

A System Biology Study of the Molecular Mechanism of Metformin on Breast Cancer

SHAYMAA ISMAEL KADHIM AL-JUBOORI

**A thesis submitted in partial fulfilment of the requirements of
Nottingham Trent University for the degree of Doctor of Philosophy**

July 2019

Copyright Statement

This work is the intellectual property of the author. You may copy up to 5% of the work for private study, or personal, non-commercial research. Any re-use of the information contained within this document should be fully referenced, quoting the author, title, university, degree level and pagination. Queries or requests for any other use, or if a more substantial copy is required, should be directed in the owner(s) of the Intellectual Property Rights.

This thesis is dedicated to the souls of my parents and brothers.

This thesis is also dedicated to my family who supported me and prayed for my success throughout my studies.

[Dedication in Arabic]

من لم يشكر المخلوق لم يشكر الخالق...

بعد الحمد لله والصلاة والسلام على سول الله وأله الأطهار..

أهدي ثمرة جهدي المتواضع هذا الى أرواح أبي وأمي وأخوتي رحمهم الله تعالى كصدقة جارية.

كما وأهدي هذا العمل الى زوجي فراس وابنتي تقى وابني محمد والى إخوتي وأخواتي ووالدي وزوجي الذين طالما صلوا ودعوا من اجلي طوال فترة دراستي.

كل الحب والامتنان لكم جميعا.....

شيماء...

Acknowledgements

It is my pleasure to thank my Director of Studies, **Professor Graham Ball** for giving me the opportunity to work in the exciting field of cancer research and drug discovery. Your guidance and endless support are much appreciated. I want to express my gratitude and great thanks to my second supervisor **Dr Tarik Regad** for the leadership, help, support, encouragement and inspiration that he provided throughout the entire journey of my PhD. I learned from **Dr Regad** how to give for free, how to be motivated, and how to love my work.

My great thanks to **Professor Graham Pockley** for all his encouragement, advice and help throughout the entire research period.

I would like to thank **Dr Jayakumar Vadakekolathu** for the enormous help with my experimental work and analysis of results. Jay has excellent skills, different expertise, scientific ideas and kind spirit. It was my pleasure to work with him and improve my skills with his help.

I would also like to thank **Dr Gemma Foulds** for scientific support and help with the Flow Cytometry work and data analysis. Gemma is a lovely person with a kind spirit.

I would also like to thank **Dr David Boocock** and **Dr Amanda Miles** for their scientific support and help with the mass spectrometry analysis.

I would like to heartily thank two special people; **Anne Schneider** and **Stephen Reeder**. I would like to thank them for all the help, the support they have given me over the past four years to facilitate my work.

A special thanks to **Dr Amanda Coutts** for her help in using the xCELLigence System.

Great thanks for **Dr Graham Hickman** to facilitate my work with the plate reader.

I would also thank **Professor Robert Rees** and **Professor Sergio Rutella**, and all other post-docs, research fellows, senior research fellows and research technicians, **Dr Stephanie McArdle**, **Dr Murrium Ahmad**, **Dr Clare Coveney**, **Dr Simon Hood**, and **Catherine Johnson** for their constant help and encouragement.

I would forever be grateful to my dearest friend **Dr Rukaia Almshayakhchi** for all her support, help, encouragement in both social and practical life. Rukaia has been the real sister when I felt lonely, sad, and weak. I do thank God to meet you and be working together in the same place.

My heartfelt thanks to my fellow PhD colleagues; **Dr Devika Agarwal, Dr Magdalena Buczek, Dr Matthew Nicklin, Anna Di Biase, Joshua Pearson, Sarra Idri, Abdullah Al-Omari, Dimitrios Kapsoulis, Divya Nagarajan, Luisa Barbato, Sarah Wagner, Pauline LeVu, Jenny Ashforth, Melissa Courtney, Marco Bocchetti**. Doing my PhD with such amazing people has been fantastic. I wish you all the best for the future.

Finally, I extend my sincere thanks and gratitude to my husband **Firas Al-allak** and my children **Tuqa and Mohammed** for their continuous support, encouragement, understanding and patience during this entire journey.

Without you, I could not do.

Great love, thanks, and gratitude to my **brothers, sisters** and my **parents in law** for their support, encouragement, and prayers upon my PhD study period.

Thank you all for having faith in my capabilities and for being always proud of me.

Finally, my great thanks and gratitude to **Iraqi Government/Ministry of Higher Education and Scientific Research**, to offer this scholarship and fund this project.

TABLE OF CONTENTS

List of Figures.....	xiii
List of Tables	xviii
Abbreviations	xxi
Abstract.....	1
Chapter 1	
1. Introduction	2
1.1. Cancer	2
1.2. Cancer Statistics for 2018.....	3
1.3. Hallmarks of cancer	7
1.3.1. Maintaining proliferation signalling	8
1.3.2. Avoiding growth suppressors	8
1.3.3. Avoid apoptosis (programmed cell death)	8
1.3.4. Enabling replicative immortality	9
1.3.5. Inducing of angiogenesis.....	9
1.3.6. Promoting tissue invasion and metastasis.....	9
1.4. Additional Hallmarks of Cancer.....	10
1.5. Breast cancer.....	12
1.6. The Mammary Gland Structure and development.....	13
1.7. Breast cancer classification.....	15
1.8. Breast cancer molecular subtypes.....	17
1.8.1. luminal A (hormone receptor positive/HER2 negative)	17
1.8.2. luminal B (hormone receptor /HER2 positive)	17
1.8.3. HER2-enriched (hormone receptor negative/HER2 positive)	18
1.8.4. Basal-like (hormone receptor negative/HER2 negative)	18
1.8.5. Claudin-low (hormone receptor/HER2 negative)	18

1.8.6.	Normal breast-like (hormone receptor negative/HER2 negative)	19
1.9.	Role of receptors in Breast cancer.....	19
1.9.1.	Hormone receptors.....	19
1.9.1.1.	Oestrogen receptors.....	20
1.9.1.2.	Progesterone receptors (PRs)	20
1.9.2.	Human Epidermal Growth Factor Receptor 2 (HER2) or HER2/neu.....	20
1.10.	Molecular mechanisms of breast cancer.....	21
1.11.	Breast cancer treatments.....	24
1.12.	Metformin hydrochloride.....	26
1.13.	Metformin and cancer treatment.....	27
1.14.	Metformin and breast cancer.....	33
1.15.	Antitumor Effect of Metformin.....	37
1.16.	Cancer-related targets of Metformin.....	40
1.16.1.	Insulin-like growth factor-1 (IGF-1)	40
1.16.2.	IGF-1 receptor (IGF-1R)	40
1.16.3.	Adenosine monophosphate protein kinase (AMPK)	40
1.16.4.	Mammalian target of Rapamycin (mTOR)	41
1.16.5.	Phosphoinositide 3-kinase (PI3K)	42
1.17.	Systems biology.....	42
1.18.	Systems biology and biological networks	43
1.19.	Machine learning and Artificial Neural Networks.....	44
1.20.	Stepwise ANN approach.....	45

1.21.	ANN interaction algorithm.....	46
1.22.	Filtering of interactions and visualisation of model.....	47
1.23.	The aim of the project	47

Chapter 2

2. MATERIALS AND METHODS	50
2.1. Materials	50
2.1.1. Reagents.....	50
2.1.2. Buffers and gels.....	53
2.1.3. Equipment.....	56
2.1.4. Cell line growth media.....	58
2.2. Methods	59
2.2.1. Cell culture.....	59
2.2.1.1. Routine Cell culture maintenance	59
2.2.2. Interrogation Metformin mode of action in different breast cancer molecular subtypes.....	60
2.2.2.1. Metformin preparation.....	60
2.2.2.2. Cell viability assay (MTT)	61
2.2.2.3. Cell proliferation assay.....	61
2.2.2.4. Flow cytometry analysis of Cell apoptosis.....	62
2.2.2.5. Metformin treatment.....	63
2.2.2.6. Annexin staining.....	63
2.2.2.7. Morphological effect of Metformin.....	64
2.2.3. Microarrays.....	64
2.2.3.1. RNA-Extraction with STAT 60 and Qiagen RNeasy Mini Kit	64
2.2.3.2. Labelling with a fluorescent dye.....	65
2.2.3. Purification of the labelled/amplified RNA.....	67
2.2.3.4. cRNA quantification	67
2.2.3.5. Hybridisation.....	68
2.2.3.6. Microarray Data Analysis	69
2.2.3.6.1. Artificial neural networks-based approach.....	69
2.2.3.6.2. Regression-based method and Fold change (Fc) -based method.....	70
2.2.4. Validation of Gene Expression Microarray data	71
2.24.1. cDNA synthesis.....	71

2.2.4.2. The application of Real-time quantitative PCR (qRT-PCR) and primers preparatio.....	72
2.2.4.3. Western Blot analysis of protein lysates from MDA-MB-468 and SkBr3 cell lines.....	73
2.2.4.3.1. preparation of the cell lysates	73
2.2.4.3.2. Protein quantification.....	73
2.2.4.4 Preparation of Gel and electrophoresis.....	75
2.2.4.5. Wet transfer of proteins from gel to a membrane.....	75
2.2.4.6. Blocking and probing membranes with antibodies.....	76
2.2.4.7. Immunofluorescence analysis of proteins.....	76
2.2.4.8. NanoString n Counter XT Gene Expression Assay for gene expression profiling in Basal-like and HER2 phenotypes samples:	77
2.2.5. Evaluation of PTK2B (Protein Tyrosine Kinase 2 Beta) role in breast carcinoma	78
2.2.5.1 PTK2B Plasmid bulking	78
2.2.5.2. Plasmid isolation.....	79
2.2.5.3. HEK-293T transfection.....	79
2.2.5.4. Infection of target breast cancer cell lines.....	80
2.2.5.5. Puromycin selection.....	80
2.2.5.6. whole cell lysate preparation for knockdown validation.....	80
2.2.5.7. Cell Migration or Wound-healing assay (Scratch assay)	81
2.2.5.8. Cultrex® BME Cell Invasion Assay.....	81
2.2.5.9. xCELLigence System (RTCA) for Cell Proliferation.....	82
2.2.5.10. Mass spectrometry analysis.....	82
2.2.5.10.1. whole cell lysate preparation.....	82
2.2.5.10.2. Cell lysate protein quantification and MS analysis.....	83

Chapter 3

3. Investigating the effect of Metformin on different breast cancer molecular subtypes	
3.1. Introduction.....	84
3.2. Investigating the effect of Metformin on cell survival (viability) using (MTT assay)	85
3.3. The computing of EC50 doses of Metformin for all breast cancer subtypes.....	88

3.4. Investigating the effect of Metformin on cell proliferation of the different breast cancer phenotypes.....	92
3.5. Investigating the effect of Metformin on apoptosis of breast cancer subtypes.....	95
3.6. Investigating the effect of Metformin on the morphology of breast cancer subtypes.....	102
3.7. Discussion.....	103

Chapter 4

4. Identification of novel markers and cellular pathways associated with BASAL-LIKE and HER2 breast cancer subtypes

4.1. Introduction.....	105
4.2. The Applications of microarray in gene expression.....	105
4.3. Identification of novel markers associated with BASAL-LIKE and HER2 breast cancer subtypes by applying the microarrays technique:	106
4.3.1. RNA isolation and integrity determination for MDA-MB-468 and SkBr3 samples.....	107
4.4. Labelling and cRNA quality assessment	109
4.5. Microarray data analysis.....	111
4.5.1. Artificial Neural Network (ANN) stepwise analysis-based approach.....	112
4.5.2. Interactome Network Map for both BASAL-LIKE (MDA-MB-468) and HER2 (SkBr3) cell lines.....	112
4.5.3. Interactome Network Map for BASAL-LIKE (MDA-MB-468) and HER2 (SkBr3) cell lines assessed each separately.....	116
4.5.4. Regression-based method and fold change- based method for analysis of data from BASAL-LIKE (MDA-MB-468) and HER2 (SkBr3) cell lines.....	121
4.6. Pathway analysis.....	135
4.7. Discussion.....	138

Chapter 5

5. Confirmation of Gene Expression Microarray data by Real-time quantitative PCR (qRT-PCR), Immunoblotting and Immunofluorescence analyses

5.1. Introduction.....	141
5.2. Assessment of efficiency, sensitivity and reproducibility using a standard curve.....	142
5.3. Real-time quantitative PCR (qRT-PCR) of investigated genes.....	142
5.4. IRF-9, PTK2B and SERPINB4 protein expression in Basal-like and HER2 breast cancer cell lines.....	148
5.5. Immunofluorescence analysis of PTK2B, IRF-9, and SERPINB4 expression in Basal-like (MDA-MB-468), Claudin-Low (MDA-MB-231) and HER2 (SkBr3 and MDA-MB-453) cell lines.....	151
5.6. The application of NanoString technology on Basal-like and HER2 phenotypes samples for pan-cancer pathways detection.....	161
5.7. The application of Real-time quantitative PCR (qRT-PCR) on common genes obtained from NanoString analysis.....	168
5.8. Validation of GADD45A protein expression in Basal-like and HER2 breast cancer cell lines.....	170
5.9. Discussion.....	172

Chapter 6

6. Evaluation of the role of PYK2 in breast cancer invasion and metastasis

6.1. Introduction.....	177
6.2. Generation of PTK2B knockdown cell lines from Basal-like (MDA-MB-468), Claudin -Low (MDA-MB-231) and HER2 (SkBr3 and MDA-MB-453) breast cancer celllines.....	178
6.3. Analyses of PTK2B protein expression in newly generated Basal-like, Claudin – Low, and HER2 PTK2B-knockdown ce lines.....	180
6.4. Investigating the effects of PTK2B knockdown on cell migration, invasion, and proliferation of breast cancer cell lines in the absence or presence of treatment	182

6.4.1. Effect of PTK2B knockdown on cell migration and in response to Metformin treatment.....	182
6.4.2. Effect of PTK2B knockdown on cell invasion and in response to Metformin treatment.....	187
6.4.3. Effect of PTK2B knockdown on cell invasion and in response to Metformin treatment using the xCELLigence System (RTCA).....	190
6.5. Employment of MetaCore™ software analysis to interrogate <i>PYK2</i> related pathways and biological processes.....	195
6.6. Protein expression profiling of samples from Her2 cell lines expressing pLKO.1 (control) and <i>PTK2B</i> shRNAs.....	201
6.7. Pathways Enrichment analysis.....	212
6.8. Discussion.....	214

Chapter 7

7. DISCUSSION

7.1. Summary of discussion.....	220
7.2. Future work.....	234

BIBLIOGRAPHY	235
APPENDIX	254

List of Figures

Figure 1. 1. Figure 1.1. Percentage of total cancer incidents by cancer site comparing the increasing rates of cancer cases in three different periods in 1993 (spotted), 2014 (spotted) and 2035 (estimated), and divided by sex.....	3
Figure 1. 2 Cancer diagnoses incidence rates in the UK 2016.....	4
Figure 1. 3. The number of cancer registrations in England, 2016.....	5
Figure 1. 4. The total number of deaths from cancer in England and Wales from 2001- 2017....	6
Figure 1. 5. The hallmarks of cancer.....	11
Figure 1. 6. Age-standardised cancer mortality and incidence rates for female breast cancer in England 2016.....	12
Figure 1. 7. Mammary Gland anatomy and histology.....	14
Figure 1. 8. Schematic representation of breast carcinomas progression from benign to malignant.....	16
Figure 1. 9. Graph representing the percentage of breast cancer patients according to different regimens in the UK.....	25
Figure 1. 10. The number of main clinical trials assessing Metformin anticancer and prevention capacities against several types of cancer... ..	32
Figure 1. 11. Metformin Mode of action in tumour cells... ..	39
Figure 1. 12. The schematic representation of the main stages of the project framework.....	49
Figure 3.1. Graph representing cell survival responses of different breast cancer cell lines to Metformin at different concentrations, determined using the MTT assay for 24h (A) and 48h (B) of treatment.....	87
Figure 3.2. The EC50 of the five breast cancer phenotypes as determined using the MTT assay after 24h (A) and 48h(B) of treatments with Metformin.....	89
Figure 3. 3. The actual EC50 of Metformin concentrations (mean) for the five breast cancer cell lines showing the significant values between the two time points (24 H, 48 H) of treatment with Metformin that was obtained from Cell Viability Assay.....	91
Figure 3.4. The proliferation status of different breast cancer phenotypes after treatment with different concentrations of Metformin, as determined using a cell proliferation assay.....	94
Figure 3.5. Flow cytometry analysis of Metformin resistance – gating strategy for cell viability staining.	96
Figure 3.6. Flow cytometry analysis of Metformin resistance – gating strategy for cell viability staining.	97

Figure 3.7. Flow cytometry analysis of Metformin resistance – gating strategy for cell viability staining.	98
Figure 3.8. Representative graph of the total percentage of combined early apoptotic and necrotic cells after 24h of treatment with Metformin.....	100
Figure 3.9. Representative graph of the total percentage of combined early apoptosis and necrosis cells after 48h of Metformin.....	101
Figure 3.10. Representative micrographs showing the morphological influence of Metformin on the different breast cancer cell lines: Luminal B (BT-474), Luminal A (MCF-7), Claudin-low (MDA-MB-231), Basal-like (MDA-MB-468), and HER2 (SkBr3)	102
Figure 4.1. Schematic representation of the main steps involved in a microarray experime.....	109
Figure 4.2. Interactome of the top 100 interactions between the 60 common genes in BASAL-LIKE (MDA-MB-468) and HER2 (SkBr3) cell lines for the two treatments concentrations compared to control. The red arrows represent the upregulated genes, while the blue arrows represent downregulated genes.....	114
Figure 4.3. Interactome of the top 100 interactions between the 48 common genes in BASAL-LIKE (MDA-MB-468) cell line. The red arrows represent up-regulated genes, while the blue arrows represent down-regulated genes.....	117
Figure 4.4. Interactome of the top 100 interactions between the 77 common genes in HER2 (SkBr3) cell line. The red arrows represent upregulated genes, while the blue arrows represent downregulated genes.....	119
Figure 4.5. Schematic diagram showing the second approach of Gene expression microarrays data analysis.....	121
Figure 4. 6. Heat map representing the difference in the gene expression pattern between BASAL-LIKE (MDA-MB-468) cell line samples that were treated with different concentrations of Metformin and compared to control with $F_c \geq -2$, Bonferroni Correction ≤ 0.05 (*)	123
Figure 4. 7. Heat map representing the difference in gene expression patterns between BASAL-LIKE (MDA-MB-468) cell line samples treated with half and double EC50 doses using Regression test, with Bonferroni Correction ≤ 0.05 (*)	126
Figure 4. 8. The heat map represents the difference in gene expression pattern between HER2 (SkBr3) cell line samples using different concentrations of Metformin and compared to control with $F_c \geq -2$, $p < 0.001-0.0001$ (**-.****)	129
Figure 4. 9. Heat map representing the difference in gene expression pattern between HER2 (SkBr3) cell line samples treated with half and double EC50 concentrations and Regression test, with $p < 0.0001$ (****)	132

Figure 4.10. Pathway ontology of the consensus BASAL-LIKE (MDA-MB-468) cell line associated genes.....	136
Figure 4.11. Pathway ontology of the consensus HER2 (SkBr3) cell line associated genes....	137
Figure 5.1. Schematic representation of the standard curve for primers efficiency test.....	143
Figure 5. 2. QRT-PCR data for Basal-like (MDA-MB-468) cell line representing up and downregulated genes, using different doses (A corresponds to half and B to double of EC50 of Metformin and when compared to controls. The threshold was set on 0.1642, while the Tm° varied according to the different primers.....	146
Figure 5.3. QRT-PCR data for HER2 (SkBr3) cell line representing up and downregulated genes, using different doses (A corresponds to half and B to double of EC50 of Metformin and when compared to controls.....	147
Figure 5.4. Representative micrographs of immunoblots showing <i>IRF-9</i> , <i>PTK2B</i> and <i>SERPINB4</i> expression in MDA-MB-468 and MDA-MB-231 (left panel), and SkBr3 and MDA-MB-453 (right panel)	149
Figure 5.5. Micrograph representing the densitometry ratio (AU) of <i>IRF-9</i> protein expression in Metformin untreated and treated (1/2 and 2 EC50) in Basal-like (MDA-MB-468), Claudin-Low (MDA-MB-231), HER2 (SkBr3) and (MDA-MB-453) breast cancer cell lines.....	150
Figure 5.6 Micrograph representing the densitometry ratio (AU) of <i>PTK2B</i> protein expression in Metformin untreated and treated (1mM and 4mM) in MDA-MB-468, MDA-MB-231, SkBr3 and MDA-MB-453 breast cancer cell lines(n=1).	151
Figure 5. 7. Micrographs showing expression of <i>IRF-9</i> in Basal-like (MDA-MB-468) cell line using specific antibodies against <i>IRF-9</i> (red).	152
Figure 5. 8. Micrographs showing expression of <i>IRF-9</i> in Claudin-Low (MDA-MB-231) using specific antibodies against <i>IRF-9</i> (red).....	153
Figure 5. 9. Micrographs showing expression of <i>IRF-9</i> in HER2 (SkBr3) cell line using specific antibodies against <i>IRF-9</i> (red).....	154
Figure 5. 10. Micrographs showing expression of <i>IRF-9</i> in HER2 (MDA-MB-453) cell line using specific antibodies against <i>IRF-9</i> (red).....	155
Figure 5. 11. Micrographs illustrating expression of <i>PTK2B</i> in Basal-like (MDA-MB-468) cell line against <i>PTK2B</i> (green).	156
Figure 5. 12. Micrographs illustrating expression of <i>PTK2B</i> in Claudin-Low (MDA-MB-231) cell line against <i>PTK2B</i> (green).....	157
Figure 5. 13. Micrographs illustrating expression of <i>PTK2B</i> in HER2 (SkBr3) cell line against <i>PTK2B</i> (green).....	158

Figure 5. 14. Micrographs illustrating expression of PTK2B in HER2 (MDA-MB-453) cell line line against PTK2B (green).	159
Figure 5.15. Classification of driver genes in cancer cell signalling and cellular processes.	162
Figure 5.16. Graphs representing qRT-PCR data for (A) Basal-like (MDA-MB-468) and (B) HER2 (SkBr3) cell lines.....	169
Figure 5.17. Representative micrographs of immunoblots showing GADD45A protein expression in MDA-MB-468 and MDA-MB-231 (left panel), and SkBr3 and MDA-MB-453 (right panel).	170
Figure 5.18. Graph representing densitometry ratio AU of GADD45A protein expression in different breast cancer cell lines that were Metformin treated and untreated with 1mM and 4mM concentrations.....	171
Figure 6.1. Micrographs showing <i>EGFP</i> expression in breast cancer cell lines.....	179
Figure 6.2. Representative images PTK2B expression in Control and PTK2B knockdown breast cancer cell lines.	180
Figure 6.3. Micrograph representing the densitometric ratio of PTK2B protein expression in control and shRNA knockdown cells.	181
Figure 6.4. PTK2B depletion prevents cell migration in MDA-MB-468 cells.....	183
Figure 6.5. PTK2B depletion does not affect cell migration in MDA-MB-231 cells.	184
Figure 6.6. PTK2B depletion prevents cell migration in SkBr3 cells.	185
Figure 6.7- PTK2B depletion prevents cell migration in MDA-MB-453 cells.	186
Figure 6.8. PTK2B knockdown prohibited cell invasiveness in MDA-MB-468 cells.	187
Figure 6.9. PTK2B knockdown prohibited cell invasiveness in MDA-MB-231 cells.	188
Figure 6.10. PTK2B knockdown prohibited cell invasiveness in SkBr3 cells.	188
Figure 6.11. PTK2B knockdown prohibited cell invasiveness in MDA-MB-453 cells.	189
Figure 6.12. Micrograph representing the effect of PTK2B downregulation on cell proliferation in MDA-MB-468 cells.	191
Figure 6.13. Micrograph representing the effect of PTK2B downregulation on cell proliferation in MDA-MB-231 cells.	192
Figure 6.14. Micrograph representing the effect of PTK2B downregulation on cell proliferation in SkBr3 cells.....	193
Figure 6.15. Micrograph representing the effect of PTK2B downregulation on cell proliferation in MDA-MB-453 cells.	194
Figure 6.16. <i>PYK2</i> involvement in Autocrine Somatotropin signalling pathway in breast cancer.....	199

Figure 6.17. Differentially expressed proteins from untreated and treated SkBr3 control and *PTK2B* cells.....202

Figure 6.18. Differentially expressed proteins from untreated and treated SkBr3 control and *PTK2B* cells.....207

List of Tables

Table 1.1. Molecular classification of breast carcinoma, associated cell lines and therapies.....	22
Table 1.2. Studies of anticancer potential of Metformin.....	28
Table 1.3. Inventory of the completed clinical trials encompassed breast cancer and Metformin.....	35
Table 2.1. Molecular classification of breast carcinoma.....	59
Table 2. 2. T7 Promoter Master Mix.....	65
Table 2. 3. cDNA Master Mix.....	66
Table 2. 4. RNA Transcription Master Mix.....	66
Table 2. 5. cDNA Master Mix.....	71
Table 2. 6. Real-time quantitative PCR (qRT-PCR) Master Mix.....	72
Table 2. 7. Generation of protein standards.....	74
Table 2. 8. List of the antibodies that used in Western Blot analysis and their dilutions.....	76
Table 2. 9. List of the antibodies that used in Immunofluorescence staining and their dilutions.....	77
Table 2. 10. NanoString probe hybridisation Master Mix.....	78
Table 4.1. The results of the analysis of RNA concentration and integrity.....	108
Table 4.2. Table representing cRNA yield and specific activity results (cRNA yield >5; SP Act>6) for the first set of samples HER2 (SkBr3) which met higher than recommended specific activity and cRNA yields.....	110
Table 4.3. Table representing cRNA yield and specific activity results (cRNA yield >5; SP Act>6) for the second set of samples BASAL-LIKE (MDA-MB-468) which met higher than recommended specific activity and cRNA yields.....	111
Table 4.4. BASAL-LIKE (MDA-MB-468) and HER2 (SkBr3) associated top 12 driver genes, for the two treatment concentrations compared to control	115
Table 4.5. BASAL-LIKE (MDA-MB-468) associated top 12 driver genes for the two treatment concentrations compared to control.....	118
Table 4.6. HER2 (SkBr3) associated top 12 driver genes for the two treatment concentrations compared to control.....	120
Table 4. 7. List of genes upregulated and downregulated in BASAL-LIKE (MDA-MB-468) cell line.....	124
Table 4. 8. List of upregulated and downregulated genes in BASAL-LIKE (MDA-MB-468) cell line.....	127

Table 4. 9. List of upregulated and downregulated genes in HER2 (SkBr3) cell line.....	130
Table 4. 10. The list of genes upregulated and downregulated in HER2 (SkBr3) cell line.....	132
Table 4.11. The SUM of rank for commonalities.....	133
Table 4.12. List of genes that were selected for validation by qRT-PCR.....	134
Table 4.13. List of genes that were selected for validation by qRT-PCR.....	134
Table 5.1. Percentage of efficiency and amplification dynamic range.....	144
Table 5.2. The top 20 significant P value genes and corresponding pathways.....	163
Table 5.3. The top 20 significant P value genes and corresponding pathways.....	164
Table 5.4. The top 20 significant P value genes and corresponding pathways.....	166
Table 5.5. The top 20 significant P value genes and corresponding pathways.....	167
Table 5.6. Representing percentages of efficiency and dynamic range of GADD45 A, B, and G primers.....	168
Table 6.1. List of <i>PYK2</i> engaged Pathways in both Normal and Pathological conditions.....	195
Table 6.2. list of biological processes that are related to <i>PYK2(FAK2)</i> gene.	200
Table 6.3. Top 25 upregulated proteins and 25 downregulated proteins in HER2 (SkBr3) cell line.....	203
Table 6.4. Top 25 upregulated proteins and 25 downregulated proteins in HER2 (SkBr3) cell line.....	205
Table 6.5. Top 25 upregulated proteins and 25 downregulated proteins in HER2 (SkBr3) cell line.....	208
Table 6.6. Top 25 upregulated proteins and 25 downregulated proteins in HER2 (SkBr3) cell lin.....	210
Table 6.8. Pathways regulated by proteins that are differentially expressed between treated and untreated PLKO-1 and <i>PYK2</i> KD cells (SkBr3 cell line)	213

ABBREVIATIONS

ABC8	ATP Binding Cassette Subfamily B Member 8
ACC	Acetyl-CoA carboxylase
AD	Atopic Dermatitis
Akt	serine/threonine kinase
ALS	Acid Labile Subunit
AMPK	Adenosine Monophosphate-activated Protein Kinase
ANNs	Artificial Neural Networks
Arg	Abl-related gene
AU	Arbitrary Unit
BRCA1 & BRCA2	Breast Cancer genes
BCL2	B cell lymphoma 2
BET1P1	Bet1 Golgi Vesicular Membrane Trafficking Protein Pseudogene 1
Bim	Bcl-2 interacting mediator of cell death
BM	Basement Membrane
Bmf	Bcl-2 modifying factor
BNs	Bayesian Networks
CaMKK	Calcium/calmodulin-dependent protein Kinase
CCK	Gastrin and cholecystokinin
CK	Cytokeratin's
COX-2	Cyclooxygenase-2
CSCs	Cancer Stem Cells
Ct	Threshold Cycle
DAPK	Death-Associated Protein Kinase
DAPK3	Death-Associated Protein Kinase 3
DCIS	Ductal Carcinoma in Situ
DDFS	Disease-Free Survival
Debtor	Domain-containing mTOR-interacting protein
DLD	Dihydrolipoamide Dehydrogenase
DPBS	Dulbecco's Phosphate Saline
DTs	Decision Trees
DTT	Dithiothreitol
ECM	Extra Cellular Matrix
EF1-α-EGFP	Eukaryotic translation elongation factor 1
EGFP	Enhanced Green Fluorescence Protein
EMEM	Minimum Essential Medium Eagles
EMT	Epithelial to Mesenchymal Transition
ER	Oestrogen Receptor
ERK	Extracellular signal-Regulated Kinases
F & R	Forward and Reverse primers
FBS	Fetal Bovine Serum
Fc	Fold change
FDA	Food and Drug Administration
FDR	False Discovery Rate

FFPE	Formalin-fixed Paraffin-embedded
FGFs	Fibroblast growth factors
FOXA1	Hepatocyte Nuclear Factor 3 Alpha
GATA3	GATA Binding Protein 3
GFs	Growth Factors
GH	Growth hormone
GO	Gene Ontology
GPCR	G protein-coupled receptors
GPCR	G-protein coupled receptors
GSEA	Gene Set Enrichment Analysis
GUSB	Glucuronidase Beta
HEGFR	Human Epidermal Growth Factor Receptor
HPV	Human Papillomavirus Virus
IDA	Independent Data Acquisition
ICD-10	International Classification of Diseases Tenth Revision
IF	ImmunoFluorescence
IGF	Insulin-like Growth Factor
IGF-1	Insulin-like growth factor-1
IGF1/2	Insulin-like growth factor 1 or 2
IGF-1R	Insulin-like growth factor 1 receptor
IL-1β	Inter Leukin 1 beta
IL-3	Interleukin 3
INSR	Insulin Receptor
IR	Insulin Receptor
JAK2	Janus kinases
KD	KnockDown
KEGG	Kyoto Encyclopedia of Genes and Genomes
KIAA0232	KIAA0232
LCIS	Lobular Carcinoma In situ
LKB1	liver kinase B1
M	produced in Mice
MA	Microarrays
MAPK	Mitogen-Activated Protein Kinase
MCCV	Monte Carlo Cross Validation
MDM2	Mouse double minute 2 homolog
MERS	membrane oestrogen receptors
ML	Machine learning
mLST8	mammalian lethal with SEC13 protein 8
MS	Mass Spectrometry
MSE	Mean Squared Error
mTOR	mammalian Target of Rapamycin
n	Number of cases
N F W	Nuclease-Free Water
NADPH	Nicotinamide Adenine Dinucleotide Phosphate Hydrogen
NLP	Natural language Processing
NSAIDs	Non-Steroidal Anti-Inflammatory Drugs

OGP	N-Octyl-Beta-Glucopyranoside
OS	Overall Survival
P	Parameters
PAI-1	Plasminogen Activator Inhibitor-1
PCA	Principle Component Analysis
PCOS	Polycystic Ovary Syndrome
PI	Protease Inhibitors
PI3K	The phosphatidylinositol 3-kinase
PIP3	Phosphatidylinositol-3,4,5-trisphosphate
PR	Progesterone Receptor
PRAS40	Proline-Rich AKT substrate 40 kDa
PS	phosphatidylserine
PTEN	phosphatase and tensin homolog
PTK2B	Protein Tyrosine Kinase 2 Beta
R	produced in Rabbit
RAPTOR	Regulatory Associated Protein Of mTOR
Rb	Retinoblastoma-associated
RTK	Receptor Tyrosine kinase
Skida1	SKI/DACH Domain Containing 1
SLC39A6	Solute Carrier Family 39 Zinc Transporter Member 6
SR	Steroidal hormone nuclearReceptor
SVMs	Support Vector Machines
SWATH	sequential window acquisition of all theoretical fragment ion spectra
T2DM	Type 2 Diabetes Mellitus
TGF-β	Transforming Growth Factor β
TKIs	Tyrosine kinase Inhibitors
Tm	Melting Temperature
TNBC	Triple-Negative Breast Cancer
TNF	Tumour Necrosis Factor
TNFα	Tumour Necrosis Factor alpha
TNM	TNM Classification of Malignant Tumours
TSC2	TSC Complex Subunit 2
VEGF	Vascular endothelial growth factor
VEGF	Vascular Endothelial Growth Factor
WT1	Wilms's tumour suppressor
XBP1	X-box Binding Protein 1
YLPM1	YLP Motif Containing 1
ZIP6	Zinc Transporter

Abstract

Breast cancer is a complicated, heterogeneous and diversified disease that comprises of a mixture of various subtypes. The emergence of high throughput technologies such as gene expression profiling and DNA copy number analysis has allowed a profound awareness of this complex disease. Breast cancer molecular classification began with ER, PR and HER2 based stratification of patients, and further classified into various intrinsic subtypes such as Luminal A, Luminal B, Her-2 enriched, Basal-like, and Claudin-low. These subtypes were proven to have significant divergences in amplified and mutated genes, survival rates, prognosis and response to therapy.

The anti-diabetic drug Metformin has been prescribed to treat type 2 diabetes patients since 1957 with a well-established side effects and safety profile. Additionally, cancer researchers have studied the anti-tumour effects of this drug since 2005 and have determined Metformin benefit in the reduction of cancers incidence. Thus, Metformin has been suggested as an ideal candidate to treat and prevent diverse types of cancers, including breast carcinoma.

The purpose of this study is to elucidate the biological and molecular effects of Metformin on breast cancer cell lines such as BT-474, MCF-7, MDA-MB-231, MDA-MB-468, and SkBr3.

Interestingly, Metformin treatments reduced the viability and proliferation of examined breast cancer cell lines, while induced cells apoptosis. Basal-like (MDA-MB-468) was the most sensitive phenotype to Metformin treatments, whereas HER2 (SkBr3) was the least sensitive subtype.

We performed Gene Expression Microarray and NanoString analysis of Metformin treated MDA-MB-468 and SkBr3 cells and found that the upregulated Protein Tyrosine Kinase 2 Beta (*PTK2B*) was predominantly involved in cell proliferation, cell survival, cell migration and cell invasion.

We found that *PTK2B* promotes invasion and migration, while prevents the proliferation of breast cancer cells. It is also, playing a vital role in Autocrine Somatotropin signalling pathway in breast cancer. Additionally, Mass Spectrometry (MS) analyses and pathways enrichment analysis that has emphasised the role of *PTK2B* in breast cancer invasion and metastasis. Additionally, the combined action of both selected markers and Metformin treatment on fundamental biological functions in breast cancer were also assessed.

These data showed that Metformin promotes Her-2 enriched breast cancer invasion through mechanisms involving PYK2, and that future treatments should consider potential complications resulting from metformin-based therapies.

CHAPTER 1

Introduction

1.1. Cancer

Cancer is a general term that describes a large group of diseases, which affects any part of the body. Other terminologies have been used to describe cancer including, malignant tumours and neoplasms. Cancer is a disease caused by an uncontrollable division of abnormal cells in a part of the body. Sometimes cancer cells can spread to other parts of the body beyond their usual boundaries, which is called a secondary tumour or a metastasis. This process of dissemination, which is known as metastasis has been recorded as the major cause of cancer death. More than 200 distinct types of cancer are occurring, and 1 in 2 people in the UK will develop cancer during their lives. Among women, the most common types of cancers are breast, colorectal, lung, uterine cervix, and stomach cancer. However, in men Lung, prostate, colorectal, stomach, and liver cancer are the most common types of cancer. The increase of cancer survival rates can be achieved by early detection, correct diagnosis, and effective treatment (<https://www.who.int/cancer>), (<http://www.cancerresearchuk.org/>).

Intensive research efforts have been directed to improve our understanding of this complex disease and have resulted in some improvement in treatment and the survival rates of cancer. For instance, the US Food and Drug Administration (FDA) has approved Gardasil 9, as a new vaccine for Human Papillomavirus (HPV), a virus that causes cervical cancer in women. Likewise, Imatinib (Gleevec) is a targeted cancer drug (biological therapy) that was approved in 2001, as targeted treatment of the rare chronic myelogenous leukaemia. Moreover, during the past few years, there has been a growing emphasis on early cancer detection, neoplastic stem cells, microRNAs and cancer biomarkers. However, the effect of cancer on individuals' life is still high, and the incidence of cancer is predicted to double within the next 20 years due to the ageing global population (*Elsevier Community, 2016*). Despite advances, there is still a great need for further research into areas which are focused on understanding the molecular mechanisms underlying cancer initiation and progression. Considering the many types of cancer and taking into account the genetic diversity of patients, there are still multiple routes that need to be explored to promote more effective, and maybe even more personalised, treatments.

1.2. Cancer Statistics for 2018

Cancer is a global and considerable health problem as reflected by its yearly increasing rate of morbidity and mortality. Worldwide, an estimated 14.1 million people were diagnosed with cancer and 8.2 million related deaths in 2012. The UK alone recorded about 360,000 new cases of cancer in 2015, and around 164,000 deaths from cancer in 2016, which makes it the first leading cause of death in the UK, above heart disease and dementia. In the UK, a cancer is diagnosed every two minutes and a related death occurs every four minutes. In addition, more than 980 people are diagnosed with cancer, and more than 440 people die from this disease every day. Since the early 1990s, the incidence rates were increased by 12%, consequential to a growing and ageing population. Moreover, the estimated number of cancer cases are projected to increase by more than 40% to about 514,000 new cases per year in 2035, with a great rising in men than women (Figure 1.1) (*Cancer Research UK, 2018*).

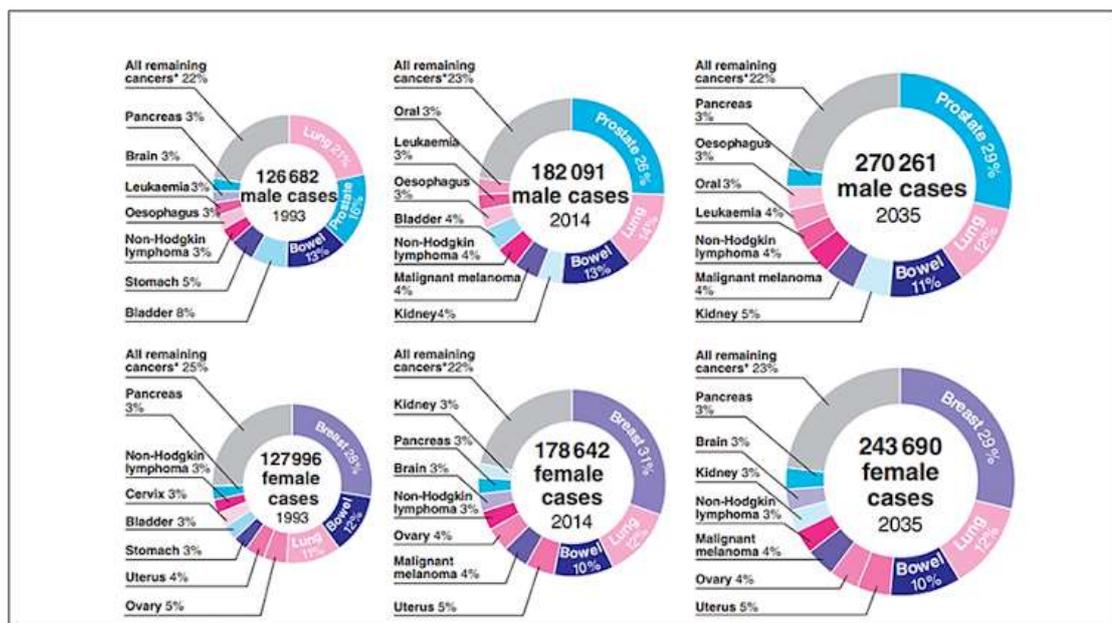


Figure 1.1. Percentage of total cancer incidents by cancer site comparing the increasing rates of cancer cases in three different periods in 1993 (observed), 2014 (observed) and 2035 (estimated), and divided by sex. Adapted from (*Smittenaar et al., 2016 via Cancer Research UK September 2018*).

In the UK, the most commonly diagnosed cancers are breast, prostate, lung and bowel cancers, which account for more than a half (53%) of the total number of cancer cases in 2015. About a fifth of all cancer deaths is caused by lung cancer. The Figure below describes the 24 different cancers in major sites of the body, for both females and males. The International Classification of Diseases Tenth Revision (ICD-10) was used to classify cancer sites, and the Non-melanoma

skin cancers (ICD-10 C44) were excluded. Moreover, all ages were combined in this statistical bulletin, according to the National Cancer Registration and Analysis Service within Public Health England; Office for National Statistics, (*Cancer Research UK, 2018*).

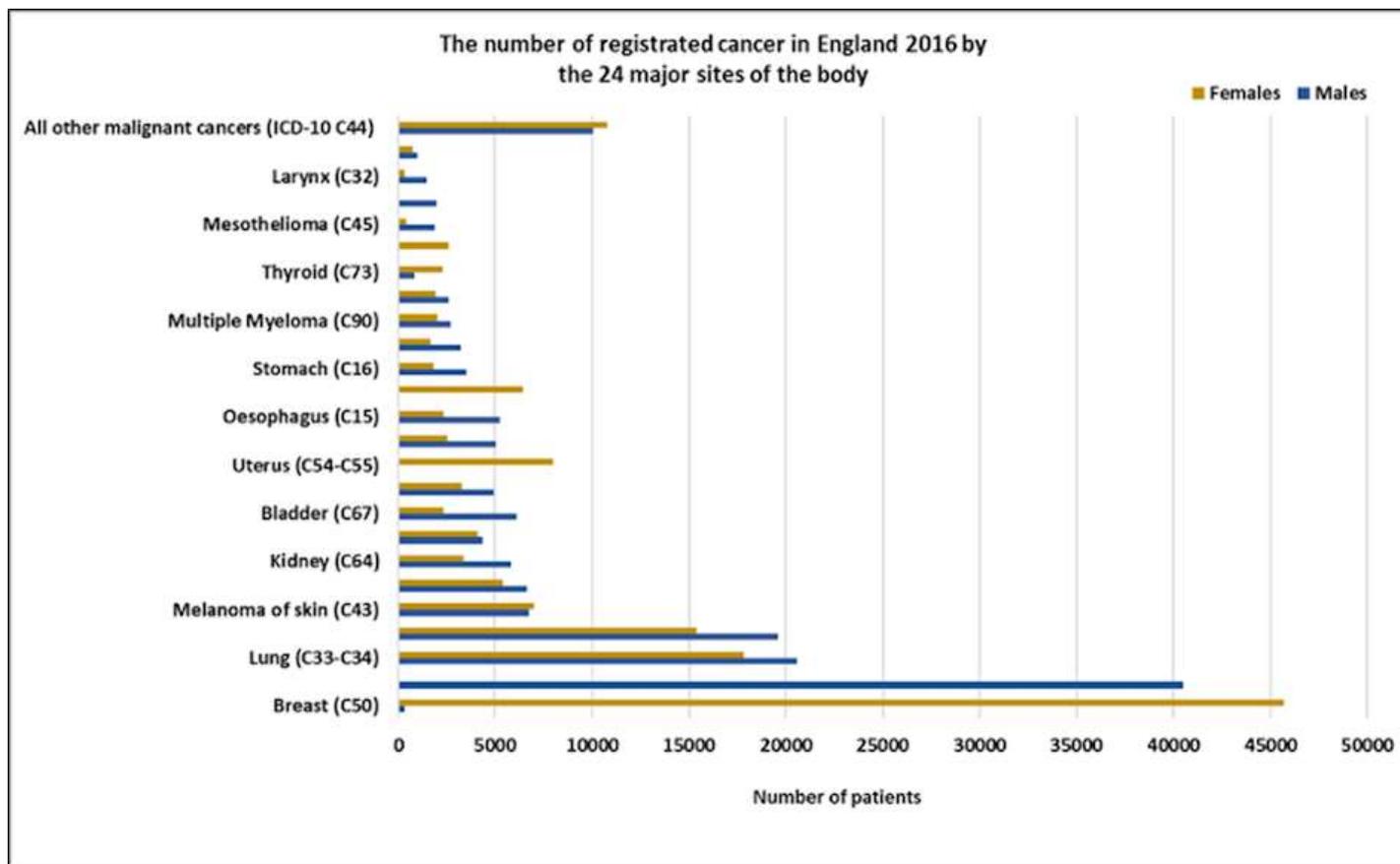


Figure 1.2. Cancer diagnoses incidence rates in the UK 2016. The diagram represents the number of diagnosed male patients (in blue) and female patients (in yellow) for each type of organ-specific cancer. This data refers to the year 2016 by National Cancer Registration and Analysis Service within Public Health England; Office for National Statistics (*Cancer Research UK, 2018*).

In addition, there is a clear correlation between cancer incidence and mortality with ageing. In 2013 – 2015, incidence rates were highest in people aged 85 – 89 for all cancers and for both males and females. During this period, more than a third (36%) of all cancer cases in the UK were diagnosed in people aged 75 and over each year, and more than half (53%) of all cancer deaths in the UK are related to people aged 75 and over. Moreover, significant divergences were noticed between sexes across age groups. Cancer incidence rates were higher in females aged 15 to 59 years compared with males in the same age group, and more than doubled in females aged 40 to 44 years. That is highly likely related to the highest incidence rates of breast cancer which registered among females aged 30 to 59. Whilst cancer rates were higher in males aged 60 years and over, which can be explained by the incidence of prostate cancer in males aged 65 to 79 years (*Cancer Research UK, 2018*). The diagram below illustrates the increase in cancer incidence according to getting old. The source of this data is the National Cancer Registration and Analysis Service within Public Health England; Office for National Statistics, (*Cancer Research UK, 2018*).

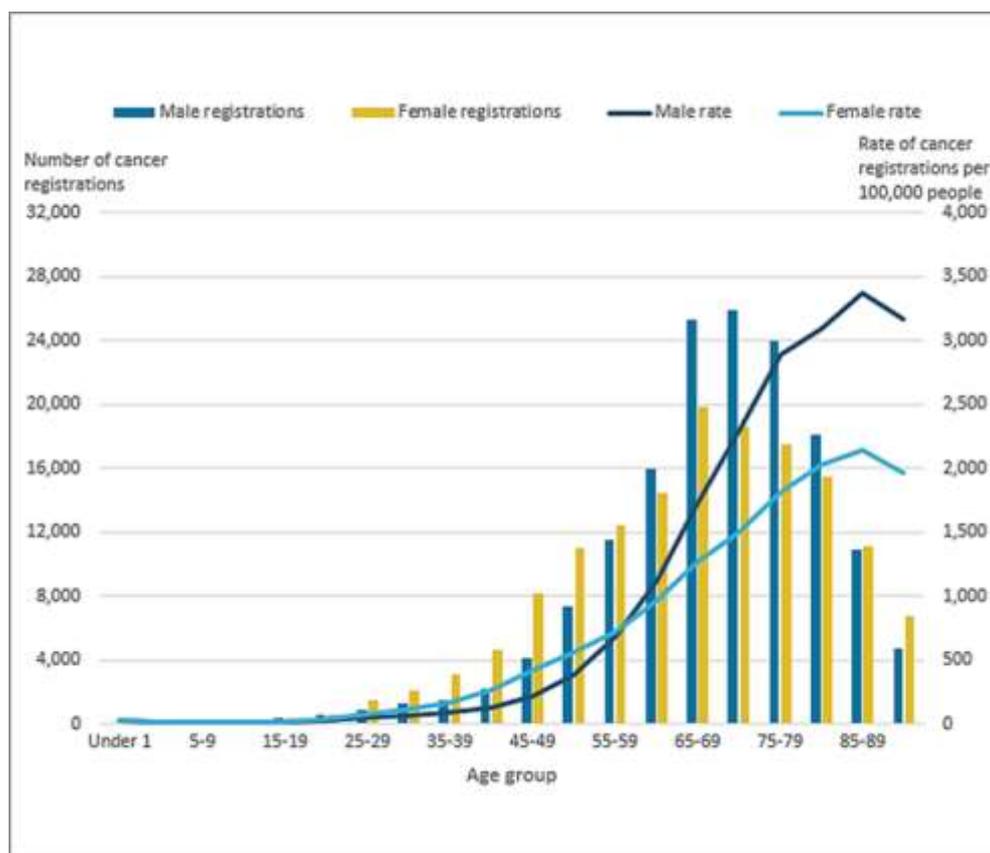


Figure 1.3. The number of cancer registrations in England, 2016. The diagram shows the age-specific cancer incidence rates (per 100,000) and the differences between males and females for each age group. This statistical bulletin excludes non-melanoma skin cancer (ICD-10 C44). The Figure was adapted from the National Cancer Registration and Analysis Service within Public Health England; Office for National Statistics 2016 (*Cancer Research UK, 2018*).

Even though cancer incidence rates are showing increasing trends over time, the mortality rates of cancer are continuously decreasing as well. In England, the death from cancer was accounted

for 28.5% of all registered deaths in 2016. However, the highest annual number of cancer deaths in England and Wales since 2003 was recorded in 2017, with a 1.6% increase in the number of deaths. This number is expected to be elevated as a consequence of increasing in both size and age of the population over the time. The Figure below 1.4. shows the total number of deaths in England and Wales from 2001 to 2017 for both sexes (*Cancer Research UK, 2018*).

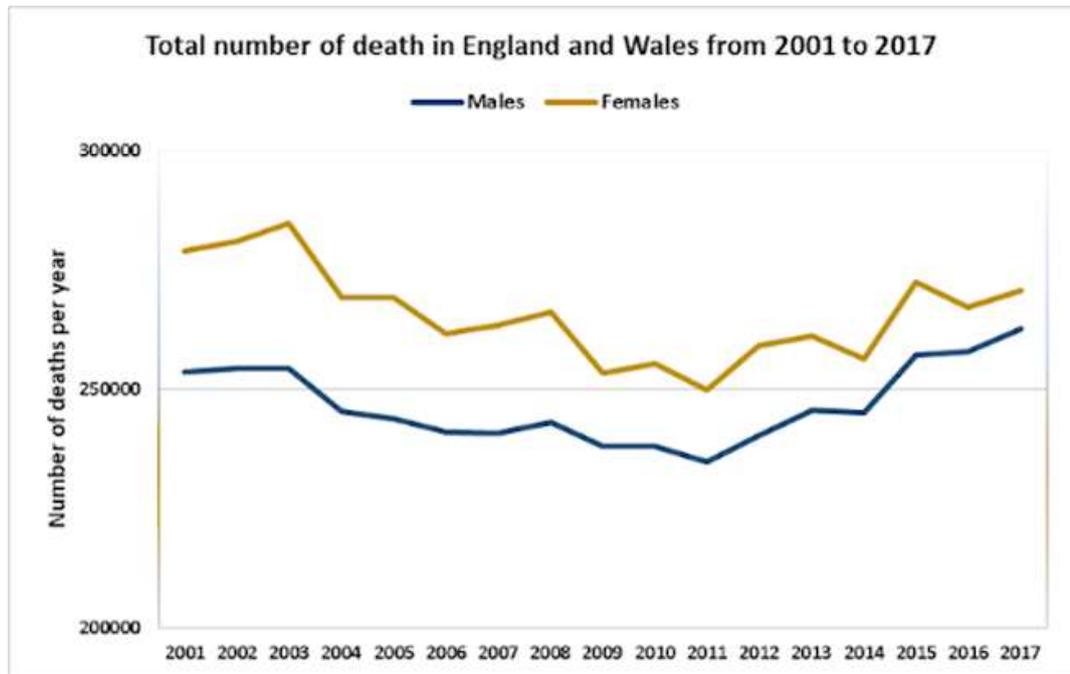


Figure 1.4. The total number of deaths from cancer in England and Wales from 2001- 2017. The trend lines are displaying cancer deaths number in men and women over different years. The highest annual number of deaths was in 2017 according to the office of National Statistics (*Cancer Research UK, 2018*).

Contradictory, cancer survival has improved and has doubled in the last 40 years in the UK. Generally, cancer survival is higher in people who diagnosed aged under 40 years old, except for breast, bowel and prostate cancers, where survival is highest in middle age. The percentage of cancer survival was half for 10 or more years in 2010-2011 for England and Wales. Survival rates vary between different cancer types, which ranged from 98% for testicular cancer to just 1% for pancreatic cancer. In addition, cancer survival in women is higher than in men. Moreover, 38% of cancer cases were prevented in 2015. The reduction in mortality rates and the growing number of surviving patients correlated with progress in technologies and medications. Besides, the improvement in early detection of cancers has reduced the number of people who died from cancer (*Cancer Research UK, 2018*).

Despite the remarkable progress in cancer research, which has been made over the last few decades, the principles behind cancer induction are still not fully interpreted. Generally, cancer is described as accumulative mutations and epimutations, that result in cellular genomic and epigenomic alterations. It is also defined as an uncontrollable proliferation of any type of cells in the body. The genetic principles and pathways involved in the generation of cancer cells have not been satisfactorily elucidated. DNA mutations result in considerable damage in mitotic events that leads to a disordered proliferation in malignant cells. However, exposure to environmental agents including carcinogens, radiations and pathogenetic infections, may contribute to the initiation of most types of cancer through different mechanisms (*Belizário, 2018*). Besides, most known cancer risk factors are ageing, smoking, sun and radiation exposure, chemicals and other substances, certain hormones, family history of cancer, excessive alcohol consumption, poor diet, lack of exercise, and excess body weight. Treating cancer can be achieved by preventing it through making changes in individuals' lifestyle. The two largest preventable causes of cancer in the UK are smoking and obesity (*Cancer Research UK, 2018*).

1.3. Hallmarks of cancer

The genesis of cancer is known as carcinogenesis, oncogenesis or tumorigenesis in which normal cells switch to cancer cells. This process can be characterised by alterations in cellular, genetic, epigenetic levels and could also be due to abnormal cell division. The multistep development process of human cancers (from normal cells to cancer cells) is governed by eight common traits "hallmarks" and two Enabling Characteristics, which are combined between all cancer types. A peer-reviewed article published in 2000 by *Hanahan and Weinberg* entitled "The Hallmarks of Cancer", highlighted the six common "hallmarks", which allow cancer cells to survive, proliferate and invade (*Hanahan and Weinberg, 2000*). Then, followed by 'Hallmarks of Cancer: The Next Generation' in 2011 by *Hanahan and Weinberg* highlighted two new Enabling Characteristics, which are Genomic Instability, and Tumour-Promoting Inflammation. They also added two Emerging Hallmarks that are Reprogramming of Energy Metabolism and Evading Immune Destruction (*Hanahan and Weinberg, 2011*).

The main hallmarks of cancer include:

1.3.1. Maintaining proliferation signalling

Sustaining proliferative signalling enables cancer cells to proliferate and multiply in the absence of stimulation signals from external growth factors (GFs). Typically, normal cells of the body need stimulation by external mitogenic growth signals, which bind into cell surface receptors to switch the cells from a quiet status to an active proliferative status. In addition, cell division in normal cells is tightly controlled. However, cancer cells have this capability to proliferate in an uncontrollable way by producing the growth signals by themselves, in which permanently triggers the signalling pathways that respond to these signals, or by destroying 'off switches' that are preventing immoderate growth from these signals (*Hanahan and Weinberg, 2000*).

1.3.2. Avoiding growth suppressors

To support cellular dormancy and tissue homeostasis, multiple growth suppressors such as Retinoblastoma-associated (Rb) and p53 proteins signal operate within the healthy tissue as central control points within two critical complementary cellular regulatory circuits that control the decisions of cells to proliferate or activate senescence and apoptotic programs. Those signals include soluble growth inhibitors and immobilised inhibitors, which have been embedded onto the surfaces of nearby cells and into the extracellular matrix (*Sherr, 2004*). Generally, cancer cells are impervious to antigrowth signals from the surrounding environment by altering their tumour suppressor proteins. An alternative way to prevent cells from over-division in normal cells is known as "contact inhibition", which cancer cells do not have, so they keep dividing even if there is a limited body cavity space. Evading growth suppressor in cancer cells can be achieved by strengthening the adhesion of cadherin-mediated attachments between cancer cells (*Curto et al. 2007; and Hanahan and Weinberg, 2011*).

1.3.3. Avoid apoptosis (programmed cell death)

Apoptosis is the mechanism by which cells are programmed to die as a regular and controlled part of an organism's growth or development, or as a consequent of experiencing damages. Cancer cells confront different physiological stresses during tumorigenesis, including imbalances in the cells signalling and DNA damages that result in the enhancement of cell death. The apoptotic machinery can be bypassed in cancer cells even if the cells are becoming grossly abnormal due to several mechanisms such as inactivation of the p53 tumour suppressor gene (*Horn and Vousden, 2007*). Or through increased expression of anti-apoptotic regulators (BCL2) or by the PI3 kinase–AKT/PKB pathway, which transmits antiapoptotic survival signals, this

pathway activates by extracellular factors such as IGF-1/2 (Insulin-like growth factor 1 or 2) or IL-3. Or by intracellular signals released from Ras, or by loss of the tumour suppressor pTEN, which is a phospholipid phosphatase that normally reduces the AKT survival signal (*Hanahan and Weinberg, 2000*).

1.3.4. Enabling replicative immortality

Normal mammalian cells have an intrinsic program, which is known as "Hayflick limit" that enables a limited number cell division before reaching senescence (non-proliferative but the viable state) stage or crisis (cell death). The sensor of this doubling programme is the telomere, which is a non-coding tandem hexanucleotide sequence repeats DNA that is found at the end of the chromosome. This telomere decreases in size during each cell cycle (division) till it becomes too small, resulting in induced senescence. Cancer cells are capable of growing indefinitely and achieving immortality by escaping this limit (*Hanahan and Weinberg, 2000; Collado and Serrano, 2010*).

1.3.5. Inducing angiogenesis

Normal tissues need oxygen and nutrients supplied by blood vessels that are crucial for cell function and survival. The formation of new blood vessels is an exclusive event during the development of embryos, wound repair, and during the female reproductive cycle. In cancer tissues, the most obvious feature is the formation of new blood vessels that allow tumour access to nutrients and oxygen, and for evacuating waste metabolic products and carbon monoxide. This process is constitutively active during tumour progression to help sustain neoplastic growth (*Bertram, 2000*). An increased transcription of angiogenic regulators causes the induction of angiogenesis in tumours and include the vascular endothelial growth factor (VEGF) and/or fibroblast growth factors (FGFs) (*Ferrara, 2009; Bertram, 2000*).

1.3.6. Promoting tissue invasion and metastasis

One of the characteristic features of cancer cells is their ability to invade adjacent tissues and thence spread to distant sites where they might find new supplies of un limiting nutrients and space to colonise, after leaving their primary tumour mass. Metastasis is the process of physical dissemination of cancer cells from a primary site of cancer to a secondary malignant growth at a distance, and their adaptation to the new microenvironment of hosting tissues. The multistep process of metastasis includes local invasion of surrounding tissues, then intravasation by cancer cells into the nearby blood and lymphatic vessels with survival within the harsh environment of the circulatory system, followed by crossing through the lymphatic and hematogenous systems, and then escape of cancer cells from the lumina of such vessels into the parenchyma of distant tissues (extravasation). Then, the formation of small nodules of

cancer cells (micrometastases), and finally the growth of a micrometastatic tumour into macroscopic tumours, which is termed as “colonisation.” It is now well documented as a scientific knowledge that cancer cells acquire migratory abilities through the process of epithelial to mesenchymal transition (EMT) (*Kalluri and Weinberg, 2009; Lamouille et al. 2014*). EMT is a process that encompasses substantial changes in cell morphology, acquired migratory capabilities, and the abilities to invade and evade apoptosis of cancer cells. To that end, multiple transcription factors including Snail, Slug, Twist and Zeb1/2 are required to promote the acquisition of characteristic traits such as the expression of specific signalling membrane proteins, reorganisation of cytoskeletal proteins, increased production of molecules that enable migration and significant changes in the expression of particular microRNAs (*Hanahan and Weinberg, 2000; Thiery and Sleeman, 2006; Lamouille et al. 2014*).

1.4. Additional Hallmarks of Cancer

Two more emerging hallmarks of cancer, deregulating or changing and reprogramming cellular metabolism and evasion of the immune response by immune cells (T and B lymphocytes) were added to this list after an increased body of research. One is involved in sustaining and supporting the neoplastic proliferation growth, while the second enables the cancer cells to escape immunological destruction by T and B lymphocytes, macrophages, and natural killer cells. Two other enabling properties that allow cancer cells to survive and proliferate, development of genomic instability and tumour stimulated inflammation were also added to the previous hallmarks of cancer list. Genomic instability development that allows mutations in cancer cells with genetic alterations to drive the progression of cancer was the first property. The second attribute is tumour-promoting inflammation via the innate and adaptive immune systems that contribute to multiple hallmark capabilities by supplying proliferative growth factors, EMT inducing signals as well as extracellular matrix changing enzymes to the tumour microenvironment which helps in angiogenesis, invasion, and metastasis (*Hanahan and Weinberg, 2011*). These ten biological capabilities that are acquired during the multistep development of tumours are known as “Cancer Hallmarks” are illustrated in Figure 1.5.

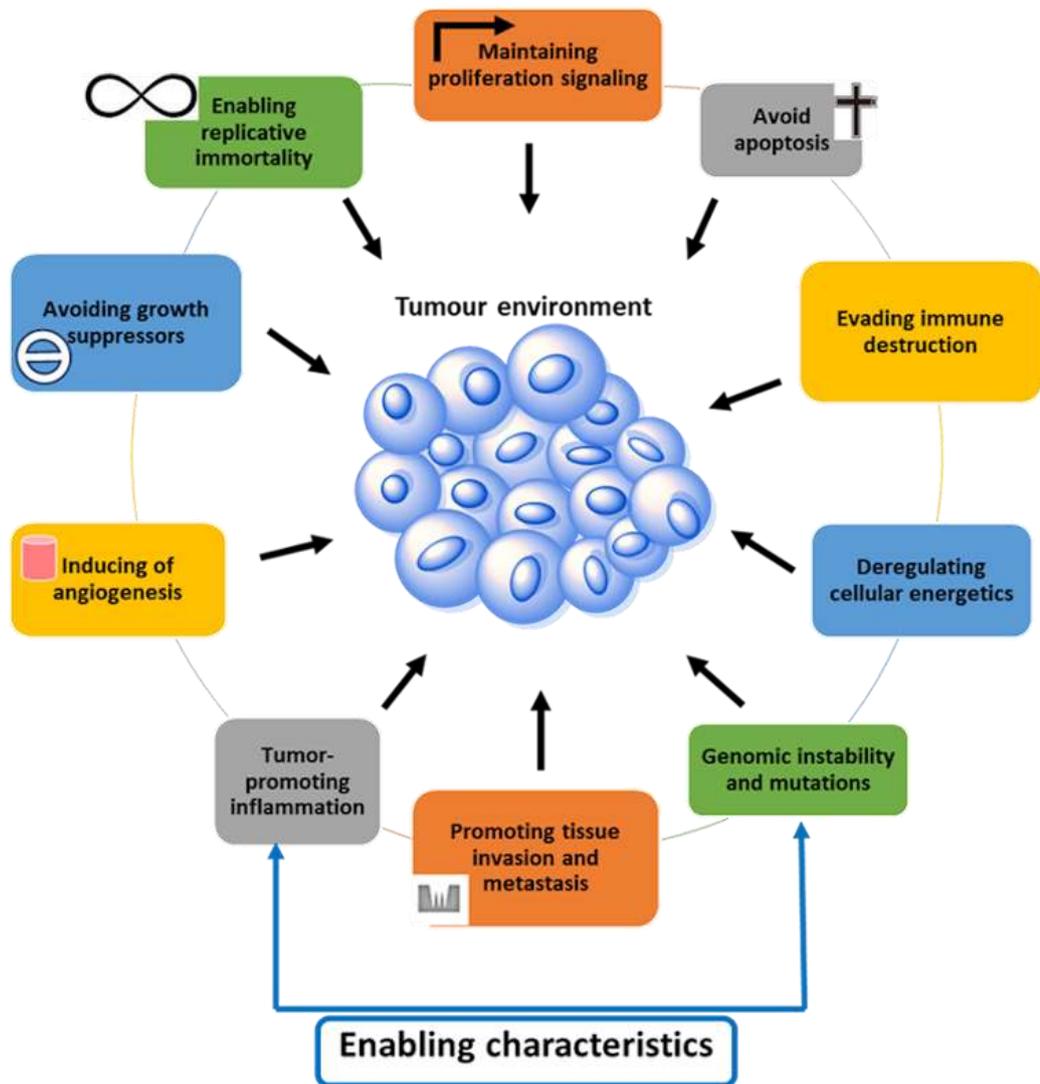


Figure 1.5. The hallmarks of cancer. The schematic representation depicts the six original hallmarks of cancer as proposed by Hanahan & Weinberg (*Hanahan & Weinberg in 2000*). Additionally, *Hanahan and Weinberg* in their next generation of "Hallmarks of Cancer" in 2011, have listed two new emerging hallmark characteristics, including Evading immune destruction and Reprogramming of energy metabolism (*Hanahan and Weinberg, 2011*). In addition, the other two subordinate enabling properties that allow cancer cells to survive and proliferate are the Development of genome instability and Tumour-promoting inflammation. *This image was adapted from (Hanahan and Weinberg, 2011).*

1.5. Breast cancer

Breast cancer is the most commonly diagnosed cancer in the West and the predominant malignant tumour in females, which accounts as one of the main causes of cancer death and its incidence rate is increasing throughout the globe (Masumi, et al., 2012; Arau 'jo et al., 2017). Breast cancer represented the highest registered cancer among other types of cancer in women in the UK. About 1 in 7 (14.2%) women and 1 in 870 (0.001%) men in the UK will develop breast cancer during their lives (<http://www.cancerresearchuk.org/>). Approximately 55,200 people are diagnosed with breast cancer in the UK each year, which is around 150 people a day. An estimated 1.38 million new cases and about 458,000 deaths from breast cancer worldwide are reported every year (Taherian-Fard, et al., 2014). In England (in 2016), mortality rates were decreased, in contrast to increased incidence rates, showing that the number of patients who survived from breast cancer has improved. The advantage of advancement in medication, technology and early diagnostic allowed the reduction of the number of people who die from cancer (Cancer Research UK, 2018). The incidence and mortality rates of breast cancer are shown in Figure 1.6 below represented by European Age-Standardised Rates per 100,000 Females, UK, 1993-2016.

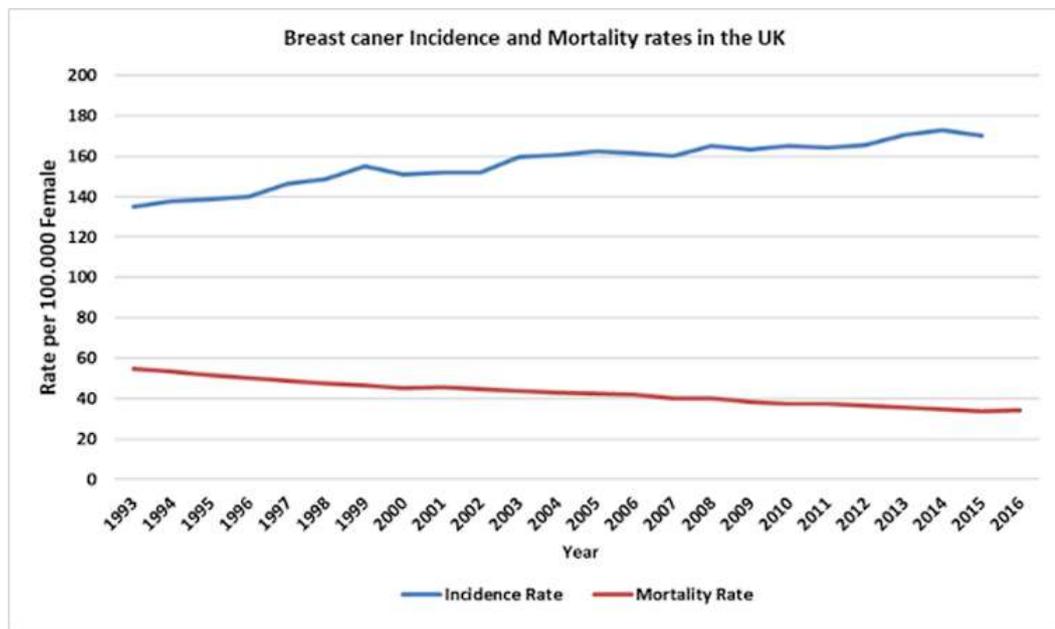


Figure 1.6. Age-standardised cancer mortality and incidence rates for female breast cancer in the UK in 2016. The graph shows the incidence and mortality trends in breast cancer, which is the most common cancer in females. It also displays rates per 100,000 persons from 1993 to 2016. The source of this data is the National Cancer Registration and Analysis Service within Public Health England; Office for National Statistics (Cancer Research UK, 2018).

1.6. The Mammary Gland Structure and development

The mammary gland (breast) is a unique anatomical structure that secretes milk during lactation stage. It is a complex secretory organ, composed of glandular (epithelium) and stromal (parenchyma) tissues (Figure 1.7). The stroma physically and nutritionally supports the epithelium. The stroma is formed by fatty cells, connective cells, blood vessels and lymphatic vessels, which influence the development of the mammary gland. The architecture of a mature human epithelium resembles a tree-like branching system with about 15 to 20 different lobes (milk producers). The lobe, which is also known as an acinus is assembled by several terminal lobular units of secretory alveoli and the converging duct (milk carriers). The epithelium is composed of a bi-layer epithelial cells structure (luminal and basal). The luminal epithelium forms the ducts and the secretory alveoli. The luminal epithelium also contains a population of cells that characterised by their hormone receptor status. The basal epithelium comprises of myoepithelial cells that generate the outer layer of the gland, in addition to a small population of stem cells, which supply the different cell types. (Macias and Hinck, 2012; Cyr et al., 2016).

The mammary gland develops through three major stages, which are embryonic, pubertal and reproductive. Embryonic development proceeds in the absence of hormone regulation, while customised hormonal inputs control the pubertal and reproductive stages. During embryogenesis, local epithelial/mesenchymal interactions are directing many developing processes and are responsible for determining the location of the mammary gland and also control cell fate, such that at birth the tissue compartments are precisely placed, and the budding structure is accurately established. Epithelial/mesenchymal interactions have also controlled the development during postnatal stages of mammary gland morphogenesis. However, the nature of these interactions will enormously change after puberty by regulation hormones and growth factors. The hormones generate complex signalling networks that impact the epithelial/mesenchymal interactions by adjusting the production of secondary signalling pathways that drive cross-talk between and within compartments. The subsequent development stages such as pubertal growth, pregnancy, lactation and involution – occur postnatally are under the hormonal control. The branching morphogenesis initiated in puberty and require the growth hormone, oestrogen, and IGF1, to generate a ductal tree and fill the fat pad. During pregnancy, the joint action of progesterone and prolactin are fundamental for generating the alveoli that secreted milk during lactation. At weaning period when no more milk is needed, the process initiates involution whereby the gland is reconstructed back to its pre-pregnancy status. Many signalling pathways, which have distinct regulatory functions at different stages of gland development are needed during this phase (Macias and Hinck, 2012).

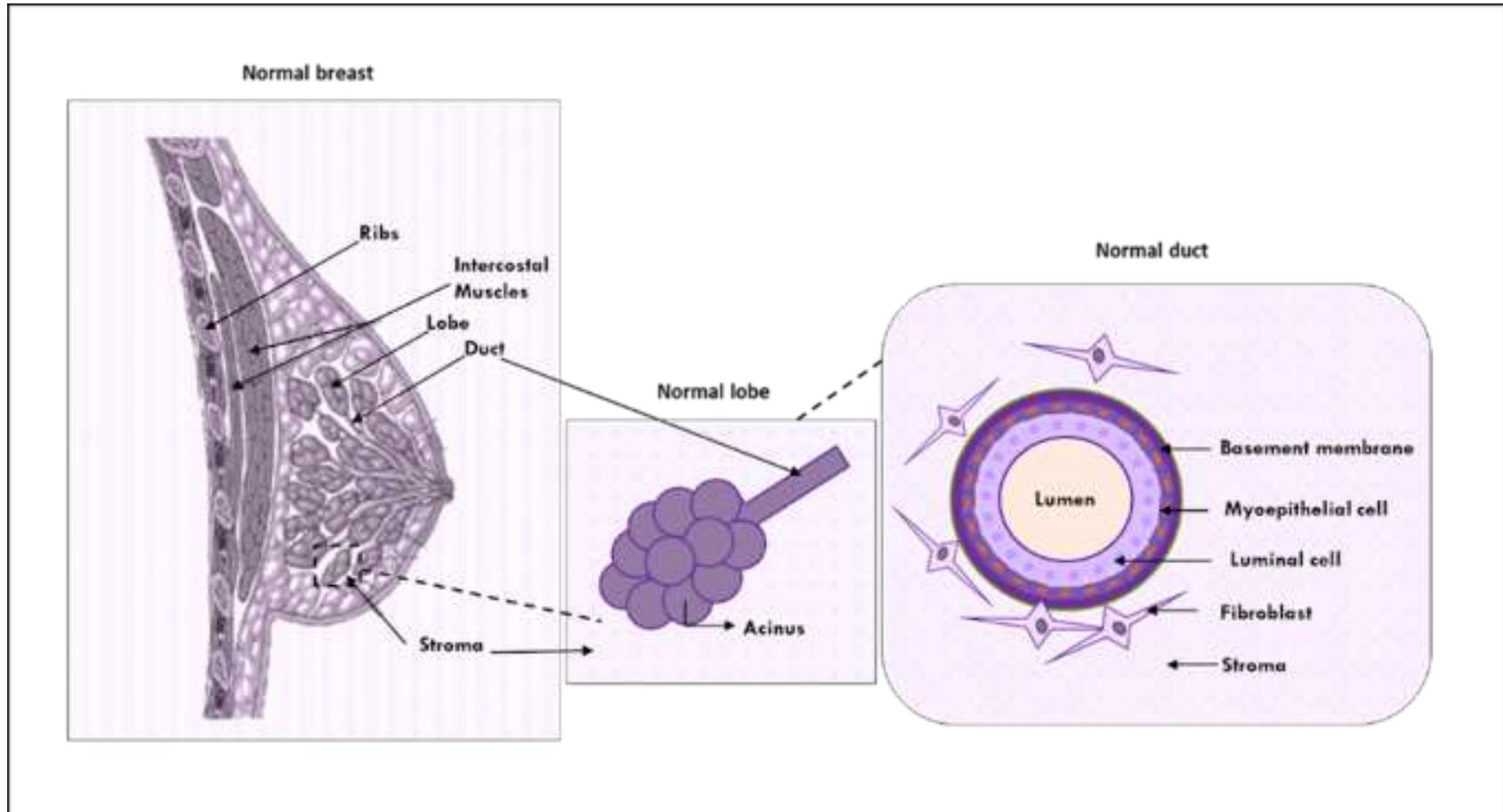


Figure 1.7. Mammary Gland anatomy and histology. Schematic representation of normal female breast *adapted from (Cyr et al., 2016)*. At the histological level, the breast is built of epithelial and stromal cells divided by a basement membrane. The epithelial layer includes two types of cells: myoepithelial cells (basal) and luminal cells, which are enveloped by the basement membrane, whilst, at the functional level, the mammary gland was formed from the glandular and stroma tissues. The glandular tissues are composing of lobes (milk factories) and ducts (milk transporter) that match up the branching system. The stroma forms the area between the lobes.

1.7. Breast cancer classification

The term breast cancer refers to a malignant tumour, which has developed from breast cells (*Cancer Research UK, 2018*). Breast cancer is a complicated heterogeneous disease, which exhibits a variety of phenotypically distinct tumour types that have different histological and pathological features, clinical outcomes, risk factors, response to therapy and prognosis (*Weigelt et al., 2010; Yersal et al., 2014*).

There are four schematic bases for classifying breast cancers and according to different criteria, and which include histological appearance (pathology), grade, stage of cancer, as well as protein and gene status (<https://www.news-medical.net/health/Breast-Cancer-Classification.aspx> ; *Vuong, et al., 2014*).

Under the histopathological characteristics, breast cancer can be divided into ductal carcinoma in situ (DCIS) and lobular carcinoma in situ (LCIS). Both are also divided into pre-invasive and invasive breast cancer Figure 1.8 below (<https://www.news-medical.net/health/Breast-Cancer-Classification.aspx>). The cancer grade is categorised as low grade “well-differentiated” to high grade “poorly differentiated”. In addition, the classification, which is based on the stage of cancer considers tumour size, lymph node involvement and metastasis of cancer, and that is known as "TNM". Similarly, hormone receptor expression status-based classification stratifies breast cancer based on the presence of oestrogen receptor (ER), progesterone receptor (PR) and HER2/neu amplification or overexpression, which facilitates certain novel therapeutic treatments (<https://www.news-medical.net/health/Breast-Cancer-Classification.aspx>).

Recently, more advanced classification has been raised according to the molecular profiling of breast cancer, which was originally proposed by *Sørli et al., 2001*. Gene Expression Microarray studies contributed to a new classification that is supported by new molecular markers. This categorized breast cancer patients into four molecular subtypes, which are luminal A, luminal B, HER2+, and basal types (*Godfrey, et al., 2016*).

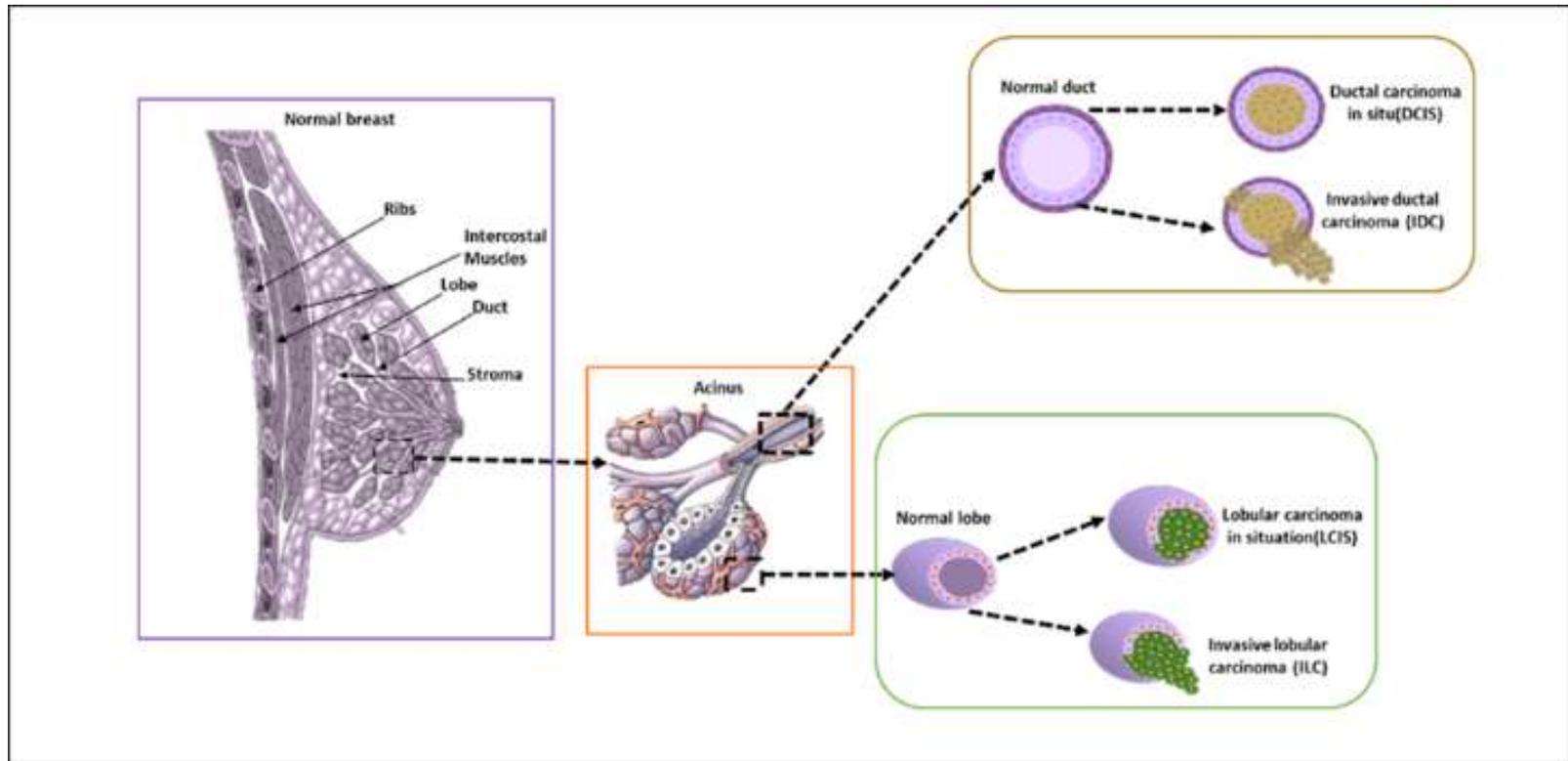


Figure 1.8. Schematic representation of breast carcinomas progression from benign to malignant. Cross section of the breast duct and lobe that compares several types of cancer. The epithelial cells will misplace and fill the ducts and lobes with abnormal cells following several stages of breast carcinoma progression. 1 in 5 of breast cancers new cases is ductal carcinoma in situ (DCIS) while spreading of cancer cells into surrounding breast tissue is known as invasive breast cancer. Invasive carcinoma is divided into Invasive ductal carcinoma (IDC) and Invasive lobular carcinoma (ILC). *This Figure was adapted from <https://www.cancer.org/cancer/breast-cancer/understanding-a-breast-cancer-diagnosis/types-of-breast-cancer/dcis.html>.*

1.8. Breast cancer molecular subtypes

Perou et al., 2000 suggested the existence of four subtypes based on a high throughput gene expression profiling. Following studies, further divided breast cancer into luminal subtypes A and B (*Perou et al. 2000; Sørlie et al., 2001*). Additionally, breast cancer may be divided into six molecular subtypes: luminal A, luminal B, Basal-like, HER2-enriched, Claudin-low, and Normal-like breast cancer (*Eroles et al. 2012, Previati et al. 2013*).

1.8.1. Luminal A (hormone receptor positive/HER2 negative)

Luminal A is the most common breast cancer subtype which includes 50-60% of all breast cancers and starts in the inner (luminal) cells lining the mammary ducts. Generally, this tumour has a low histological grade, a low degree of nuclear pleomorphism, low mitotic activity and good prognosis. The Luminal A subtype is characterised by the positive expression of oestrogen receptor (ER) and/or progesterone receptor (PR) and in higher levels, with low expression of proliferation-related genes such as *Ki-67*. It is also characterised by the expression of luminal epithelial cytokeratin's (CK) 8 and 18, and other luminal markers that are associated with ER function such as *LIV1* (zinc transporter *ZIP6* or *SLC39A6*; solute carrier family 39 zinc transporter, member 6), *hepatocyte nuclear factor 3 alpha (FOXA1)*, *X-box binding protein 1 (XBP1)*, *GATA binding protein 3 (GATA3)*, *B cell lymphoma 2 (BCL2)*, *erbB3* and *erbB4*. In this subtype, patients have a good prognosis and a significantly low rate of relapse when compared to other subtypes. Treatment of Luminal A breast cancer metastases is based on hormonal therapy (*Yersal et al., 2014; Godfrey, et al, 2016*).

1.8.2. Luminal B (hormone receptor positive /HER2 positive or negative)

The Luminal B subtype represents 15%-20% of breast cancers, which is comparable to Luminal A. In comparison to the earlier subtype (Luminal A), it has a more aggressive phenotype and higher histological grade. In addition, it has a worse prognosis and is more proliferative. Moreover, it has a higher recurrence rate and a lower survival rate after relapse. This subtype is also characterised by positive expression of ER and PR receptors and includes HER2 positive and negative subgroups associated with increased expression of growth receptor signalling genes index (*Yersal et al., 2014; Godfrey, et al, 2016*). Moreover, another distinguisher between Luminal A and B subtypes is the proliferation marker *Ki-67* (*Cheang et al., 2009*). Overall, the survival rates in untreated luminal-B breast cancers are similar to the survival rates of Basal-like and HER2-positive subtypes, which are widely recognised as high-risk tumours. Although

Luminal B tumours may have poor outcomes following hormonotherapy, it responds better to neo-adjuvant chemotherapy. The relapse rates are limited to the five first years after diagnosis in the Luminal B breast cancer subtype (Yersal et al., 2014).

1.8.3. HER2-enriched (hormone receptor negative/HER2 positive)

HER2-positive tumours are well defined, due to the ERBB2 gene being amplified, with an exclusive high expression of the HER2 receptor, combined with low or absent expression of ER and PR receptors. The HER2+ subtype stands for 15% to 20% of all breast cancer subtypes. HER2+ classify as more aggressive biologically and clinically. In terms of morphology, it is a highly proliferative tumour and 75% of HER2+ have a high histological and nuclear grade, while 40% have p53 mutations. HER2-positive has a poor prognosis in the absence of treatment however, they have increased sensitivity to certain cytotoxic agents like doxorubicin and are resistant to hormonal therapy. In addition, a targeted therapy drug, Trastuzumab, which is a monoclonal antibody used to treat HER2 positive breast cancer (Yersal et al., 2014; Godfrey, et al, 2016).

1.8.4. Basal-like (hormone receptor negative/HER2 negative)

This breast cancer subtype is characterised by the high expression levels of basal myoepithelial markers, such as CK5, CK 14, CK 17 and laminin and negative expression of ER, PR and HER2, thus referred to as triple-negative breast cancer (TNBC) regarding immune histochemical classification of breast tumours. Whereas the Basal-like subtype resembles the outer (basal) cells surrounding the mammary ducts, which is defined by the gene expression microarray analysis. This type expresses elevated levels of basal myoepithelial markers such as CK5, CK 14, CK 17 and laminin. The Basal-like breast cancer subtype represents 8-37% of all subtypes. The Basal-like subtype correlates with a high histological and nuclear grade and a remarkably high mitotic and proliferative index. This subtype is clinically aggressive and highly metastatic to the brain and lung and has a poor prognosis. It only responds to chemotherapy reagents (Yersal et al., 2014).

1.8.5. Claudin-low (hormone receptor negative/HER2 negative)

Another triple-negative breast cancer subtype is claudin-low, or normal like which is characterised by low expression levels of tight junctions and cell-cell adhesions molecules such as claudins 3, 4 and 7, Occludin and E-cadherin. It has also high expression levels of epithelial to mesenchymal transition genes and stem cell features. Indeed, it has been reported that patients with claudin-low tumours also have poor clinical outcomes (*Yersal et al., 2014*).

1.8.6. Normal breast-like (hormone receptor negative/HER2 negative)

This type of a tumour represents about 5–10% of all breast carcinomas and is poorly characterised. It is being grouped into the inherent subtypes with fibroadenomas and normal breast samples classification. Normal breast-like cancer is presenting an intermediate prognosis between luminal and Basal-like subtypes. It expresses characteristic genes of adipose tissue and does not respond to a neo-adjuvant chemotherapy. Lack of ER, HER2 and PGR expression, and a CK5 and EGFR negative statuses, make those tumours classified as triple-negative cancers. They are rare tumours and there is a doubt that they could be an artefact due to a technical contamination with normal tissue during microarrays analysis (*Eroles et al., 2012*).

1.9. Role of receptors in Breast cancer

Key receptors of breast cancer are HER2, oestrogen receptor and progesterone receptor, which have been used as biomarkers and play a critical role in breast cancer classification and treatment (*Sekar et al. 2011*).

1.9.1. Hormone receptors

Oestrogen receptor and PR are both predictive and prognostic markers for endocrine therapy outcomes and responses. Oestrogen receptor and PR bind hormones exert their effects in the nucleus. Therefore, immunostaining for both receptor proteins is demonstrated in the nucleus of healthy breast tissues, which is usually used as an internal control. In addition, the progesterone receptor serves as an index marker to decide the oestrogen level status. Thus, the expression of both PR and ER is usually reported together to determine the choice and response of therapy among patients. The patients with progesterone receptor expression (PR+) have better survival rates than those who have (PR-) in breast cancer tumours (*Cornejo et al. 2014*).

1.9.1.1. Oestrogen receptors

Two distinct types of oestrogen Receptors (ERs), which are ER α and ER β are found expressed in various tissues including breast, ovary, endometrium tissues that express ER α ; whilst kidneys, the brain, lungs, and several other organs express ER β . The contributory role of ER α protein has been proved in carcinogenesis, while this role is still controversial for ER β (Yager *et al.*, 2006). Both ER receptors, which carry a DNA binding domain are nuclear receptors. However, membrane oestrogen receptors (MERS) are mostly G protein-coupled receptors (GPCR) such as GPER, ER-X, and Gq-mER (Yager *et al.*, 2006). Once the oestrogen enters the cells, those receptors are activated, and a complex of oestrogen hormone and ER receptors will be formed. This complex will translocate into the nucleus, bind to the DNA and start to regulate the activity of different genes to produce transcriptional proteins. This complex has another function that is independent of DNA binding. Patients who express significant quantities of these intracellular receptors have increased the risk associated with proliferation and increased oestrogen-related cellular stimulation. Two major hypotheses were attempted to clarify the role of oestrogen in breast cancer development and progression. Firstly, the generation of radicals (initiator) through the genotoxic effects of oestrogen metabolites and secondly, the capacity of oestrogens hormonal property to induce the proliferation of cancers and premalignant cells (promoter) (Yager *et al.*, 2006).

1.9.1.2. Progesterone receptors (PRs)

The progesterone receptor is a ligand-activated transcription factor, which is a steroidal hormone nuclear receptor (SR) family member. ER and PR steroid hormones receptors are closely related in their biological actions. PR is the main ER target gene, which acts as a leading effector of oestrogen downstream action. More recent studies have implicated progesterone in normal breast proliferation and are a breast cancer long-lasting risk factor. Similarly, to ER, PR has an extensive cross-talk with the same signal transduction pathways, which are needed for the development of mammary gland, as those are most highly expressed in breast cancer (Daniel *et al.*, 2011).

1.9.2. Human Epidermal Growth Factor Receptor 2 (HER2) or HER2/neu

The HER2/neu receptor belongs to the human epidermal growth factor receptor (HER/EGFR/ERBB) family of proto-oncogenes. This protein plays a vital role in the development and progression of malignant breast tumours *via* its amplification and over-expression. Recently, HER2/neu has become an essential biomarker and therapeutic target for about 30% of breast cancer patients. The underlying mechanism of HER2/neu carcinogenesis is still not fully

understood, however, an overexpression or clusters of this protein were associated with rapid tumour growth, shortened survival time increased relapse risk following surgery, and poor response to chemotherapeutic treatments (Ross *et al.*, 2003).

1.10. Molecular mechanisms of breast cancer development and progression

The cellular and molecular mechanisms of breast cancer development and progression were particularly associated with the proliferative response of breast cells toward several peptide growth factors or involvement of the oestrogen receptor (ER) and IGF (Insulin-like growth factor) signalling pathways. It has been shown that insulin-like growth factor (IGF) pathway has a leading role in the development of breast cancer by mediating a chain of events that phenotypically switches normally to neoplastic cells. Experimental, clinical, and epidemiological evidence has revealed IGFs directly impact on the proliferation and survival of the cells (Oh *et al.* 2008; Werner and Bruchim 2012). It also interacts with environmental and genetic factors that are implicated in cancer initiation. IGF is the essential axis of secreted ligands network such as insulin, IGF-1, IGF-2 and IGF-1 receptor (IGF-1R), that regulate metabolism, nutrition, endocrines, growth and ageing events. Besides, the IGF-1R that facilitates the biological actions of IGF-1 and IGF-2, is also considered as a key player in cancer development due to its antiapoptotic effects and transforming activities. Moreover, there is a clear association between increased risk of several types of cancer including breast and prostate, and excessive levels of circulating IGF-1 (Werner and Bruchim 2012). Furthermore, "Breast Cancer genes" (BRCA1 and BRCA2) are well-known 'caretakers of the genome' and play a crucial role in DNA damage identification and DNA repair mediated processes. They have another vital role in the progression of inherited and sporadic breast and ovarian cancer, that is caused by mutations altered by genetic and environmental factors including high IGF-1 concentrations (Werner and Bruchim 2012).

In vitro, breast cancer is often modelled using established cell lines reflecting the different molecular classifications of breast cancer (Holliday and Speirs, 2011). The emergence of high throughput technologies such as gene expression profiling and DNA copy number analysis has allowed a profound awareness of this complex disease (Prat *et al.*, 2015). During the last 15 years, several studies have identified and further validated the classification of breast carcinoma into various intrinsic subtypes such as Luminal A, Luminal B, Her-2 enriched, Basal-like, Claudin-low and normal breast-like. These subtypes were proved to have significant divergences in amplified and mutated genes, survival rates, risk factors, prognosis and response to therapy (Prat *et al.*, 2010; de Macedo Andrade *et al.*, 2014). Table (1.1) below provides examples of breast carcinoma cell lines according to the molecular classification of breast cancer.

Table 1. 1.Molecular classification of breast carcinoma, associated cell lines and therapies

Molecular Subtype	Frequency %	Receptor Expression	CK5/6 & EGFR	Proliferation Genes (Ki67)	Histologic Grade	TP53 Mutation	Prognosis	Relapse rate	Therapy	Example Cell Line	Site of origin
Luminal A	50 – 60	ER ⁺ , PR ^{+/-} , HER2 ⁻ ,	-	Low	Low	Small deletions	Excellent	27.8%	Endocrine therapy	MCF-7, T47D, SUM185	The luminal epithelium of mammary ducts
Luminal B	10 – 20	ER ⁺ , PR ^{+/-} , HER2 ^{+/-} ,	-	High	Intermediate / High	Insertions	Intermediate /Poor	1.6 years	Endocrine therapy, chemotherapy	BT474, ZR-75	The luminal epithelium of mammary ducts
Basal-like	10 – 20	ER ⁻ , PR ⁻ , HER2 ⁻	+	High	High	Complex mutations	Poor	Initial 3 years	Chemotherapy, PARP inhibitors, cytotoxic agents e.g. platinum salts	MDA-MB-468, SUM190	Myoepithelial breast cells
Claudin-low	12 – 14	ER ⁻ , PR ⁻ , HER2 ⁻	+/-	High	High	Complex mutations	Poor	-	Poor response to chemotherapy, U-I	BT549, MDA-MB-231, Hs578T, SUM1315	Epithelial breast cells

HER2-enriched	10 – 15	ER ⁻ , PR ⁻ , HER2 ⁺	+/-	High	High	Complex mutations	Poor	-	Chemotherapy, anti-HER2+ therapy	SKBR3, MDA-MB-453	The epithelium of breast duct
Normal breast-like	5– 10	ER ⁻ , PR ⁻ , HER2 ⁻ ,	+	Low	Low	Small deletions	Intermediate	-	Poor response to chemotherapy, U-I	-	Adipose tissue

The table shows breast cancer subtypes, their receptors expression status, and gives examples of breast cancer cell lines according to their expression of molecular classification. Abbreviations: EGFR= epidermal growth factor receptor, ER=oestrogen receptor, HER2=human epidermal growth factor receptor 2, PR=progesterone receptor. It also displays the therapeutic strategies being used in breast cancer treatments. *Adapted from (Holliday and Speirs, 2011; Eroles et al., 2012; Makhoul, 2018).*

1.11. Breast cancer treatments

Traditionally, the main forms of treatment in breast cancer are Surgery, Radiation therapy, Endocrine therapy, Chemotherapy, and Biological Treatments. Varied factors are considered for breast cancer treatment choice including the type of breast cancer, size of a tumour, stage of cancer, grade of cancer, and protein receptors statuses (*Maughan et al., 2010*). The first type of treatment for breast cancer is surgery, which results in the physical removal of the tumour. Usually, surgery is followed by chemotherapy or radiotherapy. In some cases, it is also followed by hormone therapy or biological treatments. Radiation and Chemotherapy can induce DNA damage, by either inhibiting the repair mechanisms and leading to cell death, or by interrupting the cell cycle. Currently, both regimens are more efficient and less toxic than the old-fashioned methods. Chemotherapy is most effective when combinations of drugs are used for adjuvant and neoadjuvant treatments. the most commonly used drugs in this therapy are Anthracyclines, Taxanes, 5-fluorouracil, Cyclophosphamide, and Carboplatin. Endocrine therapy is an effective treatment for most ER-positive and PR-positive tumours. However, modern therapeutic strategies have been developed to eradicate cancer cells such as Anti-Angiogenic Therapy, and Immunotherapy. Anti-angiogenesis drugs failed to target angiogenesis in the adjuvant/neoadjuvant setting. That leads to excluding this mechanism from the adjuvant setting in the 2000s. In addition, the thriving of immunotherapy in various solid tumours has put forward the potential employment of this treatment in breast cancer. Many strategies have been used to use the capability of the immune system and redirect it to eradicate cancer or to persuade immune inertness. Those strategies include breast cancer vaccines, checkpoint inhibitors, monoclonal antibodies, that enhance the immune-mediated effect of chemotherapy (*Makhoul, 2018*). Immune-mediated mechanisms still need to be further investigated before applying it as a new method of treatment (*Makhoul, 2018; Makhoul et al., 2018*).

Currently, strategies that have been developed to manage breast cancer are focused on pursuing receptor-mediated signalling that includes hormone receptors or HER2 (*Gomez-Martin et al. 2013*). Tyrosine kinase inhibitors (TKIs) such as Lapatinib have been considered as a potential agent in HER2+ breast cancer for targeting EGFR and HER2 receptors (*Gomez-Martin et al. 2013; Vogelstein et al. 2013*). The phosphatidylinositol 3-kinase/AKT/mammalian target of rapamycin (PI3K/AKT/mTOR) pathway plays a critical role in the intracellular signalling system that drives cellular growth and survival. Hyperstimulation of this pathway is involved in the tumorigenesis of ER+ breast cancer and resistance to endocrine therapy. However, PI3K/AKT/mTOR pathway inhibition can augment the benefit of endocrine treatment in ER+ breast cancer. (*Gil, 2014*).

Indeed, Luminal A and luminal B subtypes are amenable to hormone therapy because of their ER expression, which is a therapeutic target. In an equivalent way, the HER2 group is a prospective candidate for trastuzumab therapy. On the other hand, the basal phenotype is characterised by the lack of ER α , PR and HER2 expression, which make these tumours difficult to treat due to the absence of expression of a recognized therapeutic target. It is more aggressive in terms of biology and often has a poor prognosis (Holliday and Speirs, 2011). Treatment regimens are variable between distinct types of cancer and the various stages of the disease. Figure 1.9. below shows the proportion of breast cancer patients under different treatment patterns.

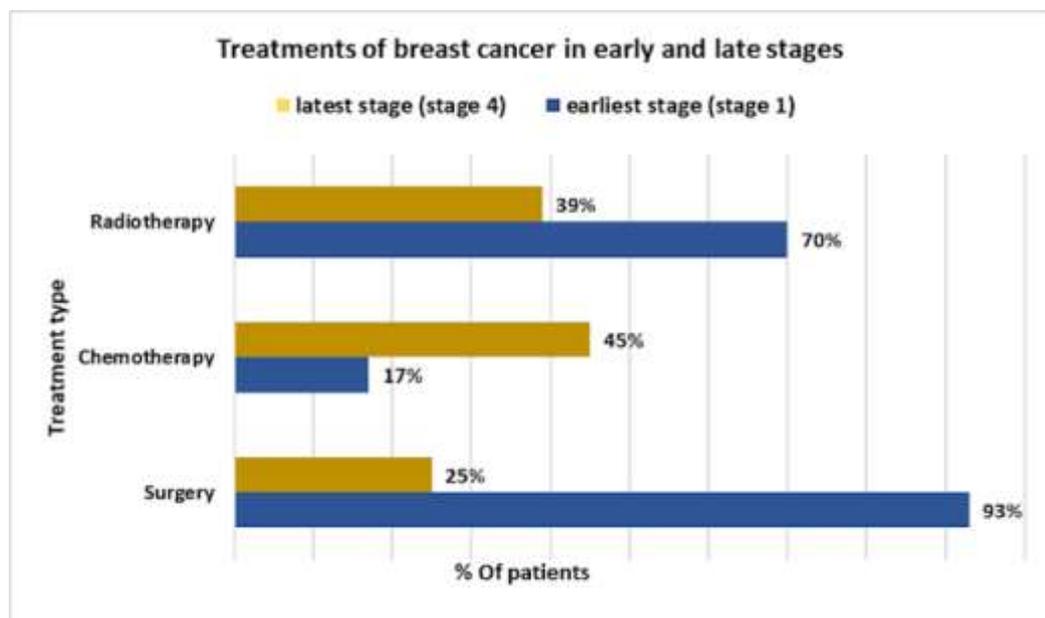


Figure 1.9. Graph representing the percentage of treatments regimens in breast cancer patients for both early and late stages in the UK. The mainstays of primary cancer treatments are Surgery, radiotherapy, and chemotherapy. However, Surgery was the most commonly used treatment in the UK in 2013 and 2014 as the first line of treatment. Whilst, radiotherapy is also varying between the site of a tumour and its stage upon diagnosing (Cancer Research UK, 2018). Early diagnosed patients are more likely to have surgery than chemotherapy, while the pattern for radiotherapy is varied. This data is according to (Cancer Research UK, 2018).

