Using AML Cell lines with induced IFNG phenotypes to produce prognostic index scores for stratification of AML patients

Melissa Anne Courtney

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

July 2024

Acknowledgements

Contents

Table of figures

Table of tables

Abstract

Acute myeloid leukaemia (AML) is a haematological malignancy which inhibits the production and maturation of functional immune cells. Patients are treated with induction chemotherapy and haematopoietic stem cell transplant (HSCT). HSCT replaces the patients defunct immune system using donor stem cells. Therefore, disruption via immune suppression is detrimental to patient prognosis. At present, AML patient's prognosis is predicted using cytogenetic abnormalities and genetic mutations, however, 50-70% of AML patients are labelled as 'intermediate' risk. This group displays variation in response to frontline chemotherapy and HSCT, indicating a need for improved stratification to assign patients tailored treatments. Prognostic indicators (PI) are crucial in guiding therapeutic decisions and improving outcomes for AML patients. This work focuses on generating prognostic index scores based on methylation modulated IFNG driven immune evasion in AML.

PI scores were generated using AML cell lines (Kasumi-1 and KG-1) subjected to treatment with IFNG, and demethylation agent 5AzaC. Pairwise linear regression identified significant treatment-induced transcriptomic changes and a shortlist of candidate transcripts associated with IFNG signalling and demethylation was created. PI scores were computed using normalised transcription data from IFNG, 5AzaC and IFNG5AzaC treated cell lines and β-values generated using cox proportional hazards forward selection model. The study employed the TCGA patient dataset for discovery and BeatAML, HOVON, German-AML, and CN-AML for validation. PI score performance was compared to established prognostic methods.

All PI scores split adult patients in the European LeukemiaNet cytogenetic risk category into subgroups with good and poor survival in the TCGA dataset, with the 5AzaC and IFNG PI scores association with OS validated in the CN-AML (intermediate risk group) data set (n = 242). Comparing the area under the curve (AUC) for PI scores (5AzaC PI AUC = 0.599, IFNG PI AUC = 0.637, and IFNG5AzaC PI AUC = 0.657), with established prognostic scores revealed comparable performance to LSC17 score (AUC=0.65) and the ELN cytogenetic risk categories (AUC=0.66). However, they were outperformed by other established scores. The study demonstrated the potential of cell line-derived PI scores to predict AML patient survival.

1 Introduction

1.1 Cancer

Cancer describes a group of diseases in which cells demonstrate abnormally increased growth, invade local tissues, and spread to distant secondary sites by metastasis. The transition from a normal cell to a cancerous one is achieved through an accumulation of mutations, which bestow a growth advantage over normal cells (Hanahan and Weinberg 2000). Cancer has a major impact on health worldwide and is a leading cause of morbidity and mortality. Consequently, a significant amount of research and resources is dedicated to cancer research. The global effort to investigate cancer has significantly advanced our understanding of the events that initiate and drive its pathogenesis, leading to the development of better tools for diagnosis, prognosis, and treatment. As cancer develops through mutation, it unsurprisingly becomes a highly heterogeneous disease, exhibiting genomic and epigenetic differences both within and between tumours. This heterogeneity means that biomarker signatures for prognosis and treatment response vary in effectiveness among patients, even with the same cancer type. This inherent variability also contributes to common issues such as treatment resistance and relapse. Therefore, investigating individual cancers to characterise mechanisms of treatment resistance is crucial for improving future patient outcomes.

1.2 Acute myeloid leukaemia

Acute myeloid leukaemia (AML) is a fast-developing haematological malignancy, characterised by infiltration of the blood and bone marrow by immature myeloid cells. This occurs through the accumulation of mutations in myeloid blasts and their progenitor myeloid stem cells, which prevent differentiation into specialised blood cells (De Kouchkovsky and Abdul-Hay 2016). The bone marrow consequently over produces immature non-functional monocytes and granulocytes, eventually causing bone marrow failure and then death. In addition, the leukemic cells enter the peripheral blood to disperse across the body, particularly to the lymph nodes, spleen and in rare cases, the brain (Estey 2018). Symptoms arise as blood conditions, which are caused by the loss of specific functional blood cells; they develop over a few weeks and quickly become severe. For example, depletion of red blood cells causes anaemia, resulting in the patient feeling fatigued. A reduced neutrophil count (Neutropenia) causes patients to suffer from more infections, and a loss of platelets prevents blood clotting. AML patients also experience

breathlessness, fevers, weight loss, easy bruising, aches and are particularly vulnerable to infection (National Health Service 2019).
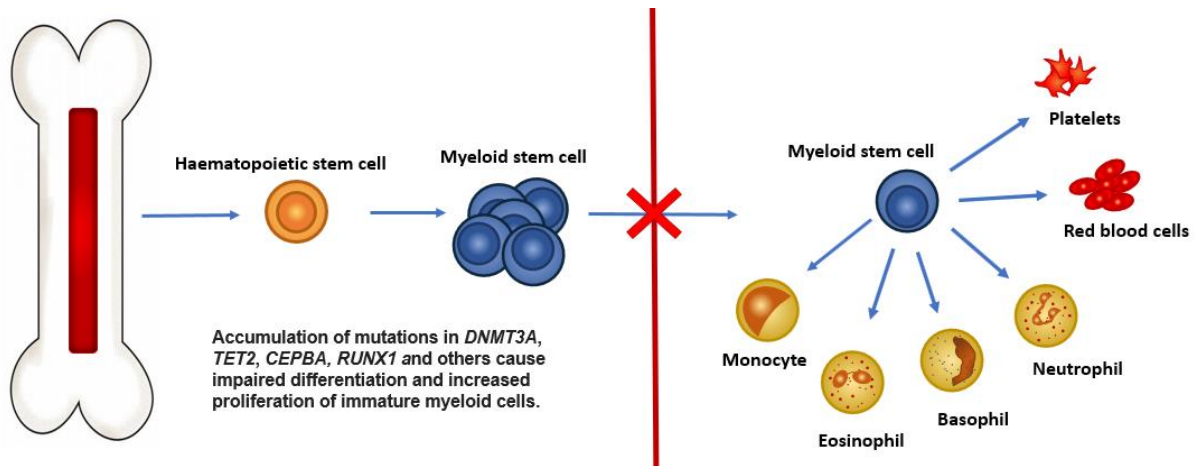
Figure 1: The transition of haematopoietic stem cells into AML blasts via mutations. The bone marrow produces haematopoietic stem cells which can become a myeloid stem cell. These stem cells ordinarily will mature into either red blood cells, platelets, or a range of white blood cells. In AML, due to accumulation of mutations, the myeloid stem cell does not develop into these specialised cells, and instead usually becomes an immature white blood cell called a myeloblast but can also become abnormal red blood cells or platelets, which are collectively termed leukaemia or blasts. These abnormal cells 'crowd out' normal healthy cells, causing symptoms from lack of platelets, white blood cells and red blood cells (De Kouchkovsky and Abdul-Hay 2016).

1.2.1 Incidence and mortality rate of AML in the UK

Acute Myeloid Leukaemia (AML) is a rare disease, accounting for less than 1% of all cancer cases with approximately 3,200 people diagnosed in the UK every year (Cancer Research 2017b). It is an age-related disease, as demonstrated by 42% new cases occurring in people aged 75 and over (Cancer Research 2017b) A gender bias is also seen in older groups, with males having a significantly higher incidence rate; most evident in the 75-79 age group where there is a 1.9-fold increase compared to females (Cancer Research 2017b).
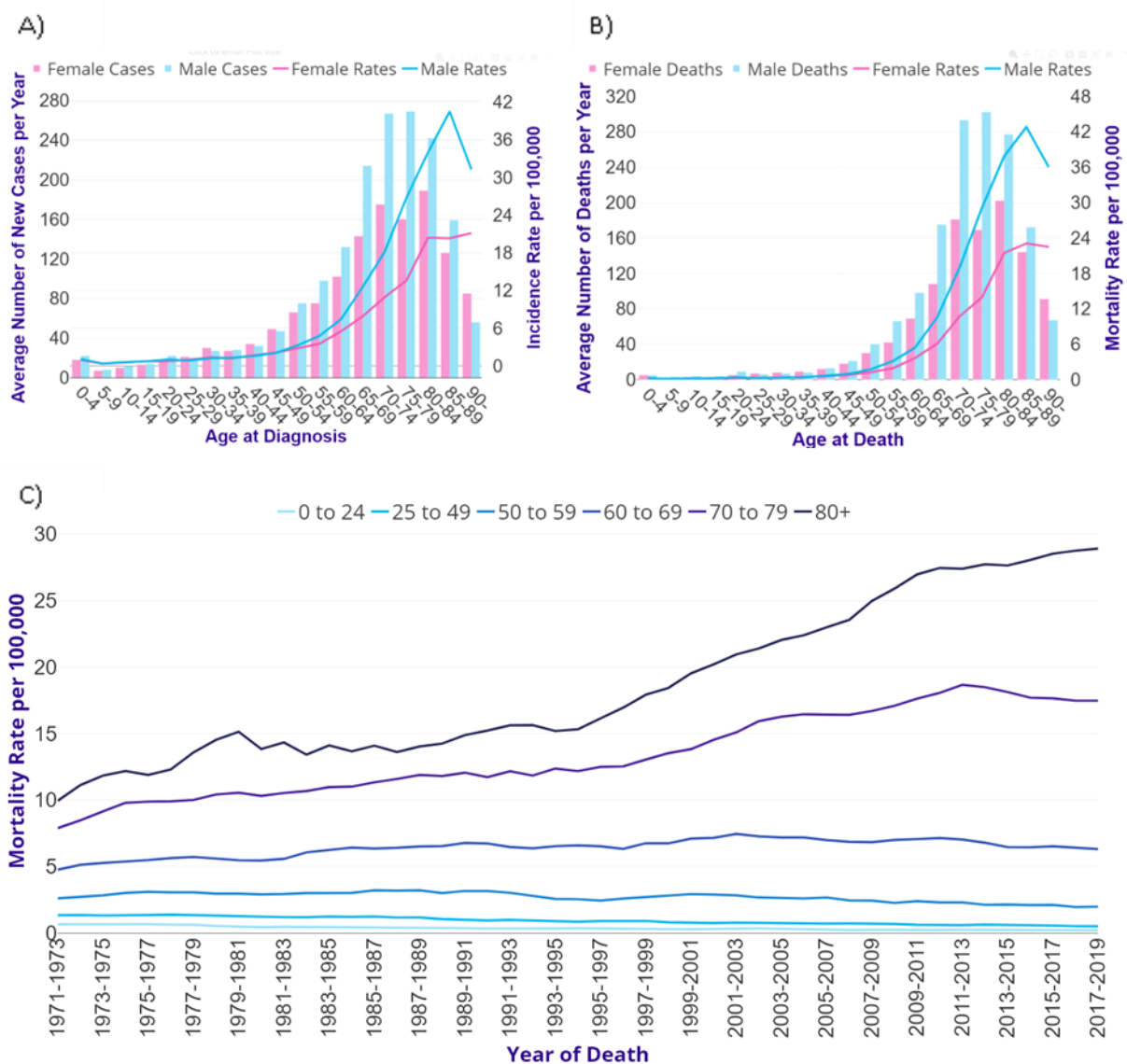
Figure 2: Incidence and Mortality Rates of AML in the UK circa 2017. A) Incidence rate of AML diagnosis per 100,000 in the UK circa 2017 (Cancer Research 2017b) *across age groups and gender; Male (blue) and female (pink). B) Mortality of AML diagnosis per 100,000 in the UK circa 2017* (Cancer Research 2017a) *across age groups and gender; Male (blue) and female (pink). C) Mortality rate of AML using European age-standardised mortality Rates per 100,000 Population in the UK circa 2017* (Cancer Research 2017a).

As depicted in Figure 2C) low mortality rate has observed for people below 50 and increased for those above. Increases in mortality rate are most dramatic in the over 80s population, who are more likely to have comorbidities, with mortality rates increased by 189% in 2017 compared to 1971 (Cancer Research 2017a).

1.3 Drivers of AML development and progression

AML is characterised by various mechanisms that allow the disease to proliferate, evade destruction by the immune system and progress. As AML originates from hematopoietic stem cells or myeloid progenitor cells, it inherently possesses self-renewal capabilities. Mutations and environmental factors, such as increased bone marrow vascularisation and altered immune environments, significantly contribute to disease progression and patient prognosis (Shih, T. T., et al. 2009, Hussong, Rodgers and Shami 2000, Trendowski 2015). AML evades cell death by overexpressing anti-apoptotic genes such as BCL2 and through TP53 mutations, which increase proliferation, leading to increased resistance to chemotherapy and poorer survival rates (Pfeffer and Singh 2018, Bories, et al. 2020). Furthermore, AML cells often grow independently of external growth signals by upregulating their own signal receptors or continuously activating intracellular signalling pathways (Hyrien 2016, Bartek and Lukas 2003).

Immune evasion is a critical factor in AML progression, where leukaemic cells alter the immune microenvironment to suppress immune responses. They achieve this by downregulating HLA molecules, expressing immune checkpoint inhibitors, and releasing immunosuppressive cytokines. These actions create a tolerant immune environment that favours regulatory T cells over effector cells (Cornel, Mimpen and Nierkens 2020, Anderson, Stromnes and Greenberg 2017). These adaptations allow AML to persist and progress despite the host's immune defence mechanisms.

1.4 AML evades the immune system by manipulating its immune environment

Genomic instability, inflammation, and reprogramming of cellular metabolism all contribute to the transformation of healthy cells into cancerous ones. Normally, the immune system detects and destroys abnormal cells. Somatic cells present antigens via major histocompatibility complexes (MHC) to immune cells. If the antigen is not recognised as self, it triggers a chain of events leading to the elimination of the abnormal cells (Chen, D. and Mellman 2017, Chen, Daniel S. and Mellman 2013). However, AML can circumvent this recognition process to escape destruction and continue to grow (Houghton and Guevara-Patiño 2004). AML uses various mechanisms to causes a paradigm shift of the immune environment from a responsive

phenotype to a tolerant phenotype, which favours regulatory T-cells over effector cells, thereby reducing immune function and preventing AML destruction (Anderson, Stromnes and Greenberg 2017). Furthermore, AML blasts downregulate the production of immune signalling molecules, including MHC, and produce high levels of immune-suppressing cytokines, such as TGFβ1, to inhibit immune activity (Cornel, Mimpen and Nierkens 2020). As AML progresses, it employs various mechanisms to evade the immune system's effector response and remodel the leukaemic microenvironment, facilitating immune escape and resistance to therapies. AML blasts use several strategies to avoid immune destruction, including:

1.4.1 AML downregulates HLAs to prevent antigen presentation to T-cells and NK-cells

HLA-A, HLA-B and HLA-C are classic major histocompatibility complexes (MHC) that present antigens for recognition by T-cells and NK-cells to identify and initiate destruction of unhealthy cells (Cornel, Mimpen and Nierkens 2020). This resistance mechanism is particularly effective against stem cell transplants, which are dependent on T and NK-cells being able to recognise antigen-MHC complexes to AML cells (Jan, et al. 2019). AML blasts genetically delete HLAs and downregulate them through epigenetic alterations. Additional methods of immunosuppression include upregulating T-cell inhibitory ligands and release of immunosuppressive molecules such as PDL-1 and IDO-1 (Christopher, et al. 2018).

1.4.2 AML induces T-cell exhaustion via immune check point inhibitors

Ordinarily, immune checkpoints maintain the immune system's balance and promote self-tolerance. AML exploits these checkpoints to suppress effector cells in its immediate environment. A well-studied checkpoint in AML is the programmed cell death protein 1 (PD-1), expressed on the surface of T and B cells (Taghiloo and Asgarian-Omran 2021). AML cells express the inhibitory ligand PDL1 on their surface, which can bind to PD-1 and transmit a co-inhibitory signal causing T-cell exhaustion. Galectin-9 is also expressed on AML blasts, where it binds to T-cell immunoglobulin and mucin domain 3 (TIM-3) on effector T cells and NK cells. This interaction promotes the self-renewal of AML blasts through stimulatory β-catenin and NFκB signalling, while simultaneously inhibiting immune cells from releasing pro-inflammatory cytokines, thus impairing their function (Silva, et al. 2017). Additionally, high mRNA levels of LAG-3, a ligand for immune checkpoint cytotoxic T-lymphocyte associated protein 4 (CTLA-4), have been found to be highly expressed in AML patients, and correlate with unfavourable prognosis (Radwan, et al. 2020). DNAM-1, a receptor on T-cells that binds to ligands CD115 and CD112 on cancer cells to regulate cytotoxic activity, is expressed at low levels on T-cells in AML patients. Conversely, its competitor, the inhibitory receptor T-cell immunoglobulin and ITIM domain (TIGIT), is

upregulated. The interaction between TIGIT and CD115/CD112 is theorised to be a mechanism of immune evasion, supported by clinical studies linking high expression of CD115 and CD112 with poor outcomes (Wang, F., et al. 2022). Although immune checkpoint inhibitors are effective in solid tumours, clinical trials are ongoing in haematological malignancies such as AML, and as such are not currently approved for use.

1.4.3 AML blasts direct tumour-associated macrophages to immunosuppressive M2 phenotype

Macrophages within the tumour microenvironment are known as tumour associated macrophages (TAMs). AML blasts cause TAMs to change from the tumour resistant M1 phenotype to the immunosuppressive M2 phenotype. M2 phenotype TAMs release immunosuppressive cytokines such as TGF-β and IL-10, which promote leukaemic cell proliferation, inhibit T-cell proliferation, and repress inflammation signalling while inducing tissue repair and angiogenesis (Mantovani, et al. 2002, Al-Matary, et al. 2016, Mussai, et al. 2013). In AML patients, higher levels of M2-type macrophages have been reported in the bone marrow and spleen compared to healthy donors, suggesting their contribution to AML progression (Al-Matary, et al. 2016). Furthermore, AML leukemic stem cells (AML LSCs) highly express CD47 on their surface, which binds to Signal Regulatory Protein Alpha (SIRPα) on macrophages. This interaction inhibits M1 macrophages from carrying out phagocytosis, further suppressing cancer immunity (Zhang, Wenting, et al. 2020).

1.4.4 AML sabotages T-cell activity by preventing the formation of functional immune synapses

A study found AML patients had an increase in number of T-cells in their peripheral blood when compared to age matched healthy donors. However, these T-cells formed faulty immune synapses with AML blasts. Gene expression profiling of those T-cells revealed differential expression of genes involved in actin cytoskeleton creation compared to the healthy group. While the T-cells can form cell conjugates, the required downstream signalling to form the immunological synapse are inhibited, preventing T-cell effector function (Le Dieu, et al. 2009).

1.4.5 AML blasts release soluble factors which polarise T-cells to a Treg phenotype

In addition to preventing effector function, AML blasts also dysregulate cytokine signalling to favour T-regulatory (Treg) cells over T-effector cells (Curti, et al. 2007). AML blasts release soluble immune inhibitory factors such as IL-10, IL-35, transforming growth factor-beta (TGF-β), and indoleamine 2,3-dioxygenase 1 (IDO1). IDO1 is discussed in more detail in Chapter 3. Through

various mechanisms these factors encourage T-cells to polarise towards the T-reg phenotype, creating an environment tolerant of the AML blasts and enabling disease progression (Platten, et al. 2015, Folgiero, et al. 2014, Locafaro, et al. 2014, Cools, et al. 2008, Walker, M. R., et al. 2003). AML patients are reported to have high numbers Tregs, with higher numbers correlating to poor outcomes. Immunosuppressive Treg phenotype has been observed to be stronger the closer in proximity to the bone marrow of the AML niche (Shenghui, et al. 2011). The disruption to cytokine profiles by AML ultimately decreases inflammation signalling by downregulating IL-15 and (Interferon gamma) IFNG, further pushing T-cells away from effector phenotypes.

1.4.6 AML deregulate cellular metabolism to fuel growth while suppressing the immune response.

Arginase II (ARG2) is a protein that catalyses the hydrolysis of the amino acid arginine into ornithine and urea. Increased ARG2 expression has been observed in the plasma of AML patients, where it inhibits T-cell proliferation and polarises macrophages to the M2 phenotype. Additionally, ARG2, in combination with inducible nitric oxide synthase (iNOS), decreases the NK cell population in the AML blast environment (Jacamo, et al. 2017). Furthermore, this group found the increase in the metabolism of arginine also reduced proliferation of hematopoietic progenitor cells (Mussai, et al. 2013). The metabolism of fatty acids and lipolysis has emerged in recent years as a mechanism AML uses to gain growth advantage. Adipocytes are common in the stroma of bone marrow, and their population increases with age. As AML patients are generally in the older population, this cell type Is abundant in the most vulnerable AML populations (Justesen, et al. 2001). In elderly patients, it has been found that AML remodels the BM niche to promote cell survival through lipolysis of adipocytes. AML uses fatty acid oxidation (FAO) to make acetyl-CoA from fatty acids provided by its adipocyte heavy environment, which then feeds into the tricarboxylic acid cycle (TCA cycle) and to create additional ATP in the oxidatively stressed environment of the AML BM (Beloribi-Djefaflia, Vasseur and Guillaumond 2016, Tabe, Konopleva and Andreeff 2020).

1.4.7 AML evades NK cells by sabotaging activating receptor mechanisms.

NK cells express receptor NKG2D on their surface, which when bound to by its ligand (NKG2DL) and other co-stimulatory factors such as MICA, ULBP1/2/3, activate NK cells. However, in AML cells lines, high methylation of the promoter for NKG2DL silences its expression on AML cells, allowing them to escape NK cells recognition (Baragaño Raneros, et al. 2015). Additionally, AML blasts release a soluble version of NKG2DL which decreases expression of NKG2D on nearby NK cells, inhibiting cytotoxicity (Tettamanti, et al. 2022). Along the same lines, NK cells have another activating receptor called DNAM-1 which activates cytotoxic activity through granulation. AML

blasts upregulate DNAM-1 ligands CD112 and CD155 on their surfaces, which in turn downregulates DNAM-1 expression on NK cells. As with T-cells, AML blasts upregulate expression of TIGIT on NK cells, inhibiting activity and IFNG secretion, this high expression has been correlated with reduced NK cell population in BM of poor prognosis SCT (Stem Cell Transplantation) treated AML patients (Hattori, et al. 2019).

1.4.8 T-cell tolerance through myeloid suppressor cells and tumour associated macrophages

Myeloid derived suppressor cells (MDSCs) are immunosuppressive cells that induce T-cell tolerance through expression of PD-L1, IDO1, Arginase II, ROS production and immunosuppressive cytokines such as TGFβ and IL-10 (Yang, Y., et al. 2020). MDSCs numbers are elevated in the peripheral blood and bone marrow of AML patients when compared to healthy controls (Pyzer, et al. 2017). MUC-1 is an oncoprotein which is released by AML blasts in extracellular vesicles (EVs). The MUC-1 containing EVs are absorbed by myeloid progenitor cells, impairing differentiation, leading to expansion of MDSCs (Groth, et al. 2019, Pyzer, et al. 2017).

AML is a complex disease which utilises multiple mechanisms to promote self-renewal and evade destruction by the immune system. Complexity in diagnosing and treating patients comes from the many mutations that can occur in AML resulting in different phenotypes. Classification of AML has changed over the years from morphological characterisation, to incorporate molecular indicators such as chromosomal abnormalities and mutations.

1.5 AML subtypes

The two systems used to classify AML are the French American-British (FAB) classification and the World Health Organisation (WHO) classification. The FAB classification was developed in the 1970s and assigns subtype to AML based on the cell of origin and its maturity. FAB subtypes are identified based on the morphology of leukaemia cells when observed with a microscope. These subtypes are summarised in the appendix (7.1), along with names, origin cells and an example picture of their appearance. This method was developed in the 1980's and has been since replaced with genetic and cytogenetic markers by the world health organisation (WHO) and the European LeukemiaNet (ELN) panel.

The WHO classification system builds further on the FAB system and is regularly updated with consideration of genetics and other known factors that contribute towards prognosis. This is summarised in Table 3.

Table 1:The categories of AML based on WHO classification criteria (American Cancer Society 2018). Information gathered from elsewhere is referenced in the table.

Category name	Details
AML with certain genetic abnormalities (gene or chromosome changes)	AML: [t(8;21)] AML: [t(16;16) or inv(16)] APL with the PML-RARA fusion gene AML: [t(9;11)] AML: [t(6:9)] AML: [t(3;3) or inv(3)] AML (megakaryoblastic): [t(1:22)] AML with the BCR-ABL1 (BCR-ABL) fusion gene AML with mutated NPM1 gene AML with dominant biallelic mutations of the CEBPA gene AML with mutated RUNX1 gene
AML with myelodysplasia-related changes	A diagnosis requires more than or equal to 20% of bone marrow cells are blasts and additionally either a history of myelodysplastic syndrome (MDS), MDS cytogenetic abnormalities or multi lineage dysplasia (Weinberg, et al. 2009).
AML related to previous chemotherapy or radiation	This is a therapy related AML that occurs following treatment with chemotherapy or radiation therapy for a prior malignancy (Kayser, et al. 2011).
AML not otherwise specified. (AML that does not clearly fit into the other groups and correlated with FAB classification)	AML with minimal differentiation (FAB M0) AML without maturation (FAB M1) AML with maturation (FAB M2) Acute myelomonocytic leukaemia (FAB M4) Acute monoblastic/monocytic leukaemia (FAB M5) Pure erythroid leukaemia (FAB M6) Acute megakaryoblastic leukaemia (FAB M7) Acute basophilic leukaemia Acute panmyelosis with fibrosis
Myeloid sarcoma	Diagnosis for when there is a proliferation of myeloid lineage blasts at extramedullary sites which disrupts the structure of the tissue it is found in. Clinical representations are variable, it has been found in soft tissues, lymph nodes and bones. It has also been called 'Chloroma' due to a distinctive green colour caused by high expression of myeloperoxidase (Solh, et al. 2016, Avni and Koren-Michowitz 2011).
Myeloid proliferations related to Down syndrome	The group involved a spectrum of MDS which evolve in AML in children with down syndrome (Trisomy 21). A subset of new-borns with down syndrome are born with erythro-megakaryocytic myeloproliferative disorder which can develop into acute megakaryoblast leukaemia. It is also characterised by GATA-1 mutation as well as trisomy 21 and additional genetic events (Cantor 2015).
Mixed phenotype acute leukaemia (MPALs)	A rare set of leukaemia that present with mixed lineage that express myeloid and lymphoid antigens. It is characterised by therapy resistance and adverse cytogenetics (Sharma, et al. 2017, Wolach and Stone 2015).

1.5.1 Molecular markers for risk stratification of AML

A combination of morphology, cytogenetics and genetic mutations can be used to identify patient disease. Correct identification of disease is essential for patient prognosis and assigning treatments. Molecular markers are regularly reviewed for their impact on clinical risk, Table 21 shows the recent categorisation of markers determined by the European LeukemiaNet (ELN) panel (Dohner, et al. 2017). Molecular cytogenetics spans techniques that study variations in chromosome structure and function and genome variation. Some standard cytogenetic techniques to study these features include but are not limited to:

- Karyotyping- examining an individual's chromosome complement to reveal missing, broken and rearranged chromosomes.
- Chromosome banding – different aspects of the chromosome can be stained for visualisation, for example T-banding stains telomeres and C-banding for centromeres.
- Fluorescence activated cell sorting (FACS) -sorts and separates chromosomes based on the intensity and size of fluorescent signals (Montazerinezhad, Emamjomeh and Hajieghrari 2020).

There are other approaches which report more detailed differences in chromosomes and genes such as PCR based methods, in-situ hybridisation methods and comparative genomic vectors methods (Groth, et al. 2019, Pyzer, et al. 2017).

Table 2: Table of molecular markers for genetic risk classification of AML as determined by ELN (Döhner, et al. 2022).

Prognostic	Genetic abnormality
Favourable	-t(8;21)(q22;q22.1)/*RUNX1::RUNX1T1* -inv(16)(p13.1q22) or t(16;16)(p13.1;q22)/ *CBFB::MYH11* -Mutated *NPM1* without *FLT3*-ITD -bZIP in-frame mutated *CEBPA*
Intermediate	-Mutated NPM1 with FLT3-ITD -Wild-type NPM1 with FLT3-ITD (without adverse-risk genetic lesions) -t(9;11)(p21.3;q23.3)/MLLT3::KMT2A†, -Cytogenetic and/or molecular abnormalities not classified as favourable or adverse
Unfavourable	-t(6;9)(p23.3;q34.1)/DEK::NUP214 -t(v;11q23.3)/KMT2A-rearranged -t(9;22)(q34.1;q11.2)/BCR::ABL1 -t(8;16)(p11.2;p13.3)/KAT6A::CREBBP -inv(3)(q21.3q26.2) or t(3;3)(q21.3;q26.2)/ GATA2, MECOM(EVI1) -t(3q26.2;v)/MECOM(EVI1)-rearranged --5 or del(5q); −7; −17/abn(17p) -Complex karyotype monosomal karyotype -Mutated ASXL1, BCOR, EZH2, RUNX1, SF3B1, SRSF2, STAG2, U2AF1, and/or ZRSR2 -Mutated TP53a

1.5.2 Genetic heterogeneity in AML

As indicated by the molecular markers used for risk stratification in Table 2, genetic heterogeneity takes many forms in AML from chromosomal abnormalities to molecular mutations. Risk factor assessment based on cytogenetics must also consider the complex interactions that can take place when the mutations co-occur, and how high and low allele ratios (*NPM1* and *FLT3-ITD* as an example) impact on disease. Mutations can be split into functional categories dependent on their role in AML development. These mutations occur in receptor tyrosine kinase (RTK) family members, transcription factors (TF), nucleophosmin, tumour suppressors and the spliceosome complex (DiNardo and Cortes 2016). Emphasis has been placed on the role of RTK family mutations and TF mutations which are thought to be key players in malignant transformation but are not necessarily always present. RTK mutations such as *FLT3*, *KIT*, *VRAS, KRAS*, *PTPN11*, *NF1* and *NRAS*, are present in approximately 66% of cases and promote the proliferation and survival of cells (DiNardo and Cortes 2016). Mutations in TF such as *RUNX1*, *CEPBA, GATA2* and *RARA*, occur in between 5 and 25% of cases dependent on TF and prevent cell differentiation, aiding in building an immature progenitor population (DiNardo and Cortes 2016, Li, Sheng, Mason and Melnick 2016a). Additionally, concurrent mutations in RTK and TFs interact synergistically to produce a new epigenetic and transcriptional profile, different to any produced by the individual categories, further complicating evaluation (Shih, A. H., et al. 2015).

1.5.3 Epigenetic modifications contribute to AML pathogenesis and progression

Epigenetics describes changes to chromatin structure through modification of histones by methylation, acetylation, phosphorylation and more, which change how genes are expressed (Itzykson, Kosmider and Fenaux 2013). It is thought that epigenetic mutations are key to inciting leukaemogenesis and promoting clonal expansion but cannot initiate transformation without the aid of additional mutational events in the population (DiNardo and Cortes 2016, Li, Sheng, Mason and Melnick 2016a). Heterogeneity is especially prevalent in AML epigenetic profiles, with phosphorylation, acetylation and methylation all recorded to contribute to pathogenesis (Dhall, et al. 2019, Seo, et al. 2022, Shih, A. H., et al. 2015). Somatic mutation of epigenetic regulators such as DNA methyltransferases (DNMT), Isocitrate dehydrogenase (IDH) and Ten-eleven translocation methylcytosine dioxygenases (TET) are identified in over half of AML cases. A variety of cytosine methylation landscapes have been recorded (Figueroa, Abdel-Wahab, et al. 2010). Mutations in enzymes DNMT3A, IDH1/2 and TET2 disturb haematopoiesis preventing differentiation into specialised blood cells and is key in initiating and progressing AML (Yang, X., Wong and Ng 2019a).

1.5.3.1 Acetylation status disrupting mutations

One of the most frequent karyotypic abnormalities in AML is the t(8;21)(q22;q22) translocation which fuses *RUNX1* to *RUNX1T1* (Goldman, et al. 2019). The Leukaemogenesis ability of the fusion product AML1-ETO, may be reliant on lysine acetylation. Murine models with lysine acetyltransferase p300 knockdown display decreased AML1-ETO acetylation, which correlates with an increase in mouse median survival. However, as p300 decreases acetylation on a broad range of targets, it is unknown if the effect on survival is linked specifically to AML1-ETO (Link, et al. 2016, Wang, L., et al. 2011).

1.5.3.2 Phosphorylation in AML is associated with drug resistance

Phosphorylation is the addition of a phosphate group to a protein at the serine (Ser), threonine (Thr), or tyrosine (Tyr) residues by a protein kinase. The addition and removal of phosphate groups (by phosphatases) regulate the function of the protein (Bachegowda, et al. 2016). The PI3K/AKT signalling pathway is activated by phosphorylation and promotes cell growth and proliferation while inhibiting apoptosis, and so is tightly regulated. Research has demonstrated that a heterozygous deletion mutation in PTEN, a tumour suppressor gene, leads to increased AKT phosphorylation in AML cell lines. This overactivation of the PI3K/AKT pathway results in uncontrolled cell growth. Additionally, AML patients with low PTEN levels experience higher relapse rates within one year compared to those with normal PTEN levels (Chen, Ping, et al. 2016). Not only can P-AKT phosphorylate direct downstream targets involved in drug resistance, but the excessive signalling increases JNK-p38 MAPK pathways activity causing c-jun dependent resistance (Roszak, Smok-Pieniążek and Stępnik 2017). Furthermore, overactivation of PI3L/AKT upregulates P-gp to transport drugs out of cell nuclei (Chen, Ping, et al. 2016), further contributing to chemoresistance.

1.5.3.3 Mutations in methylation proteins contribute to AML pathogenesis

DNA methylation is an epigenetic modification where a methyl group is added to the 5th position of the cytosine ring in DNA. It acts as an off switch for genes, and regulates many developmental processes, including haematopoiesis, where haematopoietic stem cells differentiate into specific blood cells. DNA methylation profiles vary considerably between AML cases with patients presenting with hypermethylation, hypomethylation and intermediate methylation profiles (Figueroa, Lugthart, et al. 2010). Studies have further investigated AML methylation to understand the biological and functional implications of methylation status (Figueroa, Lugthart, et al. 2010, Li, Sheng, Mason and Melnick 2016a). Mutations of genes involved in methylation, such as *DNMT3*, *TET2* and *IDH1* prevent function and produce hypo and hyper methylation profiles.

1.5.3.4   DNMT3A

DNMT3A is a *de novo* methylation enzyme and catalyses the addition of new methyl groups on DNA molecules. This activity regulates HSC differentiation and self-renewal by epigenetic methylation. Mutation in *DNMT3A* disrupt and impair this function, preventing DNA methylation and causing a hypo methylated environment. In turn HSC proliferation increases and differentiation into specialised blood cells is impaired. Mutation of *DNMT3A* commonly occurs alongside *NPM1*, *FLT3-ITD*, and *IDH1* mutations and correlates with relapse and chemotherapy resistance (Mayle, et al. 2015).

1.5.3.5   IDH1/2

Mutations in *IDH1* and *IDH2* active sites at the R132 locus of *IDH1* and the R140 or R172 locus of *IDH2* cause a loss-of-function. In the Kreb's cycle isocitrate is normally converted to alpha-ketoglutarate (αKG), however, mutations in IDH1/2 reverse the reaction converting αKG into 2-hydroxyglutarate (2-HG) (Ward, et al. 2010). 2HG then competitively inhibits αKF dependent enzymes involved in regulating chromatin structure and DNA repair systems including TET enzymes and lysin demethylase (Moran-Crusio, et al. 2011, Figueroa, Abdel-Wahab, et al. 2010). This causes a hypermethylated phenotype, silencing many genes including tumour suppressor genes, inducing increased DNA damage and impaired myeloid differentiation. The prognosis of AML with IDH mutations is poor with IDH1 mutations, but favourable with IDH2 mutations, when patients are treated with standard intensive chemotherapy (Aslanyan, et al. 2014, Kroeze, et al. 2014, Ahn, et al. 2017).

1.5.3.6   TET2

TET2 along with α-KG as a cofactor, catalyses the conversion of 5-methylcytosine (5-mC) to 5-hydroxymethylcytosine (5-hmC) which is a crucial component needed for DNA demethylation. Mutations that disrupt TET2 activity therefore impair DNA demethylation, ultimately leading to down-regulation of genes that promote myeloid differentiation and genes that inhibit self-renewal in HSCs (Chan and Majeti 2013). TET mutations can lead to hypermethylated DNA signatures which have been linked to poorer outcome (Shlush, et al. 2014). Another route for AML to suppress immune cells in its local environment is through repurposing IFNG signalling.

1.6 The role of IFNG in general immunity

Interferon Gamma (IFNG) is a cytokine involved in both the innate and adaptive immune responses to viral and bacterial infections. This cytokine is primarily produced during the innate immune response by natural killer cells (NK) and natural killer T-cells (NKT cells). It is also secreted after the development of adaptive immunity by CD4+ T Helper 1 (TH1) cells and CD8+ Cytotoxic T lymphocytes (CTL) effector cells (Rybka, Stephanou and Townsend 2009). IFNG

signals by binding to the IFNG receptor, which is expressed on various cell types. This cytokine plays a central role in numerous biological processes, most notably in coordinating the immune response. One of its earliest discovered functions is the upregulation of major histocompatibility complex (MHC) molecules (Amaldi, et al. 1989) enabling increased antigen presentation (Minn 2015) and improved immune response.

IFNG acts as a key regulator of the immune system, influencing the activity of many immune cells at the innate and adaptive immunity level (Figure 3). IFNG induces B-cells to favour producing opsonising and complement fixing antibodies over IL-4 dependent isotypes such as IgE thereby promoting inflammation and enabling phagocytes to clear more microbes and damaged cells (Gonzales-van Horn and Farrar 2015). Working in tandem with the isotype switch, IFNG enhances macrophage activity through 1) upregulating expression of FcγRI, enabling a larger uptake of pathogens by the macrophage, 2) stimulating them to synthesise more reactive oxygen species (ROS) and nitric oxide (NO) (Robinson, C. M., et al. 2010). These molecules are produced within lysosomes of the macrophage which fuse with phagosomes that contain microbes, thereby destroying the microbes (Thakur, Mikkelsen and Jungersen 2019). IFNG secretion causes a paradigm shift from innate to adaptive immune response favouring cell mediated immunity; inducing naïve CD4$^+$ T-cells to favour differentiation into T$_H$1 effector cells over T$_H$2 helper cells. This promotes specific cytotoxic immunity through upregulation of both MHC classes as well as antigen processing, presentation, and costimulatory molecules, increasing pathogen peptide presentation and T-cell – APC interactions. Furthermore, a positive feedback loop is formed by T$_H$1 cells producing more IFNG and IL-12 while inhibiting IL-4 secreting T$_H$2 cells, creating an environment which favours the T$_H$1 effector phenotype (Schroder, et al. 2004).
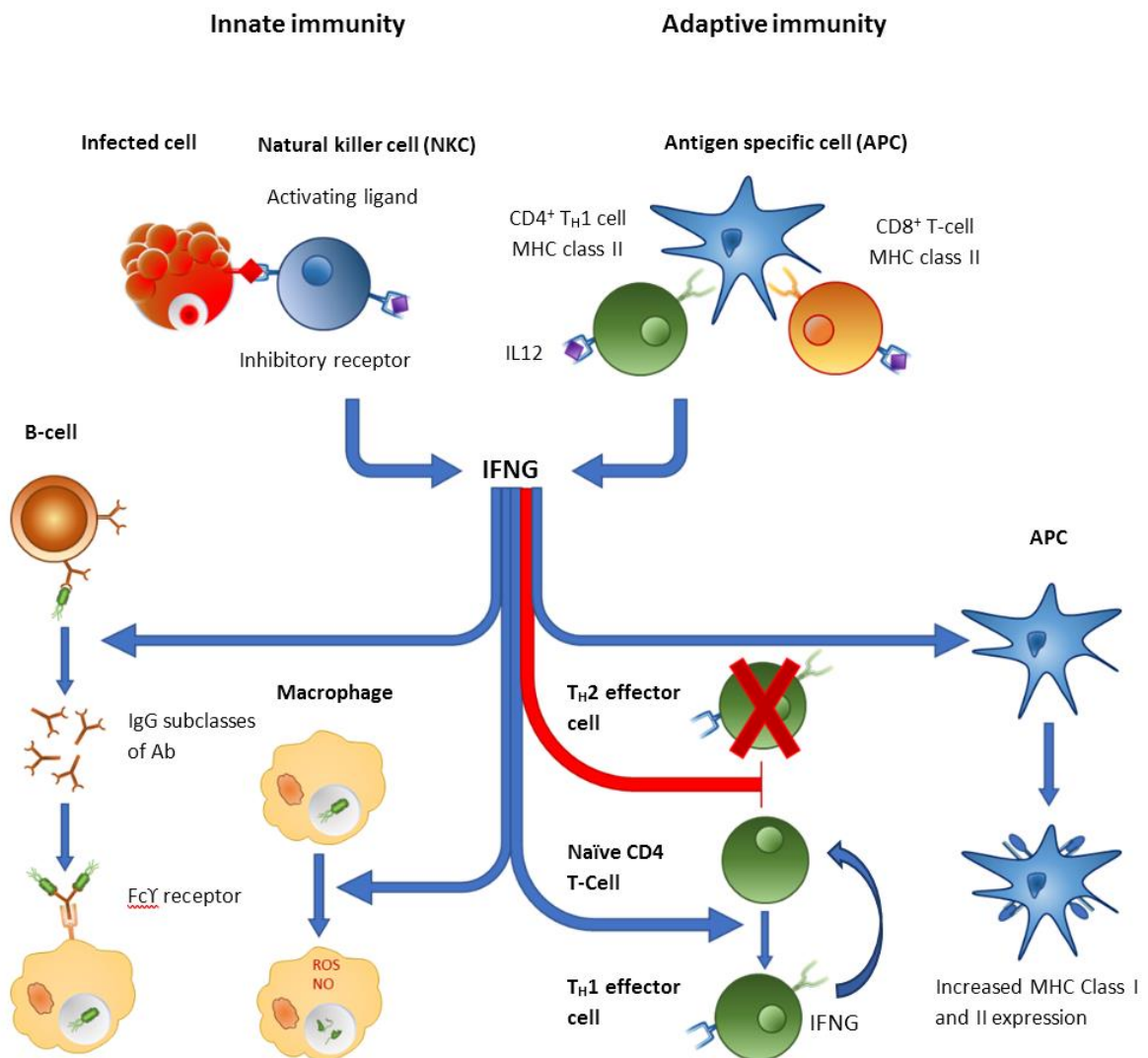
Figure 3: The roles of IFNG in the immune response. IFNG is produced first by NKC and NK T-cells during the innate response and CD4+ / CD8+ T-cells once antigen specific immunity has developed. IFNG is involved in many immune response mechanisms, some illustrated above; left to right. Outer left: IFNG induces B-cells to produce IgG Ab subclasses which protect against infectious pathogens. Inner left: IFNG activates macrophages to increase lysosomal activity. Inner right: IFNG promotes differentiation of naïve CD4+ T-cells into TH1 cells over TH2 effector cells, TH1 cell also produce IFNG resulting in a positive feedback loop. Outer right: IFNG induces the expression of MHC class I and II expression on both APC and normal cells, allowing more expression of foreign antigen and bolstering immune response.

1.7 IFNG; a key cytokine for anti-cancer activity

IFNG is also involved in immune surveillance, inflammation and known for general anti-tumour activity (Zaidi and Merlino 2011). IFNG inhibits proliferation of cancer cells (Kotredes and Gamero 2013), induces dose dependent apoptosis (Cheon, Yang and Stark 2011), and the blockage of angiogenesis resulting in tumour starvation (Sidky and Borden 1987).

In AML, it has been shown that the reduced ability of CD4+ T-cells to produce IFNG at the time of diagnosis can be restored by the time of relapse following stem cell transplantation (Schnorfeil, et al. 2015, Lamble and Lind 2018). Kornblau et al also observed reduced IFNG production by CD4+ cells, identifying lower IFNG levels in the serum of untreated AML patients compared to healthy controls (Kornblau, et al. 2010). AML appears to downregulate IFNG expression to prevent a T-cell effector response and shift towards a Treg type. However, in some AML patients, IFNG-related genes are highly expressed, which correlates with poor overall survival (Corradi, et al. 2022).

1.7.1 IFNG signalling upregulates immunosuppressive factors in AML

Although IFNG has direct anti-tumour activity, and in a functional environment, can prevent cancers including AML from taking root, it can also enable cancer growth (Ribas 2015). Adaptive immune resistance (AIR) is the process by which a cancer responds to immune surveillance by changing its phenotype to evade the immune system. AML can use IFNG to upregulate escape mechanisms, and cause immunosuppression to promote immune escape. Some of the methods used by AML to evade immune cells were outlined in section 1.4.

IFNG upregulates numerous molecules in AML, including but not limited to PD-L1, IDO1, non-classical HLA's, and BST2. These molecules contribute to remodelling the bone marrow niche and an immunosuppressive environment by inhibiting cytotoxic immune cell activity, polarising T-cells towards tolerogenic phenotypes, inhibiting memory cell survival, reducing immune cell protein synthesis, growth and HSC niche re-localisation.

In AML blasts, IFNG upregulates PD-L1 levels by various pathways including the STAT1/3 and MAPK pathways. PD-L1 prevents AML blasts from being destroyed by cytotoxic T-cells by transmitting an inhibitory signal (Berthon, et al. 2010, Moshofsky, et al. 2019). IFNG released from various sources including NK cells has been found to upregulate IDO1 expression on AML blasts. In short IDO1 catalyses the breakdown of the essential amino acid tryptophan (trp), which creates Kynurenine (Kyn). The depletion of TRP is sensed by CD4+ T-cells, which respond by reducing protein synthesis to inhibit cell growth. Furthermore, Kyn binds to the aryl hydrocarbon receptor on T-cells and dendritic cells (DCs), inducing signalling that shifts naïve CD4+ T-cells to

differentiate into Treg cells, and DCs to favour a tolerogenic phenotype (Folgiero, et al. 2014, Platten, et al. 2015). AML patients with high IFNG expression have been associated with poorer overall survival compared to those with lower IFNG expression. IFNG release by AML correlates with elevated Treg levels in the bone marrow. Furthermore, transcriptomic analysis of mesenchymal stromal cells (MSCs) co-cultured with IFNG-expressing cells showed an induction of an IFNG-dependent program of Treg-related genes in the MSCs. This indicates that AML blasts can use IFNG to upregulate immunosuppressive factors in themselves, such as IDO1 and PDL1, and in the local bone marrow niche by activating Treg induction programs in MSCs (Folgiero, et al. 2014, Platten, et al. 2015).

IFNG further contributes to a tolerant environment by upregulating the expression of nonclassical HLA such as HLA-G and HLA-E in AML (Mizuno, et al. 2000, Nguyen, et al. 2009). Expression of HLA-G promotes maternal tolerance of the foetus in pregnancy; however, AML and other cancers utilise HLA-G expression on their surface to induce immune tolerance by binding to inhibitory receptors on immune cells such as ILT-2, ILT-4 and KIR2DL4 (Kren, et al. 2010, Gallegos, et al. 2016). HLA-E interacts with NK and T-cells to regulate their activity, importantly, HLA-E can bind to inhibitory receptor NKG2A on NK cells. In AML IFNG upregulation of HLA-E on AML blasts has been found to use this mechanism to inhibit NK cells cytolysis activity in patients posts haplo-mismatched Haematopoietic stem cell transplant (HSCT) (Nguyen, et al. 2009).

HSCs usually reside in the bone marrow in an inactive quiescent state, but can be activated in response to injury or to restore blood cells through proliferation and differentiation (Loeffler and Schroeder 2021). Long term IFNG signalling due to chronic infection has been shown to activate HSC, causing IFNG dependent reduction in HSC in the bone marrow (Bogeska, et al. 2022). The migration away from the bone marrow is hypothesised to be in part due to IFNG upregulation of BST-2 on HSCs, and its binding to E-selectin activating HSCs. Ultimately this results in IFNG dependent localisation of HSC from a quiescent and inactive niche to an E-selectin positive and active niche, contributing towards AML progression. Increased BST2 expression has also been associated with poor overall survival in AML patients (Matatall, et al. 2018, Florez, et al. 2020).

In summary AML uses IFNG to upregulate AML blast proliferation and upregulates a host of immunosuppressive molecules to directly inhibit immune cells. Additionally, it induces immune cells in its environment to shift to tolerant phenotypes, both through AML blast direct interaction, and induction of immunosuppressive gene programmes in surrounding cells such as MSC. The hijacking of inflammation signalling which usually would help to control cancer progression, adds layers to how effective treatments can be via standard chemotherapy, and highlights the

importance in understanding the AML immune environment when developing new treatment and prognostic strategies for AML.

1.8 Treatment strategies for AML

Therapeutic strategies for treating AML patients have made limited progress over the past 30 years, with intensive chemotherapy remaining the predominant treatment method (Döhner, Weisdorf and Bloomfield 2015). AML is cured in 35 to 40% of adult patients, 70% of child patients and, in only 5 to 15% of elderly patients (Döhner, Weisdorf and Bloomfield 2015). Elderly patients are most at risk as they are unable to receive intensive chemotherapy due to endangering side effects. Therefore, this category of patients can only receive low intensity therapies or symptom management (Döhner, Weisdorf and Bloomfield 2015). AML accounts for approximately 25% of childhood leukaemia (Lonetti, Pession and Masetti 2019). For children under the age of 15, overall survival rates have plateaued at approximately 70% with intensive chemotherapy. In general, younger patients respond better and have higher survival rates than older patients. However, the median age of diagnosis is 68, in the age bracket with the lowest survivability. Therefore, prioritising the development of alternative strategies for high-risk AML in older patients is high priority (Howlader N, Noone AM, Krapcho M, Miller D, Brest A, Yu M, Ruhl J, Tatalovich Z, Mariotto A, Lewis DR, Chen HS, Feuer EJ, Cronin KA (eds). 2018). Treatment for AML takes many forms, in addition to the classical chemotherapy regiment, there are also molecularly targeted therapies, immune based treatments and epigenetic strategies.

1.8.1 7+3 induction chemotherapy

The front-line treatment for AML is the "7+3" chemotherapy regimen used to induce remission. This regimen is comprised of 7 days of dosing with Cytarabine, followed by 3 days treatment with an anthracycline antibiotic: Daunorubicin, Idarubicin or Mitoxantrone, depending on the patients age and general health.

Cytarabine is a prodrug which is an analogue of the nucleoside cytidine; a pyrimidine that is incorporated into nucleic acids (Rechkoblit, et al. 2018). Once Cytarabine is in the cell, it is transformed into a triphosphate form which competes with cytidine to be incorporated into DNA. The structure of cytarabine includes an arabinose sugar that causes steric hindering which prevents DNA extension, inhibiting the S-phase of cell division (Rechkoblit, et al. 2018, National Center for Biotechnology Information. 2020a). Additionally, Cytarabine also inhibits DNA polymerase, causing a direct decrease in DNA replication and repair (Blair 2018). For this reason, it is most effective in fast cycling cells, where the cell cycle has become dysregulated, which is a hallmark of cancer.

Anthracyclines are a class of cytotoxic drugs able to enter the cell membrane to intercalate between DNA base pairs and interact with topoisomerase II. Examples of anthracyclines used in combination with cytarabine include Doxorubicin, Daunorubicin, and Idarubicin.

Anthracyclines work by inhibiting mitosis and cell division through intercalation between DNA base pairs. This causes the DNA helix to uncoil, inhibits topoisomerase II activity, and leads to both single and double strand breaks in the DNA, ultimately preventing DNA synthesis. Additionally, daunorubicin can inhibit DNA polymerase, which disrupts gene expression and causes free radical-induced DNA damage, resulting in cell death (Saleem and Kasi 2021, Blair 2018, National Center for Biotechnology Information. 2020b). Idarubicin is an analogue of daunorubicin missing the methoxy group at position 4 which increases its lipophilicity compared to Daunorubicin. This increases the rate of its uptake into cells and enables better DNA binding, improving on the cytotoxicity of Doxorubicin and Daunorubicin (Assi, et al. 2016). In 2017, the FDA approved CPX-351, a 100 nm bilamellar liposomal formulation of cytarabine and daunorubicin, for the treatment of high-risk AML (Assi, et al. 2016). This formulation maintains a 5:1 drug ratio with higher bioavailability, resulting in more effective drug exposure to leukemic cells compared to standard therapy. CPX-351 was found to be less toxic than standard delivery methods and significantly improved median overall survival (OS) rates in high-risk elderly patients from 2 months to 5.1 months (Assi, et al. 2016, Lin, T. L., et al. 2019).

1.8.1.1 Allogenic hematopoietic stem cell transplant (HSCT)

Allogenic HSCT provide the best chance of preventing AML reoccurrence, but also carry a greater risk of treatment-related morbidity and mortality (TRM). HSCT may be administered in first or second remission, depending on if the patients AML is assessed as favourable or unfavourable (Assi, et al. 2016, Lin, T. L., et al. 2019, Koreth, et al. 2009, Cornelissen, et al. 2007). The incidence of TRM occurring in elderly patients after HSCT has been reduced in recent years, using reduced intensity conditioning (RIC). One study reported that patients over the age of 45 survived longer when treated with RIC transplantation than those treated with chemotherapy (Russell, et al. 2015). Two further studies evidenced that RIC transplantation can be favourable in elderly patients compared to conventional myeloablative conditioning (MAC) transplantation, which has primarily found to be beneficial to younger patients (Passweg, et al. 2015, Cornelissen, et al. 2015).

1.8.2 Molecularly targeted treatments

Targeted therapies are designed to block the growth and spread of cancer by interfering with specific proteins, pathways, or genes that are involved in the growth, progression, and spread of

cancer cells. By targeting the specific molecules or pathways that allow cancer to grow, targeted therapies can be more effective than traditional chemotherapy, which often kills healthy cells along with cancer cells.

1.8.2.1 FLT3 targeting therapies

Some therapies target specific mutations, such as Midostaurin and Gilteritinib for FLT-3 mutated AML. Mutations in *FLT3* are the most common type of mutation seen in AML, with it appearing in approximately one third of patients. There are two categories of mutation in FLT3, internal tandem duplicate (*FLT3*-ITD) and point mutation in the tyrosine kinase domain FLT3 (FMS-like tyrosine kinase 3) is a receptor, which when bound to by the FLT3 ligand activates a signalling cascade which results in inhibited apoptosis and differentiation. Both types of *FLT3* mutation result in constitutive activation of the receptor, and therefore inhibition of apoptosis and differentiation occur, promoting cancer survival (Kennedy and Smith 2020). Midostaurin is an oncogenic FLT-3 inhibitor used to treat newly diagnosed patients, in combination with classic cytarabine and daunorubicin treatment (Levis 2017). Gilteritinib is a dual FLT3-AXL tyrosine kinase inhibitor which is approved for the treatment of relapsed and refractory FLT3^{mut} AML (Lai, C., Doucette and Norsworthy 2019). There is currently an ongoing trial (Trial number: NCT03836209) to compare the efficacy of Midostaurin to Gilteritinib combined with induction and consolidation chemotherapy (Luger 2020).

1.8.2.2 IDH1 targeting therapies

Another identified drug target is isocitrate dehydrogenase type 1/2 (IDH1/2). Mutations in IDH1/2 cause overproduction of 2-hydroxyglutarate (2HG) which inhibits DNA and protein demethylation. AML cases with IDH1/2^{mut} display lack of differentiation ability, inhibiting IDH mutants can restore differentiation ability (Schvartzman, et al. 2019). Two drugs which target this molecule are Ivosidenib and Enasidenib mesylate, which are used to treat adult patients with relapsed or refractory AML, who have specific mutations of IDH1 or IDH2 respectively (Kim, E. S. 2017). While Ivodidenib achieves a respectable CR rate of 28.6% alone and 57% in combination with Azacytidine, differentiation syndrome (DS) was observed to occur as a common side effect in 25% of patients (Lai, C., Doucette and Norsworthy 2019).

1.8.2.3 BCL2 targeting therapy

Venetoclax is an alternative therapy available for IDH1^{mut} AML which inhibits B cell lymphoma 2 (BCL-2). BCL-2 suppresses apoptosis by controlling mitochondrial membrane permeability which it achieves by sequestering pro-apoptotic BAX. Venetoclax is approved for use in combination with azacytidine, decitabine or low dose cytarabine for novel AML in adults ≥75 or with comorbidities that prevent intensive chemotherapy (Winer and Stone 2019). Clinical trials

have shown that Venetoclax benefits from combining with either demethylating agents or low dose cytarabine, but currently demethylation is the favoured approach (Lai, C., Doucette and Norsworthy 2019).

1.8.3 Immunotherapy based treatments

Immunotherapy uses the body's immune system to fight cancer. Commonly used types of cancer immunotherapy include Monoclonal antibodies (mAb, Chimeric antigen receptor (CAR) T-cell therapy and vaccines.

1.8.3.1 Monoclonal antibodies (mAbs)

mAbs are manufactured proteins, designed to carry chemotherapy or other treatment and deliver it to cancer cells. At present, Gemtuzumab ozogamicin (Mylotarg) is the only FDA approved monoclonal antibody treatment for AML. Mylotarg consists of two parts. The first is an antibody specific to CD33; which is expressed on leukemic blasts >90% in AML patients, but not on normal haematopoietic stem cells and so will specifically target the malignant cells. The mAb is covalently bonded to the cytotoxic N-acetyl gamma calicheamicin, which will come in close contact with the AML cell and destroy it (Fostvedt, et al. 2019). Additional antibody targets under investigation for AML specific delivery are CD123, CD13, CLL-1 and CD38 (Williams, et al. 2019).

1.8.3.2 Chimeric antigen receptor (CAR) T-cell therapy

CAR T-cell therapy involves modifying a patient's T cells to target cancer cells, expanding them in a lab, and reinfusing them to attack the cancer (Boyiadzis, et al. 2018). While promising in early trials for other leukaemia, its efficacy in AML remains unclear due to the challenge of identifying suitable target antigens, as AML antigens often coincide with normal hematopoietic cells, leading to potential hematotoxicity therapy (Hofmann, et al. 2019). CAR T-cell and bi-specific T-cell therapies have shown some success in preclinical AML studies (Mardiros, Forman and Budde 2015). Although no CAR T-cell treatments for AML are FDA-approved, several clinical trials are ongoing (NCT03766126). However, severe side effects, such as high fevers and very low blood pressure, remain a significant concern (Brudno and Kochenderfer 2016).

1.8.3.3 Vaccines for treatment of AML

Another avenue being explored for the treatment of AML is vaccines. Dendritic cell and peptide vaccines have been tested with the purpose of reinvigorating or priming the immune system against leukaemia cells. Dendritic cells are trained *in vitro* to process and present specific tumour antigens, so that when they are administered, they can successfully prime native T-cells and restore their anti-leukaemia activity (Stanchina, et al. 2020). DCP-001 is an allogeneic DC vaccine, meaning it consists of allogeneic DCs which present preselected tumour specific

antigens. In this case the antigens were WT-1, PRAME, MAGE-A3 and NY-ESO-1 (van de Loosdrecht, et al. 2018). In this phase I study DCP-001 was found to be safe in elderly patients and to generate improved immune responses, with much longer OS in patients where there were no detectable blasts in circulation (van de Loosdrecht, et al. 2018). Other DC based vaccines have seen success with improving induction of tumour associated antigen carrying CD8[+] T cells and improving outcome (Trial number: NCT01686334 and NCT01096602) (Anguille, et al. 2017).

An additional method to target AML is through targeting epigenetic modifications, such as methylation, which is used in AML to inhibit expression of tumour suppressor genes and other anti- cancer molecules.

1.8.4 Epigenetic treatments for AML

Epigenetic dysregulation contributes to the pathogenesis and progression of AML. Due to the reversible nature of epigenetic modifications, they make attractive targets for therapeutics. Those therapies include histone deacetylase inhibitors (HDACi), histone demethylation inhibitors (HDMi) and DNMTi.

1.8.4.1 HDMi

Histone methylation takes place the arginine and lysine amino acids in the histone tail, depending on the residue and location, methylation can activate or repress transcription (Ramadoss, Guo and Wang 2017, Margueron, Trojer and Reinberg 2005). HDM such as lysine demethylase 1 (LSD1) demethylate histone-3 at the K4 residue, downregulating H3K4me3, which regulates gene transcription. In mixed lineage leukaemia (MLL) driven AML, inhibition of LSD1 by drug or genetic deletion, induced differentiation of AML stem cells (Maiques-Diaz, et al. 2018). Other HDMi such as tranylcypromine (TCP), NCD35 and NCD38 have been demonstrated to induce expression of myeloid differentiation genes and inhibit leukaemia growth (Schenk, et al. 2012, Sugino, et al. 2017). HDACi Panobinostat has shown limited anti-leukaemia activity in clinical trials for myeloid malignancy growth (Schenk, et al. 2012), but performs better when combined with HDMi SP2509 in primary leukaemia and cell lines (Fiskus, et al. 2014). Additionally, combination of HDMis GSK2879552 and IMG-7289, in combination with all-trans retinoic acid therapy proved effective *in vitro* (Smitheman, et al. 2019).

1.8.4.2 HDACi

Histone acetylation (HAC) and deacetylation (HDAC) enzymes regulate gene expression by adding and removing acetyl groups at lysine residues in the histone tail. Addition of the acetyl group weakens the interaction between the histone and DNA, opening the DNA up for transcriptions (Gujral, et al. 2020). HDAC inhibitors (HDACi) induce apoptosis and activate DNA

damage response pathways and anti-proliferation pathways in cancers (Insinga, et al. 2005, Nebbioso, et al. 2005, Bali, et al. 2005). First generation HDACi (Vorinostat and Panobinostat) have been approved by the FDA for cutaneous T cell lymphoma and multiple myeloma, while second generation HDACi such as Givinostat have been investigated *in vitro* and in clinical trials for relapsed AML (Finazzi, et al. 2013, Li, Ying, et al. 2016). Valproic acid (VPA) is another HDACi under investigation, which performs hyperacetylation at H3 and H4 histone tails to inhibit activity of class I HDACs. VPA treatment has been found to upregulate genes involved in cell cycle arrest, DNA repair and apoptosis in AML patient blasts (Rücker, et al. 2016). Other HDACi such as omidepsin/depsipeptide, mocetinostat, and entinostat have been rested in clinical trials as monotherapies, but in general work best in combination with other treatments (Gambacorta, et al. 2019).

1.8.4.3 DNMTi

DNMT inhibitors (DNMTi), such as 5-Azacytidine (5AzaC) and decitabine have been found to increase survival in elderly patients who are unsuitable for stem cell transplants or intensive chemotherapy. DNMTi promote degradation of WT DNMT, facilitate DNA demethylation, and induce cytotoxicity. Both 5AzaC and decitabine have been approved by the European Medicines Agency as AML therapeutics in 2008 and 2012 respectively (Al-Ali, Jaekel and Niederwieser 2014).

Both drugs are analogues of the nucleoside cytidine, and upon entry into the cell are activated by successive phosphorylation steps where they are incorporated into DNA and RNA (Diesch, et al. 2016). Decitabine becomes 5-aza-dCTP following modification in the cell, where it then incorporates into DNA and impairs DNA methylation by irreversibly binding to DNMT1, causing it to be degraded. At higher doses it also targets DNMT3A/B (Leonhardt, et al. 1992). Methylation is then lost in subsequent rounds of DNA replication, altering the methylation profile (Liu, K., et al. 2003). Additionally, DNMT-5-aza-dCTP-DNA adducts induce apoptosis by activation of DNA damage response pathways, upregulating and activation of DNA repair proteins CHK1/2 and RAD-51 (Orta, et al. 2014). Low doses of DNMTi are reported to induce DNA hypomethylation of tumour suppressor genes promoters, including CDKN2B (p15INK4B) and E-cadherin, reactivating them. Unlike Decitabine, 5-Azacytidine is processed into 5-aza-CTP and preferentially incorporated into RNA (80-90%), where it inhibits tRNA methylation inhibiting production of functional messenger and transfer RNAs (mRNA, tRNA) (Diesch, et al. 2016). Inhibition of mRNA and tRNA further inhibits production of proteins, triggering apoptosis (Schaefer, et al. 2009).

A pooled analysis of five randomised clinical trials of AML patients showed significantly better overall survival and remission in the DNMTi-treated cohorts than cohorts treated with 7+3 chemotherapy regiment. However, 50% of the AML patients developed resistance and/ or relapsed (Yang, X., Wong and Ng 2019a). DNMTi are often co-administered with other treatments, such as with low dose cytarabine in elderly patients, venetoclax or HDACi such as Vorinostat to synergistically improve survival outcomes (Atalay and Ateşoğlu 2016, Laloi, et al. 2022, Kirschbaum, et al. 2014).

Despite the large array of therapies used and under investigation for AML, all treatments face challenges in overcoming resistance.

1.9 Mechanisms of drug Resistance in AML

Resistance to chemotherapy and other treatments is a common problem in the cancer research field, often leading to relapse and poor prognosis. It is of relevance in AML, as even after more than 30 years, the first line treatment for most patients is an induction chemotherapy combination of cytarabine and an anthracycline (Döhner, Weisdorf and Bloomfield 2015).

There are two types of drug resistance: Primary and acquired. Primary drug resistance describes tumour cells that are fundamentally not sensitive to chemotherapy drugs and resist the first treatment. Acquired resistance refers to tumour cells which are initially sensitive to treatment but develop resistance following induction therapy. Through either route, the resistant clone can proliferate and further mutate to become the dominant clone, preventing the initial therapy being a viable future treatment. Acquired resistance can occur through mutations, which develop via genomic instability and external factors.

Mechanisms of drug resistance in AML include but are not limited to: 1) Drug resistance related proteins and enzymes, 2) genetic alterations, 3) aberrant activation of drug-resistance related pathways (Zhang, Jing, Gu and Chen 2019).

1.9.1 Resistance to chemotherapy drugs through efflux pumps

Resistance achieved through physical relocation of the drug away from the site of action can be performed by p-glycoprotein (P-gp), multidrug resistance related protein 1 (MRP1) and lung resistance protein (LRP) (Megías-Vericat, et al. 2015).

1.9.2 Glutathione S-transferases (GST)

Glutathione transferases are a family of isozymes that catalyse the conjugation of glutathione to xenobiotic substrates for removal from the cell. One of the glutathione S-transferases (GST), GST π is reported to be highly expressed in AML and thought to work synergistically with MRP1 to

promote drug resistance (Hatem, El Banna and Huang 2017). Its primary function is to catalyse the conjugation of GSH to chemical drugs to reduce their cytotoxicity by: 1) facilitating them to be transported out of the cell by efflux pumps and 2) inhibiting drug action by binding GSH to electrophilic sites which would otherwise be used to attack DNA (Hatem, El Banna and Huang 2017, Li, Shuyi, et al. 2017).

1.9.3 MRP1

Multidrug resistance-related protein 1 (MRP1) is a GSH specific transporter that removes GSH and its bound substrates, from the cell. It also confines drugs to perinuclear vesicles, preventing their cytotoxic action, and is associated with poor prognosis (Paprocka, et al. 2017). Ji et al. developed doxorubicin micelles that restored sensitivity in MRP1-overexpressing, doxorubicin-resistant AML cell lines, with in vivo studies showing increased drug accumulation and improved cytotoxicity. The micelles worked by depleting ATP to inhibit efflux pumps like MRP1 and competitively binding with GSH to reduce drug efflux (Ji and Qiu 2016).

1.9.4 P-gp and PKC

P-glycoprotein (P-gp) is a cross-membrane ion pump associated with poor outcomes and shorter overall survival (OS) in both novel and relapsed AML (Do, et al. 2007, Megías-Vericat, et al. 2015). Drug-resistant AML cell lines, such as SKM-1 and MOLM-13, upregulate P-gp and downregulate Bcl-2, but resistance can be reversed with P-gp inhibitors (Imrichova, et al. 2015). P-gp expression is linked to NF-κB and PI3K/AKT/mTOR pathway activation. Inhibition of these pathways by Balaglitazone decreases P-gp expression and restores sensitivity to doxorubicin (Yousefi, et al. 2017). Protein kinase C (PKC), a calcium phospholipid-dependent protein kinase, regulates drug transporter protein expression, internalisation, and phosphorylation. PKC activation correlates with P-gp phosphorylation and drug resistance (Robinson, Kianna and Tiriveedhi 2020).

1.9.5 LRP

LRP mediates resistance to chemotherapies by blocking nuclear pores to prevent drug entry to the cell, and, transporting drugs out of the nucleus via exocytosis (Zhang, Jing, Gu and Chen 2019). The impact of LRP1 expression on prognosis is unclear, with conflicting studies: one found high LRP expression in bone marrow improved prognosis, while another reported that LRP expression on leukemic blasts reduced the chances of achieving complete remission (Kulsoom, Shamsi and Afsar 2019, Paprocka, et al. 2017).

1.9.6 Resistance to chemotherapy by replication of damaged cells

DNA topoisomerases, such as topoisomerase II (Topo II), facilitate DNA replication and transcription. Anthracyclines inhibit Topo II by stabilising the Topo II-DNA complex, preventing DNA replication and initiating tumour cell death (Michelson, et al. 2020). Topo II is highly expressed in AML, enabling replication of damaged cells despite drug presence. Mutations in Topo II, such as K798L and K798P, hinder drug binding, contributing to resistance (Zhang, Jing, Gu and Chen 2019).

1.9.7 Resistance to chemotherapy and molecular targeting drugs through gene mutations

Resistance can also be achieved by mutation of protooncogenes with roles in cell proliferation and survival. FMS-like tyrosine kinase (FLT3) mutations play a critical role in AML, indicating prognosis and treatment response. FLT3-ITD suggests poorer outcomes, while FLT3-TKD mutations, especially with NPM1 mutations, suggest a better prognosis (Boddu, Prajwal, et al. 2017, Wakita, et al. 2013). FLT3-ITD inhibitors like midostaurin, gilteritinib, and quizartenib target this mutation. Wilms tumour (WT1) mutations are associated with AML relapse, with high WT1 expression signalling poor prognosis (Quek, et al. 2016, Pospori, et al. 2011). Vaccine-based strategies targeting WT1 mutations are currently under investigation. WT1 mutations also confer resistance to imatinib by upregulating quinolinate phosphoribosyl transferase (QPRT) (Ullmark, et al. 2017). RAS family mutations, particularly KRAS, drive constant cell proliferation signalling, giving cancer cells a competitive advantage (Burgess, et al. 2017). In AML, RAS mutations, including KRAS, correlate with decreased overall and event-free survival during induction chemotherapy (Ball, et al. 2019).

Characterising AML patients by their mutations and immunosuppressive features is crucial for determining the optimal treatment. Over the years, diagnostic methods, stratification into treatment groups, and outcome prediction based on AML features have significantly improved.

1.10 Current methods for diagnosis and prognosis

Since mutations of genes and chromosomes are a driving factor for AML, cytogenetics is commonly used for diagnosis. Cytogenetics is the examination of chromosomes, particularly looking for changes to structure where parts or whole chromosomes may be deleted, broken, rearranged, or duplicated. Observed changes to chromosomes can be used to diagnose patients with a condition and assign appropriate treatments. A high percentage of patients who accomplish CR still relapse, and, while CR can be achieved again, with each successive treatment the period of successful remission shortens (Bryan and Jabbour 2015). Despite AML

being well characterised, no single biomarker can predict chemotherapy resistance or outcome (Walter, et al. 2015), although it has been long observed that responders and non-responders display different gene expression profiles (GEP) (Heuser, et al. 2005, Tagliafico, et al. 2006, Raponi, et al. 2008, Visani, Giuseppe, et al. 2017, Visani, G., et al. 2014). Cytogenetics, copy number alterations, driver gene mutational status, indels, SNVs, methylation status and microRNA gene expression have also been noted to change between AML diagnosis and relapse (Hackl, Astanina and Wieser 2017). Identification of the patient subtype and cytogenetics is crucial for assigning the correct treatment for the best outcome.

1.10.1 Stratification of patients into treatment groups

Patients are stratified for initial treatment by their age, cytogenetic profiles, and mutational status. There is not a clear-cut method due to variability in cytogenetic risks being presented together. For example, core binding factor (CBF) translocations in AML are associated with better prognosis (Appelbaum, et al. 2006). However, if the patient also has FLT3-ITD; a driver of mutation, it confers a poor prognosis. Consequently, assigning treatment becomes more complicated. The European LeukemiaNet (ELN) panel (Dohner, et al. 2017) discussed in 1.5.1, sorts patients into favourable, intermediate and unfavourable risk groups. The intermediate group presents the most challenge in treating, as this group has unpredictable response to treatment (Prada-Arismendy, Arroyave and Röthlisberger 2017).

1.10.2 Prediction of patient relapse by cytomorphology

Following the initial induction chemotherapy period, cytogenetics and mutational profiling are also used to estimate the risk of relapse. Predictions are based on databases of large patient cohort profiles and visual assessment of samples (Fey and Buske 2013). Remission is assessed using cytomorphology and examination of bone marrow and peripheral blood samples. If more than 5% of cells are found to be blasts in the bone marrow, and the build-up cannot be attributed to another cause, the patient is defined as in relapse (Chen, Xueyan, et al. 2015). However, only a small number of cells are examined using this method, and the reliability of the prognosis given is dependent on the quality of the sample, as well as the pathologists experience (DeAngelo, Stein and Ravandi 2016).

1.10.3 Prediction of patient relapse by Minimal residual disease

Alternatively, 'minimal residual disease' (MRD) markers could be used to predict relapse. MRD is defined as the presence of leukemic cells which are below the level of traditional morphological detection. Instead, more sensitive methods including multiparameter flow cytometry (MFC), RT

quantitative PCR (RT-qPCR) and next generation sequencing (NGS) are used to identify leukemic cells in a sample population.

1.10.3.1 MFC to detect MRD

MFC is used to identify Leukaemia-associated immunophenotypes (LAIPS) which differ from healthy haematopoietic cells. LAIP cells are identified through aberrant expression of antigens (Ravandi, Walter and Freeman 2018). There are different categories for LAIP in AML. There is 'antigen over-expression' which has abnormally high expression of cd33 and cd34 on myeloid blasts, and 'lack of antigen expression' such as reduced HLA-DR expression. Then 'asynchronous antigen expression' refers to co-expression of antigens typically associated with specific and different maturation stages, for example, expression of early markers such as CD34 or CD117 and later markers such as CD65 and CD14 on myeloid cells. Finally, there is 'cross lineage antigen expression' which is the expression of antigens normally found on other lineage cells, for example lymphoid markers, on myeloid blasts, examples include CD2, CD5 and CD56 (Ravandi, Walter and Freeman 2018, Kern, et al. 2010). The panel for LAIP can include progenitor, myelomonocytic, erythro- megakaryocytic, lymphoid lineage and none-lineage markers indicative of cell maturation, a full panel for detecting all abnormal LAIP could include up to 100 monoclonal antibodies (Bewersdorf, et al. 2019, Ravandi, Walter and Freeman 2018).

1.10.3.2 RT-qPCR to detect MRD markers

RT-qPCR is used to identify some markers of MRD. Markers that can be detected with this method include core-binding factor subunit β myosin heavy chain 11 (CBFB-MYH11), runt-related transcription factor 1 (RUNX1)/RUNX1 translocated to 1 (RUNX1T1), and mutant NPM1. Markers detected by RT-qPCR such as mutations in NRAS, KRAS and IDH2 are used for initial diagnosis only and are not recommended to be used as single markers for MRD, but to be used in combination with a second marker for prognostic significance (Schuurhuis, et al. 2018). While PCR techniques are more sensitive to detecting lowly expressed markers, experiments for individual markers need to be optimised and it can be time consuming to perform a panel, compared to other approaches that identify multiple markers in one experiment, such as NGS.

1.10.3.3 NGS to detect MRD markers

NGS represents a step up from RT-qPCR as it can map out the development of malignant clones. Either the genetic profile can be obtained by sequencing an organism's DNA, or RNA can be sequenced for insight into sequences being expressed by the organism at time of extraction. It can monitor samples for treatment stratification purposes by detecting mutations such as FLT3-ITD and IDH1/2 alongside relapse prognosis markers such as CEBPA and NPM1 (Bibault, et al. 2015, Debarri, et al. 2015, Dohner, et al. 2017). Routine NGS has lower sensitivity than PCR and

can miss detection of mutations, however the more recently developed 'error corrected' NGS approach has made strides towards higher sensitivity detection of mutations. An advantage of NGS over PCR is that multiple mutations can be identified and analysed in one experiment, allowing a more comprehensive patient profile to be built, but it does also require additional bioinformatic processing steps (Schuurhuis, et al. 2018).

1.10.3.4 MRD markers in the future

MRD is already used to assess and monitor patients with acute lymphoblastic leukaemia (ALL) acute promyelocytic leukaemia (APL) and chronic myeloid leukaemia (CML)(van Dongen, et al. 2015, De Angelis and Breccia 2015, Paschka, et al. 2003). There is a drive to develop and refine tests to be applicable to AML, as there is already evidence of its use for relapse prognosis (Jongen-Lavrencic, et al. 2018). However, it is not part of routine clinical practice and there is no standard protocol in methods or markers used, nor are there established cut off points for said markers. The inherent heterogenous nature of AML adds complexity to creating a standard protocol to assess MRD, and it may be a case of developing several protocols dependent on patients known mutations, cytogenetics, and past treatments.

1.11 Prognostic signatures for AML

Biomarkers are measurable molecules that can distinguish between different stages and types of cancer. A prognostic signature, which relates these biomarkers to patient outcomes, can include gene, transcript, or protein expression, gene mutations, chromosomal abnormalities, and epigenetic modifications. These signatures serve multiple purposes, including disease diagnosis, prediction of treatment response, and monitoring of disease progression (Bensalah, Montorsi and Shariat 2007). The ELN cytogenetic risk categories already mentioned are one way to categorise patients but are unable to predict patients that will resist chemotherapy or predict better or poorer outcomes in the CN-AML group. Many researchers are developing prognostic signatures that work independently and in combination with current strategies to improve assessing AML patients for survival, risks of relapse and response to treatments.

In the context of AML these signatures have been developed using transcriptomic, genetic, methylation, and epigenetic profiles of patients, alongside survival data and clinical information. Suitable predictors for survival are commonly selected through regression models, such as the multivariate Cox regression hazard models, the least absolute shrinkage and selection operator (LASSO) model, and, more recently, machine learning and advanced neural networks (ANN) (Walker, C. J., et al. 2021, Lai, Y., et al. 2021, Wagner, et al. 2019, Awada, et al. 2021). For example, the LSC17 stemness score, a 17-gene signature derived from leukaemia stem cells (LSCs) in AML

patients, was tested in five independent patient cohorts and found to predict prognosis for chemotherapy and allogeneic stem cell transplantation (Ng, et al. 2016). Sha et al developed a 5 gene score using TCGA dataset via a multivariate logistic regression model (Sha, et al. 2021). While Wagner et al created the 'Parsimonious 3 gene score' using an artificial neural network trained on the HOVON data set and validated in various AML patient data sets (Wagner, et al. 2019).

1.11.1 Immune system-related signatures show promise for stratifying AML patients

Recently, AML scores have been developed that consider the interactions between AML, the immune system, and the bone marrow niche. Advances in immunogenomic analysis have highlighted the significant role of the immune system in AML progression and relapse. Furthermore, signatures linking AML survival to immune-related genes and immune activity states have emerged and demonstrated similar or improved stratification of patients compared to the ELN risk categories (Zhu, et al. 2020, Chen, Yongyu, Qiu and Liu 2024, Tang, et al. 2019, Wang, J., et al. 2024). Recent research in AML has identified immune-infiltrated and immune-depleted subtypes based on bone marrow transcriptomic profiles (Knaus, et al. 2019, Austin, et al. 2023). An 'immune dysfunction signature' (IED), derived from patient transcriptional data, has linked a transcriptional program associated with immune senescence to poorer outcomes for patients receiving immunotherapy (Rutella, et al. 2022).

Another signature of interest in the IFN-related DNA damage resistance signature (IRDS). The IRDS signature was originally developed to differentiate between breast cancer tumours that were sensitive or resistant to adjuvant chemotherapy treatment (Weichselbaum, et al. 2008). They demonstrated that cells with a high IRDS score (IRDS+) exhibited constitutive STAT1/IFN signalling and postulated chronic IFNG activation may cause failure to transmit the cytotoxic signal and instead promote cancer survival. Further to this, Weichselbaum et al found that knock down of STAT1 decreased expression of IFIT1 and ISG15, which re-sensitised their IRDS (+) cell line Nu61 to doxorubicin. Implying a connection between those genes and chemotherapy resistance (Bernasconi and Borsani 2019). Other groups have found high expression of IRDS genes associated with poorer outcome for chemotherapy and radiotherapy treated patients in other cancers including glioblastoma (Duarte, et al. 2012), glioma, breast, and prostate cancers (Tsai, et al. 2007) as well as head and neck cancers (Khodarev, N. N., et al. 2004)and melanoma (Khodarev, Nikolai N., et al. 2009). In recent years, IRDS genes have been linked to immune-infiltrated AML, and IFNG-related signatures has also been found to correlate with chemotherapy resistance and poor survival in AML (Vadakekolathu, et al. 2020, Corradi, et al. 2022).

These new signatures highlight the critical relationship between the immune system and AML in patient prognosis. IFNG, a key immune-regulating cytokine, enhances cellular immunity and triggers anti-tumour responses. However, it also contributes to immune evasion and resistance to chemotherapy (Mojic, Takeda and Hayakawa 2017, Vadakekolathu, Minden, et al. 2020). While treatment post-hematopoietic stem cell transplant (HSCT) reportedly restores immune activity, prolonged IFNG signalling is known to suppress immune responses, promote cancer growth, and cause T-cell exhaustion (Qiu, et al. 2023). A recent study on newly diagnosed AML patients identified a 47-gene IFNG-related signature that links IFNG signalling scores with immune activation pathways and resistance to venetoclax treatment. Additionally, the study found that IFNG scores varied significantly between FAB subclasses of AML, with late maturity AML (FAB M4/M5) exhibiting the highest scores (Wang, B., et al. 2024).

The varying responses of AML patients to IFNG suggest distinct IFNG response phenotypes, where IFNG can activate either supportive or suppressive immune pathways. This shift in IFNG response might correlate with AML progression, activating immune programs in early AML subtypes, while triggering immune suppressive pathways in mature AML subtypes. The role of IFNG in cancer progression and immunotherapy response underscores the need to develop prognostic signatures that focus on the interplay between IFNG, the immune system, and AML.

1.11.2 AML cell lines for generation of a novel signature associated with IFNG signalling and demethylation status

The examples above were developed using various methods but commonly used clinical patient samples to derive and validate their scores. In this thesis, a score was developed using both cell lines and patient data, which was subsequently validated across multiple patient datasets. The advantages and disadvantages of using cell lines over clinical patient samples are discussed below.

Biomarkers are measurable indicators of normal or pathogenic biological processes, in addition, they can be used to assess success of treatments for diseases. Prognostic scores summarise how the expression of various biomarkers are associated with a potential outcome or biological state. In cancer, a prognostic score could be associated with overall survival or response to therapy and used to stratify patients into the correct treatment groups (Qian, et al. 2021). In many cases, prognostic scores are developed using patient samples, however, they can be developed using cell lines as models. Prognostic scores incorporate biomarkers of disease (e.g. blood pressure, cholesterol, proteins, mRNA) to identify actionable targets to treat disease, define

outcomes, assign prognosis, and predict responses to treatments (Califf 2018, Prada-Arismendy, Arroyave and Röthlisberger 2017).

In this thesis, a prognostic index score was developed using changes to transcriptomics profiles induced by IFNG and demethylation agent 5AzaC. As RNA seq is expensive, only two cell lines were studied with three repeats of each treatment (IFNG alone, 5AzaC alone and IFNG and 5AzaC in combination). Therefore, cell lines response to both agents was first profiled using other methods (chapter 3), before proceeding to a transcriptomic study (chapter 4). Cell lines as models for cancer research have been discussed at length. A condensed, but not exhaustive, summary of the strengths and shortcomings of cell lines as models for cancer research when compared to primary human tissue samples are discussed below (Mirabelli, Coppola and Salvatore 2019).

1.11.2.1 Advantages of Cell Lines versus human tissue samples for Cancer Research:

Cell lines are easy to purchase, and well regulated. Easy growth and maintenance of cell lines enables low risk evaluation of how induced genetic mutations, epigenetics, exposure to environment hazards can affect cancer cell behaviour (Ferreira, Adega and Chaves 2013, Mirabelli, Coppola and Salvatore 2019). Cell lines are well characterised, and consistent, allowing more reproducible research than tissue samples, which can vary enormously between patients (Grizzle, Bell and Sexton 2011). Importantly, novel treatments can be evaluated more freely than with primary tissues, due to limited supply, before continuing to animal models and crucially, patients (Mirabelli, Coppola and Salvatore 2019).

In contrast, obtaining human primary tissue for cancer research can be difficult and expensive. Primary tissue collection is a challenge to standardise, and the tissue may be of variable quality, which can make it difficult to compare results across studies and rarely are specimens collected and stored consistently, leading to variation (Grizzle, Bell and Sexton 2011). Furthermore, maintenance of primary cells required more expertise and yields limited growth at larger costs than the well-established protocols used for immortal cell lines. Additionally, there are ethical and legal considerations when collecting tissue from living patients, leading to long waits before samples are released for research, compared to cell lines, which are immediately available with no ethical quandaries (Richter, et al. 2021).

1.11.2.2 Disadvantages of Cell Lines versus human tissue samples for Cancer Research:

While cell lines do enable study of mechanisms and signalling cascades, they are limited in their biological relevance compared to primary tissues. Although isolated from primary cells at some point, cell lines undergo genetic drift, accumulating many genetic aberrations to enable

unlimited proliferation and atypical cell (Miserocchi, et al. 2017). These aberrations only increase with continued passages, leading to diminishing returns on biological relevance to the original tumour tissue (Miserocchi, et al. 2017). Further, while heterogeneity is present between cell lines, there are fewer differences between cells of the same cell line due to growth from a limited pool of progenitors, which is not representative of the variation seen in a primary cancer tissue (Miserocchi, et al. 2017). Additionally, cell lines do not accurately model the complex interactions between the cancer and surrounding tissues or the immune system, as they are only a singular component grown in isolation (Ali, et al. 2017). A counter to this is to grow cell lines in artificial 3D cultures, which use biomaterials to form scaffolds that imitate the cells natural environment; however, these cultures can be complex and expensive, requiring careful experimental design (Habanjar, et al. 2021).

In comparison, primary cells are of more biological relevance, with the original genome present, preserving original cell behaviours and molecular properties. Human tissues are the most biologically accurate representation of cancer in its environment, with high heterogeneity between cells tissue (Miserocchi, et al. 2017). Primary tissue sample biopsies also remove healthy cells, retaining interactions with the tumour microenvironment, and enabling study of the cross talk between cancer and its environment.

1.11.2.3 AML is heterogenous and presents further challenges for cell line models

While cell lines provide an inexpensive and easily manageable model for studying cancers, they do not accurately represent the subclonal architecture of AML. In AML, newly proliferated blasts continually accumulate additional somatic mutations, evolving over time from the original cell. This dynamic is not present in cell lines, where all cells are clonal. Additionally, AML is highly heterogenous, as evidenced by the M1- through M7 FAB classification system and WHO subdivision into even more groups. Heterogeneity is influenced by many factors, including but not limited to, HSC cell of origin, mutational heterogeneity, epigenetic alterations and changes due to treatments administered (Gu, Dickerson and Xu 2020, Horibata, et al. 2019, Li, Sheng, Mason and Melnick 2016b). Heterogeneity is important to model for, patients with identical genetic mutations may respond differently to the same drugs used to treat their tumours.

In this thesis, four cell lines were used, with a wide spread of cytogenetic characteristics taken at different stages of development sources from bone marrow or peripheral blood, from young and older patients. Ideally, more cell lines would be used to encompass a wider array of possible variations, however time and resources available must also be accounted for. The molecular descriptions, as well as age and sample source for each cell lines are presented in Table 3 below.

Table 3: Details of cell lines used in chapter 3 for studying response to IFNG and 5AzaC treatment, including molecular description, the source of the original cells derived from and age of the patient donor.

Cell line	Molecular description	Sampled from	Age
Kasumi-1	AML (FAB M2) 2nd relapse following bone marrow transplant. RUNx1-RUNX1T1 (AML1-ETO) fusion gene; KIT mutation N822	Peripheral blood	7
KG-1	Erythroleukemia that developed into AML following relapse	Bone marrow	59
SIG-M5	Monocytic AML (FAB M5a) at diagnosis	Bone marrow	63
THP-1	AML at relapse carries t (9;11) (p21;q23) leading to KMT2A-MLLT3 (MLL-MLLT3; MLL-AF9) fusion gene	Peripheral blood	1

With a limited number of cancer cell lines available, it is challenging to encompass the heterogeneity of all genetic and epigenetic variations present in patients without primary *ex vivo* samples (Richter, et al. 2021). For this reason, response to IFNG and 5AzaC treatment is also highly heterogenous and cannot be generalised across all AML. Therefore, discoveries in the AML cell line models may be limited.

The goal of this work was to develop a novel prognostic index using cell lines treated with IFNG and demethylation agent 5-AzaC, that might represent patients with IFNG-related immunosuppressive phenotypes and methylation profiles.

1.11.2.4 Rationale for IFNG treatment

IFNG is known to activate a suite of immunosuppressive molecules which upregulate immune escape mechanisms and remodel the bone marrow niche to encourage T-cells to polarise to tolerogenic phenotypes (Ribas 2015, Matatall, et al. 2018, Florez, et al. 2020). Immune escape mechanisms such as these prevent immune cells recognising and destroying cancer cells, but also reduce the efficacy of chemotherapy which works in conjunction with immune cells to be most effective. When chemotherapy destroys AML cells, antigens are released which can be used to train the immune cells to find and destroy the AML cells (Chen, D. and Mellman 2017). In AML the immune system is already at a disadvantage due to the inhibited differentiation of blood cells and bone marrow producing immature non-functional monocytes, creating a weak or none functioning immune system, therefore chemotherapy cannot amplify the immune systems effects (De Kouchkovsky and Abdul-Hay 2016). The rationale to treat AML cell lines with IFNG and use the RNA seq data to generate a signature, was that the changes induced would be representative of patients with IFNG inducible immunosuppressive AML and may predict patients for favourable response to HSCT and reduced response to induction chemotherapy.

DNA methylation patterns have been used to stratify acute myeloid leukaemia (AML) risk groups. Previously a study in AML patients with CEBPA mutations revealed patients could be divided into two clusters based on DNA methylation variations. A hypermethylated profile was observed in a cluster made of exclusively *CEBPA*-double mutations and demonstrated improved prognosis than patients deemed favourable by cytogenetic risk category (Figueroa, Lugthart, et al. 2010). Similarly, the authors found *NPM1*-mutated patients can be categorised into four methylation clusters, each associated with distinct clinical outcomes. In CN-AML patients without specific features for stratification, five DNA methylation clusters were identified, highlighting the prognostic potential for this group (Figueroa, Abdel-Wahab, et al. 2010). In paediatric patients two DNA methylation signatures associated with cytogenetics were found to significantly correlate with event free survival (Bolouri et al., 2018). Additionally, Luskin et al applied a previously developed method for assessing DNA methylation status at 17 prognostic loci simultaneously they dubbed the M-score (xMELP) (Wertheim, et al. 2014). Application of the M-score to 166 *de novo* AML patients showed improvement on cytogenetics, *FLT3-ITD* status and genetic lesions for predicting which patients would achieve complete remission (CR) from induction chemotherapy (Luskin, et al. 2016). The validity of the M-score was confirmed in multiple independent AML cohorts, underscoring its potential as a prognostic tool (Dinardo, et al. 2017). Therefore, treatment of cell lines with 5AzaC could identify expression signatures associated with hyper or hypo methylation, that have the potential to better stratify patients in the CN-AML category.

1.12 Overview of study

IFNG signalling and demethylation has been associated with an immune-suppressing environment in AML patients. Since AML impedes the generation of mature specialised immune cells, it is commonly managed with a combination of chemotherapy and HSCT, therefore AML exhibiting immune suppression is particularly unfavourable for patients' prospects. This work aimed to test if AML cell lines demonstrating an IFNG-triggered immune-suppressing phenotype could be used as models for constructing scores that effectively categorise the overall survival of AML patients. Higher scores were expected to be associated with immune suppression and correlate with worse overall survival. Moreover, the combination of IFNG signalling and demethylation was hypothesised to be more detrimental to patient outcomes when occurring together rather than individually. The signature generated here is known as a prognostic index score (PI score). A PI score is a composite measure that predicts disease outcomes by integrating multiple prognostic factors, each assigned a specific weight based on its relative contribution.

These weights are determined through statistical analysis, often using regression models, to produce a robust prediction of patient outcomes. Examples of PI scores include the Nottingham Prognostic Index (NPI) for breast cancer and the International Prognostic Index for (IPI) non-Hodgkin's lymphoma (International Non-Hodgkin's Lymphoma Prognostic Factors Project 1993, Galea, et al. 1992). Despite these scores being produced in the 1990's, they are still relevant 30 years later; a testament to their predictive power and the methods used to generate them (Kerin, et al. 2022, Maurer 2023).

In this study AML cell lines (THP-1, KG-1, SIG-M5 and Kasumi-1) were treated with IFNG and 5AzaC to induce changes in IFNG signalling and methylation profiles. The effect of this treatment on expression of immunosuppressive molecules and chemotherapy efficacy were measured using PCR, western blot, flow cytometry and SWATH-MS (Chapter 3). Two cell lines were chosen for further study (Kasumi-1 & KG-1), where effects of IFNG and demethylation treatment were characterised using RNA sequencing and ontology-based pathway analysis (Chapter 4). Finally, the transcriptomic data set produced in chapter 4 was used to create a prognostic index based on treatment induced changes in cell lines (Chapter 5). Transcriptomic data was analysed using multiple pairwise linear regression to identify transcripts most differentially expressed in response to IFNG and 5AzaC. Transcripts were shortlisted for scores based on frequency of appearances in regression analysis and PI scores finalised using a forward selection cox proportional hazards regression model using overall survival (OS) data from the TCGA patient data set, following method outlined by Wagner and Blamey (Blamey, et al. 2007, Wagner, et al. 2019). PI scores were tested in the TCGA data set for stratification and correlation with survival in clinical subgroups and then validated in the Beat-AML, German-AML, CN-AML and HOVON data sets. Finally, performance was evaluated against existing signatures (Chapter 5). Full details on methods used for characterisation and generation of PI score are described in chapter 2.

2 Methods

Lists of reagents and equipment used can be found in the appendix, see 7.2.

2.1 Cell culture

Ordinarily cells divide only a finite number of times before stopping growth completely. However, transformation of cells can be induced or occur spontaneously causing cells to become immortalised and divide indefinitely. Cell lines are useful tools in biology that can be used as models to study pathology of diseases, screen drugs and test mutations. Cell lines used are summarised in Table 4.

2.1.1 Routine cell culture of AML cell lines

This study utilised four AML suspension cell lines, full descriptions are detailed in Table 4. All cell lines were grown at 37°C in a humidified atmosphere with 5% (v/v) CO_2. Growing cells were checked 3 times per week for growth and potential contamination. At ~70% confluency cells were either split for further growth or used experimentally. To split cells, the suspension was transferred to a 50 mL falcon tube and centrifuged at 300 x g for 5 minutes. Supernatant was then removed, and the pellet resuspended in fresh medium. This newly prepared suspension was split between flasks at a density of 0.5 x 10^6 cell/mL.

Table 4: Summary of the four AML cell lines used throughout the study. Details important characteristics, company obtained from, growth conditions used and biological source of the cell line (PB = Peripheral blood and BM = Bone Marrow).

Cell line	Description	Company	Growth conditions	Cell source	Age	M/F
Kasumi-1	AML (FAB M2) 2nd relapse following bone marrow transplant. RUNx1-RUNX1T1 (AML1-ETO) fusion gene; KIT mutation N822	ATCC	RPMI 1640 20 % FCS (v/v), 2 mM L-Glutamine	PB	7	M
KG-1	AML (FAB M6) which progressed into less differentiated AMLfollowing relapse	DSMZ	Iscove's MDM 20 % FCS (v/v), 2 mM L-Glutamine	BM	59	M
SIG-M5	Monocytic AML (FAB M5a) at diagnosis	DSMZ	Iscove's MDM 20 % FCS (v/v), 2 mM L-Glutamine	BM	63	M
THP-1	AML at relapse carries t (9;11) (p21;q23) leading to KMT2A-MLLT3 (MLL-MLLT3; MLL-AF9) fusion gene	DSMZ	RPMI 1640 20 % FCS (v/v), 2 mM L-Glutamine	PB	1	M

2.1.1.1 Trypan blue assay

Cells were counted using a haemocytometer and 0.4 % Trypan blue solution (93595, Sigma Aldrich). Cells which took on the blue stain of Trypan blue were dead and therefore excluded from the cell count. The volume for the desired number of cells was calculated according to the experiment being performed. The calculated volume was either transferred to a flask for further growth or used experimentally, this step was recorded as a passage.

2.1.1.2 Viability and cell count assay

Additionally, cells being used for flow cytometry experiments were counted using an automated cell counter and Solution 18 AO.DAPI (910-3018, Chemometec). Solution 18 contains DAPI and Acridine Orange. DAPI stains dead cells and Acridine orange counter stains both living and dead cells. NucleoView™ software was used to visualise cells and calculate the concentration of viable and none-viable cells in samples.

2.1.1.3 Generating cell stocks

Stocks were made at the earliest possible passage once cells had been established as healthy and reached appropriate volume. The desired number of cells were centrifuged at 300 x g for 5 minutes, the cell pellet was then resuspended in DMSO (67-68-5, Santa Cruz Biotechnology) and FCS (SH30073, GE Healthcare Hyclone) (10%). DMSO is vital to preventing the destruction of cells through the formation of ice crystals, however, it is also toxic and therefore diluted in FCS. The resulting suspension was distributed to cryogenic vials which were then transferred to a CoolCell Freezing System for 24 hours; to control cell freezing and improve cell viability for storage at -80°C.

2.1.1.4 Treatment with IFNG – Preliminary work characterising IFNG response

To characterise the AML cell lines response to IFNG; cells were seeded at 0.5×10^6 cells/mL and suspended in medium containing 5 or 100 ng/mL IFNG for 48 hrs, before further experimental work. Following treatment, cells were either used immediately for extracellular staining by FACs. Or the cell pellet and supernatant were harvested for further downstream analysis by qRT-PCR, colorimetric assay, or SWATH-MS (Sequential Window Activation of All Theoretical Mass Spectra).

2.1.1.5 Cell pellet and supernatant collection

Samples were kept on ice where possible. Following the treatment described above, the cell suspension was transferred to a 15 mL falcon and pelleted by centrifugation at 300 x g for 5 minutes. The cells were then washed with 5 mL PBS (BE17-512F, BioWhittaker), suspended in 1 mL PBS, and transferred to a 1 mL Eppendorf for centrifuging at 250 x g for 5 minutes. Medium

was removed and stored at -20°C for use in the kynurenine assay, and pellets stored at -80°C until RNA extraction.

