The Role of HDAC9 in JAK2V617F-driven myeloproliferative neoplasms.

Dominic Lowen

Submitted in accordance with the requirements for the degree of Doctor of Philosophy.

The University of Leeds

School of Molecular and Cellular Biology

Faculty of Biological Sciences

Submitted **September 2021**

The candidate confirms that the work submitted is his own and that appropriate credit has been given where reference has been made to the work of others.

Acknowledgements

I would also like to acknowledge Dr Natalia Riobo-Del Galdo for taking me in during my final year and for her help and support during the difficult COVID-19 pandemic.

I would like to thank Professor Westhead for his help with Chapter 4.

I want to thank Dr Alex Timmis and Dr Cintli Morales Alcala for their support and friendship throughout the PhD.

Finally, I would like to thank my family for their support.

Abstract

Janus kinase 2 (JAK2) encodes a protein tyrosine kinase that is critical for haematopoiesis. JAK2 functions by modulating signal transduction between the receptor and the JAK-STAT, ERK1/2 and AKT phosphorylation signalling pathways. The mutation, JAK2V617F, loses the ability to inhibit auto-phosphorylation, the result of which is a chronically active kinase which causes aberrant signalling. The influence of JAK2V617F has yet to be fully elucidated in haematopoietic stem cells (HSCs). In this work I present a new model for JAK2V617F in HSCs and have identified a gene, HDAC9, that has potential as a possible target for treatment.

In Chapter 3, I describe the experiments that explain the construction of my model and how I probed it with the aim to characterise it with a focus on signalling and DNA damage/DNA damage repair. An investigation into DNA damage/repair suggests that JAK2V617F exerts both increased DNA damage and an increased ability to repair such damage.

In Chapter 4, I provide a cross-publication dataset analysis to compare the RNA-sequencing analysis of my data to current models and patient data to gauge the relevance of my data. I utilise my RNA-sequencing data alongside bioinformatic programs to explain how the data relays back to Chapter 3. I have provided a comprehensive validation programme to validate my RNA-sequencing data through RT-PCR.

Finally in Chapter 5, I investigate the role of HDAC9 in cell signalling, cell proliferation, and DNA damage response. The data presented within this chapter suggests that HDAC9 is essential for JAK-STAT signalling in HSCs and through STAT regulation, has a role in cell proliferation and DNA damage.

Altogether, these findings reveal new insight into the role of HDAC9 in JAK2V617F expressing HSCs.

Table of contents

List of Tables

List of Figures

Abbreviations

(M)SCF – (murine) stem cell factor

4-OHT – 4-hydroxytamoxifen

AKT also known as PKB – protein kinase B

BCL – B cell lymphoma

BCR-ABL – breakpoint cluster region- Abelson murine leukaemia

BFU-E – burst-forming unit-erythroid

BSA – bovine serum albumin

CALR – calreticulin

CBX5 – Chromobox protein homolog 5

CD – cluster of differentiation

CLP - common lymphoid progenitors

CML – chronic myeloid leukaemia

CMP - common myeloid progenitors

CpG – cytosine-phosphate-guanine

CRM – cytokine-receptor molecules

CTBP – c-terminal binding protein

DDR – DNA damage repair

DGE – differential gene expression

DNMT3a – DNA methyltransferase 3a

ECD – extra cellular domain

EEC – endogenous erythroid colony

eGFP – enhanced green fluorescent protein

EGFR – endothelial growth factor receptor

EPO - erythropoietin

EPOR – erythropoietin receptor

ERK – extracellular signal-related kinase

ET – essential thrombocythaemia

EZH2 – enhancer of zeste homologue 2

FBS – foetal Bovine Serum

FERM - four-point-one, ezrin, radixin, moeisin

FSC – forward scatter

G6PD – glucose-6-phosphate dehydrogenase

G-CSF – granulocyte – colony-stimulating factor

GEO – gene expression omnibus

GM-CSF – granulocyte-macrophage Colony-Stimulating factor

GMP - granulocyte-macrophage progenitors

GSEA – gene set enrichment analysis

H3K9Ac – histone 3 lysine 9 acetylation

H4K16Ac - histone 4 lysine 16 acetylation

HDAC – histone deacetylase

HEK – human embryonic kidney

HPC7- haematopoietic precursor cell line 7

HSC - haematopoietic stem cell

IFN – interferon

IGF1 – insulin growth-like factor 1

IL – interleukin

IP- immunoprecipitation

JAK - Janus kinase

JAK2V617F – Janus kinase valine 617 phenylalanine

JAK2WT – Janus kinase wild-type

LH2 – Lim-homeobox 2

LIF – leukaemia inhibitory factor

LNK – lymphocyte adaptor protein

LOH – loss of homozygosity

LSK – lineage negative, stem cell antigen 1 positive and c-Kit positive

LT-HSC – long-term haematopoietic stem cell

MAPK – mitogen-activating protein kinase

MEF – myocyte-enhancing factor

MEP - megakaryocyte-erythrocyte progenitors

MEP50 – methylosome protein 50

MPL – c-myeloproliferative leukaemia

MPN – myeloproliferative neoplasm

MPP - multipotent progenitors

mRNA – messenger RNA

mTOR – mammalian target of rapamycin

MTT – 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide

NGS – next generation sequencing

NK - natural killer

ORF – open-reading frame

Penstrep – penicillin/streptomycin

PI3K – phosphoinositide 3-kinase

PIAS – protein inhibitors of activated STATs

PBS – phosphate-buffered saline

PKC – protein kinase C

PMF – primary myelofibrosis

PRC2 – polycomb repressive complex 2

PRMT5 – protein arginine N-methyltransferase 5

PTP1B – protein tyrosine phosphatase 1B

PV – polycythaemia vera

ROS – reactive oxygen species

rRNA -ribosomal RNA

RT-PCR – reverse transcriptase - polymerase chain reaction

SDS-PAGE – sodium dodecyl sulphate - polyacrylamide gel electrophoresis

SHPs - SH2 domain containing protein phosphatases

shRNA – short-hairpin RNA

siRNA – short-interfering RNA

SOC – super optimal broth with catabolite repression

SOCS – suppressors of cytokine signalling

SSC – side scatter

STAT - signal transducer and transcription factor

ST-HSC - short-term haematopoietic stem cell

SUMO – small ubiquitin-like modifying proteins

TAD – trans activation domain

TC-PTP - T cell protein tyrosine phosphatases

TET2 – ten-eleven translocation 2

TM – transmembrane

TORF – truncated open-reading frame

TPO – thrombopoietin

tRNA – transfer RNA

TRP53 – tumour-repressor protein 53

TTO – terminal tagging oligo

TYK – tyrosine kinase

UV – ultraviolet

Chapter 1 Literature review

Chapter 1

Literature Review

1.1 Overview of haematopoiesis

Haematopoiesis refers to the dynamic process of blood cell formation and replacement throughout the life of an organism. This conserved and hierarchical process begins with haematopoietic stem cells (HSCs) (Lorenz et al., 1951). HSCs can instigate self-renewal to give rise to more daughter HSCs or differentiate. Upon differentiation, HSCs lose self-renewal potential and terminally differentiate into specialised blood cells of a specific haematopoietic lineage **(Figure 1.1)** (Orkin, 2000). HSCs give rise to two main haematopoietic lineages known as the myeloid and the lymphoid lineages (Traver and Akashi, 2004).

The myeloid lineage includes erythrocytes, megakaryocytes, neutrophils, macrophages, eosinophils, and mast cells. The lymphoid lineage is comprised of natural killer (NK) cells, B lymphocytes and T lymphocytes. Terminally differentiated haematopoietic cell types are highly specialised and as a result can participate in distinct biological processes and perform different functions. Most of the time HSCs maintain a quiescent state in the G0 phase of the cell cycle (Rossi et al., 2007). External/internal stimuli can trigger the activation of global epigenetic changes that stimulate HSCs to differentiate and supplement the demand for new blood cells, when this requirement has been achieved HSCs return to quiescence (Wilson et al., 2008).

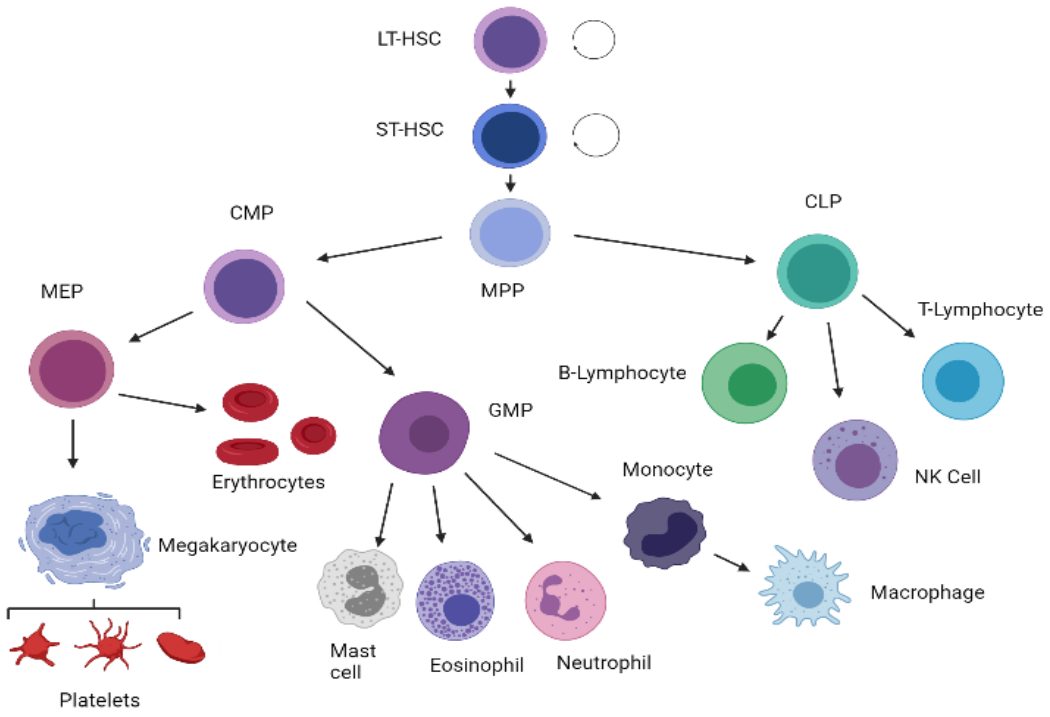


Figure 1.1 Overview of the haematopoietic system.

Classical view of the haematopoietic system. The LT-HSCs and ST-HSCs can self-renew or differentiate into multipotent progenitors which can either become CMPs or CLPs. CMPs have the potential to terminally differentiate into megakaryocytes, erythrocytes, granulocytes, or macrophages. CLPs have the potential to differentiate into B lymphocytes, T lymphocytes or NK cells. The lineage to which a progenitor differentiates to is determined by many factors, mainly through the activation of receptors via cytokine-receptor stimulation by different cytokines.

The abbreviations for the figure are LT-HSC for Long-Term Haematopoietic Stem Cells, ST-HSC for Short-Term Haematopoietic Stem Cells, MPP for MultiPotent Progenitors, CMP for Common Myeloid Progenitors, CLP for Common Lymphoid Progenitors, MEP for Megakaryocyte-Erythrocyte Progenitor, GMP for Granulocyte-Monocyte Progenitors and NK - Natural Killer cell.

Created with Biorender.com

Dysregulation of steady-state haematopoiesis is associated with the development of haematological malignancies whose phenotype can be defined by a saturation of blast cells or terminally differentiated cells in the haematopoietic system, or fibrosis of the bone marrow (Warr et al., 2011).

1.2 The role of JAK2 in haematopoiesis

1.2.1 JAK2-STAT signalling

Haematopoiesis is cytokine-dependent process with cytokines being integral for prompting a myriad of responses that drive self-renewal, proliferation, and terminal differentiation (Metcalf, 2008). Since the discovery of certain specific cytokines, it is now understood that some cytokines assist the formation of a specific lineage whilst others have a more generalised range of functions, either alone or in combination with other cytokines to stimulate the desired cellular output. For cytokines to exert their response, they must bind to their complementary cellular membrane-bound receptors on the extracellular domain which transduces signals intracellularly. Broadly, there are two types of receptor families that are structurally related and are involved in haematopoiesis.

Type I receptors are known to bind to cytokines, these cytokines include: most interleukins (IL), colony stimulating factors, Erythropoietin (EPO) and Thrombopoietin (TPO). Type II receptors bind to interferons (IFN-α, β, and γ) and members of the IL-10 family (Gadina et al., 2001). Many cytokine receptors cannot transduce a signal by themselves and therefore require additional cytoplasmic tyrosine kinases. One of the most important kinase families in haematopoiesis is the Janus Kinase family (JAKs). JAKs are required to transduce a signal and stimulate a signalling cascade. Activation of both type I and type II receptors is triggered following appropriate ligand binding to the relevant ligand-binding domains on the extracellular domain of the receptor, which results in either receptor oligomerisation, homodimerisation or induction

of conformational change in preformed receptor dimers/multimers, depending on the receptor being activated (Robb, 2007). Receptor-ligand binding leads to recruitment of JAKs and the activation of JAKs (Staerk and Constantinescu, 2012). In 2020, Wilmes et al. proposed a molecular mechanism where ligand-induced dimerisation is the fundamental switch that initiates activation of EPO, TPO and growth hormone receptors. This contradicts the current theory that preformed dimers undergo structural conformational changes upon ligand binding. In addition to this, Wilmes et al. reported that truncations in the JH2-JH1 of the JAK2 protein significantly reduced the ligand-induced dimerisation of TPO/EPO/growth hormone receptors, whereas a JH1 truncation did not affect dimerisation. This suggests that the JH2 domain of JAK2 is required for effective ligand-induced dimerisation and stabilisation (Wilmes et al., 2020).

There are four mammalian JAKs that are essential for transducing receptor signals: JAK1, JAK2, JAK3 and TYK2 (Yamaoka et al., 2004). The JAK family of proteins are characterised by four domains: (1) the 4.1 protein, Ezrin, Radixin, Moesin (FERM) domain located in the N-terminus (Ihle, James N., 1995); (2) an SH2-like domain which is located in the N-terminal part/region of the protein; (3) a JH2 pseudokinase regulatory domain and (4) a JH1 kinase domain.

The current theory is that in a non-stimulated state, the cytoplasmic domains of receptors remain distant from one another which results in an inactive form. When activated, JAK proteins are recruited to the intracellular domain of cytokine receptor signalling chains and bind through the JAK protein's FERM domain. Ligand-cytokine receptor complex formation triggers 3D structural change in the receptor's cytoplasmic domain or brings monomers of receptors towards each other and the JAK proteins bound on the receptor are brought in close proximity to one another, allowing for the transphosphorylation of neighbouring JAK molecules which leads to JAK activation (Rawlings et al., 2004) (**Figure 1.2**).

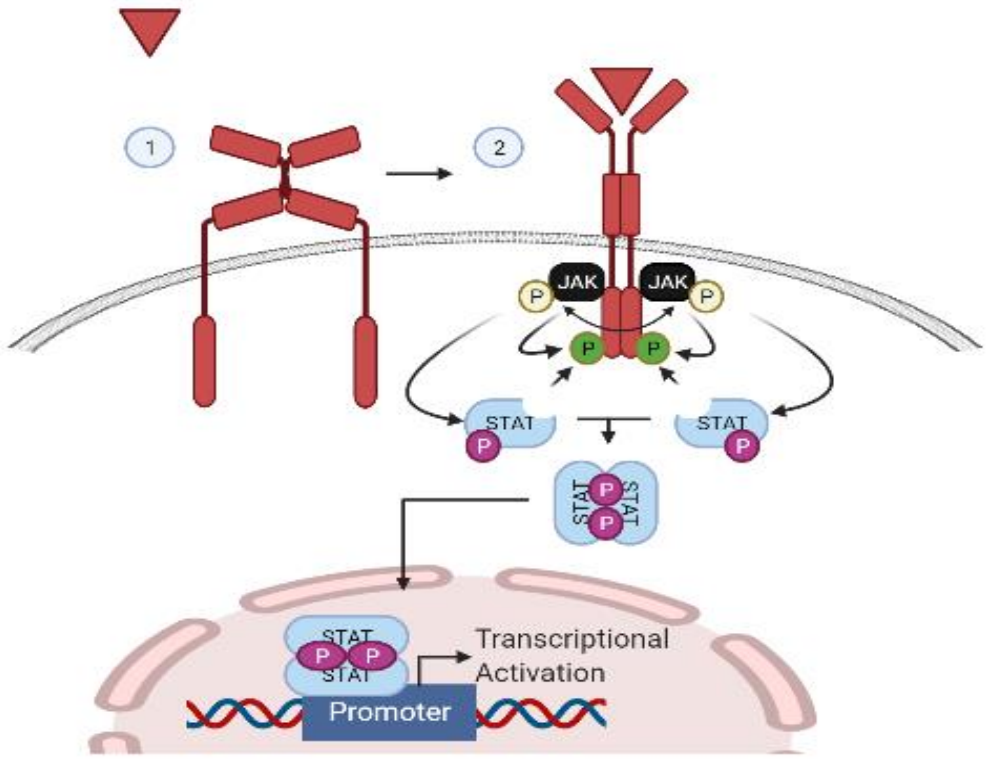


Figure 1.2 The activation of the JAK-STAT pathway.

The receptor in this image represents erythropoietin Receptor (EPOR). The unstimulated form of EPOR is represented by (1, on the left) and the stimulated form is represented by (2, on the right).

Upon stimulation of EPOR with EPO, EPOR undergoes 3D structural changes and simultaneously recruits JAK proteins to its cytoplasmic domains. JAK kinases transphosphorylate each other which allows the cytoplasmic domains of the receptors to become phosphorylated. The resulting phosphorylation of the receptor acts as a binding site that allows recruitment of STATs. Upon STAT binding, JAK kinases phosphorylate STATs. STAT phosphorylation stimulates dimerisation and translocation to the nucleus where STATs bind promoters of genes, instigating transcriptional activation.

Created with Biorender.com

JAK proteins are transphosphorylated by other JAK proteins within the JH1 domain, specifically in the activation loop (Ihle, J. N. and Gilliland, 2007). This phosphorylation reduces the phosphorylated activation loop's ability to interact with and regulate the JH2 pseudokinase domain. The JH2 pseudokinase domain keeps the JH1 in an inactive state and phosphorylation within the JH1 activation loop reduces JH2-induced repression and allows continued JAK activation. JAK is further be regulated by phosphorylated tyrosine residues located within in the FERM domain (Funakoshi-Tago et al., 2006).

Activated JAK kinases at the receptor phosphorylate key tyrosine residues within the intracellular domain of the cytokine receptor. These phosphorylated tyrosine residues then act as docking sites for STAT proteins.
In mammals, seven signal transducers and activator of transcription (STAT) proteins exist (STAT1, STAT2, STAT3, STAT4, STAT5A, STAT5B and STAT6). Upon recruitment of STATs to the receptor, STAT proteins are phosphorylated on one tyrosine residue by receptor-bound JAKs within their C-terminal transactivation domain (TAD) (Benekli et al., 2003). STAT phosphorylation triggers heterodimerisation and homodimerization of STAT proteins which facilitates their translocation from the cytoplasm into the nucleus to act as transcription factors and activate or represses transcription of target genes. Specific combinations of cytokines, receptors, JAKs and STATs are critical in regulating haematopoiesis. Generally, STAT1, STAT3 and STAT5a/b regulate signal transduction in factor-regulated control of erythropoiesis, granulopoiesis

and megakaryopoiesis, whereas STAT4 and STAT6 are required for immune cells.

The JAK/STAT signalling pathway is tightly regulated through the balance of cytokines that activate the signalling cascade and various mechanisms that antagonise these signals. Cytokine signalling is attenuated by negative regulators that dephosphorylate JAKs and inhibits STAT activation. In addition to the intrinsic inhibitory activity of JH2 domain of JAK in attenuating signalling, several additional mechanisms have evolved to regulate the duration and intensity of JAK-STAT signalling. These are mainly mediated by three classes: (i) protein phosphatases; (ii) the protein inhibitors of activated STAT (PIAS) family; and (iii) the suppressor of cytokine signalling (SOCS) family (Starr and Hilton, 1999; Naka et al., 1997). Protein phosphatases modulate signalling by dephosphorylating tyrosine residues to down-regulate JAK kinase activity.

Several phosphatases that have been previously identified include SH2 domain containing tyrosine phosphatases (SHPs), T-cell protein tyrosine phosphatase (TC-PTP), protein tyrosine phosphatase 1B (PTP1B) and transmembrane protein CD45 (Murphy et al., 2010). PIAS proteins attenuate signalling through SUMOylation, where a small ubiquitin-like modifier protein (SUMO) is attached to a lysine residue and targets proteins for proteasomal degradation (Ungureanu et al., 2003). Finally, SOCS proteins can attenuate signalling directly by binding to activated JAKs and binding to phosphorylated receptors to outcompete STAT binding to the receptor (Starr and Hilton, 1999). Altogether,

this demonstrates that regulation of JAK-STAT signalling can be mediated through inhibition of the receptors, JAKs and STATs.

STAT proteins are regulated by a plethora of post-translational modifications, for the purposes of the following chapters I will briefly discuss STAT5 acetylation and SUMOylation and the relationship with STAT5 phosphorylation, referred to as the phospho-acetyl/SUMOylation switch. Upon stimulation of the receptor with the appropriate ligand the receptor recruits JAK kinases which results in the phosphorylation of STATs. STAT5 phosphorylation is followed by acetylation of residue lysine 696 by CBP/P300. STAT5 acetylation in Treg cells has been found to be necessary for effective STAT5-driven transcriptional activation through the stabilisation of the STAT5 dimer (Beier et al., 2012). STAT5 SUMOylation is catalysed by PIAS3 on lysine 696 and competes with acetyl modifications: the addition of a SUMO molecule blocks phosphorylation and acetylation modifications and occurs subsequent to the transcriptional activation by STAT5 as a way of removing the activation signal and returning STAT5 to an unactive form or signalling for degradation (Ma et al., 2010). Interestingly, the SUMOylation only occurs in the lymphoid compartment and does not occur in the HSC or myeloid compartment with levels of SUMOlyation being beneath detectable levels and acetylation not being necessary for differentiation in HSC/myeloid compartments (Beier et al., 2012).

1.2.2 The non-canonical role of JAK2

In addition to JAK2-STAT signalling, it has been uncovered that JAK2 also has non-canonical roles in epigenetic signalling. In 2009, Dawson et al. observed via immunofluorescence studies, that in HEL, K562, UKE1 cell lines and primary $CD34^+$ cells, fluorescently tagged JAK2/JAK2V617F was present in the nucleus as well as the cytoplasm, indicating a nuclear role. Further studies revealed that JAK2/JAK2V617F was bound to histone H3, specifically histone H3 tyrosine 41 (H3Y41); as JAK2 is a kinase, the group hypothesised that JAK2 phosphorylates H3Y41. They verified this hypothesis using an antibody they generated against H3Y41 phosphorylation (H3Y41Ph) and TG101209 (a JAK2 inhibitor). The inhibition of JAK2 was found to prevent the binding of Chromobox protein homolog 5 (CBX5) to histone H3 (Dawson et al., 2009). CBX5 is a component of the polycomb repressive complex 2 (PRC2) complex which through the enhancer of zeste homolog 2 (EZH2) subunit acts as a transcriptional repressive complex through the trimethylation of histone H3 lysine 27 (H3K27).

JAK2 has also been reported to bind protein arginine N-methyltransferase 5 (PRMT5). In 2010, Liu et al. observed that JAK2 can bind PRMT5 and notably, JAK2 mutants JAK2V617F and JAK2K539L, bind to PRMT5 with greater affinity than their wild-type counterpart. PRMT5 was found to have higher levels of phosphorylation in JAK mutant-expressing HEL cells which resulted in PRMT5 being unable to methylate arginine residues on all histone proteins and this phosphorylation blocks PRMT5 binding methylosome protein 50 (MEP50). The reduction in arginine methylation of histones and binding MEP50 instigated

increases in gene expression of genes associated with erythroid expansion (Liu, F. et al., 2011).

1.3 The role of JAK2 in differentiation

1.3.1 JAK2 signalling in granulocyte development

Granulopoiesis is the process of developing granulocytes which gives rise to neutrophils, eosinophils, and basophils (Bainton et al., 1971). The cytokines that direct granulopoiesis are primarily colony-stimulating factors such as stem cell factor (SCF), granulocyte colony-stimulating factor (G-CSF), granulocyte-macrophage colony-stimulating factor (GM-CSF) and interleukin-3 (IL3) (Kawahara, 2007). G-CSF and GM-CSF have redundant functions in granulopoiesis and are dependent on signalling pathways via their complementary receptors, G-CSFR and GM-CSFR, respectively. Interestingly, only G-CSF depletion results in a granulocytic differentiation block which indicates that G-CSF must play a more significant role than GM-CSF in promoting granulocyte development (Lieschke et al., 1994).

Studies of the molecular mechanisms of G-CSF and GM-CSF activity revealed that these cytokines activate JAK/STAT, PI3K/AKT, and MAPK pathways downstream of their receptors (Marino and Roguin, 2008; Shi et al., 2006). Binding of these cytokines initiates receptor conformational changes that lead to autophosphorylation of JAK2 and activation of STATs, PI3K and MAPK (Dijkers et al., 1999; Jenkins et al., 1998).

These signalling pathways drive cell proliferation, differentiation, and resistance to apoptosis (Shimoda et al., 1994). Dysregulation of these pathways frequently leads to malignancies of the myeloid lineages.

1.3.2 JAK2 signalling in megakaryocyte development

TPO is the main regulator of megakaryopoeisis, a process which refers to the proliferation and development of megakaryocytes and thrombocytes (Kaushansky, 1997; Eaton and de Sauvage, 1997). Mouse models where the receptor for TPO, the c-myeloproliferative leukaemia receptor (MPL), has been depleted exhibit thrombocytopaenia with a deficiency in megakaryocytes (Ng et al., 2014), suggesting that TPO is required for regulating the megakaryocytic population.

Once the haematopoietic progenitor is committed to the megakaryocyte lineage, TPO also mediates the expression of specific cell surface proteins that are involved in platelet function, explicitly, glycoprotein IIb/IIIa and glycoprotein Ib (French and Seligsohn, 2000; Doubeikovski et al., 1997). In addition to regulating megakaryocytic maturation, there is also evidence suggesting that TPO is not constrained to megakaryocytes but is actually a far broader cytokine required for many other process. For example, TPO has
also been demonstrated to operate with EPO to stimulate proliferation of erythroid cells *in vitro* (Yoshihara et al., 2007; Kobayashi et al., 1995; Kaushansky et al., 1995).

The biochemical mechanism of TPO activity is by binding directly to its cell surface receptor MPL. TPO-MPL binding induces receptor dimerisation and subsequent JAK2 activation (Hitchcock and Kaushansky, 2014), that leads to phosphorylation of tyrosine residues Tyr-626 and Tyr-631 within the cytoplasmic domain of the receptor itself. This phosphorylation allows recruitment of STAT1, STAT3 and STAT5a/b proteins that lead to the intracellular signal transduction cascade that gives rise to mature megakaryocytes and subsequent platelet production (Fishley and Alexander, 2004).

Thrombopoiesis can also be regulated by an adapter protein referred to as LNK. LNK has increased activity upon TPO stimulation and acts as a negative regulator by restricting TPO-MPL signalling. The mechanism of how LNK can negatively regulate signalling is still under investigation, but the current hypothesis is that LNK regulates signalling by binding phosphorylated tyrosine resides of the juxta membrane domains of receptors. The binding of LNK to receptors is theorised to block downstream signalling by preventing the recruitment of signalling molecules to the phosphorylated tyrosine residues, thus preventing the transduction of a signalling cascade (Gery and Koeffler, 2013; Bersenev et al., 2008).

1.3.3 JAK2 signalling in erythrocyte development

Erythropoiesis is a highly regulated process that produces mature erythrocytes (red blood cells) from an erythroid progenitors in the bone marrow (Zivot et al., 2018; Orkin, 2000). This process is primarily controlled by a cytokine called EPO, a ligand specifically for EPOR which is found on the cell surface of erythroid progenitors (Spivak, 2005). EPO is produced in the kidney and is secreted into the blood upon exposure to stimuli such as bleeding or hypoxia (Chateauvieux et al., 2011). The binding of EPO and EPOR triggers 3D structural changes of the extracellular and intracellular domains of EPOR and consequently autophosphorylation and activation of JAK2 (Witthuhn et al., 1993). JAK2 activation leads to phosphorylation of eight specific tyrosine residues localised on the cytoplasmic region of EPOR that serve as docking sites to recruit various SH2-domain containing proteins to initiate four main signalling pathways that are associated with EPOR activation: (i) JAK2/STAT5, (ii) protein kinase C (PKC) and (iv) phosphoinositide 3-kinase (PI3K)/AKT and (iii) extracellular signal-regulated kinase (ERK) (Hedley et al., 2011; Frank, 2002; Lacombe and Mayeux, 1999; Bao et al., 1999). These signalling pathways regulate erythroid proliferation, differentiation and also suppresses pro-apoptotic signals. EPO suppresses apoptosis through EPOR and the stimulation of the JAK2-STAT5 pathway which upregulates anti-apoptotic genes such as *BCL-X*$_L$ and through increased phosphorylation of ERK which can inhibit caspase 3, a pro-apoptotic signal.

Altogether, these studies highlight the role of JAK2 signalling transduction downstream of various receptors that are essential for directing myeloid

differentiation. Dysregulation of these pathways is frequently leads to malignancies of the myeloid lineages.

1.4 Myeloproliferative neoplasms

1.4.1 A brief history of Myeloproliferative Neoplasms (MPNs)

From the early 1800s to the early 1900s, clinical cases had been reported concerning patients suffering from hyperactivity of haematopoietic organs. The reported cases noted common symptoms such as the splenomegaly and the increased viscosity of the blood (Osler, 2008; Means, 2008; Epstein and Goedel, 1934; Heuck, 1879). In 1879, a patient was first reported and described as having Primary myelofibrosis (PMF) with extramedullary haematopoiesis, osteosclerosis and bone marrow fibrosis (Heuck, 1879). The first description of polycythaemia vera (PV) reported symptoms of excessive and persistent erythrocytosis (Vaquez, 1892). This was confirmed to be a separate disorder compared with the later observed secondary polycythaemia (Osler, 2008; Means, 2008). Essential thrombocythaemia (ET) was the last MPN to be classed as a discrete clinical disorder and was associated with haemorrhagic thrombocythemia and elevated platelets along with evidence of a sclerotic spleen (Epstein and Goedel, 1934). These diseases presented diverse phenotypes with highly diverse symptoms and their cause was unknown. However, these disorders were assumed to be related diseases from a similar origin due to the resemblances in clinical pathogenesis (Vaughan and Harrison, 1939; Hirsch, 1935).

In 1951, William Dameshek collectively categorised the four clinicopathologic disorders, chronic myeloid leukaemia (CML), ET, PV and PMF under the generalised term 'myeloproliferative disorders' (MPD) (Dameshek, 1950). Dameshek primarily highlighted that bone marrow cells such as erythroblasts and megakaryocytes were saturating the blood which indicated that these cells were being chronically produced in these patients. Dameshek postulated that the overlapping clinical and morphological features of these disorders were attributed to the myeloproliferation in the bone marrow. Furthermore, Dameshek suggested that an unknown factor was responsible for governing the development of these malignancies (Tefferi, 2008).

MPN are regarded as a group of clonal, haematopoietic disorders originating in HSCs and are identified by their overproduction of terminally differentiated cells of the myeloid lineage (Nangalia and Green, 2014; Levine, 2012). The four classical MPNs are ET, PV, CML and MF. CML is often separated from the other MPNs; this is because CML is defined by a reciprocal translocation of chromosomes 9 and 22, referred to as being Philadelphia chromosome positive. This translocation translates into a BCR-ABL fusion protein being expressed. For this thesis, I will only focus on the three Philadelphia chromosome negative MPNs.

ET is characterised by an elevated platelet count accompanied by an overproduction of megakaryocytes. PV is characterised by a chronic production of erythroid cells with some hyperplasia of the granulocytic/megakaryocytic lineages. PMF presents a heterogeneous disorder defined by megakaryocytic

hyperplasia and fibrotic bone marrow. In all diseases an HSC acquires an MPN-driving mutation(s) that provides the MPN HSCs with a clonal advantage over normal HSCs to drive terminal myeloid differentiation and leads to increased production of a specific myeloid lineage. Additionally, depending on secondary mutations in epigenetic regulators ,such as DNMT3a or EZH2, all three disorders have the potential to transform into secondary acute myeloid leukaemia (AML).

1.4.2 Cytokine signalling in MPNs

As haematopoiesis is a cytokine-dependent process, it was commonly believed that MPNs could be a disorder stemming from hypersensitivity or independence to cytokine stimulation. In 1974 it was shown that *in vitro* culture of bone marrow cells from PV patients could give rise to erythroid colonies in the absence of cytokines such as EPO highlighting cytokine independence (Prchal and Axelrad, 1974). This observation of cytokine-independent growth of endogenous erythroid colony (EEC) was characteristic of PV and was replicated with primary cells of a subset of patients with ET and PMF (Kawasaki et al., 2001; Griesshammer et al., 1998; Juvonen et al., 1993; Gewirtz et al., 1983; Zanjani et al., 1977; Adamson, 1968).

These reports were followed by evidence of hypersensitivity of MPN cells to other cytokines such as interleukin-3 (IL-3), granulocyte-macrophage colony-stimulating factor (GM-CSF) and insulin growth-like factor 1 (IGF1) (Dai, W.J. et al., 1997; Montagna et al., 1994; Correa et al., 1994; Dai, C.H. et al., 1992; de Wolf et al., 1989).

1.5 Driver mutations of MPNs

The cytokine hypersensitivity/independent nature of MPNs led to the theory that mutations in cytokine receptors could be responsible for MPNs. Early studies however revealed no mutations in the coding and non-coding regions of the *EPOR* gene or the *MPL* gene in a small sample of ET, PV, and MF patients (Randi et al., 2005; Wang, J.C. and Hashmi, 2003; Mittelman et al., 1996; Couedic et al., 1996; Hess et al., 1994). This discovery directed research towards the JAK-STAT pathway itself and the notion that signalling has become independent from cytokines and cytokine receptors.

1.5.1 JAK2V617F

In 2005, a breakthrough came when four independent groups identified a somatic gain-of-function mutation in the JAK2 gene that is present in a majority of MPN subtypes (Levine et al., 2005; Kralovics et al., 2005; James et al., 2005; Baxter et al., 2005). The lesion was a G>T transversion in nucleotide 1849 in exon 14 of the *JAK2* gene. This resulted in a substitution from valine to phenylalanine in codon 617 (JAK2V617F) (**Figure 1.3**). Over 90% of PV patients and ~50-60% of ET and MF patients were found to harbour the JAK2V617F mutation (Levine et al., 2005; Kralovics et al., 2005; James et al., 2005; Baxter et al., 2005), making it the most frequent mutation in MPN patients regardless of the type of MPN. In the remaining 9-10% of PV patients, patients harbour mutations in exon 12 of *JAK2*, mutations in exon 12 will be discussed in more detail in Chapter 1, Section 1.5.3.

The effect of the V617F mutation on JAK2 activity is constitutive hyperactivation of JAK2 kinase activity. This is thought to be achieved by disrupting the normal activity of valine-617 which is located within the JH2 pseudokinase domain and is required to inhibit the activity of the adjacent JH1 kinase domain (Saharinen and Silvennoinen, 2002) to ensure that JAK2 is kept at an inactive state in the absence of phosphorylation. The exact nature of how JH2 inhibits JH1 is unknown.

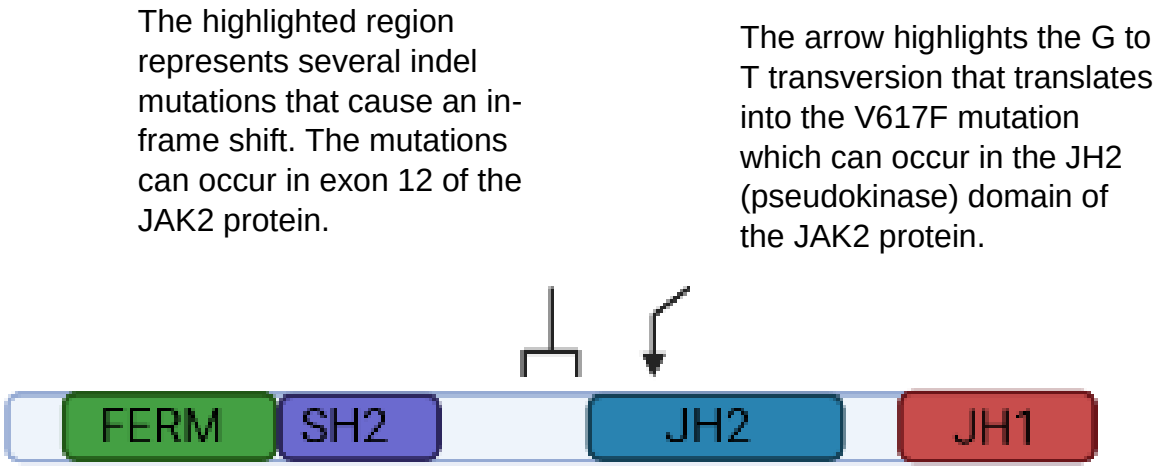


Figure 1.3 A protein schematic of the *JAK2* gene.

This figure illustrates a protein schematic of the *JAK2* gene. The figure highlights the JAK2 protein domains: FERM, SH2, JH2 (pseudokinase) and JH1 (kinase). The schematic also indicates the common mutations that can occur in the *JAK2* gene, the frequent mutation V617F mutation in the SH2 domain and the less frequent exon 12 mutations.

Created with Biorender.com

A follow-up report by Bandaranayake et al. focusing on X-ray crystallography studies performed on JAK2 and JAK2V617F concluded with structures that were very similar, with a slight difference of 0.76Å between the N lobe located in the JH2 domain of the two structures. The result of the X-ray structure illustrates that phenylalanine residues in the one helix structure located within the N lobe is rotated with Phe-594 undergoing a significant rotation rigidifying the JH2 domain overall. This suggests that mutant JAK2V617F-JH2 may promote stabilisation of the JH1 in the constitutively active, mutant form of JAK2. The same paper noted that the JAK2V617F-JH1 domain may impair the catalytic activity of JH2 through the loss of phosphorylation of Ser-523 and Tyr-570 on JH2 which are crucial negative regulators of JH1 activity (Bandaranayake et al., 2012). Overall, the data suggests that the mutation may contribute to constitutive activation of JAK2V617F by a mixture of loss of JH2 inhibitory function and a promotion of the JH1 active structure.

In 2020, Wilmes et al. reported that mutations in the JH2 domain of JAK2, including JAK2V617F, resulted in an increased interaction in the JH2-JH2 domains of the dimerisation interface between two JAK2 molecules which stabilises the receptor dimer. They hypothesise that this increased stabilisation removes the requirement of cytokine binding to the receptor and in doing so, supports cytokine independence (Wilmes et al., 2020).

Further analysis of the JAK2V617F mutant, subsequently revealed the crucial mediators downstream that are essential for MPN pathogenesis. A crucial role of STAT5 in the pathogenesis of PV was confirmed in mouse studies whereby JAK2V617F could not drive a MPN phenotype in STAT5-depleted cells (Yan, D.

et al., 2012; Walz et al., 2012). However, constitutively activated ERK, AKT and STAT3 were shown to be irrelevant for haematopoietic transformation (Yan, D. et al., 2012; Zou et al., 2011). Further publications conclude that deletion of STAT3 reduced the overall survival of JAK2V617F knock-in mice (Yan, D. et al., 2015; Grisouard et al., 2015; Yan, Dongqing and Mohi, 2013).

These data suggest that phosphorylation of STAT5 is necessary for JAK2V617F-mediated transformation whilst STAT3, ERK1/2 and AKT are not.

1.5.2 JAK2V617F in haematopoietic stem cells

To uncover the origins of MPNs and understand how JAK2V617F is acquired, several groups have examined HSCs. The logic for this is that HSCs are the source of all blood cells within the haematopoietic system: MPNs are diseases that span decades and therefore it seems reasonable to assume that JAK2V617F must have originated in a single HSC clone that has self-renewal capability to propagate and sustain itself (Fialkow, 1979).

To investigate if MPNs originated from a single HSC clone, research groups exploited polymorphisms in the X-linked glucose-6-phosphate dehydrogenase (G-6-PD) locus. The polymorphism studies utilised polymorphic G-6-PD isozymes as a marker for identifying clonality and relies on the arbitrary nature of X-chromosome inactivation (Chen, G.L. and Prchal, 2007). Tissues from non-diseased patients are predicted to exhibit similar expression of both G-6-PD isozymes due to random X-inactivation. However, MPNs derived from a single progenitor cell would exhibit a skewed expression of a singular G-6-PD isozyme due to domination of a progenitor, and MPNs with a multicellular origin should

also express both isozymes equivalently (Fialkow et al., 1981; Jacobson et al., 1978; Adamson et al., 1976; Fialkow et al., 1967).

Diseased patient samples presented highly skewed levels of a singular G-6-PD isozyme expression, supporting the hypothesis that JAK2V617F originated within a single clone (Adamson et al., 1976).

The polymorphism studies were later complemented by xenotransplantation studies. The xenotransplantation studies involved the transplantation of primary $CD45^+CD34^+$ cells of PV, post-PV MF, and primary MF patients into nonobese diabetic severe combined immunodeficient (NOD/SCID) mice. The xenotransplantation experiment revealed that the diseased $CD45^+CD34^+$ cells have the capability to self-renew long term. In further studies, $CD34^+$cells of patients were transplanted into NOD/SCID mice which revealed that $CD34^+$cells expressing JAK2V617F can also reconstitute the haematopoietic system (Ishii et al., 2007).

Much effort has gone into understanding how JAK2V617F contributes to initiation and progression of MPNs. Many more groups have investigated reconstitution capacity through the use of transplantation assays which all have indicated that mice develop a PV phenotype upon exposure to JAK2V617F. However, these results did not explain how JAK2V617F could give rise to three different diseases.

In 2008, to model JAK2V617F-driven MPNs *in vivo*, mouse models were generated by injecting the entire coding region for JAK2V617F (under the control of a human *Vav* promoter) into the pronuclei of oocytes in mice. The offspring of the transgenic mice had varying levels of JAK2V617F expression

and generated either a mild or severe PV phenotype with signs of MF. From these data, it was concluded from the experiments that the ratio of JAK2V617F to JAK2 wild type (JAK2WT) determined disease outcome (Xing et al., 2008; Shide et al., 2008). Furthermore, in 2010, Akada et al. found that LSK cells (Lineage negative, Stem cell antigen positive and c-KIT positive) and the myeloid compartment of mice underwent a significant expansion when expressing JAK2V617F (Akada et al., 2010).

Mullaly et al. reported that in their JAK2V617F knock-in mouse models, the mice expressing JAK2V617F developed a severe PV phenotype, and from serial transplantation experiments, they suggested that for an MPN to develop, JAK2V617F must be expressed in the LSK cell population. The expression of JAK2V617F in these mice caused a significant expansion of MEPs and gave a slight competitive advantage in over mice expressing wild type JAK2. The competitive advantage was defined by the LSKs of JAK2V617F knock-in mice having a slight increase in cell cycle progression relative to LSKs expressing the normal JAK2 protein. Mullaly et al. also commented that the haematopoietic stem cell compartment has a “unique” capacity for disease favouring erythroid proliferation/differentiation (Mullally et al., 2010).

In 2010, Li et al. generated mice with knock-in human JAK2V617F cDNA. After 12 weeks the newly generated mice had almost 50% fewer LSKs and the LSKs that remained had significant DNA damage, apoptosis, and reduced cell cycling. This contrasts with the theory that after patients acquire the JAK2V617F mutation, HSCs expressing JAK2V617F might acquire a slight selective advantage that allows clonal growth over time (Li, J. et al., 2010).

Furthermore, in the model generated by Li et al., non-competitive bone marrow transplantation experiments showed a progressive reduction of donor JAK2V617F-expressing cells in recipients, indicating decreased stem cell activity in cells expressing JAK2V617F compared with wild-type controls. Consistent with this, direct comparison of the wild type versus JAK2V617F bone marrow cells in competitive transplantation experiments demonstrated that JAK2V617F confers a mild but significant disadvantage to HSCs, this selective effect was significantly increased upon secondary transplantation (Li, J. et al., 2010).

In 2014, Li et al. reported that expression of either knock-in heterozygous or homozygous human *JAK2V617F* under a mouse *JAK2* promoter in mouse models compromised self-renewal of HSCs preferentially polarising HSCs towards differentiation away from self-renewal triggering an exhaustive effect in HSCs. Consistent with this, homozygous *JAK2V617F* had a much more potent effect on self-renewal than heterozygous expression. Clones derived from the *JAK2V617F*-homozygous CD45+EPCR+CD150+CD48- HSCs were reported to enter the cell cycle at a much higher rate and in the short term had a much larger proliferation advantage relative to the wild type but had a reduced composition of stem cells suggesting that JAK2V617F sacrifices self-renewal to proliferate and differentiate terminally which in turn leads to exhaustion (Li, J. et al., 2014).

The mouse models above demonstrate contrasting effects of JAK2V617F on HSCs where the Akada et al. mouse model is different to that of Li and Mullaly. The reasons for these differences are currently unclear but suggest that JAK2V617F expression generates highly heterogenous MPNs.

Several studies have directly focused on the patients with the JAK2V617F mutation. Jamieson et al. isolated whole blood samples from PV patients and noted that compared to healthy donors, PV patients had a higher number of HSPCs and common myeloid progenitors (CMPs) with a skew towards the erythroid lineage (Jamieson et al., 2006). In contrast to this report, Anand et al. and James et al. separately reported no differences in the HSPC or the CMP compartment of MPN patients. However, Anand et al. did observe that the erythroid compartment ($CD34^{-}CD71^{+}GPA^{+}$) in these patients was enriched (Anand et al., 2011; James et al., 2008).

In 2007, Ishii et al. performed xenograft assays in mice using primary $CD34^{+}$ cells from a PV patient. It was concluded from the xenograft assays that samples from PV patients could reconstitute the haematopoietic system, but only if the patients exhibited a high JAK2V617F burden (Ishii et al., 2007).

The mouse models and patient data exhibit variable phenotypes between reports. A possible explanation for this variability is the quantitative difference between low-burden and high-burden JAK2V617F HSC. A second possibility is that high-burden JAK2V617F cells contain higher numbers of authentic long-term HSCs, while $CD34^{+}$ cell fractions from patients with low JAK2V617F burden contain fewer HSCs that engraft in immuno-deficient mice. The level of JAK2V617F may determine whether or not JAK2V617F-positive HSCs have an engraftment and repopulating advantage compared with healthy HSC.

Finally, the level of STAT5 expression and STAT5 phosphorylation in the model should be considered. STAT5 phosphorylation has been reported to drive self-

renewal through expression of hypoxia-Induced Factor 2α (HIF2α) (Fatrai et al., 2011). In 2008, Wierenga et al. investigated STAT5a by inducing STAT5a activation through a dose-dependent activation of a 4-hydroxytamoxifen(OHT)-inducible STAT5a(1*6)-ER fusion protein. The increasing gradient of 4-OHT concentrations suggests that different levels of STAT5a activation can trigger different biological effects. High levels of STAT5a activation can trigger differentiation, whilst low levels of STAT5a activation can trigger self-renewal. The STAT5a protein can also regulate the expression of *C/EBPα*. The C/EBPα protein functions by negatively regulating phospho-STAT5 signalling and is significantly downregulated in MPN patients expressing JAK2V617F. This suggests that higher expression of STAT5a will promote JAK2-STAT5 signalling which, in turn, promotes a more severe MPN phenotype (Wierenga et al., 2008).

In conclusion, the phenotypes obtained by the different JAK2V617F expression studies are highly variable and it is clear that there is more work to be done to understand the heterogenous phenotypes caused by JAK2V617F.

1.5.3 JAK2 exon 12 mutations

Of the 10% of PV patients that do not express JAK2V617F, it was postulated that mutations might be acquired in either other JAK family members or elsewhere in the *JAK2* gene.

JAK2V617F-negative PV patient granulocyte DNA was sequenced to explore all of the 25 JAK2 exons for mutations. This revealed a sequence change in exon 12, CAA to ATT in positions 1614 through 1616 which resulted in H538Q and K539L mutations (Scott, Linda M. et al., 2007).

Sequence analysis of granulocyte DNA from patients with JAK2V617F-positive PV, JAK2V617F-negative ET and JAK2V617F-negative cases of idiopathic myelofibrosis showed no evidence of JAK2 exon 12 mutations (Williams et al., 2007; Scott, Linda M. et al., 2007; Pardanani, A. et al., 2007; Colaizzo et al., 2007).

A separate group sequenced JAK2 exons in a different JAK2V617F-negative patient; this yielded a 6 bp in frame deletion at positions 1611 to 1616 (also in exon 12), resulting in an F537K539delinsL mutation.

Despite the heterogeneity with respect to sequence changes in these cases, both variants targeted H538 and K539 residues. Analysis of eight additional JAK2V617F-negative PV patients subsequently revealed four additional exon 12 allele variants that all presented with changes affecting conserved residues between K537 and E543 whilst three of the alleles identified contained a K539L substitution. The mechanism of action for these mutations is thought to be similar to that of JAK2V617F, specifically via de-repression of the pseudokinase inhibitory activity.

To confirm that the exon 12 mutations could drive MPN development, functional studies were performed. Bone marrow transplantation of mouse bone marrow cells expressing JAK2K539L resulted in mice which exhibited high levels of haemoglobin, reticulocytosis and leukocytosis (Scott, Linda M. et al., 2007), and transgenic mice with JAK2-N542-E543del (the most frequent exon 12 variant) also presented with elevated haemoglobin and reduced overall survival but normal platelet and neutrophil counts (Grisouard et al., 2016). Molecular characterisation of these mice demonstrated increased signalling of phospho-STAT3 and phospho-ERK1/2 pathway (Grisouard et al., 2016; Godfrey, Anna L. et al., 2016).

These reports predicted that mutations in exon 12 function in a manner similar to that of the mutant *JAK2V617F*. In fact, in 2007 Scott et al. commented that, in the fully folded JAK2 protein, the SH2-domain is close to the loop of the JH2 pseudokinase domain where the *JAK2V617F* mutation takes place. This suggests that perhaps the exon 12 mutations are disrupting the JH2-JH1 interaction and emulate that of the JAK2V617F mutation.

1.5.4 Thrombopoietin receptor (MPL) mutations

MPL and its ligand, TPO, are essential for megakaryopoiesis. MPL is comprised of three protein domains, the extracellular domain (ECD), the transmembrane domain (TM) and the cytoplasmic domain (CD) (**Figure 1.4**). In 2006, exon 10 mutations of the gene encoding *MPL* were reported in non-JAK2 mutated patients with ET and MF at a frequency of 3-5% and 8-11%, respectively (Pikman et al., 2006; Pardanani, A.D. et al., 2006). This mutation is located in the juxta-membrane region of MPL with the most frequent mutations being substitutions in codon 515 leading to tryptophan-to-leucine (W515L) and tryptophan-to-lysine (W515K) in the CD (Beer et al., 2008).

Functional studies confirmed that *MPL* variants are true MPN driver mutations. MPLW515L-mutant mice exhibit disease phenotype similar to ET and MF (Li, J. et al., 2011; Pecquet et al., 2010), but this mutation does not cause PV-like phenotypes as seen in JAK2V617F murine models. These differences however accord with the fact that MPL mutations are only present in ET and MF patients and may suggest that the MPLW515L exerts is primary effects on megakaryocytic lineages *in vivo*.

The ECD of the MPL receptor is composed of two copies of cytokine receptor modules (CRM), each of which is composed of a fibronectin III-like (FNIII) domain characterised by four conserved cysteine residues, a hinge region and a WSXWS motif. The membrane distal CRM (CRM-1) binds TPO, as its deletion creates a mutant receptor that cannot bind TPO (Sabath et al., 1999; Deane et al., 1997).

The TM domain of MPL is responsible for anchoring the receptor to the cell surface. This domain also plays critical roles in governing receptor activation, such as the regulation of receptor dimerisation, and allowing 3D structural changes to occur which allow for signal transduction (Leroy et al., 2016). The TM domain can regulate MPL activation through a 5-residue amphipathic helical motif RWQFP neighbouring the C-terminal tail of the TM domain. This motif is essential for restricting receptor activity by maintaining the structural tilt of MPL whilst anchored to the lipid bilayer and this keeps the receptor inactive in the absence of TPO stimulation (Varghese et al., 2017). Deletion of the motif results in chronic activation of MPL (Staerk et al., 2006) and the MPLW515 mutations disrupt this motif (Defour et al., 2013).

The CD of MPL is the region where signalling intermediates such as JAK2 tyrosine kinases are constitutively bound to act as signal transducers, or where other intracellular proteins such as STATs can be recruited to be phosphorylated. The intracellular domain of MPL share two conserved elements: Box1 and Box2. Box1 contains a PxxP motif and is required for JAK2 binding, while the function of Box2 remains unclear. In addition, there are three important cytoplasmic tyrosine residues (Tyr591, Tyr626 and Tyr631) located in the intracellular domain that when phosphorylated, can act as docking sites for the SH2-domain-containing proteins, primarily STATs, MAPK, PI3K and protein

kinase C (PKC) to promote cell proliferation and cell survival (Rojnuckarin et al., 1999; Drachman et al., 1995).

The disruption of the protein structure of MPL leads to dysfunctional signalling, in MPNs, the most frequently mutated residue in MPL is W515. The presence of the W515 mutation disrupts the RWQFP motif promoting chronic signalling of the receptor. The MPLW515 mutant receptor no longer requires TPO to bring the cytoplasmic domains of the receptor together. The consequence of this mutant receptor structure, implies that the MPL-bound JAK2 molecules are in a chronic, close proximity to one another and, in effect, are triggering an increase in JAK2 catalysed signalling. The signalling output of the MPLW515 mutation is similar to that of the JAK2V617F mutation, however the MPLW515 mutation is polarised towards an ET and MF phenotype rather than PV phenotype due to the role of MPL in megakaryopoiesis.

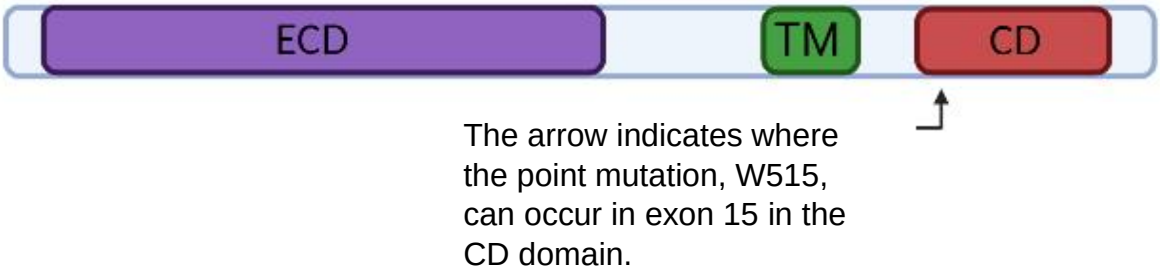


Figure 1.4 A protein schematic of the *MPL* gene.

The Figure illustrates a protein schematic of the *MPL* gene. The figure highlights the three main domains: the ECD, the TM and the CD. In exon 15 of the CD, tryptophan substitution mutations at position 515 can occur. The output of these mutations is constitutively active MPL signalling, this translates into either ET or MF.

Created with Biorender.com

1.5.5 Calreticulin (CALR) mutations

In 2013, two groups reported calreticulin (CALR) mutations in ~25-30% of ET patients and represents the second most common genetic mutation in MPN patients after JAK2V617F (Nangalia et al., 2013; Klampfl et al., 2013). CALR mutations reside in exon 9 and are insertion/deletion (indel) mutations that give rise to a frameshift that results in truncation of the C-terminal tail of CALR protein. Of the mutations identified, type I mutations encoded a 52-bp deletion (CALRdel52), and type II mutations encoded a 5-bp insertion (K385fs*47) are the most frequent although others do exist (Nangalia et al., 2013).
CALR mutations in MPN are indel mutations that introduce a frameshift in the CALR reading frame that leads to the creation of a novel, mutant C-terminal tail that lacks the amino acid motif: lysine, aspartic acid, glutamic acid, and leucine (KDEL). The loss of the endoplasmic reticulum (ER) retention motif KDEL results in the inability for CALR mutants to re-enter the ER and protein mislocalisation.

To investigate CALRdel52, multiple groups have utilised cytokine-dependent cell lines, specifically the Ba/F3 cell line. The Ba/F3 cell line is dependent on IL-3 for proliferation in culture, but Ba/F3 cells can also be re-engineered to rely on MPL/EPOR/G-CSFR and JAK2 signalling. In 2016, Chachoua et al. reported that the expression CALRdel52 alone in Ba/F3 cells was not enough to confer cytokine-independent proliferation in culture. It was actually found that for CALRdel52 to confer IL-3 independence co-expression of MPL is required. This indicated cooperation between these two proteins was necessary for IL-3 independence and proliferation in culture (Elf et al., 2016; Chachoua et al., 2016; Araki et al., 2016). The other type I receptors such as EPOR and G-

CSFR could not instigate MPN development through CALRdel52 co-expression (Elf et al., 2016; Araki et al., 2016).

Next, attempts were made to understand the role of the novel mutant-specific C-terminus of CALRdel52 in its transforming capacity. CALR mutagenesis assays were performed that generated a range of mutated forms of CALRdel52 (Elf et al., 2016). From the mutagenesis assays, it was shown that the complete truncation of the mutant-specific C-terminus resulted in an inability to confer cytokine-independence. Therefore, the C-terminus of CALRdel52 appears essential for CALRdel52 transforming activity. The removal of 8-10 successive amino acid blocks of the mutant C-terminus failed to abrogate cytokine independent growth in any mutant analysed, suggesting that the transforming ability of CALRdel52 is not sequence specific. Finally, mutating all of the 18 basic K/R amino acid residues within the C-terminal tail to glycine abrogated CALRdel52 oncogenic activity. The mutation of all non-basic K/R residues did not affect oncogenic activity. From the data it was concluded from the experiments that CALRdel52 activity was caused by the C-terminus's positive electrostatic charge (Elf et al., 2016). Furthermore, in 2018 it was reported that three specific tyrosine residues within the MPL receptor are crucial for CALRdel52-MPL activity (Elf et al., 2018).

In 2021, Rivera et al. performed site-directed mutagenesis on CALRdel52 transforming every residue within the lectin domain into alanine. This identified three zinc-binding histidine residues within the CALRdel52 N-terminal zinc-binding domain as crucial for malignant transformation (H99, H145 and H170). H99, H145 and H170 regulate CALRdel52 activity through the formation of multimers of CALRdel52 (Rivera et al., 2021). It was further shown that zinc

chelators are effective abrogators of the CALRdel52 phenotype through the disruption of the multimer formation.

The mutated CALR proteins can form multimers and bind MPL, by binding MPL, it is suggested that CALR mutant proteins can emulate cytokine binding by TPO. The binding of MPL by CALR mutants stimulates the JAK2 signalling pathway which is similar to that of MPLW515 and JAK2V617F. CALR mutants chronically activate differentiation and proliferation through increased JAK2 signalling. Notably, CALR mutations, like the MPLW515 mutation, are only found in ET and MF and this is likely due to the role of MPL in megakaryopoiesis.

In essence, the mutations in JAK2, MPL and CALR, all chronically activate similar pathways, this is because the signal transducing molecule is usually the same, JAK2 kinase.

1.6 Overview of the Thesis

JAK2V617F is the most prevalent mutation found in MPN patients occurring in ET, PV, and MF. The current main treatment of JAK2V617F-driven MPNs involves the use of Ruxolitinib, a JAK1 and JAK2 inhibitor. The use of Ruxolitinib has proven to be effective but patients with JAK2V617F quickly gain resistance to the drug and cell lines expressing JAK2 show signs of resistance vicariously through JAK1 signalling after just six hours (Koppikar et al., 2012). The discovery of new targets/pathways for targeted inhibitors is crucial for developing an arsenal with which to either treat the MPN or slow down the growth/development of MPN.

To investigate new target genes of JAK2V617F signalling, this thesis aims to:

1) Develop and test a model for JAK2V617F-driven MPNs in a haematopoietic stem cell model.

2) Interrogate the RNA-sequencing dataset of this model to look for dysregulated genes.

3) Investigate the identified gene from 2) and determine how it contributes to JAK2V617F MPNs.

Chapter 3 will discuss the creation of the disease model, a HPC7-JAK2V617F cell line, and experiments where I identified dysregulated phospho-STAT/ERK/AKT signalling, increased DNA damage and increased DNA damage repair. This chapter describes the experiments characterising HPC7-JAK2V617F cells and testing them under exposure to UV and Mitomycin C, identifying JAK2V617F expression as a mutation that contributes to both DNA damage itself and stabilising the genome against further DNA damage.

Chapter 4 will discuss RNA-sequencing data acquired from the Next-Generation Sequencing of HPC7-JAK2V617F. This chapter describes the genes upregulated and the biological validation of 40 of these genes. This chapter also discusses and compares the results gained to those of other groups to gauge how close my dataset is to patient data and mouse data. Finally, this chapter discusses HDAC9 as a potential gene target and its dysregulation in MPNs.

Chapter 5 will discuss HDAC9 and its contribution to MPNs. Chapter 5 describes the effect of HDAC9 depletion from the HPC7-JAK2V617F cell line and identifies HDAC9 as being crucial for the phosphorylation of ERK, AKT and STAT3. This chapter also presents evidence that HDAC9 contributes to DNA damage and DNA damage repair. This chapter identifies the MEF-2 binding domain, the HDAC glutamine-rich N-terminal domain, the CTBP binding domain and the disordered C-terminal as being necessary for function.