1.12. Metformin hydrochloride

Metformin hydrochloride is a well-known antidiabetic medicine, part of the biguanides family of molecules. Metformin is derived from the plant *Galega officinalis* and has been used in medicine since medieval times. Metformin was originally developed for the treatment of hyperglycemia and Type 2 Diabetes Mellitus (T2DM) (Sośnicki *et al.*, 2016). Metformin (1, 1-dimethyl biguanide hydrochloride) was introduced for the first time in 1958 in the United Kingdom and in 1995 in the United States (Daugan *et al.*, 2016). Currently, it is the first prescribed medication for type 2 diabetes Patients (T2D) (Song *et al.*, 2012) and has been recommended as first-line of therapy for all type 2 diabetes mellitus (T2DM) patients that are newly diagnosed (American Diabetes Association, 2014). Metformin lowers blood sugar levels in people with T2DM through increasing the sensitivity of muscle cells to insulin to be used effectively and to reduce the amount of sugar (glucose) production in the liver (<https://www.diabetes.org.uk/Type-2-diabetes>). Due to Metformin's favourable benefit-risk profile, it has been extensively used with about 120 million prescriptions around the world every year (Daugan *et al.*, 2016). The broad spectrum of pleiotropic effects and good tolerability by patients is the most characterised features for this drug. In addition, Metformin endothelium functions and serum lipid profile has been upgraded. Consequently, this reduced the both-sided intricacy (micro- and macrovascular) of T2DM (Sośnicki *et al.*, 2016). It is also approved for treatment of polycystic ovary syndrome (PCOS) and obesity due to its insulin resistance capability and counteraction of hormonal imbalances. Moreover, Metformin is being used as a viral and cancer inhibitor (Provinciali *et al.*, 2015). Recently, further potential indications about Metformin, have been added as results of clinical and preclinical studies. Metformin has been used as a cardiovascular protective agent, anti-inflammatory, neuroprotective, and anticancer agent. Besides, Metformin is not involved in hypoglycaemia induction or weight gain. Current studies have referred that the pleiotropic effects of Metformin are due to its mitochondrial action (Daugan *et al.*, 2016). The positive charge of Metformin allows this drug to interact with the mitochondrial membrane potential and accumulates in mitochondria to inhibit the mitochondrial electron transport chain complex I. This results in a decrease in cellular ATP concentration, allowing rise of ADP/ATP and AMP/ATP ratios which leads to low levels of energy production. The 5' adenosine monophosphate-activated protein kinase (AMPK) is the major energetic sensor in the cells. Cells can restore the energetic homeostasis by activating this protein kinase. This activation leads to an increase in the catabolic reactions and decreases in the anabolic reactions in order to balance the cellular energy and evade bioenergetic crisis and cell death (Daugan *et al.*, 2016).

1.13. Metformin and cancer treatment

Intensive evidence from *in vitro* and *in vivo* studies demonstrated the anticancer effects of Metformin. These are based on a dual mode of action: indirect effect *via* reducing glucose in the blood and insulin levels, and direct effect through AMPK activation (*Daugan et al., 2016*). In 2005, and for the first time *Evans* and colleagues suggested in their study that Metformin may reduce the risk of cancer in patients with type 2 diabetes (*Evans et al., 2005*). Additional epidemiological, clinical and laboratory studies further demonstrated the pleiotropic effects of Metformin on various types of cancers including breast (*Liu et al., 2009; Bodmer et al., 2010; Col et al., 2012; Marinello et al., 2016; Pandiri et al., 2016, Tang et al., 2018*), lung (*Libby et al., 2009; Lai et al., 2012; Noto et al., 2012*), colon (*Currie et al., 2009; Tseng, 2012*), pancreas (*Currie et al., 2009; Lee et al., 2011; Hsieh et al., 2012; Soranna et al., 2012*) and liver cancer (*Lee et al., 2011; Hsieh et al., 2012; Noto et al., 2012*). Furthermore, Metformin suppressed the growth of human head and neck squamous cell carcinoma through comprehensive inhibition of protein translation (*Sikka et al., 2012*). Metformin treatment in diabetic patients resulted in a significant decrease in cancer risk and decreased cancer-related and overall mortality (*Libby et al., 2009*). These observations of Metformin effects were mostly noticed in patients treated with Metformin for a prolonged period or with considerable dosage (*Libby et al., 2009; Bodmer et al., 2010; Col et al., 2012; Tseng, 2012; Lin et al., 2014*). In addition, this drug decreased the risk of cancer and cancer mortality in type 2 diabetes patients (*Rizos and Elisaf, 2013*). Patients that have been treated with Metformin for more than four years, the cancer incidence has been decreased by about 50% (*Evans et al., 2005*). Another meta-analysis in T2DM patients who were receiving Metformin has confirmed that cancer incidence was significantly reduced by 30–50% in pancreatic cancer, hepatocellular carcinoma and colon cancer (*Decensi et al., 2010*). Moreover, in a prospectively followed cohort study, it has been found that Metformin use was associated with a lower cancer mortality in comparison with non-user of Metformin in a dose-dependent manner. (*Landman, et al., 2010*). Moreover, a peer-reviewed article by *Ikhlas and Ahmad* has been published in 2017 to survey the underlying mechanisms of Metformin anticancer activity alongside with preclinical and clinical studies outcomes, and that highlighted the involvement of AMPK dependent and AMPK independent pathways (*Ikhlas and Ahmad, 2017*). Table 1.2. below is showing data that support the potential antitumor effect of Metformin in diverse types of cancer.

Table 1.2. Studies of anticancer potential of Metformin

Type of cancer	Metformin in combination with	Model	Results that led to anticancer effects	References
Pancreatic cancer	Rapamycin	Mice	Reduction in the growth of pancreatic cancer	<i>Cifarelli et al., 2015</i>
	Gemcitabine	MIA, SU86.86 and AsPC1 cell lines	Downregulation of ABCB1 gene	<i>Lyn-Cook et al., 2015</i>
	TRAIL	MIA PaCa-2, PANC-1 cell lines	Expression of DR5 and Bim	<i>Tanaka et al., 2015</i>
	Boswellic acid NP	MiaPaCa-2 cell line	DNA fragmentation	<i>Snima et al., 2015</i>
Breast cancer	S31-301	MDA-MB-468, HCC70, MDA231, BT20 cell lines	Inhibition of Stat3 at Tyr-705 and Ser-727	<i>Deng et al., 2012</i>
	Doxorubicin and paclitaxel	MDAMB-231, HCC-70, HCC-1937, MCF-7 cell lines	ERK 1/2 and AKT phosphorylation inhibition	<i>Guimaraes et al., 2015</i>

	5-Fluorouracil, epirubicin, and cyclophosphamide	MDA-MB-468 & 231, HCC1937, MCF7, SKBR3, T47D cell lines and CSC	Inhibition of ATP production and DNA repair	<i>Soo et al., 2015</i>
	Aspirin, atenolol	Mice	Mitochondrial complex 1 inhibition	<i>Talarico et al., 2015</i>
	Fulvestrant	MCF7 cell line	Cyclin G2 leads to the cell cycle arrest	<i>Horn et al., 2015</i>
	Rapamycin	MCF7 cell line	Reduction in aromatase activity and reduced androgens to E2 conversion	<i>Rice et al., 2015,</i> <i>Tang et al., 2018</i>
	Everolimus	SCID mice	Obstruction of S6 and 4EBP1 phosphorylation	<i>Wang et al., 2015</i>
Colorectal cancer	Vitamin D3	Rat and mouse	The decrease in S6P expression as well as c-Myc and Cyclin D1 protein expression	<i>Wang et al 2015</i>
	5-Fluorouracil and oxaliplatin	HT-29 and HCT-116 and SCID mice xenograft	Wnt/ β -catenin signalling pathway downregulation along with significant increase in miRNA 145 and reduction in miRNA 21	<i>Nangia-Makker et al., 2014</i>

	DMFO	BALB/c nu mice and HCT 116, HT 29, MEL1861, SK-23 cell lines	Upregulation of Beclin-1 and downregulation of p70S6 and 4EBP1	<i>Zhang et al., 2014</i>
	DCA	CT26-CL25, CT26-WT cancer cell lines, and BALB/c mice	Reduction in tumour growth and an increase in apoptotic cells	<i>Sticca and Murphy, 2014</i>
Prostate cancer	P1K1 inhibitor (BI2536)	LNCaP, C4-2, DU145, PC3, 293A and RWPE-1 cell lines	p53-Dependent apoptosis and inhibition of glutamine anaplerosis	<i>Shao et al., 2015</i>
	CDK inhibitor (AZD5438)	Du145, PC3, PPC1 cell lines	The remarkable increase in β -galactosidase, p27, and p16 expression	<i>Blute et al., 2015</i>
	Pyruvate kinase M2 activator (DASA-58)	PC3, DU145 cell lines	Metabolic inactivation of PKM2 via both oxidation and Src-mediated phosphorylation	<i>Giannoni et al., 2015</i>

			leading to PKM2 nuclear translocation	
	Rapamycin	Hi-Myc mice	Reduced prostate tissue inflammation and inhibition of mTORC1 signalling	<i>Saha et al., 2015</i>
	Benzylserine	LNCaP and PC-3 cell lines	Reduced basal oxygen consumption and fatty acid synthesis along with the reduced E2F expression	<i>Wang et al., 2015</i>

This table shows the anticancer potential of Metformin and when in combination with other drugs that were reported by various studies. *This Table was adapted from (Ikhlas and Ahmad, 2017)*. The abbreviations are: ABCB1 = ATP-binding cassette B1, TRAIL = tumour necrosis factor related apoptosis-inducing ligand, NP = nanoparticle, PKM2 = pyruvate kinases M1/M2, DR5 = death-receptor 5.

Furthermore, according to <https://clinicaltrials.gov> data from 2018 about 318 ongoing clinical trials were registered, which aimed at evaluating Metformin role in cancer treatment and cancer prevention procedures. Figure 1.10. provides a list of the number of principal ongoing clinical trials according to cancer types.

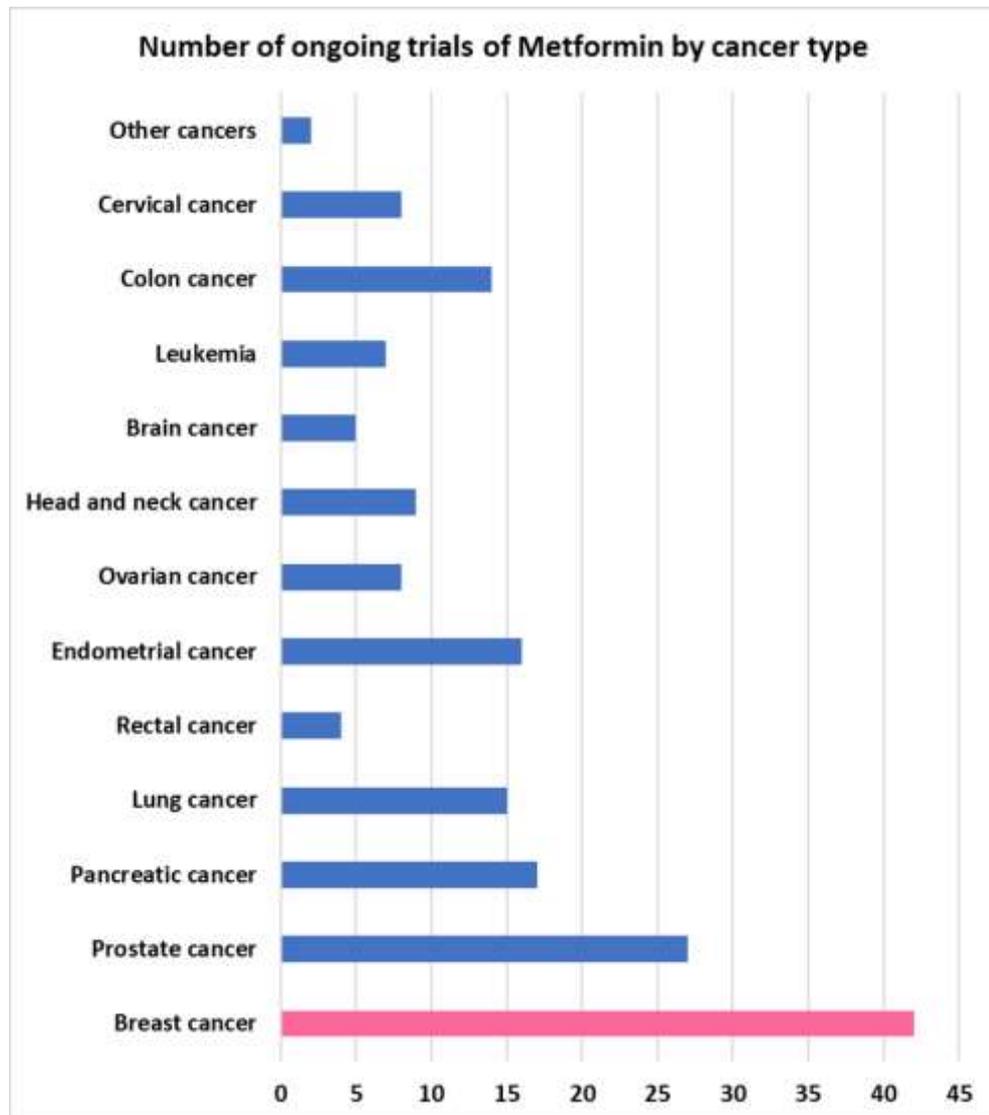


Figure 1.10. The number of main clinical trials assessing Metformin anticancer and prevention capacities against several types of cancer. This Table is displaying the number of ongoing clinical trials in several types of cancer. Breast cancer is top of this list (42 studies in breast cancer alone were registered). The source of this data is <https://clinicaltrials.gov>, 2018.

1.14. Metformin and breast cancer

The cytotoxic effect of Metformin on breast cancer has been shown using MCF-7 human breast cancer cells and FSaII mouse fibrosarcoma cells at a low concentration of the drug. Importantly, Metformin treatment caused a significant clonogenic death in FSaII mouse fibrosarcoma cells and MCF-7 human breast cancer cells. Besides, Metformin was selectively cytotoxic to cancer stem cells relative to non-cancer stem cells in a dose and time-dependent manner (*Song et al., 2012*). Another study has determined the efficiency of Metformin on the inhibition of proliferation and induction of apoptosis in different breast cancer cell lines (*Liu et al., 2012*). Trilla-Fuertes and colleagues 2018 study revealed a heterogeneous effect of Metformin treatment on cell proliferation that is consistent with the cell cycle disruption in the G2 / M phase in breast cancer cell lines. In addition, they suggested an increase in the response enzymes of reactive oxygen species (ROS) due to Metformin therapy. In MCF7 cells, nitric oxide synthase was predicted to increase, and MDA-MB-468 cells showed susceptibility to Metformin treatment (*Trilla-Fuertes et al., 2018*). A clinical study integrating the measurement of markers of systemic metabolism, dynamic FDG-PET-CT, transcriptomics, and metabolomics at paired time points to profile the bioactivity of Metformin in primary breast cancer, showed that Metformin treatment reduced the levels of mitochondrial metabolites, activated multiple mitochondrial metabolic pathways, and increased the 18-FDG flux in tumours. The study concluded that mitochondrial response to Metformin in primary breast cancer may define anti-tumour effect (*Lord, et al., 2018*).

Moreover, Metformin has increased radiosensitivity of human and mouse cancer cells (*Song et al., 2012*). This is important as radiotherapy is widely used as a co-therapy with drug interventions in breast cancer treatment. Furthermore, Metformin can intensify the inhibition of cell proliferation when co-administrated with chemotherapy (*Liu et al., 2012*). Metformin treatment for 24h before and after radiation sensitises MCF-7 cancer cell lines to irradiation (*Song et al., 2012*). Another study has shown that the combination of Metformin and chemotherapeutic agents and/or the mTOR inhibitor RAD001 might be a favourable approach for breast cancer treatment (*Liu et al., 2012*). Other studies demonstrated beneficial effects of Metformin as a coordinator, with other drugs or therapeutic methods, for cancer treatments. *Hirsch et al., 2013* suggested that Metformin-based combinatorial therapy could be more efficient in xenograft cell lines (*Hirsch et al., 2013*). Metformin also suppresses the cellular transformation and selectively kills cancer stem cells in the breast cancer cell line MCF10A-ER-*Src* (*Hirsch et al., 2013*). A finding of *Dallaglio et al., 2014* elucidated the contradictory effects of Metformin on endothelial and tumour cells, as well as on angiogenesis in breast cancer. They have shown that Metformin inhibits the ability of the endothelial cell to organize into capillary-

like networks through the energy sensor AMPK. Metformin anti-angiogenic activity was exerted through the inhibition of ERK1/2 activation, even in the presence of VEGF *via* AMPK activation (Dallaglio *et al.*, 2014). In addition, Talarico and Colleagues in 2015, found in their study that Aspirin and Atenolol enhanced Metformin targeting action against both neoplastic and the microenvironment of breast cancer cells (Talarico *et al.*, 2015). Moreover, Marinello and her team in 2016, highlighted the probable clinical utility of Metformin during treatment of luminal and triple-negative breast cancer as an adjuvant (Marinello *et al.*, 2016). A meta-analysis published in 2015 by Yang and his colleagues revealed that Metformin use does not reduce the incidence of breast cancer, however, Metformin therapy has significantly reduced the mortality rates in patients with luminal-type breast cancer (Yang *et al.*, 2015). Furthermore, *ClinicalTrials.gov* adapted data has found that 42 clinical trials involving breast cancer employ Metformin treatment alone or in combination with other drugs. However, only 13 studies have been completed, 15 are recruiting patients, 3 ended, 5 actives, 1 withdrawn and 5 were unknown studies. Table 1.3 below is showing the complete 13 studies in details (<https://clinicaltrials.gov>, 2018).

Table 1.3. Inventory of the completed clinical trials encompassed breast cancer and Metformin

Ranked	Title	Conditions	Metformin in combination with
1	Clinical and Biologic Effects of Metformin in Early Stage Breast Cancer	•Breast Cancer	
2	Effect of Metformin on Breast Cancer Metabolism	•Breast Cancer	
3	Metformin in Breast Cancer, Visualized with Positron Emission Tomography	•Breast Cancer	Radiation
4	A Trial of Standard Chemotherapy with Metformin (vs Placebo) in Women with Metastatic Breast Cancer	•Metastatic Breast Cancer	Placebo
5	Study of Erlotinib and Metformin in Triple Negative Breast Cancer	•Breast Cancer	Erlotinib
6	Use of Metformin to Reduce Cardiac Toxicity in Breast Cancer	•Breast Cancer	Doxorubicin
7	Metformin Hydrochloride vs. Placebo in Overweight or Obese Patients at Elevated Risk for Breast Cancer	•Breast Cancer	placebo

8	Neoadjuvant Letrozole Plus Metformin vs Letrozole Plus Placebo for ER-positive Postmenopausal Breast Cancer	•Hormone Receptor Positive Malignant Neoplasm of Breast	Placebo
9	Exercise and Metformin in Colorectal and Breast Cancer Survivors	•Colorectal Cancer •Breast Cancer	Exercise training
10	Efficacy and Safety of Adjuvant Metformin for Operable Breast Cancer Patients	•Breast Cancer	Placebo
11	Myocet + Cyclophosphamide + Metformin Vs Myocet + Cyclophosphamide in 1st Line Treatment of HER2 Neg. Metastatic Breast Cancer Patients	•Human Epidermal Growth Factor 2 Negative Carcinoma of Breast	Myocet + Cyclophosphamide
12	The Impact of Obesity and Obesity Treatments on Breast cancer	•Breast Cancer	Exemestane
13	Metformin and Temezirolimus in Treating Patients with Metastatic or Unresectable Solid Tumor or Lymphoma	•Breast Cancer	Temezirolimus

In this table 13 clinical trials have been completed, however, their data is not available except study number 9 which has obtained data. The condition of cancer and drugs combination were also presented. The data was collected from <https://clinicaltrials.gov> in 12/09/2018.

1.15. Antitumor Effect of Metformin

Considering several epidemiological studies that showed Metformin as a proposed antineoplastic drug, a question was raised concerning Metformin mode of action in cancer treatment. As an answer to this query, different explanations have been advanced to clarify the considerable assumption of Metformin anticancer activity. This activity is based on reduced circulating insulin and insulin growth factor I (IGF-1) levels and activated cellular pathways that counteract various mitogenic stimuli (*Sośnicki et al., 2016*). Indeed, the antitumor effects of Metformin are correlated with direct and indirect mode of action of this drug, which is also known as insulin-dependent and–independent mechanisms that are directly acting on the cancer cell growth and proliferation. Both the systemic (indirect) and intracellular (direct) mode of action of Metformin depend on its activation of the 5' adenosine monophosphate-activated protein kinase (5' AMP-activated protein kinase or AMPK). Metformin indirect mode of action can be abridged by an insulin level reduction. The two potential growth factors are insulin and IGF-1, which receptors are expressed on many cancer cells including breast cancer, and that could stimulate cell survival and mitogenesis. As a result, using Metformin for treatment could lower insulin levels and IGF-1 in the serum, which leads to the reduction of cell growth (*Viollet et al., 2012; Morales and Morris, 2015*). Different mechanisms could explain Metformin direct mode of action at the cellular level. Firstly, the stimulation of LKB/AMPK signalling might block aerobic glycolysis in cancerous cells with efficient LKB/AMPK pathways, that enhances cell death. It also could induce malignant cell death even if those cells did not contain functional LKB/AMPK pathways, by reducing ATP levels and suppressing the response to energy stress. Metformin directly works on mitochondria to alter cellular bioenergy and restrict breathing and the cell sensitivity to Metformin depends on its ability to overcome active stress (*Andrzejewski et al., 2014*). While, the second potential effect of Metformin is on chronic inflammation, which may be a substantial factor in carcinogenesis and progression. AMPK stimulation seems to prohibit the synthesis of pro-inflammatory cytokines in different types of cells such as macrophages and adipocytes. The activation of 5' adenosine monophosphate-activated protein kinase (AMPK), which is linked with the phosphatidylinositol 3-kinase (PI3K)/phosphatase and tensin homolog (PTEN)/protein kinase B (AKT) pathway and mitogen-activated protein kinase (MAPK)/extracellular signal-regulated kinases (ERK)cascades, all known for being often dysregulated in breast cancer. Therefore, simultaneously targeting AMPK through Metformin and the PI3K/AKT/mTOR pathway by a mTOR inhibitor could become a therapeutic approach" (*Liu et al., 2012*). Metformin efficacy could also mediate the downstream regulation of fatty acid synthesis AMPK-dependent. Fatty acid synthase (FAS) is a crucial enzyme in fatty acid biosynthesis, which is related to a higher risk of malignant conversion (*Pernicova and Korbonits,*

2014). Though obesity is usually characterised by the development of a chronic proinflammatory condition, that results in increased immune cytokines filtration. These cytokines like leptin, adiponectin, interleukin 1 beta (IL-1 β), IL-6, plasminogen activator inhibitor-1(PAI-1) and tumour necrosis factor alpha (TNF α), are involved in proliferation and progression of cancers. Metformin might affect chronic inflammation *via* AMPK activation, which has an inhibitory effect on the pro-inflammatory cytokine's synthesis in diverse types of cells such as macrophages and adipocytes (Morales and Morris, 2015). Accordingly, it has been proposed that Metformin could potentially target the pro inflammatory cytokines within the tumour microenvironment and inhibits the growth in the sensitive cancers. Finally, Metformin might have a significant role in limiting the tumour growth and metastasis through the inhibition of endothelial cell migration and neoplastic angiogenesis by decreasing the levels of VEGF (vascular endothelial growth factor) and PAI-1 levels of VEGF and PAI-1 (Morales and Morris, 2015). The double effect of Metformin is based on the reduction of insulin levels and its cell molecular activity on AMPK and mTOR (Ben Sahra et al., 2010). This positive action may be the key factor in cancer treatment. The schematic representation of Metformin both direct and indirect mechanism of inhibition in cancer cells is shown in Figure 1.10.

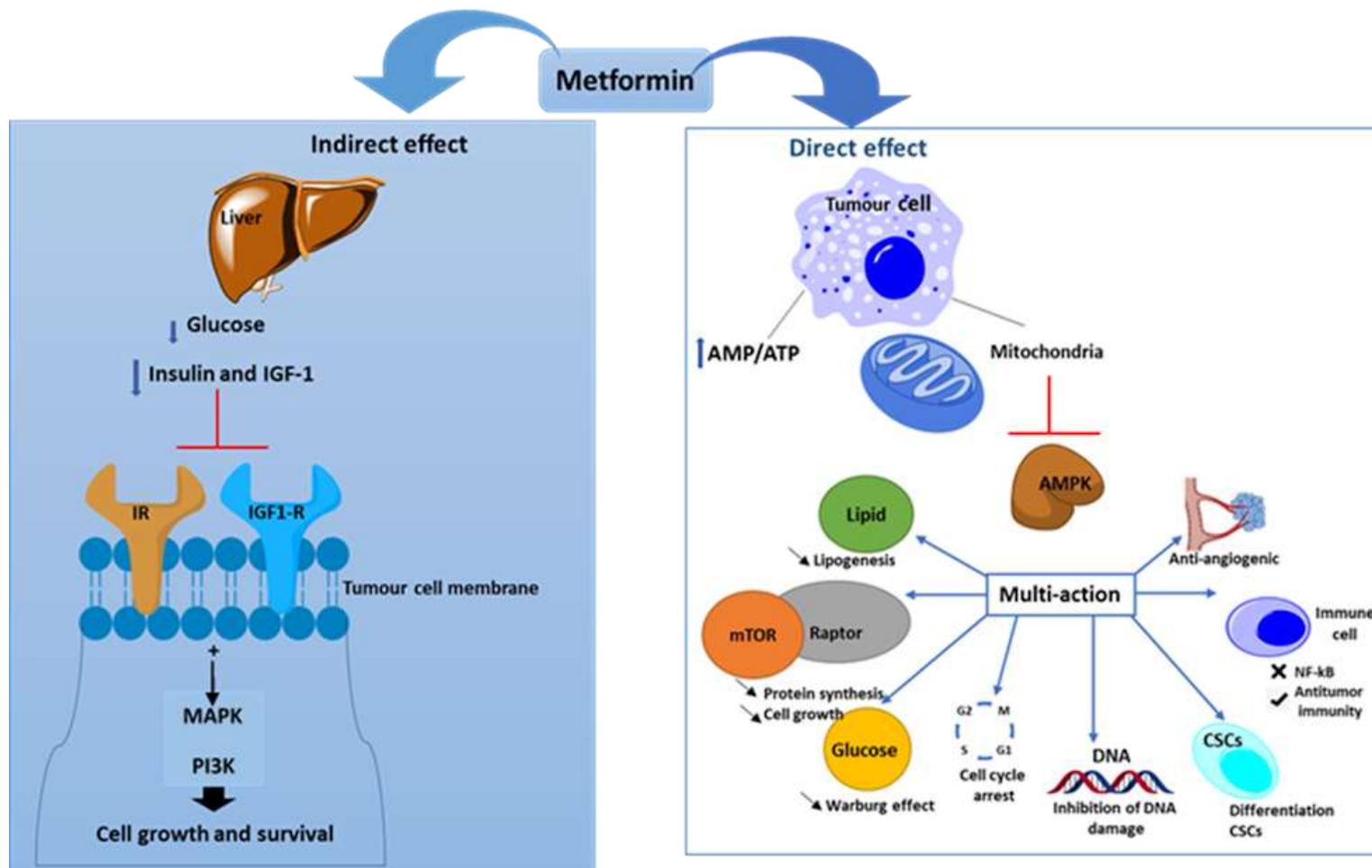


Figure 1.11. Metformin Mode of action in tumour cells. Adapted from (Daugan et al., 2016). The schematic diagram shows the unique mode of action of Metformin, which might give a potential double-strike against an anabolism-addicted tumour and its critical suppliers (insulin/glucose). The indirect action of Metformin is summarised by lowering levels of blood insulin and glucose persistently. While the indirect action of Metformin is to promptly suppresses the pivotal AMPK/mTOR/S6K1 axis and several protein kinases. Abbreviations: IR =insulin receptor, IGF-1R = insulin-like growth factor receptor, PI3K=phosphoinositide 3-kinase, MAPK=mitogen-activated protein-kinase, mTOR= mammalian target of rapamycin, CSCs=cancer stem cells, AMPK=activated protein kinase.

1.16. Cancer-related targets of Metformin

1.16.1. Insulin-like growth factor-1 (IGF-1)

Insulin-like growth factor (IGF-1) is a single polypeptides chain that comprises of 70 amino acids, with a reciprocal 70% homology in humans and shares an approximate 50% homology with the insulin, which is in line with associated insulin-like features of the IGFs. The systemic growth factors (IGF-1 and IGF-2) can be synthesised by most of the cells throughout the body. However, their levels are mainly determined by production in the liver. In addition, the main inducer of IGF synthesis is the growth hormone (GH), which regulates the concentration of circulating IGF-1 (*Heemskerk et al., 1999*). Insulin-like growth factor (IGF) has a fundamental role in the regulation of normal human tissues and growth *via* prompts the proliferation and differentiation of cells and prohibits cell apoptosis. Whilst, it has the capability to influence several underlying phenotypes of cancer. The regulation of cell processes can be achieved by the interaction of IGF-1 and its specific receptors on the cell surface, the IGF-1 receptor (IGF-1R). An increased level of both IGF-1R and/or circulating IGF ligands has been seen in several cancers such as Ewing sarcoma, breast, prostate, pancreatic, melanoma and many other types of cancer. This overexpression can be caused by loss of tumour suppressors such as p53, breast cancer gene-1 (BRCA1), von Hippel-Lindau protein and Wilms's tumour suppressor WT1. The dysregulation of the IGF axis could contribute to many hallmarks of cancer. (*Simpson et al., 2017*).

1.16.2. IGF-1 receptor (IGF-1R)

The IGF-1R is a heterotetrameric receptor, which has two extracellular binding domains ligands including alpha subunit, and two other transmembrane beta subunits. The beta subunits are holding the kinase domain, the alpha and beta domains are linked through disulfide bonds. Both IGF-1R and the insulin receptor (INSR) are members of the receptor tyrosine kinase (RTK) family class 2 (*Simpson et al., 2017*). The IGF1R complex is involved in mitogenic abilities, potent antiapoptotic and oncogenic transformation. In addition, it has an essential role in angiogenesis, invasion, and metastasis (*Werner and Bruchim, 2012*).

1.16.3. 5' Adenosine monophosphate protein kinase (AMPK)

5' AMPK is a highly conserved serine/threonine protein kinase that is composed of a catalytic subunit (α) and two regulatory subunits (β and γ), which is expressed in several tissue types such as the liver and skeletal muscles. This kinase is a leading sensor in sustaining the homeostasis of

cells energy. Activation of AMPK can boost ATP production and adjust metabolic energy, due to AMPKs essential role in several metabolic processes such as glucose uptake, oxidation of fatty acid in the muscles, synthesis of fatty acids, gluconeogenesis of liver and the regulation of food intake at the hypothalamus level. The dysregulation of this kinase has been related to several pathological conditions such as diabetes, neurodegeneration, cardiac hypertrophy and cancers (Vallianou *et al.*, 2013). One of the acquired hallmarks of cancer is reprogramming of energy metabolism. Thus, the activation of AMPK might have a metabolic suppressor activity on the tumour by regulating the levels of energy in the cells and implementing metabolic checkpoints to inhibit the growth of cells (Li *et al.*, 2015). Additionally, AMPK could modulate inflammation and directly influence tumorigenesis. The consequent events of AMPK activation could be processed through targeting several events and pathways such as mammalian target of rapamycin (mTOR), cell cycle arrest, inflammation, glucose metabolism, angiogenesis and cancer stem cells (CSCs). AMPK can be activated in response to different kinases including liver kinase B1 (LKB1), calcium/calmodulin-dependent protein kinase (CaMKK) and transforming growth factor β (TGF- β)-activated kinase (TAK1). In addition, it acts on extracellular changes including depletion of ATP (shortage of energy), low glucose levels, and alteration in Nicotinamide Adenine Dinucleotide Phosphate Hydrogen (NADPH) levels. Besides, many drugs can also activate AMPK such as Metformin, some non-steroidal anti-inflammatory drugs (NSAIDs), traditional Chinese herbs and certain natural products. As a consequent to AMPK activation, several effectors proteins that are involved in various regulatory processes will be affected and will contribute to the pathogenesis of cancer. AMPK is a well-known target in the treatment of type-2 diabetes and metabolic syndrome. Moreover, AMPK is emerging as a potential metabolic tumour suppressor and as a possible target for cancer treatment and prevention. The cancer-related targets of AMPK include a Mammalian target of Rapamycin (mTOR), Cyclooxygenase-2 (COX-2) proinflammatory enzyme, the p53 tumour suppressor, Acetyl-CoA carboxylase (ACC) and the Akt Signalling pathway (Li *et al.*, 2015; Daugan *et al.*, 2016; Ikhlas and Ahmad, 2017).

1.16.4. Mammalian target of Rapamycin (mTOR)

mTOR is a serine/threonine protein kinase, which is formed of five significant components including mammalian target of rapamycin (mTOR), regulatory associated protein of mTOR (RAPTOR), mammalian lethal with SEC13 protein 8 (mLST8) or G β L, proline-rich AKT substrate 40 kDa (PRAS40) and domain-containing mTOR-interacting protein (Debtor). Functionally, mTOR composes two distinct complexes, which are mTORC1 (mTOR complex 1) and mTORC2 (mTOR complex 2). mTOR signalling is an important pathway that regulates different cellular processes

such as growth, proliferation, motility, survival, protein synthesis and transcription. Indeed, the inhibition of mTORC1 leads to the suppression in cell growth and proliferation, which directly or indirectly regulates autophagy and apoptosis *via* ribosomal biogenesis, mRNA translation, lipid synthesis, and other metabolic events. AMPK is the dependent and independent mTORC1 inhibitor (Laplante and Sabatini, 2009; Li et al., 2015; Ikhlas and Ahmad, 2017).

1.16.5. Phosphoinositide 3-kinase (PI3K)

Phosphoinositide 3-kinase (PI3K) is one of the significant intracellular transduction cascades that is implicated in cell growth and cell survival. Besides, the PI3K signalling pathway regulates several steps in glucose metabolism, and cytoskeletal functions like cell motility and adhesion. The direct coordination between glycolysis and cytoskeletal dynamics by PI3K has been shown to be AKT-independent both *in vitro* and *in vivo* (Hu et al., 2016). Accordingly, PI3K influences Glucose avidity and cytoskeletal plasticity hallmarks of cancer. This pathway is involved in fatal diseases like diabetes and cancer. Insulin and growth factors (GFs) are the main stimulators of PI3K. Activation of PI3K resulting in phosphatidylinositol-3,4,5-trisphosphate (PIP3) production, and by which are initiated much signalling, that controls the metabolism of glucose, growth of cells and cell movement (Daugan et al., 2016; Hu et al., 2016). Moreover, the PI3K/AKT signalling pathway plays an essential role in several types of tumour progression as it is involved in cell growth, proliferation, and apoptosis. PI3K/AKT controls the downstream targets of many pathways including the indirect inhibition of mTOR causing cell cycle arrest and apoptosis. The AKT activates the Mouse double minute 2 homolog (MDM2) indirectly which leads to a decrease in the p53 level and activity that resulted in promoting of p53 translation and protein stability. (Zhang et al., 2018).

1.17. Systems biology

The complicated, adaptive, dynamic, and non-linear nature of human biology presents a challenge in drug discovery and development of novel, safe, and effective medications. The definitive aim of systems biology is to assimilate the physiology and the disease through multiple hierarchical levels of an organism, starting from chemical and molecular interactions linking to pathways and pathway networks. This understanding begins at the level of the cell to cell and tissue, going up to organs, organ systems, and eventually to the elaborate of the entire organisation. Besides, Systems biology research includes the integration of high-throughput datasets of system elements (omics data), experimental methods of analysis, and application of network approaches and computationally derived models (Hood and Perlmutter, 2004; Butcher et al., 2004; Berg, 2014). Therefore, Systems Biology can be defined as a more comprehensive

alternative approach to life science, in comparison with reductionism, which has dominated and influenced the biomedicine and clinical medicine through carrying out the diagnosis, treatment and prevention of diseases. Despite the successfulness of reductionism, it is also having many limitations (Ahn *et al.*, 2006). Also, earnest efforts have been directed through systems biology approaches towards the identification of drug targets or the development of novel therapeutics and new indications for existing drugs in pharmaceutical research. Current studies are inclined towards the small molecule's identification and characterisation, which are selectively inhibiting (or activating) specific pathway mechanisms or molecular targets. Thus, special attention has been focused toward the studies that are involving drug mechanisms of action and supporting drug development goals, for example, patient stratification and clinical care choice. Moreover, global information at the levels of cell and tissue in specific time points, checking dynamic changes can be provided by using omics tools in samples from cell-based assays, preclinical animal models or human studies. Furthermore, omics data is a set of data that are obtained from mRNA transcriptomics, proteomics, and metabolomics, which were integrated with genomics information and other data types. This integration has employed to structure a model of cell signalling pathway and as disease networks either to find novel targets and/or to help better understand and expect a drug action *in vivo*. Besides, there are resources of literature knowledge and accumulated information in addition to experimentally derived data, which can be combined and converted to a formal form. Indeed, this can be achieved by processing a natural language processing (NLP) – based method or defined by an ontology expert into sequences of semantic statements. Systems biology studies in biomedical research termed as 'network medicine' or network pharmacology, in which researchers challenged the integration of data sets and started to change the way of medicine preparation (Barabási, *et al.*, 2011; Silverman and Loscalzo, 2012; Berg, 2014).

1.18. Systems biology and biological networks

Systems biology approaches needed an integral grouping of both experimental and computational tools, to understand the biological processes, which are regulating the living system (Heath and Kavasaki, 2009). A wide range of biological databases and visualisation tools, which have access for the public were developed for different purposes and were categorised into various levels and according to different methods (Zou *et al.*, 2015). There are many databases that index the information based on biological networks and pathways. Including good examples like the KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway (Kanehisa *et al.*, 2008), STRING (database of known and predicted protein-protein interactions) (Szklarczyk *et al.*, 2014), Reactome (an open source and peer reviewed pathway database, and others (Croft *et al.*, 2013; Zou *et al.*, 2015). However, the limited information about the nature and direction

of the interactions and reductionist-based experiment construction are the main disadvantage of these databases.

1.19. Machine learning and Artificial Neural Networks

Machine learning (ML) is described as a field of computer science in which utilising statistical techniques enables the computer systems to "learn" with data by improving the performance on a specific task gradually without programming. ML has been employed in a variety of computing tasks where creating and programming clear algorithms with decent functioning is problematic or infeasible. In other words, ML is a branch of Artificial Intelligence that connected the general concept of inference, from the question of learning in samples data (*Kourou et al., 2015*). An example of this application includes email filtering, detection of network intruders and computer vision. Likewise, in the field of data analysis, ML has been used as a method of interrogating complex patterns in data and producing algorithms that advanced themselves to prognostication. The predictive analytics example enables the researchers, data scientists, engineers, and analysts to make an effective and right decision, as well as to detect the hidden insights via learning from historical relationships and the tendency of the data (*Kourou et al., 2015*). Further, these algorithms can be distributed into unsupervised learning, supervised learning, and reinforcement learning procedures, which all have a prospective implementation in biology. The employment of a variety of ML techniques, including Artificial Neural Networks (ANNs), Bayesian Networks (BNs), Support Vector Machines (SVMs) as well as Decision Trees (DTs) have been widely used to develop a predictive model in cancer research, which resulting in an efficient and precise decision action (*Kourou et al., 2015*).

Indeed, Artificial Neural Networks (ANNs) are a form of machine learning, which have been designed simulating the human brain in the way that processing information and learning capabilities (*Bishop, 2006*). Besides, the main characteristics of ANNs are highly complicated, nonlinear, and with equivalent information processing abilities. However, the unique features of ANNs including the highly fault and failure tolerant, scalable, and consistent generalisation ability that enabled them to predict or classify a new and unlearned data (*Livingstone, 2008, Lancashire et al., 2009*). As already mentioned, ANNs were inspired by the information-processing system in the human brain to handle several classification and pattern identification problems (*Haykin, 2009; Lancashire et al., 2009*). The basic structural units of the human brain are the neurons which are interconnected systematically to effectively transfer and processing the information. Neurons have the characters of plasticity - adapting to the surrounding environment. Similarly, artificial neurons are the primary processing unit of a computational neural network and can alter their internal weights, just like synapses, to optimise the network

outputs for a given training dataset. Moreover, ANNs have been trained to combine between the input variables to “generate output data. Despite ANNs having been used in several classifications tasks as a “gold standard system”, they still have some drawbacks. ANNs proved to be a time-consuming method with a very poor performance, which results from their generic layered structure. This specific technique has been described as “black-box” technology. It is almost impossible to detect how did ANN performed the classification process or why it did not work (Ayer et al., 2010; Kourou et al., 2015).

1.20. Stepwise ANN approach

It has been shown that ANNs have extraordinary abilities for data mining and model recognition (Bishop, 1995). While, the complexity of high throughput data generated for biomedical research made it challenging to process. One of the master restrictions for ANNs enforcement in such technologies, particularly microarray data, is the Curse of Dimensionality. Bellman firstly states this term as, “the exponential growth of the input space as a function of dimensionality” (Bellman, 1961), which means that the value of a specific trait or gene can be invisible according to the vast number of other gene expression vectors. This phenomenon occurs when the number of parameters (P) is extremely higher than the number of cases (n), which allows the added noise from irrelevant inputs in the data space to reduce model performance on blind data (Bishop, 1995). Several pre-processing and data-reductionist methods such as Principle Component Analysis (PCA) have been used to tackle the dimensionality aspect. Yet, simplification and feature withdrawal aspects are still challenging (Clarke et al., 2008). A Stepwise ANN approach developed in-house and already published (Lancashire et al., 2008), allows the identification of patterns within the datasets in which finding the set of individual variables that gives the best predictive performance to classify the dataset for a task. The learning in this model was carried out through a repeated manner along with the following addition of variables. Initially, the variation of each gene (n) from a microarray experiment was used as a single input node for the model, to produce n number of models, each n being a gene of the dataset. Afterwards, all the models can be compared and sorted correspondingly to their predictive performances for the hidden cases from the tested samples. Then, the learned weights and model specification can be applied for the further (n-1) example of input, to help the stepwise-additive features and contribute to the most predictive performance during each repetition. This process is repeated until no more advances in predictive performance can be seen in the model, or until the model reach the extent of best performance (Lancashire et al., 2008). To produce a better-generalised model with an improved predictive ability for the hidden

case from the validation subset, a Monte Carlo Cross Validation (MCCV) strategy can be applied to each model. In this design, the samples are randomly partitioning into a ratio of 60:20:20 for 50 iterations or bootstraps, which is, training (for model learning), test (assessing model performance during training by early stopping) and validation subsets (to independently test model on hidden data) respectively. This method has been proved as the most consistent models when no more significant improvement was observed with more loops (*Lemetre, 2010*). In addition, all the variables in the dataset have been scaled between 0 to 1 prior to being used an input for the algorithm. The parameters have been set to 3000 as the largest number of epochs, with a 1000 epochs window time, learning rate of 0.1, a mean squared error (MSE) threshold of 0.01, and momentum of 0.5. This arrangement was selected according to earlier studies, as given the best learning time and generalised models when applied successfully to gene expression datasets (*Lancashire et al., 2008, Powe et al., 2014*). Furthermore, the first weights are set randomly between -1 and 1, in addition to a constrained variance of 0.1. Only 1 step of the algorithm has been running, to produce a single gene model over 10 independent runs. To get a consistent result, a rank order of all the generated genes is constructed on the least average RMS error for the experimental subset across 10 independent runs.

1.21. ANN interaction algorithm

The exponential increase in the application of gene expression profiling technologies allowed the identification of biomarkers, and signatures associated with specific disease, or disease state through analysing high throughput data using many computational techniques. However, relevance and reliability of those selected biomarkers to biological processes are still controversial, especially for those biomarkers, which been found by such approaches and currently are employed routinely in clinical use. Additionally, it has been understood that the cause of any process can rely on a complex network of interacting genes, rather than one dependent single molecule. This has led to the application of gene expression-based analysis for a selected process as a systems perspective of a disease (*Barabasi et al., 2011*). The gene set enrichment analysis (GSEA) (*Subramanian et al., 2005*) and web-based tools such as DAVID (*Huang et al., 2009*) are examples of a gene set or a pathway analysis. More developed protein and gene interaction networks and regulatory pathways were being created including, Bayesian approaches (*Hartemink et al., 2002*), likelihood approaches (*Liu et al., 2005*), dynamic ordinary differential equations (*Christley et al., 2009*) and more recent recurrent neural network models (*Tong et al., 2014*). All the above-mentioned methodologies have the main limitation: they can identify limited information regarding the interactions only, rather than extract all the potential information contained with the data. A demand for more iterative quantification and prediction of the influence of multiple genes have been put forward. As an alternative method, the

Interaction Algorithm Model has been developed to convince the quantification of the entire genes within the dataset. Unlike the ANN, stepwise analysis can find the influence of the most important genes within a system or phenotype, rather than decide the best predicTable variables in a categorical outcome.

1.22. Filtering of interactions and visualisation of model

Finally, the interactions matrix is obtained as the output of the algorithm, which is the pairwise organisation combinations between all the variables in a specific subset. To select the relevant associations only between the genes out of the enormous number of interactions ($n(n-1)$) that can be generated, the interactions with the highest absolute value (positive or negative to support the nature of interaction) only are considered after filtering out all the non- significant associations (*Lemetre et al., 2009*). Although, selected biological pathways or genes of interest can be implemented by applying advanced defined filtering strategies, which are be based on the standard error across the 10 loops based on the consistency of interactions. Cytoscape (Version 3.2.1), which is an open source platform for visualisation and complex network analysis has been used to visualise the interactome of the pairwise interactions between the components (*Smoot et al., 2011*). In the interactome map, each node stands for a gene probe (source or target), while, the interaction between a pair of genes is represented by an arrowhead as a directed edge or link, which pointed to the target. Moreover, the colour of each edge is coded as red (positive) and blue (negative), as well as the width of the edge, is referring to the weight score of the interaction. Furthermore, *Barabási and Oltvai* have outlined the theoretical description of the interactive mapping concept, in which each gene is symbolised as a node, and the correlation between the markers as an edge (*Barabasi and Oltvai, 2004*). As well as the highly influenced genes or target hubs are those who have multiple (>5) in-degree interaction edges with the other nodes.

1.23. The aim of the project

This project aims to confirm the mechanism of Metformin action on breast cancer systems biology approach by integrating computational methods and laboratory data. The framework was structured using techniques of cell biology, molecular biology and systems biology.

Aims and Objectives

- To investigate the effect of Metformin on cellular viability, proliferation, and apoptosis using several cellular assays such as cell survival (MTT) assay, proliferation, and apoptosis assays.
- To inspect the drug mode of action on the level of gene expression in selected cell lines (MDA-MB-468 and SkBr3) through implementing the Gene Expression Microarray and NanoString techniques and various bioinformatics tools including Artificial Neural Networks (ANN) and Cytoscape among others.
- To interrogate the most influence gene in selected phenotypes of breast cancer (HER2+ and BASAL LIKE) *via* different statistical analysis and confirm the expression of *PTK2B* marker in separate ways utilising qRT-PCR, Western Blotting and immunofluorescence techniques.

Hypothesis:

Metformin has an antitumour effect in different breast cancer subtypes.

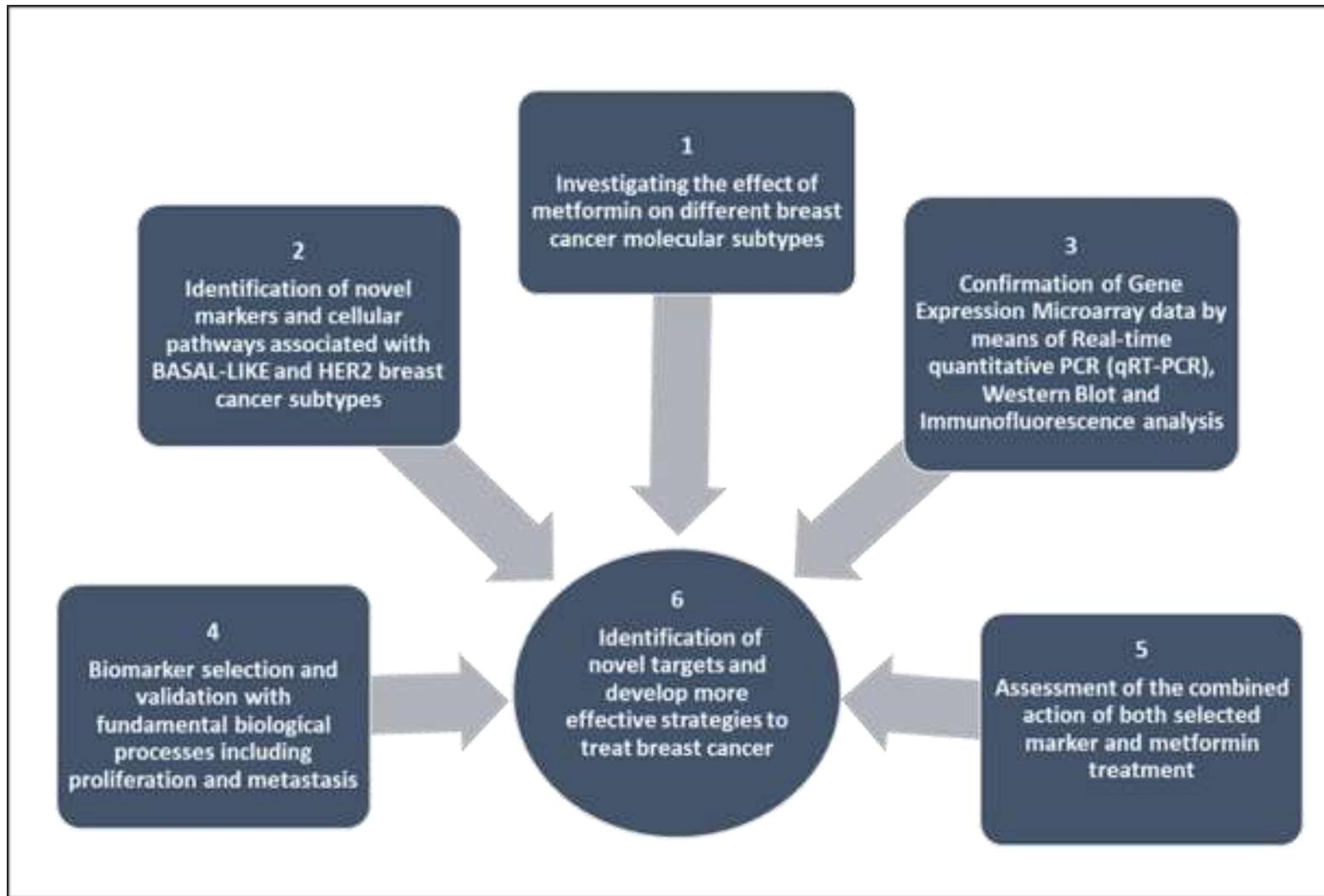


Figure 1. 12. The schematic representation of the main stages of the project framework.

CHAPTER 2

MATERIALS AND METHODS

2.1. Materials

2.1.1. Reagents

CELL CULTURE MEDIA	PROVIDER
EMEM	SLS (Lonza)
DMEM	SLS (Lonza)
Hybri-Care	ATCC® 46-X
Leibovitz's (L-15)	SLS (Lonza)
McCoy's	SLS (Lonza)
Opti-MEM®	Thermo Fisher Scientific

CULTURE MEDIA SUPPLEMENTS	PROVIDER
Foetal calf serum (FCS)	Fisher (GE Healthcare)
L-Glutamine	SLS (Lonza)
Insulin-Solution from bovine pancreas	SIGMA ALDRICH
Sodium bicarbonate	SIGMA ALDRICH

OTHER CELL CULTURE REAGENTS	PROVIDER
Dimethyl sulfoxide (DMSO)	Insight Biotechnology and Fisher Chemical
Dulbecco's phosphate buffered saline (DPBS)	SLS (Lonza)
Trypan Blue solution 0.4%	SIGMA ALDRICH
Trypsin/Versene	SLS (Lonza)
EDTA 0.5M	Ambion

DRUG USED IN THIS PROJECT	PROVIDER
1,1-Dimethylbiguanide hydrochloride 97%	SIGMA ALDRICH

ANTIBIOTICS	PROVIDER
Ampicillin	SIGMA ALDRICH
Pen/strep antibiotic solution	SLS (Lonza)
Puromycin	Life Technologies

CHEMICAL REAGENTS	PROVIDER
Agar	Bioline
Ammonium Per sulphate (APS)	Geneflow
β- mercaptoethanol	SIGMA ALDRICH
Bovine serum albumin (BSA)	Merck
Bromophenol blue	Arcos Organics
Calcium chloride (CaCl ₂)	SIGMA ALDRICH
Clarity Western ECL Substrate	Bio Rad
Chloroform	SIGMA ALDRICH
DAPI VECTASHIELD Mounting media	Vector Laboratories
DC™ protein assay reagent A, B, S	Bio-Rad
Dithiothreitol (DTT)	SIGMA ALDRICH
dNTPs	SIGMA ALDRICH
Double distilled water (ddH ₂ O)	Barnstead, Nanopure Diamond
Ethanol	Fisher Scientific
Ethyl alcohol absolute	PROLAB, Fisher chemical
Glycerol	SIGMA ALDRICH
HEPES	SIGMA ALDRICH
Hexadimethrine bromide solution(polybrene)	SIGMA ALDRICH
Isopropanol	Fisher chemical
ISOTON sheath fluid	Beckman Coulter
I TAQ™ Universal SYBR® Green Supermix	BIO-RAD
Lipofectamine 2000 Transfection Reagent	Invitrogen
Liquid nitrogen	BOC
Magnesium chloride (MgCl ₂)	SIGMA ALDRICH
Marvel skimmed milk	Co-operative
Methanol	Fisher Scientific
M-MLV Reverse transcriptase	Promega
OligodT15 primers	Promega
Paraformaldehyde	Arcos
Phosphate Buffer Saline (PBS)	Bio Whittaker Europe
Potassium chloride (KCl)	SIGMA ALDRICH
Presept (SPR25)	Johnson & Johnson
Propidium Iodide (PI)	SIGMA ALDRICH Aldrich
Protein Assay Dye Reagent Concentrate	Bio-Rad
Protease Inhibitor Cocktail	SIGMA ALDRICH
Protogel (30% Acrylamide mix)	Geneflow
RNA STAT-60	TEL-TEST-ING
RNasin	Promega
RT buffer(5x)	Promega
Sodium chloride (NaCl)	Calbiochem
Sodium dodecyl sulphate (SDS)	SIGMA ALDRICH
Solution 18-AO. DAPI	Chemometec
TEMED	SIGMA ALDRICH
Thiazolyl Blue Tetrazolium Bromide	SIGMA ALDRICH
Triton-X-100	SIGMA ALDRICH
1M Tris-HCl	Invitrogen
Trizma (Tris) base	SIGMA ALDRICH
Tween-20	SIGMA ALDRICH
Urea	SIGMA ALDRICH
Yeast extract	SIGMA ALDRICH

IMMUNOCHEMICAL REAGENTS	PROVIDER
Rabbit anti-human GADD45 α	Cell signalling
Rabbit anti-human IRF-9	Cell signalling and Abcam
Rabbit anti-human PYK2	Thermo Fisher Scientific or Invitrogen and Cell signalling
Rabbit anti-human SERPINB4	SIGMA-ALDRICH
anti-human β -actin	SIGMA ALDRICH
Rabbit anti-GFP	Abcam
Mouse anti-GFP	Abcam
Precision Plus Protein Western C Standards	Bio-Rad
Precision Protein™ Strep Tactin-HRP Conjugate	Bio-Rad
Anti-Rabbit IgG HRP-linked Ab	Cell Signalling
Anti-Mouse IgG HRP-linked Ab	Cell Signalling
Goat anti-Mouse IgG (H+L) Secondary Antibody, Alexa Fluor® 488 conjugate	Thermo Fisher Scientific
Goat anti-Rabbit IgG (H+L) Secondary Antibody, Alexa Fluor® 568 conjugate	Thermo Fisher Scientific
Goat anti-Mouse IgG (H+L) Secondary Antibody, Alexa Fluor® 568 conjugate	Thermo Fisher Scientific
Goat anti-Rabbit IgG (H+L) Secondary Antibody, Alexa Fluor® 488 conjugate	Thermo Fisher Scientific

REAGENT KITS	PROVIDER
Agilent RNA 6000 Nano Kit with RNA Nano Chips	Agilent
Alexa Fluor 647 Annexin V	Biologend
Annexin V Binding Buffer	Biologend
Cultrex® BME Cell Invasion Assay, 96 well	R&D Systems
CyQUANT® NF Cell Proliferation Assay Kit	Thermo Fisher Scientific
DC™ protein assay reagent A, B, S	Bio-Rad
Gene expression Hybridisation Kit	Agilent
Human Gene Expression 4x44K v2 Microarray Kit	Agilent
nCounter® Pan-Cancer Pathways Panel kit	NanoString® Technologies
One Color Spike-In Mix kit	Agilent
QIAGEN QIAfilter Plasmid Midi kit	QIAGEN
RNeasy Mini Kit (50)	QIAGEN

CELL LINES	PROVIDER
BT-474	ATCC
MCF-7	ATCC
MDA-MB-231	ATCC
MDA-MB-468	ATCC
MDA-MB-453	ATCC
HEK-293T c	ATCC
SKBR3	ATCC

PLASMIDS AND COMPETENT CELLS	PROVIDER
MISSION SHRNA PLASMID DNA TRCN0000000769(1)	MISSION™ /SIGMA-ALDRICH
MISSION SHRNA PLASMID DNA TRCN0000199334(2)	MISSION™ /SIGMA-ALDRICH
MISSION SHRNA PLASMID DNA TRCN0000231521(3)	MISSION™ /SIGMA-ALDRICH
MISSION SHRNA PLASMID DNA TRCN0000231523(4)	MISSION™ /SIGMA-ALDRICH
MISSION SHRNA PLASMID DNA TRCN0000199771(5)	MISSION™ /SIGMA-ALDRICH
PL-SIN-EF1 α -EGFP	Addgene
pLKO.1 puro	Addgene
psPAX2	Addgene
pMD2.G	Addgene
XL-1 blue competent cells.	NEB

2.1.2. Buffers and gels

LB AGAR PLATE WITH AMPICILLIN	FOR 500 ML
NaCl	5 g
Tryptone	5 g
Yeast Extract	2.5 g
Agar	7.5 g
ddH ₂ O	Up to 500 mL
Autoclaved, cooled down to 50°C	
Ampicillin	50 mg
Poured on Petri dishes, left to solidify and stored at 4°C for up to a week.	

LB BROTH	FOR 500 ML
NaCl	5 g
Tryptone	5 g
Yeast Extract	2.5 g
Autoclaved, cooled down to 50°C	
Ampicillin	50 mg
Stored at 4°C for up to a week	

1X ANNEXIN V buffer	FOR 0.5 L
HEPES	SIGMA ALDRICH
NaCl (SODIUM CHLORIDE)	OMNIPUR
KCl (POTASSIUM CHLORIDE)	SIGMA ALDRICH
CaCl ₂	SIGMA ALDRICH
DPBS Up to 1 L	LONZA
Adjust pH to 7.4 and stored at 4° C.	

TRIS-EDTA (TE) BUFFER	FOR 500 ML
1 M Tris pH 8	5 mL
0.5 M EDTA pH 8	1 mL
ddH ₂ O	Up to 500 mL

RIPA BUFFER	FOR 500 ML
1 M NaCl (150mM)	7.5 mL
1 M Tris-HCL, pH 8.0 (50Mm)	2.5 mL
10% sodium deoxycholate (0.5%)	2.5 mL
10% SDS (0.1%)	0.5 ml
10%TritonX100 (1%)	5 ml
100mM EDTA(1mM)	0.5 ml
dH ₂ O	31.5 ML
Add 1-10% protease inhibitor cocktail (sigma) freshly before use.	

LAEMMLI BUFFER (2X)	Volume
10% SDS (w/v) (4% final)	4mL
Glycerol (20%)	2mL
1M Tris-HCL, pH6.8 (125mM)	1.2mL
10% 2-mercaptoethanol	1mL
Distilled water	2.8mL
Add bromophenol blue to a final concentration of 0.02%.	
Add 1-10% protease inhibitor cocktail (Sigma) freshly before use.	

ERIKA'S BUFFER WITH PROTEASE INHIBITOR (EB + PI)	FOR 50 ML
Urea	28.5 g
Dithiothreitol (DTT)	1 g
N-Octyl-Beta-Glucopyranoside	0.5 g
ddH ₂ O	Up to 50mL
Solution was mixed vigorously and sonicated until dissolved	
Proteinase Inhibitor cocktail	500 µL
Aliquots were stored at -80°C.	

4X SDS-PAGE LOADING BUFFER	FOR 10 ML
1M Tris-HCl pH 6.8	2.4 mL
Sodium dodecyl sulfate (SDS)	0.8 g
Glycerol	4 mL
DTT	0.5 mL
Bromophenol blue	4 mg
ddH ₂ O	3.1 mL
Aliquots were stored at -80°C.	

5% STACKING GEL	FOR 6 ML (1X)
dd H ₂ O	4.1 mL
30% Acrylamide mix	1.0 mL
1M Tris (pH 6.8)	0.75 mL
10% SDS	0.06 mL
10% ammonium persulfate	0.06 mL
TEMED	0.006 mL

15% RESOLVING GEL	FOR 10 ML (1X)
H ₂ O	2.3 mL
30% Acrylamide mix	5.0 mL
1.5 M Tris (pH 8.8)	2.5 mL
10% SDS	0.1 mL
10% ammonium persulfate	0.1 mL
TEMED	0.004 mL

10% RESOLVING GEL	FOR 10 ML (1X)
H ₂ O	4 mL
30% Acrylamide mix	3.3 mL
1.5 M Tris (pH 8.8)	2.5 mL
10% SDS	0.1 mL
10% ammonium persulfate	0.1 mL
TEMED	0.004 mL

8% RESOLVING GEL	FOR 10 ML (1X)
H ₂ O	4.6 mL
30% Acrylamide mix	2.7 mL
1.5 M Tris (pH 8.8)	2.5 mL
10% SDS	0.1 mL
10% ammonium persulfate	0.1 mL
TEMED	0.006 mL

10X SDS RUNNING BUFFER	FOR 1 L
Glycine	144g
Trizma base	30.3 g
SDS	10 g
ddH ₂ O	Up to 1 L
10X Running buffer was diluted with ddH ₂ O to 1X working concentration prior use. Running buffer was stored at 4°C.	

1X TRANSFER BUFFER	FOR 2 L
Glycine	5.8 g
Trizma base	11.6 g
10% SDS	0.75 g
Methanol	400 mL
ddH ₂ O	Up to 2 L
Transfer buffer was stored at 4° C.	

10 X TRIS-BUFFERED SALINE (10 X TBS)	FOR 1 L
Trizma base	24.2 g
NaCl	80 g
ddH ₂ O	Up to 1 L
Adjust pH to 7.6 with concentrated HCl.	

1X TRIS-BUFFERED SALINE WITH TWEEN (TBST)	FOR 1 L
10 X TBS	100 mL
ddH ₂ O	900 mL
Tween-20	1 mL

2.1.3. Equipment

LABORATORY PLASTICS, GLASSWARE AND SHARPS	PROVIDER
Cell culture flasks (T25, T75, T175)	Sarstedt, UK
Coverslips	SLS
Conical flasks (50 ml,100 ml)	Pyrex
Corning™ Disposable Vacuum Filter/Storage Systems, 0.45µm PES, 500 mL	Fisher Scientific
E-Plate 16 PET	ACEA Biosciences
Eppendorf tubes (0.5 ml, 1.5 ml, 2 ml)	Sarstedt, UK
FACS tubes	Tyco healthcare group
Falcon tubes (50 ml, 15 ml)	Sarstedt, UK
Filter tips (0.5-10 µl, 2-20 µl, 20-200 µl, 200-1000 µl)	Greiner bio-one/ Sarstedt
Flat-bottom culture dishes (6, 24, 96-well)	Sarstedt, UK
Glass coverslips & slides	SLS
HyperSep™ SpinTip Microscale SPE Extraction Tips	Thermo Fisher Scientific
Micro tips (0.5-10 µL, 20-200 µL, 200-1000 µL)	Sarstedt, UK
MS Separation Columns	Miltenyi Biotech
Magnetic cell separators Mini MACS	Miltenyi Biotech
NC-slide A8	Chemometec
Nitrocellulose Membranes 0.2µm	GE Water & Process Technology
Pasteur pipettes	Sarstedt, UK
Petri dishes	Sarstedt, UK
Pipettes (5mL, 10mL, 25mL)	Sarstedt, UK
Cell Scraper	SLS (Swann Morton)
Screw-top tubes (15mL, 50mL)	Sarstedt
Serological pipettes	Sarstedt
Superfrost™ Microscope Slides	Thermo Fisher Scientific
Syringes (10ml,20ml)	Becton Dickenson
Timer	Cell signalling
Universal tubes (20ml)	Greiner
Western blot filter paper	Schleicher-Schuell

0.45 µm syringe filter	Sartorius
0.22 µm syringe filter	Sartorius
40 µm nylon strainer	Greiner
70 µm nylon strainer	Greiner

LABORATORY EQUIPMENT	
4°C refrigerators	Lec
-20°C freezers	Lec
-80°C freezers	Revco/ Sanyo
96-well plate reader	Tecan
Autoclave	Rodwell
Bacterial cell orbital incubator	Stuart
Bacterial cell culture plate incubator	Genlab
Cell culture incubator	Sanyo
Centrifuges	Sanyo, Eppendorf
CCD camera -Western blot imager	Syngene
Class II safety cabinets	Walker
Electrophoresis equipment	Bio-Rad
Fluorescence microscope	ZEISS
Fluorescence microplate reader	Tecan ULTRA
Flow cytometer	Beckman Coulter
Freeze vacuum dryer	Virtis
Gene Pix®4100A Microarray Scanner	GE Healthcare
Haemocytometers	SLS
Heat blocks	Lab-Line
Light microscope	Nikon/Olympus
Mass spectrometer Triple-TOF 6600	Sciex
Nucleo Counter®NC-250™	Chemometec
Vacuum drier	Eppendorf
Hybridisation oven	SHEL LAB
Microcentrifuge	MSE
Nanodrop 8000 Spectrophotometer	Thermo Scientific
NanoDrop ND UV-VIS Spectrophotometer version 3.2.1	Thermo Scientific
NanoString	Technologies
pH meters	Metler Toledo
Pipettes and multichannel pipettes	Gilson, Star Labs, Eppendorf
Plate rocker	VWR, Stuart
Real-time cell analyser System	xCELLigence
Rotor-Gene Q	QIAGEN
Measuring scale	Fisher Scientific
Sonicator	VWR
Spectrophotometer for 96-well plate	Tecan ULTRA
Transfer tank	Bio-Rad
Ultracentrifuge Optima TLX	Beckman
Ultrapure water dispenser	Barnstead
Vacuum filtration unit	Sarstedt
Vortex	Scientific Industries
Water baths	Clifton

SOFTWARE	
Agilent Feature Extraction Software	Agilent
Axiovision Microscopy Software 4.7.1. version	ZEISS
Cytoscape	open source bioinformatics software
Gene Pix Pro-software	Agilent
GraphPad Prism 7	Graph Pad software
Image J 1.50c	Image J software
Kaluza 3.1 version	Beckman Coulter
MetaCore™	Thomson Reuters
Morpheus	Broad Institute online software
OneOmics™ software	Sciex
Panther	Online databases
Protein Pilot version 5	Sciex
Q-Rex Software	QIAGEN
Real-time cell analysis software	Xcelligence
TIBCO Statistica 13.3	TIBCO Statistica software

2.1.4. Cell line growth media

BT-474 COMPETE MEDIA	CONCENTRATIONS
Hybri-Care (powder)	-
FCS	10 %
Sodium bicarbonate	1.5 g/L
Cell-culture-grade water	1L

MCF-7 COMPETE MEDIA	CONCENTRATIONS
EMEM	-
FCS	10 %
Insulin-Solution	0.01 mg/ml (1µL/ml)

MDA-MB-468, 231, 453 COMPETE MEDIA	CONCENTRATIONS
Leibovitz's (L-15)	-
FCS	10 %
L-glutamine	1 %

SKBR3 COMPETE MEDIA	CONCENTRATIONS
McCoy's	-
FCS	10 %

HEK-293T COMPETE MEDIA	CONCENTRATIONS
DMEM	-
FCS	10 %
L-glutamine	1 %

2.2. Methods

2.2.1. Cell culture

2.2.1.1. Routine cell culture maintenance

BT-474 (breast ductal carcinoma derived from solid invasive ductal carcinoma of the breast). MCF-7, MDA-MB-231, MDA-MB-468, SkBr3 (breast adenocarcinoma derived from metastatic site: pleural effusion) and MDA-MB-453 (metastatic carcinoma derived from metastatic site: pleural effusion), representing five molecular breast cancer subtypes (Luminal B, Luminal A, Claudin-low, Basal-like, and HER2) respectively, were used from (*Jon Van Geest Cancer Research Centre /Nottingham Trent University*) where they were purchased from American Type Culture Collection (ATCC), to investigate the efficacy of Metformin on breast cancer. As well as, HEK-293T (human embryonic kidney) cell line was used for *PTK2B* transfection. Table 2.1. provides examples of breast carcinoma cell lines according to the molecular classification of breast cancer.

Table 2.1. Molecular classification of breast carcinoma

Classification	Receptor expression status	Example cell lines
Luminal A	ER ⁺ , PR ^{+/-} , HER2 ⁻	MCF-7
Luminal B	ER ⁺ , PR ^{+/-} , HER2 ⁺	BT474
Basal	ER ⁻ , PR ⁻ , HER2 ⁻	MDA-MB-468
Claudin-low	ER ⁻ , PR ⁻ , HER2 ⁻	MDA-MB-231
HER2	ER ⁻ , PR ⁻ , HER2 ⁺	SKBR3, MDA-MB-453

The abbreviations are: ER, oestrogen receptor; HER2, human epidermal growth factor receptor 2; PR, progesterone receptor.

All cell lines were cultured in their dedicated media. BT-474 cell line was cultured in Hybri-Care media. Minimum Essential Medium Eagles (EMEM) was used to culture MCF-7 cell line with 0.01mg/ml insulin solution. LEIBOVITZ (L-15) with 1% L-Glutamine was used for MDA-MB-231, MDA-MB-468 and MDA-MB-453 cell lines, while Mc Coy's 5A was used for culturing SkBr3 cell line. In addition, DMEM with 1% L-Glutamine media was used for HEK-293T, 10% fetal bovine

serum (FBS) was added to all types of media as a supplementary agent according to ATCC culture methods. The 1% penicillin/streptomycin were used as antibiotics. BT-474, MCF-7, and SkBr3 were incubated at 37°C, in a humidified atmosphere with 5% CO₂, while MDA-MB-231, MDA-MB-468 and MDA-MB-453 were incubated at 37°C, in a humidified atmosphere without CO₂. To maintain the cells, they were routinely passaged at 70-80% confluence. During passage, cells were washed twice with Dulbecco's Phosphate Saline (DPBS) and detached through incubation with 0.25% (w/v) Trypsin- 0.53mM EDTA solution for 5-10 min at 37°C. Equal amounts of cell-specific media were added immediately upon cell detachment and cells were then centrifuged at 260 g for 5 min. Cell counting was carried out, by re-suspending a harvested cell pellet in 1-3 mL of cell-dedicated media and re-suspending cell solution in Trypan blue 1:10. The haemocytometer was applied to count the total number of living cells and excluded the number of dead cells (blue stained cells) from the count. The cell pellet was re-suspended in fresh media and cells were re-cultured in culture flasks by passaging. The stock of each cell line was prepared in a freezing media (1 mL FCS + 10% DMSO) at approximately 1 x10⁶ cell number and stored at -80°C. Cells were thawed, gently resuspended in 10 mL cell-dedicated media and centrifuged at 150 g for 5 min, upon need. The previous step of media change was performed to ensure that DMSO removed from frozen cells sample and to increase the viability of thawed samples. Then, cell pellets were gently re-suspended in a fresh amount of their dedicated media and plated in a suiTable flask (T25 or a T75) according to the pellet size, and then cells were incubated at 37°C with or without 5% CO₂ appropriately to their requirements.

2.2.2. Interrogation Metformin mode of action in different breast cancer molecular subtypes

2.2.2.1. Metformin preparation

In brief, 0.165g of Metformin (Sigma-Aldrich, UK) powder was completely dissolved in 1ml of cell-dedicated culture media according to different cell lines by vortex, then filtered through a 0.22µm syringe filtering device immediately before use. Concentrations of 1M and 10⁻²M were used as stock concentrations. 1M stock solution was used to make the different concentrations of (1, 2, 5, 10, 15, 20, and 25mM), while (0.01, 0.05, 0.1 and 0.5mM) concentrations were prepared from 10⁻²M stock solution consecutively. The drug was freshly prepared each time before applying to the assay.

2.2.2.2. Cell viability assay (MTT)

Five human breast cancer cell lines (BT-474, MCF-7, MDA-MB-231, MDA-MB-468, and SkBr3) representing the five molecular breast cancer subtypes (Luminal B, Luminal A, Claudin-low, Basal-like, and HER2) respectively, were chosen to serve as in vitro models of breast cancer in the upcoming assays.

Cell viability assay was determined by the MTT (3-[4, 5-dimethylthiazol-2-yl]-2, 5-diphenyltetrazolium bromide; thiazolyl blue) (Sigma-Aldrich, UK) which is a colourimetric assay that measures the absorbance of formazan molecules produced by the reduction of the tetrazolium salt by a dehydrogenase enzyme. This assay is widely used to test the toxicity of various drugs or substances by assessing the mitochondrial activity. In brief, cells were seeded into a 96-well plate at a density of (1×10^3 - 1×10^4) cell/well in 200 μ L of a dedicated media per well, with 4 replicates for each concentration point. Plates were incubated for 24h. A cell-free media was added to another 4 wells to serve as a negative control. The media was removed carefully using a multichannel pipette, then the cells were exposed to different concentrations of Metformin and incubated for 24h and 48h respectively. The MTT dye was prepared at a concentration of 5 mg/mL by adding 250mg (0.25g) to 50 ml of DPBS and sterilised using a 0.22 μ m syringe filter device (this can be stored at 4°C in protected from light for up to 1 month). 20 μ L MTT reagent was added to each well and incubated in 37°C with or without 5% CO₂ for 2 to 4 h. All culture media was carefully removed using a multichannel pipette. Then, 250 μ L of DMSO was used to dissolve the produced insoluble formazan. Absorbance measurement was made at 570 nm using the TECAN ULTRA spectrophotometer. The results were transported directly to Excel. The average absorbance values from the control wells are then subtracted from corresponding test wells to remove background signals. (*Mosmann, 1983; Van Meerloo et al., 2011; Stockert et al., 2012*). Acquired data were then analysed in Excel, Dell Statistica 13 software and following statistical analysis was performed using Graph Pad Prism 7. The EC₅₀, P value, SD, SEM were calculated in this assay using Two-way ANOVA (multiple comparisons).

2.2.2.3. Cell proliferation assay

The CyQUANT® NF assay is based on the measurement of cellular DNA content via fluorescent dye binding. Cellular DNA content is highly regulated; it is closely proportional to cell number. Therefore, the extent of proliferation is determined by comparing cell counts for samples treated with drugs and untreated cells (controls). In this assay, cells were plated in a 96 microplate at a density of 1000–10,000 cells per well, allowing 24 hours for adhesion before proceeding to the next step in the protocol. The cells were treated with Metformin in different

concentrations and incubated for 24h and 48h respectively and following manufacturer recommendations. Fluorescence intensity was measured using a fluorescence microplate reader (TECAN ULTRA fluorescence spectrophotometer) at ~485 nm excitation and ~530 nm emission. The results were transported directly to Excel and analysed. Considering the subtraction of the average absorbance values from the control wells to remove background signals. Obtained data were then analysed in Excel, Dell Statistica 13 software and statistical analysis were performed using Graph Pad Prism 7. The P value, SD, SEM were calculated in this assay using Two-way ANOVA (multiple comparisons).

2.2.2.4. Flow cytometry analysis of Cell apoptosis

Flow cytometry is a popular laser-based technology that determines the expression of cell surface and intracellular granularity, characterizes and define different cell types in a diversified cell population. It also assesses the purity of isolated subpopulations and analyses size, shape and complexity of cells. In addition, it allows a multi-parameter analysis of single cells simultaneously, using lasers and fluorescence detectors. Cell characteristics are measured through the detection of the scatters and the light emitted by each cell passing the laser beam resulting in quantifiable electrol signals. Flow cytometry is predominantly used to measure fluorescence intensity, which produced by fluorescent-labelled antibodies to detect proteins or ligands that bind to specific cell-associated molecules such as propidium iodide (PI) binding to DNA. Beckman Coulter Gallios flow cytometer that is equipped with three lasers (Blue 488 nm, Red 638 nm and Violet 405 nm) and 10 detection channels (FL-1 to FL-10) was used in this study. Apoptosis, or programmed cell death, is a normal physiological process for the removal of unwanted cells. The loss of plasma membrane asymmetry is one of earliest features of apoptosis. In apoptotic cells, the membrane phospholipid phosphatidylserine (PS) is translocated from the inner to the outer leaflet of the plasma membrane. Annexin V, a Ca²⁺-dependent phospholipid-binding protein, has a high affinity to PS. So, fluorochrome-labelled Annexin V can be used for the detection of exposed PS using flow cytometry. For this reason, the apoptotic cells were analysed via Annexin V binding to phosphatidylserine (PS) at the surface of the cell by utilising two detection channels (FL-3 & FL-6).

2.2.2.5. Metformin treatment

A confluent BT-474, MCF-7, MDA-MB-231, MDA-MB-468, and SkBr3 cells were grown in 6-well plates with 2ml of cell-dedicated media. In addition, 24 hours of incubation has been allowed for adhesion. The media was carefully aspirated and replaced with 2ml of cell-dedicated media, additionally supplemented with a range of Metformin concentrations (0.01, 0.05, 0.1, 0.5, 1, 2, 5, 10, 15, 20, and 25mM) for 24 and 48 h. After the incubation period, all treated cells were gathered. First, floating cells(dead) in the supernatants were collected into a 50-mL tube. Remaining cells were washed twice with DPBS and detached as previously described in section (2.2.1.1). All cells (floating and adherent) were centrifuged at 260 g for 5 min, resuspended and diluted in DPBS to 1×10^6 /mL roughly and counted using the Nucleo Counter® and Solution 18-AO. DAPI dye, by mixing 50 μ L of cells suspension to 2.5 μ L of Solution 18, adds a 20 μ L of the mixture onto counting slides (sample/ chamber) and inserted into Nucleo Counter. The percentage of viable cells was obtained from the automatic calculations made by the software.

2.2.2.6. Annexin staining

A set of aliquots of 100,000 cells of BT-474, MCF-7, MDA-MB-231, MDA-MB-468, and SkBr3 treated with different concentrations of Metformin were prepared in triplicate (one cell line per experiment) and resuspended in 1 mL of DPBS. Then, cells were transferred into labelled FACS tubes (include 'unstained' and 'PI-only' controls than one tube per treatment condition). Cells were washed twice with 2 mL of cold Annexin V Binding Buffer and centrifuged at 400 g for 5 min. Following centrifugation, the supernatant was removed by flicking off a tube gently to avoid losing floating cells. Then, the cells were resuspended in 100 μ L of cold Annexin Binding Buffer, and for each tube 5 μ L of Alexa Fluor 647, Annexin V was added. Cells were vortexed gently and incubated at RT for 15 min in the dark. PI working stock (50 μ g/mL) was prepared during the incubation time. Finally, cells were resuspended in 40 μ L Annexin V Binding Buffer and 10 μ L of working stock PI, gently vortexed and analysed. The flow cytometry BECKMAN COULTER GALLIOS FLOW CYTOMETER was utilised to sort the proportion of living cells, early apoptotic cells, and necrotic cells in this assay. Gained data was then analysed in Kaluza 3.1 software and statistical analysis was performed using Graph Pad Prism 7. The P value, SD, SEM were calculated in this assay using Two-way ANOVA (multiple comparisons).

2.2.2.7. Morphological effect of Metformin

The morphological effect of Metformin on different breast cancer cells was observed after 24h and 48h respectively. Bright field images of treated cells with different concentrations of Metformin have been captured using the light microscope (X10) to evaluate Metformin efficacy.

2.2.3. Microarrays

2.2.3.1. RNA-Extraction with STAT 60 and Qiagen RNeasy Mini Kit

In this assay and following assays, two concentrations of Metformin were considered, which are $\frac{1}{2}$ EC50 and 2 EC50 (1mM and 4mM) respectively. All assays were applied in triplicates for three different time points.

MDA-MB-468 and SkBr3 cells were cultured in T75 flasks in the density of 4×10^6 and 13 ml of cell-dedicated media for 24h. Then, media was replaced with a 13ml cell-dedicated media supplemented with a range of Metformin concentrations (1mM and 4mM), while the control flasks media was replaced to free-drug media. All flasks were incubated in their proper incubation condition (as described previously in Section 2.2.1.1) for 48h. After treatment with Metformin for 48h, cells were washed twice with DPBS (all the DPBS should remove carefully) and collected with 500 μ L of RNA STAT-60 by cell Scraper into a 2ml Eppendorf tubes and mixed well before incubation on ice for 5 minutes (pellet can be used immediately or frozen down in -80 C°). Next, total RNA was isolated by using the Qiagen RNeasy Mini Kit as described in manufacturer protocol. The RNA was eluted in 30 μ L of RNase free water. The concentration and integrity of RNA were determined using the Nanodrop and the Bioanalyser instruments, respectively. In addition, RNA Agilent Nano Kit with RNA Nano Chips was used as recommended by manufacturer protocol.

2.2.3.2. Labelling with a fluorescent dye

As has been recommended by the manufacturer, Agilent One Color Spike Mix was prepared. Concisely, One-Color Spike Mix stock solution was thoroughly mixed by a vortex mixer and heated for 5 min at 37°C and vortexed again. Serial dilutions of Spike Mix stock were prepared. A 'First Dilution', has been made by thoroughly mixing 38µL of Dilution Buffer with 2µL of Spike Mix stock. The 'Second Dilution', was also prepared by adding 2µL of Spike Mix stock to 48µL of Dilution Buffer and mixed thoroughly. While, the 'Third Dilution' has been prepared as a mixture of 4µL Spike Mix stock and 36µL of Dilution Buffer, which also has mixed very well. Distilled RNase-free water has been used to dilute the RNA samples to the final concentration of 134 ng/µL and mixed thoroughly. Into a fresh tube, 1.5µL (200 ng) of each sample was aliquoted, then 2µL of diluted Spike mix was added to each tube and mixed well. Master mix T7 Promoter Primer was prepared as shown in Table 2.2. For each tube, 1.8µL of T7 Promoter Primer Mix was added, then, incubated the reaction at 65°C in a rotary water bath for 10 min to denature both primer and template.

Table 2. 2. T7 Promoter Master Mix

Component	1X	18X	20X
T7 Promoter Primer (Green Cap)	0.8 µL	14.4 µL	16 µL
Nuclease-Free water (Whitecap)	1 µL	18 µL	20 µL
Total volume	1.8 µL	32.4 µL	36 µL

Table specifies the volume of reagents (experimental tubes) for 1x, 18x and 20x reactions. All reagents listed in this Table are included in Agilent low input QuickAmp one-colour labelling kit.

After the 10-min incubation at 65°C, tubes were immediately incubated on ice for another 5 min and spun down briefly. In the meantime, cDNA Master Mix was prepared as shown in Table 2.3 below. Firstly, 5X of First Strand buffer has pre-warmed for 3-4min at 80°C to ensure sufficient resuspension of the buffer components. The Affinity Script RNase Block mix was incubated on ice and was added immediately prior use to cDNA Master Mix.

Table 2. 3. cDNA Master Mix

Component	1X	18X	20X
5x First Strand Buffer (Green Cap)	2 μ L	36 μ L	40 μ L
0.1M DTT (White Cap)	1 μ L	18 μ L	20 μ L
10mM dNTP mix (Green Cap)	0.5 μ L	9 μ L	10 μ L
Affinity Script RNase Block Mix (Violet Cap)	1.2 μ L	21.6 μ L	24 μ L
Total volume	4.7 μ L	84.6 μ L	94 μ L

Table specifies the volume of reagents (experimental tubes) for 1x, 18x and 20x reactions. All reagents listed in this Table are included in Agilent low input QuickAmp one-color labelling kit.

The following step, for each tube a volume of 4.7 μ L cDNA Master Mix was added, pipetting up and down gently to mix and incubate in a water bath for 2h at 40°C. After the incubation time, samples were transferred to another water bath at 70°C and incubated for further 15min. Finally, the samples were moved onto the ice to cool down, they were briefly spun down. Table 2.4. below shows the preparation of RNA Transcription Master Mix, by adding a 6 μ L of Transcription Master Mix to each tube, and gently were mixed by pipetting then incubated at 40°C circulating water bath for 2h.

Table 2. 4. RNA Transcription Master Mix

Component	1X	18X	20X
Nuclease-Free water (Whitecap)	0.75 μ L	13.5 μ L	15 μ L
5X Transcription Buffer (Blue Cap)	3.2 μ L	57.6 μ L	64 μ L
0.1M DTT (White Cap)	0.6 μ L	10.8 μ L	12 μ L
NTP mix (Blue Cap)	1 μ L	18 μ L	20 μ L
T7 RNA Polymerase Blend (Red Cap)	0.21 μ L	3.78 μ L	4.2 μ L
Cyanine 3-CTP	0.24 μ L	4.32 μ L	4.8 μ L
Total volume	6 μ L	108 μ L	120 μ L

Table specifies the volume of reagents (experimental tubes) for 1x, 18x and 20x reactions. All reagents listed in this Table are included in Agilent low input QuickAmp one-color labelling kit.

2.2.3. Purification of the labelled/amplified RNA

To purify the amplified cRNA (RNA that labelled with cy3 dye) samples, QIAGEN RNasy mini spin columns were used. Firstly, for each cRNA samples, 84µL of Nuclease-free water was added (to make a volume of 100µL). Then, RLT buffer was added in 350µL, after mixing 250µL of 100% ethanol was added and mixed by pipetting up and down. Each cRNA sample in a volume of 700µL was transferred to an RNasy spin column, which was set on 2 mL collection tube. Then, the samples were centrifuged for 30 seconds at 10,000 g and 4°C. The flow-through was discarded. The RNasy column was transferred onto a new set of collection tubes, and 500µL of RPE buffer was added, which contained ethanol, to each column. Columns then were spun for 30 seconds at 10,000 g and 4°C and discarded the flow through. Another 500µL of RPE buffer (containing ethanol) was added to the column and columns were spun down for 60 seconds at 10,000 g and 4°C and the flow through was discarded. Again, columns were centrifuged briefly to dry the membrane.

Finally, for each RNasy filter membrane, 30µL of RNase-free water was added directly and centrifuged for 30 seconds at 10,000 g and 4°C. The purified RNA was collected as the flow through.

2.2.3.4. cRNA quantification

Quantification of cRNA was made using NanoDrop ND UV-VIS Spectrophotometer version 3.2.1. Measurement has taken by selecting a sample type of 'Microarray Measurement' and 'RNA-40'. The sample loading area was cleaned by wiping with nuclease free water. Then, 1 µL of nuclease-free water was loaded to each pedal to initialize the measurement. 1µL cRNA sample was used for the measurement. Data provided by this measurement was related to Cyanine 3 dye concentration (pmol/µL), the RNA absorbance ratio (260/280) and cRNA concentration (ng/µL). The cRNA yield and specific activity of each reaction were calculated based on those results, and according to the formulas below:

$$\text{cRNA yield: } \mu\text{g of cRNA} = \frac{(\text{Concentration of cRNA}) \times 30\mu\text{L (elution volume)}}{1000}$$

$$\text{Specific activity: pmol Cy3 per } \mu\text{g cRNA} = \frac{\text{Concentration of Cy3}}{\text{Concentration of cRNA}} \times 1000$$

According to manufacturer protocol: Sample concentration and volume needed for 1.65 µg yield were calculated based on the formulas below:

$$\text{Concentration: } (\mu\text{g}/\mu\text{L}) = \frac{\text{cRNA yield } (\mu\text{g})}{30\mu\text{L (sample vol)}}$$

$$\text{Vol needed for 1.65}\mu\text{g } (\mu\text{L}) = \frac{1.65 \mu\text{g}}{\text{Concentration } (\mu\text{g}/\mu\text{L})}$$

2.2.3.5. Hybridisation

All reagents used in these sections were supplied in the Agilent low input QuickAmp one-colour labelling kit and were used according to manufacturer protocol. Firstly, 10x Blocking Reagent solution was prepared by adding 500 µL of Nuclease-free water to lyophilised 10x Gene Expression Blocking Reagent vial. Secondly, Fragmentation Master Mix was prepared by adding 1.65 µg cRNA with 11 µL 10X Blocking Agent, and 2.2 µL Fragmentation Buffer and filled up to 55 µL volume with Nuclease Free water. Then, Fragmentation Mix containing samples were incubated for 30 min at 60°C, and immediately after incubation cooled down for 1 min on ice, and 55 µL of 2x GEx Hybridisation Buffer HI-RPM was added to each sample to stop fragmentation reaction. Samples mixed by gentle pipetting, then centrifugate at 10,000 g for 1 min at RT. Finally, samples rested on ice, and load onto the array.

Nine microarray slides (4-pack each) of Agilent Technologies Sure Print Technology G4845A Human GE 4x44 V2 Microarray Kit were used exactly as recommended by the manufacturer. Gasket slides were removed from the protective packaging (handled from the edges) and immediately placed in the chamber base to avoid any potential contamination from surrounding surface material. Slide was held with the label reading “Agilent” face up, which put the gasket slide of the slide side up. The chamber base guide points and the rectangular barcode guide were located adequately in the chamber base. Gasket slide was gently placed into the chamber base. 100 µL of Hybridisation sample mixture was loaded onto the gasket slide surface by adding a drop in the middle and avoiding touching the sides. The microarray slide was grabbed from the numeric barcode side facing up and “Agilent” label facing down and lowered on the top of the gasket slide carefully. The chamber cover was placed on the side facing up the chamber base and which contains both “sandwiched” slides. The assembled chambers were loaded into the hybridisation oven rotator rack. Finally, samples were hybridised for 17 h at 10 rpm at 65°C. The Gene Expression Wash Buffer was prepared by adding 2 mL of the provided 10% Triton X-102 into the wash buffers. Gene Expression Wash Buffer 2 was warmed overnight at 37°C. First

three staining dishes, magnetic stirrers and magnetic stir plates were thoroughly and repeatedly washed with 100% acetonitrile, air dried in the fume hood and filled up with Mili-Q water and emptied. This cleaning procedure was repeated five times. Hybridisation chambers were removed from the oven at the record time, and the array-gasket sandwich was removed and quickly transferred into the slide staining dish 1 with Gene Expression Wash Buffer 1 at RT. The gasket was gently separated from the array using the blunt end of forceps and submerged in the buffer. The wet array was quickly put on to slide rack and placed in slide staining dishes 2 filled with Gene Expression Wash Buffer 1 at RT and incubated for 1 min stirring at RT. The rack with arrays was quickly transferred to the slide staining dishes 3 filled with pre-warmed Gene Expression Wash Buffer 2 and incubated for 1min stirring at 37°C. Finally, the slide rack was removed from the staining dish 3 and the residual liquid was removed by dabbing in a cleaning tissue and the microarray slides were put on a dry slide holder.

2.2.3.6. Microarray Data Analysis

Gene Pix® Pro 4100A Scanner was used to carry out the Image acquisition at 5 µm diameters. Feature extraction, quality docking and raw data generation were carried out using Agilent feature extraction software. Then, Partek genomic suite software was employed to Normalize the raw data, and to compute the statistical analysis and differential fold change (expression). Multiple testing correction using Student T-DIST, Bonferroni correction, and Benjamini Hochberg has been utilised to compute the P-value for each gene, which allows minimising the false discovery rate (FDR).

Two approaches have been used in this project to analyse the Microarrays data.

2.2.3.6.1. Artificial neural networks-based approach

Firstly, an in-house developed integrative Stepwise-ANN algorithm (*Lancashire et al., 2010*) based approach was applied to the entire data to identify the differentially expressed genes in both breast cancer phenotypes in terms of treatment with Metformin (the question). The initial set consisted of 36 samples each sample profile was associated with 33,128 gene probes. The input data was coded as 0 for untreated (control) samples, and 1 for treated samples. Prior to training the ANN, sample data were randomly divided into three subsets; 60% for training, 20% for validation (to assess the performance of the model during the training process) and 20% for

testing (to test the model independently on data completely blind to the model), and the run has performed for 20 loops. The initial multilayer perceptron (MLP)-ANN weights were randomised between -1 and 1 with a constrained variance of 0.1. By using a three-layered feedforward –backpropagation algorithm, these weights were updated using a learning rate of 0.1 and a momentum of 0.5 (Lancashire et al., 2009). For individual genes predictors, the median expression-based class split was carried out by (1) averaging the expression values for the multiple probes of the same gene across samples, when multiple probes of the same gene were present. (2) To classify a sample as high or low for the gene expression, the median expression value for the average expression has been calculated. The average across 10 independent runs was calculated, in order to identify the top predictive gene in the ANN model. Then, the stepwise results were imported to interaction algorithm and visualised as molecular interaction networks and integrate with gene expression profiles by utilising Cytoscape software, which is an open source bioinformatics software platform.

2.2.3.6.2. Regression-based method and Fold change (Fc) -based method

Secondly, to differentiate the expressed genes in each phenotype of breast cancer separately, the regression-based method has been followed by using the Excel's Regression tool provided by the Data Analysis add-in. Regression analysis was performed by utilising control values for each sample as Input X range and treated samples as Input Y range. A new analysis sheet was created by clicking all Residuals options. Then, the Standard Residuals was averaged across all the samples replicates and sorted according to the highest T. DIST and Bonferroni corrected values. Data obtained from this approach represented the two cell lines (MDA-MB-468 and SkBr3) in two concentrations (1 and 4 mM of Metformin) separately. The cut-off criteria were calculated as Bonferroni corrected value ≤ 0.05 .

A similar Fold change (Fc) -based method was also used to identify the genes that differentially expressed in each phenotype of breast cancer in different concentrations separately. Fc analysis was performed by utilising the following equation ($Fc = T/C$) after average all control and treatment values. Then, the genes were ranked according to highest T-test (P values) and Bonferroni corrected values. The Fold change (Fc) was calculated using one cut-off criterion ($Fc \geq 2$), p-value ($p \leq 0.05$).

The Data were clustered using hierarchical clustering heat maps on both entities and conditions with Euclidean distance metric and centroid linkage rule. Presented heat maps shown the differential expression of MDA-MB-468 and SkBr3 cell line samples. The Morpheus-Broad Institute online software was utilised to create the Hierarchical clustering heat maps.

2.2.4. Validation of Gene Expression Microarray data

2.24.1. cDNA synthesis

The previously isolated 36 RNA samples were used to synthesise the cDNA templates. The purified samples were quantified again by the Nanodrop, and the required volume for cDNA synthesis was taken. The first master mix was prepared in 0.5ml Eppendorf tubes by mixing 1 μ L of Oligo-dT15 primers with mRNA and nuclease-free water according to different samples concentrations. Samples were mixed gently and incubated in a thermal block at 70°C for 5 minutes. Then, tubes were transferred immediately onto the ice box for 5 minutes. Meanwhile, the second master mix was prepared according to Table 2.5. below.

Table 2. 5. cDNA Master Mix

Reagents	Volume for 1X	Volume for 18X
RT Buffer	5 μ L	90 μ L
MMLV RT	1 μ L	18 μ L
RNA sin	0.7 μ L	12.6 μ L
dNTPs	1 μ L	18 μ L
NF Water	7.3 μ L	131.4 μ L
Total	15 μ L	270 μ L

Table specifies the volume of reagents (experimental tubes) for 1x and 18x reactions. All reagents listed in this Table are included in the cDNA synthesis kit.

Then, 15 μ L of the second master mix was added to each tube of the first master mix (RNA oligo dT) and mixed thoroughly by gently pipetting up and down. Tubes were incubated in a water bath for 60 minutes at 40°C. After the incubation time, the reaction was inactivated by incubating the tubes in a thermal block at 95°C for 5 minutes. Finally, cDNA samples were stored at -20°C until use.

2.2.4.2. The application of Real-time quantitative PCR (qRT-PCR) and primers preparation

First: Preparation of the primers

All the primers vials were labelled according to the label sheet, as F & R (forward and reverse primers). Newly purchased primers were resuspended according to the manufacturer's recommendations by adding the accurate amount of nuclease-free water (N F W) to each vial up to 100 p moll and vortexed. The vials were kept for 30 minutes to dissolve completely. Working solutions were prepared by adding 10 µl from the primer stock to 90 µl of NFW (1:10) dilution. The stocks were stored in -20. SYBR Green was aliquoted into 500µl Eppendorf tubes to avoid freezing and thawing cycle and stored in -20.

Second: Preparation of PCR reaction Master Mix.

A master mix was prepared by mixing the reagent in Table 2.6 below. The master mix was calculated according to the number of samples + two extra (each sample in triplicates).

Table 2. 6. Real-time quantitative PCR (qRT-PCR) Master Mix

Reagents	1 x	24x	28x
CYBR Green	6.25 µL	150 µL	175 µL
N F W	0.5 µL	12 µL	14 µL
Forward Primer	0.5 µL	12 µL	14 µL
Reverse Primer	4.25 µL	102 µL	119 µL
Total volume	11.5 µL	276 µL	322 µL

Table specifies the volume of reagents (experimental tubes) for 1x, 24x, and 28x reactions.

This experiment was achieved using the first set of cDNA template samples followed this order (SkBr3 control, 1mM, 4mM, and MDA-468 control, 1mM, 4mM). The threshold was set on 0.1642 for the entire run, while the TM varied according to the different primers TM. 11.5 µL of the master mix was pipetted carefully into labelled PCR tubes (kept on ice) and 1µL of cDNA sample templates was added. The tubes were closed tightly and placed into the thermal cycle rotter. Then, PCR was carried out using the primers temperature profile. Finally, obtained data were analysed by utilising Rotor-Gene Q software and performed with Graph Pad Prism 7 software.

2.2.4.3. Western Blot analysis of protein lysates from MDA-MB-468 and SkBr3 cell lines

2.2.4.3.1. Preparation of the cell lysates

MDA-MB-468, MDA-MB-231, MDA-MB-453, and SkBr3 cells were cultured and treated with Metformin as previously mentioned in section (2.2.3.1, RNA-Extraction with STAT 60 and Qiagen RNeasy Mini Kit). Following the incubation period (48h), cells were washed twice with ice-cold DPBS buffer at 4°C and detached as previously described in section (2.2.1.1). Then, the cells were centrifuged at 260 g for 5 min at 4°C, resuspended in ice-cold DPBS. Cells were counted and diluted to 1×10^6 /mL roughly and spun down again at 260 g for 5 min at 4°C (all the DPBS was removed carefully). Pellet was placed immediately on dry ice for 10 minutes and stored in -20 °C. Pellets were dissolved in 100µl of Laemelli buffer with 10% protease inhibitor cocktail (PI) by vortexing vigorously. Then, the pellets were boiled in the thermal block at 99°C for 15 minutes. Protein quantification was carried out using the Bio-Rad Protein Assay.

2.2.4.3.2. Protein quantification

Prior to Western blot analyses, the total protein concentration in all lysate samples was assessed by using the Bio-Rad Dc protein assay and according to the supplier's protocol. The use of a standard curve generated from known protein standards is the traditional method to calculate a protein concentration of an unknown sample. Hence, a standard curve was generated according to Table 2.7. below.

Tube # (1) is a 2 mg/ml of Bovine Serum Albumin (BSA) solution in Laemelli buffer, which was prepared by dissolving 0.002g of BSA in one ml of Laemmli buffer.

Table 2. 7. Generation of protein standards

Tube #	Standard (μl)	Buffer of Choice (μl)	Final Concentration (mg/ml)	Final Volume (μl)
1	500μl of Starting 2mg/ml standard	0	2	500
2	750μl of Starting 2mg/ml standard	250	1.5	750
3	500μl of Starting 2mg/ml standard	500	1	500
4	250μl of Tube #2 Standard	250	0.75	500
5	500μl of Tube #3 Standard	500	0.5	500
6	500μl of Tube #5 Standard	500	0.25	500
7	500μl of Tube #6 Standard	500	0.125	1000
8	-	500	0	500

The table specifies the volume of reagents used to create the serial dilution of the standards in Protein Assay.

The required volume of working reagent has been calculated previously according to the formula below:

$$\text{The total volume of required working volume (mL)} = (N) \text{ of standards} + (N) \text{ of samples} \times (N) \text{ of replicates} \times 200$$

10μL of each standard and unknown sample were pipetted into a 96 well plate according to plate map. Then, 25μL of working reagent was added to each well-containing standards and unknown protein samples and mixed carefully by pipetting up and down. 200μL of reagent B was added to each well, then the plate was incubated in the dark at 37 °C for 15-30 minutes. The absorbance was measured at 570 nm using TECAN ULTRA spectrophotometer. The average absorbance values from control wells were then subtracted from corresponding test wells. Finally, sample protein concentrations were calculated as a reference to the standard curve.

To denature protein samples, a reducing agent such as dithiothreitol (DTT) or 2-mercaptoethanol was added in ratio 1:3. As well as, bromophenol blue served as tracking dye (~0.05 mg/ml). Samples were boiled for at 95°C for 5 minutes in the thermal block before loading into the gel.

2.2.4.4 Preparation of Gel and electrophoresis

The SDS PAGE gel was prepared as previously mentioned in section (2.1.2) and placed in the running tank, which was filled with 1X SDS running buffer (see section 2.1.2). Then, the denatured samples (section 2.2.4.3.2) were loaded to (30µg) each well carefully, alongside 5 µL of Precision Plus Protein Western Standards onto a polyacrylamide gel. The gel was run at 70V for 10-20 min through the stacking gel, and then the voltage was increased to 100V through the resolving gel using electrophoresis.

2.2.4.5. Wet transfer of proteins from gel to a membrane

After about 90 min the electrophoresis was stopped, and gels were carefully removed from the tank. Then, proteins were gently transferred onto a Nitrocellulose blotting membrane via “wet transfer”. The Nitrocellulose membrane was already cut into 8.5 x 5 cm pieces, which fitted the gel size, besides two pieces of filter paper of the same size (per membrane).