2.2 Molecular biology techniques

2.2.1 RNA-extraction for quantitative real-time PCR

Cells were grown and treated as described in 2.1 and harvested as described in 2.1.1.5. Cell pellet RNA was extracted using the RNeasy Mini Kit (Qiagen) as per manufacturers direction. RLT buffer was prepared by adding 10 µl of β-mercaptoethanol (M3148, Sigma) per 1 ml of Buffer RLT to improve the quality of RNA extracted by denaturing RNases in the sample. RNA samples were stored at -80 °C until further use.

2.2.2 Quantification and quality check of extracted RNA

Extracted RNA was checked for quantity and quality using the NanoDrop™ 8000 Spectrophotometer (Thermofisher Scientific). For RNA to pass the quality check; the A260/A280 value had to be above 1.8, and the A260/A230 value had to be between 1.8 and 2.2. The volume of sample required for 1 µg was calculated for cDNA generation by reverse transcription.

2.2.3 Reverse transcription to generate cDNA

RNA was converted to cDNA using reverse transcription. For each sample, 1 µg RNA was combined with 1µL oligo dT (C1101, Promega) and the total volume adjusted to 10 µL with nuclease free water (NFH$_2$O) (AM9930, Ambion). Mixtures were incubated at 70°C for 5 minutes to promote primers annealing to single stranded RNA efficiently. Samples were then transferred to ice immediately. After which, 5 µl 5x RT buffer (M531A, Promega) to maintain a favourable pH for the reaction, 0.7 µl RNasin® (N2515, Promega) to prevent RNA degradation, 1 µl M-MLV Reverse Transcriptase (M1705, Promega), 1 µl dNTPs (U1511, Promega), and 7.3 µl NF H$_2$O was added to each sample. The samples were then incubated at 40 °C in a water bath for 1 hour, following which, samples were heated to 95 °C for 5 minutes to deactivate the reaction and the new cDNA stored at -20 °C.

2.2.4 Quantitative real-time PCR (qRTPCR)

Quantitative real-time PCR was used to measure the levels of mRNA expression of a particular target in samples. Per reaction, 1 µL of cDNA was combined with 5.75 µL of SYBR Green (172-5124, Bio-Rad), 0.5 µL of forward and reverse primer to a concentration of 10 µM and 3.75 µL NFH$_2$O. Samples were analysed with three technical repeats using a Rotor-Gene Q real-time PCR cycler (Qiagen), a minimum of three biological repeats were used for each primer. Denaturation was carried out at 95 °C for 5 minutes followed by a 40-cycle program. Each cycle consisted of a

10 second denaturation (95°C), 20 seconds annealing (all genes were optimised to 58°C) and 20 seconds extension (72°C). Experiments were designed to comply with MIQE guidelines (Bustin, et al. 2009). Samples had to pass melt curve analysis to confirm the amplification of the desired gene had occurred and results seen were not caused by primer dimers. Samples were checked for contamination by using NTCs (no template control) in each batch ran. Biological replicates were used to ensure the differences in qPCR results were a consequence of treatment and not a product of batch variation.

2.2.5 Primer efficiency testing

Primers were tested for efficiency prior to use. A test sample of cDNA was diluted 5-fold to create a serial dilution. All primers were analysed using the method for qRT-PCR described above. A standard curve was generated from the serial dilutions with the equation $y = mx + b$, where m (the slope) indicates the primer efficiency. The m value -3.32 gives 100% efficiency and so a slope as close to this value as possible is desirable. For primers to pass efficiency testing they had to display between 90 and 110% efficiency. An example is given in Figure 4.

M= -3.33284
Efficiency = 99.5%
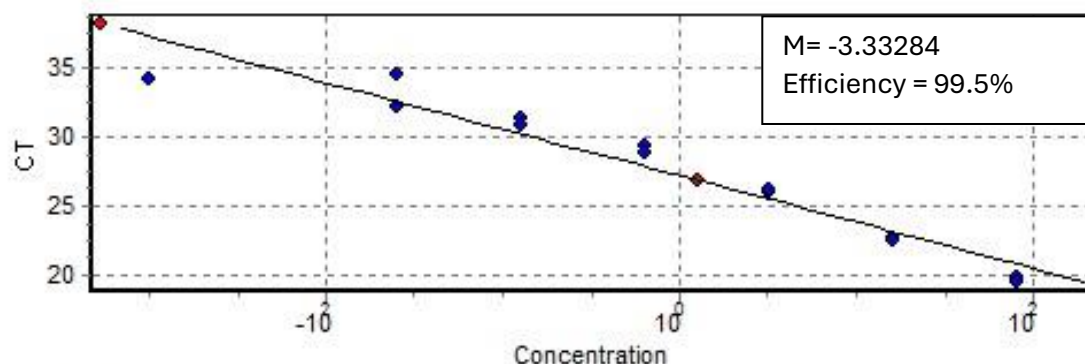
Figure 4: Example of a standard curve generated to test for primer efficiency.

2.2.6 RNA extraction for Next Generation Sequencing (NGS)

Cells were grown until confluent and treated with either a single dose 5 ng/mL IFNG or 0.5 µM 5-Azacytidine every 24 hrs, alone or in combination, over the course of 48 hrs. Cells were harvested as described in 2.1.1.5. Cell pellet RNA was extracted using the RNeasy Mini Kit (Qiagen) following the instruction manuals extraction process for cell line samples, with the addition of the on column-DNase digestion step to eliminate DNA and ensure RNA purity. RLT buffer was prepared by adding 10 µl of β-mercaptoethanol (M3148, Sigma) per 1 ml Buffer RLT to improve quality of RNA extracted by denaturing RNases in the sample. DNase I stock solution was prepared by dissolving the lyophilised DNase I in 550 uL RNAse free water provided in the kit using a syringe. RNA samples were stored at -80 °C until further use.

2.2.7 Preparing RNA for sequencing by Novogene

Prior to sending the samples to Novogene for analysis they were assessed for quantity and quality. First the samples were assessed using the NanoDrop™ 8000 Spectrophotometer, following instructions of the manufacturer. This was to ensure that the sample concentration fell within the range of 5-500 ng/µL required by the Agilent bioanalyser 2100 (Agilent). Samples were then measured using the Bioanalyser 2100 with the RNA Agilent kit and RNA Nano chips following the manufactures instructions, to ensure concentration was appropriate and RNA integrity number (RIN) was the required minimum 6.8 or above.

2.3 Metabolic and Protein profiling assays

2.3.1 Kynurenine assay

A standard curve was generated by diluting a 50,000 µM L-Kynurenine (K8625, Sigma-Aldrich) stock in medium corresponding to the cell line being tested, to make 14 standards ranging from 0 µM to 200 µM concentration. Cell medium without cells or kynurenine was used as a control. Then, 150 µL 30% trichloroacetic acid (T6399, Sigma Aldrich) was added to 300 µL of each standard, control, and sample, before vortexing and centrifuging at 8000 x g for 5 minutes at 4°C and supernatant then removed. Ehrlich's reagent was prepared by adding 20 mg p-dimethylaminobenzaldehyde (156477, Sigma Aldrich) per 1 mL glacial acetic acid (A2683, Sigma Aldrich). 75 µL of standard, control and samples were loaded onto 96 well plate in triplicate and an equal volume of Ehrlich's reagent added and mixed. Plates were read after 15 minutes incubation at 492 nm, on a Tecan infinite m200 Pro plate reader. Sample absorbance at 492 nm was compared to a standard curve of kynurenine concentrations, and sample concentrations estimated using $y = mx + C$.

2.3.2 Protein extraction for Western blot and mass spectrometry analysis

Cells were grown for 48 hrs with and without 100 ng/mL IFNG treatment, then harvested as previously described in 2.1. To each sample, 300 µL of RIPA buffer (89900, Thermo Scientific), spiked with 1 in 100 Halt protease and phosphatase inhibitor cocktail 100x (1861281, Thermo Scientific) was added to prevent proteolysis and loss of phosphate groups. In addition, 0.5 M EDTA Solution 100x (Thermo Scientific) was added to further inhibit proteases by eliminating free divalent cations. Samples were left on ice for 30 minutes, passed through a 29G fine needle 10 times, sonicated for 5 minutes, and then centrifuged at 12,000 x g for 15 minutes at 4°C. Supernatant were stored at -80°C until further use.

2.3.3 Measurement of protein concentration using Pierce assay

Protein concentration for SWATH-MS samples was measured using Pierce protein assay; a colorimetric assay that exhibits less variation than dye binding methods. To prepare the reagent, 1g of Ionic detergent compatibility reagent (22663, Thermo Scientific) was added per 20 mL Pierce™ 660 nm protein assay reagent (22660, Thermo Scientific). A pre-diluted protein assay standard BSA set (23208, Thermo Scientific) was used with the following concentrations: 0, 125, 250, 500, 750, 1000, 1500, 2000 µg/mL. Samples were prepared for the assay by diluting 1 in 10 in DDH2O to eliminate background interference from the lysis buffer. RIPA lysis buffer was diluted 1 in 10 in DDH2O to act as a background control. In addition, 10% 100x Triton (T8787, Sigma) was added to each sample to a total concentration of 0.8% Triton to prevent RNA/DNA in the samples forming a precipitate. 10 µL of each sample, standard and background control was added to a 96 well plate in triplicate. Then 150 µL of protein assay reagent was added to each well, the plate was left to incubate for 5 minutes at room temperature and absorbance measured at 660 nm using a Tecan infinite m200 Pro plate reader.

2.3.4 Measurement of protein concentration using Bio-Rad assay

Protein concentrations for use in western blot were measured using the Bio-Rad protein assay dye reagent concentrate (5000006, Bio-Rad). The dye reagent was prepared by diluting dye reagent concentrate with DDH2O water 1 in 5. Protein standards were generated by diluting the 1000 µg/mL protein assay standard from the BSA set (23208, Thermo Scientific) with DDH2O water to concentrations: 0, 50, 100, 200, 300, 400 and 500 µg/mL. Samples were diluted 1 in 10 in DDH2O to eliminate background interference from the lysis buffer. RIPA lysis buffer was diluted 1 in 10 in DDH2O to act as a background control. Then, 10 µL of each sample, standard and background control was added to a 96 well plate in triplicate. To complete the reaction, 200 µL of diluted dye reagent was added to each well and mixed using a multipipette. The plate was incubated for 5 minutes at room temperature, before absorbance was measured at 595 nm using a Tecan infinite m200 Pro plate reader.

2.3.5 SDS-PAGE

In preparation for detection of target proteins by western blot, 30 µg of cell lysates were separated using SDS-PAGE and transferred to a nitrocellulose (NC) membrane for staining. First samples were reduced for SDS-PAGE.

2.3.5.1 Reducing samples for SDS-PAGE
Denaturation of sample protein is necessary for efficient separation by SDS-PAGE. Laemmli buffer is a reducing agent which reduces inter and intra-molecular disulphide bonds resulting in

protein denaturation and providing proteins with negative charge. Sample were prepared for separation by molecular weight by adding samples at a 3 to 1 ratio to 4 x Laemmli buffer (1610747, Bio-Rad). Lysates were then incubated at 95°C for 10 minutes using a heating block, once cooled to RT samples were immediately used.

2.3.5.1.1 Running SDS-PAGE

Premade gels 4–20% Mini-PROTEAN® TGX™ Precast Protein Gels (4561093, Bio-Rad) were inserted into the appropriate running module and placed into the transfer tank. The tank was filled with running buffer up to the line indicated by the apparatus being used. In each experiment, 30 µg of sample was loaded alongside 5 µL of Precision Plus Protein™ WesternC™ Standard (1610385, Bio-Rad). To separate proteins the gel was run at 50 V for 5 minutes to check for even running and then ran at a constant 100 v for 1 hr.

2.3.6 Western blot

Once proteins in samples were separated by SDS-PAGE they were probed for targets of interest using western blot.

2.3.6.1 *Transfer of proteins on to nitrocellulose membrane*

Following separation by SDS-PAGE, proteins were transferred from the gel to a NC membrane. For this process, a "sandwich" was constructed from sponge, filter paper, the gel and NC membrane. All components were pre-soaked in transfer buffer prior to assembly and the order of stacking is shown in Figure 5. As the proteins have a negative charge, they move out of the gel, towards the positively charged anode, and on to the membrane during the transfer.
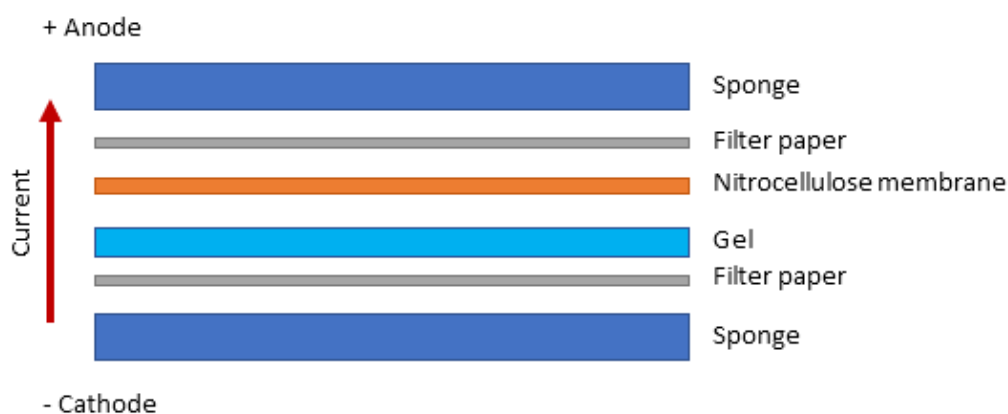
Figure 5: A diagram showing the orientation of components for the "sandwich" used to transfer proteins from gel to membrane.

The assembled sandwich was then placed in the transfer tank with electrodes matching the tanks orientation to ensure current flows. Then, the sandwich was submerged in transfer buffer and a

current of 200 mA applied for 1 hour. To prevent sticking of the gel to the membrane due to heat caused by the current, the tank was kept cold in a polystyrene boxed filled with ice.

2.3.6.2   *Detection of target proteins on nitrocellulose membrane*

Successful transfer of proteins was indicated by a visible ladder on the membrane. Following transfer, the membrane was cut into sections according to the molecular weight of proteins of interest (POI). Sections were incubated in blocking buffer for 1 hour at RT while shaking. For standard proteins, blocking buffer was made up of 1 x TBST and 5 % fat free milk powder. For phosphorylated protein, blocking buffer was made of 1 x TBST and 5 % BSA. The blocking step is essential as it prevents none-specific binding of antibodies. After the blocking step, blocking buffer was replaced with fresh blocking buffer combined with specific antibody for the POI added at manufactures recommended concentration. Membranes were incubated with the primary antibody overnight at 4°C on a shaker.

After overnight incubation, primary antibodies were drained off the membranes which were then washed 3 times for 10 minutes with 1 x TBST at RT on a shaker to remove unbound antibody. Host specific secondary antibody and conjugate for the molecular weight ladder were added to blocking buffer at manufactures recommended concentration. The membranes were incubated at RT on a shaker for 2 hours. Membranes were washed 3 times for 10 minutes with 1 x TBST at RT again. Following the wash step, membranes were placed individually on a dark backing board and clarity western ECL substrate (1705060, Bio-Rad) was added at a 1:1 ratio. Imaging of blots was carried out on a Syngene G:Box with exposure time ranging from 30 seconds to 3 mins.

2.3.7 Flow cytometry

Flow cytometry is a technique which can be used to analyse expression of multiple protein targets on large volumes of cells rapidly. Cells are 'labelled' with fluorochromes which bind to the POI. Cells are injected into the flow cytometry instrument where they enter the flow chamber. Here, hydrodynamic focusing occurs by laminar sheath flow facilitating cells to be positioned central to the laser. When fluorochromes are hit by the laser they absorb its wavelength of light and emit at a different wavelength in response. This signal is used to quantify the target protein present on the cell, and hundreds of cells can be analysed per second. The high quantities of data are collected and processed by computer.

The phenotyping of cells for adaptive immune response (AIR) targets and IFNγ response was carried out using a Beckman Coulter Gallios™ flow cytometer instrument. The Gallios has a capacity of 10 fluorescent channels; the laser wavelengths and filter details are shown in Table 5.

Laser	Excitation wavelength	Channel	Emission wavelength
Blue	488 nm	FL1	525/40
		FL2	575/30
		FL3	620/30
		FL4	695/30
		FL5	755 LP
Red	638 nm	FL6	660/20
		FL7	725/20
		FL8	755 LP
Violet	410 nm	FL9	450/40
		FL10	550/40

2.3.7.1 Extracellular staining by flow cytometry for cell immunophenotyping

Prior to running samples, compensation was carried out per core facility guidance using cell populations which were negative for the utilised fluorochromes. Each cell line was grown with and without 100 ng/mL IFNG, 2 million cells were taken per treatment sample and split into 4 FACS tubes containing 0.5 million cells each. Samples were centrifuged at 400 x g for 5 minutes at 4°C and medium decanted. Then washed with 2 mL cold pbs and centrifuged again and kept on ice. 100 μL FACS buffer was added to all samples along with recommended volume of antibodies according to manufacturer. For each treatment, 3 staining conditions were used; unstained control, L/D viability control and stained. Samples were vortexed and incubated with labelling antibodies for 30 minutes at 4°C protected from light. 200 μL of Isoton was added to each FACS tube and samples were analysed using a Gallios flow cytometer (Beckman Coulter).

2.3.7.2 Determining Daunorubicin EC50 for Kasumi-1

Cells were grown until confluent; centrifuged at 300 x g for 5 minutes and supernatant removed. Cells were then suspended in fresh medium and seeded at 5 million cells per 10 mL medium in T25 flasks. They were then dosed with Daunorubicin (30450, Sigma-Aldrich) at 0, 0.2, 0.4, 0.6, 0.8 and 1 μM. The flasks were incubated at 37°C for 48 hrs and viability determined by Gallios flow cytometer (Beckman Coulter).

2.3.7.3 Testing cell line viability post exposure to IFNG, 5AzaC and Daunorubicin

Cells were grown until confluent; centrifuged at 300 x g for 5 minutes and supernatant removed. Cells were then suspended either in fresh medium only, or fresh medium containing 0.5 ng/mL IFNG or 0.5 uM 5-Azacytidine (A2385, Sigma Aldrich) and were seeded 5 million cells per 10 mL medium in T25 flasks. They were then dosed with Daunorubicin (30450, Sigma-Aldrich) at 0, 0.1,

0.2, 0.4 and 2 µM. The flasks were incubated at 37°C for 48 hrs and viability determined by Gallios flow cytometer (Beckman Coulter).

2.3.7.4   Annexin V staining

In live healthy cells phosphatidyl serine (PS) is located on the inner cytoplasmic layer of the plasma membrane. When apoptosis occurs, the plasma membrane structurally shifts, and PS translocates to the out layer of the membrane. Annexin V is a cellular protein which has high affinity for PS and can be used to detect apoptosis. To act as a dye, Annexin V is conjugated to a fluorescent or enzymatic label, where signal is proportional to number of PS bound. The difference in signal produced between live and apoptotic cells is typically 100-fold, allowing for easy distinction between populations during analysis. As compromised plasma membranes can provide Annexin V passage into the inner leaflet of the cell, it is advised to use a live cell stain in tandem to avoid false positives.

2.3.7.5   LIVE/DEAD staining

Using a live cell stain such as LIVE/DEAD fixable stain, in combination with Annexin V both prevents false positive readings and provides more information to distinguish between cell populations which are live, apoptotic, and dead. Annexin V alone runs the risk of staining the inner leaflet PS in compromised cells, which may be necrotic rather than apoptotic. The LIVE/DEAD stain is based on fluorescent reactive dye that detects amine groups on proteins. As this dye does not penetrate intact live cell membranes, it only binds outer cell membranes producing a dim signal for live cells. Dead cells have damaged membranes, allowing the stain access past the membrane and bind to both exterior and interior proteins, causing a much brighter signal, typically greater than 50-fold compared to live cells. Therefore, when used in combination, cells testing as Annexin V positive, and LIVE/DEAD negative, can be more confidently assigned as apoptotic. As it can be assumed the annexin staining is from the outer leaflet and the membrane has been shown to be uncompromised by low LIVE/DEAD signal.

2.3.7.6   Viability staining by flow cytometry

For viability assays, 2 million cells were taken per sample type and split into 4 FACS tubes containing 0.5 million cells each. For each treatment, 4 staining conditions were used; unstained control, L/D viability control, annexin control and stained. Samples were centrifuged at 400 x g for 5 minutes at 4°C and medium decanted. Then washed with 2 mL cold PBS and centrifuged again and kept on ice. 100 µL FACS buffer was added to all samples, except for the annexin control. Annexin V is dependent on calcium for binding, therefore 100 µL of calcium free PBS was added to ensure a correct baseline. Samples were vortexed and incubated for 30 minutes at 4°C protected from light with 0.5 µL L/D viability dye. After this, 2 mL PBS was added to all samples

which were centrifuged again at 400 x g for 5 minutes at 4°C. Medium was removed and 200 µL of cold Annexin v binding buffer (422201, BioLegend) was added to all samples except for the annexin control, which was suspended in cold PBS. Samples were incubated at RT for 15 minutes protected from light, then analysed using a Gallios flow cytometer (Beckman Coulter).

2.3.7.7   Gating Strategy for analysing extracellular staining

Flow cytometry data was analysed using Kaluza 2.0 (Beckman Coulter), the gating strategy is outlined in Figure 6. The first step was to identify the live single cell population of interest. To exclude debris, forward scatter (FS) was plotted against side scatter (SS) in step 1. Step 2 was to plot FS height (FS TOF) against FS area (FS INT), to remove doublets from the analysis. This is important as doublets are read as one event, when they are in fact two events, and this can skew results. Live/dead stain (3) was used to further ensure the population gated was live, by excluding the dead population, which auto-fluoresce. The cell viability dye used was an amine dye, which fluoresces when it binds to amine groups on proteins. Live and dead populations were distinguished by intensity of the signal. As live cells have their membranes intact, there are fewer proteins to be bound by the dye therefore producing a lower signal. In contrast, dead cells membranes are damaged and so the dye can permeate through and bind to more proteins inside, resulting in a higher intensity signal. The low signal intensity population (Live cells) was gated (3) and other targets of interest were exclusively measured from this population (4).
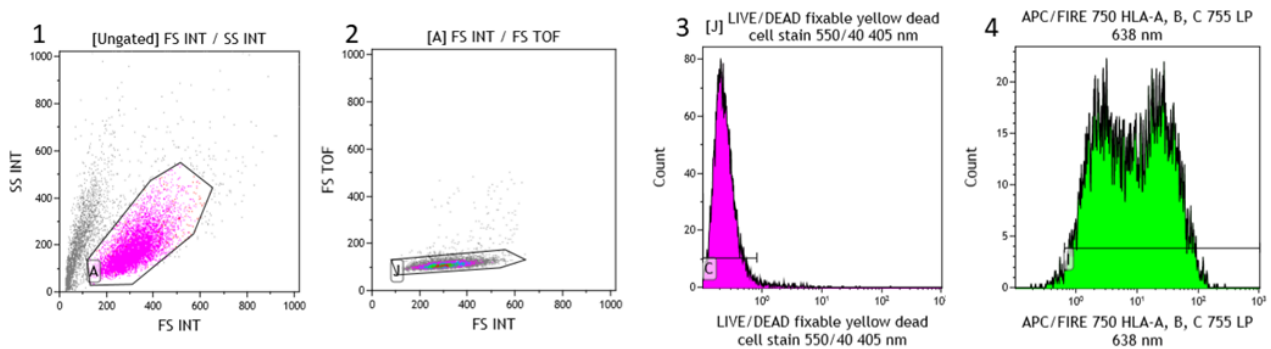
Figure 6: Gating strategy for cell surface staining data obtained through flow cytometry. 1) Shows gate A being placed on a FS/SS plot excluding dead cells and debris. 2) FS TOF plotted against FS INT and gate J placed to remove doublets in data. 3) Live/dead viability dye was used to exclude dead cells from analysis and gate C placed. 4) Example of HLA-A,B,C target MFI measured in gate C population

2.3.7.8   Gating Strategy for analysing living populations

The gating strategy for the annexin V and Live dead staining experiments was the same as above, except for step 1, where all cells were gated in the FS vs SS plot. More detail on gating is shown in chapter 3.

2.4 Mass Spectrometry

Correlation between transcriptome and proteome is not absolute, and can be influenced by transcript and protein half-life, mRNA degradation and post-translational modifications (Chakraborty, et al. 2018). Proteins are the final functional product of the transcription to translation mechanisms. As proteins control most cellular processes, it is important to understand how they are disrupted in cancer cells. Proteomics is the identification and quantification of proteins present at time of sampling and is key to understanding functional changes in cell. Proteomics studies are commonly carried out by protein mass spectrometry, which has advanced in recent years to accommodate identification of more than 12,000 proteins and 10,000 PTM sides in a single sample. MS has been used to generate proteomes of cancer patients, but the workflows required to do this are very specialised and unlikely to be clinical routine for a long time (Doll, Gnad and Mann 2019).

Typically, the sample is vaporised, then ionised by bombardment with electrons, which causes whole molecules to become charged or break into charged fragments. Samples are then accelerated and exposed to an electronic or magnetic field. This causes separation, as depending on the samples mass to charge ratio, they will be 'deflected' to different extents. Ions are detected by an electron multiplier and results are displayed as spectra. Components of the sample can be identified by comparing masses obtained to masses of known molecule or known fragmentation patterns.

The primary challenge of using mass spectrometry for proteomics, is that proteins usually exist in complex mixtures within a biological medium. The first problem to be addressed is that the ionisation techniques used for large molecules require equal amounts of each molecule present. This is rarely the case in biological samples which can wildly differ in volumes of component proteins present. If a mixture like this is ionised then proteins in higher abundance will supress signals of low abundance proteins, leading to a loss of important data. The second problem is that the mass spectra produced by protein samples is highly complex and therefore hard to correctly interpret. To combat these problems high-performance liquid chromatography (HPLC) is applied prior to MS. This technique separate mixtures into peaks, the contents of which can be identified based on the mass spectrum they produce.

2.4.1 The advantages and limitations of the SWATH-MS approach to proteomic profiling

Proteomic profiles were generated and compared between IFNG treated and untreated cell lines using mass spectrometry (MS), specifically the SWATH-MS method. When a compound is analysed by MS, the first step is for the compound to be transformed into gas phase ions, for example by bombarding it with electrons. The ion produced then fragments into smaller charged molecules such as ions and radicals, which themselves can further fragment. This produces ions of different sizes and charges, which can then be separated by their mass to charge ratio (m/z) and detected according to their abundance (De Hoffmann and Stroobant 2007). There are three approaches to generating proteomic profiles using mass spectrometry analysis, which have been developed to meet different research needs (Domon and Aebersold 2010) These methods are data-dependent acquisition (DDA), selected reaction monitoring (SRM) and data-independent acquisition (DIA) (Liu, Y., et al. 2013). All these analysis types are performed using two MS in tandem, designated MS1 and MS2. Samples are ionised and passed through MS1 for screening and separation based on their m/z ratio, then ions of a particular m/z ratio are chosen for further fragmentation. These fragments are fed to MS2 for further separation by their m/z ratio and detected to produce a fragment ion spectrum which is matched to a premade library for identification of peptides (De Hoffmann and Stroobant 2007).

The first approach is data dependent acquisition (DDA). This approach is suited to bias free discovery research as it requires no prior knowledge on the analyte and selects proteins of interest based on ions with the highest abundancy (Sidoli, et al. 2015). The main drawback of this approach is that the most abundant ions can vary between samples, which leads to varied quantitation and inconsistency in peptide detection alongside reduced reproducibility of results (Wu, J. X., et al. 2016). In the event of too many peptides co eluting in a single MS1 screening, DDA preferentially samples highly abundant peptides, and low abundancy peptides are left unreported. A further drawback is that peptides are only sampled once or twice, preventing precise quantification (Hu, Noble and Wolf-Yadlin 2016).

Some of the drawbacks of DDA are countered by the selected reaction monitoring (SRM) method, which is sometimes referred to as targeted proteomics. SRM accurately and reproducibly analyses samples multiple times compared to DDAs once or twice, to quantify a set of preselected proteins (Hu, Noble and Wolf-Yadlin 2016, Liu, Y., et al. 2013). SRM can achieve a higher sensitivity than DDA and detect low abundance proteins, with the caveat that analysis is restricted by prior knowledge and detection is limited to a known predefined list of proteins. In

addition, SRM is limited to up to one hundred proteins per run, compared to DDA thousands (Hu, Noble and Wolf-Yadlin 2016).

The data independent acquisition (DIA) method differs from the DDA and SRM methods. Firstly, unlike DDA where precursor ions are chosen by MS1 scan to be detected by MS2, fragment ions are continuously acquired without bias by MS2. Furthermore, unlike SRM, no prior knowledge of peptide precursor m/z values is used, and is therefore based on targeted data extraction, as opposed to SRM which is based on targeted MS acquisition (Ludwig, et al. 2018). In SWATH-MS, the fragment ion spectra within a defined window of m/z are measured, this is repeated over several cycles across the complete m/z range (Ludwig, et al. 2018). This method gives three pieces of information for a fragment, which are retention time, m/z and abundance (Gillet, et al. 2012). Which can then be matched to a spectral library for identification. SWATH-MS allows for a mid-ground between DDA and SRM, identifying large quantities of peptides with a higher reproducibility than DDA, but not to the same sensitivity as SRM.

SWATH-MS peptide quantification is still three to ten-fold less sensitive than SRM methods, and so is not the best option for any research seeking quantification of low abundance proteins accurately (Gillet, et al. 2012). Furthermore, data analysis is more challenging than DDA. Peptide query parameters (PQP) are set up using spectral libraries generated from previous experiments. PQPs include chromatographic elution times of peptides, optimal peptides for a protein and most intense fragment ions associated with target protein under applied fragment conditions (Ludwig, et al. 2018). The parameters set allow for successful identification of peptides by peptide centric scoring (PCS). PCS uses a pre-defined list of peptides and assess if those peptides are in the acquired data and to what confidence. This method of analysis is used in SRM as well as DIA methods such as SWATH-MS (Ludwig, et al. 2018). SWATH-MS was used to compare cell line responses to IFNG based on their proteomic profiles and select cell lines for further study with NGS.

2.4.2 Generating samples for cell line characterisation by SWATH-MS

Samples for all four AML cell lines were stimulated for 48 hrs with 100 ng/mL IFNG and harvested and pellets were lysed. Protein concentrations of lysates were estimated using the pierce protein assay, as described in section 2.3.3. Finally, the calculated volume for 50 µg of each sample was distributed to eppendorfs in duplicate and stored at -20 °C until samples were analysed.

2.4.3 Sample analysis by SWATH-MS

Analysis of samples was performed using a SCIEX TRIPLE TOF® 6600 mass spectrometry instrument linked to an Eksignent NanoLC 425 HPLC system following Lambert et als method (Lambert, et al. 2013) by Dr Amanda Miles, Dr David Boocock and Dr Clare Coveney. Samples were loaded into the LC system for fractionation by reverse phase HPLC. To do this 3 µL of each sample was injected on a YMC (15 cm by 0.3 mm) Triaer-C18 column and ran using the microflow setting (5 µL min). Independent Data Acquisition (IDA) was used to generate the spectral library. This was then passed through targeted analysis by SWATH-MS data acquisition. Files generated by IDA analysis of cell line lysates were searched, first separately, then pooled together using Protein Pilot 5.0 (SCIEX) and the Human Swissprot database. The OneOmics assembler was used to assemble data and generate fold change with confidence data for each protein change and normalisation was carried out by Dr Amanda Miles according to method outlined by Jean-Phillipe (Lambert, et al. 2013). A 1% False Discovery Rate (FDR) was used as a cut-off point.

2.5 Transcriptomics

Transcriptomics is a technique used to detect and quantify the complete set of coding and none-coding RNA transcribed from the genome at a given time. The transcriptome is more complex and transient in nature than the genome as it can be influenced by cellular, environmental, and external stimuli (Chakraborty, et al. 2018). The first widely used technique to build transcriptomic databases was the microarray, which uses a chip spotted with cDNA molecules that bind to complimentary sequences present in the sample. However, microarrays are limited by the need for prior knowledge of the transcriptome to design probes and can therefore produce bias results. Consequently, microarrays were superseded by RNA sequencing (RNA-seq), which does not use transcript specific probes and so is unbiased in its detection. Additionally, RNA-seq has a higher specificity and sensitivity allowing it to detect low expression transcripts (Chakraborty, et al. 2018). The main application of this technology is to identify genes which exhibit differential expression between cancer and non-cancer states and provide insight into the proteome.

2.5.1 RNA sequencing of cell line samples

The RNA sequencing was carried out by third party Novogene using their Illumina NovaSeq platforms, which use a paired-end 150 bp sequencing strategy. The statistical analysis for differential expression between samples and FDR correction were also carried out by Novogene using the dseq2 analysis package.

2.6 *In silico* analysis of experimentally acquired and publicly available data sets

In silico analysis was used to further investigate and compare the proteomic and transcriptomic data sets generated from cell lines treated with and without IFNG.

2.6.1 Filtering of mass spectrometry analysis data

All proteins with a confidence above 60% were considered to have undergone a significant change in expression. To further consolidate targets, only changes of two-fold or more were further considered for proteins with confidence level between 60 and 75% confidence. Finally, the resulting significant proteins were presented as a heat map.

2.6.2 Metascape analysis

Mass spectrometry data was uploaded to Metascape as a gene list. A 'one click' analysis was used to assess enriched pathways per cell line. Biological processes associated with IFNG were selected and presented as a heat maps and tables. NGS data was treated in the same way, with statistically significant genes predetermined by Novogene analysis and uploaded to metascape.

2.6.3 Identification of a gene signature using cell line transcriptomic data and patient data sets

Patient data sets are an excellent alternative resource to overcome limited availability of patient samples. For the most part, patient data sets that are publicly available are derived from transcriptomic or genomic experiments, it is rarer to find proteomic data bases. These data sets provide quantified expression of target genes or mRNA, but also clinical data for the patients. They provide an alternative no lab-based resource for validation of signatures, with clinical parameters that can be used to study specific patient groups. In this study, cell lines treated with IFNG, 5AzaC or both were used to generate a transcriptomic dataset, which was then applied to the TCGA-AML data set to generate signatures associating the treatment to patient survival. These signatures were further validated in the BeatAML, German-AML and CN-AML data sets for clinical validation. Data sets were obtained from the cbioportal platform (Table 6).

Table 6: Table of patient data bases used for discovery and validation of prognostic index scores in chapter 5. Title is given, along with number of samples, if the data set was used for discovery or validation and a link.

Title	Number of samples	Training or validation set	Link
Acute Myeloid Leukaemia (TCGA, NEJM 2013)	200	Training	https://www.cbioportal.org/study/summary?id=laml_tcga_pub
Beat AML	672	Validation	https://www.cbioportal.org/study/summary?id=aml_ohsu_2018
Acute myeloid leukaemia samples =< 60yrs on HG-U133 plus 2 (German series)	537	Validation	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE6891
Prognostic gene signature for normal karyotype AML (CN-AML)	405	Validation	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE12417
E-MTAB-3444 (HOVON)	662	Validation	https://www.ebi.ac.uk/biostudies/arrayexpress/studies/E-MTAB-3444

2.6.4 Statistical analysis of omics data

There are three components to transcriptomic analysis of RNA seq, these are pre-processing of data, statistical analysis and functional interpretation or application. Pre-processing of data involves short read alignment and assembly, as well as artifact filtering (Fang, Martin and Wang 2012). For this thesis, pre-processing of data was performed by the company Novogene using Dseq2 and data was received as RNA-seq data set where transcript levels were reported as discrete counts. For any data set there will be more than one type of statistical analysis that is appropriate, each with their own strengths and limitations. The same data set analysed via different techniques could produce similar or markedly different results. If the computing power, expertise, and resources are available, it is recommended to use multiple methods and investigate transcripts commonly reported (Mou, et al. 2020, Liu, X., et al. 2022).

2.6.5 Methods of statistical analysis used to assign significance to variation between samples

2.6.5.1 T-Test

T-tests are commonly used statistical tests which evaluate if the mean of a variable is significantly different between two sample datasets. Different T-tests are used depending on the number of samples in each group, distribution and variance of the populations and relationships

between groups (Student 1908, Welch 1947, Kim, T. K. 2015). The null hypothesis is that the mean is the same in both samples. In the case of this thesis, this was used to compare the differential expression of a gene or protein between two samples. The T-test is a basic test, which is limited by its capacity to only compare two data sets and is not appropriate for data with lots of variation or noise, which is common in large omic data sets. Furthermore, if multiple variables are also being tested and exceed the number of samples, as in the case of patient data sets with thousands of variables to only a few hundred samples, then type 1 errors (false positives) likelihood are increased (Korthauer, et al. 2019). When performing multiple comparisons, posthoc tests such as the stringent Bonferroni procedure can be applied to limit false positives, however this is at the cost of statistical power resulting in reduced detection of true positives (Grandhi, Guo and Peddada 2016).

2.6.5.2   Wald test with Benjamini-Hochberg correction

Novogene provided statistical analysis of differential expression of transcripts using the Dseq2 analysis package and Wald test. The Wald test is used for hypothesis testing to compare two groups. First the parameters for the model are estimated. In this case the log fold change is the coefficient for each parameter and is estimated by maximum likelihood. The log fold change is divided by its standard error to produce the z-statistic. The z-statistic is then compared to a standard normal distribution, where a p-value is calculated. The p-value is the probability of the z-statistic being observed by random chance. The lower the p-value, the smaller the probability of the z-statistic being obtained by chance is, and the more likely the change in expression levels is significant. As the test is being carried out on >20,000 transcripts, and the Wald-test used a significance cut-off of P<0.05, there is a 5% chance of false positives, resulting in thousands of 'positives' being by chance. DSeq2 applies the Benjamini -Hochberg (BH) correction to control the false discovery rate (FDR), where transcripts were ranked by reported p-value and then multiplied by n/rank, where n=total number of tests.

BH correction = $p \; x \; \left(\dfrac{n}{rank} \right)$

2.6.5.3   Holm-Sidak correction for multiple comparisons

The Holm-Sidak method corrects multiple comparisons by adjusting the significance level for each test to control the Type I error rate. It ranks p-values, progressively tightening significance levels as comparisons are made. The most significant test is evaluated at the standard significance level (e.g., 0.05), while the subsequent tests are evaluated at more conservative levels to account for the multiple comparisons. If a test's p-value is lower than its adjusted significance level, you can conclude that it's statistically significant. This method increases

chances of avoiding false positives when conducting multiple tests, enabling more reliable inferences while maintaining controlled statistical significance (Holm 1979).

2.6.5.4   Principal component analysis (PCA)

In large data sets, such as in omics, where there are thousands of variables, interpretation is not straight forward. Principle component analysis is a statistical technique which reduces the number of variables of large-scale data, so that it is more interpretable, but the original variance and information of the data is preserved (Jolliffe and Cadima 2016, Ringnér 2008). PCA uses the original data set to produce new variables 'Principal components' (PC), which are linear functions of groups of variables in the starting dataset. PCA identifies which variables explain the largest variance in the data set, with PC1 being responsible for the most, and each subsequent PC representing less variation. The top PCs can be plotted against one another to separate the samples of the cohort. For example, separating healthy controls from a diseased group, or a cohort who have received a treatment vs those who have received a placebo (Zhang, Ping, et al. 2019). Principal component analysis was used to analyse proteomics data to confirm treatment with IFNG and 5AzaC had caused significant changes in protein expression profiles of cell lines compare to untreated.

2.6.5.5   Pairwise linear regression

The relationship between data sets can be examined using pairwise linear regression. Pairwise linear regression can offer more statistical power over the standard T-test, as it utilises multiple comparisons, as opposed to one over all tests. Using the example of transcriptomics for this thesis, the transcriptome expression counts of a treated cell line can be plotted as the dependent variable against the transcriptome of an untreated cell line as the independent variable, example given in Figure 7:
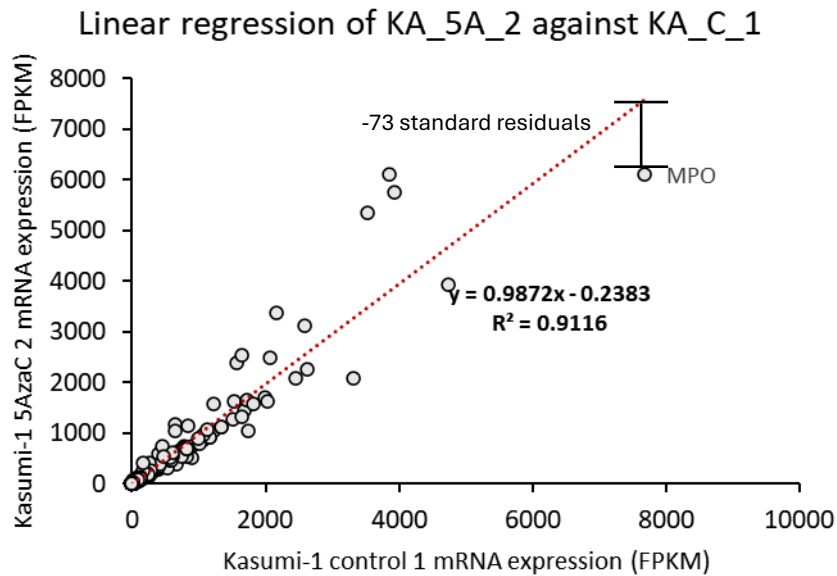
Figure 7: Pairwise linear regression of Kasumi-1 5AzaC repeat 2 against Kasumi-1 untreated control sample 1. Expression of each transcript is from each sample are plotted against each other as x,y coordinates, and a line of best fit is drawn.

From this plot, a line of best fit described by Y=mx+b can be calculated, where m is the slope of the line and b is the Y intercept where the line crosses the Y axis. The coefficient of determination, R^2 indicates the percentage of variance in in the dependent variable (Y axis) that is predicted by the independent variable (X axis). The residual output is then calculated for each observed variable where Residual output= observed y value – predicted y value. Residuals measure the difference between the predicted and observed value, with larger residuals indicating larger deviations from the expected value. A residual can then be divided by the standard deviation of all residuals in the data set to produce a standardised residual (Kim, H. Y. 2019). An observation of 2-3 standardised residuals or more away from the predicted value indicate further investigation warranted to determine if the difference is a result of causation or just chance. In the example above, MPO is 73 standard residuals below predicted value, which could be caused by treatment of cell line with 5AzaC. To further narrow down the list of variables of interest, the regression can be performed multiple times with data generated from biological repeats. A cut off point can be determined, for example the top 500 most deviating variables according to standard residual output, or anything above 3 standard residuals depending on the data obtained. Variables in these lists can be compared, and anything that is repeatedly reported above the cut-off point in every single comparison can be shortlisted for further evaluation. This process is used in chapter 5.

2.6.5.6 Regression for survival analysis

Regression analysis can be used to establish if there is a relationship between a variable, for example smoking, and an event, such as death. Two common types of regression are logistical which deals with categorical variables, and linear, for continuous variables. Linear regression is used for survival analysis, where a continuous variable, for example expression of a gene, can be plotted against survival time. The residuals between the observed data and line of best fit, can indicate how likely it is that the variable influences survival.

2.6.5.7 Cox proportional hazard regression model

The cox proportional hazard regression model (CPH), introduced in 1972, is a popular survival analysis in the biomedical field (Cox 1972). Essentially, the model compares the length of time between study start and an event, such as death, disease relapse or recovery, between two or more groups of participants. The CPH is a semiparametric method that does not make assumptions based on survival time distribution and does not assume covariates impact hazard function, and therefore survival. The hazard function describes the probability that an event will occur at a given time point, it is used to calculate the hazard ratio; which is the ratio of event occurrences at a particular time between two groups. For example, in a drug trial, a hazard ratio of 0.5 could indicate that half as many patients experience an event at any time point, than those in the untreated group.

CPH is also able to assess impact of multiple covariates on outcome and, unlike Kaplan-Meier and the log rank test, can handle censored data. Censoring occurs in survival data if a) the study ends before all participants encounter the event being measured, b) A participant leaves the study c) A participant dies of causes unrelated to the study (Schober and Vetter 2018). Censoring ultimately means an event has not been observed during the duration of the study, and if the patient were to experience the event, it would be to the right of where the patient is censored in the timeline. This is the most common type of censor and is called right 'censoring'. It is accommodated in CPH by including the estimates of survival at time points prior to censoring, and by excluding from analysis afterwards (Schober and Vetter 2018). Left censoring is when a patient experiences the event being observed prior to the study, and therefore is not relevant when death is the event.

Cox proportional hazard regression can be performed with one predictor, termed 'univariate', and multiple predictors, often termed 'multivariate'. Multiple regression is preferred as it gives improved chance of evaluating impact of variables, while also testing interactions with other variables. When performing multiple regression, a common rule is that analysis should be limited to 1 variable per 10-20 events, although other methods to determine the ratio have been

suggested (van Smeden, et al. 2019). Studies have reported that below 10 events per variable caused bias, variability, and unpredictable confidence interval coverage (Vittinghoff and McCulloch 2007).

Ergo, a logical method must be used to narrow down the pool of potential contributors to a CPH model. This is often performed by statistical testing and selecting an acceptable cut off, for example t-tests, P<0.05, or using pairwise linear regression to identify candidates by large standard residuals. From here a model can be built, commonly using one of three entry methods for selection (Smith 2018).

Forward selection: The model starts with no variables and adds them one by one based on which is most statistically significant, until no more variables in the pool are statistically significant.

Reverse selection: The model includes all variables, and then eliminates the least statistically significant variables one by one, only stopping when every remaining variable in the model is statistically significant.

Stepwise selection: A combination of forward and reverse selection. As with forward selection this method starts with no variables and adds the most significant variable to the model. However, after every addition, the model re-evaluates the model, removing any variables that are no longer significant.

2.6.5.8 Kaplan-Meier plots

Kaplan-Meier plots are graphical representations of survival table data, where time is plotted in many small intervals, against survival of patients as a percentage (Kaplan and Meier 1958). Each sample group is plotted as a line, and when a death occurs in the group, the line decreases, creating a stairs effect. Kaplan-Meier is a parametric method, where the following assumptions are made: 1) Censored patients have the same chance of survival as uncensored patients 2) Probability of survival is the same for a patient regardless of time recruited into study, 3) Time of event recorded is accurate (Goel, Khanna and Kishore 2010). At any timepoint, the probability of survival (S_t) is calculated as:

$$St = \frac{Number\ of\ subjects\ alive\ at\ t0 - Number\ of\ subjects\ dead}{Number\ of\ subjects\ alive\ at\ t0}$$

Where t0 is the start of the study. The probability of survival until that time point is calculated by multiplying the current timepoint survival probability by the survival probability of at all time points prior to it (Goel, Khanna and Kishore 2010). In Kaplan-Meier, the null hypothesis is that

there is no statistically significant difference in the survival of the groups compared and is tested by the log-rank test and cox proportional hazard test.

2.6.5.9 Log rank test

The Log rank test is used to compare survival data between different groups, testing the null hypothesis that the probability of a death occurring at a given time point is the same between those groups (Peto and Peto 1972). For all time points the sum of expected number of events in a group (E_x) and observed number of events in each group (O_x) is compared (Goel, Khanna and Kishore 2010).

Log-rank test statistic = $\dfrac{(O_1 - E_1)^2}{E_1} + \dfrac{(O_2 - E_2)^2}{E_2}$

Using the Chi-squared (X^2) table, the value generated in the test, and the degrees of freedom, the significance of the difference between the groups can be calculated. If the test statistic is greater than the critical value in the X^2 table for the desired p-value, then the null hypothesis can be rejected (Goel, Khanna and Kishore 2010). As with CPH the Log-rank test accounts for the whole time of survival, rather than a specific time point, however unlike CPH, it is limited by only being able report significance between groups and not the size or trend of that difference. Plotting the survival curve is beneficial for understanding the data, as the test may report no significance if curves overlap.

2.6.5.10 Receiver operating characteristic analysis

Receiver operating characteristic (ROC) analysis is used to evaluate the accuracy of diagnostic tests, which is defined by their sensitivity, ability to detect true positives, and their specificity, ability to detect true negatives. To plot a ROC curve, sensitivity (Y axis) is plotted against 1-specificity (X axis) across different cut off values of the diagnostic test. Where diagnostic tests values reporting below or equal to the given cut off are classed as negative tests (Zou, O'Malley and Mauri 2007). Potential ROC curves are presented in Figure 8. AUC is used to summarise a diagnostics tests accuracy, with the aim to achieve as close to 1 as possible. In this manner, tests can be compared to one another in their performance, by contrasting the AUC for each ROC, to see which is more accurate.
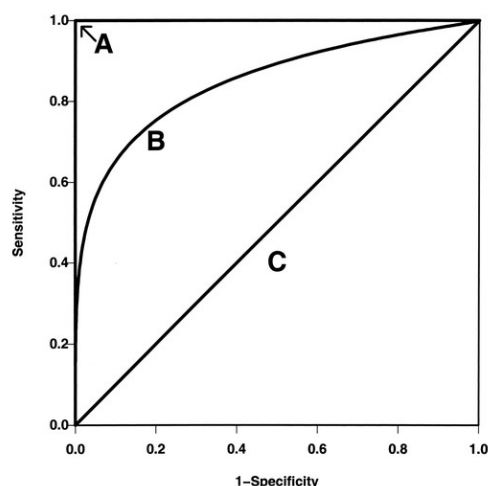
Figure 8: Three example ROC curves are shown. Line A shows the hypothetical 'gold standard' of accuracy where area under curve (AUC) = 1 and hugs the upper left corner of the graph. Line B shows a typical ROC curve with AUC=0.85. Line C is a straight diagonal line at a 45oC angle, which shows what random chance looks like, AUC=0.5. Adapted from Mauri et al.

2.7 Statistical analysis

All figures were generated using GraphPad prism V7. Error bars represent standard deviation between repeats ($p \leq 0.05$ = *, $p \leq 0.01$ = **, $p \leq 0.001$ = *** and $p \leq 0.0001$ = ****). In most cases a minimum of 3 biological repeats were used, with any deviations stated in the text of the experiment. In figure legend, n = x refers to the number of biological replicates used.

2.7.1 Data visualisation by heat map Clustering

Proteins deemed significant were uploaded to MORPHEUS https://software.broadinstitute.org/morpheus/) to visualise results by heat maps. Hierarchal clustering was performed for each cell line set of proteins, using Euclidean distance and complete linkage. This clustering approached helped identification of outliers in repeats.

2.7.2 Online tools and databases

Table 7: Table of databases with online tool name, where it was applied and the link to visit.

Online tool	Application	Link
Metascape	Identification of enriched biological processes	https://metascape.org/gp/index.html#/main/step1
MORPHEUS	Generation of heatmaps and clustering	http://software.broadinstitute.org/morpheus/
PCA online tool	Principal component analysis	https://biit.cs.ut.ee/clustvis/

3 Investigation of immunophenotypes of AML cell lines using IFNG and 5-Azacytidine

3.1 Introduction

Currently, patients under the age of 60 with AML are treated with induction chemotherapies. This treatment involves a 7-day infusion of Cytarabine combined with an anthracycline, such as Daunorubicin, administered on days 1 to 3 (National Comprehensive Cancer Network 2013). This regimen leads to better outcomes in patients below the age of 60 (Boddu, Prajwal Chaitanya, et al. 2017). Patients who achieve remission following chemotherapy then receive post-remission treatments to prevent relapse, commonly including HSCT (de Latour, et al. 2012). The aim of HSCT is to replace AML cells with healthy blood cells, which can repopulate the immune system However, HSCT carries risks, including graft-versus-host disease and high transplant-related mortality. Unfortunately, about half of HSCT patients relapse due to mechanisms that override the antileukemic activity of the transplant, leading to a poor prognosis (Rautenberg, et al. 2019). One potential reason for HSCT failure, is immune escape, which can be induced by multiple pathways, including IFNG signalling.

AML utilises IFNG signalling to upregulate immunosuppressive factors, aiding in immune evasion. This process includes the enhanced expression of molecules such as PD-L1, IDO1, non-classical HLAs, and BST2, which inhibit cytotoxic immune cells and promote tolerogenic phenotypes. Such adaptive immune resistance results in poorer overall survival for patients, underscoring the importance of understanding the impact of IFNG signalling on patient outcomes. AML cell lines with an IFNG-inducible immunosuppressive phenotype could serve as models for developing a related prognostic score to stratify overall survival of patients. It was expected that patients with higher IFNG PI scores would have more immunosuppressive AML and consequently have poorer outcomes. The roles of IFNG and 5AzaC are discussed in subsequent sections.

3.2 IFNG is utilised by AML to escape the immune system

IFNG signalling is activated when the IFNG cytokine binds to its corresponding receptors on cell surfaces. This binding causes the receptor subunits to dimerise and rearrange, leading to the activation of receptor-associated JAKs, which auto-phosphorylate and subsequently phosphorylate STAT proteins. The phosphorylated STAT proteins form homodimers, translocate

to the nucleus, and initiate the transcription of IFNG-stimulated genes (ISGs). These ISGs then direct the IFNG signalling cascade as depicted and described in Figure 9.
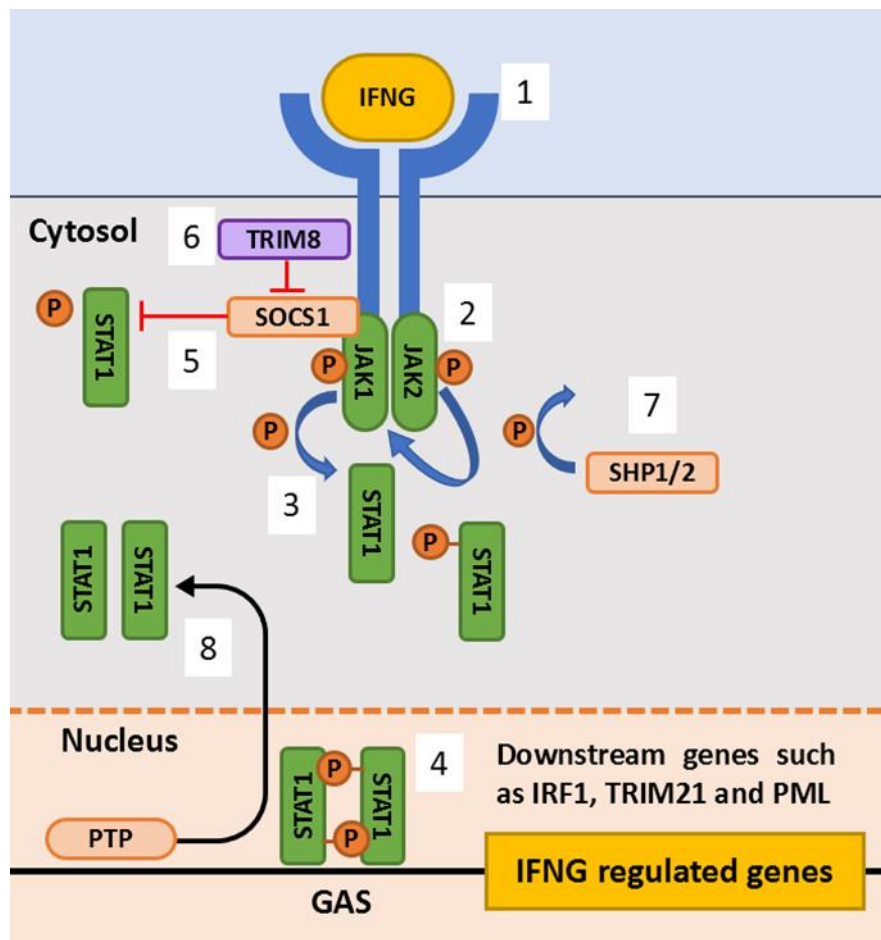
Figure 9: Diagram depicting the IFNG signalling process. IFNG signal transduction is initiated by the IFNG cytokine binding to the IFNG receptor (IFNGR), causing a conformational change (1). Shape change of the receptor triggers the Jak2 kinase to auto phosphorylate and subsequently phosphorylate Jak1 (2). Once activated, Jak1 creates two adjacent docking sites for STAT1 on the IFNGR by phosphorylating tyrosines on each IFNGR1 chain. STAT1 then docks here and is phosphorylated, it then forms a homodimer (3). The phosphorylated STAT1 homodimer undocks from the receptor and relocates to the nucleus where it binds the 'Gamma-interferon-activation sites' (GAS) elements to initiate the IFNG regulated transcription program, which produces genes such as IRF1, TRIM21, TRIM8 and PML (4). SOCS1/2 are proteins which bind to JAK1/2 to prevent STAT1 phosphorylation and thereby inhibit IFNG signalling (5). TRIM8 which is also a product of the IFNG signalling pathways targets SOCS1 for degradation to control its repressive effects on signalling (6). Finally, protein tyrosine phosphatases such as SHP1 and SHP2 disrupt IFNG signalling in the cytosol by preventing STAT1 phosphorylation by dephosphorylating JAK1/2 and IFNGR1 (7). Additionally, PTPs in the nucleus dephosphorylate incoming STAT1 homodimers and export them from the nucleus (8). Figure adapted from **(Schroder, et al. 2004, Toniato, et al. 2002),** *blue arrows show movement of phosphoryl groups between proteins, black arrow indicates export and red flathead shows inhibitory effect.*

AML manipulates IFNG signalling to upregulate escape mechanisms, creating an immunosuppressive environment and promoting immune escape, a process termed 'adaptive immune resistance (AIR).' These escape mechanisms contribute to drug resistance by disabling the immune system and preventing aspects of the cancer immune cycle, such as T-cell activation. For example, chemotherapy destroys AML cells and releases antigens, which the immune system can use to target remaining AML cells. However, if T-cells primed with these antigens are inhibited or unable to proliferate, many AML cells can escape, reducing the efficacy

of chemotherapy. Furthermore, the escaped AML cells may be adapted to evade the immune system and proliferate, creating more chemotherapy-resistant AML cells (Chen, D. S. and Mellman 2017).

Chronic IFNG signalling is known to induce immunosuppressive pathways in AML. This chapter investigates three key mechanisms: the programmed cell death protein 1, programmed death-ligand 1 (PD-1/PD-L1) pathway, Indoleamine 2,3-dioxygenase 1 (IDO1), and non-classical human leukocyte antigen (HLA) class I molecules.

3.2.1 Programmed death ligand 1 (PD-L1) mediated immune escape

PD-1 is a well-studied negative checkpoint in cancer. PD-L1 expression is induced by IFNG signalling when the STAT1 homodimer binds to the gamma interferon activate sites on DNA in the nucleus to activate transcription of several genes (Figure 9). Among those is IRF1, which then binds to the promoter of the PD-L1 gene, activating its transcription (Garcia-Diaz, et al. 2017). This increases PD-L1 expression on the surface of AML cells which then bind to PD-1 on tumour-specific T-cells. This interaction transmits an inhibitory signal that suppresses T-cell proliferation, facilitating immune evasion (Ribas 2015). Modest constitutive PD-L1 expression has been identified in most myeloid and lymphoid cell lines, excluding THP-1. Elevated PD-L1 levels have been reported in AML patients with poor survival (Wang, F., et al. 2022). Furthermore, IFNG-inducible PD-L1 expression has been shown to reduce CTL-mediated lysis and promote immune escape (Berthon, et al. 2010).
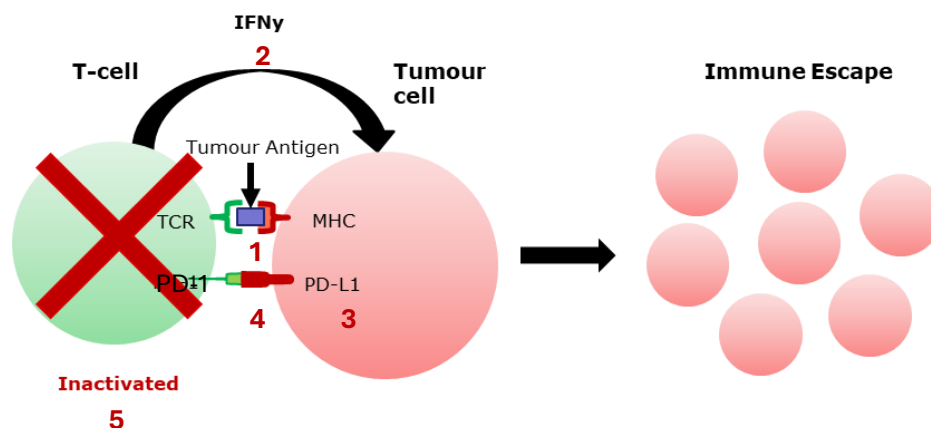
Figure 10: Diagram of Interferon mediated upregulation of PDL-1 . When T-cells bind via their TCR to the antigen presenting MHC of the tumour cell (1), T-cells secrete IFNG (2), resulting in abnormal upregulation of PD-L1 on the tumour cell (3). PD-L1 on tumour cells binds to the PD1 molecules expressed on T-cell (4). This initiates a signalling cascade that inhibits the T-cell from functioning (5) and allows the tumour cell to escape T cell mediated cytotoxicity.

3.2.2 Indoleamine 2,3-dioxygenase-1 (IDO1) mediated immune escape

IFNG induces IDO1 expression primarily through the JAK/STAT pathway. JAK phosphorylates STAT1, which then dimerises and moves to the nucleus to bind the GAS-2 and GAS-3 sites upstream of the IDO1 gene, activating its expression (Huang, et al. 2022). Additionally, IFNG and STAT1 indirectly boost IDO1 expression by inducing IRF-1 synthesis, which binds to the ISRE-1 and ISRE-2 sites. Notably, IRF-1 has been found to be more impactful for inducing IDO1 expression than STAT1 (Robinson, Cory M., Hale and Carlin 2005). IDO1 is an enzyme involved in the kynurenine (Kyn) production pathway and is commonly expressed by solid tumours. Kynurenine binds to the aryl hydrocarbon receptor on T-cells and dendritic cells, prompting naïve CD4+ T-cells to differentiate into regulatory T-cells and causing dendritic cells to adopt a tolerogenic phenotype, thereby facilitating immune escape (Platten, et al. 2015). A study on cultured leukemic blast cells from childhood AML found that IFNG induced IDO1 in approximately half of the samples, which was associated with a poorer prognosis (Folgiero, et al. 2014)Additionally in adult AML patients, a higher serum kynurenine/tryptophan (Kyn/Trp) ratio, indicative of elevated IDO activity, has been linked to decreased survival (Corm, et al. 2009). IDO1 overexpression has also been associated with regulatory T-cell phenotypes in CN-AML, correlating with poor prognosis (Arandi, et al. 2018).
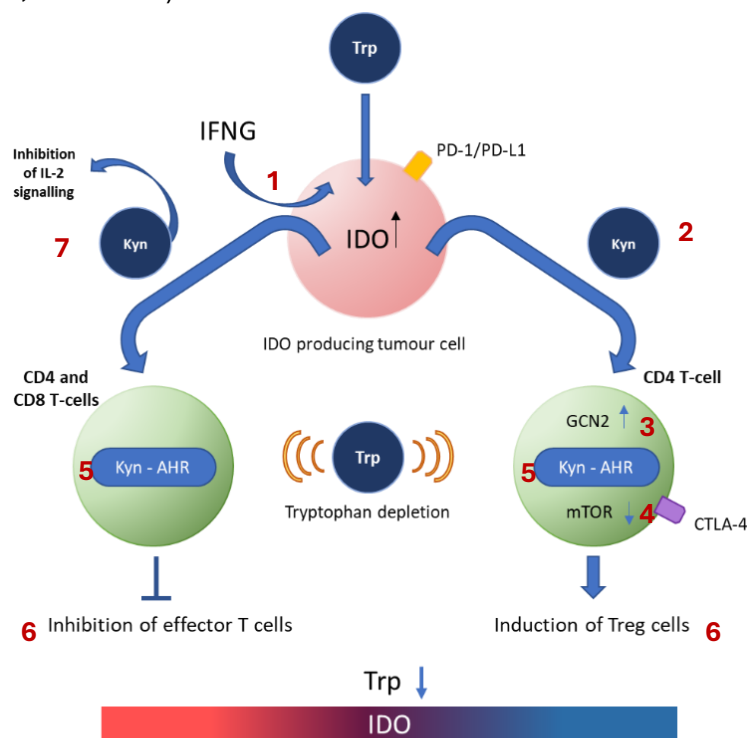
Figure 11: Diagram showing how IDO1 expression on myeloid cells promotes an immune suppressive micro-environment: IFNG induces the production of IDO/TDO in the tumour cell (1). IDO/TDO catalyse the first step in the breakdown of Trp; producing Kynurenine (Kyn) (2). CD4+ T-cells react to decreased Trp by preventing protein synthesis through increased GCN2 (3) and inhibiting cell growth by decreasing mTOR (4). CD4+ T-cells sense the Kyn increase by its binding to AHR, and start inflammation signalling (5), triggering favoured differentiation into Treg cells over T-effector cells (6). In addition, Kyn inhibits IL-2 signalling which inhibits survival of memory CD4 T-cells (7).

3.2.3 Non-classical HLA-G mediated immune escape

Classic major histocompatibility complexes (MHC) such as HLA-A, HLA-B and HLA-C present antigens for recognition by T-cells to mount immune responses to pathogens or other threats. Meanwhile, nonclassical MHC such as HLA-G and HLA-E are often overexpressed by AML to prevent immune cells from killing them (Kochan, et al. 2013, Halenius, Gerke and Hengel 2015). HLA-G inhibits CD8+ T-cells and NK cells, induces CD4+ immunosuppression, and triggers apoptosis of activated CD8+ T-cells and NK cells (Lin, A. and Yan 2018). Its expression has been linked to immune tolerance in AML patients, with significantly higher levels of shed HLA-G reported in the serum of those with more advanced AML, particularly in the FAB-M4 and M5 subtypes (Gros, et al. 2006). Moreover, stimulating cells from FAB-M4 AML patients with IFNG was observed to increase the secretion of HLA-G (Gros, et al. 2006). AML patients expressing HLA-G have been found to also have significantly higher leukaemic blasts in bone marrow than patients that were negative for HLA-G. Furthermore, the percentage of T cells is lower in HLA-G positive patients, indicating a poorer prognosis (Yan, et al. 2008).

3.2.4 Non-classical HLA-E mediated immune escape in AML

Surface expression of HLA-E is induced by IFNG signalling via the JAK/STAT1 pathway and further promoted by overactive proteosome processing (Zheng, et al. 2023). HLA-E promotes immune suppression in AML by interacting with the inhibitory receptor NKG2A on NK cells, deactivating them (Sullivan, et al. 2008). In healthy cells, HLA-E expression depends on HLA class I leader peptides, and a decrease in HLA class I reduces HLA-E expression. However, in AML, HLA-E is expressed even without HLA class I. Normally, NK cells detect HLA class I through HLA-E expression, and a lack of HLA-E indicates reduced HLA class I, triggering NK cell-mediated lysis. AML cells exploit this by downregulating HLA class I and upregulating HLA-E to evade NK cell-mediated lysis (Nguyen, et al. 2009). Increased HLA-E expression has been observed on primary blasts isolated from AML patients when exposed to IFNG producing NK cells (Nguyen, et al. 2009). Additionally, high IFNG pathway signalling in AML patients with the del7/7q mutation and monocytic differentiation correlates with HLA-E expression, and these HLA-E expressing AML are found closer to T cells than their HLA-E negative counterparts (Wang, B., et al. 2024).

3.2.5 Disruption of methylation is associated with AML development and outcomes

Disruption of enzymes involved in DNA methylation pathways affects haematopoiesis and contributes to initiation and progression of AML (Yang, X., Wong and Ng 2019b). Clonal haematopoiesis refers to the presence of genetically distinct hematopoietic stem cells (HSCs)

within an individual, often harbouring somatic mutations. Specifically, enzymes involved in DNA methylation, such as DNMT3A and TET2, are frequently mutated (Figueroa, Lugthart, et al. 2010, Yang, X., Wong and Ng 2019b). These mutations impair differentiation, enhance self-renewal, and promote clonal expansion(Tadokoro, et al. 2007, Yang, X., Wong and Ng 2019b). Consequently, mutated HSCs outcompete healthy HSCs, leading to clonal expansion and the emergence of pre-leukemic clones. Over time, these pre-leukemic clones accumulate additional mutations, including those affecting cellular signalling pathways (e.g., JAK2, TP53). The competitive advantage of mutated HSCs allows them to persist, leading to the transformation into leukemic stem cells (LSCs)(Yang, X., Wong and Ng 2019b). LSCs outcompete HSCs through multiple strategies. They flourish in a pro-inflammatory environment by releasing TNF-α, which boosts NF-κB activity, supporting their survival and growth (Kagoya, et al. 2014). They also evade the immune system by increasing TIM3 expression, emitting chemokines, and, inducing inflammatory secretome helping them avoid immune detection and enabling them to proliferate (Niu, Peng and Liu 2022). Furthermore, LSCs manipulate the bone marrow niche by interacting with BMSCs (Bone Marrow Stromal Cells) to enhance their survival and resistance to treatment, while metabolic reprogramming enables them to efficiently utilise resources, promoting their growth and resistance to therapy (Moschoi, et al. 2016, Chen, Wen-Lian, et al. 2014, Niu, Peng and Liu 2022). LSCs drive the production of leukemic blasts, ultimately culminating in AML(Yang, X., Wong and Ng 2019b).

5AzaC is an analogue of the nucleoside-based cytidine. It is a drug which, administered in low doses, inhibits DNA methyltransferase 1 (DNMT1) and in high doses incorporates into DNA and RNA, causing cell death (Frosig 2015). DNMT1 manages methylation during DNA replication, it favours methylation of hemi-methylated sites and is known as the 'maintenance' enzyme as it restores methylation sites post DNA replication (Ambrosi, Manzo and Baubec 2017). Although DNMT1 is the highest expressed DNMT in dividing cells, DNMT3A and DNMT3B need to be considered as they are not targeted by 5AzaC and play a more active role in methylation. DNMT3A/B are *de novo* DNMTs, meaning they catalyse addition of methyl groups to unmethylated DNA, as opposed to DNMT1, which maintains methylated sites. Therefore, while treatment with 5AzaC prevent maintenance of already established methylation sites, DNMT3A/B are still able to continue creating new methylation sites (Figure 12). Studies have shown that AML methylation profiles are linked to patient outcomes (Yang, Wong and Ng 2019).
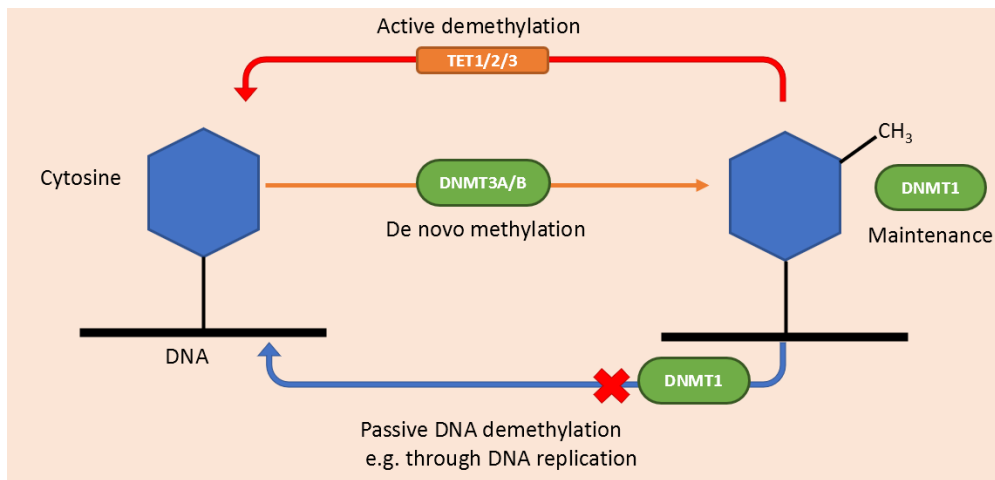
Figure 12:A diagram of methylation status maintenance by DNMT1, DNMT3A and DNMT3B and TET1/2/3. Adapted from Ambrosi et al paper (Ambrosi, Manzo and Baubec 2017). Methyl groups are added to the fifth carbon atom on cytosine bases of DNA by DNMT3A and DNMT3B. Pre-established methylated sites are maintained by DNMT1. In the absence of DNMT1, methylated sites can be lost passively through cell division. TET enzymes actively remove methyl groups through conversion to oxidised derivatives, which are consequently removed by DNA repair or lost passively.

To develop prognostic signatures, cell lines were treated with 5AzaC to inhibit methylation and observe changes in mRNA expression. Cell lines were treated with 5AzaC alone and in combination with IFNG to determine if demethylation affected the cell lines' responses to IFNG at the transcript level. The expectation was that demethylation would induce further inflammatory responses to IFNG treatment and could be related to poorer outcomes in patients.

3.3 Objectives of this thesis

Cytarabine and Daunorubicin are still the backbone of AML therapy after many decades (Blair 2018). Further research into how to improve patient outcome using these chemotherapies and to identify patients who are most likely to respond is needed. IFNG is an essential component of the anticancer response which could have negative interplay with chemotherapy in treating AML, as well as the success of HSCT (Bernasconi and Borsani 2019, Kong, et al. 2016). Cell line models provide an easily accessible way to study the relationship between AML, IFNG, chemotherapy and HSCT.

This thesis seeks to generate prognostic signatures based on AML cell lines treated with IFNG and 5AzaC and test if 'high' IFNG scores are associated with poorer outcome in patients. To do this, cell lines must first be selected that exhibit IFNG signalling and IFNG mediated immunosuppressive phenotypes, before proceeding to RNA seq (chapter 4) and score generation and testing (chapter 5).

In this chapter, four AML cell lines were treated with IFNG to induce IFNG signalling and upregulate expression adaptive immune resistance (AIR) molecules. Cells were also treated with 5AzaC, a hypomethylating drug used to treat AML in combination with cytarabine for two reasons (Huls 2015). Firstly, hypomethylating agents are used in standard clinic for AML and MDS treatment, especially in older and less medically fit patients (Stomper, et al. 2021). Secondly, it is known that the process used to immortalise cell lines induces significant DNA methylation, and therefore silencing at gene promoters, causing global shifts in gene expression and divergence from the original transcriptional state. The resultant changes in RNA and protein expression measured post treatment, give insight into possible pathways activated by IFNG and 5AzaC treatment, acting as a model for hypomethylation and inflammation within AML cells. The final aim was to use AML cell lines models to derive a prognostic score to predict patient response to induction therapy.

3.4 Results

3.4.1 IFNG and demethylation treatment validated prior to characterising cell lines

IFNG signalling and 5AzaC mediated demethylation were validated by western blot- prior to further characterisations (Figure 13). Cells were seeded at 0.5 x 10^6 cells/mL and suspended in media with either:

1. 5 ng/mL IFNG and harvested after 30 minutes.

2. 0.5 µM of 5AzaC at 0 and 24 hours, and then harvested after 48 hours of 5AzaC treatment.

3. 0.5 µM of 5AzaC at 0 and 24 hours, followed by 5 ng/mL IFNG at 48 hours, and then harvested 30 minutes later.

IFNG was expected to increase the expression of STAT1, indicating the activation of the IFNG JAK/STAT1 signalling pathway. Treatment with 5AzaC was expected to reduce the expression of DNMT1, demonstrating that the drug effectively degraded DNMT1. Results showed an increase in STAT1 protein expression with IFNG treatments and decrease in DNMT1 expression with 5AzaC treatment in both cell lines.
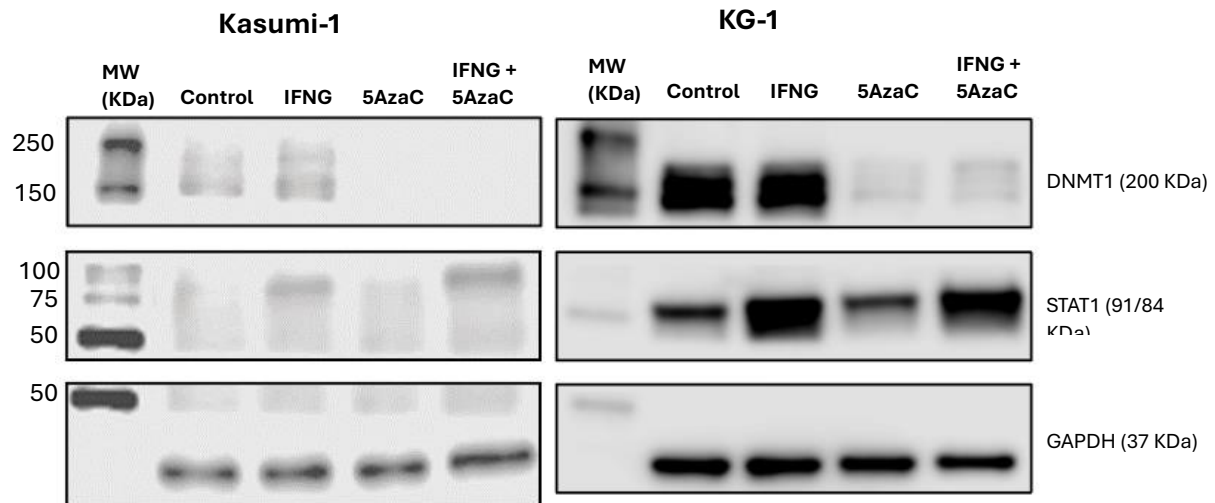
Figure 13: Western blot showing IFNG induced STAT1 expression and 5AzaC dependent degradation of DNMT1 in AML cell lines. Cell lysates were generated from Kasumi-1 and KG-1 cells both untreated and treated with 5 ng/ml IFNG, 0.5 µm 5AzaC at 0 and 24 hrs, or a combination of both. Expression of DNMT1 and STAT1 are shown alongside the loading control GAPDH. Protein was loaded at 30 µg per sample.

3.4.2 Analysis shows AML cell Lines express differential IFNG response and immunosuppressive phenotypes

IFNG response and immunosuppressive phenotypes was assessed by measuring known IFNG regulated molecules using various methods. This served to characterise the AML cell lines in the context of an inflammatory IFNG producing environment and assess them as models for immune resistance concurrently. Molecules measured included the AIR molecules as well as a downstream product of IDO1/TDO2, kynurenine. First cell lines were assessed for induction of genes for IDO1 (*IDO1*), TDO2 (*TDO2*) and PDL1 (*CD274*) in response to treatment with 100ng/mL IFNG using RT-PCR (Figure 14).
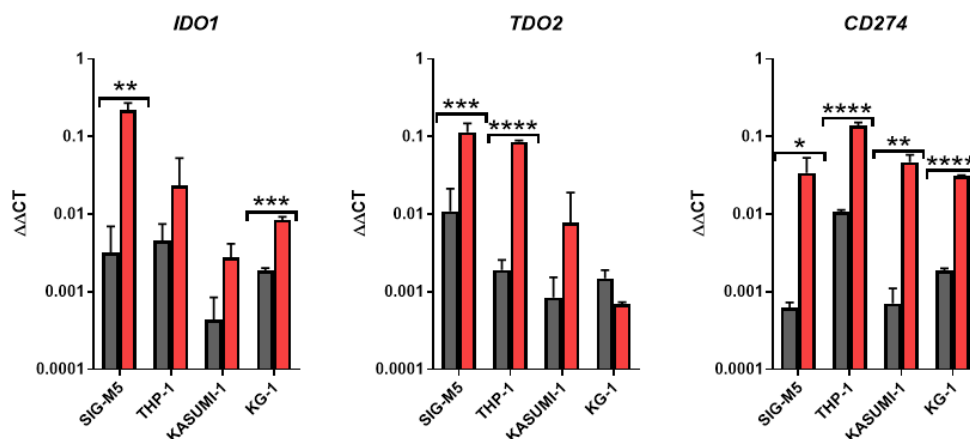
Figure 14: Expression of IDO1, TDO2 and CD274 in AML cell lines after treatment with IFNG. Cell lines were treated with IFNG (100 ng/mL) for 48 hrs. All genes were measured using the quantitative real-time PCR and 2-ΔΔT method (Schmittgen and Livak 2008). Gene expression was normalised against the housekeeping gene, GUSB. Grey = Control, Red = IFNG treated samples. Statistical testing done by the Holm-Sidak multiple comparisons method, * = P< 0.05, ** = P< 0.01, *** = P<0.001 and **** = P<0.0001, n=3-5. Error bars indicate Standard deviation.

All of the cell lines reported an IFNG induced upregulation of *IDO1* (SIG-M5 Ctrl = 0.0030, IFNG = 0.2176 ΔΔCT, THP-1 Ctrl = 0.0044, IFNG = 0.0233 ΔΔCT, KAS-1 Ctrl = 0.0004, IFNG = 0.0027 ΔΔCT, KG-1 Ctrl = 0.0018, IFNG = 0.0080 ΔΔCT). However, upregulation of *IDO1* was only significant in SIG-M5 and KG-1 cell lines (SIG-M5 P<0.01 and KG-1 P<0.001). The increase in IDO1 levels in the SIG-M5 and KG-1 cell lines was confirmed by a significant rise in kynurenine production, a downstream product of IDO1, in both lines (SIG-M5 Ctrl = 1.9 μM, IFNG = 37.46 μM ,P<0.01, and, KG-1 Ctrl = 0.12 μM, IFNG = 1.07 μM, P<0.01) (Figure 15). *TDO2* was significantly upregulated in

both SIG-M5, and THP-1 cell lines (SIG-M5 Ctrl = 0.0105 ΔΔCT, IFNG = 0.1068 ΔΔCT, P<0.001, and THP-1 Ctrl = 0.0018 ΔΔCT, IFNG = 0.0837 ΔΔCT, P<0.0001).
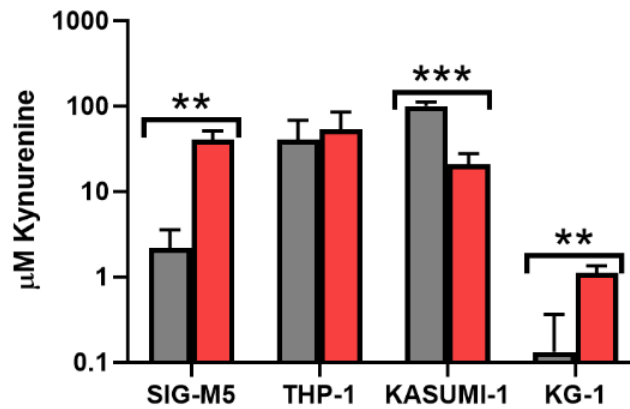
*Figure 15: Expression of Kynurenine (uM) in AML cell lines after treatment with IFNG. Cell lines were treated with 100 ng/mL IFNG for 48 hrs, measured by Ehrlich's reagent, n=3. Kynurenine was extracted using 30% TCA, from samples, controls, and standards. The concentration of kynurenine samples was estimated using a standard curve of absorbance values for set kynurenine concentrations. Grey = Control, Red = IFNG treated samples. Statistical testing done by the Holm-Sidak multiple comparisons method, * = P< 0.05, ** = P< 0.01, *** = P<0.001 and **** = P<0.0001, n=3. Error bars represent standard deviation.*

A significant increase in expression of *CD274* in response to IFNG treatment was observed in all cell lines, see Figure 14 (SIG-M5 Ctrl = 0.0006 ΔΔCT, IFNG = 0.0321 ΔΔCT, P<0.05, THP-1 Ctrl = 0.0010 ΔΔCT, IFNG = 0.1310 ΔΔCT, P<0.0001, KAS-1 Ctrl = 0.0007 ΔΔCT, IFNG = 0.0438 ΔΔCT, P<0.01, KG-1 Ctrl = 0.0017 ΔΔCT, IFNG = 0.0287 ΔΔCT, P<0.0001). This change was also reported by cell surface staining using flow cytometry, Figure 16.
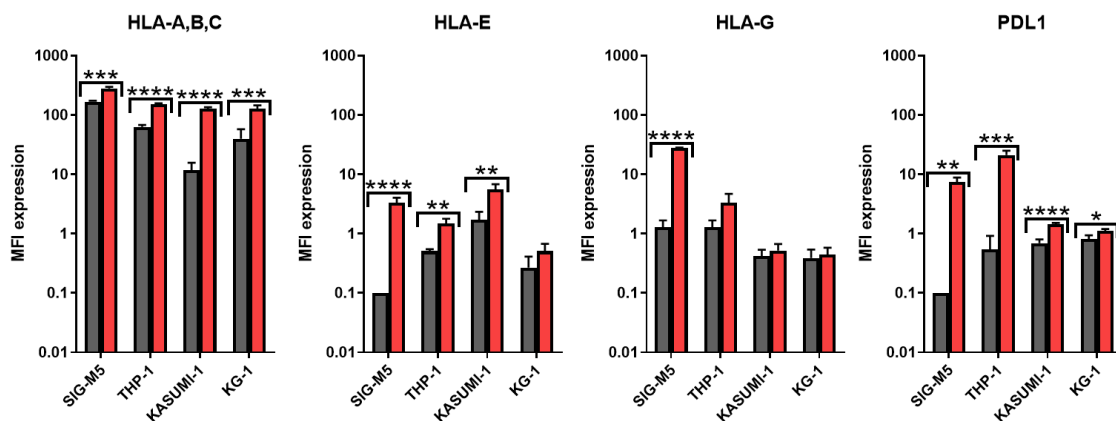
*Figure 16: Expression of surface proteins HLA-A,B,C, HLA-E, HLA-G and PDL1 on AML cell lines before and after IFNG treatment. Cell lines were treated with IFNG (100 ng/mL) for 48 hrs, expression was measured by cell surface staining flow cytometry, n=3. Grey = Control, Red = IFNG treated samples. Statistical test: Holm-Sidak multiple comparisons method, * = P< 0.05, ** = P< 0.01, *** = P<0.001 and **** = P<0.0001, n=3-4. Error bars represent standard deviation.*

All cell lines significantly upregulated surface expression of the PD-L1 protein after IFNG treatment (SIG-M5 Ctrl = 0.10 MFI, IFNG = 6.96 MFI, P<0.01, THP-1 Ctrl = 0.51 MFI, IFNG = 19.70 MFI, P<0.001, Kasumi-1 Ctrl = 0.62 MFI, IFNG = 1.35 MFI, P<0.0001 and KG-1 Ctrl = 0.77 MFI, IFNG = 1.04 MFI, P<0.05). In addition, highly significant upregulation of the classical immune response

MHC HLA-A, B and C was detected in all cell lines, indicating functional classic IFNG response in all samples (SIG-M5 Ctrl = 158.11 MFI, IFNG = 270.10 MFI, P<0.001, THP-1 Ctrl = 61.03 MFI, IFNG = 148.98 MFI, P<0.0001, Kasumi-1 Ctrl =11.20 MFI, IFNG = 120.97 ,P<0.0001 and KG-1 Ctrl = 35.72 MFI, IFNG = 120.97, P<0.001). The largest increase was seen in the Kasumi-1 cell line, with an approximate 10-fold increase. All cell lines except for KG-1 significantly upregulated HLA-E in response to IFNG (SIG-M5 Ctrl = 0.09 MFI, IFNG = 3.12 MFI, P<0.0001, THP-1 Ctrl = 0.46 MFI, IFNG = 1.40 MFI , P<0.01, Kasumi-1 Ctrl = 1.62 MFI, IFNG = 5.17 MFI, P<0.01) and HLA-G was only significantly upregulated in the SIG-M5 cell line (SIG-M5 Ctrl = 1.24 MFI, IFNG = 26.53 MFI, P<0.0001). These results show all cell lines displayed IFNG response, but with varying degrees of induction of AIR targets. The gating strategy used to obtain this data can be found in the methods section. Example of histogram results are shown in Figure 17.
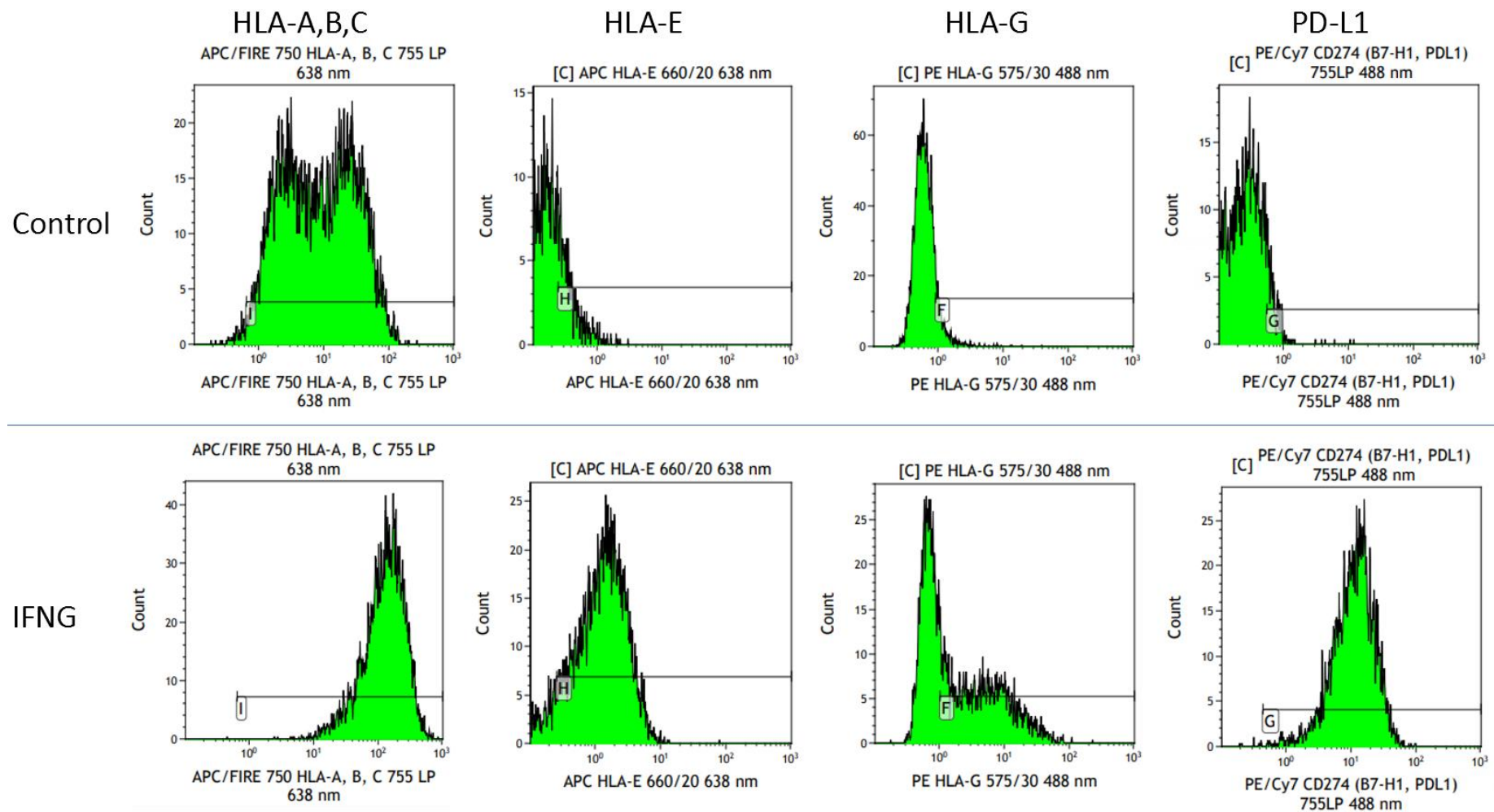
Figure 17: Examples of Histograms and Gating Strategy Used to Generate Data in Figure 19 . Gates were set on unstained samples as described in methods section. This figure shows histograms for HLA-A,B,C (Kasumi-1), HLA-E (KG-1), HLA-G (SIG-M5) and PD-L1 (THP-1) in untreated control cells (top row) and cells treated with IFNG (100 ng/mL) for 48 hrs (bottom row). Antibodies, band pass filter and excitation wavelength used are displayed on the X-axis for each graph.

3.4.3 Global Proteomic profiles demonstrate varied response to IFNG treatment

Assessment of immunosuppressive molecules upregulated by IFNG reported all cell lines to have immunosuppressive AIR responses, with the SIG-M5 and THP-1 cell lines displaying similar profiles. The SIG-M5 cell line was the only cell line to express all AIR molecules, suggesting a more immunosuppressive IFNG induced phenotype than its counterparts. However, looking at only specific molecules associated with a particular response paints a limited picture. To complement this data and gain a broader and more detailed view of responses, proteomic profiles of IFNG treated cell lines were compared to their untreated control counterparts, as well as each other. Cell lines were treated with 100 ng/mL of IFNG, and harvested at 48 hrs to be analysed by SWATH-MS.

Cell lines reported between 2,775 and 3,390 proteins differentially expressed in response to IFNG treatment (Figure 18). The SIG-M5 and THP-1 cell lines showed the most proteins to be differentially expressed between control and IFNG treated condition, with a similar number of proteins up and down regulated. The Kasumi-1 cell line reported approximately 300 fewer proteins differentially expressed than SIG-M5 and THP-1, followed by KG-1 with the smallest breadth of response at approximately 600 less proteins expression changed than the leading THP-1 cell line.
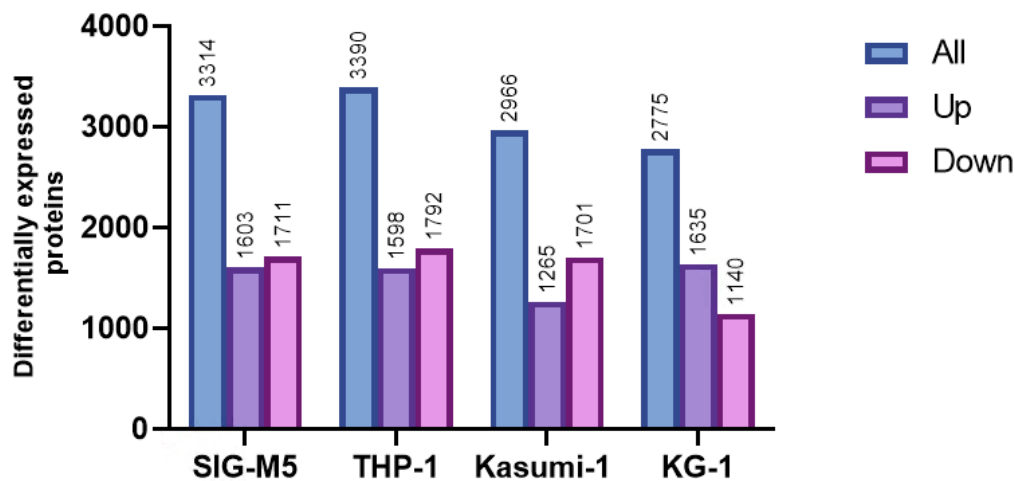
Figure 18: Differential protein expression in AML cell lines in response to IFNG treatment. Cell lines were treated with IFNG (100 ng/mL) for 48 hrs and protein expression measured by SWATH-MS and compared to untreated controls. Total number was also broken down into how many of those proteins are up or down regulated with IFNG treatment, n=5/6.

3.4.3.1 Pearson correlation and principal component analysis of proteomics show variation between cell lines and treatment

The normalised proteomic profiles reported for samples were assessed for correlation to determine variability between cell line response to IFNG, as well as identify any replicate outliers. The Pearson correlation coefficient was calculated in Graphpad Prism V using pairs of samples and visualised as a heat map in Morpheus (Figure 19). The clustering groups cell lines together based on if cells were untreated or treated with IFNG. All cell line and treatments clustered neatly except for KG-1.
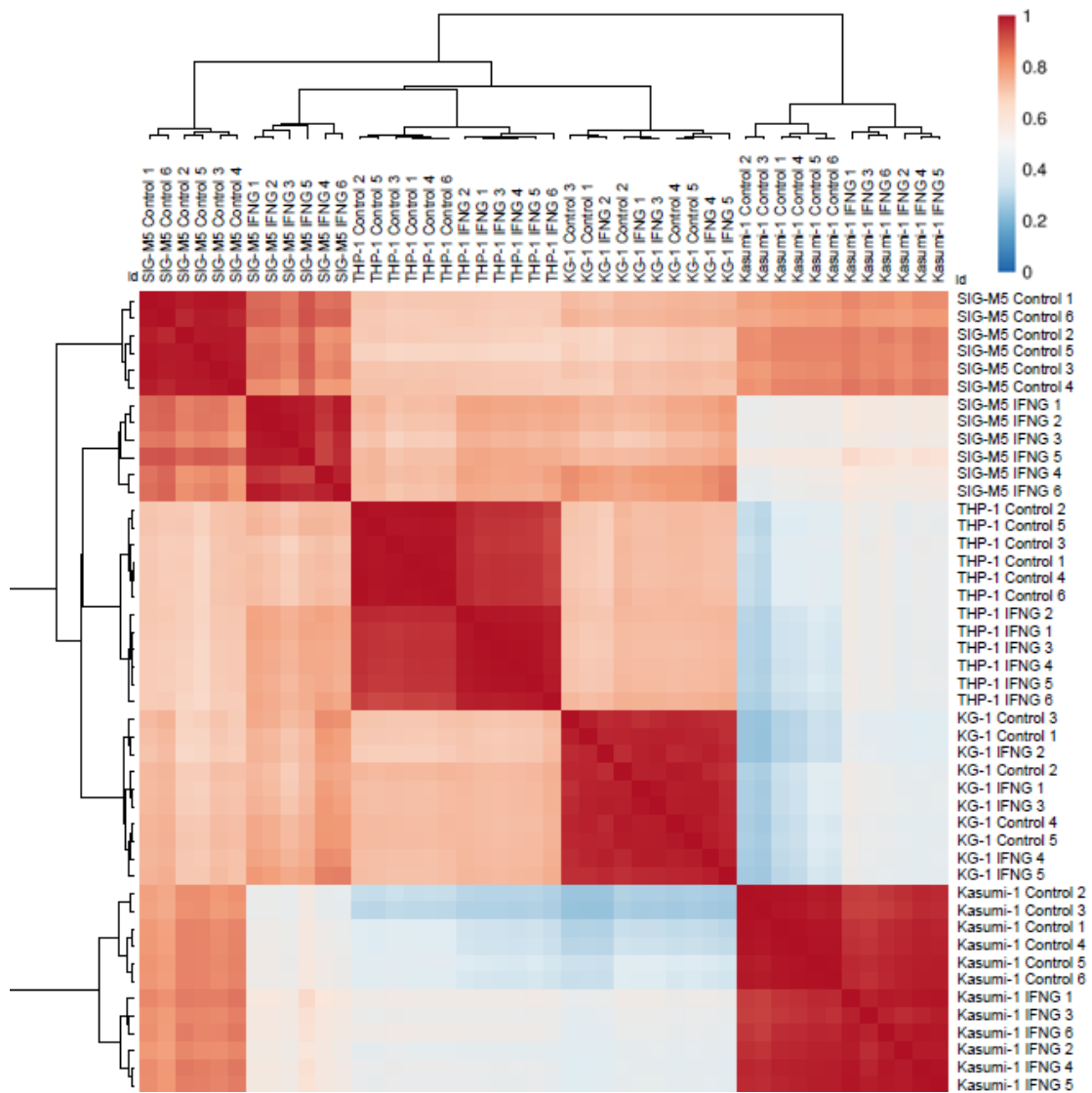
Figure 19: Pairwise comparisons of Pearson correlation coefficients between proteomics profiles of untreated control and IFNG treated AML cell lines. Coefficients were calculated in GraphPad prism V5, then visualised and clustered using the Morpheus online tool https://software.broadinstitute.org/morpheus/). Clustering was performed with Euclidean distance and complete linkage. Blue indicates low correlation and red indicates higher correlation between groups.

