DEVELOPING NOVEL THERAPEUTIC STRATEGIES FOR THE TREATMENT OF *TET2*-MUTATED ACUTE MYELOID LEUKAEMIA

A thesis submitted in part requirement for the
degree of Doctor of Philosophy

Devi Nandana Suchitra Devi

Translational and Clinical Research Institute,
Centre for Cancer, Faculty of Medical Sciences
Newcastle University

January 2022

ABSTRACT

Acute myeloid leukaemia (AML) is a clinically heterogeneous disease driven by somatic mutations and chromosomal abnormalities. *TET2* mutations occur in 7%-23% of AML, with approximately 75% monoallelic and 25% biallelic mutations. *TET2* encodes a protein involved in regulating DNA demethylation and *TET2*-mutated AML has a significantly lower event-free survival and higher relapse rate compared to disease with wild-type *TET2*. As such, there is an urgent clinical need to develop novel methods to treat patients with *TET2*-mutated AML.

Hypomethylating agents such as azacytidine that inhibits DNA methylation are now established as a therapeutic strategy for AML. This study evolved from an index patient with biallelic *TET2*-mutated AML, that was chemoresistant to anthracycline and cytarabine, but acutely sensitive to azacytidine monotherapy resulting in long-term morphological remission with overall survival of 850 days. The primary focus of my research is to interrogate the role of *TET2* mutant allele dosage in determining AML cell response to azacytidine. Using an isogenic cell model system, we demonstrated that biallelic *TET2* mutation leads to complete loss of TET2 protein and acquisition of a hypermethylated phenotype. Of note, *TET2* mutant allele dosage significantly affected the azacytidine sensitivity of cells. *TET2* biallelic mutants were hypersensitive to azacytidine with significantly lower proliferation in liquid media ($P<0.0001$) and cloning efficiency ($P=0.0056$) compared to their isogenic *TET2* monoallelic mutants.

RNA-sequencing identified ABCB1, a major azacytidine efflux transporter, to be downregulated in cells with *TET2* biallelic mutation. Inhibition of ABCB1 using selective inhibitors (Tariquidar and Verapamil) increased sensitivity to azacytidine. Also, azacytidine resistant AML cell clones were developed through long-term exposure of *TET2* mutated cells to azacytidine. ABCB1 is upregulated in resistant clones compared to the azacytidine sensitive cells and inhibition of ABCB1 re-sensitised the resistant clones to azacytidine. In addition, gene ontology analysis identified ribosomal pathway components to be downregulated in *TET2* biallelic mutated cells compared to their isogenic *TET2* monoallelic mutated counterparts. As such, the findings of this investigation suggest the potential involvement of multiple genes and pathways in sensitizing TET2 null cells to azacytidine. Finally, using ten different AML cell lines, we demonstrated that there was significant correlation between TET2 protein expression and azacytidine IC90 ($R^2=0.77$, $P=0.0008$) and IC50 ($R^2=0.88$, $P<0.0001$). Together, our findings argue in favour of using azacytidine for the treatment of *TET2* biallelic mutated AML and highlights the importance of *TET2* mutant allele dosage in developing precision medicine in AML.

Dedication

I dedicate this thesis to my parents, for teaching me to chase my dreams beyond limits and for always supporting me with their unconditional love and faith.

Acknowledgement

Firstly, I would like to extend my heartfelt gratitude to my supervisor Prof. James Allan for his invaluable guidance, support, and ideas throughout this project. His timely advice, positive outlook, and faith in my ability, gave me confidence and nurtured my development as an independent researcher. I would also like to thank my co-supervisor Dr. Gordon Strathdee for his expert guidance, help and support during this project. Many thanks also to the present and past members of the Molecular Carcinogenesis group – Dr. Claire Elstob, Dr. Sarah Fordham, Nurettin Ayvali, Dr. Mohammed Fadly, Abrar Alharbi, Dr. Rachel Piddock, Dr. Wei-Yu Lin and Dr. Catherine Park. Our lab group was more like a family to me. Thanks for all the shared moments and especially for those teatime conversations, which was a great relief during stressful times. My sincere gratitude to my panel members Prof. Craig Robson and Dr. Ruchi Shukla, for the valuable suggestions and ideas they provided during my annual progression meetings, which played an important role in the development of my project to the current stage. I would also like to thank numerous colleagues and staff members at the Centre for Cancer, Newcastle University, and our collaborators for all their support and assistance throughout this project.

I am deeply indebted to my family and friends for the unconditional love, care and understanding that they always offered me. Special thanks to my parents Sasidharan Nair and Suchitra Devi, who are the sole reason for who I am today. I can't thank them enough for the countless sacrifices they've made, to help me achieve my dreams. I would also like to thank my sister Janaki, brother Anand and other family members Gelu and Neethiya for their love and support during my studies and especially during my thesis writing. Most importantly, a very special thanks to my friends – Marie Claire, Amal, Jamie and Nurettin for those games nights, endless chats, and many more memorable moments to cherish.

I am extremely grateful to my funders – Newcastle University Overseas Research Scholarship (NUORS), Funds for Women Graduates (FFWG) and Bloodwise for the financial support throughout my research. Finally, thanks to the almighty for all the blessings and opportunities that came my way to help me achieve my goals.

Table of Contents

List of Figures

Figures of Chapter 1

Figures of Chapter 2

Figures of Chapter 3

Figures of Chapter 4

Figures of Chapter 5

Figures of Chapter 6

Figures of Chapter 7

List of Tables

Tables of Chapter 1

Tables of Chapter 2

Tables of Chapter 3

Tables of Chapter 4

Tables of Chapter 5

Tables of Chapter 6

List of Equations

Abbreviations

-5/5q-	Monosomy 5/chromosome 5q deletions
-7/7q-	Monosomy 7/chromosome 7q deletions
+ 8	Trisomy 8
2-HG	2-hydroxyglutarate
5caC	5-carboxylcytosine
5fC	5-formylcytosine
5-hmC	5-hydroxymethyl cytosine
5-hmU	5-Hydroxymethyl uracil
5-mC	5-methylcytosine
α-KG	α-Ketoglutarate
ABC	ATP binding cassette
ABCB1	ATP Binding Cassette Subfamily B Member 1
ABL	Abelson Tyrosine-Protein Kinase 1
ADC	Antibody drug conjugates
ADP	Adenosine diphosphate
AID/APOBEC	Activation-induced cytidine deaminase/ apolipoprotein B mRNA-editing enzyme complex
AML	Acute Myeloid Leukaemia
ANOVA	Analysis of variance
APL	Acute Promyelocytic Leukaemia
Ara-C	Cytarabine
ASXL1	Additional Sex Combs Like 1, Transcriptional Regulator
ATP	Adenosine triphosphate
ATRA	All-trans retinoic acid
AU	Arbitrary units
BAF	B-allele frequency
BCA	Bicinchoninic acid
BCR	Breakpoint Cluster Region
BER	Base excision repair
BLAST	Basic Local Alignment Search Tool
bp	Base pairs
BRD4	Bromodomain Containing 4
C4/C5	Carbon 4/5

CAS9	CRISPR associated protein 9
CBF	Core-binding factor
CD	Cluster of differentiation
CDA	Cytidine deaminases
CDK7/9	Cyclin-Dependent Kinase 7/9
cDNA	Complementary DNA
CEBPA	CCAAT/enhancer-binding protein-alpha
CHIP	Clonal haematopoiesis of indeterminate potential
Chr	Chromosome
CML	Chronic myeloid leukaemia
CMML	Chronic myelomonocytic leukaemia
CN-AML	Cytogenetically normal acute myeloid leukaemia
CpG	Cytosine phosphate-link Guanine dinucleotides
CR	Complete remission
CRISPR	Clustered regularly interspaced short palindromic repeats
CSF	Colony Stimulating Factor
CSF1R	Colony stimulating factor 1 receptor
C-terminus	Carboxy terminus
D	Aspartate
DAC	Decitabine, 5-aza-2'-deoxycytidine
ddH_2O	Deionised distilled water
det*P*	Detection *P* values
DMP	Differentially methylated probes
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
DNMT1/2/3A/3B	DNA methyltransferase 1/2/3A/3B
DNR	Daunorubicin
DSBH	Double-stranded β-helix
EFS	Event-free survival
EPO	Erythropoietin
ERK	Extracellular Regulated Kinase
ET	Essential thrombocythemia
Evi1	Ecotropic Viral Integration Site 1
FAB	French, American, British

FBS	Foetal Bovine Serum
FC	Fold change
FDR	False discovery rate
FLAG-IDA	Fludarabine, Ara-C, G-CSF and Idarubicin
FLT3	FMS-like tyrosine kinase 3
G	Guanine
G1/G2	Gap or growth phase 1/2
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GATA-1 ,2,3	GATA binding protein
G-CSF	Granulocyte colony-stimulating factor
GDP	Guanosine diphosphate
GFP	Green Fluorescent Protein
GO	Gemtuzumab ozogamicin
gRNA	Guide RNA
GTP	Guanosine triphosphate
H3K4me3	Histone 3 Lysine 4 trimethylation
HDR	Homology-directed repair
HEL	Human erythroleukaemia
HLA-DR	Human Leukocyte Antigen – DR isotype
HMRN	Haematological Malignancy Research Network
HOXA	Homeobox A
HRAS	Harvey Rat Sarcoma Viral Oncogene Homolog
HSC	Haematopoietic Stem Cell
IDAX	Inhibition of the Dvl and Axin Complex Protein
IDH1/2	Isocitrate dehydrogenase 1/2
IL	Interleukins
inv	inversion
ITD	Internal tandem duplication
JAK2	Janus Kinase 2
KMT2A	Lysine Methyltransferase 2A
KRAS	Kirsten rat sarcoma viral oncogene homolog.
M4EO	M4 subtype with eosinophilia
MAPK	Mitogen activated protein kinase
Mb	Megabases

MBD	Methyl CpG binding domain
MDS	Myelodysplastic syndrome
MECOM	MDS1 And EVI1 Complex Locus Protein
MeCP2	Methyl CpG binding protein 2
Meis1	Myeloid Ecotropic Viral Integration Site 1 Homolog
miRNA	Micro RNA
MKL1	Megakaryoblastic Leukaemia 1 Protein
MLL	Mixed-Lineage Leukaemia
MOI	Multiplicity of infection
MPL	Myeloproliferative Leukaemia Protein
MPN	Myeloproliferative Neoplasms
MPO	Myeloperoxidase
MTT	3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide
MUSCLE	Multiple Sequence Comparison by Log- Expectation
MYH11	Myosin Heavy Chain 11
NCoR	Nuclear receptor corepressor
ncRNA	Noncoding RNA
NHEJ	Non-homologous end joining
NPM1	Nucleophosmin 1
NRAS	Neuroblastoma RAS viral (v-ras) oncogene homolog
NSUN1/3	Nucleolar protein 2/Sun Domain Family, Member 1/3
N-terminus	Amino terminus
NUP214	Nucleoporin 214
OGA	O-GlcNAcase
O-GlcNAc	O-linked N-acetylglucosamine
OGT	O-GlcNAc transferase
ORR	Overall response rate
OS	Overall survival
*P*adj	Adjusted *P* value
PAM	Protospacer-adjacent motif
PAR	Putative azacytidine resistant
PBS	Phosphate buffered saline
PC1/2	Principal component 1/2
PCA	Principal component analysis

PCR	Polymerase Chain Reaction
PDGFR	Platelet-derived growth factor receptors
PI3K	Phosphatidylinositol 3 kinase
PLZF	Promyelocytic leukaemia zinc finger
PMF	Primary myelofibrosis
PML	Promyelocytic Leukaemia Protein
POLR2A S2P	RNA Pol II subunit A phosphorylated at serine 2
Poly-A	Poly adenylation
PTD	Partial tandem duplication
PTP	Protein tyrosine phosphatases
PTPN11	Protein Tyrosine Phosphatase Non-Receptor Type 11
PV	Polycythemia vera
RAF	Rapidly Accelerated Fibrosarcoma family kinases
RARA	Retinoic Acid Receptor Alpha
RAS	Rat sarcoma
RBM15	RNA Binding Motif Protein 15
RIPA	Radio-Immunoprecipitation Assay
RNA	Ribonucleic acid
RNAseq	RNA sequencing
RNR	Ribonucleoreductase
RPL	Ribosomal protein large subunit
RPS	Ribosomal protein small subunit
rRNA	Ribosomal RNA
RTK	Receptor tyrosine kinase
RUNX1	RUNT-related transcription factor 1
RUNX1T1	RUNX1 Partner Transcriptional Co-Repressor 1
RXRA	Retinoic X receptor alpha
SAM	S-adenyl methionine
sAML	Secondary AML
SCF	Stem cell factor
SDS	Sodium dodecyl sulphate
Ser	Serine
SETD1A	SET Domain Containing 1A Histone Lysine Methyltransferase
SH-2	Src Homology-2

SHP-2	Src homology 2 domain containing protein tyrosine phosphatase-2
SMRT	Silencing mediator of retinoic acid and thyroid hormone receptor
SNP	Single nucleotide polymorphism
S-phase	Synthesis phase
STAT3/5	Signal Transducer and Activator of Transcription 3/5
T	Thymine
t	translocation
T-AML	Therapy related AML
TBE	Tris/Borate/EDTA
TBS	Tris-buffered saline
TCGA	The Cancer Genome Atlas
TCRb	T-cell receptor beta chain
TDG	Thymine DNA glycosylase
TET1/2/3	Ten-eleven translocation methylcytosine dioxygenases 1/2/3
TET2 bi	*TET2* biallelic mutated
TET2 mono	*TET2* monoallelic mutated
Thr	Threonine
TKD	Tyrosine kinase domain
TMD	Transmembrane domain
T-MDS	Therapy related MDS
TPO	Thrombopoietin
UHRF	Ubiquitin-like, containing PHD and RING finger domain
UTR	Untranslated region
V	Valine
VEGFR	Vascular endothelial growth factor receptor
Vpr	Viral protein R
vs	versus
WHO	World Health Organisation
WT1	Wilm’s tumour 1

Chapter 1: Introduction

1.1. Haematopoiesis

Haematopoiesis is the biological process of generation of differentiated mature blood cells from multipotent haematopoietic stem cells (HSCs) through asymmetric division. Each cell division results in self-renewal of the HSC and generation of a more differentiated lineage-restricted myeloid or lymphoid progenitor. These progenitors further differentiate into more lineage specific erythroid, megakaryocytic, lymphocytic, granulocytic and monocytic lineage cells (Rieger et al. 2012).

This process which starts from early embryonic stages and continues throughout life, is tightly regulated by a family of extracellular ligands known as cytokines (growth factors) and cytokine receptors that are structurally and functionally conserved. The cytokines involved in this signalling cascade include colony-stimulating factors (CSF), interleukins (ILs), interferons, thrombopoietin (TPO) and erythropoietin (EPO) (Robb 2007). Some of these cytokines direct the growth of a stem cell to a particular lineage such as EPO regulating erythrocytic progenitors, whereas cytokines such as IL-3 regulate multiple lineages. These cytokines along with various other transcription factors and signalling molecules play an important role in lineage determination (Robb 2007).

Transcription factors such as GATA-1 and PU.1 are co-expressed in common myeloid progenitors, whereas their mutually exclusive expression pattern is key to the transition of these myeloid progenitors to megakaryocyte-erythroid progenitors or granulocyte-monocyte progenitors (Zhu et al. 2002). PU.1 is downregulated in granulocyte-monocyte progenitors whereas GATA-1 is downregulated in megakaryocyte-erythroid progenitors. Targeted silencing of these transcription factors leads to impairment in development of respective lineage determination and showed that PU.1 and GATA-1 are absolutely essential for the development of megakaryocyte-erythroid progenitors and granulocyte-monocyte progenitors respectively (Zhu et al. 2002). An example for the role of signalling molecules in haematopoiesis is the Notch receptor signalling which is fundamental to the T-cell fate initiation and determination. During the transition from common lymphoid progenitors to early T-cell progenitors, NOTCH1 induces the expression of GATA-3 and other T-cell lineage restricting proteins, thereby supporting the T-cell lineage determination (Zhu et al. 2002). Further, NOTCH1 represses the B-cell fate supporting transcription factors such as EBF1 and E2A and also the myeloid fate by inhibiting PU.1 and CEBPα (Zhu et al. 2002).

As such, the regulatory network of transcription factors and signalling molecules drive the fate of an HSC towards self-renewal or a particular lineage commitment. The process of haematopoiesis along with the key players in its molecular regulation is summarised based on current evidence (Figure 1.1)

1.2. Acute Myeloid Leukaemia

Acute Myeloid Leukaemia (AML) is characterised by uncontrolled proliferation of clonal undifferentiated myeloid blast cells that occupy the bone marrow and inhibit normal haematopoiesis that ultimately results in bone marrow failure (Medinger et al. 2017). Some of the common symptoms shown by majority of patients with AML include leucocytosis, anorexia, fatigue, weight loss and signs of bone marrow failure like thrombocytopenia and anaemia (De Kouchkovsky et al. 2016).

1.2.1. Epidemiology of AML

AML is a clinically heterogeneous disease which generally has a poor prognosis. AML is the most common acute leukaemia worldwide and incidence has increased by 73% in the UK over the last 40 years (Shysh et al. 2017). During this time there has been only very modest improvements in treatment strategies for AML. The median age at diagnosis is 72 years and the 5-year relative survival rate of patients diagnosed with AML is 14.5% with the 5-year survival of patients aged >70 years being as low as ~1% (HMRN(2010-2016)). Outcome is better in younger patients where the majority will achieve complete remission (CR), although relapse is common, and the 5-year overall survival (OS) rate is 40-50% (HMRN(2010-2016)).

1.2.2. Pathophysiology of AML

AML can occur as a *de novo* malignancy or subsequent to an underlying haematological disorder (Secondary AML; Section 1.2.4), such as myelodysplastic syndrome (MDS), or as a consequence of cytotoxic therapy given for another condition (Therapy related AML; Section 1.2.3), and can include exposure to radiation or alkylating agents, which are established human leukaemogens. One model of leukaemogenesis (the two-hit model of leukaemogenesis), suggests that leukaemia develops when class I mutations, which activate pro-proliferative pathways, occur together with class II mutations, which deregulate normal haematopoietic differentiation (De Kouchkovsky et al. 2016; Gilliland et al. 2002; Kihara et al. 2014; Takahashi 2011). The pathogenesis and nature of AML in terms of treatment response, would depend on

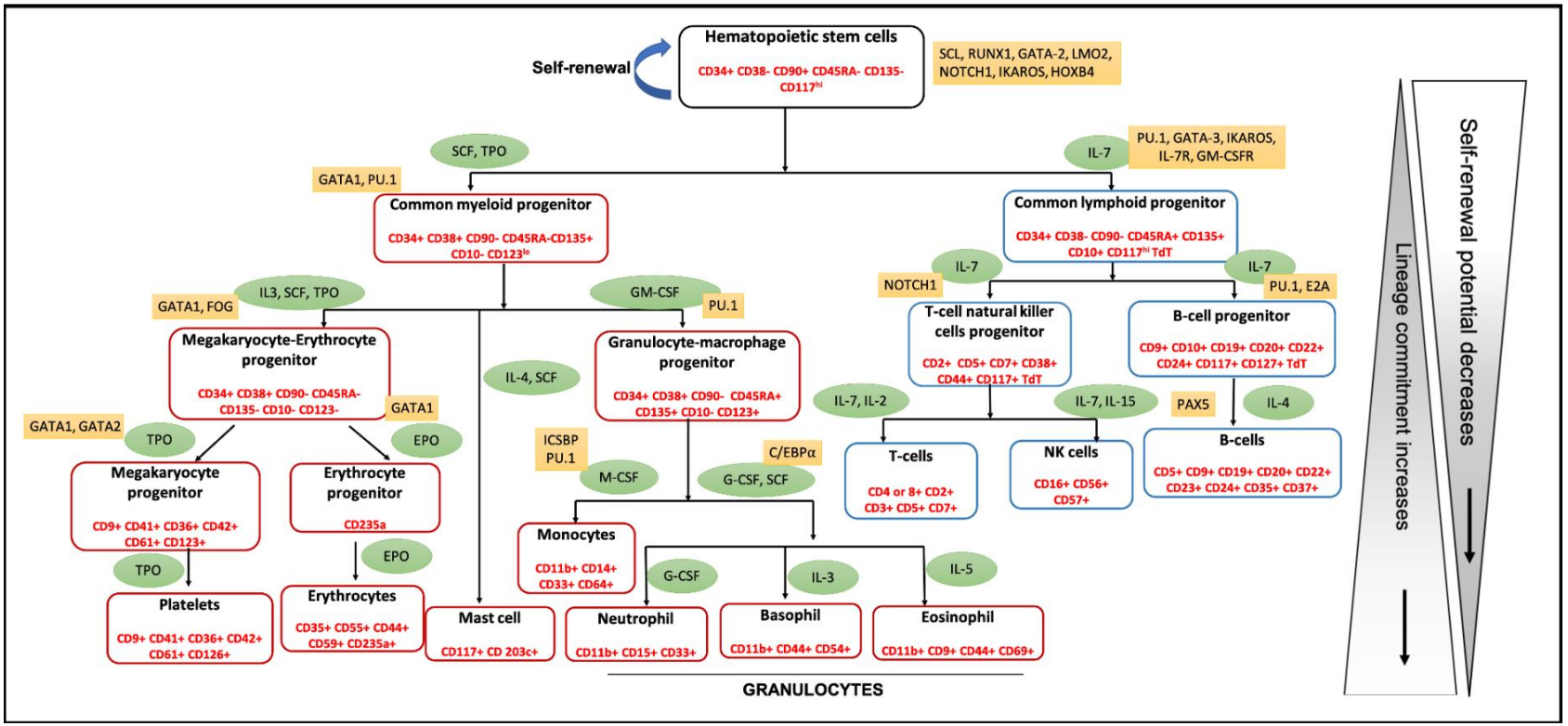

Figure 1.1. Human haematopoiesis and its molecular regulation. Haematopoietic stem cells differentiate into lineage committed myeloid or lymphoid progenitor cells. These progenitor cells further differentiate into lineage restricted cells and the process is driven by a regulatory network composed of various cytokines (Green circles) and transcription factors (yellow boxes) whose expression levels are critical for lineage determination at each stage of differentiation. Myeloid lineage cells are red boxed while lymphoid lineage cells are blue boxed. Towards the bottom of the flow chart, the lineage commitment increases whereas the self-renewal potential decreases. The surface markers specific to each cell type are shown in red.

CD, Cluster of differentiation; EPO, erythropoietin; ICSBP, Interferon Consensus Sequence Binding Protein 1; GM/G/M-CSF, Granulocyte-macrophage/granulocyte/macrophage-colony stimulating factor; IL, interleukin; NK cells, natural killer cells; SCF, stem cell factor; SCL, stem cell protein; TPO, Thrombopoietin

the particular somatic alterations, chromosomal translocations, and the interaction between the genomic abnormalities. The various mutations and chromosomal abnormalities are discussed in detail in section 1.2.5.

1.2.3. AML classifications and the major subtypes

The French, American, British (FAB) classification was established in 1976 and identifies seven AML subtypes named M0-M7, based on morphological features. AML subtypes are distinct in terms of molecular features and surface marker expression (Table 1.1). However, this type of classification does not completely depict the mutational complexity of AML. Therefore, the WHO published a revised classification of AML encompassing the different genomic abnormalities which can be used to delineate AML subtypes (Table 1.2) (Arber et al. 2016). Moreover, some of these genomic abnormalities can be prognostically relevant and can be used in determining therapeutic choices (Döhner et al. 2017). This section aims to introduce some of the major subtypes of AML, their cytogenetic features and existing treatment approaches.

Cytogenetically Normal AML

Cytogenetically normal AML (CN-AML) which is characterised by no cytogenetically visible abnormal genetic features, forms the largest subset of AML, accounting for 45%-60% of all AML cases (Yang et al. 2021). The prognostic stratification classifies AML into three risk-based categories, namely favourable, intermediate, and unfavourable (poor-risk), primarily based on the cytogenetic features of AML (Döhner et al. 2017). The genetic mutations specific to each of the risk categories are discussed in section 1.2.5. As such, patients in the favourable risk group are treated with conventional chemotherapy (incorporating cytarabine and daunorubicin) whereas those in poor risk group are treated with allogeneic hematopoietic stem cell transplantation. The intermediate risk group mostly comprises of patients with CN-AML (Martelli et al. 2013) and identifying the optimal treatment regimens has been challenging and outcomes are diverse, primarily due to the high molecular heterogeneity of this group (Yang et al. 2021). For instance, *Nucleophosmin 1 (NPM1)* mutations occur at a high frequency and are seen in 45.7% to 63.8% of CN-AML (Falini et al. 2007). Likewise, *FMS-like tyrosine kinase 3* (*FLT3)* mutations are seen in 28%–34% of CN-AML (Meyer et al. 2014; Yang et al. 2021). Patients with CN-AML with *NPM1* mutation tend to benefit from conventional chemotherapy and hence this particular group are classified in the favourable risk category, whereas patients

Subtype	Description	Occurrence in AML	Immunophenotype	Distinctive features
M0	Minimally differentiated AML	5%	CD13+, CD33+, CD11b+, CD11c+, CD14+, CD15+	Undifferentiated blast cells with morphological features of both myeloblasts and lymphoblasts.
M1	Acute myeloblastic leukaemia without maturation	< 20%	MPO+, CD13+, CD33+, CD117+	Round nuclei, lightly granulated, Auer rods in cytoplasm in 5-10% of cases.
M2	Acute myeloblastic leukaemia with maturation	30%	MPO+, CD13+, CD15+, Sudan black+	More obvious granulation and prominent auer rods. Up to 25% of patients have t(8;21) translocation.
M3	Acute promyelocytic leukaemia	10%	CD13+, CD33+, HLA-DR -, CD34-	Presence of promyelocytes (hyper-granular cells) and t(15;17) translocation.
M4 M4EO	Acute myelomonocytic leukaemia	25-30%	CD13+, CD4+, CD15+,CD33+, CD11b+, CD11c+, CD14+, CD64+,	A mixture of myeloid and monocytic elements. Nucleoli are prominent. M4EO (M4 with eosinophilia) is when 5% of cells are abnormal eosinophils.
M5	Acute monocytic leukaemia	10%	CD14+, CD68+, CD4+, CD11c+, HLA-DR+, CD64+	M5a is when monoblasts predominate whereas M5b is when there is a presence of promonocytes and monocytes in addition to monoblasts. t(9;11), t(8;11), t(8,16) translocations and 11q23 rearrangements are common.
M6	Acute erythroid leukaemia	4%	CD13+, CD33+, CD15+, Glycophorin A+, Glycophorin C+	Presence of unusual megaloblasts and at times multinucleated erythroid precursors. Often there is a change in morphology of erythrocytes to schistocytes or mushroom-shaped cells. Loss of part or all of chromosome 5 or 7 is another feature.
M7	Acute megakaryocytic leukaemia	1-3%	CD41+, CD61+, CD42+, CD13+, CD33+, CD34+	Presence of micromegakaryocytes, multinucleated cells and pseudopods. Chromosome 3 abnormalities are common.

Table 1.1. French, American, British (FAB) classification system of AML. This system classifies AML into seven subtypes (M0-M7) based on distinct morphological and phenotypic features. The frequency of occurrence of each subtype in AML, the molecular features such as the presence of specific surface markers and the morphological features of individual subtypes are summarised (Schiffer CA 2003; Ladines-Castro et al. 2016; Boyer 1996).

CD, Cluster of differentiation; HLA-DR, Human Leukocyte Antigen – DR isotype; M4EO, M4 subtype with eosinophilia; MPO, Myeloperoxidase; + or - symbols indicate presence or absence respectively.

AML with recurrent genetic abnormalities
AML with t(8;21)(q22;q22.1);RUNX1-RUNX1T1
AML with inv(16)(p13.1q22) or t(16;16)(p13.1;q22);CBFB-MYH11
APL with PML-RARA
AML with t(9;11)(p21.3;q23.3);MLLT3-KMT2A
AML with t(6;9)(p23;q34.1);DEK-NUP214
AML with inv(3)(q21.3q26.2) or t(3;3)(q21.3;q26.2); GATA2, MECOM
AML (megakaryoblastic) with t(1;22)(p13.3;q13.3);RBM15-MKL1
Provisional entity: AML with BCR-ABL1
AML with mutated NPM1
AML with biallelic mutations of CEBPA
Provisional entity: AML with mutated RUNX1
AML with myelodysplasia-related changes
Therapy-related myeloid neoplasms
AML, not otherwise categorised
AML with minimal differentiation
AML without maturation
AML with maturation
Acute myelomonocytic leukaemia
Acute monoblastic/monocytic leukaemia
Pure erythroid leukaemia
Acute megakaryoblastic leukaemia
Acute basophilic leukaemia
Acute panmyelosis with myelofibrosis
Myeloid sarcoma
Myeloid proliferations related to Down syndrome
Transient abnormal myelopoiesis
Myeloid leukaemia associated with Down syndrome

Table 1.2. WHO classification of AML (Arber et al. 2016). This system considers various cytogenetic features, morphologic, immunophenotypic and prognostic data in the classification of AML.

ABL, Abelson Tyrosine-Protein Kinase 1; AML, Acute myeloid leukaemia; APL, Acute promyelocytic leukaemia; BCR, Breakpoint Cluster Region; CBFB, Core-Binding Factor Subunit Beta; CEBPA, CCAAT/enhancer-binding protein-alpha; GATA2, GATA Binding Protein 2; inv, inversion; KMT2A, Lysine Methyltransferase 2A; MECOM, MDS1 And EVI1 Complex Locus Protein; MKL1, Megakaryoblastic Leukaemia 1 Protein; MLL3, Mixed-Lineage Leukaemia Protein 3; MYH11, Myosin Heavy Chain 11; NPM1, Nucleophosmin 1; NUP214, Nucleoporin 214; PML, Promyelocytic Leukaemia Protein; RARA, Retinoic Acid Receptor Alpha; RBM15, RNA Binding Motif Protein 15; RUNX1, RUNT-related transcription factor 1; RUNX1T1, RUNX1 Partner Transcriptional Co-Repressor 1; t, translocation

with CN-AML with *FLT3* mutations do not benefit from conventional chemotherapy and hence this group are classified as poor-risk (Martelli et al. 2013). Molecular therapy targeting *FLT3*-mutated AML such as midostaurin might potentially benefit this category of patients (Martelli et al. 2013). As such, clinical management, and prognosis of patients with CN-AML has been greatly dependent on the genetic mutations present in this group.

Acute Promyelocytic Leukaemia

Acute Promyelocytic Leukaemia (APL) is most commonly defined by a balanced translocation between chromosomes 15 and 17 (t(15;17)) leading to the fusion of promyelocytic leukaemia (*PML*) gene with the retinoic acid receptor alpha (*RARA*) gene forming the fusion protein PML-RARA, which results in a block in myeloid differentiation at the promyelocytic stage (Stahl et al. 2019). The WHO recommends that the patients with t(15;17) be considered to have a diagnosis of AML even if the blast percentage is lower than 20% (Sangle et al. 2011). The *PML-RARA* fusion is seen in 98% of patients with APL although there are other translocations that give rise to APL, including t(11;17) which generates a fusion gene between the promyelocytic leukaemia zinc finger (*PLZF*) gene and *RARA* (de Thé et al. 2015). The strong association of *RARA*-related alterations in APL suggests a role of RARA in APL pathogenesis (de Thé et al. 2015).

APL accounts for 10-15% of AML (Jimenez et al. 2020) and is by far the most curable subtype of AML with the introduction of all-trans retinoic acid and (ATRA) and arsenic trioxide-based treatment targeting PML-RARA (Stahl et al. 2019). ATRA triggers rapid differentiation of APL cells into granulocytes whereas arsenic trioxide induces apoptosis and differentiation *in vivo* (de Thé et al. 2015). As such, this combination regimen has a cure rate as high as 80%-90% in APL (Stahl et al. 2019) and is considered an excellent example precision medicine in AML.

Core Binding Factor AML

Core Binding Factor AML (CBF-AML) is characterised by disruption of genes coding for the CBF protein complex, which is a regulator of normal haematopoiesis (Sangle et al. 2011). CBF-AML includes those cases with cytogenetic rearrangements such as translocation between *RUNX1T1* on chromosome 8 and *RUNX1* on chromosome 21 (t(8;21)), forming the *RUNX1-RUNX1T1* fusion gene (Sangle et al. 2011). Another rearrangement that causes CBF-AML is the pericentric inversion of chromosome 16 (inv(16)) leading to the formation of CBFB/MYH11 fusion protein (Borthakur et al. 2021; Sangle et al. 2011). The formation of the fusion proteins alters the normal DNA binding of the CBF protein complex which is a

haematopoietic transcription factor and negatively impacts the myeloid differentiation (Borthakur et al. 2021; Speck et al. 2002).

CBF-AML accounts for 12-15% of all AML cases (Borthakur et al. 2021) and has been prognostically classified in the favourable category due to sensitivity to cytarabine-based therapy (Borthakur et al. 2021). Similar to APL, the WHO recommends that patients with t(8;21) or inv(16) be given a diagnosis of AML regardless of the blast percentage (Sangle et al. 2011).

Therapy-related AML

Therapy-related AML (T-AML) is classified as a distinct entity separate from the other AML subtypes as per the WHO classification system (Table 1.2) (Arber et al. 2016). Therapy-related AML (t-AML) is described as a direct consequence of mutations induced by pro-mutagenic agents used for the treatment of a prior neoplasm, often breast cancer or lymphoid malignancies (Bertani et al. 2020; Godley et al. 2008). T-AML accounts for 5%-20% of all AML (Samra et al. 2020) and is associated with a more aggressive phenotype compared to *de novo* AML due to various patient-related and disease-related features (Belitsky et al. 2020). The patient-related features include old age at presentation which makes the patient less fit for intensive chemotherapy, as well as the presence of comorbidities and sequelae from prior cancer and treatment (Samra et al. 2020). Disease-related features include complex cytogenetics, chromosomal abnormalities such as chromosome 5q deletions (5q-) and chromosome 7 deletions (-7/7q-) and *TP53* mutations which are all associated with chemoresistance to conventional chemotherapy (Samra et al. 2020). As a consequence of these high-risk features, the median OS of patients with t-AML is as low as 8-10 months with a 5-year survival rate of 10%-20% (Belitsky et al. 2020).

T-AML is categorised into three subtypes based on the causative pro-mutagenic agent used in prior treatment – alkylating agent-related t-AML, topoisomerase inhibitor-related t-AML and radiation-related t-AML (Vardiman et al. 2002). T-AML resulting from alkylating agent exposure often presented with cytogenetic alterations including del(5q), monosomy 7 and del(7q), and have latency period averaging 5 years between therapeutic exposure and t-AML diagnosis (Belitsky et al. 2020; Vardiman et al. 2002). The median survival time of this category is approximately 8 months after disease onset (Belitsky et al. 2020). Topoisomerase inhibitor-related t-AML typically present with balanced translocations involving chromosome bands 11q23 (affecting *MLL*) and 21q22 (affecting *RUNX1*) and the other less frequent translocations

include inv(16) (affecting *CBFB*) and t(15;17) (affecting *RARA*) (Vardiman et al. 2002). The latency period is however shorter ranging from 6 months to 5 years (median 2-3 years) for this subtype of t-AML (Vardiman et al. 2002). Although, t-AML induced by topoisomerase inhibitor have an overall poor prognosis, this subtype is relatively more sensitive to remission induction chemotherapy compared to alkylating agent-related t-AML and the survival rates are comparable to that of *de novo* AML with similar chromosomal rearrangements (Godley et al. 2008). Radiation therapy has been identified as another risk-factor for the development of T-AML. Ionising radiation used in radiotherapy can induce extensive DNA damage including double strand breaks leading to genetic alterations affecting genes critical for the growth and differentiation of HSCs and can contribute to leukaemic transformation (Guenova et al. 2013; Bhatia 2013). As such, radiation-related t-AML does not possess a unique clinical or cytogenetic pattern and are similar to alkylating agent-induced AML (Giuseppe et al. 2007; Smith et al. 2003).

1.2.4. Major haematological disorders that predispose patients to AML

Secondary AML (sAML) is a subtype of AML that occurs as a consequence of certain antecedent haematological disorders and such disorders that predispose to developing AML are discussed in this section. sAML has been associated with inferior outcome compared to *de novo* AML (Granfeldt Østgård et al. 2015).

Myelodysplastic Syndrome (MDS)

MDS is a group of heterogeneous clonal haematological disorders characterised by pancytopenia in the peripheral blood (reduction in the number of mature blood cells) and bone marrow dysplasia (presence of abnormal cells in the bone marrow) due to impaired haematopoiesis (Syed et al. 2020). MDS is often referred to as pre-leukaemia due to its propensity to transform into AML. MDS is driven by a complex combination of genomic alterations including mutations in epigenetic regulators such as *TET2*, *DNMT3A* and *ASXL1*, RNA-splicing factors such as *SRSF2*, and signalling genes such as *NRAS* as well as chromosomal aberrations such as chromosome 5q deletions (5q-) and chromosome 7 deletions (-7/7q-) (Sperling et al. 2017). The factors that increase the risk of MDS include increased age and exposure to prior chemotherapy or radiation (Sperling et al. 2017). MDS is typically diagnosed when at least one of the three following criteria are satisfied - dysplasia ≥10%, blasts count from 5%-19% or MDS-related karyotyping (5q-, +8, -7/7q-, 20q-) (Mohammad 2018).

MDS can be broadly classified as *de novo* MDS or therapy-related MDS (t-MDS), where the latter typically occurs 5-7 years post chemotherapy/radiotherapy and tends to have a poor prognosis compared to *de novo* MDS (Dotson et al. 2021). The risk of transforming into AML is relatively higher for patients with t-MDS (55%) compared to *de novo* MDS (30%) (Corey et al. 2007). The initial stages of MDS are characterised by an increase in apoptosis of myeloid cells and their precursors, whereas the later stages of MDS during progression to AML are characterised by a block in differentiation and accumulation of immature blast cells. Allogenic stem-cell transplantation has proven to be beneficial for high risk-MDS (Corey et al. 2007). However, as MDS primarily affects older patients who have a higher risk of treatment-related mortality, other strategies using lenalidomide and hypomethylating agents such as azacytidine and decitabine have been approved for the treatment of MDS and are effective at reducing disease symptoms (Corey et al. 2007).

Myeloproliferative Neoplasms (MPN)

MPN are a group of myeloproliferative malignancies characterised by an impaired haematopoiesis and defective regulation of myeloid cell proliferation resulting in an over-production of mature myeloid lineage cells (Kim et al. 2012). Broadly, MPNs are classified as *BCR-ABL1*-positive MPN such as chronic myeloid leukaemia (CML) and *BCR-ABL1*-negative MPN such as polycythemia vera (PV), essential thrombocythemia (ET) and primary myelofibrosis (PMF) (Kim et al. 2012). Of note, *BCR-ABL1* is a fusion gene, that results from a reciprocal translocation between the breakpoint cluster region gene (*BCR*) at chromosome 22q11 and the Abelson gene (*ABL1*) at chromosome 9q23 and the resulting fusion product is the primary pathogenic driver in most cases of CML (Quintás-Cardama et al. 2009). The BCR-ABL fusion oncoprotein has constitutively active ABL tyrosine kinase function, which in turn activates multiple downstream pathways such as RAS, SHP2 and PI3K–AKT signalling pathways to promote cell proliferation and survival (Ren 2005). As such, inhibitors of tyrosine kinase activity of BCR-ABL including imatinib constitute an important strategy for the treatment of CML (Ren 2005).

BCR-ABL1-negative MPN is characterised by mutations in the *JAK2*, *calreticulin* or *MPL* genes that results in constitutional activation of Janus kinase 2 (JAK2)/signal transducer and activator of transcription (STAT) pathway and confer a pro-survival and proliferation advantage to myeloid cells (Cacemiro et al. 2018; Etheridge et al. 2014). *JAK2* mutations are extremely frequent in PV and are seen in more than 97% of PV cases, whereas these mutations are seen

in 55% and 65% of ET and PMF, respectively (Kim et al. 2012). Approximately 10-15% of patients with PV, ET and PMF progress to acute myeloid leukaemia (Cacemiro et al. 2018).

1.2.5. Landscape of genomic abnormalities in AML

AML is a heterogeneous disease driven by co-occurrence of somatic mutations, epigenetic changes and chromosomal abnormalities (De Kouchkovsky et al. 2016; Sill et al. 2011). Mutations include structural alterations visible cytogenetically, such as translocations, inversions, chromosomal monosomy/trisomy and large insertions/deletions, as well as sub-cytogenetic base substitutions and small insertions/deletions. Currently, data from the TCGA project shows that the highest proportion of mutations reported in AML occur in signalling pathway genes (59%), followed by mutations in DNA methylation related genes (44%), chromatin modifiers (30%), myeloid transcription factors (22%), gene fusions (18%), tumour suppressor genes (16%), spliceosome complex genes (14%), and cohesion complex genes (13%). Mutations in *NPM1* gene are seen in 27% of AML cases (Figure 1.2A) (Coombs, Tallman, et al. 2016; Ley et al. 2013). Broadly, there are three classes of mutations that drive leukaemia in-line with the two-hit hypothesis (Figure 1.2B).

Class I mutations

Class I mutations are mutations that result in activation of pro-proliferative signalling pathways culminating in an increased survival advantage to cells. This class of mutations accounts for the most frequent subset of AML driver mutations (De Kouchkovsky et al. 2016). The common mutational targets of this class include *FLT3*, RAS oncogene family (*KRAS/NRAS), KIT*, *PTPN11*, *JAK2* and *TP53*.

The *FLT3* gene is located at chromosomal locus 13q12 and encodes a receptor tyrosine kinase (RTK), which is normally expressed in lymphoid and myeloid progenitor cells and become silent as the cells differentiate into more lineage committed stages. *FLT3* mutations are found in about 30% of AML (Daver et al. 2019) and include (a) internal tandem duplications (ITD) or insertions or deletions affecting the juxtamembrane domain and (b) point mutations in the activation loop of tyrosine kinase domain (TKD) (Renneville et al. 2008). *FLT3-ITD* mutations occur in approximately 25% of all AML cases whereas *FLT3-TKD* mutations occur in 7-10% of AML (Daver et al. 2019). *FLT3* mutations result in constitutive phosphorylation of receptors and activation of downstream pro-survival signalling pathways such as Ras/mitogen activated protein kinase (MAPK) pathway, phosphatidylinositol 3 kinase (PI3K)/AKT pathway and Janus kinase 2 (JAK2)/STAT5 pathway (Grundler et al. 2005; Reindl et al. 2006).

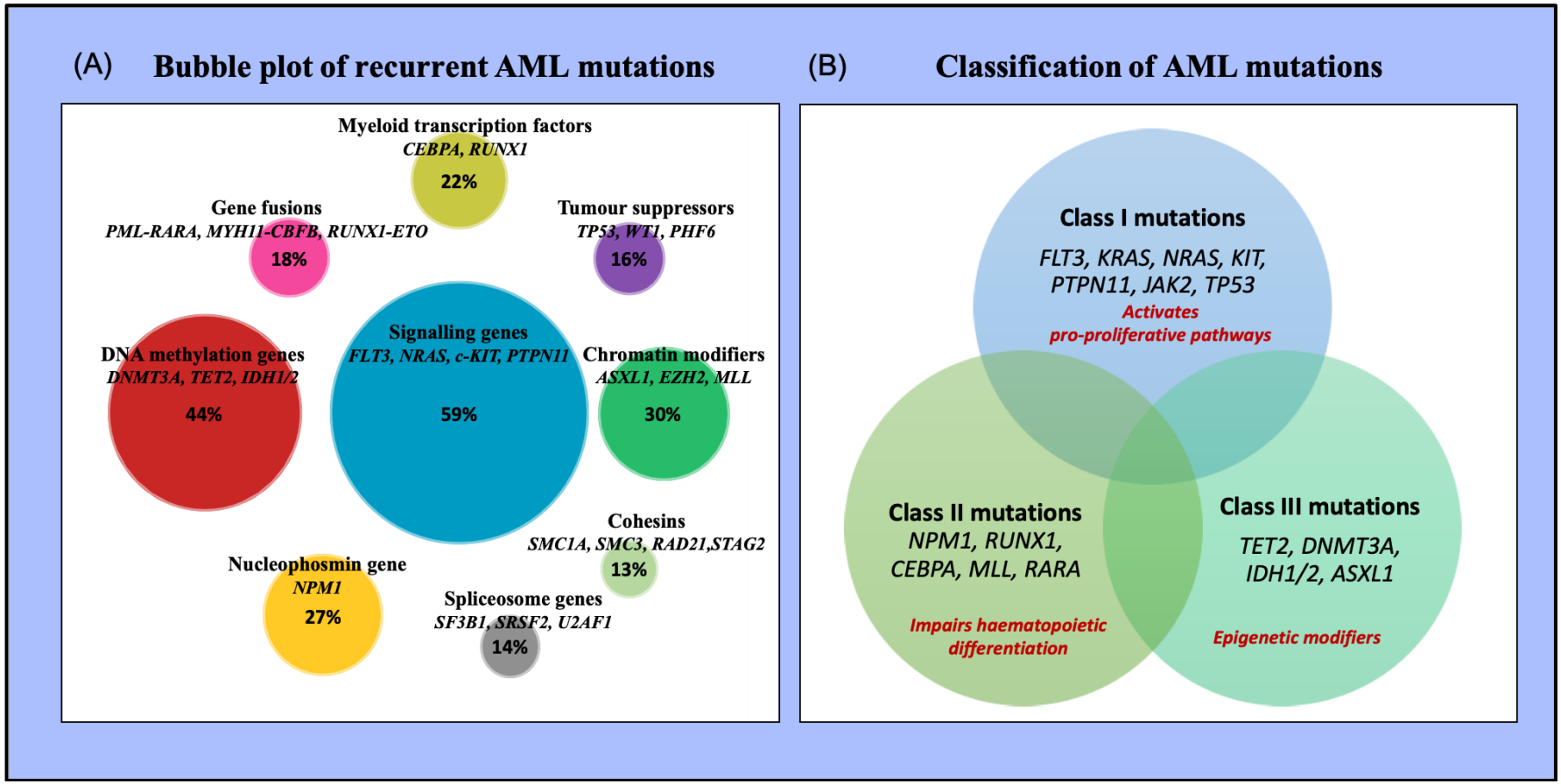

Figure 1.2. Mutational spectrum of AML (A) The bubble plot shows the distribution of genomic aberrations according to The Cancer Genome Atlas (TCGA) in AML (data obtained from (Ley et al. 2013). Bubble size represents the frequency of respective mutation group in AML. This shows that the highest percentage of recurrent aberrations in AML are found in signalling pathway genes (59%) with second highest being DNA methylation related genes (44%). (B) Generally, mutations in AML are classified into three classes. Class I mutations results in the activation of pro-proliferative pathways, Class II mutations impairs haematopoietic differentiation and Class III mutations play a role in epigenetic regulation. Based on two-hit hypothesis of leukaemogenesis, mutations in genes that belong to at least two different classes co-operate to drive leukaemia. This is represented by the Venn diagram. Data obtained from (Kelly et al. 2002; Lagunas-Rangel et al. 2017; De Kouchkovsky et al. 2016)

Of note, nearly 20% of patients have a change in *FLT3* mutation status at relapse, either loss or gain of *FLT3-ITD* or *FLT3-TKD* or change in length of *FLT3-ITD* (Daver et al. 2019). Precisely, *FLT3-TKD* mutations are more frequently lost at relapse compared to *FLT3-ITD* (7% vs 4%) whereas *FLT3-ITD* mutations are more frequently newly detected at relapse compared to *FLT3-TKD* (8% vs 2%) (McCormick et al. 2010; Daver et al. 2019). Approximately 6% of cases acquire a change in number or length of *FLT3-ITD* at relapse (Renneville et al. 2008). *FLT3-ITD* mutations in particular are correlated to a significantly shorter time to relapse compared to *FLT3* point mutations (Cloos et al. 2006).

RAS oncogenes, namely *HRAS, KRAS* and *NRAS,* encode for a family of guanine nucleotide binding proteins that bind to various receptors such as FLT3 or c-KIT and regulate signal transduction via downstream targets such as MAPK/ERK/RAF (Renneville et al. 2008). RAS proteins play an important role in regulating differentiation, proliferation, and apoptosis and generally these proteins maintain an equilibrium between the GDP-bound inactive state and GTP-bound active state (Mitin et al. 2005; Renneville et al. 2008). Point mutations in *NRAS* and *KRAS* occur in approximately 11-16% and 4-5% of AML, respectively (Ball et al. 2019) and lead to constitutive activation of RAS proteins and pro-survival downstream signalling pathways. *RAS* mutations generally co-occur with inv(16)/t(16;16) and inv(3)/t(3;3) in 35% and 27% of AML, respectively, and are less frequent (2%) in t(15;17) AML (Renneville et al. 2008). *RAS* and *FLT3* mutations are generally mutually exclusive and only rarely (2%) co-occur in AML consistent with mutations in these genes both affecting proliferation (Bacher et al. 2006; Bowen et al. 2005; Renneville et al. 2008). *RAS* mutations are involved in the progression from MDS to AML (Shih et al. 2004; Christiansen et al. 2005). Also, *RAS* mutational status does not change between diagnosis and relapse in the majority (95%) of AML patients (Bacher et al. 2006).

The *c-KIT* oncogene is located at chromosomal locus 4q12 and encodes a transmembrane glycoprotein (RTK subclass III family) which acts as a receptor for stem cell factor (SCF) protein. SCF ligand binding leads to receptor dimerization of c-KIT which results in transphosphorylation and activation of downstream pathways that function in differentiation, survival, proliferation, and migration of cells, particularly HSCs (Abbaspour Babaei et al. 2016; Renneville et al. 2008). Gain-of-function mutations affecting the extracellular domain (indels of exon 8), the juxtamembrane domain (exon 11 ITD) or the activation loop of TKD (D816; mostly D816V substitution in exon 17) lead to ligand-independent dimerization of c-KIT receptors and constitutive activation of downstream pro-proliferative signalling pathways such

as *PI3K* and *STAT3* (Abbaspour Babaei et al. 2016; Renneville et al. 2008). *c-KIT* mutations commonly occur in CBF-AML and AML with trisomy 4 and are rare in t(15;17) AML. c-KIT D816 and exon 8 mutations occur in 2% and 6-8% of AML respectively (Schnittger et al. 2006; Renneville et al. 2008).

Protein tyrosine phosphatase non-receptor 11 (PTPN11) is located at chromosomal locus 12q24 and encodes a cytoplasmic protein tyrosine phosphatase (SHP-2) which is associated with activating *RAS/MAPK* signalling cascade through cytokine receptors such as IL-3, IL-6, GM-CSF and EPO (Renneville et al. 2008). SHP-2 has three functional domains such as two Src Homology-2 (SH-2) domains (N-SH2 and C-SH2) and a protein tyrosine phosphatases (PTP) domain (catalytic domain). The N-SH2 domain interacts with the PTP domain to block and auto-inhibit the catalytic site and regulate SHP-2 mediated signalling activation (Renneville et al. 2008). Most *PTPN11* mutations occur at the interacting region of PTP and N-SH2 domains and are more frequent in paediatric AML (4.4%) compared to adult AML (2.6%) (Tartaglia et al. 2005; Tartaglia et al. 2004; Renneville et al. 2008).

Janus Kinase 2 (JAK2) is a cytoplasmic tyrosine kinase that functions in signalling downstream of cytokine receptors such as IL-3R, IL-5R, GM-CSF-R, EPOR, TPOR, the class II receptor cytokine family and the glycoprotein Gp-130 receptor family. JAK2 consists of two domains — tyrosine kinase domain (JH1) and pseudo-kinase domain (JH2), where JH2 domain interacts with JH1 domain and negatively regulates JAK2 function (Renneville et al. 2008). The most common *JAK2* mutation in haematological diseases is JAK2V617F, which occurs in JH2 domain. The mutated enzyme is constitutively activated with an increased kinase activity resulting in hyperactivation of the downstream signalling cascade involving PI3K/AKT, STAT5 and ERKs (Ungureanu et al. 2011). This particular mutation is very frequent (70%) in AML developing from myeloproliferative neoplasms (MPN) as compared to *de novo* AML (1.6%) (de Noronha et al. 2019; Kisseleva et al. 2002; Renneville et al. 2008).

TP53 (located on chromosome 17p13) is a tumour suppressor gene that is commonly referred to as the ‘guardian of the genome’ (Levine 1997) and encodes a transcription factor (TP53) that controls the expression levels of various downstream targets involved in cell cycle arrest, DNA damage response and apoptosis (Hunter et al. 2019). Loss-of-function mutations (deletions and point mutations) affecting exon 4 to 8 of *TP53* results in genomic instability and impaired DNA repair and apoptosis. *TP53* mutations are more frequent (30%) in therapy related leukaemia comprising t-AML and t-MDS as compared to *de novo* AML (less than 10%) (Hunter et al. 2019; Barbosa et al. 2019; Renneville et al. 2008).

Class II mutations

The mutations that impair the differentiation process of cells during haematopoiesis leading to leukaemogenesis are generally classified as class II mutations. The target genes of class II mutations include but are not limited to *NPM1, RUNX1, CEBPA, MLL* and *RARA.*

Nucleophosmin 1 (NPM1) located at 5q35, encodes a nucleolar phosphoprotein which shuttles between the cytoplasm and nucleus. Depending on its interacting partners, expression level and localisation, NPM1 functions as an oncoprotein or tumour suppressor (Federici et al. 2013; Renneville et al. 2008). *NPM1* is one of the top three mutated genes in AML, others being *FLT3* and *DNMT3A* (DiNardo et al. 2016). Most *NPM1* mutations affect exon 12 and more than 95% of *NPM1* mutations are 4-bp deletions at position 960. These mutations affect the nuclear export signal motif at the C-terminus resulting in aberrant cytoplasmic localisation of the NPM1 protein triggering leukaemogenesis (Heath et al. 2017). *NPM1* mutations are generally stable during AML evolution from diagnosis to relapse. They are present in approximately 30% of adult *de novo* AML and 50% of CN-AML, with a majority of them co-occurring with mutations in epigenetic regulators such as *DNMT3A, TET2, IDH1 or IDH2. NPM1* mutation is relatively rare (2-6%) in childhood AML (DiNardo et al. 2016; Heath et al. 2017; Renneville et al. 2008)

Runt-related transcription factor 1 (*RUNX1* or *AML1* or *CBFA2*) is located at 21q22 and encodes a subunit of core-binding factor (CBF) that regulates expression of various genes such as *IL-3, M-CSF* receptor, T-cell receptor beta chain (*TCRβ*) and myeloperoxidase (*MPO*) essential for haematopoiesis and differentiation of myeloid progenitor cells (Lam et al. 2012). *RUNX1* is deregulated through point mutations, chromosomal translocations and amplifications. The common translocations involving *RUNX1* are t(8;21) which forms the *RUNXI-RUNX1T1* fusion product in AML, t(12;21) which forms *RUNX1-ETV* in childhood ALL and t(3;21) which forms *RUNX1-MDS1* in MDS and CML. The RUNX1-RUNX1T1 fusion protein includes the N-terminal region of RUNX1 that has the DNA-binding Runt1 domain and the entire RUNX1T1 protein. The RUNX1T1 protein consists of NHR3 and NHR4 domains that can recruit transcriptional repressors such as NCoR/SMRT resulting in transcriptional repression of RUNX1 targets such as *PTEN, BCL-2, ARF* and *CEBPA* (Zhang et al. 2001; Lam et al. 2012). However, full-length *RUNX1- RUNX1T1* (752 amino acids) is not sufficient for leukaemogenesis, and additional mutations are required for AML (Lam et al. 2012). The most frequent co-operating mutation with t(8;21) are c-KIT, which occurs in up to 50% of AML cases with t(8;21) (Reikvam et al. 2011; Wang et al. 2005) and has been associated with a very unfavourable prognosis and high relapse rate (Schnittger et al. 2011;

Care et al. 2003). Other co-operating genomic alterations in t(8;21) AML include mutations in *RAS* genes (17%), *FLT3-ITD* (5%), *FLT3 D853* (3%-7%), *JAK2 V617F* (6%-8%) and cytogenetic abnormalities such as loss of X chromosome in females (30%-40%), loss of Y chromosome in males (50-60%), deletion of part of chromosome 9 (del(9q)) (15%-35%) and trisomy 8 (8%). (Reikvam et al. 2011). The *RUNX1-RUNX1T1* fusion occurs in approximately 12% of AML and about 40% of AML-M2 subtype. *RUNX1* point mutations are more frequent (15-50%) in AML-M0 FAB type (Lam et al. 2012).

CCAAT Enhancer Binding Protein Alpha (*CEBPA*) is located at 19q13 and encodes a transcription factor that has a pivotal role in regulating the equilibrium between terminal differentiation and cell proliferation. It plays an important role in myeloid (granulocytic) differentiation by downregulating *c-MYC*, upregulating granulocytic factors and blocking monocytic lineage specific factors (Nerlov 2004; Renneville et al. 2008). *CEBPA* defective mice showed an absence of mature granulocytes with other lineage cells being unaffected (Zhang et al. 1997). Inactivation of *CEBPA* in AML can be through three mechanisms including (i) RUNX1-RUNX1T1 fusion transcriptional repression (ii) CEBPA translational inhibition by HNRNPE2 (heterogeneous nuclear ribonucleoprotein E2) induced by the BCR-ABL fusion and (iii) *CEPBA* point mutations leading to loss-of-function (Renneville et al. 2008). *CEBPA* mutations occur in approximately 10% of AML with a majority of them in M1 (19%), M2 (16%) and M4 (11%) subtypes of AML (Mannelli et al. 2017; Renneville et al. 2008)

The *Mixed Lineage Leukaemia (MLL)* gene is located at 11q23 and encodes a histone methyltransferase which is involved in epigenetic regulation of transcriptional activation through H3K4 methylation and haematopoietic differentiation (Krivtsov et al. 2007). *MLL* translocates and fuses with an array of partner genes such as *AF4, AF9, AF10, ENL* and *ELL*. *MLL* fusions occur in about 10% of AML, 30% of sAML and are more frequent (35-50%) in childhood AML (Prada-Arismendy et al. 2017; Winters et al. 2017). Another aberration in the *MLL* gene is the *MLL* partial tandem duplication (*MLL-PTD*) that results in duplication of exon 3-9 or exons 3-11. *MLL-PTD* occurs in 5-10% of AML and requires other co-occurring mutations to drive leukaemia (Sun et al. 2017; Choi et al. 2018).

