.953058	3.74E-25	zinc finger protein 467 [HGNC:23154]
MAOB	2.153823	4.32E-25	monoamine oxidase B [HGNC:6834]
MATN2	0.591124	6.56E-25	matrilin 2 [HGNC:6908]
EMB	1.078963	9.20E-25	embigin [HGNC:30465]
NNMT	1.062351	1.00E-24	nicotinamide N-methyltransferase [HGNC:7861]
ABCA1	0.612110	1.46E-24	ATP binding cassette subfamily A member 1 [HGNC:29]
SNHG8	0.630667	3.54E-24	small nucleolar RNA host gene 8 [HGNC:33098]
ABCA2	0.663801	6.82E-24	ATP binding cassette subfamily A member 2 [HGNC:32]
NCS1	0.581922	8.04E-24	neuronal calcium sensor 1 [HGNC:3953]
CDKN1C	1.047333	1.52E-23	cyclin dependent kinase inhibitor 1C [HGNC:1786]
DYSF	0.607958	2.17E-23	dysferlin [HGNC:3097]
SCD	1.043706	3.33E-23	stearoyl-CoA desaturase [HGNC:10571]
FBXO32	1.011905	3.41E-23	F-box protein 32 [HGNC:16731]
HMGCR	0.584216	3.59E-23	3-hydroxy-3-methylglutaryl-CoA reductase [HGNC:5006]
DHRS3	2.403693	4.32E-23	dehydrogenase/reductase 3 [HGNC:17693]
IQCE	0.791036	4.96E-23	IQ motif containing E [HGNC:29171]
IDUA	0.929390	5.21E-23	iduronidase, alpha-L- [HGNC:5391]
CDK7	0.613422	6.68E-23	cyclin dependent kinase 7 [HGNC:1778]
ANOS1	0.768537	7.80E-23	anosmin 1 [HGNC:6211]
TCEA2	0.640634	9.83E-23	transcription elongation factor A2 [HGNC:11614]
BTG1	0.591568	1.01E-22	BTG anti-proliferation factor 1 [HGNC:1130]
ZNF608	0.886730	1.11E-22	zinc finger protein 608 [HGNC:29238]
CCNG2	0.639809	1.43E-22	cyclin G2 [HGNC:1593]
IL1RN	2.127881	1.49E-22	interleukin 1 receptor antagonist [HGNC:6000]
MRC2	0.843617	1.77E-22	mannose receptor C type 2 [HGNC:16875]
NOL4L	1.186714	2.93E-22	nucleolar protein 4 like [HGNC:16106]
SLC29A2	0.792753	3.48E-22	solute carrier family 29 member 2 [HGNC:11004]
CERS4	1.042865	3.55E-22	ceramide synthase 4 [HGNC:23747]
NAGLU	0.706196	3.56E-22	N-acetyl-alpha-glucosaminidase [HGNC:7632]
TBC1D9	0.589598	4.66E-22	TBC1 domain family member 9 [HGNC:21710]
FOXQ1	0.761233	4.83E-22	forkhead box Q1 [HGNC:20951]
LTBP3	0.917919	7.82E-22	latent transforming growth factor beta binding protein 3 [HGNC:6716]
RETSAT	0.631481	7.93E-22	retinol saturase [HGNC:25991]
HSPG2	1.095026	8.07E-22	heparan sulfate proteoglycan 2 [HGNC:5273]
SLC1A1	1.765149	1.10E-21	solute carrier family 1 member 1 [HGNC:10939]
ADAMTS9	1.203404	1.38E-21	ADAM metallopeptidase with thrombospondin type 1 motif 9 [HGNC:13202]

DNAH2	0.611969	2.04E-21	dynein axonemal heavy chain 2 [HGNC:2948]
SLC6A9	0.940079	2.07E-21	solute carrier family 6 member 9 [HGNC:11056]
PMEPA1	0.876187	4.56E-21	prostate transmembrane protein, androgen induced 1 [HGNC:14107]
SLCO4C1	1.203243	4.74E-21	solute carrier organic anion transporter family member 4C1 [HGNC:23612]
RHOU	1.293058	4.91E-21	ras homolog family member U [HGNC:17794]
SEMA3D	1.784501	5.40E-21	semaphorin 3D [HGNC:10726]
DHRSX	0.784988	5.59E-21	dehydrogenase/reductase X-linked [HGNC:18399]
TNS3	0.597024	5.92E-21	tensin 3 [HGNC:21616]
LRFN4	0.781169	6.69E-21	leucine rich repeat and fibronectin type III domain containing 4 [HGNC:28456]
PDE5A	1.535971	7.31E-21	phosphodiesterase 5A [HGNC:8784]
SNHG7	0.587580	8.35E-21	small nucleolar RNA host gene 7 [HGNC:28254]
CAMK2D	0.625849	8.36E-21	calcium/calmodulin dependent protein kinase II delta [HGNC:1462]
MT-TP	0.769954	9.24E-21	mitochondrially encoded tRNA proline [HGNC:7494]
FAM84B	0.676914	9.24E-21	family with sequence similarity 84 member B [HGNC:24166]
EPB41L1	0.709608	1.15E-20	erythrocyte membrane protein band 4.1 like 1 [HGNC:3378]
RARRES2	0.690597	1.42E-20	retinoic acid receptor responder 2 [HGNC:9868]
TBC1D17	0.626705	2.63E-20	TBC1 domain family member 17 [HGNC:25699]
EXD3	0.946302	2.74E-20	exonuclease 3'-5' domain containing 3 [HGNC:26023]
SLC22A18	1.190481	2.77E-20	solute carrier family 22 member 18 [HGNC:10964]
MVK	0.644184	3.90E-20	mevalonate kinase [HGNC:7530]
GADD45B	0.793420	4.08E-20	growth arrest and DNA damage inducible beta [HGNC:4096]
SOX12	0.670092	4.24E-20	SRY-box 12 [HGNC:11198]
METRNL	1.390873	4.41E-20	meteorin like, glial cell differentiation regulator [HGNC:27584]
RNPEPL1	0.619485	4.87E-20	arginyl aminopeptidase like 1 [HGNC:10079]
IL2RB	5.400564	5.12E-20	interleukin 2 receptor subunit beta [HGNC:6009]
KLHL24	0.845282	5.74E-20	kelch like family member 24 [HGNC:25947]
MAPRE3	0.976724	6.33E-20	microtubule associated protein RP/EB family member 3 [HGNC:6892]
TRUB2	0.709147	8.10E-20	TruB pseudouridine synthase family member 2 [HGNC:17170]
XYLT2	0.787841	8.68E-20	xylosyltransferase 2 [HGNC:15517]
SUSD3	0.997100	8.93E-20	sushi domain containing 3 [HGNC:28391]
ACKR3	2.523892	9.60E-20	atypical chemokine receptor 3 [HGNC:23692]
C2orf72	2.974147	1.01E-19	chromosome 2 open reading frame 72 [HGNC:27418]
NAV1	0.623313	1.33E-19	neuron navigator 1 [HGNC:15989]
MAP2	1.216278	2.40E-19	microtubule associated protein 2 [HGNC:6839]
DRAXIN	2.681510	3.06E-19	dorsal inhibitory axon guidance protein [HGNC:25054]
GRAMD1A	0.617345	3.34E-19	GRAM domain containing 1A [HGNC:29305]
LPCAT2	0.659681	3.66E-19	lysophosphatidylcholine acyltransferase 2 [HGNC:26032]
DAB2	0.742164	6.22E-19	DAB2, clathrin adaptor protein [HGNC:2662]
JAG2	0.990968	6.77E-19	jagged 2 [HGNC:6189]
ZFP36	0.868237	8.55E-19	ZFP36 ring finger protein [HGNC:12862]
GALNT3	0.612943	1.08E-18	polypeptide N-acetylgalactosaminyltransferase 3 [HGNC:4125]
REEP6	1.219814	1.11E-18	receptor accessory protein 6 [HGNC:30078]
F11R	0.921967	1.96E-18	F11 receptor [HGNC:14685]
ROR1	0.603863	2.05E-18	receptor tyrosine kinase like orphan receptor 1 [HGNC:10256]
LIPG	0.798646	2.89E-18	lipase G, endothelial type [HGNC:6623]

CXorf38	0.624723	3.42E-18	chromosome X open reading frame 38 [HGNC:28589]
SYT12	1.726184	3.57E-18	synaptotagmin 12 [HGNC:18381]
SMAD6	1.270933	4.00E-18	SMAD family member 6 [HGNC:6772]
TCEA3	0.687508	4.77E-18	transcription elongation factor A3 [HGNC:11615]
KCP	2.154570	5.24E-18	kielin/chordin-like protein [HGNC:17585]
SLC27A1	0.667572	6.13E-18	solute carrier family 27 member 1 [HGNC:10995]
KRCC1	0.758129	6.75E-18	lysine rich coiled-coil 1 [HGNC:28039]
PIM2	0.694854	7.38E-18	Pim-2 proto-oncogene, serine/threonine kinase [HGNC:8987]
MAP2K6	1.642082	1.01E-17	mitogen-activated protein kinase kinase 6 [HGNC:6846]
SCN1B	1.142868	1.18E-17	sodium voltage-gated channel beta subunit 1 [HGNC:10586]
C4orf19	1.087592	1.20E-17	chromosome 4 open reading frame 19 [HGNC:25618]
PIK3IP1	1.407645	1.75E-17	phosphoinositide-3-kinase interacting protein 1 [HGNC:24942]
IFITM10	1.826823	1.92E-17	interferon induced transmembrane protein 10 [HGNC:40022]
CACFD1	0.840180	1.95E-17	calcium channel flower domain containing 1 [HGNC:1365]
APOL1	0.939269	2.16E-17	apolipoprotein L1 [HGNC:618]
HMGCL	0.939269	3.10E-17	3-hydroxy-3-methylglutaryl-CoA lyase [HGNC:5005]
NEK8	1.111295	3.12E-17	NIMA related kinase 8 [HGNC:13387]
FIBCD1	1.638182	4.50E-17	fibrinogen C domain containing 1 [HGNC:25922]
PLCD1	0.906398	4.92E-17	phospholipase C delta 1 [HGNC:9060]
IFI35	0.692840	5.08E-17	interferon induced protein 35 [HGNC:5399]
MAFB	1.893661	5.22E-17	MAF bZIP transcription factor B [HGNC:6408]
MGAT5B	0.681934	5.79E-17	alpha-1,6-mannosylglycoprotein 6-beta-N-acetylglucosaminyltransferase B [HGNC:24140]
COL4A5	1.103097	7.82E-17	collagen type IV alpha 5 chain [HGNC:2207]
CXCL14	4.084543	8.14E-17	C-X-C motif chemokine ligand 14 [HGNC:10640]
TSC22D3	1.010369	8.74E-17	TSC22 domain family member 3 [HGNC:3051]
TXNRD2	0.880086	1.33E-16	thioredoxin reductase 2 [HGNC:18155]
VTN	2.482261	1.35E-16	vitronectin [HGNC:12724]
EPHA4	1.010630	1.81E-16	EPH receptor A4 [HGNC:3388]
THBS3	1.329680	2.53E-16	thrombospondin 3 [HGNC:11787]
TBX1	1.160543	2.55E-16	T-box 1 [HGNC:11592]
PCDH18	2.732656	3.21E-16	protocadherin 18 [HGNC:14268]
SNCG	1.004964	3.87E-16	synuclein gamma [HGNC:11141]
DPM3	0.642395	4.00E-16	dolichyl-phosphate mannosyltransferase subunit 3 [HGNC:3007]
FOS	1.199554	4.24E-16	Fos proto-oncogene, AP-1 transcription factor subunit [HGNC:3796]
GJA1	1.520957	4.30E-16	gap junction protein alpha 1 [HGNC:4274]
COL9A2	1.003566	7.28E-16	collagen type IX alpha 2 chain [HGNC:2218]
MMAB	0.630942	7.68E-16	methylmalonic aciduria (cobalamin deficiency) cblB type [HGNC:19331]
SGCE	0.604649	8.79E-16	sarcoglycan epsilon [HGNC:10808]
COL4A1	0.646916	9.97E-16	collagen type IV alpha 1 chain [HGNC:2202]
GJB2	0.871038	1.19E-15	gap junction protein beta 2 [HGNC:4284]
FKBP7	0.882548	1.28E-15	FK506 binding protein 7 [HGNC:3723]
GBP2	0.643750	1.28E-15	guanylate binding protein 2 [HGNC:4183]
GALNT14	2.232877	1.48E-15	polypeptide N-acetylgalactosaminyltransferase 14 [HGNC:22946]
CLCN6	0.597101	1.56E-15	chloride voltage-gated channel 6 [HGNC:2024]
SUMF1	0.741015	1.91E-15	sulfatase modifying factor 1 [HGNC:20376]

TTC39C	0.601447	2.34E-15	tetratricopeptide repeat domain 39C [HGNC:26595]
SCN9A	1.562242	2.59E-15	sodium voltage-gated channel alpha subunit 9 [HGNC:10597]
DOK4	1.049787	2.69E-15	docking protein 4 [HGNC:19868]
FAM198B	1.485221	3.20E-15	family with sequence similarity 198 member B [HGNC:25312]
RALGDS	0.947023	3.84E-15	ral guanine nucleotide dissociation stimulator [HGNC:9842]
SLC16A13	1.168144	4.14E-15	solute carrier family 16 member 13 [HGNC:31037]
UBA7	0.819008	4.14E-15	ubiquitin like modifier activating enzyme 7 [HGNC:12471]
RBM43	1.271941	4.27E-15	RNA binding motif protein 43 [HGNC:24790]
HHIP	1.387495	4.48E-15	hedgehog interacting protein [HGNC:14866]
SHANK3	1.034364	4.97E-15	SH3 and multiple ankyrin repeat domains 3 [HGNC:14294]
MTX1P1	1.436310	5.67E-15	metaxin 1 pseudogene 1 [HGNC:7505]
ALDOC	1.303986	6.63E-15	aldolase, fructose-bisphosphate C [HGNC:418]
VWA1	0.668378	6.67E-15	von Willebrand factor A domain containing 1 [HGNC:30910]
MTHFR	0.757236	7.97E-15	methylenetetrahydrofolate reductase [HGNC:7436]
ABTB1	0.738431	8.99E-15	ankyrin repeat and BTB domain containing 1 [HGNC:18275]
ATP2B2	2.214836	1.00E-14	ATPase plasma membrane Ca2+ transporting 2 [HGNC:815]
PSMG3-AS1	1.387122	1.10E-14	PSMG3 antisense RNA 1 (head to head) [HGNC:22230]
KCNK6	1.039891	1.43E-14	potassium two pore domain channel subfamily K member 6 [HGNC:6281]
GEMIN8	1.139719	1.45E-14	gem nuclear organelle associated protein 8 [HGNC:26044]
RHOB	0.723686	1.60E-14	ras homolog family member B [HGNC:668]
SSH3	0.664847	2.21E-14	slingshot protein phosphatase 3 [HGNC:30581]
DYRK1B	0.865504	3.41E-14	dual specificity tyrosine phosphorylation regulated kinase 1B [HGNC:3092]
NTNG2	1.077582	3.73E-14	netrin G2 [HGNC:14288]
TMTC4	0.803249	3.89E-14	transmembrane and tetratricopeptide repeat containing 4 [HGNC:25904]
LINC01443	1.458515	4.16E-14	long intergenic non-protein coding RNA 1443 [HGNC:50768]
ESPN	1.695694	4.38E-14	espin [HGNC:13281]
PSPH	0.597042	4.66E-14	phosphoserine phosphatase [HGNC:9577]
ZBTB47	0.662878	4.91E-14	zinc finger and BTB domain containing 47 [HGNC:26955]
FER1L4	1.778914	6.09E-14	fer-1 like family member 4, pseudogene [HGNC:15801]
EHD3	1.359660	6.12E-14	EH domain containing 3 [HGNC:3244]
GGPS1	0.833157	7.70E-14	geranylgeranyl diphosphate synthase 1 [HGNC:4249]
FAM117A	1.197761	7.75E-14	family with sequence similarity 117 member A [HGNC:24179]
CELSR2	0.649859	9.76E-14	cadherin EGF LAG seven-pass G-type receptor 2 [HGNC:3231]
FOSB	2.047930	1.03E-13	FosB proto-oncogene, AP-1 transcription factor subunit [HGNC:3797]
CNIH3	0.730493	1.37E-13	cornichon family AMPA receptor auxiliary protein 3 [HGNC:26802]
FOLR1	1.469675	1.39E-13	folate receptor 1 [HGNC:3791]
HNMT	0.638592	1.59E-13	histamine N-methyltransferase [HGNC:5028]
AZIN2	0.801669	1.60E-13	antizyme inhibitor 2 [HGNC:29957]
ZDHHC1	0.958391	1.80E-13	zinc finger DHHC-type containing 1 [HGNC:17916]
GHDC	0.673318	1.97E-13	GH3 domain containing [HGNC:24438]
NRP2	0.665510	2.01E-13	neuropilin 2 [HGNC:8005]
ALDH1A3	1.422595	2.01E-13	aldehyde dehydrogenase 1 family member A3 [HGNC:409]
ASIC1	0.912379	2.03E-13	acid sensing ion channel subunit 1 [HGNC:100]
MAGEB2	4.150332	2.19E-13	MAGE family member B2 [HGNC:6809]
PCDHB10	1.505659	2.80E-13	protocadherin beta 10 [HGNC:8681]

PLXNB1	0.986205	3.13E-13	plexin B1 [HGNC:9103]
DNM1	0.860150	4.18E-13	dynamin 1 [HGNC:2972]
ARHGEF3	0.599332	4.23E-13	Rho guanine nucleotide exchange factor 3 [HGNC:683]
PRELP	1.537835	4.38E-13	proline and arginine rich end leucine rich repeat protein [HGNC:9357]
PNRC1	0.725825	4.44E-13	proline rich nuclear receptor coactivator 1 [HGNC:17278]
FHOD3	0.624480	4.53E-13	formin homology 2 domain containing 3 [HGNC:26178]
MDFI	0.873430	4.58E-13	MyoD family inhibitor [HGNC:6967]
GRIK2	1.496976	4.69E-13	glutamate ionotropic receptor kainate type subunit 2 [HGNC:4580]
FBXO2	0.938105	5.25E-13	F-box protein 2 [HGNC:13581]
FO681492.1	1.406813	7.46E-13	synaptotagmin-15 [NCBI :102724488]
LUM	1.949692	8.05E-13	lumican [HGNC:6724]
SCARF2	1.108379	9.11E-13	scavenger receptor class F member 2 [HGNC:19869]
PRICKLE2	0.706081	9.69E-13	prickle planar cell polarity protein 2 [HGNC:20340]
SLC37A1	0.612699	1.07E-12	solute carrier family 37 member 1 [HGNC:11024]
PLPP4	0.668568	1.26E-12	phospholipid phosphatase 4 [HGNC:23531]
CPE	0.782372	1.33E-12	carboxypeptidase E [HGNC:2303]
PRLR	0.832036	1.36E-12	prolactin receptor [HGNC:9446]
CPPED1	0.590765	1.41E-12	calcineurin like phosphoesterase domain containing 1 [HGNC:25632]
FAM86C2P	1.259965	1.71E-12	family with sequence similarity 86 member C2, pseudogene [HGNC:42392]
DOLK	0.600912	2.60E-12	dolichol kinase [HGNC:23406]
BAIAP2-DT	0.712965	3.13E-12	BAIAP2 divergent transcript [HGNC:44342]
HAGHL	0.716129	3.33E-12	hydroxyacylglutathione hydrolase like [HGNC:14177]
HBP1	0.607847	3.61E-12	HMG-box transcription factor 1 [HGNC:23200]
AJM1	0.841388	4.58E-12	apical junction component 1 homolog [HGNC:37284]
MEGF8	0.621272	4.62E-12	multiple EGF like domains 8 [HGNC:3233]
DNAH1	1.062218	5.68E-12	dynein axonemal heavy chain 1 [HGNC:2940]
SLC43A2	1.185070	5.80E-12	solute carrier family 43 member 2 [HGNC:23087]
RGS5	1.171023	7.04E-12	regulator of G protein signaling 5 [HGNC:10001]
OLFML3	2.323502	7.13E-12	olfactomedin like 3 [HGNC:24956]
BAIAP3	1.119984	7.80E-12	BAI1 associated protein 3 [HGNC:948]
MANBA	0.899251	8.47E-12	mannosidase beta [HGNC:6831]
SHISA4	0.834837	8.49E-12	shisa family member 4 [HGNC:27139]
NATD1	0.835162	8.66E-12	N-acetyltransferase domain containing 1 [HGNC:30770]
BEST1	2.107496	1.00E-11	bestrophin 1 [HGNC:12703]
OTUD1	0.621563	1.01E-11	OTU deubiquitinase 1 [HGNC:27346]
TRPS1	0.985988	1.04E-11	transcriptional repressor GATA binding 1 [HGNC:12340]
LAMTOR2	0.643130	1.25E-11	late endosomal/lysosomal adaptor, MAPK and MTOR activator 2 [HGNC:29796]
MMP11	1.149952	1.33E-11	matrix metallopeptidase 11 [HGNC:7157]
NREP	0.688014	1.36E-11	neuronal regeneration related protein [HGNC:16834]
TCP11L2	1.436438	1.39E-11	t-complex 11 like 2 [HGNC:28627]
HES2	1.081645	1.44E-11	hes family bHLH transcription factor 2 [HGNC:16005]
RAMP1	0.600233	1.52E-11	receptor activity modifying protein 1 [HGNC:9843]
SNED1	1.326254	1.63E-11	sushi, nidogen and EGF like domains 1 [HGNC:24696]
SLC27A3	1.064627	1.64E-11	solute carrier family 27 member 3 [HGNC:10997]
PYCARD	1.063247	1.76E-11	PYD and CARD domain containing [HGNC:16608]

YPEL3	0.759124	1.80E-11	yippee like 3 [HGNC:18327]
ANKRD29	0.707289	1.83E-11	ankyrin repeat domain 29 [HGNC:27110]
MLXIPL	1.344207	1.90E-11	MLX interacting protein like [HGNC:12744]
GXYLT2	0.776328	2.03E-11	glucoside xylosyltransferase 2 [HGNC:33383]
PLSCR4	0.728014	2.05E-11	phospholipid scramblase 4 [HGNC:16497]
ORAI3	0.685625	2.07E-11	ORAI calcium release-activated calcium modulator 3 [HGNC:28185]
PCDHB14	1.083037	2.79E-11	protocadherin beta 14 [HGNC:8685]
CREBRF	0.627705	3.02E-11	CREB3 regulatory factor [HGNC:24050]
DOCK6	0.753032	3.71E-11	dedicator of cytokinesis 6 [HGNC:19189]
ECH1	0.611512	3.96E-11	enoyl-CoA hydratase 1 [HGNC:3149]
GPER1	0.752737	3.98E-11	G protein-coupled estrogen receptor 1 [HGNC:4485]
CADPS2	0.781999	4.12E-11	calcium dependent secretion activator 2 [HGNC:16018]
AK1	0.766501	4.54E-11	adenylate kinase 1 [HGNC:361]
FAHD1	0.692901	4.85E-11	fumarylacetoacetate hydrolase domain containing 1 [HGNC:14169]
BFSP1	0.925966	5.25E-11	beaded filament structural protein 1 [HGNC:1040]
SPACA9	1.294428	5.55E-11	sperm acrosome associated 9 [HGNC:1367]
DSEL	0.908169	5.60E-11	dermatan sulfate epimerase like [HGNC:18144]
TCTN1	0.834477	8.08E-11	tectonic family member 1 [HGNC:26113]
PITX2	2.929018	9.42E-11	paired like homeodomain 2 [HGNC:9005]
ADAM12	1.370834	1.01E-10	ADAM metallopeptidase domain 12 [HGNC:190]
KYAT1	0.736978	1.24E-10	kynurenine aminotransferase 1 [HGNC:1564]
PCNX2	0.613714	1.49E-10	pecanex 2 [HGNC:8736]
FAM83H	0.637539	1.52E-10	family with sequence similarity 83 member H [HGNC:24797]
PNPLA4	0.702670	1.65E-10	patatin like phospholipase domain containing 4 [HGNC:24887]
PRR7	0.695986	1.78E-10	proline rich 7, synaptic [HGNC:28130]
NLRC5	0.590862	2.20E-10	NLR family CARD domain containing 5 [HGNC:29933]
SLC16A4	1.587165	2.36E-10	solute carrier family 16 member 4 [HGNC:10925]
YPEL2	1.081444	2.84E-10	yippee like 2 [HGNC:18326]
CDK18	1.579359	2.90E-10	cyclin dependent kinase 18 [HGNC:8751]
OLFML2B	1.511388	3.04E-10	olfactomedin like 2B [HGNC:24558]
SEMA3A	0.683802	3.12E-10	semaphorin 3A [HGNC:10723]
ST6GALNAC4	0.591227	3.28E-10	ST6 N-acetylgalactosaminide alpha-2,6-sialyltransferase 4 [HGNC:17846]
CCDC88B	1.013119	3.36E-10	coiled-coil domain containing 88B [HGNC:26757]
DUXAP8	1.495041	3.44E-10	double homeobox A pseudogene 8 [HGNC:32187]
ASS1	0.769030	4.20E-10	argininosuccinate synthase 1 [HGNC:758]
ACOT1	1.163341	4.42E-10	acyl-CoA thioesterase 1 [HGNC:33128]
PSG4	0.892217	4.47E-10	pregnancy specific beta-1-glycoprotein 4 [HGNC:9521]
DGLUCY	0.589115	4.49E-10	D-glutamate cyclase [HGNC:20498]
VSTM4	2.820515	4.77E-10	V-set and transmembrane domain containing 4 [HGNC:26470]
HDAC11	0.655947	4.80E-10	histone deacetylase 11 [HGNC:19086]
NKX3-2	0.758851	5.41E-10	NK3 homeobox 2 [HGNC:951]
RWDD2B	0.691767	5.41E-10	RWD domain containing 2B [HGNC:1302]
RAB7B	2.064927	5.48E-10	RAB7B, member RAS oncogene family [HGNC:30513]
RAP1GAP	1.297222	5.48E-10	RAP1 GTPase activating protein [HGNC:9858]
TSPAN15	0.626564	5.74E-10	tetraspanin 15 [HGNC:23298]

TMEM8B	0.873069	5.93E-10	transmembrane protein 8B [HGNC:21427]
POLD4	0.888271	6.41E-10	DNA polymerase delta 4, accessory subunit [HGNC:14106]
DUXAP10	1.428954	6.96E-10	double homeobox A pseudogene 10 [NCBI :503639]
PDGFRB	0.747438	7.27E-10	platelet derived growth factor receptor beta [HGNC:8804]
RASSF9	2.169550	8.33E-10	Ras association domain family member 9 [HGNC:15739]
MYL5	0.822420	8.53E-10	myosin light chain 5 [HGNC:7586]
CAPN5	0.603383	8.75E-10	calpain 5 [HGNC:1482]
ERV3-1	1.075516	9.72E-10	endogenous retrovirus group 3 member 1, envelope [HGNC:3454]
CALHM2	0.661067	9.88E-10	calcium homeostasis modulator family member 2 [HGNC:23493]
ZNF462	0.637597	1.05E-09	zinc finger protein 462 [HGNC:21684]
SLC38A7	0.607721	1.06E-09	solute carrier family 38 member 7 [HGNC:25582]
B3GNT9	0.743543	1.18E-09	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 9 [HGNC:28714]
TNFRSF14	0.726460	1.24E-09	TNF receptor superfamily member 14 [HGNC:11912]
CNGA1	1.265422	1.72E-09	cyclic nucleotide gated channel alpha 1 [HGNC:2148]
SSPO	2.348601	1.77E-09	SCO-spondin [HGNC:21998]
ATP8B3	0.610781	1.86E-09	ATPase phospholipid transporting 8B3 [HGNC:13535]
VSIR	0.715744	1.87E-09	V-set immunoregulatory receptor [HGNC:30085]
SLC43A1	0.800291	2.14E-09	solute carrier family 43 member 1 [HGNC:9225]
PDE4B	1.060191	2.20E-09	phosphodiesterase 4B [HGNC:8781]
PPP1R3B	0.601561	2.26E-09	protein phosphatase 1 regulatory subunit 3B [HGNC:14942]
CDC42BPG	0.861816	2.27E-09	CDC42 binding protein kinase gamma [HGNC:29829]
DHRS7	0.583027	2.88E-09	dehydrogenase/reductase 7 [HGNC:21524]
ADRA2C	2.551570	3.18E-09	adrenoceptor alpha 2C [HGNC:283]
PIGV	0.614914	3.33E-09	phosphatidylinositol glycan anchor biosynthesis class V [HGNC:26031]
CNTNAP3	0.958408	3.38E-09	contactin associated protein like 3 [HGNC:13834]
SPINT1	1.232162	3.41E-09	serine peptidase inhibitor, Kunitz type 1 [HGNC:11246]
CBR4	0.701335	3.55E-09	carbonyl reductase 4 [HGNC:25891]
JOSD2	0.591712	3.55E-09	Josephin domain containing 2 [HGNC:28853]
MYL1	2.255322	3.57E-09	myosin light chain 1 [HGNC:7582]
ESYT3	1.298923	4.24E-09	extended synaptotagmin 3 [HGNC:24295]
TBKBP1	0.994150	4.43E-09	TBK1 binding protein 1 [HGNC:30140]
ACAD10	0.645405	4.86E-09	acyl-CoA dehydrogenase family member 10 [HGNC:21597]
PCDHB13	0.879714	5.52E-09	protocadherin beta 13 [HGNC:8684]
SPTBN2	0.758437	5.67E-09	spectrin beta, non-erythrocytic 2 [HGNC:11276]
NEK10	1.657843	6.73E-09	NIMA related kinase 10 [HGNC:18592]
SATB1	0.676028	7.34E-09	SATB homeobox 1 [HGNC:10541]
GPNMB	1.453058	7.36E-09	glycoprotein nmb [HGNC:4462]
JDP2	0.706044	7.74E-09	Jun dimerization protein 2 [HGNC:17546]
P2RX6	0.970583	8.18E-09	purinergic receptor P2X 6 [HGNC:8538]
AC108062.1	1.736943	8.82E-09	uncharacterized LOC100507487 [NCBI :100507487]
SLC2A12	0.745100	9.65E-09	solute carrier family 2 member 12 [HGNC:18067]
MARC1	0.912158	1.11E-08	mitochondrial amidoxime reducing component 1 [HGNC:26189]
AC102953.2	0.801771	1.24E-08	novel transcript
LAMA4	1.295146	1.33E-08	laminin subunit alpha 4 [HGNC:6484]
BBC3	0.784793	1.48E-08	BCL2 binding component 3 [HGNC:17868]

GSPT2	0.749033	1.55E-08	G1 to S phase transition 2 [HGNC:4622]
C16orf74	0.995285	1.63E-08	chromosome 16 open reading frame 74 [HGNC:23362]
SPEF2	1.142840	1.63E-08	sperm flagellar 2 [HGNC:26293]
TKFC	0.765411	1.72E-08	triokinase and FMN cyclase [HGNC:24552]
PIP5KL1	1.080818	1.72E-08	phosphatidylinositol-4-phosphate 5-kinase like 1 [HGNC:28711]
LYPD6B	2.691942	1.80E-08	LY6/PLAUR domain containing 6B [HGNC:27018]
FGFR4	1.039530	2.31E-08	fibroblast growth factor receptor 4 [HGNC:3691]
LINC01505	0.992865	2.41E-08	long intergenic non-protein coding RNA 1505 [HGNC:51186]
SLC22A3	0.634551	2.43E-08	solute carrier family 22 member 3 [HGNC:10967]
LINC00639	0.963331	2.45E-08	long intergenic non-protein coding RNA 639 [HGNC:27502]
BBS10	0.627474	2.55E-08	Bardet-Biedl syndrome 10 [HGNC:26291]
ANK2	0.717598	2.57E-08	ankyrin 2 [HGNC:493]
PTGS1	0.693946	2.82E-08	prostaglandin-endoperoxide synthase 1 [HGNC:9604]
TPGS1	0.828804	2.91E-08	tubulin polyglutamylase complex subunit 1 [HGNC:25058]
SAMD11	2.994556	3.07E-08	sterile alpha motif domain containing 11 [HGNC:28706]
PAQR6	1.064358	3.23E-08	progestin and adipoQ receptor family member 6 [HGNC:30132]
RNF224	1.213513	3.68E-08	ring finger protein 224 [HGNC:41912]
NEBL	1.051785	3.80E-08	nebulette [HGNC:16932]
ITPKB	0.927066	3.96E-08	inositol-trisphosphate 3-kinase B [HGNC:6179]
DENND2D	0.807292	4.10E-08	DENN domain containing 2D [HGNC:26192]
CTSO	0.711576	4.31E-08	cathepsin O [HGNC:2542]
SLC2A10	1.405621	5.89E-08	solute carrier family 2 member 10 [HGNC:13444]
LZTR1	0.842245	5.96E-08	leucine zipper like transcription regulator 1 [HGNC:6742]
GLB1L2	0.790925	5.96E-08	galactosidase beta 1 like 2 [HGNC:25129]
TRPM4	0.689477	5.96E-08	transient receptor potential cation channel subfamily M member 4 [HGNC:17993]
KBTBD11	1.065351	6.02E-08	kelch repeat and BTB domain containing 11 [HGNC:29104]
TPM2	0.599738	6.49E-08	tropomyosin 2 [HGNC:12011]
CRIP2	1.172527	6.61E-08	cysteine rich protein 2 [HGNC:2361]
OBSCN	0.873925	6.75E-08	obscurin, cytoskeletal calmodulin and titin-interacting RhoGEF [HGNC:15719]
LHX6	0.968309	7.39E-08	LIM homeobox 6 [HGNC:21735]
FAM171B	0.709401	7.41E-08	family with sequence similarity 171 member B [HGNC:29412]
PHGDH	1.277627	7.57E-08	phosphoglycerate dehydrogenase [HGNC:8923]
COL8A2	2.256303	8.41E-08	collagen type VIII alpha 2 chain [HGNC:2216]
COQ8B	0.589867	8.72E-08	coenzyme Q8B [HGNC:19041]
PRICKLE1	0.592965	9.14E-08	prickle planar cell polarity protein 1 [HGNC:17019]
CDK5	0.637339	9.35E-08	cyclin dependent kinase 5 [HGNC:1774]
SLC35E2A	0.647690	1.07E-07	solute carrier family 35 member E2A [HGNC:20863]
CPM	1.137494	1.09E-07	carboxypeptidase M [HGNC:2311]
PLEKHF1	0.638775	1.17E-07	pleckstrin homology and FYVE domain containing 1 [HGNC:20764]
PLA2R1	1.038676	1.34E-07	phospholipase A2 receptor 1 [HGNC:9042]
SLC25A25-AS1	1.172135	1.63E-07	SLC25A25 antisense RNA 1 [HGNC:27844]
PSAT1	0.705108	1.65E-07	phosphoserine aminotransferase 1 [HGNC:19129]
ERMAP	0.862843	1.69E-07	erythroblast membrane associated protein (Scianna blood group) [HGNC:15743]
GPRC5B	1.135041	1.70E-07	G protein-coupled receptor class C group 5 member B [HGNC:13308]
FAM13C	2.165017	1.86E-07	family with sequence similarity 13 member C [HGNC:19371]

GMFG	1.513896	2.05E-07	glia maturation factor gamma [HGNC:4374]
HCAR1	1.043820	2.42E-07	hydroxycarboxylic acid receptor 1 [HGNC:4532]
PALM	1.213137	2.65E-07	paralemmin [HGNC:8594]
LRRC75B	1.068494	3.00E-07	leucine rich repeat containing 75B [HGNC:33155]
ING4	0.712671	3.37E-07	inhibitor of growth family member 4 [HGNC:19423]
PTP4A3	0.755911	3.60E-07	protein tyrosine phosphatase type IVA, member 3 [HGNC:9636]
PCSK1N	0.875545	3.64E-07	proprotein convertase subtilisin/kexin type 1 inhibitor [HGNC:17301]
UNC5B	2.525270	3.64E-07	unc-5 netrin receptor B [HGNC:12568]
SIGIRR	0.630545	3.86E-07	single Ig and TIR domain containing [HGNC:30575]
GREB1	1.364411	4.04E-07	growth regulating estrogen receptor binding 1 [HGNC:24885]
SLC16A2	1.020392	4.10E-07	solute carrier family 16 member 2 [HGNC:10923]
ZNF117	0.983199	4.16E-07	zinc finger protein 117 [HGNC:12897]
C9orf116	1.019596	4.27E-07	chromosome 9 open reading frame 116 [HGNC:28435]
MMP16	1.122224	4.38E-07	matrix metallopeptidase 16 [HGNC:7162]
SHC2	1.443927	4.49E-07	SHC adaptor protein 2 [HGNC:29869]
COLGALT2	0.802382	4.62E-07	collagen beta(1-O)galactosyltransferase 2 [HGNC:16790]
SH3TC1	0.802613	4.67E-07	SH3 domain and tetratricopeptide repeats 1 [HGNC:26009]
SNAP91	1.133238	5.18E-07	synaptosome associated protein 91 [HGNC:14986]
MT1X	0.777151	5.27E-07	metallothionein 1X [HGNC:7405]
MRPL23	0.596729	5.31E-07	mitochondrial ribosomal protein L23 [HGNC:10322]
PIGP	0.719760	5.40E-07	phosphatidylinositol glycan anchor biosynthesis class P [HGNC:3046]
HIC1	0.810002	5.76E-07	HIC ZBTB transcriptional repressor 1 [HGNC:4909]
C1QTNF6	0.786221	6.35E-07	C1q and TNF related 6 [HGNC:14343]
DNMT3A	0.870251	6.72E-07	DNA methyltransferase 3 alpha [HGNC:2978]
TLCD2	0.822704	6.81E-07	TLC domain containing 2 [HGNC:33522]
NINJ2	0.769746	7.52E-07	ninjurin 2 [HGNC:7825]
DNAL4	0.804974	7.61E-07	dynein axonemal light chain 4 [HGNC:2955]
FBXL19	0.580886	7.61E-07	F-box and leucine rich repeat protein 19 [HGNC:25300]
NR1D1	0.684958	8.16E-07	nuclear receptor subfamily 1 group D member 1 [HGNC:7962]
PCDHB11	0.922404	8.34E-07	protocadherin beta 11 [HGNC:8682]
PCDHB9	1.286274	8.45E-07	protocadherin beta 9 [HGNC:8694]
IRF5	1.065302	8.49E-07	interferon regulatory factor 5 [HGNC:6120]
AC068580.3	0.882715	8.91E-07	novel transcript
PADI3	3.323788	9.24E-07	peptidyl arginine deiminase 3 [HGNC:18337]
CYSRT1	0.813286	9.53E-07	cysteine rich tail 1 [HGNC:30529]
SEMA3F	0.597837	9.76E-07	semaphorin 3F [HGNC:10728]
CASC10	0.792202	1.02E-06	cancer susceptibility 10 [HGNC:31448]
ARSA	0.630227	1.04E-06	arylsulfatase A [HGNC:713]
HTR1D	1.058277	1.07E-06	5-hydroxytryptamine receptor 1D [HGNC:5289]
MT1F	1.066007	1.11E-06	metallothionein 1F [HGNC:7398]
HCP5	0.827890	1.13E-06	HLA complex P5 [HGNC:21659]
OBSCN-AS1	2.032542	1.19E-06	OBSCN antisense RNA 1 [HGNC:32047]
HTRA3	0.848154	1.20E-06	HtrA serine peptidase 3 [HGNC:30406]
FBXO44	0.698087	1.26E-06	F-box protein 44 [HGNC:24847]
BICDL1	1.669032	1.26E-06	BICD family like cargo adaptor 1 [HGNC:28095]

TANGO2	0.751285	1.54E-06	transport and golgi organization 2 homolog [HGNC:25439]
BAMBI	0.716769	1.89E-06	BMP and activin membrane bound inhibitor [HGNC:30251]
IFT140	0.646720	1.95E-06	intraflagellar transport 140 [HGNC:29077]
SPATC1L	0.633497	1.95E-06	spermatogenesis and centriole associated 1 like [HGNC:1298]
PARP8	0.678315	2.23E-06	poly(ADP-ribose) polymerase family member 8 [HGNC:26124]
RNF144A	0.613217	2.24E-06	ring finger protein 144A [HGNC:20457]
ASIC3	0.809144	2.36E-06	acid sensing ion channel subunit 3 [HGNC:101]
ABCA3	0.596389	2.42E-06	ATP binding cassette subfamily A member 3 [HGNC:33]
EXOC3L2	2.545731	2.47E-06	exocyst complex component 3 like 2 [HGNC:30162]
TMC5	1.084584	2.51E-06	transmembrane channel like 5 [HGNC:22999]
CRTAM	1.400341	2.62E-06	cytotoxic and regulatory T cell molecule [HGNC:24313]
GPR137B	0.751455	2.71E-06	G protein-coupled receptor 137B [HGNC:11862]
CYB5D2	0.603906	2.85E-06	cytochrome b5 domain containing 2 [HGNC:28471]
NR1H3	0.626549	2.88E-06	nuclear receptor subfamily 1 group H member 3 [HGNC:7966]
AC090204.1	0.924170	2.94E-06	novel transcript
SLC4A3	0.858354	3.19E-06	solute carrier family 4 member 3 [HGNC:11029]
CCDC113	1.353533	3.33E-06	coiled-coil domain containing 113 [HGNC:25002]
TMEM79	0.989909	3.47E-06	transmembrane protein 79 [HGNC:28196]
PRR16	0.754801	3.48E-06	proline rich 16 [HGNC:29654]
KNDC1	1.037931	3.78E-06	kinase non-catalytic C-lobe domain containing 1 [HGNC:29374]
FBXL16	1.604226	3.92E-06	F-box and leucine rich repeat protein 16 [HGNC:14150]
TRAPPC6A	1.285093	4.14E-06	trafficking protein particle complex 6A [HGNC:23069]
PCDHGB2	0.624502	4.23E-06	protocadherin gamma subfamily B, 2 [HGNC:8709]
SLC2A5	2.946977	4.27E-06	solute carrier family 2 member 5 [HGNC:11010]
FBXW9	0.713807	4.47E-06	F-box and WD repeat domain containing 9 [HGNC:28136]
EPN3	1.331060	4.49E-06	epsin 3 [HGNC:18235]
VLDLR	0.695601	4.73E-06	very low density lipoprotein receptor [HGNC:12698]
FAAH	1.265002	5.00E-06	fatty acid amide hydrolase [HGNC:3553]
TMEM131L	0.631061	5.75E-06	transmembrane 131 like [HGNC:29146]
PIFO	1.578101	5.88E-06	primary cilia formation [HGNC:27009]
ZNF74	0.855798	6.22E-06	zinc finger protein 74 [HGNC:13144]
CST2	0.661395	6.48E-06	cystatin SA [HGNC:2474]
APOL3	0.979841	6.64E-06	apolipoprotein L3 [HGNC:14868]
CSF1R	0.936166	6.67E-06	colony stimulating factor 1 receptor [HGNC:2433]
CAPS	1.025344	7.08E-06	calcyphosine [HGNC:1487]
TMEM144	0.789012	7.16E-06	transmembrane protein 144 [HGNC:25633]
OSBPL6	0.585746	7.19E-06	oxysterol binding protein like 6 [HGNC:16388]
SH3BP5	0.866353	7.54E-06	SH3 domain binding protein 5 [HGNC:10827]
PNPLA7	1.369851	7.59E-06	patatin like phospholipase domain containing 7 [HGNC:24768]
TMEM53	0.704328	7.63E-06	transmembrane protein 53 [HGNC:26186]
CPVL	0.600666	7.89E-06	carboxypeptidase, vitellogenic like [HGNC:14399]
TRIB1	0.789980	8.19E-06	tribbles pseudokinase 1 [HGNC:16891]
CLBA1	0.767253	8.27E-06	clathrin binding box of aftiphilin containing 1 [HGNC:20126]
CILP2	0.988430	8.47E-06	cartilage intermediate layer protein 2 [HGNC:24213]
AC092171.3	1.492226	8.72E-06	uncharacterized LOC100129484 [NCBI :100129484]

TNFSF12	0.628417	8.74E-06	TNF superfamily member 12 [HGNC:11927]
LBHD1	1.292048	8.85E-06	LBH domain containing 1 [HGNC:28351]
LINC02475	1.130700	9.33E-06	long intergenic non-protein coding RNA 2475 [HGNC:53418]
C3AR1	1.807920	9.35E-06	complement C3a receptor 1 [HGNC:1319]
DMKN	1.044037	1.03E-05	dermokine [HGNC:25063]
NMRK1	0.897621	1.15E-05	nicotinamide riboside kinase 1 [HGNC:26057]
MAFG-DT	0.917356	1.17E-05	MAFG divergent transcript [HGNC:43649]
RAB3IL1	0.645736	1.20E-05	RAB3A interacting protein like 1 [HGNC:9780]
TUB	0.722943	1.24E-05	tubby bipartite transcription factor [HGNC:12406]
SLC9A3-AS1	0.755333	1.27E-05	SLC9A3 antisense RNA 1 [HGNC:40550]
ETFB	0.613564	1.35E-05	electron transfer flavoprotein subunit beta [HGNC:3482]
PCDHB8	1.712280	1.35E-05	protocadherin beta 8 [HGNC:8693]
FN3K	1.046963	1.35E-05	fructosamine 3 kinase [HGNC:24822]
ATOH8	0.870652	1.37E-05	atonal bHLH transcription factor 8 [HGNC:24126]
TFAP2E	1.027356	1.40E-05	transcription factor AP-2 epsilon [HGNC:30774]
ZMIZ1-AS1	1.383929	1.40E-05	ZMIZ1 antisense RNA 1 [HGNC:27433]
TMEM150A	0.784082	1.44E-05	transmembrane protein 150A [HGNC:24677]
HECW1	0.981784	1.60E-05	HECT, C2 and WW domain containing E3 ubiquitin protein ligase 1 [HGNC:22195]
SDHAF4	0.593683	1.61E-05	succinate dehydrogenase complex assembly factor 4 [HGNC:20957]
ZNF432	0.622944	1.66E-05	zinc finger protein 432 [HGNC:20810]
SSBP2	0.657270	1.76E-05	single stranded DNA binding protein 2 [HGNC:15831]
IL31RA	0.609122	1.79E-05	interleukin 31 receptor A [HGNC:18969]
PGGHG	0.612308	1.96E-05	protein-glucosylgalactosylhydroxylysine glucosidase [HGNC:26210]
AGPAT4	0.596789	1.99E-05	1-acylglycerol-3-phosphate O-acyltransferase 4 [HGNC:20885]
RGS14	0.886822	2.06E-05	regulator of G protein signaling 14 [HGNC:9996]
ZSWIM4	0.853761	2.15E-05	zinc finger SWIM-type containing 4 [HGNC:25704]
LMBR1L	0.639274	2.17E-05	limb development membrane protein 1 like [HGNC:18268]
RHOJ	0.733290	2.56E-05	ras homolog family member J [HGNC:688]
ERVMER34-1	1.194835	2.74E-05	endogenous retrovirus group MER34 member 1, envelope [HGNC:42970]
FBXL19-AS1	0.885644	2.94E-05	FBXL19 antisense RNA 1 [HGNC:27557]
RNASEL	0.856172	2.94E-05	ribonuclease L [HGNC:10050]
AF121898.1	1.289727	2.98E-05	novel transcript, antisense to CNBD1
TSPAN10	1.211691	3.07E-05	tetraspanin 10 [HGNC:29942]
CDK20	0.943886	3.36E-05	cyclin dependent kinase 20 [HGNC:21420]
SIRPA	0.733149	3.50E-05	signal regulatory protein alpha [HGNC:9662]
AC089983.1	0.831833	3.80E-05	novel transcript, antisense to TXNRD1
LY96	1.121128	4.27E-05	lymphocyte antigen 96 [HGNC:17156]
SLC29A4	0.833116	4.27E-05	solute carrier family 29 member 4 [HGNC:23097]
NFATC1	0.922758	4.33E-05	nuclear factor of activated T cells 1 [HGNC:7775]
PRELID3A	0.844334	5.09E-05	PRELI domain containing 3A [HGNC:24639]
ICA1	0.757187	5.20E-05	islet cell autoantigen 1 [HGNC:5343]
GNG7	0.667792	5.30E-05	G protein subunit gamma 7 [HGNC:4410]
ZDHHC11	1.236491	5.41E-05	zinc finger DHHC-type containing 11 [HGNC:19158]
KLHL3	1.213120	5.45E-05	kelch like family member 3 [HGNC:6354]
AC068580.1	0.945205	5.48E-05	novel transcript

HCN2	1.655286	5.56E-05	hyperpolarization activated cyclic nucleotide gated potassium and sodium channel 2 [HGNC:4846]
PCDHB4	2.425229	5.63E-05	protocadherin beta 4 [HGNC:8689]
TTYH2	1.819959	5.93E-05	tweety family member 2 [HGNC:13877]
FAM86HP	1.136000	5.93E-05	family with sequence similarity 86 member H, pseudogene [HGNC:42359]
GATA6-AS1	0.964377	5.93E-05	GATA6 antisense RNA 1 (head to head) [HGNC:48840]
NWD1	1.176458	6.00E-05	NACHT and WD repeat domain containing 1 [HGNC:27619]
WISP2	1.502280	6.21E-05	WNT1 inducible signaling pathway protein 2 [HGNC:12770]
LMNTD2	0.776158	6.40E-05	lamin tail domain containing 2 [HGNC:28561]
BDH2	0.710705	6.56E-05	3-hydroxybutyrate dehydrogenase 2 [HGNC:32389]
MRAS	0.719414	6.80E-05	muscle RAS oncogene homolog [HGNC:7227]
TTC30B	1.011176	7.16E-05	tetratricopeptide repeat domain 30B [HGNC:26425]
CCDC121	1.083874	7.43E-05	coiled-coil domain containing 121 [HGNC:25833]
CEBPA	6.061307	7.54E-05	CCAAT enhancer binding protein alpha [HGNC:1833]
JAKMIP3	2.845941	7.85E-05	Janus kinase and microtubule interacting protein 3 [HGNC:23523]
NAT8L	0.980148	8.83E-05	N-acetyltransferase 8 like [HGNC:26742]
NUDT18	0.855099	9.54E-05	nudix hydrolase 18 [HGNC:26194]
LINC02593	2.154491	9.63E-05	long intergenic non-protein coding RNA 2593 [HGNC:53933]
PCSK4	0.885559	9.79E-05	proprotein convertase subtilisin/kexin type 4 [HGNC:8746]
DENND6B	0.958892	9.86E-05	DENN domain containing 6B [HGNC:32690]
LRRC3	0.620146	0.000101	leucine rich repeat containing 3 [HGNC:14965]
NDP	1.306348	0.000114	NDP, norrin cystine knot growth factor [HGNC:7678]
PDE1A	3.123258	0.000122	phosphodiesterase 1A [HGNC:8774]
ADGRB1	2.903997	0.000123	adhesion G protein-coupled receptor B1 [HGNC:943]
IGHV3-15	0.640462	0.000136	immunoglobulin heavy variable 3-15 [HGNC:5582]
PCOLCE2	1.051301	0.000151	procollagen C-endopeptidase enhancer 2 [HGNC:8739]
MAPT	1.508765	0.000156	microtubule associated protein tau [HGNC:6893]
PLXDC1	1.626168	0.000159	plexin domain containing 1 [HGNC:20945]
DUXAP9	1.350010	0.000165	double homeobox A pseudogene 9 [HGNC:32188]
HMCN1	1.087968	0.000178	hemicentin 1 [HGNC:19194]
LENG9	0.851764	0.000181	leukocyte receptor cluster member 9 [HGNC:16306]
COL6A3	0.633254	0.000182	collagen type VI alpha 3 chain [HGNC:2213]
FRMD4B	0.641979	0.000183	FERM domain containing 4B [HGNC:24886]
PINK1-AS	0.859882	0.000183	PINK1 antisense RNA [HGNC:38872]
OPRL1	1.128496	0.000187	opioid related nociceptin receptor 1 [HGNC:8155]
CAMSAP3	0.882835	0.000208	calmodulin regulated spectrin associated protein family member 3 [HGNC:29307]
AL645608.2	2.090764	0.000225	novel transcript
KIF26B	0.980956	0.000227	kinesin family member 26B [HGNC:25484]
AC022784.1	0.961371	0.000237	uncharacterized LOC101929128 [NCBI :101929128]
TOGARAM2	0.838534	0.000239	TOG array regulator of axonemal microtubules 2 [HGNC:33715]
MAPK8IP2	2.033873	0.000248	mitogen-activated protein kinase 8 interacting protein 2 [HGNC:6883]
DNAJC27-AS1	2.245662	0.00025	DNAJC27 antisense RNA 1 [HGNC:42943]
RHBDL1	0.919105	0.000261	rhomboid like 1 [HGNC:10007]
ST8SIA1	0.950639	0.000283	ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 1 [HGNC:10869]
ADSSL1	1.621558	0.000284	adenylosuccinate synthase like 1 [HGNC:20093]
LINC01910	1.681520	0.000285	long intergenic non-protein coding RNA 1910 [HGNC:52729]

NBEA	0.647294	0.000286	neurobeachin [HGNC:7648]
PCDHB12	1.025099	0.000288	protocadherin beta 12 [HGNC:8683]
CPLX1	1.320736	0.000294	complexin 1 [HGNC:2309]
FNDC4	0.587181	0.000294	fibronectin type III domain containing 4 [HGNC:20239]
RIDA	0.592696	0.000319	reactive intermediate imine deaminase A homolog [HGNC:16897]
PODXL2	0.623696	0.00033	podocalyxin like 2 [HGNC:17936]
PRRT4	3.129121	0.000345	proline rich transmembrane protein 4 [HGNC:37280]
FAM86FP	0.751969	0.000353	family with sequence similarity 86 member F, pseudogene [HGNC:42357]
BBS12	0.726801	0.000356	Bardet-Biedl syndrome 12 [HGNC:26648]
IQCA1	1.951708	0.000365	IQ motif containing with AAA domain 1 [HGNC:26195]
CYP2S1	0.755267	0.000372	cytochrome P450 family 2 subfamily S member 1 [HGNC:15654]
GALNT12	0.673950	0.000385	polypeptide N-acetylgalactosaminyltransferase 12 [HGNC:19877]
UST	0.585454	0.000385	uronyl 2-sulfotransferase [HGNC:17223]
CCDC78	0.676361	0.000386	coiled-coil domain containing 78 [HGNC:14153]
PRR5	1.497374	0.000389	proline rich 5 [HGNC:31682]
FLJ20021	0.583567	0.000393	uncharacterized LOC90024 [NCBI :90024]
CCDC28B	0.812464	0.000396	coiled-coil domain containing 28B [HGNC:28163]
AP001362.2	0.827944	0.000398	uncharacterized LOC105369347 [NCBI :105369347]
DPCD	0.650162	0.000417	deleted in primary ciliary dyskinesia homolog (mouse) [HGNC:24542]
BMP6	1.378184	0.000439	bone morphogenetic protein 6 [HGNC:1073]
TMPRSS9	1.034340	0.000446	transmembrane serine protease 9 [HGNC:30079]
BOC	0.994021	0.000461	BOC cell adhesion associated, oncogene regulated [HGNC:17173]
PTCH2	1.275496	0.000466	patched 2 [HGNC:9586]
AL590004.3	1.135852	0.000481	novel transcript
SH3YL1	0.612966	0.000523	SH3 and SYLF domain containing 1 [HGNC:29546]
CCDC18-AS1	0.600487	0.000532	CCDC18 antisense RNA 1 [HGNC:52262]
RBP1	1.422944	0.000563	retinol binding protein 1 [HGNC:9919]
SNAI3-AS1	1.377336	0.000573	SNAI3 antisense RNA 1 [HGNC:28327]
DIRAS1	1.720077	0.00059	DIRAS family GTPase 1 [HGNC:19127]
AP001107.9	1.105739	0.000597	uncharacterized LOC102724064 [NCBI :102724064]
GAPLINC	1.535015	0.000626	gastric adenocarcinoma associated, positive CD44 regulator, long intergenic non-coding RNA [HGNC:51308]
RHOBTB1	0.691410	0.00063	Rho related BTB domain containing 1 [HGNC:18738]
CLDN15	0.597441	0.000639	claudin 15 [HGNC:2036]
AP006621.3	0.731984	0.000677	uncharacterized LOC171391 [NCBI :171391]
WBP1	0.730257	0.000681	WW domain binding protein 1 [HGNC:12737]
ZNF287	1.057261	0.000687	zinc finger protein 287 [HGNC:13502]
BCO1	2.229970	0.00069	beta-carotene oxygenase 1 [HGNC:13815]
TNFRSF18	1.864684	0.000693	TNF receptor superfamily member 18 [HGNC:11914]
THNSL2	3.997404	0.000697	threonine synthase like 2 [HGNC:25602]
DIXDC1	0.767507	0.000701	DIX domain containing 1 [HGNC:23695]
CLUAP1	0.590971	0.000794	clusterin associated protein 1 [HGNC:19009]
AP001453.3	1.856818	0.000824	novel transcript
ADPRH	1.462745	0.000845	ADP-ribosylarginine hydrolase [HGNC:269]
TBC1D3D	1.175577	0.000871	TBC1 domain family member 3D [HGNC:28944]
IGFBP2	2.972759	0.000876	insulin like growth factor binding protein 2 [HGNC:5471]