Proteomic data was also analysed by principal component analysis (PCA), see Figure 20. Conversion of proteomics data to principle components (PCs) confirmed the grouping of cell lines that were observed by the Pearson correlation heat map. Proteomic profile data was used

to generate a 2D PCA plot, where PC1 accounted for 23.2 % and PC2 for 17.3 % of variance in the samples (Figure 20 A). The plot confirmed cell line replicate fidelity, where samples only overlapped with samples from the same cell line. There was no overlap between SIG-M5 control cells, and IFNG treated SIG-M5 cells, indicating they are different in their profiles. There was some overlap of control and IFNG treated Kasumi-1 and KG-1, indicating while IFNG induced a change in these cell lines, the change was not as large as SIG-M5, when measured as a function of PC1 and PC2. Lastly, IFNG treated THP-1 overlapped with its untreated control suggesting that treatment with IFNG did not induce a large shift in its proteomic profile. However, PC1 and PC2 only represent 40.5% of variability in the whole data sample which includes 4 cell lines. PC1 and PC2 may better demonstrate the variability between SIG-M5 and IFNG treated SIG-M5, than between the other cell lines and their IFNG treated counterparts. When cell lines were compared by PC3 and PC4 (Figure 20 B), then Kasumi-1, SIG-M5 and KG-1 control and IFNG treated profiles were separated.

Figure 20: Principle component analysis of proteomic profiles of untreated control and IFNG treated cell lines. A) 2D Principle component analysis of cell line proteomic profiles performed using ClustVis (https://biit.cs.ut.ee/clustvis/). A) Shows proteomic profiles as summarised by PC1 and PC2, which account for a total of 40.5% of the data variability. B) Shows proteomic profiles summarised by PC3 and PC4, which represent a total of 15.9% of the data variability. Prediction ellipses are 95% confident that any new observation from cell lines of this treatment, will fall within the ellipse.

3.4.3.2 *IFNG induced changes to all four AML cell lines proteomic profiles*

All four cell lines tested (Kasumi-1, KG-1, SIG-M5 and THP-1) reported changes to proteomic profiles in response to IFNG treatment, numbers of differentially expressed proteins are shown in Figure 18. These proteins were assessed for fold change and 'confidence' of change. Methods outlined by Lambert et al were used to determine significant proteins of interest using their 'confidence' cut off (Lambert, et al. 2013). Using this method, proteins with confidence above 75% were considered significant, but for the change in expression of a target to be considered 'affected' by IFNG treatment, a fold change >1.5 was required. Proteins between 60 and 75% 'confidence' were also be considered significant if they displayed a fold change of 2 or above. Between all four cell lines, only 485 of the 3,744 proteins detected were significantly differentially expressed according to confidence and fold change cut off. Expression of those proteins was compared in a heat map in Figure 21. Normalised protein expression was Log2 transformed, and z-score values were used to generate heat maps. Hierarchical clustering by Euclidean distance of rows and columns was performed to produce Figure 21. Control and IFNG treated cell line samples clearly cluster based on this selection of proteins. This data has also been presented with row clustering to show each respective cell line control samples and IFNG treated samples side by side, for ease of visual comparison (Figure 21 B). Side by side comparisons between untreated and treated samples of the same cell line show a large breadth of changes in the SIG-M5 and THP-1 cell lines post IFNG treatment. Comparatively, the Kasumi-1 and kg-1 cell lines show fewer visually striking differences between IFNG treated and untreated samples. The four cell lines demonstrate that there are different IFNG induced proteomic profiles present.
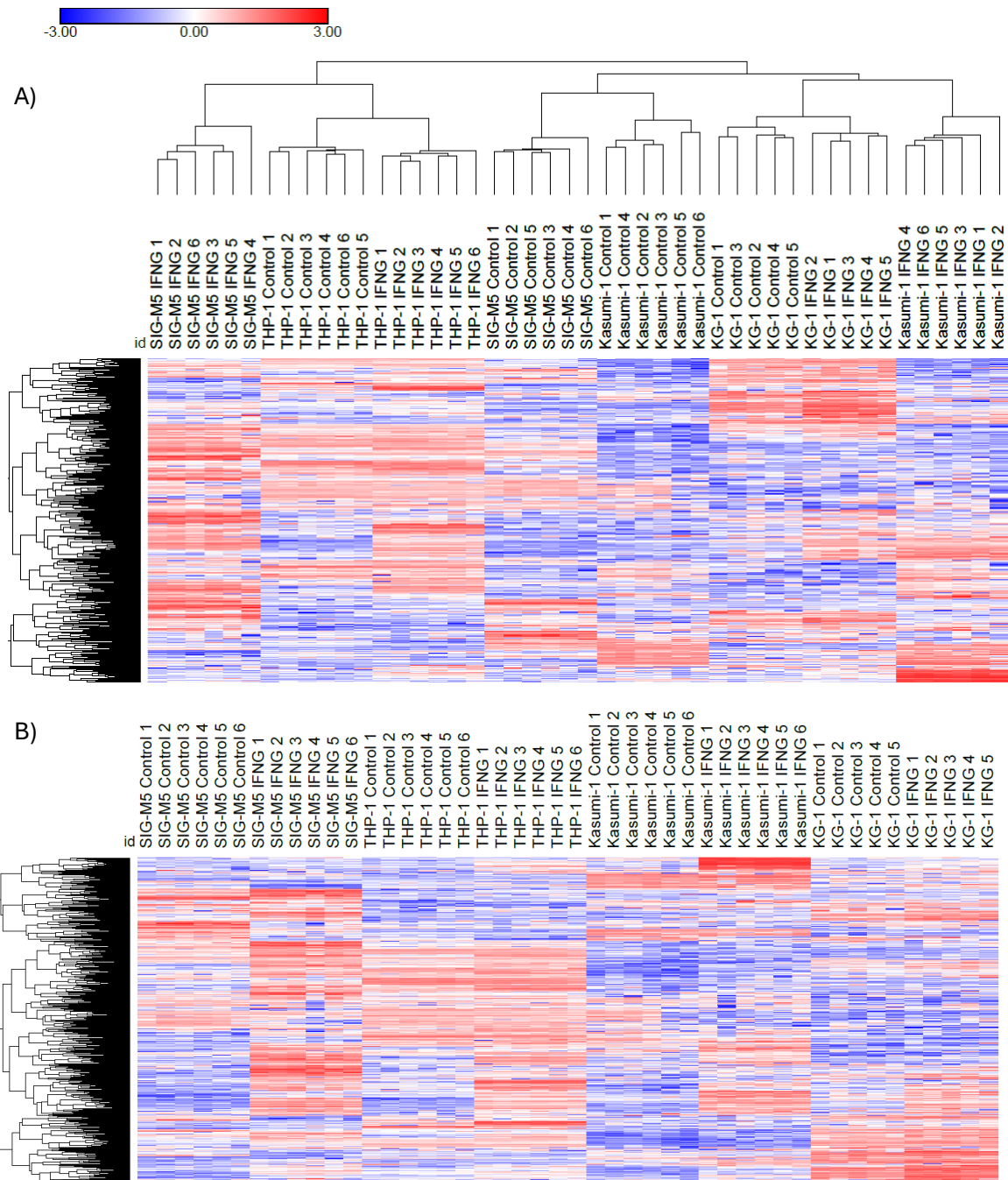
Figure 21: Heat map comparison of differentially expressed proteins across untreated control and IFNG treated AML cell lines. A) Heat map comparing 485 proteins differentially expressed according to confidence and fold change cut offs (>60% confidence, 2-fold change, >75% confidence, 1.5-fold change), Cell lines were clustered by rows and columns to show cell line and treatments successfully cluster, B) Proteomics data of the 485 proteins clustered only on rows, to allow for easier visual comparison between untreated and IFNG treated samples of the same cell line n=5/6. Key for heat maps shown top left for expression values. Heat maps generated in Morpheus (https://software.broadinstitute.org/morpheus/). Blue indicates lower expression; red indicates higher expression.

The number of proteins uniquely differentially expressed in each cell line, as well as commonly changed in multiple cell lines, is shown by Venn diagram in Figure 22. In all cell lines, more proteins were significantly upregulated than downregulated in response to IFNG. The SIG-M5 cell line had the largest number of significantly differentially expressed proteins, followed by SIG-M5, Kasumi-1 and finally KG-1 (341, 141, 101 and 48 respectively). The Venn diagram of upregulated

proteins shown in Figure 22 A shows 177 of the 251 proteins upregulated in SIG-M5 are unique to the cell line. The remaining 74 proteins were upregulated in other cell lines too. SIG-M5 and THP-1 share the most in common upregulated proteins out of all cell line combinations (28 proteins), suggesting they are more similar in response to IFNG, than the other combinations of cell lines. Overall, there were only 13 proteins significantly upregulated in all IFNG treated cell lines, demonstrating the AML cell lines to have unique IFNG response profiles. Far fewer proteins were significantly downregulated in all cell lines, and even fewer shared. The most downregulated proteins in common between cell lines was between SIG-M5 and THP-1, which have 4 proteins in common. However, each cell line had a distinct set of uniquely regulated proteins, indicating their unique response patterns to interferon exposure. To further characterise these cell lines beyond the number of proteins differentially expressed by IFNG treatments, lists of proteins significantly upregulated and down regulated per cell line was subject to enrichment analysis by metascape (https://metascape.org).

Upregulated   Downregulated
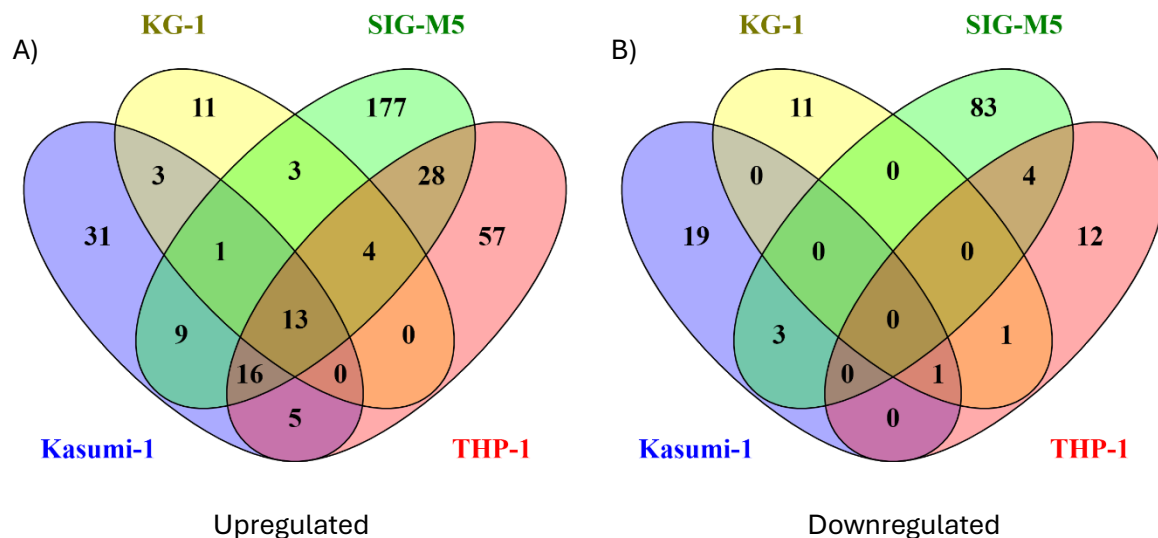
Figure 22: Venn diagram depicting unique and common proteins significantly differentially expressed in response to IFNG across all four cell lines. A) Shows upregulated proteins, while B) shows downregulated proteins.

3.4.3.3   Differentially expressed proteomics profiles were input to Metascape for pathway enrichment analysis

Significantly differentially expressed proteins were converted to gene lists and input to enrichment analysis by metascape to identify which biological pathways were most changed due to IFNG treatment. The Metascape tool draws on multiple databases to perform its analysis including: KEGG Pathway, GO Biological Processes, Reactome Gene Sets, Canonical Pathways,

CORUM, TRRUST, DisGeNET, PaGenBase, Transcription Factor Targets, WikiPathways, and PANTHER Pathway. Metascape uses the whole genome as background for enrichment (Zhou, et al. 2019). To be considered significant, a process had to be reported from the input gene list a minimum of 3 times, with a ratio of more than 1.5 times observed counts compared to counts by chance. The processes that pass were grouped into clusters by similarities, and the most statistically significant member was reported to represent the cluster. As too few significantly differentially expressed proteins were unique in some cell lines to produce a meaningful analysis, for example only 11 proteins were uniquely upregulated in KG-1, full lists of all proteins significantly upregulated or downregulated regardless of commonalties were submitted to metascape. The most significant processes in relation to input up or downregulated lists are reported in Tables below:

Table 8: The topmost significantly enriched pathways in KG-1 treated with IFNG according to up and down regulated protein lists, sorted by significance. Proteins were converted into gene IDs and so are represented by 'gene' lists. Column 'gene' specifies the number of 'genes' differentially expressed in the data, from the corresponding pathway.

Upregulated				**Downregulated**			
Summary process ID	**Description**	**Log10(P)**	**Gene**	**Summary process ID**	**Description**	**Log10(P)**	**Gene**
GO:0002479	Antigen processing and presentation of exogenous peptide antigen via MHC class I, tap-dependent	-17.78	10/75	GO:0043299	Leukocyte degranulation	-2.77	3/537
R-HSA-1280215	Cytokine signalling in immune system	-17.73	17/715				
GO:0019885	Antigen processing and presentation of endogenous peptide antigen via MHC class I	-13.54	6/17				
R-HSA-196807	Nicotinate metabolism	-7.21	4/31				
WP619	Type II interferon signalling (IFNG)	-4.88	3/37				
GO:0008285	Negative regulation of cell population proliferation	-2.65	5/753				
GO:0046649	Lymphocyte activation	-2.64	5/754				

Table 9: The topmost significantly enriched pathways in Kasumi-1 treated with IFNG according to up and down regulated protein lists, sorted by significance post FDR. Proteins were converted into gene IDs and so are represented by 'gene' lists. Column 'gene' specifies the number of 'genes' differentially expressed in the data, from the corresponding pathway.

Upregulated				Downregulated			
Summary process ID	Description	Log10(P)	Gene	Summary process ID	Description	Log10(P)	Gene
R-HSA-913531	Interferon signalling	-38.17	27/199	R-HSA-72766	Translation	-7.00	6/291
GO:0034341	Response to interferon-gamma	-27.15	21/197	CORUM:324	39s ribosomal subunit, mitochondrial	-5.10	3/48
GO:0031347	Regulation of defence response	-19.37	24/697	GO:0051640	Organelle localization	-3.80	5/655
GO:0019883	Antigen processing and presentation of endogenous antigen	-16.75	9/26	WP4223	RAS signalling	-3.35	3/185
GO:0001817	Regulation of cytokine production	-13.59	20/782	GO:0007005	Mitochondrion organization	-2.01	3/548
WP619	Type ii interferon signalling (IFNG)	-13.06	8/37				
R-HSA-1169410	Antiviral mechanism by IFN-stimulated genes	-11.93	9/80				
GO:0002683	Negative regulation of immune system process	-11.27	14/403				
WP5039	Sars-cov-2 innate immunity evasion and cell-specific immune response	-9.09	7/68				
GO:0002366	Leukocyte activation involved in immune response	-8.96	15/720				
R-HSA-8983711	OAS antiviral response	-8.17	4/9				
GO:0032612	Interleukin-1 production	-7.19	7/127				
GO:0009617	Response to bacterium	-6.13	12/728				
GO:0035456	Response to interferon-beta	-5.79	4/31				
GO:1903706	Regulation of hemopoiesis	-5.63	9/417				
GO:0051259	Protein complex oligomerization	-5.46	7/229				
GO:0002718	Regulation of cytokine production involved in immune response	-5.43	5/84				
GO:0008285	Negative regulation of cell population proliferation	-5.15	11/753				
GO:0060759	Regulation of response to cytokine stimulus	-4.84	6/189				
GO:0035455	Response to interferon-alpha	-4.58	3/21				

Table 10: The topmost significantly enriched pathways in SIG-M5 treated with IFNG according to up and down regulated protein lists, sorted by significance. Proteins were converted into gene IDs and so are represented by 'gene' lists. Column 'gene' specifies the number of 'genes' differentially expressed in the data, from the corresponding pathway.

Upregulated				Downregulated			
Summary process ID	Description	Log10(P)	Gene	Summary process ID	Description	Log10(P)	Gene
GO:0045055	Regulated exocytosis	-39.03	61/780	R-HSA-69190	DNA strand elongation	-17.34	10/32
R-HSA-1280215	Cytokine signalling in immune system	-25.46	46/715	R-HSA-69183	Processive synthesis on the lagging strand	-9.03	5/15
R-HSA-1280218	Adaptive immune system	-19.74	41/763	R-HSA-15869	Metabolism of nucleotides	-8.90	8/100
hsa04145	Phagosome	-18.76	22/168	R-HSA-8953854	Metabolism of RNA	-7.47	14/673
R-HSA-913531	Interferon signalling	-17.15	22/199	M66	Pid myc active pathway	-6.65	6/79
GO:0050778	Positive regulation of immune response	-16.39	37/761	GO:0034655	Nucleobase-containing compound catabolic process	-6.61	12/557
WP3888	Vegfa-vegfr2 signalling pathway	-14.45	27/439	GO:0043484	Regulation of RNA splicing	-6.33	7/145
hsa04142	Lysosome	-12.48	15/123	CORUM:1150	Histone h3.3 complex	-5.95	3/7
R-HSA-2262752	Cellular responses to stress	-12.29	30/676	GO:0071897	DNA biosynthetic process	-5.43	7/198
WP619	Type ii interferon signalling (IFNG)	-12.14	10/37	M46	Pid atr pathway	-5.11	4/39
GO:0042110	T cell activation	-11.90	25/475	M195	Pid cmyb pathway	-5.10	5/84
GO:0001817	Regulation of cytokine production	-11.42	31/782	GO:0045930	Negative regulation of mitotic cell cycle	-5.01	8/321
R-HSA-9716542	Signalling by rho GTPases, Miro GTPases and rhobtb3	-10.88	29/719	GO:0006412	Translation	-4.70	11/717
GO:0051345	Positive regulation of hydrolase activity	-10.75	30/779	GO:0002366	Leukocyte activation involved in immune response	-4.68	11/720
GO:0002221	Pattern recognition receptor signalling pathway	-10.06	16/212	GO:0009991	Response to extracellular stimulus	-4.61	9/477
GO:0042060	Wound healing	-9.95	24/538	WP2525	Trans-sulphuration and one carbon metabolism	-3.82	3/32
GO:0030029	Actin filament-based process	-9.86	29/794	GO:0140053	Mitochondrial gene expression	-3.70	5/165
GO:0006954	Inflammatory response	-9.39	28/778	GO:0033120	Positive regulation of RNA splicing	-3.63	3/37
GO:0060627	Regulation of vesicle-mediated transport	-8.76	22/518	R-HSA-9013407	Rho GTPase cycle	-3.60	3/38
GO:0043068	Positive regulation of programmed cell death	-8.43	23/590	M14	PID aurora b pathway	-3.56	3/39

Table 11: The topmost significantly enriched pathways in THP-1 treated with IFNG according to up and down regulated protein lists, sorted by significance. Proteins were converted into gene IDs and so are represented by 'gene' lists. Column 'gene' specifies the number of 'genes' differentially expressed in the data, from the corresponding pathway.

Upregulated				**Downregulated**			
Summary process ID	**Description**	**Log10(P)**	**Gene**	**Summary process ID**	**Description**	**Log10(P)**	**Gene**
R-HSA-913531	Interferon signalling	-28.88	25/199	R-HSA-6798695	Neutrophil degranulation	-6.43	6/480
GO:0048002	Antigen processing and presentation of peptide antigen	-25.95	23/194	hsa03050	Proteasome	-5.51	3/45
GO:0045088	Regulation of innate immune response	-17.11	20/315	GO:0090305	Nucleic acid phosphodiester bond hydrolysis	-3.04	3/306
WP619	Type ii interferon signalling (IFNG)	-13.35	9/37	GO:0016049	Cell growth	-2.50	3/473
GO:0002366	Leukocyte activation involved in immune response	-13.19	23/720				
GO:0050792	Regulation of viral process	-13.07	14/186				
GO:0001817	Regulation of cytokine production	-12.45	23/782				
R-HSA-1236977	Endosomal/vacuolar pathway	-11.58	6/11				
hsa04621	Nod-like receptor signalling pathway	-9.66	11/170				
GO:0051701	Biological process involved in interaction with host	-8.47	11/220				
GO:0038061	Nik/NF-Kappa beta signalling	-8.00	10/188				
GO:0030099	Myeloid cell differentiation	-7.41	13/421				
WP3937	Microglia pathogen phagocytosis pathway	-6.09	5/40				
GO:1990668	Vesicle fusion with endoplasmic reticulum-Golgi intermediate compartment (ergic) membrane	-5.80	3/6				
GO:0052548	Regulation of endopeptidase activity	-5.56	11/429				
WP4197	The human immune response to tuberculosis	-5.56	4/23				
GO:0009617	Response to bacterium	-5.47	14/728				
GO:0002221	Pattern recognition receptor signalling pathway	-5.38	8/212				
R-HSA-168255	Influenza infection	-5.26	7/156				
GO:0045639	Positive regulation of myeloid cell differentiation	-5.23	6/103				

3.4.4 Metascape reports enrichment in IFNG related processes

Metascape reported enrichment in protein expression for processes associated with IFNG response pathways, such as antigen processing, immune responses to viruses and bacteria, lymphocyte activation, inflammatory response, and negative regulation of cell proliferation in all cell lines. SIG-M5 reported upregulation of proteins in the positive regulation of cell death pathways and exocytosis. No other cell lines showed enrichment for proteins involved in cell death, however, expression of proteins part of cell growth and proliferation pathways such as translation, organelle localisation, DNA synthesis and regulation of mitosis were downregulated across cell lines. This could imply an anti-proliferative effect of IFNG; however, this was not tested. Metascape analysis showed enrichment for IFNG regulated processes but did not report enrichment for any pathways involved in response to chemotherapy. IFN-related DNA damage resistance signature (IRDS) genes, which promote resistance to DNA damage base therapies including chemotherapy, have been associated with poorer outcome across many cancer types (Padariya, et al. 2021). IRDS at present have not been investigated for association with poor outcome to front line chemotherapy in AML. Expression of protein products of the IRDS genes were investigated in all four cell lines proteomics data.

3.4.4.1 Proteomic expression of the IRDS signature

AML cell lines exhibiting high and low IFNG signalling could be used to generate IFNG signalling prognostic indexes, to investigate if overexpression of the IFNG pathway results in poorer response to induction chemotherapy and HSCT and thus overall outcome. The purpose was to investigate whether AML cell lines exhibit high expression of the IRDS signature, as seen in other cancers where high IRDS expression has been associated with poorer outcomes. Therefore, the protein products of the IRDS signature were compared across cell lines, using the proteomics data as detected by SWATH-MS (Figure 23). IFI44 was not detected by SWATH-MS, and so not included in further commentary.

Untreated control AML cell lines expressed comparable levels of IRDS protein products. However, treatment with IFNG revealed varying IRDS protein profiles across the cell lines. Surprisingly, although the SIG-M5 cell line exhibited the largest change in the number of deregulated proteins in response to IFNG treatment, it did not upregulate IRDS protein products as highly as the THP-1 or Kasumi-1 cell lines. Notably, the Kasumi-1 cell line was the only one to significantly upregulate all IRDS proteins upon IFNG treatment, reaching much higher expression levels than the other cell lines (MX1 Ctrl =97,079 NPA, IFNG =7,075,961 NPA, P<0.0001, OAS1 Ctrl = 32,495 NPA, IFNG =185,232 NPA, P< 0.0001, IRF7 Ctrl = 22,205 NPA, IFNG = 126,579 NPA, P<0.0001, ISG15 Ctrl = 58,738 NPA, IFNG = 2,952,021 NPA, P<0.0001, IFIT1 Ctrl = 2,952,021 NPA,

IFNG = 1,668,101 NPA, P<0.0001, IFITM1 Ctrl = 12,889 NPA, IFNG = 69,586 NPA, P<0.0001, STAT1 Ctrl = 125,437, IFNG = 1691790 NPA, P<0.001) (Figure 23). This data indicates that the IFNG-treated Kasumi-1 cell line uniquely and disproportionately upregulated these IRDS proteins compared to the other cell lines. The THP-1 cell line significantly upregulated four out of the seven IRDS targets (MX1 Ctrl = 73,612 NPA, IFNG = 147,022 NPA, P< 0.001, IRF7 Ctrl = 12,451 NPA, IFNG = 18,277 NPA, P< 0.05, ISG15 Ctrl = 92,075 NPA, IFNG = 354,650 NPA, P<0.0001, STAT1 Ctrl = 169,530 NPA, IFNG = 1,302,040 NPA, P<0.0001) (Figure 23). Meanwhile, SIG-M5 and KG-1 only significantly upregulated STAT1 (SIG-M5 Ctrl = 192,568 NPA, IFNG = 2,462,092 NPA, P< 0.0001 , KG-1 Ctrl = 782,133 NPA, KG-1 IFNG = 2,033,763 , P<0.05) and ISG15 (SIG-M5 Ctrl = 42,607 NPA, IFNG = 98,182 NPA, P<0.001, KG-1 Ctrl = 55,085 NPA, IFNG = 186,602 NPA, P<0.001) (Figure 23).

STAT1 acts as a transducer for the IFNG signalling pathway. Therefore, higher expression of STAT1 could relate to a larger or more intense response to IFNG. However, all cell lines upregulated STAT1 to similar levels, indicating that the disparity in IFNG response was not due to baseline or inducible STAT1 levels. Another explanation could be that cell lines more responsive to IFNG expressed higher levels of the IFNG receptor, thereby amplifying signalling. To test this, the untreated control cell lines were stained for CD119 (IFNG receptor 1) cell surface expression using flow cytometry (Figure 23B). Although SIG-M5 and THP-1 demonstrated the largest breadth of response to IFNG, staining showed all cell lines to have similar receptor expression (SIG-M5 = 2.92 MFI, THP-1 = 2.72 MFI, KAS-1 = 1.63 MFI and KG-1 = 3.86 MFI). Surprisingly, even though KG-1 had the smallest breadth of response to IFNG, as quantified by the number of differentially expressed proteins and affected IRDS protein products, it expressed CD119 higher than all other cell lines. CD119 expression was lowest in Kasumi-1. Therefore, differences in response to IFNG could not be attributed to receptor expression or to baseline or IFNG-induced levels of STAT1. Differences in IFNG signalling among cell lines were explored more thoroughly in chapter 4. IFNG induced expression of IRDS has been linked to chemotherapy resistance in other cancer types. IFNG-induced cells were tested for viability in the presence of chemotherapy to determine whether the IFNG signalling pathways induced cell death or conferred resistance.
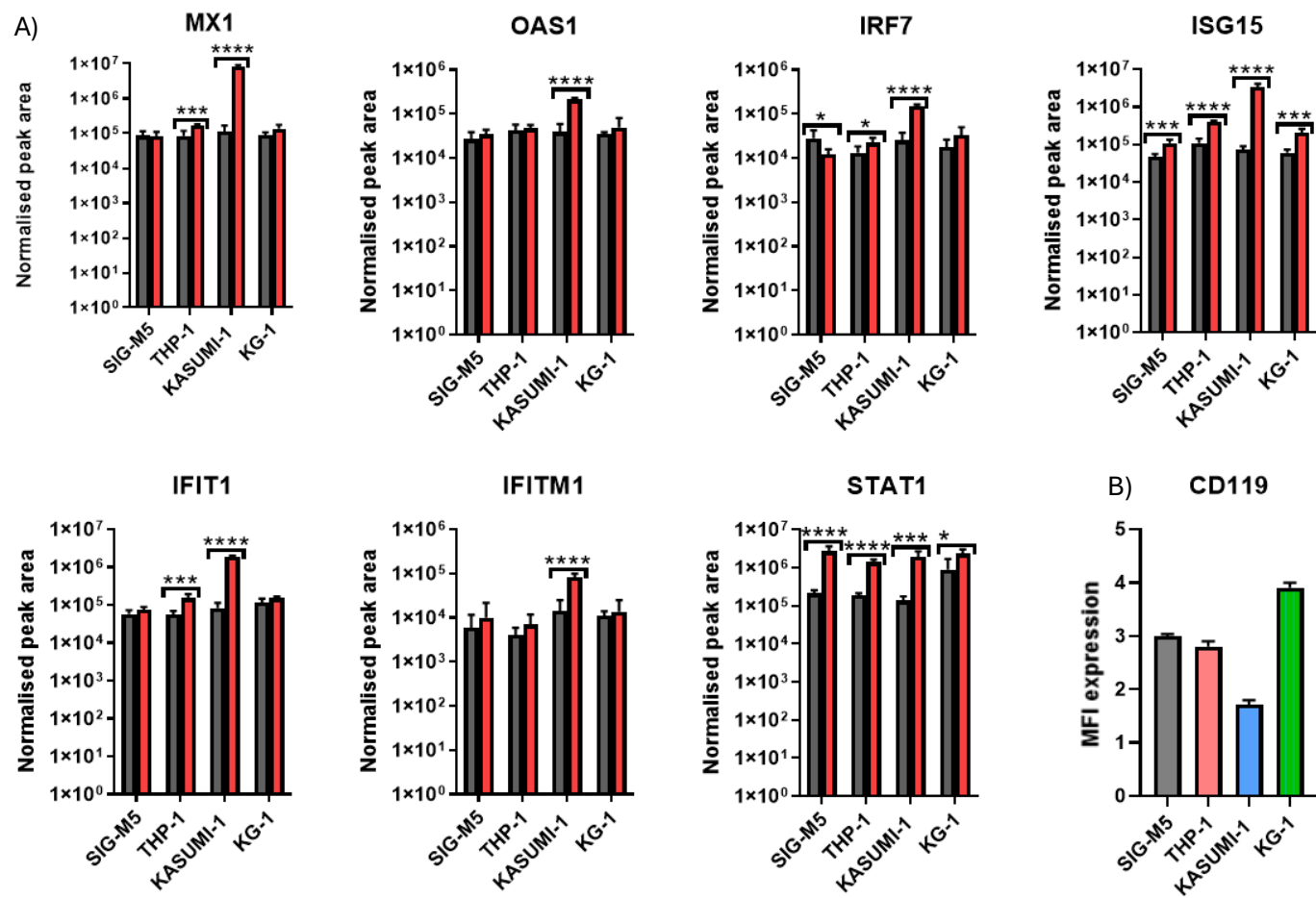
Figure 23:Normalised peak area expression of protein products of IRDS genes: MX1, OAS1, IRF7, ISG15, IFIT1, IFITM1 and STAT1 across all four AML cell lines. Control compared to 48 hrs treatment 100 ng/ml IFNG (n=5-6). Grey = Control, Red = IFNG. B) Base line expression of CD119 (IFNG receptor 1) in all cell lines measured by cell surface staining flow cytometry, n=3. Statistical test: Holm-Sidak multiple comparisons method, * = P< 0.05, ** = P< 0.01, *** = P<0.001 and **** = P<0.0001.

3.4.5 IFNG treatment drives cell death in Kasumi-1

The Kasumi-1 and KG-1 cell lines exhibited different responses to IFNG, as evidenced by their varying expression levels of IRDS proteins. This indicates they may have different biological responses to IFNG that could be explored further. The viability of cells when treated with IFNG was compared to those without.

3.4.6 IFNG induced a cell death response to chemotherapy in the Kasumi-1 and KG-1 cell lines

Cells were assessed using flow cytometry by staining for Annexin V and an amine reactive dye (LIVE/DEAD™) to determine if IFNG or 5AzaC exacerbated Daunorubicin-activated cell death. Cells were pre-treated with IFNG, 5AzaC, or both before Daunorubicin administration, allowing time for the treatments to alter signalling pathways and gene expression. The hypothesis was that demethylation by 5AzaC could modify IFNG signalling and affect the cells' viability in response to chemotherapy. To test this, cells received 5AzaC three times at 24-hour intervals to ensure DNMT inhibition. This allowed multiple rounds of DNA replication to occur without active DNMT, resulting in demethylated DNA. After three days of treatment, cells were exposed to IFNG, Daunorubicin, or both to evaluate the impact of DNA demethylation on cell viability. As this was a simple *in vitro* study with no co-culturing of immune cells, the expected result was for IFNG to induce apoptosis of cell lines, and further activate Daunorubicin induced cell death.

Using the gating strategy described in chapter 2, an experiment was conducted to determine an appropriate Daunorubicin concentration for a time course experiment. The IFNG and 5AzaC dosing and schedule is described in Table 12. An initial experiment was run with 0, 0.2, 0.4 and 0.8 μM Daunorubicin after 48 hrs of treatment to choose the dose that would reduce live cell population by approximately 50%.

Table 12: Outlines the dosing and measurement schedule used to assess cell viability of cell lines treated with IFNG, 5AzaC and Daunorubicin.

Day	Treatment	Cell viability measurement taken after Daunorubicin dosing
0	Cells seeded, 5 ng/mL IFNG added, 0.5 µM 5AzaC added	N/A
1	0.5 µM 5, AzaC added, Daunorubicin added	N/A
2	0.5 µM 5AzaC added	24 hrs
3	None	48 hrs
4	None	72 hrs

The results (Figure 24) demonstrated the impact of increasing concentrations of Daunorubicin on the number of cells in live, early apoptosis, necrotic, and dead states. At Daunorubicin concentrations of 0.2 and 0.4 µM, a higher percentage of dead cells was observed on average in the Kasumi-1 cell line compared to the KG-1 cell line (Kasumi-1: 34.29% and 60.29%; KG-1: 26.98% and 42.78%). Additionally, more cells in early apoptosis were seen in the Kasumi-1 cell line than in the KG-1 cell line across all Daunorubicin doses. Both cell lines exhibited similar percentages of necrotic cells at all chemotherapy doses used. Based on this data, a concentration of 0.4 µM Daunorubicin was chosen, as it reduced the live cell populations in both cell lines to approximately 50% of their untreated counterparts (Kasumi-1 live: untreated: 75.26%, 0.4 µM: 43.12%; KG-1 live: untreated 86.26%, 0.4 µM: 47.54%).

3.4.6.1 Flow cytometry time course experiment reveals different response phenotypes of cell lines

Cell lines were primed with IFNG and 5AzaC following the time course described in Table 12. After priming, the cells were either left untreated or were treated with 0.4 µM Daunorubicin. The cells were harvested 24, 48, and 72 hours post-Daunorubicin treatment, and cell viability was measured by staining with Annexin V and LIVE/DEAD stain, allowing separation of dead cells from cells in apoptosis or necrosis. The results are summarised in Figure 25.

IFNG treatment alone induced cell death Kasumi-1 cells, but did not affect KG-1 cells, this was most prominently seen at the 72-hour time point (KAS-1 IFNG dead population = 40.99%, KG-1 IFNG dead population = 7.24%, Figure 26 and Figure 27). Staining with Annexin V and LIVE/DEAD stain showed that priming Kasumi-1 with IFNG also increased the number of cells in early apoptosis at 72 hours compared to the untreated control (14.61% vs 3.6%, averaged across 3 repeats). Priming Kasumi-1 with both IFNG and 5AzaC induced more cell death than IFNG alone at 72 hours (Kas-1 IFNG5AzaC dead population =50.14%, IFNG dead population = 39.98, P<0.01). Furthermore, combining IFNG with Daunorubicin was significantly more effective at killing

Kasumi-1 cells than Daunorubicin alone at both 48 (Kas-1 IFNG Daunorubicin dead population = 57.02%, Daunorubicin dead population = 37.65%, P<0.05) and 72-hour time points (Kas-1 IFNG Daunorubicin dead population = 80.07%, Daunorubicin dead population = 53.57%, P<0.001, respectively). Priming with both 5AzaC and IFNG followed by Daunorubicin was no more effective at inducing cell death than IFNG combined with Daunorubicin.

In KG-1 cells, treatment with IFNG or 5AzaC alone made no significant difference to the percentage of cells in the 'Live,' 'Early apoptosis,' 'Necrotic,' or 'Dead' categories compared to the untreated cells. Differences in behaviour based on priming were only visible once Daunorubicin was added to the cells. In KG-1. Combining 5AzaC priming with Daunorubicin increased the dead cell population in KG-1 in comparison to cells treated with only Daunorubicin (KG-1 5AzaC Daunorubicin dead population = 62.03%, KG-1 Daunorubicin dead population = 49.99%, P<0.01). Similarly to Kasumi-1, combining IFNG priming with Daunorubicin was the most effective condition for killing KG-1 cells. At 24, (KG-1 IFNG Daunorubicin dead cell population = 26.58%, KG-1 Daunorubicin dead population = 15.19 %, P<0.01) 48, (KG-1 IFNG Daunorubicin dead cell population = 70.89% , KG-1 Daunorubicin dead population = 43.04%, P<0.01) and 72 hours, (KG-1 IFNG Daunorubicin dead cell population = 86.71%, KG-1 Daunorubicin dead population = 50.63, P<0.0001) the combination of IFNG and Daunorubicin induced significantly more cell death compared to Daunorubicin alone. Combining IFNG with 5AzaC priming did not improve the killing efficiency of Daunorubicin in KG-1 cells.

Representative scatter plots for cell viability for both cell lines exposed to all treatment types for 72 hours are shown in Figure 26 and Figure 27.
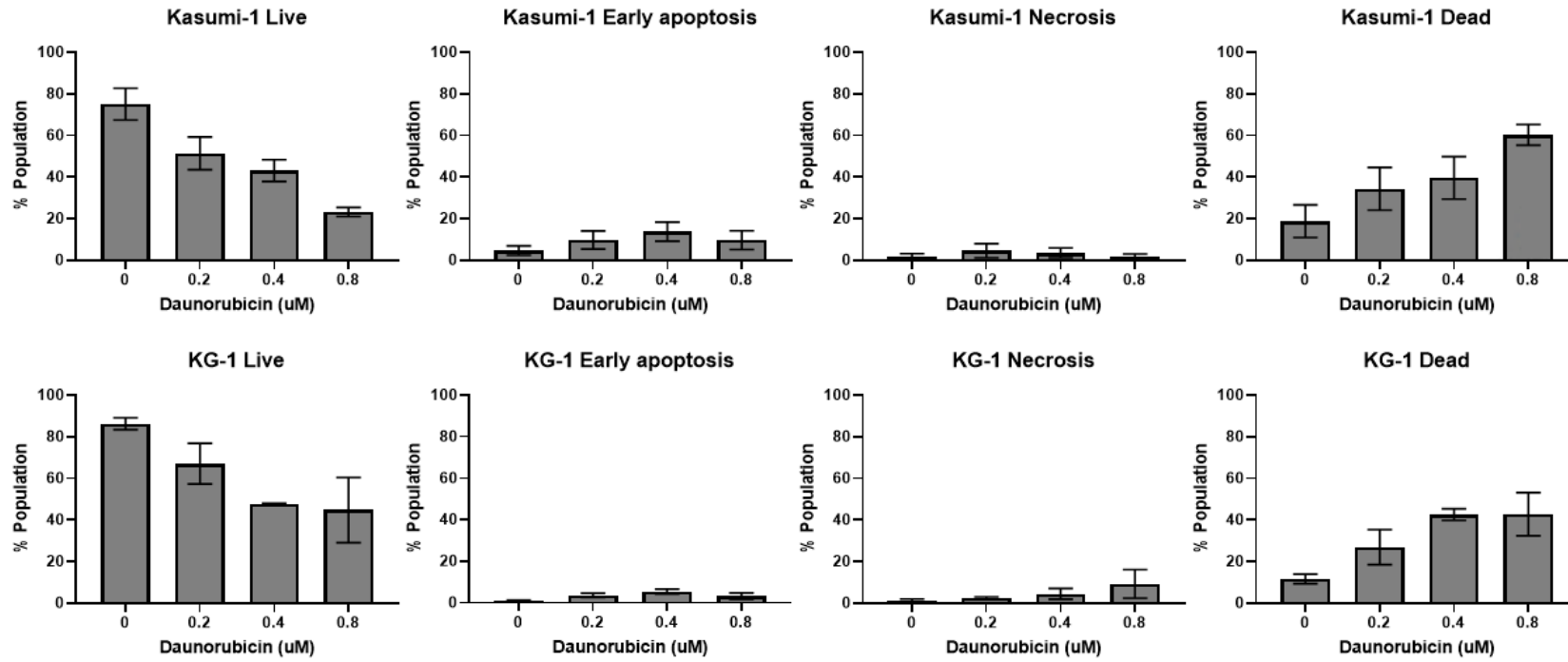
Figure 24: The population of cells 'Live', in 'Early apoptosis', 'Necrosis or 'Dead' in both cell lines after treatment with increasing doses of Daunorubicin for 48 hrs. Cell viability was determined by staining with Annexin V and LIVE/DEAD stain, n=3-5. Error bars = standard deviation.
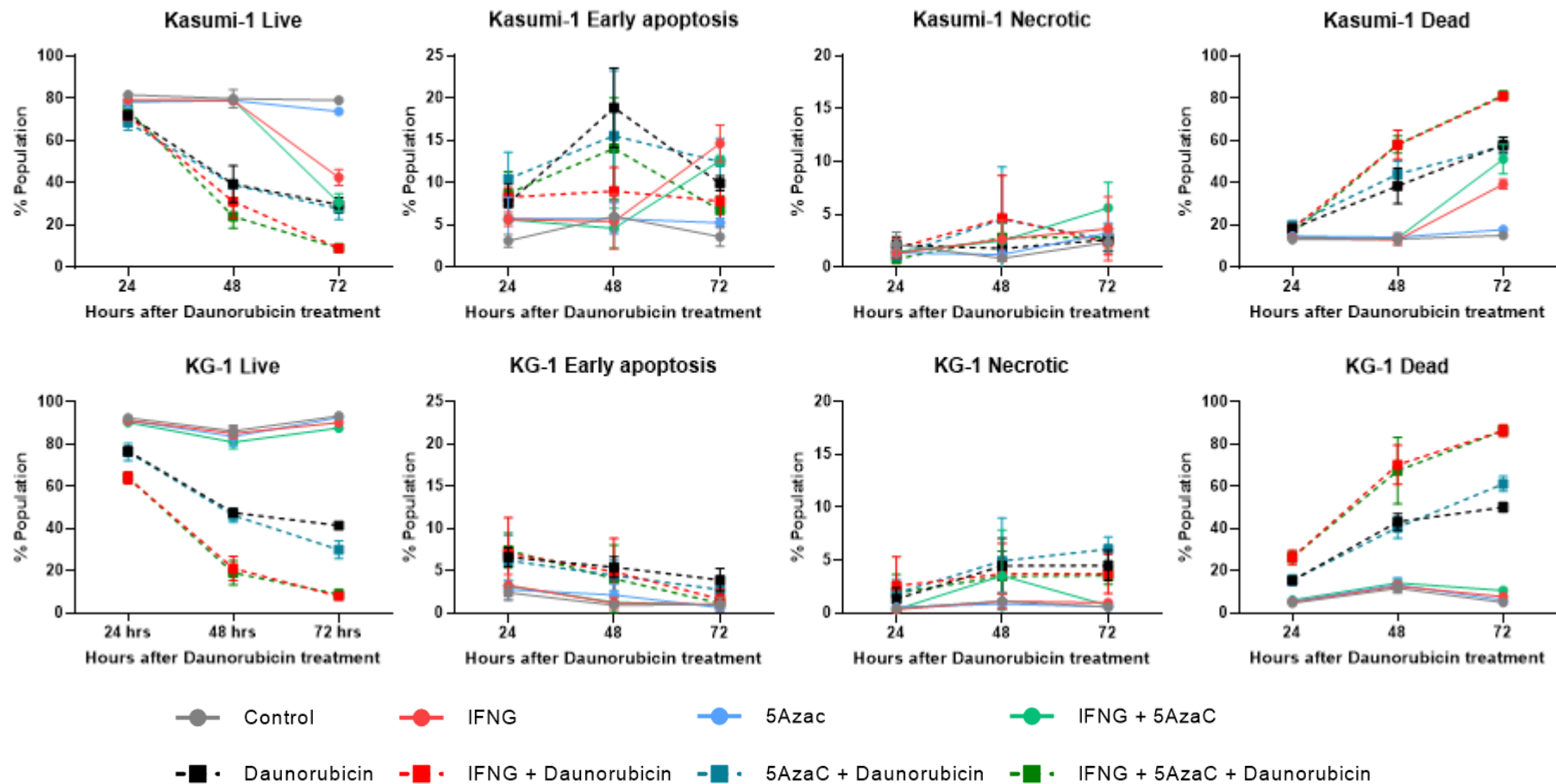
Figure 25: The population of cells 'Live', in 'Early apoptosis', 'Necrosis or 'Dead' in both cell lines after treatment with IFNG /5AzaC/Daunorubicin at 24, 48 and 72 hrs. Cell viability was determined by staining with Annexin V and LIVE/DEAD stain, n=3-6, error bars = standard deviation.
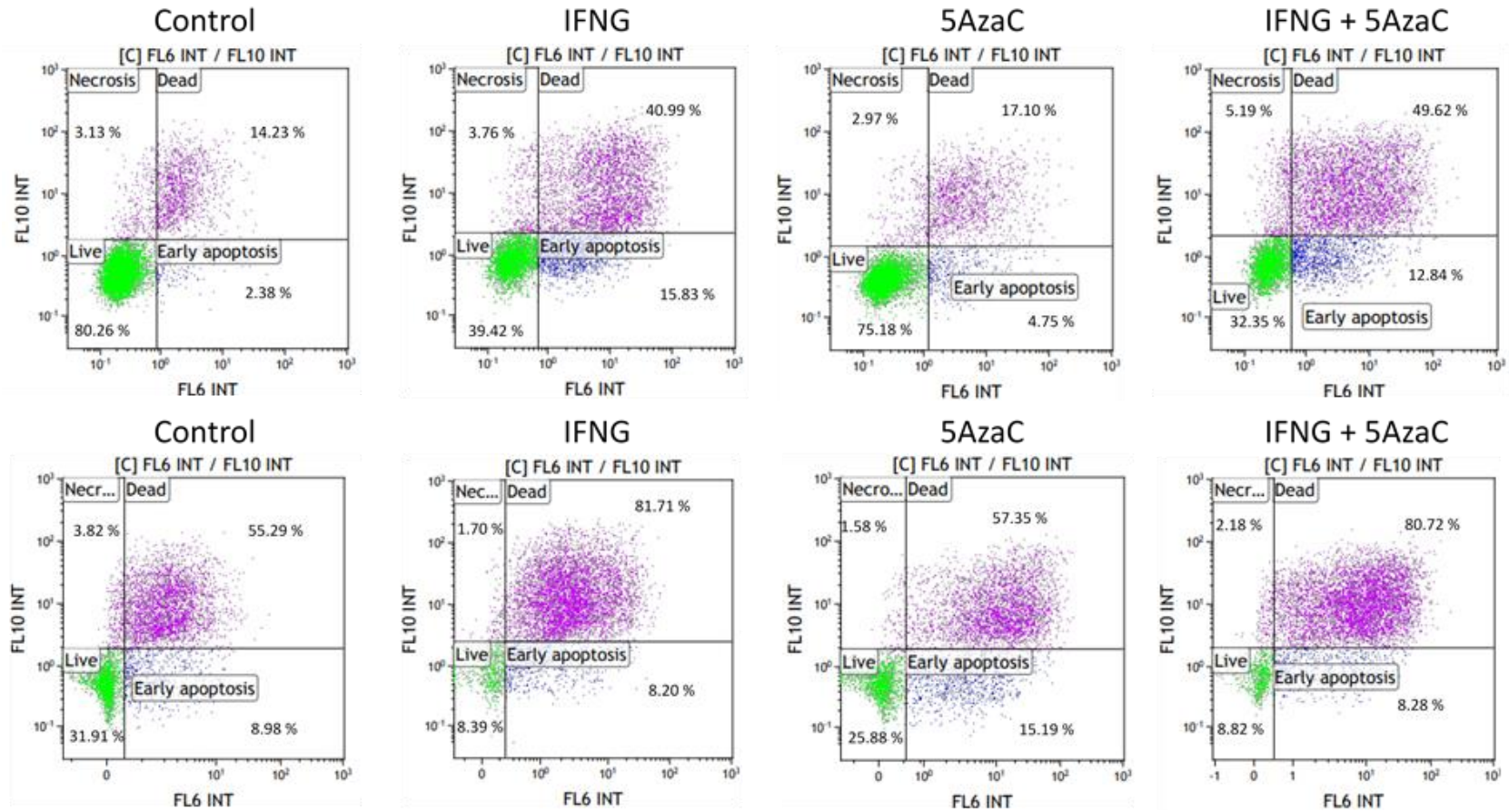
Figure 26: Example scatter plots of Kasumi-1 cells after priming with IFNG and 5AzaC, and treatment Daunorubicin for 72 hrs. Cell viability was determined by staining with Annexin V and LIVE/DEAD stain, Annexin V staining is plotted on the X axis labelled 'FL6 INT', and LIVE/DEAD stain is plotted on the Y-axis labelled 'FL10 INT'. Scatters are from the n2 repeat of data. Top row shows Kasumi-1 cells without Daunorubicin and with IFNG, 5AzaC or both, bottom row are Kasumi-1 cells with 0.4 uM Daunorubicin and IFNG, 5AzaC or both.
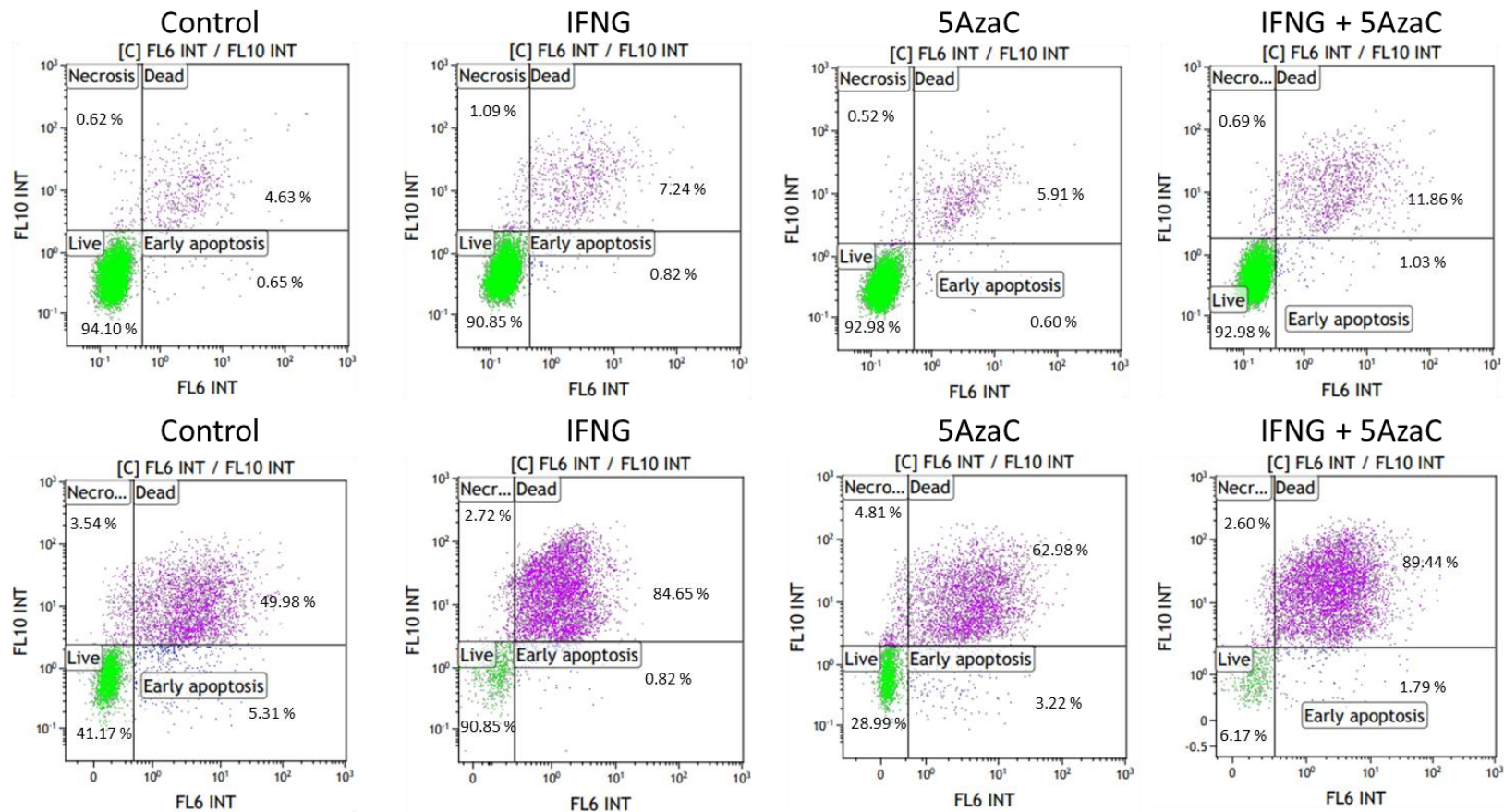
Figure 27: Example scatter plots of KG-1 cells after priming with IFNG and 5AzaC, and treatment Daunorubicin for 72 hrs. Cell viability was determined by staining with Annexin V and LIVE/DEAD stain, Annexin V staining is plotted on the X axis, labelled 'FL6 INT', and LIVE/DEAD stain is plotted on the Y-axis labelled 'FL10 INT'. Scatters are from the n2 repeat of data. Top row shows KG-1 cells without Daunorubicin and with IFNG, 5AzaC or both, bottom row are KG-1 cells with 0.4 uM Daunorubicin and IFNG, 5AzaC or both.

3.5 Discussion

Cell lines were categorised based on their response to treatment with IFNG and chemotherapy. Cell lines showed a spectrum of IFNG response as determined by known IFNG inducible AIR molecules and by breadth of differentially expressed proteins identified by SWATH-MS. First, expression of IFNG inducible molecules (*IDO1, TDO2, CD274,* HLA-A,B,C, HLA-E, HLA-G, PD-L1 and Kynurenine) were measured by RT-PCR, flow cytometry and colourimetric assays as appropriate to establish basic IFNG response and immunosuppressive phenotypes. SIG-M5 was the only cell line to significantly upregulate all AIR molecules with IFNG treatment, while all other cell lines upregulated five out of the eight AIR molecules (Summary in Table 13).

Measurement of kynurenine, a downstream product of IDO1/TDO2 signalling, allowed assessment of whether the immunosuppressive IDO1/TDO2 pathway had increased activity corresponding to IFNG induction of *IDO1/TDO2* genes. Out of the four AML cell lines, only SIG-M5 and KG-1 have a functioning IDO immunosuppressive AIR response to IFNG, supported by significantly upregulated *IDO1*(SIG-M5 Ctrl = 0.0030, IFNG = 0.2176 ΔΔCT, P<0.01, KG-1 Ctrl = 0.0018, IFNG = 0.0080 ΔΔCT), P<0.01, Figure 14) and kynurenine (SIG-M5 Ctrl = 1.9 µM, IFNG = 37.46 µM ,P<0.01, and, KG-1 Ctrl = 0.12 µM, IFNG = 1.07 µM, P<0.01, Figure 15) expression in both cell lines. However, the increase in kynurenine production by SIG-M5 was substantially larger than by KG-1. Although *TDO2* was upregulated in THP-1 (THP-1 Ctrl = 0.0018 ΔΔCT, IFNG = 0.0837 ΔΔCT, P<0.0001, Figure 14), kynurenine production did not increase with IFNG treatment, suggesting impairment in the pathway. This result was in line with Hoffmann et al, who found a combination of 100ng/mL IFNG and 1ug/mL LPS (lipopolysaccharides) induced a significant increase in TDO2 expression, but a reduced kynurenine expression compared to THP-1 treated with LPS alone. They hypothesised that the combined treatment of IFNG with LPS also highly induced IDO1, which resulted in complete removal of tryptophan, which subsequently resulted in signalling for degradation of IDO1 and TDO2 proteins. Hence, when samples were measured after 2 days for kynurenine, the negative feedback had already occurred, causing reduced kynurenine to be recorded (Hoffmann, et al. 2019). It is possible a similar negative feedback loop occurred here.

Flow cytometry was used to measure expression of the surface proteins HLA-A,B,C, HLA-E, HLA-G and PD-L1 . In all cell lines HLA-A,B,C was significantly upregulated by IFNG which was expected (SIG-M5 Ctrl = 158.11 MFI, IFNG = 270.10 MFI, P<0.001, THP-1 Ctrl = 61.03 MFI, IFNG = 148.98 MFI, P<0.0001, Kasumi-1 Ctrl =11.20 MFI, IFNG = 120.97 ,P<0.0001 and KG-1 Ctrl = 35.72 MFI, IFNG = 120.97, P<0.001, Figure 16). HLA-E was expression was induced in all cell lines

except for KG-1 (SIG-M5 Ctrl = 0.09 MFI, IFNG = 3.12 MFI, P<0.0001, THP-1 Ctrl = 0.46 MFI, IFNG = 1.40 MFI , P<0.01, Kasumi-1 Ctrl = 1.62 MFI, IFNG = 5.17 MFI, P<0.01, Figure 16), and HLA-G was only IFNG inducible in SIG-M5 line (SIG-M5 Ctrl = 1.24 MFI, IFNG = 26.53 MFI, P<0.0001, Figure 16). Finally, IFNG induced increased expression of PD-L1, the protein product of *IDO1/TDO2* in all cell lines (SIG-M5 Ctrl = 0.10 MFI, IFNG = 6.96 MFI, P<0.01, THP-1 Ctrl = 0.51 MFI, IFNG = 19.70 MFI, P<0.001, Kasumi-1 Ctrl = 0.62 MFI, IFNG = 1.35 MFI, P<0.0001 and KG-1 Ctrl = 0.77 MFI, IFNG = 1.04 MFI, P<0.05, Figure 16). This data showed active expression of various immunosuppressive proteins in all four cell lines.

Cell lines were then treated with IFNG and analysed using SWATH-MS to further characterise their IFNG responses. All cell lines reported close to 3,000 proteins differentially expressed. When confidence and fold-change cut offs were applied, this decreased numbers of proteins of interest. SIG-M5 reported the most differentially expressed proteins in response to IFNG (341), while KG-1 reported the least (48). Table 13 summarises chapter 3 AIR and IFNG response characterisation results below:

Table 13: Summary of characterisation data from chapter 3. Ticks represent significant upregulation of RNA or protein by IFNG. Also lists the number of proteins including IRDS proteins significantly differentially expressed or upregulated with IFNG treatment.

	IDO1	TDO2	CD274	HLA-A,B,C	HLA-E	HLA-G	PD-L1	Kynurenine	Proteins significantly dysregulated	IRDS proteins upregulated
SIG-M5	✓	✓	✓	✓	✓	✓	✓	✓	341	3/7
THP-1	✗	✓	✓	✓	✓	✗	✓	✗	141	5/7
Kasumi-1	✗	✗	✓	✓	✓	✗	✓	✓	101	7/7
KG-1	✓	✗	✓	✓	✗	✗	✓	✓	48	2/7

Lists of 'significantly' upregulated and downregulated proteins were uploaded to metascape for process enrichment. Enrichment analysis did not show upregulation of proteins involved in immunosuppressive pathways, contradicting the hypothesis that IFNG induces immunosuppressive phenotypes in AML cell lines. However, SWATH-MS is 3-10 times less sensitive than SRM, as discussed in Chapter 1, so low-abundance proteins might have been missed. Additionally, this was an in vitro test where AML cell lines were treated with IFNG in isolation from an immune microenvironment, which might have triggered immunosuppressive responses if present. Moreover, this was a short-term assay, so the impact of chronic IFNG exposure within a more complex immune microenvironment could not be determined.

Overall, the characterisation showed cell lines had developed varying degrees of IFNG induced AIR, with SIG-M5 upregulating the most AIR molecules, while all other cell lines only upregulated five out of the eight tested. Additionally, the SIG-M5 cell line reported the largest number of differentially expressed proteins while KG-1 expressed the least. Suggesting there are differences between the cell lines IFNG response mechanisms. This was corroborated by the pairwise comparisons of Pearson correlation coefficients preformed on SWATH-MS proteomics profiles of untreated control and IFNG treated AML cell lines. It was found that KG-1 untreated control and IFNG samples overlapped during clustering (Figure 19), suggesting either noneffective IFNG treatment, or a low and muted response to IFNG resulting in a similar expression profile to KG-1 control. This overlap of IFNG treated and untreated KG-1 was also seen in the PCA analysis (Figure 20). In contrast, the SIG-M5, THP-1 and Kasumi-1 cell lines all clustered into defined untreated control and IFNG treated groups. This suggested that IFNG was not activating a robust signalling response in KG-1 as it was in the other cell lines.

To further characterise IFNG response, expression of IRDS gene protein products were examined as they have been associated with chemotherapy and radiation resistance in other cancer types. The SIG-M5 cell line reported only three of seven IRDS proteins as significantly altered, while the Kasumi-1 cell line reported all. Furthermore, the expression of the IRDS proteins in IFNG treated Kasumi-1 was higher than any other cell line by 10 to 100 times, suggesting an IFNG response unique to this cell line, which might have been missed if only gauging IFNG response on AIR induction or number of significantly deregulated proteins. For this reason, Kasumi-1 was carried forward for analysis, along with KG-1 which showed a comparatively muted response to IFNG, expressing the lowest number of differentially expressed proteins (48), and upregulation of only 2 IRDS proteins. Finally, to investigate how IFNG influenced cell lines response to chemotherapy, cell line viability was tested in the presence of Daunorubicin, IFNG and 5AzaC treatments. IFNG alone induced cell death in Kasumi-1 (40.99%) and KG-1 (7.24%) cell lines after 72 hrs (Figure 25). Furthermore, IFNG combined with Daunorubicin induced further cell death after 72 hrs in both cell lines (Kasumi-1=80.07%, KG-1= 86.71%) than treatment with just Daunorubicin (Kasumi-1 = 53.57%, KG-1 = 50.63%) (Figure 25). Combining IFNG with Daunorubicin was more effective at killing both the AML cell lines than Daunorubicin alone. The effect of combining IFNG with Daunorubicin also induced increased cell death in KG-1, even though IFNG itself did not induce cell death in KG-1 as it did in Kasumi-1.

All AML cell lines demonstrated IFNG induced upregulation of immunosuppressive mechanisms, suggesting IFNG rich environments could contribute to poorer outcomes in AML patients. In summary, this chapter investigated if IFNG treatment upregulated expression of adaptive

immune resistance molecules in AML cell lines and generated proteomic profiles of untreated and IFNG treated AML cell lines. Cell lines exhibited different degrees of response to IFNG in terms of intensity and breadth of proteins differentially expressed, as measured by flow cytometry and SWATH-MS. It was found that cell lines upregulated immunosuppressive molecules (IDO1, PDL1, HLA-E/HLA-G), in response to IFNG. These molecules are known to aid immune escape and inhibit immune cells, indicating these cell lines could represent different types of immunosuppressive AML. This study was limited by its *in vitro* nature. Therefore, when cell lines Kasumi-1 and KG-1 were treated with IFNG and Daunorubicin, cell death was observed as opposed to proliferation. The theory of IFNG induced AML resistance to chemotherapy is dependent on an *in vivo* environment with immune cells present for immunosuppressive molecules such as IDO1 to take effect. The next step was to generate transcriptional profiles of cell lines in response to IFNG and 5AzaC treatment to analyse how influenced the AML transcriptome. Furthermore, the profiles created were then used to generate prognostic scores associated with each treatment and applied to patient databases to assess correlation with patient survival and clinical categories such as cytogenetic risk. In chapter 4, AML cell lines Kasumi-1 and KG-1 were treated with IFNG and 5AzaC, and transcriptional profiles were created using RNAseq.

4 A comparative RNA sequencing analysis of demethylation effects on IFNG induced changes to immunosuppressive molecules in Kasumi-1 and KG-1 cell lines

4.1 Introduction

All the possible transcripts that an organism can express are encoded in its genome. Under certain circumstances, these transcripts are transcribed into mRNA, which serves as an intermediary molecule between the transcript and the protein product. Changes in transcript expression can be quantified by measuring RNA transcript levels. Because transcriptomes are dynamic and respond to various stimuli, measuring RNA expression provides a snapshot of the genes being expressed in a cell at any given time. Thus, transcriptomics is a valuable tool for analysing how stimuli affect transcript expression. In this case, it involves quantifying transcripts to assess the effects of treatments such as IFNG or 5AzaC (Lowe, et al. 2017).

In Chapter 3, cell viability assays revealed that IFNG treatment increased cell death in response to Daunorubicin in both cell lines. This Daunorubicin-induced cell death was further amplified when IFNG was combined with 5AzaC, showing a synergistic effect that was more pronounced in Kasumi-1 cells than in KG-1 cells. The differing responses of the two cell lines to IFNG and 5AzaC treatments warrant further molecular-level investigation to uncover the underlying mechanisms. To this end, next-generation RNA sequencing was performed on the cell lines after treatment with IFNG, 5AzaC, or their combination, alongside a non-treated control.
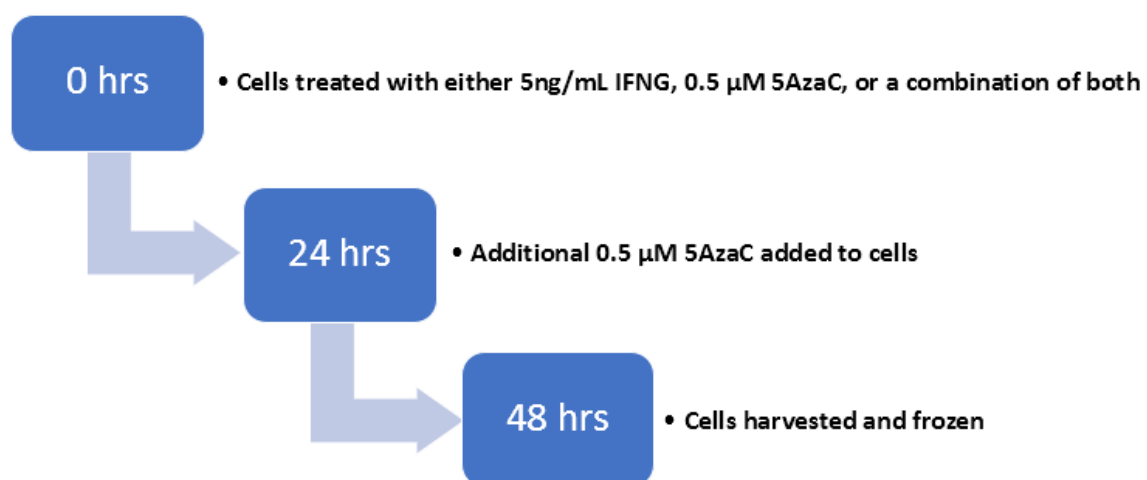
Figure 28:The Workflow and time points for treating cells prior to sending to Novogene for analysis.

4.2 Results

First, the transcriptomes of untreated Kasumi-1 and KG-1 cells were compared to determine their similarity prior to treatment with IFNG or 5AzaC. Transcripts were reported in FPKM (Fragments Per Kilobase of exon per Million mapped fragments), a metric which shares similarity with RPKM (Reads Per Kilobase of exon per Million reads mapped) but specifically used in paired-end RNA-seq investigations. After comparing the untreated cell lines, the impact of treatment with IFNG, 5AzaC, or a combination of both on transcript expression was examined. Only transcripts with an initial expression of more than 1 FPKM and significant differential expression between treatments and cell lines after FDR correction were selected for further investigation. The 1 FPKM cut-off was used because relying solely on fold change can be misleading; it excludes baseline expression levels, and significant fold changes might not reflect biologically meaningful differences if initial expression is extremely low or zero. Statistical analysis and FDR correction were performed by Novogene using the DESeq2 analysis package. Data were transformed using Log2(FPKM+1), and Z score values were calculated to generate heat maps.

In addition to differential transcript expression, enrichment analysis was conducted to identify the biological pathways most significantly associated with differentially expressed transcripts (DETs). This analysis was also provided by Novogene. As in Chapter 3, lists of significant DETs between selected treatment groups were applied to pathway analysis using Metascape (Zhou, et al. 2019). NGS seq identified far more significantly altered transcripts, than SWATH MS detected at the protein level.

4.2.1 Kasumi-1 and KG-1 transcriptomes indicate different AML phenotypes

NGS revealed 47% of transcripts were significantly (Padj < 0.05) differentially expressed between the cell lines (Figure 29). Results showed that 14,481 transcripts were differentially expressed between the cell lines, with 6,979 and 7,502 upregulated and downregulated respectively in Kasumi-1 compared to KG-1. The degree of differences in transcript expression between cell lines lays the foundation for cell lines to function in biologically unique ways to one another and showcases the heterogenous nature of AML.
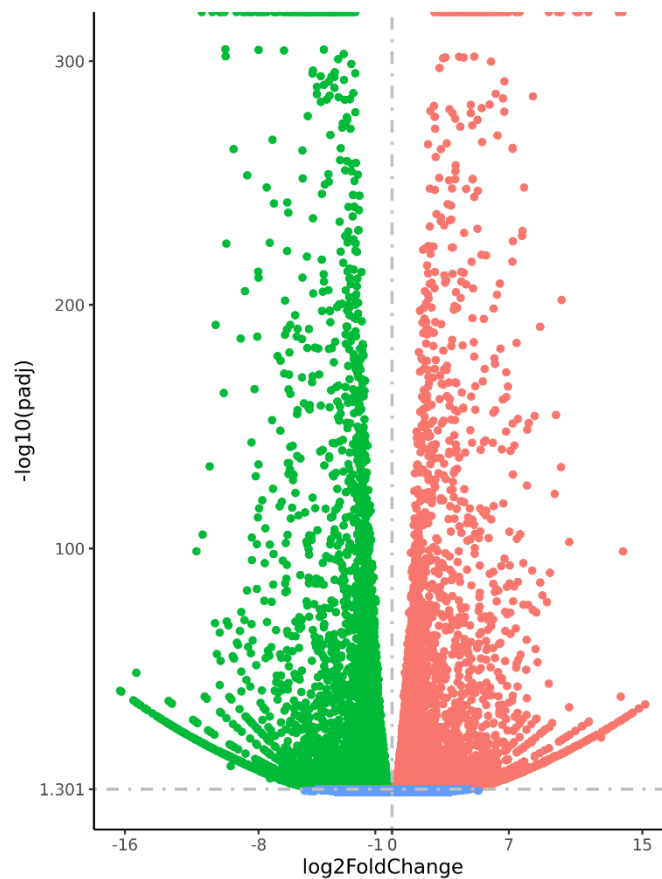
Figure 29: Volcano plot of differentially expressed transcripts (DET) between cell lines Kasumi-1 and KG-1. Upregulated and down regulated transcripts are reported as red and green dots, respectively. Unchanged transcripts are represented by blue dots. Padj threshold < 0.05.

Differences between cell lines were further evaluated by identifying the number of transcripts expressed exclusively in each cell line. An FPKM > 1 cut-off was used for downstream analysis, following the transcriptomic analysis method published by Mortazavi in Nature Methods (Mortazavi, et al. 2008). Lists of expressed transcripts for both cell lines were filtered accordingly and compared by Venn diagram (Figure 30). This comparison revealed that approximately 10% of the transcripts were uniquely expressed in each cell line.
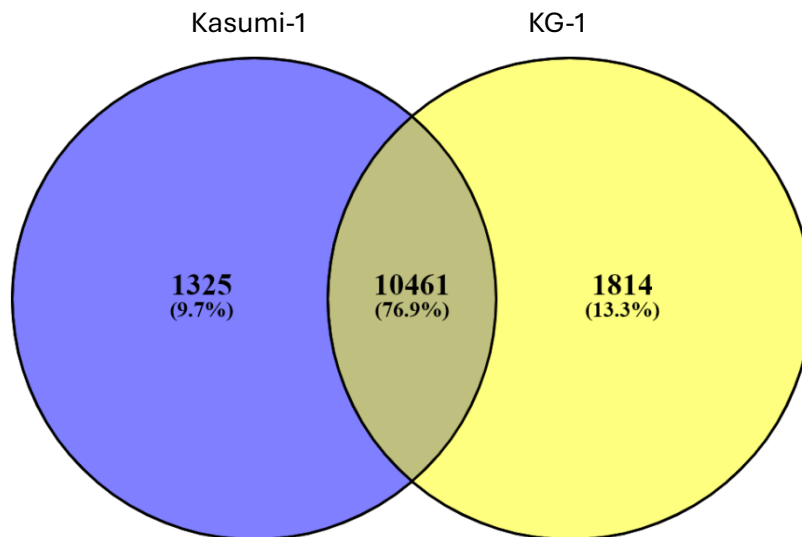
Figure 30: Venn diagram showing number of transcripts expressed specifically the Kasumi-1 or KG-1 cell line or shared between both.

The transcriptomic analysis revealed variations in the directionality of expression for shared transcripts between the cell lines, as evidenced by the volcano plot, as well as in the expression of transcripts exclusive to each individual cell line. The cell lines were derived from patients exhibiting specific morphologic and genetic features, resulting in some of the divergence in transcript expression observed. Consequently, these cell lines, as discussed in chapter 3, can serve as models to represent different and specific populations of AML patients, with discoveries in either model being clinically relevant.

4.2.1.1 *Pearson correlation and principle component analysis of treated cell lines indicate no outliers among samples and show IFNG and 5AzaC induced changes in transcription*

The transcriptomes of samples were compared using Pearson correlation to assess variability between treatments and cell lines, as well as to identify any replicate outliers. The Pearson correlation coefficient was calculated between pairs of samples and visualised as a heat map (Figure 31). Clustering clearly demonstrated that treatment with IFNG and 5AzaC resulted in successful replicates with no obvious outliers. This perfunctory form of analysis also reported global differences in treatment impact on transcriptomes between cell lines. Treatment of Kasumi-1 with IFNG, or combination treatment, resulted in reduced correlation to its control (0.9 and 0.86 respectively) and 5AzaC treated counterparts (0.9 and 0.87 respectively). In contrast, correlation coefficients of KG-1 differed very little between treatments and control. Coefficients were high (0.97-1) indicating the KG-1 transcriptome was largely unaffected by IFNG or 5azaC treatment. When cell lines were compared to one another, correlation stood at around 0.75, regardless of treatments.
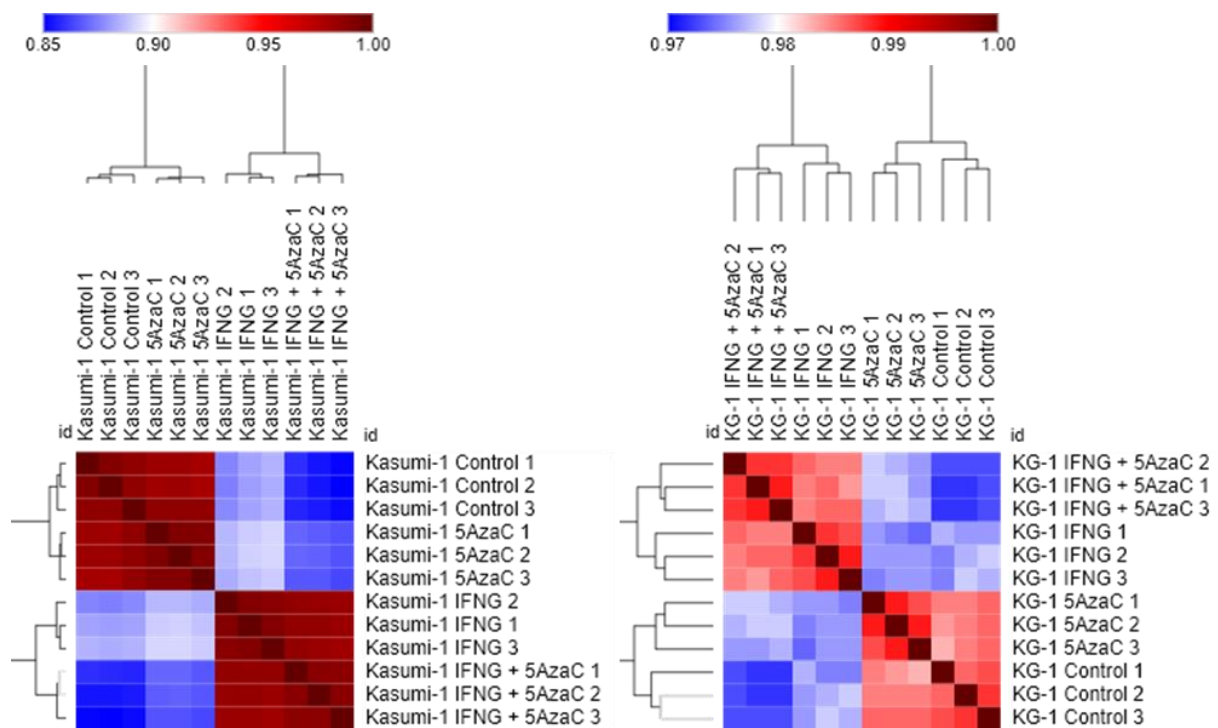
Figure 31: Pairwise comparisons of Pearson correlation coefficients between untreated and IFNG/5AzaC treated cell lines transcriptomic profiles. Left, Kasumi-1, Right, KG-1. Pearson coefficients were calculated by Novogene, and then visualised and clustered using Morpheus online tool https://software.broadinstitute.org/morpheus/). Clustering was performed with Euclidean distance and complete linkage. Blue indicates low correlation and red indicates higher correlation between groups.

Principle component analysis (PCA) confirmed the expected grouping among cell lines and treatment group replicates, when data was transformed to principal components (PC). A 3D PCA plot was generated where PC1 accounts for 71.95 % of variance, PC2 for 19.04 % and PC3 for a minor 1.50 % of variance. The plot further confirmed replicate fidelity whilst also showcasing intergroup variance as a function of cell line and treatment received.
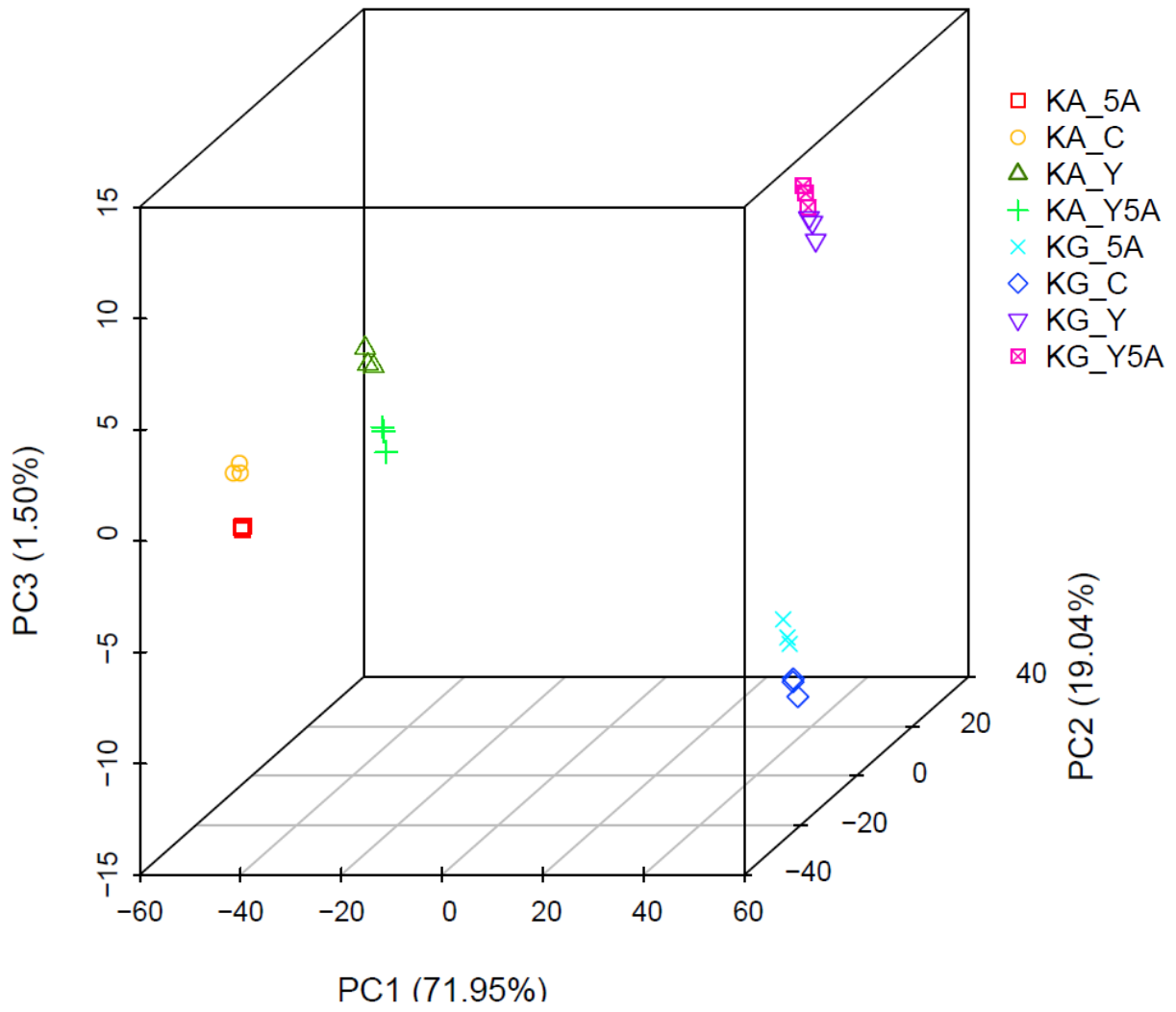
Figure 32: PC analysis plot displaying all 24 samples as a function of PC1, PC2 and PC3 within the complete transcriptomics expression data set.

4.2.1.2 All treatments induced differential expression of transcripts in both cell lines

The data revealed a range of expression changes after treatment with IFNG and 5AzaC in both cell lines. The most pronounced upregulation was observed for GBP1 (Log2 FC = +18.81) in Kasumi-1 cells treated with a combination of IFNG and 5AzaC. Conversely, the largest decrease was noted in OLFM4 (Log2 FC = -2.18) in Kasumi-1 cells treated with 5AzaC alone. Kasumi-1 exhibited the most pronounced response to both treatments compared to KG-1, which showed a more subdued response. Nonetheless, KG-1 still demonstrated significant upregulation in certain transcripts, such as PNMA5 (Log2 FC = +11.68). A comparison of the number of significantly differentially expressed transcripts between the two cell lines is illustrated as a bar chart (Figure 33). The Kasumi-1 cell line was consistently more responsive to each treatment than its KG-1 counterpart, with 2 to 4-fold difference in number of transcripts effected. For example, IFNG treatment deregulated 11,377 transcripts in Kasumi-1 compared to 2,564 in KG-1. According to this data, Kasumi-1 was more responsive to treatment with IFNG and 5AzaC compared to KG-1. The NGS and proteomics data sets were used to investigate the IFNG signalling pathway in both cell lines and identify the fundamental molecular differences.
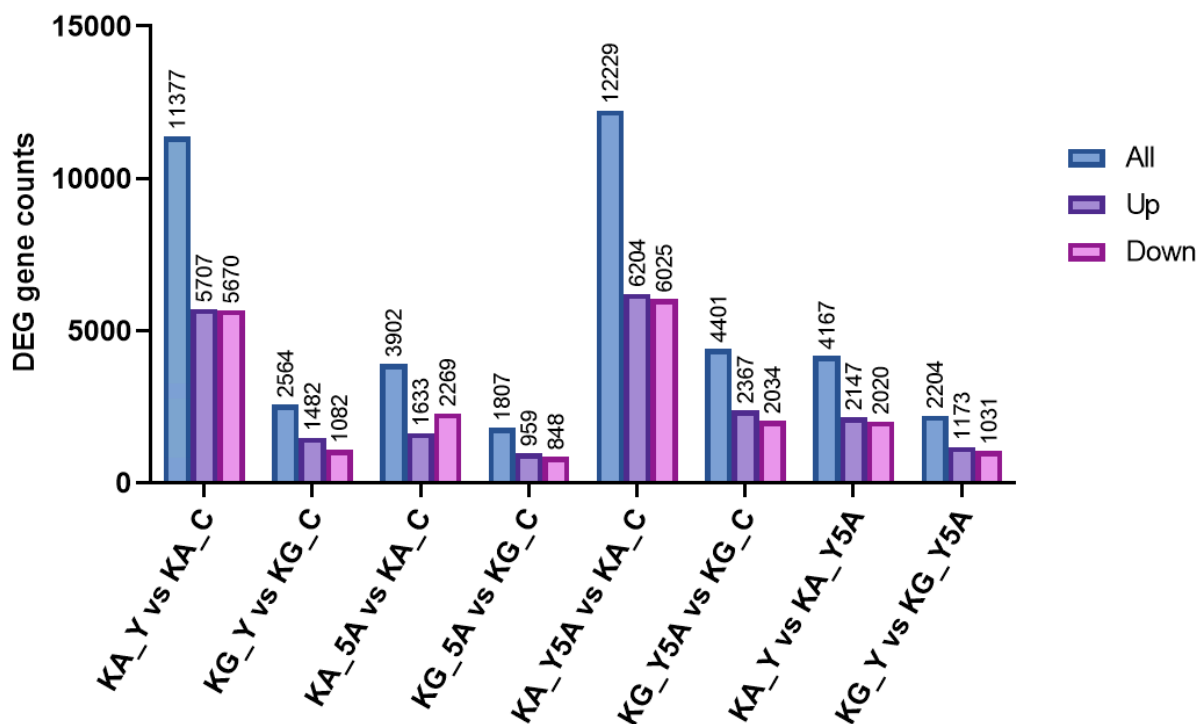
Figure 33: The number of transcripts significantly differentially expressed between treated and untreated Kasumi-1 and KG-1 cell lines. KA = Kasumi-1, KG = KG-1, C = Untreated control, Y = IFNG, 5A = 5AzaC, and Y5A = IFNG and 5AzaCytidine treatment. Blue = Total number of transcripts significantly differentially expressed (Includes up and down regulation), Purple = Number of transcripts which were significantly upregulated, and Pink = Number of transcripts significantly down regulated.

4.2.2 IFNG induced four times more differentially expressed transcripts in Kasumi-1 than KG-1

Kasumi-1 and KG-1 have shown different sensitivity to IFNG induction, with Kasumi-1 displaying a hyper response and KG-1 a muted response. This deviation was further evident in the transcriptomics data. Analysis of significantly differentially expressed transcripts between cell lines and treatments revealed a total of 11,377 and 2,564 transcripts altered in Kasumi-1 and KG-1 with IFNG treatment, respectively. These differences were visualised as a heat map (Figure 34).

Error! Reference source not found.
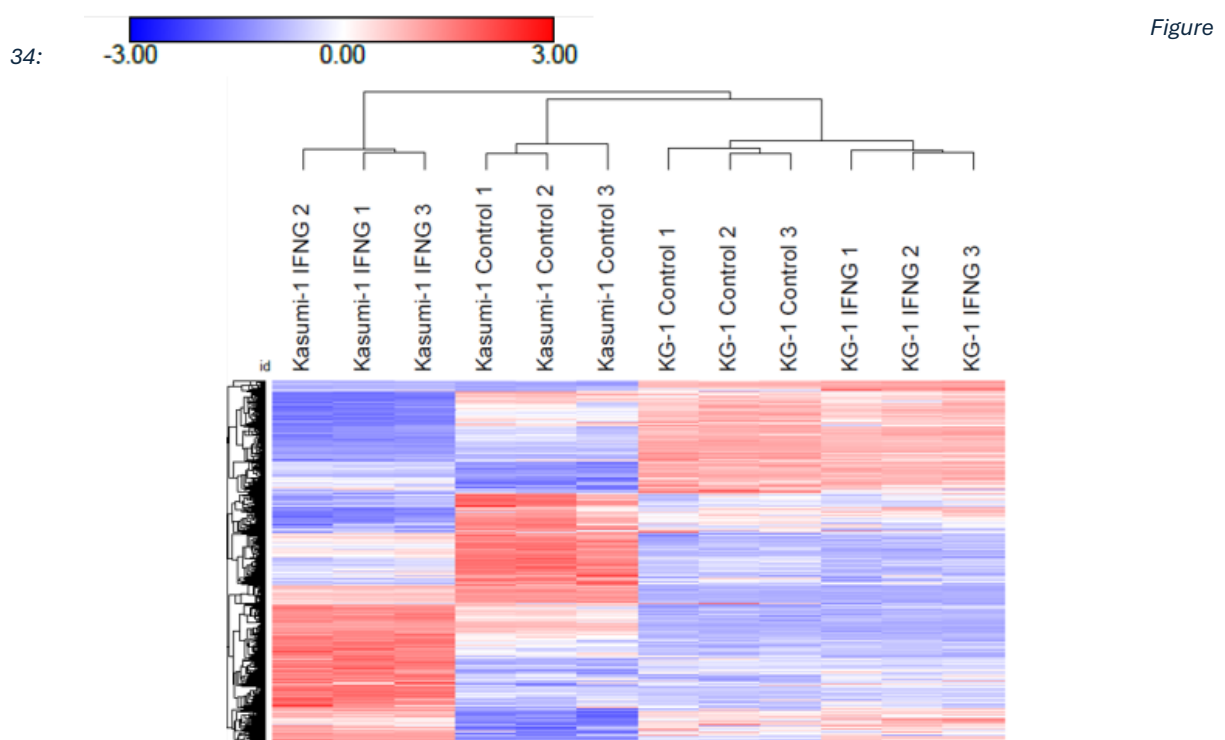
Figure 34:

Hierarchical clustering of the significantly differentially expressed transcripts between untreated and IFNG treated cell lines. Clustered using Euclidean distance and complete linkage (n = 3 per sample type). Key for heat maps shown top left for expression values. Heat maps generated in Morpheus (https://software.broadinstitute.org/morpheus/). Blue indicates lower expression; red indicates higher expression.

Transcript expression was transformed using Log2(FPKM+1) and Z values used, data was then subject to hierarchical clustering using Euclidean distance and complete linkage. Distinct clusters were assigned to treatment groups of cell lines, indicating consistent treatment across samples. Kasumi-1 showed clear clusters where IFNG induced upregulation and downregulation of transcripts compared to untreated control cells. In KG-1 these differences were harder to see by heatmap as far fewer transcripts were differentially expressed in response to the treatments.

The IFNG signalling cascade was illustrated in chapter 3 (Figure 9). In short IFNG binds to the IFNGR, which activates the Jak1 and Jak2 kinases to phosphorylate STAT1, which then translocates to the nucleus to initiate IFNG regulated transcription. The NGS data was examined for expression of early IFNG induced transcripts *TRIM21*, *IRF1*, *IFR8* and *PML* in Kasumi-1 and KG-1. These transcripts are induced early in IFNG signalling and *TRIM21*, *IRF1* and *IRF8* are required to form the IRF1-IRF8 complex that initiates the next set of transcript expression in the IFNG signalling cascade (Ozato, et al. 2008). Therefore, reduction in transcripts at this level could explain the reduced IFNG response seen in KG-1.

In both cell lines, IFNG treatment significantly upregulated expression of *TRIM21,* (Kas-1, Ctrl = 14.53 FPKM, IFNG = 82.45 FPKM, KG-1, Ctrl = 15.60 FPKM, IFNG = 36.94 FPKM) *IRF1,* (Kas-1, Ctrl = 3.47 FPKM, IFNG = 79.39 FPKM, KG-1, Ctrl = 6.24 FPKM, IFNG = 28.97 FPKM) and *PML,* (Kas-1, Ctrl = 5.22 FPKM, IFNG = 35.83 FPKM, KG-1, Ctrl = 8.25 FPKM, IFNG = 12.18 FPKM) (All comparisons, Padj < 0.0001, Wald test- BH). Comparisons are shown as graphs in Figure 35.

IFNG induced expression of these transcripts was higher in Kasumi-1 than in KG-1. In Kasumi-1 combination of IFNG and 5AzaC induced increased expression of *TRIM21* (Kas-1 IFNG5AzaC = 92.20 FPKM) and *PML* (Kas-1 IFNG5AzaC = 92.20 FPKM), than IFNG treatment on its own did (Both, Padj < 0.0001, Wald test- BH). *IRF8* was uniquely and significantly downregulated in Kasumi-1 in response to IFNG and further significantly downregulated by combining IFNG with 5AzaC when compared to only IFNG treatment (Ctrl = 4.07 FPKM, IFNG = 2.49 FPKM, IFNG5AzaC = 1.59 FPKM, Both, Padj < 0.0001, Wald test- BH).

The disparity in transcript expression was present in the first wave of IFNG induced transcripts, where IFNG induced expression of *IRF1, TRIM21* and *PML* was much higher in Kasumi-1 than in KG-1. Combined with low transcription of *IRF8* in KG-1 (<1 FPKM), decreased transcription of *TRIM21* and *IRF1* could have a knock-on effect that decreases transcription of downstream targets such as PML compared to Kasumi-1.
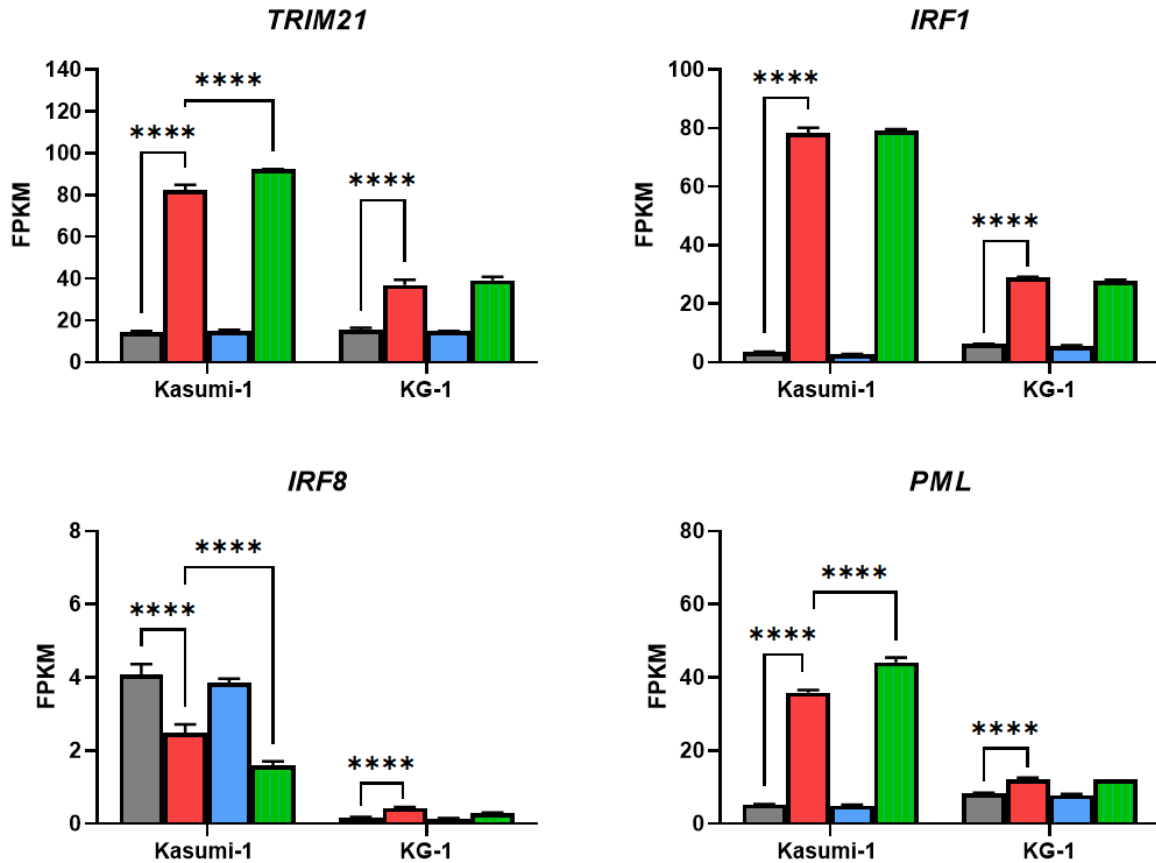
*Figure 35: Expression TRIM21, IRF1, IRF8 and PML in Kasumi-1 and KG-1 in response to treatments. Data given as FPKM (n=3). Grey = Control, Red = IFNG, Blue = 5AzaC, and Green = IFNG + 5AzaC. Statistical tests by Novogene Wald test – BH correction. * = Padj<0.05, ** = Padj< 0.01, *** = Padj< 0.001 and **** = Padj< 0.0001.*

4.2.2.2   IFNG signalling docking components transcripts are upregulated in Kasumi-1 by IFNG

Consistent with the IFNGR1 protein expression observed via flow cytometry (Figure 23), IFNGR1 mRNA was significantly higher in KG-1 compared to Kasumi-1 (KG-1 Ctrl = 33.89 FPKM and Kas-1 Ctrl = 33.89 FPKM, Padj < 0.0001, Wald test- BH). KG-1 also showed elevated levels of JAK1 mRNA (KG-1 Ctrl = 38.99 FPKM) and JAK2 mRNA (KG-1 Ctrl = 7.04 FPKM) compared to Kasumi-1. Treatment of Kasumi-1 with IFNG led to significant increases in IFNGR1 (Kas-1 Ctrl = 33.89 FPKM, IFNG = 14.13 FPKM, Padj < 0.05, Wald test-BH), IFNGR2 (Kas-1 Ctrl = 12.87 FPKM, IFNG = 16.11 FPKM, Padj < 0.0001), JAK1 (Kas-1 Ctrl = 11.92 FPKM, IFNG = 16.55 FPKM, Padj < 0.0001, Wald test-BH), and JAK2 (Kas-1 Ctrl = 3.95 FPKM, IFNG = 25.62 FPKM, Padj < 0.0001, Wald test-BH).