Finally, Chapter 6 will summarise the thesis discussing the results briefly and future work that should be undertaken to fully understand HDAC9 in JAK2V617F-driven MPNs.

Chapter 2
Materials and methods

Chapter 2

Materials and methods

2.1 Cell lines and tissue culture

HEK293T human embryonic kidney tissue cells were obtained from the American Type Culture Collection (ATCC) (Pear et al., 1993; DuBridge et al., 1987) HEK293T cells were maintained in Dulbecco's Modified Eagle's Medium (DMEM) (Lonza, Catalog #: BE12-726F) supplemented with 10% Foetal Bovine Serum (FBS), 1% L-glutamine and 1% penicillin/streptomycin. HEK293T cells were sub-cultured every 2-3 days. Briefly, cells were rinsed in 1X phosphate-buffered saline (PBS), 1 mL of trypsin-EDTA solution was added and incubated for approximately 5 minutes to allow cell detachment and seeded in fresh complete growth medium at a sub cultivation ratio of 1:10. Cells were grown at 37°C in 5% CO_2.

HPC7 cells are haematopoietic precursor cells that are murine stem cell factor (MSCF) dependent and were a generous gift from Doctor Edwin Chen (University of Leeds). HPC7 cells were maintained in Iscove's Modified Dulbecco's Medium (IMDM) (Lonza) supplemented with 10% FBS, 1% L-glutamine and 1% penicillin/streptomycin and 50 µg/ml of recombinant MSCF (PeproTech). HPC7 cells were sub-cultured at a ratio of 1:7 every 2-3 days. HPC7 cells were grown in T-25 vented flasks at 37°C in 5% CO_2.

HPC7-JAK2WT and HPC7-JAK2V617F are derivatives of HPC7 cells and were generated in the lab by lentiviral spin infection of HPC7 cells with lentiviruses expressing human *JAK2WT* cDNA and *JAK2V617F* cDNA. Following infection,

cells were cultured with the addition of Geneticin (350 ng/ml) to select the *JAK2WT* or *JAK2V617F* expressing cells.

2.2 Freezing and thawing cells

Cells were stored for long-term storage at -80^{o}C. Confluent cells were centrifuged at 180 x g for 5 minutes and the cell pellet was resuspended in 1 mL of freezing media. Freezing media for 293T cells are made up of 70% medium used for culturing the cell type, 20% FBS and 10% DMSO. The freezing media for HPC7 cells consists of 70% IMDM media, 20% FBS and 10% DMSO. The cells are then transferred into cryovials and placed into Mr. Frosty freezing container (ThermoFisher Scientific) to allow a rate of cooling of -1°C/minute, the optimal rate for cell preservation. For cell thawing, the cryovials were thawed in gel bead bath at 37°C and were suspended in pre-warmed media. Cells were centrifuged at 180 x g for 5 minutes, and the cell pellet was resuspended with new media and transferred into a T-25 flask and allowed to recover in the incubator.

2.3 RNA extraction, RT-PCR and RT-qPCR

Total RNA is extracted using RNeasy® Mini Kit (Qiagen®, #74106). To produce complementary DNA (cDNA) iScript™ cDNA Synthesis Kit (Bio-Rad) was used according to the manufacturer's instructions.

PCR was carried out using the DreamTaq Green PCR Master mix (X2) (Thermo Scientific™, catalogue number: K1071) according to the manufacturer's recommendation. The reaction mixture consisted of 100 ng of cDNA, 0.5 µM of

forward primers and 0.5 µM reverse primers with the rest being made up with water and the PCR master mix (X2).

The PCRs were run using two thermocyclers: PCRmax Thermocycler (PCRmax) and the T100™ Thermal Cycler (BIO-RAD). After amplification 10 µl of PCR product was run on a 2% Agarose gel (made using 1X TAE buffer) and exposed to the BioRAD Universal Hood II ChemiDoc™.

The RT-qPCR was performed using the SsoFast™ EvaGreen® Supermix (BIO-RAD) according to the manufacturer's instructions using 100 ng of cDNA and 0.5 µM of both forward and reverse primers. The qPCR was performed on a CFX Connect™ Real-Time System (BIO-RAD).

2.4 Genomic DNA extraction and PCR Purification of JAK2WT and JAK2V617F

Genomic DNA was extracted using a Purelink Genomic DNA Mini Kit (Invitrogen, catalogue number: K1820-01) according to the manufacturer's instructions.

For purification of JAK2WT and JAK2V617F amplicon products, the QIAquick® PCR Purification kit (QIAGEN, catalogue number: 28104) was used, following the manufacturer's instructions. After purification, samples were sent to GeneWiz® for sequencing.

2.5 Bacterial transformation

TOP10 chemically competent *E.coli* (ThermoFisher Scientific) were thawed from the -80ºC and 25 µl was added to a clean microcentrifuge tube. 2 µl of the appropriate plasmid was added to the *E.coli* and mixed by gentle flicking. The mixture was incubated on ice for 30 minutes. After 30 minutes cells were incubated in a water bath, heated to 42ºC, for 1 minute and then quickly replaced on ice for 2 minutes. 400 µL of SOC media (ThermoFisher Scientific catalogue number: 15544034) was added to the mixture and the tube was incubated at 37°C on a shaker for 1 hour. The mixture was then plated onto agar plates containing ampicillin (50 µg/ml). The next day colonies were picked from the plates and inoculated into in-house Lysogeny Broth containing 50 µg/ml of ampicillin and incubated overnight on a shaker at 37°C.

The next day the broth was taken, and the DNA was extracted using QIAprep® Spin Miniprep Kit (Qiagen, Cat No. 27106) according to the manufacturer's instructions.

Colonies were plucked and sent for sequencing, using the appropriate primer for the promoter encoded in the plasmid, to identify clones and individually re-plated as before.

2.6 Lentiviral infection of HPC7 cell line

Lentivirus generation was performed using HEK293T cells. HEK293T cells were seeded in 6 well dishes at a density of 1.5 x 10^5 cells/mL. After 24 hours, each well was transfected with a total of 2.5 µg plasmid DNA, which includes 1.2 µg of Vector encoding the gene/shRNA of interest, 0.65 µg of pCMV-pVSV-G

(pCMV-pVSV-G was a gift from Akitsu Hotta (Addgene plasmid # 138479; http://n2t.net/addgene:138479; RRID:Addgene_138479)) and 0.65 µg of psPAX2 (a gift from Didier Trono (Addgene plasmid # 12260; http://n2t.net/addgene:12260; RRID:Addgene_12260)) and mixed with 7.5 µL TransIT-LT1 transfection reagent (Mirus Bio). The DNA:cationic lipid solution was incubated for 15 minutes at room temperature and then added to HEK293T cells in a dropwise manner. Cells were incubated at 37^{o}C for 5 hours and then supplemented with fresh media. Twenty-four hours post-transfection, media from the HEK293T cells containing viral supernatants was collected, passed through a 0.22 µm filter and used to transduce HPC7-JAK2WT or HPC7-JAK2V617F cells. For each transduction, 1.5 x10^{5} cells HPC7 JAK2WT or HPC7-JAK2V617F cells were added into a well of a 6-well dish, combined with 1 mL of the viral supernatant, supplemented with 8 ng/mL polybrene and a spin-infection performed at 720 x g with slow acceleration and deceleration for 2 hours. After spin-infection, viral supernatant was removed, and cells were grown in fresh media supplemented with appropriate cytokines for 24-48 hours prior to being selected with puromycin (2 µg/ml).

2.7 Cell counting

Cell proliferation of HPC7-JAK2WT or HPC7JAK2V617F cells following MSCF withdrawal or MSCF concentration gradient were performed using Trypan Blue dye exclusion by mixing cells with a 0.4% Trypan Blue solution (Sigma Aldrich) at a 1:1 ratio and cell numbers were counted under an inverted phase contrast microscope using a haemocytometer.

2.8 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) assays

Cells were starved of MSCF or plated on a MSCF concentration gradient (0-50 µg/ml), and rates of proliferation were measured using an MTT assay. The MTT assay was conducted as follows: HPC7 cells were centrifuged at 180 x g for 5 minutes and washed three times with 1X PBS to remove traces of MSCF. The HPC7 cells were then plated into a 96-well plate and left overnight. To analyse the cells, the 96-well plates were centrifuged at 180 x g, and the media was carefully aspirated. After the media was aspirated, 100 µl of 1 mg/ml MTT diluted in phenol red-free IMDM was added to each individual well and incubated for 3 hours at 37ºC in the CO_2 incubator. Post-incubation the cells were centrifuged at 180 x g for 4 minutes to remove the media/MTT solution. After the solution was aspirated, the remaining MTT crystals were resuspended in DMSO by pipetting up and down. The absorbance was then measured using a Hidex sense plate reader at 570 nm.

2.9 Western blots.

Cells were harvested and washed with 1x PBS and lysed with 100 µL of radioimmunoprecipitation assay (RIPA+) buffer (150 mM sodium chloride, 1.0% NP-40, 0.5% sodium deoxycholate, 0.1% sodium dodecyl sulfate (SDS) and 50 mM Tris, pH 8.0). Samples were incubated on ice for 30 minutes and centrifuged at 11200 x g for 10 minutes and supernatants were then collected and transferred to a new microcentrifuge tube. The protein concentration was determined using a Bradford assay with BSA standards at known protein

concentration ranging from 1-4 mg/mL. The samples were measured in triplicate and absorbance was measured using an iMark™ spectrophotometer plate reader (Bio-Rad) at 595 nm. Prior to loading the SDS-gel, an equal volume of 2X SDS Laemmli buffer (Bio-Rad) was added to the protein lysates. The samples were boiled at 95^{o}C for 10 minutes and stored in -20^{o}C until required. Protein samples were subjected to SDS-PAGE on a 10% polyacrylamide gel for 90 minutes at 100 V or until the marker dyes reach the end of the gel. Proteins were then transferred to nitrocellulose membrane using the Trans-blot Turbo transfer system (Bio-Rad) for 30 minutes.

The membranes were stained with 1X Ponceau S to ensure efficient protein transfer. Membranes were blocked using 5% skimmed milk in 1x PBST (PBS supplemented with 0.05% Tween-20) at room temperature for 1 hour. The membranes were incubated with appropriate antibodies overnight at 4°C (antibody dilutions indicated in Table 2). The membranes were washed three times with PBST, then incubated with the appropriate secondary antibodies in PBST (1:5000 dilution) for 1 hour at room temperature, washed three times PBST, and blots were visualised using SuperSignal West Pico ECL Plus Chemiluminescent substrate kit (Thermo Fisher Scientific) according to the manufacturer's instructions. The blots were exposed using Universal Hood II ChemiDoc (Bio-Rad). Analysis of the blots were performed using Image Lab software (Schneider et al., 2012). The dilutions of antibodies are found in Table 2.

2.10 Annexin V and propidium iodide (PI) staining

500,000 cells were washed with FACS staining buffer 1 (PBS containing 2% FCS), followed by Annexin V staining buffer. The cell pellet was re-suspended in Annexin V staining buffer (BioLegend® catalogue number: 422201) before being incubated on ice and in the dark with propidium iodide (PI) and an anti-Annexin V antibody conjugated to Allophycocyanin (APC) (BioLegend® catalogue number: 640941) for 15 minutes. The cells were then analysed by flow cytometry using the CytoFlex V0-B3-R1 flow cytometer (Beckman Coulter, Product No: B53015). The procedure for gating was as follows: the events detected by the flow cytometer were plotted onto a dot plot with the forward scatter (FSC) plotted on the x-axis and side scatter (SSC) plotted on the y-axis. The FSC-SSC dot plot was used to define events representing living cells and remove events representing cellular debris. A gate was drawn around the events representing living cells and the gated population was plotted onto a second dot plot with Annexin V- APC on the x-axis and PI on the y-axis. Data analysis was performed with CytExpert V2.2.

2.11 Cell cycle analysis

For the cell cycle analysis, cells were harvested from every sample and washed in FACS staining buffer 2 (PBS containing 2% BSA). The cells were permeabilised and fixed by re-suspending the cell pellet in 70% ethanol for 30 minutes. Cell pellets were washed further with FACS staining buffer 2 and re-suspended in the same buffer supplemented RNAse A (10 mg/ml) from

BioLegend® for 15 minutes to degrade RNA. After the RNAse A treatment, cells were washed once more with FACS staining buffer 2 and then resuspended in FACS staining buffer 2 supplemented with 0.5 mg/ml PI (BioLegend®, catalogue number: 421301). After 5 minutes, the cells were analysed by flow cytometry using the CytoFlex V0-B3-R1 flow cytometer (Beckman Coulter Life Sciences, Product No: B53015). The procedure for gating was as follows: the events detected by the flow cytometer were plotted onto a dot plot with the forward scatter (FSC) plotted on the x-axis and side scatter (SSC) plotted on the y-axis. A gate was drawn around the events representing living cells and the gated population was plotted onto a histogram with PI on the x-axis and “count” on the y-axis. Data analysis was performed with CytExpert V2.2.

2.12 Immunoprecipitation (IP) and IP-mass spectroscopy

HPC7-WT/JAK2V617F cells were harvested and washed with 1X PBS, pelleted, and resuspended in 450 µL NP40 lysis buffer (20 mM Tris HCl pH 8, 137 mM NaCl, 2 mM EDTA 1% NP-40 supplemented with Roche cOmplete™ protease inhibitors) and incubated on ice for 30 minutes. The samples were centrifuged at 180 x g for 10 minutes at 4°C, and the supernatants were retained for immunoprecipitation. Separately, 50 µL of Protein G magnetic beads (BioRad) were washed with IP Wash Buffer (PBS with 0.1% Tween-20) three times and combined with the recommended amount of the relevant antibody (Table 2) and incubated for 15 minutes by gentle rotation at room temperature. Bead-antibody mixtures were then washed with IP Wash Buffer and applied to 400 µL lysate supernatants for 2 hours at room temperature with gentle rotation. Beads were washed on the Surebead magnetic bead system (BioRad), and washed beads

were resuspended in 100 µL of 1x Laemmli buffer and eluted by incubating the beads for 10 minutes at 70°C.

The processing of the IP sample was performed by the Mass spectroscopy department, and this is a summary of their protocol provided by their facility. The sample was digested using the mass spectroscopy protocol for S-Trap digest columns. The sample was digested with trypsin and the resulting peptides were detected and fragmented in the mass spectroscopy. Peak's software was used to find which proteins these peptides match to and was searched against all of the mouse entries on Uniprot.

2.13 Intracellular phosphoprotein flow cytometry analysis

HPC7-JAK2WT and HPC7-JAK2V617F cells were transferred into flow cytometry tubes and washed with FACS buffer (1X PBS + 2% FCS) twice to remove residual IMDM culture media and centrifuged at 180 x g for 5 minutes. Fixation and permeabilisation of cells were performed using Fix&Perm Kit (Nordic MUbio, GAS-002-1). The cells were fixed with 100 µL of reagent A (fixation medium) and incubated for 15 minutes at room temperature and washed with 5 mL of FACS Buffer by centrifuging for 5 minutes at 180 x g and discarding supernatant. The fixed cell pellet was resuspended with reagent B (permeabilisation medium) and 5 µL each of phycoerythrin (PE) mouse anti-STAT3 (pY705) and Alexa Fluor 647 anti-STAT5 (pY694) (BD Biosciences Phosphoflow). Staining was carried out for 1 hour at room temperature in the dark. Cells were washed after staining using the FACS Buffer, resuspended in

500 µL of FACS Buffer and analysed using a CytoFLEX flow cytometer (Beckman Coulter). Data analysis was performed with CytExpert V2.2.

2.14 UV treatment of HPC7 cell lines

HPC7 cells were plated at a density of 1.5 x 10^5 cells/ml and exposed to 254 nm UV radiation at 0-100 mJ/cm^2 for 15 minutes using an Analytik Jena crosslinker (legacy model) and left to incubate overnight. The next day cells were counted/ measured by a Trypan Blue exclusion assay or an MTT assay as described in sections 2.7 and 2.8.

2.15 UV treatment of eGFP vectors and eGFP quantification

The eGFP vector was a generous gift from Dr A Timmis. This experiment was taken from (Qiao et al., 2002) and changed to use eGFP expressing vector instead of a luminescence-based vector. The eGFP vector stock was diluted with TE buffer to 50 µg/ml and 2 mL were aliquoted onto a 60-mm culture dish on ice. The plasmid was irradiated at 0-800 J/cm^2 (254 nm) using the Analytik Jena crosslinker (legacy model). Post-UV treatment the aliquots were stored at -80ºC to preserve the damaged plasmids until they were needed for transfection into the desired cell line. To quantify the expression of eGFP, the HiDex plate reader was used.

2.16 Inhibitor treatments

HPC7 cell lines were counted and cultured at densities of 2.5 X 10^5 cells per ml alongside the appropriate drug. The drugs and the conditions in which they were used are listed below.

For JAK inhibition experiments, HPC7-JAK2V617F were treated with 500 nM of Ruxolitinib for 6 hours or 500 nM of Fedratinib for 6 hours.

For DNA damage response experiments, HPC7 cell lines were treated with mitomycin C at concentrations between 1-4 µg/ml for 24 hours. The treated cells were then counted via a Trypan Blue exclusion assay.

For DNA damage response experiments, HPC7 cell lines were treated with Olaparib at concentrations between 0-5 µM for 96 hours. The treated cells were counted via a Trypan Blue exclusion assay.

For STAT inhibition experiments, HPC7-JAK2V617F cells were treated with either: Fludarabine for 24 hours at concentrations of 1.5-4.0 µg/ml, Stattic for 24 hours at a concentration of 20 µM or STAT5 inhibitor for 24 hours at a concentration of 100 µM.

For EGFR inhibitor experiments, the HPC7-JAK2V617F cells with were treated with Erlotinib for 72 hours at 5-10 µM.

2.17 Transient transfection

Transient transfections were accomplished using the protocol recommended in the TransIT®-LT1 (MIR 2304) transfection reagent manual on the Mirus Bio website (https://www.mirusbio.com/). In brief, 4.0 x 10^5 HPC7 cells were plated into 12-well plates and left overnight to achieve maximum confluency. The next

day, a mixture was prepared in a 5 ml polypropylene tube containing: 100 µl of Opti-MEM™ (ThermoFisher), 1 µl of a 1 µg/µl stock of DNA and 3 µl of the TransIT®-LT1 transfection reagent and left to incubate for 20 minutes. Post-incubation, the mixture was added to the well, making sure to swirl the plate afterwards to distribute the complex and left in the incubator for 24 hours before harvesting.

2.18 Site-directed mutagenesis of *HDAC-ORF*

Site-directed mutagenesis was achieved using the protocol provided by the New England BioLabs® manual (https://international.neb.com/). In brief, 5 ng of HDAC-ORF template DNA was combined with 0.5 µM of primers (designed on the NEBasechanger™ on the NEB website (https://nebasechanger.neb.com/)), water and the Q5 hot start high-fidelity 2X master mix provided in the kit. The DNA/mastermix was placed into a thermocycler and amplified under the recommended conditions.

Post-amplification, the amplified mixture was subjected to a KLD (kinase, ligase and DpnI) reaction for 5 minutes to ligate/circularise and remove template DNA. The KLD-treated DNA was then introduced to chemically competent bacteria as described in section 2.5.

2.19 Dot blots

Dot blots were carried out by first isolating genomic DNA from HPC7 cell lines using the PureLink™ genomic DNA mini kit. Post isolation, the concentration of genomic DNA was measured using a NanoDrop™ spectrophotometer and

subsequently diluted with nuclease-free water to make separate stocks of 1 ng/µl, 3 ng/µl and 9 ng/µl. The dilutions of DNA were added to an Amersham Hybond $^{H+}$ -nylon membrane (RPN2250B, Cytiva) which was sandwiched between two components of a slot blot device (a gift from Dr Ron Chen) which through a vacuuming motion suctions the moisture from sample allowing the genomic DNA to dry on the nylon membrane. The genomic DNA samples were then washed three times with Tris-buffered saline with 0.1% tween-20 (TBST), each time allowing the liquid to be suctioned through the dot blot apparatus. After washing the suction device was left on for a further 5 minutes to make sure the membrane had dried.

The dry nylon membrane was taken to a hybridisation oven and baked for 2 hours at 65ºC to link the DNA to the nylon membrane.

Post-hybridisation, the membrane was blocked with 5% skimmed milk powder dissolved in TBST for 45 minutes before being stained with an anti-DNA:RNA hybrid [S9.6] antibody at a 1:1000 dilution overnight. The next day the anti-DNA:RNA hybrid [S9.6] antibody was washed off and the membrane was washed three times with TBST for 5 minutes each and stained with a secondary mouse antibody in TBST at a 1:5000 dilution for 1 hour and then washed again before being covered with ECL and visualised on the ChemiDoc.

After visualisation the membrane was stained with 0.05% methylene blue diluted in methanol and left to stain on a shaker for 30 minutes. After staining the methylene blue was washed away and the membrane was again washed three times with TBST.

2.20 Subcellular fractionation

To determine the cellular location of HDAC9 subcellular fractionation was performed using the protocol and the buffer recipes found on the Abcam website (https://www.abcam.com/protocols/subcellular-fractionation-protocol). A pellet containing 1.0 X 10^6 of either HPC7-JAK2WT or HPC7-JAK2V617F cells was resuspended in fractionation buffer (10mM KCl, 2 mM $MgCl_2$, 1mM EDTA, 1 mM EGTA and 20 mM HEPES, pH 7.4), before use, 1mM of DTT and a Roche cOmplete™ protease inhibitor tablet was added to the buffer. Cells suspended in the fractionation buffer were then passed through a syringe to lyse them, releasing their protein lysates, which were left on ice for 20 minutes. The sample was then centrifuged for 5 minutes at 1610 x g. The supernatant (containing the cytoplasm) was transferred to a different tube and kept on ice whilst the pellet was washed with fractionation buffer, passed through a syringe five more times and centrifuged again at the 1610 x g but for 10 minutes. Post-centrifugation, the supernatant was discarded, and the pellet was resuspended in TBS with 1% Tween and gently sonicated. Afterwards, 2X Laemmli buffer was added to the suspended pellet and cytoplasmic fractions at a 1:1 ratio and a western blot was run with the samples.

2.21 RNA-sequencing, bioinformatics, and gene set enrichment analysis (GSEA)

Total RNA was extracted as described previously in section 2.3 and sent to Novogene for sequencing on the HiSeq-PE150 platform.

RNA-Seq analysis was performed using my local graphics processing unit (GPU) machine to access my high-performance computing arc3 station account bs12d3l@arc3.ac.uk logged in to MobaXterm UNIX platform program and the workflow is displayed on page 54. The blue text represents the code and X represents either HPC7-JAK2WT or HPC7-JAK2V617F data.

To convert the ENSEMBL gene ID prefixes into Entrez gene IDs the online platform BioMart was used (Howe et al., 2021).

For comparison with online data, I utilised datasets on the NCBI Gene Expression Omnibus (GEO) website (Barrett et al., 2013; Edgar and Lash, 2002). The following datasets were used: GSE 120595 (Celik et al., 2018), GSE 103237 (Zini et al., 2017), GSE55976 (Čokić et al., 2015), GSE79198 (Yang, Y. et al., 2016), and GSE 62302 (Kameda et al., 2015).

The BFU-E dataset was provided by Dr Edwin Chen from his previous patient dataset which can be accessed on ArrayExpress under the code E-MTAB-384 (Chen, E. et al., 2010).

The DNMT3a dataset was acquired from the supplemental figures of the paper by Jacquelin et al. that describes JAK2V617F and DNMT3a loss cooperate to induce myelofibrosis through activated enhancer-driven inflammation (Jacquelin et al., 2018).

Program citations:

FastQC (Andrews, 2010)

TrimGalore (Kreuger, 2021)

STAR (Dobin et al., 2013)

Deseq2 (Love et al., 2014)

R (Team, 2013)

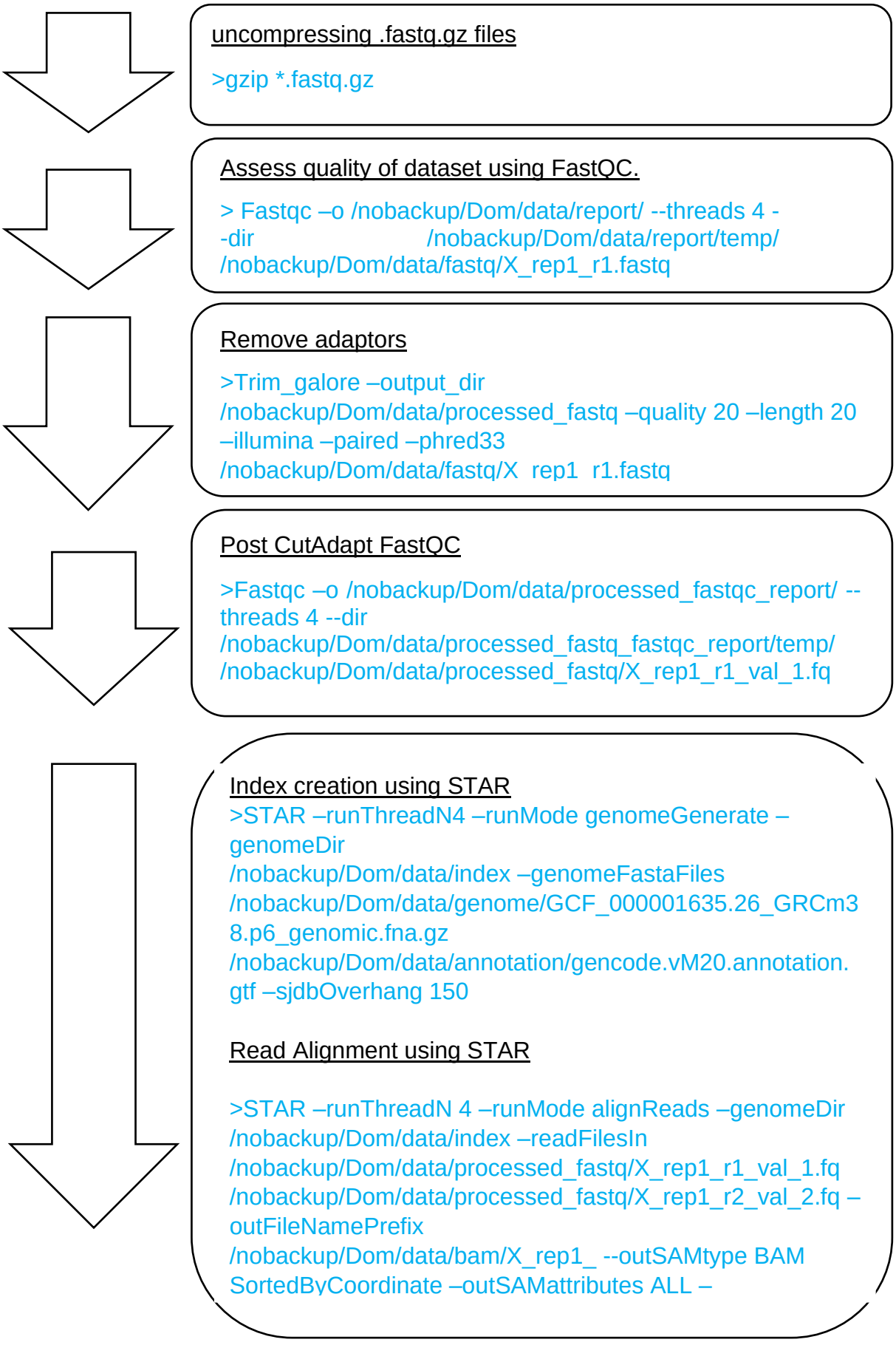
NOISeq (Tarazona et al., 2015)

GSEA (Liberzon et al., 2011; Subramanian et al., 2005)

KEGG (Kanehisa et al., 2021; Kanehisa, 2019; Kanehisa and Goto, 2000)

Reactome (Jassal et al., 2020)

All non-results figures were created with Biorender.com


uncompressing .fastq.gz files
>gzip *.fastq.gz
Assess quality of dataset using FastQC.
> Fastqc –o /nobackup/Dom/data/report/ --threads 4 --dir /nobackup/Dom/data/report/temp/ /nobackup/Dom/data/fastq/X_rep1_r1.fastq
Remove adaptors
>Trim_galore –output_dir /nobackup/Dom/data/processed_fastq –quality 20 –length 20 –illumina –paired –phred33 /nobackup/Dom/data/fastq/X_rep1_r1.fastq
Post CutAdapt FastQC
>Fastqc –o /nobackup/Dom/data/processed_fastqc_report/ --threads 4 --dir /nobackup/Dom/data/processed_fastq_fastqc_report/temp/ /nobackup/Dom/data/processed_fastq/X_rep1_r1_val_1.fq
Index creation using STAR
>STAR –runThreadN4 –runMode genomeGenerate –genomeDir /nobackup/Dom/data/index –genomeFastaFiles /nobackup/Dom/data/genome/GCF_000001635.26_GRCm38.p6_genomic.fna.gz /nobackup/Dom/data/annotation/gencode.vM20.annotation.gtf –sjdbOverhang 150
Read Alignment using STAR
>STAR –runThreadN 4 –runMode alignReads –genomeDir /nobackup/Dom/data/index –readFilesIn /nobackup/Dom/data/processed_fastq/X_rep1_r1_val_1.fq /nobackup/Dom/data/processed_fastq/X_rep1_r2_val_2.fq –outFileNamePrefix /nobackup/Dom/data/bam/X_rep1_ --outSAMtype BAM SortedByCoordinate –outSAMattributes ALL –

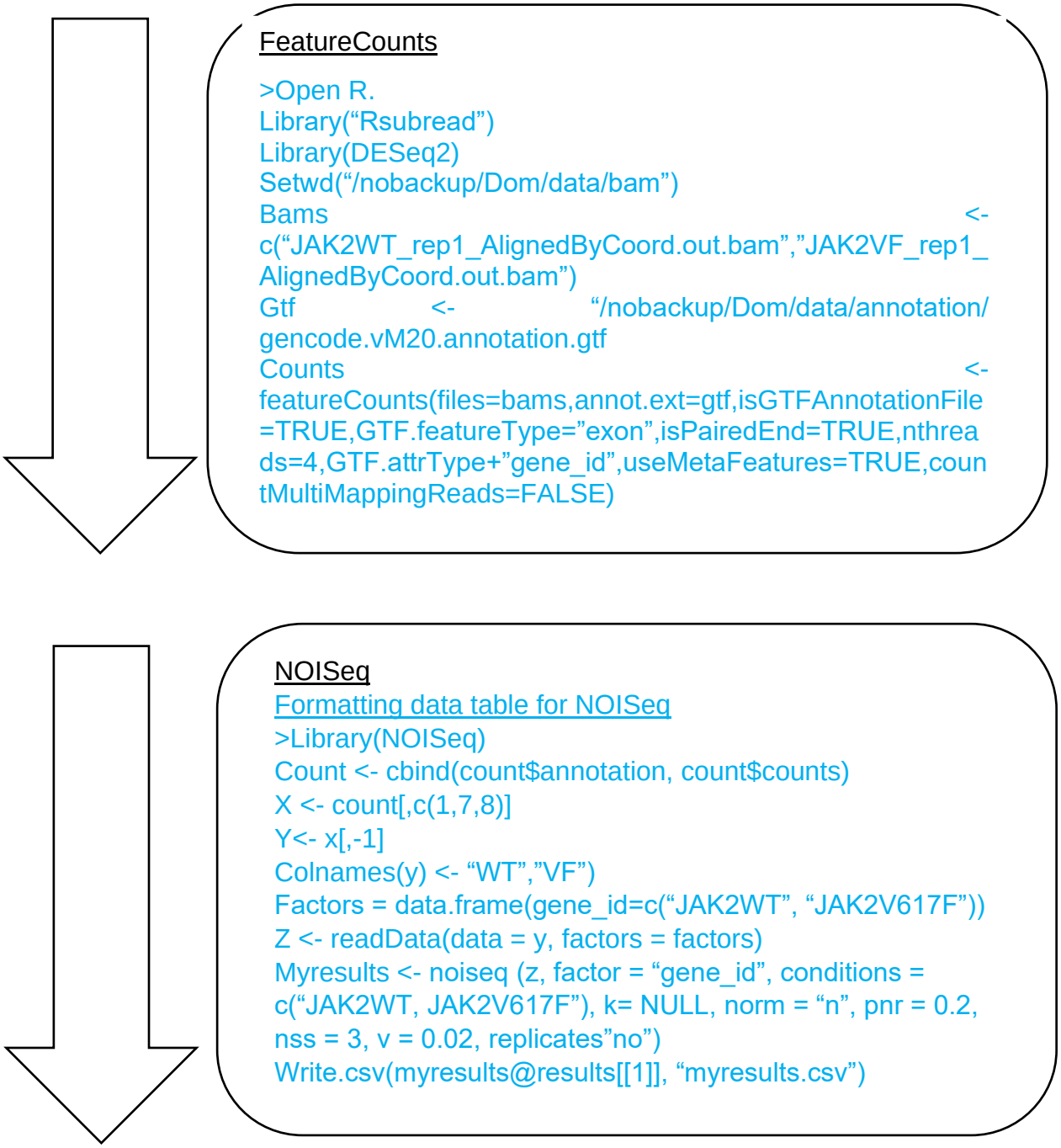


Figure 2.1 My bioinformatic workflow

A bioinformatic workflow detailing the code used to generate the output of the raw data seen in this thesis.

2.22 Inventory of antibodies, vectors, RNAi, and inhibitors

Included in Table 2 are the details of all the antibodies used in this thesis and the antibody dilutions used for western blots and IPs. Ruxolitinib and Fedratinib are not included as they were gifts from Dr Edwin Chen. For primer sequences see Table 3. For vector maps see Section 2.23.

The shRNA/siRNAs and inhibitors used in this thesis are in Table 1:

RNAi	Product code, company
HDAC9 mouse siRNA	EMU191511, Merck
Vectors	-
HDAC9 ORF	NM_024124, Origene
HDAC9 shRNA	sc-35551-SH, SCBT
STAT3 shRNA	TRCN0000071453
STAT5a shRNA clone 8	TRCN0000012549
STAT5a shRNA clone 9	TRCN0000012548
Inhibitors	-
Fludarabine	NSC 118218, SeleckChem
Stattic	S7024, SeleckChem
STAT5 inhibitor	S67884, SeleckChem

Table 1 A table of the shRNA, siRNA and inhibitors used in this thesis.

Inventory			
Antibodies	**Product code, company**	**Antibody dilutions**	
-	-	**Western blot**	**Immunoprecipitation**
JAK2	**#3230, CST**	**1:1000**	**1:200**
Pan-tyrosine phosphorylation	**GT0021, GeneTex**	**1:1000**	**1:100**
P-STAT1 (Y701)	**#7649, CST**	**1:1000**	**-**
P-STAT3 (Y705)	**#9145, CST**	**1:1000**	**-**
P-STAT5(Y694)	**#4322, CST**	**1:1000**	**-**
STAT1	**#14994, CST**	**1:1000**	**1:200**
STAT3	**#9139, CST**	**1:1000**	**1:200**
STAT5	**#94205, CST**	**1:1000**	**1:200**
P-ERK1/2 (T202/Y204)(P44/42 MAPK)	**#9101, CST**	**1:1000**	**-**
Total ERK1/2	**#4695, CST**	**1:1000**	**1:200**
P-AKT (S473)	**#4060, CST**	**1:1000**	**-**
Total AKT	**#4685, CST**	**1:1000**	**1:200**
P-STAT1 (Y701)(Flow Cytometry)	**A17012A, Biolegend**	**1:20**	**-**
P-STAT3 (Y705)(Flow Cytometry)	**A13A3-1, Biolegend**	**1:20**	**-**
P-STAT5 (Y694)(Flow Cytometry)	**A17016B.Rec, Biolegend**	**1:20**	**-**
Annexin V	**640941, BioLegend**	**1:20**	**-**
P-γH2AX(S139)	**#9718, CST**	**1:1000**	**-**
P-CHEK1(S345)	**#2348, CST**	**1:1000**	**-**
CHEK1	**#2360, CST**	**1:1000**	**-**
HDAC9	**sc-398003, SCBT**	**1:500**	**1:100**
FLAG-tag	**20543-1-AP, ProteinTech**	**1:1000**	**1:500**
Cyclin D2	**#3741, CST**	**1:1000**	**-**
C-MYC	**ab32072, Abcam**	**1:1000**	**-**
DNA-RNA hybrid [S9.6]	**MABE1094, Merck/Millipore**	**1:500**	**-**
H3K9Ac	**#9649, CST**	**1:1000**	**-**
Histone H3	**#4499, CST**	**1:1000**	**-**
H4K16Ac	**#13534, CST**	**1:1000**	**-**
Histone H4	**sc-25260, SCBT**	**1:500**	**-**
P-EGFR (Y1068)	**#3777, CST**	**1:1000**	**-**
β-actin	**A2228, Sigma-Aldrich/Merck**	**1:10000**	**-**
Total-EGFR	**#4267, CST**	**1:1000**	**-**
Anti-rabbit IgG, HRP-linked	**#7074S, CST**	**1:5000**	**-**
Anti-mouse IgG, HRP-linked	**#7076S, CST**	**1:5000**	**-**

Table 2 An inventory of antibodies and antibody dilutions.

A list of antibodies used, and the antibody dilution used. For IPs, I have only included the antibody dilutions for the antibodies that were used for IPs in this thesis.

Gene	Forward primer 5'->3'	Reverse primer 5'->3'
ACSS1	GATGCTGGTCGTTACTGG	AGGCTTCGTGGTTGATAG
AIM2	AAGAGAGCCAGGGAAACTCC	TGTCTCCTTCCTCGCACTTT
ANGPT1	ATGGACTGGGAAGGGAACCGAGC	GGGGCCACAAGCATCAAACCACC
APLP2	CTCAGCGGATGATAATGAGCAC	GGTTCTTGGCTTGAAGTTCTGC
BACTIN	GGCACCACACCTTCTACAATG	GGGGTGTTGAAGGTCTCAAAC
BHLHE40	GTGGTTCTGGAGCTTACGTT	GATGAGTGACGAGCTGGGAA
BRCA1	GGGCTGGAAGTAAGGAAACATG	CAGGATGAAGGCCTGATGTAGG
C3AR1	CACCATGGAGTCTTTCGATGCTGACACC	CACATCTGTACTCATATTG
CCND1	ATGGAACATCAGCTGCTGT	TCAGATGTCCACATCCCGC
CCND2	CAGAAGGACATCCAGCCGTAC	TCGGGACTCCAGCCAAGAA
CD45	TGGAGGCTGAATACCAGAGAC	TGCTCATCTCCAGTTCATGC
CDKN2A	TGATGATGATGGGCAACG	ACGGGAACGCAAATATCG
CISH	AAGGTGCTAGACCCTGA	CTCGCTGGCTGTAATAGAA
CLEC11A	AGGTCCTGGGAGGGAGTG	GGGCCTCCTGGAGATTCTT
COL11A2	GGCCTCAGCCTAGCAGATGG	GGCTTATGAAGTCTTGCTGG
CORO2A	TTTGACTCCCAGCGAACTCT	AAAGAGACTATACAGGGGACCG
CPA3	CAACAAACCATGCCTCAATG	CGTTTCAAAACAAGGGCAAT
DEPTOR	ACAAGCCTTCCTGGTGGTTC	CTTTTCTTCATGGAGCCGCAG
DUSP2	TTTGAGGGCCTTTTCCGCTACAAGAG	GCCTCCGCTGTTCTTCACCCAGTC
DUSP4	ATTCCGCCGTCATCGTCTAC	ATAGCCACCTTTCAGCAGGC
DZIP3	CCCCAAGTCAGCACAAGGTT	GCATGTGCCCTGTTGCATCA
EFNB1	ACCAGGAAATCCGCTTCACCATCA	ACAGCATTTGGATCTTGCCCAACC
EPHB4	AATGTCACCACTGACCGTGA	TCAGGAAACGAACACTGCTG
EPOR	AAACTCAGGGTGCCCCTCTGGCCT	GATGCGGTGATAGCGAGGAGAACC
ERO1L	ACCTGAAGAGGCCTTGTCCTTT	TCCATCAGGAACTTCATCAGATTG
ETV5	CCCGGATGCACTCTTCTCTATG	TCGGATTCTGCCTTCAGGAA
FOXO1	ACATTTCGTCCTCGAACCAGCTCA	ATTTCAGACAGACTGGGCAGCGTA
GCH1	TGCTTACTCGTCCATTCTGC	CCTTCACAATCACCATCTCG
G-CSFR	GCAGGGTCCACCAACAGTACAG	AGCAGAGCCAGGTCACTACAC
GFI1	AGGAGGCACCGAGAGACTCA	GGGAGGCAGGGAAGACATC
GM-CSFR	GGAAGCATGTAGAGGCCATCAA	CCTCTTCATTCAACGTGACAGG
HDAC1	CTGTCCGGTATTTGATGGCT	CACGAACTCCACACACTTGG
HDAC10	TAGCAGCCAAACATGCCAAGCAGA	ATGCTCATAGCGGTGCCAAGAGAAA
HDAC11	AATGGGG CAAGGTGATCAAC	AGCCACCACCAACATT GATG
HDAC2	GGCGGCAAGAAGAAAGTGTGC	GGCATCATGTAGTCCTCCAGC
HDAC3	TCTGAGGACTACATCGACTCC	GTCGCCATCATAGAACTCATTG
HDAC4	CAGATGGACTTTCTGGCCG	CTTGAGCTGCTGCAGCTTC
HDAC5	GAAGCACCTCAAGCAGCAGCAGG	CACTCTCTTTGCTCTTCTCCTTGTT
HDAC6	ACGCTGACTACATTGCTGCT	TCTCAACTGATCTCTCCAGG
HDAC7	AGCTGGCTGAAGTGATCC	TCACCATCAGCCTCTGAG
HDAC8	AACACGGCTCGATGCTGG	CCAGCTGCCACTTGATGC
HDAC9	GCGGTCCAGGTTAAAACAGA	GCCACCTCAAACACTCGCTT

HDAC9	**GGTGATGATTCTCGGAAATTCT**	**GAAGCCAGCTCAATGACACA**
HOXD8	**TTCCCTGGATGAGAC CACAAGCAGC**	**GTCTCTCCGTGAGGG CCAGAGT**
IFIH1	**TGATGCACTATTCCAAGAACTAACA**	**TCTGTGAGACGAGTTAGCCAAG**
IL-5RB	**TCTAGGGCCACCAGGGTTTGAGGA**	**CCAAGGGGGCAGGGACAGGTAGT**
ITGA2B	**GGGCCTTCCTTCGGGAT**	**AGCACAATTGGGCTCAGCTT**
KLF6	**GGACCAAATTCATTCTAGCTCGGG**	**AGGCGTCGCCATTACCCTTG**
LIG4	**AGTCCACCATCGCTCTGC**	**GATGCAACAGTTTGTGAAGTTTG**
LYZ2	**AATGGCTGGCTACTATGGAGTCA**	**TGCTCTCGTGCTGAGCTAAACA**
MDGA1	**GGCTGTGTGGCGTTTCAAAG**	**TTGGAGACGCTGCACTCGTA**
ME2	**TTCTTAGAAGCTGCAAAGGC**	**TCAGTGGGGAAGCTTCTCTT**
MPL (ENDO)	**TATTGGCAGCAGCCCTGAA**	**TGGATGGTGTTGAGGATGGATA**
MPL(TOTAL)	**CTGTATGCCTACCGAGGAGAGAAG**	**GTTCCAAAGGTGGGCACACT**
MPO	**TCCCACTCAGCAAGGTCTT**	**TAAGAGCAGGCAAATCCAG**
MTUS1	**CTGAAGC AACACAAAACCCTCTCTC**	**TGTCTG ATGCTGCTGGTTTAGTTTC**
NDRG1	**ACCCTGAGATGGTAGAGGGTCTC**	**CCAATTTAGAATTGCATTCCACC**
NEDD4L	**TCCAATGGTCCTCAGCTGTTTA**	**ATTTTCCACGGCCATGAGA**
NRGN	**AGAGAGGCTGGTTCTGCAAG**	**AATATCGTCGTCTGGCTTGG**
P2RX3	**ACAAGATGGAGAATGGCAGC**	**GCAGGATGATGTCACAGAGAAC**
PAIS3	**AGTGGCTACTACAAGTCTCCG**	**GATCTCATCACAATCCGAAC**
PAOX	**GGGAAGATACATCGCCCTTACA**	**GGACCAAAAATCCAATGAGCTT**
PF4	**CAGTCCTGAGCTGCTGCTTCT**	**TCCAGGCTGGTGATGTGCTTA**
PFKP	**CGCCTATCCGAAGTACCTGGA**	**CCCCGTGTAGATTCCCATGC**
PGK1	**ATGTCGCTTTCCAACAAG**	**CTAAACATTGCTGAGAGCATCC**
PIAS1	**TCCTGCTGTAGATACAAGCTAC**	**TGCCAAAGATGGACGCTGTGTC**
PIASX	**CAGATTTGCCAGGTTTGGAT**	**GCTTGCTGTTAAGGGTGAGG**
PIASY	**GTTCGACTGTAGCCCAGAGC**	**GTACCTCCGGCTTAAGGGTC**
PLPPR3	**GCATAGTGTACATGGAGAAGGAGG**	**GTGGCAGCATGTATAAGCAAGTG**
PPM1L	**ATGATGAAGCAGGCACAACG**	**GTCCAGGTCAAAGGTCAAGATG**
PTPN2	**CCAGTTTAGTTGACATAGAAGAGG**	**GCAGCATGTGTTAGGAAGT**
RAD51	**TTGCTGACTGCCAGTCTAGC**	**GATTGCACGCAGGTT**
RBBP7	**CTTCTAAGCCGAGCCATTT**	**GATTACGCAGGTCCCATA**
RECQL5	**CAGGCTGACTGTGAAGG3**	**TTGGAGGCTCTGTGGAT3**
RGS1	**ACGAGCAGCCATCTCCATGCC**	**CCCAGATTCCAGATGTGGGAT**
RNASEH1	**ATGAGCTGGCTTCTGTTCCT**	**AACAAGTCTTCGGTGCGTTC**
RNASEH2	**CTGCCAGAAAACTGACACCC**	**ACAAAGTCCGTGTCCTCCAT**
SHP-1	**GTGACCCATATTCGGATCCAG**	**CTTGAAATGCTCCACCAGGTC**
SIRT1	**CGGCTACCGAGGTCCATATAC**	**CAGCTCAGGTGGAGGAATTGT**
SIRT2	**GAGCCGGACCGATTCAGAC**	**AGACGCTCCTTTTGGGAACC**
SIRT3	**GGATTCGGATGGCGCTTGA**	**CACCTGTAACACTCCCGGAC**
SIRT4	**GAGCATTCTTACTAGGGATTCCA**	**AACGGCTAAACAGTCGGGTT**
SIRT5	**GCCACCGACAGATTCAGGTT**	**CCACAGGGCGGTTAAGAAGT**
SIRT6	**CCAAATCGTCAGGTCAGGGA**	**CAGAGTGGGGTACAGGGATG**
SIRT7	**CTAAGCGAAGCGGAGCCTAC**	**GTGGAGCCCATCACAGTTCT**
SLC2A1	**AACTGGGCAAGTCCTTTG**	**TTCTTCTCCCGCATCATCTG**
SOCS1	**CACTCACTTCCGCACCTTCC**	**TCCAGCAGCTCGAAAAGGCA**
SOCS2	**AAGACATCAGCCGGGCCGACTA**	**GTCTTGTTGGTAAAGGTAGTC**

SOCS3	**TCACCCACAGCAAGTTTCCCGC**	**GTTGACAGTCTTCCGACAAAGATGC**
SOD2	**CATTCTCCCAGTTGATTACATTCC**	**ACAAGCACAGCCTCCCAGACC**
ST8SIA4	**GCACCAAGAGACGCAACTCATC**	**CAGAGCTGTTGACAAGTGATCTGC**
STAT3	**AATGGAAATTGCCCGGATC**	**AGGCGAGACTCTTCCCACAG**
STAT5A	**CGCTGGACTCCATGCTTCTC**	**GACGTGGGCTCCTTACACTGA**
STAT5B	**GGACTCCGTCCTTGATACCG**	**TCCATCGTGTCTTCCAGATCG**
TCN2	**CTTTGCTGGATCTTCCTTGG**	**TCCTGGGGTTTGTAGTCAGC**
TFR2	**CCTCTATGAACAAGTGGCA**	**GATCATCCTCCATGAAG**
TIMELESS	**CGCCAGAGAGAGATCCTGC**	**CCCTTCCCCTTGAATTTCGTCT**
TNS1	**AGAGGAAGGAGAGACAGGC**	**TGCTGTAGTGCATATAAGC**
TPI1	**TACAAAGTGACTAATGGGGC**	**TCGAAAACAACCTTCTCAGT**
UNC13A	**GCTGTGCGTGGGAGTCAAA**	**CAGCTATGGTAGTGCTCTTCAC**
VWF	**TAAGTCTGAAGTAGAGGTGG**	**AGAGCAGCAGGAGCACTGGT**
ZFP36	**GGTACCCCAGGCTGGCTTT**	**ACCTGTAACCCCAGAACTTGGA**
ZFP503	**AGGTGCTGAAGATGCTGACG**	**CGATGTCGCTCAGCTTCAAG**

Table 3 A table of the primers used in this thesis and their sequences.

2.23 Vector maps

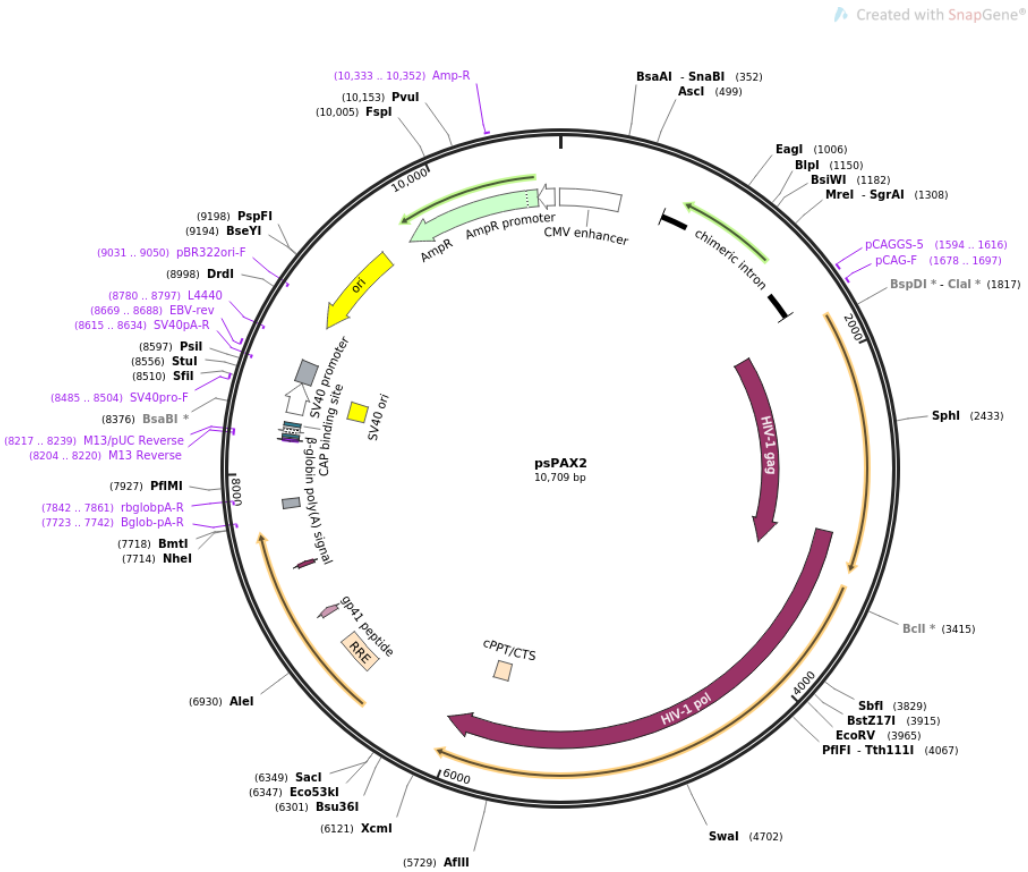


Figure 2.2 A vector map of psPAX2.

A vector map of psPAX2, a vector used in the production of lentiviral particles for the lentiviral spin infection protocol. psPAX2 is a vector that encodes components for viral packaging.

psPAX2 was a gift from Didier Trono (Addgene plasmid # 12260 ; http://n2t.net/addgene:12260 ; RRID:Addgene_12260)

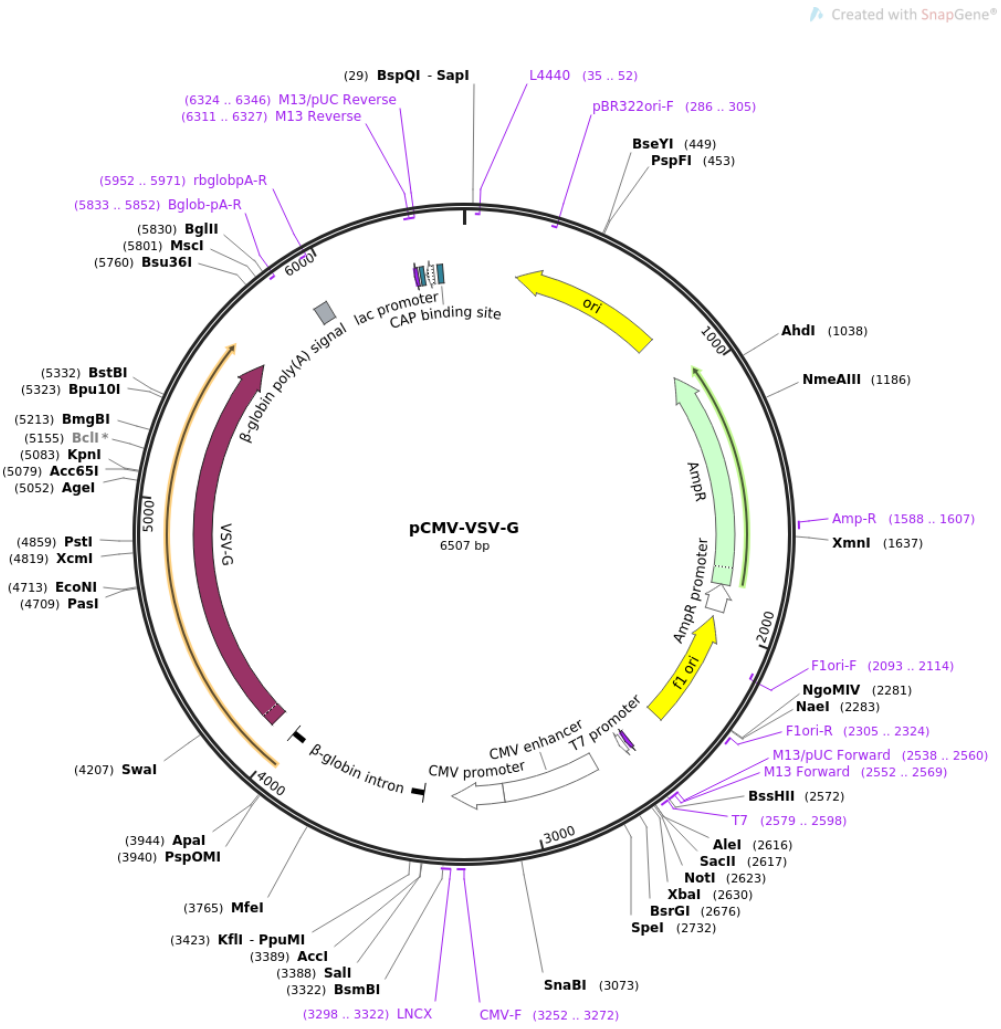


Figure 2.3 A vector map of pCMV-VSV-g.

A vector map of pCMV-VSV-g, a vector used in the production of lentiviral particles for the lentiviral spin infection protocol. pCMV-VSV-g is a vector that encodes the viral envelope.

pCMV-VSV-g was a gift from Bob Weinberg (Addgene plasmid #8454; http://n2t.net/addgene:8454; RRID: Addgene_8454)

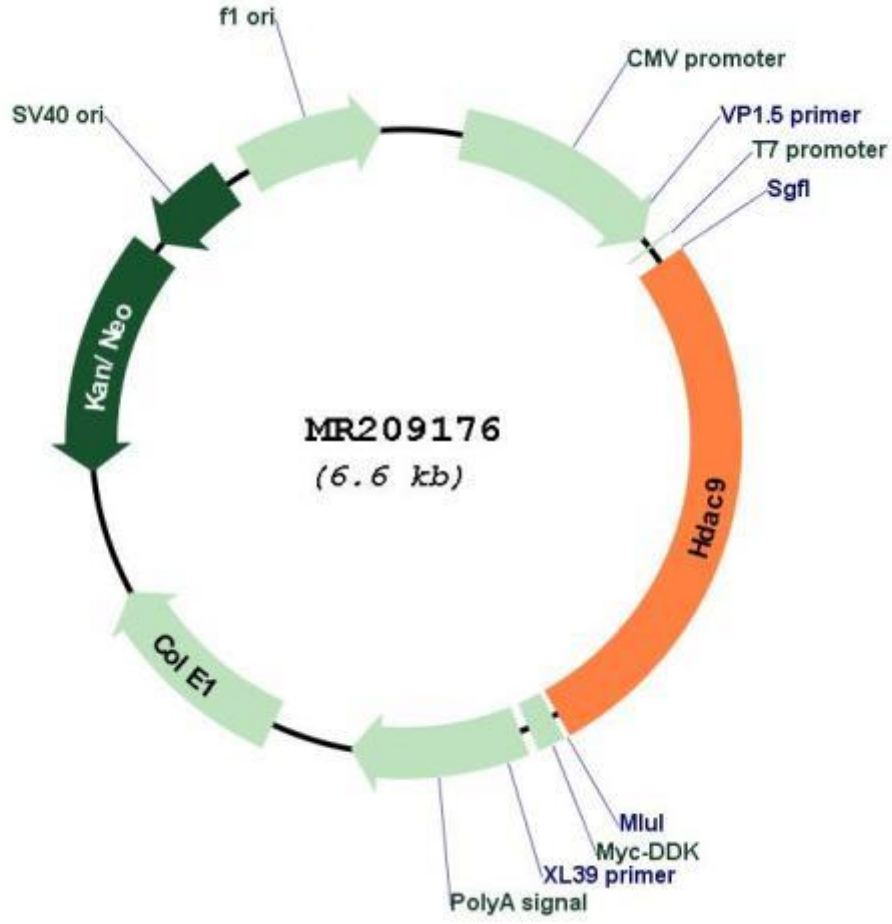


Figure 2.4 A vector map of MR209176 (*HDAC9* ORF)

A vector map of MR209176, a pCMV6-entry vector engineered to express full-length HDAC9 ORF with a MYC and FLAG tag attached.

This vector was purchased from Origene, product code NM_024124.

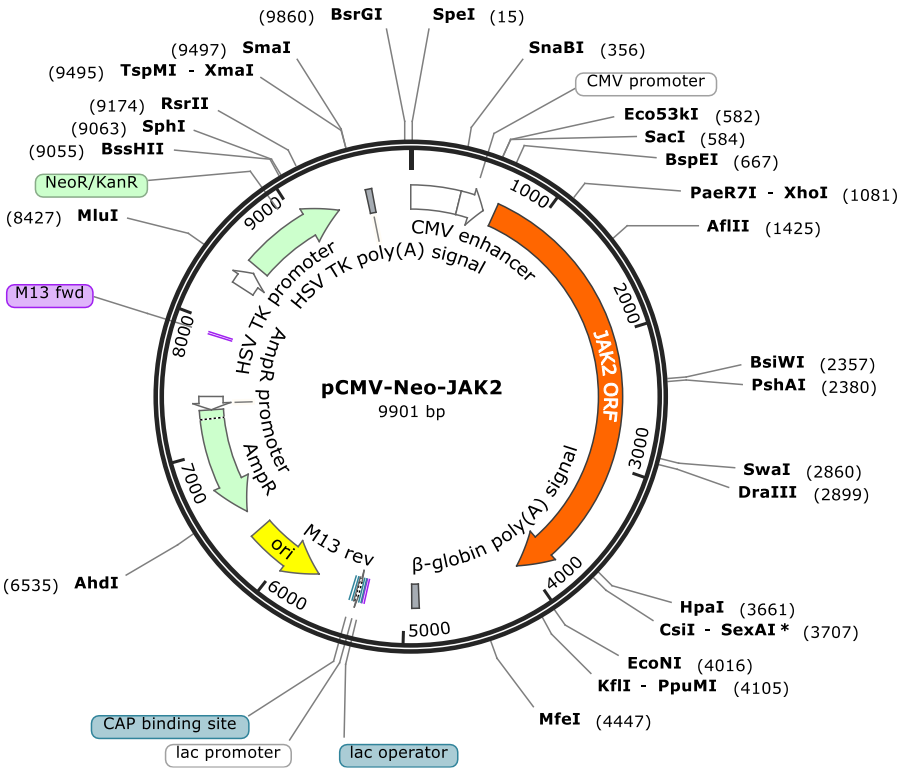


Figure 2.5 A vector map of pCMV-NEO-JAK2

A vector map, a pCMV plasmid backbone containing an open reading frame for JAK2. This plasmid was used to produce HPC7-JAK2WT cells and was mutated by site-directed mutagenesis to produce HPC7-JAK2V617F cells.

This vector was a gift from Dr Edwin Chen.