Retinoic acid receptor alpha (RARA) gene is located at 17q21 and encodes a nuclear receptor that upon binding the retinoic acid ligand, leads to transcriptional activation of target genes involved in promyelocytic differentiation. Transcriptional activation is mediated by heterodimerisation of RARA protein with the retinoic X receptor alpha (RXRA) subunit, which can then bind retinoic acid response elements. In the absence of ATRA (all-trans retinoic acid)

or 9-cis retinoic acid ligands, the RAR-RXR complex binds and recruits transcriptional repressors such as NCoR/SMRT or polycomb repressive complex 1/2 (PRC1/2) that contain histone methyltransferases or histone deacetylases (HDAC) leading to transcription silencing. The binding of a ligand triggers the dissociation of repressor complex from RAR-RXR heterodimers to initiate transcription (Tomita et al. 2013). The most frequent aberration affecting *RARA* in AML is the translocation t(15;17), which results in the formation of the fusion protein PML-RARA. In this scenario, however, the physiological concentrations of ATRA become insufficient to trigger the release of co-repressors from RAR-RXR complex, leading to a shut-down of transcription and impaired differentiation. This translocation is a hallmark that is present in 98% of APL (De Braekeleer et al. 2014; Venci et al. 2017) and APL blast cells are sensitive to ATRA leading to differentiation of immature promyelocytes to mature granulocytes (Prada-Arismendy et al. 2017).

Class III mutations

Mutations in epigenetic regulators such as *TET2, DNMT3A, IDH1*, *IDH2* and *ASXL1*, encoding proteins that regulate downstream targets involved in proliferation and differentiation, are classified as the class III mutations. *TET2* mutations will be discussed in detail in Section 1.4.

DNMT3A is located at 2p23 and encodes for a DNA methyltransferase involved in the *de novo* methylation of DNA, others being *DNMT3B* and *DNMT1*. *DNMT3A* is mutated in 18-22% of AML and is more frequent (30-37%) in patients with normal karyotype AML. The most frequent mutation is an R882 substitution that leads to a missense mutation (Ley et al. 2010). Impaired methyltransferase activity and reduced DNA methylation levels result in upregulation of self-renewal potentiating genes, downregulation of differentiation factors and impaired differentiation of HSCs driving AML. *DNMT3A* mutations often co-occur with *NPM1*, *FLT3* and *IDH1* mutations (Prada-Arismendy et al. 2017).

Isocitrate dehydrogenases, IDH1 and IDH2 catalyse the formation of α-ketoglutarate (α-KG) from isocitrate. Mutations in *IDH1* are exclusively at position R132 (arginine residue at position 132) and mutations affecting *IDH2* are usually R140 or R172 substitutions (Ward et al. 2010). As such, these mutations affect the catalytic site of IDH1 and IDH2, resulting in the formation of the oncometabolite 2-hydroxyglutarate (2-HG) instead of α-KG. This has a tumorigenic effect as 2-HG functions as a competitive inhibitor of various tumour suppressors such as TET proteins that require α-KG as a co-factor for its DNA demethylating function (DiNardo et al. 2016). AML with mutations in *IDH1/2* have global DNA hypermethylation, impaired

differentiation and an upregulation of stem cell markers (Figueroa et al. 2010; Prada-Arismendy et al. 2017). Mutations in *IDH1* and *IDH2* occur in up to 14% and 19% of AML respectively (Megías-Vericat et al. 2019).

Additional sex-comb like-1 (ASXL1) is located at 20q11 and belongs to the *ASXL* family of genes that regulates homeobox (*HOX)* gene expression. ASXL1 interacts with polycomb repressors (PRC2) to establish an H3K27me3 (histone H3 at lysine 27 trimethylation) repressive histone mark and silences *HOXA* gene expression. However, *ASXL1* mutation in AML results in the upregulation of HOX proteins and skewed myeloid differentiation (Abdel-Wahab, Adli, et al. 2012). *Asxl1* knockout murine models showed impaired differentiation and development of myeloid malignancies. *ASXL1* mutations occur in 5-11% of AML, confer poor prognosis and often co-occur with *RUNX1* or *IDH2* mutations (Asada et al. 2019).

Other genomic abnormalities

In addition to the three classes of mutations, there are various chromosomal aberrations such as inv(16), trisomy 8 and abnormalities of chromosome 5 or 7 that drive leukaemogenesis.

Inv(16) which occurs in 5-8% of AML, is observed commonly in AML-M4EO and rarely in M0, M1, M2 and M5 subtypes (Lv et al. 2020). The pericentric inversion of chromosome 16, inv(16) (p13q22) forms the fusion gene *CBFβ-MYH11* that encodes for the fusion oncoprotein CBFβ-SMMHC (Schwind et al. 2013). There are three major domains in the fusion protein including the Runx binding domain (RBD) of CBFβ as well as the high affinity binding domain (HABD) and assembly competence domain (ACD) of SMMHC. The presence of both HABD and RBD domains increases the binding affinity of fusion protein to RUNX1 compared to wild-type CBFβ alone. Also, the ACD domain of the fusion protein is critical for leukaemia initiation as it binds to histone deacetylase HDAC8 which deacetylates and inactivates *TP53*, regulates cell cycle and the expression of CBF targets giving rise to inhibition of myeloid differentiation (Pulikkan et al. 2018; Kim et al. 2017; Qi et al. 2015).

Trisomy 8 (+8) is the most frequent chromosomal change in AML and is reported in 10-15% of AML and is reported as the sole cytogenetic abnormality in as many as 5% of cases (Paulsson et al. 2007; Hemsing et al. 2019). The likely explanation for the molecular mechanism behind leukaemogenesis in +8 is the gain of genes present on chromosome 8 as a result of the trisomy. There are approximately 800 genes on chromosome 8, including *MYC* and *RUNX1*, and microarray analysis showed that many are overexpressed in +8 AML (Virtaneva et al. 2001). +8 is also seen in myelodysplastic syndrome, chronic myeloproliferative diseases, acute

lymphoblastic leukaemia and even some solid tumours suggesting a potential association with neoplasia in other settings (Bakshi et al. 2012).

Cytogenetic abnormalities affecting chromosomes 5 and 7 include monosomy 5 (-5), monosomy 7 (-7) and deletion of the long arms of chromosomes 5 (5q-) and 7 (7q-). Chromosome 5 aberrations are critical drivers of leukaemogenesis because many of the genes located on the long arm are important regulators of haematopoiesis, such as the myeloid surface antigen *CD14*, CSF1 receptor (*CSF1R*), *CSF2*, early growth response 1 (*EGR-1*), *IL-4*, *IL-5*, *IL-9* and interferon regulator factor 1 (*IRF1*). The critical genes on chromosome 7 include *EPO*, *IL6*, plasminogen activator inhibitor (*PLANH1*), apoptosis regulator (*CASP2*), *HOXA3* and *HOXA4* (Mhawech et al. 2001; Schoch et al. 2005).

Based on cytogenetics and somatic mutations AML is classified into three major risk groups, which are favourable risk, poor risk and intermediate risk (Döhner et al. 2017). The favourable risk AML group includes those with alterations that include t(8;21), t(15;17), inv(16) and cytogenetically normal AML (CN-AML) with *NPM1* mutation. Favourable risk AML has the highest complete remission rate of 90% (Döhner et al. 2017). Poor risk AML, also known as adverse risk, includes abnormalities such as t(6;9), inv(3), del(5q), del(7q) or AML with complex karyotype and CN-AML with *FLT3-ITD* mutation. Poor risk AML is associated with chemoresistance, high relapse rate and low event free survival (Döhner et al. 2017). Intermediate risk AML includes CN-AML, which accounts for the largest category of AML patients (Döhner et al. 2017). This excludes some specific CN-AML patients defined as having favourable and poor risk disease. In addition, those patients with a t(9;11) translocation are also classified in the intermediate risk group (De Kouchkovsky et al. 2016; Estey 2014; Mrozek et al. 2012; Dohner et al. 2010) Therefore, an understanding of underlying cytogenetic features and mutations play an important role in predicting the treatment outcome and developing personalised treatment approaches.

1.2.6. Current state of play in AML research and the role of murine models

Our understanding of the biology and therapy of AML has been improving rapidly drive by progress in research methodologies. Technologies such as next generation sequencing and FISH have paved the way for the detection and characterisation of various novel and recurrent cytogenetic abnormalities in AML. For example, through a comprehensive NGS study on a cohort of 1540 AML patients, Papaemmanuil and colleagues identified 5234 AML driver mutations across 76 genes, among which two or more of these driver mutations were seen in

up to 86% of AML patients (Papaemmanuil et al. 2016). Despite progress in modern technologies, it has been challenging to functionally characterise the role of genetic mutations in the pathogenesis of AML. Disease heterogeneity further complicates the study of a common mechanism driving AML (Almosailleakh et al. 2019). Nevertheless, functional characterisation of somatic genetic driver mutations has been aided by the development of *in vivo* murine models.

Murine models in AML research include carcinogen (chemically, irradiation or viral-infection) induced tumours in mice, transgenic mice expressing AML-associated proto-oncogenes and patient-derived xenotransplant models. Many of the major chemotherapeutic agents, including cytarabine, were initially tested for efficacy in AML using the L1210 cell line isolated from DBA/2 mice after exposure to the carcinogen 3-methylcholantrene (Almosailleakh et al. 2019). In another study, irradiation in mouse models were shown to accelerate the development of leukaemia, mediated by expansion of ETV6-RUNX1 fusion protein expressing cells (Li et al. 2013). Several transgenic AML mouse models have since been generated, whereby foreign genetic material was introduced into the germline of mice either by direct injection into mouse oocytes or by electroporation into mouse embryonic stem cells (Almosailleakh et al. 2019). This strategy was used to establish transgenic models for AML carrying the *PML-RARA* fusion gene, which demonstrated a role for the encoded oncoprotein in APL pathogenesis and the efficacy of ATRA/arsenic trioxide-based treatments (Brown et al. 1997; de Thé et al. 2010; Grisolano et al. 1997; He et al. 1997; Rego et al. 2000). Furthermore, conditional gain-of-function mouse models of AML driver fusion proteins were developed using integration of spatially and temporally controllable promoters such as *Vav1-iCre*, which facilitate sensitive control of transgene expression (Speck et al. 2009; Almosailleakh et al. 2019). Transgenic mouse models were further improved with the introduction of expression cassettes that are sensitive to chemical inducers such as tetracycline (Tet) or doxycycline. These strategies use on an on-off mechanism whereby the expression of the transgene can be induced or inhibited as necessary (Almosailleakh et al. 2019; Furth et al. 1994; Gossen et al. 1995). The Tet-based expression system was successfully exploited to bypass the embryonic lethality associated with the constitutive expression of the *RUNX1-RUNXITI* fusion oncogene (Rhoades et al. 2000; Higuchi et al. 2002).

All these mouse models have been vital in our understanding of leukaemogenesis and the role played by individual genes as well as co-operating mutations. However, many of these mouse models could not fully recapitulate human disease, primarily due to the complex interactions between leukaemic cells and the bone marrow microenvironment and the differences between

the mouse and human bone marrow microenvironment (Almosailleakh et al. 2019). This drawback was partially overcome via the generation of xenograft mouse models whereby human AML cells were directly transplanted into immunocompromised mouse, including severe combined immunodeficient (SCID), non-obese diabetic (NOD) and NOD-SCID mice (Almosailleakh et al. 2019; Theocharides et al. 2016). Several patient-derived xenograft (PDX) models have also been established, including several used to study the role of *TET2* mutations in clonal evolution of AML (Kawashima et al. 2022; Morita et al. 2020).

The efficacy and safety of anti-leukaemic therapies have also been tested using animal models. For example, cytarabine/doxorubicin combination therapy was tested using PDX models derived by transplantation of human MLL-AF9+ leukaemia in immunocompromised mice (Saland et al. 2015; Wunderlich et al. 2013; Almosailleakh et al. 2019). Many other chemotherapeutic agents that have been tested *in vitro* failed to demonstrate efficacy in clinical trials, which was likely due to the lack of a bone marrow environment (Almosailleakh et al. 2019). In this regard, PDX models have proved to be an excellent experimental platform to test the efficacy of chemotherapeutic agents.

1.2.7. Diagnosis of AML

AML is diagnosed when blasts constitute at least 20% of cells in the bone marrow, determined morphologically using a marrow smear (Arber et al. 2016) and a blast percentage of 5%-20% is generally categorised as MDS (Arber et al. 2016). Immunophenotyping is used to confirm or refine a diagnosis with common AML markers including CD33 and CD34, present on approximately 64%-98% and 45-68% of observed cases with AML, respectively (Ortolani 2011). Conventional cytogenetics is also used to confirm or refine an AML diagnosis via the use of fluorescence in situ hybridisation (FISH) to identify chromosomal abnormalities such as gene rearrangements (Dohner et al. 2010). The presence of fusion genes is also often confirmed using RT-PCR and other molecular techniques.

1.2.8. Treatment of AML

Based on the aim of the therapy, AML treatment can be broadly classified into three phases — Remission induction therapy, consolidation therapy and palliation.

Remission Induction chemotherapy

Conventional remission induction therapy which is aimed at destruction of maximum number of cancer cells, is based on a “7+3 regimen” which includes cytarabine for the first seven days

along with daunorubicin for the first three days (Dombret et al. 2016). Cytarabine (1-β -D-arabinofuranosylcytosine; Ara-C) is a nucleoside analogue that exerts its pharmacological effect by impairing DNA synthesis when cells are in the S-phase, through inhibition of DNA polymerase and termination of DNA elongation. It can also function by blocking the cell-cycle progression from G1 to S phase (Momparler 2013). Cytarabine treatment leads to extensive chromosomal damage and cytotoxicity in leukaemia cells (Momparler 2013; Xiaoguang et al. 2018). Daunorubicin (DNR) is an anthracycline that exerts its anti-mitotic and cytotoxic effect by intercalating between DNA base pairs and inhibiting topoisomerase II and polymerase activity resulting in single and double strand breaks (Murphy et al. 2017; Momparler 2013).

Induction therapy is usually given to intermediate and favourable risk groups with a low risk of treatment-related mortality. This mode of treatment induces complete remission in 65%-73% of young patients with AML (Saultz et al. 2016). However, in patients over 60 years of age, only 38%-62% achieve complete remission following standard remission induction therapy (Saultz et al. 2016). Generally, 14 days after the start of induction therapy, treatment response is evaluated using a bone-marrow aspirate or biopsy. At this point, cytological evidence of residual disease is observed in as many as 50% of patients, to whom additional treatment is given, including another cycle of "7+3 regimen", high dose single-agent Ara-C alone or FLAG-IDA. FLAG-IDA is a combination of fludarabine (a purine analogue that can increase the uptake of Ara-C), Ara-C, Granulocyte colony-stimulating factor (G-CSF) and idarubicin (an anthracycline) which is usually given at relapse (De Kouchkovsky et al. 2016).

Consolidation therapy using targeted therapeutic measures

Induction therapy is followed by consolidation therapy (often termed post-induction therapy) which is administered with an aim to prevent or reduce the risk of relapse. Targeted therapies and haematopoietic stem cell transplantation are two strategies used either alone or in combination during consolidation therapy (Dombret et al. 2016).

Given the very high frequency and prognostic impact of *FLT3* mutations in AML, FLT3 inhibitors have emerged as a viable treatment approach. Sorafenib and midostaurin are FLT3 inhibitors that targets multiple tyrosine kinases including FLT3. In a phase II trial, combination of sorafenib with a hypomethylating agent (azacytidine) achieved an overall response rate (ORR) of 46% with 16% CR (Ravandi et al. 2013). In the phase III RATIFY trial for treatment naïve *FLT3* mutated (ITD and TKD) AML (n=717), midostaurin vs placebo was administered as maintenance therapy for a year after "7+3 regimen" induction therapy and high dose Ara-C consolidation therapy. Midostaurin-treated patients had a significantly higher OS and EFS

compared to placebo-treated, although the CR rates were similar (Stone et al. 2012). However, sorafenib and midostaurin lack FLT3 selectivity as they target other kinases such as c-KIT, PDGFR and VEGFR. More selective second-generation FLT3 inhibitors include quizartinib, crenolanib and gilteritinib (Bose et al. 2017; De Kouchkovsky et al. 2016). In a phase II study, on relapse/refractory AML patients aged >60 years, quizartinib as single-agent yielded an ORR of 66% and a CR rate of 54% in *FLT3-ITD*+ patients (Cortes et al. 2012). A similar phase II study of crenolanib for relapse/refractory *FLT3* mutated AML resulted in an ORR of 50% and a CR rate of 39% (Bose et al. 2017). Likewise, a phase II study of gilteritinib on relapse/refractory *FLT3* mutated AML had an ORR of 52% (Altman et al. 2015).

Other major targeted therapeutic agents include venetoclax (BCL2 inhibitor), navitoclax (BCL-XL inhibitor), idasanutlin (MDM2 inhibitor), ivosidenib (IDH1 inhibitor), enasidenib (IDH2 inhibitor) and OPB31121 (STAT3/5 inhibitor) (Bose et al. 2017; De Kouchkovsky et al. 2016). Alternatively, there are antibody drug conjugates (ADC) such as Gemtuzumab ozogamicin (GO) that specifically targets CD33, a surface marker expressed in myeloid lineage cells (De Kouchkovsky et al. 2016). Cytotoxic response is driven by the DNA-cleaving antibiotic drug calicheamicin conjugated to the antibody (De Kouchkovsky et al. 2016). This ADC approach adds an extra layer of specificity as only CD33+ cells internalise the cytotoxic drug. Despite initial trial failures due to safety concerns, many new phase III trials showed an increased overall survival (primarily in patients with favourable and intermediate cytogenetics) and no toxicity (Hills et al. 2014; De Kouchkovsky et al. 2016). Developing novel targeted therapeutic agents is a domain that has been constantly evolving with many approved efficacious drugs and many others in clinical trial phases.

Palliation using hypomethylating agents

Hypomethylating agents such as azacytidine and decitabine that inhibits DNA methylation and activate epigenetically silenced tumour suppressors are now established as a therapeutic strategy for AML. These drugs are often used as palliative care for those patients who are not fit for intensive chemotherapy. One of the major hypomethylating drug azacytidine is discussed in detail in section 1.5.

Decitabine (5-aza-2'-deoxycytidine) is a pyrimidine nucleoside analogue, which functions by inhibiting DNA methyltransferases (DNMTs), resulting in direct or indirect activation of gene function. Some genes can be activated by demethylation of the gene itself, which is a direct effect. In contrast, the effect is considered indirect when an upstream target of the gene is

demethylated leading to activation of the entire pathway. Upon cellular uptake of the drug, decitabine gets phosphorylated into a triphosphate derivative which is incorporated into newly synthesised DNA (He et al. 2017; Seelan et al. 2018; Stresemann et al. 2008). The mechanism of action of hypomethylating agents is detailed in section 1.5.1.

A randomised phase III trial (NCT00260832) for elderly AML demonstrated better response rates with decitabine single agent in terms of increased median overall survival and complete response rates compared to low dose cytarabine or supportive care (Kantarjian et al. 2012). Decitabine is currently used in many clinical trials in combination with other targeted therapies to evaluate its potential in AML treatment. A drawback of hypomethylating agents is their instability in plasma due to the activity of cytidine deaminases (CDA) (Roboz et al. 2018). Guadecitabine is a newly developed hypomethylating agent, which is a dinucleotide of decitabine and deoxyguanine connected by a phosphodiester bond. Guadecitabine is not a substrate for CDA and with the gradual release of decitabine through enzymatic cleavage of the phosphodiester bond, the relative exposure to active decitabine is increased with the use of guadecitabine (Roboz et al. 2018). A phase II study with guadecitabine (NCT01261312) revealed better outcomes in terms of higher complete response rates and reduced mortality in relapsed or refractory AML (Garcia-Manero et al. 2019). Various other clinical trials are ongoing to better understand the safety and potential of this drug in the treatment of AML (Roboz et al. 2018).

1.3. DNA methylation defects in AML

Mutations in epigenetic modifiers are one of the early events of leukaemogenesis and DNA-methylation related genes are mutated in up to 44% of AML (Ley et al. 2013). Aberrant methylation leading to promoter hypermethylation and silencing of tumour suppressor genes is considered as a hallmark of haematological disorders including AML. Section 1.3 provides an overview of the mechanism of DNA methylation, the genes involved and the basic function, which is key to aetiology, prognosis, and treatment of AML.

1.3.1. Epigenetics

Epigenetics is the branch of biology which deals with the study of heritable changes (inherited from parent cell to daughter cells through mitosis) in gene expression that are not associated with changes in the genetic code itself (Dupont et al. 2009). In a multicellular organism, a variety of epigenetic mechanisms function to regulate gene expression and give rise to cellular heterogeneity in gene expression, despite all the cells having the same primary DNA sequence

(Moore et al. 2013). The mechanisms used to regulate gene expression epigenetically include DNA methylation, histone modification and RNA associated silencing (Egger et al. 2004).

1.3.2. Mechanism of DNA methylation

Gene expression in mammalian cells is regulated at the transcriptional level via control of genomic methylation, including primarily methylation at 5' CpG 3' sites (cytosine phosphate-link guanine dinucleotides) but also at non-CpG sites. Whereas the role of non-CpG methylation remains largely unknown, methylation at the carbon 5 (C5) position of cytosine residues of CpG dinucleotides serves as the major heritable epigenetic mark (Moore et al. 2013)

DNA methylation is initiated by the action of DNA methyltransferases, referred to as "writers," which transfer a methyl group onto unmethylated cytosine residues. Methylated cytosine can then be recognised by the "readers," a class of enzymes which recruit various transcriptional repressor complexes to block the transcription of methylated DNA. Finally, the third class of enzymes, termed "erasers," function in the demethylation of DNA in which methylated cytosine in converted back into unmethylated cytosine. The detailed mechanisms through which writers, readers and erasers regulate DNA methylation is as follows:

(i) Writers constitute a family of DNA methyltransferases (DNMT1, DNMT3A and DNMT3B) that catalyse the addition of a methyl group to either a completely unmodified CpG residue (*de novo* methylation) or hemimethylated DNA, where the cytosine on one strand is methylated. During the methylation process, a methyl group is transferred from S-adenyl methionine (SAM) to the C5 of a cytosine residue to form 5-methylcytosine (5-mC). DNMT3A and DNMT3B transfers methyl group to unmodified DNA to establish a new methylation pattern and hence they are known as *de novo* methyltransferases (Figure 1.3) (Moore et al. 2013; Okano et al. 1999). In contrast, DNMT1 copies an existing methylation pattern from the parental DNA strand to the daughter strand post-replication and is defined as a maintenance methyltransferase (Figure 1.3) (Moore et al. 2013; Pradhan et al. 1999; Ramsahoye et al. 2000)

(ii) Readers play a role in recognising DNA methylation and recruiting transcription repressor complexes into methylated DNA to finally silence the gene itself. They consist of three families of proteins: methyl CpG binding domain proteins (MBD), ubiquitin-like, containing PHD and RING finger domain proteins (UHRF) and zinc-finger domain proteins (Moore et al. 2013). MBD proteins also possess a

transcriptional repression domain and can bind to different repressor complexes and recruit them to methylated DNA resulting in transcriptional repression (Moore et al. 2013; Nan et al. 1998; Ng et al. 1999; Sarraf et al. 2004). MeCP2 (a member of the MBD family) has the unique role of binding to DNMT1 through its transcriptional repression domain, recruiting DNMT1 to hemimethylated DNA to maintain DNA methylation (Kimura et al. 2003; Moore et al. 2013). In a similar mechanism, UHRF proteins also function in maintaining DNA methylation by recruiting DNMT1 (Achour et al. 2008; Bostick et al. 2007; Sharif et al. 2007). Zinc finger domain proteins function via a similar mechanism to MBD proteins and repress transcription (Filion et al. 2006; Lopes et al. 2008; Moore et al. 2013; Prokhortchouk et al. 2001; Yoon et al. 2003)

(iii) Erasers are DNA demethylating proteins that perform either active or passive DNA demethylation. Passive demethylation occurs when DNMT1 is either inhibited or dysfunctional and can no longer promote methylation in the daughter strand during replication. The daughter strand remains unmethylated and global methylation levels reduce with each successive cell division (Moore et al. 2013). Active DNA demethylation is mediated via a series of chemical reactions including deamination and/or oxidation reactions with the help of TET (Ten-eleven translocation methylcytosine dioxygenases) proteins and base excision repair pathway enzymes (Figure 1.4). Furthermore, 5-mC is sequentially modified to an intermediate product which can be targeted by the base excision repair (BER) pathway and the modified base is replaced with an unmodified cytosine. 5-mC can be modified at two positions: the methyl group at carbon 5 (C5) and the amine group at carbon 4 (C4). AID/APOBEC (activation-induced cytidine deaminase/ apolipoprotein B mRNA-editing enzyme complex) deaminates the amine group at C4 of 5-mC to a carbonyl group forming thymine. This G: T mismatch is further targeted by the BER pathway and converted to unmodified cytosine (Moore et al. 2013) (Figure 1.4). Another proposed mechanism of demethylation is the modification of the methyl group by the TET enzymes in 5-mC to 5-hydroxymethyl cytosine (5-hmC) followed by a series of oxidation reactions in which 5-hmC is converted to 5-formylcytosine (5-fC), which is subsequently converted to 5-carboxylcytosine (5-caC). 5-caC activates the base excision repair initiated by thymine DNA glycosylase (TDG), which converts 5-caC to cytosine, and hence the demethylation process is complete (Ito et al. 2011; Kaelin et al. ; Moore et al. 2013; Solary et al. 2014). Each of the intermediate products is recognized and repaired by the BER pathway resulting in

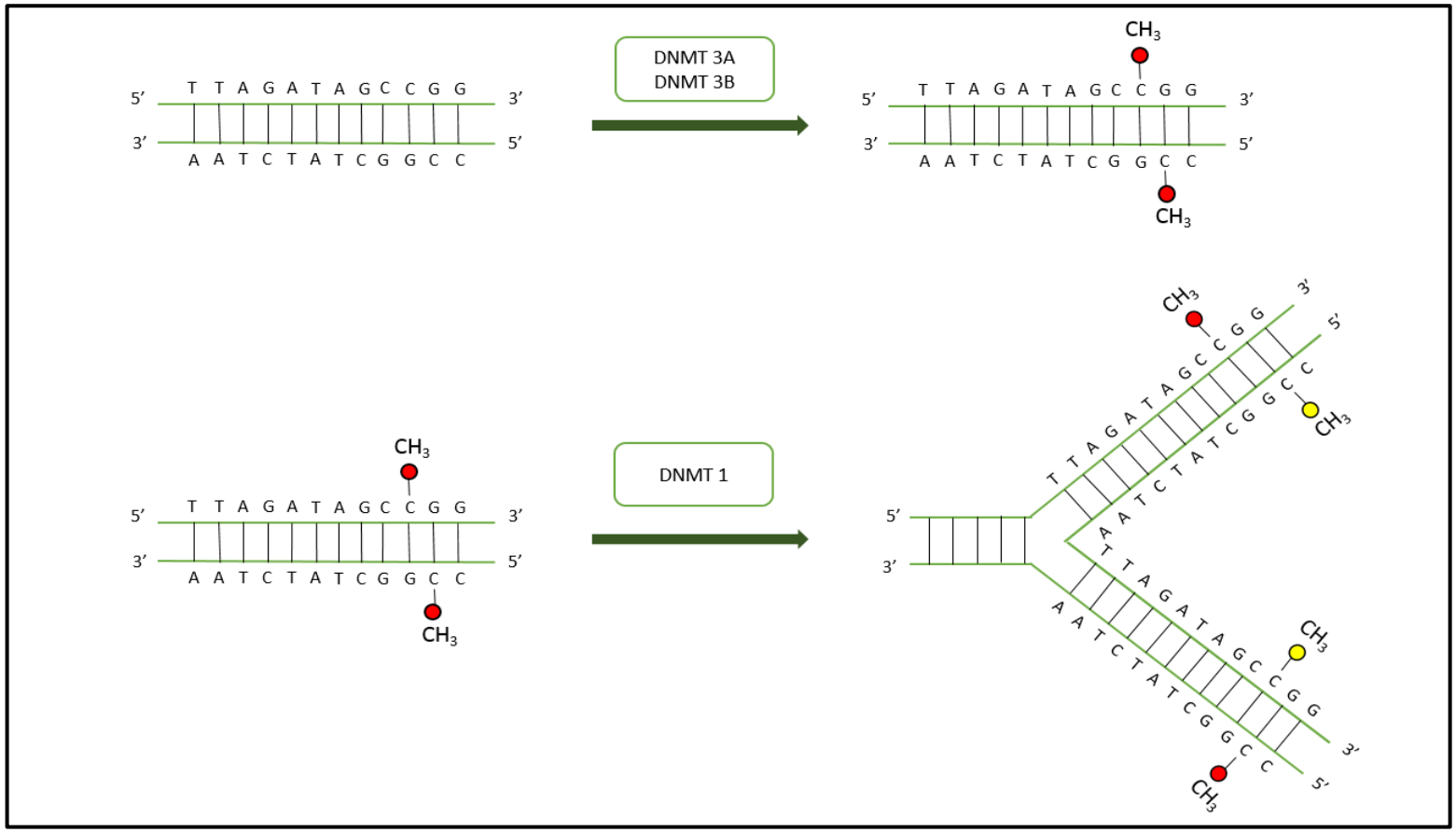

Figure 1.3. Role of DNA methyltransferases in the methylation of DNA. DNMT3A and DNMT3B (also known as *de novo* methyltransferases) transfer methyl groups (red) from S-adenosylmethionine to unmodified DNA. During replication DNMT1 (defined as a maintenance methyltransferase) copies the methyl group (yellow) from hemimethylated parental DNA into the daughter strand (modified from (Moore et al. 2013).

CH_3, methyl group; DNMT; DNA methyltransferase

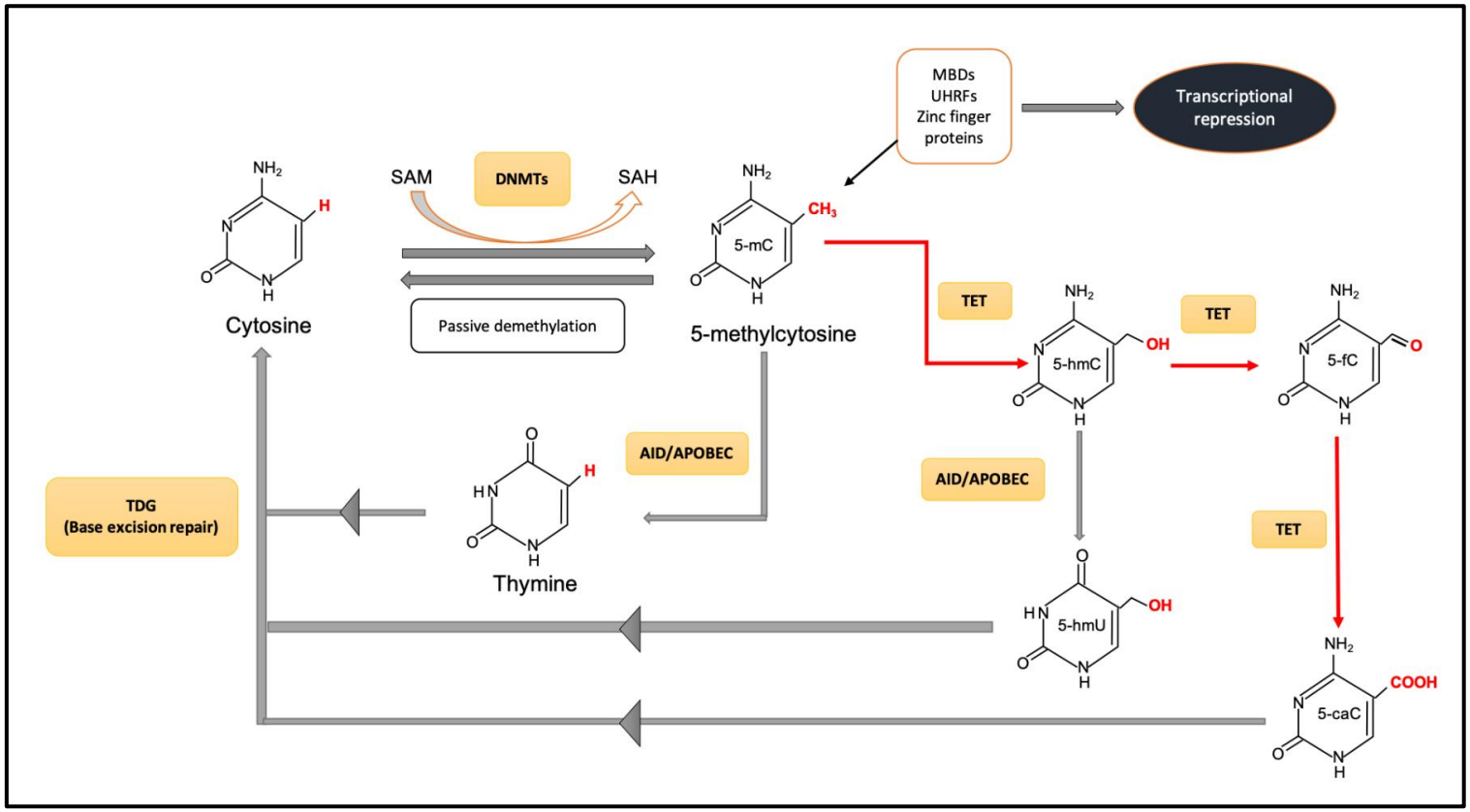

Figure 1.4. Mechanism of epigenetic regulation. During methylation, DNMTs transfer a methyl group from SAM to the C5 of cytosine to form 5-mC. 5-mC is recognised by various readers like MBD, UHRF and zinc finger proteins which targets the methylated DNA to transcriptional repression. There are three different pathways for active DNA demethylation. Primarily, 5-mC is converted to thymine by AID/ APOBEC and the resulting G: T mismatch is recognised and repaired by TDG in BER pathway. Second, 5-mC is converted to 5-hmC by TET proteins which is further converted to 5-hmU by AID/APOBEC and targeted for base excision repair. Third, TET proteins convert 5-mC in a sequential pattern to form 5-hmC, 5-fC and 5-caC and the resulting 5-caC is recognised and targeted by the BER pathway. All the three processes result in the restoration of cytosine.

5-caC, 5-carboxylcytosine; 5-hmC, 5-hydroxymethylcytosine; 5-hmU; 5-hydroxymethyl uracil; 5-fC, 5-formylcytosine; 5-mC, 5-methylcytosine; AID/APOBEC, activation-induced cytidine deaminase/ apolipoprotein B mRNA-editing enzyme complex; BER, Base excision repair; DNMT, DNA methyl transferase; MBD, methyl DNA binding domain protein; SAH S-adenosyl homocysteine; SAM, S-adenosylmethionine; TET, Ten-eleven translocation TDG, Thymine DNA Glycosylase; UHRF, ubiquitin-like, containing PHD and RING finger domain

the insertion of unmodified cytosine (Figure 1.4) (Bhutani et al. 2011; Moore et al. 2013).

1.3.3. Basic functions of DNA methylation

The basic functions of DNA methylation include silencing retroviral elements, genomic imprinting, regulating tissue-specific gene expression and X-chromosome inactivation.

Methylation can occur in different positions including intergenic regions, transcriptional start sites, promoters, gene bodies and repeat regions such as centromeres. Depending on the location of methylation, its effect on gene regulation also varies (Moore et al. 2013). Transposable and viral elements that account for approximately 45% of the mammalian genome are inactivated by DNA methylation (Schulz et al. 2006). If these elements are demethylated/expressed, they can replicate and insert into other parts of the genome causing mutation and gene disruption (Gwynn et al. 1998; Kuster et al. 1997; Michaud et al. 1994; Moore et al. 2013; Ukai et al. 2003; Wu et al. 1997). As such, an important role of DNA methylation in intergenic regions is to silence potentially harmful elements.

Likewise, methylation of gene promoters is a major determinant of gene expression status. Approximately 70% of the gene promoters include CpG islands, defined as those regions, typically between 500 and 1500 bp in length, with higher-than-average CpG density compared to the rest of the genome (Moore et al. 2013; Pfeifer 2018; Saxonov et al. 2006). Many transcription factors have consensus binding sites that are GC rich. In general, methylation of CpG islands in promoters leads to impaired transcription factor binding and silencing of gene expression (Moore et al. 2013).

Methylation within the gene body can also impact on gene expression, although the effects are gene specific and less well understood than promoter methylation. In dividing cells, gene body methylation can increase gene expression whereas in slowly dividing or non-dividing cells gene body methylation does not (Aran et al. 2011; Ball et al. 2009; Guo et al. 2011; Hellman et al. 2007; Moore et al. 2013; Xie et al. 2012).

In general, hypermethylation of CpG islands leads to silencing of gene expression, whereas DNA demethylation/hypomethylation often leads to increased gene expression, although the effect of methylation is CpG specific and gene specific.

1.4. Ten-eleven translocation (TET) methylcytosine dioxygenase 2 (*TET2*)

The TET family of dioxygenases comprises an important family of epigenetic modifiers having a role in haematopoiesis. The TET family consists of three members namely *TET1, TET2* and *TET3*. The name TET is derived from the rare "ten-eleven translocation" that occurs when the *MLL* gene in chromosome 11 fuses with the *TET1* gene in chromosome 10 (Solary et al. 2014). The chromosomal translocation t(10;11) (q22;q23) is reported in both lymphoid and myeloid malignancies (An et al. 2017).

The basic function of TET enzymes is demethylation of DNA through a sequential oxidation process (Figure 1.4). Despite having similar structural domains and functionality, these proteins exhibit very distinct pattern of expression such that TET1 is primarily expressed in embryonic stages whereas TET2 and TET3 are more ubiquitous in nature with TET2 being predominantly expressed in haematopoietic and neuronal lineages (Solary et al. 2014). Among the various isoforms of TET proteins, TET2 has a unique role in the differentiation of haematopoietic stem cells and is frequently mutated in haematological malignancies (Melamed et al. 2018; Moran-Crusio et al. 2011). Silencing of *TET2* using shRNA impairs differentiation of haematopoietic precursors (Ko et al. 2013).

1.4.1. An insight into the structural features and mechanism of TET2

TET2 is located at chromosomal locus 4q24. A schematic representation of the structures of TET proteins- TET1, TET2 and TET3 are shown in Figure 1.5.

TET2 encodes for a protein 2002 amino acids in length that consists of a catalytic core made up of a double-stranded β-helix (DSBH) domain and a cysteine rich domain at its C-terminal (Solary et al. 2014) (Figure 1.5A). The DSBH domain (also known as jelly roll fold) consists of catalytic motifs such as HxD motifs (2-His-1-carboxylate triad; two histidine and one aspartate/glutamate) where Fe(II) binds. This is followed by the binding of α-Ketoglutarate (α-KG) on the other side of DSBH domain and oxidation of Fe(II) by molecular oxygen. This further trigger the decarboxylation of α-KG into succinate and results in substrate oxidation (Solary et al. 2014; Hu et al. 2013; Shen et al. 2014; An et al. 2017). A large low complexity insert is present within the DSBH (Figure 1.5A), of which the function is unknown. As such, a large number of post-translational modifications affecting TET2 predominantly occurs at this low-complexity region and therefore this region potentially serves as the regulatory region of TET2 (Bauer et al. 2015).

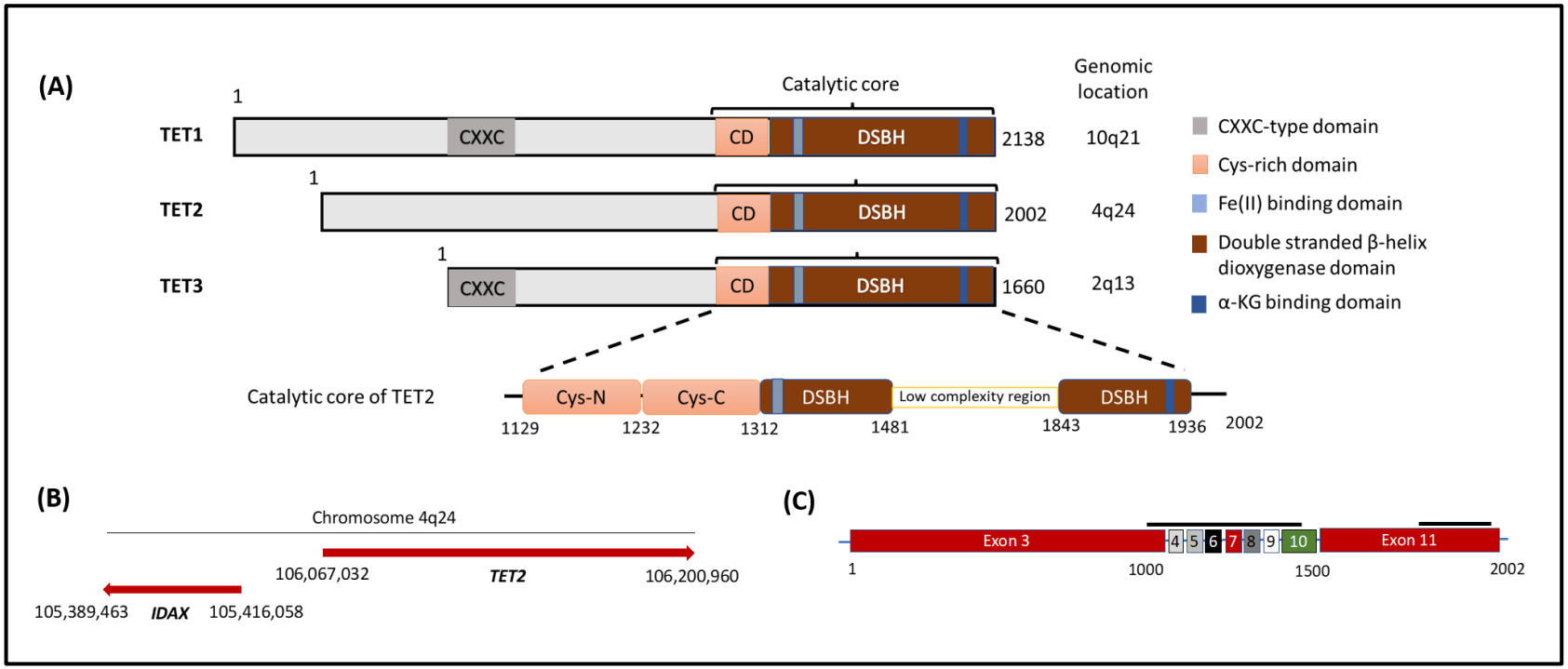

Figure 1.5. Schematic representation of TET2 structure and domain architecture (A) The three members of TET family of dioxygenases (TET1, TET2 and TET3) share close structural similarities. Unlike other TET family members, TET2 lacks the CXXC domain (dark grey boxes), which can autoregulate TET protein function. The catalytic core is evolutionarily conserved in all three TET family members. The expanded view of catalytic core shows the major functional domains of TET2. The catalytic core is comprised of a cysteine rich domain that has Cys-N and Cys-C subdomains (orange) and the double stranded β-helix (DSBH) dioxygenase domain (brown) at the carboxy terminal of the protein that contains binding sites for Fe (II) and α-ketoglutarate (α-KG) (dark and light blue boxes respectively), which are essential for the oxidation function of TET proteins. There is also a low complexity insert region (white box) within the catalytic core that serves as a site for post translational modifications. The numbering shows amino acid positions. (B) The region from chromosome 4 showing nucleotide locations of *TET2* and *IDAX* genes (red arrows) transcribing in opposite directions. The CXXC domain containing IDAX negatively regulates TET2. (C) The coding exons of TET2 and the respective amino acid positions are shown. The locations corresponding to the two conserved domains are depicted through black bars (Hu et al. 2013; Weissmann et al. 2012; Nakajima et al. 2014)

α-KG, α-ketoglutarate; CD, Cys-rich domain; CXXC, CXXC type domain; DSBH, double stranded β-helix dioxygenase domain.

The cysteine rich domain is comprised of two subdomains Cys-N and Cys-C, which wraps around the DSBH core and functions in regulating chromatin targeting. The globular structure of TET2 is stabilised by the coordination of three zinc cations by residues from cysteine rich domain and DSBH, which results in bringing the two domains together into close proximity (Hu et al. 2013) (Figure 1.5A). The region between amino acid residues 1129 and 1936 (Figure 1.5A) forms the catalytically active fragment and truncation outside this region has minor impact on TET2 activity (Bowman et al. 2017). Nevertheless, full length TET2 possesses the highest enzymatic activity compared to any truncated protein. As such, the catalytic core region is evolutionarily conserved through all three TET proteins (Hu et al. 2013).

However, unlike TET1 and TET3, TET2 lacks the zinc finger CXXC domain that consists of two cysteine rich clusters that can bind and coordinate zinc ions. The gene segment encoding the CXXC domain was lost from *TET2* and formed an independent CXXC domain-containing gene known as *IDAX* (also known as *CXXC4*), which is transcribed in the opposite direction to *TET2* (Figure 1.5B). The CXXC domain in TET1, TET3 and IDAX is comprised of a zinc finger motif and functions as a DNA binding domain. This domain shows preferential binding to unmethylated CpG rich regions and exerts a regulatory role over *TET2* (Ko et al. 2013; Nakajima et al. 2014).

Despite having close structural and functional similarities, there is no evidence of *TET1* and *TET3* compensating functionally for loss of *TET2*. Moreover, *Tet1* and *Tet3* mRNA levels were not altered in response to a deletion of *Tet2* in mice and *TET2* mutated AML patients had significantly lower levels of 5-hmC levels (Kroeze et al. 2014; Moran-Crusio et al. 2011; Inoue et al. 2016).

1.4.2. Functions of TET2

Demethylation of DNA

As discussed in Section 1.3.2 and Figure 1.4, the primary role of TET2 is demethylation of DNA and regulation of gene expression. TET2 catalyses these conversions through sequential oxidation reactions in an α-KG and Fe(II)-dependent manner. Each of the intermediate products such as 5-hmC, 5-fC and 5-caC act as stable epigenetic marks with distinct roles in regulation of gene expression by recruiting specific readers, transcription factors and repair proteins (Rasmussen et al. 2016). 5-hmC inhibits binding of transcriptional repressors to gene promoters. TET2 also facilitates passive demethylation of DNA by gradually substituting the

pool of 5-mC with 5-hmC, which is a poor substrate for DNMT1 (Valinluck et al. 2004; Shi et al. 2017; Moore et al. 2013). Therefore, with each replication cycle, the 5-mC mark is essentially diluted and the increasing levels of 5-hmC affects the functionality of DNMT1 (Rasmussen et al. 2016). As such, TET2 plays an important role in fine-tuning the very dynamic epigenetic state of a cell.

Regulates histone modifications through O-GlcNAcylation

O-GlcNAcylation is a post-translational modification that involves addition of an O-linked N-acetylglucosamine (O-GlcNAc) group to serine and threonine residues of cellular proteins. This is a cyclic process mediated by two enzymes namely O-GlcNAc transferase (OGT) and O-GlcNAcase (OGA) in which OGT catalyses the addition of an O-GlcNAc group through glycosyltransferase reaction and OGA catalyses the removal of the group from Ser and Thr residues (Yang and Qian 2017).

The implications of this process and the specific role played by TET2 can be explained by breaking the concept into three sequential events. (a) The DSBH domain of TET2 interacts with and recruit OGT to chromatin and promoter regions. This promotes OGT activity and O-GlcNAcylation on proteins such as HCF1 (Solary et al. 2014). (b) The GlcNAcylated HCF1 recruits SET1/COMPASS complex (SETD1A) and as such it plays an important role in maintaining the integrity of the complex (Solary et al. 2014). (c) SETD1A is a H3K4 methyltransferase, which causes trimethylation mark at H3K4 (H3K4me3) and results in transcriptional activation (Solary et al. 2014). TET2-OGT interaction supports SETD1A complex formation as well as the binding of the complex on chromatin (Figure 1.6B).

Tet2-/- mice showed reduced global GlcNAcylation and H3K4me3 levels (Deplus et al. 2013). In a similar way, TET2-OGT interaction activates Histone 2B O-GlcNAcylation at Ser112 and the associated genes show high level of transcription (Chen et al. 2013; Solary et al. 2014). In addition, the functions of O-GlcNAcylation include interaction with histone modifying machineries involved in acetylation, methylation and ubiquitylation, regulation of polycomb and trithorax complexes and modifications of RNA polymerase II. O-GlcNAcylation of RNA pol II occurs at the threonine 4 of the C-terminal domain (CTD) and this prevents CTD phosphorylation resulting in inhibition of RNA pol II activation and elongation (Vella et al. 2013). The major functions of TET2 are summarised in Figure 1.6.

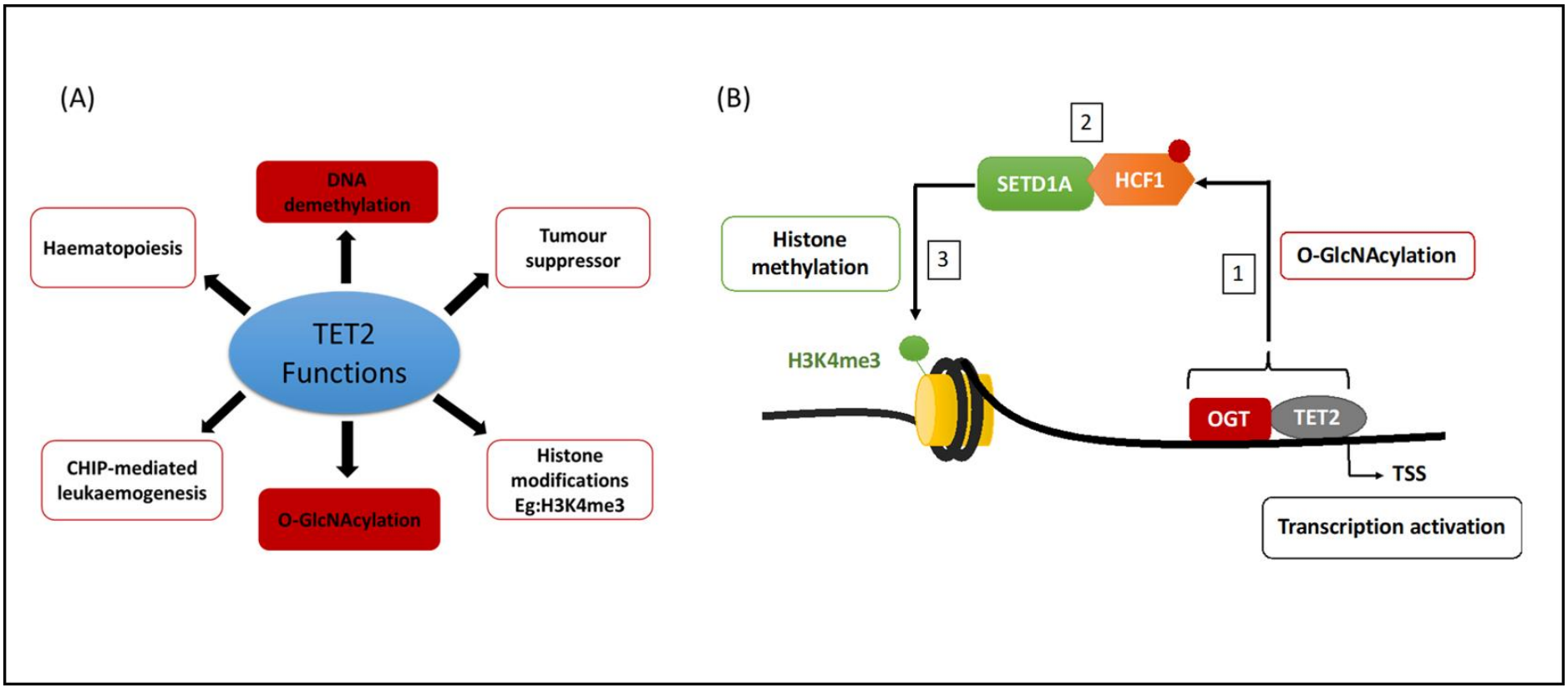

Figure 1.6. Major functions of TET2. (A)The basic functions of TET2 are DNA demethylation and O-GlcNAcylation (red boxes). The demethylation function of TET2 results in activation of various silenced genes involved in haematopoiesis. Mutations in *TET2* result in clonal haematopoiesis of indeterminate potential (CHIP)-mediated leukaemogenesis and hence is recognized as a CHIP-driver gene. The O-GlcNAcylation function of TET2 results in histone modification which often acts as a transcriptional activation mark. (B) During TET2-mediated histone modification the interaction of TET2 with OGT triggers the process of O-GlcNAcylation of HCF1. In the second step, the GlcNAcylated (red circle) HCF1 recruits and maintains the integrity of SETD1A complex which is a histone methyl transferase. The third and final step is the SETD1A mediated H3K4 trimethylation (H3K4me3) histone modification (green circle) which serves as a transcriptional activation mark.

CHIP, Clonal haematopoiesis of indeterminate potential; H3K4me3, Histone 3 Lysine 4 trimethylation mark; HCF1, Host Cell Factor C1; O-GlcNAc, O-linked N-acetylglucosamine; OGT, O-GlcNAc transferase; SETD1A, SET Domain Containing 1A Histone Lysine Methyltransferase; TET2, Tet Methylcytosine Dioxygenase 2; TSS, Transcription start site

1.4.3. Clinical implications of TET2 functions

TET2 in haematopoiesis and haematopoietic malignancies

The basic demethylating function of TET2 has implications over multiple physiological processes such as haematopoiesis and perturbation of TET2 activity often results in leukaemogenesis or other haematological disorders. The pleiotropic effects of TET2 during haematopoiesis include regulating HSC self-renewal, lineage commitment and differentiation into specific lineages. Based on the role of *TET2* in these functions and tight regulation of cellular homeostasis, the gene is widely accepted as a tumour suppressor.

TET2 expression is high in HSCs and *Tet2* deletion results in an expansion of c-kit+ Lin- immature progenitor cells, which can be attributed to the aberrant self-renewal and impaired differentiation of HSCs. Downregulation of *TET2 in vitro* resulted in decreased expression of *HOXA* genes, which are critical for haematopoietic differentiation (Bocker et al. 2012). *Tet2* silencing in mice induces a differentiation bias towards myeloid lineage cells suggesting a link between *TET2* mutations and myeloid malignancies (Moran-Crusio et al. 2011). In addition, *TET2* deletion in CD34+CD38+ myeloid progenitors resulted in an increase in monocytic lineage cells (Bowman et al. 2017; Li et al. 2011). This suggests a regulatory role of TET2 on monocyte lineage differentiation, which is in agreement with the relatively higher frequency (50%) of *TET2* mutations observed in chronic myelomonocytic leukaemia (CMML) (Solary et al. 2014). Moreover, self-renewal regulators such as *Meis1* and *Evi1* were upregulated and myeloid specific factors such as *Cebpa, Mpo* and *Csf1* were downregulated in *TET2* null HSCs. This reiterates the importance of *TET2* in self-renewal and differentiation of HSCs towards myeloid lineage cells (Moran-Crusio et al. 2011; Li et al. 2011).

TET2 mutations are also implicated in lymphoid malignancies. *Tet2* deficient mice showed impaired B-cell and T-cell development (Solary et al. 2014). Among the various haematological malignancies, *TET2* mutations occur in 33% of angioimmunoblastic T-cell lymphoma, 30% of blastic plasmacytoid dendritic cell neoplasm, 25% of MDS, 23% of AML, 20% of T-cell lymphoma, 13% of MPN and 10% of B-cell lymphoma (Feng et al. 2019).

Role of TET2 in clonal haematopoiesis

Clonal haematopoiesis of indeterminate potential (CHIP) is defined as the clonal expansion of haematopoietic cells carrying certain 'driver' mutations that are associated with a higher risk of haematological malignancies. It is often considered as a preleukemic condition in normal individuals (Busque et al. 2018; Lichtman 2015). CHIP is widely considered an ageing related event (Busque et al. 2018). *TET2* is among the most frequently mutated candidate CHIP driver genes, with others including *DNMT3A* and *ASXL1*. Individuals with CHIP have a 10-fold increased risk of progressing to develop a haematological malignancy (Genovese et al. 2014; Busque et al. 2018). Mutations in *TET2* which include frameshifts, missense and nonsense mutations, confer increased self-renewal potential to HSCs, leading *TET2*-mutated HSCs to outcompete wild-type HSCs and populate the bone marrow (Busque et al. 2018). However, additional co-operating events are necessary for progression to AML (Figure 1.7) (Abdel-Wahab, Tefferi, et al. 2012). Considering the role of *TET2* in CHIP, it is not surprising that *TET2* mutations alone appear to be insufficient to drive leukaemogenesis (Shih et al. 2015) and they commonly co-occur with other mutations in AML, such as *NPM1, RUNX1-RUNX1T1* and *FLT3-ITD,* which co-operate to drive leukaemogenesis Table 1.3 (Weissmann et al. 2012). Of note, mutations in genes that phenocopies TET2, such as *IDH1*, *IDH2* and *WT1* were highly infrequent or mutually exclusive to *TET2* mutations (Wang, Xiao, et al. 2015; Weissmann et al. 2012).