IFNG treatment of Kasumi-1 and KG-1 induced changes in IFNG signalling transcripts involved in creating the docking sites for STAT1, data is shown in Figure 36. In Kasumi-1, combination of IFNG and 5AzaC significantly increased expression of *JAK1* (IFNG5AzaC = 21.95 FPKM) and *JAK2* (IFNG5AzaC = 31.28 FPKM) in comparison to IFNG only treatment (Both, Padj < 0.0001, Wald test-BH). *STAT1* transcript was expressed higher in KG-1 than Kasumi-1 (KG-1 Ctrl = 96.98 FPKM, Kas-

1 Ctrl = 19.62 FPKM, Padj < 0.0001, Wald test-BH) as was STAT1 protein expression (KG-1 Ctrl = 136,478 NPA, Kas-1 Ctrl = 547,440 NPA)(Figure 36). IFNG treatment induced significant upregulation of *STAT1* in both cell lines compared to untreated control (Kas-1 IFNG = 744.40 FPKM, KG-1 IFNG = 966.11 FPKM, both, Padj < 0.0001). The combination of IFNG and 5AzaC reduced expression of *STAT1* in comparison to IFNG only treatment in KG-1 at the transcript level (KG-1 IFNG5AzaC = 897.92 FPKM, Padj < 0.05, Wald test-BH. In summary IFNG significantly upregulated transcripts for IFNG signalling components *STAT1, IFNGR1*, *IFNGR2*, *JAK1* and *JAK2* in Kasumi-1, and minorly *JAK2* in KG-1 (KG-1 Ctrl = 7.03 FPKM, IFNG = 7.91 FPKM, Padj < 0.05, Wald test-BH). STAT1 protein was also significantly induced with IFNG in both cell lines as determined by SWATH-MS in chapter 3 (Kas-1 Ctrl = 125,437 NPA, IFNG = 1,959,620 NPA, KG-1 Ctrl = 782,133 NPA, IFNG = 2,378,399 NPA, both, Padj <0.001, Holm-sidak). In summary, IFNG treatment significantly upregulated key signalling molecules (IFNGR1, IFNGR2, JAK1, JAK2, STAT1) in Kasumi-1 cells, but only JAK2 was significantly upregulated in the KG-1 cells. These results align with the proteomics data in Chapter 3, which showed Kasumi-1 had a robust response to IFNG, evidenced by 141 differentially expressed proteins and 7 upregulated IRDS proteins, compared to 48 proteins and 2 IRDS proteins in KG-1
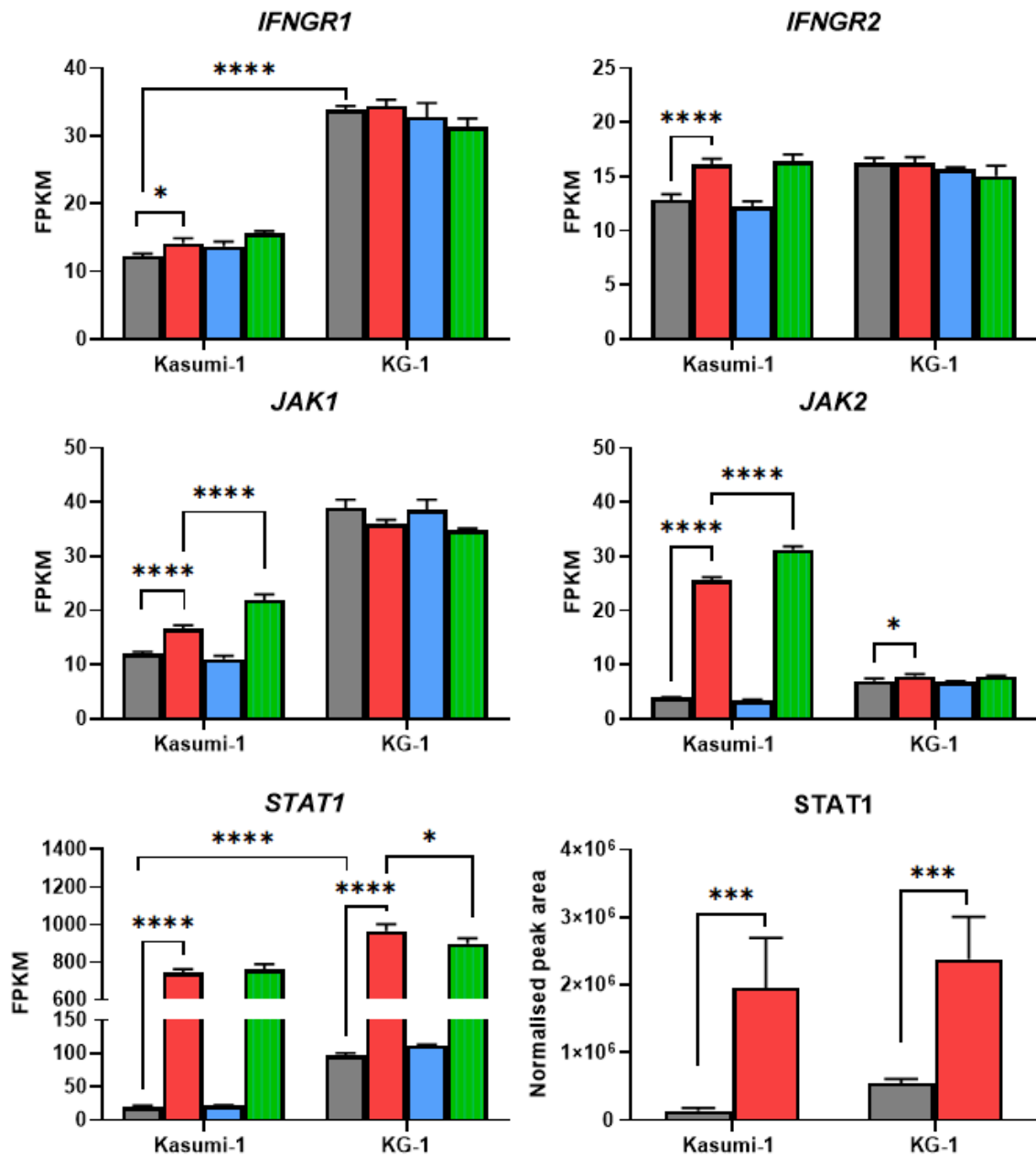
Figure 36: Expression of IFNGR1, IFNG2, JAK1, JAK2 and STAT1 transcripts and STAT1 protein in Kasumi-1 and KG-1 in response to treatments. Transcript data given as FPKM (n=3) and protein data as NPA (n=5/6). Grey = Control, Red = IFNG, Blue = 5AzaC, and Green = IFNG + 5AzaC. Statistical tests by Novogene Wald test – BH correction. Bottom right (STAT1): Normalised peak area expression of STAT1 protein in cell lines, as measured by SWATH-MS, control compared to 48 hrs treatment 100 ng/ml IFNG (n=5-6). Grey; Control, Red; IFNG. Holm-Sidak method was used to calculate statistical significance between IFNG treated and untreated cells. * = Padj< 0.05, ** = Padj< 0.01, *** = Padj< 0.001 and **** = Padj< 0.0001.

4.2.2.3   SOCS1 expression was upregulated by IFNG in both cell lines

As the differences in signalling intensity occur before the first wave of transcription, the disparity in signalling must occur between signal transduction and the STAT1 homodimer binding to DNA. SOCS1 is a protein which reduces IFNG signalling by binding to Jak1/2 to prevent phosphorylation

of STAT1 (Schroder, et al. 2004). TRIM8 is also an IFNG inducible protein whose interaction with SOCS-1 has been shown to degrade it (Toniato, et al. 2002). Data gathered by NGS showed that *SOCS1* transcripts were expressed by KG-1 but not by Kasumi-1 (<1 FPKM). Figure 37 shows IFNG induced SOCS*1* expression significantly in both cell lines (Kas-1 Ctrl = 0.02 FPKM, IFNG = 15.93 FPKM, KG-1 Ctrl = 5.57 FPKM, IFNG = 20.01 FPKM, Padj < 0.0001, Wald-test BH). In Kasumi-1, combining IFNG with 5AzaC further increased *SOCS1* expression than IFNG alone (Kas-1 IFNG5AzaC = 21.81 FPKM, Padj < 0.0001, Wald-test BH). *TRIM8* expression (Figure 37) was similar between untreated cell lines and was decreased in response to IFNG in both cell lines to a similar level (Kas-1 Ctrl = 8.09 FPKM, IFNG = 6.39 FPKM, Padj < 0.0001, Wald-test BH, KG-1 Ctrl = 8.21 FPKM, IFNG = 7.01 FPKM, Padj < 0.01, Wald-test BH). Treatment with 5AzaC uniquely upregulated TRIM8 in Kasumi-1 (Kas-1 5AzaC = 10.82 FPKM, Padj < 0.0001, Wald-test BH). *SOCS1* and *TRIM8* were expressed at similar levels in the presence of IFNG in both cell lines. It is, therefore, unlikely the large differences in IFNG signalling were SOCS1 dependent. However, transcript abundance is not necessarily an indicator of protein abundance, as many factors such as RNA degradation can prevent translation to protein. As SOCS1 and TRIM8 data were not present in the SWATH-MS data, no conclusions could be drawn about what was happening in the cell lines at the protein level.
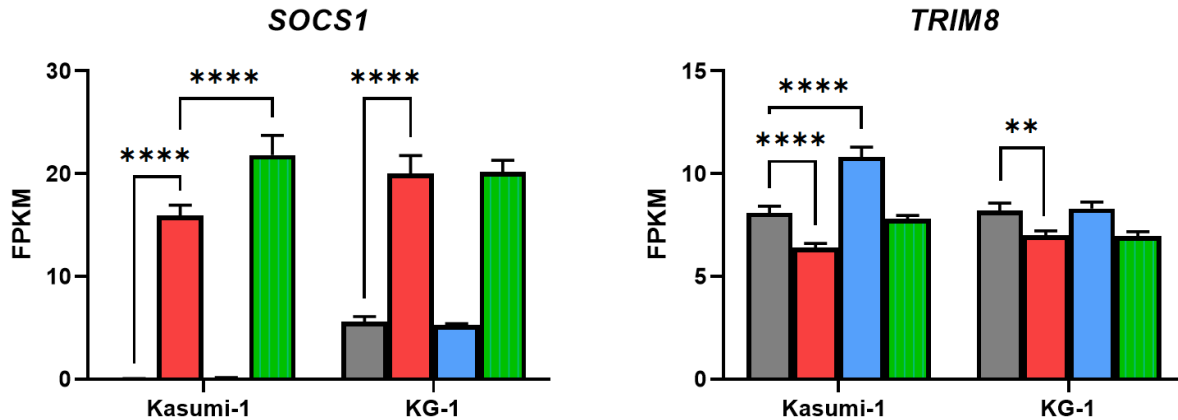
*Figure 37: Expression of SOCS1 and TRIM8 transcripts in Kasumi-1 and KG-1 in response to treatments. Data is given as FPKM (n=3). Grey = Control, Red = IFNG, Blue = 5AzaC, and Green = IFNG + 5AzaC. Statistical tests by Novogene Wald test – BH correction. * = Padj<0.05, ** = Padj< 0.01, *** = Padj< 0.001 and **** = Padj< 0.0001.*

4.2.2.4 High expression of SHP1 by KG-1 could inhibit IFNG signalling

Protein tyrosine phosphatases (PTPs) like SHP1 and SHP2, disrupt IFNG signalling by obstructing STAT1 phosphorylation (Schroder, et al. 2004). PTPs in the cytosol can dephosphorylate Jak1/2 and IFNGR1, thereby disrupting the phosphorylation chain and preventing STAT1 phosphorylation, as shown in Figure 9. This reduces the number of STAT1 proteins successfully phosphorylated, stopping homodimers forming to activate transcription. PTPs in the nucleus can also directly dephosphorylate STAT1 homodimers, rendering them inactive and causing their nuclear export (Schroder, et al. 2004).

NGS and proteomics data showed KG-1 expressed more SHP2 protein than Kasumi-1 (Kas-1 = 36,615 NPA, KG-1 = 104,270 NPA, Padj <0.01, Holm-Sidak). However, expression of its transcript *PTPN11* was similar between the cell lines (Kas-1 = 39.45 FPKM, KG-1 = 39.71 FPKM, NS, Wald-test, BH) (Figure 38). Expression of SHP1 differed between cell lines (Figure 38). Kasumi-1 expressed lower levels of SHP1 protein compared to KG-1 (Kas-1= 114,481 NPA and KG-1= 1,708,676 NPA, Padj < 0.0001, Holm-Sidak) along with its transcript *PTPN6* (Kas-1= 5.09 FPKM and KG-1= 68.59 FPKM, Padj < 0.0001, Wald-test, BH). It is possible high baseline levels of SHP1 in KG-1 could prevent IFNG signalling occurring to the same degree as Kasumi-1, which has low baseline SHP1 expression. Furthermore, Kasumi-1 upregulated the *PTPN6* transcript and corresponding SHP1 protein (*PTPN6* expression, Control = 5.09 FPKM, IFNG = 11.89 FPKM, Padj < 0.0001, Wald-test, BH, SHP1 expression, Control = 114,481 NPA, IFNG = 297,412 NPA, Padj< 0.01, Holm-Sidak) in response to IFNG, thereby displaying a potential negative feedback loop for IFNG signalling, indicative of a functional IFNG response. Meanwhile, IFNG significantly downregulated *PTPN6* in KG-1 (Control = 68.59 FPKM, IFNG = 60.14 FPKM, Padj < 0.001, Wald-test, BH) and the protein product was decreased but not reported as significant in the SWATH-MS data (Control = 1,708,676 NPA, IFNG = 1,581,148 NPA, NS, Holm-Sidak). High SHP1 expression could have allowed KG-1 to remain IFNG-resistant, while its low expression in Kasumi-1 might have enabled a hyper-response to IFNG. According to SWATH-MS data from Chapter 3, even with IFNG treatment, Kasumi-1 SHP1 levels remained below those in KG-1 and the other cell lines (data presented in appendix 7.6), indicating a dampened negative feedback loop, which could enable the more intense and prolonged IFNG response, displayed by Kasumi-1.
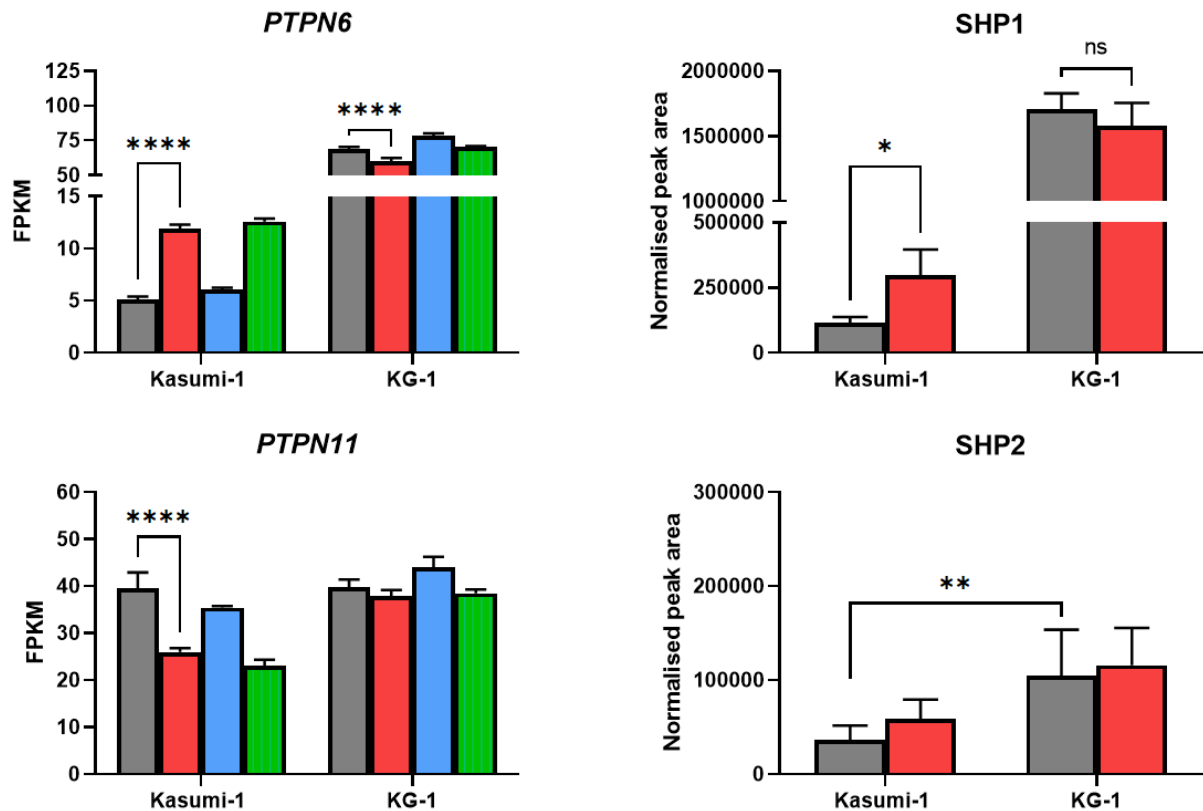
Figure 38: Expression of PTPN6 and PTPN11 transcripts and matching SHP1 and SHP2 proteins in Kasumi-1 and KG-1 in response to treatments. Data given as FPKM (n=3). Grey = Control, Red = IFNG, Blue = 5AzaC, and Green = IFNG + 5AzaC. Statistical tests by Novogene Wald test – BH correction. B) Normalised peak area expression of SHP1 and SHP2 across all four cell lines, control compared to 48 hrs treatment 100 ng/ml IFNG (n=5-6). Holm-Sidak method was used to calculate statistical significance between IFNG treated and untreated cells. * = Padj< 0.05, ** = Padj< 0.01, *** = Padj< 0.001 and **** = Padj< 0.0001.

4.2.3 Metascape analysis showed IFNG induced a unique biological response in each cell line

The IFNG signalling pathway was operating differently between cell lines, possibly due to differential expression of SHP1. The intensity of signalling was distinct between cell lines. To investigate how the divergent IFNG signalling intensity effected biological response to IFNG, a comparison of differentially expressed transcripts was performed using Metascape.

Transcript lists for each cell line were filtered for differential expression Padj < 0.05 and a fold change of 1.5 or more, then separated into two lists of increased (up-regulated) or decreased (down-regulated) expression. Venn diagrams were drawn comparing the lists of increased and decreased transcripts between each cell line in Figure 39 A and B, respectively. The large gap in IFNG signalling intensity had a subsequent impact on breadth of IFNG response, illustrated when comparing the number of transcripts differentially expressed following IFNG treatment between

cell lines. Venn diagrams clearly show IFNG treatment more significantly impacted transcript expression in Kasumi-1 than KG-1, with thousands of transcripts differentially expressed compared to KG-1's hundreds.
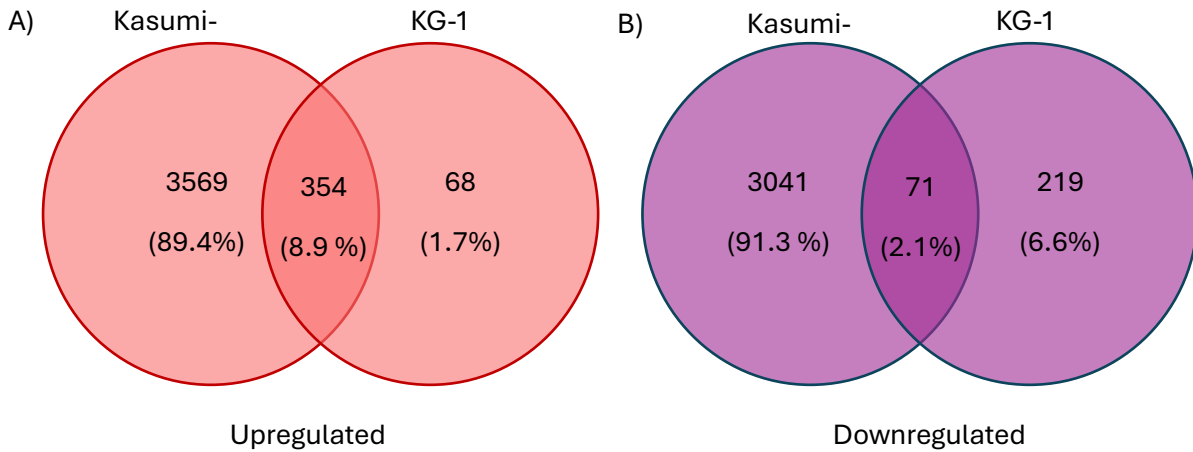
Figure 39: Venn diagrams comparing transcript lists of differentially expressed transcripts of Kasumi-1 and KG-1 treated with IFNG. Differential expression was defined as a change of 1.5-fold or more that was also statistically significant (Padj < 0.0.5). Comparison of up-regulated transcripts depicted in A, and down regulated transcripts displayed in B.

Transcript lists unique to each cell line, as defined by Venn diagram, were uploaded to Metascape for pathway enrichment. The top 10 significantly enriched pathways according to the increased and decreased transcript lists are displayed for Kasumi-1 and KG-1 in Figure 40, top row and bottom row respectively. The Metascape tool used various databases, including KEGG Pathway, GO Biological Processes, Reactome Gene Sets, Canonical Pathways, CORUM, TRRUST, DisGeNET, PaGenBase, Transcription Factor Targets, WikiPathways, and PANTHER Pathway, to conduct its analysis. The tool employed the entire genome as the background for its enrichment. For a process to be considered significant, it had to be identified in the input gene list at least three times, with a ratio of observed counts to chance counts being greater than 1.5 times. The identified processes meeting these criteria were then grouped into clusters based on similarities, and the most statistically significant member was selected to represent the cluster.
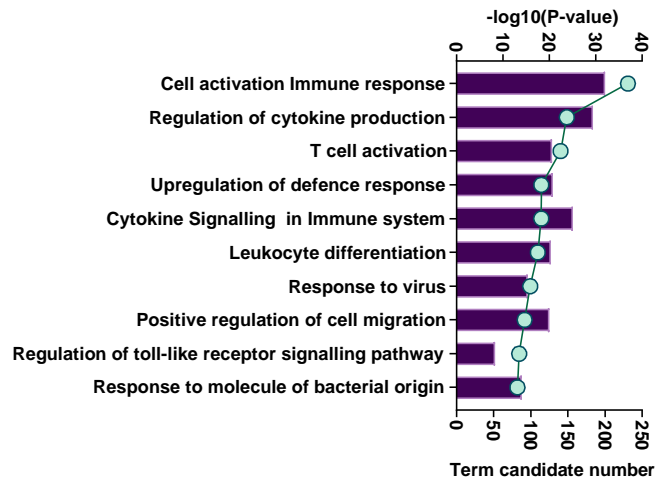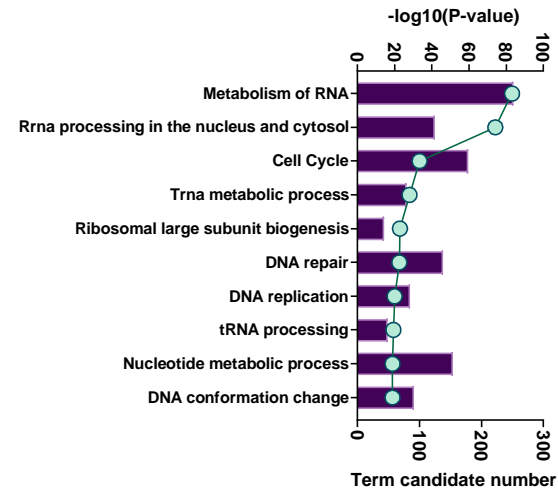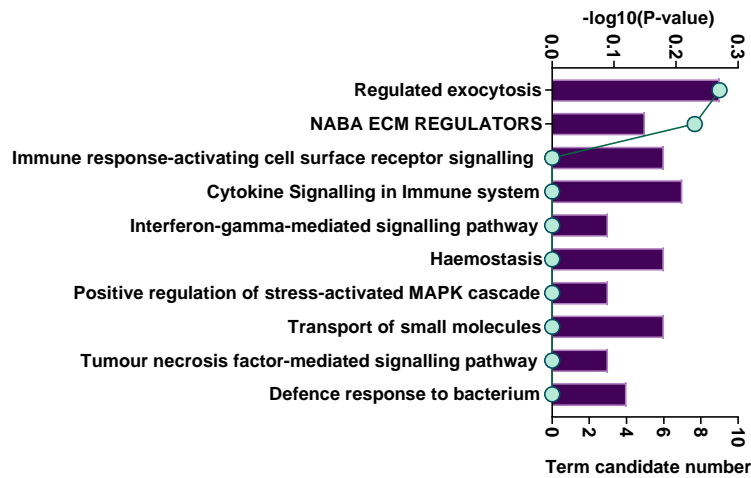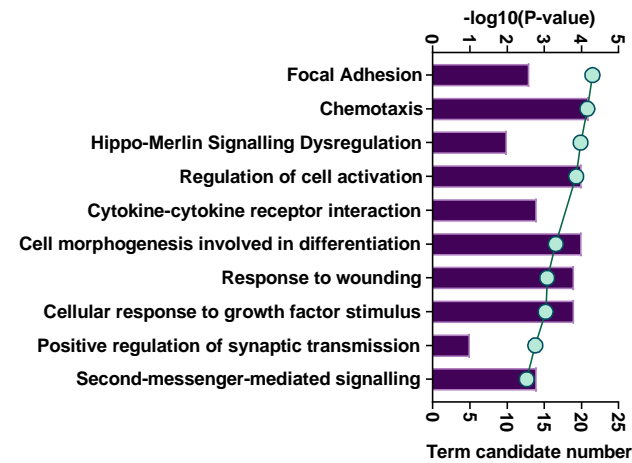
Figure 40: The top 10 most significantly enriched pathways in Kasumi-1 and KG-1 treated with IFNG. Enriched pathways were identified by the 'unique' up and down regulated transcript lists and sorted by significance post FDR. See appendix for table summary process id's and number of transcripts significantly changed in each pathway.

Pathway analysis showed that changes in Kasumi-1 induced by IFNG were more robust with large numbers of transcripts in pathways being altered, for example as many as 252 out of 673 transcripts for metabolism of RNA. As opposed to in KG-1 where a maximum of 21 out of 644 transcripts were significantly deregulated in the chemotaxis process. Cell lines shared some similarities in responses, both upregulating IFNG viral and immune defence mechanisms. Enriched pathways reported for Kasumi-1 included immune response-based processes, mostly for mobilisation and activation of immune cells, as well as cytokine signalling and inflammation response (Figure 40).

4.2.4 Investigation of IFNG Induced apoptosis marker transcripts in AML cell lines

This section of work focuses on the analysis of the Kasumi-1 cell line's response to IFNG treatment and its potential role in inducing apoptosis as well as an immunosuppressive phenotype. The hypothesis for this work is that IFNG induces expression of immunosuppressive molecules in AML. In Chapter 3, it was observed that IFNG induced immunosuppressive molecules in both cell lines. However, IFNG also triggered cell death specifically in Kasumi-1 cells. In this study, cells were treated with 5 ng/mL of IFNG, and the expression of transcripts involved in apoptosis was analysed. The chosen IFNG dose of 5 ng/mL was based on a dose selection study, which showed no increase in cell death in Kasumi-1 cells when treated with concentrations higher than 5 ng/mL. To characterise IFNG induced apoptosis, the expression of stress sensors ATF3 and BMF, as well as transcripts of genes involved in the P53 pathway for growth arrest, such as CDKN1A and GADD45, were analysed. Additionally, IFIT-mediated apoptosis pathway transcripts were also investigated. Dose experimental data is presented in graphs in the appendix (see 7.4).

4.2.4.1 Treatment with IFNG induced transcription of stress sensors exclusively in Kasumi-1

ATF3 is a stress responding transcription factor which regulates transcripts in response to DNA damage, furthermore it has been found to co-localise with p53 influencing its DNA damage transcriptional program (Zhao, et al. 2016). BMF has been described as a 'sentinel', that upon detecting cytoskeletal damage, initiates apoptosis via cytochrome C release from the mitochondria (Hausmann, et al. 2011). In Kasumi-1, IFNG caused a significant increase in expression of stress sensors *ATF3* and *BMF* (*ATF3*, Ctrl = 1.05 FPKM, IFNG = 12.07 FPKM, and *BMF*, Ctrl = 1.55 FPKM, IFNG = 12.62 FPKM, both, Padj < 0.0001, Wald-test BH). Combination of IFNG with 5AzaC proved to be synergistic, causing a further significant increase compared to IFNG treatment alone in Kasumi-1 (*ATF3*, IFNG5AzaC = 21.25 FPKM, *BMF*, IFNG5AzaC = 17.68 FPKM, both, Padj < 0.0001, Wald test- BH) (Figure 41).
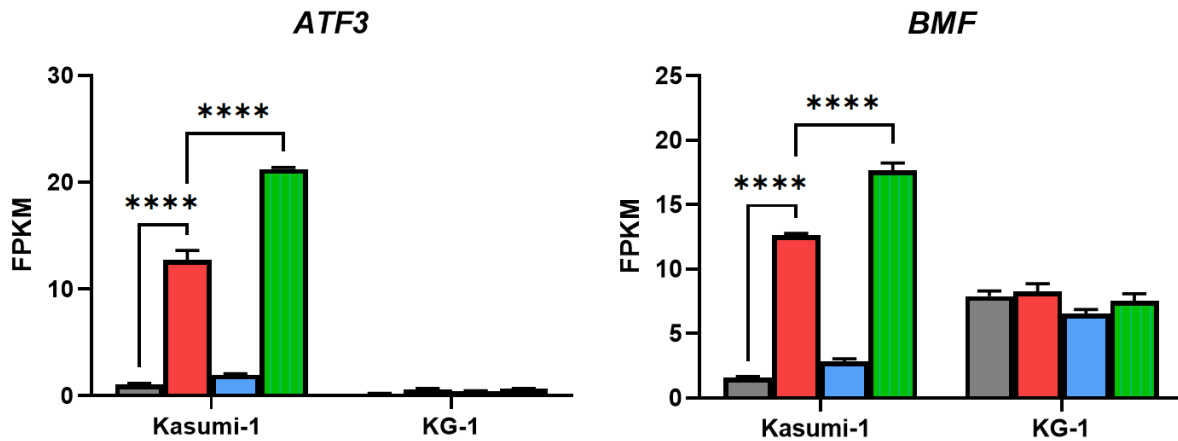
Figure 41: Transcript expression changes of ATF3 and BMF inKasumi-1 and KG-1 in response to treatments. Data given as FPKM (n=3). Grey = Control, Red = IFNG, Blue = 5AzaC, and Green = IFNG + 5AzaC. Statistical tests by Novogene Wald test – BH correction. * = Padj< 0.05, ** = Padj< 0.01, *** = Padj< 0.001 and **** = Padj< 0.0001.

4.2.4.2   IFNG induced expression of transcripts involved in p53 mediated growth arrest in Kasumi-1

P53 stops cell division by activating transcript programs for cell cycle arrest and apoptosis. Key transcripts transactivated by p53 for cell cycle arrest include *CDKN1A*, *SFN* and the *GADD45* transcripts (Benchimol 2001). Transcriptomics data showed (Figure 42) significant upregulation of *CDKN1A* with IFNG (Kas-1 Ctrl = 0.34 FPKM, IFNG = 11.39 FPKM, Padj < 0.0001, Wald test- BH), which was further increased by combination with 5zaC treatment in Kasumi-1 (Kas-1 IFNG5AzaC = 16.69 FPKM, Padj < 0.0001, Wald test- BH), but not KG-1. Expression of *SFN* is not presented as it remained low under all conditions (< 1 FPKM) in both cell lines and was unaffected by treatments. In Kasumi-1 IFNG induced significant increased expression of *GADD45A* (Kas-1 Ctrl = 12.78 FPKM, IFNG = 18.68 FPKM, Padj < 0.0001, Wald test- BH), *GADD45B* (Kas-1 Ctrl = 10.95 FPKM, IFNG = 30.48 FPKM, Padj < 0.0001, Wald test- BH) and *GADD45G* (Kas-1 Ctrl = 0.05 FPKM, IFNG = 1.19 FPKM, Padj < 0.0001, Wald test- BH). Combining IFNG with 5AzaC appeared to neutralise the inducing effects of IFNG for *GADD45A* in Kasumi-1, returning expression to a similar level to untreated cells (Kas-1 IFNG5AzaC = 13.08 FPKM). In Kasumi-1, *GADD45GIP1* expression was significantly decreased by IFNG treatment and 5AzaC treatment (Kas-1 IFNG = 30.16 FPKM, 5AzaC = 43.43 FPKM, both, Padj < 0.0001, Wald test- BH), but there was no synergistic effect when both treatments were combined (Kas-1 IFNG5AzaC = 30.27 FPKM). IFNG decreased expression of *GADD45B* in KG-1 IFNG (Padj < 0.0001, Wald test- BH).
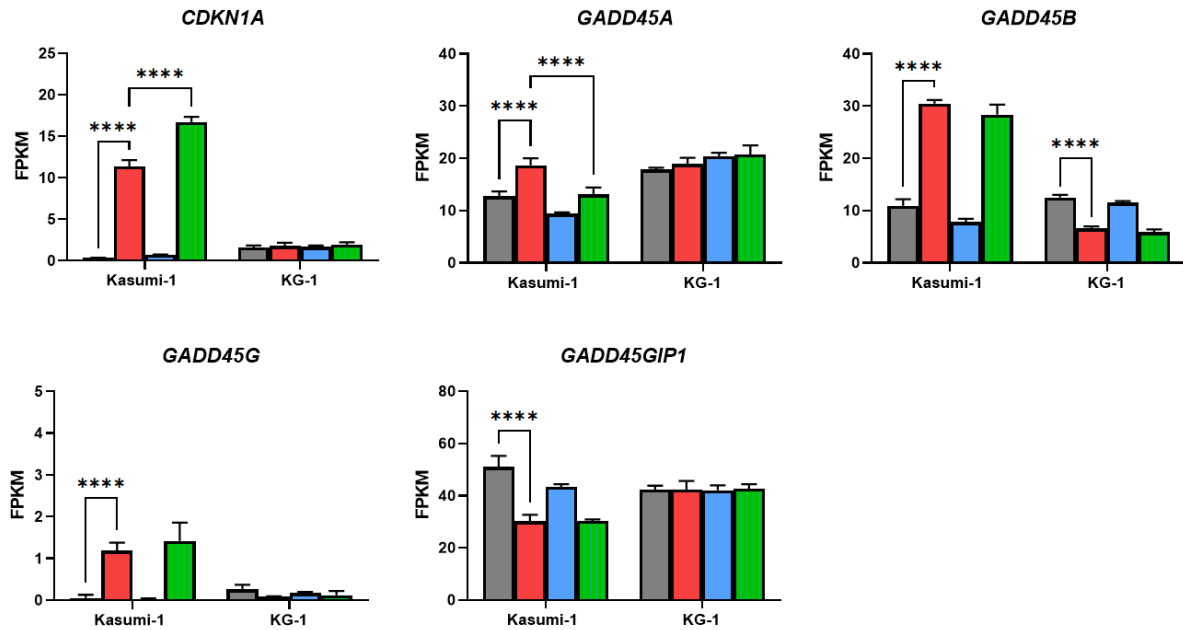
Figure 42: Expression changes of p53 regulated cell cycle arrest transcripts in Kasumi-1 and KG-1 cell lines in response to treatments. Data given as FPKM (n=3). Grey = Control, Red = IFNG, Blue = 5AzaC, and Green = IFNG + 5AzaC. Statistical tests by Novogene Wald test – BH correction. * = Padj< 0.05, ** = Padj< 0.01, *** = Padj< 0.001 and **** = Padj< 0.0001.

Given the significant IFNG-inducible expression of transcripts *CDKN1A, GADD45A, GADD45B, GADD45G,* and *GADD45GIP1* in the Kasumi-1 cell line, the SWATH-MS proteomic data from Chapter 3 was revisited to examine the expression of related proteins. This analysis revealed that, among these transcripts, protein products were detected only for CDKN1A (P21) and GADD45GIP1 (G45IP), as shown in Figure 43. This aligns with the previous FACS data in chapter 3, which reported no cell cycle arrest. No significant increase in P21 protein expression was observed in either cell line due to high variation between results, which could be attributed to several reasons outlined in chapter 2. Conversely, while IFNG induced reduction of *GADD45GIP1*, its protein counterpart, G45IP, was increased in the SWATH MS data (Kas-1 Ctrl = 137,417 NPA, IFNG = 1,164,951 NPA, KG-1 Ctrl =8,073,761 NPA, IFNG = 8,811,154 NPA). This discrepancy could be due to differences in IFNG doses causing varied responses or differences at the translation stage.
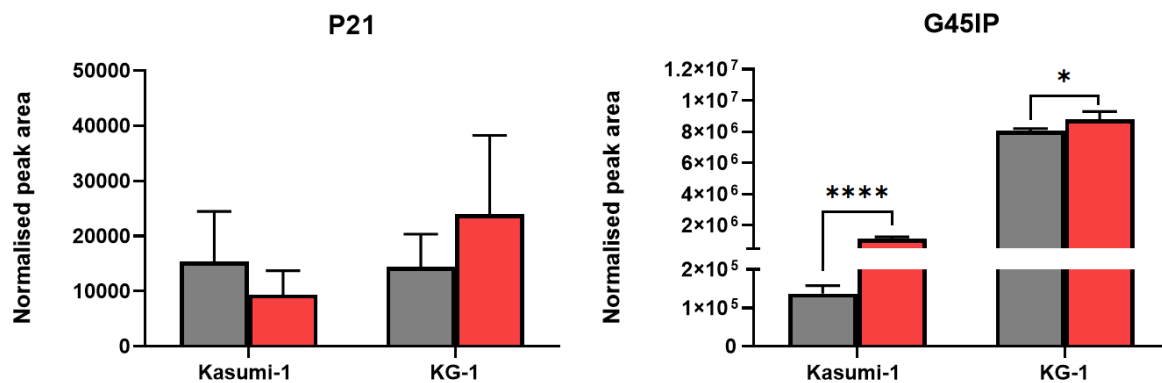
Figure 43: Expression of CDKN1A and GADD45GIP1 proteins in untreated and IFNG treated Kasumi-1 and KG-1. Untreated control was compared to IFNG treated samples 48 hrs after treatment with 100 ng/ml IFNG (n=5-6). Grey = Control, Red = IFNG. Holm-Sidak method was used to calculate statistical significance between IFNG treated aned untreated cells. * = Padj< 0.05, ** = Padj< 0.01, *** = Padj< 0.001 and **** = Padj< 0.0001.

These results suggested that low doses of IFNG might induce a p53 DNA damage response in Kasumi-1, resulting in cell cycle arrest, despite its mutational status. Transcript expression for *TP53* was reported in the NGS data to be higher for Kasumi-1 than in KG-1, which is unsurprising as KG-1 is a p53 null cell line (Figure 44 B). The average expression of the p53 protein was higher in KG-1 than in Kasumi-1 according to SWATH data (Figure 44 C). However, greater variation was observed, likely due to SWATH-MS's struggle to accurately report low abundance proteins, suggesting a possible false reading for KG-1 (See 7.6). Furthermore, the SWATH-MS data set was generated by treating cells with 100 ng/ml IFNG, in contrast to the NGS data set, which used a low dose of 5 ng/ml. Expression was also tested by western blot using the same treatment concentrations and schedules as those used to generate the NGS data. Results clearly showed that p53 protein was expressed in Kasumi-1, while only faint bands were visible in KG-1 (Figure 44). All three techniques indicated that treatment with IFNG had no impact on p53 expression at either the genetic or protein level.
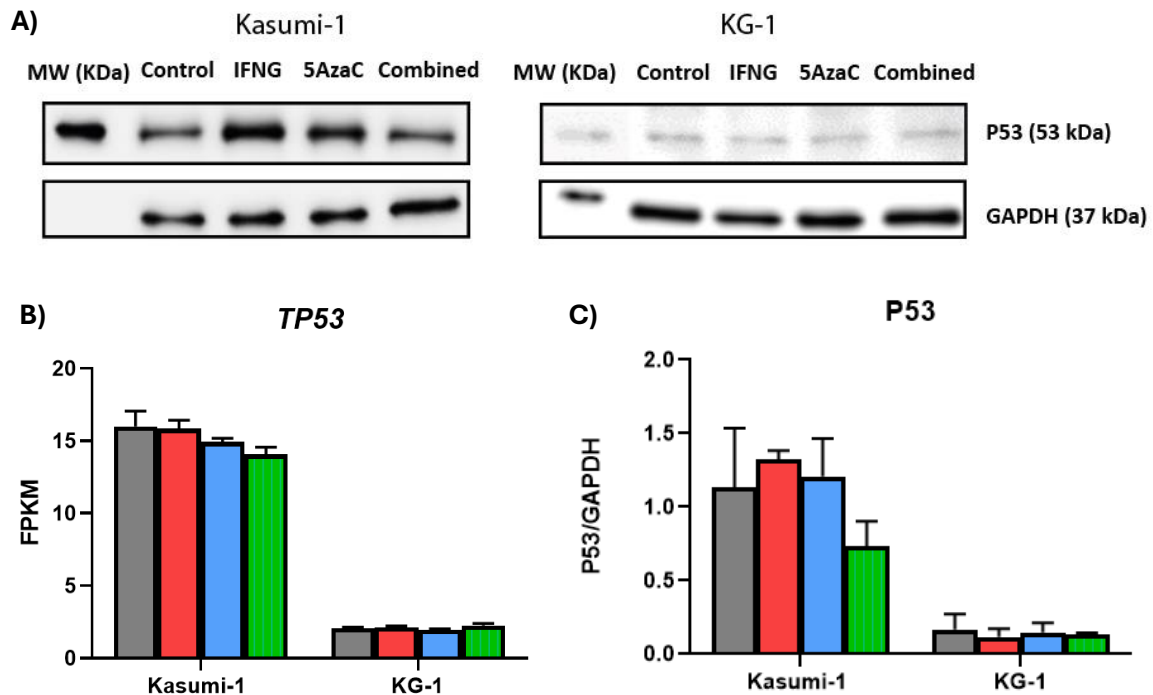
Figure 44: Expression of TP53 transcript and P53 protein by both cell lines following IFNG and 5AzaC treatments A) Western blots comparing p53 protein expression after 48 hrs of 5AzaC treatment and 30 minutes of IFNG, with GAPDH as an internal control (n=2). IFNG; 5ng/ml at for 30 minutes after 48 hrs of 5AzaC, 5AzaC; 0.5 μm at 0 and 24 hrs, combined; treatments together as stated. B) Expression of TP53 in both cell lines under all treatment conditions (n=3), data given as FPKM. C) Quantification of protein abundance of p53 as per western blot conditions stated in A. Grey = Control, Red = IFNG, Blue = 5AzaC, and Green = IFNG + 5AzaC. B) Quantitative densitometry of P53 expression for western blots shown in A).

4.2.4.3   Investigation of p53 dependent apoptosis in response to IFNG in Kasumi-1

Several downstream mediators of apoptosis contain p53 response elements and are transcribed by p53 upon stress detection. These transcripts can be categorised based on the localisation of their protein products: cell membrane (FAS, PERP, TNFRSF10B), cytosolic (PIDD1, EI24), and mitochondrial (BAX, PMAIP1, BBC3, TP53AIP1) (Benchimol 2001).

4.2.4.4   P53 death receptors may be activated independent of p53 in Kasumi-1

IFNG induced an up regulation of the death domain receptor *FAS* in Kasumi-1 (Ctrl = 0.05 FPKM, IFNG = 10.33 FPKM, $P_{adj} < 0.0001$, Wald test- BH) but decreased expression in KG-1 (Ctrl = 4.14 FPKM, IFNG = 2.45 FPKM, $P_{adj} < 0.0001$, Wald test- BH), while *PERP* (FPKM < 1) was unaffected and therefore was not presented here. *TNFRSF10B* was slightly downregulated in Kasumi-1 when treated with IFNG (Ctrl = 7.53 FPKM, IFNG = 6.47 FPKM, $P_{adj} < 0.05$, Wald test- BH), but upregulated by 5AzaC (5AzaC = 9.99 $P_{adj} < 0.0001$, Wald test- BH). Meanwhile, *TNFSF10* the p53 independent *TNFRSF10B* ligand, was highly upregulated in Kasumi-1 by IFNG, and even higher when combined with 5AzaC (Ctrl = 0.70 FPKM, IFNG = 138.95 FPKM, IFNG5AzaC = 191.18 FPKM, both $P_{adj} < 0.0001$, Wald test- BH). TNFSF10 was upregulated in KG-1 upon IFNG treatment,

though not to the extent observed in Kasumi-1 (Ctrl = 19.02 FPKM, IFNG = 26.99 FPKM, Padj < 0.05, Wald test- BH). The increase in *TNFSF10* from no mRNA (<0 FPKM) to over 140 FPKM, if reflected in protein expression, may induce activation of apoptosis by death domain TNFRSF10B, independent of p53 transcription. Further evidence this may be the case in Kasumi-1 is the significant increase in induction of *CASP3* (Ctrl = 23.42 FPKM, IFNG = 30.14 FPKM, Padj < 0.0001, Wald test- BH) and *CASP8* (Ctrl = 7.12 FPKM, IFNG = 14.40 FPKM, Padj < 0.0001, Wald test- BH) indicating apoptosis induction. Meanwhile *FADD,* ligand to FAS, was unaffected by IFNG treatment (Ctrl = 9.08 FPKM, IFNG = 9.27 FPKM, Padj < 0.05, Wald test- BH) (Figure 45).
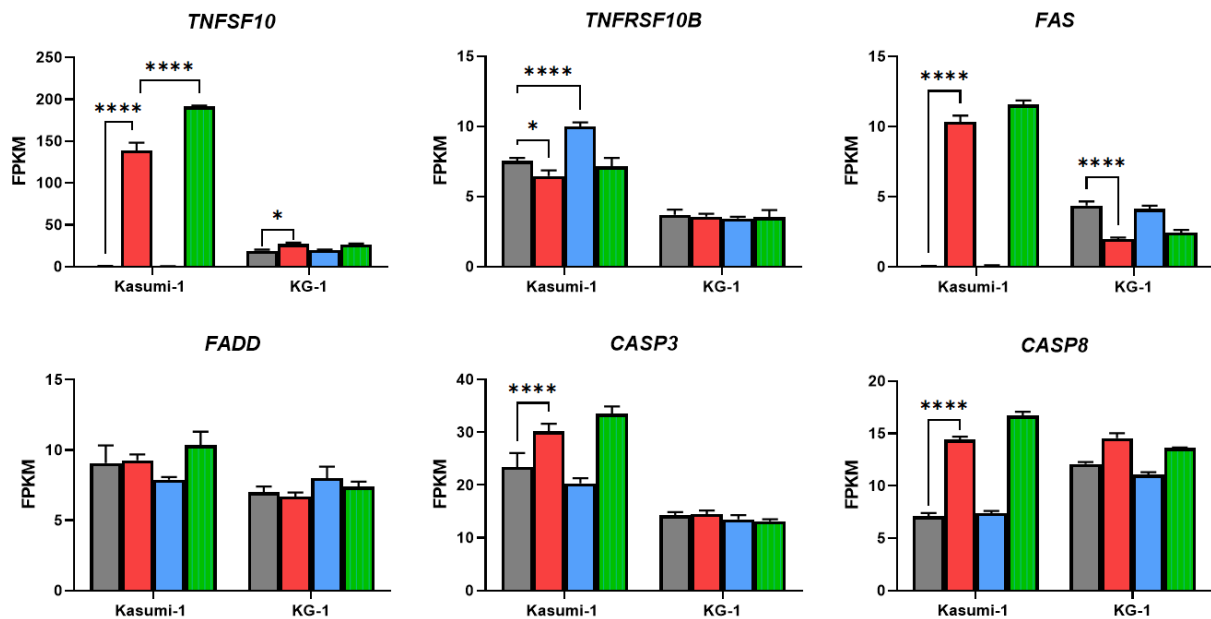
Figure 45: Expression of TNFSF10, TNFRSF10B, FAS, CASP3, CASP8 and FADD in Kasumi-1 and KG-1 under all treatment conditions. Data given as FPKM (n=3). Grey = Control, Red = IFNG, Blue = 5AzaC, and Green = IFNG + 5AzaC. Statistical tests by Novogene Wald test – BH correction. * = Padj<0.05, ** = Padj< 0.01, *** = Padj< 0.001 and **** = Padj< 0.0001.

4.2.4.5   IFNG highly upregulated expression of pro-apoptosis molecule PMAIP1 in Kasumi-1

PIDD1 and *EI24* overexpression has been linked to cell cycle arrest and induced apoptosis (Benchimol 2001). *PIDD1* expression was unaffected by treatments in both cell lines, while IFNG downregulated *EI24* in Kasumi-1 only (Ctrl = 33.05 FPKM, IFNG = 16.78 FPKM, Padj < 0.0001, Wald test-BH). In addition, the transcripts for mitochondrial proteins *BBC3* and *TP53AIP1* (<1FPKM) were very lowly expressed in both cell lines (Figure 46). Treatment with IFNG increased expression of *BBC3* (Ctrl = 0.39 FPKM, IFNG = 1.08 FPKM, Padj < 0.05, Wald test-BH) and combination treatment further increased its expression in Kasumi-1 (IFNG5AzaC = 1.98 FPKM, Padj < 0.0001, Wald test- BH). *PMAIP1,* which is a proapoptotic molecule (Janus, et al. 2020), was significantly induced by IFNG in Kasumi-1 and even further upregulated by combined IFNG with 5AzaC

treatment (Ctrl = 15.04 FPKM, IFNG = 49.77 FPKM, IFNG5AzaC = 72.44 FPKM, both, Padj < 0.0001, Wald test- BH). Meanwhile, IFNG downregulated *PMAIP1* in KG-1 (Ctrl = 18.53 FPKM, IFNG = 13.26 FPKM, Padj < 0.01, Wald test-BH) (Figure 46). The overexpression of PMAIP1 has been linked to apoptosis in other cancer cell lines, via induction of APAF1 (Kuroda, J., et al. 2010). Despite induction of *PMAIP1* (Padj < 0.0001, Wald test- BH)*, APAF1* was only minorly induced in Kasumi-1 (Ctrl = Padj < 0.05, Wald test- BH). Conversely, despite IFNG downregulating *PMAIP1* in KG-1, *APAF1* was upregulated with IFNG treatment in KG-1 (Both, Padj < 0.01, Wald test- BH).
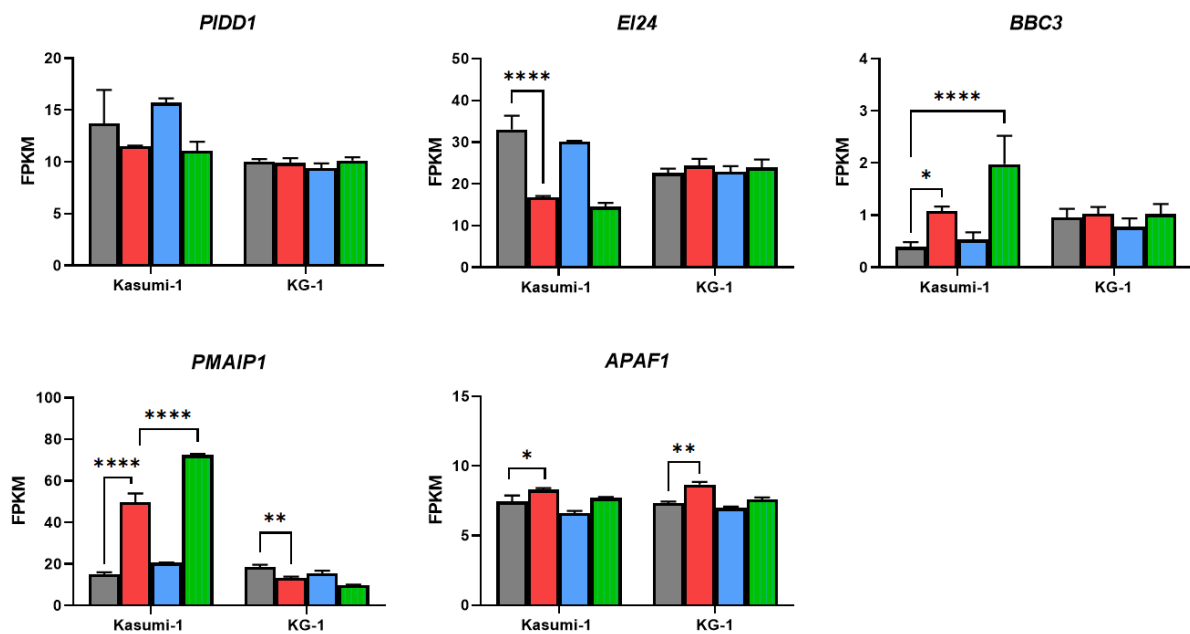
*Figure 46: Expression of PIDD1, EI24, BBC3, PMAIP1 and APAF1 in Kausmi-1 and KG-1 under all treatment conditions. Data given as FPKM (n=3). Grey = Control, Red = IFNG, Blue = 5AzaC, and Green = IFNG + 5AzaC. Statistical tests by Novogene Wald test – BH correction. * = Padj<0.05, ** = Padj< 0.01, *** = Padj< 0.001 and **** = Padj< 0.0001.*

4.2.4.6 Kasumi-1 apoptosis could be mediated through IFNG induced expression of IFIT proteins

Kasumi-1 displayed extremely high induction of all three IFIT transcripts and proteins in response to IFNG (Figure 47). IFIT1 and IFIT2 have been identified as IFNG inducible proteins which bind to MITA (mediator of IRF3 activation) to initiate apoptosis via the mitochondrial pathway (Stawowczyk, et al. 2011, Ohsugi, et al. 2017). Furthermore, apoptosis can occur even in the presence of non-functional p53 via these mechanisms (Stawowczyk, et al. 2011).
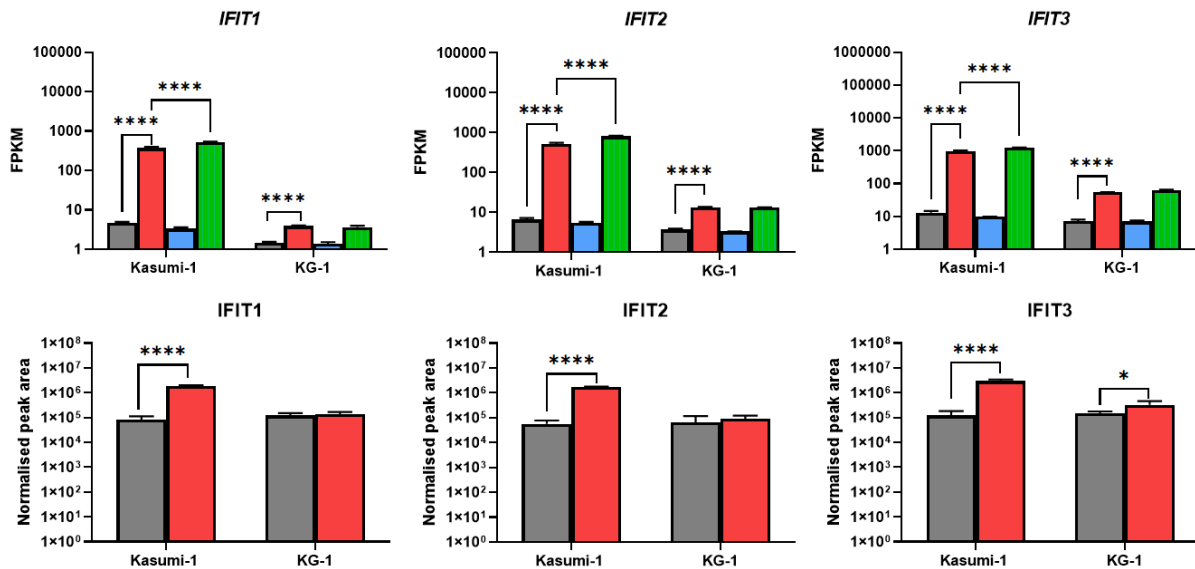
Figure 47: Transcript expression of IFIT1/2/3 transcripts and proteins in Kasumi-1 and KG-1 under all treatment conditions. Data given as FPKM (n=3). Grey = Control, Red = IFNG, Blue = 5AzaC, and Green = IFNG + 5AzaC. Statistical tests by Novogene Wald test – BH correction. B) Normalised peak area expression of IFIT1/2/3 protein in cell lines, control compared to 48 hrs treatment 100 ng/ml IFNG (n=5-6). Holm-Sidak method was used to calculate statistical significance between IFNG treated and untreated cells. * = Padj< 0.05, ** = Padj< 0.01, *** = Padj< 0.001 and **** = Padj< 0.0001.

IFIT2 forms complexes with IFIT1 and IFIT3, the latter of which significantly reduces the ability of IFIT2 to induce apoptosis (Kotredes and Gamero 2013). Expression of *IFIT1*, (Kas-1 ctrl = 4.73 FPKM, IFNG = 371.42 FPKM, KG-1 ctrl = 1.47 FPKM, IFNG = 3.88 FPKM) *IFIT2* (Kas-1 ctrl = 6.53 FPKM, IFNG = 520.63 FPKM, KG-1 ctrl = 3.68 FPKM, IFNG = 13.46 FPKM) and *IFIT3* (Kas-1 ctrl = 12.98 FPKM, IFNG = 970.06 FPKM, KG-1 ctrl = 7.25 FPKM, IFNG = 54.62 FPKM) was significantly upregulated in both cell lines with IFNG treatment, but to a lower extent in KG-1 (All, Padj < 0.0001, Wald test- BH). The combination of IFNG and 5AzaC further increased expression of *IFIT1/2/3* in Kasumi-1 compared to IFNG treatment alone (Kas-1 IFNG5AzaC *IFIT1* = 531.83 FPKM, *IFIT2* = 814.07 FPKM, *IFIT3* = 1233.68 FPKM, Padj < 0.0001, Wald test- BH). The SWATH-MS data from chapter 3 was searched for expression of the protein counterparts. IFIT proteins were also highly upregulated in Kasumi-1 (IFIT1 Ctrl =82,613 NPA, IFNG = 1.89×10^6 NPA, IFIT2 Ctrl = 56,064 NPA, IFNG = 1.72×10^6 NPA, IFIT3 Ctrl = 120,500 NPA, IFNG = 3.02×10^6, all, Padj< 0.0001, Holm Sidak),

and IFIT3 was significantly increased in KG-1 (Ctrl = 150,940 NPA, IFNG = 326,124 NPA, Padj<0.05, Holm Sidak) when treated with IFNG. As IFIT3 was also highly significantly upregulated with IFNG in Kasumi-1, there is a chance that some IFIT2 was trapped in IFIT2/3 complexes. Although IFIT3 also binds with IFIT1, it has not been reported if this decreases its ability to induce apoptosis, although it has been found to increase IFIT1 half-life and regulate its antiviral response (Johnson, et al. 2018).

4.2.4.7 *IFNG induced Annexin 1 expression could promote apoptosis of Kasumi-1 via IFIT1*

Annexin A1 (ANXA1) is a protein which binds to phospholipids in a Ca^{2+} dependent manner and has been identified to play roles in cell proliferation, apoptosis and inflammation when expressed by immune cells (Shao, et al. 2019, Sheikh and Solito 2018, Biaoxue, Xiguang and Shuanying 2014). As cell lines are not immune cells their expression of ANXA1 likely helps to resolve inflammation by constitutively activating caspase 3 activity to promote apoptosis of themselves. Additionally, ANXA1 may be translocated to the cell surface of apoptotic cells to act as an 'eat me' signal for removal by macrophages (Sheikh and Solito 2018).
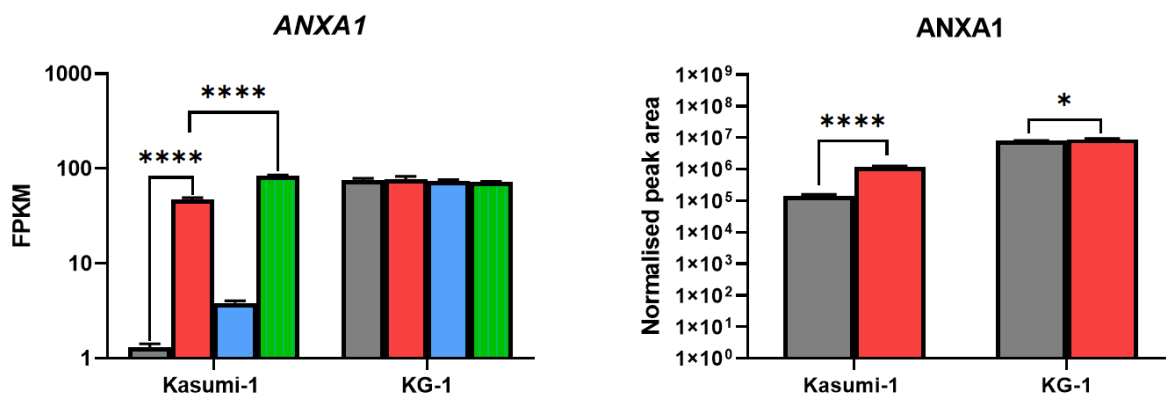
Figure 48: Expression of ANXA1 transcript and protein in Kasumi-1 and KG-1 under stated treatment conditions. Transcript data given as FPKM (n=3) and protein data was given in NPA (n=3). Grey = Control, Red = IFNG, Blue = 5AzaC, and Green = IFNG + 5AzaC. Statistical tests by Novogene Wald test – BH correction. B) Normalised peak area expression of ANXA1 protein in cell lines, control compared to 48 hrs treatment 100 ng/ml IFNG (n=5-6). Grey; Control, Red; IFNG. Holm-Sidak method was used to calculate statistical significance between IFNG treated and untreated cells. * = Padj< 0.05, ** = Padj< 0.01, *** = Padj< 0.001 and **** = Padj< 0.0001.

Transcriptomics and proteomics data showed high baseline expression of *ANXA1* mRNA (Kas-1 Ctrl = 1.30 FPKM, KG-1 Ctrl = 75.82 FPKM) and protein (Kas-1 Ctrl = 137,417 NPA, KG-1 Ctrl = 8.07 x 10^6 NPA) in KG-1 compared to Kasumi-1 (Both Padj < 0.0001, Wald-test BH and Holm Sidak respectively)(Figure 48). This high expression of ANXA1 prior to treatment is unique to KG-1, as SWATH-MS data generated in chapter 3 showed it was lowly expressed in SIG-M5 and THP-1 too (Appendix 7.6). Expression of *ANXA1* mRNA and protein was significantly upregulated by IFNG treatment in Kasumi-1 (mRNA, Kas-1 IFNG = 47.04 FPKM, Padj < 0.0001, Wald-test BH, and, protein, Kas-1 IFNG = 1.16 x 10^6 NPA, Padj< 0.0001, Holm-Sidak) and combining IFNG and 5AzaC

further upregulated *ANXA1* in Kasumi-1 compared to IFNG alone (Kas-1 IFNG5AzaC = 84.20 FPKM, Padj < 0.0001 Wald-test BH). ANXA1 expression has been associated with increased RIG-1 protein (*DDX58*) expression in lung epithelial cells. RIG-1 activation was observed to induce cell death via IFIT1, and reduced ANXA1 expression resulted in decreased IFIT1 expression, thereby reducing cell death (Yap, et al. 2020). IFNG induced upregulation of DDX58 in both cell lines (Kas-1 Ctrl = 1.41 FPKM, IFNG = 113.56 FPKM, KG-1 Ctrl = 3.28 FPKM, IFNG = 7.18 FPKM, Both, Padj < 0.0001, Wald-test BH), combining IFNG with 5AzaC further significantly upregulated DDX58 expression in Kasumi-1 compared to IFNG alone (Figure 54, Kas-1 IFNG 5AzaC = 132.00 FPKM, Padj < 0.0001, Wald-test BH). The *DDX58* protein product RIG-1 was only upregulated in Kasumi-1 in response to IFNG (Kas-1 Ctrl = 20,193 NPA, Kas-1 IFNG = 333,485 NPA, Padj< 0.0001). In Kasumi-1, IFNG induced the expression of ANXA1, IFIT1 and RIG-1 transcripts and proteins, and subsequently could initiate apoptosis via this route.
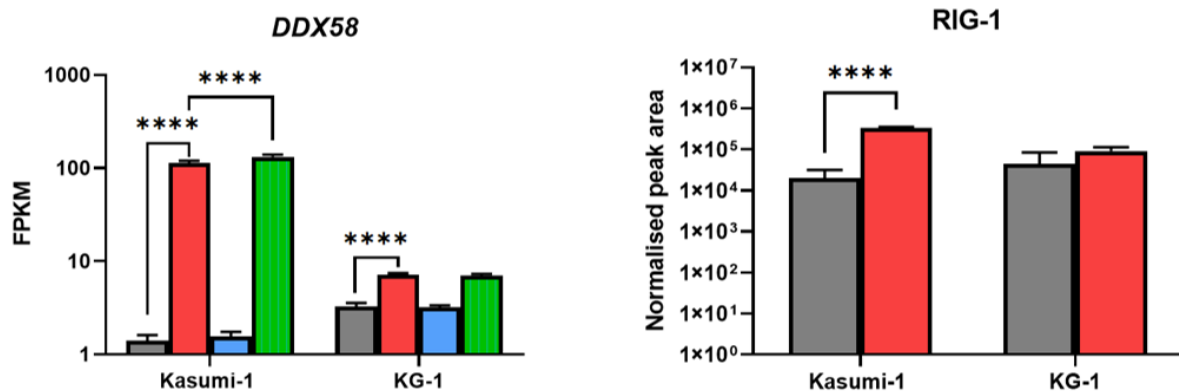
*Figure 49: Transcript expression of DDX58 transcript and protein in Kasumi-1 and KG-1 under all treatment conditions. Transcript data given as FPKM (n=3) and protein data as NPA (n=5/6). Grey = Control, Red = IFNG, Blue = 5AzaC, and Green = IFNG + 5AzaC. Statistical tests by Novogene Wald test – BH correction. B) Normalised peak area expression of RIG-1 protein in cell lines, control compared to 48 hrs treatment 100 ng/ml IFNG (n=5-6). Grey; Control, Red; IFNG. Holm-Sidak method was used to calculate statistical significance between IFNG treated and untreated cells. * = Padj< 0.05, ** = Padj< 0.01, *** = Padj< 0.001 and **** = Padj< 0.0001.*

4.2.5 IFNG induction of Immune Evasion Mechanisms in AML cell lines

The following section explores the expression of immune-related molecules and their potential implications in IFNG mediated immunosuppression and immune evasion within AML cell lines. Here, expression of known immune resistance molecules such as *CD274* (PD-L1) and *HLA-E* were investigated in the transcriptomics data generated in this chapter, and proteomics data generated in chapter 3. Other immune evasion molecules, including *TNFSF10* (TRAIL), *NCR3*, *LGALS9*, *HAVCR2*, *PVR*, and *SIRPα*, were also explored to understand their impact on immune response regulation. Kasumi-1 was vulnerable to IFNG induced cell death through several potential mechanisms. To survive immune response-mediated cell death, cells must employ adaptive resistance mechanisms to evade it. Further investigation focused on transcripts associated with immune resistance, comparing their expression across different cell lines.

4.2.5.1 AIR molecules CD274 and HLA-E reported as IFNG inducible by transcriptomics data

Adaptive immune resistance, discussed in Chapter 3, involves the expression of immune escape and resistance molecules. Examination of the transcriptomics data revealed that IFNG treatment of Kasumi-1 activated the transcription of certain adaptive immune resistance transcripts and unveiled additional mechanisms of immune resistance. Out of the transcripts discussed in chapter 3 for adaptive immune resistance, only *CD274* (PD-L1) and *HLA-E* were detected above 1 FPKM in both cell lines (*CD274* Kas-1 Ctrl = 0.01 FPKM, KG-1 Ctrl = 0.10 FPKM, *HLA-E* Kas-1 Ctrl = 90.63 FPKM, KG-1 Ctrl = 131.32 FPKM) (Figure 50). Both *CD274* (Kas-1 IFNG = 7.42 FPKM, KG-1 IFNG = 1.67 FPKM) and *HLA-E* (Kas-1 IFNG = 821.15 FPKM, KG-1 IFNG = 468.07 FPKM) were significantly upregulated with IFNG (both, Padj < 0.0001, Wald test- BH). In Kasumi-1 combination of IFNG with 5AzaC further upregulated *HLA-E* expression in comparison to IFNG only (Kas-1 IFNG5AzaC = 936.71 FPKM, Padj < 0.0001, Wald-test BH). Transcript expression levels of both targets agreed with the protein expression trends detected by flow cytometry in Figure 16.
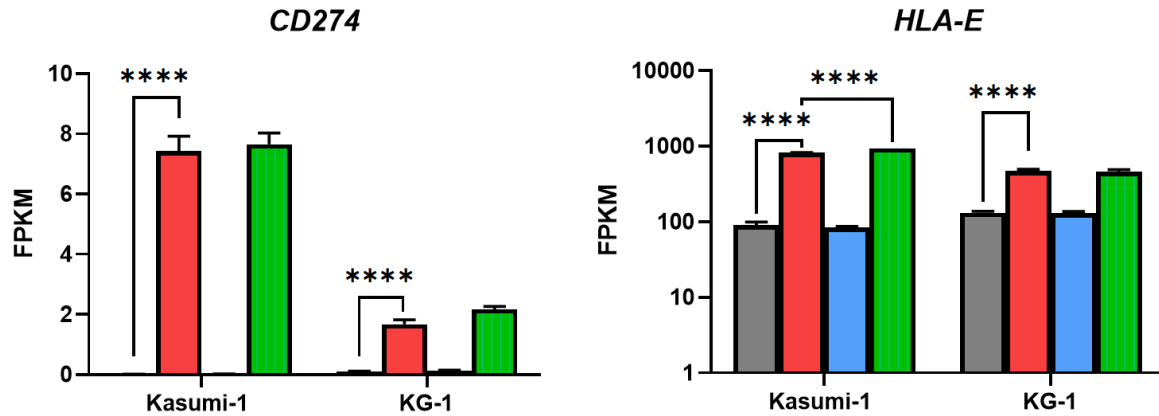
Figure 50: Expression of CD274 and HLA-E transcripts in Kasumi-1 and KG-1 under all treatment conditions. Data given as FPKM (n=3). Grey = Control, Red = IFNG, Blue = 5AzaC, and Green = IFNG + 5AzaC. Statistical tests by Novogene Wald test – BH correction, * = Padj< 0.05, ** = Padj< 0.01, *** = Padj< 0.001 and **** = Padj< 0.0001.

4.2.5.2 The immunosuppressive molecule TRAIL (TNFSF10) was significantly upregulated in Kasumi-1 cells following IFNG treatment

TNFSF10 (TRAIL) binds to death receptors DR4 and DR5 (*TNFRSF10A/TNFRSF10B)*, which have death domains, and induce apoptosis via caspases. TRAIL is expressed by NK cells and cytotoxic T cells of the immune system. As TRAIL selectively induces apoptosis in tumour cells over normal cells, recombinant TRAIL is a favourable anti-tumour treatment (Beyer, et al. 2019). *TNFSF10* was significantly highly upregulated by IFNG in Kasumi-1 when treated with IFNG (Kas-1 Ctrl = 0.70 FPKM, IFNG = 138.95 FPKM, Padj < 0.0001, Wald test- BH), and even further when treated with both IFNG and 5AzaC (Kas-1 IFNG5AzaC = 191.18 FPKM, Padj < 0.0001, Wald test- BH) (Figure 51). In comparison, KG-1 baseline expression of *TNFSF10* was higher, but IFNG only slightly upregulated expression (KG-1 Ctrl = 19.02 FPKM, KG-1 IFNG = 26.99 FPKM, Padj < 0.05), with no further increase seen when IFNG was combined with 5AzaC (KG-1 IFNG5AzaC = 26.80 FPKM). The upregulation of *TNFSF10* could contribute to Kasumi-1 immunosuppression, to induce apoptosis of NK and CD8 T cells which express TRAIL receptors (Mirandola, et al. 2004).

Receptors for TNFSF10 were consequently looked for in the transcriptomics data. *TNFRSF10A* was expressed at low levels in both cell lines but upregulated by IFNG in Kasumi-1 (Kas-1 Ctrl = 3.36 FPKM, IFNG = 4.47 FPKM, Padj< 0.01). *TNFRSF10B* was downregulated in response to IFNG and upregulated by 5AzaC in Kasumi-1 only (Kas-1 ctrl = 7.53 FPKM, IFNG = 6.47 FPKM and 5AzaC = 9.99 FPKM, Padj < 0.05 and Padj <0.0001 respectively, Wald-test BH). Some cancers upregulate 'decoy' receptors for TRAIL, which do not contain death domains such as *TNFRSF10C, TNFRSF10D* and *TNFRSF11B* (Beyer, et al. 2019). Out of the decoy receptors, only *TNFRSF10D* was expressed above 1FPKM in the cell lines and was upregulated by IFNG in Kasumi-1 only (Ctrl = 5.61 FPKM, IFNG = 7.97 FPKM, Padj < 0.0001, Wald-test BH). Increased expression of *TNFSF10*

could be an adaptive mechanism to attack Th1 cells and promote Tregs, thereby creating an immunosuppressive environment, and enabling tumour growth.
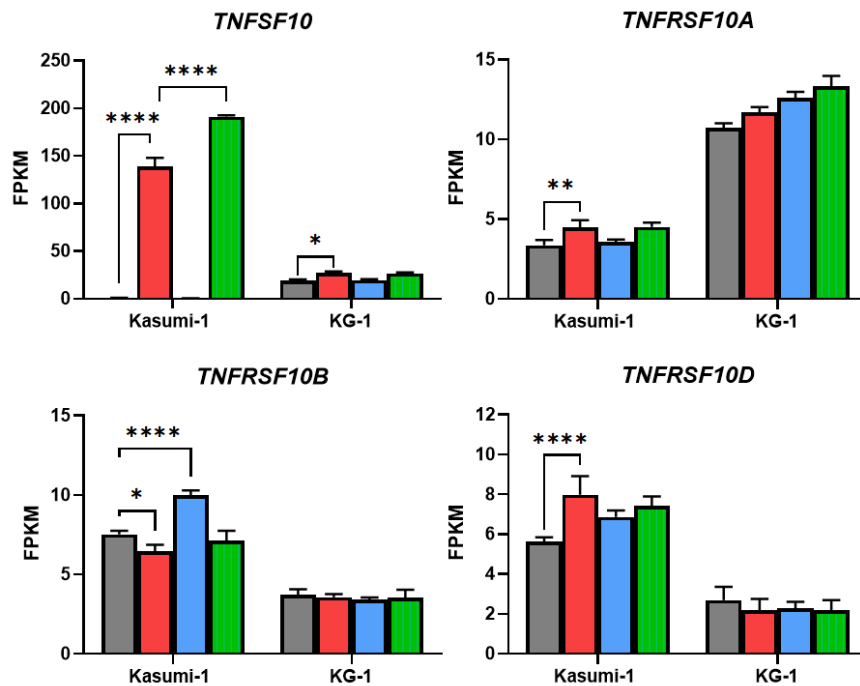
*Figure 51: Expression of TNFSF10, TNFRSF10A/B and TNFRSF10D in Kasumi-1 and KG-1 cell lines under all treatment conditions. Data given in FPKM (n=3). Grey = Control, Red = IFNG, Blue = 5AzaC, and Green = IFNG + 5AzaC. Statistical tests by Novogene Wald test – BH correction, * = Padj< 0.05, ** = Padj< 0.01, *** = Padj< 0.001 and **** = Padj< 0.0001.*

4.2.5.3   IFNG induced transcription of LGALS9 and HAVCR2 which can impair NK cell and T-cell mediated lysis, while downregulating NCR3LG1 which can prevent recognition by NK cells

NCR3 is a 'natural cytotoxicity receptor' expressed on NK cells, which, when activated in the absence of an inhibitory signal, induces tumour cell lysis. Tumour cells express NCR3LG1, enabling their recognition and destruction by NK cells (Brandt, et al. 2009). Both cell lines expressed low levels of NCR3LG1 (Kasumi-1 Ctrl = 7.07 FPKM, KG-1 Ctrl = 1.65 FPKM). IFNG treatment significantly decreased NCR3LG1 expression in Kasumi-1 (Kasumi-1 IFNG = 3.73 FPKM, Padj < 0.0001, Wald test- BH), and this reduction was further enhanced by treatment with a combination of IFNG and 5AzaC compared to IFNG alone (Kasumi-1 IFNG+5AzaC = 2.74 FPKM, Padj < 0.0001, Wald test- BH)(Figure 52). Low expression may allow cell lines to escape this mechanism of immune mediated destruction.
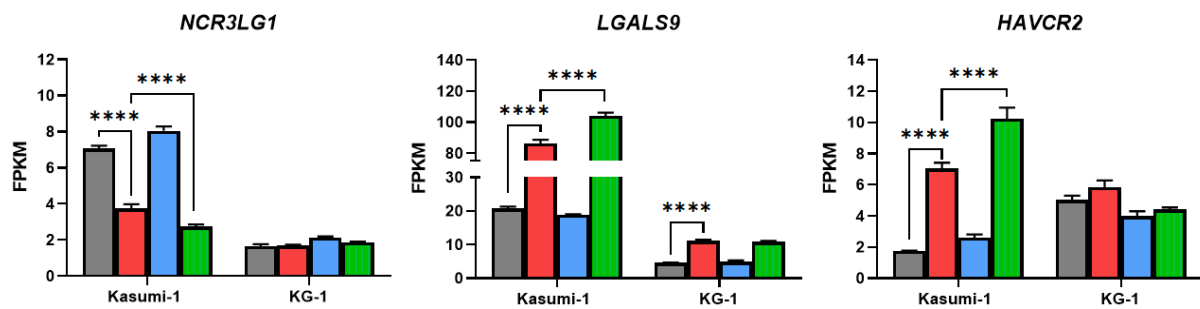
Figure 52: Expression NCR3LG1, LGALS9 and HAVCR2 (TIM-3) in Kasumi-2 and KG-1 cell lines under all treatment conditions. Data given as FPKM (n=3). Grey = Control, Red = IFNG, Blue = 5AzaC, and Green = IFNG + 5AzaC. Statistical tests by Novogene Wald test – BH correction, * = Padj< 0.05, ** = Padj< 0.01, *** = Padj< 0.001 and **** = Padj< 0.0001.

Galectin-9 (*LGALS9*) is known for its role in developing the acquired immune system through negative selection of T cells in the thymus (Wada and Kanwar 1997). Other studies have suggested its role as an anti-cancer agent in Myeloma and Chronic Myelogenous Leukaemia (Kobayashi, et al. 2010, Kuroda, Junya, et al. 2010). More recently, a study of using primary AML patient samples found high levels of Galectin-9 secretion impaired NK cells ability to kill AML. In addition, soluble Tim-3 (*HAVCR2*) decreased secretion of IL-2 by T cells, preventing activation of NK and cytotoxic T-cells (Silva, et al. 2017). *HAVCR2* expression was significantly upregulated in Kasumi-1 in response to IFNG and IFNG combined with 5AzaC (Kas-1 Ctrl = 1.75 FPKM, IFNG = 7.03 FPKM, IFNG5AzaC = 10.23 FPKM, both, Padj < 0.0001, Wald-test BH). Transcriptomics data showed IFNG significantly upregulated *LGALS9* (Kas-1 Ctrl = 20.68 FPKM, IFNG = 86.27 FPKM, KG-1 Ctrl = 4.58 FPKM, IFNG = 11.15 FPKM, both, Padj < 0.0001, Wald test- BH) in both cell lines. The combination of IFNG with 5AzaC pushed *LGALS9* expression higher than IFNG alone in Kasumi-1 (Kas-1 IFNG5AzaC = 103.97 FPKM, Padj<0.0001, Wald-test BH). Protein expression levels of Galectin-9 detected by SWATH-MS reflected the increases seen at the transcript level (see 7.6). It is possible both cell lines could use IFNG to upregulate galectin-9 to exploit this mechanism to avoid immunosurveillance. However, SWATH MS detects peptides available in a lysed sample, and so there is no data on how much of the galectin-9 protein was secreted.

4.2.5.4 Kasumi-1 may use PVR and Nectin-2 suppress immune cells through binding receptor TIGIT

PVR and Nectin-2 are molecules that regulate immune cell functions and are expressed by tumour and myeloid cells. Their counterparts, DNAM-1 and TIGIT, are expressed on immune effector cells. Binding of PVR or Nectin-2 to DNAM-1 induces NK and cytotoxic CD8[+] T cells to promote anti-tumoral action, whereas binding to TIGIT causes an immunosuppressive response (Gorvel and Olive 2020).
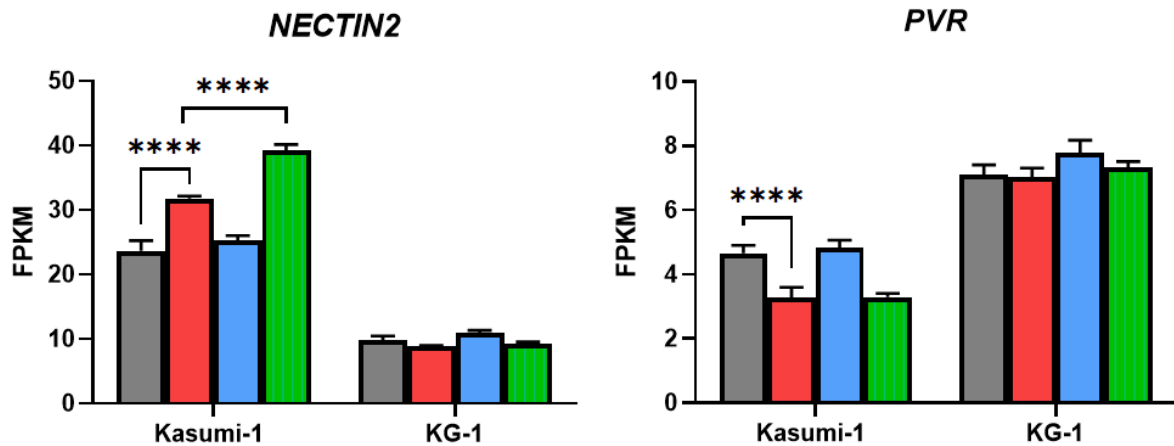
Figure 53: Expression of NECTIN2 and PVR in Kasumi-1 and KG-1 cell lines under all treatment conditions. Data given as FPKM (n=3). Grey = Control, Red = IFNG, Blue = 5AzaC, and Green = IFNG + 5AzaC. Statistical tests by Novogene Wald test – BH correction, * = Padj< 0.05, ** = Padj< 0.01, *** = Padj< 0.001 and **** = Padj< 0.0001.

Transcriptomic data showed Kasumi-1 expressed more *NECTIN2* than KG-1 (Kas-1 Ctrl = 23.64 FPKM, KG-1 Ctrl = 9.75 FPKM), while KG-1 expressed more *PVR* than Kasumi-1 (Kas-1 Ctrl = 4.65 FPKM, KG-1 Ctrl = 7.10 FPKM)(**Error! Reference source not found.**). PVR has higher affinity for the immunosuppressive TIGIT target than NECTIN2, but NECTIN2 can still contribute towards immunosuppression via TIGIT (Gorvel and Olive 2020). *NECTIN2* was upregulated with IFNG in Kasumi-1 (Kas-1 Ctrl = 23.64 FPKM, IFNG = 31.77 FPKM, Padj< 0.0001, Wald-test BH) and further upregulated when treated with IFNG and 5AzaC than IFNG alone (Kas-1 IFNG5AzaC = 39.23 FPKM, Padj < 0.0001, Wald-test BH), while PVR was downregulated with IFNG treatment in Kasumi-1 (Kas-1 Ctrl = 4.65 FPKM, IFNG = 3.28 FPKM, Padj < 0.0001). Treatments did not affect expression of either target in KG-1. *NECTIN2* was upregulated upon treatment with interferon gamma in Kasumi-1, and this may offer a potential checkpoint target for investigation in interferon dominant AML phenotypes.

4.2.5.5   Cell lines may use CD47 and SIRPα complexes to avoid phagocytosis by macrophages
CD47 is a membrane receptor highly expressed on tumours, forming a complex with signal-regulatory protein α (SIRPA). This signalling complex acts as a 'don't eat me' signal, enabling tumours to escape phagocytosis by macrophages (Zhang, Wenting, et al. 2020). *CD47* expression was higher in KG-1 than Kasumi-1 (KG-1 Ctrl = 26.98 FPKM, Kas-1 Ctrl = 5.85 FPKM), as was the protein counterpart (KG-1 Ctrl = 294,812 NPA, Kas-1 Ctrl = 73,069 NPA). IFNG induced CD47 in both cell lines at mRNA level (KG-1 IFNG = 30.95 FPKM, Kas-1 IFNG = 17.17 FPKM, Padj < 0.0001, Wald-test BH), and protein level (KG-1 IFNG = 419,110 NPA, Kas-1 IFNG = 141,881 NPA, Padj < 0.001, Padj < 0.0001, respectively, Holm-Sidak). *SIRPA* was only inducible in Kasumi-1 (Kas1 Ctrl = 14.87 FPKM, Kas-1 IFNG = 18.56 FPKM, Padj < 0.0001, Wald-test BH, Figure 54). While SIRPA

was lowly expressed at the transcript level in KG-1 (KG-1 Ctrl = 1.63 FPKM), its protein expression was comparable to Kasumi-1. Unfortunately, SWATH-MS data for SIRPA showed large variation, making the data less reliable. Both cell lines expressed CD47 and SIRPA and therefore might utilise the complex for phagocyte evasion in AML cell lines.
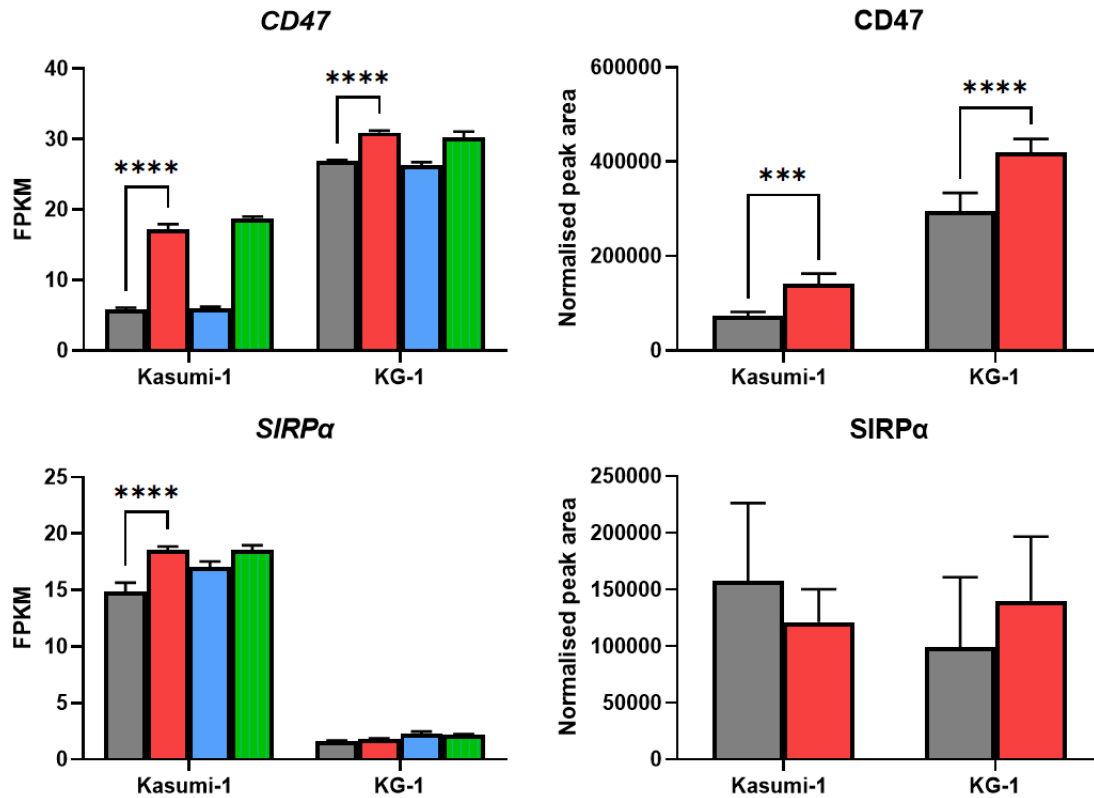
Figure 54: Transcript expression of CD47 and SIRPA in Kasumi-1 and KG-1 under all treatment conditions, and corresponding protein expression post IFNG. Transcript data given as FPKM (n=3), protein data given as NPA (n=5-6). Grey = Control, Red = IFNG, Blue = 5AzaC, and Green = IFNG + 5AzaC. Statistical tests by Novogene Wald test – BH correction. B) Normalised peak area expression of CD47 and SIRPA protein in cell lines, control compared to 48 hrs treatment 100 ng/ml IFNG (n=5-6). Grey; Control, Red; IFNG. Holm-Sidak method was used to calculate statistical significance between IFNG treated and untreated cells. * = Padj< 0.05, ** = Padj< 0.01, *** = Padj< 0.001 and **** = Padj< 0.0001.

4.2.5.6   Cell lines did not express stemness transcripts other than NOTCH1 > 1 FPKM

Another factor in relapse is stemness of cells and their capacity for self-renewal. Stemness transcripts *NANOG, OCT4, REX1, NOTCH1* and, *NESTIN* were investigated in the transcriptomics data set (Gonzalez-Garza, et al. 2018). Except for *NOTCH1,* the transcripts were expressed very lowly and reported at below 0.1 FPKM. KG-1 expressed *NOTCH1* to higher levels than Kasumi-1 (KG-1 Ctrl = 20.91 FPKM, Kas-1 Ctrl = 8.61 FPKM) and expression was decreased by IFNG significantly in Kasumi-1 but not KG-1 (Kas-1 IFNG = 6.66 FPKM, Padj<0.0001, Wald-test BH, Figure 55). While 5AzaC treatment decreased *NOTCH1 expression* in both cell lines (Kas-1 5AzaC = 6.63 FPKM, KG-1 IFNG = 16.10 FPKM, both, Padj < 0.0001, Wald-test BH). Due to the lack of expression, the stemness transcripts were not useful in characterising the cell lines behaviours.
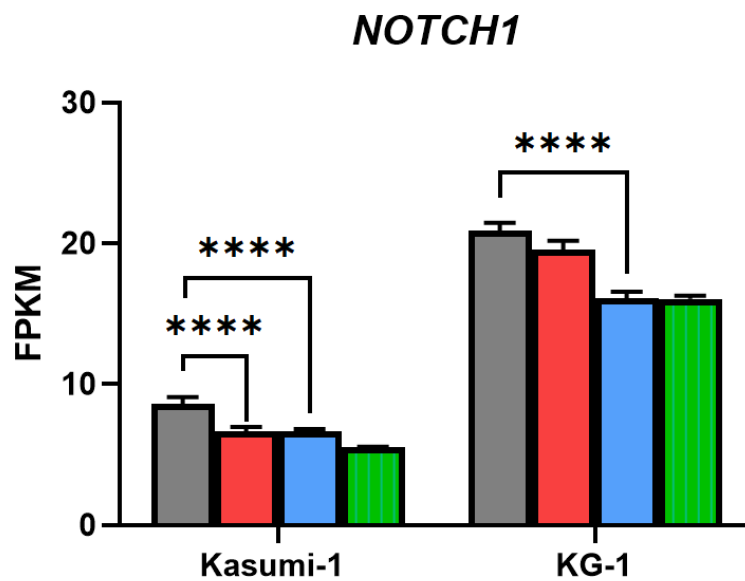
*Figure 55: Expression of NOTCH1 in Kasumi-1 and KG-1 cell lines under all treatment conditions. Data given as FPKM (n=3). Grey = Control, Red = IFNG, Blue = 5AzaC, and Green = IFNG + 5AzaC. Statistical tests by Novogene Wald test – BH correction. * = Padj< 0.05, ** = Padj< 0.01, *** = Padj< 0.001 and **** = Padj< 0.0001.*

4.2.6 5AzaC induced three times more differentially expressed transcripts in Kasumi-1 than KG-1

To investigate the effects of demethylation on each cell line, lists of significantly differentially expressed transcripts were composed and submitted to metascape. The breadth of altered expression seen was smaller than when cells were treated with IFNG. Again Kasumi-1 was more receptive to changes by 5AzaC than KG-1, with changes to 769 transcripts compared to 219 differentially expressed transcripts in KG-1. Data was transformed to log2(FPKM+1) and Z-values calculated before complete hierarchal clustering. Data summarised in heat map (Figure 56).
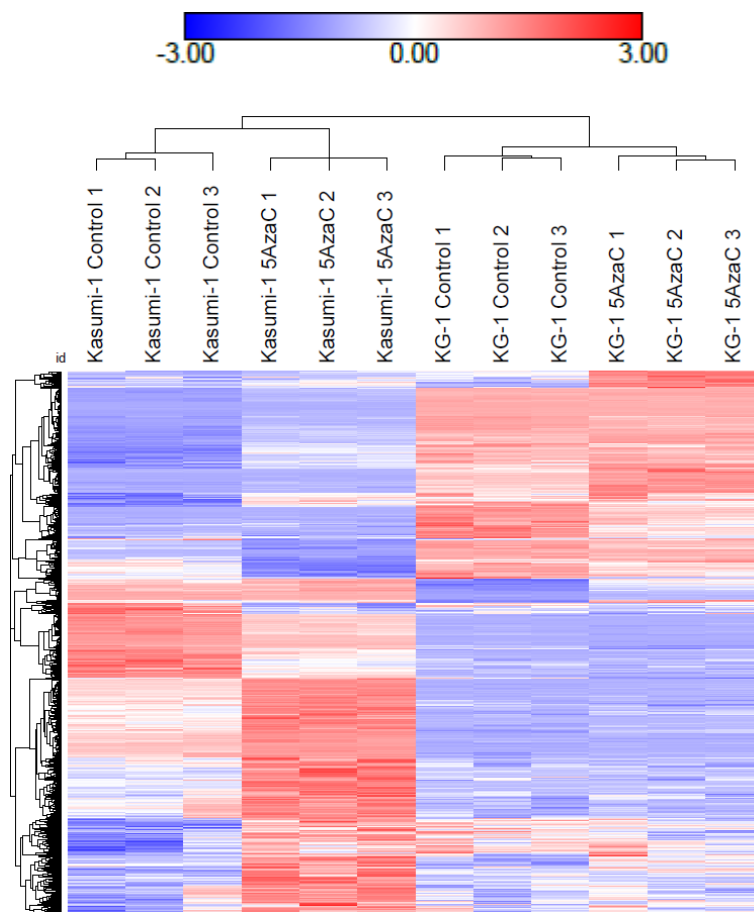
Figure 56: Hierarchical clustering of the significantly differentially expressed transcripts between untreated and 5AzaC-treated Kasumi-1 and KG-1 cell lines. Hierarchical clustering performed using Euclidean distance and complete linkage (n = 3 per sample type). Key for heat maps shown top left for expression values. Heat maps generated in Morpheus (https://software.broadinstitute.org/morpheus/). Blue indicates lower expression; red indicates higher expression.

Transcript lists for each cell line were filtered for significance Padj < 0.05 and a minimum 1.5-fold change, then separated into two lists by fold change direction. Venn diagrams were generated to compare transcript expression between cell lines, Figure 57A and B, respectively.
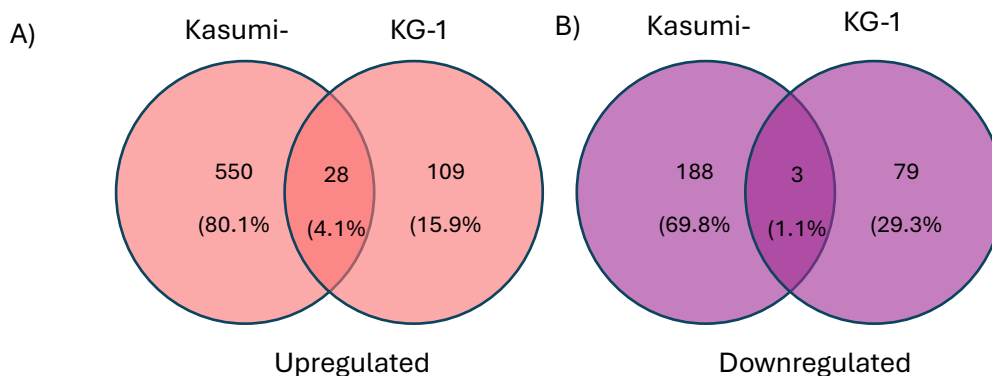
Figure 57: Venn diagrams comparing transcript lists of differentially expressed transcripts of Kasumi-1 and KG-1 treated with 5AzaC. Differential expression was defined as a change of 1.5-fold or more that was also statistically significant (Padj < 0.05). Comparison of up-regulated transcripts depicted in A, and down regulated transcripts displayed in B.

Transcript lists were uploaded to Metascape for pathway enrichment, the top 10 significantly enriched pathways are presented in Figure 58, respectively with more details available in the appendix (See 7.3).