Chapter 3

JAK2V617F overexpression in HPC7 cell lines instigates changes in cell signalling, DNA damage and cell proliferation.

Chapter 3

JAK2V617F overexpression in HPC7 cell lines instigates changes in cell signalling, DNA damage and cell proliferation.

3.1 Introduction to Chapter 3

3.1.1 Introduction

In 2005 there was a landmark finding that the majority of MPNs contained a mutation in the *JAK2* gene, specifically, a somatically acquired point mutation guanine to thymine transversion in exon 14 that results in a replacement of valine with phenylalanine at amino acid position 617 of the JAK2 protein (Levine et al., 2005; Kralovics et al., 2005; James et al., 2005; Baxter et al., 2005). The *JAK2* gene encodes Janus kinase 2, a kinase which associates with a plethora of receptor proteins to transduce a signal and begin a signalling cascade relevant to the receptor and the ligand binding stimulus received by the receptor, resulting in increased STAT phosphorylation, activation of the AKT and MAPK pathways and deregulation of appropriate cellular behaviour (Yan, D. et al., 2012; Zou et al., 2011).

Cells carrying the JAK2V617F mutation can persist in patients for long periods of time before appearance of overt disease, indicating that the mutations likely occur within the LT-HSC population (Gale et al., 2007). Evidence to support this idea came from studies that identified JAK2V617F in LT-HSCs of MPN patients, sorted using cell surface markers CD34+ CD38-CD90+Lin- (Delhommeau et al.,

2007; Jamieson et al., 2006; Ishii et al., 2006), suggesting that the JAK2V617F mutation was carried in a small population of stable LT-HSCs that propagated the disease over time (Xie et al., 2014; Jaiswal et al., 2014; Genovese et al., 2014).

JAK2V617F-driven MPNs are a complex disease with their phenotype being defined by many contributing factors such as allelic burden, additional mutations, gender, age, and the haematopoietic stem cell niche. JAK2V617F commonly presents as a homozygous mutation in MPNs (Kralovics et al., 2002) and has been shown to exert its influence depending on the proportion of its homozygosity between clones with highly homozygous MPNs promoting erythropoiesis, self-renewal/proliferation and generally a more deleterious phenotype. Low homozygosity however, loosely defined as >50%, sharply contrasts a malignant phenotype with a much more nuanced phenotype polarised to megakaryocytic differentiation. In mouse models high homozygosity results in a far more progressive and severe disease due to a more pronounced activation of STATs, ERK1/2, and MAPK pathways (Akada et al., 2014).

Recent publications have shown that the haematopoietic stem cell niche plays an integral part in the pathogenesis of MPNs by promoting increased self-renewal. For example, stem cell expansion and MPN disease was triggered in transgenic mice expressing JAK2V617F exclusively in megakaryocytes (Zhan et al., 2016). Endothelial cells harbouring the JAK2V617F mutation also promote expansion of non-mutated Lin-/c-KIT+ cells, which can contribute to disease relapse following bone marrow serial transplantation (Zhan et al., 2018). This data suggests that analysis of JAK2V617F on HSCs utilising mouse models or patient samples may be very difficult due to the multi-layers of complexity, the

interactions between differing cell types and various passenger mutations naturally acquired over time.

3.1.2 Haematopoietic Precursor 7 cell line

Haematopoietic Precursor 7 (HPC7) is a cell line first established in 1998 by Pinto do et al. (Pinto do et al., 1998). This study aimed to investigate the role of the Lim homeobox 2 (*LH2*) gene in haematopoiesis. The *LH2* gene is crucial for haematopoiesis within the foetal liver (Xu et al., 1993) and *LH2* knockouts led to impaired erythropoiesis which contributed to mortality via severe anaemia in mice (Porter et al., 1997).

The HPC7 cell line was generated by the retroviral infection of the mouse *LH2* gene into mouse embryonic stem cells (ESCs). Post-introduction of *LH2*, ESCs were expanded with leukaemia inhibitory factor (LIF) and differentiated into embryoblasts. The transduced cells were stimulated with SCF and erythropoietin which generated the SCF-dependent HPC7s which are erythroid cells with a remarkably similar immunophenotype and genetic profile to that of a LSK cell (Schütte et al., 2016; Pinto do et al., 1998).

HPC7 cells are an uncommonly used cell line which have been shown in recent publications to have a highly homologous genetic profile to the profile of a murine HSPC (Schütte et al., 2016). Therefore, this provides an excellent opportunity to investigate the influence of JAK2V617F on immature haematopoietic progenitors through use of HPC7 cells as a model.

3.2 Aims and Hypothesis.

There is contradicting information regarding the effect of the JAK2V617F mutation on haematopoietic stem cells. Most of the current published data is guided by mouse models or patient data and as previously mentioned, the complexities of patient data and mouse models make it difficult to gauge the exact effect on HSCs regarding cellular signalling and DNA damage. I wanted to examine the effect of JAK2V617F overexpressed on an HSC model, Hematopoietic precursor 7 (HPC7). The HPC7 cell line has a genetic profile highly similar to that of a genuine HSC making them an ideal system.

I hypothesise that overexpressing JAK2V617F in HPC7 cells will generate a phenotype close to those seen in MPN patients and mouse models but without the results being contaminated by other factors. In this chapter I will discuss the generation, validation, and characterisation of HPC7-JAK2WT or HPC7-JAK2V617F cell lines.

1) Validate the JAK2WT and JAK2V617F HPC7 cell lines.

2) Characterisation of the JAK2WT and JAK2V617F HPC7 cell lines with a focus on JAK-STAT signalling and DNA damage.

3.3 Results

3.3.1 Generation and validation of HPC7-JAK2WT and HPC7-JAK2V617F cell lines.

To produce the cell line model, I took vectors expressing human *JAK2* cDNA and performed a site directed mutagenesis to introduce the V617F mutation. The vectors expressing either the wild-type *JAK2* cDNA or the *JAK2V617F* cDNA were subsequently introduced into the HPC7 cell line by lentiviral spin-infection and were designated HPC7-JAK2 Wild Type (HPC7-JAK2WT) or HPC7-JAK2V617F. Post-infection, the cells were selected using geneticin which is expressed under the control of the herpes simplex virus tyrosine kinase promoter in the lentiviral vector (**Figure 2.5**). Post-selection, I wanted to validate the introduction of the gene into the cells. Therefore, I extracted genomic DNA (gDNA) from the parental HPC7, HPC7-JAK2WT and HPC7-JAK2V617F cells and amplified the gDNA with human *JAK2* primers, which were a generous gift from Dr Edwin Chen.

To validate that the amplified human JAK2V617F and JAK2WT were present in the genomic DNA of the HPC7 cells, I sent the PCR products for Sanger sequencing using the reverse primer that was used previously for the RT-PCR experiment (**Figure 3.1A**). The HPC7-JAK2WT and HPC7-JAK2V617F cell lines had incorporated the appropriate *JAK2* cDNA.

The now validated HPC7-JAK2WT and HPC7-JAK2V617F cell lines were then analysed to determine if the newly introduced *JAK2WT* or *JAK2V617F* cDNA was effectively conferring a mutant phenotype. In patients and mouse models,

the JAK2V617F protein is chronically phosphorylated, therefore we measured JAK2 phosphorylation. To this end, I performed immunoprecipitations by pulling down the JAK2 protein with an antibody against JAK2 and staining with a pan phospho-tyrosine antibody (**Figure 3.1B**). The increased phosphorylation of the JAK2V617F mutant confirms that the cells express active JAK2.

A consequence of JAK2V617F expression is the chronic phosphorylation of STATs, therefore I investigated STAT phosphorylation in HPC7-JAK2V617F cells. Specifically, I investigated STAT1, STAT3 and STAT5 phosphorylation (P-STAT or phospho-STAT) in the HPC7-JAK2WT and HPC7-JAK2V617F cell lines by western blotting (**Figure 3.2A**) and intracellular flow cytometry (**Figure 3.2B**). The data in Figure 3.2 suggests that in HPC7-JAK2V617F cells, STAT1, STAT3 and STAT5 all have much higher levels of phosphorylation at the tyrosine residue associated with the activated JAK-STAT pathway, relative to HPC7-JAK2WT or parental cells. For even further confirmation of active signalling, I probed the HPC7-JAK2WT and HPC7-JAK2V617F cell lines for phospho-AKT, phospho-ERK1/2 and expression levels of cyclin D2 and c-MYC by western blot (**Figure 3.3A**). A series of commonly upregulated genes in MPN patients were also quantified by RT-qPCR (**Figure 3.3B)** to additionally confirm the relevance of the new disease model. Figure 3.3A suggests that there is a much higher level of phosphorylated AKT, ERK1/2, cyclin D2 and a much lower amount of c-MYC in the HPC7-JAK2V617F cells relative to the HPC7-JAK2WT. The results of the RT-PCR suggest that the genes that are commonly dysregulated in patients are also dysregulated in my model. To assess that the changes are due to increased JAK2 signalling, I used Ruxolitinib, a JAK1 and JAK2 inhibitor as a negative control in Figure 3.2 and Figure 3.3.

The HPC7-JAK2WT and parental cell lines had high levels of phospho-STAT3. I hypothesise that either FBS in the media or MSCF may be causing this, if I am going to investigate JAK-STAT signalling it is important that I determine this, so I performed starvation experiments with the HPC7-JAK2WT/V617F removing either FBS or serum or both and measuring Phospo-STAT3 levels via a western blot (**Figure 3.4**). Figure 3.4 demonstrates that when starved HPC7-JAK2WT loses its phospo-STAT3 signature whilst HPC7-JAK2V617F retains a strong signal.

Figures 3.1 – 3.4 all corroborate that the HPC7-JAK2V617F cell line is a functional, robust model for haematopoietic with the stereotypical cell signalling cascades and genetic profile associated with JAK2V617F-driven MPN patients also being reproduced in my HPC7-JAK2V617F cell line.

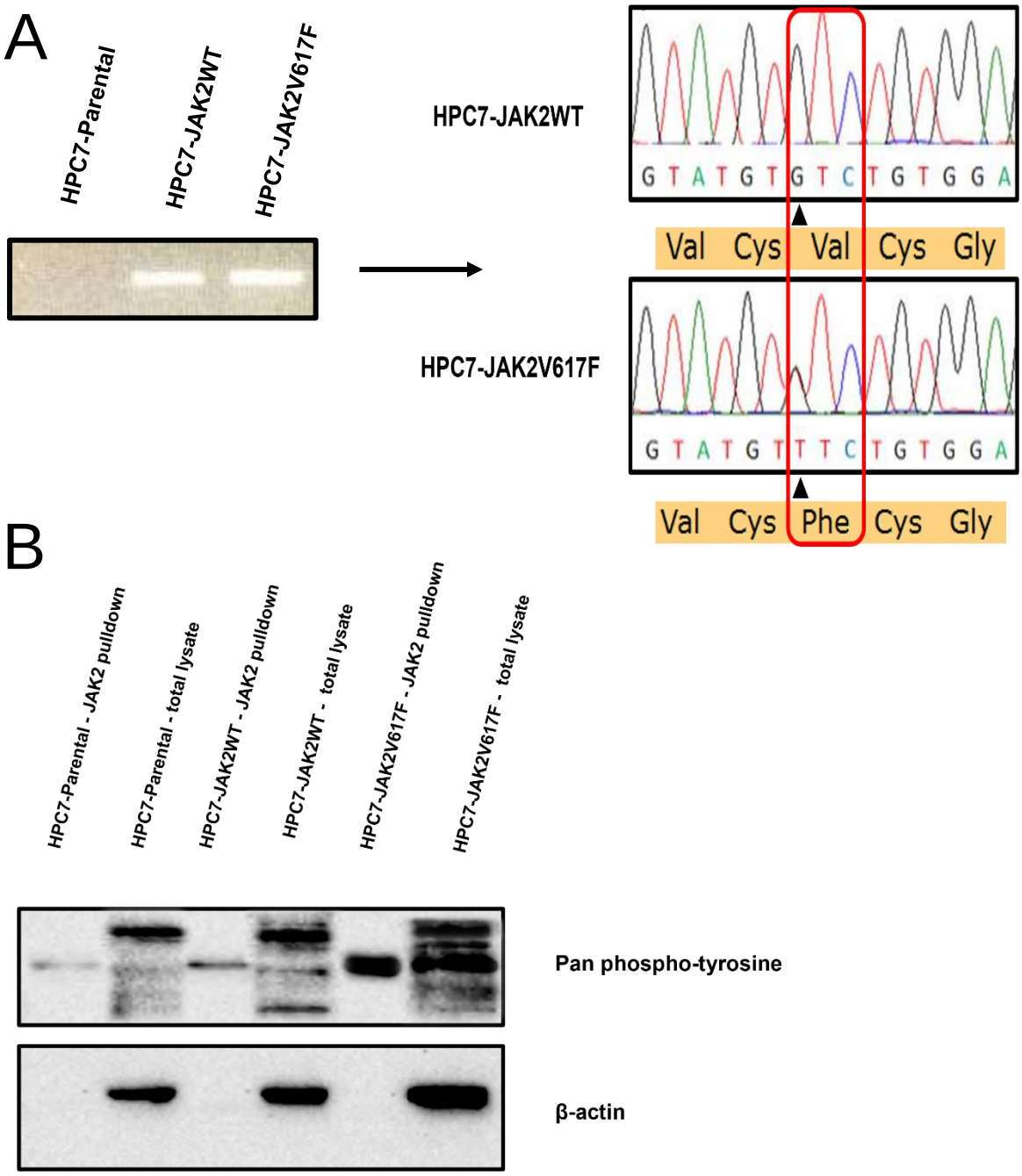


Figure 3.1 Validation of HPC7-JAK2WT and HPC7-JAK2V617F cell lines.

(A) Amplified human *JAK2* cDNA visualised on an agarose gel demonstrating the expression of the human *JAK2* cDNA Sanger sequencing data illustrates (right) the presence of *JAK2WT* or *JAK2V617F* cDNA with the G to T transversion and the subsequent amino acid translation highlighted. (B) Immunoprecipitation experiment "pulling down" JAK2 protein and staining for pan tyrosine phosphorylation. HPC7-JAK2V617F pull-down demonstrates a much higher phosphorylation relative to HPC7-JAK2WT and the parental cell line.

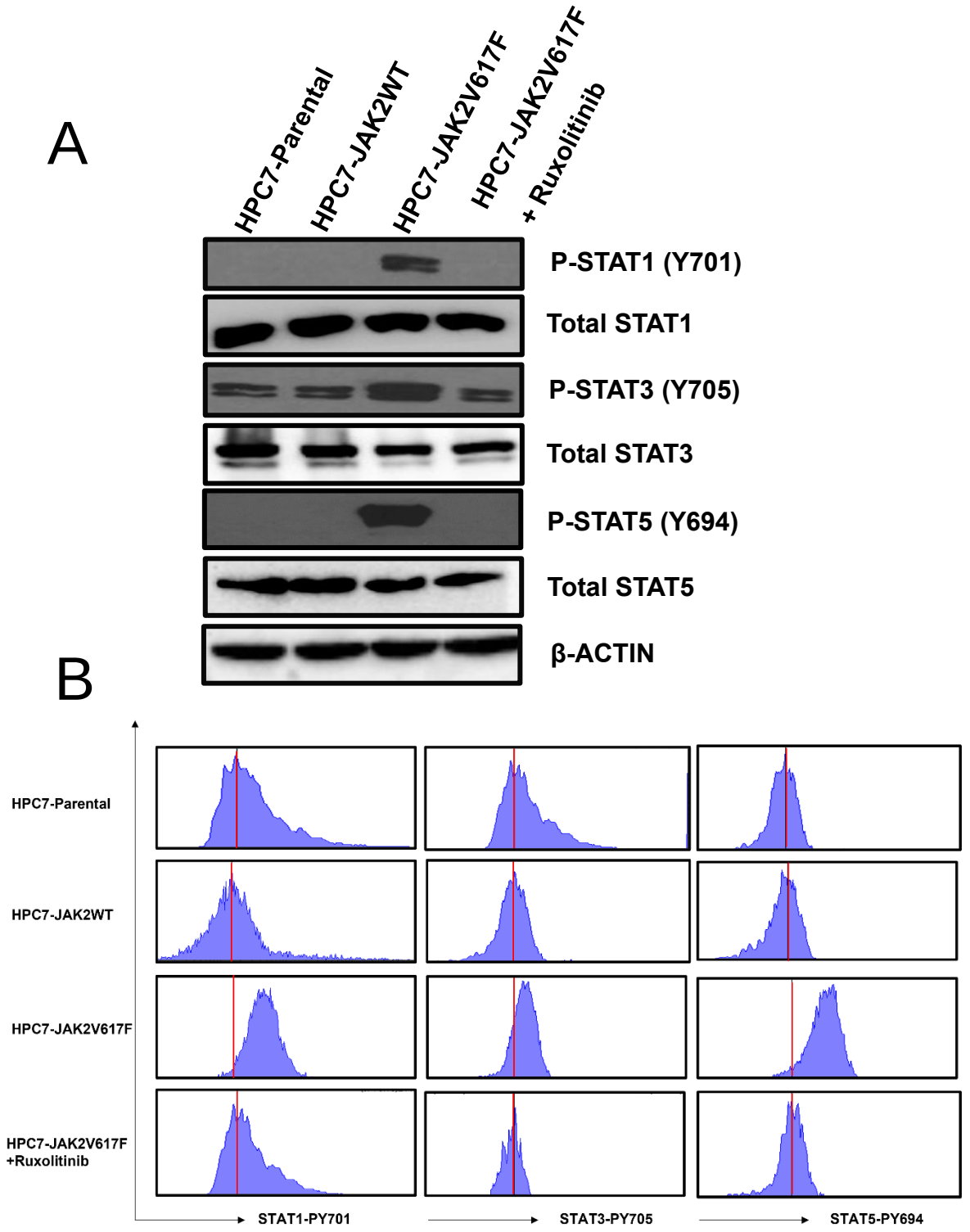


Figure 3.2 STAT phosphorylation in HPC7-JAK2WT and HPC7-JAK2V617F cell lines.

(A-B) Immunoblotting (A) and intracellular flow cytometry (B) of phosphorylated STATs 1, 3 and 5 demonstrate highly increased phosphorylation of STATs 1, 3 and 5 in HPC7-JAK2V617F cells relative to the HPC7 JAK2WT cells. Ruxolitinib was added for 6 hrs at 500 nM concentration .This indicates a highly upregulated JAK-STAT signalling cascade.

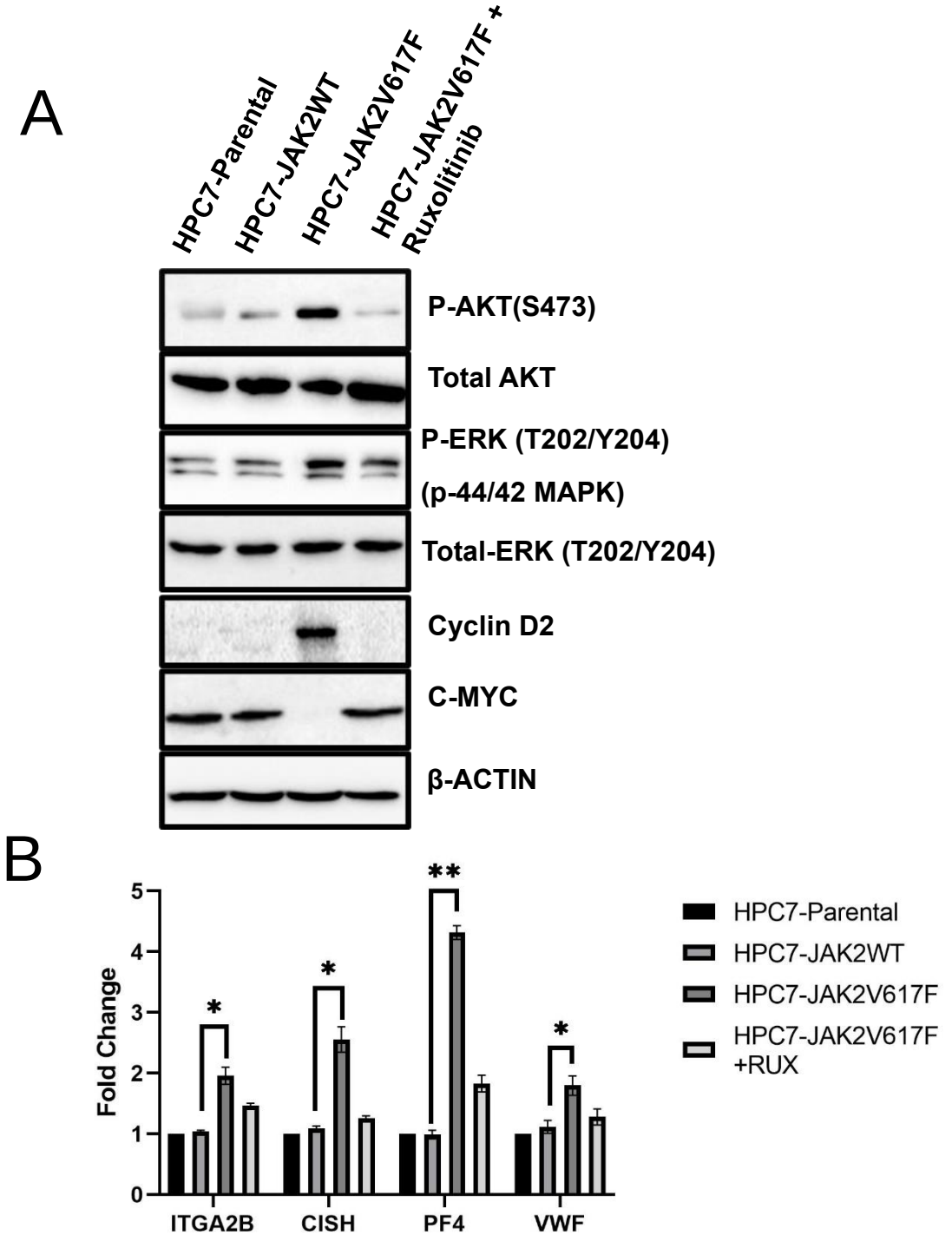


Figure 3.3 Activated cell signalling cascades and gene dysregulation in HPC7-JAK2V617F cell lines

(A) Western immunoblots demonstrating a higher amount of P-ERK1/2, P-AKT, cyclin D2 and c-MYC in the HPC7-JAK2V617F cells relative to the HPC7-JAK2WT cells. (B) RT-qPCRs of HPC7-JAK2WT/V617F cells assessing mRNA levels of genes that are commonly dysregulated in MPN patients and mouse models. (A-B) Ruxolitinib is included as a negative control of JAK-STAT signalling. In B, testing for statistical significance was performed using a student's t-test (*: $p<0.05$; **: $p<0.01$; *** $p<0.001$). Three biological replicates were performed.

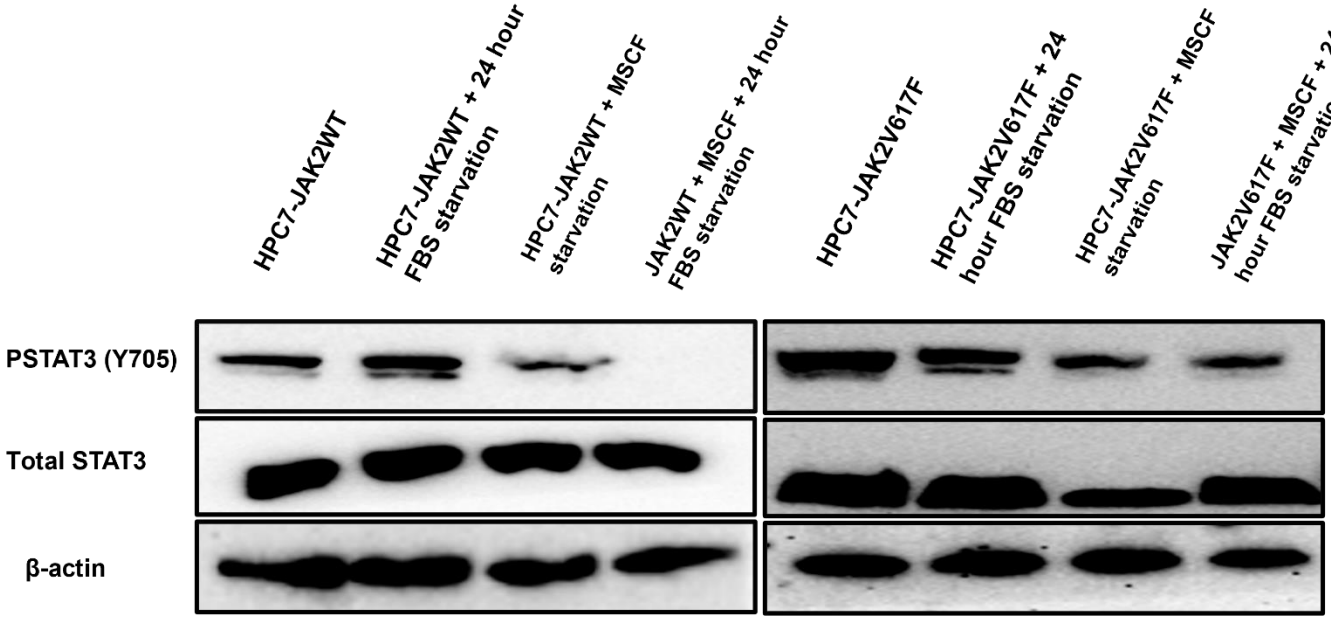


Figure 3.4 HPC7-JAK2WT/V617F FBS and MSCF starvation.

Western blots of HPC7-JAK2WT cell lysate post 24-hour starvation from MSCF, FBS or both. This Figure illustrates that the MSCF and FBS contributes to the strong presence of Phospho-STAT3 in the HPC7-JAK2WT cell line. HPC7-JAK2V617F phopho-STAT3 levels are also affected by serum and MSCF removal but even when starved show a strong signal.

3.3.2 HPC7-JAK2V617F cells demonstrate a higher cell proliferation and increased cell cycle progression

There has been conflicting evidence over whether JAK2V617F imparts a clonal expansion advantage (Staerk and Constantinescu, 2012).

In Section 3.3.1, I performed western blots for phospho-STAT3, phospho-ERK1/2, phospho-AKT, cyclin D2 and c-MYC which are all proliferation/self-renewal factors. Notably, conditional knockouts of *c-MYC* in mice were published in 2004 and it was found that *c-MYC* knockdown led to an expansion of HSCs and a drive towards increased self-renewal and away from quiescence leading to HSC exhaustion (Wilson et al., 2004).

The expression of JAK2V617F in the HPC7 cell line led to a decreased level of c-MYC protein (**Figure 3.3A**). Therefore, I hypothesise that the HPC7-JAK2V617F cell line could have an increased level of cell proliferation relative to the HPC7-JAK2WT cell line.

To test my hypothesis, I performed an MTT assay to measure viable cells. This colorimetric assay determines the conversion of a yellow MTT solution taken by live cells through endocytosis, into reduced purple formazan crystals by mitochondrial oxidoreductases in the cell. Therefore, the higher the metabolic activity or number of viable cells, the higher the reduction to formazan, which can be measured at 570 nm on a plate reader. Trypan Blue exclusion assay to count viable cells was used to validate the MTT assay (**Figure 3.5)**. Briefly, after washing the cells to remove all MSCF, HPC7-JAK2WT or HPC7-JAK2V617F cells were plated and treated with an increasing concentration of MSCF, or no

growth factor, and left overnight. MTT assays and Trypan Blue exclusion assays were set up to address the following questions 1) Can JAK2V617F could provide growth independence from MSCF? (**Figure 3.5)** and 2) Does expression of JAK2V617F cause an increase in cell proliferation at lower levels of MSCF compared to cells expressing JAK2WT? (**Figure 3.6**).

The Trypan Blue exclusion assay and MTT assay analysis that were performed over 7 days during which no MSCF was added to the media, demonstrate that JAK2V617F overexpression does not confer MSCF independence. Indeed, the proliferative capacity was reduced for all cell lines, and addition of the JAK1/2 inhibitor Ruxolitinib further reduced HPC7-JAK2V617F proliferation (**Figure 3.5**). From the MTT assay shown in Figure 3.6, it was concluded that the proliferation of HPC7-JAK2V617F cells is significantly higher than HPC7-JAK2WT cells at MSCF concentrations up to 10 µg/ml. However, at 40 µg/ml MSCF and above, the proliferation levels for both cell lines are the same suggesting that between 10-40 µg/ml there is a threshold at which the abundance of MSCF saturates the signalling cascades triggering cell proliferation.

Cyclin D2 is a well-known cell cycle regulator responsible for regulating the G1/S phase of the cell cycle and is essential for entry into the G2 phase (Koyama-Nasu et al., 2013). In MPNs, cyclin D2 is upregulated by the transcription factors STAT3 and STAT5 when activated (Park et al., 2019; Martino et al., 2001). In my HPC7-JAK2V617F cell lines cyclin D2 is highly upregulated (**Figure 3.3A**) and since their proliferation capacity is also increased relative to HPC7-JAK2WT, I predicted that more cells will be situated in later cell cycle phases.

To validate this prediction, I prepared the HPC7-JAK2WT and HPC7-JAK2V617F cell lines in the same manner as Figures 3.5 and 3.6 but plated them with 1 µg/ml of MSCF and cultured them for a further 24 hours. The nuclear DNA in HPC7-JAK2WT/V617F cells was then stained with propidium iodide (**Figure 3.7**), and the results were separately confirmed by BrdU/DAPI staining (data not shown) of the HPC7s, followed by measurement of the number of cells within each cell cycle phase using flow cytometry. As predicted, a higher percentage of HPC7-JAK2V617F cells are present in G2 compared to HPC7-JAK2WT cells.

During S phase, DNA is replicated which can lead to DNA damage, whereas G2 allows the cell the opportunity to repair damaged chromatids using their sister chromatids as a template for homology-directed repair (Burgoyne et al., 2007). My hypothesis is that the HPC7-JAK2V617F cells may have more DNA damage which may explain the increased number of HPC7-JAK2V617F cells in G2. This hypothesis is investigated in Section 3.3.3.

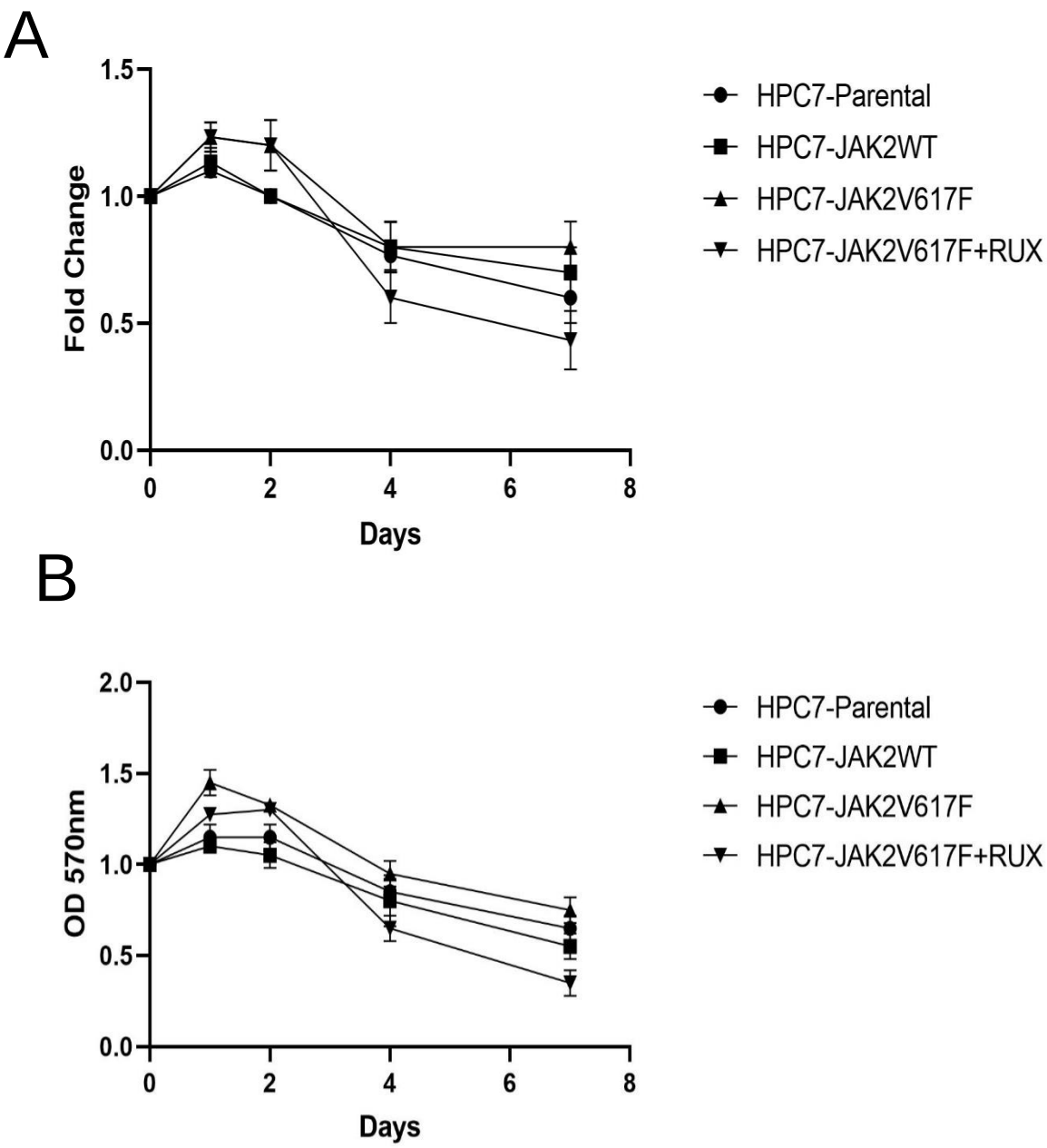


Figure 3.5 JAK2V617F does not confer MSCF independence.

(A) Trypan Blue exclusion assay and (B) MTT assay measuring fold change/OD 570 nm over 7 days. Both methods indicate that even with JAK2V617F being expressed, HPC7 cells still require MSCF to proliferate. Testing for statistical significance was performed using a student's t-test (*: $p<0.05$; **: $p<0.01$; *** $p<0.001$). Three biological replicates were performed.

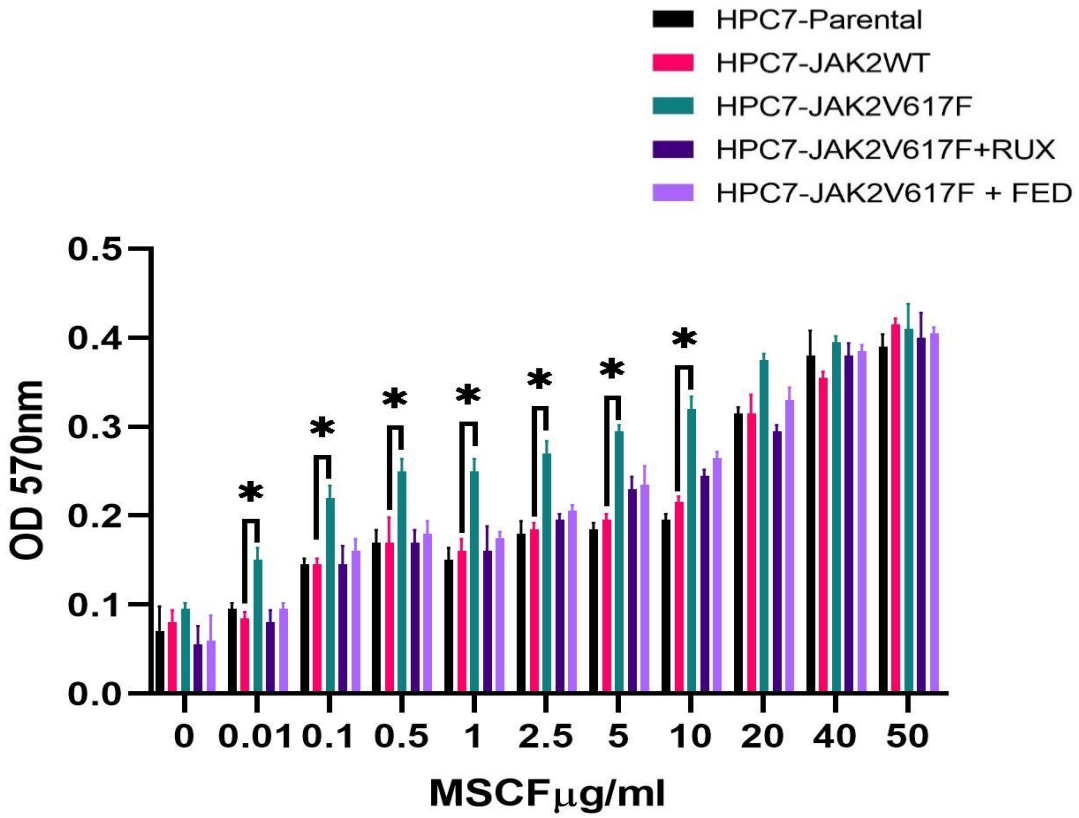
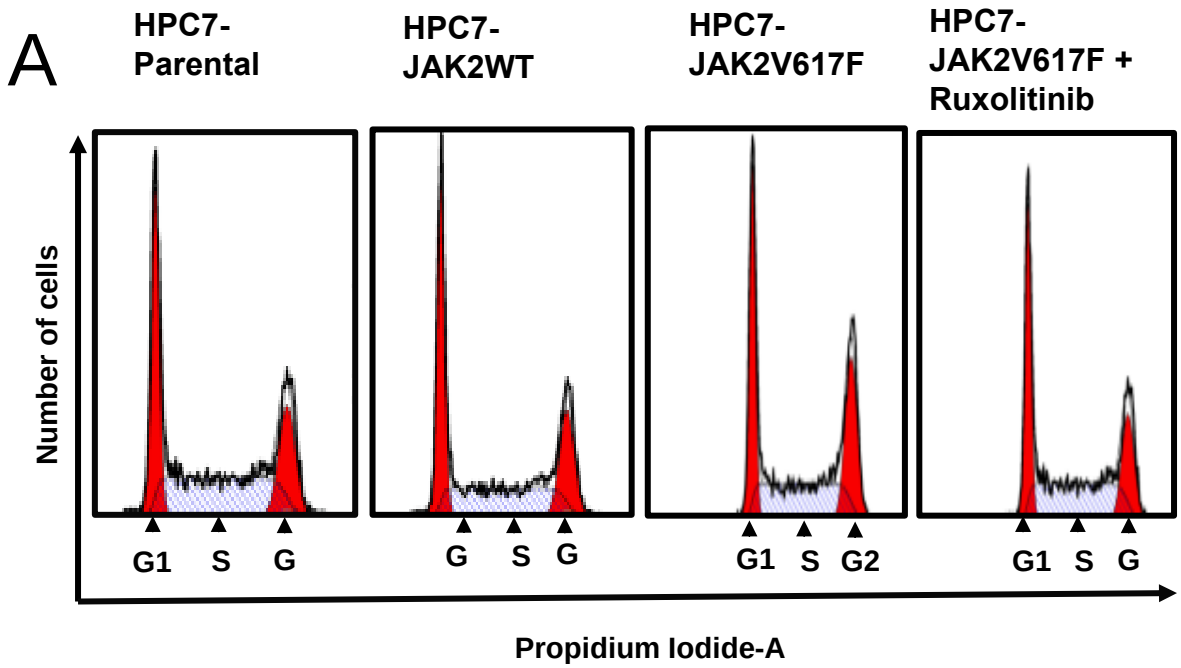


Figure 3.6 JAK2V617F confers a proliferative advantage at lower concentrations.

MTT assay demonstrates that JAK2V617F expression in HPC7 cells confers a proliferative advantage relative to the HPC7-JAK2WT cells only at lower concentrations of MSCF, in the range of 0.01 to 10 µg/ml. Fedratinib (FED), a JAK2 specific inhibitor, and Ruxolitinib (RUX), a dual JAK1/JAK2 specific inhibitor, were applied and had a final concentration of 500 nM for 6 hours.

Testing for statistical significance was performed using a student's t-test (*: $p<0.05$; **: $p<0.01$; *** $p<0.001$). Three biological replicates were performed.



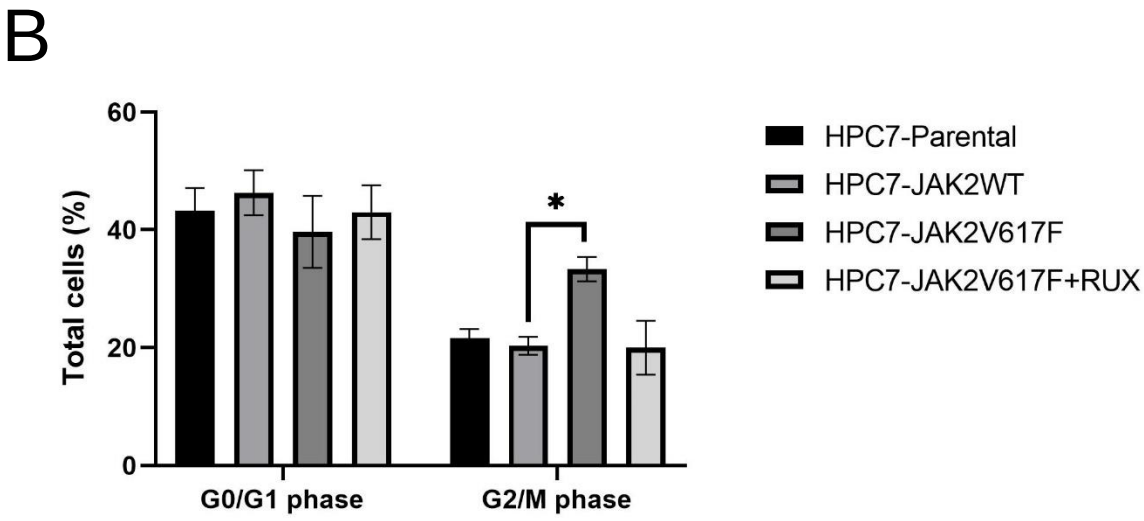


Figure 3.7 JAK2V617F results in more cells in the later cell cycle growth phases.

A) Propidium iodide staining demonstrates that JAK2V617F expression leads to a higher number of cells entering the G2/M phase of the cell cycle. The two red peaks represent G1/S phase on the left and G2/M phase on the right with the S phase represented by diagonally shaded middle part between the two red peaks. The results shown are representative of three different biological replicates B) Percentage of cells in G0/G1 phases and G2/M phases (n=3 biological replicates).

Testing for statistical significance was performed using a student's t-test (*: $p<0.05$; **: $p<0.01$; *** $p<0.001$).

3.3.3 JAK2V617F expression in HPC7s instigates DNA damage and increased reactive oxygen species whilst retaining genomic and cellular stability.

There has been a large amount of work published in recent years investigating the genomic burden of JAK2V617F expression within MPNs. HSC clones expressing JAK2V617F are surprisingly highly cytogenetically stable lasting for decades before any serious disease phenotype arises (Scott, L. M. and Rebel, 2012). The progression of the disease often depends on whether there are any other secondary mutations present such as mutations within the *DNMT3a* gene.

The necessity of a secondary mutation co-presenting alongside JAK2V617F to exhibit a more potent phenotype like acute myeloid leukaemia or primary/secondary myelofibrosis suggests that the presence of JAK2V617F may instigate DNA damage. Despite this, however, the HSC clone has been documented to be extremely stable in patients over decades with JAK2V617F expression/ allelic burden being proportional to the DNA damage impact.

The high allelic burden of JAK2V617F can be caused by loss of heterozygosity (LOH), a key driver of disease progression. LOH was demonstrated by Godfrey et al. to be instigated by overexpression of *JAK2V617F* in HEK293T cells which resulted in large increases in reactive oxygen species (ROS) and abnormal inter-chromosomal homology-directed repair capacity (Godfrey, A. L. et al., 2012). In turn, this resulted in increased DNA damage, secondary mutations and increases in JAK2V617F allelic burden. In 2013, the relationship between

ROS and JAK2V617F expression was also documented in *JAK2V617F* knock-in mouse models by Marty et al. LSK cells from mice expressing JAK2V617F, displayed a significant increase in 8-oxo-guanines (a DNA lesion formed from ROS reacting with guanine) and increases in DNA damage (Marty et al., 2013).

It is possible that JAK2V617F expression provides a protective effect by upregulating ROS buffering pathways (Marty et al., 2013) and DNA damage repair pathways, which may counteract short-term genetic instability. Recently, publications utilising human UT7 cells expressing JAK2V617F have shown an increase DNA damage response pathway with a focus on CHEK1 phosphorylation (Nagao et al., 2011) and Chen et al. demonstrated in BFU-E cells (Burst Forming Units - Erythroid) that expression of JAK2V617F caused high levels of PI3K which in turn resulted in upregulation of *RECQL5*, a DNA helicase, which prevents DNA replication fork stalling (Chen, E. et al., 2015).

In Section 3.3.2, I noted the increase of HPC7-JAK2V617F cells in the G2/M phase of the cell cycle. The G2 phase has many roles and arguably one of the most important is to provide the cell with time to repair nuclear DNA to protect genomic fidelity. The previous data combined with the published data suggest there may be increases in ROS, DNA damage and possibly DNA damage repair mechanisms in HPC7-JAK2V617F cells. I hypothesise that my HPC7-JAK2V617F cells contain higher levels of phospho-γH2AX (serine 139) and phospho-CHEK1 (serine 345), which are phosphorylated by ATM/ATR in the presence of double stranded DNA breaks, and higher levels of ROS.

To validate my hypothesis, I performed immunoblots staining for phospho-γH2AX (serine 139) and P-CHEK1 (serine 345) (**Figure 3.9A**) and measured

ROS by staining freshly plated cell lines with dihydroethidium and incubating for 30 minutes before measuring emission using a Hidex plate reader.

The immunoblot shown in Figure 3.9A demonstrates an increase in both phospho-γH2AX (serine 139) and phospho-CHEK1 (serine 345) suggesting an increase in DNA damage and DNA repair response. However, total CHEK1 levels were upregulated in HPC7-JAK2V617F cells relative to HPC7-JAK2WT, which makes it difficult to conclude whether the increase in phosphorylation represents ATR activity or simply is a consequence of higher total protein levels. The levels of ROS within the HPC7-JAK2V617F cells were also measured (**Figure 3.9B**). This showed that JAK2V617F induces a 1.6-fold increase in ROS levels, thereby identifying a source of DNA damage and potential secondary mutations.

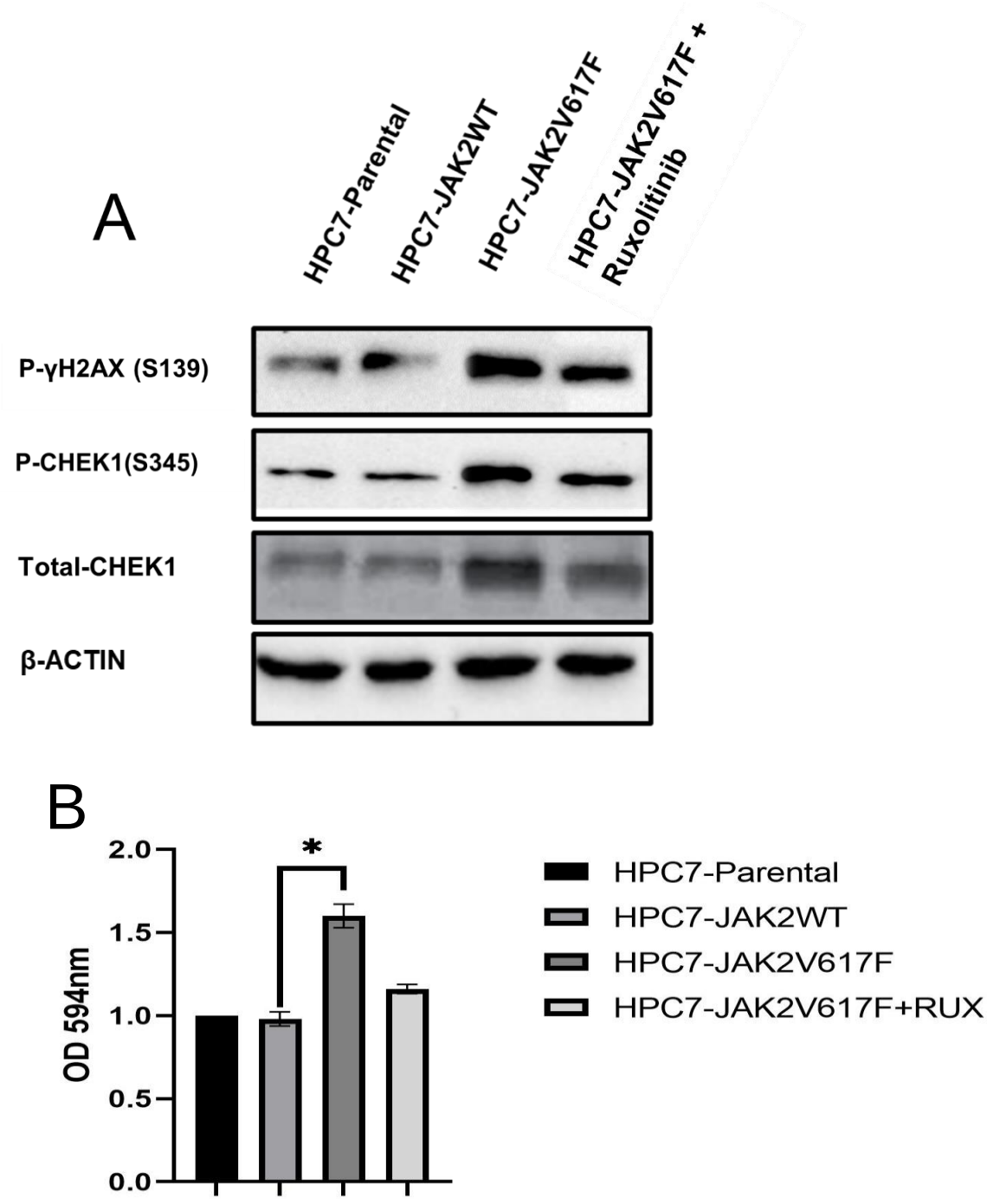


Figure 3.8 Expression of JAK2V617F increases DNA damage and levels of ROS.

(A) HPC7-JAK2V617F has a higher level of phospho-γH2AX and phospho-CHEK1 which suggests JAK2V617F causes DNA damage and upregulates the DNA damage repair protein CHEK1. (B) JAK2V617F increases levels of ROS which may drive or at least contribute to the increases in DNA damage seen in figure A.

Testing for statistical significance was performed using a student's t-test (*: $p<0.05$; **: $p<0.01$; *** $p<0.001$). Three biological replicates were performed.

As mentioned previously, HSCs expressing JAK2V617F are relatively cytogenetically stable despite increases in DNA damage and ROS levels. In my HPC7-JAK2V617F cell line both DNA damage and DNA repair appear to be increased (**Figure 3.8**). I wanted to investigate if this increase in DNA damage translates into increased levels of apoptosis or if my cell line model has levels of genome stability consistent with that which has been published (**Figure 3.9)**. As can be seen in Figure 3.9, no difference in apoptosis is detectable between samples which corroborates that this is a cytogenetically stable cell line.

Various publications have presented evidence that the expression of JAK2V617F triggers higher levels of DNA damage repair (Nieborowska-Skorska et al., 2017; Ueda et al., 2013). Therefore, I hypothesise that within my HPC7-JAK2V617F cell line there is a higher-level DNA damage repair in response to genotoxic insults. To test DNA damage repair capacity, I used three different methods: 1) A Trypan Blue exclusion assay measuring cell viability after treatment with either Mitomycin C or Olaparib. Mitomycin C is a potent cross-linker which leads to replication fork-stalling; Olaparib is a PARP1 inhibitor and both drugs lead to increased DNA damage (**Figure 3.10A+B**). 2) Apply UV radiation (254 nm) to damage the HPC7-JAK2WT/JAK2V617F DNA and measure cell viability via the Trypan Blue exclusion assay post 24 hours (**Figure 3.11A**). 3) Apply UV damage to an eGFP vector which is transduced into the HPC7-JAK2WT/V617F cell lines and quantifying fluorescence (**Figure 3.11B)**. If the JAK2V617F has a greater repair capacity, then more eGFP will be repaired resulting in increased emission.

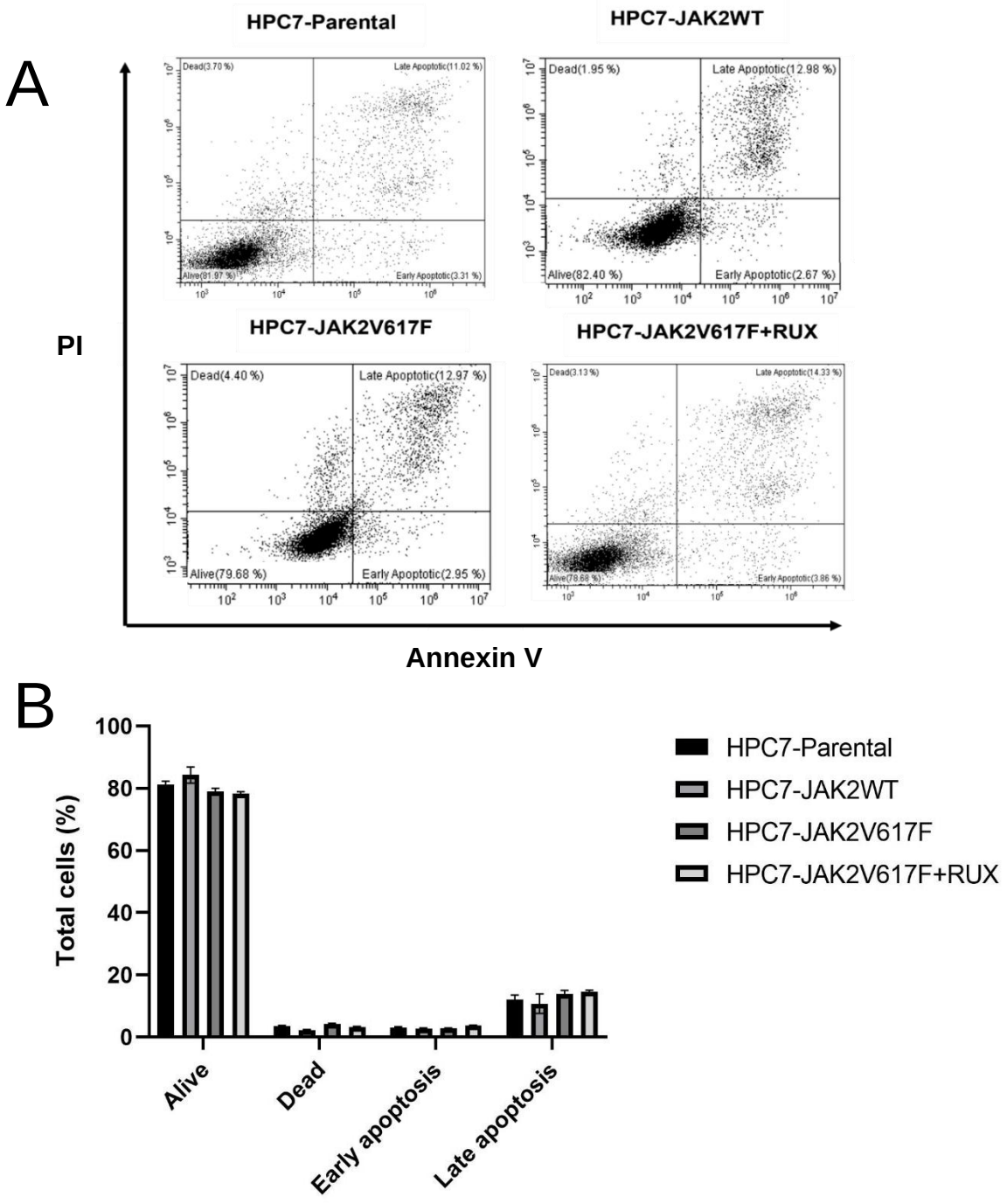


Figure 3.9 JAK2V617F expression in HPC7 cell line has no impact on levels of apoptosis.

A) Flow cytometric analysis of HPC7-JAK2V617F cells with propidium iodide and annexin V indicates that JAK2V617F expression does not reduce cell viability despite increasing DNA damage. This image is the most representative of the average. B) A graph summarising the percentage of total cells in each category.

Testing for statistical significance was performed using a student's t-test (*: $p<0.05$; **: $p<0.01$; *** $p<0.001$).

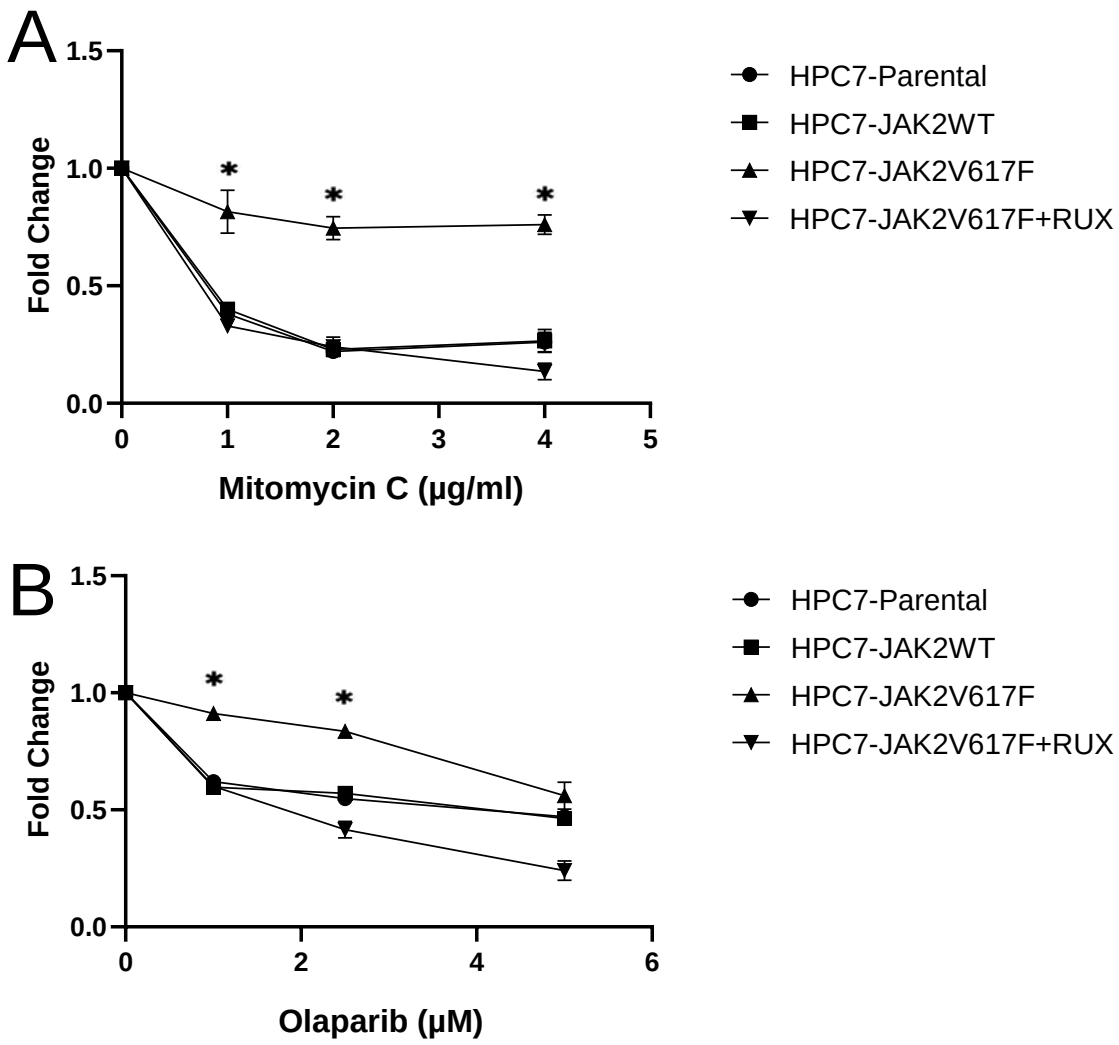


Figure 3.10 – JAK2V617F expression confers protection from mitomycin C and Olaparib represented by increasing cell survival.

Drug treatment assays in which parental/JAK2WT/JAK2V617F/ Ruxolitinib controls are subjected to either Mitomycin C treatment for 24 hours or Olaparib for 96 hours at increasing concentrations. Post-treatment, cell survival is measured by performing an MTT assay.

The JAK2V617F mutation confers a greater resistance to 1-4 µg/ml concentration of Mitomycin C (A) and 1 and 2.5 µM concentration of Olaparib (B) treatment suggesting that JAK2V617F upregulates genes that can overcome the fork stalling, a reduction in base excision repair or the DNA damage caused by it.

Testing for statistical significance was performed using a student's t-test (*: $p<0.05$; **: $p<0.01$; *** $p<0.001$). Three biological replicates were performed.

