A role for histone H3, histone H4 and histone associating proteins DNMT3A and PHF6 in *JAK2V617F* positive myeloproliferative neoplasms

---

Nisha Rao

Bachelor of Science (Hons) Molecular Biology

November 2014

Thesis submitted for the degree of Doctor of Philosophy in the discipline of Genetics at the University of Adelaide

Table of Contents

ABSTRACT ..... **IX**

DECLARATION ..... **XI**

ERRATA SHEET ..... **XIII**

ACKNOWLEDGEMENTS ..... **XV**

ABBREVIATIONS ..... **XVII**

CHAPTER 1: INTRODUCTION **1**

1.1 HAEMATOPOIESIS 1

1.1.1 Regulation of myelopoiesis *3*

A. The bone marrow (BM) niche 3

B. Haematopoietic growth factors (HGFs) and their cognate receptors 5

C. Transcription factors in myelopoiesis 8

1.2 MALIGNANT MYELOID DISORDERS 10

1.3 PHILADELPHIA CHROMOSOME-NEGATIVE MPN: PV, ET AND MF 11

1.3.1 PV *13*

1.3.2 ET *15*

1.3.3 MF: PMF and post-PV/ET-MF *18*

1.3.4 Identification of Janus Kinase 2 (JAK2) mutations in MPN *21*

1.4 JANUS KINASE (JAK) FAMILY PROTEINS 22

1.4.1 JAK2 structure and function *22*

1.4.2 JAK2 as a chromatin modifier *25*

1.4.3 JAK2 mutations in MPN *26*

1.5. MUTATIONS HIGHLY SPECIFIC TO ET AND PMF 28

1.5.1 MPL 28

1.5.2 CALR 29
1.6 OTHER SOMATIC MUTATIONS IN PV, ET AND PMF 30
1.6.1 TET2 31
1.6.2 IDH1 and IDH2 32
1.6.3 DNMT3A 33
1.6.4 ASXL1 36
1.6.5 EZH2 37
1.7 EVIDENCE FOR PRE-JAK2 MUTATION EVENT 38
1.7.1 JAK2V617F allele burden drives MPN disease phenotype: the gene dosage hypothesis. *38*
1.7.2 Additional events which may be important in driving the JAK2V617F disease phenotype: transformation to MF and AML *40*
1.7.2.1 Evidence that clonal haematopoiesis is a pre-JAK2 mutation event . 41
1.7.2.2 Evidence from familial studies 42
1.7.2.3 JAK2 haplotype: a pre-disposition loci in sporadic MPN 44
1.8 HISTONES AND MPN DISEASE PATHOGENESIS 46
1.9 HISTONES AND HISTONE GENES 47
1.9.1 Organisation of histone genes and regulation of transcription *48*
1.9.2. Histone post-translational modifications and the histone code hypothesis *49*
1.10 HISTONE GENE MUTATIONS IN GLIOMAS 51
1.11 HYPOTHESIS 53
1.12 PROJECT AIMS 53
CHAPTER 2: IDENTIFICATION OF NOVEL *HISTONE H4* VARIANTS IN *JAK2V617F* POSITIVE POLYCYTHEMIA VERA 54
2.1 INTRODUCTION 57

2.2 MATERIALS AND METHODS 59
2.2.1 MPN, normal and other cohorts *59*
2.2.2 Cell lines and culture conditions *60*
A. Human embryonic kidney-293 and -293T (HEK293/HEK293T) 60
B. Human erythromyeloblastoid leukaemia cell line (K-562) 61
C. Human Erythroleukaemia cell line 92.1.7 (HEL) 61
D. Mouse Foetal Derived Myeloid cell line (FDM) 61
2.2.3 Cloning, transfection and transduction of cell lines *61*
A. Cloning and lentiviral expression of HIST1H4C in HEK293 61
B. Cloning and retroviral expression of HIST1H4C in FDM 64
2.2.4 Primary Cell separation techniques *64*
A. Isolation of peripheral blood mononuclear cells (PBMNC), granulocytes and bone marrow mononuclear cells (BMMNC) fractions 65
B. Isolation of monocyte and granulocyte enriched cell fractions from PBMNC and haematopoietic stem cell/progenitor cells (HSPC) enriched cell fraction from BMMNC 65
2.2.5 Nucleic Acid Extraction *66*
A. Genomic DNA (gDNA) *66*
B. RNA extraction and cDNA preparation from primary cells and cell lines *66*
2.2.6 Genomic PCR amplification and Sanger Sequencing *67*
2.2.7 Flow cytometry and fluorophore conjugated antibodies *68*
2.2.8 Propidium Iodide (PI) staining of FDM cells *68*
2.2.9 Acid extraction of histones *69*
2.2.10 Western blot and antibodies *69*
2.2.11 Statistics *70*
2.3 RESULTS 70

2.3.1 Identification of novel histone H4 variants in patients with Polycythemia vera *70*

2.3.2 Characterisation of the HIST1H4C: p.R4C and p.R56Q variants detected in two patients *72*

2.3.3 Frequency of HIST1H4C p.R4C substitution in DNA samples from control populations *75*

2.3.4 HIST1H4C is a major contributor of the histone H4 mRNA in myeloid leukaemic cell lines *78*

2.3.5 HIST1H4C is expressed at significant levels in normal bone marrow CD34$^+$ cells and peripheral blood CD14$^+$ cells *79*

2.3.6 Generation of an inducible expression system for the functional analysis of p.R4C *84*

2.3.7 Analysis of histone H4 following induction of HIST1H4C variants in HEK293 inducible cells *88*

2.3.8 Changes in gene expression following over-expression of HIST1H4C:p.R4C in FDM cells *90*

2.4 DISCUSSION 94

2.5 SUPPLEMENTARY INFORMATION 97

CHAPTER 3: *PHF6* AND *HISTONE H3* MUTATION SCREENING OF MPN PATIENTS 112

STATEMENT OF AUTHORSHIP 112

3.1 INTRODUCTION 115

3.1.1 PHF6 ..... *115*

3.1.2 Histone H3 ..... *118*

3.2 MATERIALS AND METHODS: 120

3.2.1 Patients and genomic DNA (gDNA) preparation *120*

3.2.2 PCR amplification and Sanger sequencing of PHF6 exon 9 and exon 10 ..... *120*

3.2.3 Blast Forming Units Erythroid (BFUE) colony forming assays *121*

3.2.4 Roche 454 sequencing platform *121*

A. Sample preparation for PCR amplification 121

B. Primer Design and PCR amplification 122

C. Amplicon quantitation, pooling and purification 122

D. Next generation amplicon library sequencing 122

3.3 RESULTS 123

3.3.1 Sanger sequencing identified a PHF6R335fs mutation in exon 10 in a JAK2V617F^{+} PV patient *123*

3.3.2 Patient PV108 BFUE colony genotyping *125*

3.3.3 Next generation sequencing approach: mapping and filtering for potential variants *130*

3.3.4 Identification and validation of novel variants *132*

3.4 DISCUSSION: 135

3.5 SUPPLEMENTARY DATA: 139

CHAPTER 4: CLONAL AND LINEAGE ANALYSIS OF SOMATIC *DNMT3A* AND *JAK2* MUTATIONS IN A CHRONIC PHASE POLYCYTHEMIA VERA PATIENT 153

4.1 SUMMARY 156

CHAPTER 5: SUMMARY AND SIGNIFICANCE 159

5.1 ROLE OF HISTONE H4 AND DNMT3A VARIANTS IN MPN 159

5.2 IDENTIFICATION OF HISTONE H3 AND PHF6 VARIANTS USING SANGER AND NEXT-GENERATION SEQUENCING APPROACH 163

5.2.1 Histone H3 163

5.2.2 PHF6 168

RESPONSE TO EXAMINER'S COMMENTS: 171

REFERENCES: 172

Abstract

The Philadelphia chromosome negative myeloproliferative neoplasms (MPN); polycythemia vera, essential thrombocythemia and primary myelofibrosis, are clonal disorders harbouring the specific Janus kinase 2 (*JAK2*) lesion (*JAK2V617F*) at a high frequency. Accumulating evidence from pedigrees of MPN together with the identification of a plethora of heterogeneous lesions identified in sporadic MPN patients suggest that *JAK2V617F*, and other acquired changes in *JAK2*, cooperate with mutations in other genes to generate clonal disease. The nature of the other mutations dictates the disease phenotype and contributes to the potential for transformation to acute leukaemia. Emerging research is focussed on understanding the contribution of these other changes to MPN pathogenesis. As many of the other recurrent mutations reported in MPN affect genes involved in epigenetic regulation, studies have focused on identifying the role of epigenetic changes. Many epigenetic regulators mediate their effects via interaction with post-translationally modified histone H3 and H4 and, given the findings that pathogenic mutations are present in histone H3 in other tumours, the focus here was on the role of histone H3 and H4 variants in MPN pathogenesis.

Thus, a key aim of this PhD project was to identify pathogenic coding variants in the *histone H3* and *histone H4* genes in MPN. In the first study, MPN peripheral blood mononuclear cells or granulocyte patient samples were screened using Sanger sequencing for *histone H4* coding region variants. The screen identified previously unidentified sequence variants in several of the 15 *histone H4* genes. A coding variant of *HIST1H4C*, resulting in the substitution of cysteine for arginine [R, (*HIST1H4C:p.R4C*)], affects a known key residue involved in epigenetic regulation

(R3 residue on the mature protein). This gene was also shown to make a major contribution to the *histone H4* mRNA pool in several haemopoietic cell types, further indicating a potential for this variant to confer functional consequences. This was tested using enforced expression of the variant in two cell lines, HEK293 and a myeloid cell line (FDM cells). Further, it was demonstrated by RNA microarray and QPCR analyses of FDM cells expressing *HIST1H4C*:*p.R4C* that this variant conferred selective differential expression of five genes.

We extended this analysis of *Histone H4* genes to screen for disease-associated variants in *histone H3* genes (*n*=17) and the histone-H3 interacting protein *PHF6* (consisting of 9 coding exons). For this, we used an amplicon-based next generation sequencing (NGS) approach and the Roche 454 sequencing platform. This identified a coding region variant in *HIST1H3E* (*HIST1H3E:p.A96V*), the presence of which was confirmed by Sanger sequencing. A number of other changes identified by the NGS approach were not confirmed by Sanger re-sequencing, however the possibility that these are present in the original patient sample at a level below the detection limit for Sanger sequencing cannot be excluded. Sanger sequencing of the *PHF6* terminal exons 9 and 10 identified a somatic mutation (*PHF6R335fs*) in a PV patient.

Finally a Sanger-based sequencing screen of the terminal exon of gene encoding DNA (cytosine-5-)-methyltransferase 3 alpha (*DNMT3A*) identified somatic R882C and M880V substitutions in two PV patients. Clonal analysis for these mutations in *DNMT3A* indicated that their acquisition can either precede or follow the acquisition of *JAK2V617F*.

.

Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution. To the best of my knowledge and belief it contains no material previously published or written by another person, except where due reference has been made in the text.

In addition, I certify that no part of this work will, in the future, be used in a submission for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I give consent to a copy of my thesis, when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

I acknowledge that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library catalogue and also through web search engines, unless permission has been granted by the University to restrict access for a period of time. (*Please see errata sheet on the following page*)

Nisha Rao

November 2014

Errata Sheet

Due to lack of copyright and authorship certification, the following figures and sections from the thesis are not available publicly in the electronic version of this thesis:

Chapter 1

Figure 1.1, *Figure.1 2*, *Figure 1. 3*, *Figure 1. 4*, *Figure 1.5*, *Figure 1. 6*, *Figure 1. 7*, *Figure 1.8*, *Figure1. 9* and *Figure 10*

Chapter 2

Figure 2.5, *Section 2.3.5* and *Section 2.3.6*

Chapter 4

Figure 4.1

Chapter 5

Figure 5.1

Acknowledgements

I would like to thank the Queen Elizabeth Hospital and the University of Adelaide for their continual financial support during this PhD candidature. I would also like to thank the staff and students of the Queen Elizabeth Hospital Foundation, Royal Adelaide Hospital, and SA Pathology. I am grateful for the scientific communications that made my PhD candidature possible.

I would like to thank my supervisor; Prof. Richard D'Andrea, who continually supported and guided me throughout this PhD. I am truly grateful for all the support and encouragement that you have provided. I would also like to thank my co-supervisor; Ms Carolyn Butcher, whose many scientific discussions made this PhD candidature possible.

I am grateful for all of the friendships with several members of the Haematology department at the Queen Elizabeth Hospital, the Acute Leukaemia Laboratory at SA Pathology, and within the Basil Hetzel Institute for Translational Research. Thank you for all the conversations and after-hours socialising, which made this PhD candidature enjoyable.

I also wish to thank members at work especially Prof. Stephen Nicholls and my wonderful colleagues for being patient and understanding towards my study commitments. I know it would have been very difficult otherwise.

I wish to thank my parents and sister for motivating me continuously. I certainly would not have come this far without your constant encouragement and love.

To friends (Alpana, Jordan, Teresa, and Donna to name only a few), thank you for your emotional support and motivation throughout this PhD candidature. I know I could not have been encouraged more if it had not been for your constant words of support.

And lastly, but certainly not least, I wish to thank Shannon. You have inspired and motivated me to move forward, even in times of a complete lack of positivity or motivation. Thank you for all your love and support.

Abbreviations

2-HG	2-hydroxyglutarate
3´ UTR	3 prime untranslated region
4-OHT	4-Hydroxy tamoxifen
α-KG	α-ketoglutarate
Ac	Acetylated
AID	Activated Induced cytidine Deaminase
AML	Acute Myeloid Leukaemia
Allo-SCT	Allogeneic-Stem Cell Therapy
ALL	Acute lymphoblastic leukaemia
AKT	Protein kinase B
BCR-ABL1	Break point Cluster Region-Abelson murine leukaemia viral oncogenes, homolog 1
BER	Base Excision Repair
BFUE	Blast Forming Colonies Erythroid
BM	Bone Marrow
BMMNC	Bone marrow mononuclear cell
BTG2	BTG family, member 2
$CXCL_{12}$	Chemokine (C-X-C motif) ligand 12
CAR	$CXCL_{12}$-abundant reticular
CALR	Calreticulin
CEL	Chronic Eosinophilic Leukaemia
CMP	Committed Myeloid Progenitor
CLP	Committed Lymphoid Progenitor
Chr	Chromosome

CML	Chronic Myeloid Leukaemia
CMML	Chronic Myelomonocytic Leukaemia
CNL	Chronic Neutrophilic Leukaemia
C/EBPα	CCAAT/enhancer binding protein, alpha
CD	Cluster of Differentiation
CFU	Colony Forming Unit
COSMIC	Catalogue of Somatic Mutations in Cancer database
DNMT3A	*DNA (cytosine-5-)-methyltransferase 3 alpha*
Dock10	*Dedicator of cytokinesis 10*
e-BFUE	endogenous Blast Forming Colonies Erythroid (EPO independent colonies)
EEC	Endogenous erythroid colony
EPO	Erythropoietin
EED	Embryonic Ectoderm Development protein
EGR1	Early growth response protein-1
ERK	Extracellular signal-regulated kinase
ET	Essential thrombocythemia
EP	Erythroid Progenitor
ER-LBD	Estrogen Receptor-Ligand Binding Domain
EVS	Exome Variant Server
FACS	Fluorescence Activated Cell Sorting
FBS	Foetal Bovine Serum
FBN-III	Fibronectin type III
FDM	Foetal Derived Murine
FERM	4.1, ezrin, radixin, moesin homology domain

GAPDH	*Glyceraldehyde 3-phosphate dehydrogenase*
GATA1	GATA binding protein 1
GATA2	GATA binding protein 2
*g*DNA	Genomic DNA
GEV16	Gal4-Estrogen receptor-VP16 transactivation domain fusion protein
GAL4-DBD	GAL4-DNA binding domain
GFI-1	Growth factor independence-1
GFP	Green Fluorescent Protein
G-CSF	Granulocyte- Colony Stimulating factor
GM-CSF	Granulocyte Macrophage-Colony Stimulating factor
GMP	Granulocyte Monocyte Progenitor
H3K27me^3	Histone H3 lysine 27 trimethylation
H3Y41ph	Histone H3 tyrosine 41 phosphorylation
H4R3me^{2}a	Histone H4 arginine 3 asymmetric di-methylation
H4R3me^{2}s	Histone H4 arginine 3 symmetric di-methylation
HDE	Histone Downstream Element
HEK293	Human Embryonic Kidney 293
HEL	Human Erythro-Leukaemia 92.1.7
HGF	Haematopoietic growth factor
HIF	Hypoxia-inducible factor
HIST1	*Histone cluster 1*
HIST2	*Histone cluster 2*

HIST3	*Histone cluster 3*
HiNF-P	Histone nuclear factor–P
HIRA proteins	Histone cell cycle regulation defective homolog A (*Saccharomyces cerevisiae*)
HMR	Hidden Mat Right locus
HP-1α	Heterochromatin protein-1alpha
HSC	Haematopoietic Stem Cell
HSPC	Haematopoietic Stem/Progenitor cell
IDH	Isocitrate dehydrogenase
IGV2.1.2	Integrative Genomics Viewer version 2.1.2 software
IL	Interleukin
IL-3	Interleukin-3
IL-5	Interleukin-3
Il4ra	*Interleukin-4 receptor alpha*
INF-α	Interferon alpha
JAK2	Janus Kinase 2
JAK2V617F	Janus Kinase 2 Valine 617 Phenylalanine
JAK2V617F^{+}	*JAK2V617F*-positive
JH domain	JAK Homology domain
K-562	Human erythromyeloblastoid leukaemia cell
KIT	v-kit Hardy-Zucherman 4 feline sarcoma viral oncogene homolog
LIMMA	Linear Modelling for Microarray Analysis
LDH	Lactate dehydrogenase

LTHSC	Long-term repopulating haematopoietic stem cell
LMO2	*LIM (Lin11, Isl-1 & Mec-3) domain only 2 (rhombotin-like 1)*
LNK	SH2B adaptor protein 3
LIF-1	Leukaemia Inhibitory Factor-1
LTR	Long Terminating Repeats
µ	Micro (10^{-6})
µM	micro Molar
µg	microgram
µL	microlitre
M	Methionine
MACS	Magnetic Activated Cell Sorting
MAPK	Mitogen-activated protein kinase
MDS	Myelodysplastic syndrome
MDP	Macrophage/dendritic progenitor cell
MEP	Megakaryocyte Erythroid Progenitor
Me	Methylated
MEP50	Methylosome Protein 50
MkP	Megakaryocyte Progenitor
Mk	Megakaryocyte
MF	Myelofibrosis
MPL	Myeloproliferative leukaemia virus oncogene, also known as thrombopoietin receptor
MPN	Myeloproliferative neoplasm
MID	Multiplex Identification adaptors

MIG	MSCV-IRES-GFP plasmid/vector
miR	micro-RNA
Mir340	*MicroRNA 340*
mTOR	Mammalian target of rapamycin
NUSE Plot	Normalised Unscaled Standard Error plot
$NADP^+$	Nicotinamide Adenine Dinucleotide Phosphate
NADPH	reduced $NADP^+$
NGS	Next generation sequencing
Ndrg1	*n-Myc downstream regulated gene, 1*
PAX5	Paired box 5
PBMNC	Peripheral Blood Mononuclear Cell
PCR	Polymerase Chain Reaction
Ph	Philadelphia chromosome (*BCR-ABL* translocation)
ph	phosphorylated/phosphorylation
PI3K	Phosphoinositide-3-kinase
PHF6	Plant Homeodomain Finger, 6
PIAS	Protein Inhibitors of Activated STATs
Plekho2	*Pleckstrin homology domain containing family O member 2*
PMN	Polymorphonuclear
PMF	Primary myelofibrosis
PolyPhen-2	polymorphism phenotyping v2
PRC2	Polycomb Repressive Complex 2
PRMT	Peptidyl arginine methyltransferase
PRMT5	Peptidyl arginine methyltransferase, 5

PSG	Penicillin Streptomycin Glutamine
PTM	Post translational modification
PU.1	Spleen focus-forming virus (SFFV) proviral integration oncogene
PV	Polycythemia vera
QPCR	Quantitative PCR
p.R4C	Arginine 4 Cysteine, *HIST1H4C*:*c.10C>T, p.R4C*
p.R56Q	Arginine 56 Glutamine, *HIST1H4C*:*c.167C>T, p.R56Q*
RARA	Retinoic Acid Receptor, Alpha
RBAP48	Retinoblastoma binding protein 4 or RBBP4
RUNX1	Runt-related transcription factor 1
RBC	Red Blood Cell
S-phase	Synthesis-phase in DNA replication
S1	Serine 1
SAM	*S*-adenosylmethionine
SAH	*S*-adenosylhomocysteine
SCF	Stem Cell Factor
SCL	Stem Cell Leukaemia
SFFV	Spleen focus-forming virus
SNO	Spindle-shaped N-cadherin^{+}CD45^{-}Osteoblastic cell
SNP	Single nucleotide polymorphism
SLBP	Stem Loop Binding Protein
SIFT	Sorting-Intolerant-From-Tolerant

STHSC	Short-term repopulating haematopoietic stem cell
STAT	Signal Transducer and Activator of Transcription
SUZ12	Suppressor of zeste 12 homologue
T-ALL	T-cell acute lymphoblastic leukaemia
TDG	Thymine DNA glycosylase
TET2	Ten Eleven Translocation, 2
TF	Transcription factor
TPO	Thrombopoietin
TYK2	Tyrosine Kinase 2
U7snRNA	U7 small nuclear RNA
Ub	Ubiquitinated
UniProt	Universal Protein Resource database
U-MPN	MPN unclassifiable
UCSC genome browser	University of California Santa Cruz Genome Bioinformatics, Human (*Homo sapiens*) Genome Browser Gateway
5xUAS	5x *U*pstream *A*ctivation *S*equences
VP16-TD	VP16- Transcriptional activation Domain
WHO	World Health Organisation
WT	wildtype
W	Tryptophan (amino acid)
WSxWS	Tryptophan Serine, non-conserved residue, Tryptophan Serine

Chapter 1: Introduction

1.1 Haematopoiesis

The haematopoietic system serves as a paradigm for understanding tissue stem cell biology and oncogenesis. It is a highly regulated process and generates a number of mature cell types with specialised functions such as granulocytes, red blood cells (RBCs), and platelets. The constant replenishment of short-lived mature haematopoietic cells is maintained by a pool of haematopoietic stem cells (HSCs) which can undergo self-renewal and also give rise to progeny or precursor cells. These precursor cells in turn can undergo coordinated proliferation and differentiation along the various lineages giving rise to all types of haematopoietic cells. HSCs are located in the highly specialised and relatively hypoxic bone marrow (BM) microenvironment (BM niche) where they are maintained in a relatively quiescent state, retaining the characteristics of multipotency and self-renewal [(Parmar, *et al* 2007) and, reviewed in (Li 2011)].

Quiescent HSCs can be stimulated into cell cycle and the production of multipotent haematopoietic stem/progenitor cells (HSPCs) which have diminished self-renewing capacity. The HSPCs constitute a significant proportion of the adult BM and coordinate the production of large numbers of mature blood cells of diverse functional lineages [reviewed in (Weissman, *et al* 2001)]. Many models of haematopoiesis have been proposed which suggest an irreversible determination of fate between myeloid and lymphoid lineages (classical model shown in ***Figure 1.1***). However, new models suggest a greater level of plasticity of the intermediate progenitors wherein both myeloid and lymphoid lineages may be reconstituted [reviewed in (Ceredig, *et al* 2009; Rosenbauer, *et al* 2007)]..

Figure 1.1. **Overview of classical haematopoiesis.** The classical model of haematopoiesis *[see text for details, figure taken from Rosenbauer and Tenen, 2007 (Rosenbauer, et al 2007)]*. LTHSC, Long-term repopulating haematopoietic stem cell; STHSC, short-term repopulating haematopoietic stem cell; MPP, multipotent progenitor cell; CLP, common lymphoid progenitor cell; CMP, common myeloid progenitor cell; MEP, megakaryocyte-erythroid progenitor cell; GMP, granulocyte monocyte progenitor cell; MDP, macrophage/dendritic progenitor cell.

The regulation and determination of lineages is coordinated in large part by the BM niche, factors such as haematopoietic growth factors [HGFs; also known as haematopoietic cytokines] and transcription factors (TFs).

Given that the aims in this project are focussed on investigating the myeloid lineage, this literature review will be a deliberate focus on factors involved in myeloid haematopoiesis or myelopoiesis.

1.1.1 Regulation of myelopoiesis

A. The bone marrow (BM) niche

The BM niche is a highly specialised microenvironment which provides necessary regulatory signals essential for the maintenance, proliferation and differentiation of HSCs/HSPCs to produce the appropriate numbers of mature blood cells throughout life. These regulatory signals originate from a number of cell types which include osteoblasts, osteoclasts, nestin filament expressing cells, chemokine $CXCL_{12}$-abundant reticular (CAR) cells and stem cell factor (SCF) expressing cells, [(Sugiyama, *et al* 2012) and reviewed in (Lo Celso, *et al* 2011)]. In addition, other factors such as hypoxia, temperature and calcium ion gradients within the niche have been postulated to affect HSPC properties such as quiescence and self renewal [reviewed in (Lo Celso, *et al* 2011)] (***Figure 1.2***). The specialised cells in the BM niche not only regulate haematopoiesis through interaction with HSCs but also secrete HGFs which determine the fate of HSCs (Yoshihara, *et al* 2007; Zheng, *et al* 2011).

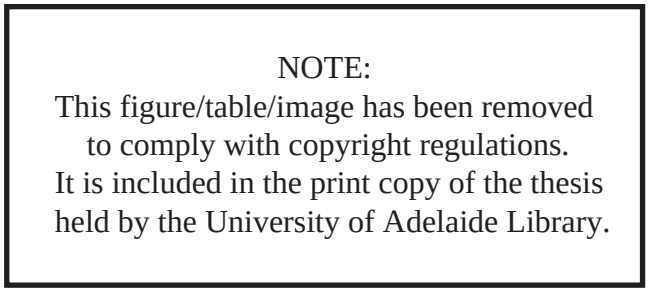
Figure 1.2. The bone marrow niche. The various cell types and factors which comprise BM niche are shown [*taken from Lo Celso, 2011 (Lo Celso, et al 2011)*].

B. Haematopoietic growth factors (HGFs) and their cognate receptors

HGFs constitute a large group of cytokines shown to play a critical role in regulating survival, proliferation, differentiation, commitment, and the end-stage function of the haematopoietic lineages. HGFs mediate regulatory signals by activating various signalling pathways and transcription factors. Important groups of HGFs in myelopoiesis include interleukin-3 (IL-3) [(Donahue, *et al* 1988) and reviewed in (Martinez-Moczygemba, *et al* 2003)], stem cell factor (SCF) (Pietsch, *et al* 1992), interleukin-5 (IL-5) (Campbell, *et al* 1987; Lopez, *et al* 1988), granulocyte macrophage colony stimulating factor (GM-CSF) [reviewed in (Martinez-Moczygemba, *et al* 2003)], thrombopoietin (TPO) (Kaushansky, *et al* 1994; Ng, *et al* 2012), erythropoietin (EPO) (Gasson, *et al* 1985; Miyake, *et al* 1977) and granulocyte colony-stimulating factor (G-CSF) [(Nicola, *et al* 1983) and reviewed in (Molineux 2011)] belong to the cytokine receptor class I super-family (***Figure 1.3*** shows examples of HGFs which determine cell lineage or type).

HGFs mediate regulatory affects by binding to specific receptors (HGF receptors) which transduce signals to the nucleus via activated downstream signalling molecules. The HGF receptors have been divided into several families, such as type I and type II cytokine receptors, based on their structure and activities. In particular, type I cytokine receptors, for example the receptors for EPO and G-CSF, are characterised by the presence of four conserved cysteine residues in the N-terminal fibronectin type III (FBN-III) repeat and a WSxWS motif in the C-terminal FBN-III repeat. In comparison, type II cytokine receptors, such as receptors for IL-10 and IL-20, are characterised by the presence of FBN-III domains, four conserved cysteines and lack of WSxWS motif

A.



B.

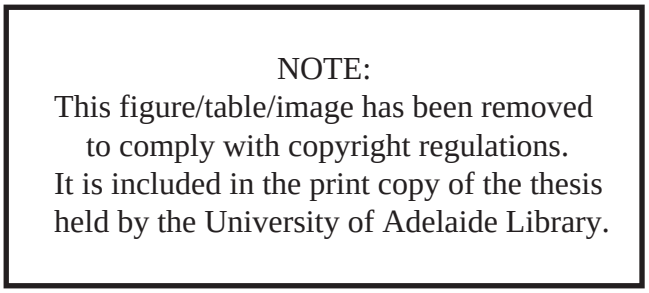


Figure 1.3. Haematopoietic growth factors in myelopoiesis. **A.** Simplified view of roles of HGFs in inducing differentiation at different cellular stages of myelopoiesis. CMP, common myeloid progenitor; MEP, megakaryocyte-erythroid progenitor; GMP, granulocyte monocyte progenitor; MkP, megakaryocyte progenitor; Mk, megakaryocyte; EP, erythroid progenitor. **B.** Summarises properties of HGFs.

[(Bazan 1990; Thoreau, *et al* 1991) and reviewed in (Gadina, *et al* 2001)]. Furthermore, signalling by HGF receptors can either utilise intrinsic kinases [for example c-Kit which is the receptor for SCF (Blume-Jensen, *et al* 1991)] or specific signal transduction molecules with kinase activity (for example in IL-3, IL-5 and EPO signalling). These downstream signalling pathway molecules include phosphoinositide-3-kinase/protein kinase B (PI3K/AKT) (Coffer, *et al* 1998; Sato, *et al* 1993; Tilton, *et al* 1997), mitogen-activated protein kinase/extracellular signal-regulated kinase (MAPK/ERK) (Coffer, *et al* 1998; Okuda, *et al* 1992) and importantly the Janus Kinase [JAK, (Quelle, *et al* 1994; Shuai, *et al* 1993; Silvennoinen, *et al* 1993)]. In the context of this thesis, the JAK-associated signal transduction pathways are probably the most relevant as this pathway is utilised by several HGFs to induce myeloid specific responses, [see ***sections 1.3*** and ***1.4*** of this chapter, reviewed in (Ihle 1995; Ihle, *et al* 1994)].

The ability of progenitors to undergo proliferation and differentiation in the presence of specific HGFs and signal transduction molecules is the basis for *in-vitro* colony forming assays, allowing investigation of HGF requirements in normal and disease haematopoietic progenitors (Metcalf, *et al* 1991). Along with HGF receptors, several other surface molecules such as glycoproteins, transmembrane proteins, and adhesion molecules also show lineage and/or stage-specific expression on myeloid cells. Together, these molecules known as cluster of differentiation (CD) markers have allowed for the extensive immuno-phenotyping of haematopoietic cells such as, CD34 (HSPCs), CD15 (granulocytes) and CD14 (monocytes) (Zola 2007).

Thus, pathways activated by HGFs and their receptors often act synergistically and induce overlapping functions forming a highly complex level of regulation which is

not only mediated by signal transduction molecules (as mentioned previously) but also transcription factors (TFs).

C. Transcription factors in myelopoiesis

There are a number of myeloid lineage-specific transcription factors (TFs) which direct the fate of the HSC/HSPC to a committed myeloid progenitor. Single TF or interplay of a number of TFs and their level of expression modulates the proliferation and differentiation of various haematopoietic cell stages. For example, master regulators of myeloid differentiation at the HSC/HSPC level include runt-related transcription factor 1 (RUNX1) and stem cell leukaemia factor (SCL). It has been reported that ablation of *RUNX1* has a myeloproliferative effect, thus suggesting that RUNX1 acts as a myeloid lineage suppressor (Growney, *et al* 2005). In contrast, SCL has been shown to have a critical role in determining the erythroid and megakaryocytic lineages (Mikkola, *et al* 2003).

In addition, there are other TFs which also determine myeloid fates, for example, the ETS transcription family member, spleen focus forming virus (SFFV) proviral integration 1 (PU.1), CCAAT/enhancer binding protein alpha (C/EBPα), growth factor independence-1 (GFI-1), and members of the Signal Transducers and Activators of Transcription (STAT) family.

PU.1 has been shown to be critical for myelomonocytic differentiation. High expression directs commitment of GMPs to neutrophils and monocytes whereas PU.1 down-regulation in MEPs is essential for the development of megakaryocytes and RBCs (Back, *et al* 2004; Laslo, *et al* 2006; Nutt, *et al* 2005; Spooner, *et al* 2009). Ablation of PU.1 has been shown to reduce HSC numbers, CMPs, CLPs, and GMPs

in *Pu.1* knockout models (Iwasaki, *et al* 2005; Rosenbauer, *et al* 2004; Scott, *et al* 1994).

C/EBPα is also a key regulator of granulocyte/macrophage differentiation. Ablation of C/EBPα leads to the absence of neutrophils, eosinophils and reduced monocytes due to a block in differentiation at the CMP and GMP stage (Heath, *et al* 2004; Zhang, *et al* 1997). The phosphorylation of specific serine residues on C/EBPα (at positions 21 and 248, S^{21} and S^{248} respectively) has been reported to have an impact on either suppression or induction of granulocyte/macrophage differentiation (Behre, *et al* 2002; Ross, *et al* 2004).

GFI-1, another key regulator of haematopoiesis, has been reported to antagonise the effects of PU.1 expression thereby enhancing differentiation of granulocytes at the cost of macrophages (Dahl, *et al* 2007; Spooner, *et al* 2009). In HSCs, expression of GFI-1 has also been associated with self-renewal and maintenance of the HSC numbers (Zeng, *et al* 2004).

Members of the STAT family such as STAT1, STAT3 and STAT5 play important roles in myelopoiesis. For example STAT5 is essential for myeloid progenitor differentiation in the production of monocytes, granulocytes, erythrocytes and platelets [reviewed in (Dorritie, *et al* 2013)] and will be further described in ***section 1.4***.

Overall, the complex regulatory signals include levels of redundancy with regard to activated molecules and the related downstream functional effects. Importantly, abnormal expression or function of HGFs, HGF receptors and TFs can lead to

dysregulated haematopoiesis resulting in disorders such as myeloproliferative neoplasms or myeloid leukaemia [reviewed in (Jatiani, *et al* 2010; Tenen, *et al* 1997)].

1.2 Malignant myeloid disorders

Malignant myeloid disorders or myeloid neoplasms are a broad group of clonal diseases that may affect a single myeloid cell lineage [for example, mast cell disease or acute myeloid leukaemia (AML)] or, multiple myeloid lineages (for example acute mixed lineage leukaemia and the myeloproliferative neoplasm, polycythemia vera). The clinical course of such disorders may be acute (short and aggressive) or chronic. Acute and chronic disorders are genetically heterogeneous clonal disorders characterised by the accumulation of somatic genetic alterations in haematopoietic progenitor cells [reviewed in (Dohner, *et al* 2008)]. However, acute disorders are characterised by a block in differentiation in progenitors, which leads to an excess of un-differentiated cells (also known as blasts). On comparison, progenitors in chronic disorders *may not* have a block in differentiation, thereby leading to an excess of differentiated cells.

The acute and chronic myeloid neoplasms have been further classified into several categories including AML, myeloproliferative neoplasms (MPN), myelodysplastic syndrome (MDS), MDS/MPN overlap disorders [see ***Figure 1.4A*** (Vardiman, *et al* 2009)]. Since the focus here is MPN, only the sub-classifications of MPN have been described.

MPN is broadly classified depending on the presence or absence of the *BCR-ABL1* translocation (Vardiman, *et al* 2009). The *BCR-ABL1* fusion oncogene generated due to the Philadelphia chromosome (*Ph*) translocation, t(9; 22)(q34; q11), generates a

fusion oncoprotein, BCR-ABL1 protein kinase, which has constitutive kinase activity. CML is defined by the presence of this translocation and is known as the *BCR-ABL1* positive or Philadelphia chromosome positive (Ph^+) CML [(de Klein, *et al* 1982; Rowley 1973) and reviewed in (Jabbour, *et al* 2012)].

Ph-negative MPN, on the other-hand, are the focus of this work and include polycythemia vera (PV), essential thrombocythemia (ET), myelofibrosis [MF, which includes primary myelofibrosis (PMF) and post-ET- or post-PV-MF], chronic neutrophilic leukaemia (CNL), chronic eosinophilic leukaemia, not otherwise specified (CEL-NOS), myeloproliferative neoplasms unclassifiable (U-MPN) and mastocytosis (***Figure 1.4A***).

1.3 *Philadelphia chromosome*-negative MPN: PV, ET and MF

PV, ET and MF (including PMF and post-PV- or post-ET-MF), referred to as MPN henceforth, are clonal disorders that arise in the HSPC, resulting in excessive production of mature cells of the myeloid lineage. This feature is typically attributed to hypersensitivity to or independence from normal regulation by HGFs such as EPO or TPO, which utilise non-receptor kinases for their activity (Adamson, *et al* 1976; Asimakopoulos, *et al* 1996; El Kassar, *et al* 1995; Gilliland, *et al* 1991; Tsukamoto, *et al* 1994). This is consistent with the finding of highly penetrant mutations in the genes encoding the receptor for TPO [myeloproliferative leukaemia virus oncogene (MPL)], and the kinase involved in the signal transduction via these receptors, Janus Kinase 2 [JAK2, (see ***sections 1.3.4*** and ***1.4***)]. The overall quality of life and survival in MPN patients is affected by a spectrum of complications (such as splenomegaly, thrombosis, haemorrhage, microvascular symptoms, pruritis, hepatomegaly, anaemia, and severe constitutional symptoms) with the more severe symptoms

Figure 1.4A. Classification of myeloid neoplasms and MPN. *Adapted from (Tefferi, et al 2014c; Vardiman, et al 2009).* Frequency for *CALR* mutations has been taken from *(Klampfl, et al 2013; Nangalia, et al 2013).* PV, polycythemia vera; ET, essential thrombocythemia; PMF, primary myelofibrosis; *MPL, myeloproliferative leukaemia virus oncogene*; *JAK2, Janus Kinase 2*; *JAK2V617F, JAK2 valine-617-phenylalanine* (denotes the mutant allele); *CALR, calreticulin.*

typically associated with MF (Mesa, *et al* 2007a; Mesa, *et al* 2009; Scherber, *et al* 2011). Furthermore, there is significant risk of transformation to leukaemia [2-5% risk in PV and ET (Finazzi, *et al* 2005; Wolanskyj, *et al* 2006), 15-30% in PMF (Barosi, *et al* 2007; Dupriez, *et al* 1996)]. Although leukaemic progression in PV, ET and PMF has been shown to occur in patients without any prior therapy (Finazzi, *et al* 2005), transformation to leukaemia is increased following treatment with cytoreductive therapies (Merlat, *et al* 1999; Sterkers, *et al* 1998). In addition, the presence of heritable genotypes (pre-*JAK2* event) may pre-dispose HSPCs to acquire clinically relevant mutations in multiple clones (Beer, *et al* 2009; Kralovics, *et al* 2006; Landgren, *et al* 2008; Nussenzveig, *et al* 2007). This possibility is consistent with the observations of the 46/1 haplotype that is prevalent in MPN (discussed in ***section 1.7***) and also with the occurrence of bi-clonal disease (for example, co-existence of mutations in *JAK2* and *MPL* in a single patient) reported in some cases (Hussein, *et al* 2009; Lasho, *et al* 2006; Li, *et al* 2008). Thus, the underlying mechanisms driving MPN transformation are likely to involve the acquisition of a plethora of mutations in genes which collaborate in oncogenesis (Abdel-Wahab, *et al* 2010; Zhang, *et al* 2012).

Although the diseases can be managed and symptoms alleviated with current treatments, therapies are not curative. Current diagnostic and management therapies for PV, ET and PMF are outlined in the following sections.

1.3.1 PV

PV is diagnosed using the 2008 World Health Organisation (WHO) criteria, either at the time of a routine examination which may reveal an elevated haematocrit, or due to the presence of disease symptoms [see ***Figure 1.4B*** for 2008 WHO diagnostic criteria (Passamonti 2012; Tefferi, *et al* 2008; Vardiman, *et al* 2009)]. Diagnosis for

PV typically also reveals increased leukocyte counts, increased platelet counts and, an enlarged spleen (Campbell, *et al* 2005; Tefferi 2008; Tefferi, *et al* 2008). The 10 year median survival for PV is >75% and prolonged survival is maintained if elevated haematocrit, leukocytosis, and platelets can be managed (Crisa, *et al* 2010). Interestingly, pruritis, which can also be a symptom of PV, has been reported as a favourable risk factor for survival in PV patients (Tefferi, *et al* 2013). Whether this association indicates bias in lead-time for diagnosis in patients with pruritis or relates to underlying biology is yet to be understood.

Risk factors associated with increased mortality in PV patients include age-at-diagnosis, thrombotic events, leukocytosis and genetic abnormalities (for example Janus Kinase *JAK2V617F*) (Bonicelli, *et al* 2013; Gangat, *et al* 2007a; Passamonti, *et al* 2010b; Passamonti, *et al* 2004). Depending on these risk factors, the classification of diagnosis in risk groups has been proposed to aid in predicting occurrence of complications, especially risk-of-thrombosis [***Figure 1.4C***, (Tefferi 2013a) and reviewed in (Finazzi, *et al* 2008)]. The primary management for PV includes regular venesections but this regimen can alter depending on the risk group. The return of haematocrit to normal values solely by venesections has been reported to lead to a reduction in the number of thrombotic events thereby increasing the overall median survival of patients by approximately 13.9 years (Berk, *et al* 1986). However, the appropriate haematocrit level required to reduce thrombotic events (either <45% or <50%) in patients has been in debate (Barbui, *et al* 2011a; Di Nisio, *et al* 2007; Fruchtman, *et al* 1997; Spivak 2002; Streiff, *et al* 2002). A recent randomised study by Marchioli and colleagues has supported current practice of maintaining a low haematocrit level (<45%) by reporting a significant reduction of cardiovascular events due to thrombosis in PV patients (Marchioli, *et al* 2013). In addition, agents

such as hydroxyurea, interferon-alpha (INF-α), busulfan and low dose of aspirin [100mg/day (Passamonti 2012)] are also included in the management regimen (see ***Figure 1.4C***). Most recently, the use of Ruxolitinib, an oral inhibitor of tyrosine kinases including JAK2, has also shown promise in phase II clinical trials by improvement of severe symptoms such as enlarged spleen and thrombosis [(Passamonti 2012; Verstovsek, *et al* 2012) and other emerging drugs have been reviewed in (Tibes, *et al* 2013)].

1.3.2 ET

ET is also diagnosed using the 2008 WHO diagnostic criteria, and characterisation includes assessing for sustained proliferation of megakaryocytes leading to an increased number of circulating platelets [see ***Figure 1.4B*** for 2008 WHO diagnostic criteria (Tefferi, et al 2008)]. Clinical course typically presents with thrombotic episodes, profound BM megakaryocyte hyperplasia and splenomegaly (Tefferi 2013a; Tefferi, *et al* 2008; Vardiman, *et al* 2009).

ET was considered a relatively indolent condition with a favourable prognosis (compared to PV and other myeloid malignancies) and with complications arising largely due to thrombosis and/or age (Barbui, *et al* 2011b; Cervantes, *et al* 1991; Mesa, *et al* 1999; Passamonti, *et al* 2004; Rozman, *et al* 1991). In addition to the risk of thrombosis, age is also a predictor for the occurrence of transformation. A report by Wolanskyj and colleagues reported a cumulative probability of 9.8% at 10 years and 28.3% at 20 years, for transformation of ET to any myeloid neoplasm [which included AML or PV, (Wolanskyj, *et al* 2006)]. Thus, the two variables have formed the basis for classifying ET into risk groups [see ***Figure 1.4C*** (Barbui, *et al* 2004; Elliott, *et al* 2005; Gangat, *et al* 2007b; Passamonti, *et al* 2008)].

	PV	ET	PMF
Major Criteria	Haemoglobin of >185g/L (men) and >165g/L (women)	Platelet count ≥450× 10^9/L	Megakaryocyte proliferation and atypia accompanied by either reticulin and/or collagen fibrosis,
	Or Haematocrit >99th percentile of reference range for age, sex or altitude of residence	Megakaryocyte proliferation with large and mature morphology. No or little granulocyte or erythroid proliferation.	*Or* In the absence of reticulin fibrosis, the megakaryocyte changes must be accompanied by increased marrow cellularity, granulocytic proliferation and often decreased erythropoiesis
	Or Haemoglobin >170g/L (men), or >150g/L (women) if associated with a sustained increase of ≥20g/L from baseline that cannot be attributed to correction of iron deficiency	Not meeting WHO criteria for CML, PV, PMF, MDS or other myeloid neoplasm	
	Or Elevated red cell mass >25% above mean normal predicted value	Demonstration of *JAK2V617F* *Or* Other clonal marker	
	And Presence of *JAK2V617F* or similar mutation	*Or* No evidence of reactive thrombocytosis	
Minor Criteria	BM trilineage myeloproliferation	None	Leukoerythroblastosis, Palpable spleenomegaly, Anaemia, Increased serum LDH
	Sub-normal serum EPO level		
	EEC growth		
Requirements for diagnosis	Fulfilling both major criteria and one minor criterion *Or* The first major criterion and two minor criteria	Fulfilling all four major criteria	Fulfilling all major criteria and minor criteria

Figure 1.4B. 2008 WHO diagnostic criteria for PV, ET and PMF. *Adapted from (Tefferi, et al 2008).* It should be noted that with the recent identification of mutations in *CALR* in *JAK2V617F*-negative and *MPL*-negative ET and PMF patients, additional diagnostic criteria may be included in the future (Tefferi, *et al* 2014b). BM, Bone marrow; EEC, endogenous erythroid colony; LDH, Lactate dehydrogenase; EPO, Erythropoietin, WHO, World Health Organisation; CML, Chronic myelogenous leukaemia; MDS, Myelodysplastic syndrome.

Risk Categories	PV	ET
Low-risk without extreme thrombocytosis (age <60 years and no thrombosis history)	Low-dose aspirin + phlebotomy	Low-dose aspirin
Low-risk with extreme thrombocytosis (platelets >1,000 × 10^9/L)	Low-dose aspirin provided ristocetin cofactor activity >30% + phlebotomy	Low-dose aspirin provided ristocetin cofactor activity >30%
High-risk (age ≥60 years and/or presence of thrombosis history)	Low-dose aspirin + phlebotomy + hydroxyurea	Low-dose aspirin + hydroxyurea
High-risk disease that is refractory or intolerant to hydroxyurea	Low-dose aspirin + phlebotomy + interferon-alpha (age <65 years) or busulfan (age ≥65 years)	Low-dose aspirin + interferon-alpha (age <65 years) or busulfan (age ≥65 years)

Figure 1.4C. Risk associated therapy in PV and ET. *Adapted from (Tefferi 2013a).*

High-risk ET patients, who have a history of thrombosis and are ≥60 years, receive cytoreductive treatments such as hydroxyurea, INF-α and busulfan (Harrison, *et al* 2005; Tefferi 2013a). In addition, anagrelide, an imidazoquinazoline derivative, may also be used to reduce platelet count instead of hydroxyurea. However, the use of anagrelide is currently a matter of debate as it has been associated with transformation in high-risk ET patients (Gisslinger, *et al* 2013; Harrison, *et al* 2005; Tefferi 2013a).

1.3.3 MF: PMF and post-PV/ET-MF

MF can either develop *de novo* (PMF) or be a severe complication of PV and ET [post-PV MF/post-ET MF, (Mesa, *et al* 2007b)]. The current protocol for diagnosis of PMF is based on 2008 WHO criteria for comprehensive assessment of clinical and laboratory features [see ***Figure1.4B*** (Tefferi, *et al* 2008; Vardiman, *et al* 2009)], whereas, the diagnosis of post-PV- or post-ET-MF is based on the criteria proposed by the International Working Group for MPN Research and Treatment (Barosi, *et al* 2008). With respect to this thesis, the following sections will be focussed on symptoms, classification and management of PMF.

A central feature of PMF is the presence of increased clonal myeloproliferation which is characterised by megakaryocyte proliferation and atypia, reticulin and/or collagen BM fibrosis, and presence of mutations (for example in *JAK2* or *MPL*). Additional important features such as decreased erythropoiesis (resulting in severe anaemia), hepato-splenomegaly, and symptoms not meeting the 2008 WHO criteria for PV, CML, MDS or any other myeloid neoplasm, aid in the diagnosis of PMF (Tefferi, *et al* 2007; Vardiman, *et al* 2009). It is believed that abnormal cytokine-production by clonal cells may lead to BM fibrosis, extra-medullary haematopoiesis, hepatomegaly and other PMF-associated symptoms (Tefferi 2005; Tefferi 2013b).

PMF was initially classified into four risk groups depending on the presence of one or more of five independent predictors of survival which were age (>65years), haemoglobin (<100g/L), leukocyte count (>25×10^9/L), ≥1% circulating blasts, and the presence of constitutional symptoms (Cervantes, *et al* 2009). The presence of none, 1, 2 and ≥3 adverse factors defined low, intermediate-1, intermediate-2 and high-risk disease with corresponding median survivals of approximately 11.3, 7.9, 4.0 and 2.3 years respectively (Cervantes, *et al* 2009). Interestingly, the actual survival for these risk groups ranged from approximately ≥20years to 8years for low to high risk groups respectively. Further to the previous risk classification, the identification and inclusion of additional risk factors is the basis of the current prognostic model which is known as the modified Dynamic International Prognostic Scoring System-plus or DIPSS-plus. The DIPSS-plus suggests the use of additional predictors such as unfavourable karyotype, platelet count of <100 × 10^9/L and the need for RBC transfusion (Caramazza, *et al* 2011; Elena, *et al* 2011a; Gangat, *et al* 2011; Hussein, *et al* 2010; Passamonti, *et al* 2010a; Patnaik, *et al* 2010a; Tefferi, *et al* 2010b). ***Figure 1.4D*** depicts the PMF risk groups and the associated median survivals.

Irrespective of the prognostic model applied, mortality in the high risk groups has been suggested to be associated with a high risk of transformation to acute leukaemia (approximately 20% of the cases), progressive cachexia and acute vascular events (Barosi, *et al* 2007; Mesa, *et al* 2005). Current disease management is largely palliative and includes the use of erythropoiesis-stimulating agents, prednisone, thalidomide, lenalidomide, and hydroxyurea (Tefferi 2013b). The combinations of therapy applied may change, depending on the PMF risk group, and patients in the

Risk Categories of PMF	Median survival	Management
Low-risk (no risk factors)	~15.4 years	Observation and Conventional drugs
Intermediate-1 risk (1 risk factor)	~6.5 years	Observation and Conventional drugs; or Experimental drugs
Intermediate-2 risk (2-3 risk factors)	~2.9 years	Allo-SCT and Experimental drugs
High-risk (≥4 risk factors)	~1.3 years	Allo-SCT and Experimental drug

Figure 1.4D. Currently proposed risk-associated therapy in PMF. Conventional drugs include androgen preparations or thalidomide with prednisone for anemia and, hydroxyurea for symptomatic splenomegaly. Current experimental drugs include pomalidomide, JAK2 inhibitor ATP mimetics (for example Ruxolitinib and CYT387) and mammalian target of rapamycin (mTOR) inhibitors. Allo-SCT, allogenic stem cell transplant. *Adapted from (Tefferi 2013b).*

low risk groups may be managed with thalidomide and/or hydroxyurea (conventional drugs). However, patients in the high risk group may be managed with additional experimental drugs such as pomalidomide, JAK2 inhibitor ATP mimetics, or mammalian target of rapamycin (mTOR) inhibitors along with allogeneic stem cell transplantation (Tefferi 2013b).

1.3.4 Identification of Janus Kinase 2 (JAK2) mutations in MPN

Although PV, ET and PMF have unique clinical courses, they also share several similar patho-physiological features including evidence of clonal stem cell expansion, altered HSPC growth properties, and overall disease progression to MF and AML. These similar disease features have provided the basis in searching for common genetic lesions associated with these three MPN.

In 2005, the observations of several investigators converged with the identification of an acquired mutation in *JAK2* (Baxter, *et al* 2005; James, *et al* 2005; Kralovics, *et al* 2005; Levine, *et al* 2005; Steensma, *et al* 2005; Zhao, *et al* 2005). A single *G* to *T* nucleotide change at position 1849 in exon 14 of *JAK2* resulting in the substitution of valine-to-phenylalanine at amino acid position 617 in JAK2 (JAK2V617F), was observed in the majority of PV (95%), ET (50–60%) and PMF (50–60%) patients. Mutations in exon 12 of *JAK2* are observed with lesser frequency (approximately 5%) and are restricted to patients with PV (Butcher, *et al* 2008; Pardanani, *et al* 2007a; Scott, *et al* 2007b). The important discovery of *JAK2* mutations in MPN has led to the development of several JAK2 inhibitors which are either approved or in clinical trials [(Pardanani, *et al* 2009; Santos, *et al* 2010; Verstovsek, *et al* 2010) and reviewed in (Sonbol, *et al* 2013)].

1.4 Janus Kinase (JAK) family proteins

JAK2, along with JAK1, JAK3 and TYK2, belongs to the JAK family of proteins which are cytoplasmic non-receptor tyrosine kinases. These proteins are critical in cytokine receptor signalling pathways [reviewed in (O'Shea, *et al* 2002)]. Upon HGF binding to a receptor, a conformational change in the receptor brings two JAK2 proteins into proximity resulting in trans-phosphorylation. Phosphorylated JAK2 acts as an activated tyrosine kinase, phosphorylating the cytoplasmic domain of the type I cytokine receptors which become the docking site of STAT and other proteins. STAT and other signal-transducing proteins are, in turn, also phosphorylated by activated JAK2 when bound to phosphorylated cytoplasmic domains of the cytokine receptors. The activated signal-transducing proteins may form dimers (for example STAT5 dimers) and enter the nucleus, where they act as transcription factors to regulate target myeloid genes [reviewed in (Ihle, *et al* 2007)]. Thus, JAK2 plays a critical role in myeloid-specific cytokine signalling and has been shown to be the predominant JAK activated in response to type-I cytokine ligands. These include EPO, GM-CSF and TPO signalling (Drachman, *et al* 1995; Drachman, *et al* 1997; Watanabe, *et al* 1996; Witthuhn, *et al* 1993).

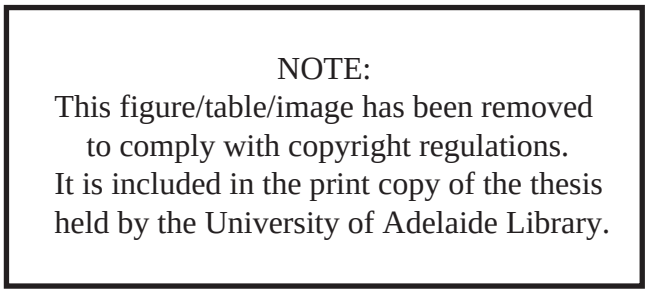
1.4.1 JAK2 structure and function

Human *JAK2* is located on chromosome 9p24 and has 25 exons (www.ensembl.org) encoding a 1132 amino acid protein. The JAK2 protein has several important domains known as the JAK homology (JH) domains [***Figure 1.5.A***, reviewed in (Smith, *et al* 2008)]. The JH1 domain, at the carboxy-terminus (C-terminus) of the protein, is the active tyrosine kinase domain where the phosphorylation of tyrosine (Y) residues (Y1007 and Y1008) is believed to activate JAK2 protein kinase activity (Feng, *et al* 1997). The JH2 domain is a pseudokinase which functions as a negative regulatory domain exerting inhibitory effects on JH1 (Dusa, *et al* 2010; Saharinen, *et*

al 2002; Saharinen, *et al* 2003). JH3 and JH4 are homologous to the Src homology 2 motif (SH2) and may play a role in the oncogenic activation of JAK2V617F (Giordanetto, *et al* 2002; Gorantla, *et al* 2010). JH5, JH6 and JH7 are known to contain the four point one, ezrin, radixin, moesin homology or FERM domain which plays a role in the non-covalent binding to the juxtamembrane cytoplasmic region of the type I cytokine receptors (Funakoshi-Tago, *et al* 2006; Funakoshi-Tago, *et al* 2008).

Given that the kinase activity of JAK2 is a crucial step in signalling for numerous HGFs (***Figure 1.5.B***) such as EPO, TPO, IL3 and IL5 [(Parganas, *et al* 1998; Silvennoinen, *et al* 1993; Witthuhn, *et al* 1993) and reviewed in (Aaronson, *et al* 2002)], regulation of JAK2 activity is performed at various levels and is an important step in maintaining normal cellular processes (Funakoshi-Tago, *et al* 2008; Saharinen, *et al* 2002). These regulatory mechanisms may include auto-inhibition by the JH2 domain of JAK2 (mentioned previously), the activation of members of the suppressors-of-cytokine-signalling (SOCS) proteins which suppress signalling by feed-back mechanisms [(Flowers, *et al* 2004; Waiboci, *et al* 2007) and reviewed in (Croker, *et al* 2008)], or regulation by tyrosine phosphatases such as protein-tyrosine phosphatase 1B (Myers, *et al* 2001). It has been suggested that intrinsic regulation of JAK2 activity depends on the phosphorylation of specific tyrosine residues such as Y221, Y570 and Y813 (Feener, *et al* 2004; Kurzer, *et al* 2004). For example, Y570 phosphorylation (which is within JH2) has been reported to have a negative regulatory effect on the kinase activity where a mutation of this amino acid (Y570F) leads to constitutively active JAK2-dependent signalling in the absence of EPO (Feener, *et al* 2004). Additional mechanisms may include the indirect regulation of

A.



B.

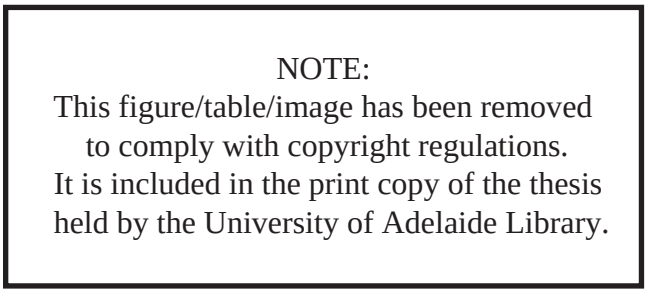


Figure 1.5. The structure of JAK2 and the spectrum of cytokines (including HGF) involved in JAK signalling. **A**. The molecular structure of JAK2 [*adapted from Smith and Fan, 2008 (Smith, et al 2008)*] showing JAK Homology (JH) domains; JH1, which has tyrosine kinase activity; JH2, has a regulatory effect on JH1. **B**. The role of JAK2 in the regulation of a number of HGFs [*adapted from Tefferi, 2010 (Tefferi 2010)*] such as EPO, Erythropoietin; TPO, Thrombopoietin and Interleukins (IL) such as IL-3, IL-4, IL5, IL-7, IL-9, IL-12, IL-21, IL-23 etc.; GM-CSF, Granulocyte macrophage colony stimulating factor; G-CSF, Granulocyte colony stimulating factor.

the STAT proteins by tyrosine phosphatases (ten Hoeve, *et al* 2002) and protein inhibitors of activated STATs (PIAS) (Chung, *et al* 1997; Liu, *et al* 1998).

1.4.2 JAK2 as a chromatin modifier

In addition to the cytoplasmic roles in HGF signalling, JAK2 has been shown to influence chromatin structure (Dawson, *et al* 2009; Shi, *et al* 2006; Shi, *et al* 2008). For example, in haematopoietic cells, JAK2 has been shown to translocate to the nucleus and displace heterochromatin protein-1alpha (HP-1α) by phosphorylating histone H3 at a tyrosine residue (Y41). This has been reported to up-regulate the expression of several downstream genes, including the *LIM (Lin11, Isl-1 & Mec-3) domain only 2* or *LMO2* oncogene (Dawson, *et al* 2009). Further, nuclear JAK2 was reported to be phosphorylated and the state of phosphorylation was dependent on the presence of an appropriate cytokine/growth factor such as leukaemia inhibitory factor (LIF-1). On addition of JAK2 inhibitors such as TG101209 and AT9283, a decrease in the levels of histone H3Y41 phosphorylation (H3Y41ph) strongly supported the chromatin modifying role of JAK2 (Dawson, *et al* 2009).

In 2010, Rinaldi and colleagues reported that MPN patients with *JAK2V617F* mutation showed a preferential accumulation of JAK2 in the nucleus of $CD34^+$ cells, as compared to MPN patients with wildtype JAK2 (Rinaldi, *et al* 2010). This finding was suggestive of a nuclear role for JAK2. Additionally, JAK2V617F was reported to phosphorylate protein arginine methyltransferase 5 (PRMT5) with greater affinity compared to wildtype JAK2 (Liu, *et al* 2011). This specific feature of the mutant JAK2 protein has been shown to disrupt the interaction of PRMT5 with its cofactor methylosome protein 50 (MEP50) leading to decreased methyltransferase activity. As knockdown of PRMT5 was also shown to increase the number of erythroid colony forming cells and erythroid differentiation of human $CD34^+$ cells,

JAK2V617F induced phosphorylation of PRMT5 may be important in the pathogenesis of MPN (***Figure 1.6*** shows the various roles of JAK2 in cytokine signalling).

1.4.3 JAK2 mutations in MPN

The *JAK2V617F* mutation has been shown to be acquired, clonal (James, *et al* 2005), present in the HSC population (Jamieson, *et al* 2006), and also confers EPO-independent growth of erythroid progenitors (Baxter, *et al* 2005). The highest frequency of this mutation is observed in PV (approximately 95%), while for ET and PMF this mutation is found at a frequency of approximately 50-60%. In addition, the *JAK2V617F* mutation has been found at lower frequencies in other myeloid neoplasms such as MDS (<5%) (Szpurka, *et al* 2006), juvenile myelomonocytic leukaemia (<20%) (Tono, *et al* 2005), and chronic myelomonocytic leukaemia (<5%) (Steensma, *et al* 2005), which indicates that acquisition of *JAK2V617F* may be heterogeneous. It may also be suggestive of distinct disease-initiating events with a potential common manner of acquisition of mutations. In the rare development of a lymphoid proliferative neoplasm from an MPN, *JAK2V617F* has occasionally been detected in lymphoid cells (Larsen, *et al* 2007; Stijnis, *et al* 2012).

In the cases of rare *JAK2V617F*-negative PV, somatic gain-of-function mutations in exon 12 of *JAK2* have been identified [(Butcher, *et al* 2008; Pardanani, *et al* 2007a; Pietra, *et al* 2008; Scott, *et al* 2007a; Scott, *et al* 2007b) and reviewed in (Scott 2011)]. These mutations occur in virtually all of *JAK2V617F*-negative PV cases (which constitute 3-4% of all PV cases), span the linker region between SH2 and JH2 domain of the protein, and may affect the interaction of the pseudokinase with the kinase domain. This results in the constitutive activation of JAK2 leading to a phenotype characterised by erythrocytosis (Pietra, *et al* 2008; Scott, *et al* 2007b).

Figure 1.6 JAK2 is central in the pathogenesis of *Ph*-negative MPN. JAK/STAT and other JAK2 signalling pathways are activated in most patients with the classic *Ph*-negative MPN. This process can be due to direct activation of JAK2 as a result of *JAK2* exon 12 and/or *JAK2V617F* mutations (represented in red). JAK2 can also be activated indirectly by mutations in MPL such as MPLW515L/K; mutations in LNK (SH2B adaptor protein 3). A number of other pathways are also activated and enhanced due to deregulated JAK2 such as PI3K/Akt/mTOR and MAPK pathways. Nuclear JAK2 leads to epigenetic deregulation by phosphorylating Histone H3 at tyrosine 41 which blocks the binding of epigenetic regulators such as heterochromatin protein-1alpha. Wildtype JAK2 is represented in orange. mTOR, mammalian target of rapamycin; Ras, rat sarcoma viral oncogenes homolog; RAF, murine leukaemia viral oncogenes homolog1. [*Schematic adapted from Santos and Verstovsek (Santos, et al 2012)]*

In addition, other mutations in exon 14 of *JAK2* and deletion of exon 14 in the *JAK2* transcript have also been reported (Ma, *et al* 2010; Warshawsky, *et al* 2010). Rare mutations in exon 14 of *JAK2* have been reported to co-exist with *JAK2V617F* and include substitutions of cysteine at position 618 to either an arginine (*JAK2:*c.*T1852C,* p.C618R) or phenylalanine (*JAK2:*c.*G1853T,* p.C618F) (Warshawsky, *et al* 2010). Each of these mutations has been reported in a single PV and ET case.