Molecular mutations	**Percentage of other gene mutations co-occurring in *TET2*-mutated AML**
NPM1	55.9%
FLT3-ITD	30.6%
RUNX1	22.2%
KRAS	10.3%
RUNX1-RUNX1T1	10.3%
JAK2	9.3%
CEBPA	7.5%
FLT3-TKD	5.2%
CBL	4.6%
IDH1	2.9%
IDH2	0.0%

Table 1.3. Molecular mutations that co-occur with *TET2* mutations in AML. The percentage of other mutations that were also present in 87 *TET2*-mutated AML cases are shown. Modified from (Weissmann et al. 2012)

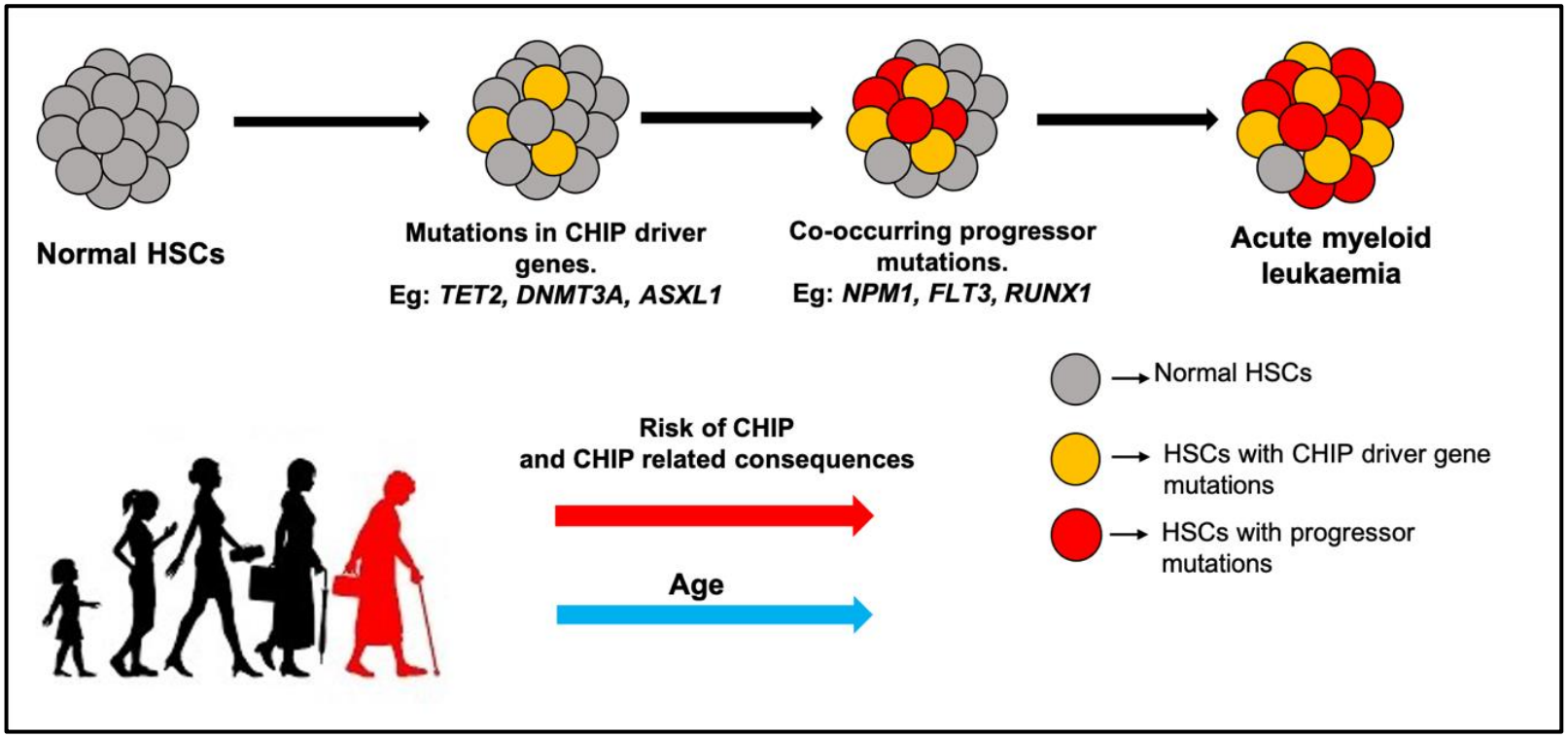

Figure 1.7. Clonal Haematopoiesis of Indeterminate Potential (CHIP). Normal HSCs acquire the so-called driver mutations (yellow) in genes such as *TET2*, *DNMT3A* or *ASXL1*. This confers increased self-renewal potential to mutated cells and results in clonal expansion of cells with 'driver' mutations. But these driver mutations alone are insufficient to cause AML and hence is referred to as pre-leukaemic condition. Further, pre-leukaemic cells acquire additional co-occurring mutations referred to as the progressor mutations (red) in genes such as *FLT3*, *NPM1* or *RUNX1* to cause leukaemia. This concept is known as CHIP and is widely accepted as an age-related event and risk of CHIP and CHIP related consequences increases with age.

ASXL1, Additional Sex Combs Like 1, Transcriptional Regulator; CHIP, Clonal Haematopoiesis of Indeterminate Potential; DNMT3A, DNA methyltransferase 3A; FLT3, Fms like Tyrosine Kinase 3; HSC, Haematopoietic Stem Cell; NPM1, Nucleophosmin 1; RUNX1, Runt-Related Transcription Factor 1; TET2, Tet Methylcytosine Dioxygenase 2

TET2 mutations in AML and the effect of mutant allele dosage

Despite *TET2* being a frequently mutated epigenetic modifier in AML, there is little known about the of allele dosage on leukemogenesis, and *TET2* biallelic mutations in particular. *TET2* is somatically mutated in 7%-23% of AML, with approximately 75% monoallelic mutations and 25% biallelic mutations (Ahn et al. 2015). Generally, *TET2* mutations are loss-of-function mutations leading to inactivation of the protein and are considered as a poor prognostic marker in AML. The event free survival of *TET2*-mutated AML is significantly lower than AML with wild-type *TET2*. Moreover, there is a higher relapse rate in patients with homozygous *TET2* mutation when compared to wild-type *TET2*, single *TET2* mutations and *TET2* heterozygous double mutation (compound heterozygosity) (Ahn et al. 2015; Chou et al. 2011; Fernandez-Mercado et al. 2012; Gaidzik et al. 2012; Metzeler, Maharry, Radmacher, Mrozek, et al. 2011; Nibourel et al. 2010; Tefferi et al. 2009).

Haploinsufficiency of *TET2* drives haematological malignancies to a similar extent as homozygous *TET2* null cells in mice. However, heterozygous *TET2* mutants retained partial mRNA level expression of *TET2* whereas homozygous mutants had complete loss of *TET2* mRNA expression. Consequently, the 5-hmC levels were significantly lower in homozygous compared to heterozygous *TET2* mutants, suggesting the progressive loss of TET2 functionality with increasing mutant allele dosage (Ahn et al. 2015). The frequency of biallelic *TET2* mutations was 7.8% in a cohort of 1045 patients with myeloid neoplasia and about 6% in AML patients within the same cohort. These mutations were more prevalent in older patients, which is in line with the role of *TET2* in ageing associated CHIP phenomenon (Awada et al. 2019). Also, *TET2* biallelic mutations most frequently co-occurred with mutations such as *ASXL1* (63%), *RUNX1* (25%) and *SRSF2* (25%). Mutations resulting in truncation of TET2 protein was significantly more frequent in biallelic than monoallelic *TET2* mutants (Awada et al. 2019).

Weissman et al. reported the landscape of *TET2* mutations in a cohort of 318 AML patients (Weissmann et al. 2012). The various mutations included point mutations (61.1%), deletions (22.1%), duplications (16%) and insertions (0.8%) and were distributed across all exons with the majority (56.5%) hitting the largest exons 3 and 11 and the highest density of mutations occurred within the exons 5 and 11 (Weissmann et al. 2012). Most of the missense mutations occurred in the two conserved domains and were predicted to have a damaging effect on the structure and function of TET2 protein. Other *TET2* mutations such as frameshift mutations (34.4%) and non-sense mutations (37.4%) were seen across all exons (Weissmann et al. 2012).

As such, *TET2* mutations co-occurring with *FLT3-ITD* resulted in hypermethylation of various regulatory elements of genes involved in haematopoietic differentiation and self-renewal such as *GATA1, GATA2, ID1, SOCS2* and *MPL* (Shih et al. 2015). Similarly, *TET2* mutations in cells with *RUNX1-RUNX1T1* translocation resulted in hypermethylation of enhancer elements of various genes and most of these enhancers regulated the expression of tumour suppressors, suggesting a mechanism for *TET2*-mediated leukaemogenesis (Rasmussen et al. 2015).

TET2 mutations in other cancers

Although *TET2* mutations are more prevalent in haematological malignancies, there is evidence implicating these mutations in other cancers (Solary et al. 2014; Bowman et al. 2017). *TET2* mutations are identified in melanoma and squamous cell carcinoma and loss of 5-hmC is widely accepted as an epigenetic hallmark of melanoma. Moreover, re-introduction of active *TET2* re-established the 5-hmC levels and resulted in suppression of tumour invasion and growth in melanoma (Lian et al. 2012). In addition, *TET2* mutations were identified in metastatic castrate-resistant prostate cancer (Nickerson et al. 2013). *TET2* plays a role in androgen receptor signalling by binding to androgen receptor and altering the expression of various genes in the pathway. Downregulation of *TET2* was associated with increase in cell proliferation and invasion, increased prostate specific antigen expression, decreased patient survival and worse prognosis in prostate cancer patients (Nickerson et al. 2017). Similarly, *TET2* was downregulated in glioblastoma and ectopic overexpression of *TET2* resulted in decrease in cell proliferation rate (García et al. 2018). Overall, dysregulated *TET2* expression is implicated in the pathology of various malignancies and is usually tightly regulated in order to maintain normal cellular homeostasis.

1.4.4. Regulators of TET2 activity

Mutations in other components of the methylation pathway are also reported in AML, including *IDH1* (11%), *IDH2* (11.9%) and *WT1* (8.3%) (Aref et al. 2014; DiNardo et al. 2015). Isocitrate dehydrogenases (IDH1 and IDH2) catalyse the formation of α-KG from isocitrate, and this process regulates TET2 activity (Figure 1.8A). Mutations in *IDH1/2* result in the production of 2-hydroxyglutarate (2-HG) instead of α-KG, which inhibits TET2 and blocks demethylation activity. Furthermore, Wilm's tumour (WT1) binds to TET2 and recruits it to target genes to be demethylated. As such, mutation in *WT1* prevents binding of TET2 to target DNA and phenocopies *TET2* mutation. As such, mutations in *IDH1, IDH2, WT1* and *TET2* all result in the acquisition of a hypermethylation phenotype (Figure 1.8A). Therefore, it is not surprising

that mutations in *WT1, IDH1* and *IDH2* very rarely co-occur with *TET2* mutation in AML and are considered to be mutually exclusive (Wang, Xiao, et al. 2015).

TET2 is also regulated through proteolysis mediated by IDAX, also known as CXXC4. The CXXC domain of IDAX specifically binds the regions with unmethylated CpG residues and recruit caspases thereby leading to cleavage and degradation of TET2 (Ko et al. 2013). Similarly, CXXC5 is another protein that can promote proteolytic degradation of TET2 (Solary et al. 2014).

Vitamin C (ascorbic acid) functions as another co-factor for Fe(II)-α-KG dependent dioxygenase enzymes activity. Specifically, vitamin C promotes the activity of TET2 through binding to the C-terminal catalytic domain of TET2 and facilitating the folding process and also recycling of Fe(2+) (Cimmino et al. 2017; Yin et al. 2013). In mice, the global 5-hmC level increased due to DNA demethylation following vitamin C treatment and the same effect was not observed in *TET1/TET2* double knockout mice showing that the effect of vitamin C is TET dependent (Yin et al. 2013). As such, vitamin C treatment restores TET2 function and suppresses perturbed self-renewal and leukaemia progression in *TET2* deficient mice (Cimmino et al. 2017; Chong et al. 2019; Solary et al. 2014; Feng et al. 2019; Yin et al. 2013).

TET2 is regulated as part of immune suppression during viral infections such as HIV-1. HIV Viral protein R (Vpr) is an accessory protein of HIV-1 that is critical for the efficient viral replication and disease progression (Lv et al. 2018). Vpr targets TET2 for proteasomal degradation through polyubiquitination mediated by $CRL4^{VprBP}$ (VprBP-DDB1-CUL4-RING E3 ubiquitin ligase complex) (Lv et al. 2018; Nakagawa et al. 2015). This is considered as one of the early events (as early as 6 hours) after HIV-1 infection (Lv et al. 2018). The hallmark of HIV-1 infection itself is persistent inflammatory response. To mitigate this, as part of inflammation resolution phase, NFKB Inhibitor Zeta recruits TET2 to *IL-6* promoter. TET2 binds and recruits HDAC1/2 which deacetylates and inactivates *IL-6* transcription (Lv et al. 2018). However, the proteasomal degradation of TET2 by Vpr, inhibits this inflammation resolution pathway and results in a sustainable IL-6 expression and persistent inflammation. IL-6 has pro- and anti-inflammatory functions and elevated IL-6 is implicated in progression of HIV-1 infection via stimulation of HIV-1 replication. CRISPR-CAS9 mediated deletion of *TET2* increased HIV-1 replication whereas *TET2* overexpression resulted in decreased HIV-1 replication (Lv et al. 2018). This reveals a role for TET2 in immune regulation and a role for Vpr in facilitating viral infection by post translationally silencing TET2 expression (Lv et al. 2018; Breen et al. 1990; Feng et al. 2019).

TET2 is also regulated by micro-RNA, miR-22 which directly targets the 3'-UTR of *TET2* and downregulates TET2 activity. miR-22 is frequently overexpressed in MDS and haematological malignancies including AML. Overexpression of miR-22 resulted in global reduction of 5-hmC levels and impaired differentiation and increased self-renewal potential in mice (Song et al. 2013; Solary et al. 2014). The other micro RNAs that target *TET2* are miR-7, miR-26, miR-29, miR-101 and miR-125, which are all upregulated in *TET2* wild-type AML (Cheng et al. 2013; Feng et al. 2019).

As such, there are several factors that regulate TET2 through a tight regulatory network (Figure 1.8B) and therefore it might be clinically relevant to consider the differential expression of these TET2 regulators along with *TET2* mutant allele dosage, in order to completely decipher the role of TET2 in AML pathology and treatment response.

1.5. 5-Azacytidine mediated demethylation therapy for AML

Hypomethylating/demethylating agents such as azacytidine and decitabine, first synthesized in 1964, are gaining use in the treatment of AML, either as single agents or in combination with other drugs. Both of these hypomethylating agents have similar structure, mechanism of action and function which is discussed in this section with a particular focus on azacytidine (Christman 2002).

1.5.1. Cellular uptake of azanucleosides

Azacytidine and decitabine are azanucleosides which function as cytidine analogue and specifically inhibits DNA methylation. The C5 position in the pyrimidine ring of cytidine is substituted with a nitrogen atom in azanucleosides (Figure 1.9). In azacytidine, the nitrogen substituted base is attached to a ribose sugar whereas in decitabine, it is attached to a deoxyribose sugar (which lacks one oxygen atom at the 2' position of ribose ring). Due to this particular structural difference, azacytidine is mostly incorporated into RNA and decitabine into DNA (Stresemann et al. 2008). The cellular uptake of nucleoside analogues is generally regulated by various transporters that belong to solute carrier (SLC) and ATP binding cassette (ABC) family of proteins (Damaraju et al. 2012; Rius et al. 2009; Stresemann et al. 2008; Stomper et al. 2021). Following cellular uptake, azanucleosides are activated through a sequential cascade of ATP-dependent phosphorylation reactions. Azacytidine (5-azacytidine) is converted to 5-azacytidine monophosphate by the enzyme uridine-cytidine kinase whereas decitabine is converted to 5-aza-2'-deoxycytidine monophosphate by deoxycytidine kinase.

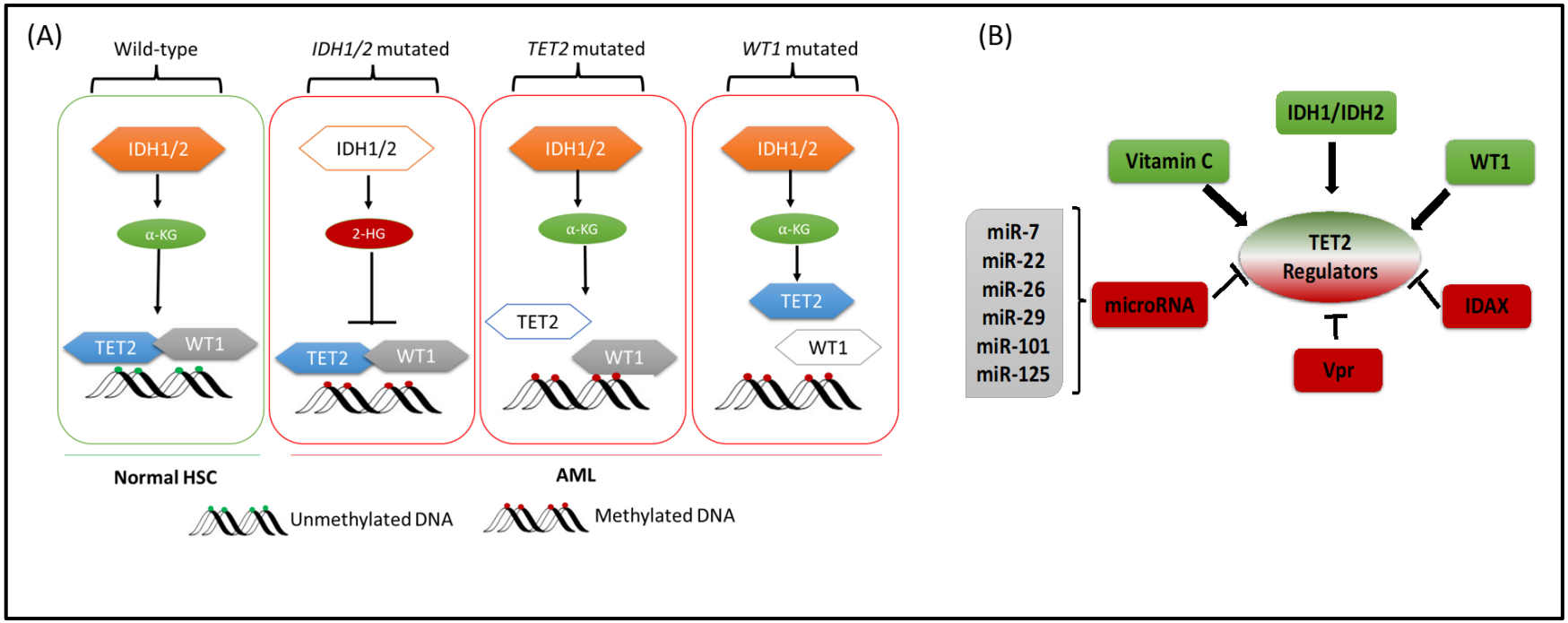

Figure 1.8. TET2 Regulatory network (A) IDH1/2-TET2-WT1 pathway in AML**.** In normal HSC, IDH1 and IDH2 catalyses the formation of α-ketoglutarate (α-KG) which is required for the catalytic activity of TET2. In this process, WT1 binds and recruits TET2 to the target gene to be demethylated. In AML, when there is a mutation in *IDH1* or *IDH2*, 2-hydroxyglutarate (2-HG) is formed instead of α-KG and this inhibits the functional activity of TET2. Similarly, in AML, mutations in *TET2* or *WT1* can also inhibit the binding of TET2 to target DNA and hence block the demethylation process. Figure modified from (Wang, Xiao, et al. 2015) (B) IDH1/IDH2, WT1 and Vitamin C positively regulates TET2 (green boxes) whereas Vpr, IDAX and micro RNAs such as miR-7, miR-22, miR-26, miR-29, miR-101 and miR-125 negatively regulates (red boxes) TET2. Figure modified from (Feng et al. 2019)

5-hmC, 5-hydroxymethylcytosine; 5-mC, 5-methylcytosine; α-KG, α-ketoglutarate; AML, Acute Myeloid Leukaemia; 2-HG, 2-hydroxyglutarate; IDAX, Inhibition of the Dvl and Axin Complex Protein; HSC, Hematopoietic Stem Cells; IDH, Isocitrate dehydrogenase; miR, micro-RNA; SIRT1, Sirtuin 1; TET2, Ten-Eleven-Translocation-2; UHRF2, Ubiquitin Like with PHD And Ring Finger Domains 2; VprBP, HIV-1 Vpr-Binding Protein; WT1, Wilm's Tumour gene

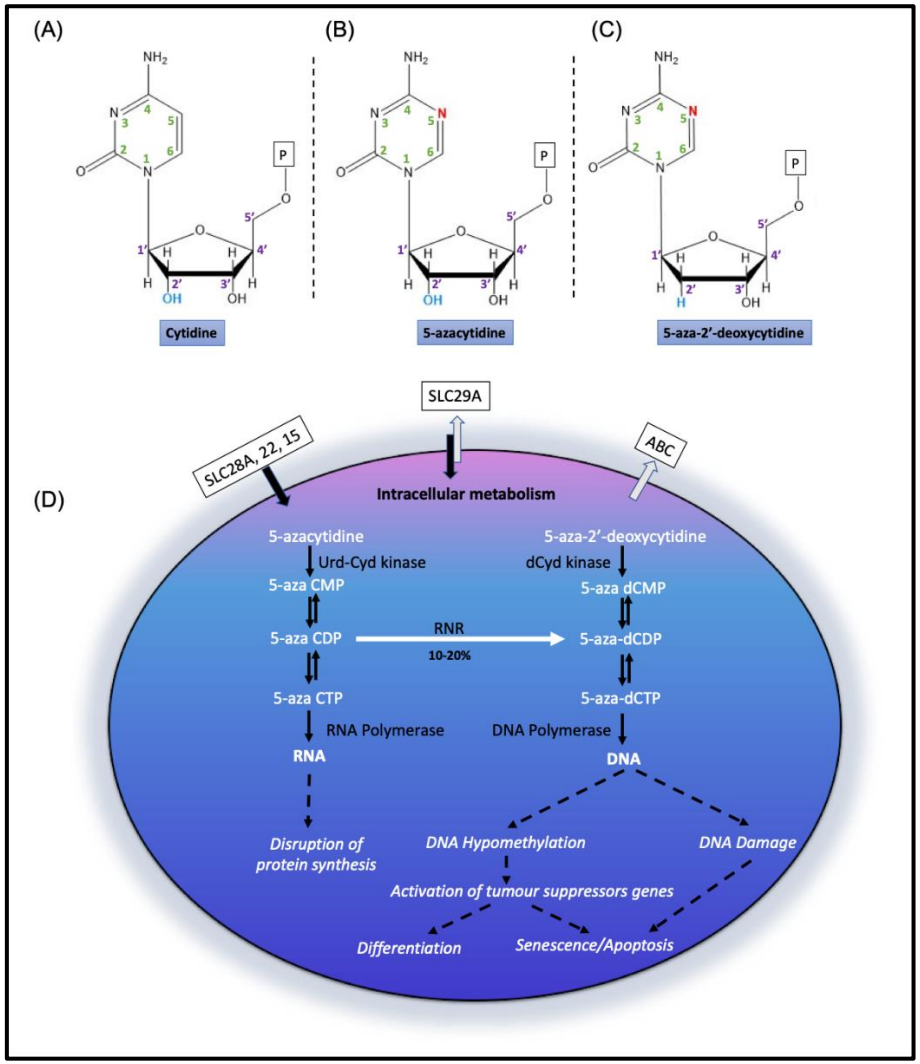

Figure 1.9. Azacytidine structure and mechanism of cellular uptake (A) Structure of Cytidine ribonucleotide where cytosine pyrimidine base is attached to a 5-carbon ribose sugar. (B) Azacytidine (5-azacytidine) structure where C5 in the pyrimidine ring is replaced by a nitrogen atom and the base is attached to a 5-carbon ribose sugar. (C) Decitabine (5-aza-2'-deoxycytidine) structure where C5 in the pyrimidine ring is replaced by a nitrogen atom and the base is attached to a deoxyribose sugar where the hydroxyl group in the 2'position of ribose sugar (in blue) is replaced by a hydrogen atom. The numbers in green depicts pyrimidine ring atoms and those in purple depict ribose/deoxy ribose sugar ring atoms. (D) The mechanism of cellular uptake of azacytidine and decitabine involves various transporter proteins such as SLC carriers and ABC transporters. Azacytidine through sequential phosphorylation is converted to 5-aza-CMP, -CDP and -CTP and is incorporated into RNA. This entire process is mediated by enzymes such as Urd-cyd kinases and RNA polymerase, which results in disruption of protein synthesis. Similarly, decitabine is converted into 5-aza-dCMP, -dCDP and dCTP and is incorporated into DNA, which is mediated by dCyd kinase and DNA polymerase. 5-aza CDP can also be converted to 5-aza dCDP by RNR and subsequently incorporated into DNA. As such, drug incorporation into DNA causes DNA hypomethylation mediated activation of tumour suppressor genes and results in differentiation, senescence or apoptosis. Another mode of cytotoxicity is induced by triggering DNA damage through drug incorporation that results in apoptosis. Data obtained from (Duchmann et al. 2019; Stresemann et al. 2008; Stomper et al. 2021).

ABC, ATP Binding Cassette transporters; 5-Aza, azacytidine; CDP, cytidine diphosphate; CMP, cytidine monophosphate; CTP, cytidine triphosphate; dCDP, deoxy CDP; dCMP, deoxy CMP; dCTP, deoxy CTP; dCyd kinase, deoxy cytidine kinases; RNR, ribonucleoreductase; SLC, solute carrier family of transporters; Urd-Cyd kinase, uridine-cytidine kinase

Thereafter, 5-azacytidine monophosphate (5-aza CMP) is sequentially converted to 5-azacytidine diphosphate (5-aza CDP) and then to 5-azacytidine triphosphate (5-aza CTP), which is incorporated into RNA via the action of RNA polymerases. Similarly, 5-aza-2'-deoxycytidine triphosphate (5-aza dCTP) which is the activated form of decitabine gets incorporated into DNA. A smaller fraction (10-20%) of 5-aza CDP is converted to 5-aza dCDP by ribonucleotide reductases (RNR) and then to 5-aza dCTP, which is incorporated into DNA, where cytosine residues are substituted by azacytosine residues (Figure 1.9) (Stresemann et al. 2008; Stomper et al. 2021)

1.5.2. Pharmacodynamics of Azacytidine

The azacytidine incorporated into DNA is misread as a natural cytosine substrate by DNA methyltransferase DNMT1. DNMT1 initiates the methylation reaction by a nucleophilic attack which forms a covalent bond between the carbon-6 of azacytidine and Cys^{1226} of DNMT1 which drives the methyl group transfer from SAM. The final step is a β-elimination reaction of the H5 atom (attached to C5) by which the covalent complex at C6 is normally resolved. However, the presence of a nitrogen atom in place of C5 blocks this enzyme release step and traps DNMT1 which facilitates its degradation, resulting in a general depletion of DNMT pool in cells and gradual DNA hypomethylation after each round of replication (Figure 1.10) (Du et al. 2016; Stresemann et al. 2008; Schermelleh et al. 2005).

The various mechanisms of cytotoxicity induced by azanucleosides include hypomethylation induced activation of various tumour suppressor genes that lead to apoptosis, senescence, and differentiation. Specifically, azacytidine induces differentiation of leukaemic cells in *TET2*-mutated AML *in vivo* (Shih et al. 2017; Duchmann et al. 2019). Another mode of cytotoxicity is mediated by the DNA damage triggered as a result of the covalently trapped DNMT1 adduct and incorporation of azacytidine (Palii et al. 2008; Stresemann et al. 2008). Finally, a majority of azacytidine drug that is incorporated into RNA interferes with RNA methylation by inhibiting DNMT2 RNA methyltransferase, which disrupts cellular protein synthesis which further promotes apoptosis. Azacytidine treated cells have a significantly lower metabolic rate compared to decitabine treated cells, which is consistent with the unique role of azacytidine in RNA metabolism (Duchmann et al. 2019; Schaefer et al. 2009). Figure 1.9 summarises the mode of action of azacytidine and decitabine.

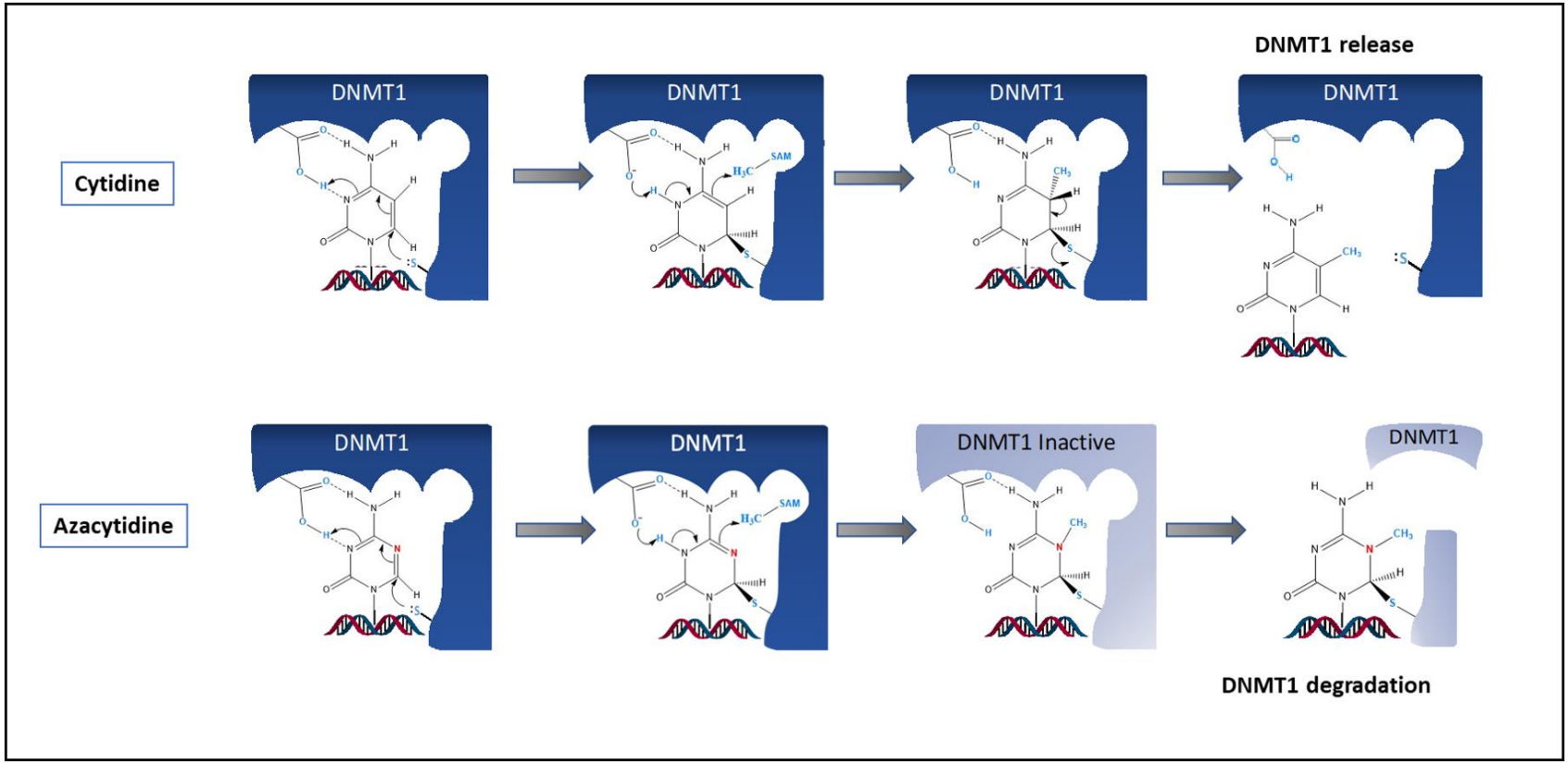

Figure 1.10. Pharmacodynamics of azacytidine. Azacytidine incorporated in the DNA acts as a substrate for DNMT1 (in blue) in a similar way as the inherent cytidine residue. This triggers the enzymatic action of DNMT1 whereby through a nucleophilic attack, the thiol group in Cys^{1226} forms a covalent bond with the C6 atom of cytidine and azacytidine. This drives the methyl group transfer from SAM (in blue) to the C5 position of the substrate. The reaction is completed through a β-elimination reaction at the H5 position whereby the covalent complex at C6 is resolved and enzyme is released (top panel). However, β-elimination reaction is blocked due to the presence of nitrogen atom at the C5 position (in red) and DNMT1 is inactivated and trapped resulting in DNMT1 degradation and gradual DNA hypomethylation (bottom panel). The atoms in blue are from DNMT1 and those in black are substrate. The chemical structures were generated using Chemdraw Figure adapted from *(Stresemann et al. 2008; Du et al. 2016)*

DNMT1, DNA methyltransferase 1; SAM, S-adenosyl-L-methionine

1.5.3. Azacytidine response biomarkers

An understanding of potential biomarkers for drug response is critical for optimal treatment for individual patients. *TET2* mutations are associated with a better response to azacytidine in MDS, AML and CMML (Itzykson et al. 2011; Cedena et al. 2017; Bejar et al. 2014; Duchmann et al. 2018) although some studies did not replicate these findings (Döhner et al. 2018). Moreover, none of these studies investigated the role of *TET2* mutant allele dosage on response to azacytidine. In addition, *DNMT3A* and *IDH1/2* mutations correlates with a higher response rate in patients with MDS and AML (Coombs, Sallman, et al. 2016; Emadi et al. 2015). Chromatin organisation mediated by RNA methylation and RNA methyltransferases plays a role in response to azacytidine treatment. The RNA methyltransferase NSUN1 binds BRD4 and RNA polymerase II to form an active chromatin structure that is resistant to azacytidine. RNA methylation and NSUN1/BRD4 mediated active chromatin were significantly higher in azacytidine resistant cells lines as well as MDS/AML patients (Cheng et al. 2018). Also, mutations in genes involved in cellular uptake and metabolism of azacytidine such as the transporters and kinases have been implicated in regulation of drug response (Valencia et al. 2014; Duchmann et al. 2019).

1.5.4. Major clinical trials and treatment regimens with azacytidine

The phase III AZA-001 study (NCT00071799) was a benchmark in the development of azacytidine as a treatment in MDS and some haematological malignancies. This study involved administration of either azacytidine at a dose of 75mg/m^2 for 7 days in every 28 days or conventional care (best supportive care, low-dose cytarabine, or intensive chemotherapy) for control group. The trial recruited MDS patients with a high risk of transformation to AML and who were not eligible for allogeneic stem cell transplantation (ASCT). The major outcome of the study was a better response rate in terms of increased overall survival and delayed transformation to AML in the azacytidine-treated cohort (Fenaux et al. 2009; Gore et al. 2013). These studies resulted in azacytidine treatment being approved for use in intermediate-2/high risk MDS in Europe (Duchmann et al. 2019). Analysis on a subset of AZA-001 study involved elderly AML patients (median age 70 years) with low blast count (20%-30%), of which 86% of the cohort were unfit for intensive chemotherapy. Azacytidine-treated patients had significantly higher median survival rates and two-year survival rates compared to patients treated with conventional care regimens. Moreover, azacytidine-treated patients with adverse karyotype had a significantly higher overall survival. In particular, patients with chromosomal

abnormalities such as -5/5q-, -7/7q-, or 17p and monosomal or complex karyotypes had a reduced risk of death when treated with azacytidine. (Fenaux et al. 2010; Döhner et al. 2018).

Another phase III trial (NCT01074047) recruited elderly AML patients (≥ 65 years) with a blast count >30% and evaluated azacytidine efficacy and safety. Azacytidine-treated patients had an overall higher survival rate compared to patients treated with conventional care regimens (standard induction chemotherapy, low-dose ara-c, or supportive care only) (Dombret et al. 2015).

In addition to single agent therapy, azacytidine is also being trialed in combination with other targeted therapies. A phase Ib trial (NCT02203773) involving elderly AML patients demonstrated an improved response rate when azacytidine was combined with anti-apoptotic BCL-2 inhibitor venetoclax compared to azacytidine single-agent treatment (DiNardo et al. 2019). An insight into the molecular mechanism of this regimen revealed that the metabolic activity of leukaemic stem cells is altered with this combination treatment. This regimen induced a better response when given as salvage therapy for relapsed or refractory AML patients compared to venetoclax and low dose ara-C combination regimen (DiNardo et al. 2019). There are several other ongoing clinical trials evaluating various combination regimens involving azacytidine coupled with checkpoint inhibitors, HDAC inhibitors (entinostat, pracinostat, valproic acid, vorinostat), IDH1/IDH2 inhibitors and other therapeutics such as lenalidomide (a drug that is commonly used in MDS with del(5q)) and sapacitabine (an analogue of cytarabine) (Gardin et al. 2017; Duchmann et al. 2019).

The increasing number of ongoing clinical trials incorporating azacytidine as a single agent or in combination regimens demonstrates the potential of this drug in the treatment of AML. Considering the fact that older patients do not often benefit from intensive chemotherapy and that mutations in DNA methylation related genes associate with older age at diagnosis, demethylating agents in combination with targeted therapies is likely to emerge as standard of care for this group of patients (Burnett 2018; Lubbert et al. 2005; Muller et al. 2006; Yu and Zheng 2017).

1.6. Thesis background

An AML index patient, 76 years old, who presented with a t(4;12) translocation (46,XY,t(4;12)(q2?;q13)[12]46,XY[10]) was initially treated with cytarabine (Ara-C) and daunorubicin as part of remission induction chemotherapy. However, the disease was chemoresistant showing a leukaemia blast reduction from 90% at presentation to only 29% at

day 30. Moreover, the patient developed pancytopenia and acute septicaemia and required intensive care incorporating intravenous antibiotics and vasopressor therapy. Further to recovering from infection, the patient was treated with single agent azacytidine as palliative care, which unexpectedly resulted in prolonged complete morphological remission for a period of 24 months. However, the patient experienced a relapse which was resistant to Ara-C and sorafenib and died 28 months after first diagnosis Figure 1.11A).

In depth analysis of cytogenetic and molecular features of leukaemic cells at presentation, remission and relapse through next generation sequencing, FISH and SNP array experiments, revealed that the patient carried a biallelic *TET2* mutation Figure 1.11B). The mutation was characterised by a large focal 1.1Mb deletion in the region spanning as the *TET2, CXXC4* and *PPA2* genes present in 95% of cells, and a *TET2* nonsense base substitution mutation (c.2815C>T) in exon 3 (codon 939 (Q939*)) in the retained allele that generated a stop codon, which was present in 96% of reads when sequenced. In addition, the presentation sample carried a heterozygous *NPM1* mutation characterised by c.863_864ins, TCTG, which was present in 50% of reads at presentation when sequenced. Although the remission bone marrow appeared to be morphologically normal, both *TET2* gene deletion and base substitution mutation were retained at high levels in the remission bone marrow. The *NPM1* mutation was not detectable in remission bone marrow but can be presumed to have persisted at below detection levels as this mutation was a prominent feature of relapse disease. As such, the relapse was characterised by a *TET2* biallelic mutation (deletion and base substitution), *NPM1* mutation and an additional mutation, *FLT3* internal tandem duplication (ITD) which was not observed at presentation or remission. *FLT3-ITD* was characterised by a c.1780_1800 duplication (TTCAGAGAATATGAATATGAT, p.F594_D600 duplication) which was present in 40% of reads. An analysis of the tumour phylogeny suggested that the *TET2* nonsense mutation with subsequent deletion of the second *TET2* allele was an ancestral AML-driving event which was followed by *NPM1* mutation during the pathogenesis of AML.

As such, there was a complete elimination of *TET2/NPM1*-mutated clones, a reduction of *TET2* biallelic mutated_*NPM1* wild-type ancestral cells and a restoration of normal haematopoiesis leading to cytomorphological remission, solely driven by the single-agent azacytidine treatment. Therefore, it was hypothesised that the *TET2* mutant allele dosage can have a positive impact on cellular response and sensitivity to azacytidine.

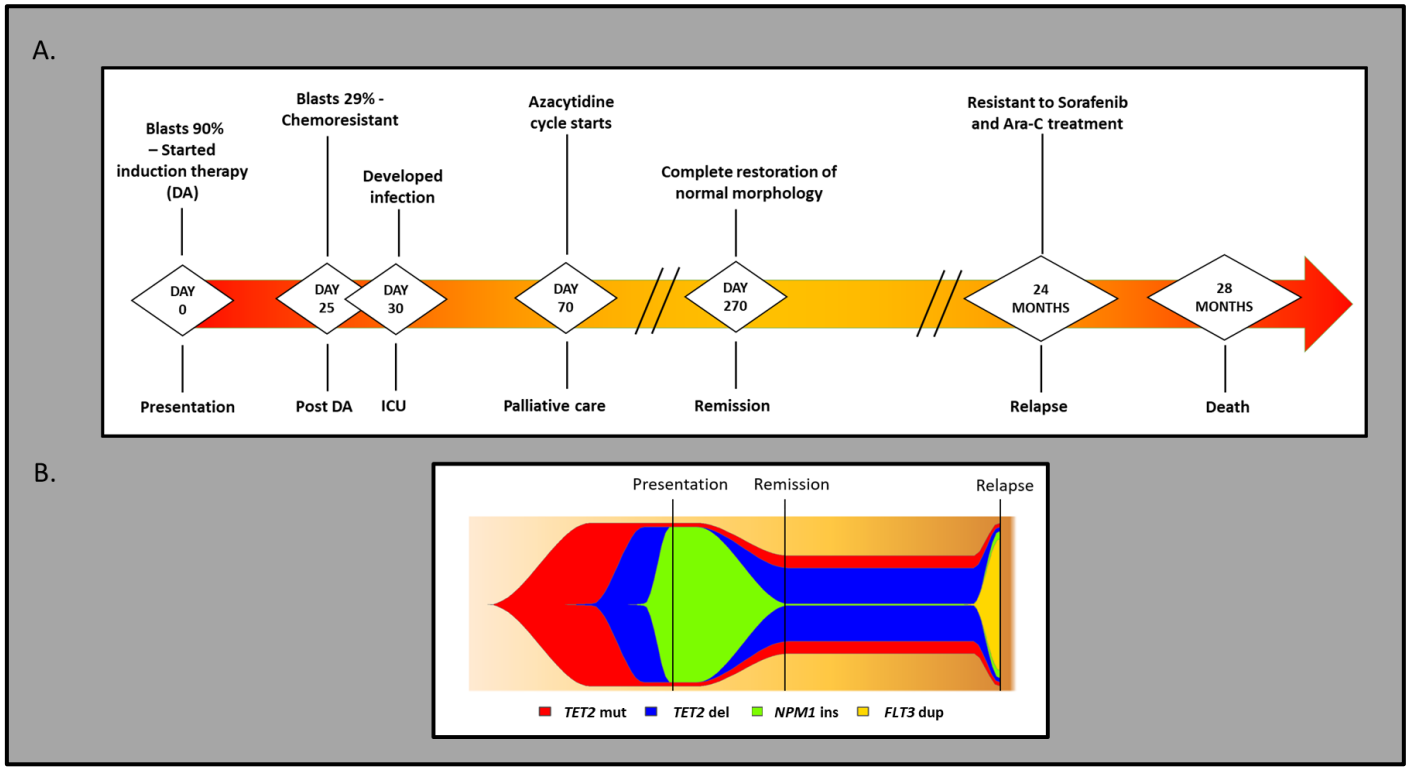

Figure 1.11. Clinical and mutational timeline of *TET2* biallelic mutated AML patient.

(A) Timeline summarises the major events in the patient's clinical history. At presentation (day 0), patient was diagnosed with 90% blast count and started induction therapy. At day 25 the blast count was still 29% indicating chemoresistance. Further, at day 30, the patient developed an infection and was moved to intensive care. At day 70, having recovered from infection, the patient was given single agent 5-azacytidine as palliative care, which led to a complete morphological remission for 24 months. The patient subsequently experienced a relapse, which was resistant to Sorafenib and Ara-C and died 28 months after initial diagnosis. (B) Fish plot showing the temporal acquisition of mutations during AML development (generated by Dr. Claire Elstob, Newcastle University). At presentation the patient had biallelic *TET2* mutation characterised by a *TET2* deletion (1.1Mb) and a *TET2* nonsense mutation, plus a *NPM1* insertional mutation. At remission, the *NPM1* mutation was not discernible but the biallelic *TET2* mutation was still detectable. At relapse the patient developed a *FLT3-ITD* mutation in addition to the already existing *TET2* biallelic mutation and *NPM1* mutation.

Ara-C, Cytarabine; DA, Daunorubicin-Cytarabine based treatment; del, deletion; dup, duplication; FLT3, FMS-like Tyrosine Kinase 3; ICU, Intensive care unit; ins, insertion; mut, mutation; NPM1, Nucleophosmin 1; TET2, Tet Methylcytosine Dioxygenase 2

1.7. Aims and objectives

AML is a disease that mainly affects older patients, and a majority of these patients are not fit for conventional chemotherapy. *TET2* mutations being a CHIP driver gene mutation, are more prevalent in the elderly population. Moreover, *TET2* mutations are a poor prognostic marker, with biallelic *TET2* mutation associated with a significantly higher relapse rate compared to monoallelic *TET2* mutation or wild-type *TET2*. Therefore, there is an urgent clinical need for better treatment strategies for *TET2* null AML patients. Precisely, the aim of this study is as follows:

- Overall Aim: To develop novel therapeutic strategies to treat *TET2*-mutated AML, with a specific focus on the effect of *TET2* mutant allele dosage on cellular response to azacytidine. In addition, the potential mechanisms behind sensitivity of *TET2* null cells to azacytidine will be investigated.

- Specific Aims:
  (i) To phenotypically characterise the effect of *TET2* mutant allele dosage using an isogenic cell model system *in vitro*.
  (ii) To evaluate the cellular response of *TET2* monoallelic and biallelic mutated AML cells to various established chemotherapy agents that will be tested using *in vitro* models and to investigate whether *TET2* mutant allele dosage has a role on sensitising AML cells to azacytidine.
  (iii) To specifically identify differentially methylated/transcribed regions in *TET2* biallelic mutant AML (compared to isogenic *TET2* monoallelic mutant AML) and identify those with a role in cellular uptake or metabolism of azacytidine.
  (iv) Combinatorial methylation/expression profiling and functional analysis will be exploited to predict the mechanism behind hypersensitivity of *TET2* null cells to azacytidine.

Chapter 2: Materials and Methods

2.1. Chemicals and Reagents

The chemicals and reagents used in this research were of Analar grade and were purchased from Merck Life Science UK Limited (Dorset, UK), unless otherwise stated. Phosphate buffered saline (PBS) was prepared by dissolving one PBS tablet (Life technologies – Paisley – UK) in 500 ml of deionised distilled water (ddH_2O). This solution was autoclaved prior to use and stored at room temperature. The other chemicals and reagents used as part of specific experiments are detailed in this chapter in relevant sections.

2.2. Cell Lines

Human Erythroleukaemia (HEL) and SKM1 AML cell lines were obtained from DSMZ (Braunschweig, Germany). All other AML cell lines such as AML2, AML3, HL-60, Kasumi-1, MV4-11, NB-4, THP1 and U937 were validated stocks at Newcastle University. The characteristics of cell lines are described in relevant results chapters. The identity of cell lines was validated by short tandem repeat profiling (NewGene, Newcastle University, UK) and cultures were tested for mycoplasma every three months, using a MycoAlert kit (Lonza, Slough, UK).

2.3. General cell culture methods

2.3.1. Routine cell culture

All cell culture experiments were performed in a class II microbiological safety cabinet (BIOMAT-2, Medical Air Technology Ltd., Oldham, UK) using aseptic technique. The cell culture plasticware were obtained from Corning-Costar (supplied by VWR international Ltd, Leicestershire, UK). All cell lines were maintained at optimum cell density (Table 2.1) as suspension cultures in sterile 25 cm^3 cell culture flasks in RF10 media (RPMI-1640 medium (Sigma-Aldrich)) supplemented with L-glutamine, sodium bicarbonate, 10% (v/v) FBS (Gibco®, life technologies) and 1% penicillin-streptomycin (v/v) (Sigma-Aldrich)). Cells were incubated in sterile conditions at 37°C and 5% CO_2. A routine cell culture schedule was followed (Table 2.1) in order to maintain the cells in exponential growth phase.

Cell line	Optimum cell density ($x10^6$ cells/ml)	Approximate doubling time (hours)	Split ratio	Time between passage (days)
AML2	0.7-1.5	40-50	1:5	3-4
AML3	0.5-2.0	30-40	1:5	2-3
HEL	0.2-1.0	36	1:10	2-3
HL-60	0.5-1.5	24	1:10	2-3
Kasumi-1	0.5-1	48-72	1:3	2-3
MV4-11	0.4-1.0	50	1:5	2-3
NB4	0.5-1	35-45	1:5	2-3
SKM1	0.4-2.0	48	1:5	3-4
THP1	0.2-1.0	40-50	1:5	3-4
U937	0.2-1.0	30-40	1:5	2-3

Table 2.1. Routine cell culture conditions of AML cell lines

Once established, the CRISPR-targeted *TET2* biallelic mutated clones (HELbi1, HELbi2 and HELbi3) as well as the CRISPR-untreated *TET2* monoallelic mutated control clones (HELmono1, HELmono2 and HELmono3) were maintained under the same conditions as that of HEL parental cells (Table 2.1).

Cell cultures were regularly monitored, and exponential growth of cells was ensured prior to use in experiments. Cell lines were maintained for up to 25 passages, after which fresh stock of cryopreserved cells were resuscitated from liquid nitrogen storage.

2.3.2. Determination of cell density using trypan blue exclusion method

This method is based on the principle that viable cells possess an intact cell membrane and hence are impermeable to dyes such as trypan blue, whereas the dye penetrates the membrane of non-viable cells. Therefore, under the microscope, non-viable cells appear blue in colour and the unstained cells can be easily identified by dye exclusion and counted.

Cell suspensions were mixed thoroughly by pipetting and 1 ml of cell suspension was aliquoted into a 5 ml bijou container (Life Technologies), out of which 10 µl of cell suspension was mixed with 10 µl 0.4% solution of trypan blue (Life Technologies). Finally, 10 µl of the mixture was loaded into a Neubauer haemocytometer (VWR international Ltd). The verification slide

system in the haemocytometer that consists of a well-defined grid system (Figure 2.1) enabled accurate cell counting using a manual cell counter (VWR international Ltd). Consistent cell counting procedures were ensured for all cell lines and a minimum of two replicates were recorded for each cell count. Cell density of each cell line was calculated based on the total number of cells counted (Equation 2.1).

$$Cell\ density = \frac{Total\ no{:}\ of\ cells\ counted\ x\ 2}{No{:}\ of\ major\ squares\ counted}\ x10^4\ cells/ml$$

Equation 2.1. Formula for calculating cell density

2.3.3. Cryopreservation of cells for long-term storage

Following determination of cell density, the corresponding volume of cell suspension containing $5x10^6$ cells was dispensed into a 15 ml or 50 ml sterile BD Falcon™ centrifuge tube (BD Biosciences, Oxford, UK) and centrifuged at 250 x g for 5 min. The supernatant was discarded, and the resultant cell pellet was re-suspended in 1 ml freezing medium [FBS supplemented with 10% (v/v) dimethyl sulphoxide (DMSO)]. The cell suspension was transferred into a sterile polypropylene cryovial (Invitrogen, Life Technologies) and temporarily placed at -80ºC in a Mr. Frosty™ Freezing Container (Life Technologies) which facilitates slow freezing of cells (approximately 1ºC per minute). After 24-48 hours, the frozen vials were moved to liquid nitrogen for long term storage.

2.3.4. Resuscitation of cryopreserved cell stocks

Frozen vials were rapidly thawed in 37ºC water bath, sprayed with 70% ethanol and the contents of the vial were re-suspended in 5 ml of pre-warmed RF10 media in a 15 ml sterile BD Falcon™ centrifuge tube. The tube with the cell suspension was centrifuged at 250 x g for 5 min. The supernatant was discarded, and the pellet was resuspended in 5 ml of fresh RF10 media. The cell suspension was transferred to a 25 cm^3 cell culture flask and incubated at 37ºC 5% CO_2. Cells were monitored daily and expanded to a larger volume with increasing cell number as required. Freshly thawed cells were not used in an experiment until exponential growth phase was achieved. In addition, two vials of the freshly thawed cell line were re-frozen and cryopreserved at an earlier passage to ensure that the cell stocks are adequately replenished.

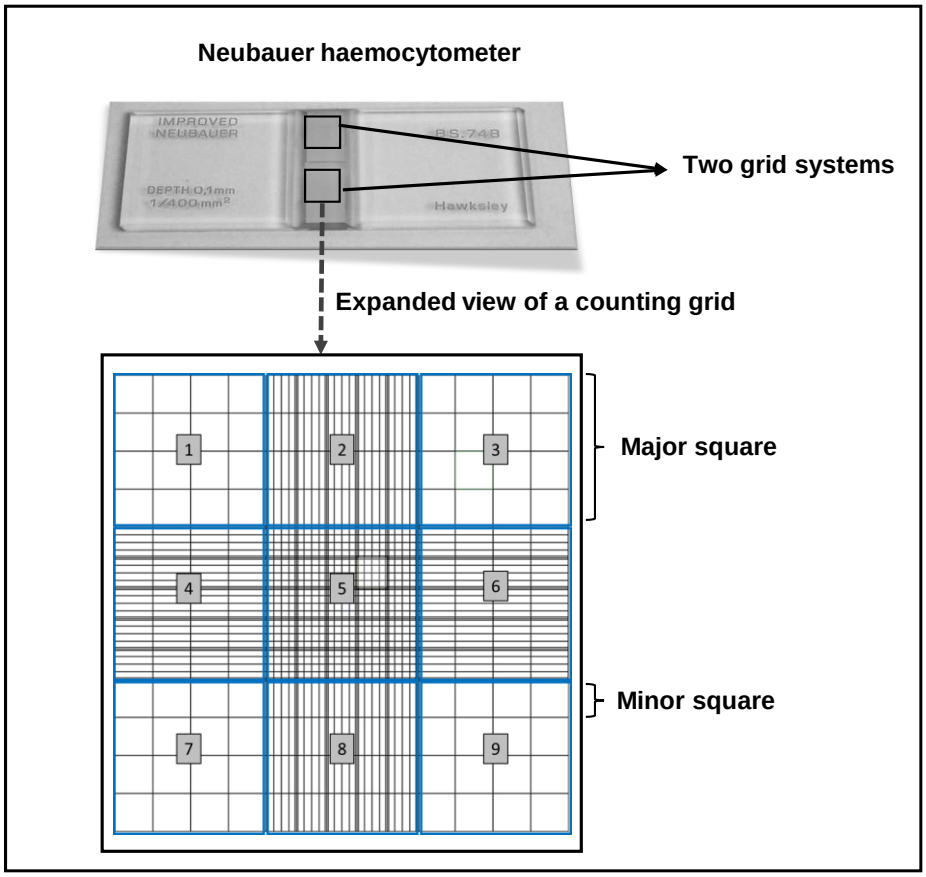

Figure 2.1. Counting grid of a haemocytometer. Each Neubauer haemocytometer (top) consists of two grid systems in the two sides which allows counting cells in replicates simultaneously. The expanded view of a counting grid (bottom) shows one grid system which consists of 9 major squares (blue) and each major square consists of 16 minor squares (black). The cell density is calculated as the total no: of cells counted divided by the total no: of major squares from which the cells were counted. Cells were included in the count of major squares if they intersected the left and top borders, but not the right or bottom borders.

2.3.5. Preparation of cell pellets

Following determination of cell density, an adequate volume of cell suspension containing five million cells was dispensed into a sterile BD Falcon™ centrifuge tube and centrifuged at 250 x g for 5 min. The supernatant was discarded, and the resultant cell pellet was washed with 5 ml PBS and centrifuged at 250 x g for 5 min. Further, the supernatant was discarded, and the pellet was resuspended in 1 ml PBS and centrifuged at 250 x g for 5 min. Finally, after complete removal of supernatant, the resultant cell pellet was stored in -80ºC for long-term storage.

2.4. Proliferation assay

Proliferation assay was performed to determine if *TET2* monoallelic and biallelic mutant cells differed significantly in terms of their proliferation rate in liquid media. An accurate viable cell count was carried out for each cell line using trypan blue exclusion method (Section 2.3.2) using a haemocytometer. Cells were then suspended at a density of $5x10^4$ cells/ml in a total of 12 ml and the flasks were incubated in sterile conditions at 37°C and 5% CO_2. Cells were counted at 24-hour intervals for the next 120 hours. The cell counts of individual cell lines as well as the mean cell counts of *TET2* monoallelic clones and *TET2* biallelic clones were plotted and analysed using two-way ANOVA in Prism (GraphPad version 9). The doubling time for each cell clone was estimated using the non-linear regression option in Prism (GraphPad version 9).

2.5. Growth Inhibition Assay

Growth inhibition assay was performed with an aim to evaluate the role of *TET2* mutant allele dosage on response to various AML therapeutics. *TET2* monoallelic and biallelic mutant cells were treated with increasing doses of different cytotoxic agents and the cell survival percentage was recorded as a measure of the cellular response to the respective drug. In addition, a panel of 10 different AML cell lines were exposed to increasing doses of azacytidine and the dose responses were measured. In general, the various stages of a growth inhibition assay are as follows:

2.5.1. Drug reconstitution

The stock solution of all the cytotoxic agents were prepared using an appropriate solvent (Table 2.2) and were aliquoted and stored at -20ºC or -80ºC, as per the storage requirements of the respective drug.

2.5.2. Preparation of cells

Exponentially growing cells were counted using trypan blue exclusion method (Section 2.3.2. and the cell density was determined. Cells were then seeded at a density of $1x10^5$ cells/ml in RF10 media and made up to 5 ml in a BD Falcon. After thoroughly mixing the cell suspension by repeated pipetting, 250 µl each from the resulting cell suspension was dispensed into six wells (within the same row) of the 24-well plate (Corning) in a horizontal pattern such that each row of the plate which consists of 6 wells corresponds to a single cell line (Figure 2.2). The same process was repeated for all other cell lines used in the experiment. The plates were then incubated at 37°C and 5% CO_2 until the cytotoxic agents were prepared.

2.5.3. Treatment of cells

Working solutions of the cytotoxic agents were prepared by dilution of the stocks (Section 2.5.1) in RF10 media immediately prior to use in experiments. A serial dilution of the working solution was performed by adding an appropriate volume of RF10 media to prepare five different doses within the dose range specified (Table 2.2). A vehicle control was prepared for each drug where an appropriate solvent (without drug) was diluted in RF10 media. Here, the solvent used was the same as the solvent in which the respective stock solution was prepared, and the volume of solvent added was equivalent to the volume of stock solution added in the highest cytotoxic dose within the dose range used for the respective drug. DMSO concentration never exceeded 1% v/v. Drug concentrations at each dose were prepared at double the required concentration, to account for the 1:2 dilution factor (while mixing the drug with the cell suspension in plates).