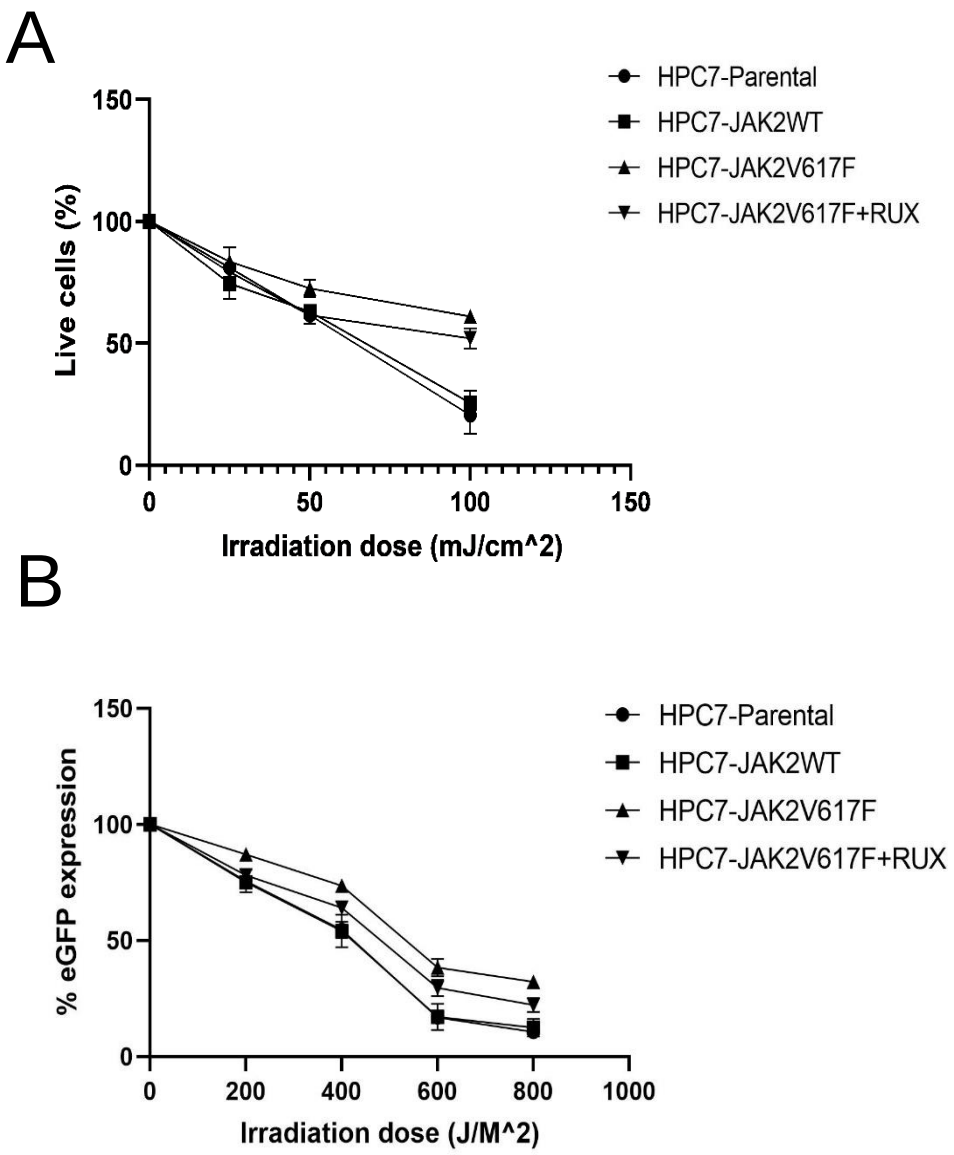


Figure 3.11 – JAK2V617F expression may confer protection from UV-driven DNA damage.

UV treatment assays which either target genomic DNA directly or the eGFP vector that was subsequently transfected.

(A) JAK2 mutation may grant a minor resistance to direct UV radiation stimulus allowing for a higher percentage of HPC7 cells to survive. (B) JAK2V617F does not increase the capability to repair the eGFP vector.

In Figure 3.10B, expression of JAK2V617F in HPC7 cells does not provide significant resistance to high concentrations of Olaparib (5 µM) but does provide resistance to lower concentrations (1 µM and 2.5 µM), relative to HPC7-JAK2WT cells. This suggests that expression of JAK2V617F in HPC7s can provide a minor resistance to Olaparib treatment in comparison with HPC7-JAK2WT cells.

The level of both resistance to DNA damage and repair suggest that JAK2V617F-driven upregulations of the signalling pathways may have a protective effect against genomic instability; however, additional experimentation is needed to confirm this, since Ruxolitinib did not abrogate the effect of JAK2V617F in Figure 3.11.

3.3.4 HPC7-JAK2V617F cells do not display increases in R-loops but do display increases in the expression of R-loop resolving genes.

R-loops are DNA: RNA hybrids formed during transcription as a complex between the newly transcribed RNA moiety and one strand of the genomic DNA. An R-loop forms when RNA is transcribed, spliced, and then subsequently hybridises with genomic DNA. When an R-loop forms the RNA sections complementary to the DNA strand's introns are spliced out and the single-stranded intronic genomic DNA projects outwards forming a loop-like structure. The R-loop formation can leave the loop of single-stranded intronic DNA, and the other strand of genomic DNA unprotected, increasing the exposure of this DNA to environmental damage-causing stimulants such as ROS or UV radiation. R-loops are believed to contribute to DNA damage/genomic instability in cancer and malignancies in general (Skourti-Stathaki and Proudfoot, 2014).

As a final investigation into DNA damage in JAK2V617F-expressing cells I wanted to pursue the following objectives: 1) I wanted to analyse the expression of genes that have been published to contribute to the observed genetic stability provided by the JAK2 mutant, namely *RECQL5, RAD51, BRCA1* and *LIG4* to confirm if these genes could contribute to the observed increased DNA repair phenotype (**Figure 3.12A**) and 2) I wanted to quantify the level of R-loops in HPC7-JAK2V617F cells via a dot-blot using isolated genomic DNA (**Figure 3.12B**).

The RT-qPCR of *RECQL5, RAD51, BRCA1* and *LIG4* yielded mixed results with *RAD51* and *BRCA1* being moderately upregulated within the HPC7-JAK2V617F cells relative to the wild-type. By contrast, *LIG4* and *RECQL5*

expression levels were unchanged (**Figure 3.12A**). Previous publications examining changes in expression of these genes were performed in BAF/3 cells (a pro-murine B cell line), which may explain the difference in results.

The dot-blot in **Figure 3.12B** suggests no significant changes in R-loop quantity between the mutant and the wild-type which may suggest that either JAK2V617F has upregulated genes to quickly resolve these RNA: DNA structures or there is no difference. To investigate possible high turnover of R-loops, I performed RT-qPCR reactions to examine the expression level of *RNASEH1* and *RNASEH2*, which are the primary resolving enzymes of R-loops (Skourti-Stathaki and Proudfoot, 2014) (**Figure 3.12C**) together with *BRCA1* (**Figure 3.12A**).

The RT-qPCRs for *RNAseH1* and *RNAseH2* show negligible difference between their expression levels. However, *BRCA1/2* is moderately upregulated in HPC-JAK2V617F which could explain a possible high turnover of R-loops.

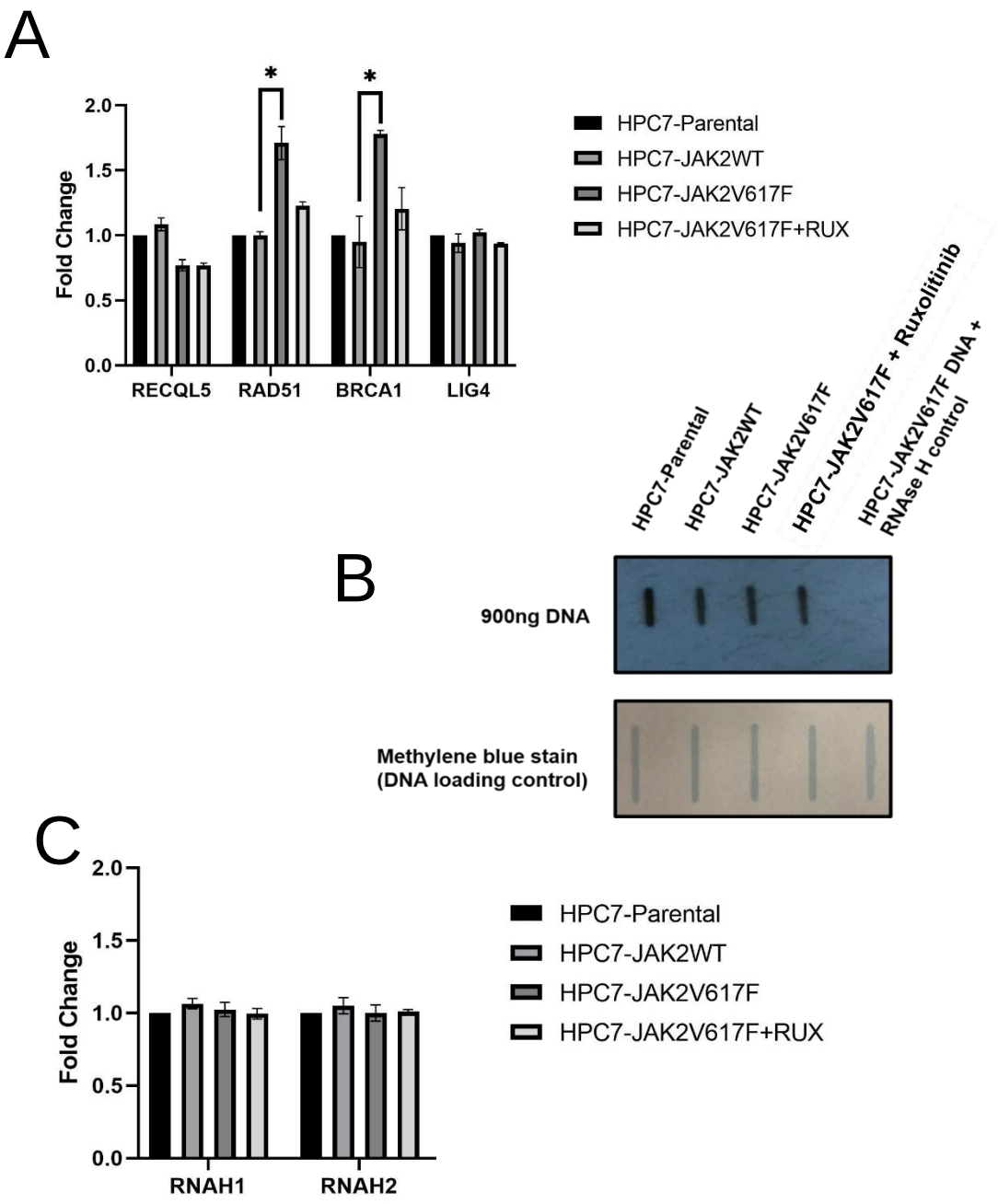


Figure 3.12 JAK2V617F upregulates some homology-directed repair associated genes but has limited influence on R-loops.

(A) RT-qPCR investigating levels of expression of DNA damage repair genes indicates that *BRCA1* and *RAD51* are moderately upregulated in JAK2V617F expressing HPC7 cells however *LIG4* and *RECQL5* are unchanged. (B) Dot-blot for R-loops suggests JAK2V617F does not contribute to R-loop formation or (C) upregulate genes associated with high R-loop turnover.

For A and C, Testing for statistical significance was performed using a student's t-test (*: $p<0.05$; **: $p<0.01$; *** $p<0.001$). Three biological replicates were performed.

Discussion 3.4

The influence of JAK2V617F on HSCs is a widely debated topic. In recent years, there has been a clear focus on JAK2V617F and DNA damage in an effort to understand the genetic burden of JAK2V617F, how JAK2V617F contributes to disease phenotype and more importantly disease progression. In this section, I attempted to create my own disease model by overexpressing the JAK2V617F mutant in the HPC7 cell line.

The data collected in Section 3.3 provides sufficient evidence that JAK2V617F is both being expressed within the HPC7 cell line and is producing a significant activation of cellular signalling pathways, cell proliferation and instigating both DNA damage and its repair.

JAK2 is a signal transduction kinase. MPNs with JAK2V617F are essentially a disease of chronic downstream signalling driven by an aberrant kinase. In my model, I successfully showed that STAT, AKT and ERK1/2 proteins were all highly phosphorylated compared to the controls which is something that has been verified in mouse models, cell lines and patient data. Interestingly, the phosphorylation status of STATs in patients varies between the diseases with ET showing a high phospho-STAT3 and low phospho-STAT5, MF demonstrating low amounts of both phospho-STATs and PV showing high levels of both phospho-STATs. MPNs are differentiated by JAK2V617F expression which indicates that my HPC7-JAK2V617F cells represent a phenotype similar to that of a PV phenotype.

I decided to measure the level of c-MYC protein in my HPC7-JAK2V617F model and found that c-MYC is downregulated relative to HPC7-JAK2WT (**Figure 3.5**) which parallels the data observed in *c-MYC* knockouts by Wilson,. 2004. In the Wilson et al. publication, downregulated expression of *c-MYC* in HSCs drives HSC expansion whilst polarising the HSCs away from quiescence. In some mouse models, conditional expression of JAK2V617F leads to loss of terminal differentiation potential and eventually a loss of HSCs (Wilson et al., 2004). I hypothesise that *c-MYC*, perhaps alongside other regulators, may drive HSC proliferation which could possibly drive the HSC exhaustion observed in the Li et al.'s published model (Li, J. et al., 2014). A likely reason why I do not see cellular exhaustion or differentiation within my model may be caused by MSCF dependence, the overexpression of *LH2* in the cell line, or even the cell line's erythroid nature. A further investigation of c-MYC in my model could entail overexpressing c-MYC in the HPC7-JAK2V617F cells using lower MSCF concentrations and measuring if the rate of cell proliferation returns to HPC7-JAK2WT levels.

The c-KIT receptor can drive self-renewal/proliferation, survival, and a wide range of other cellular processes in response to SCF (Liang et al., 2013) and is a type III receptor tyrosine kinase composed of homodimer subunits. Upon ligand binding, the two c-KIT monomers to migrate into close proximity, forming a homodimer and allowing transphosphorylation and activation of the signalling pathway (Liang et al., 2013).

It has been shown that SCF binding to c-KIT activates the self-renewal pathways, by recruiting JAK kinases which activate the STAT3/ERK/AKT pathways (Liang et al., 2013). I predicted that JAK2V617F expression may

confer MSCF independence through c-KIT the receptor. In my HPC7-JAK2V617F cell line, I observed no such independence (**Figure 3.5**). I did however find that when cultured at lower concentrations of MSCF, the JAK2V617F cell line had a higher proliferation capacity than the HPC7-JAK2WT cell line (**Figure 3.6**). To explain this, I theorise two possibilities, 1) JAK2V617F binds other receptors such as MPL or EGFR which both have published roles in the signalling which trigger proliferation. 2) JAK2V617F potentiates c-KIT signalling through an unknown mechanism.

JAK2V617F increases the proliferative capacity in HPC7 cells relative to the JAK2WT cell line. The increase in proliferation would favour the Akada et al. mouse model of increased LSK propagation against Mullaly/Li and colleague's models. I predict that this result could be explained by the HPC7 cell line being an inherently erythroid cell line that is similar to HSCs rather than a true HSC. The decreased expression of c-MYC and its relationship to HSC exhaustion and expansion should be looked into further, especially in the above mouse models as it may explain the HSC-JAK2V617F exhaustion and polarisation towards differentiation noted in these models. This finding should be further investigated to fully elucidate how c-MYC expression and MSCF contributes to mutant HSC clone expansion, exhaustion and to the larger MPN progression picture.

It has been shown in patients that JAK2V617F can linger within HSCs for decades without any noticeable negative phenotype suggesting a highly genetically stable mutant clone (Gale et al., 2007). The data in Section 3.3.3 clearly suggest that JAK2V617F confers an advantageous DNA damage repair ability whilst itself seemingly instigating DNA damage.

From **Figure 3.8**, we can clearly see an increase in double stranded DNA breaks represented by phospho–γH2AX (S139) and an increase in the DNA repair pathway protein CHEK1. CHEK1 functions by integrating ATM and ATR signalling, causing CHEK1 to phosphorylate CDC25A, which in turn triggers inhibition of the cell cycle which confines cells in the G2/M phase. I suggest that increase in phospho-CHEK1 may contribute to the increased number of HPC7-JAK2V617F cells within G2/M phase of the cell cycle.

The methods by which JAK2V617F can provide genetic stability against DNA damage are still under investigation. In my model I found that the HPC7-JAK2V617F cell line had a greater resistance towards mitomycin C than the HPC7-JAK2WT. Mitomycin C is a potent crosslinker of DNA preventing DNA synthesis, and to combat the effect of Mitomycin C embryonic stem cells can upregulate genes of the homology-directed repair pathway to repair the damage (Plo et al., 2008). This data presented in Figure 3.10A suggests that HPC7-JAK2V617F has a superior resistance to Mitomycin C relative to HPC7-JAK2WT cells. When this result is applied to the patient it may explain a possible clonal advantage regarding the toleration of DNA damage.

R-loops are a relatively new and interesting topic. Recent papers have shown that oncogenes such as *KRAS* can contribute to an increase in R-loops which consequently cause double-stranded DNA breaks (Kotsantis et al., 2018). In my model cell line, no increase in R-loops was observed. This is unexpected as R-loops are often observed in highly transcriptionally active cells such as those with increased phospho-STATs (Kotsantis et al., 2018). This incongruity could be explained by the increases seen in both *BRCA1* which has been reported to resolve R-loops or it could even be explained by method of detection as the dot

blot may not be sensitive enough to detect any difference. Another possibility is that perhaps that the genomic content of R-loops has not changed but rather that the genomic distribution has changed. This change of distribution could be identified using a DRIP-sequencing experiment which would highlight these genomic locations.

The HPC7-JAK2V617F model has both disagreed and agreed with mouse models and patient data, however overall, my model is robust and provides a sound basis for continual investigation into phospho-signalling and DNA damage/repair. For further investigation, I decided to perform RNA-sequencing on the HPC7-JAK2V617F and HPC7-JAK2WT cell lines to further investigate the changes observed in my data.

Chapter 4

JAK2V617F expression in the HPC7 cell line dysregulates global gene expression comparable to Polycythaemia Vera patients and reveals HDAC9 as a target for further investigation.

Chapter 4 JAK2V617F expression in the HPC7 cell line dysregulates global gene expression comparable to Polycythaemia Vera patients and reveals HDAC9 as a target for further investigation.

4.1 Introduction

RNA-sequencing was developed during the late 2000s (Stark et al., 2019) and since then it has become ubiquitously used in all forms of biological research to investigate differential gene expression (DGE) between cells expressing cancer-related mutations and cells exposed to genotoxic conditions providing an insight into how genes are regulated in response to such conditions. The benefit of RNA-sequencing cannot be understated. In this introduction I will outline a basic RNA-sequencing workflow and briefly, what I hope to achieve with this chapter.

RNA-sequencing workflow alters depending on various factors; therefore, I will present a workflow that matches the one used in this thesis, Illumina NGS (https://emea.illumina.com/techniques/sequencing/ngs-library-prep.html). RNA-sequencing begins in the lab by isolating and extracting total RNA from the cells being investigated. The mRNA content of RNA isolation is around 1-5% with the remaining 95-99% being ribosomal RNA (rRNA) (80-85%) and transfer RNA (tRNA) (15-20%). Therefore, to produce reliable data it is essential to enrich the mRNA or deplete ribosomal RNA.

The process of mRNA enrichment involves removal tRNA and rRNA from the samples through utilising the 20-300 nucleotide 3’ long poly(A) tail of the mRNA which hybridises with oligo dT molecules on a carrier, such as a cellulose matrix

or biotinylated oligo dT probe. The samples are washed to remove unbound RNA (rRNA and tRNA) and then displaced from the oligo dT carrier with a low-salt buffer providing a mostly pure mRNA content. This method, however, is limited to eukaryotes. rRNA depletion/capture was developed for prokaryotes to supplement the lack of a long 3' poly(A) tail found on prokaryotic mRNA. RRNA capture involves the use of probes attached to metal beads that bind the 16S and 23S regions of ribosomes which can then be removed.

The resulting mRNA content post mRNA enrichment or rRNA capture is very low therefore the next step is fragmenting and amplifying mRNA through the use of reverse transcriptase enzymes which convert mRNA into cDNA (with the sequence complementary to the original RNA) which allows for amplification through PCR using non-specific primers, usually random hexamers with a tag. Once the cDNA has been amplified, the starting RNA is removed, and the cDNA is annealed to a terminal-tagging oligo (TTO) to block synthesis of the 3' end.

The resulting TTO cDNA is then once again amplified via PCR to form DNA containing the sequence of the original mRNA. The tagged DNA is then purified for the next step which is adaptor ligation. Adaptor ligation is the process of attaching known synthetic oligonucleotides to both or one end of the cDNA fragments. This acts a readable motif allowing computer programmes to distinguish the beginning of the DNA fragment read.

For Illumina sequencing, amplification begins with cDNA being attached to a flow cell. A flow cell is a core sequencing reaction vessel where sequencing takes place and a channel that binds mobile DNA fragments. In the flow cell

channel are oligonucleotide adaptors which bind complementary adaptors ligated on the DNA fragments in the previous stage. In the flow cell, cDNA clusters are generated through cycles of bridge amplification and denaturing, and sequence-by-synthesis reactions are performed in cycles of complementary strand synthesis and laser excitation of bases with reversible nucleotide terminators (Illumina, 2014).

The next part of RNA-sequencing is purely computational. The raw input data of reads is processed removing the adaptor sequence from the read which is then mapped to the relevant genome and overlapping fragments are assembled. The transcripts are counted, attached to relevant information i.e., gene identifier and are ready for statistical analysis.

RNA-sequencing has become widely used within molecular biology, providing an insight into possible gene targets for therapeutic intervention and has significantly contributed to the understanding of haematological malignancies. JAK2V617F is present in three MPNs, all of which have a different transcriptional profile and in addition to this, individual patients with the same MPN often have high variability between their gene expression profiles which is determined by factors such as age and gender.

JAK2V617F is rarely present alone in the progressive forms of MPNs and in fact multiple other mutations can be found alongside it; these mutations are regularly found in epigenetic regulators and often tend to be loss of function mutations resulting in aberrant distribution of histone marks like H3K27 methylation and DNA methylation. Histone marks are often found to correlate with the transcriptional output of gene expression therefore RNA-sequencing is an invaluable tool in highlighting DGE.

In this chapter I aim to investigate how similar my data is to published data to gauge the relevance of my HPC7-JAK2V617F cell line to a clinical setting and which of the three JAK2V617F-related MPNs my model represents. I predict my model represents Polycythaemia Vera because my model has high amounts of phosphorylated-STAT5 and phosphorylated-STAT3. I decided to co-validate my data utilising RT-PCR measuring the top 20 and bottom 20 dysregulated genes identified in the HPC7-JAK2V617F cell line, relative to that of the HPC7-JAK2WT cell line.

I have utilised Gene Set Enrichment Analysis (GSEA), a bioinformatic programme, to analyse which pathways are upregulated in my data and whether the same pathways are shared by MPN patients and if the same genes are contributing to these pathways.

Lastly, I will focus on HDAC9 expression with a view to investigate if HDAC9 upregulation is a feature shared between other HSC models and clinical data, briefly exploring the effect of secondary mutations on HDAC9 expression.

4.2 Aims and Hypothesis

In Chapter 3 I showed via a western blot that phosphorylation of STAT1/3/5 is significantly increased in the presence of the JAK2V617F mutation, STATs are transcription factors which upon phosphorylation, dimerise and translocate to the nucleus driving transcription of genes.

In my HPC7-JAK2V617F cell lines, I observed greater phospho-STAT signalling and increased DNA damage. STATs are transcription factors; therefore, I hypothesise that JAK2V617F may alter the transcriptional profile of the HPC7 cell line to up or downregulate genes which allow these observed differences to occur.

Aims:

1. I will discuss my data and use it to try and explain the results of Chapter 3 and I will provide a comprehensive validation of the dataset using RT-PCR. I will compare values generated by image densitometry with the fold change generated by the RNA-sequencing to understand the degrees of agreement between the two values.

2. I will compare my data with other published data to gauge the relevance of my data comparing both the genes and the pathways involved. I will identify a dysregulated gene commonly found in these datasets.

3. I will highlight HDAC9 expression in clinical models and its expression in patient data, I will also briefly explore if secondary mutations effect HDAC9 expression.

Chapter 4.3 Results

4.3.1 JAK2V617F instigates global transcriptional dysregulation in the HPC7 cell line

The results of Chapter 3 strongly suggest that transcriptional changes occur within the HPC7-JAK2V617F cells because of the increase in STAT phosphorylation relative to the HPC7-JAK2WT cell lines. *STAT* genes encode transcription factors which when activated dimerise, translocate to the nucleus, and regulate transcription of genes. To visualise these changes, I extracted the total RNA of HPC7-JAK2WT and HPC7-JAK2V617F cells and sent these for RNA-sequencing. Unfortunately, I could not send replicates, and this greatly affects the gene expression analysis that can be performed. The data presented on my samples cannot be probed statistically therefore it will be used as a transcriptomic profile impression and that is why later on in this section I will provide data that validates the RNA-sequencing value generated and will compare it to other published data.

Initially, I took the raw data files and processed them utilising high-performance computing (see Figure 2.1 of methods and materials). A normal RNA-sequencing workflow requires replicates to generate an output and to circumvent this, I used NOISeq (Tarazona et al., 2015). NOISeq is a programme that can export the output to an Excel file without the need for replicates. The data was annotated through the Ensembl BioMart programme (Howe et al., 2021). The RNA-sequencing data revealed that HPC7-JAK2V617F had many upregulated genes and after curating the data (method of curation outlined below) the RNA-

sequencing data suggested that 385 upregulated genes and 390 downregulated genes relative to the HPC7-JAK2WT.

The numbers of genes that are dysregulated is likely to be far more than previously mentioned because to overcome the challenge that a lack of biological replicates imposed, I decided to remove “extreme” values from the data. I removed these “extreme” values by implementing a filter to remove any gene with less than 200 reads mapping; there was no limit to the maximum number of reads but any values where the HPC7-JAK2WT and HPC7-JAK2V617F were dramatically different and had a log((fold change) greater than 4)) were temporarily excluded from the data until being investigated by RT-PCR analysis on a gel, after which they were included. The top and bottom 25 dysregulated genes are shown in **Table 4**.

Top 25 Upregulated Genes			
Gene ID	Gene Symbol	Log (2) Fold Change	KEGG Pathway
56744	*PF4*	4.371577507	Chemokine Signaling Pathway
14165	*FGF10*	3.81056181	Bladder Cancer
13380	*DKK1*	3.484086406	Alzheimer Disease
68659	*FAM198B*	3.450923719	Adherens junction
18612	*ETV4*	3.443693509	Transcriptional Misregulation In Cancer
67951	*TUBB6*	3.354973964	Gap Junction
107684	*CORO2A*	3.05863199	N/A
65964	*MAP3K20*	3.024633629	MAPK Signalling
18019	*NFATC2*	3.003870771	cGMP-PKG Cellular Signalling
12515	*CD69*	2.987563099	IL3 Signalling
104156	*ETV5*	2.977904315	Transcriptional Misregulation in Cancer
319520	*AXL*	2.916981228	EGFR Tyrosine Inhibitor Resistance
320292	*RASGEF1B*	2.627483318	N/A
12578	*CDKN2A*	2.60023272	Chronic Myeloid Leukaemia
22021	*TPST1*	2.478457018	N/A
13641	*EFNB1*	2.44971004	Axon guidance
56421	*PFKP*	2.436903515	AMPK Signalling
19703	*RENBP*	2.422320272	Metabolic Pathway
64011	*NRGN*	2.383694382	N/A
20893	*BHLHE40*	2.233528338	Circadian Rhythm
12444	*CCND2*	2.21571689	Cell Cycle
12700	*LILR4B*	2.158334343	N/A
383619	*AIM2*	2.145939926	NOD-like Receptor Signalling
79221	*HDAC9*	1.960324462	Carcinogenesis

Top 25 Downregulated Genes			
Gene ID	Gene Symbol	Log (2) Fold Change	KEGG Pathway
102103	*MTUS1*	-4.375692125	N/A
14255	*FLT3*	-3.89055567	Acute Myeloid Leukemia
11488	*ADAM11*	-3.829776098	N/A
14924	*MAGI1*	-3.781299464	RAP1 Signaling
12606	*CEBPA*	-3.753704234	Acute Myeloid Leukemia
19152	*DUSP4*	-3.739543626	MAPK Signaling
227357	*ESPNL*	-3.587314954	N/A
16523	*MPO*	-3.350834958	Transcriptional Mis regulation in Cancer
208618	*ETL4*	-3.182102351	N/A
12289	*CACNA1D*	-3.181617445	Alzheimer Disease
97998	*DEPTOR*	-3.090336544	mTOR Signaling Pathway
22177	*TYROBP*	-3.054416974	Osteoclast differentiation
332579	*CARD9*	-2.999760279	C-type lectin receptor signaling pathway
52552	*PARP8*	-2.93686206	N/A
232174	*CYP26B1*	-2.863617335	Metabolic Pathway
216152	*PLPPR3*	-2.810378304	N/A
217328	*MYO15B*	-2.797673049	N/A
380172	*MILR1*	-2.789147563	N/A
27428	*SHROOM3*	-2.748313847	N/A
218820	*ZFP503*	-2.711524879	N/A
214855	*ARID5A*	-2.696733189	N/A
210741	*KCNK12*	-2.688255592	N/A
16409	*ITGAM*	-2.618462124	Acute Myeloid Leukemia
12873	*CPA3*	-2.57816498	Renin-Angiotensin System

Table 4 JAK2V617F expression in HPC7 cell line dysregulates gene expression.

Top table represents a list of the top 25 up-regulated genes in HPC7-JAK2V617F relative to HPC7-JAK2WT and the bottom table represents the top 25 downregulated genes.

To understand how the transcriptomic profile could be contributing to the results seen in Chapter 3, I input the data from the RNA-sequencing into Gene Set Enrichment Analysis software (GSEA) (Subramanian et al., 2005; Mootha et al., 2003). GSEA groups genes together based on the pathway to which the genes contribute. Usually, GSEA requires biological replicates to calculate a P-value; to circumvent this I used the pre-ranked analysis function of GSEA which ranks the genes based on the value given by the user, for this work; I used the log_2(fold change) of HPC7-JAK2V617F compared to HPC7-JAK2WT.

I grouped the genes based on Kyoto Encyclopaedia of Genes and Genomes (KEGG), Reactome, C2 curated and Hallmark grouping pathways. When performing GSEA, I used multiple categorical group classifications (KEGG etc) as each categorical group defines a pathway through expression of specific genes. The category is curated by the team that manages GSEA on the basis of what group of genes are dysregulated in sequencing data submitted by other lab groups. The basing of categorical groups on other lab groups data can translate into bias as each cell line or clinical model will have variation in the genes being expressed.

The GSEA of my RNA-sequencing dataset provided over 100 differentially activated pathways and therefore I have curated the data to specifically layout the relevant pathways. Firstly, I analysed the results of the GSEA to specifically focus on pathways that should be upregulated in the HPC7-JAK2V617F cells based on Chapter 3. Hallmark and C2 Curated grouping reveal that STAT1, STAT3, STAT5, MAPK and ERK1/2 targets are upregulated which corroborates

the results from Chapter 3 and supports that JAK2-targets are being phosphorylated and acting as transcription factors or activating downstream targets that act as transcription factors (**Figure 4.1**).

In Chapter 3 I suggested that HPC7-JAK2V617F had a greater DNA damage repair capacity relative to HPC7-JAK2WT under external genotoxic stress, increased ROS levels and an increased number of cells in the G2 phase of the cell cycle. In the Hallmark group GSEA dataset, I find that HPC7-JAK2V617F cells have an increased expression of genes associated with the: DNA repair pathway, G2 to M transition, oxidative phosphorylation, and increased response to UV; this can be built upon further by the KEGG dataset which highlights increase genes involved in homologous recombination, NHEJ pathway and base-excision repair. A full list of the GSEA results is provided in an appendix.

I further explored the datasets and highlighted pathways and the genes involved with the highest rank (log_2(fold change)) (**Figure 4.2**). From the cell cycle analysis performed in Chapter 3 it was implied that a greater number of HPC7-JAK2V617F cells were present in the G2 to M phase of the cell cycle relative to HPC7-JAK2WT cells which is further corroborated by the GSEA analysis. GSEA analysis highlights 179 genes with possible involvement in this pathway with the cyclin A and cyclin D protein family being present alongside CHEK1 and other factors involved with DNA repair.

In Chapter 3, Section 3.3, I investigated DNA damage repair by testing levels of ROS, and homologous recombination capacity. The oxidative phosphorylation and

glycolysis analysis revealed an increased expression of 167 and 121 genes respectively (**Figure 4.2**); this increase suggests that HPC7-JAK2V617F cells have a much greater energy demand relative to HPC7-JAK2WT, and this may explain the 1.6-fold increase in ROS levels that I observed in Chapter 3. The increased energy demand by HPC7-JAK2V617F has been observed before in primary haematopoietic progenitor cells expressing JAK2V617F and it has been suggested that JAK2V617F expression reprogrammes the metabolic pathways to favour a high-glucose breakdown (Rao et al., 2019) which may explain the increased ROS.

Further investigation into the genes dysregulated highlighted the UV response, homologous recombination, and DNA damage repair pathways. This GSEA reveals that 101 genes associated with the UV response are elevated in the HPC7-JAK2V617F cells (**Figure 4.3**). The top 10 genes in the UV response category however have not yet been investigated and the genes identified have not been investigated in MPNs yet. DNA damage has been reported in both the literature (Nieborowska-Skorska et al., 2017; Chen, E. et al., 2015; Ueda et al., 2013) and this thesis to be high in JAK2V617F expressing cells; DNA damage is thought to be caused by high ROS levels and stalled replication forks (Chen, E. et al., 2014) and is a therapeutic target for MPN therapy. The genes highlighted in the DNA repair group are mostly characterised as protein adaptors and therefore a further investigation into them may add to the current therapies of inhibiting PARP1 (Nieborowska-Skorska et al., 2017).

Category_PATHWAY	Number of Genes involved
HALLMARK_TNFA_SIGNALING_VIA_NFKB	121
HALLMARK_IL2_STAT5_SIGNALING	115
HALLMARK_INTERFERON_GAMMA_RESPONSE	115
WIERENGA_STAT5A_TARGETS_UP	99
HALLMARK_INTERFERON_ALPHA_RESPONSE	61
AZARE_NEOPLASTIC_TRANSFORMATION_BY_STAT3_UP	51
PID_PI3KCI_AKT_PATHWAY	33
REACTOME_MAP2K_AND_MAPK_ACTIVATION	32
HALLMARK_IL6_JAK_STAT3_SIGNALLING	29
WP_PI3KAKTMTOR_SIGNALING_PATHWAY	27

AKT- related pathways
MAPK-related pathways
STAT1-related pathways
STAT3-related pathways
STAT5-related pathways

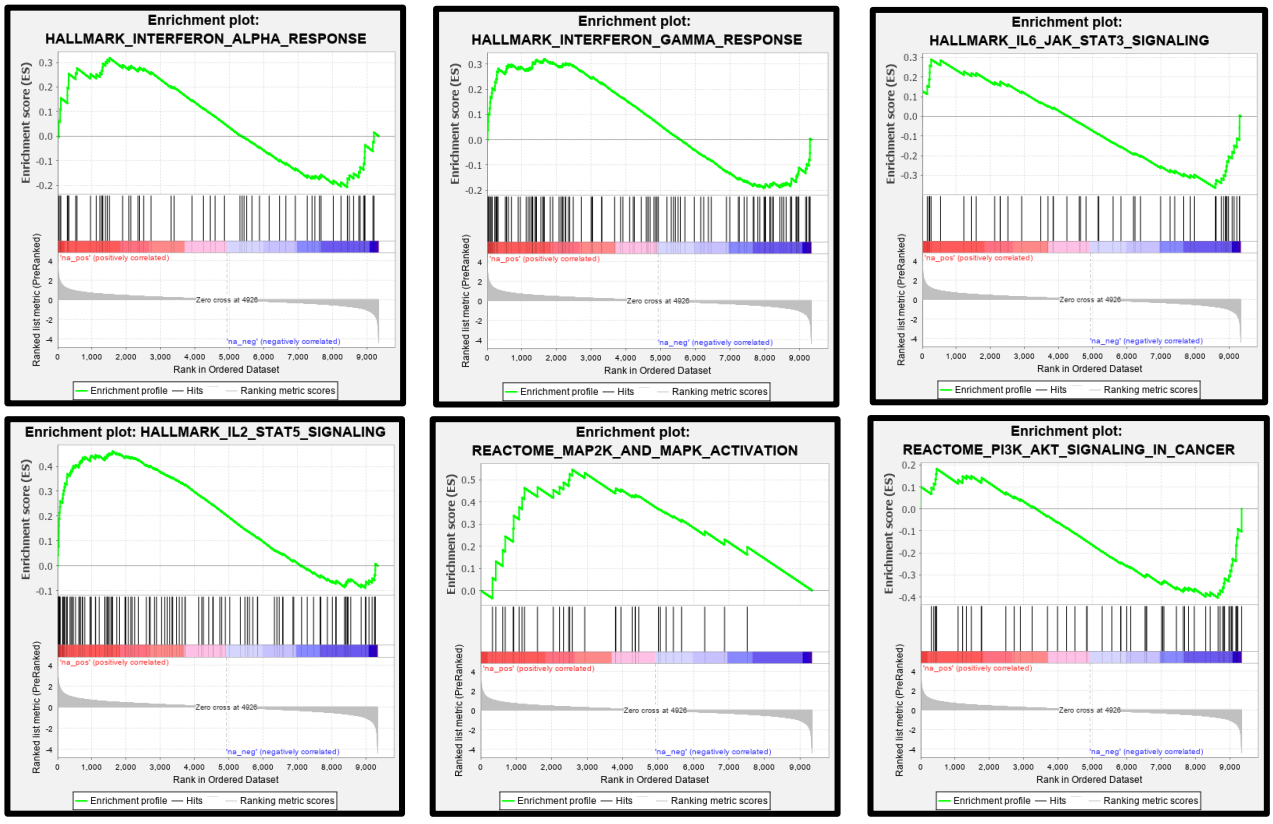
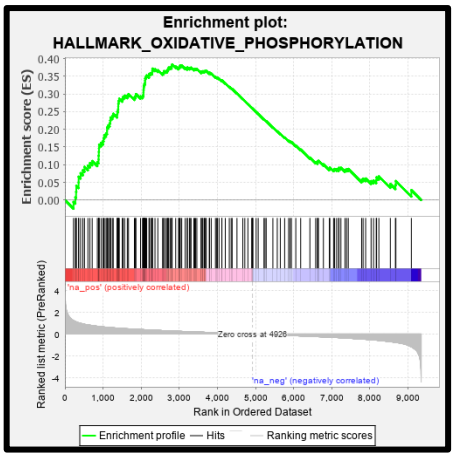


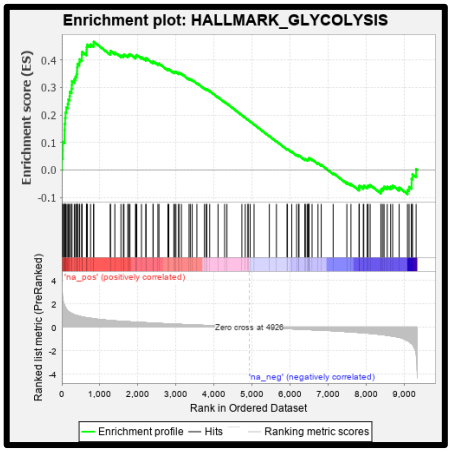
Figure 4.1 GSEA of the HPC7-JAK2V617F cell line validates an increase JAK signalling phenotype.

GSEA of HPC7-JAK2V617F cells implies that the increased phosphorylation of signalling factors observed in the western blots in Chapter 3 has translated into upregulation of relevant pathways. Interferon alpha and interferon gamma pathways are indicative of STAT1 activation.

HALLMARK_PATHWAY	Number of genes involved
HALLMARK_OXIDATIVE PHOSPHORYLATION	167
HALLMARK_GLYCOLYSIS	121



Upregulated Gene	Log2(Fold Change)
HSD17B10	1.55
COX8A	0.6
COX6A1	1.02
LDHA	0.98
COX7B	0.93
SLC25A12	0.89
COX7A2	0.83
COX6C	0.54
VDAC1	0.52

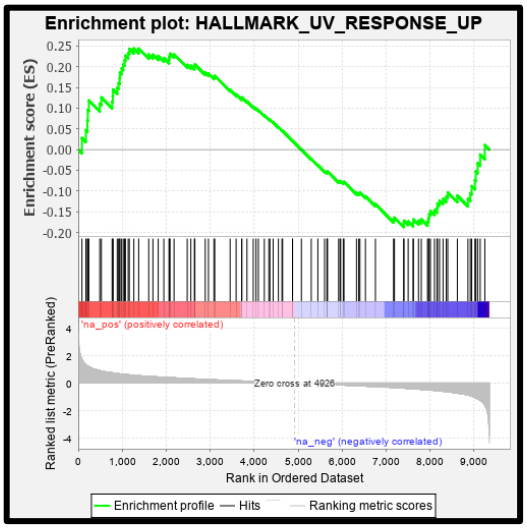


Upregulated Gene	Log2(Fold Change)
AK4	3.33
TPST1	2.47
PFKP	2.44
ISG20	1.86
FAM162A	1.86
PGK1	1.83
TPI1	1.73
ME2	1.66
PDK3	1.57

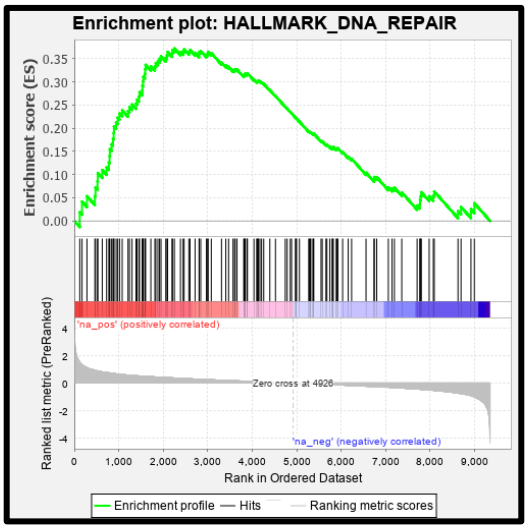
Figure 4.2 GSEA of the HPC7-JAK2V617F cell line suggests increases in pathways associated with metabolism and energy demand.

GSEA of HPC7-JAK2V617F cells reveals upregulated pathways associated with glycolysis and oxidative phosphorylation which can possibly explain the increase in ROS found in Chapter 3. Tables illustrate the seven upregulated pathways that may explain the results in Chapter 3 and the Tables to the right of the GSEA graphic are a list of the top upregulated genes contributing to the pathway.

PATHWAY	Number of genes involved
REACTOME_DNA_REPAIR	264
HALLMARK_DNA_REPAIR	131
HALLMARK_UV_RESPONSE	101
REACTOME_HOMOLOGY_DIRECTED_REPAIR	97



upregulated Gene	Log2(Fold Change)
GCH1	1.75
RFC4	1.34
TFRC	1.27
HMOX1	1.24
CLTB	1.2
AP2S1	0.91
FKBP4	0.88
PDAP1	0.75
CYB5B	0.744



Upregulated Gene	Log2(Fold Change)
GSDME	1.13
NME1	0.93
CHEK1	0.9
AK1	0.89
NFX1	0.87
CETN2	0.81
APRT	0.75
RAD51	0.72
RAD52	0.55

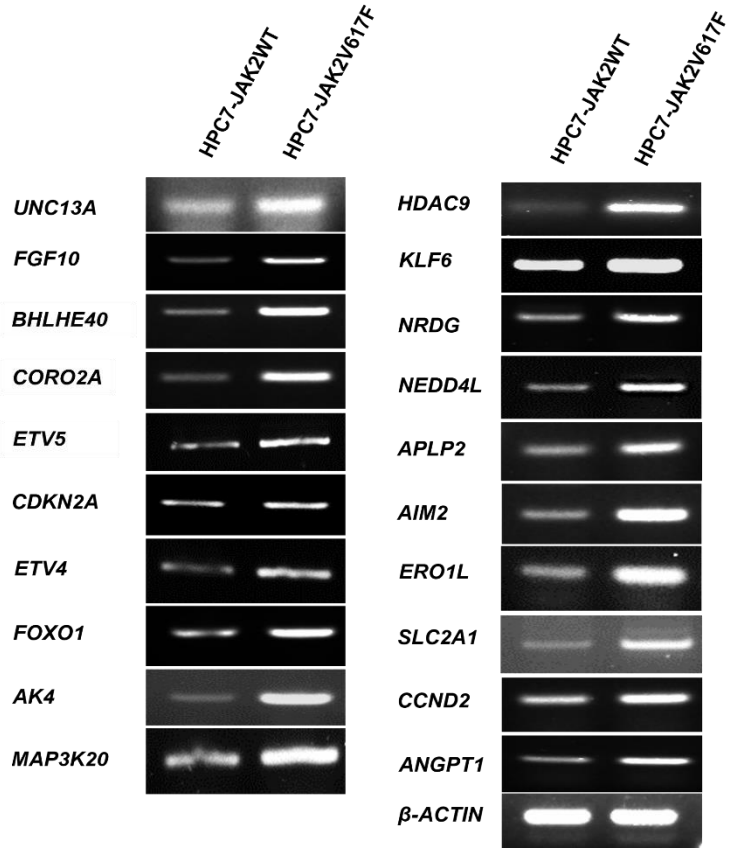
Figure 4.3 GSEA of the HPC7-JAK2V617F cell line suggests increases in pathways associated with DNA repair.

GSEA of HPC7-JAK2V617F cells reveals upregulated pathways associated with DNA repair and UV response which may explain the proposed increased DNA repair capacity of the HPC7-JAK2V617F cells. Tables illustrate the 7 upregulated pathways that may explain the results in chapter 3 and the Tables to the right of the GSEA graphic are a list of the top upregulated genes contributing to the pathway.

Analysis of genes involved in homologous recombination revealed that *BRCA1* and *RAD51* were the top ranked upregulated genes in that group. The *XRCC1, BLM1, SEM1 and RPA1* genes were also upregulated. I have highlighted these genes because their increased expression negatively regulates R-loop accumulation (Stelzer et al., 2016) which I investigated in Chapter 3, and this may explain why I did not see any significant increases in R-loop formation within the HPC7-JAK2V617F cell line which I originally predicted would be present due to the increases in global transcription.

Due to the lack of biological replicates for my sequencing data, I wanted to validate my RNA-sequencing data to confirm the transcriptomic profile of the HPC7-JAK2WT and HPC7-JAK2V617F. To achieve this validation, I took a random assortment of 40 dysregulated genes from a wide range of log2(fold change) differences in the HPC7-JAK2V617F relative to the HPC7-JAK2WT cells and performed RT-PCR with the relevant primers on extracted RNA from these samples. Post-PCR, I ran the samples on an agarose gel with Midori Green (DNA-binding dye) to visualise bands and compare image density between HPC7-JAK2WT and HPC7-JAK2V617F cells via ImageJ software (**Figure 4.4 + Figure 4.5)**

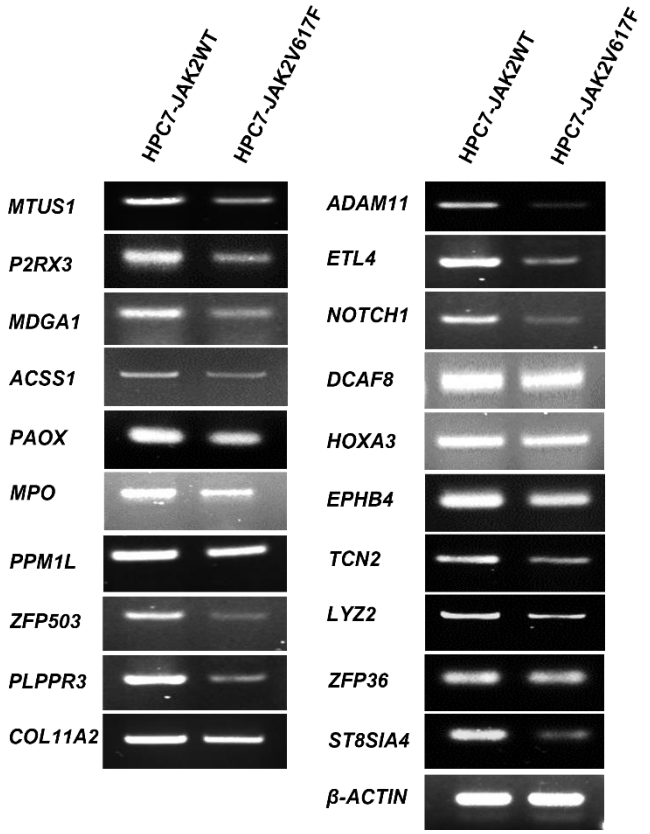
The biological validation revealed that all of the genes tested were indeed dysregulated with RT-PCRs showing that tested genes are dysregulated in the HPC7-JAK2V617. However, the image densitometry values generated were variable compared the values generated by the RNA-sequencing.



	Log2 (Fold Change)					Log2 (Fold Change)		
Gene	RNA-seq value	Mean RT-PCR Value ± SEM	Difference		Gene	RNA-seq value	Mean RT-PCR Value ± SEM	Difference
UNC13A	1.918	1.357 ± 0.177	-0.561		*HDAC9*	1.963	2.8 ± 0.174	0.837
FGF10	3.808	2.966 ± 0.182	-0.843		*KLF6*	1.138	1.371 ± 0.199	0.234
BHLHE40	2.233	3.012 ± 0.073	0.779		*NDRG1*	2.170	1.771 ± 0.205	-0.399
CORO2A	3.058	3.23 ± 0.140	0.171		*NEDD4l*	1.669	2.092 ± 0.193	0.423
ETV5	2.963	2.333 ± 0.211	-0.630		*APLP2*	2.070	1.499 ± 0.119	-0.572
CDKN2A	2.585	1.569 ± 0.054	-1.016		*AIM2*	2.144	2.253 ± 0.220	0.109
ETV4	3.433	2.25 ± 0.113	-1.183		*ERO1L*	1.836	2.171 ± 0.155	0.335
FOXO1	0.941	1.941 ± 0.088	1.000		*SLC2A1*	1.824	1.674 ± 0.139	-0.150
AK4	3.324	2.328 ± 0.224	-0.997		*CCND2*	2.214	1.437 ± 0.080	-0.777
MAP3K20	3.018	1.715 ± 0.187	-1.303		*ANGPT1*	1.42	2.015 ± 0.134	0.595
						Average log2(fold change) difference = -0.197		
						Average standard deviation of the log2(fold change) = 0.946		

Figure 4.4 RNA-Sequencing data validation by RT-PCR demonstrates that the upregulated genes found in the HPC7-JAK2V617F are genuine.

The top part of the figure shows RT-PCR analysis of samples from HPC7-JAK2WT/V617F that have been amplified with primers specific to the gene expression being measured. Three biological replicates were performed, and the images are the most representative of the average values**.** The Table underneath contains: the RNA-sequencing data and the average fold change generated by image densitometry of three different biological replicates of HPC7-JAK2WT and HPC7-JAK2V617F and the "difference" which is the RT-PCR value – the RNA-seq value.



	Log2 (Fold Change)					Log2 (Fold Change)		
Gene	RNA-seq value	Mean RT-PCR Value ± SEM	Difference		Gene	RNA-seq value	Mean RT-PCR Value ± SEM	Difference
MTUS1	-4.37	-1.816	2.554		*ADAM11*	-3.82	-3.863	-0.043
P2RX3	-1.91	-2.032	-0.122		*ETL4*	-3.18	-2.492	0.688
MDGA1	-1.94	-1.703	0.237		*NOTCH1*	-3.82	-2.475	1.345
ACSS1	-1.93	-1.408	0.522		*DCAF8*	-1.09	-0.490	0.6
PAOX	-2.1	-1.538	0.562		*HOXA3*	-1.4	-1.080	0.32
MPO	-3.35	-2.6	0.750		*EPHB4*	-1.9	-1.716	0.184
PPM1L	-2.57	-1.673	0.897		*TCN2*	-2.38	-1.800	0.58
ZFP503	-2.7	-1.793	0.907		*LYZ2*	-2.41	-2.205	0.205
PLPPR3	-2.81	-2.527	0.283		*ZFP36*	-1.94	1.658	3.598
COL11A2	-1.99	-1.501	0.489		*ST8SIA4*	-2.05	-2.936	-0.886
						Average fold change difference = 0.683		
						Average standard deviation of the log2(fold change) = 0.704		

Figure 4.5 RNA-Sequencing data validation by RT-PCR demonstrates that the downregulated genes found in the HPC7-JAK2V617F are genuine.

The top part of the Figure shows RT-PCR analysis of HPC7-JAK2WT/V617F that have been amplified with primers specific to the gene expression being measured. Three biological replicates were performed, and the images are the most representative of the average values**.** The Table underneath contains: the RNA-sequencing data and the average fold change generated by image densitometry of three different biological replicates of HPC7-JAK2WT and HPC7-JAK2V617F and the "difference" which is the RT-PCR value – the RNA-seq value.

4.3.2 HPC7-JAK2V617F genetic profile resembles that of a PV patient

Following successful validation of the RNA-sequencing data, I wanted to compare my data with published data to determine how similar my data is to other datasets. I furthermore wished to filter out clinically irrelevant genes to focus on genes with a possible novel or therapeutic value and explore how the genes may contribute to the effects of the JAK2V617F mutation in HPC7 cells.

To compare my data, I searched the Gene Expression Omnibus (GEO) on the National Centre Biotechnology Information (NCBI) website for relevant datasets. For the comparison, I used GSE120595 which was produced by sequencing $CD34^+$ cells taken from 10-week-old mice with either JAK2V617F or a control vector (five from both males and females) (Celik et al., 2018), GSE103237 which was produced by taking $CD34^+$ cells from ET, PV, and MF patients (Zini et al., 2017), GSE79198 which was produced by sequencing long term haematopoietic stem cells (LT-HSC, Lin-Sca-1+c-kit+CD34-CD135-) isolated from control and MxCre JAK2V617F/+ mice (Yang, Y. et al., 2016) and finally, I used data generated from Burst-Forming Unit-Erythrocytes (BFU-E) with wild-type JAK2 or heterozygous JAK2V617F cultured from PV patients. The BFU-E dataset was provided by Dr Edwin Chen; generated in Chen,. 2010 and deposited onto the ArrayExpress website under the ascension code E-MTAB-384.

The genes and biological pathways being dysregulated in my data parallel that of PV patients and the BFU-E datasets (**Figure 4.6**) which is possibly due to HPC7 cell line having an erythroid phenotype. However, I found that my data had more pathways activated than patient data. The published datasets were

also, very different from each other regarding overlapping GSEA profiles (**Figure 4.6**) suggesting that JAK2V617F influence on the genome is regulated by many factors. The patient data in GSE103237 suggested either a more nuanced gene dysregulation in JAK2V617F PV patients or possibly the PV patients have an unidentified secondary mutation. Another possibility is that the HPC7 cell line is not a true HSC, and this may be reflected through the sequencing data.

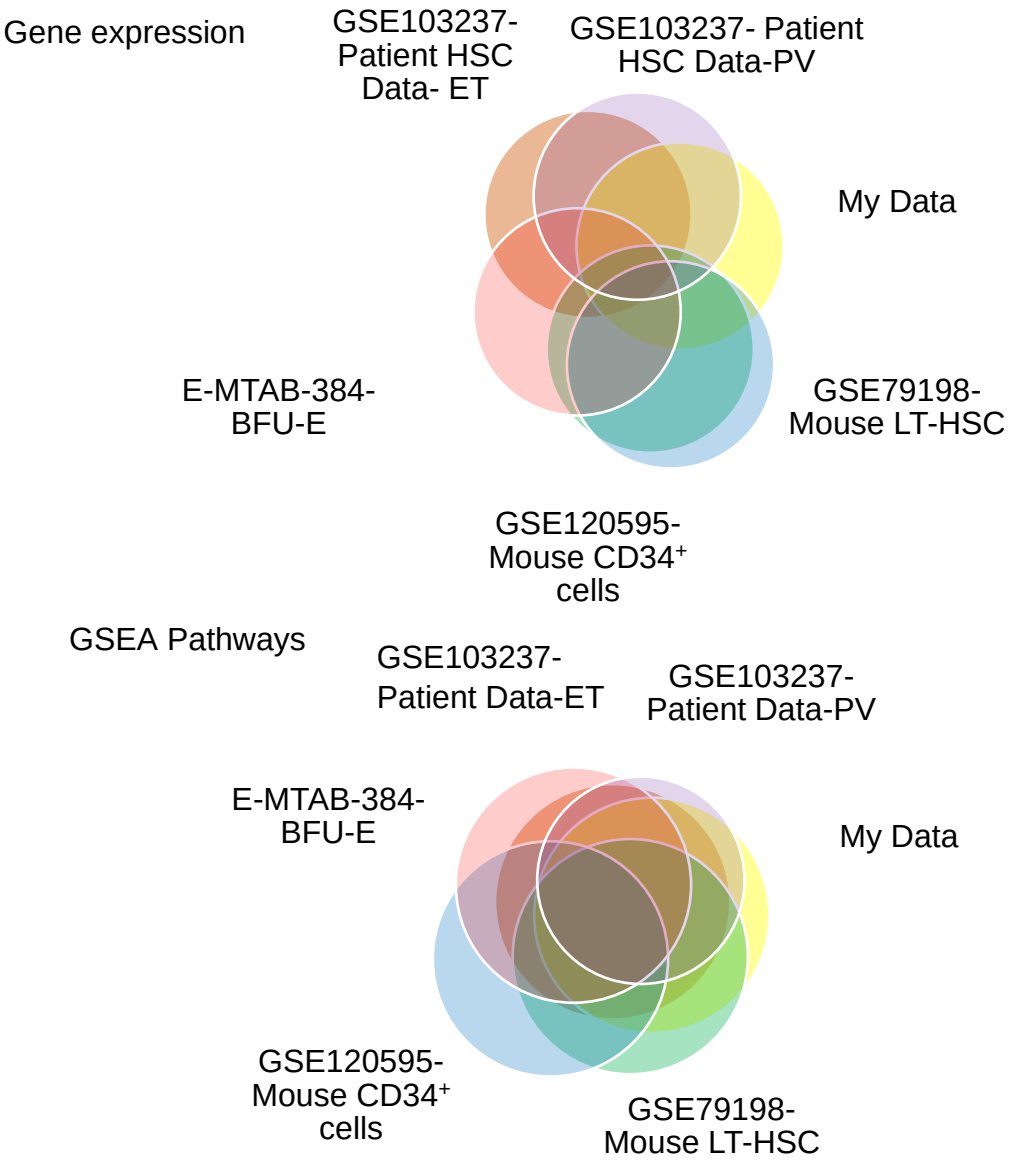


Figure 4.6 JAK2V617F expression in the HPC7 cell line upregulates similar pathways to the published data but through expression of different genes.

The top Venn diagram represents the overlap between the genes dysregulated in all five datasets (dysregulation was defined as being >1.5 or <-1.5 $\log_2$(Fold Change)). The bottom Venn diagram represents pathways dysregulated presented by the GSEA result using the previous definition provided for gene expression. The overlapping Venn diagrams for gene expression hint that my data has more in common with BFU-E gene expression. The GSEA indicates that all datasets have very similar pathways upregulated.

The datasets show variation between each other despite sharing the same mutation with the greatest variation being between the datasets for mice and those for human. To remain relevant to patient clinical data I decided to further filter my data to only show genes dysregulated in all datasets. I further took the genes that had been validated and filtered them again but prioritised novelty and $\log_2$(fold change). This identified *HDAC9* as the prime candidate.

4.3.3 *HDAC9* is a commonly dysregulated gene in JAK2V617F-expressing cells

HDAC9 has not been widely examined in haematological malignancies and has been found to be necessary for haematopoiesis to occur (Wang, P. et al., 2020). A few recent publications focused on the role of *HDAC9* in cancer and DNA damage and in 2012, *HDAC9* was found by Skov et al. to be significantly upregulated in whole blood samples of ET, PV, and primary MF patients relative to the control samples, the control samples were taken from non-diseased subjects (Skov et al., 2012).

In my sequencing data, *HDAC9* was found to be ~1.9 log_2 fold upregulated in the HPC7-JAK2V617F cells whilst Skov et al. observed an average increase of ~1.6 in MPN patients compared to their control. In the patient-based datasets *HDAC9* was upregulated ~1.41,~1.8 and ~1.34 log_2 fold in ET, PV, and MF respectively, relative to the control subjects. In the BFU-E data, *HDAC9* expression was upregulated ~1.62 log_2 fold. In the non-patient-based GSE datasets, HDAC9 was found to be increased by around ~2-fold. In the GSE dataset 103237 (patient data) PV and MF had large variations in increased *HDAC9* expression ranging from decreased expression or no altered expression to ~3-fold increased expression. Notably, patients with no increases of expression of *HDAC9* were being treated with Ruxolitinib, a JAK2 inhibitor, which may explain this. As mentioned previously, severe MPNs have multiple mutations therefore it's possible that secondary mutations may alter *HDAC9* expression and explain the variation.

In the final part of this chapter, I will briefly investigate *HDAC9* expression in HSCs expressing JAK2V617F alone or in combination with secondary mutations that are frequently expressed in MPN patients .

The aim of this part of Chapter 4 is understand if secondary mutations may contribute to *HDAC9* expression which is necessary when exploring therapeutic intervention in diseases that express multiple mutations simultaneously.

To achieve this aim, I will again use the GSE database. To explore alterations in *HDAC9* expression when *EZH2* is also mutated I used the GSE79472 dataset (Yang, Y. et al., 2016); for *TET2* I used the GSE62302 dataset (Kameda et al., 2015) and for *DNMT3a*, I used Jacquelin et al. DNA microarray data for a *JAK2V617F+DNMT3a* knockout (Supplemental Figures, (Jacquelin et al., 2018)). All datasets explored in this chapter investigate the gene of interest's loss-of-function mutations through the use of gene silencing in either JAK2WT or JAK2V617F-expressing cells.