The transfer sandwich was assembled according to below order:

Black plate
Sponge
Filter paper
Gel
Nitrocellulose membrane
Filter paper
Sponge
Red plate

Air bubbles were carefully removed from between the layers, in order not to affect the transfer process. Assembly parts were kept soaked with transfer buffer all the time. Then, the sandwich cassette was inserted into a transfer tank immediately and the tank was filled with an appropriate amount of ice-cold transfer buffer. Lastly, the lid was placed on the transfer tank and the run was set at 100v for about 1h in the cold room(4°C).

2.2.4.6. Blocking and probing membranes with antibodies

After the transfer, the membrane was removed from the assembly cassette and washed 3 times with DPBS, rocking for 10 min. The membrane was then blocked with 5% Marvel skimmed milk in DPBS for 1 h rocking at RT. The membrane was then incubated with primary antibodies at 4°C on the rocking platform overnight. The following day, the membrane was washed three times with DPBS + 0.1% Tween-20 and incubated with the corresponding secondary antibodies for 1 h on the rocking platform at RT. Following the incubation, the membrane was washed with DPBS + 0.1% Tween-20. Finally, the membrane was incubated with Clarity Western ECL Substrate for 1 minute prior to exposure and developed using CCD camera. All antibody concentrations are shown in Table 2.8.

Table 2. 8. List of the antibodies that used in Western Blot analysis and their dilutions

Primary antibody	Dilution	Secondary antibody	Dilution
Anti- β -actin (M)	1:5000	Anti-Mouse IgG HRP-linked Antibody	1:1000
PYK2 momoclonal antibody (M).	1:1000		
IRF-9 momoclonal antibody (R)	1:1000	Anti-Rabbit IgG HRP-linked Antibody	1:1000
Anti-SERPINB4 polyclonal antibody (R)	1:1000		
GADD45 α monoclonal antibody(R)	1:1000		
Precision Plus Protein Western C Standards	5 μ L	Precision Protein™ Strep Tactin-HRP Conjugate	1:5000
Antibodies were diluted in 5% (w/v) Marvel skimmed milk in TBS			

(M) = produced in Mice; and (R) = produced in Rabbit.

2.2.4.7. Immunofluorescence analysis of proteins

The MDA-MB-468, MDA-MB-231, MDA-MB-453, and SkBr3 cells were cultured at a density of 5×10^4 in 24 well plates with glass coverslips at the bottom of each well and treated with Metformin for 48h, as previously described in section (2.2.6 Metformin treatment). However, 1mL of dedicated media was used, as well as two concentrations (1 and 4mM) of Metformin were used in this experiment. The following day, cells were washed three times with 1ml DPBS and fixed with 500 μ l of 4% Paraformaldehyde for 15 minutes. Cells were washed again with 1ml of (1x DPBS) for 10 min 3 times. Then, cells were blocked with 500 μ l of 10% BSA in (1x PBS+0.1% Tween20). 200 μ l of antibody was added to the blocking solution and incubated overnight at 4°C.

The following day, cells were washed with 1ml of (1x DPBS) for 10 min 3 times. 200µl of relevant secondary antibodies, which were diluted in the blocking solution, were added and incubated at RT for 1 h. Then, cells were washed with 1ml (1x DPBS) for 10 min 3 times. The remaining wash was left in the wells to ease the coverslips removal. Glass slides were cleaned with ethanol and paper towels to remove dust or debris and labelled properly. A small drop of mounting solution (vectashield+DAPI) was added in the centre of the glass slide and each coverslip was removed slowly from the well and was mounted on the top. Finally, the edges of the coverslip and the glass slide were sealed gently with nail varnish. Slides were viewed under a fluorescent microscope and the exposure for each fluorophore was adjusted to the controls (with secondary antibodies only). Slides were stored in the dark at 4°C. The names and dilutions of antibodies used in these experiments are shown in Table 2.9 below.

Table 2. 9. List of the antibodies that used in Immunofluorescence staining and their dilutions

Primary antibody	Dilution	Secondary antibody	Dilution
Anti-Interferon regulatory factor 9 antibody momoclonal antibody (R)	1:100	Goat anti-Rabbit IgG (H+L) Secondary Antibody, Alexa Fluor® 568 conjugate (RED & GREEN)	1:1000
Anti-SERPINB4 polyclonal antibody (R)	1:100		
FAK2/PYK2 momoclonal antibody (R).	1:100		
The diluent is: 10% BSA in (1x PBS+0.1% Tween20).			

(M) = produced in Mice; and (R) = produced in Rabbit.

2.2.4.8. NanoString n Counter XT Gene Expression Assay for gene expression profiling in Basal-like and HER2 phenotypes

Gene expression profile in 12 samples was analysed using the Pan-Cancer Pathways Panel, which consists of 770 genes from 13 cancer-associated canonical pathways, and 20 housekeeping genes. This assay was carried out according to the manufacturer instructions. All RNA samples were quality controlled using Nanodrop 8000, and 150ng of total RNA from each sample were used for setting up nanostring probe hybridisation overnight for (20hrs) at 65°C. The reaction Master Mix was prepared according to Table 2.10. below.

Table 2. 10. NanoString probe hybridisation Master Mix

Reagents	Volume for 1X	Volume for 12X
RNA solution	5µL (up to 150ng),	60 µL
Reporter probe	8µL	96 µL
Capture probe	2µL	24 µL

Table specifies the volume of reagents (experimental tubes) for 1x and 12x reactions. All reagents listed in this Table are included in the Pan-Cancer Pathways Panel kit.

After the overnight hybridisation completed, excess probes were removed using n Counter Prep Station, magnetic beads and hybridised mRNA/probe were immobilised on a streptavidin-coated cartridge. The processed cartridge was subsequently scanned using an n Counter digital analyser platform for the generation of the raw data with a high-resolution scan 555 fields of view (fov). Raw data were processed with nSolver Analysis Software (V.4.0), imaging quality control (QC), mRNA positive control QC and Normalization QC checked. All samples were within the quality parameters of nanoString gene expression assays. Differential expression, pathway scoring, and cell type scoring were performed using nSolver advance analysis module V. 2.0.115. Normalisation of the data was performed using the geNorm algorithm for the selection of the best housekeeping genes. Genes which showed ≥ 2 , fold change in their expression with a BY (Benjamini yekutieli) P value ≤ 0.05 were considered significant between the group.

2.2.5. Evaluation of *PTK2B* (Protein Tyrosine Kinase 2 Beta) role in breast carcinoma

2.2.5.1 *PTK2B* Plasmid bulking

Four individual clones from MISSION™ shRNA Target Set NM_004103 in plasmid format were purchased from SIGMA. A mandatory procedure to scale up plasmid vectors for long-term storage in glycerol (to provide a continuous supply), and mini-preparation was done using XL-1 blue competent cells, which is a strain of *E. coli* used for cloning procedures. Prior to transformation, XL-1 blue competent bacterial were defrosted on ice and a 5µl of 20 ng/µL diluted shRNA plasmid was added and incubated on ice for 30 minutes. Then, the cells were heated at 42°C for 3 minutes and immediately placed on ice for 10 minutes. 250 µL of RT Luria Bertini (LB) medium was added. The tubes were taped tightly and shaken horizontally (200rpm) at 37°C for 1h. 200 µL from each transformation was spread on 2 pre-warmed selective plates with 50µg/mL ampicillin and were incubated overnight at 37°C. The following day, a single colony was picked and inoculated into 50 mL LB Miller with ampicillin(50µg/mL) and incubated

at 37°C for 16h with shaking. The following day, 4 (0.5mL) cryovials from each bacterial solution were collected and 0.5 mL of glycerol (50%) was added for long-term storage at -80° C.

2.2.5.2. Plasmid isolation

The Isolation of *PTK2B* Plasmid was carried out using the QIAGEN QIAfilter Plasmid Midi Kit following the manufacturer instructions. The 50 mL LB tubes with inoculated bacteria, which (section 2.2.5.1.) were centrifugated at 6000 rpm in ultracentrifuge for 45 min at 4°C. The pellet was resuspended in 4 mL of (P1) buffer, then 4 mL of (P2) buffer was added and mixed by inverting the sealed tube 4-6 times. Tubes were incubated at RT for 5 min. Meantime, the cap was screwed into the outlet nozzle of the QIAfilter Cartridge and was placed in a convenient tube or a QIArack. Then, 4 mL of prechilled (P3) buffer was added to the lysate and was mixed thoroughly by inverting 4-6 times. The lysate was poured into a barrel of the QIAfilter Cartridge and incubated at RT for 10min without inserting the plungers. The QIAGEN-tip was equilibrated by applying 4 mL of QBT buffer and allowed the column to empty by gravity flow. The cap was removed from the QIAfilter Cartridge outlet nozzle. The plunger was gently inserted into the QIAfilter Cartridge and the cell lysate was filtered into the equilibrated QIAGEN-tip. The lysate was entered into the resin by gravity flow. The QIAGEN-tip was washed twice with 10mL of (QC) buffer. Then, DNA was eluted with 5 mL prewarmed (QF) buffer (at 65°C). DNA was precipitated by adding 3.5 mL of RT isopropanol, mixed and centrifugated at 6,000 rpm for 60 min at 4°C. The supernatant was carefully decanted. DNA pellet was washed with 1 mL RT (70% EtOH) and centrifugated at 15,000 g for 10 min. The supernatant was carefully decanted, and the pellet was dried for 5-10 min and then DNA was dissolved in 30 µl of TE buffer. Isolated plasmids were quantified using Nanodrop 8000 Spectrophotometer and stored in TE buffer at -20°C.

2.2.5.3. HEK-293T transfection

The HEK-293T (human embryonic kidney cells containing SV40 T-antigen) cells were cultured in T25 flasks and 4 mL of cell-dedicated media until 90% confluence. A mixture of 20µL Lipofectamine 3000 and 500 µL OPTIMEM media was prepared in a 1,5 mL tubes and incubated for 30 min at RT. Meanwhile, another mix of 8µg of target plasmid and 6µg packaging plasmid psPAX2 with 2µg envelope plasmid pMD2, plus 12µL P3000 and 500µL OPTIMEM medium was prepared.

The first mixture was combined with the second mixture and the final solution was added to each of HEK-293T T25 flasks with 4 mL of cell-dedicated media and incubated in the viral incubator at 37°C for 16h. The following day, the medium was changed in each HEK293T flask with 5mL of fresh HEK293T-dedicated media. One day after, the 1st fraction of media (F1) from transfected HEK-293T cell culture was collected in fresh tubes, spun to remove all unwanted cells and filtered through 40 µm nylon strainer. Fractions were aliquoted in 1 mL Eppendorf tubes and stored at -20°C. 5ml of fresh dedicated media was added to the 25 HEK cells flasks. The following day, the 2nd fraction (F2) was collected in a similar manner.

2.2.5.4. Infection of target breast cancer cell lines

The target cells (MDA-MB-468, MDA-MB-231, MDA-MB-453, and SkBr3) were diluted to 1×10^5 cells and cultured in 6-well plates with 2mL of cell dedicated media. A combination of 1mL of 1st fraction and 1mL of cell dedicated media with 16 µL of Hexadimethrine bromide solution (HB) was prepared and added to each well of target cells and incubated for 24h at 37°C. The day after, early in the morning, the media was changed to fresh cell dedicated media and in the late afternoon, the antibiotic was added. To assess the efficiency of the infection, the expression of the Enhanced Green Fluorescence Protein (EGFP) was measured by using a Carl Zeiss PALM MicroBeam fluorescent microscope and emission wavelength at 488nm.

2.2.5.5. Puromycin selection

Newly infected cell lines were treated with cell-dedicated media supplemented with Puromycin (3 µg/mL) with the aim of selecting cells containing lentiviral construct.

2.2.5.6. Whole cell lysate preparation for knockdown validation

Cells were harvested from three passages of post-antibiotic-selection, and protein lysates from the shRNA-infected cells were analysed by Western blotting using a *PTK2B* antibody following the previously stated protocol in section (2.2.4.3).

2.2.5.7. Cell Migration or Wound-healing assay (Scratch assay)

A wound-healing assay (Scratch assay) was applied to different breast cancer *PYK2* knockdown cell lines. Briefly, the cells were counted to 1×10^6 cells and seeded in 6-well plates with 2mL of cell dedicated media supplemented with Puromycin ($3 \mu\text{g/mL}$) as non-treated cells and supplemented with Puromycin and Metformin ($1 \mu\text{L/mL}$) as treated cells. The following day the media was replaced, and serum-free media was added to starve the cells (to prevent proliferation). After 24h of starvation, the media was replaced, and cells washed with 1mL (DPBS) per well. Scratches were performed using 200 μL pipette tip and ruler. The cells were washed with 0.5mL (DPBS) twice and fresh serum-free media was added to each well (with or without Metformin). Measurement was directly carried out at time 0 by utilising the LCM and Axiovision software for imaging the scratches. The measurement was repeated after 24h and 48h of treatment respectively. Finally, the percentage of gap closure was calculated according to the formula below:

$$\% \textit{Gap closure} = \frac{\text{Measurement at 24h}}{\text{Measurement at 0h}} \times 100$$

2.2.5.8. Cultrex® BME Cell Invasion Assay

The invasion assay was implemented according to the Cultrex® BME Cell Invasion Assay kit protocol. This assay was applied on MDA-MB-468, MDA-MB-231, MDA-MB-453, and SkBr3 *PYK2* depleted gene cell lines. Cells were starved in Serum-Free Medium for 24h. Meanwhile, the top chamber of the 96 well plate cell invasion device was coated with 50 μL of 1X BME Solution (already diluted) and incubated at 37°C for 4 hours. The following day, cells were harvested and diluted to 1×10^6 cells/mL in serum-free medium (with or without Metformin). 50 μL of cell suspension was added for each well at the top chamber and 150 μL of the medium was added to each well in the bottom chamber. The plate was incubated at 37°C (with or without CO_2 according to medium requirements) for 48 hours. The following day, wells in the top chamber were washed with 100 μL of 1X washing buffer and wells in the bottom chamber were washed twice with 200 μL of 1X washing buffer. Then, 100 μL of diluted Calcein AM was added to the bottom chamber wells and incubated at 37°C for 1 hour. Finally, the top chamber was removed, and fluorescence intensity was measured using a fluorescence microplate reader (TECAN ULTRA fluorescence spectrophotometer) at $\sim 485 \text{ nm}$ excitation and $\sim 520 \text{ nm}$ emission. Results were transported directly to Excel and analysed considering the subtraction of the average absorbance values from the control wells to remove background signals. The percentage of invaded cells were calculated regarding the control condition.

2.2.5.9. xCELLigence System (RTCA) for Cell Proliferation

Real-time monitoring of cell -viability was carried out *via* the xCELLigence system, which has been designed with special microtiter plates that contain interdigitated gold microelectrodes. This assay was also applied to the same cell line that mentioned in section (2.2.5.7. Scratch assay) using the 16-well plates (E-Plate 16), which are integrated with microelectronic cell sensor arrays. Firstly, background measurement was done prior to cell culture with free cell-media as one sweep for 1 minute (step 1). Then, cells were harvested and diluted to 1×10^6 cell/mL. 100 μ l of cell suspension was added to each well and filled with 100 μ l of cell-dedicated media. E-Plate 16 was incubated for 30 min at RT under sterile condition. Then, the E-Plate 16 was placed into the cradle of the RTCA Station in the incubator and step 2 was started (proliferation curve overnight). The following day, cells were treated with cell-dedicated media supplemented with Metformin, and the third step was started. Cell Index (CI) was measured every 20 min for 2h and then every hour for 48h. Finally, the obtained results were analysed *via* RTCA software.

2.2.5.10. Mass spectrometry analysis

2.2.5.10.1. Whole cell lysate preparation

HER2 positive (SkBr3 and MDA-MB-453) cell lines were cultured in T75 flasks until 90% confluency. Cells were treated with serum-free medium (with or without Metformin) for 48h prior to harvesting. Media was removed, and cells were washed three times with cold DPBS. The remaining DPBS was carefully removed, and 200 μ L of Erika's lysis buffer, which contains 9.5 M Urea, Dithiothreitol (DTT), N-Octyl-Beta-Glucopyranoside (OGP), (details in section 2.1.2. Buffers and gels) supplemented with 1% PI was added directly. Lysis buffer was spread across the entire cells and incubated for 2 minutes. Cells lysate was collected into Eppendorf tubes and the tubes were transferred to an ice water bath for sonication at max power for 5 min. Then, samples were stored on ice for another 5 min. This step was repeated twice, then samples were centrifugated for 10min at 12,000g and 4°C. The supernatant (cell lysate) was removed carefully and stored in fresh tubes at -80°C.

The lysis buffer choice is decisive, to ensure a high yield of proteins for mass spectrometry (MS) analyses. The buffers usually used contain strong detergents such as SDS or Tween-20 as lysing agents, and that is inappropriate for mass spectrometry sample preparation. For this, cells were mechanically dissociated, lysed and denatured in Erika's buffer supplemented with 1% Protease

Inhibitors (PI). In this buffer, the high concentration of urea induces protein denaturation, DTT breaks down the secondary structure of the protein and increases the solubility of proteins.

2.2.5.10.2. Cell lysate protein quantification and MS analysis

Prior to MS analysis, the protein assay was applied to all samples to quantify protein concentrations. Bio-Rad Protein Assay, which is compatible with Erika buffer, was used as described in the supplier protocol. In brief, the dye was diluted to 1:5 in ddH₂O and samples were diluted to 1:10 in TEAB. 10 µL of standards and samples were pipetted to each well (in triplicates), then 200 µL of diluted dye was added per well. The well was mixed and incubated at room temperature in the dark for 5 min. Absorption was measured at 595nm using TECAN ULTRA spectrophotometer. Protein concentrations were then calculated according to the standard curve. 50 µg of each sample were aliquoted to fresh tubes and be subjected to MS analysis. SCIEX Triple TOF® 6600 mass spectrometers linked to an Eksigent nanoLC 425 HPLC system, was employed to analyse the protein lysate samples. The LC system was operating in microflow (5 µL/min) and 3 µl of each sample was directly injected on a YMC 25 cm × 0.3 mm Triart-C18 column (12 nm, 3 µm particle size). The MS analysis was carried out *via* two acquisition methods; Information Dependent Acquisition (IDA) for spectral library generation and by Sequential Window Activation of All Theoretical Mass Spectra (SWATH-MS) data acquisition. The tandem mass spectrometry spectra were searched using ProteinPilot 5.0 (SCIEX) with a SwissProt database containing human species at 1 % False Discovery Rate (FDR) cut off.

CHAPTER 3

Investigating the effect of Metformin on different breast cancer molecular subtypes:

3.1. Introduction

Metformin is an oral antidiabetic drug with well-established side effect and safety profiles. This drug helps diabetics to respond to insulin by lowering and maintaining blood sugar levels. It controls blood sugar in several ways such as helping type 2 diabetics respond better to their own insulin by lowering the amount of sugar created by the liver and decreasing the amount of sugar absorbed by the intestines. Metformin is unlikely to cause hypoglycemia or weight gain if taken as a monotherapy. Currently, several research studies indicated that Metformin might be beneficial in reducing cancer incidence of a variety of cancers including breast cancer. The anti-proliferative activity of Metformin, and thus its possible use as an adjuvant in traditional cancer therapies has been highlighted in various types of cancer including breast cancer (*Wurth et al., 2014; Lin et al., 2014*). All the above-mentioned features nominated Metformin as an ideal candidate for breast cancer treatment.

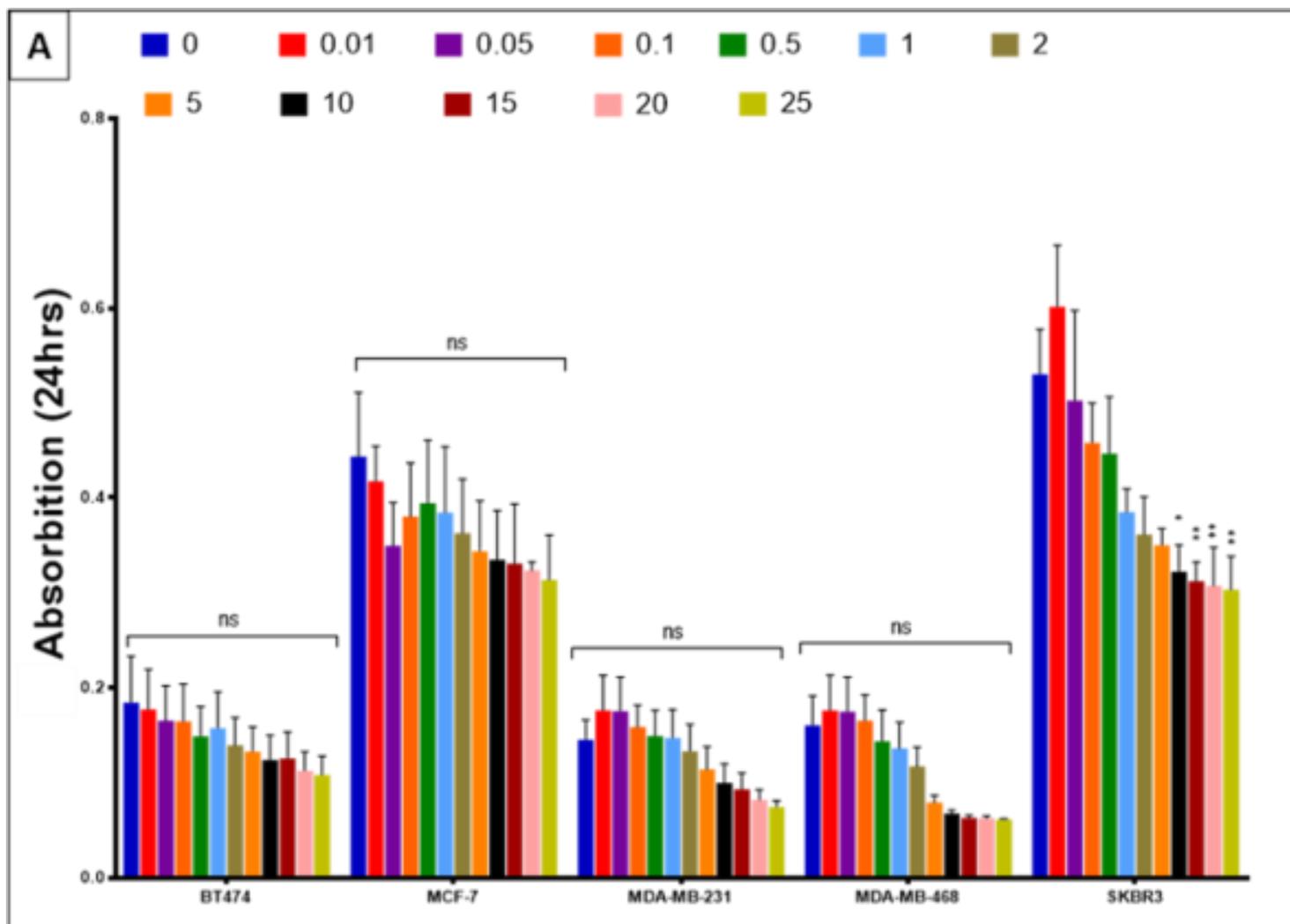
Breast cancer is a complicated multifaceted heterogeneous disease, which shows a wide spectrum of structures that display diversity in clinical, morphological and molecular features. Breast cancer can be classified under different categories. Traditional breast cancer molecular classification began with ER, PR and HER2 based stratification of patients, and that is currently applied in clinical practice (*Prat and Perou, 2011; Vuong et al., 2014*). These different subtypes have distinct clinical outcomes and responses to endocrine therapy and chemotherapy (*Yersal et al., 2014*).

The focus of this chapter was to study the cytostatic effect of Metformin on different phenotypes of breast cancer: Luminal B, Luminal A, Claudin-low, Basal-like, and HER2. Consistently with breast cancer phenotypes diversity, five different human breast cancer cell lines (BT-474, MCF-7, MDA-MB-231, MDA-MB-468, and SkBr3) representing the five molecular breast cancer subtypes (Luminal B, Luminal A, Claudin-low, Basal-like, and HER2) were used to investigate the effect of Metformin on breast carcinoma. Consequently, eleven doses of Metformin (0.01, 0.05, 0.1, 0.5, 1, 2, 5, 10, 15, 20, 25mM) were tested in vitro to determine the effect of Metformin on various cellular processes including cell morphology, survival, proliferation and apoptosis.

3.2. Investigating the effect of Metformin on cell survival (viability) using (MTT assay)

The purpose of the MTT (3- [4, 5-dimethylthiazol-2-yl]-2, 5 diphenyl tetrazolium bromide) assay is to measure cell viability in a relatively high throughput model (e.g. 96-well plates) without the need for elaborate cell counting. This commonly used assay is employed to determine the cytostatic effect of a drug at different concentrations. The MTT assay measures cellular metabolic activity by measuring the activity of reduced nicotinamide adenine dinucleotide (phosphate) or NAD (P) H-dependent oxidoreductases. This activity occurs in mitochondria but may also happen in the presence of non-mitochondrial pyridine nucleotide-dependent enzymes. The amount of MTT formazan determined by measuring absorbance is directly proportional to the number of cells (*Aleshin et al., 2015; Ito et al., 2015*).

The MTT cell viability assay has been used to determine the effect of increasing concentration of Metformin (0.01, 0.05, 0.1, 0.5, 1, 2, 5, 10, 15, 20, 25mM) on the five different breast cancer subtypes, for 24h and 48h. This experiment was independently repeated three times for each cell line. The results showed that Metformin significantly reduced the cellular metabolic activity of MDA-MB-468 and MDA-MB-231 triple negative breast cancer cell lines (Basal-like and Claudin-low, respectively), and at the different Metformin concentrations, followed by BT-474 breast cancer cells (Luminal B). The lethal concentration was $\geq 10\text{mM}$ after 24h and $\geq 5\text{mM}$ after 48h. The MCF-7 (Luminal A) and the SkBr3 (HER2) cell lines were less sensitive to Metformin treatment after 24h and 48h of incubation. The lethal concentrations were 15 to 25mM for both cell lines for the 24h and 48h incubation periods. Graph Pad Prism 7 software was utilised to reflect the responses of the five cell lines to Metformin and at the different concentrations (Figure 3. 1, A and B). There were statistically significant differences in the viability of the cells ($P \leq 0.001 - 0.0001$ ***-****) between MCF-7 and SkBr3 and when compared to the other cell lines. In addition, there was a significant statistical difference between the two-time points (24h and 48h) of treatment ($P = 0.0001$ ****) for the entire assay and when using TIBCO Statistica 13.3 software (Factorial ANOVA, Metformin concentrations versus cell lines and time).



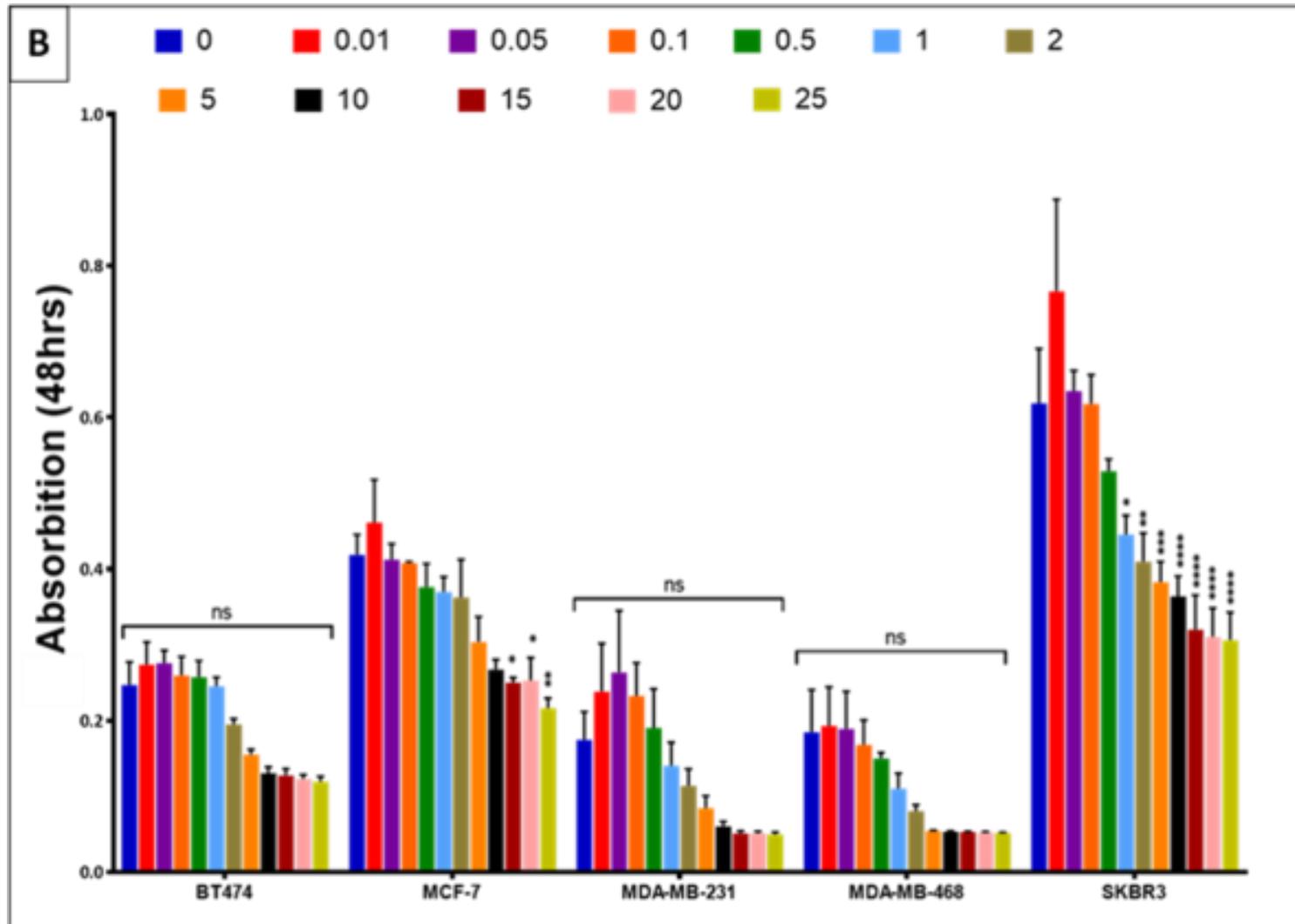
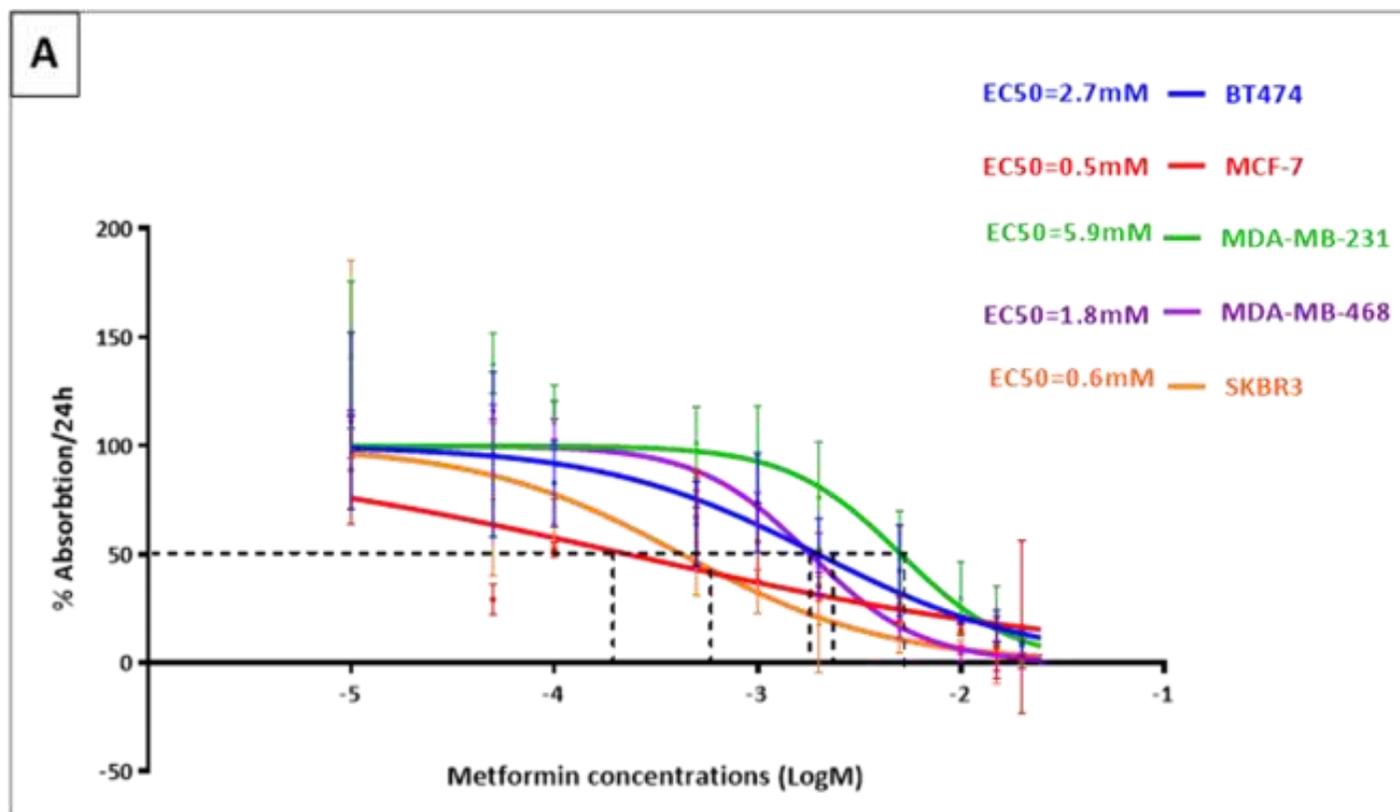


Figure 3.1. Graph representing cell viability status of different breast cancer cell lines to Metformin at different concentrations, determined using the MTT assay for 24h (A) and 48h (B) of treatment. The 2way ANOVA (multiple comparisons) has been used in this analysis. There was a significant difference $P \leq 0.05$ - 0.0001 (*-****) between MCF-7 and SkBr3 cell lines in comparison with the other cell lines within each group of treatment (0-25 mM) individually after 24 and 48 H of exposure to Metformin. As well as, there was an extremely significant difference ($P=0.0001$ ****) between the two-time points (24h and 48h) of treatment with Metformin.

3.3. The computing of EC50 doses of Metformin for all breast cancer subtypes

Further analysis was conducted to calculate the EC50 that corresponds to the lethal concentration of drug required to kill 50% of cell populations and based on the MTT assay results. The non-linear regression curve fit and log of inhibitor (Metformin) versus normalised response-variable slopes (absorptions) based on graphs obtained by Graph Pad Prism 7 software, was used to determine the dose inhibition response of Metformin for the breast cancer cell lines after 24h and 48h of treatments respectively (Figures 3.2, A and B). These doses varied between -4.7 and -2.0 Log M of Metformin concentration after 24h of treatments and -3.2 and -2.1 Log M after 48h of Metformin treatments and according to the different breast cancer subtypes.



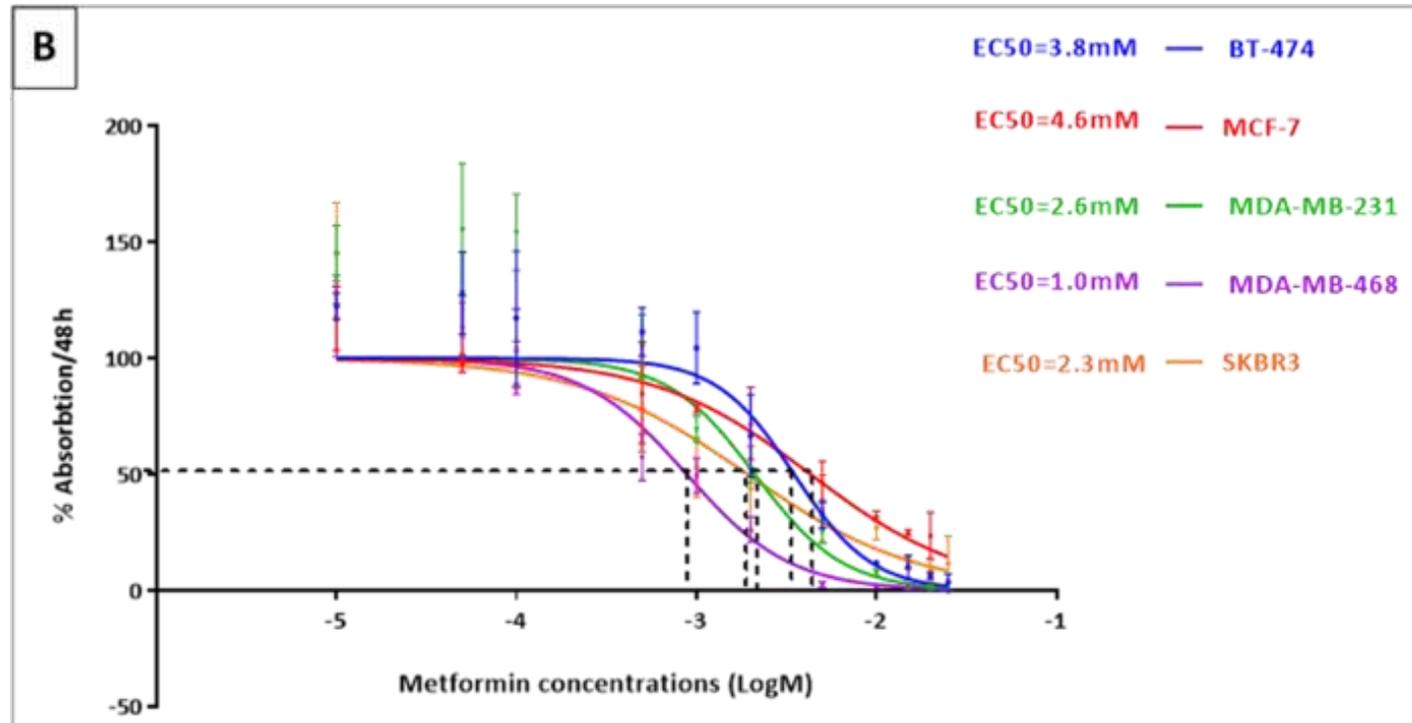


Figure 3.2. The EC50 of the five breast cancer phenotypes as determined using the MTT assay after 24h (A) and 48h(B) of treatments with Metformin. The EC50 mean varied between (-4.7 to -2.0 Log M) after 24H and (- 3.2 to -2.1 Log M) after 48H of Metformin concentrations according to different breast cancer subtypes. Representative graph from three independent experiments.

To compute the original concentrations of Metformin that reflect EC50 values (1mM), the antilog of the EC50 mean values have been calculated according to the following equation: $10^{\text{Log EC50}}$ after the calculation of the EC50 for each cell line and for each individual time point. This was to compare between the time points and for further selection to the one-time point of treatment with Metformin to carry out the next microarray assay. The results showed a significant difference (P value <0.0001 ****) between the 24h and 48h period of exposure to Metformin for the Claudin-low (MDA-MB-231) cell line. The EC50 value was dropped from 5.9mM after 24h to 2.6mM after 48h of Metformin treatment respectively. However, there was a significant difference (P < 0.0001- P ≤ 0.05****-*) between 24h and 48h of exposure time to Metformin in Luminal A (MCF-7), Luminal B (BT-474) and HER2(SkBr3) cell lines in which the concentrations were increased after 48h of treatments, that correlated to an increase in cell proliferation after 48h. On the other hand, Basal-like (MDA-MB-468) cell line showed an insignificant difference between 24h and 48h of treatment with Metformin (1.8mM/24h) and (0.9mM/48h) respectively, representing the most sensitive phenotype to Metformin treatments after 48h of exposure time (Figure 3. 3).

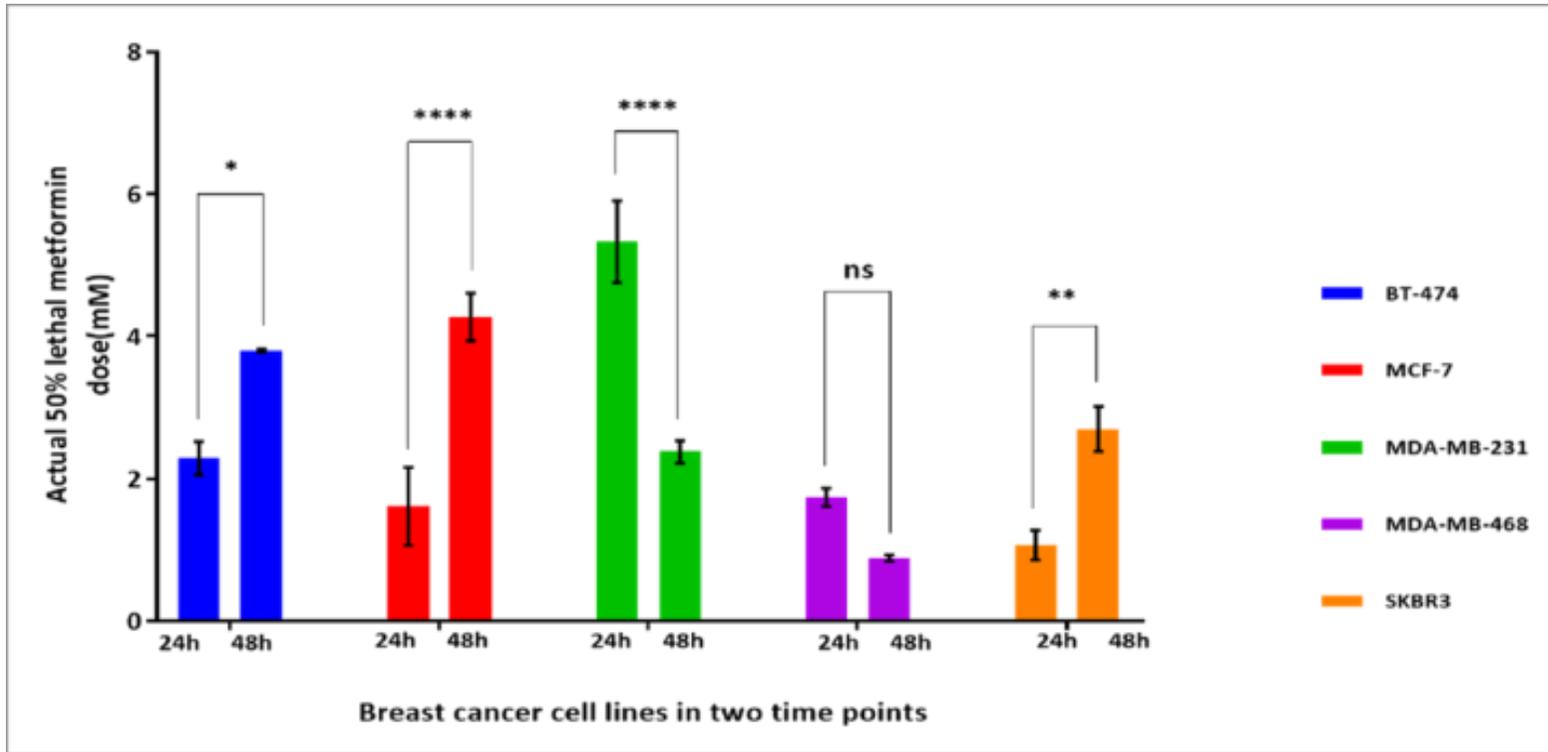


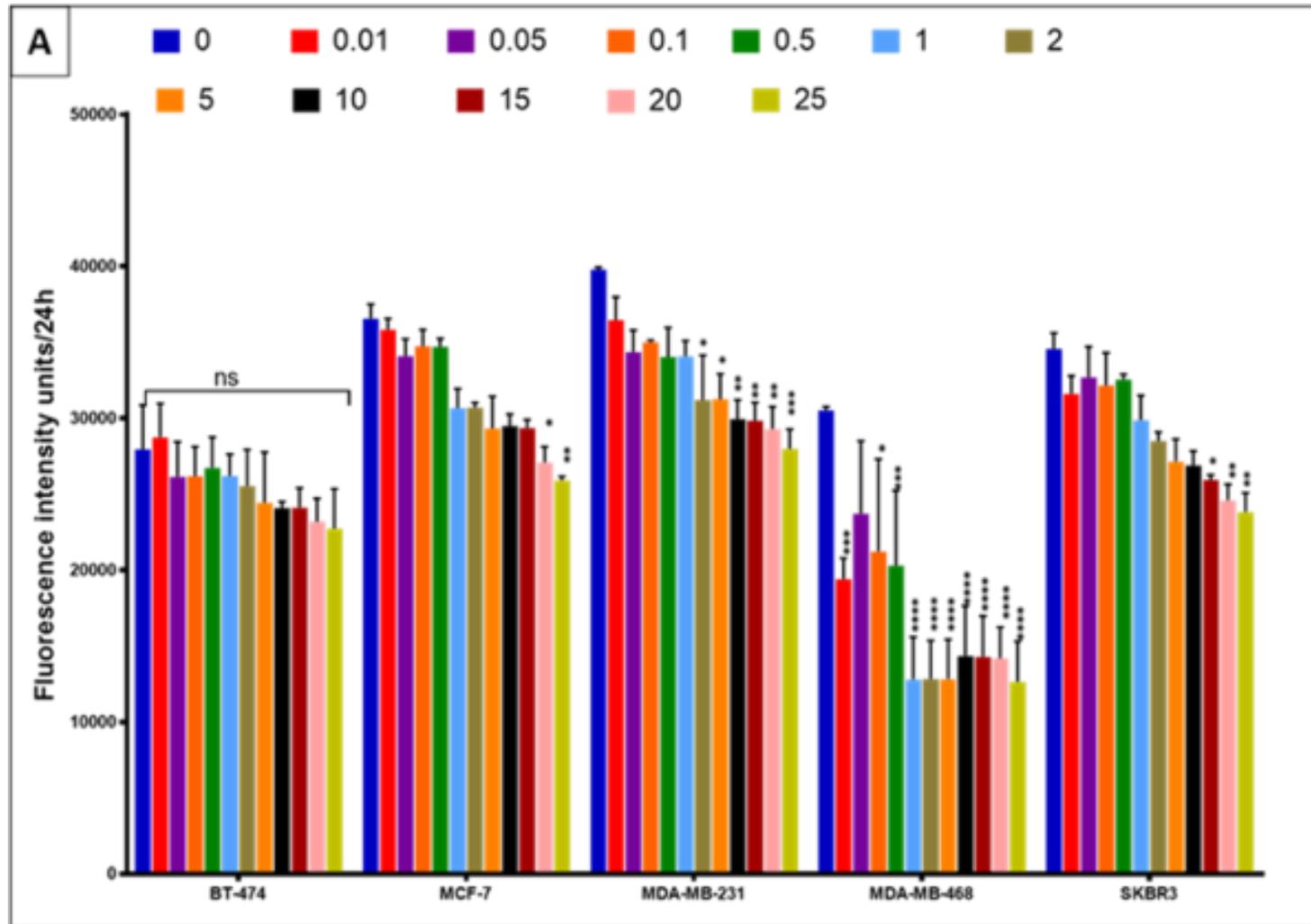
Figure 3. 3. The actual EC50 of Metformin concentrations (mean) for the five breast cancer cell lines showing the significant values between the two-time points (24 H, 48 H) of treatment with Metformin that was obtained from Cell Viability Assay. The Metformin concentrations varied between (0.5 -5.9 mM) and according to different breast cancer subtypes and time points. Luminal B (BT-474), Luminal A (MCF-7), Claudin-low (MDA-MB-231), Basal-like (MDA-MB-468), and HER2 (SkBr3).

3.4. Investigating the effect of Metformin on cell proliferation of the different breast cancer phenotypes

Cell proliferation is a process of cell number increase through cell divisions and is at the interface between cell loss (through cell death) and differentiation. The increase in cell proliferation can be observed in tumour cells at a higher rate than normal cells. A cell proliferation assay, which has been described in the methods chapter (section 2.2.2.3. Cell proliferation assay) was based on thymidine incorporation measures the number of cells synthesizing DNA during cell divisions. This was used to measure the effect of different concentrations of Metformin on the proliferation of breast cancer cells. The cell lines and the concentrations of applied Metformin are identical to the ones that were used for the cell survival assay.

This experiment was independently repeated three times, which demonstrated that the level of cell proliferation in response to different concentrations of Metformin was extremely significant (P value = 0.0001****) for the entire assay and between the two time points of treatment (24h and 48h) by using TIBCO Statistica 13.3 software. The statistical method that has been used was Factorial ANOVA (Metformin concentrations versus cell lines and time).

However, utilising Graph Pad Prism 7 software and Two-way ANOVA (multiple comparisons) analysis showed that the Basal-like (MDA-MB-468) cell line was the less proliferated cell line after 24h of exposure to Metformin in all concentrations (P=*-****). While, after 48h of treatment with Metformin Claudin-low (MDA-MB-231) cell line was the least proliferative phenotype in the highest concentrations (15-25) mM. Clearly, the HER2 (SkBr3) cell line was the highest proliferated cell line in the concentrations of (1-25) mM after 48h of Metformin treatment. In addition, the three other cell lines representing the phenotypes Luminal A (MCF-7), Luminal B (BT-474), and Basal-like (MDA-MB-468) had a median proliferative rate after 48h of treatment with Metformin. The responses of the five cell lines to Metformin and at the different concentrations for the two-time points are shown in Figure 3.5 A and B respectively.



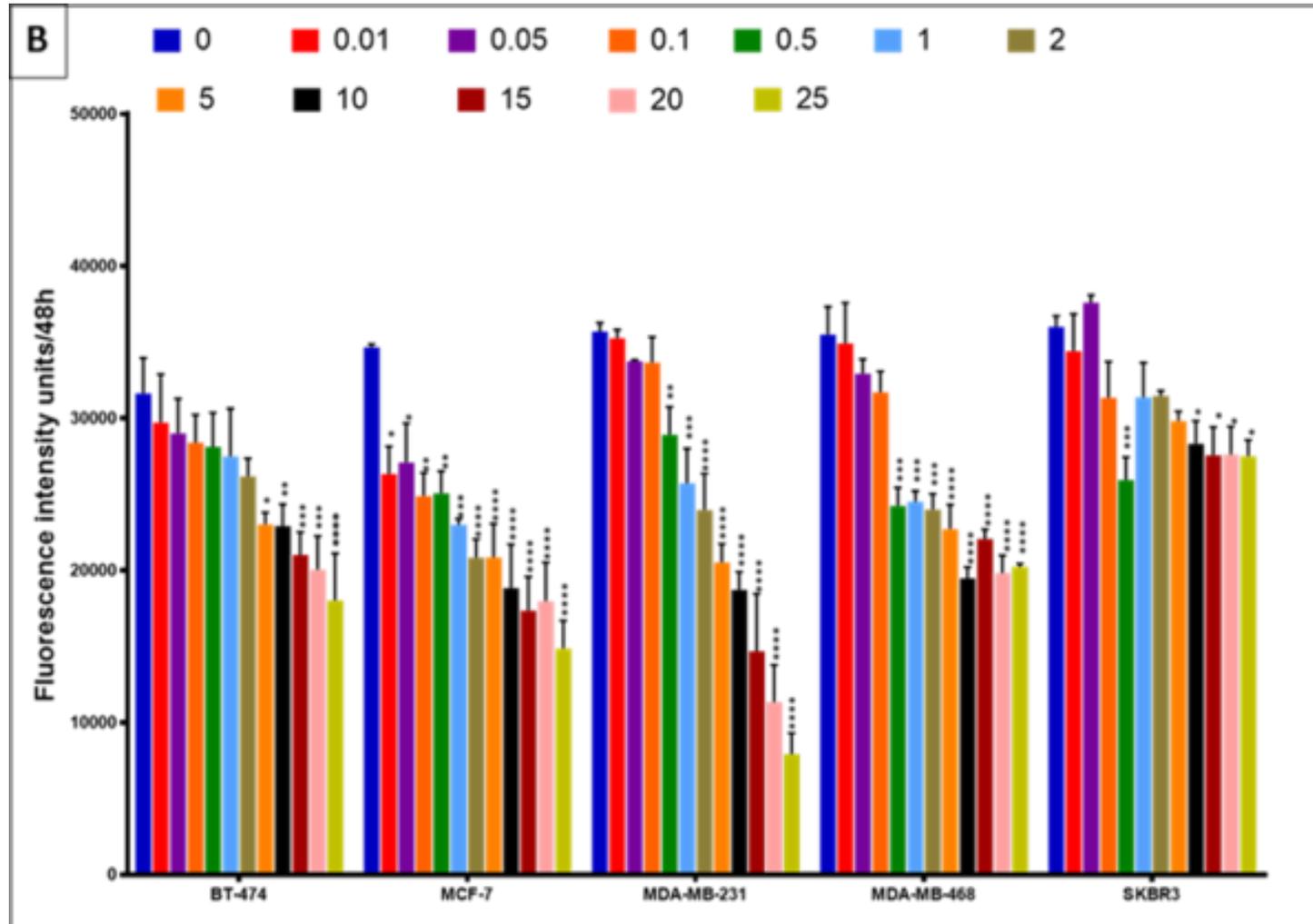


Figure 3.4. The proliferation status of different breast cancer phenotypes after treatment with different concentrations of Metformin, as determined using a cell proliferation assay. Non-significant (ns) differences between the different phenotypes in proliferation ratio were noticed after 24H of treatment. However, the MDA-MB-468 cell line was the less proliferative cell line. Further, after 48H of treatment with Metformin, significant differences have been noticed between each group of cells upon different concentration. Two-way ANOVA ($P=*-****$).

3.5. Investigating the effect of Metformin on apoptosis of breast cancer subtypes

Apoptosis, or programmed cell death, is cell death that is generally triggered by normal, healthy processes in the body. Necrosis is the premature death of cells and living tissue, caused by external factors, such as infection, toxins, or trauma.

Metformin resistance analyses in different breast cancer cell lines were determined using Flow cytometry-based apoptosis assay, which involved the co-staining of cells with Alexa Fluor 647 Annexin V (apoptotic cells) and propidium iodide (PI) for (non-viable cells).

This study demonstrated that Metformin treatment significantly increased apoptosis of Basal-like (MDA-MB-468), Claudin-low (MDA-MB-231) and Luminal B (BT-474) cell lines. The Luminal A (MCF-7) and HER2 (SkBr3) cell lines were less apoptotic at high concentrations of Metformin. PI staining demonstrated that Metformin treatments induced necrosis in all cell lines at higher concentrations. Flow cytometry (FACS) and Kaluza 3.1 software were used in this study to derive the data. Figures 3.5, 3.6, and 3.7 show the different percentages of four gates in scatter diagram which are Ann V- PI- (living cells), Ann V+ PI- (apoptosis cells), Ann V+ PI+ (apoptosis and dead cells), and Ann V- PI+ (dead cells) for different breast cancer phenotypes. The results showed that (MDA-MB-231, MDA-MB-468) cell lines were more affected by the Metformin after 48h of treatment. The proportion of apoptotic cells were 22.22% and 19.97% respectively and when compared with the control cells which had 0.09% and 0.05% respectively. The BT-474 cell line was the second most affected cell line, while MCF-7 was the third most affected cell line after 48h of treatment. SkBr3 was the least apoptotic at 48h post-treatment. The percentage of apoptotic cells was 10.33 % after 48h of treatment and when compared to the percentages of non-treated cells which was 0.0%.

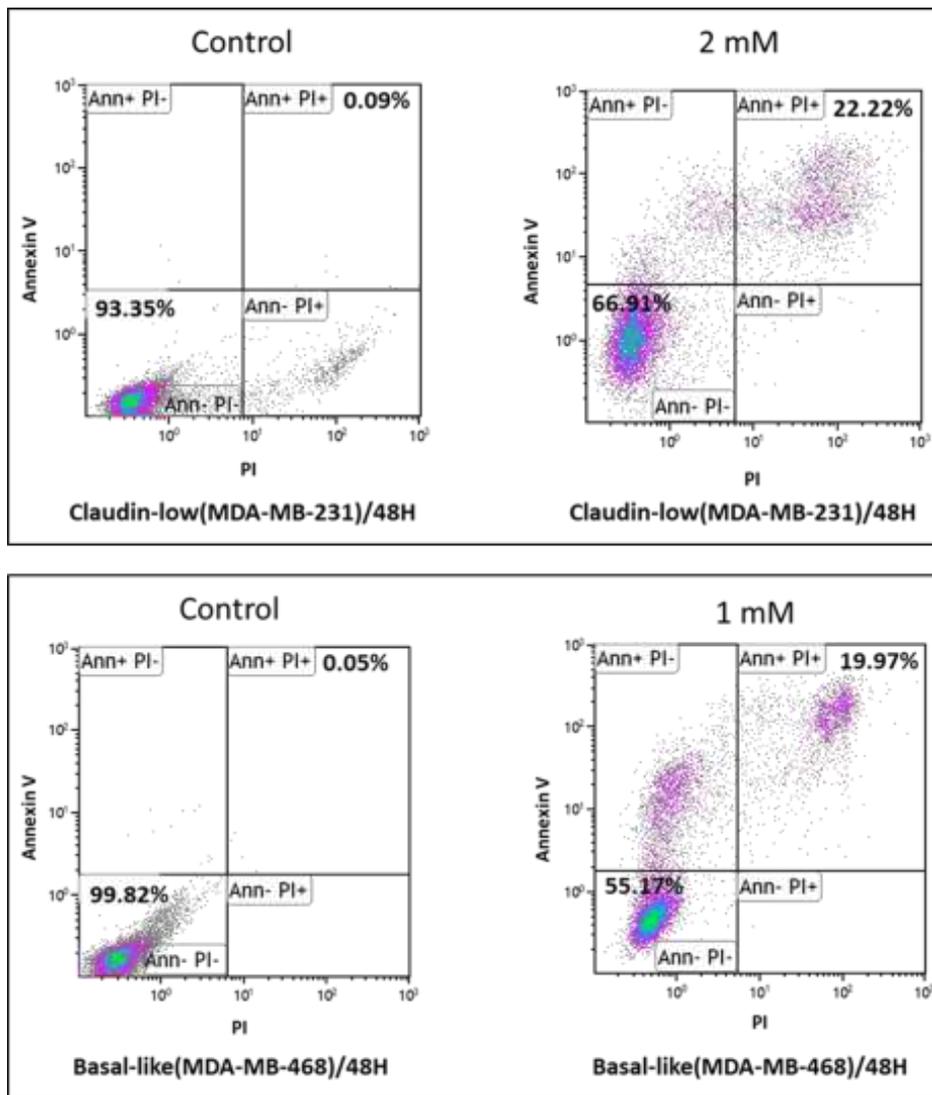


Figure 3.5. Flow cytometry analysis of Metformin resistance – gating strategy for cell viability staining. Demonstrating the effect of 2 and 1 mM Metformin (EC50) on Claudin-low (MDA-MB-231) and Basal-like (MDA-MB-468) cell lines following 48 hours of incubation correspondingly. Dot plot represents gating of 'Cells' population, the left down square of the four gates in the scatter diagram represents the Ann V- PI- (living cells); left up square, Ann V+ PI- (apoptosis cells); right up square, Ann V+ PI+ (apoptosis and dead cells); and right down square, Ann V- PI+ (dead cells). Representative graph out of the average of three independent experiments.

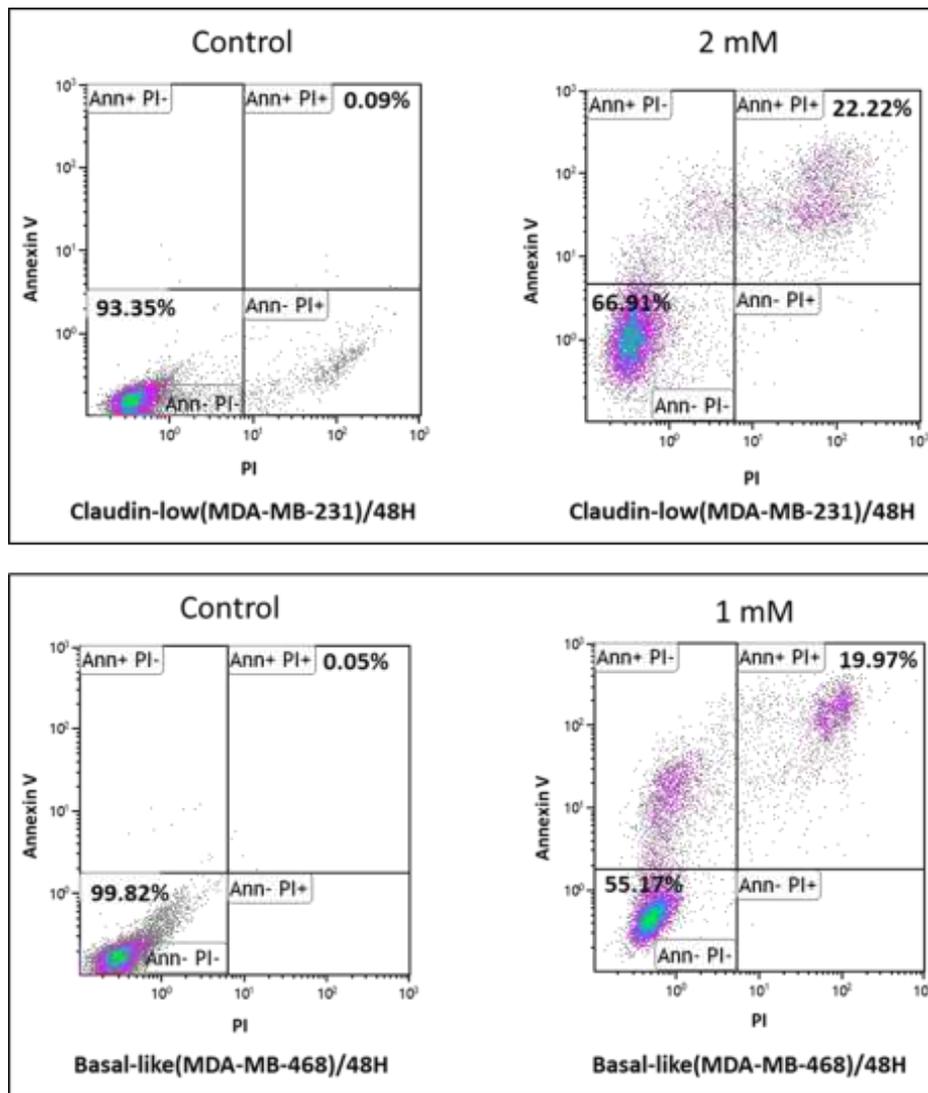


Figure 3.6. Flow cytometry analysis of Metformin resistance – gating strategy for cell viability staining. Showing the effect of 5mM Metformin (EC50) on Luminal B (BT-474) and Luminal A (MCF-7) cell lines respectively following 48 h of incubation. Dot plot represents gating of ‘Cells’ population, the left down square of the four gates in the scatter diagram represents the Ann V- PI- (living cells); left up square, Ann V+ PI- (apoptosis cells); right up square, Ann V+ PI+ (apoptosis and dead cells); and right down square, Ann V- PI+ (dead cells). Representative graph out of the average of three independent experiments.

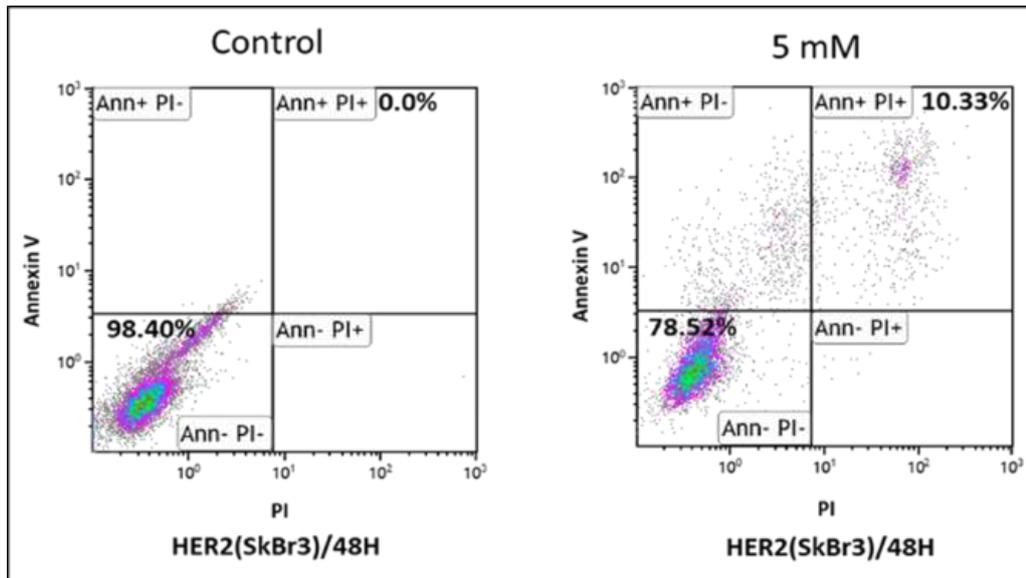


Figure 3.7. Flow cytometry analysis of Metformin resistance – gating strategy for cell viability staining. Showing the effect of 5mM Metformin (EC50) on HER2 + cell line (SkBr3) following 48 h of incubation. Dot plot represents gating of 'Cells' population, the left down square of the four gates in the scatter diagram represents the Ann V- PI- (living cells); left up square, Ann V+ PI- (apoptosis cells); right up square, Ann V+ PI+ (apoptosis and dead cells); and right down square, Ann V- PI+ (dead cells). Representative graph out of the average of three independent experiments.

Additionally, the percentage of gated cells were calculated as a combined average of early apoptosis cells (Ann V⁺ PI⁻) and necrosis cells (Ann V⁺ PI⁺), to compare between all breast cancer subtypes at different concentrations for 24h and 48 hours of treatment, respectively. This analysis has been completed using Graph Pad Prism 7 software (Grouped Analyses, Two-way ANOVA). In addition, the Bonferroni test has been used for multiple comparison correction. There was a significant reduction in the percentage of viable cells following increased Metformin concentrations in Basal-like (MDA-MB-468) and Claudin-low (MDA-MB-231) cell lines after 24h and 48h of treatment with Metformin respectively, (Figures 3.8; 3.9). The Basal-like (MDA-MB-468) cell line was significantly affected by the Metformin treatments starting from the concentration of 2mM. A lower proportion of dead cells has been observed in Luminal A (MCF-7), Luminal B (BT-474) and HER2 (SkBr3) cell lines after 24h and 48h of incubation with Metformin and using the same concentrations as above. Although, non-significant values were noticed between the other three breast cancer cell lines (Luminal B, Luminal A, and HER2) at different concentrations if compared to the control (untreated) cells after 24h and 48h of Metformin treatment. Besides, HER2 (SkBr3) cell line was the less apoptotic cell line after 48h of exposure to Metformin as shown in Figure (3.9) below.

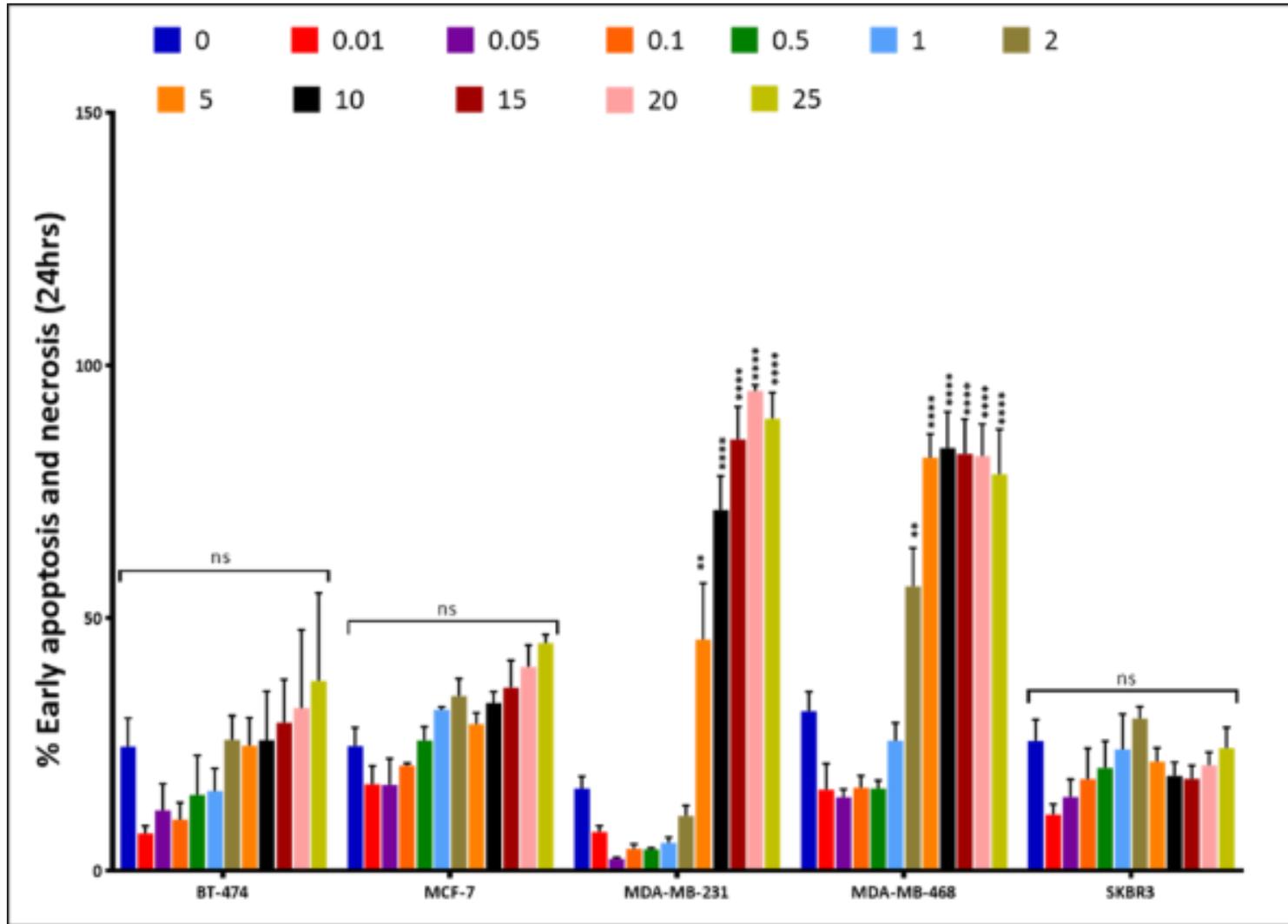


Figure 3.8. Representative graph of the total percentage of combined early apoptotic and necrotic cells after 24h of treatment with Metformin. There were significant differences in Basal-like (MDA-MB-468) and Claudin-low (MDA-MB-231) cell lines different Metformin treatments in contrast with control and starting from the concentration of 2mM and 5mM respectively, ($P \leq 0.01$ **). However, non-significant changes ($P=ns$) have been observed between in Luminal A (MCF-7), Luminal B (BT-474) and HER2 (SkBr3) cell lines at the different concentrations.

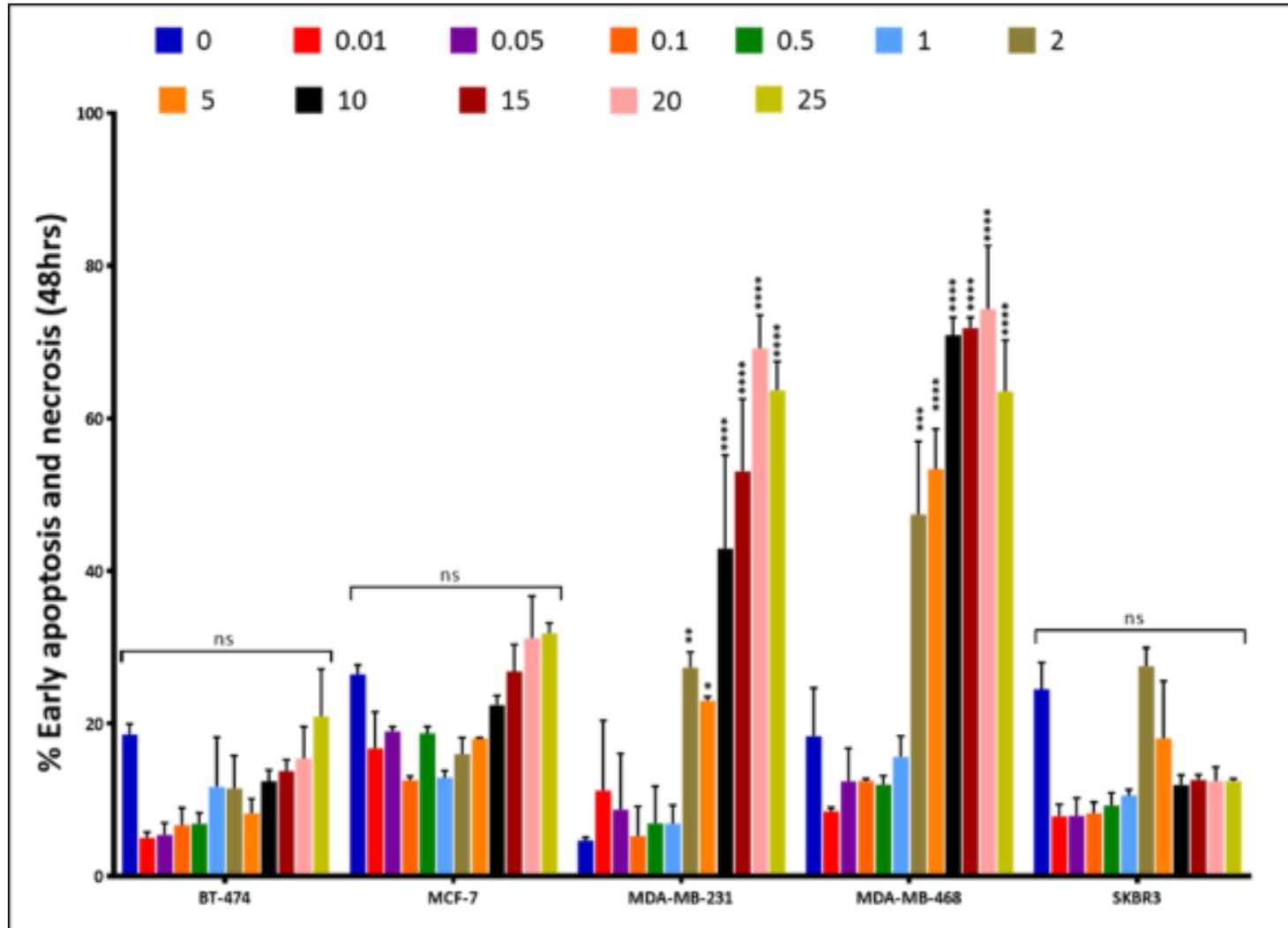


Figure 3.9. Representative graph of the total percentage of combined early apoptosis and necrosis cells after 48h of Metformin. An extreme significant dissimilarity has been noticed when compared between the Basal-like (MDA-MB-468) and the other three breast cancer cell lines (Luminal B, Luminal A, and HER2) starting at concentrations of 2 mM and up to 25 mM ($P \leq 0.0001$ ****). While, when comparing the Claudin-low (MDA-MB-231) cell line, significant differences were observed ($P \leq 0.05$ *) at concentrations of 2mM and up to 25mM. However, non-significant changes ($P=ns$) have been observed between in Luminal A (MCF-7), Luminal B (BT-474) and HER2 (SkBr3) cell lines at the different concentrations.

3.6. Investigating the effect of Metformin on the morphology of breast cancer subtypes

The effect of Metformin on cell morphologies of breast cancer cell lines was observed at different concentrations. The cells were treated with the same previously stated concentrations of Metformin and their morphologies were evaluated using light microscopy. All observed cells became rounded and detached. In addition, the number of cells decreased significantly at high concentrations of Metformin treatment and when compared with untreated controls. Figure 3.11 shows the morphological effect of Metformin (at the concentration of 25mM) particularly on the breast cancer cell lines after 48 hours of treatment.

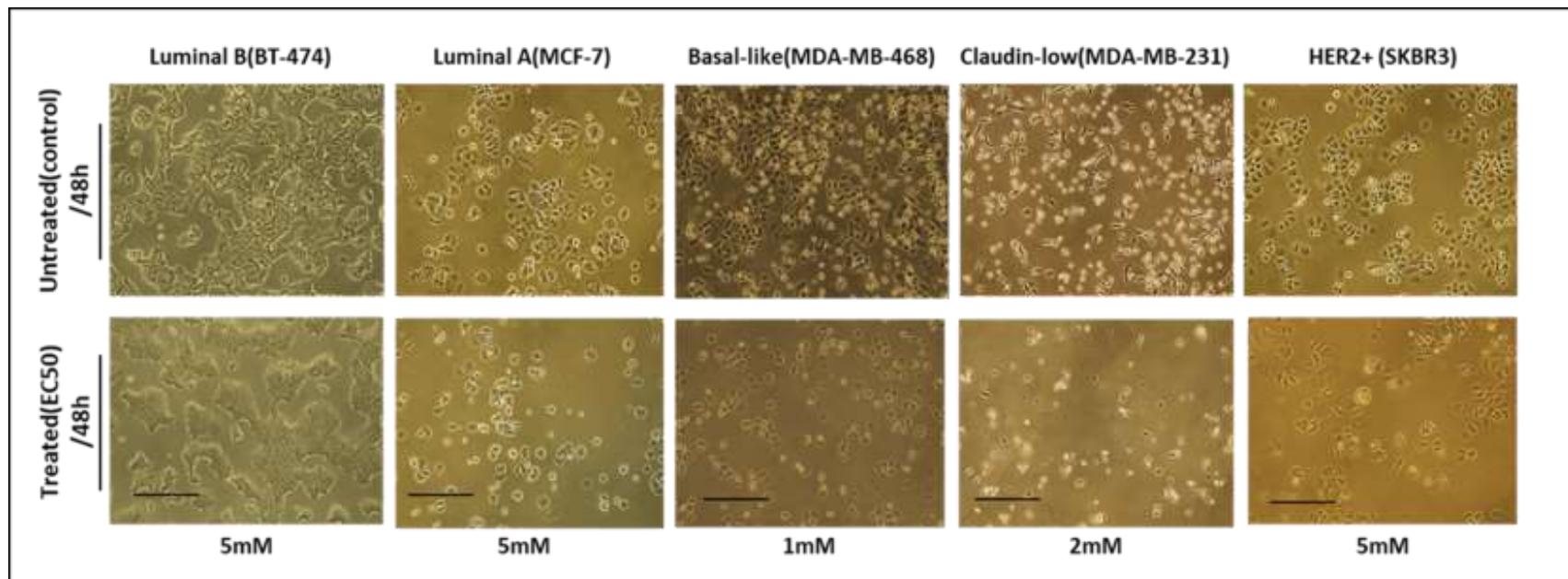


Figure 3.10. Representative micrographs showing the morphological influence of Metformin on the different breast cancer cell lines: Luminal B (BT-474), Luminal A (MCF-7), Claudin-low (MDA-MB-231), Basal-like (MDA-MB-468), and HER2 (SkBr3). The upper panel represented untreated cells, while the bottom panel refers to treated cells with EC50 of Metformin concentration for each cell line after 48h of treatment. Representative images from three experiments were taken at 10X magnification with scale bars indicating 10µm.

From the all above results, it can be concluded that the most sensitive phenotype to Metformin is the Basal-like (MDA-MB-468) cell line while the less sensitive subtype is the HER2 (SkBr3) cell line. It can also be suggested that 1mM and 4mM (1/2 EC50 and 2EC50) are optimal doses for future work, respectively and following 48h of treatment. The 1/2 EC50 (1mM) concentration was selected for the Gene Expression Microarray assay, in order to keep as many as possible living cells to extract sufficient RNA. However, the 2EC50 (4mM) dose was used to compare the effect of high concentration of Metformin treatment.

3.7. Discussion

The work presented in this chapter was aimed at investigating the cytostatic effect of Metformin on different breast cancer cell lines, that represent the breast cancer cell phenotypes: Luminal B, Luminal A, Claudin-low, Basal-like and HER2. This effect was assessed by determining cell viability, cell proliferation, and apoptosis of the cell lines. In addition, the morphological effect of Metformin on the cells was also considered. These findings enabled the selection of doses and cell lines to be determined for the following experimental work.

Interestingly, Metformin treatment significantly reduced cell survival and proliferation, whilst inducing cell apoptosis and enhance cell necrosis. Besides, the Basal-like cell line (MDA-MB-468) has been noticed as much affected cell line. The lethal dose (EC) was $\leq 2\text{mM}$ and the EC50 concentration was 1mM after 48h of treatment, respectively. In addition, Metformin affected the appearance of the cells, which became a rounded shape and floated. Comparatively, the HER2 subtype (SkBr3 cell line) was the least affected phenotype and the EC50 concentration was much higher 2.3mM after 48h of treatment.

These observations were correlated with previous findings (*Liu et al., 2009*) that showed that Metformin inhibits *in vitro*, cell proliferation and survival of triple negative breast cancer cells. Another study (*Liu et al., 2012*) demonstrated that using Metformin alone or in combination with chemotherapeutic agents inhibited cell proliferation in triple negative breast cancer cell lines. Moreover, Metformin has previously been shown to inhibit the growth of breast cancer cell lines (MCF-7, MDA-MB- 231 and MDA-MB-435) *in vitro*, and *via* AMPK induction and mTOR inhibition (*Zakikhani et al., 2006; Dowling et al., 2007; Phoenix et al., 2009; Hadad et al., 2009; Hadad et al., 2011*).

In addition, Metformin potential to reduce the proliferation rate of a subset of cancers can only be achieved if this cancer has particular molecular characteristics and if patient-specific pharmacokinetic variables lead to an adequately high level of drug exposure. For instance, it has been reported that Metformin reduced mortality for specific-HER2+-breast-cancer users versus non-users (*Klil-Drori et al., 2017*). However, the outcomes of this study raised a substantial concern regarding Metformin exposure time was imbalance over the entire observation time (type II diabetes mellitus (T2DM) long-time users and new-users) and referred to as biased (*Klil-Drori et al., 2017*). Similarly, *Sonnenblick* and colleagues in 2017 stated in their study that Metformin may improve the worse prognosis, that is cancer associated with diabetes and insulin treatment, at most in patients with primary HER2-positive and hormone receptor-positive breast cancer (*Sonnenblick et al., 2017*).

In summary, the results obtained in this study guided the project to select two breast cancer subtypes for further studies; the MDA-MB-468 Basal-like and the SkBr3 HER2 cell lines. It also allowed the selection of two different concentrations of Metformin: $\frac{1}{2}$ EC50 and 2EC50 (1 and 4mM, respectively) which will be used in the experiment detailed in chapter 4. Finally, the aim of this part was to interrogate the Metformin effect on different subtypes of breast cancer and the next chapter will present a step-wise process of identification of involved genes and cellular pathways.

CHAPTER 4

Identification of novel markers and cellular pathways associated with BASAL-LIKE and HER2 breast cancer subtypes

4.1. Introduction

One of the most important applications of DNA microarrays is the analysis of cancer cells transcripts, which can make a comparison of gene expression with normal cells. Using powerful tools including transcript profiling and clustering to classify a tumour into subtypes might lead to a preferable diagnosis and remedy of cancer (*Dadkhah et al., 2015*). Moreover, it offers the capability to design and develop personalised therapeutic treatments by improving understanding of the molecular basis of cancer and supporting the clinical decision (*Mustapha et al., 2017*). Furthermore, great efforts have been made to characterise and identify new molecules, which could improve the prediction of metastasis risk and increase the probability of therapy response (*Yersal and Barutca, 2014*).

4.2. The applications of microarray in gene expression

DNA microarrays or Nucleic acid arrays are a robust and cost-effective tool for a large-scale analysis of parallel gene expression studies, using a specific predesigned DNA sequence immobilised or bound on a solid surface such as glass which is known as a chip. This technology was used to measure the proportional concentration of nucleic acid sequences in a mixture of labelled nucleic acids through hybridisation properties and further detection of their outcomes. Rapid progress has been achieved in DNA arrays technology since the late '90s and 2000s. In addition to this, the Human Genome Project completion in 2001 provided the raw information needed to confirm that the arrays could be made entirely representing the genes in a genome, which raised challenges from biological exploration to medicine applications (*Bumgarner, 2014*).

The high throughput nature and the powerful combination of probe design algorithms and validation methods, of the technological evolution of surface chemistry and novel spotting techniques of DNA materials enabled the success of this technique. This technology has been

developed and combined as an ordinary tool in research laboratories and now has been converted into the clinic. Hence, it can be a worthwhile tool to understand the mechanisms behind a given disease system such as cancer.

DNA microarray is divided into two types, depending on the probe material spotted on the array surface: cDNA microarrays and oligonucleotide microarrays (Govindarajan *et al.*, 2012). In microarrays technology, 1000's of nucleic acids is bound to a surface (chip) and are used to measure the relative concentration of nucleic acid sequences in a mixture via hybridization and following the detection of the hybridization events (Cooper, 2001).

The focus of this chapter is to identify differences in gene expression between the two phenotypes of breast cancer. Up and down-regulated genes in BASAL-LIKE (MDA-MB-468) vs. HER2, (SkBr3) cell lines will be considered candidate markers for further investigations (chapter 5).

4.3. Identification of markers associated with BASAL-LIKE and HER2 breast cancer subtypes by applying the microarrays technique:

DNA microarray techniques have been used to identify the variables between Metformin treated and untreated samples. Regarding the previous findings, two cell lines have been chosen for further investigations. Those are the more sensitive cell line BASAL-LIKE (MDA-MB-468) and the less sensitive one HER2, (SkBr3). Moreover, two concentrations of Metformin have been used, which are the $\frac{1}{2}$ EC50 and the 2 EC50 concentrations (1 and 4 mM, respectively) and in addition to the control.

RNA extracts were obtained from three independent experiments in duplicate. The purity and integrity of RNA were determined using the Nanodrop and the Bioanalyser instruments, respectively. cRNAs samples were hybridised to Human GE 4x44 V2 Microarray slides, each containing 4 arrays with 27,958 RNA probes representing whole human genome and 10 x 32 E1A spike-in control probes according to the Agilent Technologies protocol as previously mentioned in materials and methods chapter section (2.2.3).

Data were extracted via Agilent Feature Extraction Software and were normalised *via* Partek Genomics Suite software. Later, the data were imported to TIBCO Statistica 13 software after labelling as zero for control and one for treatment; and transposed to a txt file. The entire data was run through the Artificial Neural Network (ANN) stepwise analysis, as previously mentioned in section (2.2.3.6.) Microarray Data Analysis.

4.3.1. RNA isolation and integrity determination for MDA-MB-468 and SkBr3 samples

In this study, RNAs were isolated from three independent duplicates. STAT 60 and Qiagen RNeasy Mini Kit were used for RNA-Extraction, following the manufacturer's protocol. RNA samples' quality was examined using Agilent RNA 6000 Nano Kit with RNA Nano Chips. Visual results for the quality analysis of each RNA sample are presented by an electropherogram trace (diagrams of fluorescence vs. length). The RNA integrity number (RIN), and the sample specific quality grade was computed from several features of the RNA electropherogram trace such as the ratio of 28S to 18S rRNA and the height of the 28S peak.

RIN was expressed on a scale ranging from 0 to 10 (with 10 being the best quality). Electropherograms for each analysed sample are shown in (Figures A.4.1. in Appendix). All tested RNA samples showed high RNA concentration and maximum RIN numbers; therefore, they have been adopted for further processing (Table 4.1).

Table 4.1. Results of the analysis of RNA concentration and integrity

Sample	Definition	Concentration (ng/ μ L)	RNA Integrity Number (RIN)
1	SkBr3 control_1	243.3186	9.8
2	SkBr3 control_2	217.423	9.9
3	SkBr3 1mM_1	362.953	10
4	SkBr3 1mM_2	375.5218	9.8
5	SkBr3 4mM_1	197.5103	9.6
6	SkBr3 4mM_2	386.245	9.8
7	SkBr3 control_3	641.6473	9.7
8	SkBr3 control_4	400.7732	9.8
9	SkBr3 1mM_3	506.3795	9.8
10	SkBr3 1mM_4	364.0004	9.8
11	SkBr3 4mM_3	442.9556	9.3
12	SkBr3 4mM_4	325.0985	9.3
13	SkBr3 control_5	250.5866	9.9
14	SkBr3 control_6	207.1992	9.5
15	SkBr3 1mM_5	347.4929	10
16	SkBr3 1mM_6	275.6745	9.9
17	SkBr3 4mM_5	273.0768	9.8
18	SkBr3 4mM_6	281.2303	9.8
19	MDA-MB-468 control_7	180.3707	10
20	MDA-MB-468 control_8	613.668	10
21	MDA-MB-468 1mM_7	104.704	10
22	MDA-MB-468 1mM_8	264.6074	10
23	MDA-MB-468 4mM_7	785.6243	8.5
24	MDA-MB-468 4mM_8	461.2615	9.3
25	MDA-MB-468 control_9	333.4336	10
26	MDA-MB-468 control_10	342.4385	10
27	MDA-MB-468 1mM_9	320.1664	9.4
28	MDA-MB-468 1mM_10	204.9831	9.9
29	MDA-MB-468 4mM_9	329.6005	8.3
30	MDA-MB-468 4mM_10	419.5941	10
31	MDA-MB-468 control_11	369.41	9
32	MDA-MB-468 control_12	456.615	9.9
33	MDA-MB-468 1mM_11	435.664	9.4
34	MDA-MB-468 1mM_12	500.4548	9.7
35	MDA-MB-468 4mM_11	322.7923	7.5
36	MDA-MB-468 4mM_12	439.95	9.1

Results are shown for RNA samples derived from two sets of cell lines (Samples 1-18) from HER2 (SkBr3) and (Samples 19-36) from BASAL-LIKE (MDA-MB-468) cell lines.

4.4. Labelling and cRNA quality assessment

After the quality control check has been passed successfully, the next step was to reverse transcribe all mRNA samples into cDNA. Double-stranded cDNAs were then transcribed into RNA using transcription mix containing a Cy3 dye to create Cy3-labelled complementary RNA (cRNA). Then, the generated cRNAs were assessed in terms of their concentration and the concentration of the incorporated Cy3 dye, (Figure 4.1).

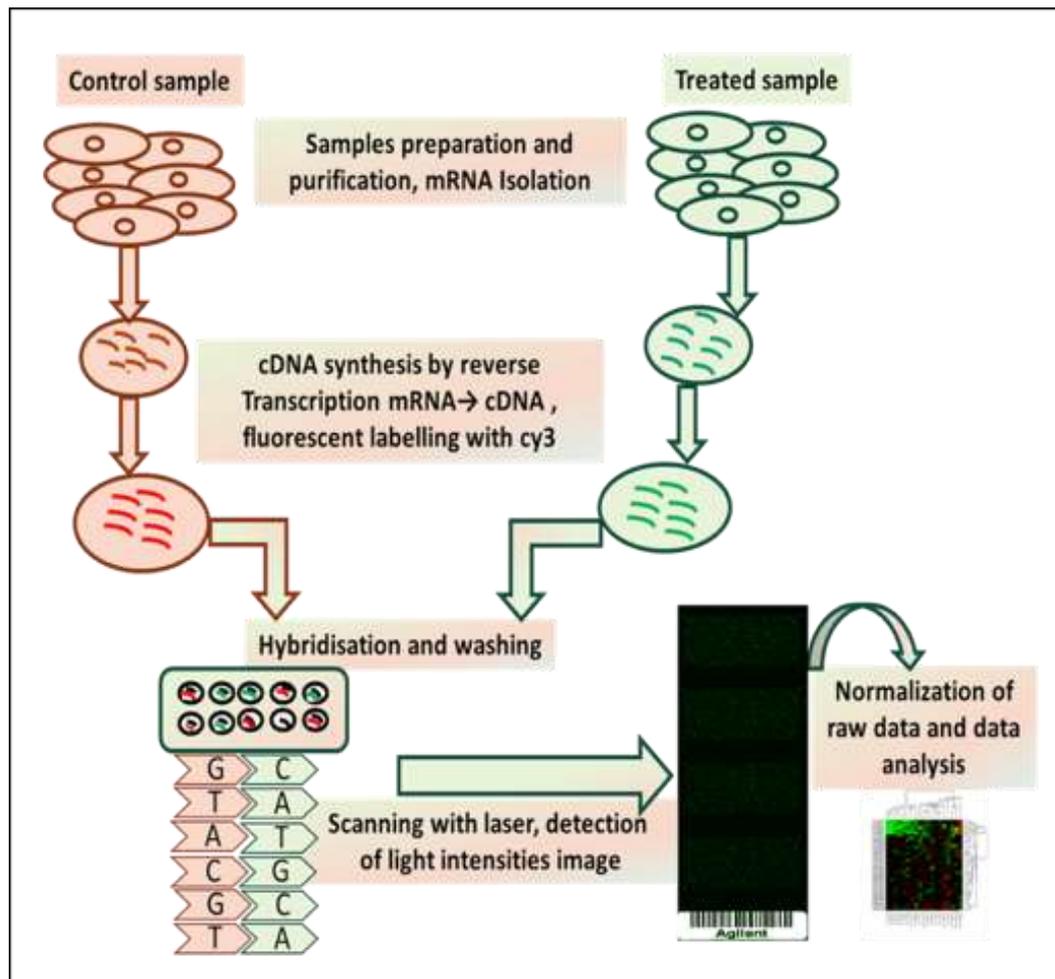


Figure 4.1. Schematic representation of the main steps involved in a microarray experiment. Adapted from (Duggan et al., 1999). The various steps involved in a typical cDNA microarray experiment are shown above. This includes preparation of cDNA probes and differential fluorescent tagging of the respective reference and test samples, the simultaneous hybridization of the labelled cDNA by complementary base pairing to the probes on the microarray chip and measuring fluorescent intensity as the level of gene expression.

These results were used to determine cRNA yield and Specific activity to ensure the sufficient integration of the Cy3 labelling. Both cRNA yield and Specific Activity (SA) parameters have met the quality standard and as recommended by the manufacturer (cRNA yield >5, SA >6). According to these data, the cRNA amount was adjusted to 1.65µg as also recommended by the manufacturer and 4-pack Microarray slide formats have been used (Tables 4.2, 4.3). The cRNAs samples were hybridised to Human GE 4x44 V2 Microarray slides, each containing 4 arrays with 27,958 RNA probes, representing the complete human genome and 10 x 32 E1A spike-in control probes. Quality Control (QC) Report obtained together with raw data for each sample includes a Spike-in check that shows the dose/response curve of the spike-ins from the detection limit to the saturation point. Spike-in Linearity check plots derived from the QC Report for each sample analysed and are showed in the Appendix (A.4. 2.). All samples examined presented linearity of 99%, which indicates an accurate and uniform array scanning and a satisfactory level of sensitivity of signal detection and saturation.

Table 4.2. Table representing cRNA yield and specific activity results (cRNA yield >5; SP Act>6)

Sample ID	Cy3conc. (pmol/ µl)	cRNA conc. (µg/µL)	260/280 Ratio	cRNA Yield(µg)	Specific activity/pmol	Vol per 1.65 µg
1	5.54	360.2	2.25	10.806	15.380	4.580
2	5.08	311.9	2.3	9.357	16.287	5.290
3	3.68	270.7	2.25	8.121	13.594	6.095
4	4.22	274.6	2.26	8.238	15.367	6.008
5	4.02	274.1	2.25	8.223	14.666	6.019
6	3.2	239.9	2.26	7.197	13.338	6.877
7	2.72	190.8	2.21	5.724	14.255	8.647
8	2.45	165.3	2.21	4.959	14.821	9.981
9	4.98	343.6	2.28	10.308	14.493	4.802
10	4.48	313	2.26	9.39	14.313	5.271
11	4.09	274	2.24	8.22	14.927	6.021
12	4.49	284.9	2.25	8.547	15.759	5.791
13	2.6	198.1	2.25	5.943	13.124	8.329
14	3.1	228.8	2.26	6.864	13.548	7.211
15	2.68	180.3	2.21	5.409	13.865	8.483
16	2.5	194.5	2.19	5.835	13.778	9.151
17	2.19	158.6	2.18	4.758	13.808	10.403
18	0.77	79.25	2.19	2.3775	9.7160	20.820

For the first set of samples HER2 (SkBr3) which met higher than recommended specific activity and cRNA yields.

Table 4.3. Table representing cRNA yield and specific activity results (cRNA yield >5; SP Act>6)

Sample ID	Cy3conc. (pmol/ μl)	cRNA conc. (μg/μL)	260/280 Ratio	cRNA Yield(μg)	Specific activity/pmol	Vol per 1.65 μg
19	4.26	316.4	2.22	9.492	13.464	5.215
20	8.87	435.9	2.29	13.077	20.349	3.785
21	3.05	264.4	2.18	7.932	11.536	6.241
22	3.49	252.7	2.24	7.581	13.811	6.529
23	2.49	167.8	2.19	4.455	15.892	11.111
24	2.2	183.2	2.16	5.496	12.009	9.007
25	4.01	278.6	2.21	8.358	14.393	5.922
26	2.55	194.4	2.23	5.832	13.117	8.488
23	2.36	148.5	2.19	4.704	11.926	10.523
27	1.87	156.8	2.19	5.547	13.953	8.924
28	2.58	184.9	2.24	3.945	13.840	12.548
29	1.82	131.5	2.21	9.219	15.197	5.369
30	4.67	307.3	2.24	5.238	12.658	9.450
31	2.21	174.6	2.25	5.058	12.040	9.786
32	2.03	168.6	2.24	3.534	9.677	14.007
33	1.14	117.8	2.17	5.217	13.744	9.488
34	2.39	173.9	2.29	3.141	9.169	15.759
35	0.96	104.7	2.12	3.096	14.826	15.988
29	1.53	103.2	2.24	5.709	6.831	8.671
36	1.3	190.3	2.07	4.686	5.570	10.563

For the second set of samples BASAL-LIKE (MDA-MB-468) which met higher than recommended specific activity and cRNA yields.

4.5. Microarray data analysis

It was necessary to summarise and transform all the probe set IDs into genes and perform Normalization using the Partek Genomics Suite analysis software (trial version). Then, the differential fold change (expression) was statistically computed using the previously mentioned software. The P-value was processed using unpaired T-test with Benjaminin Hochberg and FDR multiple corrections. Then, two type approaches have been used to analyse the Microarrays data. Firstly, the Artificial Neural Network (ANN) stepwise analysis has been applied to the entire data to visualise differentially expressed genes for both breast cancer phenotypes. In contrast, the regression-based method and Fold change (Fc) -based method have been used to differentiate expressed genes in each phenotype of breast cancer separately.

4.5.1. Artificial Neural Network (ANN) stepwise analysis-based approach

Artificial neural networks (ANNs) are intelligent thinking machines, working in non-linear mapping structures, based on the function of the human brain. They have been shown to be comprehensive and employable in a wide range of application. In addition, they are multivariate analysis methods. This powerful tool can treat complicated problems, identify and learn correlated patterns between input data sets and corresponding target values. Moreover, ANNs can be used to predict the output of new independent input data with high predictive accuracies (*Lancashire et al., 2009*).

The first approach has been achieved by importing the entire data into TIBCO Statistica 13.3 software and running with the Artificial Neural Network (ANN) stepwise analysis as previously mentioned in the methods chapter section (2.2.3.6.1. Artificial neural networks-based approach). To classify the samples as high or low gene expression, the median expression-based class split for individual genes predictors was carried out by calculating the median expression value for the average expression.

4.5.2. Interactome Network Map for both BASAL-LIKE (MDA-MB-468) and HER2 (SkBr3) cell lines

Interactome network maps have been used to demonstrate the interactions between the expressed genes. In the network map, a node symbolizes a single gene, and the link between genes is known as an edge, which can be presented with an arrow to indicate the direction of the link from a source node to a target node. Cytoscape, which is an open source bioinformatics software platform, has been utilised to visualize the molecular interaction networks and integrate with gene expression profiles (*Tong et al., 2014*).

Figure 4.2. demonstrated the actual interaction of the top 100 genes into a visual network map. The red arrows indicated to the up-regulated genes while the blue arrows represented the down-regulated genes. In addition, the width of the arrow shows the strength of the relation between the interacted genes. For a better identification to the system drivers within the data for both the source (influencers) and the target (the influenced), further analysis was carried out based on the entire matrix of interactions. Each gene was used as a source and a target separately; thus, the sum of all the interactions leading from or towards it were computed and the genes then ranked based on the highest absolute value. This analysis revealed, the highest influenced and influencer genes within this system (Table 4.4).

On comparison of the top 100-genes interactions map, the highest-ranking hubs from Figure 4.2. were also found to be top ranking target drivers (Table 4.4). As seen, the strongest influencers (source hubs) of the system are *Bet1 Golgi Vesicular Membrane Trafficking Protein Pseudogene 1 (BET1P1)*, *YLP Motif Containing 1(YLPM1)*, *Death-Associated Protein Kinase 3(DAPK3)* which, are all found in the map but not easy to identify as the key influencers. While the strongest influenced (target hubs) were *SKI/DACH Domain Containing 1 (Skida1)*, *KIAA0232 (KIAA0232)*, and *ATP Binding Cassette Subfamily B Member 8(ABCB8)* which, easily can be identified as the key influencers. In addition, *Skida1* has been noticed to be a positive source and a negative target at the same time. *Dihydrolipoamide Dehydrogenase (DLD)* has been found as a positive and negative target in the same analysis. Moreover, the top 100 gene map for the BASAL-LIKE (MDA-MB-468) and HER2 (SkBr3) cell lines when taken together, were generally negatively regulated.