APOBEC3G	0.831545	0.000882	apolipoprotein B mRNA editing enzyme catalytic subunit 3G [HGNC:17357]
CNKSR1	0.784379	0.000889	connector enhancer of kinase suppressor of Ras 1 [HGNC:19700]
LEPR	0.623777	0.000898	leptin receptor [HGNC:6554]
PCDHB16	1.600255	0.000907	protocadherin beta 16 [HGNC:14546]
KCNQ1	2.140397	0.000911	potassium voltage-gated channel subfamily Q member 1 [HGNC:6294]
TMEM198B	0.760648	0.000929	transmembrane protein 198B (pseudogene) [HGNC:43629]
ITLN1	5.169627	0.000955	intelectin 1 [HGNC:18259]
AC004980.1	0.612274	0.000961	uncharacterized LOC100133091 [NCBI :100133091]
CEACAM19	0.635392	0.00097	carcinoembryonic antigen related cell adhesion molecule 19 [HGNC:31951]
LYNX1	1.199366	0.001061	Ly6/neurotoxin 1 [HGNC:29604]
SLC34A3	0.783871	0.001071	solute carrier family 34 member 3 [HGNC:20305]
PCYOX1L	0.626101	0.001142	prenylcysteine oxidase 1 like [HGNC:28477]
GPC6	0.698050	0.00115	glypican 6 [HGNC:4454]
SPON1	1.468581	0.001166	spondin 1 [HGNC:11252]
AC015909.1	2.653612	0.00118	novel transcript
LPAR5	1.103720	0.001184	lysophosphatidic acid receptor 5 [HGNC:13307]
AC091729.3	0.814111	0.001214	uncharacterized LOC101927021 [NCBI :101927021]
AP002807.1	0.720838	0.00124	novel transcript, antisense to CHKA
COBL	1.494083	0.001246	cordon-bleu WH2 repeat protein [HGNC:22199]
LINC02041	2.411524	0.001318	long intergenic non-protein coding RNA 2041 [HGNC:52881]
ELFN1	2.311205	0.001379	extracellular leucine rich repeat and fibronectin type III domain containing 1 [HGNC:33154]
TRPM3	1.649162	0.001382	transient receptor potential cation channel subfamily M member 3 [HGNC:17992]
AP006284.1	1.082370	0.001455	novel transcript
TTC30A	0.709981	0.001508	tetratricopeptide repeat domain 30A [HGNC:25853]
C11orf45	0.757411	0.001516	chromosome 11 open reading frame 45 [HGNC:28584]
CCDC159	0.783140	0.001525	coiled-coil domain containing 159 [HGNC:26996]
LPP-AS2	0.911174	0.001593	LPP antisense RNA 2 [HGNC:27952]
SYT15	0.856323	0.00162	synaptotagmin 15 [HGNC:17167]
TMEM91	1.190748	0.001641	transmembrane protein 91 [HGNC:32393]
PKD1L2	1.522690	0.001662	polycystin 1 like 2 (gene/pseudogene) [HGNC:21715]
AP001469.1	1.130745	0.001683	novel transcript
KRT14	2.596667	0.001688	keratin 14 [HGNC:6416]
LINC01139	1.125087	0.00169	long intergenic non-protein coding RNA 1139 [HGNC:27924]
TMEM67	0.663287	0.001703	transmembrane protein 67 [HGNC:28396]
RAB26	0.711002	0.001704	RAB26, member RAS oncogene family [HGNC:14259]
RN7SL689P	2.263897	0.00178	RNA, 7SL, cytoplasmic 689, pseudogene [HGNC:46705]
OLMALINC	0.820826	0.001836	oligodendrocyte maturation-associated long intergenic non-coding RNA [HGNC:28060]
PRSS35	1.499408	0.001858	serine protease 35 [HGNC:21387]
AL391988.1	0.940517	0.001869	novel transcript, antisense to SLC18A2
SYT1	0.670377	0.002121	synaptotagmin 1 [HGNC:11509]
PEX11G	1.253255	0.002233	peroxisomal biogenesis factor 11 gamma [HGNC:20208]
TJP3	1.400678	0.002236	tight junction protein 3 [HGNC:11829]
PCDHGA4	0.832131	0.00235	protocadherin gamma subfamily A, 4 [HGNC:8702]
SNAI3	1.606017	0.002387	snail family transcriptional repressor 3 [HGNC:18411]
GBGT1	1.195003	0.002411	globoside alpha-1,3-N-acetylgalactosaminyltransferase 1 (FORS blood group) [HGNC:20460]

BRSK2	1.545628	0.002412	BR serine/threonine kinase 2 [HGNC:11405]
FAM117B	0.738423	0.002433	family with sequence similarity 117 member B [HGNC:14440]
ADAM28	0.647169	0.002496	ADAM metallopeptidase domain 28 [HGNC:206]
RPARP-AS1	0.687388	0.002525	RPARP antisense RNA 1 [HGNC:45238]
CD33	3.416168	0.002531	CD33 molecule [HGNC:1659]
ITGAX	4.292470	0.002557	integrin subunit alpha X [HGNC:6152]
AC105942.1	0.669611	0.002583	hCG2028352-like [NCBI :729970]
CCDC191	1.178509	0.002592	coiled-coil domain containing 191 [HGNC:29272]
CASP1	1.202287	0.002659	caspase 1 [HGNC:1499]
DISP2	0.621257	0.002679	dispatched RND transporter family member 2 [HGNC:19712]
PAPPA	0.790053	0.002696	pappalysin 1 [HGNC:8602]
MSX2	1.192243	0.002708	msh homeobox 2 [HGNC:7392]
SLC9A3	1.024535	0.002778	solute carrier family 9 member A3 [HGNC:11073]
U62631.1	1.962068	0.002794	novel transcript
CACNB3	0.594880	0.00282	calcium voltage-gated channel auxiliary subunit beta 3 [HGNC:1403]
AC005476.2	0.905783	0.00283	novel transcript, antisense to SIPA1L1
CCDC85A	2.598019	0.002873	coiled-coil domain containing 85A [HGNC:29400]
ATP6V0D2	1.300374	0.002903	ATPase H+ transporting V0 subunit d2 [HGNC:18266]
SCARF1	0.685937	0.002972	scavenger receptor class F member 1 [HGNC:16820]
AC004890.2	0.861984	0.002974	AI894139 pseudogene [NCBI :155060]
YJEFN3	0.649183	0.003047	YjeF N-terminal domain containing 3 [HGNC:24785]
SPON2	0.885046	0.003086	spondin 2 [HGNC:11253]
AP000695.1	1.711276	0.003165	novel transcript
C17orf100	1.201848	0.003167	chromosome 17 open reading frame 100 [HGNC:34494]
IGHV1-14	1.237974	0.00328	immunoglobulin heavy variable 1-14 (pseudogene) [HGNC:5547]
ZBTB7C	0.925595	0.003494	zinc finger and BTB domain containing 7C [HGNC:31700]
TCAF1P1	0.752175	0.00353	TRPM8 channel associated factor 1 pseudogene 1 [HGNC:33604]
CD8A	1.130569	0.003541	CD8a molecule [HGNC:1706]
AC097359.2	0.902855	0.00363	novel transcript, antisense to GOLGA4
MAN1B1-DT	0.656619	0.00363	MAN1B1 divergent transcript [HGNC:48715]
CALHM3	0.871786	0.003633	calcium homeostasis modulator 3 [HGNC:23458]
PCAT7	1.144837	0.003733	prostate cancer associated transcript 7 [HGNC:48824]
CCL5	1.884049	0.003886	C-C motif chemokine ligand 5 [HGNC:10632]
RIMBP3	1.073340	0.003889	RIMS binding protein 3 [HGNC:29344]
KRT34	1.232657	0.003907	keratin 34 [HGNC:6452]
CCDC188	1.245825	0.003912	coiled-coil domain containing 188 [HGNC:51899]
C4B	2.122212	0.003924	complement C4B (Chido blood group) [HGNC:1324]
AC110285.3	1.903668	0.004122	novel transcript
LINC00346	0.844601	0.004269	long intergenic non-protein coding RNA 346 [HGNC:27492]
NUDT7	1.726903	0.004361	nudix hydrolase 7 [HGNC:8054]
FAM86B3P	0.769857	0.004389	family with sequence similarity 86 member B3, pseudogene [HGNC:44371]
TBC1D3G	1.632826	0.004424	TBC1 domain family member 3G [HGNC:29860]
EFNA1	0.640558	0.004432	ephrin A1 [HGNC:3221]
ANGPTL4	0.905967	0.004493	angiopoietin like 4 [HGNC:16039]
FAM110B	1.072543	0.004511	family with sequence similarity 110 member B [HGNC:28587]

PPARGC1A	0.601403	0.004524	PPARG coactivator 1 alpha [HGNC:9237]
RCN1P2	0.860486	0.004622	reticulocalbin 1 pseudogene 2 [HGNC:39204]
KLF7	0.645682	0.004628	Kruppel like factor 7 [HGNC:6350]
COL1A2	0.989112	0.004667	collagen type I alpha 2 chain [HGNC:2198]
LMX1B	2.044183	0.004676	LIM homeobox transcription factor 1 beta [HGNC:6654]
SLC30A3	1.248415	0.004714	solute carrier family 30 member 3 [HGNC:11014]
GP1BA	1.260816	0.004772	glycoprotein Ib platelet subunit alpha [HGNC:4439]
SYT8	1.060602	0.005015	synaptotagmin 8 [HGNC:19264]
CLEC2D	1.188014	0.005049	C-type lectin domain family 2 member D [HGNC:14351]
KRT16	2.539254	0.005189	keratin 16 [HGNC:6423]
TTLL1	0.803085	0.005208	tubulin tyrosine ligase like 1 [HGNC:1312]
AC026471.4	3.254289	0.005356	novel transcript
GTF2IP20	0.745071	0.005363	general transcription factor IIi pseudogene 20 [HGNC:51732]
SERPINA5	2.709440	0.005379	serpin family A member 5 [HGNC:8723]
LINC01085	0.592707	0.005439	long intergenic non-protein coding RNA 1085 [HGNC:27198]
RAB17	0.943741	0.005494	RAB17, member RAS oncogene family [HGNC:16523]
ARVCF	0.809483	0.005505	ARVCF, delta catenin family member [HGNC:728]
NSFP1	4.319941	0.005622	N-ethylmaleimide-sensitive factor pseudogene 1 [HGNC:31405]
TMPRSS2	1.288558	0.005628	transmembrane serine protease 2 [HGNC:11876]
GRIN1	1.123483	0.005628	glutamate ionotropic receptor NMDA type subunit 1 [HGNC:4584]
MT-TV	1.482057	0.00566	mitochondrially encoded tRNA valine [HGNC:7500]
ZFPM2	0.711977	0.005685	zinc finger protein, FOG family member 2 [HGNC:16700]
LINC00887	4.649740	0.00572	long intergenic non-protein coding RNA 887 [HGNC:48574]
IQCK	1.103880	0.00577	IQ motif containing K [HGNC:28556]
CARD16	1.077650	0.005784	caspase recruitment domain family member 16 [HGNC:33701]
TAF12	0.635971	0.005834	TATA-box binding protein associated factor 12 [HGNC:11545]
KCNE1B	1.816123	0.005871	potassium voltage-gated channel subfamily E regulatory subunit 1B [HGNC:52280]
ECEL1P1	3.005870	0.005928	endothelin converting enzyme like 1 pseudogene 1 [HGNC:14017]
ENTPD8	1.451393	0.006189	ectonucleoside triphosphate diphosphohydrolase 8 [HGNC:24860]
LRFN1	1.985037	0.006227	leucine rich repeat and fibronectin type III domain containing 1 [HGNC:29290]
SEPT1	0.950421	0.006227	septin 1 [HGNC:2879]
KRT15	0.630108	0.006237	keratin 15 [HGNC:6421]
TPD52L1	0.730976	0.006248	tumor protein D52 like 1 [HGNC:12006]
AL928654.4	0.747737	0.00641	novel transcript
DLL1	1.543596	0.006642	delta like canonical Notch ligand 1 [HGNC:2908]
RBMS3-AS3	0.911867	0.007083	RBMS3 antisense RNA 3 [HGNC:39989]
SEPT4	0.805481	0.007125	septin 4 [HGNC:9165]
SCG2	0.778971	0.007127	secretogranin II [HGNC:10575]
AL772337.3	3.002915	0.007147	cathepsin L (CTSL) pseudogene
AC008555.5	1.510886	0.007273	uncharacterized LOC102723540 [NCBI :102723540]
AL139220.2	0.786330	0.007325	uncharacterized LOC107984948 [NCBI :107984948]
MROH6	0.603387	0.007338	maestro heat like repeat family member 6 [HGNC:27814]
ZNF773	1.103199	0.007598	zinc finger protein 773 [HGNC:30487]
SLC22A23	0.595938	0.007703	solute carrier family 22 member 23 [HGNC:21106]
LYRM9	1.272157	0.007721	LYR motif containing 9 [HGNC:27314]

AC009414.2	0.669693	0.007739	novel transcript
AZGP1	2.565935	0.007837	alpha-2-glycoprotein 1, zinc-binding [HGNC:910]
SELENOP	1.816860	0.00789	selenoprotein P [HGNC:10751]
BTBD8	0.594847	0.007902	BTB domain containing 8 [HGNC:21019]
CARF	0.604853	0.007903	calcium responsive transcription factor [HGNC:14435]
DAW1	1.061043	0.00802	dynein assembly factor with WD repeats 1 [HGNC:26383]
HLF	1.352261	0.008117	HLF, PAR bZIP transcription factor [HGNC:4977]
SP2-AS1	0.818902	0.008189	SP2 antisense RNA 1 [HGNC:51341]
SNAI1	1.592372	0.008307	snail family transcriptional repressor 1 [HGNC:11128]
MCF2L	1.139136	0.008316	MCF.2 cell line derived transforming sequence like [HGNC:14576]
TMEM191B	0.634948	0.008396	transmembrane protein 191B [HGNC:33600]
AC132812.1	0.629040	0.008698	U5 small nuclear ribonucleoprotein 200 kDa helicase pseudogene [NCBI :101929240]
MT1A	0.964513	0.008739	metallothionein 1A [HGNC:7393]
TNFRSF14-AS1	1.261637	0.008803	TNFRSF14 antisense RNA 1 [HGNC:26966]
SCX	0.851906	0.00881	scleraxis bHLH transcription factor [HGNC:32322]
FHDC1	0.892622	0.008832	FH2 domain containing 1 [HGNC:29363]
AC025569.1	0.880492	0.008924	uncharacterized LOC105369827 [NCBI :105369827]
ARHGEF37	0.887871	0.009044	Rho guanine nucleotide exchange factor 37 [HGNC:34430]
EFNB3	0.841150	0.009078	ephrin B3 [HGNC:3228]
TRIM73	1.476297	0.009277	tripartite motif containing 73 [HGNC:18162]
PPP1R3G	1.214427	0.009307	protein phosphatase 1 regulatory subunit 3G [HGNC:14945]
LINC02600	3.374551	0.009393	long intergenic non-protein coding RNA 2600 [HGNC:53177]
AC062029.1	1.225288	0.009638	uncharacterized LOC101928403 [NCBI :101928403]
ATP6V1FNB	0.982613	0.009812	ATP6V1F neighbor [HGNC:52392]
NALT1	1.340927	0.009906	NOTCH1 associated lncRNA in T cell acute lymphoblastic leukemia 1 [HGNC:51192]
FBXO36	0.703344	0.010054	F-box protein 36 [HGNC:27020]
AL512408.1	1.205298	0.01022	uncharacterized LOC101928728 [NCBI :101928728]
ACY3	2.415391	0.010382	aminoacylase 3 [HGNC:24104]
SLC35F4	4.458362	0.010468	solute carrier family 35 member F4 [HGNC:19845]
AC005256.1	1.002384	0.010475	novel transcript
AC098614.1	0.902701	0.010507	tropomyosin 4 (TPM4) pseudogene
GSTM2	0.770809	0.010626	glutathione S-transferase mu 2 [HGNC:4634]
AL162727.1	4.430140	0.010726	novel transcript
GPX7	0.678449	0.010791	glutathione peroxidase 7 [HGNC:4559]
LRGUK	1.541466	0.010918	leucine rich repeats and guanylate kinase domain containing [HGNC:21964]
TMC3-AS1	1.415596	0.010943	TMC3 antisense RNA 1 [HGNC:51424]
AC090578.2	3.119019	0.01096	novel transcript
TMEM202-AS1	1.759792	0.010962	TMEM202 antisense RNA 1 [HGNC:53265]
ALG1L8P	1.163684	0.011281	asparagine-linked glycosylation 1-like 8, pseudogene [HGNC:44377]
INSYN2	1.556509	0.011536	inhibitory synaptic factor 2A [HGNC:33859]
FBXW4P1	1.544869	0.011579	F-box and WD repeat domain containing 4 pseudogene 1 [HGNC:13609]
EXOC3-AS1	0.605489	0.011579	EXOC3 antisense RNA 1 [HGNC:25175]
TMEM163	0.801275	0.011708	transmembrane protein 163 [HGNC:25380]
LAMA1	1.116441	0.011747	laminin subunit alpha 1 [HGNC:6481]
KIAA0825	1.510726	0.011753	KIAA0825 [HGNC:28532]

MYO7A	0.908437	0.01177	myosin VIIA [HGNC:7606]
RPS10P7	0.891726	0.011841	ribosomal protein S10 pseudogene 7 [HGNC:36423]
AC137630.3	1.838503	0.011899	novel transcript, antisense to IMPDH2 and QRICH1
KU-MEL-3	2.816876	0.0128	uncharacterized LOC497048 [NCBI :497048]
HLX	0.835510	0.012837	H2.0 like homeobox [HGNC:4978]
RALY-AS1	0.752473	0.012957	RALY antisense RNA 1 [HGNC:50743]
SRRM2-AS1	0.977125	0.013144	SRRM2 antisense RNA 1 [HGNC:44162]
LINC01503	1.434987	0.01323	long intergenic non-protein coding RNA 1503 [HGNC:51184]
MT-CYB	0.596072	0.013319	mitochondrially encoded cytochrome b [HGNC:7427]
PLEKHA8P1	0.761475	0.013584	pleckstrin homology domain containing A8 pseudogene 1 [HGNC:30222]
C18orf32	1.008655	0.013637	chromosome 18 open reading frame 32 [HGNC:31690]
AL133355.1	0.613698	0.013892	novel transcript, overlapping to OBFC1
TLR8-AS1	0.704344	0.014005	TLR8 antisense RNA 1 [HGNC:40720]
CIART	1.548507	0.014232	circadian associated repressor of transcription [HGNC:25200]
IHH	1.680035	0.014396	indian hedgehog [HGNC:5956]
ADM2	0.825165	0.014456	adrenomedullin 2 [HGNC:28898]
SPATC1	2.635072	0.014621	spermatogenesis and centriole associated 1 [HGNC:30510]
IGFBP5	2.754196	0.014695	insulin like growth factor binding protein 5 [HGNC:5474]
MYCBPAP	1.662346	0.014864	MYCBP associated protein [HGNC:19677]
AC120057.3	1.682727	0.014887	TEC
APBA1	0.730192	0.014891	amyloid beta precursor protein binding family A member 1 [HGNC:578]
LINC00861	1.900356	0.014894	long intergenic non-protein coding RNA 861 [HGNC:45133]
APOE	1.224751	0.014896	apolipoprotein E [HGNC:613]
PCDH9	0.682148	0.014941	protocadherin 9 [HGNC:8661]
ADAM11	0.903247	0.015179	ADAM metallopeptidase domain 11 [HGNC:189]
MELTF-AS1	0.664142	0.01528	MELTF antisense RNA 1 [HGNC:40373]
AK4P1	0.619435	0.015512	adenylate kinase 4 pseudogene 1 [HGNC:364]
PANX2	0.652412	0.015522	pannexin 2 [HGNC:8600]
MAGED4	4.514616	0.01558	MAGE family member D4 [HGNC:23793]
MFNG	0.953124	0.015614	MFNG O-fucosylpeptide 3-beta-N-acetylglucosaminyltransferase [HGNC:7038]
LRCOL1	3.737053	0.015665	leucine rich colipase like 1 [HGNC:44160]
AC010980.1	1.825225	0.01576	uncharacterized LOC440934 [NCBI :440934]
SYNPO2	1.513397	0.015916	synaptopodin 2 [HGNC:17732]
MIR4787	1.999875	0.015963	microRNA 4787 [HGNC:41653]
BOK-AS1	2.094143	0.015989	BOK antisense RNA 1 [HGNC:35125]
RNF165	1.345810	0.016219	ring finger protein 165 [HGNC:31696]
RMDN2	0.592712	0.01645	regulator of microtubule dynamics 2 [HGNC:26567]
AC096733.2	0.740085	0.016523	novel transcript
HEXIM2	0.878077	0.01692	hexamethylene bisacetamide inducible 2 [HGNC:28591]
CEACAM1	0.944503	0.017023	carcinoembryonic antigen related cell adhesion molecule 1 [HGNC:1814]
LRRC23	0.746128	0.017078	leucine rich repeat containing 23 [HGNC:19138]
AC233723.2	0.994094	0.017404	TEC
LNCOC1	2.157284	0.017481	lncRNA associated with ovarian cancer 1 [HGNC:53947]
COL14A1	2.974628	0.017563	collagen type XIV alpha 1 chain [HGNC:2191]
NEK5	1.716763	0.017659	NIMA related kinase 5 [HGNC:7748]

SLC7A10	3.689624	0.017701	solute carrier family 7 member 10 [HGNC:11058]
FAM66B	1.199746	0.017971	family with sequence similarity 66 member B [HGNC:28890]
AC126755.2	1.111446	0.018073	nuclear pore complex interacting protein family member A5 pseudogene [NCBI :105376752]
CSRNP3	1.114543	0.018478	cysteine and serine rich nuclear protein 3 [HGNC:30729]
PPIL6	1.921912	0.018738	peptidylprolyl isomerase like 6 [HGNC:21557]
CARD9	0.793658	0.019091	caspase recruitment domain family member 9 [HGNC:16391]
ELF3	0.856050	0.019189	E74 like ETS transcription factor 3 [HGNC:3318]
VASH1	0.580471	0.019225	vasohibin 1 [HGNC:19964]
AC026471.1	0.913839	0.019586	novel transcript, antisense to ARMC5
THAP7-AS1	1.698152	0.019588	THAP7 antisense RNA 1 [HGNC:41013]
ADAMTS13	1.205284	0.019593	ADAM metallopeptidase with thrombospondin type 1 motif 13 [HGNC:1366]
AL355353.1	1.896521	0.019603	novel transcript
TMEM221	1.655950	0.01979	transmembrane protein 221 [HGNC:21943]
IGHVIII-13-1	1.367773	0.020139	immunoglobulin heavy variable (III)-13-1 (pseudogene) [HGNC:5693]
HPN	1.916026	0.020142	hepsin [HGNC:5155]
PTGFRN	0.628185	0.020165	prostaglandin F2 receptor inhibitor [HGNC:9601]
ITGB3	0.744630	0.020905	integrin subunit beta 3 [HGNC:6156]
DRC3	0.717583	0.021043	dynein regulatory complex subunit 3 [HGNC:25384]
SLC22A31	1.170181	0.021098	solute carrier family 22 member 31 [HGNC:27091]
AC092919.2	0.976845	0.021098	TEC
AC007743.1	3.600219	0.02138	uncharacterized LOC100129434 [NCBI :100129434]
BCDIN3D	0.755642	0.021474	BCDIN3 domain containing RNA methyltransferase [HGNC:27050]
LYPD1	1.198981	0.021717	LY6/PLAUR domain containing 1 [HGNC:28431]
AC245452.1	0.619970	0.022151	novel transcript, antisense to PPM1F and TOP3B
COL24A1	1.263916	0.022502	collagen type XXIV alpha 1 chain [HGNC:20821]
AC092142.1	2.425743	0.022516	uncharacterized LOC105371363 [NCBI :105371363]
RTN4RL1	2.335701	0.022593	reticulon 4 receptor like 1 [HGNC:21329]
XAF1	0.645301	0.022599	XIAP associated factor 1 [HGNC:30932]
CD300C	3.670040	0.022689	CD300c molecule [HGNC:19320]
LMF1	0.693374	0.02269	lipase maturation factor 1 [HGNC:14154]
AL365203.2	0.821782	0.022931	novel transcript
ABHD1	2.056142	0.023031	abhydrolase domain containing 1 [HGNC:17553]
P2RX1	3.632527	0.023216	purinergic receptor P2X 1 [HGNC:8533]
RFLNB	1.020534	0.023228	refilin B [HGNC:28705]
RORC	1.427365	0.023737	RAR related orphan receptor C [HGNC:10260]
CFAP300	1.086875	0.024096	cilia and flagella associated protein 300 [HGNC:28188]
POT1-AS1	0.859843	0.02417	POT1 antisense RNA 1 [HGNC:49459]
AL353150.1	0.755462	0.02421	novel transcript
SLC16A8	2.340528	0.024639	solute carrier family 16 member 8 [HGNC:16270]
AC139149.1	2.330292	0.02472	novel transcript, antisense to ACTG1
TET1	0.696669	0.02489	tet methylcytosine dioxygenase 1 [HGNC:29484]
MIR1915	2.862731	0.024985	microRNA 1915 [HGNC:35399]
TTC21A	0.684029	0.025017	tetratricopeptide repeat domain 21A [HGNC:30761]
C10orf67	2.802232	0.025108	chromosome 10 open reading frame 67 [HGNC:28716]
ZNF382	0.739282	0.025117	zinc finger protein 382 [HGNC:17409]

WAS	1.077918	0.025479	Wiskott-Aldrich syndrome [HGNC:12731]
NETO2	2.626864	0.025512	neuropilin and tolloid like 2 [HGNC:14644]
STARD5	1.240072	0.025658	StAR related lipid transfer domain containing 5 [HGNC:18065]
MZF1-AS1	0.724945	0.025882	MZF1 antisense RNA 1 [HGNC:51271]
CNTNAP3B	1.193452	0.026198	contactin associated protein like 3B [HGNC:32035]
NPIPA5	3.405411	0.026347	nuclear pore complex interacting protein family member A5 [HGNC:41980]
CHI3L2	1.399582	0.026586	chitinase 3 like 2 [HGNC:1933]
ABCC6	0.652219	0.02701	ATP binding cassette subfamily C member 6 [HGNC:57]
LINC00239	1.598564	0.027154	long intergenic non-protein coding RNA 239 [HGNC:20119]
AL161937.2	1.781312	0.02748	novel transcript
MST1P2	1.166705	0.027643	macrophage stimulating 1 pseudogene 2 [HGNC:7383]
IL3RA	1.100445	0.02766	interleukin 3 receptor subunit alpha [HGNC:6012]
ENPP5	1.761039	0.028049	ectonucleotide pyrophosphatase/phosphodiesterase 5 (putative) [HGNC:13717]
FAM78A	1.121689	0.028116	family with sequence similarity 78 member A [HGNC:25465]
CPLANE2	0.731377	0.028349	ciliogenesis and planar polarity effector 2 [HGNC:28127]
MT-TT	1.020374	0.028618	mitochondrially encoded tRNA threonine [HGNC:7499]
SCN4B	3.707827	0.028696	sodium voltage-gated channel beta subunit 4 [HGNC:10592]
ZNF815P	0.962378	0.028982	zinc finger protein 815, pseudogene [HGNC:22029]
COPZ2	0.591190	0.029374	coatomer protein complex subunit zeta 2 [HGNC:19356]
CFAP70	0.864793	0.029783	cilia and flagella associated protein 70 [HGNC:30726]
AZU1	1.379020	0.030182	azurocidin 1 [HGNC:913]
TCTEX1D2	0.773862	0.030379	Tctex1 domain containing 2 [HGNC:28482]
TESC	1.937644	0.030411	tescalcin [HGNC:26065]
AL135905.1	1.956981	0.03066	novel transcript, antisense to PTP4A1
SIGLEC15	1.487526	0.030825	sialic acid binding Ig like lectin 15 [HGNC:27596]
SIRT4	1.112378	0.031521	sirtuin 4 [HGNC:14932]
MCMDC2	1.784538	0.031978	minichromosome maintenance domain containing 2 [HGNC:26368]
SEMA4G	0.787773	0.032043	semaphorin 4G [HGNC:10735]
CHST4	0.733283	0.03306	carbohydrate sulfotransferase 4 [HGNC:1972]
RETREG1	0.769536	0.033806	reticulophagy regulator 1 [HGNC:25964]
TMEM150C	0.890826	0.03381	transmembrane protein 150C [HGNC:37263]
GOLGA7B	0.721827	0.033827	golgin A7 family member B [HGNC:31668]
CCDC183	0.722856	0.03384	coiled-coil domain containing 183 [HGNC:28236]
GGACT	1.110275	0.034039	gamma-glutamylamine cyclotransferase [HGNC:25100]
AL356583.3	2.778487	0.034374	pseudogene similar to part of tumor necrosis factor receptor superfamily member 13B (TNFRSF13B)
LRRTM3	3.463148	0.034489	leucine rich repeat transmembrane neuronal 3 [HGNC:19410]
NHLRC4	1.375497	0.034513	NHL repeat containing 4 [HGNC:26700]
RAB11B-AS1	1.083357	0.034828	RAB11B antisense RNA 1 [HGNC:44178]
C3orf49	1.437359	0.034902	chromosome 3 open reading frame 49 [HGNC:25190]
CCDC87	2.124540	0.035201	coiled-coil domain containing 87 [HGNC:25579]
IL17D	1.232741	0.03555	interleukin 17D [HGNC:5984]
AL645608.6	3.436053	0.035881	novel transcript
BMX	2.054927	0.036504	BMX non-receptor tyrosine kinase [HGNC:1079]
CACNA1I	3.118511	0.037572	calcium voltage-gated channel subunit alpha1 I [HGNC:1396]
MYCL	1.409839	0.039215	MYCL proto-oncogene, bHLH transcription factor [HGNC:7555]

PRR36	0.964247	0.03952	proline rich 36 [HGNC:26172]
THSD1	0.692334	0.039923	thrombospondin type 1 domain containing 1 [HGNC:17754]
AL139260.1	1.626989	0.0401	uncharacterized LOC105378663 [NCBI :105378663]
DIRAS3	1.625332	0.040137	DIRAS family GTPase 3 [HGNC:687]
PLIN5	1.341810	0.040452	perilipin 5 [HGNC:33196]
HRAT5	1.708173	0.040921	heart tissue-associated transcript 5 [NCBI :102467073]
PDE11A	0.670431	0.041869	phosphodiesterase 11A [HGNC:8773]
LINC01137	0.676478	0.044067	long intergenic non-protein coding RNA 1137 [HGNC:49453]
CACNG7	1.869755	0.044266	calcium voltage-gated channel auxiliary subunit gamma 7 [HGNC:13626]
AL354872.2	1.439998	0.044465	novel transcript
TIMP4	0.639067	0.044524	TIMP metallopeptidase inhibitor 4 [HGNC:11823]
AC027307.3	0.768144	0.044595	novel transcript
BX470102.2	0.602049	0.044983	
PDE6A	1.558279	0.045679	phosphodiesterase 6A [HGNC:8785]
GDPD1	0.845412	0.045884	glycerophosphodiester phosphodiesterase domain containing 1 [HGNC:20883]
AL118558.3	1.086544	0.046217	novel transcript
AL513477.1	0.610345	0.046218	small nuclear ribonucleoprotein polypeptide N pseudogene [NCBI :100129534]
VPS37D	0.637609	0.046447	VPS37D, ESCRT-I subunit [HGNC:18287]
TAC4	2.602869	0.04655	tachykinin 4 [HGNC:16641]
AC017104.1	1.817866	0.046696	novel transcript
NOS3	0.888930	0.046745	nitric oxide synthase 3 [HGNC:7876]
FGFBP1	1.411413	0.04727	fibroblast growth factor binding protein 1 [HGNC:19695]
LINC02280	2.328892	0.047591	long intergenic non-protein coding RNA 2280 [HGNC:53196]
AC124016.2	2.195029	0.047927	novel transcript
MGP	1.876537	0.048735	matrix Gla protein [HGNC:7060]
SLC18A2	1.096618	0.048781	solute carrier family 18 member A2 [HGNC:10935]
ADHFE1	1.457488	0.048851	alcohol dehydrogenase, iron containing 1 [HGNC:16354]
AL365356.5	1.158605	0.048941	uncharacterized LOC105376382 [NCBI :105376382]
GSTT2	1.754229	0.049065	glutathione S-transferase theta 2 (gene/pseudogene) [HGNC:4642]
DEPTOR	1.138230	0.049539	DEP domain containing MTOR interacting protein [HGNC:22953]
CCDC183-AS1	0.670604	0.049746	CCDC183 antisense RNA 1 [HGNC:44105]
AC040160.1	0.996180	0.050066	novel transcript
GTF2IP13	0.679612	0.050478	general transcription factor IIi pseudogene 13 [HGNC:51725]
PKDCC	0.694121	0.050912	protein kinase domain containing, cytoplasmic [HGNC:25123]
C17orf97	0.961790	0.050929	chromosome 17 open reading frame 97 [HGNC:33800]

Gene ID	log2FC	Pvalue	Gene Description
IL11	-3.099759	0	interleukin 11 [:HGNC:5966]
SPAG5	-2.766527	0	sperm associated antigen 5 [:HGNC:13452]
MAPRE1	-1.618089	0	microtubule associated protein RP/EB family member 1 [:HGNC:6890]
IRAK1	-1.172485	2.12E-211	interleukin 1 receptor associated kinase 1 [:HGNC:6112]
MCM3	-1.090979	2.88E-210	minichromosome maintenance complex component 3 [:HGNC:6945]
RRM2	-1.224563	1.47E-193	ribonucleotide reductase regulatory subunit M2 [:HGNC:10452]
NONO	-0.877384	2.16E-181	non-POU domain containing octamer binding [:HGNC:7871]

PPME1	-1.327902	1.61E-169	protein phosphatase methylesterase 1 [:HGNC:30178]
SLC25A5	-0.975241	8.65E-164	solute carrier family 25 member 5 [:HGNC:10991]
ANKRD1	-1.119606	3.96E-163	ankyrin repeat domain 1 [:HGNC:15819]
SCN5A	-2.501719	2.80E-161	sodium voltage-gated channel alpha subunit 5 [:HGNC:10593]
UCA1	-1.698795	2.89E-152	urothelial cancer associated 1 [:HGNC:37126]
IER3	-1.337718	6.28E-150	immediate early response 3 [:HGNC:5392]
BUB1	-1.024233	5.61E-144	BUB1 mitotic checkpoint serine/threonine kinase [:HGNC:1148]
LAMB3	-0.871459	1.85E-142	laminin subunit beta 3 [:HGNC:6490]
TUBB	-0.640114	5.90E-142	tubulin beta class I [:HGNC:20778]
KIFC1	-1.128666	7.25E-142	kinesin family member C1 [:HGNC:6389]
DEK	-0.834153	1.77E-133	DEK proto-oncogene [:HGNC:2768]
KIF4A	-1.249472	4.69E-132	kinesin family member 4A [:HGNC:13339]
NCAPG2	-0.891990	4.87E-131	non-SMC condensin II complex subunit G2 [:HGNC:21904]
PGK1	-0.730775	2.91E-129	phosphoglycerate kinase 1 [:HGNC:8896]
RPL7L1	-0.782776	6.76E-128	ribosomal protein L7 like 1 [:HGNC:21370]
BRPF3	-1.022460	1.74E-123	bromodomain and PHD finger containing 3 [:HGNC:14256]
IDS	-1.178269	2.06E-122	iduronate 2-sulfatase [:HGNC:5389]
RANBP9	-1.161014	9.81E-120	RAN binding protein 9 [:HGNC:13727]
PRRC2A	-0.654856	1.94E-111	proline rich coiled-coil 2A [:HGNC:13918]
TOP2A	-0.873702	1.96E-110	DNA topoisomerase II alpha [:HGNC:11989]
AL161431.1	-1.093144	2.32E-110	novel transcript
TCF19	-1.084088	4.15E-106	transcription factor 19 [:HGNC:11629]
NUP153	-0.980893	6.07E-106	nucleoporin 153 [:HGNC:8062]
DLGAP5	-0.955538	7.59E-105	DLG associated protein 5 [:HGNC:16864]
SRPK1	-0.848304	3.63E-104	SRSF protein kinase 1 [:HGNC:11305]
SMIM13	-1.497763	2.05E-103	small integral membrane protein 13 [:HGNC:27356]
CENPM	-1.877953	4.13E-102	centromere protein M [:HGNC:18352]
DSP	-1.166197	1.99E-97	desmoplakin [:HGNC:3052]
GTSE1	-1.001743	3.75E-96	G2 and S-phase expressed 1 [:HGNC:13698]
AIDA	-1.707073	1.87E-95	axin interactor, dorsalization associated [:HGNC:25761]
FAT3	-2.898650	9.89E-94	FAT atypical cadherin 3 [:HGNC:23112]
MKI67	-0.810659	1.95E-93	marker of proliferation Ki-67 [:HGNC:7107]
ANP32E	-0.756654	7.24E-90	acidic nuclear phosphoprotein 32 family member E [:HGNC:16673]
STMN1	-0.669029	2.03E-88	stathmin 1 [:HGNC:6510]
CDK1	-0.866238	5.10E-88	cyclin dependent kinase 1 [:HGNC:1722]
NCAPD3	-0.884326	1.03E-86	non-SMC condensin II complex subunit D3 [:HGNC:28952]
SSH1	-0.940242	1.33E-86	slingshot protein phosphatase 1 [:HGNC:30579]
SAA1	-2.480347	1.64E-85	serum amyloid A1 [:HGNC:10513]
C6orf106	-0.739097	8.22E-85	chromosome 6 open reading frame 106 [:HGNC:21215]
MYEOV	-0.773701	2.30E-83	myeloma overexpressed [:HGNC:7563]
CDC25A	-1.707647	8.15E-83	cell division cycle 25A [:HGNC:1725]
TMPO	-0.658921	8.69E-83	thymopoietin [:HGNC:11875]
RPS4X	-0.593405	7.67E-82	ribosomal protein S4 X-linked [:HGNC:10424]
ZNF185	-1.116817	2.39E-81	zinc finger protein 185 with LIM domain [:HGNC:12976]
SLC16A3	-0.732744	9.25E-81	solute carrier family 16 member 3 [:HGNC:10924]

ATAD2	-0.856048	1.05E-80	ATPase family, AAA domain containing 2 [:HGNC:30123]
OCRL	-1.285219	2.06E-79	OCRL, inositol polyphosphate-5-phosphatase [:HGNC:8108]
RRM1	-0.735952	9.45E-77	ribonucleotide reductase catalytic subunit M1 [:HGNC:10451]
RACGAP1	-0.863730	1.88E-76	Rac GTPase activating protein 1 [:HGNC:9804]
MTCH1	-0.593245	2.13E-76	mitochondrial carrier 1 [:HGNC:17586]
TPX2	-0.588166	1.69E-74	TPX2, microtubule nucleation factor [:HGNC:1249]
KIF2C	-0.803470	3.73E-74	kinesin family member 2C [:HGNC:6393]
GIT1	-1.108582	5.19E-73	GIT ArfGAP 1 [:HGNC:4272]
SLCO4A1	-1.773128	7.85E-72	solute carrier organic anion transporter family member 4A1 [:HGNC:10953]
ZMPSTE24	-1.002172	7.85E-72	zinc metallopeptidase STE24 [:HGNC:12877]
HMGB3	-1.028110	2.30E-70	high mobility group box 3 [:HGNC:5004]
CCT2	-0.734957	1.18E-69	chaperonin containing TCP1 subunit 2 [:HGNC:1615]
ACTR3	-0.623637	1.29E-69	ARP3 actin related protein 3 homolog [:HGNC:170]
MAMLD1	-1.105623	2.26E-68	mastermind like domain containing 1 [:HGNC:2568]
NCAPH	-1.053758	6.08E-68	non-SMC condensin I complex subunit H [:HGNC:1112]
MEA1	-0.760716	6.61E-68	male-enhanced antigen 1 [:HGNC:6986]
CENPI	-1.447839	7.14E-68	centromere protein I [:HGNC:3968]
MAP7D3	-0.979968	8.29E-68	MAP7 domain containing 3 [:HGNC:25742]
F2RL1	-0.877700	6.88E-66	F2R like trypsin receptor 1 [:HGNC:3538]
ABCF1	-0.729355	8.04E-66	ATP binding cassette subfamily F member 1 [:HGNC:70]
RAD23A	-0.944453	1.16E-65	RAD23 homolog A, nucleotide excision repair protein [:HGNC:9812]
ZWINT	-0.771517	3.12E-65	ZW10 interacting kinetochore protein [:HGNC:13195]
FAM111B	-1.216423	8.98E-65	family with sequence similarity 111 member B [:HGNC:24200]
UBE2A	-0.805745	8.48E-64	ubiquitin conjugating enzyme E2 A [:HGNC:12472]
NCAPG	-0.763969	1.66E-63	non-SMC condensin I complex subunit G [:HGNC:24304]
RAD23B	-0.601452	2.22E-63	RAD23 homolog B, nucleotide excision repair protein [:HGNC:9813]
FKBP5	-0.899067	1.04E-62	FK506 binding protein 5 [:HGNC:3721]
MAGT1	-0.886212	3.45E-62	magnesium transporter 1 [:HGNC:28880]
CCND3	-0.693169	9.37E-62	cyclin D3 [:HGNC:1585]
UHRF1	-0.741615	1.29E-61	ubiquitin like with PHD and ring finger domains 1 [:HGNC:12556]
HK1	-0.639234	1.52E-61	hexokinase 1 [:HGNC:4922]
CDC6	-0.832243	4.52E-61	cell division cycle 6 [:HGNC:1744]
SLC2A4RG	-0.812965	9.44E-61	SLC2A4 regulator [:HGNC:15930]
CCHCR1	-1.040228	2.45E-60	coiled-coil alpha-helical rod protein 1 [:HGNC:13930]
HNRNPAB	-0.626100	3.08E-60	heterogeneous nuclear ribonucleoprotein A/B [:HGNC:5034]
SCAMP1	-1.275944	5.25E-59	secretory carrier membrane protein 1 [:HGNC:10563]
ATL3	-0.649000	1.41E-58	atlastin GTPase 3 [:HGNC:24526]
EZH2	-0.799932	1.84E-58	enhancer of zeste 2 polycomb repressive complex 2 subunit [:HGNC:3527]
TRIM37	-1.045074	4.60E-58	tripartite motif containing 37 [:HGNC:7523]
ERCC6L	-1.337294	5.96E-58	ERCC excision repair 6 like, spindle assembly checkpoint helicase [:HGNC:20794]
TMEM87B	-1.206549	6.36E-58	transmembrane protein 87B [:HGNC:25913]
MIS18BP1	-0.864067	1.16E-57	MIS18 binding protein 1 [:HGNC:20190]
SERPINE1	-0.712584	1.16E-57	serpin family E member 1 [:HGNC:8583]
PKMYT1	-1.653418	3.74E-56	protein kinase, membrane associated tyrosine/threonine 1 [:HGNC:29650]
FAM122B	-0.994633	6.26E-56	family with sequence similarity 122B [:HGNC:30490]

PCNA	-0.684048	4.52E-55	proliferating cell nuclear antigen [:HGNC:8729]
IQGAP3	-0.769117	1.09E-54	IQ motif containing GTPase activating protein 3 [:HGNC:20669]
CKAP2	-0.703855	1.09E-54	cytoskeleton associated protein 2 [:HGNC:1990]
KIF11	-0.711051	1.47E-54	kinesin family member 11 [:HGNC:6388]
STAG2	-0.736126	1.52E-54	stromal antigen 2 [:HGNC:11355]
SHCBP1	-0.860784	2.02E-54	SHC binding and spindle associated 1 [:HGNC:29547]
UNG	-0.840794	4.03E-54	uracil DNA glycosylase [:HGNC:12572]
EHF	-1.495605	5.47E-54	ETS homologous factor [:HGNC:3246]
CKAP2L	-0.971723	7.27E-54	cytoskeleton associated protein 2 like [:HGNC:26877]
POLH	-1.009689	1.16E-53	DNA polymerase eta [:HGNC:9181]
SEMA5A	-3.303242	1.25E-53	semaphorin 5A [:HGNC:10736]
SMARCA1	-0.855386	1.57E-53	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 1 [:HGNC:11097]
SLC29A1	-0.722163	7.23E-53	solute carrier family 29 member 1 (Augustine blood group) [:HGNC:11003]
USP1	-0.653186	1.09E-52	ubiquitin specific peptidase 1 [:HGNC:12607]
C6orf62	-0.581691	1.36E-52	chromosome 6 open reading frame 62 [:HGNC:20998]
GPX3	-0.780913	1.24E-50	glutathione peroxidase 3 [:HGNC:4555]
PARP1	-0.614942	1.50E-50	poly(ADP-ribose) polymerase 1 [:HGNC:270]
CDC5L	-0.732162	5.58E-50	cell division cycle 5 like [:HGNC:1743]
RLIM	-0.965249	1.82E-49	ring finger protein, LIM domain interacting [:HGNC:13429]
MTMR1	-0.856412	2.16E-49	myotubularin related protein 1 [:HGNC:7449]
PHF6	-0.979856	3.26E-49	PHD finger protein 6 [:HGNC:18145]
CUL2	-0.993725	6.05E-49	cullin 2 [:HGNC:2552]
VEGFA	-0.777265	1.21E-48	vascular endothelial growth factor A [:HGNC:12680]
SUPT16H	-0.598232	2.35E-48	SPT16 homolog, facilitates chromatin remodeling subunit [:HGNC:11465]
COL8A1	-0.678317	6.67E-48	collagen type VIII alpha 1 chain [:HGNC:2215]
UBE2C	-0.647295	1.07E-47	ubiquitin conjugating enzyme E2 C [:HGNC:15937]
HELLS	-1.053571	1.71E-47	helicase, lymphoid specific [:HGNC:4861]
PPP2R5D	-0.783004	2.30E-47	protein phosphatase 2 regulatory subunit B'delta [:HGNC:9312]
OPHN1	-0.855904	3.43E-47	oligophrenin 1 [:HGNC:8148]
KRT81	-0.778072	4.61E-47	keratin 81 [:HGNC:6458]
MCM5	-0.749913	4.99E-47	minichromosome maintenance complex component 5 [:HGNC:6948]
CHML	-0.886054	7.19E-47	CHM like, Rab escort protein 2 [:HGNC:1941]
TRAM2	-0.591026	8.19E-47	translocation associated membrane protein 2 [:HGNC:16855]
AC138392.1	-1.803210	1.07E-46	novel pseudogene
MCM6	-0.811787	2.69E-46	minichromosome maintenance complex component 6 [:HGNC:6949]
CDKN3	-0.905910	3.02E-46	cyclin dependent kinase inhibitor 3 [:HGNC:1791]
ARHGAP11A	-0.696055	8.35E-46	Rho GTPase activating protein 11A [:HGNC:15783]
TTC9	-2.415141	9.89E-46	tetratricopeptide repeat domain 9 [:HGNC:20267]
ENO2	-0.773178	1.41E-45	enolase 2 [:HGNC:3353]
DIAPH3	-0.776168	1.43E-45	diaphanous related formin 3 [:HGNC:15480]
GMNN	-0.914956	1.64E-45	geminin, DNA replication inhibitor [:HGNC:17493]
EMD	-0.853087	1.85E-45	emerin [:HGNC:3331]
GDI1	-0.696686	2.95E-45	GDP dissociation inhibitor 1 [:HGNC:4226]
ANKRD13A	-1.028312	3.66E-45	ankyrin repeat domain 13A [:HGNC:21268]
PBK	-0.748470	4.66E-45	PDZ binding kinase [:HGNC:18282]

SPANXB1	-1.330888	5.14E-45	SPANX family member B1 [:HGNC:14329]
MYBL2	-0.864667	6.07E-45	MYB proto-oncogene like 2 [:HGNC:7548]
AURKB	-0.745812	1.16E-44	aurora kinase B [:HGNC:11390]
FOXM1	-0.676283	1.66E-44	forkhead box M1 [:HGNC:3818]
TTL	-0.755631	1.98E-44	tubulin tyrosine ligase [:HGNC:21586]
XPO5	-0.677275	2.36E-44	exportin 5 [:HGNC:17675]
PTGES	-0.909667	2.69E-44	prostaglandin E synthase [:HGNC:9599]
CLSPN	-0.998563	2.78E-44	claspin [:HGNC:19715]
SERINC2	-0.653457	3.51E-44	serine incorporator 2 [:HGNC:23231]
FIGNL1	-1.335176	4.32E-44	fidgetin like 1 [:HGNC:13286]
KNL1	-0.875664	1.05E-43	kinetochore scaffold 1 [:HGNC:24054]
PLS3	-0.609151	1.25E-43	plastin 3 [:HGNC:9091]
DDA1	-0.747854	2.44E-43	DET1 and DDB1 associated 1 [:HGNC:28360]
WDHD1	-0.920631	2.52E-43	WD repeat and HMG-box DNA binding protein 1 [:HGNC:23170]
THOC2	-0.753636	1.02E-42	THO complex 2 [:HGNC:19073]
INCENP	-0.721173	1.07E-42	inner centromere protein [:HGNC:6058]
HBEGF	-0.791157	1.22E-42	heparin binding EGF like growth factor [:HGNC:3059]
CIP2A	-0.794710	1.32E-42	cell proliferation regulating inhibitor of protein phosphatase 2A [:HGNC:29302]
NFYA	-0.782635	1.34E-42	nuclear transcription factor Y subunit alpha [:HGNC:7804]
OGT	-0.744487	2.30E-42	O-linked N-acetylglucosamine (GlcNAc) transferase [:HGNC:8127]
VAMP7	-0.823790	6.78E-42	vesicle associated membrane protein 7 [:HGNC:11486]
PYGL	-0.683149	1.05E-41	glycogen phosphorylase L [:HGNC:9725]
UBR7	-0.790235	1.23E-41	ubiquitin protein ligase E3 component n-recognin 7 (putative) [:HGNC:20344]
LIFR	-0.874584	1.41E-41	LIF receptor alpha [:HGNC:6597]
E2F1	-0.661084	1.72E-41	E2F transcription factor 1 [:HGNC:3113]
FANCE	-1.126577	8.62E-41	FA complementation group E [:HGNC:3586]
HJURP	-0.706806	1.69E-40	Holliday junction recognition protein [:HGNC:25444]
RBMS2	-0.830711	2.09E-40	RNA binding motif single stranded interacting protein 2 [:HGNC:9909]
AC116533.1	-0.741263	2.75E-40	ribosomal protein L36a (RPL36A) pseudogene
TAF8	-0.698713	3.01E-40	TATA-box binding protein associated factor 8 [:HGNC:17300]
GTF2A1	-0.667204	3.02E-40	general transcription factor IIA subunit 1 [:HGNC:4646]
BCL2L13	-1.072945	3.55E-40	BCL2 like 13 [:HGNC:17164]
MAPK14	-0.586806	4.56E-40	mitogen-activated protein kinase 14 [:HGNC:6876]
EBNA1BP2	-0.616271	7.84E-40	EBNA1 binding protein 2 [:HGNC:15531]
MAFF	-0.919080	1.22E-39	MAF bZIP transcription factor F [:HGNC:6780]
VMA21	-0.740326	1.76E-39	VMA21, vacuolar ATPase assembly factor [:HGNC:22082]
TMEM14B	-0.672535	1.82E-39	transmembrane protein 14B [:HGNC:21384]
C2CD3	-0.858134	2.95E-39	C2 calcium dependent domain containing 3 [:HGNC:24564]
MCTS1	-0.908699	4.71E-39	MCTS1, re-initiation and release factor [:HGNC:23357]
BRD9	-0.823717	8.48E-39	bromodomain containing 9 [:HGNC:25818]
GBE1	-0.726819	1.02E-38	1,4-alpha-glucan branching enzyme 1 [:HGNC:4180]
SNX12	-0.713924	1.56E-38	sorting nexin 12 [:HGNC:14976]
SRF	-0.611225	1.86E-38	serum response factor [:HGNC:11291]
STARD7	-0.590361	2.21E-38	StAR related lipid transfer domain containing 7 [:HGNC:18063]
SEPHS1	-0.676213	2.54E-38	selenophosphate synthetase 1 [:HGNC:19685]