The datasets indicate that reduced expression of either DNMT3a, EZH2 or TET2 alongside JAK2V617F alter *HDAC9* expression. Indeed, *DNMT3a* knockdowns in JAK2V617F-expressing cells correlate with reduced *HDAC9 expression* relative to JAK2V617F expression alone. Depleted *EZH2* levels in JAK2V617F-expressing HSCs reduced *HDAC9* expression levels relative to *JAK2V617F* expression alone. *TET2* depletion reduced *HDAC9* expression levels in JAK2V617F-expressing cells, however, this was not statistically significant. Notably, the reduction of *DNMT3a* and *EZH2* in JAK2WT-expressing cells had no significant effect on *HDAC9* expression and *TET2* repression suggests an increase in *HDAC9* expression. The *HDAC9* expression changes are represented graphically in **Figure 4.7**.

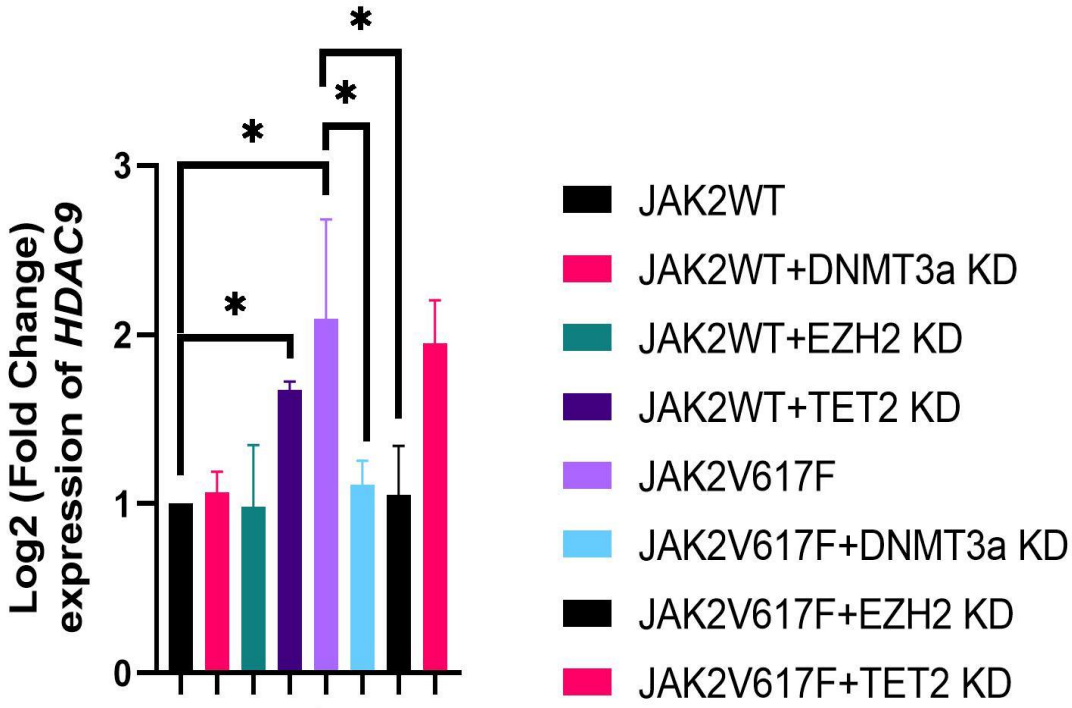


Figure 4.7 *HDAC9* expression in *JAK2V617F* expressing cells is ameliorated by the loss of *DNMT3a* and *EZH2.*

A graph representing the average *HDAC9* expression values of published RNA-sequencing datasets. The published datasets were generated by utilising cell lines that express either *JAK2WT* or *JAK2V617F* alone or in combination with knockdowns of genes commonly found mutated in MPN patients, *DNMT3a*, *EZH2* and *TET2*.

For *JAK2V617F* value, I took the average of all the averages found from the datasets mentioned so far. The datasets indicate that *HDAC9* is significantly upregulated in *JAK2V617F*-expressing cells relative to *JAK2WT*-expressing cells, although there is high variation, possibly due to unidentified secondary mutations or patients being on JAK2 inhibitor treatment which may affect the transcriptional output of *HDAC9*. The knockdown of *TET2* in *JAK2WT* expressing cells seems to increase the expression of *HDAC9.* The *DNMT3a* and *EZH2* knockdown in *JAK2V617F* expressing cells reduces the expression of *HDAC9,* however, in *JAK2WT* expressing cells, *EZH2* and *DNMT3a* knockdown did not reduce *HDAC9* expression suggesting that *DNMT3a* and *EZH2* expression is only important for *HDAC9* regulation in *JAK2V617F* expressing cells.

Testing for statistical significance was performed using a student's t-test (*: $p<0.05$; **: $p<0.01$; *** $p<0.001$).

4.4 Discussion

The purpose of this chapter was to explore my data, validate it with RT-PCR and identify genes whose expression is altered in *JAK2V617F*-expressing cells that might be targets for therapeutic intervention.

The RNA-sequencing data from the HPC7-JAK2V617F cell line was difficult to analyse due to the lack of replicates, making statistical analysis impossible. However, by extensive RT-PCR validation, I can be confident that the genetic profile suggested at by the data is true with a variance of around ±0.895 of the $\log_2$(Fold Change). Further statistical analysis with student's T-tests on the upregulated genes generated a value of 0.364 which is greater than 0.05 therefore hinted that there is no statistically significant difference between the RNA-sequencing data and the RT-PCR. Student T-tests for downregulated genes produced a value of 0.033 which would suggest a statistically significant difference, but this is skewed by large difference between the RNA-sequencing and RT-PCR values for *MTUS1* expression and exclusion of this gene from the dataset actually reverses the significance giving a value of 0.07. The statistical analysis indicates that any gene with a $\log_2$(Fold Change) 1> or <-1 should be treated with caution and validated before further investigation, especially downregulated genes.

Comparison of the pathways that are upregulated with published data suggests that my data matches patient and mouse data with a polarity towards the BFU-E dataset. I suggest this could be explained by HPC7 cells being inherently erythroid-like cells. It should be noted that the datasets were all quite different despite them being HSCs expressing *JAK2V617F* and the difference was

especially noticeable within patients. I suggest this could be caused by secondary mutations in epigenetic regulators.

The secondary aim of this Chapter was to identify a gene for investigation. The datasets were compared and filtered out to find a gene that is consistently upregulated in *JAK2V617F*-driven MPNs and that had not been examined in detail in haematological malignancies. This identified *HDAC9*. However, a further reason to choose *HDAC9* was that two other genes, *CORO2A* and *NCOR2,* are upregulated, and these genes encode proteins which form a repressive protein complex with HDAC9 suggesting that these proteins are upregulated to form a complex to repress certain genes.

Finally, I explored published data which described secondary mutations and investigated if *HDAC9* was dysregulated. *HDAC9* was dysregulated in the presence of the mutant *DNMT3a, EZH2* and *TET2* in combination with JAK2V617F expression. The loss of *DNMT3a* correlated with significant reduction in *HDAC9* expression even without *JAK2V617F* expression meaning that DNMT3a may a direct or indirect role in *HDAC9* expression. *EZH2* knockout in *JAK2V617F*-expressing cells had a negative effect on *HDAC9* expression to a similar extent as *DNMT3a* (Yang, Y. et al., 2016). The dataset was generated by Yang et al. (2016) who suggested that *EZH2* and *DNMT3a* mutations may have similar outputs with the individual loss of either triggering a reduction of repression at certain promoters, specifically at inflammatory and self-renewal pathways.

TET2 knockouts generated a mild increase in *HDAC9* and only a small reduction in collaboration with *JAK2V617F* expression. This is interesting as

TET2 mutations are thought to contribute to disease in a manner that contrasts *DNMT3a* mutations i.e., *TET2* depletion results in a reduction of 5-hydroxymethylcytosine (5-hmc) on Cytosine-phospho-Guanine (CpG) dinucleotides; 5-hmc is required to trigger demethylation and therefore *TET2* depletion is expected to result in an increase of methylated DNA compared to wild-type. By contrast DNMT3a adds methyl groups, therefore the loss of this protein leads to DNA demethylation.

Mutations in epigenetic regulators dysregulate *HDAC9* and therefore by investigating the role of *HDAC9* in HPC7-JAK2V617F cells we can infer what the reduction may do or contribute to in the severe diseases when these mutations are present. To date the role of *HDAC9* has been yet to be investigated in *JAK2V617F* expressing cells.

Chapter 5

Histone deacetylase 9 is essential for cellular signalling which contributes to cell proliferation and DNA damage repair.

Chapter 5

Histone deacetylase 9 is essential for cellular signalling which contributes to cell proliferation and DNA damage repair.

5.1 Introduction

5.1.1 An overview of the HDAC family

Histone deacetylases (HDACs) are a large family of proteins whose function is to deacetylate histones. Over time however, it has been published that HDACs can also deacetylate other proteins which alters their function and plays an active part in regulating protein activity (Wang, P. et al., 2020).

There are four distinct categories of HDACs: Class I which are comprised of HDAC1, HDAC2, HDAC3 and HDAC8, Class II, encompassing HDAC4, HDAC5, HDAC6, HDAC7, HDAC9 and HDAC10, Class III proteins SIRT1, SIRT2, SIRT3, SIRT4, SIRT5, SIRT6 and SIRT7 and finally Class IV which contains HDAC11. HDAC11 is absent in yeast and bears little resemblance to just one class of HDAC; instead HDAC11 has homology to Class I HDAC deacetylase domains and to Class II structure, and hence is in its own category. Class II HDACs can further broken down into type IIa including HDAC4, HDAC5, HDAC7 and HDAC9 and type IIb including HDAC6 and HDAC10. Class IIb HDACs are classified by additional catalytic domains relative to their Class IIa type counterparts and Class IIa are noted for having little deacetylase activity (Wang, P. et al., 2020).

The main function of HDACs is to deacetylate lysine residues found on histone tails which allow for genomic regulation and dynamic shifting of the location of

chromatin to allow repair, transcription, or repression of transcription. Generally, HDACs are thought to counter Histone acetyl transferases (HATs). HATs transfer acetyl residues to histone tails, and HDACs are thought of as “erasers” which remove acetyl groups from lysine residues. Acetyl-modified histone tails are generally associated with transcriptional upregulation as the negatively charged acetyl ($-CH_3CO^-$) reduces interactions of the negatively charged DNA with the now neutrally charged histone tail, allowing binding of basal transcription machinery. HDAC-catalysed removal of acetyl groups reinstates the positively charged lysine and DNA-histone interactions are strengthened. Therefore, HDACs are generally associated with downregulation of transcription.

5.1.2 The role of HDACs in haematopoiesis

As regulators of transcription, HDACs are widely involved in haematopoiesis. HDACs have been published to regulate multi-lineage development and are expressed in all haematopoietic cell lineages. Specific HDACs are upregulated in the appropriate cell types where they form relevant protein complexes with themselves and/or other epigenetic regulatory complexes to regulate the genes necessary for cell maintenance, survival and or differentiation (Wang, P. et al., 2020). For this review I will only focus on myeloid relevant functions.

Class I HDACs 1 and 2 are essential for haematopoiesis and mouse knockout studies demonstrated that HSCs and early haematopoietic progenitors are lost, implying that *HDAC1* and *HDAC2* are necessary for stem cell maintenance and survival. In fact, genes associated with stem cell maintenance were downregulated in the *HDAC1* and *HDAC2* knockouts, specifically *DMKN*, *NURCKS1* and *TPT1* (Heideman et al., 2014). *HDAC1* is a dynamically

expressed gene whose expression is closely linked to the prevalence of haematopoietic transcription factors such as GATA2 and GATA1. GATA2 represses HDAC1 in early haematopoietic progenitors and CMPs which is required for differentiation into granulocytes and GATA1 upregulates HDAC1 in CMPs to promote differentiation into MEPs. Downregulation of *HDAC1* in general seems to polarise haematopoiesis towards the myeloid system suggesting a role in determining cell fate (Heideman et al., 2014).

HDACs 3 and 8 are involved in stem cell maintenance and formation. HDAC3 is a negative regulator of LT-HSC expansion by cooperating with NCOR2, a repressive complex, which catalyses the deacetylation of the *FOS* promoter which regulates HSC self-renewal (Wei et al., 2014). HDAC3 can regulate *GATA2* directly possibly through deacetylation (Ozawa et al., 2001). HDAC8 has yet to be fully elucidated in haematopoiesis, but what is known currently is that HDAC8 is critical for HSC integrity by regulating P53 acetylation (Hua et al., 2017).

Class II HDACs have a myriad of roles in haematopoiesis. HDAC4 function is still under investigation. To date, HDAC4 has been shown to be recruited by BCL6 to repress genes (Lemercier et al., 2002). HDAC5 is critical for HSC migration and engraftment through the binding of HDAC3 and the subsequent deacetylation of P65 (Watamoto et al., 2003). HDAC5 was found to be critical for erythrocyte development by forming a large protein complex with GATA1, ERK, EKLF to regulate γ-globin gene expression (Delehanty et al., 2012).

The role of HDAC6 is largely unknown. Generally, HDAC6 associated with chromatin condensation and erythrocytic enucleation through cooperation with HDAC2 (Li, X. et al., 2017; Ji et al., 2010). HDAC6 can bind HDAC11 which

allows repression and deactivation of the *IL-10* gene (Cheng et al., 2014). HDAC7 is integral to the lymphoid system including B-cell and T-cell development (Azagra et al., 2016). I will discuss HDAC9 individually in Section 5.1.3. HDAC10 has gone largely under-investigated and currently its roles are only associated generally with activities such as transcriptional regulation and cell cycle regulation.

Interestingly, many of the Class IIa HDACs are either devoid of HDAC activity or have weak deacetylase activity. This reduced Class IIa HDAC catalytic activity is caused by a histidine replacing a tyrosine (found in Class I HDACs) within the catalytic domain which sacrifices catalytic potential for oligomerisation (Wang, P. et al., 2020). Class IIa HDACs are thought to exert their influence through binding of other proteins and channel their deacetylase activity through Class I HDACs, which suggests that Class IIa HDACs may function as “scaffold” proteins that form the basis for large protein complexes to form and direct the complex’s activity.

Finally, Class III HDACs revolve around the SIRTs (1-7) which are distinct from classical HDACs as SIRT deacetylase activity is NAD^+ dependent rather Zn^{2+}dependent (Wang, P. et al., 2020). The role of the SIRT family in haematopoiesis has gone largely under-investigated and demands further understanding. Recent publications have demonstrated that SIRT1 and SIRT3 are necessary for HSC maintenance, aging and protection from ROS generation. SIRT1 negatively regulates MTOR signalling through localisation of FOXO3 which protects and maintains older HSCs (Rimmelé et al., 2014; Leko et al., 2012). SIRT3 regulates ROS generated by mitochondria thus preventing additional damage to HSCs (Brown et al., 2013).

5.1.3 Histone Deacetylase 9 (HDAC9)

The HDAC9 protein is a Class IIa histone deacetylase the gene for which resides on chromosome 7. Full-length human HDAC9 is a typical Class IIa HDAC with a regulation domain, a nuclear localisation domain and a deacetylase domain. HDAC9 shuttles between the nucleus and the cytoplasm deacetylating proteins and histones (Alchini et al., 2017). HDAC9 has been the subject of many recent publications relating to cancer and disease in general (Li, H. et al., 2020; Xiong et al., 2019; Shroff et al., 2019; Liu, Feng et al., 2016; Gil et al., 2016; Smith, 2014).

HDAC9 encodes several alternative transcripts. *MITR* (MEF2-Interacting Transcriptional Regulator) is an alternative *HDAC9* transcript which lacks any deacetylase activity and functions through the binding of other proteins such as MEF2 transcription factors (The UniProt, 2021). In mice, *MITR* is the predominant isoform and whilst isoforms encoding deacetylase activity have been computationally predicted they have yet to have been observed biologically (Petrie et al., 2003).

In 2010, Z.Yuan et al. showed that HDAC9 is a negative regulator of *TRIM29* a proliferation promoting gene which normally functions by inhibiting P53 through direct binding and repression of P53 associated genes. However, upon HDAC9 binding to TRIM29 and subsequent deacetylation of lysine 116, TRIM29 undergoes a structural change which blocks its ability to bind to P53 resulting in increased expression of P53 response genes (Yuan et al., 2010).

In 2016, HDAC9 was identified as a driver of B-cell non-Hodgkin lymphoma (B-NHL), as shown by Gil et al.'s publication in which they found that HDAC9 is

frequently overexpressed in B-NHL and is recruited by BCL-6 (which is also frequently overexpressed in B-NHL) through the zinc finger domain of BCL-6. The HDAC9-BCL-6 complex can modulate both P53 and BCL-6 protein activity through HDAC9-driven deacetylation. The HDAC9-BCL-6 complex deacetylates the promoter of *TP53* resulting in reduced P53 expression. P53 expression is integral for the regulation of DNA damage-induced apoptosis and drives high proliferation whilst also blocking DNA damage-induced apoptosis as previously observed in such tumours (Gil et al., 2016).

In 2003, Petrie et al. measured global expression levels for HDAC9 and its alternative splicing variants. They found that in most human cells *HDAC9/ΔCD* (*MITR/HDRP*) is expressed more than its full length *HDAC9* counterpart, but the dominantly expressed isoform was the *HDAC9/ΔCD/Δexon7* which is MITR but without exon 7. *HDAC9/ΔCD/Δexon7* encodes a much shorter form of MITR that was far more highly expressed than the other variants in most cells of the haematopoietic system apart from myeloid progenitors where expression was low. HDAC9 seems to be relatively more expressed in B-cells, T-cells, and monocytes (Petrie et al., 2003).

In 2011, Skov et al. took whole blood samples from ET, PV and MF patients containing oncogenic mutations in either CALR, MPL or JAK2 and measured expression levels of HDACs 1-11 (Class I, II and IV) excluding SIRTs using RT-qPCR. The results showed that in all three diseases HDAC9 is upregulated, indicating that HDAC9 could potentially play a role in MPNs.

This result is consistent with the increased *HDAC9* expression observed in my data therefore I decided to investigate the role of HDAC9 in HPC7-JAK2V617F cells.

5.2 Aims and hypothesis.

In my RNA-sequencing dataset HDAC9 is upregulated which matches both MPN patient data (Skov et al., 2012) and the other datasets discussed in Chapter 4. HDAC9 has been reported to be essential for haematopoiesis and is expressed in HSCs. However, it remains unclear exactly how HDAC9 contributes to both normal haematopoiesis and MPNs. In papers it has been published that HDAC9 has a role in DNA damage repair (Kotian et al., 2011) and cellular signalling. I predict that HDAC9 may be upregulated to allow the increase in cellular signalling and to repair the DNA damage propagated through JAK2V617F expression. In this section I will attempt to explain how HDAC9 contributes to JAK2V617F driven MPNs in HSCs using the following aims:

1) Validate the increased HDAC9 levels in HPC7-JAK2V617F; deplete HDAC9 from HPC7-JAK2V617F cells and investigate differences in cellular signalling and DNA damage repair.

2) To investigate which sections of HDAC9 protein are responsible for the changes seen in aim 1) by removing protein binding sections of the HDAC9 protein binding domain.

5.3 Results

5.3.1 HDAC9 is necessary for STAT3, ERK1/2 and AKT phosphorylation

To investigate the role of HDAC9 in HPC7-JAK2V617F cells, I first wanted to fully validate the RNA-sequencing data to confirm both the change and quantify the extent of the change both in terms of *HDAC9* mRNA and protein levels via a RT-qPCR and a western blot (**Figure 5.1A**). The RT-qPCR and densitometry of the western blot suggest an mRNA increase of 2.8-fold and a protein increase of approximately 2.5-fold which is higher than the log fold change suggested by the RNA-sequencing. I included a negative control using Ruxolitinib which inhibits JAK-STAT signalling and a positive control where I stimulated HPC7-JAK2WT cells for 30 minutes with 50 ng/ml of TPO; TPO activates JAK/STAT signalling through the MPL receptor.

The increased expression levels of HDAC9 in HPC7-JAK2V617F seemed robust and I therefore decided to deplete HDAC9 using shRNA infection and siRNA transfection. Then, I measured mRNA levels via RT-qPCR and protein levels with a western blot; HPC7-JAK2V617F-EV is an Empty Vector control (Figure 5.1B). The result of the shRNA and siRNA transduction yielded an effective knockdown with protein levels being significantly reduced to below detectable levels with the shRNA and the RT-qPCR detecting a reduction of ~90% for the shRNA#1 and ~80% for the siRNA.

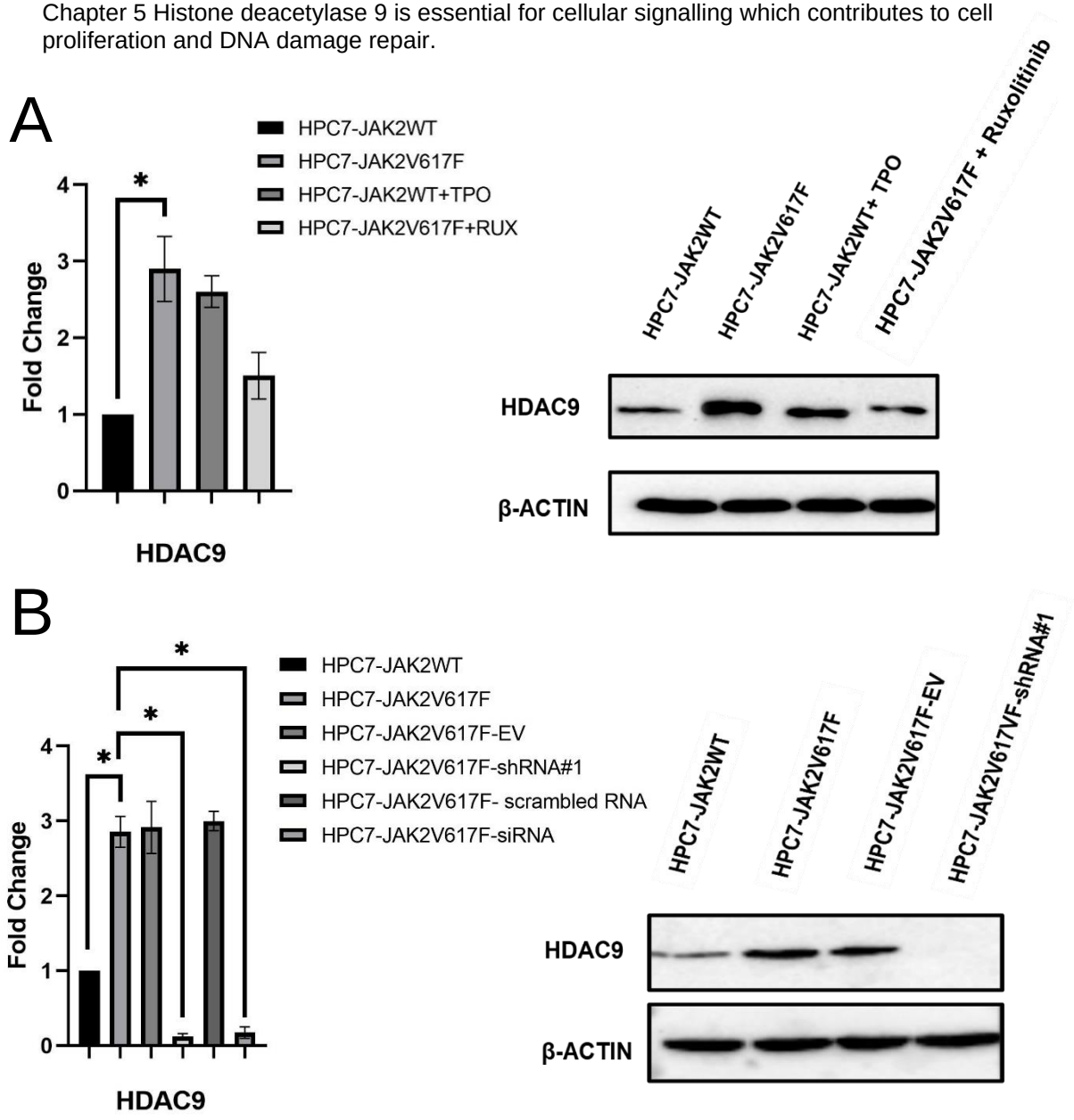


Figure 5.1 *HDAC9* is upregulated in HPC7-JAK2V617F cells and the depletion of *HDAC9* from HPC7-JAK2V617F cell lines.

(A) left image illustrating the results of an RT-qPCR which shows mRNA expression of HDAC9 is higher in HPC7-JAK2V617F cells and HPC7-JAK2WT stimulated with TPO than the HPC7-JAK2WT. Right panel, western blot showing protein levels of HDAC9 are higher in HPC7-JAK2V617F. (B) Left panel, RT-qPCR validating a depletion of HDAC9 in the HPC7-JAK2V617F cells of ~90% for shRNA and ~80% for siRNA. Right panel, western blot validating HDAC9 knockdown in HPC7-JAK2V617F cells.

Testing for statistical significance was performed using a student's t-test (*: $p<0.05$; **: $p<0.01$; *** $p<0.001$). Three biological replicates were performed.

I wanted to characterise the effect of HDAC9 depletion on the HPC7-JAK2V617F cell line beginning with cellular signalling caused by the mutant kinase. The HDAC family has recently been reported to have roles in STAT, ERK1/2 and AKT signalling (Anastas et al., 2019; Lu et al., 2018; Yang, R. et al., 2015; Pinz et al., 2015). I predict that as a Class II HDAC which binds and coordinates multiple members of its family, that HDAC9 may have a role in effective intracellular signalling.

To confirm my hypothesis, I performed western blots analysing phospho-STAT1, STAT3 and STAT5, Phospho-ERK1/2 and phospho-AKT in the HDAC9 depleted JAK2V617F cells (**Figure 5.2A**) and performed RT-qPCRs for genes regulated by these pathways (**Figure 5.2B**). The HDAC9 depletion in HPC7-JAK2V617F cells reduced phosphorylation of the factors tested except phosphorylation of STAT1 and STAT5 which remain unchanged. Likewise, expression of STAT3-target genes was ameliorated back towards HPC7-JAK2WT expression levels, apart from *AIM2* which is a gene regulated by phosphorylated STAT1. Curiously, the depletion of HDAC9 affected the expression of *PF4* whilst not affecting the levels of phosphorylated STAT5. This observation has been seen before in 2015 where Pinz et al. targeted human Treg cells with pan-HDAC inhibitors and the same phenotype presented itself. However, the effect was greater in the Treg cells than shown here, possibly due to the pan-HDAC inhibition rather than just loss of HDAC9. Pinz et al. hypothesised that this observation was caused by the hyperacetylation of basal transcriptional machinery such as BRD4 and this reduces the efficiency of STAT5 transcription.

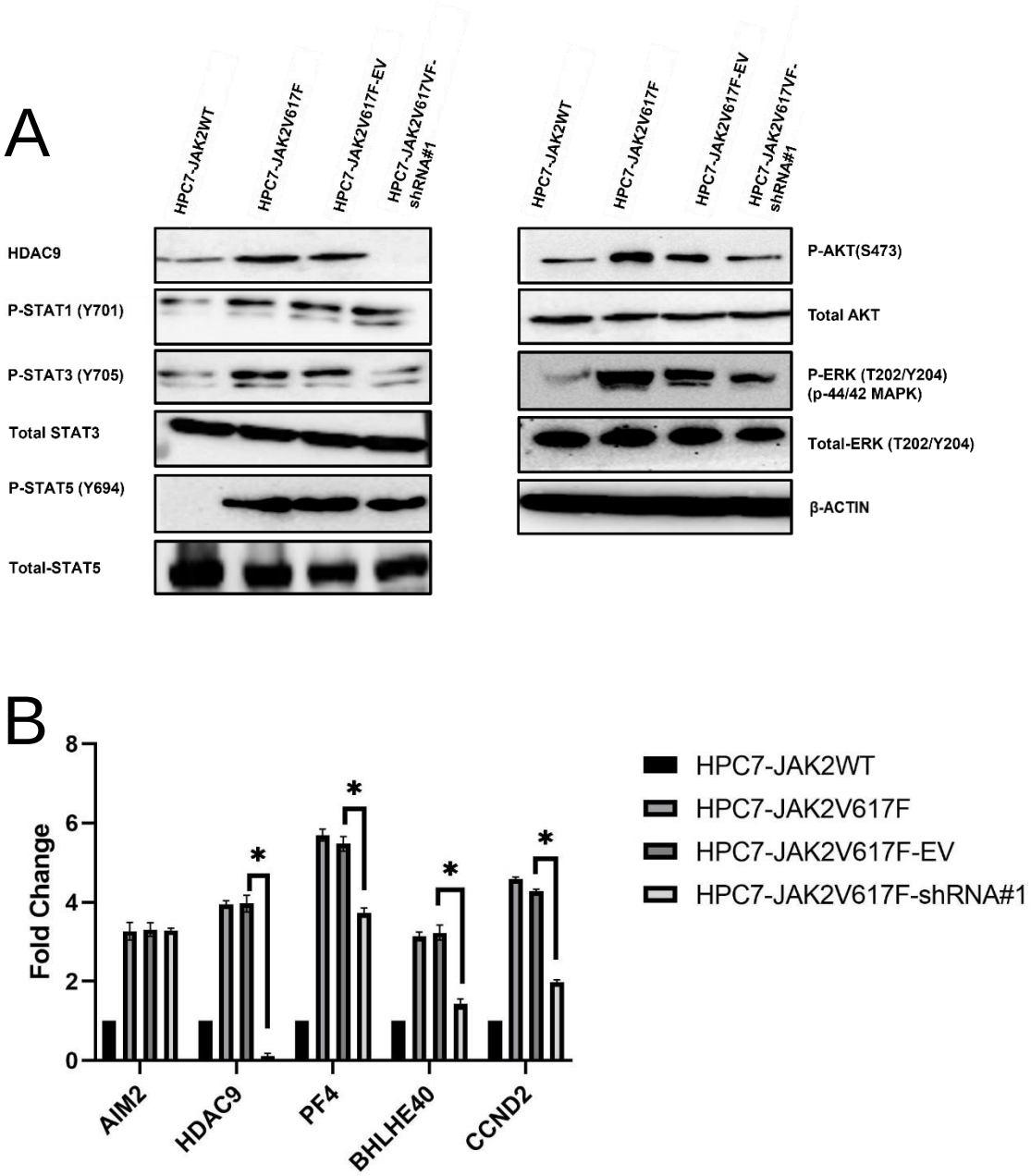


Figure 5.2 HDAC9 is a necessary component for cellular signalling except for STAT1.

(A) Immunoblotting reveals that HDAC9 depletion triggers a reduction in phospho-STAT3/AKT/ERK1/2 signalling but not in STAT1. (B) RT-qPCR showing that expression of the genes associated with these pathways is reduced corroborating a loss of phosphorylation and deactivation, except *AIM2* expression which is mainly regulated by STAT1. STAT5 showed no change in phosphorylation, but the reduction in *PF4* expression suggests a reduction in transcriptional efficacy.

Testing for statistical significance was performed using a student's t-test (*: $p<0.05$; **: $p<0.01$; *** $p<0.001$). Three biological replicates were performed.

Before I investigated the role of HDAC9 in signalling I wanted to be sure that the expression of the other *HDAC* family members is unaffected by HDAC9 and so I tested this by RT-qPCR (**Figure 5.3**). *HDAC9* depletion had no effect on the expression of other *HDAC* family members. Therefore, HDAC9 appears to cause the observed changes in Figure 5.2 rather than the decreased expression of a different HDAC family member.

To elucidate the role of HDAC9 in HPC7-JAK2V617F cells I wanted to understand both the exact component of JAK2 signalling that upregulates HDAC9 expression and the manner in which HDAC9 contributes to the signal. I hypothesise that due to murine HDAC9 being devoid of catalytic activity, it probably influences phospho-signalling through the formation of large complexes with other proteins which in turn deactivate signalling.

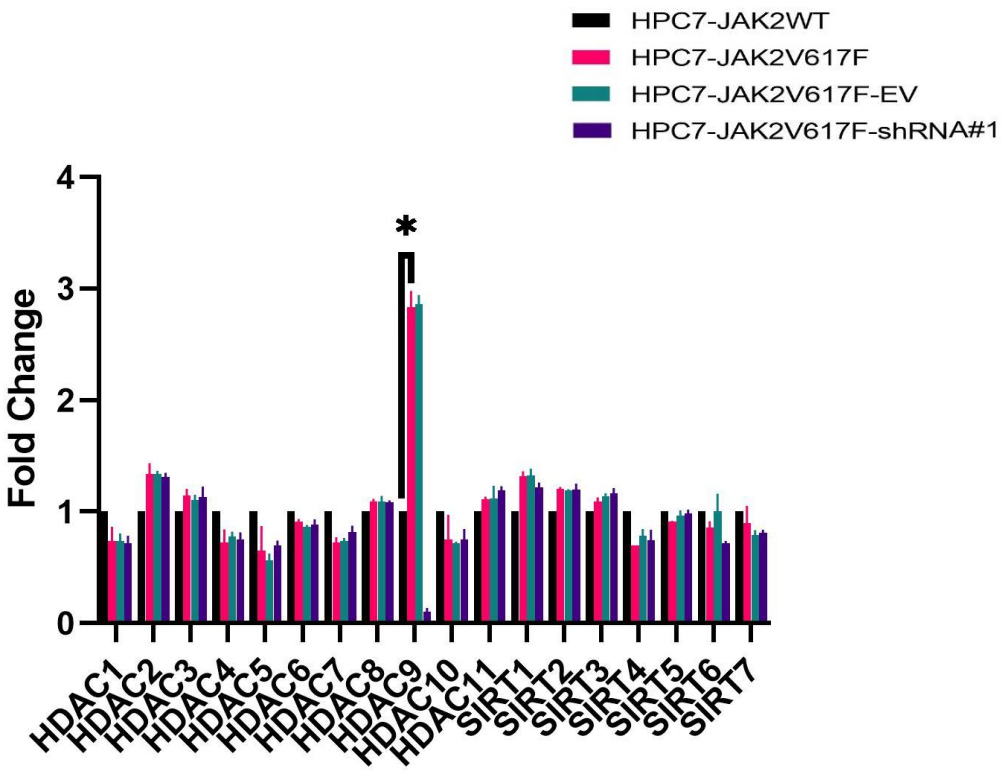


Figure 5.3 HDAC9 is the only HDAC family member dysregulated in HPC7-JAK2V617F and HDAC9-depleted HPC7-JAK2V617F cells.

Figure 5.3 represents a RT-qPCR experiment that tests the levels of every known member of the *HDAC* family to determine if HDAC9 or JAK2V617F influences expression of other *HDAC* family members. HDAC9 depletion has no effect on the expression of its fellow family members but JAK2V617F causes small changes in the expression of *HDAC2, HDAC7* and *SIRT4*.

Testing for statistical significance was performed using a student's t-test (*: $p<0.05$; **: $p<0.01$; *** $p<0.001$). Three biological replicates were performed.

To better understand the cause of *HDAC9* upregulation, I performed experiments utilising inhibitors/shRNAs that target JAK2, STAT1, STAT3 and STAT5a/b which are the main transcription factors activated by JAK2.

I began by using Ruxolitinib which inhibits both JAK1 and JAK2, and then I tested Fedratinib which is a specific inhibitor of JAK2 only. I measured the expression of the JAK-STAT target gene *PF4* after treatment to validate inhibitor efficacy then measured *HDAC9* mRNA levels via a RT-qPCR (**Figure 5.4A**). To validate this on a protein level, I firstly measured phospho-STAT5 to verify the inhibition and HDAC9 expression via a western blot in the HPC7-JAK2V617F cells (**Figure 5.4B)**. As can be seen in Figure 5.4, *HDAC9* upregulation is indeed JAK2 specific. I therefore followed up this result by investigating the STATs performing RT-qPCR. I measured genes associated with each specific STAT under inhibition or with shRNA silencing to validate whether the inhibition/silencing had been successful; I then measured *HDAC9* in the HPC7-JAK2V617F samples.

The inhibitors used and outputs measured were: Fludarabine for phospho-STAT1 by measuring *AIM2* expression, Stattic for phospho-STAT3 by measuring *BHLHE40* expression, STAT5 shRNA9 for phospho-STAT5a and STAT5 inhibitor for both STAT5a and STAT5b by measuring *PF4* expression (**Figure 5.5A).** I then sought to validate this on a protein level via a western blot (**Figure 5.5B)**. For Figures 5.5A and 5.5B before measuring phospho-STAT3 or *BHLHE40*, I reduced MSCF concentration to 0.5 µg/ml in the HPC7-JAK2V617F cells in reference to Figure 3.4.

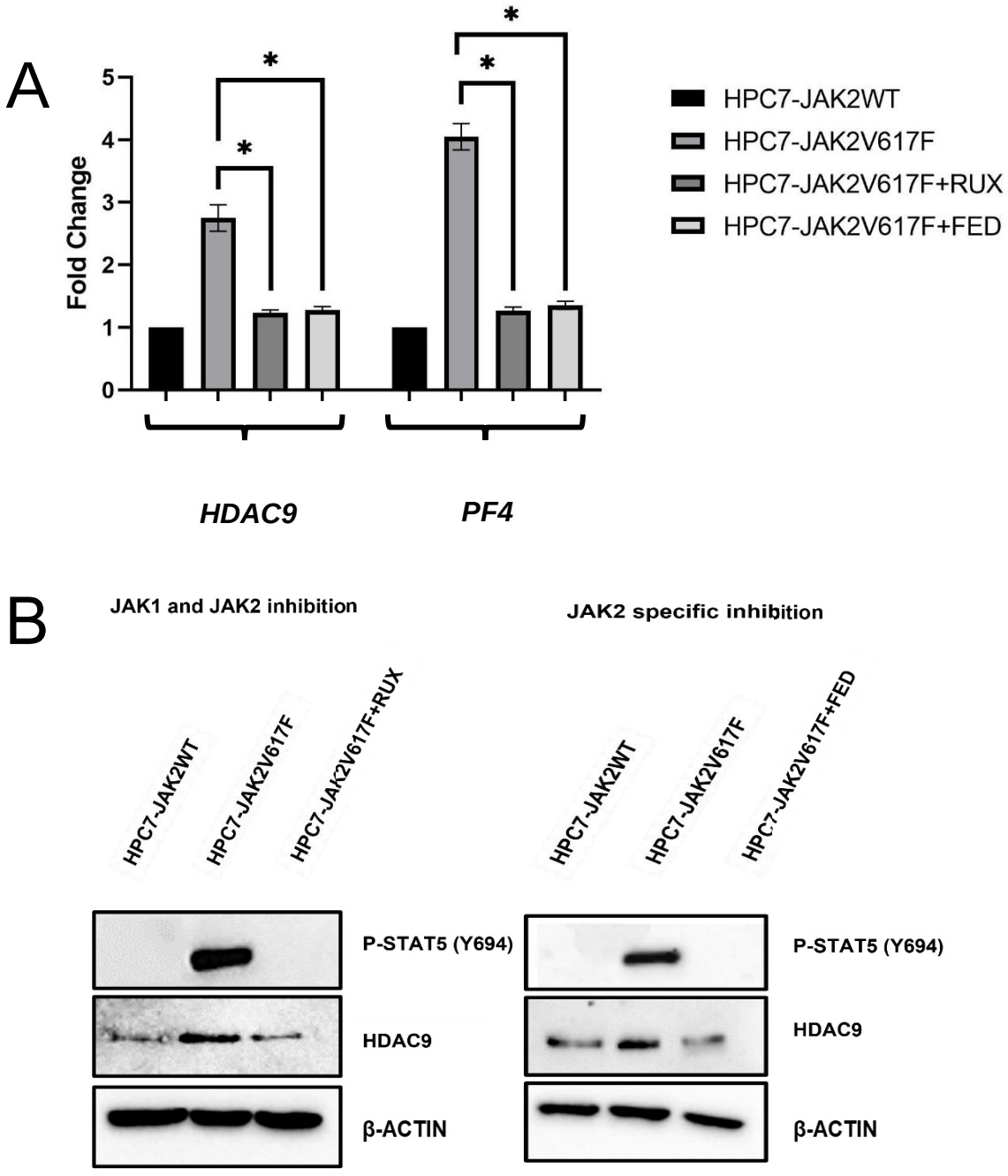


Figure 5.4 JAK2 specifically upregulates *HDAC9*.

(A) RT-qPCR measuring expression of *PF4* and *HDAC9* upon Ruxolitinib and Fedratinib treatment, the reduction of *PF4* expression validates the efficacy of inhibitor treatment and *HDAC9* reduction under Fedratinib demonstrates that *HDAC9* expression is JAK2 specific. (B) Western blot verifying (A) but on a protein level with phospho-STAT5 reduction validating successful inhibitor treatment.

For A, testing for statistical significance was performed using a student's t-test (*: $p<0.05$; **: $p<0.01$; *** $p<0.001$). Three biological replicates were performed.

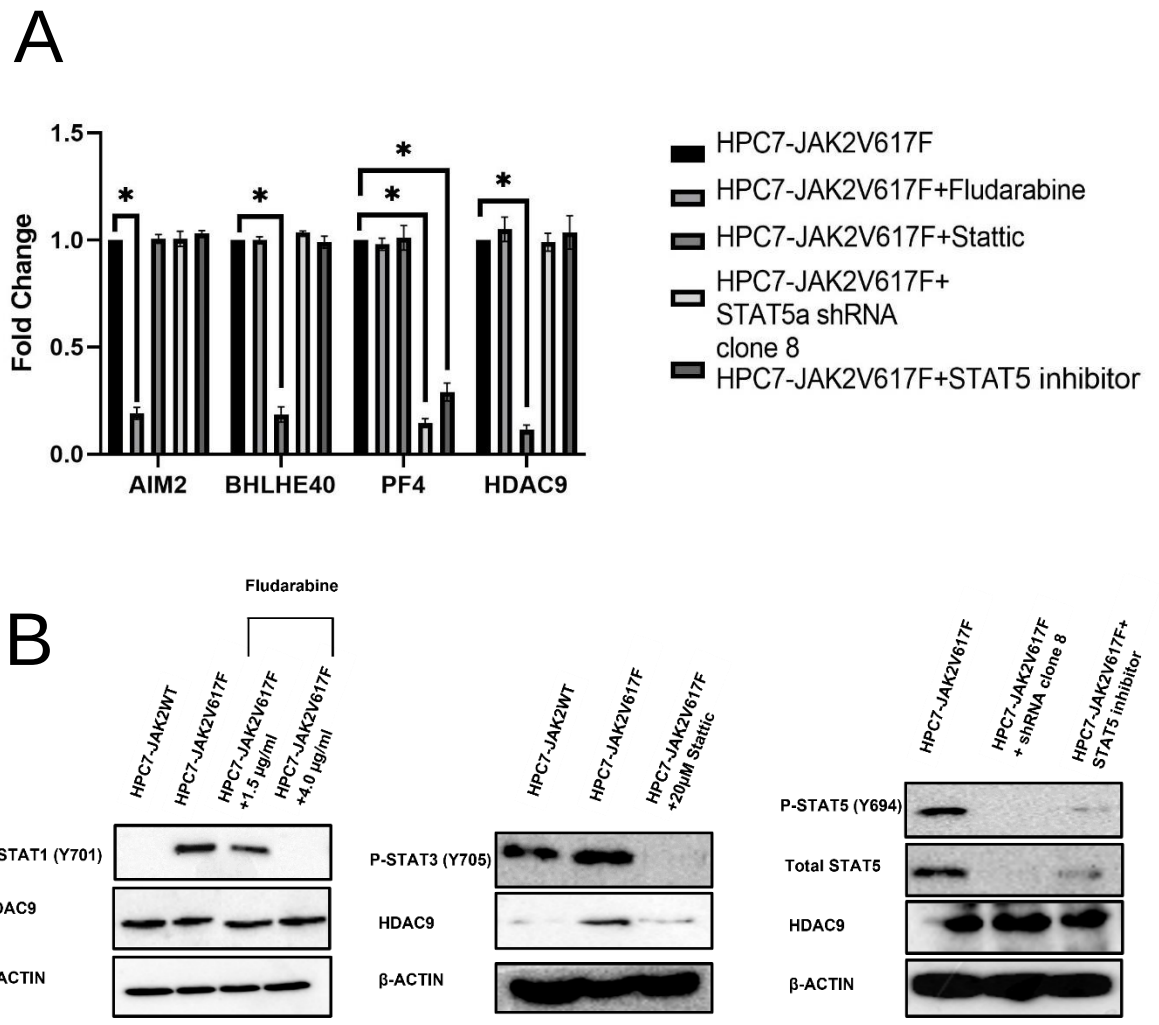


Figure 5.5 JAK2-STAT3 specifically upregulates HDAC9

(A) RT-qPCR measuring HDAC9 gene expression under STAT specific inhibition or shRNA induced silencing (STAT1=*AIM2*, STAT3=*BHLHE40* and STAT5=*PF4*). *HDAC9* is downregulated under Stattic treatment only. (B) Western blots showing the effect of each STAT inhibitor on specific STAT phosphorylation and their effect on HDAC9 expression on a protein level.

The results of Figure 5.5 suggest that STAT3 phosphorylation by JAK2 is a mediator of HDAC9 expression.

For A, testing for statistical significance was performed using a student's t-test (*: $p<0.05$; **: $p<0.01$; *** $p<0.001$). Three biological replicates were performed.

The inhibitor/shRNA depletion results imply that the JAK2-STAT3 pathway drives HDAC9 expression in HPC7-JAK2V617F cells.

The data in Figure 5.2 imply that the upregulation of *HDAC9* in HPC7-JAK2V617F cells is required for cell signalling through the regulation of phosphorylation. I therefore next asked if HDAC9 upregulation is sufficient to drive an increase in signalling in the HPC7-JAK2WT cells. Therefore, I overexpressed a mouse *HDAC9* Open Reading Frame (ORF) vector (Origene) in HPC7-JAK2WT cells by transduction using the TransLT™ reagent and measured phospho-STAT3, phospho-STAT5, phospho-ERK1/2 and phospho-AKT (**Figure 5.6**).

After overexpressing HDAC9, I found that phospho-STAT3, phospho-AKT and phospho-ERK1/2 increased but phospho-STAT1/5 did not. From this, it can be suggested that *HDAC9* overexpression alone is not enough to generate an increase in STAT phosphorylation and that the STAT phosphorylation must already be present which is then amplified through the increased presence of HDAC9. Consistent with this, when HPC7-JAK2WT was starved of MSCF for 48 hours and transduced with *HDAC9*-ORF no increased phosphorylation signal was seen in any of the pathways (data not shown)**.**

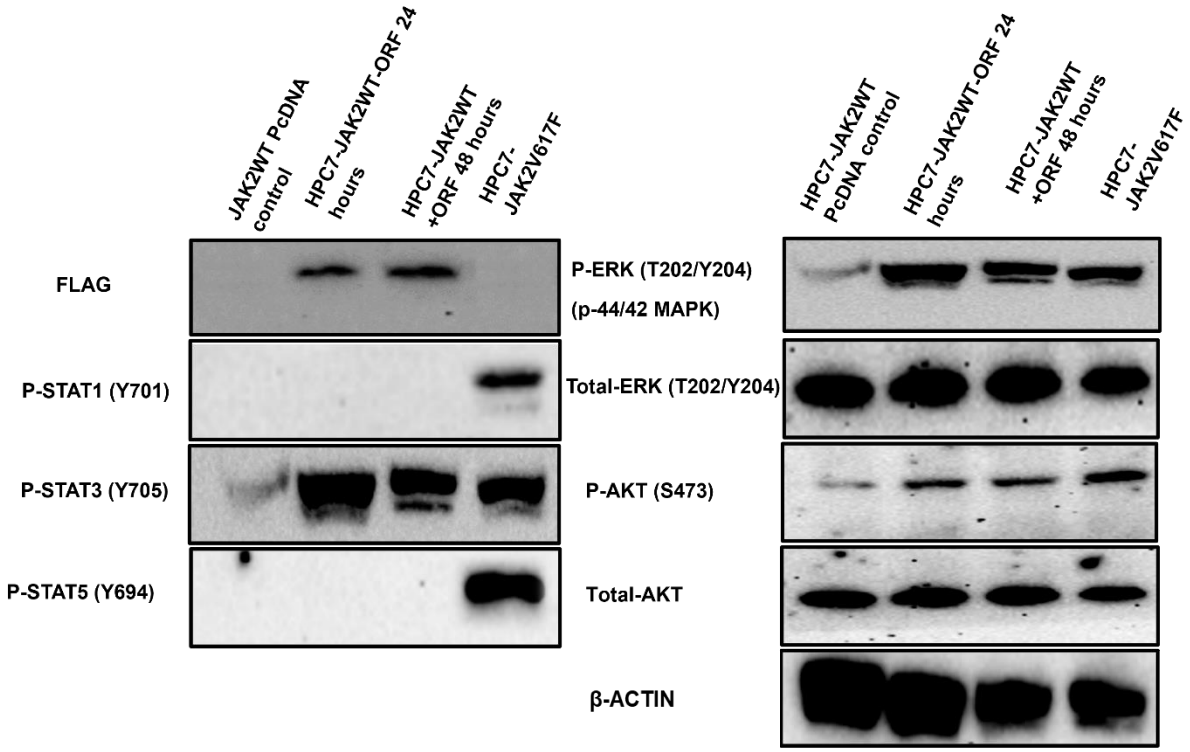


Figure 5.6 HDAC9 overexpression in HPC7-JAK2WT amplifies existing phospho-signalling.

A western blot shows that overexpressing HDAC9 in HPC7-JAK2WT results in increased phosphorylation of STAT3, ERK1/2 and AKT but not STAT1 or STAT5. This result suggests that HDAC9 does not create a signal but allows for it to become amplified in its increased presence.

To explain the change in phosphorylation in the signalling cascades, I first investigated whether total JAK2 phosphorylation is changed (**Figure 5.7**). The unchanged total phosphorylation levels of JAK2 in the HPC7-JAK2WT+ORF cell line or JAK2V617F in the HPC7-JAK2V617F-shRNA#1 cell line suggests that HDAC9 does not regulate signalling through JAK2. Figure 5.6 indicates that the difference in phosphorylation of STAT3, ERK and AKT is post-JAK2V617F phosphorylation of these proteins. Post-JAK2 phosphorylation regulation by HDAC9 is further supported by STAT1/STAT5 being unchanged as a reduction in JAK2V617F phosphorylation would have also reduced STAT1/STAT5 phosphorylation.

I wanted to examine whether HDAC9 directly or indirectly binds to JAK2, STATs, ERK1/2 or AKT as protein deacetylation can regulate catalytic activity. To this end, I performed coimmunoprecipitation experiments with the appropriate STAT/ERK or AKT antibody and stained with HDAC9. I could not use the HDAC9 antibody for immunoprecipitation as it was incompatible with pull-downs. In human Treg cells human HDAC9 can bind STAT5 to regulate acetylation of STAT5 (Beier et al., 2012); in the mouse HPC7 cell lines, however, STAT5 and HDAC9 do not interact (**Figure 5.8**). HDAC9 does not bind STAT3 either. This suggests that it regulates STAT3 phosphorylation indirectly. I predict that the effect of HDAC9 depletion may regulate STAT phosphorylation through indirect means, probably through the formation of protein complexes which regulate genes involved in cell signalling.

HDACs are generally considered transcriptional repressors and I therefore decided to investigate the expression of genes that negatively regulate STAT3 phosphorylation post-JAK2 modification in the HPC7-JAK2V617F-shRNA#1 cell line via RT-qPCR (**Figure 5.9**). The results of the RT-qPCR suggest that *SOCS3* is downregulated in the HPC7-JAK2V617F cells, which has been observed before (Fourouclas et al., 2008); *SOC3* is returned to near HPC7-JAK2WT levels upon HDAC9 depletion.

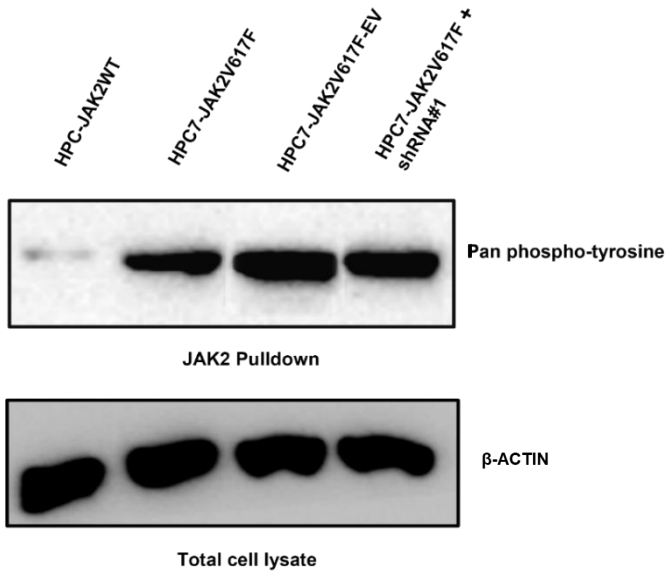


Figure 5.7 HDAC9 does not change levels of JAK2 phosphorylation.

A pulldown of JAK2 followed by a staining with a pan-tyrosine phosphorylation antibody in HPC7-JAK2V617F-shRNA#1 implies that HDAC9 does not influence total JAK2 levels and therefore likely influences STAT phosphorylation through an alternate pathway.

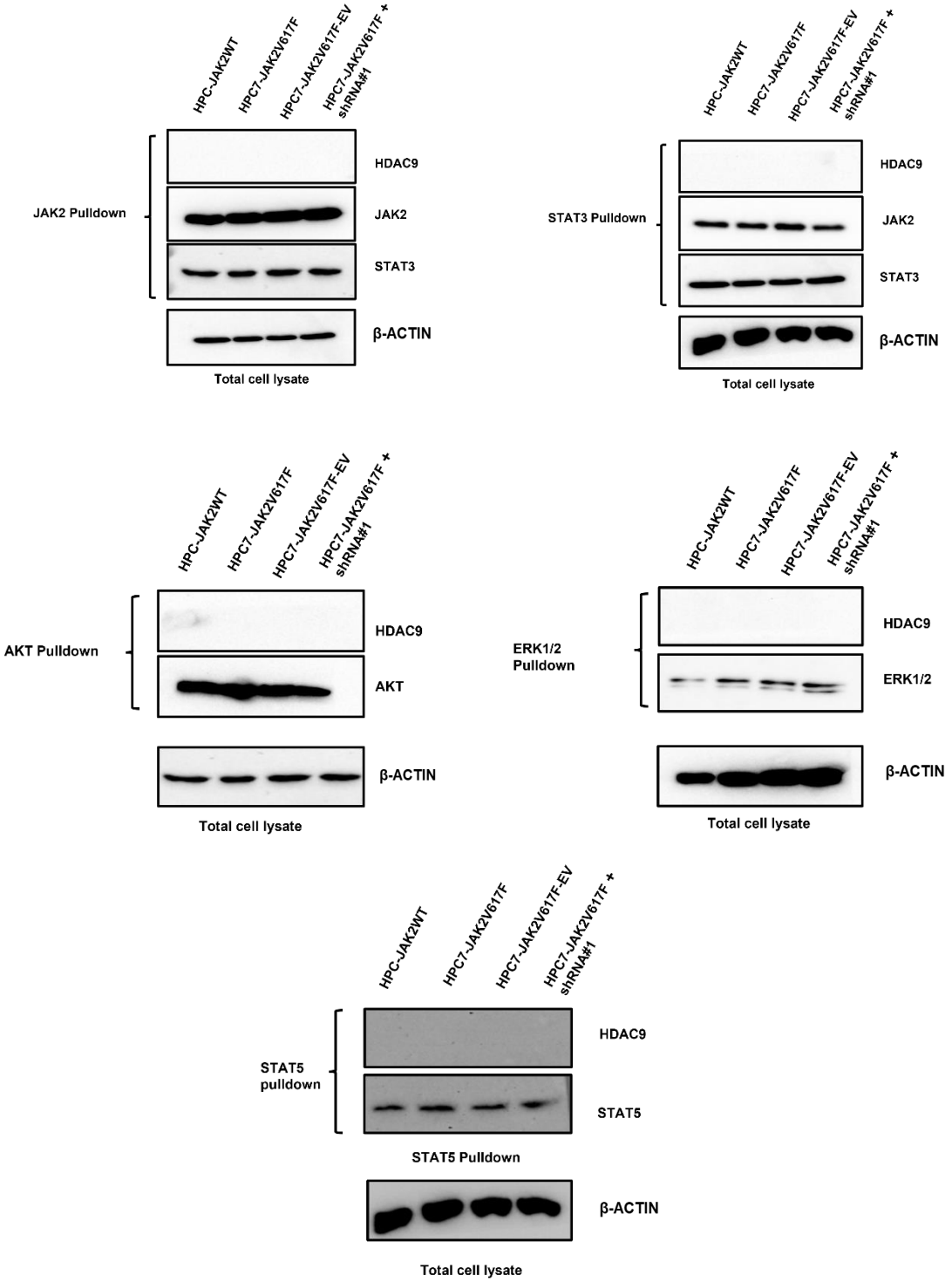


Figure 5.8 HDAC9 does not bind JAK2-related factors.

Immunoprecipitation experiments demonstrate that HDAC9 does not interact with JAK2, STAT3, STAT5 , AKT or ERK1/2. This result suggests HDAC9 could regulate factors through indirect pathways.

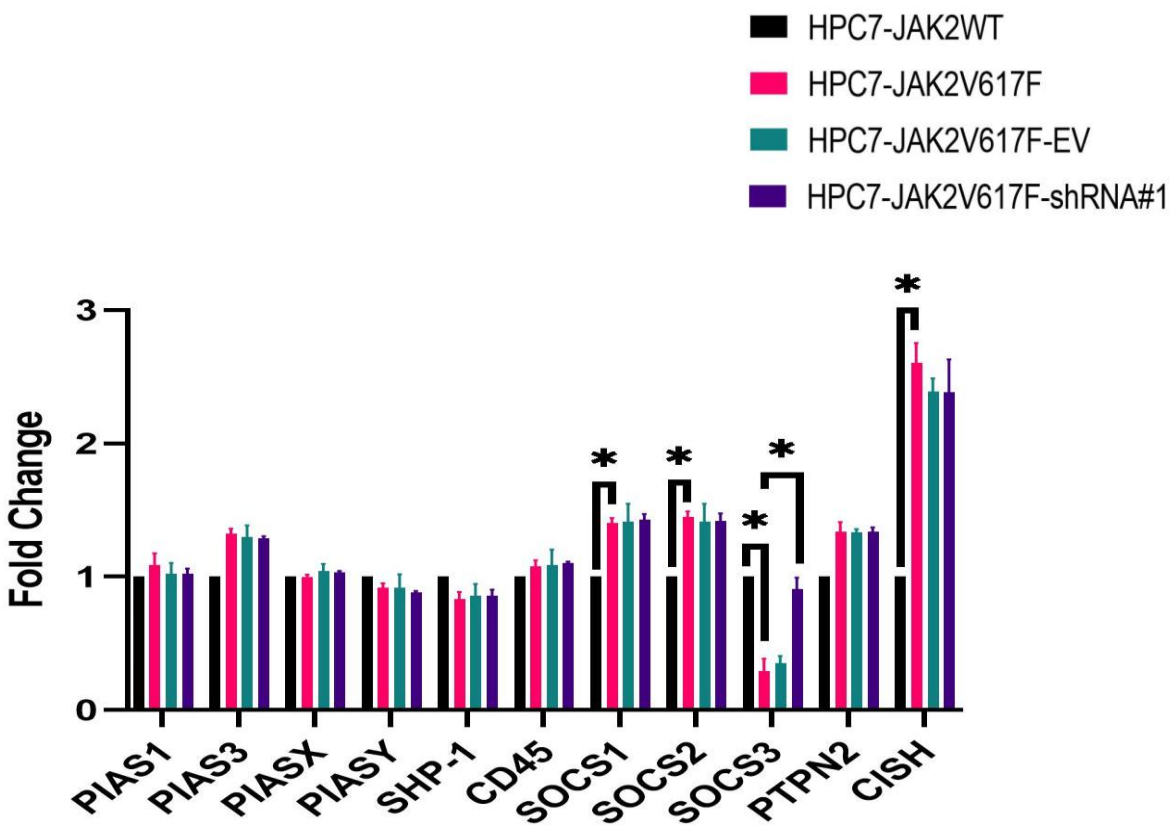


Figure 5.9 HDAC9 depletion in HPC7-JAK2V617F returns *SOCS3* expression towards HPC7-JAK2WT levels.

An RT-qPCR experiment measuring expression of known regulators of STAT phosphorylation. The RT-qPCR illustrates that the expression of *SOCS1, SOCS2, PIAS3,* and *CISH* are upregulated, and *SOCS3* is downregulated in HPC7-JAK2V617F cells. Only *SOCS3* is returns to HPC7-JAK2WT levels upon *HDAC9* depletion.

Testing for statistical significance was performed using a student's t-test (*: $p<0.05$; **: $p<0.01$; *** $p<0.001$). Three biological replicates were performed.

Although this might explain how HDAC9 impacts STAT3 phosphorylation, it does not explain how HDAC9 may be contributing to phospho-AKT and phospho-ERK1/2 pathways. Yang et al. found that increased HDAC9 expression could promote endothelial growth factor receptor (EGFR) signalling and subsequently Phospho-AKT and ERK1/2 signalling through the regulation of *WWTR1* expression which encodes TAZ, a regulator of the Hippo pathway. EGFR contributes to JAK2V617F-driven MPNs through recruitment of JAK2V617F to EGFR and the subsequent chronic activation of EGFR. In fact, in 2006 Li et al. found that inhibition by Erlotinib (an EGFR inhibitor) in Human erythroleukemic (HEL) cells and PV patient BFU-E colonies could dramatically ameliorate JAK2V617F/EGFR signalling and cell viability of PV patient cell colonies and HEL cells through the reduction of STAT3, ERK1/2 and AKT signalling (Li, Z. et al., 2007).