1.5. Mutations highly specific to ET and PMF

Subsequent to the discovery of mutations in *JAK2,* mutations in other genes have also been reported in MPN.

1.5.1 MPL

Somatic mutations in *myeloproliferative leukaemia virus oncogene* or *MPL* were initially described in 2006 and are virtually exclusive to ET and PMF patients (Pikman, *et al* 2006; Schnittger, *et al* 2007). *MPL* is located on chromosome 1p34, includes 12 exons and encodes for the TPO receptor (635–680 amino acids). Mutations in *MPL* mostly affect tryptophan (W) at position 515 (W515L/S/N due to a G to T/C/A transition at nucleotide 1544) and, occur in approximately 3% to 15% of ET and PMF patients (Beer, *et al* 2008; Pardanani, *et al* 2006a; Schnittger, *et al* 2007). *MPL* mutation allele burden has been reported to be higher than *JAK2V617F* allele burden in some PMF patients with mutations in both *JAK2* and *MPL* (Lasho, *et al* 2006).

Mutations in *MPL* are clonal, stem cell derived and involve myeloid and lymphoid lineage cells (Chaligne, *et al* 2007; Pardanani, *et al* 2007b). These mutations induce oncogenesis by constitutive activation of a number of signalling pathway proteins which include JAKs and STATs (Pikman, *et al* 2006). This constitutive activity of mutant MPL may require specific residues such as Y112 (Pecquet, *et al* 2010). In murine bone marrow transplantation assays, *MPLW515L* mutation was found to result in a PMF-like disorder with co-existing thrombocytosis (Pikman, *et al* 2006). MPL is the key growth and survival factor for megakaryocytes consistent with the affects of the mutations being associated with higher platelet count, low haemoglobin counts and arterial thrombosis (Beer, *et al* 2008; Vannucchi, *et al* 2008a)

1.5.2 CALR

Recently, somatic mutations in exon 9 of calreticulin or *CALR* gene were identified in ET and PMF patients and these were usually mutually exclusive to mutations in *JAK2* and *MPL* (Klampfl, *et al* 2013; Nangalia, *et al* 2013). However, there is at least one known report of a rare co-occurrence of *CALR* and *JAK2V617F* in one ET patient which may reflect the possibility of bi-clonal evolution of the disease (Lundberg, *et al* 2014). The mutational frequencies described in ET and PMF cases, which were non-mutated for *JAK2* or *MPL*, were 67-71% and 56-88% respectively (Klampfl, *et al* 2013; Nangalia, *et al* 2013). Mutations in *CALR* were less frequently observed in other myeloid malignancies and not detected in lymphoid malignancies (Broseus, *et al* 2014; Hou, *et al* 2014; Nangalia, *et al* 2013).

CALR is present on chromosome 19p13.3 and the mutations, which occur in MPN, have predominantly been heterozygous insertions and deletions in exon 9. CALR is a complex calcium-binding protein chaperone localised in the endoplasmic reticulum and possibly the nucleus. Mutations in *CALR* have been reported to express a mutant

protein with an altered C-terminus which lacks the endoplasmic reticulum targeting sequence (Klampfl, *et al* 2013; Nangalia, *et al* 2013). Consequences of over-expressing one of the most frequently occurring *CALR* deletion mutants (52-bp deletion; *c.1092_1143del*) was associated with increased STAT5 signalling and hypersensitivity to IL-3 (Klampfl, *et al* 2013). A recent study comparing expression of *CALR* between different sub-types of haematopoietic cells from normal and *CALR*-mutated individuals has suggested preferential expression in the megakaryocytic lineage (Vannucchi, *et al* 2014). However, further studies are required of CALR activity in MPN.

ET and PMF patients with mutations in *CALR* have been associated with higher platelet counts and lower haemoglobin levels when compared to patients with mutated *JAK2* (Nangalia, *et al* 2013). However in ET, mutations in *CALR* were associated with a lower risk of thrombosis and cardiovascular events when compared to *JAK2* mutated patients (Rotunno, *et al* 2013; Rumi, *et al* 2013). In addition, *CALR*-mutated PMF patients have been associated with a favourable survival compared to *JAK2*-mutated PMF (Guglielmelli, *et al* 2014; Tefferi, *et al* 2014a).

1.6 Other somatic mutations in PV, ET and PMF

In addition to the mutations identified in *JAK2*, *MPL* and *CALR*, a number of other mutations have been described in PV, ET and PMF. These mutations include a number of genes which encode TET2, IDH1, IDH2, DNMT3A, ASXL1, and EZH2, all of which have known epigenetic roles. However, mutations in these genes are not exclusive to MPN and can occur in addition to mutations in *JAK2*, *MPL* or *CALR*. The frequencies and significance of the mutations are described in the following sections.

1.6.1 TET2

Somatic mutations in *Ten Eleven Translocation 2* or *TET2* were initially described by Delhommeau and colleagues, and by Langemeijer and colleagues as occurring in MPN, MDS and AML (Delhommeau, *et al* 2009; Langemeijer, *et al* 2009). *TET2* is located on chromosome 4q24 and has multiple isoforms. One of its isoforms; *isoform A*, is the most affected by the mutations described so far. The spectrum of mutations includes missense, frameshift and nonsense scattered across several of its 12 exons.

Mutations in *TET2* have been observed in 17% of *JAK2V617F*$^{+}$ and 7% of *JAK2V617F*-negative MPN patients (Tefferi, *et al* 2009c). Furthermore, in MPN, mutations in *TET2* have been described as either preceding or following the acquisition of mutations in *JAK2* (in exon 12 and 14), or as occurring in an independent clone leading to a bi-clonal pattern of mutation acquisition (Abdel-Wahab, *et al* 2009; Beer, *et al* 2009; Delhommeau, *et al* 2009; Schaub, *et al* 2010). In addition, mutations in *TET2* have been identified in MDS, MDS/MPN, AML and CMML and, have been shown to co-exist with other disease-relevant mutations such as mutations in *KIT* and *RARA* (Abdel-Wahab, *et al* 2009; Langemeijer, *et al* 2009; Tefferi, *et al* 2006; Tefferi, *et al* 2009a; Tefferi, *et al* 2009b).

TET2 is one of three homologous human proteins belonging to the TET family which includes TET1 and TET3. These proteins have a highly conserved catalytic domain. In recent years both TET1 and TET2 have been shown to catalyse the conversion of 5-methylcytosine to 5-hydroxymethylcytosine resulting in the modification of DNA (Koh, *et al* 2011; Tahiliani, *et al* 2009). Also, it has been reported that TET proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxymethylcytosine (Ito, *et al* 2011). These products are speculated to be

intermediates in a mechanism of DNA demethylation. Consistent with this, mutations in *TET2* have been associated with DNA hyper-methylation phenotype in AML (Figueroa, *et al* 2010).

MDS/MPN, MDS and AML patients with mutations in *TET2* have been shown to possess lower amounts of 5-hydroxymethylcytosine and altered DNA methylation patterns consistent with defects in the DNA demethylation pathway (Ko, *et al* 2010). Furthermore, functional evidence for a key role of TET2 was recently described wherein *Tet2*-deficient mice had higher numbers of HSPCs as compared to wildtype mice (Shide, *et al* 2012). The progenitors isolated from the *Tet2*-deficient mice had increased self renewal capacity and growth advantage over the wildtype mice in competitive transplant experiments. This suggested that Tet2 played a crucial role in maintaining the HSPC pool (Shide, *et al* 2012).

1.6.2 IDH1 and IDH2

Mutations in *isocitrate dehydrogenase 1*, or *IDH1*, and, *isocitrate dehydrogenase 2*, or *IDH2*, were initially described in gliomas and AML, following which they have been found to occur less frequently in other tumours (Kang, *et al* 2009; Mardis, *et al* 2009; Park, *et al* 2009; Parsons, *et al* 2008). Mutations in *IDH* have been described in transformed MPN such as post-MPN AML (secondary AML), PMF, and less frequently in PV or ET [(Abdel-Wahab, *et al* 2010; Green, *et al* 2010; Pardanani, *et al* 2010; Tefferi, *et al* 2010a) and reviewed in (Tefferi 2010)]. The mutations identified in MPN are heterozygous, somatic, and affect three specific arginine residues (which are identical to the residues affected in gliomas): R132 (IDH1), R172 (IDH2) and R140 (IDH2).

IDH1 (10 exons; located on chromosome 2q33.3, 414 amino acids) and *IDH2* (11 exons; located on chromosome 15q26.1, 452 amino acids) are located in the cytoplasm and peroxisomes, and mitochondria respectively. However, both proteins are homodimeric $NADP^+$-dependent enzymes that catalyse the oxidative decarboxylation of isocitrate to α–ketoglutarate generating NADPH from $NADP^+$. Functional characterisation of the mutations of *IDH1* has suggested a decreased affinity of the mutant enzyme for isocitrate (loss of function) and gain-of-function in catalysing the NADPH-dependent reduction of α–ketoglutarate to 2-hydroxyglutarate (Dang, *et al* 2009; Yan, *et al* 2009; Zhao, *et al* 2009b). Further, mutations affecting IDH1 and IDH2 have been associated with accumulation of 2-hydroxyglutarate in AML (Gross, *et al* 2010). Accumulation of 2-hydroxyglutarate (*R*-enantiomer) has reported to inhibit TET2 activity thereby mediating leukaemic transformation [refer to ***Figure 1.7*** (Figueroa, *et al* 2010; Losman, *et al* 2013)].

1.6.3 DNMT3A

Recurrent somatic mutations of *DNA (cytosine-5-)-methyltransferase 3 alpha*, or *DNMT3A*, have been described in 13-22% of AML and 8% of MDS cases (Ley, *et al* 2010; Walter, *et al* 2011; Yan, *et al* 2011). Mutations in *DNMT3A* have been associated with poor prognosis, altered methylation and gene expression profiles in AML and MDS (Ley, *et al* 2010; Shivarov, *et al* 2013; Walter, *et al* 2011; Yan, *et al* 2011). In the case of MDS, these mutations are associated with risk of progression to AML (Walter, *et al* 2011). A single mutation in *DNMT3A* (R882H) has been reported in a case of post MPN-AML which was negative for the *JAK2V617F* and *MPL* mutations (Yamashita, *et al* 2010). In chronic-phase MPN, mutations in *DNMT3A* have been reported at frequencies of approximately 2.6-7% [see chapter 3, (Abdel-Wahab, *et al* 2011; Rao, *et al* 2012; Stegelmann, *et al* 2011)].

De novo DNA methylation is carried out either by DNMT3A (DNA Methyltransferase 3 Alpha) or by DNMT3B together with DNMT3L at CpG dinucleotides (Chen, *et al* 2003; Hsieh 2005; Okano, *et al* 1999; Suetake, *et al* 2004). The roles played by DNMT3A, IDH1, IDH2 and TET2 in maintaining DNA methylation is summarised in ***Figure 1.7***.

In AML, missense mutations in *DNMT3A* are clustered within the methyltransferase domain of the protein, most commonly affecting an arginine (R) at position 882 (R882). Furthermore, DNMT3A R882 is analogous to the R823 residue of DNMT3B that is mutated in the inherited Immunodeficiency, Centromere instability, Facial anomalies (ICF) Syndrome (Ehrlich 2003; Ehrlich, *et al* 2001; Yamashita, *et al* 2010). Importantly this residue is involved in homo-dimerisation and functional activation of the protein (Jia, *et al* 2007). *In-vitro* studies with missense mutations affecting R882 suggest that mutant proteins are associated with reduced methyltransferase activity (Yamashita, *et al* 2010; Yan, *et al* 2011). Further, Challen and colleagues have demonstrated that the loss of *Dnmt3a* alters global DNA methylation patterns in the HSC and impairs differentiation [block in differentiation, (Challen, *et al* 2012)]. This may contribute to the establishment of tumour-initiating cells and may result in the expansion of immature cells that acquire additional mutations for the onset of myeloid disorders. Consistent with this theory, mutations in *DNMT3A* have also been recently reported to occur early in AML transformation in pre-leukaemic HSCs (Shlush, *et al* 2014).

Figure 1.7. The process of DNA methylation and the roles played by TET2, IDH1 and IDH2 and, DNMT3A. The enzymes mutated in MPN and AML have been marked with an asterisk (*****). 2-HG, 2-hydroxyglutarate; α-KG, α-ketoglutarate, AID, activated induced cytidine deaminase; BER, base excision repair; DNMT3A, DNA methyltransferase 3A; IDH, isocitrate dehydrogenase; $NADP^+$, nicotinamide adenine dinucleotide phosphate; SAH, *S*-adenosylhomocysteine; SAM, *S*-adenosylmethionine; TDG, thymine DNA glycosylase. *Figure taken from Shih, 2012 (Shih, et al 2012).*

1.6.4 ASXL1

Additional Sex Combs Like 1 or *ASXL1* (chromosome 20q11.1) is a member of the polycomb gene family. Somatic exon 12 mutations of *ASXL1*, which affect the C-terminal plant homeodomain (PHD), have been identified in MPN at an approximate frequency of 8% (Carbuccia, *et al* 2009). In addition, exon 12 mutations of *ASXL1* have been identified in 11% of patients with MDS, 43% of CMML and 47% with secondary AML [where AML was secondary to PMF, CMML or MDS, (Carbuccia, *et al* 2010; Gelsi-Boyer, *et al* 2009)]. Furthermore, in a study of 63 patients with secondary AML, mutations in *ASXL1* did not appear to be acquired during leukaemic transformation, were mutually exclusive to mutations in *IDH1* and were shown to co-exist with and even predate mutations in *JAK2* or *TET2* (Abdel-Wahab, *et al* 2010).

Mutations in *ASXL1* are predominantly found as nonsense and frameshift alterations in exon 12 leading to truncation of the C-terminal PHD domain. Whether these are loss-of-function or gain-of-function mutations is not completely understood. An early constitutive *Asxl1* knockout murine study reported that loss of *Asxl1* did not develop MDS- or leukaemia-like phenotype (Fisher, *et al* 2010). Nevertheless, the associated phenotype displayed mild haematopoietic defects and did not alter long term reconstitution in competitive repopulation studies (Fisher, *et al* 2010). However, other reports have suggested that mutations in *ASXL1* may lead to leukaemogenesis via the loss of polycomb repressive complex (PRC2)-mediated trimethylation of histone H3K27 (Abdel-Wahab, *et al* 2012; Abdel-Wahab, *et al* 2013). In further support of a role in leukemogenesis, Inoue and colleagues reported that C-terminal truncating mutations of Asxl1, in mice, were associated with an MDS-like disease development (Inoue, *et al* 2013). It was also reported that these truncating mutations of Asxl1 may inhibit PRC2-mediated methylation of H3K27 leading to derepression

of genes such as *homeobox a9* (*Hoxa9*) and *miR-125a* thereby promoting myeloid transformation (Inoue, *et al* 2013).

1.6.5 EZH2

Enhancer of zeste homolog 2 or *EZH2*, located on chromosome 7q36.1, encodes a member of the polycomb group of proteins that forms the catalytic subunit of the PRC2 complex [(Cao, *et al* 2002) and reviewed in (Cao, *et al* 2004)]. The PRC2 complex consists of four core members: EZH1 or EZH2, embryonic ectoderm development protein (EED), suppressor of zeste 12 homologue (SUZ12) and RBAP48 (also known as retinoblastoma binding protein 4 or RBBP4). EZH2 is a histone H3K27 methyltransferase, the activity of which is associated with transcriptional repression (Cao, *et al* 2002). Mutations in *EZH2* occur throughout the open-reading frame and have been detected in patients with *Ph*-negative MPN, MDS (approximately 5-15%) and MPN/MDS overlap disorders [approximately 10-12%, (Ernst, *et al* 2010; Makishima, *et al* 2010; Nikoloski, *et al* 2010)]. Mutations in *EZH2* have been observed in MF [includes both PMF and post-PV- or post-ET-MF, approximately 13%], PV (approximately 3%) and, appear to be virtually absent in ET (Ernst, *et al* 2010). Of interest is the discovery of germline mutations in *EZH2* in Weaver's syndrome, which overlap considerably with the somatic mutations identified in myeloid malignancies (Cross 2012; Tatton-Brown, *et al* 2011). Two separate individuals with germline mutations in *EZH2* were also observed to progress to neuroblastoma along with acute lymphoblastic leukaemia (ALL) and lymphoma respectively (Tatton-Brown, *et al* 2011).

Previous studies indicate that mutations in *EZH2* confer loss-of-function, thereby suggestive of a role as a tumour suppressor (Ernst, *et al* 2010; Nikoloski, *et al* 2010). However, a study by Herrera-Merchan and colleagues demonstrated that ectopic

over-expression of *Ezh2* can lead to myeloid transformation (MPN-like phenotype), suggesting that, on over-expression, *EZH2* may act as an oncogene in the myeloid compartment (Herrera-Merchan, *et al* 2012). Other studies have demonstrated that deletion of *Ezh2* can attenuate the leukaemogenicity of AML in mouse models (Neff, *et al* 2012; Tanaka, *et al* 2012). More recently, Muto and colleagues confirmed that loss of *Ezh2* can lead to an MDS-like disorder in mice and this was more severe with a concomitant loss of *Tet2* (Muto, *et al* 2013). Thus, the above and other studies are indicative of both tumour suppressor and oncogene associated roles of EZH2 (Khan, *et al* 2013; Morin, *et al* 2010; Ntziachristos, *et al* 2012; Simon, *et al* 2012). However, additional studies need to be performed to address the specific role of mutations in *EZH2* in MPN.

1.7 Evidence for pre-*JAK2* mutation event

It is now well established that somatic mutations of *JAK2*, particularly *JAK2V617F* are a central and significant feature of MPN. This is evident in the high frequencies observed of *JAK2V617F* mutations in MPN and supported by the constitutive activity observed in *in-vitro* (James, *et al* 2005) and murine models [as reviewed in (Li, *et al* 2011a)]. However, a number of questions remain with regards to the role of mutations of *JAK2* in MPN. Described below are theories which suggest that mutations affecting JAK2 are the main lesions to drive MPN disease, and other theories which suggest that additional events may contribute to the MPN disease phenotype.

1.7.1 JAK2V617F allele burden drives MPN disease phenotype: the gene dosage hypothesis.

Results from murine models support the notion that expression of *JAK2V617F* leads to an MPN-like disease phenotype (Akada, *et al* 2010; Bumm, *et al* 2006; James, *et*

al 2005; Lacout, *et al* 2006; Shide, *et al* 2008; Tiedt, *et al* 2008; Wernig, *et al* 2008; Xing, *et al* 2008; Zaleskas, *et al* 2006). Initial studies in bone marrow transplantation models, using retrovirus-transduced *JAK2V617F* BM cells reported MPN-like disease development (Bumm, *et al* 2006; Lacout, *et al* 2006; Wernig, *et al* 2008; Zaleskas, *et al* 2006). However, the extent of transformation either to PV- or MF-like disease phenotype was reported to be dependent on strain-specific differences. The differences observed may also have been due to different viral integration sites, level of expression and cells targeted by the viruses.

Later transgenic and *knock-in* murine models suggested that dosage of *JAK2V617F* defines MPN disease sub-type (Akada, *et al* 2010; Shide, *et al* 2008; Tiedt, *et al* 2008). Tiedt and colleagues reported that both ET- and PMF-like phenotypes were induced following manipulation of the *JAK2V617F* allele burden (Tiedt, *et al* 2008). In this study, mice with *JAK2V617F* expression lower than endogenous wildtype *Jak2* displayed an ET-like phenotype. However, a PV-like phenotype was reported when *JAK2V617F* expression was equal to endogenous wildtype *Jak2* (Tiedt, *et al* 2008). *Knock-in* models have shown that heterozygous and homozygous *Jak2V617F* leads to PV-like phenotype, with homozygosity resulting in additional MF-like features (Akada, *et al* 2010). Recently Li and colleagues have confirmed that homozygosity of *JAK2V617F* acts as a *switch*, leading to a PV-like phenotype from heterozygous *JAK2V617F*$^{+}$ ET-like disease phenotype in a murine model of MPN (Li, *et al* 2014).

Analysis of single cell progenitors from PV and ET patients by Scott and colleagues showed that all PV patients have a subset of homozygous *JAK2V617F*$^{+}$ cells (Scott, *et al* 2006). However, homozygous *JAK2V617F*$^{+}$ cells were absent from patients

with ET. This raised the possibility that *JAK2V617F* allele burden may drive disease phenotype (*gene-dosage* hypothesis). The *gene-dosage* hypothesis has been supported by other studies which implicate allele burden and strength of JAK2V617F signalling in different MPN disease phenotypes (Passamonti, *et al* 2009; Passamonti, *et al* 2010b). Further, the initial report of loss-of-heterozygosity (somatically acquired uniparental disomy) on chromosome 9p (9pLOH) being observed more frequently in PV as compared to ET also supports the *gene-dosage* hypothesis (Kralovics, *et al* 2002; Kralovics, *et al* 2005).

Thus, several clinical and experimental data support the hypothesis that *JAK2V617F* expression is the primary cause for the emergence of MPN. However, the *gene-dosage* hypothesis does not entirely explain the presence of low *JAK2V617F* allele burdens in PV patients, or the finding that *JAK2V617F* allele burdens in PV is often similar to those in PMF [(Barosi, *et al* 2007; Baxter, *et al* 2005; James, *et al* 2005; Kralovics, *et al* 2005; Levine, *et al* 2005; Vannucchi, *et al* 2007; Vannucchi, *et al* 2008b) and reviewed in (Jones, *et al* 2013)]. In addition it has also been reported that *JAK2V617F* may not entirely be sufficient to drive clonal expansion of HSPCs (Hasan, *et al* 2014; Li, *et al* 2014; Mullally, *et al* 2010; Mullally, *et al* 2012). This may suggest the presence of a *JAK2V617F* disease sub-clone, which does not always account for transformation to post-MPN-MF (Levine, *et al* 2006) and leukaemia (Beer, *et al* 2009; Jamal, *et al* 2008; Theocharides, *et al* 2007).

1.7.2 Additional events which may be important in driving the JAK2V617F disease phenotype: transformation to MF and AML

With regards to post-PV/post-ET-MF, it is apparent that transformation to MF (and AML) may require additional factors (Beer, *et al* 2009). A number of acquired lesions have been described to associate with transformation to MF and AML [for

example *ASXL1*, *IDH1/IDH2*, *EZH2* etc., and see ***section 1.6***]. However, these described lesions are neither exclusive to MPN nor follow a specific disease pattern. In addition, it has been observed that approximately 50% of *JAK2V617F*$^{+}$ MPN cases can progress to *JAK2V617F*-negative AML independently of cytoreductive therapy (Bjorkholm, *et al* 2011; Campbell, *et al* 2006; Merlat, *et al* 1999; Sterkers, *et al* 1998; Theocharides, *et al* 2007). Of note, progression in *MPL*-mutated MPN patients to *MPL* mutation-negative sAML has also been observed to occur in the absence of prior cytoreductive therapy (Beer, *et al* 2010).

Given that *JAK2V617F* is always acquired somatically there are some contradicting views suggesting that *JAK2V617F* mutation may not entirely dictate MPN pathogenesis (Godfrey, *et al* 2012; Li, *et al* 2014; Nussenzveig, *et al* 2007). These views, however, may highlight differences and the presence of multiple factors in disease pathogenesis in the presence of *JAK2V617F*. One suggested mechanism proposed include the preferential activation of STAT1 and STAT5 which may contribute to the differences between PV and ET. Increased STAT1 activity in normal progenitors has been reported to produce an ET-like phenotype, whereas down-regulation of STAT1 activity in *JAK2V617F* heterozygous ET progenitors produced a PV-like phenotype (Chen, *et al* 2010). Other theories are described below.

1.7.2.1 Evidence that clonal haematopoiesis is a pre-JAK2 mutation event

Late disease onset and clonal haematopoiesis are central features of MPN. Clonal haematopoiesis has been reported to occur with aging in the absence of *JAK2V617F* mutation (Kralovics, *et al* 2006; Nussenzveig, *et al* 2007). The features of an ageing haematopoietic system may include bias to differentiation along the myeloid lineage and an increased risk of malignancies (Beerman, *et al* 2010; Pang, *et al* 2011). Such

clonal haematopoiesis has been postulated to occur even in *sub-clinical* or *normal* individuals as a result of an aging haematopoietic system which has accumulated somatic chromosomal lesions [(Busque, *et al* 2012; Forsberg, *et al* 2012; Jacobs, *et al* 2012; Laurie, *et al* 2012) and reviewed in (Them, *et al* 2013)]. Genome-wide association studies on the general population performed by Jacobs and colleagues, and Laurie and colleagues are consistent with the increased frequency of acquiring chromosomal aberrations with age, and these changes correlated with changes observed in myeloid malignancies (Jacobs, *et al* 2012; Laurie, *et al* 2012). The study by Laurie and colleagues (Laurie, *et al* 2012) also showed a 10-fold increased risk of developing a haematological malignancy for those individuals with the presence of any detectable aberrations. In addition, Busque and colleagues identified 5.5% of the elderly population (>65 years) with somatic mutations in *TET2* exhibiting clonal haematopoiesis but no haematological malignancy (Busque, *et al* 2012). These reports suggest that an ageing haematopoietic system can shift to clonal haematopoiesis which may constitute an event preceding acquisition of mutations in *JAK2* and development of MPN in humans [reviewed in (Them, *et al* 2013)]. Studies have suggested populations of pre-leukaemic HSC may be associated with mutations in *TET2*, *DNMT3A* and *IDH1/IDH2* [reviewed in (Chan, *et al* 2013)]. This has led to a theory that certain combinations of mutations can induce expansion of pre-leukaemic HSC clones which may be resistant to cyto-reductive therapy and possibly the target of further leukaemogenic mutations.

1.7.2.2 Evidence from familial studies

Familial MPN mirrors sporadic MPN with regard to features such as risk of thrombosis, risk of haemorrhage, presence of clonal haematopoiesis and the ability to form EPO-independent erythroid colonies *in-vitro* (Bellanne-Chantelot, *et al* 2006; Kralovics, *et al* 2003; Rumi, *et al* 2006). Additionally, familial MPN has also been

observed to progress to MF and leukaemia (Bellanne-Chantelot, *et al* 2006; Rumi, *et al* 2007). Thus, it is believed that inherited risk/susceptibility factors may also play a role in the acquisition of *JAK2V617F* in sporadic MPN. Evidence for this is supported by studies of familial MPN in which mutations in *JAK2* (*JAK2V617F* and exon 12 mutations*)* and *CALR* are acquired (Bellanne-Chantelot, *et al* 2006; Maffioli, *et al* 2014; Pardanani, *et al* 2006b; Pietra, *et al* 2008; Rumi, *et al* 2014; Rumi, *et al* 2007; Rumi, *et al* 2006).

Analysis of families with MPN has shown that the inheritance of disease subtype and somatic acquisition of *JAK2V617F* is heterogeneous where most families display an autosomal dominant mode of inheritance with low penetrance (Bellanne-Chantelot, *et al* 2006; Cario, *et al* 2005; Kralovics, *et al* 2005; Kralovics, *et al* 2003; Landgren, *et al* 2008; Levine, *et al* 2005; Pardanani, *et al* 2006b; Rumi, *et al* 2007; Rumi, *et al* 2006). While 60% of these families typically either present with PV or ET, the remaining 40% may have a different myeloid disorder, such as mastocytosis or CML, at diagnosis (Bellanne-Chantelot, *et al* 2006; Rumi, *et al* 2007). This suggests multiple factors/lesions can predispose to the development of MPN. In a single large population-based study from Sweden, Landgren and colleagues reported a 5- to 7-fold increased risk of developing MPN in first degree relatives of MPN patients compared to the general Swedish population with no known family history of MPN (Landgren, *et al* 2008). This inherited risk was specific to PV (relative risk was 5.7) and ET (relative risk was 7.4) but was not observed in CML where the relative risk was 1.9 with a *p*-value of 0.09. However, it was not until 2009 that a risk single nucleotide polymorphism (SNP) was identified for the development of MPN.

1.7.2.3 JAK2 haplotype: a pre-disposition loci in sporadic MPN

In 2009, three studies reported that a major component of population-level predisposition for MPN relates to the *JAK2* locus itself, specifically a haplotype referred to as 46/1 (GGCC) (Goldin, *et al* 2009; Jones, *et al* 2009a; Kilpivaara, *et al* 2009; Olcaydu, *et al* 2009). The studies demonstrated that heterozygotes for this haplotype had a high probability to acquire *JAK2V617F* mutation in *cis* to the risk SNP allele than on the other chromosome. Furthermore, it was reported that the GGGC genotype frequency was much higher in MPN patients compared with matched controls, suggesting the G allele (characterised by the presence of the SNP rs10974944) is a dominant predisposition allele associated with a 3- to 4- fold increased risk of developing MPN (Jones, *et al* 2009a; Kilpivaara, *et al* 2009; Olcaydu, *et al* 2009). It is speculated that 46/1 predisposes to the development of mutations in *JAK2* or *MPL* (*hypermutability hypothesis*) or alternatively, confers a selective advantage to cells that have acquired mutations in *JAK2* (*fertile ground hypothesis*) (Campbell 2009; Patnaik, *et al* 2010b; Pietra, *et al* 2012).

Other emerging studies have associated other SNPs with *JAK2V617F*$^{+}$ MPN disease development (Ohyashiki, *et al* 2012; Wang, *et al* 2012) however, none have demonstrated a disease association as significant as the *JAK2* 46/1 haplotype. The mechanisms associated with these MPN risk alleles are not yet understood.

Thus, together the theories and related evidence for pre-clinical clonal haematopoiesis and pre-leukaemic HSCs suggests a model where acquisition of mutations in *JAK2*, *MPL* and *CALR* occurs on a background of changes that initiate clonal haematopoiesis and myeloid expansion. In this setting, the mutations in *JAK2/MPL/CALR* are proposed to specify the MPN disease phenotype (see ***Figure 1.8***).

Figure 1.8. Schematic representation of a hypothetical view of MPN disease initiation and progression to AML (see text for details). The model suggests a polyclonal population of HSC/HSPCs which harbour disease susceptibility alleles. Clonal haematopoiesis may occur depending on the nature of the disease initiating event within the HSC/HSPC. Acquisition of further somatic mutations in genes such as *JAK2, MPL* or *CALR* will define the phenotype of the disease (such as PV, ET or PMF) and potential transformation to AML (or MF in the case of PV and ET). Leukaemic transformation can also occur in pre-leukaemic cells that have not acquired mutations in *JAK2*, *MPL* or *CALR*. *Figure adapted from(Tefferi 2010)]*.

1.8 Histones and MPN disease pathogenesis

Mutations in genes known or suspected to encode epigenetic regulators have been shown to be common in MPN. These include *TET2*, *IDH1/IDH2*, *DNMT3A*, *ASXL1*, *EZH2* and *JAK2* (as described in previous sections). As a result, there has been growing interest in understanding epigenetic mechanisms in MPN. Epigenetics refers to changes in phenotype or gene expression that are heritable through cell divisions but are caused by mechanisms other than changes in the underlying DNA sequence [reviewed in (Goldberg, *et al* 2007)]. In MPN, the epigenetic mechanisms investigated include DNA methylation of CpG-dinucleotides (for example due to mutations in *DNMT3A* or *TET2* which have been described in ***section 1.6***), regulation by micro-RNAs and alterations to chromatin function via histone modifications [(Bruchova, *et al* 2008; Bruchova, *et al* 2007; Dawson, *et al* 2009; Guglielmelli, *et al* 2007; Liu, *et al* 2011; Slezak, *et al* 2009; Zhao, *et al* 2009a) and reviewed in (Cross 2011)]. Interestingly, JAK2, DNMT3A, EZH2 and ASXL1 have been shown to interact with and/or modify histones, particularly histone H3 and H4, at specific residues. Specific residue modifications of interest include histone H3 lysine 27 trimethylation ($H3K27me^3$) by EZH2, histone H3Y41ph by JAK2 and histone H4 arginine 3 symmetric di-methylation ($H4R3me^2s$) by PRMT5 (Cao, *et al* 2002; Dawson, *et al* 2009; Liu, *et al* 2011). These modifications have been shown to have a direct effect on the regulation of a number of specific downstream genes such as *LMO2* (JAK2 mediated H3Y41ph) (Dawson, *et al* 2009) and *early growth response protein-1* or *Egr1* (EZH2 mediated $H3K27me^3$) (Tanaka, *et al* 2012). Also of interest is the role played by mutant ASXL1, which may lead to loss of activity of the PRC2 complex leading to reduction in the level of histone H3K27 methylation and promote myeloid leukaemogeneis (Abdel-Wahab, *et al* 2012). Histone H4R3 can be di-methylated by PRMT5, a target of JAK2 kinase activity, which may affect

binding and activity of DNMT3A (Dawson, *et al* 2009; Liu, *et al* 2011; Zhao, *et al* 2009a). Histone H4R3me^{2}s, mediated by PRMT5, has been shown to recruit DNMT3A in human erythromyeloblastoid leukaemia cell line (K-562) and also lead to silencing of globin genes in erythroid progenitors from adult human bone marrow cells (Zhao, *et al* 2009a).

Taken together the above evidence suggests that histone H3 and histone H4 modifications have an important role in haematopoiesis and specifically in mediating the effects of mutated proteins in MPN. Below is an introduction to the histones including the roles played by the variety of modifications on the histone N-terminal tails.

1.9 Histones and histone genes

Histones are highly conserved basic proteins that package DNA into functional units know as nucleosomes. An octamer of histone proteins constitute the nucleosome core which is the basic unit of chromatin [reviewed in (Alberts, *et al* 2002)]. This nucleosome consists of a tetramer of $(H3\text{-}H4)_2$ flanked by two H2A-H2B dimers. In higher eukaryotes, these nucleosomes are joined together by variable lengths of linker DNA (18-65 base pairs). Linker histones (H1) shield the excess negative charge of the linker DNA to promote folding of nucleosomes into a higher order structure known as chromatin [reviewed in (Kornberg, *et al* 1999; Luger, *et al* 1998)]. The overall structural state of the chromatin impacts DNA replication, expression of genes and DNA repair [(Dawson, *et al* 2009; Petruk, *et al* 2013) and reviewed in (Zentner, *et al* 2013)]. For example, chromatin structure is altered following DNA damage which activates DNA repair processes and proteins which recognize specific modifications on histone N-terminal tails of core histones (such as

histone H4 lysine 20 methylation) and activate specific histone variants such as gamma H2A.X (Botuyan, *et al* 2006; Hanasoge, *et al* 2007; Rogakou, *et al* 1998).

1.9.1 Organisation of histone genes and regulation of transcription

Histone gene expression is tightly coupled to DNA replication. At the onset of the DNA synthesis phase (*S*-phase), *de novo* synthesis of histone proteins is required to package nascent DNA into chromatin [reviewed in (Alberts, *et al* 2002)]. This tight control of histone protein synthesis is essential to ensure proper chromatin formation. Over-expression of histones has been shown to increase the frequency of chromosome loss in yeast (Meeks-Wagner, *et al* 1986), whereas, inhibition of histone gene-expression during *S*-phase leads to DNA damage (Nelson, *et al* 2002; Ye, *et al* 2003). Inappropriate expression of histones outside of *S*-phase has been linked to developmental arrest in *Drosophila* (Sullivan, *et al* 2001). Deleting one of the two *histone H4* genes in *Saccharomyces cerevisiae* (*S.cerevisiae*) has been reported to increase the rate of homologous recombination and genomic instability (Prado, *et al* 2005). Thus, changes in histone levels have a potential impact on gene expression and genomic stability, which may be of importance in disease.

In all organisms, histone encoding genes are clustered together into two major clusters [reviewed in (Marzluff, *et al* 2002)]. The large cluster on chromosome 6 in the human genome [*HIST1* cluster, present on chromosome 13 in mouse] contains more than 80% (55 genes) of the histone encoding genes which includes all six of the *histone H1* genes. Two smaller clusters present on human chromosome 1q21 [*HIST2* cluster, present on chromosome 3 in mice] and 1q42 [*HIST3* cluster, present on chromosome 11 in mice] contain six and three of the four core histone encoding genes (the complete list of histone genes is shown in ***Figure 1.9***). The nature of histone gene organisation suggests a selective pressure to conserve histone gene

arrangement. It also indicates that gene-clustering may have an important role in the co-ordinated and controlled expression of histone encoding genes through a number of sequence elements and transcription factors [(Das, *et al* 1993; Holmes, *et al* 2005; Mitra, *et al* 2003; Palko, *et al* 2004; Schild-Poulter, *et al* 2003; Spector, *et al* 1997; van Wijnen, *et al* 1991; Xie, *et al* 2009)].

The mRNA encoding the replication-dependent histones are characterised by the lack of a poly-A tail and the presence of a unique 3´ untranslated region (3´UTR). The 3´UTR of histones contains a stem loop, a histone downstream element (HDE) and U7 small nuclear RNA (U7snRNA) binding site (within the HDE) [reviewed in (Marzluff, *et al* 2008)]. The stem loop is the binding region for the stem loop binding protein (SLBP) which is an essential protein in the histone mRNA metabolism (Salzler, *et al* 2009; Sullivan, *et al* 2001). SLBP aids in the binding of other histone processing factors, such as ribonucleoprotein U7snRNP and forms a complex which enables cleavage and export of the histone mRNA [(Azzouz, *et al* 2005; Jaeger, *et al* 2006; Wagner, *et al* 2006) and reviewed in (Marzluff, *et al* 2008)].

1.9.2. Histone post-translational modifications and the histone code hypothesis

The core histones are the main structural proteins of chromatin, and they mediate the stability and regulation of the genome. Each of the core histones has a related globular domain that mediates histone–histone interactions within the octamer and organizes the two turns of nucleosomal DNA [reviewed in (Alberts, *et al* 2002)]. Each histone has an N-terminal 20–35 residue segment that is rich in basic amino acids and extends from the surface of the nucleosome. These histone tails are the site of a variety of post-translational modifications (PTMs) that affect interactions with chromatin binding, remodelling proteins and regulatory processes. Some of the key

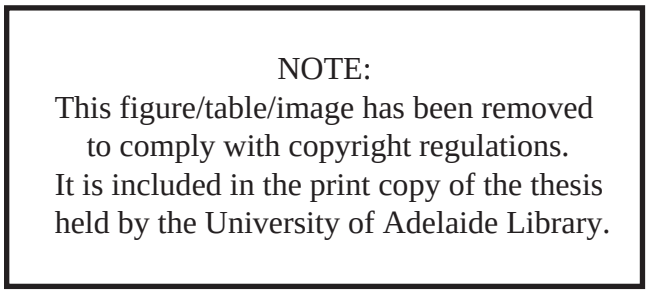
Figure 1.9. Histone genes in the human and mouse genome. Some genes for which no homologous gene has been found in either species have been denoted with 'Not present'; pseudogenes have been depicted with Ψ. *The figure has been adapted from (Marzluff, et al 2002b).*

modifications are listed in ***Figure 1.10A***. The variety of PTMs include acetylation, phosphorylation, methylation (me), deimination, ADP ribosylation, ubiquitylation, sumolyation, and the addition of β–N-acetylglucosamine [as reviewed by (Bannister, *et al* 2011)]. Overall, histone PTMs provide a layer of complex codes that mediate genome stability and gene regulation. It is not surprising that de-regulation of histone PTMs are found to occur in various forms of cancer (Elsheikh, *et al* 2009; Fraga, *et al* 2005; Kwon, *et al* 2010; Schwartzentruber, *et al* 2012; Seligson, *et al* 2005; Wei, *et al* 2008; Wu, *et al* 2012). These residues include histone H3 lysines at positions 4, 9, 18 and 27 and, histone H4 lysines at positions 12, 16 and 20 of the mature protein (see ***Figure 1.10B*** for summary). In addition to deregulated PTMs at histone tails in cancers, two groups were the first to report and highlight a role of mutations in genes encoding histone H3 (Schwartzentruber, *et al* 2012; Wu, *et al* 2012).

1.10 Histone gene mutations in gliomas

Two groups reported mutations in *histone H3* genes; *H3F3A* and *HIST1H3B*, in paediatric gliomas (Schwartzentruber, *et al* 2012; Wu, *et al* 2012). Schwartzentruber and colleagues performed whole-exome sequencing in paediatric diffuse gliomas, a lethal tumour occurring more frequently in children than in adults, and identified acquired mutations in *H3F3A* (encodes for variant histone H3; H3.3) in 53% of all cases [including 3% adult and 50% of the paediatric cases, (Schwartzentruber, *et al* 2012)]. The mutations were highly specific and resulted in amino acid substitutions of either lysine at position 27 to a methionine (K27M) or, glycine at position 34 to an arginine (G34R) or valine (G34V). In a concurrent study by Wu and colleagues, mutations in *H3F3A* were found to occur in 36% of the glioma cases and, interestingly, K27M somatic mutations were also found to occur in *HIST1H3B* at a frequency of 9% of all cases (Wu, *et al* 2012). All mutations identified were mutually exclusive,

A.



B.

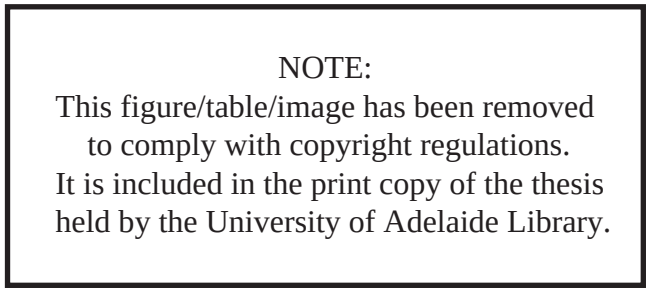


Figure 1.10. A. Post-translational modifications on the core histone residues. Histone PTMs, which occur mostly on N-terminal residues have been depicted. Those affected in disease are highlighted with yellow arrows. Note that disrupted modifications in disease affect Histone H3 and Histone H4 residues. *Figure taken from (Rodriguez-Paredes, et al 2011)* Me, Methylated; Ac, Acetylated; P, phosphorylated and Ub, ubiquitinated. **B. Deregulated PTMs on the residues of histone H3 and H4.** The table lists most studied key residues and the associated modifications of histones H3 and H4 which are deregulated in cancers.

heterozygous and acquired, which was suggestive of a single mutation in *HIST1H3B* or *H3F3A* possibly driving the pathogenesis of disease. Subsequently, other groups have reported somatic mutations in *H3F3A* (Behjati, *et al* 2013; Behjati, *et al* 2014; Huether, *et al* 2014; Je, *et al* 2013; Je, *et al* 2014) and *H3F3B* [mutations affecting K36 (Behjati, *et al* 2013; Behjati, *et al* 2014)] in other malignancies, such as rare acute leukaemias, lymphomas, and solid tumours, indicating a role for histone mutations in driving various cancer phenotypes.

Given the growing body of evidence implicating roles for histone H3 and histone H4 and associated proteins in various myeloid malignancies including MPN, studies were undertaken to investigate the direct role of mutations of histone encoding genes in MPN disease pathogenesis.

1.11 Hypothesis

Mutations in genes encoding for histone H3, histone H4, and associated proteins occur in MPN and play a role in MPN pathogenesis.

1.12 Project Aims

Aim 1: To screen for novel histone H4 variants associated with MPN and functionally characterise identified variants.

Aim 2: To identify and characterise novel MPN-associated somatic mutations in histone H3 and chromatin-associating proteins; PHF6 and DNMT3A.

Chapter 2: Identification of novel *histone H4* variants in *JAK2V617F* positive Polycythemia vera

Statement of Authorship

Title of Paper	**Identification of novel *histone H4* variants in *JAK2V617F* positive Polycythemia vera**
Publication Status	○ Published ○ Accepted for Publication ◉ Submitted for Publication ○ Publication Style
Publication Details	Rao N, Butcher CM, Kok CH, Neufing PJ, Lewis ID, Hahn CN, Scott HS, Bardy PG, D'Andrea RJ.

Author Contributions

By signing the Statement of Authorship, each author certifies that their stated contribution to the publication is accurate and that permission is granted for the publication to be included in the candidate's (***Nisha Rao***) thesis.

Name of Principal Author (Candidate)	Nisha Rao		
Contribution to the Paper	70%, performed the research, analysed and interpreted data and wrote the publication		
Signature		Date	08/04/2013

Name of Co-Author	Carolyn M. Butcher		
Contribution to the Paper	5%, supervised the research study and critical input of interpreted data		
Signature		Date	22/03/2013

Name of Co-Author	Chung H Kok		
Contribution to the Paper	5%, Bioinformatics processing of Microarray data		
Signature		Date	22/03/2013

Name of Co-Author	Petra J Neufing		
Contribution to the Paper	3%, design of QPCR primers and supervision of the initial screening of MPN samples		
Signature	*Not Available*	Date	Not applicable

Name of Co-Author	Ian D Lewis		
Contribution to the Paper	3%, provided clinical information on patients		
Signature		Date	21/03/2013

Name of Co-Author	Chris N Hahn		
Contribution to the Paper	2%, provided 4-OHT vectors and critical feedback on study		
Signature		Date	21/03/2013

Name of Co-Author	Hamish S Scott		
Contribution to the Paper	2%, provided access to Molecular pathology diagnostic samples		
Signature		Date	16/03/2013

Name of Co-Author	Peter G Bardy		
Contribution to the Paper	3%, provided clinical information and coordinated collection of MPN patient samples		
Signature		Date	08/04/2013

Name of Co-Author	Richard J D'Andrea		
Contribution to the Paper	7%, Principal supervision of the study, gave critical input on writing		
Signature		Date	22/03/2013

2.1 Introduction

The genes which encode for the four core histones are highly conserved across species with greater than 95% DNA sequence conservation between mice and human. In the human genome there is one major histone gene cluster, the *HIST1* cluster, present on chromosome 6 (6p21-22), and two minor histone clusters on chromosome 1 (1q21 and 1q42) [reviewed in (Marzluff, *et al* 2002)]. The major cluster on chromosome 6 has 55 histone genes of which 12 are *histone H4* genes. There are two identical *histone H4* genes in the *HIST2* cluster, *HIST2H4A* and *HIST2H4B.* In addition to the *histone H4* genes on chromosomes 6 and 1, one *histone H4* gene, *HIST4H4,* is present on chromosome 12 (12p13.1) ***(Figure 2.1)***. Of the 15 known *histone H4* genes, 14 encode for the identical 103 amino acid histone H4 protein, while *HIST1H4G* encodes for a 98 amino acid form of the histone H4 protein.

Residues on the N-terminal tail of histone H4 undergo a number of post-translational modifications, including acetylation and methylation of Lysine (K) and methylation of arginine (R) residues. These post-translational modifications of specific amino acids form the histone code, wherein combinations of modified residues on the same or separate histone tails (mostly N-terminal) can lead to recruitment of a range of chromatin-associating proteins resulting in either the repression or activation of genes (see ***chapter 1 section 1.8.2*** and ***1.9***). *S. cerevisiae* mutagenesis studies of the amino terminal tail residues in histone H4 have been reported in the HistoneHits database (http://histonehits.org). Mutations affecting key residues such as K20 and R3 have been shown to affect silencing of ribosomal genes and the *Hidden Mat Right* or *HMR* locus (Huang, *et al* 2009).

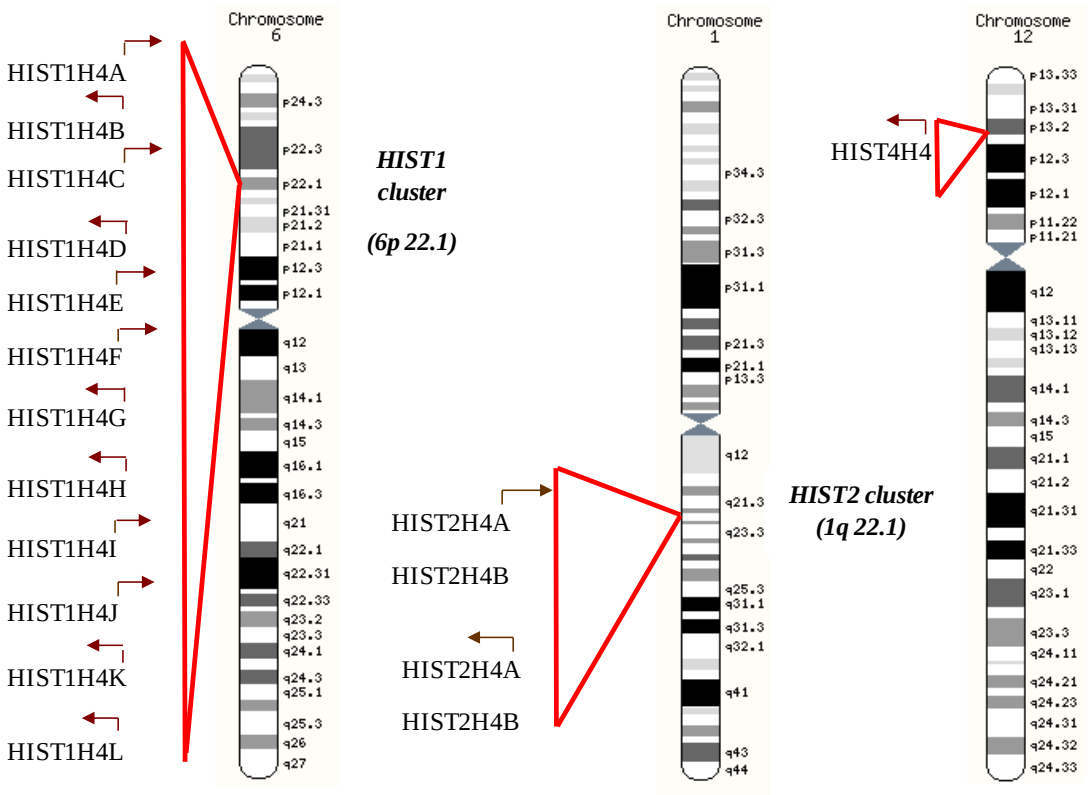


Figure 2.1. Organisation of *Histone H4* genes in the human genome. The *HIST1* cluster, on human chromosome 6 (6p21–p22) contains 55 histone genes including 12 *Histone H4* genes. There are two smaller clusters on human chromosome 1:*HIST2* (1q21), which has 6 *Histone* genes of which 2 genes encode for Histone H4 and *HIST3* (1q42), which has 3 *Histone* genes none encoding for Histone H4. Of note are the two identical genes (*HIST2H4A* and *HIST2H4B*) which have been represented with both annotations at both loci in the *HIST2* cluster due to the inability to differentiate between the two genes. In addition to the *Histone H4* genes in *HIST1* and *HIST2*, one other gene has been identified; *HIST4H4* which is present on human chromosome 12. The organization of the Histone genes in the *HIST1* cluster is conserved across the human and mouse genomes and is spread over approximately 2Mb. There are two large regions (>250 kb each) within the cluster where there are no Histone genes, but many other genes. ⮡ indicates 5`-3` transcription and ⮠ indicates 3`-5` transcription of the gene. (*Chromosomes/gene positions have been adapted from NCBI, USA*).

In higher eukaryotes histone modifications represent one of the major epigenetic events which regulate transcriptional activity. Proteins that perform these modifications are affected in MPN [(Dawson, *et al* 2009; Makishima, *et al* 2010; Nikoloski, *et al* 2010) and reviewed in (Abdel-Wahab 2011)]. In PV, ET and PMF, studies with JAK2 and EZH2 have highlighted the role of histone H3 and H4 post-translational modifications.

Histone modifications can be directly affected by altered JAK2 activity in MPN. JAK2 has been reported to translocate to the nucleus and phosphorylate histone H3 at a tyrosine residue at position 41 (Y41), thereby up-regulating expression of oncogenes such as *LMO2* and *BCL2L1* (Dawson, *et al* 2009). In addition, JAK2V617F has been reported to phosphorylate specific residues of PRMT5, leading to decreased PRMT5-MEP50 mediated methyltransferase activity and decreased levels of methylation on histone residues (Liu, *et al* 2011). Furthermore, EZH2 trimethylates histone H3 ($H3K27me^3$) and is mutated in MPN patients (see ***section 1.6.5 in chapter 1***). Given the ability of JAK2 and JAK2V617F to affect, either directly or indirectly, histone H3 and histone H4 modifications, a search for novel pathogenic variants in the genes encoding the histone H3 and histone H4 proteins was undertaken. In this first study, a detailed screen of the genes encoding for *histone H4* by Sanger sequencing was conducted. This identified novel germline coding variants, following which, their functional significance was examined further.

2.2 Materials and Methods

2.2.1 MPN, normal and other cohorts

Human Research Ethics Approval for this study was obtained from the Royal Adelaide Hospital (*RAH ethics #991104c*), The Queen Elizabeth Hospital *(CNAHS*

ethics #2007159) and Flinders Medical Centre (*FMC ethics #349/09*) Research Ethics Committees. 110 PV, 20 ET and 3 PMF patients and 239 haematologically normal volunteers were recruited over a period of 12 years (from 2000-2012). All PV patients were screened for *JAK2V617F* or exon 12 mutations of *JAK2* as described previously (Butcher, *et al* 2008). Access to a Melbourne MPN cohort [Prof. Andrew Roberts (WEHI, Melbourne)] of 60 patient *g*DNA samples from a combination of PV, ET and PMF] was also granted under the *CNAHS ethics # 2007159*.

Molecular Pathology diagnostic *g*DNA samples [*n*=850 (IMVS, SA Pathology diagnostic facility, Adelaide)] were also screened for this study (*CNAHS ethics# 2007159*). These Molecular Pathology *g*DNA samples were from individuals with non-haematological disorders and had been tested for a number of genetic lesions by the diagnostic facility. The tissue source of these samples (buccal, bone marrow or whole blood) and other patient details (except age and gender) were not accessible under the ethics approval.

2.2.2 Cell lines and culture conditions

A. Human embryonic kidney-293 and -293T (HEK293/HEK293T)

HEK293 and HEK293T cell lines (Graham, *et al* 1977) were obtained from American Type Culture Collection (ATCC; VA, USA; Cryosite Distribution Pty Ltd, Australia) and were maintained in Dulbecco's Modified Eagle Medium (DMEM; Sigma-Aldrich Pty Ltd, Australia), supplemented with 10% (v/v) Foetal Bovine Serum (FBS; Sigma-Aldrich Pty Ltd, Australia) and Penicillin-Streptomycin-Glutamine (PSG; 100x PSG, Gibco®, Invitrogen, Life Technologies Australia Pty Ltd, Australia). HEK293 cells containing the 4-Hydroxytamoxifen (4-OHT) inducible plasmid constructs were maintained in 500 µg/mL Hygromycin-B (Invitrogen, Life Technologies Australia Pty Ltd, Australia) and 2µg/mL of

Puromycin dihydrochloride (will be referred as Puromycin, Invitrogen, Life Technologies Australia Pty Ltd, Australia).

B. Human erythromyeloblastoid leukaemia cell line (K-562)

K-562 is a Ph^+ leukaemia cell line (Lozzio, *et al* 1975) which was obtained from ATCC and maintained in RPMI 1640 (Sigma-Aldrich Pty Ltd, Australia) supplemented with PSG and 10% heat inactivated FBS.

C. Human Erythroleukaemia cell line 92.1.7 (HEL)

HEL cells display amplification of *JAK2V617F* (approximately 4 copies of *JAK2V617F* allele and also have loss-of-function mutation in *TP53* (Martin, *et al* 1982; Zhao, *et al* 2012). These obtained from ATCC and were maintained in Iscove's Modified Dulbecco's Medium (IMDM; Sigma-Aldrich Pty Ltd, Australia) supplemented PSG and 10% heat inactivated FBS.

D. Mouse Foetal Derived Myeloid cell line (FDM)

The *HoxB8* transformed murine foetal derived myeloid (FDM) cell line was generously provided by Prof. Paul G. Ekert (The Walter and Eliza Hall Institute for Medical Research, Victoria, Australia). FDM cells were cultured in low glucose DMEM (1000mg glucose; Sigma-Aldrich Pty Ltd, Australia) supplemented with 10% FBS and 10ng/mL recombinant murine IL-3 (rmIL-3) (Ekert, *et al* 2004; Jabbour, *et al* 2010). rmIL-3 was produced from baculoviral vectors and generously supplied by Dr Andrew Hapel (John Curtin School of Medical Research, Canberra).

2.2.3 Cloning, transfection and transduction of cell lines

A. Cloning and lentiviral expression of HIST1H4C in HEK293

Human *HIST1H4C* [wild-type (WT) and mutant] or *Green Fluorescent Protein (GFP)* cDNA constructs were introduced into *BamHI* and *HpaI* cleaved lentiviral expression vector *pF5x_UAS_SV40puroW (pF5x_UAS_SV40*). For more information

of the lentiviral plasmid maps and cloning primer details, see ***Supplementary Figure S2.1 and Supplementary Table S2.1***. Plasmid stocks and information were generously provided by A/Prof. Chris Hahn and Prof. Hamish S.Scott.

To generate inducible cell lines, HEK293 were transfected with *pF5x_UAS_SV40* and *ap_pF_U_G147EV16_ PGK_ Hygro* (*GEV16*) plasmids separately, along with the packaging constructs, using the transfection reagent Lipofectamine™ 2000 (Invitrogen, Life Technologies Australia Pty Ltd, Australia). Viral supernatant thus generated was harvested at 48 hours post-transfection. Following this HEK293 cells were infected [using the spinfection method (Persons, *et al* 1998)], firstly, with *GEV16* lentivirus containing supernatant. This was followed by the selection of clonal Hygromycin-B resistant colonies by sorting for transduced cells, using Becton Dickinson Fluorescence Activated Cell Sorter (FACS) Aria III (BD Biosciences, USA), after 72 hours of culturing. A clonal population expanded from a single well was then transduced with *pF5x_UAS_SV40* containing lentivirus supernatant produced using either without (empty vector) or, with *HIST1H4C* (*WT*, *p.R4C* or *p.R56Q*) or *GFP* constructs. Cells resistant to both Hygromycin-B and Puromycin were selected and tested for gene expression. Expression of *HIST1H4C* or *GFP* was induced by treatment with 400nM 4-OHT for 24 hours (***Figure 2.2*** describes the work flow). Cells were harvested by vigorous washing of the flask with Dulbecco's modified phosphate buffered saline (DPBS, Invitrogen, Life Technologies Australia Pty Ltd, Australia) or DMEM.

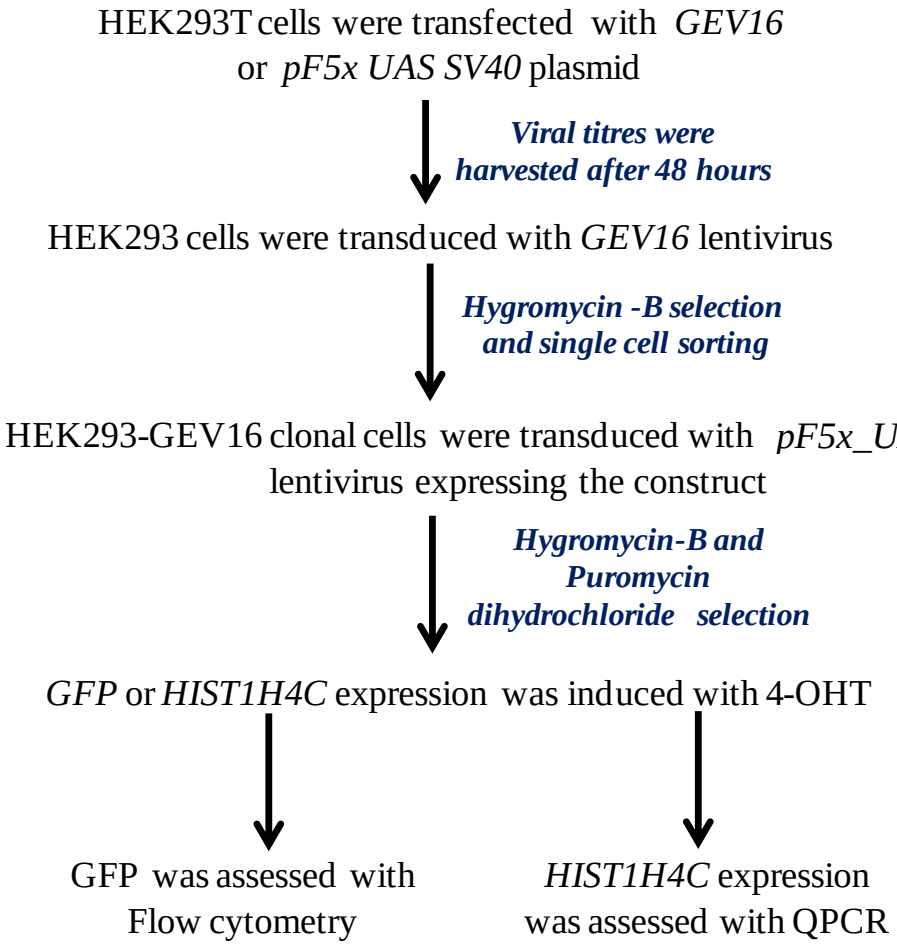


Figure 2.2. **Work flow for the generation of 4-OHT inducible HEK293 cells.** $5x10^5$ HEK293T cells in 5mLs media (without PSG) were seeded 24 hours prior to transfection. 8µg of total DNA (4µg each of lentiviral packaging vector and plasmid (*pF5x_UAS_SV40* expression vector and *GEV16* fusion construct expressing plasmid was transfected into HEK293T cells. Lentiviral containing supernatant was harvested from the HEK293T cells at 48 hours and used to transduce HEK293 cells using the spinfection method (Persons, Mehaffey et al. 1998). Briefly, cells and viral supernatant were centrifuged at 800*g* for 1 hour at 37°C with 4µg/mL polybrene, 0.25µg/mL fungizone and incubated for 24 hours before addition of 500µg/ml Hygromycin-B and/or 2µg/ml Puromycin dihydrochloride.

The cells were then washed twice by centrifugation at 250*g* for 5 minutes and resuspended in DPBS or DMEM for further analysis. For flow cytometric analysis, cells were fixed in 500 µl FACS FIX (also see ***section 2.2.7***). Cells were resuspended in 1mL Triton Extraction Buffer for extraction of histones (described in ***section 2.2.9***).

B. Cloning and retroviral expression of HIST1H4C in FDM

The cDNA for human *HIST1H4C* (WT and the variant p.R4C) was introduced into *EcorRI* cleaved of *MSCV-IRES-GFP* (*MIG*) expression plasmid (see ***Supplementary Table S2.1*** for primer details). Briefly, $5x10^5$ cells in 5mLs media (without PSG) were seeded 24 hours prior to transfection. *MIG* expression plasmid without (empty vector) or expressing the *HIST1H4C* (*WT* or *p.R4C*) construct (and mouse viral packaging plasmids) was transfected using Lipofectamine^{TM} 2000 (Invitrogen, Life Technologies Australia Pty Ltd, Australia). Media was harvested up to 48 hours after transfection. For FDM cells, 3mLs of each of the viral supernatant (empty, *WT* or *p.R4C* construct expressing) was applied to $1x10^5$ FDM cells. Cells were centrifuged at 800*g* for 1 hour at 37°C. This was followed by incubation for 6 hours in a 37°C and 5% CO_2 incubator. The FDM cells were centrifuged a second time at 800*g* for 1 hour at 37°C after replacing the viral supernatant with 1mL of fresh media and 2mLs of an additional viral supernatant. Cells were incubated overnight before performing FACS for separating GFP positive (GFP^+) cells using the Becton Dickinson FACS Aria III (BD Biosciences, USA). Sorted GFP^+ cells were recovered and maintained for 14 days in FDM media. The non-adherent FDM cells were harvested by spinning the cells at 250*g* for 10 minutes and resuspended in DPBS or FDM media. RNA prepared from FDM cells was assessed for quality using the Agilent 2100 Bioanalyser (Agilent Technologies Australia).

2.2.4 Primary Cell separation techniques

A. Isolation of peripheral blood mononuclear cells (PBMNC), granulocytes and bone marrow mononuclear cells (BMMNC) fractions

30mls of whole anti-coagulated blood was incubated for 30 minutes at room temperature with 15mls of a 5% solution of ultrapure grade Dextran Sulfate (MW 500,000, AMRESCO, Astral Scientific Pty Ltd, Australia) in sterile DPBS. Following the dextran-based sedimentation of blood density, centrifugation was performed at 800*g* to isolate PBMNC (middle furry layer) and granulocytes (pellet) using Ficoll-Paque PLUS^{TM} (GE Healthcare Inc., Uppsala, Sweden). The layers were aspirated and washed in sterile DPBS before proceeding to further separation of cell types or freezing in liquid nitrogen. BMMNC were isolated from bone marrow aspirate samples by density centrifugation, as described above.