After thoroughly mixing the serially diluted drug solutions, 250 µl each from the five different drug doses was dispensed into corresponding wells (Dose 2 to 6) (Figure 2.2). The same volume of the vehicle control was dispensed into the 1st well (Dose 1) (Figure 2.2). Given the 1:2 dilution factor, the final cell density after adding the drugs was $5x10^4$ cells/ml. The plates were then incubated at 37°C and 5% CO_2 for 96 hours.

Drug Name	Catalogue No:	Provider	Molecular weight (g)	Solvent	Stock Concentration	Dose range used
Azacytidine	A1287	Merck	244.2	RF-10 media	500 µM	0-20 µM
Decitabine	A3656	Merck	228.21	Sterile ddH_2O	1 mM	0-5 µM
Cytarabine	C6645	Merck	279.68	DMSO	5 mM	0-200 nM
Daunorubicin	30450	Merck	563.98	Sterile ddH_2O	10 mM	0-50 nM
Fludarabine	F9813	Merck	365.21	DMSO	2 mM	0-10 µM
Idarubicin	A2476	Stratech	533.95	DMSO	5 mM	0-80 nM
Calicheamicin	HY19609	MedChem Express	1368.35	DMSO	5 mM	0-0.6125 nM
Ivosidenib	AG-120	Selleckchem	582.96	DMSO	50 mM	0-100 µM
Enasidenib	AG-221	Selleckchem	473.38	DMSO	50 mM	0-200 µM
Tariquidar	SML1790	Merck	646.73	DMSO	10 mM	0-5 µM
Verapamil	V4629	Merck	491.06	Sterile ddH_2O	10 M	0-100 µM

Table 2.2. Cytotoxic agents used in growth inhibition assays

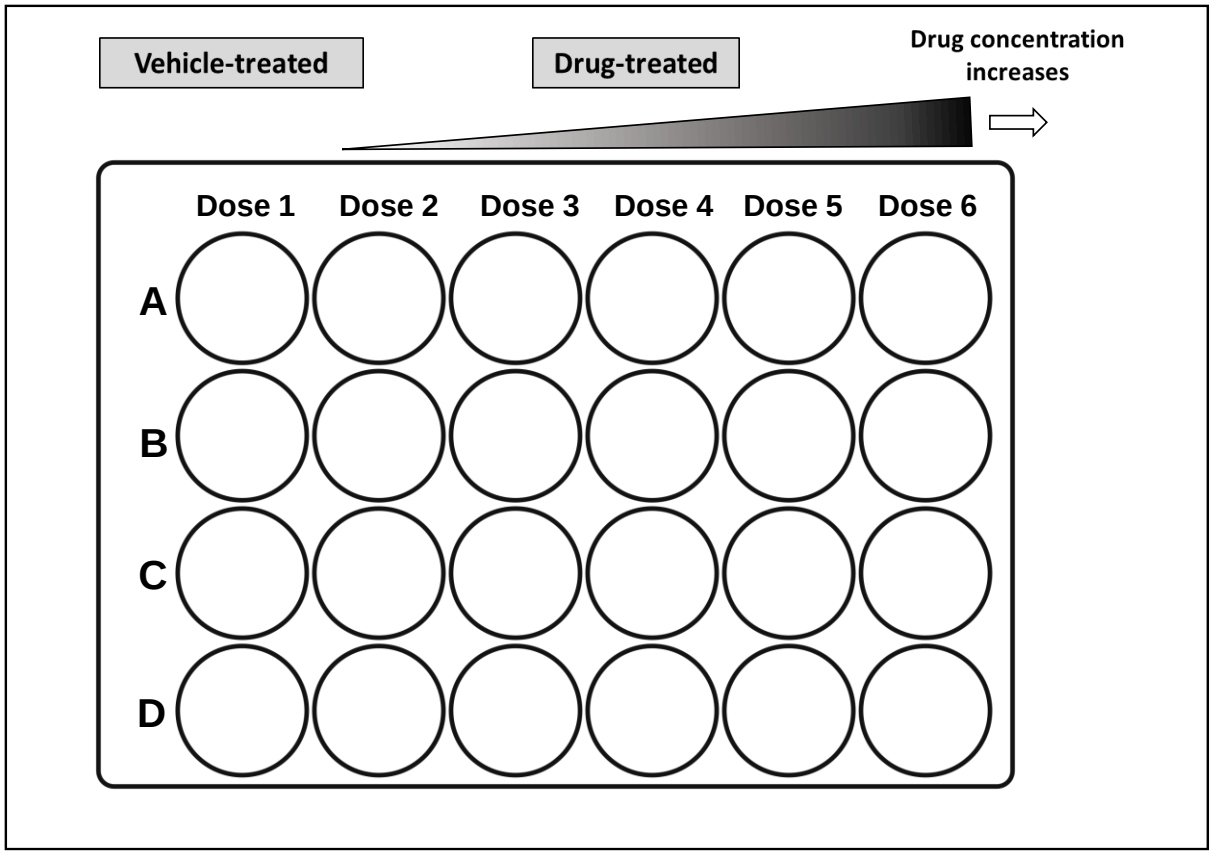

Figure 2.2. Plate design for single drug growth inhibition assay. Each row in the plate corresponds to a different cell clone and each column corresponds to a different drug dose. 250 µl each of cell suspension at a density of $1x10^5$ cells/ml was dispensed into six wells (horizontally) in a row of the plate. Wells in column 1 of the plate (Dose 1) were supplemented with 250 µl of vehicle control and wells in columns 2 to 6 (Dose 2 to Dose 6) with 250 µl each of escalating doses of cytotoxic agent.

2.5.4. Assessment of Cell Viability

After 96 hours of incubation, the number of viable cells and the cell density were determined using trypan blue exclusion method (Section 2.3.2) and cell survival percentage was calculated (Equation 2.2).

$$Cell\ Survival\ (\%) = \frac{Cell\ density\ at\ a\ given\ dose}{Cell\ density\ of\ vehicle\ control}\ x\ 100$$

Equation 2.2. Formula to calculate cell survival percentage

Alternatively, cell survival percentage was determined using CellTiter-Glo Luminescent Cell Viability Assay (Promega, Southampton, UK), which is a metabolic assay that determines the number of viable cells based on quantitation of ATP present (Crouch et al. 1993). The addition of CellTiter Glo reagent results in cell lysis and generation of luminescence that is proportional to the amount of ATP present. Briefly, CellTiter-Glo reagent was reconstituted by transferring the entire contents of the CellTiter-Glo buffer (100 ml) into the lyophilised CellTiter-Glo substrate (both provided by the manufacturer in ready-to-use format) and the resulting enzyme-substrate mixture was gently vortexed, aliquoted, and stored at -20ºC.

This assay was based on a simple 'add-mix-measure' format, where 100 µl each of cell suspension to be counted was added into an opaque-walled 96-well plate (in triplicates). Further, 30 µl of CellTiter-Glo reagent was dispensed into each well of the 96-well plate using a multi-channel pipette and mixed well. Following this, the plate was covered with an aluminium foil and was placed in a shaker for 10 minutes at room temperature. After incubation, the luminescence was recorded using an Omega Plate Reader (BMG Labtech, Aylesbury, UK) and the cell survival percentage was calculated (Equation 2.3). The mean cell survival percentage of *TET2* monoallelic and *TET2* biallelic cell clones were calculated and analysed using two-way ANOVA in Prism (GraphPad version 9).

$$Cell\ Survival\ (\%) = \frac{Luminescence\ signal\ intensity\ at\ a\ given\ dose}{Luminescence\ signal\ intensity\ of\ vehicle\ control} x\ 100$$

Equation 2.3. Formula to calculate cell survival percentage using CellTiter Glo assay

2.6. Clonogenic assay

Three technical replicates each of exponentially growing HEL parental, HELmono1, HELmono2, HELmono3, HELbi1, HELbi2 and HELbi3 were seeded at a density of 500 cells/plate in soft agar [RF10 media supplemented with 0.2% agarose] in a 100mmx20mm dish (Corning). The plates were incubated for 25 days until mature colonies were visible, and colonies were counted following staining with 3-(4,5-dimethylthiazolyl-2)-2,5-diphenyltetrazolium bromide (MTT). The cloning efficiency of individual cells were calculated (Equation 2.4). The mean cloning efficiency of *TET2* monoallelic clones and *TET2* biallelic clones were calculated and analysed using two-way ANOVA in Prism (GraphPad version 6.07).

$$Cloning\ efficiency\ (\%) = \frac{Colonies\ counted}{Cells\ seeded}\ x\ 100$$

Equation 2.4. Formula to calculate cloning efficiency

For determination of cloning efficiency in the presence of azacytidine, *TET2* monoallelic and biallelic mutated cells were seeded at a density of 500 cells/plate in 100mmx20mm dish (Corning) in soft agar supplemented with 5 μM of azacytidine or equivalent volume of vehicle control. Three technical replicates each of aza-treated and vehicle control-treated plates were prepared. Macroscopically visible colonies were counted at day 25 and cloning efficiencies were calculated (Equation 2.4). Relative cloning efficiencies were calculated (Equation 2.5) and plotted as bar graph.

$$Relative\ cloning\ efficiency\ (\%) = \frac{Cloning\ efficiency\ of\ aza-treated\ plates}{Cloning\ efficiency\ of\ vehicle-treated\ plates}$$

Equation 2.5. Formula to calculate relative cloning efficiency in the presence of azacytidine

2.7. 450k methylation array and differential methylation analysis

DNA was extracted from cell pellets using QIAamp DNA Mini kit (Qiagen) and hybridised to the Infinium® HumanMethylation450 Beadchips (Illumina) by Eurofins Genomics (Galten, Denmark) according to the manufacturer's protocol. Data processing and analysis were performed by Dr. Wei-Yu Lin (Newcastle University) based on an established workflow (Maksimovic et al. 2016). Briefly, the methylated (M) and unmethylated (U) probe intensity measurements stored in the raw intensity data (IDAT) files were imported into R, a platform for statistical computing. The minfi Bioconductor package (Aryee et al. 2014) was used for the

calculation of detection *P* (det*P*) values, data normalisation (using the preprocessFunnorm function) and calculation of the β (β = M/(M+U+100)) and M (M = $\log_2$(M/U)) values of the individual CpG probes. Detection *P* values were calculated by comparing the total signal intensity (M+U) with the background signal intensity and the poorly performing probes with det*P* > 0.01 and those having SNPs at the CpG sites were filtered out leaving 410,811 probes in the final dataset.

The significantly differentially methylated probes based on *TET2* mutant allele dosage (*TET2* monoallelic vs biallelic), were identified based on the M values using the limma Bioconductor package (Ritchie et al. 2015) and the resulting *P*-values were adjusted such that the false discovery rate (FDR) is limited to 5%. As such, the probes with resulting FDR-adjusted P-values (P_{adj}) < 0.05 and $|Log_2FC| \geq 2$ were considered as significantly differentially methylated and were represented using volcano plot. Further, a heatmap was generated using unsupervised hierarchical clustering based on M values scaled by standard deviation using R.

2.8. RNA sequencing and differential gene expression analysis

Exponentially growing cells ($5x10^6$ cells) were pelleted (section 2.3.5. and snap-frozen using ice-cold methanol and dry ice. RNA was extracted from the frozen pellets using the RNeasy micro kit (Qiagen) and quantified using a Qubit 2.0 Fluorometer with Qubit RNA BR assay kit (Thermo Fisher Scientific, MA, USA) by the Edinburgh Clinical Research Facility (Edinburgh, UK). The quality control of the samples, library preparation and RNA sequencing were performed on the NextSeq 550 platform (Illumina) by the Edinburgh Clinical Research Facility (Edinburgh, UK). The initial processing of RNAseq data was performed by Dr. Wei-Yu Lin (Newcastle University). STAR aligner was used to accurately align, map and annotate the RNAseq reads to human genome build 37 (hg19). The gene features of the aligned reads were characterised using the Rsubread package (using featureCounts function) (Liao et al. 2019) in R (v3.5.1). Read counts were normalised and expressed as counts per million (CPM) (Equation 2.6).

$$Counts\ per\ million\ (CPM) = \frac{Raw\ counts}{library\ size} x10^6$$

Equation 2.6. Formula for counts per million (CPM)

Differential gene expression analysis was performed using an interactive platform, Differential Expression Browser (DEBrowser) that uses multiple Bioconductor packages such as DEseq2 (Love et al. 2014), EdgeR (Robinson et al. 2010), limma (Ritchie et al. 2015) and shiny (Chang

2016) to bring out real-time changes in the plots and helps visualise the data with a variety of analysis tools. Briefly, pre-processed data stored in the form of count data and metadata file was uploaded as a .TSV file. The count data contains the read count data for each gene mapped to the human genome, whereas the metadata establishes the batch effect of multiple conditions, in this case the conditions are *TET2* monoallelic vs biallelic.

Initially an unsupervised hierarchical clustering was performed using the 'Go to QC plots' option in the DEBrowser, which visualizes the quality control (QC) matrix of the data by generating heatmaps and principal component analysis (PCA) plots based on the top 1500 transcripts with the highest variance across all samples. Following the unsupervised clustering, a supervised analysis was performed to identify the differentially expressed genes or pathways based on *TET2* status using DEseq2 Bioconductor package (Love et al. 2014) in DEBrowser and this was done by manually grouping the samples to be compared under the two conditions (*TET2* monoallelic vs biallelic). The resulting *P* values were adjusted (P_{adj}) to limit the FDR to 5% and those genes with P_{adj} value < 0.05 and $|Log_2FC| \geq 0.3$ were defined to be significantly differentially expressed. Heatmap and volcano plot were generated to represent the significantly up and downregulated genes. Selecting any gene from the volcano plot retrieved a bar plot showing the gene expression profile (read counts) of the selected gene across the sample set. Alternatively, a gene name can be typed in the 'search' column to retrieve the gene expression profile as a bar plot.

In addition, the 'enrichGO' function under the 'GO Term' tab in the DEBrowser was used to perform gene ontology analysis and to identify critical pathways associated with the significantly differentially expressed transcripts. From the resulting list of pathways, a pathway name was selected and the 'DE Genes' option was selected to view and export the read count data of differentially expressed genes of the selected pathway into an excel file. Further, the gene names in the resulting excel file was used to construct pathway specific heatmaps by pasting the list of gene names in the 'Search' column in the 'QC Plots' tab and selecting 'searched' under the 'Choose dataset' option.

2.9. Identification of interesting differentially expressed transcripts

The 'Download data' option in the DEBrowser allowed for export of data including differentially expressed gene names, read counts across the sample set, *Padj* and $|Log_2FC|$ values into an excel file. Further, a literature search was performed using NCBI PubMed and Google, on the resulting differentially expressed genes to identify evidence for a role in determining response to azacytidine. Keywords used for search were "gene name and

azacytidine", "gene name and chemoresistance", "gene name and chemosensitivity", "gene name and leukaemia", "gene name and chemotherapy" and "gene name and cancer". Those significantly differentially expressed genes with evidence for a role in determining azacytidine response were taken forward for protein expression analysis using western blot and further functional studies.

2.10. Western immunoblotting

Western blot was performed to determine if the levels of specific proteins differ in *TET2* monoallelic vs biallelic mutated HEL cells (Chapter 3, Chapter 4, and Chapter 5), azacytidine resistant clones vs their parental counterparts (Chapter 5) and to test the TET2 protein levels in different AML cell lines (Chapter 6). As such, the amount of antibody binding the target protein and hence the signal intensity of the protein bands are directly proportional to the amount of respective protein present in the sample. Protein levels as determined by western blot were compared with the corresponding transcript level expression as determined by RNA sequencing, where possible. The various stages of a western immunoblot are as follows:

2.10.1. Preparation of protein samples for western blot

Viable cell counts were obtained using trypan blue exclusion method. $5x10^6$ cells were pelleted by centrifugation at 250g and washed twice with pre-cooled PBS. Cell lysis was achieved through either of the three lysis buffers such as Sodium dodecyl sulphate (SDS) buffer, Radio-Immunoprecipitation Assay (RIPA) buffer or PhosphoSafe buffer (when it is required to preserve the phosphorylated state of the protein of interest).

(i) Cell lysis by SDS buffer

Washed pellets were resuspended in 100 µl SDS sample buffer (62.5mM Tris-HCl pH 6.8, 2% (w/v) SDS, 20% (v/v) glycerol). Samples were homogenised by passing through a 21-gauge needle and heated at 100°C for 5 minutes. The samples were then centrifuged at 14000 x g for 10 minutes at 4ºC and the supernatant was transferred to a labelled fresh Eppendorf.

(ii) Cell lysis by RIPA buffer

Washed pellets were resuspended in 1 ml of RIPA buffer (Scientific Laboratory Supplies, Nottingham, UK) (25 mM Tris•HCl pH 7.6, 150 mM NaCl, 1% NP-40, 1% sodium deoxycholate, 0.1% SDS) and mixed by pipetting. The resulting mixture was kept in shaker for

30 minutes at 4ºC, following which, the tubes were centrifuged at 14000 x g for 20 minutes at 4ºC and the supernatant was transferred to a labelled fresh Eppendorf.

(iii) Cell lysis by PhosphoSafe buffer

Washed pellets were resuspended in 150 µl PhosphoSafe buffer (Merck) and incubated at room temperature for 10 minutes. The tubes were centrifuged at 14000 x g for 5 minutes at 4ºC. The supernatant was transferred to a labelled fresh Eppendorf.

Following protein extraction, the total protein concentration was determined using the Pierce BCA kit and the unused samples were stored at -20°C. Samples were stored at -80ºC in case of PhosphoSafe buffer extracted protein samples.

2.10.2. Estimation of protein concentration through Pierce BCA Assay

The Pierce BCA assay is based on the reduction of copper cations (Cu^{+2} to Cu^{+1}) by protein and each Cu^{+1} forms a coloured coordination complex with four to six peptide bonds. Further, the bicinchoninic acid (BCA) which is a highly sensitive and selective colorimetric detection reagent, forms a water-soluble BCA-copper complex (purple colour) with the Cu^{+1} ions formed from the first reaction and shows a strong absorbance at 562 nm with increasing protein concentrations. As such, the intensity of colour is directly proportional to the protein concentration.

The assay was performed according to the manufacturer's protocol. Briefly, protein samples were thawed on ice and an aliquot of the sample was diluted 1:5 using ddH_2O to bring the protein concentrations within the detection range of the assay (20 to 2,000 µg/mL). After diluting, 10 µl of the sample was dispensed in triplicates into a 96-well optical plate (VWR International Ltd.). BCA working solution (green coloured) is a two-component system which was prepared by combining 50 parts of Reagent A (carbonate buffer containing BCA reagent) with 1 part of Reagent B (a cupric sulphate solution). After mixing, 190 µl of this solution was added into each well of the 96-well optical plate and the well contents were mixed with pipette. The colour of the BCA working solution changed from green to purple when mixed with protein samples. The plate was incubated at 37ºC for 30 minutes, following which the absorbance at 562 nm was measured using an Omega plate reader (BMG Labtech). A bovine serum albumin (BSA) standard curve was prepared each time the assay was performed, and the protein concentrations were estimated from the mean absorbance values by comparison with the

standard curve. The estimated protein concentrations were multiplied by 5 to account for the dilution factor to obtain the protein concentration of each extract.

2.10.3. SDS-PAGE and Electrophoretic Transfer

Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) is used to separate the proteins based on molecular weight via electrophoresis through a polyacrylamide gel. Protein samples were diluted to a final concentration of 20 to 40 µg/well (depending on efficiency of the antibody used) using ddH_2O and was made up to 13 µl total volume. The other components for sample preparation were added according to Table 2.3. After gentle vortexing, the tubes with the samples were pre-heated to 70ºC for 10 minutes in a heating block, following which, the samples were pulse centrifuged. Finally, 20 µl of the resulting sample was loaded into the wells of NuPAGE 4-12% Bis-Tris gel (Invitrogen) and 10 µl of either PageRuler™ Plus Prestained Protein Ladder (ThermoFisher Scientific) or Precision Plus Protein Dual Color Standards (Bio-Rad) was added into the first well. Electrophoresis was performed at 80V for the first 20 minutes (until all the bands align at the same position) and then 150 V for another 1 hour in a tank filled with Bis-Tris running buffer (5% (v/v) NuPAGE MOPS SDS running buffer (50 mM MOPS, 50 mM Tris Base, 0.1% SDS, 1 mM EDTA, pH 7.7), 0.1% (v/v) NuPAGE antioxidant and dH2O made up to a total volume of 500ml).

The gel with separated proteins was carefully placed on a nitrocellulose membrane (Novex 0.45µM nitrocellulose membrane filter paper sandwiches) within a transfer cassette. The cassette was placed in a transfer tank filled with transfer buffer (5% (v/v) NuPAGE 20X SDS transfer buffer, 0.1% (v/v) NuPAGE antioxidant, 10% (v/v) methanol and dH2O made up to a total volume of 1000ml) and electrophoresed at a constant voltage of 100V using a magnetic stirrer to maintain ion distribution in the buffer.

Component	Volume per well	Final concentration
Protein extract	13 µl	20 - 40 µg *
NuPAGE LDS Sample buffer (4X) (NP0007)	5 µl	1X
NuPAGE Reducing agent (10X) (NP0004)	2 µl	1X
Total volume	**20 µl**	

Table 2.3. Sample preparation for SDS-PAGE. The sample buffer contains lithium dodecyl sulphate (LDS) at pH 8.5 with SERVA Blue G250 and phenol red. The reducing agent contains 500 mM dithiothreitol (DTT)).
* The protein concentration was decided based on the antibody efficiency.

2.10.4. Antibody detection of bound proteins and visualisation

After transfer, the membrane with bound proteins was removed from the transfer cassette and was incubated in appropriate blocking buffer (5% (w/v) skimmed milk powder or BSA dissolved in TBS-Tween (0.01M Tris-HCl pH 7.5, 0.1M NaCl, 0.05% (v/v) Tween-20) for 1 hour at room temperature. The same buffer was used for blocking as well as dilution of primary and secondary antibodies and was selected as per the antibody manufacturer's recommendations. The complete list of antibodies used, and the respective dilutions are provided in Table 2.4.

Following blocking, the membrane was transferred to a 50 ml BD Falcon containing 5 ml (adequate volume to cover the membrane) of primary antibody and was subsequently probed using the respective primary antibody overnight at 4°C with gentle agitation (using a roller mixer). All the primary antibody solutions were stored at -20ºC after each use and were re-used for a maximum of 5 times, after which fresh dilutions were made.

After primary antibody incubation, the membranes were washed three times with 5 ml of TBS-Tween for 10 minutes each. The membrane was subsequently probed with an appropriate secondary antibody (specific to the host species in which the primary antibody was raised) for 1 hour at room temperature followed by another set of washes. Finally, the membrane was coated with Amersham ECL detection agent (GE healthcare Life Sciences) for 1 minute and the bands were visualised using ChemiDoc XRS+ system (Bio-Rad Laboratories Ltd.). Exposure time varied depending on the efficiency of antibody.

2.10.5. TET2 western blot

For TET2 western blot analysis, protein extracts were separated using NuPAGE® 3-8% Tris acetate gels (Invitrogen) and electrophoresis in Tris acetate running buffer (5% (v/v) NuPAGE® Tris acetate SDS running buffer, 0.1% NuPAGE® antioxidant and dH_2O made up to a total volume of 500ml). HiMark™ pre-stained protein standard (Invitrogen) was used as marker. Since the molecular weight of TET2 is high (224kDa), electrophoresis was carried out for a longer duration until the high molecular weight bands of the marker were well separated. For transfer, PVDF membranes were used because these are optimised for higher molecular weight proteins and also possess higher protein binding capacity compared to nitrocellulose membrane. Transfer was performed overnight at 4°C at a constant voltage of 20V using a magnetic stirrer to maintain ion distribution in the buffer. The membrane was blocked with 5% (w/v) skimmed milk in TBS-Tween and then the blot was incubated with anti-TET2 antibody overnight at 4°C,

and later with secondary antibody for 1 hour at room temperature. Protein bands were visualised using ECL detection agent (Amersham) and ChemiDoc XRS+ system (Bio-Rad Laboratories Ltd.).

2.10.6. Protein expression quantification and data analysis

Protein quantification was performed on immunoblots using the Image Lab Software (Bio-Rad Laboratories Ltd.). Target protein intensity was quantified and normalised to the respective intensities of loading control (α-TUBULIN/GAPDH). Protein intensity in *TET2* monoallelic and biallelic mutated clones were grouped and represented through a dot plot. Statistical significance was tested using student's *t*-test (GraphPad Prism).

2.11. Functional validation of interesting targets

The genes and pathways that showed significant differential expression based on *TET2* mutant allele dosage at transcript and protein levels were taken forward for functional validation using the following strategies:

2.11.1. Estimation of the functional impact of ribosomal pathway downregulation in TET2 biallelic mutants

Considering the role of ribosomal pathway in protein synthesis, total protein concentration was quantified using the Pierce BCA assay in *TET2* monoallelic (n=2) and biallelic (n=2) mutant HEL cell clones. Briefly, $5x10^6$ exponentially growing cells were pelleted via centrifugation, washed with PBS, and lysed using RIPA buffer (Section 2.10.1). Total protein concentration was estimated using the BCA assay (Section 2.10.2). Five independent replicates of the experiment were performed, and each independent replicate consisted of three technical replicates for each cell clone. The mean protein concentrations were calculated based on *TET2* mutant allele dosage (*TET2* monoallelic vs biallelic) and the resulting *P*-values were calculated using paired *t*-test (GraphPad Prism).

Protein	Molecular Weight (kDa)	Antibody Type	Clone	Host	Supplier	Catalogue No.	Dilution	Dilution buffer*
Primary antibodies								
α-Tubulin	50	Monoclonal	DM1A	Mouse	Sigma-Aldrich	T9026	1:5000	Milk
ABCB1	170	Monoclonal	G-1	Mouse	Santa cruz	SC-13131	1:100	Milk
CDK7	42/37	Monoclonal	C-4	Mouse	Santa cruz	SC-7344	1:500	BSA
GAPDH	37	Monoclonal	411	Mouse	Santa cruz	SC-47724	1:1000	Milk
HnRNPK	65	Monoclonal	D-6	Mouse	Santa cruz	SC-28380	1:5000	Milk
IDH1	46	Polyclonal	-	Rabbit	Abcam	ab113232	1:1000	Milk
IDH2	48	Monoclonal	5F11	Mouse	Abcam	ab55271	1:1000	Milk
LSM8	10	Monoclonal	F-8	Mouse	Santa cruz	SC-390542	1:200	BSA
NSUN1	89	Monoclonal	E-7	Mouse	Santa cruz	SC-398884	1:500	Milk
POLR2A S2P	250	Monoclonal	E1Z3G	Rabbit	Cell Signaling	13499S	1:1000	BSA
RPL19	23	Monoclonal	K-12	Mouse	Santa cruz	SC-100830	1:500	BSA
RPL22	15	Monoclonal	D-7	Mouse	Santa cruz	SC-373993	1:200	BSA
RPS6	32	Monoclonal	5G10	Rabbit	Cell Signaling	2217S	1:1000	BSA
RPS14	16	Monoclonal	3G5	Mouse	Santa cruz	SC-293478	1:500	BSA
RPS19	16	Monoclonal	WW-4	Mouse	Santa cruz	SC-100836	1:500	BSA
TET2	224	Monoclonal	Mab-179-050	Mouse	Diagenode	C15200179	1:500	Milk
WT1	51	Polyclonal	-	Rabbit	Sigma-Aldrich	HPA053848	1:500	Milk
Secondary antibodies								
Goat Anti-Mouse immunoglobulins/HRP	-	Polyclonal	-	Goat	Agilent Technologies	P044701-2	1:1000	-
Goat Anti-Rabbit immunoglobulins/HRP	-	Polyclonal	-	Goat	Agilent Technologies	P044801-2	1:1000	-

Table 2.4. List of antibodies used for western blot.* The dilution buffer was selected based on antibody manufacturer's recommendations and was prepared by dissolving 5% (w/v) of skimmed milk/BSA in TBS-Tween. The secondary antibodies were dissolved in the same dilution buffer as that of the respective primary antibodies.

BSA, Bovine Serum Albumin; kDa, kilodalton

2.11.2. ABCB1 Functional study 1: Testing the effect of ABCB1 inhibition on azacytidine sensitivity

The aim of the experiment was to investigate if ABCB1 inhibition leads to an increase in azacytidine sensitivity. To test this hypothesis, a multi-drug growth inhibition assay was performed, where *TET2* monoallelic (n=2) and biallelic (n=2) mutant cell clones were treated with a combination of increasing doses of azacytidine and an ABCB1 inhibitor and the drug synergy was evaluated. Verapamil and Tariquidar were the ABCB1 inhibitors used in this assay.

(i) Setting up of 24-well plates

TET2 monoallelic (n=2) and biallelic (n=2) mutant cell clones were treated with a combination of increasing doses of azacytidine and an ABCB1 inhibitor. Verapamil and Tariquidar were the ABCB1 inhibitors used in this assay and drug stocks were prepared according to Table 2.2. Exponentially growing cells were seeded at a density of $1x10^5$ cells/ml in RF10 media and made up to 10 ml in a BD Falcon tube. After mixing, 250 µl of the resulting cell suspension was dispensed into all the wells of a 24-well plate (Corning), such that each plate was used for a single cell clone. The plates were then incubated at 37°C and 5% CO_2 until the cytotoxic agents were prepared.

Working solutions of cytotoxic agents were prepared by dilution of the stocks in RF10 media followed by preparation of a drug serial dilution by adding an appropriate volume of RF10 media to prepare six different doses of azacytidine and 4 doses each of verapamil (Figure 2.3A) and tariquidar (Figure 2.3B). A vehicle control was prepared for each drug where an appropriate solvent (without drug) was diluted in RF10 media. After thoroughly mixing the serially diluted drug solutions, 125 µl each from the respective drug solutions were added into the wells of the plate according to the plate design (Figure 2.3). The final cell density after adding the drugs was $5x10^4$ cells/ml. The plates were incubated at 37°C and 5% CO_2 for 96 hrs.

(ii) Assessment of cell viability and data analysis

After 96 hours of incubation, cell viability was estimated using CellTiter-Glo Luminescent Cell Viability Assay (Promega) (Section 2.5.4.). The resulting dose-response matrix was used to calculate drug synergy using SynergyFinder 2.0 (Ianevski et al. 2020). Briefly, the excel file consisting of the dose-response matrix was uploaded to the SynergyFinder 2.0 platform and viability was selected as the readout from the drop-down list. The synergy scores and the most synergistic area plots were generated for each drug combinations (Azacytidine-Verapamil and

Azacytidine-Tariquidar), and the synergy scores were tabulated. Further, the mean synergy scores were calculated based on *TET2* mutation status and the resulting *P*-values were estimated using ANOVA (GraphPad Prism). Finally, the pattern of synergy scores for both the drug combinations were compared with the pattern of ABCB1 protein expression based on *TET2* mutant allele dosage.

2.11.3. ABCB1 Functional study 2: Investigating the ABCB1 protein expression in azacytidine resistant clones

This experiment was aimed at generating azacytidine resistant clones from *TET2* monoallelic and biallelic mutant cell clones and testing the ABCB1 expression pattern in the resistant clones compared to the parental cells from which the clones were made. The experiment consists of the following stages:

(i) Generation of azacytidine resistant clones

HEL cells (*TET2* monoallelic and biallelic mutants) were plated at a starting density of 1000 cells/plate in soft agar in the presence of high dose azacytidine (10 µM; which corresponds to 90%-95% cytotoxicity). Colonies were picked after 28 days and expanded in drug-free liquid media in 96-well dishes (Figure 2.4 and Table 2.5). Cells were subsequently expanded and maintained in 5 µM azacytidine in order to establish putative chemoresistant clones.

Cell clone	No. of colonies		No. of colonies picked	No. of colonies successfully expanded
	Plate 1	Plate 2		
HELmono1	0	0	0	0
HELmono2	31	19	10	3
HELmono3	22	20	10	1
HEL Parental	17	-	10	2
HELbi1	4	5	9	3
HELbi2	0	0	0	0
HELbi3	6	3	9	0

Table 2.5. Number of colonies that appeared, picked and successfully expanded after 28 days of azacytidine treatment.

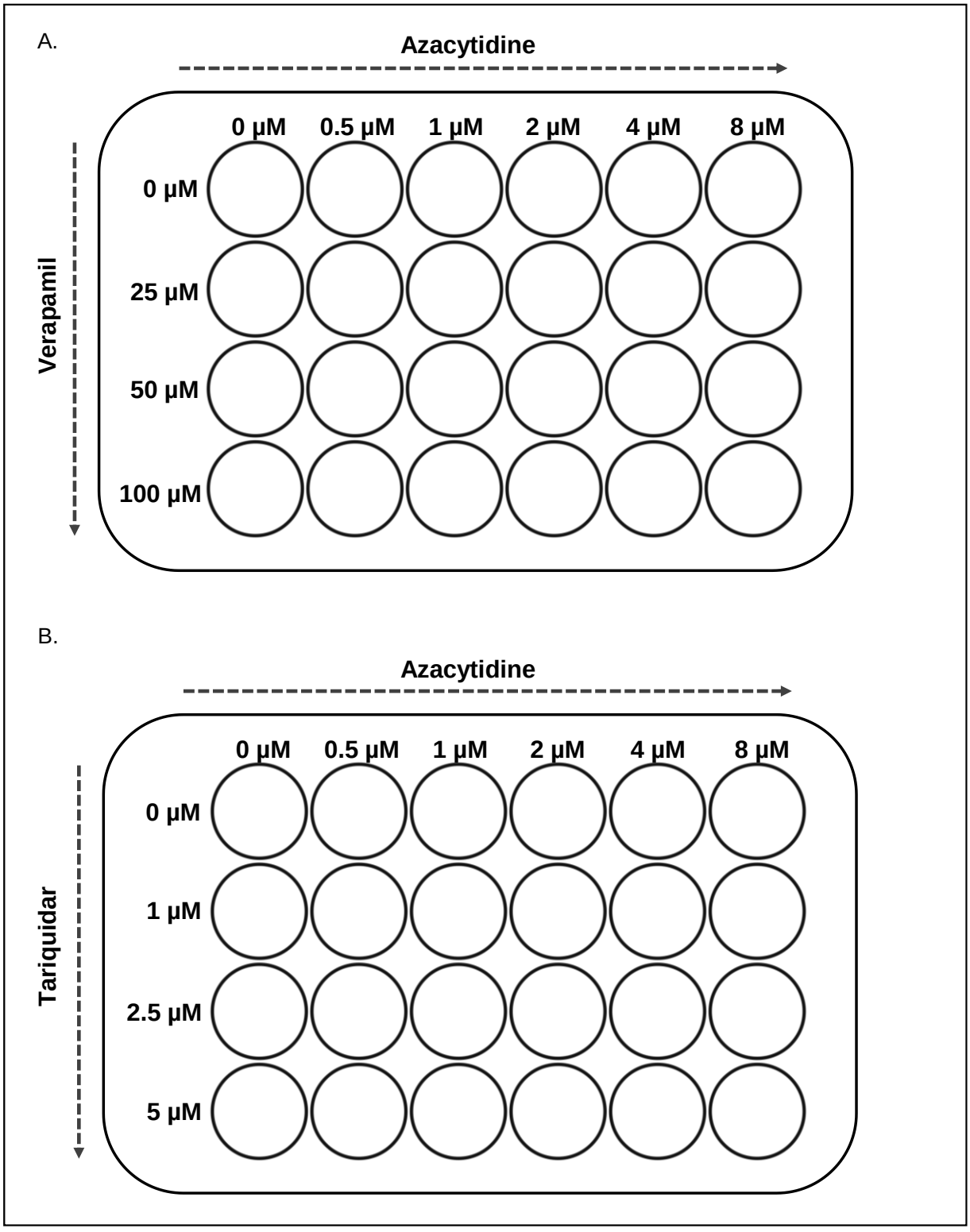

Figure 2.3. Plate design for multi-drug growth inhibition assay. Cells were treated with (A) a combination of azacytidine (0-20 µM) and verapamil (0-100 µM) and (B) a combination of azacytidine (0-20 µM) and tariquidar (0-5 µM). Each 24 well plate corresponds to a single cell line. Cell viability and drug synergy were calculated 96 hours post drug treatment.

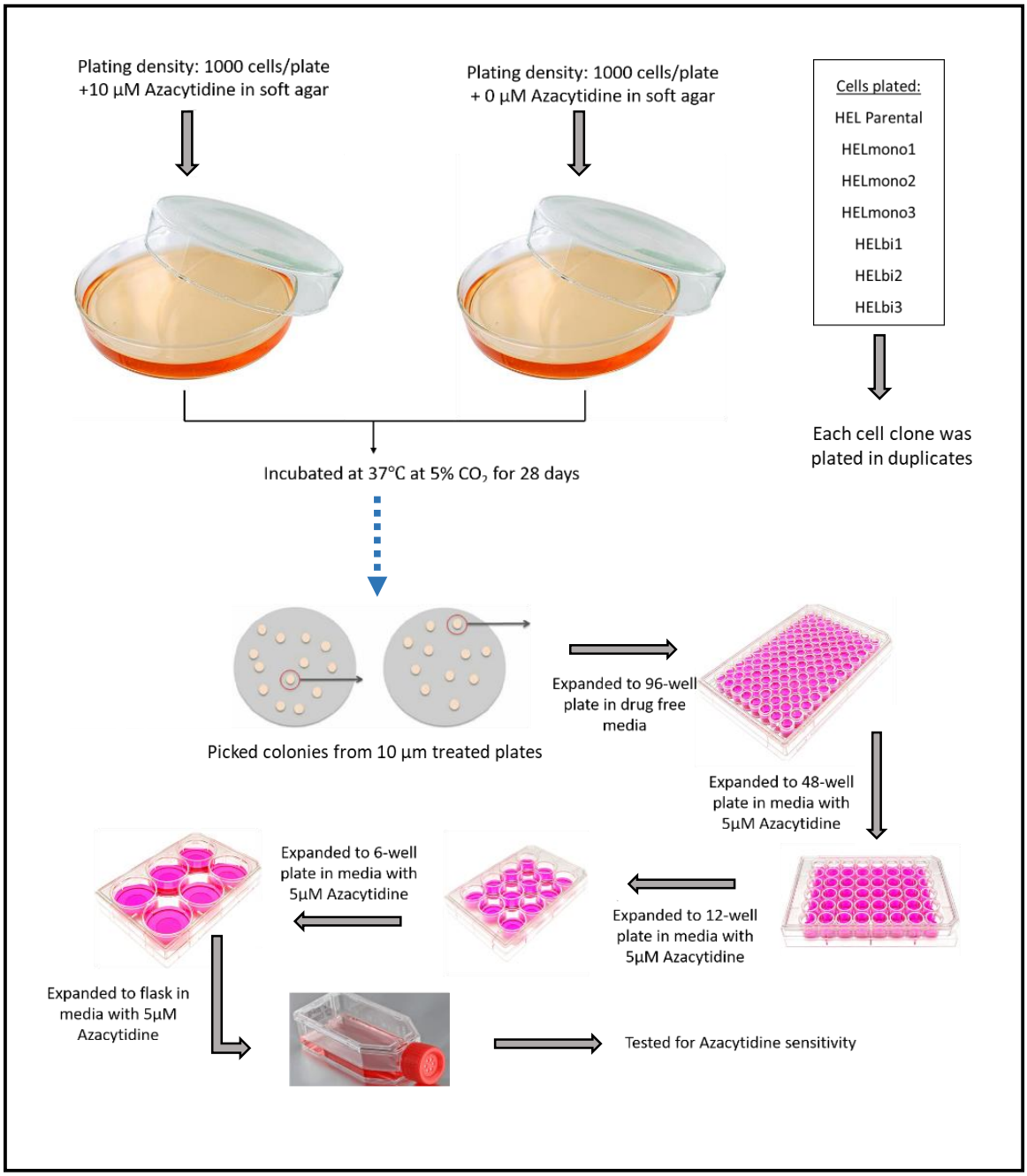

Figure 2.4. Protocol used to develop azacytidine resistant clones

Following expansion into an adequate volume, each cell clone was tested for sensitivity to azacytidine along with the parental cell line from which the putative azacytidine resistant (PAR) clone was developed. Growth inhibition assay (Section 2.5) was performed in the presence of escalating azacytidine doses to test the azacytidine resistance phenotype of the PAR clones. The cell viability was estimated using trypan-blue exclusion method, and the dose-response graphs and the respective IC50 values were plotted using GraphPad Prism. Two-way ANOVA (GraphPad Prism) was used to test for significant difference in azacytidine-response between the PAR clones and the respective parental cells.

(i) Western blot to characterise the ABCB1 expression in azacytidine resistant clones and their parental cells

After testing the azacytidine resistance phenotype, protein level expression of ABCB1 was tested in azacytidine resistant clones along with their respective parental cells using western immunoblotting (Section 2.10). ABCB1 expression was quantified using Image Lab Software (Bio-Rad Laboratories Ltd.) and the resulting ABCB1 band intensities in PAR clones were normalised to the ABCB1 band intensity in the respective parental cells.

(ii) Growth inhibition assay to test the effect of ABCB1 inhibition on azacytidine response of azacytidine resistant clones

Exponentially growing *TET2* monoallelic (n=3) and biallelic (n=3) mutated PAR clones along with their respective parental cells were seeded at a cell density of $1x10^5$ cells/ml using RF10 media. After mixing, 500 µl of the resulting cell suspension was dispensed into each well of a 12 well plate. Stock solutions of azacytidine and verapamil were prepared according to Table 2.2. Working solutions were prepared in four 15 ml BD Falcon tubes, one each for vehicle control (without drug), single agent azacytidine (10 µM), single agent verapamil (100 µM) and combination of azacytidine (10 µM) and verapamil (100 µM). After mixing, 500 µl of the resulting drug solutions were dispensed into respective wells (in triplicates) of the 12 well plate, according to the plate design (Figure 2.5). The plates were then incubated at 37°C and 5% CO_2. Cell viability was determined after 96 hours incubation using trypan blue exclusion method and the cell survival percentage was calculated as the percentage of viable cells under each treatment condition compared to the vehicle-treated cell viability. Three independent replicates of the experiment were performed and the resulting mean cell survival percentage for each cell

clone along with the standard deviation from the mean was depicted as a bar plot using GraphPad prism and the *P*-value was computed using two-way ANOVA (GraphPad Prism).

2.12. Estimation of correlation between TET2 protein levels and azacytidine sensitivity

A panel of 10 AML cell lines (AML2, AML3, HEL, HL-60, KASUMI-1, MV4-11, NB4, SKM1, THP1 and U937) was used to determine whether 5'-azacytidine sensitivity correlates with TET2 protein expression. Western immunoblotting (Section 2.10.5) was used to determine the TET2 protein levels in the 10 AML cell lines. The resulting protein bands were quantified using Image Lab Software (Bio-Rad Laboratories Ltd.). Growth inhibition assay was performed (Section 2.5) with escalating doses of azacytidine. After 96 hours of drug treatment, cell viability was determined using CellTiter-Glo Luminescent Cell Viability Assay (Promega) (Section 2.5.4). IC50 and IC90 values, corresponding to 50% and 90% growth inhibition respectively in response to azacytidine, were determined by calculating the interpolated X-values (drug doses) using Fit Spline-Lowess algorithm in GraphPad Prism. The resulting azacytidine IC50 and IC90 values were plotted against the corresponding TET2 protein intensity. The respective correlation co-efficient (R^2 values) along with the statistical significance *P*-value were computed using the linear regression program in GraphPad Prism.

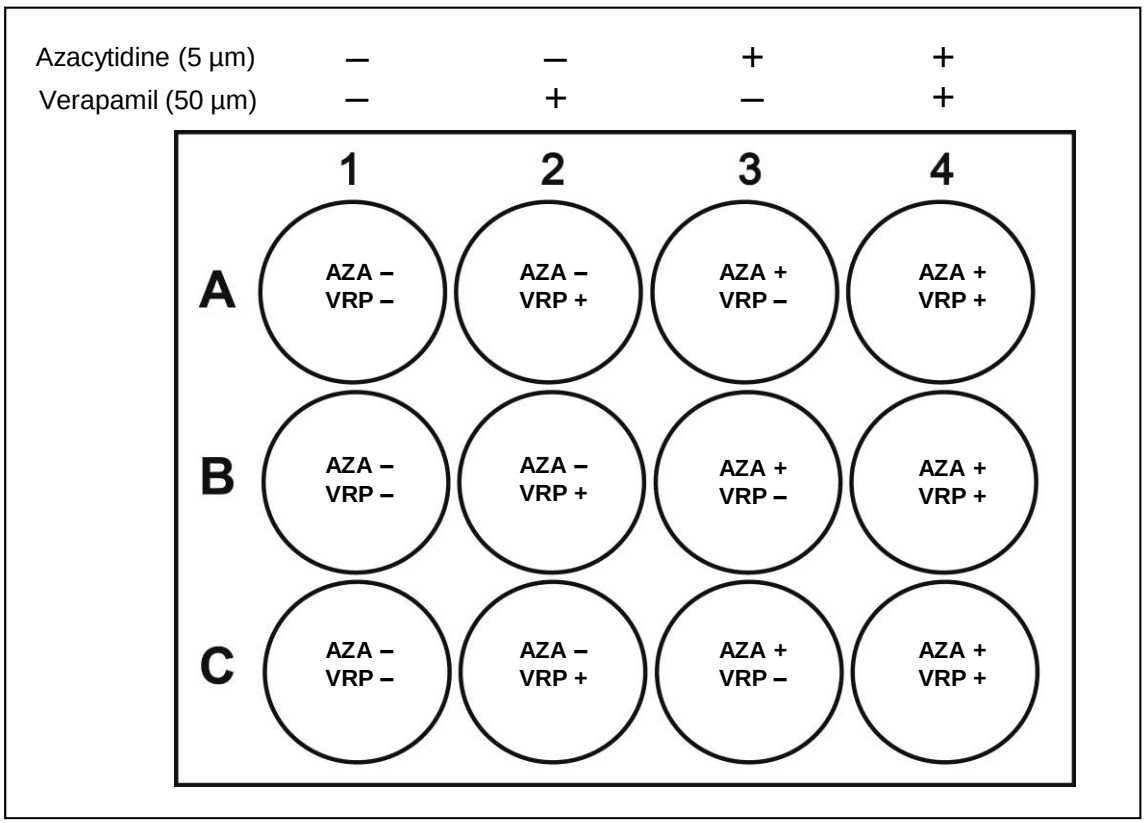

Figure 2.5. Plate design for azacytidine-verapamil (single dose) growth inhibition assay. Rows A, B and C denote different cell lines whereas columns 1, 2, 3 and 4 denote different drug treatment conditions. Cells in column 1 of the plate were treated with vehicle control (no drug), column 2 with single-agent azacytidine (5 µM), column 3 with single-agent verapamil (50 µM) and column 4 with a combination of azacytidine (5 µM) and verapamil (50 µM).

Aza, azacytidine; VRP, verapamil

Chapter 3. Phenotypic characterisation of *TET2* biallelic mutated cells

3.1. Introduction

Acute myeloid leukaemia (AML) is a clinically heterogeneous disease driven by a plethora of somatic mutations and chromosomal abnormalities. The epigenetic landscape of cells is often perturbed during leukaemogenesis with alterations in DNA methylation and histone modifications leading to oncogene expression or silencing of tumour suppressors (Venney et al. 2021). Among the various epigenetic modifiers, TET2 has a key role in DNA demethylation by catalysing the sequential conversion of 5-mC to 5-hmC which serves to de-repress silenced genes (Feng et al. 2019) (Section 1.4.2 and 1.4.3). Considering the important role played by TET2 in regulating haematopoiesis, it is not surprising that *TET2* is often inactivated in myeloid and lymphoid malignancies through loss-of-function mutations, deletions, or inactivating mutations in TET2 regulators such as *IDH1/IDH2* (Solary et al. 2014).

3.1.1. TET2 mutant allele dosage in AML

TET2 is mutated in up to 23% of AML, with the majority of these being monoallelic and approximately 25% biallelic mutation (Ahn et al. 2015). *TET2* mutations mostly affect the conserved domains and N-terminus region of TET2 (Jankowska et al. 2009). Homozygous mutations are typically associated with uniparental disomy and heterozygous mutations with CN-AML (Feng et al. 2019; Jankowska et al. 2009).

The effect of *TET2* mutant allele dosage on AML cell phenotype and pathogenesis has already been discussed in detail (Section 1.4.3). Briefly, methylation profiling in isogenic AML cell lines generated by *TET2* knockout in THP1, and KG-1 cells revealed acquisition of a hypermethylated phenotype with complete loss of *TET2* compared to their *TET2* wild-type counterpart (Morinishi et al. 2020). Animal models demonstrate a role for *TET2* deletions in driving myeloid and lymphoid transformations. Although heterozygous *TET2*+/- mice, which leads to ~50% reduction in *TET2* gene expression (compared to *TET2* wild-type), develop haematological malignancies similar to null *TET2*-/- mice, the latency to disease development is longer in the heterozygous animals (Solary et al. 2014; Feng et al. 2019), suggesting a role for mutant allele dosage in malignant transformation. Despite this, there is limited evidence on the impact of mutant allele dosage on leukaemogenesis and AML therapeutic response. Therefore, this chapter aims to phenotypically characterise the impact of *TET2* mutant allele dosage based on an isogenic cell model system.

3.1.2. Modelling a biallelic TET2 mutation in vitro

(i) Identification of an appropriate cell line for study

TET2 mutations are often reported as an unfavourable prognostic marker associated with early relapse and poor overall survival (Metzeler, Maharry, Radmacher, Mrózek, et al. 2011; Abdel-Wahab et al. 2009). As such, there is an urgent clinical need to develop better therapeutic interventions to treat *TET2* mutated AML. In order to investigate this further, we required an AML cellular model which allows for comparison between monoallelic and biallelic *TET2* mutants in an isogenic setting.

Establishing such a cellular model requires a cell line derived from a leukaemia which developed in the background of a *TET2* monoallelic mutation, such that targeting of the retained wild-type *TET2* allele would generate an isogenic cell line pair. Another requirement for cell line selection was the absence of mutations reported in DNA methylation related genes such as *IDH1, IDH2, WT1, DNMT3A or DNMT3B,* which might phenocopy the *TET2* loss and impact our studies. Therefore, HEL (Human erythroleukaemia) cell line was identified as the best candidate for this study because of the large deletion reported on chromosome 4 capturing the *TET2* locus. Moreover, there is absence of other DNA methylation related gene mutations in HEL AML cells (Martin et al. 1982; Abdel-Wahab et al. 2009).

(ii) CRISPR-Cas9 based gene editing to generate biallelic TET2 mutation

The Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR-Cas9) type II system is a novel and well characterised gene editing tool used for precise gene editing in human cells (Khatodia et al. 2016). The system consists of two major components, namely the CRISPR guide RNA (gRNA) which is specific to target sequence within the gene of interest, and the Cas9 endonuclease which causes double strand breaks in the target site when guided by the gRNA (Khatodia et al. 2016). The gRNA is followed by a protospacer-adjacent motif (PAM) which is specific to the CRISPR system used. For example, for CRISPR derived from *Streptococcus pyogenes*, the PAM sequence is 5'-NGG-3' and the Cas9 induced double strand break will be within the immediate vicinity (within 3-4 base pairs upstream) of the PAM (Ran et al. 2013). Upon base pairing of the CRISPR gRNA with the target sequence, the Cas9 endonuclease is directed to the target site, which causes double strand breaks adjacent to the PAM sequence. Following cleavage, the cell attempts to repair the double strand breaks through either of the two major DNA repair pathways, namely the homology-directed repair (HDR) pathway or non-homologous end joining (NHEJ). If a specific mutation is required, a

homologous repair template with the desired mutation can be provided and the particular mutation will be incorporated as part of the HDR pathway. However, in the absence of a repair template, the target site undergoes re-ligation through the NHEJ repair pathway, which is known to be error-prone resulting in random insertions or deletions (Indels) (Ran et al. 2013). Indels within coding exons can result in frameshift mutations and formation of premature stop codons which can lead to gene editing and silencing (Ran et al. 2013). For the purpose of this project, it is sufficient to generate indels (without providing a repair template) such that the retained *TET2* allele is made non-functional, thus generating an isogenic cell line pair with monoallelic *TET2* mutation (in parental HEL cells) or biallelic *TET2* mutation, with complete loss of TET2 protein expression in the latter (in CRISPR treated HEL cells).

3.2. Aims of the chapter

Despite the occurrence of *TET2* biallelic mutations in AML and its key impact on clonal haematopoiesis, there have been no published study interrogating the effect of *TET2* mutant allele dosage on sensitising AML to any particular chemotherapeutic agent. Therefore, this chapter aims to phenotypically characterise the effect of *TET2* biallelic mutation on chemosensitivity to established therapeutic strategies of AML. This was accomplished using an isogenic cell model system based on *TET2* monoallelic mutated HEL cells. *TET2* biallelic mutated HEL cells were generated by completely inactivating the retained *TET2* allele using CRISPR-CAS9 gene editing.

Specifically, the experimental aims of this chapter are as follows:

- Generate *TET2* biallelic mutated HEL cells using CRISPR-CAS9 and to validate mutation status using Sanger sequencing.
- Determine the effect of *TET2* biallelic mutation on cellular TET2 protein levels and its impact on cell proliferation and clonogenicity.
- Evaluate the cytotoxic response of *TET2* biallelic mutated cells to established chemotherapeutic drugs widely used in the treatment of AML.
- Determine if *TET2* biallelic mutation specifically sensitises AML cells to a particular chemotherapeutic agent and hence to develop an efficient therapeutic strategy.

3.3. Materials and Methods

The experimental methods that are specific to chapter 3 are outlined in this section.

3.3.1. CRISPR-CAS9 TET2 targeting in HEL cells

All the experiments outlined in section 3.3.1 were performed by Daniel Allsop (Newcastle University) except the flow cytometry, which was performed by Dr. Sarah Fordham (Newcastle University). The CRISPR lentivirus construct was designed by Dr. Thahira Rahman (Newcastle University) and supplied by Sigma-Aldrich.

A one vector system (pLV-U6-gRNA/EF1a-puro-2A-Cas9-2A-GFP packaged in lentiviral particles (Merck Life Science UK Limited, Dorset, UK)) with the guide RNA (gRNA) sequence 5'GTTTGGTGCGGGAGCGAGC-3' was used to target exon 6 of *TET2* (*TET2* transcript ID ENST00000540549.5) in HEL cells. The lentivirus vector design and the target location will be discussed in detail (Section 3.4.1). In general, the CRISPR-CAS9 gene editing protocol can be divided into the following stages.

Lentiviral transduction

Lentiviral particles were introduced into the HEL cells through the process of spinfection, where HEL cells were centrifuged at 800 g for 30 minutes at 32ºC in the presence of lentiviral particles. Hexadimethrine bromide (Polybrene) at 8 µg/ml was added to increase the transduction efficiency. The cells were initially transduced at a target multiplicity of infection (MOI) of 2, which appeared to be too low following puromycin selection and therefore the MOI was increased to 100 in subsequent transductions. The appropriate volume of lentiviral particles to give the required MOI was calculated using Equation 3.1. The mock-treated cells were supplemented with an equivalent volume of RF-10 media to the volume of lentiviral particles added to the CRISPR-treated cells.

$$Volume\ required\ (ml) = \frac{total\ no.of\ cells\ \times\ required\ MOI}{viral\ titre\ \ (TU/ml)}$$

Equation 3.1. Formula to calculate volume of lentivirus corresponding to the target MOI. The viral titre for each vial of lentiviral particles was provided by the manufacturer.

Selection of transduced population of cells

The lentivirus vector was designed with markers for puromycin resistance and GFP expression. The puromycin sensitivity of parental HEL cells was evaluated through a growth inhibition assay and the lowest cytotoxic dose which is the minimum puromycin dose required to kill all the cells was determined to be 1 µg/ml (personal communication, Daniel Allsop, Newcastle University). The lentivirus treated cells were cultured in the presence of 1 µg/ml puromycin for 1 week and the puromycin resistant cells that survived the cytotoxic dose were subsequently tested for GFP positivity using flow cytometry.

Isolation of clonal populations using semi-solid agar

Lentivirus treated cells were plated out in 6-well plates (Corning) to isolate clonal populations of cells. The plating density was 10 cells per well in 2 ml of 0.2% soft agar (2 mg/ml) media. The plates were incubated at 37°C and 5% CO_2 for 2-3 weeks until the colonies were visible. As the cloning efficiency of HEL cells appeared to be very low, the plating density was increased to 20 cells per well in the subsequent replicates. The individual colonies that appeared were picked and expanded in RF10 media for further flow cytometric analysis.

Flow cytometry to test GFP expression

Flow cytometry was performed by Dr. Sarah Fordham using BD FACSCalibur flow cytometer (BD Biosciences) to determine the GFP positivity of transduced population of cells. Gating was performed to reliably detect viable cells based on forward scatter (cell size) and side scatter (cell granularity). Pre-optimised settings for detection of GFP fluorescence were used and fluorescence of the gated population was measured at wavelength 530/30 nm using the fluorescence light height (FL1-H) channel. The resulting flow cytometry data was analysed using the Cyflogic tool (CyFlo limited, Finland).

3.3.2. TET2 mutation screening

TET2 Primer designing

Primer pairs spanning the CRISPR-target region of exon 6 in *TET2* were designed (Table 3.1) using Primer- Basic Local Alignment Search Tool (Primer-BLAST) (Ye et al. 2012) and was supplied by Merck (Merck Life Science UK Limited). The target specificity of the primers were tested *in silico* using the BLAST tool and the desired region of *TET2* was confirmed to be the only strong target of the primer pairs in the human genome. The expected product size was 534 base pairs (bp).

DNA extraction

DNA was extracted from cell pellets using QIAamp DNA Mini kit (Qiagen, Manchester, UK) according to manufacturer's protocol. The extracted DNA was quantified using a Nanodrop ND-1000 spectrophotometer and all samples were diluted to 50 ng/µl using nuclease free water to make up to a total volume of 20 µl.