RNA-sequencing data (Chapter 4) suggest that *WWTR1* is two-fold upregulated in HPC7-JAK2V617F. Therefore, I firstly validated this through an RT-qPCR including the HDAC9-depleted HPC7-JAK2V617F cell line and the HPC7-JAK2WT+ORF cell line. RT-qPCR implies that *WWTR1* is around 1.7-fold upregulated in the HPC7-JAK2V617F cells and the HPC7-JAK2WT+ORF cells further also support this idea with an increase in *WWTR1* expression (**Figure 5.10A**).

To compare my model with published data even further, I performed a western blot looking at phospho-EGFR levels in the various HPC7-JAK2 cell lines. The results show that Phospho-EGFR is higher in HPC7-JAK2V617F cells and reduced upon HDAC9 depletion but is not increased in HPC7-JAK2WT-ORF.

This implies that HDAC9 overexpression in HPC7-JAK2WT alone is not enough to increase EGFR signalling in these cells (**Figure 5.10B**). The data suggests HDAC9 protein upregulation may contribute to phospho-AKT and phospho-ERK1/2 signalling through EGFR signalling in JAK2V617F-expressing cells.

As the role of TAZ in EGFR signalling has yet to be understood I wanted to investigate different avenues by which HDAC9 may contribute to ERK1/2 and AKT signalling. In 2016, Gil et al. investigated HDAC9 in B-Cell Lymphomas (Gil et al., 2016) using mouse models that overexpressed *HDAC9* cDNA. The mice expressing the human *HDAC9* cDNA were sacrificed, and their B-cells were examined for differential gene expression. The results revealed that many DUSP proteins are downregulated in HDAC9 overexpressing cells.

Dual specificity phosphatases (DUSP) are phosphatases that regulate cell signalling pathways by deactivating phosphorylated protein by catalysing the removal of phosphates. One of the pathways that DUSP proteins regulate is the MAPK/ERK1/2 pathway. I compared my RNA-sequencing data to that generated by Gil et al. and found that *DUSP2* is downregulated to around 2-fold change which is similar to that found by Gil et al. (2.1-fold change). I decided to test *DUSP2* expression by performing RT-qPCR in the HDAC9-depleted HPC7-JAK2V617F cells. (**Figure 5.11**). *DUSP2* is downregulated in HPC7-JAK2V617F and returns to HPC7-JAK2WT levels upon HDAC9-depletion suggesting HDAC9 regulates *DUSP2* expression.

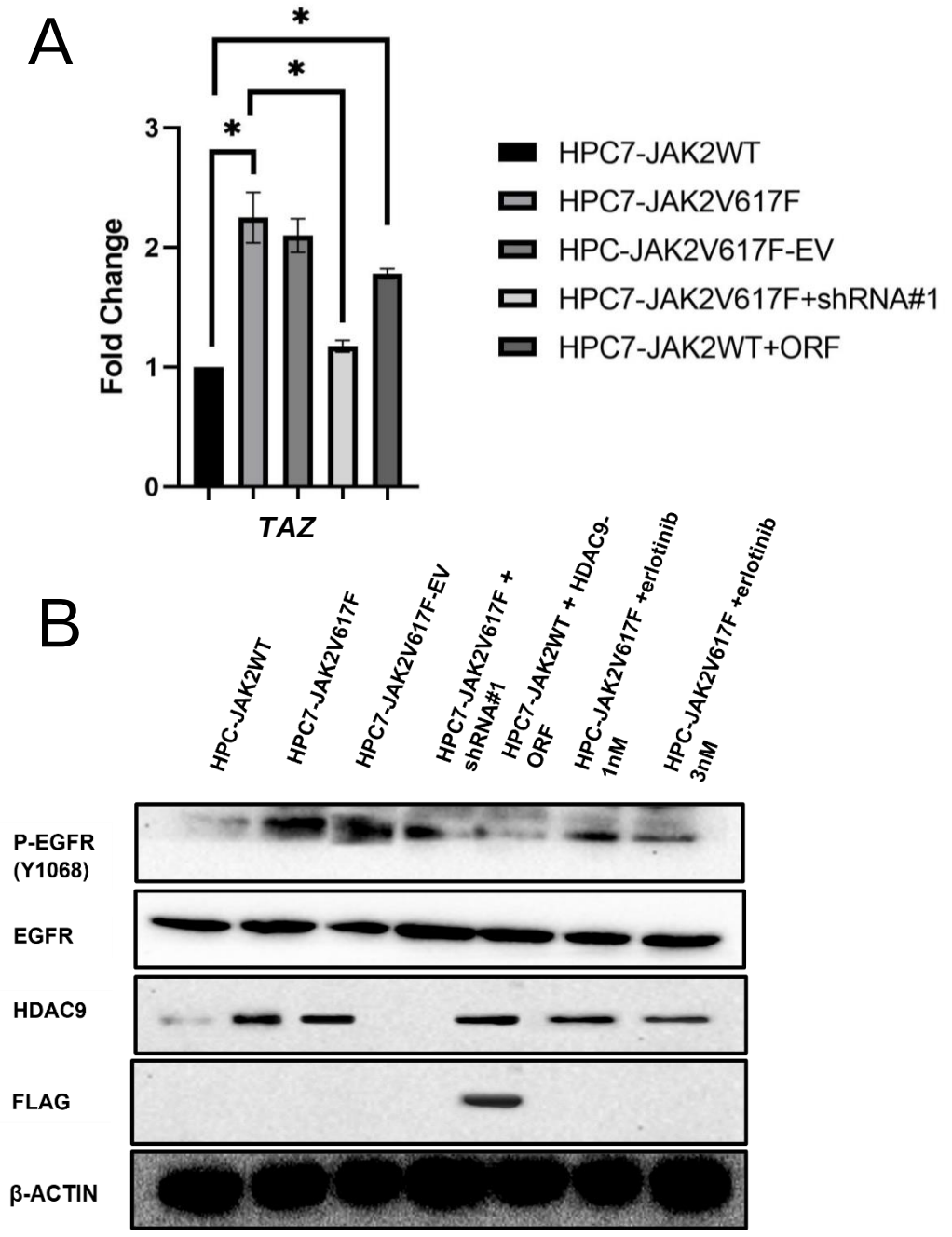


Figure 5.10 Levels of EGFR phosphorylation are linked to HDAC9 expression.

(A) RT-qPCR measuring *TAZ* expression. *TAZ* is overexpressed in HPC7-JAK2V617F cell lines and reduced in HDAC9-depleted cells. (B) Western blot, phospho-EGFR levels are significantly higher in the JAK2V617F expressing cells but are ameliorated upon HDAC9 depletion. Overexpression of HDAC9 in HPC7 JAK2WT cells does not increase EGFR activation, despite the increases in *TAZ* expression.

For A, testing for statistical significance was performed using a student's t-test (*: $p<0.05$; **: $p<0.01$; *** $p<0.001$). Three biological replicates were performed.

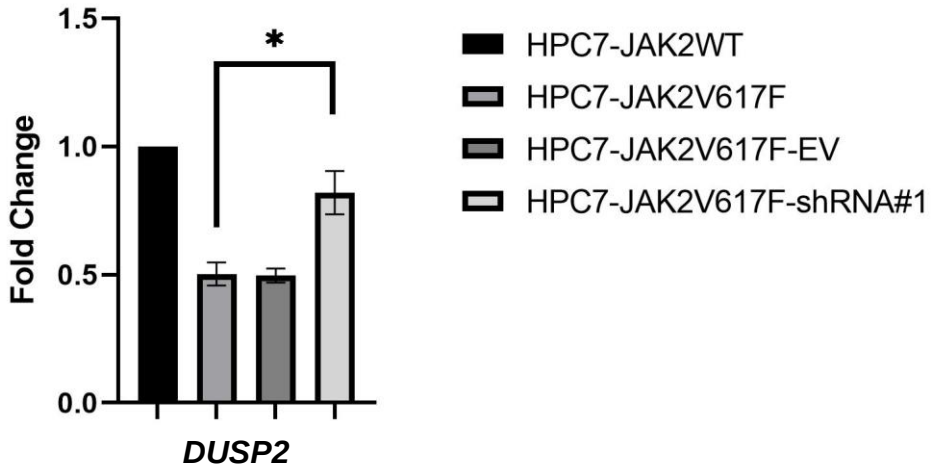


Figure 5.11 *DUSP2* expression is reduced in JAK2V617F cells and returns to base levels upon HDAC9-depletion

RT-qPCR of *DUSP2* expression in HPC7-JAK2V617F cells with and without *HDAC9* shRNA. JAK2V617F negatively regulates *DUSP2* expression possibly through HDAC9-driven deacetylation of histones or binding repressor proteins. The silencing of HDAC9 allows *DUSP2* to be re-expressed.

Testing for statistical significance was performed using a student's t-test (*: $p<0.05$; **: $p<0.01$; *** $p<0.001$). Three biological replicates were performed.

Lastly, I wanted to investigate two other factors to see if they are different in the HPC7-JAK2V617F cells, namely, the phosphorylation status of HDAC9 tyrosine residues and the cellular location of HDAC9.

The phosphorylation of HDAC9 serine residues has been shown to alter the activity and location of HDAC9 which dictates the functional output of the protein (Alchini et al., 2017). However, the role of tyrosine residue phosphorylation has not been investigated. A pulldown with a pan-tyrosine phosphorylation antibody and staining with HDAC9 antibody suggests that the total phosphorylation of tyrosine residues in HDAC9 is higher in the HPC7-JAK2V617F cells than in HPC7-JAK2WT cells (**Figure 5.12A**). The role of tyrosine phosphorylation of HDAC9 should be investigated further to fully understand how it may contribute to HDAC9 function.

Finally, I wanted to explore the cellular location of HDAC9 in JAK2V617F cells compared to wild type cells. I performed subcellular fractionation experiments to isolate the nuclear and the cytoplasmic fractions and then carried out a western blot (**Figure 5.12B**). The results imply that in HPC7-JAK2V617F cell lines, HDAC9 is more expressed in both the cytoplasm and nucleus, being much more prevalent in the nucleus than in HPC7-JAK2WT.

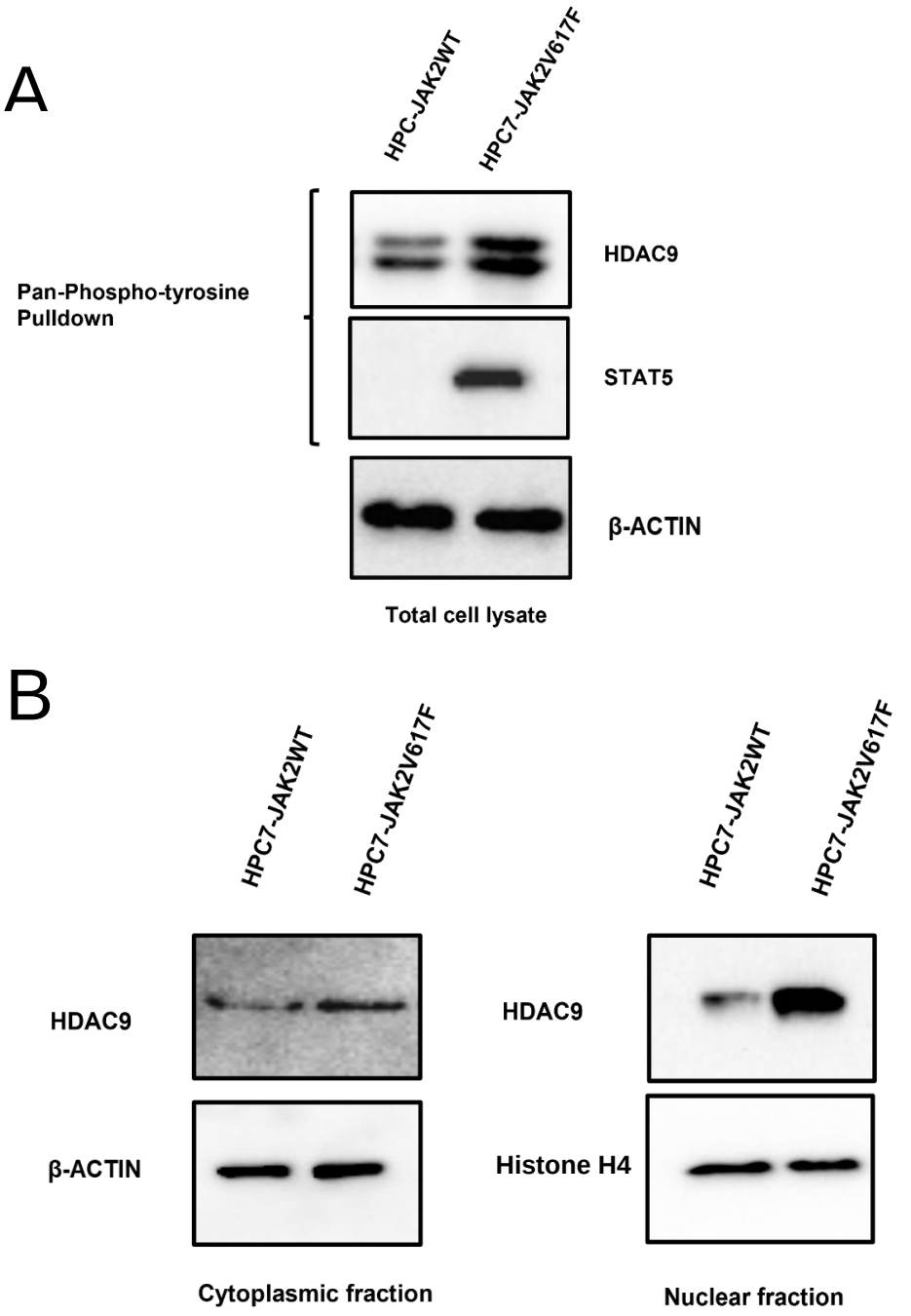


Figure 5.12 Tyrosine phosphorylation and cellular location of HDAC9 is altered by JAK2V617F expression.

(A) Pulldown with a pan-tyrosine antibody followed by probing the western blot with HDAC9 antibody reveals increased tyrosine phosphorylation in the HPC7-JAK2V617F cell line. The two bands may represent different isoforms. (B) Western blot of sub-cellular fractions. HDAC9 staining of the cytoplasm and nucleus reveals that HDAC9 is more prevalent in the nucleus of HPC7-JAK2V617F cells compared to HPC7-JAK2WT cells but is also increased in the cytoplasm to a much lesser extent.

5.3.2 HDAC9 overexpression contributes to cell proliferation in HPC7 cells and DNA damage/DNA damage repair.

As shown in Section 5.3.1, HDAC9 depletion can reduce STAT3/AKT and ERK1/2 signalling. Therefore, I wanted to investigate if HDAC9 has a role in mediating the other phenotypes observed in the HPC7-JAK2V617F cells in Chapter 3. To achieve this, I repeated the same experiments conducted in Chapter 3, but this time including HPC7-JAK2WT-ORF cell line and/or HPC7-JAK2V617F-shRNA#1 cell line.

Firstly, I performed a western blot to examine c-MYC levels as the Class II HDAC family has been shown to repress *c-MYC* expression (Barneda-Zahonero et al., 2015) (**Figure 5.13A**). The western blot shows an increase in c-MYC expression in HPC7-JAK2V617F cells when *HDAC9* is silenced. In Chapter 3, I hypothesised that the downregulation of *c-MYC* could contribute to HSC expansion; I therefore measured the cell proliferation of the HDAC9-depleted HPC7-JAK2V617F cell line using an MTT assay at a MSCF concentration of 1 µg/ml (**Figure 5.13B**). From this it might be concluded that HDAC9 depletion contributes to cell proliferation in the HPC7 cell line as the HPC7-JAK2V617F-shRNA#1 cell line had a lower rate of proliferation than the HPC7-JAK2V617F cell line. Likewise, the HPC7-JAK2WT-ORF cell line had a more modest increase in proliferation compared to HPC7-JAK2WT cells which, may indicate that phospho-STAT3, phospho-ERK1/2 and phospho-AKT, alone or in combination, could have a role in HSC expansion.

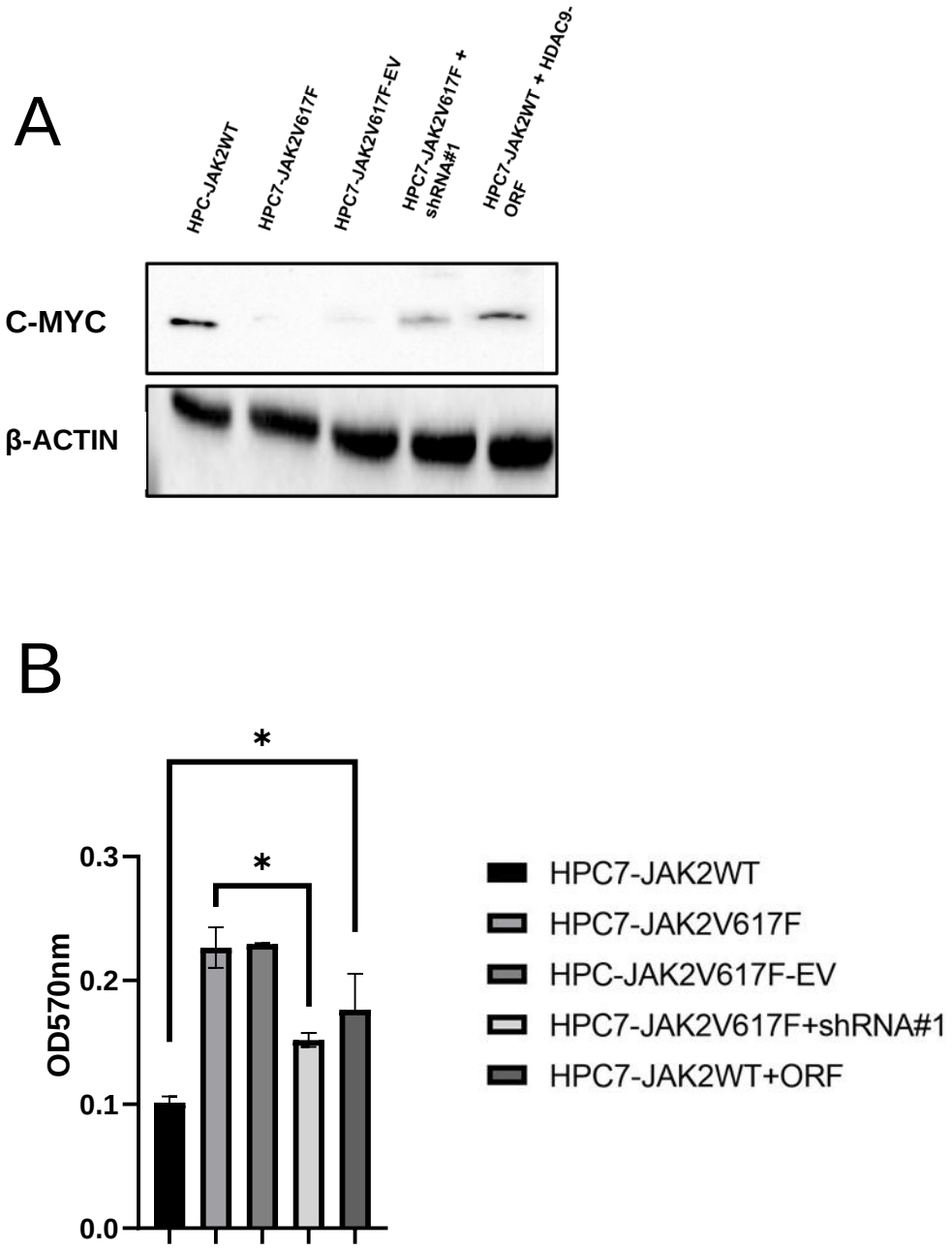


Figure 5.13 HDAC9 depletion increases levels of c-MYC protein expression and reduces cell proliferation.

(A) Immunoblot of c-MYC indicates HDAC9 silencing results in increases in c-MYC protein. (B) MTT assay suggests HDAC9-depletion reduces proliferative capacity of the HPC7-JAK2V617F cells.

For B, testing for statistical significance was performed using a student's t-test (*: $p<0.05$; **: $p<0.01$; *** $p<0.001$). Three biological replicates were performed.

In Section 3.3.5, I found that the HPC7-JAK2V617F cell line has a 1.6-fold increase in ROS and is more tolerant to DNA damage through increased DNA damage repair. I therefore measured ROS levels in the HDAC9 depleted HPC7-JAK2V617F cell line and the HPC7-JAK2WT-ORF cell line (2 days post-transduction) (**Figure 5.14A**). ROS levels were decreased in the HPC7-JAK2V617F-shRNA#1 cell line to around 1.25-fold higher than HPC7-JAK2WT. However, increased expression of HDAC9 in HPC7-JAK2WT cells increased ROS levels by only 1.15-fold.

To explain this discrepancy, I predict that phospho-STAT1 and phospho-STAT5 may also contribute to ROS levels. To validate this, I re-performed the dihydroethidium-based assay with HPC7-JAK2V617F cells incubated with either STAT1 inhibitor (Fludarabine), STAT3 inhibitor (Stattic), STAT5 inhibitor or both STAT3 and STAT5 inhibitors (**Figure 5.14B**). The inhibition of STAT3 caused the greatest reduction on ROS levels, with STAT1 and STAT5 inhibition alone barely altering ROS levels. However, none of the inhibitors caused a statistically significant reduction in ROS levels, this implies that other factors are contributing to the increases in ROS levels. In Chapter 4, increases in metabolic pathways were indicated by the GSEA which may explain the increases in ROS levels.

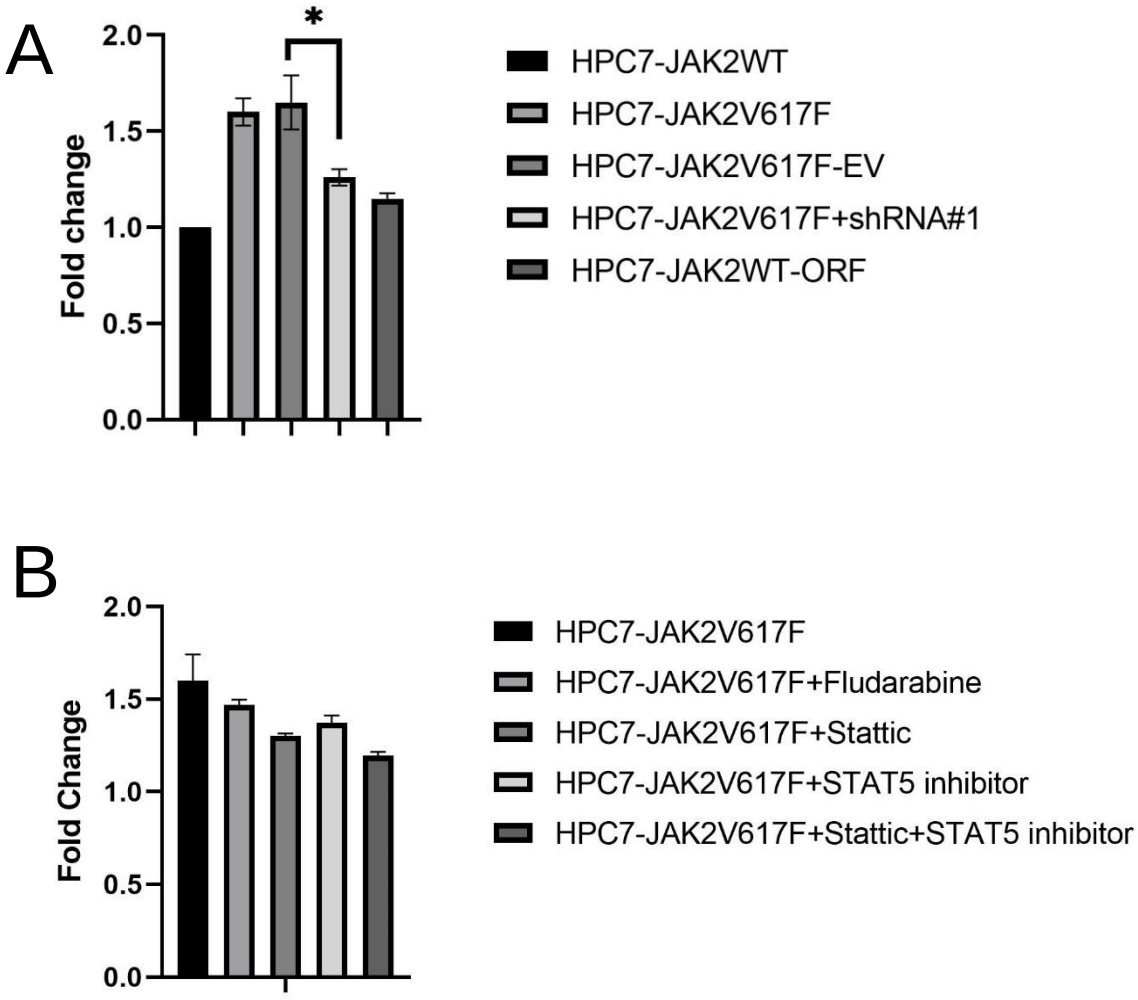


Figure 5.14 HDAC9 depletion reduces ROS levels through a mechanism other than STAT activation.

(A) Dihydroethidium assay measuring ROS levels in HDAC9 depleted HPC7-JAK2V617F cells shows that HDAC9 reduction leads to a reduction in ROS levels but overexpression of HDAC9 led to only a very small increase in ROS (B) A further ROS assay investigating STAT contribution to ROS levels indicates that none of the STATs are significant contributors to ROS levels.

Testing for statistical significance was performed using a student's t-test (*: $p<0.05$; **: $p<0.01$; *** $p<0.001$). Three biological replicates were performed.

In 2011, Kotian et al. provided evidence that HDAC9 is essential for homologous recombination. They found this by silencing each member of the HDAC family individually in HeLa-DR-13-9 cells (HeLa cells that stably express GFP) and then introduced a I-SceI expressing vector (which introduces DNA damage breaks in GFP); the amount of GFP expressed was equal to the effectiveness of the DNA being repaired by homologous-directed repair. They also found that HDAC9 removal re-sensitised cells to Mitomycin C treatment (Kotian et al., 2011). They hypothesised that HDAC9 may be indirectly involved with repair as after DNA damage was induced with ionising radiation, the location of HDAC9 in the nucleus remained unchanged.

The data from Kotian et al. are consistent with that in Chapter 3, and it has also been published that STAT3 has a role in DNA repair through the regulation of phospho-CHEK1 (Park et al., 2019). I therefore first wanted to investigate DNA damage repair in both the HDAC9-depleted HPC7-JAK2V617F cells and HPC7-JAK2WT-ORF cells using the same methods used in Section 3.3.5, Figure 3.11(**Figure 5.16**). The results suggest that HDAC9 may contribute to Mitomycin C/Olaparib sensitivity and DNA damage repair. Specifically, HPC7-JAK2V617F-shRNA#1 cells show a reduced capacity to repair DNA damage (**Figure 5.15**). However, overexpression of *HDAC9* in HPC7-JAK2WT cells only provided a small increase to cell viability following treatment with Olaparib and Mitomycin C relative to the HPC7-JAK2WT cells. This implies that overexpressing *HDAC9* in HPC7-JAK2WT is not enough to increase DNA repair capacity and the result observed in the HPC7-JAK2V617F-shRNA#1 cell line implies that removing *HDAC9* is reducing DNA repair capacity provided by expressing JAK2V617F through some unknown mechanism. This may be

because of the lack of phospho-STAT1 and phospho-STAT5 which have been shown to have roles in DNA repair (Wingelhofer et al., 2018; Meissl et al., 2017).

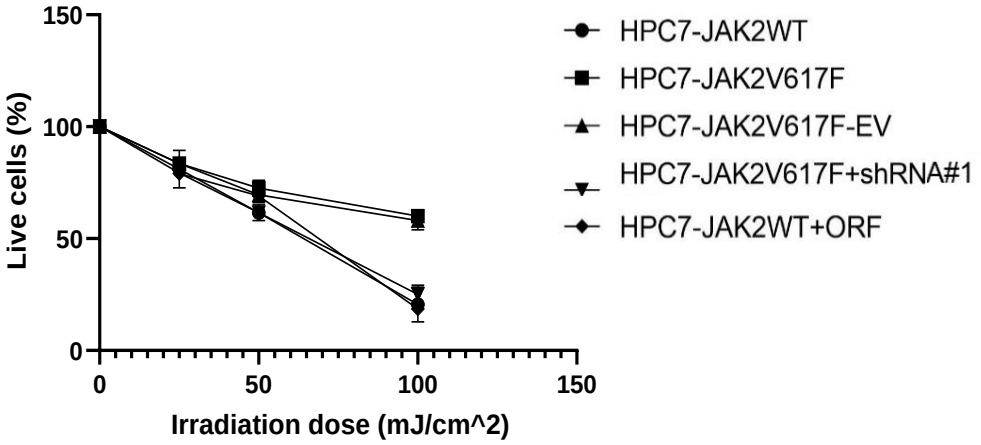


Figure 5.15 HDAC9 depletion removes the protection from UV exposure conferred by JAK2V617F.

HDAC9 depletion re-sensitises HPC7-JAK2V617F cells to an increased susceptibility to UV treatment. The lack of effect observed by overexpression of HDAC9 in HPC7-JAK2WT cells may be because of a lack of phospho-STAT1 and phospho-STAT5 which regulate genes involved in DNA repair.

Testing for statistical significance was performed using a student's t-test (*: $p<0.05$; **: $p<0.01$; *** $p<0.001$). Three biological replicates were performed.

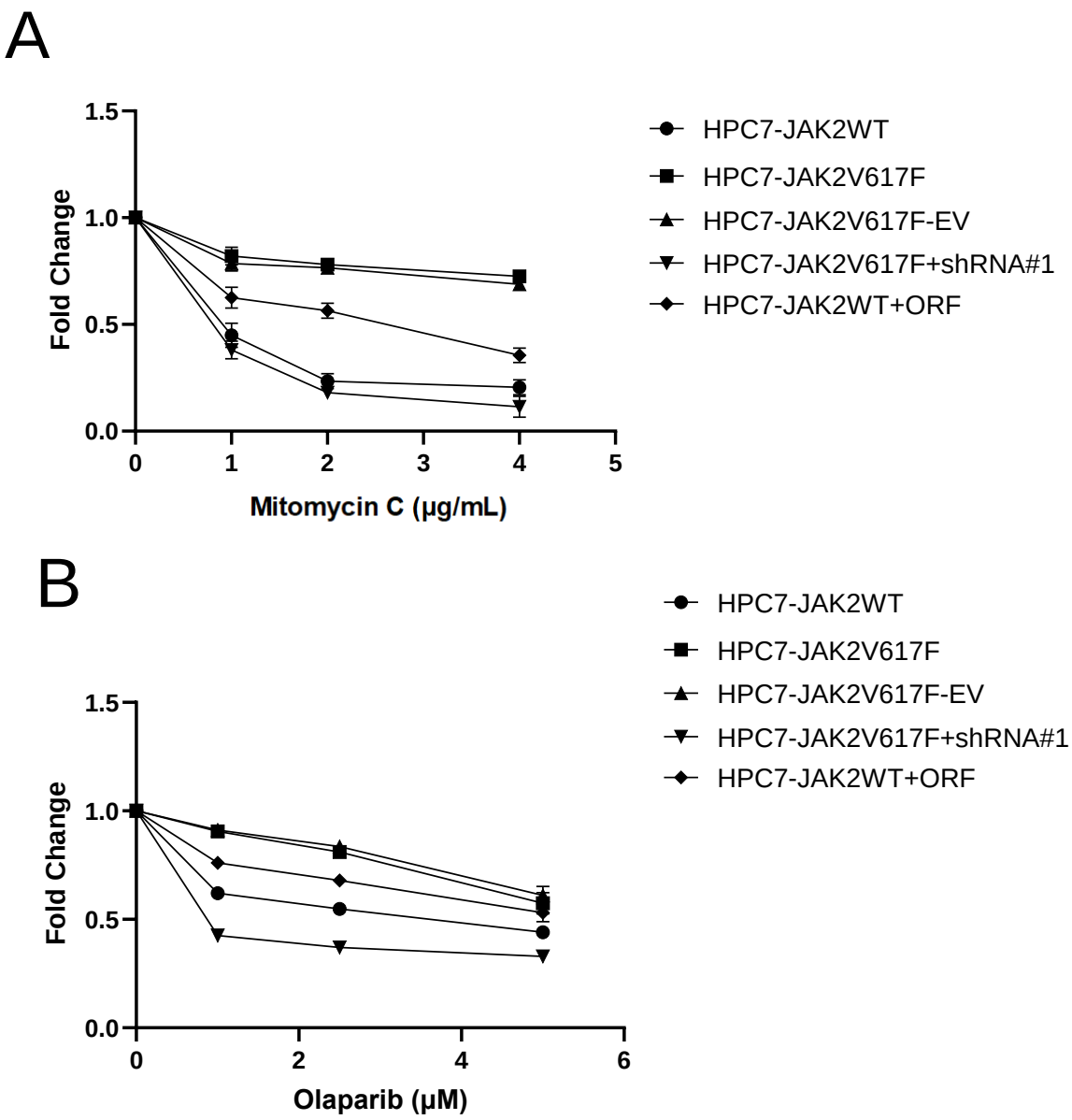


Figure 5.16 HDAC9 depletion eliminates the protection from Olaparib, and Mitomycin C conferred by JAK2V617F

HDAC9 depletion re-sensitises HPC7-JAK2V617F cells to (A) Mitomycin C and (B) Olaparib treatment resulting in increased levels of cell death relative to both HPC7-JAK2V617F and HPC7-JAK2WT. Overexpression of HDAC9 in HPC7-JAK2WT increased the levels of resistance to Olaparib.

Testing for statistical significance was performed using a student's t-test (*: $p<0.05$; **: $p<0.01$; *** $p<0.001$). Three biological replicates were performed.

Kotian et al. hypothesised that histone acetylation may contribute to the DNA damage repair response. There are many acetylation markers associated with DNA damage, but I decided to test the levels of H3K9Ac and H4K16Ac as they have been shown to be increased upon HDAC9 loss (Horikoshi et al., 2019; Meyer et al., 2016). To this end, I performed a western blot staining for H4K16Ac and H3K9Ac (**Figure 5.17A**). This indicates that in the HPC7-JAK2V617F cell line, there is a much lower level of H3K9Ac and H4K16Ac relative to the HPC7-JAK2WT and HPC7-JAK2V617F-shRNA#1 cell lines.

H4K16Ac and H3K9Ac are histone marks associated with transcriptional activation and I therefore next asked if DNA damage associated genes are dysregulated in the absence of HDAC9. I performed a western blot probing the levels of Phospho-γH2AX and phospho-CHEK1 (**Figure 5.17A**) and performed RT-qPCR to examine the transcript levels of *BRCA1* and *RAD51* **(Figure 5.17B**). The effects on phospho-γH2AX were negligible but, total CHEK1 and phospho-CHEK1 are significantly reduced in HPC7-JAK2V617F-shRNA#1 and increased in HPC7-JAK2WT+ORF.

RAD51 expression decreased but not *BRCA1* in the HPC7-JAK2V617F+shRNA#1 cells. Neither *RAD51* nor *BRCA1* gene expression is increased in the HPC7-JAK2WT-ORF cells suggesting that only HDAC9-depletion in HPC7-JAK2V617F alters *RAD51* or *BRCA1* expression. From these results, it can be concluded that HDAC9 can regulate increased expression in *CHEK1* which may explain the increased DNA repair.

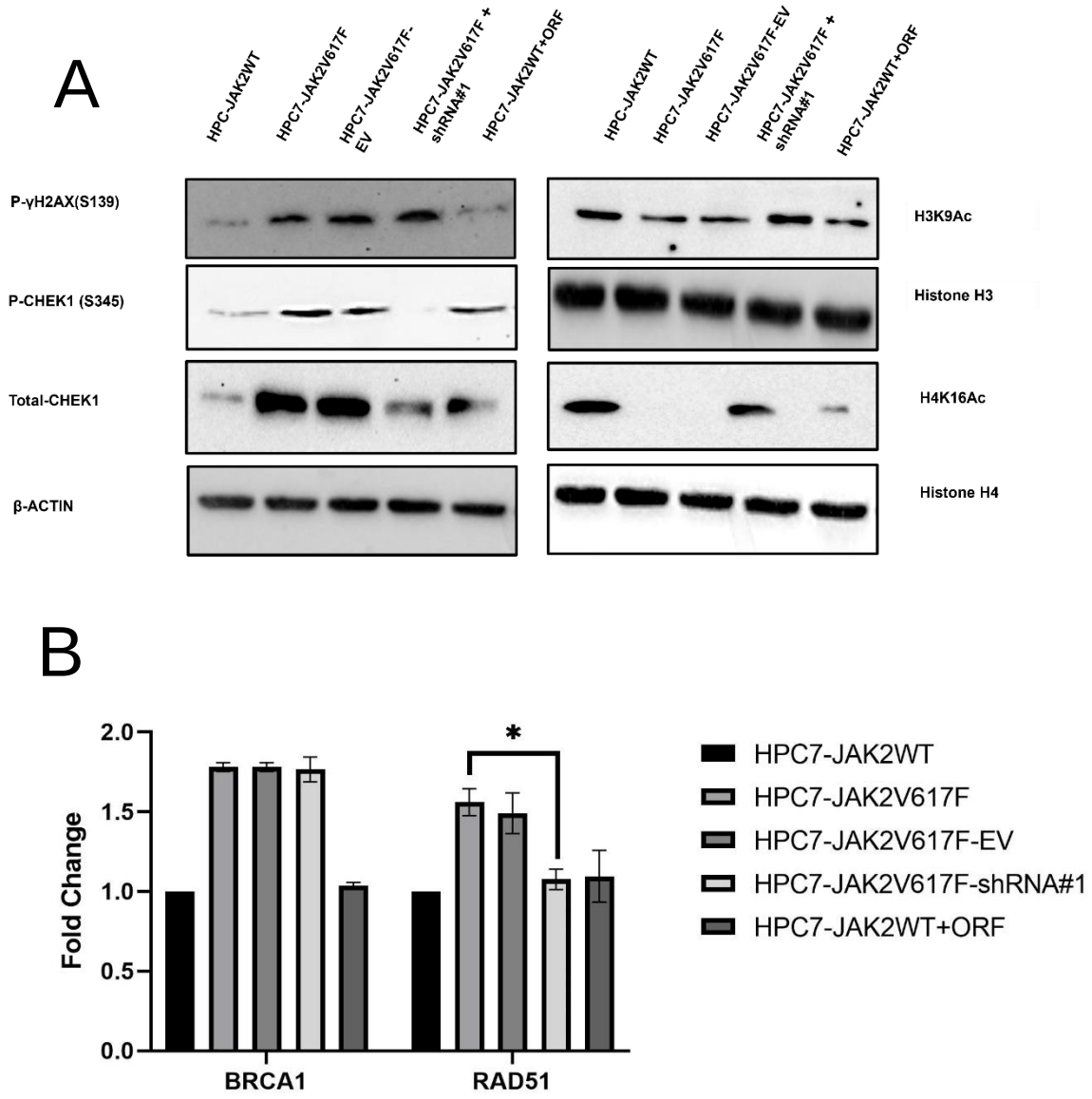


Figure 5.17 HDAC9 depletion reduces *CHEK1* and *RAD51* expression.

(A) Western blot probing the levels of phospho-γH2AX, CHEK1 and H4K16AC, CHEK1 and CHEK1 phosphorylation decrease when HDAC9 is silenced and CHEK1 is increased through the overexpression of HDAC9 in HPC7-JAK2WT. phospho-γH2AX is slightly increased. Levels of H4K16Ac increase when HDAC9 is knocked down and decrease when HDAC9 is overexpressed.

(B) RT-qPCR measuring expression levels of *RAD51* and *BRCA1*. *BRCA1* is unchanged by HDAC9 expression and *RAD51* expression is depleted upon HDAC9 reduction, but overexpression of HDAC9 alone is not enough to increase *RAD51* levels.

Testing for statistical significance was performed using a student's t-test (*: $p<0.05$; **: $p<0.01$; *** $p<0.001$). Three biological replicates were performed.

The presence of low levels of H4K16Ac/H3K9Ac in HPC7-JAK2V617F cells coupled with the increased expression of DNA damage repair genes *RAD51, BRCA1* and *CHEK1* in HPC7-JAK2V617F suggests that H4K16Ac may not be required for expression of these genes and that may be instead regulated by different histone modifications or through binding of a negative transcriptional regulator. CHEK1 is activated by phosphorylated ERK1/2 and AKT (Ray-David et al., 2013), which is consistent with my model, but more data are required. H3K9Ac is also reduced in the HPC7-JAK2V617F cell line although not to the levels of H4K16Ac and reverts back to HPC7-JAK2WT levels upon HDAC9 depletion.

Regarding DNA repair, it has been shown that H4K16Ac/H3K9Ac deacetylation is necessary for DNA repair to occur, possibly acting through crosstalk with other modifications (Horikoshi et al., 2019). Another possible theory may be that H4K16Ac/H3K9Ac is associated with transcriptionally active DNA and, whilst being transcribed, DNA is difficult to access for repair thereby warranting the need for reducing transcriptional activity via removal of an acetyl group by HDAC9/Class I HDAC complex.

CHEK1 activation allows stalling in G2-M, a phase in the cell cycle that allows repair of DNA damage accumulated during DNA replication. In Chapter 3, I found that more of my HPC7-JAK2V617F cells are in this cell cycle phase than HPC7-JAK2WT cells. Therefore, I repeated the cell cycle analysis on the HPC7-JAK2V617F-shRNA#1 cell line to test if this change is abrogated by loss of *HDAC9*. The depletion of *HDAC9* reduced the number of cells in the G2-M phase of the cell cycle (**Figure 5.18**).

Finally, the HDAC family forms protein complexes that can regulate proteins through catalysing deacetylation and HDAC9 is a necessary scaffold protein component required for these complexes to form. I therefore performed an IP using anti-FLAG to pull down FLAG-HDAC9 from HPC7-JAK2WT-ORF cells. After the “pull-down”, the samples were analysed by the mass spectrometry facility to reveal the binding partners of HDAC9. As a control, to remove non-specific binding, I overexpressed a vector that expressed a FLAG tag only in HPC7-JAK2WT (a generous gift from Dr Cintli Morales Alcala). The IP-mass spectrometry analysis, in **Table 5,** revealed that HDAC9 does not bind DNA damage proteins in HPC7-JAK2WT-ORF cells but binds other HDAC members and transcriptional regulators. **Table 5** suggests that HDAC9 contributes to DNA damage repair through binding other proteins which either deacetylate histones or bind epigenetic regulators that alter gene transcription.

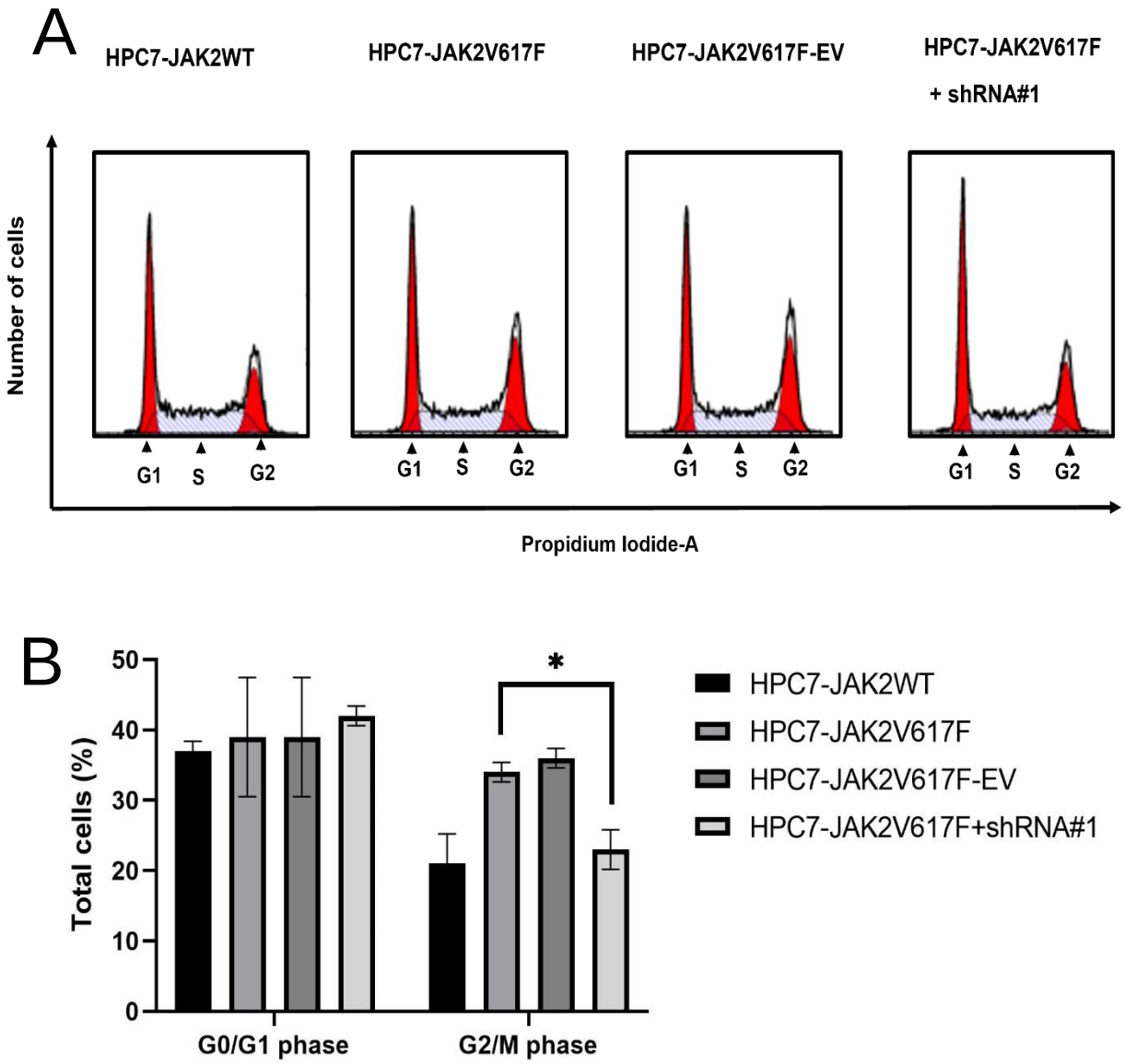


Figure 5.18 HDAC9 depletion decreases number of cells in the G2 phase of the cell cycle.

Cell cycle analysis reveals a reduction in the number of cells in the G2 phase of HDAC9 depleted cells. The reduction could be explained by the reduced expression of *CCND2* and/or *CHEK1* which can both regulate the G2 phase. The G2 phase is required to allow DNA to repair after possible damage during the S-phase**.** A) Three biological replicates were performed, and the images are the most representative of the average. B) A graph representing a summary of the three biological replicates split between the G0/G1 phase and the G2/M phase.

Testing for statistical significance was performed using a student's t-test (*: $p<0.05$; **: $p<0.01$; *** $p<0.001$).

Accession	#Unique peptides	Gene Name	Description
E9Q6J5|MEF2_MOUSE	39	MEF2D	Myocyte specific enhancer factor 2D
Q6PDK2|KMT2D_MOUSE	31	KMT2D	Histone-lysine N-methyltransferase 2D
O70546|KDM6A_MOUSE	29	KDM6A	Lysine-specific demethylase 6A
tr|Q8CBV1|Q8CBV1_MOUSE	26	ETV6	ETS-variant transcription factor 6
tr|Q5PPQ7|Q5PPQ7_MOUSE	21	CORO1C	Coronin 1C
Q8VIJ6|SFPQ_MOUSE	19	SFPQ	Splicing factor proline- and glutamine-rich
E9PYH6|SET1A_MOUSE	18	SETD1A	Histone-lysine N-methyltransferase
tr|E9PU93|E9PU93_MOUSE	18	ASH2L	Set1/Ash2 histone methyltransferase complex subunit ASH2
Q61069|USF1_MOUSE	17	USF1	Upstream stimulating factor 1
Q9CWW7|CXXC1_MOUSE	16	CXXC1	CXXC-type zinc finger protein 1
P02301 | H3_MOUSE	16	H3.3	Histone H3.3
tr|G3UWD2|G3UWD2_MOUSE	15	RUNX1	Runt-related transcription factor
Q5RIM6| NCOR1_MOUSE	15	NCOR1	Nuclear co-repressor 1
P70288| HDAC2_MOUSE	13	HDAC2	Histone deacetylase 2
P15864|H12_MOUSE	13	H1.2	Histone H1.2
A0A0A0MQ79| PRRC2C_MOUSE	12	PRRC2C	Protein PRRC2C
P62806|H4_MOUSE	11	H4C1	Histone H4 C1
Q9Z2X1|HNRPF_MOUSE	10	HNRNP	Heterogeneous nuclear ribonucleoprotein
O08550|KMT2B_MOUSE	10	KMT2B	Histone-lysine N-methyltransferase 2B
Q6NZM9 | HDAC4_MOUSE	10	HDAC4	Histone deacetylase 4
O09106| HDAC1_MOUSE	8	HDAC1	Histone deacetylase 1
O88712|CTBP_MOUSE	8	CTBP	C-terminal binding protein
P63101|1433Z_MOUSE	8	YWHAZ	14-3-3 protein zeta/delta
Q3UM33| HDAC3_MOUSE	6	HDAC3	Histone deacetylase 3
Q61033|LAP2A_MOUSE	3	TMPO	Lamina-associated polypeptide 2 isoforms alpha/zeta
Q91ZR3|ZHANG_MOUSE	3	CREBZF	CREB/ATF bZIP transcription factor
Q99N13 | HDAC9_MOUSE	2	HDAC9	Histone deacetylase 9

Table 5 List of proteins that bind to FLAG-tagged HDAC9.

The results of the IP-Mass Spectrometry were achieved by pulling down HDAC9-FLAG with a FLAG antibody. The results indicate a large list of transcriptional regulators/regulators of histone markers. The list shows that no DNA damage proteins are bound and therefore any effect on DNA damage repair is indirect.

Interestingly, HDAC9 is low on the list of identified proteins by IP-Mass Spectrometry analysis. This may be explained by the methodology by which the programme identifies the proteins that have been pulled down. The PEAKs programme identifies proteins based on unique peptide sequences generated post protein degradation by trypsin. HDAC9 has high homology to other Class II HDACs, and this may explain an underrepresentation of unique peptides for identification.

5.3.3 MEF2 binding, CTBP1 binding, HDAC glutamine-rich N-terminus domain and a highly disordered C-terminal region are integral for HDAC9-Class I HDAC activity.

Mouse histone deacetylase 9 is devoid of histone deacetylase activity and instead, functions as a “scaffold” that binds other proteins and directs them to the target destination. HDAC9 has only one named domain, referred to as the histone deacetylase glutamine-rich N-terminal domain which, as the name suggests, is a glutamine-rich domain near the N-terminus that takes the shape of a coiled-coil. Coiled-coils can promote the formation of protein complexes, although this has been shown to be uncommon, and they are thought to be more of a molecular spacer between binding proteins (Truebestein and Leonard, 2016). The rest of the HDAC9 sequence is comprised of disordered regions that undergo conformational changes upon binding to other more structurally ordered proteins.

I wanted to investigate where Class I HDAC proteins bind on HDAC9. To achieve this, I used data from the UniProt database (The UniProt, 2021) to highlight the regions in which proteins are predicted to bind to HDAC9, took the HDAC9-ORF vector and introduced truncations into the vector using the Q5® site-directed mutagenesis kit from New England Biolabs® inc. (NEB) referring to the truncations as Truncated Open-Reading Frames (TORFs); I have represented this graphically (**Figure 5.19A**). I took the truncated vectors and verified them through sequencing (data not shown) and transiently transfected them into the HPC7-JAK2WT cell line.

Overexpression of the full-length HDAC9-ORF in HPC7-JAK2WT cells resulted in reduced H4K16Ac levels upon transfection (**Figure 5.17**), and this is likely to

be caused by binding other HDACs. Therefore, the truncation(s) that reveal an increase in H4K16Ac levels will indicate where proteins necessary for histone deacetylation bind, including the Class I HDAC family, with H4K16Ac (**Figure 5.19B**).

The western blot in Figure 5.19 suggests that truncated HDAC9 variants TORF2, TORF4 and TORF5 are important regions for H4K16Ac deacetylation. Following on from Figure 5.19, I used TORF2, TORF4 and TORF5 mutants to investigate if they are necessary for the increased phosphorylation of STAT3, AKT or ERK1/2. To this end, I transfected the HPC7-JAK2WT cells with either TORF2, TORF4 or TORF5 cDNA and then performed western blots for phospho-STAT3, phospho-ERK1/2 and Phospho-AKT. The HPC7-JAK2WT cells were cultured in reduced MSCF conditions (1 µg/ml of MSCF) during the transfection experiment to allow for the differences in STAT3 phosphorylation to be more identifiable (**Figure 5.20**).

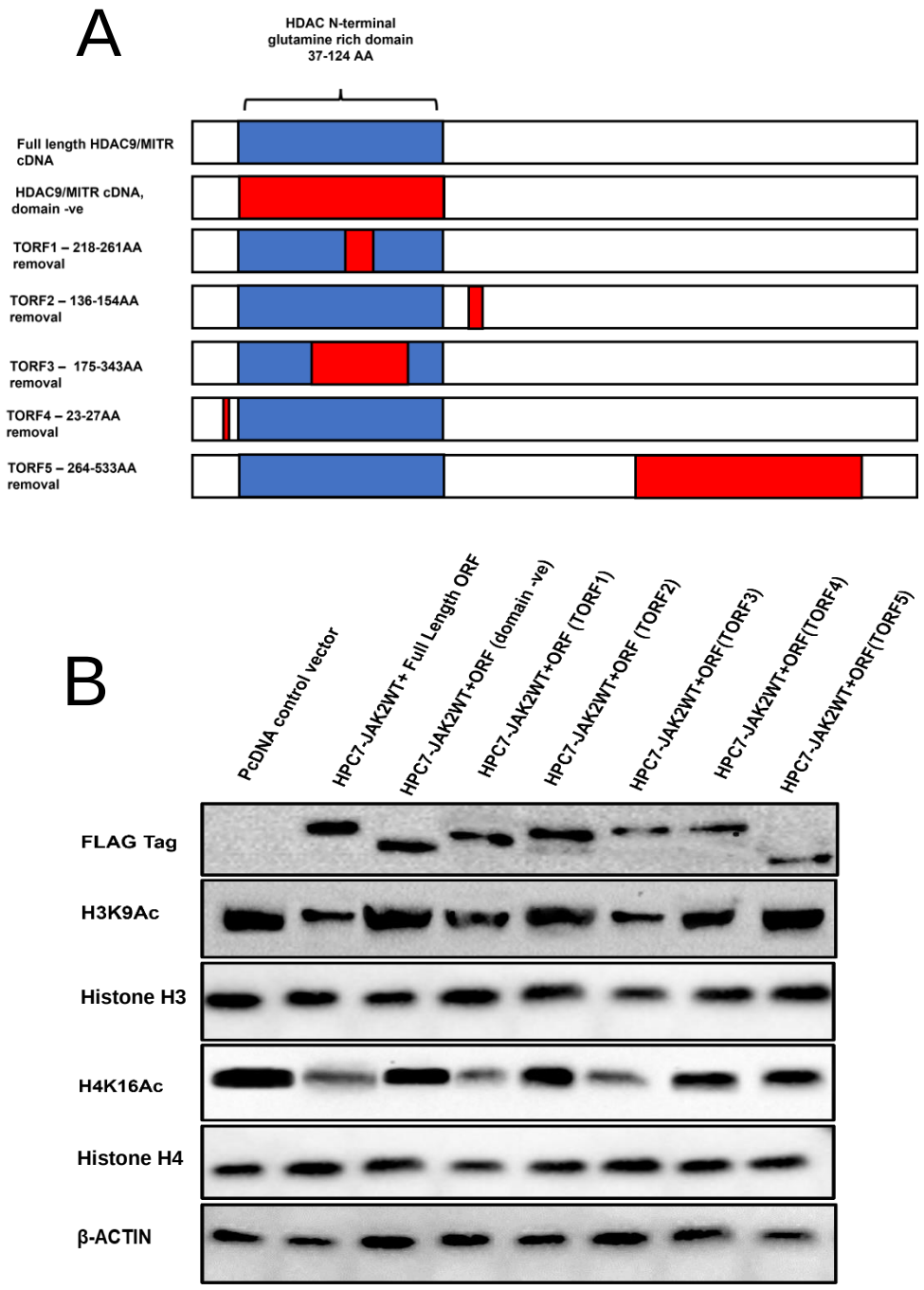


Figure 5.19 TORF 2, 4, 5 and the domain of the HDAC9-ORF are necessary for histone deacetylation

(A) A graphical illustration representing the areas of the HDAC9-ORF I have removed (red) and the domain (blue). (B) The histone deacetylase glutamine-rich N-terminal domain is important for HDAC activity. This activity relies on regions deleted in TORF2, TORF3 and TORF5.

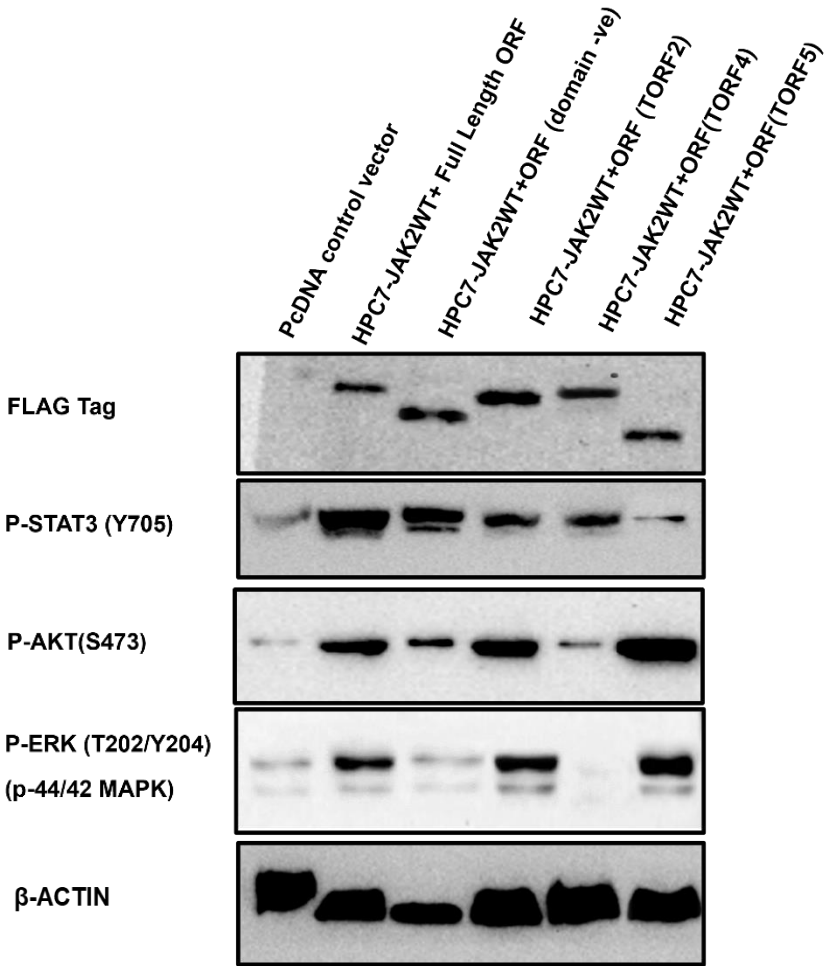


Figure 5.20 Different regions of HDAC9 are required for phospho-signalling.

Western blot of phospho-STAT3, ERK1/2 and AKT. Different parts of HDAC9 appear to be necessary for different signals. TORFs 2, 4 and 5 were unable to induce high levels of STAT3 phosphorylation, with the TORF5 deletion mutant showing the greatest reduction. Phosphorylation of AKT and ERK1/2 appears to require the residues 23-27 missing in TORF4.

The results of Figure 5.19 suggest that regions deleted in second, fourth and fifth TORFs are important for histone deacetylation. TORF2 has been shown to lack regions that bind MEF2 proteins which are required for binding DNA and repressor complexes by Class II HDAC members (Han et al., 2005), TORF4 encompasses the area known to bind CTBP1 (Han et al., 2005), a protein which allows the binding of NCOR2 protein complexes to DNA. I hypothesise that CTBP1 and MEF2 contribute to signalling through the binding of DNA and/or NCOR2 repressor complex, which directs HDAC9-bound Class I HDAC members towards the target histone residue. TORF1 and TORF3 also have a modest effect on histone acetylation but this may be due to the deletion being in the middle of the domain and disrupting the 3D-structure of the protein and, by default, deacetylation activity.

TORF5 represents a highly disordered region (predicted by InterPro online tool (Blum et al., 2021)) towards the C-terminus of HDAC9. This region is under-investigated in HDAC9 but has been shown in other Class II HDACs to serve as a protein-interaction interphase. In Figure 5.19, I show that when this disordered region is removed histone deacetylation does not occur indicating that it likely binds proteins with deacetylase activity. TORF1 and TORF3 lack regions within the HDAC glutamine-rich N-terminus domain and from Figure 5.19 it can be concluded that TORF1 and TORF3 does not contribute to H4K16Ac deacetylation, but the entire removal of the domain does. I hypothesise that the coiled-coil domain is contributing to the overall 3D structure of the protein, protein complex formation or the molecular “spacer” rather than being involved directly with deacetylases.