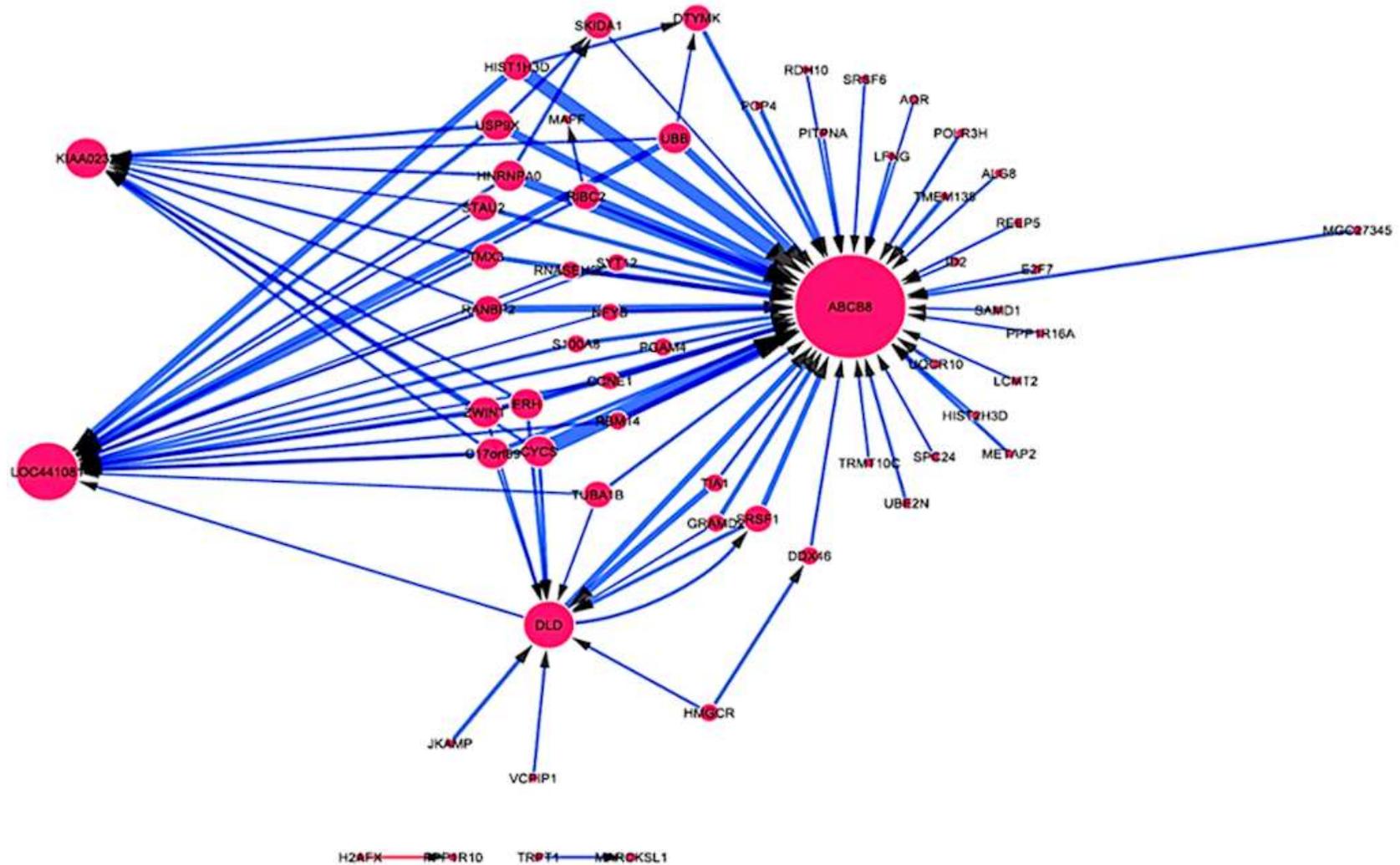


Figure 4.2. Interactome of the top 100 interactions between the 60 common genes in BASAL-LIKE (MDA-MB-468) and HER2 (SkBr3) cell lines for the two treatments concentrations compared to control. Red arrow represents the upregulated gene, while the blue arrows represent downregulated genes. List of the 60 interacted genes can be found in Appendix Table A. 4.1.

Table 4.4. BASAL-LIKE (MDA-MB-468) and HER2 (SkBr3) associated top 12 driver genes, for the two treatment concentrations compared to control.

Rank	source	Gene Name	Average interaction sum	Absolute value
1	BET1P1	Bet1 Golgi Vesicular Membrane Trafficking Protein Pseudogene 1	-331.26	331.26
2	YLPM1	Nucleophosmin 1	-322.36	322.36
3	DAPK3	Death-Associated Protein Kinase 3	-310.6	310.6
4	FLII	FLII, Actin Remodeling Protein	-297.96	297.96
5	TTC7B	Tetratricopeptide Repeat Domain 7B	-270.87	270.87
6	KIAA1161	Myogenesis Regulating Glycosidase (Putative)	-251.68	251.68
7	GOLGA2	Golgin A2	-242.12	242.12
8	TMEM248	Transmembrane Protein 248	-238.74	238.74
9	KIAA0040	Uncharacterise d Protein KIAA0040	-224.14	224.14
10	RIBC2	RIB43A Domain With Coiled-Coils 2	-201.36	201.36
11	Skida1	SKI/DACH Domain Containing 1	122.24	122.24
12	ETHE1	ETHE1, Persulfide Dioxygenase	142.28	142.28

Rank	Target	Gene Name	Average interaction sum	Absolute value
1	Skida1	SKI/DACH Domain Containing 1	-331.26	331.26
2	KIAA0232	Uncharacterise d Protein KIAA0232	-322.36	322.36
3	ABCB8	ATP Binding Cassette Subfamily B Member 8	-310.6	310.6
4	LOC441081	POM121 Membrane Glycoprotein (Rat) Pseudogene	-297.96	297.96
5	SRSF1	Serine And Arginine Rich Splicing Factor 1	-270.87	270.87
6	STIM1	Stromal Interaction Molecule 1	-251.68	251.68
7	DLD	Dihydrolipoamide Dehydrogenase	-242.12	242.12
8	METAP2	Methionyl Aminopeptidase 2	-238.74	238.74
9	EIF4E	Eukaryotic Translation Initiation Factor 4E	-224.14	224.14
10	MAFF	MAF BZIP Transcription Factor F	-201.36	201.36
11	DLD	Dihydrolipoamide Dehydrogenase	122.24	122.24
12	Uqcr10	Ubiquinol-Cytochrome C Reductase, Complex III Subunit X	142.28	142.28

The top absolute value ranked source (influencers) and target (influenced) drivers were identified by the driver analysis, a method that utilises the entire matrix of interactions with each gene used as a source and a target separately, to calculate the sum of all the interactions leading from or towards it.

4.5.3. Interactome Network Map for BASAL-LIKE (MDA-MB-468) and HER2 (SkBr3) cell lines assessed separately

The same approach (ANN) stepwise analysis has been applied to the same population after separating the data into two groups, BASAL-LIKE (MDA-MB-468) and HER2 (SkBr3) to compare gene expression of each cell line independently. Again, the results were presented as molecular interaction networks for BASAL-LIKE (MDA-MB-468) and HER2 (SkBr3) cell lines, respectively (Figures 4.3. and 4.4.). In addition, the analysis of entire matrix interactions was also carried out to identify the highest influenced and influencer genes within this system (Table 4.5 and Table 4.6).

As before, the top 100 highest intensity interactions between the 48 probes in BASAL-LIKE (MDA-MB-468) cells were selected for a better identification of key driver hubs within these cells. From the map in Figure 4.3, it can be observed that the top target hubs or drivers based on the highest number of in-degree edges in order are *IZUMO Family Member 2 (IZUMO2)*, *Serpin Family E Member 1 (SERPINE1)*, *Coagulation Factor VIII Associated 2 (F8A2)*, *EGF Like Domain Multiple 7 (EGFL7)*, and *B-Raf Proto-Oncogene, Serine/Threonine Kinase (BRAF)* while, the top target hubs were *Limb Development Membrane Protein 1 (LMBR1)*, *Bromodomain Adjacent To Zinc Finger Domain 2B (BAZ2B)*, *ABR*, *RhoGEF And GTPase Activating Protein (ABR)*, and *NK3 Homeobox 2 (NKX3-2)*. The overall interaction feedback loops were negative except for a few positive loops.

On the other hand, the top 100 strongest interactions between the 77 combined genes in HER2 (SkBr3) cells, were found to be between *G Protein-Coupled Receptor 62 (GPR62)*, *Chromosome 1 Open Reading Frame 159(C1orf159)*, *Synapsin II(SYN2)* and *Glucuronidase Beta (Inc-GUSB-5)* as a source. However, *Doublecortin Domain Containing 1 (DCDC1)*, *HIG1 Hypoxia Inducible Domain Family Member 1A Pseudogene 1 (HIGD1AP1)*, *ADAM Metallopeptidase With Thrombospondin Type 1 Motif 9 (ADAMTS9)*, and *Iduronate 2-Sulfatase (IDS)* were ranked as the robustly interacted targets, as shown in Table 4.6. In this map, both positive and negative feedback loops can be observed. Additionally, it was noticeable that the *F-Box And Leucine-Rich Repeat Protein 8 (FBXL8)* probe has been found as a “common gene” as a source and target in the same time in HER2 (SkBr3) cells (Table 4.6).

Moreover, *Synapsin II (SYN2)* was noticed as "common" between the two populations. The former played as a target in BASAL-LIKE (MDA-MB-468) phenotype and as a source in HER2 (SkBr3) cell line. Over and above, the top four ranked sources and targets were negative feedback loops (Figure 4.4).

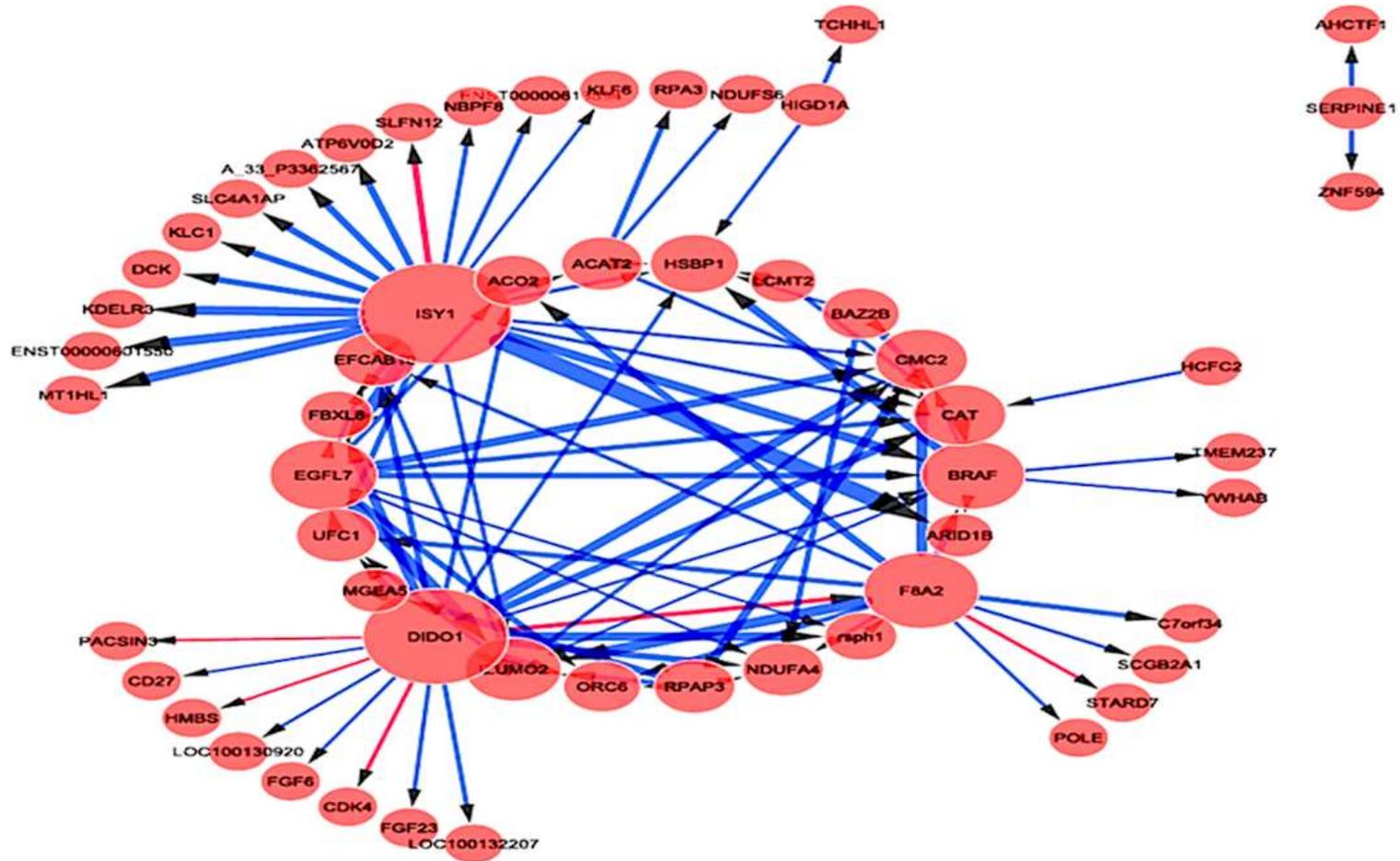


Figure 4.3. Interactome of the top 100 interactions between the 48 common genes in BASAL-LIKE (MDA-MB-468) cell line. The red arrows represent up-regulated genes, while the blue arrows represent down-regulated genes. List of the 48 interacted genes can be found in Appendix Table A.4.2.

Table 4.5. BASAL-LIKE (MDA-MB-468) associated top 12 driver genes for the two treatment concentrations compared to control

Rank	Source	Gene Name	Average interaction sum	Absolute value
1	LMBR1	Limb Development Membrane Protein 1	-227.05	227.05
2	BAZ2B	Bromodomain Adjacent To Zinc Finger Domain 2B	-220.88	220.88
3	ABR	ABR, RhoGEF And GTPase Activating Protein	-215.39	215.39
4	NKX3-2	NK3 Homeobox 2	-215.03	215.03
5	CNTNAP5	Contactin Associated Protein-Like 5	-204.03	204.03
6	ECT2L	Epithelial Cell Transforming 2 Like	-198.38	198.38
7	PKDCC	Protein Kinase Domain Containing, Cytoplasmic	61.58	61.58
8	STARD7	StAR-Related Lipid Transfer Domain Containing 7	69.59	69.59
9	F2	coagulation factor II	71.22	71.22
10	THC2670954	Unknown	80.26	80.26
11	TBC1D5	TBC1 Domain Family Member 5	122.26	122.26
12	RGS7	Regulator Of G Protein Signaling 7	142.28	142.28

Rank	Target	Gene Name	Average interaction sum	Absolute value
1	IZUMO2	IZUMO Family Member 2	-227.05	227.05
2	SERPINE1	Serpin Family E Member 1	-220.88	220.88
3	F8A2	Coagulation Factor VIII Associated 2	-215.39	215.39
4	EGFL7	EGF Like Domain Multiple 7	-215.03	215.03
5	BRAF	B-Raf Proto-Oncogene, Serine/Threonine Kinase	-204.03	204.03
6	SYN2	Synapsin II	-198.38	198.38
7	TNFRSF12A	TNF Receptor Superfamily Member 12A	61.58	61.58
8	NACC1	Nucleus Accumbens Associated 1	69.59	69.59
9	SSBP4	Single-Stranded DNA Binding Protein 4	71.22	71.22
10	TOM1L1	Target Of Myb1 Like 1 Membrane Trafficking Protein	80.26	80.26
11	THC2673793	Unknown	122.26	122.26
12	SLC4A1AP	Solute Carrier Family 4 Member 1 Adaptor Protein	142.28	142.28

The top absolute value ranked source (influencers) and target (influenced) drivers. These were identified by the driver analysis, a method that utilizes the entire matrix of interactions with each gene used as a source and a target separately, to calculate the sum of all the interactions leading from or towards it.

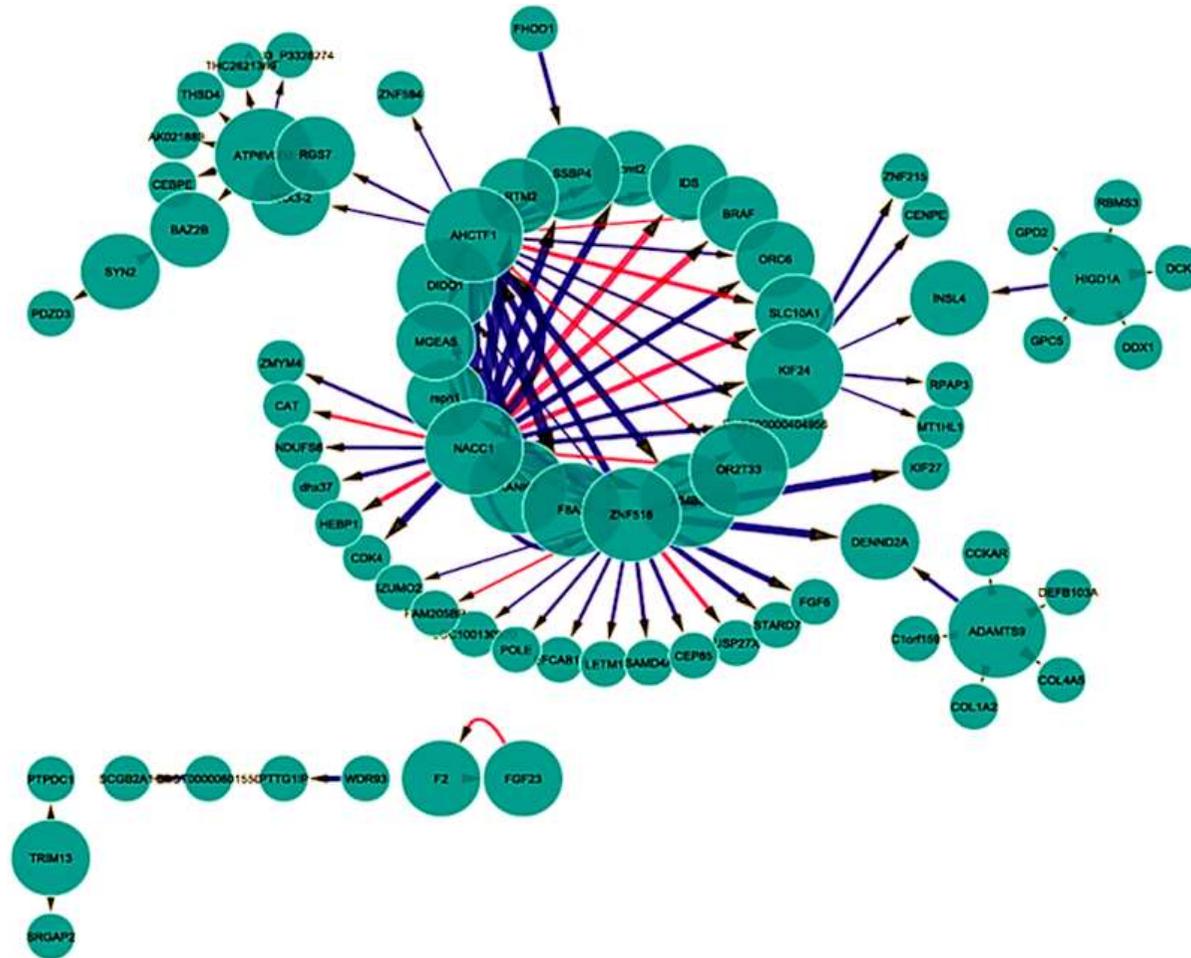


Figure 4.4. Interactome of the top 100 interactions between the 77 common genes in HER2 (SkBr3) cell line. The red arrows represent upregulated genes, while the blue arrows represent downregulated genes. List of the 77 interacted genes can be found in Appendix Table A.4.3.

Table 4.6. HER2 (SkBr3) associated top 12 driver genes for the two treatment concentrations compared to control

Rank	Source	Gene Name	Average interaction sum	Absolute value
1	GPR62	G Protein-Coupled Receptor 62	-213.14	213.14
2	C1orf159	Chromosome 1 Open Reading Frame 159	-209.3	209.3
3	SYN2	Synapsin II	-201.76	201.76
4	Inc-GUSB-5	Glucuronidase Beta	-184.4	184.4
5	SLFN12	Schlafen Family Member 12	41.77	41.77
6	COL1A2	Collagen Type I Alpha 2 Chain	44.5	44.5
7	A_33_P3266889	Unknown	45.69	45.69
8	UNC5CL	Unc-5 Family C-Terminal Like	46.21	46.21
9	FBXL8	F-Box And Leucine-Rich Repeat Protein 8	49.36	49.36
10	MAPK11	Mitogen-Activated Protein Kinase 11	52.03	52.03
11	FAM205BP	Family With Sequence Similarity 205 Member B, Pseudogene	53.26	53.26
12	ATP6V0D2	ATPase H+ Transporting V0 Subunit D2	70.78	70.78

Rank	Target	Gene Name	Average interaction sum	Absolute value
1	DCDC1	Doublecortin Domain Containing 1	-213.14	213.1399779
2	HIGD1AP1	HIG1 Hypoxia Inducible Domain Family Member 1A Pseudogene 1	-209.3	209.3
3	ADAMTS9	ADAM Metallopeptidase With Thrombospondin Type 1 Motif 9	-201.76	201.76
4	IDS	Iduronate 2-Sulfatase	-184.4	184.4
5	A_33_P3338788	Unknown	41.77	41.77
6	FBXL8	F-Box And Leucine-Rich Repeat Protein 8	44.5	44.5
7	CAT	Catalase	45.69	45.69
8	ACO2	Aconitase 2	46.21	46.21
9	DDX1	DEAD-Box Helicase 1	49.36	49.36
10	A_33_P3326898	Unknown	52.03	52.03
11	NOXA1	NADPH Oxidase Activator 1	53.26	53.26
12	FAM122A	Family With Sequence Similarity 122A	70.78	70.78

The top absolute value ranked source (influencers) and target (influenced) drivers. These were identified by the driver analysis, a method that utilizes the entire matrix of interactions with each gene used as a source and a target separately, to calculate the sum of all the interactions leading from or towards it.

The number of samples that have been used for the Gene Expression Microarray experiment was 36 samples only, which were sufficient to run the ANN. As previously cited, prior to the ANN training, the sample will randomly be divided into three subsets; 60% for training, 20% for validation and 20% for testing. However, to compare each cell line individually and compare between the two concentrations, as well as separately, means that the number of samples will be reduced. Hence, an alternative method was needed. For this aim, the Regression-based method and fold change- based method was utilised for data analytics.

4.5.4. Regression-based method and fold change- based method for analysis of data from BASAL-LIKE (MDA-MB-468) and HER2 (SkBr3) cell lines

This approach has been conducted to compare and analyse the gene expression levels between BASAL-LIKE (MDA-MB-468) and HER2 (SkBr3) cell lines samples individually, as well as to compare the different doses of Metformin treatments that have been used 1 and 4 mM). The two-dimensional comparison was conducted in the second approach, between the different breast cancer subtypes as the main comparison and between the different doses of Metformin as sub comparison.

To apply this approach, an experimental group was created with the MDA-MB-468 cell line as a first group, and SkBr3 cell line as a second group. Each group was separately analysed. Moreover, two other subgroups were created according to the two treatments concentrations (EC 25 and 2 EC50) as shown in Figure 4.5.

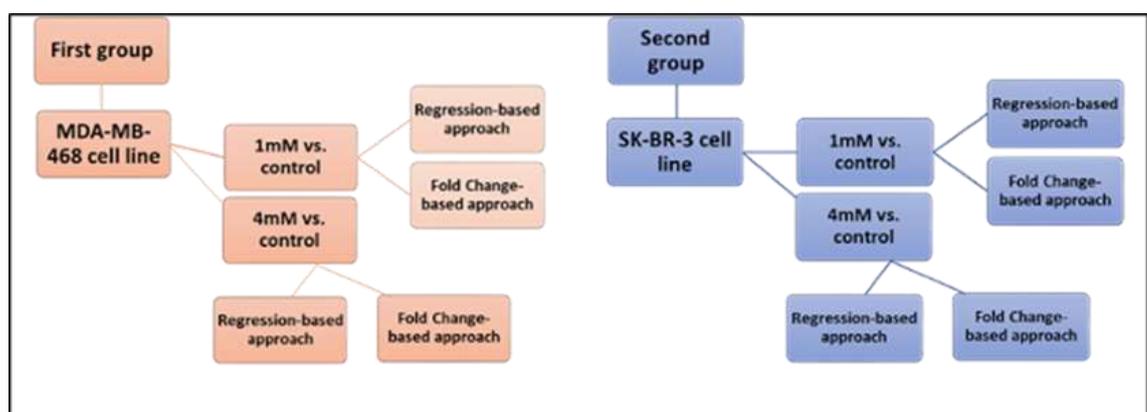


Figure 4.5. Schematic diagram showing the second approach of Gene expression microarray data analysis. Each cell line was divided into two subgroups and two statistical methods were applied to each subgroup.

This analysis was performed with a regression-based method, which calculated the residuals and the standard residuals. The average of standard residuals was calculated, and the genes were ranked based on their expression values. Moreover, the fold change (Fc) for the same set of samples was calculated using one cut-off criterion ($2 Fc \geq -2$), p-value ($p \leq 0.05$). The P value for each gene was calculated with multiple testing correction using Student T-DIST, Bonferroni correction, and Benjamini Hochberg. A detailed description was provided in the method chapter section (2.2.3.6.2. Regression-based method and fold change (Fc)- based method).

This approach allowed the choice of highly significant Bonferroni values and minimise the false discovery rate (FDR). The reported differential expressions in this study were obtained from both the regression-based approach and Fold change-based approach, respectively.

The Morpheus-Broad Institute online software was utilised to create the Hierarchical clustering heat maps. The heat map is a two-dimensional representation of data in which values are symbolized by colours. The heat maps in Figures 4.6, 4.7 show a graphical representation of the total number of differentially expressed genes (where individual values are represented as colours) out of the previously mentioned approaches (Fold change- based and regression-based, respectively in BASAL-LIKE (MDA-MB-468). In addition, Tables 4.7 and 4.8 display the list of upregulated and downregulated genes in BASAL-LIKE (MDA-MB-468) cell line samples and in different concentrations when compared to control.

Figures 4.8 and 4.9 referred to HER2 (SkBr3) fold change and regression analysis data representation individually, while Tables 4.9 and 4.10 show the different genes expression patterns that are significantly upregulated or downregulated. This data is obtained from both (Fc), and Regression analysis in HER2 (SkBr3) cell line samples, the Bonferroni Correction was $\leq 0.05(*)$.

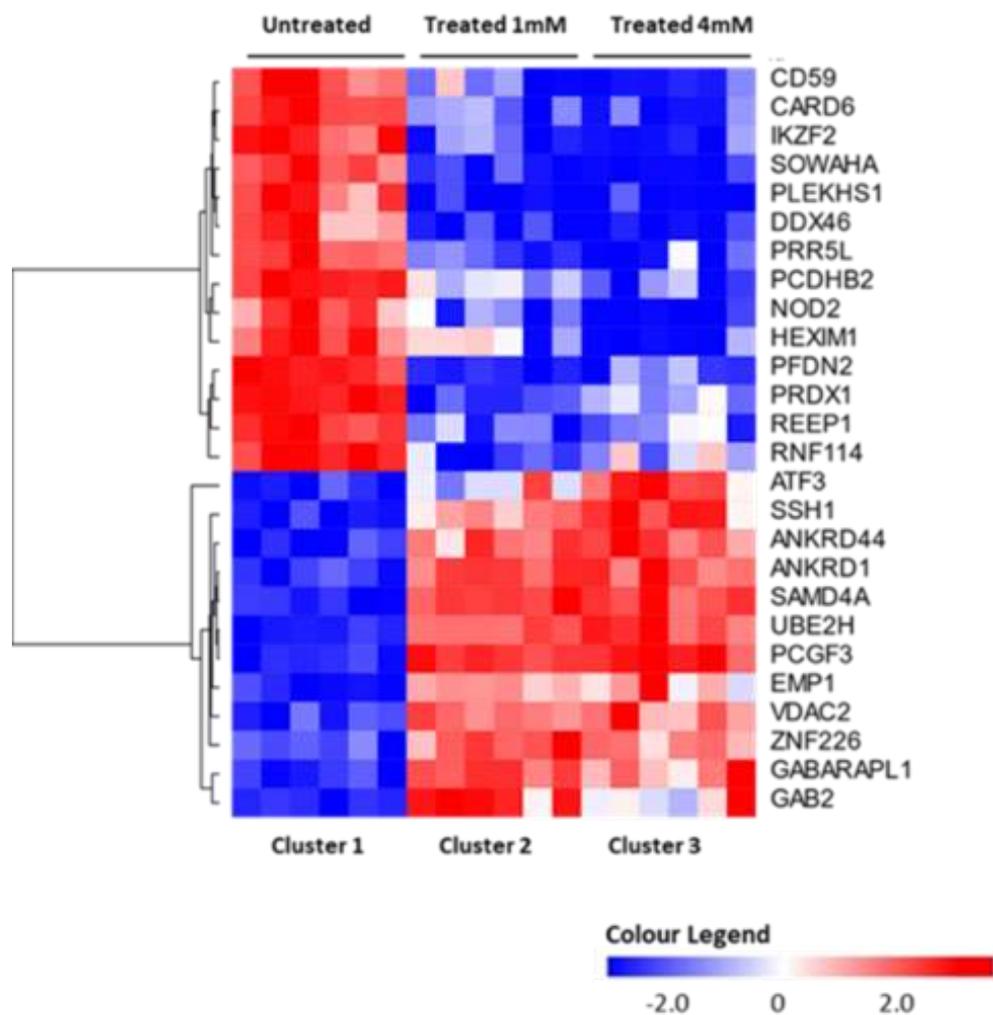


Figure 4. 6. Heat map representing the difference in the gene expression pattern between BASAL-LIKE (MDA-MB-468) cell line samples that were treated with different concentrations of Metformin and compared to control with $Fc \geq -2$. The colour gradient is between red and blues with white in the middle. The heat map shows 26 genes differently expressed according to different treatments.

Table 4.7 shows the number of genes that satisfied the cut-off criteria $Fc \geq 2$, Bonferroni Correction ≤ 0.05 (*). Gene expression pattern revealed differential expression of 26 genes (with cut-off $FC \geq 2$, Bonferroni Correction ≤ 0.05 (*), including 9 upregulated and 17 downregulated in BASAL-LIKE (MDA-MB-468) cell line samples treated with different concentrations and compared to controls.

Table 4. 7. List of genes upregulated and downregulated in BASAL-LIKE (MDA-MB-468) cell line Samples that were treated with different concentrations of Metformin and compared to controls with $Fc \geq 2$, Bonferroni Correction ≤ 0.05 (*).

MDA-MB-468 4mM				
Gene Symbol	Description	Fold Change	P-Value	Bonferroni corrected
SSH1	<i>Slingshot protein phosphatase 1</i>	2.366	9.47327E-07	0.032
ATF3	<i>Activating transcription factor 3</i>	2.135	1.61397E-06	0.055
ANKRD1	<i>Ankyrin repeat domain 10</i>	2.115	9.13559E-08	0.003
ANKRD44	<i>Ankyrin repeat domain 44</i>	2.050	1.39402E-07	0.005
NOD2	<i>Nucleotide binding oligomerization domain containing 2</i>	-2.065	3.292E-07	0.011
DDX46	<i>DEAD-box helicase 46</i>	-2.067	1.39636E-06	0.048
HEXIM1	<i>Hexamethylene bisacetamide inducible 1</i>	-2.129	9.57046E-07	0.033
CD59	<i>CD59 molecule</i>	-2.163	1.56252E-07	0.005
SOWAHA	<i>Sosondowah ankyrin repeat domain family member A</i>	-2.190	6.81005E-09	0.000
PCDHB2	<i>Protocadherin beta 2</i>	-2.242	2.79865E-07	0.010
CARD6	<i>Caspase recruitment domain family, member 6</i>	-2.344	1.32221E-07	0.005
PLEKHS1	<i>Pleckstrin homology domain containing S1</i>	-2.475	2.27157E-07	0.008
IKZF2	<i>IKAROS family zinc finger 2</i>	-2.485	2.72316E-07	0.009

MDA-MB-468 1mM				
Gene Symbol	Description	Fold Change	P-Value	Bonferroni corrected
SAMD4A	<i>Sterile alpha motif domain containing 4A</i>	2.166	6.2123E-11	2.12007E-06
PRDX1	<i>Peroxiredoxin 1</i>	2.162	3.25844E-10	1.11201E-05
PCGF3	<i>Polycomb group ring finger 3</i>	2.135	4.25531E-11	1.45221E-06
UBE2H	<i>Ubiquitin-conjugating enzyme E2 H</i>	2.067	3.25644E-11	1.11132E-06
PFDN2	<i>Prefoldin subunit 2</i>	2.043	3.74805E-10	1.2791E-05
ZNF226	<i>Zinc finger protein 300</i>	-2.041	8.95164E-07	0.030
VDAC2	<i>Voltage-dependent anion channel 2</i>	-2.042	5.2118E-08	0.001
RNF114	<i>Ring finger protein 114</i>	-2.071	8.40503E-07	0.028
REEP1	<i>Receptor accessory protein 1</i>	-2.078	1.55661E-06	0.053
EMP1	<i>Epithelial membrane protein 1</i>	-2.107	2.44773E-09	8.35337E-05
GAB2	<i>GRB2 associated binding protein 2</i>	-2.370	7.34812E-07	0.025
GABARAPL1	<i>GABA (A) receptor-associated protein like 1</i>	-2.475	3.01892E-09	0.0001
PRR5L	<i>Proline-rich 5 like</i>	-2.505	1.52578E-07	0.005

Gene expression pattern has revealed differential expression of 14 genes, including 4 upregulated genes and 10 downregulated genes in BASAL-LIKE (MDA-MB-468) cell line samples using 4mM Metformin treatment. In addition, 5 upregulated genes and 8 downregulated genes in 1mM Metformin-treated samples and when compared to controls. Red colour indicated upregulated genes and blue colour downregulated genes.

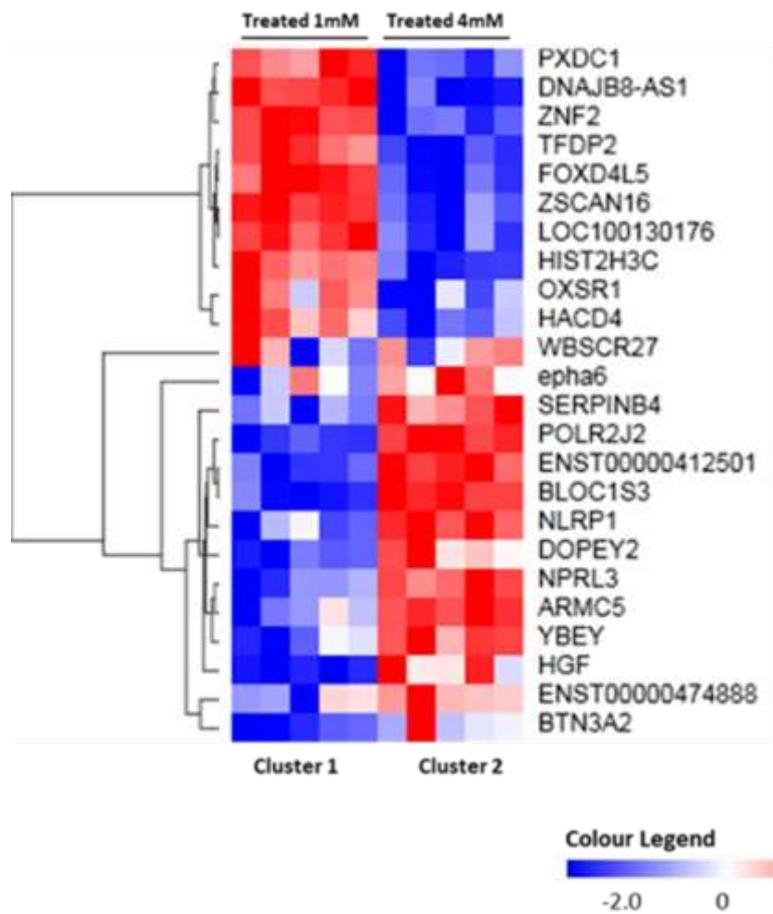


Figure 4. 7. Heat map representing the difference in gene expression patterns between BASAL-LIKE (MDA-MB-468) cell line samples treated with 1mM and 4mM doses using Regression test. The colour gradient is between red and blues with white in the middle. A heat map shows 25 genes differently expressed according to different treatments concentrations.

Table 4.8 below shows the number of genes that achieved the highly significant Bonferroni Correction ≤ 0.05 (*). Gene expression pattern has exposed differential expression of 25 genes including 10 upregulated and 15 downregulated in BASAL-LIKE (MDA-MB-468) cell line samples using different concentrations and compared to controls.

Table 4. 8. List of upregulated and downregulated genes in BASAL-LIKE (MDA-MB-468) cell line Samples using different concentrations and compared to controls in Regression test, Bonferroni Correction ≤ 0.05 (*).

MDA-MB-468 4mM				
Gene Symbol	Description	Average of standard residuals	p-value	Bonferroni Corrected
<i>WBSCR27</i>	<i>Williams Beuren syndrome chromosome region 27</i>	6.754	1.45394E-11	4.96185E-07
<i>SERPINB4</i>	<i>Serpin family B member 4</i>	5.035	4.78565E-07	0.016
<i>DNAJB8-AS1</i>	<i>DNAJB8 antisense RNA 1, long non-coding RNA</i>	-4.817	1.45931E-06	0.049
<i>PXDC1</i>	<i>PX domain containing 1(PXDC1).</i>	-5.373	7.75787E-08	0.002

MDA-MB-468 1mM				
Gene Symbol	Description	Average of standard residuals	p-value	Bonferroni Corrected
<i>WBSCR27</i>	<i>Williams Beuren syndrome chromosome region 27</i>	6.706	2.03126E-11	6.93209E-07
<i>ZNF2</i>	<i>Zinc finger protein 2</i>	6.275	3.52788E-10	1.20396E-05
<i>TFDP2</i>	<i>Transcription factor Dp-2</i>	5.536	3.11691E-08	0.001
<i>ZSCAN16</i>	<i>Zinc finger and SCAN domain-containing 16</i>	5.430	5.65294E-08	0.001
<i>OXSRI</i>	<i>Oxidative stress responsive 1</i>	5.422	5.91336E-08	0.002
<i>FOXD4L5</i>	<i>Forkhead box D4-like 5</i>	5.184	2.17484E-07	0.007
<i>LOC100130176</i>	<i>Transmembrane protein 188 pseudogene</i>	5.151	2.60258E-07	0.008

HIST2H3C	Histone cluster 2 H3 family member c	5.131	2.89666E-07	0.009
HACD4	3-hydroxy acyl-CoA dehydratase 4	4.910	9.13949E-07	0.031
Gene Symbol	Description	Average of standard residuals	p-value	Bonferroni Corrected
SLC12A9	solute carrier family 12, member 9, transcript variant 1	-4.815	1.47357E-06	0.050
BTN3A2	Butyrophilin subfamily 3-member A2	-4.816	1.46393E-06	0.049
YBEY	YbeY metalloproteinase (putative), transcript variant 1.	-4.834	1.34178E-06	0.045
ARMCS	Armadillo repeat containing 5, transcript variant 2.	-4.894	9.92157E-07	0.033
DOPEY2	Dopey family member 2	-4.905	9.3853E-07	0.032
ENST00000474888	Small nuclear ribonucleoprotein polypeptide C pseudogene 3.	-5.033	4.83771E-07	0.016
HGF	Hepatocyte growth factor.	-5.064	4.10966E-07	0.014
NLRP1	NLR family pyrin domain containing 1	-5.246	1.55979E-07	0.005
BLOC1S3	Biogenesis of lysosomal organelles complex-1.	-5.366	8.07266E-08	0.002
ENST00000412501	Olfactory receptor, family 13.	-5.525	3.30497E-08	0.001
epha6	EPH receptor A6	-5.744	9.30484E-09	0.0003
NPRL3	NPR3 like GATOR1 complex subunit	-6.407	1.50243E-10	5.12734E-06
POLR2J2	RNA polymerase II subunit J2	-7.626	2.46892E-14	8.42568E-10

Analysis was applied in BASAL-LIKE (MDA-MB-468) cell line samples treated with 4mM Metformin, 4 genes were highly significant using Bonferroni Correction ≤ 0.05 (*), including 2 upregulated and 2 downregulated genes. In addition, 9 were upregulated and 13 downregulated genes when cells were treated with 1mM of Metformin and when compared to controls. *WBCR27* was the common upregulated gene in cells treated with 2EC50 and ½ EC50 concentrations of Metformin. Red colour indicated upregulated genes and blue colour downregulated genes.

The 21 differentially expressed genes in HER2 (SkBr3) cell line, which was treated with 4mM and 1mM Metformin concentrations are shown in the heat map (4.8) below.

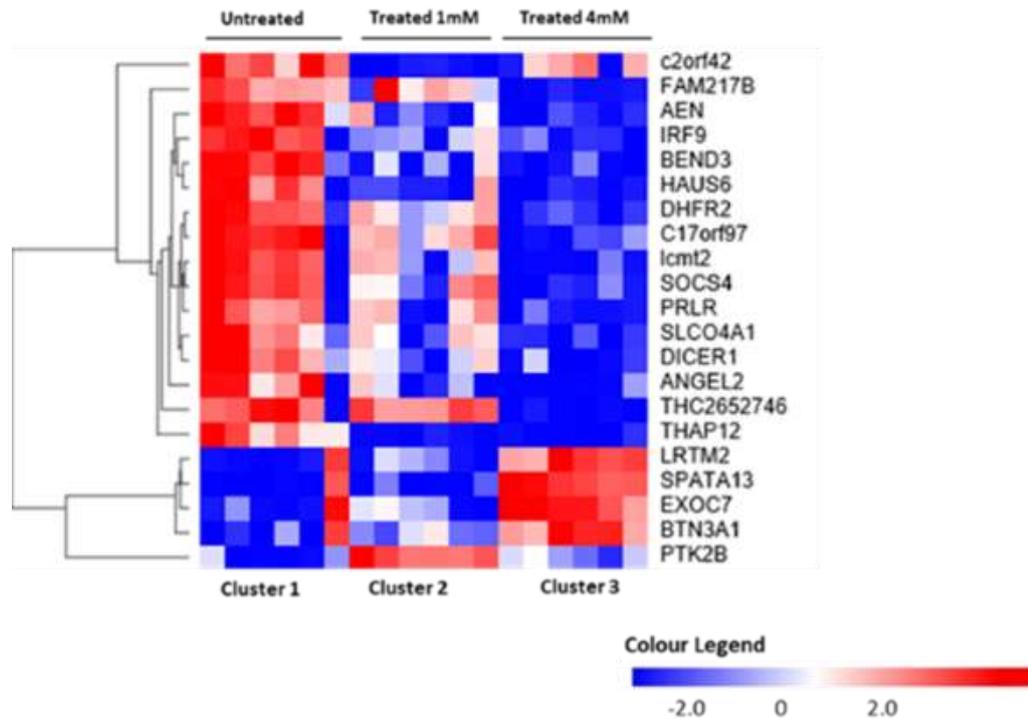


Figure 4. 8. The heat map represents the difference in gene expression pattern between HER2 (SkBr3) cell line samples using different concentrations of Metformin and compared to control with $Fc \geq -2$. The colour gradient is between red and blues with white in the middle. 21 genes were differentially expressed according to different treatments concentrations.

Table 4.9 shows the number of genes that satisfied the cut-off criteria $Fc_2 \geq -2$, $p\text{-value } (p) \leq 0.05$. The gene expression pattern revealed 21 differentially expressed genes (with cut-off $Fc_2 \geq -2$, $p < 0.001\text{-}0.0001$ (***-****)). This includes 4 upregulated and 17 downregulated genes in HER2 (SkBr3) cell line samples treated with different concentrations and compared to controls.

Table 4. 9. List of upregulated and downregulated genes in HER2 (SkBr3) cell line Samples treated with different concentrations and compared to control with $Fc_2 \geq -2$, $p < 0.001\text{-}0.0001$ (*-****).**

SkBr3 4mM			
Gene Symbol	Description	Average of standard residuals	p-value
<i>SPATA13</i>	<i>Spermatogenesis associated 13</i>	2.303	0.0003
<i>EXOC7</i>	<i>Exocyst complex component 7</i>	2.093	0.002
<i>IRF-9</i>	<i>Interferon regulatory factor 9</i>	2.014	0.001
<i>BTN3A1</i>	<i>Homo sapiens butyrophilin, subfamily 3, member A1, transcript variant 1</i>	2.008	0.003
<i>DICER1</i>	<i>Dicer 1, ribonuclease type III</i>	-2.009	0.0002
<i>PRLR</i>	<i>Prolactin receptor (PRLR).</i>	-2.012	0.001
<i>SOCS4</i>	<i>Suppressor of cytokine signalling 4</i>	-2.022	0.001
<i>c2orf42</i>	<i>Chromosome 2 open reading frame 42</i>	-2.031	5.83448E-08
<i>LRTM2</i>	<i>Leucine-rich repeats and transmembrane domains 2</i>	-2.044	5.60099E-06
<i>BEND3</i>	<i>BEN domain containing 3</i>	-2.045	0.002
<i>C17orf97</i>	<i>Chromosome 17 open reading frame 97</i>	-2.079	0.002

FAM217B	Family with sequence similarity 217-member B	-2.091	0.0005
AEN	Apoptosis enhancing nuclease	-2.106	0.001
ANGEL2	Angel homolog 2	-2.107	0.003
HAUS6	HAUS augmin like complex subunit 6	-2.124	0.001
lcmt2	Leucine carboxyl methyltransferase 2	-2.154	0.0007
THAP12	THAP domain containing 12	-2.211	6.24919E-06
SLCO4A1	Solute carrier organic anion transporter family member 4A1.	-2.253	0.0001
THC2652746	tc Q5SY11_HUMAN (Q5SY11) Collagen, type V, alpha 1.	-2.316	0.0007

SkBr3 1mM			
Gene Symbol	Description	Average of standard residuals	p-value
PTK2B	Protein tyrosine kinase 2 beta	2.201257043	8.36492E-06
DHFR2	Dihydrofolate reductase 2	-2.041615903	1.9546E-07
THAP12	THAP domain containing 12	-2.205949006	5.78134E-06

With a 4mM Metformin concentration, 4 genes were upregulated, and 15 genes were downregulated. However, in 1mM Metformin concentration, only 1 gene was upregulated, and 2 genes were down-regulated. In addition, *THAP12* gene was shown as downregulated in both treatment conditions. Red colour indicated upregulated genes and blue colour downregulated genes.

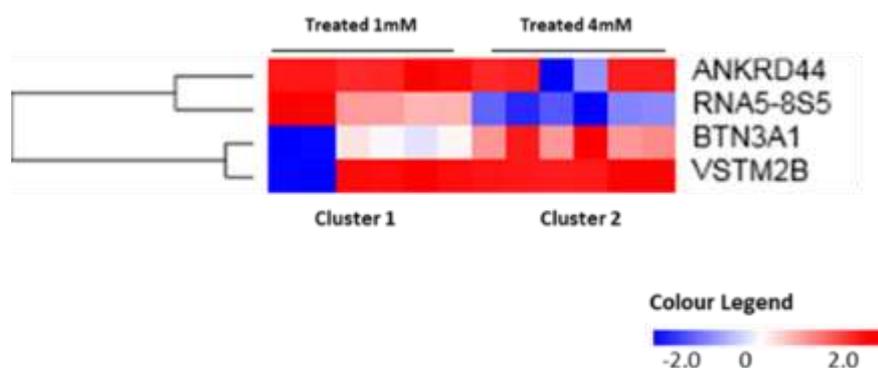


Figure 4. 9. Heat map representing the difference in gene expression pattern between HER2 (SkBr3) cell line samples treated with 1mM and 4mM concentrations in Regression test. The colour gradient is between red and blues with white in the middle. The heat map shows 4 differently expressed genes and according to different treatment concentrations.

Table 4.10 shows the number of genes that achieved the highly significant P values, in the Regression based method, $p < 0.0001$ (****). Gene expression pattern has exposed differential expression of four genes including one upregulated and one downregulated in each concentration in HER2 (SkBr3) cell line samples.

Table 4. 10. The list of genes upregulated and downregulated in HER2 (SkBr3) cell line Samples treated with different concentrations of Metformin and compared to control in Regression test, $p < 0.0001$ (****).

SkBr3 4mM			
Gene Symbol	Description	Average of standard residuals	p-value
BTN3A1	<i>Butyrophilin, subfamily 3, member A1.</i>	4.466	7.95913E-06
ANKRD44	<i>Ankyrin repeat domain 44</i>	-5.878	4.17474E-09

SkBr3 1mM			
Gene Symbol	Description	Average of standard residuals	p-value
RNA5-8S5	<i>RNA, 5.8S ribosomal 5</i>	5.028	4.96422E-07
VSTM2B	<i>V-set and transmembrane domain containing 2B</i>	-4.901	9.54479E-07

Gene expression model displayed one upregulated and one downregulated gene in each of Metformin treatment conditions. Red colour indicated upregulated genes and blue colour downregulated genes.

The summation of rank for all the highly significant genes based on the second approach analysis outcomes has been calculated and presented in Table 4.11. This Table demonstrates the position of each gene and to carefully select the genes that will be considered for further investigations.

Table 4.11. The SUM of rank for commonalities

SUM of rank	Gene Symbol
1	<i>c2orf42</i>
2	<i>PXDC1</i>
2	<i>VSTM2B</i>
3	<i>WBSCR27</i>
3	<i>SERPINB4</i>
3	<i>NPRL3</i>
4	<i>DNAJB8-AS1</i>
4	<i>ZNF2</i>
4	<i>LRTM2</i>
5	<i>epha6</i>
6	<i>TFDP2</i>
7	<i>ENST00000412501</i>
8	<i>ZSCAN16</i>
8	<i>DHFR2</i>
9	<i>OXSRI</i>
10	<i>BLOC1S3</i>
11	<i>NLRP1</i>
12	<i>FOXD4L5</i>
13	<i>LOC100130176</i>
14	<i>HIST2H3C</i>
15	<i>HGF</i>
15	<i>GABARAPL1</i>
17	<i>HACD4</i>
17	<i>SOWAHA</i>
18	<i>DOPEY2</i>
18	<i>EMP1</i>
19	<i>ARMCS</i>
20	<i>YBEY</i>
21	<i>BTN3A2</i>
27	<i>PCGF3</i>
35	<i>PFDN2</i>
54	<i>SLCO4A1</i>
63	<i>ANKRD1</i>
75	<i>CARD6</i>
75	<i>VDAC2</i>
78	<i>THAP12</i>
80	<i>ANKRD44</i>
84	<i>CD59</i>
87	<i>PTK2B</i>
92	<i>AEN</i>
102	<i>PLEKHS1</i>
103	<i>DICER1</i>
107	<i>IKZF2</i>
117	<i>NOD2</i>
121	<i>PRR5L</i>
172	<i>HEXIM1</i>
174	<i>SSH1</i>
183	<i>PRDX1</i>
197	<i>DDX46</i>
207	<i>SOCS4</i>
223	<i>GAB2</i>
238	<i>RNF114</i>
241	<i>ZNF226</i>
244	<i>FAM217B</i>
245	<i>IRF-9</i>
277	<i>PCDHB2</i>
370	<i>SAMD4A</i>
395	<i>PRLR</i>
478	<i>Icmt2</i>
551	<i>HAUS6</i>
563	<i>THC2652746</i>
575	<i>C17orf97</i>
669	<i>BEND3</i>
677	<i>EXOC7</i>
698	<i>ANGEL2</i>
983	<i>BTN3A1</i>
1100	<i>SPATA13</i>

Ranking of 67 genes out of 34127 genes in both different analysis tests (Regression and Fold Change) for both BASAL-LIKE (MDA-MB-468) and HER2 (SkBr3) using different concentrations of Metformin.

Commonalities between the different analyses (Regression-based method and fold change-based method) for both cell lines BASAL-LIKE (MDA-MB-468) and HER2 (SkBr3) cell lines and for different concentrations (1 and 4 mL) of Metformin shown in Appendix Figure A.4. 4.

The outcomes of the second approach that was previously applied, nominated 16 genes to be a subject of investigation and validation. An in-depth investigation took place for those 16 genes via different databases such as PANTHER- GENE LIST ANALYSIS, DAVID Bioinformatics Resources, GEO Profiles Results-NCBI, Gene MANIA, KEGG: Kyoto Encyclopedia of Genes and Genomes, and Gene Cards suite (Tables 4.12. and 4.13.).

Table 4.12. List of genes that were selected for validation by qRT-PCR

Input ID	Cell line	Concentration	Statistic test	Regulation
<i>BTN3A2</i>	MDA-MB-468	1mM	Regression test	Down
<i>EMP1</i>	MDA-MB-468	1mM	Fold Change	Down
<i>HGF</i>	MDA-MB-468	1mM	Regression test	Down
<i>NPRL3</i>	MDA-MB-468	1mM	Regression test	Down
<i>OXSRI</i>	MDA-MB-469	1mM	Regression test	Up
<i>PRDX1</i>	MDA-MB-468	1mM	Fold Change	Up
<i>ANKRD44</i>	MDA-MB-468	4mM	Fold Change	Up
<i>PCDHB2</i>	MDA-MB-468	4mM	Fold Change	Down
<i>PXDC1</i>	MDA-MB-468	4mM	Regression test	Down
<i>SERPINB4</i>	MDA-MB-468	4mM	Regression test	Up
<i>WBCSR27</i>	MDA-MB-468	4mM	Regression test	Up

BASAL-LIKE (MDA-MB-468) cell line upregulated and downregulated genes after treatment with Metformin out of Regression test and Fold Change test. Red colour indicated upregulated genes and blue colour downregulated genes.

Table 4.13. List of genes that were selected for validation by qRT-PCR

Input ID	Cell line	Concentration	Statistic test	Regulation
<i>DHFR2</i>	SkBr3	1mM	Fold Change	Down
<i>PTK2B</i>	SkBr3	1mM	Fold Change	Up
<i>VSTM2B</i>	SkBr3	1mM	Regression test	Down
<i>IRF-9</i>	SkBr3	4mM	Fold Change	Up
<i>ANKRD44</i>	SkBr3	4mM	Regression test	Down
<i>C2orf42</i>	SkBr3	4mM	Fold Change	Down

HER2 (SkBr3) cell line upregulated and downregulated genes after treatment with Metformin and using Regression test and Fold Change test. Red colour indicated upregulated genes and blue colour downregulated genes.

Those sixteen genes (*ANKRD44*, *BTN3A1*, *C2orf42*, *DHFR2*, *EMP1*, *HGF*, *IRF-9*, *NPRL3*, *OXSRI*, *PCDHB2*, *PRDX1*, *PTK2B*, *PXDC1*, *SERPINB4*, *VSTM2B*, and *WBSCR27*) were chosen to quantify their gene expression statuses by Real-time quantitative PCR (qRT-PCR). Further validation involved immunoblotting and immunofluorescence (IF) to determine their expression at the protein level.

4.6. Pathway analysis

Upregulated and downregulated genes that were associated with BASAL-LIKE (MDA-MB-468) and HER2 (SkBr3) cell lines were investigated for their association with other molecular pathways using the Panther online databases (*Mi et al., 2013*). The main upregulated pathways in BASAL-LIKE (MDA-MB-468) cell line was associated with apoptosis signalling pathway, cytoskeletal regulation by Rho GTPase, gonadotropin-releasing hormone receptor pathway and the ubiquitin-proteasome pathway; with each representing (13 %) of each pathway. The down-regulated pathways were related to Wnt signalling, EGF receptor signalling, Cadherin signalling, and PDGF signalling pathways (12%). The pathways were represented as a percentage of contribution in Figure 4.10.

In HER2 (SkBr3) cell line, upregulated genes were associated with integrin signalling pathway (12%), CCKR signalling (12%), inflammation mediated by chemokine and cytokine signalling pathway (13%), and Gonadotropin-releasing hormone receptor pathway (13%). The down-regulated pathways were involved in formyltetrahydrofolate biosynthesis, tetrahydrofolate biosynthesis, inflammation mediated by chemokine and cytokine signalling, Gonadotropin-releasing hormone receptor, and Interferon-gamma signalling pathways as 10 %, as displayed in Figure 4.11. Gonadotropin-releasing hormone receptor pathway was found as a common pathway between the up- and down-regulated pathways in SkBr3 and also with the up-regulated pathways in MDA-MB-468. Whilst, the inflammation mediated by chemokine and cytokine signalling pathway was found as a common pathway in both up-regulated and down-regulated pathways in SkBr3 cell line.

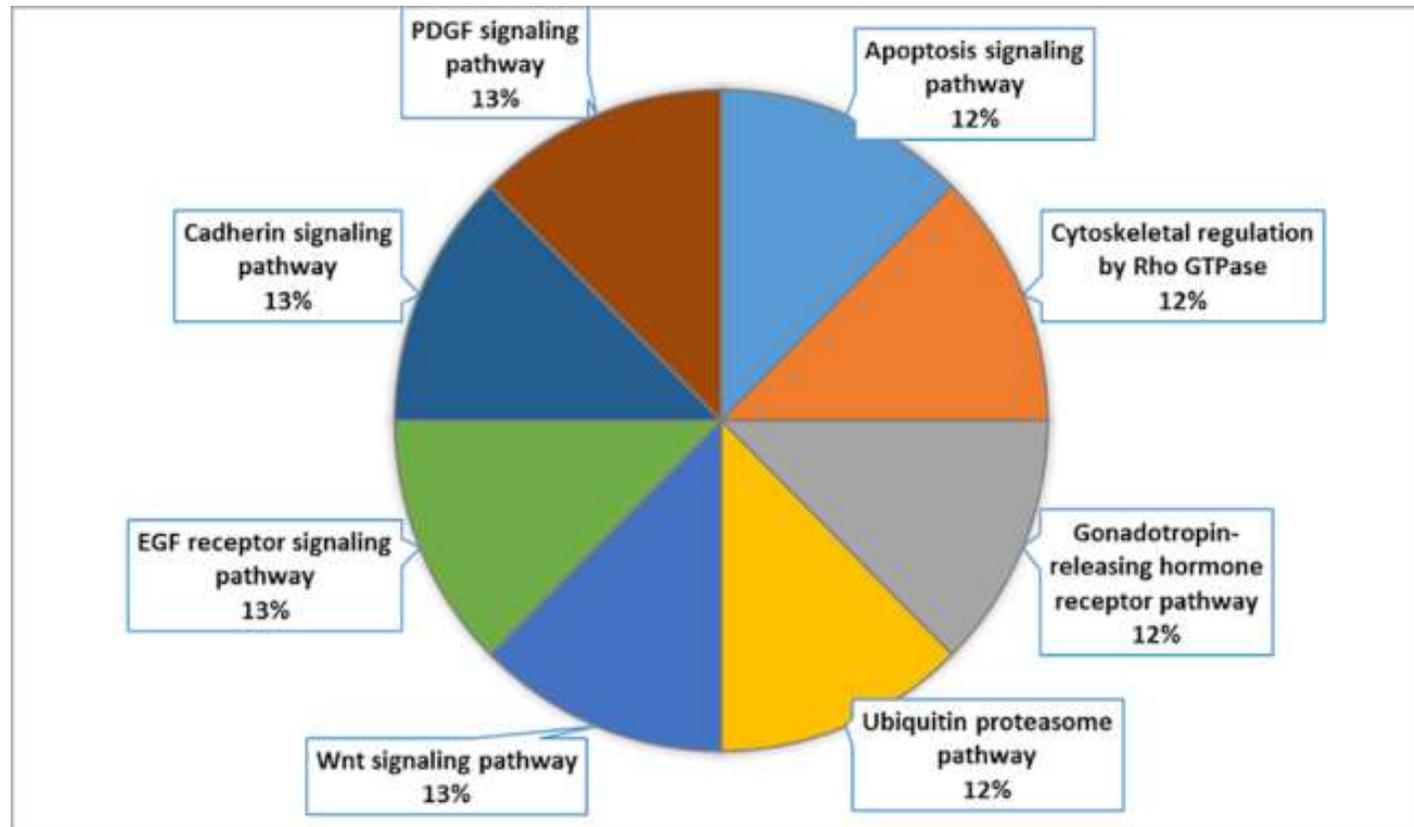


Figure 4.10. Pathway ontology of the consensus BASAL-LIKE (MDA-MB-468) cell line associated genes. Pie chart representing the percentage of contribution of each gene to Panther GO. The genes are assigned to the different molecular and biological pathways based on their functions in developmental process, cellular process, metabolic process, cellular component organisation or biogenesis, and localization, either directly or indirectly. This diagram was constructed using the online pathway databases, Panther (*Mi et al., 2013*).

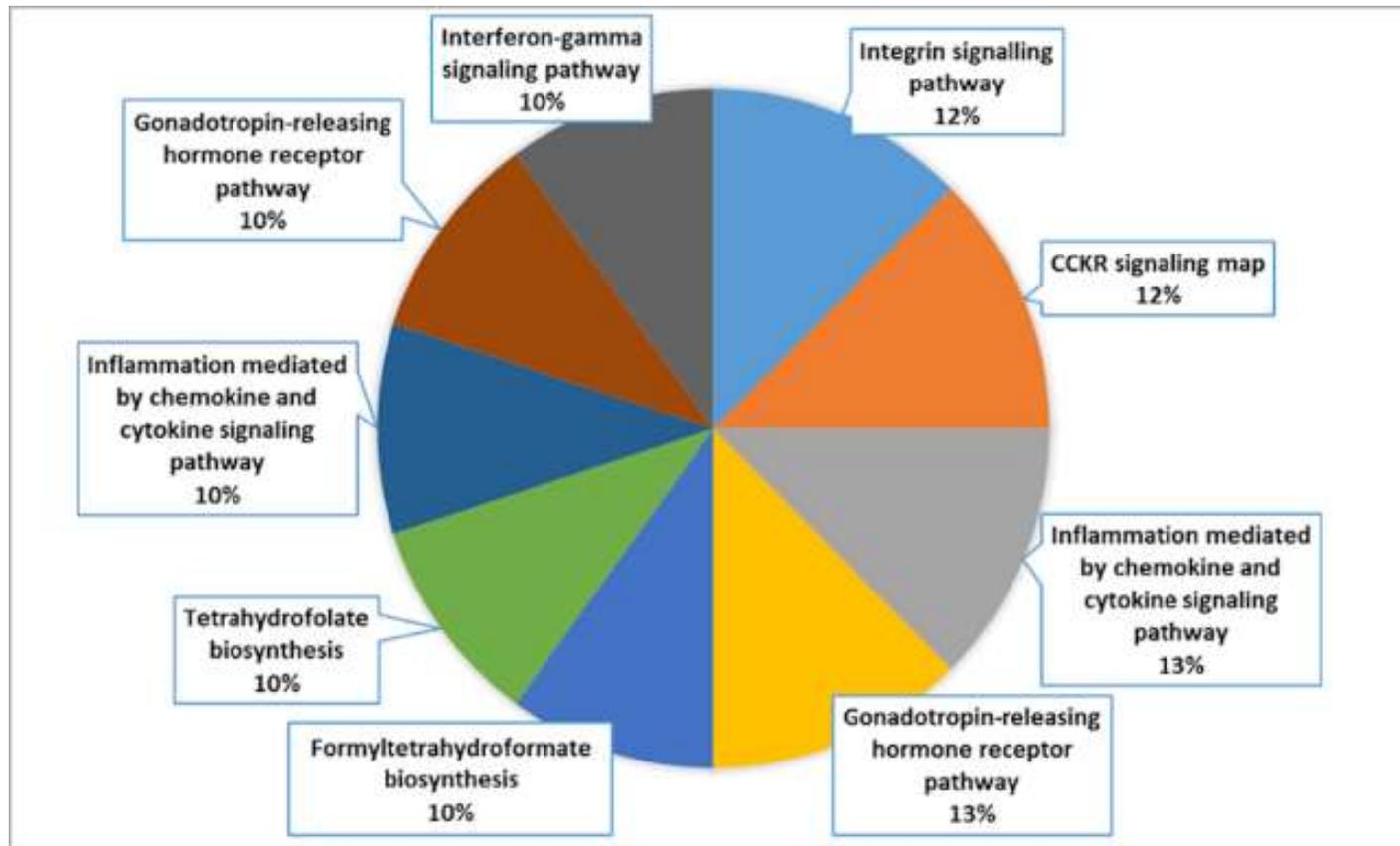


Figure 4.11. Pathway ontology of the consensus HER2 (SkBr3) cell line associated genes. Pie chart representing the percentage of contribution of each gene to Panther GO. The genes are assigned to the different molecular and biological pathways based on their functions in developmental process, cellular process, multicellular organismal process, metabolic process, biological regulation, and localization, either directly or indirectly. This diagram was constructed using the online pathway databases, Panther (Mi et al., 2013).

4.7. Discussion

The main aims of this chapter were to investigate gene expression between BASAL-LIKE and HER2 breast cancer cell subtypes following treatment with Metformin. These allowed finding gene expression commonalities between the two breast cancer phenotypes. Furthermore, up-regulated genes may be associated with the resistance to Metformin treatment that was previously observed with the HER2 breast cancer phenotype (SkBr3 cell line).

Interestingly, it has been noticed that, on a comparison of the top 100 interaction maps of the BASAL-LIKE and HER2 breast cancer subtypes populations, identical genes were not immediately identified as influential hubs. Although, from the driver analysis it was easy to notice the most influential influencers (source and target hubs) of the system.

It was also noticed that in Basal-Like (MDA-MB-468) and HER2 (SkBr3) cell lines, the hubs were found to be “globally” negatively regulated and no positive centres were identified. Comparatively, the overall interaction feedback loops of BASAL-LIKE cell line were negative with few positive loops. However, the HER2 breast cancer subtype cell line interactions map showed several positive feedback loops between some molecules.

The Artificial Neural Network (ANN) stepwise analysis-based approach has been applied to a total of 36 Microarray samples. Besides, the interactome Network Map has also been utilised for both Basal-Like (MDA-MB-468) and HER2 (SkBr3) cell lines simultaneously and then individually. The random segregation of samples into three different groups; 60% for training, 20% for validation and 20% for testing, before the ANN training, required a large number of samples, to compare a small number of samples each cell line individually and distinguish between the two concentrations as well separately another approach has been implemented. At this end, the Regression-based method and fold change- based method was utilised.

The second approach revealed a cluster of information for each breast cancer subtype independently. Data collected from Microarrays have provided evidence of differential gene expression. The outcomes of these methods nominated sixteen genes that are *ANKRD44*, *BTN3A1*, *C2orf42*, *DHFR2*, *EMP1*, *HGF*, *IRF-9*, *NPRL3*, *OXSRI*, *PCDHB2*, *PRDX1*, *PTK2B*, *PXDC1*, *SERPINB4*, *VSTM2B*, and *WBSCR27*, and that were selected for quantification by real-time quantitative PCR (qRT-PCR) next chapter. Those genes have been selected according to in-depth investigative annotation *via* different databases including PANTHER- GENE LIST ANALYSIS, DAVID Bioinformatics Resources, GEO Profiles Results-NCBI, Gene MANIA, KEGG: Kyoto Encyclopedia of Genes and Genomes, and Gene Cards suite. The results highlighted their influence on biological functions. Also, they had highly significant P values assessed by both

statistical methods and were mostly up-regulated genes. The Panther online databases (*Mi et al., 2013*) has been used to investigate the differentially expressed genes that were obtained from both Basal-Like (MDA-MB-468) and HER2 (SkBr3) cell lines and out of fold change approach to determine their associations with other molecular pathways. This analysis highlighted a few combined pathways such as gonadotropin-releasing hormone receptor pathway, as a common pathway in both Basal-Like (MDA-MB-468) cell line-up-regulated pathways, and HER2 (SkBr3) cell line, in both up and downregulated channels. Indeed, apoptosis signalling pathway was found as up-regulated in Basal-Like (MDA-MB-468), confirming previous data in chapter 3 Figures (3.6, 3.9, and 3.10). Likewise, the Wnt signalling pathway was downregulated in MDA-MB-468 cells. This pathway is a critical pathway in the regulation of cell migration, cell polarity, neural patterning and organogenesis during fetal development. It is also involved in tumour initiation, and disease retrogression, by which retaining the inherent drug-resistance features. However, in HER2 (SkBr3) cell line, four critical upregulated pathways were found: integrin signalling pathway, CCKR signalling, inflammation mediated by chemokine and cytokine signalling pathway, and Gonadotropin-releasing hormone receptor pathway. Interestingly, *Protein Tyrosine Kinase 2 Beta (PTK2B)* has been found as signalling through these pathways. *PTK2B* is also known as *PYK2* or *FAK2*, which is a non-receptor tyrosine kinase that has been detected as an upregulated gene in the HER2 cell line after treatment with Metformin. Its encoded cytoplasmic protein tyrosine kinase is involved in calcium-induced regulation of ion channels and activation of the map kinase signalling pathway. This protein-tyrosine kinase also regulates reorganisation of the actin cytoskeleton, cell growth, proliferation, survival, migration, invasion and bone remodelling. Also, *PYK2* is implicated in the regulation of cell migration through its dual action in the integrin signalling pathway, and Gonadotropin-releasing Hormone (GnRH) transmitting pathway receptor(Gaq), (*Naor et al., 2000; Kraus and Naor, 2001; Huvneers and Danen, 2009*). Besides, *PYK2* is mediating the inflammation by integrating growth factor and cytokine receptors signalling (*Selitrennik and Lev, 2015*). Moreover, *PYK2* has a contributory role in motility and migration through cholecystokinin /gastrin receptor signalling pathway (CCKR), (*Smith et al., 2016*). While the Interferon-gamma signalling pathway was downregulated in HER2 (SkBr3) cell line, it is also found that *PYK2* is critical for the Jak-mediated MAPK and Stat1 activation by IFN-gamma (*Takaoka et al., 1999*). Furthermore, integral of signalling pathways initiated by receptor tyrosine kinases and integrins, are essential for growth-factor intermediated the biological responses.

Up-regulated pathways including integrin signalling pathway, CCKR signalling, Inflammation mediated by chemokine and cytokine signalling pathway, and Gonadotropin-releasing hormone receptor pathway that involved *PYK2* in their signals were proved to be associated with cancer

progression, drug resistance, and cell survival. HER2(SkBr3) cells Metformin resistance is likely to be related to these up-regulated pathways.

However, the down-regulated pathways were related to cell cycle arrest and apoptosis could explain the high proportion of dead cells in Basal-like cell line.

Collectively, the objectives of this part were to screen the differentiation in gene expression pattern between both Basal-like and HER2 breast cancer phenotypes, in addition to detecting highly-expressed genes, which might be associated with the resistance to Metformin treatment in the HER2 subtype. This has been attained, and *PYK2* has been nominated as a target gene. However, the validation of gene expression will be assessed by qRT-PCR, Western Blot and Immunofluorescence analysis. As well, the NanoString technique will be applied on Basal-like and HER2 phenotypes samples for pan-cancer pathways detection, in the next chapter (chapter 5).

CHAPTER 5

Confirmation of Gene Expression Microarray data by Real-time quantitative PCR (qRT-PCR), Immunoblotting and Immunofluorescence analyses

5.1. Introduction

Real-time quantitative PCR has become one of the most powerful technologies, especially in biologically oriented laboratories. It is a sensitive and reliable method to detect and quantify nucleic acids (DNA and RNA (cDNA) levels. It relies on the detection and quantification of fluorescence emitted from a reporter molecule at a real time. This detection occurs during the accumulation of the PCR output with each cycle of amplification, thus allowing the monitoring of the PCR reaction during the early and exponential phase, where the first significant increase in the amount of the PCR product correlates to the initial amount of target template.

There are two different methods to present quantitative gene expression: absolute and relative quantification. The absolute quantification usually calculates the copy number of the gene by relating the PCR signal to a standard curve, while the relative gene expression presents the data of the gene of interest relative to some calibrator or internal control gene.

The comparative CT method (relative quantification) also referred to as the $2^{-\Delta\Delta Ct}$ method is a very popular technique that compares results from experimental samples with both a calibrator (Untreated or wild-type sample) and a Normalizer (housekeeping gene). In this method, which used for this study, Ct values for the gene of interest in both test and calibrator samples were adjusted in relation to a normalizer gene Ct from the same two samples. The generated $\Delta\Delta Ct$ value was combined to determine the fold change difference in gene expression (*Schmittgen1 and Livak, 2008*).

The aim of this chapter was to investigate the expression of *ANKRD44*, *BTN3A1*, *C2orf42*, *DHFR2*, *EMP1*, *HGF*, *IRF-9*, *NPRL3*, *OXSRI*, *PCDHB2*, *PRDX1*, *PTK2B*, *PXDC1*, *SERPINB4*, *VSTM2B*, and *WBSCR27* in Basal-like and HER2 phenotypes to confirm the Gene Expression Microarray data (chapter 4). In addition, this work allowed the selection of a marker for further studies.

5.2. Assessment of efficiency, sensitivity and reproducibility using a standard curve

PCR amplification efficiency is the most consistent factor in a real-time PCR reaction. Reaction efficiency can be assessed by creating a standard curve. This curve was plotted by generating a 5-fold serial dilution of nucleic acid samples and performing real-time PCR. The obtained results were plotted with input nucleic acid quantity on the x-axis and Ct on the y-axis. The slope of the curve was used to determine the reaction efficiency, which should be between 90% and 110% (Figure 5.1). In addition, a good primer should have a length of about 18-22 bases, a melting temperature (T_m) of below 65°C, and a GC-content between 40-60% (Schmittgen and Livak, 2008).

5.3. Real-time quantitative PCR (qRT-PCR) of investigated genes

The nominated 16 genes (*ANKRD44*, *BTN3A1*, *C2orf42*, *DHFR2*, *EMP1*, *HGF*, *IRF-9*, *NPRL3*, *OXSRI*, *PCDHB2*, *PRDX1*, *PTK2B*, *PXDC1*, *SERPINB4*, *VSTM2B*, and *WBSCR27*) have been a subject of investigation and validation through qRT-PCR. First, the standard curve was performed to assess the amplification efficacy of all primers used for this assay. Table 5. 1 shows the percentage of efficiency for the 16 primers and the housekeeping gene *Glucuronidase Beta (GUSB)*. The percentage of efficiency has previously calculated for the *GUSB* housekeeping gene on threshold level set on 0.1324, which gave the best efficiency. While the threshold was set on 0.164 for the other primers. The threshold level is the signal level that statistically reflects a significant increase over the calculated baseline (threshold should be higher than the baseline, which is higher than the noise level). The setting of threshold allows distinguishing the relevant amplification background signal from real signals. The point where the reaction curve intersects with the threshold line is known as Ct, or "threshold cycle", which shows the number of cycles that were taken to detect a real signal from the samples.

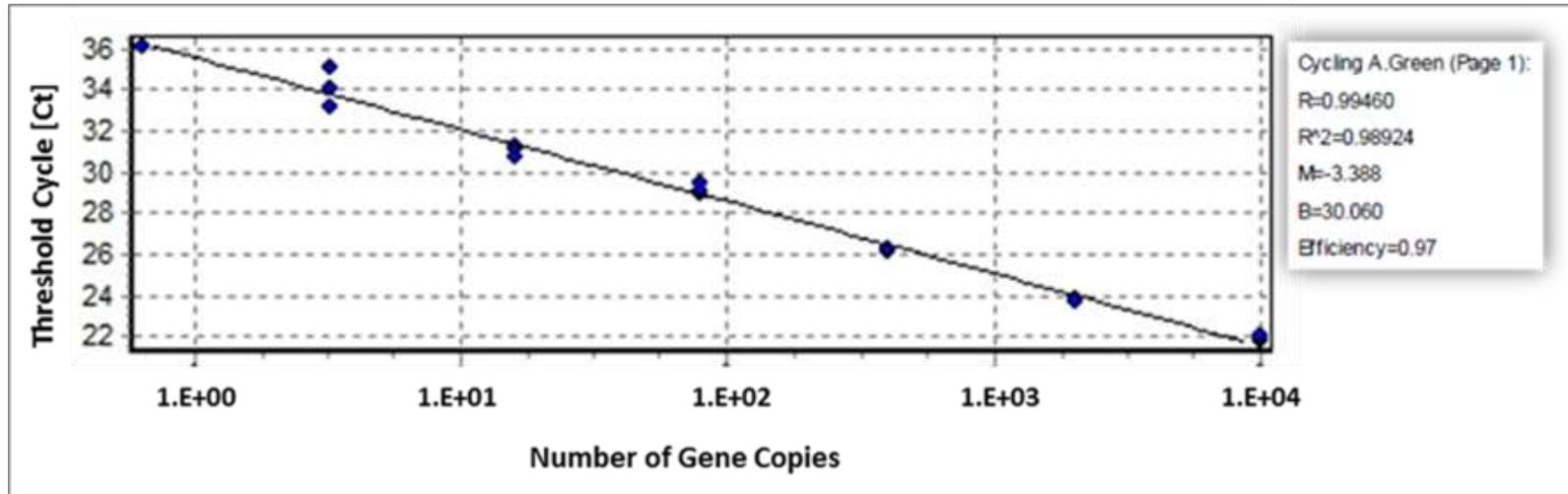


Figure 5.1. Schematic representation of the standard curve for primers efficiency test. The X-axis represented the number of copies for each gene, while the Y-axis represented the Threshold Cycle (Ct). The slope of the curve determined the reaction efficiency, which should be between 90% and 110%. This curve was plotted by creating a dilution series of cDNA template samples according to this order: SkBr3 control, ½ Ec50, 2Ec50, and MDA-468 control, ½ Ec50, 2Ec50; and performing real-time PCR.

Table 5.1. Percentage of efficiency and amplification dynamic range

Oligo Name	Sequence: (5' to 3')	Tm°	Efficiency
FH1_ANKRD44	GTGCTTTGTGTCTTGAATTG	55° C	1.07
RH1_ANKRD44	TCAATTTACCTCCATTCTG		
FH1_BTN3A1	AGTCTCTGCTTTCTTTTCC	55° C	1.1
RH1_BTN3A1	AAATCACCTCTTAGCATTCC		
FH1_C2orf42	AGCTTTTGTTCGAAAGATG	56° C	0.97
RH1_C2orf42	GCATCTCTGGGGTATCTAAG		
FH1_DHFR2	CGCTGTGTCCAAAACATGG	56° C	1.09
RH1_DHFR2	GAATTCATTCTGAGCGGCG		
FH1_EMP1	CTTCAGAACTCTCTTGCTC	54° C	1.07
RH1_EMP1	CATAATAACAGTAGCGATGTGG		
FH1_HGF	CAAGGACCTACGAGAAAATTAC	58 AND 58° C	NOT GOOD
RH1_HGF	ATCACAGTTTGAATTTGGG		
FH1_IRF-9	CTCAGAAAGTACCATCAAAGC	58° C	1
RH1_IRF-9	TCATTATTGAGGGAGTCCTG		
FH1_NPRL3	CGATTCCAGTTTTTCAGATG	54° C	0.91
RH1_NPRL3	CAGGTTATGCAGACAGTTTATC		
FH1_OXSR1	AGGTTTTAATGCTGACACTG	54° C	1
RH1_OXSR1	TTCTGGATCTTTTTGAAGGC		
FH1_PCDHB2	ACCCAGATACTCAGCTAAAGAAGC	58° C	0.95
RH1_PCDHB2	TCTCTACCAAACGTGCTCC		
FH1_PRDX1	GGGTCAATACACCTAAGAAAC	54° C	1.02
RH1_PRDX1	CTTCATCAGCCTTTAAGACC		
FH1_PTK2B	AATGCACTTGACAAGAAGTC	54° C	1.05
RH1_PTK2B	GCTTTAAGTTCTCCTGCATC		
FH1_PXDC1	AGAAATCATGAGGTCCAATG	58° C	0.98
RH1_PXDC1	AATAAATGCTCTGTTGGGTC		
FH1_SERPINB4	CATGTTGATAGGTCAGGAAATG	55° C	1.09
RH1_SERPINB4	ATTGATACGTCTTTTCTCCG		
FH1_VSTM2B	ACAGAAGTCCCAAAGATG	55 AND 57° C	NOT GOOD
RH1_VSTM2B	ATCCTTATTTGTTACCTTGCTC		
FH1_WBSCR27	CCGGACTACGACCAGGATGTG	58 AND 60° C	NOT GOOD
RH1_WBSCR27	GCTTGTGTGAGGCAGTCCAC		
FH1_GUSB	CTC ATT TGG AAT TTT GCC GAT T	58° C	0.99
RH1_GUSB	CCG AGT GAA GAT CCC CTT TTT A		

Percentage of efficiency for the 16 primers, and the housekeeping gene *GUSB*. The Table displays the Oligo name and its Sequence (5' to 3'), melting temperature (Tm°), and the Efficiency of each primer. The results showed that three of these primers (*HGF*, *VSTM2B*, and *WBSCR27*) were not efficient for further study. While the other thirteen primers were efficient and used for the next study. The PCR amplification efficiencies were between 90%-110%.

The efficiency test results showed that three of these primers (*HGF*, *VSTM2B*, and *WBSCR27*) are not efficient for further study. The other thirteen primers were efficient and thus selected for the next study. QRT-PCR was performed for those thirteen primers and with the housekeeping gene *GUSB*. This experiment was achieved according to the protocol in section 2.2.4.2 using different sets of cDNA template samples, and according to this order: SkBr3 control, 1mM, 4mM, and MDA-468 control, 1mM, 4mM. The threshold was set on 0.1642 for all qRT-PCR runs and primers, while the T_m° varied according to the different primers. This experiment was independently repeated three times, and the run was performed in triplicates for each sample. The expression profile for each gene is presented in Figures 5.2 and 5.3 for both Basal-like (MDA-MB-468) and HER2 (SkBr3) cell lines, respectively.

In this study, the comparative threshold method (Relative quantification) has been used to quantify gene expression in qRT-PCR, which provide gene expression values as fold change in comparison with reference samples or calibrators and assuming that PCR efficiencies of all investigated samples are the same. The Fold change for each sample was calculated according to this equation: ***Fold change due to treatment = $2^{-\Delta\Delta CT}$*** .

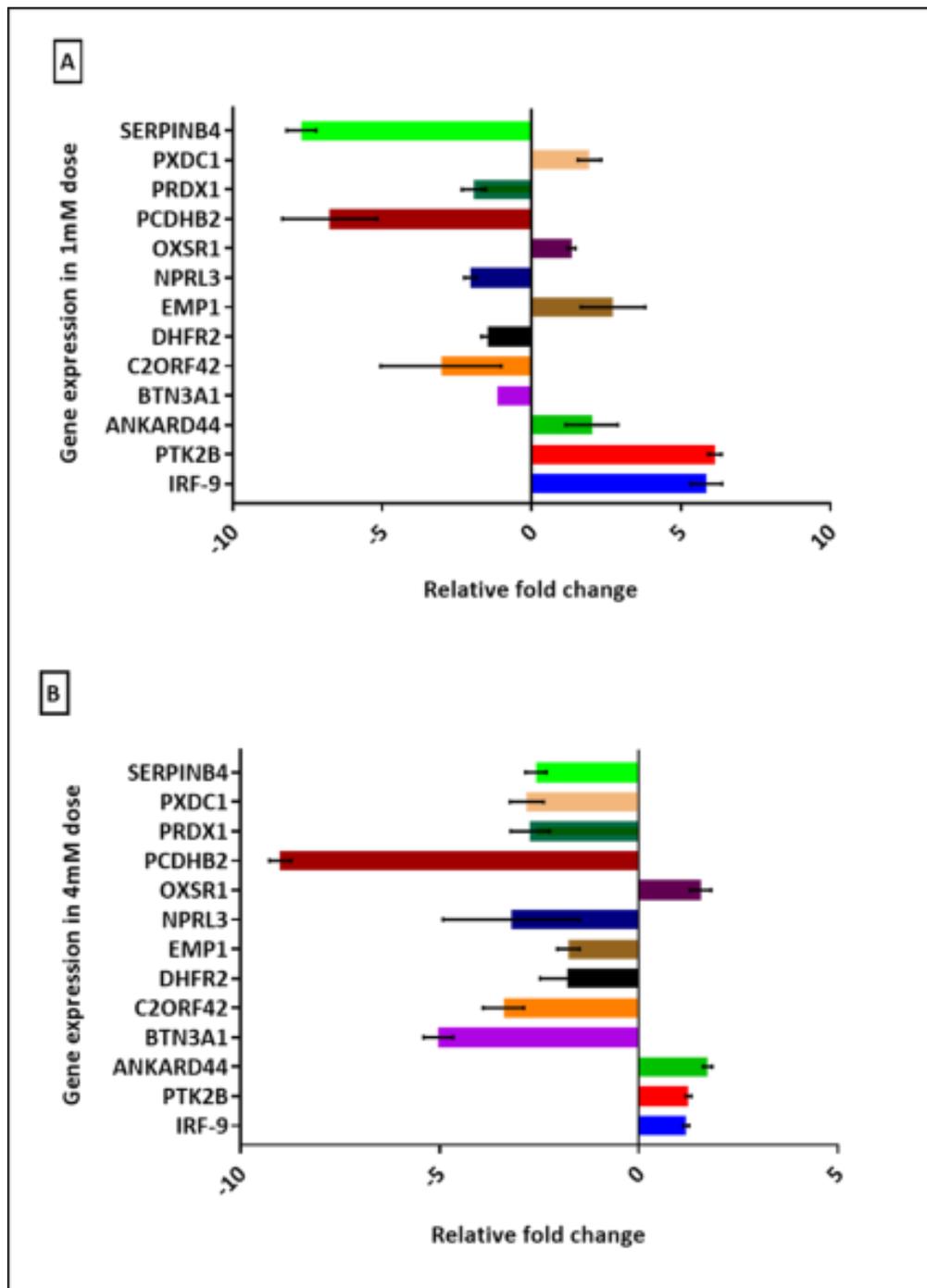


Figure 5.2. QRT-PCR data for Basal-like (MDA-MB-468) cell line representing up and downregulated genes, using different doses (A corresponds to 1mM and B to 4mM of Metformin treatments and when compared to controls. The threshold was set on 0.1642, while the T_m° varied according to the different primers. This experiment was independently repeated three times and in triplicates for each sample and following this order: SkBr3 control, 1mM, 4mM, and MDA-468 control, 1mM, 4mM. In addition, the Fold change for each sample was calculated according to this equation: Fold change due to treatment = $2^{-\Delta\Delta CT}$.

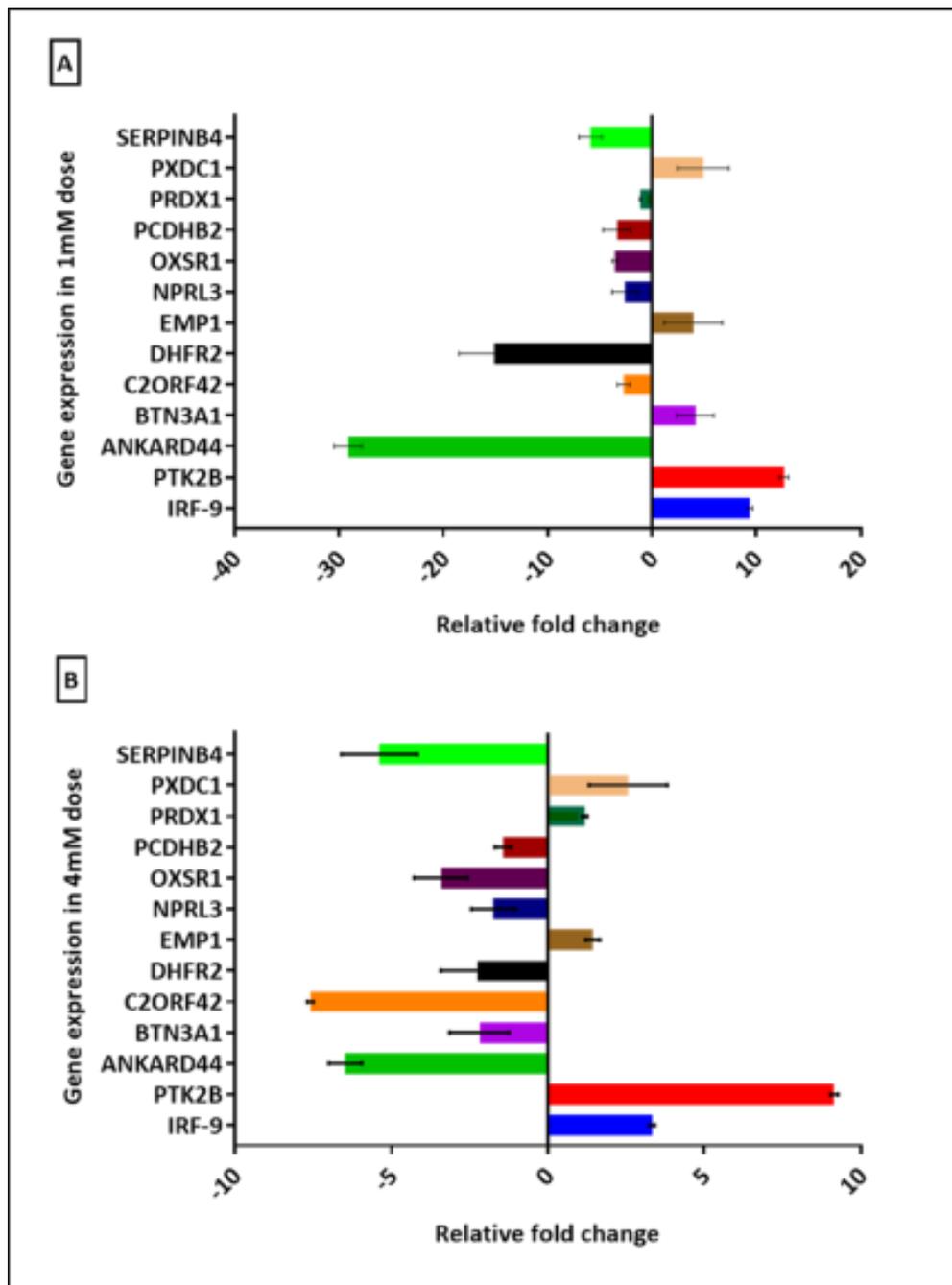


Figure 5.3. QRT-PCR data for HER2 (SkBr3) cell line representing up and downregulated genes, using different doses (A corresponds to 1mM and B to 4mM of Metformin treatments and when compared to controls. The threshold was set on 0.1642, while the T_m varied according to the different primers. This experiment was independently repeated three times and in triplicates for each sample and following this order: SkBr3 control, 1mM, 4mM, and MDA-468 control, 1mM, 4mM. In addition, the Fold change for each sample was calculated according to this equation: Fold change due to treatment = $2^{-\Delta\Delta CT}$.

The gene expression performance that was determined by qRT-PCR data analysis confirmed previous gene expression profiling that was established by gene expression microarray analysis. However, three genes (*SERPINB4*, *EMP1*, and *PRDX1*) displayed opposite expression by qRT-PCR. These genes were detected in the Basal-like cell line (MDA-MB-468) in a Fold Change-based method (*EMP1*, *PRDX1*) and regression-based method (*SERPINB4*).

After confirming gene expression profiles, it was essential to determine their expression at the protein level. Consequently, two upregulated genes (*IRF-9* and *PTK2B*) and one downregulated gene (*SERPINB4*) were selected for western blotting and Immunofluorescence studies.

The **Interferon Regulatory Factor 9 (IRF-9)**, which is also known as **p48**, **ISGF3** or **ISGF3G**, is an interferon regulatory factor that mediates cell signalling by type I IFNs (IFN- α and IFN- β). *IRF-9* is also a specific transcription factor found downstream the JAK/STAT pathway, which plays a role in body immune balance (Zhao et al., 2017).

Protein Tyrosine Kinase 2 Beta (PTK2B) or PYK2, is a non-receptor protein tyrosine kinase that regulates reorganisation of the actin cytoskeleton, cell polarisation, migration, adhesion and spreading, and is also involved in bone remodelling (Selitrennik and Lev, 2015).

Moreover, **Serpin Family B Member 4 (SerpinB4) or SCCA2 (Squamous cell carcinoma antigen2)**, is a member of the ovalbumin family of serine proteinase inhibitors. Originally, it was discovered as a tumour-specific antigen and a tumour marker for various kinds of squamous cell carcinomas (de Koning et al., 2011).

The above-mentioned genes have been selected according to two factors: first, they showed significantly high P values in both statistical methods, which represented 5.61E+01 for *IRF-9*, 2.85E-01 for *PTK2B*, and 0.0163 for *SERPINB4*; secondly, their association with biologically important functions. However, *SERPINB4* demonstrated an opposite expression in (qRT-PCR) and it has been found as down-regulated in an MDA-MB-468 cell line in ANN stepwise analysis approach (see chapter 4 section 4.5.3).

5.4. Protein expression in Basal-like and HER2 breast cancer cell lines

For western blotting analyses, protein lysates of Basal-like (MDA-MB-468), Claudin -Low (MDA-MB-231) and HER2 (SkBr3, and MDA-MB-453) cell were used at 30 μ g (per well). Extracts from two cell lines have been used for each phenotype and for more reliability. The data showed higher expression of IRF-9 in 1mM treatments for all cell lines and when compared with untreated and 4mM treatments (Figure 5.4).

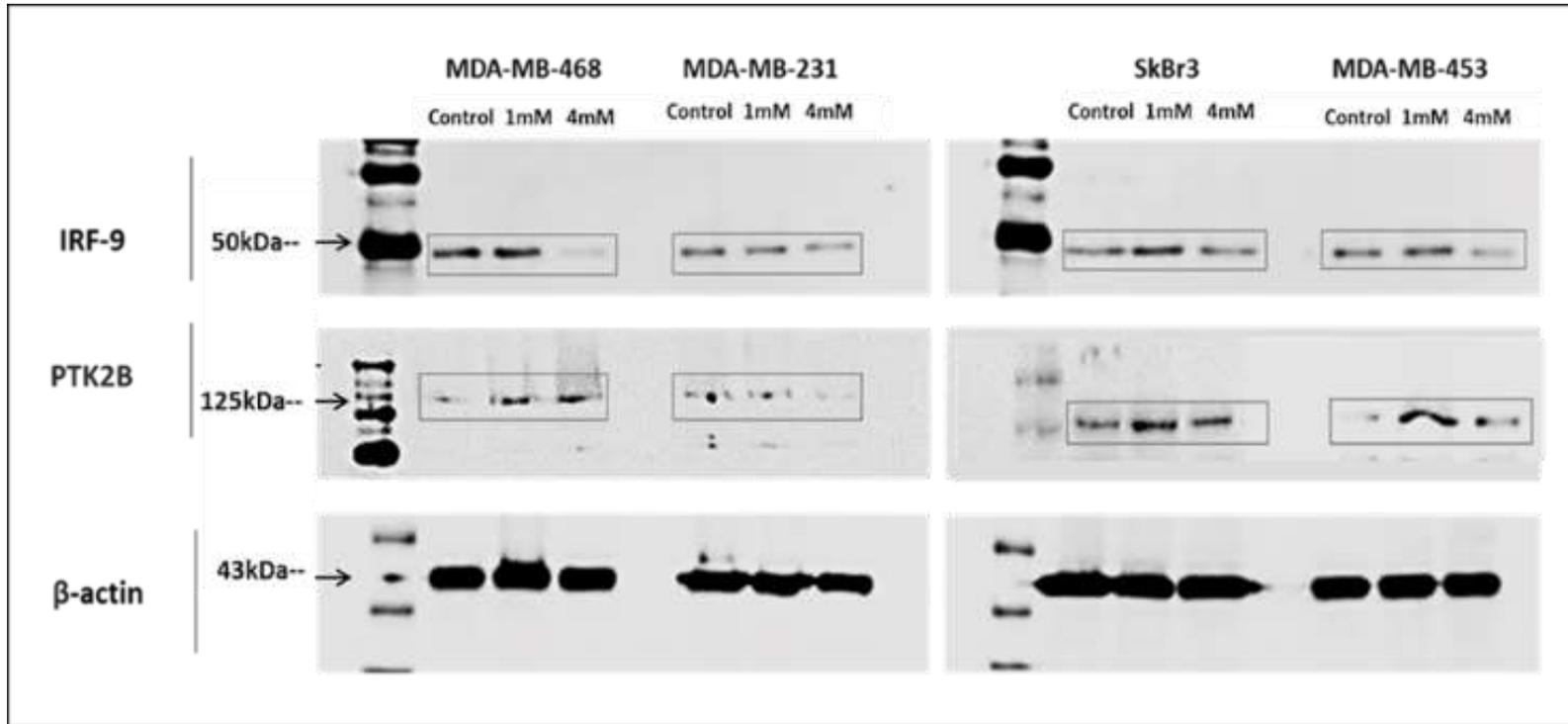


Figure 5.4. Representative micrographs of immunoblots showing IRF-9 and PTK2B expression in MDA-MB-468 and MDA-MB-231 (left panel), and SkBr3 and MDA-MB-453 (right panel). β -actin is used as a loading control. Cell extracts for immunoblotting were obtained from untreated and Metformin treated cells with 1mM and 4mM concentrations.

PTK2B expression was the higher expression, which shown in 1mM Metformin treatments in all cell lines when compared with control and 4mM. However, the expression of PTK2B was higher in 1mM dosage in SkBr3 cell line than the other cell lines.

Additionally, SERPINB4 showed negative expression in 1mM and 4mM Metformin doses in both MDA-MB-468 and MDA-MB-231 populations. Moreover, SKBR3 and MDA-MB-453 cell lines shown a low expression of SERPINB4 in both doses as demonstrated in Appendix (Figure A.5.1).

Moreover, ImageJ (v1.50c) software was used to calculate the density of each sample band as demonstrated by measuring the densitometric ratio in Arbitrary Unit (AU) for IRF-9 and PTK2B compared to β -actin expression in untreated Basal-like, Claudin-Low, and HER2 cell lines, and after treatment with Metformin 1 and 4mM Figure 5.5 and Figure 5.6. SERPINB4 densitometric ratio was shown in Appendix Figure A.5.2. These findings confirmed the previous results from qRT-PCR data analysis. The Arbitrary Unit (AU) is a relative unit of measurement to show the ratio of the amount of substance, intensity, or other quantities, to a predetermined reference measurement.

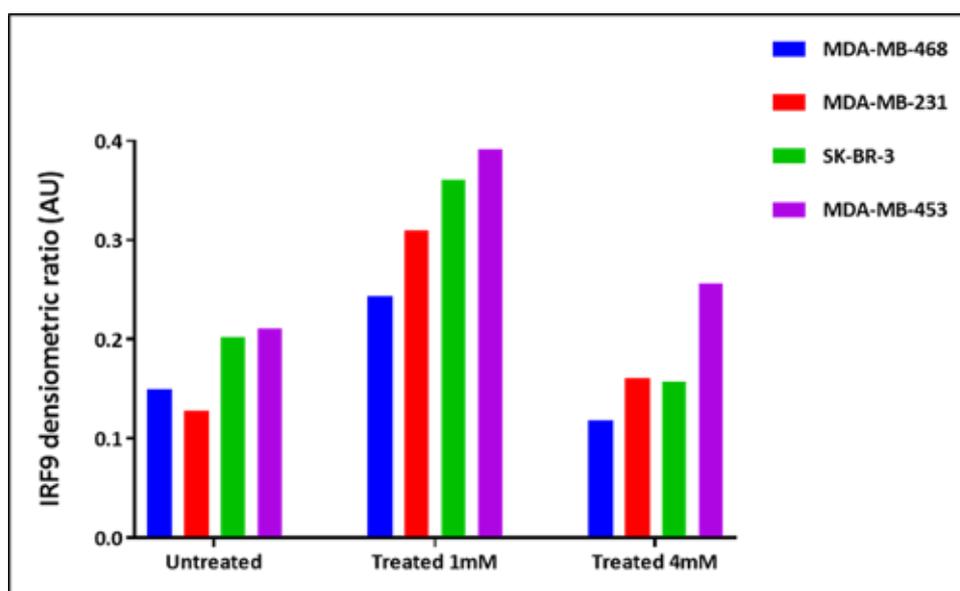


Figure 5.5. Micrograph representing the densitometry ratio (AU) of IRF-9 protein expression in Metformin untreated and treated (1mM and 4mM) in Basal-like (MDA-MB-468), Claudin-Low (MDA-MB-231), HER2 (SkBr3) and (MDA-MB-453) breast cancer cell lines(n=1). Higher protein expression level was observed in 1mm Metformin-treated cell lines.

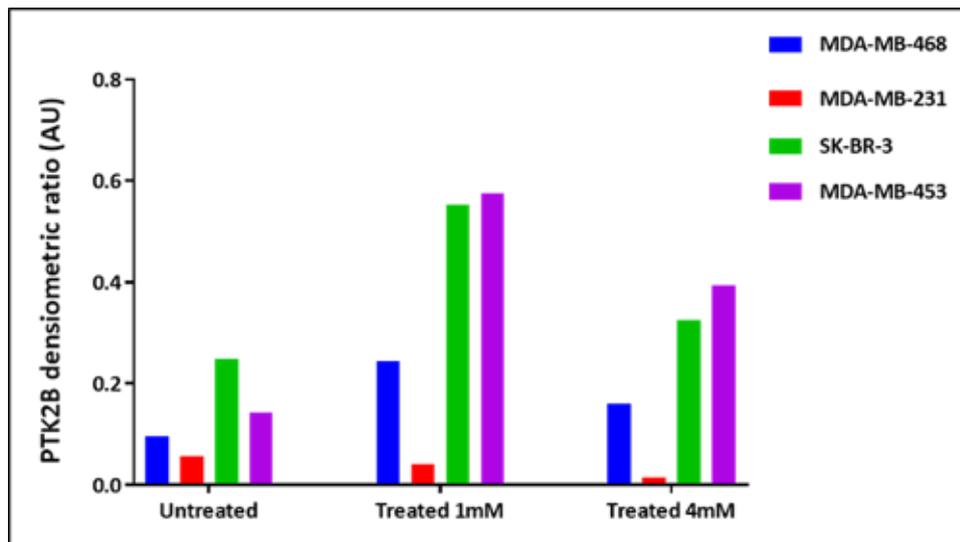


Figure 5.6. Micrograph representing the densitometry ratio (AU) of PTK2B protein expression in Metformin untreated and treated (1mM and 4mM) in MDA-MB-468, MDA-MB-231, SkBr3 and MDA-MB-453 breast cancer cell lines(n=1). Higher protein expression level was observed in 1mM Metformin-treated cell lines.

5.5. Immunofluorescence analysis of PTK2B and IRF-9 expression in Basal-like (MDA-MB-468), Claudin-Low (MDA-MB-231) and HER2 (SkBr3 and MDA-MB-453) cell lines

Confirming the protein expression of IRF-9, PTK2B and SERPINB4 in untreated and treated cells, immunofluorescence (IF) staining was performed. For more reliability, two cell lines have been used for each phenotype in this staining, which is Basal-like (MDA-MB-468), Claudin-Low (MDA-MB-231) and HER2 (SkBr3 and MDA-MB-453) cell lines. Specific antibodies have been used in this assay, which details are in chapter 2 materials and methods (section 2.2. Reagents, page 2). IF analysis showed a higher level of IRF-9 and PTK2B expression in almost all cell lines (MDA-MB-468, MDA-MB-231, SkBr3 and MDA-MB-453) in 1mM Metformin treated cells compared to a lower level of expression in control samples and a very low or undetected expression in 4mM samples. On the other hand, the SERPINB4 expression was noticeably higher in control samples and all cell lines compared to 1mM and 4mM samples, which were expressing a low or undetectable level of expression. These confirm previously obtained results by western blot analyses and further validate qRT-PCR and microarrays results. IRF-9 expression (red fluorescence) was shown in Figures 5.7, 5.8, 4.9 and 5.10 below.