B. Isolation of monocyte and granulocyte enriched cell fractions from PBMNC and haematopoietic stem cell/progenitor cells (HSPC) enriched cell fraction from BMMNC

Magnetic Activated Cell Sorting (MACS) based isolation of cells was performed to enrich for HSPCs ($CD34^+$) from BMMNC. PBMNC and granulocytes pellet were isolated from peripheral blood following dextran sedimentation and Ficoll-Paque PLUS^{TM} (GE Healthcare Inc., Uppsala, Sweden) based density centrifugation procedure. BMMNC were isolated from BM aspirates as described in the previous section. Anti-human CD34 (Catalog number 130-048-702), CD14 (Catalog number 130-050-201) and CD15 (Catalog number 130-046-601) conjugated magnetic beads (MACS, Miltenyi Biotec Inc., USA) were used to separate $CD34^+$, $CD14^+$ and $CD15^+$ enriched haematopoietic cells (***Supplementary Figure S2.2***). Briefly cells were incubated with beads at 4ºC for 15 minutes (for $CD14^+$ and $CD15^+$ enrichment) or 30 minutes (for $CD34^+$) and washed twice in MACS buffer (filter sterilised solution of 0.5% BSA and 2mM EDTA in PBS) at 300*g* for 10 minutes. The cells

were then resuspended in 500µl MACS buffer and passed through a MACS LS column. The column was washed three times with 3mLs of MACS buffer and the unbound cells (negative fraction) were pelleted at 300*g* and resuspended in TRIzol® (200-500µl) and stored in -80 ºC. The cells bound to beads were eluted with 5mLs of MACS buffer. This fraction (positive fraction) was washed twice in MACS buffer, resuspended in TRIzol® and stored at -80 ºC for further use (see ***section 2.2.5 for RNA and cDNA preparation***).

2.2.5 Nucleic Acid Extraction

A. Genomic DNA (gDNA)

*g*DNA from PBMNC, granulocytes and buccal mucosal cells (buccal swabs were obtained from air-dried cytologic brushes) samples were prepared using the QIAmp™ DNA Blood Mini Kit (QIAGEN™, QIAGEN™ Pty Ltd, Australia, Australia) according to the manufacturer's instructions.

B. RNA extraction and cDNA preparation from primary cells and cell lines

RNA was extracted using a modified TRIzol® (Invitrogen, Life Technologies Australia Pty Ltd, Australia) extraction method. Briefly, cells were lysed in TRIzol® and chloroform, then centrifuged for 15 minutes at 12000*g* at 4°C. The topmost chloroform layer was isolated and purified using isopropanol-based centrifugation. The remaining RNA pellet was air-dried and resuspended in 20-40µl of nuclease-free water. cDNA was prepared using the QuantiTect™ Reverse Transcription kit (QIAGEN™, QIAGEN™ Pty Ltd, Australia, Australia) as per manufacturer's instructions.

Real time quantitiative PCRs (QPCR) were performed using QuantiTect ™ SYBR green PCR kit (QIAGEN™, QIAGEN™ Pty Ltd, Australia, Australia). The design

and optimisation of *histone H4* QPCRs was performed by Dr. Petra Neufing and Ms. Carolyn Butcher. The cycling conditions for *histone H4* QPCR primers were as follows; 5 minutes at 98°C, followed by 25 cycles of 30 seconds at 68°C. Primer product melts were performed at 80°C to 90°C. The QPCR primer details for validation of Mouse Gene_ST microarray expression data have been listed in ***Supplementary Table S2.6***.

2.2.6 Genomic PCR amplification and Sanger Sequencing

Primers for the human *histone H4* genes were designed using the Primer3 Input (v.0.4.0) online software and optimised for specificity using the HEL *g*DNA (*histone H4* primer details and the electrophoresis images of the products generated have been listed in ***Supplementary Table S2.2***). *g*DNA samples were amplified using FastStart Taq DNA polymerase (Roche Applied Science, Roche Diagnostics Australia Pty. Ltd, Australia) and the cycling conditions for all the primer pairs were as follows: 5 minutes at 95°C; 35 cycles of 30 seconds at 94°C; 30 seconds at 59°C annealing temperature; 45 seconds at 72°C; a final extension at 72°C for 5 minutes. Following agarose gel electrophoresis, products were developed with Gel Red stain (Biotium, Jomar Biosciences, Australia) and visualised using BioRad UV Light & Gel Documentation system (Bio-Rad Laboratories Pty., Ltd., Australia). PCR products were purified using the QIAquick™ PCR purification kit (QIAGEN™ Pty Ltd, Australia, Australia) or ExoSAP-IT™ (GE Healthcare Life Sciences, GE Healthcare Australia Pty Ltd, Australia). Sanger sequencing was performed using BigDye® (Applied Biosystems**,** Life Technologies Australia Pty Ltd**,** Australia) terminator cycle sequencing protocol as described previously (Wen 2001). Sequence analysis and detection of variants was performed using SoftGenetics® Mutation Surveyor™ Version 3 and ChromasLite 2.1.1 software. The sequences were further analysed using the University of California Santa Cruz Genome Bioinformatics, Human

(*Homo sapiens*) Genome Browser Gateway (UCSC genome browser), Ensembl|1000 Genomes database (*hg18* and *hg19*), Catalogue of Somatic Mutations in Cancer [COSMIC, Wellcome Trust Sanger Institute, UK, (Bamford, *et al* 2004)] and the Exome Variant Server database [EVS; NHLBI Exome sequencing Project].

2.2.7 Flow cytometry and fluorophore conjugated antibodies

For flow cytometry cells were harvested, washed in phosphate buffered saline (PBS), resuspended in 100µl FACS wash [PBS containing 5% newborn calf serum and 0.2%Sodium Azide (w/v)] and incubated for 30 minutes in the dark at 4°C with fluorophore-conjugated primary antibody (see ***Supplementary Table S2.3*** for antibody details). Cells were washed in FACS wash, centrifuged at 250*g* for 5 minutes and the cell pellet resuspended in FACS FIX [PBS containing 2% glucose (w/v), 1% formaldehyde (v/v), 0.02% sodium azide]. Cells were analysed using the Becton Dickinson FACSCanto II flow cytometer (BD Biosciences, USA) and data analysis was performed using the FCS Express V3 software (De Novo software, USA).

2.2.8 Propidium Iodide (PI) staining of FDM cells

Cellular DNA content was determined by permeablizing and fixing cells in 70% ethanol for 16 hours at 4°C followed by staining with Propidium iodide solution (PI; Sigma-Aldrich Pty Ltd, Australia) supplemented with ribonuclease A (Sigma-Aldrich Pty Ltd, Australia) for 30 minutes at room temperature. These cells were analysed on Beckman Coulter Cytomics FC500 (Beckman Coulter, Miami, FL, USA) analyser.

2.2.9 Acid extraction of histones

Cells were washed twice with cold PBS for 5 minutes to remove culture media, followed by lysis in Triton Extraction Buffer [TEB; PBS containing 0.5% Triton X 100 (v/v), 2mM phenylmethylsulfonyl fluoride (PMSF), 0.02% (w/v) sodium azide (NaN_3), at a concentration of 10^7 cells /mL]. Histones were acid extracted overnight by incubation in 0.2N HCl, followed by centrifugation at 500*g* for 10 minutes at 4°C to remove debris. The supernatant containing the histones was aliquoted and stored at -80°C for further use.

2.2.10 Western blot and antibodies

Histone fractions were separated by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) and transferred to nitrocellulose membrane. Specific histones and post-translationally modified forms of histones were detected by western blotting with mouse- or rabbit-derived primary antibodies. Rabbit polyclonal anti-histone H4 total (ab10158), anti-histone H4 monomethyl R3 (ab17339), anti-histone H4 dimethyl R3 [asymmetric, 39705, Active Motif], anti-histone H3 total (ab1791), anti-histone H2A.X (ab11175) and mouse monoclonal anti-histone H2A.X gamma (phosphor-serine139) (ab18311) were purchased from Abcam (Sapphire Bioscience Pty Ltd, Australia). Incubation with primary antibodies was for 16 hours at 4°C. Anti-histone H3 antibody was incubated for 1 hour at room temperature. Goat anti-rabbit/mouse Horse Radish Peroxidase (HRP)-conjugated antibody (Thermo Fisher Scientific Australia Pty Ltd, Australia) was used as secondary antibody and detected with Super Signal West Pico or Super Signal West Dura substrates (Thermo Fisher Scientific Australia Pty Ltd, Australia) using the Luminescent Image Analyser (LAS-4000, Fujifilm, Australia). Quantitation of protein bands on the blots was performed using the Multi Guage Ver 3.0 software (Fujifilm, Australia).

2.2.11 Statistics

Student's unpaired *t*-test and one-way ANOVA were (GraphPad Prism version 6.01) used to determine significance of values. Microarray Data analysis, Linear modelling of microarray (LIMMA) and generation of NUSE plots were performed by Dr. Chung Hoow Kok.

2.3 Results

2.3.1 Identification of novel histone H4 variants in patients with Polycythemia vera

The 15 *histone H4* genes were amplified from 88 PV patient PBMNC or granulocyte *g*DNA samples as single amplicons using specific forward and reverse primers. For specificity, these primers were designed to match intergenic regions (see ***Supplementary Table S2.2***). All 15 *histone H4* genes were successfully amplified from the patient *g*DNA samples [the complete list of coding region alterations identified in PV has been listed in ***Table 2.1***]. An additional 22 PV, 15 ET, and 3 PMF PBMNC or granulocyte patient *g*DNA samples were amplified for *HIST1H4C*. No coding region variants were found in *HIST1H4C* in the small number of ET and PMF samples screened. For all variants the Human Genome Browser (University of California Santa Cruz Genome Bioinformatics, Human (*Homo sapiens*) Genome Browser Gateway or UCSC genome browser), Ensembl|1000 Genomes (*hg18* and *hg19*) and the Exome Variant Server (EVS) databases were searched to separate previously unidentified or novel variants from known and rare population variants. Several 5`UTR, 3`UTR (***Supplementary Table S2.4***) and synonymous coding region variants were identified in *HIST1H4B, HIST1H4D, HIST1H4E, HIST1H4K* and *HIST2H4*. The focus of further studies was non-synonymous coding region variants identified in *HIST1H4B* and *HIST1H4C*.

Gene	Chr:bp	Alleles (Reference SNP ID)	Coding Region Variant	Variant type	Freq. in PV (%)	Pop^{n}. Freq (%)	Source
HIST1H4A	6:26021924	G/A	c. 27G>A, p.K8K	synonymous	1.136	ND	NA
HIST1H4B	6:26027399	A/G (rs201003635)	c.82G>T, p.Q27*	stop gained	1.136	0.700	EVS
HIST1H4C	6:26104185	C/T	c.10C>T, p.R4C	**missense**	0.980	ND	NA
	6:26104342	G/A	c.167G>A, p.R56Q	**missense**	0.980	ND	NA
HIST1H4E	6:26205178	C/A (rs116339316)	c.306C>A, p.G102G	synonymous	1.204	0.730 & 0.900	dbSNP & EVS
	6:26205147	G/A	c.276G>A ,p.K92K	synonymous	1.204	ND	NA
	6:26204884	C/T (rs139616312)	c.12C>T, p.R4R	synonymous	1.204	ND	dbSNP & EVS
HIST1H4H	6:26285425	G/A	c.303C>T, p.F101F	synonymous	1.200	ND	NA
	6:26285428	G/A (rs143923288)	c.300C>T, p.G100G	synonymous	1.200	0.090	dbSNP
	6:27792100	G/A	c.198G>A, p.V66V	synonymous	1.136	ND	NA

Table 2.1. Summary of histone H4 coding region variants identified in PV. 15 *Histone H4* genes were screened in 88 PV patient samples (102 were screened for *HIST1H4C*). Ensembl|1000 genomes, dbSNP (build 135, 137 and 138) and the Exome Variant Server (EVS) databases were used to identify novel variants from known variants. Freq.,frequency in the PV cohort; Pop^{n}. freq., population frequency ascertained from databases; ND, Not Detected; NA, Not Applicable.

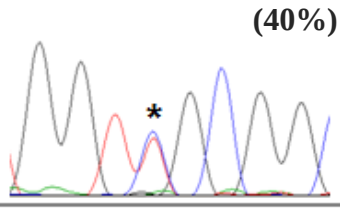
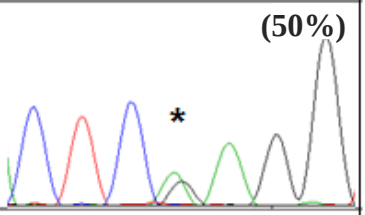
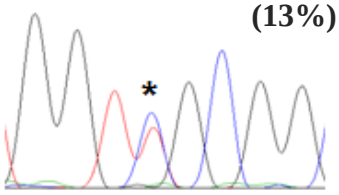
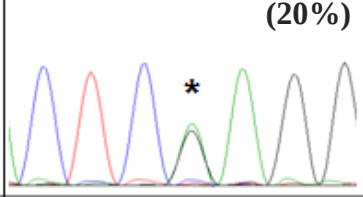
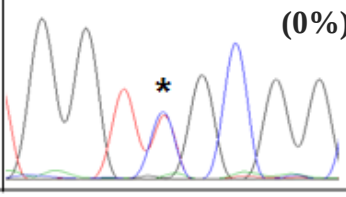
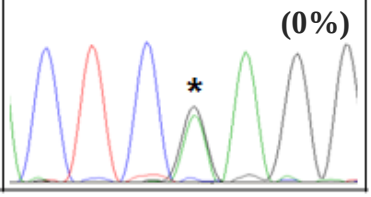
A heterozygous protein-truncating variant was identified in *HIST1H4B* [*HIST1H4B:c.82C>T, p.Q28**] in the patient PV103, (Male, 84 years at diagnosis) (***Supplementary Figure S2.3***). As shown in the figure, this variant was detected in the germline *g*DNA sample from the patient and has been found in EVS, at a frequency of 0.7% in the North American population.

Two novel heterozygous *HIST1H4C* variants were identified in granulocytes isolated from two PV patients: a *HIST1H4C:c.10C>T, p.R4C* (referred as p.R4C) detected in patient PV109 and *HIST1H4C:c.167G>A, p.R56Q* (referred as p.R56Q) detected in patient PV115 (***see section 2.3.3***). p.R4C and p.R56Q affect the R3 and R55 residues on the mature histone H4 protein (as the methionine is removed after translation). The R3 residue on the histone H4 N-terminal tail is a critical residue as it is methylated by a number methyltransferases such as PRMT1 (Huang, *et al* 2005; Strahl, *et al* 2001; Wang, *et al* 2001), PRMT5 (Zhao, *et al* 2009a), PRMT7 (Lee, *et al* 2005), or modified to citrulline by the peptidyl arginine deiminase; PAD4 (Wang, *et al* 2004). Thus, this novel coding variant was of particular interest and was investigated further (particularly in ***section 2.3.8***).

2.3.2 Characterisation of the HIST1H4C: p.R4C and p.R56Q variants detected in two patients

To determine somatic or germline status of the variants, buccal *g*DNA samples from each patient were screened by PCR amplification and Sanger sequencing to determine whether the changes were germline or somatic. ***Figure 2.3*** shows patient characteristics and sequence traces for the changes in buccal, PBMNC and granulocyte *g*DNA samples confirming germline transmission. p.R4 and p.R56 are conserved across a number of eukaryotes (***Figure 2.4***) and were predicted to be *damaging* using the web based polymorphism phenotyping tool:

A.

Sample type	PV109 G G T C G C G G	PV115 C T C G A G G
Granulocyte *g*DNA (*JAK2V617F*%)		
PBMNC *g*DNA (*JAK2V617F*%)		
Buccal *g*DNA (*JAK2V617F*%)		

B.

Patient characteristics	PV109	PV115	Normal Reference range
Gender/Age at diagnosis	M/37	M/72	NA
JAK2V617F mutation	positive	positive	NA
Haemoglobin	139g/L	179g/L	135-175g/L
Red Blood Cell	7.15 x10^{12}/L	6.26 x10^{12}/L	4.50-6.00 x10^{12}/L
White Cell	10.8 x10^{9}/L	8.95 x10^{9}/L	4.00-11.0 x10^{9}/L
Serum EPO	<1 mU/ml	2 mU/ml	3.3-16.6 mU/ml
EPO-independent BFUE (e-BFUE)	detected	detected	NA

Figure 2.3. Clinical characteristics of Polycythemia vera patients PV109 and PV115. A. Sequence traces of *HIST1H4C* in the two patients depicting the germline variants; *HIST1H4C:c.10C>T, p.R4C* and *HIST1H4C:c.167G>A, p.R56Q*. The underlined nucleotides depict WT sequence. The sequence trace in patient PV109 and PV115 shows the nucleotide affected (as depicted by *) in granulocyte, PBMNC and buccal *g*DNA sample for *HIST1H4C* (% indicates the *JAK2V617F* allele load). **B.** Patient characteristics. Samples were collected at the time of diagnosis and the value for each patient characteristic is from the year of diagnosis. The values in the normal reference range column were used at the time of diagnosis. NA, not applicable.

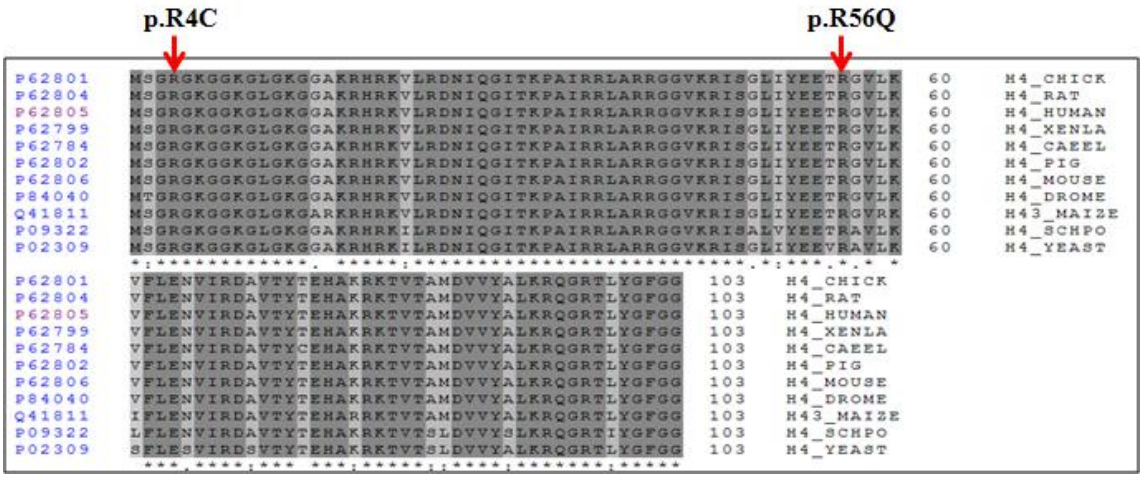


Figure 2.4. Histone H4 amino acid conservation across eukaryotes. Affected residues have been indicated with red arrows. Grey shading of residues, across species, indicates to degree of conservation (Image aligned and captured using UniProt bioinformatics tool). *Histone H4* genes are highly conserved across species with a reported ~90% evolutionary conservation (Huang, *et al* 2009).

PolyPhen-2 [PolyPhen (Adzhubei, *et al* 2010)]. Subsequently, Sanger sequencing of *g*DNA samples from a further 82 MPN patient *g*DNA samples [22 PV, and 60 MPN (a combination of ET, MF and PV which were generously provided by Prof. Andrew Roberts from WEHI)]. These revealed no other non-synonymous coding region variants.

2.3.3 Frequency of HIST1H4C p.R4C substitution in DNA samples from control populations

As shown in ***Table 2.2***, a number of other somatic coding region variants in cancers have been identified in *HIST1H4C*. In addition variants affecting the p.R4 have been identified in two different *histone H4* genes; *HIST1H4B: c.10C>T, p.R4C* and *HIST1H4F:c.11G>A, p.R4K*. To establish the frequency of *HIST1H4C* variants in the population, sequencing from 236 buccal *g*DNA samples (included 109 females and 114 males, age range 19-73 years, mean age of 59 years) was performed. Sequence analysis from this small cohort did not show coding variants in *HIST1H4C.* However additional sequencing of the same identified a number of germline 5`- and 3`-UTR variants in several of the *histone H4* genes. ***Supplementary Table S2.5*** lists the *histone H4* genes identified with UTR changes. Sequencing of a further 804 diagnostic *gDNA* samples (available through Molecular Pathology, IMVS, *CNAHS ethics# 2007159*, see *section 2.2.1*) identified two additional heterozygous non-synonymous coding region variants in *HIST1H4C* (*HIST1H4C: c.166C>T, p.R56** and *HIST1H4C:c.268G>T, p.A90S*) in two individual *g*DNA samples, giving a frequency of 0.248% in this cohort. Interestingly, *HIST1H4C:c.268G>T, p.A90S* has now been assigned with a SNP identifier (rs144469714, approximate population frequency of 0.03% for the minor allele as reported in UCSC genome browser's dbSNP build 138). While these results allowed an assessment of frequency for *HIST1H4C* variants in an independently acquired set of samples, this analysis cannot

Gene	**Table 2.2 Somatic non-synonymous coding region variants of histone H4**					
HIST1H4A	c.182T>C, p.V61A	c.271C>T, p.L91F	c.292C>T, p.L98F			
HIST1H4B	**c.10C>T, p.R4C**	c.74A>T, p.D25V	c.77A>T, p.N26I	c.82C>G, p.Q28E	c.91A>T, p.T31S	c.110G>T, p.R37L
	c.119G>T, p.R40M	c.120G>T, p.R40S	c.137G>A, p.R46Q	c.162G>C, p.E54D	c.172G>A, p.V58I	c.175C>T, p.L59F
	c.181G>A, p.V61M	c.202C>T, p.R68W	c.208G>T, p.A70S	c.215C>T, p.T72I	c.229G>A, p.A77T	c.240G>C, p.K80N
	c.256G>A, p.D86N	c.256G>T, p.D86Y				
HIST1H4C	c.4T>G, p.S2A	c.106C>A, p.R36S	c.115G>T, p.A39S	c.116C>T, p.A39V	c.124G>C, p.G42R	c.203G>C, p.R68P
	c.208G>A, p.A70T	c.242C>T, p.T81I				
HIST1H4D	c.26A>G, p.K9R	c.29G>A, p.G10D	c.43G>T, p.G15C	c.77A>G, p.N26S	c.90C>G, p.I30M	c.92C>T, p.T31I
	c.97C>A, p.P33T	c.100G>A, p.A34T	c.179A>T, p.K60I	c.182T>G, p.V61G	c.205G>T, p.D69Y	c.221C>T, p.T74M
	c.233A>T, p.K78I	c.239A>G, p.K80R	c.247A>G, p.T83A	c.250G>A, p.A84T	c.295T>C, p.Y99H	c.310T>C, p.*104R
	c.223G>A, p.E75K	c.178A>T, p.K60*				
HIST1H4E	c.13G>T, p.G5C	c.14G>A, p.G5D	c.29G>C, p.G10A	c.58C>A, p.R20S	c.77A>T, p.N26I	c.148C>T, p.L50F
	c.159G>T, p.E53D	c.162G>C, p.E54D	c.197T>C, p.V66A	c.205_206GA>TT, p.D69F	c.236G>A, p.R79H	c.240G>C, p.K80N
	c.276G>C, p.K92N	c.278G>C, p.R93T	c.290C>T, p.T97I			
HIST1H4F	**c.11G>A, p.R4K**	c.170G>A, p.G57D	c.211G>A, p.V71I	c.277C>T, p.R93C	c.286C>T, p.R96C	c.297C>G, p.Y99*
	c.299G>A, p.G100D					
HIST1H4G	c.34G>C, p.G12R	c.43G>A, p.G15S	c.68T>C, p.L23P	c.106C>T, p.R36W	c.141C>G, p.I47M	c.200T>G, p.I67S
	c.205T>C, p.Y69H	c.234G>T, p.K78N	c.248C>T, p.T83I	c.250G>T, p.A84S	c.259G>T, p.V87L	c.278G>A, p.R93H
	c.128G>A, p.G43D	c.129C>T, p.G43G	c.190G>C, p.E64Q	c.250G>A, p.A84T	c.311G>T, p.*104L	

Continued on the next page...

HIST1H4H	c.5C>G, p.S2C	c.16A>G, p.K6E	c.47C>T, p.A16V	c.50A>G, p.K17R	c.63G>C, p.K21N	c.91A>G, p.T31A
	c.118C>G, p.R40G	c.244G>C, p.V82L	c.262G>T, p.V88F	c.292C>T, p.L98F		
HIST1H4J	c.5C>T, p.S2F	c.28G>C, p.G10R	c.128G>T, p.G43V	c.103A>G, p.I35V	c.220A>G, p.T74A	c.223G>A, p.E75K
	c.235C>T, p.R79C	c.286C>T, p.R96C	c.287G>T, p.R96L	c.299G>A, p.G100D		
HIST1H4K	c.3G>C, p.M1I	c.5C>G, p.S2C	c.8G>C, p.G3A	c.23G>T, p.G8V	c.286C>T, p.R96C	c.304G>T, p.G102C
	c.248C>A, p.T83N	c.268G>A, p.A90T	c.311G>C, p.*104S			
HIST1H4L	c.1A>G, p.M1V	c.3G>T, p.M1I	c.38A>T, p.K13I	c.73G>A, p.D25N	c.83A>G, p.Q28R	c.107G>A, p.R36Q
	c.5C>G, p.S2C	c.32T>C, p.L11P	c.58C>T, p.R20C	c.73G>T, p.D25Y	c.74A>G, p.D25G	c.115G>T, p.A39S
	c.119G>T, p.R40L	c.128G>A, p.G43D	c.184T>C, p.F62L	c.191A>G, p.E64G	c.190G>C, p.E64Q	c.250G>A, p.A84T
	c.311G>T, p.*104L	c.255G>A, p.M85I				
HIST4H4	c.109C>T, p.R37C	c.119G>A, p.R40Q	c.119G>T, p.R40L	c.283G>A, p.G95S		

Table 2.2. Somatic non-synonymous coding region variants of histone H4. The COSMIC database was used to identify all variants in the 15 *histone H4* genes. Non-synonymous coding region variants in *Histone H4* genes; *HIST1H4I* and *HIST2H4A/B* were not found in the COSMIC database Variants highlighted in bold depict the identical amino acid residues affected in *HIST1H4C* in the PV cohort. The data in table was updated in April 2014.

be considered as being representative of a healthy population. This cohort consisted of *g*DNA samples for genetic lesion testing and information was limited to the gender and age of the sample source, thus the significance of the identified variants was difficult to interpret. However, given the identification of novel *HIST1H4C* variants, the possibility of the identified variants in MPN being rare SNPs cannot be excluded. Moreover, the presence of variants present at analogous positions in *HIST1H4F* [*HIST1H4F:c.11G>T, p.R4I (rs370419237)* and *HIST1H4F:c.166C>T, p.R56C (rs146739153)*] and *HIST1H4H* (*HIST1H4H:c.166C>T, p.R56C, rs367980184*) further suggests that the *p.R4C* and *p.R56Q* variants may be normal sequence variants. Thus, to further understand the relevance of these variants, the expression of *HIST1H4C* was assessed in haematopoietic cell populations.

2.3.4 HIST1H4C is a major contributor of the histone H4 mRNA in myeloid leukaemic cell lines

Given that there are 15 *histone H4* genes, 14 encoding for the identical protein and a number of variants associated in disease for most of the *histone H4* genes, it was important to determine the relative contribution of the *HIST1H4C* gene to the *histone H4* mRNA pool in haematopoietic populations. A significant role for this gene in haematopoietic cells has been suggested by a previous report of targeting in the K562 cell line to prevent expression of *HIST1H4C* using DNA analogs which led to inhibition of proliferation and increased apoptosis (Minoshima, *et al* 2010). Analysis of expression using real time QPCR for all the *histone H4* genes was therefore performed on cDNA isolated from K-562 and HEL haematopoietic cell lines and from the non-haematopoietic HEK293 cell line. The relative contribution of each *histone H4* gene was determined by normalising expression to the reference gene *GAPDH*, and converted to a percentage. For this, each expression value was divided by the total of the additive mean expression values for all the 15 *histone H4* genes.

These values were then used to calculate a percentage value (±SD) for each gene. The relative expression of *HIST1H4C* was the highest in all cell lines followed by *HIST1H4E, HIST2H4A/B* and *HIST1H4H* respectively (***Figure 2.5***). Evaluation of *HIST1H4C* expression values following extraction of the normalised expression data from the study of Greshock and colleagues (Greshock, *et al* 2010). ***Figure 2.6*** also identified *HIST1H4C* as a major contributor to the *histone H4* mRNA pool in a further 15 leukaemic cell lines (Chung Hoow Kok, unpublished).

2.3.5 HIST1H4C is expressed at significant levels in normal bone marrow CD34$^+$ cells and peripheral blood CD14$^+$ cells

NOTE:
This thesis section 2.3.5 has been removed to comply with copyright regulations.
It is included in the print copy of the thesis held by the University of Adelaide Library.

Figure 2.5 Expression of *HIST1H4C* was significantly higher compared to the other *histone H4* genes in transformed human cell lines. Two-step QPCR using long primers was used to differentiate between transcripts and to measure relative expression levels of the *Histone H4* genes. The figures depict percentage expression of total *Histone H4* mRNA contributed by each of the 15 *Histone H4* genes in: **A.** K-562, **B.** HEL and **C.** HEK293 cell lines. All PCRs were performed in triplicate (n=3). Student's unpaired two-tailed t-test was performed to assess significance. Values are the mean ±SD.

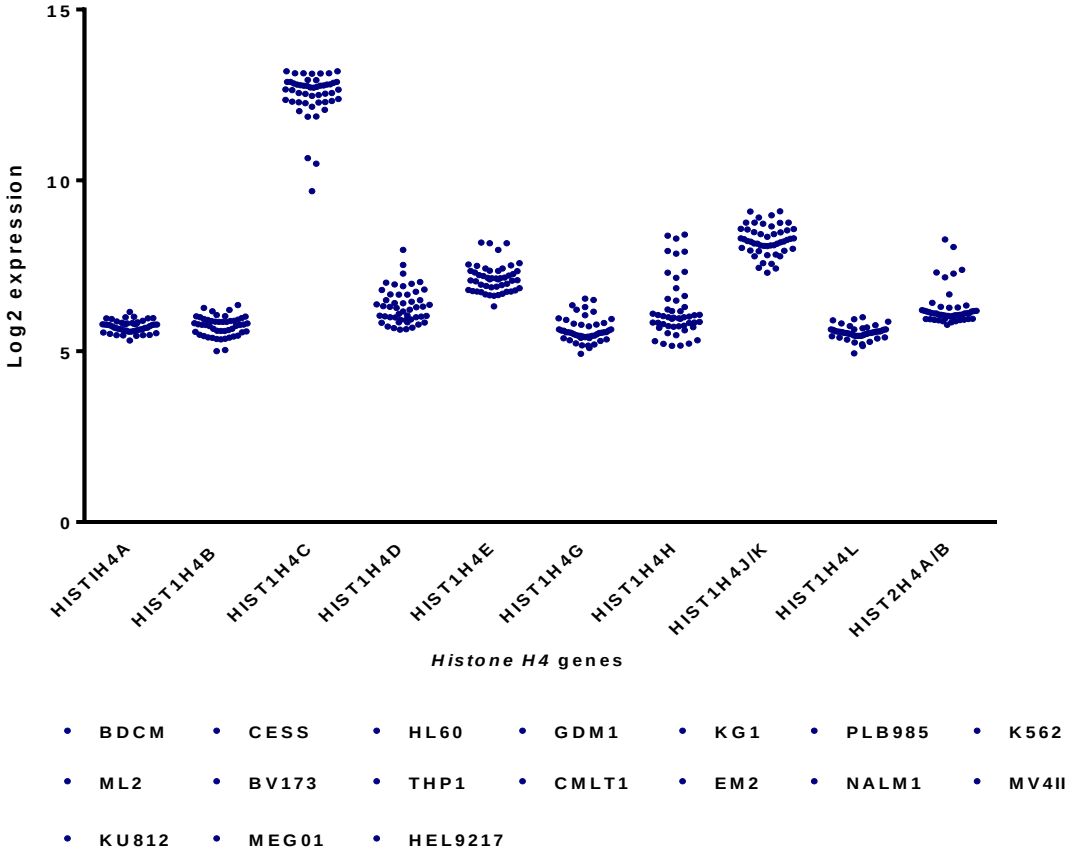


Figure 2.6. **Analysis of microarray expression data of myeloid leukaemia cell lines.** Normalised mRNA expression for *Histone H4* genes shows that *HIST1H4C* is highly expressed in myeloid leukaemic cell types. The extracted data lacked values for *Histone H4* genes, *HIST1H4F*, *HIST1H4I* and *HIST4H4*. Data from Greshock and colleagues. (Greshock, *et al* 2010) study was acquired from the GlaxoSmithKline Cancer Cell Line genomic Profiling Data set (National Cancer Institute's cancer Bioinformatics Grid). Acute leukaemia cell lines and morphologies were as follows- BDCM, B-lymphoblast; CESS, lymphoblast; HL60, promyleoblast; GDM1, myelomonoblast; KG1, macrophage/myeloblast; PLB985, myeloblast; ML2, monocytic; THP1, monocytic; NALM1, lymphoblast; MV411, lymphoblast; KU812, myeloblast; MEG01, megakaryoblast. Chronic myeloid leukaemia cell lines and morphologies were as follows- K562, *Ph+* erythromyeloblast; BV173, *Ph+* myeloblast;, CMLT1, *Ph+* lymphoblast; EM2, *Ph+* myeloblast (Mahon, Deininger et al. 2000); HEL9217, erythroblast. *(Information about cell lines were taken from ATCC unless cited)*

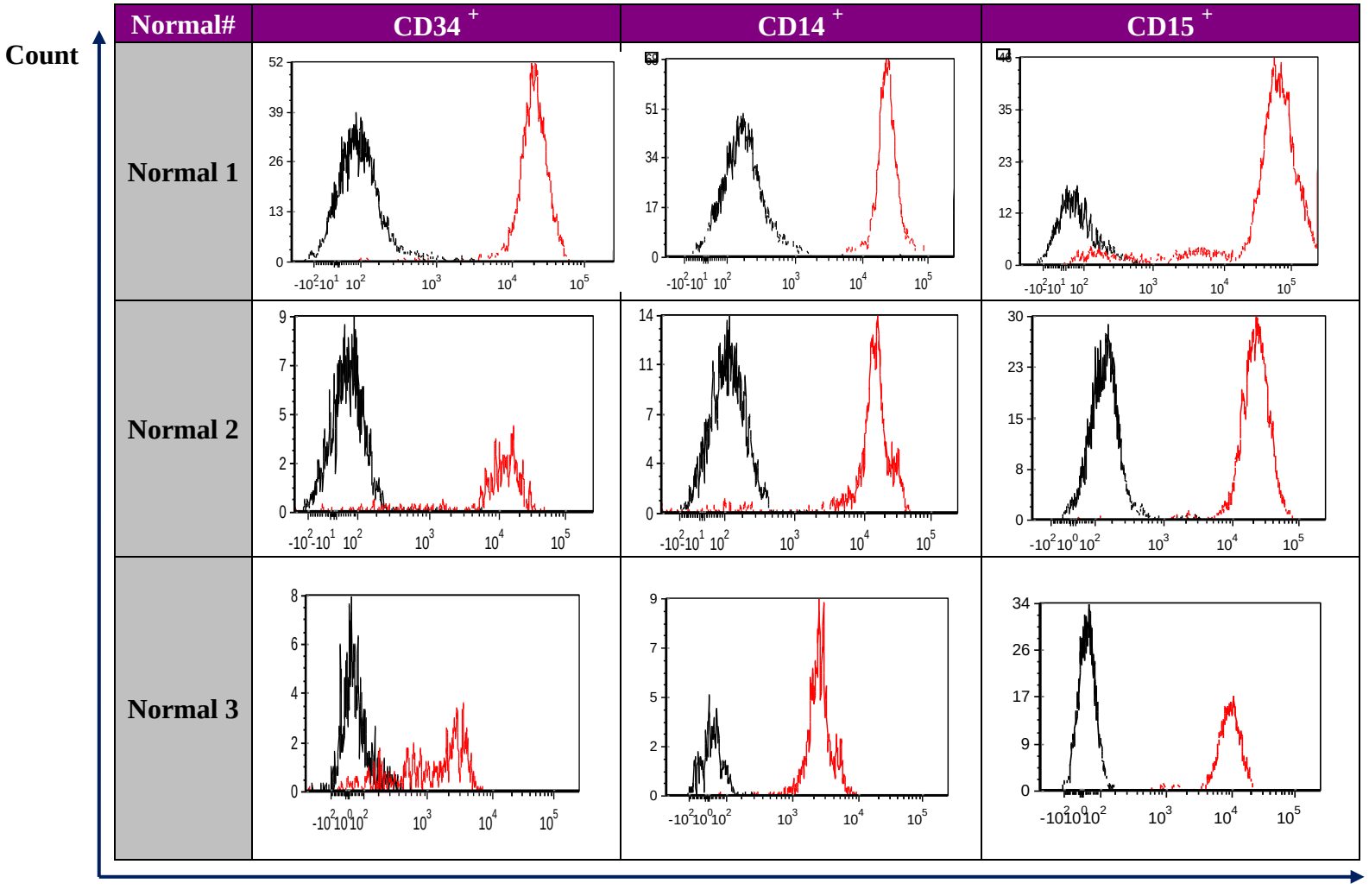


Figure 2.7 Purity of MACS enriched haematopoietic sub-populations assessed by flow cytometry. The isotype control (black histogram) and positive samples (red histograms) for $CD34^{+}$ BMMNC; $CD14^{+}$ (monocyte) enriched fraction from PBMNC and $CD15^{+}$ (granulocyte) enriched fraction from three independent normal samples. Also see ***Supplementary Figure S2.4*** for flow cytometric analysis of CD14 negative fractions.

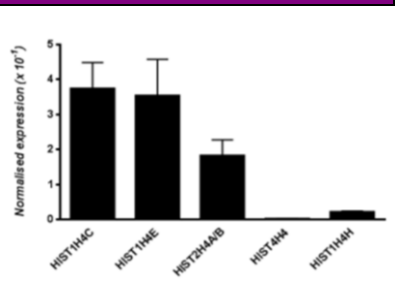
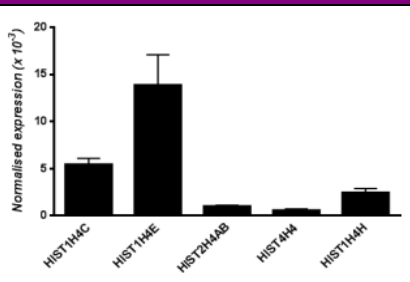
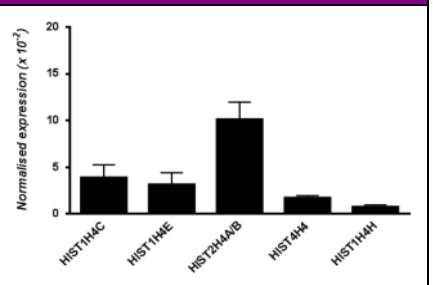
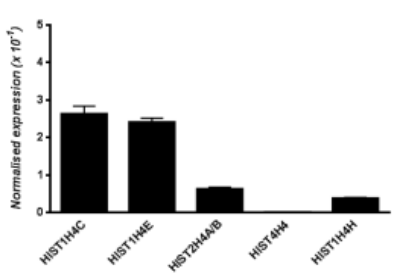
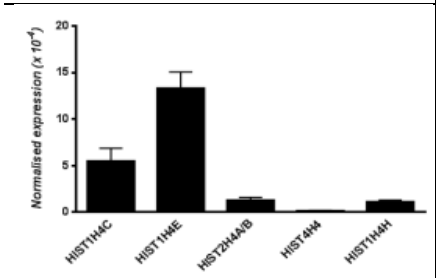
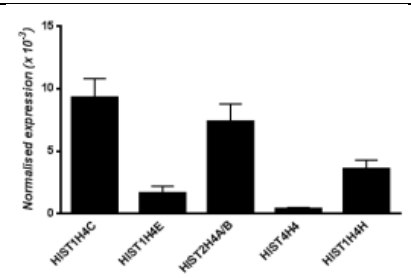
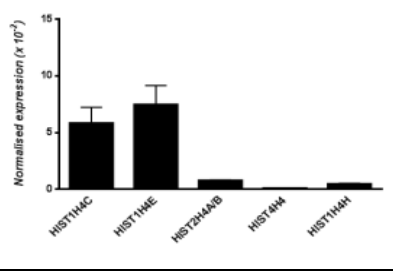
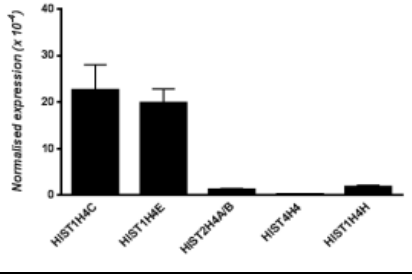
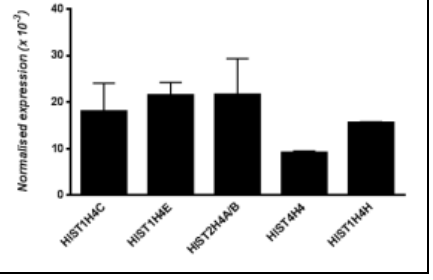
Normal#	HSPC enriched (CD34^{+})	Monocyte enriched (CD14^{+})	Granulocyte enriched (CD15^{+})
Normal 1			
Normal 2			
Normal 3			

Figure 2.8. The expression of *HIST1H4C* in normal human haematopoietic HSPC, monocyte and granulocyte enriched cell fractions. On comparison to the expression of *HIST2H4A/B*, *HIST4H4* and *HIST1H4H*, the expression of *HIST1H4C* and *HIST1H4E* was relatively higher in three independently extracted cell fractions (*n*=3). CD15+ cell fraction showed *HIST2H4A/B* as the predominantly expressed gene in 2 of the three samples other than *HIST1H4C*. Student's unpaired *t*-test was performed. Values are the mean ±SEM.

2.3.6 Generation of an inducible expression system for the functional analysis of p.R4C

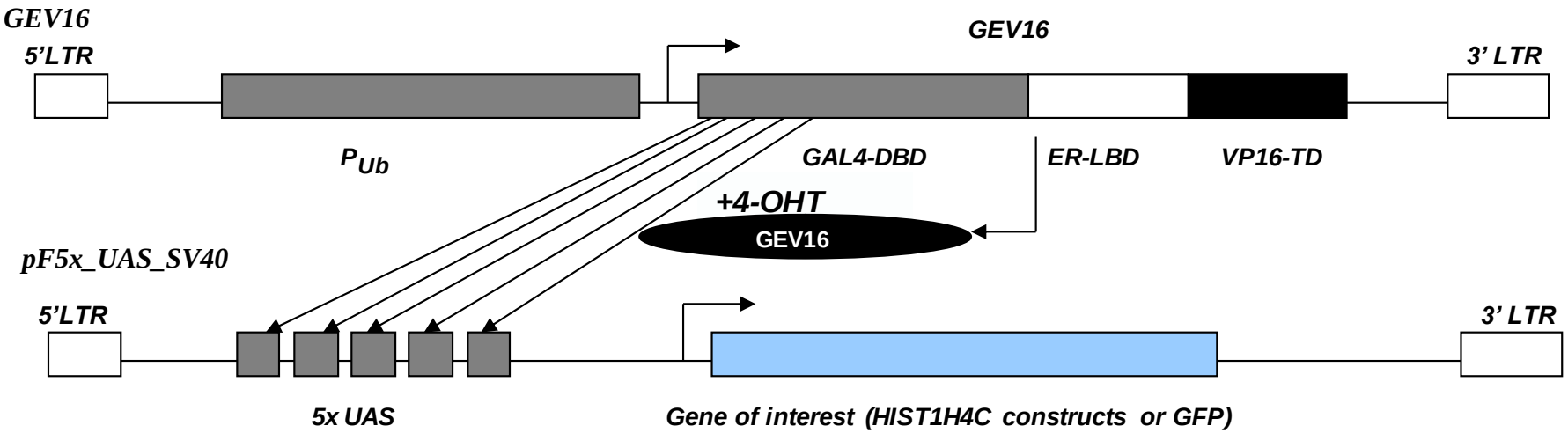
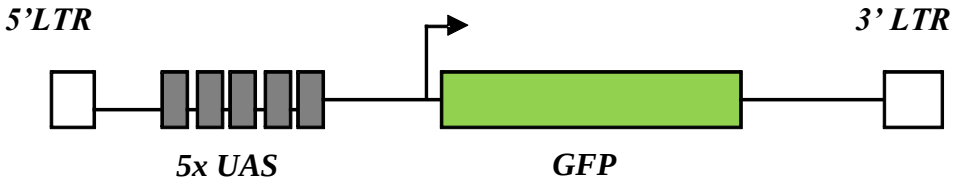


Figure 2.9. Schematic representation of the dual vector lentiviral system in mammalian cells. After stable integration in the host genome, GAL4-Estrogen Receptor-VP16 (GEV16) fusion protein is produced constitutively and is present in the cytoplasm in abundance. On addition of 4-hydroxy tamoxifen (4-OHT) to the cells, 4-OHT binds to GEV16 via the estrogen receptor part of the GEV16 fusion protein and the 4-OHT bound GEV16 fusion protein is translocated into the nucleus. In the nucleus, the GAL4 domain of GEV16 binds to the *U*pstream *A*ctivation *S*equences (*5xUAS*) of the expression vector to up-regulate the transcription of the downstream construct (s); *HIST1H4C* or *GFP* construct. The figure has been adapted from (Callus, *et al* 2008). GAL4-DBD, GAL4-DNA Binding Domain; ER-LBD, Estrogen Receptor-Ligand Binding Domain; VP16-TD, VP16- Transcriptional activation Domain; LTR, Long Terminating Repeats.

A.



B.

Time	0h	4h	6h	8h	14h	24h
(%GFP$^+$)	15.5%	52.1%	52.2%	45.5%	80.0%	76.8%

4-OHT	0nM	50nM	100nM	200nM	300nM	400nM
(%GFP$^+$)	12.1%	45.6%	45.2%	67.5%	74.0%	82.0%

C.

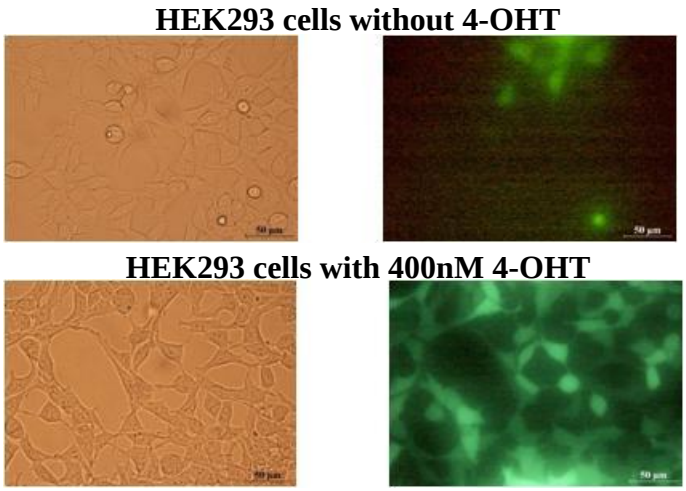


Figure 2.10. GFP constructs expression in HEK293 was optimal following incubation with 400nM of 4-OHT for 24 hours. A., Schematic of the integrated *pF5x_UAS_SV40* containing the *GFP* construct; **B.**, The several incubation times and dose of 4-OHT used to determine the optimum induction of *GFP* expression. Optimum time was determined using 400nM of 4-OHT and dose was determined after 24 hours. % indicates to percentage GFP$^+$ cells. **C**. GFP$^+$ cells following induction with 400nM of 4-OHT after 24 hours depicted using fluorescence microscopy (at 40x magnification) using the Nikon Eclipse TE2000-U (Nikon, Coherent Scientific Pty Ltd). Image was analysed using the Nikon NIS Elements version 2.3.4 (Nikon) software. h, hours; LTR, Long Terminating Repeats.

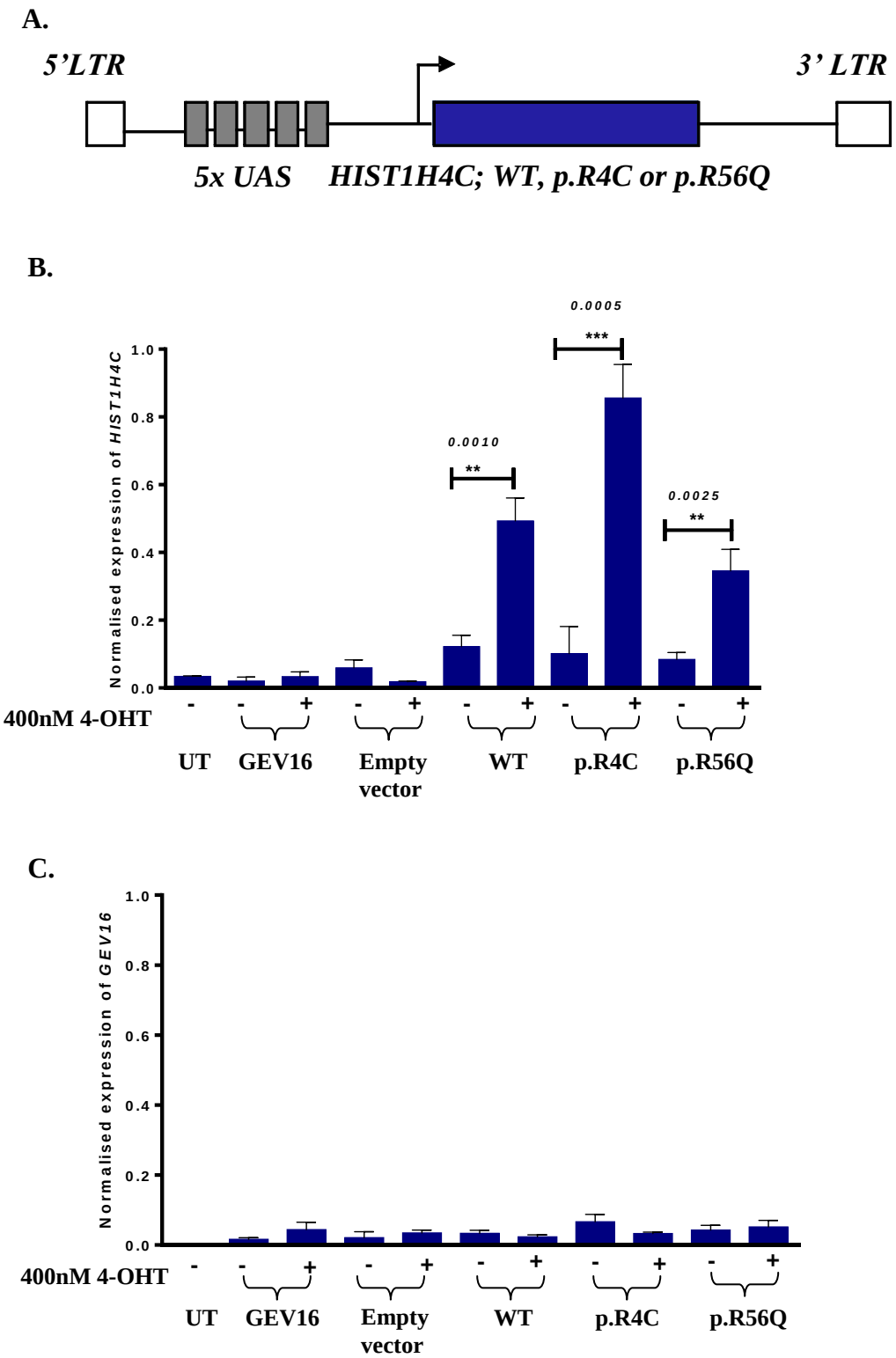


Figure 2.11. HEK293 cells expressing *HIST1H4C* constructs were observed with a minimum two-fold induction. A. Schematic of the *PF5XUAS SV40-HIST1H4C* construct; **B.** mRNA expression of *HIST1H4C* normalised to *GAPDH*. The endogenous expression of *HIST1H4C* in the untransduced HEK293 cells is comparable to the empty vectors only expressing cells; GEV16 and empty vector. HEK293 cells with empty vector co-expressed empty *p_F5xUAS_SV40* and *GEV16*. All the *HIST1H4C* constructs also expressed *GEV16* construct. **C.**, Depicts the expression of *GEV16* in the transduced HEK293 cells. Predictably, *GEV16* was not expressed in untransduced HEK293 cells. UT, untransduced. Values are represented as normalised expression values (mean±SD). Values above the horizontal line represent *p*-value between the two depicted groups. LTR, Long Terminating Repeats

2.3.7 Analysis of histone H4 following induction of HIST1H4C variants in HEK293 inducible cells

To identify alterations to histone H4, histone H4 levels and histone post-translational modifications (H4R3me^2 and histone H4K20me^3) western blotting was performed on histones extracted from the different HEK293 cell lines (***Figure 2.12 A, B & C)***. Phosphorylation of serine at position 139 on histone H2A.X was also measured, as there is evidence that excess core histones including histone H4 can cause genomic instability (for example DNA double strand breaks) which can be indicated by increased formation of gamma H2A.X foci [(Herrero, *et al* 2011; Liang, *et al* 2012) and reviewed in (Podhorecka, *et al* 2010)]. Alterations to the level of histone H4K20me^3 was of particular interest, as H4K20me^3 has been reported to be a repressive mark in gene silencing (Schotta, *et al* 2004) and has also been shown to be down-regulated in cancers (Fraga, *et al* 2005). Western blotting showed consistently higher levels of histone H4 in the cells containing WT and p.R4C histone H4 constructs when untreated with 4-OHT, although this increase did not reach significance when compared to the empty vector. This increased histone H4 level was not observed in the presence of 4-OHT, suggesting that induced levels of histone H4 may not be tolerated. Furthermore, it is also possible that 4-OHT has an independent effect on the chromatin as has been observed in the studies performed by Liu and colleagues (Liu, *et al* 2004). Levels of histone modification or gamma H2AX were compared within the 4-OHT untreated and treated groups, however this did not reveal consistent changes between samples. Furthermore, histone H4 levels in the p.R56Q variant expressing cells remained largely unchanged before and after treatment with 4-OHT. Given the variation in levels of the modifications observed

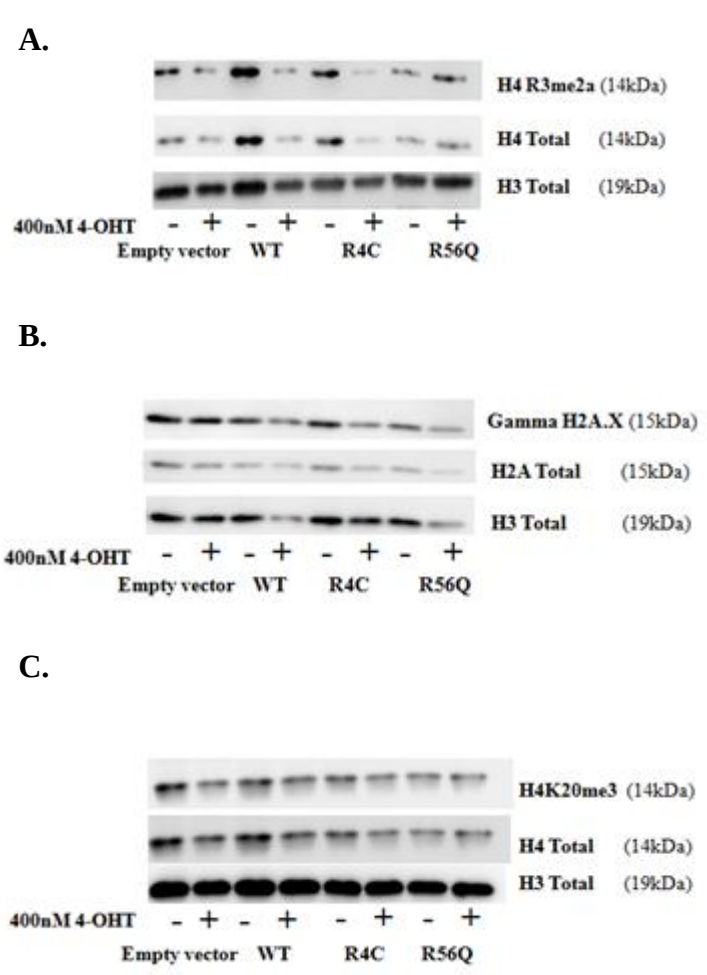


Figure 2.12. Western blotting analysis of the level of histone H4 post-translational modifications. (A) Histone H4R3me2a (asymmetric di-methylation), **(B)** Gamma H2A.X (phospho-serine 139) and, **(C)** histone H4K20me3. Modifications were assessed by western blotting following independent inductions (*n*=3) and quantitation is shown in **D-F**. There was no significant change detected in the level of these modifications with the mutant constructs when compared to empty vector or wildtype constructs. A-C show representative western blots for each modification and D-F shows the quantitated results (data is representative of three independent inductions, *n*=3). *p*-value <0.05, unpaired two-tailed students's *t*-test, values are plotted as mean±SD. Values above the horizontal line represent *p*-value between the two depicted groups.

between experiments, it is likely that alteration of these markers is too subtle and difficult to detect using this system and technique.

2.3.8 Changes in gene expression following over-expression of HIST1H4C:p.R4C in FDM cells

To further investigate subtle changes induced by the expression of *HIST1H4C:p.R4C* variant, global gene expression analysis was performed. As the interest was in assessing changes in the myeloid progenitor cell compartment, studies were performed in the IL3-dependent foetal mouse *HOXB8*-transformed myeloid cell line [FDM; (Ekert, *et al* 2004)].

Given that histone H4 encoded by the human *HIST1H4C* gene is identical to the mouse *Hist1h4c*, human *HIST1H4C* cDNA constructs (*p.R4C* or *WT*) were cloned into the *MSCV-IRES-GFP* plasmid (MIG). HEK293T cells were transfected and viral supernatant was harvested at 48 hours after transfection. FDM cells were transduced and cultured for 17 days before harvesting the cells for flow cytometry and isolation of RNA. The quality of the RNA, expression of *HIST1H4C*, and expression of GFP were assessed before proceeding to microarray analysis (***Figures 2.13 A, B and C***). Cell cycle analysis of FDM cells was also performed and this did not show any significant change ($p<0.05$) between the groups (***Supplementary Figure S2.5B***).

The Mouse Gene 1.0_ST Array platform (Mouse Gene_1_0-st.v1; Affymetrix) was chosen for the analysis of gene expression in four biological replicates (microarray was performed by the Adelaide Microarray Centre). The raw data (CEL format) was converted to a readable format, assessed for quality (***Supplementary Figure S2.5A***) and further processed by Dr. Chung Kok. The data was exported to an Excel™ file

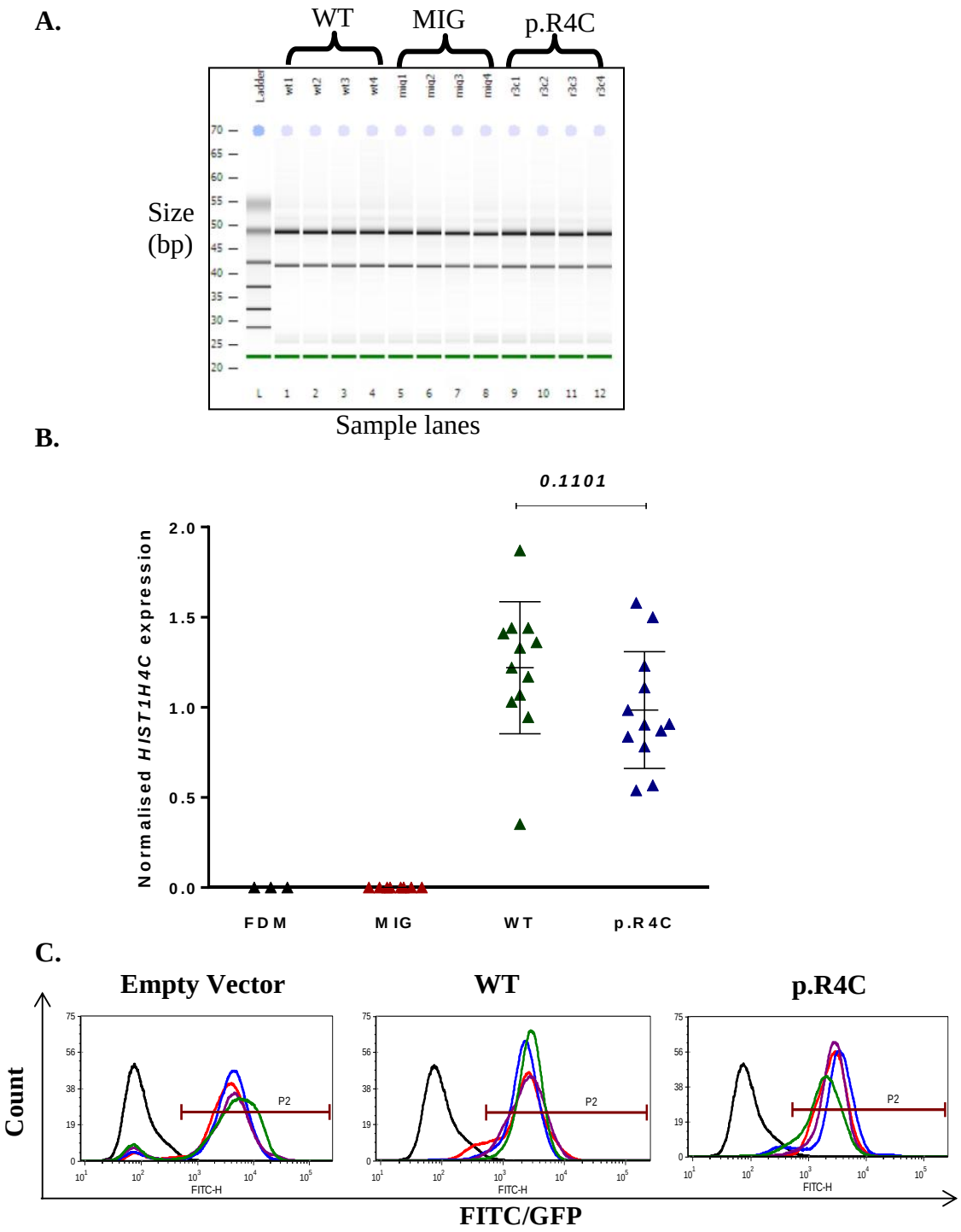


Figure 2.13. FDM cells transduced with MIG vector expressing human *HIST1H4C;* WT or R3C constructs. (A) Quality of total RNA prepared from the cell lines as assessed by the Agilent 2100 Bioanalyser. The RIN scores for all the cell lines ranged from 9.5-10.0. **(B)** *HIST1H4C* expression in untransduced and transduced FDM cells. *HIST1H4C* expression was normalised using the housekeeping gene *mouse β-actin* and QPCR was performed in triplicates (values are mean±SD). Value above the horizontal line indicates to the *p*-value. **(C)** GFP fluorescence as assessed by flow cytometry. Each histogram on a plot represents an independent transduction for a particular construct. UT, Untransduced FDM cells; Empty vector, MIG without *HIST1H4C* construct; WT, wildtype *HIST1H4C*; p.R4C, *HIST1H4C* with the p.R4C sequence variant.

format and further analysed for genes of potential interest. Genes were ranked using adjusted *p*-values for differences in expression between *MIG vs. p.R4C* and *WT vs. p.R4C*. Gene annotation and functions were analysed using the online database GeneGo Metacore™ bioinformatics analysis tool (Thomsons Reuters, NY, USA).