4.2.6.1 Demethylation of Kasumi-1 reported pathways which promote regulatory T-cell environment as enriched

Demethylation of Kasumi-1 with 5AzaC promotes IFN type I signalling, which in turn may induce signalling through the ERK1/2 cascade and induce cell motility (Tanimura and Takeda 2017). Flagged pathways included leukocyte and lymphocyte differentiation, which could be ignored as the model consisted solely of AML cells. Myeloid leukocyte activation and macrophage activation pathways were upregulated. Changes to cell structure were reflected in increased transcripts for extracellular matrix reorganisation and supramolecular fibre organisation. While cell locomotion increased, one of the top decreased pathways was cellular extravasation, preventing exit through vessel walls. There was also a decrease in synthesis and metabolic pathways, as well as pathways involved in the regulation of cell development and morphogenesis involved in differentiation. In summary, treatment of Kasumi-1 cells with 5AzaC promoted pathways involved in myeloid differentiation and cell mobility. Example transcripts driving differentiation reported as part of leukocyte differentiation pathways (totalling 46) included *ANXA1* and *CCR2*. However, none of these targets reported significant changes in expression after 5AzaC treatment in the transcriptomics data. *CCR2, DUSP10* and *ZAP70* transcripts reported <1FPKM with and without 5AzaC treatment.
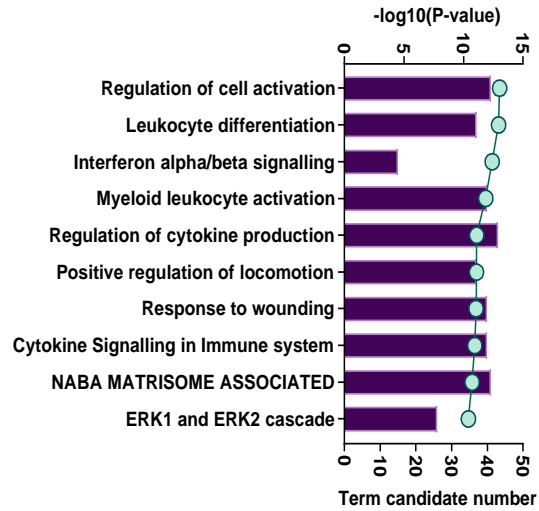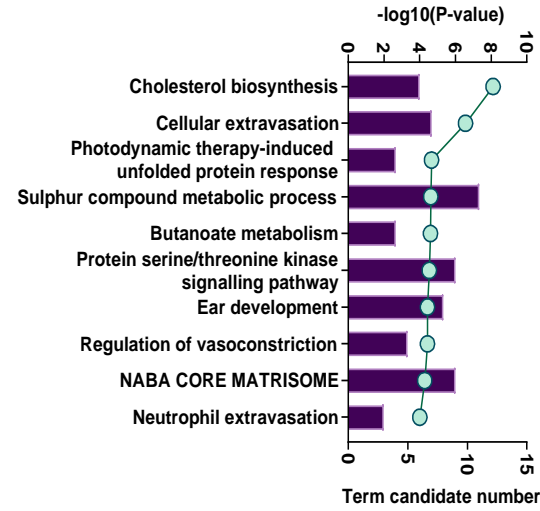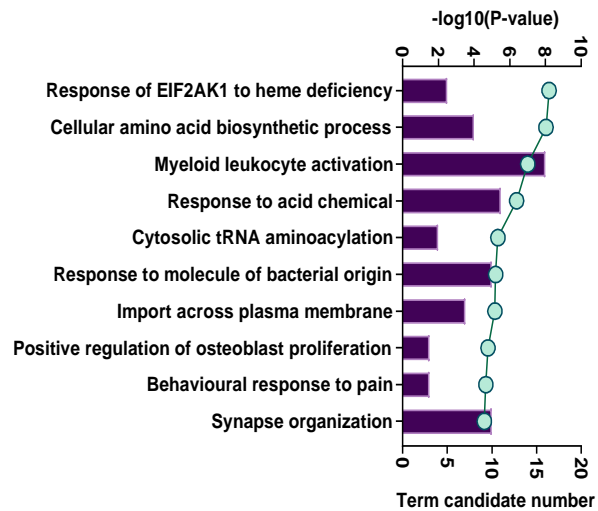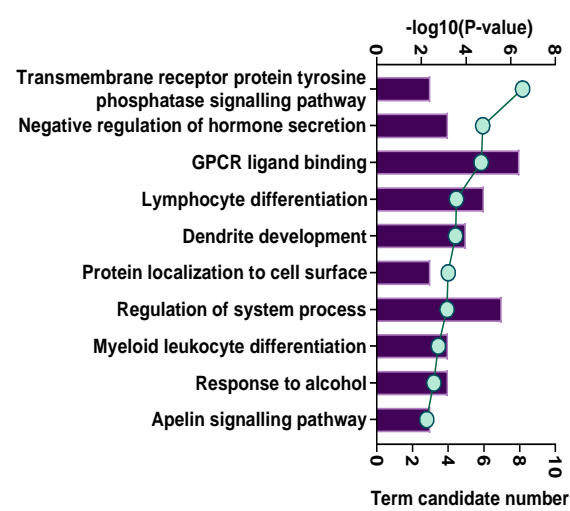
Figure 58: The top 10 most significantly enriched pathways in Kasumi-1 and KG-1 treated with 5AzaC. Enriched pathways were identified by the 'unique' up and down regulated transcript lists and sorted by significance post FDR. See appendix for table summary process id's and number of transcripts significantly changed in each pathway

Treatment of KG-1 with 5AzaC enriched for amino acid biosynthesis and aminoacylation processes, as well as amino acid import across plasma membranes. The most upregulated pathway was response of EIF2AK1 (HRI) to heme deficiency; a pathway related to erythrocyte generation. As KG-1 cells are macrophage derived cells and not erythroid precursors, this pathway is not able to generate erythrocytes. This pathway has however been shown to effect maturation and some functions of macrophages in mice (Liu, S., et al. 2007). KG-1 shared upregulation of the pathway 'myeloid leukocyte activation' in common with Kasumi-1. The process of lymphocyte differentiation was downregulated in KG-1, unlike in Kasumi-1, where it was upregulated with the same treatment. Additionally, a top upregulated pathway for KG-1 was the extrinsic apoptotic signalling pathway. However, this was contested by the significant downregulation of the 'positive regulation of cell death' pathway. As seen in the IFNG data, enrichment in KG-1 was weaker, as fewer transcripts passed input specifications compared to Kasumi-1 (Fold change > 1.5, Padj < 0.05). This led to some pathways being reported as 'enriched' despite very few transcripts in the pathway being altered. For example, in the 'positive regulation of cell death' pathway, only 6 out of 725 transcripts were altered, yet it was still reported as significant, despite changes to less than 1% of the pathway, which could be coincidental and inconsequential.

4.2.7 DNA methylation plays a key role in haematopoiesis

DNA methylation is an epigenetic modification that influences gene expression and is crucial for haematopoiesis, the process of blood cell development. Disturbances in methylation modifiers like DNMTs or TETs can disrupt haematopoiesis and contribute to haematological malignancies such as AML. Abnormal DNA methylation, including mutations in DNMT3A and TET2, is a hallmark of AML. It is therefore unsurprising AML cells treated with a demethylation agent report enrichment for leukocyte differentiation. Kasumi-1 was more responsive to changes inflicted by 5AzaC, as demonstrated in Figure 59 which presents a heat map of significantly altered transcripts of leukocyte differentiation (Genes taken from GO:0002521).
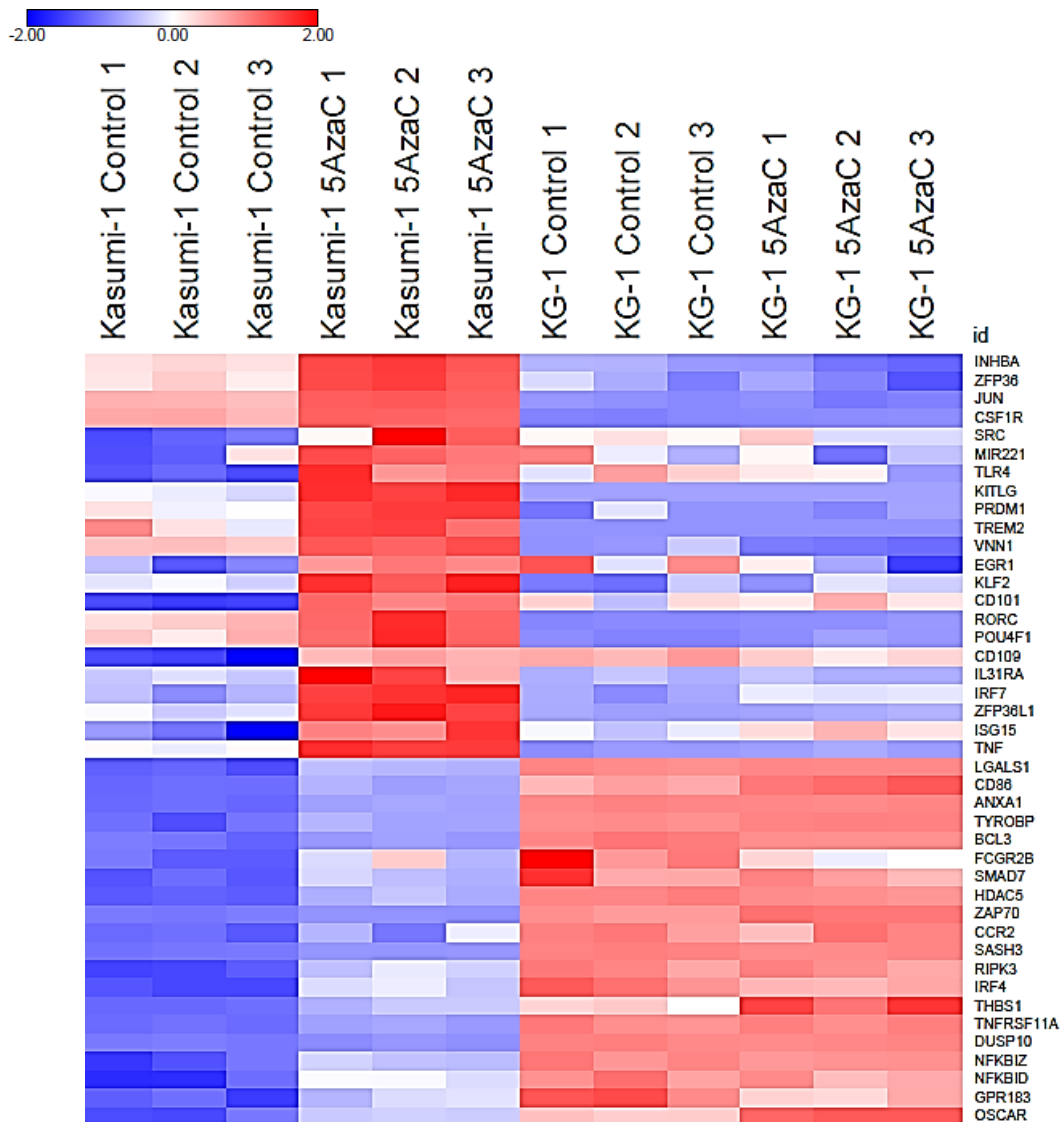
Figure 59: Heat map of differentially expressed transcripts related to leukocyte differentiation between untreated and 5AzaC Kasumi-1 and KG-1. Hierarchal clustering performed using Euclidean distance and complete linkage (n = 3 per sample type). Key for heat maps shown top left for expression values. Heat maps generated in Morpheus (https://software.broadinstitute.org/morpheus/). Blue indicates lower expression; red indicates higher expression.

The roles DNMT1, DNMT3A/B and TET2 play in methylation of DNA are described and illustrated by diagram in chapter 3, Figure 12. Resistance to 5AzaC has been associated with higher DNMT1 expression in AML cell lines (Wong, Lawrie and Green 2019). To summarise, 5AzaC is a drug that is analogous to the nucleoside cytidine, it can be administered as a drug to inhibit DNMT1. DNMT1 is a 'maintenance' methylation enzyme, as it tends to methylation of hemi-methylated sites, to restore methylation following DNA replication. 5AzaC does not target DNMT3A/B which are *de* novo DNMTs that oversee methylation of unmethylated DNA. While DNMT1 is inhibited by 5AzaC, DNMT3A/B are unaffected, and continue to catalyse methylation of new sites. The western blot shown in chapter 3 (Figure 13) showed KG-1 had higher DNMT1 expression than Kasumi-1. However, treatment with 5AzaC also successfully depleted DNMT1 expression in both cell lines.

4.2.7.2   Differences in response to 5AzaC may be explained by cell line mutations

DNMT3A mutations are very common in AML and found in between 22-33% of AML cases. They are associated with worse survival outcome (Wong, Lawrie and Green 2019). *DNMT3A* mutation consequences are not well documented, but the most well studied mutation is the R882 missense mutation, which causes loss of methylation activity (Bera, et al. 2020). DNMT3B function is redundant of DNMT3A and higher expression of *DNMT3B* is associated with lower event free survival and overall survival (Wong, Lawrie and Green 2019). High *DNMT3B* expression has also been associated with higher genome wide methylation in paediatric AML (Lamba, et al. 2018). Interestingly, in older patients *DNMT3B* expression has not been associated with methylation (Russler-Germain, et al. 2014). TET2 is also frequently mutated, with 17% of AML patients harbouring loss of function mutations, which leads to hypermethylation (Yang, X., Wong and Ng 2019a). Expression of enzymes involved in DNA methylation was assessed using the transcriptomics data. Corresponding proteins were searched for in the SWATH-MS data set but were not found. Expression of demethylation enzyme transcripts is shown in Figure 60.

At base line, expression of *DNMT1 (Kas-1 Ctrl* = 40.16 FPKM, *KG-1 Ctrl* = 41.98 FPKM) was similar between cell lines (Figure 60). *DNMT3A* was expressed more highly in Kasumi-1 than KG-1 (Kas-1 Ctrl = 16.53 FPKM, KG-1 Ctrl = 1.64 FPKM), and *DNMT3B* was expressed very highly in KG-1 (Kas-1 Ctrl = 5.86 FPKM, KG-1 Ctrl = 43.23 FPKM). *TET2* expression was lower in Kasumi-1 (Kas-1 Ctrl = 3.39 FPKM, KG-1 Ctrl = 6.28 FPKM) but significantly induced by IFNG in Kasumi-1 (Kas-1 IFNG = 6.59 FPKM, Padj < 0.0001, Wald-test BH). IFNG caused significant downregulation of all the *DNMT* transcripts in Kasumi-1 (*DNMT1*= 31.59 FPKM, *DNMT3A* = 11.64 FPKM, *DNMT3B* = 3.73 FPKM, all, Padj < 0.0001, Padj < 0.0001 and Padj < 0.01 respectively, Wald-test BH), while 5AzaC treatment

only significantly downregulated *DNMT3B* in KG-1 (KG-1 5AzaC = 36.06 FPKM, Padj < 0.0001, Wald-test BH) which was further decreased when treated with IFNG and 5AzaC compared to IFNG alone (KG-1 IFNG5AzaC = 30.61 FPKM, Padj < 0.01, Wald-test BH). Both cell lines expressed DNMT1, which is involved in methylation maintenance. Each cell line also expressed one of the de novo methylation DNMTs (DNMT3A or DNMT3B) above 10 FPKM, and both had similar expression levels of TET2. In Kasumi-1 IFNG reduced *DNMT1, DNMT3A* and *DNMT3B* expression, while increasing *TET2* expression. If this translated to protein expression levels, IFNG could cause a decrease in active methylation and maintenance of sites by decreasing expression of DNMT's, while increasing active demethylation through TET2. However, this is assuming that enzymes are all wild type. As DNMT mutation is common in AML patients, the mutational status of DNA methylation mediators was searched for in the cell lines project COSMIC database (Tate, et al. 2019).
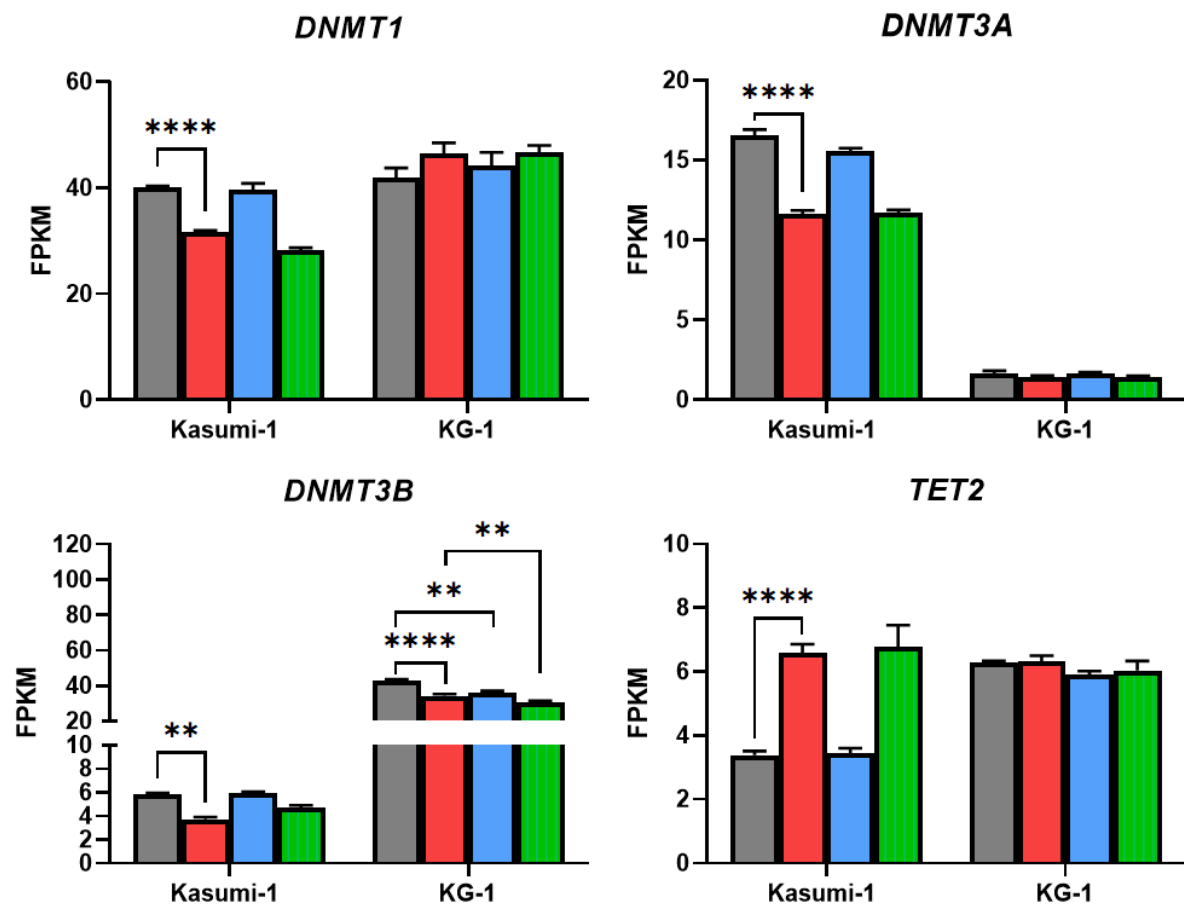
*Figure 60: Expression of DNMT1/3A/3B and TET2 in Kasumi-1 and KG-1 under all treatment conditions. Data given as FPKM (n=3). Grey = Control, Red = IFNG, Blue = 5AzaC, and Green = IFNG + 5AzaC. Statistical tests by Novogene Wald test – BH correction. * = Padj< 0.05, ** = Padj< 0.01, *** = Padj< 0.001 and **** = Padj< 0.0001.*

The COSMIC database reported Kasumi-1 harboured heterozygous mutations of *DNMT3A/B* at multiple codons, but not the well-studied loss of function R882 missense mutation (Russler-Germain, et al. 2014). The type of mutation was listed as unknown for all instances. The database also reported substitution mutations in *TET1* and *TET2* in the KG-1 cell line. Mutation of *TET*2 also usually leads to loss of function (Rasmussen, et al. 2015). As Kasumi-1 has mutated DNMT3A/B, de novo methylation of DNA might be inhibited, and methylation status could be low and maintained through DNMT1. Therefore, treatment with 5AzaC may be effective in Kasumi-1, as it targets DNMT1 preventing methylation maintenance, and methylated sites are lost either actively through TET proteins or passively by cell division. In KG-1, DNMT3A/B are not mutated and may still carry out de novo methylation. In addition, TET1 and TET2, which would normally catalyse active demethylation are mutated, which could prevent their function. In combination, these factors could mean KG-1 is in a hypermethylated state compared to Kasumi-1. Fewer changes were seen when KG-1 was treated with 5AzaC, and this could be because DNMT1 is not the main source of methylation mediation as could be in Kasumi-1.

4.2.7.3 Combination treatment further increases IFNG signalling in Kasumi-1

In section 4.2.5, it was observed that the combination of IFNG and 5AzaC induced more significant changes in immunosuppressive transcripts compared to IFNG or 5AzaC alone. Transcripts such as *HLA-E, TNFRSF10B,* and *HAVCR2* were further upregulated (See Figure 50, Figure 51 and Figure 52, respectively) while *NC3LG1* (Figure 52) was further downregulated by the combination treatment in Kasumi-1. This synergistic effect was unique to Kasumi-1 and was also seen in the expression of IFNG-inducible transcripts such as *IFIT1/2/3* (Figure 47) as well as *ANXA1* (Figure 48) and *DDX58 (*Figure 49*)*.

To study how combining demethylation with IFNG signalling changed cell lines response, transcriptomes were compared between control and combination treated cells. Lists of transcripts significantly differentially expressed between control cell lines and treated with IFNG alone and in combination with 5AzaC were generated and submitted to metascape. Then a comparison between untreated cells and combination treated cells was performed, to see how pathway enrichment correlated to IFNG treated enrichment. Data was transformed to log2(FPKM+1) and Z-values calculated before complete hierarchal clustering summarised by heat map in Figure 61.
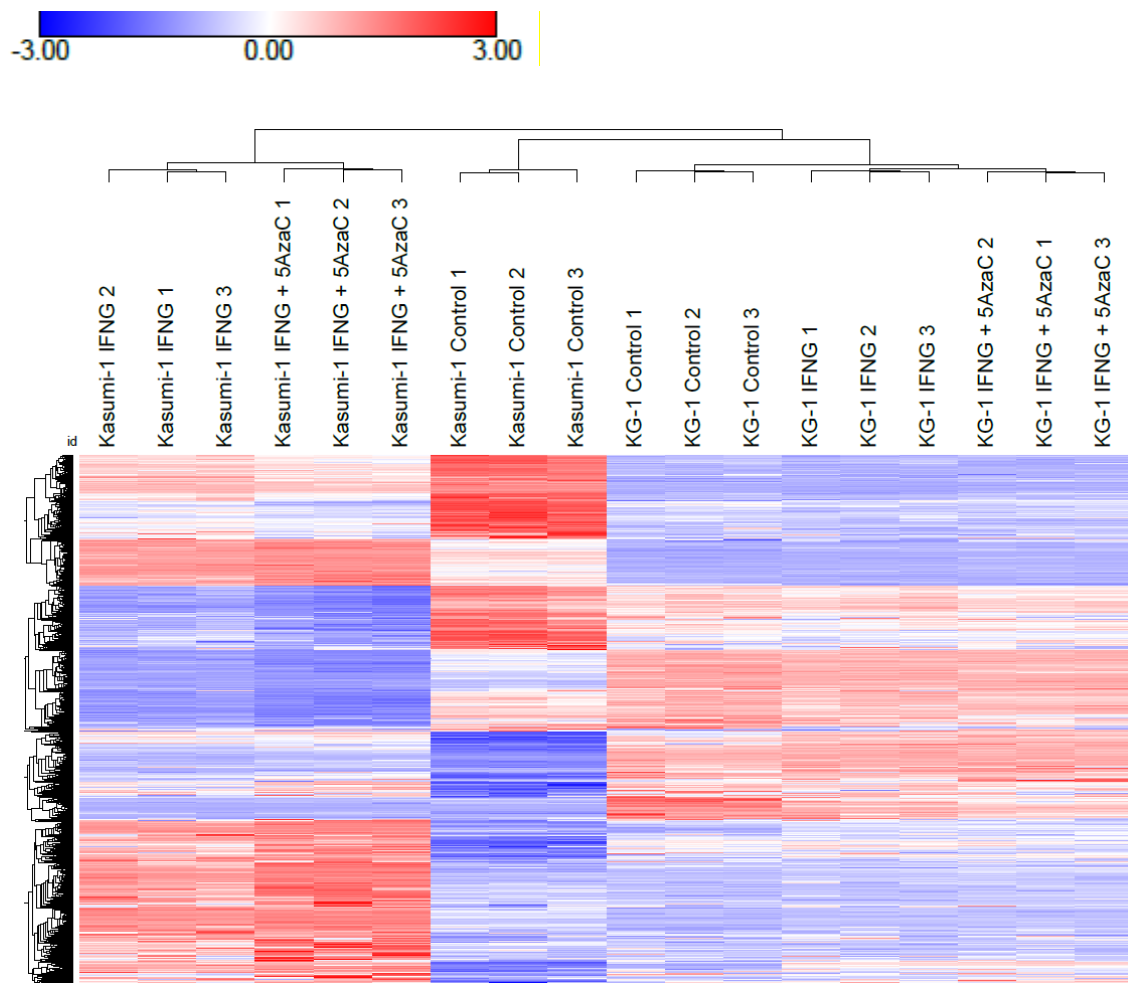
Figure 61: Heat map of the significantly differentially expressed transcripts between Kasumi-1 and KG-1 under all treatment conditions. Hierarchical clustering performed using Euclidean distance and complete linkage (n = 3 per sample type). Key for heat maps shown top left for expression values. Heat maps generated in Morpheus (https://software.broadinstitute.org/morpheus/). Blue indicates lower expression; red indicates higher expression.

Transcript lists for each cell line were filtered for significance Padj < 0.05 and a minimum 1.5-fold change and compared according to if they were up or down regulated. Results are presented as Venn diagrams in Figure 62, A and B, respectively, and results for pathway enrichment are shown in Figure 62. To find differences unique to combination treatment, transcripts that were found to be significantly up or down-regulated in either IFNG only or 5AzaC only treated cells were removed, the number of transcripts left in each instance is shown in the Venn diagrams in Figure 62, C and D. The top 10 significantly enriched pathways were compared to pathways enriched by IFNG treatment. Results can be seen in Figure 63.
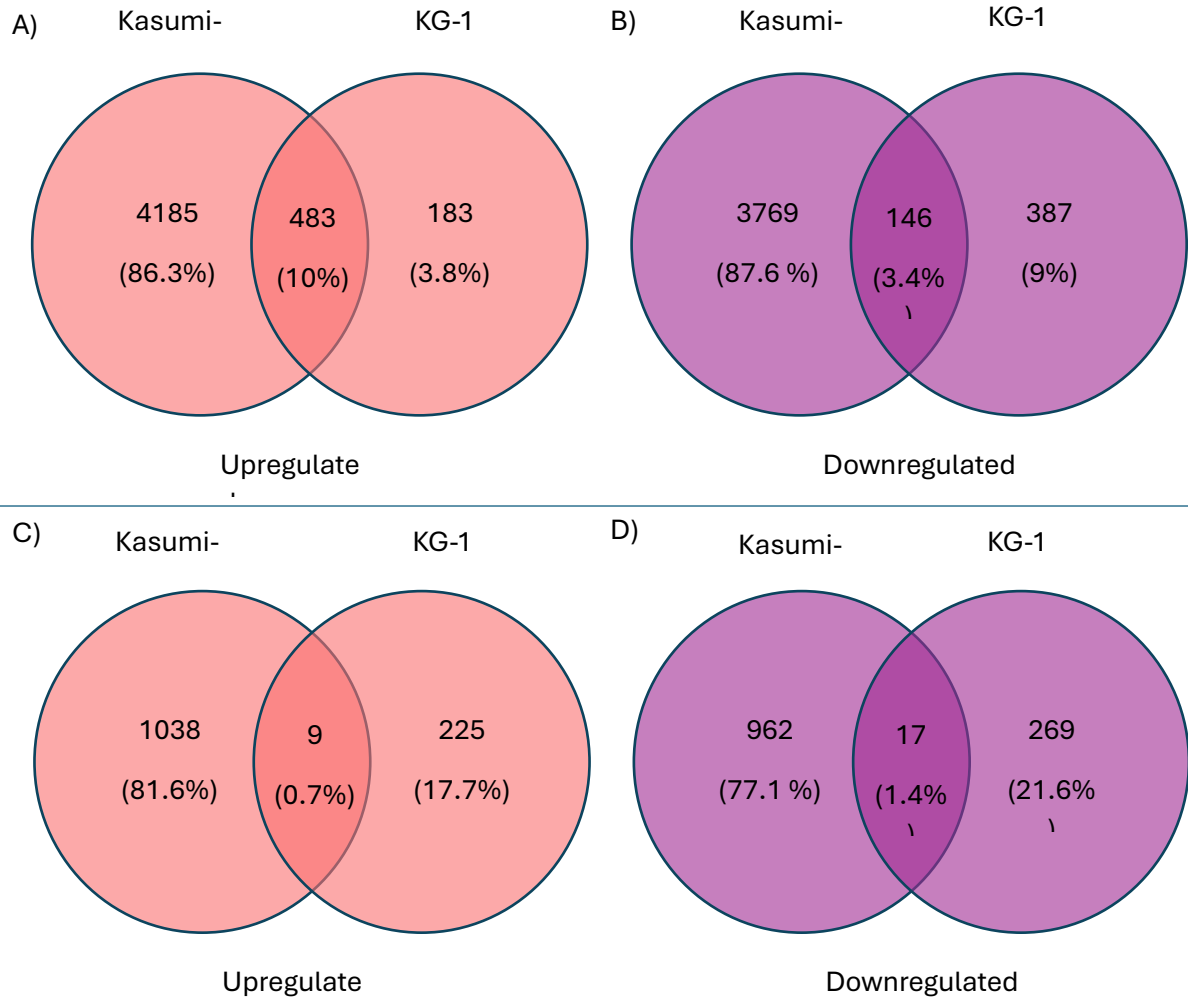
Figure 62: Venn diagrams comparing transcript lists of differentially expressed transcripts of Kasumi-1 and KG-1 under all treatment conditions. Differential expression was defined as a change of 1.5-fold or more that was also statistically significant (Padj<0.05).Comparison of increased transcripts in combination treated vs untreated cells B) Comparison of decreased transcripts in combination treated vs untreated cells C) Comparison of increased transcripts in combination treated vs untreated cells with significant transcripts from IFNG and 5AzaC mono treatment removed D) Comparison of decreased transcripts in combination treated vs untreated cells with significant transcripts from IFNG and 5AzaC mono treatment removed.
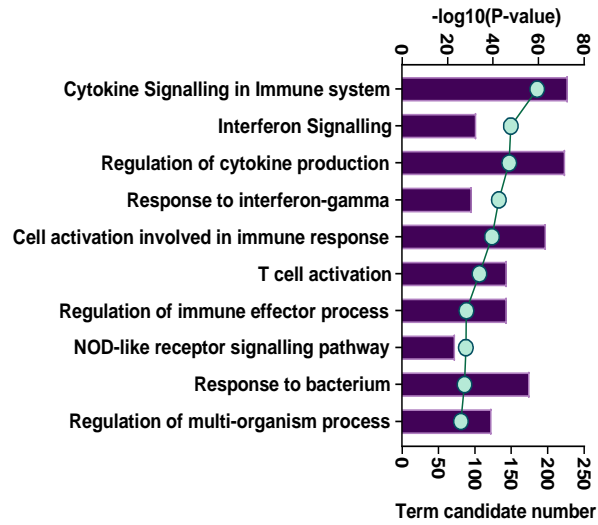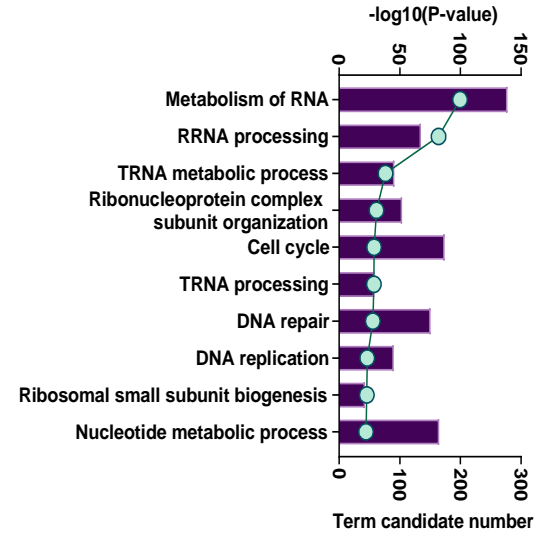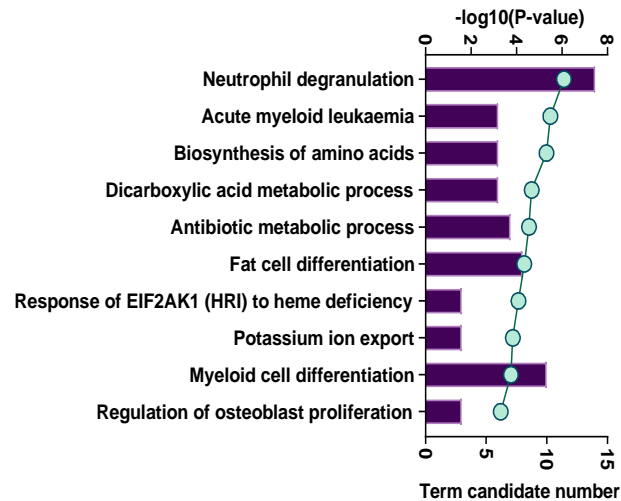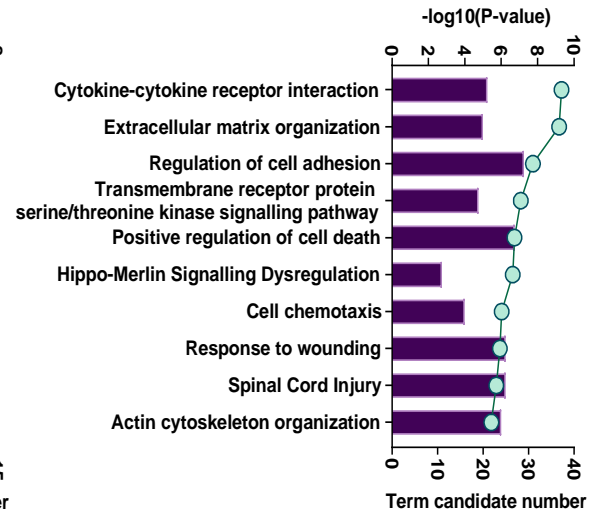
Figure 63: The top 10 most significantly enriched pathways in Kasumi-1 and KG-1 treated with IFNG and 5AzaC. Enriched pathways were identified by the 'unique' up and down regulated transcript lists, sorted by significance post FDR. See appendix for table summary process ids and number of transcripts significantly changed in each pathway.

Prior to removing the IFNG-only and 5AzaC-only transcripts, changes in the combination treatment were driven by interferon signalling. The combination with 5AzaC was shown to enhance the effects of IFNG in both cell lines. In Kasumi-1, altered pathways were thematically the same as IFNG treated cells; pathways reported as upregulated related to interferon signalling and several were immune based responses. Pathway's enrichment shows pre-treatment with 5AzaC bolstered the IFNG response, increasing the number of transcripts significantly altered in reported pathways. Indicating that demethylation synergised with IFNG to promote more IFNG related signalling in Kasumi-1. This was most obvious in 'cytokine signalling in immune system' process, where transcripts upregulated were increased from 157 in IFNG only treatment, to 228 with combination treatment. Detailed tables with number of transcripts upregulated and downregulated compared between IFNG5AzaC combination treatment and IFNG treatment alone are available in the appendix (See 7.3).

In KG-1 reported changes to upregulated processes were related to cell differentiation, cell remodelling and metabolic processes. More pathways were reported as significantly upregulated than when treated with IFNG alone. This could simply be a result of the transcript list having more transcripts. There were only 68 upregulated transcripts 'unique' to KG-1 reported for IFNG alone, and 387 'unique' transcripts reported for IFNG combined with 5AzaC treatment, presenting a more robust data set. Five of the 20 decreased pathways were shared between combination treated and IFNG treated. Compared to Kasumi-1, increases in number of transcripts altered in pathways was very low. Both treatment types caused downregulation of chemotaxis and cell to cell adhesion. There was a decrease in transcripts involved with 'negative regulation of cell proliferation' and 'positive regulation of cell death', indicating a possible increase in cell proliferation signalling, that might link to increased effectiveness of Daunorubicin with combination therapy.

While combination treatment causes more differential expression of transcripts in KG-1 than each treatment alone, demethylation was unable to restore the capacity of KG-1 to respond more fully or traditionally to IFNG. This might be due to KG-1 expressing WT DNMT3A/B to *de novo* methylate DNA, while TET1/2 LOF mutations prevent removal of methyl groups, creating a hypermethylated pattern compared to Kasumi-1.
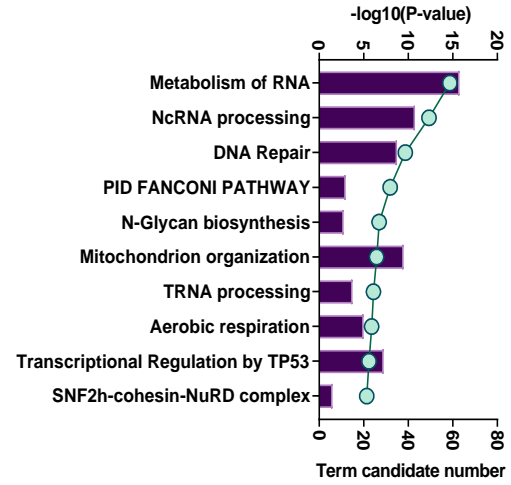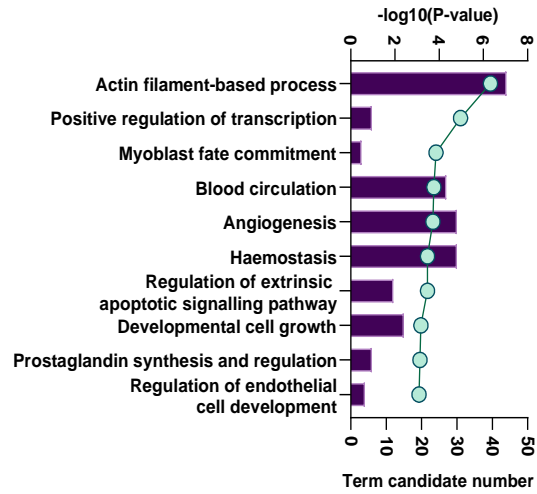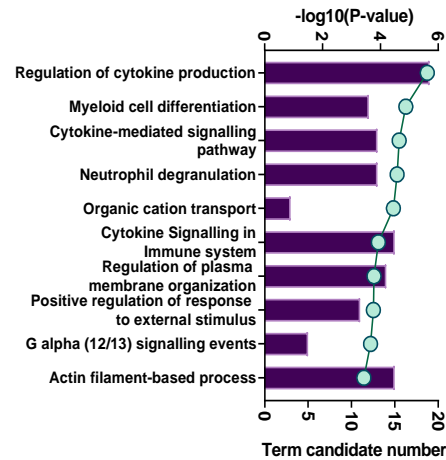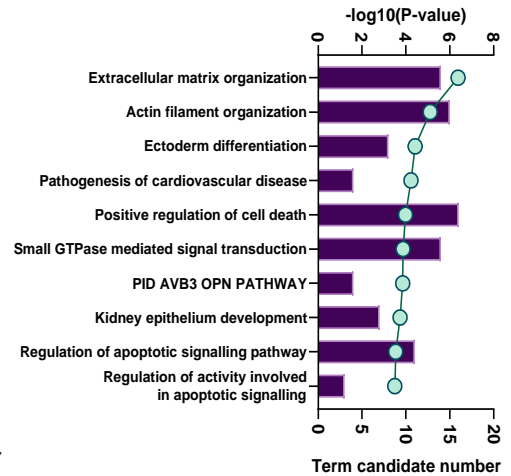
Figure 64: The top 10 most significantly enriched pathways in Kasumi-1 and KG-1 treated with IFNG and 5AzaC. Enriched pathways were identified by the 'unique' up and down regulated transcript lists, sorted by significance post FDR. See appendix for table summary process ids and number of transcripts significantly changed in each pathway

After removing IFNG and 5AzaC monotreatment-specific transcripts, unique pathways to the combination treatment emerged. In Kasumi-1 cells, the most upregulated process was 'Actin filament-based process,' involving transcripts like *MAP1B*, *SPTAN1*, and *TMOD3*. Other upregulated pathways included positive regulation of RNA polymerase II promoter activity in response to stress and angiogenesis due to hypoxia. Additionally, there was increased expression of negative regulation pathways for extrinsic apoptotic signalling, reflecting the dying status of the cells. Conversely, downregulated processes in Kasumi-1 indicated declining cell health, including RNA metabolism, DNA repair pathways, and cell cycle processes. This downregulation, along with inhibited DNA replication, translation, and protein processing, confirms that the combination treatment exacerbates IFNG-induced cell death by breaking down pathways for cell maintenance, growth, and repair.

In KG-1 cells, the response to combination treatment differed. The top upregulated process was 'regulation of cytokine production,' including pathways for negative regulation of lymphocyte activation and positive regulation of immune response. Upregulated pathways also involved antiviral responses and differentiation processes. The most downregulated pathway was 'Extracellular matrix organisation,' with a decrease in transcripts involved in ECM-receptor interactions, such as CD44. KG-1 cells also showed downregulation of apoptotic activity, including pathways for positive regulation of cell death and apoptotic signalling. However, unlike Kasumi-1, KG-1 did not enter a state of cell death. Furthermore, when KG-1 received the combination treatment, a more immunosuppressive response was seen, with downregulation of IL-18 signalling pathways and chemotaxis-related processes. There was no significant decrease in cell growth, DNA repair, transcription, and translation, indicating that KG-1 cells were more successful in mitigating the effects of the combination treatment. The response in KG-1 was more oriented towards immune modulation, with enriched immunosuppressive pathways and decreased expression of transcripts involved in cell motility and adhesion. This differential response suggests that the combination treatment of IFNG and 5AzaC may be further investigated for its potential to prevent egression in blood circulation of AML cells, particularly those resembling the KG-1 phenotype, by targeting cell motility and adhesion mechanisms. IFNG induced sensitivity to Daunorubicin

4.2.7.4   Cell cycle arrest

Daunorubicin, an anthracycline antibiotic, is effective against both solid and haematological cancers. Normally, topoisomerase II binds to DNA and catalyses its cutting and unwinding, making it accessible to transcription enzymes. Daunorubicin traps topoisomerase II in a covalent complex, preventing transcription and DNA replication, which ultimately causes cell death

(Tyleckova, et al. 2012). IFNG treatment increased sensitivity of both cell lines to Daunorubicin, however pathway analysis through metascape did not report any obvious routes through which this was achieved. Expression of markers such as *MKI67, PCNA* and *MCM2* correlate with faster proliferation (Whitfield, et al. 2006). In Kasumi-1, IFNG decreased expression of *MKI67* (Kas-1 Ctrl = 40.15 FPKM, IFNG = 25.91 FPKM, Padj < 0.0001, Wald-test BH), *PCNA* (Kas-1 Ctrl = 129.55 FPKM, IFNG = 82.13 FPKM, Padj < 0.0001, Wald-test BH) and *MCM2* (Kas-1 Ctrl = 53.64 FPKM, IFNG = 30.00 FPKM, Padj < 0.0001, Wald-test BH)(Figure 65).
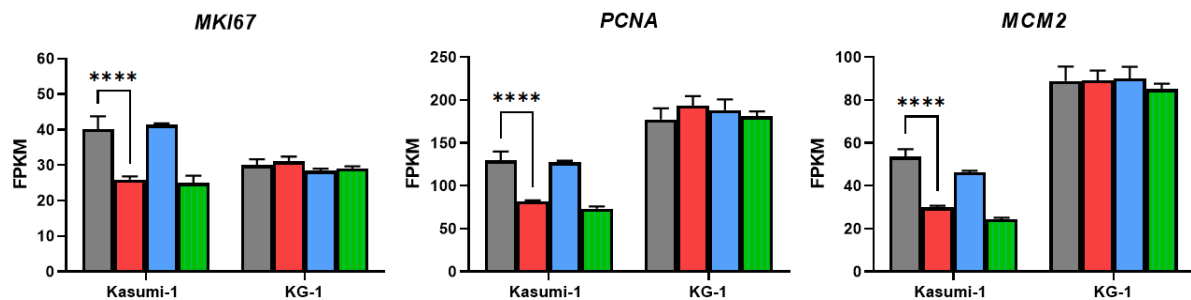
*Figure 65: Expression of transcripts for markers of proliferation in Kasumi-1 and KG-1 under all treatment conditions. Data given as FPKM (n=3). Grey = Control, Red = IFNG, Blue = 5AzaC, and Green = IFNG + 5AzaC. Statistical tests by Novogene Wald test – BH correction. * = Padj< 0.05, ** = Padj< 0.01, *** = Padj< 0.001 and **** = Padj< 0.0001.*

Daunorubicin is most effective in rapidly proliferating cells entering the S-phase of the cell cycle, where DNA is replicated. Despite significant downregulation of DNA replication transcripts in Kasumi-1 cells treated with IFNG, these cells become vulnerable to Daunorubicin. This suggests that IFNG induces a mechanism that overcomes resistance caused by cell cycle arrest. It is possible that IFNG arrests the cells, and the subsequent addition of Daunorubicin inflicts further damage, which cannot be repaired due to downregulated DNA repair pathways. The combination of cell cycle arrest, damaged DNA, inhibited DNA replication, and additional stress from Daunorubicin treatment could commit the cells to death.

4.2.7.5   IFNG treatment reduced expression of two drug resistance transcripts

Out of the mechanisms of drug resistance discussed in section 1.9, only *GSTP1* and *ABCC1* were significantly affected by IFNG treatment (Figure 66). *ABBC1* encodes for multi-drug resistance protein 1, an ATP dependent efflux pump. *GSTP1* is the transcript for a glutathione s-transferase which attaches glutathione to substrates, such as drugs, enabling their removal from the cell by efflux pumps. IFNG treatment decreased expression of *ABCC1* (Kas-1 Ctrl = 22.11 FPKM, IFNG = 12.88 FPKM, Padj< 0.0001, KG-1 Ctrl = 45.51 FPKM, IFNG = 41.32 FPKM, Padj<0.001, Wald-test BH) and *GSTP1* (Kas-1 Ctrl = 236.74 FPKM, IFNG = 114.48 FPKM, Padj< 0.0001, KG-1 Ctrl = 263.74

FPKM, IFNG = 217.39 FPKM, Padj<0.0001, Wald-test BH) in both cell lines. In both cases, the biggest differences were seen in Kasumi-1. Decreased expression of GSTP1 protein could result in reduced inhibition of drug action through GSH conjugation and reduced drug export from the cell. This could lead to higher accumulation of the drug in cells compared to the control, potentially contributing to the observed sensitivity to Daunorubicin in both cell lines when IFNG is present.
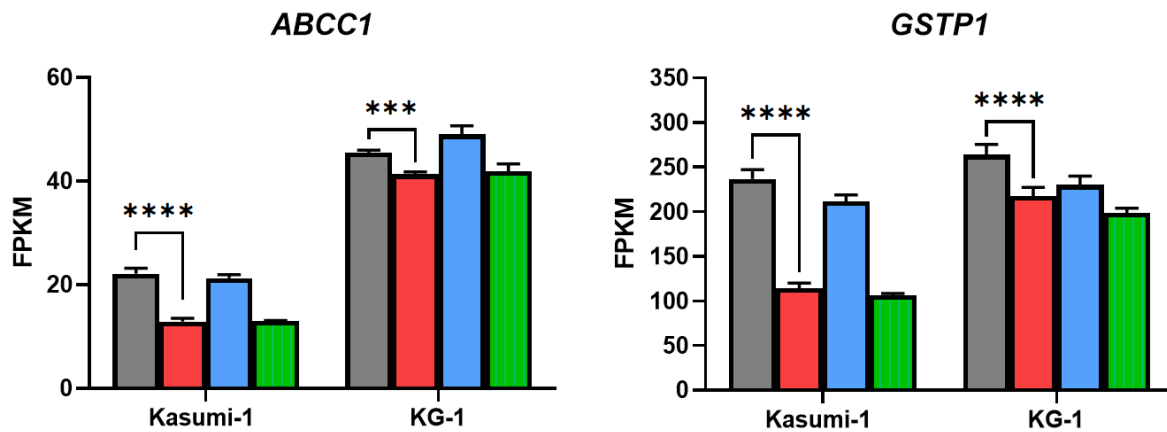
*Figure 66: Expression of ABCC1 and GSTP1 in Kasumi-1 and KG-1 under all treatment conditions. Data given as FPKM (n=3). Grey = Control, Red = IFNG, Blue = 5AzaC, and Green = IFNG + 5AzaC. Statistical tests by Novogene Wald test – BH correction. * = Padj< 0.05, ** = Padj< 0.01, *** = Padj< 0.001 and **** = Padj< 0.0001.*

4.2.7.6 *P2X7 receptor transcript expression increased with IFNG treatment*

The P2X7 receptor is an ATP gated ion channel usually expressed on the cell surface, which following high stimulation by ATP, opens a large none-selective pore and triggers cell death. This pore opening has previously been shown to increase uptake of Doxorubicin and linked to Daunorubicin sensitivity in AML (Pegoraro, et al. 2020). IFNG significantly increased expression of the transcript (*P2RX7*) in both cell lines (Kas-1 Ctrl = 8.39 FPKM, Kas-1 IFNG = 98.13 FPKM, KG-1 Ctrl = 2.47 FPKM, IFNG = 10.55 FPKM, both, Padj < 0.0001, Wald test- BH) (Figure 67). In Kasumi-1 combination of IFNG with 5AzaC further increased expression compared to IFNG alone (Kas-1 IFNG5AzaC = 111.82 FPKM, Padj<0.001, Wald-test BH). This shows another mechanism for further nonspecific uptake of Daunorubicin and a potential reason for increased sensitivity to the drug, especially in Kasumi-1. There are two isoforms for the P2X7 receptor, P2X7RA and P2X7RB. Increased Daunorubicin sensitivity has been associated with increased expression of P2X7RA, and resistance to the drug has been associated with higher P2X7RB levels (Pegoraro, et al. 2020). However, Information for the protein product is unavailable in the proteomics data, therefore it is unknown how affected protein levels are by treatment and which isoform for P2RX7R is present in either cell line.
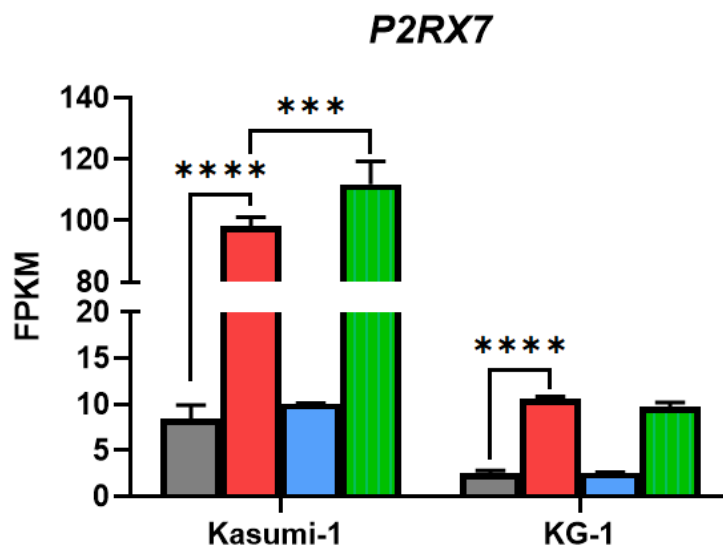
*Figure 67: Expression P2RX7 transcript in both cell lines under all treatment conditions. Data was given as FPKM (n=3). Grey = Control, Red = IFNG, Blue = 5AzaC, and Green = IFNG + 5AzaC. Statistical tests by Novogene Wald test – BH correction. * = Padj< 0.05, ** = Padj< 0.01, *** = Padj< 0.001 and **** = Padj< 0.0001.*

Pathway enrichment by metascape did not overtly provide an explanation for the sensitivity to Daunorubicin observed in IFNG treated cell lines. Some potential mechanisms could be inferred from the transcript expression data and literature available on the subject. Unfortunately, protein data was not available to compare against for many of these molecules. Mechanisms for cell death may have been identified had cells treated with Daunorubicin and IFNG combined with Daunorubicin been analysed by NGS as well.

4.3 Discussion

The hypothesis of this work was that IFNG induces immunosuppressive phenotype in AML cell lines, and that 5AzaC treatment modifies the strength of the IFNG response. Further, that the transcriptomes generated from these changes could be used to create a signature using patient survival data, that stratifies AML patients who receive HSCT or chemotherapy induction treatment. Proteomics data presented in chapter 3 (Figure 18) suggested that Kasumi-1 and KG-1 had different strength responses to IFNG treatment. KG-1 was found to be the least effected by IFNG in terms of number significantly differentially expressed proteins (SIG-M5 = 341, THP-1= 141, Kasumi-1 = 101, KG-1 = 48), where significance was determined by confidence and fold-change method (Lambert, et al. 2013). In this chapter, transcriptomics analysis echoed that Kasumi-1 was more responsive than KG-1 to IFNG with 11,377 and 2,564 transcripts significantly differentially expressed respectively. It found that IFNG induced transcripts associated with immunosuppressive phenotype in both AML cell lines. Additionally, treatment with 5AzaC prior

to IFNG was found to further induce expression of these transcripts in Kasumi-1, but not in KG-1. The divergence in response might be due to differences in mutations of methylation machinery, however, this was unable to be confirmed. Cell line phenotypes were examined by pathway enrichment and characterised based on expression of specific transcripts involved in processes of interest.

Metascape analysis of both cell lines revealed enriched pathways such as NF-Kappa Beta signalling and the MAPK cascade. Depending on molecular crosstalk from other cell types like immune cells (absent in the model), these pathways could either prevent or promote apoptosis (Zhang, Wei and Liu 2002, Hoesel and Schmid 2013). Specifically in Kasumi-1, there was downregulation of cell cycle, DNA repair, and DNA replication pathways, alongside increased vesicle-mediated transport, consistent with observed cell death in viability experiments (Figure 24). The upregulation of autophagy suggests an effort to protect against IFNG related stress, but the effort might then be superseded by enrichment of several apoptosis pathways (Mariño, et al. 2014). Importantly, 'Positive regulation of cell death' was markedly enriched in Kasumi-1 but absent in KG-1, aligning with the results of cell death assays in Chapter 3. In contrast, KG-1 exhibited significantly upregulated processes such as cytokine signalling activity, immune response, and IFNG response. The MAPK stress-activated cascade was also upregulated, but unlike Kasumi-1, no other processes associated with cell death were reported. Upregulation of exocytosis and transport of small molecules may suggest mechanisms to eliminate IFNG from cells. While Kasumi-1 responded to IFNG by halting the cell cycle, KG-1 displayed a more immunosuppressive and anti-inflammatory response. Moreover, there was a trend of downregulating pathways involved in cell chemotaxis and adhesion, potentially to evade immune response by minimising interactions with immune cells. These pathways indicated that Kasumi-1 was vulnerable to an active immune response involving IFNG delivery by T-cells and NK cells, whereas KG-1 was less susceptible. Further investigation into pathways regulating apoptosis compared expression of specific molecules between cell lines.

Cell lines were found to have distinct responses to both IFNG and 5AzaC treatment. Kasumi-1 displayed a hyper IFNG response, seen by induced changes to thousands of transcripts. Demethylation was also effective and found to synergise with IFNG to boost IFNG response. The low dose of 5 ng/mL IFNG was enough to induce apoptosis programmes and cell cycle arrest. Suggesting that these cells would be vulnerable in an inflamed microenvironment with immune cells. However, the high IFNG response also caused massive upregulation of several

immunosuppressive transcripts (*TRAIL, CD274, CD47, LGALS9* and *HAVCR2)*. It may be that a lower dose of IFNG would not induce apoptosis, and only the immune suppression program. Low SHP1/2 expression could indicate hyper IFNG response, and a vulnerability to treatment by HSCT to restore local IFNG production, combined with immune blockade, to combat immune resistance mechanisms. SHP2 null cells have been shown to be vulnerable to apoptosis when treated with IFNG compared to SHP2-expressing cells. You et al previously reported that SHP2 null mouse fibroblast cells had high STAT1 phosphorylation (You, Yu and Feng 1999). In turn, the phosphorylated STAT1 had enhanced DNA binding capability, which was able to suppress cell growth and induce cell death. This may be the mechanism for IFNG induced cell death in Kasumi-1. Summary of Kasumi-1 key features are presented as a diagram in Figure 68.
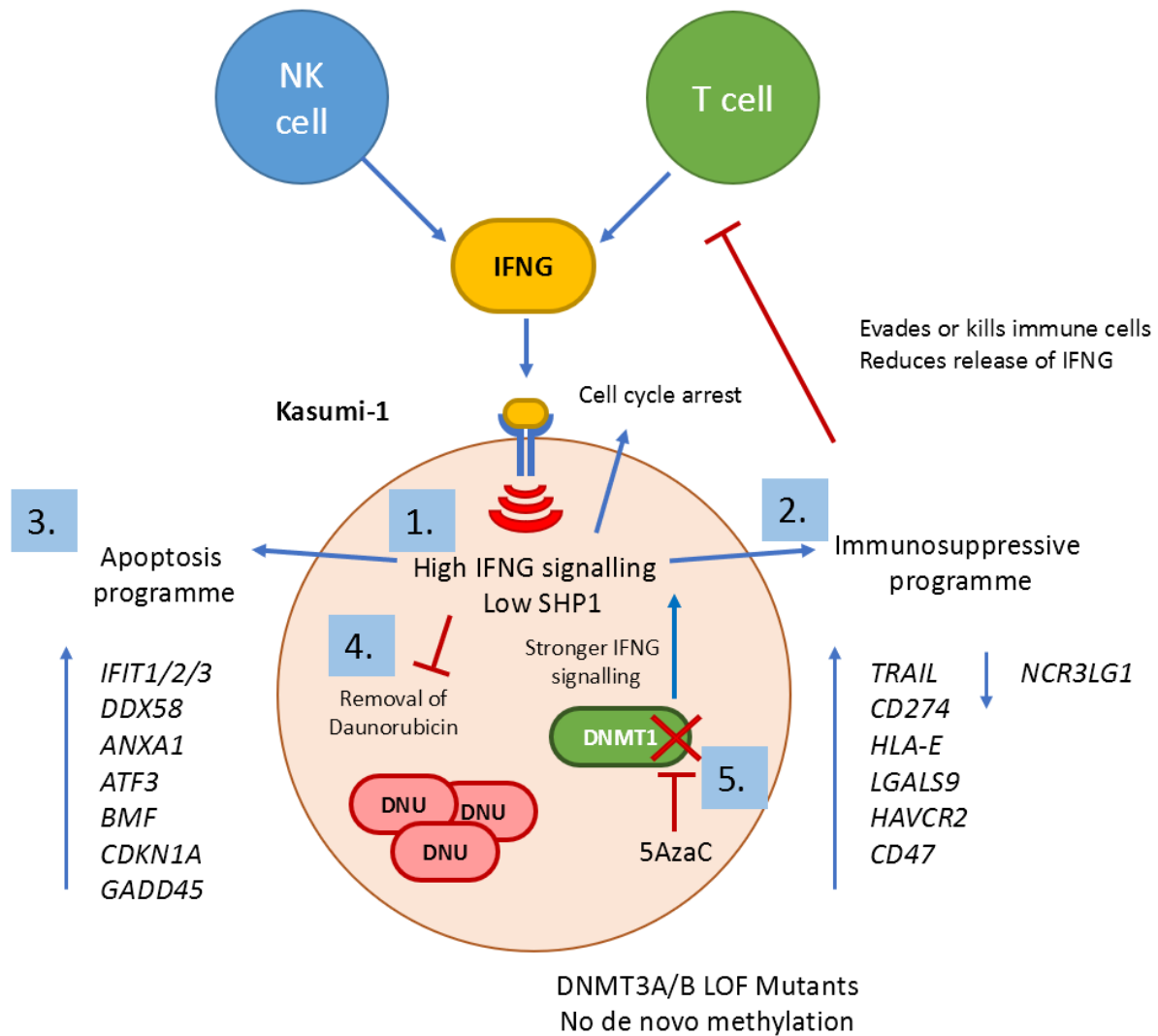
Figure 68: Figure summarising key observed effects of Kasumi-1 treatment with IFNG or 5Azac. 1) Due to Low SHP1 expression, negative feedback is lowered, causing higher intensity signalling. 2) IFNG upregulates transcription of several immunosuppressive transcripts whose products either evade or kill immune cells. 3) Cell cycle arrest and apoptosis is induced by IFNG dosing. 4) IFNG decreases expression of genes for efflux pumps enzymes involved in removal of xenobiotics from the cell, causing an accumulation of Daunorubicin. 5) Demethylation by 5AzaC is effective due to mutations in DNMT3A/B preventing de novo methylation, and wild type TET1/2/3 still actively removing sites. DNMT1 is the only means of maintaining methylation. Successful demethylation reduces silencing at IFNG related gene promoter sites, increasing expression of genes seen in the apoptosis and immunosuppressive programme. DNU = Daunorubicin.

KG-1 on the other hand could be described as resistant to IFNG, possibly due to the high expression of SHP-1, with IFNG response transcripts displaying muted expression. Treatment with IFNG did not induce the apoptosis programme seen in Kasumi-1, but was still able to induce cell cycle arrest, and a muted immunosuppressive response. The response of KG-1 to 5AzaC was also minimal. DNMT3A/B are reported as not mutated in this cell line and thus still able to methylate new sites. Combined with TET1/2 LOF mutations which might prevent active removal of methyl sites, this cell line could be hypermethylated, and reduction of DNMT1 might not have as much of an impact on transcription as it does in Kasumi-1. Treatment with 5AzaC did not holistically increase IFNG signalling as it did in Kasumi-1 either. Despite the low IFNG response (relative to Kasumi-1), IFNG still induced sensitisation to Daunorubicin driven apoptosis in KG-1. Importantly, KG-1 expressed high baseline levels of ANXA1 protein (See 7.6), which signalled for inflammation resolution, dampening immune response. It also expressed high levels of *CD47* to avoid immune cells. The high expression of SHP1 and ANXA1 proteins (See 7.6), along with *CD47* mRNA (Figure 54) could make KG-1 recalcitrant to recognition and removal by immune cells. AML with these features could be resistant to HSCT, and immune blockade, and better treated with chemotherapy. In KG-1, while ANXA1 expression was constitutively high, IFIT1 and RIG-1 expression remained at low levels compared to Kasumi-1, post IFNG treatment. This could explain why KG-1 was not prone to apoptosis, despite high ANXA1 expression. Restoration of the IFNG signalling pathway in KG-1, by blocking or downregulation of SHP-1, may restore IFIT1 and RIG-1 induction, and possibly induce apoptosis via this pathway. A summary of KG-1 key features is presented in Figure 69.
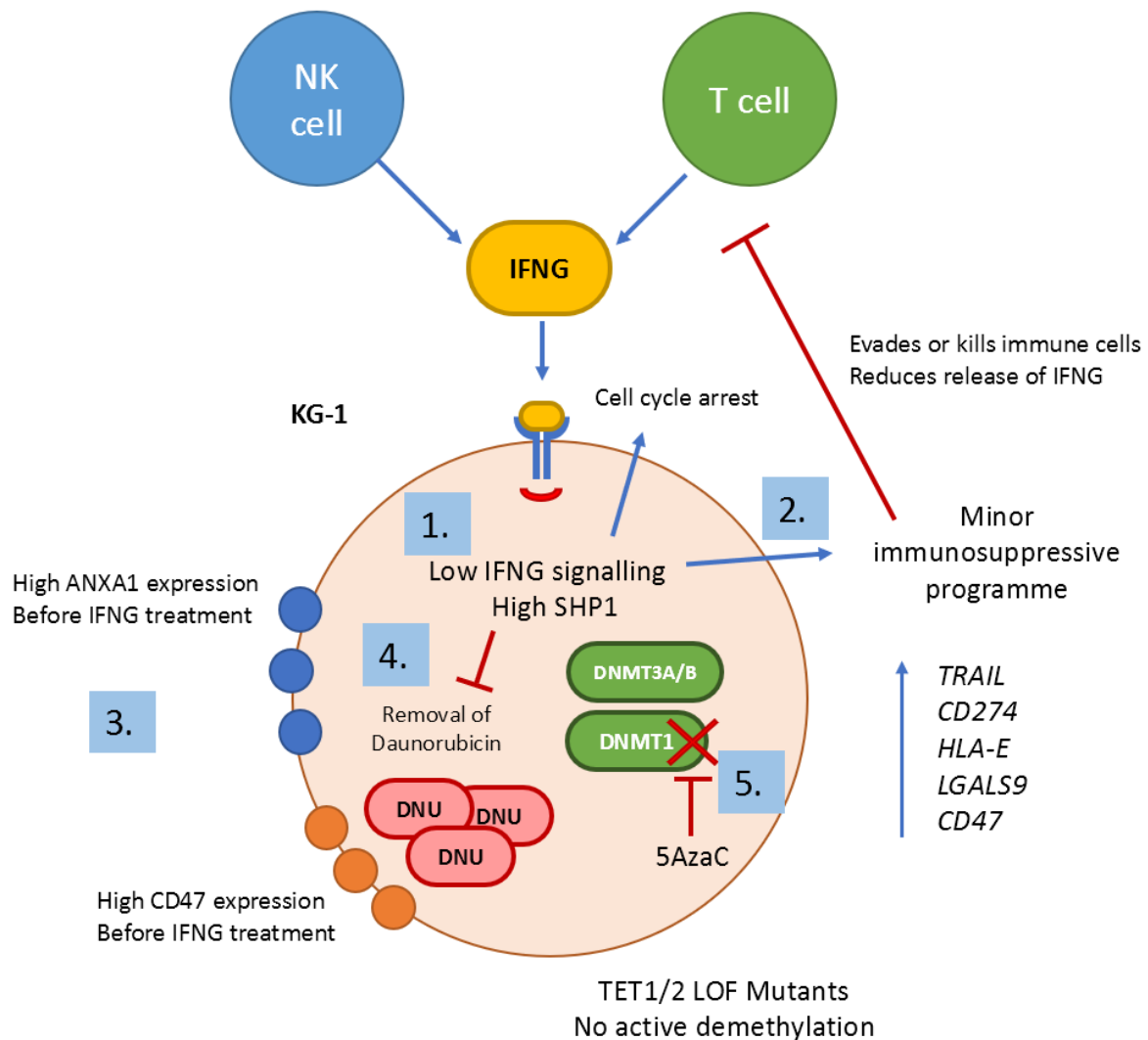
Figure 69: Figure summarising KG-1 before and after treatment with IFNG or 5Azac. 1) Due to high SHP1 expression, negative feedback already in place, causing reduced IFNG signalling, and a muted IFNG response. 2) IFNG upregulates transcription of immunosuppressive transcripts, but not to the same degree as Kasumi-1. Additionally, IFNG does not induce apoptosis. 3) High basal expression of ANXA1 and CD47 aid in immune evasion, IFNG treatment further increases expression of CD47. 4) IFNG mildly decreases of transcripts for efflux of xenobiotics from the cell, causing an accumulation of Daunorubicin. 5) Demethylation by 5AzaC is less effective due to wild type DNMT3A/B still adding new methylation sites, and mutant TET1/2 no longer removing methyl groups. DNU = Daunorubicin.

There are limitations to the conclusions that can be drawn from this study. Firstly, the data was generated from individual AML cell lines without co-cultures involving immune cells. Secondly, while NGS is sensitive and reproducible, it does not directly reflect protein expression.

As this data was generated from cell lines in isolation, with no immune cells present, some of the pathways enriched may be limited or irrelevant. The reported activation of immune cells and inflammatory response was limited in relevance as no immune cells were present in the model. While inflammatory molecules may be upregulated in the cell lines, without immune cells present, we cannot say that IFNG treatment would induce mobilisation and activation of immune cells from this model.

Secondly this work used Next-generation sequencing (NGS) to obtain the transcript expression data. The transcriptome includes protein-coding RNAs, non-coding RNAs, and regulatory microRNAs (Djebali, et al. 2012). RNA sequencing (RNA-seq) holds advantages over alternatives like DNA microarrays, offering heightened sensitivity for detecting low-expression transcripts, independence from prior genome knowledge, and strong reproducibility (Wang, Z., Gerstein and Snyder 2009). Despite these merits, transcriptomic expression levels do not universally mirror eventual protein expression. Protein-mRNA correlation is often modest (R < 0.4), suggesting mRNA accounts for only 40% of protein variance (Chakraborty, et al. 2018). The residual 60% discrepancy stems from processes such as transcription, translation, mRNA and protein degradation. Translation, notably, exerts the most influence on protein expression levels (Schwanhäusser, et al. 2011). Consequently, combining transcriptomics with proteomics yields a more comprehensive cellular snapshot. While cell lines may report significant induction of transcripts, these mRNA may not be translated into proteins. Translation is a highly regulated process and can be prevented by mechanisms including but not limited to post-transcriptional modification such as splicing, RNA degradation by microRNAs (mRNAi), RNA binding proteins which prevent ribosome access and low availability of tRNAs or amino acids (Baker and Coller 2006). Furthermore, NGS probes a sample's transcriptome recording all its RNA transcripts at a specific time, introducing a temporal variable into data collection.

In this study, the Pearson correlation coefficient was employed to quantify the linear relationship between transcript expression and its corresponding protein expression. This method was used to assess RNA and protein abundance by measuring the strength and direction of association between them. Essentially, it determines if increases in RNA levels correlate with corresponding increases in protein levels. In this study the hypothesis being tested is that IFNG induces immunosuppressive phenotype in AML cell lines, and that this phenotype could be present in

AML patients that respond poorly to HSCT or chemotherapy induction. For this phenotype to be expressed immunosuppressive transcripts must be translated into functional immunosuppressive proteins. Pearson correlation was used here to assess if IFNG induced changes in transcripts correlate to induced protein expression.

Notably, in this work, proteomics yielded smaller datasets; baseline proteomics detected 3,744 proteins across cell lines, contrasting with 58,735 transcripts detected by NGS. Upon IFNG treatment, 138 and 67 proteins, along with 11,377 and 2,564 transcripts in Kasumi-1 and KG-1, respectively, experienced significant differential expression. Altered proteins were scrutinised alongside corresponding transcripts via Pearson correlation, with only significantly modified targets considered, reducing background noise. In both datasets, only 137 targets in Kasumi-1 and 67 in KG-1 demonstrated significant alterations.

Subsequent correlation significance testing (Figure 70) encompassed untreated and IFNG-treated data. Outliers inflating R^2 coefficients were highlighted. Significance was established for both untreated and IFNG-treated Kasumi-1, albeit with low R^2 values (0.087 and 0.043). Notably, KG-1 exhibited significance only in the IFNG-treated context, with low R^2 values (0.0386 for untreated and 0.0706 for IFNG-treated). These low correlation coefficients indicated a minor portion of protein variance is attributed to mRNA variance.

Comparing KG-1 and Kasumi-1, the observed correlation between transcripts and proteins fell below Chakraborty's R < 0.4 benchmark. The conclusion from these results is that the increase in transcript abundance induced by IFNG was not reflected in equivalent increased protein expression. Although the transcriptomes indicate an immunosuppressive phenotype, the proteome might not. However, Notably, the comparison's validity was compromised due to differential IFNG dosing (5 ng/mL for NGS and 100 ng/mL for proteomics), introducing unaccounted variance.

However, protein abundance does not necessarily reflect protein activity, as various factors besides its expression level can influence protein function. These factors include post-translational modifications, protein conformation, localisation, and degradation mechanisms, all of which can hinder interactions with other molecules (Lee, et al. 2023, Tan, et al. 2019). Additionally, some protein mutations that cause loss of protein activity do not affect protein abundance, but rather alter the catalytic or substrate-binding properties of the protein (Cagiada, et al. 2021). Furthermore, protein abundance and sample complexity can influence the availability of proteins for mass spectrometric quantitation, which may introduce biases in the measurement of protein. Therefore, protein abundance is not a reliable indicator of protein

activity, and other methods are needed to assess the functional state of proteins in biological systems. Signatures derived from IFNG treated AML cell lines may indirectly confirm an immunosuppressive phenotype if high expression of signatures correlates with poor survival in patients treated with HSCT.
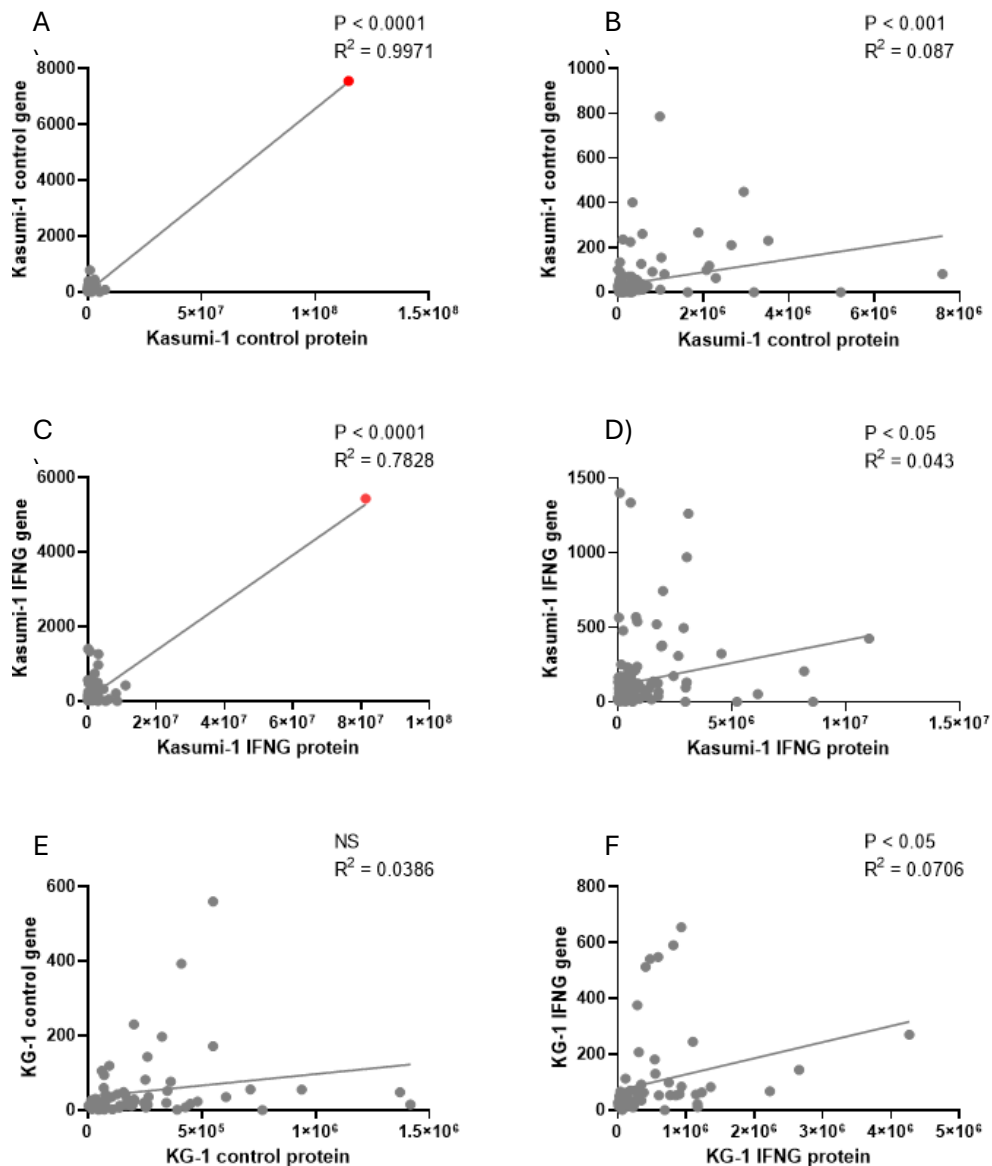
Figure 70: Pearson correlation plots comparing transcript and protein expression in Kasumi-1 and KG-1, before and after IFNG treatment. Kasumi-1 untreated and IFNG treated (Protein: n=6, transcript: n=3). KG-1 untreated and IFNG treated (Protein: n=5, transcript: n=3). A) Comparison of Kasumi-1 untreated transcript and protein expression before outlier removed, B) Kasumi-1 untreated transcript and protein comparison after outlier was removed, C) Kasumi-1 IFNG treated transcript and protein expression before outlier removed, D) Kasumi-1 IFNG treated transcript and proteins expression after outlier removed, E) KG-1 untreated comparison between transcript and protein expression, no outliers, F) KG-1 IFNG treated comparison between transcript and protein expression, no outliers. Red dots indicate outliers.

Transcriptomics data showed kasumi-1 to be hyper responsive to IFNG, while KG-1 held a muted response. Despite the disparity in IFNG responsiveness, both cell lines were equally vulnerable to Daunorubicin when treated with IFNG in advance. Showing a fundamental change occurs, even with low level signalling. Differences in cell line response also suggest there may be a 'threshold' of IFNG signalling, that leads to apoptosis. Cells like Kasumi-1, which exceed the IFNG signalling threshold by decreasing negative feedback through low SHP1 expression, are vulnerable to immune cells, and so also increase immunosuppressive molecule production to protect themselves. In contrast, KG-1 protects itself from immune cells by limiting IFNG signalling via negative feedback though high SHP1 expression, and high expression of CD47 and ANXA1 to evade immune cells and decrease local inflammation.

As IFNG and 5AzaC treated AML cell lines were found to have IFNG inducible immunosuppressive phenotypes, it was determined could be used as models to represent AML which might be resistant to chemotherapy and HSCT treatment. It is important to note, that while cell lines provide an inexpensive and easy to manage model for studying cancers, discoveries made in them are not always applicable to patients. Cell lines change when they undergo immortalisation and are not a direct representation of the primary sample they were taken from. KG-1 cell lines originate from AML with minimal maturation (FAB M1), while Kasumi-1 represents a more mature AML (FAB M2). AML is highly heterogenous, as evidenced by the M1- through M7 FAB classification system and WHO subdivision into even more groups. Heterogeneity is influenced by many factors, including but not limited to, HSC cell of origin, mutational heterogeneity, epigenetic alterations and changes due to treatments administered (Gu, Dickerson and Xu 2020, Li, Sheng, Mason and Melnick 2016a, Horibata, et al. 2019). For this reason, the IFNG response is also highly heterogenous, and cannot be generalised across all AML. Therefore, discoveries in the AML cell line models may only be applicable to a small subsection of AML and not representative of all AML. Markers found in cell lines therefore require further testing for validity, which can be done using the large publicly available cancer patient data sets such as TCGA.

5 In silico investigation of markers of interest in publicly available datasets

5.1 Introduction

Prognostic signatures are quantifiable profiles composed of multiple biomarkers whose combined expression can be used to distinguish a biological state or malignancy, such as cancer. Different types of components in prognostic signatures include gene expression, gene mutations, protein expression, and patterns of epigenetic modification. The uses of prognostic signatures are varied and can be applied to diagnose a disease, predict treatment response, or monitor a disease state (Prada-Arismendy, Arroyave and Röthlisberger 2017).Bensalah et al. summarises cancer-related signatures as profiles composed of molecules produced by the tumour or by the body in response to the tumour (Bensalah, Montorsi and Shariat 2007). Their paper highlights six types of biomarker applications in cancer summarised in Table 14.

Table 14: summary of biomarker applications as adapted from Bensalah et al (Bensalah, Montorsi and Shariat 2007).

Category	Application
Early detection	Used for screening of patients to reveal early-stage cancer.
Diagnostic	Identifies if a tissues mass or other sample type contains cancer.
Prognostic	Categorises a patient into risk groups and guides clinicians on best treatment.
Predictive	Predicts response to treatment or can be used to monitor patient response to treatment.
Therapeutic target	A molecule that is targeted by specific therapy and whose presence indicates response to said therapy.
Surrogate end point	A marker that can substitute for clinical end points such as disease related mortality, relapse, and disease reoccurrence.

Successful prognostic signatures display high sensitivity and specificity for the condition of interest. They should be detectable in blood or tissue samples through routine clinical tests. To be implemented in clinic, they must exceed the accuracy of currently used methods of stratification or successfully identify patient groups that existing methods cannot. There are three core phases for the development of prognostic signatures as highlighted by the Broad Institute's

"team approach" and Bensalah et al. (Bensalah, Montorsi and Shariat 2007, Broad Institute 2016).

1. Discovery: Multiomics analysis of cell lines, animal models, or clinical samples curates a list of potential signature candidates for further investigation.
2. Verification: Signatures are assessed by various techniques such as immunohistochemistry, ELISAs, QRTPCR in clinical samples, or by in silico analysis in patient databases. The relationship of the signature to clinical outcomes or biological status must be verifiable in a relevant patient population.
3. Validation: The robustness of the signature is tested in a large patient group separate from the population used for discovery. This stage identifies if the signature correctly picks out true positives (sensitivity) and rejects false positives (specificity), usually by AUROC. Additionally, the signature is compared to other methods of clinical stratification currently in use.

Prognostic signatures are an essential tool for drug development in an industrial setting. Many drugs are discarded at early stages of trials due to inefficiency compared to the current standard. The development of prognostic signatures to detect patients more suited to certain treatments could improve patient care and make it more financially viable for industrial groups to invest in novel drug pipelines. Developing signatures to identify valid targets for current therapies could be more cost-effective than blindly developing new drugs and administering them non-selectively.

5.1.1 Prognostic Signatures in AML

Many types of prognostic signatures are implemented for the assessment of AML patients. These include genetic markers, protein markers, methylation patterns, and cytogenetics. Cytogenetics refers to the state of chromosomes, which may be missing, rearranged, or broken. A combination of cytogenetic and molecular markers is used to stratify patients into risk groups (Table 15). In AML, cytogenetics is routinely used to differentiate between patients in favourable, intermediate, and unfavourable risk groups.

Patients with chromosomal abnormalities such as t(8;21), t(15;17), inv.(16), and t(16;16) are considered favourable, in addition to those presenting normal cytogenetics and an NPM1 mutation with WT-FLT3 or isolated biallelic CEBPA mutation. The favourable group achieves remission 90% of the time and has an overall survival (OS) of 60% (Prada-Arismendy, Arroyave and Röthlisberger 2017). Patients who have a poor-risk karyotype, including inv(3), t(3;3), t(6;9), −5, 5q-, −7, 7q-, or a complex karyotype, as well as those with normal cytogenetics and an FLT3-ITD mutation, exhibit high treatment resistance to induction chemotherapy. This group also has

an increased likelihood of relapse, as well as low disease-free survival and overall survival rates between 5 and 15% (Prada-Arismendy, Arroyave and Röthlisberger 2017).

Table 15:Summary of cytogenetic and molecular abnormalities used for risk stratification in AML patients, adapted from (Döhner, et al. 2022).

1. Prognostic	Genetic abnormality
Favourable	-t(8;21)(q22;q22.1)/*RUNX1*::*RUNX1T1* -inv(16)(p13.1q22) or t(16;16)(p13.1;q22)/ *CBFB*::*MYH11* -Mutated *NPM1* without *FLT3*-ITD -bZIP in-frame mutated *CEBPA*
Intermediate	-Mutated NPM1 with FLT3-ITD -Wild-type NPM1 with FLT3-ITD (without adverse-risk genetic lesions) -t(9;11)(p21.3;q23.3)/MLLT3::KMT2A†, -Cytogenetic and/or molecular abnormalities not classified as favourable or adverse
Unfavourable	-t(6;9)(p23.3;q34.1)/DEK::NUP214 -t(v;11q23.3)/KMT2A-rearranged -t(9;22)(q34.1;q11.2)/BCR::ABL1 -t(8;16)(p11.2;p13.3)/KAT6A::CREBBP -inv(3)(q21.3q26.2) or t(3;3)(q21.3;q26.2)/ GATA2, MECOM(EVI1) -t(3q26.2;v)/MECOM(EVI1)-rearranged -−5 or del(5q); −7; −17/abn(17p) -Complex karyotype monosomal karyotype -Mutated ASXL1, BCOR, EZH2, RUNX1, SF3B1, SRSF2, STAG2, U2AF1, and/or ZRSR2 -Mutated TP53a

The largest group of patients (~45%) displays normal cytogenetics and is categorised as intermediate risk, with variable treatment responses. It is probable that some of these patients may belong in the 'unfavourable' risk subgroup; development of molecular markers may help to categorise these patients to receive more suitable therapy (Prada-Arismendy, Arroyave and Röthlisberger 2017).

As mentioned in Chapter 1, a prognostic index (PI) is a robust type of prognostic signature created through regression analysis to integrate multiple prognostic factors. Each factor is assigned a weight, calculated through regression analysis and denoted as the B-value, which is then converted into a hazard ratio to describe how each unit increase in the prognostic factor contributes to risk. In this chapter, a PI was developed using transcriptomic data from IFNG and 5AzaC-treated cell lines and verified and validated in silico with clinical and survival data from patient datasets. An IFNG-related PI score was anticipated to be linked with immunosuppressive AML, with high IFNG and IFNG5AzaC PI scores expected to correlate with poorer overall survival.

The workflow for this chapter is as follows:

1. The Kasumi-1 and KG-1 cell line transcriptomic data generated in chapter 4 was analysed for treatment induced changes by multiple pairwise linear regression analysis.

2. A short-list of transcripts for each treatment type was generated based on magnitude of change and frequency of appearances in the top 500 most differentially expressed transcripts for each pairwise comparison performed.

3. For each treatment type, the shortlist was analysed by cox proportional hazards regression using the expression and clinical overall survival (OS) data from the TCGA patient data set.

4. Models were built using the forward step selection method by CPH regression. Based on Blamey et al, each transcript was weighted by B-value to generate final PI scores.

5. Scores were assessed in TCGA for stratification and association with clinical subgroups.

6. PI scores were validated in the Beat-AML, German-AML, CN-AML and HOVON data sets.

Performance of new PI scores was critically evaluated against existing signatures and stratification methods.

5.2 Method to develop Prognostic Index (PI) scores

Signatures were generated from the cell line RNAseq data obtained from treatment of Kasumi-1 and KG-1 cell lines with IFNG, 5AzaC or IFNG and 5AzaC in combination. To evaluate which transcripts were most changed by treatment, pairwise linear regression was performed between transcriptomes of treated and untreated cell lines. For each treatment, this resulted in 36 pairwise comparisons. From there, the top 500 transcripts with the largest standard residuals ((observed y value – predicted y value)/standard error) (SR) were chosen. A shortlist of candidates was created where only transcripts that appeared in the top 500 of every pairwise comparison was carried forward. Those candidates were then looked for in the TCGA patient database, where CPH regression analysis was performed with each list against overall survival with forward selection used to build the models. The process is summarised below with each step described in more detail in the following pages:

Treatment condition		
IFNG	5AzaC	IFNG + 5AzaC
58,734 transcripts		

Step 1: 36 Pairwise linear regressions of each cell line under listed treatment condition vs untreated. Standard residuals ordered largest to smallest, top 500 targets for each list carried forward and duplicates removed.

| 1176 transcripts | 1411 transcripts | 1162 transcripts |

Step 2: Only transcripts that appear in the top 500 of every pairwise comparison carried forward

| 55 transcripts | 15 transcripts | 56 transcripts |

Step 3: Transcripts appear in TCGA-LAML data set

| 50 transcripts | 8 transcripts | 50 transcripts |

Step 4: Cox proportional hazard regression with forward selection in TCGA-LAML data set to build a model for each treatment type

| 9 transcript score | 2 transcript score | 5 transcript score |

Figure 71: Flow chart showing the steps involved in generating a signature from transcriptomics data and TCGA-LAML patient data set.

5.2.1 Step 1: Pairwise comparisons

To narrow down the pool of candidates from the transcriptomics data for survival analysis, pairwise comparisons were performed by linear regression. For each pair, regression was performed on the expression of all 58,735 transcripts. The analysis was performed in excel using the 'Excel data analysis toolpak' add-on. For each treatment type, regression was performed between treated cell lines and untreated cell lines. For example, data from 5AzaC treated Kasumi-1 biological repeat 2 (KA_5A_2) was regressed against untreated Kasumi-1 biological repeat 1 (KA_C_1), and then KA_C_2, KA_C_3 and so on with each treatment type. To generalise the model more, regressions were performed across cell lines too, eliminating cell lines as a contributing factor to change. Control was not compared to control and treated was not compared to treated. Figure 72 below shows the possible 36 comparisons between each treated cell line and control.
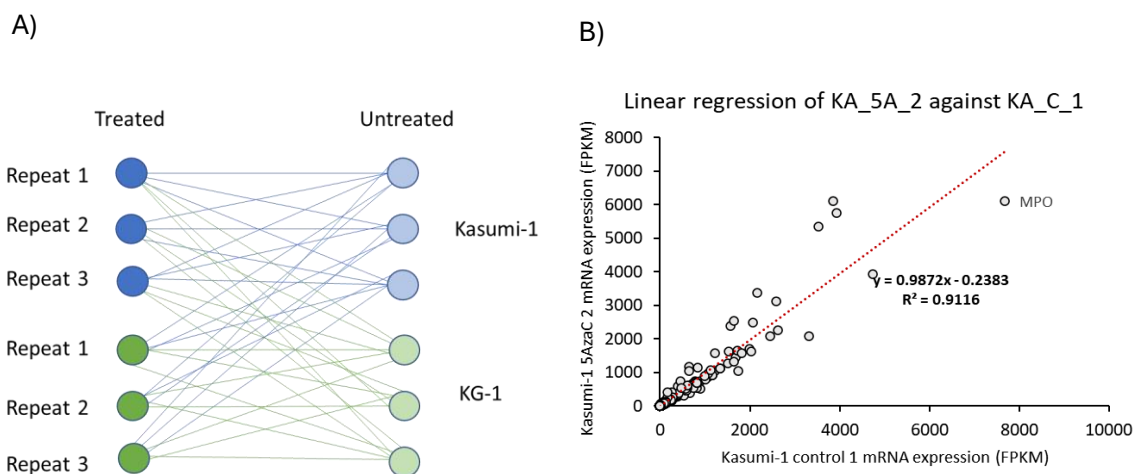
Figure 72:A) Shows possible comparisons between each treated and untreated cell lines where treated could be IFNG, 5AzaC or a combination of IFNG and 5AzaC, 3 biological repeats were available for each condition. B) Pairwise linear regression of Kasumi-1 5AzaC repeat 2 against Kasumi-1 untreated control sample 1. Expression of each transcript is from each sample are plotted against each other as x,y coordinates, and a line of best fit is drawn. Repeated figure from chapter 2, presented again here for ease of reader.

An example of linear regression is given in Figure 72 B) for 5AzaC treated Kasumi-1 biological repeat 2 (KA_5A_2) regressed against untreated Kasumi-1 biological repeat 1 (KA_C_1). As described in more detail in chapter 2, the regression fits a linear line describing the relationship between the two data sets as Y=mx+b. This slope is used to predict expression of transcripts of variable y (KA_5A_2) based on variable x (KA_C_1). The residual output is calculated for all observations and converted to standard residual (SR) output. The larger the standard residual reported for an observation, the more it has deviated from the expected trend.

To determine a cut-off point; where an observation has deviated by enough SR to be considered of interest, the distribution of SR as absolute values was plotted for all regressions. Examples of these plots are given in Figure 73. By plotting from highest to lowest SR, it can be visualised that about where observations noticeably differ from predicted (red dotted line).
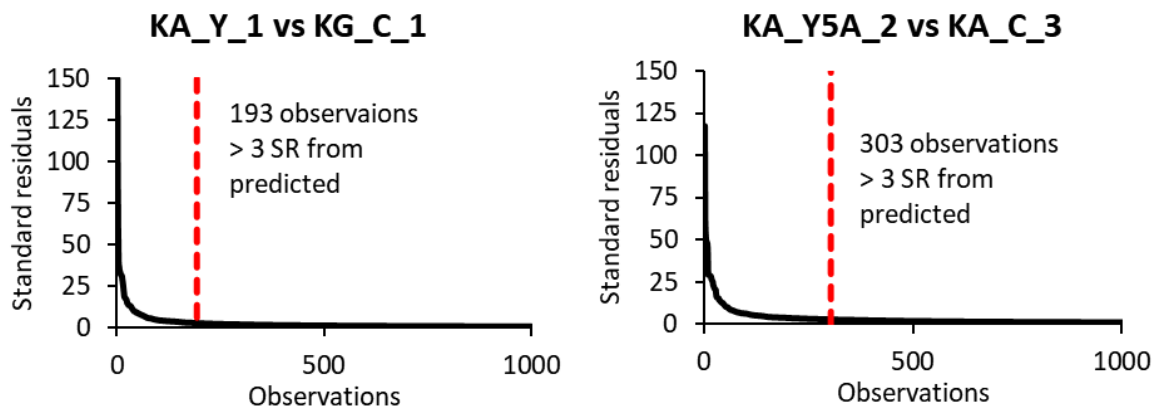
Figure 73: Two examples of regression analysis, with the top 1000 largest standardised residuals (SR) plotted for each comparison. Representative for all data sets, around 200 to 300 of residuals compared this way output values 3 or more SR away from their predicted outcomes. Left: Kasumi-1 IFNG treated sample 1 vs KG-1 untreated sample 1, Right: Kasumi-1 IFNG and 5AzaC treated sample 2 vs Kasiumi1 untreated sample 3.

5.2.2 Step 2: Shortlisting candidates by frequency of appearances in top 500 lists

In comparisons performed 100 to 300 observations exceeded deviation of 3 SR. The top 500 observations with the largest SR for each regression were carried forward. The top 500 transcripts in for each treatment were compiled into lists and duplicates removed, producing lists of 1411, 1176 and 1162 transcripts for 5AzaC, IFNG and IFNG and 5AzaC treatment respectively. To narrow down the lists further, only transcripts that appeared in the top 500 list of all 36 pairwise comparisons were selected for further investigation. This left lists of 15, 55 and 56 genes respectively, displayed in Table 16 for each treatment type. As IFNG and 5AzaC treatments were combined in the final treatment type there is overlap between lists. The IFNG & 5AzaC list shares 38 transcripts in common with the IFNG only list, and 4 with the 5AzaC only list, while 3 transcripts are present in every list and 11 are unique to the combined treatment. This has been visualised by venn diagram (See appendix 7.7), with a table listing the overlapping transcripts for each treatment condition.

5.2.3 Step 3: Removal of transcripts not in the TCGA database

Prior to performing CPH univariate analysis, transcript lists generated from pairwise linear regression were searched for in the TCGA database. Not all transcripts were present, these were mostly mitochondrial transcripts that were lost and in one case, a novel transcript (AL713998.1)

was found with no entrez entry number to search for currently set. The 5AzaC list was shortened from 15 to 8, IFNG was shortened from 55 to 50 and IFNG & 5AzaC was shortened from 56 to 50. Candidates that were removed are highlighted in red in Table 16.

Table 16:Lists of transcripts after pairwise linear regression and ranking by number of appearances in the top 500 of all 36 comparisons for every treatment type. Transcripts shown in red in table are mitochondrial and ribosomal transcripts which are routinely filtered out from data as a quality control step, they were also unavailable in data sets for the same reason (Subramanian, et al. 2022).

5AzaC

MT-CO1	MT-CO2	MT-ND4L	CALR	RPL41P1	MT-RNR1	YBX1	GAPDH	MT-TN	NPM1
MT-ND3	TUBB	LGALS1	S100A11	BST2					

IFNG

IFI6	EEF1A1	TMSB10	FTL	IFITM1	HLA-B	TMSB4X	IFIT3	HLA-E	STAT1
MT-CO2	MT-CO3	LY6E	RPL8	B2M	RPS2	SRGN	MT-ATP6	BST2	WARS
ACTB	GBP1	AL713998.1	RPL7	HLA-A	LAP3	UBE2L6	TAP1	HLA-C	GAPDH
PARP10	GBP5	UBA7	PSME1	NPM1	VAMP8	GBP4	PABPC1	GPX1	PSMB9
RSAD2	PLEK	GLUL	GBP2	PSMB8	ADAR	HLA-DRA	PRDX1	NUCB1	EPSTI1
CALHM6/FAM26F	PSME2	GSTP1	ACTG1	PIM1					

IFNG & 5AzaC

IFI6	TMSB10	IFITM1	PABPC1	HLA-B	IFIT3	RPL7	MT-CO2	HLA-DRA	HLA-E
RPL8	LY6E	RPL41P1	STAT1	ACTB	MT-ND4	MT-RNR2	MT-ND3	GAPDH	WARS
GBP1	BST2	CLEC11A	B2M	EEF1A1	AL713998.1	LAP3	GPX1	SRGN	GBP5
UBE2L6	TAP1	GSTP1	HLA-A	PARP10	TRIM22	GBP4	CHCHD2	S100A11	UBA7
PSME1	PSMB9	HLA-C	DTX3L	GBP2	PLEK	NCOA4	PIM1	RSAD2	CALR
EPSTI1	APOL6	ATF4	PSMB8	CD74	SPI1				

5.2.4 Step 4: Cox proportional hazard regression (CPH) to build PI

The CPH method is described in more detail in the chapter 2 but will be explained here in short. CPH is a semi-parametric method which is used to verify a relationship between a variable, termed a predictor, and an event, such as death, by calculating hazard ratio (HR) over time, between two or more groups.

The hazard function is the statistical likelihood that at a given time point an event will occur given that up to that time point, said event has not occurred. The HR is the ratio of an event happening between groups being compared. For example, when looking at survival of patients based on high or low expression of a gene signature, a ratio of 0.5 would tell us that half the number of patients with high expression of the signature experiences an event compared to those in the low expression group. In this thesis, CPH is used to determine if expression of transcripts selected thus far significantly relate to AML patient overall survival using CPH univariate analysis, and then to build a signature by CPH forward selection model.

5.3 The TCGA patient data set

The Cancer Genome Atlas (TCGA) program is a joint venture between the National Cancer Institute (NCI) and National Human Genome Research Institute (NHI). The program has sequenced and molecularly characterised 33 cancer types generating genomic, transcriptomic, epigenomic and proteomic data bases. These data bases are public domain for anyone in research to access and use. For this thesis, the Acute Myeloid Leukaemia TCGA (Cancer Genome Atlas Research Network 2013) RNAseq database (with data from 179 samples) was downloaded from cbioportal and used to generate initial signatures for testing. This data included clinical data for patients, such as overall survival time, expression of 20,531 mRNA transcripts, disease reoccurrence and induction therapy used and more.

5.4 CPH univariate determined significance of individual transcripts

To investigate the relationship between each transcript in the lists generated from cell lines and overall survival, CPH univariate analysis was performed using the TCGA data base in Tibco Statistica software. Each transcript that was reported as significant is listed in Table 17 below, the full set of results for each list is available in the appendix (see 7.8).

Table 17: Presents reported results from CPH univariate analysis of each transcript of interest and patient overall survival with P-value < 0.05, and thus deemed 'significant', a full list for all transcripts is available in the appendix (see 7.8).

	Transcripts	Parameter Estimate (B-value)	Standard Error	Chi-square	P value	Hazard Ratio
5AzaC list	CALR	-0.00002	0.00001	11.34799	0.00076	0.99998
	BST2	0.00043	0.00015	8.83214	0.00296	1.00043
	LGALS1	0.00007	0.00003	7.25487	0.00707	1.00007
IFNG list	PIM1	0.00014	0.00004	11.50584	0.00069	1.00014
	LY6E	0.00012	0.00004	10.69635	0.00107	1.00012
	PARP10	0.00027	0.00009	9.64581	0.00190	1.00027
	BST2	0.00043	0.00015	8.83214	0.00296	1.00043
	PRDX1	0.00014	0.00005	7.50531	0.00615	1.00014
	PSMB8	0.00052	0.00019	7.45553	0.00632	1.00052
	IFITM1	0.00024	0.00009	6.55459	0.01046	1.00024
	IFIT3	0.00047	0.00019	6.08997	0.01360	1.00047
	GSTP1	0.00012	0.00005	5.14428	0.02332	1.00012
	GPX1	0.00006	0.00003	4.81383	0.02823	1.00006
	ADAR	0.00009	0.00004	4.58753	0.03221	1.00009
	EPSTI1	0.00062	0.00030	4.33206	0.03740	1.00062
IFNG & 5AzaC list	CLEC11A	-0.00008	0.00002	11.52432	0.00069	0.99992
	PIM1	0.00014	0.00004	11.50584	0.00069	1.00014
	CALR	-0.00002	0.00001	11.34799	0.00076	0.99998
	LY6E	0.00012	0.00004	10.69635	0.00107	1.00012
	PARP10	0.00027	0.00009	9.64581	0.00190	1.00027
	BST2	0.00043	0.00015	8.83214	0.00296	1.00043
	PSMB8	0.00052	0.00019	7.45553	0.00632	1.00052
	IFITM1	0.00024	0.00009	6.55459	0.01046	1.00024
	IFIT3	0.00047	0.00019	6.08997	0.01360	1.00047
	GSTP1	0.00012	0.00005	5.14428	0.02332	1.00012
	GPX1	0.00006	0.00003	4.81383	0.02823	1.00006
	NCOA4	-0.00005	0.00002	4.42300	0.03546	0.99995
	EPSTI1	0.00062	0.00030	4.33206	0.03740	1.00062
	APOL6	0.00013	0.00006	4.04828	0.04422	1.00013

5.4.1 Interpretation of CPH univariate analysis results

The key results for univariate analysis are displayed in Table 17, with significance denoted by p-value, which is determined by the chi-squared test. The parameter estimates, also sometimes referred to as the β-value, shows the increase in the log(hazard) for each single unit of the variable being tested, in this case, each additional 1 unit of transcript expression. To make interpretation easier, the parameter estimate was transformed by exponentiation, and termed the hazard ratio

(HR). For example, in Table 17, 5AzaC CALR reports -2.3E-05 parameter estimate, when exponentiated becomes: exp(-2.3E-05) = 0.999977, which in turn can be expressed in percentage as 99%. Therefore, for every increase by 1 unit of CALR mRNA, there is ~ 0.0023% decrease in expected hazard, in this case death. Thus, an increase in CALR expression is suggested to be protective, while a decrease would increase risk of death. To generalise, for every increase of 1 unit of the variable being tested a HR:

-Above 1 indicates increased risk of hazard

-Below 1 indicates decreased risk of hazard

-Exactly 1 indicates no impact on risk of hazard

The trend for the list above is that for most candidates HR increases with increased expression of mRNA transcript, with the exceptions of CALR, CLEC11A and NCOA4, which all decrease. This suggested that in general, increased expression of the above lists indicates poorer survival. Univariate analysis only assessed a single risk factor at a time, whereas the outcome is likely influenced by multiple factors. While increased expression of LGALS1 indicated an increased risk of death, it could have a stronger predictor ability if combined with other variables, which univariate analysis had reported as insignificant. To identify a signature of multiple variables a CPH regression model was used.