DEPDC1B	-0.996656	3.44E-38	DEP domain containing 1B [:HGNC:24902]
PBRM1	-0.603547	5.00E-38	polybromo 1 [:HGNC:30064]
KNTC1	-0.679731	6.63E-38	kinetochore associated 1 [:HGNC:17255]
VBP1	-0.758428	8.63E-38	VHL binding protein 1 [:HGNC:12662]
LSM2	-0.863886	1.11E-37	LSM2 homolog, U6 small nuclear RNA and mRNA degradation associated [:HGNC:13940]
MRPL14	-0.701458	2.51E-37	mitochondrial ribosomal protein L14 [:HGNC:14279]
CDCA5	-0.629294	3.94E-37	cell division cycle associated 5 [:HGNC:14626]
DKC1	-0.617370	4.84E-37	dyskerin pseudouridine synthase 1 [:HGNC:2890]
FANCI	-0.720967	6.08E-37	FA complementation group I [:HGNC:25568]
MCM8	-0.824011	8.79E-37	minichromosome maintenance 8 homologous recombination repair factor [:HGNC:16147]
BRIP1	-0.839164	2.77E-36	BRCA1 interacting protein C-terminal helicase 1 [:HGNC:20473]
CDK2	-0.777824	3.45E-36	cyclin dependent kinase 2 [:HGNC:1771]
MCUR1	-0.606416	3.81E-36	mitochondrial calcium uniporter regulator 1 [:HGNC:21097]
YIPF6	-0.772442	5.79E-36	Yip1 domain family member 6 [:HGNC:28304]
HIPK3	-0.946623	1.33E-35	homeodomain interacting protein kinase 3 [:HGNC:4915]
SERTAD2	-0.718729	1.85E-35	SERTA domain containing 2 [:HGNC:30784]
FEN1	-0.599859	2.35E-35	flap structure-specific endonuclease 1 [:HGNC:3650]
HPRT1	-0.845492	3.88E-35	hypoxanthine phosphoribosyltransferase 1 [:HGNC:5157]
HTATSF1	-0.600353	4.07E-35	HIV-1 Tat specific factor 1 [:HGNC:5276]
SMC2	-0.588809	4.58E-35	structural maintenance of chromosomes 2 [:HGNC:14011]
AGTPBP1	-0.905683	5.07E-35	ATP/GTP binding protein 1 [:HGNC:17258]
MCM10	-0.922115	5.97E-35	minichromosome maintenance 10 replication initiation factor [:HGNC:18043]
CTGF	-0.673358	6.01E-35	connective tissue growth factor [:HGNC:2500]
ERO1A	-0.620505	8.04E-35	endoplasmic reticulum oxidoreductase 1 alpha [:HGNC:13280]
NOL11	-0.752295	1.07E-34	nucleolar protein 11 [:HGNC:24557]
BIRC3	-1.111440	1.16E-34	baculoviral IAP repeat containing 3 [:HGNC:591]
APBB1IP	-1.552588	1.30E-34	amyloid beta precursor protein binding family B member 1 interacting protein [:HGNC:17379]
CMTR1	-0.597124	1.47E-34	cap methyltransferase 1 [:HGNC:21077]
CABLES2	-0.903749	1.50E-34	Cdk5 and Abl enzyme substrate 2 [:HGNC:16143]
ADGRF1	-1.598543	3.39E-34	adhesion G protein-coupled receptor F1 [:HGNC:18990]
GCLM	-0.619412	4.21E-34	glutamate-cysteine ligase modifier subunit [:HGNC:4312]
PPIL1	-0.641681	7.57E-34	peptidylprolyl isomerase like 1 [:HGNC:9260]
DIP2B	-1.029089	8.20E-34	disco interacting protein 2 homolog B [:HGNC:29284]
SLC9A6	-0.850937	1.37E-33	solute carrier family 9 member A6 [:HGNC:11079]
NGFR	-1.098270	2.05E-33	nerve growth factor receptor [:HGNC:7809]
NFKBIA	-0.770505	4.41E-33	NFKB inhibitor alpha [:HGNC:7797]
SAP130	-0.835703	1.21E-32	Sin3A associated protein 130 [:HGNC:29813]
SAMD4A	-0.697508	1.59E-32	sterile alpha motif domain containing 4A [:HGNC:23023]
DOT1L	-0.772954	4.71E-32	DOT1 like histone lysine methyltransferase [:HGNC:24948]
BUB1B	-0.779524	7.04E-32	BUB1 mitotic checkpoint serine/threonine kinase B [:HGNC:1149]
ASPM	-0.607091	7.35E-32	abnormal spindle microtubule assembly [:HGNC:19048]
TMED8	-0.952222	8.79E-32	transmembrane p24 trafficking protein family member 8 [:HGNC:18633]
RAD51AP1	-0.924927	8.98E-32	RAD51 associated protein 1 [:HGNC:16956]
POLE2	-1.149175	1.12E-31	DNA polymerase epsilon 2, accessory subunit [:HGNC:9178]
CDCA4	-0.645346	1.14E-31	cell division cycle associated 4 [:HGNC:14625]

GAS2L3	-0.700903	1.51E-31	growth arrest specific 2 like 3 [:HGNC:27475]
WDR62	-0.730928	2.03E-31	WD repeat domain 62 [:HGNC:24502]
TYMS	-0.786076	2.29E-31	thymidylate synthetase [:HGNC:12441]
AGO2	-0.722335	2.91E-31	argonaute 2, RISC catalytic component [:HGNC:3263]
MYH10	-0.618721	3.27E-31	myosin heavy chain 10 [:HGNC:7568]
KISS1	-3.481904	3.55E-31	KiSS-1 metastasis suppressor [:HGNC:6341]
PRPS1	-0.767212	4.54E-31	phosphoribosyl pyrophosphate synthetase 1 [:HGNC:9462]
GLTP	-0.699225	4.65E-31	glycolipid transfer protein [:HGNC:24867]
FAM111A	-0.680993	4.83E-31	family with sequence similarity 111 member A [:HGNC:24725]
TSC22D4	-0.735343	4.87E-31	TSC22 domain family member 4 [:HGNC:21696]
SGO2	-0.670938	5.55E-31	shugoshin 2 [:HGNC:30812]
ETV4	-0.761136	6.58E-31	ETS variant 4 [:HGNC:3493]
CHEK1	-0.676640	7.26E-31	checkpoint kinase 1 [:HGNC:1925]
E2F8	-0.989853	1.36E-30	E2F transcription factor 8 [:HGNC:24727]
MYO19	-0.812788	2.07E-30	myosin XIX [:HGNC:26234]
RBBP8	-0.670404	3.66E-30	RB binding protein 8, endonuclease [:HGNC:9891]
PSME2	-0.613943	4.84E-30	proteasome activator subunit 2 [:HGNC:9569]
FAM50A	-0.650128	4.92E-30	family with sequence similarity 50 member A [:HGNC:18786]
FAM102B	-0.626742	5.53E-30	family with sequence similarity 102 member B [:HGNC:27637]
E2F3	-0.709383	5.97E-30	E2F transcription factor 3 [:HGNC:3115]
MMGT1	-0.884085	1.16E-29	membrane magnesium transporter 1 [:HGNC:28100]
TRIM26	-0.645425	1.31E-29	tripartite motif containing 26 [:HGNC:12962]
FUBP1	-0.625553	1.53E-29	far upstream element binding protein 1 [:HGNC:4004]
FAM199X	-0.796315	4.49E-29	family with sequence similarity 199, X-linked [:HGNC:25195]
BYSL	-0.694539	4.77E-29	bystin like [:HGNC:1157]
LAGE3	-0.955289	5.11E-29	L antigen family member 3 [:HGNC:26058]
ERAP1	-0.664744	5.81E-29	endoplasmic reticulum aminopeptidase 1 [:HGNC:18173]
KLHL23	-1.153116	6.50E-29	kelch like family member 23 [:HGNC:27506]
APOBEC3B	-1.103249	6.60E-29	apolipoprotein B mRNA editing enzyme catalytic subunit 3B [:HGNC:17352]
MOCS1	-1.597276	7.51E-29	molybdenum cofactor synthesis 1 [:HGNC:7190]
KIF15	-0.810264	7.92E-29	kinesin family member 15 [:HGNC:17273]
RFWD3	-0.595462	7.95E-29	ring finger and WD repeat domain 3 [:HGNC:25539]
CBL	-0.623058	1.45E-28	Cbl proto-oncogene [:HGNC:1541]
CD83	-1.540964	1.77E-28	CD83 molecule [:HGNC:1703]
NDC80	-0.692618	1.89E-28	NDC80, kinetochore complex component [:HGNC:16909]
AQP1	-1.990854	2.18E-28	aquaporin 1 (Colton blood group) [:HGNC:633]
ESPL1	-0.695862	2.19E-28	extra spindle pole bodies like 1, separase [:HGNC:16856]
KIF18B	-0.599674	2.31E-28	kinesin family member 18B [:HGNC:27102]
C1GALT1	-0.697911	3.55E-28	core 1 synthase, glycoprotein-N-acetylgalactosamine 3-beta-galactosyltransferase 1 [:HGNC:24337]
SEMA7A	-0.681534	3.64E-28	semaphorin 7A (John Milton Hagen blood group) [:HGNC:10741]
TIMELESS	-0.634196	6.36E-28	timeless circadian regulator [:HGNC:11813]
ZNF318	-0.632772	6.56E-28	zinc finger protein 318 [:HGNC:13578]
HABP4	-1.158838	7.47E-28	hyaluronan binding protein 4 [:HGNC:17062]
ADGRG1	-0.632799	8.28E-28	adhesion G protein-coupled receptor G1 [:HGNC:4512]
DENND2A	-2.611720	9.50E-28	DENN domain containing 2A [:HGNC:22212]

CSTF2	-0.921376	9.66E-28	cleavage stimulation factor subunit 2 [:HGNC:2484]
CDC45	-0.840981	1.38E-27	cell division cycle 45 [:HGNC:1739]
DDX58	-0.691819	1.91E-27	DExD/H-box helicase 58 [:HGNC:19102]
RBL1	-0.725715	1.97E-27	RB transcriptional corepressor like 1 [:HGNC:9893]
SLAMF7	-1.172723	2.78E-27	SLAM family member 7 [:HGNC:21394]
NEK2	-0.688204	2.88E-27	NIMA related kinase 2 [:HGNC:7745]
TAP2	-0.620991	3.40E-27	transporter 2, ATP binding cassette subfamily B member [:HGNC:44]
POLA1	-0.770088	3.86E-27	DNA polymerase alpha 1, catalytic subunit [:HGNC:9173]
GXYLT1	-0.791921	4.11E-27	glucoside xylosyltransferase 1 [:HGNC:27482]
FGF5	-1.793386	4.28E-27	fibroblast growth factor 5 [:HGNC:3683]
CENPU	-0.585489	5.85E-27	centromere protein U [:HGNC:21348]
RFC4	-0.805771	8.39E-27	replication factor C subunit 4 [:HGNC:9972]
PLK4	-0.714529	9.68E-27	polo like kinase 4 [:HGNC:11397]
SLC10A3	-0.687095	1.59E-26	solute carrier family 10 member 3 [:HGNC:22979]
MZT1	-0.695712	2.39E-26	mitotic spindle organizing protein 1 [:HGNC:33830]
XIAP	-0.655608	2.65E-26	X-linked inhibitor of apoptosis [:HGNC:592]
AKR1C1	-1.352362	2.87E-26	aldo-keto reductase family 1 member C1 [:HGNC:384]
CTPS1	-0.603946	3.09E-26	CTP synthase 1 [:HGNC:2519]
TAF9B	-0.850535	4.83E-26	TATA-box binding protein associated factor 9b [:HGNC:17306]
SPC25	-1.029707	5.38E-26	SPC25, NDC80 kinetochore complex component [:HGNC:24031]
GLA	-0.921515	8.62E-26	galactosidase alpha [:HGNC:4296]
CEP78	-0.815164	1.17E-25	centrosomal protein 78 [:HGNC:25740]
EXO1	-0.888416	1.29E-25	exonuclease 1 [:HGNC:3511]
FAM83D	-0.586634	1.36E-25	family with sequence similarity 83 member D [:HGNC:16122]
WDR76	-0.973548	2.38E-25	WD repeat domain 76 [:HGNC:25773]
PAK1IP1	-0.807536	2.47E-25	PAK1 interacting protein 1 [:HGNC:20882]
DNAJC9	-0.723169	3.30E-25	DnaJ heat shock protein family (Hsp40) member C9 [:HGNC:19123]
NFKB2	-0.610649	3.83E-25	nuclear factor kappa B subunit 2 [:HGNC:7795]
AGFG2	-1.041706	5.16E-25	ArfGAP with FG repeats 2 [:HGNC:5177]
HAS2	-3.146252	6.60E-25	hyaluronan synthase 2 [:HGNC:4819]
PLXNA2	-1.402240	8.61E-25	plexin A2 [:HGNC:9100]
TMEM71	-1.050491	9.28E-25	transmembrane protein 71 [:HGNC:26572]
RHNO1	-0.829419	9.37E-25	RAD9-HUS1-RAD1 interacting nuclear orphan 1 [:HGNC:28206]
ARHGAP19	-0.805281	1.00E-24	Rho GTPase activating protein 19 [:HGNC:23724]
SKA3	-0.777633	1.05E-24	spindle and kinetochore associated complex subunit 3 [:HGNC:20262]
GATA2	-1.306483	1.52E-24	GATA binding protein 2 [:HGNC:4171]
C4orf46	-0.652003	1.54E-24	chromosome 4 open reading frame 46 [:HGNC:27320]
SUZ12	-0.593537	2.24E-24	SUZ12, polycomb repressive complex 2 subunit [:HGNC:17101]
CYR61	-0.938019	3.03E-24	cysteine rich angiogenic inducer 61 [:HGNC:2654]
PARP2	-0.831317	7.46E-24	poly(ADP-ribose) polymerase 2 [:HGNC:272]
PPP1R37	-0.879314	7.72E-24	protein phosphatase 1 regulatory subunit 37 [:HGNC:27607]
DTL	-0.628398	9.17E-24	denticleless E3 ubiquitin protein ligase homolog [:HGNC:30288]
MCAM	-0.667536	1.06E-23	melanoma cell adhesion molecule [:HGNC:6934]
ARNTL2	-0.628902	1.16E-23	aryl hydrocarbon receptor nuclear translocator like 2 [:HGNC:18984]
DUSP5	-0.686957	1.59E-23	dual specificity phosphatase 5 [:HGNC:3071]

SYTL4	-1.636421	2.80E-23	synaptotagmin like 4 [:HGNC:15588]
ESCO2	-0.961587	3.15E-23	establishment of sister chromatid cohesion N-acetyltransferase 2 [:HGNC:27230]
CDH24	-0.862471	3.76E-23	cadherin 24 [:HGNC:14265]
C6orf132	-0.733924	3.91E-23	chromosome 6 open reading frame 132 [:HGNC:21288]
USP49	-1.048440	4.01E-23	ubiquitin specific peptidase 49 [:HGNC:20078]
SMTN	-0.608840	4.27E-23	smoothelin [:HGNC:11126]
ADRB2	-0.882321	4.37E-23	adrenoceptor beta 2 [:HGNC:286]
KLHL15	-0.737654	4.96E-23	kelch like family member 15 [:HGNC:29347]
CHRNB1	-0.784602	6.19E-23	cholinergic receptor nicotinic beta 1 subunit [:HGNC:1961]
CDCA3	-0.607565	8.12E-23	cell division cycle associated 3 [:HGNC:14624]
SOCS4	-0.707001	1.22E-22	suppressor of cytokine signaling 4 [:HGNC:19392]
TRMU	-0.683952	1.63E-22	tRNA 5-methylaminomethyl-2-thiouridylate methyltransferase [:HGNC:25481]
PBDC1	-0.867405	2.19E-22	polysaccharide biosynthesis domain containing 1 [:HGNC:28790]
IL7R	-1.083689	2.25E-22	interleukin 7 receptor [:HGNC:6024]
PRMT5	-0.644371	2.50E-22	protein arginine methyltransferase 5 [:HGNC:10894]
DLG3	-0.852576	3.00E-22	discs large MAGUK scaffold protein 3 [:HGNC:2902]
MASTL	-0.617960	3.33E-22	microtubule associated serine/threonine kinase like [:HGNC:19042]
VPS37B	-0.806852	3.76E-22	VPS37B, ESCRT-I subunit [:HGNC:25754]
MAN1A1	-1.618369	4.40E-22	mannosidase alpha class 1A member 1 [:HGNC:6821]
TP53INP2	-0.637399	9.55E-22	tumor protein p53 inducible nuclear protein 2 [:HGNC:16104]
CNOT6	-0.594688	9.77E-22	CCR4-NOT transcription complex subunit 6 [:HGNC:14099]
GINS1	-0.730908	1.26E-21	GINS complex subunit 1 [:HGNC:28980]
DLC1	-0.639999	1.49E-21	DLC1 Rho GTPase activating protein [:HGNC:2897]
ICAM1	-0.926712	1.55E-21	intercellular adhesion molecule 1 [:HGNC:5344]
PBX2	-0.589017	1.87E-21	PBX homeobox 2 [:HGNC:8633]
SPC24	-0.711466	1.90E-21	SPC24, NDC80 kinetochore complex component [:HGNC:26913]
MIR137HG	-3.440254	2.04E-21	MIR137 host gene [:HGNC:42871]
LIG1	-0.589992	2.04E-21	DNA ligase 1 [:HGNC:6598]
ZNF219	-0.970430	3.06E-21	zinc finger protein 219 [:HGNC:13011]
POLD3	-0.721701	4.65E-21	DNA polymerase delta 3, accessory subunit [:HGNC:20932]
PRIM2	-0.721256	4.75E-21	DNA primase subunit 2 [:HGNC:9370]
RIOK1	-0.737548	7.67E-21	RIO kinase 1 [:HGNC:18656]
RAB27A	-0.879853	8.02E-21	RAB27A, member RAS oncogene family [:HGNC:9766]
BARD1	-0.823578	1.02E-20	BRCA1 associated RING domain 1 [:HGNC:952]
TNFSF15	-2.424330	1.07E-20	TNF superfamily member 15 [:HGNC:11931]
HSPA4L	-0.607159	1.08E-20	heat shock protein family A (Hsp70) member 4 like [:HGNC:17041]
GLCE	-0.690009	1.24E-20	glucuronic acid epimerase [:HGNC:17855]
KDM1B	-0.798475	1.43E-20	lysine demethylase 1B [:HGNC:21577]
HS6ST1	-0.640408	2.19E-20	heparan sulfate 6-O-sulfotransferase 1 [:HGNC:5201]
ATP11C	-0.778324	2.52E-20	ATPase phospholipid transporting 11C [:HGNC:13554]
CD274	-1.102677	2.77E-20	CD274 molecule [:HGNC:17635]
CGAS	-0.590562	2.81E-20	cyclic GMP-AMP synthase [:HGNC:21367]
PGBD1	-0.898925	2.99E-20	piggyBac transposable element derived 1 [:HGNC:19398]
MBNL3	-1.053747	3.29E-20	muscleblind like splicing regulator 3 [:HGNC:20564]
MED20	-0.614961	3.41E-20	mediator complex subunit 20 [:HGNC:16840]

RBX1	-0.592354	4.16E-20	ring-box 1 [:HGNC:9928]
GREB1L	-1.762716	5.73E-20	GREB1 like retinoic acid receptor coactivator [:HGNC:31042]
B3GALT6	-0.754375	5.85E-20	beta-1,3-galactosyltransferase 6 [:HGNC:17978]
PAG1	-0.711189	5.89E-20	phosphoprotein membrane anchor with glycosphingolipid microdomains 1 [:HGNC:30043]
TGFBR3	-0.769198	6.83E-20	transforming growth factor beta receptor 3 [:HGNC:11774]
NFKBIE	-0.703400	8.10E-20	NFKB inhibitor epsilon [:HGNC:7799]
SCML2	-0.904109	8.61E-20	Scm polycomb group protein like 2 [:HGNC:10581]
NUSAP1	-1.013623	9.55E-20	nucleolar and spindle associated protein 1 [:HGNC:18538]
NEURL1B	-0.843085	1.07E-19	neuralized E3 ubiquitin protein ligase 1B [:HGNC:35422]
ARHGEF40	-0.799593	2.42E-19	Rho guanine nucleotide exchange factor 40 [:HGNC:25516]
NEDD4	-0.782991	2.74E-19	neural precursor cell expressed, developmentally down-regulated 4, E3 ubiquitin protein ligase [:HGNC:7727]
PHKA1	-0.994178	2.98E-19	phosphorylase kinase regulatory subunit alpha 1 [:HGNC:8925]
PURA	-0.656035	3.79E-19	purine rich element binding protein A [:HGNC:9701]
ADM	-0.732828	6.38E-19	adrenomedullin [:HGNC:259]
TMEM164	-0.615424	7.76E-19	transmembrane protein 164 [:HGNC:26217]
FANCA	-0.676257	7.85E-19	FA complementation group A [:HGNC:3582]
SRGN	-1.118677	9.89E-19	serglycin [:HGNC:9361]
CTDSPL2	-0.702227	1.23E-18	CTD small phosphatase like 2 [:HGNC:26936]
RFC3	-0.731169	1.23E-18	replication factor C subunit 3 [:HGNC:9971]
LGMN	-0.652351	1.44E-18	legumain [:HGNC:9472]
FBXO34	-0.622607	1.92E-18	F-box protein 34 [:HGNC:20201]
RASGRF1	-1.264370	1.94E-18	Ras protein specific guanine nucleotide releasing factor 1 [:HGNC:9875]
POLR2D	-0.588924	1.95E-18	RNA polymerase II subunit D [:HGNC:9191]
CYFIP2	-1.574019	2.44E-18	cytoplasmic FMR1 interacting protein 2 [:HGNC:13760]
UQCC2	-0.616835	2.47E-18	ubiquinol-cytochrome c reductase complex assembly factor 2 [:HGNC:21237]
AEN	-0.687655	3.02E-18	apoptosis enhancing nuclease [:HGNC:25722]
LINC01224	-0.933106	4.06E-18	long intergenic non-protein coding RNA 1224 [:HGNC:49676]
CENPQ	-0.856904	4.39E-18	centromere protein Q [:HGNC:21347]
MPHOSPH9	-0.615094	4.87E-18	M-phase phosphoprotein 9 [:HGNC:7215]
CSRP2	-1.219838	5.44E-18	cysteine and glycine rich protein 2 [:HGNC:2470]
ARMCX4	-1.447717	6.30E-18	armadillo repeat containing X-linked 4 [:HGNC:28615]
MED12	-0.775652	6.94E-18	mediator complex subunit 12 [:HGNC:11957]
AL109918.1	-1.125475	8.25E-18	uncharacterized LOC730101 [Source:NCBI gene;Acc:730101]
APCS	-4.008266	8.36E-18	amyloid P component, serum [:HGNC:584]
TBC1D22B	-0.732162	9.27E-18	TBC1 domain family member 22B [:HGNC:21602]
CENPA	-0.644257	1.17E-17	centromere protein A [:HGNC:1851]
MARCH4	-1.445766	1.38E-17	membrane associated ring-CH-type finger 4 [:HGNC:29269]
FANCD2	-0.726152	1.50E-17	FA complementation group D2 [:HGNC:3585]
TNFRSF10D	-0.828533	1.70E-17	TNF receptor superfamily member 10d [:HGNC:11907]
BRCA2	-0.806334	1.77E-17	BRCA2, DNA repair associated [:HGNC:1101]
TELO2	-0.706091	1.80E-17	telomere maintenance 2 [:HGNC:29099]
NANOS1	-0.771818	1.97E-17	nanos C2HC-type zinc finger 1 [:HGNC:23044]
C9orf40	-0.637493	2.53E-17	chromosome 9 open reading frame 40 [:HGNC:23433]
MPP1	-0.596979	2.77E-17	membrane palmitoylated protein 1 [:HGNC:7219]
PCLAF	-0.709559	3.24E-17	PCNA clamp associated factor [:HGNC:28961]

LINC02009	-0.705242	3.29E-17	long intergenic non-protein coding RNA 2009 [:HGNC:52845]
NFATC2	-0.962429	3.31E-17	nuclear factor of activated T cells 2 [:HGNC:7776]
ENPP4	-0.801870	3.33E-17	ectonucleotide pyrophosphatase/phosphodiesterase 4 [:HGNC:3359]
HMGXB4	-0.603582	4.31E-17	HMG-box containing 4 [:HGNC:5003]
FN3KRP	-0.637979	4.42E-17	fructosamine 3 kinase related protein [:HGNC:25700]
METTL17	-0.674323	6.66E-17	methyltransferase like 17 [:HGNC:19280]
DNA2	-0.800854	8.22E-17	DNA replication helicase/nuclease 2 [:HGNC:2939]
LRR1	-0.752469	8.86E-17	leucine rich repeat protein 1 [:HGNC:19742]
TRAF1	-1.041137	9.54E-17	TNF receptor associated factor 1 [:HGNC:12031]
RAB3IP	-1.084789	9.95E-17	RAB3A interacting protein [:HGNC:16508]
PROSER2	-0.691511	1.03E-16	proline and serine rich 2 [:HGNC:23728]
NEDD9	-1.240701	1.18E-16	neural precursor cell expressed, developmentally down-regulated 9 [:HGNC:7733]
LRRC8C	-0.890939	1.20E-16	leucine rich repeat containing 8 VRAC subunit C [:HGNC:25075]
FBXO5	-0.609193	1.26E-16	F-box protein 5 [:HGNC:13584]
PAQR3	-0.637864	1.44E-16	progestin and adipoQ receptor family member 3 [:HGNC:30130]
CHM	-0.676401	1.62E-16	CHM, Rab escort protein 1 [:HGNC:1940]
TICRR	-0.831776	1.66E-16	TOPBP1 interacting checkpoint and replication regulator [:HGNC:28704]
SUV39H2	-0.909487	1.81E-16	suppressor of variegation 3-9 homolog 2 [:HGNC:17287]
RAD18	-0.609258	2.38E-16	RAD18, E3 ubiquitin protein ligase [:HGNC:18278]
E2F2	-1.054698	2.51E-16	E2F transcription factor 2 [:HGNC:3114]
CEP97	-0.593003	3.21E-16	centrosomal protein 97 [:HGNC:26244]
MAFK	-0.666872	3.59E-16	MAF bZIP transcription factor K [:HGNC:6782]
HERC5	-0.749584	3.63E-16	HECT and RLD domain containing E3 ubiquitin protein ligase 5 [:HGNC:24368]
TIMM8A	-1.010189	3.79E-16	translocase of inner mitochondrial membrane 8A [:HGNC:11817]
ZBED6CL	-0.619682	4.11E-16	ZBED6 C-terminal like [:HGNC:21720]
MITF	-0.917808	4.29E-16	melanogenesis associated transcription factor [:HGNC:7105]
LINC00472	-0.807645	5.24E-16	long intergenic non-protein coding RNA 472 [:HGNC:21380]
CCNF	-0.673631	5.69E-16	cyclin F [:HGNC:1591]
B3GNT5	-0.719634	7.02E-16	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 5 [:HGNC:15684]
CCDC93	-0.591178	8.12E-16	coiled-coil domain containing 93 [:HGNC:25611]
LCP1	-1.300348	8.47E-16	lymphocyte cytosolic protein 1 [:HGNC:6528]
KLF5	-0.786660	9.34E-16	Kruppel like factor 5 [:HGNC:6349]
TRMT2B	-0.860159	1.10E-15	tRNA methyltransferase 2 homolog B [:HGNC:25748]
RFTN1	-0.642351	1.15E-15	raftlin, lipid raft linker 1 [:HGNC:30278]
ATAD5	-0.870639	1.34E-15	ATPase family, AAA domain containing 5 [:HGNC:25752]
POU2F2	-1.527423	1.76E-15	POU class 2 homeobox 2 [:HGNC:9213]
HASPIN	-0.796529	2.35E-15	histone H3 associated protein kinase [:HGNC:19682]
SKA1	-0.665053	2.68E-15	spindle and kinetochore associated complex subunit 1 [:HGNC:28109]
SLC7A2	-0.756825	2.83E-15	solute carrier family 7 member 2 [:HGNC:11060]
BTG3	-0.613305	3.54E-15	BTG anti-proliferation factor 3 [:HGNC:1132]
MERTK	-0.872011	4.44E-15	MER proto-oncogene, tyrosine kinase [:HGNC:7027]
WDR44	-0.619891	4.89E-15	WD repeat domain 44 [:HGNC:30512]
PPT2	-1.052565	5.46E-15	palmitoyl-protein thioesterase 2 [:HGNC:9326]
FARP1	-0.597285	6.38E-15	FERM, ARH/RhoGEF and pleckstrin domain protein 1 [:HGNC:3591]
C1orf198	-0.650289	7.88E-15	chromosome 1 open reading frame 198 [:HGNC:25900]

CXorf56	-0.795644	7.98E-15	chromosome X open reading frame 56 [:HGNC:26239]
AFAP1-AS1	-1.054306	8.03E-15	AFAP1 antisense RNA 1 [:HGNC:28141]
TAF5	-0.757374	8.25E-15	TATA-box binding protein associated factor 5 [:HGNC:11539]
SFTA1P	-0.986571	1.20E-14	surfactant associated 1, pseudogene [:HGNC:18383]
CHAF1B	-0.640242	1.22E-14	chromatin assembly factor 1 subunit B [:HGNC:1911]
CHTF18	-0.735461	1.38E-14	chromosome transmission fidelity factor 18 [:HGNC:18435]
UPF3B	-0.762323	1.46E-14	UPF3B, regulator of nonsense mediated mRNA decay [:HGNC:20439]
CXCL8	-1.485130	1.51E-14	C-X-C motif chemokine ligand 8 [:HGNC:6025]
MMS22L	-0.715728	2.00E-14	MMS22 like, DNA repair protein [:HGNC:21475]
LRRC20	-0.597021	2.17E-14	leucine rich repeat containing 20 [:HGNC:23421]
PPP1R3E	-0.946153	2.27E-14	protein phosphatase 1 regulatory subunit 3E [:HGNC:14943]
RMI1	-0.623855	2.65E-14	RecQ mediated genome instability 1 [:HGNC:25764]
RTKN2	-0.699391	2.82E-14	rhotekin 2 [:HGNC:19364]
TPMT	-0.760025	2.87E-14	thiopurine S-methyltransferase [:HGNC:12014]
NMRAL1	-0.666891	3.74E-14	NmrA like redox sensor 1 [:HGNC:24987]
RAD54L	-0.874725	3.80E-14	RAD54 like [:HGNC:9826]
DSCC1	-0.758587	4.06E-14	DNA replication and sister chromatid cohesion 1 [:HGNC:24453]
SYK	-2.968613	4.28E-14	spleen associated tyrosine kinase [:HGNC:11491]
G2E3	-0.724613	4.40E-14	G2/M-phase specific E3 ubiquitin protein ligase [:HGNC:20338]
HPSE	-0.705970	4.49E-14	heparanase [:HGNC:5164]
DNAJC15	-0.635621	4.52E-14	DnaJ heat shock protein family (Hsp40) member C15 [:HGNC:20325]
TEX30	-0.755992	5.89E-14	testis expressed 30 [:HGNC:25188]
STARD8	-0.671470	7.00E-14	StAR related lipid transfer domain containing 8 [:HGNC:19161]
DCLRE1B	-0.756829	9.76E-14	DNA cross-link repair 1B [:HGNC:17641]
CELF2	-0.998672	1.09E-13	CUGBP Elav-like family member 2 [:HGNC:2550]
FAM110A	-0.956616	1.09E-13	family with sequence similarity 110 member A [:HGNC:16188]
EPN2	-0.717485	1.25E-13	epsin 2 [:HGNC:18639]
CCP110	-0.621760	1.28E-13	centriolar coiled-coil protein 110 [:HGNC:24342]
SNRNP48	-0.659836	2.72E-13	small nuclear ribonucleoprotein U11/U12 subunit 48 [:HGNC:21368]
CDKAL1	-0.684897	2.78E-13	CDK5 regulatory subunit associated protein 1 like 1 [:HGNC:21050]
CLIC5	-1.306771	2.86E-13	chloride intracellular channel 5 [:HGNC:13517]
IL6	-1.706871	3.01E-13	interleukin 6 [:HGNC:6018]
FBXO33	-0.717674	3.05E-13	F-box protein 33 [:HGNC:19833]
DDX12P	-1.002629	3.41E-13	DEAD/H-box helicase 12, pseudogene [:HGNC:2737]
GINS4	-0.691237	3.60E-13	GINS complex subunit 4 [:HGNC:28226]
PDE4D	-0.786334	4.49E-13	phosphodiesterase 4D [:HGNC:8783]
MAP3K9	-0.720976	4.89E-13	mitogen-activated protein kinase kinase kinase 9 [:HGNC:6861]
AOX1	-0.697122	5.25E-13	aldehyde oxidase 1 [:HGNC:553]
CENPW	-0.730735	5.59E-13	centromere protein W [:HGNC:21488]
PRTFDC1	-0.956652	5.98E-13	phosphoribosyl transferase domain containing 1 [:HGNC:23333]
MUC5AC	-2.697661	6.41E-13	mucin 5AC, oligomeric mucus/gel-forming [:HGNC:7515]
RALGAPA1	-0.789523	6.53E-13	Ral GTPase activating protein catalytic alpha subunit 1 [:HGNC:17770]
MORC3	-0.627659	8.71E-13	MORC family CW-type zinc finger 3 [:HGNC:23572]
HMGN5	-1.143112	1.28E-12	high mobility group nucleosome binding domain 5 [:HGNC:8013]
MXD3	-0.585778	1.30E-12	MAX dimerization protein 3 [:HGNC:14008]

ROBO1	-1.511855	1.31E-12	roundabout guidance receptor 1 [:HGNC:10249]
RBM41	-0.833462	1.41E-12	RNA binding motif protein 41 [:HGNC:25617]
FAM83A	-1.675917	1.77E-12	family with sequence similarity 83 member A [:HGNC:28210]
TOPORS	-0.581096	2.12E-12	TOP1 binding arginine/serine rich protein [:HGNC:21653]
CHRDL1	-1.429902	2.24E-12	chordin like 1 [:HGNC:29861]
NCF2	-0.985249	2.31E-12	neutrophil cytosolic factor 2 [:HGNC:7661]
CCNT1	-0.607998	3.03E-12	cyclin T1 [:HGNC:1599]
IL32	-0.933015	3.09E-12	interleukin 32 [:HGNC:16830]
NUDT15	-0.586714	3.09E-12	nudix hydrolase 15 [:HGNC:23063]
PARL	-1.028627	3.44E-12	presenilin associated rhomboid like [:HGNC:18253]
ABCB7	-0.695697	4.46E-12	ATP binding cassette subfamily B member 7 [:HGNC:48]
C18orf54	-0.769996	4.98E-12	chromosome 18 open reading frame 54 [:HGNC:13796]
JRK	-0.648952	5.23E-12	Jrk helix-turn-helix protein [:HGNC:6199]
CDK5RAP1	-0.596068	5.36E-12	CDK5 regulatory subunit associated protein 1 [:HGNC:15880]
ORC1	-0.731364	6.26E-12	origin recognition complex subunit 1 [:HGNC:8487]
TBC1D31	-0.655120	6.61E-12	TBC1 domain family member 31 [:HGNC:30888]
APOOL	-0.756814	6.66E-12	apolipoprotein O like [:HGNC:24009]
CSPG5	-1.941479	7.40E-12	chondroitin sulfate proteoglycan 5 [:HGNC:2467]
GRB7	-1.913278	1.10E-11	growth factor receptor bound protein 7 [:HGNC:4567]
ZNF542P	-2.527179	1.15E-11	zinc finger protein 542, pseudogene [:HGNC:25393]
VWA5A	-1.438355	1.58E-11	von Willebrand factor A domain containing 5A [:HGNC:6658]
TP73	-1.351065	1.58E-11	tumor protein p73 [:HGNC:12003]
ZNF714	-0.650587	1.72E-11	zinc finger protein 714 [:HGNC:27124]
EDA	-2.080416	2.15E-11	ectodysplasin A [:HGNC:3157]
BRI3BP	-0.633404	2.49E-11	BRI3 binding protein [:HGNC:14251]
DDIAS	-0.692945	2.52E-11	DNA damage induced apoptosis suppressor [:HGNC:26351]
MAD2L1BP	-0.620900	2.87E-11	MAD2L1 binding protein [:HGNC:21059]
GPSM3	-0.728903	3.38E-11	G protein signaling modulator 3 [:HGNC:13945]
SYTL2	-0.875547	3.52E-11	synaptotagmin like 2 [:HGNC:15585]
RBM15	-0.658660	4.31E-11	RNA binding motif protein 15 [:HGNC:14959]
SLIT2	-0.833140	4.81E-11	slit guidance ligand 2 [:HGNC:11086]
PIN4	-0.623957	5.69E-11	peptidylprolyl cis/trans isomerase, NIMA-interacting 4 [:HGNC:8992]
PNMA8A	-0.653907	5.97E-11	PNMA family member 8A [:HGNC:25578]
NNT-AS1	-0.616671	6.54E-11	NNT antisense RNA 1 [:HGNC:49005]
CCDC190	-0.883008	6.69E-11	coiled-coil domain containing 190 [:HGNC:28736]
MUC4	-1.599505	6.80E-11	mucin 4, cell surface associated [:HGNC:7514]
AFF3	-1.321203	6.96E-11	AF4/FMR2 family member 3 [:HGNC:6473]
FHL1	-0.731821	7.04E-11	four and a half LIM domains 1 [:HGNC:3702]
ATP8B2	-0.788899	7.09E-11	ATPase phospholipid transporting 8B2 [:HGNC:13534]
XK	-1.216953	7.48E-11	X-linked Kx blood group [:HGNC:12811]
DHRS1	-0.582539	7.52E-11	dehydrogenase/reductase 1 [:HGNC:16445]
EPB41L5	-0.733795	8.01E-11	erythrocyte membrane protein band 4.1 like 5 [:HGNC:19819]
CCNE2	-0.798407	9.92E-11	cyclin E2 [:HGNC:1590]
SAMD12	-1.068660	1.05E-10	sterile alpha motif domain containing 12 [:HGNC:31750]
HERPUD2	-0.601287	1.07E-10	HERPUD family member 2 [:HGNC:21915]

EEF1E1	-0.645392	1.07E-10	eukaryotic translation elongation factor 1 epsilon 1 [:HGNC:3212]
ZNF850	-0.835459	1.20E-10	zinc finger protein 850 [:HGNC:27994]
FAM133A	-2.259495	1.62E-10	family with sequence similarity 133 member A [:HGNC:26748]
PTK2B	-0.789797	2.33E-10	protein tyrosine kinase 2 beta [:HGNC:9612]
RAD51C	-0.852463	2.34E-10	RAD51 paralog C [:HGNC:9820]
SPDEF	-1.387619	2.42E-10	SAM pointed domain containing ETS transcription factor [:HGNC:17257]
SYNE3	-0.597415	2.63E-10	spectrin repeat containing nuclear envelope family member 3 [:HGNC:19861]
NAA10	-0.633288	3.06E-10	N(alpha)-acetyltransferase 10, NatA catalytic subunit [:HGNC:18704]
L2HGDH	-0.706130	3.07E-10	L-2-hydroxyglutarate dehydrogenase [:HGNC:20499]
RAB30	-0.724223	3.09E-10	RAB30, member RAS oncogene family [:HGNC:9770]
F8A1	-0.718017	3.27E-10	coagulation factor VIII associated 1 [:HGNC:3547]
SAYSD1	-0.581386	3.42E-10	SAYSVFN motif domain containing 1 [:HGNC:21025]
IGDCC4	-0.749752	3.45E-10	immunoglobulin superfamily DCC subclass member 4 [:HGNC:13770]
BRMS1L	-0.748189	3.51E-10	BRMS1 like transcriptional repressor [:HGNC:20512]
NEGR1	-0.842282	3.78E-10	neuronal growth regulator 1 [:HGNC:17302]
MYO5C	-0.857496	3.93E-10	myosin VC [:HGNC:7604]
ZNF22	-0.591639	3.99E-10	zinc finger protein 22 [:HGNC:13012]
EDN1	-0.856826	4.19E-10	endothelin 1 [:HGNC:3176]
KHDRBS3	-0.871610	4.63E-10	KH RNA binding domain containing, signal transduction associated 3 [:HGNC:18117]
CXCR4	-1.709426	4.79E-10	C-X-C motif chemokine receptor 4 [:HGNC:2561]
RNF169	-0.599039	4.79E-10	ring finger protein 169 [:HGNC:26961]
SAA2	-2.442969	4.93E-10	serum amyloid A2 [:HGNC:10514]
HDAC8	-0.991941	5.53E-10	histone deacetylase 8 [:HGNC:13315]
FBXL7	-3.939028	6.08E-10	F-box and leucine rich repeat protein 7 [:HGNC:13604]
GCH1	-0.938366	6.50E-10	GTP cyclohydrolase 1 [:HGNC:4193]
FOXA2	-1.034452	7.42E-10	forkhead box A2 [:HGNC:5022]
NRGN	-0.580474	7.45E-10	neurogranin [:HGNC:8000]
NEU3	-0.626933	8.26E-10	neuraminidase 3 [:HGNC:7760]
AC006452.1	-2.719529	8.65E-10	novel transcript
FBN1	-0.597676	1.05E-09	fibrillin 1 [:HGNC:3603]
NUP62CL	-1.632327	1.06E-09	nucleoporin 62 C-terminal like [:HGNC:25960]
CGN	-1.790164	1.07E-09	cingulin [:HGNC:17429]
MND1	-0.630301	1.11E-09	meiotic nuclear divisions 1 [:HGNC:24839]
CEP152	-0.688541	1.20E-09	centrosomal protein 152 [:HGNC:29298]
C1orf112	-0.619592	1.27E-09	chromosome 1 open reading frame 112 [:HGNC:25565]
PRIM1	-0.875399	1.30E-09	DNA primase subunit 1 [:HGNC:9369]
MYSM1	-0.719129	1.31E-09	Myb like, SWIRM and MPN domains 1 [:HGNC:29401]
AL161891.1	-0.880703	1.35E-09	novel transcript, sense intronic to RFC3
RNF144B	-1.099262	1.90E-09	ring finger protein 144B [:HGNC:21578]
USP37	-0.817686	2.09E-09	ubiquitin specific peptidase 37 [:HGNC:20063]
LINC02001	-0.942693	2.25E-09	long intergenic non-protein coding RNA 2001 [:HGNC:52836]
ZNF275	-0.688495	2.32E-09	zinc finger protein 275 [:HGNC:13069]
GALNT6	-0.625252	2.36E-09	polypeptide N-acetylgalactosaminyltransferase 6 [:HGNC:4128]
UBE2D4	-0.832848	2.37E-09	ubiquitin conjugating enzyme E2 D4 (putative) [:HGNC:21647]
DCLRE1A	-0.683425	2.91E-09	DNA cross-link repair 1A [:HGNC:17660]

ARHGEF26	-0.599459	3.24E-09	Rho guanine nucleotide exchange factor 26 [:HGNC:24490]
FGFR1OP	-0.737570	3.32E-09	FGFR1 oncogene partner [:HGNC:17012]
FBXO48	-1.028843	4.20E-09	F-box protein 48 [:HGNC:33857]
DPF3	-1.069329	4.23E-09	double PHD fingers 3 [:HGNC:17427]
PIGA	-0.647931	4.24E-09	phosphatidylinositol glycan anchor biosynthesis class A [:HGNC:8957]
SNHG18	-0.830866	5.05E-09	small nucleolar RNA host gene 18 [:HGNC:49007]
NES	-0.793788	5.75E-09	nestin [:HGNC:7756]
GGT5	-0.751550	5.91E-09	gamma-glutamyltransferase 5 [:HGNC:4260]
ULBP2	-0.804943	6.91E-09	UL16 binding protein 2 [:HGNC:14894]
FANCM	-0.784044	6.91E-09	FA complementation group M [:HGNC:23168]
LINC00973	-0.699591	8.72E-09	long intergenic non-protein coding RNA 973 [:HGNC:48868]
PEAR1	-1.121793	8.87E-09	platelet endothelial aggregation receptor 1 [:HGNC:33631]
ANO5	-2.477668	8.90E-09	anoctamin 5 [:HGNC:27337]
NFKBIZ	-0.723621	9.75E-09	NFKB inhibitor zeta [:HGNC:29805]
SGO1	-0.584148	1.02E-08	shugoshin 1 [:HGNC:25088]
SLCO4A1-AS1	-1.403603	1.11E-08	SLCO4A1 antisense RNA 1 [:HGNC:40537]
CXorf57	-1.207983	1.31E-08	chromosome X open reading frame 57 [:HGNC:25486]
PGM5	-3.568989	1.44E-08	phosphoglucomutase 5 [:HGNC:8908]
ZSCAN9	-0.711260	1.49E-08	zinc finger and SCAN domain containing 9 [:HGNC:12984]
SHQ1	-0.671132	1.53E-08	SHQ1, H/ACA ribonucleoprotein assembly factor [:HGNC:25543]
CCDC15	-0.918010	1.53E-08	coiled-coil domain containing 15 [:HGNC:25798]
MIR503HG	-1.222844	1.58E-08	MIR503 host gene [:HGNC:28258]
KHDC1	-0.748248	1.62E-08	KH domain containing 1 [:HGNC:21366]
PIF1	-0.704173	1.92E-08	PIF1 5'-to-3' DNA helicase [:HGNC:26220]
GCAT	-0.611049	2.74E-08	glycine C-acetyltransferase [:HGNC:4188]
NGDN	-0.607851	2.86E-08	neuroguidin [:HGNC:20271]
MYO1D	-1.047758	2.94E-08	myosin ID [:HGNC:7598]
ANKRD18EP	-1.009585	3.25E-08	ankyrin repeat domain 18E, pseudogene [:HGNC:43609]
CERS6	-0.640774	3.34E-08	ceramide synthase 6 [:HGNC:23826]
HOMER1	-0.770591	3.68E-08	homer scaffold protein 1 [:HGNC:17512]
TIPIN	-0.703529	4.62E-08	TIMELESS interacting protein [:HGNC:30750]
POLH-AS1	-1.808847	4.67E-08	POLH antisense RNA 1 [:HGNC:40459]
PDE1C	-3.392569	4.71E-08	phosphodiesterase 1C [:HGNC:8776]
NKRF	-0.604312	4.82E-08	NFKB repressing factor [:HGNC:19374]
ZNF804A	-2.062582	5.18E-08	zinc finger protein 804A [:HGNC:21711]
TEDC2	-0.685150	5.84E-08	tubulin epsilon and delta complex 2 [:HGNC:25849]
RNF125	-1.247444	5.89E-08	ring finger protein 125 [:HGNC:21150]
LONRF3	-0.668956	5.91E-08	LON peptidase N-terminal domain and ring finger 3 [:HGNC:21152]
MMD	-0.625948	6.71E-08	monocyte to macrophage differentiation associated [:HGNC:7153]
ZNF641	-0.611554	7.09E-08	zinc finger protein 641 [:HGNC:31834]
MFSD2A	-0.769983	7.34E-08	major facilitator superfamily domain containing 2A [:HGNC:25897]
IL24	-1.620372	7.57E-08	interleukin 24 [:HGNC:11346]
TXNDC16	-0.794241	7.60E-08	thioredoxin domain containing 16 [:HGNC:19965]
RAB23	-0.592891	7.75E-08	RAB23, member RAS oncogene family [:HGNC:14263]
C2orf49	-0.652782	8.61E-08	chromosome 2 open reading frame 49 [:HGNC:28772]

RAB2B	-0.737161	9.07E-08	RAB2B, member RAS oncogene family [:HGNC:20246]
BLM	-0.700977	1.09E-07	Bloom syndrome RecQ like helicase [:HGNC:1058]
RIMS2	-0.852968	1.19E-07	regulating synaptic membrane exocytosis 2 [:HGNC:17283]
EXPH5	-0.744385	1.22E-07	exophilin 5 [:HGNC:30578]
ING2	-0.588015	1.28E-07	inhibitor of growth family member 2 [:HGNC:6063]
DIAPH2	-0.603171	1.33E-07	diaphanous related formin 2 [:HGNC:2877]
AC244153.1	-1.691361	1.44E-07	uncharacterized LOC101929494 [Source:NCBI gene;Acc:101929494]
OXTR	-0.794774	1.48E-07	oxytocin receptor [:HGNC:8529]
MBIP	-0.609882	1.51E-07	MAP3K12 binding inhibitory protein 1 [:HGNC:20427]
IGFL2-AS1	-2.336233	1.58E-07	IGFL2 antisense RNA 1 [:HGNC:52559]
MB21D2	-1.056286	1.76E-07	Mab-21 domain containing 2 [:HGNC:30438]
MTM1	-0.716646	1.79E-07	myotubularin 1 [:HGNC:7448]
PELI2	-1.492077	1.89E-07	pellino E3 ubiquitin protein ligase family member 2 [:HGNC:8828]
CXCL1	-1.563691	1.92E-07	C-X-C motif chemokine ligand 1 [:HGNC:4602]
BTN2A2	-0.636969	1.93E-07	butyrophilin subfamily 2 member A2 [:HGNC:1137]
ZNF718	-0.651804	2.00E-07	zinc finger protein 718 [:HGNC:26889]
BLOC1S5	-0.624089	2.16E-07	biogenesis of lysosomal organelles complex 1 subunit 5 [:HGNC:18561]
PDE9A	-0.884446	2.33E-07	phosphodiesterase 9A [:HGNC:8795]
CHST7	-0.610162	2.47E-07	carbohydrate sulfotransferase 7 [:HGNC:13817]
DPF1	-1.089131	2.54E-07	double PHD fingers 1 [:HGNC:20225]
MFAP2	-1.262737	2.56E-07	microfibril associated protein 2 [:HGNC:7033]
PARM1	-1.542906	3.17E-07	prostate androgen-regulated mucin-like protein 1 [:HGNC:24536]
RALGAPA2	-0.649874	3.44E-07	Ral GTPase activating protein catalytic alpha subunit 2 [:HGNC:16207]
TMEM170B	-0.900422	3.62E-07	transmembrane protein 170B [:HGNC:34244]
HES4	-0.893328	3.85E-07	hes family bHLH transcription factor 4 [:HGNC:24149]
TSPAN8	-2.529325	4.50E-07	tetraspanin 8 [:HGNC:11855]
ZNF280C	-0.825008	4.64E-07	zinc finger protein 280C [:HGNC:25955]
IPO5P1	-1.060779	5.52E-07	importin 5 pseudogene 1 [:HGNC:49687]
DGKH	-0.656649	5.53E-07	diacylglycerol kinase eta [:HGNC:2854]
CNEP1R1	-0.658413	6.04E-07	CTD nuclear envelope phosphatase 1 regulatory subunit 1 [:HGNC:26759]
GATA2-AS1	-0.982625	6.89E-07	GATA2 antisense RNA 1 [:HGNC:51108]
CXCL2	-0.907486	7.05E-07	C-X-C motif chemokine ligand 2 [:HGNC:4603]
NXT2	-0.654364	7.09E-07	nuclear transport factor 2 like export factor 2 [:HGNC:18151]
RAPH1	-0.590719	7.27E-07	Ras association (RalGDS/AF-6) and pleckstrin homology domains 1 [:HGNC:14436]
USP31	-0.684297	8.25E-07	ubiquitin specific peptidase 31 [:HGNC:20060]
FAM162A	-0.647801	1.01E-06	family with sequence similarity 162 member A [:HGNC:17865]
CEP128	-0.644698	1.07E-06	centrosomal protein 128 [:HGNC:20359]
DLK2	-0.885923	1.11E-06	delta like non-canonical Notch ligand 2 [:HGNC:21113]
DAPK1	-3.662839	1.17E-06	death associated protein kinase 1 [:HGNC:2674]
FRS3	-0.818456	1.21E-06	fibroblast growth factor receptor substrate 3 [:HGNC:16970]
GPR68	-0.582975	1.25E-06	G protein-coupled receptor 68 [:HGNC:4519]
EAF1	-0.593990	1.32E-06	ELL associated factor 1 [:HGNC:20907]
KCNK1	-1.604464	1.77E-06	potassium two pore domain channel subfamily K member 1 [:HGNC:6272]
PHC1	-0.843386	1.81E-06	polyhomeotic homolog 1 [:HGNC:3182]
ATP10A	-1.249414	1.81E-06	ATPase phospholipid transporting 10A (putative) [:HGNC:13542]

ZNF620	-1.185998	1.95E-06	zinc finger protein 620 [:HGNC:28742]
AC009533.1	-0.888723	2.00E-06	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide like pseudogene
INSYN2B	-1.140006	2.09E-06	inhibitory synaptic factor family member 2B [:HGNC:37271]
RFXAP	-1.150365	2.14E-06	regulatory factor X associated protein [:HGNC:9988]
PDE6D	-0.599456	2.27E-06	phosphodiesterase 6D [:HGNC:8788]
WNK3	-0.747445	2.35E-06	WNK lysine deficient protein kinase 3 [:HGNC:14543]
ADAL	-0.901961	2.41E-06	adenosine deaminase like [:HGNC:31853]
SLC30A4	-0.734452	2.41E-06	solute carrier family 30 member 4 [:HGNC:11015]
PSMC3IP	-0.673543	2.51E-06	PSMC3 interacting protein [:HGNC:17928]
CCDC138	-0.673905	2.59E-06	coiled-coil domain containing 138 [:HGNC:26531]
UPRT	-0.690195	2.64E-06	uracil phosphoribosyltransferase homolog [:HGNC:28334]
GPR3	-0.714750	2.75E-06	G protein-coupled receptor 3 [:HGNC:4484]
N4BP2	-0.678837	2.78E-06	NEDD4 binding protein 2 [:HGNC:29851]
PKDREJ	-2.589074	2.90E-06	polycystin family receptor for egg jelly [:HGNC:9015]
FBXO43	-0.847545	3.03E-06	F-box protein 43 [:HGNC:28521]
AC002480.1	-2.218766	3.09E-06	uncharacterized LOC100506178 [Source:NCBI gene;Acc:100506178]
BX470102.1	-1.095985	3.27E-06	novel transcript
DPY19L2P2	-0.704758	3.71E-06	DPY19L2 pseudogene 2 [:HGNC:21764]
MGAT4A	-0.788763	3.72E-06	alpha-1,3-mannosyl-glycoprotein 4-beta-N-acetylglucosaminyltransferase A [:HGNC:7047]
SLC48A1	-0.623426	3.99E-06	solute carrier family 48 member 1 [:HGNC:26035]
F13A1	-1.331254	3.99E-06	coagulation factor XIII A chain [:HGNC:3531]
SUPT3H	-0.756606	5.20E-06	SPT3 homolog, SAGA and STAGA complex component [:HGNC:11466]
STK19	-0.803487	5.24E-06	serine/threonine kinase 19 [:HGNC:11398]
PRRG4	-0.773830	5.61E-06	proline rich and Gla domain 4 [:HGNC:30799]
SLC16A12	-1.891733	5.92E-06	solute carrier family 16 member 12 [:HGNC:23094]
EDNRB	-0.637103	5.94E-06	endothelin receptor type B [:HGNC:3180]
NEMP2	-0.633590	6.23E-06	nuclear envelope integral membrane protein 2 [:HGNC:33700]
CT83	-0.880699	6.32E-06	cancer/testis antigen 83 [:HGNC:33494]
TRAC	-2.239596	6.52E-06	T cell receptor alpha constant [:HGNC:12029]
ACOXL	-2.720463	6.74E-06	acyl-CoA oxidase like [:HGNC:25621]
NLGN1	-0.828650	7.10E-06	neuroligin 1 [:HGNC:14291]
TMEM182	-1.152244	7.15E-06	transmembrane protein 182 [:HGNC:26391]
CADM1	-1.631484	7.85E-06	cell adhesion molecule 1 [:HGNC:5951]
PDCD1LG2	-1.069743	8.49E-06	programmed cell death 1 ligand 2 [:HGNC:18731]
MOSMO	-0.654654	8.55E-06	modulator of smoothened [:HGNC:27087]
LINC01293	-1.169638	8.74E-06	long intergenic non-protein coding RNA 1293 [:HGNC:50362]
INAFM1	-0.730795	9.14E-06	InaF motif containing 1 [:HGNC:27406]
KYNU	-2.139560	9.73E-06	kynureninase [:HGNC:6469]
ATG4A	-0.787783	9.80E-06	autophagy related 4A cysteine peptidase [:HGNC:16489]
RNF113A	-0.660247	9.89E-06	ring finger protein 113A [:HGNC:12974]
KCNH5	-3.390873	1.05E-05	potassium voltage-gated channel subfamily H member 5 [:HGNC:6254]
PLCXD2	-1.482592	1.07E-05	phosphatidylinositol specific phospholipase C X domain containing 2 [:HGNC:26462]
MISP	-1.326581	1.08E-05	mitotic spindle positioning [:HGNC:27000]
ATF7IP2	-4.628458	1.17E-05	activating transcription factor 7 interacting protein 2 [:HGNC:20397]
LRRCC1	-0.589719	1.20E-05	leucine rich repeat and coiled-coil centrosomal protein 1 [:HGNC:29373]