Based on the differential expression and function, the expression of 11 genes was further investigated by QPCR (see ***Supplementary Table S2.6*** for primer details). Of the 11 genes, the expression of five genes: *MicroRNA 340* (*Mir340), Pleckstrin homology domain containing family O member 2* (*Plekho2*), *n-Myc downstream regulated gene, 1 (Ndrg1), Dedicator of cytokinesis 10 (Dock10), Interleukin-4 receptor alpha* (*Il4ra*); was validated in the FDM cells expressing the constructs [*n*=4 for each construct, ***Figure 2.14***]. Of these, *Dock10* and *Il4ra* were of particular interest as these genes were significantly up-regulated in FDM cells expressing the *HIST1H4C:p.R4C* variant relative to the wildtype *HIST1H4C* [1.35- and 1.44- mean fold-change respectively, and mean normalised values ±SD for *Dock10* (0.9317 ±0.24 and 0.6917 ±0.21 respectively), and *Il4ra* (8.064 ± 2.69 and 5.598 ±1.64 respectively)]. There was also an increasing trend in the expressions of *Mir340* (1.6 mean fold change), *Plekho2* (1.24 mean fold change), and *Ndrg1* (1.51 mean fold change) in *HIST1H4C:p.R4C* variant relative to the wildtype *HIST1H4C*. However, these did not reach significance possibly due to the high variability within the *HIST1H4C:p.R4C* variant- and wildtype- expressing groups [mean normalised values ±SD for *Mir340* (2.26±1.3 and 1.41±0.78 respectively), *Plekho2* (6.60±2.04 and 5.29 ±1.36 respectively), and *Ndrg1* (2.23±1.60 and 1.46±0.73 respectively)]. This selective deregulated gene expression in the cells over-expressing the *HIST1H4C:p.R4C* construct indicated the capacity of this variant to impact on gene expression in myeloid progenitors.

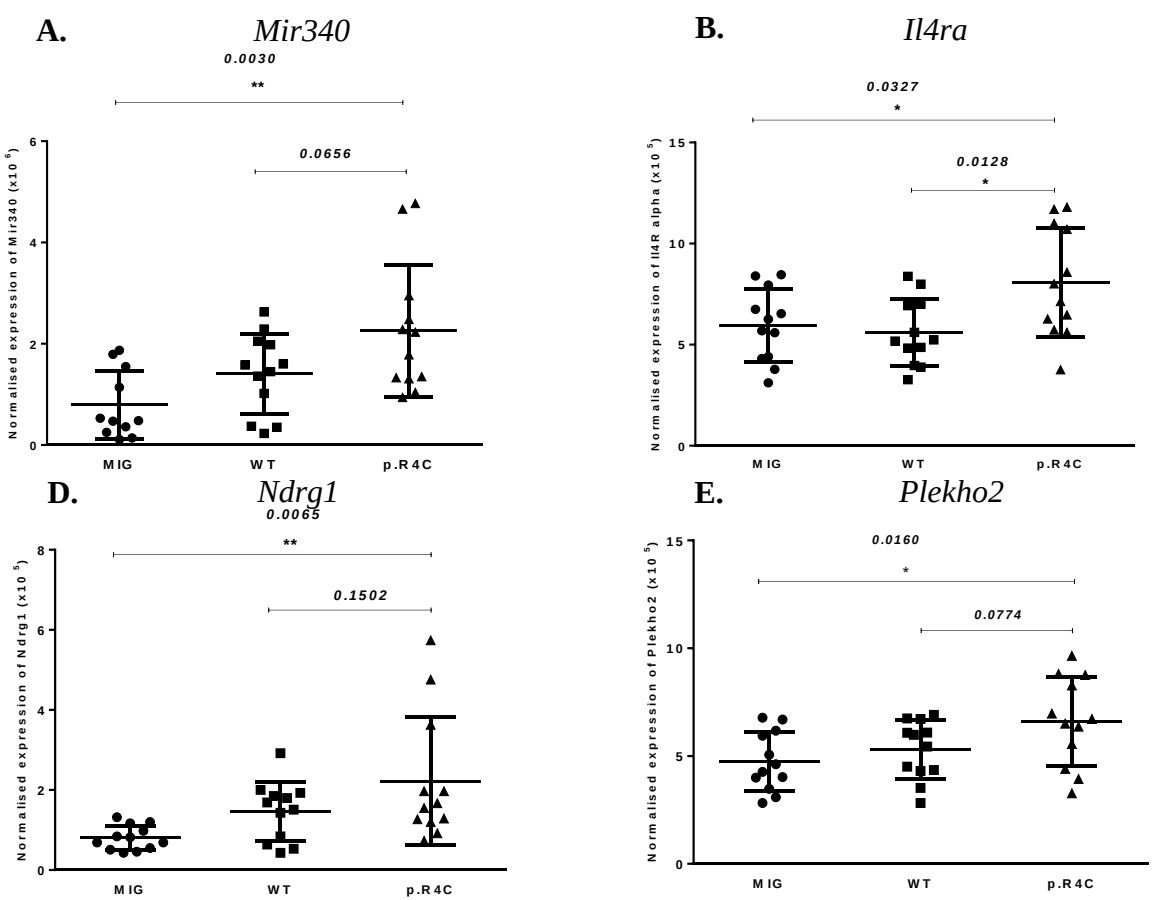


Figure. 2.14. Expression of 5 genes was significantly different in the *HIST1H4C*: p.R4C expressing FDM cells. QPCR validated expression of **A.** *Mir340*, **B.** *Il4ra*, **C.** *Dock10*, **D.** *Ndrg1* and **E.** *Plekho2*. QPCR were performed in triplicates (n=3) for each construct. * and ** denotes statistical significance with a p-value of <0.05 and <0.01 respectively, based on unpaired two-tailed student's t-test. Values have been plotted as normalised mean ±SD. Values represented above horizontal lines represent p-values between two groups.

2.4 Discussion

The aim of this study was to identify novel variants in the *histone H4* genes of importance in MPN pathogenesis. In total, 188 MPN and 1040 non-MPN (which included normal and diagnostic) *g*DNA samples were sequenced. The screen of *histone H4* genes identified two coding variants in *HIST1H4C*: *HIST1H4C:c.10C>T, p.R4C*, in PV109, and, *HIST1H4C:c.167G>A, p.R56Q*, in PV115. As mentioned previously, p.R56Q affects the R55 residue on the mature protein. This arginine residue resides on the outer surface of the nucleosome where it is predicted to be an accessible target for modifications. Based on modelling of the nucleosome, the R55 residue has been predicted to be mono- or di-methylated (Beck, *et al* 2006). Evidence from yeast studies suggests that the R55 residue on the mature histone H4 protein, may play a role in genomic stability wherein alterations at this residue can lead to DNA silencing (Huang, *et al* 2009). Due to the lack of known interacting proteins for this residue, the *HIST1H4C:* p.R56Q variant was not investigated further.

The rare and novel p.R4C variant in patient PV109 was investigated further. R3 on Histone H4 is di-methylated (asymmetric or symmetric) by a variety of PRMTs (Fabbrizio, *et al* 2002; Huang, *et al* 2005; Wang, *et al* 2001; Zhao, *et al* 2009a). The type of di-methylation on R3 can dictate either activation or repression of transcription. The R3 residue can also be modified to citrulline which is a mark of transcriptional repression [(Hagiwara, *et al* 2005) and reviewed in (Jones, *et al* 2009b)]. Interestingly, histone H4R3 is symmetrically di-methylated by PRMT5 (Zhao, *et al* 2009a), which is a phosphorylation target of JAK2 (Liu, *et al* 2011) (described in ***Chapter 1***, ***sections 1.4.3*** and ***1.8***). This raises the possibility that the p.R4C variant, observed in the samples from PV109, may contribute to MPN pathogenesis. Also, it is interesting to note that histone H4R3me^{2}s is an interacting

target for DNMT3A, thereby allowing the subsequent methylation of DNA (Zhao, *et al* 2009a), and it can be speculated that reduced histone H4R3me^2 (symmetric) may lead to altered DNA methylation and deregulation of gene expression. Thus, it was expected that a mutation at this residue would abrogate both symmetric and asymmetric di-methylation. It is important to note that variants affecting the R3 residue exist in other *histone H4* genes (***Table 2.2***), however, given the high relative expression of *HIST1H4C* observed in haematopoietic cell lines and cell populations (***Figure 2.5*** and ***2.6***), it was speculated that the *HIST1H4C* p.R4C could contribute to altered growth and survival of the MPN clone.

The assessment of alterations to methylation on p.R4C via western blotting did not detect consistent changes in methylation levels between the cells expressing histone H4 or the histone variants. It may be possible that the activity of methyltransferases, such as PRMT5, may have been inconsistent due to the differential expression of histones in the cell lines. The results in ***Figure 2.12A*** also suggest that 4-OHT treatment, or higher levels of WT or the p.R4C histone H4 variant, resulted in overall lower histone H4 levels. Thus the over-expression of histones in HEK293 cells was problematic, and it cannot be ruled out that increased histone H4 levels are associated with activation of a pathway for degradation of excess histones (Gunjan, *et al* 2003; Singh, *et al* 2009). Future experiments may require co-expression of histone H3 which may stabilise the excess histone H4 through formation of histone H3-H4 dimers and tetramers (personal communication with Dr. Fan Liu, Sloan-Kettering Institute, USA). Additionally, use of proteosomal inhibitors such as MG132 may prevent the rapid degradation of histones which may help in understanding the effect of histone degradation in experiments. Experiments may

also need to include assessing the levels of relevant methyltransferases in uninduced cells to determine the effect of 4-OHT.

As microarray gene expression profiling is an unbiased and sensitive method for detecting changes in response to stimuli, a gene expression analysis was performed in the IL-3 dependent murine myeloid cell line; FDM (Ekert, *et al* 2004; Jabbour, *et al* 2010). In FDM cells, five genes (*Mir340*, *Dock10*, *Il4ra*, *Plekho2* and *Ndrg1*) showed significant differential expression in the presence of *p.R4C* relative to *WT* or empty vector (***Figure 2.14*** and ***Supplementary Table S2.7***). However, a limitation to these experiments is that histone H4 protein levels were not ascertained. Thus, while the expression of *HIST1H4C* mRNA and the level of GFP fluorescence in cells with *p.R4C* or *WT* were not significantly different (see ***Figure 2.13***), it cannot be excluded that the gene expression changes observed are due to differences in total histone H4 protein levels between these two cell populations.

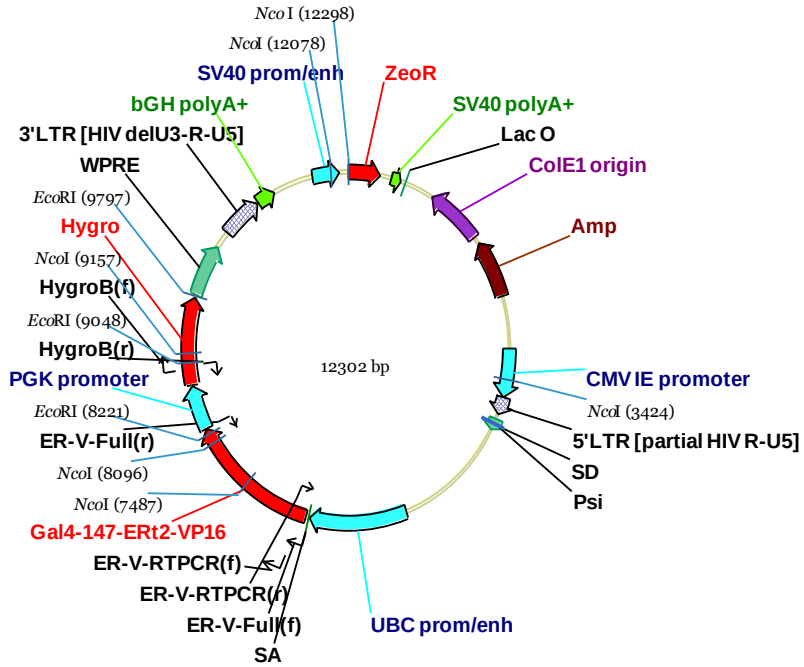
Nevertheless, some interesting target genes were identified including *Dock10* (Alcaraz-Garcia, *et al* 2011) and *Mir340*. Predicted targets of *Mir340* included *Myeloid Leukaemia Factor 1* (*MLF1)* [MicroCosm targets version 5 (EMBL-European Bioinformatics Institute, Wellcome Trust Genome, UK)], which was of interest (Arber, *et al* 2003; Bras, *et al* 2012; Lee, *et al* 2012). In the FDM cells, *Mlf1* displayed a trend toward reduced expression, although this still needs to be validated.

Future experiments will include further investigations of selected target genes in MPN pathogenesis. For example, given the link between histone H4R3 modifications and DNMT3A, it is of interest to know whether changes in expression of these genes is associated with changes to *DNMT3A*. Given the importance of histones and PTMs

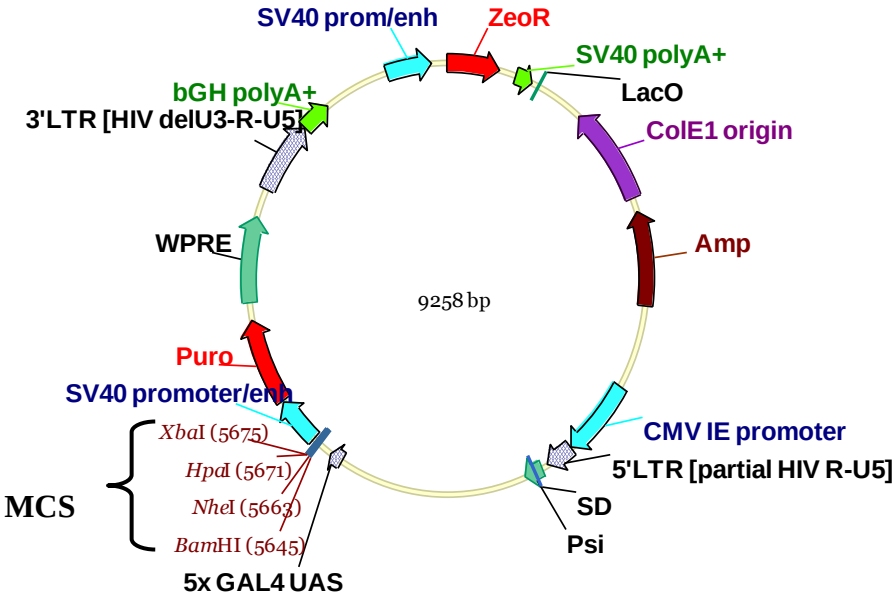
in epigenetics, it would also be interesting to assess alterations in global histone modifications and DNA methylation in myeloid cells as a consequence of *HIST1H4C:p.R4C* and *HIST1H4C:p.R56Q* mutations. Finally it will be important to assess alterations to JAK2 signalling in patient samples and experimental models with the histone H4 MPN variants. These studies are beyond the scope of this project.

2.5 Supplementary information

A. ***ap pF U G147EV16 PGK Hygro W (GEV16)***



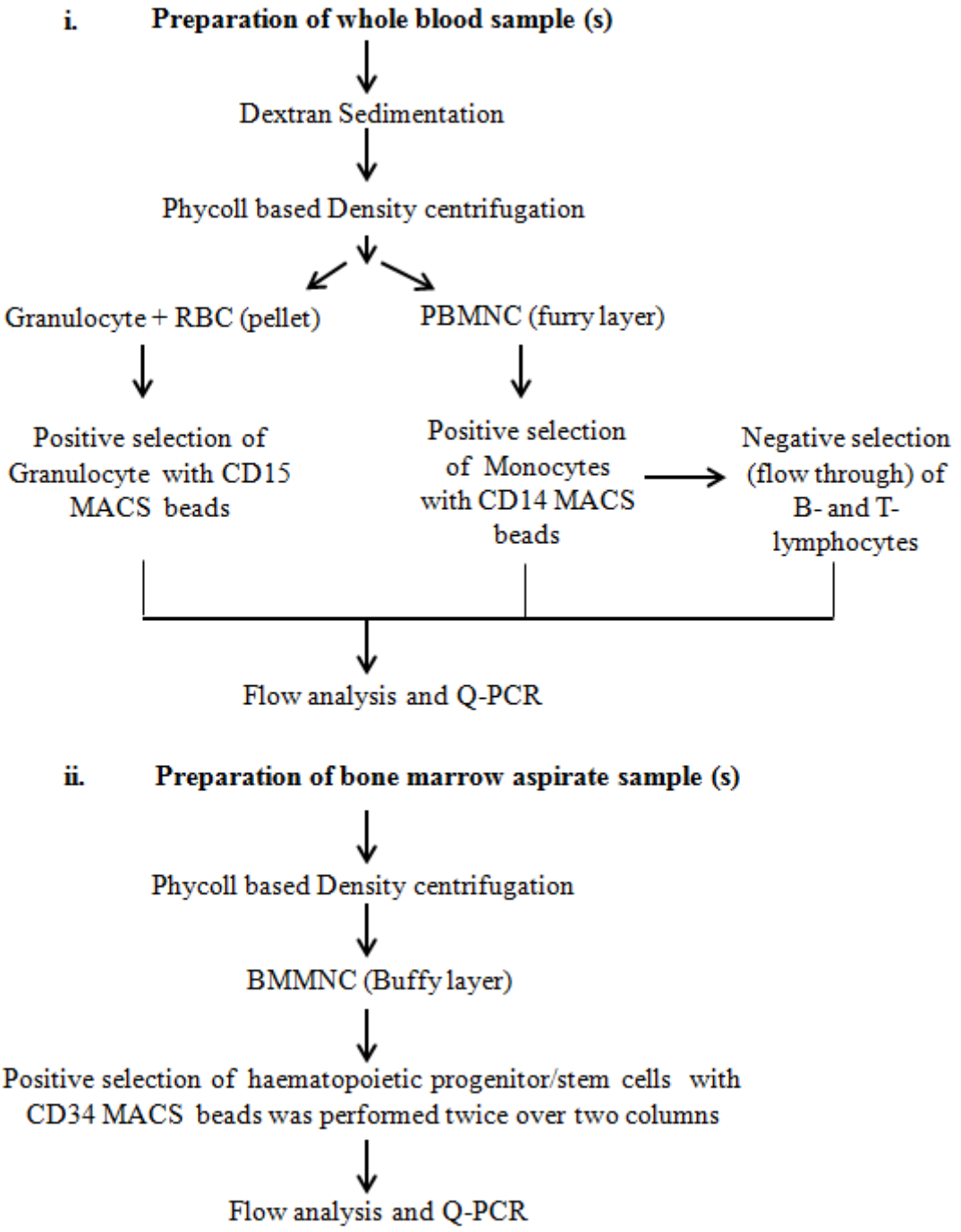
B. ***pF 5x UAS SV40 puro W (pF 5x UAS SV40 puro)***



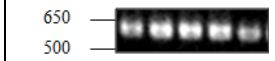
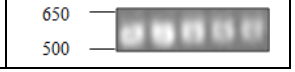
Supplementary Figure S2.1 Schematic representation of the 4-hydroxy tamoxifen (4-OHT) regulatable lentiviral vectors. A and **B** depict maps of the plasmids expressing the *GAL4-Estrogen receptor-VP16* fusion construct and expression vector with multiple cloning sites (MCS) respectively. Plasmid stocks and information was generously provided by A/Prof. Chris Hahn and Prof. Hamish S.Scott.

Histone H4 primer	*RE site introduced*	*Primer (5`-3`)*	*Primer Tm (°C)*	*Product size (bp)*
HIST1H4C- Forward Arg4-BamHI	*BamHI*	ccatccggatccatgtctggtcgcggcaaa	65.87	323
HIST1H4C- Forward Cys4-BamHI	*BamHI*	ccatccggatccatgtctggttgcggcaaa	65.87	
HIST1H4C-Reverse-HpaI	*HpaI*	tctagagttaactcagccgccgaagccata	66.15	
HIST1H4C- Forward Arg4-EcoRI	*EcoRI*	cgccggaattcatgtctggtc	67.00	333
HIST1H4C- Forward Cys4-EcoRI	*EcoRI*	gccggaattcatgtctggttgc	67.00	
HIST1H4C-Reverse-EcoRI	*EcoRI*	ttaacgaattccagccgccg	66.00	
GFP-Forward-BamHI	*BamHI*	ccatccggatccatggtgagcaagggcgaggag	68.12	744
GFP-Reverse-HpaI	*HpaI*	tctagagttaacttacttgtacagctcgtccatgcc	64.6	

Supplementary Table S2.1. *HIST1H4C* cloning primers. Primers for the human *HIST1H4C* cDNA was designed using the Primer3 Input (v.0.4.0) online software. *g*DNA samples were amplified using Roche Hi-fidelity DNA polymerase kit (Roche Applied Science, Roche Diagnostics Australia Pty. Ltd, Australia) and the cycling conditions for all the primer pairs were as follows: 5 minutes at 95°C, followed by 35 cycles of 30 seconds at 94°C, 30 seconds at 62°C and 45 seconds at 72°C, followed by a final extension at 72°C for 5 minutes. Ligations were performed at a construct to vector ratio of 4:1 along with T4 DNA ligase (Roche Applied Science, Roche Diagnostics Australia Pty. Ltd, Australia) for 30 minutes following which the reaction volumes were used to transform 100 µL of ultra-competent *E.coli* (Promega, Australia). RE, Restriction Enzyme



Supplementary Figure S2.2. Isolation of primary human haematopoietic cell sub-populations. i. Monoyctes and granulocytes from peripheral blood; and **ii.** $CD34^{+}$ cells from bone marrow aspirate. Briefly, the cells are incubated with magnetic beads conjugated to a specific antibody. Unbound cells flow through the column (negatively selected fraction) and the bound cells are retained in the column and are forced out of the column by a syringe.

GENE ID (RefSeq ID)	**Forward primer (5`→3`)**	**Forward primer Tm (C°)**	**Reverse primer (5`→3`)**	**Reverse primer Tm (C°)**	**Product size (bp)**	**Products observed on 1.5%TAE agarose gel**
HIST1H4A (NM_003538)	ggactgtaggcgtcacattt	57.67	cccctaatctttccttttagg	57.48	610	
HIST1H4B (NM_003544)	tcccgctttcagttctcaat	59.81	ggcagaatagccgagtaaagc	60.37	531	
HIST1H4C (NM_003542)	caggtccgccaagtttgtat	59.17	ggtcatggcaaacacaattc	58.83	577	
HIST1H4D (NM_003539)	caaacaggtccgtcatgcta	59.72	gccaaggcgtaataagacca	60.10	551	
HIST1H4E (NM_003545.3)	cgaattcccggcatataaga	59.41	agagaccaaccaatcgtctaaca	60.05	504	
HIST1H4F (NM_003540.3)	ctccaaaaaggtccgcataa	60.07	acaaaacgaaacccaagtcg	60.01	511	
HIST1H4G (NM_003547.2)	tataaaaggctgcggaaagc	59.48	aacagcttttgggcacctac	59.24	527	
HIST1H4H (NM_003543.3)	gcatataagggcgttgcttt	59.41	ctcagcaaaaatctggaacaaa	59.38	500	
HIST1H4I (NM_003495.2)	ccgcagaggttacccataaa	59.95	ttcggcgtcgagttaatctt	59.85	585	
HIST1H4J (NM_021968.3)	gtccgccaactgtcgtataaag	60.01	aactgactcagaaccaagcaaaat	60.55	511	
HIST1H4K (NM_003541.2)	gtccgccaactgtcgtataaa	60.01	gatagcagcacaaactcagcag	60.22	517	
HIST1H4L (NM_003546.2)	cccaatcaggtccgatttatt	60.03	gccaaaactgccatacgttta	60.01	514	
HIST2H4A/B (NM_003548.2)	gctttcggttttcaatctggt	59.69	ctgttgcactttacgcagtca	60.10	507	
HIST4H4 (NM_175054.2)	tcgagagactataagccctggt	59.39	cgaatatcccatttccggtat	59.89	543	

Supplementary Table S2.2 *Histone H4* primer sequence with the respective annealing temperatures, size of the product and an example of the PCR products obtained (PCRs of select MPN *g*DNA samples) as run on a 1.5% TAE gel based (see ***section 2.2.6***).

Antibody	Product details		Lot #	Dilution
Anti Human CD34-PE	IO TEST (BECKMAN COULTER)	PN IM1871U	NA	1:20
CD14-PE	eBiosciences	12-0149-73 (CLONE 61D3)	E026669	1:20
CD15-PE	BD Pharmingen	323006 (CLONE W6D3)	B110765	1:20
CD19-PE	BD Pharmingen	555413	36211	1:20
CD3-PE	Becton Dickinson (BD Biosciences)	347347	NA	1:20
PE Mouse IgG1, κ Isotype Control	BD Pharmingen	550617	NA	1:50

Supplementary Table S2.3. Summary of the fluorophore-conjugated antibodies used to measure purity of CD34 (HSPCs), CD14 (Monocyte), CD15 (granulocytes/polymorphonuclear cells), CD19 (B-lymphocytes) and CD3 (T-lymphocytes) cell fractions isolated by MACS. NA, not available.

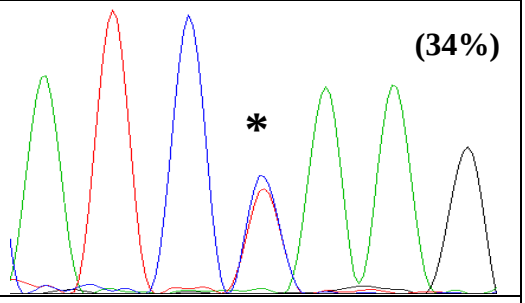
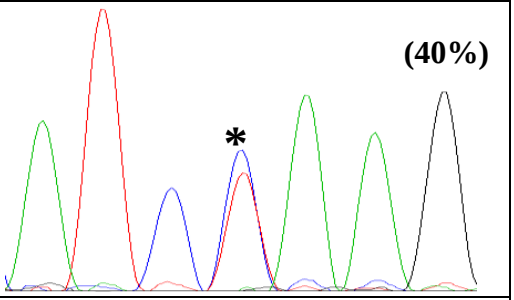
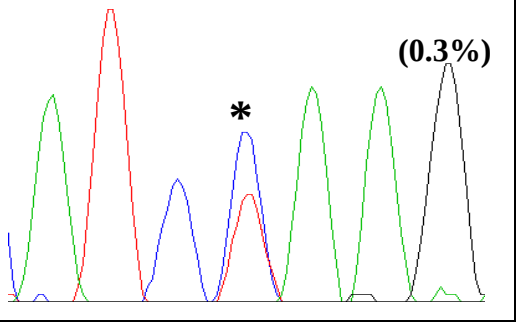
Gene	Chr:position	Alleles (Reference SNP ID)	Position, Minor Allele	Freq. in PV (%)	Pop^{n}. Freq (%)	Source
5`UTR variants						
HIST1H4B	6:26027507	G/C	c. -27 G>C, C	1.136	ND	NA
	6:26189335	C/T	c.-30C>T, T	1.136	ND	NA
3`UTR variants						
HIST1H4B	6:26027135	G/A (rs3752420)	c. *34C>T, T	9.090	42.30	dbSNP
HIST1H4F	6:26241008	C/T (rs201595082)	c.*43C>T, T	1.136	ND	dbSNP
HIST1H4G	6:26240976	G/A	c.*11G>A,A	1.136	ND	NA
HIST1H4J	6:27792300	G/C (rs141884356)	c.*86G>C ,C	1.136	0.280	dbSNP
HIST1H4K	6:27798937	G/A (rs116719739)	c.*57C>T, T	12.50	1.330	dbSNP
HIST2H4A/B	1:149832366	A/G	c.*20T>C, C	18.18	ND	NA

Supplementary Table S2.4 Summary of *Histone H4* 5`UTR, 3`UTR variants identified in PV patients. 15 *histone H4* genes were screened in 88 PV patient *g*DNA samples (102 were screened for *HIST1H4C*). Ensembl|1000 genomes and dbSNP (build 135, 137 and 138) databases were used to identify novel variants from known variants. Freq. in PV, frequency in the PV cohort; Pop^{n}. freq., population frequency ascertained from databases; ND, Not Detected; NA, Not Applicable; UTR, Untranslated region; *, Stop codon.

Gene (sample size)	Chr: position	Alleles	Position, Minor Allele	Number of individuals detected with the minor allele	Freq. in normal cohort (%)
5` UTR variants					
HIST1H4B (n=228)	6:26027507	G/C	c. -27 G>C,C	00	00.00
HIST1H4D (n=233)	6:26189312	G/C (rs79868973)	c. -8 G>C,C	34	22.74
	6:26189313	G/C (rs78735537)	c.-9G>C,C	19	08.15
	6:26189322	T/C	c.-18 T>C,C	01	00.42
3`UTR variants					
HIST1H4B (n=228)	6:26027135	G/A (rs3752420)	c. *34 C>T , T	10	04.82
HIST1H4C (n=236)	6:26104535-26104538	AAAA>del_AAAA (rs58741377)	c.*48 AAAA> del_AAAA	01	00.42
	6:26104554	C/T (rs138617314)	c.*67 C>T,T	01	00.42
HIST1H4F (n=237)	6:26241008	C/T (rs201595082)	c.*43 C>T,T	01	00.42
HIST1H4K (n=233)	6:27798937	G/A (rs116719739)	c.*57 C>T,T	19	08.15
	6:27798929	A/C	c.*65 T>C,C	04	01.17
HIST2H4A/B (n=227)	1:149832366	A/G	c.*20 T>C, C	39	16.70

Supplementary Table S2.5 Summary of *Histone H4* 5' UTR, 3' UTR variants identified in a normal cohort. 15 *histone H4* genes were screened in disease free normal buccal *g*DNA samples. Freq.,frequency in the cohort;.UTR, Untranslated region;*, Stop codon; ND, Not Detected; NA, Not Applicable.

A.

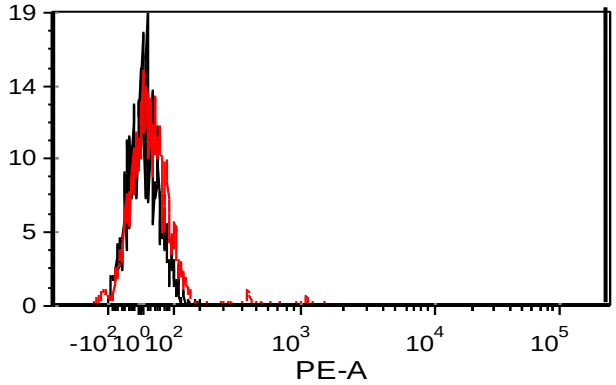
Sample type	PV103 A T C C A A G
Granulocyte *g*DNA (*JAK2V617F*%)	
PBMNC *g*DNA (*JAK2V617F*%)	
Buccal *g*DNA (*JAK2V617F*%)	

B.

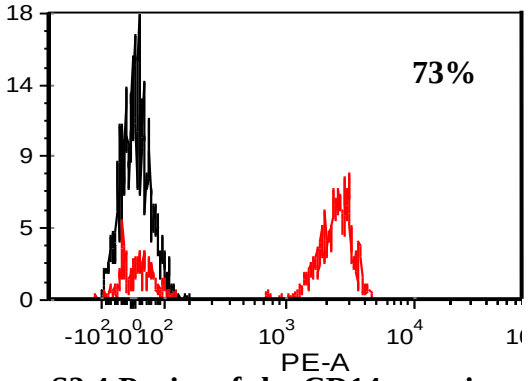
Patient characteristics	PV103	Normal Reference range
Gender/Age at diagnosis	M/83	NA
JAK2V617F mutation	positive	NA
Haemoglobin	130g/L	135-175 g/L
Red Blood Cell	6.62 $x10^{12}$/L	4.50-6.00 $x10^{12}$/L
White Cell	13.0 $x10^{9}$/L	4.00-11.0 $x10^{9}$/L
Serum EPO	2mU/ml	3.3-16.6 mU/ml
EPO-independent BFUE (e-BFUE)	detected	NA

Supplementary Figure S2.3. Clinical characteristics of the Polycthemia vera patient PV103. A. Sequence traces of *HIST1H4B* in PV103 depicting the germline variants; *HIST1H4B*:*c.82C>T, p.Q28**. The underlined nucleotides depict wild-type sequence. The sequence trace also shows the nucleotide affected (as depicted by *) in granulocyte, PBMNC and buccal *g*DNA for *HIST1H4B* (% indicates to the *JAK2V617F* allele load). **B.** Patient characteristics. Samples were collected at the time of diagnosis (2008) and values for each patient characteristic is from the year 2008. The values in the normal reference range column were used at the time of diagnosis. NA, not applicable.

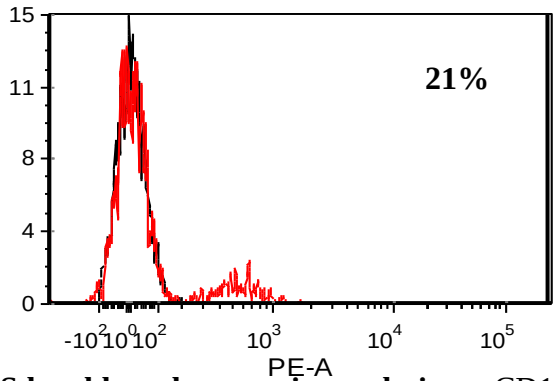
i. CD14 negative fraction with anti-CD14 staining



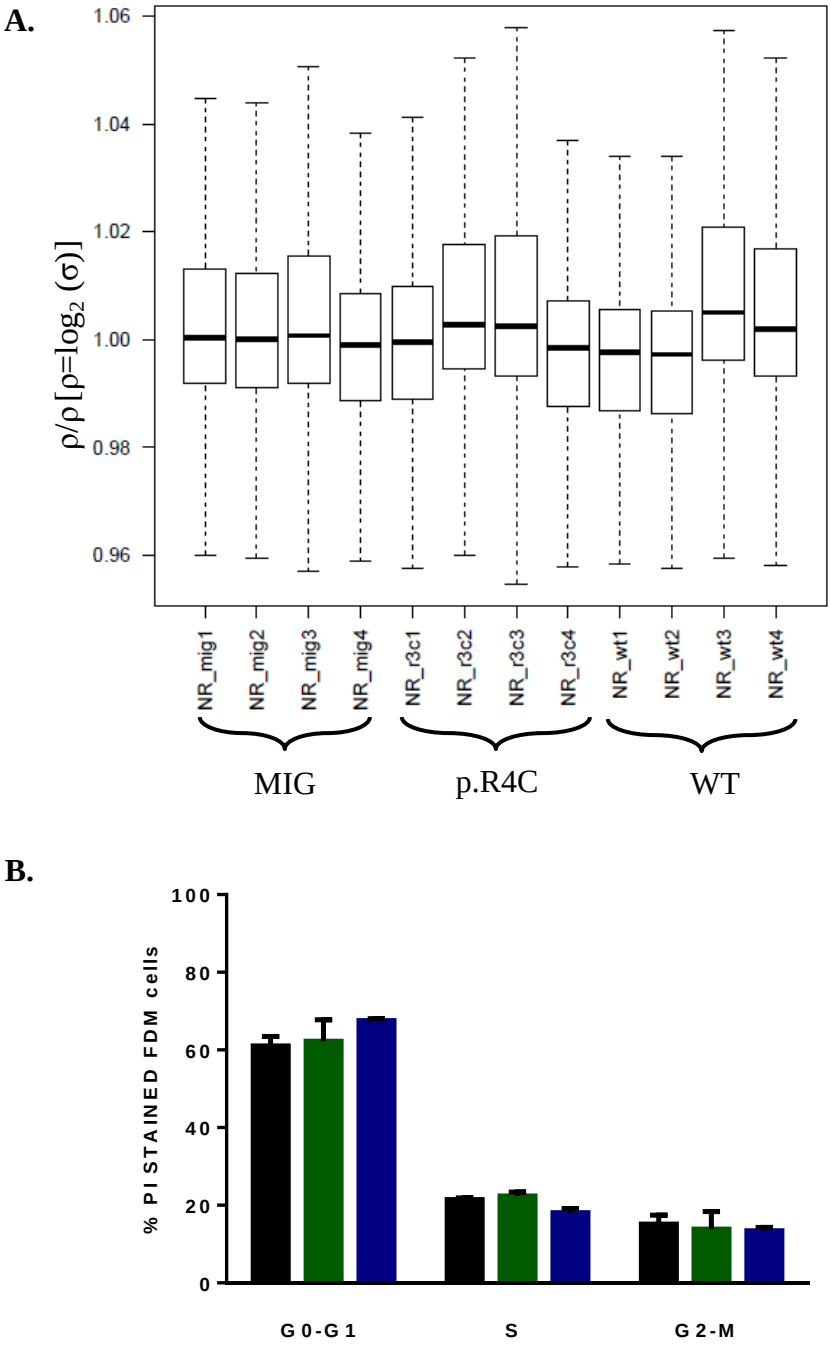
ii. T-lymphocyte (anti-CD3)



iii. B-lymphocyte (anti-CD19)



Supplementary Figure S2.4 Purity of the CD14 negative cell fraction following MACS bead based separation technique. CD14 negative cells were incubated either with isotype control, CD14-PE, CD3-PE or CD19-PE antibodies and then analysed by flow cytometry. $CD3^+$ and $CD19^+$ cells constituted the CD14 negative fractions as depicted in plots **ii** and **iii** as red histograms compared to the isotype control (black histograms). As observed in plot **i** there were virtually no $CD14^+$ cells in this fraction. The plots represent an example of flow cytometric analysis following MACS based separation of a normal PBMNC sample.



Supplementary Figure. S2.5. A. Normalised Unscaled Standard Errors (NUSE) plot. The NUSE values are standardized at the probe-set level across the Mouse Gene_ST arrays: Typically values centered around 1 are considered as good quality arrays. A low quality array in which the median value of a box-plot is plotted further away from 1 or has a bigger distribution is removed from analysis. The quality of each of the Mouse Gene_St array was plotted using a CEL file format and were of good quality. **B. Cell cycle analysis of FDM cells.** There was no significant difference observed in the cells expressing the *p.R4C* variant with either *MIG* or *WT* cells. The values represent n=4, values are plotted as mean±SD, and two-way ANOVA was performed [F (2, 27) =0.0882, *p*=0.9158]. ■, *MIG* (empty vector); ■, *WT*; ■, *p.R4C* expressing FDM cells. G0-G1, S and G2-M depict phases of cell cycling.

Gene (mouse RefSeq ID)	**Forward Primer (5`→3`)** / **Reverse Primer (5`→3`)**	**Primer Tm (C°)** / **Primer Tm (C°)**	**Product size (bp)**	**Exon-Exon junction**	**R^2 value**
Pmf1 (NM_025928.3)	atccatccgcgaggaaatc	57.39	71	Yes	0.9695
	tatccagggagttcaggaca	57.06			
Slc38a1 (NM_134086.4)	gcgtacttgtctcagacgtt	57.95	77	Yes	0.9943
	gagaccccgttttctagcag	57.99			
Ndrg1 (NM_008681)	ccatcttcgtctccctcttc	57.12	94	Yes	0.9991
	cacacacctaatgagcaagc	57.09			
Fancd2 (NM_001033244.3)	ttgggcagaccatcttcaat	57.09	79	Yes	0.9961
	aatggaaagtcaccctctgg	57.11			
H2afx (NM_010436)	cagctcttctacctcgtacac	57.59	82	No	0.9787
	tgaagagcgcgacttgg	57.25			
Plekho2 (NM_153119.2)	acgagaatgaggatgagcag	57.11	55	Yes	0.9844
	tcatagctcccaagttccac	56.92			
Pten (NM_008960)	actgcagagttgcacagtat	57.15	54	Yes	0.9938
	gttctagctgtggtgggtta	56.86			
Dock10 (NM_175291)	gaacttgaaaagacctatcggg	56.79	77	Yes	0.9995
	gtggctgttgagaagtcatc	56.45			
Il4ra (NM_001008700)	gtggtaatgtgaagccccta	56.91	85	Yes	0.9924
	aggttattccaggtcagca	56.91			
Bard1 (NM_007525)	agtgctctatcgtgaacacc	57.34	54	Yes	0.9924
	ctataaataccagcggccca	57.15			
Mir340 (NR_029769)	ttggtgtgattataaagcaatga	57.30	62	No	0.9757
	actgagacggatcccaca	57.36			

Supplementary Table S2.6. Forward and reverse primers designed to assess the changes in gene expression observed in FDM constructs. Primers have been depicted with the respective annealing temperatures, size of the product and the R^2-value ascertained by DNA standards ratios ranging from 1- 1:10 (where 1 represents a cDNA dilution of 100ng/µL). The cycling conditions for all the primer pairs was 5 minutes at 98°C, followed by 25 cycles of 30 seconds at 57°C, and primer product melts were performed from 70-90°C. One of the primer pair of each primer set spanned the exon-exon junction except for genes *Mir340* and *H2afx* which had no introns.

Supplementary Table S2.7. Top 50 differentially regulated genes in FDM cells expressing p.R3C compared to WT or MIG

ProbesetID	Adjusted p-value	*Gene Symbol*	Chromosome	strand	Start	Stop	Gene_assignment
10501072	1.77E-06	*Olfr266*	chr3	-	106624474	106625500	olfactory receptor 266
10584458	8.27E-05	*Olfr920*	chr9	+	38563275	38564213	olfactory receptor 920
10375677	9.70E-06	*Mir340*	chr11	+	49883204	49883301	microRNA 340
10594538	5.54E-05	*Plekho2*	chr9	-	65399491	65399898	pleckstrin homology domain containing, family O member 2
10586240	0.00015501	*Dennd4a*	chr9	+	64709710	64709854	DENN/MADD domain containing 4A
10482695	0.00013003	*Cacnb4*	chr2	-	52283840	52532101	calcium channel, voltage-dependent, beta 4 subunit
10429140	0.00012556	*Ndrg1*	chr15	-	66760880	66801201	N-myc downstream regulated gene 1
10493666	1.68E-05	*Nup210l*	chr3	+	89975971	89976099	nucleoporin 210-like
10363312	2.60E-05	*Ccdc138*	chr10	+	57960696	58038990	coiled-coil domain containing 138
10499366	0.00013293	*Pmf1*	chr3	-	88198060	88214256	polyamine-modulated factor 1
10559446	0.00014126	*Lilrb3*	chr7	-	3664134	3671984	leukocyte immunoglobulin-like receptor, subfamily B (with TM and ITIM domains), member 3
10431872	0.00020623	*Slc38a1*	chr15	-	96402753	96403121	solute carrier family 38, member 1
10432511	0.00020571	*Racgap1*	chr15	-	99450722	99482062	Rac GTPase-activating protein 1
10501164	8.32E-05	*Csf1*	chr3	-	107543975	107563387	colony stimulating factor 1 (macrophage)
10499904	0.00025879	*Ivl*	chr3	-	92374824	92377637	involucrin
10606064	0.00021096	*Rgag4*	chrX	-	99261883	99266643	retrotransposon gag domain containing 4
10374060	4.75E-05	*Ankrd36*	chr11	+	5528765	5545156	ankyrin repeat domain 36 Ankrd36
10406982	0.00025319	*Adamts6*	chr13	+	105077915	105286775	a disintegrin-like and metallopeptidase (reprolysin type) with thrombospondin type 1 motif, 6

10554325	0.00027187	*5730590G19Rik*	chr7	+	86805082	86843246	RIKEN cDNA 5730590G19 gene
10583203	0.00033547	*Phxr4*	chr9	+	13234581	13235971	per-hexamer repeat gene 4
10382321	0.00017856	*Kcnj2*	chr11	+	110927478	110936965	potassium inwardly-rectifying channel, subfamily J, member 2
10577326	0.00034532	*Xkr5*	chr8	-	18932729	18950975	X Kell blood group precursor-related family, member 5
10566574	0.0005248	*Gvin1*	chr7	-	113300049	113358854	GTPase, very large interferon inducible 1
10542249	0.00038397	*Tas2r123*	chr6	+	132797160	132798161	taste receptor, type 2, member 123
10519857	0.00029055	*Hgf*	chr5	+	16059368	16125257	hepatocyte growth factor
10544171	0.0002866	*Slc37a3*	chr6	-	39285970	39327671	solute carrier family 37 (glycerol-3-phosphate transporter), member 3
10513805	0.00031017	*Dbc1*	chr4	-	68422406	68615431	deleted in bladder cancer 1 (human)1
10469867	0.00038246	*Pnpla7*	chr2	+	24831582	24909577	patatin-like phospholipase domain containing 7
10449672	0.00022989	*Tmprss3*	chr17	-	31316210	31335304	transmembrane protease, serine 3
10374043	2.66E-05	*Ankrd36*	chr11	+	5469688	5514689	ankyrin repeat domain 36
10487340	0.00056632	*Ncaph*	chr2	-	126929562	126959679	non-SMC condensin I complex, subunit H
10576403	0.00034189	*Urb2*	chr8	+	126547373	126572404	URB2 ribosome biogenesis 2 homolog (S. cerevisiae)
10472128	0.00054218	*Arl6ip6*	chr2	+	53051145	53078060	ADP-ribosylation factor-like 6 interacting protein 6
10430825	0.00049154	*Cenpm*	chr15	-	82064191	82075178	centromere protein M
10548859	0.00044136	*Wbp11*	chr6	-	136762175	136776649	WW domain binding protein 11
10519472	0.00047196	*Gm9897*	chr5	+	4192367	4197652	predicted gene 9897
10598507	0.00045125	*Slc38a5*	chrX	+	7848520	7857300	solute carrier family 38, member 5
10557326	0.00056141	*Il4ra*	chr7	+	132695796	132722988	interleukin 4 receptor, alpha

10506680	0.00063168	*Tmem48*	chr4	+	107040389	107086946	transmembrane protein 48
10601011	0.00061743	*Kif4*	chrX	+	97821076	97922553	kinesin family member 4
10491300	0.00052821	*Skil*	chr3	+	30993983	31021499	SKI-like
10359890	0.00060161	*Nuf2*	chr1	-	171428065	171461595	NDC80 kinetochore complex component, homolog (S. cerevisiae)
10586184	0.00065827	*Tipin*	chr9	+	64129414	64152601	timeless interacting protein
10530156	0.00065143	*Tmem156*	chr5	-	65454457	65482794	transmembrane protein 156
10363845	0.00071348	*Ccdc6*	chr10	+	69560048	69652026	coiled-coil domain containing 6
10385822	0.00048864	*Gm10447*	chr11	-	53269795	53270202	predicted gene 10447
10407435	7.89E-05	*Akr1c18*	chr13	-	4131861	4149877	aldo-keto reductase family 1, member C18
10604832	0.00091577	*Mir505*	chrX	-	57647578	57647667	microRNA 505
10384423	0.00040848	*Cobl*	chr11	-	12136611	12364807	cordon-bleu
10373918	0.00072172	*Lif*	chr11	+	4157571	4172517	leukaemia inhibitory factor
10366335	0.00078603	*Bbs10*	chr10	+	110735735	110738781	Bardet-Biedl syndrome 10 (human)

Chapter 3: *PHF6* and *histone H3* mutation screening of MPN patients

Statement of Authorship

Title of Paper	***PHF6* and *histone H3* mutation screening of MPN patients**
Publication Status	○ Published ○ Accepted for Publication ○ Submitted for Publication ○ Publication Style
Publication Details	Rao N, Butcher CM, Schreiber AW, Hahn CN, Corbett M, Gécz J, Lewis ID, Bardy PG, Scott HS, D'Andrea RJ

Author Contributions

By signing the Statement of Authorship, each author certifies that their stated contribution to the publication is accurate and that permission is granted for the publication to be included in the candidate's (***Nisha Rao***) thesis.

Name of Principal Author (Candidate)	Nisha Rao		
Contribution to the Paper	60%, performed the research, analysed and interpreted data and wrote the publication		
Signature		Date	9/04/2013

Name of Co-Author	Carolyn M Butcher		
Contribution to the Paper	5%, designed the research study and interpreted data		
Signature		Date	22/03/2013

Name of Co-Author	Andreas W Schreiber		
Contribution to the Paper	6%, Bioinformatics Analysis of the Roche 454 sequencing		
Signature		Date	16/03/2013

Name of Co-Author	Chris N Hahn		
Contribution to the Paper	5%, gave critical input for experimental design and setup		
Signature		Date	21/03/2013

Name of Co-Author	Mark Corbett		
Contribution to the Paper	2% provided reagents and technical advice		
Signature		Date	14/3/2013

Name of Co-Author	Jozef Gécz		
Contribution to the Paper	2% technical advice and background for PHF6		
Signature		Date	14/3/2013

Name of Co-Author	Ian D Lewis		
Contribution to the Paper	3%, provided clinical information on patients		
Signature		Date	8/3/2013

Name of Co-Author	Peter G Bardy		
Contribution to the Paper	3%, provided clinical information and coordinated collection of MPN patient samples		
Signature		Date	9/04/2013

Name of Co-Author	Hamish S Scott		
Contribution to the Paper	5%, critical input on experimental setup		
Signature		Date	26/03/2013

Name of Co-Author	Richard J D'Andrea		
Contribution to the Paper	9%, Principal supervision of the project and critical input in writing		
Signature		Date	26/03/2013

3.1 Introduction

MPN and adult AML have a considerable overlap in mutations indicating the involvement of common pathways which include JAK2, EZH2 and ASXL1. However, the identification of mutations associated specifically with MPN or leukaemic transformation indicates that specific genetic changes direct disease course [(Rocquain, *et al* 2010) and reviewed in (Tefferi 2010)]. The aim of the following study was to investigate the presence and frequency of mutations in *PHF6*, commonly mutated in leukaemia, and histone H3 encoding genes which have been reported to be mutated in glioma (Schwartzentruber, *et al* 2012; Wu, *et al* 2012). These two proteins may potentially be interacting in a chromatin modifying complex; *NU*cleosome *R*emo*D*elling (NURD) complex which has an important role in cancer [(Todd, *et al* 2012) and reviewed in (Lai, *et al* 2011)].

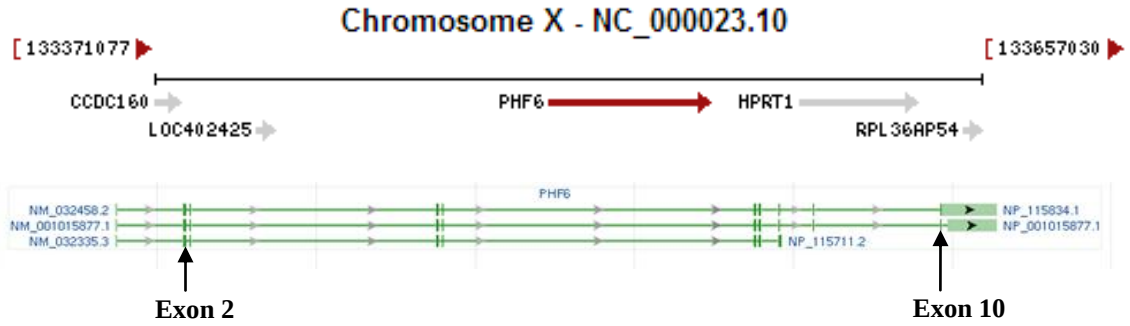
3.1.1 PHF6

The *plant homeodomain finger-6* (*PHF6*) gene was recently identified as a novel tumour suppressor gene mutated in T-cell acute lymphoblastic leukaemia (T-ALL) and AML at a frequency of 20% and approximately 3% respectively (Chao, *et al* 2010; Van Vlierberghe, *et al* 2010; Van Vlierberghe, *et al* 2011). 70% of the mutations observed in *PHF6* were frameshift and the remaining 30% were missense mutations. 2.4% of Ph^+ blast crisis CML has also been found to harbour a mutation in *PHF6* (Li, *et al* 2012). Additionally, a whole genome sequencing approach for 24 AMLs identified one mutation in *PHF6* in one patient sample (Welch, *et al* 2012), however, no screens have been reported in *Ph*-negative chronic phase MPN.

PHF6 was initially reported to harbour mutations in the inherited intellectual disability disorder Borjeson Forssman Lehman syndrome [BFLS, (Lower, *et al* 2002)] (***Figure 3.1*** shows the gene location, protein structure and the mutations identified thus far in BFLS, AML, and T-ALL). The protein encoded by this gene belongs to a large family of zinc finger genes known as the plant homeodomain-like family (PHF) and consists of two PHD domains, suggesting a role in chromatin regulation. Chromatin modifying proteins are of particular interest in myeloid malignancies given that a number of these proteins have been shown to be mutated in MPN such as ASXL1 (has a PHD domain) and EZH2 (interacts and trimethylates histone H3 at K27).

The PHF family includes factors which have a role in chromatin remodelling, for example PHF1 (Musselman, *et al* 2012), PHF2, and PHF8 [reviewed in (Fortschegger, *et al* 2011)]. PHF1 is involved in maintaining global histone H3K27me^3 and has been suggested to be a component of the polycomb repressive complex 2 (PRC2 complex), in addition to EZH2 (Sarma, *et al* 2008). PHF2 and PHF8 are known histone lysine demethylases which catalyse the removal of di-methyl groups on lysine residues (K4 and K9) of histone H3 and histone H4 (K20). Although the role of PHF6 is not completely determined, it has been speculated it may play a role in transcriptional regulation possibly as a component of the NURD complex. Todd and colleagues have reported the co-purification of PHF6 with NURD complex proteins along with several ribosomal proteins, spliceosomal proteins and histone H3 (Todd, *et al* 2012).

A.



B.

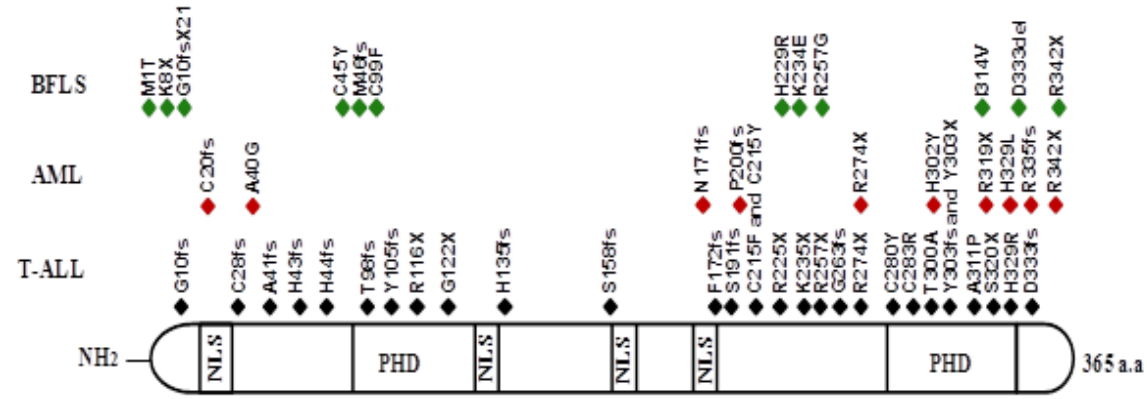


Figure 3.1 Genomic location and protein structure of PHF6. (**A**) human *PHF6* is present on X chromosome (ChrXq26.2) and is a member of a large family of zinc-finger genes. *PHF6* is transcribed as a 4.5kb mRNA and has two major isoforms which differ in the inclusion or exclusion of intron 10 from the mRNA. This alternative splicing does not change the PHF6 protein. The figure depicts a minor isoform along with the two major isoforms. *PHF6* is highly conserved in vertebrates but has no obvious ortholog (s) in lower organisms [i.e., insect or yeast, *adapted from NCBI genome browser*]. (**B**.) PHF6 is 365 amino acids long and has two PHD domains, which suggests an important cellular (nucleus and nucleolus) role of the protein. Figure depicts known mutations of PHF6 in BFLS, AML, and T-ALL. The diamonds indicate to position and residue affected. Different coloured diamonds represent the mutations in the three separate disorders. NLS, nuclear localisation sequence; X, Stop codon; fs, frameshift; del, deletion.

3.1.2 Histone H3

To complement the study focussed on *PHF6*, mutations in *histone H3* genes were also investigated. A potential role for histone H3 in MPN is supported by the discovery of a number of somatic mutations in genes encoding histone H3-interacting-proteins such as *EZH2*, *ASXL1* and *JAK2*. The discovery of somatic mutations in the *histone H3* genes; *H3F3A* (encoding for H3.3) and *HIST1H3B* (encoding for H3.1) in glioblastomas (Schwartzentruber, *et al* 2012; Wu, *et al* 2012) and subsequently in a number of other studies (Behjati, *et al* 2013; Behjati, *et al* 2014; Huether, *et al* 2014; Je, *et al* 2013; Je, *et al* 2014) demonstrates potential for the occurrence of pathogenic mutations in genes encoding histone H3 in MPN. There are 17 known *histone H3* genes in the human genome which encode for three histone H3 variant proteins (H3.1, H3.2 and H3.3). (***Figure 3.2*** shows the conservation of amino acid residues across the H3 variants). 11 of the *histone H3* genes encode for histone H3.1 (including the testicular cell specific histone H3.1t), 3 for histone H3.2 and 3 for histone H3.3. Histone H3.3 is a unique histone which, unlike the classic histone H3 proteins, has a poly-A tail and has specialised functions which are independent of cell cycle. These specialised roles include chromatin assembly, chromatin maintenance, and telomere maintenance (Ahmad, *et al* 2002; Lewis, *et al* 2010; Lin, *et al* 2013; Tagami, *et al* 2004). The telomere maintenance role of H3.3 is particularly interesting as it has been previously reported that telomeres are shorter in MPN patients compared to control samples (Bernard, *et al* 2009; Elena, *et al* 2011b).

In this study, a somatic mutation was identified in an initial screen of MPN patient *g*DNA samples using Sanger sequencing of *PHF6* targeting exons 9 and 10 (terminal exons). The mutation screening was subsequently extended using a custom amplicon

A.

Entry	Entry name	Protein names	Gene names
P84243	H33_HUMAN	**Histone H3.3**	**H3F3A** H3.3A H3F3 PP781 **H3F3B** H3.3B
Q71DI3	H32_HUMAN	**Histone H3.2**	**HIST2H3A** **HIST2H3C** H3F2 H3FM **HIST2H3D**
P68431	H31_HUMAN	**Histone H3.1**	**HIST1H3A** H3FA **HIST1H3B** H3FL **HIST1H3C** H3FC **HIST1H3D** H3FB **HIST1H3E** H3FD **HIST1H3F** H3FI **HIST1H3G** H3FH **HIST1H3H** H3FK **HIST1H3I** H3FF **HIST1H3J** H3FJ
Q16695	H31T_HUMAN	**Histone H3.1t**	**HIST3H3** H3FT
Q6NXT2	H3C_HUMAN	**Histone H3.3C**	**H3F3C**

B.

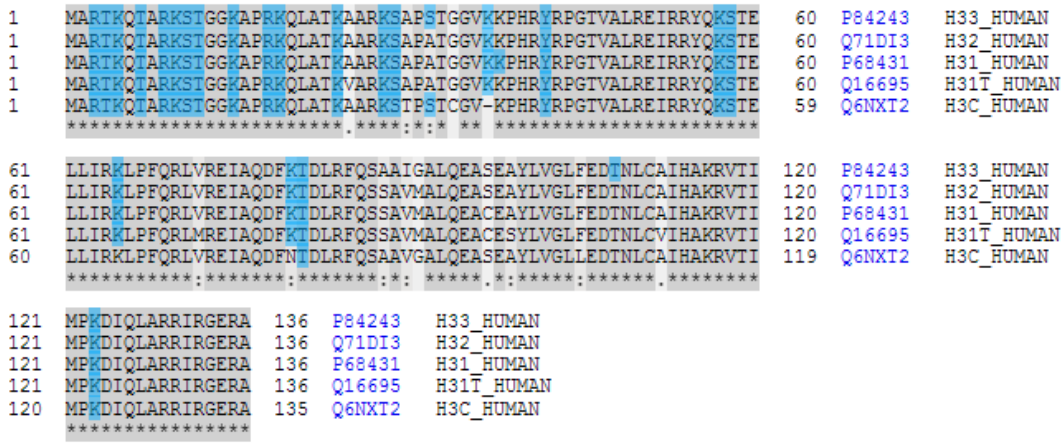


Figure 3.2. Conservation of human Histone H3 variants, H3.1, H3.2 and H3.3. A. Universal Protein Resource (UniProt) database annotations for the Histone H3 variants along with *Histone H3* genes encoding the variants. UniProt database predicts ~89.706% homology between the variants with 122 of the 135 amino acids identical between the proteins. **B**. Amino acid conservation across human Histone H3 variants. Blue boxes represent post-translationally modified residues and grey boxes indicate similarity of residues at specific positions. ***** indicates conserved amino acids; '**:**' indicates one dissimilar amino acid; '**.**' indicates >1 dissimilar amino acid.

based next generation sequencing approach to include the remaining coding exons of *PHF6* and also the 17 *histone H3* genes.

3.2 Materials and Methods:

3.2.1 Patients and genomic DNA (gDNA) preparation

Approval for this study was obtained from the Human Research Ethical Review Committees of the Royal Adelaide Hospital, The Queen Elizabeth Hospital and Flinders Medical Centre (*ethics approval number details have been given in* ***Chapter 2 section 2.2.1***). Informed written consent, peripheral blood, and buccal swab samples were collected from MPN patients, and samples were screened for the presence of *JAK2V617F* or exon 12 mutations of *JAK2*. U-MPN patients were individuals at an advanced or early stage of MPN (did not meet some of the diagnostic criteria and/or were being tested for mutations in *JAK2* exon 12/*MPL*), were *JAK2V617F*-negative, and had other underlying conditions which could have obscured diagnosis. PBMNC, granulocyte, and buccal cell isolation techniques have been described in ***Chapter 2***, ***section 2.2.4***. *g*DNA from cell populations was prepared using the QIAmp™ spin procedure (QIAGEN™, USA).

3.2.2 PCR amplification and Sanger sequencing of PHF6 exon 9 and exon 10

MPN patient *g*DNA samples (*n*=124) were investigated for mutations in *PHF6* exons 9 and 10. 77 of the 97 PV patient PBMNC or granulocyte *g*DNA samples were confirmed for the presence of *JAK2V617F*. Additionally, 9 ET (*n*=15), 3 PMF (*n*=5) and 2 U-MPN (*n*=7) patient *g*DNA samples had the *JAK2V617F* mutation. Primers and PCR amplification conditions for coding exons of *PHF6* (exon 9 and exon 10) (Genbank # NM_175629.1) were as previously described (Crawford, *et al* 2006). PBMNC or granulocyte *g*DNA sample amplification, purification, and sequencing were as described previously in ***Chapter 2, section 2.2.6***. Analysis of the sequences

was performed using SoftGenetics® Mutation Surveyor™ Version 3 and ChromasLite 2.1.1 software. The sequences containing novel variants were further verified using the Santa Cruz Human Genome Browser (UCSC genome browser), Ensembl|1000 Genomes (*hg18* and *hg19*), and the Exome Variant Server (EVS, NHLBI Exome sequencing Project).

3.2.3 Blast Forming Units Erythroid (BFUE) colony forming assays

EPO-independent and EPO-dependent Blast forming units-Erythroid (BFUE) were scored and isolated from PBMNC as described previously (Butcher, *et al* 2008). Briefly, cells were plated at a concentration of $1.6x10^6$-$1.6x10^7$ cells/mL in methylcellulose-based media (Methocult; Catalog number 4230; STEMCELL Technologies Australia Pty Ltd, Australia) supplemented with 20ng/mL recombinant human Interleukin-3 (rhIL-3; Peprotech, Abacus-ALS, Australia), 50ng/mL human stem cell factor (hSCF, Peprotech, Abacus-ALS, Australia) with or without 2U/mL human EPO (EPREX®40,000 IU/mL prefilled Syringe, JANNSEN-CILAG Pty. LTD., Australia). Erythroid colonies were haemoglobinised with a characteristic red colour and were composed of small circular cells. Colonies were scored and picked from culture on days 10 to 14 depending on the morphology and size of the colony.

3.2.4 Roche 454 sequencing platform

A. Sample preparation for PCR amplification

70 PV (64 were *JAK2V617F*$^+$), 20 ET (8 were *JAK2V617F*$^+$) and 10 PMF (3 were *JAK2V617F*$^+$) PBMNC or granulocyte *g*DNA samples (see ***Chapter2***, ***sections 2.2.5*** for preparation of *g*DNA) were separately diluted to 2ng/µL and pooled to generate 10 pools of *g*DNA samples (10 patients in each pool) (see ***Supplementary Table S3.1*** for pool summaries**)**. The *JAK2V617F*-negative PV *g*DNA samples were included in the study as these samples were in the process of being updated for the

presence of exon 12 mutations of *JAK2* and clinical information. The final concentration of *g*DNA in each sample pool was verified using the Nanodrop 2000 spectrophotometer V1.0 (Thermo Scientific, Australia).