5.4.2 Final signatures determined by CPH forward selection model

CPH regression models were generated using the full lists of transcripts of interest listed in Table 16, using Tibco Statistica and the forward selection setting with acceptance criteria set to P<0.05. As described in chapter 2 methods, this method started with no variables in the model, and added variables to the model one by one based on which was the most statistically significant, first in isolation, and then in combination with the variables thus far chosen for the model. This way a signature of multiple variables was built, with more predictive power than individual variables. The summary of the models built are shown in Table 18, while the full regression analysis can be found in the appendix (See appendix 7.9)

Table 18: CPH models created by forward selection for each treatment type (AzaC, IFNG and IFNG & 5AzaC)

		Summary of Forward Selection CPH models				
	Step	Parameter Estimate (β values)	Effect Entered	Score Chi-Square	P value	Hazard ratio
5AzaC list	1	0.00045	BST2	9.71333	0.00183	1.00045
	2	-0.00002	CALR	11.75729	0.00061	0.99998
	Step	Parameter Estimate (β values)	Effect Entered	Score Chi-Square	P value	Hazard ratio
IFNG	1	-0.00004	B2M	23.94877	0.00000	0.99996
	2	-0.00002	GAPDH	11.41136	0.00073	0.99998
	3	0.00033	IFIT3	1.65957	0.19766	1.00033
	4	0.00012	LY6E	4.89105	0.02700	1.00012
	5	0.00012	PIM1	5.99830	0.01432	1.00012
	6	0.00124	PSMB8	23.10665	0.00000	1.00124
	7	0.00001	SRGN	8.46569	0.00362	1.00001
	8	0.00010	TMSB10	11.91718	0.00056	1.00010
	Step	Parameter Estimate (β values)	Effect Entered	Score Chi-Square	P value	Hazard ratio
IFNG & 5AzaC	1	0.00047	BST2	6.40053	0.01141	1.00047
	2	-0.00001	CALR	1.84682	0.17415	0.99999
	3	-0.00008	CLEC11A	8.26981	0.00403	0.99992
	4	0.00011	LY6E	7.01487	0.00808	1.00011
	5	0.00010	PIM1	4.43433	0.03522	1.00010

5.4.3 Transcripts selected by the CPH model for PI scores were biologically relevant

As Prognostic scores were generated through purely statistical and data driven methods, variables reported for the models were checked for biological relevance, see Table 19.

Table 19: Summary of transcripts selected by CPH modelling and the biological relevance in context of AML and IFNG signalling or demethylation.

PI score	Transcript	Relevance
5AzaC, IFNG5AzaC	*BST2*	BST2 (Bone marrow stromal cell antigen) is expressed on the surface of HSCs, where it plays a part in relocalisation to the bone marrow through binding of E-selectin. It is known to be induced by IFNG (Florez, et al. 2020). Demethylation facilitated by RUNX1 overexpression has been linked to demethylation at CPG in the promoter sites for various HSC genes, including BST2, increasing its expression (Suzuki, et al. 2017).
5AzaC, IFNG5AzaC	*CALR*	CALR (Calreticulin) is a chaperon protein localised in the ER which relocates to the cell surface when exposed to stress, acting as an 'eat me' signal to phagocytes, bolstering immune recognition of AML cells (Chen, Xiufen, et al. 2013).
IFNG	*B2M*	B2M is a protein found in the heavy chain of MHCI on cell surfaces, it has been found to be prognostic for survival of patients with myelodysplastic syndrome and associated with risk of transformation into AML. IFNG upregulates expression of MHC on cell surfaces, and therefore B2M (Neumann, et al. 2009).
IFNG	*GAPDH*	GAPDH (Glyceraldehyde 3-phosphate dehydrogenase) has been found to be upregulated in response to interfere with IFNG translation and supress chronic signalling (Siska and Rathmell 2016).
IFNG	*IFIT3*	IFIT3 (Interferon Induced Protein with Tetratricopeptide Repeats 3) is an interferon stimulated gene (ISG) which is upregulated by IFNG via the jak-stat signalling pathway and involved in antiproliferative pathways (Wang, W., et al. 2017, Bhat, et al. 2018).
IFNG, IFNG5AzaC	*LY6E*	LY6E, a glycophosphatidylinositol-linked glycoprotein which is inducible by interferon, is displayed on the surfaces of numerous solid tumours, and has been found to be marker of hematopoietic stem cells in mice (Dela Cruz Chuh, et al. 2021, Virtaneva, et al. 2001).
IFNG, IFNG5AzaC	*PIM1*	Up-regulation of PIM1, a protooncogene and serine/threonine kinase, has long been known to be upregulated by IFNG signalling, has been observed in FLT3-ITD AML and implicated in FLT3-associated leukemogenesis (Yip-Schneider, Horie and Broxmeyer 1995, Fathi, et al. 2012).
IFNG	*PSMB8*	PSMB8 (proteasome 20S subunit beta 8) is a component of the immunoproteasome upregulated by IFNG in AML cells. PSMB8 inhibition triggers a build-up of polyubiquitinated proteins in AML cells with high concentration of immunoproteasomes, leading to cell death but not in low expressing AML cells (Rouette, et al. 2016, Niewerth, et al. 2014).
IFNG	*SRGN*	The SRGN gene encodes a hematopoietic cell granule proteoglycan, used to neutralise hydrolytic enzymes, and has been reported to be overexpressed in AML patients of the M1 and M2 FAB subtypes in comparison to healthy volunteers (Handschuh, et al. 2018)
IFNG	*TMSB10*	TMSB10 (Thymosin beta-10) encodes a protein that inhibits actin polymerisation and is highly expressed in AML progenitor cells (Wu, Junqing, et al. 2020).
IFNG5AzaC	*CLEC11A*	CLEC11A a growth factor protein for primitive hematopoietic progenitor cells. High expression of CLEC11A has been reported in leukaemia and as a favourable prognostic biomarker in (Yin, et al. 2021).

As summarised in Table 19, all transcripts selected by the CPH model made sense in the context of AML or in the presence of IFNG signalling or demethylation agents.

5.4.4 Derivation of prognostic indexes using CPH models

A prognostic index (PI) was generated for each set of transcripts generated from the CPH models in Table 18, following Blamey et als published formula (Blamey, et al. 2007). As per this method, expression of transcripts was normalised between 0 and 1 using the min-max approach, and then multiplied by the β values generated in the CPH models. In instances where there were multiple probes for the same transcript, expression was averaged across all probes and then normalised. A general formula for each PI was:

PI = (βT1 x NT1) + (βT2 x NT2) + (βT3 x NT3)

Where βT1 is the β value for T1 (Transcript 1) and NT1 was the normalised value of transcript 1 and so on for each transcript in the signature. This formula for generating a PI was previously used to successfully generate a 3 gene signature for AML, although starting with patient samples and using an advanced neural network approach for identifying the initial signature (Wagner, et al. 2019). As PI's generated were very small figures, and some negative, as a final step PI were multiplied by 10,000 for ease of use. Final formulas are in the appendix (see 7.11)

5.5 Results

5.5.1 PI scores contain transcripts capable of stratifying patients with short and long OS

First for every PI score, a Kaplan-Meier plot was generated for each individual transcript using the TCGA patient data set. This was performed using a median split to define high and low expression, with a log-rank (Mantel -cox) test to compare survival between groups. Results for individual transcripts for OS and EFS can be found in the appendix (see 7.9). Results reported that expression of transcripts in the same PI score varied in significance and hazard ratio. For example, in Appendix 7.12.4 high expression of *PIM1* was associated with poor OS (P<0.05), while high expression of *CALR* was associated with be (P<0.01).

5.5.2 PI scores are strong in univariate analysis

PI scores were assessed by univariate and multivariate cox proportional hazard models and compared to other categorical and continuous variables for outcome in the TCGA data set. When evaluated with a univariate model, all PI scores were deemed significant indicators for OS, with HR >1 indicating higher expression of PIs to be associated with worse outcomes. When multivariate analysis was performed, only sex was significant (P<0.05).

Table 20: Variables available in the TCGA databased evaluated by univariate Cox proportional hazard models for overall survival included: Each PI generated, age at diagnosis, bone marrow blast %, peripheral blood (PB) blast %, ELN cytogenetic risk group (1 = good, 2 = intermediate, 3 = poor), FAB subtype, sex (female =1, male = 0) and white blood cell count (WBC). Hazard ratios are given along with 95% confidence interval (CI), statistically significant p-values are in bold text.

Multivariate Cox proportional hazard (OS)				Univariate Cox proportional hazard (OS)			
Variable	P value	Hazard ratio	HR 95% CI	Variable	P value	Hazard ratio	HR 95% CI
5AzaC PI	0.2119	0.4083	0.09978 to 1.660	5AzaC PI	**0.0008**	1.401	1.144 to 1.698
IFNG PI	0.1026	0.8741	0.7393 to 1.022	IFNG PI	**0.0007**	1.117	1.046 to 1.189
IFNG5AzaC PI	0.0678	3.478	0.8944 to 12.98	IFNG5AzaC PI	**<0.0001**	1.526	1.281 to 1.804
Diagnosis Age	0.1146	1.02	0.9952 to 1.046	Diagnosis Age	**<0.0001**	1.036	1.022 to 1.052
Bone Marrow Blast Percentage	0.6901	0.997	0.9825 to 1.012	Bone Marrow Blast Percentage	0.9798	0.9999	0.9902 to 1.010
FAB	0.8536	0.9722	0.7146 to 1.302	FAB	0.4846	1.043	0.9263 to 1.171
PB Blast Percentage	0.3291	1.008	0.9923 to 1.023	PB Blast Percentage	0.5543	1.002	0.9957 to 1.008
Risk (Cyto)	0.819	1.058	0.6404 to 1.697	Risk (Cyto)	**<0.0001**	2.051	1.514 to 2.785
Sex	**0.0374**	1.98	1.042 to 3.795	Sex	0.9132	1.021	0.6965 to 1.493
WBC	0.5097	1.003	0.9928 to 1.013	WBC	**0.0208**	1.005	1.000 to 1.008

5.5.2.1 Kaplan-Meier analysis was performed using x-tile software to calculate optimal cut offs

To determine the PI scores cut-offs that could provide the best OS log-rank p-value, the software X-tile, a package developed at Yale school of medicine, was used. This software identifies the optimal cut-off point within a given data set for the PI to split patient groups into subgroups with significant differences in survival, and works across different gene expression platforms, making it suitable for use in all data sets used for this study. More information about this software can be found here: (https://medicine.yale.edu/lab/rimm/research/software.aspx)(Camp, Dolled-Filhart and Rimm 2004).

5.5.3 PI scores demonstrated strong prognostic ability in discovery set (TCGA)

All PI scores generated were reported to stratify patients in the TCGA data set when Kaplan-Meier's were performed (Figure 74). Survival curves were compared using a log-rank (Mantel-Cox) test. When PI were applied to Kaplan-Meier analysis of OS and EFS of TCGA patients, all PI scores were significant (minimum P<0.01) and hazard ratio were all reported above 2. In all cases, high PI expression was seen to be associated with poorer survival outcome. The most significant difference in survival time was seen for IFNG PI OS, where median survival times was four times longer for patients with low IFNG scores (46.5 months versus 11 months, PI < 0.0001).
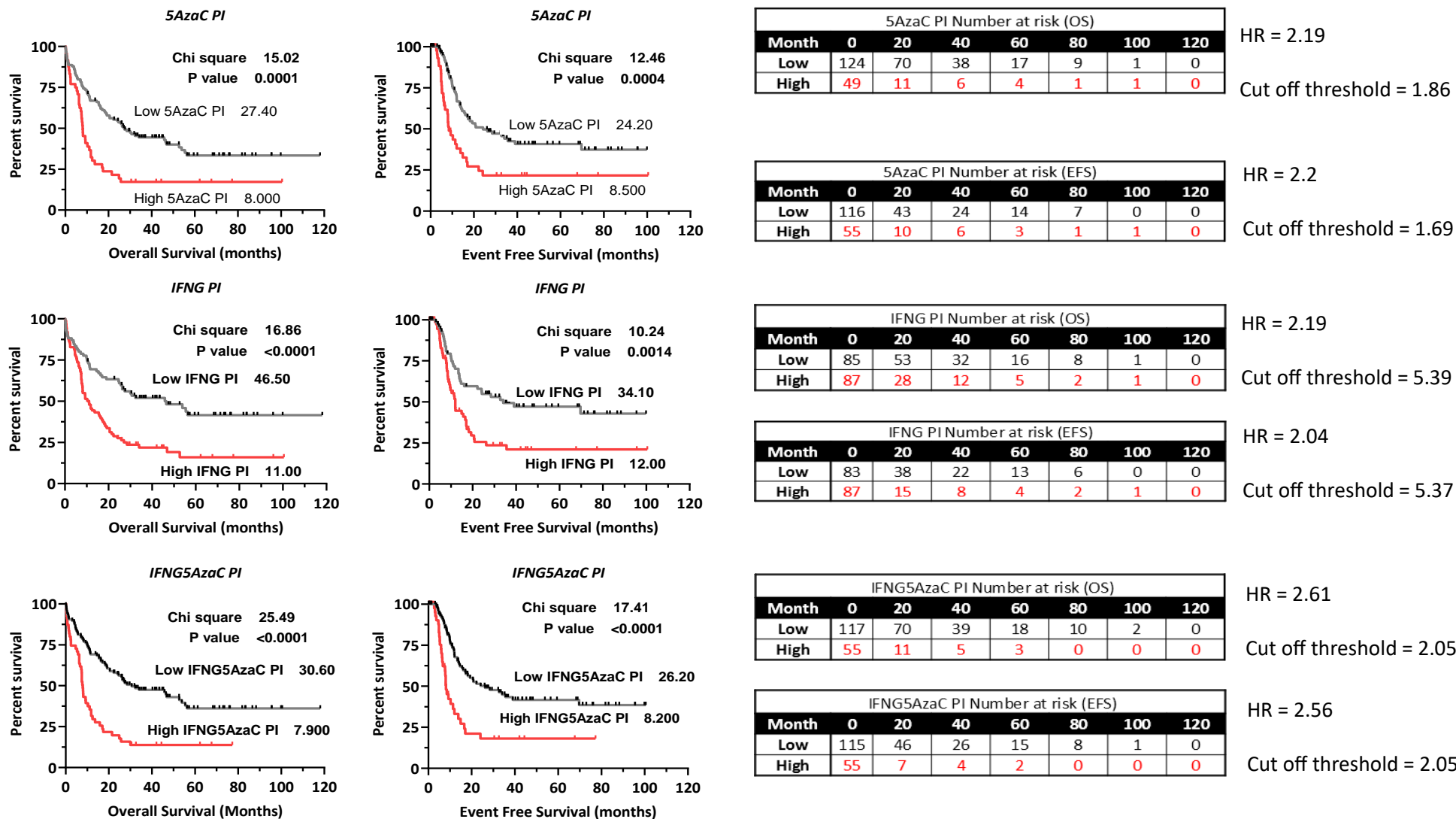
Figure 74: Kaplan-Meier estimates of OS and EFS for each prognostic index generated in the TCGA data set, with optimal cut off selected by X-tile software. Survival curves were compared using a log-rank (Mantel-Cox) test. Right of each set of graphs shows table with number of patients at risk at stated months. Median survival (months) is recorded next to each curve. Cut off threshold for each PI low and high stated, HR = Hazard ratio. Chi square and P-value are stated in the top right of each KM plot. HR = Hazard ratio.

5.5.4 High PI scores were significantly associated with age and poor ELN cytogenetic risk in the discovery series (TCGA)

Distribution of PI scores within subcategories of the TCGA patient database were examined using the optimum cut off thresholds calculated by x-tile. A high IFNG PI score was associated with increased age at diagnosis (P<0.05) but the other PI scores were not significantly different between patients above and below the age of 60 (Figure 75A). Additionally, high IFNG PI was significantly associated with patients who received chemotherapy, but not HSCT treatments (P<0.05, Figure 75B). Finally, the proportion of patients expressing high PI scores for 5AzaC, IFNG and IFNG5AzaC (P<0.00001, P<0.01 and P<0.01 respectively) were seen to significantly increase between patients in the good (5AzaC: 10%, IFNG: 20% and IFNG5AzaC: 10%), intermediate (5AzaC: 70%, IFNG: 43%, IFNG5AzaC: 66%) and poor (5AzaC: 56%, IFNG: 44% and IFNG5AzaC: 50%) ELN cytogenetic risk categories (Figure 75C). Note that the TCGA naming convention is different but refers to the same ELN risk categories, where favourable = good, intermediate remains intermediate and unfavourable = poor.
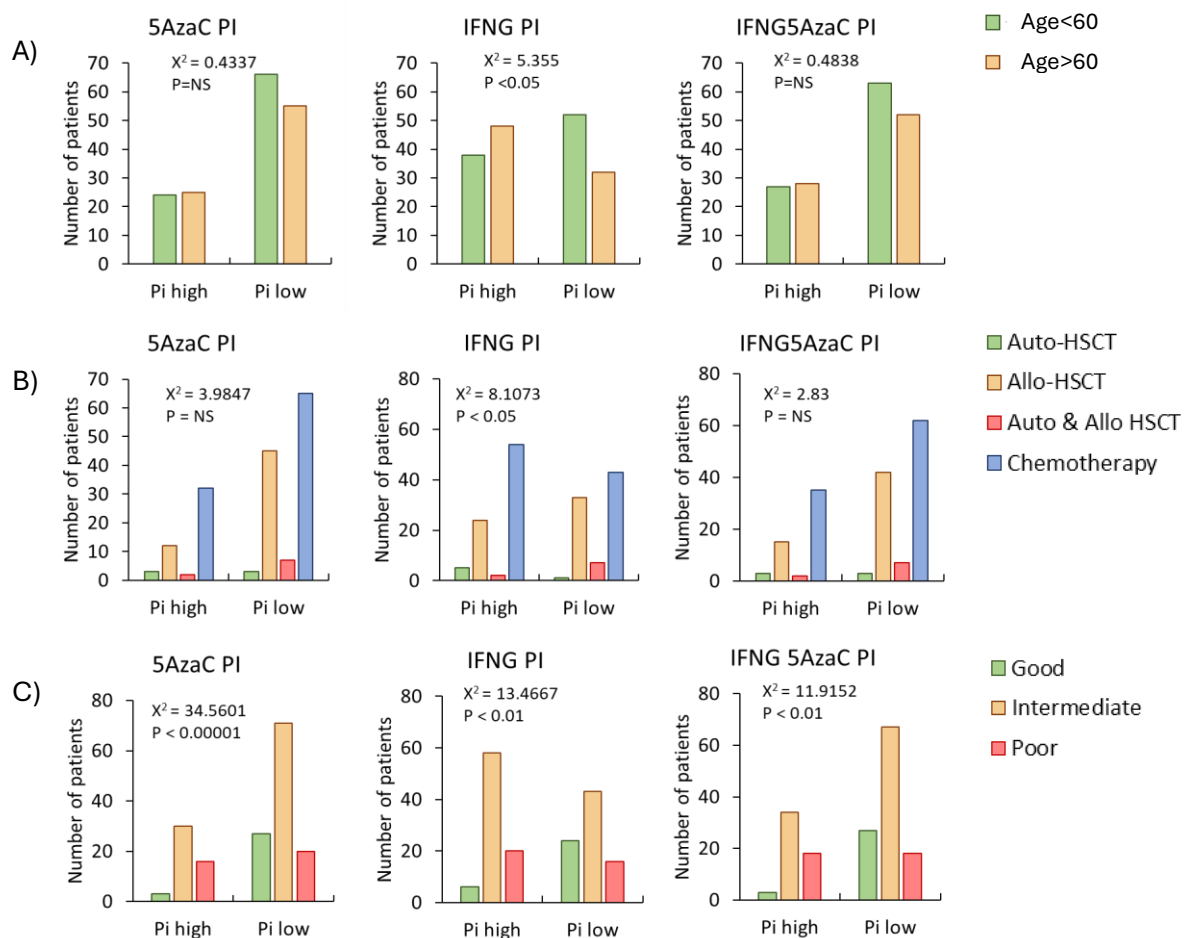
Figure 75: Distribution of PI scores in different AML groups. A) Distribution of AML patients with low and high PI ≥ 60 and < 60 years old. B) Distribution of AML patients with low and high PI according to treatment received, Auto HSCT, Allo HSCT, Auto and Allo HSCT or chemotherapy. C) Distribution of AML patients with low and high PI by ELN cytogenetic risk group; good, intermediate or poor. In all cases comparisons were carried out by the x^2 test. NS = Not significant.

Median PI scores increased with worsening ELN cytogenetic risk in the discovery series (TCGA)

As distribution of patients with high PI scores increased in intermediate and high-risk patients, PI scores were examined by violin plot. For each score, median PI expression showed a trend of increasing with worsening risk category, suggesting a relationship between PI scores and ELN cytogenetic risk. Significance increases in medium PI score were present between good and intermediate (5AzaC P<0.05, IFNG P<0.01, IFNG5AzaC P<0.01) and good to poor patients (5AzaC P<0.01, IFNG P<0.05, IFNG5AzaC P<0.0001).

	5AzaC PI			IFNG PI			IFNG5AzaC PI		
	Good	**Intermediate**	**Poor**	**Good**	**Intermediate**	**Poor**	**Good**	**Intermediate**	**Poor**
	1.072	1.524	1.675	4.601	5.977	5.996	1.127	1.786	2.193
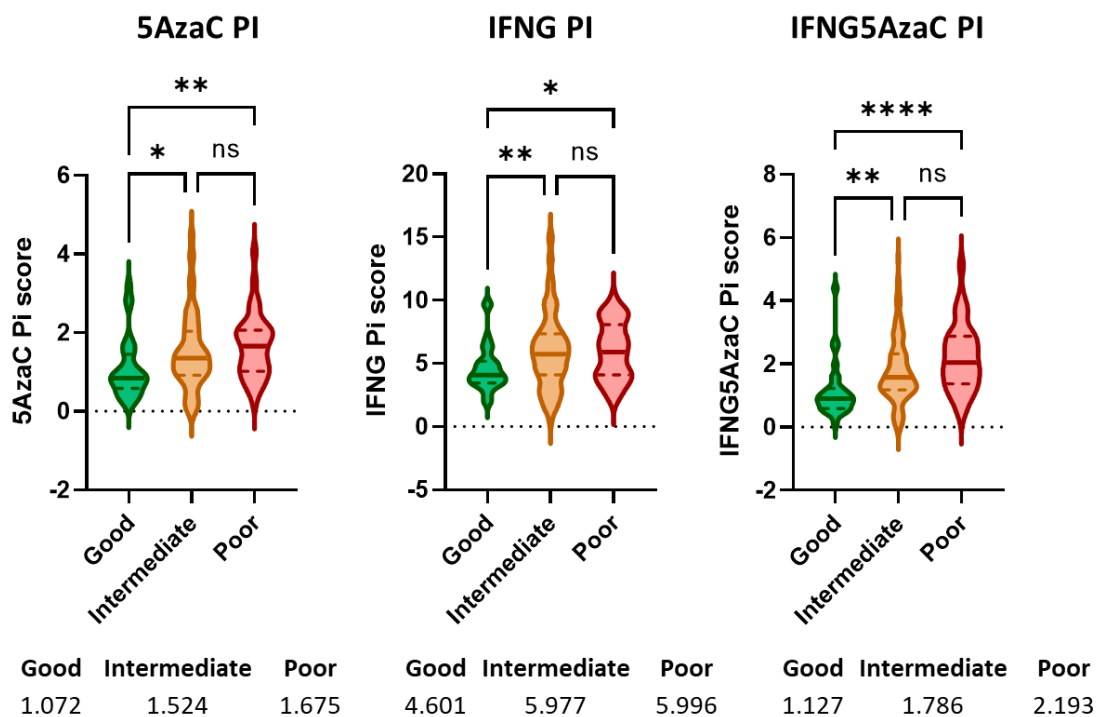
Figure 76: Violin plot of patient PI scores in the TCGA data set split by ELN cytogenetic risk group, left to right, 5AzaC, IFNG and IFNG5AzaC. Median is depicted by a solid line, upper and lower quartiles denoted by dashed lines. Significance calculated using Dunnett's T3 multiple comparison. Median PI scores for each risk category are tabulated below each graph.

The PI scores were next evaluated for their ability to stratify patients within the good, intermediate, and poor risk groups. As seen in Figure 77, patients stratified by ELN cytogenetic risk groups produced an excellent split as expected for OS (P<0.0001), however, and overlap between intermediate and poor risk patients is seen in EFS. The intermediate patients exhibit both good and poor outcome, and there is a need for an alternative method to identify which patients among the intermediate group will have poorer survival.
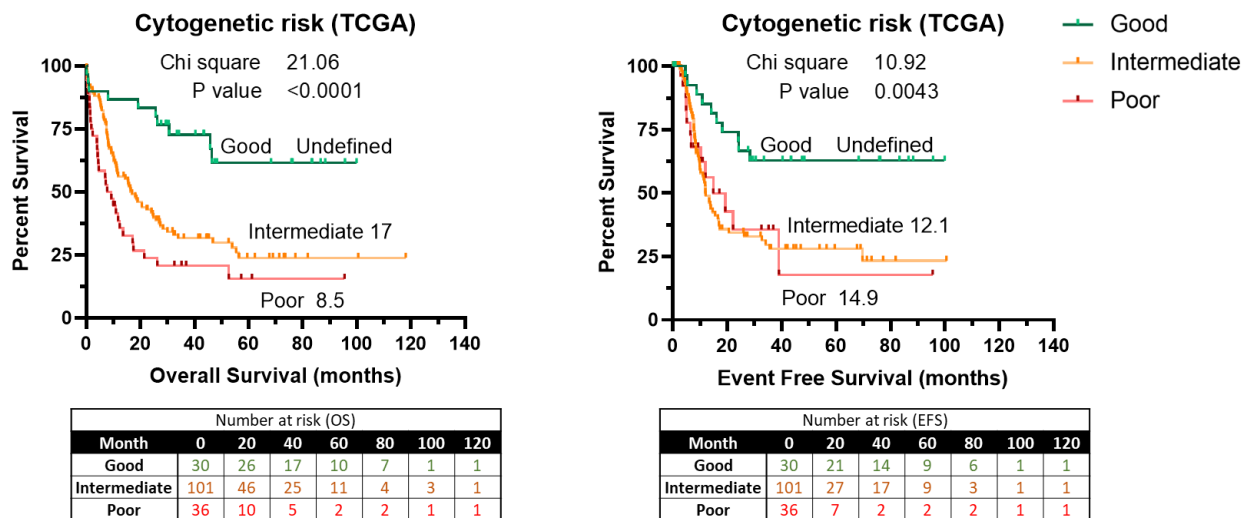
Figure 77: Kaplan-Meier estimates of OS and EFS for ELN cytogenetic risk categories generated in the TCGA data set, with optimal cut off selected by X-tile software. Survival curves were compared using a log-rank (Mantel-Cox) test. Under each graph is a table with number of patients at risk at stated months. Median survival (months) is recorded next to each curve. Chi square and P-value are stated in the top right of each KM plot.

5.5.4 All PI successfully stratify intermediate risk patients in the discovery series (TCGA)

Each PI score was applied to each set of patients in the good, intermediate, and poor risk categories. The results for the good and poor ELN cytogenetic risk patient groups are not reported as too few patients (<40 patients) were available overall, so the reliability of these analysis are not adequate to draw conclusions from. The intermediate risk group contained 101 patients, making for a more reliable data set to analyse.

All PI scores stratified intermediate patients with significance, where a high PI score was consistently significantly associated with poorer outcomes in all cases (Figure 78). Median survival was around three times as long in low-scoring PI patients for OS, and double for EFS. The biggest difference seen for median survival was for 5AzaC PI in OS (low score = 27 months, high score = 8.20 months, P<0.01). All PI scores significantly differentiated between intermediate patients with poorer and improved outcomes in the TCGA discovery series (P-scores reported in Figure 78).

5AzaC PI (OS)

Chi square 8.016
P value 0.0046

Low 5AzaC PI 27.00
High 5AzaC PI 8.20

OS (Months)

5AzaC PI (EFS)

Chi square 5.842
P value 0.0156

Low 5AzaC PI 14.90
High 5AzaC PI 8.20
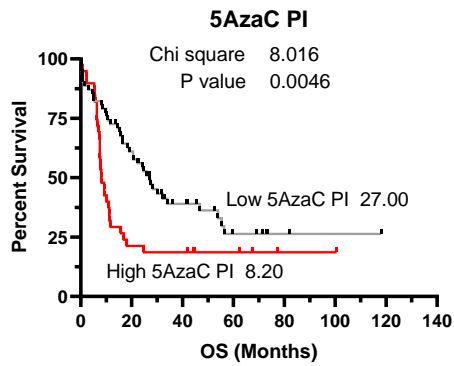
EFS (Months)

5AzaC PI Number at risk (OS)							
Month	0	20	40	60	80	100	120
Low	62	38	18	7	3	2	1
High	39	9	8	5	2	2	1

HR = 1.92

Cut off threshold = 1.51
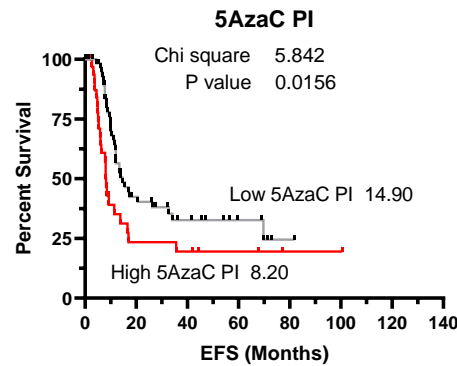
5AzaC PI Number at risk (EFS)							
Month	0	20	40	60	80	100	120
Low	66	21	12	6	2	1	1
High	35	7	6	4	2	2	1

HR =1.88

Cut off threshold = 7.28

IFNG PI (OS)

Chi square 6.382
P value 0.0115

Low IFNG PI 30.00
High IFNG PI 11.10

OS (Months)

IFNG PI (EFS)

Chi square 3.882
P value 0.0488

Low IFNG PI 13.40
High IFNG PI 9.30
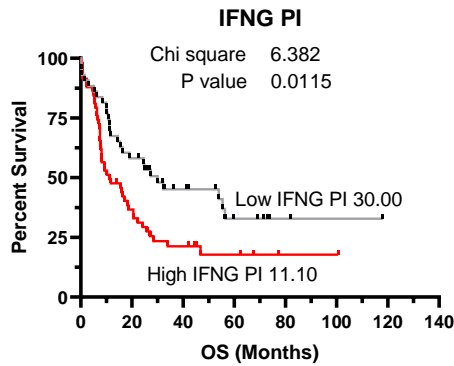
EFS (Months)

IFNG PI Number at risk (OS)							
Month	0	20	40	60	80	100	120
Low	43	26	15	7	3	2	1
High	58	21	11	5	2	2	1

HR = 1.83

Cut off threshold = 5.39
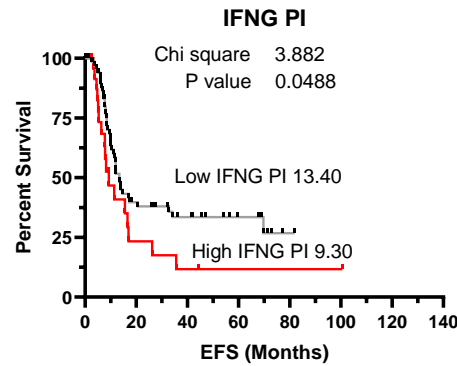
IFNG PI Number at risk (EFS)							
Month	0	20	40	60	80	100	120
Low	75	23	15	8	2	1	1
High	26	5	3	2	2	1	1

HR = 1.75

Cut off threshold = 7.28

IFNG5AzaC PI (OS)

Chi square 10.17
P value 0.0014

Low IFNG5AzaC PI 25.80
High IFNG5AzaC PI 7.90

OS (Months)

IFNG5AzaC PI (EFS)

Chi square 8.120
P value 0.0044

Low IFNG5AzaC PI 15.60
High IFNG 5AzaC PI 8.20
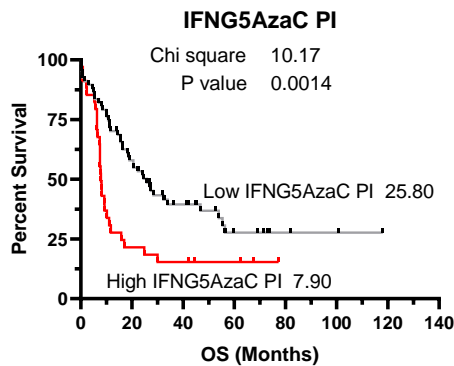
EFS (Months)

IFNG 5AzaC PI Number at risk (OS)							
Month	0	20	40	60	80	100	120
Low	67	39	20	8	4	3	1
High	34	8	6	3	1	1	1

HR = 2.11

Cut off threshold = 2.05
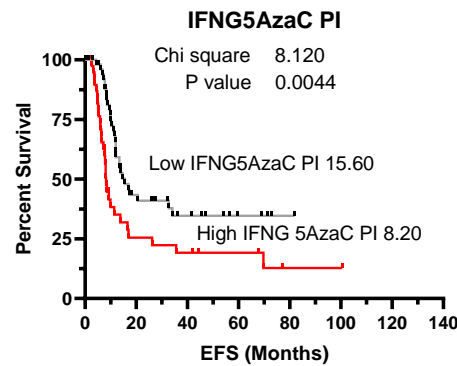
IFNG5AzAC PI Number at risk (EFS)							
Month	0	20	40	60	80	100	120
Low	58	19	11	5	2	1	1
High	43	9	7	5	2	2	1

HR =2.07

Cut off threshold = 1.72

Figure 78: Kaplan-Meier estimates of OS and EFS for patients in the intermediate ELN cytogenetic risk group of the TCGA data set, with optimal cut off selected by X-tile software. Survival curves were compared using a log-rank (Mantel-Cox) test. To the side of each graph is a table with number of patients at risk at stated months. Median survival (months) is recorded next to each curve. Chi square and P-value are stated in the top right of each KM plot. HR = Hazard ratio.

5.5.5 High PI scores strongly associated with poor outcome in chemotherapy treated patients in the discovery series (TCGA)

In the TCGA data set, 97 patients received induction chemotherapy and no HSCT treatment. When PI scores were applied, significant splits were observed for OS and EFS (Figure 79). In the case of EFS, the median survival of patients with low PI scores were undefined, meaning that more than 50% of patients were alive by the end of the study (5AzaC PI $P<0.001$, IFNG PI $P<0.0001$, and IFNG5AzaC PI $P<0.0001$). In addition to hazard ratios reported at 3.180, 4.341 and 4.646 for 5AzaC, IFNG and IFNG5AzaC PI respectively, these results indicated a strong relationship between PI scores and chemotherapy induction outcome.

5.5.6 High PI scores associated with poor outcome in patients receiving HSCT in the discovery series (TCGA)

Next, patients who were treated with any kind of HSCT (allo, auto, MUD -72 patients) were assessed for outcome using the PI scores. In this case, high 5AzaC and IFNG5AzaC PI scores were associated with significantly poorer outcomes in OS whereas only a high IFNG5AzaC PI score was significantly associated with poorer EFS (P-scores reported in Figure 80).

5AzaC PI (OS)

Chi square 10.97
P value 0.0009

Low 5AzaC PI 18.10

High 5AzaC PI 6.30

5AzaC PI (EFS)

Chi square 12.07
P value 0.0005

Low 5AzaC PI Undefined

High 5AzaC PI 8.200

IFNG PI (OS)

Chi square 11.58
P value 0.0007

Low IFNG PI 46.500

High IFNG PI 7.7000

IFNG PI (EFS)

Chi square 21.35
P value <0.0001

Low IFNG PI Undefined

High IFNG PI 8.2000

IFNG5AzaC PI (OS)

Chi square 15.60
P value <0.0001

Low IFNG5AzaC PI 24.955

High IFNG5AzaC PI 6.3000

IFNG5AzaC PI (EFS)

Chi square 19.17
P value <0.0001

Low IFNG5AzaC PI Undefined

High IFNG5AzaC PI 10.300

5AzaC PI Number at risk (OS)						
Month	0	20	40	60	80	100
Low	65	31	19	12	6	1
High	32	5	3	2	1	1

HR = 2.152

Cut off threshold = 1.86

5AzaC PI Number at risk (EFS)						
Month	0	20	40	60	80	100
Low	65	24	18	12	6	1
High	32	4	3	2	1	1

HR = 3.180

Cut off threshold = 1.86
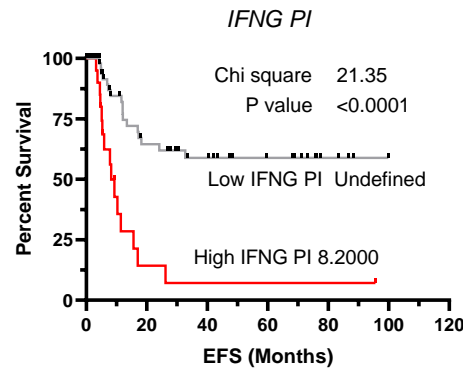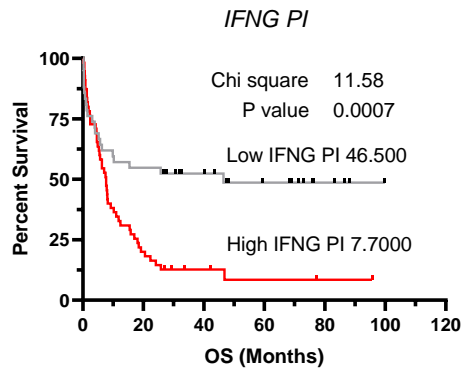
IFNG PI Number at risk (OS)						
Month	0	20	40	60	80	100
Low	42	24	18	11	5	1
High	55	12	5	3	2	1

HR = 2.317

Cut off threshold = 5.23

IFNG PI Number at risk (EFS)						
Month	0	20	40	60	80	100
Low	70	26	18	12	5	1
High	27	3	2	2	2	1

HR = 4.341

Cut off threshold = 7.16
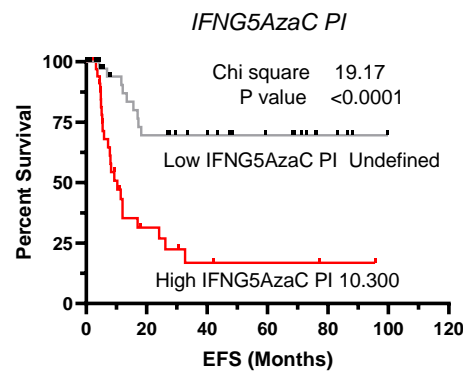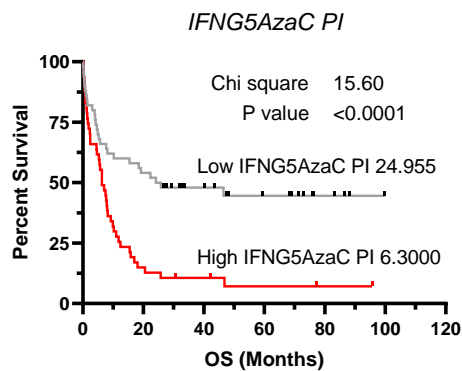
IFNG5AzaC PI Number at risk (OS)						
Month	0	20	40	60	80	100
Low	50	28	17	11	5	1
High	47	8	5	3	2	1

HR = 2.491

Cut off threshold = 1.64

IFNG5AzaC PI Number at risk (EFS)						
Month	0	20	40	60	80	100
Low	48	21	17	11	5	1
High	49	8	4	3	2	1

HR = 4.646

Cut off threshold = 1.58

Figure 79: Kaplan-Meier estimates of OS and EFS for patients receiving induction chemotherapy and no HSCT treatment in the TCGA data set, with optimal cut off selected by X-tile software. Survival curves were compared using a log-rank (Mantel-Cox) test. To the side of the graphs are tables with the number of patients at risk at stated months. Median survival (months) is recorded next to each curve. Chi square and P-value are stated in the top right of each KM plot. HR = Hazard ratio.

5AzaC PI

Percent Survival vs OS (Months)
Chi square 7.719
P value 0.0055
Low 5AzaC PI 52.70
High 5AzaC PI 6.30

5AzaC PI

Percent Survival vs EFS (Months)
Chi square 3.333
P value 0.0679
Low 5AzaC PI 16.10
High 5AzaC PI 8.500

IFNG PI

Percent Survival vs OS (Months)
Chi square 3.795
P value 0.0514
Low IFNG PI 86.000
High IFNG PI 26.300

IFNG PI

Percent Survival vs EFS (Months)
Chi square 3.688
P value 0.0548
Low IFNG PI 24.200
High IFNG PI 13.800

IFNG5AzaC PI

Percent Survival vs OS (Months)
Chi square 8.026
P value 0.0046
Low IFNG5AzaC PI 52.700
High IFNG5AzaC PI 13.60

IFNG5AzaC PI

Percent Survival vs EFS (Months)
Chi square 5.585
P value 0.0181
Low IFNG5AzaC PI 19.300
High IFNG5AzaC PI 8.0000
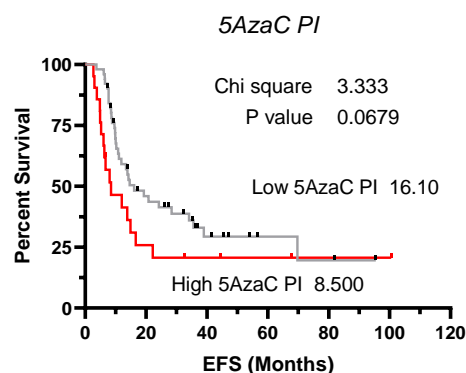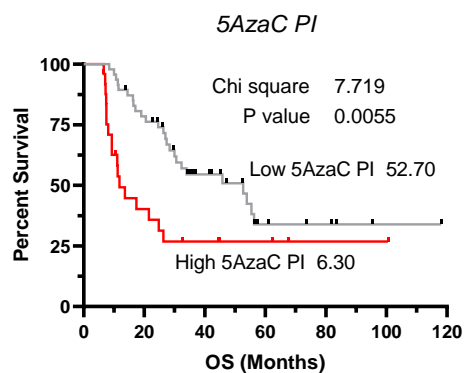
5AzaC PI Number at risk (OS)							
Month	0	20	40	60	80	100	120
Low	47	36	19	7	5	2	1
High	25	10	6	4	2	2	1

HR = 2.320

Cut off threshold = 1.53

5AzaC PI Number at risk (EFS)						
Month	0	20	40	60	80	100
Low	51	20	9	4	3	1
High	21	6	4	2	2	1

HR = 1.729

Cut off threshold = 1.69
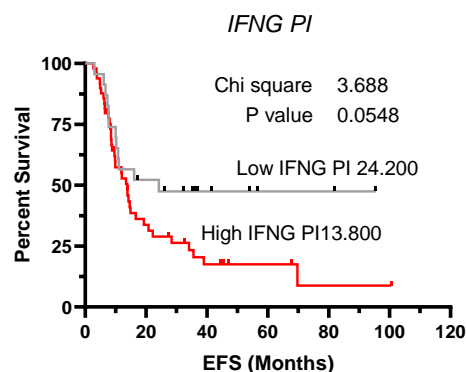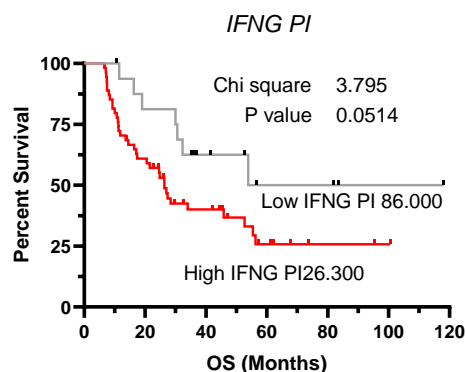
IFNG PI Number at risk (OS)							
Month	0	20	40	60	80	100	120
Low	17	14	8	4	4	2	1
High	55	33	17	7	3	2	1

HR = 2.186

Cut off threshold = 3.05

IFNG PI Number at risk (EFS)						
Month	0	20	40	60	80	100
Low	23	12	6	3	3	1
High	49	15	7	4	2	1

HR = 1.865

Cut off threshold = 4.075
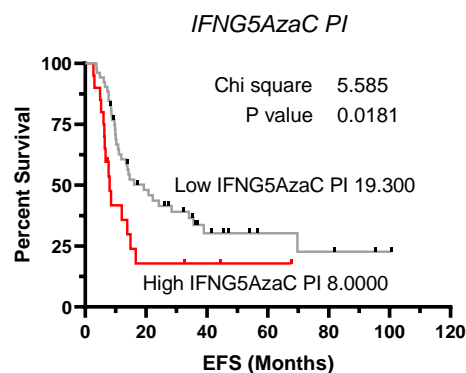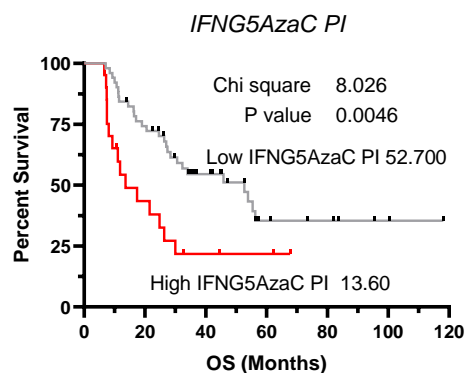
IFNG5AzaC PI Number at risk (OS)							
Month	0	20	40	60	80	100	120
Low	51	38	21	8	6	2	1
High	21	9	4	3	1	1	1

HR = 2.403

Cut off threshold = 2.04

IFNG5AzaC PI Number at risk (EFS)						
Month	0	20	40	60	80	100
Low	52	23	10	5	4	2
High	20	4	3	1	1	1

HR = 2.034

Cut off threshold = 2.05

Figure 80: Kaplan-Meier estimates of OS and EFS for patients receiving HSCT treatment alone or in conjunction with induction chemotherapy in the TCGA data set, with optimal cut off selected by X-tile software. Survival curves were compared using a log-rank (Mantel-Cox) test. To the side of the graphs are tables with the number of patients at risk at stated months. Median survival (months) is recorded next to each curve. Chi square and P-value are stated in the top right of each KM plot. HR = Hazard ratio.

5.5.7 PI scores show variable success in stratifying patients with molecular lesions in validation data sets (HOVON and BeatAML)

There were not enough patients with information for molecular lesions present in the TCGA data set for meaningful survival analysis; instead, PI were tested in the HOVON and BeatAML data sets. Categories of patients investigated included those with WT and mutated *NPM1*, WT *FLT3* and *FLT3-ITD* and finally *KMT2A* rearranged.

5.5.7.1 PI scores show variable success in stratifying patients with WT NPM1 and mutated NPM1 validation data sets (HOVON and BeatAML)

In the HOVON data set, patients with mutant *NPM1* could not be significantly stratified by any of the PI scores for OS or EFS (P>0.05, all, appendix see 7.14.1). When tested in the BeatAML data set, 5AzaC and IFNG5AzaC PI scores successfully stratified patients with mutated *NPM1* (OS available only, Figure 81), P<0.05. Interestingly, this analysis showed high 5AzaC and high IFNG5AzaC PI scores were associated with improved outcome for this subset of patients, the opposite to the trend shown in HOVON (7.14.1, Figure 81). Furthermore, the IFNG PI, although not significant, exhibits the trend of high PI score being associated with poorer outcome. It is unclear why the PI scores exhibit opposite trends, although not carried out here, a comparison of other clinical features or treatment schedules between the two subsets might reveal more.
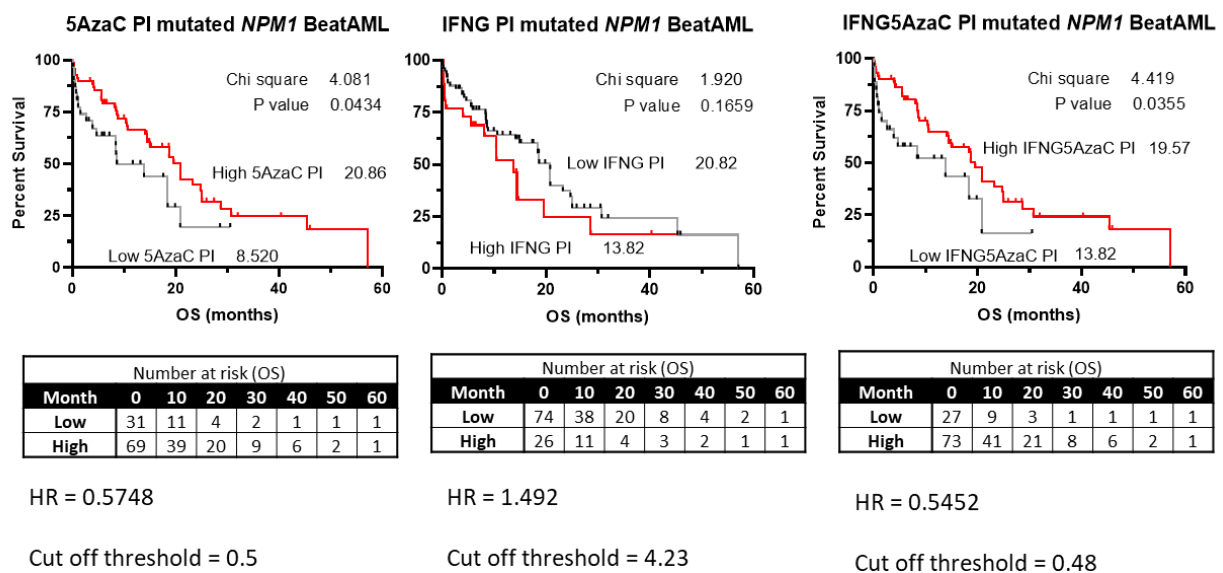
Figure 81: Kaplan-Meier estimates of OS of patients with mutated NPM1 in the BeatAML data set, with optimal cut off selected by X-tile software. Survival curves were compared using a log-rank (Mantel-Cox) test. Number of patients at risk at stated months below each graph. Median survival (months) is recorded next to each curve. Chi square and P-value are stated in the top right of each KM plot. HR = Hazard ratio.

When PI scores were applied to patients with WT *NPM1* in both data sets, contradicting outcomes were observed. In the BeatAML data set, only the IFNG PI score significantly stratified patients (P<0.001), with high score associated with poor outcome in OS only (Figure 82).

IFNG PI *NPM1* WT BeatAML

Chi square 13.96
P value 0.0002

Low IFNG PI 18.26
High IFNG PI 10.53

NPM1 WT Number at risk (OS)						
Month	0	40	80	120	160	200
Low	181	10	3	2	2	1
High	121	6	3	1	1	1

HR = 1.730

Cut off threshold = 4.62
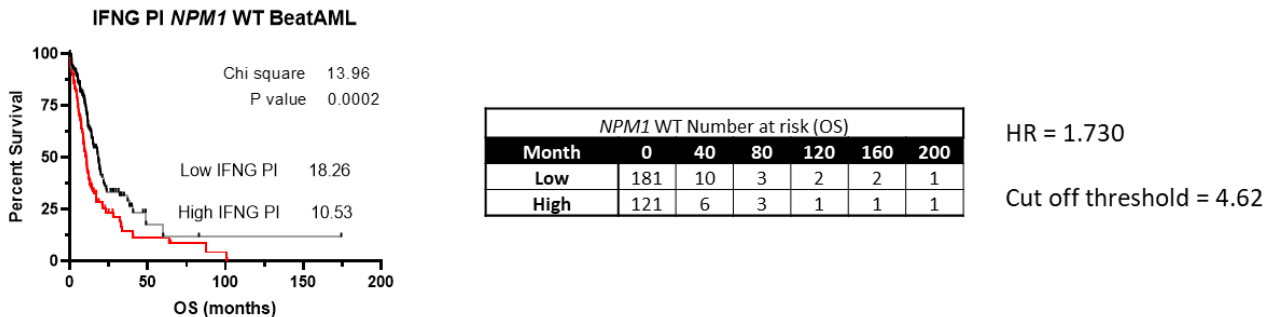
Figure 82: Kaplan-Meier estimates of OS of patients with WT NPM1 in the BeatAML data set, with optimal cut off selected by X-tile software. Survival curves were compared using a log-rank (Mantel-Cox) test. Number of patients at risk at stated months next to graph. Median survival (months) is recorded next to each curve. Chi square and P-value are stated in the top right of each KM plot. HR = Hazard ratio.

This contrasts with the HOVON data set, where IFNG had no significant stratification power (P>0.05) (See 7.14.2). In addition, 5AzaC and IFNG5AzaC were significant outcome predictors in the HOVON data set for WT *NPM1* patients OS and EFS (Figure 83), where they lacked any power in the BeatAML data set (See 7.14.3).

5AzaC PI Number at risk (OS)						
Month	0	40	80	120	160	200
Low	293	118	67	41	11	1
High	112	26	17	10	2	1

HR = 1.510, Cut off threshold = 2.98

5AzaC PI Number at risk (EFS)						
Month	0	40	80	120	160	200
Low	186	62	37	27	7	1
High	219	48	30	18	3	1

HR = 1.438, Cut off threshold = 2.66

IFNG5AzaC PI Number at risk (OS)						
Month	0	40	80	120	160	200
Low	285	118	71	47	12	1
High	120	25	13	5	1	1

HR = 1.716, Cut off threshold = 3.72

IFNG 5AzaC PI Number at risk (EFS)						
Month	0	40	80	120	160	200
Low	115	46	28	22	6	1
High	290	64	39	23	4	1

HR = 1.708, Cut off threshold = 2.94

IFNG5AzaC PI *NPM1* WT OS HOVON

Chi square 19.52
P value <0.0001

Low IFNG5AzaC PI 20.07
High IFNG5AzaC PI 12.39

Chi square 10.62
P value 0.0011

Low 5AzaC PI 19.55
High 5AzaC PI 10.76

IFNG5AzaC PI *NPM1* WT EFS HOVON

Chi square 16.37
P value <0.0001

Low IFNG5AzaC PI 13.96
High IFNG5AzaC PI 8.710

Chi square 10.12
P value 0.0015

Low 5AzaC PI 12.45
High 5AzaC PI 8.710

Figure 83:Kaplan-Meier estimates of OS of patients with mutated NPM1 in the HOVON data set, with optimal cut off selected by X-tile software. Survival curves were compared using a log-rank (Mantel-Cox) test. Number of patients at risk in tables next to graphs. Median survival (months) is recorded next to each curve. Chi square and P-value are stated in the top right of each KM plot. HR = Hazard ratio.

5.5.7.2 PI scores show variable success in stratifying patients with FLT3-ITD and WT FLT3 in validation data sets (HOVON and BeatAML)

In the HOVON data set, patients with *FLT3*-ITD could not be significantly stratified by any of the PI scores for OS or EFS (P>0.05, all, see 7.14.4).In the Beat-AML data set, patients with *FLT3*-ITD were best stratified by the 5AzaC and IFNG5AzaC PIs (P<0.01 and P<0.001 respectively). While IFNG PI did not report significance (P>0.05), it followed the same trend as the other PI scores where high expressions indicated improved OS outcome. This trend was flipped in the patients with WT *FLT3* where higher scores indicated poorer OS, however, only the IFNG PI scores was significant (P<0.0001). Kaplan-Meiers comparing PI scores in WT *FLT3* and *FLT3-ITD* can be seen in Figure 84.
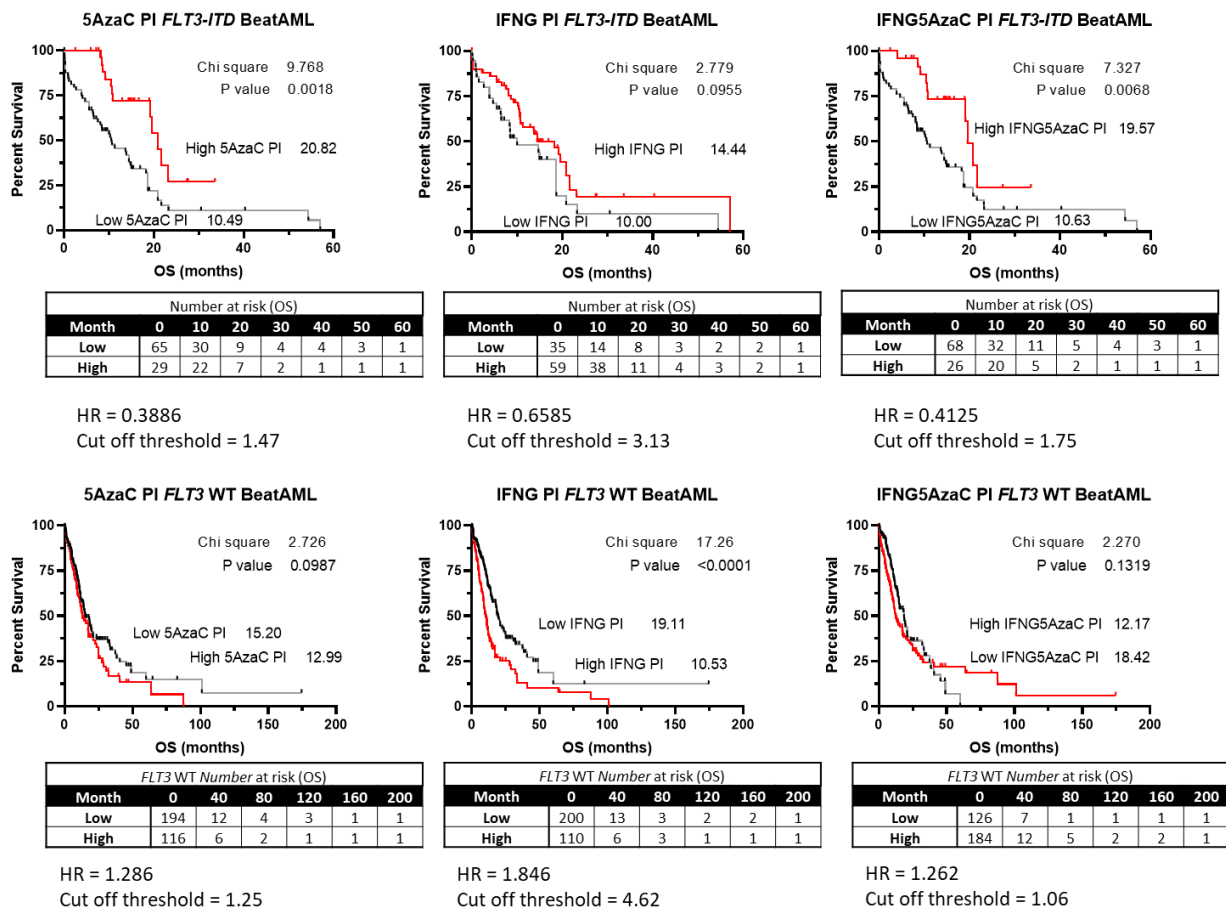
Figure 84: Kaplan-Meier estimates of OS of patients with FLT3-ITD in the BeatAML data set, with optimal cut off selected by X-tile software. Survival curves were compared using a log-rank (Mantel-Cox) test. Number of patients at risk in tables under graphs. Median survival (months) is recorded next to each curve. Chi square and P-value are stated in the top right of each KM plot. HR = Hazard ratio.

Conversely, when patients with WT *FLT3* were analysed with PI scores, only the 5AzaC and IFNG5AzaC PI scores produced significant stratification for OS and EFS (P<0.01, P<0.01 for 5AzaC PI OS ,EFS respectively and P<0.01 and P<0.001 for IFNG5AzaC PI OS and EFS respectively). High PI scores also indicated poorer outcome, in agreement with the trend shown in the BeatAML patients with WT *FLT3* (Figure 85).

5AzaC PI FLT3 WT OS HOVON

Chi square 7.365
P value 0.0066
Low 5AzaC PI 56.31
High 5AzaC PI 18.00

5AzaC PI FLT3 WT EFS HOVON

Chi square 7.111
P value 0.0077
Low 5AzaC PI 16.59
High 5AzaC PI 10.64

5AzaC PI Number at risk (OS)						
Month	0	40	80	120	160	200	240
Low	105	53	34	19	4	1	1
High	314	124	70	48	15	4	1

HR = 1.496, Cut off threshold = 2.15

5AzaC PI Number at risk (EFS)						
Month	0	40	80	120	160	200	240
Low	162	63	38	26	6	1	1
High	257	75	42	28	8	2	1

HR = 1.381. Cut off threshold = 2.48

IFNG5AzaC PI FLT3 WT OS HOVON

Chi square 9.323
P value 0.0023
Low IFNG5AzaC PI 51.71
High IFNG5AzaC PI 17.87

IFNG5AzaC PI FLT3 WT EFS HOVON

Chi square 10.99
P value 0.0009
Low IFNG5AzaC PI 17.81
High IFNG5AzaC PI 10.35

IFNG5AzaC PI Number at risk (OS)						
Month	0	40	80	120	160	200	240
Low	135	66	39	26	7	1	1
High	284	111	65	41	12	4	1

HR = 1.520, Cut off threshold = 2.94

IFNG5AzaC PI Number at risk (EFS)						
Month	0	40	80	120	160	200	240
Low	135	57	32	22	6	1	1
High	284	81	48	32	8	2	1

HR = 1.532, Cut off threshold = 2.94

Figure 85:Kaplan-Meier estimates of OS of patients with FLT3 WT in the HOVON data set, with optimal cut off selected by X-tile software. Survival curves were compared using a log-rank (Mantel-Cox) test. Number of patients at risk in tables under graphs. Median suvival (months) is recorded next to each curve. Chi square and P-value are stated in the top right of each KM plot. HR = Hazard ratio.

5.5.7.3 IFNG5AzaC PI score was a strong prognostic indicator in patients with KMT2A rearrangements

Next, patients classified as having a *KMT2A* rearrangement (also known as MLL translocation) were assessed in the HOVON data set. In all cases, high PI scores were associated with improved outcomes. IFNG and IFNG5AzaC PI scores showed high expression to significantly associate with improved median survival for OS and EFS (P<0.05 for IFNG OS and EFS, P<0.001 and p< 0.01 for IFNG5AzaC OS and EFS respectively, Figure 86).
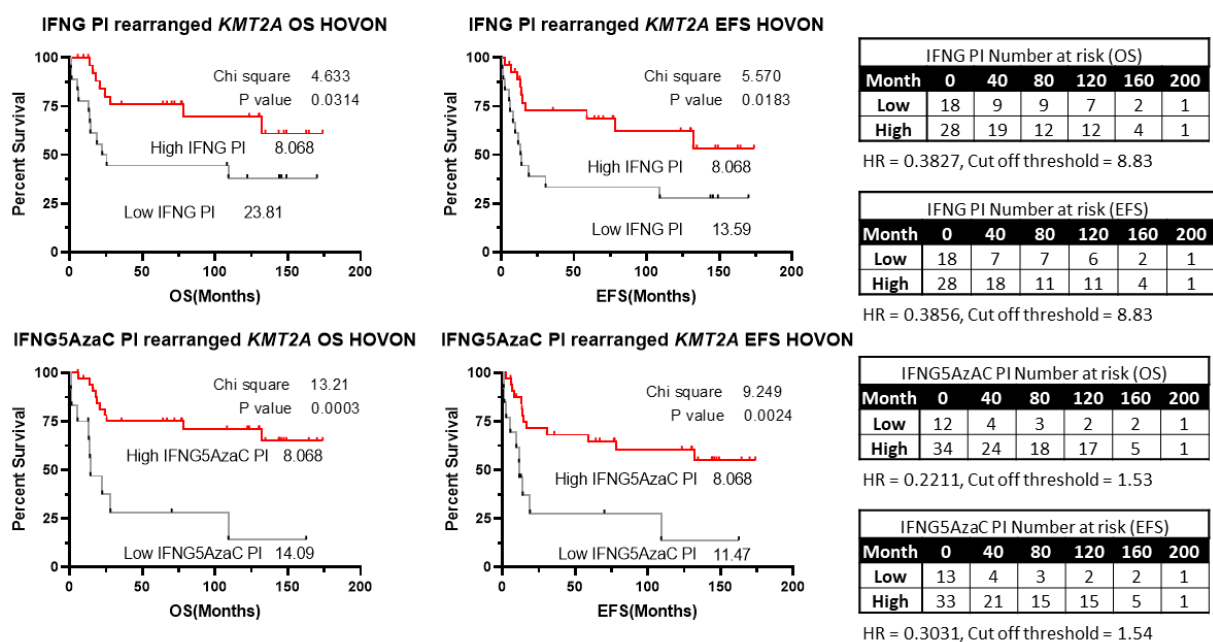
Figure 86: Kaplan-Meier estimates of OS and EFS of patients with KMT2A rearranged in the HOVON data set, with optimal cut off selected by X-tile software. Survival curves were compared using a log-rank (Mantel-Cox) test. Number of patients at risk in tables next to graphs. Median survival (months) is recorded next to each curve. Chi square and P-value are stated in the top right of each KM plot. HR = Hazard ratio.

5.5.8 PI scores stratify patients across validation data sets

Finally, PI scores were validated in the HOVON, BeatAML, German AML data sets and the German CN-AML subset. As expected, there was variation in performance of PI scores across data sets, but validation was achieved for at least one score in each data set by Kaplan-Meier. Figure 87 shows Kaplan-Meier plots which validated scores in the BeatAML and HOVON data sets. In Beat AML, the IFNG PI was the only significant score, with patients in the low PI score category achieving improved median OS (18.59 months vs 10.53 months, P<0.001).

In the HOVON data set, for all PI scores, a high score indicated a poorer outcome, see Figure 87. In contrast the BeatAML data set, IFNG PI score was the only score to not be reported as significant in stratifying patients, although only just (P=0.0552). In HOVON the 5AzaC PI and IFNG5AzaC PI scores significantly stratified patients into poor and improved outcomes for OS, (5AzaC PI low OS = 27.99 months vs 5AzaC PI high OS = 14.88 months, P< 0.01. IFNG5AzaC PI low OS = 22.31 months vs IFNG5AzaC PI high = 13.11 months, P< 0.001).

5AzaC PI Beat AML OS

P value 0.1018
Chi square 2.677

Low 5AzaC PI 13.95
High 5AzaC PI 19.57

Percent survival / OS (Months)
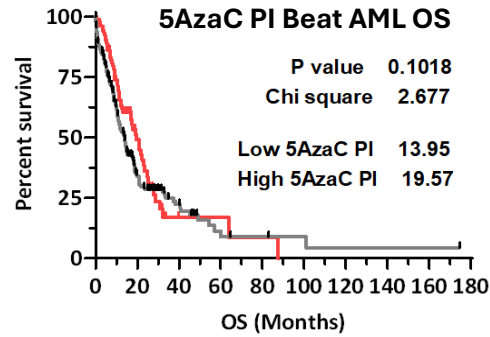
Month	0	20	40	60	80	100	120	140	160	180
Low	297	49	17	5	3	2	1	1	1	0
High	108	23	3	2	1	0	0	0	0	0

Number at risk (OS)

HR = 0.78
Cut off threshold = 1.46

IFNG PI Beat AML OS

P value 0.0003
Chi square 13.40

Low IFNG PI 18.59
High IFNG PI 10.53
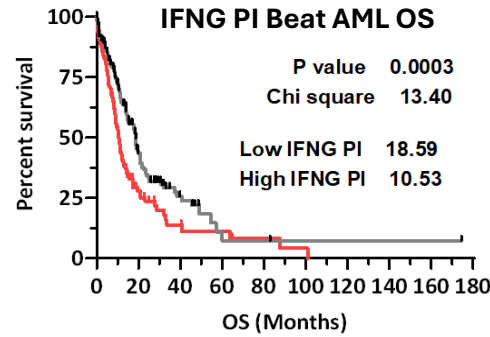
Month	0	20	40	60	80	100	120	140	160	180
Low	266	52	14	3	2	1	1	1	1	0
High	139	20	6	4	2	1	0	0	0	0

Number at risk (OS)

HR = 1.61
Cut off threshold = 4.62

IFNG5AzaC PI Beat AML OS

P value 0.3333
Chi square 0.9361

Low IFNG5AzaC PI 18.42
High IFNG5AzaC PI 12.70
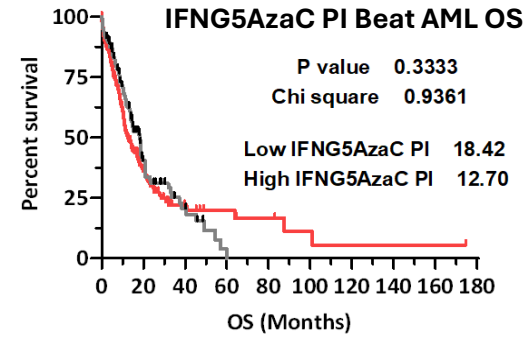
Month	0	20	40	60	80	100	120	140	160	180
Low	166	30	9	1	0	0	0	0	0	0
High	239	42	11	6	4	2	1	0	0	0

Number at risk (OS)

HR = 1.14
Cut off threshold = 1.05

5AzaC PI HOVON OS

Chi square 9.314
P value 0.0023

Low 5AzaC PI 27.99
High 5AzaC PI 14.88
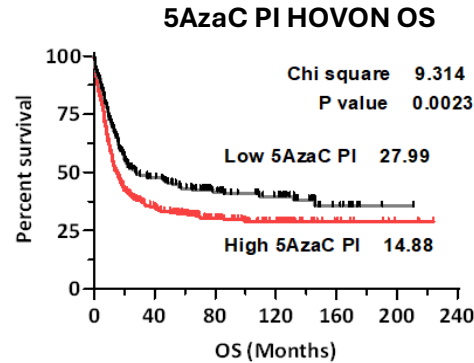
Month	0	20	40	60	80	100	120	140	160	180	200	220
Low	196	109	87	75	54	39	34	20	9	2	1	0
High	397	167	138	114	77	59	46	18	11	6	4	1

Number at risk (OS)

HR = 1.4
Cut off threshold = 2.39

IFNG PI HOVON OS

Chi square 3.677
P value 0.0552

Low IFNG PI 19.98
High IFNG PI 15.47
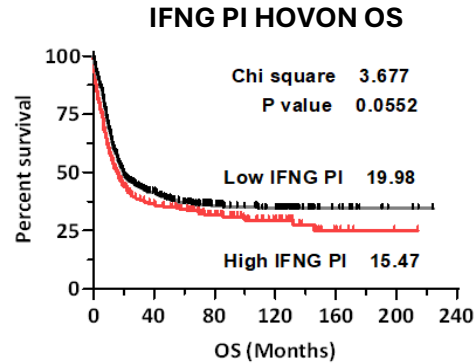
Month	0	20	40	60	80	100	120	140	160	180	200	220
Low	330	162	131	111	75	59	51	26	14	6	3	1
High	263	114	94	78	56	39	29	11	6	2	1	0

Number at risk (OS)

HR = 1.21
Cut off threshold = 10.15

IFNG5AzaC PI HOVON OS

Chi square 14.84
P value 0.0001

Low IFNG5AzaC PI 22.31
High IFNG5AaC PI 13.11
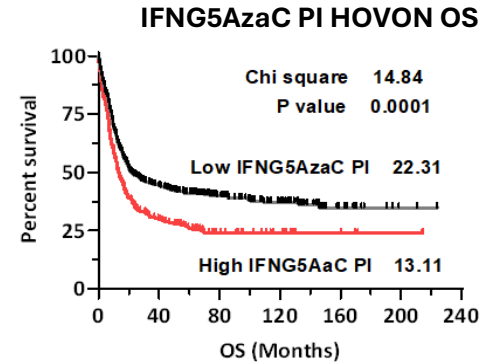
Month	0	20	40	60	80	100	120	140	160	180	200	220
Low	375	192	160	139	98	74	62	32	15	6	2	1
High	218	84	65	50	33	24	18	5	5	2	2	0

Number at risk (OS)

HR = 1.48
Cut off threshold = 3.64

Figure 87:Kaplan-Meier estimates of OS of patients in the BeatAML and the HOVON validation data sets, with optimal cut off selected by X-tile software. Survival curves were compared using a log-rank (Mantel-Cox) test. Number of patients at risk in tables under graphs. Median survival (months) is under the graphs. Chi square and P-value are stated in the top right of each KM plot. HR = Hazard ratio.

5AzaC PI German AML OS

Chi square 8.876
P value 0.0029

Percent survival vs OS (Months)

Low 5AzaC PI 13.97
High 5AzaC PI 8.515
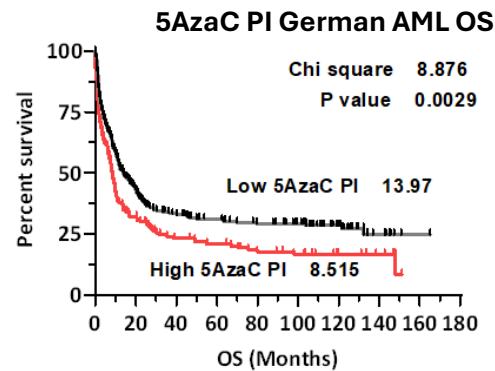
Number at risk (OS)>										
Month	0	20	40	60	80	100	120	140	160	180
Low	262	109	81	65	56	47	24	8	1	1
High	155	46	31	27	21	17	9	6	0	0

HR = 1.41
Cut off threshold = 2.88

IFNG PI German AML OS

Chi square 1.294
P value 0.2554

High IFNG PI 10.70
Low IFNG PI 11.20
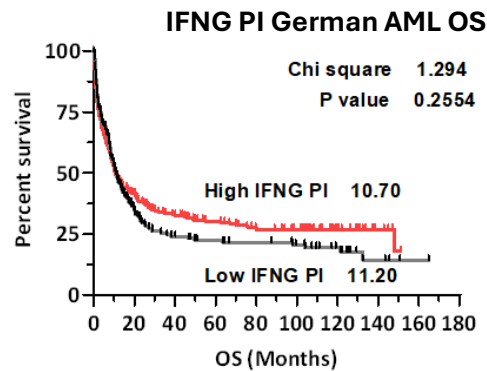
Number at risk (OS)										
Month	0	20	40	60	80	100	120	140	160	180
Low	151	50	34	27	23	19	12	5	1	1
High	266	105	78	65	54	45	21	10	0	0

HR = 0.88
Cut off threshold = 8.06

IFNG5AzaC PI German AML OS

Chi square 14.60
P value 0.0001

Low IFNG5AzaC PI 16.20
High IFNG5AaC PI 8.600
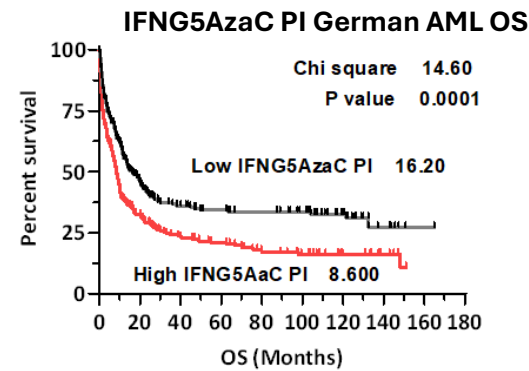
Number at risk (OS)										
Month	0	20	40	60	80	100	120	140	160	180
Low	199	88	66	53	49	42	22	5	2	2
High	218	67	46	39	29	22	12	10	1	1

HR = 1.55
Cut off threshold = 3.08

5AzaC PI CN AML OS

Chi square 6.312
P value 0.0120

Low 5AzaC PI 24.36
High 5AzaC PI 9.630
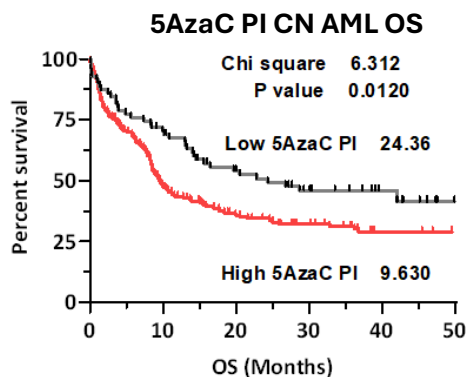
Number at risk (OS)						
Month	0	10	20	30	40	50
Low	71	49	34	26	10	0
High	171	73	49	37	3	0

HR = 1.6
Cut off threshold = 2.26

IFNG PI CN AML OS

Chi square 5.459
P value 0.0195

Low IFNG PI 25.87
High IFNG PI 9.633
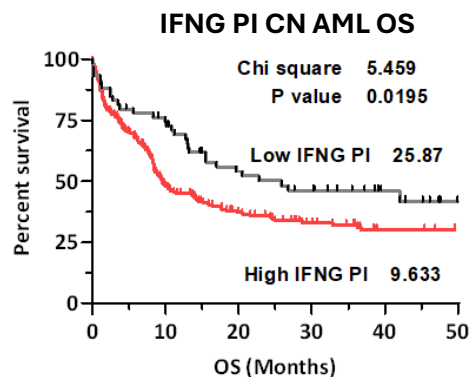
Number at risk (OS)						
Month	0	10	20	30	40	50
Low	59	42	28	23	10	0
High	183	80	55	40	3	0

HR = 1.6
Cut off threshold = 7.62

IFNG5AzaC PI CN AML OS

Chi square 3.387
P value 0.0657

Low IFNG5AzaC PI 15.55
High IFNG5AaC PI 8.646
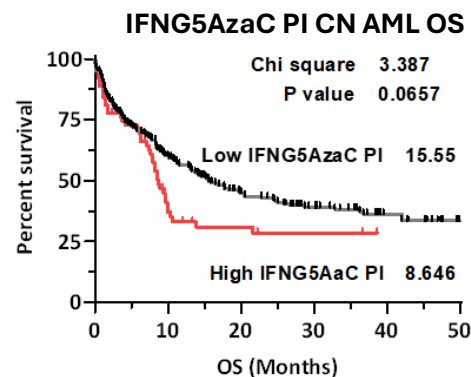
Number at risk (OS)						
Month	0	10	20	30	40	50
Low	179	103	70	52	13	0
High	63	19	13	11	0	0

HR = 1.4
Cut off threshold = 4.13

Figure 88: Kaplan-Meier estimates of OS of patients in the German AML and the CN-AML validation data sets, with optimal cut off selected by X-tile software. Survival curves were compared using a log-rank (Mantel-Cox) test. Number of patients at risk in tables under graphs. Median survival (months) is under the graphs. Chi square and P-value are stated in the top right of each KM plot. HR = Hazard ratio.

Next, validation was carried out in the German AML data set, and then PI scores were also tested in the cytogenetically normal subset of patients from that data set, termed CN-AML (Figure 88). CN-AML patients are considered intermediate risk, so this dataset was used to validate findings in the intermediate subset of the TCGA discovery series (Figure 78).

In the German AML data set, the 5AzaC PI and IFNG5AzaC PI scores were effective in predicting OS (Median survival, 5AzaC PI low OS = 13.97 months vs 5AzaC PI high OS = 8.515 months, $P < 0.01$. IFNG5AzaC PI low OS = 16.2 months vs IFNG5AzaC PI high = 8.6 months, $P < 0.001$). The IFNG PI score was ineffective in this data set at stratifying patients, with low score patients' median survival of 11.20 months vs high scoring patients 10.70 months ($P > 0.05$), see Figure 88.

In the CN-AML subset of patients, the 5AzaC and IFNG PI scores were successful in stratifying patient outcome for OS. (Median survival, 5AzaC PI low OS = 24.36 months vs 5AzaC PI high OS = 9.630 months, $P < 0.05$. IFNG PI low OS = 25.87 months vs IFNG PI high = 9.633 months, $P < 0.001$). The final IFNG5AzaC PI score was not significant but did show a similar trend of high scores associated with poor OS outcome. These results resemble those reported by the 86 probe set gene expression score generated from the same CN-AML dataset previously (OS, HR = 1.85; $P = 0.002$ and EFS, HR = 1.73; $P = 0.001$) (Metzeler, et al. 2008).

5.5.8.1 AUC curves of PI scores report similar predictive ability to established scores

Finally, all PI scores predictive ability for death was assessed using the area under receiver operator characteristics (AUC) curve in the TCGA data set. AUC reported for PI scores respectively were: 5AzaC PI AUC = 0.599, IFNG PI AUC = 0.637 and IFNG5AzaC PI AUC = 0.657, see Figure 89. The PI score's prognostic ability was compared to other established AML scores in the TCGA data set.
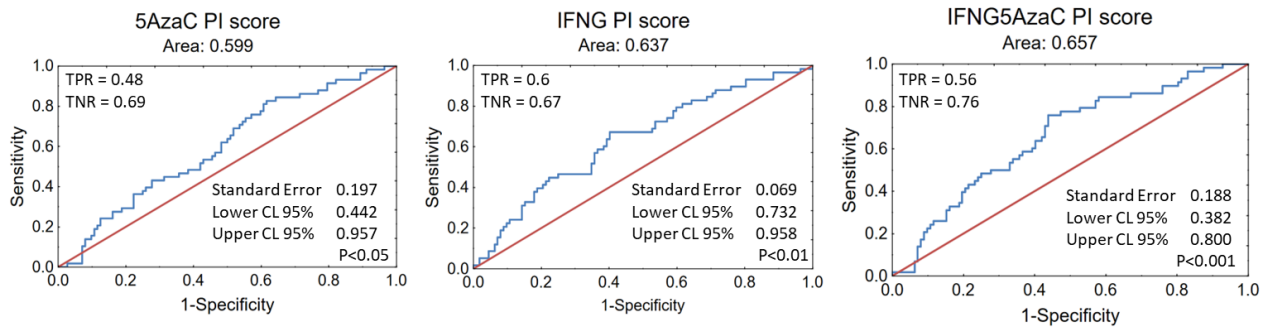
Figure 89: AUC curves quantify the ability of each PI score to predict outcome in individual patients by their specificity and sensitivity in the TCGA data set. TPR = True positive rate, TNR = True negative rate, Lower CL = Lower confidence limit, Upper CL = Upper confidence limit, Wald χ2 test used to compare curves. AUC = 1 denotes perfect prediction, AUC = 0.5 denotes no predictive ability above chance.

IFNG and IFNG5AzaC PI scores showed comparable performance to the LSC17 score in the TCGA data set (AUC = 0.65) (Ng, et al. 2016), which was generated using leukaemia stem cells, and the ELN cytogenetic risk categories (AUC = 0.66) (Wang, M., et al. 2018). Scores were outperformed by the 5 gene score (AUC = 0.74) developed by Sha et al by multivariate logistic regression model using clinical and expression data from the TCGA data set (Sha, et al. 2021). The Parsimonious 3 gene score by Wagner et al, generated by applying an artificial neural network (ANN) based machine learning approach to the HOVON data set, also proved stronger than the PI scores generated here (AUC = 0.71) (Wagner, et al. 2019).

5.6 Discussion

In this study, novel PI scores were generated from the transcriptomic data of the AML cell lines Kasumi-1 and KG-1 under different treatment conditions (IFNG, 5AzaC or IFNG and 5AzaC combined). Pairwise linear regression of treated to untreated cell lines revealed transcripts changed (indicated by standard residuals) and created a shortlist of candidates associated with IFNG signalling and demethylation individually and in combination. Candidates were selected that appeared in the top 500 list of all 36 pairwise comparisons for each treatment condition. These transcripts were then fed to a CPH forward selection model utilising expression and OS data in the TCGA patient set. PI scores were generated following the method outlined by blamey et al, weighting transcripts contribution to scores by the β-parameter reported in the CPH analysis (Blamey, et al. 2007).

The PI scores stratified patients for OS and EFS in the discovery series (TCGA). Then, the scores were validated in other data sets (BeatAML, HOVON, German-AML series and CN-AML subset), where success in patient stratification varied. Due to the heterogeneity of AML, it is unlikely that any single score would prove effective in stratifying patients in all data sets, therefore this result is expected. Comparison of prognostic strength of generated scores by AUC analysis (5AzaC PI AUC = 0.599, IFNG PI AUC = 0.637 and IFNG5AzaC AUC PI = 0.657) showed similar performance to the LSC17 signature (AUC = 0.65), and the established ELN cytogenetic risk categories (AUC = 0.66), but were outperformed by the parsominous-3 gene score by Wagner et al (AUC = 0.71), and the 5-gene score by Sha et al (AUC = 0.74).

In the discovery series, TCGA, analysis of the PI scores revealed that there was a significantly higher proportion of patients with high IFNG PI scores in patients over the age of 60 (P<0.05, Figure 75). Additionally, more patients who received only induction chemotherapy expressed high IFNG PI scores, compared to those receiving HSCT as part of their treatment course (P<0.05, Figure 75). It was also seen that the proportion of patients expressing high PI scores for 5AzaC, IFNG and IFNG5AzaC increased significantly between good, intermediate, and poor ELN cytogenetic risk categories (P<0.00001, P<0.01 and P<0.01 respectively, Figure 75). This was corroborated by plotting PI scores from patients in each risk category, where median expression of all PI scores rises with increasing risk groups (Figure 76).

Importantly, scores were effective in stratifying two groups of patients in need of improved outcome categorisation; patients treated with just chemotherapy, and patients in the intermediate risk category (Prada-Arismendy, Arroyave and Röthlisberger 2017).

High PI scores were strongly and significantly associated with poor outcome in chemotherapy treated patients in the TCGA data set for OS (5azaC: P<0.001, HR = 2.152, IFNG: P<0.001, HR = 2.317, IFNG5AzaC: P<0.0001, HR = 2.491). PI scores were particularly effective in stratifying patients who received chemotherapy induction treatment for EFS (5azaC: P<0.001, HR = 3.180; IFNG: P<0.0001, HR = 4.341; IFNG5AzaC: P<0.0001, HR = 4.646). Patients in the low PI score categories had a median survival that was undefined, indicating that more than 50% survived the duration of the study (Figure 79).

High PI scores were strongly and significantly associated with poor outcome in intermediate risk patients OS (5azaC: P<0.01, HR =1.92 , IFNG: P<0.05, HR = 1.83, IFNG5AzaC: P<0.01, HR = 2.11), and EFS (5AzaC: P<0.05, HR = 1.88, IFNG: P<0.05, HR = 1.75, IFNG5AzaC: P<0.01, HR = 2.07)(Figure 78). Patients with low PI scores were shown to have median OS between 2-3 times longer than patients with high PI scores, while median EFS was between 1.5 and 2 times longer in low scoring compared to high scoring patients. Overall, patients in the intermediate ELN cytogenetic risk group that also expressed low PI had improved OS and EFS compared to those with high PI scores. Importantly, a high PI predicted shorter OS, irrespective of whether patients received chemotherapy alone or chemotherapy followed by HSCT.

When scores were assessed by Kaplan-Meier tests in the HOVON, BeatAML, German-AML and CN-AML data sets, performance varied. The results of tests in the validation data sets are summarised in Figure 90 with hazard ratios (HR) and P-values plotted. The 5AzaC PI could be validated in the German-AML, CN-AML and HOVON data sets, IFNG PI was validated in only the BeatAML and CN-AML data sets and finally, IFNG5AzaC PI was significant in the German-AML and HOVON data set.
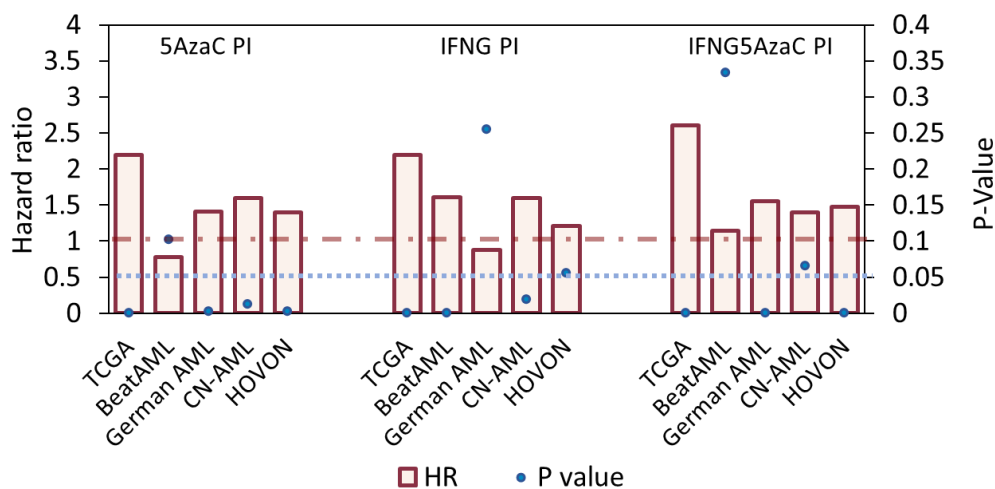
Figure 90: Summary of results of Kaplan-Meier tests across discovery series (TCGA) and validation data sets (BeatAML, German AML, CN-AML and HOVON). Hazard ratios (HR) are given as bar charts, significance as calculated by log rank (Mantel-Cox) test is denoted as blue circles (P-value).

Importantly, the effectiveness of the 5AzaC and IFNG PI scores seen in the TCGA data set could be validated in the CN-AML data set (Figure 88) (5AzaC, HR = 1.6, P < 0.05 , IFNG, HR = 1.6 P < 0.05). The 5AzaC and IFNG PI scores improve on the 86-probe set generated by Metzeler et al by consisting of only 2 or 8 transcripts; a more manageable number of targets to test for in clinical practice (Metzeler, et al. 2008).

In all cases where scores were validated, hazard ratios were >1, indicating increased risk with rising expression. Suggesting a negative link between 5AzaC demethylation patterns and IFNG signalling with survival in generalised populations. The pattern of high PI scores associated with poorer outcomes was only disrupted when scores were applied to patients with specific mutations. The discovery series did not have enough examples of mutated patients to examine for meaningful conclusions, so the HOVON and BeatAML data sets were investigated instead.

PI scores achieved no stratification for OS in the HOVON data set for patients with mutated *NPM1*. In the BeatAML data set the 5AzaC and IFNG5AzaC scores proved more effective for patients with mutated *NPM1* (both P<0.05). Interestingly, there is an inverse relationship with the scores in these patients, where unlike in the general populations, high PI scores were associated with improved outcomes by the 5AzaC and IFNG5AzaC scores, HR = 0.5748 and 0.5452 respectively (Figure 81).

When WT *NPM1* was assessed in the BeatAML data, only the IFNG PI scores was effective (P<0.001) and a high score indicated poor outcome (HR = 1.730) (Figure 12). In the HOVON data sets, only the 5AzaC and IFNG5AzaC PI scores were significant for WT *NPM1 (OS P<0.01 & P<0.0001 respectively, EFS P< 0.01 & P< 0.0001 respectively)*, and showed the same pattern of high PI scores associated with poorer outcome for both OS (HR = 1.510 & 1.716 respectively) and EFS (HR = 1.438 & 1.708 respectively) (Figure 83). These results show a potential relationship between PI scores and *NPM1* mutation status, where in mutated *NPM1* high PI scores were associated with improved outcomes, but in WT *NPM1* this outcome worsened. However, this inverse relationship could not be confirmed in the HOVON data set where patients with mutated *NPM1* reported HR>1 and PI scores were unable to stratify patients (P>0.05) (See 7.14.1).

In the BeatAML data set an inverse relationship of hazard ratios was observed for all PI scores in patients with *FLT3* WT (HR>1) and *FLT3-ITD (HR<1)* in OS. *FLT3-ITD* and high 5Azac or IFNG5AzaC PI scores were associated with improved outcomes (P<0.01 both), whereas WT *FLT3* and high IFNG PI score was associated with poorer outcome (P<0.0001) (Fig14). The trend of WT *FLT3* and high PI scores indicating poorer outcome was confirmed in the HOVON data set, however, this time only the 5AzaC and IFNG5AzaC scores were significant (Figure 85).

Finally, patients with rearranged *KMT2A* and IFNG and IFNG PI scores reported improved outcome for both OS (P<0.05 & P<0.001 respectively) and EFS (P<0.05 & P<0.01 respectively), shown in Figure 86.

Overall, PI stratified subgroups of patients with molecular abnormalities and identified individual patients with poorer prognostic outcomes such as patients with WT *NPM1* and high PI scores. Additionally, patients with *FLT3-ITD* belong in the unfavorable ELN cytogenetic risk group. Patients with *FLT3-ITD* and high PI scores exhibit improved outcomes (Figure 84).

This study demonstrated the advantages and disadvantages of cell line derived prognostic signatures. As expected, PI scores performed best in the discovery TCGA data set than the validation data sets, this is due to the bias of the transcripts being chosen using the TCGA OS survival and transcript data. Success of scores in validation data sets varied, likely due to differing patient characteristics or treatments between data sets due to the heterogeneous nature of AML.

The PI scores were able to match some existing methods for stratification of patients, including ELN cytogenetic risk and the LSC17 signature. An advantage of the PI scores presented here is their ability to stratify patient outcomes for those who have received induction chemotherapy as well as for patients in the intermediate ELN cytogenetic risk group. However, the PI scores lacked the predictive power of other established scores, such as the 5 gene score, or the Parsimonious 3 gene score. This may be because the PI scores here are generated from a limited number of repeats (n=3), from cell lines, which are not perfect reflections of *in* vivo biological processes, whereas the other scores mentioned are generated directly from hundreds of patient samples. The model could be improved by increasing the number of cell line repeats and unique cell lines used could create more robust and powerful signatures using this method. With more data, advanced neural networks (ANN) could be applied to find relationships between expression data across cell lines as performed by Wagner et al and others (Kathad, et al. 2021). Additionally, co-culturing AML cell lines with immune cells would improve to model by providing more data on the AML immunosuppressive mechanisms that effectively suppress immune cell activity.

6 Discussion

AML impairs the production and maturation of functional immune cells and is typically managed with a combination of chemotherapy and allogenic-HSCT (allo-HSCT). However, after treatment by allo-HSCT, there is still a high chance of relapse (50%) within the first 6 months following treatment (Yang, J., et al. 2018). Currently, patients are stratified for outcome by the ELN risk classification panel. This stratification method uses AML cytogenetic profiles, mutational status, blast percentage and minimal residual disease to place patients into favourable, intermediate and unfavourable risk groups(Döhner, et al. 2022). Approximately 50% of patients are labelled as 'intermediate' risk, and this group exhibits a wide array of responses to chemotherapy and allo-HSCT.

The goal of allo-HSCT is to provide the patient with a "reset" immune system. The elimination of AML relies on interactions between the host's antigen-presenting cells (APCs) and donor cytotoxic lymphocytes (Kolb 2017). Therefore, immune suppression mechanisms are particularly harmful to a patient's overall survival and risk of relapse. Although AML employs various mechanisms to evade the immune system, this study focused on IFNG-facilitated evasion. IFNG is known to upregulate molecules that inhibit cytotoxic immune cells, polarise T-cells towards tolerogenic phenotypes, deplete amino acids to hinder cell growth, and generally foster an immunosuppressive environment (Taghiloo and Asgarian-Omran 2021, Cornel, Mimpen and Nierkens 2020, Locafaro, et al. 2014, Folgiero, et al. 2014).

The variation in outcome seen in the intermediate ELN group indicates a need for an improved stratification method to assign these patients more suited treatments. The hypothesis of this work was that IFNG may induce an immunosuppressive phenotype in AML, which drives immune escape and leads to poorer patient outcomes. Additionally, the methylation status of AML may influence IFNG-driven immune suppression. Furthermore, different IFNG-driven immunosuppressive phenotypes could emerge depending on the maturity of the AML. Therefore, a prognostic index score based on methylation-modulated IFNG-driven immune suppression could stratify patients for survival, with higher scores expected to be associated with increased immune suppression and consequently poorer outcomes. To test this hypothesis, AML cell lines were used to generate scores related to IFNG-induced immunosuppression. AML cell lines were tested and characterised to ensure they presented with IFNG inducible immunosuppression phenotypes prior to transcriptomics being performed.

First, the cell lines were treated with IFNG and assessed for the induction of immunosuppressive molecules using RT-PCR, flow cytometry, colorimetric assay, and SWATH-MS. Among the four

AML cell lines studied, Kasumi-1 was the most responsive to IFNG, while KG-1 was the least (Number of significantly differentially expressed proteins: SIG-M5 = 341, THP-1= 141, Kasumi-1 = 101, KG-1 = 48, Figure 18). Pearson correlation and PCA analyses confirmed the weak response of KG-1 compared to the robust responses of SIG-M5, THP-1, and Kasumi-1 (Figure 20). IFNG induced increased expression of several AIR molecules in Kasumi-1 (*CD274*, HLA-ABC, HLA-E, HLA-G, PD-L1 and Kynurenine, Figure 14,Figure 15,Figure 16) and significant upregulation of all 7 IRDS proteins to the highest expression levels out of all cell lines, indicating a strong IFNG response (Figure 23). Conversely, IFNG induced changes to only 2 IRDS proteins in KG-1, but still upregulated some AIR molecules (Table 13). Consequently, Kasumi-1 and KG-1 were selected for further transcriptomic analysis due to their IFNG-inducible immunosuppressive phenotypes, despite their differing IFNGresponse strengths.

The selected cell lines were treated with IFNG and 5AzaC, followed by RNA sequencing. Kasumi-1 showed a strong response with 11,377 differentially expressed transcripts, while KG-1 had a weaker response with 2,564 transcripts, reflecting Chapter 3 findings. Transcription data identified high PTPN6 expression in KG-1 compared to Kasumi-1 (Kas-1 = 5.09 FPKM, KG-1 = 68.59 FPKM, Padj < 0.0001, Wald-test, BH, Figure 38) and its protein product SHP-1 (KG-1 = 1,708,676 NPA vs. Kas-1 = 114,481 NPA, Padj < 0.0001, Holm-Sidak, Figure 38), which likely accounts for the differences in IFNG signalling. SHP-1 prevents phosphorylation of STAT1, limiting IFNG-induced transcription (Schroder, et al. 2004). High SHP-1 in KG-1 likely conferred IFNG resistance, while low SHP-1 in Kasumi-1 led to a hyper-response. Even with IFNG treatment, Kasumi-1's SHP-1 levels remained lower than those in untreated KG-1 (data in Appendix 7.6), enabling a prolonged IFNG response (Schroder, et al. 2004). Metascape analysis of both cell lines reported enrichment in IFNG pathways (Table 8 and Table 9), confirming that IFNG signalling pathways were still activated in KG-1 despite its high SHP-1 levels.