RBPMS2	-1.864413	1.23E-05	RNA binding protein, mRNA processing factor 2 [:HGNC:19098]
EXOG	-0.659389	1.24E-05	exo/endonuclease G [:HGNC:3347]
TMEM98	-2.166434	1.25E-05	transmembrane protein 98 [:HGNC:24529]
STXBP6	-2.161382	1.33E-05	syntaxin binding protein 6 [:HGNC:19666]
FAM161A	-0.588484	1.39E-05	FAM161A, centrosomal protein [:HGNC:25808]
THUMPD3-AS1	-0.633498	1.43E-05	THUMPD3 antisense RNA 1 [:HGNC:44478]
ARID5A	-0.777359	1.43E-05	AT-rich interaction domain 5A [:HGNC:17361]
PAK6	-1.794844	1.49E-05	p21 (RAC1) activated kinase 6 [:HGNC:16061]
LTB	-2.131019	1.50E-05	lymphotoxin beta [:HGNC:6711]
FANCB	-0.783873	1.56E-05	FA complementation group B [:HGNC:3583]
PTGS2	-3.019994	1.67E-05	prostaglandin-endoperoxide synthase 2 [:HGNC:9605]
TNFRSF9	-0.821130	1.76E-05	TNF receptor superfamily member 9 [:HGNC:11924]
C3orf52	-0.888663	1.93E-05	chromosome 3 open reading frame 52 [:HGNC:26255]
POLR3G	-0.598133	2.07E-05	RNA polymerase III subunit G [:HGNC:30075]
APLN	-0.594818	2.87E-05	apelin [:HGNC:16665]
PRSS3	-1.165085	2.98E-05	serine protease 3 [:HGNC:9486]
S100A5	-0.777181	3.17E-05	S100 calcium binding protein A5 [:HGNC:10495]
ERFE	-0.618229	3.34E-05	erythroferrone [:HGNC:26727]
AGR2	-0.765821	3.56E-05	anterior gradient 2, protein disulphide isomerase family member [:HGNC:328]
ARG2	-0.756148	3.60E-05	arginase 2 [:HGNC:664]
STYK1	-0.830573	3.68E-05	serine/threonine/tyrosine kinase 1 [:HGNC:18889]
RPS6KA5	-0.712176	4.17E-05	ribosomal protein S6 kinase A5 [:HGNC:10434]
FTX	-0.695892	4.18E-05	FTX transcript, XIST regulator [:HGNC:37190]
PCDH1	-0.713119	4.47E-05	protocadherin 1 [:HGNC:8655]
RASD1	-0.702486	4.81E-05	ras related dexamethasone induced 1 [:HGNC:15828]
ZNF681	-0.758855	4.95E-05	zinc finger protein 681 [:HGNC:26457]
GPR137C	-0.917407	5.25E-05	G protein-coupled receptor 137C [:HGNC:25445]
RELN	-1.922791	5.30E-05	reelin [:HGNC:9957]
CSF2	-1.139691	5.35E-05	colony stimulating factor 2 [:HGNC:2434]
KCNS3	-0.710392	5.69E-05	potassium voltage-gated channel modifier subfamily S member 3 [:HGNC:6302]
FMO3	-2.505823	5.88E-05	flavin containing monooxygenase 3 [:HGNC:3771]
PLEKHH2	-0.656720	6.01E-05	pleckstrin homology, MyTH4 and FERM domain containing H2 [:HGNC:30506]
FBLN7	-0.657570	6.18E-05	fibulin 7 [:HGNC:26740]
NMNAT2	-0.610854	6.24E-05	nicotinamide nucleotide adenylyltransferase 2 [:HGNC:16789]
CORO7	-0.817107	6.25E-05	coronin 7 [:HGNC:26161]
PASK	-0.618529	6.99E-05	PAS domain containing serine/threonine kinase [:HGNC:17270]
TFAP2C	-0.599603	7.79E-05	transcription factor AP-2 gamma [:HGNC:11744]
EFCAB2	-1.035670	7.80E-05	EF-hand calcium binding domain 2 [:HGNC:28166]
CDKL5	-1.085607	8.20E-05	cyclin dependent kinase like 5 [:HGNC:11411]
NUDT3	-0.589331	9.38E-05	nudix hydrolase 3 [:HGNC:8050]
TYMSOS	-1.033283	9.86E-05	TYMS opposite strand [:HGNC:29553]
LINC02407	-1.763588	0.000104	long intergenic non-protein coding RNA 2407 [:HGNC:53336]
RPS17	-0.799000	0.000105	ribosomal protein S17 [:HGNC:10397]
GINS3	-0.612044	0.000107	GINS complex subunit 3 [:HGNC:25851]
SEL1L3	-0.911824	0.000113	SEL1L family member 3 [:HGNC:29108]

ASRGL1	-2.374296	0.000114	asparaginase like 1 [:HGNC:16448]
CHEK2	-0.582331	0.000145	checkpoint kinase 2 [:HGNC:16627]
AC139769.1	-1.850162	0.000157	zinc finger protein pseudogene
AC138393.1	-0.762412	0.000161	septin 7 (SEPT7) pseudogene
JPH1	-0.627125	0.000162	junctophilin 1 [:HGNC:14201]
CENPP	-0.645571	0.000166	centromere protein P [:HGNC:32933]
FCMR	-0.632777	0.000212	Fc fragment of IgM receptor [:HGNC:14315]
FEZF1	-1.082448	0.000212	FEZ family zinc finger 1 [:HGNC:22788]
BMP8B	-0.935016	0.00022	bone morphogenetic protein 8b [:HGNC:1075]
RPL36A	-0.996374	0.000221	ribosomal protein L36a [:HGNC:10359]
INHBB	-1.235768	0.000227	inhibin subunit beta B [:HGNC:6067]
SDR16C5	-1.508711	0.000229	short chain dehydrogenase/reductase family 16C member 5 [:HGNC:30311]
GPR17	-1.221213	0.000231	G protein-coupled receptor 17 [:HGNC:4471]
FEZF1-AS1	-1.091212	0.000246	FEZF1 antisense RNA 1 [:HGNC:41001]
PSG1	-1.629525	0.000262	pregnancy specific beta-1-glycoprotein 1 [:HGNC:9514]
TMPPE	-0.854236	0.000264	transmembrane protein with metallophosphoesterase domain [:HGNC:33865]
ADAMTSL4	-0.604531	0.000265	ADAMTS like 4 [:HGNC:19706]
SP4	-0.669158	0.000266	Sp4 transcription factor [:HGNC:11209]
TMEM156	-1.003450	0.000282	transmembrane protein 156 [:HGNC:26260]
AKAP6	-0.801795	0.000282	A-kinase anchoring protein 6 [:HGNC:376]
PDE3B	-0.870857	0.000285	phosphodiesterase 3B [:HGNC:8779]
C9orf72	-0.659182	0.000298	chromosome 9 open reading frame 72 [:HGNC:28337]
CXorf40A	-0.619077	0.000299	chromosome X open reading frame 40A [:HGNC:28089]
LINC02535	-0.985650	0.000308	long intergenic non-protein coding RNA 2535 [:HGNC:53569]
SAMD5	-1.408027	0.000322	sterile alpha motif domain containing 5 [:HGNC:21180]
SLCO1B3	-0.940124	0.000327	solute carrier organic anion transporter family member 1B3 [:HGNC:10961]
FLJ22447	-1.202256	0.000335	uncharacterized LOC400221 [Source:NCBI gene;Acc:400221]
RF00003	-3.241397	0.000347	
OR4F4	-5.705536	0.00035	olfactory receptor family 4 subfamily F member 4 [:HGNC:8301]
PLEKHS1	-0.805697	0.000377	pleckstrin homology domain containing S1 [:HGNC:26285]
RNF152	-0.758149	0.000387	ring finger protein 152 [:HGNC:26811]
PSTK	-0.631573	0.000387	phosphoseryl-tRNA kinase [:HGNC:28578]
LAT2	-0.899098	0.000402	linker for activation of T cells family member 2 [:HGNC:12749]
PDXP	-0.738543	0.000403	pyridoxal phosphatase [:HGNC:30259]
KCNQ4	-1.844240	0.000448	potassium voltage-gated channel subfamily Q member 4 [:HGNC:6298]
ADGRV1	-0.884678	0.000453	adhesion G protein-coupled receptor V1 [:HGNC:17416]
SP140	-0.914949	0.000458	SP140 nuclear body protein [:HGNC:17133]
GRHL1	-2.233998	0.00049	grainyhead like transcription factor 1 [:HGNC:17923]
FIRRE	-2.665062	0.000504	firre intergenic repeating RNA element [:HGNC:49627]
AC093724.1	-0.663039	0.000511	translocase of outer mitochondrial membrane 40 (TOMM40) pseudogene
AC073130.1	-0.849016	0.000514	uncharacterized LOC102724434 [Source:NCBI gene;Acc:102724434]
MID2	-0.708845	0.000537	midline 2 [:HGNC:7096]
TAGLN	-0.742035	0.000539	transgelin [:HGNC:11553]
RNU6-1	-4.575536	0.000573	RNA, U6 small nuclear 1 [:HGNC:10227]
MYO7B	-2.010133	0.000585	myosin VIIB [:HGNC:7607]

EMC3-AS1	-0.725968	0.000597	EMC3 antisense RNA 1 [:HGNC:49223]
AC068831.7	-0.768825	0.000621	novel transcript
RAB37	-1.150526	0.000632	RAB37, member RAS oncogene family [:HGNC:30268]
SNORA73B	-1.053554	0.00071	small nucleolar RNA, H/ACA box 73B [:HGNC:10116]
IGFBPL1	-0.801109	0.000728	insulin like growth factor binding protein like 1 [:HGNC:20081]
RAPGEF3	-0.965412	0.000729	Rap guanine nucleotide exchange factor 3 [:HGNC:16629]
ZNF521	-1.255557	0.000786	zinc finger protein 521 [:HGNC:24605]
ZNF699	-0.765299	0.000795	zinc finger protein 699 [:HGNC:24750]
PXMP2	-0.685676	0.000815	peroxisomal membrane protein 2 [:HGNC:9716]
AC020916.1	-0.896831	0.000851	novel transcript
ACP7	-1.844891	0.000858	acid phosphatase 7, tartrate resistant (putative) [:HGNC:33781]
TRIML2	-1.341888	0.000906	tripartite motif family like 2 [:HGNC:26378]
U91328.1	-0.807052	0.000927	novel transcript
ZNF586	-0.707810	0.001041	zinc finger protein 586 [:HGNC:25949]
TRIM36	-0.586919	0.001093	tripartite motif containing 36 [:HGNC:16280]
SH3BP1	-1.199342	0.001139	SH3 domain binding protein 1 [:HGNC:10824]
SPRYD4	-0.644563	0.00115	SPRY domain containing 4 [:HGNC:27468]
AL365273.1	-1.086112	0.001193	novel transcript, sense intronic to ENTPD1
LINC01638	-1.515648	0.001203	long intergenic non-protein coding RNA 1638 [:HGNC:52425]
RNF180	-0.875373	0.001232	ring finger protein 180 [:HGNC:27752]
ALPP	-0.674878	0.001235	alkaline phosphatase, placental [:HGNC:439]
CYP24A1	-2.482364	0.001359	cytochrome P450 family 24 subfamily A member 1 [:HGNC:2602]
TC2N	-1.907815	0.001362	tandem C2 domains, nuclear [:HGNC:19859]
ZSCAN16	-0.695939	0.001366	zinc finger and SCAN domain containing 16 [:HGNC:20813]
AC019205.1	-1.093687	0.001383	novel transcript, antisense to KHDC1
RPS6KL1	-0.882039	0.001484	ribosomal protein S6 kinase like 1 [:HGNC:20222]
ICAM5	-0.798720	0.001521	intercellular adhesion molecule 5 [:HGNC:5348]
TRABD2A	-0.841768	0.001567	TraB domain containing 2A [:HGNC:27013]
KRT87P	-1.574379	0.001575	keratin 87 pseudogene [:HGNC:30198]
TJP2	-0.700527	0.001608	tight junction protein 2 [:HGNC:11828]
RASGRP3	-0.606536	0.001612	RAS guanyl releasing protein 3 [:HGNC:14545]
AC005831.1	-0.949025	0.00162	TEC
EML1	-0.768100	0.001638	echinoderm microtubule associated protein like 1 [:HGNC:3330]
FAM122C	-0.830459	0.001736	family with sequence similarity 122C [:HGNC:25202]
GPS2	-0.664931	0.00184	G protein pathway suppressor 2 [:HGNC:4550]
KCNQ3	-1.278438	0.001853	potassium voltage-gated channel subfamily Q member 3 [:HGNC:6297]
CYP26B1	-0.667319	0.002017	cytochrome P450 family 26 subfamily B member 1 [:HGNC:20581]
NOSTRIN	-0.836244	0.002071	nitric oxide synthase trafficking [:HGNC:20203]
DTNA	-1.064980	0.00211	dystrobrevin alpha [:HGNC:3057]
EFR3B	-0.689419	0.002136	EFR3 homolog B [:HGNC:29155]
CACNA2D4	-1.039098	0.002243	calcium voltage-gated channel auxiliary subunit alpha2delta 4 [:HGNC:20202]
AC112907.3	-0.894697	0.002312	novel transcript, antisense to eukaryotic translation initiation factor 4A, isoform 2 EIF4A2
RNFT2	-0.666866	0.00247	ring finger protein, transmembrane 2 [:HGNC:25905]
TRBC2	-3.142027	0.002525	T cell receptor beta constant 2 [:HGNC:12157]
IPO5P1	-1.003279	0.002639	importin 5 pseudogene 1 [Source:NCBI gene;Acc:100132815]

C2orf48	-1.537090	0.002723	chromosome 2 open reading frame 48 [:HGNC:26322]
KLC3	-0.709882	0.002744	kinesin light chain 3 [:HGNC:20717]
AC005180.2	-1.840149	0.002789	novel transcript
BTN2A3P	-0.657346	0.002836	butyrophilin subfamily 2 member A3, pseudogene [:HGNC:13229]
ENTPD1-AS1	-0.659243	0.002889	ENTPD1 antisense RNA 1 [:HGNC:45203]
KIF7	-0.711451	0.003044	kinesin family member 7 [:HGNC:30497]
ZDHHC23	-0.604870	0.003118	zinc finger DHHC-type containing 23 [:HGNC:28654]
F2RL3	-1.382537	0.003148	F2R like thrombin or trypsin receptor 3 [:HGNC:3540]
SERF1B	-1.117741	0.003429	small EDRK-rich factor 1B [:HGNC:10756]
SLC25A53	-0.859513	0.003451	solute carrier family 25 member 53 [:HGNC:31894]
PPM1L	-1.032641	0.003592	protein phosphatase, Mg2+/Mn2+ dependent 1L [:HGNC:16381]
DSC2	-1.301559	0.003638	desmocollin 2 [:HGNC:3036]
MFAP3L	-0.631802	0.003942	microfibril associated protein 3 like [:HGNC:29083]
NHSL2	-1.131749	0.00401	NHS like 2 [:HGNC:33737]
PSMA6	-0.662949	0.004066	proteasome subunit alpha 6 [:HGNC:9535]
AL365184.1	-1.886853	0.004225	novel transcript
SLC25A14	-0.599684	0.004255	solute carrier family 25 member 14 [:HGNC:10984]
ZNF582	-1.891676	0.00448	zinc finger protein 582 [:HGNC:26421]
FAM169A	-0.623034	0.004727	family with sequence similarity 169 member A [:HGNC:29138]
FLJ31356	-2.639135	0.004847	uncharacterized protein FLJ31356 [Source:NCBI gene;Acc:403150]
HIST1H2BJ	-1.039586	0.004915	histone cluster 1 H2B family member j [:HGNC:4761]
MBOAT1	-0.721532	0.005041	membrane bound O-acyltransferase domain containing 1 [:HGNC:21579]
AC083843.2	-0.837771	0.005042	novel transcript
WWOX	-0.628169	0.005049	WW domain containing oxidoreductase [:HGNC:12799]
PHOSPHO2	-0.851713	0.00507	phosphatase, orphan 2 [:HGNC:28316]
MARCH1	-1.091372	0.005112	membrane associated ring-CH-type finger 1 [:HGNC:26077]
PCLO	-0.589882	0.005131	piccolo presynaptic cytomatrix protein [:HGNC:13406]
NLRP3	-0.851395	0.005484	NLR family pyrin domain containing 3 [:HGNC:16400]
SNORD67	-1.667368	0.005595	small nucleolar RNA, C/D box 67 [:HGNC:32728]
GFAP	-0.617728	0.005628	glial fibrillary acidic protein [:HGNC:4235]
ZNRD1ASP	-0.798717	0.005678	zinc ribbon domain containing 1 antisense, pseudogene [:HGNC:13924]
ARHGAP28	-1.463741	0.00572	Rho GTPase activating protein 28 [:HGNC:25509]
GPR85	-2.737391	0.005749	G protein-coupled receptor 85 [:HGNC:4536]
PLCE1-AS1	-3.210833	0.00597	PLCE1 antisense RNA 1 [:HGNC:45193]
BEX5	-4.557289	0.006064	brain expressed X-linked 5 [:HGNC:27990]
OR10Y1P	-4.557946	0.006205	olfactory receptor family 10 subfamily Y member 1 pseudogene [:HGNC:15140]
FZD8	-0.615005	0.006303	frizzled class receptor 8 [:HGNC:4046]
SDHAP3	-1.193566	0.006334	succinate dehydrogenase complex flavoprotein subunit A pseudogene 3 [:HGNC:18781]
ZNF503	-0.824908	0.006456	zinc finger protein 503 [:HGNC:23589]
C9orf152	-2.774693	0.006461	chromosome 9 open reading frame 152 [:HGNC:31455]
FOXD3-AS1	-0.896121	0.006469	FOXD3 antisense RNA 1 [:HGNC:40241]
HSP90AB3P	-0.762487	0.006607	heat shock protein 90 alpha family class B member 3, pseudogene [:HGNC:5259]
MBLAC1	-0.830338	0.006695	metallo-beta-lactamase domain containing 1 [:HGNC:22180]
CDKL1	-0.655981	0.006709	cyclin dependent kinase like 1 [:HGNC:1781]
HNRNPA3P6	-0.828248	0.006747	heterogeneous nuclear ribonucleoprotein A3 pseudogene 6 [:HGNC:48495]

ATP6V0E2-AS1	-0.582926	0.006992	ATP6V0E2 antisense RNA 1 [:HGNC:44180]
EML5	-1.189359	0.00711	echinoderm microtubule associated protein like 5 [:HGNC:18197]
AC083799.1	-0.622586	0.007195	novel transcript, sense intronic to TMCC1
AL683807.1	-1.905590	0.007204	novel transcript
GDAP1	-0.638611	0.007246	ganglioside induced differentiation associated protein 1 [:HGNC:15968]
MIPOL1	-0.583180	0.007387	mirror-image polydactyly 1 [:HGNC:21460]
ARL2BP	-0.636541	0.007482	ADP ribosylation factor like GTPase 2 binding protein [:HGNC:17146]
FES	-2.536711	0.007821	FES proto-oncogene, tyrosine kinase [:HGNC:3657]
IGFL1P1	-3.968738	0.00788	IGF like family member 1 pseudogene 1 [:HGNC:32956]
AC092807.3	-0.675744	0.007894	novel transcript
ZNF204P	-0.695757	0.008002	zinc finger protein 204, pseudogene [:HGNC:12995]
AL138724.1	-0.960895	0.008206	uncharacterized LOC105374952 [Source:NCBI gene;Acc:105374952]
NLRP10	-1.651633	0.00821	NLR family pyrin domain containing 10 [:HGNC:21464]
BX640514.2	-0.873940	0.008331	novel transcript
ICAM2	-0.993191	0.008665	intercellular adhesion molecule 2 [:HGNC:5345]
LINC00894	-1.085534	0.008803	long intergenic non-protein coding RNA 894 [:HGNC:48579]
LAMB4	-1.605119	0.008851	laminin subunit beta 4 [:HGNC:6491]
MIR17HG	-1.579597	0.009063	miR-17-92a-1 cluster host gene [:HGNC:23564]
EGFL8	-1.262551	0.009434	EGF like domain multiple 8 [:HGNC:13944]
PPM1E	-1.394147	0.009469	protein phosphatase, Mg2+/Mn2+ dependent 1E [:HGNC:19322]
AC099850.1	-0.974053	0.009519	novel transcript
HFM1	-1.783634	0.009763	HFM1, ATP dependent DNA helicase homolog [:HGNC:20193]
RNF128	-0.971811	0.009952	ring finger protein 128, E3 ubiquitin protein ligase [:HGNC:21153]
AL035461.2	-1.629360	0.010082	novel transcript
ADAMTS17	-1.110891	0.01023	ADAM metallopeptidase with thrombospondin type 1 motif 17 [:HGNC:17109]
AC234772.2	-1.052508	0.010302	novel transcript
ENTPD1	-0.807018	0.01032	ectonucleoside triphosphate diphosphohydrolase 1 [:HGNC:3363]
AC090833.1	-0.864465	0.010403	uncharacterized LOC105376603 [Source:NCBI gene;Acc:105376603]
SPARC	-0.776640	0.010892	secreted protein acidic and cysteine rich [:HGNC:11219]
MLLT11	-0.839752	0.010903	MLLT11, transcription factor 7 cofactor [:HGNC:16997]
AC117422.1	-1.499289	0.011176	novel transcript
TSPAN12	-0.583220	0.011249	tetraspanin 12 [:HGNC:21641]
LINC01291	-1.003070	0.011377	long intergenic non-protein coding RNA 1291 [:HGNC:50358]
GOLGA8H	-1.837163	0.011447	golgin A8 family member H [:HGNC:37443]
LY6G5B	-0.583846	0.011571	lymphocyte antigen 6 family member G5B [:HGNC:13931]
ALG10	-0.831860	0.011799	ALG10, alpha-1,2-glucosyltransferase [:HGNC:23162]
NRK	-3.029491	0.011824	Nik related kinase [:HGNC:25391]
Z68871.1	-1.085377	0.011986	novel transcript
FLT1	-4.358703	0.012018	fms related tyrosine kinase 1 [:HGNC:3763]
SLITRK5	-1.455870	0.01231	SLIT and NTRK like family member 5 [:HGNC:20295]
ZNF551	-0.582481	0.012488	zinc finger protein 551 [:HGNC:25108]
EDN2	-1.136340	0.01272	endothelin 2 [:HGNC:3177]
FAS	-0.725863	0.013021	Fas cell surface death receptor [:HGNC:11920]
TMEM236	-1.036433	0.013216	transmembrane protein 236 [:HGNC:23473]
FAM184A	-1.572751	0.013372	family with sequence similarity 184 member A [:HGNC:20991]

LINC02474	-2.999106	0.013387	long intergenic non-protein coding RNA 2474 [:HGNC:53417]
BX255923.1	-1.880482	0.013401	novel transcript
C6orf223	-1.206094	0.013571	chromosome 6 open reading frame 223 [:HGNC:28692]
ANKRD36B	-0.685684	0.013695	ankyrin repeat domain 36B [:HGNC:29333]
RIBC2	-0.892271	0.013803	RIB43A domain with coiled-coils 2 [:HGNC:13241]
SCG5	-0.940531	0.013822	secretogranin V [:HGNC:10816]
AC069499.1	-0.991578	0.01397	ribosomal protein L13 (RPL13) pseudogene
PSG9	-1.598643	0.014005	pregnancy specific beta-1-glycoprotein 9 [:HGNC:9526]
CTAGE1	-3.243776	0.014421	cutaneous T cell lymphoma-associated antigen 1 [:HGNC:24346]
CARD14	-1.202660	0.014488	caspase recruitment domain family member 14 [:HGNC:16446]
LINC01234	-1.303632	0.01471	long intergenic non-protein coding RNA 1234 [:HGNC:49757]
ZNF887P	-1.201635	0.014931	zinc finger protein 887, pseudogene [:HGNC:38700]
LGALS12	-0.798135	0.015058	galectin 12 [:HGNC:15788]
LMTK3	-0.615738	0.016144	lemur tyrosine kinase 3 [:HGNC:19295]
AC004233.3	-1.083598	0.016189	novel transcript
SH3PXD2A-AS1	-2.052483	0.016234	SH3PXD2A antisense RNA 1 [:HGNC:45242]
AC011479.2	-2.496517	0.01641	novel transcript, sense intronic to SIPA1L3
AC092645.1	-0.859243	0.016487	TEC
PDGFD	-1.128681	0.016788	platelet derived growth factor D [:HGNC:30620]
RPS6KA6	-1.392400	0.017016	ribosomal protein S6 kinase A6 [:HGNC:10435]
CSF3	-3.213687	0.017214	colony stimulating factor 3 [:HGNC:2438]
LINC01389	-1.032787	0.017718	long intergenic non-protein coding RNA 1389 [:HGNC:50661]
MORF4L2-AS1	-2.498128	0.018022	MORF4L2 antisense RNA 1 [:HGNC:27991]
PRSS51	-0.661685	0.01807	serine protease 51 [:HGNC:37321]
AP000577.1	-2.268227	0.0181	TEC
AC037486.1	-3.628287	0.018252	novel transcript
RPSAP52	-0.969285	0.018417	ribosomal protein SA pseudogene 52 [:HGNC:35752]
PALM2	-2.537322	0.018562	paralemmin 2 [:HGNC:15845]
AMPH	-2.716089	0.01871	amphiphysin [:HGNC:471]
WDR72	-1.421218	0.018712	WD repeat domain 72 [:HGNC:26790]
MAP3K15	-0.873306	0.018914	mitogen-activated protein kinase kinase kinase 15 [:HGNC:31689]
PDLIM3	-0.745336	0.018917	PDZ and LIM domain 3 [:HGNC:20767]
RPS10	-0.998044	0.018982	ribosomal protein S10 [:HGNC:10383]
MTMR8	-1.252748	0.019452	myotubularin related protein 8 [:HGNC:16825]
AC087741.1	-0.702947	0.019496	novel transcript, antisense to CARD14
TIGD7	-0.820152	0.019564	tigger transposable element derived 7 [:HGNC:18331]
SERBP1P5	-0.584042	0.020028	SERPINE1 mRNA binding protein 1 pseudogene 5 [:HGNC:44632]
PIH1D2	-1.146962	0.020277	PIH1 domain containing 2 [:HGNC:25210]
MESP1	-0.650584	0.020584	mesoderm posterior bHLH transcription factor 1 [:HGNC:29658]
IRX2	-3.597823	0.020696	iroquois homeobox 2 [:HGNC:14359]
DCLK1	-1.049780	0.02101	doublecortin like kinase 1 [:HGNC:2700]
AL627230.2	-1.678278	0.021152	family with sequence similarity 27-like (FAM27L) pseudogene
SLC25A19	-0.691082	0.021785	solute carrier family 25 member 19 [:HGNC:14409]
AC018463.1	-1.432373	0.021855	chromosome 1 open reading frame 80 (C1orf80) pseudogene
RHEBL1	-0.782190	0.022063	RHEB like 1 [:HGNC:21166]

RGS7	-1.224511	0.022386	regulator of G protein signaling 7 [:HGNC:10003]
PDZK1IP1	-1.586895	0.022964	PDZK1 interacting protein 1 [:HGNC:16887]
AC108463.2	-1.450721	0.022984	novel transcript
ALMS1-IT1	-0.906722	0.023015	ALMS1 intronic transcript 1 [:HGNC:41305]
AC092115.1	-2.873954	0.023238	non-POU domain containing, octamer-binding (NONO) pseudogene
KIAA1755	-0.629783	0.023701	KIAA1755 [:HGNC:29372]
EPM2A	-0.596956	0.024034	EPM2A, laforin glucan phosphatase [:HGNC:3413]
SALL2	-1.022765	0.024038	spalt like transcription factor 2 [:HGNC:10526]
HS6ST3	-1.518308	0.024128	heparan sulfate 6-O-sulfotransferase 3 [:HGNC:19134]
AC004943.2	-0.637840	0.024403	novel transcript, antisense to ZFHX3
PGM5P2	-0.789947	0.025231	phosphoglucomutase 5 pseudogene 2 [:HGNC:18965]
SLC4A4	-0.723434	0.02586	solute carrier family 4 member 4 [:HGNC:11030]
PHETA2	-0.888589	0.025863	PH domain containing endocytic trafficking adaptor 2 [:HGNC:27161]
RAD51B	-0.632265	0.025979	RAD51 paralog B [:HGNC:9822]
ANKRD18B	-3.147448	0.026499	ankyrin repeat domain 18B [:HGNC:23644]
AC245140.3	-1.749211	0.027258	novel transcript, antisense to FLNA
CNKSR2	-1.547758	0.027451	connector enhancer of kinase suppressor of Ras 2 [:HGNC:19701]
DHFRP1	-1.647298	0.027453	dihydrofolate reductase pseudogene 1 [:HGNC:2862]
PHACTR1	-1.384711	0.027581	phosphatase and actin regulator 1 [:HGNC:20990]
AC133644.3	-1.870412	0.028052	novel transcript
ERC2	-1.941253	0.028332	ELKS/RAB6-interacting/CAST family member 2 [:HGNC:31922]
FAM155A	-1.324882	0.028399	family with sequence similarity 155 member A [:HGNC:33877]
FLI1	-2.347189	0.028618	Fli-1 proto-oncogene, ETS transcription factor [:HGNC:3749]
KIF17	-1.061022	0.028789	kinesin family member 17 [:HGNC:19167]
AL137003.2	-0.689456	0.029034	novel transcript
AC107308.1	-1.323285	0.029079	novel transcript, sense intronic to HMGA2
POC1B-AS1	-1.163112	0.029162	POC1B antisense RNA 1 [:HGNC:52949]
ALG1L5P	-0.963820	0.029297	asparagine-linked glycosylation 1-like 5, pseudogene [:HGNC:44374]
CYB5R2	-0.808198	0.029369	cytochrome b5 reductase 2 [:HGNC:24376]
FGF1	-0.693835	0.029652	fibroblast growth factor 1 [:HGNC:3665]
KBTBD8	-1.449202	0.03027	kelch repeat and BTB domain containing 8 [:HGNC:30691]
FAM111A-DT	-0.890996	0.03084	FAM111A divergent transcript [:HGNC:53752]
LINC00630	-0.672469	0.032341	long intergenic non-protein coding RNA 630 [:HGNC:44263]
MS4A4A	-3.450009	0.033032	membrane spanning 4-domains A4A [:HGNC:13371]
SLC35A1	-0.589332	0.033259	solute carrier family 35 member A1 [:HGNC:11021]
BBOF1	-0.684260	0.033393	basal body orientation factor 1 [:HGNC:19855]
MIR4258	-1.630577	0.033609	microRNA 4258 [:HGNC:38281]
MNS1	-0.596948	0.033618	meiosis specific nuclear structural 1 [:HGNC:29636]
Z83843.1	-0.974519	0.033757	novel transcript, sense intronic FTX
LINC00624	-0.832830	0.034004	long intergenic non-protein coding RNA 624 [:HGNC:44254]
PRRT1	-1.061820	0.03464	proline rich transmembrane protein 1 [:HGNC:13943]
AC026412.1	-0.603431	0.035002	programmed cell death 6 pseudogene [Source:NCBI gene;Acc:728613]
PTGES3L	-1.320827	0.035552	prostaglandin E synthase 3 like [:HGNC:43943]
AP003392.4	-0.646216	0.036109	novel transcript
ARID3A	-0.714877	0.036405	AT-rich interaction domain 3A [:HGNC:3031]

FAXC	-0.936758	0.036433	failed axon connections homolog [:HGNC:20742]
AL355512.1	-1.052136	0.03701	novel transcript
STX11	-1.014469	0.037203	syntaxin 11 [:HGNC:11429]
CXCL11	-1.157138	0.037436	C-X-C motif chemokine ligand 11 [:HGNC:10638]
SNORD100	-1.705415	0.03749	small nucleolar RNA, C/D box 100 [:HGNC:32763]
TMEM217	-1.036014	0.037771	transmembrane protein 217 [:HGNC:21238]
HIST1H3B	-1.749339	0.037954	histone cluster 1 H3 family member b [:HGNC:4776]
C1QL1	-0.614859	0.038256	complement C1q like 1 [:HGNC:24182]
AC005180.1	-1.875631	0.038293	novel transcript
KDM4D	-0.800656	0.039156	lysine demethylase 4D [:HGNC:25498]
AP000866.2	-0.976686	0.039697	uncharacterized LOC101929340 [Source:NCBI gene;Acc:101929340]
IRAK1BP1	-0.824583	0.040269	interleukin 1 receptor associated kinase 1 binding protein 1 [:HGNC:17368]
AC121761.1	-0.730343	0.040412	novel transcript, antisense to GLIPR1
AC087752.4	-2.185120	0.04056	novel transcript, antisense to CCNE2
BOLA3-AS1	-1.300481	0.04064	BOLA3 divergent transcript [:HGNC:42922]
SNORA33	-1.230960	0.040783	small nucleolar RNA, H/ACA box 33 [:HGNC:32623]
MYLK2	-1.822488	0.041389	myosin light chain kinase 2 [:HGNC:16243]
AL121832.3	-0.598535	0.041475	novel transcript, sense intronic to CABLES2
LINC00702	-0.662701	0.041999	long intergenic non-protein coding RNA 702 [:HGNC:44676]
SYCP2L	-0.970051	0.042089	synaptonemal complex protein 2 like [:HGNC:21537]
DMRT2	-2.946029	0.043151	doublesex and mab-3 related transcription factor 2 [:HGNC:2935]
MAGEC1	-2.479210	0.043151	MAGE family member C1 [:HGNC:6812]
PTPRQ	-2.660841	0.045175	protein tyrosine phosphatase, receptor type Q [:HGNC:9679]
AC135178.3	-0.901364	0.045205	novel transcript, antisense to KRBA2 and RPL26
HOOK1	-1.081032	0.045658	hook microtubule tethering protein 1 [:HGNC:19884]
ADCY10P1	-0.685318	0.045827	adenylate cyclase 10, soluble pseudogene 1 [:HGNC:44143]
CHRFAM7A	-1.367130	0.046378	CHRNA7 (exons 5-10) and FAM7A (exons A-E) fusion [:HGNC:15781]
AL590644.1	-0.590214	0.046731	novel transcript
PAX9	-1.780388	0.047238	paired box 9 [:HGNC:8623]
SOST	-1.339376	0.047307	sclerostin [:HGNC:13771]
UCP3	-0.933419	0.048222	uncoupling protein 3 [:HGNC:12519]
OR2B6	-2.871363	0.048502	olfactory receptor family 2 subfamily B member 6 [:HGNC:8241]
RPS18P9	-0.862135	0.048524	ribosomal protein S18 pseudogene 9 [:HGNC:36483]
HAUS7	-0.971862	0.048599	HAUS augmin like complex subunit 7 [:HGNC:32979]
AC245140.2	-0.896459	0.048882	novel transcript, antisense to RPL10
AP001893.1	-1.919090	0.050088	novel transcript
IKBKGP1	-1.189916	0.050475	inhibitor of nuclear factor kappa B kinase subunit gamma pseudogene 1 [:HGNC:24455]
AC090589.3	-0.902876	0.050607	novel transcript

Table 3.2 – Differentially expressed genes (DEGs) in DU145 SPAG5 knockdown. The tables present the list of the 435 the most upregulated and 472 downregulated genes out of 907 identified, obtained applying Log2FC and a cut-off of 0.58. Gene listed in orange are upregulated in knockdown vs control cell populations, whereas the blue table are genes downregulated in knockdown vs control cell populations.

Gene ID	log2FC	Pvalue	Gene Description
MYL9	1.577447	1.07E-93	myosin light chain 9 [HGNC:15754]
MFAP5	3.764184	7.91E-63	microfibril associated protein 5 [HGNC:29673]
FSTL3	1.240095	4.94E-58	follistatin like 3 [HGNC:3973]
CCND3	1.295482	2.57E-57	cyclin D3 [HGNC:1585]
COL4A2	0.902212	1.07E-48	collagen type IV alpha 2 chain [HGNC:2203]
CCNB1	1.154629	7.88E-44	cyclin B1 [HGNC:1579]
TPM1	1.034945	1.26E-43	tropomyosin 1 [HGNC:12010]
HYOU1	1.110194	2.75E-42	hypoxia up-regulated 1 [HGNC:16931]
L1CAM	1.618426	1.14E-37	L1 cell adhesion molecule [HGNC:6470]
KIF20A	0.994810	3.99E-34	kinesin family member 20A [HGNC:9787]
SDC2	0.816378	2.11E-32	syndecan 2 [HGNC:10659]
TAGLN	1.737152	3.32E-32	transgelin [HGNC:11553]
MANF	1.498763	6.18E-32	mesencephalic astrocyte derived neurotrophic factor [HGNC:15461]
TGFB2	0.973249	2.34E-31	transforming growth factor beta 2 [HGNC:11768]
KLF10	1.131694	3.77E-30	Kruppel like factor 10 [HGNC:11810]
FLNA	0.668684	9.9E-30	filamin A [HGNC:3754]
ANXA8	2.040325	1.81E-28	annexin A8 [HGNC:546]
AURKA	1.155469	3.31E-28	aurora kinase A [HGNC:11393]
CPT1A	0.947283	1.51E-27	carnitine palmitoyltransferase 1A [HGNC:2328]
ADAMTSL4	1.155428	2.19E-27	ADAMTS like 4 [HGNC:19706]
DNAJB11	0.870588	6.08E-27	DnaJ heat shock protein family (Hsp40) member B11 [HGNC:14889]
PLK1	1.087892	7.83E-27	polo like kinase 1 [HGNC:9077]
CDC20	0.969655	1.1E-26	cell division cycle 20 [HGNC:1723]
BMI1	0.827339	2.21E-26	BMI1 proto-oncogene, polycomb ring finger [HGNC:1066]
PTTG1	1.159973	2.3E-25	pituitary tumor-transforming 1 [HGNC:9690]
ANXA6	0.756524	8.38E-25	annexin A6 [HGNC:544]
IQGAP3	0.923058	1.35E-24	IQ motif containing GTPase activating protein 3 [HGNC:20669]
HES1	1.081622	2.13E-24	hes family bHLH transcription factor 1 [HGNC:5192]
SLC1A1	1.087195	4.71E-24	solute carrier family 1 member 1 [HGNC:10939]
CDHR1	1.354170	8.48E-24	cadherin related family member 1 [HGNC:14550]
DEPDC1	0.914623	1.36E-23	DEP domain containing 1 [HGNC:22949]
NID2	1.381226	2.01E-23	nidogen 2 [HGNC:13389]
ESPL1	0.883108	3.04E-23	extra spindle pole bodies like 1, separase [HGNC:16856]
WNT5A	1.234124	5.06E-23	Wnt family member 5A [HGNC:12784]
APLP1	1.110070	7.22E-23	amyloid beta precursor like protein 1 [HGNC:597]
TPX2	0.846801	8.51E-22	TPX2, microtubule nucleation factor [HGNC:1249]
KIF14	0.962606	1.49E-20	kinesin family member 14 [HGNC:19181]
F13A1	1.070358	1.53E-20	coagulation factor XIII A chain [HGNC:3531]
UBE2S	0.696981	4.59E-20	ubiquitin conjugating enzyme E2 S [HGNC:17895]
ARHGAP11A	0.724092	1.09E-18	Rho GTPase activating protein 11A [HGNC:15783]

COL5A1	0.957377	3.11E-18	collagen type V alpha 1 chain [HGNC:2209]
CCNB2	0.900736	5.81E-18	cyclin B2 [HGNC:1580]
VTN	2.036796	2.81E-17	vitronectin [HGNC:12724]
KIF23	0.783961	2.81E-17	kinesin family member 23 [HGNC:6392]
COL4A1	0.658129	4.02E-17	collagen type IV alpha 1 chain [HGNC:2202]
FBLN1	1.082728	4.28E-17	fibulin 1 [HGNC:3600]
COL12A1	0.713313	5.57E-17	collagen type XII alpha 1 chain [HGNC:2188]
CEP55	0.790080	7.79E-17	centrosomal protein 55 [HGNC:1161]
PAPSS2	0.733276	7.89E-17	3'-phosphoadenosine 5'-phosphosulfate synthase 2 [HGNC:8604]
TROAP	0.867124	1.06E-16	trophinin associated protein [HGNC:12327]
CRELD2	1.010806	1.25E-16	cysteine rich with EGF like domains 2 [HGNC:28150]
PTPRU	0.710550	1.97E-16	protein tyrosine phosphatase, receptor type U [HGNC:9683]
TACC3	0.645677	2.48E-16	transforming acidic coiled-coil containing protein 3 [HGNC:11524]
TGFB1I1	1.360809	2.84E-16	transforming growth factor beta 1 induced transcript 1 [HGNC:11767]
LAMA5	0.681509	3.21E-16	laminin subunit alpha 5 [HGNC:6485]
PDIA4	0.733855	3.28E-16	protein disulfide isomerase family A member 4 [HGNC:30167]
GALNS	0.815269	3.59E-16	galactosamine (N-acetyl)-6-sulfatase [HGNC:4122]
KIF4A	0.887816	4.45E-16	kinesin family member 4A [HGNC:13339]
HMMR	0.956844	4.75E-16	hyaluronan mediated motility receptor [HGNC:5012]
SDF2L1	1.236556	6.68E-16	stromal cell derived factor 2 like 1 [HGNC:10676]
KYNU	0.794587	7.59E-16	kynureninase [HGNC:6469]
MFGE8	0.581630	1.48E-15	milk fat globule-EGF factor 8 protein [HGNC:7036]
KIF2C	0.904733	4.52E-15	kinesin family member 2C [HGNC:6393]
CACNG4	1.361348	4.83E-15	calcium voltage-gated channel auxiliary subunit gamma 4 [HGNC:1408]
BMP7	0.777704	5.86E-15	bone morphogenetic protein 7 [HGNC:1074]
SORL1	0.666424	7.67E-15	sortilin related receptor 1 [HGNC:11185]
DLGAP5	0.874457	1.83E-14	DLG associated protein 5 [HGNC:16864]
RHOB	0.644178	3.08E-14	ras homolog family member B [HGNC:668]
UBE2C	0.753451	4.04E-14	ubiquitin conjugating enzyme E2 C [HGNC:15937]
MIR22HG	1.016559	6.32E-14	MIR22 host gene [HGNC:28219]
EDEM2	0.746020	6.7E-14	ER degradation enhancing alpha-mannosidase like protein 2 [HGNC:15877]
C15orf48	1.106880	9.76E-14	chromosome 15 open reading frame 48 [HGNC:29898]
FSTL1	0.612331	1.26E-13	follistatin like 1 [HGNC:3972]
GTSE1	0.852407	1.41E-13	G2 and S-phase expressed 1 [HGNC:13698]
STC1	1.259924	1.62E-13	stanniocalcin 1 [HGNC:11373]
IFT81	0.762185	2.02E-13	intraflagellar transport 81 [HGNC:14313]
MYPN	1.958707	2.04E-13	myopalladin [HGNC:23246]
BIRC5	0.631376	2.44E-13	baculoviral IAP repeat containing 5 [HGNC:593]
MLPH	0.712739	2.53E-13	melanophilin [HGNC:29643]
JPH2	1.694032	3.98E-13	junctophilin 2 [HGNC:14202]
KLF2	1.304591	7.86E-13	Kruppel like factor 2 [HGNC:6347]
C1QTNF6	0.661745	8.49E-13	C1q and TNF related 6 [HGNC:14343]
NUF2	0.941960	8.55E-13	NUF2, NDC80 kinetochore complex component [HGNC:14621]
SLC1A3	1.170434	1.76E-12	solute carrier family 1 member 3 [HGNC:10941]
CCNA2	0.732189	1.89E-12	cyclin A2 [HGNC:1578]

SPDL1	0.718517	2.06E-12	spindle apparatus coiled-coil protein 1 [HGNC:26010]
AP003119.3	0.830266	2.23E-12	novel transcript, overlapping to TSKU
PDGFA	0.960047	3.06E-12	platelet derived growth factor subunit A [HGNC:8799]
PCSK1N	0.717357	1.21E-11	proprotein convertase subtilisin/kexin type 1 inhibitor [HGNC:17301]
FAM83D	0.769015	1.46E-11	family with sequence similarity 83 member D [HGNC:16122]
EIF4A1	0.951924	1.51E-11	eukaryotic translation initiation factor 4A1 [HGNC:3282]
PDIA3	0.601702	1.71E-11	protein disulfide isomerase family A member 3 [HGNC:4606]
MYBL2	0.640941	3.15E-11	MYB proto-oncogene like 2 [HGNC:7548]
NEK2	0.801270	3.64E-11	NIMA related kinase 2 [HGNC:7745]
CDC25C	1.248774	4.22E-11	cell division cycle 25C [HGNC:1727]
JPH3	1.261639	4.55E-11	junctophilin 3 [HGNC:14203]
BUB1	0.738979	4.98E-11	BUB1 mitotic checkpoint serine/threonine kinase [HGNC:1148]
PIMREG	0.907464	7.73E-11	PICALM interacting mitotic regulator [HGNC:25483]
CENPE	0.899009	8.01E-11	centromere protein E [HGNC:1856]
CDCA3	0.758574	8.19E-11	cell division cycle associated 3 [HGNC:14624]
AURKB	0.693197	8.31E-11	aurora kinase B [HGNC:11390]
CPA4	1.237109	8.85E-11	carboxypeptidase A4 [HGNC:15740]
PPIB	0.719939	9.43E-11	peptidylprolyl isomerase B [HGNC:9255]
PIF1	1.133292	1.19E-10	PIF1 5'-to-3' DNA helicase [HGNC:26220]
MDGA1	1.655782	1.36E-10	MAM domain containing glycosylphosphatidylinositol anchor 1 [HGNC:19267]
GPC4	0.731452	1.36E-10	glypican 4 [HGNC:4452]
ADGRG2	0.673876	2.53E-10	adhesion G protein-coupled receptor G2 [HGNC:4516]
PCYOX1L	0.664453	2.91E-10	prenylcysteine oxidase 1 like [HGNC:28477]
OGDHL	1.100729	3.53E-10	oxoglutarate dehydrogenase like [HGNC:25590]
GGT5	1.086182	3.73E-10	gamma-glutamyltransferase 5 [HGNC:4260]
DNAJB2	0.612689	4.23E-10	DnaJ heat shock protein family (Hsp40) member B2 [HGNC:5228]
FOXM1	0.612689	4.26E-10	forkhead box M1 [HGNC:3818]
KIF18B	0.691552	5.31E-10	kinesin family member 18B [HGNC:27102]
FGB	1.311934	6.36E-10	fibrinogen beta chain [HGNC:3662]
FBXL16	0.662561	8.64E-10	F-box and leucine rich repeat protein 16 [HGNC:14150]
LARGE1	1.278810	9.85E-10	LARGE xylosyl- and glucuronyltransferase 1 [HGNC:6511]
ERO1B	0.704845	1.15E-09	endoplasmic reticulum oxidoreductase 1 beta [HGNC:14355]
AGR2	0.780541	1.28E-09	anterior gradient 2, protein disulphide isomerase family member [HGNC:328]
KIF3A	0.604693	1.7E-09	kinesin family member 3A [HGNC:6319]
SRGN	2.112596	1.74E-09	serglycin [HGNC:9361]
GPRC5B	0.745429	2.37E-09	G protein-coupled receptor class C group 5 member B [HGNC:13308]
CENPA	0.738220	2.5E-09	centromere protein A [HGNC:1851]
CA2	0.783867	2.7E-09	carbonic anhydrase 2 [HGNC:1373]
KNSTRN	0.681578	3.17E-09	kinetochore localized astrin (SPAG5) binding protein [HGNC:30767]
NABP1	0.610273	5.8E-09	nucleic acid binding protein 1 [HGNC:26232]
B4GAT1	0.605328	6.56E-09	beta-1,4-glucuronyltransferase 1 [HGNC:15685]
CDYL2	1.698690	7.21E-09	chromodomain Y like 2 [HGNC:23030]
CRELD1	0.706801	8.09E-09	cysteine rich with EGF like domains 1 [HGNC:14630]
SLCO3A1	1.096220	1.12E-08	solute carrier organic anion transporter family member 3A1 [HGNC:10952]
IFITM10	0.818397	1.19E-08	interferon induced transmembrane protein 10 [HGNC:40022]

ARL10	0.652927	1.98E-08	ADP ribosylation factor like GTPase 10 [HGNC:22042]
MICAL1	0.582303	2.02E-08	microtubule associated monooxygenase, calponin and LIM domain containing 1 [HGNC:20619]
CACNA1B	0.599753	2.44E-08	calcium voltage-gated channel subunit alpha1 B [HGNC:1389]
CHST6	0.769395	2.91E-08	carbohydrate sulfotransferase 6 [HGNC:6938]
C1S	0.954278	3.04E-08	complement C1s [HGNC:1247]
LIMS2	1.103955	3.66E-08	LIM zinc finger domain containing 2 [HGNC:16084]
PLCG2	0.996651	3.73E-08	phospholipase C gamma 2 [HGNC:9066]
PIGP	0.587777	3.82E-08	phosphatidylinositol glycan anchor biosynthesis class P [HGNC:3046]
UACA	0.595434	4.54E-08	uveal autoantigen with coiled-coil domains and ankyrin repeats [HGNC:15947]
CDKN3	0.716382	4.81E-08	cyclin dependent kinase inhibitor 3 [HGNC:1791]
PSRC1	0.862130	5.11E-08	proline and serine rich coiled-coil 1 [HGNC:24472]
AC091057.1	0.752461	6.4E-08	OTU deubiquitinase 7A pseudogene [Source:NCBI gene;Acc:100288637]
TGFB2-AS1	1.528421	6.9E-08	TGFB2 antisense RNA 1 (head to head) [HGNC:50628]
BVES	0.640130	7.56E-08	blood vessel epicardial substance [HGNC:1152]
CIT	0.644837	8.01E-08	citron rho-interacting serine/threonine kinase [HGNC:1985]
HSP90B1	1.252341	8.19E-08	heat shock protein 90 beta family member 1 [HGNC:12028]
TMEM178B	1.314061	8.39E-08	transmembrane protein 178B [HGNC:44112]
CDCA8	0.644503	9.76E-08	cell division cycle associated 8 [HGNC:14629]
OLFM1	0.888464	1.02E-07	olfactomedin 1 [HGNC:17187]
ROR2	1.154521	1.02E-07	receptor tyrosine kinase like orphan receptor 2 [HGNC:10257]
PRR11	0.612310	1.23E-07	proline rich 11 [HGNC:25619]
PIEZO2	1.814232	1.52E-07	piezo type mechanosensitive ion channel component 2 [HGNC:26270]
MXD3	0.851801	1.57E-07	MAX dimerization protein 3 [HGNC:14008]
DNER	1.246062	2.95E-07	delta/notch like EGF repeat containing [HGNC:24456]
SHCBP1	0.602629	2.95E-07	SHC binding and spindle associated 1 [HGNC:29547]
WNT10A	1.590658	3.2E-07	Wnt family member 10A [HGNC:13829]
IQCA1	1.508737	3.21E-07	IQ motif containing with AAA domain 1 [HGNC:26195]
ADAMTS2	1.172996	3.54E-07	ADAM metallopeptidase with thrombospondin type 1 motif 2 [HGNC:218]
DISP2	1.466851	3.8E-07	dispatched RND transporter family member 2 [HGNC:19712]
TCTA	0.737985	5.89E-07	T cell leukemia translocation altered [HGNC:11692]
ABCG1	0.910617	8.73E-07	ATP binding cassette subfamily G member 1 [HGNC:73]
TMOD2	0.590439	9.07E-07	tropomodulin 2 [HGNC:11872]
PDGFB	0.625187	9.3E-07	platelet derived growth factor subunit B [HGNC:8800]
CTSV	0.761568	9.68E-07	cathepsin V [HGNC:2538]
APOBEC3B	1.226365	9.93E-07	apolipoprotein B mRNA editing enzyme catalytic subunit 3B [HGNC:17352]
ISM2	1.553165	1.07E-06	isthmin 2 [HGNC:23176]
ZNF469	1.670064	1.26E-06	zinc finger protein 469 [HGNC:23216]
SMOC1	0.876124	1.26E-06	SPARC related modular calcium binding 1 [HGNC:20318]
KIF26A	0.595901	1.3E-06	kinesin family member 26A [HGNC:20226]
NPTX2	0.878898	1.33E-06	neuronal pentraxin 2 [HGNC:7953]
SLC9A3-AS1	0.655729	1.39E-06	SLC9A3 antisense RNA 1 [HGNC:40550]
ZNRF1	0.630668	1.5E-06	zinc and ring finger 1 [HGNC:18452]
STK36	0.622457	1.57E-06	serine/threonine kinase 36 [HGNC:17209]
C14orf132	0.665647	1.57E-06	chromosome 14 open reading frame 132 [HGNC:20346]
MYO7B	2.699282	1.81E-06	myosin VIIB [HGNC:7607]

GBP1	1.904376	2.01E-06	guanylate binding protein 1 [HGNC:4182]
AL161431.1	0.717330	2.03E-06	novel transcript
SDK2	2.153069	2.14E-06	sidekick cell adhesion molecule 2 [HGNC:19308]
AP3B2	1.344077	2.26E-06	adaptor related protein complex 3 subunit beta 2 [HGNC:567]
EPHB3	0.584061	2.41E-06	EPH receptor B3 [HGNC:3394]
ARHGEF39	0.663413	3.45E-06	Rho guanine nucleotide exchange factor 39 [HGNC:25909]
LAMP3	0.793268	3.45E-06	lysosomal associated membrane protein 3 [HGNC:14582]
GYG2	0.734615	4.08E-06	glycogenin 2 [HGNC:4700]
VASH1	0.589448	4.66E-06	vasohibin 1 [HGNC:19964]
SGK1	2.063809	4.8E-06	serum/glucocorticoid regulated kinase 1 [HGNC:10810]
STRA6	2.263711	5.11E-06	stimulated by retinoic acid 6 [HGNC:30650]
RGS19	0.628162	5.35E-06	regulator of G protein signaling 19 [HGNC:13735]
CXCR4	0.696363	5.56E-06	C-X-C motif chemokine receptor 4 [HGNC:2561]
CCDC3	0.950306	6.52E-06	coiled-coil domain containing 3 [HGNC:23813]
JAZF1	0.796509	7.03E-06	JAZF zinc finger 1 [HGNC:28917]
SPSB4	1.284441	7.43E-06	splA/ryanodine receptor domain and SOCS box containing 4 [HGNC:30630]
IQGAP2	0.835464	7.6E-06	IQ motif containing GTPase activating protein 2 [HGNC:6111]
RSPO1	1.907397	8.07E-06	R-spondin 1 [HGNC:21679]
HJURP	0.618156	8.14E-06	Holliday junction recognition protein [HGNC:25444]
CRLF1	0.885059	8.97E-06	cytokine receptor like factor 1 [HGNC:2364]
PCBP3	0.666450	9.26E-06	poly(rC) binding protein 3 [HGNC:8651]
GNAZ	0.675822	9.37E-06	G protein subunit alpha z [HGNC:4395]
PCDHA12	0.757952	9.8E-06	protocadherin alpha 12 [HGNC:8666]
DIAPH3	0.592263	1.08E-05	diaphanous related formin 3 [HGNC:15480]
CCNA1	1.278766	1.17E-05	cyclin A1 [HGNC:1577]
ANXA8L1	3.537973	1.24E-05	annexin A8 like 1 [HGNC:23334]
NFASC	1.515927	1.24E-05	neurofascin [HGNC:29866]
HEMK1	0.659241	1.3E-05	HemK methyltransferase family member 1 [HGNC:24923]
HUNK	0.890349	1.35E-05	hormonally up-regulated Neu-associated kinase [HGNC:13326]
SYCP2	1.026410	1.88E-05	synaptonemal complex protein 2 [HGNC:11490]
COMMD3	0.681583	1.88E-05	COMM domain containing 3 [HGNC:23332]
ADAM19	0.652973	1.9E-05	ADAM metallopeptidase domain 19 [HGNC:197]
FREM2	0.608182	1.97E-05	FRAS1 related extracellular matrix protein 2 [HGNC:25396]
ADRA2C	0.750245	1.98E-05	adrenoceptor alpha 2C [HGNC:283]
PARPBP	0.695014	1.99E-05	PARP1 binding protein [HGNC:26074]
CALR	1.136629	2E-05	calreticulin [HGNC:1455]
CENPN	0.616165	2.19E-05	centromere protein N [HGNC:30873]
CENPI	0.953730	2.5E-05	centromere protein I [HGNC:3968]
ATP8A2	0.743271	2.55E-05	ATPase phospholipid transporting 8A2 [HGNC:13533]
SLC9A2	0.793129	2.85E-05	solute carrier family 9 member A2 [HGNC:11072]
FAM72B	0.935424	3E-05	family with sequence similarity 72 member B [HGNC:24805]
HSPA5	0.940678	3E-05	heat shock protein family A (Hsp70) member 5 [HGNC:5238]
DBF4B	0.643335	3.09E-05	DBF4 zinc finger B [HGNC:17883]
NRG2	0.835333	3.26E-05	neuregulin 2 [HGNC:7998]
RASGRF2	1.261211	3.41E-05	Ras protein specific guanine nucleotide releasing factor 2 [HGNC:9876]