B. Primer Design and PCR amplification

Pooled *g*DNA at a concentration of 2ng/µL was used as template in PCRs in a total reaction volume of 25µL. *Histone H3* primers were designed using online software tools (***Supplementary Table S3.2***) and the specificity of the primers was validated by amplifying and sequencing HEL *g*DNA. All *histone H3* primers were designed such that the amplicon size was 250-350 bp with overlapping amplicon regions so as to allow sequencing of a complete *histone H3* gene. Primers for *PHF6* coding exons (Genbank # NM_175629.1) were designed as described by Crawford and colleagues, and these primers fulfilled all criteria of primer design (Crawford, *et al* 2006). PCR products were sequenced using the BigDye® (Applied Biosystems, Life Technologies Australia Pty Ltd, Australia) terminator cycle sequencing protocol as described previously (Wen 2001). The sequences were verified using the UCSC genome browser's BLAT search tool.

C. Amplicon quantitation, pooling and purification

PCR products generated from each PCR were quantified using the MultiNA™ (Shimadzu-Biotech Australia Pty Ltd, Australia) microchip-based electrophoresis system. MultiNA™ viewer software was used to determine the size estimation and amplicon concentration. In addition, products were also detected by agarose gel electrophoresis and a concentration estimate was assessed for pooling the products.

D. Next generation amplicon library sequencing

A sequencing library of the amplicons covering all coding exons and splice junctions of *histone H3* genes and *PHF6* was generated by pooling the 59 amplicons in equal

concentrations to generate a sequencing master library at a concentration of 31ng/µL (***Figure 3.3*** shows the workflow). DNA for each library was purified using the QIAquick™ PCR Purification Kit. The purified amplicon pools were re-amplified to attach multiplex identification (MID) adaptors and re-purified (***Supplementary Figure S3.1*** shows an example of purified amplicon library from sample pool#1 before and after MID ligations). The 10 resultant MID-ligated libraries were amplified on DNA beads followed by bidirectional pyrosequencing using the Roche 454 GS Roche Junior technology. Roche raw sequencing data files were converted to '.bam' and '.bam.bai' file formats by Dr. Andreas Schreiber. Data analysis was performed using Integrative Genomics Viewer version 2.1.2 software [(Robinson, *et al* 2011), IGV2.1.2, Broad Institute] comparing to *hg19* genomic references. Data analysis was also performed in parallel with Dr. Andreas Schreiber, who assisted with further variant filtering.

3.3 Results

3.3.1 Sanger sequencing identified a PHF6R335fs mutation in exon 10 in a JAK2V617F^{+} PV patient

A screen of 124 MPN (97 PV, 15 ET, 5 PMF and 7 U-MPN) patients was initially screened using Sanger sequencing of *PHF6* exons 9 and 10 (spanning amino acids from position 278-365). This detected a heterozygous mutation in two independent PBMNC samples from a single patient (the first sample was obtained at the time of diagnosis in 2008 and the second in 2011). The mutation, a four nucleotide frameshift (fs) deletion causes a frameshift from the arginine (R) at position 335 (*PHF6*:*c.1004GAGA>GAGA_del, p.R335fs*; referred as *PHF6R335fs*) and results in

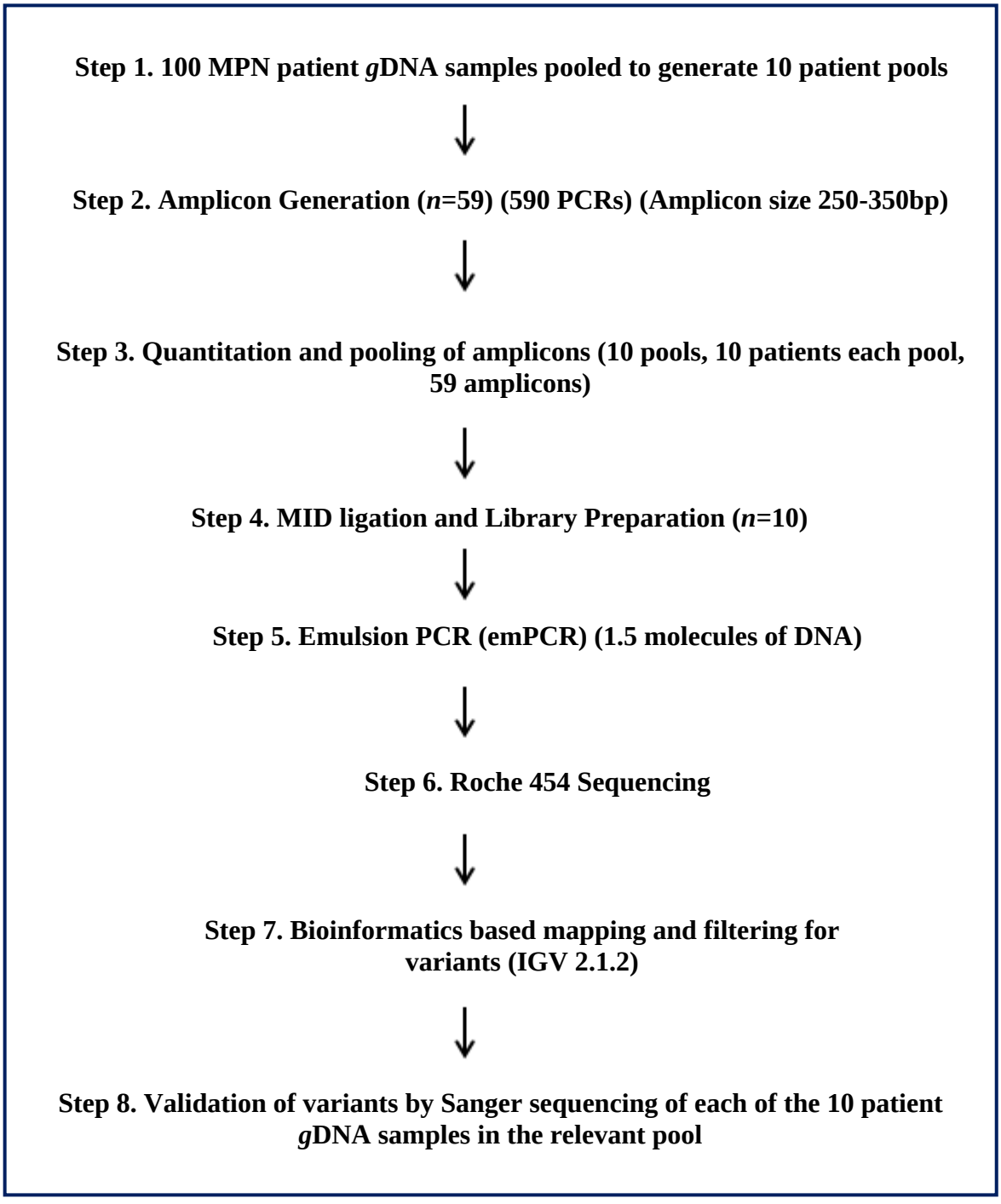


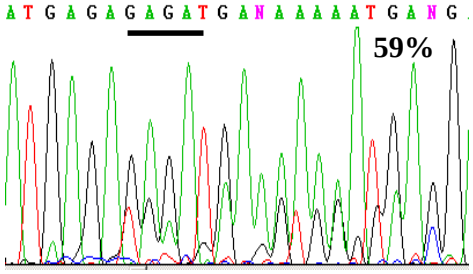
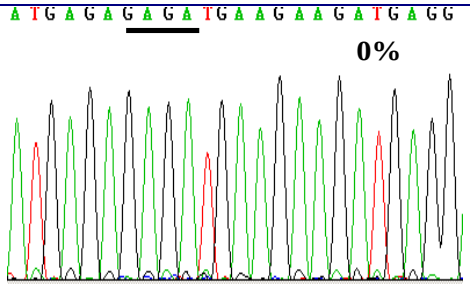
Figure 3.3. Workflow for Roche 454 sequencing.

a truncated protein (***Figure 3.4A***). The *PHF6R335fs* mutation was found to be somatic as it was not observed by Sanger sequencing of two independent buccal samples from this patient. Although this mutation has been reported in a single case of AML (Van Vlierberghe, *et al* 2011), the functional significance is undetermined since the mutation occurs distal to the second PHD domain at the C-terminus of the protein (***Figure 3.4B***). However, given that the mutation affects the last 30 amino acids, it can be speculated that the mutation may lead to altered potential of the mutant protein to localise in the nucleus, associate with the chromatin or mediate transcriptional regulation as recently suggested by Liu and colleagues (Liu, *et al* 2014; Lower, *et al* 2002). Additionally, the mutation may also affect the structural stability of the protein.

3.3.2 Patient PV108 BFUE colony genotyping

An investigation into the timing of acquisition of the *PHF6R335fs* mutation with regard to *JAK2V617F* was performed by genotyping BFUE colonies (*n*=15) isolated from patient PV108 (***Figure 3.5***). The presence of the *PHF6R335fs* mutation in colony analysis did not correlate with the presence or absence of *JAK2V617F*. Interestingly, colonies with a high allele load of *JAK2V617F* (presumed homozygous) were observed with no mutations in *PHF6*, and colonies with heterozygous mutations in *PHF6* were observed with low *JAK2V617F* allele load. While based on a small number of colonies, these observations are suggestive of mutations having arisen independently of each other with the homozygous *JAK2V617F* mutation arising in an independent clone that is wildtype for *PHF6*. The two scenarios have been illustrated in ***Figure 3.6***. Scenario 1 suggests that, in PV108, the *PHF6* mutation may have been acquired in a wildtype *JAK2* HSPC clone followed by the acquisition of *JAK2V617F*. Scenario 2 suggests *JAK2V617F* may have occurred first in a separate HSPC, followed by either the mutation in *PHF6* or

A.

Patient characteristics	PV108	Normal Reference range
Gender/Age at diagnosis (years)	F/82	NA
JAK2V617F mutation	positive	NA
Haemoglobin	115g/L	115-155 g/L
Red Blood Cell count	5.50 x 10^{12}/L	3.8-5.20 x 10^{12}/L
White Cell count	7.21 x 10^{9}/L	4.0- 11.0 x 10^{9}/L
Serum EPO	4mU/ml	3.3-16.6 mU/ml
Presence of EPO-independent erythroid colonies (e-BFUE)	Yes	NA
sequence trace depicting the heterozygous *PHF6R335fs* change in PV108 PBMNC and *JAK2V617F* %		
PHF6 sequence trace in buccal And *JAK2V617F* %		

B.

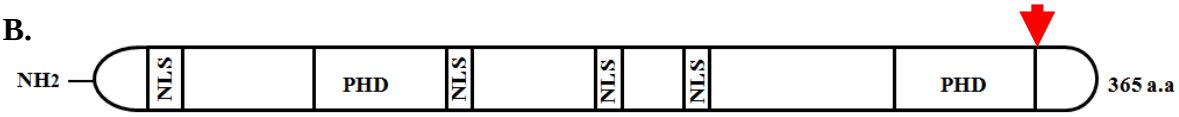
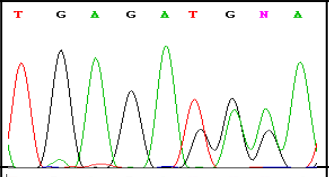
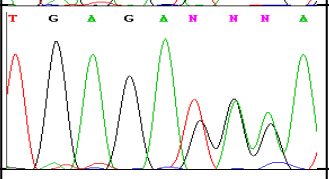
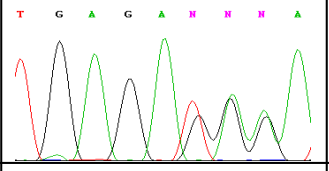
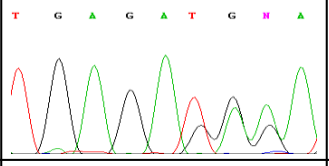
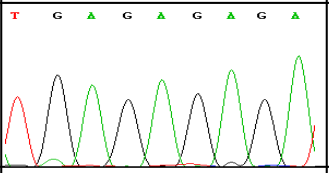
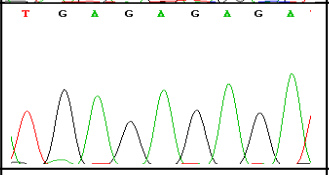
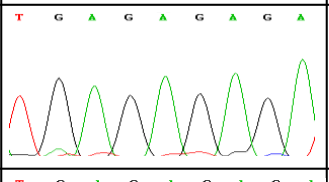
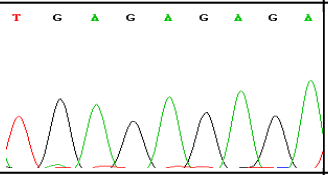
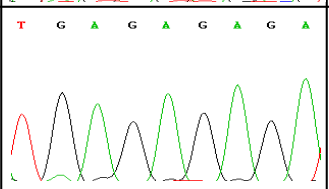
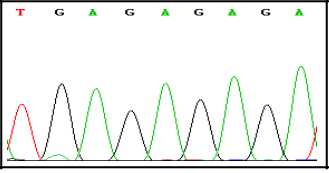


Figure 3.4. A. Clinical characteristics of PV108 with the *PHF6*R335fs variant. PV108 was diagnosed with PV based on the 2008 WHO diagnostic criteria. Before 2012 this patient was having regular venesections to manage PV. In 2012 the management therapy included hydroxyurea and anagrelide to control high platelet count and neutropenia. Values in normal reference range column indicate normal female range used at diagnosis. Patient values are from the year of diagnosis (2008). The sequence traces for PBMNC and buccal *g*DNA for *PHF6* is suggestive of the *R335fs* mutation being somatic. The *JAK2V617F* allele loads have also been depicted above the sequence trace for each sample. Granulocyte samples from this patient were not amplifiable and additional samples were not available. NA, not applicable. **B. Structure of PHF6 and the location of the p.R335fs** as depicted by the arrow.

3.5 A

Colony #	*PHF6R335fs* T G A G A G A G A	*JAK2V617F%*
BFUE#1		0
BFUE#2		0
BFUE#3		0
BFUE#4		0
BFUE#5		82
BFUE#6		92
BFUE#7		91
BFUE#8		94
BFUE#9		100
BFUE#10		87

Continued on the next page

3.5 B

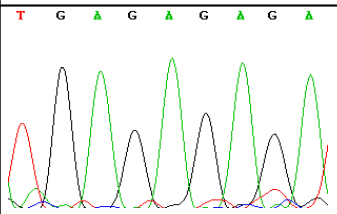
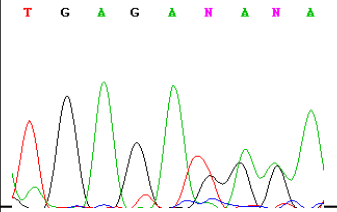
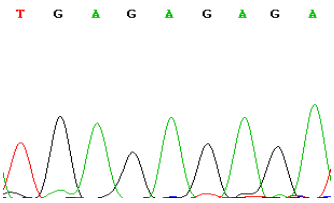
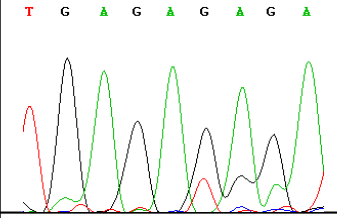
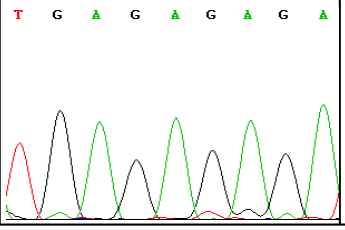
Colony #	*PHF6R335fs* T G A G A G A G A	*JAK2V617F%*
e-BFUE #1		23
e-BFUE #2		25
e-BFUE #3		72
e-BFUE #4		19
e-BFUE #5		65

Figure 3.5. Genotyping of BFUE from PV108. PBMNC from PV108 were plated in methylcellulose media (**A**) with and (**B**) without EPO and colonies isolated. The colonies were genotyped for *JAK2V617F* and *PHF6R335fs* using Sanger sequencing. The sequence traces are representative of two independent PCRs from each colony.

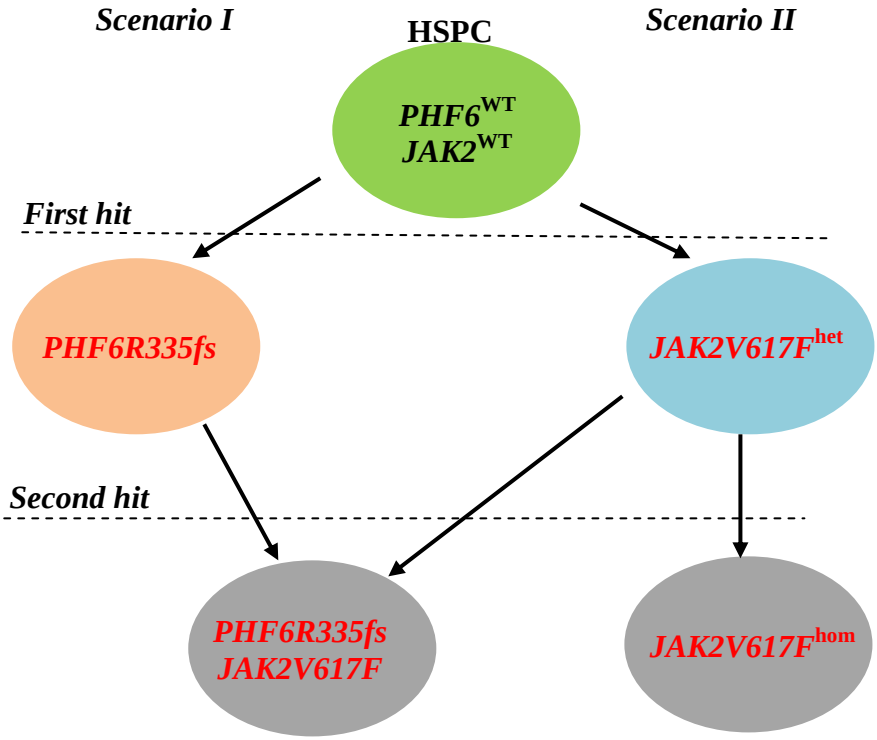


Figure 3.6 Possible scenarios for BFUE colonies with and without *PHF6R335fs*. See text for details. WT, wildtype; *JAK2V617F*^{het}, heterozygous *JAK2V617F*; *JAK2V617F*^{hom}, homozygous *JAK2V617F*.

another *JAK2V617F* event. These two scenarios may explain the allele burden observed in colonies grown with EPO; for example, a mutually exclusive presence of either a heterozygous *PHF6* mutation (BFUE#1-4) or homozygous *JAK2V617F* mutations (BFUE#5-10). With regards to the allele burden of BFUE# 5-8 and 10, inconsistencies in the picking of single cell colonies most likely explain the range of allele load (82%-97%) observed. However, given that the allele burdens were close to approximately 100%, it was inferred that these were homozygous for *JAK2V617F*%.

With regards to e-BFUE#1-5, the technical issues described above may not entirely explain the low *JAK2V617F* allele burdens in some colonies. It is possible that additional mechanisms may need to be evaluated such as the possibility of acquisition of an additional chromosome 9. This would result in either an additional wildtype *JAK2* or *JAK2V617F*, and could explain the allele loads for colonies 1, 2 and 4 (presence of two wildtype *JAK2* alleles and one *JAK2V617F* allele) and, 3 and 5 (presence of one wildtype *JAK2* and two *JAK2V617F* alleles) respectively.

It may also be worth noting that none of the colonies displayed homozygous *PHF6R335fs* possibly suggesting that homozygous mutations in *PHF6* may not be tolerated.

3.3.3 Next generation sequencing approach: mapping and filtering for potential variants

The above screen was extended to the remaining exons of *PHF6* and to 17 *histone H3* genes using a next generation sequencing approach. 100 MPN (70 PV, 10 PMF and 20 ET) PBMNC and granulocyte *g*DNA samples were pooled to make 10 pools. PV, ET, and PMF patient gDNA were grouped separately to make up pools of only

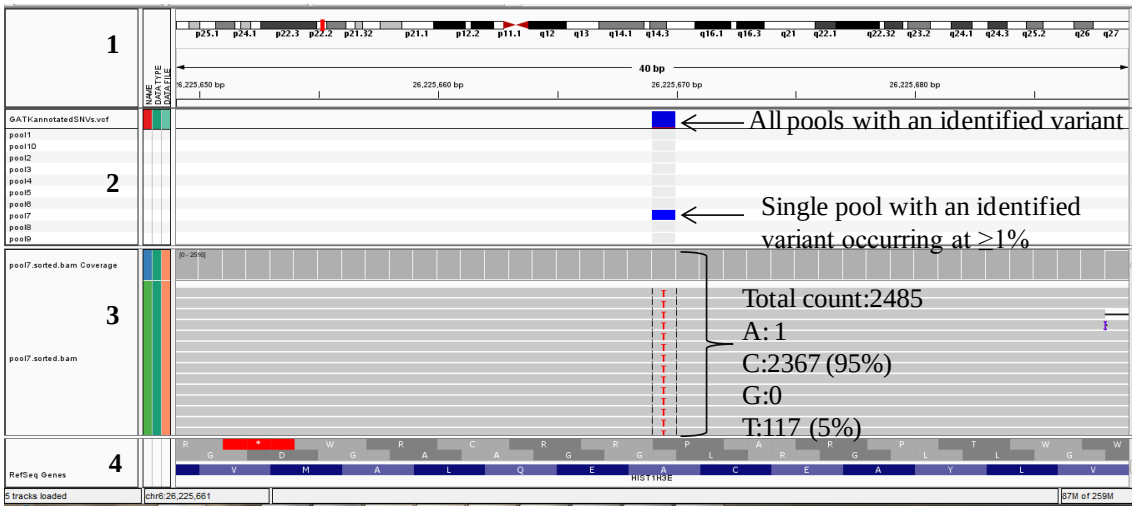
one disease type. PBMNC and granulocyte samples were also kept separate wherever possible. It is important to note that the individual *g*DNA samples used overlapped with the samples already screened in ***section 3.3.1*** and included a PV108 granulocyte *g*DNA sample. Samples were chosen depending on the availability of germline *g*DNA to allow for validation of possible changes detected through sequencing. Primers were designed to amplify 250-350bp of a genomic region, which resulted in 59 PCRs to amplify the 17 *histone H3* genes and 9 *PHF6* coding exons from each of the 10 *g*DNA pools (total PCRs=590). Each of the 59 amplicons generated for a sample pool were combined in equal concentrations and purified, following which MID were ligated to all the amplicon pools. These MID-tagged amplicon pools (or amplicon libraries) were then sequenced using the Roche 454 sequencing platform (see ***Figure 3.3*** for the workflow). The number of passed filter beads were visualised to determine the success of the sequencing runs. The read lengths and output quality were also assessed by calculating the percent of control DNA match to the reference sequences. All the 10 libraries contained reads predominantly in the range of 250-300bp which represented most of the input amplicon DNA (***Supplementary Figure S3.2*** represents an example output). Quality scores were assigned to all bases and were generated using a Phred base-calling algorithm (***Supplementary Figure S3.3*** shows a plot of quality scores of the reads generated using the Roche 454 sequencing technology). Read files above the quality score of 15 were mapped to the human genome (*hg19*). SNP and variant calling was performed and potential novel variants were further analysed using the Polymorphism Phenotype (PolyPhen-2) and Sorting-Intolerant-From-Tolerant (SIFT) algorithms, to indicate significance. Three inclusion criteria were used to accept valid variant sequences for further analysis: 1. variants found in either 1% of forward or reverse sequences, 2. at least 1% of sequences had to have the variant and; 3. at least 20 times sequence coverage of the amplicons with

the suspected novel variant of the total reads. Two exclusion criteria were also used: 1. variants detected in a homopolymeric tract of ≥4, and 2. mutations found in the last nucleotide of the sequence with frequencies of less than 1%. A sequencing run of healthy individual DNA samples was not performed as there was a specific interest in detecting *somatic* non-synonymous changes occurring in MPN patients and any potential somatic variants identified would be confirmed by Sanger sequencing of the individual germline *g*DNA samples, which was available for a number of patients.

3.3.4 Identification and validation of novel variants

Based on the above outlined criteria, a number of potential candidate variants were identified using the IGV 2.1.2. The IGV 2.1.2 browser allowed visual analysis of reads aligned to *hg19* aligned reads (forward and reverse reads) along with chromosome position. The IGV 2.1.2 browser also allowed viewing of specific pool-related details such as pool ID, depth of reads, and depth of a single base at specific position (see ***Figure 3.7A***). So that the validation process was manageable, 71 of the total 94 potential variants were selected. These included the mutations identified in the studies performed by Schwartzentruber and colleagues, and Wu and colleagues (Wu, *et al* 2012);(Schwartzentruber, *et al* 2012) [these have been listed in ***Supplementary Table S3.3***]. From this screen a single heterozygous coding region variant was validated by Sanger sequencing in a *histone H3* gene; *HIST1H3E:c.287 C>T; p.A96V* ***(Figure 3.7A and B)*** in the PBMNC *g*DNA sample from a PV patient (PV44 in Pool #7, see ***Figure 3.8***). The somatic or germline status of this variant was not ascertained as unaffected tissue was not available for this patient. However the PBMNC sequence trace showing a 50% presence of the novel variant which was suggestive of the germline presence of the variant. This novel variant was not detected in EVS, dbSNP or Ensembl|1000 Genomes databases. However, time

A.



B.

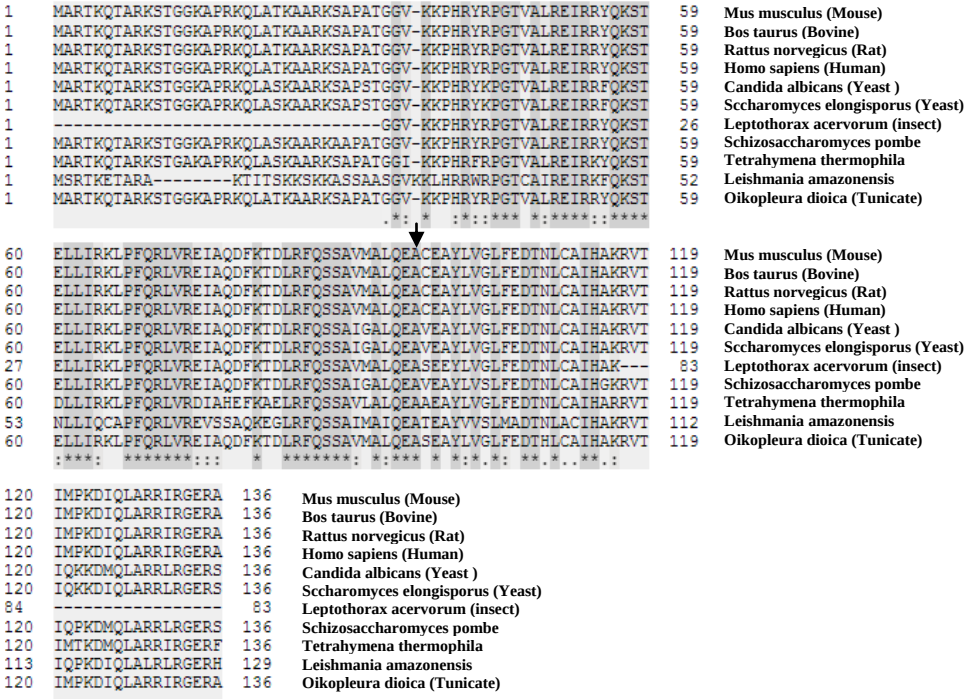
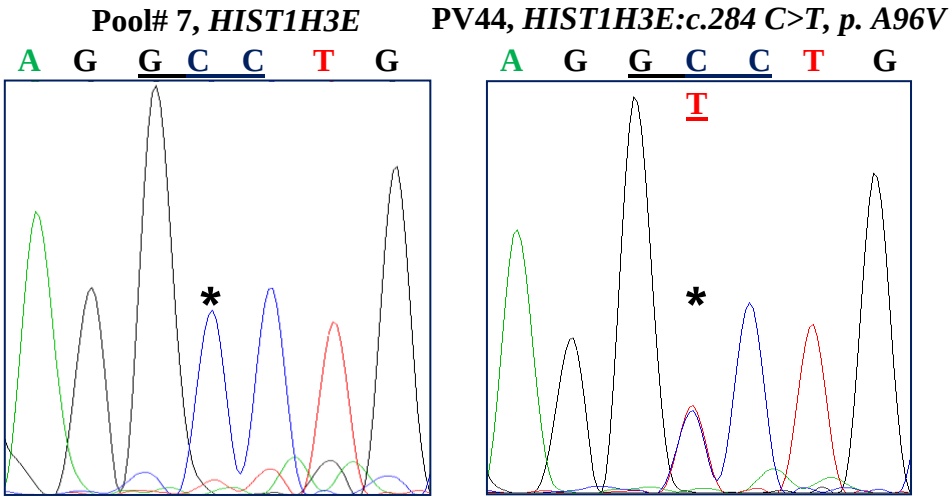


Figure 3.7. Identification of a coding variant in *HIST1H3E* in the sequencing output from pool #7. A. view in the IGV.2.1.2 genome browser which was used to visualise and analyse mapped reads from sequencing runs. **1** shows the position of the gene on the chromosome, **2** depicts the pools 1 to 10 and blue box indicates the identification of a heterozygous variant in a pool. Here the blue box represents a variant in pool #7. **3** indicate reads from a pool and also includes statistics such as read quality and depth. The statistics for pool #7 depicted that the number of reads with the novel substitution of 'T' constituted approximately 5% of the total 2485 reads at that specific position. **4** shows the RefSeq gene and the codon details for the gene *HIST1H3E*. **B.** Shows the conservation of the alanine at position 95 in Histone H3.1 across a number of species and has been highlighted with an arrow. The residues highlighted in grey are conserved residues with high conservation (indicated with *).

A.



B.

Patient characteristics	Patient PV44	Normal Reference range
Gender/Age at diagnosis	F/75	NA
JAK2V617F mutation	positive	NA
Haemoglobin	137g/L	115-155g/L
Red Blood Cell	3.56 x 10^{12}/L	3.8-5.20 x 10^{12}/L
White Cell	5.12 x 10^{9}/L	4.0 - 11.0 x 10^{9}/L
Serum EPO	1.7 mU/ml	3.3-16.6 mU/ml
Presence of EPO-independent BFUE	Yes	NA

Figure 3.8. *HIST1H3E* sequence trace and clinical characteristics of patient PV44. A. *HIST1H3E* sequence trace of Pool#7 and PV44 PBMNC depicting the affected nucleotide (by *). Sanger re-sequencing of Pool#7 did not detect the change however sequencing of the patient PBMNC sample identified the heterozygous *c.287 C>T, p.A96V* variant. Germline *g*DNA sample was not available for this patient. **B. Clinical characteristics of PV44**. PV44 was diagnosed with PV. Values in the right column represent the normal female range used at diagnosis. This patient was diagnosed in the year 2009 and the values in the table depict the values at the time of diagnosis.

constraints did not permit functional assessment of this change and therefore, currently it is difficult to interpret the significance of this mutation. No further variants of *PHF6* identified in the Roche 454 sequencing were confirmed by Sanger sequencing.

3.4 Discussion:

The *PHF6* tumour suppressor gene encodes a PHD protein containing four nuclear localization signals and two zinc-finger domains (Voss, *et al* 2007) with a proposed role in transcriptional regulation and/or chromatin remodelling (Lower, *et al* 2002). Since the identification of mutations in *PHF6* in T-ALL, adult AML, and blast-crisis Ph^+ CML [frequency of approximately 2.4%, (Li, *et al* 2012)], one of the aims of this study was to perform a screen in MPN for mutations in *PHF6*. Identification of novel coding variants in genes encoding for histone H3 was the second aim of the study. Standard Sanger sequencing and Roche 454 sequencing identified a single somatic variant in *PHF6* (in PV108) and a coding region variant in *HIST1H3E* in a separate patient (PV44). The *PHF6R335fs* mutation was identified in a *JAK2V617F*$^+$ chronic phase PV patient and was heterozygous (*PHF6* is located on the X chromosome).

The current study is the first to report a screen of *PHF6* and the identification of a mutation in chronic MPN. This study suggests that mutations in *PHF6* are rare but do occur in chronic phase MPN. To further understand the relevance of this mutation, functional studies will be important.

This is also the first known study to report a screen of *histone H3* genes in MPN. This screen identified a novel coding variant in the histone H3.1 (*HIST1H3E*); involving substitution of alanine at position 96 to valine (a list of known somatic coding region variants from the COSMIC database are listed in ***Table 3.1***). Somatic

mutations in *histone H3* genes have been identified and associated with aggressive paediatric gliomas and a number of other cancers (Behjati, *et al* 2013; Huether, *et al* 2014; Je, *et al* 2013; Je, *et al* 2014; Schwartzentruber, *et al* 2012; Wu, *et al* 2012). Recurrent mutations in gliomas were observed in two genes; *H3F3A* and *HIST1H3B*. Although this study in this chapter did not validate changes in either of these *histone H3* genes, the Roche 454 sequencing did identify a 1% occurrence of G35R variant in *H3F3A* in pool#10. While this was not detected in the disease *g*DNA samples from that pool using Sanger sequencing, we cannot exclude that the mutation may be present in a patient sample at a level below the detection of Sanger sequencing (that is ≤10%). To exclude this possibility it will be necessary to establish a sensitive assay for detecting specific mutations in individual samples from this pool.

The Roche 454 sequencing approach was undertaken in this study because of the increased depth of coverage achieved when compared with Sanger sequencing. The flexibility with the use of MID-tagged pools allowed for pooling of 59 amplicons generated from 10 patients to achieve a minimum read depth of >200 reads. Roche 454 sequencing protocol allowed 100 samples (total of 590 PCR products) to be sequenced in 10 sequencing runs. Although these advantages were significant, there were a number of issues with the output data and filtering process. Given that *Histone H3* genes are highly conserved with a high degree of sequence similarity, the next generation sequencing output depicted a high occurrence of homopolymer tracts in the sequencing output. Occurrence of homopolymer tracts in a sequencing output can result in false-positive nucleotide deletions/insertions. This can be problematic if small size reads (approximately 50bp) map to multiple genes which made the filtering (for variants) problematic. The ability to differentiate between errors and

Table 3.1

Gene	**Somatic non-synonymous coding region variants of histone H3**					
HIST1H3A	c.60G>T p.Q20H	c.125A>G p.Y42C	c.190C>A p.R64S	c.236T>G p.F79C	c.247C>A p.L83M	c.331T>A p.T119A
	c.377A>C p.Q126P					
HIST1H3B	**c.83A>T p.K28M**	c.95C>A p.A32D	c.97A>G p.T33A	c.139G>A p.V47M	c.151G>C p.E51Q	c.220G>A p.E74K
	c.274G>A p.A92T	c.282G>T p.Q94H				
HIST1H3C	c.7C>G p.R3G	c.16C>G p.Q6E	c.34A>G p.T12A	c.69T>A p.T23T	c.91C>T p.P31S	c.121C>T p.R41C
	c.151G>A p.E51K	c.308G>A p.G103E	c.361A>C p.M121L	c.379C>A p.L127M	c.396_397insG p.E134fs*>4	
HIST1H3D	c.7C>G p.R3G	c.79C>G p.R27G	c.110A>T p.K37M	c.131C>A p.P44H	c.178G>T p.E60*	c.199C>T p.P67S
	c.240G>C p.K80N	c.244G>C p.D82H				
HIST1H3E	c.5C>T p.A2V	c.54C>A p.R18R	c.142G>T p.A48S	c.200C>A p.P67Q	c.340C>A p.H114N	c.350G>A p.R117H
HIST1H3F	c.115C>A p.P39T	c.142G>A p.A48T	c.168G>T p.Q56H	c.195G>C p.K65N	c.237C>G p.F79L	c.292G>A p.E98K
	c.299A>T p.Y100F					
HIST1H3G	c.127C>T p.R43C	c.130C>A p.P44T	c.137C>A p.T46N	c.161G>A p.R54H	c.166C>T p.Q56*	c.290G>T p.C97F
	c.302_303insACTT p.V102fs*6					
HIST1H3H	c.10A>G p.T4A	c.18G>T p.Q6H	c.25C>G p.R9G	c.80G>A p.R27Q	c.81G>T p.R27R	c.109A>G p.K37E
	c.110A>T p.K37M	c.250C>T p.R84C	c.259A>T p.S87C	c.382G>C p.A128P	c.406G>T p.A136S	

Continued on the next page....

Gene	**Somatic non-synonymous coding region variants of histone H3**					
HIST1H3I	c.79C>T p.R27C	c.97A>T p.T33S	c.218G>T p.R73L	c.240G>C p.K80N	c.244G>A p.D82N	c.292G>A p.E98K
	c.298T>C p.Y100H	c.322A>G p.T108A	c.371A>G p.D124G	c.411A>C p.*137Y		
HIST1H3J	c.37G>T p.G13C	c.47C>T p.A16V	c.95C>T p.A32V	c.202T>C p.F68L	c.332G>C p.C111S	c.364C>A p.P122T
	c.386G>A p.R129H					
HIST2H3D	c.23C>T p.A8V	c.24C>T p.A8A	c.82A>T p.K28*	c.120C>G p.H40Q	c.127C>T p.R43W	c.166C>T p.Q56*
	c.355A>G p.T119A	c.381G>T p.L127F				
HIST3H3	c.103G>A p.G35S	c.137C>T p.T46M	c.149G>A p.R50H	c.218G>A p.R73H	c.220G>A p.E74K	c.305T>C p.V102A
	c.385C>T p.R129C					
H3F3A	c.7C>T p.R3C	c.26G>A p.R9H	c.60A>C p.Q20H	**c.83A>T p.K28M**	**c.103G>A p.G35R**	**c.104G>T p.G35V**
	c.110A>G p.K37R					
H3F3B	c.52_53insC p.R18fs*80	c.79A>G p.R27G	c.128G>A p.R43K	c.260G>T p.S87I	c.262G>A p.A88T	c.315C>G p.F105L
	c.358A>G p.I120V	c.379T>G p.L127V				
H3F3C	c.89C>T p.T30I	c.97A>G p.T33A	c.103G>A p.G35R	c.118C>T p.R40C	c.119G>A p.R40H	c.140C>A p.A47E
	c.145C>T p.R49*	c.169T>A p.S57T	c.239C>T p.T80I	c.240T>C p.T80T	c.265G>A p.V89I	c.289G>A p.E97K
	c.364A>G p.K122E	c.386G>T p.R129L	c.392G>T p.R131L	c.401G>A p.R134K		

Table 3.1. Somatic non-synonymous coding region variants of histone H3. The COSMIC database was used to identify variants in the 17 *histone H3* genes. *Histone H3* genes; *HIST2H3A* and *HIST2H3C* were not found in the database possibly due to the lack of any coding region variants in these genes. Changes highlighted in bold have been identified in glioma.

true changes will need to be optimised to improve detection of variants. Another important variable in the experimental approach was the possible preferential amplification of samples within a *g*DNA pool. This may explain why PV108 *PHF6R335fs* was not detected in the Roche 454 sequencing output data. It is important to note that the granulocyte *g*DNA sample used for Roche 454 sequencing did not amplify when amplified by standard PCR followed by Sanger sequencing, suggesting that the granulocyte sample may have been of low quality or had impurities which prevented successful amplification.

In conclusion, Roche 454 sequencing allowed the rapid sequencing of 100 MPN patients (for 59 genomic regions) at a depth ranging from 250 to 3500 reads. However, preferential *g*DNA amplification, presence of homopolymer tracts in sequencing reads and difficulties with the variant filtering process limited the success of this approach. This approach did detect a large number of potential low-level novel variants in genes encoding PHF6 and histone H3. Given the detection of a number of coding variants at a low level and the difficulties in using Sanger sequencing, further validation with sensitive and specific assays would be required to determine significance of these coding variants.

3.5 Supplementary data:

Patient pool	**POOL 1**	**POOL 2**	**POOL 3**	**POOL 4**	**POOL 5**
ID of Patient *g*DNA Samples in each pool ↓	PMF4 granulocyte	ET20 granulocyte	PV23 granulocyte	PV105 granulocyte	PV124 granulocyte
	PMF5 granulocyte	ET21 granulocyte	PV25 granulocyte	PV108 granulocyte	PV125 granulocyte
	PMF6 granulocyte	ET23 granulocyte	PV34 granulocyte	PV109 granulocyte	PV126 granulocyte
	PMF7 granulocyte	ET24 granulocyte	PV38 granulocyte	PV110 granulocyte	PV129 granulocyte
	PMF8 PBMNC	ET25 granulocyte	PV40 granulocyte	PV111 granulocyte	PV133 granulocyte
	PMF9 granulocyte	ET26 granulocyte	PV47 granulocyte	PV114 granulocyte	PV134 granulocyte
	PMF3 granulocyte	ET27 granulocyte	PV57 granulocyte	PV115 granulocyte	PV136 granulocyte
	PMF 10 granulocyte	ET28 granulocyte	PV97 granulocyte	PV116 granulocyte	PV137 granulocyte
	PMF12 granulocyte	ET29 granulocyte	PV101 granulocyte	PV120 granulocyte	PV139 granulocyte
	PMF13 granulocyte	ET30 granulocyte	PV104 granulocyte	PV123 granulocyte	PV140 granulocyte
Patient pool	**POOL 6**	**POOL 7**	**POOL 8**	**POOL 9**	**POOL 10**
ID of Patient *g*DNA Samples in each pool ↓	ET1 PBMNC	PV29 PBMNC	PV63 PBMNC	PV7 PBMNC	PV76 PBMNC
	ET2 PBMNC	PV35 PBMNC	PV67 PBMNC	PV66 PBMNC	PV77 PBMNC
	ET3 PBMNC	PV36 PBMNC	PV106 PBMNC	PV68 PBMNC	PV78 PBMNC
	ET4 PBMNC	PV37 PBMNC	PV99 PBMNC	PV83 PBMNC	PV80 PBMNC
	ET5 PBMNC	PV41 PBMNC	PV54 PBMNC	PV88 PBMNC	PV91 PBMNC
	ET6 PBMNC	PV42 PBMNC	PV96 PBMNC	PV89 PBMNC	PV60 PBMNC
	ET7 PBMNC	PV44 PBMNC	PV62PBMNC	PV90 PBMNC	PV55 PBMNC
	ET8 PBMNC	PV48 PBMNC	PV53 PBMNC	PV92 PBMNC	PV45 PBMNC
	ET9 PBMNC	PV49 PBMNC	PV8 PBMNC	PV95 PBMNC	PV52 PBMNC
	ET10 PBMNC	PV58 PBMNC	PV82 PBMNC	PV103 PBMNC	PV107 PBMNC

Supplementary Table S3.1. 10 patient pools and patient *g*DNA samples comprising each pool. Similar MPN subtypes and *g*DNA samples were pooled in one pool wherever possible.

Supplementary Table S3.2. Primers and conditions for amplification of the 17 *histone H3* genes.					
Gene-amplicon#	Accession ID	Forward Primer (bp) Reverse Primer (bp)	Primer Tm (C°)	Genome position	Product size (bp)
H3F3A-1	NM_002107	cctccattgtgtgtgattgg gctcctcacctccatttctg	59.81 59.8	chr1:226250337-226250639	322
H3F3A-2	NM_002107	atcgtggcaggaaaagttgt caagagagactttgtcccattttt	59.6 60.04	chr1: 226251962-226252221	260
H3F3A-3	NM_002107	tttctttgaagctgcccact ctgttcttggtacacaacaggaat	59.99 59.49	chr1: 226253281-226253567	287
H3F3A-4	NM_002107	ttgggtcttaactattggaaataaca cgttcagaactaccaataacagga	59.31 59.61	Chr1: 226258993-226259263	271
H3F3A-5	NM_002107	gaacaagtttcagcggttca tggaatataattgtagggcaagg	59.13 59.18	Chr1: 226259433- 226259749	317
H3F3A-6	NM_002107	tgccaaaagacatccagcta atgctttacgtctccgcatt	59.42 59.72	Chr1: 226259131- 226259408	278
H3F3A-7	NM_002107	caggtattggcagtttttcca tgtacttgcatgtaggacaactca	59.98 59.78	Chr1: 226259339 - 226259636	298
H3F3C-1	NM_001013699	gttggaccagtgtcggttct aacgacgaatctctcgaagc	60.01 59.58	Chr12: 31944941 - 31945209	269
H3F3C-2	NM_001013699	aggaaaagcacccccctctac gggcatgatggtgactctct	59.58 60.08	Chr12: 31944738 - 31945022	285
H3F3C-3	NM_001013699	tggtgggtctgttggaagat tcctgagcaacagtgctgac	60.36 60.19	Chr12: 31944507 - 31944802	296
H3F3C-4	NM_001013699	ggtccttttgcaataaaactgg	59.88	Chr12: 31944099 - 31944380	282

		atgacaagactccccaccac	59.82		
HIST1H3A-1	NM_003529	ggcgtatttgcgctagtgtt	60.29	Chr6: 26020670 - 26020977	308
		ctctggaaacgcaggtctgt	60.44		
HIST1H3A-2	NM_003529	ttgcgcaggactttaaaaca	59.49	Chr6: 26020941 - 26021250	310
		aacgaaaagtgctcccaatg	60.11		
HIST1H3B-1	NM_003537	tttcctttcctccacagacg	60.22	Chr6: 26032027 -26032340	313
		gctctggaagcgaagatcg	61.18		
HIST1H3B-2	NM_003537	agaaatcgcccaagacttca	59.81	Chr6: 26031752 - 26032070	319
		gggggtccttagaaatttgttc	60.06		
HIST1H3C-1	NM_003531	aggccacttgctctcagttc	59.60	Chr6:26045589 - 26045886	298
		aggtcggttttgaagtcctg	59.18		
HIST1H3C-2	NM_003531	gaaatccgtcgctaccagaa	60.21	Chr6:26045789 - 26046134	346
		caaccgaatgggttacaactg	60.27		
HIST1H3D-1	NM_003532	tgccttgtccaatcagaaca	60.24	Chr6:26197265 - 26197575	311
		ctagacgctggaatggcagt	60.42		
HIST1H3D-2	NM_003532	ctgccattccagcgtctagt	60.42	Chr6:26196966 - 26197283	318
		ccagtttacggacagcaaca	59.76		
HIST1H3D-3	NM_003532	accaaggctgctcgaaagag	61.98	Chr6:26197145 - 26197412	268
		cgcataggttggtgtcctca	62.06		
HIST1H3E-1	NM_003532.2	tcgtccacaatccaatcaga	60.05	Chr6:26225298 - 26225627	330
		tcggtcttgaagtcctgagc	60.53		
HIST1H3E-2	NM_003532.2	ggtgcgagaaatagctcagg	59.18	Chr6: 26225595 - 26225943	349
		gagggaactgaatgccactatt	59.49		
HIST1H3E-3	NM_003532.2	actaaggcagctcgcaagag	59.92	Chr6: 26225449 - 26225725	277

Supplementary Table S3.2

		catgaatagcgcacaggttg	60.28		
HIST1H3F-1	NM_021018.2	ttcggaccaatcagaaggac	60.05	Chr6: 26250629 - 26250922	294
		ggaatggtagcttgcgaatc	59.67		
HIST1H3F-2	NM_021018.2	aggacttcaagaccgacctg	59.3	Chr6:26250304 - 26250604	301
		cctccaccctccttgtttta	59.02		
HIST1H3F-3	NM_021018.2	gctacaggcctggtactgtc	57.41	Chr6:26250693 -26250459	254
		gatgtccttgggcatgatag	57.96		
HIST1H3G-1	NM_003534.2	gccaagtggctgagtttttc	59.86	Chr6:26271377 - 26271648	272
		aagtcctgagcgatttctcg	59.57		
HIST1H3G-2	NM_003534.2	agattcgccgctatcagaag	59.57	Chr6:26271184 - 26271461	278
		attactgcccggaaacctct	59.96		
HIST1H3G-3	NM_003534.2	ggctctttgaggataccaacc	59.95	Chr6:26271056 - 26271305	250
		tttagccaatcaaatgaggactt	59.17		
HIST1H3H-1	NM_003536.2	ttgggtaggccttcagcata	60.6	Chr6:27777748 - 27778032	285
		gctcggtggacttctggtag	59.87		
HIST1H3H-2	NM_003536.2	ctatcggcctggtacagtgg	60.27	Chr6:27777974 - 27778227	254
		ggatgtccttgggcatgata	58.96		
HIST1H3H-3	NM_003536.2	gctctttgaggacaccaacc	59.7	Chr6:27778160- 27778409	250
		gattgtataacgtctggaaatgg	57.59		
HIST1H3I-1	NM_003533.2	acgcgagggcagctataaat	61.09	Chr6:27839879- 27840170	292
		accaagcgctgaaaaggtag	59.52		
HIST1H3I-2	NM_003533.2	agaagtcgaccgagctgcta	60.3	Chr6:27839879- 27839927	274
		ggacagacttcttgggctga	60.39		

Supplementary Table S3.2

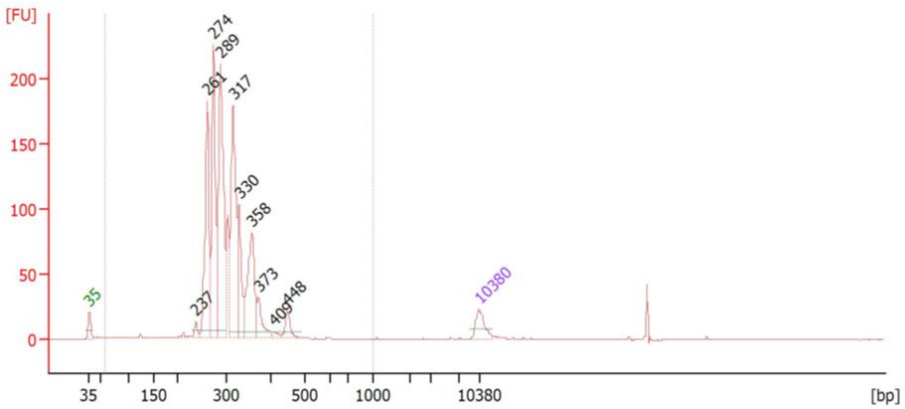
HIST1H3I-3	NM_003533.2	cctacctggtggggctattt	60.2	Chr6: 27839543- 27839798	256
		taccacgttttgtccacagc	59.62		
HIST1H3J-1	NM_003535	ttcagctgcttgctttcaga	60.01	Chr6: 27858316- 27858607	292
		gaaacgaaggtcggttttga	60.09		
HIST1H3J-2	NM_003535	gagatccgccgttatcagaa	60.18	Chr6: 27858114 - 27858420	307
		tttggaagcttggaagcaat	59.82		
HIST1H3J-3	NM_003535	tctgtgctattcacgccaag	60.01	Chr6: 27857993 - 27858242	250
		gaatgaacaacccttgacgac	59.44		
HIST2H3D-1	NM_001123375.2	ttttcaggcccaatcaagag	60.18	Chr1:120904601 - 120904863	264
		ggatcagcagctccgtagac	59.98		
HIST2H3D-2	NM_001123375.2	gcaggactttaagacggacctg	62.37	Chr1: 120904901 -120905170	264
		gagagccggtacagctgctc	62.58		
HIST2H3D-3	NM_001123375.2	gatcgcgcaggactttaaga	60.49	Chr1: 120904895 -120905165	265
		ccggtacagctgctcttgat	60.42		
HIST3H3-1	NM_003493	acgcggcaagctctataaaa	60.01	Chr1: 228612830 - 228613104	275
		aacttgcggattagcagctc	59.62		
HIST3H3-2	NM_003493	ccgctaccagaagtccactg	60.84	Chr1: 228612591 - 228612868	278
		cctctcaggtggcgagatag	59.97		
HIST3H3-3	NM_003493	gctgtttgaggacaccaacc	60.55	Chr1: 228612447 - 228612718	272
		agcctcccaccctaatcaga	60.98		
HIST2H3A/C-1	NM_001005464.2	ccaatcgggaaaagactgtg	60.49	Chr1: 149812549 - 149824397	282
		gctccgtggacttctggtag	59.87		
HIST2H3A/C-2	NM_001005464.2	cgctggtaagcctgtgtttt	60.3	Chr1: 149824189 - 149824456	268

Supplementary Table S3.2

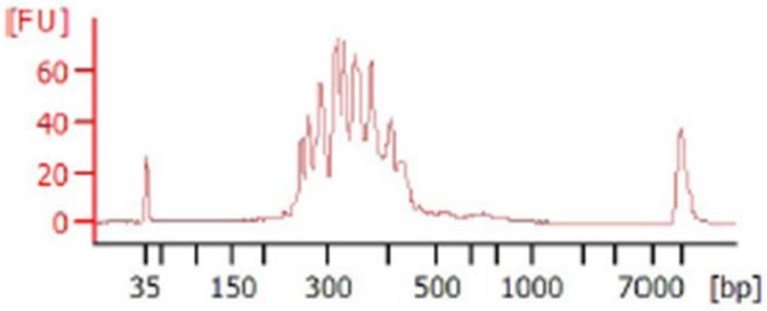
		cttaaagtcctgcgcgatct	59.61		
HIST2H3A/C-3	NM_001005464.2	gcccttccagcggctggt	68.25	Chr1: 149812549 - 149824399	319
		aggagcacggagggcaagc	67.35		
H3F3B-1	NM_005324.3	gagcgcggcggccttatctt	69.78	Chr17: 73775048 - 73775302	255
		cagcgagcaggggaggagtga	69.53		
H3F3B-2	NM_005324.3	gatttcaaaaccgacctgag	57.24	Chr17: 73775303-73774941	266
		ttaagctctctctccccgta	56.79		
H3F3B-3	NM_005324.3	aggaaaagcgctccctctac	59.98	Chr17: 73774874-73775177	304
		ccaggcctttgtcttacctg	59.73		
H3F3B-4	NM_005324.3	cgtacctggtgggtctgttc	60.42	Chr17: 73774474-73774791	318
		agcaacttgtcactcctgagc	59.66		

Supplementary Table S3.2. *Histone H3* primer details. All 17 *Histone H3* genes were amplified as 250-350bp PCR products. This required 2-7 primer sets to cover each *Histone H3* gene. The primers were designed to have similar annealing temperature and the optimized annealing temperature for PCRs was 57C°. Cycling conditions for all Histone H3 amplicons was 94C° for 2 minutes followed by 30 cycles of 94C° for 15 seconds; 57C° for 30 seconds (annealing); extension at 72C° for 30 seconds. The final extension was performed at 72C°for 7 minutes. All primers were designed using the Primer3 Output online software.

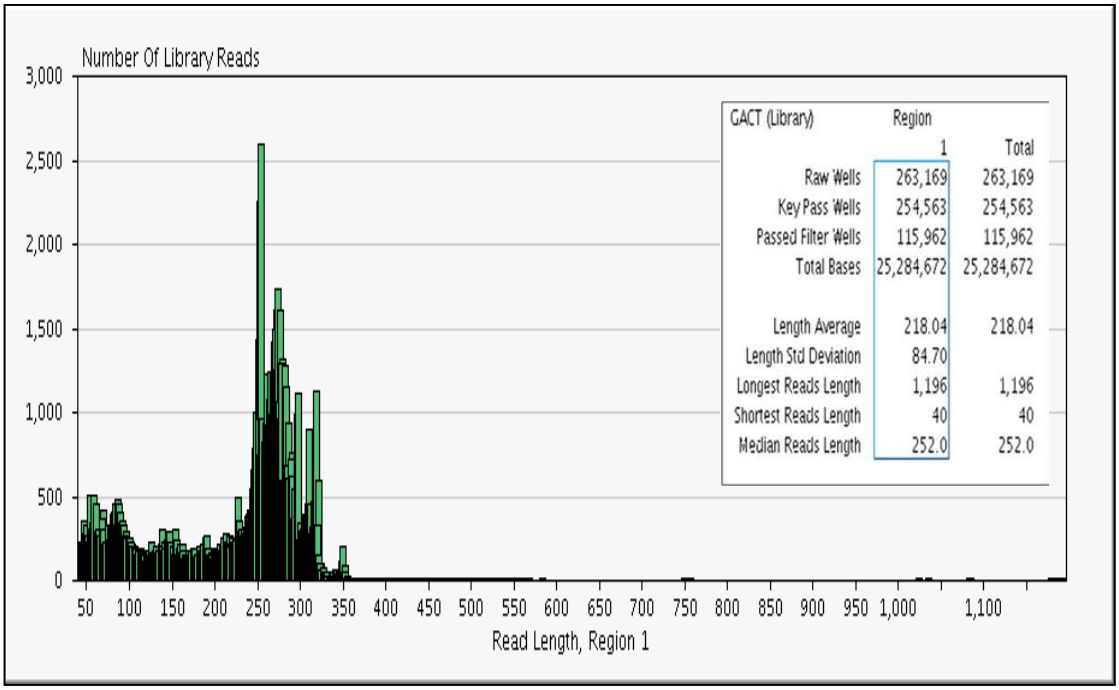
A.



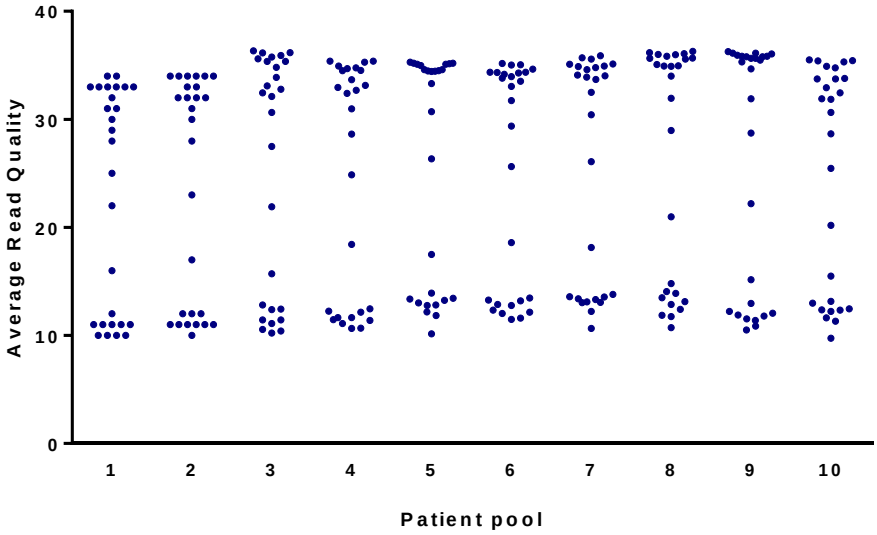
B.



Supplementary Figure S3.1. Agilent Bioanalyser analysis of A. pooled amplicons and B. MID ligated amplicon libraries for pool#1. Plots depict purity of library pools following four column purifications. All samples were washed three times with the ethanol based wash buffer from the QIAquick™ PCR Purification kit (QIAGEN™ Australia Pty Ltd, Australia) and double loaded on separate columns during elution.



Supplementary Figure S3.2. Data output from a sequencing run. The plot depicts number of reads (y-axis) versus read length (x-axis) following sequencing of sample pool#1. The maximum number of reads was observed for read lengths of 250-300 bases which were the predicted size for the amplicons.



Supplementary Figure S3.3. Read qualities (linear phred scale) of each pool output. The plot shows the average read or base phred quality for each sequencing run output from different pools. The x-axis depicts pools # 1-10 and the y-axis is the read output quality scale where a scale of 0-15 is poor, 15-20 are moderate quality reads and ≥ 25 are good quality reads.

Supplementary Table S3.3. *Histone H3* and *PHF6* gene changes identified in the Roche 454 sequencing chosen for validation with Sanger sequencing.

Chr	Position	Sample	Bases at this position	Number of Variant Allele Reads	Variant Allele Base Quality (ABQ) score	Number of Reference Allele Reads	Reference Allele Quality (RBQ) score	Reference Allele	Alternate Allele	*Gene ID*	Codon Variant	Amino Acid Variant
1	226252155	pool10	G/A	14	21	2041	22	G	[A]	*H3F3A*	Ggg/Agg	G35R
1	226252148	pool5	T/C	32	35	1346	34	T	[C]	*H3F3A*	tcT/tcC	S32S
1	226252085	pool7	G/C	12	17	1105	35	G	[C]	*H3F3A*	tcG/tcC	S11S
1	226252084	pool7	C/A	12	18	1107	34	C	[A]	*H3F3A*	tCg/tAg	S11*
17	73775174	pool1	T/C	13	20	734	22	T	[C]	*H3F3B*	Aaa/Gaa	K28E
17	73775221	pool1	G/T	25	18	889	25	G	[T]	*H3F3B*	aCc/aAc	T12N
17	73774689	pool3	C/T	20	18	1140	24	C	[T]	*H3F3B*	gGa/gAa	G133E
17	73774724	pool3	C/G	27	20	1514	34	C	[G]	*H3F3B*	atG/atC	M121I
17	73775141	pool7	G/C	25	18	2242	32	G	[C]	*H3F3B*	Cct/Gct	P39A
17	73774707	pool8	A/C	29	27	1061	35	A	[C]	*H3F3B*	tTg/tGg	L127W
17	73774928	pool9	T/C	53	18	1927	30	T	[C]	*H3F3B*	gAc/gGc	D82G
12	31945037	pool10	C/T	7	20	641	27	C	[T]	*H3F3C*	Gcc/Acc	A22T
12	31945066	pool3	G/T	22	16	1605	26	G	[T]	*H3F3C*	aCc/aAc	T12N
12	31944733	pool7	T/C	8	22	327	32	T	[C]	*H3F3C*	gAc/gGc	D123G
6	26020959	pool3	C/A	61	19	2669	29	C	[A]	*HIST1H3A*	aCa/aAa	T81K
6	26045750	pool2	A/G	77	22	5638	31	A	[G]	*HIST1H3C*	Aaa/Gaa	K38E

6	26045876	pool3	A/C	203	23	4287	28	A	[C]	*HIST1H3C*	Aaa/Caa	K80Q
6	26045879	pool3	A/C	46	22	4199	28	A	[C]	*HIST1H3C*	Acc/Ccc	T81P
6	26197367	pool10	T/C	10	18	885	30	T	[C]	*HIST1H3D*	Aag/Gag	K38E
6	26197364	pool10	G/C	23	18	771	24	G	[C]	*HIST1H3D*	Ccc/Gcc	P39A
6	26197444	pool10	G/T	7	17	513	30	G	[T]	*HIST1H3D*	aCg/aAg	T12K
6	26197136	pool3	C/G	53	18	2597	36	C	[G]	*HIST1H3D*	Gcc/Ccc	A115P
6	26225752	pool3	G/A	40	20	2053	36	G	[A]	*HIST1H3E*	Gac/Aac	D123N
6	26225497	pool3	C/G	13	20	841	27	C	[G]	*HIST1H3E*	Ccc/Gcc	P39A
6	26225669	pool7	C/T	117	35	2455	36	C	[T]	*HIST1H3E*	gCc/gTc	A96V
6	26271216	pool10	C/A	39	18	2009	22	C	[A]	*HIST1H3G*	Ggg/Tgg	G133W
6	27777919	pool3	C/A	10	18	521	35	C	[A]	*HIST1H3H*	aCc/aAc	T23N
6	27840030	pool10	C/T	8	20	528	31	C	[T]	*HIST1H3I*	Gcc/Acc	A22T
6	27840056	pool10	C/T	10	17	524	31	C	[T]	*HIST1H3I*	gGc/gAc	G13D
6	27839741	pool10	A/T	28	19	2196	34	A	[T]	*HIST1H3I*	gTc/gAc	V118D
6	27839742	pool10	C/A	29	19	2197	34	C	[A]	*HIST1H3I*	Gtc/Ttc	V118F
6	27839990	pool8	C/T	17	18	762	34	C	[T]	*HIST1H3I*	gGc/gAc	G35D
6	27858326	pool3	T/C	45	17	2096	31	T	[C]	*HIST1H3J*	gAc/gGc	D82G
6	27858327	pool3	C/G	45	16	2455	30	C	[G]	*HIST1H3J*	Gac/Cac	D82H
6	27858489	pool5	T/C	7	36	363	22	T	[C]	*HIST1H3J*	Aaa/Gaa	K28E
6	27858503	pool6	G/C	24	18	1104	33	G	[C]	*HIST1H3J*	aCc/aGc	T23S
1	149812693	pool10	C/G	11	17	802	33	C	[G]	*HIST2H3C*	Ggc/Cgc	G13R
1	149812689	pool10	C/G	11	19	830	30	C	[G]	*HIST2H3C*	gGc/gCc	G14A

Supplementary Table S3.3

1	149812695	pool10	G/T	18	19	810	32	G	[T]	*HIST2H3C*	aCc/aAc	T12N
1	149812620	pool2	T/C	13	18	1188	30	T	[C]	*HIST2H3C*	aAg/aGg	K37R
1	149812662	pool3	G/T	23	19	1053	34	G	[T]	*HIST2H3C*	aCc/aAc	T23N
1	149812627	pool6	C/G	15	20	1426	25	C	[G]	*HIST2H3C*	Ggg/Cgg	G35R
1	228612617	pool3	T/A	23	16	2155	31	T	[A]	*HIST3H3*	tAg/tTg	*137L
1	228612616	pool3	C/G	39	16	2086	32	C	[G]	*HIST3H3*	taG/taC	*137Y
1	228612684	pool3	C/G	29	18	2468	37	C	[G]	*HIST3H3*	Gcc/Ccc	A115P
X	133547613	pool1	A/T	8	15	481	16	A	[T]	*PHF6*	Aat/Tat	N137Y
X	133549119	pool1	T/G	12	20	1092	34	T	[G]	*PHF6*	gTa/gGa	V268G
X	133511751	pool10	A/G	9	26	744	26	A	[G]	*PHF6*	gAa/gGa	E35G
X	133511702	pool10	T/C	17	19	741	23	T	[C]	*PHF6*	Ttt/Ctt	F19L
X	133559249	pool10	T/A	20	19	1837	33	T	[A]	*PHF6*	caT/caA	H295Q
X	133549208	pool10	A/G	12	20	1199	29	A	[G]	*PHF6*	Aga/Gga	R299G
X	133559250	pool10	A/T	21	18	1819	33	A	[T]	*PHF6*	Agt/Tgt	S296C
X	133549202	pool2	A/T	7	22	567	26	A	[T]	*PHF6*	Aaa/Taa	K297*
X	133547658	pool3	G/T	12	20	763	35	G	[T]	*PHF6*	Gga/Tga	G152*
X	133547625	pool3	C/G	40	19	751	31	C	[G]	*PHF6*	Ctg/Gtg	L141V
X	133512072	pool3	A/T	13	24	665	36	A	[T]	*PHF6*	aAt/aTt	N59I
X	133511679	pool3	C/G	34	18	3338	30	C	[G]	*PHF6*	cCt/cGt	P11R
X	133512108	pool3	A/C	8	23	495	26	A	[C]	*PHF6*	cAa/cCa	Q71P
X	133547895	pool4	C/A	11	20	767	32	C	[A]	*PHF6*	Cct/Act	P176T
X	133549223	pool5	T/A	19	25	525	27	T	[A]	*PHF6*	Ttt/Att	F304I

Supplementary Table S3.3

X	133547927	pool6	A/G	16	19	863	21	A	[G]	*PHF6*	gaA/gaG	E186
X	133559309	pool7	A/G	24	24	1696	23	A	[G]	*PHF6*	aaA/aaG	K315
X	133512122	pool7	A/G	10	21	267	27	A	[G]	*PHF6*	Aga/Gga	R76G
X	133559310	pool7	G/T	23	21	2072	28	G	[T]	*PHF6*	Gta/Tta	V316L
X	133547937	pool8	G/A	10	16	662	33	G	[A]	*PHF6*	Gca/Aca	A190T
X	133547912	pool8	C/G	15	20	653	32	C	[G]	*PHF6*	tgC/tgG	C181W
X	133547936	pool8	A/G	10	16	624	33	A	[G]	*PHF6*	aA/gaG	E189
X	133512096	pool8	T/A	9	17	551	37	T	[A]	*PHF6*	aTt/aAt	I67N
X	133547898	pool8	A/T	7	17	592	29	A	[T]	*PHF6*	Aaa/Taa	K177*
X	133512114	pool9	A/G	16	20	681	20	A	[G]	*PHF6*	Aa/gGa	E73G
X	133549047	pool9	T/G	18	24	1194	31	T	[G]	*PHF6*	tTg/tGg	L244W

Supplementary Table S3.3. Changes identified in the Roche 454 sequencing chosen for validation with Sanger sequencing. 71 changes were sequenced in individual patient *g*DNA samples for validation. * indicates to Stop codon; Chr, chromosome. The *HIST1H3E* variant in Pool#7 has been highlighted in red.