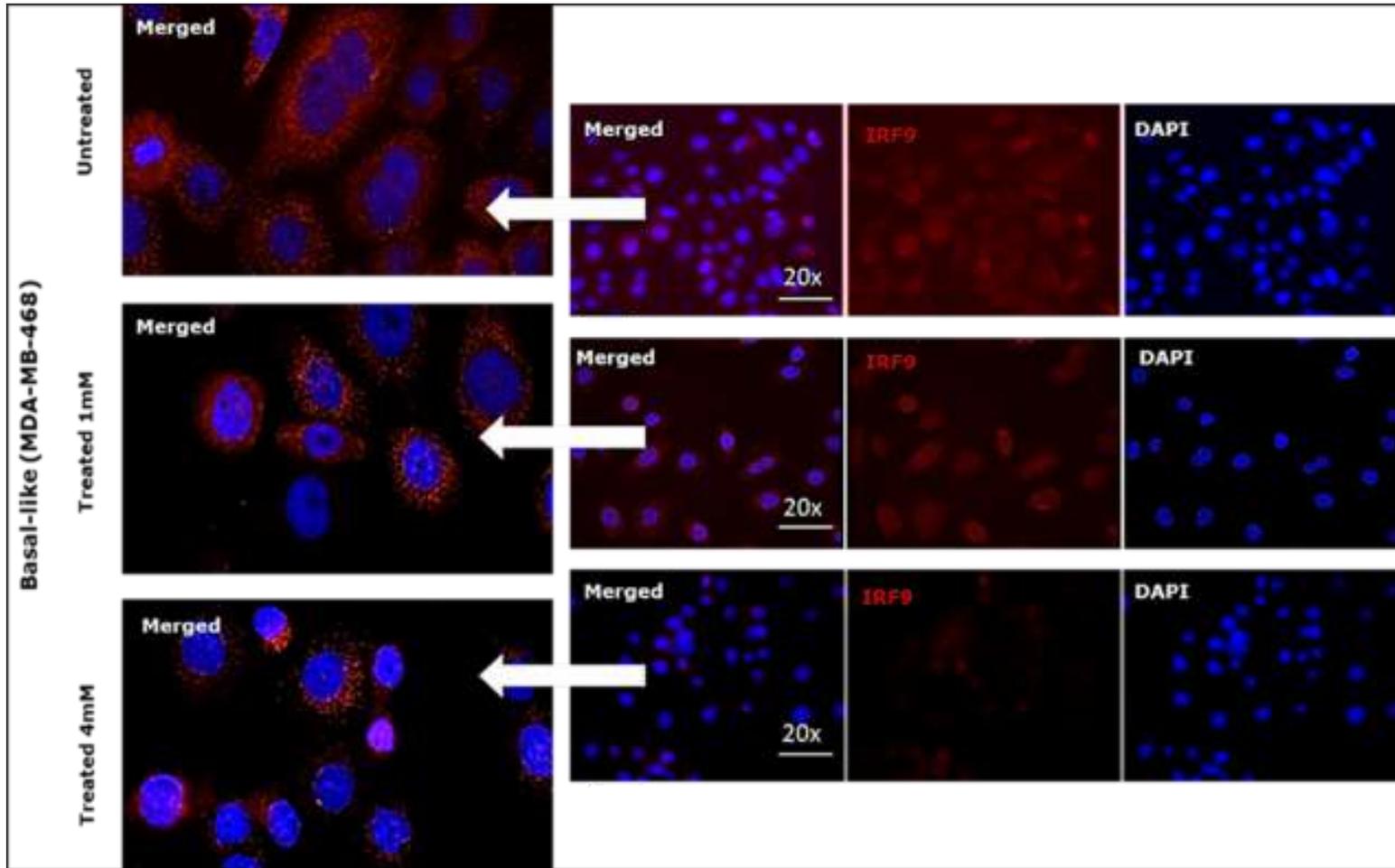


Figure 5. 7. Micrographs showing expression of IRF-9 in Basal-like (MDA-MB-468) cell line using specific antibodies against IRF-9 (red). Nuclei were stained with DAPI (blue). the images were taken at 20X magnification and scale bars indicate 100µm.

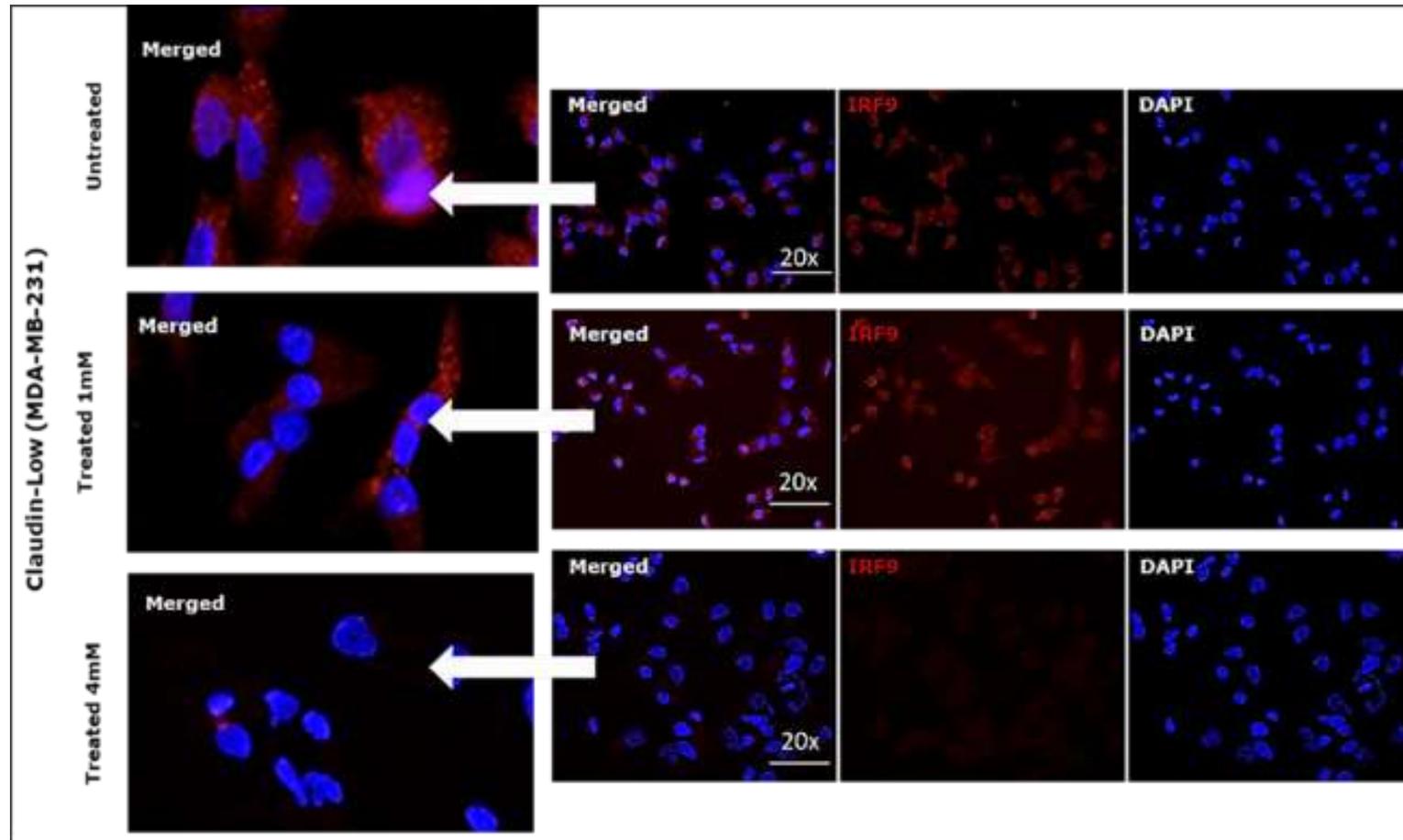


Figure 5. 8. Micrographs showing expression of IRF-9 in Claudin-Low (MDA-MB-231) using specific antibodies against IRF-9 (red). Nuclei were stained with DAPI (blue). the images were taken at 20X magnification and scale bars indicate 100µm.

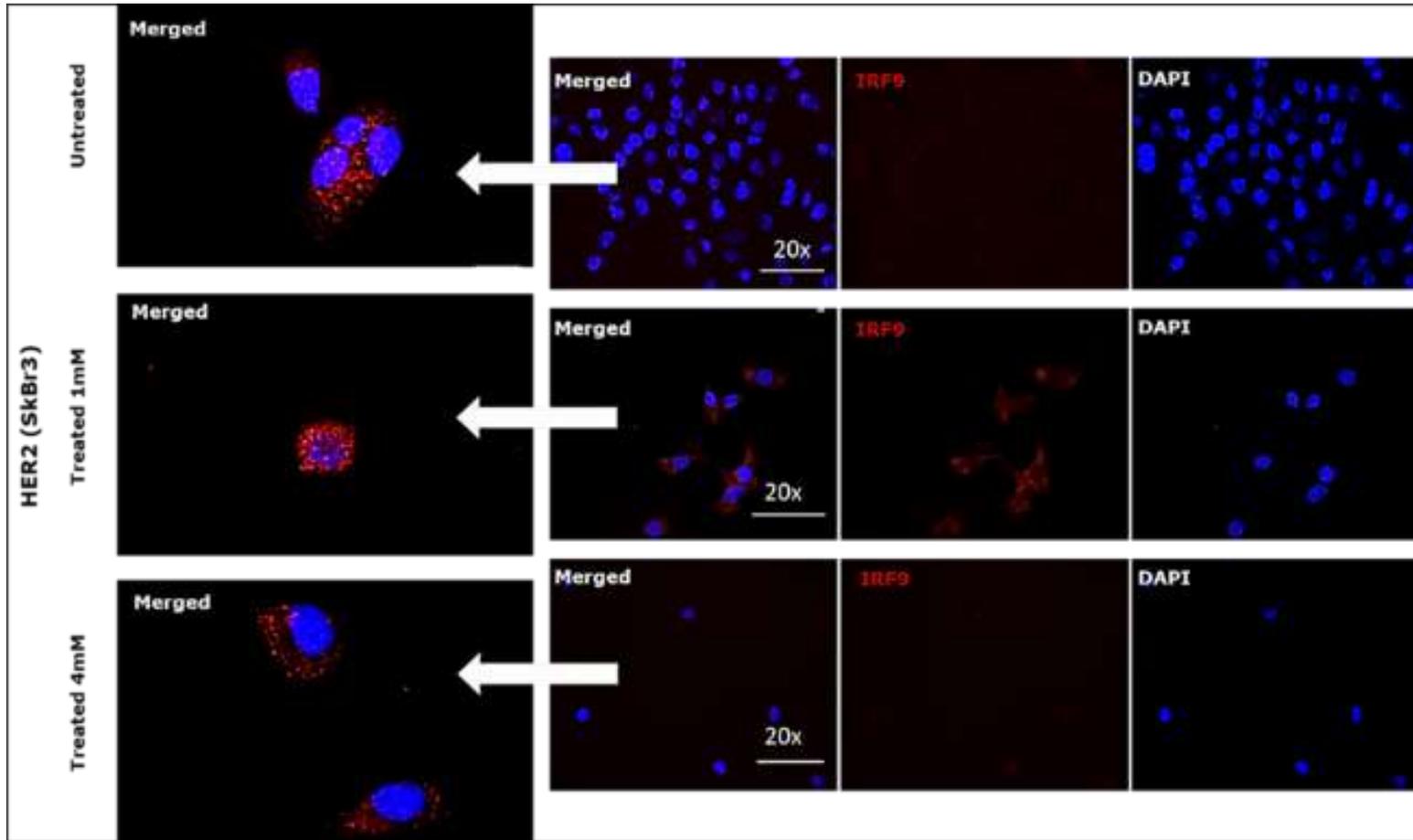


Figure 5. 9. Micrographs showing expression of IRF-9 in HER2 (SkBr3) cell line using specific antibodies against IRF-9 (red). Nuclei were stained with DAPI (blue). the images were taken at 20X magnification and scale bars indicate 100µm.

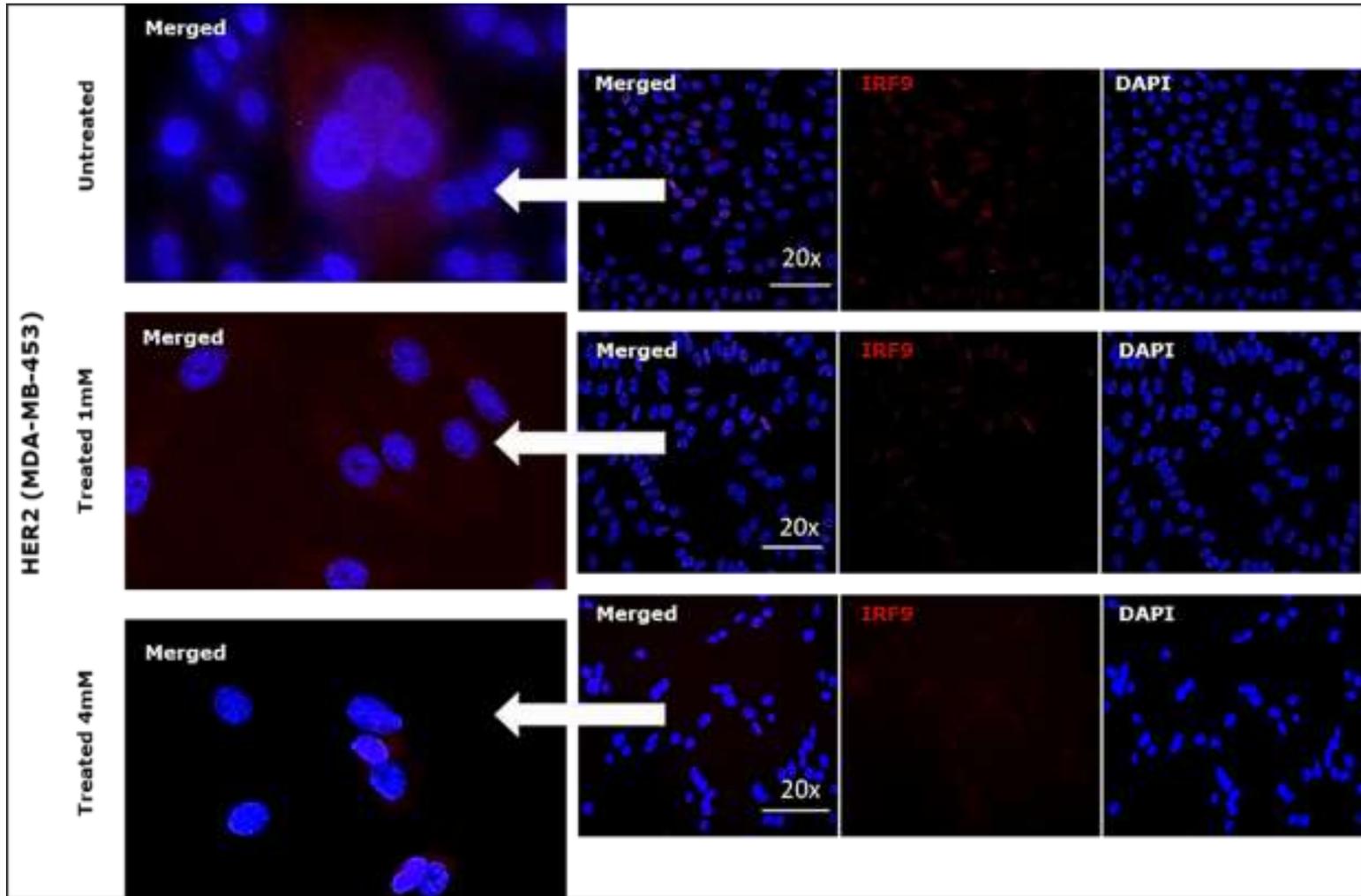


Figure 5. 10. Micrographs showing expression of IRF-9 in HER2 (MDA-MB-453) cell line using specific antibodies against IRF-9 (red). Nuclei were stained with DAPI (blue). the images were taken at 20X magnification and scale bars indicate 100µm.

A stronger PTK2B expression (green fluorescence) was detected in all cell lines treated with 1mM and when compared to the controls and 4mM treatments that showed a very low or undetectable expression of PTK2B Figures (5.11, 5.12, 4.13, and 5.14). These IF results confirmed previous findings on a PTK2B expression whose expression was noticeably increased and correlated with 1mM Metformin treatments.

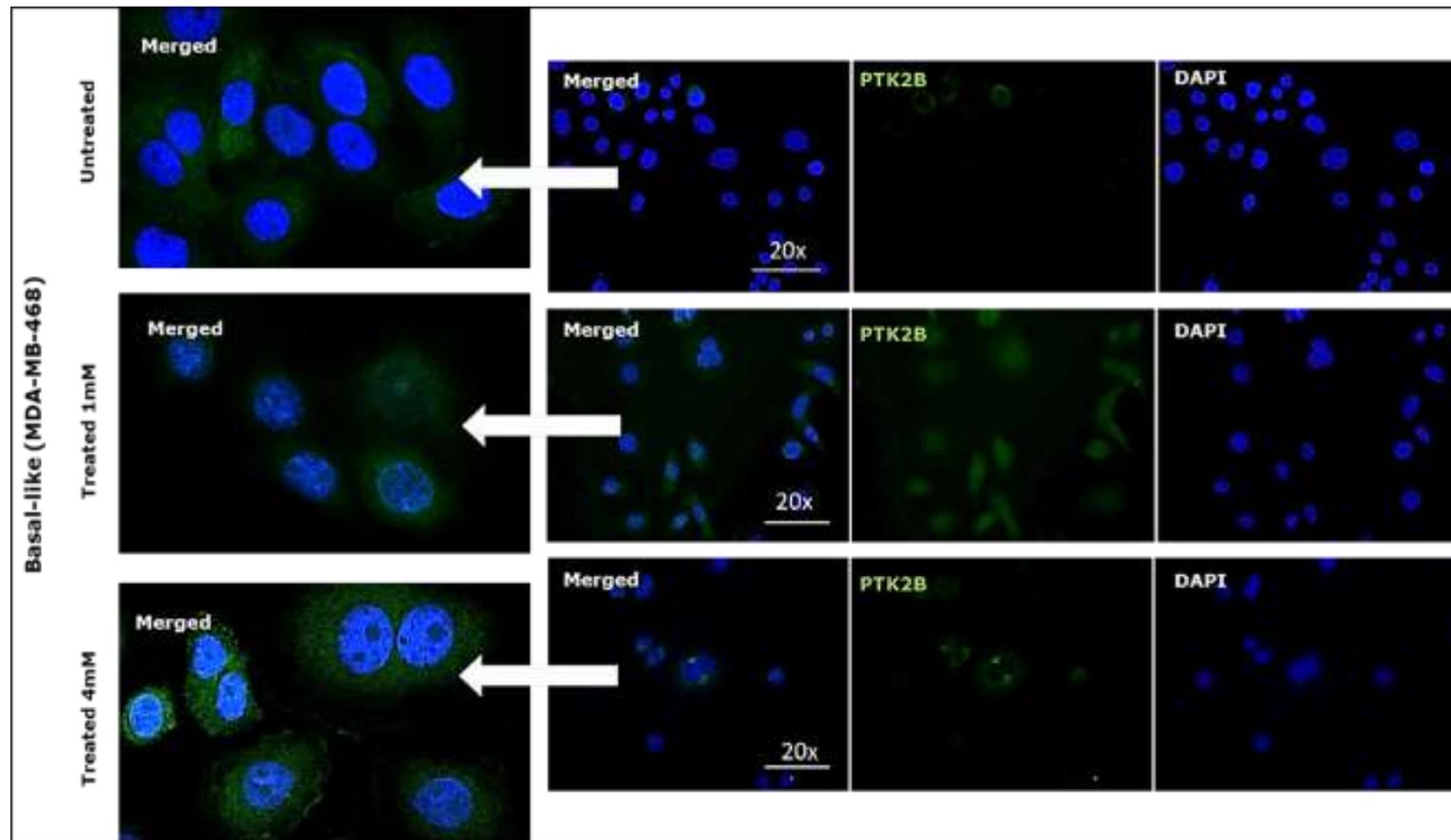


Figure 5. 11. Micrographs illustrating expression of PTK2B in Basal-like (MDA-MB-468) cell line against PTK2B (green). Nuclei were stained with DAPI (blue). the images were taken at 20X magnification and scale bars indicate 100µm.

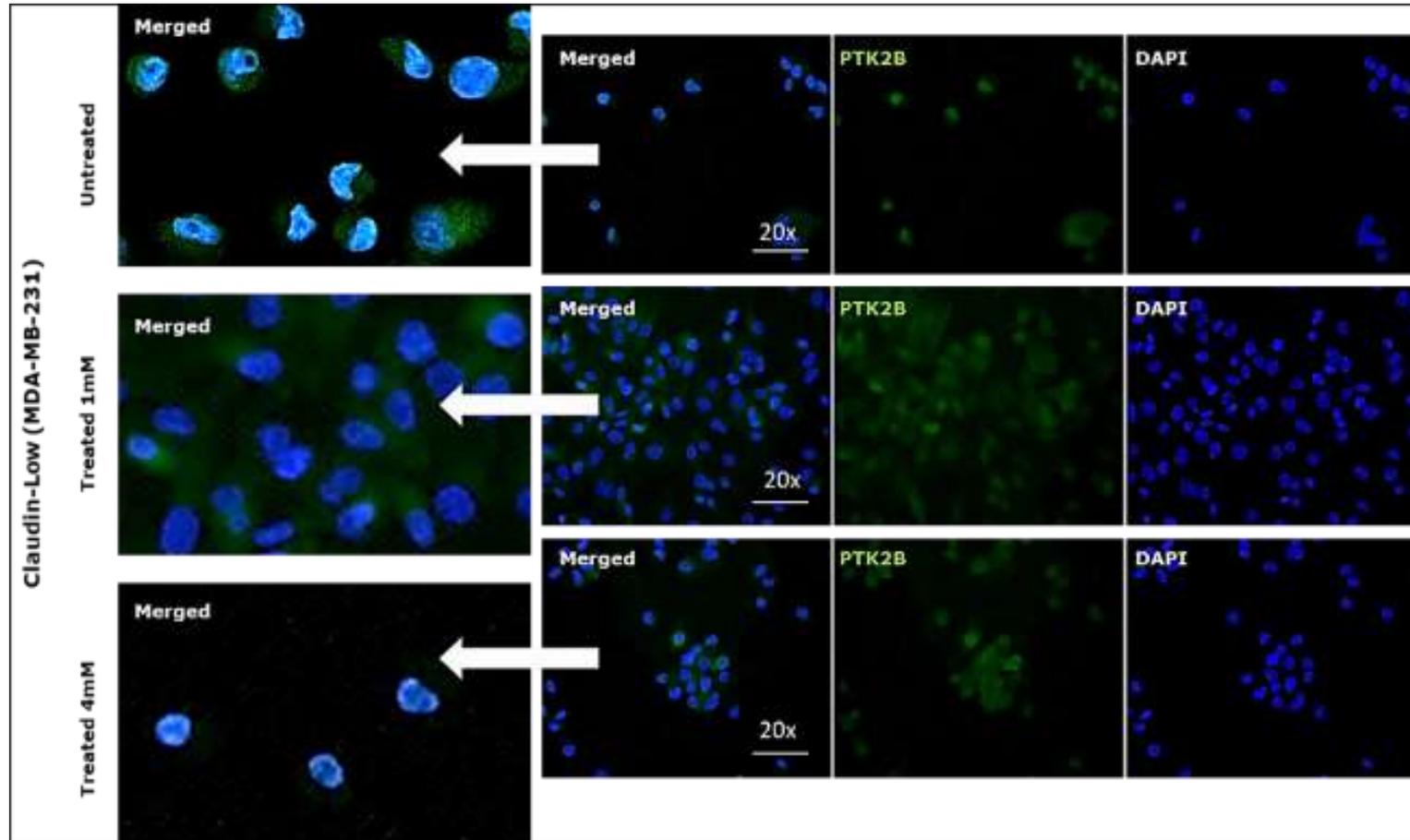


Figure 5. 12. Micrographs illustrating expression of PTK2B in Claudin-Low (MDA-MB-231) cell line against PTK2B (green). Nuclei were stained with DAPI (blue). the images were taken at 20X magnification and scale bars indicate 100µm.

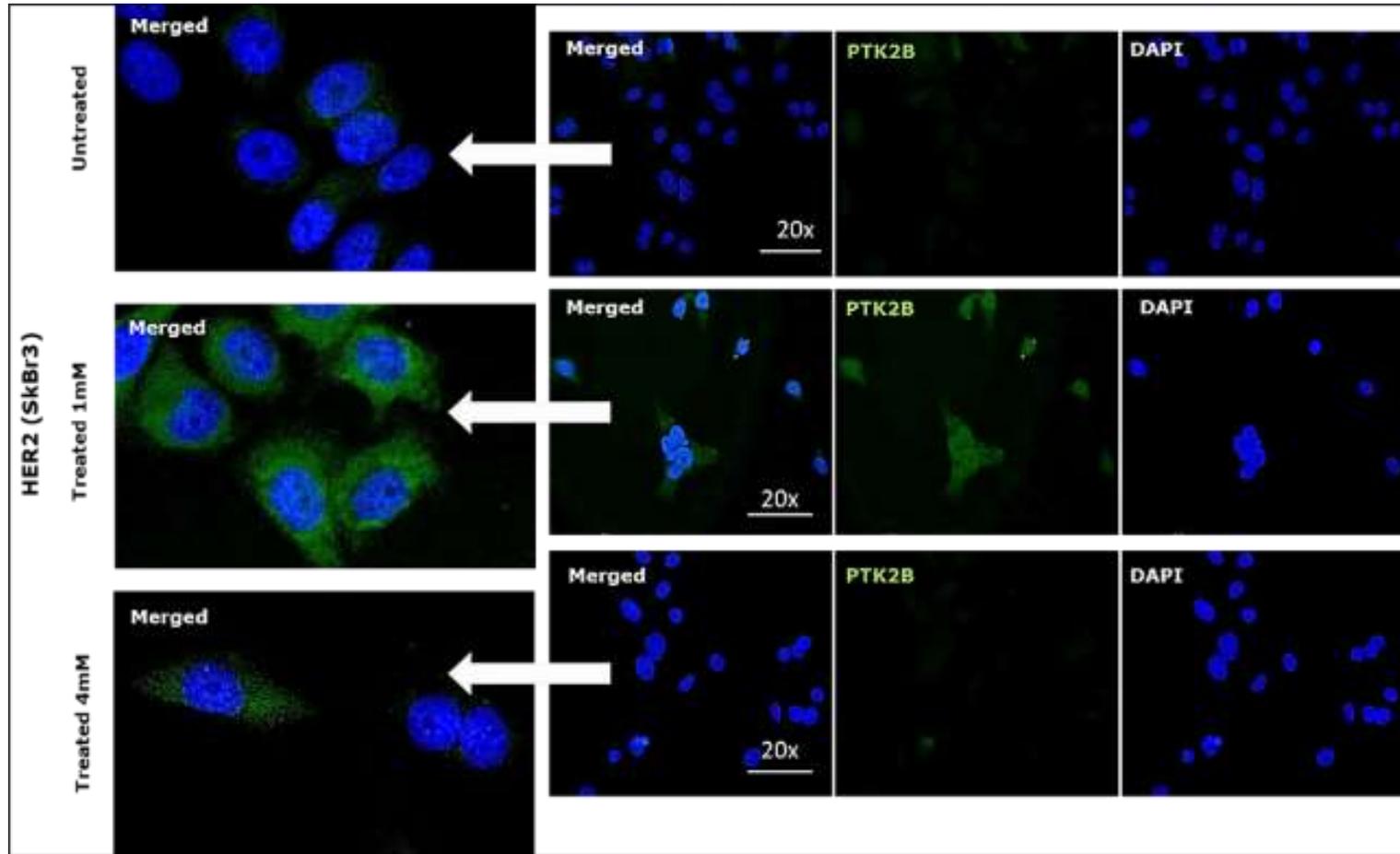


Figure 5. 13. Micrographs illustrating expression of PTK2B in HER2 (SkBr3) cell line against PTK2B (green). Nuclei were stained with DAPI (blue). the images were taken at 20X magnification and scale bars indicate 100µm.

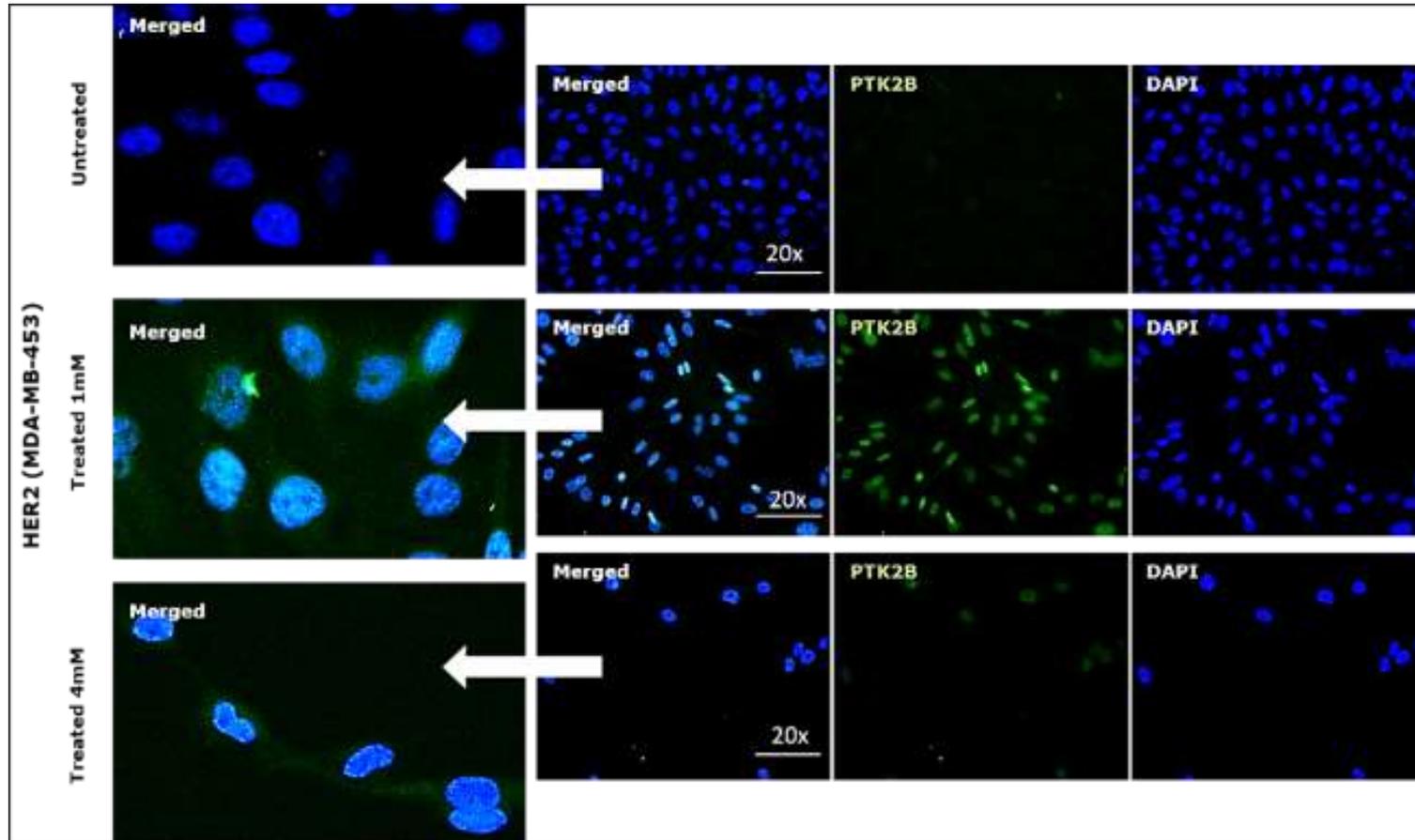


Figure 5. 14. Micrographs illustrating expression of PTK2B in HER2 (MDA-MB-453) cell line against PTK2B (green). Nuclei were stained with DAPI (blue). the images were taken at 20X magnification and scale bars indicate 100 μ m.

SERPINB4 expression was highly expressed in controls in all cell lines compared to 1mM and 4mM samples, and which showed low or undetectable levels of expression (Figures A.5.3, A.5.4, A.5.5 and A.5.6) in Appendix.

The IF results confirmed previously obtained results by Western blot and qRT-PCR analyses, however, *SERPINB4* displayed a different level of expression in Gene Expression Microarray analysis when Regression test has been used and only in BASAL-LIKE (MDA-MB-468) 4mM doses, while this gene has been found as down-regulated in an MDA-MB-468 cell line in (ANN) stepwise analysis approach (see chapter 4 section 4.5.3)

The *Protein Tyrosine Kinase 2 Beta (PTK2B)* gene has been found as highly expressed in HER2 (SkBr3) 1mM concentration of Metformin, with this subtype being less sensitive to Metformin treatment. *PTK2B* is also represented at a lower level in the other breast cancer subtypes after treatment with similar concentrations of Metformin.

PTK2B was selected for further study. The selection of *PTK2B* as a target for small molecule therapeutic approach was based on three criteria: (i) The high level of expression in low doses of Metformin treatment in HER2+ breast cancer cell line, which was the less sensitive breast cancer phenotype (with significantly high P value). (ii) The essential role in different biological processes including regulating the reorganisation of the actin cytoskeleton, cell polarisation, migration, adhesion and spreading, and is also involved in bone remodelling; (iii) The oncogenic role in the development of cancer generally, and in breast cancer specifically.

5.6. The application of NanoString technology on Basal-like and HER2 phenotypes samples for pan-cancer pathways detection

The nCounter SPRINT Profiler is an advanced instrument, which is designed to analyse RNA, DNA, and proteins from difficult samples, such as tissue, impure cell lysates and biofluid samples or Formalin-fixed Paraffin-embedded (FFPE) extracts. The Counter[®] Technology Analysis System employs a novel digital barcode technology for the direct multiplexed measurement of analysis. In addition, it offers high levels of accuracy and sensitivity (< 1 copy per cell). This technology utilises molecular "barcodes" and single molecule imaging for direct hybridisation and detection of hundreds of unique transcripts in a single reaction. Each colour-coded barcode is tied to a single specific- target probe correspondent to an analyte of interest. Simultaneously, it is combined with an invariant control; the probes form a multiplexed Code Set (*Saunus et al., 2018*).

The nCounter[®] Pan-Cancer Pathways Panel comprises a multiplex gene expression analysis with 770 genes from 13 cancer-associated canonical pathways including MAPK, STAT, PI3K, RAS, Cell Cycle, Apoptosis, Hedgehog HH, Wnt APC, DNA Damage Control, Transcriptional Regulation, Chromatin Modification and TGF- β . Signalling pathways in tumours can be further organized into three core cellular processes: (i) Cell fate: the pathways that function through this process include Wnt, HH, and NOTCH. (ii) Cell survival: pathways like MAPK, STAT, PI3K, RAS, Cell Cycle Apoptosis, and TGF- β are well known to control cell survival. (iii) Genome maintenance: including the DNA damage control pathway. Figure 5.15 below summarises the signalling pathways and the cellular processes they regulate in cancer cells (*Vogelstein et al., 2013*). Furthermore, driver genes are genes in which acquired mutations are causally linked to cancer progression. Cancer driver genes can be classified as tumour suppressor genes (*TSGs*) or oncogenes (*OGs*), functionally based on their role in disease formation (*Waks et al., 2016*).

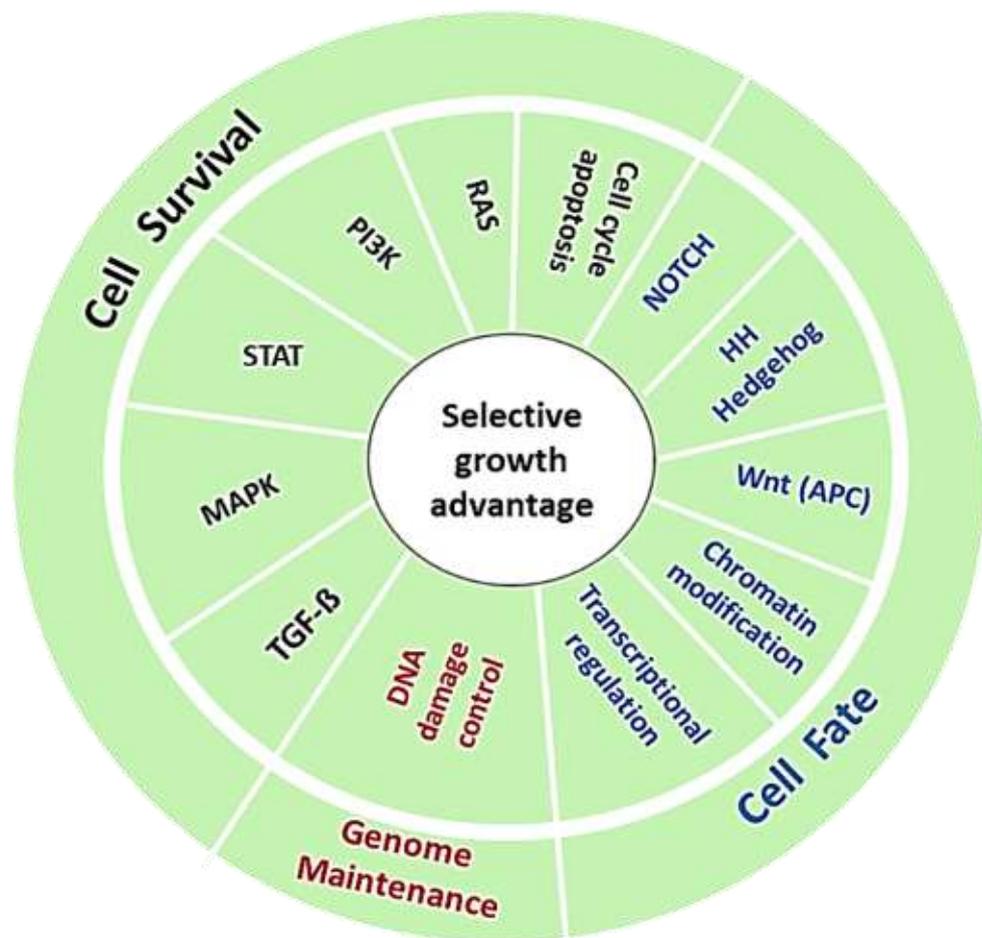


Figure 5.15. Classification of driver genes in cancer cell signalling and cellular processes. Adapted from Vogelstein et al., (2013). All driver genes can be classified into one or more of 12 pathways represented in the middle circle. This can be further organized into three core cellular processes (Cell fate, Cell survival, and Genome maintenance), represented in the outer circle.

The above-mentioned panel was used to measure the effect of Metformin treatments on the Pan-Cancer pathways with highly multiplexed analysis of basic cancer biology and pathway deregulation activity. In addition, this experiment was aimed at confirming the previously discussed results (Last chapter).

RNA samples that were previously extracted using the RNeasy kit (QIAGEN), and quantified by the Nanodrop, were used in this assay. Twelve samples that had the highest quality (six samples for each) of both Basal-like (MDA-MB-468) and HER2 (SkBr3) cell lines were applied to Pan-Cancer Pathways Panel and according to the manufacturer's instructions using (100ng/5ul) of total RNA. Data were collected using the nCounter_ Dx Digital Analyser and processed using nSolver Software. The top 20 significant P values affected pathways are displayed in the Tables 5.2 and 5.3, and which show the Log2 fold change for Basal-like (MDA-MB-468) cell line that was

Metformin treated with concentrations of 1mM and 4mM and compared to the untreated (controls).

Table 5.2. The top 20 significant P value genes and corresponding pathways

Gene ID	Log 2-fold change	P-value	Gene. Sets
GADD45G-mRNA	3.01	0.0004	Cell Cycle - Apoptosis, MAPK
IL6R-mRNA	2.62	1.55E-05	JAK-STAT, PI3K
FOSL1-mRNA	2.59	0.0012	Wnt
GADD45A-mRNA	2.58	0.0010	Cell Cycle - Apoptosis, MAPK
LAMB3-mRNA	2.36	0.0001	PI3K
DUSP5-mRNA	2.35	0.0001	MAPK
KLF4-mRNA	1.86	0.0001	Driver Gene
NF2-mRNA	1.8	6.29E-05	Driver Gene
SPRY2-mRNA	1.64	0.0007	JAK-STAT
POLB-mRNA	0.98	0.0003	DNA Damage - Repair
AXIN1-mRNA	0.90	0.0004	Driver Gene, Wnt
DTX4-mRNA	-1.10	0.0008	Notch
FEN1-mRNA	-1.14	0.0002	DNA Damage - Repair
HSPB1-mRNA	-1.25	0.0001	MAPK
MYD88-mRNA	-1.30	0.0003	Cell Cycle - Apoptosis, Driver Gene
PRKACA-mRNA	-1.42	0.0011	Cell Cycle - Apoptosis, Hedgehog, MAPK, Ras, Wnt
CREB3L4-mRNA	-1.52	0.0006	PI3K
MCM2-mRNA	-1.62	0.0010	Cell Cycle - Apoptosis
PBX1-mRNA	-2.06	5.42E-06	Transcriptional Misregulation
MMP7-mRNA	-2.52	0.0014	Wnt

According to the Log2 fold change out of Nanostring analysis in Basal-like (MDA-MB-468) cell line using 1mM doses of Metformin. The colour gradient is between red (up-regulated genes) and blues (down-regulated genes).

Table 5.3. The top 20 significant P value genes and corresponding pathways

Gene ID	Log2 fold change	P-value	Gene. Sets
GADD45A-mRNA	4.72	0.0004	Cell Cycle - Apoptosis, MAPK
TNFAIP3-mRNA	3.19	0.0012	Driver Gene
IL6R-mRNA	2.49	9.01E-05	JAK-STAT, PI3K
KLF4-mRNA	1.60	0.0013	Driver Gene
POLB-mRNA	1.43	0.0003	DNA Damage - Repair
MAPK1-mRNA	1.17	0.0008	MAPK, PI3K, Ras, TGF-beta
FANCC-mRNA	0.81	0.0006	DNA Damage - Repair
PBX1-mRNA	-1.16	0.0002	Transcriptional Misregulation
ABL1-mRNA	-1.27	0.0003	Cell Cycle - Apoptosis, Driver Gene, Ras
TMPRSS2-mRNA	-1.52	0.0008	Transcriptional Misregulation
AMER1-mRNA	-1.92	0.0010	Driver Gene
DTX4-mRNA	-2.11	0.0002	Notch
BAMBI-mRNA	-2.21	0.0004	TGF-beta, Wnt
MEN1-mRNA	-2.35	0.0012	Driver Gene, Transcriptional Misregulation
IGFBP3-mRNA	-2.36	0.0001	Transcriptional Misregulation
MYD88-mRNA	-2.43	0.0001	Cell Cycle - Apoptosis, Driver Gene
HIST1H3B-mRNA	-4.52	0.0010	Driver Gene, Transcriptional Misregulation
H2AFX-mRNA	-4.93	5.86E-05	DNA Damage - Repair
ID1-mRNA	-5.03	0.0002	TGF-beta
HIST1H3G-mRNA	-6.04	0.0007	Transcriptional Misregulation

According to the Log2 fold change out of Nanostring analysis in Basal-like (MDA-MB-468) cell line using 4mM doses of Metformin. The colour gradient is between red (up-regulated genes) and blues (down-regulated genes).

The up-regulated genes in Basal-like (MDA-MB-468) cell line were mainly involved in Cell Cycle - Apoptosis, MAPK; JAK-STAT, PI3K; Wnt; PI3K; MAPK; Driver Gene; JAK-STAT; Driver Gene, Wnt; and DNA Damage – Repair, as well as MAPK- PI3K-Ras- TGF- β , which are directed to the cell survival.

Whilst, the down-regulated genes were implicated in Notch; DNA Damage – Repair; MAPK; Cell Cycle – Apoptosis- Driver Gene; Cell Cycle – Apoptosis- Hedgehog- MAPK- Ras- Wnt; PI3K; Cell Cycle – Apoptosis; Transcriptional Misregulation; and Wnt in addition to Driver Gene; TGF- β -Wnt; and TGF- β pathways that govern cell growth advantage.

This result with the previous results in chapter 4, which displayed a few common pathways including Apoptosis signalling as an up-regulated pathway. In addition, there were few combined downregulated pathways such as; Wnt signalling pathway; Apoptosis signalling pathway; PI3 kinase pathway; and TGF- β signalling pathway between the current results and the previous results in chapter 4.

Likewise, several pathways including Cell Cycle – Apoptosis- MAPK; Ras; MAPK; Transcriptional Misregulation; Driver Gene; Driver Gene- PI3K; JAK-STAT; and MAPK- PI3K, were found as up-regulated pathways in HER2 (SkBr3) cell line population at both concentrations of 1mM and 4mM. Some of these pathways are involved in cell survival selective growth feature (Tables 5.4 and 5.5). The down-regulated pathways were; Cell Cycle – Apoptosis- Driver Gene- JAK-STAT- MAPK- PI3K- Ras; Transcriptional Misregulation; Driver Gene- MAPK- PI3K- Ras; PI3K; DNA Damage – Repair; Cell Cycle – Apoptosis; Cell Cycle – Apoptosis- TGF-beta; and Driver Gene as well as, MAPK; PI3K; Ras; Notch; and Driver Gene, JAK-STAT in both (1mM and 4mM) treatments of Metformin. Few pathways are involved in cell survival and some pathways are function in cell fat selective growth advantages. Once again, consistent results with earlier results (chapter 4) represent one common pathway identified by down-regulated genes, which is Cell Cycle – Apoptosis- TGF- β pathway.

Table 5.4. The top 20 significant P value genes and corresponding pathways

Gene ID	Log2 fold change	P-value	Gene. Sets
GADD45A-mRNA	2.77	0.0008	Cell Cycle - Apoptosis, MAPK
ETS2-mRNA	1.94	0.0014	Ras
RASA4-mRNA	1.80	0.0026	Ras
GADD45B-mRNA	1.59	0.0014	Cell Cycle - Apoptosis, MAPK
HSPB1-mRNA	1.36	0.0001	MAPK
TMPRSS2-mRNA	1.20	0.0005	Transcriptional Misregulation
KLF4-mRNA	1.08	0.0014	Driver Gene
IDH2-mRNA	0.86	0.0019	Driver Gene
TSC1-mRNA	0.84	0.0020	Driver Gene, PI3K
SPRY1-mRNA	0.81	0.0027	JAK-STAT
NR4A1-mRNA	0.75	0.0018	MAPK, PI3K
AKT1-mRNA	-0.72	0.0027	Cell Cycle - Apoptosis, Driver Gene, JAK-STAT, MAPK, PI3K, Ras
IGFBP3-mRNA	-0.92	0.0017	Transcriptional Misregulation
HRAS-mRNA	-0.99	0.0013	Driver Gene, MAPK, PI3K, Ras
SYK-mRNA	-1.05	0.0020	PI3K
FEN1-mRNA	-1.08	0.0002	DNA Damage - Repair
CDC6-mRNA	-1.25	0.0021	Cell Cycle - Apoptosis
TFDP1-mRNA	-1.38	0.0003	Cell Cycle - Apoptosis, TGF-beta
SRSF2-mRNA	-1.64	0.0018	Driver Gene
CDC25A-mRNA	-2.04	0.0010	Cell Cycle - Apoptosis

According to the Log2 fold change out of Nanostring analysis in HER2 (SkBr3) cell line using 1mM doses of Metformin. The colour gradient is between red (up-regulated genes) and blues (down-regulated genes).

Table 5.5. The top 20 significant P value genes and corresponding pathways

Gene ID	Log2 fold change	P-value	Gene. Sets
TNFAIP3-mRNA	3.68	0.0001	Driver Gene
JUN-mRNA	3.36	0.0003	MAPK, Wnt
GADD45A-mRNA	3.31	0.0003	Cell Cycle - Apoptosis, MAPK
GADD45B-mRNA	2.29	0.0003	Cell Cycle - Apoptosis, MAPK
HSPA1A-mRNA	2.05	0.0011	MAPK
NFKBIA-mRNA	2.05	0.0021	Cell Cycle - Apoptosis
DUSP10-mRNA	1.86	0.0014	MAPK
LAMB3-mRNA	1.76	0.0004	PI3K
TSC1-mRNA	1.37	0.0003	Driver Gene, PI3K
HSPB1-mRNA	1.17	0.0002	MAPK
HSPA2-mRNA	0.93	0.0021	MAPK
AKT1-mRNA	-0.73	0.0025	Cell Cycle - Apoptosis, Driver Gene, JAK-STAT, MAPK, PI3K, Ras
FEN1-mRNA	-0.88	0.0006	DNA Damage - Repair
NF2-mRNA	-0.95	0.0007	Driver Gene
PDGFB-mRNA	-1.08	0.0023	MAPK, PI3K, Ras
CREB3L4-mRNA	-1.13	0.0020	PI3K
DTX4-mRNA	-1.14	0.0007	Notch
H2AFX-mRNA	-1.31	0.0022	DNA Damage - Repair
SOCS1-mRNA	-1.47	0.0022	Driver Gene, JAK-STAT
CDC25A-mRNA	-1.55	0.0027	Cell Cycle - Apoptosis

According to the Log2 fold change out of Nanostring analysis in HER2 (SkBr3) cell line using 4mM doses of Metformin. The colour gradient is between red (up-regulated genes) and blues (down-regulated genes).

Furthermore, from Tables 5.2, 5.3, 5.4, and 5.5, it has been noticed that *Growth Arrest And DNA Damage-Inducible Alpha (GADD45A mRNA)*, and *Kruppel Like Factor 4 (KLF4-mRNA)*, were commonly upregulated genes between both breast cancer subtypes in 1mM dose, and *GADD45A-mRNA* and *TNFAIP3-mRNA* were upregulated in 4mM dose. *GADD45A* is a member of a group of genes whose transcript levels are increased following stressful growth arrest conditions and treatment with DNA-damaging agents (Fabregat et al., 2014). In addition, Tumor necrosis factor α induced protein 3 (*TNFAIP3*) was identified as a protein whose expression is rapidly induced by the tumour necrosis factor TNF-mediated NF- κ B activation. This protein (*TNFAIP3*) is related to inflammatory carcinogenesis in several cancer types (Hadisaputri et al., 2017).

5.7. The application of Real-time quantitative PCR (qRT-PCR) on common genes obtained from NanoString analysis

Common up-regulated genes between both breast cancer subtypes Basal-like (MDA-MB-468) and HER2 (SkBr3) cell lines have been selected for gene expression by qRT-PCR and for protein expression by western blotting. The amplification efficiency was measured for the newly ordered primers, which was quite efficient for further studies. Table 5.6 below displayed the efficiency of (*GADD45 A, B, and G*) isoforms respectively.

Table 5.6. Table representing percentages of efficiency and dynamic range of *GADD45 A, B, and G* primers

Oligo Name	Sequence (5'-3')	Tm°	Efficiency
FH1_GADD45A	GCTCAACGTAATCCACATTC	59.6	1.08
RH1_GADD45A	GAGATTAATCACTGGAACCC	57.5	
FH1_GADD45B	GTTGATGAATGTGGACCC	58.6	1.03
RH1_GADD45B	CGATGTTGATGTCGTTGTC	60.2	
FH1_GADD45G	CATTTTACGCTGATCCAGG	60.7	1.17
RH1_GADD45G	GGGTTCGAAATGAGGATG	60.5	

The table displays the Oligo name and its Sequence (5' to 3') and melting temperatures (Tm°). The results showed that all three primers were efficient for further study. The Efficiency of each primer was between 90% and 110%.

QRT-PCR was performed using three *GADD45* isoform's (A, B, and G) primers together with the housekeeping gene *GUSB*. This experiment was using different sets of cDNA template samples in this order SkBr3 control, 1mM, 4mM, and MDA-468 control, 1mM, 4mM. The threshold was set on 0.1642 for all qRT-PCRs run, while the Tm° varied according to the different primers melting temperature. Three replicates were independently performed and for each sample. Significant differences were found between *GADD45 A, B* and *G* expressions upon different treatments, and also between different cell lines. The performance of each gene as up or down-regulated are shown in Figures 5.16 for both Basal-like (MDA-MB-468) and HER2 (SkBr3) cell lines.

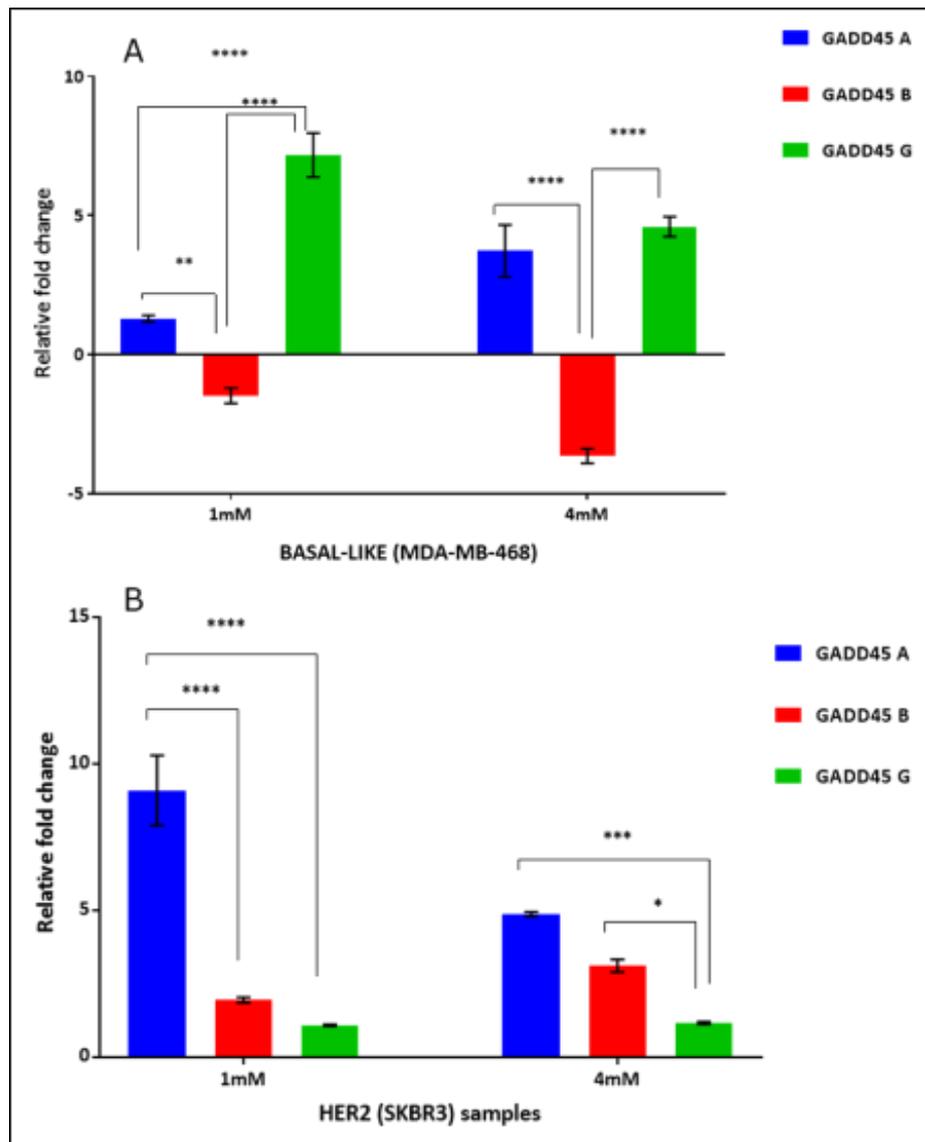


Figure 5.16. Graphs representing qRT-PCR data for (A) Basal-like (MDA-MB-468) and (B) HER2 (SkBr3) cell lines. The graphs show the relative fold change, which represented up and downregulated genes (*GADD45A*, *B*, and *G*), using different Metformin concentrations (1mM and 4mM) and when compared to untreated (control) samples. In addition, the Fold change for each sample was calculated according to this equation: Fold change due to treatment = $2^{-\Delta\Delta CT}$.

5.8. Validation of GADD45A protein expression in Basal-like and HER2 breast cancer cell lines

This assay was achieved using whole cell lysates from the Basal-like (MDA-MB-468), Claudin -Low (MDA-MB-231) and HER2 (SkBr3 and MDA-MB-453) cell lines to detect the expression of *GADD45A* protein. Two cell lines have been used for HER2 phenotype for more reliability. A higher expression of *GADD45A* protein in response to 1mM and 4mM treatments for Basal-like (MDA-MB-468) and HER2 (SkBr3) cell lines was observed and when compared to the control Figure 5.17. However, Claudin-Low (MDA-MB-231) showed low expression of the protein in response to 4mM treatment and when compared to control and 1mM treatment concentrations. Additionally, in the HER2 cell line MDA-MB-453, negative expression of *GADD45A* protein was noticed for both treatments and control.

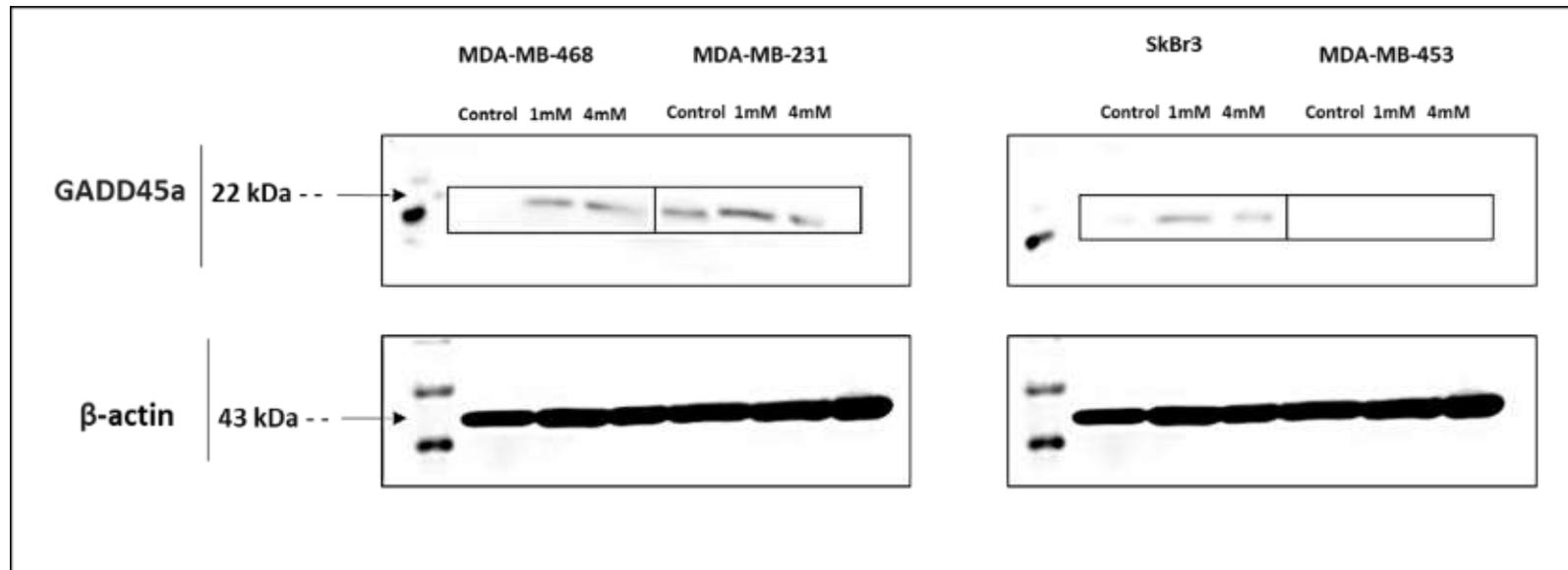


Figure 5.17. Representative micrographs of immunoblots showing GADD45A protein expression in MDA-MB-468 and MDA-MB-231 (left panel), and SkBr3 and MDA-MB-453 (right panel). β -actin is used as a loading control. Cell extracts for immunoblotting were obtained from untreated and Metformin treated cells with 1mM and 4mM concentrations.

Moreover, the protein intensity of each band was calculated using ImageJ software and was represented as Arbitrary Unit (AU). Figure 5.18 demonstrates the protein intensity ratio in AU for *GADD45A* protein. Clearly, these data confirmed previous results obtained from the NanoString and qRT-PCR techniques.

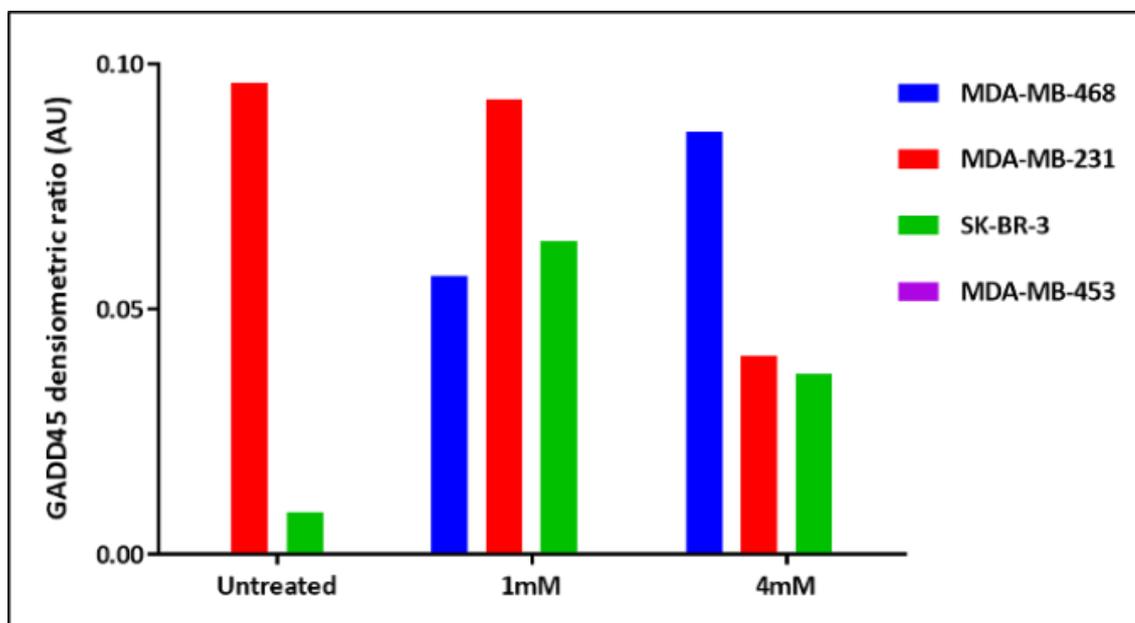


Figure 5.18. Graph representing densitometry ratio AU of *GADD45A* protein expression in different breast cancer cell lines that were Metformin treated and untreated with 1mM and 4mM concentrations. Higher protein expression in 1mM and 4mM concentrations in Basal-like (MDA-MB-468), and HER2 (SkBr3), and when compared to the control. However, less protein expression in both (1mM and 4mM) doses in Claudin-Low (MDA-MB-231) in comparison with the control. No protein expression was detected in the HER2 (MDA-MB-453) cell.

5.9. Discussion

The main objective of this chapter was to investigate the expression of several genes (*ANKRD44*, *BTN3A1*, *C2orf42*, *DHFR2*, *EMP1*, *HGF*, *IRF-9*, *NPRL3*, *OXSRI*, *PCDHB2*, *PRDX1*, *PTK2B*, *PXDC1*, *SERPINB4*, *VSTM2B*, and *WBSCR27*) that were previously identified (see Chapter 4) using microarray gene expression analysis from Basal-like and HER2 cell lines. This study also aimed to select markers for further studies. This objective has been met and evidence was provided using qRT-PCR, western blot, Immunofluorescence techniques. NanoString technique has also been used for pan-cancer pathways detection, attempting to find common pathways across different techniques. The qRT-PCR data analysis revealed that *ANKRD44*, *BTN3A1*, *C2orf42*, *DHFR2*, *IRF-9*, *NPRL3*, *OXSRI*, *PCDHB2*, *PTK2B*, and *PXDC1* confirmed the previous Gene expression profiling data. Three other genes including *SERPINB4*, *EMP1*, and *PRDX1* showed opposite gene expression in this analysis. In addition, three primers were inefficient for further study, which are (*HGF*, *VSTM2B*, and *WBSCR27*).

Additionally, two upregulated genes (*IRF-9* and *PTK2B*) and one downregulated gene (*SERPINB4*) have been chosen for further studies using western blotting and immunofluorescence assays to investigate their protein expression statuses. Those markers were selected according to their significant P values, and substantial biological and pathological functions. A brief summary of the function of the identified genes is provided below:

IRF-9 (Interferon Regulatory Factor 9), also known as *p48*, *ISGF3* or *ISGF3G*, is an interferon regulatory factor that mediates cell signalling by type I IFNs (IFN- α and IFN- β). *IRF-9* is also a specific transcription factor found downstream the JAK/STAT pathway (Janus kinase (JAK)/Signal Transducers and Activators of Transcription (STATs)), and which plays a role in body immune balance (*Zhao et al., 2017*). In addition, *Luker* and her colleagues, observed in their study published in 2001, that transient overexpression of *IRF-9* confers a drug-resistance phenotype and stimulates the expression of IFN- reacting genes in breast and uterine tumours. Their data also identified a novel IFN-independent role for *IRF-9* in the development of resistance to antimicrotubular agents by breast tumour cells, and which might link downstream mediators of IFN signalling to drug resistance in human cancers. This suggests that *IRF-9* upregulation following the Metformin treatment is probably associated with drug resistance in breast cancer (*Luker et al., 2001*).

PTK2B (Protein Tyrosine Kinase 2 Beta) or *PYK2*, is a non-receptor protein tyrosine kinase that regulates reorganisation of the actin cytoskeleton, cell polarisation, migration, adhesion and spreading, and is also involved in bone remodelling. Amongst its related pathways are the immune response-IFN gamma signalling pathway and G-protein-coupled receptors (GPCRs) that

mediate the sympathetic nervous system. It also functions in downstream signalling of several cellular pathways involving integrin and collagen receptors, immune receptors, G-protein coupled receptors (GPCR), cytokine, chemokine and growth factor receptors, and cellular stress. Moreover, *PTK2B* plays an oncogenic role in cancer development by regulating cell motility and invasion in glioma cells, modulating cell proliferation and tumour growth in multiple myeloma cells, and promoting cell proliferation and invasiveness through c-Src and ERK (extracellular signal-regulated kinases) activation in hepatocellular carcinoma. c-Src is a 60-kDa non-receptor protein tyrosine kinase, which is the protein product of the proto-oncogene *c-src*. In breast cancer cells, the kinase activity of *PYK2* and its Tyrosine 402 phosphorylation are required for cell invasion. It also can mediate ErbB-2 (Her2 Neu) signalling to upregulate the adhesive ability of androgen receptor (AR-positive) in prostate cancer (PCa) cells and is involved in RhoC-promoted invasiveness. In cells, *PYK2* via *S6K1* (*Ribosomal protein S6 kinase beta-1*) activation modulates AR function and growth properties (Hsiao *et al.*, 2016). Interestingly, Selitrennik and Lev, 2015 found in their study, that *PYK2* merges EGFR/ HER2- and Interleukin-8 (IL-8 -receptor) signalling to promote cell invasiveness in breast cancer cells. *PYK2* has also been found to be activated in response to both EGF and heregulin (HRG) in different breast cancer cell lines. Finally, the depletion of *PYK2* leads to significant inhibition of EGF/HRG-mediated cell spreading, migration and invasion in breast carcinoma (Selitrennik and Lev, 2015).

***SerpinB4* (Serpin Family B Member 4)** or *SCCA2* (Squamous cell carcinoma antigen2), is a member of the ovalbumin family of serine proteinase inhibitors. Originally, it was discovered as tumour-specific antigen and a tumour marker for various kinds of squamous cell carcinomas. The mechanism of how *SCCA2* enhance tumour growth has been established through its involvement in cell death inhibition, cell growth enhancement, epithelial-mesenchymal transition (EMT) induction, and defence against tumours inhibition (Izuhara *et al.*, 2018). Furthermore, *SCCA2* was shown to be involved in the pathogenesis of several inflammatory diseases including asthma, psoriasis, and atopic dermatitis (AD). In addition, *SCC2* has been suggested as a novel biomarker for skin inflammatory diseases (Izuhara *et al.*, 2018). Moreover, the study by (de Koning *et al.*, 2011) demonstrate that overexpression of *SERPINB4* in HeLa cells inhibits recombinant (Granzyme M) GrM-induced and (Natural killer) NK cell-mediated cell death indicating the beneficial expression of *SERPINB4* for tumour cells. This might represent a novel mechanism by which tumour cells avoid the cytotoxic lymphocyte-induced GrM-mediated cell death.

Analyses using NanoString technique and the nCounter® Pan-Cancer Pathways Panel, which encodes multiplex gene expression analysis with 770 genes from 13 cancer-associated canonical pathways such as MAPK, STAT, PI3K, RAS, Cell Cycle, Apoptosis, Hedgehog, Wnt, DNA Damage

Control, Transcriptional Regulation, Chromatin Modification, and TGF- β has been utilised to emphasise previously discussed results in the last chapter.

The commonly downregulated genes in 1mM and 4mM were *Flap endonuclease 1(FEN1-mRNA)*, *Deltex E3 Ubiquitin Ligase 4 (DTX4-mRNA)* and *H2A Histone Family Member X (H2AFX-mRNA)*. *Flap Structure-Specific Endonuclease 1(FEN1)* is a Protein-Coding gene. The annotations related to this gene according to Gene Ontology (GO) include magnesium ion binding and damaged DNA binding. This gene expression was found to be inducible during cell proliferation for DNA synthesis, while it was down-regulated during cell differentiation (*Kim et al., 2000*). As well as, *Deltex E3 Ubiquitin Ligase 4 (DTX4)* is a Protein-Coding gene, which is a regulator of the Notch signalling pathway (this pathway is involved in cell-cell communications that regulate a broad spectrum of cell-fate determinations by similarity), (*Moretti and Brou, 2013*). *H2A Histone Family Member X (H2AFX-mRNA)* belongs to the histones family that are basic nuclear proteins responsible for the nucleosome structure of the chromosomal fibre in eukaryotes. Essentially, it contributes to genome stability *via* its signalling role in DNA damage events and acts as a foundation for the association of repair foci (*Pinto and Flaus, 2010*). Another observation was made, that *HSPB1-mRNA (Heat Shock Protein Family B (Small) Member 1)* was up-regulated gene in both 1mM and 4mM treatment concentrations in HER2 (SkBr3) cell line and down-regulated in Basal-like (MDA-MB-468) cell line (1mM concentration). *HSP27* is a significant member of the HSP family, which plays a key role in tumours prominence and progression. This protein is highly expressed in variance cancers and is related to aggressive tumour behaviour, metastasis, poor prognosis and resistance to chemotherapy (*Konda et al., 2017*).

Additionally, *Transmembrane Protease, Serine 2 (TMPRSS2-mRNA)* was shown to be strongly expressed in HER2 (SkBr3) 1mM treated samples and weakly expressed in Basal-like (MDA-MB-468) 4mM treated samples. This gene encodes a protein that belongs to the serine protease family, which is known to be involved in many physiological and pathological processes. It is also having an important role in mammalian tissues development and homeostasis including (heart, skin, inner ear, placenta, and digestive tract). Apparently, the irregular expression of these genes is associated with several aetiology disorders, including cancer (*Szabo and Bugge, 2008*).

In this regard, the common upregulated gene *GADD45A* between Basal-like and HER2 breast cancer cell lines in the NanoString was validated for gene expression (qRT-PCR) and protein expression (western blot). These results also explored the link between previously mentioned pathways and the ones obtained by NanoString. For instance, the apoptosis signalling pathway and Wnt signalling pathway were found up-regulated and down-regulated respectively in Basal-like (MDA-MB-468) cell line (See chapter 4, Figure 4.10. Pathway ontology of the consensus

Basal-like (MDA-MB-468) cell line associated genes). However, in HER2 (SkBr3) population, a few common pathways have been found. Once more, consistent results between different validation methods have been established in this study and outcomes of these findings revealed interactions between all these pathways, that might explain the sensitivity of Basal-like (MDA-MB-468) and the less sensitivity of HER2 (SkBr3) breast cancer phenotypes to Metformin treatment.

The **Apoptosis signalling pathway** was significantly up-regulated in both breast cancer phenotypes, with increased cells death in Basal-like (MDA-MB-468) but a smaller proportion of apoptotic cells in HER2 (SkBr3) population (chapter 3). Likewise, the **Wnt signalling pathway** is a critical pathway that regulates cell migration, cell polarity, neural patterning and organogenesis during foetal development, and plays a crucial role in cell fate determination (*Komiya and Habas, 2008*).

Similarly, the **phosphatidylinositol 3-kinase (PI3K) pathway** is a regulator of various cellular processes such as metabolism, apoptosis, growth, proliferation, survival, and cell migration and participates in specialised context-dependent functions (*Chalhoub and Baker, 2009*). The involvement of this pathway in various biological processes within the cell could clarify its up and down-regulation expression upon the different treatment conditions and in both cell lines (MDA-MB-468, and SkBr3). In addition, the resistance to Metformin treatment is likely linked to the deregulation of this pathway.

Finally, the **transforming growth factor-beta (TGF- β)** is a member of a superfamily of cytokines that act on protein kinase receptors found on the plasma membrane and that induce an excess of biological signals that regulate cell growth, death and differentiation, immune response, angiogenesis, and inflammation. Furthermore, TGF- β regulates cell cycle by arresting cell division at the early G1 phase and by downstream activation of pro-apoptotic factors including *death-associated protein kinase (DAPK)*, growth arrest and *DNA damage-inducible 45 (GADD45)* and Bim or Bmf limit, therefore, cancer formation (*Fabregat et al., 2014*). (*Kake et al., 2017*), a study demonstrated for the first time that *death-associated protein kinase 3 (DAPK3)* controls proliferation, migration and invasion of A549 cells via ERK MAPK/c-Myc signalling, as well as it possibly being responsible for tumour growth and metastasis.

Clearly, in Basal-like (MDA-MB-468) population samples, downstream pathways were mostly involved in apoptosis, proliferation, survival and cell migration, which might explain the sensitivity of this population to Metformin treatment and the significant death (apoptosis) of these cells after treatment. Yet, in the HER2 (SkBr3) cell line, one common pathway (Cell Cycle – Apoptosis- TGF- β) has been found downregulated and which is also involved in the regulation

of cell growth and death through G1 phase cell cycle arrest and limits the formation of cancer *via* negative stimulation of *DAPK*, *GADD45*. The latter (*GADD45*) was upregulated in both breast cancer cell lines, which linked to the TGF- β pathway downstream.

In summary, protein expression status agrees with the data obtained by qRT-PCR and gave another indicator that helped the choice of one candidate for further experimental work. In addition, the NanoString data analysis has also emphasized the previous finding involving *PTK2B* in several pathways that contribute to tumour progression and invasion. *PTK2B* represents a significantly up-regulated marker in HER2 phenotype, which is the least sensitive breast cancer subtype to Metformin treatment. *PTK2B* has been involved in invasion and migration in breast cancer cells and might be a potential therapeutic target in HER2 breast cancer. *PTK2B* was selected for further study, according to its oncogenic role in the development of cancer generally, and in breast carcinoma specifically. The next chapter (chapter 6) will examine *PTK2B*'s role in proliferation, invasion, and migration.

CHAPTER 6

Evaluation of the role of *PYK2* in breast cancer invasion and metastasis

6.1. Introduction

Protein Tyrosine Kinase 2 Beta (PTK2B), which is also known as *Proline-rich tyrosine kinase 2 (PYK2)* and *focal adhesion kinase (FAK2)*, is a member of the focal adhesion kinase (*FAK*) family. The expression and phosphorylation of *PYK2* are activated by various transmembrane receptors such as integrin, G-protein -coupled and cytokine receptors activation, and growth factors. In addition, *PYK2* can be stimulated by changes in intracellular Ca^{2+} concentrations (*Ni et al., 2018*). The involvement of *FAK* family kinases in several diseases including cancer, cardiovascular disease, bone disease, fibrosis, rheumatoid arthritis, and neurological disorders has been shown in numerous studies. Besides, accumulating evidence has proved the contributory role of *FAK* and *PYK2* to other proliferative and degenerative diseases. Additionally, many preclinical and clinical trials are testing small molecule inhibitors for *FAK* and *PYK2* as options for cancer treatment (*Murphy et al., 2016*). Moreover, this protein tyrosine Kinase is controlling essential cellular processes such as adhesion, migration, proliferation and cell survival in different cell types. Another study by *Ni et al., 2018*, suggested that breast cancer metastasis could be blocked *via* targeting of either one (*FAK* or *PYK2*) or both kinases. Furthermore, *Genna and Gil-Henn, 2018*, reported that *PYK2* and its closely related focal adhesion *FAK* are involved in the invasion of cancer cells through organising the balance between invadopodia-dependent extracellular matrix invasion and *FAK*-mediated migration. They proposed that *FAK* organized the formation of invadopodium precursor indirectly by sequestering Src to focal adhesions and consequently controlling the tyrosine phosphorylation balance between invadopodia and focal adhesions in invasive cancer cells. *Genna* and his group (2018) found that *PYK2* was expressed in a high level in invasive breast carcinoma. The underlying mechanism that potentiates tumour cell invasiveness by *PYK2* could rely on cortactin, an interactor of *PYK2*. In invasive breast cancer cells, *PYK2* colocalises with cortactin to invadopodia, which interposes, directly or indirectly, epidermal growth factor-induced cortactin tyrosine phosphorylation by *Src-mediated Abl-related gene (Arg)* activation. This results in actin polymerisation in invadopodia, extracellular matrix degradation and tumour cell invasion. Indeed, cell migration is a fundamental process, from simple organisms such as amoeba to a complex organism such as mammals. In multicellular organisms, cell migration is a central

process in development and maintenance of these organisms, including tissue formation during embryonic development, wound healing and immune responses. All above required the coordinated movement of cells in certain directions to specified locations (*Vicente-Manzanares et al., 2005*).

In fact, effective cell invasion is mostly dependent upon cell migration, which determines the ability of cells to become mobile and to navigate through the extracellular matrix within a tissue or to infiltrate neighbouring tissues. Cancer cells that became invasive may spread to secondary sites and form metastases, which is a hallmark of cancer and the most leading cause of mortality among cancer patients. This Phenotype is correlated to high expression levels in several genes that involved in cells motility. This can stimulate the invasion of carcinoma cells by responding to signals from the microenvironment. For that, the molecules that involved in cell migration could be probable therapy targets (*Yamaguchi et al., 2005; Clark and Vignjevic, 2015*).

The focus of this chapter was to determine the role of *PYK2* on cell migration, invasion and proliferation of breast cancer cell lines and in response to Metformin. Downstream pathways that are *PYK2*-dependent involved were also identified. For these purposes, Basal-like (MDA-MB-468), Claudin -Low (MDA-MB-231), and HER2 (SkBr3, and MDA-MB-453) cell lines were chosen for the experimental studies.

6.2. Generation of *PTK2B* knockdown cell lines from Basal-like (MDA-MB-468), Claudin -Low (MDA-MB-231) and HER2 (SkBr3 and MDA-MB-453) breast cancer cell lines

The introduction of small interfering RNAs (siRNAs) or short interfering RNAs (shRNA) into cultured cells provides fast and efficient means of knocking down genes. siRNA has been shown to be effective for short-term gene knockdown, while shRNA provides a more stable gene silencing. However, long-term gene silencing can be achieved by standard clonal selection methodologies using empty control or shRNA lentiviral expressing constructs and selection antibiotic selection such as puromycin. In these experiments, two recommended negative controls have been used, which are the untreated cells (provide a reference point for comparing all other samples) and the lentiviral carrying the empty vector (TRC1.5-pLKO.1-puro). See materials and methods (2.2.5.1 and 2.2.5.2). Cells expressing the empty vector provide an essential reference point for comparison with a gene-specific knockdown in cells expressing the shRNAs.

The Basal-like (MDA-MB-468, Claudin -Low (MDA-MB-231) cell HER2 (SkBr3) and (MDA-MB-453) cell lines were selected for *PTK2B* knockdown. The Her2 phenotype represented the less sensitive breast cancer subtype towards Metformin treatments, while the Basal-like and Claudin -Low was the most sensitive phenotypes (see results in chapter 3).

In these experiments, four individual clones from MISSION™ shRNA Target Set NM_004103 were co-transfected with a lentivirus packaging plasmid into HEK 293T cells. The resulting lentiviral particles were used to infect the human breast cancer cell lines. The lentiviral delivery system ensures a high efficiency and permanent incorporation of the construct in the genome of infected cells. A library of infected clones was created from the MDA-MB-468, MDA-MB-231, SkBr3 and MDA-MB-453 cell lines. Each cell line was transfected with four different *PTK2B* shRNA constructs, and the lentiviral construct *EF1- α -EGFP* (Eukaryotic translation elongation factor 1) was used as a positive control to evaluate the efficiency of lentiviral production and transduction.

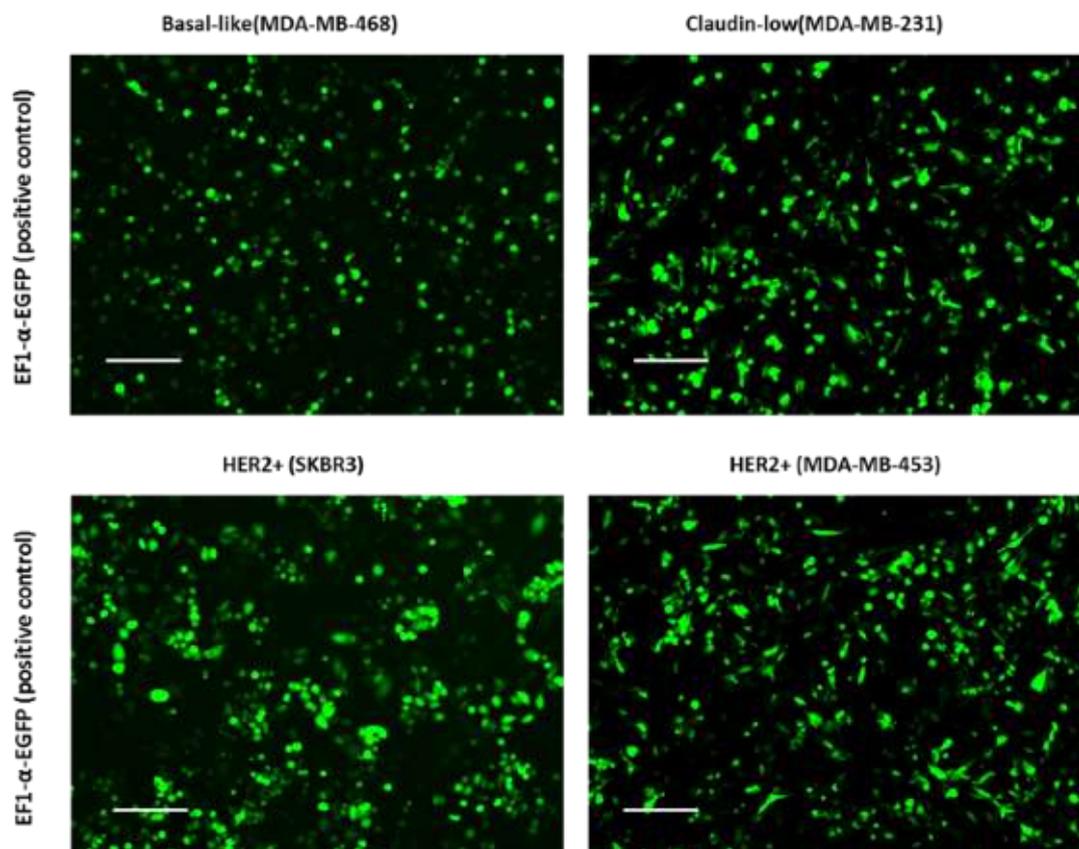


Figure 6.1. Micrographs showing *EGFP* expression in breast cancer cell lines. Cells were transduced with *EF1- α -EGFP* (positive control), Images were taken 48 hours post-transduction. Representative images were taken at 5X magnification. scale bar represents 100 μ m.

Antibiotic selection (puromycin) showed that 3µg/mL is the minimum concentration that killed 100% of the cells after 72h (Figure A.6.1 in Appendix) and three clones out of four MISSION™ shRNA Target Set were efficiently transfected. For this, three clones were used for the upcoming validation. In addition, western blotting experiments were performed to determine the efficiency of PTK2B knockdown in transduced cell lines.

6.3. Analyses of PTK2B protein expression in newly generated Basal-like , Claudin – Low, and HER2 *PTK2B*-knockdown cell lines

The aim of the following section was to determine *PTK2B* protein expression pattern following *PTK2B* knockdown in breast cancer cell lines. Whole protein lysates from shRNA-infected cells were analysed by western blotting using a *PTK2B* antibody (Figure 6.2).

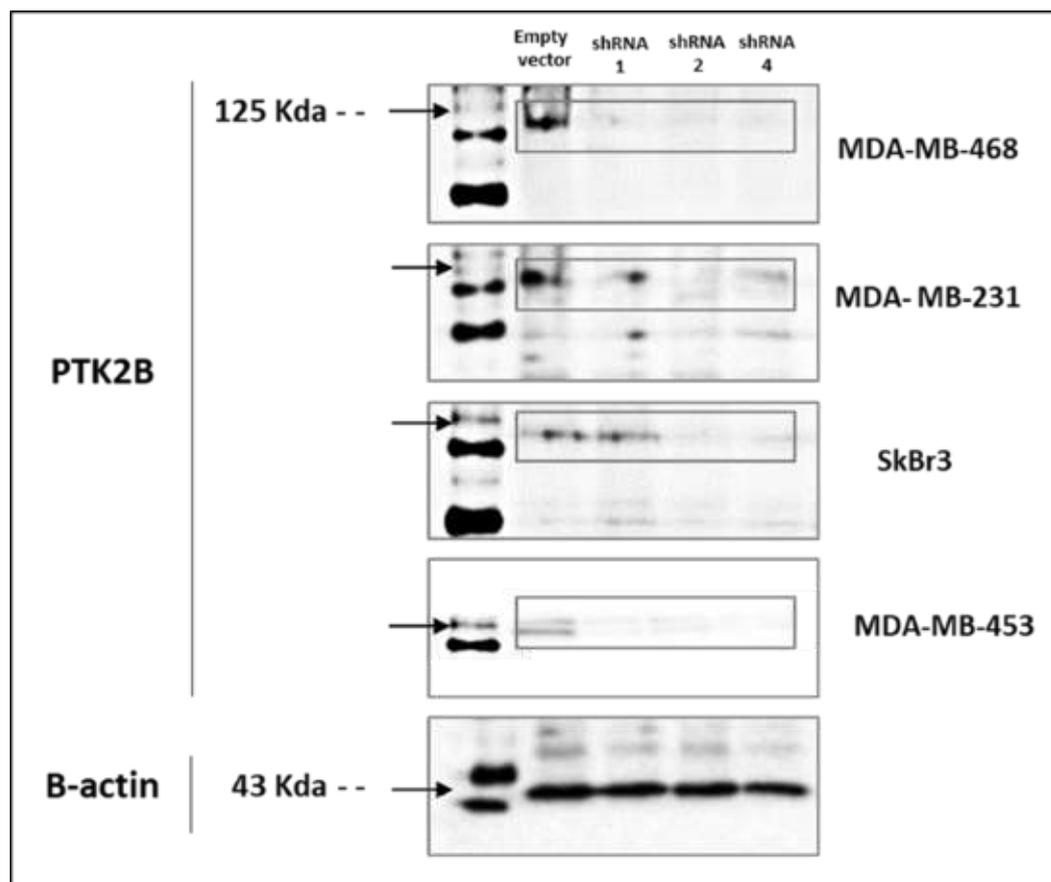


Figure 6.2. Representative images PTK2B expression in Control and PTK2B knockdown breast cancer cell lines. Western blot analysis of whole cell lysates from control and *PTK2B* knockdown from Basal-like, Claudin-Low and HER2 phenotypes. pLKO.1-puro corresponds to the empty vector negative control. β -actin was used as a loading control.

PTK2B was expressed at its highest level in the pLKO.1-puro (Empty vector negative control) in all breast cancer cell lines. Although the expression of *PTK2B* was downregulated in shRNA 2 and 4 expressing cells, its expression was also downregulated in shRNA1 expressing cells except for the HER2 (SkBr3) cell line that also expressed *PTK2B* at a level similar to the control (Figure 6.2). In addition, protein densities of each shRNA bands have been measured in relation to the pLKO.1-puro control bands. Figure (6.3) below reflects the densitometric ratio in Arbitrary Unit (AU) for *PTK2B* protein expression.

These results confirmed that the knockdown system using *PYK2* constructs is efficient and that shRNA2 (clone2) and shRNA4 (clone4) expressing cells can be used for next studies. In the subsequent experimental work, clone2 is referred to as shRNA1 and clone4 as shRNA2.

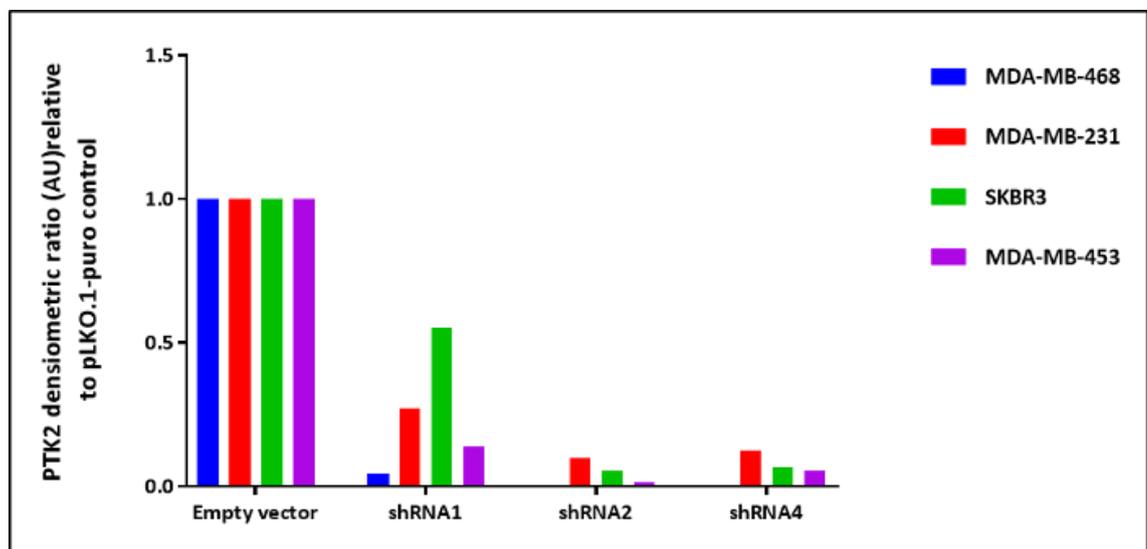


Figure 6.3. Micrograph representing the densitometric ratio of PTK2B protein expression in control and shRNA knockdown cells. PTK2B expression in Arbitrary Unit (AU) in Basal-like (MDA-MB-468), Claudin -Low (MDA-MB-231), HER2 (SkBr3) and (MDA-MB-453) cell lines.

6.4. Investigating the effects of PTK2B knockdown on cell migration, invasion and proliferation of breast cancer cell lines in the absence or presence of treatment

6.4.1. Effect of PTK2B knockdown on cell migration and in response to Metformin treatment

To determine the role of PTK2B in cell migration, a Wound-healing (Scratch) assay was used. This assay was applied to MDA-MB-468, MDA-MB-231, SkBr3, and MDA-MB-453 controls and shRNAs (shRNA1 and 2) in the absence or presence of Metformin. The results demonstrated that PTK2B knockdowns in MDA-MB-468, HER2 (SkBr3) and (MDA-MB-453) cells resulted in decreased migration 24 and 48h following scratching. This reduced migration was also significant when Metformin has used. However, non-significant differences on cell migration capability were observed in MDA-MB-231 cell line when comparing untreated and Metformin treated empty vector and PTK2B expressing cells (Figure 6.4 A, B, Figure 6.5 A, B, Figure 6.6 A, B and Figure 6.7 A, B). In addition, the Wound-healing (Scratch) assay images for MDA-MB-468, MDA-MB-231 and HER2+ breast cancer cell lines (SkBr3, MDA-MB-453) can be found in Appendix Figures A.6.2, A.6.3, A.6.4, A.6.5., A.6.6, A.6.7, A.6.8, A.6.9.

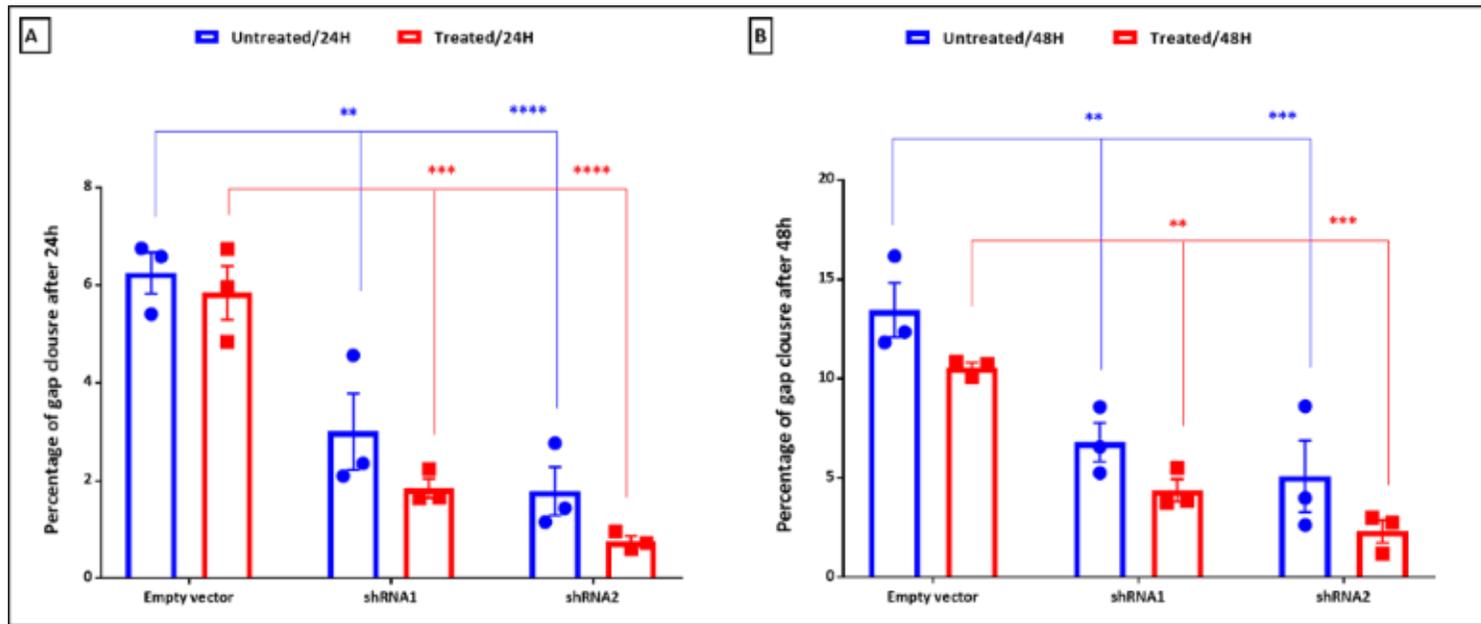


Figure 6.4. PTK2B depletion prevents cell migration in MDA-MB-468 cells. Wound-healing assay (Scratch assay) using MDA-MB-468 cells expressing empty vector, PTK2B- shRNA1 and PTK2B- shRNA2 and corresponding data quantification of gap closure at time point 24 h(A) and 48 h(B) following scratching. Two-way ANOVA P (**-****) $\leq 0.01-0.0001$.

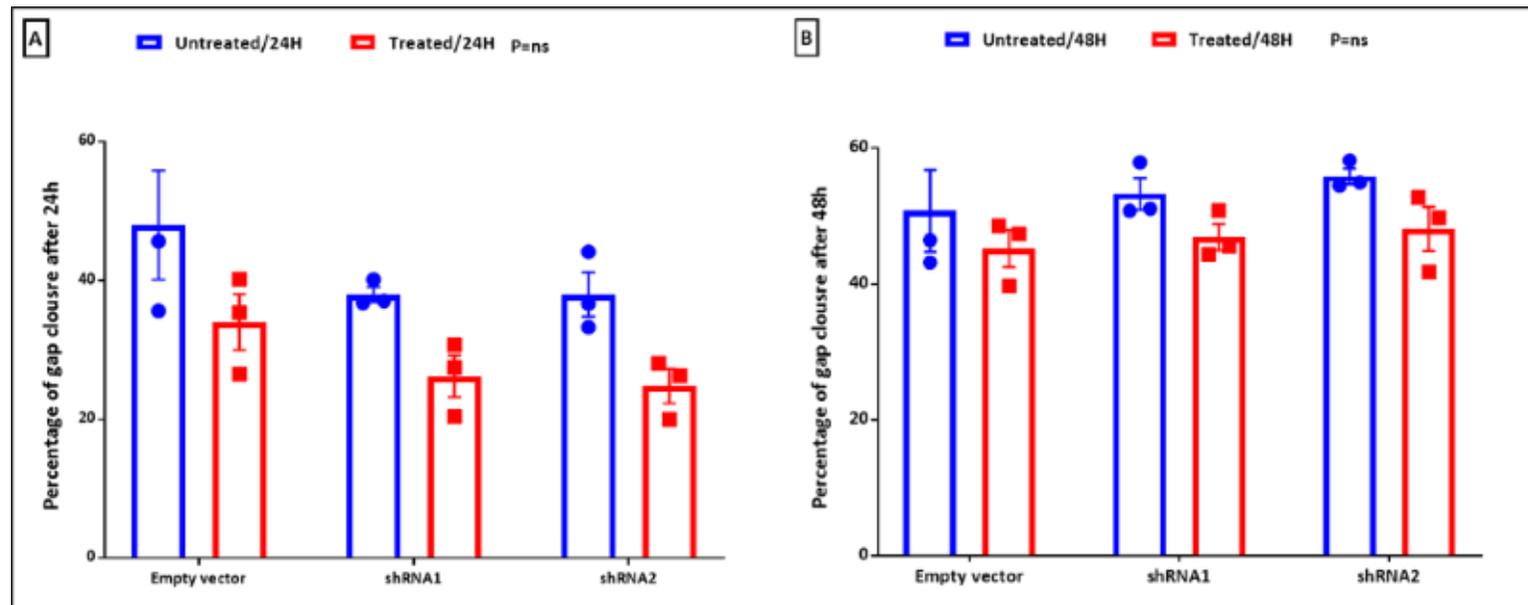


Figure 6.5. PTK2B depletion does not affect cell migration in MDA-MB-231 cells. Wound-healing assay (Scratch assay) using MDA-MB-231 cells expressing empty vector, PTK2B- shRNA1 and PTK2B- shRNA2 and corresponding data quantification of gap closure at time point 24 h(A) and 48 h(B) following scratching. Two-way ANOVA P =(ns).

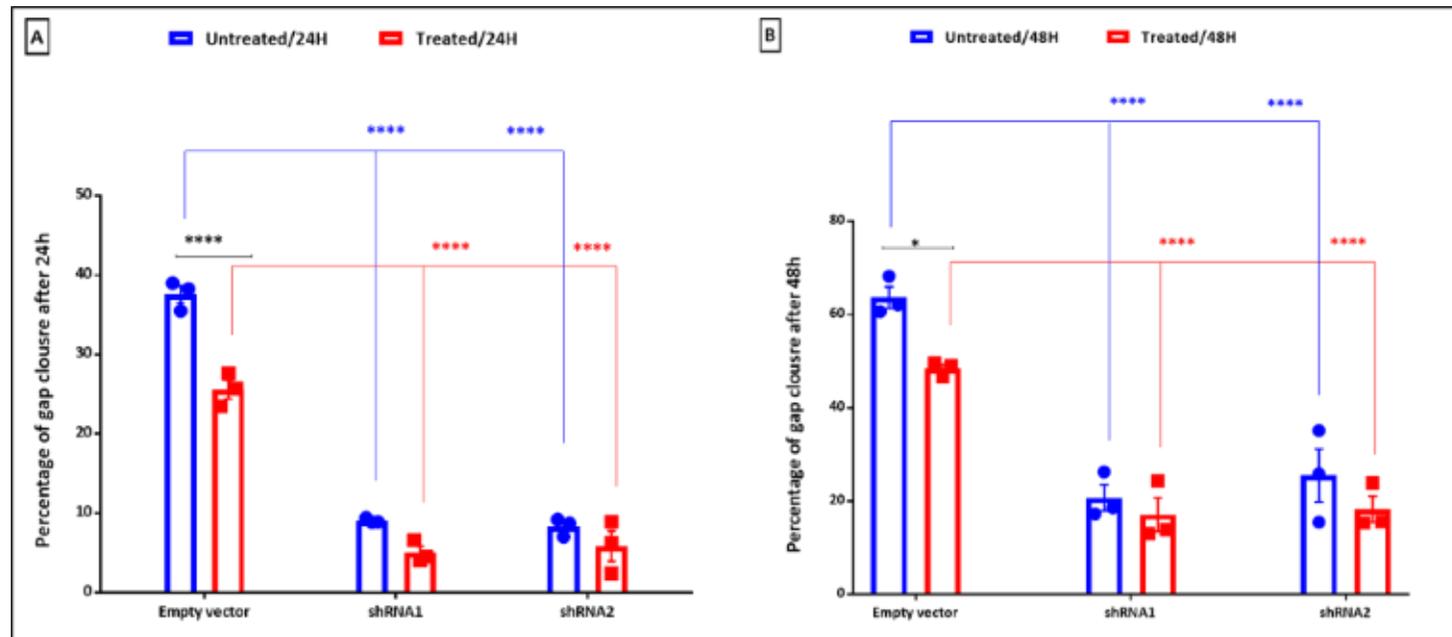


Figure 6.6. PTk2B depletion prevents cell migration in SkBr3 cells. Wound-healing assay (Scratch assay) using SkBr3 cells expressing empty vector, PTk2B- shRNA1 corresponding data quantification of gap closure at time points 24 h(A) and 48 h(B) following scratching. Two-way ANOVA P (****) \leq 0.0001.