ACSS1	0.876009	3.52E-05	acyl-CoA synthetase short chain family member 1 [HGNC:16091]
ARHGEF6	1.396936	3.64E-05	Rac/Cdc42 guanine nucleotide exchange factor 6 [HGNC:685]
CSRNP1	0.606529	4.7E-05	cysteine and serine rich nuclear protein 1 [HGNC:14300]
AC005696.4	2.588961	6.13E-05	novel transcript
DLGAP3	0.782292	6.66E-05	DLG associated protein 3 [HGNC:30368]
VWA5B2	0.720673	6.66E-05	von Willebrand factor A domain containing 5B2 [HGNC:25144]
PCDHB5	0.915429	6.83E-05	protocadherin beta 5 [HGNC:8690]
HAS2	1.764182	6.93E-05	hyaluronan synthase 2 [HGNC:4819]
EFHC1	0.657381	7.36E-05	EF-hand domain containing 1 [HGNC:16406]
EMC9	0.633848	9.34E-05	ER membrane protein complex subunit 9 [HGNC:20273]
PRIMA1	1.464495	0.000104	proline rich membrane anchor 1 [HGNC:18319]
NEURL1B	0.690430	0.000104	neuralized E3 ubiquitin protein ligase 1B [HGNC:35422]
EVA1C	0.606744	0.000109	eva-1 homolog C [HGNC:13239]
ABCA1	0.703068	0.00011	ATP binding cassette subfamily A member 1 [HGNC:29]
KCNK13	2.369310	0.000112	potassium two pore domain channel subfamily K member 13 [HGNC:6275]
STK32B	2.198164	0.000124	serine/threonine kinase 32B [HGNC:14217]
GPR1	2.675521	0.000137	G protein-coupled receptor 1 [HGNC:4463]
PSMG3-AS1	0.613651	0.000142	PSMG3 antisense RNA 1 (head to head) [HGNC:22230]
AC025259.3	1.288306	0.000144	novel transcript
RNF165	1.242587	0.000159	ring finger protein 165 [HGNC:31696]
SLC4A8	0.870625	0.000169	solute carrier family 4 member 8 [HGNC:11034]
AC008429.1	1.194149	0.000184	uncharacterized LOC100268168 [Source:NCBI gene;Acc:100268168]
TMC7	0.677576	0.000211	transmembrane channel like 7 [HGNC:23000]
PLA2R1	0.700437	0.000218	phospholipase A2 receptor 1 [HGNC:9042]
MAP2K6	0.829058	0.00022	mitogen-activated protein kinase kinase 6 [HGNC:6846]
REEP1	0.787132	0.000222	receptor accessory protein 1 [HGNC:25786]
FAM72C	1.022774	0.000256	family with sequence similarity 72 member C [HGNC:30602]
AMPH	0.717730	0.000264	amphiphysin [HGNC:471]
POC1A	0.727427	0.000267	POC1 centriolar protein A [HGNC:24488]
GPR17	2.184271	0.000313	G protein-coupled receptor 17 [HGNC:4471]
CEP19	0.754661	0.000342	centrosomal protein 19 [HGNC:28209]
LINC01006	0.874294	0.000351	long intergenic non-protein coding RNA 1006 [HGNC:48971]
KDELC1	0.603857	0.000352	KDEL motif containing 1 [HGNC:19350]
LHX2	0.601398	0.00036	LIM homeobox 2 [HGNC:6594]
SEPT4	0.961705	0.00036	septin 4 [HGNC:9165]
SLC34A3	0.880487	0.000396	solute carrier family 34 member 3 [HGNC:20305]
CMPK2	1.374976	0.000404	cytidine/uridine monophosphate kinase 2 [HGNC:27015]
OXTR	0.894038	0.000409	oxytocin receptor [HGNC:8529]
TBX1	0.797198	0.000444	T-box 1 [HGNC:11592]
GNB1L	0.695961	0.000445	G protein subunit beta 1 like [HGNC:4397]
LIMD2	0.603792	0.000447	LIM domain containing 2 [HGNC:28142]
FRMPD1	1.274548	0.000454	FERM and PDZ domain containing 1 [HGNC:29159]
SAA1	1.522207	0.000471	serum amyloid A1 [HGNC:10513]
CENPF	0.809486	0.000486	centromere protein F [HGNC:1857]
WNT6	2.084784	0.0005	Wnt family member 6 [HGNC:12785]

AC010168.2	1.370514	0.000509	novel transcript, overlapping HIST4H4
SCUBE1	0.787325	0.000548	signal peptide, CUB domain and EGF like domain containing 1 [HGNC:13441]
SGO1	0.806209	0.000558	shugoshin 1 [HGNC:25088]
TSHZ3	1.039190	0.000562	teashirt zinc finger homeobox 3 [HGNC:30700]
RPL23AP82	0.644475	0.000572	ribosomal protein L23a pseudogene 82 [HGNC:33730]
AC244153.1	1.604948	0.000643	uncharacterized LOC101929494 [Source:NCBI gene;Acc:101929494]
ALDH3A1	0.606531	0.000692	aldehyde dehydrogenase 3 family member A1 [HGNC:405]
LGALS1	0.942561	0.000694	galectin 1 [HGNC:6561]
C12orf73	0.583827	0.000694	chromosome 12 open reading frame 73 [HGNC:34450]
ONECUT3	1.562242	0.000696	one cut homeobox 3 [HGNC:13399]
CHST8	1.880335	0.000752	carbohydrate sulfotransferase 8 [HGNC:15993]
SUSD4	0.824742	0.000754	sushi domain containing 4 [HGNC:25470]
KCNK12	2.635792	0.00076	potassium two pore domain channel subfamily K member 12 [HGNC:6274]
LOX	0.951869	0.000806	lysyl oxidase [HGNC:6664]
RNF224	0.909891	0.000825	ring finger protein 224 [HGNC:41912]
GTF2H2	0.966374	0.000831	general transcription factor IIH subunit 2 [HGNC:4656]
SYT11	1.486923	0.000833	synaptotagmin 11 [HGNC:19239]
CIP2A	0.638576	0.000862	cell proliferation regulating inhibitor of protein phosphatase 2A [HGNC:29302]
NUDT1	0.639029	0.00088	nudix hydrolase 1 [HGNC:8048]
ITIH5	1.472695	0.00107	inter-alpha-trypsin inhibitor heavy chain family member 5 [HGNC:21449]
TMEM151A	0.778914	0.001125	transmembrane protein 151A [HGNC:28497]
KLRG2	0.753649	0.001156	killer cell lectin like receptor G2 [HGNC:24778]
AC018553.1	1.505205	0.001168	novel transcript
CBFA2T3	0.714234	0.001272	CBFA2/RUNX1 translocation partner 3 [HGNC:1537]
LINC00689	1.045602	0.001298	long intergenic non-protein coding RNA 689 [HGNC:27217]
ERCC6L	0.719304	0.001331	ERCC excision repair 6 like, spindle assembly checkpoint helicase [HGNC:20794]
HCN2	0.840094	0.001355	hyperpolarization activated cyclic nucleotide gated potassium and sodium channel 2 [HGNC:4846]
HES7	0.617419	0.001394	hes family bHLH transcription factor 7 [HGNC:15977]
CACNA1G	0.784093	0.001418	calcium voltage-gated channel subunit alpha1 G [HGNC:1394]
FBXL2	0.639874	0.001478	F-box and leucine rich repeat protein 2 [HGNC:13598]
WNT3A	1.343735	0.001478	Wnt family member 3A [HGNC:15983]
TMEM130	1.936504	0.001692	transmembrane protein 130 [HGNC:25429]
OLIG1	2.473408	0.001706	oligodendrocyte transcription factor 1 [HGNC:16983]
CCDC163	0.978824	0.001862	coiled-coil domain containing 163 [HGNC:27003]
PAQR5	0.597200	0.001875	progestin and adipoQ receptor family member 5 [HGNC:29645]
NMU	0.847166	0.001938	neuromedin U [HGNC:7859]
FBXL8	0.598747	0.002005	F-box and leucine rich repeat protein 8 [HGNC:17875]
ACYP1	0.594570	0.002137	acylphosphatase 1 [HGNC:179]
VIPR2	0.853822	0.002216	vasoactive intestinal peptide receptor 2 [HGNC:12695]
TP73	0.767071	0.002407	tumor protein p73 [HGNC:12003]
AC092919.2	1.270309	0.002419	TEC
BAMBI	0.643095	0.002421	BMP and activin membrane bound inhibitor [HGNC:30251]
FBXO43	1.069397	0.002585	F-box protein 43 [HGNC:28521]
DUBR	0.898224	0.002609	DPPA2 upstream binding RNA [HGNC:48569]
GRID1	0.962258	0.002731	glutamate ionotropic receptor delta type subunit 1 [HGNC:4575]

ENPP2	1.375052	0.002758	ectonucleotide pyrophosphatase/phosphodiesterase 2 [HGNC:3357]
CYR61	0.873567	0.002768	cysteine rich angiogenic inducer 61 [HGNC:2654]
LINC01278	0.642562	0.003058	long intergenic non-protein coding RNA 1278 [HGNC:28090]
HIST2H2BD	0.734914	0.003097	histone cluster 2 H2B family member d (pseudogene) [HGNC:20517]
UPK3B	1.009285	0.00313	uroplakin 3B [HGNC:21444]
LINC01730	1.084019	0.003143	long intergenic non-protein coding RNA 1730 [HGNC:52518]
VSX1	1.794466	0.003192	visual system homeobox 1 [HGNC:12723]
GPC3	0.834652	0.003257	glypican 3 [HGNC:4451]
OSER1-DT	0.821102	0.003311	OSER1 divergent transcript [HGNC:48585]
TRIM29	2.009711	0.003329	tripartite motif containing 29 [HGNC:17274]
ACBD7	0.966584	0.00337	acyl-CoA binding domain containing 7 [HGNC:17715]
FLVCR2	0.727385	0.003402	feline leukemia virus subgroup C cellular receptor family member 2 [HGNC:20105]
RBP4	1.514343	0.003518	retinol binding protein 4 [HGNC:9922]
MATN3	0.587072	0.003602	matrilin 3 [HGNC:6909]
PLXNC1	0.766074	0.003619	plexin C1 [HGNC:9106]
NKILA	1.450431	0.003718	NF-kappaB interacting lncRNA [HGNC:51599]
PCDH1	1.124218	0.003738	protocadherin 1 [HGNC:8655]
RET	1.293705	0.003781	ret proto-oncogene [HGNC:9967]
INHBB	0.616980	0.003903	inhibin subunit beta B [HGNC:6067]
ATP1B2	0.591201	0.003996	ATPase Na+/K+ transporting subunit beta 2 [HGNC:805]
GXYLT2	0.586265	0.004245	glucoside xylosyltransferase 2 [HGNC:33383]
HMCN2	1.547114	0.004319	hemicentin 2 [HGNC:21293]
NRXN2	1.647429	0.004558	neurexin 2 [HGNC:8009]
AC245140.2	1.268787	0.004628	novel transcript, antisense to RPL10
CAMK2N2	1.212982	0.004766	calcium/calmodulin dependent protein kinase II inhibitor 2 [HGNC:24197]
ARHGAP31	0.923528	0.004811	Rho GTPase activating protein 31 [HGNC:29216]
AC109322.1	1.169969	0.004908	TEC
NPPC	0.951432	0.005014	natriuretic peptide C [HGNC:7941]
DIABLO	0.876256	0.005282	diablo IAP-binding mitochondrial protein [HGNC:21528]
DGCR5	0.593111	0.005418	DiGeorge syndrome critical region gene 5 [HGNC:16757]
AC099850.3	0.740293	0.005845	novel transcript, antisense to PRR11
IRF8	1.181515	0.005853	interferon regulatory factor 8 [HGNC:5358]
SAMD10	0.624480	0.005864	sterile alpha motif domain containing 10 [HGNC:16129]
MKI67	0.638004	0.005965	marker of proliferation Ki-67 [HGNC:7107]
LCN12	1.164767	0.0061	lipocalin 12 [HGNC:28733]
PAX2	0.785387	0.0062	paired box 2 [HGNC:8616]
AC006538.1	0.658450	0.00629	novel transcript
HCN4	0.789599	0.006383	hyperpolarization activated cyclic nucleotide gated potassium channel 4 [HGNC:16882]
WNT10B	0.631847	0.006404	Wnt family member 10B [HGNC:12775]
PANO1	1.457509	0.006466	proapoptotic nucleolar protein 1 [HGNC:51237]
FAM72D	0.647056	0.006567	family with sequence similarity 72 member D [HGNC:33593]
ZNF674-AS1	0.880197	0.006592	ZNF674 antisense RNA 1 (head to head) [HGNC:44266]
MFAP3L	0.611533	0.006789	microfibril associated protein 3 like [HGNC:29083]
ANO2	0.912349	0.006998	anoctamin 2 [HGNC:1183]
SNAI2	0.753398	0.007133	snail family transcriptional repressor 2 [HGNC:11094]

DZIP1	0.728320	0.007397	DAZ interacting zinc finger protein 1 [HGNC:20908]
LPAR3	0.854038	0.007398	lysophosphatidic acid receptor 3 [HGNC:14298]
AC244154.1	0.663080	0.007671	aminopeptidase puromycin sensitive pseudogene [Source:NCBI gene;Acc:440434]
EFCAB11	0.646983	0.00768	EF-hand calcium binding domain 11 [HGNC:20357]
FAM72A	0.761473	0.008014	family with sequence similarity 72 member A [HGNC:24044]
AC026250.1	1.531795	0.008275	uncharacterized LOC440028 [Source:NCBI gene;Acc:440028]
UNC13D	0.965636	0.008285	unc-13 homolog D [HGNC:23147]
PPP1R16B	0.748612	0.008328	protein phosphatase 1 regulatory subunit 16B [HGNC:15850]
XAGE1B	1.269399	0.008719	X antigen family member 1B [HGNC:25400]
AC145098.2	0.819661	0.009042	TEC
JUNB	0.736421	0.009223	JunB proto-oncogene, AP-1 transcription factor subunit [HGNC:6205]
FGA	0.801051	0.009418	fibrinogen alpha chain [HGNC:3661]
MT-TT	1.163411	0.009876	mitochondrially encoded tRNA threonine [HGNC:7499]
LINGO1	0.755201	0.009901	leucine rich repeat and Ig domain containing 1 [HGNC:21205]
SPCS2P4	0.764497	0.009962	signal peptidase complex subunit 2 pseudogene 4 [HGNC:45237]
RPL13P5	0.804987	0.010056	ribosomal protein L13 pseudogene 5 [HGNC:30363]
SMN1	0.586402	0.010267	survival of motor neuron 1, telomeric [HGNC:11117]
ANXA10	1.831515	0.010776	annexin A10 [HGNC:534]
GPR135	0.647033	0.011408	G protein-coupled receptor 135 [HGNC:19991]
FAM19A4	0.680471	0.011961	family with sequence similarity 19 member A4, C-C motif chemokine like [HGNC:21591]
AC067930.5	0.653932	0.012032	novel transcript
TMC3-AS1	1.262973	0.012387	TMC3 antisense RNA 1 [HGNC:51424]
TMEM63C	0.755034	0.012468	transmembrane protein 63C [HGNC:23787]
SMG1P6	1.956591	0.012508	SMG1 pseudogene 6 [HGNC:49863]
PPEF1	1.900940	0.012635	protein phosphatase with EF-hand domain 1 [HGNC:9243]
TINCR	0.816638	0.012687	TINCR ubiquitin domain containing [HGNC:14607]
AC084033.3	0.598449	0.012841	uncharacterized LOC100506844 [Source:NCBI gene;Acc:100506844]
CHRNA4	1.047045	0.012881	cholinergic receptor nicotinic alpha 4 subunit [HGNC:1958]
TNFSF10	0.588212	0.012918	TNF superfamily member 10 [HGNC:11925]
BMP6	0.654309	0.013116	bone morphogenetic protein 6 [HGNC:1073]
KCNJ10	1.166382	0.013142	potassium voltage-gated channel subfamily J member 10 [HGNC:6256]
MESP1	0.633610	0.013142	mesoderm posterior bHLH transcription factor 1 [HGNC:29658]
COX6B2	0.980230	0.014997	cytochrome c oxidase subunit 6B2 [HGNC:24380]
NR1H3	0.758249	0.015438	nuclear receptor subfamily 1 group H member 3 [HGNC:7966]
LDLRAD4	0.765058	0.015621	low density lipoprotein receptor class A domain containing 4 [HGNC:1224]
RUNDC3A	1.063764	0.01567	RUN domain containing 3A [HGNC:16984]
EGR1	1.278728	0.016156	early growth response 1 [HGNC:3238]
AC092718.4	0.675042	0.016574	novel transcript
GNB3	1.014807	0.016688	G protein subunit beta 3 [HGNC:4400]
AC092757.2	1.511449	0.017598	novel transcript, sense overlapping to CCNB2
CACNA2D3	2.043530	0.018109	calcium voltage-gated channel auxiliary subunit alpha2delta 3 [HGNC:15460]
FGFR2	0.729102	0.01835	fibroblast growth factor receptor 2 [HGNC:3689]
SOWAHA	0.810261	0.018838	sosondowah ankyrin repeat domain family member A [HGNC:27033]
FAM49A	1.287927	0.019261	family with sequence similarity 49 member A [HGNC:25373]
FXYD6	0.691914	0.0195	FXYD domain containing ion transport regulator 6 [HGNC:4030]

TRAIP	0.613819	0.019569	TRAF interacting protein [HGNC:30764]
RHBDL3	0.650108	0.019763	rhomboid like 3 [HGNC:16502]
TUBB3	0.851280	0.019904	tubulin beta 3 class III [HGNC:20772]
AC068282.1	0.870058	0.020103	novel transcript
RTN4RL1	0.933449	0.020756	reticulon 4 receptor like 1 [HGNC:21329]
HR	1.027196	0.020846	HR, lysine demethylase and nuclear receptor corepressor [HGNC:5172]
PCOLCE	0.631545	0.021983	procollagen C-endopeptidase enhancer [HGNC:8738]
AC007191.1	0.722030	0.022	TEC
AC022107.1	0.625757	0.022394	TEC
EPHB1	1.478009	0.022561	EPH receptor B1 [HGNC:3392]
FGF19	1.436750	0.022873	fibroblast growth factor 19 [HGNC:3675]
CHST1	1.656946	0.024158	carbohydrate sulfotransferase 1 [HGNC:1969]
ARMC12	0.893815	0.024481	armadillo repeat containing 12 [HGNC:21099]
CDH15	0.896679	0.024653	cadherin 15 [HGNC:1754]
CDH23	0.865035	0.024808	cadherin related 23 [HGNC:13733]
GASAL1	1.213825	0.025491	growth arrest associated lncRNA 1 [HGNC:53461]
AZU1	0.839536	0.025906	azurocidin 1 [HGNC:913]
SLC9A7P1	1.625751	0.026128	solute carrier family 9 member 7 pseudogene 1 [HGNC:32679]
FAAP24	0.591703	0.026347	FA core complex associated protein 24 [HGNC:28467]
GLI1	0.888954	0.026481	GLI family zinc finger 1 [HGNC:4317]
AC002350.1	1.337005	0.026616	novel transcript
AL138781.2	1.313388	0.026989	novel transcript
IQCH-AS1	0.708961	0.027111	IQCH antisense RNA 1 [HGNC:44104]
ROBO1	0.672759	0.028227	roundabout guidance receptor 1 [HGNC:10249]
DNAI1	1.742113	0.028521	dynein axonemal intermediate chain 1 [HGNC:2954]
DPY19L2P1	0.998904	0.028671	DPY19L2 pseudogene 1 [HGNC:22305]
AC034213.1	1.040270	0.028864	novel transcript
AQP11	0.683747	0.028864	aquaporin 11 [HGNC:19940]
CTGF	0.866215	0.029101	connective tissue growth factor [HGNC:2500]
AL365205.1	1.029619	0.02913	novel transcript
ARRDC3	0.745117	0.029316	arrestin domain containing 3 [HGNC:29263]
LINC01270	0.688998	0.02933	long intergenic non-protein coding RNA 1270 [HGNC:27658]
C9orf43	1.221690	0.029753	chromosome 9 open reading frame 43 [HGNC:23570]
SNHG20	0.590822	0.029753	small nucleolar RNA host gene 20 [HGNC:33099]
DMBX1	1.337909	0.030073	diencephalon/mesencephalon homeobox 1 [HGNC:19026]
PAOX	0.623352	0.030727	polyamine oxidase [HGNC:20837]
U73166.1	0.788429	0.031273	novel transcript
FST	0.973936	0.031368	follistatin [HGNC:3971]
ETNPPL	0.631216	0.031738	ethanolamine-phosphate phospho-lyase [HGNC:14404]
PROC	0.975228	0.032968	protein C, inactivator of coagulation factors Va and VIIIa [HGNC:9451]
AL157838.1	1.412384	0.033147	uncharacterized LOC101927770 [Source:NCBI gene;Acc:101927770]
ITGA11	0.883961	0.033445	integrin subunit alpha 11 [HGNC:6136]
WNT5A-AS1	1.255942	0.034667	WNT5A antisense RNA 1 [HGNC:40616]
SPA17	0.640469	0.035528	sperm autoantigenic protein 17 [HGNC:11210]
DCLK2	0.696559	0.035551	doublecortin like kinase 2 [HGNC:19002]

PLXNB3	0.633670	0.037768	plexin B3 [HGNC:9105]
ST6GALNAC6	1.428330	0.038186	ST6 N-acetylgalactosaminide alpha-2,6-sialyltransferase 6 [HGNC:23364]
GSC	0.951960	0.03832	goosecoid homeobox [HGNC:4612]
NR4A1	1.193757	0.038974	nuclear receptor subfamily 4 group A member 1 [HGNC:7980]
AC120114.1	0.805596	0.039067	uncharacterized LOC107984836 [Source:NCBI gene;Acc:107984836]
LINC00472	0.882233	0.041639	long intergenic non-protein coding RNA 472 [HGNC:21380]
EIF1P6	1.454123	0.041886	eukaryotic translation initiation factor 1 pseudogene 6 [HGNC:49619]
AK4P1	0.889320	0.042073	adenylate kinase 4 pseudogene 1 [HGNC:364]
C9orf139	1.106823	0.042272	chromosome 9 open reading frame 139 [HGNC:31426]
AVPR2	1.331345	0.04347	arginine vasopressin receptor 2 [HGNC:897]
SAMD14	0.677914	0.043472	sterile alpha motif domain containing 14 [HGNC:27312]
AC135178.3	0.788208	0.044072	novel transcript, antisense to KRBA2 and RPL26
THBS4	0.870172	0.045263	thrombospondin 4 [HGNC:11788]
ITGB1-DT	0.646341	0.045458	ITGB1 divergent transcript [HGNC:53718]
AC243829.1	1.540936	0.04547	novel transcript, antisense CCL3L3
AC007743.1	1.609508	0.045569	uncharacterized LOC100129434 [Source:NCBI gene;Acc:100129434]
TMEM171	0.844786	0.045719	transmembrane protein 171 [HGNC:27031]
HIC1	0.933172	0.045977	HIC ZBTB transcriptional repressor 1 [HGNC:4909]
CACNA1D	0.746383	0.046233	calcium voltage-gated channel subunit alpha1 D [HGNC:1391]
NSG1	0.894695	0.046409	neuronal vesicle trafficking associated 1 [HGNC:18790]
SLC49A3	0.653343	0.047427	solute carrier family 49 member 3 [HGNC:26177]
BORCS8	0.596615	0.047666	BLOC-1 related complex subunit 8 [HGNC:37247]
CDK5R2	0.867380	0.048043	cyclin dependent kinase 5 regulatory subunit 2 [HGNC:1776]
SNAP91	0.681089	0.048674	synaptosome associated protein 91 [HGNC:14986]
AC008894.3	1.257733	0.048846	TEC
FGF18	1.691507	0.049281	fibroblast growth factor 18 [HGNC:3674]
CYSRT1	0.595612	0.050129	cysteine rich tail 1 [HGNC:30529]

Gene ID	log2FC	Pvalue	Gene Description
BGN	-7.656243	4.10606E-08	biglycan [HGNC:1044]
SLC6A17	-5.348507	2.39523E-07	solute carrier family 6 member 17 [HGNC:31399]
PLPP4	-4.870703	3.35646E-09	phospholipid phosphatase 4 [HGNC:23531]
NKX2-5	-3.954559	3.89353E-06	NK2 homeobox 5 [HGNC:2488]
RPL21P93	-3.484753	1.02655E-05	ribosomal protein L21 pseudogene 93 [HGNC:35646]
CD79A	-3.440928	1.56204E-09	CD79a molecule [HGNC:1698]
CD74	-3.437297	1.12313E-16	CD74 molecule [HGNC:1697]
ZNF503-AS1	-3.310019	5.6617E-05	ZNF503 antisense RNA 1 [HGNC:27370]
FOXB1	-3.260286	2.21428E-09	forkhead box B1 [HGNC:3799]
SLC4A4	-3.207880	1.04936E-31	solute carrier family 4 member 4 [HGNC:11030]
CCND2	-3.118093	1.74966E-10	cyclin D2 [HGNC:1583]

LINC00923	-2.888863	0.000422986	long intergenic non-protein coding RNA 923 [HGNC:28088]
TNC	-2.887328	6.81953E-36	tenascin C [HGNC:5318]
CEMIP	-2.755901	1.80777E-28	cell migration inducing hyaluronidase 1 [HGNC:29213]
AC023906.5	-2.660056	0.000218197	novel transcript, antisense to MAPK6
SHISAL1	-2.631332	1.07176E-93	shisa like 1 [HGNC:29335]
AC011330.1	-2.515530	1.48041E-07	histidine acid phosphatase domain containing 2A (HISPPD2A) pseudogene
SERPINA1	-2.502275	5.50431E-28	serpin family A member 1 [HGNC:8941]
MAGED4B	-2.251705	1.78142E-05	MAGE family member D4B [HGNC:22880]
CST6	-2.228049	0.000110729	cystatin E/M [HGNC:2478]
PDE4B	-2.172971	2.75508E-10	phosphodiesterase 4B [HGNC:8781]
CCDC187	-2.148171	0.000343226	coiled-coil domain containing 187 [HGNC:30942]
PAPPA2	-2.147583	0.001945243	pappalysin 2 [HGNC:14615]
RHBDL2	-2.105192	0.009962059	rhomboid like 2 [HGNC:16083]
KRT16P2	-2.094599	0.009662176	keratin 16 pseudogene 2 [HGNC:37807]
COL4A6	-2.068941	1.63327E-05	collagen type IV alpha 6 chain [HGNC:2208]
AL109615.3	-2.066078	0.000209365	uncharacterized LOC101929705 [Source:NCBI gene;Acc:101929705]
SFTPB	-2.046250	9.12084E-08	surfactant protein B [HGNC:10801]
AJ011932.1	-2.022488	0.003664763	novel transcript
FBXL13	-2.018996	0.001412213	F-box and leucine rich repeat protein 13 [HGNC:21658]
RORC	-2.014065	0.000195348	RAR related orphan receptor C [HGNC:10260]
AL596244.1	-2.012458	1.9391E-08	novel transcript, overlapping TTLL11
BRDT	-1.987383	1.02971E-33	bromodomain testis associated [HGNC:1105]
AL024497.2	-1.973722	0.005017509	novel transcript
SARDH	-1.957014	0.021243256	sarcosine dehydrogenase [HGNC:10536]
CACNG6	-1.944893	1.45994E-11	calcium voltage-gated channel auxiliary subunit gamma 6 [HGNC:13625]
CR769775.1	-1.926617	0.011027458	novel transcript
PLEKHG4	-1.878777	9.00456E-45	pleckstrin homology and RhoGEF domain containing G4 [HGNC:24501]
ZC3H12D	-1.837421	8.39849E-05	zinc finger CCCH-type containing 12D [HGNC:21175]
CBSL	-1.835207	0.037509043	cystathionine-beta-synthase like [HGNC:51829]
GHDC	-1.834199	1.15771E-12	GH3 domain containing [HGNC:24438]
ID2	-1.817736	2.72675E-16	inhibitor of DNA binding 2 [HGNC:5361]
TNFRSF9	-1.797222	0.001666801	TNF receptor superfamily member 9 [HGNC:11924]
BX284668.2	-1.764895	0.013141093	novel transcript
MIR1244-3	-1.728810	0.040250819	microRNA 1244-3 [HGNC:38390]
CLIC2	-1.714965	1.14521E-06	chloride intracellular channel 2 [HGNC:2063]
FBP1	-1.687792	0.001096035	fructose-bisphosphatase 1 [HGNC:3606]
IL24	-1.648329	0.010037576	interleukin 24 [HGNC:11346]
SLC43A3	-1.626199	0.001691909	solute carrier family 43 member 3 [HGNC:17466]
LINC01094	-1.622738	0.021006848	long intergenic non-protein coding RNA 1094 [HGNC:49219]
CXCL8	-1.604605	3.94483E-12	C-X-C motif chemokine ligand 8 [HGNC:6025]
C1orf61	-1.603404	1.00591E-05	chromosome 1 open reading frame 61 [HGNC:30780]
KCTD4	-1.589506	0.005058976	potassium channel tetramerization domain containing 4 [HGNC:23227]
GFAP	-1.565088	0.000429158	glial fibrillary acidic protein [HGNC:4235]
ALOX5	-1.559612	0.044314684	arachidonate 5-lipoxygenase [HGNC:435]
CNTN5	-1.556738	0.024826982	contactin 5 [HGNC:2175]

FCMR	-1.551768	4.84981E-05	Fc fragment of IgM receptor [HGNC:14315]
CNGA1	-1.546763	9.36635E-06	cyclic nucleotide gated channel alpha 1 [HGNC:2148]
AC144831.1	-1.536064	3.47977E-11	novel transcript
C11orf86	-1.531425	1.29775E-08	chromosome 11 open reading frame 86 [HGNC:34442]
RAC2	-1.529392	4.54122E-10	Rac family small GTPase 2 [HGNC:9802]
FOXL1	-1.523419	0.035720041	forkhead box L1 [HGNC:3817]
KLHDC9	-1.501001	0.044778326	kelch domain containing 9 [HGNC:28489]
KIAA1257	-1.497829	0.005713189	KIAA1257 [HGNC:29231]
LINC02009	-1.497412	0.048446446	long intergenic non-protein coding RNA 2009 [HGNC:52845]
NFIB	-1.484125	3.80069E-37	nuclear factor I B [HGNC:7785]
HRNR	-1.475425	5.5562E-08	hornerin [HGNC:20846]
TCHH	-1.464786	0.015079802	trichohyalin [HGNC:11791]
NT5E	-1.456749	1.49059E-31	5'-nucleotidase ecto [HGNC:8021]
C9orf84	-1.447655	1.59816E-16	chromosome 9 open reading frame 84 [HGNC:26535]
JCAD	-1.435306	3.49216E-18	junctional cadherin 5 associated [HGNC:29283]
DIRAS2	-1.426694	0.001272672	DIRAS family GTPase 2 [HGNC:19323]
RAB7B	-1.413981	0.000649827	RAB7B, member RAS oncogene family [HGNC:30513]
TMEM92	-1.412323	0.002938588	transmembrane protein 92 [HGNC:26579]
TNFRSF11A	-1.385872	2.37386E-05	TNF receptor superfamily member 11a [HGNC:11908]
KISS1	-1.373650	0.000300623	KiSS-1 metastasis suppressor [HGNC:6341]
SEMA3A	-1.373586	0.000149207	semaphorin 3A [HGNC:10723]
RINL	-1.370327	0.001916294	Ras and Rab interactor like [HGNC:24795]
AC130371.2	-1.364947	0.020415144	novel transcript
LCP1	-1.364837	0.014007859	lymphocyte cytosolic protein 1 [HGNC:6528]
LINC01508	-1.362757	0.000738822	long intergenic non-protein coding RNA 1508 [HGNC:51190]
APOC1	-1.348636	2.7495E-10	apolipoprotein C1 [HGNC:607]
SESN3	-1.339502	6.88463E-48	sestrin 3 [HGNC:23060]
SLCO2A1	-1.336087	2.99715E-07	solute carrier organic anion transporter family member 2A1 [HGNC:10955]
SPAG5	-1.333891	6.32092E-33	sperm associated antigen 5 [HGNC:13452]
KCNQ3	-1.329875	3.25973E-17	potassium voltage-gated channel subfamily Q member 3 [HGNC:6297]
TMEM215	-1.327664	1.44944E-06	transmembrane protein 215 [HGNC:33816]
TCN2	-1.323719	0.021455576	transcobalamin 2 [HGNC:11653]
COL13A1	-1.308735	7.57309E-09	collagen type XIII alpha 1 chain [HGNC:2190]
HLA-DRA	-1.304335	0.049916105	major histocompatibility complex, class II, DR alpha [HGNC:4947]
SCAMP1	-1.288817	1.47818E-64	secretory carrier membrane protein 1 [HGNC:10563]
BMPER	-1.281891	0.007250538	BMP binding endothelial regulator [HGNC:24154]
IL20RB	-1.273246	5.20123E-29	interleukin 20 receptor subunit beta [HGNC:6004]
AC116407.1	-1.265821	0.001196601	uncharacterized LOC105371730 [Source:NCBI gene;Acc:105371730]
SOX2	-1.265564	0.008373717	SRY-box 2 [HGNC:11195]
CHST9	-1.260743	0.03685967	carbohydrate sulfotransferase 9 [HGNC:19898]
GPNMB	-1.250750	7.49396E-06	glycoprotein nmb [HGNC:4462]
GPER1	-1.241958	2.93058E-05	G protein-coupled estrogen receptor 1 [HGNC:4485]
TBC1D3E	-1.238721	0.050386407	TBC1 domain family member 3E [HGNC:27071]
ANKRD20A10P	-1.236260	0.011376225	ankyrin repeat domain 20 family member A10, pseudogene [HGNC:39707]
HAL	-1.231796	0.034572594	histidine ammonia-lyase [HGNC:4806]

NES	-1.221202	9.15507E-07	nestin [HGNC:7756]
AC091563.1	-1.218885	0.030259988	novel transcript
AP002884.1	-1.211407	1.74561E-10	uncharacterized LOC283140 [Source:NCBI gene;Acc:283140]
NAMPT	-1.203043	2.35001E-80	nicotinamide phosphoribosyltransferase [HGNC:30092]
AC005336.1	-1.200502	7.49557E-11	inositol polyphosphate multikinase (IPMK) pseudogene
NGF	-1.196289	0.000162633	nerve growth factor [HGNC:7808]
RASL11A	-1.189993	3.57299E-11	RAS like family 11 member A [HGNC:23802]
TC2N	-1.187919	8.63614E-10	tandem C2 domains, nuclear [HGNC:19859]
CDK6	-1.176439	7.00153E-22	cyclin dependent kinase 6 [HGNC:1777]
KIF9-AS1	-1.175605	6.81037E-06	KIF9 antisense RNA 1 [HGNC:26822]
GSTM1	-1.172725	9.46572E-13	glutathione S-transferase mu 1 [HGNC:4632]
NAMPTP1	-1.168502	5.33668E-22	nicotinamide phosphoribosyltransferase pseudogene 1 [HGNC:17633]
PPIEL	-1.167916	0.011737487	peptidylprolyl isomerase E like pseudogene [HGNC:33195]
AC092645.1	-1.164551	0.000831203	TEC
CYP4F11	-1.163505	1.45592E-40	cytochrome P450 family 4 subfamily F member 11 [HGNC:13265]
BX284668.5	-1.156719	5.44727E-06	uncharacterized LOC105376805 [Source:NCBI gene;Acc:105376805]
WFDC2	-1.153696	2.7495E-10	WAP four-disulfide core domain 2 [HGNC:15939]
PPP1R37	-1.148615	1.08603E-22	protein phosphatase 1 regulatory subunit 37 [HGNC:27607]
SPATA17	-1.139350	0.005418223	spermatogenesis associated 17 [HGNC:25184]
LIPG	-1.127761	0.002730817	lipase G, endothelial type [HGNC:6623]
RASD2	-1.124659	0.001849147	RASD family member 2 [HGNC:18229]
PRSS8	-1.121844	0.002363701	serine protease 8 [HGNC:9491]
MAP1LC3A	-1.119980	0.010564741	microtubule associated protein 1 light chain 3 alpha [HGNC:6838]
POF1B	-1.116417	1.46771E-07	POF1B, actin binding protein [HGNC:13711]
IFI16	-1.108892	2.87978E-07	interferon gamma inducible protein 16 [HGNC:5395]
ATP2B4	-1.083626	4.46446E-21	ATPase plasma membrane Ca2+ transporting 4 [HGNC:817]
KCTD12	-1.071935	4.28328E-05	potassium channel tetramerization domain containing 12 [HGNC:14678]
FP565260.3	-1.071796	0.03556205	Homo sapiens ICOS ligand (LOC102723996), mRNA. [NM_001363770]
THBD	-1.067972	1.1002E-06	thrombomodulin [HGNC:11784]
AC027117.2	-1.065435	0.017155491	novel transcript
PRODH	-1.062992	0.050374455	proline dehydrogenase 1 [HGNC:9453]
AC138392.1	-1.054844	5.63381E-08	novel pseudogene
RBM11	-1.054105	0.000753218	RNA binding motif protein 11 [HGNC:9897]
CAPG	-1.052012	6.81889E-18	capping actin protein, gelsolin like [HGNC:1474]
SLAMF7	-1.051052	0.000162351	SLAM family member 7 [HGNC:21394]
COL4A5	-1.050050	2.15193E-14	collagen type IV alpha 5 chain [HGNC:2207]
IL32	-1.046571	2.1927E-06	interleukin 32 [HGNC:16830]
IFITM1	-1.041318	0.000998617	interferon induced transmembrane protein 1 [HGNC:5412]
GIT1	-1.035893	2.15272E-41	GIT ArfGAP 1 [HGNC:4272]
PRRG2	-1.032090	0.009768431	proline rich and Gla domain 2 [HGNC:9470]
CBLC	-1.031061	3.55836E-07	Cbl proto-oncogene C [HGNC:15961]
BCL2L13	-1.029693	4.19422E-72	BCL2 like 13 [HGNC:17164]
STRIP2	-1.026951	1.3846E-19	striatin interacting protein 2 [HGNC:22209]
PPP1R14C	-1.024319	6.93716E-06	protein phosphatase 1 regulatory inhibitor subunit 14C [HGNC:14952]
RAB31	-1.022854	1.24356E-58	RAB31, member RAS oncogene family [HGNC:9771]

DPYD	-1.019527	0.006789491	dihydropyrimidine dehydrogenase [HGNC:3012]
CD22	-1.019380	0.042614718	CD22 molecule [HGNC:1643]
AL354718.1	-1.017943	0.001098233	ribosomal protein L7a (RPL7A) pseudogene
S100A14	-1.015499	6.80278E-07	S100 calcium binding protein A14 [HGNC:18901]
ZFP3	-1.013130	1.76916E-05	ZFP3 zinc finger protein [HGNC:12861]
TFR2	-1.012664	0.001734422	transferrin receptor 2 [HGNC:11762]
RAD23A	-1.006474	1.5946E-44	RAD23 homolog A, nucleotide excision repair protein [HGNC:9812]
ZNF43	-1.006019	0.002495265	zinc finger protein 43 [HGNC:13109]
TRIML2	-0.997190	2.79389E-05	tripartite motif family like 2 [HGNC:26378]
HTR2B	-0.995504	0.001748448	5-hydroxytryptamine receptor 2B [HGNC:5294]
G0S2	-0.981673	3.79305E-06	G0/G1 switch 2 [HGNC:30229]
VPS37B	-0.977445	2.95694E-26	VPS37B, ESCRT-I subunit [HGNC:25754]
INSIG1	-0.977140	1.6073E-21	insulin induced gene 1 [HGNC:6083]
STMND1	-0.970543	0.015686122	stathmin domain containing 1 [HGNC:44668]
CREB5	-0.967928	4.70672E-06	cAMP responsive element binding protein 5 [HGNC:16844]
HABP4	-0.966843	4.42696E-15	hyaluronan binding protein 4 [HGNC:17062]
VDR	-0.966130	7.76696E-11	vitamin D receptor [HGNC:12679]
SATB1	-0.966094	9.00808E-10	SATB homeobox 1 [HGNC:10541]
ZNF826P	-0.958748	7.96384E-05	zinc finger protein 826, pseudogene [HGNC:33875]
TSPAN1	-0.954407	5.84469E-10	tetraspanin 1 [HGNC:20657]
HOXB8	-0.950647	0.003443033	homeobox B8 [HGNC:5119]
AIDA	-0.950401	6.32017E-14	axin interactor, dorsalization associated [HGNC:25761]
GOLT1A	-0.949299	0.000382889	golgi transport 1A [HGNC:24766]
FTL	-0.944871	1.41303E-30	ferritin light chain [HGNC:3999]
DTNA	-0.944723	5.56523E-08	dystrobrevin alpha [HGNC:3057]
B3GALT6	-0.943665	7.69034E-28	beta-1,3-galactosyltransferase 6 [HGNC:17978]
IFITM3	-0.943589	1.62581E-08	interferon induced transmembrane protein 3 [HGNC:5414]
LIMCH1	-0.937018	1.31066E-24	LIM and calponin homology domains 1 [HGNC:29191]
PROX1-AS1	-0.936937	0.017926011	PROX1 antisense RNA 1 [HGNC:43656]
SPINK1	-0.935341	0.040255142	serine peptidase inhibitor, Kazal type 1 [HGNC:11244]
ZNF395	-0.935074	1.97798E-21	zinc finger protein 395 [HGNC:18737]
COL5A2	-0.931236	7.11722E-07	collagen type V alpha 2 chain [HGNC:2210]
MLXIP	-0.930677	2.22677E-32	MLX interacting protein [HGNC:17055]
RCAN2	-0.927801	0.001474695	regulator of calcineurin 2 [HGNC:3041]
TMEM155	-0.927729	0.043320905	transmembrane protein 155 [HGNC:26418]
B3GNT3	-0.926449	2.98855E-05	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 3 [HGNC:13528]
GNGT1	-0.923322	0.033058336	G protein subunit gamma transducin 1 [HGNC:4411]
LYPLA2	-0.922946	1.49059E-31	lysophospholipase II [HGNC:6738]
PLEKHA4	-0.920878	0.006283993	pleckstrin homology domain containing A4 [HGNC:14339]
AC089983.1	-0.920808	9.43555E-06	novel transcript, antisense to TXNRD1
SPINK4	-0.920788	0.001442218	serine peptidase inhibitor, Kazal type 4 [HGNC:16646]
ZNF793	-0.916891	0.000143471	zinc finger protein 793 [HGNC:33115]
KBTBD7	-0.910479	5.89647E-06	kelch repeat and BTB domain containing 7 [HGNC:25266]
LIFR	-0.904275	1.01756E-22	LIF receptor alpha [HGNC:6597]
ATAD3C	-0.898678	0.006217473	ATPase family, AAA domain containing 3C [HGNC:32151]

NUPR1	-0.897762	4.65981E-12	nuclear protein 1, transcriptional regulator [HGNC:29990]
HYAL1	-0.892615	0.010037522	hyaluronoglucosaminidase 1 [HGNC:5320]
CCT2	-0.892557	5.39419E-50	chaperonin containing TCP1 subunit 2 [HGNC:1615]
AL354718.3	-0.887788	0.00357375	CMT1A duplicated region transcript 15 (CDRT15) pseudogene
AC027117.1	-0.885820	0.004462849	novel transcript
TRIM2	-0.884654	0.000395695	tripartite motif containing 2 [HGNC:15974]
PSD3	-0.882241	2.10326E-22	pleckstrin and Sec7 domain containing 3 [HGNC:19093]
AC096921.2	-0.881940	0.023795057	novel transcript, overlapping to TGFBR2
ERV3-1	-0.878386	1.32508E-05	endogenous retrovirus group 3 member 1, envelope [HGNC:3454]
NELL2	-0.877412	0.004779858	neural EGFL like 2 [HGNC:7751]
PLPP3	-0.876255	5.49591E-11	phospholipid phosphatase 3 [HGNC:9229]
HK1	-0.875622	1.8843E-34	hexokinase 1 [HGNC:4922]
SELENOP	-0.869637	6.14674E-05	selenoprotein P [HGNC:10751]
UGCG	-0.867247	2.96293E-43	UDP-glucose ceramide glucosyltransferase [HGNC:12524]
CFI	-0.866346	0.044778326	complement factor I [HGNC:5394]
PTP4A3	-0.865869	5.29684E-12	protein tyrosine phosphatase type IVA, member 3 [HGNC:9636]
ZNF486	-0.861596	0.014190445	zinc finger protein 486 [HGNC:20807]
FGL1	-0.857728	0.021098677	fibrinogen like 1 [HGNC:3695]
ID3	-0.857338	6.31616E-18	inhibitor of DNA binding 3, HLH protein [HGNC:5362]
TNFSF13	-0.856961	0.016718945	TNF superfamily member 13 [HGNC:11928]
CA12	-0.848187	4.35561E-05	carbonic anhydrase 12 [HGNC:1371]
AC007325.4	-0.848179	0.004960606	protein DGCR6 [Source:NCBI gene;Acc:102724770]
HMGCS1	-0.847504	2.49807E-45	3-hydroxy-3-methylglutaryl-CoA synthase 1 [HGNC:5007]
TREML3P	-0.846621	0.000230028	triggering receptor expressed on myeloid cells like 3, pseudogene [HGNC:30806]
LPCAT1	-0.843497	5.4543E-06	lysophosphatidylcholine acyltransferase 1 [HGNC:25718]
RNF144B	-0.842953	3.33194E-07	ring finger protein 144B [HGNC:21578]
ZNF616	-0.842547	0.003459179	zinc finger protein 616 [HGNC:28062]
ANK3	-0.841293	2.3066E-06	ankyrin 3 [HGNC:494]
FRG1JP	-0.836826	0.029875157	FSHD region gene 1 family member J, pseudogene [HGNC:51768]
DEPTOR	-0.830913	2.88694E-08	DEP domain containing MTOR interacting protein [HGNC:22953]
WIF1	-0.829316	0.043553096	WNT inhibitory factor 1 [HGNC:18081]
ZCCHC12	-0.827113	0.047804007	zinc finger CCHC-type containing 12 [HGNC:27273]
BMP8B	-0.825225	4.63636E-11	bone morphogenetic protein 8b [HGNC:1075]
PSAP	-0.821262	4.9425E-58	prosaposin [HGNC:9498]
TRIM37	-0.820816	4.41634E-26	tripartite motif containing 37 [HGNC:7523]
TMC4	-0.819810	2.458E-08	transmembrane channel like 4 [HGNC:22998]
SPATA6	-0.818069	0.001187914	spermatogenesis associated 6 [HGNC:18309]
CLDN7	-0.817946	2.49524E-17	claudin 7 [HGNC:2049]
CTSO	-0.816274	0.00057685	cathepsin O [HGNC:2542]
INAFM1	-0.814608	8.03987E-05	InaF motif containing 1 [HGNC:27406]
PRAG1	-0.813525	1.04467E-13	PEAK1 related, kinase-activating pseudokinase 1 [HGNC:25438]
ATL3	-0.813261	2.62032E-41	atlastin GTPase 3 [HGNC:24526]
CEACAM19	-0.807146	0.000100687	carcinoembryonic antigen related cell adhesion molecule 19 [HGNC:31951]
FAM86DP	-0.806124	0.000784007	family with sequence similarity 86 member D, pseudogene [HGNC:32659]
MEF2C	-0.805572	0.029749181	myocyte enhancer factor 2C [HGNC:6996]

AHR	-0.802358	4.21959E-13	aryl hydrocarbon receptor [HGNC:348]
CUL2	-0.802300	7.51105E-16	cullin 2 [HGNC:2552]
ZNF117	-0.797772	0.000264971	zinc finger protein 117 [HGNC:12897]
SLC16A5	-0.795069	5.31931E-10	solute carrier family 16 member 5 [HGNC:10926]
ACOX2	-0.794227	0.018442653	acyl-CoA oxidase 2 [HGNC:120]
CEACAM6	-0.793768	0.000522755	carcinoembryonic antigen related cell adhesion molecule 6 [HGNC:1818]
PIP4P2	-0.788364	5.03393E-05	phosphatidylinositol-4,5-bisphosphate 4-phosphatase 2 [HGNC:25452]
MAP2	-0.788097	8.36972E-13	microtubule associated protein 2 [HGNC:6839]
SLC16A9	-0.784975	6.31129E-05	solute carrier family 16 member 9 [HGNC:23520]
MOCS1	-0.784645	0.017468849	molybdenum cofactor synthesis 1 [HGNC:7190]
CXorf57	-0.781124	0.000585033	chromosome X open reading frame 57 [HGNC:25486]
SPRY1	-0.779081	0.000643652	sprouty RTK signaling antagonist 1 [HGNC:11269]
OGA	-0.775024	0.0012402	O-GlcNAcase [HGNC:7056]
GAS5	-0.772276	1.55352E-19	growth arrest specific 5 [HGNC:16355]
IGIP	-0.770830	0.001705735	IgA inducing protein [HGNC:33847]
MAN1A1	-0.770191	6.68774E-05	mannosidase alpha class 1A member 1 [HGNC:6821]
AL049840.4	-0.770122	0.000351969	novel transcript, sense intronic to KLC1
ETV4	-0.768646	2.58726E-07	ETS variant 4 [HGNC:3493]
HACD4	-0.768415	0.019628631	3-hydroxyacyl-CoA dehydratase 4 [HGNC:20920]
FLNC	-0.767903	4.17551E-30	filamin C [HGNC:3756]
WDR78	-0.766413	0.021864486	WD repeat domain 78 [HGNC:26252]
AL391422.4	-0.764530	0.000247107	novel transcript, antisense to PXDC1
HIPK3	-0.763632	1.34895E-28	homeodomain interacting protein kinase 3 [HGNC:4915]
TMEM173	-0.758323	3.81661E-10	transmembrane protein 173 [HGNC:27962]
LCN2	-0.755498	0.03149551	lipocalin 2 [HGNC:6526]
RPS16	-0.752794	1.3806E-30	ribosomal protein S16 [HGNC:10396]
TIAM1	-0.752623	6.27811E-08	T cell lymphoma invasion and metastasis 1 [HGNC:11805]
CSF1	-0.752463	6.86238E-30	colony stimulating factor 1 [HGNC:2432]
VWA5A	-0.751655	5.41739E-05	von Willebrand factor A domain containing 5A [HGNC:6658]
FGFBP1	-0.751446	0.003331202	fibroblast growth factor binding protein 1 [HGNC:19695]
SPDEF	-0.746975	0.009608634	SAM pointed domain containing ETS transcription factor [HGNC:17257]
PAG1	-0.744983	0.002731053	phosphoprotein membrane anchor with glycosphingolipid microdomains 1 [HGNC:30043]
TRPS1	-0.742377	5.15966E-05	transcriptional repressor GATA binding 1 [HGNC:12340]
NDRG4	-0.742291	3.92059E-06	NDRG family member 4 [HGNC:14466]
FZD3	-0.739147	4.99473E-09	frizzled class receptor 3 [HGNC:4041]
LINC02331	-0.738154	0.032974424	long intergenic non-protein coding RNA 2331 [HGNC:53251]
NEDD4	-0.738022	4.13792E-34	E3 ubiquitin protein ligase [HGNC:7727]
MBP	-0.737919	7.00153E-22	myelin basic protein [HGNC:6925]
P2RY6	-0.737546	0.001392074	pyrimidinergic receptor P2Y6 [HGNC:8543]
ZMIZ1	-0.736295	1.00075E-23	zinc finger MIZ-type containing 1 [HGNC:16493]
TTPA	-0.735648	0.023452724	alpha tocopherol transfer protein [HGNC:12404]
PXK	-0.735396	1.81223E-07	PX domain containing serine/threonine kinase like [HGNC:23326]
TNFRSF18	-0.732986	4.43828E-05	TNF receptor superfamily member 18 [HGNC:11914]
CYP4V2	-0.732765	3.72835E-06	cytochrome P450 family 4 subfamily V member 2 [HGNC:23198]
CAVIN2	-0.730863	6.32717E-05	caveolae associated protein 2 [HGNC:10690]

SFN	-0.727700	2.93044E-24	stratifin [HGNC:10773]
RPS6KA5	-0.726167	1.02176E-07	ribosomal protein S6 kinase A5 [HGNC:10434]
INPP5B	-0.724520	6.41911E-12	inositol polyphosphate-5-phosphatase B [HGNC:6077]
RNF128	-0.718938	0.020938098	ring finger protein 128, E3 ubiquitin protein ligase [HGNC:21153]
GBE1	-0.718540	2.72436E-15	1,4-alpha-glucan branching enzyme 1 [HGNC:4180]
CTHRC1	-0.717807	1.43798E-14	collagen triple helix repeat containing 1 [HGNC:18831]
SLC23A2	-0.717351	2.59343E-10	solute carrier family 23 member 2 [HGNC:10973]
C1orf116	-0.717130	4.55758E-14	chromosome 1 open reading frame 116 [HGNC:28667]
TCF3	-0.716523	6.81953E-36	transcription factor 3 [HGNC:11633]
EYA3	-0.713673	3.21478E-07	EYA transcriptional coactivator and phosphatase 3 [HGNC:3521]
MMP15	-0.713554	1.74089E-17	matrix metallopeptidase 15 [HGNC:7161]
CTH	-0.713126	1.79823E-08	cystathionine gamma-lyase [HGNC:2501]
CENPM	-0.712268	0.00055373	centromere protein M [HGNC:18352]
PADI2	-0.710903	5.84343E-13	peptidyl arginine deiminase 2 [HGNC:18341]
AC144530.1	-0.710636	0.014438199	ribosomal protein L17 (RPL17) pseudogene
MAPRE1	-0.709968	1.45853E-18	microtubule associated protein RP/EB family member 1 [HGNC:6890]
LHPP	-0.708764	0.000101816	phospholysine phosphohistidine inorganic pyrophosphate phosphatase [HGNC:30042]
FLG	-0.705571	0.005746043	filaggrin [HGNC:3748]
TSPAN7	-0.705258	3.65238E-05	tetraspanin 7 [HGNC:11854]
MCAM	-0.699520	3.531E-08	melanoma cell adhesion molecule [HGNC:6934]
ZNF800	-0.699141	8.74369E-08	zinc finger protein 800 [HGNC:27267]
ANKRD33B	-0.694256	2.48764E-15	ankyrin repeat domain 33B [HGNC:35240]
DKK3	-0.694144	0.008667108	dickkopf WNT signaling pathway inhibitor 3 [HGNC:2893]
CCDC80	-0.689801	0.000305673	coiled-coil domain containing 80 [HGNC:30649]
IPO5P1	-0.688501	0.037241929	importin 5 pseudogene 1 [HGNC:49687]
ALDH1L2	-0.686257	9.96266E-09	aldehyde dehydrogenase 1 family member L2 [HGNC:26777]
ABHD8	-0.686167	2.09878E-07	abhydrolase domain containing 8 [HGNC:23759]
TOLLIP	-0.682880	1.72655E-12	toll interacting protein [HGNC:16476]
SH2B3	-0.681214	2.8485E-11	SH2B adaptor protein 3 [HGNC:29605]
PQLC1	-0.680739	4.43354E-06	PQ loop repeat containing 1 [HGNC:26188]
ROR1	-0.680022	4.82852E-08	receptor tyrosine kinase like orphan receptor 1 [HGNC:10256]
UGT8	-0.679857	3.66486E-08	UDP glycosyltransferase 8 [HGNC:12555]
KLB	-0.679634	1.47264E-22	klotho beta [HGNC:15527]
ZMPSTE24	-0.675975	1.28267E-16	zinc metallopeptidase STE24 [HGNC:12877]
SLC25A41	-0.675905	0.0216945	solute carrier family 25 member 41 [HGNC:28533]
AKR1C2	-0.675252	2.76998E-17	aldo-keto reductase family 1 member C2 [HGNC:385]
APOE	-0.673689	6.59126E-08	apolipoprotein E [HGNC:613]
PELI1	-0.673583	3.05455E-15	pellino E3 ubiquitin protein ligase 1 [HGNC:8827]
SKI	-0.672595	7.96041E-18	SKI proto-oncogene [HGNC:10896]
ZNF790	-0.672425	0.042949259	zinc finger protein 790 [HGNC:33114]
AKR1B10	-0.671343	0.000210081	aldo-keto reductase family 1 member B10 [HGNC:382]
BMPR1B	-0.670923	0.000148154	bone morphogenetic protein receptor type 1B [HGNC:1077]
SUMO3	-0.669836	7.53706E-28	small ubiquitin-like modifier 3 [HGNC:11124]
CDK4	-0.667791	7.44974E-15	cyclin dependent kinase 4 [HGNC:1773]
PLAC9	-0.667055	0.024619334	placenta specific 9 [HGNC:19255]