Chapter 4: Clonal and lineage analysis of somatic *DNMT3A* and *JAK2* mutations in a chronic phase polycythemia vera patient

Statement of Authorship Title of Paper	Clonal and lineage analysis of somatic DNMT3A and JAK2 mutations in a chronic phase polycythemia vera patient
Publication Status	(•) Published ( ) Accepted for Publication ( ) Submitted for Publication ( ) Publication Style
Publication Details	Rao N, Butcher CM, Lewis ID, Ross DM, Melo JV, Scott HS, Bardy PG, D'Andrea RJ. *British Journal of Haematology. 2012;156(2):268-270.*

Author Contributions

By signing the Statement of Authorship, each author certifies that their stated contribution to the publication is accurate and that permission is granted for the publication to be included in the candidate's (Nisha Rao) thesis.

Name of Principal Author (Candidate)	Nisha Rao		
Contribution to the Paper	50%, performed the research, analysed and interpreted data and edited the manuscript		
Signature		Date	10/03/2013

Name of Co-Author	Carolyn M. Butcher		
Contribution to the Paper	15%, designed the research study, analysed and interpreted data and wrote the manuscript		
Signature		Date	07/03/2013

Name of Co-Author	Ian D. Lewis		
Contribution to the Paper	2%, contributed and interpreted clinical information and edited the manuscript		
Signature		Date	07/03/2013

Name of Co-Author	David M. Ross		
Contribution to the Paper	2%, contributed and interpreted clinical information and edited the manuscript		
Signature		Date	1/3/13

Name of Co-Author	Junia V. Melo		
Contribution to the Paper	2%, edited the manuscript		
Signature		Date	05/03/2013

Name of Co-Author	Hamish S. Scott		
Contribution to the Paper	2%, edited the manuscript		
Signature		Date	05/03/2013

Name of Co-Author	Peter G. Bardy		
Contribution to the Paper	7%, coordinated sample collection/contributed and interpreted clinical information		
Signature		Date	08/03/2013

Name of Co-Author	Richard J D'Andrea		
Contribution to the Paper	20%, contributed and interpreted clinical information and edited the manuscript		
Signature		Date	07/03/2013

4.1 Summary

Numerous somatic mutations have been identified in common with MPN and *de novo* AML, including somatic mutations of *TET2, ASXL1, CBL, IKZF1, EZH2, IDH1* and *IDH2* [reviewed in (Tefferi 2010)]. In 2011, a number of studies reported somatic mutations in the DNA methyltransferase, *DNMT3A*, in AML (Abdel-Wahab, *et al* 2011; Ley, *et al* 2010; Stegelmann, *et al* 2011; Walter, *et al* 2011; Yan, *et al* 2011). These mutations have subsequently been shown to alter the global methylation pattern of HSPC in AML patients (Holz-Schietinger, *et al* 2012).

DNMT3A is a DNA methyltransferase recruited to chromatin via interaction with histone H4 and histone H3 N-terminal tail residues, such as histone H4R3 (Zhao, *et al* 2009a) and histone H3K4 (Zhang, *et al* 2010). Thus, the histone H4R3 non-synonymous variant identified in a previous screen for histone gene mutations (Chapter 2) has clear potential to indirectly affect DNMT3A-induced methylation. As mutations in *DNMT3A* have been identified in AML, we wished to test whether similar mutations in *DNMT3A* may directly affect DNMT3A function in MPN. As the terminal exon of *DNMT3A* (amino acids 866–912) encodes part of the methyltransferase domain and was reported as the most commonly mutated region of *DNMT3A* in AML (Abdel-Wahab, *et al* 2011; Ley, *et al* 2010; Stegelmann, *et al* 2011; Yan, *et al* 2011), we used Sanger sequencing to screen this exon for mutations in *g*DNA samples isolated from 75 PV, 16 ET, 5 PMF and 2 U-MPN patients.

Heterozygous missense variants of *DNMT3A* affecting amino acid residues R882 and M880 were detected in PBMNC from two *JAK2V617F*$^{+}$ PV patients and these were confirmed in granulocyte samples. Sequencing of matched buccal DNA samples indicated that both variants were somatically acquired in these patients (see ***Fig 1***)

suggesting a low frequency (2.67%) of mutations for this *DNMT3A* exon in chronic-phase PV. Further, an investigation into the timing of acquisition of one of the somatic mutations of *DNMT3A* which affects R882 (R882C), suggested that the acquisition of the mutation preceded the acquisition of *JAK2V617F* (see ***Fig 2***).

These findings have been subsequently confirmed in another study of MPN patients (Brecqueville, *et al* 2011; Stegelmann, *et al* 2011) with an overall frequency of 10% for *DNMT3A* mutations in MPN which also included sAML. The frequency in chronic MPN was 7.5%. The R882 residue is the most commonly mutated residue in AML/MPN with substitution to histidine or cysteine the most frequent change. Of interest, is that the R882H mutation has been reported to reduce methyltransferase activity of DNMT3A (Yamashita, *et al* 2010) and has been associated with poor prognosis in AML (Ribeiro, *et al*).

Rao, N., Butcher, C.M., Lewis, I.D., Ross, D.M., Melo, J.V., Scott, H.S., Bardy, P.G. & D'Andrea, R.J. (2012) Clonal and lineage analysis of somatic DNMT3A and JAK2 mutations in a chronic phase polycythemia vera patient.
British Journal of Haematology, v. 156 (2), pp, 268-270

Chapter 5: Summary and significance

PV, ET and PMF are clonal disorders which are characterised by the presence of genetic lesions in *JAK2* (*JAK2V617F* and exon 12 mutations of *JAK2*), *MPL*, *CALR* and in epigenetic regulators including *DNMT3A, ASXL1*, and *EZH2*. A number of the mutated proteins in MPN are epigenetic modifiers and interact with histone PTMs, and given that histone gene mutations have been shown to be important in selected cancers, we focussed on the identification of coding region variants in MPN that could result pathogenesis via a similar mechanism (***Chapter 1, sections 1.7 and 1.9***). A combination of Sanger sequencing and next generation sequencing approaches was used to screen our MPN cohort for novel variants of histone H3, histone H4, and chromatin associated factors, DNMT3A and PHF6.

5.1 Role of histone H4 and DNMT3A variants in MPN

The first aim of the project was to identify coding variants in the *histone H4* genes (***Chapter 2***). *Histone H4* genes are highly conserved, and for histone H4 there are no known protein variants which play different functional roles such as the protein variants known for histone H3 (H3.1, H3.2 and H3.3). Given the extreme level of conservation, it was proposed that non-synonymous coding region variants in the *histone H4* genes may be of significance in MPN (***section 1.11***, ***Project Aim 1***).

Following screening of 128 MPN patients by Sanger sequencing two *JAK2V617F*$^{+}$ PV patients were identified with non-synonymous coding region changes in the *histone H4* gene *HIST1H4C; HIST1H4C:c.10C>T, p.R4C* (patient PV109) and *HIST1H4C:c.167G>A, p.R56Q* (patient PV115).

The *HIST1H4C:c.10C>T, p.R4C* affects the R3 residue on the mature histone H4 protein (as the methionine is removed after translation). The R3 residue is post translationally modified by a number of enzymes such as PRMT5 (Zhao, *et al* 2009a), PRMT1 (Wang, *et al* 2001), PRMT7 (Lee, *et al* 2005) and PAD4 (Wang, *et al* 2004). Of interest, histone H4R3 undergoes di-methylation of two types; symmetric di-methylation (H4R3me^{2}s) and asymmetric di-methylation (H4R3me^{2}a). These distinct PTMs on R3 play an important role in the recruitment of proteins. For example, P300/CBP-associated protein has been reported to interact with H4R3me^{2}a, facilitating the acetylation of lysine residues on histone H3 [K9 and K14, (Li, *et al* 2010)]. The resulting histone H3 acetylation has been reported to affect transcription of β-globin genes in erythroid cells and can be a mark of transcription activation [(Li, *et al* 2010) and reviewed in (Bannister, *et al* 2011)]. Furthermore, in a study of the developing murine cortex, the state of the histone H4R3 modification was observed to be a mark for specific differentiation states of neuronal precursors (Chittka 2010). It has also been reported that undifferentiated neuronal precursors displayed H4R3me^{2}s with the H4R3me^{2}a modification more predominant at the onset of neurogenesis (Chittka 2010). It is speculated that neuronal growth factors such as BTG2 and nerve growth factor (beta-polypeptide) may drive the specific PRMT(s) involved in directing methylation and thereby the differentiation state of the neuronal precursors (Chittka 2010). While the specific downstream effects and the proteins recruited due to H4R3me^{2}s and H4R3me^{2}a in neuronal precursors are yet to be determined in mammalian primary erythroid cells, DNMT3A has been shown to be a reader of H4R3me^{2}s. DNMT3A binds to the two methyl groups on R3 and methylates DNA leading to the silencing of the globin genes [for example γ-globin gene, (Zhao, *et al* 2009a)]. This association of DNMT3A with histone H4R3 was crucial in the reasoning for the study described in ***Chapter 4*** (section ***1.11***, ***Project***

Aim 2). The terminal exon of *DNMT3A* was screened for mutations in MPN. Two separate mutations were identified in two patients, *DNMT3A:c.2638A>G, p.M880V* (patient PV113) and *DNMT3A:c.2644C>T, p.R882C* (patient PV120). Subsequent to this study, mutations in *DNMT3A*, particularly those affecting the R882 residue, have been reported in many studies of myeloid malignancies including MPN (Abdel-Wahab, *et al* 2011; Rao, *et al* 2012; Stegelmann, *et al* 2011). In addition, given that JAK2V617F phosphorylates PRMT5, thereby reducing methyltransferase activity (Liu, *et al* 2011) and presumably histone H4R3 methylation, this may also affect recruitment of DNMT3A to chromatin. Thus several mechanisms including mutations affecting histone H4R3 have potential to impact DNMT3A recruitment; activity and gene expression (see ***Figure 5.1***).

In the human genome, there are 15 *histone H4* genes which are highly conserved and tightly regulated. Given that 14 of the 15 genes encode for the identical protein, a single missense mutation in one of the genes may not be functional in all cell types. However, two lines of supporting evidence from the studies performed in ***Chapter 2*** suggested that the p.R4C variant of *HIST1H4C* may play a functional role in haematopoietic cells (***Project Aim 1***). Firstly, the expression of *HIST1H4C* in primary and haematopoietic cell lines made a significant contribution to the total *Histone H4* mRNA, and secondly, over-expression of the p.R4C variant in FDM cells showed a significant differential expression of two genes (*Il4rα and Dock10*) relative to cells expressing the wildtype histone H4 (***Chapter 2, sections 2.3.8*** and ***Figure 2.14***). However, further studies are needed to confirm these changes in expression in independent FDM cell populations with equivalent levels of histone H4 expression. It would also be of interest to perform global DNA

Figure 5.1. **Model of cooperation between DNMT3A mutation and JAK2V617F**. *DNMT3A* mutations (red triangle) are associated with reduced methyl transferase activity and may lead to partial activation of oncogenes. The *JAK2*V617F mutation results in phosphorylation of PMRT5 (green circles) which reduces arginine methyltransferase activity, recruitment of DNMT3A and adjacent DNA methylation. The combination of reduced DNMT3A activity and reduced recruitment via PRMT5-mediated H4R3me2s is predicted to result in further oncogene activation. *Figure adapted from (Zhao, et al 2009a).*

methylation and histone modification profiles in cells expressing the p.R4C (and p.R56Q) variants to detect alterations associated with gene expression changes.

5.2 Identification of *histone H3* and *PHF6* variants using Sanger and next-generation sequencing approach

5.2.1 Histone H3

In ***Chapter 3***, 17 *histone H3* genes were also investigated for novel coding region variants. Histone H3 is the binding partner of histone H4 and forms the histone H3-H4 heterodimer (***Project Aim 2***). Also, further reasoning for such a screen was provided when mutations in *histone H3* genes, *H3F3A* and *HIST1H3B*, were identified in patients with gliomas (Schwartzentruber, *et al* 2012; Wu, *et al* 2012). A next-generation sequencing screen (Roche 454 sequencing) and Sanger sequencing based validation of all *histone H3* genes identified a coding variant in *HIST1H3E* (*HIST1H3E:c.287C>T;* p.A96V) in patient PV44 (***Chapter 3, sections 3.3.4, Figure 3.7 and 3.8***).

Using this next-generation sequencing approach, we also identified changes in histone H3 that have previously been reported in glioma. These included a G35R variant (G34 on the mature protein) in *H3F3A* in pool#10 (1% of the total reads in that pool). In addition, G35R and G35D variants were also observed in two separate sample pools [G35R (*HIST2H3A/C* in pool#6) and G35D (*HIST1H3I* in pool#8)]. The K28E variant (K27 on the mature protein) was observed at a low frequency of reads in two *histone H3* genes [*H3F3B* in pool#1 (approximately 1.7% of total reads) and *HIST1H3J* in pool#5 (approximately 1.9% of total reads)]. While these were not confirmed in the individual disease *g*DNA samples by Sanger sequencing, we cannot exclude that the mutations may be present in an individual patient sample at a level

below the detection limit of Sanger sequencing (see ***Chapter 3, Supplementary Table S3.3*** for further details and ***section 3.4*** for discussion). Thus, sensitive techniques such as allele-specific PCR will be necessary for validation of these variants of particular interest.

Interestingly, somatic coding region variants of histone H3 and H4 have been described in the COSMIC database in cancer samples such as squamous-cell carcinoma, glioma, and adenocarcinoma. A list of *histone H3* and *histone H4* genes with coding region variants that could potentially affect protein function have been listed in ***Table 5.1*** and ***Table 5.2*** respectively. Apart from the variants observed in glioma, variants affecting R2, K36 and Y41 residues were observed for a number of *histone H3* genes. For *histone H4* genes, variants affecting initiator methionine (M), S1 and termination codon were frequently observed. In addition, histone H4R3 variants were identified for *HIST1H4B* (*HIST1H4B*:*c.10C>T, p.R4C,* confirmed somatic) and *HIST1H4F* (*HIST1H4F*:*c.11G>A, p.R4K*) in cancer. Thus, recurrent somatic variants affecting the same residues of Histone H3 and H4 have been reported at a low frequency in cancer, consistent with a role of these genes in disease pathogenesis. However, given that these variants have been detected in advanced malignancies, it is not clear whether these mutations are causal or occur late in malignancy. The altered gene expression associated with forced expression of the histone H4R3 variant in myeloid cells (***section 5.1***) is consistent with a potential to contribute to changes in progenitor behaviour, although further functional studies are necessary as outlined above.

Table 5.1

Gene	Somatic non-synonymous coding region variants of histone H3	Disease type	Mutation Count	Confirmed Somatic	Altered PTM/ Effect on protein	Relevant References
HIST1H3A	c.125A>G p.Y42C	Squamous cell carcinoma	1	No	Phosphorylated (Y41)	*(Stransky, et al 2011)*
HIST1H3B	***c.83A>T p.K28M***	***Glioma***	10	Yes	Methylated and Acetylated(K27)	*(Wu, et al 2012)*
HIST1H3C	c.7C>G p.R3G c.34A>G p.T12A	Serous carcinoma non small cell carcinoma	1 1	Yes, No	Methylated (R2) Phosphorylated (T11)	*(Network 2011)* *(Liu, et al 2012)*
HIST1H3D	c.110A>T p.K37M, c.178G>T p.E60*, c.240G>C p.K80N	Adenocarcinoma, Breast cancer, Endometrioid carcinoma	1 1 1	Yes Yes Yes	Methylated and Acetylated (K36) Truncated protein (E59) Acetylated (K79)	*(Network 2012)* *(Stephens, et al 2012)* *(COSMIC study ID COSU419)*
HIST1H3F	c.195G>C p.K65N	Small cell carcinoma	1	Yes	Methylated (K64)	*(Network 2012; Rudin, et al 2012)*
HIST1H3G	c.166C>T p.Q56*, c.302_303insACTT p.V102fs*6	Squamous cell carcinoma, Plasma cell myeloma	1 1	Yes No	Truncated protein Altered a.a. sequence	*(COSMIC study ID COSU418)* *(Walker, et al 2012)*
HIST1H3H	c.10A>G p.T4A, c.25C>G p.R9G, c.109A>G p.K37E, c.110A>T p.K37M	Small cell carcinoma, Serous carcinoma, Adenocarcinoma, Malignant melanoma	1 1 1 1	No Yes No No	Phosphorylated (T3) Methylated (R8) Methylation and Acetylated (K36) Methylation and Acetylated (K36)	*(Peifer, et al 2012)* *(Network 2011)* *(Imielinski, et al 2012)* *(Krauthammer, et al 2012)*
HIST1H3I	c.79C>T p.R27C , c.240G>C p.K80N , c.411A>C p.*137Y	Squamous cell carcinoma, Endometrioid carcinoma, Adenocarcinoma	1 1 1	Yes Yes No	Methylation (R26) Methylated and Acetylated (K79) Altered termination of translation	*(COSMIC study ID COSU418)* *(COSMIC study ID COSU419)* *(Imielinski, et al 2012; Seligson, et al 2009)*

Continued on the next page

Gene	Somatic non-synonymous coding region variants of histone H3	Disease type	Mutation Count	Confirmed Somatic	Altered PTM/ Effect on protein	Relevant References
HIST2H3D	c.82A>T p.K28*, c.166C>T p.Q56*	Hepatocellular carcinoma Endometrioid carcinoma	1 1	Yes Yes	Truncated protein Truncated protein	*(Li, et al 2011b)* *(COSMIC study ID COSU419)*
H3F3A	c.7C>T p.R3C, ***c.83A>T p.K28M,*** ***c.103G>A p.G35R,*** ***c.104G>T p.G35V*** c.103G>T p.G35W c.?, p.G35L c.?, p.G35R c.110A>G p.K37R c.111G>T p.K37N c.?,p.K37M	Squamous cell carcinoma, **Glioma,** **Glioma** **Glioma** Sarcomas,Giant cell tumour of bone Giant cell tumour of bone Osteosarcoma Carcinoma (urinary tract) Endometroid Carcinoma Chondroblastoma,Chondrosarcoma	1 166 21 1 1,48 1 1 1 1 5,1	No Yes Yes Yes Yes, Yes Yes Yes Yes No Yes,Yes	Methylated (R2) Methylated and Acetylated (K27) Altered DNA Methylation Altered DNA Methylation Altered DNA Methylation Methylated and Acetylated (K36) Methylated and Acetylated (K36) Methylated and Acetylated (K36)	*(Stransky, et al 2011)* *(Schwartzentruber, et al 2012; Wu, et al 2012)* *(Behjati, et al 2013; Joseph, et al 2014)* *(Behjati, et al 2013)* *(Behjati, et al 2013)* *(COSMIC study ID COSU413)* *(COSMIC study ID COSU419)* *(Behjati, et al 2013)*
H3F3B	c.52_53insC p.R18fs*80 c.?, p.G35R c.?,p.K37M	Serous carcinoma Adenocarcinoma Osteosarcoma Chondroblastoma, Chondrosarcoma	1 1 68,1	Yes Yes Yes,Yes	Altered a.a sequence Decreased methylation of K36 Methylated and Acetylated (K36)	*(Network 2012; Network 2011) (Behjati, et al 2013)* *(Behjati, et al 2013)*
H3F3C	c.145C>T p.R49*, c.169T>A p.S57T, c.239C>T p.T80I	Prostate cancer, Adenocarcinoma, NS	1 1 1	Yes No Yes	Truncated protein (at R48) Phosphorylated (S56) Phosphorylated (T79)	*(Lindberg, et al 2013)* *(Imielinski, et al 2012)* *(COSMIC study ID COSU376)*

Table 5.1. Somatic non-synonymous coding region variants of histone H3. The COSMIC database was used to identify variants in the 17 *histone H3* genes which could potentially alter PTM or the protein. NS, Not Specified. Diseases identified with common mutations have been highlighted in bold and italized. Details on potential effect on protein were taken from the UniProt database and any cited studies. ? indicates that the nucleic acid change was not reported in the study. a.a, amino acid.

Gene	Somatic non-synonymous coding region variants of histone H4	Disease type	Mutation Count	Confirmed Somatic	Altered PTM/ Effect on protein	Relevant References
HIST1H4B	***c.10C>T, p.R4C***	Clear cell renal carcinoma	1	Yes	Methylated (R3)	*(Sato, et al 2013)*
HIST1H4C	c.4T>G, p.S2A	Plasma cell myeloma	1	No	Phosphorylated (S1)	*(Walker, et al 2012)*
HIST1H4D	c.26A>G, p.K9R, c.310T>C, p.*104R	Breast carcinoma, Endometroid carcinoma	1 1	Yes Yes	Acetylated (K8) Altered termination of translation	*(Stephens, et al 2012)* *(COSMIC study ID COSU419)*
HIST1H4F	***c.11G>A, p.R4K***	Squamous cell carcinoma	1	No	Methylated (R3)	*(Stransky, et al 2011)*
HIST1H4H	c.5C>G, p.S2C, c.16A>G, p.K6E	Serous carcinoma, Adenocarcinoma	1 1	Yes Yes	Phosphorylated (S1) Acetylated (K5)	*(Network 2011)* *(COSMIC study ID COSU417)*
HIST1H4J	c.5C>T, p.S2F	Squamous cell carcinoma	1	Yes	Phosphorylated (S1)	*(Agrawal, et al 2012)*
HIST1H4K	c.3G>C p.M1I, c.5C>G, p.S2C c.311G>C, p.*104S	Adenocarcinoma	1 1 1	No Yes No	Altered Initiation of Translation (M) Phosphorylated (S1) Altered termination of translation	*(COSMIC study ID COSU435)* *(Imielinski, et al 2012)* *(Imielinski, et al 2012)*
HIST1H4L	c.1A>G p.M1V, c.3G>T p.M1I, c.38A>T p.K13I, c.311G>T p.*104L	Adenocarcinoma, Small cell carcinoma, Squamous cell carcinoma Adenocarcinoma	1 1 1 1	No No No No	Altered Initiation of Translation (M) Altered Initiation of Translation (M) Acetylated (K12) Altered termination of translation	*(Imielinski, et al 2012)* *(Rudin, et al 2012)* *(Stransky, et al 2011)* *(Imielinski, et al 2012)*

Table 5.2. Somatic non-synonymous coding region variants of histone H4. The COSMIC database was used to identify variants in the *Histone H4* which could potentially alter PTM or the protein. Bold and italized variant affects residue R3 on the mature Histone H4 protein. Details on potential effect on protein were taken from the UniProt database.

5.2.2 PHF6

In addition to the *histone H3* screen, the coding exons of *PHF6* were also screened for mutations. *PHF6* (located on the X chromosome) encodes for a PHD domain protein with possible roles in transcription (Lower, *et al* 2002; Wang, *et al* 2013). In addition, reports of PHF6 as a possible component of the NURD complex, with an association with histone H3 (Todd, *et al* 2012) was further reasoning to perform mutational screening of *PHF6* in MPN patients.

An initial targeted sequencing screen of *PHF6* exons 9 and 10 identified a frameshift mutation, *PHF6R335fs*, in one chronic phase PV patient (PV108) (***Chapter 3, section 3.3.1***, ***Figure 3.4*** and ***Figure 3.5***). The *PHF6R335fs* mutation was identified in a chronic phase *JAK2V617F*$^{+}$ PV patient and was heterozygous. This mutation has previously been reported as a homozygous change in one adult male patient with trisomy 8 AML (Van Vlierberghe, *et al* 2011) suggesting a role for this *PHF6* mutation in the development or progression to AML.

Subsequently, a screen of all coding exons of *PHF6* using next-generation sequencing identified 26 coding region variants however, these did not validate with Sanger sequencing. Inability to validate these variants can be attributed either to technical difficulties associated with the sequencing platform or lack of sensitivity of Sanger sequencing to detect variants present at a threshold below the detection limits of the approach. However, given that mutations in *PHF6* are recurrent in AML and have been associated with poor prognosis and survival (Patel, *et al* 2012), the identification of a mutation in *PHF6* in MPN suggested that genetic alterations to *PHF6*, although rare, may have functional importance in MPN. In a recent conference report, mutations in *PHF6* have been identified by targeted next-

generation deep sequencing and Sanger sequencing in 19 of 809 cases (2.5%) (Shen, *et al* 2013). These cases included MPN/MDS, AML, MDS, and MPN. It was indicated that mutations in *PHF6* occurred at a low frequency in aggressive myeloid malignancies. Given that Shen and colleagues have speculated that mutations affecting PHF6 may act as *drivers* in aggressive myeloid malignancies (Shen, *et al* 2013), in future it will be crucial to relate *PHF6*R335fs mutation to molecular or diagnostic alterations and progression occurring (if any) in patient PV108. This would provide prognostic relevance of mutations in *PHF6* in chronic MPN and in progression to AML.

In conclusion, this project identified rare and significant coding region variants in genes encoding for histone H3, histone H4, DNMT3A and PHF6. The central role of the histone H4R3 residue in PTM and for DNMT3A recruitment highlight the potential of this novel histone variant to contribute to pathogenesis in this group of diseases, and the preliminary functional studies indicate a potential functional role. The presence of low frequency mutations affecting histone H3K27 and H3G34 may also be of significance given the highlighted role of altered global DNA methylation in gliomas (Bender, *et al* 2013; Lewis, *et al* 2013; Schwartzentruber, *et al* 2012; Sturm, *et al* 2012). Future projects will focus on more detailed functional analysis of epigenetic and gene expression changes following introduction of the novel disease-associated histone variants in *in-vitro* models of normal and malignant myeloid development.

Response to Examiner's Comments:

A. REFERENCES AND REFERENCING

(1) As identified by the examiner, all references with incorrect journal names, cases of journals and incomplete references have been modified.

(2) Care has been taken to ensure the first author's name is written correctly. Please see ***page 25*** for the specific example which was raised by the examiner.

(3) As highlighted by the examiner, the reference: Olcaydu *et al*, 2009, has now been cited on ***page 44*** of the thesis.

B. **Introduction**

(1) Factual error: As highlighted by the examiner, the factual error on ***page 26*** has been rectified. The sentence now reads as follows: "*These mutations occur in virtually all of JAK2V617F-negative PV cases (which constitute 3-4% of all PV cases)*".

Additional references have been included on ***page 40*** to denote other studies which highlight the role of *JAK2V617F* in contributing to MPN phenotypes.

(2) ***Figure 1.4A*** on ***page 12*** has been altered to depict the percentage contribution of each genotype in relation to the entire MPN population.

References:

Aaronson, D. S. & Horvath, C. M. (2002) A road map for those who don't know JAK-STAT. *Science,* **296,** 1653-1655.

Abdel-Wahab, O. (2011) Genetics of the myeloproliferative neoplasms. *Current Opinion in Hematology,* **18,** 117-123.

Abdel-Wahab, O., Adli, M., Lafave, L. M., Gao, J., Hricik, T., Shih, A. H., Pandey, S., Patel, J. P., Chung, Y. R., Koche, R., Perna, F., Zhao, X., Taylor, J. E., Park, C. Y., Carroll, M., Melnick, A., Nimer, S. D., Jaffe, J. D., Aifantis, I., Bernstein, B. E. & Levine, R. L. (2012) ASXL1 mutations promote myeloid transformation through loss of PRC2-mediated gene repression. *Cancer Cell,* **22,** 180-193.

Abdel-Wahab, O., Gao, J., Adli, M., Dey, A., Trimarchi, T., Chung, Y. R., Kuscu, C., Hricik, T., Ndiaye-Lobry, D., Lafave, L. M., Koche, R., Shih, A. H., Guryanova, O. A., Kim, E., Li, S., Pandey, S., Shin, J. Y., Telis, L., Liu, J., Bhatt, P. K., Monette, S., Zhao, X., Mason, C. E., Park, C. Y., Bernstein, B. E., Aifantis, I. & Levine, R. L. (2013) Deletion of Asxl1 results in myelodysplasia and severe developmental defects in vivo. *The Journal of Experimental of Medicine,* **210,** 2641-2659.

Abdel-Wahab, O., Manshouri, T., Patel, J., Harris, K., Yao, J., Hedvat, C., Heguy, A., Bueso-Ramos, C., Kantarjian, H., Levine, R. L. & Verstovsek, S. (2010) Genetic analysis of transforming events that convert chronic myeloproliferative neoplasms to leukemias. *Cancer Research,* **70,** 447-452.

Abdel-Wahab, O., Mullally, A., Hedvat, C., Garcia-Manero, G., Patel, J., Wadleigh, M., Malinge, S., Yao, J., Kilpivaara, O., Bhat, R., Huberman, K., Thomas, S., Dolgalev, I., Heguy, A., Paietta, E., Le Beau, M. M., Beran, M., Tallman, M. S., Ebert, B. L., Kantarjian, H. M., Stone, R. M., Gilliland, D. G., Crispino, J. D. & Levine, R. L. (2009) Genetic characterization of TET1, TET2, and TET3 alterations in myeloid malignancies. *Blood,* **114,** 144-147.

Abdel-Wahab, O., Pardanani, A., Rampal, R., Lasho, T. L., Levine, R. L. & Tefferi, A. (2011) DNMT3A mutational analysis in primary myelofibrosis, chronic myelomonocytic leukemia and advanced phases of myeloproliferative neoplasms. *Leukemia* **25,** 1219-1220.

Adamson, J. W., Fialkow, P. J., Murphy, S., Prchal, J. F. & Steinmann, L. (1976) Polycythemia vera: stem-cell and probable clonal origin of the disease. *The New England Journal of Medicine,* **295,** 913-916.

Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., Kondrashov, A. S. & Sunyaev, S. R. (2010) A method and server for predicting damaging missense mutations. *Nature Methods,* **7,** 248-249.

Agrawal, N., Jiao, Y., Bettegowda, C., Hutfless, S. M., Wang, Y., David, S., Cheng, Y., Twaddell, W. S., Latt, N. L., Shin, E. J., Wang, L. D., Wang, L., Yang, W., Velculescu, V. E., Vogelstein, B., Papadopoulos, N., Kinzler, K. W. & Meltzer, S. J. (2012) Comparative genomic analysis of esophageal adenocarcinoma and squamous cell carcinoma. *Cancer Discovery,* **2,** 899-905.

Ahmad, K. & Henikoff, S. (2002) The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. *Molecular Cell,* **9,** 1191-1200.

Akada, H., Yan, D., Zou, H., Fiering, S., Hutchison, R. E. & Mohi, M. G. (2010) Conditional expression of heterozygous or homozygous Jak2V617F from its

endogenous promoter induces a polycythemia vera-like disease. *Blood,* **115,** 3589-3597.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K. & Walter, P. (2002) *Molecular Biology of the Cell, 4th Edition.* Garland Science, New York.

Alcaraz-Garcia, M. J., Ruiz-Lafuente, N., Sebastian-Ruiz, S., Majado, M. J., Gonzalez-Garcia, C., Bernardo, M. V., Alvarez-Lopez, M. R. & Parrado, A. (2011) Human and mouse DOCK10 splicing isoforms with alternative first coding exon usage are differentially expressed in T and B lymphocytes. *Human Immunology,* **72,** 531-537.

Arber, D. A., Chang, K. L., Lyda, M. H., Bedell, V., Spielberger, R. & Slovak, M. L. (2003) Detection of NPM/MLF1 fusion in t(3;5)-positive acute myeloid leukemia and myelodysplasia. *Human Pathology,* **34,** 809-813.

Asimakopoulos, F. A., Holloway, T. L., Nacheva, E. P., Scott, M. A., Fenaux, P. & Green, A. R. (1996) Detection of chromosome 20q deletions in bone marrow metaphases but not peripheral blood granulocytes in patients with myeloproliferative disorders or myelodysplastic syndromes. *Blood,* **87,** 1561-1570.

Azzouz, T. N., Gruber, A. & Schumperli, D. (2005) U7 snRNP-specific Lsm11 protein: dual binding contacts with the 100 kDa zinc finger processing factor (ZFP100) and a ZFP100-independent function in histone RNA 3' end processing. *Nucleic Acids Research,* **33,** 2106-2117.

Back, J., Dierich, A., Bronn, C., Kastner, P. & Chan, S. (2004) PU.1 determines the self-renewal capacity of erythroid progenitor cells. *Blood,* **103,** 3615-3623.

Bamford, S., Dawson, E., Forbes, S., Clements, J., Pettett, R., Dogan, A., Flanagan, A., Teague, J., Futreal, P. A., Stratton, M. R. & Wooster, R. (2004) The COSMIC (Catalogue of Somatic Mutations in Cancer) database and website. *British Journal of Cancer,* **91,** 355-358.

Bannister, A. J. & Kouzarides, T. (2011) Regulation of chromatin by histone modifications. *Cell Research,* **21,** 381-395.

Barbui, T., Barosi, G., Birgegard, G., Cervantes, F., Finazzi, G., Griesshammer, M., Harrison, C., Hasselbalch, H. C., Hehlmann, R., Hoffman, R., Kiladjian, J. J., Kroger, N., Mesa, R., Mcmullin, M. F., Pardanani, A., Passamonti, F., Vannucchi, A. M., Reiter, A., Silver, R. T., Verstovsek, S. & Tefferi, A. (2011a) Philadelphia-negative classical myeloproliferative neoplasms: critical concepts and management recommendations from European LeukemiaNet. *Journal of Clinical Oncology,* **29,** 761-770.

Barbui, T., Barosi, G., Grossi, A., Gugliotta, L., Liberato, L. N., Marchetti, M., Mazzucconi, M. G., Rodeghiero, F. & Tura, S. (2004) Practice guidelines for the therapy of essential thrombocythemia. A statement from the Italian Society of Hematology, the Italian Society of Experimental Hematology and the Italian Group for Bone Marrow Transplantation. *Haematologica,* **89,** 215-232.

Barbui, T., Thiele, J., Passamonti, F., Rumi, E., Boveri, E., Ruggeri, M., Rodeghiero, F., D'amore, E. S., Randi, M. L., Bertozzi, I., Marino, F., Vannucchi, A. M., Antonioli, E., Carrai, V., Gisslinger, H., Buxhofer-Ausch, V., Mullauer, L., Carobbio, A., Gianatti, A., Gangat, N., Hanson, C. A. & Tefferi, A. (2011b) Survival and disease progression in essential thrombocythemia are significantly influenced by accurate morphologic diagnosis: an international study. *Journal of Clinical Oncology,* **29,** 3179-3184.

Barlesi, F., Giaccone, G., Gallegos-Ruiz, M. I., Loundou, A., Span, S. W., Lefesvre, P., Kruyt, F. A. & Rodriguez, J. A. (2007) Global histone modifications

predict prognosis of resected non small-cell lung cancer. *Journal of Clinical Oncology* **25,** 4358-4364.

Barosi, G., Bergamaschi, G., Marchetti, M., Vannucchi, A. M., Guglielmelli, P., Antonioli, E., Massa, M., Rosti, V., Campanelli, R., Villani, L., Viarengo, G., Gattoni, E., Gerli, G., Specchia, G., Tinelli, C., Rambaldi, A. & Barbui, T. (2007) JAK2 V617F mutational status predicts progression to large splenomegaly and leukemic transformation in primary myelofibrosis. *Blood,* **110,** 4030-4036.

Barosi, G., Mesa, R. A., Thiele, J., Cervantes, F., Campbell, P. J., Verstovsek, S., Dupriez, B., Levine, R. L., Passamonti, F., Gotlib, J., Reilly, J. T., Vannucchi, A. M., Hanson, C. A., Solberg, L. A., Orazi, A. & Tefferi, A. (2008) Proposed criteria for the diagnosis of post-polycythemia vera and post-essential thrombocythemia myelofibrosis: a consensus statement from the International Working Group for Myelofibrosis Research and Treatment. *Leukemia,* **22,** 437-438.

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. & Green, A. R. (2005) Acquired mutation of the tyrosine kinase JAK2 in human myeloproliferative disorders. *The Lancet,* **365,** 1054-1061.

Bazan, J. F. (1990) Structural design and molecular evolution of a cytokine receptor superfamily. *Proceedings of the National Academy of Sciences USA,* **87,** 6934-6938.

Beck, H. C., Nielsen, E. C., Matthiesen, R., Jensen, L. H., Sehested, M., Finn, P., Grauslund, M., Hansen, A. M. & Jensen, O. N. (2006) Quantitative proteomic analysis of post-translational modifications of human histones. *Molecular and Cellular Proteomics,* **5,** 1314-1325.

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. & Green, A. R. (2008) MPL mutations in myeloproliferative disorders: analysis of the PT-1 cohort. *Blood,* **112,** 141-149.

Beer, P. A., Delhommeau, F., Lecouedic, J. P., Dawson, M. A., Chen, E., Bareford, D., Kusec, R., Mcmullin, M. F., Harrison, C. N., Vannucchi, A. M., Vainchenker, W. & Green, A. R. (2009) Two routes to leukemic transformation after a JAK2 mutation-positive myeloproliferative neoplasm. *Blood,* **115,** 2891-2900.

Beer, P. A., Ortmann, C. A., Stegelmann, F., Guglielmelli, P., Reilly, J. T., Larsen, T. S., Hasselbalch, H. C., Vannucchi, A. M., Moller, P., Dohner, K. & Green, A. R. (2010) Molecular mechanisms associated with leukemic transformation of MPL-mutant myeloproliferative neoplasms. *Haematologica,* **95,** 2153-2156.

Beerman, I., Bhattacharya, D., Zandi, S., Sigvardsson, M., Weissman, I. L., Bryder, D. & Rossi, D. J. (2010) Functionally distinct hematopoietic stem cells modulate hematopoietic lineage potential during aging by a mechanism of clonal expansion. *Proceedings of the National Academy of Sciences USA,* **107,** 5465-5470.

Behjati, S., Tarpey, P. S., Presneau, N., Scheipl, S., Pillay, N., Van Loo, P., Wedge, D. C., Cooke, S. L., Gundem, G., Davies, H., Nik-Zainal, S., Martin, S., Mclaren, S., Goodie, V., Robinson, B., Butler, A., Teague, J. W., Halai, D., Khatri, B., Myklebost, O., Baumhoer, D., Jundt, G., Hamoudi, R., Tirabosco, R., Amary, M. F., Futreal, P. A., Stratton, M. R., Campbell, P. J. & Flanagan, A. M. (2013) Distinct H3F3A and H3F3B driver mutations define

chondroblastoma and giant cell tumor of bone. *Nature Genetics,* **45,** 1479-1482.

Behjati, S., Tarpey, P. S., Presneau, N., Scheipl, S., Pillay, N., Van Loo, P., Wedge, D. C., Cooke, S. L., Gundem, G., Davies, H., Nik-Zainal, S., Martin, S., Mclaren, S., Goodie, V., Robinson, B., Butler, A., Teague, J. W., Halai, D., Khatri, B., Myklebost, O., Baumhoer, D., Jundt, G., Hamoudi, R., Tirabosco, R., Amary, M. F., Futreal, P. A., Stratton, M. R., Campbell, P. J. & Flanagan, A. M. (2014) Corrigendum: Distinct H3F3A and H3F3B driver mutations define chondroblastoma and giant cell tumor of bone. *Nature Genetics,* **46,** 316.

Behre, G., Singh, S. M., Liu, H., Bortolin, L. T., Christopeit, M., Radomska, H. S., Rangatia, J., Hiddemann, W., Friedman, A. D. & Tenen, D. G. (2002) Ras signaling enhances the activity of C/EBP alpha to induce granulocytic differentiation by phosphorylation of serine 248. *The Journal of Biological Chemistry,* **277,** 26293-26299.

Bellanne-Chantelot, C., Chaumarel, I., Labopin, M., Bellanger, F., Barbu, V., De Toma, C., Delhommeau, F., Casadevall, N., Vainchenker, W., Thomas, G. & Najman, A. (2006) Genetic and clinical implications of the Val617Phe JAK2 mutation in 72 families with myeloproliferative disorders. *Blood,* **108,** 346-352.

Bender, S., Tang, Y., Lindroth, A. M., Hovestadt, V., Jones, D. T., Kool, M., Zapatka, M., Northcott, P. A., Sturm, D., Wang, W., Radlwimmer, B., Hojfeldt, J. W., Truffaux, N., Castel, D., Schubert, S., Ryzhova, M., Seker-Cin, H., Gronych, J., Johann, P. D., Stark, S., Meyer, J., Milde, T., Schuhmann, M., Ebinger, M., Monoranu, C. M., Ponnuswami, A., Chen, S., Jones, C., Witt, O., Collins, V. P., Von Deimling, A., Jabado, N., Puget, S., Grill, J., Helin, K., Korshunov, A., Lichter, P., Monje, M., Plass, C., Cho, Y. J. & Pfister, S. M. (2013) Reduced H3K27me3 and DNA hypomethylation are major drivers of gene expression in K27M mutant pediatric high-grade gliomas. *Cancer Cell,* **24,** 660-672.

Berk, P. D., Goldberg, J. D., Donovan, P. B., Fruchtman, S. M., Berlin, N. I. & Wasserman, L. R. (1986) Therapeutic recommendations in polycythemia vera based on Polycythemia Vera Study Group protocols. *Seminars in Hematology,* **23,** 132-143.

Bernard, L., Belisle, C., Mollica, L., Provost, S., Roy, D. C., Gilliland, D. G., Levine, R. L. & Busque, L. (2009) Telomere length is severely and similarly reduced in JAK2V617F-positive and -negative myeloproliferative neoplasms. *Leukemia,* **23,** 287-291.

Bjorkholm, M., Derolf, A. R., Hultcrantz, M., Kristinsson, S. Y., Ekstrand, C., Goldin, L. R., Andreasson, B., Birgegard, G., Linder, O., Malm, C., Markevarn, B., Nilsson, L., Samuelsson, J., Granath, F. & Landgren, O. (2011) Treatment-related risk factors for transformation to acute myeloid leukemia and myelodysplastic syndromes in myeloproliferative neoplasms. *Journal of Clinical Oncology,* **29,** 2410-2415.

Blume-Jensen, P., Claesson-Welsh, L., Siegbahn, A., Zsebo, K. M., Westermark, B. & Heldin, C. H. (1991) Activation of the human c-kit product by ligand-induced dimerization mediates circular actin reorganization and chemotaxis. *The EMBO Journal,* **10,** 4121-4128.

Bonicelli, G., Abdulkarim, K., Mounier, M., Johansson, P., Rossi, C., Jooste, V., Andreasson, B., Maynadie, M. & Girodon, F. (2013) Leucocytosis and thrombosis at diagnosis are associated with poor survival in polycythaemia

vera: a population-based study of 327 patients. *British Journal of Haematology,* **160,** 251-254.

Botuyan, M. V., Lee, J., Ward, I. M., Kim, J. E., Thompson, J. R., Chen, J. & Mer, G. (2006) Structural basis for the methylation state-specific recognition of histone H4-K20 by 53BP1 and Crb2 in DNA repair. *Cell,* **127,** 1361-1373.

Bras, S., Martin-Lanneree, S., Gobert, V., Auge, B., Breig, O., Sanial, M., Yamaguchi, M., Haenlin, M., Plessis, A. & Waltzer, L. (2012) Myeloid leukemia factor is a conserved regulator of RUNX transcription factor activity involved in hematopoiesis. *Proceedings of the National Academy of Sciences USA,* **109,** 4986-4991.

Brecqueville, M., Cervera, N., Gelsi-Boyer, V., Murati, A., Adelaide, J., Chaffanet, M., Rey, J., Vey, N., Mozziconacci, M. J. & Birnbaum, D. (2011) Rare mutations in DNMT3A in myeloproliferative neoplasms and myelodysplastic syndromes. *Blood Cancer Journal,* **1,** e18.

Broseus, J., Lippert, E., Harutyunyan, A. S., Jeromin, S., Zipperer, E., Florensa, L., Milosevic, J. D., Haferlach, T., Germing, U., Luno, E., Schnittger, S., Kralovics, R. & Girodon, F. (2014) Low rate of calreticulin mutations in refractory anaemia with ring sideroblasts and marked thrombocytosis. *Leukemia*.

Bruchova, H., Merkerova, M. & Prchal, J. T. (2008) Aberrant expression of microRNA in polycythemia vera. *Haematologica,* **93,** 1009-1016.

Bruchova, H., Yoon, D., Agarwal, A. M., Mendell, J. & Prchal, J. T. (2007) Regulated expression of microRNAs in normal and polycythemia vera erythropoiesis. *Experimental Hematology,* **35,** 1657-1667.

Bumm, T. G., Elsea, C., Corbin, A. S., Loriaux, M., Sherbenou, D., Wood, L., Deininger, J., Silver, R. T., Druker, B. J. & Deininger, M. W. (2006) Characterization of murine JAK2V617F-positive myeloproliferative disease. *Cancer Res,* **66,** 11156-11165.

Busque, L., Patel, J. P., Figueroa, M. E., Vasanthakumar, A., Provost, S., Hamilou, Z., Mollica, L., Li, J., Viale, A., Heguy, A., Hassimi, M., Socci, N., Bhatt, P. K., Gonen, M., Mason, C. E., Melnick, A., Godley, L. A., Brennan, C. W., Abdel-Wahab, O. & Levine, R. L. (2012) Recurrent somatic TET2 mutations in normal elderly individuals with clonal hematopoiesis. *Nature Genetics,* **44,** 1179-1181.

Butcher, C. M., Hahn, U., To, L. B., Gecz, J., Wilkins, E. J., Scott, H. S., Bardy, P. G. & D'andrea, R. J. (2008) Two novel JAK2 exon 12 mutations in JAK2V617F-negative polycythaemia vera patients. *Leukemia,* **22,** 870-873.

Callus, B. A., Ekert, P. G., Heraud, J. E., Jabbour, A. M., Kotevski, A., Vince, J. E., Silke, J. & Vaux, D. L. (2008) Cytoplasmic p53 is not required for PUMA-induced apoptosis. *Cell Death and Differentiation,* **15,** 213-215; author reply 215-216.

Campbell, H. D., Tucker, W. Q., Hort, Y., Martinson, M. E., Mayo, G., Clutterbuck, E. J., Sanderson, C. J. & Young, I. G. (1987) Molecular cloning, nucleotide sequence, and expression of the gene encoding human eosinophil differentiation factor (interleukin 5). *Proceedings of the National Academy of Sciences USA* **84,** 6629-6633.

Campbell, P. J. (2009) Somatic and germline genetics at the JAK2 locus. *Nature Genetics,* **41,** 385-386.

Campbell, P. J., Baxter, E. J., Beer, P. A., Scott, L. M., Bench, A. J., Huntly, B. J., Erber, W. N., Kusec, R., Larsen, T. S., Giraudier, S., Le Bousse-Kerdiles, M. C., Griesshammer, M., Reilly, J. T., Cheung, B. Y., Harrison, C. N. & Green, A. R. (2006) Mutation of JAK2 in the myeloproliferative disorders: timing,

clonality studies, cytogenetic associations, and role in leukemic transformation. *Blood,* **108,** 3548-3555.

Campbell, P. J. & Green, A. R. (2005) Management of polycythemia vera and essential thrombocythemia. *The American Society of Hematology Education Program Book,* 201-208.

Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R. S. & Zhang, Y. (2002) Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science,* **298,** 1039-1043.

Cao, R. & Zhang, Y. (2004) The functions of E(Z)/EZH2-mediated methylation of lysine 27 in histone H3. *Current opinion in Genetics and Development,* **14,** 155-164.

Caramazza, D., Begna, K. H., Gangat, N., Vaidya, R., Siragusa, S., Van Dyke, D. L., Hanson, C., Pardanani, A. & Tefferi, A. (2011) Refined cytogenetic-risk categorization for overall and leukemia-free survival in primary myelofibrosis: a single center study of 433 patients. *Leukemia,* **25,** 82-88.

Carbuccia, N., Murati, A., Trouplin, V., Brecqueville, M., Adelaide, J., Rey, J., Vainchenker, W., Bernard, O. A., Chaffanet, M., Vey, N., Birnbaum, D. & Mozziconacci, M. J. (2009) Mutations of ASXL1 gene in myeloproliferative neoplasms. *Leukemia,* **23,** 2183-2186.

Carbuccia, N., Trouplin, V., Gelsi-Boyer, V., Murati, A., Rocquain, J., Adelaide, J., Olschwang, S., Xerri, L., Vey, N., Chaffanet, M., Birnbaum, D. & Mozziconacci, M. J. (2010) Mutual exclusion of ASXL1 and NPM1 mutations in a series of acute myeloid leukemias. *Leukemia,* **24,** 469-473.

Cario, H., Goerttler, P. S., Steimle, C., Levine, R. L. & Pahl, H. L. (2005) The JAK2V617F mutation is acquired secondary to the predisposing alteration in familial polycythaemia vera. *British Journal of Haematology,* **130,** 800-801.

Ceredig, R., Rolink, A. G. & Brown, G. (2009) Models of haematopoiesis: seeing the wood for the trees. *Nature Reviews Immunology,* **9,** 293-300.

Cervantes, F., Dupriez, B., Pereira, A., Passamonti, F., Reilly, J. T., Morra, E., Vannucchi, A. M., Mesa, R. A., Demory, J. L., Barosi, G., Rumi, E. & Tefferi, A. (2009) New prognostic scoring system for primary myelofibrosis based on a study of the International Working Group for Myelofibrosis Research and Treatment. *Blood,* **113,** 2895-2901.

Cervantes, F., Tassies, D., Salgado, C., Rovira, M., Pereira, A. & Rozman, C. (1991) Acute transformation in nonleukemic chronic myeloproliferative disorders: actuarial probability and main characteristics in a series of 218 patients. *Acta Haematologica,* **85,** 124-127.

Chaligne, R., James, C., Tonetti, C., Besancenot, R., Le Couedic, J. P., Fava, F., Mazurier, F., Godin, I., Maloum, K., Larbret, F., Lecluse, Y., Vainchenker, W. & Giraudier, S. (2007) Evidence for MPL W515L/K mutations in hematopoietic stem cells in primitive myelofibrosis. *Blood,* **110,** 3735-3743.

Challen, G. A., Sun, D., Jeong, M., Luo, M., Jelinek, J., Berg, J. S., Bock, C., Vasanthakumar, A., Gu, H., Xi, Y., Liang, S., Lu, Y., Darlington, G. J., Meissner, A., Issa, J. P., Godley, L. A., Li, W. & Goodell, M. A. (2012) Dnmt3a is essential for hematopoietic stem cell differentiation. *Nature Genetics,* **44,** 23-31.

Chan, S. M. & Majeti, R. (2013) Role of DNMT3A, TET2, and IDH1/2 mutations in pre-leukemic stem cells in acute myeloid leukemia. *International Journal of Hematology,* **98,** 648-657.

Chao, M. M., Todd, M. A., Kontny, U., Neas, K., Sullivan, M. J., Hunter, A. G., Picketts, D. J. & Kratz, C. P. (2010) T-cell acute lymphoblastic leukemia in

association with Borjeson-Forssman-Lehmann syndrome due to a mutation in PHF6. *Pediatric Blood and Cancer,* **55,** 722-724.

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. & Green, A. R. (2010) Distinct clinical phenotypes associated with JAK2V617F reflect differential STAT1 signaling. *Cancer Cell,* **18,** 524-535.

Chen, T., Ueda, Y., Dodge, J. E., Wang, Z. & Li, E. (2003) Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. *Molecular and Cellular Biology,* **23,** 5594-5605.

Chittka, A. (2010) Dynamic distribution of histone H4 arginine 3 methylation marks in the developing murine cortex. *PLOS One,* **5,** e13807.

Chung, C. D., Liao, J., Liu, B., Rao, X., Jay, P., Berta, P. & Shuai, K. (1997) Specific inhibition of Stat3 signal transduction by PIAS3. *Science,* **278,** 1803-1805.

Coffer, P. J., Schweizer, R. C., Dubois, G. R., Maikoe, T., Lammers, J. W. & Koenderman, L. (1998) Analysis of signal transduction pathways in human eosinophils activated by chemoattractants and the T-helper 2-derived cytokines interleukin-4 and interleukin-5. *Blood,* **91,** 2547-2557.

Crawford, J., Lower, K. M., Hennekam, R. C., Van Esch, H., Megarbane, A., Lynch, S. A., Turner, G. & Gecz, J. (2006) Mutation screening in Borjeson-Forssman-Lehmann syndrome: identification of a novel de novo PHF6 mutation in a female patient. *Journal of Medical Genetics,* **43,** 238-243.

Crisa, E., Venturino, E., Passera, R., Prina, M., Schinco, P., Borchiellini, A., Giai, V., Ciocca Vasino, M. A., Bazzan, M., Vaccarino, A., Boccadoro, M. & Ferrero, D. (2010) A retrospective study on 226 polycythemia vera patients: impact of median hematocrit value on clinical outcomes and survival improvement with anti-thrombotic prophylaxis and non-alkylating drugs. *Annals of Hematology,* **89,** 691-699.

Croker, B. A., Kiu, H. & Nicholson, S. E. (2008) SOCS regulation of the JAK/STAT signalling pathway. *Seminars in Cell & Developmental Biology,* **19,** 414-422.

Cross, N. C. (2011) Genetic and epigenetic complexity in myeloproliferative neoplasms. *The American Society of Hematology Education Program Book,* **2011,** 208-214.

Cross, N. C. (2012) Histone modification defects in developmental disorders and cancer. *Oncotarget,* **3,** 3-4.

Dahl, R., Iyer, S. R., Owens, K. S., Cuylear, D. D. & Simon, M. C. (2007) The transcriptional repressor GFI-1 antagonizes PU.1 activity through protein-protein interaction. *The Journal of Biological Chemistry,* **282,** 6473-6483.

Dang, L., White, D. W., Gross, S., Bennett, B. D., Bittinger, M. A., Driggers, E. M., Fantin, V. R., Jang, H. G., Jin, S., Keenan, M. C., Marks, K. M., Prins, R. M., Ward, P. S., Yen, K. E., Liau, L. M., Rabinowitz, J. D., Cantley, L. C., Thompson, C. B., Vander Heiden, M. G. & Su, S. M. (2009) Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. *Nature,* **465,** 966.

Das, G. & Herr, W. (1993) Enhanced activation of the human histone H2B promoter by an Oct-1 variant generated by alternative splicing. *The Journal of Biological Chemistry,* **268,** 25026-25032.

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

De Klein, A., Van Kessel, A. G., Grosveld, G., Bartram, C. R., Hagemeijer, A., Bootsma, D., Spurr, N. K., Heisterkamp, N., Groffen, J. & Stephenson, J. R. (1982) A cellular oncogene is translocated to the Philadelphia chromosome in chronic myelocytic leukaemia. *Nature,* **300,** 765-767.

Delhommeau, F., Dupont, S., Della Valle, V., James, C., Trannoy, S., Masse, A., Kosmider, O., Le Couedic, J. P., Robert, F., Alberdi, A., Lecluse, Y., Plo, I., Dreyfus, F. J., Marzac, C., Casadevall, N., Lacombe, C., Romana, S. P., Dessen, P., Soulier, J., Viguie, F., Fontenay, M., Vainchenker, W. & Bernard, O. A. (2009) Mutation in TET2 in myeloid cancers. *The New England Journal of Medicine,* **360,** 2289-2301.

Di Nisio, M., Barbui, T., Di Gennaro, L., Borrelli, G., Finazzi, G., Landolfi, R., Leone, G., Marfisi, R., Porreca, E., Ruggeri, M., Rutjes, A. W., Tognoni, G., Vannucchi, A. M. & Marchioli, R. (2007) The haematocrit and platelet target in polycythemia vera. *British Journal of Haematology,* **136,** 249-259.

Dohner, K. & Dohner, H. (2008) Molecular characterization of acute myeloid leukemia. *Haematologica,* **93,** 976-982.

Donahue, R. E., Seehra, J., Metzger, M., Lefebvre, D., Rock, B., Carbone, S., Nathan, D. G., Garnick, M., Sehgal, P. K., Laston, D. & Et Al. (1988) Human IL-3 and GM-CSF act synergistically in stimulating hematopoiesis in primates. *Science,* **241,** 1820-1823.

Dorritie, K. A., Mccubrey, J. A. & Johnson, D. E. (2013) STAT transcription factors in hematopoiesis and leukemogenesis: opportunities for therapeutic intervention. *Leukemia,* **28,** 248-257.

Drachman, J. G., Griffin, J. D. & Kaushansky, K. (1995) The c-Mpl ligand (thrombopoietin) stimulates tyrosine phosphorylation of Jak2, Shc, and c-Mpl. *The Journal of Biological Chemistry,* **270,** 4979-4982.

Drachman, J. G., Sabath, D. F., Fox, N. E. & Kaushansky, K. (1997) Thrombopoietin signal transduction in purified murine megakaryocytes. *Blood,* **89,** 483-492.

Dupriez, B., Morel, P., Demory, J. L., Lai, J. L., Simon, M., Plantier, I. & Bauters, F. (1996) Prognostic factors in agnogenic myeloid metaplasia: a report on 195 cases with a new scoring system. *Blood,* **88,** 1013-1018.

Dusa, A., Mouton, C., Pecquet, C., Herman, M. & Constantinescu, S. N. (2010) JAK2 V617F constitutive activation requires JH2 residue F595: a pseudokinase domain target for specific inhibitors. *PLOS One,* **5,** e11157.

Ehrlich, M. (2003) The ICF syndrome, a DNA methyltransferase 3B deficiency and immunodeficiency disease. *Clinical immunology,* **109,** 17-28.

Ehrlich, M., Buchanan, K. L., Tsien, F., Jiang, G., Sun, B., Uicker, W., Weemaes, C. M., Smeets, D., Sperling, K., Belohradsky, B. H., Tommerup, N., Misek, D. E., Rouillard, J. M., Kuick, R. & Hanash, S. M. (2001) DNA methyltransferase 3B mutations linked to the ICF syndrome cause dysregulation of lymphogenesis genes. *Human Molecular Genetics,* **10,** 2917-2931.

Ekert, P. G., Read, S. H., Silke, J., Marsden, V. S., Kaufmann, H., Hawkins, C. J., Gerl, R., Kumar, S. & Vaux, D. L. (2004) Apaf-1 and caspase-9 accelerate apoptosis, but do not determine whether factor-deprived or drug-treated cells die. *The Journal of Cell Biology,* **165,** 835-842.

El Kassar, N., Hetet, G., Li, Y., Briere, J. & Grandchamp, B. (1995) Clonal analysis of haemopoietic cells in essential thrombocythaemia. *British Journal of Haematology,* **90,** 131-137.

Elena, C., Passamonti, F., Rumi, E., Malcovati, L., Arcaini, L., Boveri, E., Merli, M., Pietra, D., Pascutto, C. & Lazzarino, M. (2011a) Red blood cell transfusion-dependency implies a poor survival in primary myelofibrosis irrespective of IPSS and DIPSS. *Haematologica,* **96,** 167-170.

Elena, C., Rumi, E., Portolan, M., Della Porta, M. G., Pascutto, C. & Passamonti, F. (2011b) Flow-FISH evaluation of telomere length in Philadelphia-negative myeloproliferative neoplasms. *Haematologica,* **96,** 1236-1238.

Ellinger, J., Kahl, P., Von Der Gathen, J., Rogenhofer, S., Heukamp, L. C., Gutgemann, I., Walter, B., Hofstadter, F., Buttner, R., Muller, S. C., Bastian, P. J. & Von Ruecker, A. (2010) Global levels of histone modifications predict prostate cancer recurrence. *The Prostate,* **70,** 61-69.

Elliott, M. A. & Tefferi, A. (2005) Thrombosis and haemorrhage in polycythaemia vera and essential thrombocythaemia. *British Journal of Haematology,* **128,** 275-290.

Elsheikh, S. E., Green, A. R., Rakha, E. A., Powe, D. G., Ahmed, R. A., Collins, H. M., Soria, D., Garibaldi, J. M., Paish, C. E., Ammar, A. A., Grainge, M. J., Ball, G. R., Abdelghany, M. K., Martinez-Pomares, L., Heery, D. M. & Ellis, I. O. (2009) Global histone modifications in breast cancer correlate with tumor phenotypes, prognostic factors, and patient outcome. *Cancer Research,* **69,** 3802-3809.

Ernst, T., Chase, A. J., Score, J., Hidalgo-Curtis, C. E., Bryant, C., Jones, A. V., Waghorn, K., Zoi, K., Ross, F. M., Reiter, A., Hochhaus, A., Drexler, H. G., Duncombe, A., Cervantes, F., Oscier, D., Boultwood, J., Grand, F. H. & Cross, N. C. (2010) Inactivating mutations of the histone methyltransferase gene EZH2 in myeloid disorders. *Nature Genetics,* **42,** 722-726.

Fabbrizio, E., El Messaoudi, S., Polanowska, J., Paul, C., Cook, J. R., Lee, J. H., Negre, V., Rousset, M., Pestka, S., Le Cam, A. & Sardet, C. (2002) Negative regulation of transcription by the type II arginine methyltransferase PRMT5. *EMBO Reports,* **3,** 641-645.