IFNG treatment induced expression of immune evasion transcripts such as *HLA-E* (Kas-1 Ctrl = 90.63 FPKM, IFNG = 821.15 FPKM, KG-1 Ctrl = 131.32, IFNG = 468.07 FPKM, both Padj < 0.0001, Figure 50), *LGALS9* (Kas-1 Ctrl = 20.68 FPKM, IFNG = 86.27 FPKM, KG-1 Ctrl = 4.58 FPKM, IFNG = 11.15 FPKM, both, Padj < 0.0001, Figure 52) and *TNFSF10* (Kas-1 Ctrl = 0.70 FPKM, IFNG = 138.95 FPKM, Padj < 0.0001, KG-1 Ctrl = 19.02 FPKM, KG-1 IFNG = 26.99 FPKM, Padj < 0.05, Figure 45), in both cell lines. Pre-treatment with 5AzaC synergistically enhanced IFNG-induced expression of immune escape molecules in Kasumi-1 only, including *TNFSF10* (Kas-1 IFNG5AzaC = 191.18 FPKM, Padj < 0.0001, Figure 45), *HLA-E (*Kas-1 IFNG5AzaC = 936.71 FPKM, Padj < 0.0001, Figure 50) and *LGALS9* (Kas-1 IFNG5AzaC = 103.97 FPKM, Padj<0.0001, Figure 52), compared to IFNG alone. It was hypothesised that KG-1 did not exhibit a strong response to 5AzaC due to a

hypermethylation state caused by TET1/2 LOF mutations. Therefore, targeting only DNMT1 was not as impactful as it was in Kasumi-1. Chapter 4 diagrams summarise these immune suppression-related transcript changes in each cell line following IFNG and 5AzaC treatments (Figure 68 and Figure 69). The transcriptomics data generated from the IFNG and 5AzaC treatment of cell lines were used to generate PI scores.

Novel PI scores were generated from cell lines treated with IFNG, 5AzaC or both. Short lists of significantly differentiated transcripts associated with each condition were identified using pairwise linear regression. The top 500 transcripts from all comparisons were input to a CPH forward selection model which used TCGA patient survival data. PI scores were made using the Blamey et al method and validated in BeatAML, HOVON, German-AML and the CN-AML patient data sets.

PI scores effectively stratified patient outcomes in the TCGA discovery set, and each PI score was validated in at least one additional dataset. High PI scores were linked to significantly poorer outcomes among patients who received only induction chemotherapy (5AzaC PI, $P<0.001$, IFNGPI, $P<0.001$ and IFNG5AzaC PI, $P< 0.0001$,Figure 79) and those who underwent HSCT (5AzaC PI, $P<0.01$ and IFNG5AzaC PI, $P<0.01$, Figure 80) in the TCGA discovery series. Additionally, a correlation was observed between ELN risk categories and PI scores, with higher PI scores associated with worse risk categories in patients (Comparison of good to poor patients: 5AzaC PI, $P<0.05$, IFNG $P<0.01$, IFNG5AzaC $P<0.01$, Comparison of good to poor patients:5AzaC $P<0.01$, IFNG $P<0.05$, IFNG5AzaC $P<0.0001$, Figure 76). Notably, 5AzaC and IFNG PI scores were particularly effective in stratifying patients within the ELN intermediate risk category, as demonstrated in both the discovery series (5AzaC PI, $P<0.01$, IFNG PI, $P<0.05$ and IFNG5AzaC PI ,Figure 78) and the CN-AML validation dataset (5AzaC PI, $P<0.05$, IFNG PI, $P < 0.05$, Figure 88). Finally, AUC analysis (5AzaC PI AUC = 0.599, IFNG PI AUC = 0.637 and IFNG5AzaC PI AUC = 0.657, Figure 89) revealed similar performance of these scores compared to the LSC17 signature and ELN cytogenetic risk categories (AUC = 0.65 and 0.66, respectively). However, they were outperformed by other AML scores, including the parsimonious 3-gene score and the 5-gene score by Sha et al. (AUC = 0.71 and 0.74).

This study postulated that IFNG induces immune suppression in AML, resulting in poorer outcomes, and that the methylation status of AML may influence this suppression. Furthermore, AML cell lines could be used to develop PI scores associated with this suppression. Characterisation of AML cell lines revealed an IFNG-inducible immune response, which was enhanced by pretreatment with 5AzaC in one cell line but not in another. All scores were validated

using at least one additional patient dataset. The IFNG5AzaC score performed best, clearly stratifying 'high' and 'low' PI scores, in HOVON (n=593, Figure 87) and German AML (n=417, Figure 88) datasets (both P<0.0001). In contrast, the IFNG PI score was ineffective (ns, Figure 88), and the 5AzaC PI score stratified patients with lower significance (P<0.01, Figure 88) in those data sets. However, both IFNG and 5AzaC PI scores effectively stratified patients in the intermediate-risk group, while the IFNG5AzaC PI score was not significant (P<0.05, P<0.05 and ns respectively, Figure 88).

Overall, this work demonstrated that AML cell lines exhibited IFNG-inducible immune suppressive phenotypes. Additionally, PI scores derived from 5AzaC, and IFNG-treated cell line models could be used to significantly stratify AML patients. The IFNG5AzaC PI score performed best when validated in general AML patient populations, while the individual 5AzaC PI score and IFNG PI score showed significant stratification in the intermediate-risk patient population of the CN-AML dataset.

Other groups have explored IFNG-related signatures for AML prognosis. Recently, a study used single-sample gene set enrichment analysis on transcriptomics data from 672 de novo AML patients to create a parsimonious 47-gene score (Wang et al., 2024). They found that the IFNG score correlated with molecules involved in avoiding immune recognition, such as HLA-E and CD47-SIRPα, which were also highlighted as IFNG inducible in this thesis. Additionally, the score was associated with venetoclax resistance, proliferation of AML blasts, and worse overall survival for patients in the training data set. Their score was validated using Kaplan-Meier survival curve analysis in the Malani dataset (Malani, et al. 2022), where significant stratification was observed with above and below median score cutoffs (P<0.05, n=186). All scores in this thesis were validated in at least 2 data sets each, showing similar or stronger performances in terms of significance of separation than the IFNG-47 gene score. They did not test the score in intermediate risk patients or report AUC specificity and sensitivity, and so the scores cannot be compared for performance in that group.

Furthermore, Wang et al found that IFNG signalling scores had an approximately normal distribution, indicating variation in IFNG signalling between patients. IFNG signalling was highest in specific cytogenetic groups (CBF AML with inv(16) and non-CBF, non-diploid AML with del7/7q), supporting the theory that there may be strong and weak IFNG signalling phenotypes in AML patients, as observed in AML cell lines in this thesis (Wang, B., et al. 2024). Another group has reported that high IFNG expression in AML BM samples correlated with higher Treg population and expression of *IDO1*. They found that AML cells release IFNG which remodelled the BM

towards a tolerant phenotype and generated an IFNG related PI where high scores also indicated worse survival (Corradi, et al. 2022). Further supporting the idea that there is a specific type of IFNG signalling characterised by high IFNG signalling strength related to immunosuppression and worse outcome. Their score performed strongly when validated in the TCGA dataset (P<0.0001). Genes in their IFNG-based PI were found to correlate with TP53 mutations, which are known to be associated with higher IFNG expression and increased Treg populations in AML (Vadakekolathu, Lai, et al. 2020, Ragaini, et al. 2022, Sallman, et al. 2020). This suggests a link between the IFNG-induced immunosuppressive phenotype and TP53 mutational status, which was not explored in this work.

Finally, the 5AzaC and IFNG5AzaC PI scores shared *BST2* and *CALR* in common with an immune related gene (IRG) signature developed by another group for AML (Xu, et al. 2022). This group also developed their 7 gene signatures using a CPH model and the TCGA data set and validated in two patient data sets (GSE 37642, P<0.01 and GSE 146173, P<0.05). Furthermore, AUC analysis of the signature showed slightly improved specificity and sensitivity to the 5AzaC PI score developed in this thesis, but not as strong as the IFNG5AzaC PI score (7-gene signature AUC = 0.621, 5AzaC PI AUC = 0.599, IFNG5AzaC AUC = 0.657).

Identifying additional AML cell lines with strong IFNG signalling could enhance the model and investigate the relationship between IFNG, SHP1, and immune suppression. Experimental validation is needed to determine if the loss of SHP1 induces a Kasumi-1-like IFNG signalling phenotype in other AML cell lines. This would facilitate the exploration and characterisation of the IFNG-SHP1 axis and its role in AML immunosuppression, potentially leading to the development of an improved AML cell line derived IFNG-related PI score. Future studies could further validate these PI scores in other data sets and explore how these scores can be combined with other established scores for improved prognostication in AML.

In conclusion, the PI scores generated in this study stratified patients with short and long OS and EFS in the discovery series (TCGA), and each PI score could be validated in at least two other patient data sets. Cell line derived PI scores showed similar sensitivity and specificity, as determined by AUC to other AML signatures, developed from patient sample derived PI scores.
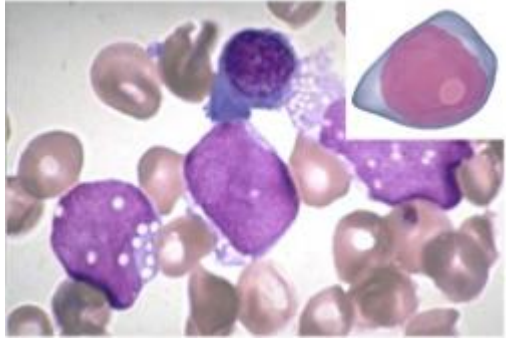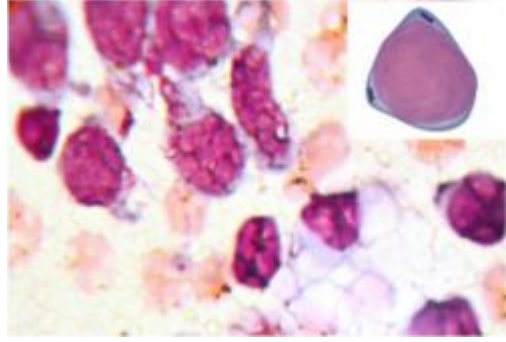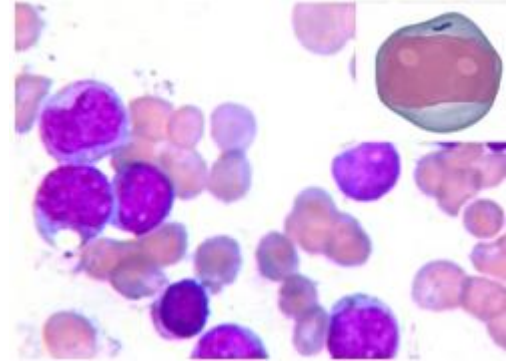
Although not as powerful as other stratification signatures for AML, the source of data is from cell lines, a more readily available resource than patient samples. Cell lines undergo fundamental changes from the patient source cells to proliferate indefinitely and be cultured, so they are not perfect models of the patient they are derived from. However, they are still biologically relevant platforms to study cancer processes, and in combination with patient data sets, this study has
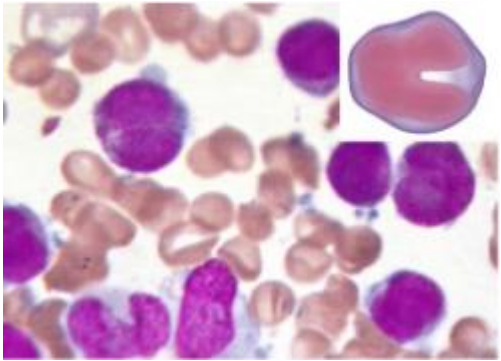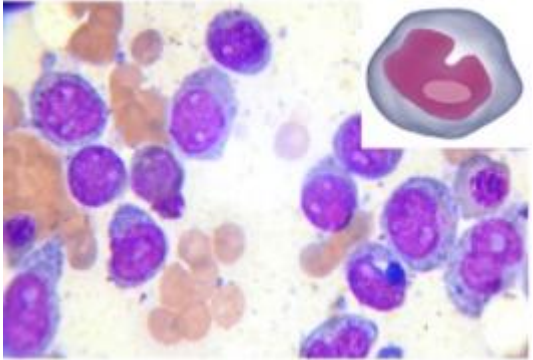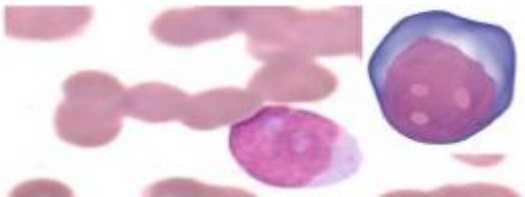
shown their potential to inform prognostic signatures with clinical relevance. The PI scores created in this study are easy to calculate, with 8 or fewer transcripts to measure and have the potential to be integrated into clinical practice. Signatures composed of a smaller number of components are easier to incorporate into clinical practice, and so these signatures offer advantages over the 47-gene IFNG signature and LSC17 gene signatures. Importantly, these scores show strength in stratifying patients in the intermediate risk group, where there is currently a lack of reliable scores for prognostication (Gerstung, et al. 2017, Döhner, et al. 2017). Furthermore, they could improve allocation of intensive chemotherapy treatments to patients with intermediate cytogenetic risk and low PI scores, while finding alternative to the subset with high PI scores, who may suffer from intense chemotherapy.
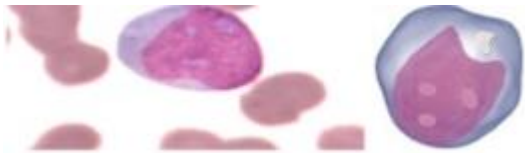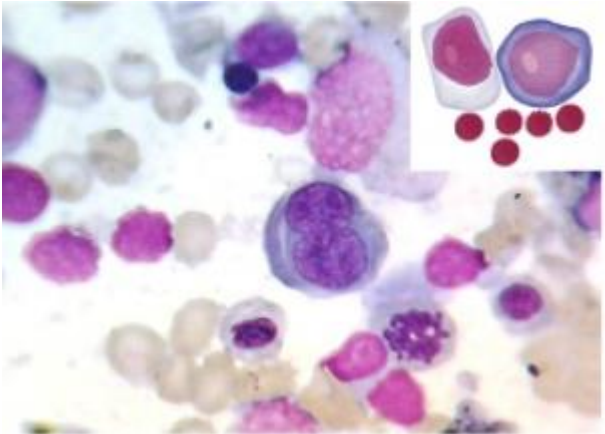
7 Appendix
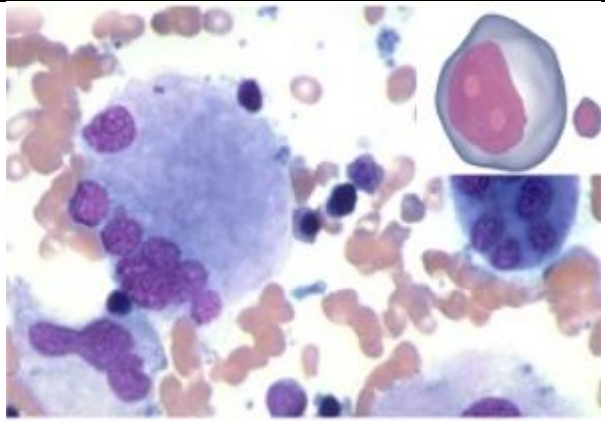
7.1 Table of AML morphology classed by the FAB classification system

Table 21:The categories of AML based on the cell of origin and its morphological characteristics according to the FAB classification system (American Cancer Society 2018). Descriptions and photos are from (Ladines-Castro, et al. 2016).

FAB	Name	Description of morphology	Example by Ladines-Castro et al 2016
M0	Acute myeloblastic leukaemia with minimal differentiation	Medium sized blasts with round nuclei, fine chromatin, a prominent nucleolus and nongranular cytoplasm.	
M1	Acute myeloblastic leukaemia without maturation	Medium sized blasts with a high nucleus to cytoplasm ration. The nuclei are round with immature dispersed chromatin and sometimes multiple nucleoli are present. Between 5 and 10% have isolated Auer rods in the cytoplasm or a fine azurophilic granulation.	
M2	Acute myeloblastic leukaemia with maturation	Blasts are small or medium sized with high nucleus to cytoplasm ratio. Immature chromatin is dispersed in the nucleus and there is one or more nucleoli. Again, there are isolate Auer rods and	

		small amounts of azurophilic granulation. The cytoplasm is basophilic.	
M3	Promyelocytic leukaemia (PML)	Has intense azurophilic granulation and scarce basophilic staining of the cytoplasm. Nucleus is irregular with a distinct cleft. Some PML have long crystal structures in the cytoplasm and form clumps that can be differentiated from Auer rods by their more tubular structure.	
M4	Acute myelomonocytic leukaemia	Blasts are large and have a moderate ratio of nucleus to cytoplasm and variable basophilic staining. The nucleus varies from a typic round shape, to irregular or kidney shape. Nucleoli are also prominent.	
M5	Acute monocytic leukaemia	There are two types of M5. The first M5a is characterised by large blasts with round nuclei and dispersed immature chromatin between 1 and 3 three	**M5a**

		nucleoli. The cytoplasm is basophilic and may also contain Auer rods, granulations, and prolongations. The second type M5b, features promonocytes with kidney or nucleus shaped nucleus. The cytoplasm is less basophilic and has more granulation than what is seen in monoblasts. Additionally, it contains some vacuoles.	**M5b**
M6	Acute erythroid leukaemia	There are two types of M6. M6a features 50% erythroid precursor cells and 30% myeloblasts. The erythroid cell morphology is changed in peripheral blood. With observations of fragmented erythrocytes (schistocytes), 'mushroom shaped' cells and spiked erythrocytes (echinocytes and acanthocytes). M6b has a composition of 80% bone marrow cells with less than 3% myeloid cells. Larger erythrocytes (macrocytes), basophilic stippling of	

		the cytoplasm and Howell-jolly bodies and Cabot rings are distinctive features of the subtype.	
M7	Acute megakaryocytic leukaemia	Immature polymorphic blasts with a dispersed irregular nucleus containing reticulated chromatin and between 1 and 3 nucleoli. The cytoplasm is basophilic and nongranular with a similar appearance to platelets.	

7.2 Reagents and equipment

All reagents stored according to manufacturer's instruction and used within defined expiry date.

Reagent	Supplier
Cell culture medium	
1640 RPMI without L-glutamine	Lonza/ Corning
Iscove's MDM without L-glutamine	Lonza/ Corning
MCcoys medium without L-glutamine	Lonza/ Corning
Cell culture additives	
5-Azacytidine	Sigma
Foetal Calf Serum	GE Healthcare Hyclone
Interferon Gamma	R&D systems
L-Glutamine	SLS (Lonza)
Other cell culture Reagents	
Dimethyl sulfoxide (DMSO)	Santa Cruz biotechnology
Dulbecco's phosphate buffered saline (DPBS)	SLS (Lonza)

Trypan blue solution 0.4 %	Sigma-Aldrich
Solution 18 AO•DAPI	Chemometec

Chemical reagent

0.5 M EDTA solution (100x)	Thermo Scientific
5-Azacytidine	Sigma-Aldrich
10x TRIS Glycine	Geneflow
10x TRIS/Glycine/SDS	Geneflow
1M Tris-HCl pH 8	Invitrogen
2-Mercaptoethanol	Sigma-Aldrich
4x Laemmli buffer	Bio-Rad
Annexin V binding buffer	BioLegend
Bovine Serum Albumin (BDS)	Sigma-Aldrich
Cytarabine	Sigma-Aldrich
Daunorubicin	Sigma-Aldrich
Deoxyribonucleotide triphosphate (dNTP)	Promega
Double distilled water (ddH$_2$O)	Barnstead
Ethanol	Fisher scientific
Glacial acetic acid	Sigma-Aldrich
Halt protease and phosphatase inhibitor cocktail (100X)	Thermo Scientific
Instant dried skimmed milk	Marvel
Ionic detergent compatibility reagent	Thermo Scientific
Liquid nitrogen	BOC
L-Kynurenine	Sigma-Aldrich
Methanol	Fisher Chemical
NaCl	Sigma
Nuclease-free water	Ambion
Oligo(dT)$_{15}$ Primer	Promega
Presept tablets	Johnson and Johnson
SDS	Sigma
Sodium deoxycholate	Sigma
Triton x 100	Sigma
Trichloroacetic acid	Sigma-Aldrich

P-dimethylaminobenzaldehyde	Sigma-Aldrich
Protein Assay Dye Reagent Concentrate	Bio-Rad
PierceTM 660 nm Protein assay reagent	Thermo Scientific
Sodium dodecyl sulphate (SDS)	Sigma-Aldrich
Reverse Transcriptase	Promega
Reverse Transcriptase 5x RT buffer	Promega
RNaseZAP	Ambion
RNasin	Promega
RT 5x Buffer	Promega
Standard BSA set	Thermo Scientific
SYBR Green	Bio-Rad
Trizma (Tris) base	Sigma-Aldrich
Tween 20	Sigma-Aldrich

Kits

Clarity western ECL substrate	Bio Rad
Cell Proliferation Kit II (XTT)	Sigma-Aldrich
RNeasy mini kit	Quiagen

7.2.1 Consumables and Equipment

Laboratory Plastics, Glassware and Sharps	Supplier
4–20% Mini-PROTEAN® TGX™ Precast Protein Gels	Bio-Rad
Bijou tubes (7 mL)	Starlab
Cell culture flasks (T25, T75, T175)	Sarstedt
Clear flat bottom 6-well plate, sterile (cell culture)	Sarstedt
Clear flat bottom 24-well plate, sterile (cell culture)	Sarstedt
Clear flat bottom 96-well plate, sterile (cell culture)	Sarstedt
Clear flat bottom 96-well plate (protein assay)	Starlab
Cryogenic vials (1.0 mL)	Starlab
Falcon tubes (15 mL, 50 mL)	Sarstedt
Filter tips (10ul, 20ul, 100ul, 200ul, 1000ul)	Starlab
Glass bottles	Duran
Glass coverslips	

Glass slides	
Nitrocellulose membrane	GE Water & Process Techn.
Rotor-Gene Strip Tubes & Caps	Starlab
Pasteur pipettes	Sarstedt
Petri dishes	Sarstedt
Scalpels	SLS
Serological pipettes (5 mL, 10 mL, 25 mL)	Sarstedt
Syringe filter 0.2µm	Sartorius
Syringes (20 mL)	Medicina
Western Blot filter paper	GE Healthcare

Equipment	Supplier
4°C Fridge	LEC Medical
-20°C Freezer	LEC Medical
-80°C Freezer	Panasonic
37°C/5% CO_2 Incubator	Scientific Laboratory Supplies
4°C Centrifuge	Eppendorf
Agilent 2100 bioanalyzer	Agilent
Autoclave	Rodwell
Benchtop vortex mixer	Scientific Industries
Chip priming station	Agilent
Class II Safety Cabinet	Walker
CoolCell Freezing System	Corning
Haemocytometer (counting chamber)	Weber
Heating block	Lab-Line
Micropipettes (2 µl, 10 µl, 100 µl, 200 µl, 1000 µl)	Gilson/Starlab
Minispin benchtop centrifuge	Eppendorf
Multichannel pipette (300 µl)	Eppendorf
Nanodrop ND-8000 spectrophotometer	Thermo Scientific
Nanopure Diamond water reservoir	Barnstead
NucleoCounter NC-250	Chemometec
PCR workstation cabinet	Grant-Bio

Real-time qPCR thermal cycler	Qiagen
Rocker	Stuart
SCIEX TripleTOF 6600	SCIEX
Syngene G:Box	Syngene
Sonicator	Fisherbrand
Tecan Ultra Microtiter Plate Reader	Tecan
Thermoblock	Biometra
Waterbath	Clifton
Weighing Scale	Fisher Scientific
Software	**Company**
GraphPad Prism v8	Graphpad software inc.
Morpheus (online)	Broadinsitute
Rotor-GeneQ Series Software v 2.3.5	Qiagen
HL image Quick spots	Western vision
Metacore	Clarivate Analytics
NucleoView	Chemometec
Protein Pilot v5.0	SCIEX

7.2.2 Composition of buffers used

7.2.2.1   Cell culture

Growth medium for Kasumi-1	**Makes 500 mL**
RPMI 1640 1 x	395 mL
Fetal Calf Serum	100 mL (20 %)
L-glutamine	5 mL (2mM)

Growth medium for KG-1	**Makes 500 mL**
Iscove's MDM	395 mL
Fetal Calf Serum	100 mL (20 %)
L-glutamine	5 mL (2mM)

Growth medium for SIG-M5	**Makes 500 mL**
Iscove's MDM	395 mL

Fetal Calf Serum	100 mL (20 %)
L-glutamine	5 mL (2mM)

Growth medium for THP-1	**Makes 500 mL**
RPMI 1640 1 x	395 mL
Fetal Calf Serum	100 mL (20 %)
L-glutamine	5 mL (2mM)

7.2.2.2   Flow cytometry

FAC's buffer	**For 1 L**
DDH$_2$O	1 L
BSA	5 g
Sodium Azide	0.2 g

7.2.2.3   Protein extraction

RIPA buffer	**For 50 mL**
1M Tris-HCl pH 8	2.5 mL
150 mM NaCl	7.5 mL
10 % SDS	0.5 mL
10 % Sodium Deoxycholate	2.5 mL
10% Triton x 100	5 mL
1 mM EDTA	0.5 mL
ddH$_2$O	31.5 mL

7.2.2.4   Western blot

4x Laemmli buffer (1610747, Bio-Rad)	**For 1 mL**
4x Laemmli buffer	900 µL
2-Mercaptoethanol	100 µL

7.2.2.5   Running buffer

	For 1L
10x TRIS/Glycine	100 mL
ddH$_2$O	900 mL

7.2.2.6   Transfer buffer

	For 1 L

10x TRIS/Glycine	100 mL
Methanol	200 mL
ddH$_{20}$	700 mL

7.2.2.7 Blocking buffer (5%)

	For 50 mL
ddH$_2$O	50 mL
Instant dried skimmed milk	**2.5 g**
10x Tris-buffered saline (TBS) pH 7.6	**For 1 L**
Trizma Base	24.2 g
Sodium Chloride	80 g
ddH$_2$0	1000 mL

7.2.2.8 1x Tris-buffered saline (TBST) pH 7.6

	For 1 L
10x Tris-buffered saline (TBS) pH 7.6	100 mL
ddH$_2$O	900 mL
Tween 20	1 mL

7.2.3 Antibodies

7.2.3.1 Antibodies used in western blot analysis

Reagent	Company	Cat no	Dilution
DNMT1 XP® Rabbit mAb	Cell signalling technology	D63A6	1:1000
p53 Antibody (DO-1): sc-126 MOUSE	Santa Cruz	Sc-126	1:1000
Human Phospho-p53 (S46) Antibody	R&D systems	AF1489-SP	1:200
STAT1 Rabbit mAb	Cell signalling technology	9172S	1:1000
Phospho-Stat1 (Ser727) Antibody Rabbit	Cell signalling technology	9177	1:1000
RNF213 Polyclonal Antibody Rabbit	Thermo Fisher Scientific	PA5-51902	1:250
GAPDH (D16H11) XP® Rabbit mAb	Invitrogen	5174S	1:1000

Reagent	Catalog number	Supplier	Clone	Volume used (µL)	Laser (nm)	Dichroic filter	Band pass filter
LIVE/DEAD fixable yellow dead cell stain	L34968	Invitrogen	-	0.5	405	-	550/40
PE antihuman CD119 (IFN-Y R a chain)	308703	Biolegend	GIR-94	2.5	488	595 SP	575/30
APC anti human HLA-E	342605	Biolegend	3D12	2.5	638	694 SP/25	660/20
PE anti human HLA-G	335906	Biolegend	87G	2.5	488	595 SP	575/30
APC/FIRE 750 anti-human HLA-A, B, C	311443	Biolegend	W6/32	2.5	638	-	755 LP
PE/Cy7 anti-human CD274	329718	Biolegend	29E.2A3	2.5	488	-	755 LP
Annexin V Alexa Fluor® 647	640943	Biolegend	-	2.5	638	-	660/20

7.2.4 Quantitative real-time PCR primers used throughout this study

Gene	Primer	Primer Sequence	Annealing temp (°C)
GUSB	FH1_GUSB	5' ACTGAACAGTCACCGAC 3'	58
GUSB	RH1_GUSB	5' AAACATTGTGACTTGGCTAC 3'	58
TDO2	FH1_TDO2	5' AAGAAAAAGAGGAACAGGTG 3'	58
TDO2	RH1_TDO2	5' CACCTTTACTAAGGAGATGTTC 3'	58
CD274	FH1_CD274	5' ATGCCCCATACAACAAAATC 3'	58
CD274	RH1_CD274	5' GACATGTCAGTTCATGTTCAG 3'	58
IDO1	FH1_IDO1	5' TTGTTCTCATTTCGTGATGG 3'	58
IDO1	RH1_IDO1	5' TACTTTGATTGCAGAAGCAG 3'	58

7.3 Tables of most significantly enriched pathways reported by metascape for each treatment

Table 22: The top 20 most significantly enriched pathways in Kasumi-1 treated with IFNG according to the 'unique' up and down regulated transcript lists, sorted by significance post FDR. Column 'transcript' specifies the number of transcripts differentially expressed in the data compared to the total number of transcripts from the corresponding pathway.

Kasumi-1 IFNG

		Upregulated			Downregulated		
Summary process ID	Description	Log10(P)	Transcript	Summary process ID	Description	Log10(P)	Transcript
GO:0002263	Cell activation involved in immune response	-36.94	200/724	R-HSA-8953854	Metabolism of RNA	-83.09	252/673
GO:0001817	Regulation of cytokine production	-23.74	184/795	R-HSA-8868773	Rrna processing in the nucleus and cytosol	-74.18	125/194
GO:0042110	T cell activation	-22.43	129/475	R-HSA-1640170	Cell Cycle	-33.40	179/692
GO:0031349	Positive regulation of defence response	-18.28	130/534	GO:0006399	Trna metabolic process	-27.95	79/193
R-HSA-1280215	Cytokine Signalling in Immune system	-18.26	157/707	GO:0042273	Ribosomal large subunit biogenesis	-22.77	43/72
GO:0002521	Leukocyte differentiation	-17.53	127/526	GO:0006281	DNA repair	-22.48	138/567
GO:0009615	Response to virus	-15.90	96/359	GO:0006260	DNA replication	-19.98	85/278
GO:0030335	Positive regulation of cell migration	-14.68	125/555	R-HSA-72306	Trna processing	-19.32	49/107
WP1449	Regulation of toll-like receptor signalling pathway	-13.49	52/145	GO:0009117	Nucleotide metabolic process	-18.81	154/726
GO:0002237	Response to molecule of bacterial origin	-13.14	88/346	GO:0071103	DNA conformation change	-18.78	91/323
GO:0002697	Regulation of immune effector process	-12.94	108/474	GO:0042274	Ribosomal small subunit biogenesis	-18.58	40/75
GO:0032103	Positive regulation of response to external stimulus	-12.94	85/331	GO:0032200	Telomere organization	-17.92	62/174
GO:0007249	I-kappab kinase/NF-kappab signalling	-12.64	76/282	R-HSA-72203	Processing of Capped Intron-Containing Pre-mrna	-16.99	74/244
GO:0010942	Positive regulation of cell death	-12.48	144/725	GO:0007005	Mitochondrion organization	-16.10	122/551
GO:0043408	Regulation of MAPK cascade	-11.47	143/740	GO:0044770	Cell cycle phase transition	-16.02	135/641
GO:0060627	Regulation of vesicle-mediated transport	-11.40	113/534	GO:0140053	Mitochondrial transcript expression	-15.67	57/165
GO:0006914	Autophagy	-11.28	114/543	WP4022	Pyrimidine metabolism	-15.61	39/83
GO:0051345	Positive regulation of hydrolase activity	-10.82	147/784	GO:0006403	RNA localization	-15.56	70/235
Ko04060	Cytokine-cytokine receptor interaction	-10.77	70/270	R-HSA-6790901	Rrna modification in the nucleus and cytosol	-14.36	32/61
GO:0043900	Regulation of multi-organism process	-10.52	96/437	GO:0090305	Nucleic acid phosphodiester bond hydrolysis	-14.32	80/306

KG-1 IFNG

Table 23: The top 20 most significantly enriched pathways in KG-1 treated with IFNG according to the 'unique' up and down regulated transcript lists, sorted by significance post FDR. Column 'transcript' specifies the number of transcripts differentially expressed in the data compared to the total number of transcripts from the corresponding pathway.

		Upregulated				Downregulated	
Summary process ID	Description	Log10(P)	Transcript	Summary process ID	Description	Log10(P)	Transcript
GO:0045055	Regulated exocytosis	-0.27	9/780	WP306	Focal Adhesion	-4.30	13/199
M3468	NABA ECM REGULATORS	-0.23	5/238	GO:0006935	Chemotaxis	-4.16	21/644
GO:0002429	Immune response-activating cell surface receptor signalling pathway	0.00	6/481	WP4541	Hippo-Merlin Signalling Dysregulation	-3.98	10/122
R-HSA-1280215	Cytokine Signalling in Immune system	0.00	7/707	GO:0050865	Regulation of cell activation	-3.86	20/636
GO:0060333	Interferon-gamma-mediated signalling pathway	0.00	3/91	Hsa04060	Cytokine-cytokine receptor interaction	-3.53	14/328
R-HSA-109582	Haemostasis	0.00	6/620	GO:0000904	Cell morphogenesis involved in differentiation	-3.31	20/723
GO:0032874	Positive regulation of stress-activated MAPK cascade	0.00	3/163	GO:0009611	Response to wounding	-3.08	19/701
R-HSA-382551	Transport of small molecules	0.00	6/728	GO:0071363	Cellular response to growth factor stimulus	-3.04	19/707
GO:0033209	Tumour necrosis factor-mediated signalling pathway	0.00	3/175	GO:0051968	Positive regulation of synaptic transmission, glutamatergic	-2.76	5/30
GO:0042742	Defence response to bacterium	0.00	4/348	GO:0019932	Second-messenger-mediated signalling	-2.53	14/444
GO:0051345	Positive regulation of hydrolase activity	0.00	6/784	GO:0002521	Leukocyte differentiation	-2.40	15/526
				GO:0016339	Calcium-dependent cell-cell adhesion via plasma membrane cell adhesion molecules	-2.139	5/43
				R-HSA-9006934	Signalling by Receptor Tyrosine Kinases	-2.04	14/505
				GO:0062009	Secondary palate development	-1.92	4/25
				GO:0002683	Negative regulation of immune system process	-1.826	13/484
				GO:0009617	Response to bacterium	-1.699	16/728
				GO:0035767	Endothelial cell chemotaxis	-1.699	4/32
				GO:0043114	Regulation of vascular permeability	-1.524	4/36
				GO:0007423	Sensory organ development	-1.509	13/534
				R-HSA-2730905	Role of LAT2/NTAL/LAB on calcium mobilization	-1.39	3/16

Kasumi-1 5AzaC

Table 24: The top 20 most significantly enriched pathways in Kasumi-1 treated with 5AzaC according to the 'unique' up and down regulated transcript lists, sorted by significance post FDR. Column 'transcript' specifies the number of transcripts differentially expressed in the data, from the corresponding pathway.

Upregulated				Downregulated			
Summary process ID	Description	Log10(P)	Transcript	Summary process ID	Description	Log10(P)	Transcript
GO:0050865	Regulation of cell activation	-13.03	41/636	R-HSA-191273	Cholesterol biosynthesis	-8.11	6/25
GO:0002521	Leukocyte differentiation	-12.94	37/526	GO:0045123	Cellular extravasation	-6.57	7/71
R-HSA-909733	Interferon alpha/beta signalling	-12.41	15/70	WP3613	Photodynamic therapy-induced unfolded protein response	-4.67	4/27
GO:0002274	Myeloid leukocyte activation	-11.87	40/660	GO:0006790	Sulphur compound metabolic process	-4.62	11/380
GO:0001817	Regulation of cytokine production	-11.10	43/795	ko00650	Butanoate metabolism	-4.61	4/28
GO:0040017	Positive regulation of locomotion	-11.09	37/607	GO:0090092	Regulation of transmembrane receptor protein serine/threonine kinase signalling pathway	-4.54	9/255
GO:0009611	Response to wounding	-11.06	40/701	GO:0043583	Ear development	-4.44	8/202
R-HSA-1280215	Cytokine Signalling in Immune system	-10.95	40/707	GO:0019229	Regulation of vasoconstriction	-4.44	5/61
M5885	NABA MATRISOME ASSOCIATED	-10.73	41/751	M5884	NABA CORE MATRISOME	-4.28	9/275
GO:0070371	ERK1 and ERK2 cascade	-10.41	26/327	GO:0072672	Neutrophil extravasation	-4.01	3/15
GO:0042116	Macrophage activation	-9.93	15/102	M5887	NABA BASEMENT MEMBRANES	-3.98	4/40
GO:0030198	Extracellular matrix organization	-7.99	25/395	GO:0050654	Chondroitin sulphate proteoglycan metabolic process	-3.86	4/43
GO:0002237	Response to molecule of bacterial origin	-7.79	23/346	GO:0000904	Cell morphogenesis involved in differentiation	-3.82	14/723
GO:0097435	Supramolecular fibre organization	-7.51	34/713	GO:1901654	Response to ketone	-3.71	7/193
GO:0032103	Positive regulation of response to external stimulus	-7.48	22/331	GO:0050804	Modulation of chemical synaptic transmission	-3.57	10/418
WP3624	Lung fibrosis	-7.15	10/63	GO:0010769	Regulation of cell morphogenesis involved in differentiation	-3.27	8/300
ko05144	Malaria	-7.06	9/49	GO:0010720	Positive regulation of cell development	-3.24	11/547
GO:0030098	Lymphocyte differentiation	-6.88	22/358	R-HSA-5173105	O-linked glycosylation	-3.21	5/111
WP2864	Apoptosis-related network due to altered Notch3 in ovarian cancer	-6.75	9/53	GO:0043087	Regulation of GTPase activity	-3.10	10/481
GO:0050920	Regulation of chemotaxis	-6.72	17/224	GO:0050808	Synapse organization	-3.00	9/411

KG-1 5AzaC

Table 25: The top 20 most significantly enriched pathways in KG-1 treated with 5AzaC according to the 'unique' up and down regulated transcript lists, sorted by significance post FDR. Column 'transcript' specifies the number of transcripts differentially expressed in the data, from the corresponding pathway.

		Upregulated				Downregulated	
Summary process ID	Description	Log10(P)	Transcript	Summary process ID	Description	Log10(P)	Transcript
R-HSA-9648895	Response of EIF2AK1 (HRI) to heme deficiency	-8.20	5/15	GO:0007185	Transmembrane receptor protein tyrosine phosphatase signalling pathway	-6.53	3/6
GO:0008652	Cellular amino acid biosynthetic process	-8.03	8/88	GO:0046888	Negative regulation of hormone secretion	-4.75	4/62
GO:0002274	Myeloid leukocyte activation	-7.00	16/660	R-HSA-500792	GPCR ligand binding	-4.68	8/467
GO:0001101	Response to acid chemical	-6.38	11/323	GO:0030098	Lymphocyte differentiation	-3.57	6/358
R-HSA-379716	Cytosolic tRNA aminoacylation	-5.33	4/24	GO:0016358	Dendrite development	-3.53	5/234
GO:0002237	Response to molecule of bacterial origin	-5.21	10/346	GO:0034394	Protein localization to cell surface	-3.20	3/67
GO:0098739	Import across plasma membrane	-5.16	7/148	GO:0044057	Regulation of system process	-3.15	7/597
GO:0033690	Positive regulation of osteoblast proliferation	-4.78	3/11	GO:0002573	Myeloid leukocyte differentiation	-2.76	4/205
GO:0048266	Behavioural response to pain	-4.66	3/12	GO:0097305	Response to alcohol	-2.56	4/234
GO:0050808	Synapse organization	-4.57	10/411	hsa04371	Apelin signalling pathway	-2.24	3/146
GO:0009064	Glutamine family amino acid metabolic process	-4.51	5/76	WP3932	Focal Adhesion-PI3K-Akt-mTOR-signalling pathway	-2.16	4/303
hsa04611	Platelet activation	-4.43	6/130	WP481	Insulin signalling	-2.13	3/160
GO:0089718	Amino acid import across plasma membrane	-4.33	4/42	R-HSA-9006934	signalling by Receptor Tyrosine Kinases	-2.07	5/505
GO:0009612	Response to mechanical stimulus	-4.29	7/202	GO:0010942	Positive regulation of cell death	-2.03	6/725
GO:0019932	Second-messenger-mediated signalling	-4.28	10/444				
GO:0043062	Extracellular structure organization	-4.18	10/457				
GO:0097191	Extrinsic apoptotic signalling pathway	-4.09	7/217				
GO:0097028	Dendritic cell differentiation	-4.07	4/49				
R-HSA-389357	CD28 dependent PI3K/Akt signalling	-3.83	3/22				
GO:0001819	Positive regulation of cytokine production	-3.45	9/461				

Kasumi-1 IFNG 5AzaC

Table 26: The top 20 most significantly enriched pathways in Kasumi-1 comparing combination treated cells to control cells, prior to removal of IFNG only and 5AzaC only treated transcripts, sorted by significance post FDR. Column 'transcript' specifies the number of transcripts differentially expressed in the data, from the corresponding pathway. For ease of comparison, number of transcripts altered by IFNG only treated cells are presented in brackets.

	Upregulated			Downregulated			
Summary process ID	Description	Log10(P)	Transcript	Summary process ID	Description	Log10(P)	Transcript
R-HSA-1280215	Cytokine Signalling in Immune system	-59.42 (-18.26)	228/707 (157/707)	R-HSA-8953854	Metabolism of RNA	-99.77 (-83.09)	278/673 (252/673)
R-HSA-913531	Interferon Signalling	-47.82	102/199	R-HSA-72312	RRNA processing	-81.94	135/204
GO:0001817	Regulation of cytokine production	-47.10 (-23.74)	224/795 (184/795)	GO:0006399	TRNA metabolic process	-38.17 (-27.95)	91/193 (79/193)
GO:0034341	Response to interferon-gamma	-42.55	96/197	GO:0071826	Ribonucleoprotein complex subunit organization	-30.80	104/290
GO:0002263	Cell activation involved in immune response	-39.50 (-36.94)	198/724 (200/724)	R-HSA-1640170	Cell Cycle	-28.75 (-33.40)	174/692 (179/692)
GO:0042110	T cell activation	-34.05 (-22.43)	144/475 (129/475)	R-HSA-72306	TRNA processing	-28.69 (-19.32)	58/107 (49/107)
GO:0002697	Regulation of immune effector process	-28.32 (-12.94)	134/474 (108/474)	GO:0006281	DNA repair	-27.81 (-22.48)	151/567 (138/567)
ko04621	NOD-like receptor signalling pathway	-28.06	73/170	GO:0006260	DNA replication	-23.08 (-19.98)	90/278 (85/278)
GO:0009617	Response to bacterium	-27.47	175/728	GO:0042274	Ribosomal small subunit biogenesis	-22.77 (-18.58)	43/75 (40/75)
GO:0043900	Regulation of multi-organism process	-25.84 (-10.52)	123/437 (96/437)	GO:0009117	Nucleotide metabolic process	-22.16 (-18.81)	165/726 (154/726)
GO:0032103	Positive regulation of response to external stimulus	-25.48 (-12.94)	103/331 (85/331)	GO:0006520	Cellular amino acid metabolic process	-20.61	102/365
WP619	Type II interferon signalling (IFNG)	-24.79	31/37	GO:0007005	Mitochondrion organization	-19.89 (-16.10)	132/551 (122/551)
ko04060	Cytokine-cytokine receptor interaction	-21.51	85/270	GO:0032200	Telomere organization	-19.87 (-17.92)	64/174 (62/174)
GO:0050777	Negative regulation of immune response	-21.26	62/159	GO:0071103	DNA conformation change	-19.82 (-18.78)	93/323 (91/323)
GO:0006935	Chemotaxis	-21.23	148/644	GO:1901137	Carbohydrate derivative biosynthetic process	-19.80	168/783
ko05152	Tuberculosis	-20.29	65/179	R-HSA-6790901	RRNA modification in the nucleus and cytosol	-19.76 (-14.36)	36/61 (32/61)
R-HSA-1280218	Adaptive Immune System	-19.51	161/756	R-HSA-72203	Processing of Capped Intron-Containing Pre-mRNA	-19.72 (-16.99)	78/244 (74/244)
GO:0010942	Positive regulation of cell death	-18.97 (-12.48)	155/725 (144/725)	GO:0140053	Mitochondrial transcript expression	-19.10 (-15.67)	61/165 (57/165)
hsa05169	Epstein-Barr virus infection	-18.90	86/299	GO:0090305	Nucleic acid phosphodiester bond hydrolysis	-18.75 (-14.32)	88/306 (80/306)
GO:0007249	I-kappaB kinase/NF-kappaB signalling	-17.80 (-12.64)	81/282 (76/282)	ko00240	Pyrimidine metabolism	-18.22	46/104

KG-1 IFNG 5AzaC

Table 27: The top 20 most significantly enriched pathways in KG-1 comparing combination treated cells control cells, prior to removal of IFNG only and 5AzaC only treated transcripts, sorted by significance post FDR. Column 'transcript' specifies the number of transcripts differentially expressed in the data, from the corresponding pathway. For ease of comparison, number of transcripts altered by IFNG only treated cells are presented in brackets.

		Upregulated				Downregulated	
Summary process ID	Description	Log10(P)	Transcript	Summary process ID	Description	Log10(P)	Transcript
R-HSA-6798695	Neutrophil degranulation	-6.09	14/480	hsa04060	Cytokine-cytokine receptor interaction	-9.32	21/328 (14/328)
hsa05221	Acute myeloid leukaemia	-5.49	6/70	R-HSA-1474244	Extracellular matrix organization	-9.18	20/301
hsa01230	Biosynthesis of amino acids	-5.32	6/75	GO:0030155	Regulation of cell adhesion	-7.74	29/734
GO:0043648	Dicarboxylic acid metabolic process	-4.67	6/97	GO:0007178	Transmembrane receptor protein serine/threonine kinase signalling pathway	-7.08	19/363
GO:0016999	Antibiotic metabolic process	-4.55	7/151	GO:0010942	Positive regulation of cell death	-6.73	27/725
GO:0045444	Fat cell differentiation	-4.34	8/222	WP4541	Hippo-Merlin Signalling Dysregulation	-6.63 (-3.98)	11/122 (10/122)
R-HSA-9648895	Response of EIF2AK1 (HRI) to heme deficiency	-4.09	3/15	GO:0060326	Cell chemotaxis	-6.03	16/307
GO:0071435	Potassium ion export	-3.84	3/18	GO:0009611	Response to wounding	-5.93 (-3.08)	25/701 (19/701)
GO:0030099	Myeloid cell differentiation	-3.76	10/421	WP2431	Spinal Cord Injury	-5.73	10/121
GO:0033688	Regulation of osteoblast proliferation	-3.31	3/27	GO:0030036	Actin cytoskeleton organization	-5.46	24/696
WP2849	Hematopoietic Stem Cell Differentiation	-3.26	4/65	R-HSA-500792	GPCR ligand binding	-5.44	19/467
GO:0048771	Tissue remodelling	-3.21	6/179	GO:0008285	Negative regulation of cell proliferation	-5.38	25/753
R-HSA-381426	Regulation of Insulin-like Growth Factor (IGF) transport and uptake by Insulin-like Growth Factor Binding Proteins (IGFBPs)	-3.10	5/125	R-HSA-9031628	NGF-stimulated transcription	-5.21	6/39
GO:0002237	Response to molecule of bacterial origin	-3.03	8/346	GO:0001503	Ossification	-5.14	17/402
GO:0009583	Detection of light stimulus	-2.99	4/77	GO:0001655	Urogenital system development	-5.12	15/321
GO:0001889	Liver development	-2.90	5/138	M159	PID AMB2 NEUTROPHILS PATHWAY	-5.08	6/41
GO:0043062	Extracellular structure organization	-2.86	9/457	GO:0099560	Synaptic membrane adhesion	-5.01	5/25
GO:0032787	Monocarboxylic acid metabolic process	-2.80	11/659	ko05202	Transcriptional mis regulation in cancer	-4.97	11/180
GO:0006734	NADH metabolic process	-2.68	3/44	GO:0009617	Response to bacterium	-4.65 (-1.699)	23/728 (16/728)
GO:0045936	Negative regulation of phosphate metabolic process	-2.58	10/601	GO:0051968	Positive regulation of synaptic transmission, glutamatergic	-4.60 (-2.76)	5/30 (5/30)

Kasumi-1 'unique' IFNG 5AzaC

Table 28: The top 20 most significantly enriched pathways in Kasumi-1 comparing combination treated cells control cells, according to the 'unique' up and down regulated transcript lists, sorted by significance post FDR. Column 'transcript' specifies the number of transcripts differentially expressed in the data, from the corresponding pathway. For ease of comparison, number of transcripts altered by IFNG only treated cells are presented in brackets.

		Upregulated				Downregulated	
Summary process ID	Description	Log10(P)	Transcript	Summary process ID	Description	Log10(P)	Transcript
GO:0030029	Actin filament-based process	-6.32	44/796	R-HSA-8953854	Metabolism of RNA	-14.67	63/673
GO:1901522	Positive regulation of transcription from RNA polymerase II promoter involved in cellular response to chemical stimulus	-4.96	6/22	GO:0034470	NcRNA processing	-12.34	43/395
GO:0048625	Myoblast fate commitment	-3.86	3/5	R-HSA-73894	DNA Repair	-9.67	35/335
GO:0008015	Blood circulation	-3.76	27/513	M1	PID FANCONI PATHWAY	-7.98	12/47
GO:0001525	Angiogenesis	-3.72	30/600	hsa00510	N-Glycan biosynthesis	-6.74	11/49
R-HSA-109582	Haemostasis	-3.47	30/621	GO:0007005	Mitochondrion organization	-6.45	38/509
GO:2001236	Regulation of extrinsic apoptotic signalling pathway	-3.47	12/151	R-HSA-72306	TRNA processing	-6.10	15/107
GO:0048588	Developmental cell growth	-3.17	15/234	GO:0009060	Aerobic respiration	-5.89	20/189
WP98	Prostaglandin synthesis and regulation	-3.13	6/45	R-HSA-3700989	Transcriptional Regulation by TP53	-5.60	29/365
GO:1901550	Regulation of endothelial cell development	-3.09	4/18	CORUM:282	SNF2h-cohesin-NuRD complex	-5.34	6/16
GO:0019934	CGMP-mediated signalling	-3.01	5/32	M84	PID ATM PATHWAY	-5.22	8/34
R-HSA-5653656	Vesicle-mediated transport	-2.91	30/673	GO:0006412	Translation	-5.15	45/736
R-HSA-373752	Netrin-1 signalling	-2.88	6/50	R-HSA-1640170	Cell Cycle	-5.11	43/692
GO:0050678	Regulation of epithelial cell proliferation	-2.78	20/392	R-HSA-3108214	SUMOylation of DNA damage response and repair proteins	-4.71	11/77
GO:0018209	Peptidyl-serine modification	-2.76	18/338	GO:0000959	Mitochondrial RNA metabolic process	-4.71	9/51
WP3414	Initiation of transcription and translation elongation at the HIV-1 LTR	-2.71	5/37	GO:0035268	Protein mannosylation	-4.44	6/22
GO:0060249	Anatomical structure homeostasis	-2.71	17/314	GO:0006260	DNA replication	-4.28	21/262
GO:0006511	Ubiquitin-dependent protein catabolic process	-2.68	28/636	R-HSA-68952	DNA replication initiation	-4.28	4/8
GO:0006575	Cellular modified amino acid metabolic process	-2.67	12/186	hsa04141	Protein processing in endoplasmic reticulum	-4.24	16/169
GO:0035162	Embryonic hemopoiesis	-2.67	4/23	GO:0042273	Ribosomal large subunit biogenesis	-4.23	10/72

KG-1 'unique' IFNG 5AzaC

Table 29: The top 20 most significantly enriched pathways in KG-1 comparing combination treated cells control cells, according to the 'unique' up and down regulated transcript lists, sorted by significance post FDR. Column 'transcript' specifies the number of transcripts differentially expressed in the data, from the corresponding pathway. For ease of comparison, number of transcripts altered by IFNG only treated cells are presented in brackets.

		Upregulated				Downregulated	
Summary process ID	Description	Log10(P)	Transcript	Summary process ID	Description	Log10(P)	Transcript
GO:0001817	Regulation of cytokine production	-5.62	19/799	R-HSA-1474244	Extracellular matrix organization	-6.38	14/301
GO:0030099	Myeloid cell differentiation	-4.88	12/383	GO:0007015	Actin filament organization	-5.11	15/441
GO:0019221	Cytokine-mediated signalling pathway	-4.65	13/472	WP2858	Ectoderm differentiation	-4.42	8/144
R-HSA-6798695	Neutrophil degranulation	-4.58	13/480	WP3668	Hypothesized pathways in pathogenesis of cardiovascular disease	-4.23	4/25
R-HSA-549127	Organic cation transport	-4.45	3/10	GO:0010942	Positive regulation of cell death	-3.98	16/618
R-HSA-1280215	Cytokine Signalling in Immune system	-3.93	15/715	GO:0007264	Small GTPase mediated signal transduction	-3.87	14/503
GO:0120035	Regulation of plasma membrane bounded cell projection organization	-3.78	14/655	M63	PID AVB3 OPN PATHWAY	-3.85	4/31
GO:0032103	Positive regulation of response to external stimulus	-3.76	11/427	GO:0072073	Kidney epithelium development	-3.73	7/136
R-HSA-416482	G alpha (12/13) signalling events	-3.66	5/80	GO:2001233	Regulation of apoptotic signalling pathway	-3.54	11/356
GO:0030029	Actin filament-based process	-3.43	15/796	GO:2001267	Regulation of cysteine-type endopeptidase activity involved in apoptotic signalling pathway	-3.49	3/16
GO:0035455	Response to interferon-alpha	-3.43	3/21	R-HSA-163685	Integration of energy metabolism	-3.45	6/108
GO:0044546	NLRP3 inflammasome complex assembly	-3.43	3/21	R-HSA-388396	GPCR downstream signalling	-3.39	15/629
GO:0002764	Immune response-regulating signalling pathway	-3.39	11/472	GO:0009101	Glycoprotein biosynthetic process	-3.31	10/319
GO:0032088	Negative regulation of NF-kappa B transcription factor activity	-3.36	5/93	GO:0071379	Cellular response to prostaglandin stimulus	-3.19	3/20
GO:0032642	Regulation of chemokine production	-3.28	5/97	GO:0040017	Positive regulation of locomotion	-3.14	14/596
hsa00340	Histidine metabolism	-3.25	3/24	WP4754	IL-18 signalling pathway	-3.12	9/279
GO:0050778	Positive regulation of immune response	-3.22	12/576	GO:0099560	Synaptic membrane adhesion	-2.90	3/25
GO:0043410	Positive regulation of MAPK cascade	-2.76	10/481	GO:0050927	Positive regulation of positive chemotaxis	-2.90	3/25
GO:0051354	Negative regulation of oxidoreductase activity	-2.73	3/36	GO:0030855	Epithelial cell differentiation	-2.82	14/643

7.4 Optimisation of IFNG Dose Selection for Differential Responses in Kasumi-1 and KG-1 Cell Lines: A Comparative Viability Study with Daunorubicin Treatment
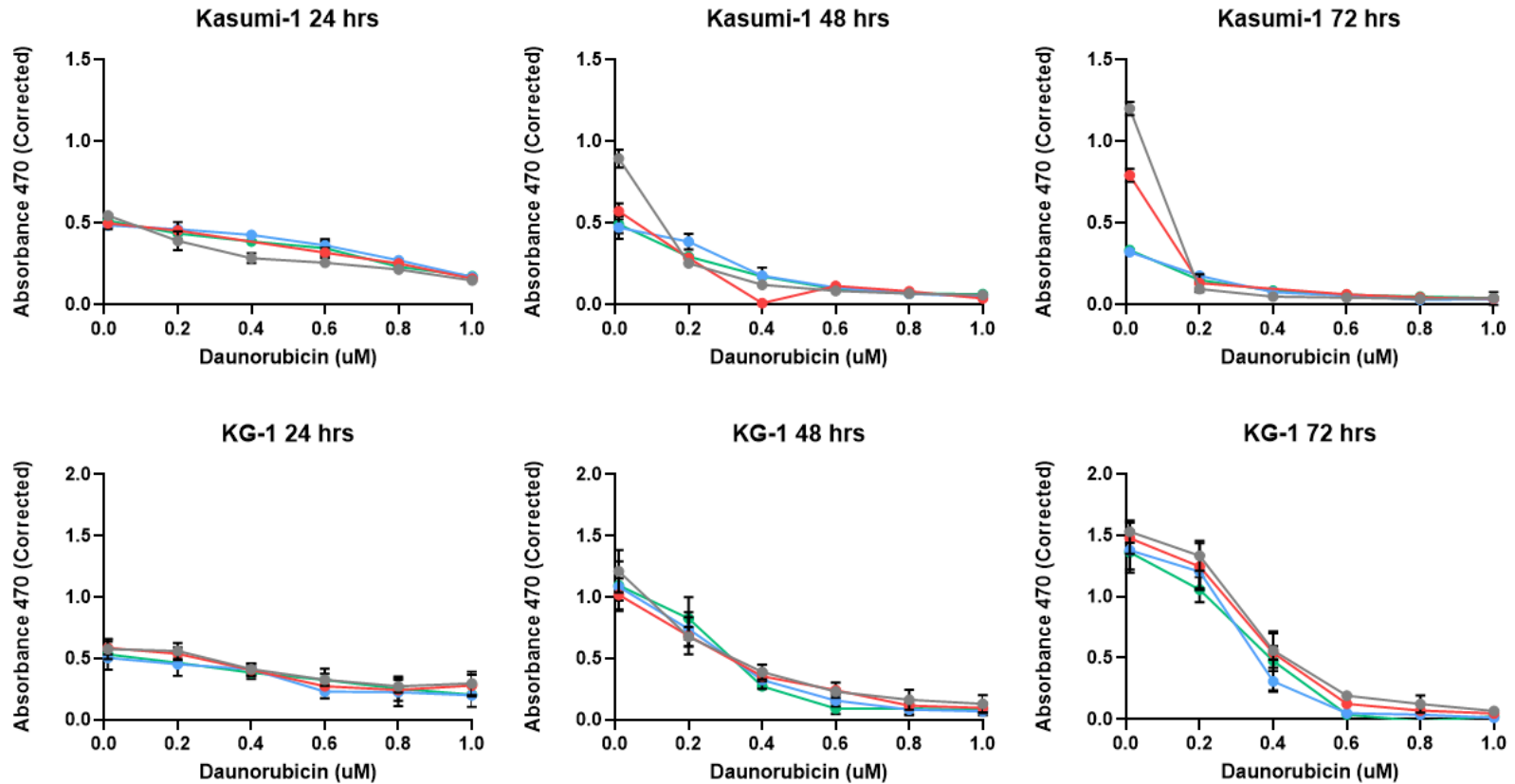
Comparison of how varying doses of IFNG changes cell line response to Daunorubicin, as measured by XTT assay at 24 hr intervals after Daunorubicin dosing for 3 days, n=3. Grey = 0 ng/mL, Red = 0.5 ng/mL, Blue = 5 ng/mL, and Green = 10 ng/mL IFNG.

7.5 Electropherograms showing RNA quality of samples used in next generation sequencing experiment generate in chapter 4

7.5.1 Gel and electropherogram of RNA extracted from untreated, IFNG treated, 5AzaC treated and IFNG and 5AzaC treated Kasumi-1 samples
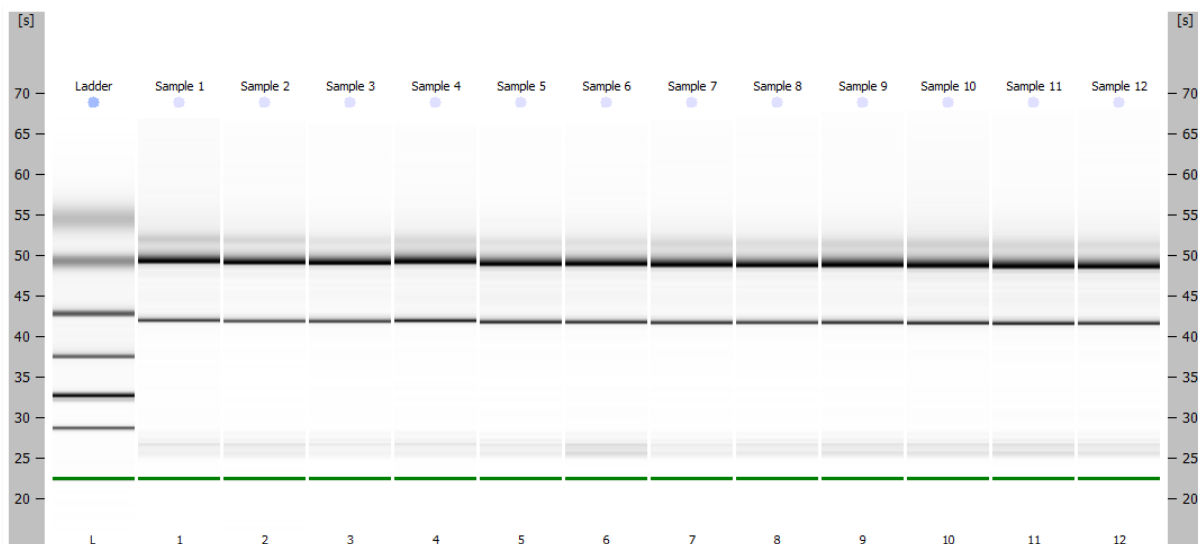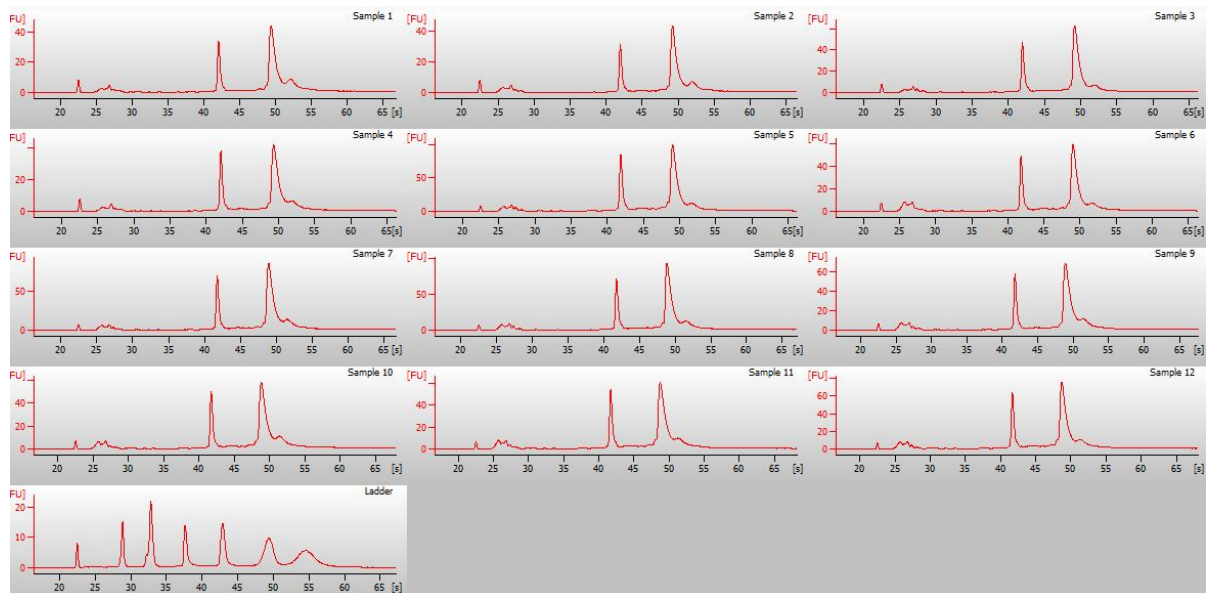
Samples 1 to 3 = Control untreated Kasumi-1

Samples 4 to 6 = IFNG treated Kasumi-1

Samples 7 to 9 = 5AzaC treated Kasumi-1

Samples 10 to 12 = IFNG and 5AzaC treated Kasumi-1

7.5.2 Gel and electropherogram of RNA extracted from untreated, IFNG treated, 5AzaC treated and IFNG and 5AzaC treated KG-1 samples
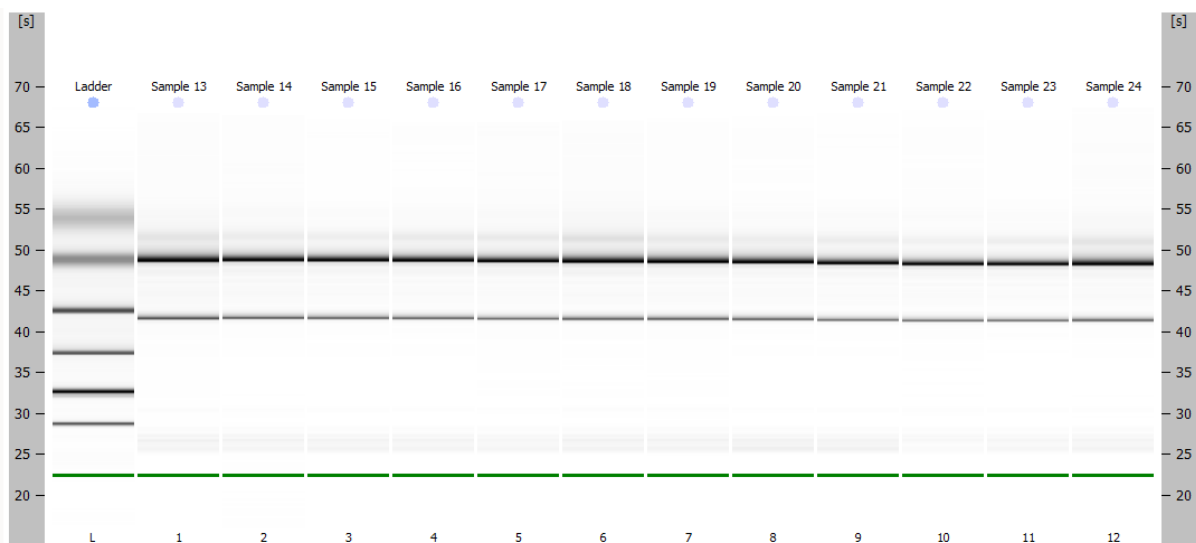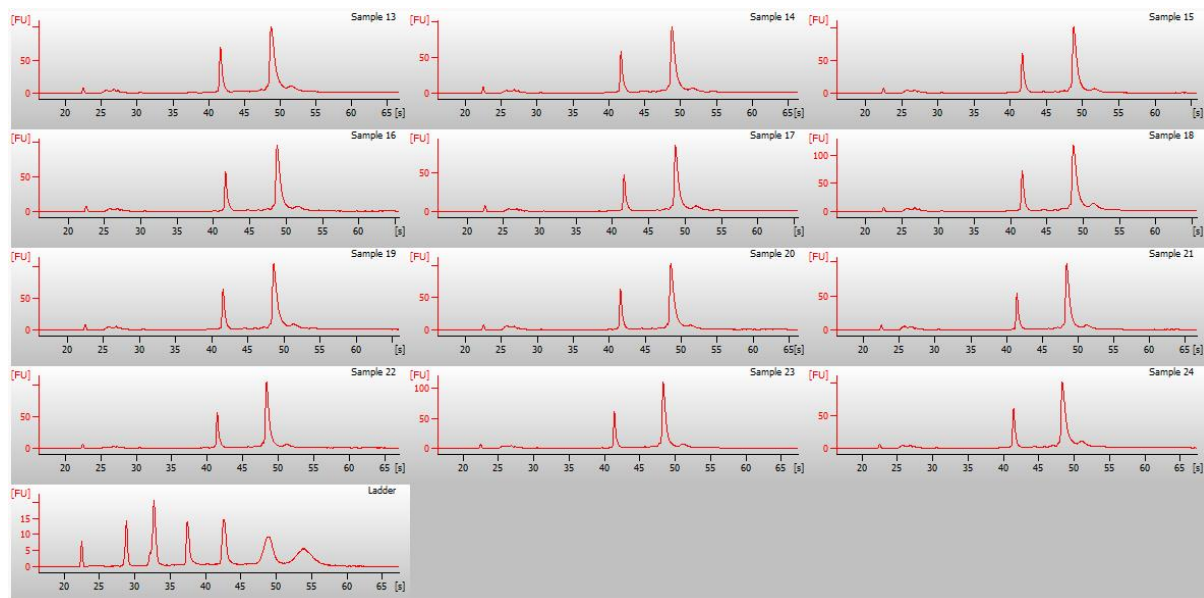
Samples 13 to 15 = IFNG and 5AzaC treated KG-1

Samples 16 to 18 = 5AzaC treated KG-1

Samples 19 to 21 = IFNG treated KG-1

Samples 21 to 24 = Control untreated KG-1

7.6 Normalised peak area expression of key proteins measured by SWATH-MS in all cell lines referenced in chapter 4
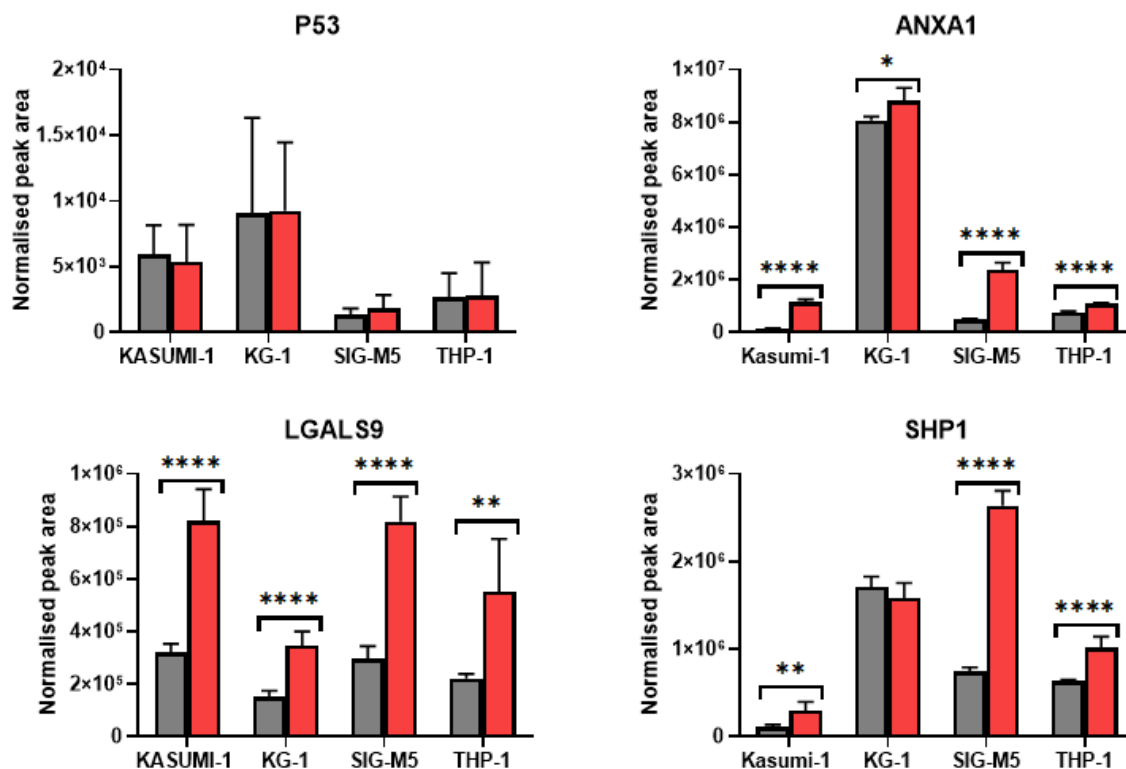
Normalised peak area expression of P53, ANXA1, LGALS9 and SHP1 proteins in cell lines, control compared to 48 hrs treatment 100 ng/ml IFNG (n=5-6). Grey; Control, Red; IFNG. Holm-Sidak method was used to calculate statistical significance between IFNG treated and untreated cells. * = P < 0.05, ** = P < 0.01, *** = P < 0.001 and **** = P < 0.0001.

7.7 Venn diagram comparison of significant transcripts determined by pairwise linear regression and ranked on number of appearances
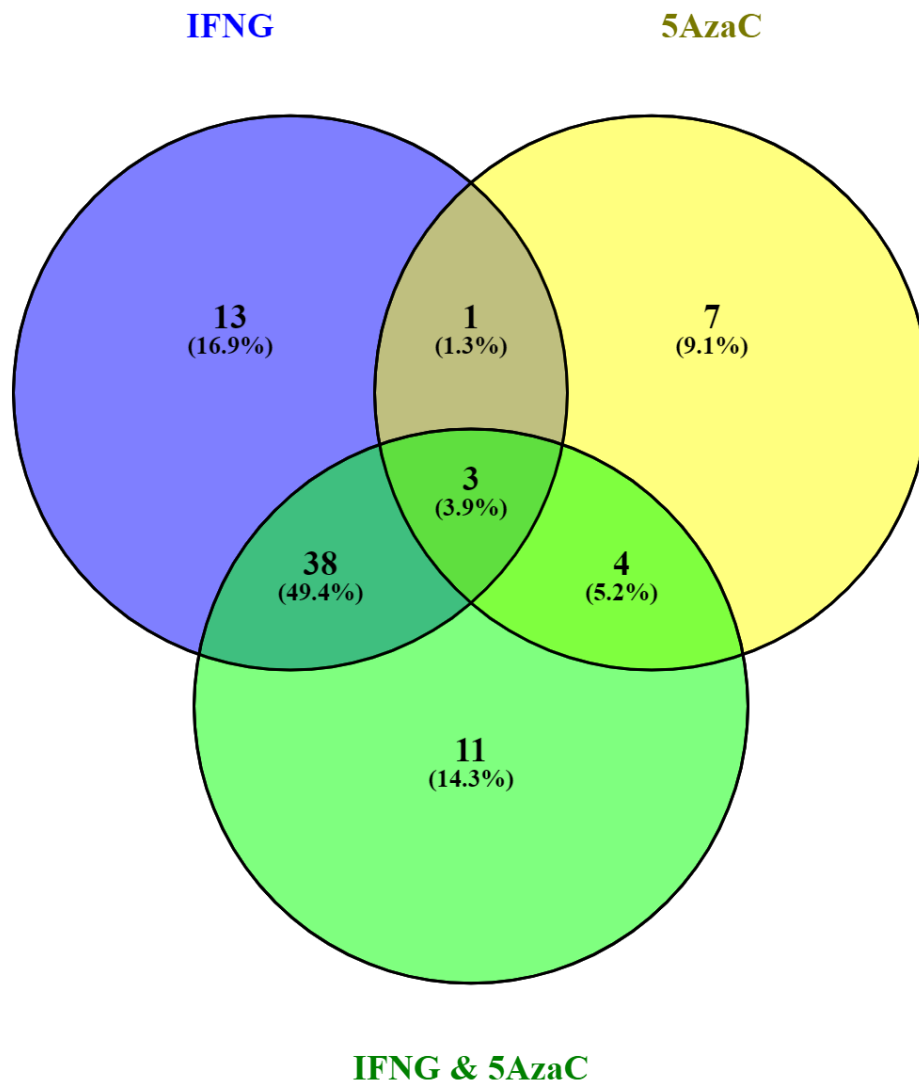
Venn diagram comparison of transcripts chosen in chapter 5 by pairwise linear regression and ranking on number of appearances

Table 30: Table of transcripts compared in venn diagram above showing which transcripts were shared between different treatments and which were unique to each treatment.

IFNG only	5AzaC only	IFNG & 5AzaC only	Shared IFNG, IFNG 5AzaC	Shared IFNG, 5AzaC	Shared 5AzaC, IFNG 5AzaC	Shared between all
FTL TMSB4X MT-CO3 RPS2 MT-ATP6 VAMP8 GLUL ADAR PRDX1 NUCB1 CALHM6 PSME2 ACTG1	MT-CO1 MT-ND4L MT-RNR1 YBX1 MT-TN TUBB LGALS1	MT-ND4 MT-RNR2 CLEC11A TRIM22 CHCHD2 DTX3L NCOA4 APOL6 ATF4 CD74 SPI1	IFI6 EEF1A1 TMSB10 IFITM1 HLA-B IFIT3 HLA-E STAT1 LY6E RPL8 B2M SRGN WARS ACTB GBP1 AL713998.1 RPL7 HLA-A LAP3 UBE2L6 TAP1 HLA-C PARP10 GBP5 UBA7 PSME1 GBP4 PABPC1 GPX1 PSMB9 RSAD2 PLEK GBP2 PSMB8 HLA-DRA EPSTI1 GSTP1 PIM1	NPM1	CALR RPL41P1 MT-ND3 S100A11	MT-CO2 BST2 GAPDH

7.8 Results of CPH univariate analysis on each list of candidates in order 5AzaC, IFNG and the IFNG & 5AzaC

Full table of results for CPH univariate analysis on each list of candidates in order 5AzaC, IFNG and the IFNG

Parameter Estimates (Data_tcga_rnaseq_5AzaC_list_Clinical-for Statistica)									
Target	Parameter Estimate	Standard Error	Chi-square	P value	95% Lower CL	95% Upper CL	Hazard Ratio	95% Hazard Ratio Lower CL	95% Hazard Ratio Upper CL
CALR	-0.000023	0.000007	11.34799	0.00076	-0.000037	-0.00001	0.999977	0.999963	0.99999
BST2	0.000434	0.000146	8.832139	0.00296	0.000148	0.00072	1.000434	1.000148	1.000721
LGALS1	0.000068	0.000025	7.25487	0.00707	0.000019	0.000118	1.000068	1.000019	1.000118
NPM1	-0.00005	0.000031	2.627771	0.10501	-0.00011	0.00001	0.99995	0.99989	1.00001
GAPDH	-0.000003	0.000004	0.554421	0.45652	-0.000012	0.000005	0.999997	0.999988	1.000005
TUBB	0.000008	0.000012	0.49476	0.48181	-0.000015	0.000032	1.000008	0.999985	1.000032
S100A11	0.000022	0.00004	0.30482	0.58088	-0.000057	0.000102	1.000022	0.999943	1.000102
YBX1	0.000005	0.000036	0.016649	0.89733	-0.000065	0.000074	1.000005	0.999935	1.000074

Parameter Estimates (Data_tcga_rnaseq_IFNG_list_Clinical-for Statistica)									
Target	Parameter Estimate	Standard Error	Chi-square	P value	95% Lower CL	95% Upper CL	Hazard Ratio	95% Hazard Ratio Lower CL	95% Hazard Ratio Upper CL
PIM1	0.000143	0.000042	11.505843	0.000694	0.000060	0.000225	1.000143	1.000060	1.000226
LY6E	0.000116	0.000036	10.696348	0.001073	0.000047	0.000186	1.000116	1.000047	1.000186
PARP10	0.000272	0.000087	9.645805	0.001898	0.000100	0.000443	1.000272	1.000100	1.000443

BST2	0.000434	0.000146	8.832139	0.002960	0.000148	0.000720	1.000434	1.000148	1.000721
PRDX1	0.000141	0.000051	7.505309	0.006152	0.000040	0.000242	1.000141	1.000040	1.000242
PSMB8	0.000515	0.000189	7.455528	0.006324	0.000145	0.000885	1.000515	1.000145	1.000885
IFITM1	0.000240	0.000094	6.554585	0.010461	0.000056	0.000424	1.000240	1.000056	1.000424
IFIT3	0.000468	0.000190	6.089966	0.013595	0.000096	0.000839	1.000468	1.000096	1.000839
GSTP1	0.000123	0.000054	5.144279	0.023323	0.000017	0.000229	1.000123	1.000017	1.000229
GPX1	0.000058	0.000027	4.813825	0.028232	0.000006	0.000111	1.000058	1.000006	1.000111
ADAR	0.000087	0.000041	4.587531	0.032205	0.000007	0.000167	1.000087	1.000007	1.000167
EPSTI1	0.000615	0.000295	4.332060	0.037401	0.000036	0.001194	1.000615	1.000036	1.001195
UBE2L6	0.000280	0.000148	3.589236	0.058155	-0.000010	0.000570	1.000280	0.999990	1.000570
B2M	-0.000010	0.000005	3.507687	0.061085	-0.000021	0.000000	0.999990	0.999979	1.000000
HLA-E	0.000045	0.000025	3.272423	0.070453	-0.000004	0.000094	1.000045	0.999996	1.000094
UBA7	0.000088	0.000051	2.952284	0.085756	-0.000012	0.000187	1.000088	0.999988	1.000187
PSME1	0.000211	0.000130	2.632272	0.104712	-0.000044	0.000465	1.000211	0.999956	1.000465
NPM1	-0.000050	0.000031	2.627771	0.105009	-0.000110	0.000010	0.999950	0.999890	1.000010
PSME2	0.000274	0.000175	2.440640	0.118228	-0.000070	0.000618	1.000274	0.999930	1.000618
RSAD2	0.000398	0.000255	2.436035	0.118576	-0.000102	0.000898	1.000398	0.999898	1.000898
EEF1A1	-0.000004	0.000003	2.158921	0.141744	-0.000010	0.000001	0.999996	0.999990	1.000001
RPL8	0.000026	0.000019	1.796913	0.180086	-0.000012	0.000063	1.000026	0.999988	1.000063
GBP2	0.000093	0.000071	1.726007	0.188922	-0.000046	0.000232	1.000093	0.999954	1.000232
LAP3	0.000119	0.000091	1.717012	0.190078	-0.000059	0.000297	1.000119	0.999941	1.000297
TAP1	0.000147	0.000117	1.591938	0.207049	-0.000082	0.000376	1.000147	0.999918	1.000376
FAM26F	0.000416	0.000373	1.242318	0.265025	-0.000315	0.001147	1.000416	0.999685	1.001148
PSMB9	0.000274	0.000258	1.123188	0.289233	-0.000233	0.000780	1.000274	0.999767	1.000781
TSMB10	0.000019	0.000021	0.810586	0.367947	-0.000022	0.000059	1.000019	0.999978	1.000059
IFI6	0.000084	0.000095	0.775994	0.378369	-0.000102	0.000269	1.000084	0.999898	1.000269

VAMP8	-0.000066	0.000076	0.765855	0.381502	-0.000214	0.000082	0.999934	0.999786	1.000082
ACTB	0.000002	0.000002	0.693469	0.404987	-0.000003	0.000007	1.000002	0.999997	1.000007
GBP4	0.000038	0.000051	0.556383	0.455721	-0.000062	0.000139	1.000038	0.999938	1.000139
GAPDH	-0.000003	0.000004	0.554421	0.456517	-0.000012	0.000005	0.999997	0.999988	1.000005
FTL	0.000004	0.000006	0.525032	0.468703	-0.000007	0.000016	1.000004	0.999993	1.000016
RPL7	0.000022	0.000033	0.425881	0.514017	-0.000044	0.000087	1.000022	0.999956	1.000087
GBP5	-0.000065	0.000108	0.364031	0.546276	-0.000278	0.000147	0.999935	0.999722	1.000147
PLEK	-0.000019	0.000032	0.358713	0.549222	-0.000082	0.000044	0.999981	0.999918	1.000044
GBP1	0.000061	0.000107	0.318115	0.572743	-0.000150	0.000271	1.000061	0.999850	1.000271
ACTG1	0.000003	0.000006	0.231360	0.630517	-0.000008	0.000014	1.000003	0.999992	1.000014
HLA-A	-0.000007	0.000014	0.211551	0.645555	-0.000035	0.000022	0.999993	0.999965	1.000022
HLA-DR	0.000002	0.000004	0.141253	0.707038	-0.000007	0.000010	1.000002	0.999993	1.000010
SRGN	-0.000001	0.000002	0.112406	0.737422	-0.000005	0.000004	0.999999	0.999995	1.000004
HLA-C	0.000005	0.000016	0.091857	0.761829	-0.000027	0.000037	1.000005	0.999973	1.000037
GLUL	0.000003	0.000011	0.084142	0.771760	-0.000019	0.000025	1.000003	0.999981	1.000025
RPS2	-0.000003	0.000015	0.043789	0.834247	-0.000032	0.000026	0.999997	0.999968	1.000026
NUCB1	0.000010	0.000059	0.028887	0.865041	-0.000106	0.000126	1.000010	0.999894	1.000126
PABPC1	0.000001	0.000006	0.025565	0.872967	-0.000011	0.000013	1.000001	0.999989	1.000013
WARS	0.000005	0.000032	0.022460	0.880870	-0.000059	0.000068	1.000005	0.999941	1.000068
STAT1	-0.000005	0.000046	0.009938	0.920592	-0.000094	0.000085	0.999995	0.999906	1.000085
HLA-B	0.000000	0.000008	0.001610	0.967999	-0.000016	0.000016	1.000000	0.999984	1.000016

Parameter Estimates (Data_tcga_rnaseq_IFNG & 5AzaC_list_Clinical-for Statistica)									
Target	Parameter Estimate	Standard Error	Chi-square	P value	95% Lower CL	95% Upper CL	Hazard Ratio	95% Hazard Ratio Lower CL	95% Hazard Ratio Upper CL
CLEC11A	-0.000083	0.000024	11.52432	0.000687	-0.000131	-0.000035	0.999917	0.999869	0.999965
PIM1	0.000143	0.000042	11.50584	0.000694	0.000060	0.000225	1.000143	1.000060	1.000226
CALR	-0.000023	0.000007	11.34799	0.000755	-3.7E-05	-0.00001	0.999977	0.999963	0.99999
LY6E	0.000116	0.000036	10.69635	0.001073	0.000047	0.000186	1.000116	1.000047	1.000186
PARP10	0.000272	0.000087	9.645805	0.001898	0.000100	0.000443	1.000272	1.000100	1.000443
BST2	0.000434	0.000146	8.832139	0.00296	0.000148	0.00072	1.000434	1.000148	1.000721
PSMB8	0.000515	0.000189	7.455528	0.006324	0.000145	0.000885	1.000515	1.000145	1.000885
IFITM1	0.000240	0.000094	6.554585	0.010461	0.000056	0.000424	1.000240	1.000056	1.000424
IFIT3	0.000468	0.00019	6.089966	0.013595	0.000096	0.000839	1.000468	1.000096	1.000839
GSTP1	0.000123	0.000054	5.144279	0.023323	0.000017	0.000229	1.000123	1.000017	1.000229
GPX1	0.000058	0.000027	4.813825	0.028232	0.000006	0.000111	1.000058	1.000006	1.000111
NCOA4	-0.000047	0.000022	4.423004	0.035458	-0.000091	-0.000003	0.999953	0.999909	0.999997
EPSTI1	0.000615	0.000295	4.332060	0.037401	0.000036	0.001194	1.000615	1.000036	1.001195
APOL6	0.000128	0.000063	4.048276	0.044216	0.000003	0.000252	1.000128	1.000003	1.000252
UBE2L6	0.000280	0.000148	3.589236	0.058155	-0.000010	0.000570	1.000280	0.999990	1.000570
B2M	-0.000010	0.000005	3.507687	0.061085	-0.000021	0.000000	0.999990	0.999979	1.000000
HLA-E	0.000045	0.000025	3.272423	0.070453	-4E-06	0.000094	1.000045	0.999996	1.000094
UBA7	0.000088	0.000051	2.952284	0.085756	-0.000012	0.000187	1.000088	0.999988	1.000187
PSME1	0.000211	0.000130	2.632272	0.104712	-0.000044	0.000465	1.000211	0.999956	1.000465
RSAD2	0.000398	0.000255	2.436035	0.118576	-0.000102	0.000898	1.000398	0.999898	1.000898
EEF1A1	-0.000004	0.000003	2.158921	0.141744	-0.000010	0.000001	0.999996	0.999990	1.000001
DTX3L	0.00013	0.000096	1.806424	0.178938	-5.9E-05	0.000319	1.00013	0.999941	1.000319
RPL8	0.000026	0.000019	1.796913	0.180086	-0.000012	0.000063	1.000026	0.999988	1.000063
GBP2	0.000093	0.000071	1.726007	0.188922	-0.000046	0.000232	1.000093	0.999954	1.000232

LAP3	0.000119	0.000091	1.717012	0.190078	-0.000059	0.000297	1.000119	0.999941	1.000297
TAP1	0.000147	0.000117	1.591938	0.207049	-0.000082	0.000376	1.000147	0.999918	1.000376
TRIM22	0.000074	0.000061	1.438299	0.230415	-0.000047	0.000194	1.000074	0.999953	1.000194
PSMB9	0.000274	0.000258	1.123188	0.289233	-0.000233	0.000780	1.000274	0.999767	1.000781
SPI1	0.000043	0.000046	0.860067	0.353720	-0.000048	0.000134	1.000043	0.999952	1.000134
ATF4	-0.000032	0.000036	0.810918	0.367849	-0.0001	0.000038	0.999968	0.999898	1.000038
TMSB10	0.000019	0.000021	0.810586	0.367947	-0.000022	0.000059	1.000019	0.999978	1.000059
IFI6	0.000084	0.000095	0.775994	0.378369	-0.000102	0.000269	1.000084	0.999898	1.000269
CD74	0.000001	0.000002	0.731103	0.392526	-0.000002	0.000005	1.000001	0.999998	1.000005
ACTB	0.000002	0.000002	0.693469	0.404987	-0.000003	0.000007	1.000002	0.999997	1.000007
GBP4	0.000038	0.000051	0.556383	0.455721	-0.000062	0.000139	1.000038	0.999938	1.000139
GAPDH	-0.000003	0.000004	0.554421	0.456517	-0.000012	0.000005	0.999997	0.999988	1.000005
RPL7	0.000022	0.000033	0.425881	0.514017	-0.000044	0.000087	1.000022	0.999956	1.000087
GBP5	-0.000065	0.000108	0.364031	0.546276	-0.000278	0.000147	0.999935	0.999722	1.000147
PLEK	-0.000019	0.000032	0.358713	0.549222	-0.000082	0.000044	0.999981	0.999918	1.000044
GBP1	0.000061	0.000107	0.318115	0.572743	-0.000150	0.000271	1.000061	0.999850	1.000271
S100A11	0.000022	0.000040	0.304820	0.580876	-0.000057	0.000102	1.000022	0.999943	1.000102
HLA-A	-0.000007	0.000014	0.211551	0.645555	-0.000035	0.000022	0.999993	0.999965	1.000022
HLA-DRA	0.000002	0.000004	0.141253	0.707038	-7E-06	0.00001	1.000002	0.999993	1.00001
SRGN	-0.000001	0.000002	0.112406	0.737422	-5E-06	0.000004	0.999999	0.999995	1.000004
HLA-C	0.000005	0.000016	0.091857	0.761829	-0.000027	0.000037	1.000005	0.999973	1.000037
CHCHD2	-0.000039	0.000131	0.089784	0.764452	-0.000297	0.000218	0.999961	0.999703	1.000218
PABPC1	0.000001	0.000006	0.025565	0.872967	-0.000011	0.000013	1.000001	0.999989	1.000013
WARS	0.000005	0.000032	0.022460	0.880870	-0.000059	0.000068	1.000005	0.999941	1.000068
STAT1	-0.000005	0.000046	0.009938	0.920592	-0.000094	0.000085	0.999995	0.999906	1.000085
HLA-B	0.000000	0.000008	0.001610	0.967999	-0.000016	0.000016	1.000000	0.999984	1.000016

7.9 Table of results for CPH regression model on each list of candidates in order 5AzaC, IFNG and the IFNG & 5AzaC using forward selection method

	Target	Parameter estimates for each model								
		Parameter Estimate	Standard Error	Chi-square	P value	95% Lower CL	95% Upper CL	Hazard Ratio	95% Hazard Ratio Lower CL	95% Hazard Ratio Upper CL
5AzaC list	BST2	0.00045	0.00015	9.71333	0.00183	0.00017	0.00074	1.00045	1.00017	1.00074
	CALR	-0.00002	0.00001	11.75729	0.00061	-0.00004	-0.00001	0.99998	0.99996	0.99999
	Target	Parameter Estimate	Standard Error	Chi-square	P value	95% Lower CL	95% Upper CL	Hazard Ratio	95% Hazard Ratio Lower CL	95% Hazard Ratio Upper CL
IFNG list	B2M	-0.000039	0.000008	23.94877	0.000001	-0.000054	-0.000023	0.999961	0.999946	0.999977
	GAPDH	-0.000022	0.000006	11.41136	0.000730	-0.000034	-0.000009	0.999978	0.999966	0.999991
	IFIT3	0.000334	0.000259	1.65957	0.197662	-0.000174	0.000841	1.000334	0.999826	1.000842
	LY6E	0.000117	0.000053	4.89105	0.026996	0.000013	0.000222	1.000117	1.000013	1.000222
	PIM1	0.000118	0.000048	5.99830	0.014320	0.000024	0.000213	1.000118	1.000024	1.000213
	PSMB8	0.001244	0.000259	23.10665	0.000002	0.000737	0.001751	1.001244	1.000737	1.001752
	SRGN	0.000007	0.000002	8.46569	0.003619	0.000002	0.000011	1.000007	1.000002	1.000011
	TMSB10	0.000095	0.000027	11.91718	0.000556	0.000041	0.000148	1.000095	1.000041	1.000148
	B2M	-0.000039	0.000008	23.94877	0.000001	-0.000054	-0.000023	0.999961	0.999946	0.999977

	Target	Parameter Estimate	Standard Error	Chi-square	P value	95% Lower CL	95% Upper CL	Hazard Ratio	95% Hazard Ratio Lower CL	95% Hazard Ratio Upper CL
IFNG & 5AzaC list	BST2	0.000470	0.000186	6.400533	0.011409	0.000106	0.000834	1.000470	1.000106	1.000834
	CALR	-0.000011	0.000008	1.846817	0.174154	-0.000027	0.000005	0.999989	0.999973	1.000005
	CLEC11A	-0.000084	0.000029	8.269813	0.004031	-0.000142	-0.000027	0.999916	0.999858	0.999973
	LY6E	0.000114	0.000043	7.014869	0.008084	0.000030	0.000198	1.000114	1.000030	1.000198
	PIM1	0.000096	0.000046	4.434329	0.035223	0.000007	0.000186	1.000096	1.000007	1.000186

7.10 Table for each patient data set used describing distribution of clinical features

Patient data sets					
	TCGA	**German AML**	**CN-AML**	**HOVON**	**Beat AML**
Gene-expression platform	RNA sequencing	Affymetrix Human Genome U133 Plus 2.0, U133A/B Array	Affymetrix Human Genome U133 Plus 2.0, U133A/B Array	Affymetrix Human Genome U133 Plus 2.0 Microarray	RNA sequencing
Patients, n	128	535	242	593	242
Males/females split, n	70/58	N.A.	108/134	306/287	109/133
Median follow-up, y	3.25	8.61	8.84	8.34	1.27
Age group, n					
0-14 y	0	0	0	0	10
15-39 y	25	93	33	188	36
40-59 y	41	185	86	327	64
≥60 y	60	257	123	78	132
N.A.	2	0	0	0	0
WHO category					
AML with minimal differentiation/without maturation	41	N.A.	N.A.	161	6
AML with t(8;21)(q22;q22)	7	N.A.	N.A.	46	9
AML with inv(16)(p13;q22)	9	N.A.	N.A.	49	17
AML with biallelic mutations of CEBPA	N.A.	N.A.	N.A.	28	17
AML with mutated *NPM1*	N.A.	N.A.	N.A.	183	64
AML with t(9;11)(p22;q23)	7	N.A.	N.A.	21	8
AML with inv(3)(q21;q26)/t(3;3)(q21;q26)	N.A.	N.A.	N.A.	15	4
Acute myelomonocytic	18	N.A.	N.A.	112	13
Acute monoblastic/monocytic leukemia	13	N.A.	N.A.	139	11

Pure erythroid leukaemia	N.A.	N.A.	N.A.	9	1
AML with myelodysplasia-related	N.A.	N.A.	N.A.	N.A	45
AML, not otherwise specified	N.A.	N.A.	N.A.	N.A	27
Median presenting WBC count	15 (1-224)	N.A.	N.A.	N.A	24.2 (0.5-
Median percentage of BM blasts	72 (30-100)	N.A.	N.A.	73 (0-98)	71 (1-98)
<td colspan="6" align="center">**ELN cytogenetic risk group, n (%)**</td>
Favourable	18 (14)	N.A.	0	201 (34)	73 (30)
Intermediate	76 (59)	223	242	263 (44)	80 (33)
Adverse	32 (25)	N.A.	0	126 (21)	89 (37)
N.A.	2 (2)	0	0	3 (1)	0
<td colspan="6" align="center">**Patients with haematopoietic stem cell transplant, n**</td>
Autologous	4	N.A.	N.A.	95	0
Allogeneic	55	N.A.	N.A.	181	75

7.11 Formulas for PI

Where N_T as a prefix to a transcript means min max normalised value of transcript stated, scores were multiplied by 10,000 for ease of use.