The western blots for phospho-STAT3, ERK1/2 and AKT show that phospho-STAT3 shows a similar pattern to the data seen for H4K16Ac where the removal of the HDAC glutamine-rich N-terminus domain, MEF2-binding region (TORF2), CTBP1-binding region (TORF4) and the fragment missing in TORF5 contribute to phospho-STAT3 signalling. The Figures 5.19 and 5.20 suggest that the region deleted in TORF5 of HDAC9 is required to bind proteins with deacetylase activity and these proteins may be what catalyses deacetylation of histone residues and this deacetylation of histones may be what regulates STAT3 phosphorylation probably through *SOCS3* expression.

5.4 Discussion

The aim of this Chapter was to understand the role of HDAC9 in HPC7-JAK2V617F cells and how this may contribute to disease. Some answers have been provided, but it is clear that many questions remain unanswered as discussed below.

In Section 5.5.1, I successfully validated that JAK2V617F drives increased expression of HDAC9 in the HPC7 cell line through phosphorylation of STAT3 and this in turn leads to increased phosphorylation of STAT3 through the suppression of SOCS3 expression and possibly EGFR signalling. HDAC9 is necessary for the phosphorylation of STAT3, ERK1/2 and AKT whilst having no effect on the level of JAK2 phosphorylation. I also confirmed that HDAC9 has a role in DNA damage repair through the upregulation of some DNA damage repair proteins.

Using coimmunoprecipitation and mass spectrometry, I have shown that HDAC9 interacts with HDAC1, HDAC2, HDAC3 and HDAC4. I can hypothesise that HDAC9 forms a protein complex (probably with the NCOR2 repressor complex) with Class I HDAC proteins and perhaps acts as a coordinator to repress global acetylation of histones H3 and H4.

In this study, I confirmed that EGFR signalling is upregulated in HPC7-JAK2V617F and HPC7-JAK2WT-ORF cells which also contributes to increased phosphorylation of STAT3, AKT and ERK1/2. The increased EGFR signalling could be regulated through increased *WWTR1* expression. The levels of ERK1/2 phosphorylation could also be promoted by the downregulation of *DUSP2* expression driven by a loss of acetylation on histones H3 and H4.

HDAC9 knock-down in HPC7-JAK2V617F cells ameliorated the proliferative capacity of HPC7-JAK2V617F. If I considered my original theory it would suggest that *c-MYC* expression is dysregulated and so is proliferation/HSC expansion, but it must be noted that phospho-STAT3, phospho-AKT and phospho-ERK1/2 are proliferation factors and are known to contribute to cell proliferation/self-renewal (Hirai et al., 2011) and it could be considered specious to assume *c-MYC* is involved without more data which could generated through the silencing of *c-MYC* expression in the HPC7-JAK2WT cell line.

Towards the end of Section 1 of Chapter 5, I briefly investigated the localisation of HDAC9 and the tyrosine phosphorylation status of HDAC9 to provide an interesting basis for further investigation into how HDAC9 tyrosine phosphorylation may contribute the cellular signalling and how JAK2V617F may contribute to cellular localisation of HDAC9. The results provide an interesting foundation for a new project as the role of HDAC9 tyrosine residue phosphorylation in JAK2V617F expressing cells has not been investigated. I would be interested to know if the decreased tyrosine phosphorylation of HDAC9 triggers increased nuclear isolation. It would be of interest to know which sites (if any) are necessary for HDAC9 function/location and whether tyrosine phosphorylation contributes to protein binding/specific protein complex formation. The Figure provided in the thesis only shows total tyrosine phosphorylation.

To date, the role of HDAC9 in the DNA damage response is unclear. Here, I show that HDAC9 upregulates *CHEK1*, which encodes a DNA repair protein. I showed that HDAC9 depletion re-sensitises the HPC7-JAK2V617F cells to Mitomycin C and Olaparib whilst also reducing the efficacy of UV-induced damage repair with expression being reduced back towards JAK2WT levels

coinciding with a reduction in ROS levels. It is difficult to conclude however if this is response is down to a reduction of HDAC9 or indirectly as a reduction of phospho-STAT/AKT/ERK1/2 which is a result of HDAC9 depletion.

It is still unclear whether H4K16Ac and H3K9Ac are necessary for DNA repair. HDAC9 can deacetylate many residues on histones H3 and H4, I have shown that HDAC9 can also bind histones H1 and H2A through mass spectrometry; therefore, currently the mechanism is incomplete since only two residues were investigated in the thesis. The reduction of HDAC9 in HPC7-JAK2V617F cells only slightly increased phospho-γH2AX levels despite a decrease in ROS which may suggest other contributing factors. It may also be that there are fewer HPC7-JAK2V617F-shRNA#1 cells in the G2 phase of the cell cycle due to the reduction of CHEK1 and as a result reducing the number of cells undergoing repair post S phase or possibly it may be that the levels of γH2AX phosphorylation are saturated in the HPC7-JAK2V617F cells and it's not possible to observe any greater increase.

The results taken together may suggest that HDAC9 is upregulated by STAT3 act as a DNA damage response gene; this is highlighted by the HDAC9-depleted HPC7-JAK2V617F cell line's reduced response to Mitomycin C/Olaparib treatment.

In my final section, I found that the histone deacetylase glutamine-rich N-terminal domain and TORFs 2,4 and 5 are integral for histone deacetylation and phosphorylation of STAT3. However, for the phosphorylation of AKT and ERK1/2, the histone deacetylase glutamine-rich N-terminal domain and the region removed in TORF4 are necessary.

Chapter 6

General discussion

Chapter 6 General discussion

MPNs are haematological malignancies originating from a single haematopoietic stem cell clone that develops mutations which drive a progression towards chronic proliferation and differentiation (Nangalia and Green, 2014). The discovery of *JAK2*, *CALR* and *MPL* mutations have allowed us to understand the mechanism by which these disorders phenotypically present themselves and how they develop over time (Nangalia et al., 2013; Klampfl et al., 2013; Beer et al., 2008; Pikman et al., 2006; Levine et al., 2005; Kralovics et al., 2005; James et al., 2005; Baxter et al., 2005). These driver mutations are mutually exclusive (Nangalia and Green, 2014) and each mutation shows different disease phenotypes despite all three activating the same JAK-STAT pathway. The mechanisms by which the driver mutations contribute to disease are still unclear and therefore it is important to elucidate the varying protein networks contributing to disease, to generate new treatments to improve patient quality of life and survival chances.

JAK2V617F is the most common mutation found in MPN patients with ET and MF patients having a 40-60% chance of having JAK2V617F and PV patients a 90% chance (Arber et al., 2016). JAK2V617F signalling can be inhibited with Ruxolitinib treatment, but resistance to Ruxolitinib accumulates in patients after 3-6 weeks (Koppikar et al., 2012). To combat this resistance to Ruxolitinib it is necessary to find new avenues of treatment.

To identify new genes, I used the HPC7 cell line which is a mouse erythroid cell line with a genetic profile and immunophenotype with high homology to a mouse HSC. The use of this model was necessary for the convenience and simplicity of emulating JAK2V617F in HSCs without the need for animal *in vivo* models or patient samples. Patient samples have the issue of expressing secondary mutations or the patient being currently under treatment, whereas HSCs in an *in vivo* environment have the additional problem of being influenced by the other cells in the haematopoietic stem cell niche or of the haematopoietic system.

The model created in Chapter 3 gave a robust cellular response, where JAK2V617F expression generated a phenotype that matched published data. Through investigation of the HPC7-JAK2V617F cells, I found that JAK2V617F expression gave some resilience to DNA damage and external genotoxic insults which would explain the cytogenetically stable nature of patients HSCs expressing this mutation. In my model, I also identified c-MYC as protein of interest. The c-*MYC* gene encodes a transcription factor associated with proliferation and HSC exhaustion. In my model, I found that *c-MYC* expression was downregulated in JAK2V617F. A deficiency of c-MYC in HSCs triggers HSC exhaustion which is a phenotype observed in JAK2V617F knock-in mice (Kent et al., 2013). Although I did not observe exhaustion in my cell line, I predict that was caused by the erythroid/MSCF-dependent nature of cell line and it may be more valuable to explore the role of c-MYC expression in a genuine HSC cell expressing JAK2V617F.

In Chapter 4, I analysed RNA-sequencing data from the HPC7-JAK2V617F cells to identify genes that may contribute to JAK2V617F-driven phenotype observed in Chapter 3. In this Chapter, I demonstrated that the gene expression

profile resembles PV patient data with many of the same genes and signalling pathways being upregulated. I identified upregulated genes that may contribute to an increased metabolic demand, and DNA damage repair in the HPC7-JAK2V617F cells. The results of Chapter 4 highlighted that the components of the NCOR2 repressor protein complex are upregulated with the HDAC9 protein being the most upregulated component in my data and in other datasets. Finally, I investigated the published datasets exploring secondary mutations commonly found alongside JAK2V617F expressing cells. The results indicate that different mutations alter HDAC9 expression and DNMT3a has a significantly negative impact on HDAC9 protein expression.

In Chapter 5, I explored the role of HDAC9 upregulation in JAK2V617F expressing HSCs. I demonstrated that activated STAT3 upregulates *HDAC9* expression and subsequently HDAC9 regulates STAT3, ERK1/2 and AKT activation. I demonstrated that HDAC9 regulates *SOCS3* expression which likely controls the levels of STAT3 phosphorylation. How HDAC9 regulates ERK1/2 and AKT is less clear. However, I have shown that EGFR activation is increased in JAK2V617F expressing cells possibly through *WWTR1* upregulation and I have also shown that HDAC9 regulates *DUSP2* expression which may contribute to increased ERK1/2 phosphorylation through the reduction of a negative feedback loop by DUSP proteins.

I have also demonstrated that HDAC9 has role in DNA damage repair through the regulation of CHEK1, *BRCA1* and *RAD51* expression. As HDAC9 is associated with transcriptional repression, I don't think that HDAC9 directly regulates the expression of *CHEK1* and *RAD51*, but rather I hypothesise that HDAC9 indirectly regulates *CHEK1* and *RAD51* expression through STAT3, AKT or ERK1/2 signalling.

HDAC9 upregulation decreases the level of histone H3 and histone H4 acetylation which would trigger global transcriptional repression. It is also possible that global deacetylation has a protective effect by condensing the chromatin and preventing chromatin related activity which may reduce DNA damage. In a recent paper, Hsiao et al. found that deacetylation of H4K16 is essential for proper binding of 53BP1 (a DNA repair protein) to a neighbouring histone marker H4K20Me2 (Hsiao and Mizzen, 2013). This suggests that histone crosstalk may play a role in DNA damage repair in the HPC7-JAK2V617F-shRNA#1 cells.

Finally, I demonstrated that HDAC9 needs its highly disordered C-terminal region, HDAC glutamine-rich N-terminal domain, CTBP and MEF2 binding domains for effective deacetylation and regulation of the changes observed in Chapter 5.

In the broader context of MPNs, the increased expression of DNA damage response proteins in JAK2V617F-expressing cells may be a result of increased *HDAC9* expression. This may explain why JAK2V617F-expressing patients are cytogenetically stable for decades and, why many MPN patients only begin to present an observable phenotype when a secondary mutation has occurred (Scott, L. M. and Rebel, 2012).

If the upregulation of HDAC9 expression is responsible for the increased resistance to Mitomycin C and Olaparib treatment and the regulation of DNA damage repair proteins, then HDAC9 inhibition may have a therapeutic role in the treatment of MPNs. The inhibition of HDAC9 is difficult as HDAC9 has no catalytic domain. In this work, Section 5.3.3 shows that different regions of the HDAC9 protein have different roles and potentiate different pathways, I

hypothesise that one of these pathways may be upregulating DNA damage repair proteins and by disrupting the region of HDAC9 responsible, and in combination with other treatments, it may be possible to trigger cell death of the JAK2V617F-expressing cells through the accumulation of DNA damage.

The most significant question left open by this thesis is that if HDAC9 has no deacetylase activity and catalyses deacetylation by binding through Class I HDACs, and there is no increase in Class I HDAC expression, then how is HDAC9 having such a significant effect on histone acetylation levels? I postulate that perhaps the HDAC family have a high binding affinity for the HDAC9/NCOR2 complex or *vice versa* which “pulls” the Class I HDAC away from its normal cellular location to function as the catalytic component of the NCOR2 repressor protein complex. To test the theory, future work could include coimmunoprecipitation experiments in the HPC7-JAK2WT and HPC7-JAK2V617F cell lines with antibodies targeting each individual member of the HDAC family, followed by probing in a western blot with an anti-HDAC9 antibody.

The second significant question posed by Chapter 4 is that HDAC9 is not always upregulated in JAK2V617F-expressing patients and is even sometimes downregulated. This raises the question of how relevant the data in this thesis is to human patients? This is a difficult question to discuss without more data but looking through the PV datasets published in chapter 4 and RT-qPCRs of patient data (Skov et al., 2012) revealed that some of the patients demonstrate increased expression of other HDAC family members with 4, 5, 6 and 8 being upregulated to the same degree as HDAC9. These HDAC members are also Class II HDACs which act through binding Class I HDACs and target the same residues as HDAC9. Therefore, it is possible that Class II HDACs alone or in

combination may have some functional overlap and providing that histones are deacetylated the same results could be generated.

One of the main challenges during this project has been the lack of an effective HDAC9 antibody for immunoprecipitation and this is probably unlikely to change due to the highly homologous nature of Class IIa HDACs. This unfortunately means that I could not perform a ChIP-qPCR/ ChIP-Sequencing and unfortunately, I did not have the budget to perform ChIP-qPCR/ChIP-Sequencing which leaves questions unanswered. Finally, I would have liked to have investigated replication fork-stalling in HDAC9-depleted JAK2V617F cells as I predict that fork-stalling would have increased due to the increases in histone deacetylation seen in HDAC9-depleted HPC7-JAK2V617F cell lines.

This thesis provides new insights into how HDAC9 may contribute to the pathology of MPNs instigated through JAK2V617F mutation and has elucidated that HDAC9 is upregulated by phosphorylated STAT3 to downregulate negative regulators of phospho-protein signalling, promote EGFR signalling, and promote DNA damage repair of cells expressing JAK2V617F.

Appendices

List of upregulated genes in HPC7-JAK2V617F (> or equal to 1 log_2 (Fold change))

Gene_ID	Log2 (fold change)
PF4	*4.371577507*
DKK1	*3.484086406*
CORO2A	*3.05863199*
ETV5	*2.977904315*
SLC2A3	*2.749286578*
RASGEF1B	*2.627483318*
CDKN2A	*2.60023272*
SLC35D3	*2.511933653*
EFNB1	*2.44971004*
PFKP	*2.436903515*
NRGN	*2.383694382*
IFIH1	*2.327793651*
NUAK1	*2.290178841*
BHLHE40	*2.233528338*
CCND2	*2.21571689*
NDRG1	*2.184683862*
DZIP3	*2.166497022*
GEM	*2.076361607*
APLP2	*2.074552154*
LILR4B	*2.002775569*
TMSB4X	*1.937322995*
PSD3	*1.921762536*
UNC13A	*1.921579195*
FAM162A	*1.861070128*
ERO1L	*1.837760305*
PGK1	*1.834864093*
SLC2A1	*1.826815604*
RGS1	*1.760612765*
GCH1	*1.751090601*
TPI1	*1.726327784*
LILRB4A	*1.688397698*
NEDD4L	*1.672780446*
AMOT	*1.667362452*
ME2	*1.655068119*
NAB2	*1.643646598*
DLG3	*1.63661666*
UBASH3B	*1.625910237*
RAD51AP1	*1.625404113*
CIP2A	*1.619088636*
ACSL4	*1.613388785*

RAI14	*1.611562961*
BBX	*1.590124692*
PDK3	*1.565119325*
FAR2	*1.535573142*
RASA2	*1.533674237*
POLA1	*1.53197534*
BNIP3	*1.528823544*
CPD	*1.522268709*
CDCA3	*1.513257994*
NGLY1	*1.512246307*
KIF4	*1.502750925*
CARD19	*1.493285193*
NAA50	*1.483497266*
LCP2	*1.47361846*
TXN1	*1.470617118*
PMVK	*1.464585999*
YBX3	*1.454838625*
GAPDH	*1.440660924*
USF3	*1.434506704*
ANGPT1	*1.429264656*
MTMR1	*1.427686643*
MCU	*1.424721095*
DUSP6	*1.415244755*
F2R	*1.406612518*
P4HA1	*1.395670286*
SLCO4A1	*1.392799596*
ATP6V1A	*1.383779286*
HHIP	*1.382772417*
NCAPG	*1.374463209*
CENPI	*1.372583364*
UCHL3	*1.368812872*
POLQ	*1.364781582*
FOXM1	*1.364710359*
ALG13	*1.359631042*
LPCAT3	*1.35802084*
EGLN3	*1.356889591*
RFC4	*1.34473478*
CAPG	*1.33839108*
HMGN5	*1.331093266*
HIGD1A	*1.329102*
TLE1	*1.322164405*
PEX26	*1.320675723*
MUC13	*1.313336124*
KIF1C	*1.305175731*
STAP1	*1.304784344*
DHFR	*1.301487528*
KNTC1	*1.298764843*
MYO7A	*1.297768625*
CDC45	*1.297006366*

SMC1A	*1.294769906*
CD9	*1.29464832*
TFRC	*1.265192082*
RRM2	*1.260446984*
PLOD2	*1.251469389*
SLC12A2	*1.237001043*
NAMPT	*1.2309314*
EPB41L2	*1.229066514*
HSD17B10	*1.221729807*
MRPL51	*1.217268368*
MPHOSPH6	*1.216719575*
ZFP57	*1.213685896*
CLCN3	*1.212889241*
IFT122	*1.208804912*
FAM102A	*1.20835034*
TTBK1	*1.207506632*
GINS2	*1.206987125*
SLC16A1	*1.205718092*
MVD	*1.205444034*
LSS	*1.204857158*
SNAP47	*1.203676863*
CLTB	*1.202761314*
HELLS	*1.200557126*
MORF4L2	*1.197720315*
PIM2	*1.19109677*
ATP2B1	*1.18805882*
SELENBP1	*1.186159078*
USP45	*1.185742399*
GK	*1.184464757*
DDIT4	*1.183432724*
RGS18	*1.183077885*
JUN	*1.180140163*
SHROOM4	*1.179477168*
CCDC58	*1.173527093*
AMMECR1	*1.160682363*
HMMR	*1.160080118*
SLC43A3	*1.155467972*
TMX3	*1.155182972*
MED21	*1.154753304*
PKM	*1.151397632*
SERPINB1 A	*1.15088236*
FOXN2	*1.143186354*
SPRED2	*1.142497012*
KLF6	*1.140558916*
CDC25C	*1.139496811*
NDUFA9	*1.139386721*
PLA2G12A	*1.138947945*
PDE4DIP	*1.136423609*

SNX25	*1.135407587*
DDX58	*1.133859584*
HOOK1	*1.13357874*
GNG11	*1.133521868*
SELL	*1.132034964*
FUBP1	*1.128603424*
UBL4A	*1.127950681*
SLC20A1	*1.126409312*
RLIM	*1.124860322*
SLC7A5	*1.12459895*
LRP8	*1.124374535*
PJA1	*1.119981506*
CPEB4	*1.119254011*
HCCS	*1.118252623*
PDZD11	*1.11384498*
RAP2A	*1.109861466*
FGD1	*1.107187332*
DMWD	*1.104055844*
RPS6KA5	*1.102260364*
FANCI	*1.101404662*
TIMM8A1	*1.092175081*
CKS1B	*1.087433715*
BIN2	*1.085955016*
LAT	*1.080910877*
CORO1C	*1.080202175*
AI504432	*1.078870052*
IFT57	*1.075766744*
ERCC6L	*1.073356618*
TAB3	*1.072040001*
CD47	*1.071978569*
SPDL1	*1.070278775*
KCTD9	*1.07005906*
RBBP7	*1.068546637*
NT5C3	*1.064695775*
TPM1	*1.064283535*
ENO1	*1.059335763*
SLC25A24	*1.059064519*
RASA1	*1.058747552*
CENPW	*1.05852084*
RNF157	*1.054706652*
NCAPD2	*1.050329733*
CD200R1	*1.049235107*
PCYT1B	*1.048194506*
SHCBP1	*1.048081198*
IQCB1	*1.047662965*
TK1	*1.04521125*
HAUS7	*1.042631918*
POP5	*1.04192402*
ORC6	*1.037209494*

ITFG2	*1.036606615*
BARD1	*1.035210454*
IQGAP3	*1.034851495*
RBPMS2	*1.034248617*
AP1S3	*1.03020871*
COX6A1	*1.028542721*
LAGE3	*1.026610044*
PHB2	*1.026162078*
POLE	*1.023611483*
HK2	*1.021681035*
SC5D	*1.02076754*
TNK2	*1.019156121*
STK3	*1.018796913*
MLKL	*1.015767878*
TXNL4A	*1.015262038*
SPIDR	*1.015207064*
PRKAR2B	*1.014913005*
DIS3	*1.013955108*
PLP2	*1.012013968*
EMG1	*1.010025762*
GLTP	*1.00726304*
CDC7	*1.004891307*
RESF1	*1.004643372*
PCLAF	*1.001639023*

List of downregulated genes in HPC7-JAK2V617F (< or equal to -1 log_2 (Fold change))

Gene_ID	Log2 (fold change)
ERP29	*-1.001061986*
PARP9	*-1.003833317*
SLC1A4	*-1.00407604*
MED27	*-1.009323799*
CALU	*-1.010156281*
PRKCI	*-1.010404385*
HYOU1	*-1.013198315*
DNAJA2	*-1.013943758*
METTL4	*-1.016236471*
DTD1	*-1.017595459*
ITGB3	*-1.017915547*
NUDT13	*-1.031231017*
PPP5C	*-1.0320195*
CROCC	*-1.043847731*
MYEF2	*-1.044078557*
MPRIP	*-1.045409219*
DDIAS	*-1.050819691*
EMC4	*-1.052464745*
PDCD11	*-1.057079676*
CCDC59	*-1.058114613*
PPIP5K1	*-1.0597213*
UBE2Q1	*-1.063993793*
SRP72	*-1.064028568*
BC052040	*-1.064745734*
AGPS	*-1.068343572*
CITED2	*-1.068889333*
ATP6V0A2	*-1.069395537*
EIF4A3	*-1.073969769*
UQCC2	*-1.074003859*
CSTB	*-1.074288174*
CANX	*-1.074609916*
UEVLD	*-1.07515792*
POMT2	*-1.075608679*
PMPCB	*-1.075701793*
WDR83OS	*-1.077101235*
TXNDC17	*-1.08021204*
CSNK1G3	*-1.083970418*
PPP1R37	*-1.089910343*
IKBKG	*-1.090037467*
ZC3H15	*-1.090890813*
YWHAE	*-1.091002611*

SEC61B	*-1.091328542*
PSAT1	*-1.091788747*
ELAC1	*-1.095179842*
CDC123	*-1.095683572*
EMC3	*-1.096034081*
USP36	*-1.097873622*
MRPL47	*-1.0990797*
TOMM7	*-1.101503204*
SNRNP27	*-1.102582806*
THADA	*-1.10542884*
NATD1	*-1.107423329*
ZFP933	*-1.111095804*
HSPA5	*-1.122637209*
GINS1	*-1.124049167*
ARL6IP5	*-1.126159219*
EME1	*-1.127018197*
GATAD2B	*-1.129215868*
NAA35	*-1.132727488*
HSPA13	*-1.133549577*
SUCLA2	*-1.134566279*
CFAP97	*-1.135015949*
ARL6IP4	*-1.138371834*
ARRDC3	*-1.138648225*
ABI2	*-1.148834266*
ASF1A	*-1.149288768*
ATXN2	*-1.151819995*
SLK	*-1.152825911*
PHIP	*-1.153316021*
EIF3J2	*-1.155437985*
TMEM181A	*-1.157020663*
XPNPEP3	*-1.163004203*
NUDT21	*-1.163246833*
SCFD2	*-1.163304419*
EXO1	*-1.163403284*
CCDC124	*-1.164640807*
ARHGAP10	*-1.164940214*
USP9X	*-1.176682708*
RAF1	*-1.180800987*
NUP88	*-1.182541375*
ERCC6L2	*-1.184521481*
TMEM168	*-1.185651226*
CD3EAP	*-1.188580891*
TIMM13	*-1.191549142*
SLC16A6	*-1.204383616*
TMEM209	*-1.206063946*
TMEM248	*-1.210134717*
CLTC	*-1.210158115*
HINFP	*-1.214264186*

SSB	*-1.21590278*
UFD1	*-1.217067309*
ACAA2	*-1.220416315*
CLPP	*-1.220550052*
NOP10	*-1.221324739*
CSTF3	*-1.228962071*
HASPIN	*-1.242323659*
UBR3	*-1.246950965*
TENT4A	*-1.249318317*
TRAPPC9	*-1.249612109*
ARMC1	*-1.250678394*
DDX19A	*-1.256900105*
FNTA	*-1.263326338*
ARHGAP2 5	*-1.265655979*
DVL1	*-1.268719371*
ENSA	*-1.27692485*
RAB8A	*-1.293416956*
SLX4	*-1.293585393*
BABAM2	*-1.296991831*
CWC15	*-1.301829154*
AP4E1	*-1.305960923*
ATP9B	*-1.307153328*
SEC13	*-1.308006493*
TAF5L	*-1.315731183*
FBXW7	*-1.324874101*
PAPSS1	*-1.327562695*
DNAJC2	*-1.332400067*
GNPAT	*-1.333390486*
REEP5	*-1.34189788*
CUL1	*-1.354566471*
FAM92A	*-1.361508157*
OTUD6B	*-1.367883516*
TAOK1	*-1.369052576*
PSMD6	*-1.371012365*
MYCN	*-1.37367327*
SRPRB	*-1.373786086*
IPP	*-1.376973554*
TMX2	*-1.377333596*
PRKACA	*-1.382221633*
MIGA2	*-1.395416451*
PCNP	*-1.401891186*
PACS2	*-1.401935204*
ETFA	*-1.404758237*
SFT2D1	*-1.407593858*
HSPA12B	*-1.411534211*
NUDCD2	*-1.416241194*
BC031181	*-1.428925838*
RNF168	*-1.436250368*

POLE3	*-1.474489868*
CEP57	*-1.476634625*
PURB	*-1.482874986*
UBE2S	*-1.485015442*
ATPIF1	*-1.488547733*
APH1A	*-1.491272693*
INPP5F	*-1.495984439*
DNPH1	*-1.496593159*
ZFP120	*-1.499040293*
THNSL1	*-1.508869691*
HEATR5B	*-1.534733619*
CYB5R4	*-1.553209137*
RAB1A	*-1.558381134*
USP16	*-1.56867963*
GART	*-1.574929431*
CAD	*-1.575976426*
NINJ1	*-1.591070748*
RPL22L1	*-1.594289574*
YIPF4	*-1.624779506*
NEMP1	*-1.627785587*
MOCS2	*-1.628328081*
NUP153	*-1.649008329*
ZCRB1	*-1.669035275*
HDAC1	*-1.71160151*
SYNGR2	*-1.740442477*
NDUFS4	*-1.741298134*
RNASEH2C	*-1.770650173*
WDR5	*-1.786381883*
ABHD2	*-1.819996691*
LUC7L2	*-1.834071462*
CCDC34	*-1.842636003*
MRPL45	*-1.866409609*
PITHD1	*-1.867418456*
ING2	*-1.872369731*
BAIAP2	*-1.874866002*
HINT1	*-1.903782009*
JDP2	*-1.912103033*
FXN	*-1.936388383*
ROCK1	*-1.944377232*
SRSF10	*-1.94750177*
TMEM161A	*-1.971476319*
PDCD5	*-1.989033664*
SHLD2	*-2.036739523*
CUL5	*-2.053373566*
ELMSAN1	*-2.070232207*
PRNP	*-2.073473514*
INKA1	*-2.088085488*
IFRD1	*-2.113992869*

GOLGA3	*-2.213168985*
ZKSCAN17	*-2.264044798*
ZFP148	*-2.386857388*
MRPL19	*-2.410177311*
IPPK	*-2.576799882*
MRPS15	*-2.57816498*
PTMS	*-2.711524879*
OXSM	*-2.810378304*
PHLPP2	*-2.863617335*
GTF2H2	*-2.916981228*
AZIN1	*-3.090336544*
RAB24	*-3.350834958*

List of GSEA of dysregulated pathways in the HPC7-JAK2V617F cell lines.

NAME	SIZE
KOBAYASHI_EGFR_SIGNALING_24HR_UP	80
ROSTY_CERVICAL_CANCER_PROLIFERATION_CLUSTER	125
FLORIO_NEOCORTEX_BASAL_RADIAL_GLIA_DN	132
SOTIRIOU_BREAST_CANCER_GRADE_1_VS_3_UP	139
DUTERTRE_ESTRADIOL_RESPONSE_24HR_UP	250
SHEDDEN_LUNG_CANCER_POOR_SURVIVAL_A6	360
BUFFA_HYPOXIA_METAGENE	42
GRAHAM_NORMAL_QUIESCENT_VS_NORMAL_DIVIDING_DN	78
GOBERT_OLIGODENDROCYTE_DIFFERENTIATION_UP	467
ZHOU_CELL_CYCLE_GENES_IN_IR_RESPONSE_24HR	102
FERREIRA_EWINGS_SARCOMA_UNSTABLE_VS_STABLE_UP	129
MEBARKI_HSC_PROGENITOR_FZD8CRD_UP	343
RUIZ_TNC_TARGETS_DN	102
FISCHER_DREAM_TARGETS	802
BERENJENO_TRANSFORMED_BY_RHOA_UP	435
KORKOLA_SEMINOMA_UP	31
CROONQUIST_IL6_DEPRIVATION_DN	82
BURTON_ADIPOGENESIS_3	85
NAKAYAMA_SOFT_TISSUE_TUMORS_PCA2_UP	45
BLANCO_MELO_BRONCHIAL_EPITHELIAL_CELLS_INFLUENZA_A_DEL_NS1_INFECTION_DN	135
HOFFMANN_LARGE_TO_SMALL_PRE_BII_LYMPHOCYTE_UP	132
VANTVEER_BREAST_CANCER_METASTASIS_DN	74
REACTOME_RESOLUTION_OF_SISTER_CHROMATID_COHESION	108
MARSON_BOUND_BY_E2F4_UNSTIMULATED	577
BASAKI_YBX1_TARGETS_UP	205
FUJII_YBX1_TARGETS_DN	162
ZHAN_MULTIPLE_MYELOMA_PR_UP	41
REACTOME_RHO_GTPASES_ACTIVATE_FORMINS	117
MARKEY_RB1_ACUTE_LOF_UP	189
MISSIAGLIA_REGULATED_BY_METHYLATION_DN	99
SARRIO_EPITHELIAL_MESENCHYMAL_TRANSITION_UP	128
ZHENG_GLIOBLASTOMA_PLASTICITY_UP	128
GRAHAM_CML_DIVIDING_VS_NORMAL_QUIESCENT_UP	146
PUJANA_XPRSS_INT_NETWORK	153
KANG_DOXORUBICIN_RESISTANCE_UP	49
KAUFFMANN_MELANOMA_RELAPSE_UP	58
CROONQUIST_NRAS_SIGNALING_DN	62
YU_MYC_TARGETS_UP	42
MORI_LARGE_PRE_BII_LYMPHOCYTE_UP	77
LI_WILMS_TUMOR_VS_FETAL_KIDNEY_1_DN	124
HARRIS_HYPOXIA	44
TANG_SENESCENCE_TP53_TARGETS_DN	36

VECCHI_GASTRIC_CANCER_EARLY_UP	288
WHITFIELD_CELL_CYCLE_LITERATURE	44
ODONNELL_TFRC_TARGETS_DN	91
SEMENZA_HIF1_TARGETS	19
CHICAS_RB1_TARGETS_GROWING	133
WP_RETINOBLASTOMA_GENE_IN_CANCER	80
RODRIGUES_THYROID_CARCINOMA_POORLY_DIFFERENTIATED_UP	524
WP_DNA_REPLICATION	40
KAUFFMANN_DNA_REPAIR_GENES	196
CHIANG_LIVER_CANCER_SUBCLASS_PROLIFERATION_UP	132
KEGG_CELL_CYCLE	108
KAN_RESPONSE_TO_ARSENIC_TRIOXIDE	55
REACTOME_HOMOLOGOUS_DNA_PAIRING_AND_STRAND_EXCHANGE	39
ZHANG_BREAST_CANCER_PROGENITORS_UP	337
REACTOME_G2_M_CHECKPOINTS	126
KEGG_MISMATCH_REPAIR	20
REACTOME_DNA_REPAIR	264
REACTOME_HOMOLOGY_DIRECTED_REPAIR	97
LINDGREN_BLADDER_CANCER_CLUSTER_3_UP	240
SHEPARD_BMYB_MORPHOLINO_DN	111
VILIMAS_NOTCH1_TARGETS_UP	21
FARMER_BREAST_CANCER_CLUSTER_2	26
REACTOME_TRANSPORT_OF_MATURE_MRNAS_DERIVED_FROM_INTRONLESS_TRANSCRIPTS	39
SMID_BREAST_CANCER_BASAL_UP	255
REACTOME_TRANSPORT_OF_THE_SLBP_DEPENDANT_MATURE_MRNA	32
CHIARADONNA_NEOPLASTIC_TRANSFORMATION_KRAS_UP	89
SHEPARD_BMYB_TARGETS	44
REACTOME_INTERACTIONS_OF_REV_WITH_HOST_CELLULAR_PROTEINS	32
REACTOME_PROCESSING_OF_DNA_DOUBLE_STRAND_BREAK_ENDS	60
BURTON_ADIPOGENESIS_PEAK_AT_16HR	35
KEGG_DNA_REPLICATION	34
WP_G1_TO_S_CELL_CYCLE_CONTROL	55
HORTON_SREBF_TARGETS	20
REACTOME_LAGGING_STRAND_SYNTHESIS	19
MIYAGAWA_TARGETS_OF_EWSR1_ETS_FUSIONS_DN	73
UDAYAKUMAR_MED1_TARGETS_UP	112
RIGGI_EWING_SARCOMA_PROGENITOR_DN	54
WP_DNA_REPAIR_PATHWAYS_FULL_NETWORK	112
WP_DNA_DAMAGE_RESPONSE	55
NUYTTEN_EZH2_TARGETS_DN	757
REACTOME_TRANSCRIPTIONAL_REGULATION_BY_SMALL_RNAS	43

BHATI_G2M_ARREST_BY_2METHOXYESTRADIOL_UP	73
DESERT_STEM_CELL_HEPATOCELLULAR_CARCINOMA_SUBCLASS_UP	166
WP_CHOLESTEROL_METABOLISM_INCLUDES_BOTH_BLOCH_AND_KANDUTSCHRUSSELL_PATHWAYS	35
REACTOME_REGULATION_OF_GLUCOKINASE_BY_GLUCOKINASE_REGULATORY_PROTEIN	26
PID_HIF2PATHWAY	23
REACTOME_HCMV_INFECTION	78
MILI_PSEUDOPODIA_HAPTOTAXIS_UP	430
PID_PRL_SIGNALING_EVENTS_PATHWAY	19
REACTOME_SUMOYLATION_OF_UBIQUITINYLATION_PROTEINS	35
REACTOME_REGULATION_OF_CHOLESTEROL_BIOSYNTHESIS_BY_SREBP_SREBF	49
SWEET_KRAS_TARGETS_DN	37
REACTOME_SUMOYLATION_OF_SUMOYLATION_PROTEINS	31
ELVIDGE_HIF1A_TARGETS_DN	43
FRASOR_RESPONSE_TO_SERM_OR_FULVESTRANT_DN	44
REACTOME_EXTENSION_OF_TELOMERES	47
WP_FERROPTOSIS	28
WP_DNA_IRDOUBLE_STRAND_BREAKS_DSBS_AND_CELLULAR_RESPONSE_VIA_ATM	47
BIOCARTA_CELLCYCLE_PATHWAY	21
KRIEG_HYPOXIA_NOT_VIA_KDM3A	461
REACTOME_G1_S_SPECIFIC_TRANSCRIPTION	28
REACTOME_POSTMITOTIC_NUCLEAR_PORE_COMPLEX_NPC_REFORMATION	25
REACTOME_TRANSLESION_SYNTHESIS_BY_POLH	19
YANG_BCL3_TARGETS_UP	168
CERVERA_SDHB_TARGETS_2	28
KEGG_NUCLEOTIDE_EXCISION_REPAIR	42
WP_NUCLEOTIDE_EXCISION_REPAIR	42
PID_FANCONI_PATHWAY	42
SCIBETTA_KDM5B_TARGETS_DN	58
PASQUALUCCI_LYMPHOMA_BY_GC_STAGE_UP	151
REACTOME_E2F_MEDIATED_REGULATION_OF_DNA_REPLICATION	21
BOQUEST_STEM_CELL_CULTURED_VS_FRESH_UP	166
REACTOME_NONHOMOLOGOUS_END_JOINING_NHEJ	32
AMIT_EGF_RESPONSE_60_HELA	33
WP_HEREDITARY_LEIOMYOMATOSIS_AND_RENAL_CELL_CARCINOMA_PATHWAY	18
REACTOME_CREB1_PHOSPHORYLATION_THROUGH_NMDA_RECEPTOR_MEDIATED_ACTIVATION_OF_RAS_SIGNALING	15
REACTOME_GLOBAL_GENOME_NUCLEOTIDE_EXCISION_REPAIR_GG_NER	83
TONKS_TARGETS_OF_RUNX1_RUNX1T1_FUSION_HSC_UP	103
NIKOLSKY_BREAST_CANCER_16Q24_AMPLICON	33

REACTOME_RESOLUTION_OF_D_LOOP_STRUCTURES	31
BIOCARTA_ATM_PATHWAY	16
AMIT_EGF_RESPONSE_480_MCF10A	21
SMITH_TERT_TARGETS_UP	105
WP_FLUOROPYRIMIDINE_ACTIVITY	19
WANG_METASTASIS_OF_BREAST_CANCER_ESR1_DN	18
MOREAUX_MULTIPLE_MYELOMA_BY_TACI_DN	152
MATSUDA_NATURAL_KILLER_DIFFERENTIATION	263
GUTIERREZ_MULTIPLE_MYELOMA_UP	28
GARCIA_TARGETS_OF_FLI1_AND_DAX1_DN	129
REACTOME_APC_C_CDH1_MEDIATED_DEGRADATION_OF_CDC20_AND_OTHER_APC_C_CDH1_TARGETED_PROTEINS_IN_LATE_MITOSIS_EARLY_G1	70
WP_EXERCISEINDUCED_CIRCADIAN_REGULATION	32
BROWNE_HCMV_INFECTION_48HR_UP	101
GHANDHI_BYSTANDER_IRRADIATION_UP	25
SEITZ_NEOPLASTIC_TRANSFORMATION_BY_8P_DELETION_UP	20
FIGUEROA_AML_METHYLATION_CLUSTER_1_UP	52
FULCHER_INFLAMMATORY_RESPONSE_LECTIN_VS_LPS_UP	352
KEGG_SPLICEOSOME	115
REACTOME_SUMOYLATION_OF_DNA_DAMAGE_RESPONSE_AND_REPAIR_PROTEINS	72
REACTOME_TP53_REGULATES_TRANSCRIPTION_OF_DNA_REPAIR_GENES	57
SESTO_RESPONSE_TO_UV_C1	45
WONG_PROTEASOME_GENE_MODULE	43
EBAUER_TARGETS_OF_PAX3_FOXO1_FUSION_DN	20
IZADPANAH_STEM_CELL_ADIPOSE_VS_BONE_DN	37
BERTUCCI_INVASIVE_CARCINOMA_DUCTAL_VS_LOBULAR_UP	15
NUYTTEN_NIPP1_TARGETS_UP	414
SESTO_RESPONSE_TO_UV_C7	45
WIERENGA_STAT5A_TARGETS_UP	99
REACTOME_REGULATION_OF_RUNX3_EXPRESSION_AND_ACTIVITY	54
SENESE_HDAC3_TARGETS_UP	291
OKUMURA_INFLAMMATORY_RESPONSE_LPS	99
HELLER_HDAC_TARGETS_UP	157
DAIRKEE_CANCER_PRONE_RESPONSE_E2	15
BAKKER_FOXO3_TARGETS_UP	44
MCBRYAN_PUBERTAL_BREAST_4_5WK_DN	121
REACTOME_ADORA2B_MEDIATED_ANTI_INFLAMMATORY_CYTOKINES_PRODUCTION	28
REACTOME_GENE_AND_PROTEIN_EXPRESSION_BY_JAK_STAT_SIGNALING_AFTER_INTERLEUKIN_12_STIMULATION	30
PID_HDAC_CLASSII_PATHWAY	28
WP_HISTONE_MODIFICATIONS	36
REACTOME_SIGNALING_BY_EGFR	35

RIZKI_TUMOR_INVASIVENESS_2D_DN	42
WP_PROLACTIN_SIGNALING_PATHWAY	67
REACTOME_REGULATED_NECROSIS	38
REACTOME_RAB_GEFS_EXCHANGE_GTP_FOR_GDP_ON_RABS	74
FIGUEROA_AML_METHYLATION_CLUSTER_4_UP	54
REACTOME_ATTENUATION_PHASE	17
REACTOME_TRANSLATION	258
REACTOME_MITOCHONDRIAL_BIOGENESIS	64
PID_IL1_PATHWAY	25
PID_ER_NONGENOMIC_PATHWAY	30
REACTOME_TRISTETRAPROLIN_TTP_ZFP36_BINDS_AND_DESTABILIZES_MRNA	16
WP_REGULATION_OF_APOPTOSIS_BY_PARATHYROID_HORMONERELATED_PROTEIN	16
KEGG_ADHERENS_JUNCTION	44
MARIADASON_RESPONSE_TO_CURCUMIN_SULINDAC_5	16
BARRIER_CANCER_RELAPSE_NORMAL_SAMPLE_DN	28
QUINTENS_EMBRYONIC_BRAIN_RESPONSE_TO_IR	45
ODONNELL_METASTASIS_UP	29
REACTOME_SIGNALING_BY_NOTCH1	54
BIOCARTA_VDR_PATHWAY	23
WP_ENDOPLASMIC_RETICULUM_STRESS_RESPONSE_IN_CORONAVIRUS_INFECTION	34
REACTOME_VEGFR2_MEDIATED_VASCULAR_PERMEABILITY	22
REACTOME_NUCLEOBASE_CATABOLISM	17
BIOCARTA_ERK_PATHWAY	21
SIG_CHEMOTAXIS	34
KEGG_TGF_BETA_SIGNALING_PATHWAY	40
HEIDENBLAD_AMPLICON_8Q24_DN	25
REACTOME_REGULATION_OF_TP53_ACTIVITY_THROUGH_ACETYLATION	29
WP_HOSTPATHOGEN_INTERACTION_OF_HUMAN_CORONA_VIRUSES_APOPTOSIS	19
LINSLEY_MIR16_TARGETS	164
PID_ERBB4_PATHWAY	23
SA_PTEN_PATHWAY	16
PARK_HSC_AND_MULTIPOTENT_PROGENITORS	30

List of Hallmark GSEA of dysregulated pathways in the HPC7-JAK2V617F cell lines.

NAME	SIZE
HALLMARK_E2F_TARGETS	185
HALLMARK_MTORC1_SIGNALING	174
HALLMARK_G2M_CHECKPOINT	179
HALLMARK_CHOLESTEROL_HOMEOSTASIS	49
HALLMARK_HYPOXIA	112
HALLMARK_MYC_TARGETS_V1	189
HALLMARK_GLYCOLYSIS	121
HALLMARK_IL2_STAT5_SIGNALING	115
HALLMARK_UNFOLDED_PROTEIN_RESPONSE	94
HALLMARK_OXIDATIVE_PHOSPHORYLATION	167
HALLMARK_HEDGEHOG_SIGNALING	16
HALLMARK_MITOTIC_SPINDLE	164
HALLMARK_DNA_REPAIR	131
HALLMARK_MYC_TARGETS_V2	56
HALLMARK_SPERMATOGENESIS	51
HALLMARK_TNFA_SIGNALING_VIA_NFKB	121
HALLMARK_ESTROGEN_RESPONSE_LATE	94
HALLMARK_FATTY_ACID_METABOLISM	108
HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION	53
HALLMARK_ADIPOGENESIS	143
HALLMARK_ALLOGRAFT_REJECTION	89
HALLMARK_PROTEIN_SECRETION	79
HALLMARK_INTERFERON_GAMMA_RESPONSE	115
HALLMARK_COAGULATION	47
HALLMARK_KRAS_SIGNALING_UP	71
HALLMARK_PI3K_AKT_MTOR_SIGNALING	86
HALLMARK_MYOGENESIS	68
HALLMARK_P53_PATHWAY	135
HALLMARK_INTERFERON_ALPHA_RESPONSE	61
HALLMARK_ANDROGEN_RESPONSE	68
HALLMARK_WNT_BETA_CATENIN_SIGNALING	25
HALLMARK_INFLAMMATORY_RESPONSE	87
HALLMARK_APICAL_JUNCTION	96
HALLMARK_XENOBIOTIC_METABOLISM	91
HALLMARK_ESTROGEN_RESPONSE_EARLY	99
HALLMARK_UV_RESPONSE_UP	101

Bibliography

Adamson, J.W. 1968. The erythropoietin-hematocrit relationship in normal and polycythemic man: implications of marrow regulation. *Blood.* **32**(4), pp.597-609.

Adamson, J.W., Fialkow, P.J., Murphy, S., Prchal, J.F. and Steinmann, L. 1976. Polycythemia vera: stem-cell and probable clonal origin of the disease. *N Engl J Med.* **295**(17), pp.913-916.

Akada, H., Akada, S., Hutchison, R.E. and Mohi, G. 2014. Loss of wild-type Jak2 allele enhances myeloid cell expansion and accelerates myelofibrosis in Jak2V617F knock-in mice. *Leukemia.* **28**(8), pp.1627-1635.

Akada, H., Yan, D., Zou, H., Fiering, S., Hutchison, R.E. and Mohi, M.G. 2010. Conditional expression of heterozygous or homozygous Jak2V617F from its endogenous promoter induces a polycythemia vera-like disease. *Blood.* **115**(17), pp.3589-3597.

Alchini, R., Sato, H., Matsumoto, N., Shimogori, T., Sugo, N. and Yamamoto, N. 2017. Nucleocytoplasmic Shuttling of Histone Deacetylase 9 Controls Activity-Dependent Thalamocortical Axon Branching. *Scientific Reports.* **7**(1), p6024.

Anand, S., Stedham, F., Beer, P., Gudgin, E., Ortmann, C.A., Bench, A., Erber, W., Green, A.R. and Huntly, B.J. 2011. Effects of the JAK2 mutation on the hematopoietic stem and progenitor compartment in human myeloproliferative neoplasms. *Blood.* **118**(1), pp.177-181.

Anastas, J.N., Zee, B.M., Kalin, J.H., Kim, M., Guo, R., Alexandrescu, S., Blanco, M.A., Giera, S., Gillespie, S.M., Das, J., Wu, M., Nocco, S., Bonal, D.M., Nguyen, Q.D., Suva, M.L., Bernstein, B.E., Alani, R., Golub, T.R., Cole, P.A., Filbin, M.G. and Shi, Y. 2019. Re-programing Chromatin with a Bifunctional LSD1/HDAC Inhibitor Induces Therapeutic Differentiation in DIPG. *Cancer Cell.* **36**(5), pp.528-544.e510.

Andrews, S. 2010. *FastQC: A Quality Control Tool for High Throughput Sequence Data.* [Online]. Available from: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Araki, M., Yang, Y., Masubuchi, N., Hironaka, Y., Takei, H., Morishita, S., Mizukami, Y., Kan, S., Shirane, S., Edahiro, Y., Sunami, Y., Ohsaka, A. and Komatsu, N. 2016. Activation of the thrombopoietin receptor by mutant calreticulin in CALR-mutant myeloproliferative neoplasms. *Blood.* **127**(10), pp.1307-1316.

Arber, D.A., Orazi, A., Hasserjian, R., Thiele, J., Borowitz, M.J., Le Beau, M.M., Bloomfield, C.D., Cazzola, M. and Vardiman, J.W. 2016. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood.* **127**(20), pp.2391-2405.

Azagra, A., Román-González, L., Collazo, O., Rodríguez-Ubreva, J., de Yébenes, V.G., Barneda-Zahonero, B., Rodríguez, J., Castro de Moura, M., Grego-Bessa, J., Fernández-Duran, I., Islam, A.B., Esteller, M., Ramiro, A.R., Ballestar, E. and Parra, M. 2016. In vivo conditional deletion of HDAC7 reveals its requirement to establish proper B lymphocyte identity and development. *J Exp Med.* **213**(12), pp.2591-2601.

Bainton, D.F., Ullyot, J.L. and Farquhar, M.G. 1971. The development of neutrophilic polymorphonuclear leukocytes in human bone marrow. *J Exp Med.* **134**(4), pp.907-934.

Bandaranayake, R.M., Ungureanu, D., Shan, Y., Shaw, D.E., Silvennoinen, O. and Hubbard, S.R. 2012. Crystal structures of the JAK2 pseudokinase domain and the pathogenic mutant V617F. *Nat Struct Mol Biol.* **19**(8), pp.754-759.

Bao, H., Jacobs-Helber, S.M., Lawson, A.E., Penta, K., Wickrema, A. and Sawyer, S.T. 1999. Protein kinase B (c-Akt), phosphatidylinositol 3-kinase, and STAT5 are activated by erythropoietin (EPO) in HCD57 erythroid cells but are constitutively active in an EPO-independent, apoptosis-resistant subclone (HCD57-SREI cells). *Blood.* **93**(11), pp.3757-3773.

Barneda-Zahonero, B., Collazo, O., Azagra, A., Fernández-Duran, I., Serra-Musach, J., Islam, A.B.M.M.K., Vega-García, N., Malatesta, R., Camós, M., Gómez, A., Román-González, L., Vidal, A., López-Bigas, N., Villanueva, A., Esteller, M. and Parra, M. 2015. The transcriptional repressor HDAC7 promotes apoptosis and c-Myc downregulation in particular types of leukemia and lymphoma. *Cell death & disease.* **6**(2), pp.e1635-e1635.

Barrett, T., Wilhite, S.E., Ledoux, P., Evangelista, C., Kim, I.F., Tomashevsky, M., Marshall, K.A., Phillippy, K.H., Sherman, P.M., Holko, M., Yefanov, A., Lee, H., Zhang, N., Robertson, C.L., Serova, N., Davis, S. and Soboleva, A. 2013. NCBI GEO: archive for functional genomics data sets--update. *Nucleic Acids Res.* **41**(Database issue), pp.D991-995.

Baxter, E.J., Scott, L.M., Campbell, P.J., East, C., Fourouclas, N., Swanton, S., Vassiliou, G.S., Bench, A.J., Boyd, E.M., Curtin, N., Scott, M.A., Erber, W.N. and Green, A.R. 2005. Acquired mutation of the tyrosine kinase JAK2 in human myeloproliferative disorders. *Lancet.* **365**(9464), pp.1054-1061.

Beer, P.A., Campbell, P.J., Scott, L.M., Bench, A.J., Erber, W.N., Bareford, D., Wilkins, B.S., Reilly, J.T., Hasselbalch, H.C., Bowman, R., Wheatley, K., Buck, G., Harrison, C.N. and Green, A.R. 2008. MPL mutations in myeloproliferative disorders: analysis of the PT-1 cohort. *Blood.* **112**(1), pp.141-149.

Beier, U.H., Wang, L., Han, R., Akimova, T., Liu, Y. and Hancock, W.W. 2012. Histone deacetylases 6 and 9 and sirtuin-1 control Foxp3+ regulatory T cell function through shared and isoform-specific mechanisms. *Sci Signal.* **5**(229), pra45.

Benekli, M., Baer, M.R., Baumann, H. and Wetzler, M. 2003. Signal transducer and activator of transcription proteins in leukemias. *Blood.* **101**(8), pp.2940-2954.

Bersenev, A., Wu, C., Balcerek, J. and Tong, W. 2008. Lnk controls mouse hematopoietic stem cell self-renewal and quiescence through direct interactions with JAK2. *J Clin Invest.* **118**(8), pp.2832-2844.

Blum, M., Chang, H.-Y., Chuguransky, S., Grego, T., Kandasaamy, S., Mitchell, A., Nuka, G., Paysan-Lafosse, T., Qureshi, M., Raj, S., Richardson, L., Salazar, G.A., Williams, L., Bork, P., Bridge, A., Gough, J., Haft, D.H., Letunic, I., Marchler-Bauer, A., Mi, H., Natale, D.A., Necci, M., Orengo, C.A., Pandurangan, A.P., Rivoire, C., Sigrist, C.J.A., Sillitoe, I., Thanki, N., Thomas, P.D., Tosatto, S.C.E., Wu, C.H., Bateman, A. and Finn, R.D. 2021. The InterPro protein families and domains database: 20 years on. *Nucleic acids research.* **49**(D1), pp.D344-D354.

Brown, K., Xie, S., Qiu, X., Mohrin, M., Shin, J., Liu, Y., Zhang, D., Scadden, D.T. and Chen, D. 2013. SIRT3 reverses aging-associated degeneration. *Cell Rep.* **3**(2), pp.319-327.

Burgoyne, P.S., Mahadevaiah, S.K. and Turner, J.M. 2007. The management of DNA double-strand breaks in mitotic G2, and in mammalian meiosis viewed from a mitotic G2 perspective. *Bioessays.* **29**(10), pp.974-986.

Celik, H., Koh, W.K., Kramer, A.C., Ostrander, E.L., Mallaney, C., Fisher, D.A.C., Xiang, J., Wilson, W.C., Martens, A., Kothari, A., Fishberger, G., Tycksen, E., Karpova, D., Duncavage, E.J., Lee, Y., Oh, S.T. and Challen, G.A. 2018. JARID2 Functions as a Tumor Suppressor in Myeloid Neoplasms by Repressing Self-Renewal in Hematopoietic Progenitor Cells. *Cancer Cell.* **34**(5), pp.741-756.e748.

Chachoua, I., Pecquet, C., El-Khoury, M., Nivarthi, H., Albu, R.-I., Marty, C., Gryshkova, V., Defour, J.-P., Vertenoeil, G., Ngo, A., Koay, A., Raslova, H., Courtoy, P.J., Choong, M.L., Plo, I., Vainchenker, W., Kralovics, R. and Constantinescu, S.N. 2016. Thrombopoietin receptor activation by myeloproliferative neoplasm associated calreticulin mutants. *Blood.* **127**(10), pp.1325-1335.

Chateauvieux, S., Grigorakaki, C., Morceau, F., Dicato, M. and Diederich, M. 2011. Erythropoietin, erythropoiesis and beyond. *Biochem Pharmacol.* **82**(10), pp.1291-1303.

Chen, E., Ahn, J.S., Massie, C.E., Clynes, D., Godfrey, A.L., Li, J., Park, H.J., Nangalia, J., Silber, Y., Mullally, A., Gibbons, R.J. and Green, A.R. 2014. JAK2V617F promotes replication fork stalling with disease-restricted impairment of the intra-S checkpoint response. *Proc Natl Acad Sci U S A.* **111**(42), pp.15190-15195.

Chen, E., Ahn, J.S., Sykes, D.B., Breyfogle, L.J., Godfrey, A.L., Nangalia, J., Ko, A., DeAngelo, D.J., Green, A.R. and Mullally, A. 2015. RECQL5 Suppresses Oncogenic JAK2-Induced Replication Stress and Genomic Instability. *Cell Rep.* **13**(11), pp.2345-2352.

Chen, E., Beer, P.A., Godfrey, A.L., Ortmann, C.A., Li, J., Costa-Pereira, A.P., Ingle, C.E., Dermitzakis, E.T., Campbell, P.J. and Green, A.R. 2010. Distinct

clinical phenotypes associated with JAK2V617F reflect differential STAT1 signaling. *Cancer Cell.* **18**(5), pp.524-535.

Chen, G.L. and Prchal, J.T. 2007. X-linked clonality testing: interpretation and limitations. *Blood.* **110**(5), pp.1411-1419.

Cheng, F., Lienlaf, M., Perez-Villarroel, P., Wang, H.W., Lee, C., Woan, K., Woods, D., Knox, T., Bergman, J., Pinilla-Ibarz, J., Kozikowski, A., Seto, E., Sotomayor, E.M. and Villagra, A. 2014. Divergent roles of histone deacetylase 6 (HDAC6) and histone deacetylase 11 (HDAC11) on the transcriptional regulation of IL10 in antigen presenting cells. *Mol Immunol.* **60**(1), pp.44-53.

Čokić, V.P., Mossuz, P., Han, J., Socoro, N., Beleslin-Čokić, B.B., Mitrović, O., Subotički, T., Diklić, M., Leković, D., Gotić, M., Puri, R.K., Noguchi, C.T. and Schechter, A.N. 2015. Microarray and Proteomic Analyses of Myeloproliferative Neoplasms with a Highlight on the mTOR Signaling Pathway. *PLoS One.* **10**(8), pe0135463.

Colaizzo, D., Amitrano, L., Tiscia, G.L., Scenna, G., Grandone, E., Guardascione, M.A., Brancaccio, V. and Margaglione, M. 2007. The JAK2 V617F mutation frequently occurs in patients with portal and mesenteric venous thrombosis. *J Thromb Haemost.* **5**(1), pp.55-61.

Correa, P.N., Eskinazi, D. and Axelrad, A.A. 1994. Circulating erythroid progenitors in polycythemia vera are hypersensitive to insulin-like growth factor-1 in vitro: studies in an improved serum-free medium. *Blood.* **83**(1), pp.99-112.

Couedic, J.-P.L., Mitjavila, M.-T., Villeval, J.-L., Feger, F., Gobert, S., Mayeux, P., Casadevall, N. and Vainchenker, W. 1996. Missense Mutation of the Erythropoietin Receptor Is a Rare Event in Human Erythroid Malignancies. *Blood.* **87**(4), pp.1502-1511.

Dai, C.H., Krantz, S.B., Dessypris, E.N., Means, R.T., Jr., Horn, S.T. and Gilbert, H.S. 1992. Polycythemia vera. II. Hypersensitivity of bone marrow erythroid, granulocyte-macrophage, and megakaryocyte progenitor cells to interleukin-3 and granulocyte-macrophage colony-stimulating factor. *Blood.* **80**(4), pp.891-899.

Dai, W.J., Bartens, W., Köhler, G., Hufnagel, M., Kopf, M. and Brombacher, F. 1997. Impaired macrophage listericidal and cytokine activities are responsible for the rapid death of Listeria monocytogenes-infected IFN-gamma receptor-deficient mice. *J Immunol.* **158**(11), pp.5297-5304.

Dameshek, W. 1950. PHYSIOPATHOLOGY AND COURSE OF POLYCYTHEMIA VERA AS RELATED TO THERAPY. *Journal of the American Medical Association.* **142**(11), pp.790-797.

Dawson, M.A., Bannister, A.J., Göttgens, B., Foster, S.D., Bartke, T., Green, A.R. and Kouzarides, T. 2009. JAK2 phosphorylates histone H3Y41 and excludes HP1alpha from chromatin. *Nature.* **461**(7265), pp.819-822.

de Wolf, J.T., Beentjes, J.A., Esselink, M.T., Smit, J.W., Halie, R.M., Clark, S.C. and Vellenga, E. 1989. In polycythemia vera human interleukin 3 and granulocyte-macrophage colony-stimulating factor enhance erythroid colony growth in the absence of erythropoietin. *Exp Hematol.* **17**(9), pp.981-983.

Deane, C.M., Kroemer, R.T. and Richards, W.G. 1997. A structural model of the human thrombopoietin receptor complex. *J Mol Graph Model.* **15**(3), pp.170-178, 185-178.

Defour, J.-P., Itaya, M., Gryshkova, V., Brett, I.C., Pecquet, C., Sato, T., Smith, S.O. and Constantinescu, S.N. 2013. Tryptophan at the transmembrane–cytosolic junction modulates thrombopoietin receptor dimerization and activation. *Proceedings of the National Academy of Sciences.* **110**(7), p2540.

Delehanty, L.L., Bullock, G.C. and Goldfarb, A.N. 2012. Protein kinase D-HDAC5 signaling regulates erythropoiesis and contributes to erythropoietin cross-talk with GATA1. *Blood.* **120**(20), pp.4219-4228.

Delhommeau, F., Dupont, S., Tonetti, C., Massé, A., Godin, I., Le Couedic, J.P., Debili, N., Saulnier, P., Casadevall, N., Vainchenker, W. and Giraudier, S. 2007. Evidence that the JAK2 G1849T (V617F) mutation occurs in a lymphomyeloid progenitor in polycythemia vera and idiopathic myelofibrosis. *Blood.* **109**(1), pp.71-77.

Dijkers, P.F., van Dijk, T.B., de Groot, R.P., Raaijmakers, J.A., Lammers, J.W., Koenderman, L. and Coffer, P.J. 1999. Regulation and function of protein kinase B and MAP kinase activation by the IL-5/GM-CSF/IL-3 receptor. *Oncogene.* **18**(22), pp.3334-3342.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. and Gingeras, T.R. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics.* **29**(1), pp.15-21.

Doubeikovski, A., Uzan, G., Doubeikovski, Z., Prandini, M.H., Porteu, F., Gisselbrecht, S. and Dusanter-Fourt, I. 1997. Thrombopoietin-induced expression of the glycoprotein IIb gene involves the transcription factor PU.1/Spi-1 in UT7-Mpl cells. *J Biol Chem.* **272**(39), pp.24300-24307.