FOXL2NB	-0.665767	0.020730088	FOXL2 neighbor [HGNC:34428]
EFEMP1	-0.665483	1.08791E-12	EGF containing fibulin extracellular matrix protein 1 [HGNC:3218]
NRN1	-0.665253	0.041207035	neuritin 1 [HGNC:17972]
HLA-DQB1	-0.665020	0.005795108	major histocompatibility complex, class II, DQ beta 1 [HGNC:4944]
GNAL	-0.664791	4.61694E-09	G protein subunit alpha L [HGNC:4388]
AC007388.1	-0.663647	0.011300639	uncharacterized LOC728730 [Source:NCBI gene;Acc:728730]
CYBA	-0.663410	3.56437E-09	cytochrome b-245 alpha chain [HGNC:2577]
H6PD	-0.663080	2.80954E-17	hexose-6-phosphate dehydrogenase/glucose 1-dehydrogenase [HGNC:4795]
ZNF585A	-0.662088	0.035972674	zinc finger protein 585A [HGNC:26305]
FKBP10	-0.661969	1.5118E-07	FK506 binding protein 10 [HGNC:18169]
PRICKLE1	-0.661060	0.000138199	prickle planar cell polarity protein 1 [HGNC:17019]
ESPN	-0.659749	0.001511522	espin [HGNC:13281]
SLC34A2	-0.659209	0.002817798	solute carrier family 34 member 2 [HGNC:11020]
PGM5	-0.659028	0.019204412	phosphoglucomutase 5 [HGNC:8908]
SPOCK2	-0.657054	0.006502172	SPARC (osteonectin), cwcv and kazal like domains proteoglycan 2 [HGNC:13564]
MAB21L4	-0.654929	0.007682512	mab-21 like 4 [HGNC:26216]
MICAL2	-0.654209	1.26113E-10	microtubule associated monooxygenase, calponin and LIM domain containing 2 [HGNC:24693]
SH3GL2	-0.653398	0.000355896	SH3 domain containing GRB2 like 2, endophilin A1 [HGNC:10831]
RNF212	-0.653189	0.000110729	ring finger protein 212 [HGNC:27729]
ZNF430	-0.652643	0.010054324	zinc finger protein 430 [HGNC:20808]
BRD9	-0.651680	1.88639E-12	bromodomain containing 9 [HGNC:25818]
SMIM13	-0.651563	1.02956E-09	small integral membrane protein 13 [HGNC:27356]
ELMO3	-0.649773	0.000875189	engulfment and cell motility 3 [HGNC:17289]
PTPN20	-0.649348	0.009839742	protein tyrosine phosphatase, non-receptor type 20 [HGNC:23423]
NKX2-1	-0.649287	6.67648E-18	NK2 homeobox 1 [HGNC:11825]
TMEM87B	-0.649201	3.27071E-09	transmembrane protein 87B [HGNC:25913]
C6orf48	-0.648968	2.07552E-14	chromosome 6 open reading frame 48 [HGNC:19078]
PPARD	-0.648751	4.39018E-12	peroxisome proliferator activated receptor delta [HGNC:9235]
FKBP7	-0.648564	0.021473256	FK506 binding protein 7 [HGNC:3723]
TNFRSF10A-AS1	-0.647314	0.018037937	TNFRSF10A antisense RNA 1 [HGNC:53164]
EIF3D	-0.645720	5.47042E-18	eukaryotic translation initiation factor 3 subunit D [HGNC:3278]
HOXB9	-0.645553	7.3573E-06	homeobox B9 [HGNC:5120]
NEK6	-0.644516	5.32877E-07	NIMA related kinase 6 [HGNC:7749]
PACS1	-0.644379	3.06963E-18	phosphofurin acidic cluster sorting protein 1 [HGNC:30032]
NAP1L2	-0.643508	0.029435664	nucleosome assembly protein 1 like 2 [HGNC:7638]
TENT4A	-0.642708	9.44768E-16	terminal nucleotidyltransferase 4A [HGNC:16705]
HLA-DRB1	-0.641585	0.019740378	major histocompatibility complex, class II, DR beta 1 [HGNC:4948]
RAB27B	-0.640306	0.00071326	RAB27B, member RAS oncogene family [HGNC:9767]
IGF2BP1	-0.640129	0.010898877	insulin like growth factor 2 mRNA binding protein 1 [HGNC:28866]
PSAT1	-0.636819	1.3804E-13	phosphoserine aminotransferase 1 [HGNC:19129]
ZNF516	-0.635727	1.0448E-07	zinc finger protein 516 [HGNC:28990]
HCP5	-0.635463	7.94152E-06	HLA complex P5 [HGNC:21659]
NOL11	-0.634747	1.58431E-15	nucleolar protein 11 [HGNC:24557]
ZC3H6	-0.634448	0.000667682	zinc finger CCCH-type containing 6 [HGNC:24762]
AC018645.2	-0.634350	0.002836792	novel transcript

NAV2	-0.632556	0.010733076	neuron navigator 2 [HGNC:15997]
SLC2A4RG	-0.630197	6.19216E-13	SLC2A4 regulator [HGNC:15930]
LRRC8A	-0.629783	8.50986E-22	leucine rich repeat containing 8 VRAC subunit A [HGNC:19027]
SERPINB8	-0.629087	0.001960924	serpin family B member 8 [HGNC:8952]
LINC02274	-0.627829	0.003484645	long intergenic non-protein coding RNA 2274 [HGNC:53190]
CLYBL	-0.627538	0.017504738	citrate lyase beta like [HGNC:18355]
FAR2P2	-0.626215	0.041160307	fatty acyl-CoA reductase 2 pseudogene 2 [HGNC:49279]
RNF145	-0.625439	2.4867E-17	ring finger protein 145 [HGNC:20853]
TGFBR2	-0.625331	2.48325E-16	transforming growth factor beta receptor 2 [HGNC:11773]
SLFN5	-0.621588	3.65868E-11	schlafen family member 5 [HGNC:28286]
ANKRD29	-0.619760	9.152E-07	ankyrin repeat domain 29 [HGNC:27110]
PHETA2	-0.618862	0.001627032	PH domain containing endocytic trafficking adaptor 2 [HGNC:27161]
ZNF551	-0.617660	0.009052005	zinc finger protein 551 [HGNC:25108]
SMIM29	-0.616965	0.004272201	small integral membrane protein 29 [HGNC:1340]
ETV5	-0.616677	1.11409E-14	ETS variant 5 [HGNC:3494]
ELOVL6	-0.615694	2.87568E-15	ELOVL fatty acid elongase 6 [HGNC:15829]
SLC39A4	-0.615422	9.11343E-06	solute carrier family 39 member 4 [HGNC:17129]
HOXA-AS2	-0.614988	0.000175101	HOXA cluster antisense RNA 2 [HGNC:43745]
ZNF680	-0.614547	0.000447451	zinc finger protein 680 [HGNC:26897]
SLC27A1	-0.614174	8.54387E-06	solute carrier family 27 member 1 [HGNC:10995]
MAML2	-0.613813	0.027153549	mastermind like transcriptional coactivator 2 [HGNC:16259]
TSC22D4	-0.613726	1.16734E-07	TSC22 domain family member 4 [HGNC:21696]
AL157935.1	-0.612260	0.01518575	novel transcript
NAB2	-0.611673	0.000289648	NGFI-A binding protein 2 [HGNC:7627]
CARD6	-0.611207	0.006742198	caspase recruitment domain family member 6 [HGNC:16394]
CDC25A	-0.610850	0.001991148	cell division cycle 25A [HGNC:1725]
DNMBP	-0.610819	3.88582E-12	dynamin binding protein [HGNC:30373]
PTGR1	-0.608029	5.36662E-30	prostaglandin reductase 1 [HGNC:18429]
CALHM2	-0.607510	0.03192465	calcium homeostasis modulator family member 2 [HGNC:23493]
SLC16A3	-0.606806	7.50108E-24	solute carrier family 16 member 3 [HGNC:10924]
RAB11B	-0.606703	2.68469E-22	RAB11B, member RAS oncogene family [HGNC:9761]
ABCA13	-0.605970	0.020120054	ATP binding cassette subfamily A member 13 [HGNC:14638]
LYN	-0.605285	1.43494E-21	LYN proto-oncogene, Src family tyrosine kinase [HGNC:6735]
MAP3K1	-0.605283	5.10422E-09	mitogen-activated protein kinase kinase kinase 1 [HGNC:6848]
MTMR12	-0.604150	1.14587E-16	myotubularin related protein 12 [HGNC:18191]
RABL3	-0.602819	4.65862E-05	RAB, member of RAS oncogene family like 3 [HGNC:18072]
CADPS2	-0.602613	7.86017E-06	calcium dependent secretion activator 2 [HGNC:16018]
MEGF9	-0.601150	3.15189E-11	multiple EGF like domains 9 [HGNC:3234]
LINC01291	-0.599806	0.008174803	long intergenic non-protein coding RNA 1291 [HGNC:50358]
SSH1	-0.599764	1.184E-17	slingshot protein phosphatase 1 [HGNC:30579]
RPS3AP26	-0.599669	0.045568855	ribosomal protein S3a pseudogene 26 [HGNC:36513]
NMD3	-0.599622	3.01976E-14	NMD3 ribosome export adaptor [HGNC:24250]
BCL7A	-0.598606	2.2546E-10	BCL tumor suppressor 7A [HGNC:1004]
MPP7	-0.597266	6.67911E-05	membrane palmitoylated protein 7 [HGNC:26542]
RPS12	-0.597071	1.40736E-13	ribosomal protein S12 [HGNC:10385]

PPP4R1	-0.594999	3.11332E-18	protein phosphatase 4 regulatory subunit 1 [HGNC:9320]
HSH2D	-0.594783	0.029061807	hematopoietic SH2 domain containing [HGNC:24920]
AC113935.1	-0.593825	0.034039627	ribosomal protein L17 (RPL17) pseudogene
SGMS2	-0.591512	1.20736E-07	sphingomyelin synthase 2 [HGNC:28395]
SMAD4	-0.591274	1.10307E-09	SMAD family member 4 [HGNC:6770]
RNF125	-0.591213	0.000439436	ring finger protein 125 [HGNC:21150]
IL15	-0.589641	0.001708081	interleukin 15 [HGNC:5977]
CYP51A1	-0.589060	0.002938588	cytochrome P450 family 51 subfamily A member 1 [HGNC:2649]
SMPDL3B	-0.588470	7.25458E-06	sphingomyelin phosphodiesterase acid like 3B [HGNC:21416]
PKIA	-0.586459	0.001944807	cAMP-dependent protein kinase inhibitor alpha [HGNC:9017]
MYO10	-0.584775	2.19192E-12	myosin X [HGNC:7593]
TFPI2	-0.584094	8.76567E-10	tissue factor pathway inhibitor 2 [HGNC:11761]
ARMCX2	-0.583713	0.000151252	armadillo repeat containing X-linked 2 [HGNC:16869]
GRB7	-0.583568	1.78834E-07	growth factor receptor bound protein 7 [HGNC:4567]
RPL37	-0.582875	3.70134E-17	ribosomal protein L37 [HGNC:10347]
SOD2	-0.581364	5.85554E-15	superoxide dismutase 2 [HGNC:11180]
RPL3P4	-0.580409	1.04038E-08	ribosomal protein L3 pseudogene 4 [HGNC:19805]
DSE	-0.580322	3.57299E-11	dermatan sulfate epimerase [HGNC:21144]
PARL	-0.580129	0.014083241	presenilin associated rhomboid like [HGNC:18253]

Table 3.10 - GO analysis of differentially gene expression in MDA-MB-231 shRNA4 SPAG5. Tables shows the enrichment genes differentially upregulated and downregulated in MDA-MB-231 shRNA4 SPAG5

Category	Term	Description	LogP	In	Symbols
GO Biological Processes	GO:0003013	circulatory system process	14.17216	56/497	ABCA2,ADRA2C,ANK2,APOE,ATP1A1,AZU1,CEACAM1,BMP6,C3AR1,CAMK2D,COL1A2,CYBA,DBH,FAAH,GAA,GJA1,GPER1,GSN,SLC29A2 ,HTR1D,ID2,KCNQ1,LEPR,LRP1,LRP3,NOS3,OPRL1,P2RX1,MAP2K6,PTGS1,SCN1B,SCN4B, SCNN1A,SLC1A1,SLC4A3,SLC2A3,SLC2A5,SLC6A9,SLC16A2,SLC22A3,SOD3,TGFB1,ABCC3,KCNK6,ABCG2,PTP4A3, TJP3,DLL1,TRPM4,ADM2,SLC2A10,VSTM4,SLC29A4,TAC4,SLC27A1,LOC102723475
GO Biological Processes	GO:0035239	tube morphogenesis	13.95735	66/663	APOE,RHOB,CEACAM1,CALD1,COL4A1,COL8A2,CSF1,CSF1R,CSPG4,EFNA1,MEGF8,EPHA4,FN1,FOLR1,GJA1,HLX,HSPG2,ID2, IHH,ITGB3,LEPR,LOX,LRP1,TACSTD2,MSX2,MTHFR,MYH9,NOS3,PBX1,PDGFRB,SLC1A1,SPINT1, TBX1,TGFB1,TGFBI,TSC1,SCG2,TNFSF12,NRP2,RAMP1,VASH1,COBL,ZFPM2,CDK20,KLHL3,DLL1,ANGPTL4,EGFL7,KIF26B,LGR4, MYDGF,ADAMTS9,ACKR3,PLXDC1,RHOJ,HHIP,IRX3,TCTN1,ADM2,ATOH8,PRICKLE1,DAB2IP,VSTM4,UNC5B,LAMA1,YJEFN3
GO Biological Processes	GO:0008610	lipid biosynthetic process	11.69121	56/570	ABCA2,ACACA,ACLY,ALDH1A3,ALDH3B1,BMX,DHCR7,DHCR24,ACSL1,FASN,GPX4,GSTM2,HMGCR,HMGCS1,ITPKB, LSS,MVD,MVK,P2RX1,PBX1,PRLR,PTGS1,PCYT2,RBP1,MSMO1,SCD,ST3GAL1,ST8SIA1,PLPP2,ST3GAL5, LIPG,FADS2,GGPS1,DOLK,LPIN1,GBGT1,SH3YL1,ST6GALNAC4,PIGP,ACSL5,BCO1,DPM3,LPCAT2,THNSL2,PIGV,ACSS2,AGPAT4,CERS4, PNPLA3,ACSS1,ABHD1,CBR4,OSBPL6,TMEM150A,PIP5KL1,SLC27A1
GO Biological Processes	GO:0001568	blood vessel development	11.66895	52/505	APOE,RHOB,CEACAM1,CALD1,COL1A2,COL4A1,COL5A1,COL8A2,CSPG4,EFNA1,MEGF8,FN1,FOLR1,GJA1,HSPG2,IHH, ITGB3,LEPR,LOX,LRP1,SMAD6,MYH9,NOS3,PDGFRB,SLC1A1,SPINT1,TBX1,TGFB1,TGFBI,SCG2, TNFSF12,NRP2,RAMP1,SEMA3C,VASH1,ZFPM2,DLL1,ANGPTL4,EGFL7,MYDGF,ADAMTS9,ACKR3,PLXDC1,RHOJ, ADM2,SLC2A10,PRICKLE1,DAB2IP,VSTM4,UNC5B,LAMA1,YJEFN3
GO Biological Processes	GO:0001944	vasculature development	11.51651	53/526	APOE,RHOB,CEACAM1,CALD1,COL1A2,COL4A1,COL5A1,COL8A2,CSPG4,EFNA1,MEGF8,FN1,FOLR1,GJA1,HSPG2, IHH,ITGB3,LEPR,LOX,LRP1,SMAD6,MYH9,NDP,NOS3,PDGFRB,SLC1A1,SPINT1,TBX1,TGFB1, TGFBI,SCG2,TNFSF12,NRP2,RAMP1,SEMA3C,VASH1,ZFPM2,DLL1,ANGPTL4,EGFL7,MYDGF,ADAMTS9,ACKR3,PLXDC1, RHOJ,ADM2,SLC2A10,PRICKLE1,DAB2IP,VSTM4,UNC5B,LAMA1,YJEFN3
GO Biological Processes	GO:0045596	negative regulation of cell differentiation	10.50431	61/699	ABCA1,NKX32,CEACAM1,CDK5,COL5A1,EFNA1,EFNB3,EPHA4,EFEMP1,GPER1,HLX,HPN,ID2,IGFBP5,IHH,ITGB3 ,ITPKB,LRP3,LTBP3,TACSTD2,SMAD6,MAP2,MELTF,MMP11,MSX2,NFATC1,PBX1,RAP1GAP,RGS4, SEMA3F,SNAI1,TBX1,TGFB1,GPR137B,ZFP36,SEMA3B,NR1D1,ABCG1,MAFB,NR1H3,TRIB1,SEMA3A,SEMA3C,TMEM131L,ZFPM2,BAMBI, DLL1,TRPM4,SEMA4G,TRIB3,RTN4R,IRX3,DIXDC1,VASN,JDP2,PRICKLE1,DAB2IP,BBS12,SEMA3D,RFLNB,DRAXIN
GO Biological Processes	GO:0048514	blood vessel morphogenesis	10.17845	44/421	APOE,RHOB,CEACAM1,CALD1,COL4A1,COL8A2,CSPG4,EFNA1,FN1,FOLR1,GJA1,HSPG2,IHH,ITGB3,LEPR, LOX,LRP1,MYH9,NOS3,PDGFRB,SLC1A1,TBX1,TGFB1,TGFBI,SCG2,TNFSF12,NRP2,RAMP1,VASH1,ZFPM2, DLL1,ANGPTL4,EGFL7,MYDGF,ADAMTS9,ACKR3,PLXDC1,RHOJ,ADM2,DAB2IP,VSTM4,UNC5B,LAMA1,YJEFN3
GO Biological Processes	GO:0003018	vascular process in circulatory system	9.825761	33/263	ABCA2,ADRA2C,APOE,AZU1,CEACAM1,BMP6,DBH,FAAH,GPER1,SLC29A2,HTR1D,LEPR,LRP1,LRP3 ,NOS3,P2RX1,SLC1A1,SLC4A3,SLC2A3,SLC6A9,SLC16A2,SLC22A3,SOD3,TGFB1,ABCC3,ABCG2,PTP4A3, TJP3,TRPM4,SLC2A10,VSTM4,SLC29A4,SLC27A1
GO Biological Processes	GO:0006066	alcohol metabolic process	9.466771	36/316	ABCA1,ABCA2,ACLY,ALDH1A3,ALDH3B1,APOE,CEBPA,DBH,DHCR7,DHCR24,HMGCR,HMGCS1, IDH1,ITPKB,LEPR,LSS,MVD,MVK,RXRA,MSMO1,VLDLR,PNPLA4,APOL1,SCARF1,PLPP2,DHRS3,ABCG1,TKFC, DHRS7,BCO1,RETSAT,ACSS2,ACSS1,CBR4,PCSK9,GDPD1
GO Biological Processes	GO:0098609	cellcell adhesion	9.198131	49/537	ARVCF,CEACAM1,CD33,CDH11,COL8A2,CELSR2,FN1,GCNT1,GP1BA,IHH,ITGAX,ITGB3,MYH9,NCAM2,NINJ2 ,PCDH9,PTPRM,CCL5,COL14A1,GPC6,CLDN15,TJP3,PODXL2,F11R,CERCAM,PCDH18,PCDHGB2 ,PCDHGA4,PCDHB14,PCDHB13,PCDHB12,PCDHB11,PCDHB10,PCDHB9,PCDHB8,PCDHB4,CRTAM,CEACAM19,CAMSAP3, PCDHB16,LRFN4,HMCN1,OBSCN,NTNG2,SCARF2,BOC,VSTM2L,EMB,SIRPA

Category	Term	Description	LogP	InTerm_InList	Symbols
GO Cellular Components	GO:0031012	extracellular matrix	24.82418	78/571	ANXA4,APOE,AZGP1,CLU,COL1A2,COL4A1,COL4A5,COL5A1,COL6A2,COL6A3,COL8A2, COL9A2,CSPG4,CTSD,EFEMP1,FN1,GP1BA,HSPG2,IGFBP7,IHH,ITGB4,ANOS1, LAMA4,LGALS3BP,LOX,LTBP3,LUM,MATN2,ADAM11,MGP,MMP11,MMP16,NDP, SERPINA5,SERPINA1,SERPINE2,PRELP,HTRA1,RARRES2,SOD3,TGFB1,TGFBI,THBS3, TIMP1,TIMP4,COL14A1,VTN,SEMA3B,SRPX,CCN5,SDC3,GPC6,EDIL3,SPON2,SPON1,ADAMTS13,SSPOP, OLFML2B,SNED1, ANGPTL4,EGFL7,SCARA3,ADAMTS9,VWA1,FRAS1,HMCN1, LOXL4,NTNG2,FIBCD1,PAPLN, ELFN2,VASN,LRRC15,RTN4RL1,COL24A1,LAMA1,LRRTM3,ELFN1
GO Cellular Components	GO:0030312	external encapsulating structure	24.77453	78/572	ANXA4,APOE,AZGP1,CLU,COL1A2,COL4A1,COL4A5, COL5A1,COL6A2,COL6A3,COL8A2,COL9A2,CSPG4, CTSD,EFEMP1,FN1,GP1BA,HSPG2,IGFBP7,IHH,ITGB4,ANOS1,LAMA4 ,LGALS3BP,LOX,LTBP3,LUM,MATN2,ADAM11,MGP,MMP11,MMP16,NDP,SERPINA5,SERPINA1, SERPINE2,PRELP,HTRA1, RARRES2,SOD3,TGFB1,TGFBI,THBS3,TIMP1, TIMP4,COL14A1,VTN, SEMA3B,SRPX,CCN5,SDC3,GPC6,EDIL3,SPON2, SPON1,ADAMTS13,SSPOP,OLFML2B,SNED1,ANGPTL4,EGFL7,SCARA3,ADAMTS9,VWA1,FRAS1, HMCN1,LOXL4,NTNG2,FIBCD1,PAPLN,ELFN2,VASN,LRRC15,RTN4RL1,COL24A1,LAMA1,LRRTM3,ELFN1
GO Cellular Components	GO:0062023	collagencontaining extracellular matrix	19.692	60/429	ANXA4,APOE,AZGP1,CLU,COL1A2,COL4A1,COL4A5,COL5A1, COL6A2,COL6A3,COL8A2,COL9A2,CSPG4,CTSD,EFEMP1,FN1, HSPG2,IGFBP7,ITGB4,LAMA4,LGALS3BP,LOX,LTBP3,LUM,MATN2,ADAM11,MGP,NDP,SERPINA5,SERPINA1,SERPINE2,PRELP,HTRA1, RARRES2,SOD3,TGFB1,TGFBI,THBS3,TIMP1,COL14A1, VTN,SEMA3B,SRPX,SDC3,GPC6,EDIL3,SPON1, ANGPTL4,EGFL7,SCARA3,ADAMTS9,VWA1,FRAS1,HMCN1,LOXL4,NTNG2,FIBCD1,LRRC15,COL24A1,LAMA1
GO Cellular Components	GO:0005788	endoplasmic reticulum lumen	9.660067	36/311	APOE,ARSA,CLU,COL1A2,COL4A1,COL4A5,COL5A1,COL6A2, COL6A3,COL8A2,COL9A2,CSF1,FN1,GPX7,IGFBP5,IGFBP7,MELTF, SERPINA1,PPIB,TIMP1,COL14A1,SCG2,APOL1,SPON1,ADAMTS13, PRSS23,COLGALT2,TSPAN15,CERCAM,FKBP7,MXRA8,MYDGF,VWA1,COL24A1,PCSK9,SUMF1
GO Cellular Components	GO:0016324	apical plasma membrane	7.40021	35/361	ANK2,ATP1A1,ATP2B2,CEACAM1,BST2,CSPG4,CYBA,DPP4,FN1, FOLR1,GJA1,HPN,IGFBP2,KCNQ1,MYO7A,SLC22A18,PDGFRB, SCNN1A,SLC1A1,SLC2A5,SLC6A9,SLC9A3,SLC16A2,SPTBN2, ABCG2,TNIK,SLC16A8,DLL1,RAB17,CYBRD1,ACY3,SLC34A3,SLC29A4,ATP6V0D2,AJM1
GO Cellular Components	GO:0030424	axon	7.356318	50/631	ATP1A1,C4B,CALB2,CALCR,CDK5,COMT,EPHA4,GPER1, GRIK2,GRIN1,MAP2,MAPT,NCAM2,ROR1,PALM,PCDH9, SEPTIN4,RAP1GAP,SCN1B,SCN9A,SLC1A1,SLC18A2, SNCG,SYT1,TSC1,SLC30A3,NRP2,BRSK2,SEMA3A,TXNRD2,CPLX1,COBL,NCS1,KLHL24,SLC38A7,CACNG7,RTN4R,IRX3, NTNG2,DIXDC1,NAV1,SYT8,BOC,TPGS1,AZIN2,VSTM2L,IL31RA,EMB,DAB2IP,BTBD8
GO Cellular Components	GO:0043025	neuronal cell body	6.933713	41/483	APOE,ASS1,ATP2B2,CDK5,CYBA,EPHA4,GRIK2,HPN, KCNN4,KCNQ1,MAP2,MAPT,PDE1A,SEPTIN4,RAP1GAP,SCN1B,SLC1A1,SLC2A3,SLC22A3,SNCG,SPTBN2,P2RX6, TXNRD2,CPLX1,COBL,MAPK8IP2,NSMF,PYCARD,PDE11A, TRPM4,KLHL24,SLC38A7,PLXDC1,CACNG7,RAB17,RTN4R,KNDC1,DIXDC1,AZIN2,RTN4RL1,DAB2IP
GO Cellular Components	GO:0045177	apical part of cell	6.514506	37/428	ANK2,ATP1A1,ATP2B2,CEACAM1,BST2,CSPG4,CYBA,DPP4,FN1,FOLR1,GJA1,HFE,HPN,IGFBP2,KCNQ1,MYO7A, SLC22A18,PDGFRB,SCNN1A,SLC1A1,SLC2A5,SLC6A9,SLC9A3, SLC16A2,SPTBN2,ABCG2,VASH1,TNIK,SLC16A8,DLL1,RAB17,CYBRD1,ACY3,SLC34A3,SLC29A4,ATP6V0D2,AJM1
GO Cellular Components	GO:0044297	cell body	6.278883	43/549	APOE,ASS1,ATP2B2,CDK5,CYBA,EPHA4,GJB2,GRIK2,HPN,KCNN4,KCNQ1,MAP2,MAPT,PDE1A,SEPTIN4,RAP1GAP, SCN1B,SLC1A1,SLC2A3,SLC22A3,SNCG,SPTBN2,P2RX6, TXNRD2,CPLX1,COBL,MAPK8IP2,NSMF, PYCARD,PDE11A,TRPM4,KLHL24,SLC38A7,PLXDC1,CACNG7,RAB17,RTN4R,KNDC1,DIXDC1, AZIN2,RTN4RL1,DAB2IP,BTBD8
GO Cellular Components	GO:0005604	basement membrane	5.970006	15/97	COL4A1,COL4A5,COL5A1,COL8A2,FN1,HSPG2,ITGB4,LAMA4,TGFBI,TIMP1, VWA1,FRAS1,HMCN1,NTNG2,LAMA1

Category	Term	Description	LogP	InTerm_InList	Symbols
GO Molecular Functions	GO:0005201	extracellular matrix structural constituent	12.67413	30/172	COL1A2,COL4A1,COL4A5,COL5A1,COL6A2,COL6A3,COL8A2,COL9A2,EFEMP1 ,FN1,HSPG2,IGFBP7,ANOS1,LAMA4,LUM,MATN2,MGP,PRELP,TGFBI, THBS3,COL14A1,VTN,SRPX,EDIL3,SPON1,VWA1,FRAS1,HMCN1,COL24A1,LAMA1
GO Molecular Functions	GO:0005509	calcium ion binding	11.96975	65/714	ANXA4,ARSA,CALB2,CAPS,CBLB,CDH11,CELSR2,MEGF6,MEGF8,EFEMP1,GALNT3,GJB2 ,GRIN1,GSN,HSPG2,IHH,JAG2,LRP1,LTBP3,MATN2,MGP,MYL1,MYL5,PCDH9,PLCD1,SYT1, THBS3,VLDLR,DYSF,BAIAP3,EDIL3,ADAMTS13,MAN1B1,TBC1D9,NCS1,SNED1,DLL1,EHD3,EGFL7,FKBP7, PADI3,PCDH18,TRPM4,LPCAT2,TESC,ITLN1,PCDHGB2,PCDHGA4,PCDHB14,PCDHB13,PCDHB12,PCDHB11, PCDHB10,PCDHB9,PCDHB8,PCDHB4,PLSCR4,PCDHB16,SLC25A23,SYT15,ESYT3,HMCN1,SYT8,SYT12,PKD1L2
GO Molecular Functions	GO:0005198	structural molecule activity	8.721433	58/719	ANK2,APOE,BFSP1,COL1A2,COL4A1,COL4A5,COL5A1,COL6A2,COL6A3,COL8A2,COL9A2,CPOX, EPB41L1,EFEMP1,FN1,HMGCL,HSPG2,IGFBP7,ANOS1,KRT14,KRT15,KRT16,KRT19,KRT34,LAMA4, LUM,MAP2,MATN2,MGP,MYL1,MYL5,SEPTIN4,PPL,PRELP,MRPL23,SNTB1,SPTBN2,TGFBI,THBS3,TPM2, COL14A1,VTN,TUBA1A,SRPX,EDIL3,SPON1,NEBL,MAPK8IP2,CLDN15,KLHL3,PANX2,VWA1,FRAS1,HMCN1, OBSCN,COL24A1,LMNTD2,LAMA1
GO Molecular Functions	GO:0030020	extracellular matrix structural constituent conferring tensile strength	6.039554	10/41	COL1A2,COL4A1,COL4A5,COL5A1,COL6A2,COL6A3,COL8A2,COL9A2,COL14A1,COL24A1
GO Molecular Functions	GO:0038024	cargo receptor activity	5.566828	13/79	ABCA1,DAB2,FOLR1,LGALS3BP,LRP1,TMPRSS2,VLDLR,VTN,SCARF1,SCARA3,ACKR3,LOXL4,SCARF2
GO Molecular Functions	GO:0008509	anion transmembrane transporter activity	5.264042	28/317	ABCC6,CLCN6,GJA1,SLC1A1,SLC4A3,SLC2A3,SLC6A9,SLC9A3,SLC18A2,SLC25A1,SLC22A3,BEST1, APOL1,ABCC3,CLIC3,ASIC3,ABCG2,SLC2A6,SLC16A8,SLC37A1,SLC38A7,SLC7A10,SLC25A23,SLC2A10, TTYH2,SLC37A2,SLCO4C1,SLC27A1
GO Molecular Functions	GO:0022804	active transmembrane transporter activity	5.169383	33/412	ABCA1,ABCA2,ABCA3,ABCC6,ATP1A1,ATP2B2,CLCN6,CYTB,SLC22A18,SLC1A1,SLC4A3, SLC6A9,SLC9A3,SLC16A2,SLC18A2,SLC25A1,SLC22A3,SURF1,ABCC3,SLC16A4,ABCG2, ABCG1,SLC35E2A,SLC16A8,SLC37A1,CYBRD1,SLC2A10,SLC34A3,SLC16A13,SLC37A2,SLC29A4,ATP6V0D2,SLCO4C1
GO Molecular Functions	GO:0030215	semaphorin receptor binding	5.08377	7/23	PLXNB1,SEMA3F,SEMA3B,SEMA3A,SEMA3C,SEMA4G,SEMA3D
GO Molecular Functions	GO:0061134	peptidase regulator activity	4.802483	22/230	ABCA2,BST2,C4B,CASP1,COL6A3,CST2,FN1,ANOS1,SERPINA5,SERPINA1,SERPINE2,SPINT1, TIMP1,TIMP4,CST7,SSPOP,PCOLCE2,PCSK1N,PYCARD,VSIR,PAPLN,CARD16
GO Molecular Functions	GO:0005044	scavenger receptor activity	4.575288	9/47	LGALS3BP,LRP1,TMPRSS2,VTN,SCARF1,SCARA3,ACKR3,LOXL4,SCARF2

Category	Term	Description	LogP	InTerm_InList	Symbols
GO Biological Processes	GO:0000278	mitotic cell cycle	33.55088	97/605	BARD1,BLM,BRCA2,BUB1,BUB1B,CCND3,CCNF,CDK1,CDC6,CDC25A,CDK2,CDKN3,CENPA,CHEK1,DNA2,E2F1,E2F3,EML1,FANCD2,FOXM1,INCENP,KIF11,KIFC1,STMN1,LIG1,MCM3,MCM6,MYBL2,MYH10,NEK2,ORC1,PCNA,POLA1,RBBP8,CDC45,CUL2,PKMYT1,CCNE2,AURKB,ESPL1,KNTC1,DLGAP5,GINS1,NDC80,SMC2,SPAG5,PLK4,STAG2,NES,KIF2C,UBE2C,ZWINT,WDHD1,CHEK2,MAPRE1,TPX2,FBXL7,NCAPD3,NCAPH,KIF4A,CKAP2,RACGAP1,NUSAP1,PPME1,GTSE1,DTL,NCAPG2,TIPIN,RFWD3,PBRM1,NUDT15,TRIM36,PBK,KIF15,SPC25,USP37,CLSPN,NCAPG,GINS3,DSCC1,E2F8,HASPIN,MASTL,TICRR,CDCA5,MISP,IQGAP3,KIF18B,SPC24,SGO2,SGO1,TUBB,SKA1,SKA3,WDR62,CENPW,MZT1
GO Biological Processes	GO:0006259	DNA metabolic process	31.62712	106/763	PARP1,BARD1,BLM,BRCA2,CDK1,CDC5L,CDC6,CDK2,CHEK1,CCN2,DKC1,DNA2,EZH2,FANCA,FANCD2,FANCE,FANCB,FEN1,FOXM1,HELLS,HMGB3,LIG1,MCM3,MCM5,MCM6,NONO,ORC1,PCNA,POLA1,POLE2,POLH,PRIM1,PRIM2,PURA,RAD23A,RAD23B,RAD51C,RAD51B,RBBP8,RFC3,RFC4,RRM1,RRM2,TOP2A,TP73,TYMS,UBE2A,UNG,USP1,RNF113A,CHAF1B,CDC45,RAD54L,TIMELESS,CCNE2,EXO1,APOBEC3B,PCLAF,GINS1,TELO2,DCLRE1A,EXOG,RBX1,PARP2,ZMPSTE24,RAD51AP1,POLD3,WDHD1,SUPT16H,CHEK2,UHRF1,PSMC3IP,GMNN,DTL,TIPIN,RFWD3,FANCI,NUDT15,MCM10,KDM4D,RAD18,FANCM,FAM111A,CHTF18,CLSPN,GINS3,DCLRE1B,DSCC1,ATAD5,RMI1,PIF1,RHNO1,BRIP1,MND1,GINS4,DOT1L,MCM8,TICRR,CDCA5,CGAS,ESCO2,HFM1,KDM1B,MMS22L,RNF169,FAM111B
GO Biological Processes	GO:1903047	mitotic cell cycle process	31.13992	86/514	BARD1,BLM,BRCA2,BUB1,BUB1B,CCND3,CCNF,CDK1,CDC6,CDC25A,CDK2,CDKN3,CENPA,CHEK1,DNA2,E2F1,E2F3,EML1,FANCD2,FOXM1,INCENP,KIF11,KIFC1,STMN1,LIG1,MCM3,MCM6,MYBL2,MYH10,NEK2,ORC1,PCNA,POLA1,RBBP8,CDC45,CUL2,PKMYT1,CCNE2,AURKB,ESPL1,KNTC1,DLGAP5,GINS1,NDC80,SMC2,SPAG5,STAG2,NES,KIF2C,UBE2C,ZWINT,CHEK2,MAPRE1,TPX2,FBXL7,NCAPD3,NCAPH,KIF4A,CKAP2,RACGAP1,NUSAP1,PPME1,GTSE1,DTL,NCAPG2,TIPIN,RFWD3,TRIM36,SPC25,USP37,CLSPN,NCAPG,GINS3,DSCC1,HASPIN,MASTL,TICRR,CDCA5,MISP,IQGAP3,KIF18B,SPC24,SGO2,SGO1,WDR62,MZT1
GO Biological Processes	GO:0006260	DNA replication	27.29154	52/199	BARD1,BLM,BRCA2,CDK1,CDC6,CDK2,CHEK1,DNA2,FEN1,LIG1,MCM3,MCM5,MCM6,ORC1,PCNA,POLA1,POLE2,POLH,PRIM1,PRIM2,PURA,RBBP8,RFC3,RFC4,RRM1,RRM2,CHAF1B,CDC45,TIMELESS,CCNE2,PCLAF,GINS1,POLD3,WDHD1,SUPT16H,GMNN,DTL,TIPIN,RFWD3,MCM10,FANCM,FAM111A,CHTF18,GINS3,DSCC1,ATAD5,RMI1,PIF1,GINS4,TICRR,MMS22L,FAM111B
GO Biological Processes	GO:0010564	regulation of cell cycle process	26.49554	96/734	BARD1,BLM,BRCA2,BUB1,BUB1B,CCND3,CCNF,CDK1,CDC5L,CDC6,CDC25A,CDK2,CHEK1,MAPK14,CCN2,DNA2,E2F1,EDN1,EZH2,FANCD2,FEN1,FHL1,GPR3,INCENP,KIF11,MKI67,TRIM37,NEK2,ORC1,PSME2,RAD51C,RAD51B,RBBP8,RBL1,RRM1,RRM2,DPF3,DPF1,NAA10,CDC45,TIMELESS,PKMYT1,AURKB,MAD2L1BP,ESPL1,KNTC1,CCP110,DLGAP5,TELO2,ZMPSTE24,BTN2A2,NDC80,PRMT5,SMC2,SPAG5,RAD51AP1,PLK4,UBE2C,ZWINT,CHEK2,TPX2,NCAPH,FBXO5,KCNH5,GIT1,RACGAP1,NUSAP1,GTSE1,DTL,NCAPG2,TIPIN,RFWD3,PBRM1,HDAC8,KIF15,PDXP,SPC25,CLSPN,NCAPG,CEP97,E2F8,ATAD5,WDR76,MYO19,FAM83D,TMEM14B,RHNO1,BRIP1,DOT1L,TICRR,CDCA5,SPC24,TTL,MBLAC1,WDR62,FBXO43
GO Biological Processes	GO:0006281	DNA repair	24.66771	75/485	PARP1,BARD1,BLM,BRCA2,CDK1,CDC5L,CDK2,CHEK1,DNA2,FANCA,FANCD2,FANCE,FANCB,FEN1,FOXM1,LIG1,MCM3,MCM5,MCM6,NONO,PCNA,POLA1,POLE2,POLH,RAD23A,RAD23B,RAD51C,RAD51B,RBBP8,RFC3,RFC4,RRM1,TP73,UBE2A,UNG,USP1,RNF113A,CHAF1B,CDC45,RAD54L,TIMELESS,EXO1,PCLAF,DCLRE1A,RBX1,PARP2,ZMPSTE24,RAD51AP1,POLD3,WDHD1,SUPT16H,CHEK2,UHRF1,DTL,RFWD3,FANCI,KDM4D,RAD18,FANCM,FAM111A,CLSPN,DCLRE1B,RMI1,PIF1,RHNO1,BRIP1,GINS4,DOT1L,MCM8,TICRR,CDCA5,CGAS,ESCO2,MMS22L,RNF169
GO Biological Processes	GO:0006974	cellular response to DNA damage stimulus	24.37608	92/725	PARP1,XIAP,BARD1,BLM,BRCA2,CBL,CDK1,CDC5L,CDK2,CHEK1,MAPK14,DNA2,E2F1,FANCA,FANCD2,FANCE,FANCB,FEN1,FOXM1,LIG1,MCM3,MCM5,MCM6,NEDD4,NFATC2,NONO,PCNA,POLA1,POLE2,POLH,RAD23A,RAD23B,RAD51C,RAD51B,RBBP8,RFC3,RFC4,RRM1,TOP2A,TP73,UBE2A,UNG,USP1,RNF113A,CHAF1B,CDC45,RAD54L,TIMELESS,EXO1,PCLAF,DCLRE1A,RBX1,PARP2,TOPORS,ZMPSTE24,RAD51AP1,POLD3,WDHD1,SUPT16H,CHEK2,RPS6KA6,MCTS1,UHRF1,GTSE1,DTL,TIPIN,RFWD3,FANCI,MCM10,KDM4D,RAD18,FANCM,FAM111A,CLSPN,AEN,DCLRE1B,ATAD5,WDR76,RMI1,PIF1,RHNO1,BRIP1,GINS4,DOT1L,MCM8,MASTL,TICRR,CDCA5,CGAS,ESCO2,MMS22L,RNF169
GO Biological Processes	GO:0051301	cell division	23.01488	74/503	BUB1,BUB1B,CCND3,CCNF,CCNT1,CDK1,CDC6,CDC25A,CDK2,CENPA,HELLS,INCENP,ING2,KIF11,KIFC1,STMN1,LIG1,MYH10,NEDD9,NEK2,RBBP8,TOP2A,TIMELESS,CCNE2,AURKB,ESPL1,KNTC1,DCLRE1A,ACTR3,NDC80,SMC2,SPAG5,STAG2,KIF2C,UBE2C,ZWINT,CHEK2,MAPRE1,TPX2,FBXL7,NCAPD3,NCAPH,KIF4A,FBXO5,CKAP2,RACGAP1,NUSAP1,ERCC6L,NCAPG2,TIPIN,MIS18BP1,TRIM36,HAUS7,KNL1,SPC25,USP37,NCAPG,FAM83D,CABLES2,CDCA3,BRIP1,MASTL,LRRCC1,CDCA5,MISP,KIF18B,SPC24,SGO2,SGO1,TUBB,SKA1,SKA3,ASPM,CENPW
GO Biological Processes	GO:0000280	nuclear division	22.88181	58/310	BRCA2,BUB1,BUB1B,FANCA,FANCD2,INCENP,ING2,KIF11,KIFC1,MYBL2,NEK2,RAD51C,RAD51B,TOP2A,RAD54L,CCNE2,AURKB,ESPL1,KNTC1,DLGAP5,ACTR3,NDC80,SMC2,SPAG5,STAG2,KIF2C,UBE2C,ZWINT,CHEK2,MAPRE1,TPX2,NCAPD3,NCAPH,KIF4A,FBXO5,RACGAP1,PSMC3IP,NUSAP1,NCAPG2,CYP26B1,FANCM,FIGNL1,NCAPG,DSCC1,RMI1,HASPIN,BRIP1,MND1,MASTL,CDCA5,MISP,KIF18B,SGO2,SGO1,HFM1,SYCP2L,ASPM,MZT1
GO Biological Processes	GO:0048285	organelle fission	21.93514	59/335	BRCA2,BUB1,BUB1B,FANCA,FANCD2,INCENP,ING2,KIF11,KIFC1,MYBL2,NEK2,RAD51C,RAD51B,TOP2A,RAD54L,CCNE2,AURKB,ESPL1,KNTC1,DLGAP5,ACTR3,NDC80,SMC2,SPAG5,STAG2,KIF2C,UBE2C,ZWINT,CHEK2,MAPRE1,TPX2,NCAPD3,NCAPH,KIF4A,FBXO5,RACGAP1,PSMC3IP,NUSAP1,GDAP1,NCAPG2,CYP26B1,FANCM,FIGNL1,NCAPG,DSCC1,RMI1,HASPIN,BRIP1,MND1,MASTL,CDCA5,MISP,KIF18B,SGO2,SGO1,HFM1,SYCP2L,ASPM,MZT1

Category	Term	Description	LogP	InTerm_InList	Symbols
GO Cellular Components	GO:0098687	chromosomal region	21.70824	63/387	PARP1,BLM,BRCA2,BUB1,BUB1B,CDK1,CDK2,CENPA,CHEK1,DNA2,EZH2,FEN1,CENPI,HELLS,INCENP,MCM3,MCM5,MCM6,NEK2,ORC1,PCNA,PURA,TOP2A,H2BC11,AURKB,KNTC1,TELO2,NDC80,SPAG5,RAD51AP1,STAG2,KIF2C,ZWINT,CHEK2,NCAPD3,NGDN,ERCC6L,CENPQ,PBRM1,MIS18BP1,HJURP,KDM4D,KNL1,THOC2,SPC25,NCAPG,DCLRE1B,CENPM,DSCC1,CENPU,SUV39H2,PIF1,PHF6,CDCA5,SPC24,SGO2,SGO1,ESCO2,SKA1,SKA3,SYCP2L,CENPW,CENPP
GO Cellular Components	GO:0000775	chromosome, centromeric region	15.39388	43/246	BUB1,BUB1B,CENPA,EZH2,CENPI,HELLS,INCENP,NEK2,TOP2A,H2BC11,AURKB,KNTC1,NDC80,SPAG5,STAG2,KIF2C,ZWINT,NCAPD3,NGDN,ERCC6L,CENPQ,PBRM1,MIS18BP1,HJURP,KDM4D,KNL1,SPC25,NCAPG,CENPM,DSCC1,CENPU,SUV39H2,PHF6,CDCA5,SPC24,SGO2,SGO1,ESCO2,SKA1,SKA3,SYCP2L,CENPW,CENPP
GO Cellular Components	GO:0005819	spindle	14.1666	55/425	SLC25A5,XIAP,BUB1B,CDK1,CDC6,MAPK14,EML1,EMD,INCENP,KIF11,KIFC1,NEDD9,NEK2,PIN4,AURKB,MAD2L1BP,ESPL1,KNTC1,DLGAP5,TOPORS,SPAG5,PLK4,STAG2,KIF2C,MAPRE1,TPX2,ARL2BP,KIF4A,FBXO5,CKAP2,GIT1,RACGAP1,NUSAP1,MBIP,HAUS7,KIF15,MAP7D3,SHCBP1,FAM83D,FAM110A,HASPIN,FAM161A,KBTBD8,MISP,CEP128,KIF18B,TTL,CKAP2L,SGO1,TUBB,SKA1,SKA3,ASPM,WDR62,MZT1
GO Cellular Components	GO:0000228	nuclear chromosome	13.93891	41/245	BLM,BRCA2,CHEK1,INCENP,ING2,MCM3,MCM5,MCM6,NEK2,ORC1,PCNA,POLA1,POLE2,PRIM1,PRIM2,TOP2A,CDC45,H3C2,OGT,TIMELESS,GINS1,SMC2,POLD3,STAG2,WDHD1,NCAPH,PSMC3IP,NCAPG2,TIPIN,PBRM1,MCM10,HDAC8,FIGNL1,NCAPG,GINS3,SAP130,ANP32E,GINS4,BRMS1L,SYCP2L,MMS22L

GO Cellular Components	GO:0005657	replication fork	13.56257	22/63	BLM,CHEK1,MCM3,PCNA,POLA1,POLH,PRIM1,PRIM2,RAD51C,RAD51B,RFC3,RFC4, TIMELESS,POLD3,WDHD1,UHRF1,TIPIN,RADX,MCM10,RAD18,PIF1,MMS22L
GO Cellular Components	GO:0000793	condensed chromosome	12.89234	42/276	BLM,BRCA2,BUB1,BUB1B,CDK2,CENPA,CHEK1,FANCD2,CENPI,INCENP,MKI67,NEK2, TOP2A,AURKB,KNTC1,NDC80,SMC2,SPAG5,STAG2,KIF2C,ZWINT,NCAPD3,NCAPH,PSMC3IP, ERCC6L,NCAPG2,CENPQ,PBRM1,HJURP,KNL1,SPC25,NCAPG,CENPM,CENPU,PHF6,SPC24,SGO2,SGO1,SKA1,SKA3,SYCP2L,CENPW
GO Cellular Components	GO:0000779	condensed chromosome, centromeric region	10.38174	30/171	BUB1,BUB1B,CENPA,CENPI,INCENP,NEK2,AURKB,KNTC1,NDC80,SPAG5,KIF2C,ZWINT,NCAPD3, ERCC6L,CENPQ,PBRM1,HJURP,KNL1,SPC25,NCAPG,CENPM,CENPU,PHF6,SPC24,SGO2,SGO1, SKA1,SKA3,SYCP2L,CENPW
GO Cellular Components	GO:0005813	centrosome	10.10142	60/620	BRCA2,CCNF,CDK1,CDK2,CHEK1,E2F1,EMD,MCM3,MPP1,NEK2,PCNA,CDKL5,CDC45, CCNE2,AURKB,ESPL1,CCP110,PCLAF,DLGAP5,MAMLD1,MPHOSPH9,NDC80,SPAG5, PLK4,KIF2C,CEP43,MAPRE1,CEP152,FBXL7,ARL2BP,C2CD3,CKAP2,GIT1,HOOK1,DTL,RAB23,HAUS7,RAD18,KIF15,ILRUN, DCLRE1B,UPF3B,CEP97,CENPU,CCDC15,HASPIN,CEP78,FAM161A,MASTL,LRRCC1,TBC1D31,MISP,CEP128,CKAP2L,SGO1, SKA1,SKA3,ASPM,WDR62,MZT1
GO Cellular Components	GO:0000776	kinetochore	8.767399	27/161	BUB1,BUB1B,CENPA,CENPI,INCENP,NEK2,AURKB,KNTC1,NDC80,SPAG5,KIF2C,ZWINT,ERCC6L, CENPQ,PBRM1,HJURP,KNL1,SPC25,CENPM,CENPU,PHF6,SPC24,SGO2,SGO1,SKA1,SKA3,CENPW
GO Cellular Components	GO:0072686	mitotic spindle	8.254078	28/182	CDK1,CDC6,EML1,KIF11,KIFC1,NEDD9,AURKB,ESPL1,SPAG5,STAG2,MAPRE1,TPX2,CKAP2, GIT1,RACGAP1,NUSAP1,MBIP,HAUS7,FAM83D,FAM161A,MISP,KIF18B,CKAP2L,TUBB,SKA1,SKA3,ASPM,WDR62

Category	Term	Description	LogP	InTerm_InList	Symbols
GO Molecular Functions	GO:0140097	catalytic activity, acting on DNA	12.05952	37/228	BLM,DKC1,DNA2,FEN1,HELLS,LIG1,MCM3,MCM5,MCM6,PCNA,POLA1 ,POLE2,POLH,RAD51C,RAD51B,RBBP8,RFC3,RFC4,SMARCA1,TOP2A,UNG, RAD54L,EXO1,DCLRE1A,EXOG,POLD3,ERCC6L,FANCM,CHTF18,DCLRE1B ,DSCC1,ATAD5,PIF1,BRIP1,MCM8,HFM1,DDX12P
GO Molecular Functions	GO:0008094	ATPdependent activity, acting on DNA	9.197679	23/110	BLM,DNA2,HELLS,MCM3,MCM5,MCM6,RAD51C,RAD51B,RFC3,RFC4, SMARCA1,TOP2A,RAD54L,ERCC6L,FANCM,CHTF18,DSCC1,ATAD5,PIF1 ,BRIP1,MCM8,HFM1,DDX12P
GO Molecular Functions	GO:0140657	ATPdependent activity	8.31263	52/546	ABCB7,BLM,DNA2,HELLS,HSP90AB3P,KIF11,KIFC1,MCM3,MCM5,MCM6,MYH10 ,MYO1D,MYO7B,ORC1,RAD51C,RAD51B,RFC3,RFC4,SMARCA1,TAP2,TOP2A,RAD54L, CCT2,SMC2,KIF2C,HSPA4L,MORC3,DDX58,KIF4A,ATAD2,ERCC6L,MYO5C,KIF15,ATP10A, ATP8B2,KIF17,FANCM,CHTF18,FIGNL1,DSCC1,ATAD5,PIF1,MYO19,BRIP1,MCM8,MFSD2A ,KIF18B,HFM1,ATP11C,NLRP10,KIF7,DDX12P
GO Molecular Functions	GO:0003697	singlestranded DNA binding	7.712723	22/119	BLM,BRCA2,MCM3,MCM5,MCM6,POLA1,PURA,RAD23A,RAD23B,RAD51B,RBBP8,SMARCA1,CDC45 ,RAD54L,FUBP1,SMC2,RAD51AP1,RADX,MCM10,RAD18,FAM111A,MCM8
GO Molecular Functions	GO:0003678	DNA helicase activity	7.302542	17/72	BLM,DNA2,MCM3,MCM5,MCM6,RFC3,RFC4,RAD54L,ERCC6L,FANCM,CHTF18,DSCC1, PIF1,BRIP1,MCM8,HFM1,DDX12P
GO Molecular Functions	GO:0015631	tubulin binding	6.927846	39/376	BRCA2,DIAPH2,EML1,EMD,FES,KIF11,KIFC1,STMN1,VBP1,DLGAP5,SPAG5,KIF2C,MID2,MAPRE1, TPX2,AGTPBP1,KIF4A,GIT1,RACGAP1,NUSAP1,HOOK1,GTSE1,TRIM36,HAUS7,KIF15,KIF17,DIP2B, MAP7D3,FAM83D,DIAPH3,FAM161A,PHF6,KIF18B,KLC3,SPC24,EML5,SKA1,GAS2L3,KIF7
GO Molecular Functions	GO:0003682	chromatin binding	6.856332	51/586	CCNT1,CDK1,CENPA,ARID3A,PHC1,EZH2,FLI1,GATA2,HELLS,ING2,MITF,TRIM37,NFATC2,NONO,ORC1, PBX2,PCNA,POLA1,RBL1,SMARCA1,SRF,TOP2A,CHAF1B,CDC45,EXO1,PCLAF,MED12,PARP2,SCML2,SMC2 ,POLR3G,STAG2,ARID5A,WDHD1,SUPT16H,NCAPH,SUZ12,ATAD2,GRHL1,GMNN,PBRM1,KDM4D,FANCM, HMGN5,BRIP1,MCM8,TICRR,CDCA5,CGAS,KDM1B,RNF169
GO Molecular Functions	GO:0017116	singlestranded DNA helicase activity	6.505726	10/23	DNA2,MCM3,MCM5,MCM6,RFC3,RFC4,CHTF18,DSCC1,PIF1,MCM8
GO Molecular Functions	GO:0140640	catalytic activity, acting on a nucleic acid	6.279278	49/575	BLM,DKC1,DNA2,FEN1,HELLS,LIG1,MCM3,MCM5,MCM6,PCNA,POLA1,POLE2,POLH,PRIM1, RAD51C,RAD51B,RBBP8,RFC3,RFC4,SMARCA1,TOP2A,UNG,RAD54L,EXO1,MED20,DCLRE1A, EXOG,POLR3G,POLD3,CMTR1,DDX58,AGO2,ERCC6L,CDKAL1,TRMU,CNOT6,FANCM,CHTF18, DCLRE1B,DSCC1,ATAD5,TRMT2B,PIF1,BRIP1,MCM8,PSTK,HFM1,MBLAC1,DDX12P
GO Molecular Functions	GO:0016887	ATP hydrolysis activity	5.942209	36/360	BLM,DNA2,HSP90AB3P,KIFC1,MCM3,MCM5,MCM6,ORC1,RFC4,RAD54L,CCT2,SMC2,KIF2C,HSPA4L, MORC3,DDX58,ATAD2,ERCC6L,KIF15,ATP10A,ATP8B2,KIF17,FANCM,CHTF18,FIGNL1,ATAD5,PIF1 ,MYO19,BRIP1,MCM8,KIF18B,HFM1,ATP11C,NLRP10,KIF7,DDX12P