Feener, E. P., Rosario, F., Dunn, S. L., Stancheva, Z. & Myers, M. G., Jr. (2004) Tyrosine phosphorylation of Jak2 in the JH2 domain inhibits cytokine signaling. *Molecular and Cellular Biology,* **24,** 4968-4978.

Feng, J., Witthuhn, B. A., Matsuda, T., Kohlhuber, F., Kerr, I. M. & Ihle, J. N. (1997) Activation of Jak2 catalytic activity requires phosphorylation of Y1007 in the kinase activation loop. *Molecular and Cellular Biology,* **17,** 2497-2501.

Figueroa, M. E., Abdel-Wahab, O., Lu, C., Ward, P. S., Patel, J., Shih, A., Li, Y., Bhagwat, N., Vasanthakumar, A., Fernandez, H. F., Tallman, M. S., Sun, Z., Wolniak, K., Peeters, J. K., Liu, W., Choe, S. E., Fantin, V. R., Paietta, E., Lowenberg, B., Licht, J. D., Godley, L. A., Delwel, R., Valk, P. J., Thompson, C. B., Levine, R. L. & Melnick, A. (2010) Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. *Cancer Cell,* **18,** 553-567.

Finazzi, G. & Barbui, T. (2008) Evidence and expertise in the management of polycythemia vera and essential thrombocythemia. *Leukemia* **22,** 1494-1502.

Finazzi, G., Caruso, V., Marchioli, R., Capnist, G., Chisesi, T., Finelli, C., Gugliotta, L., Landolfi, R., Kutti, J., Gisslinger, H., Marilus, R., Patrono, C., Pogliani, E. M., Randi, M. L., Villegas, A., Tognoni, G. & Barbui, T. (2005) Acute leukemia in polycythemia vera: an analysis of 1638 patients enrolled in a prospective observational study. *Blood,* **105,** 2664-2670.

Fisher, C. L., Pineault, N., Brookes, C., Helgason, C. D., Ohta, H., Bodner, C., Hess, J. L., Humphries, R. K. & Brock, H. W. (2010) Loss-of-function Additional sex combs like 1 mutations disrupt hematopoiesis but do not cause severe myelodysplasia or leukemia. *Blood,* **115,** 38-46.

Flowers, L. O., Johnson, H. M., Mujtaba, M. G., Ellis, M. R., Haider, S. M. & Subramaniam, P. S. (2004) Characterization of a peptide inhibitor of Janus kinase 2 that mimics suppressor of cytokine signaling 1 function. *Journal of Immunology,* **172,** 7510-7518.

Forsberg, L. A., Rasi, C., Razzaghian, H. R., Pakalapati, G., Waite, L., Thilbeault, K. S., Ronowicz, A., Wineinger, N. E., Tiwari, H. K., Boomsma, D., Westerman, M. P., Harris, J. R., Lyle, R., Essand, M., Eriksson, F., Assimes, T. L., Iribarren, C., Strachan, E., O'hanlon, T. P., Rider, L. G., Miller, F. W., Giedraitis, V., Lannfelt, L., Ingelsson, M., Piotrowski, A., Pedersen, N. L., Absher, D. & Dumanski, J. P. (2012) Age-related somatic structural changes in the nuclear genome of human blood cells. *The American Journal of Human Genetics,* **90,** 217-228.

Fortschegger, K. & Shiekhattar, R. (2011) Plant homeodomain fingers form a helping hand for transcription. *Epigenetics,* **6,** 4-8.

Fraga, M. F., Ballestar, E., Villar-Garea, A., Boix-Chornet, M., Espada, J., Schotta, G., Bonaldi, T., Haydon, C., Ropero, S., Petrie, K., Iyer, N. G., Perez-Rosado, A., Calvo, E., Lopez, J. A., Cano, A., Calasanz, M. J., Colomer, D., Piris, M. A., Ahn, N., Imhof, A., Caldas, C., Jenuwein, T. & Esteller, M. (2005) Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. *Nature Genetics,* **37,** 391-400.

Fruchtman, S. M., Mack, K., Kaplan, M. E., Peterson, P., Berk, P. D. & Wasserman, L. R. (1997) From efficacy to safety: a Polycythemia Vera Study group report on hydroxyurea in patients with polycythemia vera. *Seminars in Hematology,* **34,** 17-23.

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

Funakoshi-Tago, M., Pelletier, S., Moritake, H., Parganas, E. & Ihle, J. N. (2008) Jak2 FERM domain interaction with the erythropoietin receptor regulates Jak2 kinase activity. *Molecular and Cellular Biology,* **28,** 1792-1801.

Gadina, M., Hilton, D., Johnston, J. A., Morinobu, A., Lighvani, A., Zhou, Y. J., Visconti, R. & O'shea, J. J. (2001) Signaling by type I and II cytokine receptors: ten years after. *Current Opinion in Immunology,* **13,** 363-373.

Gangat, N., Caramazza, D., Vaidya, R., George, G., Begna, K., Schwager, S., Van Dyke, D., Hanson, C., Wu, W., Pardanani, A., Cervantes, F., Passamonti, F. & Tefferi, A. (2011) DIPSS plus: a refined Dynamic International Prognostic Scoring System for primary myelofibrosis that incorporates prognostic information from karyotype, platelet count, and transfusion status. *Journal of Clinical Oncology,* **29,** 392-397.

Gangat, N., Strand, J., Li, C. Y., Wu, W., Pardanani, A. & Tefferi, A. (2007a) Leucocytosis in polycythaemia vera predicts both inferior survival and leukaemic transformation. *British Journal of Haematology,* **138,** 354-358.

Gangat, N., Wolanskyj, A. P., Mcclure, R. F., Li, C. Y., Schwager, S., Wu, W. & Tefferi, A. (2007b) Risk stratification for survival and leukemic transformation in essential thrombocythemia: a single institutional study of 605 patients. *Leukemia,* **21,** 270-276.

Gasson, J. C., Golde, D. W., Kaufman, S. E., Westbrook, C. A., Hewick, R. M., Kaufman, R. J., Wong, G. G., Temple, P. A., Leary, A. C., Brown, E. L. & Et Al. (1985) Molecular characterization and expression of the gene encoding human erythroid-potentiating activity. *Nature,* **315,** 768-771.

Gelsi-Boyer, V., Trouplin, V., Adelaide, J., Bonansea, J., Cervera, N., Carbuccia, N., Lagarde, A., Prebet, T., Nezri, M., Sainty, D., Olschwang, S., Xerri, L., Chaffanet, M., Mozziconacci, M. J., Vey, N. & Birnbaum, D. (2009) Mutations of polycomb-associated gene ASXL1 in myelodysplastic

syndromes and chronic myelomonocytic leukaemia. *British Journal of Haematology,* **145,** 788-800.

Gilliland, D. G., Blanchard, K. L., Levy, J., Perrin, S. & Bunn, H. F. (1991) Clonality in myeloproliferative disorders: analysis by means of the polymerase chain reaction. *Proceedings of the National Academy of Sciences USA,* **88,** 6848-6852.

Giordanetto, F. & Kroemer, R. T. (2002) Prediction of the structure of human Janus kinase 2 (JAK2) comprising JAK homology domains 1 through 7. *Protein Engineering Design and Selection,* **15,** 727-737.

Gisslinger, H., Gotic, M., Holowiecki, J., Penka, M., Thiele, J., Kvasnicka, H. M., Kralovics, R. & Petrides, P. E. (2013) Anagrelide compared with hydroxyurea in WHO-classified essential thrombocythemia: the ANAHYDRET Study, a randomized controlled trial. *Blood,* **121,** 1720-1728.

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., Dohner, K., Vannucchi, A. M., Besses, C., Campbell, P. J. & 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,** 2704-2707.

Goldberg, A. D., Allis, C. D. & Bernstein, E. (2007) Epigenetics: a landscape takes shape. *Cell,* **128,** 635-638.

Goldin, L. R., Bjorkholm, M., Kristinsson, S. Y., Samuelsson, J. & Landgren, O. (2009) Germline and somatic JAK2 mutations and susceptibility to chronic myeloproliferative neoplasms. *Genome Medicine,* **1,** 55.

Gorantla, S. P., Dechow, T. N., Grundler, R., Illert, A. L., Zum Buschenfelde, C. M., Kremer, M., Peschel, C. & Duyster, J. (2010) Oncogenic JAK2V617F requires an intact SH2-like domain for constitutive activation and induction of a myeloproliferative disease in mice. *Blood,* **116,** 4600-4611.

Graham, F. L., Smiley, J., Russell, W. C. & Nairn, R. (1977) Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *The Journal of General Virology,* **36,** 59-74.

Green, A. & Beer, P. (2010) Somatic mutations of IDH1 and IDH2 in the leukemic transformation of myeloproliferative neoplasms. *The New England Journal of Medicine,* **362,** 369-370.

Greshock, J., Bachman, K. E., Degenhardt, Y. Y., Jing, J., Wen, Y. H., Eastman, S., Mcneil, E., Moy, C., Wegrzyn, R., Auger, K., Hardwicke, M. A. & Wooster, R. (2010) Molecular target class is predictive of in vitro response profile. *Cancer Research,* **70,** 3677-3686.

Gross, S., Cairns, R. A., Minden, M. D., Driggers, E. M., Bittinger, M. A., Jang, H. G., Sasaki, M., Jin, S., Schenkein, D. P., Su, S. M., Dang, L., Fantin, V. R. & Mak, T. W. (2010) Cancer-associated metabolite 2-hydroxyglutarate accumulates in acute myelogenous leukemia with isocitrate dehydrogenase 1 and 2 mutations. *The Journal of Experimental Medicine,* **207,** 339-344.

Growney, J. D., Shigematsu, H., Li, Z., Lee, B. H., Adelsperger, J., Rowan, R., Curley, D. P., Kutok, J. L., Akashi, K., Williams, I. R., Speck, N. A. & Gilliland, D. G. (2005) Loss of Runx1 perturbs adult hematopoiesis and is associated with a myeloproliferative phenotype. *Blood,* **106,** 494-504.

Guglielmelli, P., Lasho, T. L., Rotunno, G., Score, J., Mannarelli, C., Pancrazzi, A., Biamonte, F., Pardanani, A., Zoi, K., Reiter, A., Duncombe, A., Fanelli, T., Pietra, D., Rumi, E., Finke, C., Gangat, N., Ketterling, R. P., Knudson, R. A., Hanson, C. A., Bosi, A., Pereira, A., Manfredini, R., Cervantes, F., Barosi,

G., Cazzola, M., Cross, N. C., Vannucchi, A. M. & Tefferi, A. (2014) The number of prognostically detrimental mutations and prognosis in primary myelofibrosis: an international study of 797 patients. *Leukemia*.

Guglielmelli, P., Tozzi, L., Pancrazzi, A., Bogani, C., Antonioli, E., Ponziani, V., Poli, G., Zini, R., Ferrari, S., Manfredini, R., Bosi, A. & Vannucchi, A. M. (2007) MicroRNA expression profile in granulocytes from primary myelofibrosis patients. *Experimental Hematology,* **35,** 1708-1718.

Gunjan, A. & Verreault, A. (2003) A Rad53 kinase-dependent surveillance mechanism that regulates histone protein levels in S. cerevisiae. *Cell,* **115,** 537-549.

Hagiwara, T., Hidaka, Y. & Yamada, M. (2005) Deimination of histone H2A and H4 at arginine 3 in HL-60 granulocytes. *Biochemistry,* **44,** 5827-5834.

Hanasoge, S. & Ljungman, M. (2007) H2AX phosphorylation after UV irradiation is triggered by DNA repair intermediates and is mediated by the ATR kinase. *Carcinogenesis,* **28,** 2298-2304.

Harrison, C. N., Campbell, P. J., Buck, G., Wheatley, K., East, C. L., Bareford, D., Wilkins, B. S., Van Der Walt, J. D., Reilly, J. T., Grigg, A. P., Revell, P., Woodcock, B. E. & Green, A. R. (2005) Hydroxyurea compared with anagrelide in high-risk essential thrombocythemia. *The New England Journal of Medicine,* **353,** 33-45.

Hasan, S., Cassinat, B., Droin, N., Le Couedic, J. P., Favale, F., Monte-Mor, B., Lacout, C., Fontenay, M., Dosquet, C., Chomienne, C., Solary, E., Villeval, J. L., Casadevall, N., Kiladjian, J. J., Vainchenker, W. & Plo, I. (2014) Use of the 46/1 haplotype to model JAK2(V617F) clonal architecture in PV patients: clonal evolution and impact of IFNalpha treatment. *Leukemia,* **28,** 460-463.

Heath, V., Suh, H. C., Holman, M., Renn, K., Gooya, J. M., Parkin, S., Klarmann, K. D., Ortiz, M., Johnson, P. & Keller, J. (2004) C/EBPalpha deficiency results in hyperproliferation of hematopoietic progenitor cells and disrupts macrophage development in vitro and in vivo. *Blood,* **104,** 1639-1647.

Herrera-Merchan, A., Arranz, L., Ligos, J. M., De Molina, A., Dominguez, O. & Gonzalez, S. (2012) Ectopic expression of the histone methyltransferase Ezh2 in haematopoietic stem cells causes myeloproliferative disease. *Nature Communications,* **3,** 623.

Herrero, A. B. & Moreno, S. (2011) Lsm1 promotes genomic stability by controlling histone mRNA decay. *The EMBO Journal,* **30,** 2008-2018.

Holmes, W. F., Braastad, C. D., Mitra, P., Hampe, C., Doenecke, D., Albig, W., Stein, J. L., Van Wijnen, A. J. & Stein, G. S. (2005) Coordinate control and selective expression of the full complement of replication-dependent histone H4 genes in normal and cancer cells. *The Journal of Biological Chemistry,* **280,** 37400-37407.

Holz-Schietinger, C., Matje, D. M. & Reich, N. O. (2012) Mutations in DNA methyltransferase (DNMT3A) observed in acute myeloid leukemia patients disrupt processive methylation. *The Journal of Biological Chemistry,* **287,** 30941-30951.

Hou, H. A., Kuo, Y. Y., Chou, W. C., Chen, P. H. & Tien, H. F. (2014) Calreticulin mutation was rarely detected in patients with myelodysplastic syndrome. *Leukemia*.

Hsieh, C. L. (2005) The de novo methylation activity of Dnmt3a is distinctly different than that of Dnmt1. *BMC Biochemistry,* **6,** 6.

Huang, H., Maertens, A. M., Hyland, E. M., Dai, J., Norris, A., Boeke, J. D. & Bader, J. S. (2009) HistoneHits: a database for histone mutations and their phenotypes. *Genome Research,* **19,** 674-681.

Huang, S., Litt, M. & Felsenfeld, G. (2005) Methylation of histone H4 by arginine methyltransferase PRMT1 is essential in vivo for many subsequent histone modifications. *Genes and Development,* **19,** 1885-1893.

Huether, R., Dong, L., Chen, X., Wu, G., Parker, M., Wei, L., Ma, J., Edmonson, M. N., Hedlund, E. K., Rusch, M. C., Shurtleff, S. A., Mulder, H. L., Boggs, K., Vadordaria, B., Cheng, J., Yergeau, D., Song, G., Becksfort, J., Lemmon, G., Weber, C., Cai, Z., Dang, J., Walsh, M., Gedman, A. L., Faber, Z., Easton, J., Gruber, T., Kriwacki, R. W., Partridge, J. F., Ding, L., Wilson, R. K., Mardis, E. R., Mullighan, C. G., Gilbertson, R. J., Baker, S. J., Zambetti, G., Ellison, D. W., Zhang, J. & Downing, J. R. (2014) The landscape of somatic mutations in epigenetic regulators across 1,000 paediatric cancer genomes. *Nature Communications,* **5,** 3630.

Hussein, K., Bock, O., Theophile, K., Schlue, J., Ballmaier, M., Kroger, N., Gohring, G., Busche, G. & Kreipe, H. (2009) Biclonal expansion and heterogeneous lineage involvement in a case of chronic myeloproliferative disease with concurrent MPLW515L/JAK2V617F mutation. *Blood,* **113,** 1391-1392.

Hussein, K., Pardanani, A. D., Van Dyke, D. L., Hanson, C. A. & Tefferi, A. (2010) International Prognostic Scoring System-independent cytogenetic risk categorization in primary myelofibrosis. *Blood,* **115,** 496-499.

Ihle, J. N. (1995) The Janus protein tyrosine kinases in hematopoietic cytokine signaling. *Seminars in Immunology,* **7,** 247-254.

Ihle, J. N. & Gilliland, D. G. (2007) Jak2: normal function and role in hematopoietic disorders. *Current Opinion in Genetics and Development,* **17,** 8-14.

Ihle, J. N., Witthuhn, B. A., Quelle, F. W., Silvennoinen, O., Tang, B. & Yi, T. (1994) Protein tyrosine phosphorylation in the regulation of hematopoiesis by receptors of the cytokine-receptor superfamily. *Blood Cells,* **20,** 65-80; discussion 80-62.

Imielinski, M., Berger, A. H., Hammerman, P. S., Hernandez, B., Pugh, T. J., Hodis, E., Cho, J., Suh, J., Capelletti, M., Sivachenko, A., Sougnez, C., Auclair, D., Lawrence, M. S., Stojanov, P., Cibulskis, K., Choi, K., De Waal, L., Sharifnia, T., Brooks, A., Greulich, H., Banerji, S., Zander, T., Seidel, D., Leenders, F., Ansen, S., Ludwig, C., Engel-Riedel, W., Stoelben, E., Wolf, J., Goparju, C., Thompson, K., Winckler, W., Kwiatkowski, D., Johnson, B. E., Janne, P. A., Miller, V. A., Pao, W., Travis, W. D., Pass, H. I., Gabriel, S. B., Lander, E. S., Thomas, R. K., Garraway, L. A., Getz, G. & Meyerson, M. (2012) Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. *Cell,* **150,** 1107-1120.

Inoue, D., Kitaura, J., Togami, K., Nishimura, K., Enomoto, Y., Uchida, T., Kagiyama, Y., Kawabata, K. C., Nakahara, F., Izawa, K., Oki, T., Maehara, A., Isobe, M., Tsuchiya, A., Harada, Y., Harada, H., Ochiya, T., Aburatani, H., Kimura, H., Thol, F., Heuser, M., Levine, R. L., Abdel-Wahab, O. & Kitamura, T. (2013) Myelodysplastic syndromes are induced by histone methylation-altering ASXL1 mutations. *The Journal of Clinical Investigation,* **123,** 4627-4640.

Ito, S., Shen, L., Dai, Q., Wu, S. C., Collins, L. B., Swenberg, J. A., He, C. & Zhang, Y. (2011) Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science,* **333,** 1300-1303.

Iwasaki, H., Somoza, C., Shigematsu, H., Duprez, E. A., Iwasaki-Arai, J., Mizuno, S., Arinobu, Y., Geary, K., Zhang, P., Dayaram, T., Fenyus, M. L., Elf, S., Chan, S., Kastner, P., Huettner, C. S., Murray, R., Tenen, D. G. & Akashi, K. (2005) Distinctive and indispensable roles of PU.1 in maintenance of hematopoietic stem cells and their differentiation. *Blood,* **106,** 1590-1600.

Jabbour, A. M., Daunt, C. P., Green, B. D., Vogel, S., Gordon, L., Lee, R. S., Silke, N., Pearson, R. B., Vandenberg, C. J., Kelly, P. N., Nutt, S. L., Strasser, A., Borner, C. & Ekert, P. G. (2010) Myeloid progenitor cells lacking p53 exhibit delayed up-regulation of Puma and prolonged survival after cytokine deprivation. *Blood,* **115,** 344-352.

Jabbour, E. & Kantarjian, H. (2012) Chronic myeloid leukemia: 2012 update on diagnosis, monitoring, and management. *American Journal of Hematology,* **87,** 1037-1045.

Jacobs, K. B.&Yeager, M.&Zhou, W.&Wacholder, S.&Wang, Z.&Rodriguez-Santiago, B.&Hutchinson, A.&Deng, X.&Liu, C.&Horner, M. J.&Cullen, M.&Epstein, C. G.&Burdett, L.&Dean, M. C.&Chatterjee, N.&Sampson, J.&Chung, C. C.&Kovaks, J.&Gapstur, S. M.&Stevens, V. L.&Teras, L. T.&Gaudet, M. M.&Albanes, D.&Weinstein, S. J.&Virtamo, J.&Taylor, P. R.&Freedman, N. D.&Abnet, C. C.&Goldstein, A. M.&Hu, N.&Yu, K.&Yuan, J. M.&Liao, L.&Ding, T.&Qiao, Y. L.&Gao, Y. T.&Koh, W. P.&Xiang, Y. B.&Tang, Z. Z.&Fan, J. H.&Aldrich, M. C.&Amos, C.&Blot, W. J.&Bock, C. H.&Gillanders, E. M.&Harris, C. C.&Haiman, C. A.&Henderson, B. E.&Kolonel, L. N.&Le Marchand, L.&Mcneill, L. H.&Rybicki, B. A.&Schwartz, A. G.&Signorello, L. B.&Spitz, M. R.&Wiencke, J. K.&Wrensch, M.&Wu, X.&Zanetti, K. A.&Ziegler, R. G.&Figueroa, J. D.&Garcia-Closas, M.&Malats, N.&Marenne, G.&Prokunina-Olsson, L.&Baris, D.&Schwenn, M.&Johnson, A.&Landi, M. T.&Goldin, L.&Consonni, D.&Bertazzi, P. A.&Rotunno, M.&Rajaraman, P.&Andersson, U.&Beane Freeman, L. E.&Berg, C. D.&Buring, J. E.&Butler, M. A.&Carreon, T.&Feychting, M.&Ahlbom, A.&Gaziano, J. M.&Giles, G. G.&Hallmans, G.&Hankinson, S. E.&Hartge, P.&Henriksson, R.&Inskip, P. D.&Johansen, C.&Landgren, A.&Mckean-Cowdin, R.&Michaud, D. S.&Melin, B. S.&Peters, U.&Ruder, A. M.&Sesso, H. D.&Severi, G.&Shu, X. O.&Visvanathan, K.&White, E.&Wolk, A.&Zeleniuch-Jacquotte, A.&Zheng, W.&Silverman, D. T.&Kogevinas, M.&Gonzalez, J. R.&Villa, O.&Li, D.&Duell, E. J.&Risch, H. A.&Olson, S. H.&Kooperberg, C.&Wolpin, B. M.&Jiao, L.&Hassan, M.&Wheeler, W.&Arslan, A. A.&Bueno-De-Mesquita, H. B.&Fuchs, C. S.&Gallinger, S.&Gross, M. D.&Holly, E. A.&Klein, A. P.&Lacroix, A.&Mandelson, M. T.&Petersen, G.&Boutron-Ruault, M. C.&Bracci, P. M.&Canzian, F.&Chang, K.&Cotterchio, M.&Giovannucci, E. L.&Goggins, M.&Hoffman Bolton, J. A.&Jenab, M.&Khaw, K. T.&Krogh, V.&Kurtz, R. C.&Mcwilliams, R. R.&Mendelsohn, J. B.&Rabe, K. G.&Riboli, E.&Tjonneland, A.&Tobias, G. S.&Trichopoulos, D.&Elena, J. W.&Yu, H.&Amundadottir, L.&Stolzenberg-Solomon, R. Z.&Kraft, P.&Schumacher, F.&Stram, D.&Savage, S. A.&Mirabello, L.&Andrulis, I. L.&Wunder, J. S.&Patino Garcia, A.&Sierrasesumaga, L.&Barkauskas, D. A.&Gorlick, R. G.&Purdue, M.&Chow, W. H.&Moore, L. E.&Schwartz, K. L.&Davis, F. G.&Hsing, A. W.&Berndt, S. I.&Black, A.&Wentzensen, N.&Brinton, L. A.&Lissowska, J.&Peplonska, B.&Mcglynn, K. A.&Cook, M. B.&Graubard, B. I.&Kratz, C. P.&Greene, M. H.&Erickson, R. L.&Hunter, D. J.&Thomas, G.&Hoover, R. N.&Real, F. X.&Fraumeni, J. F., Jr.&Caporaso, N. E.&Tucker, M.&Rothman, N.&Perez-Jurado, L. A. & Chanock, S. J. (2012) Detectable clonal mosaicism and its relationship to aging and cancer. *Nature Genetics,* **44,** 651-658.

Jaeger, S., Martin, F., Rudinger-Thirion, J., Giege, R. & Eriani, G. (2006) Binding of human SLBP on the 3'-UTR of histone precursor H4-12 mRNA induces

structural rearrangements that enable U7 snRNA anchoring. *Nucleic Acids Research,* **34,** 4987-4995.

Jamal, R., Belisle, C., Lessard, M. C., Hebert, J., Roy, D. C., Levine, R. & Busque, L. (2008) Evidence suggesting the presence of a stem cell clone anteceding the acquisition of the JAK2-V617F mutation. *Leukemia,* **22,** 1472-1474.

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

Jamieson, C. H., Gotlib, J., Durocher, J. A., Chao, M. P., Mariappan, M. R., Lay, M., Jones, C., Zehnder, J. L., Lilleberg, S. L. & 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 USA* **103,** 6224-6229.

Jatiani, S. S., Baker, S. J., Silverman, L. R. & Reddy, E. P. (2010) Jak/STAT pathways in cytokine signaling and myeloproliferative disorders: approaches for targeted therapies. *Genes and Cancer,* **1,** 979-993.

Je, E. M., Yoo, N. J., Kim, Y. J., Kim, M. S. & Lee, S. H. (2013) Somatic mutation of H3F3A, a chromatin remodeling gene, is rare in acute leukemias and non-Hodgkin lymphoma. *European Journal of Haematology,* **90,** 169-170.

Je, E. M., Yoo, N. J. & Lee, S. H. (2014) Mutational analysis of H3F3A, a chromatin remodeling gene in common solid tumors. *Acta Pathologica, Microbiologica, et Immunologica Scandinavica,* **122,** 81-82.

Jia, D., Jurkowska, R. Z., Zhang, X., Jeltsch, A. & Cheng, X. (2007) Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation. *Nature,* **449,** 248-251.

Jones, A. V., Chase, A., Silver, R. T., Oscier, D., Zoi, K., Wang, Y. L., Cario, H., Pahl, H. L., Collins, A., Reiter, A., Grand, F. & Cross, N. C. (2009a) JAK2 haplotype is a major risk factor for the development of myeloproliferative neoplasms. *Nature Genetics,* **41,** 446-449.

Jones, A. V. & Cross, N. C. (2013) Inherited predisposition to myeloproliferative neoplasms. *Therapeutic Advances in Hematology,* **4,** 237-253.

Jones, J. E., Causey, C. P., Knuckley, B., Slack-Noyes, J. L. & Thompson, P. R. (2009b) Protein arginine deiminase 4 (PAD4): Current understanding and future therapeutic potential. *Current Opinion in Drug Discovery and Development,* **12,** 616-627.

Joseph, C. G., Hwang, H., Jiao, Y., Wood, L. D., Kinde, I., Wu, J., Mandahl, N., Luo, J., Hruban, R. H., Diaz, L. A., Jr., He, T. C., Vogelstein, B., Kinzler, K. W., Mertens, F. & Papadopoulos, N. (2014) Exomic analysis of myxoid liposarcomas, synovial sarcomas, and osteosarcomas. *Genes Chromosomes Cancer,* **53,** 15-24.

Kang, M. R., Kim, M. S., Oh, J. E., Kim, Y. R., Song, S. Y., Seo, S. I., Lee, J. Y., Yoo, N. J. & Lee, S. H. (2009) Mutational analysis of IDH1 codon 132 in glioblastomas and other common cancers. *International Journal of Cancer,* **125,** 353-355.

Kaushansky, K., Lok, S., Holly, R. D., Broudy, V. C., Lin, N., Bailey, M. C., Forstrom, J. W., Buddle, M. M., Oort, P. J., Hagen, F. S. & Et Al. (1994) Promotion of megakaryocyte progenitor expansion and differentiation by the c-Mpl ligand thrombopoietin. *Nature,* **369,** 568-571.

Khan, S. N., Jankowska, A. M., Mahfouz, R., Dunbar, A. J., Sugimoto, Y., Hosono, N., Hu, Z., Cheriyath, V., Vatolin, S., Przychodzen, B., Reu, F. J.,

Saunthararajah, Y., O'keefe, C., Sekeres, M. A., List, A. F., Moliterno, A. R., Mcdevitt, M. A., Maciejewski, J. P. & Makishima, H. (2013) Multiple mechanisms deregulate EZH2 and histone H3 lysine 27 epigenetic changes in myeloid malignancies. *Leukemia,* **27,** 1301-1309.

Kilpivaara, O., Mukherjee, S., Schram, A. M., Wadleigh, M., Mullally, A., Ebert, B. L., Bass, A., Marubayashi, S., Heguy, A., Garcia-Manero, G., Kantarjian, H., Offit, K., Stone, R. M., Gilliland, D. G., Klein, R. J. & Levine, R. L. (2009) A germline JAK2 SNP is associated with predisposition to the development of JAK2(V617F)-positive myeloproliferative neoplasms. *Nature Genetics,* **41,** 455-459.

Klampfl, T., Gisslinger, H., Harutyunyan, A. S., Nivarthi, H., Rumi, E., Milosevic, J. D., Them, N. 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., Schonegger, A., Bock, C., Malcovati, L., Pascutto, C., Superti-Furga, G., Cazzola, M. & Kralovics, R. (2013) Somatic mutations of calreticulin in myeloproliferative neoplasms. *The New England Journal of Medicine,* **369,** 2379-2390.

Ko, M., Huang, Y., Jankowska, A. M., Pape, U. J., Tahiliani, M., Bandukwala, H. S., An, J., Lamperti, E. D., Koh, K. P., Ganetzky, R., Liu, X. S., Aravind, L., Agarwal, S., Maciejewski, J. P. & Rao, A. (2010) Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. *Nature,* **468,** 839-843.

Koh, K. P., Yabuuchi, A., Rao, S., Huang, Y., Cunniff, K., Nardone, J., Laiho, A., Tahiliani, M., Sommer, C. A., Mostoslavsky, G., Lahesmaa, R., Orkin, S. H., Rodig, S. J., Daley, G. Q. & Rao, A. (2011) Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. *Cell Stem Cell,* **8,** 200-213.

Kornberg, R. D. & Lorch, Y. (1999) Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. *Cell,* **98,** 285-294.

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

Kralovics, R., Passamonti, F., Buser, A. S., Teo, S. S., Tiedt, R., Passweg, J. R., Tichelli, A., Cazzola, M. & Skoda, R. C. (2005) A gain-of-function mutation of JAK2 in myeloproliferative disorders. *The New England Journal of Medicine,* **352,** 1779-1790.

Kralovics, R., Stockton, D. W. & Prchal, J. T. (2003) Clonal hematopoiesis in familial polycythemia vera suggests the involvement of multiple mutational events in the early pathogenesis of the disease. *Blood,* **102,** 3793-3796.

Kralovics, R., Teo, S. S., Li, S., Theocharides, A., Buser, A. S., Tichelli, A. & Skoda, R. C. (2006) Acquisition of the V617F mutation of JAK2 is a late genetic event in a subset of patients with myeloproliferative disorders. *Blood,* **108,** 1377-1380.

Krauthammer, M., Kong, Y., Ha, B. H., Evans, P., Bacchiocchi, A., Mccusker, J. P., Cheng, E., Davis, M. J., Goh, G., Choi, M., Ariyan, S., Narayan, D., Dutton-Regester, K., Capatana, A., Holman, E. C., Bosenberg, M., Sznol, M., Kluger, H. M., Brash, D. E., Stern, D. F., Materin, M. A., Lo, R. S., Mane, S., Ma, S., Kidd, K. K., Hayward, N. K., Lifton, R. P., Schlessinger, J., Boggon, T. J. & Halaban, R. (2012) Exome sequencing identifies recurrent somatic RAC1 mutations in melanoma. *Nature Genetics,* **44,** 1006-1014.

Kurzer, J. H., Argetsinger, L. S., Zhou, Y. J., Kouadio, J. L., O'shea, J. J. & Carter-Su, C. (2004) Tyrosine 813 is a site of JAK2 autophosphorylation critical for

activation of JAK2 by SH2-B beta. *Molecular and Cellular Biology,* **24,** 4557-4570.

Kwon, M. J., Kim, S. S., Choi, Y. L., Jung, H. S., Balch, C., Kim, S. H., Song, Y. S., Marquez, V. E., Nephew, K. P. & Shin, Y. K. (2010) Derepression of CLDN3 and CLDN4 during ovarian tumorigenesis is associated with loss of repressive histone modifications. *Carcinogenesis,* **31,** 974-983.

Lacout, C., Pisani, D. F., Tulliez, M., Gachelin, F. M., Vainchenker, W. & Villeval, J. L. (2006) JAK2V617F expression in murine hematopoietic cells leads to MPD mimicking human PV with secondary myelofibrosis. *Blood,* **108,** 1652-1660.

Lai, A. Y. & Wade, P. A. (2011) Cancer biology and NuRD: a multifaceted chromatin remodelling complex. *Nature Reviews Cancer,* **11,** 588-596.

Landgren, O., Goldin, L. R., Kristinsson, S. Y., Helgadottir, E. A., Samuelsson, J. & Bjorkholm, M. (2008) Increased risks of polycythemia vera, essential thrombocythemia, and myelofibrosis among 24,577 first-degree relatives of 11,039 patients with myeloproliferative neoplasms in Sweden. *Blood,* **112,** 2199-2204.

Langemeijer, S. M., Kuiper, R. P., Berends, M., Knops, R., Aslanyan, M. G., Massop, M., Stevens-Linders, E., Van Hoogen, P., Van Kessel, A. G., Raymakers, R. A., Kamping, E. J., Verhoef, G. E., Verburgh, E., Hagemeijer, A., Vandenberghe, P., De Witte, T., Van Der Reijden, B. A. & Jansen, J. H. (2009) Acquired mutations in TET2 are common in myelodysplastic syndromes. *Nature Genetics,* **41,** 838-842.

Larsen, T. S., Christensen, J. H., Hasselbalch, H. C. & Pallisgaard, N. (2007) The JAK2 V617F mutation involves B- and T-lymphocyte lineages in a subgroup of patients with Philadelphia-chromosome negative chronic myeloproliferative disorders. *British Journal of Haematology,* **136,** 745-751.

Lasho, T. L., Pardanani, A., Mcclure, R. F., Mesa, R. A., Levine, R. L., Gilliland, D. G. & Tefferi, A. (2006) Concurrent MPL515 and JAK2V617F mutations in myelofibrosis: chronology of clonal emergence and changes in mutant allele burden over time. *British Journal of Haematology,* **135,** 683-687.

Laslo, P., Spooner, C. J., Warmflash, A., Lancki, D. W., Lee, H. J., Sciammas, R., Gantner, B. N., Dinner, A. R. & Singh, H. (2006) Multilineage transcriptional priming and determination of alternate hematopoietic cell fates. *Cell,* **126,** 755-766.

Laurie, C. C., Laurie, C. A., Rice, K., Doheny, K. F., Zelnick, L. R., Mchugh, C. P., Ling, H., Hetrick, K. N., Pugh, E. W., Amos, C., Wei, Q., Wang, L. E., Lee, J. E., Barnes, K. C., Hansel, N. N., Mathias, R., Daley, D., Beaty, T. H., Scott, A. F., Ruczinski, I., Scharpf, R. B., Bierut, L. J., Hartz, S. M., Landi, M. T., Freedman, N. D., Goldin, L. R., Ginsburg, D., Li, J., Desch, K. C., Strom, S. S., Blot, W. J., Signorello, L. B., Ingles, S. A., Chanock, S. J., Berndt, S. I., Le Marchand, L., Henderson, B. E., Monroe, K. R., Heit, J. A., De Andrade, M., Armasu, S. M., Regnier, C., Lowe, W. L., Hayes, M. G., Marazita, M. L., Feingold, E., Murray, J. C., Melbye, M., Feenstra, B., Kang, J. H., Wiggs, J. L., Jarvik, G. P., Mcdavid, A. N., Seshan, V. E., Mirel, D. B., Crenshaw, A., Sharopova, N., Wise, A., Shen, J., Crosslin, D. R., Levine, D. M., Zheng, X., Udren, J. I., Bennett, S., Nelson, S. C., Gogarten, S. M., Conomos, M. P., Heagerty, P., Manolio, T., Pasquale, L. R., Haiman, C. A., Caporaso, N. & Weir, B. S. (2012) Detectable clonal mosaicism from birth to old age and its relationship to cancer. *Nature Genetics,* **44,** 642-650.

Lee, J. H., Cook, J. R., Yang, Z. H., Mirochnitchenko, O., Gunderson, S. I., Felix, A. M., Herth, N., Hoffmann, R. & Pestka, S. (2005) PRMT7, a new protein

arginine methyltransferase that synthesizes symmetric dimethylarginine. *The Journal of Biological Chemistry,* **280,** 3656-3664.

Lee, W. H., Salek-Ardakani, S., Pandolfi, P. P., Brady, H. J., De Boer, J. & Williams, O. (2012) NPM-MLF1 synergizes with Npm haploinsufficiency to enhance myeloid progenitor activity. *Leukemia,* **26,** 1110-1112.

Levine, R. L., Belisle, C., Wadleigh, M., Zahrieh, D., Lee, S., Chagnon, P., Gilliland, D. G. & Busque, L. (2006) X-inactivation-based clonality analysis and quantitative JAK2V617F assessment reveal a strong association between clonality and JAK2V617F in PV but not ET/MMM, and identifies a subset of JAK2V617F-negative ET and MMM patients with clonal hematopoiesis. *Blood,* **107,** 4139-4141.

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., Frohling, S., Dohner, 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. & Gilliland, D. G. (2005) Activating mutation in the tyrosine kinase JAK2 in polycythemia vera, essential thrombocythemia, and myeloid metaplasia with myelofibrosis. *Cancer Cell,* **7,** 387-397.

Lewis, P. W., Elsaesser, S. J., Noh, K. M., Stadler, S. C. & Allis, C. D. (2010) Daxx is an H3.3-specific histone chaperone and cooperates with ATRX in replication-independent chromatin assembly at telomeres. *Proceedings of the National Academy of Sciences USA,* **107,** 14075-14080.

Lewis, P. W., Muller, M. M., Koletsky, M. S., Cordero, F., Lin, S., Banaszynski, L. A., Garcia, B. A., Muir, T. W., Becher, O. J. & Allis, C. D. (2013) Inhibition of PRC2 activity by a gain-of-function H3 mutation found in pediatric glioblastoma. *Science,* **340,** 857-861.

Ley, T. J., Ding, L., Walter, M. J., Mclellan, M. D., Lamprecht, T., Larson, D. E., Kandoth, C., Payton, J. E., Baty, J., Welch, J., Harris, C. C., Lichti, C. F., Townsend, R. R., Fulton, R. S., Dooling, D. J., Koboldt, D. C., Schmidt, H., Zhang, Q., Osborne, J. R., Lin, L., O'laughlin, M., Mcmichael, J. F., Delehaunty, K. D., Mcgrath, S. D., Fulton, L. A., Magrini, V. J., Vickery, T. L., Hundal, J., Cook, L. L., Conyers, J. J., Swift, G. W., Reed, J. P., Alldredge, P. A., Wylie, T., Walker, J., Kalicki, J., Watson, M. A., Heath, S., Shannon, W. D., Varghese, N., Nagarajan, R., Westervelt, P., Tomasson, M. H., Link, D. C., Graubert, T. A., Dipersio, J. F., Mardis, E. R. & Wilson, R. K. (2010) DNMT3A mutations in acute myeloid leukemia. *The New England Journal of Medicine,* **363,** 2424-2433.

Li, J. (2011) Quiescence regulators for hematopoietic stem cell. *Experimental Hematology,* **39,** 511-520.

Li, J., Kent, D. G., Chen, E. & Green, A. R. (2011a) Mouse models of myeloproliferative neoplasms: JAK of all grades. *Disease Models & Mechanisms,* **4,** 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. & Green, A. R. (2014) JAK2V617F-homozygosity drives a phenotypic switch between myeloproliferative neoplasms in a murine model, but is insufficient to sustain clonal expansion. *Blood*.

Li, M., Zhao, H., Zhang, X., Wood, L. D., Anders, R. A., Choti, M. A., Pawlik, T. M., Daniel, H. D., Kannangai, R., Offerhaus, G. J., Velculescu, V. E., Wang, L., Zhou, S., Vogelstein, B., Hruban, R. H., Papadopoulos, N., Cai, J.,

Torbenson, M. S. & Kinzler, K. W. (2011b) Inactivating mutations of the chromatin remodeling gene ARID2 in hepatocellular carcinoma. *Nature Genetics,* **43,** 828-829.

Li, S., Kralovics, R., De Libero, G., Theocharides, A., Gisslinger, H. & Skoda, R. C. (2008) Clonal heterogeneity in polycythemia vera patients with JAK2 exon12 and JAK2-V617F mutations. *Blood,* **111,** 3863-3866.

Li, X., Hu, X., Patel, B., Zhou, Z., Liang, S., Ybarra, R., Qiu, Y., Felsenfeld, G., Bungert, J. & Huang, S. (2010) H4R3 methylation facilitates beta-globin transcription by regulating histone acetyltransferase binding and H3 acetylation. *Blood,* **115,** 2028-2037.

Li, X., Yao, H., Chen, Z., Wang, Q., Zhao, Y. & Chen, S. (2012) Somatic mutations of PHF6 in patients with chronic myeloid leukemia in blast crisis. *Leukemia and Lymphoma*.

Liang, D., Burkhart, S. L., Singh, R. K., Kabbaj, M. H. & Gunjan, A. (2012) Histone dosage regulates DNA damage sensitivity in a checkpoint-independent manner by the homologous recombination pathway. *Nucleic Acids Research,* **40,** 9604-9620.

Lin, C. J., Conti, M. & Ramalho-Santos, M. (2013) Histone variant H3.3 maintains a decondensed chromatin state essential for mouse preimplantation development. *Development,* **140,** 3624-3634.

Lindberg, J., Mills, I. G., Klevebring, D., Liu, W., Neiman, M., Xu, J., Wikstrom, P., Wiklund, P., Wiklund, F., Egevad, L. & Gronberg, H. (2013) The mitochondrial and autosomal mutation landscapes of prostate cancer. *European Urology,* **63,** 702-708.

Liu, B., Liao, J., Rao, X., Kushner, S. A., Chung, C. D., Chang, D. D. & Shuai, K. (1998) Inhibition of Stat1-mediated gene activation by PIAS1. *Proceedings of the National Academy of Sciences USA,* **95,** 10626-10631.

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. & Nimer, S. D. (2011) JAK2V617F-mediated phosphorylation of PRMT5 downregulates its methyltransferase activity and promotes myeloproliferation. *Cancer Cell,* **19,** 283-294.

Liu, J., Lee, W., Jiang, Z., Chen, Z., Jhunjhunwala, S., Haverty, P. M., Gnad, F., Guan, Y., Gilbert, H. N., Stinson, J., Klijn, C., Guillory, J., Bhatt, D., Vartanian, S., Walter, K., Chan, J., Holcomb, T., Dijkgraaf, P., Johnson, S., Koeman, J., Minna, J. D., Gazdar, A. F., Stern, H. M., Hoeflich, K. P., Wu, T. D., Settleman, J., De Sauvage, F. J., Gentleman, R. C., Neve, R. M., Stokoe, D., Modrusan, Z., Seshagiri, S., Shames, D. S. & Zhang, Z. (2012) Genome and transcriptome sequencing of lung cancers reveal diverse mutational and splicing events. *Genome Research,* **22,** 2315-2327.

Liu, X. F. & Bagchi, M. K. (2004) Recruitment of distinct chromatin-modifying complexes by tamoxifen-complexed estrogen receptor at natural target gene promoters in vivo. *The Journal of Biological Chemistry,* **279,** 15050-15058.

Liu, Z., Li, F., Ruan, K., Zhang, J., Mei, Y., Wu, J. & Shi, Y. (2014) Structural and Functional Insights into the Human Borjeson-Forssman-Lehmann Syndrome-associated Protein PHF6. *The Journal of Biological Chemistry,* **289,** 10069-10083.

Lo Celso, C. & Scadden, D. T. (2011) The haematopoietic stem cell niche at a glance. *Journal of Cell Science,* **124,** 3529-3535.

Lopez, A. F., Sanderson, C. J., Gamble, J. R., Campbell, H. D., Young, I. G. & Vadas, M. A. (1988) Recombinant human interleukin 5 is a selective

activator of human eosinophil function. *The Journal of Experimental Medicine,* **167,** 219-224.

Losman, J. A., Looper, R. E., Koivunen, P., Lee, S., Schneider, R. K., Mcmahon, C., Cowley, G. S., Root, D. E., Ebert, B. L. & Kaelin, W. G., Jr. (2013) (R)-2-hydroxyglutarate is sufficient to promote leukemogenesis and its effects are reversible. *Science,* **339,** 1621-1625.

Lower, K. M., Turner, G., Kerr, B. A., Mathews, K. D., Shaw, M. A., Gedeon, A. K., Schelley, S., Hoyme, H. E., White, S. M., Delatycki, M. B., Lampe, A. K., Clayton-Smith, J., Stewart, H., Van Ravenswaay, C. M. A., De Vries, B. B. A., Cox, B., Grompe, M., Ross, S., Thomas, P., Mulley, J. C. & Gecz, J. (2002) Mutations in PHF6 are associated with Borjeson-Forssman-Lehmann syndrome. *Nature Genetics,* **32,** 661-665.

Lozzio, C. B. & Lozzio, B. B. (1975) Human chronic myelogenous leukemia cell-line with positive Philadelphia chromosome. *Blood,* **45,** 321-334.

Luger, K. & Richmond, T. J. (1998) The histone tails of the nucleosome. *Current Opinion in Genetics & Development,* **8,** 140-146.

Lundberg, P., Karow, A., Nienhold, R., Looser, R., Hao-Shen, H., Nissen, I., Girsberger, S., Lehmann, T., Passweg, J., Stern, M., Beisel, C., Kralovics, R. & Skoda, R. C. (2014) Clonal evolution and clinical correlates of somatic mutations in myeloproliferative neoplasms. *Blood*.

Ma, W., Kantarjian, H., Zhang, X., Wang, X., Zhang, Z., Yeh, C. H., O'brien, S., Giles, F., Bruey, J. M. & Albitar, M. (2010) JAK2 exon 14 deletion in patients with chronic myeloproliferative neoplasms. *PLOS One,* **5,** e12165.

Maffioli, M., Genoni, A., Caramazza, D., Mora, B., Bussini, A., Merli, M., Giorgino, T., Casalone, R. & Passamonti, F. (2014) Looking for CALR mutations in familial myeloproliferative neoplasms. *Leukemia*.

Makishima, H., Jankowska, A. M., Tiu, R. V., Szpurka, H., Sugimoto, Y., Hu, Z., Sauntharajah, Y., Guinta, K., Keddache, M. A., Putnam, P., Sekeres, M. A., Moliterno, A. R., List, A. F., Mcdevitt, M. A. & Maciejewski, J. P. (2010) Novel homo- and hemizygous mutations in EZH2 in myeloid malignancies. *Leukemia,* **24,** 1799-1804.

Manuyakorn, A., Paulus, R., Farrell, J., Dawson, N. A., Tze, S., Cheung-Lau, G., Hines, O. J., Reber, H., Seligson, D. B., Horvath, S., Kurdistani, S. K., Guha, C. & Dawson, D. W. (2010) Cellular histone modification patterns predict prognosis and treatment response in resectable pancreatic adenocarcinoma: results from RTOG 9704. *Journal of Clinical Oncology,* **28,** 1358-1365.

Marchioli, R., Finazzi, G., Specchia, G., Cacciola, R., Cavazzina, R., Cilloni, D., De Stefano, V., Elli, E., Iurlo, A., Latagliata, R., Lunghi, F., Lunghi, M., Marfisi, R. M., Musto, P., Masciulli, A., Musolino, C., Cascavilla, N., Quarta, G., Randi, M. L., Rapezzi, D., Ruggeri, M., Rumi, E., Scortechini, A. R., Santini, S., Scarano, M., Siragusa, S., Spadea, A., Tieghi, A., Angelucci, E., Visani, G., Vannucchi, A. M. & Barbui, T. (2013) Cardiovascular events and intensity of treatment in polycythemia vera. *The New England Journal of Medicine,* **368,** 22-33.

Mardis, E. R., Ding, L., Dooling, D. J., Larson, D. E., Mclellan, M. D., Chen, K., Koboldt, D. C., Fulton, R. S., Delehaunty, K. D., Mcgrath, S. D., Fulton, L. A., Locke, D. P., Magrini, V. J., Abbott, R. M., Vickery, T. L., Reed, J. S., Robinson, J. S., Wylie, T., Smith, S. M., Carmichael, L., Eldred, J. M., Harris, C. C., Walker, J., Peck, J. B., Du, F., Dukes, A. F., Sanderson, G. E., Brummett, A. M., Clark, E., Mcmichael, J. F., Meyer, R. J., Schindler, J. K., Pohl, C. S., Wallis, J. W., Shi, X., Lin, L., Schmidt, H., Tang, Y., Haipek, C., Wiechert, M. E., Ivy, J. V., Kalicki, J., Elliott, G., Ries, R. E., Payton, J. E.,

Westervelt, P., Tomasson, M. H., Watson, M. A., Baty, J., Heath, S., Shannon, W. D., Nagarajan, R., Link, D. C., Walter, M. J., Graubert, T. A., Dipersio, J. F., Wilson, R. K. & Ley, T. J. (2009) Recurring mutations found by sequencing an acute myeloid leukemia genome. *The New England Journal of Medicine,* **361,** 1058-1066.

Martin, P. & Papayannopoulou, T. (1982) HEL cells: a new human erythroleukemia cell line with spontaneous and induced globin expression. *Science,* **216,** 1233-1235.

Martinez-Moczygemba, M. & Huston, D. P. (2003) Biology of common beta receptor-signaling cytokines: IL-3, IL-5, and GM-CSF. *Journal of Allergy and Clinical Immunology,* **112,** 653-665; quiz 666.

Marzluff, W. F., Gongidi, P., Woods, K. R., Jin, J. & Maltais, L. J. (2002) The human and mouse replication-dependent histone genes. *Genomics,* **80,** 487-498.

Marzluff, W. F., Wagner, E. J. & Duronio, R. J. (2008) Metabolism and regulation of canonical histone mRNAs: life without a poly(A) tail. *Nature Reviews. Genetics,* **9,** 843-854.

Meeks-Wagner, D. & Hartwell, L. H. (1986) Normal stoichiometry of histone dimer sets is necessary for high fidelity of mitotic chromosome transmission. *Cell,* **44,** 43-52.

Merlat, A., Lai, J. L., Sterkers, Y., Demory, J. L., Bauters, F., Preudhomme, C. & Fenaux, P. (1999) Therapy-related myelodysplastic syndrome and acute myeloid leukemia with 17p deletion. A report on 25 cases. *Leukemia,* **13,** 250-257.

Mesa, R. A., Li, C. Y., Ketterling, R. P., Schroeder, G. S., Knudson, R. A. & Tefferi, A. (2005) Leukemic transformation in myelofibrosis with myeloid metaplasia: a single-institution experience with 91 cases. *Blood,* **105,** 973-977.

Mesa, R. A., Niblack, J., Wadleigh, M., Verstovsek, S., Camoriano, J., Barnes, S., Tan, A. D., Atherton, P. J., Sloan, J. A. & Tefferi, A. (2007a) The burden of fatigue and quality of life in myeloproliferative disorders (MPDs): an international Internet-based survey of 1179 MPD patients. *Cancer,* **109,** 68-76.

Mesa, R. A., Schwager, S., Radia, D., Cheville, A., Hussein, K., Niblack, J., Pardanani, A. D., Steensma, D. P., Litzow, M. R., Rivera, C. E., Camoriano, J., Verstovsek, S., Sloan, J., Harrison, C., Kantarjian, H. & Tefferi, A. (2009) The Myelofibrosis Symptom Assessment Form (MFSAF): an evidence-based brief inventory to measure quality of life and symptomatic response to treatment in myelofibrosis. *Leukemia Research,* **33,** 1199-1203.

Mesa, R. A., Silverstein, M. N., Jacobsen, S. J., Wollan, P. C. & Tefferi, A. (1999) Population-based incidence and survival figures in essential thrombocythemia and agnogenic myeloid metaplasia: an Olmsted County Study, 1976-1995. *American Journal of Hematology,* **61,** 10-15.

Mesa, R. A., Verstovsek, S., Cervantes, F., Barosi, G., Reilly, J. T., Dupriez, B., Levine, R., Le Bousse-Kerdiles, M. C., Wadleigh, M., Campbell, P. J., Silver, R. T., Vannucchi, A. M., Deeg, H. J., Gisslinger, H., Thomas, D., Odenike, O., Solberg, L. A., Gotlib, J., Hexner, E., Nimer, S. D., Kantarjian, H., Orazi, A., Vardiman, J. W., Thiele, J. & Tefferi, A. (2007b) Primary myelofibrosis (PMF), post polycythemia vera myelofibrosis (post-PV MF), post essential thrombocythemia myelofibrosis (post-ET MF), blast phase PMF (PMF-BP): Consensus on terminology by the international working group for

myelofibrosis research and treatment (IWG-MRT). *Leukemia Research,* **31,** 737-740.

Metcalf, D. & Nicola, N. A. (1991) Direct proliferative actions of stem cell factor on murine bone marrow cells in vitro: effects of combination with colony-stimulating factors. *Proceedings of the National Academy of Sciences USA* **88,** 6239-6243.

Mikkola, H. K., Klintman, J., Yang, H., Hock, H., Schlaeger, T. M., Fujiwara, Y. & Orkin, S. H. (2003) Haematopoietic stem cells retain long-term repopulating activity and multipotency in the absence of stem-cell leukaemia SCL/tal-1 gene. *Nature,* **421,** 547-551.

Minoshima, M., Chou, J. C., Lefebvre, S., Bando, T., Shinohara, K., Gottesfeld, J. M. & Sugiyama, H. (2010) Potent activity against K562 cells by polyamide-seco-CBI conjugates targeting histone H4 genes. *Bioorganic and Medicinal Chemistry,* **18,** 168-174.

Mitra, P., Xie, R. L., Medina, R., Hovhannisyan, H., Zaidi, S. K., Wei, Y., Harper, J. W., Stein, J. L., Van Wijnen, A. J. & Stein, G. S. (2003) Identification of HiNF-P, a key activator of cell cycle-controlled histone H4 genes at the onset of S phase. *Molecular and Cellular Biology,* **23,** 8110-8123.

Miyake, T., Kung, C. K. & Goldwasser, E. (1977) Purification of human erythropoietin. *The Journal of Biological Chemistry*, **252,** 5558-5564.

Molineux, G. (2011) Granulocyte colony-stimulating factors. *Cancer Treatment and Research,* **157,** 33-53.

Morin, R. D., Johnson, N. A., Severson, T. M., Mungall, A. J., An, J., Goya, R., Paul, J. E., Boyle, M., Woolcock, B. W., Kuchenbauer, F., Yap, D., Humphries, R. K., Griffith, O. L., Shah, S., Zhu, H., Kimbara, M., Shashkin, P., Charlot, J. F., Tcherpakov, M., Corbett, R., Tam, A., Varhol, R., Smailus, D., Moksa, M., Zhao, Y., Delaney, A., Qian, H., Birol, I., Schein, J., Moore, R., Holt, R., Horsman, D. E., Connors, J. M., Jones, S., Aparicio, S., Hirst, M., Gascoyne, R. D. & Marra, M. A. (2010) Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. *Nature Genetics*, **42,** 181-185.

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. & Ebert, B. L. (2010) Physiological Jak2V617F expression causes a lethal myeloproliferative neoplasm with differential effects on hematopoietic stem and progenitor cells. *Cancer Cell*, **17,** 584-596.

Mullally, A., Poveromo, L., Schneider, R. K., Al-Shahrour, F., Lane, S. W. & Ebert, B. L. (2012) Distinct roles for long-term hematopoietic stem cells and erythroid precursor cells in a murine model of Jak2V617F-mediated polycythemia vera. *Blood,* **120,** 166-172.

Musselman, C. A., Avvakumov, N., Watanabe, R., Abraham, C. G., Lalonde, M. E., Hong, Z., Allen, C., Roy, S., Nunez, J. K., Nickoloff, J., Kulesza, C. A., Yasui, A., Cote, J. & Kutateladze, T. G. (2012) Molecular basis for H3K36me3 recognition by the Tudor domain of PHF1. *Nature Structural and Molecular Biology,* **19,** 1266-1272.

Muto, T., Sashida, G., Oshima, M., Wendt, G. R., Mochizuki-Kashio, M., Nagata, Y., Sanada, M., Miyagi, S., Saraya, A., Kamio, A., Nagae, G., Nakaseko, C., Yokote, K., Shimoda, K., Koseki, H., Suzuki, Y., Sugano, S., Aburatani, H., Ogawa, S. & Iwama, A. (2013) Concurrent loss of Ezh2 and Tet2 cooperates in the pathogenesis of myelodysplastic disorders. *The Journal of Experimental Medicine,* **210,** 2627-2639.

Myers, M. P., Andersen, J. N., Cheng, A., Tremblay, M. L., Horvath, C. M., Parisien, J. P., Salmeen, A., Barford, D. & Tonks, N. K. (2001) TYK2 and JAK2 are substrates of protein-tyrosine phosphatase 1B. *The Journal of Biological Chemistry,* **276,** 47771-47774.

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., Harrison, C. N., Cross, N. C., Ron, D., Vannucchi, A. M., Papaemmanuil, E., Campbell, P. J. & Green, A. R. (2013) Somatic CALR mutations in myeloproliferative neoplasms with nonmutated JAK2. *The New England Journal of Medicine,* **369,** 2391-2405.

Neff, T., Sinha, A. U., Kluk, M. J., Zhu, N., Khattab, M. H., Stein, L., Xie, H., Orkin, S. H. & Armstrong, S. A. (2012) Polycomb repressive complex 2 is required for MLL-AF9 leukemia. *Proceedings of the National Academy of Sciences USA,* **109,** 5028-5033.

Nelson, D. M., Ye, X., Hall, C., Santos, H., Ma, T., Kao, G. D., Yen, T. J., Harper, J. W. & Adams, P. D. (2002) Coupling of DNA synthesis and histone synthesis in S phase independent of cyclin/cdk2 activity. *Molecular and Cellular Biology,* **22,** 7459-7472.

Network, C. G. A. (2012) Comprehensive molecular characterization of human colon and rectal cancer. *Nature,* **487,** 330-337.

Network, C. G. A. R. (2011) Integrated genomic analyses of ovarian carcinoma. *Nature,* **474,** 609-615.

Ng, A. P., Kauppi, M., Metcalf, D., Di Rago, L., Hyland, C. D. & Alexander, W. S. (2012) Characterization of thrombopoietin (TPO)-responsive progenitor cells in adult mouse bone marrow with in vivo megakaryocyte and erythroid potential. *Proceedings of the National Academy of Sciences USA* **109,** 2364-2369.

Nicola, N. A., Metcalf, D., Matsumoto, M. & Johnson, G. R. (1983) Purification of a factor inducing differentiation in murine myelomonocytic leukemia cells. Identification as granulocyte colony-stimulating factor. *The Journal of Biological Chemistry,* **258,** 9017-9023.

Nikoloski, G., Langemeijer, S. M., Kuiper, R. P., Knops, R., Massop, M., Tonnissen, E. R., Van Der Heijden, A., Scheele, T. N., Vandenberghe, P., De Witte, T., Van Der Reijden, B. A. & Jansen, J. H. (2010) Somatic mutations of the histone methyltransferase gene EZH2 in myelodysplastic syndromes. *Nature Genetics,* **42,** 665-667.

Ntziachristos, P., Tsirigos, A., Van Vlierberghe, P., Nedjic, J., Trimarchi, T., Flaherty, M. S., Ferres-Marco, D., Da Ros, V., Tang, Z., Siegle, J., Asp, P., Hadler, M., Rigo, I., De Keersmaecker, K., Patel, J., Huynh, T., Utro, F., Poglio, S., Samon, J. B., Paietta, E., Racevskis, J., Rowe, J. M., Rabadan, R., Levine, R. L., Brown, S., Pflumio, F., Dominguez, M., Ferrando, A. & Aifantis, I. (2012) Genetic inactivation of the polycomb repressive complex 2 in T cell acute lymphoblastic leukemia. *Nature Medicine,* **18,** 298-301.

Nussenzveig, R. H., Swierczek, S. I., Jelinek, J., Gaikwad, A., Liu, E., Verstovsek, S., Prchal, J. F. & Prchal, J. T. (2007) Polycythemia vera is not initiated by JAK2V617F mutation. *Experimental Hematology,* **35,** 32-38.

Nutt, S. L., Metcalf, D., D'amico, A., Polli, M. & Wu, L. (2005) Dynamic regulation of PU.1 expression in multipotent hematopoietic progenitors. *The Journal of Experimental Medicine,* **201,** 221-231.

O'shea, J. J., Gadina, M. & Schreiber, R. D. (2002) Cytokine signaling in 2002: new surprises in the Jak/Stat pathway. *Cell,* **109 Suppl,** S121-131.

Ohyashiki, J. H., Yoneta, M., Hisatomi, H., Iwabuchi, T., Umezu, T. & Ohyashiki, K. (2012) The C allele of JAK2 rs4495487 is an additional candidate locus that contributes to myeloproliferative neoplasm predisposition in the Japanese population. *BMC Medical Genetics,* **13,** 6.

Okano, M., Bell, D. W., Haber, D. A. & Li, E. (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell,* **99,** 247-257.

Okuda, K., Sanghera, J. S., Pelech, S. L., Kanakura, Y., Hallek, M., Griffin, J. D. & Druker, B. J. (1992) Granulocyte-macrophage colony-stimulating factor, interleukin-3, and steel factor induce rapid tyrosine phosphorylation of p42 and p44 MAP kinase. *Blood,* **79,** 2880-2887.

Olcaydu, D., Harutyunyan, A., Jager, R., Berg, T., Gisslinger, B., Pabinger, I., Gisslinger, H. & Kralovics, R. (2009) A common JAK2 haplotype confers susceptibility to myeloproliferative neoplasms. *Nature Genetics,* **41,** 450-454.

Palko, L., Bass, H. W., Beyrouthy, M. J. & Hurt, M. M. (2004) The Yin Yang-1 (YY1) protein undergoes a DNA-replication-associated switch in localization from the cytoplasm to the nucleus at the onset of S phase. *Journal of Cell Science,* **117,** 465-476.

Pang, W. W., Price, E. A., Sahoo, D., Beerman, I., Maloney, W. J., Rossi, D. J., Schrier, S. L. & Weissman, I. L. (2011) Human bone marrow hematopoietic stem cells are increased in frequency and myeloid-biased with age. *Proceedings of the National Academy of Sciences USA,* **108,** 20012-20017.

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. & Tefferi, A. (2006a) MPL515 mutations in myeloproliferative and other myeloid disorders: a study of 1182 patients. *Blood,* **108,** 3472-3476.

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

Pardanani, A., Lasho, T. L., Finke, C. M., Mai, M., Mcclure, R. F. & Tefferi, A. (2010) IDH1 and IDH2 mutation analysis in chronic- and blast-phase myeloproliferative neoplasms. *Leukemia,* **24,** 1146-1151.

Pardanani, A., Lasho, T. L., Finke, C., Mesa, R. A., Hogan, W. J., Ketterling, R. P., Gilliland, D. G. & Tefferi, A. (2007b) Extending Jak2V617F and MplW515 mutation analysis to single hematopoietic colonies and B and T lymphocytes. *Stem Cells,* **25,** 2358-2362.

Pardanani, A., Lasho, T., Mcclure, R., Lacy, M. & Tefferi, A. (2006b) Discordant distribution of JAK2V617F mutation in siblings with familial myeloproliferative disorders. *Blood,* **107,** 4572-4573.

Pardanani, A., Lasho, T., Smith, G., Burns, C. J., Fantino, E. & Tefferi, A. (2009) CYT387, a selective JAK1/JAK2 inhibitor: in vitro assessment of kinase selectivity and preclinical studies using cell lines and primary cells from polycythemia vera patients. *Leukemia,* **23,** 1441-1445.