Drachman, J.G., Griffin, J.D. and Kaushansky, K. 1995. The c-Mpl ligand (thrombopoietin) stimulates tyrosine phosphorylation of Jak2, Shc, and c-Mpl. *J Biol Chem.* **270**(10), pp.4979-4982.

DuBridge, R.B., Tang, P., Hsia, H.C., Leong, P.M., Miller, J.H. and Calos, M.P. 1987. Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system. *Mol Cell Biol.* **7**(1), pp.379-387.

Eaton, D.L. and de Sauvage, F.J. 1997. Thrombopoietin: the primary regulator of megakaryocytopoiesis and thrombopoiesis. *Exp Hematol.* **25**(1), pp.1-7.

Edgar, R. and Lash, A. 2002. 6. The Gene Expression Omnibus (GEO): A Gene Expression and Hybridization Repository.

Elf, S., Abdelfattah, N.S., Baral, A.J., Beeson, D., Rivera, J.F., Ko, A., Florescu, N., Birrane, G., Chen, E. and Mullally, A. 2018. Defining the requirements for the pathogenic interaction between mutant calreticulin and MPL in MPN. *Blood.* **131**(7), pp.782-786.

Elf, S., Abdelfattah, N.S., Chen, E., Perales-Patón, J., Rosen, E.A., Ko, A., Peisker, F., Florescu, N., Giannini, S., Wolach, O., Morgan, E.A., Tothova, Z., Losman, J.A., Schneider, R.K., Al-Shahrour, F. and Mullally, A. 2016. Mutant Calreticulin Requires Both Its Mutant C-terminus and the Thrombopoietin Receptor for Oncogenic Transformation. *Cancer Discov.* **6**(4), pp.368-381.

Epstein, E. and Goedel, A. 1934. Hemorrhagic thrombocythemia with a cascular, sclerotic spleen. *Virchows Arch.* **293**, pp.233-248.

Fatrai, S., Wierenga, A.T., Daenen, S.M., Vellenga, E. and Schuringa, J.J. 2011. Identification of HIF2alpha as an important STAT5 target gene in human hematopoietic stem cells. *Blood.* **117**(12), pp.3320-3330.

Fialkow, P.J. 1979. Clonal Origin of Human Tumors. *Annual Review of Medicine.* **30**(1), pp.135-143.

Fialkow, P.J., Faguet, G.B., Jacobson, R.J., Vaidya, K. and Murphy, S. 1981. Evidence that essential thrombocythemia is a clonal disorder with origin in a multipotent stem cell. *Blood.* **58**(5), pp.916-919.

Fialkow, P.J., Gartler, S.M. and Yoshida, A. 1967. Clonal origin of chronic myelocytic leukemia in man. *Proceedings of the National Academy of Sciences of the United States of America.* **58**(4), pp.1468-1471.

Fishley, B. and Alexander, W.S. 2004. Thrombopoietin signalling in physiology and disease. *Growth Factors.* **22**(3), pp.151-155.

Fourouclas, N., Li, J., Gilby, D.C., Campbell, P.J., Beer, P.A., Boyd, E.M., Goodeve, A.C., Bareford, D., Harrison, C.N., Reilly, J.T., Green, A.R. and Bench, A.J. 2008. Methylation of the suppressor of cytokine signaling 3 gene (SOCS3) in myeloproliferative disorders. *Haematologica.* **93**(11), pp.1635-1644.

Frank, S.J. 2002. Receptor dimerization in GH and erythropoietin action--it takes two to tango, but how? *Endocrinology.* **143**(1), pp.2-10.

French, D.L. and Seligsohn, U. 2000. Platelet Glycoprotein IIb/IIIa Receptors and Glanzmann's Thrombasthenia. *Arteriosclerosis, Thrombosis, and Vascular Biology.* **20**(3), pp.607-610.

Funakoshi-Tago, M., Pelletier, S., Matsuda, T., Parganas, E. and Ihle, J.N. 2006. Receptor specific downregulation of cytokine signaling by autophosphorylation in the FERM domain of Jak2. *The EMBO journal.* **25**(20), pp.4763-4772.

Gadina, M., Hilton, D., Johnston, J.A., Morinobu, A., Lighvani, A., Zhou, Y.J., Visconti, R. and O'Shea, J.J. 2001. Signaling by type I and II cytokine receptors: ten years after. *Curr Opin Immunol.* **13**(3), pp.363-373.

Gale, R.E., Allen, A.J., Nash, M.J. and Linch, D.C. 2007. Long-term serial analysis of X-chromosome inactivation patterns and JAK2 V617F mutant levels in patients with essential thrombocythemia show that minor mutant-positive clones can remain stable for many years. *Blood.* **109**(3), pp.1241-1243.

Genovese, G., Kähler, A.K., Handsaker, R.E., Lindberg, J., Rose, S.A., Bakhoum, S.F., Chambert, K., Mick, E., Neale, B.M., Fromer, M., Purcell, S.M., Svantesson, O., Landén, M., Höglund, M., Lehmann, S., Gabriel, S.B., Moran, J.L., Lander, E.S., Sullivan, P.F., Sklar, P., Grönberg, H., Hultman, C.M. and McCarroll, S.A. 2014. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N Engl J Med.* **371**(26), pp.2477-2487.

Gery, S. and Koeffler, H.P. 2013. Role of the adaptor protein LNK in normal and malignant hematopoiesis. *Oncogene.* **32**(26), pp.3111-3118.

Gewirtz, A.M., Bruno, E., Elwell, J. and Hoffman, R. 1983. In vitro studies of megakaryocytopoiesis in thrombocytotic disorders of man. *Blood.* **61**(2), pp.384-389.

Gil, V.S., Bhagat, G., Howell, L., Zhang, J., Kim, C.H., Stengel, S., Vega, F., Zelent, A. and Petrie, K. 2016. Deregulated expression of HDAC9 in B cells promotes development of lymphoproliferative disease and lymphoma in mice. *Dis Model Mech.* **9**(12), pp.1483-1495.

Godfrey, A.L., Chen, E., Massie, C.E., Silber, Y., Pagano, F., Bellosillo, B., Guglielmelli, P., Harrison, C.N., Reilly, J.T., Stegelmann, F., Bijou, F., Lippert, E., Boiron, J.-M., Döhner, K., Vannucchi, A.M., Besses, C. and Green, A.R. 2016. STAT1 activation in association with JAK2 exon 12 mutations. *Haematologica.* **101**(1), pp.e15-e19.

Godfrey, A.L., Chen, E., Pagano, F., Ortmann, C.A., Silber, Y., Bellosillo, B., Guglielmelli, P., Harrison, C.N., Reilly, J.T., Stegelmann, F., Bijou, F., Lippert, E., McMullin, M.F., Boiron, J.M., Döhner, K., Vannucchi, A.M., Besses, C., Campbell, P.J. and Green, A.R. 2012. JAK2V617F homozygosity arises commonly and recurrently in PV and ET, but PV is characterized by expansion of a dominant homozygous subclone. *Blood.* **120**(13), pp.2704-2707.

Griesshammer, M., Hornkohl, A., Nichol, J.L., Hecht, T., Raghavachar, A., Heimpel, H. and Schrezenmeier, H. 1998. High levels of thrombopoietin in sera of patients with essential thrombocythemia: cause or consequence of abnormal platelet production? *Ann Hematol.* **77**(5), pp.211-215.

Grisouard, J., Li, S., Kubovcakova, L., Rao, T.N., Meyer, S.C., Lundberg, P., Hao-Shen, H., Romanet, V., Murakami, M., Radimerski, T., Dirnhofer, S. and Skoda, R.C. 2016. JAK2 exon 12 mutant mice display isolated erythrocytosis and changes in iron metabolism favoring increased erythropoiesis. *Blood.* **128**(6), pp.839-851.

Grisouard, J., Shimizu, T., Duek, A., Kubovcakova, L., Hao-Shen, H., Dirnhofer, S. and Skoda, R.C. 2015. Deletion of Stat3 in hematopoietic cells enhances thrombocytosis and shortens survival in a JAK2-V617F mouse model of MPN. *Blood.* **125**(13), pp.2131-2140.

Han, A., He, J., wu, Y., Liu, J. and Chen, L. 2005. Han A, He J, Wu Y, Liu JO, Chen L.. Mechanism of recruitment of class II histone deacetylases by myocyte enhancer factor-2. J Mol Biol 345: 91-102. *Journal of molecular biology.* **345**, pp.91-102.

Hedley, B.D., Allan, A.L. and Xenocostas, A. 2011. The role of erythropoietin and erythropoiesis-stimulating agents in tumor progression. *Clin Cancer Res.* **17**(20), pp.6373-6380.

Heideman, M.R., Lancini, C., Proost, N., Yanover, E., Jacobs, H. and Dannenberg, J.H. 2014. Sin3a-associated Hdac1 and Hdac2 are essential for hematopoietic stem cell homeostasis and contribute differentially to hematopoiesis. *Haematologica.* **99**(8), pp.1292-1303.

Hess, G., Rose, P., Gamm, H., Papadileris, S., Huber, C. and Seliger, B. 1994. Molecular analysis of the erythropoietin receptor system in patients with polycythaemia vera. *Br J Haematol.* **88**(4), pp.794-802.

Heuck, G. 1879. Zwei fälle von Leukämie mit eigenthümlichem Blut-resp. Knochenmarksbefund. *Archiv für pathologische Anatomie und Physiologie und für klinische Medicin.* **78**(3), pp.475-496.

Hirai, H., Karian, P. and Kikyo, N. 2011. Regulation of embryonic stem cell self-renewal and pluripotency by leukaemia inhibitory factor. *Biochem J.* **438**(1), pp.11-23.

Hirsch, E.F. 1935. Generalized osteosclerosis with chronic polycythemia vera. *Arch Pathol.* **19**, pp.91-97.

Hitchcock, I.S. and Kaushansky, K. 2014. Thrombopoietin from beginning to end. *Br J Haematol.* **165**(2), pp.259-268.

Horikoshi, N., Sharma, D., Leonard, F., Pandita, R.K., Charaka, V.K., Hambarde, S., Horikoshi, N.T., Gaur Khaitan, P., Chakraborty, S., Cote, J., Godin, B., Hunt, C.R. and Pandita, T.K. 2019. Pre-existing H4K16ac levels in euchromatin drive DNA repair by homologous recombination in S-phase. *Communications Biology.* **2**(1), p253.

Howe, K.L., Achuthan, P., Allen, J., Allen, J., Alvarez-Jarreta, J., Amode, M.R., Armean, I.M., Azov, A.G., Bennett, R., Bhai, J., Billis, K., Boddu, S., Charkhchi, M., Cummins, C., Da Rin Fioretto, L., Davidson, C., Dodiya, K., El Houdaigui, B., Fatima, R., Gall, A., Garcia Giron, C., Grego, T., Guijarro-Clarke, C., Haggerty, L., Hemrom, A., Hourlier, T., Izuogu, O.G., Juettemann, T., Kaikala, V., Kay, M., Lavidas, I., Le, T., Lemos, D., Gonzalez Martinez, J., Marugán, J.C., Maurel, T., McMahon, A.C., Mohanan, S., Moore, B., Muffato, M., Oheh, D.N., Paraschas, D., Parker, A., Parton, A., Prosovetskaia, I., Sakthivel, M.P., Salam, A.I.A., Schmitt, B.M., Schuilenburg, H., Sheppard, D., Steed, E., Szpak, M., Szuba, M., Taylor, K., Thormann, A., Threadgold, G., Walts, B., Winterbottom, A., Chakiachvili, M., Chaubal, A., De Silva, N., Flint, B., Frankish, A., Hunt, S.E., GR, I.I., Langridge, N., Loveland, J.E., Martin, F.J., Mudge, J.M., Morales, J., Perry, E., Ruffier, M., Tate, J., Thybert, D., Trevanion, S.J.,

Cunningham, F., Yates, A.D., Zerbino, D.R. and Flicek, P. 2021. Ensembl 2021. *Nucleic Acids Res.* **49**(D1), pp.D884-d891.

Hsiao, K.Y. and Mizzen, C.A. 2013. Histone H4 deacetylation facilitates 53BP1 DNA damage signaling and double-strand break repair. *J Mol Cell Biol.* **5**(3), pp.157-165.

Hua, W.-K., Qi, J., Cai, Q., Carnahan, E., Ayala Ramirez, M., Li, L., Marcucci, G. and Kuo, Y.-H. 2017. HDAC8 regulates long-term hematopoietic stem-cell maintenance under stress by modulating p53 activity. *Blood.* **130**(24), pp.2619-2630.

Ihle, J.N. 1995. Cytokine receptor signalling. *Nature.* **377**(6550), pp.591-594.

Ihle, J.N. and Gilliland, D.G. 2007. Jak2: normal function and role in hematopoietic disorders. *Curr Opin Genet Dev.* **17**(1), pp.8-14.

Illumina. 2014. *A tour de force that includes a graphical abstract, a brief description and primary references for most sequencing methods.* [Online]. Available from: https://emea.illumina.com/techniques/sequencing/ngs-library-prep/library-prep-methods.html

Ishii, T., Bruno, E., Hoffman, R. and Xu, M. 2006. Involvement of various hematopoietic-cell lineages by the JAK2V617F mutation in polycythemia vera. *Blood.* **108**(9), pp.3128-3134.

Ishii, T., Zhao, Y., Sozer, S., Shi, J., Zhang, W., Hoffman, R. and Xu, M. 2007. Behavior of CD34+ cells isolated from patients with polycythemia vera in NOD/SCID mice. *Exp Hematol.* **35**(11), pp.1633-1640.

Jacobson, R.J., Salo, A. and Fialkow, P.J. 1978. Agnogenic myeloid metaplasia: a clonal proliferation of hematopoietic stem cells with secondary myelofibrosis. *Blood.* **51**(2), pp.189-194.

Jacquelin, S., Straube, J., Cooper, L., Vu, T., Song, A., Bywater, M., Baxter, E., Heidecker, M., Wackrow, B., Porter, A., Ling, V., Green, J., Austin, R., Kazakoff, S., Waddell, N., Hesson, L.B., Pimanda, J.E., Stegelmann, F., Bullinger, L., Döhner, K., Rampal, R.K., Heckl, D., Hill, G.R. and Lane, S.W. 2018. Jak2V617F and Dnmt3a loss cooperate to induce myelofibrosis through activated enhancer-driven inflammation. *Blood.* **132**(26), pp.2707-2721.

Jaiswal, S., Fontanillas, P., Flannick, J., Manning, A., Grauman, P.V., Mar, B.G., Lindsley, R.C., Mermel, C.H., Burtt, N., Chavez, A., Higgins, J.M., Moltchanov, V., Kuo, F.C., Kluk, M.J., Henderson, B., Kinnunen, L., Koistinen, H.A., Ladenvall, C., Getz, G., Correa, A., Banahan, B.F., Gabriel, S., Kathiresan, S., Stringham, H.M., McCarthy, M.I., Boehnke, M., Tuomilehto, J., Haiman, C., Groop, L., Atzmon, G., Wilson, J.G., Neuberg, D., Altshuler, D. and Ebert, B.L. 2014. Age-Related Clonal Hematopoiesis Associated with Adverse Outcomes. *New England Journal of Medicine.* **371**(26), pp.2488-2498.

James, C., Mazurier, F., Dupont, S., Chaligne, R., Lamrissi-Garcia, I., Tulliez, M., Lippert, E., Mahon, F.X., Pasquet, J.M., Etienne, G., Delhommeau, F.,

Giraudier, S., Vainchenker, W. and de Verneuil, H. 2008. The hematopoietic stem cell compartment of JAK2V617F-positive myeloproliferative disorders is a reflection of disease heterogeneity. *Blood.* **112**(6), pp.2429-2438.

James, C., Ugo, V., Le Couédic, J.P., Staerk, J., Delhommeau, F., Lacout, C., Garçon, L., Raslova, H., Berger, R., Bennaceur-Griscelli, A., Villeval, J.L., Constantinescu, S.N., Casadevall, N. and Vainchenker, W. 2005. A unique clonal JAK2 mutation leading to constitutive signalling causes polycythaemia vera. *Nature.* **434**(7037), pp.1144-1148.

Jamieson, C.H.M., Gotlib, J., Durocher, J.A., Chao, M.P., Mariappan, M.R., Lay, M., Jones, C., Zehnder, J.L., Lilleberg, S.L. and Weissman, I.L. 2006. The JAK2 V617F mutation occurs in hematopoietic stem cells in polycythemia vera and predisposes toward erythroid differentiation. *Proceedings of the National Academy of Sciences of the United States of America.* **103**(16), pp.6224-6229.

Jassal, B., Matthews, L., Viteri, G., Gong, C., Lorente, P., Fabregat, A., Sidiropoulos, K., Cook, J., Gillespie, M., Haw, R., Loney, F., May, B., Milacic, M., Rothfels, K., Sevilla, C., Shamovsky, V., Shorser, S., Varusai, T., Weiser, J., Wu, G., Stein, L., Hermjakob, H. and D'Eustachio, P. 2020. The reactome pathway knowledgebase. *Nucleic acids research.* **48**(D1), pp.D498-D503.

Jenkins, B.J., Blake, T.J. and Gonda, T.J. 1998. Saturation mutagenesis of the beta subunit of the human granulocyte-macrophage colony-stimulating factor receptor shows clustering of constitutive mutations, activation of ERK MAP kinase and STAT pathways, and differential beta subunit tyrosine phosphorylation. *Blood.* **92**(6), pp.1989-2002.

Ji, P., Yeh, V., Ramirez, T., Murata-Hori, M. and Lodish, H.F. 2010. Histone deacetylase 2 is required for chromatin condensation and subsequent enucleation of cultured mouse fetal erythroblasts. *Haematologica.* **95**(12), pp.2013-2021.

Juvonen, E., Ikkala, E., Oksanen, K. and Ruutu, T. 1993. Megakaryocyte and erythroid colony formation in essential thrombocythaemia and reactive thrombocytosis: diagnostic value and correlation to complications. *Br J Haematol.* **83**(2), pp.192-197.

Kameda, T., Shide, K., Yamaji, T., Kamiunten, A., Sekine, M., Hidaka, T., Kubuki, Y., Sashida, G., Aoyama, K., Yoshimitsu, M., Abe, H., Miike, T., Iwakiri, H., Tahara, Y., Yamamoto, S., Hasuike, S., Nagata, K., Iwama, A., Kitanaka, A. and Shimoda, K. 2015. Gene expression profiling of loss of TET2 and/or JAK2V617F mutant hematopoietic stem cells from mouse models of myeloproliferative neoplasms. *Genomics data.* **4**, pp.102-108.

Kanehisa, M. 2019. Toward understanding the origin and evolution of cellular organisms. *Protein Sci.* **28**(11), pp.1947-1951.

Kanehisa, M., Furumichi, M., Sato, Y., Ishiguro-Watanabe, M. and Tanabe, M. 2021. KEGG: integrating viruses and cellular organisms. *Nucleic Acids Res.* **49**(D1), pp.D545-d551.

Kanehisa, M. and Goto, S. 2000. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**(1), pp.27-30.

Kaushansky, K. 1997. Thrombopoietin the primary regulator of platelet production. *Trends Endocrinol Metab.* **8**(2), pp.45-50.

Kaushansky, K., Broudy, V.C., Grossmann, A., Humes, J., Lin, N., Ren, H.P., Bailey, M.C., Papayannopoulou, T., Forstrom, J.W. and Sprugel, K.H. 1995. Thrombopoietin expands erythroid progenitors, increases red cell production, and enhances erythroid recovery after myelosuppressive therapy. *The Journal of clinical investigation.* **96**(3), pp.1683-1687.

Kawahara, R. 2007. Leukopoiesis. In: Enna, S.J. and Bylund, D.B. eds. *xPharm: The Comprehensive Pharmacology Reference.* New York: Elsevier, pp.1-5.

Kawasaki, H., Nakano, T., Kohdera, U. and Kobayashi, Y. 2001. Hypersensitivity of megakaryocyte progenitors to thrombopoietin in essential thrombocythemia. *Am J Hematol.* **68**(3), pp.194-197.

Kent, D.G., Li, J., Tanna, H., Fink, J., Kirschner, K., Pask, D.C., Silber, Y., Hamilton, T.L., Sneade, R., Simons, B.D. and Green, A.R. 2013. Self-renewal of single mouse hematopoietic stem cells is reduced by JAK2V617F without compromising progenitor cell expansion. *PLoS Biol.* **11**(6), pe1001576.

Klampfl, T., Gisslinger, H., Harutyunyan, A.S., Nivarthi, H., Rumi, E., Milosevic, J.D., Them, N.C.C., Berg, T., Gisslinger, B., Pietra, D., Chen, D., Vladimer, G.I., Bagienski, K., Milanesi, C., Casetti, I.C., Sant'Antonio, E., Ferretti, V., Elena, C., Schischlik, F., Cleary, C., Six, M., Schalling, M., Schönegger, A., Bock, C., Malcovati, L., Pascutto, C., Superti-Furga, G., Cazzola, M. and Kralovics, R. 2013. Somatic Mutations of Calreticulin in Myeloproliferative Neoplasms. *New England Journal of Medicine.* **369**(25), pp.2379-2390.

Kobayashi, M., Laver, J.H., Kato, T., Miyazaki, H. and Ogawa, M. 1995. Recombinant human thrombopoietin (Mpl ligand) enhances proliferation of erythroid progenitors. *Blood.* **86**(7), pp.2494-2499.

Koppikar, P., Bhagwat, N., Kilpivaara, O., Manshouri, T., Adli, M., Hricik, T., Liu, F., Saunders, L.M., Mullally, A., Abdel-Wahab, O., Leung, L., Weinstein, A., Marubayashi, S., Goel, A., Gönen, M., Estrov, Z., Ebert, B.L., Chiosis, G., Nimer, S.D., Bernstein, B.E., Verstovsek, S. and Levine, R.L. 2012. Heterodimeric JAK–STAT activation as a mechanism of persistence to JAK2 inhibitor therapy. *Nature.* **489**(7414), pp.155-159.

Kotian, S., Liyanarachchi, S., Zelent, A. and Parvin, J.D. 2011. Histone deacetylases 9 and 10 are required for homologous recombination. *J Biol Chem.* **286**(10), pp.7722-7726.

Kotsantis, P., Petermann, E. and Boulton, S.J. 2018. Mechanisms of Oncogene-Induced Replication Stress: Jigsaw Falling into Place. *Cancer Discov.* **8**(5), pp.537-555.

Koyama-Nasu, R., Nasu-Nishimura, Y., Todo, T., Ino, Y., Saito, N., Aburatani, H., Funato, K., Echizen, K., Sugano, H., Haruta, R., Matsui, M., Takahashi, R., Manabe, E., Oda, T. and Akiyama, T. 2013. The critical role of cyclin D2 in cell cycle progression and tumorigenicity of glioblastoma stem cells. *Oncogene.* **32**(33), pp.3840-3845.

Kralovics, R., Guan, Y. and Prchal, J.T. 2002. Acquired uniparental disomy of chromosome 9p is a frequent stem cell defect in polycythemia vera. *Exp Hematol.* **30**(3), pp.229-236.

Kralovics, R., Passamonti, F., Buser, A.S., Teo, S.S., Tiedt, R., Passweg, J.R., Tichelli, A., Cazzola, M. and Skoda, R.C. 2005. A gain-of-function mutation of JAK2 in myeloproliferative disorders. *N Engl J Med.* **352**(17), pp.1779-1790.

Kreuger, F. 2021. *TrimGalore.* [Online]. Available from: https://github.com/FelixKrueger/TrimGalore

Lacombe, C. and Mayeux, P. 1999. The molecular biology of erythropoietin. *Nephrol Dial Transplant.* **14 Suppl 2**, pp.22-28.

Leko, V., Varnum-Finney, B., Li, H., Gu, Y., Flowers, D., Nourigat, C., Bernstein, I.D. and Bedalov, A. 2012. SIRT1 is dispensable for function of hematopoietic stem cells in adult mice. *Blood.* **119**(8), pp.1856-1860.

Lemercier, C., Brocard, M.P., Puvion-Dutilleul, F., Kao, H.Y., Albagli, O. and Khochbin, S. 2002. Class II histone deacetylases are directly recruited by BCL6 transcriptional repressor. *J Biol Chem.* **277**(24), pp.22045-22052.

Leroy, E., Defour, J.P., Sato, T., Dass, S., Gryshkova, V., Shwe, M.M., Staerk, J., Constantinescu, S.N. and Smith, S.O. 2016. His499 Regulates Dimerization and Prevents Oncogenic Activation by Asparagine Mutations of the Human Thrombopoietin Receptor. *J Biol Chem.* **291**(6), pp.2974-2987.

Levine, R.L. 2012. JAK-mutant myeloproliferative neoplasms. *Curr Top Microbiol Immunol.* **355**, pp.119-133.

Levine, R.L., Wadleigh, M., Cools, J., Ebert, B.L., Wernig, G., Huntly, B.J., Boggon, T.J., Wlodarska, I., Clark, J.J., Moore, S., Adelsperger, J., Koo, S., Lee, J.C., Gabriel, S., Mercher, T., D'Andrea, A., Fröhling, S., Döhner, K., Marynen, P., Vandenberghe, P., Mesa, R.A., Tefferi, A., Griffin, J.D., Eck, M.J., Sellers, W.R., Meyerson, M., Golub, T.R., Lee, S.J. and Gilliland, D.G. 2005. Activating mutation in the tyrosine kinase JAK2 in polycythemia vera, essential thrombocythemia, and myeloid metaplasia with myelofibrosis. *Cancer Cell.* **7**(4), pp.387-397.

Li, H., Li, X., Lin, H. and Gong, J. 2020. High HDAC9 is associated with poor prognosis and promotes malignant progression in pancreatic ductal adenocarcinoma. *Mol Med Rep.* **21**(2), pp.822-832.

Li, J., Kent, D.G., Chen, E. and Green, A.R. 2011. Mouse models of myeloproliferative neoplasms: JAK of all grades. *Dis Model Mech.* **4**(3), pp.311-317.

Li, J., Kent, D.G., Godfrey, A.L., Manning, H., Nangalia, J., Aziz, A., Chen, E., Saeb-Parsy, K., Fink, J., Sneade, R., Hamilton, T.L., Pask, D.C., Silber, Y., Zhao, X., Ghevaert, C., Liu, P. and Green, A.R. 2014. JAK2V617F homozygosity drives a phenotypic switch in myeloproliferative neoplasms, but is insufficient to sustain disease. *Blood.* **123**(20), pp.3139-3151.

Li, J., Spensberger, D., Ahn, J.S., Anand, S., Beer, P.A., Ghevaert, C., Chen, E., Forrai, A., Scott, L.M., Ferreira, R., Campbell, P.J., Watson, S.P., Liu, P., Erber, W.N., Huntly, B.J., Ottersbach, K. and Green, A.R. 2010. JAK2 V617F impairs hematopoietic stem cell function in a conditional knock-in mouse model of JAK2 V617F-positive essential thrombocythemia. *Blood.* **116**(9), pp.1528-1538.

Li, X., Mei, Y., Yan, B., Vitriol, E., Huang, S., Ji, P. and Qiu, Y. 2017. Histone deacetylase 6 regulates cytokinesis and erythrocyte enucleation through deacetylation of formin protein mDia2. *Haematologica.* **102**(6), pp.984-994.

Li, Z., Xu, M., Xing, S., Ho, W.T., Ishii, T., Li, Q., Fu, X. and Zhao, Z.J. 2007. Erlotinib effectively inhibits JAK2V617F activity and polycythemia vera cell growth. *The Journal of biological chemistry.* **282**(6), pp.3428-3432.

Liang, J., Wu, Y.L., Chen, B.J., Zhang, W., Tanaka, Y. and Sugiyama, H. 2013. The C-kit receptor-mediated signal transduction and tumor-related diseases. *Int J Biol Sci.* **9**(5), pp.435-443.

Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdóttir, H., Tamayo, P. and Mesirov, J.P. 2011. Molecular signatures database (MSigDB) 3.0. *Bioinformatics.* **27**(12), pp.1739-1740.

Lieschke, G.J., Grail, D., Hodgson, G., Metcalf, D., Stanley, E., Cheers, C., Fowler, K.J., Basu, S., Zhan, Y.F. and Dunn, A.R. 1994. Mice lacking granulocyte colony-stimulating factor have chronic neutropenia, granulocyte and macrophage progenitor cell deficiency, and impaired neutrophil mobilization. *Blood.* **84**(6), pp.1737-1746.

Liu, F., Zhao, X., Perna, F., Wang, L., Koppikar, P., Abdel-Wahab, O., Harr, M.W., Levine, R.L., Xu, H., Tefferi, A., Deblasio, A., Hatlen, M., Menendez, S. and Nimer, S.D. 2011. JAK2V617F-mediated phosphorylation of PRMT5 downregulates its methyltransferase activity and promotes myeloproliferation. *Cancer Cell.* **19**(2), pp.283-294.

Liu, F., Zong, M., Wen, X., Li, X., Wang, J., Wang, Y., Jiang, W., Li, X., Guo, Z. and Qi, H. 2016. Silencing of Histone Deacetylase 9 Expression in Podocytes Attenuates Kidney Injury in Diabetic Nephropathy. *Scientific Reports.* **6**(1), p33676.

Lorenz, E., Uphoff, D., Reid, T.R. and Shelton, E. 1951. Modification of Irradiation Injury in Mice and Guinea Pigs by Bone Marrow Injections. *JNCI: Journal of the National Cancer Institute.* **12**(1), pp.197-201.

Love, M.I., Huber, W. and Anders, S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology.* **15**(12), p550.

Lu, S., Li, H., Li, K. and Fan, X.D. 2018. HDAC9 promotes brain ischemic injury by provoking IκBα/NF-κB and MAPKs signaling pathways. *Biochem Biophys Res Commun.* **503**(3), pp.1322-1329.

Ma, L., Gao, J.-s., Guan, Y., Shi, X., Zhang, H., Ayrapetov, M.K., Zhang, Z., Xu, L., Hyun, Y.-M., Kim, M., Zhuang, S. and Chin, Y.E. 2010. Acetylation modulates prolactin receptor dimerization. *Proceedings of the National Academy of Sciences.* **107**(45), p19314.

Marino, V.J. and Roguin, L.P. 2008. The granulocyte colony stimulating factor (G-CSF) activates Jak/STAT and MAPK pathways in a trophoblastic cell line. *J Cell Biochem.* **103**(5), pp.1512-1523.

Martino, A., Holmes, J.H., Lord, J.D., Moon, J.J. and Nelson, B.H. 2001. Stat5 and Sp1 Regulate Transcription of the Cyclin D2 Gene in Response to IL-2. *The Journal of Immunology.* **166**(3), p1723.

Marty, C., Lacout, C., Droin, N., Le Couédic, J.P., Ribrag, V., Solary, E., Vainchenker, W., Villeval, J.L. and Plo, I. 2013. A role for reactive oxygen species in JAK2V617F myeloproliferative neoplasm progression. *Leukemia.* **27**(11), pp.2187-2195.

Means, R.T., Jr. 2008. Perspective: Osler's 1903 paper on polycythemia vera. *Am J Med Sci.* **335**(6), pp.418-419.

Meissl, K., Macho-Maschler, S., Müller, M. and Strobl, B. 2017. The good and the bad faces of STAT1 in solid tumours. *Cytokine.* **89**, pp.12-20.

Metcalf, D. 2008. Hematopoietic cytokines. *Blood.* **111**(2), pp.485-491.

Meyer, B., Fabbrizi, M.R., Raj, S., Zobel, C.L., Hallahan, D.E. and Sharma, G.G. 2016. Histone H3 Lysine 9 Acetylation Obstructs ATM Activation and Promotes Ionizing Radiation Sensitivity in Normal Stem Cells. *Stem Cell Reports.* **7**(6), pp.1013-1022.

Mittelman, M., Gardyn, J., Carmel, M., Malovani, H., Barak, Y. and Nir, U. 1996. Analysis of the erythropoietin receptor gene in patients with myeloproliferative and myelodysplastic syndromes. *Leuk Res.* **20**(6), pp.459-466.

Montagna, C., Massaro, P., Morali, F., Foa, P., Maiolo, A.T. and Eridani, S. 1994. In vitro sensitivity of human erythroid progenitors to hemopoietic growth factors: studies on primary and secondary polycythemia. *Haematologica.* **79**(4), pp.311-318.

Mootha, V.K., Lindgren, C.M., Eriksson, K.-F., Subramanian, A., Sihag, S., Lehar, J., Puigserver, P., Carlsson, E., Ridderstråle, M., Laurila, E., Houstis, N., Daly, M.J., Patterson, N., Mesirov, J.P., Golub, T.R., Tamayo, P., Spiegelman, B., Lander, E.S., Hirschhorn, J.N., Altshuler, D. and Groop, L.C. 2003. PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nature Genetics.* **34**(3), pp.267-273.

Mullally, A., Lane, S.W., Ball, B., Megerdichian, C., Okabe, R., Al-Shahrour, F., Paktinat, M., Haydu, J.E., Housman, E., Lord, A.M., Wernig, G., Kharas, M.G., Mercher, T., Kutok, J.L., Gilliland, D.G. and Ebert, B.L. 2010. Physiological Jak2V617F expression causes a lethal myeloproliferative neoplasm with differential effects on hematopoietic stem and progenitor cells. *Cancer Cell.* **17**(6), pp.584-596.

Murphy, J.M., Tannahill, G.M., Hilton, D.J. and Greenhalgh, C.J. 2010. Chapter 64 - The Negative Regulation of JAK/STAT Signaling. In: Bradshaw, R.A. and Dennis, E.A. eds. *Handbook of Cell Signaling (Second Edition).* San Diego: Academic Press, pp.467-480.

Nagao, T., Oshikawa, G., Wu, N., Kurosu, T. and Miura, O. 2011. DNA damage stress and inhibition of Jak2-V617F cause its degradation and synergistically induce apoptosis through activation of GSK3β. *PLoS One.* **6**(11), pe27397.

Naka, T., Narazaki, M., Hirata, M., Matsumoto, T., Minamoto, S., Aono, A., Nishimoto, N., Kajita, T., Taga, T., Yoshizaki, K., Akira, S. and Kishimoto, T. 1997. Structure and function of a new STAT-induced STAT inhibitor. *Nature.* **387**(6636), pp.924-929.

Nangalia, J. and Green, T.R. 2014. The evolving genomic landscape of myeloproliferative neoplasms. *Hematology Am Soc Hematol Educ Program.* **2014**(1), pp.287-296.

Nangalia, J., Massie, C.E., Baxter, E.J., Nice, F.L., Gundem, G., Wedge, D.C., Avezov, E., Li, J., Kollmann, K., Kent, D.G., Aziz, A., Godfrey, A.L., Hinton, J., Martincorena, I., Van Loo, P., Jones, A.V., Guglielmelli, P., Tarpey, P., Harding, H.P., Fitzpatrick, J.D., Goudie, C.T., Ortmann, C.A., Loughran, S.J., Raine, K., Jones, D.R., Butler, A.P., Teague, J.W., O'Meara, S., McLaren, S., Bianchi, M., Silber, Y., Dimitropoulou, D., Bloxham, D., Mudie, L., Maddison, M., Robinson, B., Keohane, C., Maclean, C., Hill, K., Orchard, K., Tauro, S., Du, M.Q., Greaves, M., Bowen, D., Huntly, B.J.P., Harrison, C.N., Cross, N.C.P., Ron, D., Vannucchi, A.M., Papaemmanuil, E., Campbell, P.J. and Green, A.R. 2013. Somatic CALR Mutations in Myeloproliferative Neoplasms with Nonmutated JAK2. *New England Journal of Medicine.* **369**(25), pp.2391-2405.

Ng, A.P., Kauppi, M., Metcalf, D., Hyland, C.D., Josefsson, E.C., Lebois, M., Zhang, J.G., Baldwin, T.M., Di Rago, L., Hilton, D.J. and Alexander, W.S. 2014. Mpl expression on megakaryocytes and platelets is dispensable for thrombopoiesis but essential to prevent myeloproliferation. *Proc Natl Acad Sci U S A.* **111**(16), pp.5884-5889.

Nieborowska-Skorska, M., Maifrede, S., Dasgupta, Y., Sullivan, K., Flis, S., Le, B.V., Solecka, M., Belyaeva, E.A., Kubovcakova, L., Nawrocki, M., Kirschner, M., Zhao, H., Prchal, J.T., Piwocka, K., Moliterno, A.R., Wasik, M., Koschmieder, S., Green, T.R., Skoda, R.C. and Skorski, T. 2017. Ruxolitinib-induced defects in DNA repair cause sensitivity to PARP inhibitors in myeloproliferative neoplasms. *Blood.* **130**(26), pp.2848-2859.

Orkin, S.H. 2000. Diversification of haematopoietic stem cells to specific lineages. *Nat Rev Genet.* **1**(1), pp.57-64.

Osler, W. 2008. Chronic cyanosis, with polycythaemia and enlarged spleen: a new clinical entity. 1903. *Am J Med Sci.* **335**(6), pp.411-417.

Ozawa, Y., Towatari, M., Tsuzuki, S., Hayakawa, F., Maeda, T., Miyata, Y., Tanimoto, M. and Saito, H. 2001. Histone deacetylase 3 associates with and represses the transcription factor GATA-2. *Blood.* **98**(7), pp.2116-2123.

Pardanani, A., Lasho, T.L., Finke, C., Hanson, C.A. and Tefferi, A. 2007. Prevalence and clinicopathologic correlates of JAK2 exon 12 mutations in JAK2V617F-negative polycythemia vera. *Leukemia.* **21**(9), pp.1960-1963.

Pardanani, A.D., Levine, R.L., Lasho, T., Pikman, Y., Mesa, R.A., Wadleigh, M., Steensma, D.P., Elliott, M.A., Wolanskyj, A.P., Hogan, W.J., McClure, R.F., Litzow, M.R., Gilliland, D.G. and Tefferi, A. 2006. MPL515 mutations in myeloproliferative and other myeloid disorders: a study of 1182 patients. *Blood.* **108**(10), pp.3472-3476.

Park, S.Y., Lee, C.J., Choi, J.H., Kim, J.H., Kim, J.W., Kim, J.Y. and Nam, J.S. 2019. The JAK2/STAT3/CCND2 Axis promotes colorectal Cancer stem cell persistence and radioresistance. *J Exp Clin Cancer Res.* **38**(1), p399.

Pear, W.S., Nolan, G.P., Scott, M.L. and Baltimore, D. 1993. Production of high-titer helper-free retroviruses by transient transfection. *Proceedings of the National Academy of Sciences of the United States of America.* **90**(18), pp.8392-8396.

Pecquet, C., Staerk, J., Chaligné, R., Goss, V., Lee, K.A., Zhang, X., Rush, J., Van Hees, J., Poirel, H.A., Scheiff, J.M., Vainchenker, W., Giraudier, S., Polakiewicz, R.D. and Constantinescu, S.N. 2010. Induction of myeloproliferative disorder and myelofibrosis by thrombopoietin receptor W515 mutants is mediated by cytosolic tyrosine 112 of the receptor. *Blood.* **115**(5), pp.1037-1048.

Petrie, K., Guidez, F., Howell, L., Healy, L., Waxman, S., Greaves, M. and Zelent, A. 2003. The histone deacetylase 9 gene encodes multiple protein isoforms. *J Biol Chem.* **278**(18), pp.16059-16072.

Pikman, Y., Lee, B.H., Mercher, T., McDowell, E., Ebert, B.L., Gozo, M., Cuker, A., Wernig, G., Moore, S., Galinsky, I., DeAngelo, D.J., Clark, J.J., Lee, S.J., Golub, T.R., Wadleigh, M., Gilliland, D.G. and Levine, R.L. 2006. MPLW515L is a novel somatic activating mutation in myelofibrosis with myeloid metaplasia. *PLoS Med.* **3**(7), pe270.

Pinto do, O.P., Kolterud, A. and Carlsson, L. 1998. Expression of the LIM-homeobox gene LH2 generates immortalized steel factor-dependent multipotent hematopoietic precursors. *Embo j.* **17**(19), pp.5744-5756.

Pinz, S., Unser, S., Buob, D., Fischer, P., Jobst, B. and Rascle, A. 2015. Deacetylase inhibitors repress STAT5-mediated transcription by interfering with bromodomain and extra-terminal (BET) protein function. *Nucleic acids research.* **43**(7), pp.3524-3545.

Plo, I., Nakatake, M., Malivert, L., de Villartay, J.P., Giraudier, S., Villeval, J.L., Wiesmuller, L. and Vainchenker, W. 2008. JAK2 stimulates homologous recombination and genetic instability: potential implication in the heterogeneity of myeloproliferative disorders. *Blood.* **112**(4), pp.1402-1412.

Porter, F.D., Drago, J., Xu, Y., Cheema, S.S., Wassif, C., Huang, S.P., Lee, E., Grinberg, A., Massalas, J.S., Bodine, D., Alt, F. and Westphal, H. 1997. Lhx2, a LIM homeobox gene, is required for eye, forebrain, and definitive erythrocyte development. *Development.* **124**(15), pp.2935-2944.

Prchal, J.F. and Axelrad, A.A. 1974. Letter: Bone-marrow responses in polycythemia vera. *N Engl J Med.* **290**(24), p1382.

Qiao, Y., Spitz, M.R., Guo, Z., Hadeyati, M., Grossman, L., Kraemer, K.H. and Wei, Q. 2002. Rapid assessment of repair of ultraviolet DNA damage with a modified host-cell reactivation assay using a luciferase reporter gene and correlation with polymorphisms of DNA repair genes in normal human lymphocytes. *Mutat Res.* **509**(1-2), pp.165-174.

Randi, M.L., Putti, M.C., Pacquola, E., Luzzatto, G., Zanesco, L. and Fabris, F. 2005. Normal thrombopoietin and its receptor (c-mpl) genes in children with essential thrombocythemia. *Pediatr Blood Cancer.* **44**(1), pp.47-50.

Rao, T.N., Hansen, N., Hilfiker, J., Rai, S., Majewska, J.-M., Leković, D., Gezer, D., Andina, N., Galli, S., Cassel, T., Geier, F., Delezie, J., Nienhold, R., Hao-Shen, H., Beisel, C., Di Palma, S., Dimeloe, S., Trebicka, J., Wolf, D., Gassmann, M., Fan, T.W.M., Lane, A.N., Handschin, C., Dirnhofer, S., Kröger, N., Hess, C., Radimerski, T., Koschmieder, S., Čokić, V.P. and Skoda, R.C. 2019. JAK2-mutant hematopoietic cells display metabolic alterations that can be targeted to treat myeloproliferative neoplasms. *Blood.* **134**(21), pp.1832-1846.

Rawlings, J.S., Rosler, K.M. and Harrison, D.A. 2004. The JAK/STAT signaling pathway. *J Cell Sci.* **117**(Pt 8), pp.1281-1283.

Ray-David, H., Romeo, Y., Lavoie, G., Déléris, P., Tcherkezian, J., Galan, J.A. and Roux, P.P. 2013. RSK promotes G2 DNA damage checkpoint silencing and participates in melanoma chemoresistance. *Oncogene.* **32**(38), pp.4480-4489.

Rimmelé, P., Bigarella, C.L., Liang, R., Izac, B., Dieguez-Gonzalez, R., Barbet, G., Donovan, M., Brugnara, C., Blander, J.M., Sinclair, D.A. and Ghaffari, S. 2014. Aging-like phenotype and defective lineage specification in SIRT1-deleted hematopoietic stem and progenitor cells. *Stem Cell Reports.* **3**(1), pp.44-59.

Rivera, J.F., Baral, A.J., Nadat, F., Boyd, G., Smyth, R., Patel, H., Burman, E.L., Alameer, G., Boxall, S.A., Jackson, B.R., Baxter, E.J., Laslo, P., Green, A.R., Kent, D.G., Mullally, A. and Chen, E. 2021. Zinc-dependent multimerization of mutant calreticulin is required for MPL binding and MPN pathogenesis. *Blood Adv.* **5**(7), pp.1922-1932.

Robb, L. 2007. Cytokine receptors and hematopoietic differentiation. *Oncogene.* **26**(47), pp.6715-6723.

Rojnuckarin, P., Drachman, J.G. and Kaushansky, K. 1999. Thrombopoietin-induced activation of the mitogen-activated protein kinase (MAPK) pathway in normal megakaryocytes: role in endomitosis. *Blood.* **94**(4), pp.1273-1282.

Rossi, D.J., Seita, J., Czechowicz, A., Bhattacharya, D., Bryder, D. and Weissman, I.L. 2007. Hematopoietic stem cell quiescence attenuates DNA damage response and permits DNA damage accumulation during aging. *Cell Cycle.* **6**(19), pp.2371-2376.

Sabath, D.F., Kaushansky, K. and Broudy, V.C. 1999. Deletion of the extracellular membrane-distal cytokine receptor homology module of Mpl results in constitutive cell growth and loss of thrombopoietin binding. *Blood.* **94**(1), pp.365-367.

Saharinen, P. and Silvennoinen, O. 2002. The pseudokinase domain is required for suppression of basal activity of Jak2 and Jak3 tyrosine kinases and for cytokine-inducible activation of signal transduction. *J Biol Chem.* **277**(49), pp.47954-47963.

Schneider, C.A., Rasband, W.S. and Eliceiri, K.W. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods.* **9**(7), pp.671-675.

Schütte, J., Wang, H., Antoniou, S., Jarratt, A., Wilson, N.K., Riepsaame, J., Calero-Nieto, F.J., Moignard, V., Basilico, S., Kinston, S.J., Hannah, R.L., Chan, M.C., Nürnberg, S.T., Ouwehand, W.H., Bonzanni, N., de Bruijn, M.F. and Göttgens, B. 2016. An experimentally validated network of nine haematopoietic transcription factors reveals mechanisms of cell state stability. *Elife.* **5**, pe11469.

Scott, L.M. and Rebel, V.I. 2012. JAK2 and genomic instability in the myeloproliferative neoplasms: a case of the chicken or the egg? *Am J Hematol.* **87**(11), pp.1028-1036.

Scott, L.M., Tong, W., Levine, R.L., Scott, M.A., Beer, P.A., Stratton, M.R., Futreal, P.A., Erber, W.N., McMullin, M.F., Harrison, C.N., Warren, A.J., Gilliland, D.G., Lodish, H.F. and Green, A.R. 2007. JAK2 Exon 12 Mutations in Polycythemia Vera and Idiopathic Erythrocytosis. *New England Journal of Medicine.* **356**(5), pp.459-468.

Shi, Y., Liu, C.H., Roberts, A.I., Das, J., Xu, G., Ren, G., Zhang, Y., Zhang, L., Yuan, Z.R., Tan, H.S., Das, G. and Devadas, S. 2006. Granulocyte-macrophage colony-stimulating factor (GM-CSF) and T-cell responses: what we do and don't know. *Cell Res.* **16**(2), pp.126-133.

Shide, K., Shimoda, H.K., Kumano, T., Karube, K., Kameda, T., Takenaka, K., Oku, S., Abe, H., Katayose, K.S., Kubuki, Y., Kusumoto, K., Hasuike, S., Tahara, Y., Nagata, K., Matsuda, T., Ohshima, K., Harada, M. and Shimoda, K. 2008. Development of ET, primary myelofibrosis and PV in mice expressing JAK2 V617F. *Leukemia.* **22**(1), pp.87-95.

Shimoda, K., Iwasaki, H., Okamura, S., Ohno, Y., Kubota, A., Arima, F., Otsuka, T. and Niho, Y. 1994. G-CSF induces tyrosine phosphorylation of the JAK2

protein in the human myeloid G-CSF responsive and proliferative cells, but not in mature neutrophils. *Biochem Biophys Res Commun.* **203**(2), pp.922-928.

Shroff, N., Ander, B.P., Zhan, X., Stamova, B., Liu, D., Hull, H., Hamade, F.R., Dykstra-Aiello, C., Ng, K., Sharp, F.R. and Jickling, G.C. 2019. HDAC9 Polymorphism Alters Blood Gene Expression in Patients with Large Vessel Atherosclerotic Stroke. *Transl Stroke Res.* **10**(1), pp.19-25.

Skourti-Stathaki, K. and Proudfoot, N.J. 2014. A double-edged sword: R loops as threats to genome integrity and powerful regulators of gene expression. *Genes Dev.* **28**(13), pp.1384-1396.

Skov, V., Larsen, T.S., Thomassen, M., Riley, C.H., Jensen, M.K., Bjerrum, O.W., Kruse, T.A. and Hasselbalch, H.C. 2012. Increased gene expression of histone deacetylases in patients with Philadelphia-negative chronic myeloproliferative neoplasms. *Leuk Lymphoma.* **53**(1), pp.123-129.

Smith, J.D. 2014. New role for histone deacetylase 9 in atherosclerosis and inflammation. *Arterioscler Thromb Vasc Biol.* **34**(9), pp.1798-1799.

Spivak, J.L. 2005. The anaemia of cancer: death by a thousand cuts. *Nat Rev Cancer.* **5**(7), pp.543-555.

Staerk, J. and Constantinescu, S.N. 2012. The JAK-STAT pathway and hematopoietic stem cells from the JAK2 V617F perspective. *Jakstat.* **1**(3), pp.184-190.

Staerk, J., Lacout, C., Sato, T., Smith, S.O., Vainchenker, W. and Constantinescu, S.N. 2006. An amphipathic motif at the transmembrane-cytoplasmic junction prevents autonomous activation of the thrombopoietin receptor. *Blood.* **107**(5), pp.1864-1871.

Stark, R., Grzelak, M. and Hadfield, J. 2019. RNA sequencing: the teenage years. *Nature Reviews Genetics.* **20**(11), pp.631-656.

Starr, R. and Hilton, D.J. 1999. Negative regulation of the JAK/STAT pathway. *Bioessays.* **21**(1), pp.47-52.

Stelzer, G., Rosen, N., Plaschkes, I., Zimmerman, S., Twik, M., Fishilevich, S., Stein, T.I., Nudel, R., Lieder, I., Mazor, Y., Kaplan, S., Dahary, D., Warshawsky, D., Guan-Golan, Y., Kohn, A., Rappaport, N., Safran, M. and Lancet, D. 2016. The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. *Curr Protoc Bioinformatics.* **54**, pp.1.30.31-31.30.33.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S. and Mesirov, J.P. 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A.* **102**(43), pp.15545-15550.

Tarazona, S., Furió-Tarí, P., Turrà, D., Pietro, A.D., Nueda, M.J., Ferrer, A. and Conesa, A. 2015. Data quality aware analysis of differential expression in RNA-seq with NOISeq R/Bioc package. *Nucleic Acids Res.* **43**(21), pe140.

Team, R.c. 2013. *R: A language and environment for statistical computing. R Foundation for Statistical Computing.* [Online]. Available from: URL http://www.R-project.org/.

Tefferi, A. 2008. Mutant Molecules of Interest in Myeloproliferative Neoplasms: Introduction. *Acta Haematologica.* **119**(4), pp.192-193.

The UniProt, C. 2021. UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Research.* **49**(D1), pp.D480-D489.

Traver, D. and Akashi, K. 2004. Lineage Commitment and Developmental Plasticity in Early Lymphoid Progenitor Subsets. *Advances in Immunology.* Academic Press, pp.1-54.

Truebestein, L. and Leonard, T.A. 2016. Coiled-coils: The long and short of it. *Bioessays.* **38**(9), pp.903-916.

Ueda, F., Sumi, K., Tago, K., Kasahara, T. and Funakoshi-Tago, M. 2013. Critical role of FANCC in JAK2 V617F mutant-induced resistance to DNA cross-linking drugs. *Cell Signal.* **25**(11), pp.2115-2124.

Ungureanu, D., Vanhatupa, S., Kotaja, N., Yang, J., Aittomaki, S., Jänne, O.A., Palvimo, J.J. and Silvennoinen, O. 2003. PIAS proteins promote SUMO-1 conjugation to STAT1. *Blood.* **102**(9), pp.3311-3313.

Vaquez, H. 1892. Sur une forme spéciale de cyanose s' accompagnant d'hyperglobulie excessive et persistante. *CR Soc Biol (Paris).* **44**, pp.384-388.

Varghese, L.N., Defour, J.P., Pecquet, C. and Constantinescu, S.N. 2017. The Thrombopoietin Receptor: Structural Basis of Traffic and Activation by Ligand, Mutations, Agonists, and Mutated Calreticulin. *Front Endocrinol (Lausanne).* **8**, p59.

Vaughan, J.M. and Harrison, C.V. 1939. Leuco-erythroblastic anæmia and myelosclerosis. *The Journal of Pathology and Bacteriology.* **48**(2), pp.339-352.

Walz, C., Ahmed, W., Lazarides, K., Betancur, M., Patel, N., Hennighausen, L., Zaleskas, V.M. and Van Etten, R.A. 2012. Essential role for Stat5a/b in myeloproliferative neoplasms induced by BCR-ABL1 and JAK2(V617F) in mice. *Blood.* **119**(15), pp.3550-3560.

Wang, J.C. and Hashmi, G. 2003. Elevated thrombopoietin levels in patients with myelofibrosis may not be due to enhanced production of thrombopoietin by bone marrow. *Leuk Res.* **27**(1), pp.13-17.

Wang, P., Wang, Z. and Liu, J. 2020. Role of HDACs in normal and malignant hematopoiesis. *Molecular Cancer.* **19**(1), p5.

Warr, M.R., Pietras, E.M. and Passegué, E. 2011. Mechanisms controlling hematopoietic stem cell functions during normal hematopoiesis and hematological malignancies. *Wiley Interdiscip Rev Syst Biol Med.* **3**(6), pp.681-701.

Watamoto, K., Towatari, M., Ozawa, Y., Miyata, Y., Okamoto, M., Abe, A., Naoe, T. and Saito, H. 2003. Altered interaction of HDAC5 with GATA-1 during MEL cell differentiation. *Oncogene.* **22**(57), pp.9176-9184.

Wei, Y., Ma, D., Gao, Y., Zhang, C., Wang, L. and Liu, F. 2014. Ncor2 is required for hematopoietic stem cell emergence by inhibiting Fos signaling in zebrafish. *Blood.* **124**(10), pp.1578-1585.

Wierenga, A.T., Vellenga, E. and Schuringa, J.J. 2008. Maximal STAT5-induced proliferation and self-renewal at intermediate STAT5 activity levels. *Mol Cell Biol.* **28**(21), pp.6668-6680.

Williams, D.M., Kim, A.H., Rogers, O., Spivak, J.L. and Moliterno, A.R. 2007. Phenotypic variations and new mutations in JAK2 V617F-negative polycythemia vera, erythrocytosis, and idiopathic myelofibrosis. *Exp Hematol.* **35**(11), pp.1641-1646.

Wilmes, S., Hafer, M., Vuorio, J., Tucker, J.A., Winkelmann, H., Löchte, S., Stanly, T.A., Pulgar Prieto, K.D., Poojari, C., Sharma, V., Richter, C.P., Kurre, R., Hubbard, S.R., Garcia, K.C., Moraga, I., Vattulainen, I., Hitchcock, I.S. and Piehler, J. 2020. Mechanism of homodimeric cytokine receptor activation and dysregulation by oncogenic mutations. *Science.* **367**(6478), pp.643-652.

Wilson, A., Laurenti, E., Oser, G., van der Wath, R.C., Blanco-Bose, W., Jaworski, M., Offner, S., Dunant, C.F., Eshkind, L., Bockamp, E., Lió, P., Macdonald, H.R. and Trumpp, A. 2008. Hematopoietic stem cells reversibly switch from dormancy to self-renewal during homeostasis and repair. *Cell.* **135**(6), pp.1118-1129.

Wilson, A., Murphy, M.J., Oskarsson, T., Kaloulis, K., Bettess, M.D., Oser, G.M., Pasche, A.C., Knabenhans, C., Macdonald, H.R. and Trumpp, A. 2004. c-Myc controls the balance between hematopoietic stem cell self-renewal and differentiation. *Genes Dev.* **18**(22), pp.2747-2763.

Wingelhofer, B., Neubauer, H.A., Valent, P., Han, X., Constantinescu, S.N., Gunning, P.T., Müller, M. and Moriggl, R. 2018. Implications of STAT3 and STAT5 signaling on gene regulation and chromatin remodeling in hematopoietic cancer. *Leukemia.* **32**(8), pp.1713-1726.

Witthuhn, B.A., Quelle, F.W., Silvennoinen, O., Yi, T., Tang, B., Miura, O. and Ihle, J.N. 1993. JAK2 associates with the erythropoietin receptor and is tyrosine phosphorylated and activated following stimulation with erythropoietin. *Cell.* **74**(2), pp.227-236.

Xie, M., Lu, C., Wang, J., McLellan, M.D., Johnson, K.J., Wendl, M.C., McMichael, J.F., Schmidt, H.K., Yellapantula, V., Miller, C.A., Ozenberger, B.A., Welch, J.S., Link, D.C., Walter, M.J., Mardis, E.R., Dipersio, J.F., Chen, F.,

Wilson, R.K., Ley, T.J. and Ding, L. 2014. Age-related mutations associated with clonal hematopoietic expansion and malignancies. *Nature Medicine.* **20**(12), pp.1472-1478.

Xing, S., Wanting, T.H., Zhao, W., Ma, J., Wang, S., Xu, X., Li, Q., Fu, X., Xu, M. and Zhao, Z.J. 2008. Transgenic expression of JAK2V617F causes myeloproliferative disorders in mice. *Blood.* **111**(10), pp.5109-5117.

Xiong, K., Zhang, H., Du, Y., Tian, J. and Ding, S. 2019. Identification of HDAC9 as a viable therapeutic target for the treatment of gastric cancer. *Experimental & Molecular Medicine.* **51**(8), pp.1-15.

Xu, Y., Baldassare, M., Fisher, P., Rathbun, G., Oltz, E.M., Yancopoulos, G.D., Jessell, T.M. and Alt, F.W. 1993. LH-2: a LIM/homeodomain gene expressed in developing lymphocytes and neural cells. *Proc Natl Acad Sci U S A.* **90**(1), pp.227-231.

Yamaoka, K., Saharinen, P., Pesu, M., Holt, V.E.T., Silvennoinen, O. and O'Shea, J.J. 2004. The Janus kinases (Jaks). *Genome Biology.* **5**(12), p253.

Yan, D., Hutchison, R.E. and Mohi, G. 2012. Critical requirement for Stat5 in a mouse model of polycythemia vera. *Blood.* **119**(15), pp.3539-3549.

Yan, D., Jobe, F., Hutchison, R.E. and Mohi, G. 2015. Deletion of Stat3 enhances myeloid cell expansion and increases the severity of myeloproliferative neoplasms in Jak2V617F knock-in mice. *Leukemia.* **29**(10), pp.2050-2061.

Yan, D. and Mohi, G. 2013. Stat3 Negatively Regulates Myeloproliferative Neoplasm Induced By Jak2V617F. *Blood.* **122**(21), pp.111-111.

Yang, R., Wu, Y., Wang, M., Sun, Z., Zou, J., Zhang, Y. and Cui, H. 2015. HDAC9 promotes glioblastoma growth via TAZ-mediated EGFR pathway activation. *Oncotarget.* **6**(10), pp.7644-7656.

Yang, Y., Akada, H., Nath, D., Hutchison, R.E. and Mohi, G. 2016. Loss of Ezh2 cooperates with Jak2V617F in the development of myelofibrosis in a mouse model of myeloproliferative neoplasm. *Blood.* **127**(26), pp.3410-3423.

Yoshihara, H., Arai, F., Hosokawa, K., Hagiwara, T., Takubo, K., Nakamura, Y., Gomei, Y., Iwasaki, H., Matsuoka, S., Miyamoto, K., Miyazaki, H., Takahashi, T. and Suda, T. 2007. Thrombopoietin/MPL signaling regulates hematopoietic stem cell quiescence and interaction with the osteoblastic niche. *Cell Stem Cell.* **1**(6), pp.685-697.

Yuan, Z., Peng, L., Radhakrishnan, R. and Seto, E. 2010. Histone deacetylase 9 (HDAC9) regulates the functions of the ATDC (TRIM29) protein. *The Journal of biological chemistry.* **285**(50), pp.39329-39338.

Zanjani, E.D., Lutton, J.D., Hoffman, R. and Wasserman, L.R. 1977. Erythroid colony formation by polycythemia vera bone marrow in vitro. Dependence on erythropoietin. *J Clin Invest.* **59**(5), pp.841-848.

Zhan, H., Lin, C.H.S., Segal, Y. and Kaushansky, K. 2018. The JAK2V617F-bearing vascular niche promotes clonal expansion in myeloproliferative neoplasms. *Leukemia.* **32**(2), pp.462-469.

Zhan, H., Ma, Y., Lin, C.H. and Kaushansky, K. 2016. JAK2(V617F)-mutant megakaryocytes contribute to hematopoietic stem/progenitor cell expansion in a model of murine myeloproliferation. *Leukemia.* **30**(12), pp.2332-2341.

Zini, R., Guglielmelli, P., Pietra, D., Rumi, E., Rossi, C., Rontauroli, S., Genovese, E., Fanelli, T., Calabresi, L., Bianchi, E., Salati, S., Cazzola, M., Tagliafico, E., Vannucchi, A.M., Manfredini, R. and on behalf of the, A.i. 2017. CALR mutational status identifies different disease subtypes of essential thrombocythemia showing distinct expression profiles. *Blood Cancer Journal.* **7**(12), p638.

Zivot, A., Lipton, J.M., Narla, A. and Blanc, L. 2018. Erythropoiesis: insights into pathophysiology and treatments in 2017. *Molecular medicine (Cambridge, Mass.).* **24**(1), pp.11-11.

Zou, H., Yan, D. and Mohi, G. 2011. Differential biological activity of disease-associated JAK2 mutants. *FEBS letters.* **585**(7), pp.1007-1013.