Table 3.11 - KEGG analysis of differentially gene expression in MDA-MB-231 shRNA4 SPAG5

Category	Term	Description	LogP	InTerm_InList	Symbols
KEGG Pathway	hsa04512	ECMreceptor interaction	7.335163	16/88	COL1A2,COL4A1,COL4A5,COL6A2,COL6A3,COL9A2,FN1,GP1BA, HSPG2,ITGB3,ITGB4,LAMA4,THBS3,VTN,FRAS1,LAMA1
KEGG Pathway	hsa04974	Protein digestion and absorption	6.353206	16/103	ATP1A1,COL1A2,COL4A1,COL4A5,COL5A1,COL6A2,COL6A3,COL8A2,COL9A2 ,DPP4,KCNN4,KCNQ1,SLC1A1,SLC9A3,COL14A1,COL24A1
KEGG Pathway	hsa04151	PI3KAkt signaling pathway	4.801948	29/354	COL1A2,COL4A1,COL4A5,COL6A2,COL6A3,COL9A2,CSF1,CSF1R,EFNA1,FGFR4,FN1,GNG7,IL2RB,IL3RA, ITGB3,ITGB4,LAMA4,KITLG,NOS3,PDGFRB,PRLR,RXRA,THBS3,TSC1,VTN,DDIT4,LPAR5,CREB3L1,LAMA1
KEGG Pathway	hsa03320	PPAR signaling pathway	4.319689	11/75	ACSL1,HMGCS1,ME1,RXRA,SCD,FADS2,NR1H3,ANGPTL4,ACSL5,SLC27A1,PLIN5
KEGG Pathway	hsa00604	Glycosphingolipid biosynthesis ganglio series	3.976474	5/15	GLB1,ST3GAL1,ST8SIA1,ST3GAL5,ST6GALNAC4
KEGG Pathway	hsa04360	Axon guidance	3.727803	17/182	CAMK2D,CDK5,EFNA1,EFNB3,EPHA4,MYL5,PLXNB1,SEMA3F,SEMA3B,SEMA3A, SEMA3C,SSH3,SEMA4G,NTNG2,BOC,UNC5B,SEMA3D
KEGG Pathway	hsa00061	Fatty acid biosynthesis	3.558376	5/18	ACACA,ACSL1,FASN,ACSL5,CBR4
KEGG Pathway	hsa00100	Steroid biosynthesis	3.325574	5/20	DHCR7,DHCR24,LIPA,LSS,MSMO1
KEGG Pathway	hsa04142	Lysosome	3.203775	13/132	ABCA2,ARSA,CTSD,CTSO,GAA,GLB1,IDUA,LIPA,MANBA,NAGLU, CTSA,ATP6V0D2,SUMF1
KEGG Pathway	hsa00900	Terpenoid backbone biosynthesis	3.120218	5/22	HMGCR,HMGCS1,MVD,MVK,GGPS1

Category	Term	Description	LogP	InTerm_InList	Symbols
KEGG Pathway	hsa03030	DNA replication	11.46258	14/36	DNA2,FEN1,LIG1,MCM3,MCM5,MCM6,PCNA,POLA1,POLE2,PRIM1,PRIM2,RFC3,RFC4,POLD3
KEGG Pathway	hsa04110	Cell cycle	11.41426	24/126	BUB1,BUB1B,CCND3,CDK1,CDC6,CDC25A,CDK2,CHEK1,E2F1,E2F2,E2F3,MCM3,MCM5,MCM6 ,ORC1,PCNA,RBL1,CDC45,PKMYT1,CCNE2,ESPL1,RBX1,STAG2,CHEK2
KEGG Pathway	hsa03460	Fanconi anemia pathway	8.772341	14/54	BLM,BRCA2,FANCA,FANCD2,FANCE,FANCB,POLH,RAD51C,USP1,TELO2,FANCI,FANCM, RMI1,BRIP1
KEGG Pathway	hsa04064	NFkappa B signaling pathway	7.243264	17/104	PARP1,BIRC3,XIAP,EDA,CXCL1,CXCL2,ICAM1,CXCL8,IRAK1,LTB,NFKB2,NFKBIA,PTGS2,SYK, TRAF1,DDX58,CARD14
KEGG Pathway	hsa04218	Cellular senescence	6.591319	20/156	SLC25A5,CCND3,CDK1,CDC25A,CDK2,CHEK1,MAPK14,E2F1,E2F2,E2F3,FOXM1,IL6,CXCL8,MYBL2, NFATC2,SERPINE1,RBL1,CCNE2,HIPK3,CHEK2

KEGG Pathway	hsa03440	Homologous recombination	5.198329	9/41	BARD1,BLM,BRCA2,RAD51C,RAD51B,RBBP8,RAD54L,POLD3,BRIP1
KEGG Pathway	hsa03410	Base excision repair	5.023206	8/33	PARP1,FEN1,LIG1,PCNA,POLE2,UNG,PARP2,POLD3
KEGG Pathway	hsa03420	Nucleotide excision repair	4.685911	9/47	LIG1,PCNA,POLE2,RAD23A,RAD23B,RFC3,RFC4,RBX1,POLD3
KEGG Pathway	hsa04115	p53 signaling pathway	4.559563	11/73	FAS,CCND3,CDK1,CDK2,CHEK1,SERPINE1,RRM2,TP73,CCNE2,CHEK2,GTSE1
KEGG Pathway	hsa05169	EpsteinBarr virus infection	4.321074	19/202	FAS,CCND3,ENTPD1,CDK2,MAPK14,E2F1,E2F2,E2F3,ICAM1,IL6,IRAK1,NEDD4, NFKB2,NFKBIA,NFKBIE,SYK,TAP2,CCNE2,DDX58

Table 3.12 - GO analysis of differentially gene expression in DU145 shRNA4 SPAG5. Tables shows the enrichment genes differentially upregulated and downregulated in DU145 shRNA4 SPAG5

Category	Term	Description	LogP	InTerm_InList	Symbols
GO Biological Processes	GO:0051301	cell division	23.58335	47/503	BIRC5,RHOB,BUB1,CCNA2,CCNB1,CCND3,CDC20,CDC25C,CENPA,CENPE,CENPF,FGFR2, NEK2,PLK1,SEPTIN4,AURKA,TGFB2,CCNA1,CCNB2,AURKB,PTTG1,KIF23,ESPL1,KIF14,KIF20A, SYCP2,TACC3,KIF2C,UBE2C,CIT,TPX2,KIF4A,UBE2S,PIMREG,ERCC6L,SPDL1,CDCA8,CEP55,FAM83D, CDCA3,NUF2,PSRC1,WNT3A,KNSTRN,EFHC1,KIF18B,SGO1
GO Biological Processes	GO:1903047	mitotic cell cycle process	23.18181	47/514	BIRC5,RHOB,BUB1,CCNA2,CCNB1,CCND3,CDC20,CDC25C,CDKN3,CENPA,CENPE,CENPF,FOXM1, FLNA,MYBL2,NEK2,PLK1,AURKA,WNT10B,CCNA1,CCNB2,AURKB,KIF23,ESPL1,DLGAP5,KIF14,KIF20A, TACC3,KIF2C,UBE2C,CIT,TPX2,KIF4A,UBE2S,GTSE1,SPDL1,CDCA8,CEP55,FBXL8,NABP1,NUF2,PSRC1 ,KNSTRN,EFHC1,IQGAP3,KIF18B,SGO1
GO Biological Processes	GO:0000278	mitotic cell cycle	21.03014	48/605	BIRC5,RHOB,BUB1,CCNA2,CCNB1,CCND3,CDC20,CDC25C,CDKN3,CENPA,CENPE,CENPF,FOXM1,FLNA ,MYBL2,NEK2,PLK1,AURKA,WNT10B,CCNA1,CCNB2,AURKB,KIF23,ESPL1,DLGAP5,KIF14,KIF20A,TUBB3, TACC3,KIF2C,UBE2C,CIT,TPX2,KIF4A,UBE2S,GTSE1,SPDL1,CDCA8,CEP55,FBXL8,NABP1,NUF2,PSRC1,KNSTRN,EFHC1, IQGAP3,KIF18B,SGO1
GO Biological Processes	GO:0140014	mitotic nuclear division	18.58305	26/169	BIRC5,CCNB1,CDC20,CENPE,FLNA,MYBL2,NEK2,PLK1,AURKA,AURKB,KIF23,ESPL1,DLGAP5,KIF14, KIF2C,UBE2C,TPX2,KIF4A,UBE2S,SPDL1,CDCA8,NUF2,PSRC1,KNSTRN,KIF18B,SGO1
GO Biological Processes	GO:0000280	nuclear division	16.49801	31/310	BIRC5,BUB1,CCNB1,CDC20,CENPE,FLNA,MYBL2,NEK2,PLK1,AURKA,CCNA1,CCNB2,AURKB,PTTG1, KIF23,ESPL1,DLGAP5,KIF14,SYCP2,KIF2C,UBE2C,TPX2,KIF4A,UBE2S,SPDL1,CDCA8,NUF2,PSRC1,KNSTRN,KIF18B,SGO1
GO Biological Processes	GO:1902850	microtubule cytoskeleton organization involved in mitosis	15.92312	21/125	BIRC5,CCNB1,CDC20,CENPA,CENPE,FLNA,MYBL2,NEK2,PLK1,AURKA,AURKB,KIF23,ESPL1, DLGAP5,TACC3,TPX2,KIF4A,SPDL1,CDCA8,NUF2,EFHC1
GO Biological Processes	GO:0048285	organelle fission	15.54443	31/335	BIRC5,BUB1,CCNB1,CDC20,CENPE,FLNA,MYBL2,NEK2,PLK1,AURKA,CCNA1,CCNB2,AURKB ,PTTG1,KIF23,ESPL1,DLGAP5,KIF14,SYCP2,KIF2C,UBE2C,TPX2,KIF4A,UBE2S,SPDL1,CDCA8,NUF2, PSRC1,KNSTRN,KIF18B,SGO1
GO Biological Processes	GO:0000070	mitotic sister chromatid segregation	15.19638	20/119	BIRC5,CCNB1,CDC20,CENPE,NEK2,PLK1,AURKB,KIF23,ESPL1,DLGAP5,KIF14,KIF2C,KIF4A, SPDL1,CDCA8,NUF2,PSRC1,KNSTRN,KIF18B,SGO1
GO Biological Processes	GO:0007052	mitotic spindle organization	14.77963	18/94	BIRC5,CCNB1,CDC20,CENPE,FLNA,MYBL2,NEK2,PLK1,AURKA,AURKB,KIF23,DLGAP5,T ACC3,TPX2,KIF4A,CDCA8,NUF2,EFHC1
GO Biological Processes	GO:0051783	regulation of nuclear division	14.57181	21/145	BIRC5,BMP7,BUB1,CALR,CCNB1,CDC20,CDC25C,CENPF,MKI67,NEK2,PDGFB,PLK1, AURKA,WNT5A,AURKB,ESPL1,DLGAP5,SPDL1,CDCA8,NUF2,FBXO43

Category	Term	Description	LogP	InTerm_InList	Symbols
GO Cellular Components	GO:0062023	collagencontaining extracellular matrix	22.78871	43/429	ANXA6,APLP1,BMP7,CALR,COL4A1,COL4A2,COL5A1,COL12A1,CCN2,F13A1,FBLN1,GPC4,FGA,FGB,FGFR2 ,GPC3,CCN1,L1CAM,LAMA5,LGALS1,LOX,MATN3,MFGE8,PCOLCE,PDGFB,SDC2,TGFB1I1,TGFB2,THBS4, HSP90B1,VTN,WNT5A,MFAP5,ADAM19,ADAMTS2,NID2,ADAMTSL4,SMOC1,CRELD1,ITIH5,HMCN2,FREM2,ANXA8
GO Cellular Components	GO:0031012	extracellular matrix	20.42661	46/571	ANXA6,APLP1,BMP7,CALR,COL4A1,COL4A2,COL5A1,COL12A1,CCN2,F13A1,FBLN1,GPC4,FGA,FGB, FGFR2,GPC3,CCN1,L1CAM,LAMA5,LGALS1,LOX,MATN3,MFGE8,PCOLCE,PDGFB,SDC2,TGFB1I1,TGFB2, THBS4,HSP90B1,VTN,WNT5A,WNT6,MFAP5,ADAM19,ADAMTS2,NID2,ADAMTSL4,SMOC1,CRELD1,ITIH5,LINGO1 ,RTN4RL1,HMCN2,FREM2,ANXA8
GO Cellular Components	GO:0030312	external encapsulating structure	20.39579	46/572	ANXA6,APLP1,BMP7,CALR,COL4A1,COL4A2,COL5A1,COL12A1,CCN2,F13A1,FBLN1,GPC4,FGA,FGB, FGFR2,GPC3,CCN1,L1CAM,LAMA5,LGALS1,LOX,MATN3,MFGE8,PCOLCE,PDGFB,SDC2,TGFB1I1,TGFB2, THBS4,HSP90B1,VTN,WNT5A,WNT6,MFAP5,ADAM19,ADAMTS2,NID2,ADAMTSL4,SMOC1,CRELD1,ITIH5,LINGO1 ,RTN4RL1,HMCN2,FREM2,ANXA8
GO Cellular Components	GO:0005788	endoplasmic reticulum lumen	17.38141	32/311	CALR,COL4A1,COL4A2,COL5A1,COL12A1,FGA,GPC3,PDIA3,HSPA5,CCN1,LGALS1,MATN3,MFGE8,PDGFA, PDGFB,PPIB,PROC,SDC2,HSP90B1,WNT5A,WNT6,MANF,PDIA4,FSTL3,HYOU1,FSTL1,SDF2L1,DNAJB11, ADAMTSL4,EDEM2,POGLUT2,WNT3A
GO Cellular Components	GO:0005819	spindle	13.49343	32/425	BIRC5,CCNB1,CDC20,CENPE,CENPF,HMMR,NEK2,PLK1,AURKA,AURKB,KIF23,ESPL1,DLGAP5,KIF14,KIF20A ,TUBB3,TACC3,KIF2C,KIF3A,TPX2,KIF4A,POC1A,SPDL1,CDCA8,SHCBP1,FAM83D,PSRC1,CEP19,KNSTRN, EFHC1,KIF18B,SGO1
GO Cellular Components	GO:0000776	kinetochore	10.6293	18/161	BIRC5,BUB1,CCNB1,CENPA,CENPE,CENPF,CENPI,NEK2,PLK1,AURKB,KIF2C,ERCC6L,SPDL1,HJURP,CENPN, NUF2,KNSTRN,SGO1
GO Cellular Components	GO:0000779	condensed chromosome, centromeric region	10.1894	18/171	BIRC5,BUB1,CCNB1,CENPA,CENPE,CENPF,CENPI,NEK2,PLK1,AURKB,KIF2C,ERCC6L,SPDL1, HJURP,CENPN,NUF2,KNSTRN,SGO1
GO Cellular Components	GO:0000922	spindle pole	10.1894	18/171	CCNB1,CDC20,CENPF,NEK2,PLK1,AURKA,AURKB,DLGAP5,TACC3,TPX2,POC1A,SPDL1,FAM83D, PSRC1,CEP19,KNSTRN,EFHC1,SGO1
GO Cellular Components	GO:0030496	midbody	9.906237	19/201	BIRC5,CENPE,CENPF,HSPA5,NEK2,PLK1,AURKA,HSP90B1,AURKB,KIF23,KIF14,KIF20A,CIT, KIF4A,CDCA8,CEP55,MICAL1,SHCBP1,PSRC1
GO Cellular Components	GO:0034663	endoplasmic reticulum chaperone complex	8.425712	6/11	HSPA5,PPIB,HSP90B1,HYOU1,SDF2L1,DNAJB11

Category	Term	Description	LogP	InTerm_InList	Symbols
GO Molecular Functions	GO:0008017	microtubule binding	10.04464	22/272	BIRC5,CENPE,CENPF,GLI1,PLK1,KIF23,DLGAP5,KIF14,KIF20A,KIF2C,KIF3A,TPX2,KIF4A,KIF26A,GTSE1 ,REEP1,FAM83D,DIAPH3,PSRC1,KNSTRN,KIF18B,DCLK2
GO Molecular Functions	GO:0005509	calcium ion binding	10.0003	36/714	ANXA6,C1S,CACNA1B,CALR,CDH15,FBLN1,HSPA5,MATN3,PCDH1,ENPP2,PPEF1,PROC,RET, THBS4,HSP90B1,MYL9,FSTL1,ANXA10,NID2,SYT11,PCDHB5,EDEM2,PCDHA12,CDH23,SMOC1, CRELD1,CRELD2,SCUBE1,EFCAB11,CDHR1,DNER,EFHC1,RHBDL3,HMCN2,ANXA8,ANXA8L1
GO Molecular Functions	GO:0015631	tubulin binding	9.511698	25/376	BIRC5,CENPE,CENPF,GLI1,PLK1,KIF23,DLGAP5,KIF14,KIF20A,KIF2C,KIF3A,TPX2,SYT11,KIF4A ,KIF26A,IFT81,GTSE1,REEP1,FAM83D,DIAPH3,PSRC1,KNSTRN,EFHC1,KIF18B,DCLK2
GO Molecular Functions	GO:0005201	extracellular matrix structural constituent	9.228349	17/172	COL4A1,COL4A2,COL5A1,COL12A1,FBLN1,FGA,FGB,CCN1,LAMA5,MATN3,MFGE8,PCOLCE, VTN,MFAP5,NID2,CRELD1,HMCN2
GO Molecular Functions	GO:1901681	sulfur compound binding	8.550467	20/269	ANXA6,APLP1,AZU1,BMP7,COL5A1,CCN2,FGFR2,CCN1,PCOLCE,SAA1,THBS4,VTN,FST ,FSTL1,OGDHL,EVA1C,SMOC1,RTN4RL1,RSPO1,ACBD7
GO Molecular Functions	GO:0008201	heparin binding	8.411494	16/170	APLP1,AZU1,BMP7,COL5A1,CCN2,FGFR2,CCN1,PCOLCE,SAA1,THBS4,VTN, FSTL1,EVA1C,SMOC1,RTN4RL1,RSPO1
GO Molecular Functions	GO:0005539	glycosaminoglycan binding	7.928374	18/236	ANXA6,APLP1,AZU1,BMP7,COL5A1,CCN2,FGFR2,HMMR,CCN1,PCOLCE ,SAA1,THBS4,VTN,FSTL1,EVA1C,SMOC1,RTN4RL1,RSPO1
GO Molecular Functions	GO:0003774	cytoskeletal motor activity	5.309946	10/111	CENPE,MYO7B,KIF23,KIF14,KIF20A,KIF2C,KIF3A,KIF4A,DNAI1,KIF18B
GO Molecular Functions	GO:0003777	microtubule motor activity	5.268117	8/67	CENPE,KIF23,KIF14,KIF20A,KIF2C,KIF3A,KIF4A,KIF18B
GO Molecular Functions	GO:0022836	gated channel activity	4.97538	17/340	ANXA6,HCN2,CACNA1B,CACNA1D,CHRNA4,GRID1,KCNJ10,CACNA1G,HCN4, CACNG4,CACNA2D3,KCNK13,KCNK12,ANO2,TMEM63C,PIEZO2,TMC7

Category	Term	Description	InTerm_InList	Symbols
GO Biological Processes	GO:0008285	negative regulation of cell population proliferation	32/786	ALOX5,APOE,BMPR1B,CDK6,SFN,GPER1,HLADRB1,CXCL8,IL15,TNFRSF9,KISS1,LYN, SMAD4,MEF2C,NFIB,NGF,PPARD,SKI,SOD2,SOX2,TGFBR2,VDR,IFITM1,FGFBP1,SH2B3, SPRY1,GPNMB,IL24,NUPR1,IL20RB,PELI1,NDRG4
GO Biological Processes	GO:0071345	cellular response to cytokine stimulus	30/710	CD74,CDK4,CSF1,CTH,CYBA,GPER1,HYAL1,IFI16,CXCL8,IL15,LIFR,SMAD4,RPS16,UGCG, IFITM1,TNFRSF18,TNFRSF11A,RPS6KA5,SPOCK2,SH2B3,IFITM3,IL24,PTP4A3,PADI2,IL20RB ,TOLLIP,RNF125,RAB7B,STING1,SLC27A1
GO Biological Processes	GO:0030155	regulation of cell adhesion	31/786	ALOX5,ANK3,CD74,CDK6,CSF1,FGL1,HLADQB1,HLADRA,HLADRB1,TNC,HYAL1,CXCL8,IL15, LYN,MBP,MYO10,CEACAM6,RAC2,SOX2,TGFBR2,PLPP3,TNFRSF18,SPOCK2,SH2B3,GPNMB, LIMCH1,IL20RB,PAG1,PELI1,ZMIZ1,CCDC80
GO Biological Processes	GO:0006029	proteoglycan metabolic process	9/76	BGN,BMPR1B,FOXL1,HYAL1,PPARD,SPOCK2,DSE,CHST9,B3GALT6
GO Biological Processes	GO:1903510	mucopolysaccharide metabolic process	9/82	BGN,HYAL1,IL15,SPOCK2,B3GNT3,DSE,CEMIP,CHST9,B3GALT6
GO Biological Processes	GO:0051272	positive regulation of cellular component movement	24/597	CD74,CSF1,GPER1,GRB7,HYAL1,CXCL8,LYN,MAP2,MCAM,CEACAM6,P2RY6,RAC2, SOD2,TGFBR2,TIAM1,PLPP3,TNFRSF18,SPOCK2,FGFBP1,SEMA3A,GPNMB,CEMIP,S100A14,JCAD
GO Biological Processes	GO:0022407	regulation of cellcell adhesion	21/477	ALOX5,ANK3,CD74,FGL1,HLADQB1,HLADRA,HLADRB1,IL15,LYN,MBP,MYO10, CEACAM6,SOX2,TGFBR2,PLPP3,SH2B3,GPNMB,IL20RB,PAG1,PELI1,ZMIZ1
GO Biological Processes	GO:0048732	gland development	19/402	CSF1,NKX25,FGL1,HOXB9,TNC,ID2,SMAD4,NFIB,PSAP,SOD2,SOX2,TGFBR2, NKX21,VDR,TNFRSF11A,ZMPSTE24,SEMA3A,FOXB1,DKK3
GO Biological Processes	GO:0030204	chondroitin sulfate metabolic process	6/30	BGN,HYAL1,SPOCK2,DSE,CHST9,B3GALT6
GO Biological Processes	GO:0045596	negative regulation of cell differentiation	26/699	BCL7A,CD74,CDK6,COL5A2,NKX25,EFEMP1,GPER1,HOXB8,ID2,ID3,INSIG1,LYN, SMAD4,MAP2,MBP,PPARD,SKI,SOD2,SOX2,NKX21,TTPA,SPRY1,SEMA3A,SPDEF,BRD9,PRICKLE1

Category	Term	Description	InTerm_InList	Symbols
GO Cellular Components	GO:0000323	lytic vacuole	29/739	ANK3,BGN,CD74,CTSO,FTL,GFAP,HLADQB1,HLADRA,HLADRB1,HYAL1,LYN,CEACAM6,PSAP,SFTPB, TCN2,IFITM1,TSPAN1,IFITM3,CCT2,PADI2,TOLLIP,PIP4P2,LRRC8A,AKR1B10,LPCAT1,MAP1LC3A,ABCA13,RAB7B,HRNR
GO Cellular Components	GO:0005764	lysosome	29/739	ANK3,BGN,CD74,CTSO,FTL,GFAP,HLADQB1,HLADRA,HLADRB1,HYAL1,LYN,CEACAM6,PSAP, SFTPB,TCN2,IFITM1,TSPAN1,IFITM3,CCT2,PADI2,TOLLIP,PIP4P2,LRRC8A,AKR1B10,LPCAT1, MAP1LC3A,ABCA13,RAB7B,HRNR
GO Cellular Components	GO:0030666	endocytic vesicle membrane	14/194	APOE,CD74,CYBA,HLADQB1,HLADRA,HLADRB1,INPP5B,LYN,RAC2,SH3GL2, RAB11B,RAB31,PIP4P2,RAB7B
GO Cellular Components	GO:0030139	endocytic vesicle	17/342	APOE,CD74,CYBA,HLADQB1,HLADRA,HLADRB1,INPP5B,LYN,RAC2,SFTPB,SH3GL2,RAB11B, RAB31,PIP4P2,CEMIP,RINL,RAB7B
GO Cellular Components	GO:0031012	extracellular matrix	22/571	APOE,BGN,COL4A5,COL4A6,COL5A2,COL13A1,MEGF9,EFEMP1,FGL1,FLG,TNC,MMP15, SERPINA1,SERPINB8,PSAP,TFPI2,SPOCK2,NAV2,CTHRC1,CCDC80,BMPER,HRNR
GO Cellular Components	GO:0030312	external encapsulating structure	22/572	APOE,BGN,COL4A5,COL4A6,COL5A2,COL13A1,MEGF9,EFEMP1,FGL1,FLG,TNC,MMP15, SERPINA1,SERPINB8,PSAP,TFPI2,SPOCK2,NAV2,CTHRC1,CCDC80,BMPER,HRNR
GO Cellular Components	GO:0062023	collagencontaining extracellular matrix	18/429	APOE,BGN,COL4A5,COL4A6,COL5A2,COL13A1,MEGF9,EFEMP1,FGL1,FLG,TNC,SERPINA1, SERPINB8,PSAP,NAV2,CTHRC1,CCDC80,HRNR
GO Cellular Components	GO:0005770	late endosome	14/287	CD74,CD79A,HLADRA,HLADRB1,PSAP,RAB27B,SFTPB,TTPA,IFITM3,PIP4P2, RNF128,VPS37B,MAP1LC3A,RAB7B
GO Cellular Components	GO:0005774	vacuolar membrane	18/458	CD74,GFAP,HLADQB1,HLADRA,HLADRB1,LYN,CEACAM6,PSAP,THBD ,IFITM1,TSPAN1,IFITM3,PIP4P2,LRRC8A,LPCAT1,MAP1LC3A,ABCA13,STING1
GO Cellular Components	GO:0030670	phagocytic vesicle membrane	7/77	CYBA,INPP5B,RAC2,RAB11B,RAB31,PIP4P2,RAB7B

Category	Term	Description	InTerm_InList	Symbols
GO Molecular Functions	GO:0005539	glycosaminoglycan binding	14/236	APOE,BGN,COL13A1,HK1,CXCL8,NELL2,TGFBR2,LIPG,SPOCK2,FGFBP1,GPNMB,CEMIP,NAV2,CCDC80
GO Molecular Functions	GO:1901681	sulfur compound binding	12/269	APOE,CBS,COL13A1,GSTM1,CXCL8,SMAD4,NELL2,LIPG,FGFBP1,GPNMB,NAV2,CCDC80
GO Molecular Functions	GO:0008201	heparin binding	9/170	APOE,COL13A1,CXCL8,NELL2,LIPG,FGFBP1,GPNMB,NAV2,CCDC80
GO Molecular Functions	GO:0005201	extracellular matrix structural constituent	9/172	BGN,COL4A5,COL4A6,COL5A2,COL13A1,EFEMP1,TNC,TFPI2,CTHRC1
GO Molecular Functions	GO:0005198	structural molecule activity	21/719	ANK3,APOE,BGN,COL4A5,COL4A6,COL5A2,COL13A1,CLDN7,EFEMP1,FLG,GFAP, HLADRB1,TNC,MAP2,MBP,PGM5,RPL37,RPS12,RPS16,TFPI2,CTHRC1
GO Molecular Functions	GO:0030020	extracellular matrix structural constituent conferring tensile strength	4/41	COL4A5,COL4A6,COL5A2,COL13A1

Table 3.13 - KEGG analysis of differentially gene expression in DU145 shRNA4 SPAG5

Category	Term	Description	LogP	InTerm_InList	Symbols
KEGG Pathway	hsa05200	Pathways in cancer	7.117598	26/531	BIRC5,CCNA2,CCND3,COL4A1,COL4A2,FGFR2,GLI1,GNB3,HES1,LAMA5, PDGFA,PDGFB,PLCG2,RET,TGFB2, HSP90B1,WNT5A,WNT6,WNT10B,CXCR4,FGF18,CCNA1,FGF19,LPAR3,WNT10A,WNT3A
KEGG Pathway	hsa04110	Cell cycle	6.516694	12/126	BUB1,CCNA2,CCNB1,CCND3,CDC20,CDC25C,PLK1,TGFB2,CCNA1,CCNB2,PTTG1,ESPL1
KEGG Pathway	hsa04390	Hippo signaling pathway	6.300872	13/157	BIRC5,BMP6,BMP7,CCND3,CCN2,SNAI2,TGFB2,TP73,WNT5A,WNT6,WNT10B,WNT10A,WNT3A
KEGG Pathway	hsa04114	Oocyte meiosis	5.482149	11/131	BUB1,CCNB1,CDC20,CDC25C,PLK1,AURKA,CCNB2,PTTG1,ESPL1,SGO1,FBXO43
KEGG Pathway	hsa04151	PI3KAkt signaling pathway	4.749862	17/354	CCND3,COL4A1,COL4A2,FGFR2,GNB3,NR4A1,LAMA5,PDGFA,PDGFB,SGK1,THBS4,HSP90B1 ,VTN,FGF18,FGF19,ITGA11,LPAR3
KEGG Pathway	hsa04010	MAPK signaling pathway	4.560482	15/294	CACNA1B,CACNA1D,FGFR2,FLNA,NR4A1,PDGFA,PDGFB,MAP2K6,RASGRF2, TGFB2,FGF18,CACNA1G,FGF19,CACNG4,CACNA2D3
KEGG Pathway	hsa05165	Human papillomavirus infection	4.539432	16/331	CCNA2,CCND3,COL4A1,COL4A2,HES1,LAMA5,THBS4,VTN,WNT5A ,WNT6,WNT10B,CCNA1,ITGA11,WNT10A,HES7,WNT3A
KEGG Pathway	hsa04512	ECMreceptor interaction	4.388718	8/88	COL4A1,COL4A2,HMMR,LAMA5,THBS4,VTN,ITGA11,FREM2
KEGG Pathway	hsa04360	Axon guidance	4.14169	11/182	BMP7,EPHB1,EPHB3,L1CAM,PLCG2,PLXNB3,ROBO1,WNT5A,CXCR4,PLXNC1,MYL9
KEGG Pathway	hsa04810	Regulation of actin cytoskeleton	4.07398	12/218	FGFR2,PDGFA,PDGFB,CXCR4,FGF18,ARHGEF6,FGF19,MYL9,IQGAP2,ITGA11,DIAPH3,IQGAP3

Category	Term	Description	LogP	Log(q-value)	Symbols
KEGG Pathway	hsa04060	Cytokine-cytokine receptor interaction	-4.96716	-2.428082461	BMP8B,BMPR1B,CSF1,CXCL8,IL15,TNFRSF9,LIFR,NGF,TGFBR2,TNFSF13, TNFRSF18,TNFRSF11A,IL32,IL24,IL20RB
KEGG Pathway	hsa05323	Rheumatoid arthritis	-4.47189	-2.233848467	CSF1,HLA-DQB1,HLA-DRA,HLA-DRB1,CXCL8,IL15,TNFSF13,TNFRSF11A
KEGG Pathway	hsa05166	Human T-cell leukemia virus 1 infection	-3.69868	-1.63672785	CCND2,CDK4,HLA-DQB1,HLA-DRA,HLA-DRB1,IL15,SMAD4,MAP3K1, TCF3,TGFBR2,CREB5
KEGG Pathway	hsa05202	Transcriptional misregulation in cancer	-3.56695	-1.629937444	CCND2,ETV4,ETV5,ID2,CXCL8,MEF2C,TCF3,TGFBR2,TSPAN7,NUPR1
KEGG Pathway	hsa04933	AGE-RAGE signaling pathway in diabetic complications	-3.42285	-1.584582023	CDK4,COL4A5,COL4A6,CXCL8,SMAD4,TGFBR2,THBD
KEGG Pathway	hsa04672	Intestinal immune network for IgA production	-3.32815	-1.584582023	HLA-DQB1,HLA-DRA,HLA-DRB1,IL15,TNFSF13
KEGG Pathway	hsa04659	Th17 cell differentiation	-3.22096	-1.584582023	AHR,HLA-DQB1,HLA-DRA,HLA-DRB1,SMAD4,RORC,TGFBR2
KEGG Pathway	hsa04550	Signaling pathways regulating pluripotency of stem cells	-3.17758	-1.584582023	BMPR1B,ID2,ID3,LIFR,SMAD4,SOX2,TCF3,FZD3
KEGG Pathway	hsa00600	Sphingolipid metabolism	-3.16942	-1.584582023	PSAP,UGCG,UGT8,PLPP3,SGMS2
KEGG Pathway	hsa04978	Mineral absorption	-2.92274	-1.416733084	ATP2B4,FTL,VDR,SLC34A2,SLC39A4

Table 3. 8 – Full list of the most upregulated gene in MDA-MB-231 and DU145 SPAG5 silencing

MDA-MB-231

Gene ID	log2FC	Pvalue
AC007743.1	3.600219	0.02138
COL14A1	2.974628	0.017563
ADRA2C	2.55157	3.18E-09
VTN	2.482261	1.35E-16
RTN4RL1	2.335701	0.022593
IQCA1	1.951708	0.000365
IFITM10	1.826823	1.92E-17
SLC1A1	1.765149	1.1E-21
HCN2	1.655286	5.56E-05
MAP2K6	1.642082	1.01E-17
FBXL16	1.604226	3.92E-06
TMC3-AS1	1.415596	0.010943
PSMG3-AS1	1.387122	1.1E-14
AZU1	1.37902	0.030182
BMP6	1.378184	0.000439
COL5A1	1.37122	4.3E-216
ABCG1	1.354695	3.8E-55
RNF165	1.34581	0.016219
LOX	1.277069	1.46E-97
RNF224	1.213513	3.68E-08
TBX1	1.160543	2.55E-16
GPRC5B	1.135041	1.7E-07
SNAP91	1.133238	5.18E-07
PLA2R1	1.038676	1.34E-07
MT-TT	1.020374	0.028618
AC092919.2	0.976845	0.021098
PCSK1N	0.875545	3.64E-07
ACSS1	0.855552	1.64E-36
CYSRT1	0.81	0.00
HIC1	0.810002	5.76E-07
SEPT4	0.805481	0.007125
C1QTNF6	0.786221	6.35E-07
SLC34A3	0.783871	0.001071
GXYLT2	0.776328	2.03E-11
SLC9A3-AS1	0.755333	1.27E-05
PIGP	0.71976	5.4E-07
BAMBI	0.716769	1.89E-06
MLPH	0.669606	4.24E-32
UACA	0.649291	3.79E-67
COL4A1	0.646916	9.97E-16
PPIB	0.62834	2.7E-35

DU145

Gene ID	logFC	Pvalue
VTN	2.036796	2.81E-17
AC007743.1	1.609508	0.045569
IQCA1	1.508737	3.21E-07
DISP2	1.466851	3.8E-07
AC092919.2	1.270309	0.002419
TMC3-AS1	1.262973	0.012387
RNF165	1.242587	0.000159
MT-TT	1.163411	0.009876
SLC1A1	1.087195	4.71E-24
SEPT4	0.961705	0.00036
COL5A1	0.957377	3.11E-18
LOX	0.951869	0.000806
RTN4RL1	0.933449	0.020756
HIC1	0.933172	0.045977
ABCG1	0.910617	8.73E-07
RNF224	0.909891	0.000825
AK4P1	0.88932	0.042073
SLC34A3	0.880487	0.000396
ACSS1	0.876009	3.52E-05
HCN2	0.840094	0.001355
AZU1	0.839536	0.025906
MAP2K6	0.829058	0.00022
IFITM10	0.818397	1.19E-08
TBX1	0.797198	0.000444
NR1H3	0.758249	0.015438
ADRA2C	0.750245	1.98E-05
GPRC5B	0.745429	2.37E-09
PPIB	0.719939	9.43E-11
PCSK1N	0.717357	1.21E-11
MLPH	0.712739	2.53E-13
ABCA1	0.703068	0.00011
PLA2R1	0.700437	0.000218
SNAP91	0.681089	0.048674
PCYOX1L	0.664453	2.91E-10
FBXL16	0.662561	8.64E-10
C1QTNF6	0.661745	8.49E-13
COL4A1	0.658129	4.02E-17
SLC9A3-AS1	0.655729	1.39E-06
BMP6	0.654309	0.013116
RHOB	0.644178	3.08E-14
BAMBI	0.643095	0.002421

NR1H3	0.626549	2.88E-06
PCYOX1L	0.626101	0.001142
DISP2	0.621257	0.002679
AK4P1	0.619435	0.015512
ABCA1	0.61211	1.46E-24
VASH1	0.580471	0.019225

PSMG3-AS1	0.613651	0.000142
CYSRT1	0.595612	0.050129
UACA	0.595434	4.54E-08
VASH1	0.589448	4.66E-06
PIGP	0.587777	3.82E-08
GXYLT2	0.586265	0.004245

Table 3. 8 – Full list of the most downregulated gene in MDA-MB-231 and DU145 SPAG5 silencing

MDA-MB-231

Gene ID	log2FC	Pvalue
SPAG5	-2.76653	0
MAPRE1	-1.61809	0
SMIM13	-1.49776	2.1E-103
CENPM	-1.87795	4.1E-102
AIDA	-1.70707	1.87E-95
SSH1	-0.94024	1.33E-86
CDC25A	-1.70765	8.15E-83
SLC16A3	-0.73274	9.25E-81
GIT1	-1.10858	5.19E-73
ZMPSTE24	-1.00217	7.85E-72
CCT2	-0.73496	1.18E-69
RAD23A	-0.94445	1.16E-65
HK1	-0.63923	1.52E-61
SLC2A4RG	-0.81297	9.44E-61
SCAMP1	-1.27594	5.25E-59
ATL3	-0.649	1.41E-58
TRIM37	-1.04507	4.6E-58
TMEM87B	-1.20655	6.36E-58
CUL2	-0.99373	6.05E-49
AC138392.1	-1.80321	1.07E-46
LIFR	-0.87458	1.41E-41
BCL2L13	-1.07295	3.55E-40
BRD9	-0.82372	8.48E-39
GBE1	-0.72682	1.02E-38
HIPK3	-0.94662	1.33E-35
NOL11	-0.7523	1.07E-34
KISS1	-3.4819	3.55E-31
TSC22D4	-0.73534	4.87E-31
ETV4	-0.76114	6.58E-31
MOCS1	-1.59728	7.51E-29

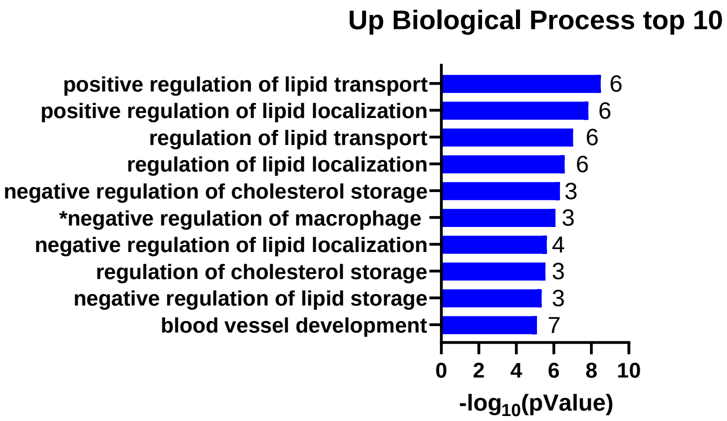
DU145

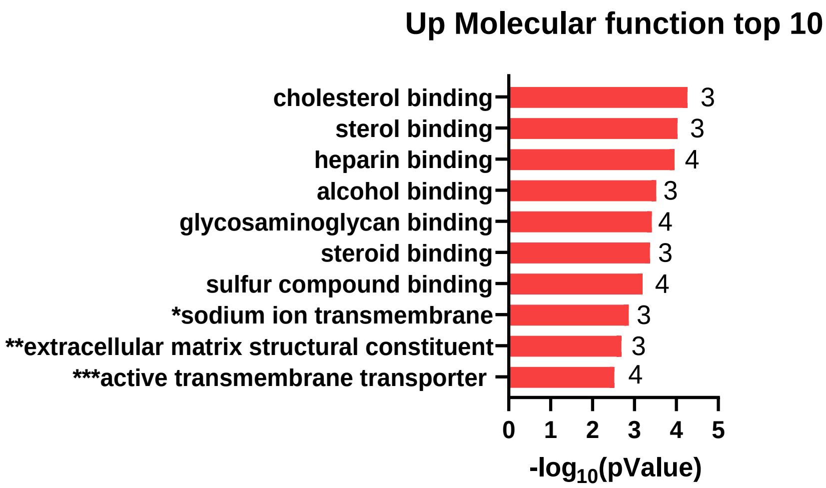
Gene ID	logFC	Pvalue
SLC4A4	-3.20788	1.05E-31
TNFRSF9	-1.79722	0.001667
IL24	-1.64833	0.010038
CXCL8	-1.60461	3.94E-12
GFAP	-1.56509	0.000429
FCMR	-1.55177	4.85E-05
LINC02009	-1.49741	0.048446
KISS1	-1.37365	0.000301
LCP1	-1.36484	0.014008
SPAG5	-1.33389	6.32E-33
KCNQ3	-1.32988	3.26E-17
SCAMP1	-1.28882	1.48E-64
NES	-1.2212	9.16E-07
TC2N	-1.18792	8.64E-10
AC092645.1	-1.16455	0.000831
PPP1R37	-1.14861	1.09E-22
AC138392.1	-1.05484	5.63E-08
SLAMF7	-1.05105	0.000162
IL32	-1.04657	2.19E-06
GIT1	-1.03589	2.15E-41
BCL2L13	-1.02969	4.19E-72
RAD23A	-1.00647	1.59E-44
TRIML2	-0.99719	2.79E-05
VPS37B	-0.97745	2.96E-26
HABP4	-0.96684	4.43E-15
AIDA	-0.9504	6.32E-14
DTNA	-0.94472	5.57E-08
B3GALT6	-0.94366	7.69E-28
LIFR	-0.90427	1.02E-22
CCT2	-0.89256	5.39E-50

HABP4	-1.15884	7.47E-28
SLAMF7	-1.17272	2.78E-27
PPP1R37	-0.87931	7.72E-24
MCAM	-0.66754	1.06E-23
VPS37B	-0.80685	3.76E-22
MAN1A1	-1.61837	4.4E-22
B3GALT6	-0.75438	5.85E-20
PAG1	-0.71119	5.89E-20
NEDD4	-0.78299	2.74E-19
LINC02009	-0.70524	3.29E-17
LCP1	-1.30035	8.47E-16
CXCL8	-1.48513	1.51E-14
IL32	-0.93302	3.09E-12
PARL	-1.02863	3.44E-12
GRB7	-1.91328	1.1E-11
VWA5A	-1.43835	1.58E-11
SPDEF	-1.38762	2.42E-10
RNF144B	-1.09926	1.9E-09
NES	-0.79379	5.75E-09
CXorf57	-1.20798	1.31E-08
PGM5	-3.56899	1.44E-08
RNF125	-1.24744	5.89E-08
IL24	-1.62037	7.57E-08
IPO5P1	-1.06078	5.52E-07
INAFM1	-0.73079	9.14E-06
TNFRSF9	-0.82113	1.76E-05
RPS6KA5	-0.71218	4.17E-05
FCMR	-0.63278	0.000212
BMP8B	-0.93502	0.00022
TRIML2	-1.34189	0.000906
TC2N	-1.90781	0.001362
KCNQ3	-1.27844	0.001853
DTNA	-1.06498	0.00211
GFAP	-0.61773	0.005628
RNF128	-0.97181	0.009952
LINC01291	-1.00307	0.011377
ZNF551	-0.58248	0.012488
AC092645.1	-0.85924	0.016487
SLC4A4	-0.72343	0.02586
PHETA2	-0.88859	0.025863

HK1	-0.87562	1.88E-34
RNF144B	-0.84295	3.33E-07
BMP8B	-0.82523	4.64E-11
TRIM37	-0.82082	4.42E-26
INAFM1	-0.81461	8.04E-05
ATL3	-0.81326	2.62E-41
CUL2	-0.8023	7.51E-16
MOCS1	-0.78464	0.017469
CXorf57	-0.78112	0.000585
MAN1A1	-0.77019	6.69E-05
ETV4	-0.76865	2.59E-07
HIPK3	-0.76363	1.35E-28
VWA5A	-0.75166	5.42E-05
SPDEF	-0.74697	0.009609
PAG1	-0.74498	0.002731
NEDD4	-0.73802	4.14E-34
RPS6KA5	-0.72617	1.02E-07
RNF128	-0.71894	0.020938
GBE1	-0.71854	2.72E-15
CENPM	-0.71227	0.000554
MAPRE1	-0.70997	1.46E-18
MCAM	-0.69952	3.53E-08
IPO5P1	-0.6885	0.037242
ZMPSTE24	-0.67598	1.28E-16
PGM5	-0.65903	0.019204
BRD9	-0.65168	1.89E-12
SMIM13	-0.65156	1.03E-09
TMEM87B	-0.6492	3.27E-09
NOL11	-0.63475	1.58E-15
SLC2A4RG	-0.6302	6.19E-13
PHETA2	-0.61886	0.001627
ZNF551	-0.61766	0.009052
TSC22D4	-0.61373	1.17E-07
CDC25A	-0.61085	0.001991
SLC16A3	-0.60681	7.5E-24
LINC01291	-0.59981	0.008175
SSH1	-0.59976	1.18E-17
RNF125	-0.59121	0.000439
GRB7	-0.58357	1.79E-07
PARL	-0.58013	0.014083

Figure 3.18 – Enrichment analysis on common upregulated in MDA-MB-231 and DU145 shRNA4 SPAG5.





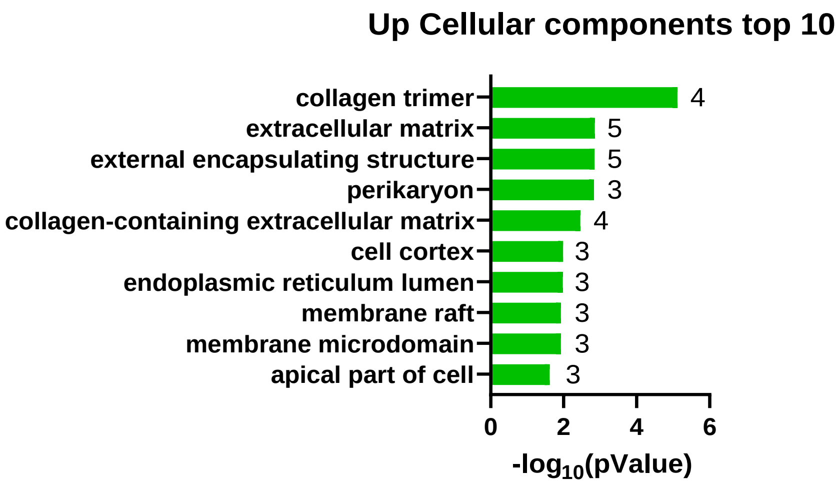


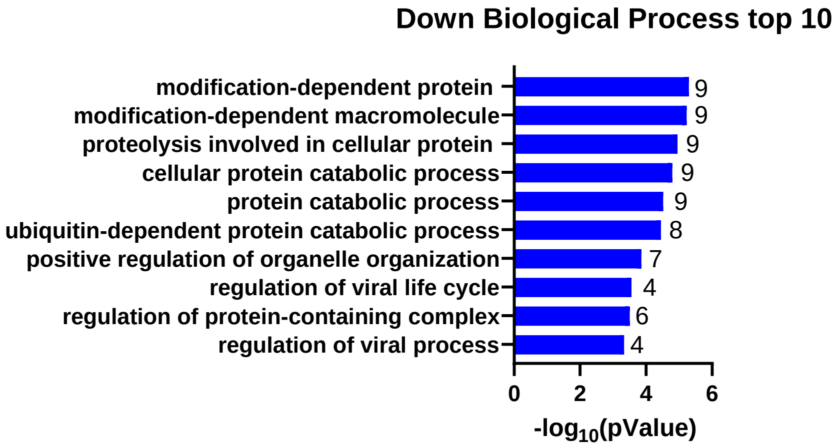
Table 3.18 – GO enrichment analysis on common upregulated in MDA-MB-231 and DU145 shRNA4 SPAG5.

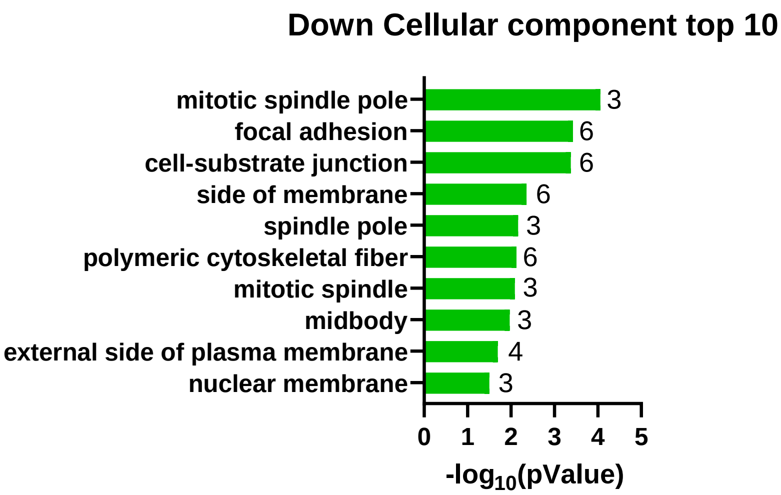
Category	Term	Description	LogP	InTerm_InList	Symbols
GO Biological Processes	GO:0032370	positive regulation of lipid transport	8.508279	6/87	ABCA1,BMP6,MAP2K6,ABCG1,NR1H3,PLA2R1
GO Biological Processes	GO:1905954	positive regulation of lipid localization	7.844446	6/112	ABCA1,BMP6,MAP2K6,ABCG1,NR1H3,PLA2R1
GO Biological Processes	GO:0032368	regulation of lipid transport	7.034797	6/153	ABCA1,BMP6,MAP2K6,ABCG1,NR1H3,PLA2R1
GO Biological Processes	GO:1905952	regulation of lipid localization	6.575178	6/183	ABCA1,BMP6,MAP2K6,ABCG1,NR1H3,PLA2R1
GO Biological Processes	GO:0010887	negative regulation of cholesterol storage	6.325777	3/11	ABCA1,ABCG1,NR1H3
GO Biological Processes	GO:0010745	negative regulation of macrophage derived foam cell differentiation	6.087779	3/13	ABCA1,ABCG1,NR1H3
GO Biological Processes	GO:1905953	negative regulation of lipid localization	5.631458	4/64	ABCA1,ABCG1,NR1H3,PLA2R1
GO Biological Processes	GO:0010885	regulation of cholesterol storage	5.560475	3/19	ABCA1,ABCG1,NR1H3
GO Biological Processes	GO:0010888	negative regulation of lipid storage	5.360605	3/22	ABCA1,ABCG1,NR1H3
GO Biological Processes	GO:0001568	blood vessel development	5.103934	7/505	RHOB,COL4A1,COL5A1,LOX,SLC1A1,TBX1,VASH1

Category	Term	Description	LogP	InTerm_InList	Symbols
GO Cellular Components	GO:0005581	collagen trimer	5.117545	4/86	COL4A1,COL5A1,LOX,C1QTNF6
GO Cellular Components	GO:0031012	extracellular matrix	2.852773	5/571	COL4A1,COL5A1,LOX,VTN,RTN4RL1
GO Cellular Components	GO:0030312	external encapsulating structure	2.849426	5/572	COL4A1,COL5A1,LOX,VTN,RTN4RL1
GO Cellular Components	GO:0043204	perikaryon	2.830421	3/154	SEPTIN4,SLC1A1,RTN4RL1
GO Cellular Components	GO:0062023	collagen-containing extracellular matrix	2.458285	4/429	COL4A1,COL5A1,LOX,VTN
GO Cellular Components	GO:0005938	cell cortex	1.982938	3/310	RHOB,SEPTIN4,MLPH
GO Cellular Components	GO:0005788	endoplasmic reticulum lumen	1.979165	3/311	COL4A1,COL5A1,PPIB
GO Cellular Components	GO:0045121	membrane raft	1.924142	3/326	ABCA1,SLC1A1,RTN4RL1
GO Cellular Components	GO:0098857	membrane microdomain	1.920576	3/327	ABCA1,SLC1A1,RTN4RL1
GO Cellular Components	GO:0045177	apical part of cell	1.612868	3/428	SLC1A1,VASH1,SLC34A3

Category	Term	Description	LogP	InTerm_InList	Symbols
GO Molecular Functions	GO:0015485	cholesterol binding	4.268243	3/50	ABCA1,ABCG1,NR1H3
GO Molecular Functions	GO:0032934	sterol binding	4.030634	3/60	ABCA1,ABCG1,NR1H3
GO Molecular Functions	GO:0008201	heparin binding	3.957058	4/170	AZU1,COL5A1,VTN,RTN4RL1
GO Molecular Functions	GO:0043178	alcohol binding	3.522467	3/89	ABCA1,ABCG1,NR1H3
GO Molecular Functions	GO:0005539	glycosaminoglycan binding	3.413152	4/236	AZU1,COL5A1,VTN,RTN4RL1
GO Molecular Functions	GO:0005496	steroid binding	3.373847	3/100	ABCA1,ABCG1,NR1H3
GO Molecular Functions	GO:1901681	sulfur compound binding	3.199551	4/269	AZU1,COL5A1,VTN,RTN4RL1
GO Molecular Functions	GO:0015081	sodium ion transmembrane transporter activity	2.863178	3/150	HCN2,SLC1A1,SLC34A3
GO Molecular Functions	GO:0005201	extracellular matrix structural constituent	2.693413	3/172	COL4A1,COL5A1,VTN
GO Molecular Functions	GO:0022804	active transmembrane transporter activity	2.521013	4/412	ABCA1,SLC1A1,ABCG1,SLC34A3

Figure 3.15 – GO Enrichment analysis on common downregulated genes in MDA-MB-231 and DU145 shRNA4 SPAG5.