PCR optimisation

Primers were originally dissolved in molecular grade nuclease free water to a concentration of 100 µM and then an aliquot was diluted to 10 µM to make up the final working concentration. DreamTaq PCR master mix (2X) (Thermo Scientific) was used for all PCR reactions. In addition, appropriate volumes of forward and reverse primers and nuclease free water was added into the reaction mix (Table 3.2) and 23 µl of the reaction mix was dispensed into 0.2 ml PCR tubes (Eppendorf, Stevenage, UK). For each sample, 2 µl of respective template DNA (100 ng) was added separately into the respective PCR reaction tube whereas an equal volume (2 µl) of nuclease free water was substituted in the negative control tube to make up to a total volume of 25µl reaction mix in all tubes. The PCR was run in a thermocycler as per the reaction conditions (Table 3.3).

Agarose Gel Electrophoresis

An aliquot of each post-PCR reaction was analysed using agarose gel electrophoresis to visualise and confirm the amplification of the target region in *TET2*. A 1% (w/v) agarose solution was made by melting 1g of Ultrapure™ agarose (Invitrogen Life Technologies) in 100 ml 1X TBE buffer [89 mM Tris-HCl pH 8, 89 mM boric acid, 2 mM EDTA] using a microwave. The solution was allowed to cool to 'hand-heat' and 10 µl of GelGreen Nucleic Acid Stain (10,000X in water) was added, thoroughly mixed, and the resultant mixture was poured into the large gel casting tray pre-assembled with the desired comb with required number of wells. The gel was allowed to set at room temperature, following which the comb was carefully removed and the cassette with the gel was placed in a Sub-Cell® GT Agarose Gel Electrophoresis System (Bio-Rad Laboratories Ltd., Watford, UK) pre-filled with 1X TBE buffer.

To precisely estimate the sizes of bands, 10 µl of Quick-Load® 100bp DNA Ladder (NEB #N0467S) was added into lanes 1 and 10. Following this, 5 µl of each post-PCR reaction was added in lanes 2 to 8 and 5 µl of PCR product of the negative control (where the template DNA was replaced with equivalent volume of nuclease free water) was added into lane 9.

Primer	Sequence (5' → 3')	Length	Melting temperature	GC%
Forward	GAATGGTGATCCACGCAGGT	20	67.3	55
Reverse	AGGAGCCGTCACTTGTAGCA	20	64.9	55

Table 3.1. PCR Primer sequences, length, melting temperature and GC%.
GC%, Guanine-Cytosine percentage

Component	Volume per reaction (µl)	Final Concentration
DreamTaq PCR master mix (2X)	12.5	1X
Forward Primer (10µM)	1.25	0.5µM
Reverse Primer (10µM)	1.25	0.5µM
Template DNA (50ng/µl) *	2.0	100ng
Nuclease free water	8.0	
Total volume	**25.0**	

Table 3.2. Preparation of PCR master mix
*Template DNA was added separately into the respective reaction tubes.

Step	Temperature (ºC)	Time	No: of cycles
Initial denaturation	95	2 min	1
Denaturation	95	25 sec	38
Annealing	60	40 sec	
Extension	72	65 sec	
Final extension	72	5 min	1
Hold	4	∞	-

Table 3.3. Thermal cycler conditions for PCR

Gel electrophoresis was performed at 100V for approximately 1.5 hour until the dye front has travelled the required distance to ensure good separation of the bands based on size. Following electrophoresis, the gel was placed in a ChemiDoc XRS+ system (Bio-Rad Laboratories Ltd.) and the bands were visualised and analysed using the Image Lab Software (Bio-Rad Laboratories Ltd.). Presence of bands at the desired size (between 500 and 600 bp) and absence of bands in the negative control confirmed successful PCR.

Sanger sequencing

The remaining PCR products (approximately 20 µl) were cleaned used QIAquick PCR purification kit (Qiagen) as per the manufacturer's protocol. Briefly, the protocol is based on a simple and fast bind-wash-elute procedure where the PCR products are initially mixed with a binding buffer in a 1:5 ratio. The reaction mixture is then applied to a spin-column containing silica-membrane which allows binding of DNA into the membrane in high-salt conditions provided by the buffer and the residual impurities are washed away. Finally, the purified PCR products are eluted using 20 µl of low-salt elution buffer [10 mM Tris-HCl pH 8.5], following which the concentration of purified DNA is quantified using a Nanodrop ND-1000 spectrophotometer. The purified PCR products were diluted to 5 ng/µl and 15 µl of the resulting sample was sent to Eurofins Scientific (Eurofins Genomics Germany GmbH) for Sanger sequencing with forward primer only (Table 3.1).

The sequencing data obtained in FASTA format, were initially compared with the published human *TET2* sequences using nucleotide BLAST (NCBI). The sequences were visualised and analysed using Finch TV (Geospiza, Inc) to check the quality of sequence reads. Multiple Sequence Comparison by Log- Expectation (MUSCLE) tool (EMBL-EBI) was used to perform multiple sequence alignment and to confirm that the sequences of *TET2* monoallelic mutated control clones (HELmono1, HELmono2 and HELmono3) completely matched with that of HEL parental cells. Further, the target region sequences of CRISPR-treated *TET2* biallelic mutated clones (HELbi1, HELbi2 and HELbi3) were manually compared with the corresponding sequence of HEL parental cells to identify the exact nature of mutation (insertion/deletion), location of mutation, change in number of base pairs and the mutated sequence. In case of HELbi1 and HELbi2, the inserted sequence was predicted by carefully comparing and correlating the mixed sequence signal with the corresponding region in HEL parental cells.

3.4. Results

3.4.1. Generation of CRISPR targeted TET2 biallelic mutant HEL cells

In order to characterise the effect of *TET2* biallelic mutation on AML cells and to develop a therapeutic strategy for *TET2* biallelic mutated AML, isogenic cell lines with biallelic *TET2* mutation were generated using CRISPR-Cas9 gene editing. HEL cells, originally established from a 30-year-old male with erythroleukaemia (AML-M6) were used as the model for this study. HEL AML cells are hypotriploid and reported to carry a large monoallelic deletion and concomitant loss of heterozygosity covering the majority of chromosome 4q, which includes the *TET2* gene (Martin et al. 1982; Abdel-Wahab et al. 2009).

(i) A brief account of CRISPR lentiviral transduction

A CRISPR-Cas9 one vector system (pLV-U6-gRNA/EF1a-puro-2A-Cas9-2A-GFP in lentiviral particles) was designed incorporating a gRNA sequence targeting exon 6 of *TET2* (Figure 3.1; this part of the study was performed by Daniel Allsop and Dr. Sarah Fordham, Newcastle University). The lentiviral particles were incubated with HEL cells at MOI of 2 and thereafter at MOI of 100 using spinfection protocol (Section 3.3.1). The presence of puromycin resistance and GFP markers added two layers to the selection process to efficiently isolate successfully transduced cells. Transduced cells were selected by culture in the presence of puromycin, subsequently cloned using plating in soft agar and finally tested for GFP positivity using flow cytometry. With multiple rounds of selection and subsequent single-cell cloning, three independently transduced clones HELbi1, HELbi2 and HELbi3 were taken forward for *TET2* mutation screening using sanger sequencing. Likewise, HELmono1, HELmono2 and HELmono3 are independently derived sub-clones of mock-treated HEL parental cells, which are *TET2* monoallelic mutated developed to act as control clones throughout this study.

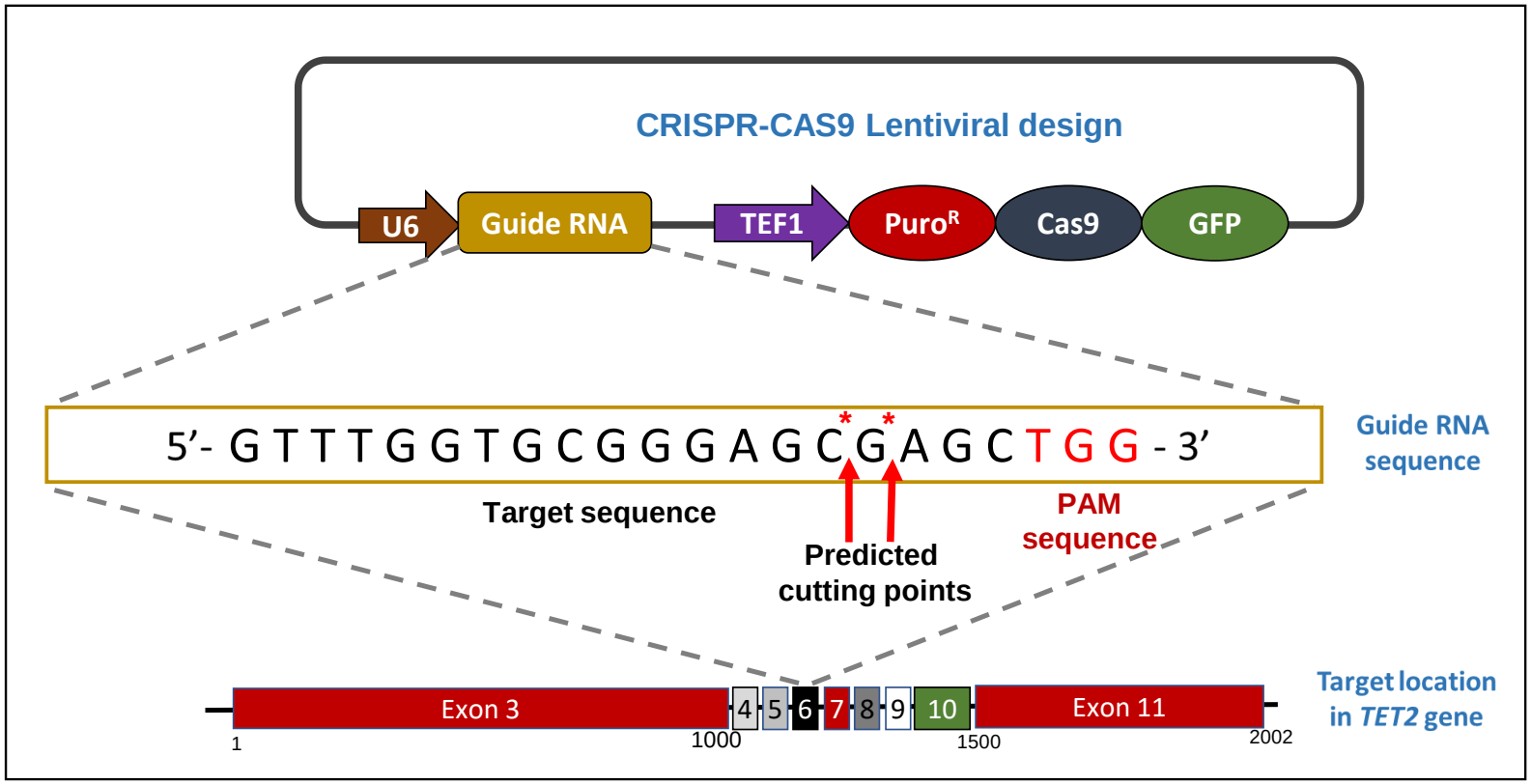

Figure 3.1. Schematic representation of the CRISPR-Cas9 lentiviral vector, showing the guide RNA sequence and *TET2* target (exon) location. The lentiviral design consists of a guide RNA region, and sequence encoding a puromycin resistance cassette, green fluorescent protein (GFP), the Cas9 endonuclease and their respective promoters. The U6 promoter is a class III RNA polymerase III promoter which has been widely used to induce expression of small RNAs and the truncated human elongation factor-1 (TEF1) promotor is commonly used to drive the expression of CAS9. The guide RNA sequence consists of a target sequence which is specific to the target location in *TET2* exon 6 and incorporates a PAM sequence. The double strand breaks induced by CRISPR-Cas9 are predicted to be three to four base pairs upstream of the PAM. The *TET2* exon distribution (bottom) shows the target location of CRISPR-Cas9 induced break in the *TET2* gene and the numbers below the exon distribution denotes the approximate amino acid locations.

CAS9, CRISPR-associated endonuclease 9; CRISPR, clustered regularly interspaced short palindromic repeats; GFP, green fluorescence protein; PAM, protospacer adjacent motif; Puro^R, puromycin resistance marker; TEF1, truncated human elongation factor-1; U6, class III RNA polymerase III promoter.

(ii) Validation of TET2 mutation using Sanger sequencing

The CRISPR targeting region in *TET2* exon 6 was PCR amplified using DNA extracted from HEL parental and HEL monoallelic clones. An amplicon between 500-600 bp was generated, consistent with the expected product size of 534 bp generated from the intact wild-type TET2 allele (Figure 3.2; lanes 2-5). PCR amplification of the exon 6 sequence in CRISPR-Cas9-targeted HEL cells also generated products between 500-600 bp, but with some evidence of larger than expected products potentially consistent with gene editing (Figure 3.2; lanes 6-10). There was no evidence of non-specific amplification, suggesting that the primer design and PCR conditions were appropriately optimised, generating an amplicon that was suitable for sanger sequencing.

Sanger sequencing and MUSCLE sequencing alignment of HEL parental cells and *TET2* monoallelic control clones confirmed the presence of an intact CRISPR targeting region on *TET2* exon 6 (Figure 3.3A and 3.3B; Full sequence data are provided in Appendix A). Sanger sequencing data of HEL CRISPR-targeted clones (HELbi1 and HELbi2) suggested the presence of a 4-base pairs deletion at nucleotide position 40 and a putative 30-base pairs insertion at nucleotide position 43 in the same clones, suggesting the presence of two CRISPR-targeted alleles in these clones (Figure 3.4; Full sequence data are provided in Appendix A). The exact sequence was difficult to decipher due to the presence of mixed signal from position 43 but suggests that parental HEL cells have two intact wild-type *TET2* alleles, and not just one as expected. As such, CRISPR-targeting appears to have generated two different mutant alleles (compound heterozygosity) in HELbi1 and HELbi2 clones. The best estimate of the inserted sequence is ACTGTGAAGAGAACTACTGCTGTTTGGTGC (5'-3'). Both HELbi1 and HELbi2 were cloned out from the same parental CRISPR treated population (which also showed mixed sequence signals). The impact of this mutation at protein level is too complex to allow prediction at this stage. However, a change in 26-base pairs is likely to generate a frame shift which is expected to have a functional impact.

In contrast to HELbi1 and HELbi2, Sanger sequencing of HELbi3 generated high quality clear sequence signal that confirmed the presence of a 4 bp deletion at position 40 (Figure 3.4). As such, the location of the deletion is exactly at the predicted CRISPR-Cas9 cutting point, which is 3-4 base pairs 5' to the PAM sequence and this mutation is expected to cause a frameshift nonsense mutation. The PolyPhen-2 protein prediction tool, which is designed to predict the functional impact of a mutation at the protein level, confirmed the generation of a premature stop codon resulting in a non-functional allele as a result of the 4 bp deletion.

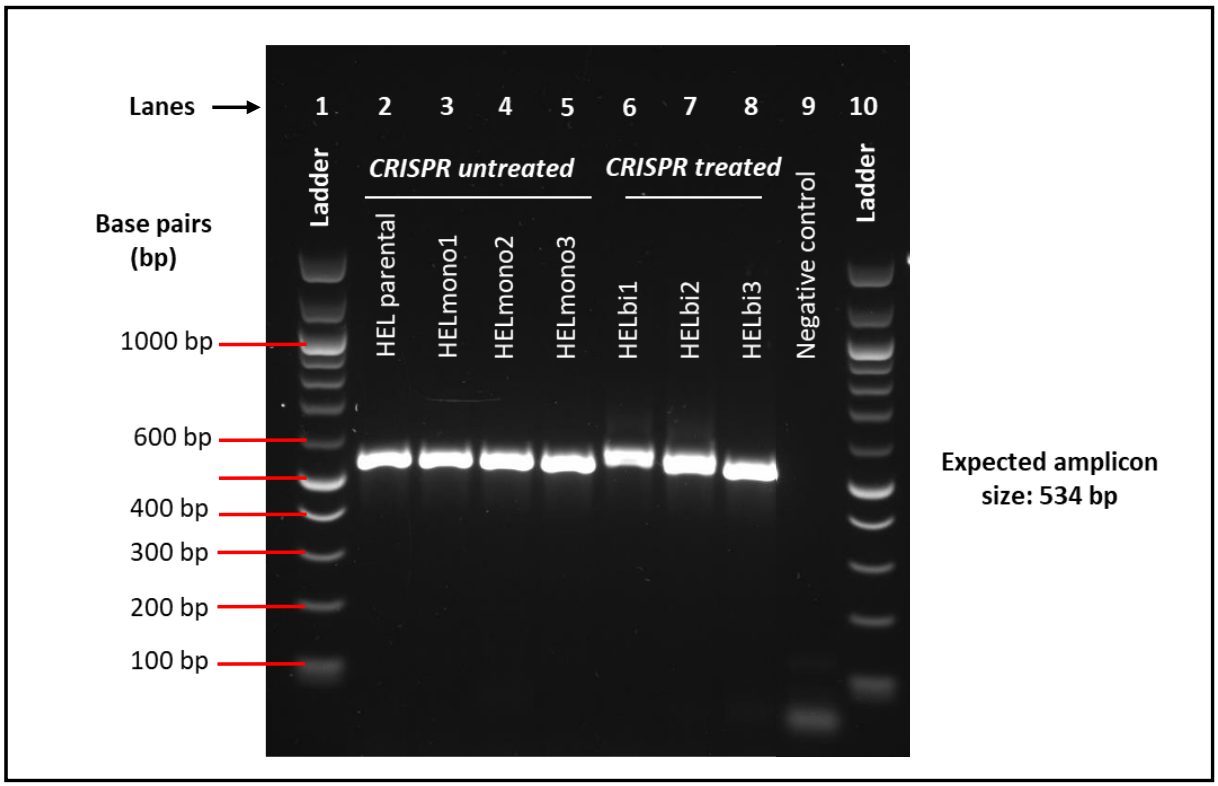

Figure 3.2. PCR optimisation gel for Sanger sequencing. The CRISPR target region within the *TET2* locus was PCR amplified on HEL parental cells and CRISPR untreated and treated clones, which were selected for mutation screening. The band observed between 500 bp and 600 bp is consistent with the expected amplicon size of 534 bp. Lane 1 and lane 10 shows the DNA ladder and lane 9 shows the negative control (template DNA was replaced with nuclease free water in the PCR reaction mix).

bp, base pairs; CRISPR, clustered regularly interspaced short palindromic repeats; PCR, polymerase chain reaction.

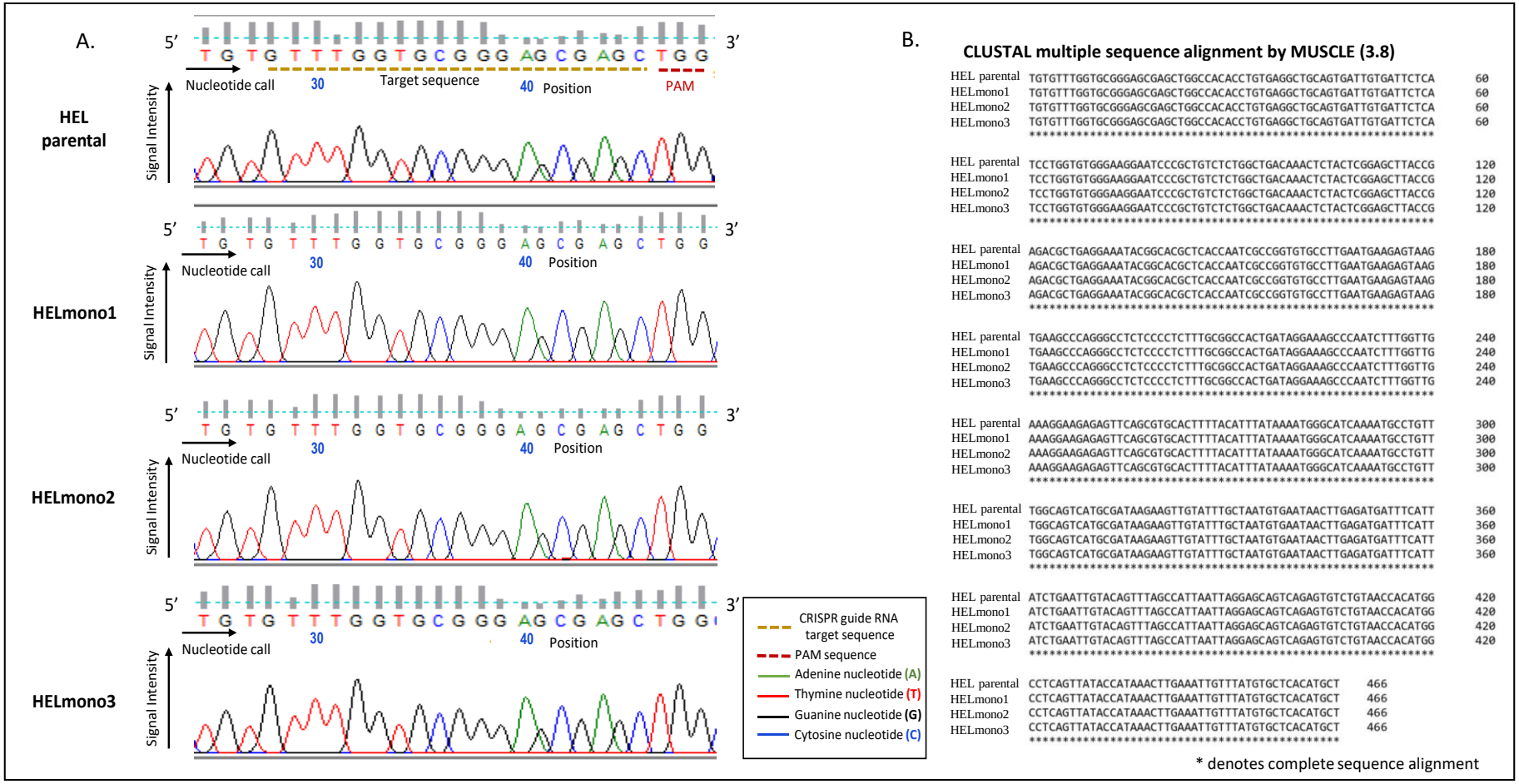

Figure 3.3. Sanger sequencing of *TET2* monoallelic mutated cell clones. A. Forward sequencing (5'-3') electropherogram of HEL parental cells and its subclones with *TET2* monoallelic mutation. The respective sequence signal intensity, nucleotide call, base position numbers (blue), CRISPR-target region (gold dotted lines), PAM sequence (red dotted lines) are shown. B. Multiple sequence alignment shows complete alignment of *TET2* sequences in HEL parental cells and its clones.

CRISPR, clustered regularly interspaced short palindromic repeats; PAM, Protospacer adjacent motif

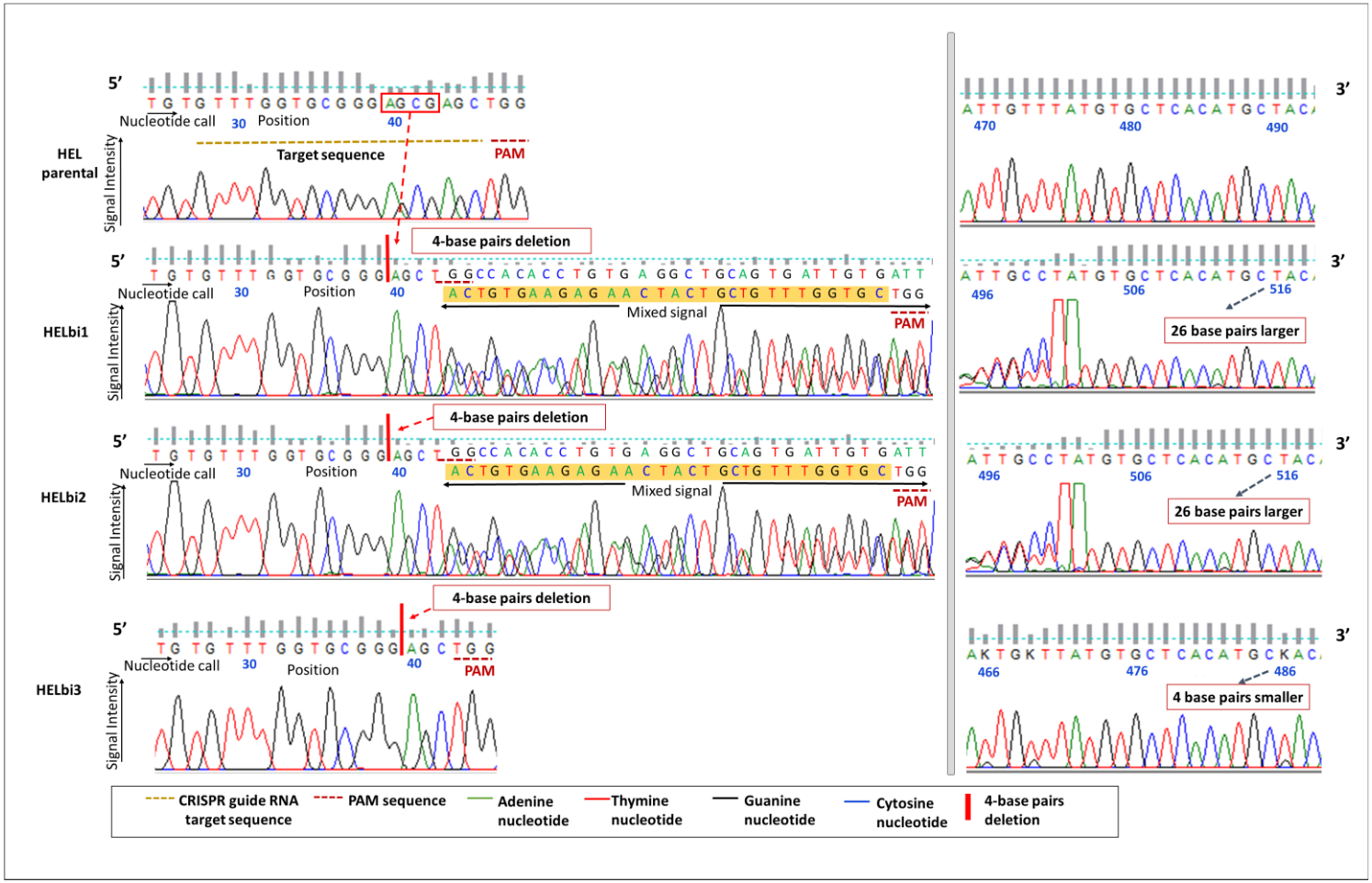

Figure 3.4. Sanger sequencing of *TET2* biallelic mutated cells. Forward sequencing (5'-3') electropherogram of the three selected CRISPR-treated clones in comparison with the HEL parental cells show the presence of a 4 bp deletion and a 30 bp insertion in HELbi1 and HELbi2 and a 4 bp deletion alone in HELbi3. The best estimate of the inserted sequence is highlighted in yellow. The right side of the figure shows the 3' end of the sequences. By comparison of the base position numbers (blue) corresponding to aligning regions, it can be inferred that sequence derived from HELbi1 and HELbi2 was 26 bp longer (base position number 516) whereas HEL bi3 was 4 bp shorter (base position number 486) when compared to the corresponding region of HEL parental (base position number 490). The respective sequence signal intensity, nucleotide call, base position numbers (blue), CRISPR-target region (gold dotted lines), PAM sequence (red dotted lines) are shown.
CRISPR, clustered regularly interspaced short palindromic repeats; PAM, protospacer adjacent motif

(iii) Characterisation of TET2 mutation in HEL parental cells

Considering the mixed sequence signal observed in two of the biallelic *TET2* mutated clones, HELbi1 and HELbi2, it was decided to further validate the exact *TET2* mutation status of HEL cells. The SNP array data on HEL cells (HEL, DSMZ accession number GSM888147) was downloaded from the Cancer Cell Line Encyclopaedia (CCLE) (data accessible at NCBI GEO database, (Edgar et al. 2002) and was analysed using Nexus software.

SNP array data comprises two major components, namely B-allele frequency (BAF) and Log R ratio. The BAF is a measure of allele intensity ratio of the two alleles A and B at polymorphic sites in the genome, such that a value of 1 or 0 shows a complete absence of one of the two alleles (homozygous for AA or BB or due to loss of one allele causing allelic imbalance), and a value of 0.5 shows equal presence of both the alleles (heterozygous AB), and an intermediate BAF value between 0 and 0.5 or 0.5 and 1 results from an allelic imbalance (AAB or ABB) (Wang et al. 2014). The log R ratio is a normalised measure of total signal intensities for the two alleles (Wang et al. 2014). In short, evidence of loss of heterozygosity (LOH) as indicated by a BAF close to 0 or 1 accompanied by a downward negative shift in log R ratio indicates a loss event or reduction in copy number. Conversely, evidence of allelic imbalance accompanied by an increase in log R ratio indicates a gain or increase in copy number (Wang et al. 2014).

Based on SNP data analysis it can be concluded that HEL cells have a highly complex hypotriploid karyotype and as three copies of most of the chromosomes (Figure 3.5A). Inspection of the BAF on chromosome 4 indicates a large deletion in the long arm (4q) which includes the *TET2* locus (denoted by a green line; Figure 3.5B), consistent with published reports (Martin et al. 1982; Abdel-Wahab et al. 2009). However, the entire chromosome 4 except the deleted region shows a triploid karyotype, as indicated by the BAF data demonstrating allelic imbalance. The corresponding increase in log R ratio is not very discernible as the normal baseline itself (which generally shows diploidy) has slightly shifted upwards because of the global triploid pattern of the probes in HEL cells. HEL cells are derived from a male with one copy of X chromosome and hence the log R ratio of this chromosome could be considered as a plausible baseline (which corresponds to a copy number of 1) (red dotted lines, Figure 3.5A), assuming there is no duplication of this chromosome. Using the X chromosome as a baseline a representing copy number of 1, it is clear that the copy number of the deleted region of chromosome 4 is diploid, (approx. baseline - red dotted lines, Figure 3.5A) suggesting the existence of two intact copies and that these are homozygous, as evidenced

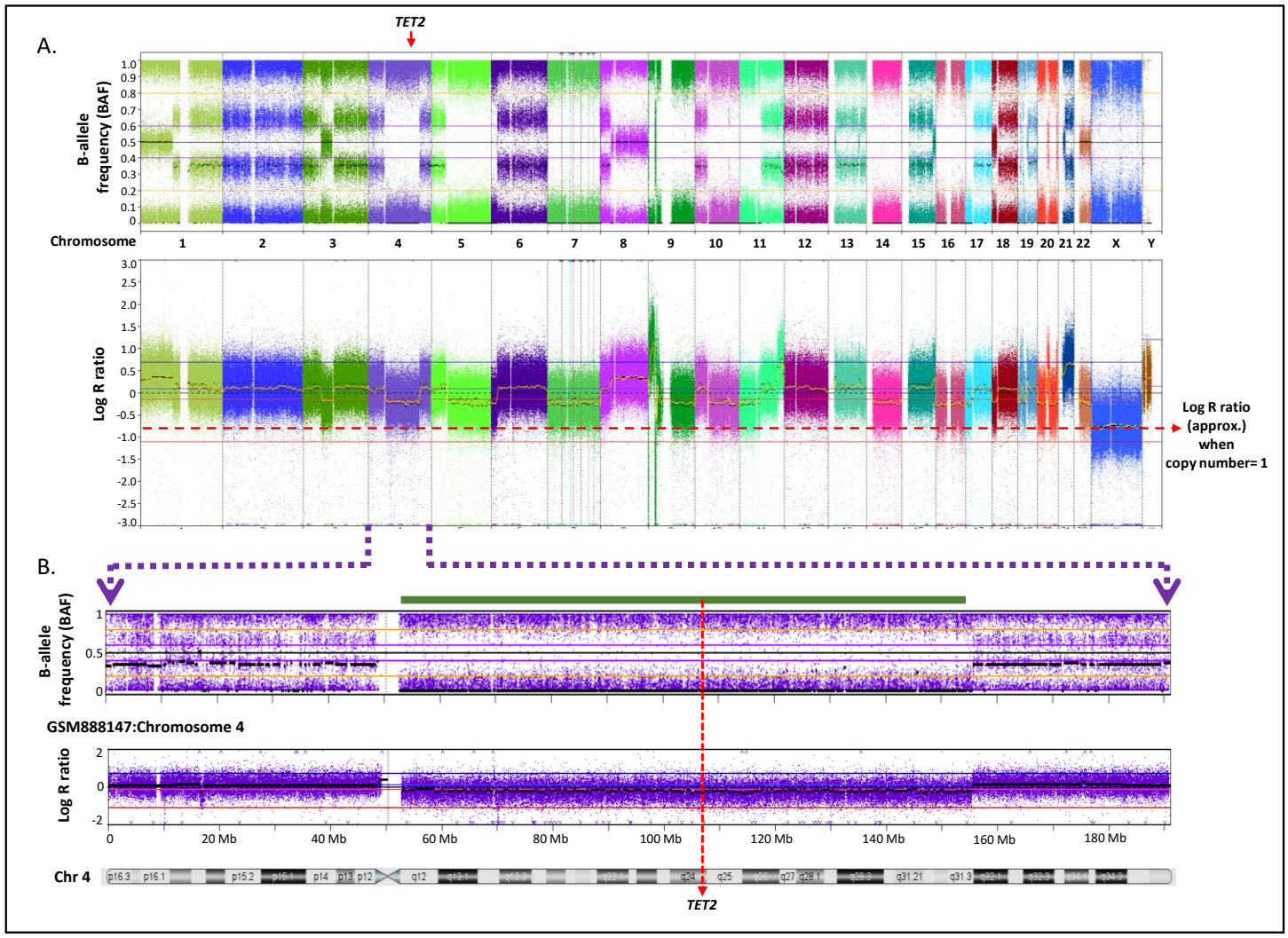

Figure 3.5. High density SNP array data on HEL cells.

A. Copy number profile of HEL cells showing a hypotriploid karyotype. B. Copy number profile of chromosome 4 from HEL cells with a large deletion (green line) affecting most of the q arm including *TET2* locus (4q24) (red dotted line). Consistently, the deleted region shows LOH as indicated by the respective BAF and a reduction in Log R ratio. Furthermore, the rest of the Chr 4 flanking the deleted region shows a gain in copy number (triploid). The data accessed from GEO database (GSM888147) was analysed using Nexus software.

BAF, B-allele frequency; Chr, Chromosome; Mb, mega base

by the BAF. As such, it can be concluded that HEL parental cells retain two intact copies of chromosome 4 (likely the result of a somatic duplication) and a third copy of chromosome 4 characterised by a large deletion encompassing the *TET2* locus. These data are consistent with the Sanger sequencing data from CRISPR-targeted clones indicating independent targeting events on two separate *TET2* loci (section 3.3.1 (ii)).

Regarding the order of somatic events, it's possible to speculate that the duplication resulting in triploidy happened first as a single mitotic cycle, as this is seen in the entire set of chromosomes, followed by localised deletion in the long arm of on the non-gained chromosome. Nevertheless, HEL cells are derived from a leukaemia that evolved in the background of a monoallelic *TET2* deletion (*TET2* -/+/+) and hence HEL parental cells and the sub-clones derived from these cells will be coined as *TET2* monoallelic mutated cells hereafter. Also, this suggests a potential explanation for the mixed signals seen in HELbi1 and HELbi2 that might have resulted from a 4 bp deletion in one *TET2* allele and 4 bp deletion along with a 30 bp insertion in the other allele.

3.4.2. TET2 protein level expression in TET2 biallelic mutated cells

To elucidate the functional impact of *TET2* mutant allele dosage, it was necessary to determine TET2 protein levels in CRISPR-targeted cells. Protein extracts from HEL cells with monoallelic *TET2* mutation (HEL parental cells and sub-clones HELmono1, HELmono2 and HELmono3) and CRISPR treated HEL cells with biallelic *TET2* mutation (HELbi1, HELbi2 and HELbi3) were subjected to western analysis with anti-TET2 antibody. Despite the differences in mutational profiles, all CRISPR targeted *TET2* biallelic mutated cell clones (HELbi1, HELbi2 and HELbi3) were consistently null for TET2 protein expression (Figure 3.6A). Furthermore, a band consistent with the predicted size of TET2 (224kDa) was observed in HEL cells with monoallelic *TET2* mutation (Figure 3.6A).

TET2 protein levels were quantified using Imagelab software and the resultant band intensity values were normalised based on the respective GAPDH band intensities and show loss of TET2 protein expression in cell clones with *TET2* biallelic mutation (Figure 3.6B). Protein quantification clearly mirrored the TET2 levels in the blot.

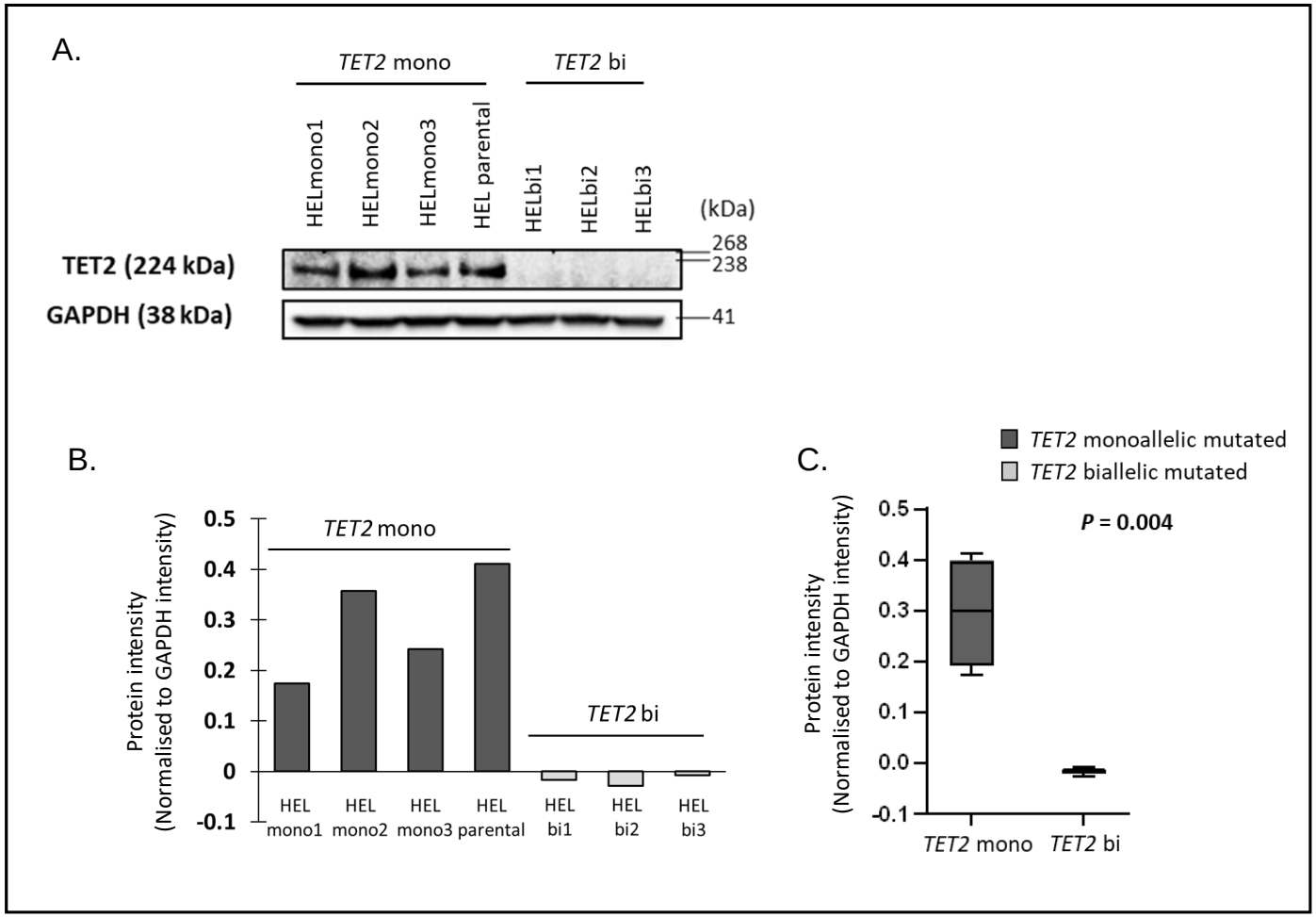

Figure 3.6. Western blot analysis of *TET2* monoallelic and biallelic mutated HEL cells. A. HEL cells with *TET2* monoallelic mutation (HELmono1, HELmono2, HELmono3 and HEL parental) expressed TET2 protein whereas *TET2* biallelic mutation (HELbi1, HELbi2 and HELbi3) leads to complete loss of TET2 protein expression. TET2 bands were developed in 22 minutes exposure using Chemidoc imaging system. GAPDH was used as a loading control and the blot was developed in 41 seconds. The negative trend observed in *TET2* biallelic mutated cells are due to the lower band intensities of the respective bands compared to that of the local background. This trend is usually seen when the expression of a particular protein is completely null in the cell sample. B. TET2 protein levels were quantified and normalised to respective GAPDH protein intensities to account for the protein sample loading variations. C. Mean TET2 protein levels of cells grouped based on *TET2* mutation status. In the box and whiskers plot, the bars in the plot represent the mean, the boxes represent the interquartile range, and the whiskers represent the range between minimum and maximum values.

TET2 bi, *TET2* biallelic mutated; *TET2* mono, *TET2* monoallelic mutated

3.4.3. Effect of TET2 biallelic mutation on growth kinetics of cells

A proliferation assay was conducted to evaluate the effect of *TET2* mutant allele dosage on growth kinetics of cells in liquid culture. The results are based on three independent replicates of the experiment. Cell proliferation experiments indicated that regardless of the *TET2* mutant allele dosage, there was no difference in proliferation rate of monoallelic and biallelic *TET2* mutated cell clones in liquid media (P = 0.99, ANOVA) (Figure 3.7 A-B). Specifically, doubling time of monoallelic clones was estimated to be 32.04 hours, compared to 32.88 hours for biallelic clones.

Likewise, there was no significant difference in mean cloning efficiency of *TET2* monoallelic and biallelic cell clones (Figure 3.7 C-D) (P = 0.18, student's *t*-test). Therefore, complete loss of TET2 in cells with biallelic *TET2* mutation did not affect either growth kinetics or clonogenic potential, in comparison to cells with monoallelic *TET2* mutation.

3.4.4. Impact of mutant TET2 allele dosage on response to AML therapeutics

(i) Response to Azacytidine

Given the response of index patient with *TET2* biallelic mutated AML to single agent azacytidine (Section 1.6), *TET2* monoallelic and biallelic mutated HEL cells were treated with increasing doses of azacytidine. *TET2* biallelic mutated cell clones had significantly lower proliferation in liquid culture in the presence of azacytidine compared to their isogenic counterparts with *TET2* monoallelic mutation (P <0.0001, ANOVA) (Figure 3.8A-B). Consistently, both IC50 and IC95 values were significantly lower for *TET2* biallelic mutated cell clones compared to *TET2* monoallelic cell clones (P = 0.007 and <0.0001 respectively, student's *t*-test) (Figure 3.8 C-D). In addition, *TET2* biallelic mutated cells had significantly lower cloning efficiency compared to monoallelic *TET2* mutant HEL cells in the presence of 5μM azacytidine (P =0.006, student's *t*-test) (Figure 3.8E).

In summary, *TET2* biallelic mutation renders cells hypersensitive to the growth inhibitory effects of azacytidine compared to cells with monoallelic *TET2* mutation.

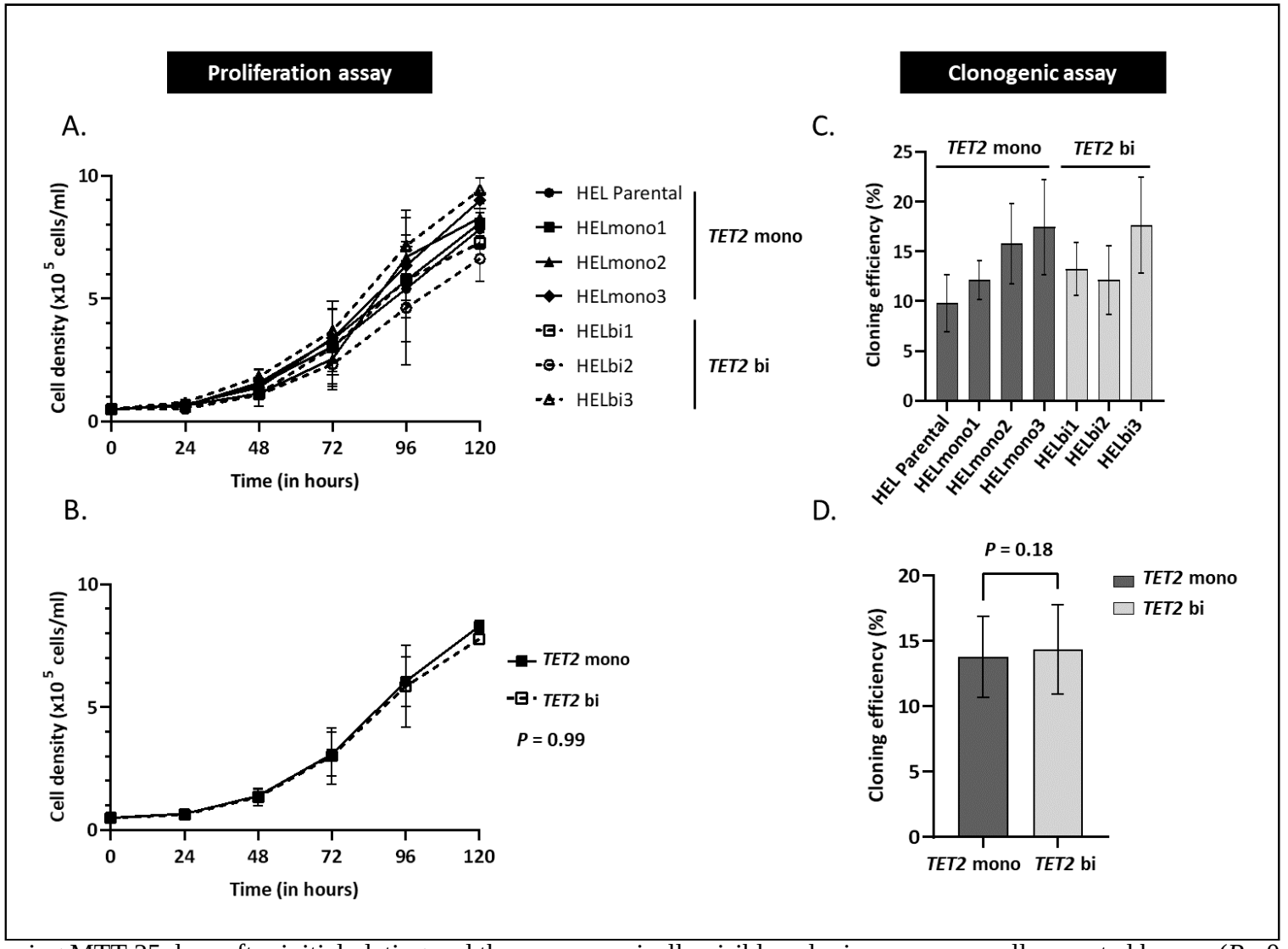

Figure 3.7. Growth kinetics and clonogenicity of *TET2* monoallelic and biallelic mutated HEL cells.
A. Viable HEL cells with *TET2* monoallelic mutation (n=4) and *TET2* biallelic mutation (n=3) were seeded at $5x10^4$ cells/ml and counted at 24-hour intervals. Proliferation rate of individual cell clones are shown. B. Mean proliferation rate over individual cell types based on their *TET2* status. *TET2* mutation status did not have a significant impact on the proliferation rate of cells (P =0.99, ANOVA) C. Colony formation efficiency of individual cell clones of HEL *TET2* mutated cells D. Mean colony forming efficiency over individual cell types based on their *TET2* status. The colonies were stained using MTT 25 days after initial plating and the macroscopically visible colonies were manually counted by eye (P =0.18, student's *t*-test).

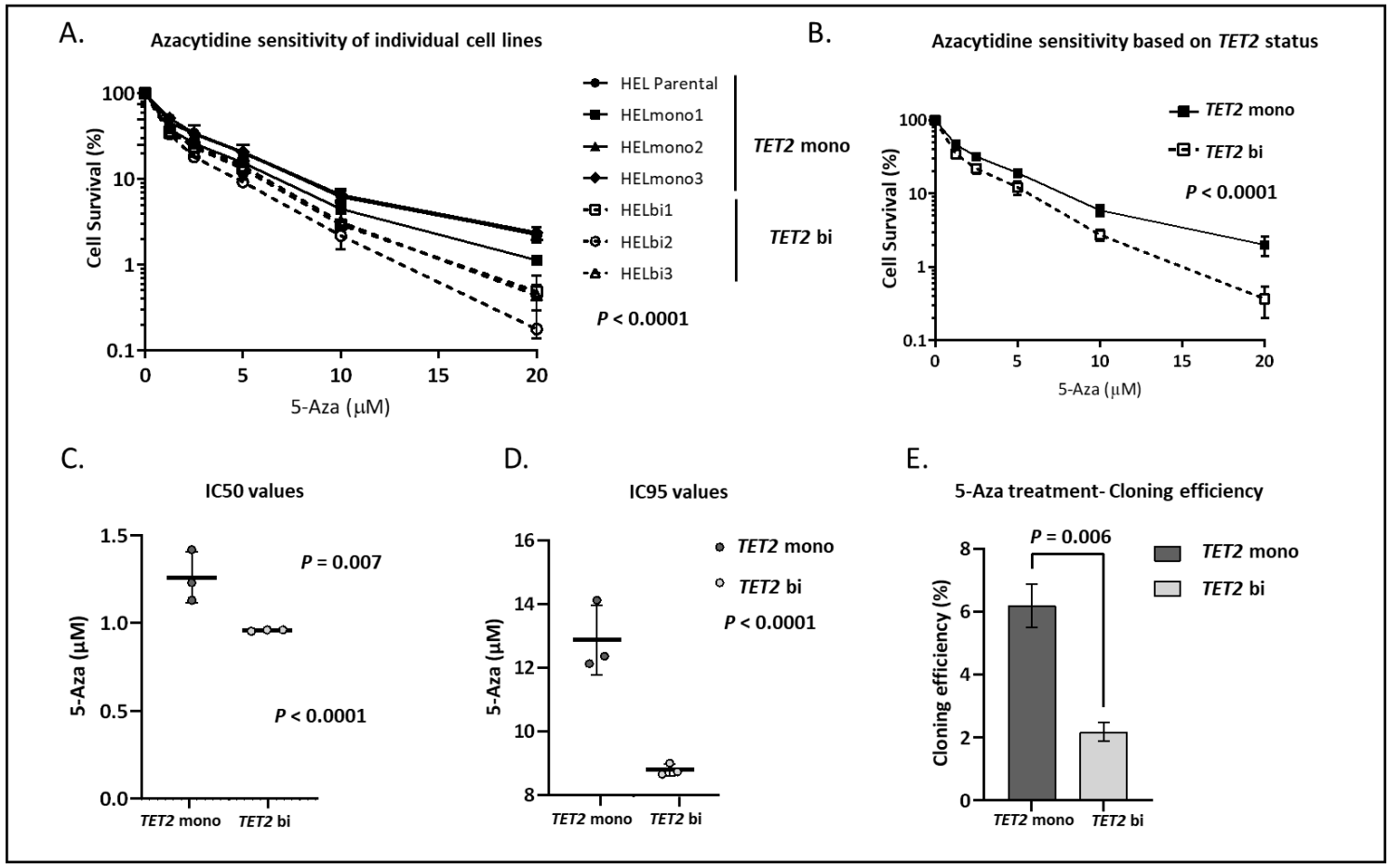

Figure 3.8. Azacytidine sensitivity of *TET2* monoallelic and biallelic HEL cells.

(A) Growth inhibition profile of individual cell types in response to increasing doses (0-20 μM) of azacytidine at 96 hours post treatment. Open symbols show the biallelic *TET2* mutants (HELbi1, HELbi2 and HELbi3) and closed symbols show the monoallelic *TET2* mutants (HEL parental, HELmono1, HELmono2 and HELmono3). *P* value (ANOVA) was calculated based on three independent replicates. (B) represents the mean azacytidine sensitivity over individual cell types based on their *TET2* status. *TET2* biallelic mutant HEL cells were hypersensitive to azacytidine compared to *TET2* monoallelic mutant HEL cells ($P < 0.0001$, ANOVA). Also, *TET2* biallelic mutant HEL cells had significantly lower IC50 (C) and IC95 (D) values with $P = 0.007$ and <0.0001, respectively (student's *t*-test). (E) *TET2* biallelic mutant HEL cells had significantly lower cloning efficiency in the presence of azacytidine compared to *TET2* monoallelic mutant HEL cells ($P = 0.006$, student's *t*-test). The colonies were stained using MTT 25 days after initial plating and the macroscopically visible colonies were manually counted by eye. All the error bars are based on standard deviation of mean of three independent replicates.

5-Aza, 5-Azacytidine

(ii) Response to Decitabine

TET2 monoallelic and biallelic mutated cells were treated with increasing doses of decitabine (DAC) to determine the sensitivity of *TET2* monoallelic and biallelic mutated cells. *TET2* biallelic mutated cell clones had significantly lower proliferation in liquid culture in the presence of decitabine compared to their isogenic counterparts with *TET2* monoallelic mutation ($P<0.0001$, ANOVA) (Figure 3.9A). However, the effect size was relatively weak compared to the differential sensitivity of these cells observed in response to azacytidine ($P=0.007$, ANOVA) (Figure 3.9B).

(iii) Response to Induction therapy drugs

Remission induction therapy which aims at complete remission is a *de facto* primary treatment approach in AML (Dombret et al. 2016). Therefore, it is important to determine the impact of *TET2* mutant allele dosage on cellular response to currently used remission induction therapies. As such, *TET2* monoallelic and biallelic mutated HEL cells were treated with increasing doses of various remission induction drugs selected based on AML15 clinical trial (NCT00006122).

AML15 compared standard cytarabine-daunorubicin-etoposide (ADE) remission induction treatment with cytarabine-daunorubicin (DA) treatment and fludarabine-cytarabine-G-CSF-idarubicin (FLAG-Ida). It also evaluated the effect of combining gemtuzumab ozogamicin (GO) along with these induction treatments (Burnett et al. 2013). Therefore, *TET2* monoallelic and biallelic mutated HEL cells were treated with increasing doses of individual agents such as cytarabine, daunorubicin, fludarabine, idarubicin and calicheamicin (the chemotherapeutic component of GO).

As such, both *TET2* monoallelic and biallelic mutated cells had similar proliferation in liquid culture in response to each remission induction therapeutics tested. *TET2* mutant allele dosage did not have a significant impact on response to cytarabine ($P=0.97$), daunorubicin ($P=0.49$), fludarabine ($P=0.99$), idarubicin ($P=0.24$) and calicheamicin ($P=0.32$) (ANOVA) (Figure 3.10 A-E). Therefore, based on our results *TET2* biallelic mutation does not particularly sensitise cells to the common remission induction regimen agents.

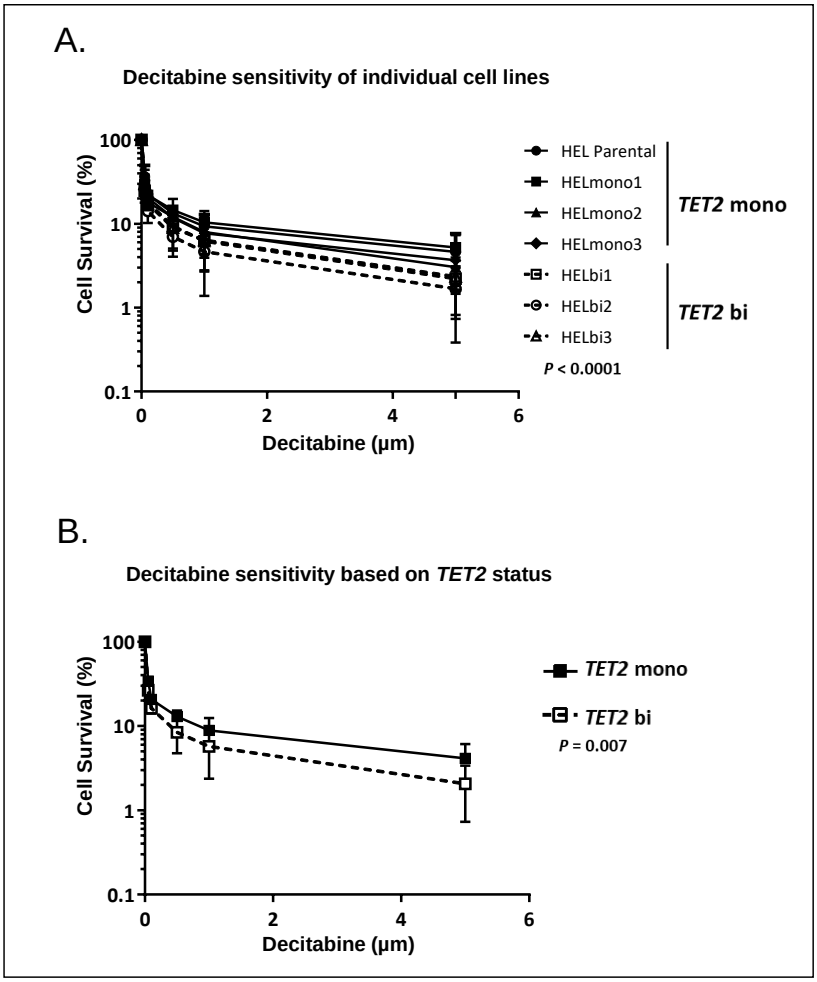

Figure 3.9. Decitabine sensitivity of *TET2* monoallelic and biallelic mutated HEL cells. A. Growth inhibition profile of individual cell types in response to increasing doses (0-5 µM) of decitabine at 96 hours post treatment (P <0.0001, ANOVA). B. represents the mean decitabine sensitivity over individual cell types based on their *TET2* status. Open symbols show the biallelic *TET2* mutants (HELbi1, HELbi2 and HELbi3) and closed symbols show the monoallelic *TET2* mutants (HEL parental, HELmono1, HELmono2 and HELmono3). *TET2* biallelic mutants were significantly more sensitive to decitabine compared to *TET2* monoallelic mutants (P =0.007, ANOVA). All the error bars are based on standard deviation of mean of three independent replicates.