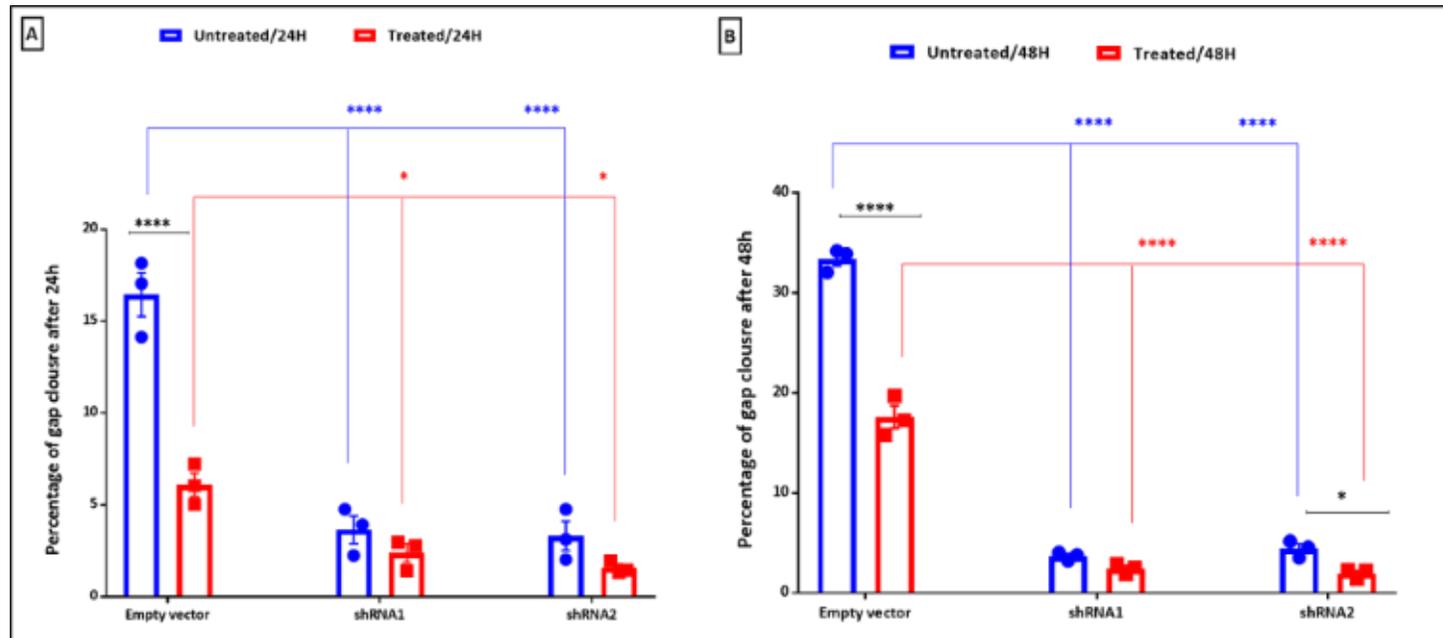


Figure 6.7- PTK2B depletion prevents cell migration in MDA-MB-453 cells. Wound-healing assay (Scratch assay) using MDA-MB-453 cells expressing empty vector, PTK2B- shRNA1 corresponding data quantification of gap closure at time points 24 h(A) and 48 h(B) following scratching. Two-way ANOVA P (*..****) $\leq 0.05 \leq 0.0001$.

6.4.2. Effect of PTK2B knockdown on cell invasion and in response to Metformin treatment

The Cultrex® BME Cell Invasion Assay is a 96-well microplate-based assay designed to accelerate the screening process for compounds that influence chemotaxis, by quantifying the degree to which invasive cells penetrate a barrier consisting of Basement Membrane Extract (BME) in response to chemoattractants and/or inhibiting compounds. This assay was used to validate the invasiveness capacity of PTK2B shRNA knockdown clones compared to the PLKO-1 (empty vector) in both conditions of non-treated and Metformin treatment after 48h of incubation. This assay has been applied to the four breast cancer phenotypes. Figures 6.8, 6.9, 6.10, and 6.11 shown the total number of invading cells in the empty vector compared to shRNA1 and shRNA 2 in the presence or of Metformin treated MDA-MB-468, MDA-MB-231, SkBr3, and MDA-MB-453 cell lines (48h of treatment) and non-treated cells. The results indicated a reduction of invasive capacities of depleted PTK2B cells in all cell lines when compared to cells expressing the empty vector (pLKO.1-puro) except for the MDA-MB-231 cell line which showed a different response. The treatment with Metformin also significantly decreased the invasiveness of treated cells when compared to the control in most cell lines. However, a non-significant difference between the empty vector and PTK2B shRNA clones was noticed in MDA-MB-468 treated cells.

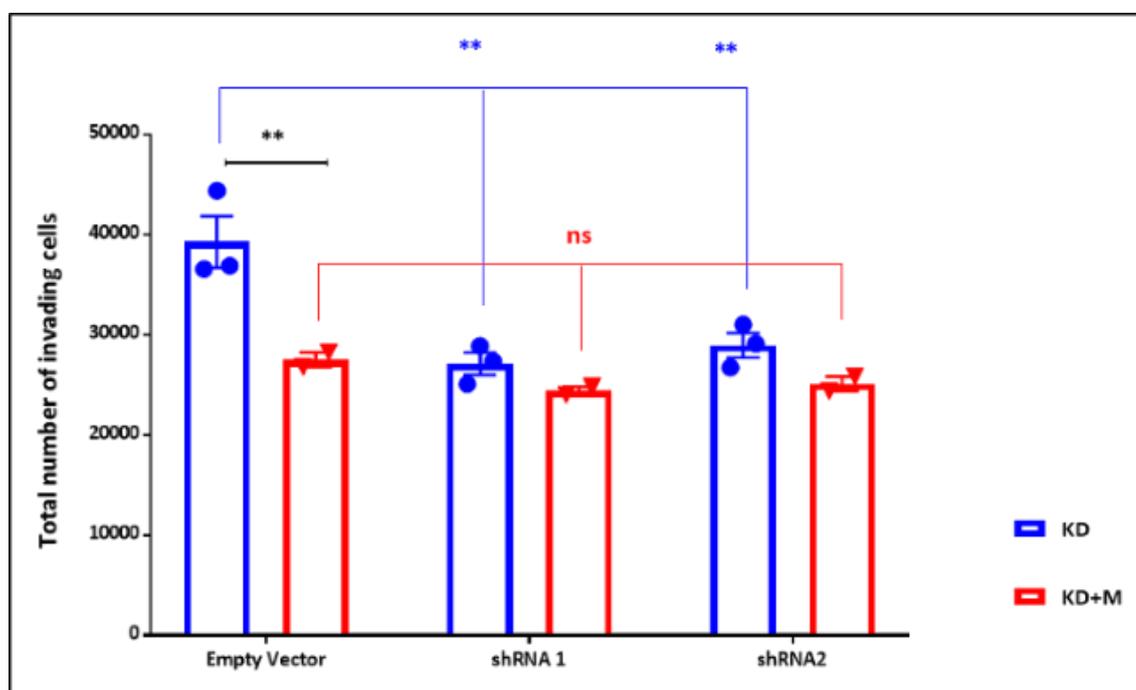


Figure 6.8. PTK2B knockdown prohibited cell invasiveness in MDA-MB-468 cells. Cultrex® BME Cell Invasion Assay using MDA-MB-468 cells expressing empty vector and PTK2B constructs, corresponding data quantification the number of invading cells at time points 48 h following culturing. Two-way ANOVA P (**) ≤ 0.001 .

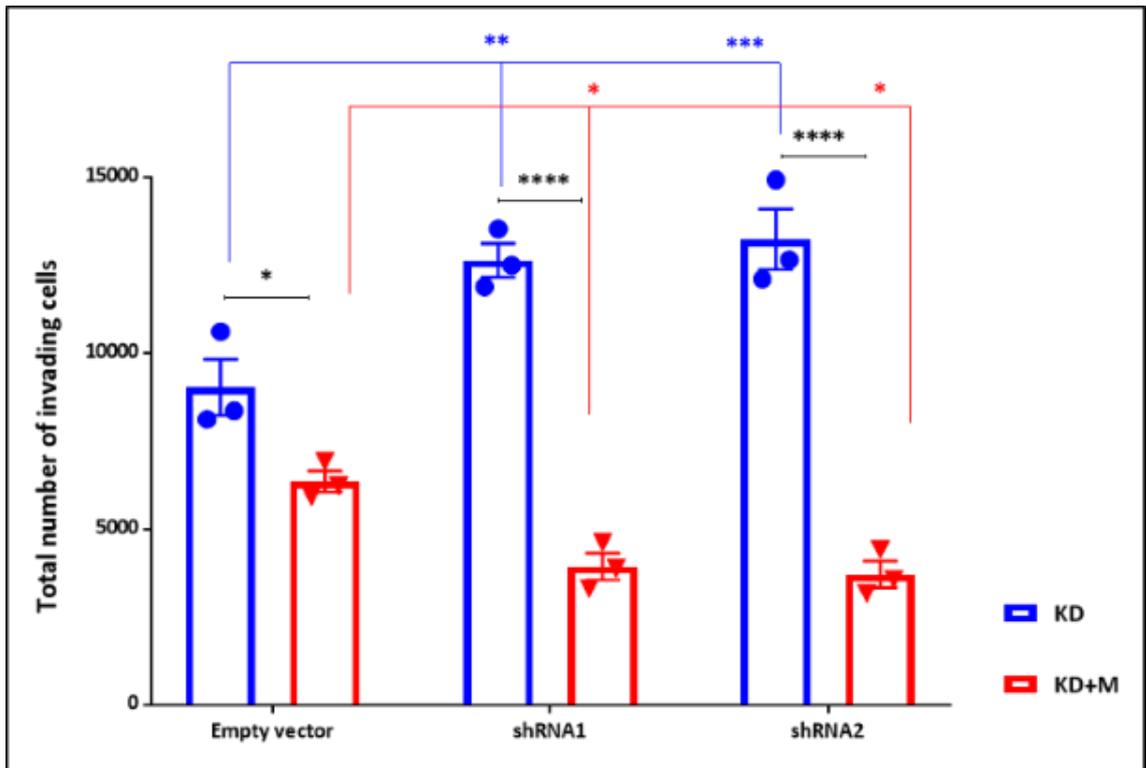


Figure 6.9. PTK2B knockdown prohibited cell invasiveness in MDA-MB-231 cells. Cultrex® BME Cell Invasion Assay using MDA-MB-231 cells expressing empty vector and PTK2B constructs, corresponding data quantification the number of invading cells at time points 48 h following culturing. ANOVA $**P \leq 0.001$, and Two-way ANOVA $P (*) \leq 0.05$.

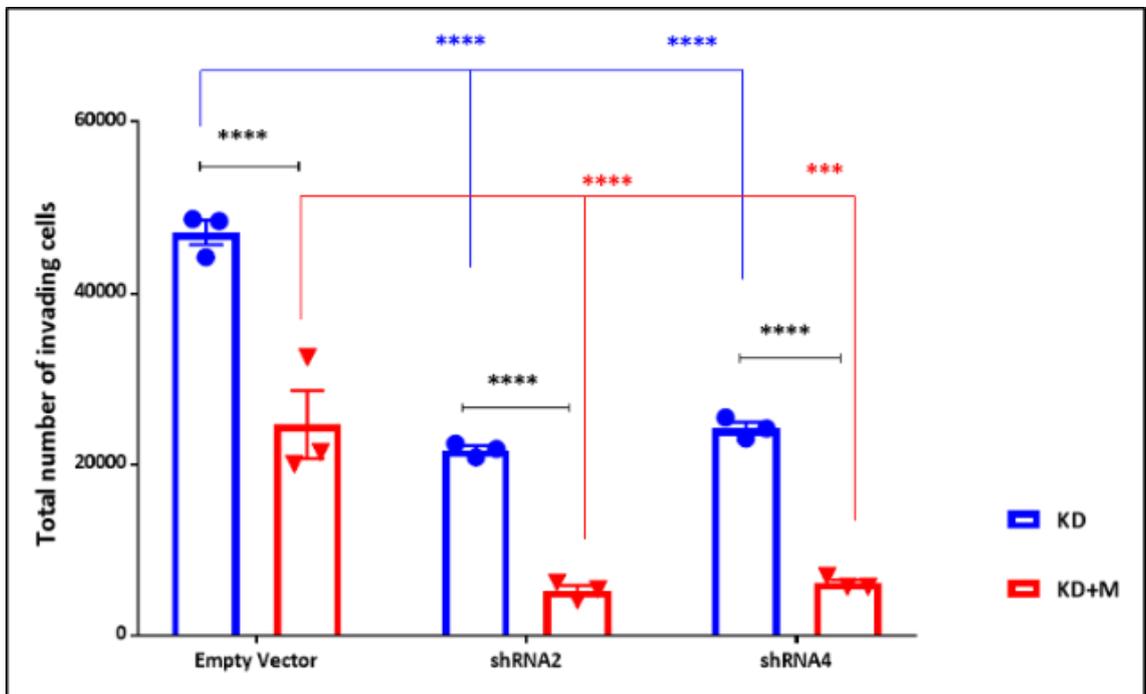


Figure 6.10. PTK2B knockdown prohibited cell invasiveness in SkBr3 cells. Cultrex® BME Cell Invasion Assay using SkBr3 cells expressing empty vector and PTK2B constructs, corresponding data quantification the number of invading cells at time points 48 h following culturing. Two-way ANOVA $P (****) \leq 0.0001$, between the empty vector and PTK2B constructs.

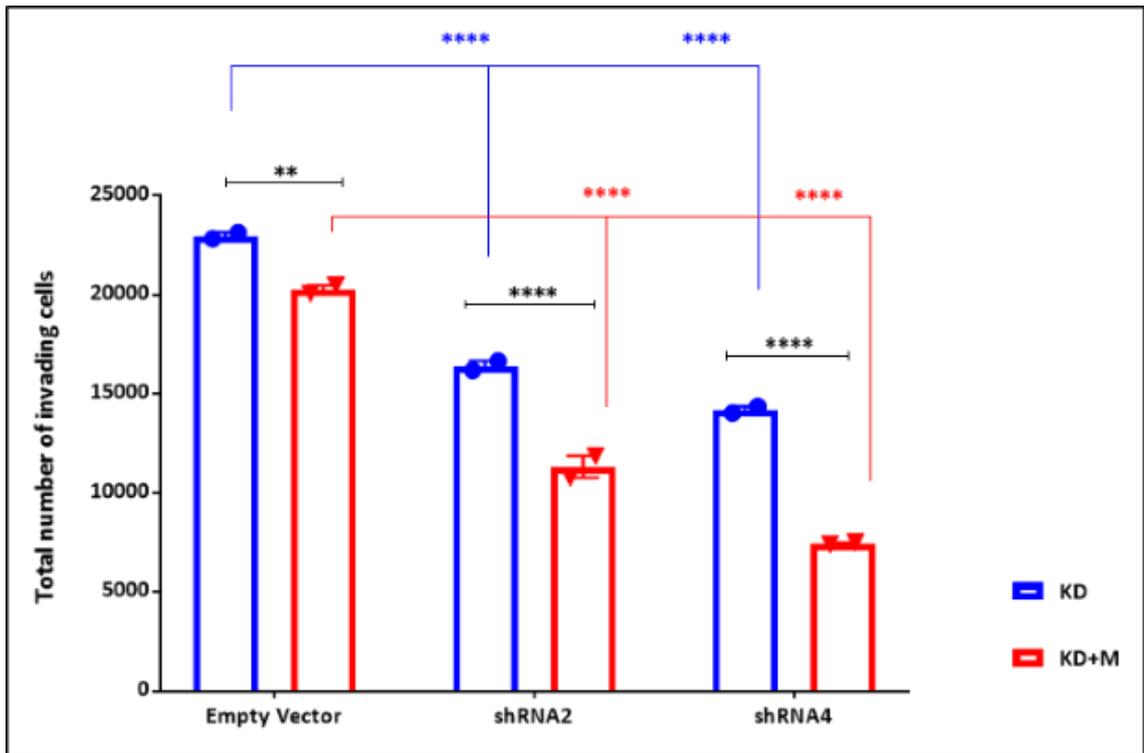
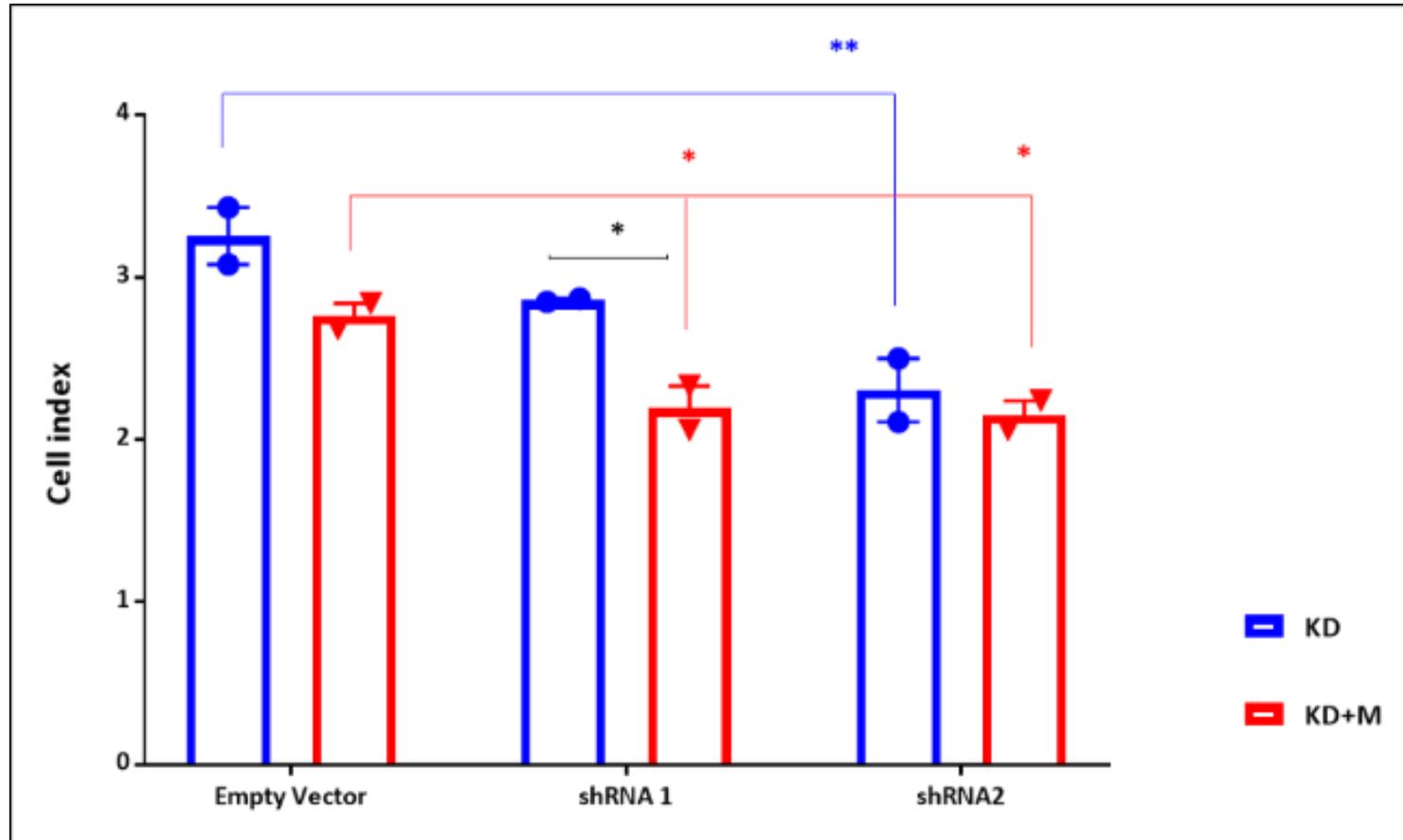


Figure 6.11. PTK2B knockdown prohibited cell invasiveness in MDA-MB-453 cells. Cultrex® BME Cell Invasion Assay using MDA-MB-453 cells expressing empty vector and PTK2B constructs, corresponding data quantification the number of invading cells at time points 48 h following culturing. Two-way ANOVA P (**-****) \leq 0.001 to 0.0001.

6.4.3. Effect of PTK2B knockdown on cell proliferation and in response to Metformin treatment using the xCELLigence System (RTCA)

The xCELLigence system is a pioneer technology method, which provides a dynamic real-time, label-free cellular analysis for a variety of research applications in drug development, toxicology, cancer, medical microbiology and virology. The xCELLigence System is a microelectronic biosensor system for cell-based assays, which monitors cellular events in real time, without the incorporation of labels by measuring electrical impedance across interdigitated microelectrodes integrated on the bottom of its special tissue culture plates (E-Plate). The impedance measurement improves conventional endpoint assays and provides quantitative information about the biological status of the cells, including cell number, adhesion, viability, and morphology. This impedance signal is converted to a specific parameter called Cell Index using a proprietary algorithm. The Cell Index is an excellent measure of how the cells are behaving over time (*Martinez-Serra et al., 2014*).

Cell Proliferation assay has been applied on control and PTK2B knockdowns breast cancer cell lines, utilising the RTCA xCELLigence System and the E-Plate (16) (see chapter 2 materials and method section 2.2.5.9). The Cell Index (CI) was measured every 20 min for 2 h and then every hour until reaching 48h. The results showed that MDA-MB-468 and MDA-MB-231 cells had reduced cell proliferation in PTK2B knockdowns when compared to the empty vector (control), in both treated and untreated conditions. However, a non-significant difference was observed between untreated and treated cells, except for the MDA-MB-468 cell line expressing PTK2B shRNA 1 (Figure. 6.12 and 6.13), while the proliferation of SkBr3 and MDA-MB-453 breast cancer cells was significantly increased following PTK2B knockdowns in comparison with control (empty vector). A similar observation has been made after Metformin treatment with significant differences between empty vector and PTK2B constructs (Fig. 6.14, 6.15) respectively. Furthermore, in this assay Metformin treatment did not appear to affect the proliferation of transduced cells in all cell lines except for the MDA-MB-453 cell line.



Figure

Micrograph representing the effect of PTK2B downregulation on cell proliferation in MDA-MB-468 cells. Data representing measurements of cell proliferation of MDA-MB-468 cells expressing PLKO-1(control) and PTK2B shRNA 1 and 2 after 48 h post treatment with Metformin. Using the xCELLigence System (RTCA) assay, the data represent the number of cells (Cell index) after 48h. Only shRNA2 untreated cells displayed a significant difference between the empty vector and PTK2B shRNAs. Both treated shRNA1 and2 showed a significant difference between treated empty vector and PTK2B constructs. Two-way ANOVA P (*-**) ≤ 0.05 - ≤ 0.01 . In addition, there was a significant difference between the treated and untreated shRNA1 only.

6.12.

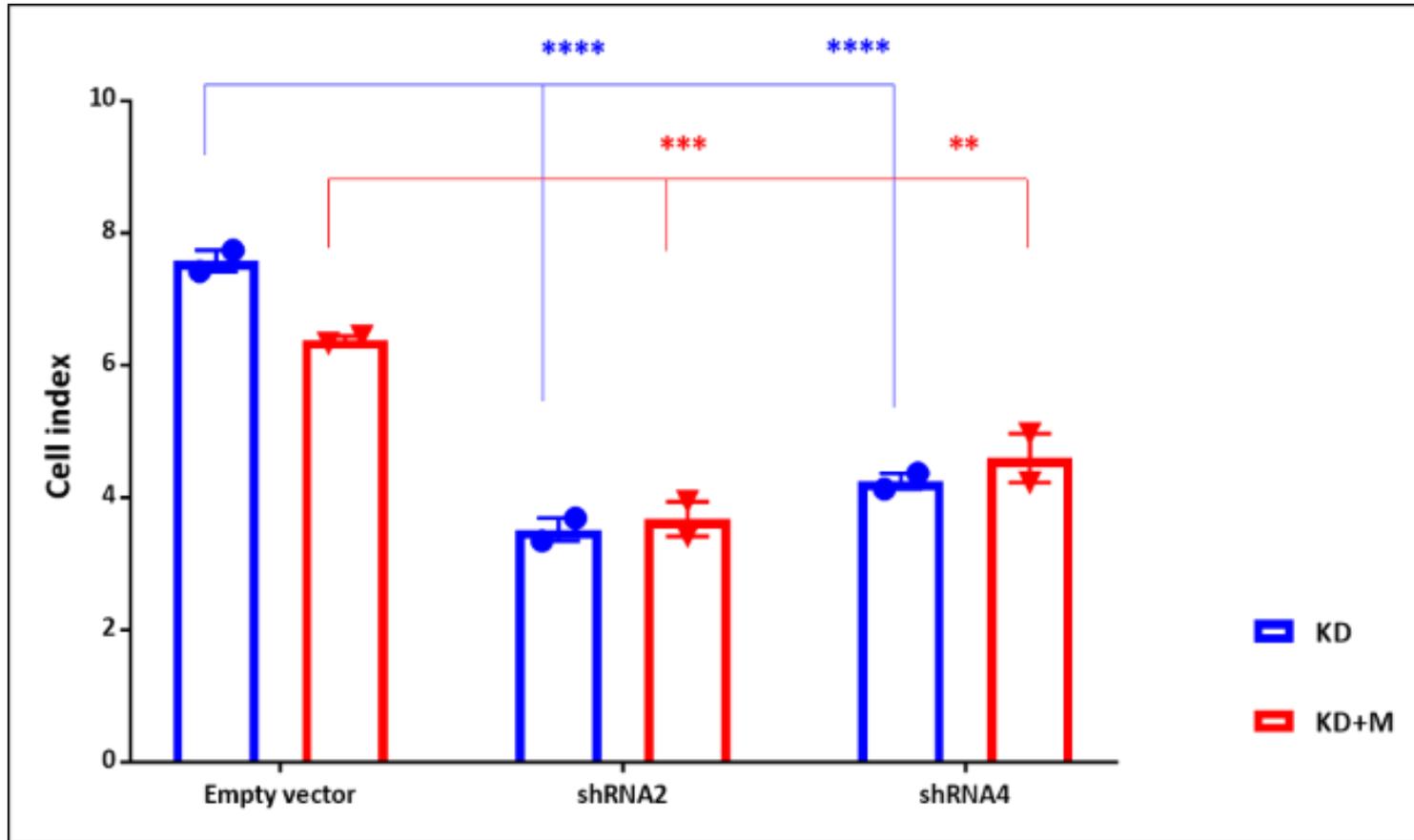


Figure 6.13. Micrograph representing the effect of PTK2B downregulation on cell proliferation in MDA-MB-231 cells. Data representing measurements of cell proliferation of MDA-MB-231 cells expressing PLKO-1(control) and PTK2B shRNA 1 and 2 after 48 h post treatment with Metformin. Using the xCELLigence System (RTCA) assay, the data represent the number of cells (Cell index) after 48h. Both shRNA1 and2 showed a significant difference between empty vector and PTK2B constructs in treated and untreated conditions. However, there were not any significant differences between untreated and treated cells. Two-way ANOVA P (**-****) $\leq 0.01 \leq 0.0001$.

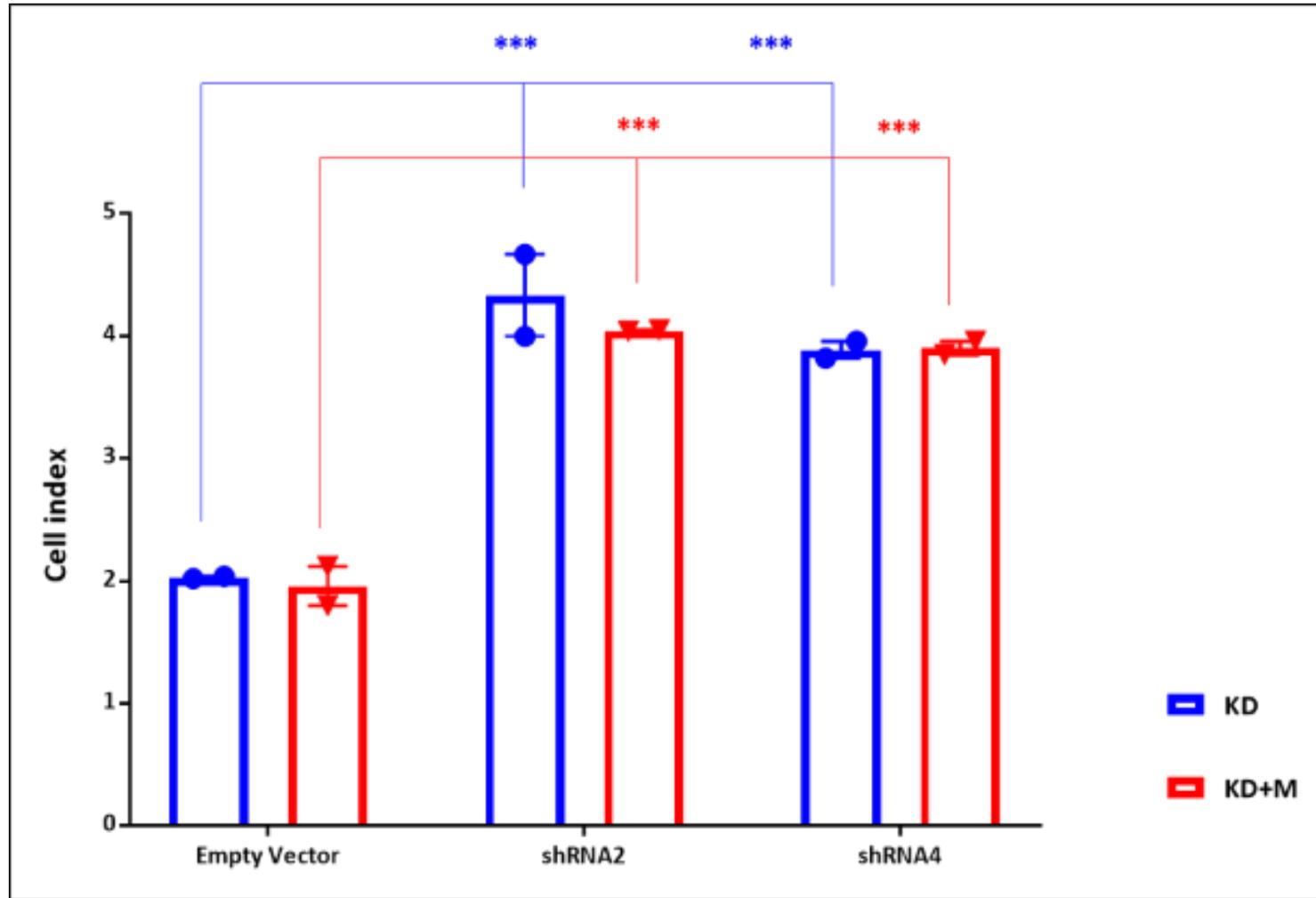


Figure 6.14. Micrograph representing the effect of PTK2B downregulation on cell proliferation in SkBr3 cells. Data representing measurements of cell proliferation of HER2 (SkBr3) cells expressing PLKO-1(control) and PTK2B shRNA 1 and 2 after 48 h post treatment with Metformin. Using the xCELLigence System (RTCA) assay, the data represent the number of cells (Cell index) after 48h. Both shRNA 1 and 2 showed a significant difference between empty vector and PTK2B constructs constructs in treated and untreated conditions. However, a non-significant difference between untreated and treated cells has been observed. Two-way ANOVA P (***) ≤ 0.0001 .

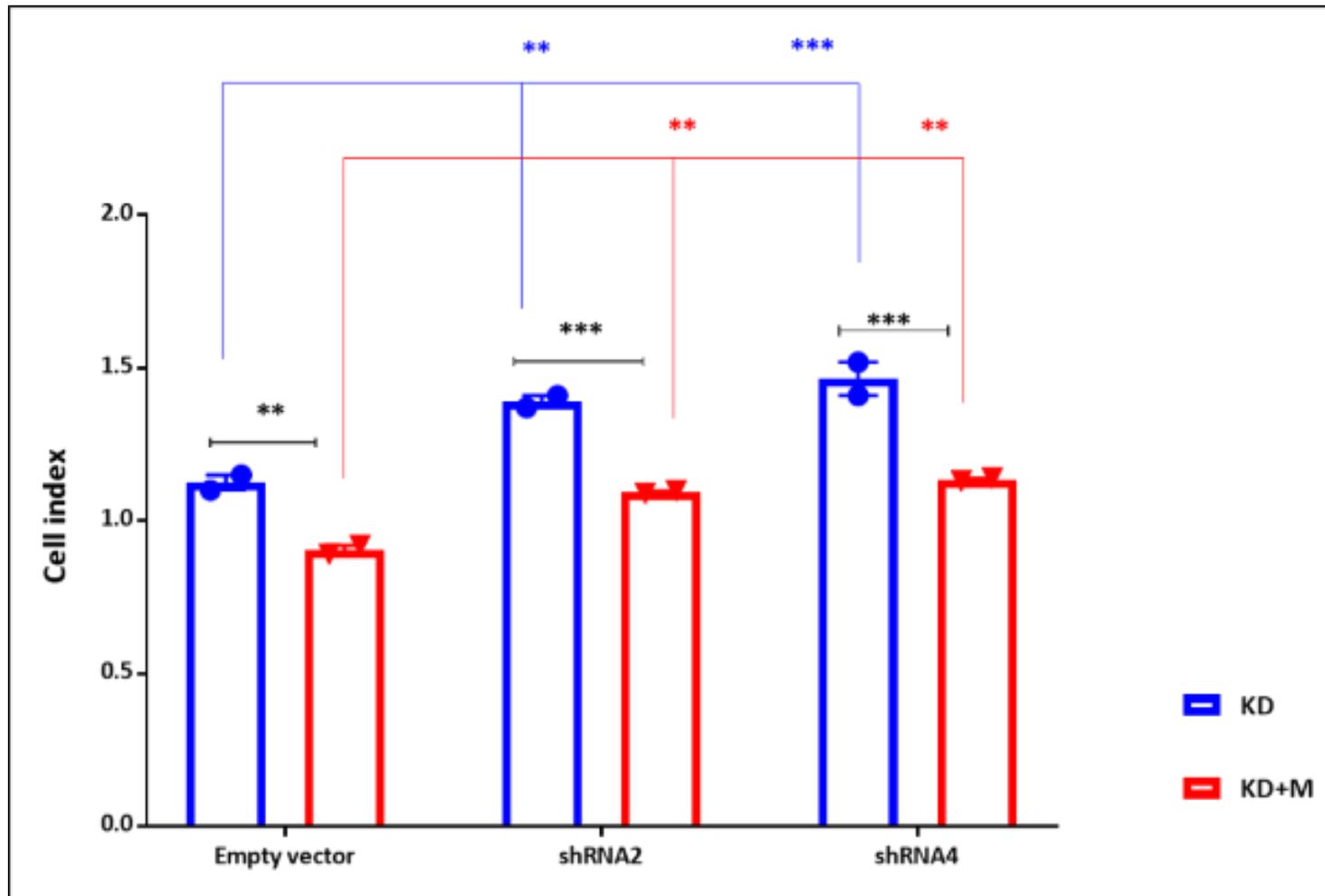


Figure 6.15. Micrograph representing the effect of PTK2B downregulation on cell proliferation in MDA-MB-453 cells. Data representing measurements of cell proliferation of HER2 (MDA-MB-453) cells expressing PLKO-1(control) and PTK2B shRNA 1 and 2 after 48 h post treatment with Metformin. Using the xCELLigence System (RTCA) assay, the data represent the number of cells (Cell index) after 48h. Both shRNA 1 and 2 showed a significant difference between empty vector and PTK2B constructs in both treated and non-treated conditions. In addition, significant differences between untreated and treated cells have been shown. Two-way ANOVA P (**-****) $\leq 0.01 \leq 0.0001$.

6.5. Employment of MetaCore™ software analysis to interrogate PYK2 related pathways and biological processes.

Meta Core™ online software search has been utilised to interrogate the biological pathways that are related to PYK2. The analysis has revealed 46 Normal pathways and 17 Pathological pathways that are related to this gene (PYK2) as presented in Table 6.1 below.

Table 6.1. List of PYK2 engaged Pathways in both Normal and Pathological conditions

Entry number	Pathway	Condition
1	Blood coagulation_ GPIIb-IX-V-dependent platelet activation	Normal
2	Cell adhesion_ Histamine H1 receptor signalling in the interruption of cell barrier integrity	Normal
3	Chemotaxis_ CCL19- and CCL21-mediated chemotaxis	Normal
4	Chemotaxis_ CCL2-induced chemotaxis	Normal
5	Chemotaxis_ CXCR4 signalling pathway	Normal
6	Chemotaxis_ SDF-1/ CXCR4-induced chemotaxis of immune cells	Normal
7	Development_ ACM2 and ACM4 activation of ERK	Normal
8	Development_ Activation of Erk by ACM1, ACM3 and ACM5	Normal
9	Development_ Activation of ERK by Alpha-1 adrenergic receptors	Normal
10	Development_ Activation of ERK by Kappa-type opioid receptor	Normal
11	Development_ Adenosine A2B receptor signalling	Normal
12	Development_ Adenosine A3 receptor signalling	Normal
13	Development_ Alpha-2 adrenergic receptor activation of ERK	Normal
14	Development_ Angiotensin activation of ERK	Normal
15	Development_ Angiotensin signalling via PYK2	Normal
16	Development_ Angiotensin signalling via STATs	Normal
17	Development_ c-Kit ligand signalling pathway during hemopoiesis	Normal
18	Development_ Delta-type opioid receptor signalling via G-protein alpha-14	Normal
19	Development_ Endothelin-1/EDNRA signalling	Normal
20	Development_ Endothelin-1/EDNRA transactivation of EGFR	Normal
21	Development_ G-protein-mediated regulation of MAPK-ERK signalling	Normal

22	Development_ Growth hormone signalling via PI3K/AKT and MAPK cascades	Normal
23	Development_ Osteopontin signalling in osteoclasts	Normal
24	Development_ SDF-1 signalling in hematopoietic stem cell homing	Normal
25	Development_ VEGF signalling via VEGFR2 - generic cascades	Normal
26	G-protein signalling_ G-Protein alpha-q signalling cascades	Normal
27	G-protein signalling_ Regulation of p38 and JNK signalling mediated by G-proteins	Normal
28	Immune response_ IFN gamma signalling pathway	Normal
29	Immune response_ CCR5 signalling in macrophages and T lymphocytes	Normal
30	Immune response_ CXCR4 signalling via the second messenger	Normal
31	Immune response_ Fc gamma R-mediated phagocytosis in macrophages	Normal
32	Immune response_ IL-7 signalling in T lymphocytes	Normal
33	Immune response_ M-CSF-receptor signalling pathway	Normal
34	Immune response_ Neurotensin-induced activation of IL-8 in colonocytes	Normal
35	Immune response_ Role of integrins in NK cells cytotoxicity	Normal
36	Neurophysiological process_ ACM1 and ACM2 in neuronal membrane polarization	Normal
37	Neurophysiological process_ Constitutive and regulated NMDA receptor trafficking	Normal
38	Neurophysiological process_ NMDA-dependent postsynaptic long-term potentiation in CA1 hippocampal neurons	Normal
39	Oxidative stress_ ROS-mediated MAPK activation via canonical pathways	Normal
40	Regulation of lipid metabolism_ Alpha-1 adrenergic receptors signalling via arachidonic acid	Normal
41	Regulation of lipid metabolism_ Stimulation of Arachidonic acid production by ACM receptors	Normal
42	Reproduction_ Gonadotropin-releasing hormone (GnRH) signalling	Normal
43	Stem cells_ Role of PKR1 and ILK in cardiac progenitor cells	Normal
44	Transcription_ Androgen Receptor nuclear signalling	Normal

45	Translation_ Translation regulation by Alpha-1 adrenergic receptors	Normal
46	Transport_ACM3 signalling in lacrimal glands	Normal
47	Autocrine Somatotropin signalling in breast cancer	Pathology
48	CCR7 signalling pathways in dendritic cells in allergic contact dermatitis	Pathology
49	CHDI_ Correlations from Replication data_ Causal network (positive correlations)	Pathology
50	FGFR3 signalling in multiple myeloma	Pathology
51	G protein-coupled receptors signalling in lung cancer	Pathology
52	HBV mediates angiogenesis in HCC	Pathology
53	HBV signalling via protein kinases leading to HCC	Pathology
54	HBV-dependent NF-kB and PI3K/AKT pathways leading to HCC	Pathology
55	HGF receptor (Met) and MSP receptor (RON) signalling pathways in SCLC	Pathology
56	HIV-1 signalling via CCR5 in macrophages and T lymphocytes	Pathology
57	Neuroprotective action of lithium	Pathology
58	The proliferative action of Gastrin in pancreatic cancer	Pathology
59	Role of neuropeptides in the pathogenesis of SCLC	Pathology
60	Stem cells_ Pancreatic cancer stem cells in tumour metastasis	Pathology
61	Tissue Factor signalling in cancer via PAR1 and PAR2	Pathology
62	The transition of HCC cells to an invasive and migratory phenotype	Pathology
63	VEGF signalling in multiple myeloma	Pathology

Pathways were derived by Meta Core™ online software analysis. The table shows 46 Normal pathways and 17 Pathological pathways. The Blue colour indicates Normal pathways and the red Pathological pathways.

The table above demonstrated several normal pathways, as well as pathological pathways that are engaged with PYK2. In addition, the Autocrine Somatotropin signalling pathway has been noticed as the main pathway that is related to breast cancer. Somatotropin, which is also known as Growth hormone (GH), or human growth hormone (HGH) is an important development peptide hormone that stimulates growth, cell reproduction, and cell regeneration in human and other animals. It is a mitogen type that is specific to only certain kinds of cells (*Subramani et al., 2017*). The somatotropic cells in the anterior pituitary gland, are the factory store and secretor of this hormone. In addition, GH is a stress hormone that boosts the concentration of glucose and free fatty acids as well as, stimulating the production of IGF-1. Epidemiological data demonstrate that GH/IGF-1 is associated with an increased risk of breast cancer directly and positively (*Subramani et al., 2017*). Besides, as shown in Figure 6.16 the overexpression of Somatotropin (GH), stimulates the JAK2 pathway, which induces the PYK2 pathway via different pathways. Some pathways lead to cell survival and cell proliferation, which are drivers of breast cancer, while, other pathways lead to migration and spreading of cells from one side, and angiogenesis from another side, which all are responsible for metastasis in breast cancer. The PYK2 is an essential driver gene in this pathologic pathway in breast cancer.

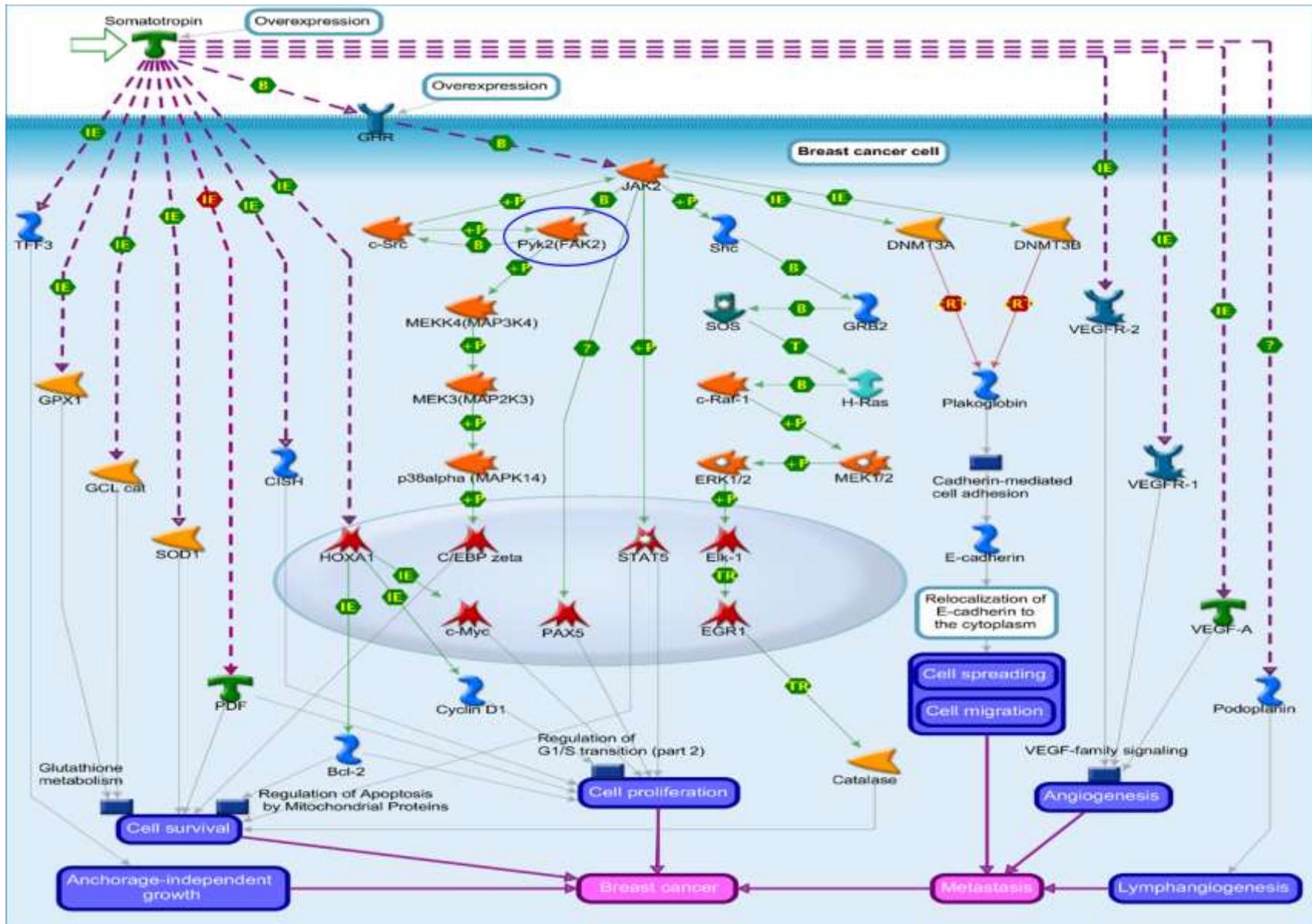


Figure 6.16. PYK2 involvement in Autocrine Somatotropin signalling pathway in breast cancer. The chart was obtained using Meta Core™ online software analysis (on 06/08/2018). Somatotropin (GH) stimulates the JAK2 pathway, which induces the PYK2 pathway via different pathways. All stimulated pathways are associated with breast cancer and metastasis. MetaCore annotation: IE=Influence on expression, B= Binding, T= Transformation, CRT= Co-regulation of transcription, Tr= Transcription regulation.

Furthermore, the biological processes that are related to PYK2 were also analysed by Meta Core™ online software analysis. This search revealed 18 biological processes that are associated to PYK2, including Apoptosis, Cell adhesion, Immune response, Inflammation, Neurophysiological process, Proliferation, Reproduction, Signal Transduction, and Transport (Table 6.2.).

Table 6.2. list of biological processes that are related to *PYK2(FAK2)* gene

Entry number	Processes
1	Apoptosis_ Apoptosis stimulation by external signals
2	Cell adhesion_ Glycoconjugates
3	Cell adhesion_ Leucocyte chemotaxis
4	Cell adhesion_ Platelet aggregation
5	Immune response_ Phagocytosis
6	Immune response_ Phagosome in antigen presentation
7	Inflammation_ IFN-gamma signalling
8	Inflammation_ NK cell cytotoxicity
9	Neurophysiological process_ GABAergic neurotransmission
10	Neurophysiological process_ Long-term potentiation
11	Proliferation_ Lymphocyte proliferation
12	Proliferation_ Positive regulation cell proliferation
13	Reproduction_ FSH-beta signalling pathway
14	Reproduction_ GnRH signalling pathway
15	Reproduction_ Gonadotropin regulation
16	Signal Transduction_ Cholecystokinin signalling
17	Signal Transduction_ TGF-beta, GDF and Activin signalling
18	Transport_ Calcium transport

18 fundamental biological and molecular processes that are crucially involved in cell fate and cancer development were identified. These were derived from Meta Core™ online software analysis.

6.6. Protein expression profiling of samples from Her2 cell lines expressing pLKO.1 (control) and PTK2B shRNAs

Mass spectrometry (MS) analyses were performed to identify and quantify proteins that are expressed in the HER2 breast cancer cell line SkBr3 PLKO-1(control) and PTK2B knockdown (KD) samples. The proteomic method allows identifying the consistency of the proteins that are differentially expressed by control and PTK2B knockdown cells that were treated or untreated with Metformin. In this experiment, whole cell protein lysates from SkBr3 and MDA-MB-453 cells expressing pLKO.1 and shRNA 1 (treated and untreated) were prepared. The protein amount of 6 µg were analysed (pLKO.1 and shRNA 1 treated and untreated) to ensure high peptide identification. The samples were analysed *via* SCIEX Triple TOF 6600 mass spectrometer, which generates data obtained in both IDA (Independent Data Acquisition) and SWATH (sequential window acquisition of all theoretical fragment ion spectra) modes.

Protein Pilot (version 5) software was utilised for a generation of a spectral library of identified proteins from the IDA data acquisition and the spectral library, and which aligned with the obtained SWATH data using the Peak View software (version 2.2). SCIEX One Omics platform was also used to analyse the data following library alignment for fold change of proteins expression definition. A fold change >2 and a confidence level > 0.65 cut-off were used to determine significantly regulated proteins between sample groups. For the generation of heat maps and pathway analysis, the confidence threshold was reduced to > 70 percentage to identify more protein IDs.

A comprehensive and complete record of all forerunner fragmentations from detected peptides and in biological samples can be obtained and examined by utilising the MS/MS SWATH independent data acquisition method, which allows a sensitive detection and quantification of several peptides including less abundant ones. The global protein profiling revealed 3284 proteins IDs in SkBr3 cell line (all replicates). Top 25 upregulated and top 25 downregulated proteins according to the Fold Change $F_c \geq 2$ and confidence of 70 %, have been considered for heat maps generation (Figure 6.17), and which displays comparison between pLKO-1 and PYK2 Knockdown samples from untreated (left-hand heat map) and pLKO-1 vs. PYK2 Knockdown (KD) from Metformin treated samples (right-hand heat map).

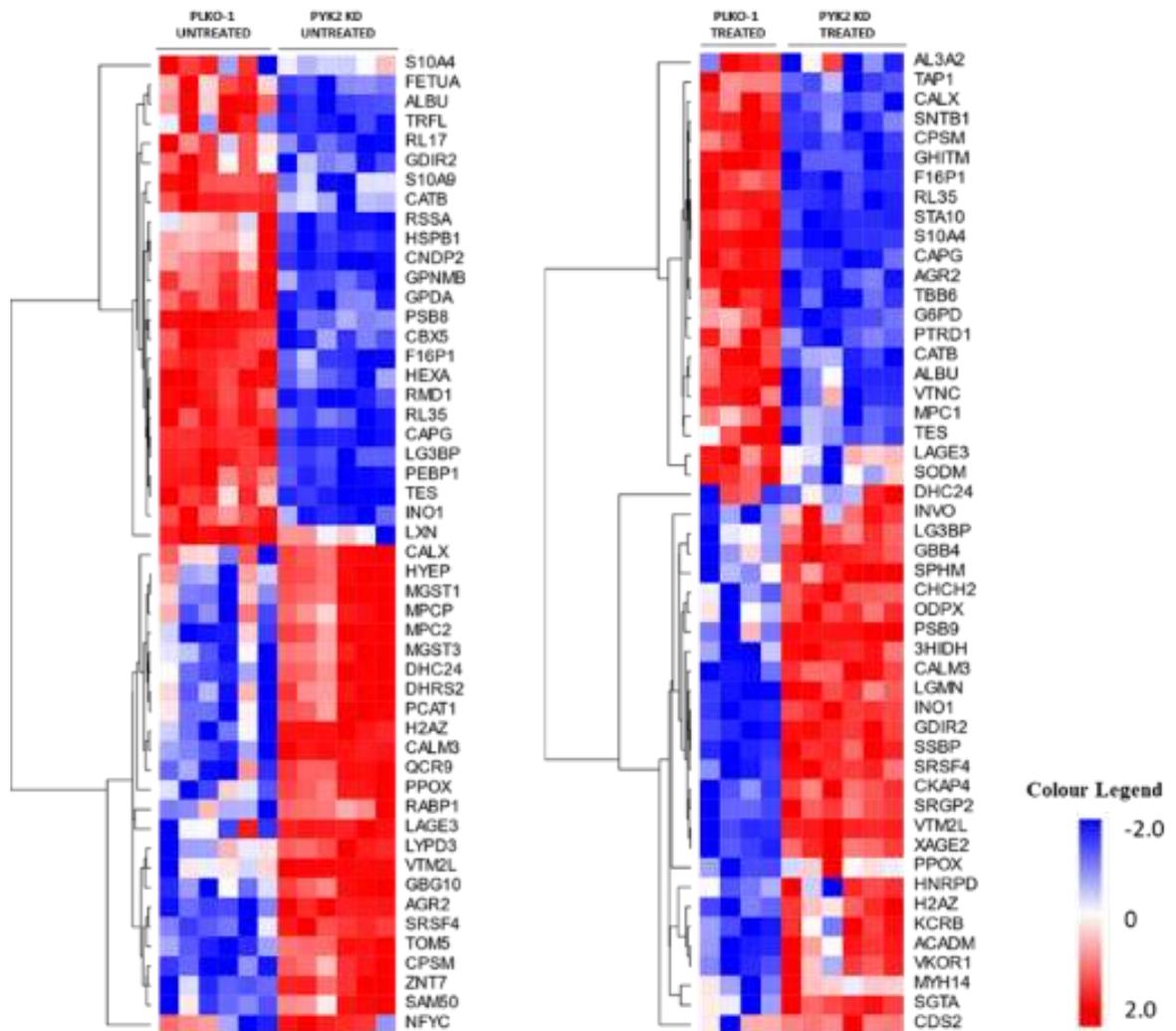


Figure 6.17. Differentially expressed proteins from untreated and treated SkBr3 control and PTK2B cells. The left-hand heat map shows the comparison between untreated PLKO-1 and PYK2 KD samples. The right-hand heat map represents the comparison between Metformin-treated pLKO-1 and PYK2 KD samples. The heat maps display the top 25 upregulated and 25 downregulated proteins according to $F_c \geq -2$, and confidence of 70%. The colour gradient is between red and blues with white in the middle.

Table 6.3 shows the list of proteins that satisfied the cut-off criteria $2 Fc \geq -2$, and confidence of 70% comparing untreated PLKO-1 and PYK2 KD.

Table 6.3. Top 25 upregulated proteins and 25 downregulated proteins in HER2 (SkBr3) cell line

Protein Symbol	Description	Fold change	Confidence level
TOM5	Mitochondrial import receptor subunit TOM5 homolog	18.349	0.762
SRSF4	Serine/arginine-rich splicing factor 4	14.853	0.758
H2AZ	Histone H2A.Z	12.249	0.895
AGR2	Anterior gradient protein 2 homolog	11.426	0.7908
CPSM	Carbamoyl-phosphate synthase [ammonia], mitochondrial	10.967	0.772
GBG10	Guanine nucleotide-binding protein G(I)/G(S)/G(O) subunit gamma-10	10.598	0.715
LAGE3	EKC/KEOPS complex subunit LAGE3	9.374	0.823
MGST1	Microsomal glutathione S-transferase 1	8.617	0.804
NFYC	Nuclear transcription factor Y subunit gamma	8.213	0.785
PPOX	Protoporphyrinogen oxidase	7.198	0.948
HYEP	Epoxide hydrolase 1	7.111	0.824
CALM3	Calmodulin-3	7.052	0.718
LYPD3	Ly6/PLAUR domain-containing protein 3	6.878	0.789
SAM50	Sorting and assembly machinery component 50 homolog	6.659	0.760
MPC2	Mitochondrial pyruvate carrier 2	6.587	0.789
RABP1	Cellular retinoic acid-binding protein 1	6.252	0.795
CALX	Calnexin	5.822	0.787
QCR9	Cytochrome b-c1 complex subunit 9	5.763	0.719
MPCP	Phosphate carrier protein, mitochondrial	5.503	0.710
DHC24	Delta (24)-sterol reductase	5.376	0.772
PCAT1	Lysophosphatidylcholine acyltransferase 1	5.083	0.816
VTM2L	V-set and transmembrane domain-containing protein 2-like protein	4.847	0.789
DHRS2	Dehydrogenase/reductase SDR family member 2, mitochondrial	4.753	0.876
MGST3	Microsomal glutathione S-transferase 3	4.579	0.776
ZNT7	Zinc transporter 7	4.540	0.875

Protein Symbol	Description	Fold change	Confidence level
PEBP1	Phosphatidylethanolamine-binding protein 1	-2.890	0.902
RSSA	40S ribosomal protein SA	-2.946	0.885
HSPB1	Heat shock protein beta-1	-2.962	0.844
HEXA	Beta-hexosaminidase subunit alpha	-3.057	0.821
RMD1	Regulator of microtubule dynamics protein 1	-3.133	0.822
GDIR2	Rho GDP-dissociation inhibitor 2	-3.153	0.765
TES	Testing	-3.195	0.843
CNDP2	Cytosolic non-specific dipeptidase	-3.228	0.753
RL17	60S ribosomal protein L17	-3.229	0.710
S10A9	Protein S100-A9	-3.563	0.741
CBX5	Chromobox protein homolog 5	-3.832	0.755
CATB	Cathepsin B	-3.832	0.806
INO1	Inositol-3-phosphate synthase 1	-3.844	0.739
GPDA	Glycerol-3-phosphate dehydrogenase [NAD (+)], cytoplasmic	-4.107	0.761
S10A4	Protein S100-A4	-4.239	0.721
GPNMB	Transmembrane glycoprotein NMB	-4.831	0.724
LXN	Latexin	-5.566	0.727
TRFL	Lactotransferrin	-5.939	0.806
FETUA	Alpha-2-HS-glycoprotein	-8.989	0.703
RL35	60S ribosomal protein L35	-9.105	0.831
PSB8	Proteasome subunit beta type-8	-10.789	0.747
LG3BP	Galectin-3-binding protein	-16.712	0.906
CAPG	Macrophage-capping protein	-17.133	0.883
ALBU	Serum albumin	-17.665	0.843
F16P1	Fructose-1,6-bisphosphatase 1	-64.565	0.816

Samples comparing untreated PLKO-1 and PYK2 KD samples with Fc ≥ -2 , and confidence of 70%. Red indicated upregulated proteins and blue downregulated proteins.

Table 6.4 shown the list of proteins that satisfied the cut-off criteria $Fc \geq -2$, and confidence of 70% comparing between PLKO-1 and PYK2 KD samples after treatment with Metformin for 48h.

Table 6.4. Top 25 upregulated proteins and 25 downregulated proteins in HER2 (SkBr3) cell line

Protein Symbol	Description	Fold change	Confidence level
ODP	Pyruvate dehydrogenase protein X component, mitochondrial	24.007	0.719
CDS2	Phosphatidate cytidyltransferase 2	23.202	0.741
AGR2	Anterior gradient protein 2 homolog	16.920	0.854
PPOX	Protoporphyrinogen oxidase	10.232	0.843
LGMN	Legumain	9.976	0.748
SRSF4	Serine/arginine-rich splicing factor 4	8.652	0.781
CALM3	Calmodulin-3	8.365	0.765
LAGE3	EKC/KEOPS complex subunit LAGE3	7.965	0.709
XAGE2	X antigen family member 2	7.816	0.710
VTM2L	V-set and transmembrane domain-containing protein 2-like protein	7.273	0.846
KCRB	Creatine kinase B-type	5.427	0.827
SRGP2	SLIT-ROBO Rho GTPase-activating protein 2	4.412	0.707
DHC24	Delta (24)-sterol reductase	4.195	0.835
CPSM	Carbamoyl-phosphate synthase [ammonia], mitochondrial	4.104	0.890
MPC1	Mitochondrial pyruvate carrier 1	3.965	0.837
HNRPD	Heterogeneous nuclear ribonucleoprotein D0	3.540	0.715
GBB4	Guanine nucleotide-binding protein subunit beta-4	3.367	0.785
H2AZ	Histone H2A.Z	3.150	0.999
SNTB1	Beta-1-syntrophin	3.095	0.894
G6PD	Glucose-6-phosphate 1-dehydrogenase	3.081	0.915
SGTA	Small glutamine-rich tetratricopeptide repeat-containing protein alpha	3.071	0.786
STA10	PCTP-like protein	3.066	0.851
AL3A2	Fatty aldehyde dehydrogenase	3.036	0.927
VKOR1	Vitamin K epoxide reductase complex subunit 1	2.997	0.912
CALX	Calnexin	2.961	0.778

Protein Symbol	Description	Fold change	Confidence level
3HIDH	3-hydroxybutyrate dehydrogenase, mitochondrial	-2.722	0.938
CATB	Cathepsin B	-2.951	0.790
VTNC	Vitronectin	-3.047	0.796
CKAP4	Cytoskeleton-associated protein 4	-3.112	0.889
TES	Testing	-3.172	0.897
SSBP	Single-stranded DNA-binding protein, mitochondrial	-3.236	0.788
GDIR2	Rho GDP-dissociation inhibitor 2	-3.283	0.774
CHCH2	Coiled-coil-helix-coiled-coil-helix domain-containing protein 2	-3.391	0.736
SODM	Superoxide dismutase [Mn], mitochondrial	-3.465	0.824
ACADM	Medium-chain specific acyl-CoA dehydrogenase, mitochondrial	-3.879	0.876
SPHM	N-sulphoglucosamine sulphohydrolase	-3.885	0.872
RL35	60S ribosomal protein L35	-3.972	0.828
S10A4	Protein S100-A4	-4.973	0.736
MYH14	Myosin-14	-5.164	0.809
INO1	Inositol-3-phosphate synthase 1	-5.182	0.830
GHITM	Growth hormone-inducible transmembrane protein	-5.510	0.789
TAP1	Antigen peptide transporter 1	-5.659	0.769
ALBU	Serum albumin	-7.230	0.856
TBB6	Tubulin beta-6 chain	-8.234	0.766
PTRD1	Putative peptidyl-tRNA hydrolase PTRHD1	-8.256	0.707
INVO	Involucrin	-9.613	0.825
PSB9	Proteasome subunit beta type-9	-10.508	0.814
LG3BP	Galectin-3-binding protein	-14.193	0.838
CAPG	Macrophage-capping protein	-20.253	0.858
F16P1	Fructose-1,6-bisphosphatase 1	-40.187	0.841

Samples comparing Metformin-treated pLKO-1 and PYK2 KD samples with $F_c \geq -2$, and confidence of 70%. Red indicated upregulated proteins and blue downregulated proteins.

In addition, Figure 6.18 shows the comparison between PLKO-1 untreated samples (control) and PLKO-1 treated (with Metformin) samples (left-hand heat map), and PLKO-1 untreated (control) samples vs. PYK2 KD treated samples (right-hand heat map).

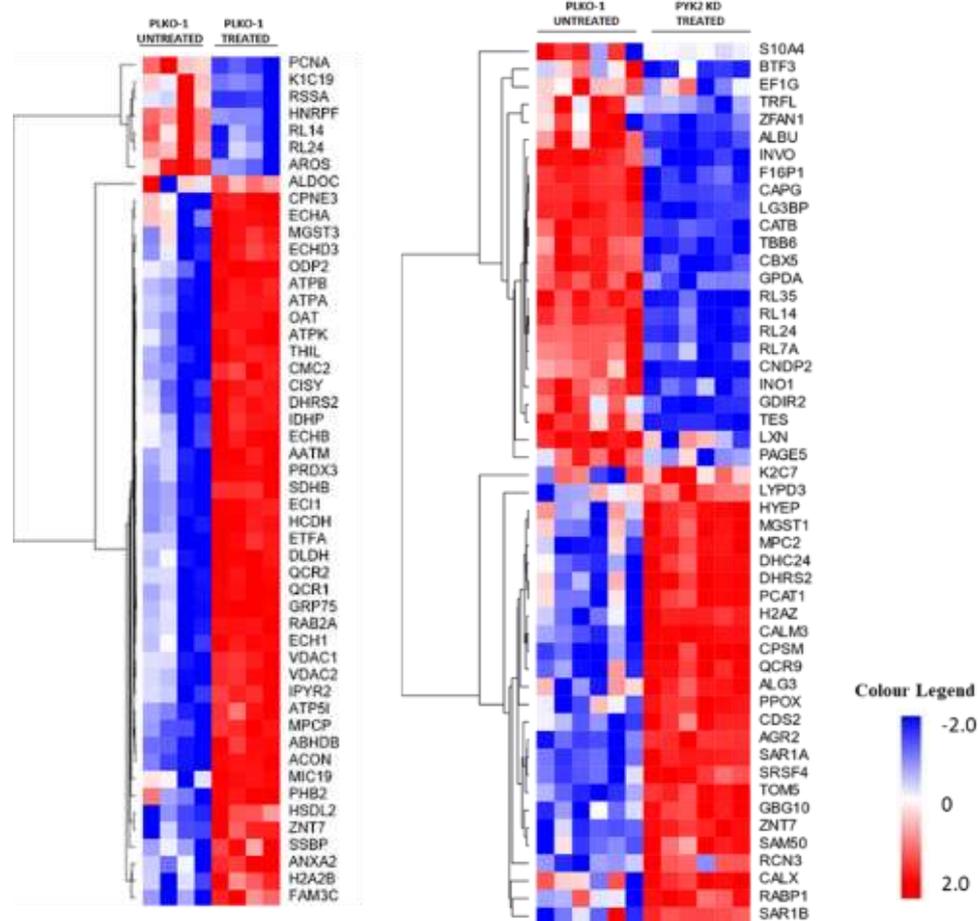


Figure 6.18. Differentially expressed proteins from untreated and treated SkBr3 control and PTK2B cells. The left-hand heat map shows the comparison between untreated and treated PLKO-1 samples. The right-hand heat map represents the comparison between untreated PLKO-1 samples and Metformin-treated PYK2 KD samples. The heat maps display the top 25 upregulated and 25 downregulated proteins according to $F_c \geq -2$, and confidence of 70%. The colour gradient is between red and blues with white in the middle.

Table 6.5 below shows the list of proteins that satisfied the cut-off criteria $1.4 F_c \geq -2$, and confidence of 70% comparing between pLKO-1 untreated (control) and PLKO-1 treated samples.

Table 6.5. Top 25 upregulated proteins and 25 downregulated proteins in HER2 (SkBr3) cell line

Protein Symbol	Description	Fold change	Confidence level
ZNT7	Zinc transporter 7	5.525	0.751
ATPK	ATP synthase subunit f, mitochondrial	5.234	0.791
MPCP	Phosphate carrier protein, mitochondrial	4.858	0.721
ATP5I	ATP synthase subunit e, mitochondrial	4.046	0.757
DHRS2	Dehydrogenase/reductase SDR family member 2, mitochondrial	3.304	0.776
CISY	Citrate synthase, mitochondrial	3.265	0.752
QCR2	Cytochrome b-c1 complex subunit 2, mitochondrial	3.187	0.793
IDHP	Isocitrate dehydrogenase [NADP], mitochondrial	3.176	0.839
ATPB	ATP synthase subunit beta, mitochondrial	3.172	0.905
SDHB	Succinate dehydrogenase [ubiquinone] iron-sulfur subunit, mitochondrial	3.119	0.73
OAT	Ornithine aminotransferase, mitochondrial	3.098	0.731
CMC2	Calcium-binding mitochondrial carrier protein Aralar2	3.086	0.758
VDAC1	Voltage-dependent anion-selective channel protein 1	2.999	0.768
ATPA	ATP synthase subunit alpha, mitochondrial	2.974	0.88
VDAC2	Voltage-dependent anion-selective channel protein 2	2.901	0.783
HCDH	Hydroxyacyl-coenzyme A dehydrogenase, mitochondrial	2.835	0.721
ECB	Trifunctional enzyme subunit beta, mitochondrial	2.825	0.842
MGST3	Microsomal glutathione S-transferase 3	2.812	0.714
ABHDB	Protein ABHD11	2.725	0.776
ACON	Aconitate hydratase, mitochondrial	2.675	0.768
ODP2	Dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex, mitochondrial	2.655	0.729
QCR1	Cytochrome b-c1 complex subunit 1, mitochondrial	2.637	0.85
H2A2B	Histone H2A type 2-B	2.624	0.811
THIL	Acetyl-CoA acetyltransferase, mitochondrial	2.612	0.832
ECHA	Trifunctional enzyme subunit alpha, mitochondrial	2.605	0.743
ECH1	Delta (3,5)-Delta (2,4)-dienoyl-CoA isomerase, mitochondrial	2.242	0.813

Protein Symbol	Description	Fold change	Confidence level
AATM	Aspartate aminotransferase, mitochondrial	2.227	0.744
GRP75	Stress-70 protein, mitochondrial	2.188	0.808
SSBP	Single-stranded DNA-binding protein, mitochondrial	2.161	0.839
DLDH	Dihydrolipoyl dehydrogenase, mitochondrial	2.134	0.76
EC11	Enoyl-CoA delta isomerase 1, mitochondrial	2.077	0.750
IPYR2	Inorganic pyrophosphatase 2, mitochondrial	2.059	0.708
PRDX3	Thioredoxin-dependent peroxide reductase, mitochondrial	2.049	0.784
MIC19	MICOS complex subunit MIC19	2.026	0.738
ETFA	Electron transfer flavoprotein subunit alpha, mitochondrial	1.941	0.816
ANXA2	Annexin A2	1.877	0.78
ECHD3	Enoyl-CoA hydratase domain-containing protein 3, mitochondrial	1.848	0.887
PHB2	Prohibitin-2	1.793	0.722
RAB2A	Ras-related protein Rab-2A	1.639	0.714
CPNE3	Copine-3	1.537	0.705
HSDL2	Hydroxysteroid dehydrogenase-like protein 2	1.522	0.732
FAM3C	Protein FAM3C	1.473	0.718

Protein Symbol	Description	Fold change	Confidence level
RSSA	40S ribosomal protein SA	-1.337	0.769
HNRPF	Heterogeneous nuclear ribonucleoprotein F	-1.365	0.754
ALDOC	Fructose-bisphosphate aldolase C	-1.417	0.764
K1C19	Keratin, type I cytoskeletal 19	-1.440	0.731
RL14	60S ribosomal protein L14	-1.509	0.885
PCNA	Proliferating cell nuclear antigen	-1.563	0.774
RL24	60S ribosomal protein L24	-1.579	0.752
AROS	Active regulator of SIRT1	-1.621	0.774

Samples comparing untreated and treated PLKO-1 samples with Fc 1.4 \geq -2, and confidence of 70%. Red indicated upregulated proteins and blue downregulated proteins.

Table 6.6 shows the list of proteins that satisfied the cut-off criteria $2 Fc \geq -2$, and confidence of 70% comparing between PLKO-1 untreated (control) and PYK2 KD treated samples.

Table 6.6. Top 25 upregulated proteins and 25 downregulated proteins in HER2 (SkBr3) cell line

Protein Symbol	Description	Fold change	Confidence level
SRSF4	Serine/arginine-rich splicing factor 4	17.913	0.749
GBG10	Guanine nucleotide-binding protein G(I)/G(S)/G(O) subunit gamma-10	12.517	0.747
CDS2	Phosphatidate cytidyltransferase 2	12.435	0.736
TOM5	Mitochondrial import receptor subunit TOM5 homolog	11.294	0.779
CPSM	Carbamoyl-phosphate synthase [ammonia], mitochondrial	10.862	0.796
H2AZ	Histone H2A.Z	10.102	0.920
AGR2	Anterior gradient protein 2 homolog	9.677	0.779
PPOX	Protoporphyrinogen oxidase	8.439	0.971
HYEP	Epoxide hydrolase 1	8.375	0.825
MGST1	Microsomal glutathione S-transferase 1	8.102	0.807
QCR9	Cytochrome b-c1 complex subunit 9	7.351	0.749
SAM50	Sorting and assembly machinery component 50 homolog	7.333	0.824
ALG3	Dol-P-Man: Man(5)GlcNAc(2)-PP-Dol alpha-1,3-mannosyltransferase	6.978	0.782
ZNT7	Zinc transporter 7	6.974	0.789
MPC2	Mitochondrial pyruvate carrier 2	6.660	0.800
CALM3	Calmodulin-3	6.641	0.729
RABP1	Cellular retinoic acid-binding protein 1	6.610	0.808
RCN3	Reticulocalbin-3	6.375	0.777
LYPD3	Ly6/PLAUR domain-containing protein 3	6.334	0.721
CALX	Calnexin	6.136	0.873
PCAT1	Lysophosphatidylcholine acyltransferase 1	5.903	0.841

SAR1A	GTP-binding protein SAR1a	5.464	0.856
DHC24	Delta (24)-sterol reductase	5.434	0.800
SAR1B	GTP-binding protein SAR1b	5.395	0.727
DHRS2	Dehydrogenase/reductase SDR family member 2, mitochondrial	5.341	0.883

Protein Symbol	Description	Fold change	Confidence level
RL7A	60S ribosomal protein L7a	-3.210	0.802
CNDP2	Cytosolic non-specific dipeptidase	-3.230	0.755
CATB	Cathepsin B	-3.267	0.781
TES	Testing	-3.283	0.852
RL24	60S ribosomal protein L24	-3.681	0.905
RL14	60S ribosomal protein L14	-3.709	0.942
GPDA	Glycerol-3-phosphate dehydrogenase [NAD (+)], cytoplasmic	-3.829	0.740
TBB6	Tubulin beta-6 chain	-3.835	0.822
GDIR2	Rho GDP-dissociation inhibitor 2	-3.881	0.770
INO1	Inositol-3-phosphate synthase 1	-4.017	0.741
K2C7	Keratin, type II cytoskeletal 7	-4.060	0.710
CBX5	Chromobox protein homolog 5	-4.163	0.831
TRFL	Lactotransferrin	-5.657	0.701
LXN	Latexin	-7.613	0.797
S10A4	Protein S100-A4	-8.067	0.727
INVO	Involucrin	-10.096	0.722
RL35	60S ribosomal protein L35	-10.733	0.820
ZFAN1	AN1-type zinc finger protein 1	-11.261	0.7195
EF1G	Elongation factor 1-gamma	-13.597	0.752
PAGE5	P antigen family member 5	-14.763	0.713
LG3BP	Galectin-3-binding protein	-16.528	0.900
ALBU	Serum albumin	-16.538	0.831
BTF3	Transcription factor BTF3	-18.724	0.730
CAPG	Macrophage-capping protein	-20.770	0.837
F16P1	Fructose-1,6-bisphosphatase 1	-26.012	0.834

Samples comparing untreated PLKO-1 and Metformin-treated PYK2 KD samples with $Fc \geq -2$, and confidence of 70%. Red indicated upregulated proteins and blue downregulated proteins.

6.7. Pathways Enrichment analysis

Differentially expressed proteins in untreated and treated pLKO-1 and PYK2 KD SkBr3 samples with Fold change $2 \geq -2$ and confidence of 70%, were subsequently analysed using Meta Core™ pathway identification online tool. A list of upregulated or downregulated proteins in pLKO-1 and PYK2 KD were clustered into several signalling pathways and in accordance with their statistical significance ($p < 0.05$). Meta Core process functional enrichment analyses led to the identification of 7 pathways with FDR (False Discovery Rate) < 0.05 . Table 6.8 shows the pathways that are differentially regulated and that are expressed between pLKO-1 and PYK2 KD cells. Four common upregulated pathways (Transcription and Chromatin modification, Protein folding and Protein folding nucleus, Reproduction and spermatogenesis, motility and copulation, Reproduction and Male sex differentiation) have been identified in untreated pLKO-p and PYK2 KD samples, and which are likely to be related to breast cancer disease due to Somatotropin overexpression in breast cancer (see details in the previous section 6.5, page 192). In addition, one up-regulated pathway has been found common between untreated pLKO-1 and PYK2 KD Control and treated pLKO-1 and PYK2 KD. This pathway is associated with transcription and chromatin modification is also found to be associated with human diseases, including cancer. However, one common pathway was downregulated (Cell cycle and Mitosis) in both untreated and treated pLKO-1 (Control) vs. PYK2 KD and that may explain the increase of proliferation in PYK2 KD samples compared to pLKO-1 samples. Moreover, two up-regulated pathways have been observed in untreated pLKO-1 vs. PLKO-1 treated, which are Apoptosis and Apoptotic mitochondria, and Response to hypoxia and oxidative stress, which is likely due to the impact of Metformin treatment.

All above-mentioned steps of protein analysis and Pathways Enrichment analysis have been applied to the other HER2 cell line. Heat maps can be found in the Appendix (A.6.10.) and Tables (A.6.1, A.6.2, A.6.3, A.6.4, A.6.5).

Table 6.8. Pathways regulated by proteins that are differentially expressed between treated and untreated PLKO-1 and PYK2

KD cells (SkBr3 cell line)

	Biological functions	p-value	FDR	Protein Names
PLKO-1 Control up-regulated pathways	Transcription_ Chromatin modification	1.80456E-05	0.0009	Histone H2A, Pc2, H2AFZ, Histone H2
	Protein folding_ Protein folding nucleus	0.003	0.0492	NFYC, SFRS4
	Reproduction_ Spermatogenesis, motility and copulation	0.003	0.0492	Histone H2A, Histone H2, Calmodulin
	Reproduction_ Male sex differentiation	0.004	0.0492	Histone H2A, Histone H2, Calmodulin
PLKO-1 Control down-regulated pathways	Cell cycle_ Mitosis	0.0022	0.0377	HP1 alpha, CAP-G/G2, HP1
PYK2 KD Control up-regulated pathways	Transcription_ Chromatin modification	1.80456E-05	0.0009	Histone H2A, Pc2, H2AFZ, Histone H2
	Protein folding_ Protein folding nucleus	0.0028	0.0492	NFYC, SFRS4
	Reproduction_ Spermatogenesis, motility and copulation	0.0033	0.0492	Histone H2A, Histone H2, Calmodulin
	Reproduction_ Male sex differentiation	0.0039	0.0492	Histone H2A, Histone H2, Calmodulin
PLKO-1 Control vs. PLKO-1 Treated up-regulated pathways	Apoptosis_ Apoptotic mitochondria	0.0001	0.0036	VDAC 2, VDAC 1, HSP70
	Response to hypoxia and oxidative stress	0.0012	0.0157	Peroxiredoxin, PRDX3, MGST3
PLKO-1 Control vs. PYK2 KD Treated up-regulated pathways	Transcription_ Chromatin modification	1.80456E-05	0.0009	Pc2, Histone H2, Histone H2A, H2AFZ
PLKO-1 Control vs. PYK2 KD Treated down-regulated pathways	Cell cycle_ Mitosis	0.0016	0.0197	HP1 alpha, CAP-G/G2, HP1

Pathways were derived by Meta Core™ online software. Table only shows significant pathways (with FDR < 0.05). Red indicated upregulated pathways and blue downregulated pathways.

6.8. Discussion:

Several studies indicate that *PTK2B* (*PYK2*) is associated with increased migration and invasion, an early event of the metastatic process. Therefore, the main objective of this chapter was to interrogate this relation by investigating the role of *PYK2* in different breast cancer subtypes. This was achieved through *PYK2* knockdown and regarding Metformin treatment. Effects of this experimental approach on fundamental biological functions such as proliferation as well as the metastasis of cells were investigated. Thorough and complete interrogation has been made in this chapter to validate the role of *PYK2* in cell migration, invasion and proliferation of breast cancer cell lines.

Collectively, the findings showed that *PYK2*-knockdown cells exhibited a dramatic reduction in migration and invasion of HER2 (SkBr3) and (MDA-MB-453) breast cancer cells, while significantly increased their ability to proliferate after 48h of treatment and when compared to PLKO-1 cells (control). Similarly, the Basal-like (MDA-MB-468) cell migration and invasion were significantly reduced following *PYK2* knockdown. However, the proliferation of cells was also decreased in comparison to the empty vector (control).

Indeed, these findings are consistent with previous studies that reported a significant decrease in 2D motility and invadopodia mediated functions that decreased the matrix invasion as a result of *PYK2* depletion in breast cancer cells. Besides, *Genna* and his colleagues have identified *PYK2* as a unique facilitator of invadopodia formation and function. As well as, they proved a novel underlying mechanism by which *PYK2* mediates tumour cell invasion (*Genna et al., 2018*). Another study by *Genna* and *Gil-Henn* (2018) confirmed the relation between Proline-rich tyrosine kinase 2 (*PYK2*) and its closely related focal adhesion kinase (*FAK*) in the regulation of cancer cell invasion. This is achieved by coordinating the balance between focal adhesion-mediated migration and invadopodia-dependent extracellular matrix invasion. It was suggested that targeting either one or both kinases could block breast cancer metastasis. *Caoa* and Colleagues (2017) showed that the loss of miR-23b played an essential role in hepatocellular carcinoma (HCC) progression and metastasis through regulating *PYK2*. They explored in their study the possible involvement of miR-23b in HCC cell proliferation and metastasis and provided insight into underlying mechanisms. They found that miR-23b, which functioned, as a tumour suppressor that inhibits HCC cell invasion and migration via EMT regulation, was downregulated in HCC tissues and cell lines. Interestingly, they identified *PYK2* as a target of miR-23b, and the

overexpression of *PYK2* could significantly restore the inhibitory effects of miR- 23b overexpression on the metastasis (Cao *et al.*, 2017).

Our results are also consistent with the “Divide or Conquer” and “go or grow” hypothesis, which proposed that dividing or migrating cells are temporally exclusive events in infiltrative gliomas (Lipinski *et al.*, 2005). This implies that cells cannot migrate and proliferate simultaneously, thus one behaviour will have a direct influence on the other. *FAK* has been shown *in vitro* as a promoter of cell cycle progression specifically in malignant astrocytoma cells. In addition, *FAK* expression in both SF767 and G112 glioma cell lines was associated with a reduction in cell migration and induction of cell cycle progression (Lipinski *et al.*, 2005). Recent evidence by Kohrman and Matus (2017) documented a functional link between cell cycle arrest and invasive activity. The cells may require a switch from a proliferative to an invasive state as a critical aspect of metastasis. They reviewed an evidence that basement membrane (BM) invasion, which is a fundamental feature of cancer metastasis, required cell cycle arrest (Kohrman and Matus, 2017).

Conversely, Claudin-Low (MDA-MB-231) cells showed a completely different response pattern, with increased migration and invasion and decreased proliferation following *PYK2* knockdown and when compared to the control. This can rely on *Pyk2* function in specific cell types migration, a study on *Pyk2*-deficient mice demonstrated that Macrophage cells fail to become polarised or to migrate despite the normal expression of *Fak* in these cells. Notably, in *Pyk2*-null cells, integrin adhesion-mediated activation of Rho and PI-3 kinase was significantly compromised. Similarly, the migration of lymphocytes B was impaired in *Pyk2*-null mice in the absence or presence of chemokines (Okigaki *et al.*, 2003; Lipinski *et al.*, 2005). A comparable study in brain microvascular endothelial cells that express both *PYK2* and *FAK*, found that expression of *PYK2* stimulated migration whereas expression of an inactive *PYK2* variant substantially inhibited cell spreading and migration (Avraham *et al.*, 2003). Additionally, *PYK2* has been reported as a convergence point between receptor tyrosine kinases such as EGFR and FGFR, which play a central activation role in critical signalling networks to cell motility and proliferation, and G protein-coupled receptors (Meyer *et al.*, 2003). Moreover, *PYK2* has also been identified as a mediator in STAT3 triggering, which enhances the proliferation of cancer cells (Shi and Kehrl, 2004). It is also implicated in the induction of heregulin stimulation, which promotes breast carcinoma invasion (McShan *et al.*, 2002).

Resistance to chemotherapeutic compounds is considered as one of the main therapeutic limitations in advanced breast cancer, where cancer cells escape the cytotoxic effects of chemotherapies by developing multiple drug resistance, which leads to cancer recurrence and

decreased survival in cancer patients. *Tavora* and his team in their 2014 study, identified a novel molecular mechanism underlying the chemo-sensitivity regulation of endothelial cells. They established that to inhibit tumour growth in mice, induction of tumour cell sensitisation to DNA-damaging therapies is required, and this can be achieved explicitly by targeting focal adhesion kinase (*FAK*) in endothelial cells. They also supported their work through the clinical observation in human lymphoma, that complete remission was associated with low *FAK* expression in blood vessel. The study showed that *FAK* deletion in endothelial cells induces increased apoptosis and decreased proliferation within perivascular tumour-cell compartment of doxorubicin- and radiotherapy-treated mice. However, *FAK* deletion did not impact blood vessel function. In another word, the loss of endothelial cell *FAK* enhanced the chemosensitisation of tumour cells to DNA-damaging therapies *in vitro* and *in vivo* by reducing DNA-damage-induced cytokine production (*Tavora et al., 2014*).

In agreement with the above study, we observed that *PYK2* depletion and Metformin treatment prevents invasion of Claudin-Low (MDA-MB-231), HER2 (SkBr3) and (MDA-MB-453) breast cancer cells. Metformin treatment did not appear to have any effect on migration and proliferation of the knockdown cells except the proliferation of MDA-MB-453 cell line, which reduced following the Metformin treatment. PI3K (phosphatidylinositol 3-kinases) /AKT (serine/threonine kinase also known as PKB) signalling pathway is also regulated by *PYK2*. Elevated levels of *PYK2* expression have been associated with poor survival and metastasis in HCC *via* the activation of the PI3K/AKT pathway in a *PYK2*-dependent phosphorylation of AKT. *PYK2* is also implicated in the enhancement of migration and invasion through the activation of the PI3K/AKT signalling pathway (*Gutenberg et al., 2004; Sun et al., 2007; Sun et al., 2008; Gong et al., 2014*). In addition, *Zrihan-Licht* and colleagues (2000) found that *PYK2* expression increased the invasive potential of MDA-MB-435 and MCF-7 breast cancer cells by activating Src and the mitogen-activated protein kinase (MAP kinase) pathways. *Behmoaram* and her group in a 2008 study, observed an increased level of *PYK2* expression in early and advanced breast cancer and when compared to benign and normal breast tissues. Moreover, the inhibition of *PYK2* resulted in reduced tumour development and metastasis in pulmonary metastases (*Fan and Guan, 2011*). Furthermore, overexpression of *PYK2* promoted cell migration and invasion, and enhanced metastasis and EMT in Hep-3B HCC cells, through phosphorylation, upregulation, and localisation of the EMT regulator transcription factor Hic-5 (*Sun et al., 2011*).

The Meta Core™ search has uncovered a role of *PYK2* in several biological processes involved in various pathological pathways that include Autocrine Somatotropin signalling pathway in breast cancer. *PYK2* plays a central role (in breast cancer) in Autocrine Somatotropin signalling pathway. This can clarify the diverse roles of this molecule in the different biological processes of breast

cancer. The exact role of *PYK2* in cancer development is varied and complex. *PYK2* plays a critical role in a variety of biological processes of carcinoma such as proliferation, and cellular migration and invasion.

Interestingly, MS data analyses revealed several up and downregulated pathways, including common pathways in untreated and treated cells expressing *PYK2* shRNAs. Most significant upregulated pathways that are associated with untreated samples (pLKO.1 and *PYK2* KD) are associated with Transcription and Chromatin modification, Protein folding and Protein folding nucleus, Reproduction and Spermatogenesis, motility and copulation, and Reproduction and Male sex differentiation. Those pathways are most likely related to the exceeding levels of growth hormone (GH) in breast cancer. *Bartke* summarised in his 2000 review, that the growth hormone (GH) has a physiological role in the control of male reproductive development and function in the normal level of expression. GH is fundamental in growth promotion, cell division and regeneration. It also exerts a diverse and widespread action on the human body after binding to its receptor (GHR). Besides, GH impacts the metabolism of carbohydrates, lipids and proteins, as well as, shapes body composition. GH has diverse roles in maintaining human development and homeostasis. Its continuous secretion stimulates the growth and participates in the equilibrium of a process that tightly organised and arranged by many organs. The deficiency in GH level is considered a medical condition, which affects all ages. This deficiency will not only have significant effect consequences in the health of the patient but also will impact on the quality of life. Besides its positive actions, it has also been involved in the genesis of several diseases including cancer and insulin-resistant diabetes, as a result of GH/GHR interaction and (IGF-I) production (*Caicedo and Rosenfeld, 2018*). Once GH is secreted by the somatotrophic cells in the pituitary gland, it will bind to its receptor (GHR) and then transported to the entire body by the circulatory system. This protein receptor (GHR) is a class 1 cytokine receptor family member that is located on the cell membrane as a constitutive dimer. As a consequence of this binding, Janus kinases (JAK2) are activated, an action followed by activation of transcription activators (STATs) that induce the IGF-I, IGFBP3 and the leucine-rich protein known as acid labile subunit synthesis (ALS).

The major mediator of GH-stimulated somatic growth, as well as a mediator of GH-independent anabolic responses in many cells and tissues, is IGF-I, which is a small peptide that consist of 70 amino acids. The main route by which the GHR exerts its physiological effects is JAK2/STAT-5/IGF-I signalling pathway. Nevertheless, an altered route *via* expression of multiple transcripts has been found in murine models with deletion of liver GHR. Additionally, GH may employ other routes such as mitogen-activated protein kinase (MAPK) to perform its activities. Following GH activation, several continues activation will occur including RAS GTPase, RAF kinase and the

MAP-ERK kinase (MEK). As a consequent of these serial activations of the transcriptional regulation of target genes, the growth enhancement and metabolism will be achieved (*Guevara-Aguirrea et al., 2018*).

Transcription and chromatin modification, which was a commonly upregulated pathway, is also related to cancer. Chromatin modification is a dynamic modification of chromatin architecture that controls gene expression by allowing access of condensed genomic DNA to regulatory transcription machinery proteins. Besides, a dynamic chromatin modification conveys an epigenetic regulatory role in several biological processes such as DNA replication and repair, apoptosis, chromosomes segregation, development and pluripotency. This re-modelling is implemented through covalent histone modifications by specific enzymes or ATP-dependent chromatin remodelling complexes. In addition, deviation in chromatin remodelling proteins, were found to be related to human diseases, including cancer. Currently, a major therapeutic strategy in the treatment of several cancers has evolving targeting chromatin remodelling pathways (*Wang et al., 2007*). The up-regulation of this pathway might be related to Metformin action as a DNA damage agent. This pathway is employing several biological processes to repair and replicate the DNA. Furthermore, *Chi* and colleagues, summarised in their review, some evidence that links oncogenesis and miswriting, misinterpretation and mis-erasing of histone modifications, suggesting that deregulated gene expression and perturbation of cellular identity can be caused by histone code mis-regulation, contributing to cancer initiation, progression and/or metastasis (*Chi et al., 2010*).

Cell Cycle and Mitosis pathway were noticed as down-regulated pathways in control samples comparing to treated samples. The activation of signalling pathways that promote cell cycle arrest and DNA repair is likely caused by Metformin mode of action.

A study by Marinello and her team, in 2016 suggested that Metformin mechanism of action involves oxidative stress generation, DNA damage, and transforming growth factor β 1 induction in MCF7 and MDA-MB-231 human breast cancer cells.

Besides, Apoptosis, Apoptotic mitochondria, Response to hypoxia and oxidative stress pathways, which were also up-regulated in PLKO-1 Control samples vs. PLKO-1 Treated samples might be also due to Metformin treatment.

In summary, this chapter aimed to elucidate the involving role of PYK2 in cell migration, cell invasion and cell proliferation of breast cancer cell lines. Our findings demonstrated that PYK2 depletion decreased the migration and invasion of HER2 (SkBr3) and (MDA-MB-453) breast cancer cells and increased their ability to proliferate after 48h of treatment with Metformin.

Meta Core™ search has revealed the involving role of PYK2 in several biological processes and pathological pathways including Autocrine Somatotropin signalling pathway in breast cancer. In addition, MS data analysis was also showed several common up and down-regulated pathways between untreated and treated cells expressing PYK2 shRNAs. These pathways were likely related to levels of growth hormone (GH) in breast cancer and Metformin action.

CHAPTER 7

DISCUSSION

7.1. Summary of discussion

This project sought to confirm the molecular mechanism of Metformin action on breast cancer using systems biology approach *via* integrating computational methods and laboratory data. This project is considered as an original study that identified novel targets that could be used to help develop more effective strategies in the treatment of breast cancer. To that end, five breast cancer cell lines including BT-474 (breast ductal carcinoma derived from solid invasive ductal carcinoma of the breast), MCF-7, MDA-MB-231, MDA-MB-468, SkBr3 (breast adenocarcinoma derived from metastatic site: pleural effusion) and MDA-MB-453 (metastatic carcinoma derived from metastatic site: pleural effusion), representing five molecular breast cancer subtypes (Luminal B, Luminal A, Claudin-low, Basal-like, and HER2) respectively, have been employed in this project to investigate the efficacy of Metformin on breast cancer. Cell survival (MTT), proliferation and apoptosis assays have revealed that Metformin provides same responses patterns in all breast cancer subtypes and at different concentrations. However, significant differences have been seen between different phenotypes at the highest concentrations of 5-25mM after 48h of treatment with Metformin. Interestingly, the treatment significantly reduced cell survival and proliferation, while induced cell apoptosis and enhanced cell necrosis. The Basal-like cell line (MDA-MB-468) was the most affected cell line. The lethal dose (EC) was \leq 2mM and the EC50 concentration was 1mM after 48h of treatment, respectively. In addition, Metformin affected the appearance of the cells (Morphology) and their adhesion ability. Comparatively, the HER2 subtype (SkBr3 cell line) was the least affected phenotype and the EC50 concentration was much higher 2.3mM after 48h of treatment.

This finding comes in line with previous results from several studies that revealed the antitumour activity of Metformin in various breast cancer types. *In vitro* studies showed that Metformin inhibits cell proliferation and survival of triple negative breast cancer cells (Liu *et al.*, 2009; Denget *et al.*, 2012). Another study by Vazquez-Martin and co-authors in 2011 confirmed that Basal-like breast cancer cells (MDA-MB-468) were significantly more sensitive to the growth-inhibitory effects of Metformin among the different molecular classes of breast cancer. Furthermore, they suggested that Metformin suppresses the metastasis-associated protein and

stem cell marker CD24 in MDA-MB-468 triple-negative breast cancer cells (*Vazquez-Martin et al., 2011*). Besides, Metformin monotherapy or in combination with chemotherapeutic agents inhibited cell proliferation in triple negative breast cancer cell lines. Additionally, it has been proved that Metformin inhibits the growth of breast cancer cell lines (MCF-7, MDA-MB-231 and MDA-MB-435) *in vitro*, and *via* AMPK induction and mTOR inhibition (*Zakikhani et al., 2006*; *Dowling et al., 2007*; *Phoenix et al., 2009*; *Hadad et al., 2009*; *Hadad et al., 2011*; *Liu et al., 2009*; *Liu et al., 2012*).

Moreover, an investigation of Metformin efficacy on human breast cancer cell lines (MCF-7 and in MDA-MB-231) using different clinical and experimental concentrations has revealed the cytotoxicity, oxidative stress, DNA damage, and intracellular pathways related to cell growth and survival after 24 h of drug exposure (*Marinello et al., 2016*). Moreover, *Zheng* and co-authors in their 2018 study, have shown that Metformin monotherapy had low inhibition rates on cell viability and apoptosis. However, co-treatment of Metformin with flavone (core structure of flavonoids), has synergistically inhibited cell viability, increased apoptosis in human breast cancer cells *via* the inhibition of PI3K/AKT pathway (*Zheng et al., 2018*). Metformin exerts anti-proliferative activity by interfering with the PI3K/AKT pathway to induce apoptosis in many cancer cells. AKT could activate eIF4E by inhibiting mTOR to stimulate cyclin D1, that leads to cell cycle arrest in G1 phase. Metformin considered as an original AMPK activator, in which increasing nuclear accumulation and protein stability of FoxO3a through the inhibition of AKT-MDM2 signalling pathway to reduce the invasive and metastatic capacity of aggressive cancer cells (*Zheng et al., 2018*).

Meta-analysis of 11 studies that involved 5,464 patients with breast cancer have shown that Metformin treatment in patients with diabetes enhanced both overall survival (OS) and cancer-specific survival (*Sonnenblick et al., 2017*). However, the described studies in this analysis, had several limitations including heterogeneous populations, anticancer treatments and inclusion criteria, with limited modification for confusing variables. Metformin effect and diabetes outcome have not been previously evaluated in patients with human epidermal growth factor receptor 2 (HER2) –positive breast cancer who have taken treatment with adjuvant anti-HER2 therapy. An exciting observation of this study is that the risk of distant disease-free survival (DDFS) and death was more than double in patients with diabetes and with hormone receptor-positive cancer. However, this effect was not seen in patients with hormone receptor-negative tumours (*Sonnenblick et al., 2017*). Another study has used a syngeneic model of murine primary and metastatic mammary cancer and found that Metformin might have a tumour-suppressing activity in some instances including a metabolic phenotype of high fuel intake, a metabolic syndrome, and diabetes, however, it may have little or no effect on the events the controlling

the metastatic niche driven by pro inflammatory events (*Phoenix et al., 2010*). It has been reported that Metformin has cytotoxic activity in high concentrations and on various malignant cells in nutrient medium with foetal bovine serum (FBS). The major growth inhibitory effects of Metformin on tumour cells are mediated *via* the suppression of mTOR signalling pathway which might be preceded by the activation of AMPK (tumour suppressor Liver Kinase B1 (LKB1) dependent signalling pathway). This complex can prevent mTOR signalling through phosphorylation and stabilisation of TSC Complex Subunit 2(TSC2). It can inhibit HER2 protein kinase activation causing decreased signalling through downstream pathways. Metformin can inhibit the cell cycle slightly through the reduction of cyclin D1 expression (*Damjanović et al., 2015*). Metformin has antiproliferative action on dividing cells, independently from the nature of target cells (malignant or normal proliferative cells), suggesting the possibility of a favourable effect of Metformin in autoimmune diseases suppression. In addition, Metformin can decrease levels of Vascular endothelial growth factor (VEGF) resulting in angiogenesis inhibition that can promote apoptosis through p53-dependent or independent pathways. It has been shown that high concentrations of Metformin directly decrease the survival of HER2+ breast cancer cells, or through immune-mediated PBMC antitumour action (*Damjanović et al., 2015*). However, immunocytochemical staining of HER2+ in treated MDA-MB-361 breast adenocarcinoma cells did not affect the expression of HER2 on the surface of target cells (*Damjanović et al., 2015*).

In conclusion, Metformin can potentially reduce the proliferation rate of a subset of cancers can only be achieved if these subsets have specific molecular characteristics and if the patient had the treatment.

Gene Expression Microarray and various bioinformatics tools including Artificial Neural Networks (ANN) and Cytoscape among others, to inspect the effect of Metformin on the level of gene expression were used. Further, Regression-based method and Fold Change-based method was also employed for data analytics. These approaches have enabled the detection of several genes that were significantly expressed in an uncontrolled manner as up or downregulated in the breast cancer cell lines MDA-MB-468 and SkBr3. Additionally, the Panther online databases (*Mi et al., 2013*) has exposed the involvement of those genes with several biological pathways. These genes were *ANKRD44, BTN3A1, C2orf42, DHFR2, EMP1, HGF, IRF-9, NPRL3, OXSR1, PCDHB2, PRDX1, PTK2B, PXDC1, SERPINB4, VSTM2B, and WBSCR27* respectively. The gonadotropin-releasing hormone receptor pathway has been found as a common upregulated pathway in both BASAL-LIKE (MDA-MB-468) and HER2 (SkBr3) cell lines. On the other hand, apoptosis signalling pathway was found up-regulated in BASAL-LIKE (MDA-MB-468) cell line, confirming previous data. Besides, the Wnt signalling pathway was downregulated in MDA-MB-468 cells, which might explain the sensitivity of this cell line to Metformin treatments.

As this pathway has been involved in drug resistance. Moreover, In HER2 (SkBr3) cell line, four critical upregulated pathways have been found: integrin signalling pathway, CCKR signalling, Inflammation mediated by chemokine and cytokine signalling pathway and Gonadotropin-releasing hormone receptor pathway. Interestingly, *Protein Tyrosine Kinase 2 Beta (PTK2B)* has been found as signalling through all these pathways. *PTK2B* is also known as *PYK2* or *FAK2*, is a non-receptor tyrosine kinase that has been confirmed as an upregulated gene in HER2 cell line following Metformin treatment. *PYK2* is a critical regulator of survival and invasiveness of cancer cells (*Naser et al., 2018*).

Several studies have identified the focal-adhesion kinases *FAK* and *Pyk2* as possible common mediators of signalling by growth factors and integrins. The co-stimulation of growth-factor receptors and integrins activates the focal-adhesion kinase (FAK) family to support the outgrowth of neurites in rat pheochromocytoma (PC12) cells and human neuroblastoma SH-SY5Y cells (*Ivankovic-Dikic et al., 2000*). Integrin is a protein that functions mechanically as an anchorage by attaching the cell cytoskeleton to the extracellular matrix (ECM), and biochemically by transferring chemical signals into the cell (outside-in signalling) to provide information on its site, local environment, adhesive status and the surrounding matrix. Integrins are a family of proteins that consist of α and β subtypes, which form transmembrane heterodimers. Integrins function as adhesion receptors for extracellular ligands and convert biochemical signals into the cell, through downstream effector proteins. Remarkably, they perform bidirectionally, by transporting information both outside and inside out (*Harburger and Calderwood, 2009*). Recently, several findings have demonstrated the involvement of integrins in stem and cancer stem cells. Integrins are also required for cancer progression and drug resistance. Besides, integrins are essential markers for identifying epithelial stem cells and progenitor cells in the mammary gland. Epithelial lineages arising from the same precursors in the breast can be discriminated according to their integrin profiles. The integrin $\alpha 6$ (CD49f) and $\beta 1$ subunit are expressed in low levels in Luminal cells, and at higher levels in the basal layer including mammary stem cells (*Seguin et al., 2015*). A cell can regulate the adhesive activity of its integrins from within. It also functions as signal transducers when activated by matrix binding, which enables various intracellular signalling pathways. Often Integrins and traditional signalling receptors cooperate to promote cell growth, cell survival, and cell proliferation (*Alberts et al., 2002*). Thus, this could clarify the susceptibility of Basal-like (MDA-MB-468) by downstream integrins and the reduced sensitivity of HER2 (SKBR3) phenotype upstream.