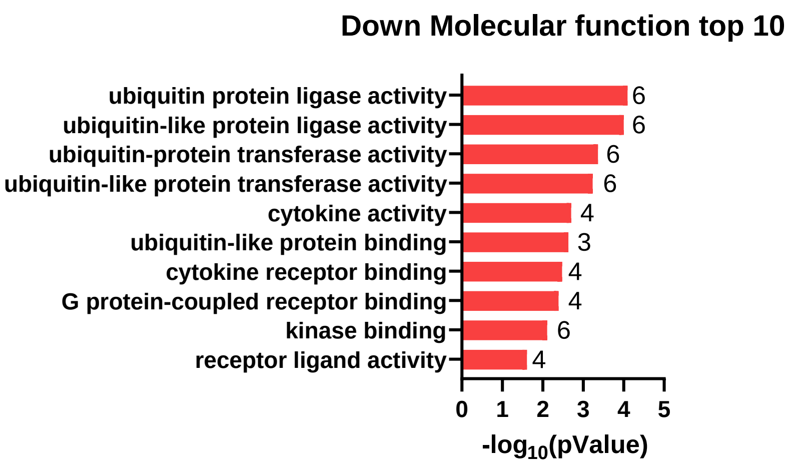


Table 3.15 – Enrichment analysis on common downregulated genes in MDA-MB-231 and DU145 shRNA4 SPAG5.

Category	Description	LogP	InTerm_InList	Symbols
GO Biological Processes	modification-dependent protein catabolic process	-5.30003	9/567	MAN1A1,NEDD4,RAD23A,CUL2,ZMPSTE24,RNF125,RNF128,VPS37B,RNF144B
GO Biological Processes	modification-dependent macromolecule catabolic process	-5.23286	9/578	MAN1A1,NEDD4,RAD23A,CUL2,ZMPSTE24,RNF125,RNF128,VPS37B,RNF144B
GO Biological Processes	proteolysis involved in cellular protein catabolic process	-4.95036	9/627	MAN1A1,NEDD4,RAD23A,CUL2,ZMPSTE24,RNF125,RNF128,VPS37B,RNF144B
GO Biological Processes	cellular protein catabolic process	-4.79484	9/656	MAN1A1,NEDD4,RAD23A,CUL2,ZMPSTE24,RNF125,RNF128,VPS37B,RNF144B
GO Biological Processes	protein catabolic process	-4.52063	9/711	MAN1A1,NEDD4,RAD23A,CUL2,ZMPSTE24,RNF125,RNF128,VPS37B,RNF144B
GO Biological Processes	ubiquitin-dependent protein catabolic process	-4.44803	8/557	MAN1A1,NEDD4,RAD23A,CUL2,RNF125,RNF128,VPS37B,RNF144B
GO Biological Processes	positive regulation of organelle organization	-3.86136	7/503	LCP1,CCT2,SPAG5,NES,MAPRE1,ATL3,GIT1
GO Biological Processes	regulation of viral life cycle	-3.55505	4/139	CXCL8,RAD23A,VPS37B,TRIML2
GO Biological Processes	regulation of protein-containing complex assembly	-3.50513	6/407	GFAP,LCP1,MAPRE1,GIT1,SSH1,AIDA
GO Biological Processes	regulation of viral process	-3.33377	4/159	CXCL8,RAD23A,VPS37B,TRIML2

Category	Description	LogP	InTerm_InList	Symbols
GO Cellular Components	mitotic spindle pole	-4.06015	3/38	SPAG5,MAPRE1,GIT1
GO Cellular Components	focal adhesion	-3.42713	6/421	GRB7,LCP1,MCAM,PGM5,MAPRE1,GIT1
GO Cellular Components	cell-substrate junction	-3.37854	6/430	GRB7,LCP1,MCAM,PGM5,MAPRE1,GIT1
GO Cellular Components	side of membrane	-2.35939	6/683	GFAP,TNFRSF9,LIFR,MCAM,PGM5,SLAMF7
GO Cellular Components	spindle pole	-2.16565	3/171	SPAG5,MAPRE1,GIT1
GO Cellular Components	polymeric cytoskeletal fiber	-2.13035	6/763	GFAP,LCP1,CCT2,SPAG5,NES,MAPRE1
GO Cellular Components	mitotic spindle	-2.09156	3/182	SPAG5,MAPRE1,GIT1
GO Cellular Components	midbody	-1.97462	3/201	SPAG5,SSH1,VPS37B
GO Cellular Components	external side of plasma membrane	-1.69602	4/460	TNFRSF9,LIFR,MCAM,SLAMF7
GO Cellular Components	nuclear membrane	-1.50407	3/304	SLC16A3,ZMPSTE24,HABP4

Category	Description	LogP	InTerm_InList	Symbols
GO Molecular Functions	ubiquitin protein ligase activity	-4.09987	6/316	TRIM37,NEDD4,RNF125,RNF128,TRIML2,RNF144B
GO Molecular Functions	ubiquitin-like protein ligase activity	-4.00391	6/329	TRIM37,NEDD4,RNF125,RNF128,TRIML2,RNF144B
GO Molecular Functions	ubiquitin-protein transferase activity	-3.3626	6/433	TRIM37,NEDD4,RNF125,RNF128,TRIML2,RNF144B
GO Molecular Functions	ubiquitin-like protein transferase activity	-3.23454	6/458	TRIM37,NEDD4,RNF125,RNF128,TRIML2,RNF144B
GO Molecular Functions	cytokine activity	-2.70523	4/235	BMP8B,CXCL8,IL32,IL24
GO Molecular Functions	ubiquitin-like protein binding	-2.63678	3/116	NEDD4,RAD23A,HABP4
GO Molecular Functions	cytokine receptor binding	-2.47665	4/272	CXCL8,LIFR,TRIM37,NES
GO Molecular Functions	G protein-coupled receptor binding	-2.38846	4/288	CXCL8,KISS1,NEDD4,NES
GO Molecular Functions	kinase binding	-2.10914	6/771	CDC25A,GFAP,GRB7,RAD23A,MAPRE1,TRIML2
GO Molecular Functions	receptor ligand activity	-1.61034	4/489	BMP8B,CXCL8,IL32,IL24

Appendix 4

Table 4.1 – Tables show the full list of proteins differentially expressed in whole cell lysate from MDA-MB-231 SPAG5 knockdown (shRNA4) vs MDA-MB-231 control (empty vector pLKO.1) cell populations

Protein ID	logFC	Pvalue	Protein names
MYL6	0.8056733	0.000045	myosin light chain 6
FASN	0.7727183	0.000045	fatty acid synthase
CALB2	0.673465	0.000294174	calbindin 2
NDUFA4	0.8102433	0.000451162	NDUFA4, mitochondrial complex associated
ACACA	0.8507667	0.000551243	acetyl-CoA carboxylase alpha
CKAP4	0.735005	0.000557335	cytoskeleton associated protein 4
THBS1	1.38988	0.000921797	thrombospondin 1
PGLS	0.706205	0.000921797	6-phosphogluconolactonase
GSTK1	0.6611033	0.000921797	glutathione S-transferase kappa 1
AHNAK2	1.7353433	0.001132068	AHNAK nucleoprotein 2
LSS	0.7825617	0.001135313	lanosterol synthase
SNX9	0.8484367	0.001268741	sorting nexin 9
CASK	0.8451483	0.001268741	calcium/calmodulin dependent serine protein kinase
DHCR7	0.7306267	0.001268741	7-dehydrocholesterol reductase
CPOX	1.1915783	0.001351428	coproporphyrinogen oxidase
MYO18A	0.6174167	0.001374104	myosin XVIIIA
NIPSNAP1	1.442075	0.001377712	nipsnap homolog 1
SCRIB	0.8392683	0.001514356	scribbled planar cell polarity protein
TACSTD2	1.318375	0.001565657	tumor associated calcium signal transducer 2
MRRF	0.9911617	0.001574805	mitochondrial ribosome recycling factor
PTTG1IP	1.1009183	0.001749585	PTTG1 interacting protein
ATP1B3	0.7288283	0.001795587	ATPase Na+/K+ transporting subunit beta 3
HMGCL	1.4836431	0.002349424	3-hydroxy-3-methylglutaryl-CoA lyase
ETFB	0.745035	0.002355012	electron transfer flavoprotein subunit beta
COMT	1.2053067	0.00237133	catechol-O-methyltransferase
CHMP1A	0.7476983	0.00241235	charged multivesicular body protein 1A
TPM1	1.4309017	0.002611892	tropomyosin 1
LGALS3BP	0.9106717	0.002611892	galectin 3 binding protein
ACSS2	0.9590017	0.002876132	acyl-CoA synthetase short chain family member 2
MGST1	1.069165	0.003061624	microsomal glutathione S-transferase 1
PPA2	0.5820217	0.003061624	pyrophosphatase (inorganic) 2
EDIL3	1.64719	0.003201382	EGF like repeats and discoidin domains 3
FKBP9	0.66788	0.003201382	FK506 binding protein 9
SUMF2	0.80889	0.003342966	sulfatase modifying factor 2
CD81	0.629395	0.003762997	CD81 molecule
ECH1	0.610095	0.004151725	enoyl-CoA hydratase 1
TAGLN2	0.6808633	0.004536582	transgelin 2

ACSS1	1.23133	0.004793199	acyl-CoA synthetase short chain family member 1
MYH9	1.2113767	0.004795626	myosin heavy chain 9
NUCB2	0.7228883	0.004795626	nucleobindin 2
UBE2L6	1.12521	0.00484794	ubiquitin conjugating enzyme E2 L6
DPM1	0.581995	0.004877593	dolichyl-phosphate mannosyltransferase subunit 1, catalytic
NPC1	1.0354567	0.005054043	NPC intracellular cholesterol transporter 1
SDHB	0.6010067	0.005209352	succinate dehydrogenase complex iron sulfur subunit B
UACA	0.6639217	0.005291533	uveal autoantigen with coiled-coil domains and ankyrin repeats
ETFA	0.595535	0.005472585	electron transfer flavoprotein subunit alpha
DAB2IP	1.2502783	0.005986445	DAB2 interacting protein
RAB9A	0.7118033	0.006064921	RAB9A, member RAS oncogene family
ATP2C1	0.6364	0.006226877	ATPase secretory pathway Ca2+ transporting 1
ACADVL	0.5927033	0.006226877	acyl-CoA dehydrogenase very long chain
PAXX	0.69896	0.006319005	PAXX, non-homologous end joining factor
F11R	1.2516533	0.006535574	F11 receptor
NDUFB1	0.5945783	0.006535574	NADH:ubiquinone oxidoreductase subunit B1
GLB1	0.6684033	0.006700144	galactosidase beta 1
MVK	0.9714967	0.007455213	mevalonate kinase
MYDGF	0.6590467	0.007704685	myeloid derived growth factor
RMDN1	0.67639	0.007821175	regulator of microtubule dynamics 1
RAB11FIP1	0.8221633	0.008180424	RAB11 family interacting protein 1
DRAP1	0.9528433	0.008209326	DR1 associated protein 1
FAM3C	0.6447967	0.008825736	family with sequence similarity 3 member C
NMES1	0.622745	0.009191447	Normal Mucosa Of Esophagus-Specific Gene 1 Protein
MTDH	0.8828433	0.009209191	metadherin
GSN	0.9441617	0.01016367	gelsolin
CTSD	1.0257767	0.010465482	cathepsin D
SCD	0.9591567	0.010655224	stearoyl-CoA desaturase
SNAP29	0.7537133	0.01103501	synaptosome associated protein 29
IDH1	0.8721267	0.011837312	isocitrate dehydrogenase (NADP(+)) 1, cytosolic
DHRS7	0.7311817	0.0120373	dehydrogenase/reductase 7
CLPP	0.6509983	0.012271009	caseinolytic mitochondrial matrix peptidase proteolytic subunit
ATP13A1	0.6065983	0.012271009	ATPase 13A1
TMX2	0.9406467	0.013006293	thioredoxin related transmembrane protein 2
LIPA	0.6733717	0.013465159	lipase A, lysosomal acid type
IGFBP7	1.0975033	0.013583119	insulin like growth factor binding protein 7
SPATS2L	0.784685	0.014506601	spermatogenesis associated serine rich 2 like
DCAF8	0.76774	0.014636254	DDB1 and CUL4 associated factor 8
BDH2	1.05977	0.01524722	3-hydroxybutyrate dehydrogenase 2
MRPL12	0.6018717	0.016272188	mitochondrial ribosomal protein L12
URM1	1.08613	0.016330209	ubiquitin related modifier 1
NDUFAB1	0.97387	0.017412416	NADH:ubiquinone oxidoreductase subunit AB1
NAPRT	0.59707	0.018293272	nicotinate phosphoribosyltransferase
CTSA	0.7921833	0.018656303	cathepsin A
CPD	0.6797967	0.018656303	carboxypeptidase D

NDUFB8	0.7310667	0.019531037	NADH:ubiquinone oxidoreductase subunit B8
ATP2B4	0.7395817	0.019702086	ATPase plasma membrane Ca2+ transporting 4
TNKS1BP1	0.62779	0.020058082	tankyrase 1 binding protein 1
MAPK1	0.7773133	0.020216484	mitogen-activated protein kinase 1
PARP4	0.649345	0.020487415	poly(ADP-ribose) polymerase family member 4
TNS3	0.86765	0.02051778	tensin 3
RCN1	0.6398817	0.021187193	reticulocalbin 1
B2M	0.621295	0.021223761	beta-2-microglobulin
RRBP1	0.8070317	0.021622929	ribosome binding protein 1
RNF181	0.847155	0.021970577	ring finger protein 181
LRPAP1	0.673045	0.022097937	LDL receptor related protein associated protein 1
FN1	0.8545782	0.022384175	fibronectin 1
ECHDC1	0.6173333	0.022384175	ethylmalonyl-CoA decarboxylase 1
CDK5	0.7056567	0.022690834	cyclin dependent kinase 5
MTX2	0.8032733	0.022825913	metaxin 2
EEA1	0.6124867	0.022972088	early endosome antigen 1
IKBIP	0.7148833	0.023265094	IKBKB interacting protein
SPINT2	0.7943917	0.023796127	serine peptidase inhibitor, Kunitz type 2
AGRN	0.8073283	0.024552747	agrin
UBXN6	0.735145	0.025498338	UBX domain protein 6
EIF3J	0.712215	0.026333688	eukaryotic translation initiation factor 3 subunit J
MRPL27	0.7076633	0.02721343	mitochondrial ribosomal protein L27
ASAH1	0.6456483	0.02721343	N-acylsphingosine amidohydrolase 1
ALAD	0.6627033	0.028037604	aminolevulinate dehydratase
PYCR3	0.616035	0.028200584	pyrroline-5-carboxylate reductase 3
NMT2	0.5833033	0.028200584	N-myristoyltransferase 2
TRMT112	0.6006967	0.028514347	tRNA methyltransferase subunit 11-2
NDUFB9	0.944025	0.028736779	NADH:ubiquinone oxidoreductase subunit B9
TBC1D5	0.7830617	0.028736779	TBC1 domain family member 5
PTRHD1	0.67954	0.028736779	peptidyl-tRNA hydrolase domain containing 1
GAA	0.80096	0.029130867	glucosidase alpha, acid
RANBP1	0.5958883	0.029130867	RAN binding protein 1
AFAP1L2	0.7460667	0.031433375	actin filament associated protein 1 like 2
P4HA1	0.687305	0.035530357	prolyl 4-hydroxylase subunit alpha 1
GOLGA2	0.695335	0.039115781	golgin A2
PPOX	0.7200447	0.040197843	protoporphyrinogen oxidase
RRAS2	0.6744117	0.040197843	RAS related 2
MYO1E	0.7309093	0.041685162	myosin IE
PIR	0.6169567	0.042183717	pirin
IST1	0.8479983	0.04258573	IST1, ESCRT-III associated factor
SLC35A2	0.7325681	0.045688848	solute carrier family 35 member A2
WBP2	0.7119317	0.046232297	WW domain binding protein 2
MELTF	0.5828983	0.047284784	melanotransferrin

Entry name	logFC	Pvalue	Protein names
HPRT1	-1.252545	5.92E-08	hypoxanthine phosphoribosyltransferase 1
MAPRE1	-1.3980017	0.00000168	microtubule associated protein RP/EB family member 1
ZMPSTE24	-0.9384217	0.000105945	zinc metallopeptidase STE24
GHITM	-0.8055733	0.000279117	growth hormone inducible transmembrane protein
GBE1	-0.9666867	0.000290266	1,4-alpha-glucan branching enzyme 1
GLO1	-0.6487933	0.000352396	glyoxalase I
DARS1	-0.59797	0.000448744	Aspartyl-TRNA Synthetase 1
RAD23A	-0.99648	0.000556365	RAD23 homolog A, nucleotide excision repair protein
ASNS	-1.0340183	0.000557335	asparagine synthetase domain containing 1
PYGL	-0.7982167	0.000892677	glycogen phosphorylase L
CUL2	-0.9719667	0.000920326	cullin 2
MCM3	-1.013495	0.001070567	minichromosome maintenance complex component 3 associated protein
GDI1	-0.7964817	0.001102588	GDP dissociation inhibitor 1
WDR44	-0.7441833	0.001268741	WD repeat domain 44
PCNA	-0.67534	0.001377712	proliferating cell nuclear antigen
POLR1C	-0.72079	0.001514356	RNA polymerase I and III subunit C
PBDC1	-1.04886	0.001574805	polysaccharide biosynthesis domain containing 1
SEPHS1	-0.791175	0.001574805	selenophosphate synthetase 1
SBDS	-0.603655	0.001574805	SBDS, ribosome maturation factor
KIF11	-0.623325	0.001659784	kinesin family member 11
NONO	-0.8818133	0.001776323	non-POU domain containing octamer binding
TUBB	-0.5979	0.001979869	tubulin beta class I
MCM4	-0.7870467	0.002020141	minichromosome maintenance complex component 4
PGK1	-0.9889233	0.002518539	phosphoglycerate kinase 1
NEDD4	-1.2932506	0.002528786	neural precursor cell expressed, developmentally down-regulated 4, E3 ubiquitin protein ligase
PPME1	-1.5909333	0.002611892	protein phosphatase methylesterase 1
SLC25A5	-0.8981	0.002611892	solute carrier family 25 member 5
ERAP1	-0.74992	0.002611892	endoplasmic reticulum aminopeptidase 1
CFL2	-0.7892917	0.003061624	cofilin 2
PSME2	-0.7973783	0.003103624	proteasome activator subunit 2
MCM5	-0.9505433	0.003201382	minichromosome maintenance complex component 5
MCM2	-0.7774667	0.003201382	minichromosome maintenance complex component 2
MCM6	-0.8342483	0.003342966	minichromosome maintenance complex component 6
OSGEP	-0.7639633	0.003428732	O-sialoglycoprotein endopeptidase
IGBP1	-0.9126433	0.003691103	immunoglobulin binding protein 1
MPP1	-1.0288117	0.004536582	membrane palmitoylated protein 1
MCM7	-0.89233	0.004656648	minichromosome maintenance complex component 7
CMTR1	-0.9283867	0.005054043	cap methyltransferase 1
ARPC1B	-0.614765	0.00506812	actin related protein 2/3 complex subunit 1B
RFC5	-0.7380467	0.005219142	replication factor C subunit 5
ADSL	-0.6029817	0.005747283	adenylosuccinate lyase
PPP2R5D	-1.115325	0.005864383	protein phosphatase 2 regulatory subunit B'delta
RIPK1	-0.614805	0.005864383	receptor interacting serine/threonine kinase 1
PCK2	-0.743765	0.005949232	phosphoenolpyruvate carboxykinase 2, mitochondrial
WDR77	-0.6777033	0.006502081	WD repeat domain 77
DDAH2	-0.66396	0.006670396	dimethylarginine dimethylaminohydrolase 2
TK1	-0.899005	0.00676251	thymidine kinase 1

HCFC1	-0.6119767	0.007176155	host cell factor C1
ABCF1	-0.7319483	0.007334494	ATP binding cassette subfamily F member 1
SMC4	-0.7500783	0.007860649	structural maintenance of chromosomes 4
NCBP1	-0.6097167	0.007947407	nuclear cap binding protein subunit 1
MSRA	-0.87977	0.008329547	methionine sulfoxide reductase A
OARD1	-0.8766433	0.008329547	O-acyl-ADP-ribose deacylase 1
XPO5	-0.809255	0.008345737	exportin 5
CDCP1	-0.6057583	0.008904969	CUB domain containing protein 1
LBR	-0.5935817	0.009191447	lamin B receptor
LSM3	-0.58072	0.00964512	LSM3 homolog, U6 small nuclear RNA and mRNA degradation associated
DUT	-0.7229233	0.010465482	deoxyuridine triphosphatase
ACSL4	-1.0184017	0.011522355	acyl-CoA synthetase long chain family member 4
FTL	-0.72126	0.012271009	ferritin light chain
RFC4	-0.69604	0.012271009	replication factor C subunit 4
NCAPG	-0.9137883	0.01251445	non-SMC condensin I complex subunit G
ERCC6L	-1.0638505	0.012514814	ERCC excision repair 6 like, spindle assembly checkpoint helicase
RRM2	-1.2197133	0.012899856	ribonucleotide reductase regulatory subunit M2
ITPR3	-0.6362467	0.013006293	inositol 1,4,5-trisphosphate receptor type 3
HK2	-0.8554183	0.013465159	hexokinase 2
DDX39A	-0.7532383	0.013465159	DExD-box helicase 39A
KIFC1	-0.8472205	0.014060697	kinesin family member C1
RFC2	-0.7494983	0.01413156	replication factor C subunit 2
SLC25A6	-0.640215	0.014636254	solute carrier family 25 member 6
GMDS	-0.8644033	0.016330209	GDP-mannose 4,6-dehydratase
PPIL1	-0.7849417	0.016475559	peptidylprolyl isomerase like 1
SMC2	-0.6956083	0.017275118	structural maintenance of chromosomes 2
TYMS	-1.0749833	0.017308893	thymidylate synthetase
CDC27	-0.70331	0.017412416	cell division cycle 27
RBM8A	-0.6605133	0.017412416	RNA binding motif protein 8A
WDHD1	-0.784215	0.017694106	WD repeat and HMG-box DNA binding protein 1
ENO2	-0.9208767	0.017717432	enolase 2
ALDH2	-0.5820817	0.018596097	aldehyde dehydrogenase 2 family member
RFC3	-0.76437	0.020058082	replication factor C subunit 3
NUDT3	-0.6714483	0.020487415	nudix hydrolase 3
BCAS2	-0.9846683	0.02051778	BCAS2, pre-mRNA processing factor
SNRNP200	-0.7706417	0.02051778	small nuclear ribonucleoprotein U5 subunit 200
LSM2	-0.7144033	0.020758335	LSM2 homolog, U6 small nuclear RNA and mRNA degradation associated
WDR17	-0.5968335	0.021024124	WD repeat domain 17
NCAPH	-0.792705	0.021220409	non-SMC condensin I complex subunit H
SMCHD1	-0.7173233	0.022435143	structural maintenance of chromosomes flexible hinge domain containing 1
SMC3	-0.7641917	0.022825913	structural maintenance of chromosomes 3
CUL4B	-0.6421567	0.024073494	cullin 4B
SNX12	-0.8172783	0.024291435	sorting nexin 12
HNRNPAB	-0.7621433	0.024370026	heterogeneous nuclear ribonucleoprotein A/B
SMC1A	-0.7610867	0.024891811	structural maintenance of chromosomes 1A

NQO2	-0.7664033	0.025498338	N-ribosyldihydronicotinamide:quinone reductase 2
RBBP7	-0.6041617	0.02721343	RB binding protein 7, chromatin remodeling factor
TSNAX	-0.704	0.02780507	translin associated factor X
OCRL	-0.7537333	0.028736779	OCRL, inositol polyphosphate-5-phosphatase
EIF4A3	-0.8871733	0.028985075	eukaryotic translation initiation factor 4A3
SLC25A4	-0.7167967	0.029499144	solute carrier family 25 member 4
DHX9	-0.5896583	0.029499144	DExH-box helicase 9
TPR	-0.6224783	0.034174691	translocated promoter region, nuclear basket protein
POLR2E	-0.6260767	0.035573157	RNA polymerase II subunit E
KCTD14	-0.5876767	0.041134122	potassium channel tetramerization domain containing 14
GCLM	-0.5965667	0.045582281	glutamate-cysteine ligase modifier subunit
UCHL3	-0.6465867	0.046072648	ubiquitin C-terminal hydrolase L3

Table 4.2– Tables show the full list of proteins differentially expressed in whole cell lysate from DU145 SPAG5 knockdown (shRNA4) vs DU145 control (empty vector pLKO.1)

Protein ID	logFC	Pvalue	Protein names
PLOD2	0.9102233	0.0000018	procollagen-lysine,2-oxoglutarate 5-dioxygenase 2
TAGLN	1.9978000	0.0000281	transgelin 2
ASS1	0.9117067	0.0002215	argininosuccinate synthase 1
SLC7A5	0.6795283	0.0002215	solute carrier family 7 member 5
P4HA2	0.6859083	0.0010722	prolyl 4-hydroxylase subunit alpha 2
SPINT1	0.7130467	0.0014671	serine peptidase inhibitor, Kunitz type 1
EHHADH	0.7213000	0.0020397	enoyl-CoA hydratase and 3-hydroxyacyl CoA dehydrogenase
IL18	0.6753500	0.0021343	interleukin 18
CNTN1	0.8171367	0.0024670	contactin 1
CDH1	0.6199767	0.0024670	cadherin 1
CPT1A	0.7492967	0.0028184	carnitine palmitoyltransferase 1A
LDHAL6A	0.8232723	0.0064972	lactate dehydrogenase A like 6A
EDIL3	0.9038333	0.0099024	EGF like repeats and discoidin domains 3
CNN2	0.7463550	0.0137128	calponin 2
DSP	0.8526917	0.0140276	desmoplakin
CBR3	0.6576333	0.0146019	carbonyl reductase 3
TUBB3	0.6077567	0.0149972	tubulin beta 3 class III
ALDH3B1	0.8438483	0.0220629	aldehyde dehydrogenase 3 family member B1
POMP	0.8479359	0.0255891	proteasome maturation protein
FKBP9	0.5843483	0.0394885	FK506 binding protein 9
CAAP1	0.9661603	0.0409898	caspase activity and apoptosis inhibitor 1
FNDC3B	0.5897333	0.0469626	fibronectin type III domain containing 3B

protein ID	logFC	Pvalue	Protein names
CCT2	-0.6710517	0.00000350	chaperonin containing TCP1 subunit 2
HK1	-0.9666550	0.0000136	hexokinase 1
NAMPT	-0.6151333	0.0000136	nicotinamide phosphoribosyltransferase
FLNC	-1.3137983	0.0000344	filamin C
ATL3	-0.8932283	0.0000846	atlastin GTPase 3

AKR1C2	-0.8908967	0.0000846	aldo-keto reductase family 1 member C2
MAPRE1	-0.7842400	0.0001412	microtubule associated protein RP/EB family member 1
SCAMP1	-2.0134300	0.0001433	secretory carrier membrane protein 1
BCL2L13	-1.5891683	0.0001446	BCL2 like 13
SLC16A3	-0.7439383	0.0001650	solute carrier family 16 member 3
CUL2	-1.5723500	0.0003184	cullin 2
SLC38A2	-1.1839450	0.0003301	solute carrier family 38 member 2
GBE1	-0.6666900	0.0004077	1,4-alpha-glucan branching enzyme 1
DBNL	-0.6274200	0.0004871	drebrin like
PADI2	-1.7476717	0.0006640	peptidyl arginine deiminase 2
NEDD4	-1.2189133	0.0009678	NEDD4 E3 ubiquitin protein ligase
NMD3	-0.6576550	0.0022002	NMD3 ribosome export adaptor
GLUL	-0.9610100	0.0027392	glutamate-ammonia ligase
NT5E	-1.3649800	0.0028184	5'-nucleotidase ecto
SPAG5	-0.8722488	0.0036863	sperm associated antigen 5
CAPG	-1.3539633	0.0052776	capping actin protein, gelsolin like
ASF1B	-0.5925217	0.0059416	anti-silencing function 1B histone chaperone
UPP1	-0.9201450	0.0099953	uridine phosphorylase 1
TJP2	-0.9140683	0.0101636	tight junction protein 2
TMED8	-1.0877333	0.0101925	transmembrane p24 trafficking protein family member 8
PPIL3	-0.6093133	0.0124470	peptidylprolyl isomerase like 3
FAR1	-0.5950900	0.0124712	fatty acyl-CoA reductase 1
ZMPSTE24	-0.8774150	0.0137276	zinc metallopeptidase STE24
NIBAN1	-0.5884317	0.0139895	Niban Apoptosis Regulator 1
ATP2B4	-0.8170267	0.0149972	ATPase plasma membrane Ca2+ transporting 4
LRRC8A	-1.0827400	0.0161831	leucine rich repeat containing 8 VRAC subunit A
PCDHGA6	-0.6672000	0.0233087	protocadherin gamma subfamily A, 6
ANKRD13A	-0.6997485	0.0234497	ankyrin repeat domain 13A
SLC12A7	-0.7562783	0.0241113	solute carrier family 12 member 7
STAT5B	-1.0075748	0.0266554	signal transducer and activator of transcription 5B
KBTBD3	-0.5980550	0.0288638	kelch repeat and BTB domain containing 3
TTC4	-0.7644283	0.0364284	tetratricopeptide repeat domain 4
INTS3	-0.6779117	0.0393744	integrator complex subunit 3
PCCB	-0.6715867	0.0398990	propionyl-CoA carboxylase subunit beta
DDRGK1	-0.6431633	0.0408783	DDRGK domain containing 1
PRC1	-0.7193650	0.0417676	protein regulator of cytokinesis 1
TNPO3	-0.5818933	0.0444988	transportin 3
ERCC2	-0.6247417	0.0448711	ERCC excision repair 2, TFIIH core complex helicase subunit

Table 4.6 – GO Enrichment analysis in MDA-MB-231 shRNA4 SPAG5

Category	Term	Description	LogP	InTerm_InList	Symbols
Biological process	GO:0006091	generation of precursor metabolites and energy	5.15E-15	24/394	ACADVL,ADSL,ALDH2,SLC25A4,ENO2,ETFA,ETFB,GAA,GBE1,HK2,HMGCL,IDH1,NDUFA4 ,NDUFAB1,NDUFB1,NDUFB8, NDUFB9,NQO2,PGK1,PYGL,SDHB,PGLS,ACSS2,ACSS1
biological process	GO:0090329	regulation of DNA-templated DNA replication	1.78E-14	12/58	MCM2,MCM3,MCM4,MCM5,MCM6,MCM7,PCNA,RFC2,RFC3,RFC4,RFC5,ZMPSTE24
biological process	GO:0000278	mitotic cell cycle	1.76E-13	27/605	CDC27,CDK5,GOLGA2,KIF11,KIFC1,MCM2,MCM3,MCM4,MCM6,PCNA ,CHMP1A,TPR,SMC1A,CUL4B,CUL2, SMC3,SMC4,SMC2,WDHD1, MAPRE1,NCAPH,RMDN1,SBDS,PPME1,SNX9,NCAPG,TUBB
biological process	GO:0032787	monocarboxylic acid metabolic process	1.41E-09	20/492	ACACA,ACADVL,ASAH1,ECH1,ENO2,ETFA,ETFB,ACSL4,FASN,HK2,IDH1, NDUFAB1,NPC1,PCK2,PGK1,SCD, ECHDC1,ACSS2,BDH2,ACSS1
biological process	GO:0044282	small molecule catabolic process	1.40E-08	16/352	ACADVL,ALDH2,ECH1,ENO2,ETFA,ETFB,GLB1,HK2,HMGCL,NQO2,OCRL,PCK2, NUDT3,DDAH2,ECHDC1,BDH2
biological process	GO:1900262	regulation of DNA-directed DNA polymerase activity	2.94E-08	5/13	PCNA,RFC2,RFC3,RFC4,RFC5
biological process	GO:0043603	cellular amide metabolic process	1.93E-07	22/792	ABCF1,ACACA,ASAH1,ASNS,CPD,DARS1,ACSL4,GCLM,GLO1,IDH1,MVK,NCBP1, MRPL12,RRBP1,EIF3J,MRPL27, ERAP1,ACSS2,BDH2,ACSS1,MRRF,GSTK1
biological process	GO:0007051	spindle organization	3.25E-07	10/155	GOLGA2,KIF11,KIFC1,MYH9,SMC1A,SMC3,MAPRE1,RMDN1,SBDS,TUBB
biological process	GO:0015931	nucleobase-containing compound transport	8.49E-07	11/215	SLC25A4,SLC25A5,SLC25A6,DHX9,NCBP1,TPR,SLC35A2,EIF4A3,RBM8A,DDX39A,XPO5
biological process	GO:0048545	response to steroid hormone	1.50E-06	12/276	ALAD,GLB1,IDH1,IGFBP7,NEDD4,NPC1,PCK2,PCNA,RBBP7,THBS1,TYMS,WBP2
biological process	GO:0009123	nucleoside monophosphate metabolic process	2.20E-06	7/78	ADSL,DUT,HPRT1,MPP1,TK1,TYMS,CASK
biological process	GO:0140021	mitochondrial ADP transmembrane transport	4.26E-06	3/5	SLC25A4,SLC25A5,SLC25A6
biological process	GO:0030855	epithelial cell differentiation	9.52E-06	16/576	ACADVL,ASAH1,FASN,MYO1E,PCK2,PCNA,PGK1,TYMS,TAGLN2,SPINT2,SCRIB ,F11R,BDH2,WDR77,TUBB,GSTK1
biological process	GO:0005975	carbohydrate metabolic process	1.10E-05	14/454	ALDH2,ENO2,GAA,GBE1,GLB1,GLO1,HK2,IDH1,OCRL,PCK2,PGK1,PYGL,SLC35A2,PGLS
biological process	GO:0006790	sulfur compound metabolic process	1.11E-05	12/336	ACACA,ACSL4,GLB1,GCLM,GLO1,IDH1,MSRA,MVK,NDUFAB1,ACSS2,ACSS1,GSTK1
biological process	GO:0009185	ribonucleoside diphosphate metabolic process	1.69E-05	6/71	ENO2,HK2,MPP1,PGK1,RRM2,CASK
biological process	GO:0016032	viral process	2.17E-05	10/248	CD81,DHX9,HCFC1,NEDD4,NPC1,CHMP1A,EEA1,NMT2,IST1,F11R
biological process	GO:0006782	protoporphyrinogen IX biosynthetic process	3.50E-05	3/9	ALAD,CPOX,PPOX
biological process	GO:0046037	GMP metabolic process	4.93E-05	4/27	ADSL,HPRT1,MPP1,CASK
biological process	GO:0006520	cellular amino acid metabolic process	5.69E-05	10/278	ASNS,DARS1,ETFA,ETFB,GCLM,HMGCL,MSRA,SEPHS1,DDAH2,PYCR3

Category	Term	Description	LogP		
cellular component	GO:0000228	nuclear chromosome	6.40E-13	18/245	DHX9,MCM2,MCM3,MCM4,MCM5,MCM6,MCM7,PCNA,CHMP1A,RBBP7,SMC1A,SMC3,SMC4, SMC2,WDHD1,SMCHD1,NCAPH,NCAPG
cellular component	GO:0034774	secretory granule lumen	5.11E-10	17/322	ALAD,B2M,CTSD,FN1,FTL,GLB1,GSN,IDH1,LGALS3BP,CTSA,MAPK1,PYGL,THBS1,IST1,FAM3C,NAPRT,TUBB
cellular component	GO:0005775	vacuolar lumen	1.00E-08	12/174	ASAH1,CTSD,FTL,GAA,GLB1,LIPA,CTSA,MAPK1,IST1,NAPRT,TUBB,AGRN
cellular component	GO:0019866	organelle inner membrane	1.12E-08	20/556	ACADVL,SLC25A4,SLC25A5,SLC25A6,CPOX,LBR,NDUFA4,NDUFAB1,NDUFB1,NDUFB8,NDUFB9,PPOX,MRPL12,SDHB, TYMS,ZMPSTE24,MTX2,GHITM,MRPL27,C15orf48
cellular component	GO:0005663	DNA replication factor C complex	1.60E-08	4/5	RFC2,RFC3,RFC4,RFC5
cellular component	GO:0072686	mitotic spindle	1.66E-08	12/182	CDC27,GOLGA2,KIF11,KIFC1,MAPK1,TPR,SMC1A,SMC3,MAPRE1,RMDN1,PYCR3,TUBB
cellular component	GO:0005759	mitochondrial matrix	3.65E-08	18/483	ACADVL,ALDH2,SLC25A5,ETFA,ETFB,HMGCL,NDUFAB1,NDUFB8,PCK2,MRPL12,TYMS,CLPP,PPA2, MRPL27,ACSS2,ACSS1,MRRF,GSTK1
cellular component	GO:0000796	condensin complex	2.20E-07	4/8	SMC4,SMC2,NCAPH,NCAPG
cellular component	GO:0005788	endoplasmic reticulum lumen	8.47E-07	13/311	B2M,FN1,IGFBP7,LRPAP1,MELTF,P4HA1,MAPK1,RCN1,THBS1,CKAP4,SUMF2,ERAP1,MYDGF
cellular component	GO:0031300	intrinsic component of organelle membrane	3.75E-06	14/413	SLC25A4,B2M,DHCR7,LBR,NPC1,PPOX,RRBP1,SCD,SLC35A2,ZMPSTE24,MTX2,ATP2C1,GHITM,ATP13A1
cellular component	GO:0071013	catalytic step 2 spliceosome	4.96E-06	7/88	EIF4A3,RBM8A,BCAS2,SNRNP200,LSM3,PPIL1,LSM2
cellular component	GO:0140535	intracellular protein-containing complex	2.00E-05	18/753	CDC27,FTL,HCFC1,MCM3,NCBP1,NEDD4,POLR2E,PPP2R5D,PSME2,RAD23A,CUL4B, CUL2,DPM1,UBE2L6,POLR1C,DCAF8,TNKS1BP1,PAXX
cellular component	GO:0005635	nuclear envelope	2.28E-05	14/485	ABCF1,DHCR7,ITPR3,LBR,NPC1,NUCB2,CHMP1A,RANBP1,TPR,IST1,ZMPSTE24,SEPHS1,MTDH,TUBB
cellular component	GO:0005793	endoplasmic reticulum-Golgi intermediate compartment	6.95E-05	7/132	FN1,GOLGA2,LRPAP1,NUCB2,IST1,MYDGF,MYO18A
cellular component	GO:0005777	peroxisome	1.10E-04	7/142	ECH1,ACSL4,HMGCL,IDH1,MGST1,MVK,GSTK1
cellular component	GO:0015629	actin cytoskeleton	1.30E-04	13/501	ACACA,CFL2,DHX9,GSN,MSRA,MYH9,MYL6,MYO1E,TPM1,CASK,ARPC1B,SNX9,MYO18A
cellular component	GO:0045121	membrane raft	9.15E-04	9/326	SLC25A5,ATP2B4,CTSD,ENO2,NPC1,PGK1,MAPK1,RIPK1,TUBB
cellular component	GO:0005925	focal adhesion	1.50E-03	10/421	B2M,CD81,GSN,MYH9,MAPK1,CASK,ARPC1B,RRAS2,MAPRE1,TNS3
cellular component	GO:0043596	nuclear replication fork	1.65E-03	3/31	MCM3,PCNA,WDHD1

cellular component	GO:0030670	phagocytic vesicle membrane	2.83E-03	4/77	B2M,OCRL,RAB9A,RAB11FIP1

Category	Term	Description	LogP	InTerm_InList	Symbols
molecular function	GO:0140657	ATP-dependent activity	3.04E-18	31/546	ATP2B4,DHX9,ACSL4,KIF11,KIFC1,MCM2,MCM3,MCM4,MCM5,MCM6,MCM7,MYH9,MYL6, MYO1E,RFC2,RFC3,RFC4,RFC5,CLPP,SMC1A,SMC3, EIF4A3,SMC4,DDX39A,SMC2,SNRNP200,SMCHD1,ATP2C1,ERCC6L,ATP13A1,MYO18A
molecular function	GO:0016491	oxidoreductase activity	2.86E-12	28/736	ACADVL,ALDH2,CPOX,DHCR7,ETFA,ETFB,FASN,IDH1,LBR,MGST1,MSRA,NDUFA4 ,NDUFB1,NDUFB8,NDUFB9,NQO2,P4HA1 ,PGK1,PPOX,RRM2,SCD,SDHB,PIR,TMX2,DHRS7,BDH2,PYCR3,GSTK1
molecular function	GO:0016627	oxidoreductase activity, acting on the CH-CH group of donors	1.54E-08	8/59	ACADVL,CPOX,DHCR7,FASN,LBR,PPOX,SDHB,BDH2
molecular function	GO:0045296	cadherin binding	4.53E-08	15/333	FASN,GOLGA2,HCFC1,IDH1,MYH9,RANBP1,TAGLN2,IST1,MAPRE1, SCRIB,F11R,PPME1,SNX9,TNKS1BP1,DAB2IP
molecular function	GO:0003689	DNA clamp loader activity	3.94E-07	4/9	RFC2,RFC3,RFC4,RFC5
molecular function	GO:0009055	electron transfer activity	4.72E-07	9/124	ALDH2,ETFA,ETFB,NDUFA4,NDUFB1,NDUFB8,NDUFB9,NQO2,SDHB
molecular function	GO:0043138	3'-5' DNA helicase activity	3.04E-06	4/14	DHX9,MCM2,MCM5,MCM6
molecular function	GO:0005471	ATP:ADP antiporter activity	4.26E-06	3/5	SLC25A4,SLC25A5,SLC25A6
molecular function	GO:0042803	protein homodimerization activity	5.72E-06	18/686	B2M,CPOX,GLB1,IDH1,MYH9,NQO2,CHMP1A,RRM2,TPM1,TPR,TYMS, EEA1,RIPK1,SMCHD1,F11R,SNX9,DAB2IP,PAXX
molecular function	GO:0016829	lyase activity	2.11E-05	9/197	ADSL,ALAD,ENO2,FASN,GLO1,GMDS,HMGCL,PCK2,ECHDC1
molecular function	GO:0016874	ligase activity	4.00E-05	8/165	ACACA,ASNS,DARS1,ACSL4,GCLM,ACSS2,ACSS1,NAPRT
molecular function	GO:0000287	magnesium ion binding	5.56E-05	9/223	COMT,DUT,ENO2,HMGCL,HPRT1,IDH1,MVK,NUDT3,PPA2
molecular function	GO:0016208	AMP binding	1.83E-04	3/15	PYGL,ACSS2,ACSS1
molecular function	GO:0003774	cytoskeletal motor activity	2.08E-04	6/111	KIF11,KIFC1,MYH9,MYL6,MYO1E,SMC3
molecular function	GO:0015662	P-type ion transporter activity	4.46E-04	3/20	ATP2B4,ATP2C1,ATP13A1
molecular function	GO:0140098	catalytic activity, acting on RNA	6.95E-04	10/380	DARS1,DHX9,POLR2E,TSNAX,POLR1C,EIF4A3,DDX39A,SNRNP200,CMTR1,PTRHD1
molecular function	GO:0019904	protein domain specific binding	8.25E-04	14/687	ATP2B4,FN1,LBR,MYH9,NEDD4,CHMP1A,RIPK1,IST1,F11R,PPIL1,AFAP1L2,TNKS1BP1,DAB2IP,TUBB
molecular function	GO:0005178	integrin binding	1.26E-03	6/156	CD81,FN1,MYH9,THBS1,EDIL3,F11R
molecular function	GO:0005198	structural molecule activity	1.27E-03	14/719	CPOX,FN1,HMGCL,IGFBP7,MYL6,MRPL12,THBS1,TPM1,TPR,EDIL3,ARPC1B,MRPL27,TUBB,AGRN
molecular function	GO:0051015	actin filament binding	1.33E-03	7/215	CFL2,GSN,MYH9,MYO1E,TPM1,ARPC1B,MYO18A

Table 4.7– GO Enrichment analysis in DU145 shRNA4 SPAG5

Category	Term	Description	LogP	InTerm_InList	Symbols
GO Biological Processes	GO:0044282	small molecule catabolic process	-4.969227227	7/352	ALDH3B1,CPT1A,EHHADH,GLUL,HK1,PCCB,UPP1,LDHAL6A,CBR3,SLC7A5 ,LRRC8A,AKR1C2,FAR1
GO Biological Processes	GO:1905039	carboxylic acid transmembrane transport	-4.941540535	5/136	CPT1A,SLC7A5,SLC16A3,SLC38A2,LRRC8A
GO Biological Processes	GO:0060341	regulation of cellular localization	-4.577688711	9/733	ATP2B4,CDH1,GLUL,NEDD4,CCT2,SPAG5,NMD3,DDRGK1,ANKRD13A
GO Biological Processes	GO:0071398	cellular response to fatty acid	-4.226940734	3/35	ASS1,CPT1A,AKR1C2
GO Biological Processes	GO:0010817	regulation of hormone levels	-4.013436264	7/498	CPT1A,AKR1C2,GLUL,STAT5B,SLC7A5,ZMPSTE24,LRRC8A
GO Biological Processes	GO:0018208	peptidyl-proline modification	-3.568853975	3/58	P4HA2,FKBP9,PPIL3
GO Biological Processes	GO:0009161	ribonucleoside monophosphate metabolic process	-3.546856455	3/59	NT5E,UPP1,TJP2
GO Biological Processes	GO:0002720	positive regulation of cytokine production involved in immune response	-3.364879465	3/68	HK1,IL18,SLC7A5
GO Biological Processes	GO:1901607	alpha-amino acid biosynthetic process	-3.364879465	3/68	ASS1,GLUL,PLOD2
GO Biological Processes	GO:0051099	positive regulation of binding	-3.279958216	4/172	ERCC2,MAPRE1,NMD3,DDRGK1
GO Biological Processes	GO:0071407	cellular response to organic cyclic compound	-3.118401964	6/505	ASS1,ATP2B4,CDH1,IL18,NEDD4,PADI2
GO Biological Processes	GO:0002028	regulation of sodium ion transport	-2.99563618	3/91	ATP2B4,CNTN1,NEDD4
GO Biological Processes	GO:0098609	cell-cell adhesion	-2.980707556	6/537	CDH1,CNTN1,DSP,NT5E,TJP2,PCDHGA6
GO Biological Processes	GO:0010927	cellular component assembly involved in morphogenesis	-2.828476266	3/104	CNTN1,ERCC2,FLNC
GO Biological Processes	GO:0000278	mitotic cell cycle	-2.71786318	6/605	CUL2,PRC1,TUBB3,SPAG5,MAPRE1,INTS3

GO Biological Processes	GO:0033365	protein localization to organelle	-2.5	5	ASS1,DSP,FLNC,TAGLN,ZMPSTE24
GO Biological Processes	GO:0061061	muscle structure development	-2.4	3	SLC12A7,DBNL,LRRC8A
GO Biological Processes	GO:0071214	cellular response to abiotic stimulus	-2.3	6	HK1,NEDD4,SPAG5,MAPRE1,TNPO3,DDRGK1
GO Biological Processes	GO:0008361	regulation of cell size	-2.2	4	CNN2,NEDD4,ZMPSTE24,SLC38A2

Category	Term	Description	LogP	InTerm_InList	Symbols
GO Cellular Components	GO:0072686	mitotic spindle	-4.333653417	5/182	CAPG,PRC1,TUBB3,SPAG5,MAPRE1
GO Cellular Components	GO:0030027	lamellipodium	-3.019297655	4/202	CAPG,CDH1,TUBB3,DBNL
GO Cellular Components	GO:0031301	integral component of organelle membrane	-2.849127246	5/383	CPT1A,SCAMP1,ZMPSTE24,LRRC8A,FAR1
GO Cellular Components	GO:0016323	basolateral plasma membrane	-2.832565778	4/227	ATP2B4,DSP,SLC7A5,SLC16A3
GO Cellular Components	GO:0042383	sarcolemma	-2.524483894	3/133	ATP2B4,FLNC,SLC38A2
GO Cellular Components	GO:0043204	perikaryon	-2.346023938	3/154	ASS1,GLUL,DBNL
GO Cellular Components	GO:0042581	specific granule	-2.299872082	3/160	ALDH3B1,CNN2,SCAMP1
GO Cellular Components	GO:0034774	secretory granule lumen	-2.287766353	4/322	CNN2,CCT2,PADI2,DBNL
GO Cellular Components	GO:0045121	membrane raft	-2.268988864	4/326	ATP2B4,CDH1,CNTN1,HK1
GO Cellular Components	GO:0005741	mitochondrial outer membrane	-2.004850762	3/205	ASS1,CPT1A,HK1

Category	GO	Description	LogP	Symbols
GO Molecular Functions	GO:0050839	cell adhesion molecule binding	-5.5	CAPG,CDH1,CNN2,CNTN1,DSP,TJP2 ,EDIL3 ,MAPRE1 ,DBNL
GO Molecular Functions	GO:0016616	oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor	-3.8	CBR3,AKR1C2,EHHADH,LDHAL6A
GO Molecular Functions	GO:0043177	organic acid binding	-3.6	ASS1,GLUL,PLOD2,P4HA2
GO Molecular Functions	GO:0008509	anion transmembrane transporter activity	-3.2	SLC7A5,SLC16A3,SLC12A7,SLC38A2,LRRC8A
GO Molecular Functions	GO:0019901	protein kinase binding	-2.4	ATP2B4,DSP,PRC1,TJP2,SLC12A7,MAPRE1
GO Molecular Functions	GO:0030674	protein-macromolecule adaptor activity	-2.3	ERCC2,CUL2,TJP2,NMD3
GO Molecular Functions	GO:0016853	isomerase activity	-2.3	EHHADH,FKBP9,PPIL3
GO Molecular Functions	GO:0016874	ligase activity	-2.3	ASS1,GLUL,PCCB
GO Molecular Functions	GO:0008022	protein C-terminus binding	-2.2	ERCC2,MAPRE1,DBNL

Table 4.8 KEGG enrichment pathway MDA-MB-231 and DU145 shRNA4 SPAG5

category	Term	Description	LogP	InTerm_InList	Symbols
KEGG	hsa03030	DNA replication	1.91E-15	11/36	MCM2,MCM3,MCM4,MCM5,MCM6,MCM7,PCNA,RFC2,RFC3,RFC4,RFC5
KEGG	hsa00010	Glycolysis / Gluconeogenesis	7.77E-07	7/67	ALDH2,ENO2,HK2,PCK2,PGK1,ACSS2,ACSS1
KEGG	hsa05012	Parkinson disease	1.02E-06	12/266	SLC25A4,SLC25A5,SLC25A6,ITPR3,NDUFA4,NDUFAB1,NDUFB1,NDUFB8,NDUFB9,SDHB,UBE2L6,TUBB
KEGG	hsa00620	Pyruvate metabolism	1.47E-06	6/47	ACACA,ALDH2,GLO1,PCK2,ACSS2,ACSS1
KEGG	hsa00100	Steroid biosynthesis	1.42E-05	4/20	DHCR7,LBR,LIPA,LSS
KEGG	hsa03040	Spliceosome	1.75E-05	8/147	NCBP1,EIF4A3,RBM8A,BCAS2,SNRNP200,LSM3,PPIL1,LSM2
KEGG	hsa04146	Peroxisome	3.86E-05	6/82	ECH1,ACSL4,HMGCL,IDH1,MVK,GSTK1
KEGG	hsa01232	Nucleotide metabolism	4.73E-05	6/85	ADSL,DUT,HPRT1,RRM2,TK1,TYMS

KEGG	hsa04142	Lysosome	6.95E-05	7/132	ASAH1,CTSD,GAA,GLB1,LIPA,NPC1,CTSA
KEGG	hsa01212	Fatty acid metabolism	7.29E-05	5/57	ACACA,ACADVL,ACSL4,FASN,SCD
KEGG	hsa00500	Starch and sucrose metabolism	1.57E-04	4/36	GAA,GBE1,HK2,PYGL
KEGG	hsa01240	Biosynthesis of cofactors	1.75E-04	7/153	ADSL,ALAD,ALDH2,CPOX,GCLM,PPOX,NAPRT
KEGG	hsa04217	Necroptosis	2.21E-04	7/159	SLC25A4,SLC25A5,SLC25A6,FTL,CHMP1A,PYGL,RIPK1
KEGG	hsa04530	Tight junction	3.21E-04	7/169	MYH9,MYL6,NEDD4,PCNA,ARPC1B,SCRIB,F11R
KEGG	hsa04114	Oocyte meiosis	5.06E-04	6/131	CDC27,ITPR3,PPP2R5D,MAPK1,SMC1A,SMC3
KEGG	hsa04210	Apoptosis	6.17E-04	6/136	PARP4,CTSD,ITPR3,MAPK1,RIPK1,DAB2IP
KEGG	hsa04810	Regulation of actin cytoskeleton	1.44E-03	7/218	CFL2,FN1,GSN,MYH9,MAPK1,ARPC1B,RRAS2
KEGG	hsa00020	Citrate cycle (TCA cycle)	1.50E-03	3/30	IDH1,PCK2,SDHB
KEGG	hsa04144	Endocytosis	3.17E-03	7/251	NEDD4,CHMP1A,EEA1,IST1,ARPC1B,SNX12,RAB11FIP1
KEGG	hsa04216	Ferroptosis	3.71E-03	3/41	ACSL4,FTL,GCLM

Category	**Term**	**Description**	**LogP**	**InTerm_InList**	**Symbols**
KEGG Pathway	hsa05230	Central carbon metabolism in cancer	-4.79298	4/70	HK1,SLC7A5,SLC16A3,LDHAL6A
KEGG Pathway	hsa00640	Propanoate metabolism	-4.34537	3/32	EHHADH,PCCB,LDHAL6A
KEGG Pathway	hsa00010	Glycolysis / Gluconeogenesis	-3.38381	3/67	ALDH3B1,HK1,LDHAL6A

Appendix 5

Table 5.2 - Complete list of cross-over gene and proteome in MDA-MB-231 SPAG5 silencing.

Downregulated	Upregulated
MAPRE1	CTSD
MCM3	IGFBP7
RRM2	PTTG1IP
NONO	GSN
PPME1	LGALS3BP
SLC25A5	TACSTD2
TUBB	LSS
KIFC1	FAM3C
PGK1	DHCR7
OCRL	CTSA
ZMPSTE24	MELTF
NCAPH	ACSS2
ABCF1	CPOX
RAD23A	MYH9
NCAPG	UACA
HNRNPAB	MGST1
ERCC6L	EDIL3
PCNA	MYDGF
KIF11	IDH1
CUL2	MYO18A
PPP2R5D	LIPA
MCM5	ACACA
MCM6	COMT
ENO2	DAB2IP
GDI1	ACSS1
XPO5	PGLS
WDHD1	FN1
PYGL	AHNAK2
GBE1	GSTK1
SNX12	FASN
SEPHS1	CALB2
LSM2	NIPSNAP1
HPRT1	GAA
SMC2	GLB1
CMTR1	SCD
GCLM	TNS3
PPIL1	MVK
TYMS	F11R
PSME2	HMGCL
ERAP1	ECH1
RFC4	DHRS7
PBDC1	CDK5
NEDD4	ETFB
RFC3	BDH2
MPP1	
WDR44	
NUDT3	