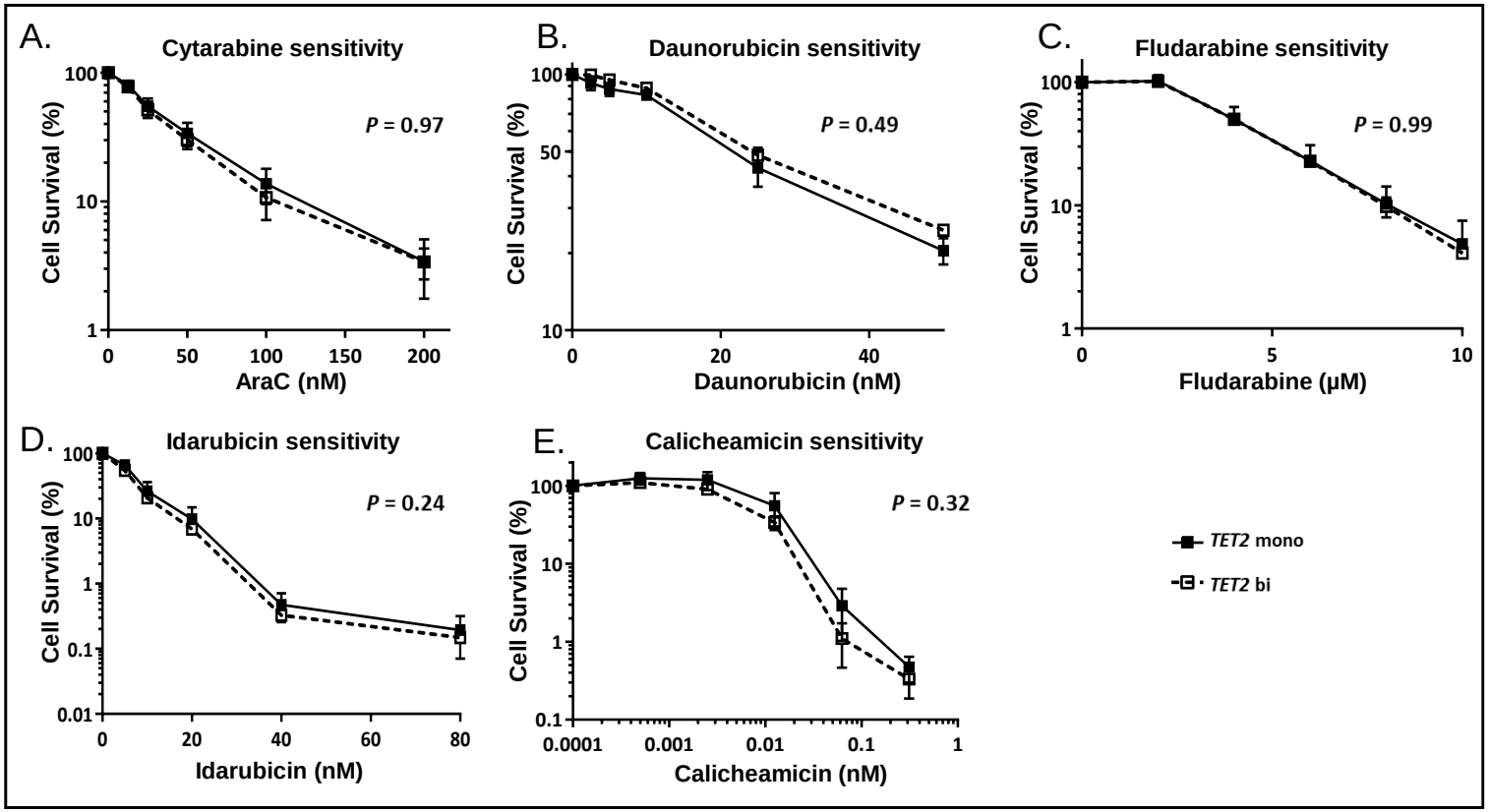

Figure 3.10. Response to various other common AML therapeutics. The figure represents the mean growth inhibition profile over individual cell types based on their *TET2* status in response to increasing doses of (A) cytarabine (0-200 nM), (B) Daunorubicin (0-50 nM), (C) Fludarabine (0-10 μM), (D) Idarubicin (0-80 nM) and (E) Calicheamicin (0-0.612 nM), which is the cytotoxic drug component of GO (the antibody-drug conjugate). The drug responses were not significantly different based on *TET2* status. Open symbols show the biallelic *TET2* mutants (HELbi1, HELbi2 and HELbi3) and closed symbols show the monoallelic *TET2* mutants (HEL parental, HELmono1, HELmono2 and HELmono3). The error bars are based on standard deviation of mean of three independent replicates.

AraC, Cytarabine; GO, Gemtuzumab ozogamicin

(iv) Response to IDH1/IDH2 inhibitors

Inhibition of TET2 induces synthetic lethality on *IDH1/IDH2* mutated cells and this has been attributed to the phenotypic redundancy of IDH1/IDH2 with that of TET2 (Pagliuca et al. 2020). In this context, it was interesting to validate if a similar synthetic lethality can be caused by pharmacologically inhibiting IDH1/IDH2 in a TET2 null background. Therefore, *TET2* biallelic mutated HEL cells were treated with increasing doses of ivosidenib (IDH1 inhibitor) and enasidenib (IDH2 inhibitor) and dose responses were evaluated. However, *TET2* biallelic mutated cells did not have significant difference in proliferation in liquid culture in the presence of enasidenib (P =0.6) or ivosidenib (P =0.99) when compared to their isogenic *TET2* monoallelic mutated counterparts (Figure 3.11). As such, IDH1/IDH2 inhibition did not induce a synthetic lethality in TET2 null cells.

3.4.5. Summary of findings

In summary, an isogenic cell model system was established based on HEL cells using CRISPR-Cas9 gene editing. The control clones were *TET2* monoallelic mutated whereas the CRISPR-treated clones were *TET2* biallelic mutated. Based on our model, *TET2* biallelic mutated cells had complete loss of TET2 protein whereas *TET2* monoallelic mutated cells partially expressed TET2 protein. *TET2* mutant allele dosage had significant impact on cellular response to azacytidine and decitabine, though the effect size of decitabine response was relatively weaker. *TET2* biallelic mutation did not have a significant effect on baseline proliferation rate, clonogenicity or cellular response to other chemotherapeutics tested, compared to *TET2* monoallelic mutants (Table 3.4)

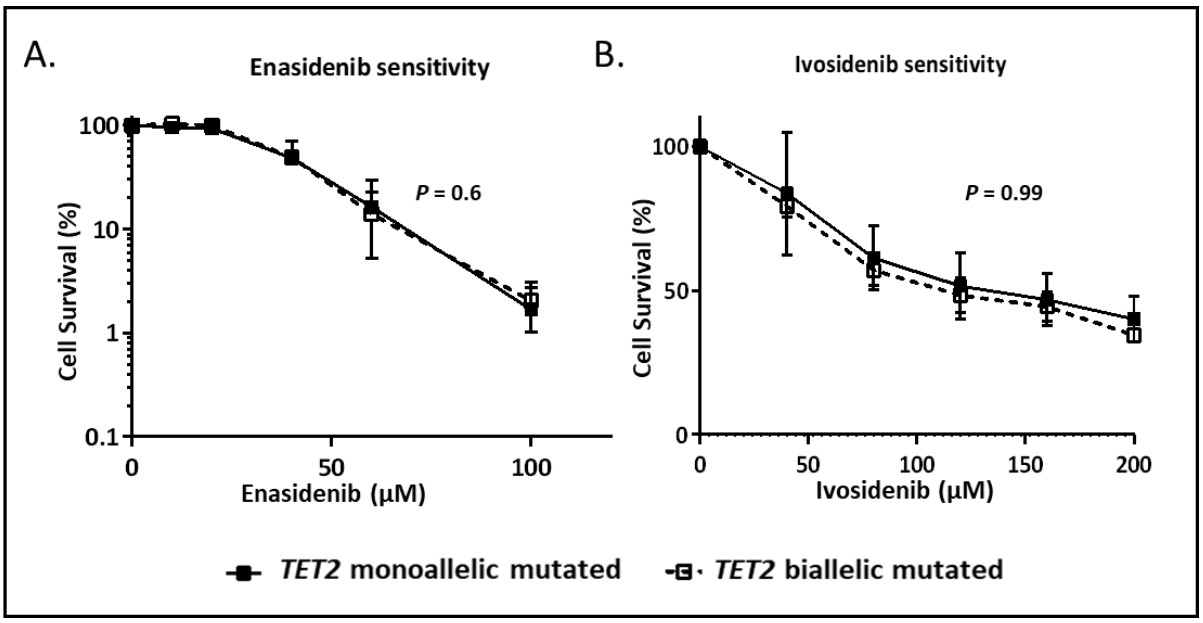

Figure 3.11. Response to IDH1/IDH2 inhibitors. The figure represents the mean growth inhibition profile over individual cell types based on their *TET2* status in response to increasing doses of (A) enasidenib (IDH2 inhibitor; 0-100 μM) and (B) ivosidenib (IDH1 inhibitor; 0-200 μM). Open symbols show the biallelic *TET2* mutants (HELbi1, HELbi2 and HELbi3) and closed symbols show the monoallelic *TET2* mutants (HEL parental, HELmono1, HELmono2 and HELmono3). The error bars are based on standard deviation of mean of three independent replicates.

IDH1/2, isocitrate dehydrogenase 1/2

Feature		*TET2* MONO				*TET2* BI			P-value
Cell lines		HEL Parental	HELmono1	HELmono2	HELmono3	HELbi1	HELbi2	HELbi3	
CRISPR treatment		untreated				treated			
TET2 mutation status		monoallelic mutation				biallelic mutation			
						4 bp deletion + 30 bp insertion	4 bp deletion + 30 bp insertion	4 bp deletion	
TET2 protein expression		Present				Completely absent			
Growth kinetics	Proliferation rate	NS				NS			0.34
	Clonogenicity	NS				NS			0.18
Drug Response Phenotype	Azacytidine	moderately sensitive				hypersensitive			**0.0001**
	Decitabine	moderately sensitive				slightly higher sensitivity (low effect size)			**0.007**
	Cytarabine	NS				NS			0.97
	Daunorubicin	NS				NS			0.49
	Fludarabine	NS				NS			0.99
	Idarubicin	NS				NS			0.24
	Calicheamicin	NS				NS			0.32
	Enasidenib	NS				NS			0.6
	Ivosidenib	NS				NS			0.99

Table 3.4. Impact of TET2 mutant allele dosage on cellular response to the anti-proliferative effects of AML therapeutics

bp, base pairs; NS, no significant difference based on *TET2* mutation status

3.5. Discussion

AML is a heterogeneous disease which has a very distinct genetic landscape of which mutations in epigenetic modifiers play a major role in pathogenesis. Therefore, therapeutic strategies aimed at exploiting epigenetic alterations in AML have now been identified as potentially efficacious. TET2 is an enzyme which actively demethylates DNA by converting 5-mC sequentially into 5-hmC, 5-fC and 5-caC, which is then converted back to cytosine by the base excision repair pathway (Moore et al. 2013).

TET2 biallelic mutations are seen in up to 25% of *TET2*-mutated AML and loss of *TET2* drives clonal expansion in pre-leukaemia (Ahn et al. 2015; Busque et al. 2018). The prognostic value of *TET2* mutations in AML have always been controversial with some studies reporting *TET2* mutations as an unfavourable prognostic marker and others as having no impact on overall survival (Feng et al. 2019; Metzeler, Maharry, Radmacher, Mrózek, et al. 2011; Nibourel et al. 2010; Chou et al. 2011). However, considering that *TET2* mutations in AML are associated with older age (P <0.001) (Metzeler, Maharry, Radmacher, Mrózek, et al. 2011) and that older patients are often not fit for traditional intensive chemotherapy, it is not surprising that the 5-year relative survival rate of patients aged above 70 years is as low as ~1% (HMRN(2010-2016)). Therefore, there is an urgent clinical requirement to develop an efficient treatment approach for *TET2*-mutated cohort of AML.

The case report of index patient described in section 1.6 suggests the potential of azacytidine as a single agent to induce long-term complete remission in the background of a *TET2* biallelic mutation. The patient was otherwise refractory to cytarabine-daunorubicin based chemotherapy. This further strengthens the hypothesis that azacytidine-based therapy might emerge as a standard of care for *TET2* biallelic mutated patients.

The first step towards testing this hypothesis was to establish a model system based on isogenic cell lines which differ only with respect to their *TET2* mutation status, and this was accomplished using CRISPR gene editing targeting exon 6 of *TET2* in the HEL AML cell line, which is reported to be *TET2* monoallelic mutated (Martin et al. 1982) and which was confirmed using SNP array analysis.

Understanding the mutational landscape of HEL cells

HEL cells are reported to possess a highly rearranged and complex genome driven by a global triploid karyotype and multiple chromosomal abnormalities suggesting a potential role for perturbed oncogenes and tumour suppressors in the sustained proliferation of these cells (Mackinnon et al. 2013). Most importantly, HEL cells carry a homozygous *JAK2V617F* mutation affecting the autoregulatory JH2 domain of JAK2 leading to constitutive activation of JAK2 tyrosine kinase, which drives cell proliferation and pro-survival signalling (Quentmeier et al. 2006). As such, the HEL cell line is growth factor independent, which is primarily attributed to the *JAK2V617F* mutation (Estes et al. 2005). Consistently, inhibition of JAK2 resulted in remarkable growth inhibition and induction of apoptosis in *JAK2*-mutant HEL cells compared to other cell lines that are wild type for *JAK2* (Estes et al. 2005; Quentmeier et al. 2006).

Despite having a large deletion on chromosome 4q capturing the *TET2* gene, our data suggest that HEL AML cells retain two intact copies of *TET2* due to chromosome 4 trisomy. Consistent with this, SNP array data confirmed the complex hypotriploid karyotype of HEL cells with three copies of most chromosomes, including chromosome 4. As such, the genetic status of monoallelic *TET2* mutants can be summarised as *TET2*-/+/+ and biallelic *TET2* mutants as *TET2*-/-/-. Specifically, the monoallelic *TET2* mutants possess two wild-type copies of *TET2* and therefore is not strictly similar to monoallelic mutation in diploid cells, where there is only one retained copy. This has been identified as a limitation of choosing HEL cells as a model system and further validation of results using an independent isogenic model system would be ideal. However, it must be noted that irrespective of the number of retained copies of *TET2*, the HEL cells are derived from a leukaemia that evolved in the background of a *TET2* deletion and this suggests the presence of characteristic phenotypic features of *TET2* mutation, such as changes in methylation profile and increased self-renewal potential in these cells. Considering the complex mutation status of HEL cells, it was necessary to validate the phenotypic impact of complete loss of *TET2* including its effect in TET2 protein expression and response to AML therapeutics.

Phenotypic characterisation of impact of TET2 mutant allele dosage

Deletions and recurrent areas of somatic copy number neutral LOH affecting chromosome 4q24 are common event in myeloid malignancies (Jankowska et al. 2009). Based on clinical data, *TET2* mutations in AML have been reported across all of the exons with the highest number of

mutations in exon 3 and 11 which are the largest exons of *TET2* (Weissmann et al. 2012). In addition, frameshift mutations within the conserved regions (between amino acids 1104-2002; exons 4-11) often generate premature stop codons resulting in protein truncation (Weissmann et al. 2012). Furthermore, the region between amino acids 1129 and 1936 (coded by exons 4-11) encodes the minimum catalytic region and any truncation affecting this region significantly reduces the catalytic activity of TET2 (Hu et al. 2013; Weissmann et al. 2012). The structure of TET2 and its various domains are discussed in detail in Section 1.4.1 and Figure 1.5. Exon 6 of *TET2* encodes for the evolutionarily conserved domain CD1 that is key to the globular structure of TET2. As such, the exon 6 mutations reported in this study, which cause premature termination, are likely to have a structural and functional impact on TET2 protein (Weissmann et al. 2012; Hu et al. 2013). Consistent with this conclusion, there was complete loss of TET2 protein expression in all the three biallelic *TET2* mutated cell clones, irrespective of the differences in their mutation profiles. Furthermore, HEL parental cells and monoallelic subclones consistently expressed TET2 protein at similar levels. Also, the antibody that was used in the western blot was raised against an antigen comprising amino acids 1-300 in the N-terminal region of the TET2 protein, which is upstream of the region coded by the CRISPR targeting site (exon 6) and therefore the antibody should potentially identify any truncated polypeptide sequence, if expressed by cell clones with biallelic mutation. However, results indicate no TET2 protein expression at all in cell clones with biallelic mutation, suggesting degradation of premature truncated TET2 polypeptide as part of protein quality control pathways (Karamyshev et al. 2018).

Irrespective of the differences in mutant allele dosage and TET2 protein status, *TET2* monoallelic and biallelic mutated cells did not differ in their proliferation rate and clonogenicity. However, Li et al. (2011) demonstrated that biallelic *TET2* mutation in mice confers an increased proliferation potential and higher cloning efficiency to $Lin^{-}Sca\text{-}1^{+}Kit^{+}$ (LSK) HSC cells from the bone marrow. It was shown that the cloning efficiency of LSK *TET2* biallelic mutant HSCs was 71% whereas that of LSK monoallelic mutant HSCs was 48% (Li et al. 2011). One possible reason for the dissimilarity in results could be the difference in experimental model (murine model and mammalian cells). Given this, it would have been ideal to test the proliferation and clonogenicity of TET2 null cells in a second independent isogenic model system based on another mammalian cell line.

Impact of TET2 mutant allele dosage on response to AML therapeutics

Nevertheless, *TET2* biallelic mutated HEL cells had significantly lower proliferation rate and colony forming efficiency in the presence of azacytidine, when compared to *TET2* monoallelic mutated HEL cells. Specifically, *TET2* biallelic mutated cells had significantly lower IC50 and IC95 values compared to *TET2* monoallelic mutated cells, despite no significant difference in baseline proliferation rate or clonogenicity in the absence of azacytidine. These data are consistent with the case history of the index patient with *TET2* biallelic mutated AML. *TET2* mutations have been associated with favourable response when treated with hypomethylating agents such as azacytidine in myelodysplastic syndrome (MDS) (Bejar et al. 2014; Lin et al. 2017; Falconi et al. 2019) despite some contrary findings (Döhner et al. 2018). In a recent phase II clinical trial on MDS patients (BMT-AZA trial; EudraCT number 2010-019673-15) *TET2* mutations were cited as a favourable prognostic factor in terms of overall survival when treated with azacytidine (Falconi et al. 2019). Likewise, Bejar et al. 2014 demonstrated that *TET2* loss sensitized MDS to azacytidine and that the response rates were the highest in the absence of *ASXL1* mutations, a gene involved in chromatin remodelling.

We further explored the relationship between *TET2* mutant allele dosage and response to hypomethylating agents by investigating decitabine, a functional analogue of azacytidine widely used in the treatment of MDS (Malik et al. 2014). Azacytidine and decitabine (5-aza-2′-deoxycytidine) function by inactivating DNMT1 which leads to demethylation of previously silenced tumour suppressor genes and results in reducing cell proliferation/ promoting apoptosis/ regulating cell cycle (Derissen et al. 2013). *TET2* biallelic mutated cells were clearly more sensitive to decitabine compared to *TET2* monoallelic mutated cells, although the effect size was weaker. Knockdown of *TET2* in KG-1 cells, which otherwise express high levels of TET2, increased the vulnerability of these cells to decitabine and the effect is predicted to be mediated by cell cycle and DNA replication related genes (Chen et al. 2015). As such, gene expression profiling demonstrated differential expression of cell cycle related genes including *CCNB2*, *RBL1* and DNA replication related genes including *PRIM1, RCF3, FEN1*, which was coupled with concomitant alteration in cell cycle profile as demonstrated by flow cytometry (Chen et al. 2015). Considering the hypomethylating function of TET2, it is mechanistically plausible that *TET2* biallelic mutated cells are more sensitive to hypomethylating agents such as azacytidine and decitabine.

The relationship between *TET2* mutant allele dosage and response to standard remission induction therapies was also investigated. Specifically, *TET2* mutant allele dosage had no

significant impact on proliferation following exposure to agents including cytarabine, daunorubicin, fludarabine, idarubicin and calicheamicin. Calicheamicin is the cytotoxic drug component of the antibody-drug conjugate (ADC) GO, which was the first approved ADC for AML induction therapy (Yu et al. 2019). A meta-analysis based on 4577 cases altogether from 16 different clinical studies, mostly treated with standard induction regimens based on agents such as cytarabine, daunorubicin, idarubicin, GO (with the exception of two studies where treatment strategy is unknown), demonstrated *TET2* mutation as an unfavourable prognostic marker (Wang et al. 2019). Furthermore, Metzeler and colleagues reported that *TET2* mutations were significantly associated with lower complete remission rate and shorter event-free survival in patients with favourable-risk CN-AML, when treated with standard induction treatment regimens (Metzeler, Maharry, Radmacher, Mrózek, et al. 2011), although this trend was not seen in other studies (Nibourel et al. 2010; Kosmider et al. 2011). However, none of these studies particularly investigated the effect of *TET2* mutant allele dosage on cellular response to induction therapeutics. As such, our results suggests that *TET2* mutant allele dosage does not have a significant impact on cellular response to standard induction therapeutic agents.

Pagliuca et al. 2020 reported synthetic lethality through TET2 inhibition in *IDH1/IDH2* mutated cells and this was attributed to the phenotypic redundancy of TET2 and IDH1/IDH2. To test whether the reverse is true, *TET2* mutated cells were exposed to IDH1/IDH2 inhibitors. However, inhibition of IDH1/IDH2 was not synthetically lethal in TET2 null cells. This could be because of the phenotypic difference between inhibition of IDH1/IDH2 and mutated *IDH1/IDH2*. Specifically, *IDH1/IDH2* mutation increases the production of 2-hydroxyglutarate (2-HG) instead of α-ketoglutarate (α-KG), whereas *IDH1/IDH2* inhibitors such as ivosidenib and enasidenib are reported to reduce the levels of 2-HG (Liu et al. 2019). Therefore, instead of targeting cells with IDH1/IDH2 inhibitors, treating cells with 2-HG (which better mimics *IDH1/IDH2* mutation) might induce synthetic lethality in *TET2* biallelic mutated AML. Consistently, a recent study reported synthetic lethality in *TET2* deficient cells when treated with *IDH1/IDH2* mutant-derived-2HG in myeloid neoplasms. (Guan et al. 2021).

In summary, *TET2* biallelic mutation (TET2 null phenotype) renders AML cells hypersensitive to azacytidine relative to cells with monoallelic mutation. Likewise, complete loss of TET2 function also sensitises AML cells to decitabine, although the differential sensitivity is less striking than seen with azacytidine. Several randomised controlled phase III trial in patients with MDS demonstrated similar response rates for azacytidine and decitabine (Hollenbach et al. 2010). However, overall survival was significantly higher for azacytidine treated patients compared to conventional care regimes (best supportive care, low-dose cytarabine, or intensive

chemotherapy) whereas decitabine did not show a significant effect on overall survival in a similar trial (Hollenbach et al. 2010; Fenaux et al. 2009; Lübbert et al. 2011). Another study that compared the sensitivity of human cancer cell lines *in vitro* to azacytidine and decitabine demonstrated low correlation between the sensitivity pattern observed in response to the two agents (Hollenbach et al. 2010). In addition, the decitabine resistant cells developed from HL60 cell line were still sensitive to azacytidine (Qin et al. 2009; Hollenbach et al. 2010). Moreover, azacytidine and decitabine had distinct impact on gene expression profiles of Kasumi-1 AML cell line such that the target genes as well as the extent of demethylation induced by the two agents varied for individual genes (Flotho et al. 2009). For instance, the homeobox gene *HOXA9* was upregulated in response to decitabine but not azacytidine (Flotho et al. 2009). Therefore, it is not surprising to observe differential response to the two hypomethylating agents.

Azacytidine and decitabine are structurally and functionally very similar and the primary differences between the two agents are related to incorporation into nucleic acids. Azacytidine is primarily incorporated in RNA (80%-90%) and RNR mediated conversion of the drug facilitates its incorporation into DNA (10%-20%) (Stomper et al. 2021; Griffiths et al. 2008). As such, the incorporated azacytidine inhibits DNA methylation, RNA methylation and protein synthesis (Griffiths et al. 2008). However, decitabine is solely incorporated into DNA and does not have an impact on RNA methylation or protein synthesis (Griffiths et al. 2008). The exact mechanism behind the differential impact of *TET2* mutant allele dosage on cellular response to the two hypomethylating agents are unclear. Given the differences in cellular uptake and mechanism of action of azacytidine and decitabine, it can be speculated that *TET2* mutant allele dosage might have a differential or greater impact on azacytidine-specific pathways (RNA-related and protein synthesis) compared to the pathways that facilitate decitabine incorporation.

Further investigations were performed to elucidate the complete mechanism behind the chemosensitivity of *TET2* biallelic mutated cells to azacytidine and this will be discussed in the upcoming chapters. Nevertheless, the findings present here clearly support the use of azacytidine for the treatment of *TET2* biallelic mutated AML.

3.5.1. Summary of Chapter

In summary, the results of this chapter demonstrate that:

(i) *TET2* biallelic mutation leads to complete loss of TET2 protein expression, based on our model system.

(ii) *TET2* mutant allele dosage did not have an impact on spontaneous proliferation rate or clonogenicity of cells.

(iii) *TET2* biallelic mutation sensitises AML cells to hypomethylating agents, in particular leads to hypersensitivity to azacytidine and the mechanism behind this will be studied in detail in upcoming chapters.

Chapter 4. The Methylomic and Transcriptomic profile of *TET2* monoallelic and biallelic mutated cells.

4.1. Introduction

Altered DNA methylation leading to aberrant silencing of tumour suppressor genes and expression of oncogenes is a common event in leukaemogenesis and has been attributed to frequent mutations in epigenetic modifiers including *TET2* (Rasmussen et al. 2015; Baylin et al. 2011; Shih et al. 2012). Considering the role of TET2 in regulating demethylation (Section 1.4.2), it is important to determine the effect of *TET2* mutant allele dosage on genomic methylation and to further investigate the impact on global gene expression. In this project, the Illumina Infinium HumanMethylation450 (450K) BeadChip array was used to explore the DNA methylome, and RNAseq the transcriptome, of *TET2* monoallelic and biallelic mutated HEL AML cells.

Illumina Infinium HumanMethylation450 (450K) BeadChip array

Infinium methylation arrays have evolved as a preferred method for genome-wide methylation analysis. The 450K methylation array evaluates the methylation status of >450000 CpG sites throughout the genome, which covers 96% of CpG islands in the genome (Wang et al. 2018; Bibikova et al. 2011). This method relies on quantitative genotyping of bisulphite converted DNA which uses target-specific oligonucleotides conjugated to beads in order to interrogate the methylation levels of each CpG site (Wilhelm-Benartzi et al. 2013; Bibikova et al. 2011). Bisulphite treatment converts unmethylated cytosine to uracil via sequential sulphonation, hydrolytic deamination and alkali desulphonation (Kristensen et al. 2009) (Figure 4.1A). In contrast, 5-mC remains unmodified as it is resistant to bisulphite treatment (Figure 4.1B). The final dual colour read out results from two fluorescently labelled nucleotides; adenine (‘A’, red) which pairs with the unmethylated (U) uracil nucleotide and guanine (‘G’; green) which pairs with the methylated (M) cytosine (Figure 4.1C) (Wilhelm-Benartzi et al. 2013; Bibikova et al. 2011; Wang et al. 2018).

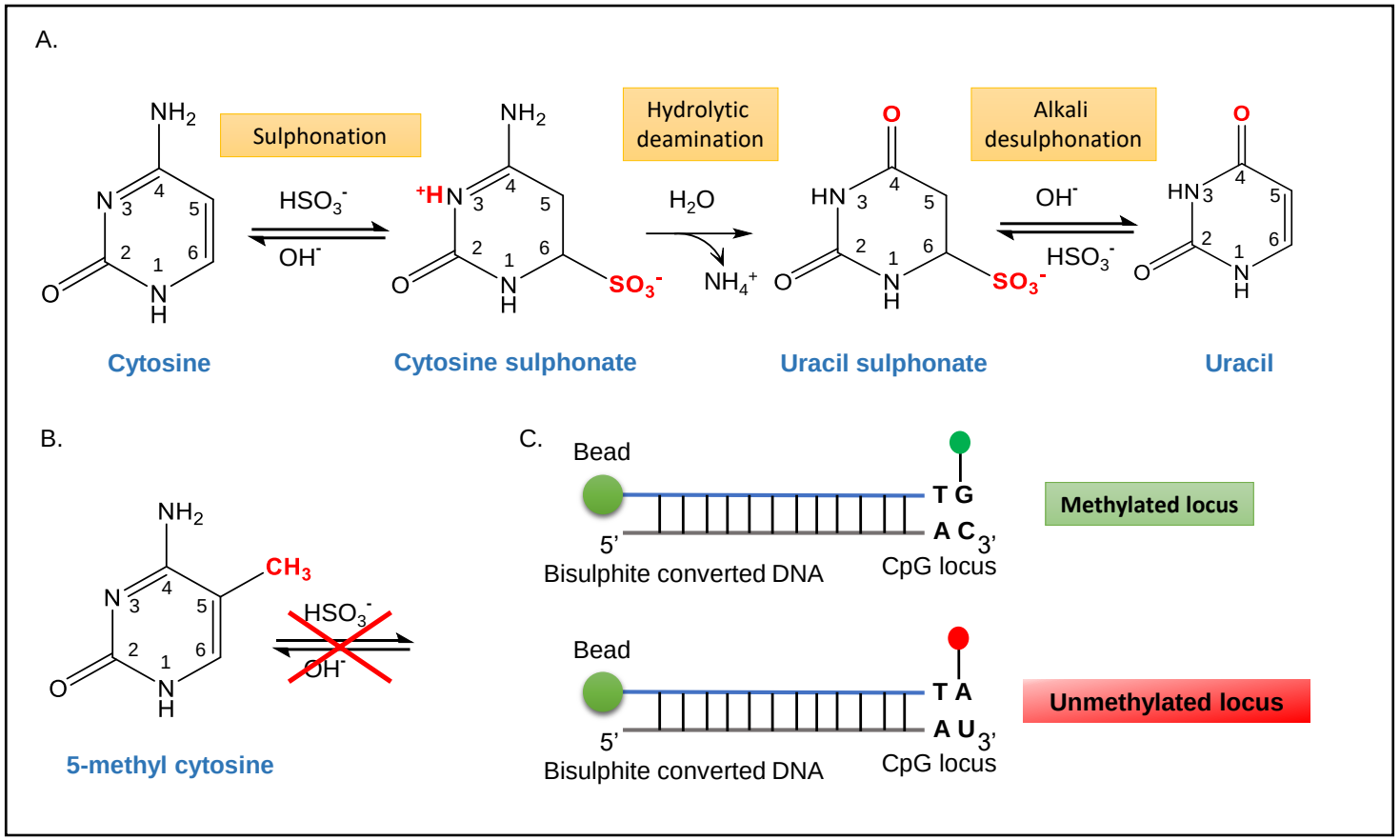

Figure 4.1. Illumina Infinium HumanMethylation450 (450K) BeadChip array. (A) Bisulphite conversion coverts unmethylated cytosine sequentially into uracil. (B) Methylated cytosine residues are resistant to bisulphite conversion. The chemical structures were generated using chemdraw tool. (C) Schematic representation of illumina 450K methylation array. Following bisulphite conversion, the unmethylated cytosines that are converted to uracil nucleotides are targeted by red fluorescently labelled adenine nucleotides, and the methylated cytosines which remains intact are targeted by green fluorescently labelled guanine nucleotides. Figures modified from (Kristensen et al. 2009; Bibikova et al. 2011).

A, adenine; C, cytosine; CpG, cytosine and guanine linked by a phosphate group; G, guanine; T, thymine; U, uracil

The β-value determines the methylation level of each CpG site and is computed as the ratio of methylated probe intensity to the overall intensity (Equation 4.1) The β-value is a number between 0 to 1, where under optimal conditions 0 indicates a completely unmethylated CpG site and 1 indicates a completely methylated site (Du et al. 2010). The M-value, which is the log transformation of the intensities of methylated probes to unmethylated probes (Equation 4.1), assumes positive values when there are more methylated probes compared to unmethylated probes and negative values when unmethylated probes outnumber methylated probes (Du et al. 2010; Bibikova et al. 2011).

$$\beta = \frac{M}{M + U + \alpha}; \mathrm{M} = \log_2 \frac{M}{U}$$

Equation 4.1. Formula to calculate β and M values. *M* is the intensity of methylated signals, *U* is the intensity of unmethylated signals and α is a constant (100 generally) (Du et al. 2010).

RNA sequencing

RNA sequencing (RNAseq) is a revolutionary tool developed for whole transcriptome profiling which substantially expanded our understanding of eukaryotic gene expression. RNAseq is a deep-sequencing tool that allows precise mapping and quantification of gene expression, which can be used to identify different RNA species including total RNA, pre-mRNA, and noncoding RNA (ncRNA), such as long ncRNA and micro RNA (miRNA) (Kukurba et al. 2015; Wang et al. 2009) (Figure 4.2).

RNAseq analysis requires the isolation of RNA from cells/biological samples, followed by the generation of a transcriptome library, often referred to as the RNAseq library. Ribosomal RNA (rRNA) accounts for almost 95% of the total RNA pool in most cell types and if not removed can constitute the bulk of the sequencing reads and therefore limit the detection of other less-abundant species (Kukurba et al. 2015). As such, it is important to isolate the desired RNA species from the total RNA pool using appropriate selection methods, such as poly-A selection and ribo-depletion which selectively removes rRNA from the total RNA pool. Although both selection methods enrich mRNA species, unlike poly-A selection, ribo-depletion also enriches for ncRNA and pre-mRNA, which are not post-translationally modified (Kukurba et al. 2015). As such, the choice of library selection method entirely depends on the particular research question. For example, if an accurate quantification of non-coding and coding RNA sequence

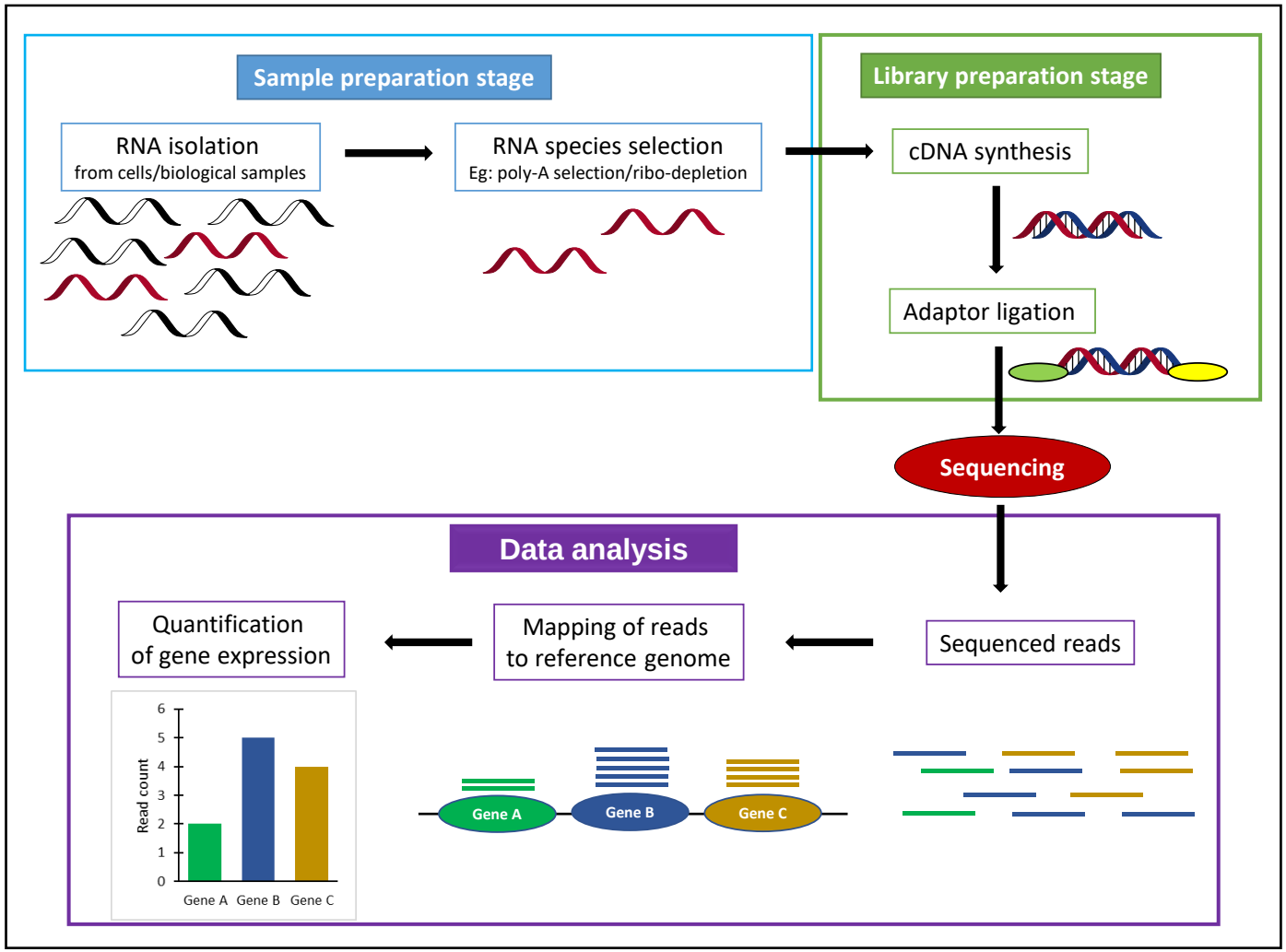

Figure 4.2. Schematic representation of RNAseq pipeline.

The various stages of RNAseq are sample preparation (blue boxed), library preparation (green boxed), sequencing (red) and data analysis (violet boxed). The sample preparation stage involves RNA isolation from cell/biological samples followed by RNA species enrichment. RNA is then converted to cDNA using a reverse transcriptase, and sequence adaptors are ligated to the ends of the cDNA strands forming the cDNA library. The library is then sequenced in a high-throughput manner. During data analysis, the sequenced reads are mapped to the respective gene regions in the reference genomes. Finally, gene expression is quantified as the number of reads aligning to a particular gene.

cDNA, complementary DNA; DNA, deoxyribonucleic acid; poly-A, poly adenylation; RNA, ribonucleic acid

is required, ribo-depletion is optimal, whereas for mRNA sequencing poly-A selection would be preferred (Kukurba et al. 2015).

The RNA preparation step is followed by complementary DNA (cDNA) library preparation where RNA is converted into cDNA with the use of a reverse transcriptase (Wang et al. 2009). Furthermore, specific sequencing adaptors are attached to one or both ends of the cDNA strands, which are specific to the sequencing platform used. The adaptors consist of critical components that aid sequencing, such as the amplification element near the 5' and/or 3' ends, and the primer sequencing priming site adjacent to the cDNA insert (Voelker et al. 2021). Such adaptor ligated cDNA sequences are then sequenced using a high throughput platform at 10-30 million reads per sample (Stark et al. 2019). The sequencing step is followed by rigorous *in silico* quality control culminating in mapping of reads to their respective regions of the reference genome. Finally, gene expression levels are quantified and expressed as the number of reads (read count) aligned to each gene (Kukurba et al. 2015) (Figure 4.2). Further downstream analysis of RNAseq data can be performed to identify differential expression of genes and pathways in experimental samples, making it an ideal platform to validate the functional impact of *TET2* mutant allele dosage on the transcriptomic profile of cells.

4.2. Aims of the Chapter

Considering the demethylation function of TET2, it is mechanistically plausible that *TET2* biallelic mutation leads to a global shift in methylomic and transcriptomic profile of cells, with multiple genes and pathways being impacted. This chapter aims to characterise the methylation and gene expression profile of monoallelic and biallelic *TET2* mutated cells using Illumina 450k methylation array and RNAseq analysis. Specifically, the experimental aims of this chapter are as follows:

- To elucidate the functional impact of *TET2* mutant allele dosage on genome-wide methylation using the illumina 450k methylation array and to identify differential methylation in *TET2* biallelic mutated cells compared to *TET2* monoallelic mutated cells.
- To understand the gene expression profile of *TET2* monoallelic and biallelic mutated cells using RNAseq.
- To identify the top differentially expressed genes and pathways in *TET2* biallelic mutated cells compared to *TET2* monoallelic mutated cells, which might potentially have a role in determining sensitivity to azacytidine.
- To validate the protein level expression pattern of selected differentially expressed pathway members and to validate the role of differentially expressed genes/pathway members in sensitising *TET2* biallelic mutated cells to azacytidine, using selective target inhibition studies or other appropriate functional assays.

4.3. Results

4.3.1. Methylation profiling through Illumina 450K methylation arrays

The Illumina 450k human methylation array was used to investigate the CpG methylation pattern in *TET2* monoallelic (n=2) and biallelic (n=4) mutated HEL AML cells. An unsupervised hierarchical clustering of top 1500 differentially methylated CpG probes across all samples demonstrated distinct clustering based on *TET2* status (Figure 4.3)

An analysis of *TET2* monoallelic and biallelic clones identified 21530 significantly differentially methylated probes [P_{adj}<0.05], with 7960 hypermethylated probes [$\log_2FC \geq 2$, P_{adj}<0.05] and 4584 hypomethylated probes [$\log_2FC \leq -2$, P_{adj}<0.05] (Figure 4.4). As such, complete loss of TET2 resulted in the acquisition of both methylated and unmethylated probes, although overall there was a clear acquisition of a hypermethylated phenotype in TET2 null cells, and this is entirely consistent with the demethylation function of TET2.

4.3.2. Transcriptome profiling using RNAseq

Transcriptome profiling of *TET2* biallelic and monoallelic mutated cells was achieved using RNAseq analysis. An unsupervised analysis was performed to cluster the cells based on the top 1500 transcripts with the highest variance across all samples (Figure 4.5A). HELmono2 and HELmono3 cells (*TET2* monoallelic mutants) clustered together in one group and HELbi1, HELbi2 and HELbi3 (*TET2* monoallelic mutants) clustered together in another group. However, HELmono1 cells (*TET2* monoallelic mutant) clustered independently but closer to the group of biallelic *TET2* mutants (Figure 4.5A).

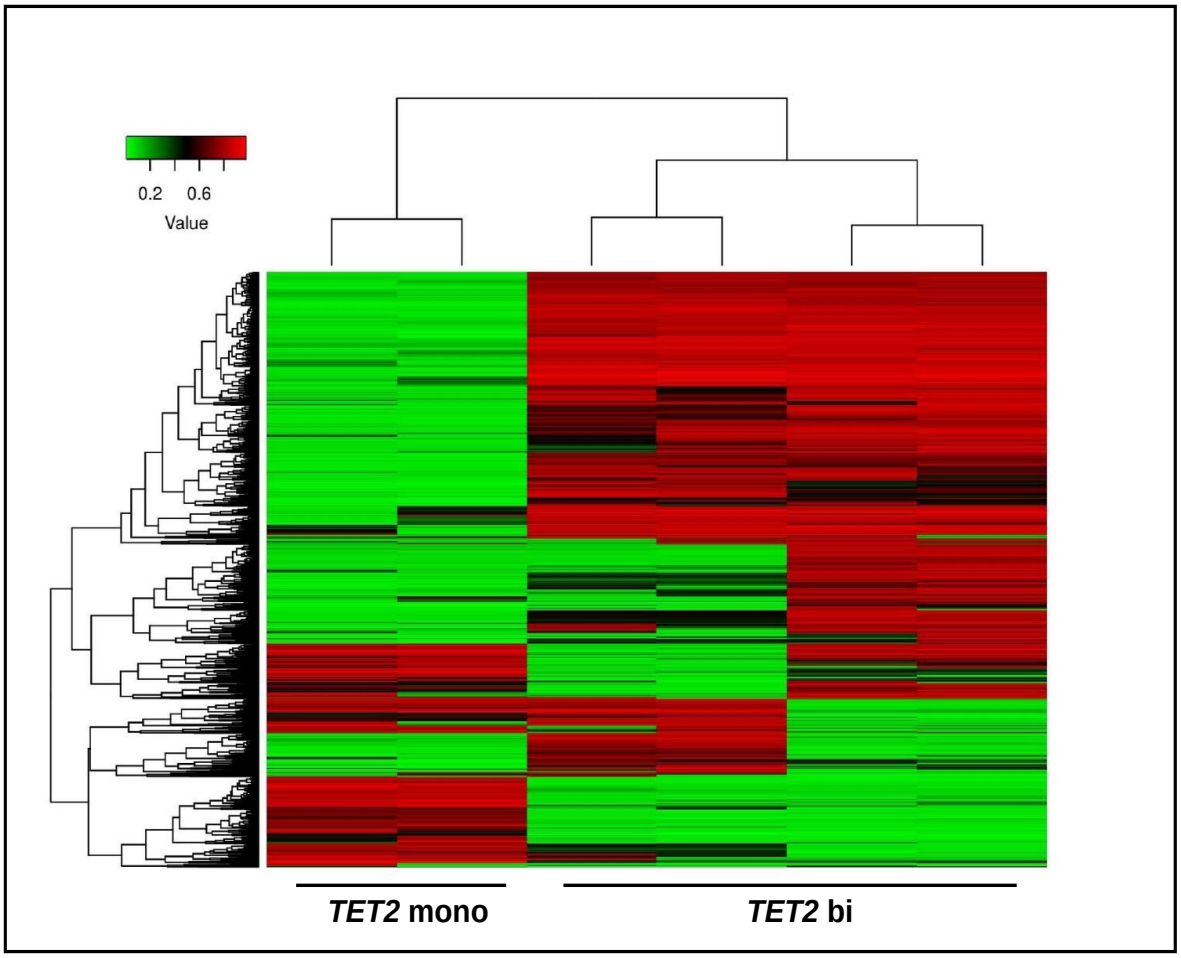

Figure 4.3. Hierarchical unsupervised clustering of the top 1500 differentially methylated CpG probes across all samples. Heatmap shows distinct clustering of monoallelic (n=2) and biallelic (n=4) *TET2* mutant cell clones. In the heatmap, the vertical columns denote each cell clone, the horizontal rows denote individual CpG probes, and the red-green colour gradient represents the relative methylation level of each locus.

CpG, cytosine and guanine linked by a phosphate group; *TET2* bi, *TET2* biallelic mutated; *TET2* mono, *TET2* monoallelic mutated

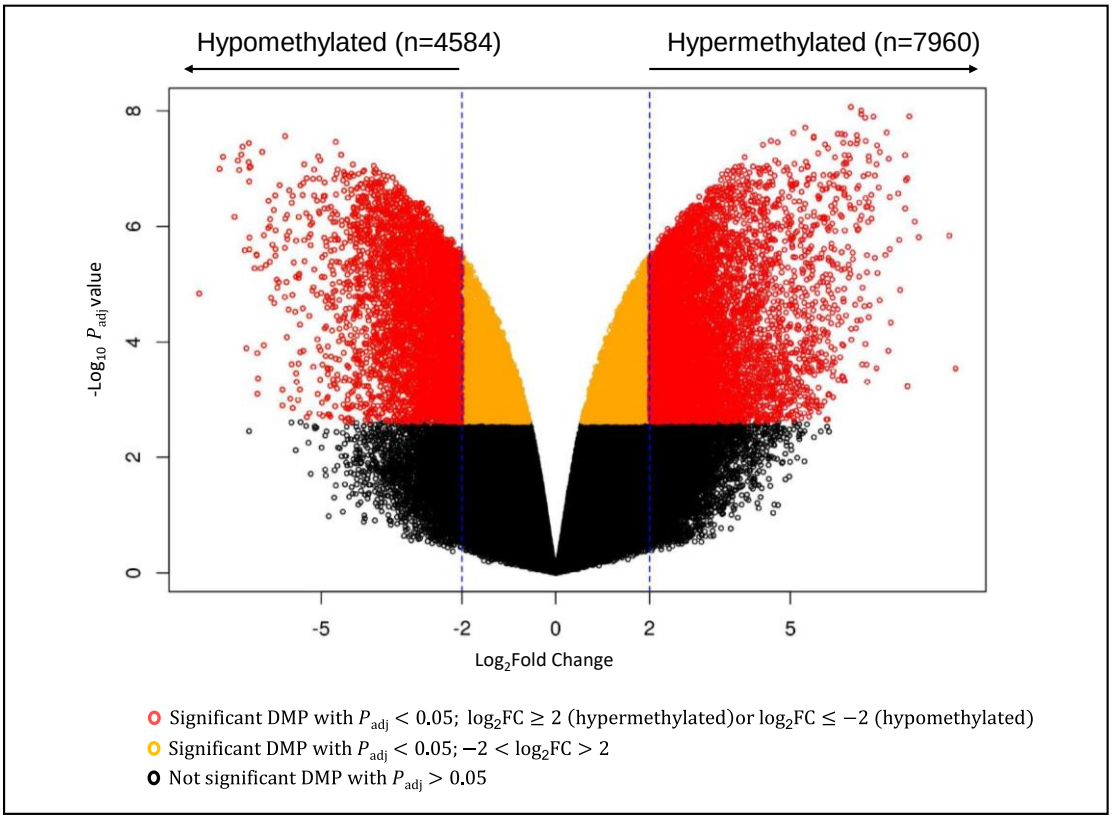

Figure 4.4. Complete loss of TET2 leads to acquisition of a hypermethylated phenotype. Volcano plot showing the differentially methylated CpG probes between *TET2* monoallelic (n=2) and biallelic mutated samples (n=4). Fold change (Log_2FC) values are plotted against *P* values ($Log_{10}P_{adj}$ value) and the coloured dots denote individual CpG probes. The red-coloured probes show the significantly differentially methylated probes that passed the fold change threshold, black-coloured probes show the non-significant probes ($P_{adj}\geq0.05$) and yellow-coloured probes show the significantly differentially methylated probes which did not pass the fold change threshold. 7960 probes were significantly hypermethylated [$log_2FC \geq 2$, $P<0.05$] and 4584 probes were significantly hypomethylated [$log_2FC \geq 2$, $P<0.05$].

CpG, cytosine and guanine linked by a phosphate group; DMP, differentially methylated probes; FC, fold change

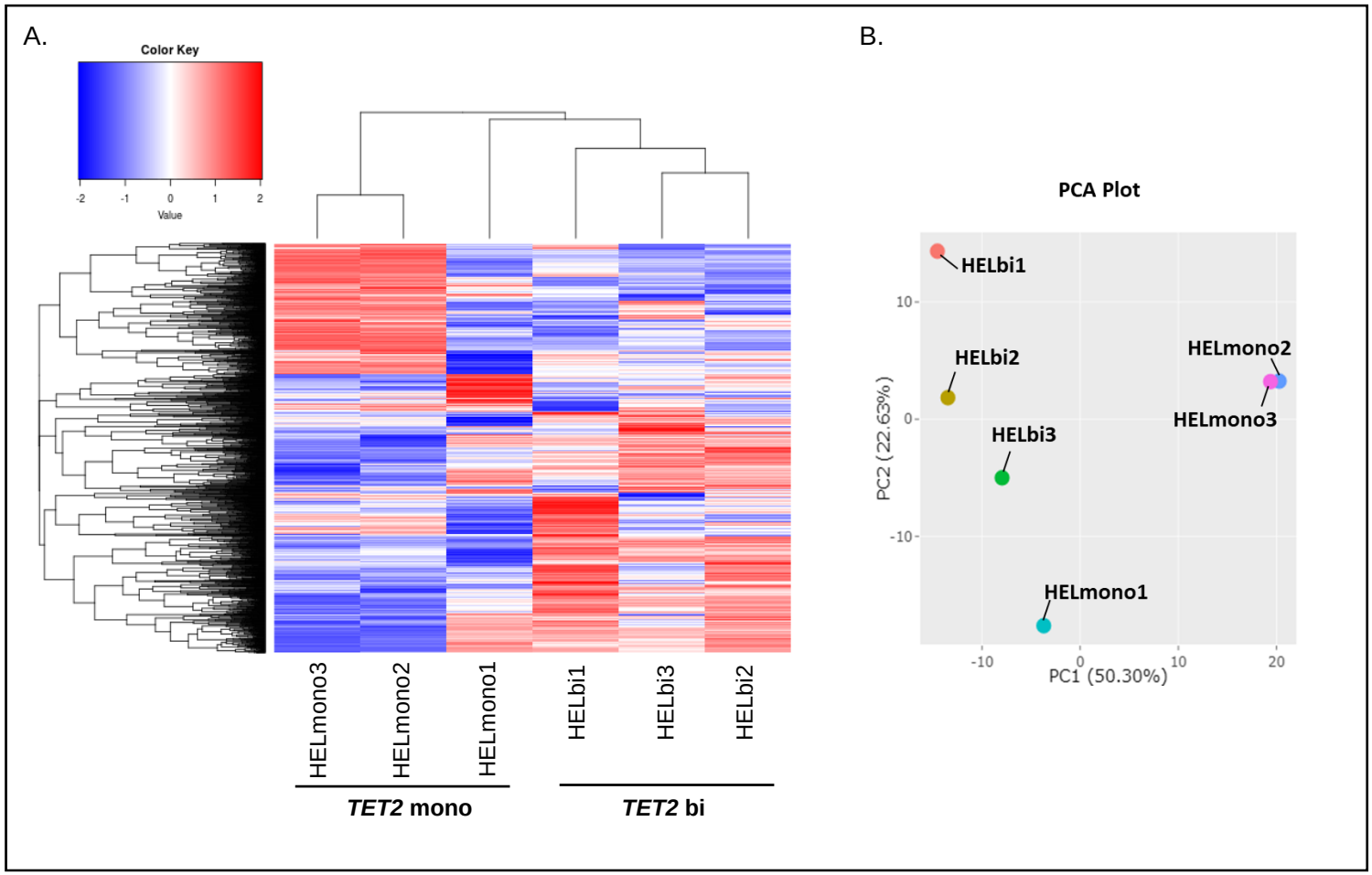

Figure 4.5. Unsupervised hierarchical clustering of *TET2* monoallelic and biallelic mutated cells. A. Heatmap showing unsupervised hierarchical clustering of the top 1500 transcripts with the highest variance across all samples. In the heatmap, the vertical columns denote each cell clone, the horizontal rows denote individual transcripts, and the red-blue colour gradient represents the relative up-down expression levels of each transcript, respectively. B. PCA plot that groups the samples based on first two principal components with highest variance. The clustering is broadly consistent with the *TET2* mutation status, except for HELmono1 cells which clustered independently, but closer to *TET2* biallelic mutated cell clones.

PC1, Principal component 1; PC2, Principal component 2; PCA, Principal component analysis

The principal component analysis (PCA) is another unsupervised analysis approach which essentially simplifies the complex dimensionality of large datasets while retaining the global trend or pattern of the data (Lever et al. 2017). This is particularly helpful in RNAseq analysis which involves measuring multiple features such as the expression of many genes for each sample (Ringnér 2008; Jolliffe et al. 2016; Lever et al. 2017). Clustering pattern was similar in PCA plot and unsupervised clustering based heatmap (Figure 4.5B). HELmono2 and HELmono3 clustered together following PCA demonstrating a similar transcriptomic profile, whereas HELmono1 clustered independently from the other *TET2* monoallelic mutated clones, with PC1 score similar to that of the *TET2* biallelic mutant HEL cells (Figure 4.5B).

A supervised analysis based on *TET2* status shows that 698 transcripts were significantly differentially expressed (P_{adj}<0.05; |Log_2FC|≥0.3) (Figure 4.6A), which included 369 significantly upregulated and 329 significantly downregulated transcripts in *TET2* biallelic mutated HEL cells compared to *TET2* monoallelic mutated HEL cells (P_{adj}<0.05; |Log_2FC|≥0.3) (Figure 4.6B). The complete list of differentially expressed transcripts is provided in Appendix B.

4.3.3. Pathway analysis based on RNA sequencing results

Gene enrichment analysis was performed to understand if the differentially expressed transcripts (P_{adj}<0.05; |Log_2FC|≥0.3) were associated with any relevant biological processes potentially involved in determining cellular response to azacytidine. Gene ontology (GO) pathway enrichment analysis was used to group the genes by their broad function and to compare the differential expression of pathways in *TET2* monoallelic and biallelic mutant HEL cells. Four out of five top hits among the differentially affected pathways were spliceosome-related (Table 4.1; Full list of pathways in Appendix C) and included the spliceosomal tri-snRNP (small nuclear ribonucleoprotein) complex, the formation of quadruple SL/U4/U5/U6 snRNP, mRNA trans-splicing via spliceosome and mRNA trans-splicing via SL (spliced leader) addition.