Additionally, it has been proved that Gastrin-releasing peptide receptor regulates the leupaxin localisation and *PYK2* activity (*Chen and Kroog, 2005*). Gastrin has two forms of the biologically active peptide, G34 and G17. The Gastrin gene is upregulated in pre-malignant conditions and

in certain cancers, which influences proliferation, angiogenesis and apoptosis, successively acting through the classical gastrin cholecystokinin B receptor CCK-BR and its isoforms, and *via* alternative receptors to stimulate signalling pathways, that influence the expression of downstream genes that affect cell survival, angiogenesis and invasion (*Tripathi et al., 2015*). This also supports this finding, which also found that (CCK) pathway was down-regulated in Basal-like (MDA-MB-468) and upregulated in HER2 (SkBr3) phenotypes. Moreover, the Gastrin and cholecystokinin (CCK) are gastrointestinal peptide hormone whose primary function is to stimulate hydrochloric acid secretion by the gastric mucosa and inhibition of gastrin formation. It also acts as a mitogenic factor for gastrointestinal epithelial cells. The former is involved in different physiological processes including digestion, appetite control and regulation of body weight. Also, they play an essential role in several diseases such as acute pancreatitis, obesity, irritable bowel syndrome gallbladder disease, as well as stimulate tumour proliferation, and osteoporosis (*Smith et al., 2016*).

It has been found that *PYK2* is critical for the Jak-mediated MAPK and Stat1 activation by IFN-gamma (*Takaoka et al., 1999*). Interferons (IFNs) are glycoproteins synthesised and released by host cells as a response to the presence of several pathogens such as viruses, bacteria, parasites or tumour cells. IFNs were named after their ability to “interfere” with viral replicate infection within the host cells. It has also been known that IFNs act as a cell proliferation inhibitor, differentiation inducer, immune system regulator and angiogenesis inhibitor through multiple signalling pathways. The biological effect of IFN-receptor signalling is mainly regulated by three factors, which are the IFN expression profile, the receptor profile, and target genes expression profile. The binding to specific receptors causes the initiation of signalling by IFNs that can directly induce gene transcription and/or multiple downstream signalling that consequently induce diverse cellular responses in tumour cells such as cell cycle arrest and apoptosis (*Galani et al., 2017*). Moreover, *Mimura* and his colleagues (2018) have found in their study, that interferon gamma increased the expression of programmed death ligand-1 in solid tumour cells out of the JAK-signal transduction and activation of transcription pathway and decreased the cytotoxicity effect of tumour antigen-specific CTL against tumour cells in gastric cancer. Despite, IFNs regulate the duration and intensity of innate and adaptive immune responses as well as control the survival, proliferation, and differentiation of responding cells. They are also instigators of several severe neuroinflammatory disorders. Consequently, this dual nature of IFNs requires that their signalling and expression be tightly regulated (*Mimura et al., 2018*). *IRF-9* plays a central role in interferon signalling and contributes to the intracellular signalling of all three classes of interferons. Specifically, type I interferons (IFN-Is) which include the IFN- α 's and IFN- β amongst others, type II interferon with IFN- γ being the only member, and the type III

interferons (IFN-IIIs), consisting of the IFN- λ s. IFN-Is have been proven to be highly effective treatments in some diseases therapy such as chronic viral hepatitis, autoimmune diseases, certain malignancies, and osteoporosis. Importantly, several studies have demonstrated that *IRF-9* inhibits the emergence of a potentially harmful IFN- γ -like response by driving signalling towards an IFN-I response. It has been found that *IRF-9* is importantly linked to the tumour suppressor protein (p53) by inducing the p53 apoptotic response *via* the induction of cell transformation and enhancement of oncogene-dependent apoptosis. Besides, it has been reported that cells lacking *IRF-9* were oversensitive to cytotoxic drugs. (Suprunenko and Hofer, 2016). Moreover, IRF heterodimers can regulate the gene expression both positively and negatively depending on the cell type, signals and target gene (Song and Schindler, 2009). The upregulation expression of *IRF-9* in both breast cancer subtypes Basal-like (MDA-MB-468) and HER2 (SkBr3) cell lines is associated with the up and down-regulation of (Interferon-gamma signalling pathway) in both phenotypes following the exposure to Metformin.

All previously identified up and down-regulated pathways; integrin signaling pathway, CCKR signalling, Inflammation mediated by chemokine and cytokine signalling pathway, Gonadotropin-releasing hormone receptor pathway, and Interferon-gamma signalling pathway that involved *PYK2* in their signals have been proved as associated to cancer progression, drug resistance and cell survival, which might have influenced the resistance to Metformin treatment of the HER2 breast cancer cell line. However, the down-regulated pathways were related to cell cycle arrest and apoptosis, that could explain the high proportion of dead cells in the Basal-like cell line.

Further analysis by Real-time quantitative PCR (qRT-PCR), Immunoblotting and Immunofluorescence have confirmed obtained data by Gene Expression Microarray analysis. Besides, subsequent analysis by NanoString technology has also confirmed obtained data. As previously mentioned, the aim of analysing total gene and protein expression was to discover novel biomarkers that are associated with Metformin resistance in HER2 positive cell lines and as potential and novel therapeutic targets, and these data will certainly contribute to these efforts.

Interestingly, it has been found that the apoptosis signalling pathway and Wnt signalling pathway were respectively upregulated and downregulated pathways in Basal-like (MDA-MB-468) population samples analysed by Gene Expression Microarray and NanoString techniques. However, in HER2 (SkBr3) population, a few common pathways have been found and that include apoptosis signalling pathway, phosphatidylinositol 3-kinase (PI3K) and transforming growth factor-beta (TGF beta) pathways.

It is evident that in Basal-like (MDA-MB-468) population samples, downstream pathways were mostly involved in apoptosis, proliferation, survival and cell migration, that might explain the sensitivity of this population to Metformin treatment and the significant death (apoptosis) of these cells after Metformin exposure. These down-regulated pathways include the apoptosis signalling pathway, the Wnt signalling pathway. The latter is a critical pathway that regulates cell migration, cell polarity, neural patterning and organogenesis during foetal development, and plays a crucial role in cell fate determination (*Komiya and Habas, 2008*). The Wnt signalling pathway is critically involved in tissue development and homeostasis by maintaining cancer stem cells pools. It is also a main player in the initiation, maintenance and development of many cancers through effects on the behaviour of cancer stem cells (CSCs). The latter is responsible for tumour initiation, drug resistance and cancer relapse (*Duchartrea et al., 2016; Zhan et al., 2017*).

Few pathways have been found as up and down-regulated such as the phosphatidylinositol 3-kinase (PI3K) pathway that is a regulator of various cellular processes such as metabolism, apoptosis, growth, proliferation, survival and cell migration, and that participates in specialised context-dependent functions (*Chalhoub and Baker, 2009*). A clear linkage has been established since the 1980s, between the PI3K pathway and cancer, in which elevated PI3K signalling is considered as a hallmark of cancer. The PI3K signalling pathway is activated by extracellular signals including diverse oncogenes and growth factor receptors such as epidermal growth factor receptor, platelet-derived growth factor receptor and the mesenchymal-epithelial transition factor (*Fruman et al., 2017*). This pathway is activated in human cancers *via* several different mechanisms. An increased level of PI3K signalling is often caused by a direct mutational activation or amplification of genes that encode key components of the PI3K pathway like PIK3CA and AKT1, or loss of PTEN. Genetic mutation, and/or amplification of the upregulation RTKs have also been shown to activate the PI3K pathway. This is also shown to be possible *via* oncogenic activation (mutations) of Ras (*Courtney et al., 2010*). Thus, clarifying the involvement of this pathway in transferring cancer-relevant signals (*Chalhoub and Baker, 2009*). The phosphatidylinositol 3-kinase (PI3K)/AKT/ (mTOR) pathway, which is one of the most common activated signalling pathways in cancer and that leads to cell proliferation, survival, and differentiation. Thus, it has become the focus of clinical research. Besides, the inhibition of the PI3K signalling pathway can reduce cell proliferation, and in some conditions, promote cell death (*Courtney et al., 2010*). Furthermore, the dysregulation expression of the central phosphatase in the PI3K/PTEN/Akt pathway led to drug resistance in breast cancer (*McCubrey et al., 2006*). The involvement of this pathway in various biological processes within the cell could clarify its up and down-regulation expression upon different treatment conditions and in both cell lines

(MDA-MB-468, and SkBr3). In addition, the resistance to Metformin treatment might refer to the deregulation of this pathway.

The transforming growth factor-beta (TGF- β) is a member of a superfamily of cytokines that act on protein kinase receptors found on the plasma membrane. TGF- β induces several biological functions that including embryonic development, wound healing, cell growth, death and differentiation, organogenesis, immune modulation, inflammation, and cancer progression. In epithelial cells, TGF- β is an essential regulatory tumour suppressor factor, where it induces early apoptosis and inhibits proliferation. The phenomenon of inhibiting cell growth in benign cells and promoting progression in cancer cells by TGF- β is known as a TGF- β paradox. Tumour cells had developed a mechanism to beat the TGF- β -induced suppressive effects. Once this happens, the cells might respond to this cytokine-inducing, which contributes to the tumour progression. In addition, TGF- β is an inducer of epithelial-mesenchymal transition (EMT), a pre-invasive process of tumour cells that facilitate cancer cell migration and invasion. TGF- β also mediates the production of mitogenic growth factors, that are involved in tumour proliferation and survival. Moreover, TGF- β is a known pro-angiogenic factor and immune suppressor that has been found overexpressed in a variety of human cancer types, and its expression correlated with angiogenesis, metastasis, tumour progression and patients' poor prognosis. Furthermore, TGF- β regulates cell cycle by arresting cell division at the early G1 phase and thus inhibit the growth of various cell types. It also activates the apoptotic pathway and limits cancer formation through downstream activation of pro-apoptotic factors including *death-associated protein kinase (DAPK)*, *growth arrest and DNA damage-inducible 45 (GADD45)* and Bim (Bcl-2 interacting mediator of cell death) or Bmf (Bcl-2 modifying factor) among others (*Fabregat et al., 2014*). Clearly, in Basal-like (MDA-MB-468) population samples downstream pathways were mostly involved in apoptosis, proliferation, survival and cell migration, which might explain the sensitiveness of this population to Metformin treatment and the significant death (apoptosis) of these cells after treatment. In the HER2 (SkBr3) cell line, one common pathway (Cell Cycle – Apoptosis- TGF- β) has been found downregulated, which is also associated with the regulation of cell growth and death through G1 phase cell cycle arrest and inhibition of cancer formation *via* negative stimulation of *DAPK* and *GADD45*. Additionally, (*GADD45*) has been noticed as upregulated in both breast cancer cell lines, which is explicitly linked to the downregulation of TGF-beta pathway. Once again, different techniques have yielded consistent results between different validation methods in this study. The outcomes of these findings revealed interactions between all these pathways and *PYK2* signalling, and that explain the susceptibility of Basal-like (MDA-MB-468) and the less sensitivity of HER2 (SkBr3) breast cancer phenotypes to Metformin

treatment. Besides, as previously emphasised, an ideal candidate marker has been selected for further studies which consist of *PYK2*.

Protein Tyrosine Kinase 2 Beta (PYK2) regulates the restructuring of the actin cytoskeleton, cell division, cell migration, adhesion, spreading and bone remodelling (Gao and Blystone, 2009; Rufanova et al., 2009; Lipinski and Loftus, 2010; Sun et al., 2011). It plays an essential role in the regulation of the humoral immune response and is required for normal levels of peripheral B-cells in the spleen and normal migration of splenic B-cells. It is also needed for macrophage divergence and migration towards sites of inflammation (Xu et al., 2008). It regulates cytoskeleton rearrangement and cell spreading in T-cells and contributes to the regulation of T-cell responses (Hjorthaug and Aasheim, 2007; Ruusala and Aspenstrom, 2008; Collins et al., 2010). It also promotes osteoclastic bone resorption, which required both *PTK2B/PYK2* and SRC. It also might inhibit the differentiation and activity of osteoprogenitor cells (Hendriks et al., 2013; Posritong et al., 2018). It functions in downstream signalling of integrin and collagen receptors, immune receptors, G-protein coupled receptors (GPCR), cytokine, chemokine and growth factor receptors, and mediates responses to cellular stress (Dylla et al., 2004; Schaller, 2010). Additionally, it forms multisubunit signalling complexes with SRC and SRC family members upon activation; which leads to the phosphorylation of additional tyrosine residues, creating binding sites for scaffold proteins, effectors and substrates (Park et al., 2004; Takahashi et al., 2005). *PYK2* regulates numerous signalling pathways (Roberts et al., 2008). It promotes activation of phosphatidylinositol 3-kinase and the AKT1 signalling cascade (Dikic, et al., 1996) and enhances the activation of NOS3. *PYK2* regulates the production of cellular messenger cGMP and promotes the activation of the MAP kinase signalling cascade including activation of MAPK1/ERK2, MAPK3/ERK1 and MAPK8/JNK1 (Tokiwa et al., 1996; Sun et al., 2008). It activates the Rho family GTPases such as RHOA and RAC1 and recruits the ubiquitin ligase MDM2 to P53/TP53 in the nucleus, and thereby regulates P53/TP53 activity, P53/TP53 ubiquitination and proteasomal degradation (Lim, 2013; Gao et al., 2015). It also acts as a scaffold, binding to both *PDPK1* and SRC, thus allowing SRC to phosphorylate *PDPK1* at Tyr-9, Tyr-373 and Tyr-376. Moreover, *PYK2* promotes phosphorylation of NMDA receptors by SRC family members, by which it contributes to the regulation of NMDA receptor ion channel activity and intracellular Ca (2+) levels (Lev, et al., 1995; Taniyama et al., 2003). Interestingly, it has been found that angiotensin II stimulates the tyrosine phosphorylation of *PDK1* in vascular smooth muscle in a calcium- and c-Src-dependent manner. The calcium-activated tyrosine kinase *PYK2* is acting as a scaffold for Src-dependent phosphorylation of *PDK1* on Tyr9 that allows phosphorylation of Tyr373 and -376 by Src. *PYK2* and tyrosine-phosphorylated *PDK1* colocalise in focal adhesions after angiotensin II stimulation (Taniyama et al., 2003). It might also regulate potassium ion

transport by phosphorylation of potassium channel subunits. Additionally, it phosphorylates SRC and increases the SRC kinase activity. It also phosphorylates *ASAP1*, *NPHP1*, *KCNA2* and *SHC1*, and promotes the phosphorylation of *ASAP2*, *RHOA* and *PXN*, which required both SRC and *PTK2/PYK2*. (Mandiyan et al., 1999 ; Kruljac-Leticic et al., 2003 ; Liebau et al., 2011).

In addition, a plethora of published studies revealed the role of *PYK2* in increased migration and invasion, an early event of the metastatic process. Lentiviral constructs transfected into breast cancer Basal-like (MDA-MB-468), Claudin -Low (MDA-MB-231), and HER2 (SkBr3, and MDA-MB-453) cell lines enabled the evaluation of *PYK2* role in breast cancer invasion and metastasis. *PYK2* silencing Basal-like, Claudin –Low, and HER2 cell lines confirmed allowed the establishment of a knockdown system that was used for the project. Wound-healing (Scratch) assay, Cultrex® BME Cell Invasion assay, and the xCELLigence System (RTCA) assay have been employed to verify the role of *PYK2* on fundamental biological functions such as proliferation, migration and invasion of breast cancer cells, and in response to Metformin treatment. *PYK2*-knockdown cells exhibited a significant reduction in migration and invasion of HER2 (SkBr3) and (MDA-MB-453) breast cancer cells, while dramatically increased their ability to proliferate after 48h of treatment and when compared to PLKO-1 cells (control). Similarly, in the Basal-like (MDA-MB-468) cell migration and invasion were significantly reduced following *PYK2* knockdown. The proliferation of cells was also decreased in comparison to the empty vector (control).

Previous studies have reported that *PYK2* depletion in breast cancer cells significantly decreases motility and invadopodia which affects invasion. Besides, *PYK2* has been found as a unique facilitator of invadopodia formation and function and further revealed a potential mechanism of *PYK2*-mediated tumour cell invasion (Genna et al., 2018). It also confirmed that *PYK2* and its closely related focal adhesion kinase (*FAK*) are actively involved in the regulation of cancer cell invasion, through balancing focal adhesion-mediated migration and invadopodia-dependent extracellular matrix invasion. This suggested that targeting either one or both kinases could block breast cancer metastasis (Genna and Gil-Henn, 2018). In summary, the results of this studies show a role of *PYK2* in preventing proliferation and promoting migration and invasion of breast cancer cell lines which is in line with evidence that showed that cells cannot migrate and proliferate simultaneously. Thus, one behaviour will have a direct influence on the other (Lipinski et al., 2005). An *in vitro* study has proved *FAK* as a promoter of cell cycle progression, particularly in malignant astrocytoma cells. *FAK* expression in glioma cell lines was implicated in the reduction of cell migration and an increase in cell cycle progression (Lipinski et al., 2005). A recent study has shown a functional link between cell cycle arrest and invasive activity and revealed a potential mechanism of switching from a proliferative to an intrusive state as a critical

aspect of metastasis. Other studies provided further evidence that confirms that invasiveness of basement membrane (BM) requires cell cycle arrest (*reviewed in Kohrman and Matus, 2017*).

On the contrary, Claudin-Low (MDA-MB-231) cells showed a completely different response pattern, with increased migration and invasion, and decreased proliferation following *PYK2* knockdown and when compared to the control which can be associated with the *PYK2* function in specific cell type migration. Evidence obtained from studies on *Pyk2*-deficient mice demonstrated *Pyk2* function in motility of particular cell types, such as macrophages which failed to become polarised or migrate despite regular expression of *Fak* in these cells. A similar study has highlighted the association of *Pyk2* depletion and decreased lymphocytes B migration in mice (*Okigaki et al., 2003; Lipinski et al., 2005*). It has also been found that the expression of an inactive *PYK2* variant substantially inhibited cell spreading and migration in brain microvascular endothelial cells (*Avraham et al., 2003*). Additionally, *PYK2* has a central activation role of critical signalling networks involved in cell motility, proliferation and G protein-coupled receptors (*Meyer et al., 2004*). Moreover, *PYK2* has also been found as a mediator in *STAT3* triggering (*Shi and Kehrl, 2004*) and was also involved in the induction of heregulin stimulation that enhances breast carcinoma invasion (*McShan et al., 2002*).

The synergistic effect of *PYK2* knockdown and Metformin treatment has been observed as significantly decreasing the invasion of Claudin-Low (MDA-MB-231), HER2 (SkBr3) and (MDA-MB-453) breast cancer cells. However, Metformin treatment did not affect the migration and proliferation of the knockdown cells apart from the proliferative capacity of the MDA-MB-453 cell line, which has been reduced following Metformin treatment. It was reported that the loss of endothelial cell *FAK* promotes the chemosensitisation of tumour cells to DNA-damaging therapies *in vitro* and *in vivo* by reducing DNA-damage-induced cytokine production (*Tavora et al., 2014*). Besides, it has been found that the invasive potential of MDA-MB-435 and MCF-7 breast cancer cells was increased in a *PYK2* expression manner and *via* activating Src and the MAP kinase pathways (*Zrihan-Licht et al., 2000*). Also, an increased level of *PYK2* expression has been observed in early and advanced breast cancer in comparison to benign and normal breast tissues (*Behmoaram et al., 2008*). Moreover, the overexpression of *PYK2* enhanced cell migration and invasion and promoted metastasis and EMT in Hep-3B HCC cells (*Sun et al., 2011*). Furthermore, *PYK2* suppression caused a reduction of tumour development and metastasis in pulmonary metastases carcinoma (*Fan and Guan, 2011*).

Meta Core™ search revealed the critical role of *PYK2* in autocrine Somatotropin signalling pathway in breast cancer. This verified the diverse functions of *PYK2* in breast cancer development and dissemination and its proliferation capacity. Interestingly, MS results strongly

indicated several common up and downregulated pathways in untreated and treated cells expressing PYK2 shRNAs. Upregulated pathways were associated with transcription and chromatin modification, protein folding and protein folding nucleus, reproduction and spermatogenesis, motility and copulation, and reproduction and male sex differentiation. Those pathways are highly likely related to the exceeding levels of growth hormone (GH) in breast cancer. It has been reported that the growth hormone (GH) has a physiological role in the control of male reproductive development and function at the normal level of expression (*Bartke, 2000*). GH is essential to promote the growth, division, and regeneration of the cells. It also exerts a diverse and widespread action on the human body following binding to its receptor (GHR). Additionally, GH has affected the metabolism of carbohydrates, lipids and proteins, and shapes body composition. It has distinct roles in supporting human development and homeostasis. GH continuous secretion stimulates the growth and takes part in the equilibrium of a process that is tightly organised and arranged by many organs. However, the deficiency in GH is a considerable medical condition, which affects all ages. Along with its good actions, it is also associated in the genesis of several diseases including cancer and insulin-resistant diabetes, as a consequence of GH/GHR interaction and (IGF-I) production (*Caicedo and Rosenfeld, 2017*).

Once GH is secreted by the somatotropic cells in the pituitary gland, it binds to its receptor (GHR) and then is transported to the entire body by the circulatory system. This protein receptor (GHR) is a class 1 cytokine receptor family member that is located on the cell membrane as a constitutive dimer. As a consequent of this binding, Janus kinases (JAK2) are activated, an action followed by activation of transcription activators (STATs) that induce the IGF-I, IGFBP3 and the leucine-rich protein known as acid labile subunit synthesis (ALS). IGF-1 is the central mediator of GH actions and is an insulinoid peptide with an amino acid sequence that is identified to that of proinsulin. The main route by which the GHR exerts its physiological effects is through the JAK2/STAT-5/IGF-I signalling pathway.

Nevertheless, an altered route via expression of multiple transcripts has been found in murine models with deletion of liver GHR. Additionally, GH may employ other paths such as mitogen-activated protein kinase (MAPK) to perform its activities. Following GH activation, several continuous activations will occur including RAS GTPase, RAF kinase and the MAP-ERK kinase (MEK). Because of these successive activations of the transcriptional regulation of target genes, the growth enhancement and metabolism will be achieved (*Guevara-Aguirrea et al., 2017*).

Another common upregulated pathway was transcription and chromatin modification, which is also related to cancer status. Chromatin modification is a useful modification of chromatin architecture that controls gene expression by exerting access of condensed genomic DNA to

regulatory transcription machinery proteins. Also, a dynamic chromatin modification carries an epigenetic regulatory role in several biological processes including DNA repair and replication, apoptosis, chromosomes division, development and pluripotency. This re-modelling is performed via covalent histone modifications by specific enzymes or ATP-dependent chromatin remodelling complexes. Likewise, variation in chromatin remodelling proteins was found to be related to human diseases that comprise cancer. Currently, a dominant therapeutic strategy in the treatment of diversified tumours was developed to targets chromatin remodelling pathways (Wang *et al.*, 2007). Up-regulation of this pathway could be related to Metformin action as a DNA damage agent. This pathway is utilising several biological processes for DNA repair and replication. Furthermore, it is evident that oncogenesis is linked to miswriting, misinterpretation and misreading of histone modifications, proposing that the deregulation in the gene expression and perturbation of cellular identity can be caused by histone code misregulation, contributing to the initiation of cancer, and its progression through metastasis (Chi *et al.*, 2010).

On the other hand, the cell cycle and mitosis pathways were downregulated in control samples compared to treated samples. Metformin action has activated signalling pathways that promote cell cycle arrest, and DNA repair as a response to DNA damage in eukaryotic cells. Metformin cytotoxicity, oxidative stress, DNA damage, and intracellular pathways related to cell growth and survival have also been shown in MCF-7 and MDA-MB-231 human breast cancer cells (Marinello *et al.*, 2016).

Apoptosis and mitochondria-mediated apoptosis, and response to hypoxia and oxidative stress pathways were also up-regulated in PLKO-1 Control vs PLKO-1 Treated, which could be caused by Metformin treatment. Metformin functions directly on mitochondria by limiting citric acid cycle activity and Oxidative phosphorylation (OXPHOS), as has been shown in isolated mitochondria and in intact cells. A compensatory increase in glycolysis accompanied the Metformin-mediated decrease in mitochondrial function. The sensitivity of Metformin is dependent on cells capabilities to engage aerobic glycolysis. Thus, Metformin could potentially be used in oncology to capitalise on the metabolic sensitivity of cancer cells (Andrzejewski *et al.*, 2014). A study recently published revealed that constant subjection to Metformin in cancer cells eventually leads to drug resistance that is related to increased PGC-1 α levels. Metformin resistant cells are metabolically flexible and have the capability to switch nutrition sources from oxidative metabolism to glycolysis and glutamine metabolism in the condition of Metformin-mediated inhibition of oxidative phosphorylation. An increased level of PGC-1 α is a crucial regulator of OXPHOS and mitochondrial biogenesis, during inhibition of OXPHOS by Metformin (Andrzejewski *et al.*, 2018).

In summary, different breast cancer phenotypes display selective sensitivity to metformin. Triple-negative (Basal-like and Claudin-low) subtypes were the most sensitive to Metformin treatment, followed by Luminal B, while HER2 positive and Luminal A were the less sensitive phenotypes. Interestingly, Metformin significantly enhanced PYK2 expression in HER2 cell lines and other phenotypes, however, less abundantly. PYK2 promoted invasion and migration and impacted the proliferation of breast cancer cell lines. Several techniques have been employed in this investigation to emphasise the previously observed phenomenon, that revealed the association of PYK2 with breast cancer progression and dissemination. PYK2 has been found to signal through several pathways that are involved in invasion and metastasis of breast cancer. In addition to others that are associated with cancer stem cells in breast cancer. PYK2 might be the suggested target in HER2 breast cancer therapy, and Metformin may be a promising candidate in triple negative breast cancer treatment. These findings will require further investigations using Immunohistochemistry staining, animal models, and preclinical trials which better reflect real patients. The outcomes of this study could influence the choice of medication according to different breast cancer phenotypes.

7.2. Future work

This PhD project aimed to interrogate Metformin efficacy on several breast cancer molecular subtypes. The aim of this project has been achieved, and Metformin treatment has induced the expression of *PYK2* in HER2 positive phenotype. *PYK2* drives crucial biological functions including cancer initiation and dissemination. The second achievement was to find a putative breast cancer marker that could be used as a biomarker and potentially as a target for breast cancer therapy. This was also met by identifying the potential marker which is *PYK2*. The expression of *PYK2* under the levels of gene and protein was validated. Besides, the involvement of *PYK2* in breast cancer progression and metastasis was also assessed. However, to fully assess *PYK2* suitability for a small molecule targeted therapy in breast cancer, further Immunohistochemistry staining, animal work and preclinical trials should be implemented. A significant number of samples and a greater variety of breast cancer tissue samples that include different molecular subtypes would allow a more detailed study. Besides, the application of the ANN integrative data mining approach for clinical datasets validation would be additionally helpful. Finally, outcomes of clinical trials using Metformin Hydrochloride for breast cancer supported by the National Cancer Institute (NCI) will provide significant information about the potential use of Metformin in breast cancer treatment.

BIBLIOGRAPHY

- Ahn, A.C., Tewari, M., Poon, C. and Phillips, R.S., 2006. The limits of reductionism in medicine: could systems biology offer an alternative? *PLoS Medicine*, 3 (6), e208.
- Albert, R., and Barabási, A., 2002. Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74 (1), 47.
- Aleshin, V.A., Artiukhov, A.V., Oppermann, H., Kazantsev, A.V., Lukashev, N.V. and Bunik, V.I., 2015. Mitochondrial impairment may increase cellular NAD (P) H: resazurin oxidoreductase activity, perturbing the NAD (P) H-based viability assays. *Cells*, 4 (3), 427-451.
- American Diabetes Association, 2014. Standards of medical care in diabetes--2014. *Diabetes Care*, 37 Suppl 1, S14-80.0
- Andrzejewski, S., Gravel, S., Pollak, M. and St-Pierre, J., 2014. Metformin directly acts on mitochondria to alter cellular bioenergetics. *Cancer & Metabolism*, 2 (1), 12.
- Andrzejewski, S., Siegel, P.M. and St-Pierre, J., 2018. Metabolic Profiles Associated with Metformin Efficacy in Cancer. *Frontiers in Endocrinology*, 9.
- Araújo, T., Aresta, G., Castro, E., Rouco, J., Aguiar, P., Eloy, C., Polónia, A. and Campilho, A., 2017. Classification of breast cancer histology images using convolutional neural networks. *PLoS One*, 12 (6), e0177544.
- Avraham, H.K., Lee, T.H., Koh, Y., Kim, T.A., Jiang, S., Sussman, M., Samarel, A.M. and Avraham, S., 2003. Vascular endothelial growth factor regulates focal adhesion assembly in human brain microvascular endothelial cells through activation of the focal adhesion kinase and related adhesion focal tyrosine kinase. *The Journal of Biological Chemistry*, 278 (38), 36661-36668.
- Ayer, T., Alagoz, O., Chhatwal, J., Shavlik, J.W., Kahn Jr, C.E. and Burnside, E.S., 2010. Breast cancer risk estimation with artificial neural networks revisited: discrimination and calibration. *Cancer*, 116 (14), 3310-3321.
- Barabási, A., Gulbahce, N. and Loscalzo, J., 2011. Network medicine: a network-based approach to human disease. *Nature Reviews Genetics*, 12 (1), 56.
- Barabasi, A., and Oltvai, Z.N., 2004. Network biology: understanding the cell's functional organisation. *Nature Reviews Genetics*, 5 (2), 101.
- BARTKE, A., 2000. Effects of Growth Hormone Review on Male Reproductive Functions. *Journal of Andrology*, 21 (2), 181-188.
- Behmoaram, E., Bijian, K., Jie, S., Xu, Y., Darnel, A., Bismar, T.A. and Alaoui-Jamali, M.A., 2008. Focal adhesion kinase-related proline-rich tyrosine kinase 2 and focal adhesion kinase are co-overexpressed in early-stage and invasive ErbB-2-positive breast cancer and cooperate for breast cancer cell tumorigenesis and invasiveness. *The American Journal of Pathology*, 173 (5), 1540-1550.
- Belizário, J., 2018. Cancer Risks Linked to the Bad Luck Hypothesis and Epigenomic Mutational Signatures. *Epigenomes*, 2 (3), 13.

Ben Sahra, I., Regazzetti, C., Robert, G., Laurent, K., Le Marchand-Brustel, Y., Auberger, P., Tanti, J.F., Giorgetti-Peraldi, S. and Bost, F., 2011. Metformin, independent of AMPK, induces mTOR inhibition and cell-cycle arrest through REDD1. *Cancer Research*, 71 (13), 4366-4372.

Berg, C.J., Haardoerfer, R., Escoffery, C., Zheng, P. and Kegler, M., 2014. Cigarette users' interest in using or switching to electronic nicotine delivery systems for smokeless tobacco for harm reduction, cessation, or novelty: a cross-sectional survey of US adults. *Nicotine & Tobacco Research*, 17 (2), 245-255.

Bertram, J.S., 2000. The molecular biology of cancer. *Molecular Aspects of Medicine*, 21 (6), 167-223.

Bishop, C.M., 2006. Graphical models. *Pattern Recognition and Machine Learning*, 4, 359-422.

Bishop, C.M., 1995. *Neural networks for pattern recognition*. Oxford university press.

Bodmer, M., Meier, C., Krahenbuhl, S., Jick, S.S. and Meier, C.R., 2010. Long-term metformin use is associated with decreased risk of breast cancer. *Diabetes Care*, 33 (6), 1304-1308.

America. American Cancer Society, 2016. Ductal Carcinoma In Situ (DCIS) - American Cancer Society. Available at: <https://www.cancer.org/cancer/breastcancer/understanding-a-breast-cancer-diagnosis/types-of-breast-cancer/dcis.html> [Accessed 30 August 2018].

Ananya Mandal, MD, 2009. Breast Cancer Classification. News-Medical.net. Available at: <https://www.cancer.org/cancer/breast-cancer/understanding-a-breast-cancerdiagnosis/types-of-breast-cancer/dcis.html> [Accessed 30 August 2018].

Bumgarner, R., 2013. Overview of DNA microarrays: types, applications, and their future. *Current Protocols in Molecular Biology*, 101 (1), 22.1. 1-22.1. 11.

Butcher, E.C., Berg, E.L. and Kunkel, E.J., 2004. Systems biology in drug discovery. *Nature Biotechnology*, 22 (10), 1253.

Caicedo, A., and Rosenfeld, R., 2018. Challenges and future for the delivery of growth hormone therapy. *Growth Hormone & IGF Research*, 38, 39-43.

Great Britain. Breast cancer statistics, 2016. Cancer Research UK. Available at: <https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer/dcis.html> [Accessed September 2018].

Great Britain. Breast cancer statistics, 2016. Cancer Research UK. Available at: <https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer/dcis.html> [Accessed September 2018].

Cao, J., Liu, J., Long, J., Fu, J., Huang, L., Li, J., Liu, C., Zhang, X. and Yan, Y., 2017. microRNA-23b suppresses epithelial-mesenchymal transition (EMT) and metastasis in hepatocellular carcinoma via targeting Pyk2. *Biomedicine & Pharmacotherapy*, 89, 642-650.

Chalhoub, N., and Baker, S.J., 2009. PTEN and the PI3-kinase pathway in cancer. *Annual Review of Pathological Mechanical Disease*, 4, 127-150.

- Chalhoub, N., Zhu, G., Zhu, X. and Baker, S.J., 2009. Cell type specificity of PI3K signaling in Pdk1- and Pten-deficient brains. *Genes & Development*, 23 (14), 1619-1624.
- Cheang, M.C., Chia, S.K., Voduc, D., Gao, D., Leung, S., Snider, J., Watson, M., Davies, S., Bernard, P.S. and Parker, J.S., 2009. Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. *JNCI: Journal of the National Cancer Institute*, 101 (10), 736-750.
- Chen, P., and Kroog, G.S., 2005. Gastrin-Releasing Peptide Receptor Regulates Leupaxin Localization and Pyk2 Activity.
- Chi, P., Allis, C.D. and Wang, G.G., 2010. Covalent histone modifications—miswritten, misinterpreted and mis-erased in human cancers. *Nature Reviews Cancer*, 10 (7), 457.
- Christley, S., Nie, Q. and Xie, X., 2009. Incorporating existing network information into gene network inference. *PLoS One*, 4 (8), e6799.
- Cifarelli, V., Lashinger, L.M., Devlin, K.L., Dunlap, S.M., Huang, J., Kaaks, R., Pollak, M.N. and Hursting, S.D., 2015. Metformin and Rapamycin Reduce Pancreatic Cancer Growth in Obese Prediabetic Mice by Distinct MicroRNA-Regulated Mechanisms. *Diabetes*, 64 (5), 1632-1642.
- Clarke, R., Ransom, H.W., Wang, A., Xuan, J., Liu, M.C., Gehan, E.A. and Wang, Y., 2008. The properties of high-dimensional data spaces: implications for exploring gene and protein expression data. *Nature Reviews Cancer*, 8 (1), 37.
- United States, National Library of Medicine at the National Institutes of Health, 2018. *ClinicalTrials.gov - Official Site*. Available at: [https:// www. clinicaltrials.gov//dcis.html](https://www.clinicaltrials.gov/dcis.html) [Accessed September 2018].
- Col, N.F., Ochs, L., Springmann, V., Aragaki, A.K. and Chlebowski, R.T., 2012. Metformin and breast cancer risk: a meta-analysis and critical literature review. *Breast Cancer Research and Treatment*, 135 (3), 639-646.
- Collado, M., and Serrano, M., 2010. Senescence in tumours: evidence from mice and humans. *Nature Reviews Cancer*, 10 (1), 51.
- Collins, M., Bartelt, R.R. and Houtman, J.C., 2010. T cell receptor activation leads to two distinct phases of Pyk2 activation and actin cytoskeletal rearrangement in human T cells. *Molecular Immunology*, 47 (9), 1665-1674.
- Cooper, C.S., 2001. Applications of microarray technology in breast cancer research. *Breast Cancer Research*, 3 (3), 158.
- Cornejo, K.M., Kandil, D., Khan, A. and Cosar, E.F., 2014. Theranostic and molecular classification of breast cancer. *Archives of Pathology and Laboratory Medicine*, 138 (1), 44-56.
- Courtney, K.D., Corcoran, R.B. and Engelman, J.A., 2010. The PI3K pathway as drug target in human cancer. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, 28 (6), 1075-1083.
- Coyle, C., Cafferty, F., Vale, C. and Langley, R., 2016. Metformin as an adjuvant treatment for cancer: a systematic review and meta-analysis. *Annals of Oncology*, 27 (12), 2184-2195.

Croft, D., Mundo, A.F., Haw, R., Milacic, M., Weiser, J., Wu, G., Caudy, M., Garapati, P., Gillespie, M. and Kamdar, M.R., 2013. The Reactome pathway knowledgebase. *Nucleic Acids Research*, 42 (D1), D472-D477.

Currie, C., Poole, C. and Gale, E., 2009. The influence of glucose-lowering therapies on cancer risk in type 2 diabetes. *Diabetologia*, 52 (9), 1766-1777.

Curto, M., Cole, B.K., Lallemand, D., Liu, C.H. and McClatchey, A.I., 2007. Contact-dependent inhibition of EGFR signaling by Nf2/Merlin. *The Journal of Cell Biology*, 177 (5), 893-903.

Cyr, D.G., Devine, P.J. and Plante, I., 2016. Immunohistochemistry and Female Reproductive Toxicology: The Ovary and Mammary Glands. In: *Immunohistochemistry and Female Reproductive Toxicology: The Ovary and Mammary Glands. Technical Aspects of Toxicological Immunohistochemistry*. Springer, 113-145.

Dadkhah, K., et al., 2015. DNA microarray, types and its application in medicine. *Scholars Academic Journal of Biosciences*, 3 (7), 598-599.

Dallaglio, K., Bruno, A., Cantelmo, A.R., Esposito, A.I., Ruggiero, L., Orecchioni, S., Calleri, A., Bertolini, F., Pfeffer, U. and Noonan, D.M., 2014. Paradoxical effects of metformin on endothelial cells and angiogenesis. *Carcinogenesis*, 35 (5), 1055-1066.

Damjanović, A., Matić, I.Z., Đorđić, M., Đurović, M.N., Nikolić, S., Roki, K., Milovanović, Z., Antić-Stanković, J., Džodić, R. and Damjanović, S., 2015. Metformin effects on malignant cells and healthy PBMC; the influence of metformin on the phenotype of breast cancer cells. *Pathology & Oncology Research*, 21 (3), 605-612.

Daniel, A.R., Hagan, C.R. and Lange, C.A., 2011. Progesterone receptor action: defining a role in breast cancer. *Expert Review of Endocrinology & Metabolism*, 6 (3), 359-369.

Daugan, M., Wojcicki, A.D., d'Hayer, B. and Boudy, V., 2016. Metformin: An anti-diabetic drug to fight cancer. *Pharmacological Research*, 113, 675-685.

de Koning, P.J., Kummer, J.A., de Poot, S.A., Quadir, R., Broekhuizen, R., McGettrick, A.F., Higgins, W.J., Devreese, B., Worrall, D.M. and Bovenschen, N., 2011. Intracellular serine protease inhibitor SERPINB4 inhibits granzyme M-induced cell death. *PLoS One*, 6 (8), e22645.

de Macêdo Andrade, Ana Cláudia, Júnior, C.A.F., Guimarães, B.D., Barros, A.W.P., de Almeida, G.S. and Weller, M., 2014. Molecular breast cancer subtypes and therapies in a public hospital of Northeastern Brazil. *BMC Women's Health*, 14 (1), 110.

Decensi, A., Puntoni, M., Goodwin, P., Cazzaniga, M., Gennari, A., Bonanni, B. and Gandini, S., 2010. Metformin and cancer risk in diabetic patients: a systematic review and meta-analysis. *Cancer Prevention Research (Philadelphia, Pa.)*, 3 (11), 1451-1461.

Deng, X., Wang, S., Deng, A., Liu, B., Edgerton, S.M., Lind, S.E., Wahdan-Alaswad, R. and Thor, A.D., 2012. Metformin targets Stat3 to inhibit cell growth and induce apoptosis in triple-negative breast cancers. *Cell Cycle*, 11 (2), 367-376.

Dikic, I., Tokiwa, G., Lev, S., Courtneidge, S.A. and Schlessinger, J., 1996. A role for Pyk2 and Src in linking G-protein-coupled receptors with MAP kinase activation. *Nature*, 383 (6600), 547.

Dowling, R.J., Zakikhani, M., Fantus, I.G., Pollak, M. and Sonenberg, N., 2007. Metformin inhibits mammalian target of rapamycin-dependent translation initiation in breast cancer cells. *Cancer Research*, 67 (22), 10804-10812.

Duchartre, Y., Kim, Y. and Kahn, M., 2016. The Wnt signaling pathway in cancer. *Critical Reviews in oncology/hematology*, 99, 141-149.

Dylla, S.J., Deyle, D.R., Theunissen, K., Padurean, A.M. and Verfaillie, C.M., 2004. Integrin engagement-induced inhibition of human myelopoiesis is mediated by proline-rich tyrosine kinase 2 gene products. *Experimental Hematology*, 32 (4), 365-374.

Elsevier Launches First Assessment of Cancer Research on Global 2016, <https://www.elsevier.com> Science and Technology - Press Info - Press Releases

Eroles, P., Bosch, A., Pérez-Fidalgo, J.A. and Lluch, A., 2012. Molecular biology in breast cancer: intrinsic subtypes and signaling pathways. *Cancer Treatment Reviews*, 38 (6), 698-707.

Evans, J.M., Donnelly, L.A., Emslie-Smith, A.M., Alessi, D.R. and Morris, A.D., 2005. Metformin and reduced risk of cancer in diabetic patients. *BMJ (Clinical Research Ed.)*, 330 (7503), 1304-1305.

Fabregat, I., Fernando, J., Mainez, J. and Sancho, P., 2014. TGF-beta signaling in cancer treatment. *Current Pharmaceutical Design*, 20 (17), 2934-2947.

Fan, H., and Guan, J.L., 2011. Compensatory function of Pyk2 protein in the promotion of focal adhesion kinase (FAK)-null mammary cancer stem cell tumorigenicity and metastatic activity. *The Journal of Biological Chemistry*, 286 (21), 18573-18582.

Ferrara, N., 2009. Vascular endothelial growth factor. *Arteriosclerosis, Thrombosis, and Vascular Biology*, 29 (6), 789-791.

Fruman, D.A., Chiu, H., Hopkins, B.D., Bagrodia, S., Cantley, L.C. and Abraham, R.T., 2017. The PI3K pathway in human disease. *Cell*, 170 (4), 605-635.

Galani, I.E., Triantafyllia, V., Eleminiadou, E., Koltsida, O., Stavropoulos, A., Manioudaki, M., Thanos, D., Doyle, S.E., Kotenko, S.V. and Thanopoulou, K., 2017. Interferon- λ mediates non-redundant front-line antiviral protection against influenza virus infection without compromising host fitness. *Immunity*, 46 (5), 875-890. e6.

Gao, C., Chen, G., Kuan, S., Zhang, D.H., Schlaepfer, D.D. and Hu, J., 2015. FAK/PYK2 promotes the Wnt/ β -catenin pathway and intestinal tumorigenesis by phosphorylating GSK3 β . *Elife*, 4, e10072.

Gao, C., and Blystone, S.D., 2009. A Pyk2-Vav1 complex is recruited to beta3-adhesion sites to initiate Rho activation. *The Biochemical Journal*, 420 (1), 49-56.

Genna, A., and Gil-Henn, H., 2018. FAK family kinases: The Yin and Yang of cancer cell invasiveness. *Molecular & Cellular Oncology*, (just-accepted), 01-10.

Genna, A., Lapetina, S., Lukic, N., Twafra, S., Meirson, T., Sharma, V.P., Condeelis, J.S. and Gil-Henn, H., 2018. Pyk2 and FAK differentially regulate invadopodia formation and function in breast cancer cells. *The Journal of Cell Biology*, 217 (1), 375-395.

Giannoni, E., Taddei, M.L., Morandi, A., Comito, G., Calvani, M., Bianchini, F., Richichi, B., Raugei, G., Wong, N., Tang, D. and Chiarugi, P., 2015. Targeting stromal-induced pyruvate kinase M2 nuclear translocation impairs oxphos and prostate cancer metastatic spread. *Oncotarget*, 6 (27), 24061-24074.

Gil, E.M.C., 2014. Targeting the PI3K/AKT/mTOR pathway in estrogen receptor-positive breast cancer. *Cancer Treatment Reviews*, 40 (7), 862-871.

Godfrey, K., Agatha, T. and Nankumbi, J., 2016. Breast Cancer Knowledge and Breast Self-Examination Practices Among Female University Students in Kampala, Uganda: A Descriptive Study. *Oman Medical Journal*, 31 (2), 129-134.

Gomez-Martin, C., Plaza, J.C., Pazo-Cid, R., Salud, A., Pons, F., Fonseca, P., Leon, A., Alsina, M., Visa, L. and Rivera, F., 2013. Level of HER2 gene amplification predicts response and overall survival in HER2-positive advanced gastric cancer treated with trastuzumab. *Journal of Clinical Oncology*, 31 (35), 4445-4452.

Gong, J., Luk, F., Jaiswal, R. and Bebawy, M., 2014. Microparticles mediate the intercellular regulation of microRNA-503 and proline-rich tyrosine kinase 2 to alter the migration and invasion capacity of breast cancer cells. *Frontiers in Oncology*, 4, 220.

Govindarajan, R., Duraiyan, J., Kaliyappan, K. and Palanisamy, M., 2012. Microarray and its applications. *Journal of Pharmacy & Bioallied Sciences*, 4 (Suppl 2), S310-2.

Guevara-Aguirre, J., Guevara, A., Palacios, I., Pérez, M., Prócel, P. and Terán, E., 2018. GH and GHR signaling in human disease. *Growth Hormone & IGF Research*, 38, 34-38.

Guimaraes, I.S., Tessarollo, N.G., Oliveira, L.F., Zampier, R.C., Silva, I.V., Sternberg, C. and Rangel, L.B., 2015. Metformin Inhibits Proliferation and Acts Synergistically with Paclitaxel and Doxorubicin in Triple Negative Breast Cancer Cell Lines.

Gutenberg, A., Brück, W., Buchfelder, M. and Ludwig, H., 2004. Expression of tyrosine kinases FAK and Pyk2 in 331 human astrocytomas. *Acta Neuropathologica*, 108 (3), 224-230.

Hadad, S.M., Baker, L., Quinlan, P.R., Robertson, K.E., Bray, S.E., Thomson, G., Kellock, D., Jordan, L.B., Purdie, C.A. and Hardie, D.G., 2009. Histological evaluation of AMPK signalling in primary breast cancer. *BMC Cancer*, 9 (1), 307.

Hadad, S., Iwamoto, T., Jordan, L., Purdie, C., Bray, S., Baker, L., Jellema, G., Deharo, S., Hardie, D.G. and Puztai, L., 2011. Evidence for biological effects of metformin in operable breast cancer: a pre-operative, window-of-opportunity, randomized trial. *Breast Cancer Research and Treatment*, 128 (3), 783-794.

Hadisaputri, Y.E., Miyazaki, T., Yokobori, T., Sohda, M., Sakai, M., Ozawa, D., Hara, K., Honjo, H., Kumakura, Y. and Kuwano, H., 2017. TNFAIP3 overexpression is an independent factor for poor survival in esophageal squamous cell carcinoma. *International Journal of Oncology*, 50 (3), 1002-1010.

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

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

- Harburger, D.S., and Calderwood, D.A., 2009. Integrin signalling at a glance. *Journal of Cell Science*, 122 (Pt 2), 159-163.
- Hartemink, A.J., Gifford, D.K., Jaakkola, T.S. and Young, R.A., 2002. Bayesian methods for elucidating genetic regulatory networks. *IEEE Intelligent Systems*, 17 (2), 37-43.
- Haykin, S.S., Haykin, S.S., Haykin, S.S. and Haykin, S.S., 2009. *Neural networks and learning machines*. Pearson Upper Saddle River.
- Heath, A.P., and Kavraki, L.E., 2009. Computational challenges in systems biology. *Computer Science Review*, 3 (1), 1-17.
- Heemskerk, V.H., Daemen, M.A. and Buurman, W.A., 1999. Insulin-like growth factor-1 (IGF-1) and growth hormone (GH) in immunity and inflammation. *Cytokine & Growth Factor Reviews*, 10 (1), 5-14.
- Hendriks, W.J., Elson, A., Harroch, S., Pulido, R., Stoker, A. and den Hertog, J., 2013. Protein tyrosine phosphatases in health and disease. *The FEBS Journal*, 280 (2), 708-730.
- Hirsch, H.A., Iliopoulos, D. and Struhl, K., 2013. Metformin inhibits the inflammatory response associated with cellular transformation and cancer stem cell growth. *Proceedings of the National Academy of Sciences of the United States of America*, 110 (3), 972-977.
- Hjorthaug, H.S., and Aasheim, H., 2007. Ephrin-A1 stimulates migration of CD8 CCR7 T lymphocytes. *European Journal of Immunology*, 37 (8), 2326-2336.
- Holliday, D.L., and Speirs, V., 2011. Choosing the right cell line for breast cancer research. *Breast Cancer Research*, 13 (4), 215.
- Hood, L., and Perlmutter, R.M., 2004. The impact of systems approaches on biological problems in drug discovery. *Nature Biotechnology*, 22 (10), 1215.
- Horn, H., and Vousden, K., 2007. Coping with stress: multiple ways to activate p53. *Oncogene*, 26 (9), 1306.
- Horne, M., Zimmermann, M., Arachchige Don, A., Donaldson, M. and Patriarchi, T., 2015. Cyclin G2 contributes to the cell cycle arrest response of breast cancer cells to estrogen signaling-antagonists and the AMPK agonist, metformin. *The FASEB Journal*, 29 (1_supplement), 576.10.
- Hsiao, Y.H., Huang, Y.T., Hung, C.Y., Kuo, T.C., Luo, F.J. and Yuan, T.C., 2016. PYK2 via S6K1 regulates the function of androgen receptors and the growth of prostate cancer cells. *Endocrine-Related Cancer*, 23 (8), 651-663.
- Hsieh, M., Lee, T., Cheng, S., Tu, S., Yen, M. and Tseng, C., 2012. The influence of type 2 diabetes and glucose-lowering therapies on cancer risk in the Taiwanese. *Experimental Diabetes Research*, 2012.
- Hu, H., Juvekar, A., Lyssiotis, C.A., Lien, E.C., Albeck, J.G., Oh, D., Varma, G., Hung, Y.P., Ullas, S. and Lauring, J., 2016. Phosphoinositide 3-kinase regulates glycolysis through mobilization of aldolase from the actin cytoskeleton. *Cell*, 164 (3), 433-446.
- Huang, D.W., Sherman, B.T. and Lempicki, R.A., 2008. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, 4 (1), 44.

Huveneers, S., and Danen, E.H., 2009. Adhesion signaling - crosstalk between integrins, Src and Rho. *Journal of Cell Science*, 122 (Pt 8), 1059-1069.

Ikhlas, S., and Ahmad, M., 2017. Metformin: Insights into its anticancer potential with special reference to AMPK dependent and independent pathways. *Life Sciences*, 185, 53-62.

Ito, H., Terai, T., Hanaoka, K., Ueno, T., Komatsu, T., Nagano, T. and Urano, Y., 2015. Detection of NAD (P) H-dependent enzyme activity with dynamic luminescence quenching of terbium complexes. *Chemical Communications*, 51 (39), 8319-8322.

Ivankovic-Dikic, I., Grönroos, E., Blaukat, A., Barth, B. and Dikic, I., 2000. Pyk2 and FAK regulate neurite outgrowth induced by growth factors and integrins. *Nature Cell Biology*, 2 (9), 574.

Izuhara, K., Yamaguchi, Y., Ohta, S., Nunomura, S., Nanri, Y., Azuma, Y., Nomura, N., Noguchi, Y. and Aihara, M., 2018. Squamous Cell Carcinoma Antigen 2 (SCCA2, SERPINB4): An Emerging Biomarker for Skin Inflammatory Diseases. *International Journal of Molecular Sciences*, 19 (4), 1102.

Take, S., Usui, T., Ohama, T., Yamawaki, H. and Sato, K., 2017. Death-associated protein kinase 3 controls the tumor progression of A549 cells through ERK MAPK/c-Myc signaling. *Oncology Reports*, 37 (2), 1100-1106.

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

Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., Itoh, M., Katayama, T., Kawashima, S., Okuda, S. and Tokimatsu, T., 2007. KEGG for linking genomes to life and the environment. *Nucleic Acids Research*, 36 (suppl_1), D480-D484.

Kim, I., Lee, M., Lee, I., Shin, S. and Lee, S., 2000. Gene expression of flap endonuclease-1 during cell proliferation and differentiation. *Biochimica Et Biophysica Acta (BBA)-Molecular Cell Research*, 1496 (2-3), 333-340.

Kim, E.H., Potretzke, A.M., Figenschau, R.S. and Perrino, C.M., 2015. Metastatic renal cell carcinoma presenting as a thyroid mass. *The Journal of Urology*, 193 (2), 677-678.

Klil-Drori, A.J., Azoulay, L. and Pollak, M.N., 2017. Cancer, obesity, diabetes, and antidiabetic drugs: is the fog clearing? *Nature Reviews Clinical Oncology*, 14 (2), 85.

Kohrman, A.Q., and Matus, D.Q., 2017. Divide or conquer: cell cycle regulation of invasive behavior. *Trends in Cell Biology*, 27 (1), 12-25.

Komiya, Y., and Habas, R., 2008. Wnt signal transduction pathways. *Organogenesis*, 4 (2), 68-75.

Konda, J.D., Olivero, M., Musiani, D., Lamba, S. and Di Renzo, M.F., 2017. Heat-shock protein 27 (HSP 27, HSPB 1) is synthetic lethal to cells with oncogenic activation of MET, EGFR and BRAF. *Molecular Oncology*, 11 (6), 599-611.

Kourou, K., Exarchos, T.P., Exarchos, K.P., Karamouzis, M.V. and Fotiadis, D.I., 2015. Machine learning applications in cancer prognosis and prediction. *Computational and Structural Biotechnology Journal*, 13, 8-17.

Kraus, S., Naor, Z. and Seger, R., 2001. Intracellular signaling pathways mediated by the gonadotropin-releasing hormone (GnRH) receptor. *Archives of Medical Research*, 32 (6), 499-509.

Kruljac-Letic, A., Moelleken, J., Kallin, A., Wieland, F. and Blaukat, A., 2003. The tyrosine kinase Pyk2 regulates Arf1 activity by phosphorylation and inhibition of the Arf-GTPase-activating protein ASAP1. *The Journal of Biological Chemistry*, 278 (32), 29560-29570.

Lai, S., Liao, K., Chen, P., Tsai, P., Hsieh, D.P.H. and Chen, C., 2012. Antidiabetes drugs correlate with decreased risk of lung cancer: a population-based observation in Taiwan. *Clinical Lung Cancer*, 13 (2), 143-148.

Lamouille, S., Xu, J. and Derynck, R., 2014. Molecular mechanisms of epithelial–mesenchymal transition. *Nature Reviews Molecular Cell Biology*, 15 (3), 178.

Lancashire, L.J., Lemetre, C. and Ball, G.R., 2009. An introduction to artificial neural networks in bioinformatics—application to complex microarray and mass spectrometry datasets in cancer studies. *Briefings in Bioinformatics*, 10 (3), 315-329.

Lancashire, L.J., Powe, D., Reis-Filho, J., Rakha, E., Lemetre, C., Weigelt, B., Abdel-Fatah, T., 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.

Lancashire, L.J., Rees, R.C. and Ball, G.R., 2008. Identification of gene transcript signatures predictive for estrogen receptor and lymph node status using a stepwise forward selection artificial neural network modelling approach. *Artificial Intelligence in Medicine*, 43 (2), 99-111.

Laplanche, M., and Sabatini, D.M., 2009. mTOR signaling at a glance. *Journal of Cell Science*, 122 (Pt 20), 3589-3594.

Lee, M., Hsu, C., Wahlqvist, M.L., Tsai, H., Chang, Y. and Huang, Y., 2011. Type 2 diabetes increases, and metformin reduces total, colorectal, liver and pancreatic cancer incidences in Taiwanese: a representative population prospective cohort study of 800,000 individuals. *BMC Cancer*, 11 (1), 20.

Lemetre, C., 2010. Artificial Neural Network Techniques to Investigate Potential Interactions between Biomarkers.

Lemetre, C., Lancashire, L.J., Rees, R.C. and Ball, G.R., 2009. Artificial neural network-based algorithm for biomolecular interactions modeling. In: *International Work-Conference on Artificial Neural Networks*, Springer, 877-885.

Lev, S., Moreno, H., Martinez, R., Canoll, P., Peles, E., Musacchio, J., Plowman, G., Rudy, B. and Schlessinger, J., 1995. Protein tyrosine kinase PYK2 involved in Ca²⁺-induced regulation of ion channel and MAP kinase functions. *Nature*, 376 (6543), 737.

Li, W., Saud, S.M., Young, M.R., Chen, G. and Hua, B., 2015. Targeting AMPK for cancer prevention and treatment. *Oncotarget*, 6 (10), 7365-7378.

Li, W., Wang, Q.L., Liu, X., Dong, S.H., Li, H.X., Li, C.Y., Guo, L.S., Gao, J.M., Berger, N.A., Li, L., Ma, L. and Wu, Y.J., 2015. Combined use of vitamin D3 and metformin exhibits synergistic

- chemopreventive effects on colorectal neoplasia in rats and mice. *Cancer Prevention Research (Philadelphia, Pa.)*, 8 (2), 139-148.
- Libby, G., Donnelly, L.A., Donnan, P.T., Alessi, D.R., Morris, A.D. and Evans, J.M., 2009. New users of metformin are at low risk of incident cancer: a cohort study among people with type 2 diabetes. *Diabetes Care*, 32 (9), 1620-1625.
- Liebau, M.C., Hopker, K., Muller, R.U., Schmedding, I., Zank, S., Schairer, B., Fabretti, F., Hohne, M., Bartram, M.P., Dafinger, C., Hackl, M., Burst, V., Habbig, S., Zentgraf, H., Blaukat, A., Walz, G., Benzing, T. and Schermer, B., 2011. Nephrocystin-4 regulates Pyk2-induced tyrosine phosphorylation of nephrocystin-1 to control targeting to monocilia. *The Journal of Biological Chemistry*, 286 (16), 14237-14245.
- Lim, S.S., 2013. Nuclear FAK: a new mode of gene regulation from cellular adhesions. *Molecules and Cells*, 36 (1), 1-6.
- Lin, H., Hsu, Y., Kachingwe, B., Hsu, C., Uang, Y. and Wang, L., 2014. Dose effect of thiazolidinedione on cancer risk in type 2 diabetes mellitus patients: a six-year population-based cohort study. *Journal of Clinical Pharmacy and Therapeutics*, 39 (4), 354-360.
- Lipinski, C.A., and Loftus, J.C., 2010. Targeting Pyk2 for therapeutic intervention. *Expert Opinion on Therapeutic Targets*, 14 (1), 95-108.
- Lipinski, C.A., Tran, N.L., Menashi, E., Rohl, C., Kloss, J., Bay, R.C., Berens, M.E. and Loftus, J.C., 2005. The tyrosine kinase pyk2 promotes migration and invasion of glioma cells. *Neoplasia*, 7 (5), 435-445.
- Liu, B., Fan, Z., Edgerton, S.M., Deng, X., Alimova, I.N., Lind, S.E. and Thor, A.D., 2009. Metformin induces unique biological and molecular responses in triple negative breast cancer cells. *Cell Cycle*, 8 (13), 2031-2040.
- Liu, S., Tu, Y., Chien, M. and Chien, K., 2012. Effect of antidiabetic agents added to metformin on glycaemic control, hypoglycaemia and weight change in patients with type 2 diabetes: a network meta-analysis. *Diabetes, Obesity and Metabolism*, 14 (9), 810-820.
- Liu, Y., Liu, N. and Zhao, H., 2005. Inferring protein–protein interactions through high-throughput interaction data from diverse organisms. *Bioinformatics*, 21 (15), 3279-3285.
- Liu, H., Scholz, C., Zang, C., Schefe, J.H., Habbel, P., Regierer, A.C., Schulz, C.O., Possinger, K. and Eucker, J., 2012. Metformin and the mTOR inhibitor everolimus (RAD001) sensitize breast cancer cells to the cytotoxic effect of chemotherapeutic drugs in vitro. *Anticancer Research*, 32 (5), 1627-1637.
- Livingstone, D.J., 2008. *Artificial neural networks: Methods and applications (methods in molecular biology)*. Humana Press.
- Lord, S.R., et al., 2018. Integrated pharmacodynamic analysis identifies two metabolic adaption pathways to metformin in breast cancer. *Cell Metabolism*, 28 (5), 679-688. e4.
- Luker, K.E., Pica, C.M., Schreiber, R.D. and Piwnica-Worms, D., 2001. Overexpression of IRF9 confers resistance to antimicrotubule agents in breast cancer cells. *Cancer Research*, 61 (17), 6540-6547.

Lyn-Cook, B.D., Osborne, T., Joseph, S., Word, B., Pang, L. and Hammons, G., 2015. Metformin Effects on ABCB1 Expression and Proliferation in Pancreatic Cancer Cell Lines with Different ABCB1 genotypes/haplotypes.

Macias, H., and Hinck, L., 2012. Mammary gland development. *Wiley Interdisciplinary Reviews: Developmental Biology*, 1 (4), 533-557.

Makhoul, I., 2018. Therapeutic strategies for breast cancer. In: *Therapeutic strategies for breast cancer. The Breast (Fifth Edition)*. Elsevier, 315-330. e7.

Makhoul, I., Atiq, M., Alwbari, A. and Kieber-Emmons, T., 2018. Breast Cancer Immunotherapy: An Update. *Breast Cancer: Basic and Clinical Research*, 12, 1178223418774802.

Mandiyan, V., Andreev, J., Schlessinger, J. and Hubbard, S.R., 1999. Crystal structure of the ARF-GAP domain and ankyrin repeats of PYK2-associated protein beta. *The EMBO Journal*, 18 (24), 6890-6898.

Marinello, P.C., da Silva, Thamara Nishida Xavier, Panis, C., Neves, A.F., Machado, K.L., Borges, F.H., Guarnier, F.A., Bernardes, S.S., de-Freitas-Junior, J.C.M. and Morgado-Díaz, J.A., 2016. Mechanism of metformin action in MCF-7 and MDA-MB-231 human breast cancer cells involves oxidative stress generation, DNA damage, and transforming growth factor β 1 induction. *Tumor Biology*, 37 (4), 5337-5346.

Martinez-Serra, J., Gutierrez, A., Munoz-Capo, S., Navarro-Palou, M., Ros, T., Amat, J.C., Lopez, B., Marcus, T.F., Fueyo, L., Suquia, A.G., Gines, J., Rubio, F., Ramos, R. and Besalduch, J., 2014. xCELLigence system for real-time label-free monitoring of growth and viability of cell lines from hematological malignancies. *OncoTargets and Therapy*, 7, 985-994.

Maughan, K.L., Lutterbie, M.A. and Ham, P.S., 2010. Treatment of breast cancer. *Chemotherapy*, 51, 53.

McCubrey, J.A., Steelman, L.S., Abrams, S.L., Lee, J.T., Chang, F., Bertrand, F.E., Navolanic, P.M., Terrian, D.M., Franklin, R.A. and D'Assoro, A.B., 2006. Roles of the RAF/MEK/ERK and PI3K/PTEN/AKT pathways in malignant transformation and drug resistance. *Advances in Enzyme Regulation*, 46 (1), 249-279.

McShan, G.D., Zagozdzon, R., Park, S., Zrihan-Licht, S., Fu, Y., Avraham, S. and Avraham, H., 2002. Csk homologous kinase associates with RAFTK/Pyk2 in breast cancer cells and negatively regulates its activation and breast cancer cell migration. *International Journal of Oncology*, 21 (1), 197-205.

Meyer zu Heringdorf, D., Liliom, K., Schaefer, M., Danneberg, K., Jaggar, J.H., Tigyi, G. and Jakobs, K.H., 2003. Photolysis of intracellular caged sphingosine-1-phosphate causes Ca²⁺ mobilization independently of G-protein-coupled receptors. *FEBS Letters*, 554 (3), 443-449.

Meyers, V.E., Zayzafoon, M., Gonda, S.R., Gathings, W.E. and McDonald, J.M., 2004. Modeled microgravity disrupts collagen I/integrin signaling during osteoblastic differentiation of human mesenchymal stem cells. *Journal of Cellular Biochemistry*, 93 (4), 697-707.

Mi, H., Muruganujan, A., Casagrande, J.T. and Thomas, P.D., 2013. Large-scale gene function analysis with the PANTHER classification system. *Nature Protocols*, 8 (8), 1551.

Mimura, K., Teh, J.L., Okayama, H., Shiraishi, K., Kua, L., Koh, V., Smoot, D.T., Ashktorab, H., Oike, T. and Suzuki, Y., 2018. PD-L1 expression is mainly regulated by interferon gamma associated with JAK-STAT pathway in gastric cancer. *Cancer Science*, 109 (1), 43-53.

Morales, D.R., and Morris, A.D., 2015. Metformin in cancer treatment and prevention. *Annual Review of Medicine*, 66, 17-29.

Moretti, J., and Brou, C., 2013. Ubiquitinations in the notch signaling pathway. *International Journal of Molecular Sciences*, 14 (3), 6359-6381.

Mosmann, T., 1983. Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays. *Journal of Immunological Methods*, 65 (1-2), 55-63.

Murphy, J.M., Park, H. and Lim, S.S., 2016. FAK and Pyk2 in disease. *Frontiers in Biology*, 11 (1), 1-9.

Nangia-Makker, P., Yu, Y., Vasudevan, A., Farhana, L., Rajendra, S.G., Levi, E. and Majumdar, A.P., 2014. Metformin: a potential therapeutic agent for recurrent colon cancer. *PloS One*, 9 (1), e84369.

Naor, Z., Benard, O. and Seger, R., 2000. Activation of MAPK cascades by G-protein-coupled receptors: the case of gonadotropin-releasing hormone receptor. *Trends in Endocrinology & Metabolism*, 11 (3), 91-99.

Naser, R., Aldehaiman, A., Díaz-Galicia, E. and Arold, S., 2018. Endogenous Control Mechanisms of FAK and PYK2 and Their Relevance to Cancer Development. *Cancers*, 10 (6), 196.

Ni, B., Farrar, J.S., Chen, S., Lownik, J.C. and Celi, F.S., 2018. A novel role for PTK2B in cultured beige adipocyte differentiation. *Biochemical and Biophysical Research Communications*.

Noto, H., Goto, A., Tsujimoto, T. and Noda, M., 2012. Cancer risk in diabetic patients treated with metformin: a systematic review and meta-analysis. *PloS One*, 7 (3), e33411.

Oh, S.H., Jin, Q., Kim, E.S., Khuri, F.R. and Lee, H.Y., 2008. Insulin-like growth factor-I receptor signaling pathway induces resistance to the apoptotic activities of SCH66336 (lonafarnib) through Akt/mammalian target of rapamycin-mediated increases in survivin expression. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, 14 (5), 1581-1589.

Okigaki, M., Davis, C., Falasca, M., Harroch, S., Felsenfeld, D.P., Sheetz, M.P. and Schlessinger, J., 2003. Pyk2 regulates multiple signaling events crucial for macrophage morphology and migration. *Proceedings of the National Academy of Sciences of the United States of America*, 100 (19), 10740-10745.

Pandiri, I., Chen, Y., Joe, Y., Kim, H.J., Park, J., Chung, H.T. and Park, J.W., 2016. Tristetraprolin mediates the anti-proliferative effects of metformin in breast cancer cells. *Breast Cancer Research and Treatment*, 156 (1), 57-64.

Park, S.Y., Avraham, H.K. and Avraham, S., 2004. RAFTK/Pyk2 activation is mediated by trans-acting autophosphorylation in a Src-independent manner. *The Journal of Biological Chemistry*, 279 (32), 33315-33322.

- Pernicova, I., and Korbonits, M., 2014. Metformin—mode of action and clinical implications for diabetes and cancer. *Nature Reviews Endocrinology*, 10 (3), 143.
- Perou, C.M., Sørlie, T., Eisen, M.B., Van De Rijn, M., Jeffrey, S.S., Rees, C.A., Pollack, J.R., Ross, D.T., Johnsen, H. and Akslen, L.A., 2000. Molecular portraits of human breast tumours. *Nature*, 406 (6797), 747.
- Phoenix, K.N., Vumbaca, F. and Claffey, K.P., 2009. Therapeutic metformin/AMPK activation promotes the angiogenic phenotype in the ER α negative MDA-MB-435 breast cancer model. *Breast Cancer Research and Treatment*, 113 (1), 101-111.
- Phoenix, K.N., Vumbaca, F., Fox, M.M., Evans, R. and Claffey, K.P., 2010. Dietary energy availability affects primary and metastatic breast cancer and metformin efficacy. *Breast Cancer Research and Treatment*, 123 (2), 333-344.
- Pinto, D.M.S., and Flaus, A., 2010. Structure and function of histone H2AX. In: *Structure and function of histone H2AX. Genome Stability and Human Diseases*. Springer, 55-78.
- Posritong, S., Hong, J.M., Eleniste, P.P., McIntyre, P.W., Wu, J.L., Himes, E.R., Patel, V., Kacena, M.A. and Bruzzaniti, A., 2018. Pyk2 deficiency potentiates osteoblast differentiation and mineralizing activity in response to estrogen or raloxifene. *Molecular and Cellular Endocrinology*.
- Powe, D.G., Dhondalay, G.K.R., Lemetre, C., Allen, T., Habashy, H.O., Ellis, I.O., Rees, R. and Ball, G.R., 2014. DACH1: its role as a classifier of long-term good prognosis in luminal breast cancer. *PLoS One*, 9 (1), e84428.
- Prat, A., Parker, J.S., Karginova, O., Fan, C., Livasy, C., Herschkowitz, J.I., He, X. and Perou, C.M., 2010. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Research*, 12 (5), R68.
- Prat, A., and Perou, C.M., 2011. Deconstructing the molecular portraits of breast cancer. *Molecular Oncology*, 5 (1), 5-23.
- Prat, A., Pineda, E., Adamo, B., Galván, P., Fernández, A., Gaba, L., Díez, M., Viladot, M., Arance, A. and Muñoz, M., 2015. Clinical implications of the intrinsic molecular subtypes of breast cancer. *The Breast*, 24, S26-S35.
- Previati, M., Manfrini, M., Galasso, M., Zerbinati, C., Palatini, J., Gasparini, P. and Volinia, S., 2013. Next generation analysis of breast cancer genomes for precision medicine. *Cancer Letters*, 339 (1), 1-7.
- Provinciali, N., Lazzeroni, M., Cazzaniga, M., Gorlero, F., Dunn, B.K. and DeCensi, A., 2015. Metformin: risk-benefit profile with a focus on cancer. *Expert Opinion on Drug Safety*, 14 (10), 1573-1585.
- Rice, S., Pellat, L., Ahmetaga, A., Bano, G., Mason, H. and Whitehead, S., 2015. Dual effect of metformin on growth inhibition and oestradiol production in breast cancer cells. *International Journal of Molecular Medicine*, 35 (4), 1088-1094.
- Rizos, C.V., and Elisaf, M.S., 2013. Metformin and cancer. *European Journal of Pharmacology*, 705 (1-3), 96-108.

Roberts, W.G., Ung, E., Whalen, P., Cooper, B., Hulford, C., Autry, C., Richter, D., Emerson, E., Lin, J., Kath, J., Coleman, K., Yao, L., Martinez-Alsina, L., Lorenzen, M., Berliner, M., Luzzio, M., Patel, N., Schmitt, E., LaGreca, S., Jani, J., Wessel, M., Marr, E., Griffor, M. and Vajdos, F., 2008. Antitumor activity and pharmacology of a selective focal adhesion kinase inhibitor, PF-562,271. *Cancer Research*, 68 (6), 1935-1944.

Ross, J.S., Fletcher, J.A., Linette, G.P., Stec, J., Clark, E., Ayers, M., Symmans, W.F., Puztai, L. and Bloom, K.J., 2003. The Her-2/neu gene and protein in breast cancer 2003: biomarker and target of therapy. *The Oncologist*, 8 (4), 307-325.

Rufanova, V.A., Alexanian, A., Wakatsuki, T., Lerner, A. and Sorokin, A., 2009. Pyk2 mediates endothelin-1 signaling via p130Cas/BCAR3 cascade and regulates human glomerular mesangial cell adhesion and spreading. *Journal of Cellular Physiology*, 219 (1), 45-56.

Ruusala, A., and Aspenstrom, P., 2008. The atypical Rho GTPase Wrch1 collaborates with the nonreceptor tyrosine kinases Pyk2 and Src in regulating cytoskeletal dynamics. *Molecular and Cellular Biology*, 28 (5), 1802-1814.

Saha, A., Blando, J., Tremmel, L. and DiGiovanni, J., 2015. Effect of Metformin, Rapamycin, and Their Combination on Growth and Progression of Prostate Tumors in HiMyc Mice. *Cancer Prevention Research (Philadelphia, Pa.)*, 8 (7), 597-606.

Saunus, J.M., Smart, C.E., Kutasovic, J.R., Johnston, R.L., Kalita-de Croft, P., Miranda, M., Rozali, E.N., Vargas, A.C., Reid, L.E. and Lorsy, E., 2018. Multidimensional phenotyping of breast cancer cell lines to guide preclinical research. *Breast Cancer Research and Treatment*, 167 (1), 289-301.

Schaller, M.D., 2010. Cellular functions of FAK kinases: insight into molecular mechanisms and novel functions. *Journal of Cell Science*, 123 (Pt 7), 1007-1013.

Schmittgen, T.D., and Livak, K.J., 2008. Analyzing real-time PCR data by the comparative C T method. *Nature Protocols*, 3 (6), 1101.

Seguin, L., Desgrosellier, J.S., Weis, S.M. and Cheresch, D.A., 2015. Integrins and cancer: regulators of cancer stemness, metastasis, and drug resistance. *Trends in Cell Biology*, 25 (4), 234-240.

Selitrennik, M., and Lev, S., 2015. PYK2 integrates growth factor and cytokine receptors signaling and potentiates breast cancer invasion via a positive feedback loop. *Oncotarget*, 6 (26), 22214-22226.

Shao, C., Ahmad, N., Hodges, K., Kuang, S., Ratliff, T. and Liu, X., 2015. Inhibition of polo-like kinase 1 (Plk1) enhances the antineoplastic activity of metformin in prostate cancer. *The Journal of Biological Chemistry*, 290 (4), 2024-2033.

Sherr, C.J., 2004. Principles of tumor suppression. *Cell*, 116 (2), 235-246.

Shi, C.S., and Kehrl, J.H., 2004. Pyk2 amplifies epidermal growth factor and c-Src-induced Stat3 activation. *The Journal of Biological Chemistry*, 279 (17), 17224-17231.

Sikka, A., Kaur, M., Agarwal, C., Deep, G. and Agarwal, R., 2012. Metformin suppresses growth of human head and neck squamous cell carcinoma via global inhibition of protein translation. *Cell Cycle*, 11 (7), 1374-1382.

- Silverman, E.K., and Loscalzo, J., 2012. Network medicine approaches to the genetics of complex diseases. *Discovery Medicine*, 14 (75), 143-152.
- Simpson, A., Petnga, W., Macaulay, V.M., Weyer-Czernilofsky, U. and Bogenrieder, T., 2017. Insulin-like growth factor (IGF) pathway targeting in cancer: Role of the IGF Axis and opportunities for future combination studies. *Targeted Oncology*, 12 (5), 571-597.
- Smith, J.P., Fonkoua, L.K. and Moody, T.W., 2016. The Role of Gastrin and CCK Receptors in Pancreatic Cancer and other Malignancies. *International Journal of Biological Sciences*, 12 (3), 283-291.
- Smoot, M.E., Ono, K., Ruscheinski, J., Wang, P. and Ideker, T., 2010. Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics*, 27 (3), 431-432.
- Snima, K., Nair, R.S., Nair, S.V., Kamath, C.R. and Lakshmanan, V., 2015. Combination of anti-diabetic drug metformin and boswellic acid nanoparticles: a novel strategy for pancreatic cancer therapy. *Journal of Biomedical Nanotechnology*, 11 (1), 93-104.
- Song, C.W., Lee, H., Dings, R.P., Williams, B., Powers, J., Dos Santos, T., Choi, B. and Park, H.J., 2012. Metformin kills and radiosensitizes cancer cells and preferentially kills cancer stem cells. *Scientific Reports*, 2, 362.
- Song, L., and Schindler, C., 2009. JAK-STAT Signaling. In: *JAK-STAT Signaling. Handbook of Cell Signaling (Second Edition)*. Elsevier, 2041-2048.
- Sonnenblick, A., Agbor-Tarh, D., Bradbury, I., Di Cosimo, S., Azim, H.A., Jr, Fumagalli, D., Sarp, S., Wolff, A.C., Andersson, M., Kroep, J., Cufer, T., Simon, S.D., Salman, P., Toi, M., Harris, L., Gralow, J., Keane, M., Moreno-Aspitia, A., Piccart-Gebhart, M. and de Azambuja, E., 2017. Impact of Diabetes, Insulin, and Metformin Use on the Outcome of Patients with Human Epidermal Growth Factor Receptor 2-Positive Primary Breast Cancer: Analysis from the ALTTO Phase III Randomized Trial. *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, 35 (13), 1421-1429.
- Soo, J.S., Ng, C., Tan, S.H., Malik, R.A., Teh, Y., Tan, B., Ho, G., See, M., Taib, N.A.M. and Yip, C., 2015. Metformin synergizes 5-fluorouracil, epirubicin, and cyclophosphamide (FEC) combination therapy through impairing intracellular ATP production and DNA repair in breast cancer stem cells. *Apoptosis*, 20 (10), 1373-1387.
- Soranna, D., Scotti, L., Zambon, A., Bosetti, C., Grassi, G., Catapano, A., La Vecchia, C., Mancina, G. and Corrao, G., 2012. Cancer risk associated with use of metformin and sulfonylurea in type 2 diabetes: a meta-analysis. *The Oncologist*, 17 (6), 813-822.
- Sorlie, T., 2016. The Impact of Gene Expression Patterns in Breast Cancer. *Clinical Chemistry*, 62 (8), 1150-1151.
- Sośnicki, S., Kapral, M. and Węglarz, L., 2016. Molecular targets of metformin antitumor action. *Pharmacological Reports*, 68 (5), 918-925.
- Sticca, R.P. and Murphy, T.C., 2014. Effects of DCA and Metformin on Murine Colon Cancer Growth.

Stockert, J.C., Blázquez-Castro, A., Cañete, M., Horobin, R.W. and Villanueva, Á., 2012. MTT assay for cell viability: Intracellular localization of the formazan product is in lipid droplets. *Acta Histochemica*, 114 (8), 785-796.

Subramani, R., Nandy, S.B., Pedroza, D.A. and Lakshmanaswamy, R., 2017. Role of growth hormone in breast cancer. *Endocrinology*, 158 (6), 1543-1555.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S. and Mesirov, J.P., 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America*, 102 (43), 15545-15550.

Sun, C.K., Man, K., Ng, K.T., Ho, J.W., Lim, Z.X., Cheng, Q., Lo, C., Poon, R.T. and Fan, S., 2008. Proline-rich tyrosine kinase 2 (Pyk2) promotes proliferation and invasiveness of hepatocellular carcinoma cells through c-Src/ERK activation. *Carcinogenesis*, 29 (11), 2096-2105.

Sun, C.K., Ng, K.T., Lim, Z.X., Cheng, Q., Lo, C.M., Poon, R.T., Man, K., Wong, N. and Fan, S.T., 2011. Proline-rich tyrosine kinase 2 (Pyk2) promotes cell motility of hepatocellular carcinoma through induction of epithelial to mesenchymal transition. *PLoS One*, 6 (4), e18878.

Sun, C., Ng, K., Sun, B., Ho, J., Lee, T., Ng, I., Poon, R., Lo, C., Liu, C. and Man, K., 2007. The significance of proline-rich tyrosine kinase2 (Pyk2) on hepatocellular carcinoma progression and recurrence. *British Journal of Cancer*, 97 (1), 50.

Suprunenko, T., and Hofer, M.J., 2016. The emerging role of interferon regulatory factor 9 in the antiviral host response and beyond. *Cytokine & Growth Factor Reviews*, 29, 35-43.

Szabo, R., and Bugge, T.H., 2008. Type II transmembrane serine proteases in development and disease. *The International Journal of Biochemistry & Cell Biology*, 40 (6-7), 1297-1316.

Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., Simonovic, M., Roth, A., Santos, A. and Tsafou, K.P., 2014. STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Research*, 43 (D1), D447-D452.

Taherian-Fard, A., Srihari, S. and Ragan, M.A., 2014. Breast cancer classification: linking molecular mechanisms to disease prognosis. *Briefings in Bioinformatics*, 16 (3), 461-474.

Takahashi, S.X., Miriyala, J., Tay, L.H., Yue, D.T. and Colecraft, H.M., 2005. A CaVbeta SH3/guanylate kinase domain interaction regulates multiple properties of voltage-gated Ca²⁺ channels. *The Journal of General Physiology*, 126 (4), 365-377.

Takaoka, A., Tanaka, N., Mitani, Y., Miyazaki, T., Fujii, H., Sato, M., Kovarik, P., Decker, T., Schlessinger, J. and Taniguchi, T., 1999. Protein tyrosine kinase Pyk2 mediates the Jak-dependent activation of MAPK and Stat1 in IFN-gamma, but not IFN-alpha, signaling. *The EMBO Journal*, 18 (9), 2480-2488.

Talarico, G., Orecchioni, S., Dallaglio, K., Reggiani, F., Mancuso, P., Calleri, A., Gregato, G., Labanca, V., Rossi, T. and Noonan, D.M., 2016. Aspirin and atenolol enhance metformin activity against breast cancer by targeting both neoplastic and microenvironment cells. *Scientific Reports*, 6, 18673.

Talarico, G., Reggiani, F., Orecchioni, S., Mancuso, P., Calleri, A., Gregato, G., Labanca, V., Noonan, D.M., Dallaglio, K. and Albini, A., 2015. Synergistic Activity of Aspirin, Atenolol and Metformin in the Inhibition of Angiogenesis, Local and Metastatic Growth of Breast Cancer by Targeting both Neoplastic and Microenvironment Cells.

Tanaka, R., Tomosugi, M., Horinaka, M., Sowa, Y. and Sakai, T., 2015. Metformin causes G1-phase arrest via down-regulation of MiR-221 and enhances TRAIL sensitivity through DR5 up-regulation in pancreatic cancer cells. *PLoS One*, 10 (5), e0125779.

Tang, G.H., Satkunam, M., Pond, G.R., Steinberg, G.R., Blandino, G., Schunemann, H.J. and Muti, P., 2018. Association of Metformin with Breast Cancer Incidence and Mortality in Patients with Type II Diabetes: A GRADE-Assessed Systematic Review and Meta-analysis. *Cancer Epidemiology, Biomarkers & Prevention: A Publication of the American Association for Cancer Research, Cosponsored by the American Society of Preventive Oncology*, 27 (6), 627-635.

Taniyama, Y., Weber, D.S., Rocic, P., Hilenski, L., Akers, M.L., Park, J., Hemmings, B.A., Alexander, R.W. and Griendling, K.K., 2003. Pyk2- and Src-dependent tyrosine phosphorylation of PDK1 regulates focal adhesions. *Molecular and Cellular Biology*, 23 (22), 8019-8029.

Tavora, B., Reynolds, L.E., Batista, S., Demircioglu, F., Fernandez, I., Lechertier, T., Lees, D.M., Wong, P., Alexopoulou, A. and Elia, G., 2014. Endothelial-cell FAK targeting sensitizes tumours to DNA-damaging therapy. *Nature*, 514 (7520), 112.

Thiery, J.P., and Sleeman, J.P., 2006. Complex networks orchestrate epithelial–mesenchymal transitions. *Nature Reviews Molecular Cell Biology*, 7 (2), 131.

Trilla-Fuertes, L., et al., 2018. Molecular characterization of breast cancer cell response to metabolic drugs. *Oncotarget*, 9 (11), 9645.

Great Britain. About Type 2 diabetes, 2018. About Type 2 diabetes | Types of diabetes | Diabetes UK. Available at: <https://www.diabetes.org.uk/Type-2-diabetes/dcis.html> [Accessed September 2018].

Tokiwa, G., Dikic, I., Lev, S. and Schlessinger, J., 1996. Activation of Pyk2 by stress signals and coupling with JNK signaling pathway. *Science (New York, N.Y.)*, 273 (5276), 792-794.

Tong, D.L., Boocock, D.J., Dhondalay, G.K.R., Lemetre, C. and Ball, G.R., 2014. Artificial neural network inference (ANNI): a study on gene-gene interaction for biomarkers in childhood sarcomas. *PLoS One*, 9 (7), e102483.

Tripathi, S., Flobak, Å., Chawla, K., Baudot, A., Bruland, T., Thommesen, L., Kuiper, M. and Lægreid, A., 2015. The gastrin and cholecystokinin receptors mediated signaling network: a scaffold for data analysis and new hypotheses on regulatory mechanisms. *BMC Systems Biology*, 9 (1), 40.

Tseng, C.H., 2012. Diabetes, metformin use, and colon cancer: a population-based cohort study in Taiwan. *European Journal of Endocrinology*, 167 (3), 409-416.

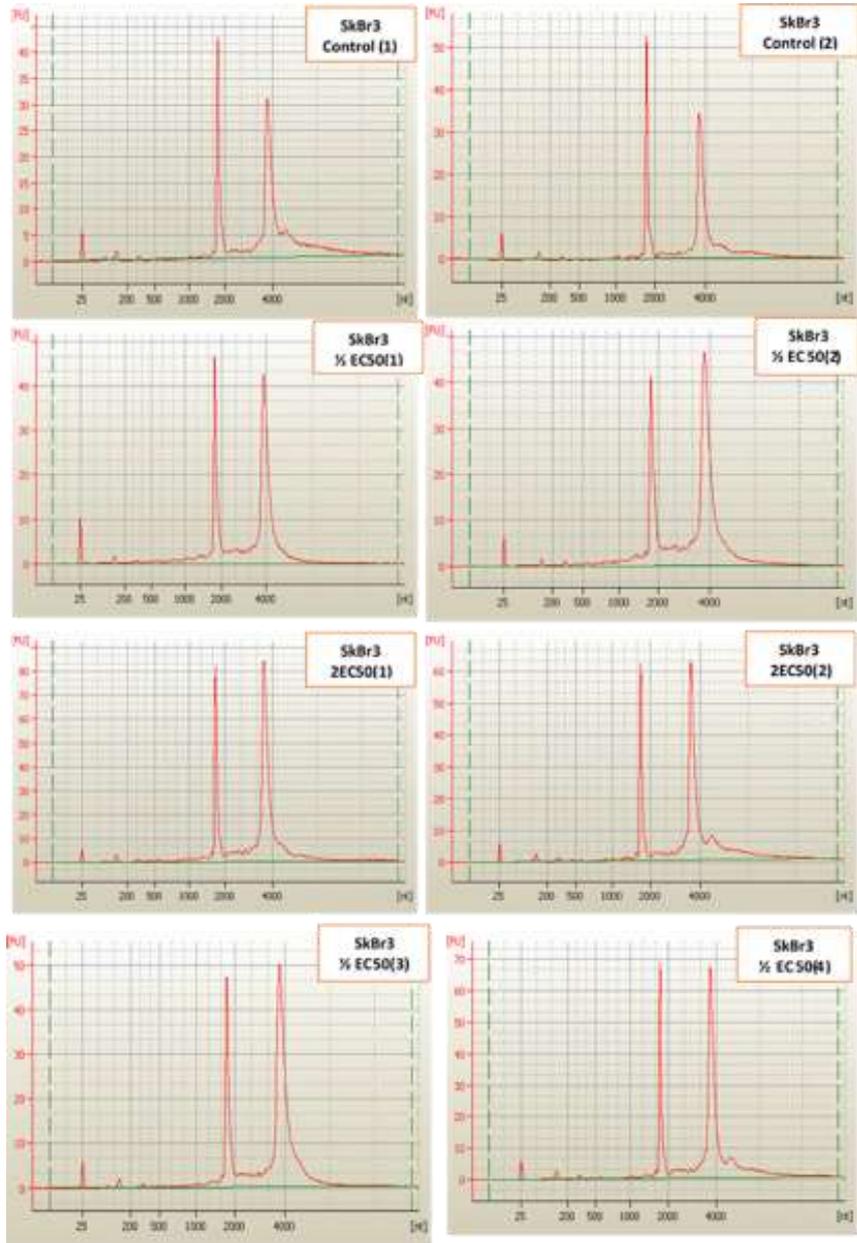
V Sekar, T., Dhanabalan, A. and Paulmurugan, R., 2011. Imaging cellular receptors in breast cancers: an overview. *Current Pharmaceutical Biotechnology*, 12 (4), 508-527.

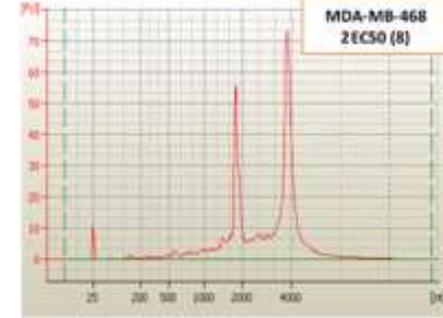
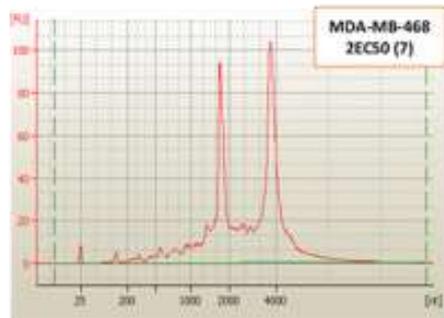
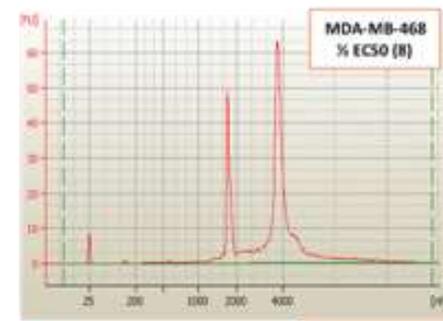
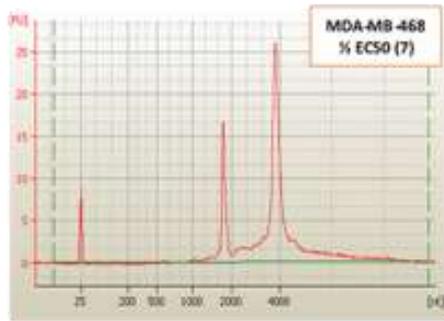
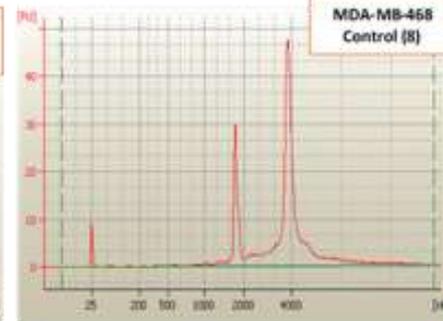
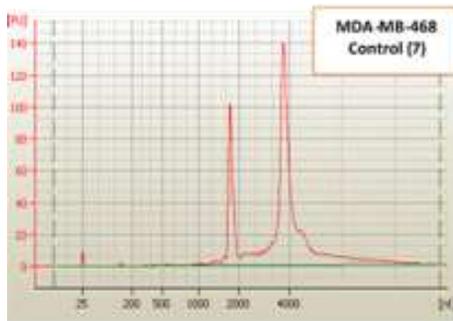
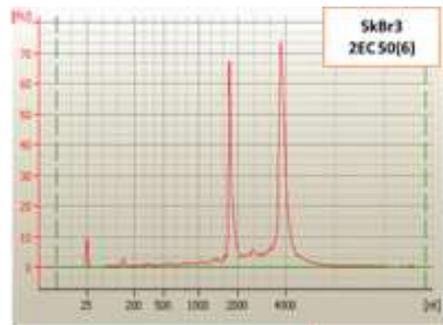
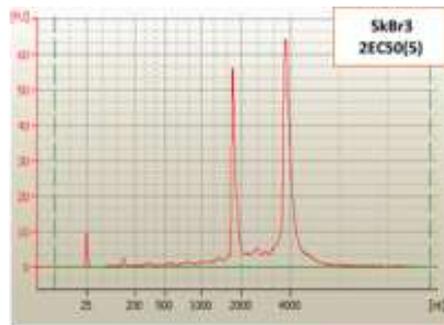
Vallianou, N.G., Evangelopoulos, A. and Kazazis, C., 2013. Metformin and cancer. *The Review of Diabetic Studies: RDS*, 10 (4), 228-235.

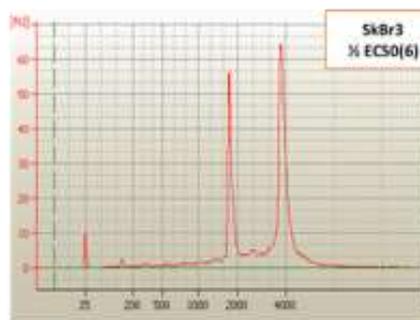
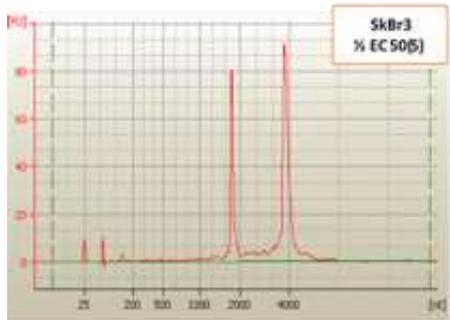
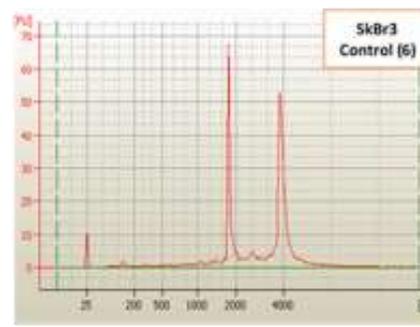
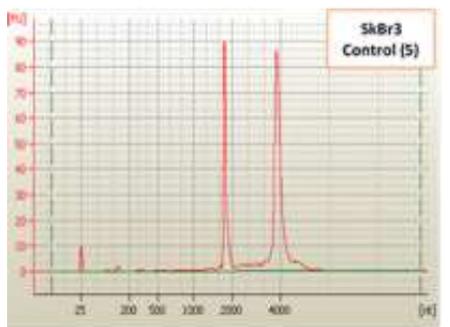
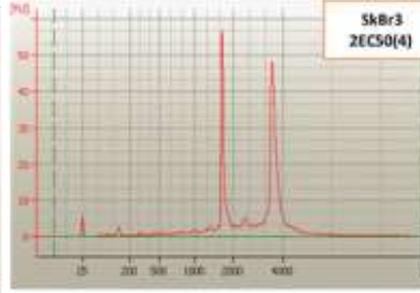
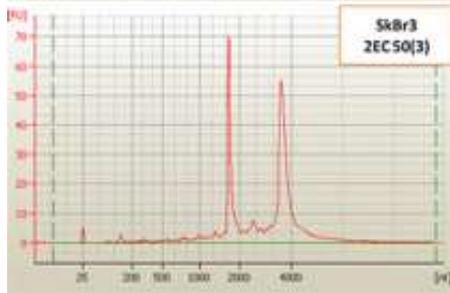
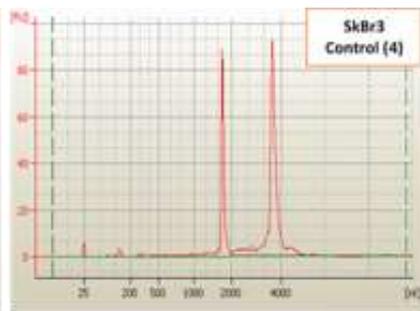
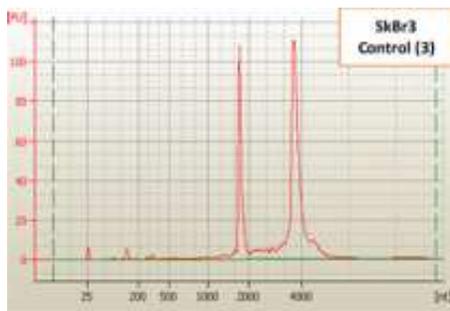
- van Meerloo, J., Kaspers, G.J. and Cloos, J., 2011. Cell sensitivity assays: the MTT assay. In: Cell sensitivity assays: the MTT assay. Cancer cell culture. Springer, 237-245.
- Vazquez-Martin, A., Oliveras-Ferraro, C., Del Barco, S., Martin-Castillo, B. and Menendez, J.A., 2011. The anti-diabetic drug metformin suppresses self-renewal and proliferation of trastuzumab-resistant tumor-initiating breast cancer stem cells. *Breast Cancer Research and Treatment*, 126 (2), 355-364.
- Vicente-Manzanares, M., Webb, D.J. and Horwitz, A.R., 2005. Cell migration at a glance. *Journal of Cell Science*, 118 (Pt 21), 4917-4919.
- Viollet, B., Guigas, B., Sanz Garcia, N., Leclerc, J., Foretz, M. and Andreelli, F., 2012. Cellular and molecular mechanisms of metformin: an overview. *Clinical Science (London, England: 1979)*, 122 (6), 253-270.
- Vogelstein, B., Papadopoulos, N., Velculescu, V.E., Zhou, S., Diaz, L.A., Jr and Kinzler, K.W., 2013. Cancer genome landscapes. *Science (New York, N.Y.)*, 339 (6127), 1546-1558.
- Vuong, D., Simpson, P.T., Green, B., Cummings, M.C. and Lakhani, S.R., 2014. Molecular classification of breast cancer. *Virchows Archiv*, 465 (1), 1-14.
- Waks, Z., Weissbrod, O., Carmeli, B., Norel, R., Utro, F. and Goldschmidt, Y., 2016. Driver gene classification reveals a substantial overrepresentation of tumor suppressors among very large chromatin-regulating proteins. *Scientific Reports*, 6, 38988.
- Wang, G.G., Allis, C.D. and Chi, P., 2007. Chromatin remodeling and cancer, Part I: Covalent histone modifications. *Trends in Molecular Medicine*, 13 (9), 363-372.
- Wang, G.G., Allis, C.D. and Chi, P., 2007. Chromatin remodeling and cancer, Part II: ATP-dependent chromatin remodeling. *Trends in Molecular Medicine*, 13 (9), 373-380.
- Wang, Q., Hardie, R., Hoy, A.J., van Geldermalsen, M., Gao, D., Fazli, L., Sadowski, M.C., Balaban, S., Schreuder, M. and Nagarajah, R., 2015. Targeting ASCT2-mediated glutamine uptake blocks prostate cancer growth and tumour development. *The Journal of Pathology*, 236 (3), 278-289.
- Wang, Y., Wei, J., Li, L., Fan, C. and Sun, Y., 2015. Combined use of metformin and everolimus is synergistic in the treatment of breast cancer cells. *Oncology Research Featuring Preclinical and Clinical Cancer Therapeutics*, 22 (4), 193-201.
- Weigelt, B., Mackay, A., A'hern, R., Natrajan, R., Tan, D.S., Dowsett, M., Ashworth, A. and Reis-Filho, J.S., 2010. Breast cancer molecular profiling with single sample predictors: a retrospective analysis. *The Lancet Oncology*, 11 (4), 339-349.
- Werner, H., and Bruchim, I., 2012. IGF-1 and BRCA1 signalling pathways in familial cancer. *The Lancet Oncology*, 13 (12), e537-e544.
- World Health Organization, 2018. World Health Organization: Cancer - who.int. Available at: <https://www.who.int/cancer/en/dcis.html> [Accessed September 2018].
- Wurth, R., Barbieri, F. and Florio, T., 2014. New molecules and old drugs as emerging approaches to selectively target human glioblastoma cancer stem cells. *BioMed Research International*, 2014, 126586.