Parganas, E., Wang, D., Stravopodis, D., Topham, D. J., Marine, J. C., Teglund, S., Vanin, E. F., Bodner, S., Colamonici, O. R., Van Deursen, J. M., Grosveld,

G. & Ihle, J. N. (1998) Jak2 is essential for signaling through a variety of cytokine receptors. *Cell,* **93,** 385-395.

Park, S. W., Chung, N. G., Han, J. Y., Eom, H. S., Lee, J. Y., Yoo, N. J. & Lee, S. H. (2009) Absence of IDH2 codon 172 mutation in common human cancers. *International Journal of Cancer,* **125,** 2485-2486.

Parmar, K., Mauch, P., Vergilio, J. A., Sackstein, R. & Down, J. D. (2007) Distribution of hematopoietic stem cells in the bone marrow according to regional hypoxia. *Proceedings of the National Academy of Sciences USA* **104,** 5431-5436.

Parsons, D. W., Jones, S., Zhang, X., Lin, J. C., Leary, R. J., Angenendt, P., Mankoo, P., Carter, H., Siu, I. M., Gallia, G. L., Olivi, A., Mclendon, R., Rasheed, B. A., Keir, S., Nikolskaya, T., Nikolsky, Y., Busam, D. A., Tekleab, H., Diaz, L. A., Jr., Hartigan, J., Smith, D. R., Strausberg, R. L., Marie, S. K., Shinjo, S. M., Yan, H., Riggins, G. J., Bigner, D. D., Karchin, R., Papadopoulos, N., Parmigiani, G., Vogelstein, B., Velculescu, V. E. & Kinzler, K. W. (2008) An integrated genomic analysis of human glioblastoma multiforme. *Science,* **321,** 1807-1812.

Passamonti, F. (2012) How I treat polycythemia vera. *Blood,* **120,** 275-284.

Passamonti, F., Cervantes, F., Vannucchi, A. M., Morra, E., Rumi, E., Pereira, A., Guglielmelli, P., Pungolino, E., Caramella, M., Maffioli, M., Pascutto, C., Lazzarino, M., Cazzola, M. & Tefferi, A. (2010a) A dynamic prognostic model to predict survival in primary myelofibrosis: a study by the IWG-MRT (International Working Group for Myeloproliferative Neoplasms Research and Treatment). *Blood,* **115,** 1703-1708.

Passamonti, F. & Rumi, E. (2009) Clinical relevance of JAK2 (V617F) mutant allele burden. *Haematologica,* **94,** 7-10.

Passamonti, F., Rumi, E., Arcaini, L., Boveri, E., Elena, C., Pietra, D., Boggi, S., Astori, C., Bernasconi, P., Varettoni, M., Brusamolino, E., Pascutto, C. & Lazzarino, M. (2008) Prognostic factors for thrombosis, myelofibrosis, and leukemia in essential thrombocythemia: a study of 605 patients. *Haematologica,* **93,** 1645-1651.

Passamonti, F., Rumi, E., Pietra, D., Elena, C., Boveri, E., Arcaini, L., Roncoroni, E., Astori, C., Merli, M., Boggi, S., Pascutto, C., Lazzarino, M. & Cazzola, M. (2010b) A prospective study of 338 patients with polycythemia vera: the impact of JAK2 (V617F) allele burden and leukocytosis on fibrotic or leukemic disease transformation and vascular complications. *Leukemia,* **24,** 1574-1579.

Passamonti, F., Rumi, E., Pungolino, E., Malabarba, L., Bertazzoni, P., Valentini, M., Orlandi, E., Arcaini, L., Brusamolino, E., Pascutto, C., Cazzola, M., Morra, E. & Lazzarino, M. (2004) Life expectancy and prognostic factors for survival in patients with polycythemia vera and essential thrombocythemia. *The American Journal of Medicine,* **117,** 755-761.

Patel, J. P., Gonen, M., Figueroa, M. E., Fernandez, H., Sun, Z., Racevskis, J., Van Vlierberghe, P., Dolgalev, I., Thomas, S., Aminova, O., Huberman, K., Cheng, J., Viale, A., Socci, N. D., Heguy, A., Cherry, A., Vance, G., Higgins, R. R., Ketterling, R. P., Gallagher, R. E., Litzow, M., Van Den Brink, M. R., Lazarus, H. M., Rowe, J. M., Luger, S., Ferrando, A., Paietta, E., Tallman, M. S., Melnick, A., Abdel-Wahab, O. & Levine, R. L. (2012) Prognostic relevance of integrated genetic profiling in acute myeloid leukemia. *The New England Journal of Medicine,* **366,** 1079-1089.

Patnaik, M. M., Caramazza, D., Gangat, N., Hanson, C. A., Pardanani, A. & Tefferi, A. (2010a) Age and platelet count are IPSS-independent prognostic factors in

young patients with primary myelofibrosis and complement IPSS in predicting very long or very short survival. *European Journal of Haematology,* **84,** 105-108.

Patnaik, M. M., Lasho, T. L., Finke, C. M., Gangat, N., Caramazza, D., Siragusa, S., Hanson, C. A., Pardanani, A. & Tefferi, A. (2010b) MPL mutation effect on JAK2 46/1 haplotype frequency in JAK2V617F-negative myeloproliferative neoplasms. *Leukemia,* **24,** 859-860.

Pecquet, C., Staerk, J., Chaligne, 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. & 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,** 1037-1048.

Peifer, M., Fernandez-Cuesta, L., Sos, M. L., George, J., Seidel, D., Kasper, L. H., Plenker, D., Leenders, F., Sun, R., Zander, T., Menon, R., Koker, M., Dahmen, I., Muller, C., Di Cerbo, V., Schildhaus, H. U., Altmuller, J., Baessmann, I., Becker, C., De Wilde, B., Vandesompele, J., Bohm, D., Ansen, S., Gabler, F., Wilkening, I., Heynck, S., Heuckmann, J. M., Lu, X., Carter, S. L., Cibulskis, K., Banerji, S., Getz, G., Park, K. S., Rauh, D., Grutter, C., Fischer, M., Pasqualucci, L., Wright, G., Wainer, Z., Russell, P., Petersen, I., Chen, Y., Stoelben, E., Ludwig, C., Schnabel, P., Hoffmann, H., Muley, T., Brockmann, M., Engel-Riedel, W., Muscarella, L. A., Fazio, V. M., Groen, H., Timens, W., Sietsma, H., Thunnissen, E., Smit, E., Heideman, D. A., Snijders, P. J., Cappuzzo, F., Ligorio, C., Damiani, S., Field, J., Solberg, S., Brustugun, O. T., Lund-Iversen, M., Sanger, J., Clement, J. H., Soltermann, A., Moch, H., Weder, W., Solomon, B., Soria, J. C., Validire, P., Besse, B., Brambilla, E., Brambilla, C., Lantuejoul, S., Lorimier, P., Schneider, P. M., Hallek, M., Pao, W., Meyerson, M., Sage, J., Shendure, J., Schneider, R., Buttner, R., Wolf, J., Nurnberg, P., Perner, S., Heukamp, L. C., Brindle, P. K., Haas, S. & Thomas, R. K. (2012) Integrative genome analyses identify key somatic driver mutations of small-cell lung cancer. *Nature Genetics,* **44,** 1104-1110.

Persons, D. A., Mehaffey, M. G., Kaleko, M., Nienhuis, A. W. & Vanin, E. F. (1998) An improved method for generating retroviral producer clones for vectors lacking a selectable marker gene. *Blood Cells, Molecules and Diseases,* **24,** 167-182.

Petruk, S., Black, K. L., Kovermann, S. K., Brock, H. W. & Mazo, A. (2013) Stepwise histone modifications are mediated by multiple enzymes that rapidly associate with nascent DNA during replication. *Nature Communications,* **4,** 2841.

Pietra, D., Casetti, I., Da Via, M. C., Elena, C., Milanesi, C. & Rumi, E. (2012) JAK2 GGCC haplotype in MPL mutated myeloproliferative neoplasms. *American Journal of Hematology,* **87,** 746-747.

Pietra, D., Li, S., Brisci, A., Passamonti, F., Rumi, E., Theocharides, A., Ferrari, M., Gisslinger, H., Kralovics, R., Cremonesi, L., Skoda, R. & Cazzola, M. (2008) Somatic mutations of JAK2 exon 12 in patients with JAK2 (V617F)-negative myeloproliferative disorders. *Blood,* **111,** 1686-1689.

Pietsch, T., Kyas, U., Steffens, U., Yakisan, E., Hadam, M. R., Ludwig, W. D., Zsebo, K. & Welte, K. (1992) Effects of human stem cell factor (c-kit ligand) on proliferation of myeloid leukemia cells: heterogeneity in response and synergy with other hematopoietic growth factors. *Blood,* **80,** 1199-1206.

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. & Levine, R. L. (2006) MPLW515L is a novel somatic activating mutation in myelofibrosis with myeloid metaplasia. *PLOS Medicine,* **3,** e270.

Podhorecka, M., Skladanowski, A. & Bozko, P. (2010) H2AX Phosphorylation: Its Role in DNA Damage Response and Cancer Therapy. *Journal of Nucleic Acids,* **2010**.

Prado, F. & Aguilera, A. (2005) Partial depletion of histone H4 increases homologous recombination-mediated genetic instability. *Molecular and Cellular Biology,* **25,** 1526-1536.

Quelle, F. W., Sato, N., Witthuhn, B. A., Inhorn, R. C., Eder, M., Miyajima, A., Griffin, J. D. & Ihle, J. N. (1994) JAK2 associates with the beta c chain of the receptor for granulocyte-macrophage colony-stimulating factor, and its activation requires the membrane-proximal region. *Molecular and Cellular Biology,* **14,** 4335-4341.

Rao, N., Butcher, C. M., Lewis, I. D., Ross, D. M., Melo, J. V., Scott, H. S., Bardy, P. G. & D'andrea, R. J. (2012) Clonal and lineage analysis of somatic DNMT3A and JAK2 mutations in a chronic phase polycythemia vera patient. *British Journal of Haematology,* **156,** 268-270.

Ribeiro, A. F., Pratcorona, M., Erpelinck-Verschueren, C., Rockova, V., Sanders, M., Abbas, S., Figueroa, M. E., Zeilemaker, A., Melnick, A., Lowenberg, B., Valk, P. J. & Delwel, R. Mutant DNMT3A: a new marker of poor prognosis in acute myeloid leukemia. *Blood*.

Richmond, T. D., Chohan, M. & Barber, D. L. (2005) Turning cells red: signal transduction mediated by erythropoietin. *Trends in Cell Biology,* **15,** 146-155.

Rinaldi, C. R., Rinaldi, P., Alagia, A., Gemei, M., Esposito, N., Formiggini, F., Martinelli, V., Senyuk, V., Nucifora, G. & Pane, F. (2010) Preferential nuclear accumulation of JAK2V617F in CD34+ but not in granulocytic, megakaryocytic, or erythroid cells of patients with Philadelphia-negative myeloproliferative neoplasia. *Blood,* **116,** 6023-6026.

Robinson, J. T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G. & Mesirov, J. P. (2011) Integrative genomics viewer. *Nature Biotechnology,* **29,** 24-26.

Rocquain, J., Carbuccia, N., Trouplin, V., Raynaud, S., Murati, A., Nezri, M., Tadrist, Z., Olschwang, S., Vey, N., Birnbaum, D., Gelsi-Boyer, V. & Mozziconacci, M. J. (2010) Combined mutations of ASXL1, CBL, FLT3, IDH1, IDH2, JAK2, KRAS, NPM1, NRAS, RUNX1, TET2 and WT1 genes in myelodysplastic syndromes and acute myeloid leukemias. *BMC Cancer,* **10,** 401.

Rodriguez-Paredes, M. & Esteller, M. (2011) Cancer epigenetics reaches mainstream oncology. *Nature Medicine,* **17,** 330-339.

Rogakou, E. P., Pilch, D. R., Orr, A. H., Ivanova, V. S. & Bonner, W. M. (1998) DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *The Journal of Biological Chemistry,* **273,** 5858-5868.

Rosenbauer, F. & Tenen, D. G. (2007) Transcription factors in myeloid development: balancing differentiation with transformation. *Nature Reviews Immunology,* **7,** 105-117.

Rosenbauer, F., Wagner, K., Kutok, J. L., Iwasaki, H., Le Beau, M. M., Okuno, Y., Akashi, K., Fiering, S. & Tenen, D. G. (2004) Acute myeloid leukemia induced by graded reduction of a lineage-specific transcription factor, PU.1. *Nature Genetics,* **36,** 624-630.

Ross, S. E., Radomska, H. S., Wu, B., Zhang, P., Winnay, J. N., Bajnok, L., Wright, W. S., Schaufele, F., Tenen, D. G. & Macdougald, O. A. (2004) Phosphorylation of C/EBPalpha inhibits granulopoiesis. *Molecular and Cellular Biology,* **24,** 675-686.

Rotunno, G., Mannarelli, C., Guglielmelli, P., Pacilli, A., Pancrazzi, A., Pieri, L., Fanelli, T., Bosi, A. & Vannucchi, A. (2013) Impact of Calreticulin Mutations on Clinical and Hematological Phenotype and Outcome in Essential Thrombocythemia. *Blood*.

Rowley, J. D. (1973) Letter: A new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining. *Nature,* **243,** 290-293.

Rozman, C., Giralt, M., Feliu, E., Rubio, D. & Cortes, M. T. (1991) Life expectancy of patients with chronic nonleukemic myeloproliferative disorders. *Cancer,* **67,** 2658-2663.

Rudin, C. M., Durinck, S., Stawiski, E. W., Poirier, J. T., Modrusan, Z., Shames, D. S., Bergbower, E. A., Guan, Y., Shin, J., Guillory, J., Rivers, C. S., Foo, C. K., Bhatt, D., Stinson, J., Gnad, F., Haverty, P. M., Gentleman, R., Chaudhuri, S., Janakiraman, V., Jaiswal, B. S., Parikh, C., Yuan, W., Zhang, Z., Koeppen, H., Wu, T. D., Stern, H. M., Yauch, R. L., Huffman, K. E., Paskulin, D. D., Illei, P. B., Varella-Garcia, M., Gazdar, A. F., De Sauvage, F. J., Bourgon, R., Minna, J. D., Brock, M. V. & Seshagiri, S. (2012) Comprehensive genomic analysis identifies SOX2 as a frequently amplified gene in small-cell lung cancer. *Nature Genetics,* **44,** 1111-1116.

Rumi, E., Harutyunyan, A. S., Pietra, D., Milosevic, J. D., Casetti, I. C., Bellini, M., Them, N. C., Cavalloni, C., Ferretti, V. V., Milanesi, C., Berg, T., Sant'antonio, E., Boveri, E., Pascutto, C., Astori, C., Kralovics, R. & Cazzola, M. (2014) CALR exon 9 mutations are somatically acquired events in familial cases of essential thrombocythemia or primary myelofibrosis. *Blood*.

Rumi, E., Passamonti, F., Della Porta, M. G., Elena, C., Arcaini, L., Vanelli, L., Del Curto, C., Pietra, D., Boveri, E., Pascutto, C., Cazzola, M. & Lazzarino, M. (2007) Familial chronic myeloproliferative disorders: clinical phenotype and evidence of disease anticipation. *Journal of Clinical Oncology,* **25,** 5630-5635.

Rumi, E., Passamonti, F., Pietra, D., Della Porta, M. G., Arcaini, L., Boggi, S., Elena, C., Boveri, E., Pascutto, C., Lazzarino, M. & Cazzola, M. (2006) JAK2 (V617F) as an acquired somatic mutation and a secondary genetic event associated with disease progression in familial myeloproliferative disorders. *Cancer,* **107,** 2206-2211.

Rumi, E., Pietra, D., Ferretti, V., Klampfl, T., Harutyunyan, A. S., Milosevic, J. D., Them, N. C., Berg, T., Elena, C., Casetti, I. C., Milanesi, C., Sant'antonio, E., Bellini, M., Fugazza, E., Renna, M. C., Boveri, E., Astori, C., Pascutto, C., Kralovics, R. & Cazzola, M. (2013) JAK2 or CALR mutation status defines subtypes of essential thrombocythemia with substantially different clinical course and outcomes. *Blood*.

Saharinen, P. & 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. *The Journal of Biological Chemistry,* **277,** 47954-47963.

Saharinen, P., Vihinen, M. & Silvennoinen, O. (2003) Autoinhibition of Jak2 tyrosine kinase is dependent on specific regions in its pseudokinase domain. *Molecular Biology of the Cell,* **14,** 1448-1459.

Salzler, H. R., Davidson, J. M., Montgomery, N. D. & Duronio, R. J. (2009) Loss of the histone pre-mRNA processing factor stem-loop binding protein in Drosophila causes genomic instability and impaired cellular proliferation. *PLOS One,* **4,** e8168.

Santos, F. P., Kantarjian, H. M., Jain, N., Manshouri, T., Thomas, D. A., Garcia-Manero, G., Kennedy, D., Estrov, Z., Cortes, J. & Verstovsek, S. (2010) Phase 2 study of CEP-701, an orally available JAK2 inhibitor, in patients with primary or post-polycythemia vera/essential thrombocythemia myelofibrosis. *Blood,* **115,** 1131-1136.

Santos, F. P. & Verstovsek, S. (2012) Therapy with JAK2 inhibitors for myeloproliferative neoplasms. *Hematology/Oncology Clinics of North America,* **26,** 1083-1099.

Sarma, K., Margueron, R., Ivanov, A., Pirrotta, V. & Reinberg, D. (2008) Ezh2 requires PHF1 to efficiently catalyze H3 lysine 27 trimethylation in vivo. *Molecular and Cellular Biology,* **28,** 2718-2731.

Sato, N., Sakamaki, K., Terada, N., Arai, K. & Miyajima, A. (1993) Signal transduction by the high-affinity GM-CSF receptor: two distinct cytoplasmic regions of the common beta subunit responsible for different signaling. *The EMBO Journal,* **12,** 4181-4189.

Sato, Y., Yoshizato, T., Shiraishi, Y., Maekawa, S., Okuno, Y., Kamura, T., Shimamura, T., Sato-Otsubo, A., Nagae, G., Suzuki, H., Nagata, Y., Yoshida, K., Kon, A., Suzuki, Y., Chiba, K., Tanaka, H., Niida, A., Fujimoto, A., Tsunoda, T., Morikawa, T., Maeda, D., Kume, H., Sugano, S., Fukayama, M., Aburatani, H., Sanada, M., Miyano, S., Homma, Y. & Ogawa, S. (2013) Integrated molecular analysis of clear-cell renal cell carcinoma. *Nature Genetics,* **45,** 860-867.

Schaub, F. X., Looser, R., Li, S., Hao-Shen, H., Lehmann, T., Tichelli, A. & Skoda, R. C. (2010) Clonal analysis of TET2 and JAK2 mutations suggests that TET2 can be a late event in the progression of myeloproliferative neoplasms. *Blood,* **115,** 2003-2007.

Scherber, R., Dueck, A. C., Johansson, P., Barbui, T., Barosi, G., Vannucchi, A. M., Passamonti, F., Andreasson, B., Ferarri, M. L., Rambaldi, A., Samuelsson, J., Birgegard, G., Tefferi, A., Harrison, C. N., Radia, D. & Mesa, R. A. (2011) The Myeloproliferative Neoplasm Symptom Assessment Form (MPN-SAF): international prospective validation and reliability trial in 402 patients. *Blood,* **118,** 401-408.

Schild-Poulter, C., Shih, A., Yarymowich, N. C. & Hache, R. J. (2003) Down-regulation of histone H2B by DNA-dependent protein kinase in response to DNA damage through modulation of octamer transcription factor 1. *Cancer Research,* **63,** 7197-7205.

Schnittger, S., Haferlach, C., Beelen, D.W., Bojko, P., Dengler, R., Diestelrath, A., Eckart, M.J., Eckert, R., Fries, S., Köchling, G., Laubenstein, H.P., Planker, M., Pihusch, R., Kern, W. & Haferlach, T. (2007) Detection of three different MPLW515 mutations in 10.1% of all JAK2 V617 unmutated ET and 9.3% of all JAK2 V617F unmutated OMF: a study of 387 patients. . *Blood (American Society of Hematology Annual Meeting Abstracts), Abstract#2546,* **110**

Schotta, G., Lachner, M., Sarma, K., Ebert, A., Sengupta, R., Reuter, G., Reinberg, D. & Jenuwein, T. (2004) A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes and Development,* **18,** 1251-1262.

Schwartzentruber, J., Korshunov, A., Liu, X. Y., Jones, D. T., Pfaff, E., Jacob, K., Sturm, D., Fontebasso, A. M., Quang, D. A., Tonjes, M., Hovestadt, V.,

Albrecht, S., Kool, M., Nantel, A., Konermann, C., Lindroth, A., Jager, N., Rausch, T., Ryzhova, M., Korbel, J. O., Hielscher, T., Hauser, P., Garami, M., Klekner, A., Bognar, L., Ebinger, M., Schuhmann, M. U., Scheurlen, W., Pekrun, A., Fruhwald, M. C., Roggendorf, W., Kramm, C., Durken, M., Atkinson, J., Lepage, P., Montpetit, A., Zakrzewska, M., Zakrzewski, K., Liberski, P. P., Dong, Z., Siegel, P., Kulozik, A. E., Zapatka, M., Guha, A., Malkin, D., Felsberg, J., Reifenberger, G., Von Deimling, A., Ichimura, K., Collins, V. P., Witt, H., Milde, T., Witt, O., Zhang, C., Castelo-Branco, P., Lichter, P., Faury, D., Tabori, U., Plass, C., Majewski, J., Pfister, S. M. & Jabado, N. (2012) Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. *Nature*, **482,** 226-231.

Scott, E. W., Simon, M. C., Anastasi, J. & Singh, H. (1994) Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages. *Science*, **265,** 1573-1577.

Scott, L. M. (2011) The JAK2 exon 12 mutations: a comprehensive review. *American Journal of Hematology*, **86,** 668-676.

Scott, L. M., Beer, P. A., Bench, A. J., Erber, W. N. & Green, A. R. (2007a) Prevalance of JAK2 V617F and exon 12 mutations in polycythaemia vera. *British Journal of Haematology,* **139,** 511-512.

Scott, L. M., Scott, M. A., Campbell, P. J. & Green, A. R. (2006) Progenitors homozygous for the V617F mutation occur in most patients with polycythemia vera, but not essential thrombocythemia. *Blood,* **108,** 2435-2437.

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. & Green, A. R. (2007b) JAK2 exon 12 mutations in polycythemia vera and idiopathic erythrocytosis. *The New England Journal of Medicine,* **356,** 459-468.

Seligson, D. B., Horvath, S., Mcbrian, M. A., Mah, V., Yu, H., Tze, S., Wang, Q., Chia, D., Goodglick, L. & Kurdistani, S. K. (2009) Global levels of histone modifications predict prognosis in different cancers. *The American Journal of Pathology,* **174,** 1619-1628.

Seligson, D. B., Horvath, S., Shi, T., Yu, H., Tze, S., Grunstein, M. & Kurdistani, S. K. (2005) Global histone modification patterns predict risk of prostate cancer recurrence. *Nature,* **435,** 1262-1266.

Shen, W., Przychodzen, B. P., Polprasert, C., Hosono, N., Dienes, B. , Miyano, S. , Shiraishi, Y., Yoshida, K., Ogawa, S., Sauntharajah, Y. , Maciejewski, J. P. & Makishima, H. (2013) Somatic PHF6 Mutations In Myeloid Malignancies. *Blood (American Society of Hematology Annual Meeting Abstracts), Abstract # 2514.*

Shi, S., Calhoun, H. C., Xia, F., Li, J., Le, L. & Li, W. X. (2006) JAK signaling globally counteracts heterochromatic gene silencing. *Nature Genetics,* **38,** 1071-1076.

Shi, S., Larson, K., Guo, D., Lim, S. J., Dutta, P., Yan, S. J. & Li, W. X. (2008) Drosophila STAT is required for directly maintaining HP1 localization and heterochromatin stability. *Nature Cell Biology,* **10,** 489-496.

Shide, K., Kameda, T., Shimoda, H., Yamaji, T., Abe, H., Kamiunten, A., Sekine, M., Hidaka, T., Katayose, K., Kubuki, Y., Yamamoto, S., Miike, T., Iwakiri, H., Hasuike, S., Nagata, K., Marutsuka, K., Iwama, A., Matsuda, T., Kitanaka, A. & Shimoda, K. (2012) TET2 is essential for survival and hematopoietic stem cell homeostasis. *Leukemia,* **26,** 2216-2223.

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. & Shimoda, K. (2008) Development of ET, primary myelofibrosis and PV in mice expressing JAK2 V617F. *Leukemia,* **22,** 87-95.

Shih, A. H., Abdel-Wahab, O., Patel, J. P. & Levine, R. L. (2012) The role of mutations in epigenetic regulators in myeloid malignancies. *Nature Reviews Cancer,* **12,** 599-612.

Shivarov, V., Gueorguieva, R., Stoimenov, A. & Tiu, R. (2013) DNMT3A mutation is a poor prognosis biomarker in AML: results of a meta-analysis of 4500 AML patients. *Leukemia Research,* **37,** 1445-1450.

Shlush, L. I., Zandi, S., Mitchell, A., Chen, W. C., Brandwein, J. M., Gupta, V., Kennedy, J. A., Schimmer, A. D., Schuh, A. C., Yee, K. W., Mcleod, J. L., Doedens, M., Medeiros, J. J., Marke, R., Kim, H. J., Lee, K., Mcpherson, J. D., Hudson, T. J., Brown, A. M., Trinh, Q. M., Stein, L. D., Minden, M. D., Wang, J. C. & Dick, J. E. (2014) Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia. *Nature,* **506,** 328-333.

Shuai, K., Ziemiecki, A., Wilks, A. F., Harpur, A. G., Sadowski, H. B., Gilman, M. Z. & Darnell, J. E. (1993) Polypeptide signalling to the nucleus through tyrosine phosphorylation of Jak and Stat proteins. *Nature,* **366,** 580-583.

Silvennoinen, O., Witthuhn, B. A., Quelle, F. W., Cleveland, J. L., Yi, T. & Ihle, J. N. (1993) Structure of the murine Jak2 protein-tyrosine kinase and its role in interleukin 3 signal transduction. *Proceedings of the National Academy of Sciences USA* **90,** 8429-8433.

Simon, C., Chagraoui, J., Krosl, J., Gendron, P., Wilhelm, B., Lemieux, S., Boucher, G., Chagnon, P., Drouin, S., Lambert, R., Rondeau, C., Bilodeau, A., Lavallee, S., Sauvageau, M., Hebert, J. & Sauvageau, G. (2012) A key role for EZH2 and associated genes in mouse and human adult T-cell acute leukemia. *Genes and Development,* **26,** 651-656.

Singh, R. K., Kabbaj, M. H., Paik, J. & Gunjan, A. (2009) Histone levels are regulated by phosphorylation and ubiquitylation-dependent proteolysis. *Nature Cell Biology,* **11,** 925-933.

Slezak, S., Jin, P., Caruccio, L., Ren, J., Bennett, M., Zia, N., Adams, S., Wang, E., Ascensao, J., Schechter, G. & Stroncek, D. (2009) Gene and microRNA analysis of neutrophils from patients with polycythemia vera and essential thrombocytosis: down-regulation of micro RNA-1 and -133a. *Journal of Translational Medicine,* **7,** 39.

Smith, C. A. & Fan, G. (2008) The saga of JAK2 mutations and translocations in hematologic disorders: pathogenesis, diagnostic and therapeutic prospects, and revised World Health Organization diagnostic criteria for myeloproliferative neoplasms. *Human Pathology,* **39,** 795-810.

Sonbol, M. B., Firwana, B., Zarzour, A., Morad, M., Rana, V. & Tiu, R. V. (2013) Comprehensive review of JAK inhibitors in myeloproliferative neoplasms. *Therapeutic Advances in Hematology,* **4,** 15-35.

Spector, M. S., Raff, A., Desilva, H., Lee, K. & Osley, M. A. (1997) Hir1p and Hir2p function as transcriptional corepressors to regulate histone gene transcription in the Saccharomyces cerevisiae cell cycle. *Molecular and Cellular Biology,* **17,** 545-552.

Spivak, J. L. (2002) Polycythemia vera: myths, mechanisms, and management. *Blood,* **100,** 4272-4290.

Spooner, C. J., Cheng, J. X., Pujadas, E., Laslo, P. & Singh, H. (2009) A recurrent network involving the transcription factors PU.1 and Gfi1 orchestrates innate and adaptive immune cell fates. *Immunity,* **31,** 576-586.

Steensma, D. P., Dewald, G. W., Lasho, T. L., Powell, H. L., Mcclure, R. F., Levine, R. L., Gilliland, D. G. & Tefferi, A. (2005) The JAK2 V617F activating tyrosine kinase mutation is an infrequent event in both "atypical" myeloproliferative disorders and myelodysplastic syndromes. *Blood,* **106,** 1207-1209.

Stegelmann, F., Bullinger, L., Schlenk, R. F., Paschka, P., Griesshammer, M., Blersch, C., Kuhn, S., Schauer, S., Dohner, H. & Dohner, K. (2011) DNMT3A mutations in myeloproliferative neoplasms. *Leukemia,* **25,** 1217-1219.

Stephens, P. J., Tarpey, P. S., Davies, H., Van Loo, P., Greenman, C., Wedge, D. C., Nik-Zainal, S., Martin, S., Varela, I., Bignell, G. R., Yates, L. R., Papaemmanuil, E., Beare, D., Butler, A., Cheverton, A., Gamble, J., Hinton, J., Jia, M., Jayakumar, A., Jones, D., Latimer, C., Lau, K. W., Mclaren, S., Mcbride, D. J., Menzies, A., Mudie, L., Raine, K., Rad, R., Chapman, M. S., Teague, J., Easton, D., Langerod, A., Lee, M. T., Shen, C. Y., Tee, B. T., Huimin, B. W., Broeks, A., Vargas, A. C., Turashvili, G., Martens, J., Fatima, A., Miron, P., Chin, S. F., Thomas, G., Boyault, S., Mariani, O., Lakhani, S. R., Van De Vijver, M., Van 'T Veer, L., Foekens, J., Desmedt, C., Sotiriou, C., Tutt, A., Caldas, C., Reis-Filho, J. S., Aparicio, S. A., Salomon, A. V., Borresen-Dale, A. L., Richardson, A. L., Campbell, P. J., Futreal, P. A. & Stratton, M. R. (2012) The landscape of cancer genes and mutational processes in breast cancer. *Nature,* **486,** 400-404.

Sterkers, Y., Preudhomme, C., Lai, J. L., Demory, J. L., Caulier, M. T., Wattel, E., Bordessoule, D., Bauters, F. & Fenaux, P. (1998) Acute myeloid leukemia and myelodysplastic syndromes following essential thrombocythemia treated with hydroxyurea: high proportion of cases with 17p deletion. *Blood,* **91,** 616-622.

Stijnis, C., Kroes, W. G., Balkassmi, S., Marijt, E. W., Van Rossum, A. P., Bakker, E. & Vlasveld, L. T. (2012) No evidence for JAK2(V617F) mutation in monoclonal B cells in 2 patients with polycythaemia vera and concurrent monoclonal B cell disorder. *Acta Haematologica,* **128,** 183-186.

Strahl, B. D., Briggs, S. D., Brame, C. J., Caldwell, J. A., Koh, S. S., Ma, H., Cook, R. G., Shabanowitz, J., Hunt, D. F., Stallcup, M. R. & Allis, C. D. (2001) Methylation of histone H4 at arginine 3 occurs in vivo and is mediated by the nuclear receptor coactivator PRMT1. *Current Biology,* **11,** 996-1000.

Stransky, N., Egloff, A. M., Tward, A. D., Kostic, A. D., Cibulskis, K., Sivachenko, A., Kryukov, G. V., Lawrence, M. S., Sougnez, C., Mckenna, A., Shefler, E., Ramos, A. H., Stojanov, P., Carter, S. L., Voet, D., Cortes, M. L., Auclair, D., Berger, M. F., Saksena, G., Guiducci, C., Onofrio, R. C., Parkin, M., Romkes, M., Weissfeld, J. L., Seethala, R. R., Wang, L., Rangel-Escareno, C., Fernandez-Lopez, J. C., Hidalgo-Miranda, A., Melendez-Zajgla, J., Winckler, W., Ardlie, K., Gabriel, S. B., Meyerson, M., Lander, E. S., Getz, G., Golub, T. R., Garraway, L. A. & Grandis, J. R. (2011) The mutational landscape of head and neck squamous cell carcinoma. *Science,* **333,** 1157-1160.

Streiff, M. B., Smith, B. & Spivak, J. L. (2002) The diagnosis and management of polycythemia vera in the era since the Polycythemia Vera Study Group: a survey of American Society of Hematology members' practice patterns. *Blood,* **99,** 1144-1149.

Sturm, D., Witt, H., Hovestadt, V., Khuong-Quang, D. A., Jones, D. T., Konermann, C., Pfaff, E., Tonjes, M., Sill, M., Bender, S., Kool, M., Zapatka, M., Becker, N., Zucknick, M., Hielscher, T., Liu, X. Y., Fontebasso, A. M., Ryzhova, M., Albrecht, S., Jacob, K., Wolter, M., Ebinger, M., Schuhmann, M. U., Van Meter, T., Fruhwald, M. C., Hauch, H., Pekrun, A., Radlwimmer, B., Niehues, T., Von Komorowski, G., Durken, M., Kulozik, A. E., Madden, J., Donson, A., Foreman, N. K., Drissi, R., Fouladi, M., Scheurlen, W., Von Deimling, A., Monoranu, C., Roggendorf, W., Herold-Mende, C., Unterberg, A., Kramm, C. M., Felsberg, J., Hartmann, C., Wiestler, B., Wick, W., Milde, T., Witt, O., Lindroth, A. M., Schwartzentruber, J., Faury, D., Fleming, A., Zakrzewska, M., Liberski, P. P., Zakrzewski, K., Hauser, P., Garami, M., Klekner, A., Bognar, L., Morrissy, S., Cavalli, F., Taylor, M. D., Van Sluis, P., Koster, J., Versteeg, R., Volckmann, R., Mikkelsen, T., Aldape, K., Reifenberger, G., Collins, V. P., Majewski, J., Korshunov, A., Lichter, P., Plass, C., Jabado, N. & Pfister, S. M. (2012) Hotspot mutations in H3F3A and IDH1 define distinct epigenetic and biological subgroups of glioblastoma. *Cancer Cell*, **22,** 425-437.

Suetake, I., Shinozaki, F., Miyagawa, J., Takeshima, H. & Tajima, S. (2004) DNMT3L stimulates the DNA methylation activity of Dnmt3a and Dnmt3b through a direct interaction. *The Journal of Biological Chemistry,* **279,** 27816-27823.

Sugiyama, T. & Nagasawa, T. (2012) Bone marrow niches for hematopoietic stem cells and immune cells. *Inflammation & Allergy-Drug Targets*, **11,** 201-206.

Sullivan, E., Santiago, C., Parker, E. D., Dominski, Z., Yang, X., Lanzotti, D. J., Ingledue, T. C., Marzluff, W. F. & Duronio, R. J. (2001) Drosophila stem loop binding protein coordinates accumulation of mature histone mRNA with cell cycle progression. *Genes and Development,* **15,** 173-187.

Szpurka, H., Tiu, R., Murugesan, G., Aboudola, S., Hsi, E. D., Theil, K. S., Sekeres, M. A. & Maciejewski, J. P. (2006) Refractory anemia with ringed sideroblasts associated with marked thrombocytosis (RARS-T), another myeloproliferative condition characterized by JAK2 V617F mutation. *Blood,* **108,** 2173-2181.

Tagami, H., Ray-Gallet, D., Almouzni, G. & Nakatani, Y. (2004) Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. *Cell,* **116,** 51-61.

Tahiliani, M., Koh, K. P., Shen, Y., Pastor, W. A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L. M., Liu, D. R., Aravind, L. & Rao, A. (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science,* **324,** 930-935.

Tanaka, S., Miyagi, S., Sashida, G., Chiba, T., Yuan, J., Mochizuki-Kashio, M., Suzuki, Y., Sugano, S., Nakaseko, C., Yokote, K., Koseki, H. & Iwama, A. (2012) Ezh2 augments leukemogenicity by reinforcing differentiation blockage in acute myeloid leukemia. *Blood,* **120,** 1107-1117.

Tatton-Brown, K., Hanks, S., Ruark, E., Zachariou, A., Duarte Sdel, V., Ramsay, E., Snape, K., Murray, A., Perdeaux, E. R., Seal, S., Loveday, C., Banka, S., Clericuzio, C., Flinter, F., Magee, A., Mcconnell, V., Patton, M., Raith, W., Rankin, J., Splitt, M., Strenger, V., Taylor, C., Wheeler, P., Temple, K. I., Cole, T., Douglas, J. & Rahman, N. (2011) Germline mutations in the oncogene EZH2 cause Weaver syndrome and increased human height. *Oncotarget,* **2,** 1127-1133.

Tefferi, A. (2005) Pathogenesis of myelofibrosis with myeloid metaplasia. *Journal of Clinical Oncology,* **23,** 8520-8530.

Tefferi, A. (2008) Essential thrombocythemia, polycythemia vera, and myelofibrosis: current management and the prospect of targeted therapy. *American Journal of Hematology,* **83,** 491-497.

Tefferi, A. (2010) Novel mutations and their functional and clinical relevance in myeloproliferative neoplasms: JAK2, MPL, TET2, ASXL1, CBL, IDH and IKZF1. *Leukemia,* **24,** 1128-1138.

Tefferi, A. (2013a) Polycythemia vera and essential thrombocythemia: 2013 update on diagnosis, risk-stratification, and management. *American Journal of Hematology,* **88,** 507-516.

Tefferi, A. (2013b) Primary myelofibrosis: 2013 update on diagnosis, risk-stratification, and management. *American Journal of Hematology,* **88,** 141-150.

Tefferi, A., Lasho, T. L., Abdel-Wahab, O., Guglielmelli, P., Patel, J., Caramazza, D., Pieri, L., Finke, C. M., Kilpivaara, O., Wadleigh, M., Mai, M., Mcclure, R. F., Gilliland, D. G., Levine, R. L., Pardanani, A. & Vannucchi, A. M. (2010a) IDH1 and IDH2 mutation studies in 1473 patients with chronic-, fibrotic- or blast-phase essential thrombocythemia, polycythemia vera or myelofibrosis. *Leukemia,* **24,** 1302-1309.

Tefferi, A., Lasho, T. L., Finke, C. M., Knudson, R. A., Ketterling, R., Hanson, C. H., Maffioli, M., Caramazza, D., Passamonti, F. & Pardanani, A. (2014a) CALR vs JAK2 vs MPL-mutated or triple-negative myelofibrosis: clinical, cytogenetic and molecular comparisons. *Leukemia,* **28,** 1472-1477.

Tefferi, A., Lasho, T. L., Schwager, S. M., Strand, J. S., Elliott, M., Mesa, R., Li, C. Y., Wadleigh, M., Lee, S. J. & Gilliland, D. G. (2006) The clinical phenotype of wild-type, heterozygous, and homozygous JAK2V617F in polycythemia vera. *Cancer,* **106,** 631-635.

Tefferi, A., Levine, R. L., Lim, K. H., Abdel-Wahab, O., Lasho, T. L., Patel, J., Finke, C. M., Mullally, A., Li, C. Y., Pardanani, A. & Gilliland, D. G. (2009a) Frequent TET2 mutations in systemic mastocytosis: clinical, KITD816V and FIP1L1-PDGFRA correlates. *Leukemia,* **23,** 900-904.

Tefferi, A., Lim, K. H., Abdel-Wahab, O., Lasho, T. L., Patel, J., Patnaik, M. M., Hanson, C. A., Pardanani, A., Gilliland, D. G. & Levine, R. L. (2009b) Detection of mutant TET2 in myeloid malignancies other than myeloproliferative neoplasms: CMML, MDS, MDS/MPN and AML. *Leukemia,* **23,** 1343-1345.

Tefferi, A. & Pardanani, A. (2014b) Genetics: CALR mutations and a new diagnostic algorithm for MPN. *Nature Reviews Clinical Oncology,* **11,** 125-126.

Tefferi, A., Pardanani, A., Lim, K. H., Abdel-Wahab, O., Lasho, T. L., Patel, J., Gangat, N., Finke, C. M., Schwager, S., Mullally, A., Li, C. Y., Hanson, C. A., Mesa, R., Bernard, O., Delhommeau, F., Vainchenker, W., Gilliland, D. G. & Levine, R. L. (2009c) TET2 mutations and their clinical correlates in polycythemia vera, essential thrombocythemia and myelofibrosis. *Leukemia,* **23,** 905-911.

Tefferi, A., Rumi, E., Finazzi, G., Gisslinger, H., Vannucchi, A. M., Rodeghiero, F., Randi, M. L., Vaidya, R., Cazzola, M., Rambaldi, A., Gisslinger, B., Pieri, L., Ruggeri, M., Bertozzi, I., Sulai, N. H., Casetti, I., Carobbio, A., Jeryczynski, G., Larson, D. R., Mullauer, L., Pardanani, A., Thiele, J., Passamonti, F. & Barbui, T. (2013) Survival and prognosis among 1545 patients with contemporary polycythemia vera: an international study. *Leukemia,* **27,** 1874-1881.

Tefferi, A., Siragusa, S., Hussein, K., Schwager, S. M., Hanson, C. A., Pardanani, A., Cervantes, F. & Passamonti, F. (2010b) Transfusion-dependency at

presentation and its acquisition in the first year of diagnosis are both equally detrimental for survival in primary myelofibrosis--prognostic relevance is independent of IPSS or karyotype. *American Journal of Hematology,* **85,** 14-17.

Tefferi, A., Thiele, J., Orazi, A., Kvasnicka, H. M., Barbui, T., Hanson, C. A., Barosi, G., Verstovsek, S., Birgegard, G., Mesa, R., Reilly, J. T., Gisslinger, H., Vannucchi, A. M., Cervantes, F., Finazzi, G., Hoffman, R., Gilliland, D. G., Bloomfield, C. D. & Vardiman, J. W. (2007) Proposals and rationale for revision of the World Health Organization diagnostic criteria for polycythemia vera, essential thrombocythemia, and primary myelofibrosis: recommendations from an ad hoc international expert panel. *Blood,* **110,** 1092-1097.

Tefferi, A., Thiele, J., Vannucchi, A. M. & Barbui, T. (2014c) An overview on CALR and CSF3R mutations and a proposal for revision of WHO diagnostic criteria for myeloproliferative neoplasms. *Leukemia,* **28,** 1407-1413.

Tefferi, A. & Vardiman, J. W. (2008) Classification and diagnosis of myeloproliferative neoplasms: the 2008 World Health Organization criteria and point-of-care diagnostic algorithms. *Leukemia,* **22,** 14-22.

Ten Hoeve, J., De Jesus Ibarra-Sanchez, M., Fu, Y., Zhu, W., Tremblay, M., David, M. & Shuai, K. (2002) Identification of a nuclear Stat1 protein tyrosine phosphatase. *Molecular and Cellular Biology,* **22,** 5662-5668.

Tenen, D. G., Hromas, R., Licht, J. D. & Zhang, D. E. (1997) Transcription factors, normal myeloid development, and leukemia. *Blood,* **90,** 489-519.

Them, N. C. & Kralovics, R. (2013) Genetic basis of MPN: Beyond JAK2-V617F. *Current Hematologic Malignancy Reports,* **8,** 299-306.

Theocharides, A., Boissinot, M., Girodon, F., Garand, R., Teo, S. S., Lippert, E., Talmant, P., Tichelli, A., Hermouet, S. & Skoda, R. C. (2007) Leukemic blasts in transformed JAK2-V617F-positive myeloproliferative disorders are frequently negative for the JAK2-V617F mutation. *Blood,* **110,** 375-379.

Thoreau, E., Petridou, B., Kelly, P. A., Djiane, J. & Mornon, J. P. (1991) Structural symmetry of the extracellular domain of the cytokine/growth hormone/prolactin receptor family and interferon receptors revealed by hydrophobic cluster analysis. *FEBS Letters,* **282,** 26-31.

Tibes, R. & Mesa, R. A. (2013) Emerging drugs for polycythemia vera. *Expert Opinion on Emerging Drugs,* **18,** 393-404.

Tiedt, R., Hao-Shen, H., Sobas, M. A., Looser, R., Dirnhofer, S., Schwaller, J. & Skoda, R. C. (2008) Ratio of mutant JAK2-V617F to wild-type Jak2 determines the MPD phenotypes in transgenic mice. *Blood,* **111,** 3931-3940.

Tilton, B., Andjelkovic, M., Didichenko, S. A., Hemmings, B. A. & Thelen, M. (1997) G-Protein-coupled receptors and Fcgamma-receptors mediate activation of Akt/protein kinase B in human phagocytes. *The Journal of Biological Chemistry,* **272,** 28096-28101.

Todd, M. A. & Picketts, D. J. (2012) PHF6 interacts with the nucleosome remodeling and deacetylation (NuRD) complex. *Journal of Proteome Research,* **11,** 4326-4337.

Tono, C., Xu, G., Toki, T., Takahashi, Y., Sasaki, S., Terui, K. & Ito, E. (2005) JAK2 Val617Phe activating tyrosine kinase mutation in juvenile myelomonocytic leukemia. *Leukemia,* **19,** 1843-1844.

Tsukamoto, N., Morita, K., Maehara, T., Okamoto, K., Sakai, H., Karasawa, M., Naruse, T. & Omine, M. (1994) Clonality in chronic myeloproliferative disorders defined by X-chromosome linked probes: demonstration of

heterogeneity in lineage involvement. *British Journal of Haematology,* **86,** 253-258.

Tzao, C., Tung, H. J., Jin, J. S., Sun, G. H., Hsu, H. S., Chen, B. H., Yu, C. P. & Lee, S. C. (2009) Prognostic significance of global histone modifications in resected squamous cell carcinoma of the esophagus. *Modern Pathology* **22,** 252-260.

Van Den Broeck, A., Brambilla, E., Moro-Sibilot, D., Lantuejoul, S., Brambilla, C., Eymin, B., Khochbin, S. & Gazzeri, S. (2008) Loss of histone H4K20 trimethylation occurs in preneoplasia and influences prognosis of non-small cell lung cancer. *Clinical Cancer Research* **14,** 7237-7245.

Van Vlierberghe, P., Palomero, T., Khiabanian, H., Van Der Meulen, J., Castillo, M., Van Roy, N., De Moerloose, B., Philippe, J., Gonzalez-Garcia, S., Toribio, M. L., Taghon, T., Zuurbier, L., Cauwelier, B., Harrison, C. J., Schwab, C., Pisecker, M., Strehl, S., Langerak, A. W., Gecz, J., Sonneveld, E., Pieters, R., Paietta, E., Rowe, J. M., Wiernik, P. H., Benoit, Y., Soulier, J., Poppe, B., Yao, X., Cordon-Cardo, C., Meijerink, J., Rabadan, R., Speleman, F. & Ferrando, A. (2010) PHF6 mutations in T-cell acute lymphoblastic leukemia. *Nature Genetics,* **42,** 338-342.

Van Vlierberghe, P., Patel, J., Abdel-Wahab, O., Lobry, C., Hedvat, C. V., Balbin, M., Nicolas, C., Payer, A. R., Fernandez, H. F., Tallman, M. S., Paietta, E., Melnick, A., Vandenberghe, P., Speleman, F., Aifantis, I., Cools, J., Levine, R. & Ferrando, A. (2011) PHF6 mutations in adult acute myeloid leukemia. *Leukemia,* **25,** 130-134.

Van Wijnen, A. J., Ramsey-Ewing, A. L., Bortell, R., Owen, T. A., Lian, J. B., Stein, J. L. & Stein, G. S. (1991) Transcriptional element H4-site II of cell cycle regulated human H4 histone genes is a multipartite protein/DNA interaction site for factors HiNF-D, HiNF-M, and HiNF-P: involvement of phosphorylation. *Journal of Cellular Biochemistry,* **46,** 174-189.

Vannucchi, A. M., Antonioli, E., Guglielmelli, P., Longo, G., Pancrazzi, A., Ponziani, V., Bogani, C., Ferrini, P. R., Rambaldi, A., Guerini, V., Bosi, A. & Barbui, T. (2007) Prospective identification of high-risk polycythemia vera patients based on JAK2(V617F) allele burden. *Leukemia,* **21,** 1952-1959.

Vannucchi, A. M., Antonioli, E., Guglielmelli, P., Pancrazzi, A., Guerini, V., Barosi, G., Ruggeri, M., Specchia, G., Lo-Coco, F., Delaini, F., Villani, L., Finotto, S., Ammatuna, E., Alterini, R., Carrai, V., Capaccioli, G., Di Lollo, S., Liso, V., Rambaldi, A., Bosi, A. & Barbui, T. (2008a) Characteristics and clinical correlates of MPL 515W>L/K mutation in essential thrombocythemia. *Blood,* **112,** 844-847.

Vannucchi, A. M., Antonioli, E., Guglielmelli, P., Pardanani, A. & Tefferi, A. (2008b) Clinical correlates of JAK2V617F presence or allele burden in myeloproliferative neoplasms: a critical reappraisal. *Leukemia,* **22,** 1299-1307.

Vannucchi, A. M., Rotunno, G., Bartalucci, N., Raugei, G., Carrai, V., Balliu, M., Mannarelli, C., Pacilli, A., Calabresi, L., Fjerza, R., Pieri, L., Bosi, A., Manfredini, R. & Guglielmelli, P. (2014) Calreticulin mutation-specific immunostaining in myeloproliferative neoplasms: pathogenetic insight and diagnostic value. *Leukemia,* **28,** 1811-1818.

Vardiman, J. W., Thiele, J., Arber, D. A., Brunning, R. D., Borowitz, M. J., Porwit, A., Harris, N. L., Le Beau, M. M., Hellstrom-Lindberg, E., Tefferi, A. & Bloomfield, C. D. (2009) The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. *Blood,* **114,** 937-951.

Verstovsek, S., Kantarjian, H., Mesa, R. A., Pardanani, A. D., Cortes-Franco, J., Thomas, D. A., Estrov, Z., Fridman, J. S., Bradley, E. C., Erickson-Viitanen, S., Vaddi, K., Levy, R. & Tefferi, A. (2010) Safety and efficacy of INCB018424, a JAK1 and JAK2 inhibitor, in myelofibrosis. *The New England Journal of Medicine,* **363,** 1117-1127.

Verstovsek, S., Passamonti, F., Rambaldi, A., Barosi, G., Rosen, P.J., He, S., Contel, N., Mookerjee, B., Rumi, E., Gattoni, E., Pieri, L., Cazzola, M., Kantarjian, H.M., Barbui, T. & Vannucchi, A.M. (2012) Long-term efficacy and safety results from a phase II study of ruxolitinib in patients with polycythemia vera. *Blood (American Society of Hematology Annual Meeting Abstracts), Abstract # 804*, **120**.

Voss, A. K., Gamble, R., Collin, C., Shoubridge, C., Corbett, M., Gecz, J. & Thomas, T. (2007) Protein and gene expression analysis of Phf6, the gene mutated in the Borjeson-Forssman-Lehmann Syndrome of intellectual disability and obesity. *Gene Expression Patterns,* **7,** 858-871.

Wagner, E. J. & Marzluff, W. F. (2006) ZFP100, a component of the active U7 snRNP limiting for histone pre-mRNA processing, is required for entry into S phase. *Molecular Cell Biology,* **26,** 6702-6712.

Waiboci, L. W., Ahmed, C. M., Mujtaba, M. G., Flowers, L. O., Martin, J. P., Haider, M. I. & Johnson, H. M. (2007) Both the suppressor of cytokine signaling 1 (SOCS-1) kinase inhibitory region and SOCS-1 mimetic bind to JAK2 autophosphorylation site: implications for the development of a SOCS-1 antagonist. *Journal of Immunology,* **178,** 5058-5068.

Walker, B. A., Wardell, C. P., Melchor, L., Hulkki, S., Potter, N. E., Johnson, D. C., Fenwick, K., Kozarewa, I., Gonzalez, D., Lord, C. J., Ashworth, A., Davies, F. E. & Morgan, G. J. (2012) Intraclonal heterogeneity and distinct molecular mechanisms characterize the development of t(4;14) and t(11;14) myeloma. *Blood,* **120,** 1077-1086.

Walter, M. J., Ding, L., Shen, D., Shao, J., Grillot, M., Mclellan, M., Fulton, R., Schmidt, H., Kalicki-Veizer, J., O'laughlin, M., Kandoth, C., Baty, J., Westervelt, P., Dipersio, J. F., Mardis, E. R., Wilson, R. K., Ley, T. J. & Graubert, T. A. (2011) Recurrent DNMT3A mutations in patients with myelodysplastic syndromes. *Leukemia,* **25,** 1153-1158.

Wang, H., Huang, Z. Q., Xia, L., Feng, Q., Erdjument-Bromage, H., Strahl, B. D., Briggs, S. D., Allis, C. D., Wong, J., Tempst, P. & Zhang, Y. (2001) Methylation of histone H4 at arginine 3 facilitating transcriptional activation by nuclear hormone receptor. *Science,* **293,** 853-857.

Wang, J., Leung, J. W., Gong, Z., Feng, L., Shi, X. & Chen, J. (2013) PHF6 Regulates Cell Cycle Progression by Suppressing Ribosomal RNA Synthesis. *The Journal of Biological Chemistry,* **288,** 3174-3183.

Wang, L., Swierczek, S., Piterkova, L., Hickman, K., Wheeler, D. A. & Prchal, J. T. (2012) Whole Exome Sequencing of Polycythemia Vera Reveals Novel Recurrent Somatic and Germline Variation. *Blood (American Society of Hematology Annual Meeting Abstracts), Abstract#705*.

Wang, Y., Wysocka, J., Sayegh, J., Lee, Y. H., Perlin, J. R., Leonelli, L., Sonbuchner, L. S., Mcdonald, C. H., Cook, R. G., Dou, Y., Roeder, R. G., Clarke, S., Stallcup, M. R., Allis, C. D. & Coonrod, S. A. (2004) Human PAD4 regulates histone arginine methylation levels via demethylimination. *Science,* **306,** 279-283.

Warshawsky, I., Mularo, F., Hren, C. & Jakubowski, M. (2010) Failure of the Ipsogen MutaScreen kit to detect the JAK2 617V>F mutation in samples with

additional rare exon 14 mutations: implications for clinical testing and report of a novel 618C>F mutation in addition to 617V>F. *Blood,* **115,** 3175-3176.

Watanabe, S., Itoh, T. & Arai, K. (1996) JAK2 is essential for activation of c-fos and c-myc promoters and cell proliferation through the human granulocyte-macrophage colony-stimulating factor receptor in BA/F3 cells. *The Journal of Biological Chemistry,* **271,** 12681-12686.

Wei, Y., Xia, W., Zhang, Z., Liu, J., Wang, H., Adsay, N. V., Albarracin, C., Yu, D., Abbruzzese, J. L., Mills, G. B., Bast, R. C., Jr., Hortobagyi, G. N. & Hung, M. C. (2008) Loss of trimethylation at lysine 27 of histone H3 is a predictor of poor outcome in breast, ovarian, and pancreatic cancers. *Molecular Carcinogenesis,* **47,** 701-706.

Weissman, I. L., Anderson, D. J. & Gage, F. (2001) Stem and progenitor cells: origins, phenotypes, lineage commitments, and transdifferentiations. *Annual Review of Cell and Developmental Biology,* **17,** 387-403.

Welch, J. S., Ley, T. J., Link, D. C., Miller, C. A., Larson, D. E., Koboldt, D. C., Wartman, L. D., Lamprecht, T. L., Liu, F., Xia, J., Kandoth, C., Fulton, R. S., Mclellan, M. D., Dooling, D. J., Wallis, J. W., Chen, K., Harris, C. C., Schmidt, H. K., Kalicki-Veizer, J. M., Lu, C., Zhang, Q., Lin, L., O'laughlin, M. D., Mcmichael, J. F., Delehaunty, K. D., Fulton, L. A., Magrini, V. J., Mcgrath, S. D., Demeter, R. T., Vickery, T. L., Hundal, J., Cook, L. L., Swift, G. W., Reed, J. P., Alldredge, P. A., Wylie, T. N., Walker, J. R., Watson, M. A., Heath, S. E., Shannon, W. D., Varghese, N., Nagarajan, R., Payton, J. E., Baty, J. D., Kulkarni, S., Klco, J. M., Tomasson, M. H., Westervelt, P., Walter, M. J., Graubert, T. A., Dipersio, J. F., Ding, L., Mardis, E. R. & Wilson, R. K. (2012) The origin and evolution of mutations in acute myeloid leukemia. *Cell,* **150,** 264-278.

Wen, L. (2001) Two-step cycle sequencing improves base ambiguities and signal dropouts in DNA sequencing reactions using energy-transfer-based fluorescent dye terminators. *Molecular Biotechnology,* **17,** 135-142.

Wernig, G., Kharas, M. G., Okabe, R., Moore, S. A., Leeman, D. S., Cullen, D. E., Gozo, M., Mcdowell, E. P., Levine, R. L., Doukas, J., Mak, C. C., Noronha, G., Martin, M., Ko, Y. D., Lee, B. H., Soll, R. M., Tefferi, A., Hood, J. D. & Gilliland, D. G. (2008) Efficacy of TG101348, a selective JAK2 inhibitor, in treatment of a murine model of JAK2V617F-induced polycythemia vera. *Cancer Cell,* **13,** 311-320.

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

Wolanskyj, A. P., Schwager, S. M., Mcclure, R. F., Larson, D. R. & Tefferi, A. (2006) Essential thrombocythemia beyond the first decade: life expectancy, long-term complication rates, and prognostic factors. *Mayo Clinic Proceedings,* **81,** 159-166.

Wu, G., Broniscer, A., Mceachron, T. A., Lu, C., Paugh, B. S., Becksfort, J., Qu, C., Ding, L., Huether, R., Parker, M., Zhang, J., Gajjar, A., Dyer, M. A., Mullighan, C. G., Gilbertson, R. J., Mardis, E. R., Wilson, R. K., Downing, J. R., Ellison, D. W., Zhang, J. & Baker, S. J. (2012) Somatic histone H3 alterations in pediatric diffuse intrinsic pontine gliomas and non-brainstem glioblastomas. *Nature Genetics,* **44,** 251-253.

Xie, R., Medina, R., Zhang, Y., Hussain, S., Colby, J., Ghule, P., Sundararajan, S., Keeler, M., Liu, L. J., Van Der Deen, M., Mitra, P., Lian, J. B., Rivera-Perez, J. A., Jones, S. N., Stein, J. L., Van Wijnen, A. J. & Stein, G. S. (2009) The

histone gene activator HINFP is a nonredundant cyclin E/CDK2 effector during early embryonic cell cycles. *Proceedings of the National Academy of Sciences USA,* **106,** 12359-12364.

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

Yamashita, Y., Yuan, J., Suetake, I., Suzuki, H., Ishikawa, Y., Choi, Y. L., Ueno, T., Soda, M., Hamada, T., Haruta, H., Takada, S., Miyazaki, Y., Kiyoi, H., Ito, E., Naoe, T., Tomonaga, M., Toyota, M., Tajima, S., Iwama, A. & Mano, H. (2010) Array-based genomic resequencing of human leukemia. *Oncogene,* **29,** 3723-3731.

Yan, H., Parsons, D. W., Jin, G., Mclendon, R., Rasheed, B. A., Yuan, W., Kos, I., Batinic-Haberle, I., Jones, S., Riggins, G. J., Friedman, H., Friedman, A., Reardon, D., Herndon, J., Kinzler, K. W., Velculescu, V. E., Vogelstein, B. & Bigner, D. D. (2009) IDH1 and IDH2 mutations in gliomas. *The New England Journal of Medicine,* **360,** 765-773.

Yan, X. J., Xu, J., Gu, Z. H., Pan, C. M., Lu, G., Shen, Y., Shi, J. Y., Zhu, Y. M., Tang, L., Zhang, X. W., Liang, W. X., Mi, J. Q., Song, H. D., Li, K. Q., Chen, Z. & Chen, S. J. (2011) Exome sequencing identifies somatic mutations of DNA methyltransferase gene DNMT3A in acute monocytic leukemia. *Nature Genetics,* **43,** 309-315.

Ye, X., Franco, A. A., Santos, H., Nelson, D. M., Kaufman, P. D. & Adams, P. D. (2003) Defective S phase chromatin assembly causes DNA damage, activation of the S phase checkpoint, and S phase arrest. *Molecular Cell,* **11,** 341-351.

Yoshihara, H., Arai, F., Hosokawa, K., Hagiwara, T., Takubo, K., Nakamura, Y., Gomei, Y., Iwasaki, H., Matsuoka, S., Miyamoto, K., Miyazaki, H., Takahashi, T. & Suda, T. (2007) Thrombopoietin/MPL signaling regulates hematopoietic stem cell quiescence and interaction with the osteoblastic niche. *Cell Stem Cell,* **1,** 685-697.

Zaleskas, V. M., Krause, D. S., Lazarides, K., Patel, N., Hu, Y., Li, S. & Van Etten, R. A. (2006) Molecular pathogenesis and therapy of polycythemia induced in mice by JAK2 V617F. *PLOS One,* **1,** e18.

Zeng, H., Yucel, R., Kosan, C., Klein-Hitpass, L. & Moroy, T. (2004) Transcription factor Gfi1 regulates self-renewal and engraftment of hematopoietic stem cells. *The EMBO Journal,* **23,** 4116-4125.

Zentner, G. E. & Henikoff, S. (2013) Regulation of nucleosome dynamics by histone modifications. *Nature Structural and Molecular Biology,* **20,** 259-266.

Zhang, D. E., Zhang, P., Wang, N. D., Hetherington, C. J., Darlington, G. J. & Tenen, D. G. (1997) Absence of granulocyte colony-stimulating factor signaling and neutrophil development in CCAAT enhancer binding protein alpha-deficient mice. *Proceedings of the National Academy of Sciences USA,* **94,** 569-574.

Zhang, S. J., Rampal, R., Manshouri, T., Patel, J., Mensah, N., Kayserian, A., Hricik, T., Heguy, A., Hedvat, C., Gonen, M., Kantarjian, H., Levine, R. L., Abdel-Wahab, O. & Verstovsek, S. (2012) Genetic analysis of patients with leukemic transformation of myeloproliferative neoplasms shows recurrent SRSF2 mutations that are associated with adverse outcome. *Blood,* **119,** 4480-4485.

Zhang, Y., Jurkowska, R., Soeroes, S., Rajavelu, A., Dhayalan, A., Bock, I., Rathert, P., Brandt, O., Reinhardt, R., Fischle, W. & Jeltsch, A. (2010) Chromatin methylation activity of Dnmt3a and Dnmt3a/3L is guided by interaction of

the ADD domain with the histone H3 tail. *Nucleic Acids Research,* **38,** 4246-4253.

Zhao, Q., Rank, G., Tan, Y. T., Li, H., Moritz, R. L., Simpson, R. J., Cerruti, L., Curtis, D. J., Patel, D. J., Allis, C. D., Cunningham, J. M. & Jane, S. M. (2009a) PRMT5-mediated methylation of histone H4R3 recruits DNMT3A, coupling histone and DNA methylation in gene silencing. *Nature Structural and Molecular Biology,* **16,** 304-311.

Zhao, R., Xing, S., Li, Z., Fu, X., Li, Q., Krantz, S. B. & Zhao, Z. J. (2005) Identification of an acquired JAK2 mutation in polycythemia vera. *The Journal of Biological Chemistry,* **280,** 22788-22792.

Zhao, S., Lin, Y., Xu, W., Jiang, W., Zha, Z., Wang, P., Yu, W., Li, Z., Gong, L., Peng, Y., Ding, J., Lei, Q., Guan, K. L. & Xiong, Y. (2009b) Glioma-derived mutations in IDH1 dominantly inhibit IDH1 catalytic activity and induce HIF-1alpha. *Science,* **324,** 261-265.

Zhao, W., Du, Y., Ho, W. T., Fu, X. & Zhao, Z. J. (2012) JAK2V617F and p53 mutations coexist in erythroleukemia and megakaryoblastic leukemic cell lines. *Experimental Hematology and Oncology,* **1,** 15.

Zheng, J., Huynh, H., Umikawa, M., Silvany, R. & Zhang, C. C. (2011) Angiopoietin-like protein 3 supports the activity of hematopoietic stem cells in the bone marrow niche. *Blood,* **117,** 470-479.

Zola, Heddy (2007) *Leukocyte and stromal cell molecules : the CD markers.* Wiley-Liss, Hoboken, N.J.