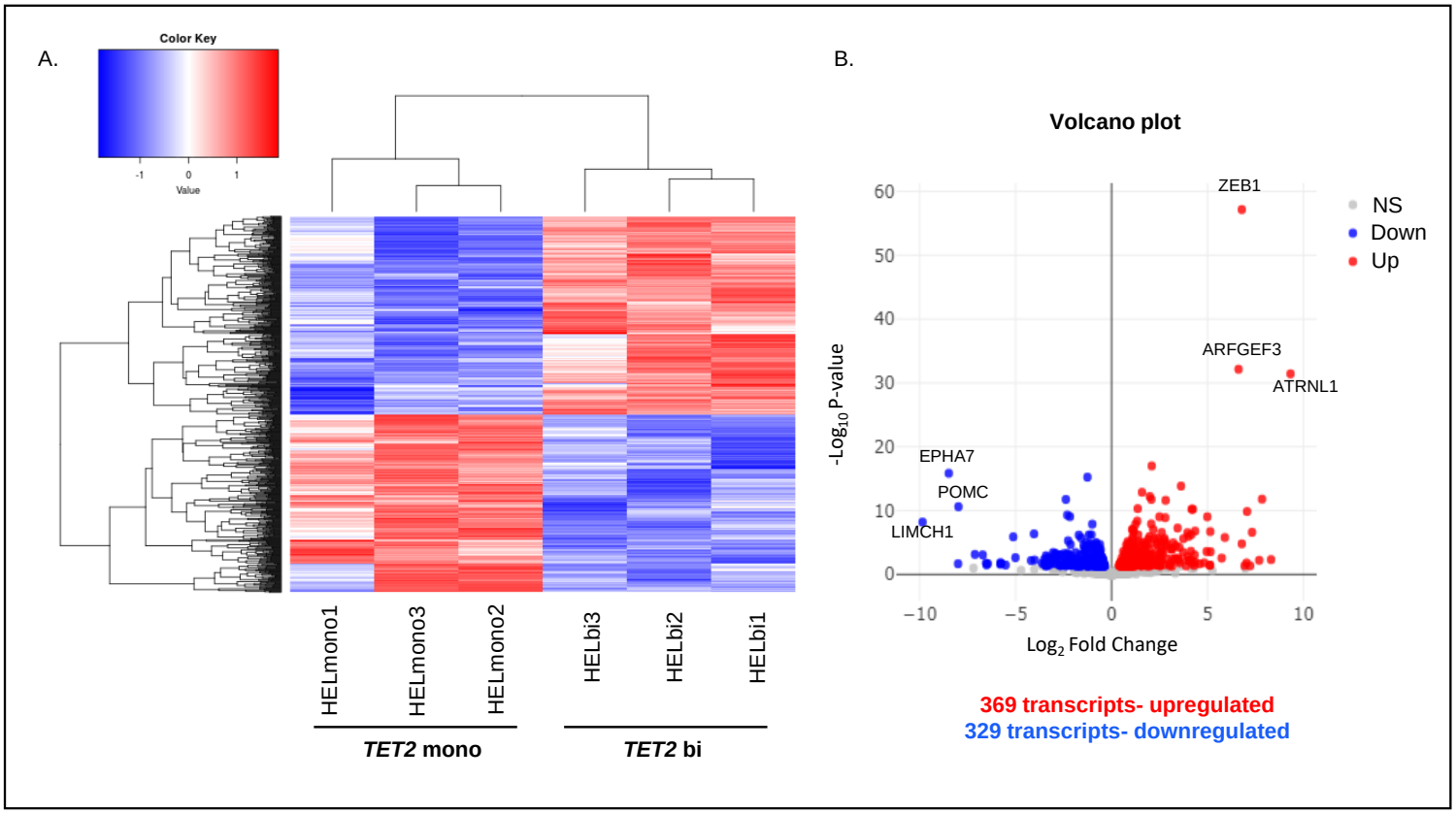

Figure 4.6. RNAseq transcriptome analysis of *TET2* monoallelic and biallelic mutated HEL cells. (A) Heatmap showing the hierarchical supervised clustering of significantly differentially expressed transcripts (P_{adj}<0.05; |Log_2FC|≥0.3) based on *TET2* mutation status. In the heatmap, the vertical columns denote each cell clone, the horizontal rows denote individual transcripts and colour indicates relative gene expression. (B) Volcano plot generated following a supervised analysis (P_{adj}<0.05; |Log_2FC|≥0.3) based on *TET2* status showing 369 upregulated and 329 downregulated transcripts in *TET2* biallelic mutant cells compared to *TET2* monoallelic mutant cells. The significantly up and down regulated transcripts (P_{adj}<0.05; |Log_2FC|≥0.3) are coloured red and blue, respectively, in the volcano plot.

Down, downregulated; FC, fold change; NS, non-significant; Up, upregulated

Pathways	Padj value
Spliceosomal tri-snRNP complex (GO:0000244)	0.00036
Formation of quadruple SL/U4/U5/U6 snRNP (GO:0000353)	0.00144
mRNA trans-splicing via spliceosome (GO:0000365)	0.00144
mRNA trans-splicing via SL addition (GO: 0045291)	0.00144
Ossification (GO: 001543)	0.00154

Table 4.1.Top five pathways retrieved from gene ontology (enrichGO) pathway analysis of differentially expressed transcripts* based on *TET2* mutation status
**(P_{adj}<0.05; |Log_2FC|≥0.3)*

Thirteen unique differentially expressed transcripts are components of the top four spliceosome related pathways, namely RNU1-1, RNVU1-6, RNVU1-7 and RNVU1-18 (part of U1 snRNA), RNU4-1 and RNU4-2 (part of U4 snRNA), RNU5A-1, RNU5B-1, RNU5D-1 and RNU5F-1 (part of U5 snRNA), RNU4ATAC and RNU6ATAC (part of minor spliceosome) and LSM8 (Figure 4.7A) (Will et al. 2011; Schneider et al. 2002; Cloutier et al. 2018). In general, these spliceosomal pathway components were downregulated in *TET2* biallelic mutated HEL cells compared to monoallelic mutated cells (Figure 4.7A). Based on antibody availability, protein expression of LSM8 was confirmed to be consistent with RNA transcript levels (Figure 4.7B-D). Specifically, *TET2* biallelic mutated HEL cells expressed lower levels of LSM8 compared to *TET2* monoallelic HEL cells (Figure 4.7E).

Cheng et al. 2018 described a unique interaction of HNRNPK (Heterogeneous Nuclear Ribonucleoprotein K) which is a component of the spliceosomal pathway to be involved in modulating response to azacytidine. Irrespective of the interaction, the individual proteins levels of HNRNPK, NSUN1, RNA Pol II subunit A phosphorylated at serine 2 (POLR2A S2P) and CDK7 were upregulated in azacytidine resistant leukaemia cells compared to their isogenic azacytidine sensitive cells (Cheng et al. 2018). As such, protein expression of HNRNPK and other RNA methylation related genes were tested in HEL AML cells (Figure 4.8A). The transcript and protein levels of NSUN1, HNRNPK, CDK7 and POLR2A S2P respectively were not significantly associated with *TET2* mutant allele dosage (Figure 4.8B-E).

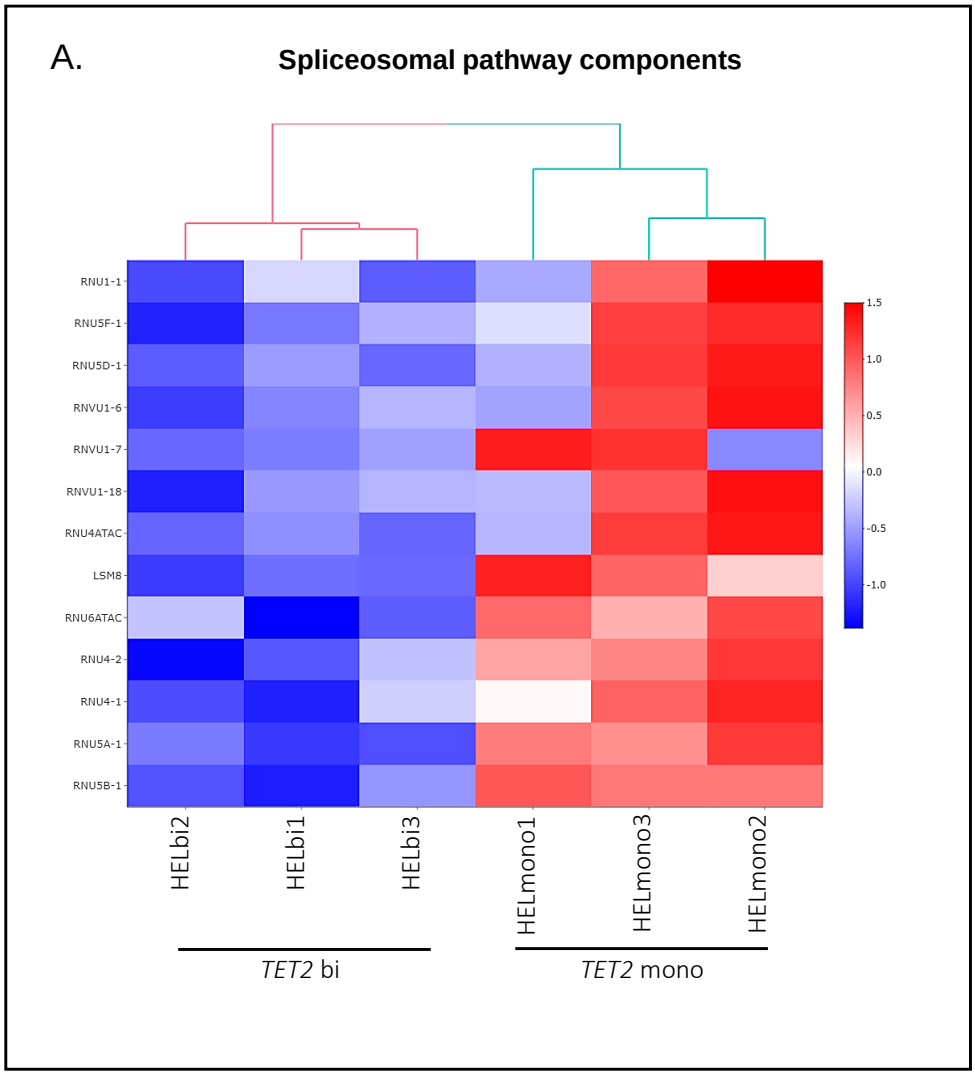

Figure 4.7. Spliceosomal pathway components are downregulated in *TET2* biallelic mutated cells compared to monoallelic mutated cells (continue on next page)

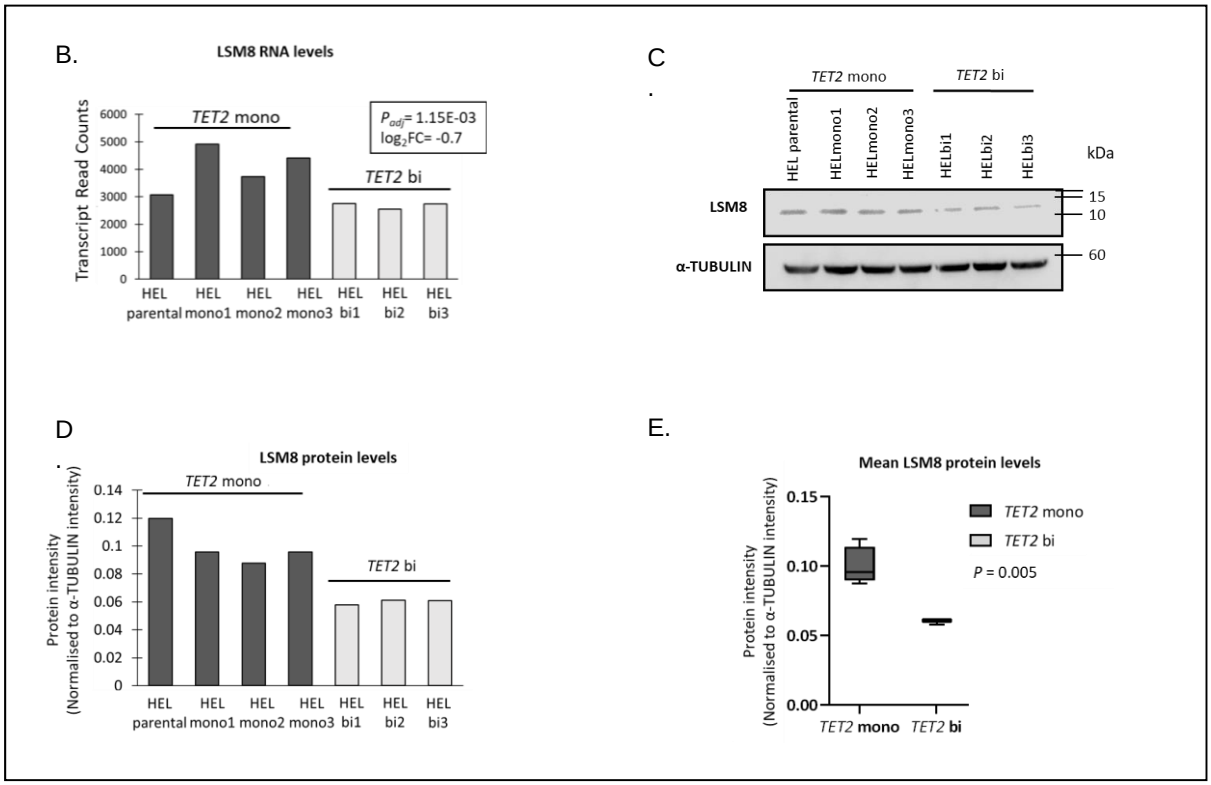

Figure 4.7. Spliceosomal pathway components are downregulated in *TET2* biallelic mutated cells compared to monoallelic mutated cells (continued from previous page) (A) Heatmap showing the expression pattern of 13 differentially expressed transcripts of the spliceosomal pathway across individual cell clones stratified by *TET2* mutation status. The vertical columns denote each cell clone, the horizontal rows denote individual transcripts and colour indicates relative gene expression. (B) LSM8 RNA level variation (C) LSM8 protein levels detected using western blot (D) quantification of protein levels normalised to α-TUBULIN. (E) Mean LSM8 protein levels of cells grouped based on *TET2* mutation status. LSM8 was significantly downregulated (P = 0.005) in *TET2* biallelic mutated cells compared to monoallelic mutated cells. In the box and whiskers plot, the bars represent the mean, the boxes represent the interquartile range, and the whiskers represent the range between minimum and maximum values. Dark grey denotes *TET2* monoallelic mutants and light grey denotes *TET2* biallelic mutants.

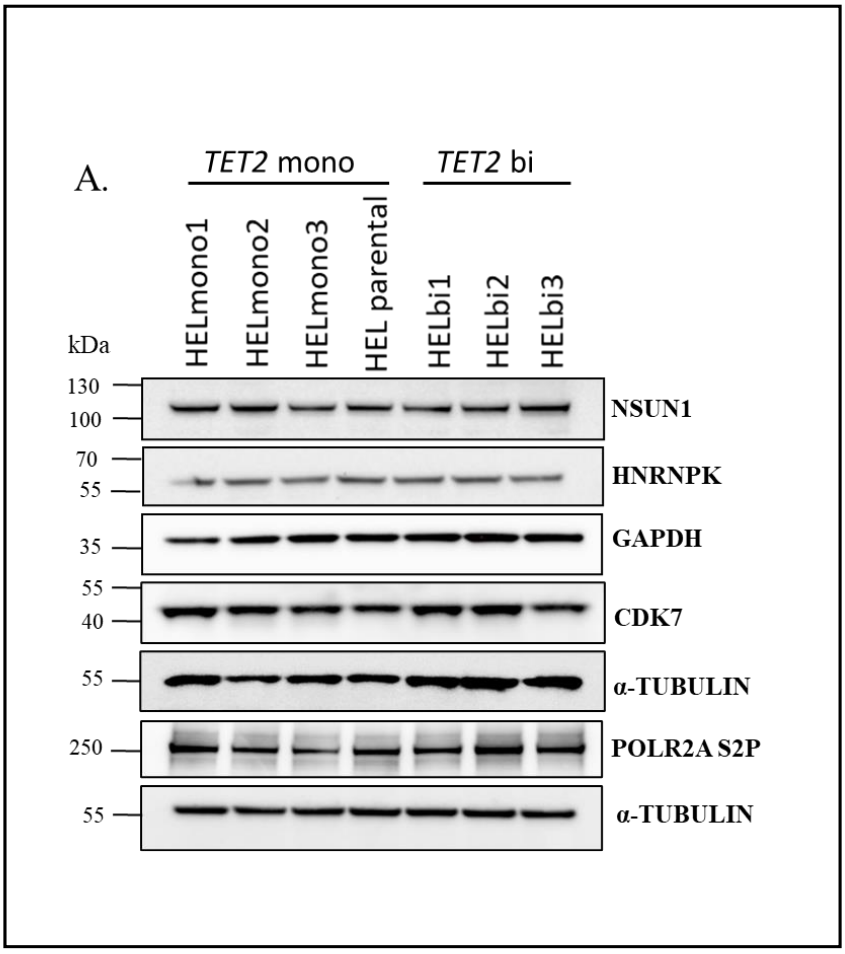

Figure 4.8. Expression of NSUN1, HNRNPK, CDK7 and POLR2A S2P (continued on next page)

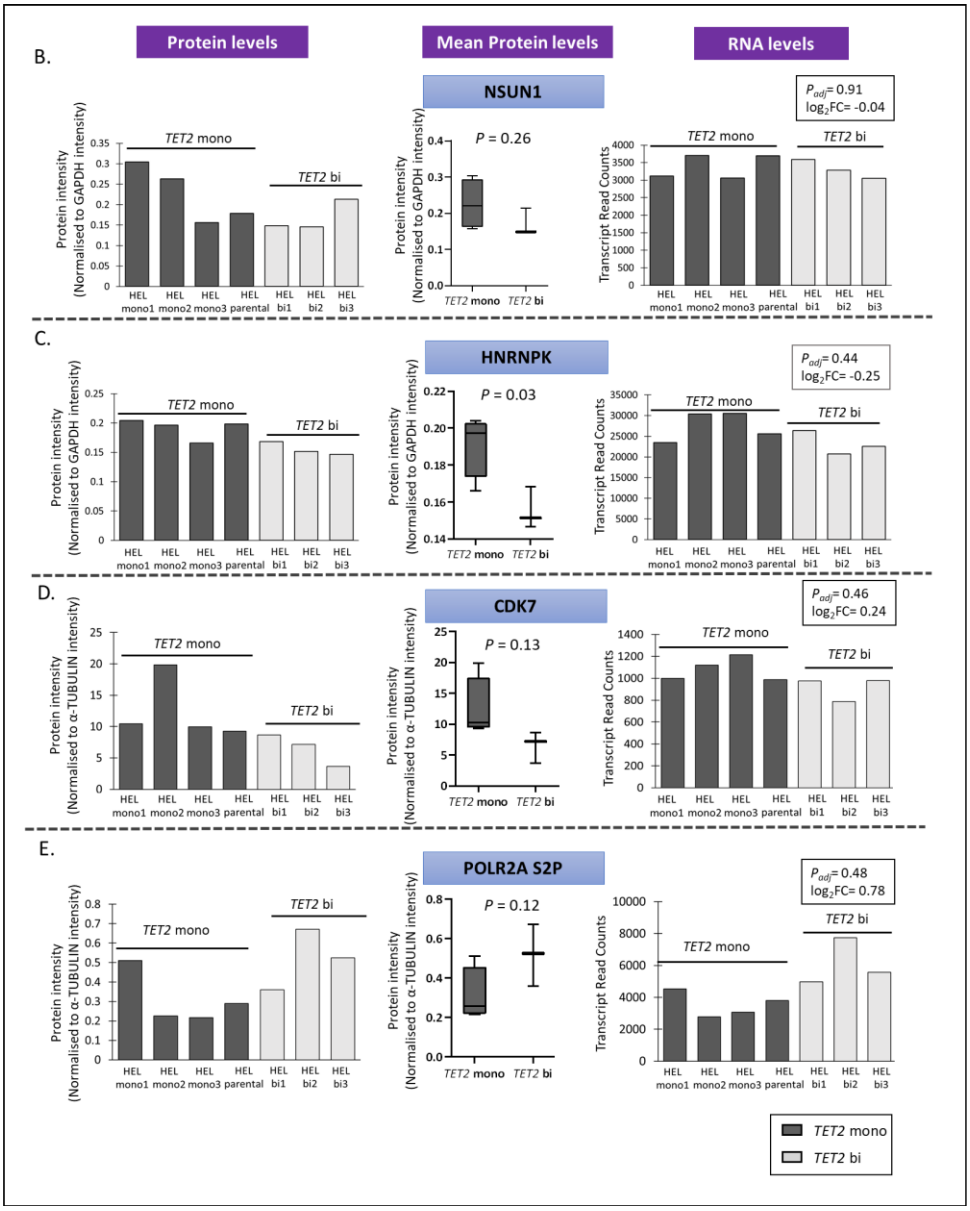

Figure 4.8. Expression of NSUN1, HNRNPK, CDK7 and POLR2A S2P (continued from previous page (A) western blot images showing the protein level expression of NSUN1, HNRNPK, CDK7 and POLR2A S2P. (B)-(E) The respective protein levels are quantified (normalised to loading control GAPDH or α-TUBULIN levels) along with mean protein levels of cells grouped based on *TET2* mutation status (*P* value, student's t-test) and respective transcript levels from RNA sequencing data. Although HNRNPK was significantly downregulated in *TET2* biallelic mutated cells, the effect size was very low and the HNRNPK levels in HEL mono3 (*TET2* monoallelic mutant) was lower than that in *TET2* biallelic mutants. In case of POLR2A S2P, the transcript level of unphosphorylated POLR2A is shown. In the box and whiskers plot, the bars represent the mean, boxes represent the interquartile range, and the whiskers represent the range between minimum and maximum values. Dark grey denotes *TET2* monoallelic mutants and pale grey denotes *TET2* biallelic mutants.

CDK7, Cyclin-dependent kinase 7; HNRNPK, Heterogeneous nuclear ribonucleoprotein K; kDa, kilodalton; NSUN1, NOP2/Sun Domain Family, Member 1; P_{adj}, adjusted *P* value; POLR2A S2P, RNA Polymerase II Subunit A Serine2 phosphorylated

4.3.4. Heterogeneity of TET2 monoallelic mutant cell clones

Given that HELmono1 cells clustered independent of the other *TET2* monoallelic cell clones, the azacytidine sensitivity of these cells was re-analysed. HELmono1 cells had a cloning efficiency lower than other *TET2* monoallelic mutant HEL cells in the presence of azacytidine (P=0.01) (Figure 4.9A-B), despite having comparable spontaneous cloning efficiencies for all cells irrespective of the *TET2* mutation status (Figure 3.7C-D). In addition, HELmono1 cells had a lower proliferation rate in presence of azacytidine compared to other monoallelic *TET2* mutant HEL cells in liquid culture (Figure 4.9C). In general, HELmono1 cells are more sensitive to azacytidine compared to other monoallelic *TET2* mutant HEL cells but are not as sensitive as cells with biallelic *TET2* mutation, which show the highest sensitivity to azacytidine.

The protein expression of other major epigenetic regulators such as IDH1, IDH2 and WT1 was investigated (Figure 4.10A). IDH1 was downregulated at transcript and protein levels in HELmono1 cells compared to other *TET2* monoallelic mutated cells (Figure 4.10B-C). In contrast, IDH2 and WT1 were not differentially expressed at protein level in HELmono1 cells compared to other *TET2* monoallelic mutated cells (Figure 4.10A). Given the role of IDH1 in generating α-KG, which is a co-factor essential for TET2 functionality (Wang, Xiao, et al. 2015), it is mechanistically plausible that HELmono1 cells have reduced TET2 functionality compared to other *TET2* monoallelic mutated cells potentially explaining why this clone clusters independently.

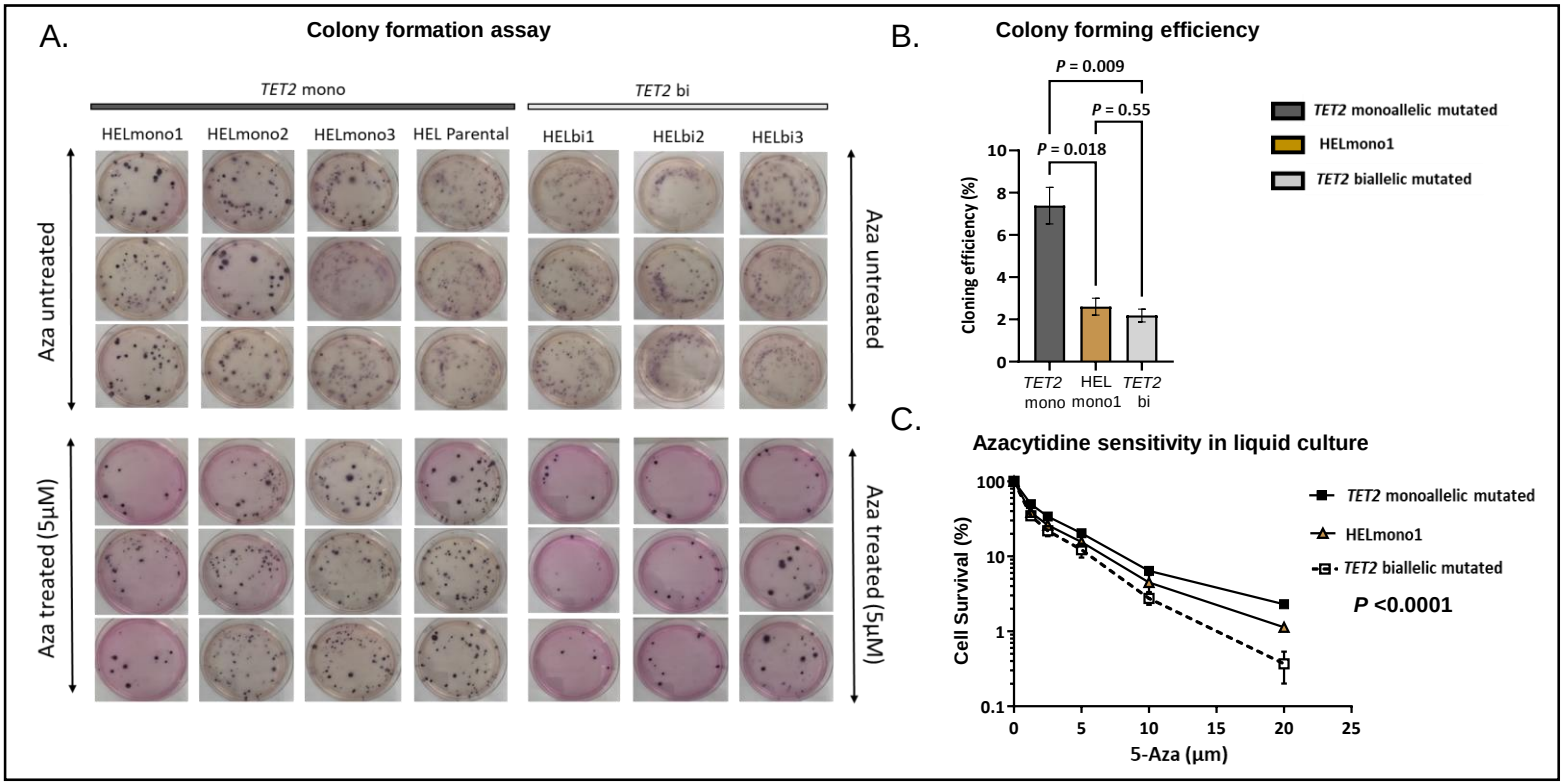

Figure 4.9. Heterogeneity of HEL TET2 monoallelic cell clones in terms of azacytidine sensitivity. (A) MTT staining of HEL AML cell colonies shows the increased sensitivity of biallelic TET2 mutant HEL cells and HELmono1 cells compared to other TET2 monoallelic mutated HEL cells. (B) Cloning efficiency of HELmono1 cells plotted separately from the mean of other TET2 monoallelic mutant HEL cells. HELmono1 cells had an intermediate colony forming efficiency which was lower than other TET2 monoallelic mutant HEL cells, in response to azacytidine. The colony formation assay data is based on three independent replicates of each cell line at each condition (azacytidine treated or untreated). P values were calculated using student's t-test. (C) Azacytidine sensitivity of HELmono1 cells plotted separately from the mean of other TET2 monoallelic mutant HEL cells. This shows that HELmono1 cells were relatively more sensitive to azacytidine compared to other TET2 monoallelic mutant HEL cells. P value (ANOVA) is based on three independent replicates. Pale grey denotes TET2 biallelic mutants, gold denotes HELmono1 cells and dark grey denotes other TET2 monoallelic mutants.
MTT, 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide

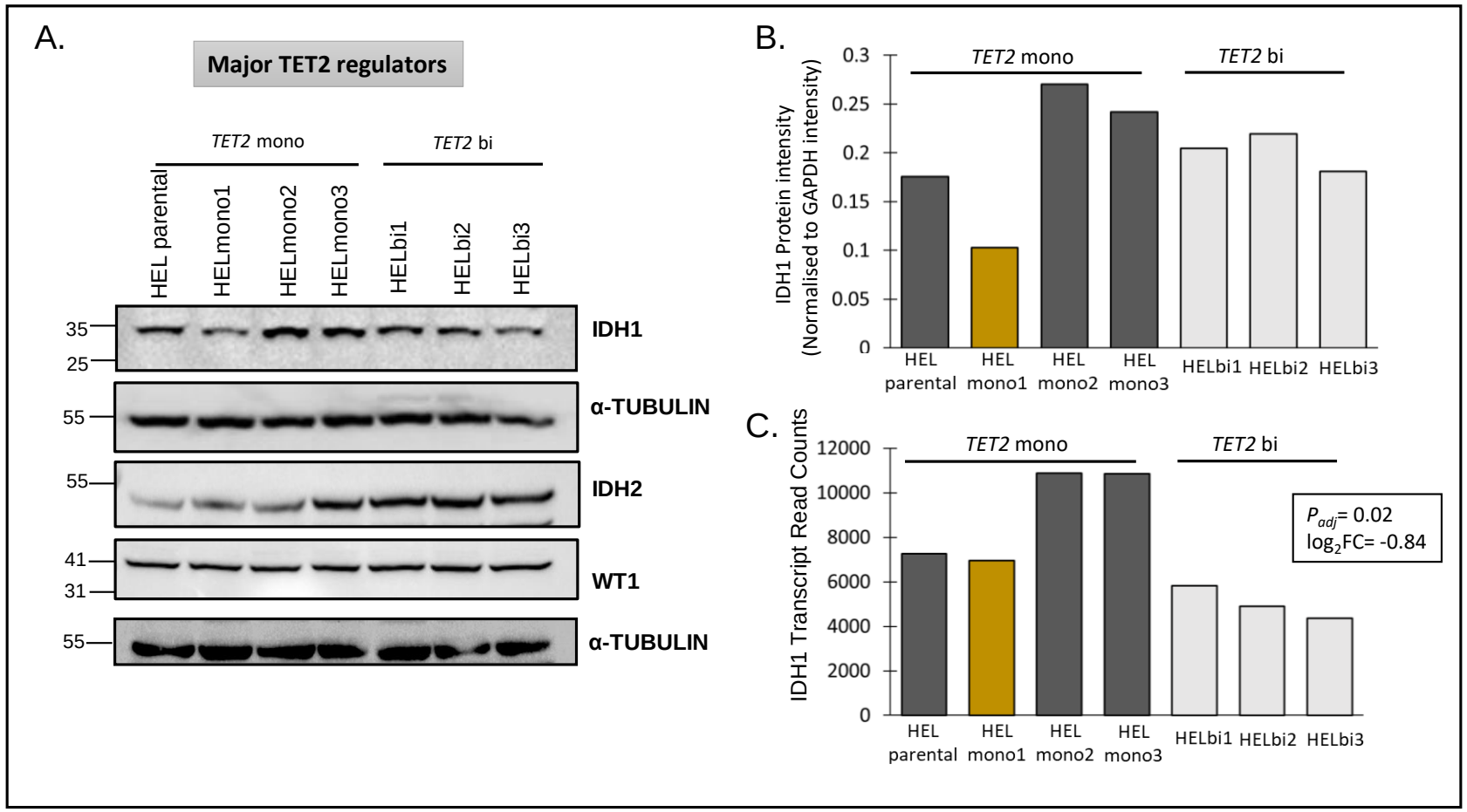

Figure 4.10. IDH1 is downregulated in HELmono1 cells. A. Protein level expression of major regulators of TET2 function such as IDH1, IDH2 and WT1. B. IDH1 protein levels were quantified using ImageLab and normalised to α-TUBULIN levels. C. IDH1 transcript levels were derived from RNAseq data (section 4.3.2. Pale grey denotes *TET2* biallelic mutants, Gold denotes HELmono1 cells and dark grey denotes other *TET2* monoallelic mutants.

4.3.5. Role of ribosomal pathway in sensitising TET2 null cells to azacytidine

In order to elucidate the possible mechanisms for azacytidine sensitivity, it was necessary to identify significantly differentially expressed pathways in azacytidine sensitive cells compared to other cells. This was done through a supervised analysis in which individual cell types were grouped based on *TET2* status. Given that HELmono1 is an outlier in terms of azacytidine sensitivity and expression of key epigenetic regulators (section 4.3.4), it was excluded from the pathway analysis. Gene ontology (GO) analysis identified 105 significantly enriched pathways associated with the differentially expressed transcripts (P_{adj}<0.05; |Log2FC|≥0.3) comparing *TET2* monoallelic and biallelic mutant HEL cell clones. Moreover, four out of top five significantly enriched pathways were ribosome related (Table 4.2; Full list of pathways in Appendix D).

Pathways	P_{adj} value
Cytosolic ribosome (GO:0022626)	1.20E-23
Ribosomal subunit (GO:0044391)	1.60E-20
Cytosolic large ribosomal subunit (GO:0022625)	1.26E-16
Cytosolic part (GO: 0044445)	1.06E-15
Focal adhesion (GO: 0005925)	7.50E-15

Table 4.2. Top five pathways retrieved from enrichGO pathway analysis of differentially expressed transcripts* based on TET2 status (excluding HELmono1)
* P_{adj}*<0.05; |Log2FC|≥0.3*

The cytosolic ribosomal pathway (GO:0022626) was the significantly most enriched pathway (P_{adj} value:1.20E-23) with 69 differentially expressed transcripts. A supervised hierarchical analysis based on *TET2* mutant allele dosage illustrates major differences in cytosolic ribosomal pathway components (Figure 4.11). HELmono1 cells were included in the heatmap for reference and despite being *TET2* monoallelic mutated these clustered with biallelic *TET2* mutant HEL cells (Figure 4.11). Ribosomal protein large (RPL) components form the larger subunit of ribosome whereas Ribosomal protein small (RPS) components form the smaller ribosomal subunit. The majority of the pathway component were downregulated in cell clones with biallelic *TET2* mutation (Figure 4.11).

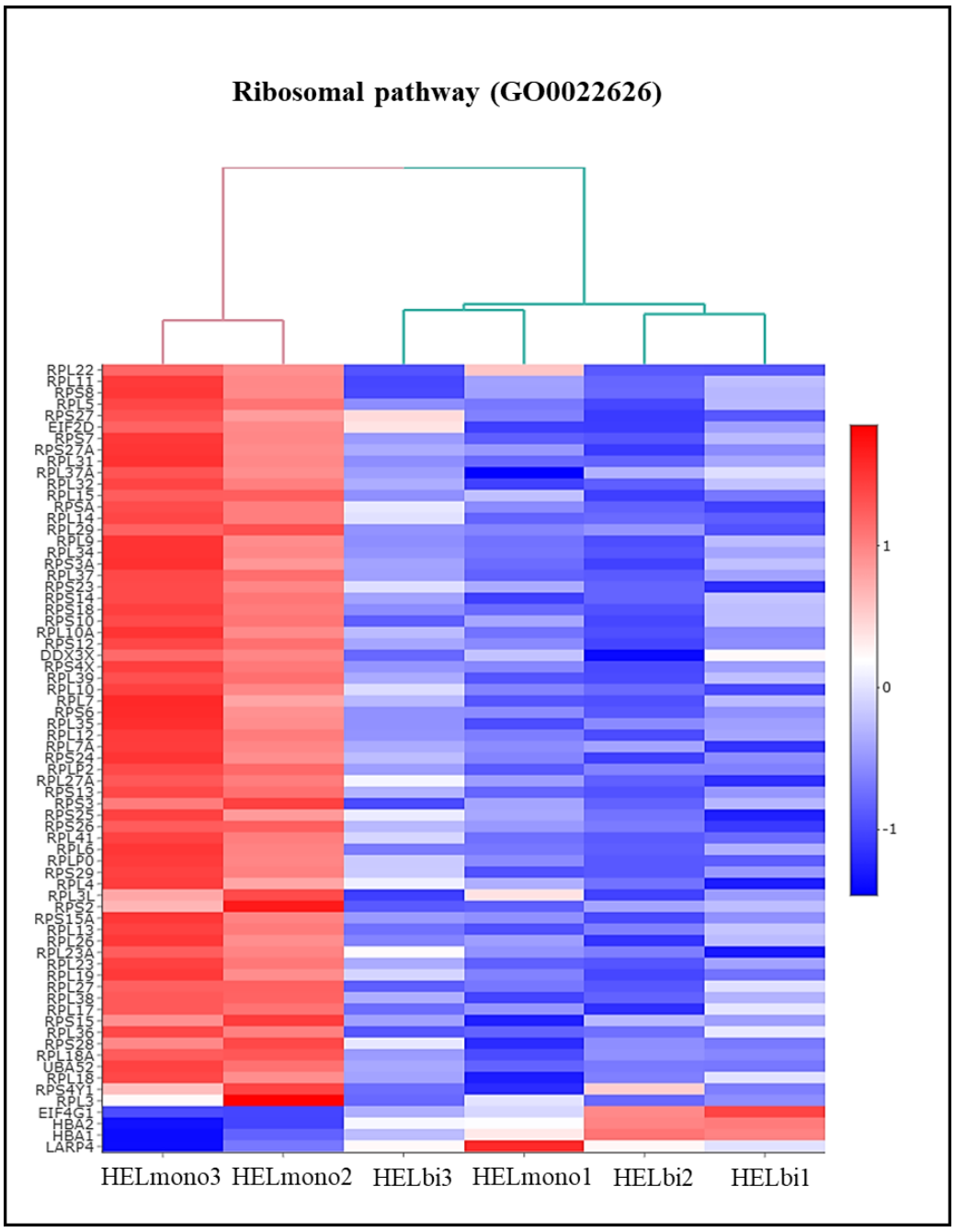

Figure 4.11. Ribosomal pathway components are downregulated in *TET2* biallelic mutant cell clones compared to *TET2* monoallelic mutant cell clones. Gene enrichment analysis identified the cytosolic ribosomal pathway (GO0022626) as the most significantly enriched with 69 differentially expressed transcripts (P_{adj}<0.05; |Log_2FC|≥0.3) in *TET2* biallelic mutants (HELbi1, HELbi2 and HELbi3) compared to *TET2* monoallelic mutants (HELmono2 and HELmono3). Heatmap shows the variance of individual ribosomal components across various cell clones. The vertical columns denote each cell clone, the horizontal rows denote individual transcripts and colour indicates relative gene expression. HELmono1 cells were excluded from the gene enrichment analysis but were included in the heatmap for reference.

4.3.6. Protein level expression of ribosomal pathway components

Protein level expression of ribosomal components was verified using western analysis and the results were broadly consistent with the RNA sequencing results. Specifically, RPS14, RPS6, RPS19, RPL22 and RPL19 were downregulated at both transcript and protein level (Figure 4.12) Though the effect size was small, there was a consistent pattern of downregulation of ribosomal subunits in TET2 null cells compared to *TET2* monoallelic mutated cells.

4.3.7. Functional validation of ribosomal pathway using BCA assay

The ribosomal complex is formed by the assembly of large and small ribosomal protein subunits which are made up of RPL and RPS protein components, respectively (Provost et al. 2013). This ribosomal machinery assembles into a "sandwich-like structure", attaches into the mRNA strand, and serves as a platform for protein synthesis (Cooper 2000; Saha et al. 2017) (Figure 4.13A).

Given the role of ribosomal pathway in protein synthesis, it was hypothesised that TET2 null cells have significantly compromised protein synthesis capacity and have lower total protein levels as a consequence. Therefore, BCA assay which measures the total protein concentration in a sample (He 2011) was exploited to test this. Consistent with this hypothesis, *TET2* biallelic mutated cells (n=2) had significantly lower total protein concentration (*P* value=0.001, paired *t*-test) compared to *TET2* monoallelic mutated cells (n=2), determined across five independent replicates (Figure 4.13B). As such, it is mechanically plausible that the downregulation of ribosomal pathway components observed in the aza-sensitive *TET2* biallelic mutated cells compromises protein synthesis giving rise to lower total protein levels.

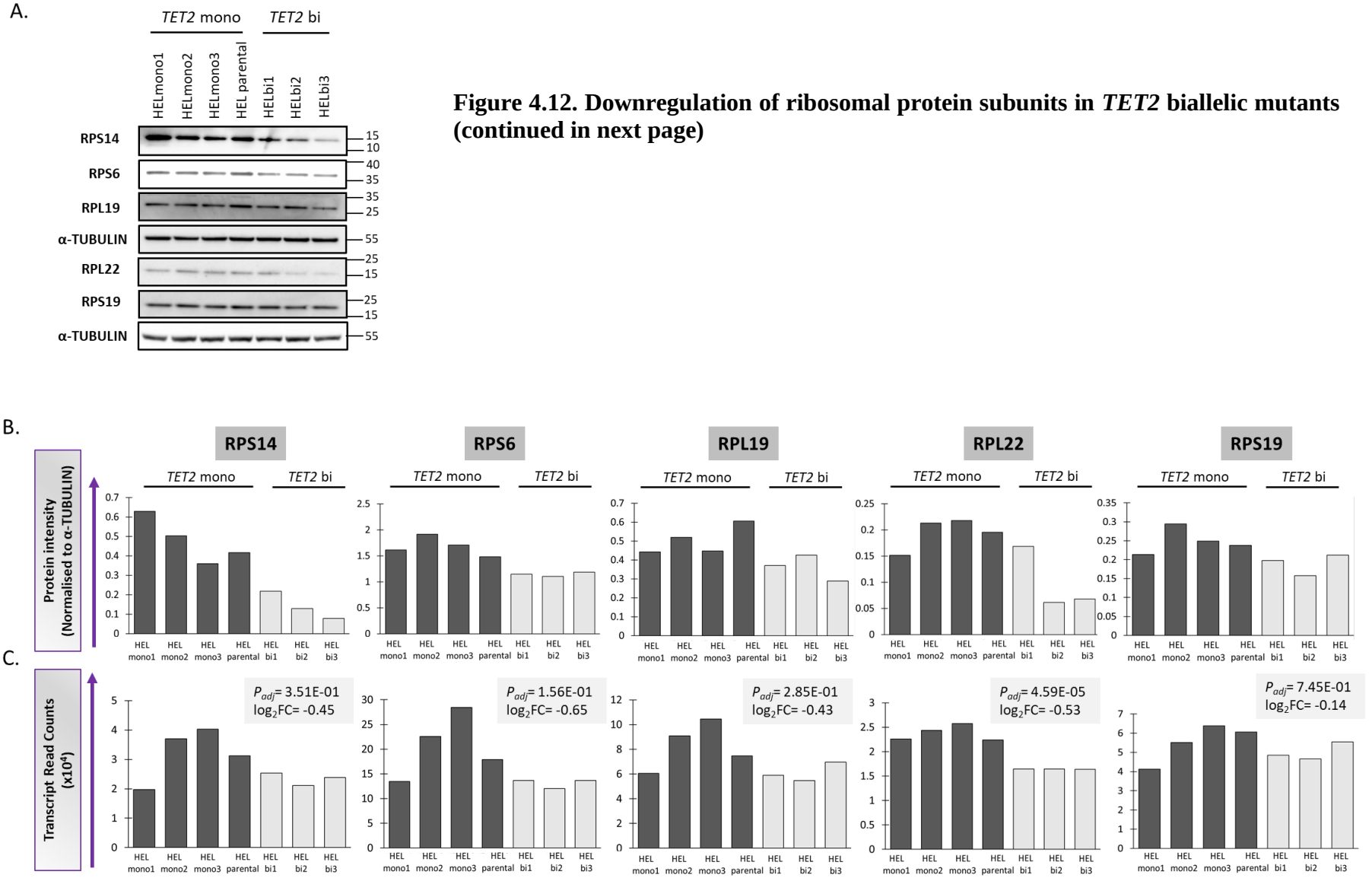

Figure 4.12. Downregulation of ribosomal protein subunits in *TET2* biallelic mutants (continued in next page)

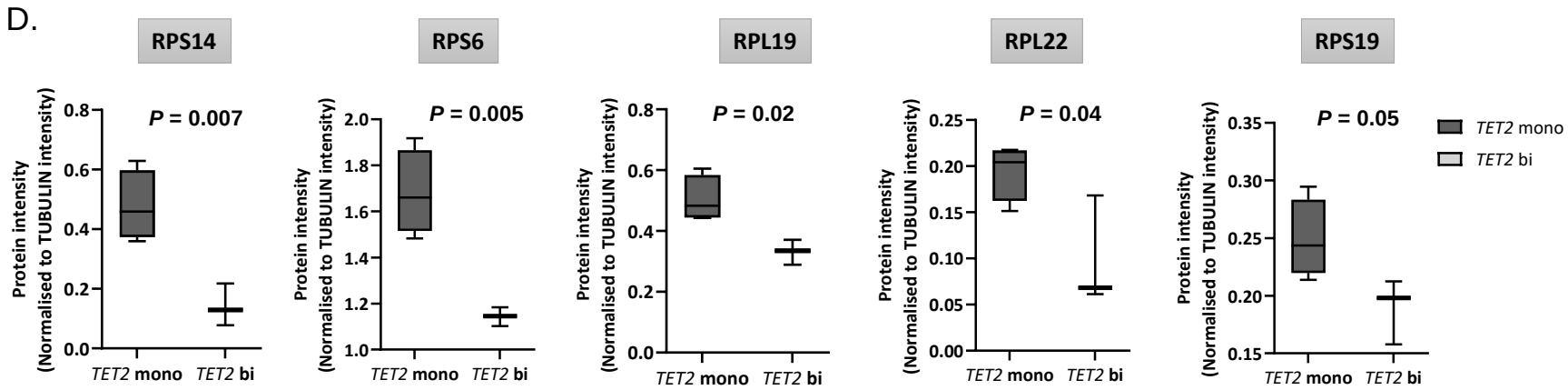

Figure 4.12. Downregulation of ribosomal protein subunits in *TET2* biallelic mutants (continued from previous page) (A) Western blot showing downregulation of ribosomal pathway components such as of RPS14, RPS6, RPL19, RPL22 and RPS19 respectively in *TET2* biallelic mutants compared to *TET2* monoallelic mutants. (B) Protein levels of respective ribosomal subunits were quantified normalised to α-TUBULIN using ImageLab. (C) Transcript level variation of the respective ribosomal subunits show a similar downregulation trend in *TET2* biallelic mutants compared to *TET2* monoallelic mutants. (D) Mean protein levels of respective ribosomal pathway component in cells when grouped based on *TET2* mutation status. The bars in the plot represent the mean, the boxes represent the interquartile range, and the whiskers represent the range between minimum and maximum values. All the ribosomal subunits tested were significantly downregulated ($P < 0.05$, student's *t*-test) in *TET2* biallelic mutated cells compared to monoallelic mutated cells, except RPS19. The weak significance in RPS19 protein levels could be driven by the low RPS19 protein expression in HELmono1 cells. Dark grey denotes *TET2* monoallelic mutants and light grey denotes *TET2* biallelic mutants.

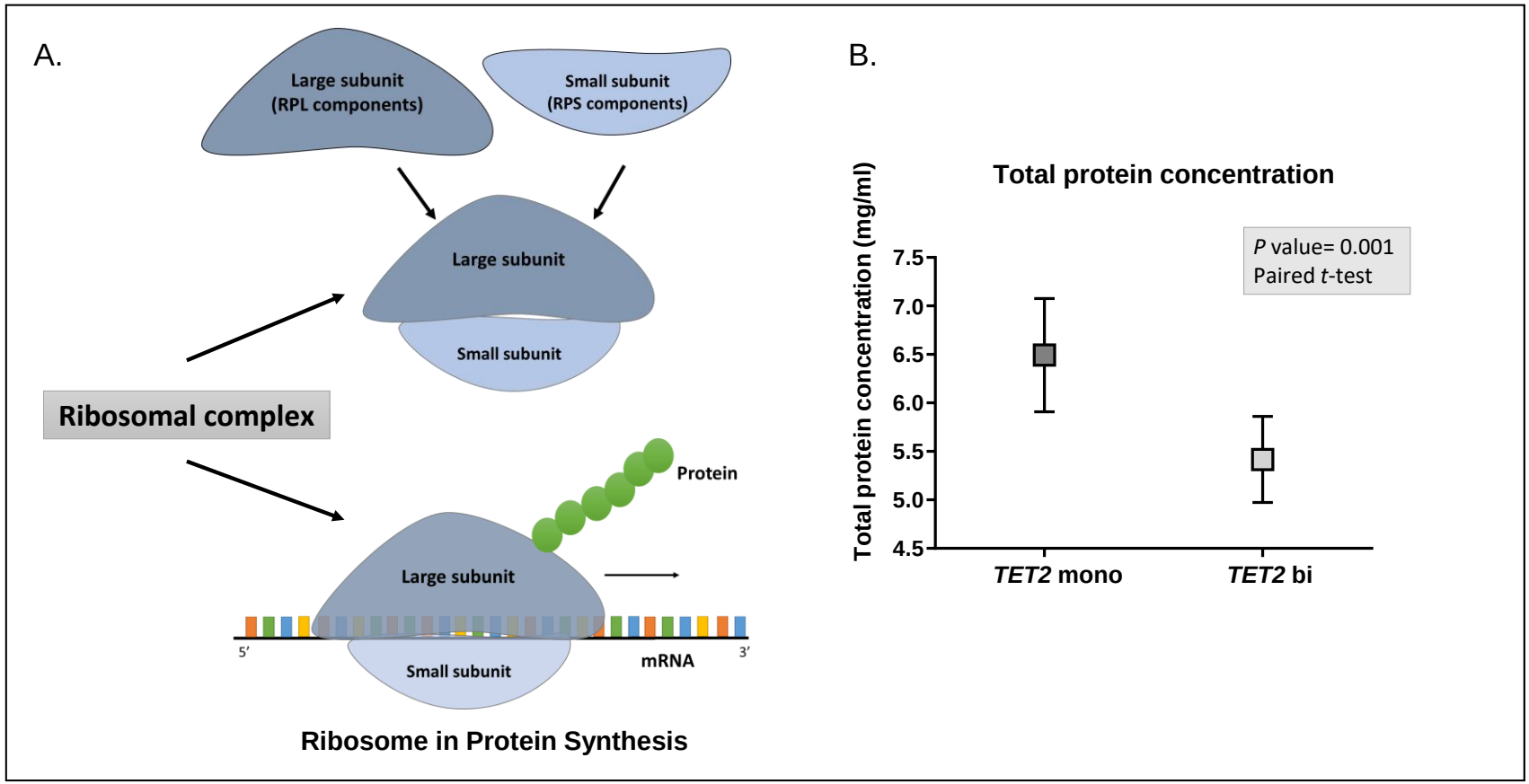

Figure 4.13. Functional validation of ribosomal pathway downregulation in *TET2* biallelic mutants. A. Schematic representation showing the assembly of large and small ribosomal subunits to form the ribosomal complex which functions as a platform for protein synthesis. B. The total protein concentration (as measured with BCA assay) was significantly lower in *TET2* biallelic mutants cell clones (n=2) compared to *TET2* monoallelic mutants (n=2) (P value=0.001, Paired t-test). The total protein concentrations shown here are derived from the mean protein concentration of cell clones grouped based on *TET2* mutation status and are derived from five independent replicates and each independent replicate consisted of three technical replicates of each cell clone. Dark grey denotes *TET2* monoallelic mutants and light grey denotes *TET2* biallelic mutants.

RPL, Ribosomal protein large subunit; RPS, Ribosomal protein small subunit

4.4. Discussion

Considering the basic function of TET2 as a driver of genomic demethylation, it is perhaps not surprising that complete loss of TET2 function via biallelic gene mutation would result in the acquisition of a global genomic hypermethylated phenotype. Illumina 450K methylation array revealed a significantly higher number of hypermethylated probes in *TET2* null cells, which is likely due to the absence of TET2-catalysed demethylation in these cells, compared to *TET2* monoallelic mutated cells. Consistent with our findings in HEL AML cells, RNAi mediated *TET2* inhibition in leukaemia cell lines and cord blood CD34+ cells resulted in significantly reduced 5-hmC levels (Pronier et al. 2011). In addition, *TET2* mutated patients with MPN have been correlated with reduction in 5-hmC levels when compared to healthy subjects (Pronier et al. 2011; Ko et al. 2010).Moreover, the 5-hmC levels in homozygous *TET2* mutant mice was dramatically lower than in heterozygous *TET2* mutant mice (Feng et al. 2019), suggesting the role of *TET2* mutant allele dosage in regulating 5-hmC levels.

Although we observed a general trend towards hypermethylation, it must be noted that the effect was relatively modest due to the presence of a large number of hypomethylated probes in TET2 null cells. Consistently, there were a large number of upregulated transcripts in biallelic *TET2* mutant cells based on RNAseq analysis. Therefore, it seems unlikely that the global hypermethylation phenotype has a profound role in conferring sensitivity to a hypomethylating agent such as azacytidine. Consistent with our results, baseline methylation levels did not correlate with better response to decitabine in patients with MDS (Shen et al. 2010). As such, accumulating evidence supports the role for specific genes or pathways in conferring sensitivity to azacytidine.

Gene ontology analysis based on biological processes identified the spliceosomal pathway as significantly enriched when comparing differentially expressed transcripts based on *TET2* mutation status. These data potentially implicate this pathway as a determinant of cellular response to azacytidine. Cheng et al. 2018 also reported data implicating this pathway as potentially being important as a determinant of azacytidine sensitivity. Specifically, they reported a unique interaction of HNRNPK with certain RNA-methyltransferases (RNMTs) and suggested that this interaction is involved in modulating response to azacytidine. Precisely, HNRNPK directly binds RNMTs such as NSUN3 and DNMT2, and CDK9 and recruits RNA Polymerase II to form an azacytidine sensitive active chromatin structure, whereas another RNA methyltransferase NSUN1 interacts with BRD4 and CDK7 and recruit RNA Pol II to form active chromatin structure that is resistant to azacytidine (Cheng et al. 2018). Irrespective

of the interaction, the individual proteins levels of HNRNPK, NSUN1, POLR2A S2P and CDK7 were upregulated in azacytidine resistant leukaemia cells compared to their azacytidine sensitive counterparts (Cheng et al. 2018). However, none of these individual components were significantly differentially expressed at the transcript or protein level in our data. Although the exact reason for this inconsistency is unclear, a possible reason could be that the differential expression of individual protein components was reported in azacytidine resistant cells that were generated from AML cell lines (OCI-M2 and SC) that are otherwise sensitive to azacytidine (Cheng et al. 2018) and in general azacytidine resistant cells are developed by long-term exposure of cells to azacytidine. It could be that the mechanism for treatment induced azacytidine resistance is different to baseline azacytidine resistance in pre-treated AML. However, further investigation is required to confirm this model.

Unsupervised clustering of RNAseq data grouped HELmono1 cells independently but closer to *TET2* biallelic mutated cells rather than to other TET2 monoallelic mutated cells. Interestingly, re-analysis of azacytidine sensitivity data showed that these cells were clearly more sensitive to azacytidine than other *TET2* monoallelic mutant cells. IDH1 was downregulated at the transcript and protein levels in HELmono1 cells compared to other *TET2* monoallelic mutated cells. IDH1 generates α-KG and is a co-factor required for TET2 functionality providing a potential explanation for the azacytidine sensitivity in the HELmono1 cell clone. Consistent with this model, downregulation of IDH1 was correlated with a reduction in α-KG levels in MiaPaCa2 (pancreatic cancer) cells (Zarei et al. 2017) and TET protein function requires α-KG to mediate genomic demethylation (Tahiliani et al. 2009; Figueroa et al. 2010). Furthermore, *IDH1* mutant primary AML displays genomic hypermethylation compared to *IDH1* wild-type AML and normal bone marrow samples (Figueroa et al. 2010). As such, it can be speculated that *IDH1* mutant AML could also be sensitive to azacytidine. Consistently, *IDH1* mutations in MDS was significantly associated with a higher overall response rate when treated with hypomethylating agents (Traina et al. 2014) although other studies have failed to replicate this finding (DiNardo et al. 2014). Therefore, the predictive value of *IDH1* mutations for response to azacytidine remains to be determined.

Role of ribosomal pathway in sensitising AML cells to azacytidine

Gene enrichment analysis based on differentially expressed transcripts in cell clones stratified by *TET2* status identified the ribosomal pathway as the most significantly downregulated pathway in *TET2* biallelic mutated cells. Eukaryotic ribosomes are also known as 80S ribosomes as they sediment at 80S (S denotes Svedberg unit for sedimentation coefficient) and

are made up of two subunits, a small 40S subunit and a large 60S subunit (Kang et al. 2021). The small subunit interacts with the anti-codon containing site of the complementary tRNA and functions as a 'decoding site' by translating the codon information of an mRNA into the respective amino acids (Kang et al. 2021). Whereas the large subunit plays a role in forming polypeptide chains by linking the amino acids with peptide bonds (Kang et al. 2021; Moore et al. 2003). The ribosome is an example of a ribonucleoprotein which is a combination of RNA and protein components. The small ribosomal subunit consists of one 18S rRNA and 33 ribosomal proteins (RPS) whereas the large subunit consists of three rRNAs (28S, 5.8S, and 5S) and 48 ribosomal proteins (RPL) (Kang et al. 2021). The various rRNA and protein components are generated and assembled into the final catalytic 40S and 60S subunits through a cascade of processes collectively known as ribosome biogenesis, tightly regulated by RNA polymerases within the nucleolus. Finally, the mature 40S and 60S subunits are exported into the cytoplasm, where the functional 80S ribosome is formed (Kang et al. 2021).

In our data, out of 81 ribosomal protein components (33 RPS and 48 RPL), 69 were downregulated in *TET2* biallelic mutated cells. In addition, ribosomal pathway components were also downregulated in both HELmono1 cells which were also concomitantly sensitive to azacytidine. Although the magnitude of ribosomal protein subunit downregulation was relatively modest at the protein level, it must be noted that the protein level validation was performed on only 5 of 69 differentially expressed ribosomal pathway components. Therefore, it will be crucial to investigate other pathway components at the protein level and also to interrogate the functional impact on azacytidine sensitivity in an isogenic setting.

Ribosomal proteins are essential for the functioning of ribosome machinery as they play an important role in rRNA processing, stabilization of rRNA secondary structures, stabilization of the subunits, pre-ribosome transport and interaction with auxiliary factors that assists ribosome assembly and mRNA translation (Kang et al. 2021). RNAi mediated knockdown of RPS components negatively affected the pre-rRNA processing and significantly reduced the 40S subunit production (O'Donohue et al. 2010).

Given the essential role of ribosomes in protein synthesis, the BCA assay was used to quantify the total protein in *TET2* biallelic mutated cells. Consistent with downregulation of ribosomal pathway components and the hypothesised attenuated of the protein machinery, *TET2* biallelic mutated cells had significantly lower total protein concentration compared to *TET2* monoallelic mutated cells. As such, ribosomal proteins are essential in shaping the catalytically active ribosomal machinery and therefore it is plausible that the global downregulation pattern of

ribosomal pathway observed in *TET2* biallelic mutated cells have contributed to the decrease in total protein concentration in these cells, due to a reduction in mRNA translation. Consistently, rapamycin treated MCF-7 cells with markedly reduced rRNA and ribosomal protein levels had a global reduction in protein synthesis, as indicated by the polysome profiling (Donati et al. 2011).