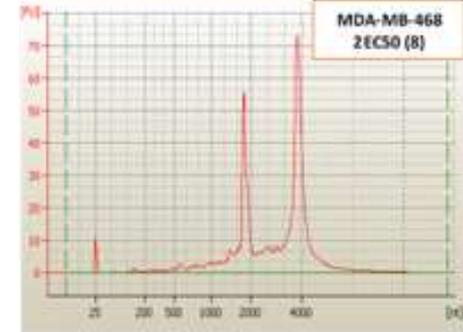
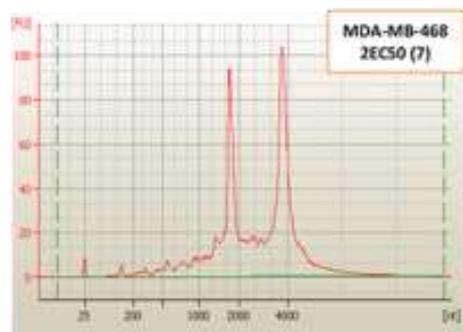
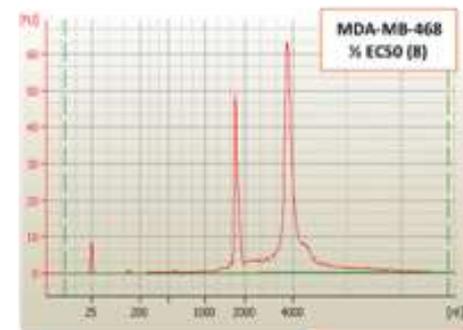
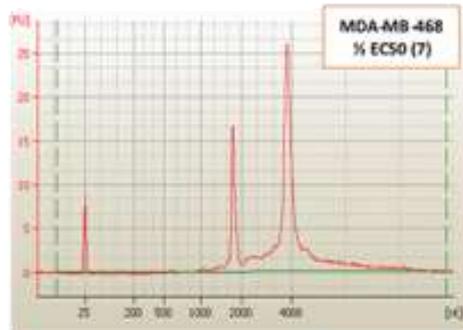
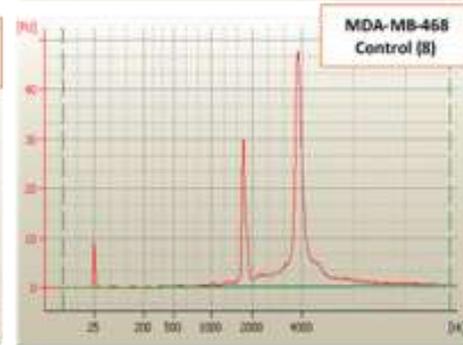
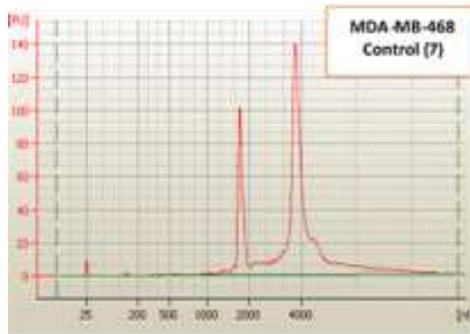
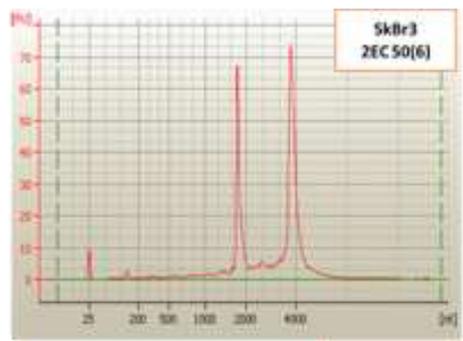
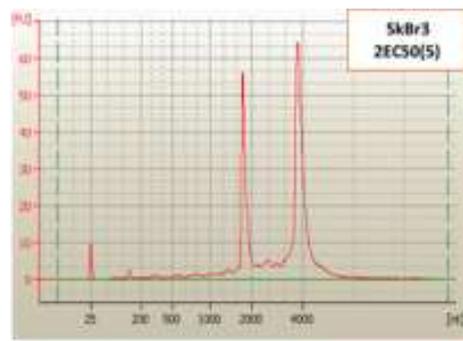
- Xu, J., Gao, X., Ramchandran, R., Zhao, Y., Vogel, S.M. and Malik, A.B., 2008. Nonmuscle myosin light-chain kinase mediates neutrophil transmigration in sepsis-induced lung inflammation by activating β 2 integrins. *Nature Immunology*, 9 (8), 880.
- Yager, J.D., and Davidson, N.E., 2006. Estrogen carcinogenesis in breast cancer. *New England Journal of Medicine*, 354 (3), 270-282.
- Yamaguchi, H., Wyckoff, J. and Condeelis, J., 2005. Cell migration in tumors. *Current Opinion in Cell Biology*, 17 (5), 559-564.
- Yanagawa, M., Ikemot, K., Kawauchi, S., Furuya, T., Yamamoto, S., Oka, M., Oga, A., Nagashima, Y. and Sasaki, K., 2012. Luminal A and luminal B (HER2 negative) subtypes of breast cancer consist of a mixture of tumors with different genotype. *BMC Research Notes*, 5 (1), 376.
- Yang, T., Yang, Y. and Liu, S., 2015. Association between metformin therapy and breast cancer incidence and mortality: evidence from a meta-analysis. *Journal of Breast Cancer*, 18 (3), 264-270.
- Yersal, O., and Barutca, S., 2014. Biological subtypes of breast cancer: Prognostic and therapeutic implications. *World Journal of Clinical Oncology*, 5 (3), 412-424.
- Zakikhani, M., Dowling, R., Fantus, I.G., Sonenberg, N. and Pollak, M., 2006. Metformin is an AMP kinase-dependent growth inhibitor for breast cancer cells. *Cancer Research*, 66 (21), 10269-10273.
- Zhan, T., Rindtorff, N. and Boutros, M., 2017. Wnt signaling in cancer. *Oncogene*, 36 (11), 1461.
- Zhang, M., Heldin, A., Palomar-Siles, M., Öhlin, S., Bykov, V.J. and Wiman, K.G., 2018. synergistic rescue of nonsense Mutant Tumor suppressor p53 by combination Treatment with aminoglycosides and Mdm2 inhibitors. *Frontiers in Oncology*, 7, 323.
- Zhang, Y., Peng, G. and Hsueh, E.C., 2014. Induction of Autophagy and Apoptosis with Polyamine Synthesis Inhibition and Metformin in Human Melanoma and Colon Cancer Cells.
- Zhao, X., Chu, Q., Cui, J., Huo, R. and Xu, T., 2017. IRF9 as a negative regulator involved in TRIF-mediated NF- κ B pathway in a teleost fish, *Miichthys miiuy*. *Molecular Immunology*, 85, 123-129.
- Zheng, W., Zhang, Q., Cai, D., Yang, X., Ungvari, G.S., Ng, C.H., Wu, R. and Xiang, Y., 2018. Combination of Metformin and Lifestyle Intervention for Antipsychotic-Related Weight Gain: A Meta-Analysis of Randomized Controlled Trials. *Pharmacopsychiatry*.
- Zou, Q., Wang, X., Liu, Y., Ouyang, Z., Long, H., Wei, S., Xin, J., Zhao, B., Lai, S. and Shen, J., 2015. Generation of gene-target dogs using CRISPR/Cas9 system. *Journal of Molecular Cell Biology*, 7 (6), 580-583.
- Zrihan-Licht, S., Fu, Y., Settleman, J., Schinkmann, K., Shaw, L., Keydar, I., Avraham, S. and Avraham, H., 2000. RAFTK/Pyk2 tyrosine kinase mediates the association of p190 RhoGAP with RasGAP and is involved in breast cancer cell invasion. *Oncogene*, 19 (10), 1318.
- Zrihan-Licht, S., Fu, Y., Settleman, J., Schinkmann, K., Shaw, L., Keydar, I., Avraham, S. and Avraham, H., 2000. RAFTK/Pyk2 tyrosine kinase mediates the association of p190 RhoGAP with RasGAP and is involved in breast cancer cell invasion. *Oncogene*, 19 (10), 1318.

APPENDIX









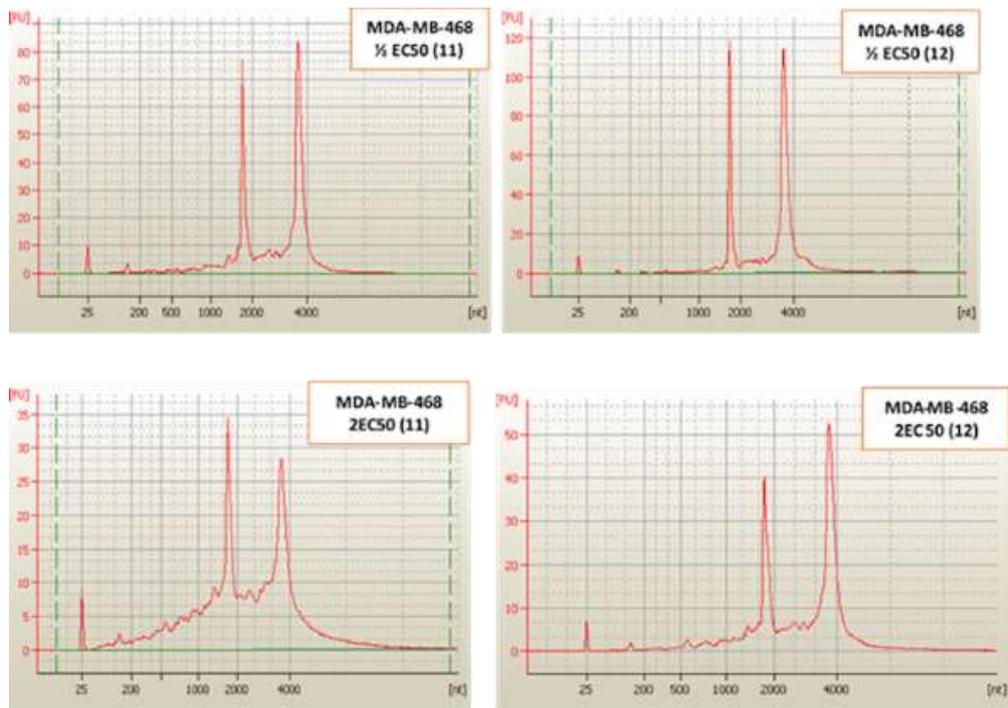
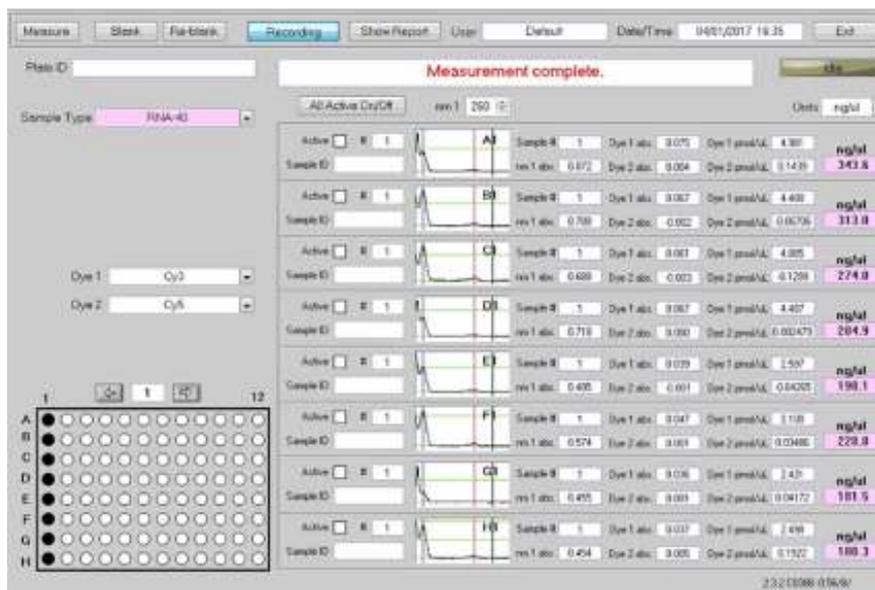


Figure A.4. 1. Electropherograms generated for RNA derived from SkBr3 and MDA-MB-468 cells in quadruplicate (Sample 1-36). Data refer to the Result section 4.3.1.



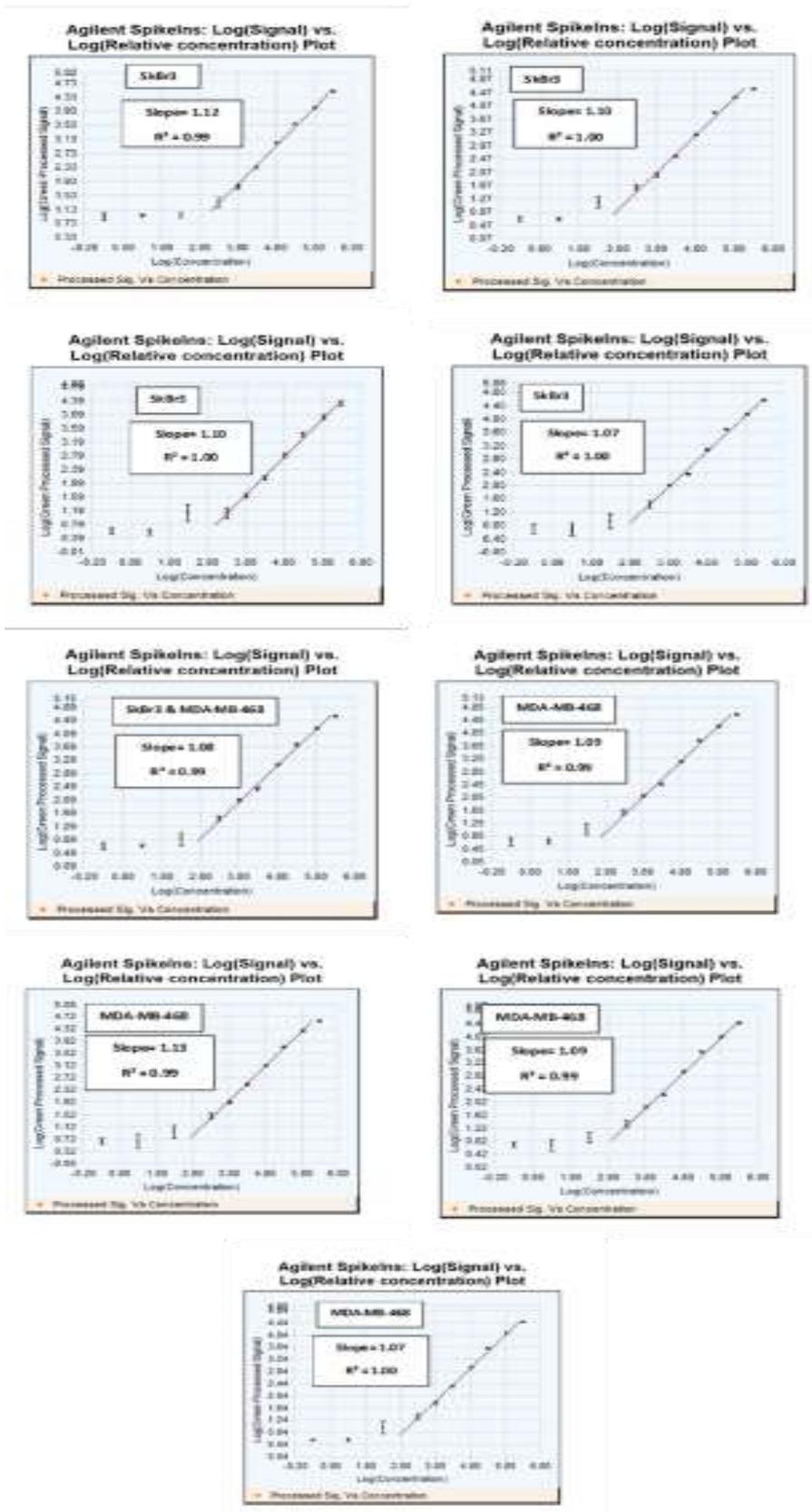


Figure A.4. 3. Quality Control (QC) Report from Data acquisition in GenePixPro microarray scanner – part 1. Plots show Agilent Spike-in Linearity check plots with Slope and R2 values for each sample analysed.

Table A. 4.1. List of the 60 common genes in BASAL-LIKE (MDA-MB-468) and HER2 (SkBr3) cell lines

Input ID	Gene symbol	Gene Name	Average of interaction
A_32_P69368	ID2	glutamate receptor, ionotropic, delta 2 (GRID2), transcript variant 1	-6.846
A_32_P25050	RDH10	lnc-RDH10-3:1 gb AK125786	-6.628
A_32_P210202	E2F7	E2F transcription factor 7 (E2F7)	-6.186
A_32_P174083	CYCS	cytochrome c, somatic (CYCS)	-6.164
A_32_P152767	SKIDA1	SKI/DACH domain containing 1 (SKIDA1)	-6.095
A_32_P105195	DDX46	DEAD (Asp-Glu-Ala-Asp) box polypeptide 46 (DDX46), transcript variant 2	-5.929
A_24_P942354	PITPNA	phosphatidylinositol transfer protein, alpha (PITPNA)	-5.926
A_24_P921933	SRSF1	serine/arginine-rich splicing factor 11 (SRSF11), transcript variant 1	-5.911
A_24_P66001	UQCR10	ubiquinol-cytochrome c reductase, complex III subunit X (UQCR10), transcript variant 1	-5.797
A_24_P405430	TIA1	TIA1 cytotoxic granule-associated RNA binding protein (TIA1), transcript variant 2	-5.794
A_24_P38895	H2AFX	H2A histone family, member X (H2AFX)	-5.751
A_24_P322847	POLR3H	polymerase (RNA) III (DNA directed) polypeptide H (22.9kD) (POLR3H), transcript variant 2	-5.679
A_24_P316305	AQR	progesterin and adipoQ receptor family member VII (PAQR7)	-5.647
A_24_P314571	SPC24	HSPC249 mRNA, complete cds.	-5.642
A_24_P217834	HIST1H3D	histone cluster 1, H3d (HIST1H3D)	-5.642
A_24_P141736	METAP2	methionyl aminopeptidase 2 (METAP2)	-5.430
A_23_P98248	TRPT1	tRNA phosphotransferase 1 (TRPT1), transcript variant 1	-5.428
A_23_P90533	POP4	processing of precursor 4, ribonuclease P/MRP subunit (S. cerevisiae) (POP4), transcript variant 1	-5.421
A_23_P8452	LFNG	LFNG O-fucosylpeptide 3-beta-N-acetylglucosaminyltransferase (LFNG), transcript variant 2	-5.417

A_23_P63789	ZWINT	ZW10 interacting kinetochore protein (ZWINT), transcript variant 2	-5.338
A_23_P53668	NFYB	nuclear transcription factor Y, beta (NFYB)	-5.336
A_23_P435941	SAMD1	ens ENST00000448179 ens ENST00000398216 linc lnc-SAMD11-1:1 linc lnc-SAMD11-1:2	-5.315
A_23_P434809	S100A8	S100 calcium binding protein A8 (S100A8)	-5.303
A_23_P421306	SYT12	synaptotagmin XII (SYT12), transcript variant 1	-5.303
A_23_P350045	REEP5	receptor accessory protein 5 (REEP5)	-5.254
A_23_P33154	STAU2	STAU2 antisense RNA 1 (STAU2-AS1), long non-coding	-5.238
A_23_P327069	KIAA0232	KIAA0232 (KIAA0232), transcript variant 1	-5.181
A_23_P310	MARCKSL1	MARCKS-like 1 (MARCKSL1), transcript variant 1	-5.160
A_23_P305977	GRAMD2	GRAM domain containing 2 (GRAMD2)	-5.153
A_23_P30495	HMGCR	3-hydroxy-3-methylglutaryl-CoA reductase (HMGCR), transcript variant 1	-5.145
A_23_P27215	UBB	Synthetic construct Homo sapiens gateway clone IMAGE:100019426 3' read TUBB	-5.117
A_23_P24723	TMEM138	transmembrane protein 138 (TMEM138), transcript variant 1	-5.103
A_23_P209200	CCNE1	cyclin E1 (CCNE1)	-5.097
A_23_P166716	TRMT10C	tRNA methyltransferase 10 homolog C (S. cerevisiae) (TRMT10C)	-5.078
A_23_P166526	RIBC2	RIB43A domain with coiled-coils 2 (RIBC2)	-5.032
A_23_P157715	PPP1R16A	protein phosphatase 1, regulatory subunit 16A (PPP1R16A)	-5.024
A_23_P156667	PPP1R10	protein phosphatase 1, regulatory subunit 10 (PPP1R10), transcript variant 1	-5.014
A_23_P150255	RBM14	RBM14-RBM4 readthrough (RBM14-RBM4), transcript variant 1	-5.001
A_23_P13554	ALG8	ALG8, alpha-1,3-glucosyltransferase (ALG8), transcript variant 2	-4.989
A_23_P128734	ERH	enhancer of rudimentary homolog (Drosophila) (ERH)	-4.980
A_23_P128147	TUBA1B	tubulin, alpha 1b (TUBA1B)	-4.960
A_23_P123974	DTYMK	lnc-DTYMK-3:1	-4.946
A_23_P116829	UBE2N	ubiquitin-conjugating enzyme E2N (UBE2N)	-4.944
A_23_P115375	HIST2H3D	ENST00000415338 linc lnc-HIST2H3D-1:1 linc lnc-FAM72B-2:1 linc lnc-HIST2H3PS2-1:1	-4.943
A_23_P106505	LCMT2	leucine carboxyl methyltransferase 2 (LCMT2)	-4.941

A_23_P103110	MAFF	v-maf avian musculoaponeurotic fibrosarcoma oncogene homolog F (MAFF), transcript variant 1	-4.927
A_24_P303193	HNRNPA0	heterogeneous nuclear ribonucleoprotein A0 (HNRNPA0)	-4.923
A_33_P3387771	USP9X	ENST00000452501 linc lnc-USP9X-1:1 linc TCONS_00016979	-4.918
A_33_P3272390	RANBP2	RANBP2-like and GRIP domain containing 5 (RGPD5), transcript variant 1	-4.912
A_33_P3344579	DLD	dihydrolipoamide dehydrogenase (DLD), transcript variant 1	-4.900
A_33_P3252141	TMX3	thioredoxin-related transmembrane protein 3 (TMX3)	-4.890
A_33_P3410935	C17orf89	chromosome 17 open reading frame 89 (C17orf89)	-4.885
A_33_P3358977	RNASEH2C	ribonuclease H2, subunit C (RNASEH2C)	-4.873
A_23_P205584	JKAMP	JNK1/MAPK8-associated membrane protein (JKAMP), transcript variant 1	-4.855
A_33_P3268343	PGAM4	phosphoglycerate mutase family member 4 (PGAM4)	-4.848
A_33_P3628481	MGC27345	uncharacterized protein MGC27345 (MGC27345), long non-coding	-4.813
A_32_P74366	VCPIP1	valosin containing protein (p97)/p47 complex interacting protein 1	-4.807
A_33_P3220530	SRSF6	serine/arginine-rich splicing factor 6 (SRSF6), transcript variant 2, non-coding	-4.797
A_33_P3336780	ABCB8	ATP-binding cassette, sub-family B (MDR/TAP), member 8 (ABCB8), transcript variant 2	-4.789
A_33_P3346048	LOC441081	POM121 membrane glycoprotein (rat) pseudogene (LOC441081), non-coding	-4.775

The 60 common genes in two treatments concentrations compared to control in BASAL-LIKE (MDA-MB-468) and HER2 (SkBr3) cell lines.

Table A. 4.2. List of the 48 common genes in BASAL-LIKE (MDA-MB-468) cell line

Input ID	Gene symbol	Gene Name	Average of interaction
A_23_P146146	ATP6V0D2	ATPase H+ transporting V0 subunit d2(ATP6V0D2)	-6.668
A_23_P100203	HSBP1	heat shock factor binding protein 1(HSBP1)	-5.407
A_23_P122228	NDUFS6	NADH:ubiquinone oxidoreductase subunit S6(NDUFS6)	-5.282
A_23_P24997	CDK4	cyclin dependent kinase 4(CDK4)	-5.265
A_24_P89080	DCK	deoxycytidine kinase(DCK)	-5.178
A_32_P210642	EGFL7	EGF like domain multiple 7(EGFL7)	-5.094
A_23_P117095	FGF23	fibroblast growth factor 23(FGF23)	-5.072
A_24_P28657	AHCTF1	AT-hook containing transcription factor 1(AHCTF1)	-5.043
A_23_P11262	F8A2	coagulation factor VIII-associated 2(F8A2)	-4.900
A_23_P211504	KDELR3	KDEL endoplasmic reticulum protein retention receptor 3(KDELR3)	-4.769
A_23_P118536	SLFN12	schlafen family member 12(SLFN12)	-4.740
A_23_P42935	BRAF	B-Raf proto-oncogene, serine/threonine kinase(BRAF)	-4.704
A_23_P105138	CAT	catalase(CAT)	-4.659
A_23_P256455	RPA3	replication protein A3(RPA3)	-4.560
A_23_P106544	CMC2	C-X9-C motif containing 2(CMC2)	-4.511
A_23_P56810	SLC4A1AP	solute carrier family 4 member 1 adaptor protein(SLC4A1AP)	-4.420
A_23_P120414	YWHAB	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein beta(YWHAB)	-4.376
A_23_P100344	ORC6	origin recognition complex subunit 6(ORC6)	-4.372
A_23_P98431	HMBS	hydroxymethylbilane synthase(HMBS)	-4.368
A_23_P103905	UFC1	ubiquitin-fold modifier conjugating enzyme 1(UFC1)	-4.355
A_23_P2873	KLC1	kinesin light chain 1(KLC1)	-4.320
A_23_P103149	ACO2	aconitase 2(ACO2)	-4.304
A_23_P105705	FGF6	fibroblast growth factor 6(FGF6)	-4.240
A_23_P157316	C7orf34	chromosome 7 open reading frame 34(C7orf34)	-4.224

A_23_P48088	CD27	CD27 molecule(CD27)	-4.214
A_23_P101521	IZUMO2	IZUMO family member 2(IZUMO2)	-4.205
A_23_P31135	ACAT2	acetyl-CoA acetyltransferase 2(ACAT2)	-4.159
A_23_P102950	rsph1	radial spoke head 1 homolog(RSPH1)	-4.145
A_24_P67898	MGEA5	meningioma expressed antigen 5 (hyaluronidase)(MGEA5)	-4.093
A_32_P815507	LOC100130920	uncharacterized LOC100130920(LOC100130920)	-4.093
A_23_P312300	SCGB2A1	secretoglobin family 2A member 1(SCGB2A1)	-4.074
A_24_P381136	PACSIN3	protein kinase C and casein kinase substrate in neurons 3(PACSIN3)	-4.026
A_23_P370097	TMEM237	transmembrane protein 237(TMEM237)	-4.000
A_23_P321160	ZNF594	zinc finger protein 594(ZNF594)	-3.999
A_23_P145777	NDUFA4	NDUFA4, mitochondrial complex associated(NDUFA4)	-3.997
A_23_P120237	STARD7	StAR related lipid transfer domain containing 7(STARD7)	-3.984
A_33_P3291877	ARID1B	AT rich interactive domain 1B (SWI1-like) (ARID1B), transcript variant 2, mRNA	-3.978
A_33_P3315314	MT1HL1	metallothionein 1H-like 1 (MT1HL1)	-3.973
A_33_P3223631	ENST00000601550	Unknown	-3.962
A_33_P3284586	FBXL8	F-box and leucine-rich repeat protein 8 (FBXL8)	-3.944
A_33_P3785051	EFCAB10	cDNA clone IMAGE:6616931, partial cds.	-3.943
A_33_P3362567	A_33_P3362567	Unknown	-3.916
A_33_P3388745	LOC100132207	cDNA FLJ41345 fis, clone BRAWH2002761.	3.869
A_33_P3372788	NBPF8	neuroblastoma breakpoint family, member 8 (NBPF8), transcript variant 1	-3.858
A_33_P3424577	ENST00000613594	T cell receptor beta constant 1 [Source:HGNC Symbol;Acc:HGNC:12156]	3.834
A_33_P3424122	TCHHL1	trichohyalin-like 1 (TCHHL1)	-3.782
A_33_P3695548	POLE	polymerase (DNA directed), epsilon, catalytic subunit (POLE), mRNA [NM_006231]	-3.774
A_33_P3367615	KLF6	Kruppel-like factor 6 (KLF6), transcript variant B	-3.744

The 49 common genes in two treatments concentrations compared to control in BASAL-LIKE (MDA-MB-468) cell line.

Table A. 4.3. List of the 77 common genes in HER2 (SkBr3) cell line

Input ID	Gene symbol	Gene Name	Average of interaction
A_23_P122228	NDUFS6	NADH:ubiquinone oxidoreductase subunit S6(NDUFS6)	-6.254
A_23_P24997	CDK4	cyclin dependent kinase 4(CDK4)	-5.351
A_24_P89080	DCK	deoxycytidine kinase(DCK)	-5.308
A_23_P117095	FGF23	fibroblast growth factor 23(FGF23)	-5.223
A_23_P131089	KANK3	KN motif and ankyrin repeat domains 3(KANK3)	-5.191
A_24_P28657	AHCTF1	AT-hook containing transcription factor 1(AHCTF1)	-5.122
A_23_P13914	dhx37	DEAH-box helicase 37(DHX37)	-4.813
A_23_P217475	IDS	iduronate 2-sulfatase(IDS)	-4.690
A_23_P105138	CAT	catalase(CAT)	-4.620
A_23_P10995	RBMS3	RNA binding motif single stranded interacting protein 3(RBMS3)	-4.610
A_23_P94879	F2	coagulation factor II, thrombin(F2)	-4.573
A_23_P98431	HMBS	hydroxymethylbilane synthase(HMBS)	-4.542
A_23_P318616	LRTM2	leucine rich repeats and transmembrane domains 2(LRTM2)	-4.447
A_24_P277934	COL1A2	collagen type I alpha 2 chain(COL1A2)	-4.428
A_23_P169017	DEFB103A	defensin beta 103A(DEFB103A)	4.330
A_23_P112874	GPC5	glypican 5(GPC5)	-4.315
A_24_P410086	SSBP4	single stranded DNA binding protein 4(SSBP4)	4.251
A_23_P117082	HEBP1	heme binding protein 1(HEBP1)	-4.208
A_23_P200874	CEP85	centrosomal protein 85(CEP85)	-4.159
A_23_P35977	PDZD3	PDZ domain containing 3(PDZD3)	-4.151
A_23_P386254	NKX3-2	NK3 homeobox 2(NKX3-2)	-4.113
A_24_P67898	MGEA5	meningioma expressed antigen 5 (hyaluronidase)(MGEA5)	-4.095
A_23_P45365	COL4A5	collagen type IV alpha 5 chain(COL4A5)	-4.035
A_23_P9280	KIF27	kinesin family member 27(KIF27)	-4.005
A_23_P253524	CENPE	centromere protein E(CENPE)	-3.963

A_24_P915007	NACC1	nucleus accumbens associated 1(NACC1)	3.918
A_23_P20876	PTPDC1	protein tyrosine phosphatase domain containing 1(PTPDC1)	-3.851
A_23_P11262	F8A2	coagulation factor VIII-associated 2(F8A2)	-3.826
A_23_P51587	RGS7	regulator of G-protein signaling 7(RGS7)	-3.773
A_23_P104146	ZMYM4	zinc finger MYM-type containing 4(ZMYM4)	-3.764
A_23_P42935	BRAF	B-Raf proto-oncogene, serine/threonine kinase(BRAF)	3.747
A_23_P106505	lcmt2	leucine carboxyl methyltransferase 2(LCMT2)	-3.741
A_23_P2683	RPAP3	RNA polymerase II associated protein 3(RPAP3)	-3.716
A_23_P94998	LETM1	leucine zipper and EF-hand containing transmembrane protein 1(LETM1)	3.673
A_23_P388190	DIDO1	death inducer-obliterator 1(DIDO1)	-3.665
A_23_P100344	ORC6	origin recognition complex subunit 6(ORC6)	-3.646
A_23_P105705	FGF6	fibroblast growth factor 6(FGF6)	3.639
A_32_P50123	SRGAP2	SLIT-ROBO Rho GTPase activating protein 2(SRGAP2)	-3.599
A_23_P53057	ZNF215	zinc finger protein 215(ZNF215)	3.595
A_23_P108554	DDX1	DEAD-box helicase 1(DDX1)	3.595
A_23_P101521	IZUMO2	IZUMO family member 2(IZUMO2)	3.543
A_23_P102950	rsph1	radial spoke head 1 homolog(RSPH1)	-3.528
A_23_P109345	PTTG1IP	pituitary tumor-transforming 1 interacting protein(PTTG1IP)	-3.496
A_23_P112220	INSL4	insulin like 4(INSL4)	-3.465
A_23_P120048	BAZ2B	bromodomain adjacent to zinc finger domain 2B(BAZ2B)	-3.461
A_23_P160828	C1orf159	chromosome 1 open reading frame 159(C1orf159)	-3.438
A_23_P312300	SCGB2A1	secretoglobin family 2A member 1(SCGB2A1)	-3.436
A_32_P815507	LOC100130920	uncharacterized LOC100130920(LOC100130920)	3.429
A_23_P25698	SLC10A1	solute carrier family 10 member 1(SLC10A1)	-3.417
A_23_P41395	CCKAR	cholecystokinin A receptor(CCKAR)	3.406
A_23_P321160	ZNF594	zinc finger protein 594(ZNF594)	-3.390
A_23_P171095	USP27X	ubiquitin specific peptidase 27, X-linked(USP27X)	-3.388
A_23_P120237	STARD7	StAR related lipid transfer domain containing 7(STARD7)	-3.387
A_33_P3248629	DENND2A	ens DENN/MADD domain containing 2A [Source:HGNC Symbol;Acc:HGNC:22212] [ENST00000492720]	-3.374

A_33_P3225690	ZNF516	zinc finger protein 516 (ZNF516)	-3.367
A_33_P3406939	KIF24	kinesin family member 24 (KIF24)	-3.364
A_24_P358131	ENST00000404956	keratin 18 pseudogene 52 [Source:HGNC Symbol;Acc:HGNC:37888]	-3.325
A_33_P3350086	OR2T33	olfactory receptor, family 2, subfamily T, member 33 (OR2T33)001004695]	-3.315
A_33_P3344292	SAMD4A	sterile alpha motif domain containing 4A [ENST00000554335]	3.304
A_33_P3387991	CEBPE	CCAAT/enhancer binding protein (C/EBP), epsilon (CEBPE)	-3.302
A_33_P3785051	EFCAB10	cDNA clone IMAGE:6616931, partial cds.	-3.296
A_23_P69089	AK021889	cDNA FLJ11827 fis, clone HEMBA1006502	3.286
A_33_P3695548	POLE	polymerase (DNA directed), epsilon, catalytic subunit (POLE)	-3.273
A_33_P3240693	THSD4	thrombospondin, type I, domain containing 4 (THSD4), transcript variant 2	-3.272
A_33_P3375086	THC2621369	tc Q39C09_BURS3 (Q39C09) Flagellar FliF M-ring protein, partial (3%)	-3.270
A_33_P3221808	FAM205BP	mRNA; cDNA DKFZp434J193 (from clone DKFZp434J193).	-3.242
A_33_P3328274	A_33_P3328274	Unknown	-3.234
A_33_P3315314	MT1HL1	metallothionein 1H-like 1 (MT1HL1)	-3.232
A_33_P3354569	GPD2	glycerol-3-phosphate dehydrogenase 2 (mitochondrial) (GPD2), transcript variant 1	-3.215
A_24_P200162	HIGD1A	HIG1 hypoxia inducible domain family, member 1A (HIGD1A), transcript variant 3	-3.206
A_32_P196263	ADAMTS9	ADAM metalloproteinase with thrombospondin type 1 motif, 9 (ADAMTS9)	3.194
A_23_P37778	FHOD1	formin homology 2 domain containing 1 (FHOD1)	3.189
A_33_P3317618	SYN2	synapsin II (SYN2), transcript variant IIb	-3.187
A_32_P34876	WDR93	WD repeat domain 93 (WDR93), transcript variant 1	-3.177
A_23_P99360	TRIM13	tripartite motif containing 13 (TRIM13), transcript variant 3	-3.164
A_33_P3223631	ENST00000601550	Unknown	-3.146
A_23_P146146	ATP6V0D2	ATPase, H ⁺ transporting, lysosomal 38kDa, V0 subunit d2 (ATP6V0D2)	-3.136

The 77 common genes in two treatments concentrations compared to control in HER2 (SkBr3) cell line.

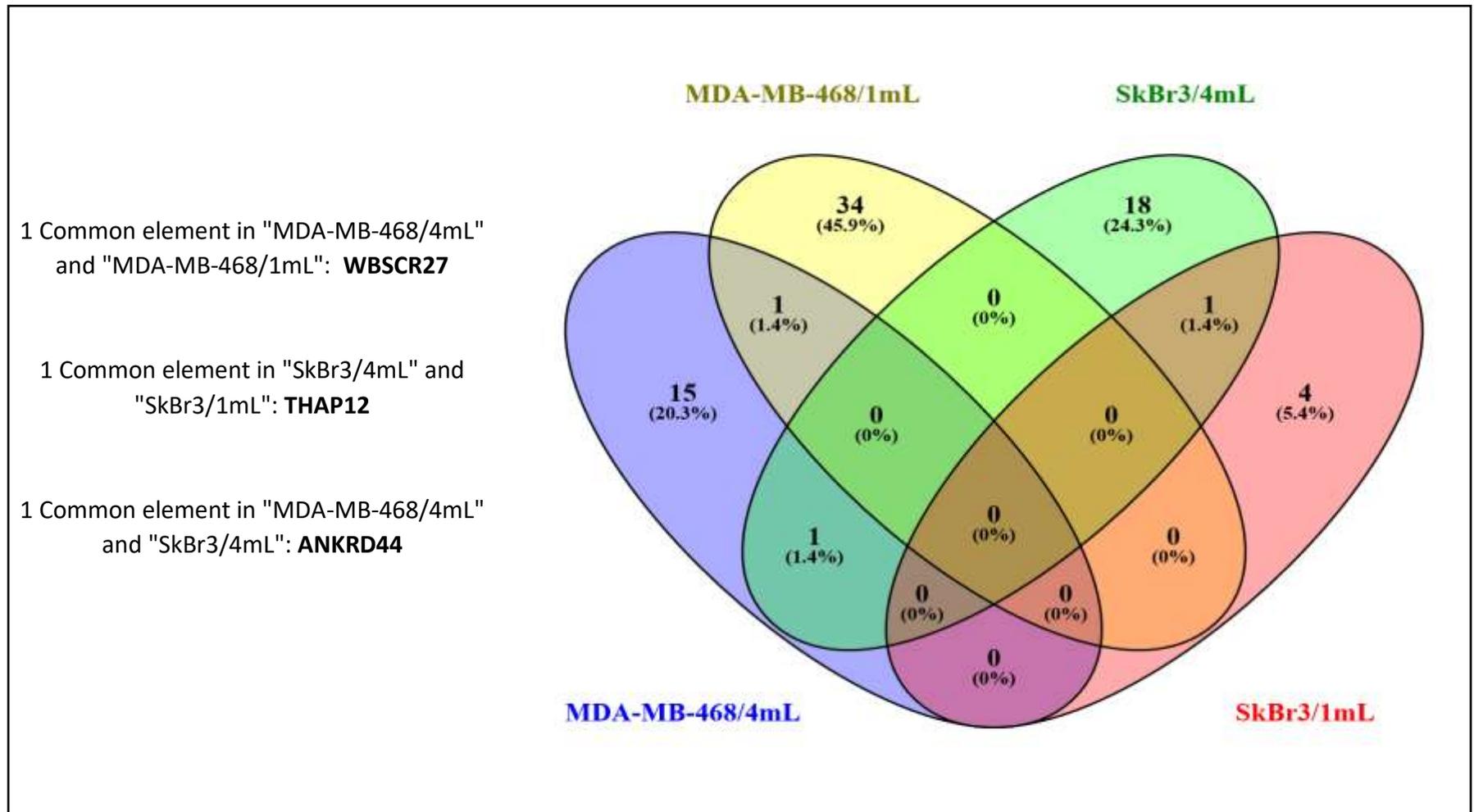


Figure A.4. 4. Venn diagram showing commonalities between the different analyses Regression-based method and fold change-based method for both cell lines BASAL-LIKE (MDA-MB-468) and HER2 (SkBr3) cell lines and different concentrations (1 and 4 mL) of Metformin.

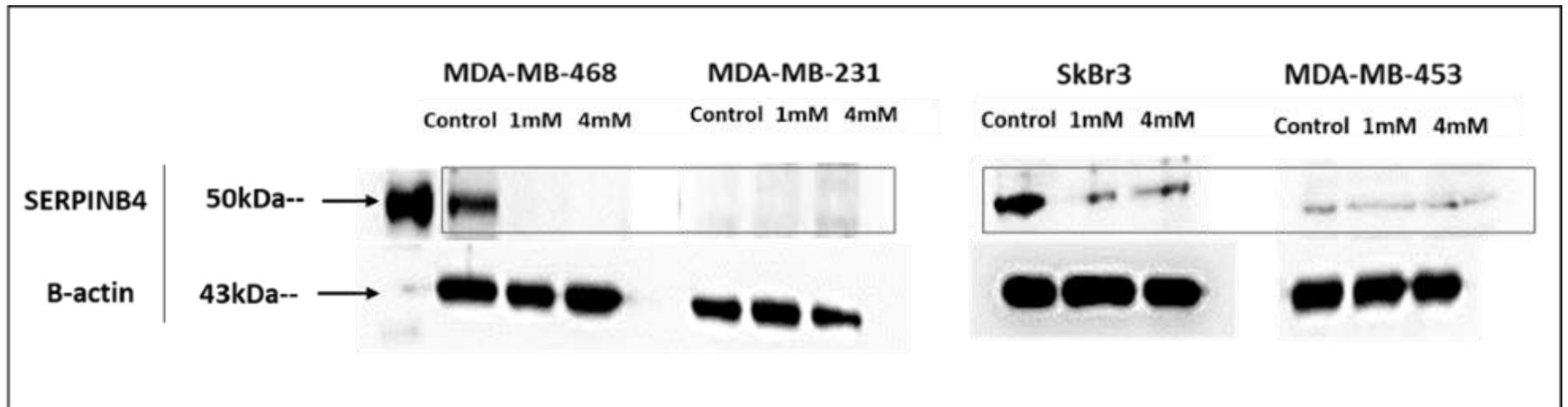


Figure A.5.1. Representative micrographs of immunoblots showing SERPINB4 expression in MDA-MB-468 and MDA-MB-231 (left panel), and SkBr3 and MDA-MB-453 (right panel). β -actin is used as a loading control. Cell extracts for immunoblotting were obtained from untreated and Metformin treated cells with 1mM and 4mM concentrations.

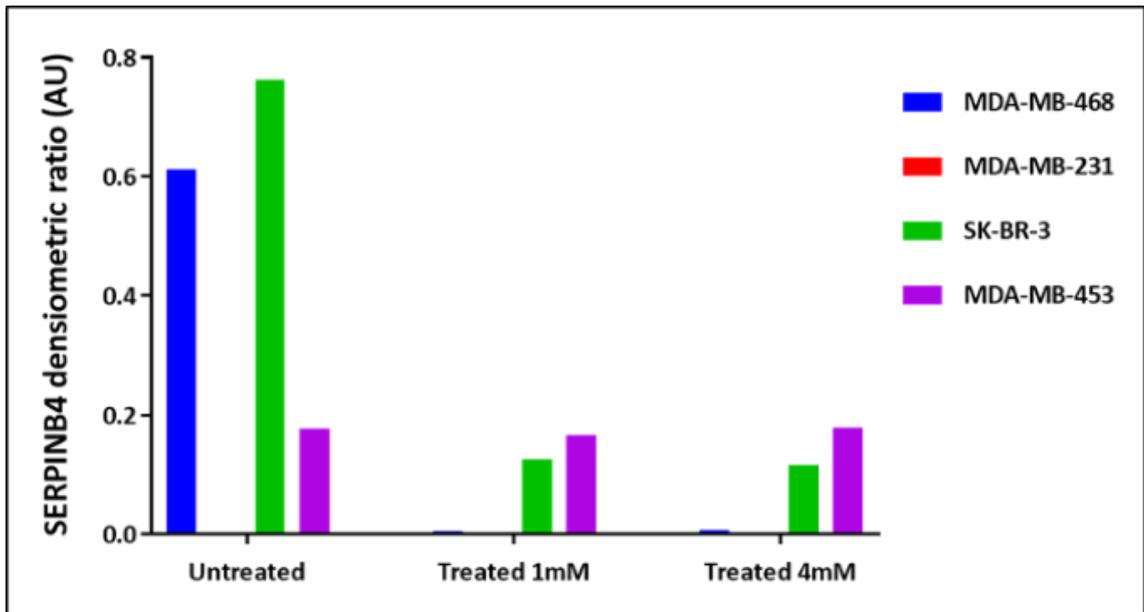


Figure A.5.2. Micrograph representing the densitometry ratio (AU) of SERPINB4 protein expression in Metformin untreated and treated (1mM and 4mM) in MDA-MB-468, MDA-MB-231, SkBr3 and MDA-MB-453 breast cancer cell lines. low protein expression level was observed in 1mM and 4mM Metformin-treated cell lines, and a non-detectable protein in both MDA-MB-468 and MDA-MB-231 cell lines.

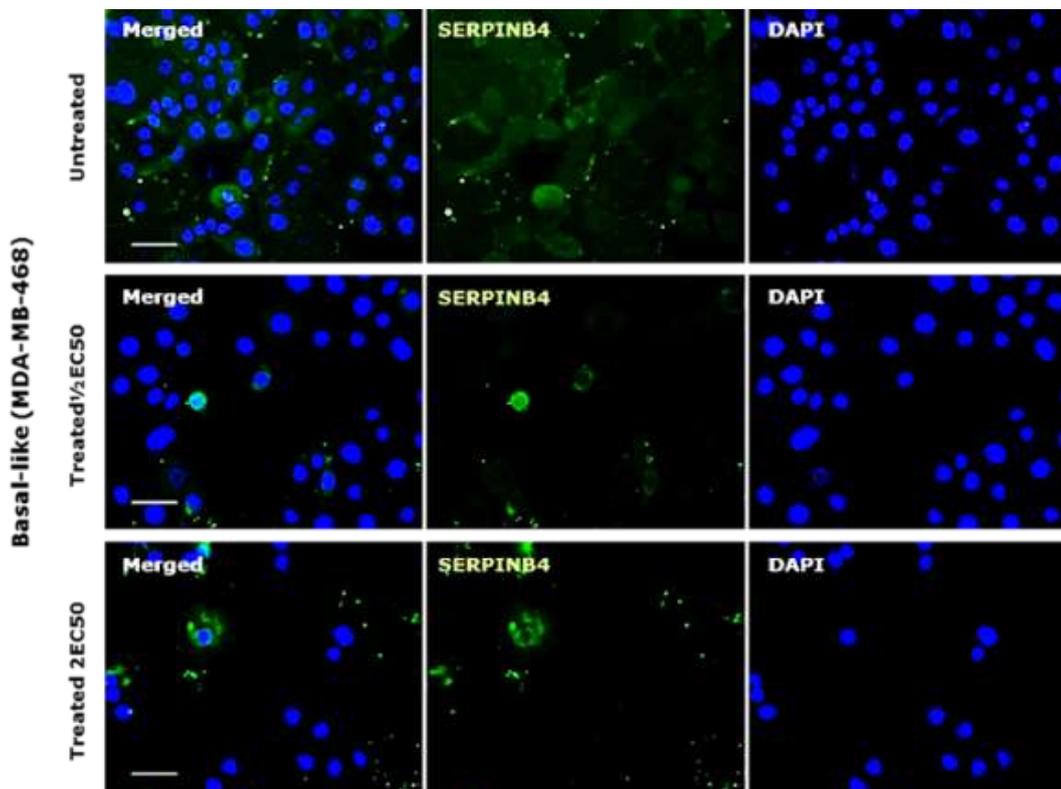


Figure A.5.3. Micrographs displaying expression of SERPINB4 in Basal-like (MDA-MB-468) cell line against SERPINB4 (green). Nuclei were stained with DAPI (blue). the images were taken at 20X magnification and scale bars indicate 100µm.

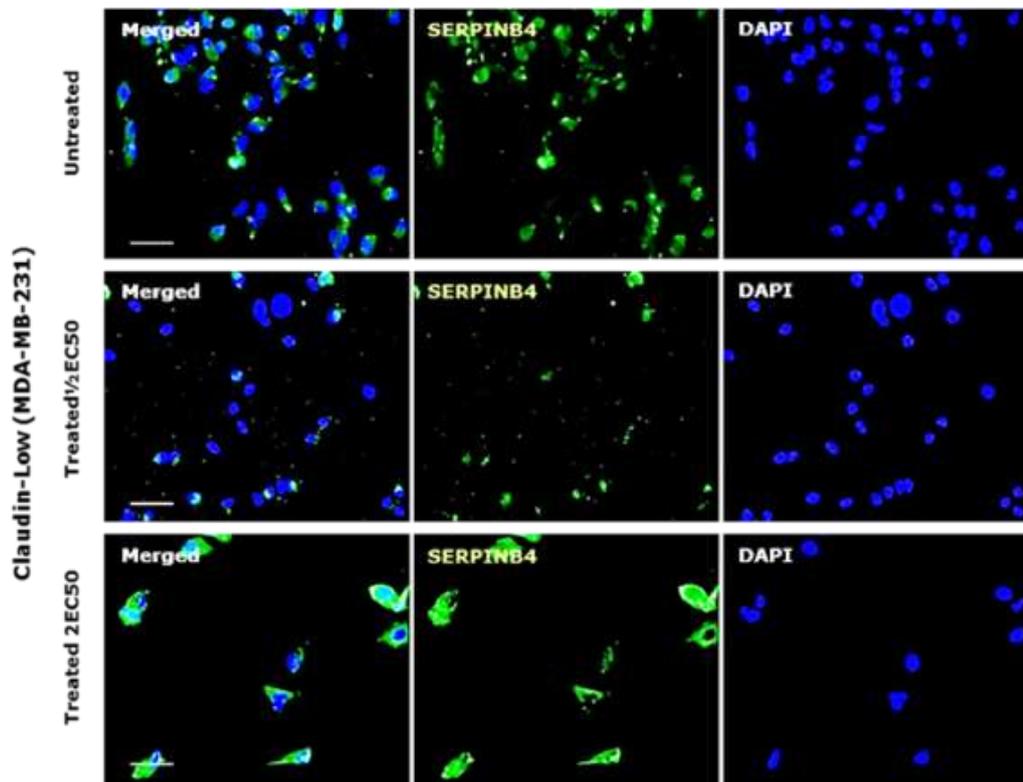


Figure A.5.4. Micrographs displaying expression of SERPINB4 in Claudin-Low (MDA-MB-231) cell line against SERPINB4 (green). Nuclei were stained with DAPI (blue). the images were taken at 20X magnification and scale bars indicate 100µm.

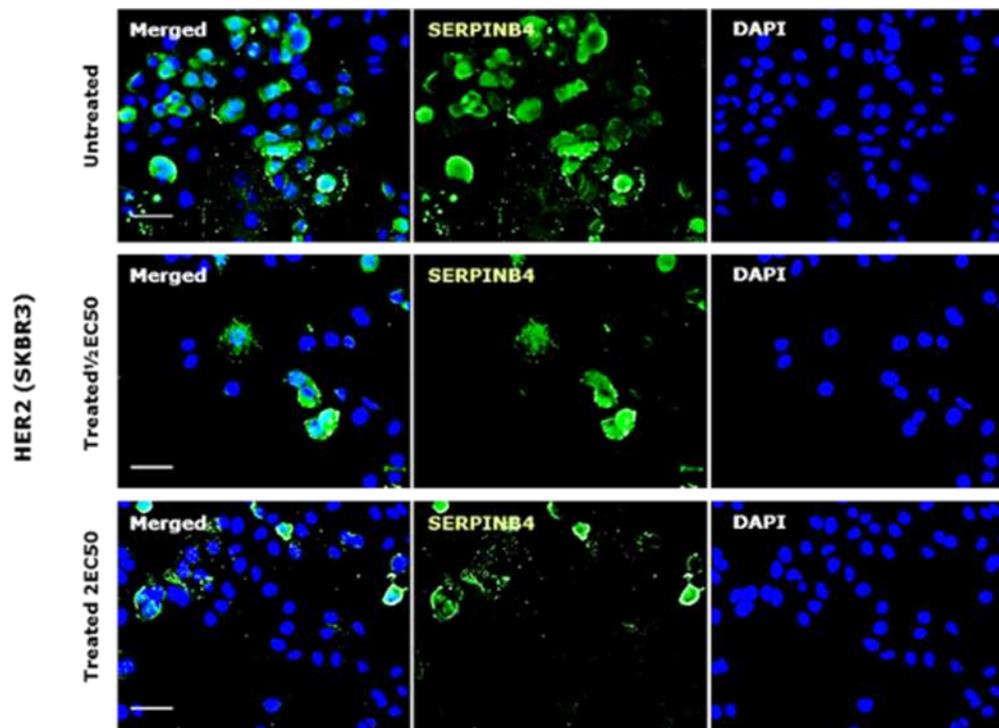


Figure A.5. 5. Micrographs displaying expression of SERPINB4 in HER2 (SkBr3) cell line against SERPINB4 (green). Nuclei were stained with DAPI (blue). the images were taken at 20X magnification and scale bars indicate 100µm.

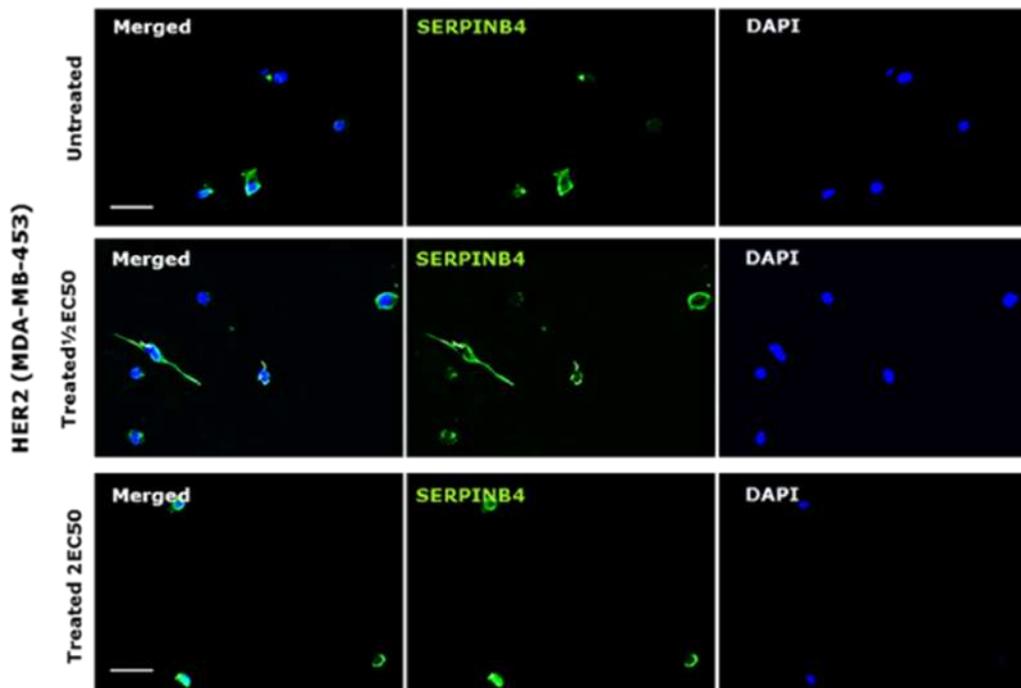


Figure A.5.6. Micrographs displaying expression of SERPINB4 in HER2 (MDA-MB-453) cell line against SERPINB4 (green). Nuclei were stained with DAPI (blue). the images were taken at 20X magnification and scale bars indicate 100µm.

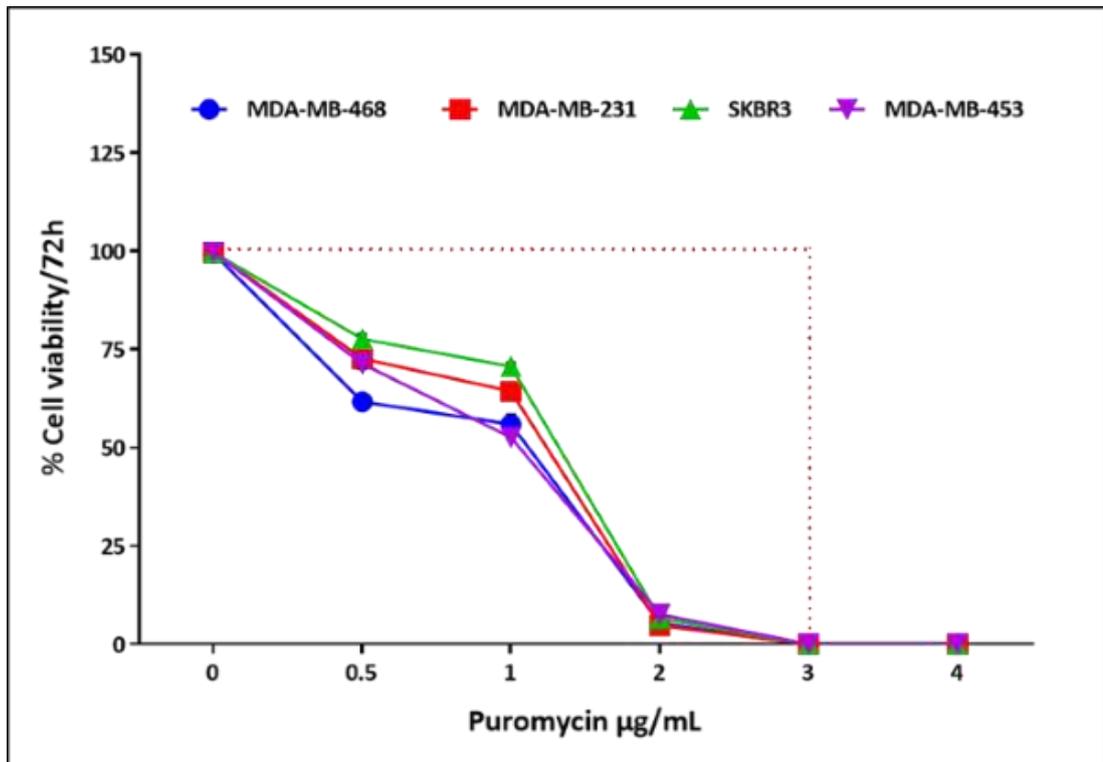


Figure A.6.1. The dose-response curve of antibiotic selection in breast cancer cell lines. Micrograph showed that 100% of the cells were killed after 72 h of exposure to 3µg/mL of puromycin.

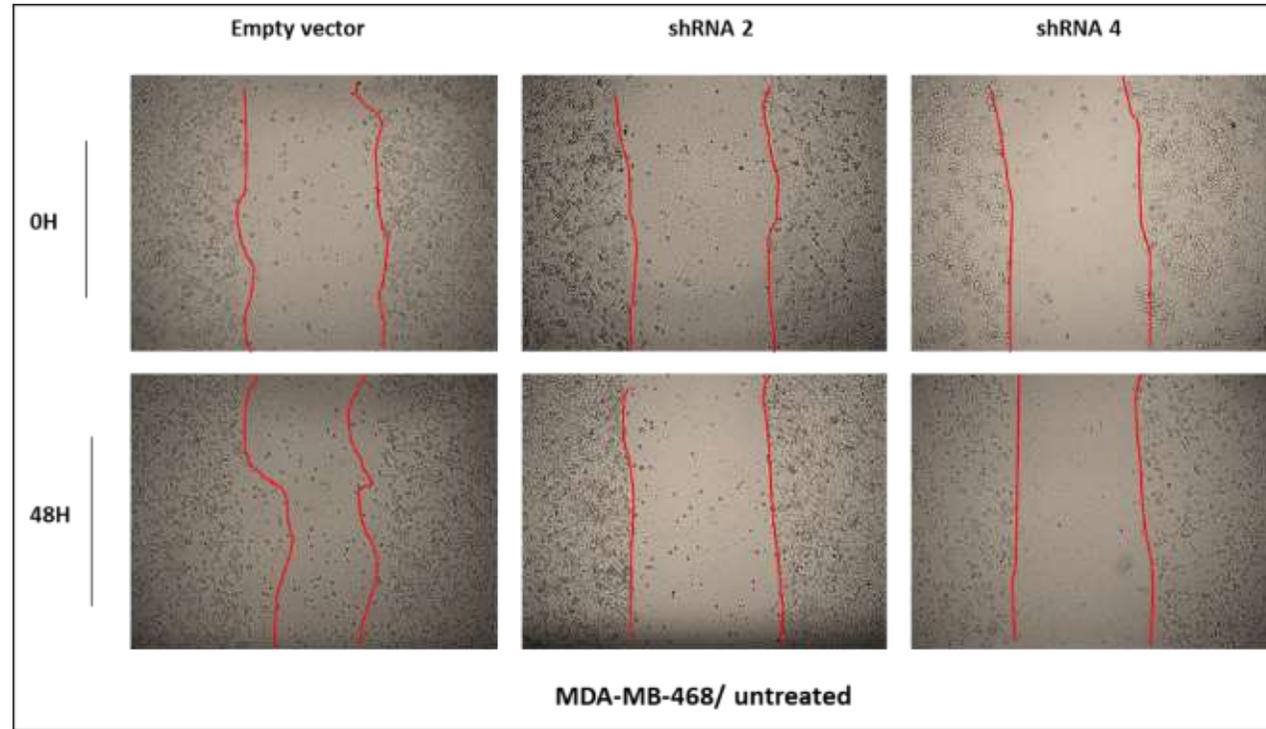


Figure.A 6.2. Wound-healing assay (Scratch assay) using MDA-MB-468 cells expressing empty vector, *PTK2B*- shRNA1 and *PTK2B*- shRNA2 untreated constructs.

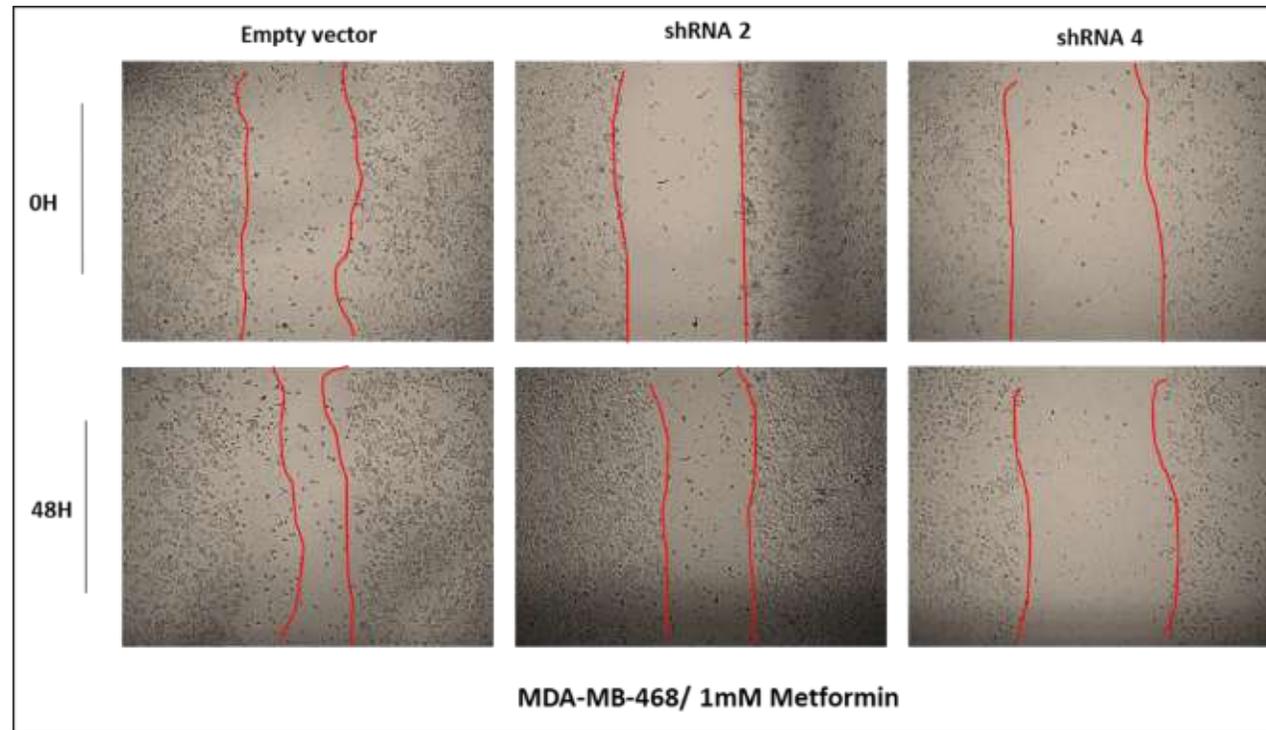


Figure.A 6.3. Wound-healing assay (Scratch assay) using MDA-MB-468 cells expressing empty vector, *PTK2B*- shRNA1 and *PTK2B*- shRNA2 1mM Metformin treated constructs.

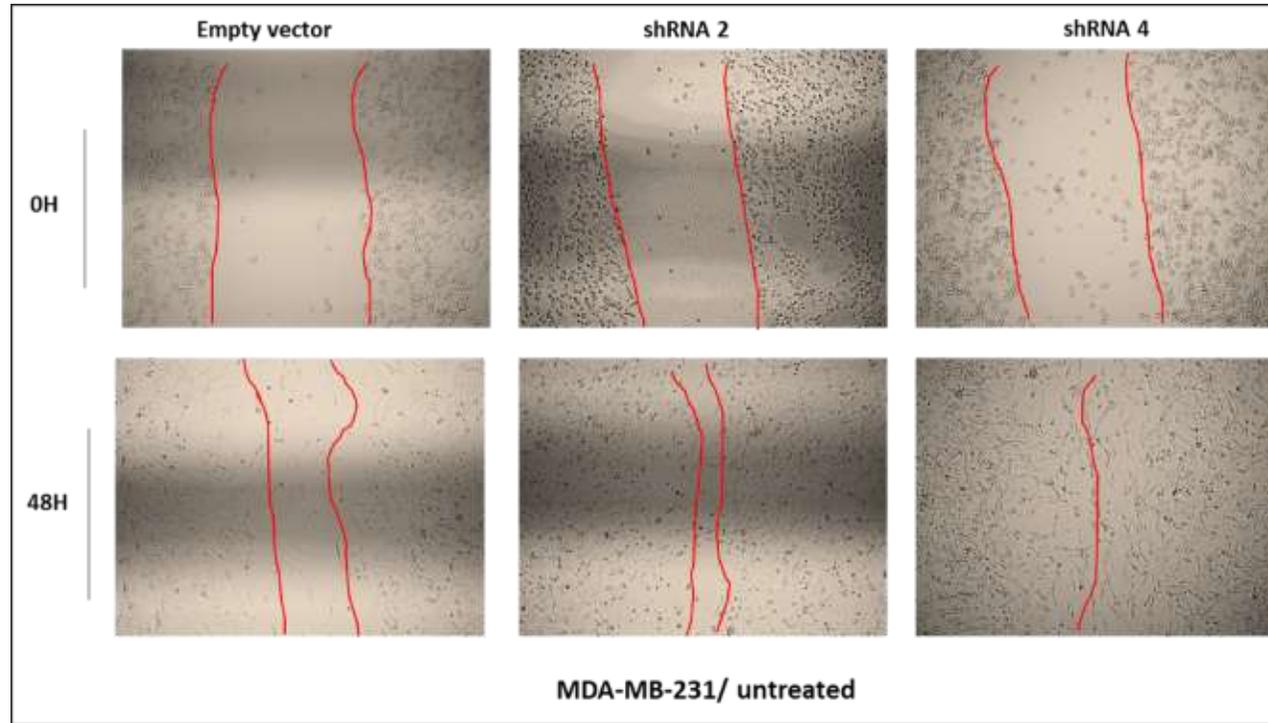


Figure.A 6.4. Wound-healing assay (Scratch assay) using MDA-MB-231 cells expressing empty vector, *PTK2B*- shRNA1 and *PTK2B*- shRNA2 untreated constructs.

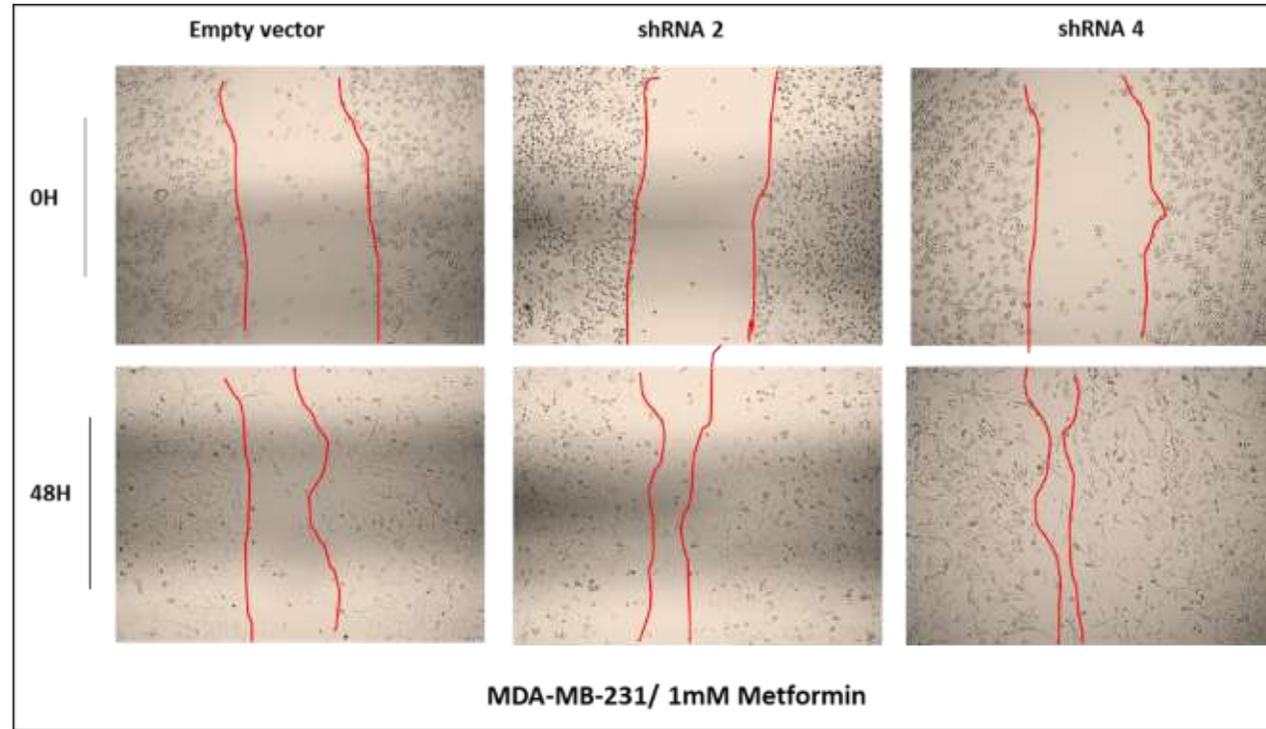


Figure.A 6.5. Wound-healing assay (Scratch assay) using MDA-MB-231 cells expressing empty vector, *PTK2B*- shRNA1 and *PTK2B*- shRNA2 1mM Metformin treated constructs.

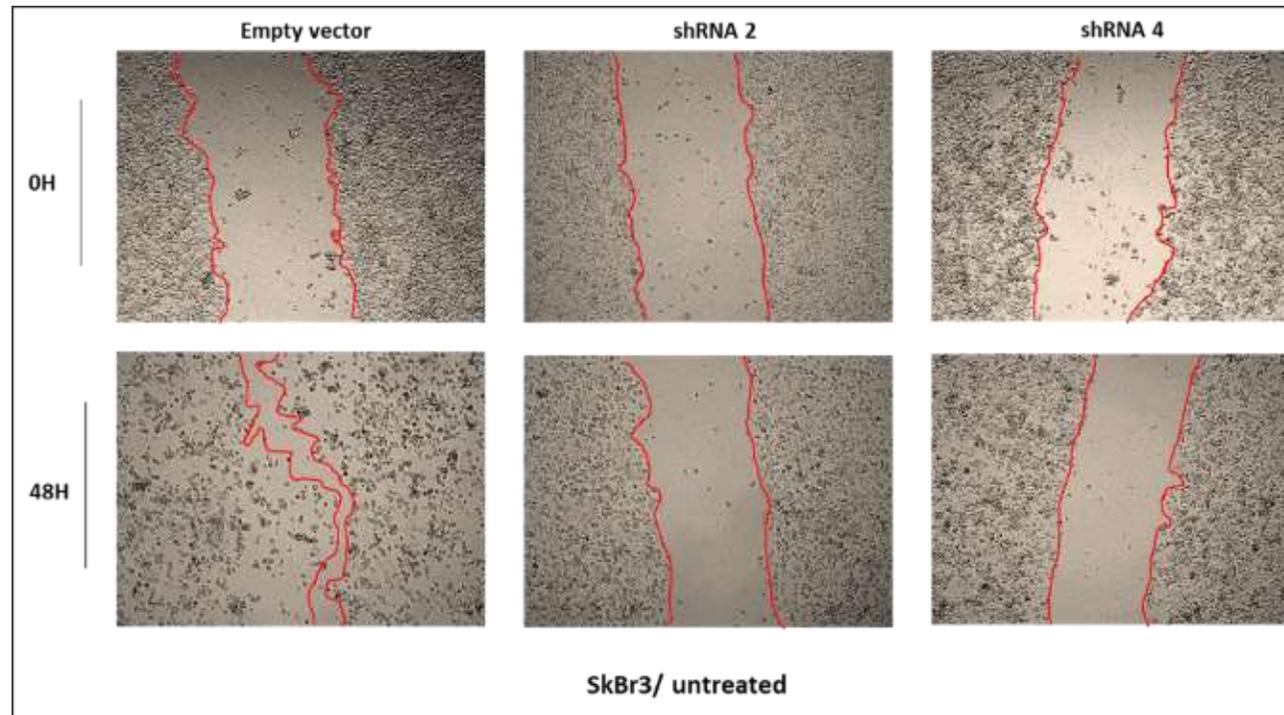


Figure.A 6.6. Wound-healing assay (Scratch assay) using SkBr3 cells expressing empty vector, *PTK2B*- shRNA1 and *PTK2B*- shRNA2 untreated constructs.

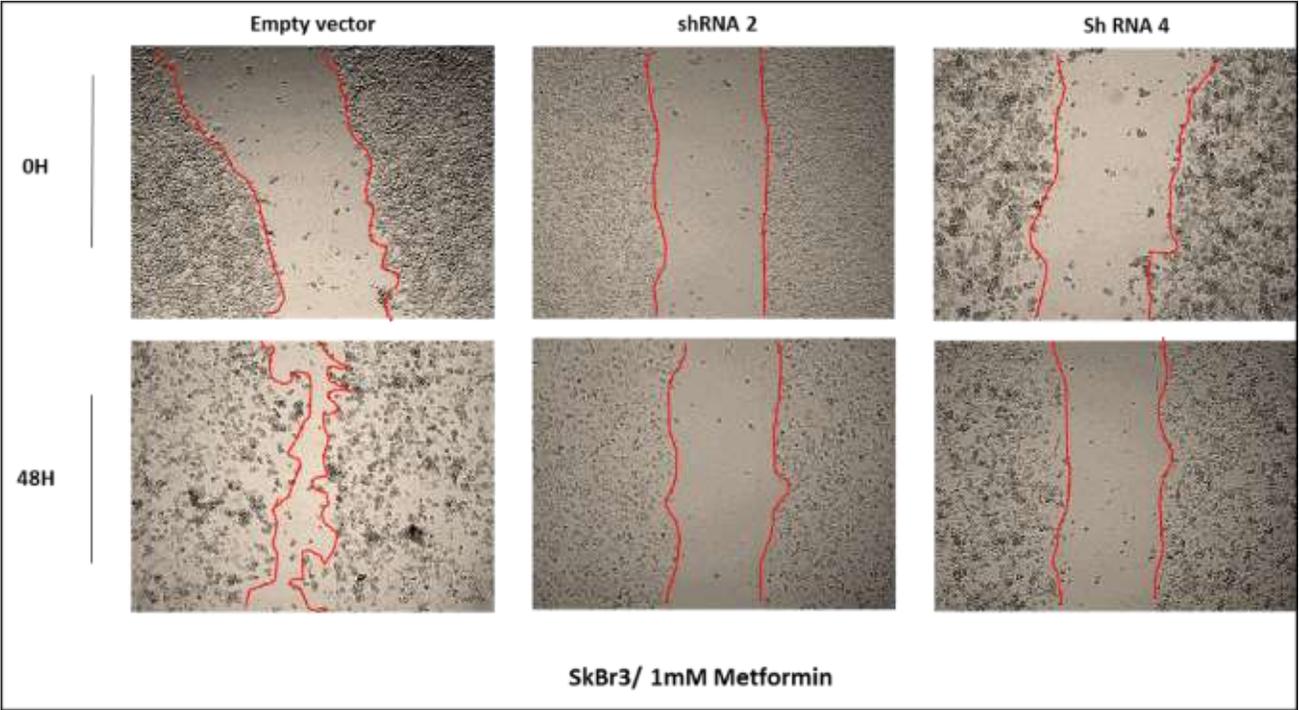


Figure.A 6.7. Wound-healing assay (Scratch assay) using SkBr3 cells expressing empty vector, *PTK2B*- shRNA1 and *PTK2B*- shRNA2 1mM Metformin treated constructs.

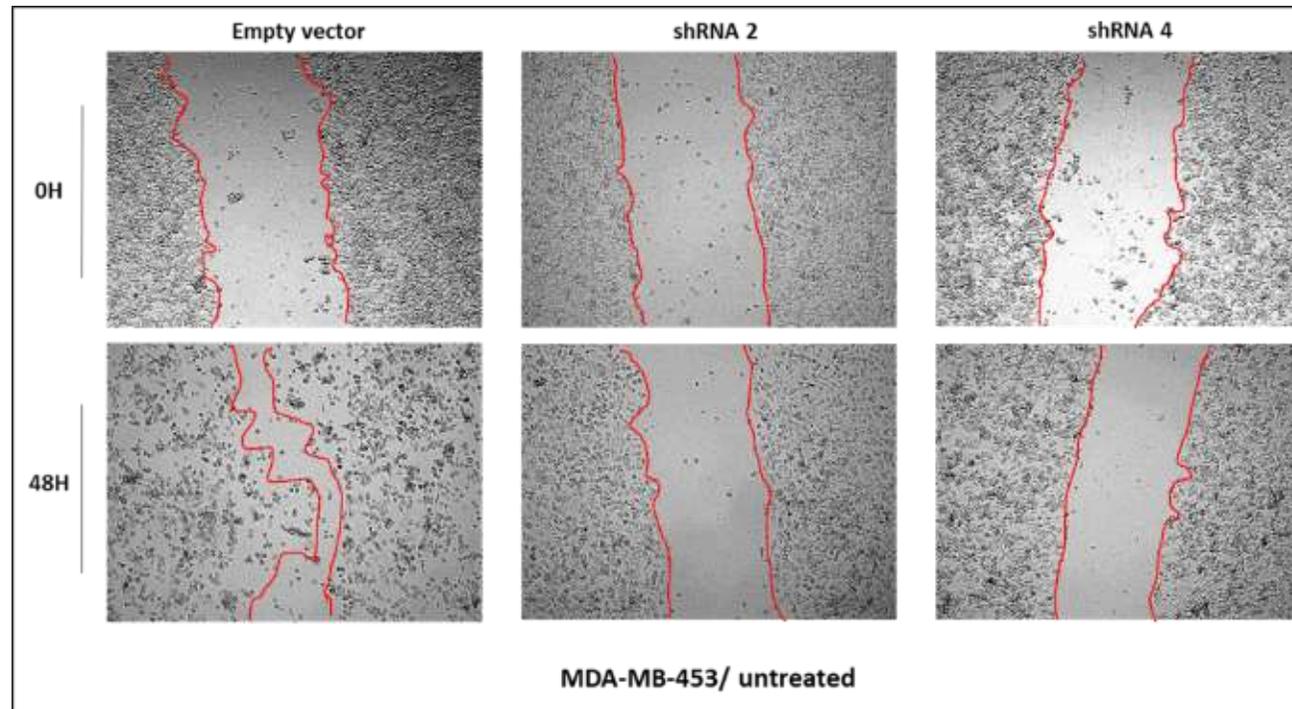


Figure.A 6.8. Wound-healing assay (Scratch assay) using MDA-MB-453 cells expressing empty vector, *PTK2B*- shRNA1 and *PTK2B*- shRNA2 untreated constructs.

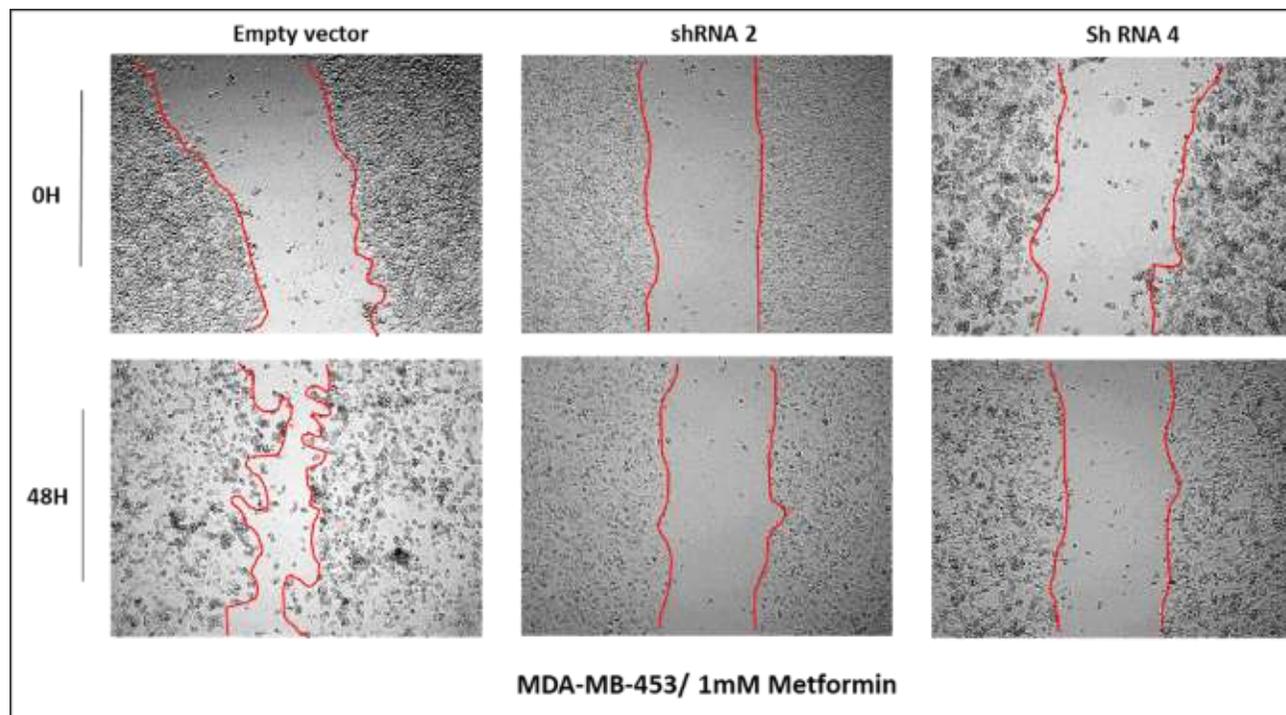


Figure.A 6.9. Wound-healing assay (Scratch assay) using MDA-MB-453 cells expressing empty vector, *PTK2B*- shRNA1 and *PTK2B*- shRNA2 1mM Metformin treated constructs.

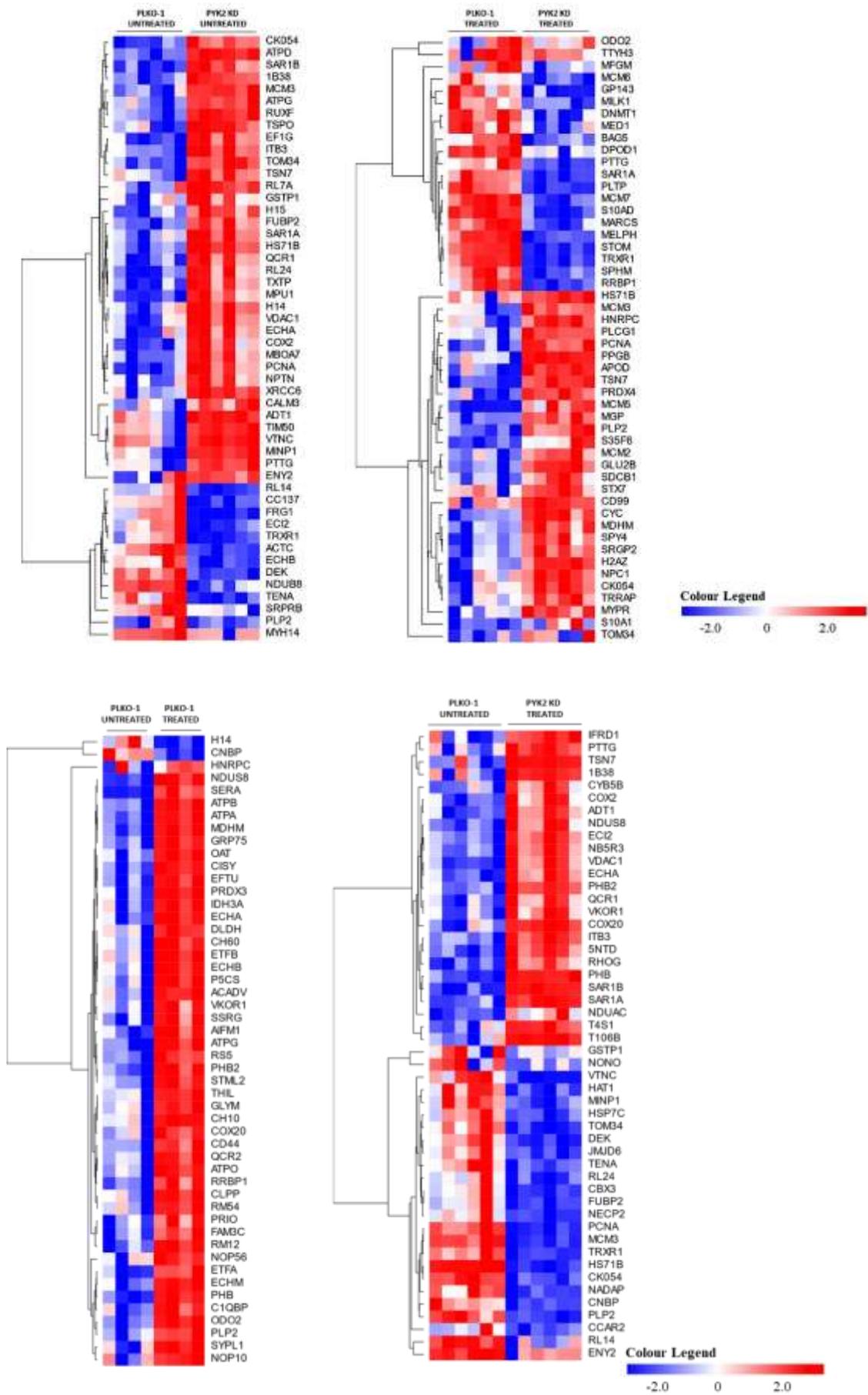


Figure A.6. 10. Differentially expressed proteins from untreated and treated MDA-MB-453 control and *PTK2B* cells. Data refers to the Results section 6.7.

Table A.6. 1. Top 50 upregulated and downregulated proteins in HER2 (MDA-MB-453) cell line samples

Protein Symbol	Description	Fold change	Confidence level
QCR2	Cytochrome b-c1 complex subunit 2, mitochondrial	9.113	0.706
ECHA	Trifunctional enzyme subunit alpha, mitochondrial	8.719	0.825
ATPG	ATP synthase subunit gamma, mitochondrial	8.566	0.762
ATPB	ATP synthase subunit beta, mitochondrial	7.050	0.859
ATPA	ATP synthase subunit alpha, mitochondrial	6.683	0.830
PHB2	Prohibitin-2	6.392	0.757
NDUS8	NADH dehydrogenase [ubiquinone] iron-sulfur protein 8, mitochondrial	6.206	0.763
P5CS	Delta-1-pyrroline-5-carboxylate synthase	6.134	0.787
ATPO	ATP synthase subunit O, mitochondrial	5.969	0.718
CLPP	ATP-dependent Clp protease proteolytic subunit, mitochondrial	5.897	0.730
EFTU	Elongation factor Tu, mitochondrial	5.788	0.825
IDH3A	Isocitrate dehydrogenase [NAD] subunit alpha, mitochondrial	5.743	0.712
VKOR1	Vitamin K epoxide reductase complex subunit 1	5.664	0.730
PHB	Prohibitin	5.541	0.873
ECHB	Trifunctional enzyme subunit beta, mitochondrial	5.508	0.775
OAT	Ornithine aminotransferase, mitochondrial	5.116	0.790
CISY	Citrate synthase, mitochondrial	4.783	0.837
ECHM	Enoyl-CoA hydratase, mitochondrial	4.738	0.828
STML2	Stomatin-like protein 2, mitochondrial	4.688	0.704
COX20	Cytochrome c oxidase protein 20 homolog	4.584	0.715
THIL	Acetyl-CoA acetyltransferase, mitochondrial	4.325	0.757
GRP75	Stress-70 protein, mitochondrial	4.145	0.921
SSRG	Translocon-associated protein subunit gamma	3.937	0.775
ACADV	Very long-chain specific acyl-CoA dehydrogenase, mitochondrial	3.886	0.764
CH60	60 kDa heat shock protein, mitochondrial	3.870	0.841
CH10	10 kDa heat shock protein, mitochondrial	3.684	0.766
C1QBP	Complement component 1 Q subcomponent-binding protein, mitochondrial	3.632	0.748
ETFA	Electron transfer flavoprotein subunit alpha, mitochondrial	3.628	0.727
ODO2	Dihydrolipoyllysine-residue succinyltransferase component of 2-oxoglutarate dehydrogenase complex, mitochondrial	3.535	0.761
GLYM	Serine hydroxymethyltransferase, mitochondrial	3.514	0.763
MDHM	Malate dehydrogenase, mitochondrial	3.477	0.873
PRDX3	Thioredoxin-dependent peroxide reductase, mitochondrial	3.330	0.812
RM12	39S ribosomal protein L12, mitochondrial	3.253	0.731
RM54	39S ribosomal protein L54, mitochondrial	3.248	0.776
ETFB	Electron transfer flavoprotein subunit beta	3.172	0.745

CD44	CD44 antigen	3.052	0.735
RS5	40S ribosomal protein S5	2.948	0.848
DLDH	Dihydrolipoyl dehydrogenase, mitochondrial	2.638	0.759
AIFM1	Apoptosis-inducing factor 1, mitochondrial	2.595	0.832
RRBP1	Ribosome-binding protein 1	2.323	0.784
NOP56	Nucleolar protein 56	2.225	0.728
PLP2	Proteolipid protein 2	2.176	0.947
SYPL1	Synaptophysin-like protein 1	2.113	0.919
FAM3C	Protein FAM3C	2.099	0.701
PRIO	Major prion protein	2.005	0.702
NOP10	H/ACA ribonucleoprotein complex subunit 3	1.810	0.769
HNRPC	Heterogeneous nuclear ribonucleoproteins C1/C2	1.698	0.707
SERA	D-3-phosphoglycerate dehydrogenase	1.348	0.730

Protein Symbol	Description	Fold change	Confidence level
H14	Histone H1.4	-2.087	0.715
CNBP	Cellular nucleic acid-binding protein	-2.443	0.740

Comparing untreated and Metformin-treated PLKO-1 samples with $Fc \geq -2$, and confidence of 70%. Red indicated upregulated proteins and blue downregulated proteins.

Table A.6. 2. Top 50 upregulated and downregulated proteins in HER2 (MDA-MB-453) cell line samples

Protein Symbol	Description	Fold change	Confidence level
ADT1	ADP/ATP translocase 1	18.424	0.822
ATPG	ATP synthase subunit gamma, mitochondrial	11.460	0.799
TXTP	Tricarboxylate transport protein, mitochondrial	11.405	0.733
TIM50	Mitochondrial import inner membrane translocase subunit TIM50	11.173	0.718
COX2	Cytochrome c oxidase subunit 2	10.555	0.779
VDAC1	Voltage-dependent anion-selective channel protein 1	10.528	0.789
SAR1B	GTP-binding protein SAR1b	10.071	0.855
MYH14	Myosin-14	9.769	0.706
SAR1A	GTP-binding protein SAR1a	9.447	0.749
1B38	HLA class I histocompatibility antigen, B-38 alpha chain	9.341	0.853
MBOA7	Lysophospholipid acyltransferase 7	9.128	0.771
ITB3	Integrin beta-3	9.060	0.776
TSPO	Translocator protein	8.325	0.747
CALM3	Calmodulin-3	8.302	0.704
ECHA	Trifunctional enzyme subunit alpha, mitochondrial	8.049	0.849
SRPRB	Signal recognition particle receptor subunit beta	7.929	0.729
NPTN	Neuroplastin	7.891	0.755
QCR1	Cytochrome b-c1 complex subunit 1, mitochondrial	7.706	0.714
NDUB8	NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 8, mitochondrial	7.284	0.730
ATPD	ATP synthase subunit delta, mitochondrial	7.075	0.735
TSN7	Tetraspanin-7	6.997	0.724
PTTG	Pituitary tumor-transforming gene 1 protein-interacting protein	6.916	0.765
ECI2	Enoyl-CoA delta isomerase 2, mitochondrial	6.855	0.778
ECHB	Trifunctional enzyme subunit beta, mitochondrial	6.719	0.794
MPU1	Mannose-P-dolichol utilization defect 1 protein	6.560	0.748

Protein Symbol	Description	Fold change	Confidence level
RUXF	Small nuclear ribonucleoprotein F	-1.646	0.735
XRCC6	X-ray repair cross-complementing protein 6	-1.652	0.736
ACTC	Actin, alpha cardiac muscle 1	-1.663	0.802
EF1G	Elongation factor 1-gamma	-1.669	0.781
GSTP1	Glutathione S-transferase P	-1.726	0.770
FUBP2	Far upstream element-binding protein 2	-1.798	0.895
FRG1	Protein FRG1	-1.817	0.716
MINP1	Multiple inositol polyphosphate phosphatase 1	-1.854	0.803
DEK	Protein DEK	-1.858	0.807
PCNA	Proliferating cell nuclear antigen	-1.962	0.796
TOM34	Mitochondrial import receptor subunit TOM34	-1.987	0.716
CK054	Ester hydrolase C11orf54	-2.084	0.854
ENY2	Transcription and mRNA export factor ENY2	-2.111	0.894
RL7A	60S ribosomal protein L7a	-2.189	0.787
MCM3	DNA replication licensing factor MCM3	-2.197	0.776
RL24	60S ribosomal protein L24	-2.291	0.878
TRXR1	Thioredoxin reductase 1, cytoplasmic	-2.615	0.816
CC137	Coiled-coil domain-containing protein 137	-2.743	0.729
RL14	60S ribosomal protein L14	-2.841	0.882
TENA	Tenascin	-3.582	0.856
HS71B	Heat shock 70 kDa protein 1B	-3.673	0.878
H15	Histone H1.5	-3.895	0.767
PLP2	Proteolipid protein 2	-4.246	1.000
H14	Histone H1.4	-4.274	0.759
VTNC	Vitronectin	-5.846	0.839

Comparing untreated PLKO-1 and *PYK2* KD samples with $Fc \geq -2$, and confidence of 70%. Red indicated upregulated proteins and blue downregulated proteins.

Table A.6. 3. Top 50 upregulated and downregulated proteins in HER2 (MDA-MB-453) cell line samples

Protein Symbol	Description	Fold change	Confidence level
SAR1A	GTP-binding protein SAR1a	6.367	0.828
PLTP	Phospholipid transfer protein	4.212	0.705
GP143	G-protein coupled receptor 143	3.626	0.895
BAG5	BAG family molecular chaperone regulator 5	3.472	0.792
TSN7	Tetraspanin-7	3.335	0.887
SRGP2	SLIT-ROBO Rho GTPase-activating protein 2	3.068	0.917
MILK1	MICAL-like protein 1	2.961	0.773
SPY4	Protein sprouty homolog 4	2.727	0.732
S10AD	Protein S100-A13	2.611	0.711
MELPH	Melanophilin	2.484	0.754
SDCB1	Syntenin-1	2.445	0.835
MFGM	Lactadherin	2.343	0.862
APOD	Apolipoprotein D	2.327	0.887
S35F6	Solute carrier family 35 member F6	2.236	0.908
NPC1	Niemann-Pick C1 protein	2.206	0.886
PTTG	Pituitary tumor-transforming gene 1 protein-interacting protein	2.043	0.722
H2AZ	Histone H2A.Z	2.035	0.910
MYPR	Myelin proteolipid protein	1.935	0.764
STOM	Erythrocyte band 7 integral membrane protein	1.919	0.817
STX7	Syntaxin-7	1.893	0.860
CD99	CD99 antigen	1.873	0.703
TTYH3	Protein tweety homolog 3	1.864	0.789
MARCS	Myristoylated alanine-rich C-kinase substrate	1.844	0.862
PPGB	Lysosomal protective protein	1.838	0.780
S10A1	Protein S100-A1	1.814	0.900

Protein Symbol	Description	Fold change	Confidence level
GLU2B	Glucosidase 2 subunit beta	-1.589	0.751
TOM34	Mitochondrial import receptor subunit TOM34	-1.640	0.766
MCM7	DNA replication licensing factor MCM7	-1.691	0.817
HNRPC	Heterogeneous nuclear ribonucleoproteins C1/C2	-1.691	0.845
MCM2	DNA replication licensing factor MCM2	-1.691	0.805
MCM5	DNA replication licensing factor MCM5	-1.729	0.823
MCM6	DNA replication licensing factor MCM6	-1.753	0.813
DPOD1	DNA polymerase delta catalytic subunit	-1.788	0.758
CYC	Cytochrome c	-1.802	0.735
CK054	Ester hydrolase C11orf54	-1.815	0.813
MDHM	Malate dehydrogenase, mitochondrial	-1.851	0.838
PCNA	Proliferating cell nuclear antigen	-1.958	0.887
PRDX4	Peroxiredoxin-4	-1.970	0.738
DNMT1	DNA (cytosine-5)-methyltransferase 1	-2.053	0.816
MCM3	DNA replication licensing factor MCM3	-2.082	0.827
ODO2	Dihydrolipoyllysine-residue succinyltransferase component of 2-oxoglutarate dehydrogenase complex, mitochondrial	-2.164	0.726
TRXR1	Thioredoxin reductase 1, cytoplasmic	-2.234	0.837
RRBP1	Ribosome-binding protein 1	-2.434	0.836
MGP	Matrix Gla protein	-3.237	0.871
SPHM	N-sulphoglucosamine sulphohydrolase	-3.722	0.785
HS71B	Heat shock 70 kDa protein 1B	-3.920	0.926
MED1	Mediator of RNA polymerase II transcription subunit 1	-7.762	0.799
TRRAP	Transformation/transcription domain-associated protein	-7.828	0.752
PLCG1	1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase gamma-1	-7.868	0.760
PLP2	Proteolipid protein 2	-11.326	0.952

Comparing Metformin-treated PLKO-1 and *PYK2* KD samples with $Fc \geq -2$, and confidence of 70%. Red indicated upregulated proteins and blue downregulated proteins.

Table A.6. 4. Top 50 upregulated and downregulated proteins in HER2 (MDA-MB-453) cell line samples

Protein Symbol	Description	Fold change	Confidence level
ADT1	ADP/ATP translocase 1	13.499	0.742
SAR1A	GTP-binding protein SAR1a	10.812	0.778
NDUAC	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 12	10.489	0.734
RHOG	Rho-related GTP-binding protein RhoG	9.006	0.706
T106B	Transmembrane protein 106B	8.776	0.713
SAR1B	GTP-binding protein SAR1b	8.292	0.846
VDAC1	Voltage-dependent anion-selective channel protein 1	8.020	0.735
1B38	HLA class I histocompatibility antigen, B-38 alpha chain	7.796	0.810
COX2	Cytochrome c oxidase subunit 2	7.624	0.733
IFRD1	Interferon-related developmental regulator 1	7.482	0.800
ITB3	Integrin beta-3	7.329	0.776
PTTG	Pituitary tumor-transforming gene 1 protein-interacting protein	6.870	0.761
QCR1	Cytochrome b-c1 complex subunit 1, mitochondrial	6.767	0.727
VKOR1	Vitamin K epoxide reductase complex subunit 1	6.217	0.868
ECHA	Trifunctional enzyme subunit alpha, mitochondrial	6.134	0.778
NDUS8	NADH dehydrogenase [ubiquinone] iron-sulfur protein 8, mitochondrial	6.110	0.789
COX20	Cytochrome c oxidase protein 20 homolog	5.844	0.841
PHB	Prohibitin	5.651	0.904
NB5R3	NADH-cytochrome b5 reductase 3	5.631	0.751
TSN7	Tetraspanin-7	5.508	0.721
ECI2	Enoyl-CoA delta isomerase 2, mitochondrial	5.405	0.757
PHB2	Prohibitin-2	5.390	0.919
5NTD	5'-nucleotidase	5.192	0.785
CYB5B	Cytochrome b5 type B	5.099	0.713
T4S1	Transmembrane 4 L6 family member 1	5.072	0.743

Protein Symbol	Description	Fold change	Confidence level
GSTP1	Glutathione S-transferase P	-1.612	0.719
JMJD6	Bifunctional arginine demethylase and lysyl-hydroxylase JMJD6	-1.613	0.729
CCAR2	Cell cycle and apoptosis regulator protein 2	-1.633	0.791
RL24	60S ribosomal protein L24	-1.659	0.726
HSP7C	Heat shock cognate 71 kDa protein	-1.681	0.785
NADAP	Kanadaptin	-1.758	0.888
NONO	Non-POU domain-containing octamer-binding protein	-1.759	0.726
NECP2	Adaptin ear-binding coat-associated protein 2	-1.800	0.775
FUBP2	Far upstream element-binding protein 2	-1.801	0.884
MINP1	Multiple inositol polyphosphate phosphatase 1	-1.820	0.814
RL14	60S ribosomal protein L14	-1.891	0.713
DEK	Protein DEK	-1.967	0.811
ENY2	Transcription and mRNA export factor ENY2	-2.013	0.773
CBX3	Chromobox protein homolog 3	-2.030	0.729
CK054	Ester hydrolase C11orf54	-2.049	0.823
TOM34	Mitochondrial import receptor subunit TOM34	-2.185	0.788
HAT1	Histone acetyltransferase type B catalytic subunit	-2.303	0.724
MCM3	DNA replication licensing factor MCM3	-2.400	0.768
PCNA	Proliferating cell nuclear antigen	-2.479	0.837
TENA	Tenascin	-2.546	0.794
TRXR1	Thioredoxin reductase 1, cytoplasmic	-2.553	0.820
CNBP	Cellular nucleic acid-binding protein	-2.648	0.711
HS71B	Heat shock 70 kDa protein 1B	-4.798	0.881
PLP2	Proteolipid protein 2	-5.274	1.000
VTNC	Vitronectin	-7.530	0.795

Comparing untreated PLKO-1 and Metformin-treated PYK2 KD samples with $Fc \geq -2$, and confidence of 70%. Red indicated upregulated proteins and blue downregulated proteins.

Table A.6.5. Pathways regulated by proteins that are differentially expressed between treated and untreated PLKO-1 and *PYK2* KD cells (MDA-MB-453 cell line)

	Biological functions	p-value	FDR	Protein Names
PLKO-1 Control vs. <i>PYK2</i> KD Treated up-regulated pathways	Proliferation_Negative regulation of cell proliferation	0.000	0.010	Prohibitin, Securin, COX-2 (PTGS2)
PLKO-1 Control vs. <i>PYK2</i> KD Treated down-regulated pathways	Cell cycle_S phase	0.003	0.069	MCM3, PCNA, HP1
PLKO-1 Control vs. PLKO-1 Treated up-regulated pathways	Cell adhesion_Cell-matrix interactions	0.002	0.058	CD44 (EXT), CD44 soluble, CD44, CD44 (ICD)
PLKO-1 Control vs. PLKO-1 Treated down-regulated pathways	Proliferation_Positive regulation cell proliferation	0.031	0.031	CNBP

Pathways were derived by Meta Core™ online software. Table only shows significant pathways (with FDR < 0.05). Red indicated upregulated pathways and blue downregulated pathways.