5AzaC PI = ((0.00045 x N_T*BST2)* + (-0.00002 x N_T*CALR*)) x 10,000

IFNG PI = ((-0.00004 x N_T*B2M*) + (0.00057 x N_T*IFIT3*) + (-0.00010 x N_T*NPM1*) + (0.00015 x N_T*PIM1*) + (0.00012 x N_T*PRDX1*) + (0.00135 x N_T*PSMB8)* + *(*0.00015 x N_T*RPL7*) + (-0.00004 x N_T*RPS2*) + (0.00008 x N_T*TMSB10*)) x 10,000

IFNG5AzaC PI = ((0.00047 x N_T*BST2*) + (-0.00001 x N_T*CALR*) + (-0.00008 x N_T*CLEC11A*) + (0.00011 x N_T*LY6E*) + (0.00010 x N_T*PIM1*)) x 10,000

7.12 KM plots for each individual transcript in each PI using a median split for OS and EFS in the TCGA data set

7.12.1 5AzaC PI transcripts KM Plots – Overall survival

BST2

Chi square 2.583
P value 0.1080

BST < median 26.30
BST > median 11.30

CALR

Chi square 8.186
P value 0.0042

CALR > median 30.60
CALR < median 14.50

BST2

Chi square 0.7557
P value 0.3847

BST < median 17.30
BST > median 14.90

CALR

Chi square 3.331
P value 0.0680

CALR > median 24.10
CALR < median 13.40

7.12.2   IFNG PI transcripts KM Plots – Overall survival

B2M

Chi square   1.041
P value   0.3076

B2M > median   18.65

B2M < median   18.50

IFIT3

Chi square   4.432
P value   0.0353

IFIT3 < median   25.75

IFIT3 > median   11.80

GAPDH

Chi square   1.428
P value   0.2321

GAPDH > median   17.75

GAPDH < median   19.00

PIM1

Chi square   6.458
P value   0.0110

PIM1 < median   26.30

PIM1 > median   11.50

LY6E

Chi square   0.8337
P value   0.3612

LY6E < median   23.20

LY6E > median   15.50

PSMB8

Chi square   9.310
P value   0.0023

PSMB8 < median   30.60

PSMB8 > median   11.10

SRGN

Chi square   0.0005311
P value   0.9816

SRGN > median   17.55

SRGN < median   19.00

TMSB10

Chi square   0.1267
P value   0.7219

TMSB10 > median   18.50

TMSB10 < median   17.40

7.12.3   IFNG PI transcripts KM Plots – Event free survival

B2M

Chi square   0.8495
P value   0.3567

B2M > median   18.20

B2M < median   14.60

IFIT3

Chi square   6.322
P value   0.0119

IFIT3 < median   26.20

IFIT3 > median   12.00

GAPDH

Chi square   2.594
P value   0.1073

GAPDH < median   14.20

GAPDH > median   24.10

PIM1

Chi square   4.974
P value   0.0257

PIM1 < median   24.10

PIM1 > median   12.00

LY6E

Chi square   0.7003
P value   0.4027

LY6E < median   14.60

LY6E > median   22.20

PSMB8

Chi square   4.123
P value   0.0423

PSMB8 < median   24.20

PSMB8 > median   14.90

SRGN

Chi square   0.2395
P value   0.6246

SRGN > median   13.80

SRGN < median   18.20

TMSB10

Chi square   0.09091
P value   0.7630

TMSB10 > median   16.60

TMSB10 < median   16.10

7.12.4   IFNG5AzaC PI transcripts KM Plots – Overall survival

PIM1

Chi square   5.656
P value   0.0174
PIM1 < median   24.10
PIM1 > median   12.00

BST2

Chi square   1.936
P value   0.1641
BST2 < median   17.30
BST2 > median   14.90

CALR

Chi square   2.621
P value   0.1055
CALR > median   24.10
CALR < median   13.40

CLEC11A

Chi square   6.091
P value   0.0136
CLC11A > median   35.60
CLEC11A < median   12.00

LY6E

Chi square   0.7003
P value   0.4027
LY6E > median   22.20
LY6E < median   14.60

7.13 Survival over timetables corresponding to KM plots

7.13.1 5AzaC PI transcript survival tables for patients with each transcript expressed above and below median levels in the TCGA data set, left table OS, right table EFS

Transcript	Number at risk (OS - median split)							
	Months	0	20	40	60	80	100	120
BST2	<Median	86	48	25	11	6	1	0
	>Median	86	33	19	10	4	1	0
CALR	<Median	86	31	14	6	2	0	0
	>Median	86	50	30	15	8	2	0

Transcript	Number at risk (EFS - median split)							
	Months	0	20	40	60	80	100	120
BST2	<Median	85	27	15	8	4	0	0
	>Median	85	26	15	9	4	1	0
CALR	<Median	85	18	9	5	2	0	0
	>Median	85	35	21	12	6	1	0

7.13.2 IFNG PI transcript survival tables for patients with each transcript expressed above and below median levels in the TCGA data set, left table OS, right table EFS

Number at risk (OS - median split)								
Transcript	Months	0	20	40	60	80	100	120
B2M	<Median	86	39	17	7	3	1	0
B2M	>Median	86	42	27	14	7	1	0
GAPDH	<Median	86	40	17	6	4	1	0
GAPDH	>Median	86	41	27	15	6	1	0
IFIT3	<Median	86	50	28	13	6	0	0
IFIT3	>Median	86	31	16	8	4	2	0
LY6E	<Median	86	45	26	11	5	1	0
LY6E	>Median	86	36	18	10	5	1	0
PIM1	<Median	86	50	25	14	7	1	0
PIM1	>Median	86	31	19	7	3	1	0
PSMB8	<Median	86	51	30	14	8	1	0
PSMB8	>Median	86	30	14	7	2	1	0
SRGN	<Median	86	39	20	10	3	0	0
SRGN	>Median	86	42	24	11	7	2	0
TMSB10	<Median	86	40	18	5	2	0	0
TMSB10	>Median	86	41	26	16	8	2	0

Number at risk (EFS - median split)								
Transcript	Months	0	20	40	60	80	100	120
B2M	<Median	85	24	11	5	1	0	0
B2M	>Median	85	29	19	12	7	1	0
GAPDH	<Median	85	21	11	4	3	0	0
GAPDH	>Median	85	32	19	13	5	1	0
IFIT3	<Median	85	35	22	12	5	0	0
IFIT3	>Median	85	18	8	5	3	1	0
LY6E	<Median	85	25	18	9	4	0	0
LY6E	>Median	85	28	12	8	4	1	0
PIM1	<Median	85	34	19	13	7	1	0
PIM1	>Median	85	19	11	4	1	0	0
PSMB8	<Median	85	35	20	11	6	0	0
PSMB8	>Median	85	18	10	6	2	1	0
SRGN	<Median	85	27	15	9	3	0	0
SRGN	>Median	85	26	15	8	5	1	0
TMSB10	<Median	85	26	11	5	2	0	0
TMSB10	>Median	85	27	19	12	6	1	0

7.13.3   IFNG5AzaC PI transcript survival tables for patients with each transcript expressed above and below median levels in the TCGA data set, left table OS, right table EFS

Transcript	Number at risk (OS - median split)							
	Months	0	20	40	60	80	100	120
CALR	<Median	86	31	14	6	2	0	0
	>Median	86	50	30	15	8	2	0
LY6E	<Median	86	45	26	11	5	1	0
	>Median	86	36	18	10	5	1	0
PIM1	<Median	86	50	25	14	7	1	0
	>Median	86	31	19	7	3	1	0
CLEC11A	<Median	86	30	12	4	3	1	0
	>Median	86	51	32	17	7	1	0
BST2	<Median	86	48	25	11	6	1	0
	>Median	86	33	19	10	4	1	0

Transcript	Number at risk (EFS - median split)							
	Months	0	20	40	60	80	100	120
CALR	<Median	85	18	9	5	2	0	0
	>Median	85	35	21	12	6	1	0
LY6E	<Median	85	25	18	9	4	0	0
	>Median	85	28	12	8	4	1	0
PIM1	<Median	85	34	19	13	7	1	0
	>Median	85	19	11	4	1	0	0
CLEC11A	<Median	85	15	6	3	2	0	0
	>Median	85	38	24	14	6	1	0
BST2	<Median	85	27	15	8	4	0	0
	>Median	85	26	15	9	4	1	0

7.14 KM plots of PI scores in subsets of AML patients

7.14.1 KM plots of all PI scores in HOVON patients with mutant *NPM1*

5AzaC PI *NPM1* Mutated OS HOVON
Chi square 0.06128
P value 0.8045
Low 5AzaC PI 21.06
High 5AzaC PI 22.64

5AzaC PI *NPM1* Mutated EFS HOVON
Chi square 5.380e-007
P value 0.9994
Low 5AzaC PI 12.65
High 5AzaC PI 13.77
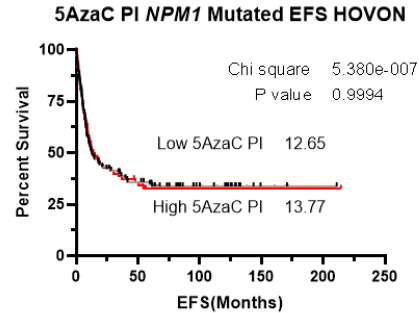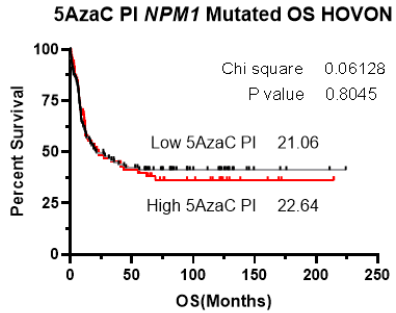
Number at risk (OS)							
Month	0	40	80	120	160	200	240
Low	111	48	32	19	5	3	1
High	72	32	17	13	6	3	1

HR = 1.049

Cut off threshold = 3.1

Number at risk (EFS)							
Month	0	40	80	120	160	200	240
Low	111	41	27	15	3	2	1
High	72	27	14	11	4	2	1

HR =1

Cut off threshold = 3.1

IFNG PI *NPM1* Mutated OS HOVON
Chi square 2.619
P value 0.1056
Low IFNG PI 33.38
High IFNG PI 13.86

IFNG PI *NPM1* Mutated EFS HOVON
Chi square 2.610
P value 0.1062
Low IFNG PI 14.92
High IFNG PI 9.430
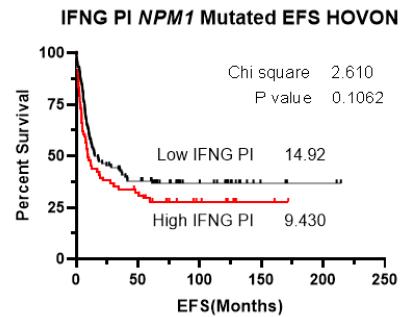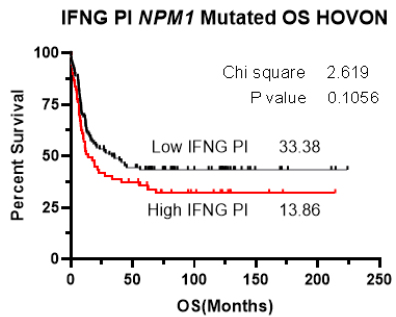
Number at risk (EFS)							
Month	0	40	80	120	160	200	240
Low	116	53	33	23	7	4	1
High	67	27	16	9	4	2	1

HR = 1.367

Cut off threshold =10.15

Number at risk (EFS)							
Month	0	40	80	120	160	200	240
Low	112	43	27	17	5	3	1
High	71	25	14	8	2	1	1

HR = 1.345

Cut off threshold = 10.02

IFNG5AzaC PI *NPM1* Mutated OS HOVON
Chi square 2.653
P value 0.1034
Low IFNG5AzaC PI 33.38
High IFNG5AzaC PI 16.20

IFNG5AzaC PI *NPM1* Mutated EFS HOVON
Chi square 2.137
P value 0.1437
Low IFNG5AzaC PI 17.81
High IFNG5AzaC PI 10.45
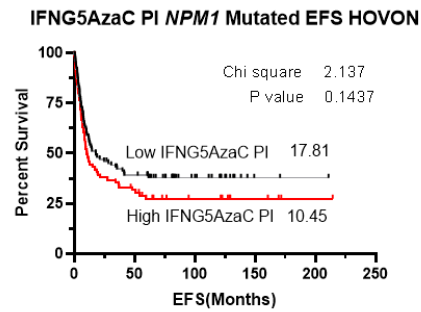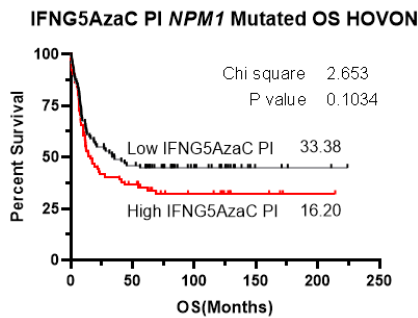
Number at risk (OS)							
Month	0	40	80	120	160	200	240
Low	102	49	31	18	5	3	1
High	81	31	18	14	6	3	1

HR = 1.362

Cut off threshold = 3.64

Number at risk (EFS)							
Month	0	40	80	120	160	200	240
Low	102	41	26	14	3	2	1
High	81	27	15	12	5	2	1

HR = 1.303

Cut off threshold = 3.64
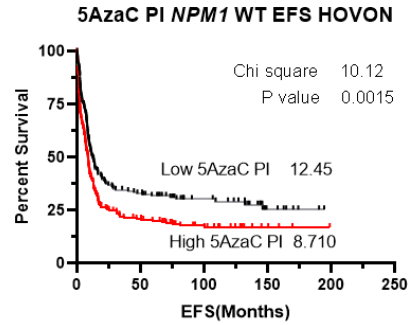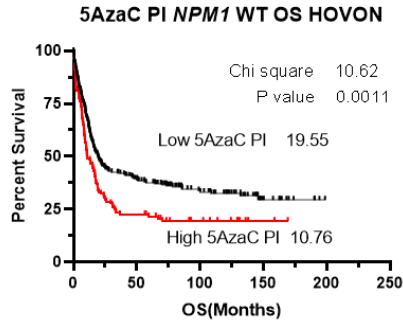
5AzaC PI *NPM1* WT OS HOVON

Chi square 10.62
P value 0.0011
Low 5AzaC PI 19.55
High 5AzaC PI 10.76

5AzaC PI *NPM1* WT EFS HOVON

Chi square 10.12
P value 0.0015
Low 5AzaC PI 12.45
High 5AzaC PI 8.710

Number at risk (OS)					
Month	0	40	80	120	160	200
Low	293	118	67	41	11	1
High	112	26	17	10	2	1

HR = 1.510

Cut off threshold = 2.98

Number at risk (EFS)					
Month	0	40	80	120	160	200
Low	186	62	37	27	7	1
High	219	48	30	18	3	1

HR = 1.438

Cut off threshold = 2.66

IFNG PI *NPM1* WT OS HOVON

Chi square 0.4384
P value 0.5079
Low IFNG PI 17.71
High IFNG PI 16.66

IFNG PI *NPM1* WT EFS HOVON

Chi square 0.4851
P value 0.4861
Low IFNG PI 9.660
High IFNG PI 9.790
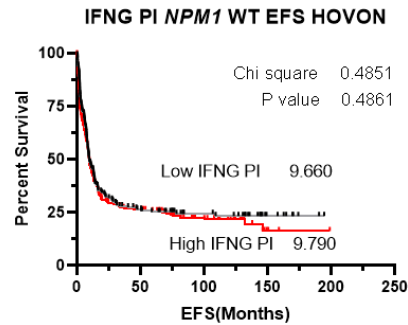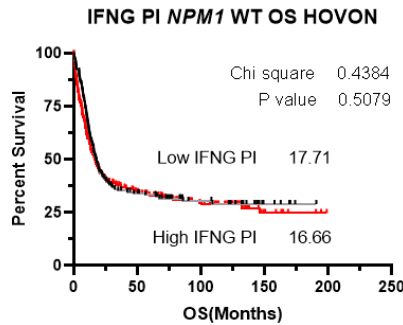
Number at risk (OS)					
Month	0	40	80	120	160	200
Low	149	52	31	21	6	1
High	256	92	53	31	7	1

HR = 1.085

Cut off threshold = 9.36

Number at risk (EFS)					
Month	0	40	80	120	160	200
Low	194	53	30	24	8	1
High	211	37	37	21	2	1

HR = 1.082

Cut off threshold = 10.11
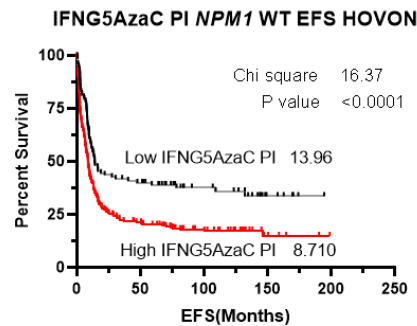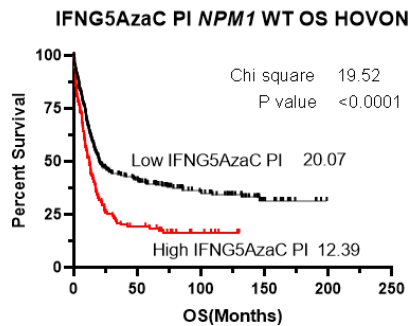
IFNG5AzaC PI *NPM1* WT OS HOVON

Chi square 19.52
P value <0.0001
Low IFNG5AzaC PI 20.07
High IFNG5AzaC PI 12.39

IFNG5AzaC PI *NPM1* WT EFS HOVON

Chi square 16.37
P value <0.0001
Low IFNG5AzaC PI 13.96
High IFNG5AzaC PI 8.710

Number at risk (OS)					
Month	0	40	80	120	160	200
Low	285	118	71	47	12	1
High	120	25	13	5	1	1

HR = 1.716

Cut off threshold = 3.72

Number at risk (EFS)					
Month	0	40	80	120	160	200
Low	115	46	28	22	6	1
High	290	64	39	23	4	1

HR = 1.708

Cut off threshold = 2.94

5AzaC PI mutated *NPM1* BeatAML

Chi square 4.081
P value 0.0434

High 5AzaC PI 20.86

Low 5AzaC PI 8.520

5AzaC PI *NPM1* WT BeatAML

Chi square 2.880
P value 0.0897

Low 5AzaC PI 14.74
High 5AzaC PI 12.60
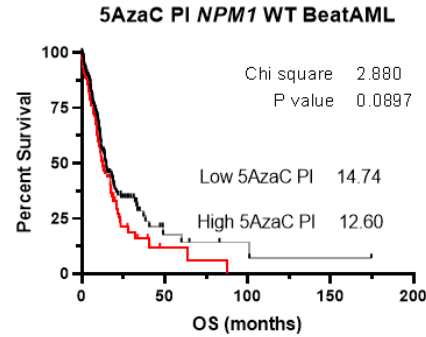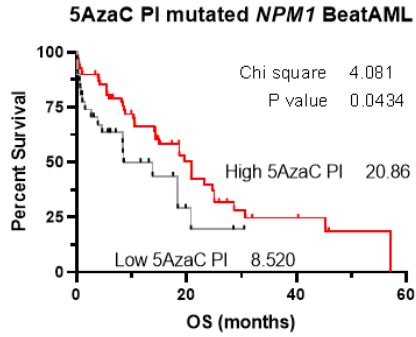
Number at risk (OS)							
Month	0	10	20	30	40	50	60
Low	31	11	4	2	1	1	1
High	69	39	20	9	6	2	1

HR = 0.5748

Cut off threshold = 0.5

Number at risk (EFS)						
Month	0	40	80	120	160	200
Low	184	11	4	2	2	1
High	118	5	2	1	1	1

HR = 1.289

Cut off threshold = 1.23

IFNG PI mutated *NPM1* BeatAML

Chi square 1.920
P value 0.1659

Low IFNG PI 20.82

High IFNG PI 13.82

IFNG PI *NPM1* WT BeatAML

Chi square 13.96
P value 0.0002

Low IFNG PI 18.26
High IFNG PI 10.53
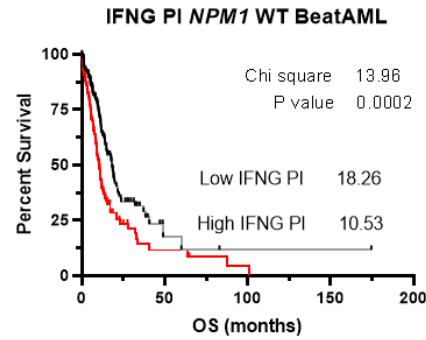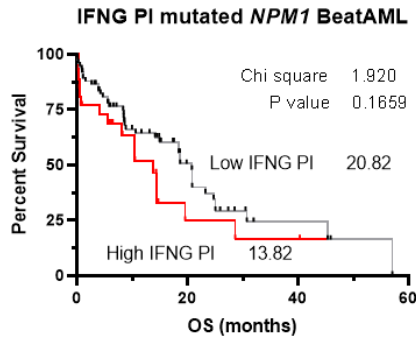
Number at risk (OS)							
Month	0	10	20	30	40	50	60
Low	74	38	20	8	4	2	1
High	26	11	4	3	2	1	1

HR = 1.492

Cut off threshold = 4.23
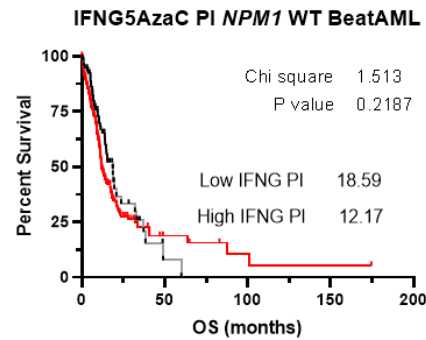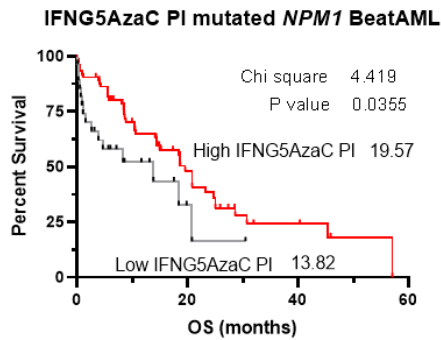
Number at risk (EFS)						
Month	0	40	80	120	160	200
Low	181	10	3	2	2	1
High	121	6	3	1	1	1

HR = 1.730

Cut off threshold = 4.62

IFNG5AzaC PI mutated *NPM1* BeatAML

Chi square 4.419
P value 0.0355

High IFNG5AzaC PI 19.57

Low IFNG5AzaC PI 13.82

IFNG5AzaC PI *NPM1* WT BeatAML

Chi square 1.513
P value 0.2187

Low IFNG PI 18.59
High IFNG PI 12.17

Number at risk (OS)							
Month	0	10	20	30	40	50	60
Low	27	9	3	1	1	1	1
High	73	41	21	8	6	2	1

HR = 0.5452

Cut off threshold = 0.48

Number at risk (EFS)						
Month	0	40	80	120	160	200
Low	90	4	1	1	1	1
High	212	12	5	2	2	1

HR = 1.232

Cut off threshold = 0.89

7.14.4 KM plots of all PI scores in HOVON patients with *FLT3-ITD (OS and EFS)*

5AzaC PI FLT3-ITD OS HOVON

Chi square 2.923
P value 0.0873

Low 5AzaC PI 11.50

High 5AzaC PI 8.068

5AzaC PI FLT3-ITD EFS HOVON

Chi square 2.060
P value 0.1512

Low 5AzaC PI 7.100

High 5AzaC PI 5.095
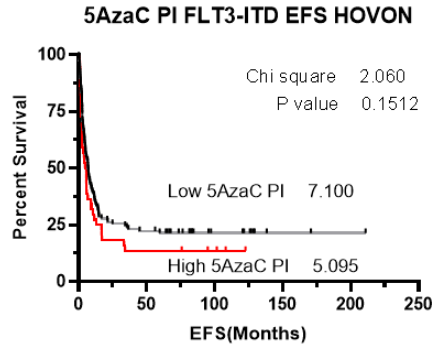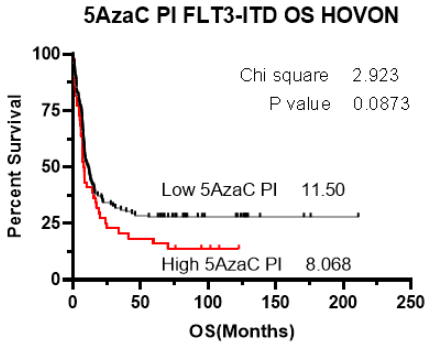
Number at risk (OS)						
Month	0	40	80	120	160	200
Low	126	39	23	14	4	1
High	44	9	6	3	1	1

HR = 1.388

Cut off threshold = 3.12

Number at risk (EFS)						
Month	0	40	80	120	160	200
Low	126	30	19	11	3	1
High	44	7	6	3	1	1

HR = 1.313

Cut off threshold = 3.12

IFNG PI FLT3-ITD OS HOVON

Chi square 3.238
P value 0.0719

Low IFNG PI 13.01

High IFNG PI 8.380

IFNG PI FLT3-ITD EFS HOVON

Chi square 3.478
P value 0.0622

Low IFNG PI 7.230

High IFNG PI 5.060
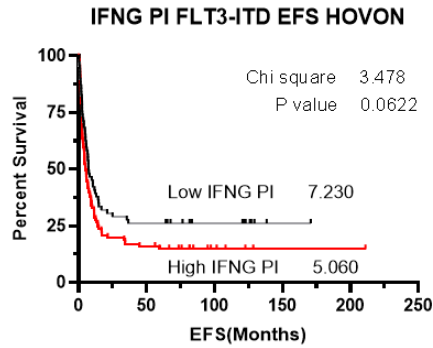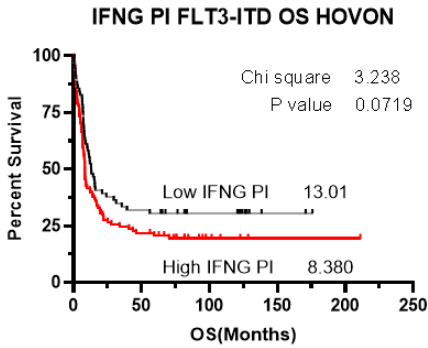
Number at risk (OS)						
Month	0	40	80	120	160	200
Low	126	30	19	11	3	1
High	44	7	6	3	1	1

HR = 1.385

Cut off threshold = 9.44

Number at risk (EFS)						
Month	0	40	80	120	160	200
Low	69	19	13	10	2	1
High	101	18	12	4	2	1

HR = 1.387

Cut off threshold = 9.44

IFNG5AzaC PI FLT3-ITD OS HOVON

Chi square 3.701
P value 0.0544

Low IFNG PI 11.04

High IFNG PI 8.310

IFNG5AzaC PI FLT3-ITD EFS HOVON

Chi square 3.174
P value 0.0748

Low IFNG5AzaC PI 7.129

High IFNG5AzaC PI 4.845
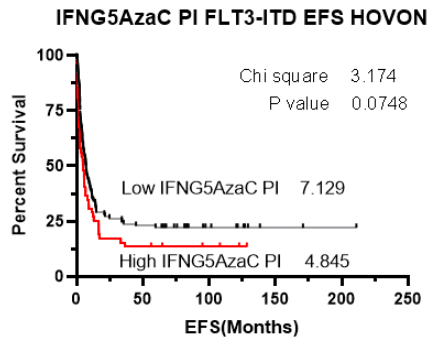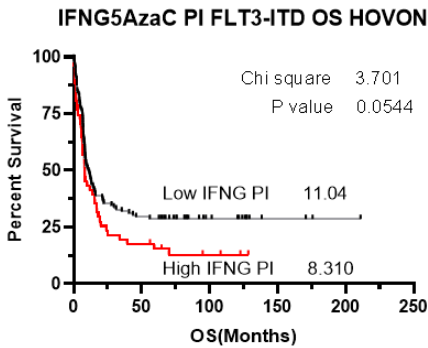
Number at risk (OS)						
Month	0	40	80	120	160	200
Low	119	39	23	13	4	1
High	51	10	6	4	1	1

HR = 1.426

Cut off threshold = 3.93

Number at risk (EFS)						
Month	0	40	80	120	160	200
Low	118	29	19	10	3	1
High	52	8	6	4	1	1

HR = 1.380

Cut off threshold = 3.92

Bibliography

Ahn, J., et al., 2017. 5-Hydroxymethylcytosine correlates with epigenetic regulatory mutations, but may not have prognostic value in predicting survival in normal karyotype acute myeloid leukemia. *Oncotarget,* 8 (5), 8305-8314.

Al-Ali, H.K., Jaekel, N. and Niederwieser, D., 2014. The role of hypomethylating agents in the treatment of elderly patients with AML. *Journal of Geriatric Oncology,* 5 (1), 89-105.

Ali, R., et al., 2017. Isolation and characterization of a new naturally immortalized human breast carcinoma cell line, KAIMRC1. *BMC Cancer,* 17 (1), 1-13.

Al-Matary, Y.S., et al., 2016. Acute myeloid leukemia cells polarize macrophages towards a leukemia supporting state in a Growth factor independence 1 dependent manner. *Haematologica,* 101 (10), 1216.

Amaldi, I., et al., 1989. Induction of HLA class II genes by IFN-gamma is transcriptional and requires a trans-acting protein. *Journal of Immunology (Baltimore, Md.: 1950),* 142 (3), 999-1004.

Ambrosi, C., Manzo, M. and Baubec, T., 2017. Dynamics and context-dependent roles of DNA methylation. *Journal of Molecular Biology,* 429 (10), 1459-1475.

American Cancer Society, 2018. *Acute Myeloid Leukemia (AML) Subtypes and Prognostic Factors* [online]. American Cancer Society. Available at: https://www.cancer.org/cancer/acute-myeloid-leukemia/detection-diagnosis-staging/how-classified.html#references [Accessed 17/04 2020].

Anderson, K.G., Stromnes, I.M. and Greenberg, P.D., 2017. Obstacles posed by the tumor microenvironment to T cell activity: a case for synergistic therapies. *Cancer Cell,* 31 (3), 311-325.

Anguille, S., et al., 2017. Dendritic cell vaccination as postremission treatment to prevent or delay relapse in acute myeloid leukemia. *Blood, the Journal of the American Society of Hematology,* 130 (15), 1713-1721.

Appelbaum, F.R., et al., 2006. The clinical spectrum of adult acute myeloid leukaemia associated with core binding factor translocations. *British Journal of Haematology,* 135 (2), 165-173.

Arandi, N., et al., 2018. Overexpression of indoleamine 2, 3-dioxygenase correlates with regulatory T cell phenotype in acute myeloid leukemia patients with normal karyotype. *Blood Research,* 53 (4), 294.

Aslanyan, M.G., et al., 2014. Clinical and biological impact of TET2 mutations and expression in younger adult AML patients treated within the EORTC/GIMEMA AML-12 clinical trial. *Annals of Hematology,* 93, 1401-1412.

Assi, R., et al., 2016. CPX-351 for the treatment of high-risk patients with acute myeloid leukemia. *Blood,* 128 (22), 4047.

Atalay, F. and Ateşoğlu, E.B., 2016. Low Dose Cytosine Arabinoside and Azacitidine Combination in Elderly Patients with Acute Myeloid Leukemia and Refractory Anemia with Excess Blasts (MDS-RAEB2). *Indian Journal of Hematology and Blood Transfusion,* 32, 46-53.

Austin, R.J., et al., 2023. Oncogenic drivers dictate immune control of acute myeloid leukemia. *Nature Communications,* 14 (1), 2155.

Awada, H., et al., 2021. Machine learning integrates genomic signatures for subclassification beyond primary and secondary acute myeloid leukemia. *Blood, the Journal of the American Society of Hematology,* 138 (19), 1885-1895.

Bachegowda, L., et al., 2016. Pexmetinib: A Novel Dual Inhibitor of Tie2 and p38 MAPK with Efficacy in Preclinical Models of Myelodysplastic Syndromes and Acute Myeloid LeukemiaTie2 Inhibition in MDS and AML. *Cancer Research,* 76 (16), 4841-4849.

Baker, K.E. and Coller, J., 2006. The many routes to regulating mRNA translation. *Genome Biology,* 7 (332).

Bali, P., et al., 2005. Inhibition of histone deacetylase 6 acetylates and disrupts the chaperone function of heat shock protein 90: a novel basis for antileukemia activity of histone deacetylase inhibitors. *Journal of Biological Chemistry,* 280 (29), 26729-26734.

Ball, B.J., et al., 2019. RAS mutations are independently associated with decreased overall survival and event-free survival in patients with AML receiving induction chemotherapy.

Baragaño Raneros, A., et al., 2015. Methylation of NKG2D ligands contributes to immune system evasion in acute myeloid leukemia. *Genes & Immunity,* 16 (1), 71-82.

Bartek, J. and Lukas, J., 2003. Chk1 and Chk2 kinases in checkpoint control and cancer. *Cancer Cell,* 3 (5), 421-429.

Beloribi-Djefaflia, S., Vasseur, S. and Guillaumond, F., 2016. Lipid metabolic reprogramming in cancer cells. *Oncogenesis,* 5 (1), e189.

Benchimol, S., 2001. No title. *P53-Dependent Pathways of Apoptosis,* .

Bensalah, K., Montorsi, F. and Shariat, S.F., 2007. Challenges of cancer biomarker profiling. *European Urology,* 52 (6), 1601-1609.

Bera, R., et al., 2020. DNMT3A mutants provide proliferating advantage with augmentation of self-renewal activity in the pathogenesis of AML in KMT2A-PTD-positive leukemic cells. *Oncogenesis,* 9 (2), 1-16.

Bernasconi, P. and Borsani, O., 2019. Immune escape after hematopoietic stem cell transplantation (HSCT): From mechanisms to novel therapies. *Cancers,* 12 (1), 69.

Berthon, C., et al., 2010. In acute myeloid leukemia, B7-H1 (PD-L1) protection of blasts from cytotoxic T cells is induced by TLR ligands and interferon-gamma and can be reversed using MEK inhibitors. *Cancer Immunology, Immunotherapy,* 59, 1839-1849.

Bewersdorf, J.P., et al., 2019. The minimal that kills: Why defining and targeting measurable residual disease is the "Sine Qua Non" for further progress in management of acute myeloid leukemia. *Blood Reviews,* , 100650.

Beyer, K., et al., 2019. Interactions of tumor necrosis factor–related apoptosis-inducing ligand (TRAIL) with the immune system: Implications for inflammation and cancer. *Cancers,* 11 (8), 1161.

Bhat, M.Y., et al., 2018. Comprehensive network map of interferon gamma signaling. *Journal of Cell Communication and Signaling,* 12 (4), 745-751.

Biaoxue, R., Xiguang, C. and Shuanying, Y., 2014. Annexin A1 in malignant tumors: current opinions and controversies. *The International Journal of Biological Markers,* 29 (1), e8-e20.

Bibault, J.E., et al., 2015. Next-generation sequencing of FLT3 internal tandem duplications for minimal residual disease monitoring in acute myeloid leukemia. *Oncotarget,* 6 (26), 22812-22821.

Blair, H.A., 2018. Daunorubicin/cytarabine liposome: a review in acute myeloid leukaemia. *Drugs,* 78 (18), 1903-1910.

Blamey, R.W., et al., 2007. Reading the prognosis of the individual with breast cancer. *European Journal of Cancer,* 43 (10), 1545-1547.

Boddu, P.C., et al., 2017. Characteristics and outcomes of older patients with secondary acute myeloid leukemia according to treatment approach. *Cancer,* 123 (16), 3050-3060.

Boddu, P., et al., 2017. Co-occurrence of FLT3-TKD and NPM1 mutations defines a highly favorable prognostic AML group. *Blood Advances,* 1 (19), 1546-1550.

Bogeska, R., et al., 2022. Inflammatory exposure drives long-lived impairment of hematopoietic stem cell self-renewal activity and accelerated aging. *Cell Stem Cell,* 29 (8), 1273-1284. e8.

Bories, P., et al., 2020. Impact of TP53 mutations in acute myeloid leukemia patients treated with azacitidine. *PLoS One,* 15 (10), e0238795.

Boyiadzis, M.M., et al., 2018. Chimeric antigen receptor (CAR) T therapies for the treatment of hematologic malignancies: clinical perspective and significance. *Journal for Immunotherapy of Cancer,* 6 (1), 137.

Brandt, C.S., et al., 2009. The B7 family member B7-H6 is a tumor cell ligand for the activating natural killer cell receptor NKp30 in humans. *Journal of Experimental Medicine,* 206 (7), 1495-1503.

Broad Institute, 2016. *A refined view of the biomarker pipeline* [online]. Broad Institute. Available at: https://www.broadinstitute.org/cptac/team-approach [Accessed 25/05 2021].

Brudno, J.N. and Kochenderfer, J.N., 2016. Toxicities of chimeric antigen receptor T cells: recognition and management. *Blood,* 127 (26), 3321-3330.

Bryan, J.C. and Jabbour, E.J., 2015. Management of relapsed/refractory acute myeloid leukemia in the elderly: current strategies and developments. *Drugs & Aging,* 32 (8), 623-637.

Burgess, M.R., et al., 2017. KRAS allelic imbalance enhances fitness and modulates MAP kinase dependence in cancer. *Cell,* 168 (5), 817-829. e15.

Cagiada, M., et al., 2021. Understanding the origins of loss of protein function by analyzing the effects of thousands of variants on activity and abundance. *Molecular Biology and Evolution,* 38 (8), 3235-3246.

Califf, R.M., 2018. Biomarker definitions and their applications. *Experimental Biology and Medicine,* 243 (3), 213-221.

Camp, R.L., Dolled-Filhart, M. and Rimm, D.L., 2004. X-tile: a new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization. *Clinical Cancer Research,* 10 (21), 7252-7259.

Cancer Genome Atlas Research Network, 2013. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *New England Journal of Medicine,* 368 (22), 2059-2074.

Cancer Research, U.K., 2017a. *Acute myeloid leukaemia (AML) death rate by year* [online]. Cancer Research UK. Available at: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/leukaemia-aml/mortality#heading-Two [Accessed 12/03 2020].

Cancer Research, U.K., 2017b. *Acute myeloid leukaemia (AML) incidence by age* [online]. Cancer Research UK. Available at: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/leukaemia-aml/incidence#heading-One [Accessed 10/03 2020].

Chakraborty, S., et al., 2018. Onco-multi-OMICS approach: a new frontier in cancer research. *BioMed Research International,* 2018.

Chan, S.M. and Majeti, R., 2013. Role of DNMT3A, TET2, and IDH1/2 mutations in pre-leukemic stem cells in acute myeloid leukemia. *International Journal of Hematology,* 98, 648-657.

Chen, D.S. and Mellman, I., 2017. Elements of cancer immunity and the cancer-immune set point. *Nature,* 541 (7637), 321-330.

Chen, D. and Mellman, I., 2017. Elements of cancer immunity and the cancer–immune set point. *Nature,* 541 (7637), 321-330.

Chen, D.S. and Mellman, I., 2013. Oncology meets immunology: the cancer-immunity cycle. *Immunity,* 39 (1), 1-10.

Chen, P., et al., 2016. Induction of multidrug resistance of acute myeloid leukemia cells by cocultured stromal cells via upregulation of the PI3K/Akt signaling pathway. *Oncology Research Featuring Preclinical and Clinical Cancer Therapeutics,* 24 (4), 215-223.

Chen, W., et al., 2014. A distinct glucose metabolism signature of acute myeloid leukemia with prognostic value. *Blood, the Journal of the American Society of Hematology,* 124 (10), 1645-1654.

Chen, X., et al., 2013. Calreticulin mediates immune recognition of acute myeloid leukemia cells in vivo. *Journal for ImmunoTherapy of Cancer,* 1, 1.

Chen, X., et al., 2015. Relation of clinical response and minimal residual disease and their prognostic impact on outcome in acute myeloid leukemia. *Journal of Clinical Oncology,* 33 (11), 1258-1264.

Chen, Y., Qiu, X. and Liu, R., 2024. Comprehensive characterization of immunogenic cell death in acute myeloid leukemia revealing the association with prognosis and tumor immune microenvironment. *BMC Medical Genomics,* 17 (1), 107.

Cheon, H., Yang, J. and Stark, G.R., 2011. The functions of signal transducers and activators of transcriptions 1 and 3 as cytokine-inducible proteins. *Journal of Interferon & Cytokine Research,* 31 (1), 33-40.

Christopher, M.J., et al., 2018. Immune escape of relapsed AML cells after allogeneic transplantation. *New England Journal of Medicine,* 379 (24), 2330-2341.

Cools, N., et al., 2008. Immunosuppression induced by immature dendritic cells is mediated by TGF-β/IL-10 double-positive CD4 regulatory T cells. *Journal of Cellular and Molecular Medicine,* 12 (2), 690-700.

Corm, S., et al., 2009. Indoleamine 2, 3-dioxygenase activity of acute myeloid leukemia cells can be measured from patients' sera by HPLC and is inducible by IFN-γ. *Leukemia Research,* 33 (3), 490-494.

Cornel, A.M., Mimpen, I.L. and Nierkens, S., 2020. MHC class I downregulation in cancer: underlying mechanisms and potential targets for cancer immunotherapy. *Cancers,* 12 (7), 1760.

Cornelissen, J.J., et al., 2015. Comparative therapeutic value of post-remission approaches in patients with acute myeloid leukemia aged 40–60 years. *Leukemia,* 29 (5), 1041.

Cornelissen, J.J., et al., 2007. Results of a HOVON/SAKK donor versus no-donor analysis of myeloablative HLA-identical sibling stem cell transplantation in first remission acute myeloid leukemia in young and middle-aged adults: benefits for whom? *Blood,* 109 (9), 3658-3666.

Corradi, G., et al., 2022. Release of IFNγ by acute myeloid leukemia cells remodels bone marrow immune microenvironment by inducing regulatory T cells. *Clinical Cancer Research,* 28 (14), 3141-3155.

Cox, D.R., 1972. Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological),* 34 (2), 187-202.

Curti, A., et al., 2007. Modulation of tryptophan catabolism by human leukemic cells results in the conversion of CD25– into CD25 T regulatory cells. *Blood,* 109 (7), 2871-2877.

De Angelis, F. and Breccia, M., 2015. Molecular monitoring as a path to cure acute promyelocytic leukemia. *Rare Cancers and Therapy,* 3 (1-2), 119-132.

De Hoffmann, E. and Stroobant, V., 2007. *Mass spectrometry: principles and applications.* John Wiley & Sons.

De Kouchkovsky, I. and Abdul-Hay, M., 2016. Acute myeloid leukemia: a comprehensive review and 2016 update. *Blood Cancer Journal,* 6 (7), e441.

de Latour, R.P., et al., 2012. Age-adjusted recipient pre transplant telomere length and treatment-related mortality after haematopoietic stem cell transplantation. *In: BONE MARROW TRANSPLANTATION,* NATURE PUBLISHING GROUP MACMILLAN BUILDING, 4 CRINAN ST, LONDON N1 9XW, ENGLAND, pp. S28.

DeAngelo, D.J., Stein, E.M. and Ravandi, F., 2016. Evolving therapies in acute myeloid leukemia: progress at last? *American Society of Clinical Oncology Educational Book,* 36, e302-e312.

Debarri, H., et al., 2015. IDH1/2 but not DNMT3A mutations are suitable targets for minimal residual disease monitoring in acute myeloid leukemia patients: a study by the Acute Leukemia French Association. *Oncotarget,* 6 (39), 42345-42353.

Dela Cruz Chuh, J., et al., 2021. Preclinical optimization of Ly6E-targeted ADCs for increased durability and efficacy of anti-tumor response. *In: Mabs,* Taylor & Francis, pp. 1862452.

Dhall, A., et al., 2019. Intersection of epigenetic and metabolic regulation of histone modifications in acute myeloid leukemia. *Frontiers in Oncology,* 9, 432.

Diesch, J., et al., 2016. A clinical-molecular update on azanucleoside-based therapy for the treatment of hematologic cancers. *Clinical Epigenetics,* 8, 1-11.

DiNardo, C.D. and Cortes, J.E., 2016. Mutations in AML: prognostic and therapeutic implications. *Hematology 2014, the American Society of Hematology Education Program Book,* 2016 (1), 348-355.

Dinardo, C.D., et al., 2017. Validation of a clinical assay of multi-locus DNA methylation for prognosis of newly diagnosed AML. *American Journal of Hematology,* 92 (2), E14-E15.

Djebali, S., et al., 2012. Landscape of transcription in human cells. *Nature,* 489 (7414), 101-108.

Do, J.H., et al., 2007. Treatment outcome of multidrug resistance related mRNA expression and c-jun-N-terminal kinase activity in patients with acute myeloid leukemia. *The Korean Journal of Laboratory Medicine,* 27 (4), 229-236.

Dohner, H., et al., 2017. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood,* 129 (4), 424-447.

Döhner, H., et al., 2017. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood, the Journal of the American Society of Hematology,* 129 (4), 424-447.

Döhner, H., et al., 2022. Diagnosis and management of AML in adults: 2022 recommendations from an international expert panel on behalf of the ELN. *Blood, the Journal of the American Society of Hematology,* 140 (12), 1345-1377.

Döhner, H., Weisdorf, D.J. and Bloomfield, C.D., 2015. Acute myeloid leukemia. *New England Journal of Medicine,* 373 (12), 1136-1152.

Doll, S., Gnad, F. and Mann, M., 2019. The Case for Proteomics and Phospho-Proteomics in Personalized Cancer Medicine. *PROTEOMICS–Clinical Applications,* 13 (2), 1800113.

Domon, B. and Aebersold, R., 2010. Options and considerations when selecting a quantitative proteomics strategy. *Nature Biotechnology,* 28 (7), 710-721.

Duarte, C.W., et al., 2012. Expression signature of IFN/STAT1 signaling genes predicts poor survival outcome in glioblastoma multiforme in a subtype-specific manner. *PloS One,* 7 (1), e29653.

Estey, E.H., 2018. Acute myeloid leukemia: 2019 update on risk-stratification and management. *American Journal of Hematology,* 93 (10), 1267-1291.

Fang, Z., Martin, J. and Wang, Z., 2012. Statistical methods for identifying differentially expressed genes in RNA-Seq experiments. *Cell & Bioscience,* 2 (1), 1-8.

Fathi, A.T., et al., 2012. A potential therapeutic target for FLT3-ITD AML: PIM1 kinase. *Leukemia Research,* 36 (2), 224-231.

Ferreira, D., Adega, F. and Chaves, R., 2013. The importance of cancer cell lines as in vitro models in cancer methylome analysis and anticancer drugs testing. *Oncogenomics and Cancer Proteomics-Novel Approaches in Biomarkers Discovery and Therapeutic Targets in Cancer,* 1, 139-166.

Fey, M.F. and Buske, C., 2013. Acute myeloblastic leukaemias in adult patients: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Annals of Oncology,* 24, vi138-vi143.

Figueroa, M.E., et al., 2010. Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. *Cancer Cell,* 18 (6), 553-567.

Figueroa, M.E., et al., 2010. DNA methylation signatures identify biologically distinct subtypes in acute myeloid leukemia. *Cancer Cell,* 17 (1), 13-27.

Finazzi, G., et al., 2013. A phase II study of G ivinostat in combination with hydroxycarbamide in patients with polycythaemia vera unresponsive to hydroxycarbamide monotherapy. *British Journal of Haematology,* 161 (5), 688-694.

Fiskus, W., et al., 2014. Highly effective combination of LSD1 (KDM1A) antagonist and pan-histone deacetylase inhibitor against human AML cells. *Leukemia,* 28 (11), 2155-2164.

Florez, M.A., et al., 2020. Interferon gamma mediates hematopoietic stem cell activation and niche relocalization through BST2. *Cell Reports,* 33 (12), 108530.

Folgiero, V., et al., 2014. Indoleamine 2,3-dioxygenase 1 (IDO1) activity in leukemia blasts correlates with poor outcome in childhood acute myeloid leukemia. *Oncotarget,* 5 (8), 2052-2064.

Fostvedt, L.K., et al., 2019. Pharmacokinetic/Pharmacodynamic Modeling To Support The Re-Approval Of Gemtuzumab Ozogamicin. *Clinical Pharmacology & Therapeutics,* .

Frosig, T.M., 2015. Elucidating the immunological effects of 5-azacytidine treatment in patients with myelodysplastic syndrome and identifying new conditional ligands and T-cell epitopes of relevance in melanoma. *Danish Medical Journal,* 62 (8), B5144.

Galea, M.H., et al., 1992. The Nottingham Prognostic Index in primary breast cancer. *Breast Cancer Research and Treatment,* 22, 207-219.

Gallegos, C.E., et al., 2016. Immunomodulation of classical and non-classical HLA molecules by ionizing radiation. *Cellular Immunology,* 303, 16-23.

Gambacorta, V., et al., 2019. Epigenetic therapies for acute myeloid leukemia and their immune-related effects. *Frontiers in Cell and Developmental Biology,* 7, 207.

Garcia-Diaz, A., et al., 2017. Interferon receptor signaling pathways regulating PD-L1 and PD-L2 expression. *Cell Reports,* 19 (6), 1189-1201.

Gerstung, M., et al., 2017. Precision oncology for acute myeloid leukemia using a knowledge bank approach. *Nature Genetics,* 49 (3), 332-340.

Gillet, L.C., et al., 2012. Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Molecular & Cellular Proteomics : MCP,* 11 (6), O111.016717.

Goel, M.K., Khanna, P. and Kishore, J., 2010. Understanding survival analysis: Kaplan-Meier estimate. *International Journal of Ayurveda Research,* 1 (4), 274-278.

Goldman, S.L., et al., 2019. Epigenetic Modifications in Acute Myeloid Leukemia: Prognosis, Treatment, and Heterogeneity. *Frontiers in Genetics,* 10.

Gonzales-van Horn, S.R. and Farrar, J.D., 2015. Interferon at the crossroads of allergy and viral infections. *Journal of Leukocyte Biology,* 98 (2), 185-194.

Gonzalez-Garza, M.T., et al., 2018. Comparing stemness gene expression between stem cell subpopulations from peripheral blood and adipose tissue. *American Journal of Stem Cells,* 7 (2), 38-47.

Gorvel, L. and Olive, D., 2020. Targeting the "PVR-TIGIT axis" with immune checkpoint therapies. *F1000Research,* 9, 10.12688/f1000research.22877.1. eCollection 2020.

Grandhi, A., Guo, W. and Peddada, S.D., 2016. A multiple testing procedure for multi-dimensional pairwise comparisons with application to gene expression studies. *BMC Bioinformatics,* 17 (1), 1-12.

Grizzle, W.E., Bell, W.C. and Sexton, K.C., 2011. Issues in collecting, processing and storing human tissues and associated information to support biomedical research. *Cancer Biomarkers,* 9 (1-6), 531-549.

Gros, F., et al., 2006. Soluble HLA-G Molecules Are Increased during Acute Leukemia, Especially in Subtypes Affecting Monocytic and Lymphoid Lineages'. *Neoplasia,* 8 (3), 223-230.

Groth, C., et al., 2019. Immunosuppression mediated by myeloid-derived suppressor cells (MDSCs) during tumour progression. *British Journal of Cancer,* 120 (1), 16-25.

Gu, Z., Dickerson, K.E. and Xu, J., 2020. Therapy Response and Outcome Explained by Leukemia Cell of Origin. *Cancer Discovery,* 10 (10), 1445-1447.

Gujral, P., et al., 2020. Histone acetylation and the role of histone deacetylases in normal cyclic endometrium. *Reproductive Biology and Endocrinology,* 18, 1-11.

Habanjar, O., et al., 2021. 3D cell culture systems: tumor application, advantages, and disadvantages. *International Journal of Molecular Sciences,* 22 (22), 12200.

Hackl, H., Astanina, K. and Wieser, R., 2017. Molecular and genetic alterations associated with therapy resistance and relapse of acute myeloid leukemia. *Journal of Hematology & Oncology,* 10 (1), 51.

Halenius, A., Gerke, C. and Hengel, H., 2015. Classical and non-classical MHC I molecule manipulation by human cytomegalovirus: so many targets—but how many arrows in the quiver? *Cellular & Molecular Immunology,* 12 (2), 139-153.

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

Handschuh, L., et al., 2018. Gene expression profiling of acute myeloid leukemia samples from adult patients with AML-M1 and-M2 through boutique microarrays, real-time PCR and droplet digital PCR. *International Journal of Oncology,* 52 (3), 656-678.

Hatem, E., El Banna, N. and Huang, M., 2017. Multifaceted roles of glutathione and glutathione-based systems in carcinogenesis and anticancer drug resistance. *Antioxidants & Redox Signaling,* 27 (15), 1217-1234.

Hattori, N., et al., 2019. Monitoring TIGIT/DNAM-1 and PVR/PVRL2 immune checkpoint expression levels in allogeneic stem cell transplantation for acute myeloid leukemia. *Biology of Blood and Marrow Transplantation,* 25 (5), 861-867.

Hausmann, M., et al., 2011. BCL-2 modifying factor (BMF) is a central regulator of anoikis in human intestinal epithelial cells. *Journal of Biological Chemistry,* 286 (30), 26533-26540.

Heuser, M., et al., 2005. Gene-expression profiles and their association with drug resistance in adult acute myeloid leukemia. *Haematologica,* 90 (11), 1484-1492.

Hoesel, B. and Schmid, J.A., 2013. The complexity of NF-κB signaling in inflammation and cancer. *Molecular Cancer,* 12 (1), 1-15.

Hofmann, S., et al., 2019. Chimeric antigen receptor (CAR) T cell therapy in acute myeloid leukemia (AML). *Journal of Clinical Medicine,* 8 (2), 200.

Holm, S., 1979. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics,* , 65-70.

Horibata, S., et al., 2019. Heterogeneity in refractory acute myeloid leukemia. *Proceedings of the National Academy of Sciences of the United States of America,* 116 (21), 10494-10503.

Houghton, A.N. and Guevara-Patiño, J.A., 2004. Immune recognition of self in immunity against cancer. *The Journal of Clinical Investigation,* 114 (4), 468-471.

Howlader N, Noone AM, Krapcho M, Miller D, Brest A, Yu M, Ruhl J, Tatalovich Z, Mariotto A, Lewis DR, Chen HS, Feuer EJ, Cronin KA (eds)., 2018. *SEER Cancer Statistics Review.* Bethesda, MD: National Cancer Institute.

Hu, A., Noble, W.S. and Wolf-Yadlin, A., 2016. Technical advances in proteomics: new developments in data-independent acquisition. *F1000Research,* 5, 10.12688/f1000research.7042.1. eCollection 2016.

Huang, X., et al., 2022. The role of indoleamine 2, 3-dioxygenase 1 in regulating tumor microenvironment. *Cancers,* 14 (11), 2756.

Huls, G., 2015. Azacitidine in AML: a treatment option? *Blood, the Journal of the American Society of Hematology,* 126 (3), 283-284.

Hussong, J.W., Rodgers, G.M. and Shami, P.J., 2000. Evidence of increased angiogenesis in patients with acute myeloid leukemia. *Blood, the Journal of the American Society of Hematology,* 95 (1), 309-313.

Hyrien, O., 2016. How MCM loading and spreading specify eukaryotic DNA replication initiation sites. *F1000Research,* 5, 10.12688/f1000research.9008.1. eCollection 2016.

Imrichova, D., et al., 2015. Selection of resistant acute myeloid leukemia SKM-1 and MOLM-13 cells by vincristine-, mitoxantrone-and lenalidomide-induced upregulation of P-glycoprotein activity and downregulation of CD33 cell surface exposure. *European Journal of Pharmaceutical Sciences,* 77, 29-39.

Insinga, A., et al., 2005. Inhibitors of histone deacetylases induce tumor-selective apoptosis through activation of the death receptor pathway. *Nature Medicine,* 11 (1), 71-76.

International Non-Hodgkin's Lymphoma Prognostic Factors Project, 1993. A predictive model for aggressive non-Hodgkin's lymphoma. *New England Journal of Medicine,* 329 (14), 987-994.

Itzykson, R., Kosmider, O. and Fenaux, P., 2013. Somatic mutations and epigenetic abnormalities in myelodysplastic syndromes. *Best Practice & Research Clinical Haematology,* 26 (4), 355-364.

Jacamo, R., et al., 2017. Up-Regulation of iNOS in AML Blasts Creates an Immunosuppressive Microenvironment, Inhibits T-Cell Proliferation and Transforms T-Cells Towards a Tumor-Tolerating Phenotype. *Blood,* 130, 2443.

Jan, M., et al., 2019. Recurrent genetic HLA loss in AML relapsed after matched unrelated allogeneic hematopoietic cell transplantation. *Blood Advances,* 3 (14), 2199-2204.

Janus, P., et al., 2020. Pro-death signaling of cytoprotective heat shock factor 1: upregulation of NOXA leading to apoptosis in heat-sensitive cells. *Cell Death & Differentiation,* 27 (7), 2280-2292.

Ji, Q. and Qiu, L., 2016. Mechanism study of PEGylated polyester and β-cyclodextrin integrated micelles on drug resistance reversal in MRP1-overexpressed HL60/ADR cells. *Colloids and Surfaces B: Biointerfaces,* 144, 203-213.

Johnson, B., et al., 2018. Human IFIT3 Modulates IFIT1 RNA Binding Specificity and Protein Stability. *Immunity,* 48 (3), 487-499.e5.

Jolliffe, I.T. and Cadima, J., 2016. Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences,* 374 (2065), 20150202.

Jongen-Lavrencic, M., et al., 2018. Molecular minimal residual disease in acute myeloid leukemia. *New England Journal of Medicine,* 378 (13), 1189-1199.

Justesen, J., et al., 2001. Adipocyte tissue volume in bone marrow is increased with aging and in patients with osteoporosis. *Biogerontology,* 2, 165-171.

Kagoya, Y., et al., 2014. Positive feedback between NF-κB and TNF-α promotes leukemia-initiating cell capacity. *The Journal of Clinical Investigation,* 124 (2), 528-542.

Kaplan, E.L. and Meier, P., 1958. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association,* 53 (282), 457-481.

Kathad, U., et al., 2021. A machine learning-based gene signature of response to the novel alkylating agent LP-184 distinguishes its potential tumor indications. *BMC Bioinformatics,* 22 (1), 1-23.

Kennedy, V.E. and Smith, C.C., 2020. FLT3 Mutations in Acute Myeloid Leukemia: Key Concepts and Emerging Controversies. *Frontiers in Oncology,* 10, 2927.

Kerin, E.P., et al., 2022. Comparison of the Nottingham Prognostic Index and OncotypeDX© recurrence score in predicting outcome in estrogen receptor positive breast cancer. *The Breast,* 66, 227-235.

Kern, W., et al., 2010. The role of multiparameter flow cytometry for disease monitoring in AML. *Best Practice & Research Clinical Haematology,* 23 (3), 379-390.

Khodarev, N.N., et al., 2004. STAT1 is overexpressed in tumors selected for radioresistance and confers protection from radiation in transduced sensitive cells. *Proceedings of the National Academy of Sciences of the United States of America,* 101 (6), 1714-1719.

Khodarev, N.N., et al., 2009. STAT1 pathway mediates amplification of metastatic potential and resistance to therapy. *PloS One,* 4 (6), e5821.

Kim, E.S., 2017. Enasidenib: first global approval. *Drugs,* 77 (15), 1705-1711.

Kim, H.Y., 2019. Statistical notes for clinical researchers: simple linear regression 3 - residual analysis. *Restorative Dentistry & Endodontics,* 44 (1), e11.

Kim, T.K., 2015. T test as a parametric statistic. *Korean Journal of Anesthesiology,* 68 (6), 540-546.

Kirschbaum, M., et al., 2014. A phase 1 clinical trial of vorinostat in combination with decitabine in patients with acute myeloid leukaemia or myelodysplastic syndrome. *British Journal of Haematology,* 167 (2), 185-193.

Knaus, H.A., et al., 2019. A Signature of T Cell Exhaustion Is Enriched in the Bone Marrow (BM) of AML Patients and Shared with Immune Exhaustion Signatures of Solid Tumors. *Blood,* 134, 2711.

Kobayashi, T., et al., 2010. Galectin-9 exhibits anti-myeloma activity through JNK and p38 MAP kinase pathways. *Leukemia,* 24 (4), 843-850.

Kochan, G., et al., 2013. Role of non-classical MHC class I molecules in cancer immunosuppression. *Oncoimmunology,* 2 (11), e26491.

Kolb, H., 2017. Hematopoietic stem cell transplantation and cellular therapy. *Hla,* 89 (5), 267-277.

Kong, Y., et al., 2016. T-cell immunoglobulin and ITIM domain (TIGIT) associates with CD8 T-cell exhaustion and poor clinical outcome in AML patients. *Clinical Cancer Research,* 22 (12), 3057-3066.

Koreth, J., et al., 2009. Allogeneic stem cell transplantation for acute myeloid leukemia in first complete remission: systematic review and meta-analysis of prospective clinical trials. *Jama,* 301 (22), 2349-2361.

Kornblau, S.M., et al., 2010. Recurrent expression signatures of cytokines and chemokines are present and are independently prognostic in acute myelogenous leukemia and myelodysplasia. *Blood,* 116 (20), 4251-4261.

Korthauer, K., et al., 2019. A practical guide to methods controlling false discoveries in computational biology. *Genome Biology,* 20 (1), 1-21.

Kotredes, K.P. and Gamero, A.M., 2013. Interferons as inducers of apoptosis in malignant cells. *Journal of Interferon & Cytokine Research,* 33 (4), 162-170.

Kren, L., et al., 2010. Production of immune-modulatory nonclassical molecules HLA-G and HLA-E by tumor infiltrating ameboid microglia/macrophages in glioblastomas: a role in innate immunity? *Journal of Neuroimmunology,* 220 (1-2), 131-135.

Kroeze, L.I., et al., 2014. Characterization of acute myeloid leukemia based on levels of global hydroxymethylation. *Blood, the Journal of the American Society of Hematology,* 124 (7), 1110-1118.

Kulsoom, B., Shamsi, T.S. and Afsar, N.A., 2019. Lung resistance-related protein (LRP) predicts favorable therapeutic outcome in Acute Myeloid Leukemia. *Scientific Reports,* 9 (1), 1-11.

Kuroda, J., et al., 2010. Targeting activating transcription factor 3 by Galectin-9 induces apoptosis and overcomes various types of treatment resistance in chronic myelogenous leukemia. *Molecular Cancer Research : MCR,* 8 (7), 994-1001.

Kuroda, J., et al., 2010. Targeting activating transcription factor 3 by Galectin-9 induces apoptosis and overcomes various types of treatment resistance in chronic myelogenous leukemia. *Molecular Cancer Research,* 8 (7), 994-1001.

Lai, C., Doucette, K. and Norsworthy, K., 2019. Recent drug approvals for acute myeloid leukemia. *Journal of Hematology & Oncology,* 12 (1), 100.

Lai, Y., et al., 2021. A novel 85-gene expression signature predicts unfavorable prognosis in acute myeloid leukemia. *Technology in Cancer Research & Treatment,* 20, 15330338211004933.

Laloi, L., et al., 2022. Retrospective, real-life study of venetoclax plus azacitidine or low-dose cytarabine in French patients with acute myeloid leukemia ineligible for intensive chemotherapy. *Cancer Medicine, .*

Lamba, J.K., et al., 2018. Integrated epigenetic and genetic analysis identifies markers of prognostic significance in pediatric acute myeloid leukemia. *Oncotarget,* 9 (42), 26711-26723.

Lambert, J., et al., 2013. Mapping differential interactomes by affinity purification coupled with data-independent mass spectrometry acquisition. *Nature Methods,* 10 (12), 1239.

Lamble, A.J. and Lind, E.F., 2018. Targeting the immune Microenvironment in Acute Myeloid Leukemia: A Focus on T Cell immunity. *Frontiers in Oncology,* 8.

Le Dieu, R., et al., 2009. Peripheral blood T cells in acute myeloid leukemia (AML) patients at diagnosis have abnormal phenotype and genotype and form defective immune synapses with AML blasts. *Blood, the Journal of the American Society of Hematology,* 114 (18), 3909-3916.

Lee, J.M., et al., 2023. Control of protein stability by post-translational modifications. *Nature Communications,* 14 (1), 201.

Leonhardt, H., et al., 1992. A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell,* 71 (5), 865-873.

Levis, M., 2017. Midostaurin approved for FLT3-mutated AML. *Blood,* 129 (26), 3403-3406.

Li, S., Mason, C.E. and Melnick, A., 2016a. Genetic and epigenetic heterogeneity in acute myeloid leukemia. *Current Opinion in Genetics & Development,* 36, 100-106.

Li, S., Mason, C.E. and Melnick, A., 2016b. Genetic and epigenetic heterogeneity in acute myeloid leukemia. *Current Opinion in Genetics & Development,* 36, 100-106.

Li, S., et al., 2017. Overcoming resistance to cisplatin by inhibition of glutathione S-transferases (GSTs) with ethacraplatin micelles in vitro and in vivo. *Biomaterials,* 144, 119-129.

Li, Y., et al., 2016. Givinostat, a type II histone deacetylase inhibitor, induces potent caspase-dependent apoptosis in human lymphoblastic leukemia. *Genes & Cancer,* 7 (9-10), 292-300.

Lin, A. and Yan, W., 2018. Heterogeneity of HLA-G expression in cancers: Facing the challenges. *Frontiers in Immunology,* 9, 2164.

Lin, T.L., et al., 2019. The Impact of Hematopoietic Cell Transplantation (HCT) on Survival: An Exploratory Analysis of a Phase 3 Study of CPX-351 Versus 7 3 in Older Patients with Newly Diagnosed, High-Risk/Secondary AML (sAML). *Biology of Blood and Marrow Transplantation,* 25 (3), S125-S126.

Link, K.A., et al., 2016. Supraphysiologic levels of the AML1-ETO isoform AE9a are essential for transformation. *Proceedings of the National Academy of Sciences of the United States of America,* 113 (32), 9075-9080.

Liu, K., et al., 2003. Endogenous assays of DNA methyltransferases: Evidence for differential activities of DNMT1, DNMT2, and DNMT3 in mammalian cells in vivo. *Molecular and Cellular Biology,* 23 (8), 2709-2719.

Liu, S., et al., 2007. The function of heme-regulated eIF2alpha kinase in murine iron homeostasis and macrophage maturation. *The Journal of Clinical Investigation,* 117 (11), 3296-3305.

Liu, X., et al., 2022. A comparison of transcriptome analysis methods with reference genome. *BMC Genomics,* 23 (1), 1-15.

Liu, Y., et al., 2013. Mass spectrometric protein maps for biomarker discovery and clinical research. *Expert Review of Molecular Diagnostics,* 13 (8), 811-825.

Locafaro, G., et al., 2014. HLA-G expression on blasts and tolerogenic cells in patients affected by acute myeloid leukemia. *Journal of Immunology Research,* 2014, 636292.

Loeffler, D. and Schroeder, T., 2021. Symmetric and asymmetric activation of hematopoietic stem cells. *Current Opinion in Hematology,* 28 (4), 262-268.

Lonetti, A., Pession, A. and Masetti, R., 2019. Targeted therapies for pediatric AML: gaps and perspective. *Frontiers in Pediatrics,* 7.

Lowe, R., et al., 2017. Transcriptomics technologies. *PLoS Computational Biology,* 13 (5), e1005457.

Ludwig, C., et al., 2018. Data-independent acquisition-based SWATH-MS for quantitative proteomics: a tutorial. *Molecular Systems Biology,* 14 (8), e8126.

Luger, S.M., 2020. *Randomized Trial of Gilteritinib vs Midostaurin in FLT3 Mutated Acute Myeloid Leukemia* [online]. Clinicaltrials.gov. Available at: https://clinicaltrials.gov/ct2/show/record/NCT03836209 https://ClinicalTrials.gov/show/NCT03588078; [Accessed 04/20 2020].

Luskin, M.R., et al., 2016. A clinical measure of DNA methylation predicts outcome in de novo acute myeloid leukemia. *JCI Insight,* 1 (9).

Maiques-Diaz, A., et al., 2018. LSD1 inhibitors disrupt the GFI1 transcription repressor complex. *Molecular & Cellular Oncology,* 5 (4), e1481813.

Malani, D., et al., 2022. Implementing a functional precision medicine tumor board for acute myeloid leukemia. *Cancer Discovery,* 12 (2), 388-401.

Mantovani, A., et al., 2002. Macrophage polarization: tumor-associated macrophages as a paradigm for polarized M2 mononuclear phagocytes. *Trends in Immunology,* 23 (11), 549-555.

Mardiros, A., Forman, S.J. and Budde, L.E., 2015. T cells expressing CD123 chimeric antigen receptors for treatment of acute myeloid leukemia. *Current Opinion in Hematology,* 22 (6), 484-488.

Margueron, R., Trojer, P. and Reinberg, D., 2005. The key to development: interpreting the histone code? *Current Opinion in Genetics & Development,* 15 (2), 163-176.

Mariño, G., et al., 2014. Self-consumption: the interplay of autophagy and apoptosis. *Nature Reviews Molecular Cell Biology,* 15 (2), 81-94.

Matatall, K.A., et al., 2018. A Novel Role for Bst2 and E-Selectin in IFNg-Stimulated HSC Niche Relocalization. *Blood,* 132, 874.

Maurer, M.J., 2023. The International Prognostic Index in aggressive B-cell lymphoma. *Haematologica,* 108 (11), 2874.

Mayle, A., et al., 2015. Dnmt3a loss predisposes murine hematopoietic stem cells to malignant transformation. *Blood, the Journal of the American Society of Hematology,* 125 (4), 629-638.

Megías-Vericat, J.E., et al., 2015. Influence of ABCB1 polymorphisms upon the effectiveness of standard treatment for acute myeloid leukemia: a systematic review and meta-analysis of observational studies. *The Pharmacogenomics Journal,* 15 (2), 109-118.

Metzeler, K.H., et al., 2008. An 86-probe-set gene-expression signature predicts survival in cytogenetically normal acute myeloid leukemia. *Blood, the Journal of the American Society of Hematology,* 112 (10), 4193-4201.

Michelson, A.P., et al., 2020. Association of immunophenotype with expression of topoisomerase ii α and β in adult acute myeloid leukemia. *Scientific Reports,* 10 (1), 1-9.

Minn, A.J., 2015. Interferons and the immunogenic effects of cancer therapy. *Trends in Immunology,* 36 (11), 725-737.

Mirabelli, P., Coppola, L. and Salvatore, M., 2019. Cancer cell lines are useful model systems for medical research. *Cancers,* 11 (8), 1098.

Mirandola, P., et al., 2004. Activated human NK and CD8 T cells express both TNF-related apoptosis-inducing ligand (TRAIL) and TRAIL receptors but are resistant to TRAIL-mediated cytotoxicity. *Blood,* 104 (8), 2418-2424.

Miserocchi, G., et al., 2017. Management and potentialities of primary cancer cultures in preclinical and translational studies. *Journal of Translational Medicine,* 15, 1-16.

Mizuno, S., et al., 2000. Aberrant expression of HLA-G antigen in interferon γ-stimulated acute myelogenous leukaemia. *British Journal of Haematology,* 111 (1), 280-282.

Mojic, M., Takeda, K. and Hayakawa, Y., 2017. The dark side of IFN-γ: its role in promoting cancer immunoevasion. *International Journal of Molecular Sciences,* 19 (1), 89.

Montazerinezhad, S., Emamjomeh, A. and Hajieghrari, B., 2020. Chromosomal abnormality, laboratory techniques, tools and databases in molecular Cytogenetics. *Molecular Biology Reports,* , 1-19.

Moran-Crusio, K., et al., 2011. Tet2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. *Cancer Cell,* 20 (1), 11-24.

Mortazavi, A., et al., 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods,* 5 (7), 621-628.

Moschoi, R., et al., 2016. Protective mitochondrial transfer from bone marrow stromal cells to acute myeloid leukemic cells during chemotherapy. *Blood, the Journal of the American Society of Hematology,* 128 (2), 253-264.

Moshofsky, K.B., et al., 2019. Acute myeloid leukemia–induced T-cell suppression can be reversed by inhibition of the MAPK pathway. *Blood Advances,* 3 (20), 3038-3051.

Mou, T., et al., 2020. Reproducibility of methods to detect differentially expressed genes from single-cell RNA sequencing. *Frontiers in Genetics,* 10, 1331.

Mussai, F., et al., 2013. Acute myeloid leukemia creates an arginase-dependent immunosuppressive microenvironment. *Blood, the Journal of the American Society of Hematology,* 122 (5), 749-758.

National Center for Biotechnology Information., 2020a. *Cytarabine, CID=6253* [online]. PubChem Database. Available at: https://pubchem.ncbi.nlm.nih.gov/compound/Cytarabine [Accessed 02/07 2020].

National Center for Biotechnology Information., 2020b. *Daunorubicin, CID=30323* [online]. PubChem Database. Available at: https://pubchem.ncbi.nlm.nih.gov/compound/Daunorubicin [Accessed 02/07 2020].

National Comprehensive Cancer Network, 2013. NCCN clinical practice guidelines in oncology (NCCN guidelines): Acute myeloid leukemia. *Fort Washington, PA: National Comprehensive Cancer Network, .*

National Health Service, 2019. *Symptoms Acute myeloid leukaemia* [online]. National Health Service. Available at: https://www.nhs.uk/conditions/acute-myeloid-leukaemia/symptoms/ [Accessed 14/04 2020].

Nebbioso, A., et al., 2005. Tumor-selective action of HDAC inhibitors involves TRAIL induction in acute myeloid leukemia cells. *Nature Medicine,* 11 (1), 77-84.

Neumann, F., et al., 2009. Levels of beta 2 microglobulin have a prognostic relevance for patients with myelodysplastic syndrome with regard to survival and the risk of transformation into acute myelogenous leukemia. *Leukemia Research,* 33 (2), 232-236.

Ng, S.W., et al., 2016. A 17-gene stemness score for rapid determination of risk in acute leukaemia. *Nature,* 540 (7633), 433-437.

Nguyen, S., et al., 2009. HLA-E upregulation on IFN-γ-activated AML blasts impairs CD94/NKG2A-dependent NK cytolysis after haplo-mismatched hematopoietic SCT. *Bone Marrow Transplantation,* 43 (9), 693-699.

Niewerth, D., et al., 2014. Interferon-γ-induced upregulation of immunoproteasome subunit assembly overcomes bortezomib resistance in human hematological cell lines. *Journal of Hematology & Oncology,* 7, 1-15.

Niu, J., Peng, D. and Liu, L., 2022. Drug resistance mechanisms of acute myeloid leukemia stem cells. *Frontiers in Oncology,* 12, 896426.

Ohsugi, T., et al., 2017. Decreased expression of interferon-induced protein 2 (IFIT2) by Wnt/beta-catenin signaling confers anti-apoptotic properties to colorectal cancer cells. *Oncotarget,* 8 (59), 100176-100186.

Orta, M.L., et al., 2014. The PARP inhibitor Olaparib disrupts base excision repair of 5-aza-2′-deoxycytidine lesions. *Nucleic Acids Research,* 42 (14), 9108-9120.

Ozato, K., et al., 2008. TRIM family proteins and their emerging roles in innate immunity. *Nature Reviews Immunology,* 8 (11), 849-860.

Padariya, M., et al., 2021. Functional interfaces, biological pathways, and regulations of interferon-related DNA damage resistance signature (IRDS) genes. *Biomolecules,* 11 (5), 622.

Paprocka, M., et al., 2017. MRP1 protein expression in leukemic stem cells as a negative prognostic marker in acute myeloid leukemia patients. *European Journal of Haematology,* 99 (5), 415-422.

Paschka, P., et al., 2003. Molecular monitoring of response to imatinib (Glivec®) in CML patients pretreated with interferon alpha. Low levels of residual disease are associated with continuous remission. *Leukemia,* 17 (9), 1687.

Passweg, J.R., et al., 2015. Conditioning intensity in middle-aged patients with AML in first CR: no advantage for myeloablative regimens irrespective of the risk group–an observational analysis by the Acute Leukemia Working Party of the EBMT. *Bone Marrow Transplantation,* 50 (8), 1063.

Pegoraro, A., et al., 2020. Differential sensitivity of acute myeloid leukemia cells to daunorubicin depends on P2X7A versus P2X7B receptor expression. *Cell Death & Disease,* 11 (10), 1-12.

Peto, R. and Peto, J., 1972. Asymptotically efficient rank invariant test procedures. *Journal of the Royal Statistical Society: Series A (General),* 135 (2), 185-198.

Pfeffer, C.M. and Singh, A.T., 2018. Apoptosis: a target for anticancer therapy. *International Journal of Molecular Sciences,* 19 (2), 448.

Platten, M., et al., 2015. Cancer immunotherapy by targeting IDO1/TDO and their downstream effectors. *Frontiers in Immunology,* 5, 673.

Pospori, C., et al., 2011. Specificity for the tumor-associated self-antigen WT1 drives the development of fully functional memory T cells in the absence of vaccination. *Blood, the Journal of the American Society of Hematology,* 117 (25), 6813-6824.

Prada-Arismendy, J., Arroyave, J.C. and Röthlisberger, S., 2017. Molecular biomarkers in acute myeloid leukemia. *Blood Reviews,* 31 (1), 63-76.

Pyzer, A.R., et al., 2017. MUC1-mediated induction of myeloid-derived suppressor cells in patients with acute myeloid leukemia. *Blood, the Journal of the American Society of Hematology,* 129 (13), 1791-1801.

Qian, Y., et al., 2021. Prognostic cancer gene expression signatures: current status and challenges. *Cells,* 10 (3), 648.

Qiu, J., et al., 2023. Cancer cells resistant to immune checkpoint blockade acquire interferon-associated epigenetic memory to sustain T cell dysfunction. *Nature Cancer,* 4 (1), 43-61.

Quek, L., et al., 2016. Mutational analysis of disease relapse in patients allografted for acute myeloid leukemia. *Blood Advances,* 1 (3), 193-204.

Radwan, S., et al., 2020. AML-273: The Immune Checkpoints CTLA-4 and LAG-3 Expression is Up-Regulated in Acute Myeloid Leukemia. *Clinical Lymphoma Myeloma and Leukemia,* 20, S198.

Ragaini, S., et al., 2022. An IDO1-related immune gene signature predicts overall survival in acute myeloid leukemia. *Blood Advances,* 6 (1), 87-99.

Ramadoss, S., Guo, G. and Wang, C., 2017. Lysine demethylase KDM3A regulates breast cancer cell invasion and apoptosis by targeting histone and the non-histone protein p53. *Oncogene,* 36 (1), 47-59.

Raponi, M., et al., 2008. A 2-gene classifier for predicting response to the farnesyltransferase inhibitor tipifarnib in acute myeloid leukemia. *Blood,* 111 (5), 2589-2596.

Rasmussen, K.D., et al., 2015. Loss of TET2 in hematopoietic cells leads to DNA hypermethylation of active enhancers and induction of leukemogenesis. *Genes & Development,* 29 (9), 910-922.

Rautenberg, C., et al., 2019. Relapse of acute myeloid leukemia after allogeneic stem cell transplantation: prevention, detection, and treatment. *International Journal of Molecular Sciences,* 20 (1), 228.

Ravandi, F., Walter, R.B. and Freeman, S.D., 2018. Evaluating measurable residual disease in acute myeloid leukemia. *Blood Advances,* 2 (11), 1356-1366.

Rechkoblit, O., et al., 2018. Structural basis for polymerase η–promoted resistance to the anticancer nucleoside analog cytarabine. *Scientific Reports,* 8 (1), 1-9.

Ribas, A., 2015. Adaptive Immune Resistance: How Cancer Protects from Immune Attack. *Cancer Discovery,* 5 (9), 915-919.

Richter, M., et al., 2021. From donor to the lab: A fascinating journey of primary cell lines. *Frontiers in Cell and Developmental Biology,* 9, 711381.

Ringnér, M., 2008. What is principal component analysis? *Nature Biotechnology,* 26 (3), 303-304.

Robinson, C.M., et al., 2010. Cytokines involved in interferon-gamma production by human macrophages. *Journal of Innate Immunity,* 2 (1), 56-65.

Robinson, C.M., Hale, P.T. and Carlin, J.M., 2005. The role of IFN-γ and TNF-α-responsive regulatory elements in the synergistic induction of indoleamine dioxygenase. *Journal of Interferon & Cytokine Research,* 25 (1), 20-30.

Robinson, K. and Tiriveedhi, V., 2020. Perplexing Role of P-Glycoprotein in Tumor Microenvironment. *Frontiers in Oncology,* 10, 265.

Roszak, J., Smok-Pieniążek, A. and Stępnik, M., 2017. Transcriptomic analysis of the PI3K/Akt signaling pathway reveals the dual role of the c-Jun oncogene in cytotoxicity and the development of resistance in HL-60 leukemia cells in response to arsenic trioxide.

Rouette, A., et al., 2016. Expression of immunoproteasome genes is regulated by cell-intrinsic and–extrinsic factors in human cancers. *Scientific Reports,* 6 (1), 34019.

Rücker, F.G., et al., 2016. Molecular dissection of valproic acid effects in acute myeloid leukemia identifies predictive networks. *Epigenetics,* 11 (7), 517-525.

Russell, N.H., et al., 2015. A comparative assessment of the curative potential of reduced intensity allografts in acute myeloid leukaemia. *Leukemia,* 29 (7), 1478.

Russler-Germain, D.A., et al., 2014. The R882H DNMT3A mutation associated with AML dominantly inhibits wild-type DNMT3A by blocking its ability to form active tetramers. *Cancer Cell,* 25 (4), 442-454.

Rutella, S., et al., 2022. Immune dysfunction signatures predict outcomes and define checkpoint blockade–unresponsive microenvironments in acute myeloid leukemia. *The Journal of Clinical Investigation,* 132 (21).

Rybka, A.E., Stephanou, A. and Townsend, P.A., 2009. Janus Kinase (JAK)-Signal Transducer and Activator of Transcription (STAT) Pathway in Heart Disease. *JAK-STAT Pathway in Disease,* , 76.

Saleem, T. and Kasi, A., 2021, Daunorubicin. *In:* Daunorubicin. *StatPearls.* Treasure Island (FL): StatPearls Publishing LLC, 2021, .

Sallman, D.A., et al., 2020. TP53 mutations in myelodysplastic syndromes and secondary AML confer an immunosuppressive phenotype. *Blood, the Journal of the American Society of Hematology,* 136 (24), 2812-2823.

Schaefer, M., et al., 2009. Azacytidine inhibits RNA methylation at DNMT2 target sites in human cancer cell lines. *Cancer Research,* 69 (20), 8127-8132.

Schenk, T., et al., 2012. Inhibition of the LSD1 (KDM1A) demethylase reactivates the all-trans-retinoic acid differentiation pathway in acute myeloid leukemia. *Nature Medicine,* 18 (4), 605-611.

Schnorfeil, F.M., et al., 2015. T cells are functionally not impaired in AML: increased PD-1 expression is only seen at time of relapse and correlates with a shift towards the memory T cell compartment. *Journal of Hematology & Oncology,* 8 (1), 93.

Schober, P. and Vetter, T.R., 2018. Survival Analysis and Interpretation of Time-to-Event Data: The Tortoise and the Hare. *Anesthesia and Analgesia,* 127 (3), 792-798.

Schroder, K., et al., 2004. Interferon-γ: an overview of signals, mechanisms and functions. *Journal of Leukocyte Biology,* 75 (2), 163-189.

Schuurhuis, G.J., et al., 2018. Minimal/measurable residual disease in AML: a consensus document from the European LeukemiaNet MRD Working Party. *Blood, the Journal of the American Society of Hematology,* 131 (12), 1275-1291.

Schvartzman, J.M., et al., 2019. 2-hydroxyglutarate inhibits MyoD-mediated differentiation by preventing H3K9 demethylation. *Proceedings of the National Academy of Sciences of the United States of America,* 116 (26), 12851-12856.

Schwanhäusser, B., et al., 2011. Global quantification of mammalian gene expression control. *Nature,* 473 (7347), 337-342.

Seo, W., et al., 2022. Targeting ERRα promotes cytotoxic effects against acute myeloid leukemia through suppressing mitochondrial oxidative phosphorylation. *Journal of Hematology & Oncology,* 15 (1), 1-6.

Sha, K., et al., 2021. Identifying a novel 5-gene signature predicting clinical outcomes in acute myeloid leukemia. *Clinical and Translational Oncology,* 23, 648-656.

Shao, G., et al., 2019. Advancements of Annexin A1 in inflammation and tumorigenesis. *OncoTargets and Therapy,* 12, 3245-3254.

Sheikh, M.H. and Solito, E., 2018. Annexin A1: uncovering the many talents of an old protein. *International Journal of Molecular Sciences,* 19 (4), 1045.

Shenghui, Z., et al., 2011. Elevated frequencies of CD4 CD25 CD127lo regulatory T cells is associated to poor prognosis in patients with acute myeloid leukemia. *International Journal of Cancer,* 129 (6), 1373-1381.

Shih, A.H., et al., 2015. Mutational cooperativity linked to combinatorial epigenetic gain of function in acute myeloid leukemia. *Cancer Cell,* 27 (4), 502-515.

Shih, T.T., et al., 2009. Bone marrow angiogenesis magnetic resonance imaging in patients with acute myeloid leukemia: peak enhancement ratio is an independent predictor for overall survival. *Blood, the Journal of the American Society of Hematology,* 113 (14), 3161-3167.

Shlush, L.I., et al., 2014. Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia. *Nature,* 506 (7488), 328-333.

Sidky, Y.A. and Borden, E.C., 1987. Inhibition of angiogenesis by interferons: effects on tumor- and lymphocyte-induced vascular responses. *Cancer Research,* 47 (19), 5155-5161.

Sidoli, S., et al., 2015. Sequential Window Acquisition of all Theoretical Mass Spectra (SWATH) Analysis for Characterization and Quantification of Histone Post-translational Modifications. *Molecular & Cellular Proteomics : MCP,* 14 (9), 2420-2428.

Silva, I.G., et al., 2017. The Tim-3-galectin-9 secretory pathway is involved in the immune escape of human acute myeloid leukemia cells. *EBioMedicine,* 22, 44-57.

Siska, P.J. and Rathmell, J.C., 2016. Metabolic signaling drives IFN-γ. *Cell Metabolism,* 24 (5), 651-652.

Smith, G., 2018. Step away from stepwise. *Journal of Big Data,* 5 (1), 1-12.

Smitheman, K.N., et al., 2019. Lysine specific demethylase 1 inactivation enhances differentiation and promotes cytotoxic response when combined with all-trans retinoic acid in acute myeloid leukemia across subtypes. *Haematologica,* 104 (6), 1156-1167.

Stanchina, M., et al., 2020. Advances in acute myeloid leukemia: recently approved therapies and drugs in development. *Cancers,* 12 (11), 3225.

Stawowczyk, M., et al., 2011. The interferon stimulated gene 54 promotes apoptosis. *Journal of Biological Chemistry,* 286 (9), 7257-7266.

Stomper, J., et al., 2021. Hypomethylating agents (HMA) for the treatment of acute myeloid leukemia and myelodysplastic syndromes: mechanisms of resistance and novel HMA-based therapies. *Leukemia,* , 1-17.

Student, 1908. The probable error of a mean. *Biometrika,* , 1-25.

Subramanian, A., et al., 2022. Biology-inspired data-driven quality control for scientific discovery in single-cell transcriptomics. *Genome Biology,* 23 (1), 1-27.

Sugino, N., et al., 2017. A novel LSD1 inhibitor NCD38 ameliorates MDS-related leukemia with complex karyotype by attenuating leukemia programs via activating super-enhancers. *Leukemia,* 31 (11), 2303-2314.

Sullivan, L.C., et al., 2008. The major histocompatibility complex class Ib molecule HLA-E at the interface between innate and adaptive immunity. *Tissue Antigens,* 72 (5), 415-424.

Suzuki, T., et al., 2017. RUNX1 regulates site specificity of DNA demethylation by recruitment of DNA demethylation machineries in hematopoietic cells. *Blood Advances,* 1 (20), 1699-1711.

Tabe, Y., Konopleva, M. and Andreeff, M., 2020. Fatty acid metabolism, bone marrow adipocytes, and AML. *Frontiers in Oncology,* 10, 155.

Tadokoro, Y., et al., 2007. De novo DNA methyltransferase is essential for self-renewal, but not for differentiation, in hematopoietic stem cells. *The Journal of Experimental Medicine,* 204 (4), 715-722.

Taghiloo, S. and Asgarian-Omran, H., 2021. Immune evasion mechanisms in acute myeloid leukemia: A focus on immune checkpoint pathways. *Critical Reviews in Oncology/Hematology,* 157, 103164.

Tagliafico, E., et al., 2006. Identification of a molecular signature predictive of sensitivity to differentiation induction in acute myeloid leukemia. *Leukemia,* 20 (10), 1751-1758.

Tan, J., et al., 2019. Regulation of protein activity and cellular functions mediated by molecularly evolved nucleic acids. *Angewandte Chemie International Edition,* 58 (6), 1621-1625.

Tang, L., et al., 2019. Signatures of Immune Dysfunction and Identification of Prognostic Immune-Related Risk Factors in Patients with Acute Myeloid Leukemia. *Blood,* 134, 1399.

Tanimura, S. and Takeda, K., 2017. ERK signalling as a regulator of cell motility. *The Journal of Biochemistry,* 162 (3), 145-154.

Tate, J.G., et al., 2019. COSMIC: the catalogue of somatic mutations in cancer. *Nucleic Acids Research,* 47 (D1), D941-D947.

Tettamanti, S., et al., 2022. Catch me if you can: How AML and its niche escape immunotherapy. *Leukemia,* 36 (1), 13-22.

Thakur, A., Mikkelsen, H. and Jungersen, G., 2019. Intracellular Pathogens: Host Immunity and Microbial Persistence Strategies. *Journal of Immunology Research,* 2019.

Toniato, E., et al., 2002. TRIM8/GERP RING finger protein interacts with SOCS-1. *Journal of Biological Chemistry,* 277 (40), 37315-37322.

Trendowski, M., 2015. The inherent metastasis of leukaemia and its exploitation by sonodynamic therapy. *Critical Reviews in Oncology/Hematology,* 94 (2), 149-163.

Tsai, M., et al., 2007. Gene expression profiling of breast, prostate, and glioma cells following single versus fractionated doses of radiation. *Cancer Research,* 67 (8), 3845-3852.

Tyleckova, J., et al., 2012. Cancer cell response to anthracyclines effects: Mysteries of the hidden proteins associated with these drugs. *International Journal of Molecular Sciences,* 13 (12), 15536-15564.

Ullmark, T., et al., 2017. Anti-apoptotic quinolinate phosphoribosyltransferase (QPRT) is a target gene of Wilms' tumor gene 1 (WT1) protein in leukemic cells. *Biochemical and Biophysical Research Communications,* 482 (4), 802-807.

Vadakekolathu, J., et al., 2020. TP53 abnormalities correlate with immune infiltration and associate with response to flotetuzumab immunotherapy in AML. *Blood Advances,* 4 (20), 5011-5024.

Vadakekolathu, J., et al., 2020. Immune landscapes predict chemotherapy resistance and immunotherapy response in acute myeloid leukemia. *Science Translational Medicine,* 12 (546), eaaz0463.

van de Loosdrecht, A.A., et al., 2018. A novel allogeneic off-the-shelf dendritic cell vaccine for post-remission treatment of elderly patients with acute myeloid leukemia. *Cancer Immunology, Immunotherapy,* 67 (10), 1505-1518.

van Dongen, J.J., et al., 2015. Minimal residual disease diagnostics in acute lymphoblastic leukemia: need for sensitive, fast, and standardized technologies. *Blood,* 125 (26), 3996-4009.

van Smeden, M., et al., 2019. Sample size for binary logistic prediction models: beyond events per variable criteria. *Statistical Methods in Medical Research,* 28 (8), 2455-2474.

Virtaneva, K., et al., 2001. Expression profiling reveals fundamental biological differences in acute myeloid leukemia with isolated trisomy 8 and normal cytogenetics. *Proceedings of the National Academy of Sciences,* 98 (3), 1124-1129.

Visani, G., et al., 2014. Low-dose lenalidomide plus cytarabine induce complete remission that can be predicted by genetic profiling in elderly acute myeloid leukemia patients. *Leukemia,* 28 (4), 967-970.

Visani, G., et al., 2017. Low-dose lenalidomide plus cytarabine in very elderly, unfit acute myeloid leukemia patients: Final result of a phase II study. *Leukemia Research,* 62, 77-83.

Vittinghoff, E. and McCulloch, C.E., 2007. Relaxing the rule of ten events per variable in logistic and Cox regression. *American Journal of Epidemiology,* 165 (6), 710-718.

Wada, J. and Kanwar, Y.S., 1997. Identification and characterization of galectin-9, a novel β-galactoside-binding mammalian lectin. *Journal of Biological Chemistry,* 272 (9), 6078-6086.

Wagner, S., et al., 2019. A parsimonious 3-gene signature predicts clinical outcomes in an acute myeloid leukemia multicohort study. *Blood Advances,* 3 (8), 1330-1346.

Wakita, S., et al., 2013. Mutations of the epigenetics-modifying gene (DNMT3a, TET2, IDH1/2) at diagnosis may induce FLT3-ITD at relapse in de novo acute myeloid leukemia. *Leukemia,* 27 (5), 1044-1052.

Walker, C.J., et al., 2021. Gene expression signature predicts relapse in adult patients with cytogenetically normal acute myeloid leukemia. *Blood Advances,* 5 (5), 1474-1482.

Walker, M.R., et al., 2003. Induction of FoxP3 and acquisition of T regulatory activity by stimulated human CD4 CD25–T cells. *The Journal of Clinical Investigation,* 112 (9), 1437-1443.

Wang, B., et al., 2024. Comprehensive characterization of IFNγ signaling in acute myeloid leukemia reveals prognostic and therapeutic strategies. *Nature Communications,* 15 (1), 1821.

Wang, F., et al., 2022. PD-L1 regulates cell proliferation and apoptosis in acute myeloid leukemia by activating PI3K-AKT signaling pathway. *Scientific Reports,* 12 (1), 11444.

Wang, J., et al., 2024. NET-related gene signature for predicting AML prognosis. *Scientific Reports,* 14 (1), 9115.

Wang, L., et al., 2011. The leukemogenicity of AML1-ETO is dependent on site-specific lysine acetylation. *Science (New York, N.Y.),* 333 (6043), 765-769.

Wang, M., et al., 2018. Development and validation of a novel RNA sequencing–based prognostic score for acute myeloid leukemia. *JNCI: Journal of the National Cancer Institute,* 110 (10), 1094-1101.

Wang, W., et al., 2017. Transcriptional regulation of antiviral interferon-stimulated genes. *Trends in Microbiology,* 25 (7), 573-584.

Wang, Z., Gerstein, M. and Snyder, M., 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics,* 10 (1), 57-63.

Ward, P.S., et al., 2010. The common feature of leukemia-associated IDH1 and IDH2 mutations is a neomorphic enzyme activity converting α-ketoglutarate to 2-hydroxyglutarate. *Cancer Cell,* 17 (3), 225-234.

Weichselbaum, R.R., et al., 2008. An interferon-related gene signature for DNA damage resistance is a predictive marker for chemotherapy and radiation for breast cancer. *Proceedings of the National Academy of Sciences,* 105 (47), 18490-18495.

Welch, B.L., 1947. The generalization of 'STUDENT'S'problem when several different population varlances are involved. *Biometrika,* 34 (1-2), 28-35.

Wertheim, G.B., et al., 2014. Microsphere-based multiplex analysis of DNA methylation in acute myeloid leukemia. *The Journal of Molecular Diagnostics,* 16 (2), 207-215.

Williams, B.A., et al., 2019. Antibody therapies for acute myeloid leukemia: unconjugated, toxin-conjugated, radio-conjugated and multivalent formats. *Journal of Clinical Medicine,* 8 (8), 1261.

Winer, E.S. and Stone, R.M., 2019. Novel therapy in acute myeloid leukemia (AML): moving toward targeted approaches. *Therapeutic Advances in Hematology,* 10, 2040620719860645.

Wong, K.K., Lawrie, C.H. and Green, T.M., 2019. Oncogenic roles and inhibitors of DNMT1, DNMT3A, and DNMT3B in acute myeloid leukaemia. *Biomarker Insights,* 14, 1177271919846454.

Wu, J.X., et al., 2016. SWATH Mass Spectrometry Performance Using Extended Peptide MS/MS Assay Libraries. *Molecular & Cellular Proteomics : MCP,* 15 (7), 2501-2514.

Wu, J., et al., 2020. A single-cell survey of cellular hierarchy in acute myeloid leukemia. *Journal of Hematology & Oncology,* 13 (1), 1-19.

Xu, Q., et al., 2022. Immune-related gene signature predicts clinical outcomes and immunotherapy response in acute myeloid leukemia. *Cancer Medicine,* 11 (17), 3364-3380.

Yan, W., et al., 2008. Unfavourable clinical implications for HLA-G expression in acute myeloid leukaemia. *Journal of Cellular and Molecular Medicine,* 12 (3), 889-898.

Yang, J., et al., 2018. Early tapering of immunosuppressive agents after HLA-matched donor transplantation can improve the survival of patients with advanced acute myeloid leukemia. *Annals of Hematology,* 97, 497-507.

Yang, X., Wong, M.P.M. and Ng, R.K., 2019a. Aberrant DNA methylation in acute myeloid leukemia and its clinical implications. *International Journal of Molecular Sciences,* 20 (18), 4576.

Yang, X., Wong, M.P.M. and Ng, R.K., 2019b. Aberrant DNA methylation in acute myeloid leukemia and its clinical implications. *International Journal of Molecular Sciences,* 20 (18), 4576.

Yang, Y., et al., 2020. Myeloid-derived suppressor cells in tumors: from mechanisms to antigen specificity and microenvironmental regulation. *Frontiers in Immunology,* 11, 1371.

Yap, G.L., et al., 2020. Annexin-A1 promotes RIG-I-dependent signaling and apoptosis via regulation of the IRF3–IFNAR–STAT1–IFIT1 pathway in A549 lung epithelial cells. *Cell Death & Disease,* 11 (6), 1-12.

Yin, C., et al., 2021. High expression of Clec11a predicts favorable prognosis in acute myeloid leukemia. *Frontiers in Oncology,* 11, 608932.

Yip-Schneider, M.T., Horie, M. and Broxmeyer, H.E., 1995. Transcriptional induction of pim-1 protein kinase gene expression by interferon gamma and posttranscriptional effects on costimulation with steel factor.

You, M., Yu, D.H. and Feng, G.S., 1999. Shp-2 tyrosine phosphatase functions as a negative regulator of the interferon-stimulated Jak/STAT pathway. *Molecular and Cellular Biology,* 19 (3), 2416-2424.

Yousefi, B., et al., 2017. Balaglitazone reverses P-glycoprotein-mediated multidrug resistance via upregulation of PTEN in a PPARγ-dependent manner in leukemia cells. *Tumor Biology,* 39 (10), 1010428317716501.

Zaidi, M.R. and Merlino, G., 2011. The two faces of interferon-γ in cancer. *Clinical Cancer Research,* .

Zhang, J., Gu, Y. and Chen, B., 2019. Mechanisms of drug resistance in acute myeloid leukemia. *OncoTargets and Therapy,* 12, 1937.

Zhang, W. and Liu, H.T., 2002. MAPK signal pathways in the regulation of cell proliferation in mammalian cells. *Cell Research,* 12 (1), 9-18.

Zhang, W., et al., 2020. Advances in anti-tumor treatments targeting the CD47/SIRPα axis. *Frontiers in Immunology,* 11, 18.

Zhao, J., et al., 2016. The common stress responsive transcription factor ATF3 binds genomic sites enriched with p300 and H3K27ac for transcriptional regulation. *BMC Genomics,* 17 (1), 1-14.

Zheng, H., et al., 2023. IFN-γ in ovarian tumor microenvironment upregulates HLA-E expression and predicts a poor prognosis. *Journal of Ovarian Research,* 16 (1), 229.

Zhou, Y., et al., 2019. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nature Communications,* 10 (1), 1-10.

Zhu, R., et al., 2020. Identification of an immune-related gene signature based on immunogenomic landscape analysis to predict the prognosis of adult acute myeloid leukemia patients. *Frontiers in Oncology,* 10, 574939.

Zou, K.H., O'Malley, A.J. and Mauri, L., 2007. Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models. *Circulation,* 115 (5), 654-657.