Mutations in genes encoding ribosomal proteins including *RPL5, RPL10, RPL11*, *RPL15* and *RPL22* are implicated in various cancers including haematological malignancies (De Keersmaecker et al. 2013; Yu, Kim, et al. 2017; Rao et al. 2012). Ribosomal proteins play a role in maintaining cellular homeostasis by regulating DNA repair, cell growth and proliferation, differentiation and apoptosis (Wang, Nag, et al. 2015). As such, it is plausible that loss of ribosomal pathway component expression could impact on azacytidine sensitivity through one or more of these functions. Upregulation of ribosomal genes were previously implicated in poor response to azacytidine in MDS, CMML and AML (Monika Belickova et al. 2016). Specifically, *RPL28, RPL31* and *RPL32*, which are part of the larger subunit of ribosome, were upregulated at RNA level in non-responders to azacytidine compared to responders (Monika Belickova et al. 2016).

To summarise, our data demonstrates that there is significant downregulation of ribosomal pathway in *TET2* biallelic mutated cells compared to *TET2* monoallelic mutated cells and the downregulation of this pathway has been previously implicated in sensitising AML to azacytidine (Monika Belickova et al. 2016)

4.4.1. Summary of Chapter

- *Methylome/Transcriptome profile of TET2 biallelic mutated cells*

(i) *TET2* biallelic mutation results in the acquisition of a global hypermethylated phenotype, with 7960 hypermethylated and 4584 hypomethylated CpG probes (P_{adj}<0.05; |Log_2FC|≥2).

(ii) RNAseq transcriptome profiling identified 369 significantly upregulated transcripts and 329 significantly downregulated transcripts (P_{adj}<0.05; |Log_2FC |≥0.3).

(iii) GO analysis on differentially expressed transcripts stratified by *TET2* status revealed the spliceosomal pathway as the most significantly enriched pathway (downregulated) in *TET2* biallelic mutated cells.

- *Heterogeneity of TET2 monoallelic mutated cell clones*

(i) Based on the transcript profile, HELmono1 cells clustered independently from other *TET2* monoallelic mutated cell clones and closer to *TET2* biallelic mutated cells.

(ii) HELmono1 cells have intermediate sensitivity to azacytidine.

(iii) IDH1 (a major TET2 regulator) was downregulated in HELmono1 cells, which might have attributed to reduced functionality of TET2 in these cells.

- *Role of ribosomal pathway in azacytidine sensitivity*

(i) GO analysis (excluding HELmono1) identified the cytosolic ribosomal pathway as the top differentially expressed pathway in *TET2* biallelic mutated cells compared to *TET2* monoallelic mutated cells.

(ii) Ribosomal pathway (which has a role in azacytidine sensitivity) is broadly downregulated in TET2 null cells.

(iii) *TET2* biallelic mutated cells had significantly lower total protein concentration compared to *TET2* monoallelic mutated cells, with downregulation of the ribosomal pathway as a possible cause.

Downregulation of ribosomal pathway could be an important mechanism responsible for the azacytidine sensitivity of *TET2* biallelic mutated AML cells.

Chapter 5. Role of ABCB1 in sensitising *TET2* biallelic mutant cells to azacytidine

5.1. Introduction

ATP Binding Cassette Subfamily B Member 1 (ABCB1)

ATP-binding cassette (ABC) transporters are a class of ubiquitously expressed integral membrane proteins that are responsible for the ATP-powered primary transport of substrates across the cellular membrane (Rees et al. 2009). Despite the presence of both export and import ABC pumps in bacteria, eukaryotes only express ABC-mediated export pumps (Wilkens 2015). The ABC superfamily in humans consists of 48 members of which *ABCB1* is the best characterised. The *ABCB1* gene, also known as *multi drug resistance 1* (*MDR1*), is located at 7q21.1 and encodes for P-glycoprotein (Zawadzka et al. 2020). Structurally, ABC transporters are characterised by the presence of four domains, consisting of two nucleotide binding domains or ATP binding cassettes (ABC) located in the cytoplasm and two transmembrane domains (TMD) located in the membrane bilayer (Figure 5.1). Each TMD consists of 6 α-helices which makes up a total of 12 α-helices for each complete ABC transporter. These α-helices form of a transmembrane pore which opens either into the cytoplasm (inward-facing) or outside of the cell (outward-facing) (Rees et al. 2009; Wilkens 2015).

The catalytic cycle is initiated by the binding of transport substrate at the TMD domains and two ATP molecules at the ABC domains resulting in ABC domain dimerization. Subsequent hydrolysis of one of the two ATP molecules into ADP and phosphate group results in a conformational change and outward opening of the TMD gate. This allows the efflux of the substrate and release of one ADP molecule. Hydrolysis of the second ATP molecule releases ADP molecule from the ABC domain leading to the dimer dissociation. This allows a conformational reset and inward closing of the TMD gate, where substrate and ATP can bind and re-initiate the cycle (Figure 5.1). Eventually there is a conformational switching of TMD domains between inward and outward-facing conformations and this facilitates the unidirectional transport of substrates across the membrane (Wilkens 2015; Rees et al. 2009). ABCB1 is structurally similar to other ABC transporters and is involved in the unidirectional efflux of substrates across the cellular membrane (Choudhuri et al. 2006; Wolf et al. 2011). The ABC region is highly conserved across the various ABC family members, whereas the TMD region is largely variable which permits the binding and translocation of a wide range of substrates including various chemotherapeutic drugs (Rees et al. 2009).

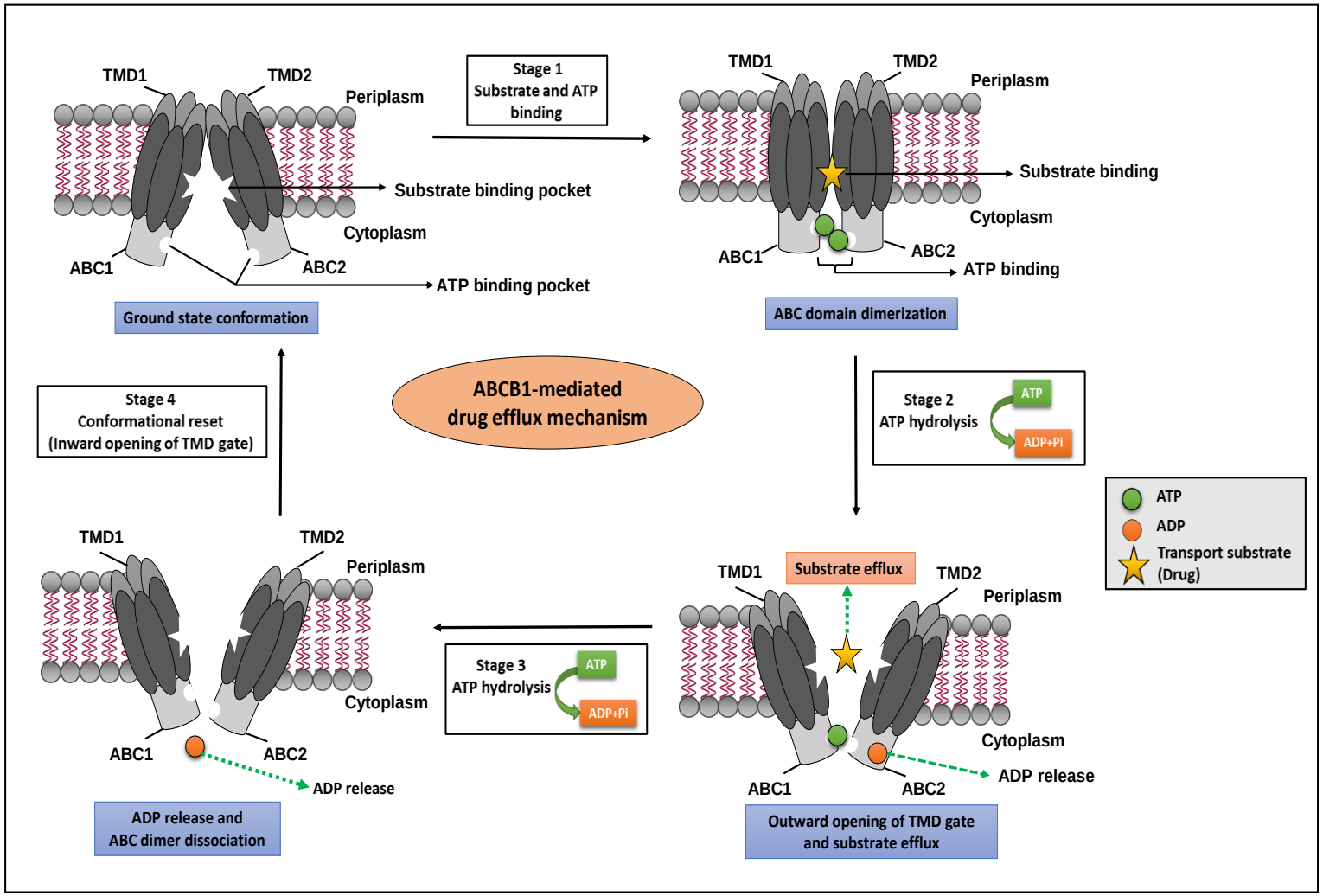

Figure 5.1. Schematic representation of the catalytic cycle of ABCB1-mediated drug efflux. ABCB1 is made of two ABC domains (ABC1 and ABC2) and two TMD domains (TMD1 and TMD2). The substrate binds to the substrate binding pocket and two ATP molecules binds to the ATP binding pocket. This is followed by ABC domain dimerization. In the next stage, there is ATP hydrolysis of one ATP molecule leading to conformational change of TMD gate and substrate efflux. The hydrolysis of the second ATP molecule results in conformational reset where the ATP molecules and substrate can bind and the cycle repeats.

ABC, ATP-binding cassette; ADP, adenosine diphosphate; ATP, adenosine triphosphate; Pi, inorganic phosphate; TMD, transmembrane domain

Clinical significance of ABCB1 in AML treatment response

Chemoresistance is a major reason behind treatment failure in various malignancies including AML. ABC transporter mediated increase in cellular drug efflux is one of the most common mechanisms for drug resistance (Steinbach et al. 2007; To et al. 2020).

ABCB1 confers chemoresistance to a wide range of AML therapeutics including cytarabine, daunorubicin, methotrexate, mitoxantrone, vincristine, vinblastine, and others (Steinbach et al. 2007; Boyer et al. 2019). Also, AML patients with low ABCB1 mRNA levels were correlated with a higher complete remission rate and a longer event free survival when compared to patients with high ABCB1 expression (Boyer et al. 2019). High ABCB1 expression is reported in 35%-70% of adult AML cases and is often attributed to the presence of multiple polymorphisms affecting *ABCB1* such as C1236T (rs1128503), G2677T / A (rs2032582), and C3435T (rs1045642) (Olarte Carrillo et al. 2021). Overall, several studies have demonstrated the unfavourable prognostic significance of high ABCB1 expression in AML (Damiani et al. 2006; Megías-Vericat et al. 2015) and its prognostic impact is more pronounced in elderly patients compared to younger patients, thus becoming an age-dependent prognostic factor (Boyer et al. 2019).

A comparison of ABCB1 activity in different risk categories showed a significant association between high ABCB1 expression and adverse risk groups as well as low ABCB1 expression and favourable risk groups (Boyer et al. 2019). Moreover, ABCB1 differential expression was often correlated with presence or absence of specific mutations in AML such that high ABCB1 expression was frequently associated with presence of *CEBPA* mutations (P =0.037) and and absence of *FLT3*-ITD (P =0.006) and *NPM1* mutations (P <0.001) (Boyer et al. 2019). In addition, there has been strong negative correlation between ABCB1 expression and CD33 expression and therefore, patients with low ABCB1 tends to benefit from GO treatment that targets CD33 (Walter et al. 2007). As such, it is evident that ABCB1 expression plays a critical role in determining therapeutic response in AML.

5.2. Aims of Chapter 5

Transcript profiling of *TET2* monoallelic and biallelic mutated cells revealed that ABCB1 is among the top differentially expressed transcripts and is significantly downregulated in *TET2* biallelic mutated cells compared to *TET2* monoallelic mutated cells. Given the role of ABCB1 in rendering chemoresistance to AML cells, it was important to investigate the role of transcript downregulation in sensitising *TET2* biallelic mutated cells to azacytidine. The specific aims of this chapter are as follows.

- To test whether the ABCB1 protein level expression is consistent with transcript level expression in *TET2* monoallelic and biallelic mutated cells.
- To investigate the effect of ABCB1 inhibition on modulating the response of *TET2* mutated cells to azacytidine treatment.
- To generate and establish azacytidine resistant cell clones of *TET2* monoallelic and biallelic mutated HEL cells using long-term exposure to azacytidine. Also, to test and quantify the azacytidine resistance phenotype of these cells using a growth inhibition assay.
- To test the ABCB1 protein levels in azacytidine resistant clones and to compare them with the ABCB1 protein levels of the respective parental cells from which each azacytidine resistant clone was generated.
- To investigate whether ABCB1 inhibition in azacytidine resistant cells has the potential to re-sensitise these cells to azacytidine.

5.3. Results

5.3.1. Effect of TET2 mutant allele dosage on ABCB1 RNA and protein expression

Transcript profiling and further shortlisting based on published evidence identified ABCB1 as a potential candidate having a role in sensitising *TET2* null cells to azacytidine. RNAseq analysis demonstrated significant downregulation of ABCB1 in *TET2* biallelic mutated cells compared to *TET2* monoallelic mutated cells (P_{adj}=0.016; Log_2FC=-0.59) (Figure 5.2 A). Consistently, ABCB1 was downregulated in *TET2* biallelic mutated cells at the protein level (Figure 5.2 B-C), which is predicted to lead to intracellular accumulation of azacytidine (Lainey et al. 2013), and hence increased chemosensitivity to azacytidine in these cells compared to their isogenic *TET2* monoallelic HEL cells.

5.3.2. ABCB1 functional study 1: Effect of ABCB1 inhibition on azacytidine sensitivity of TET2 monoallelic and biallelic mutated cells

Given the role of ABCB1 in drug efflux, it was hypothesized that an inhibition of ABCB1 using a selective inhibitor would block the ABCB1-mediated efflux of azacytidine leading to increased intracellular accumulation of azacytidine and increased azacytidine induced cytotoxicity. To test this hypothesis, two different inhibitors of ABCB1 (namely verapamil and tariquidar) were used independently in combination with azacytidine.

(i) Use of Verapamil to inhibit ABCB1

TET2 monoallelic (n=2) and biallelic (n=2) mutated cell clones were exposed to a combination of increasing concentrations of azacytidine and verapamil. The effect of ABCB1 inhibition, which is measured in terms of the synergy score, were clearly higher for *TET2* monoallelic mutants (Figure 5.3 A-B) compared to the *TET2* biallelic mutants (Figure 5.3 C-D) (Table 5.1) and this phenotype is directly proportional to the ABCB1 protein levels in these cells.

Of note, the synergy scores are interpreted as the excess response due to the drug combination which is beyond expectation, such that a synergy score of 10 corresponds to 10% of response beyond expectation (Ianevski et al. 2020). Generally, when the synergy score is below -10, the drug interaction is considered to be antagonistic, when score ranges from -10 to 10, the interaction is considered additive and when larger than 10, the interaction is considered synergistic. Based on these guidelines, the azacytidine-verapamil interaction was synergistic in *TET2* monoallelic mutants, whereas it was additive in *TET2* biallelic mutants.

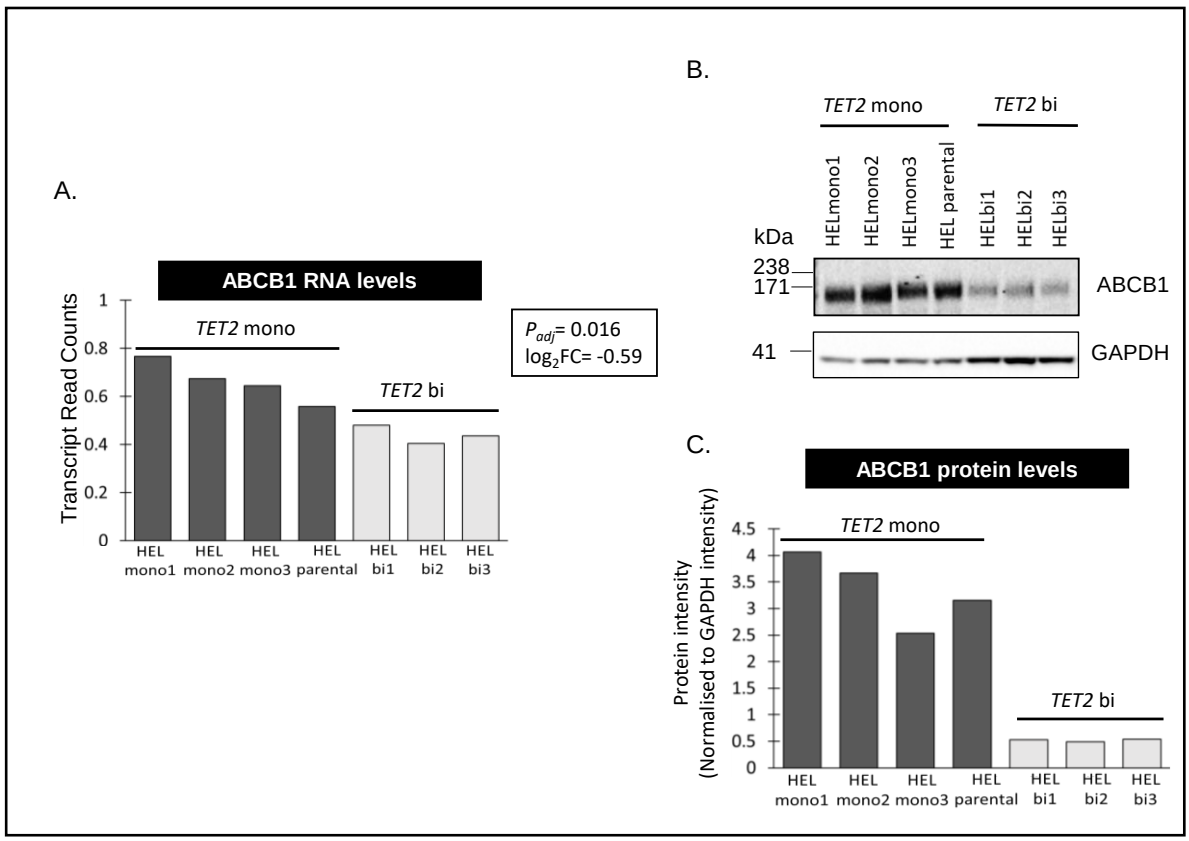

Figure 5.2. ABCB1 transcript and protein level expression. A. Transcript profile showing significant downregulation (P_{adj}=0.016; Log_2FC=-0.59) of ABCB1 in *TET2* biallelic mutated cells compared to *TET2* monoallelic mutated cells. B. Western blot showing downregulation of ABCB1 protein in *TET2* biallelic mutated cells. C. ABCB1 protein intensity is quantified and normalised to loading control (GAPDH) protein intensity for each sample.

FC, fold change

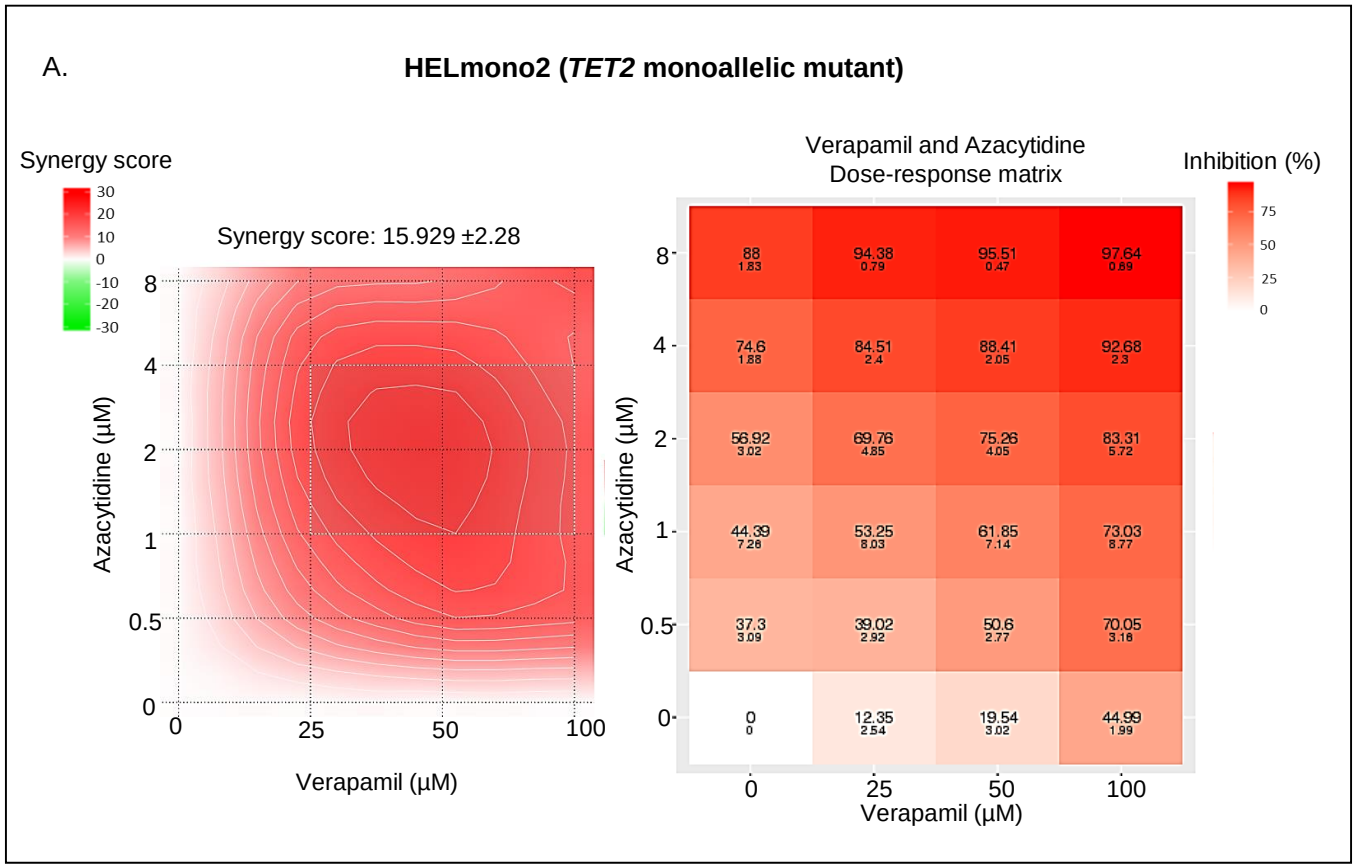

Figure 5.3. Effect of inhibition of ABCB1 using verapamil on azacytidine sensitivity (continued on next page).

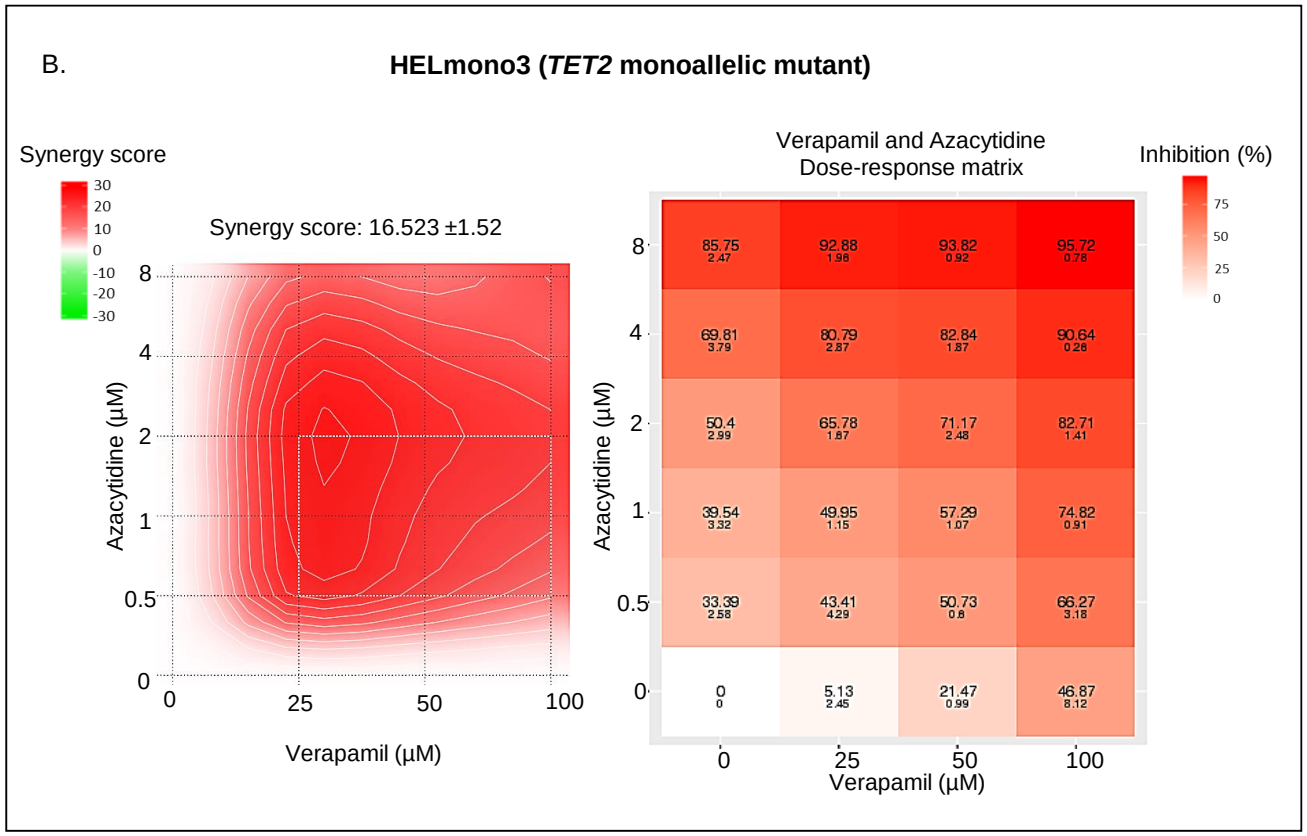

Figure 5.3. Effect of inhibition of ABCB1 using verapamil on azacytidine sensitivity (continued on next page)

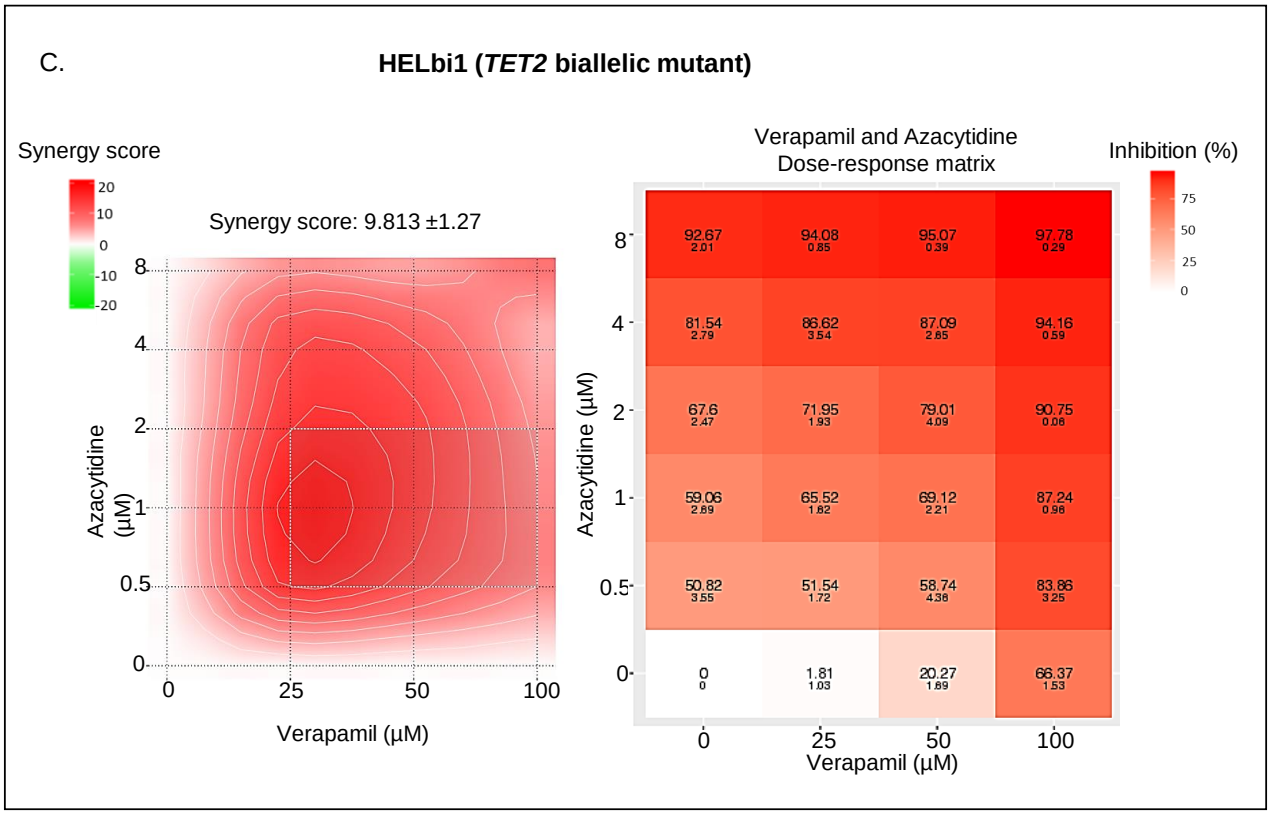

Figure 5.3. Effect of inhibition of ABCB1 using verapamil on azacytidine sensitivity (continued on next page)

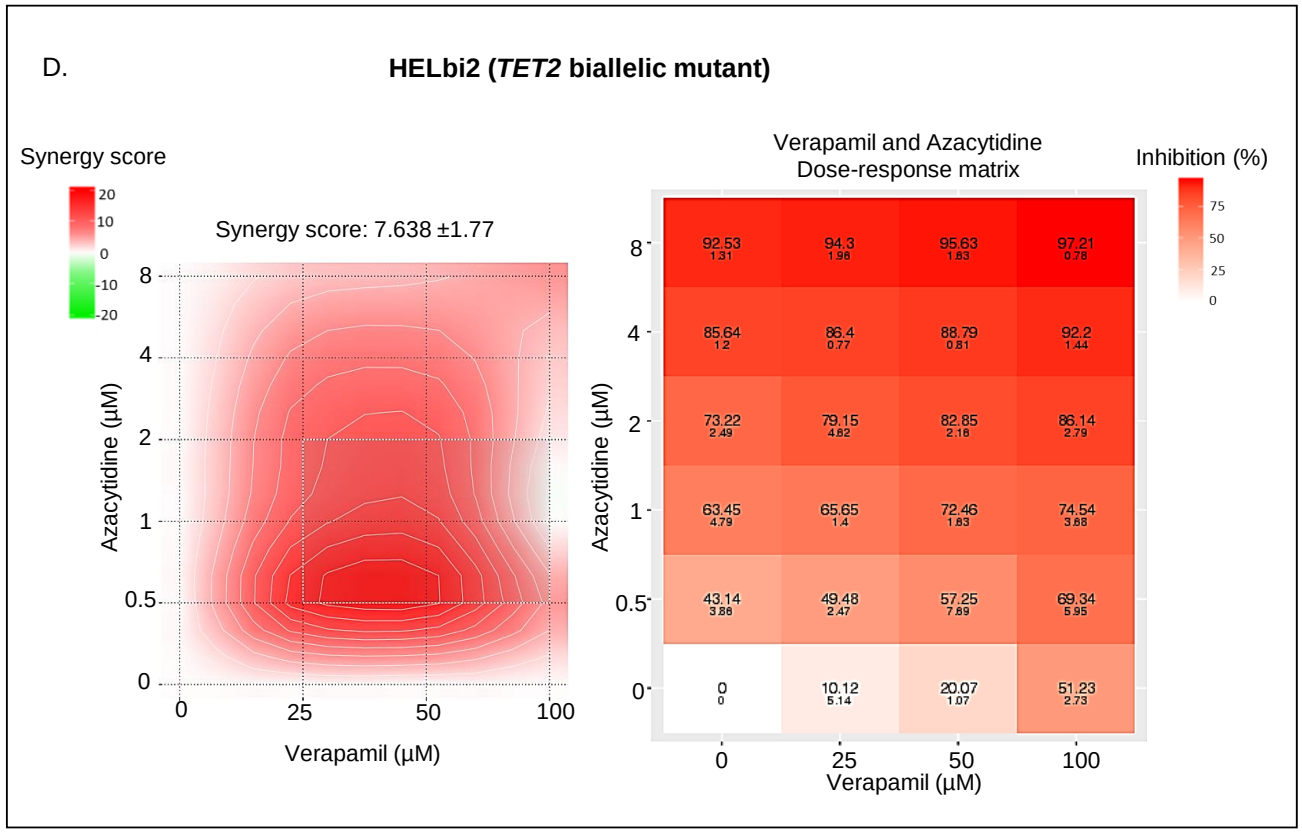

Figure 5.3. Effect of inhibition of ABCB1 using verapamil on azacytidine sensitivity (continued from previous page). The synergy map *(left)* shows the synergy score along with the most synergistic area (region shown in darker red colour) and the dose-response matrix *(right)* shows the respective cellular growth inhibition percentage, when the *TET2* monoallelic mutant cells (A) HELmono2 and (B) HELmono3 and the *TET2* biallelic mutants (C) HELbi1 and (D) HELbi2 were treated with a combination of increasing doses of verapamil and azacytidine. In general, the synergy scores and hence the effect of ABCB1 inhibition on azacytidine sensitivity was higher in *TET2* monoallelic mutants compared to *TET2* biallelic mutants. The error values of synergy scores are based on the standard deviation of mean synergy score of three independent replicates. Darker red indicates higher synergy in the synergy map *(left)* and higher growth inhibition percentage in the dose response matrix *(right).* In the dose response matrix, the numbers in larger font size *(above)* denote the growth inhibition percentage and those in smaller font size *(below)* denote the standard deviation from the mean inhibition percentage of three independent replicates.

(ii) Use of Tariquidar to inhibit ABCB1

TET2 monoallelic (n=2) and biallelic (n=2) mutated cell clones were exposed to a more selective ABCB1 inhibitor such as tariquidar in combination with azacytidine. Consistent with the previous findings, *TET2* monoallelic mutants had a higher synergy score compared to *TET2* biallelic mutants which is again directly proportional to the ABCB1 protein levels in these cells (Figure 5.4 A-D). Similar to the observed azacytidine-verapamil interaction (section 5.3.2.(i)), the azacytidine-tariquidar interaction was also synergistic in *TET2* monoallelic mutants and additive or borderline synergistic (in case of HELbi2) in *TET2* biallelic mutants. The synergy scores are summarised in Table 5.1.

Drug 1	Drug 2	Synergy scores			
		TET2 monoallelic		*TET2* biallelic	
		HELmono2	HELmono3	HELbi1	HELbi2
Azacytidine	Verapamil	15.93 ± 2.28	16.52 ± 1.52	9.81 ± 1.27	7.64 ± 1.77
	Tariquidar	18.34 ± 1.31	19.18 ± 1.16	7.24 ± 1.01	10.2 ± 1.01

Table 5.1. Synergy scores of combination treatment with azacytidine and ABCB1 inhibitor. The error values are based on standard deviation from the mean of three independent replicates.

Taken together, inhibition of ABCB1 using verapamil and tariquidar clearly sensitised the cells to azacytidine, which is denoted by the synergistic phenotype and the overall impact was significantly greater on *TET2* monoallelic mutants (higher synergy score) compared to *TET2* biallelic mutants, which is entirely consistent with the ABCB1 protein levels in these cells (Figure 5.5).

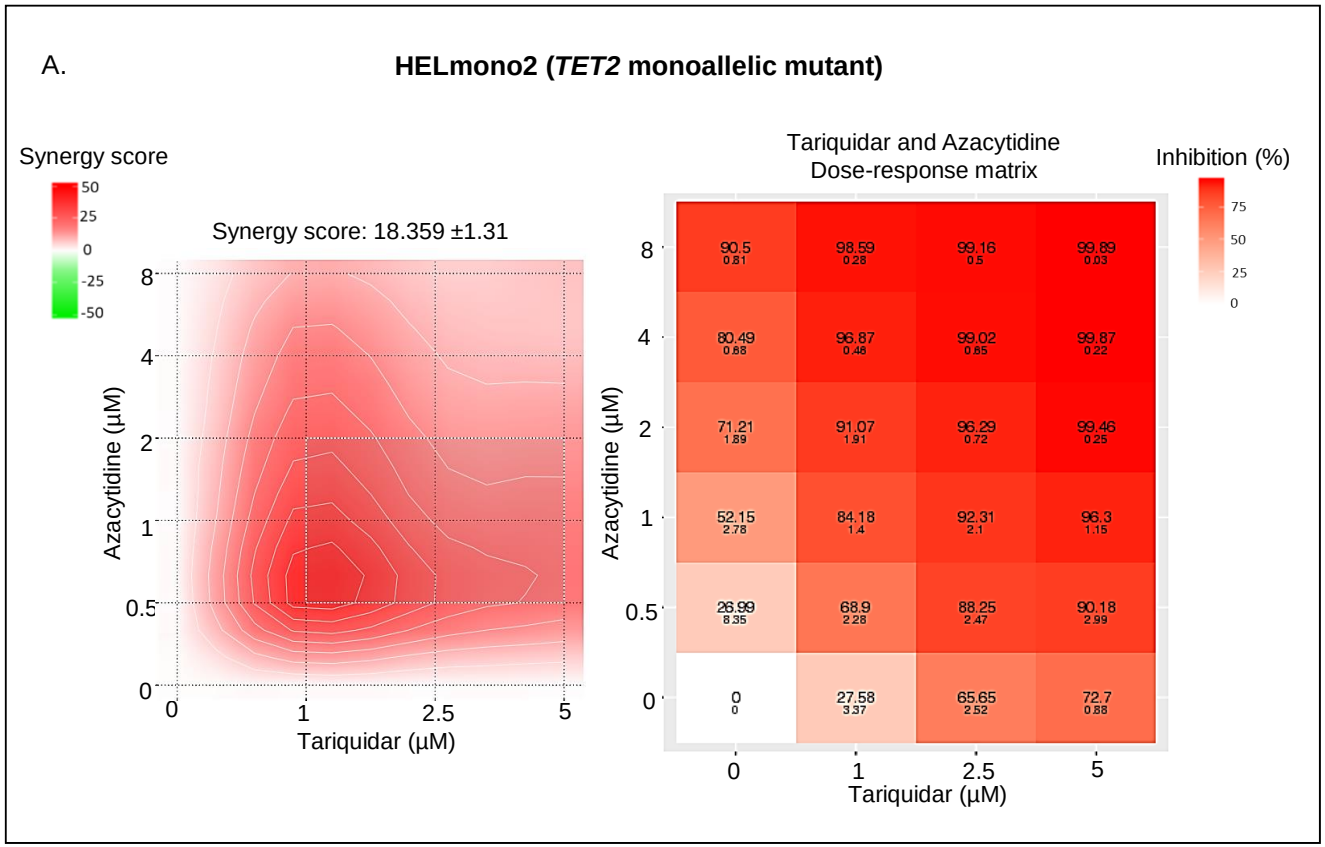

Figure 5.4. Effect of inhibition of ABCB1 using tariquidar on azacytidine sensitivity (continued on next page)

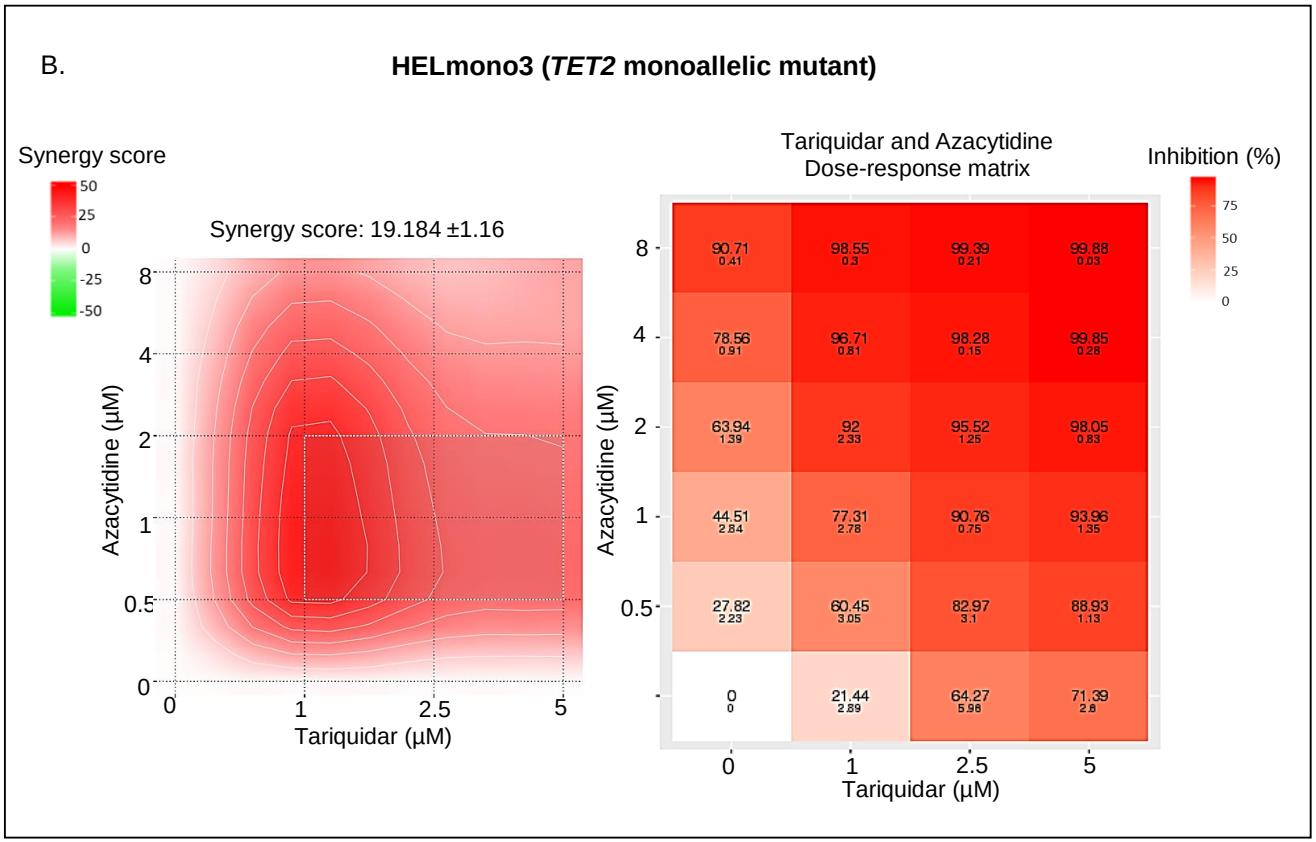

Figure 5.4. Effect of inhibition of ABCB1 using tariquidar on azacytidine sensitivity (continued on next page)

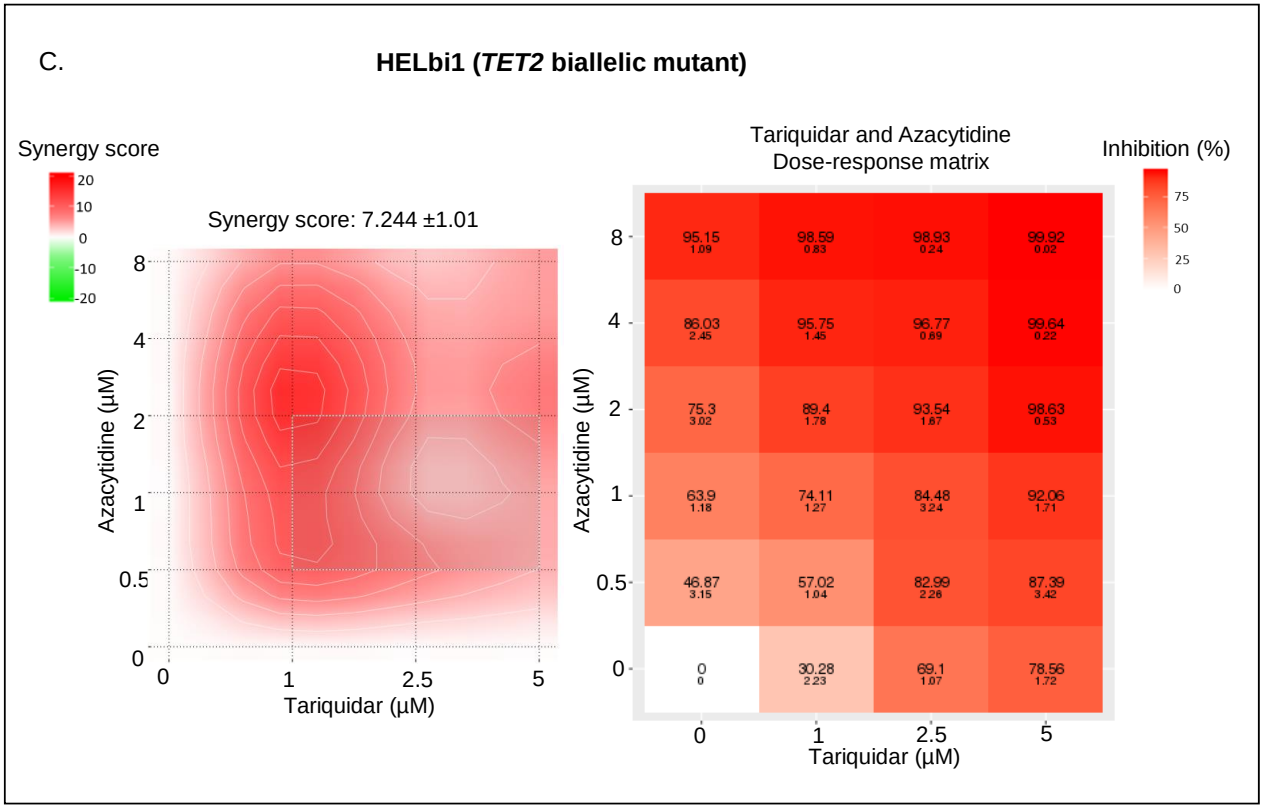

Figure 5.4. Effect of inhibition of ABCB1 using tariquidar on azacytidine sensitivity (continued on next page)

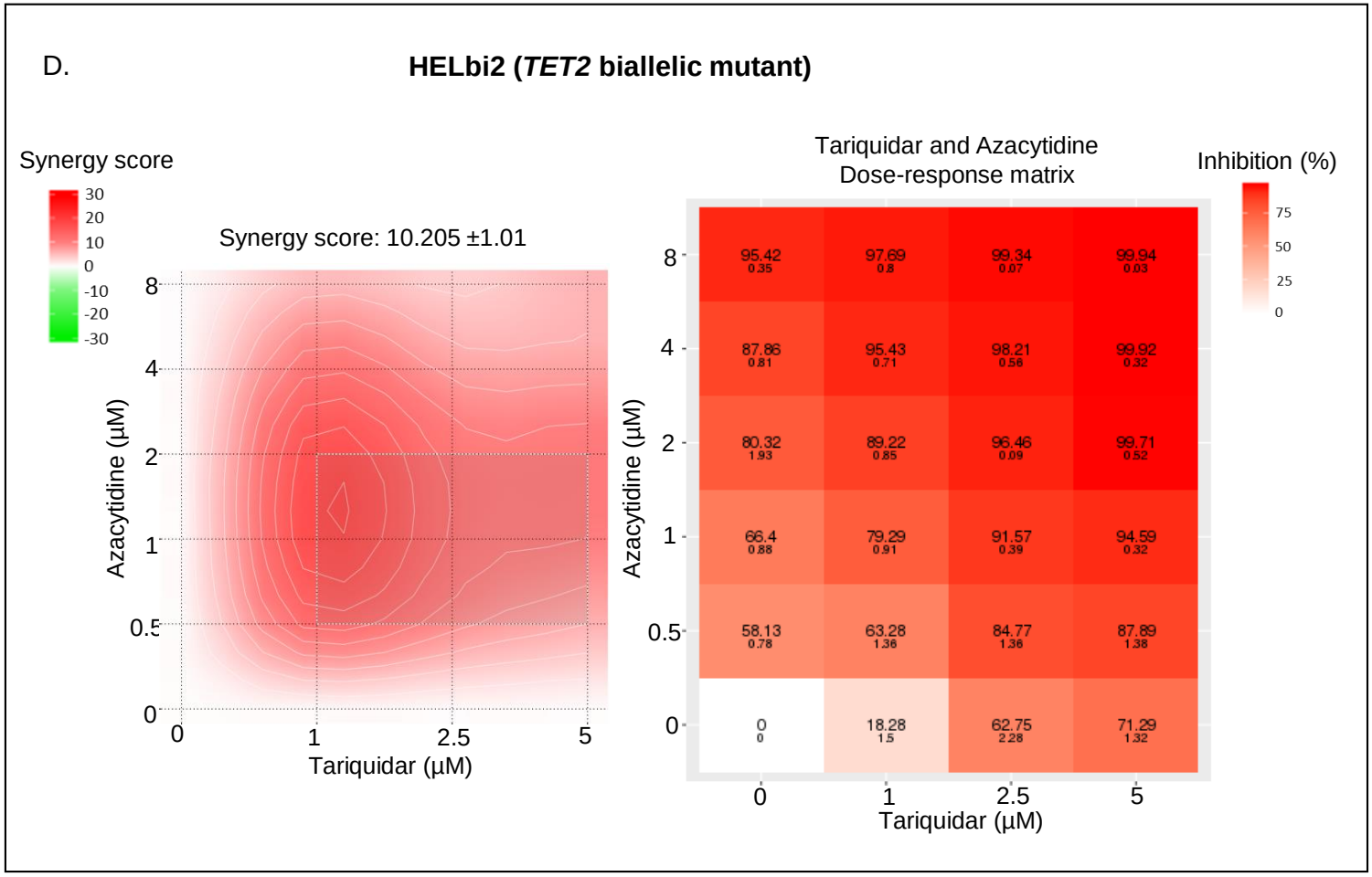

Figure 5.4. Effect of inhibition of ABCB1 using tariquidar on azacytidine sensitivity (continued from previous page). The synergy map (left) shows the synergy score along with the most synergistic area (region shown in darker red) and the dose-response matrix (right) shows the respective cellular growth inhibition percentage, when the *TET2* monoallelic mutant cells (A) HELmono2 and (B) HELmono3 and the *TET2* biallelic mutants (C) HELbi1 and (D) HELbi2 were treated with a combination of increasing doses of tariquidar and azacytidine. The synergy scores and hence the effect of ABCB1 inhibition on azacytidine sensitivity was higher in *TET2* monoallelic mutants compared to *TET2* biallelic mutants. The error values of synergy scores are based on the standard deviation of mean synergy score of three independent replicates. Darker red indicates higher synergy in the synergy map (left) and higher growth inhibition percentage in the dose response matrix (right). In the dose response matrix, the numbers in larger font size (above) denote the growth inhibition percentage and those in smaller font size (below) denote the standard deviation from the mean inhibition percentage of three independent replicates.

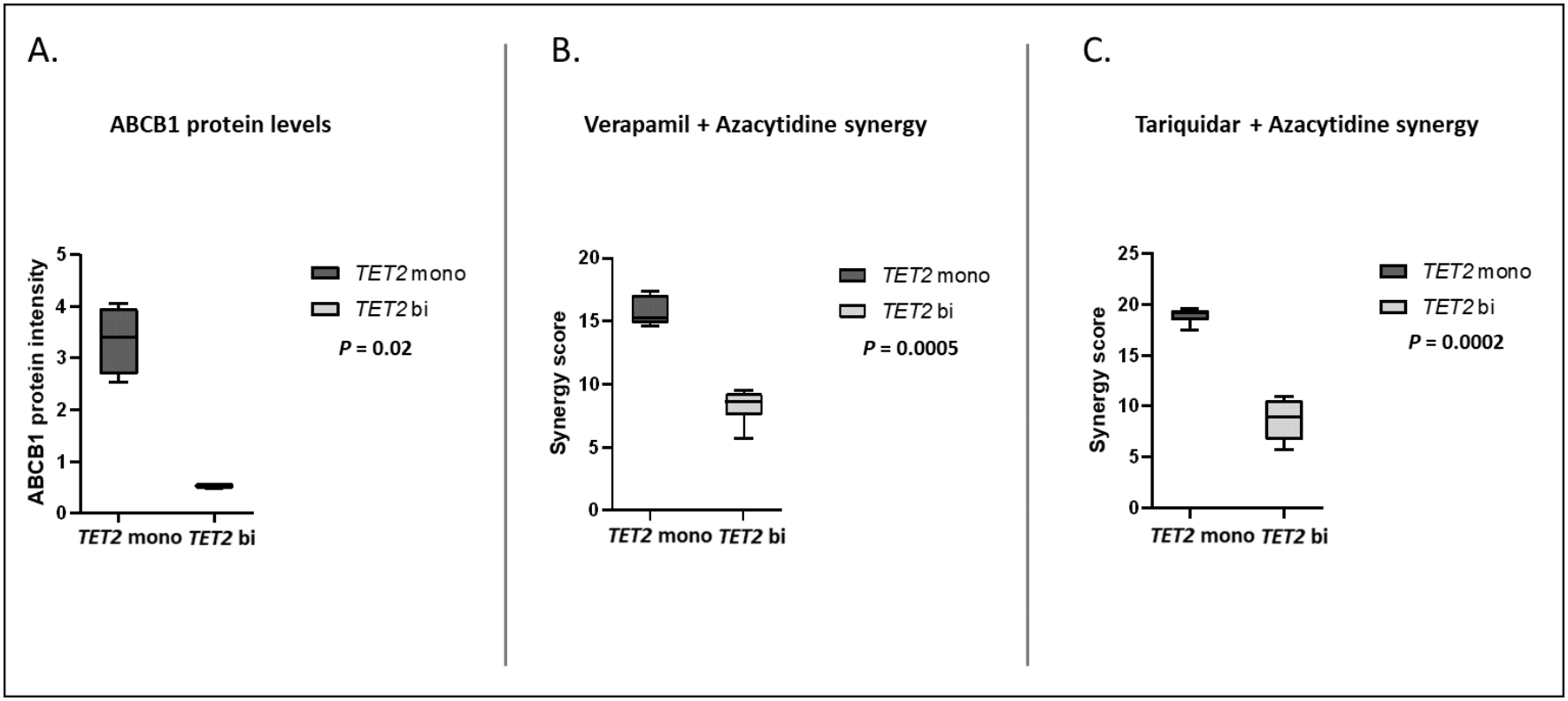

Figure 5.5. Relationship between ABCB1 protein levels and synergy scores. (A) ABCB1 protein levels quantified based on western blot. (B) Synergy scores for verapamil and azacytidine combination. (C) Synergy scores for tariquidar and azacytidine combination. ABCB1 protein levels were significantly higher (P =0.02, ANOVA) in *TET2* monoallelic mutant cell clones compared to TET2 biallelic mutant cell clones (A). Similarly, the synergy scores for both verapamil + azacytidine and tariquidar + azacytidine combinations were significantly higher (P = 0.0005 and 0.0002 respectively, ANOVA) for *TET2* monoallelic mutants compared to *TET2* monoallelic mutants (B and C). Therefore, the impact of ABCB1 inhibition on azacytidine sensitivity was higher in *TET2* monoallelic mutants which express higher levels of ABCB1 protein. The bars in the plot represents the mean, the box represents the interquartile range, and the whisker represents the range between minimum and maximum values.

5.3.3. ABCB1 functional study 2: Characterisation of ABCB1 expression in azacytidine resistant clones of TET2 mutated HEL cells

(i) Generation of azacytidine resistant clones through long-term azacytidine treatment

In order to functionally evaluate the role of ABCB1 in azacytidine resistance, cells that are resistant to azacytidine were developed through long-term exposure of *TET2* monoallelic and biallelic mutated HEL cells to high dose azacytidine giving rise to 90-95% cell death. Cells that survived azacytidine exposure were cloned, sequentially expanded, and tested for azacytidine resistance.

Putative azacytidine resistant (PAR) clones PAR1 and PAR2 were developed from parental HEL cells, PAR3, PAR4 and PAR5 were developed from HELmono2 cells and PAR6 was developed from HELmono3 cells, all of which have a monoallelic *TET2* mutation. All six PAR clones were resistant to azacytidine relative to monoallelic *TET2* mutated parental cells (Figure 5.6 A-B). PAR clones PAR7, PAR8 and PAR9 were developed from HELbi1 with biallelic *TET2* mutation and were also relatively resistant to azacytidine compared to their *TET2* biallelic mutant parental cells (Figure 5.6 C-D).

(ii) Validation of ABCB1 protein levels in azacytidine resistant clones

ABCB1 protein expression was evaluated in azacytidine-resistant clones along with their respective parental cells. When compared to the parental cells from which resistant clones were generated. Specifically, ABCB1 was upregulated at the protein level in PAR1 and PAR2 developed from HEL parental cells (*TET2* monoallelic mutated) (Figure 5.7 A), PAR3 and PAR4 developed from HELmono2 cells (*TET2* monoallelic mutated) (Figure 5.7 B), PAR6 developed from HELmono3 cells (*TET2* monoallelic mutated) (Figure 5.7 C) and PAR7, PAR8 and PAR9 developed from HELbi3 cells (*TET2* biallelic mutated) (Figure 5.7 D). ABCB1 protein upregulation ranged from 1.8-fold in PAR9 to 4.3-fold in PAR7, when compared to the respective parental cells. The consistent upregulation of ABCB1 suggests that this important drug efflux component may be responsible for the chemoresistant phenotype, although it should be noted that PAR5 did not show the same pattern of ABCB1 upregulation (Figure 5.7 B).

(iii) Effect of inhibition of ABCB1 on azacytidine sensitivity of azacytidine resistant clones

Both parental cells and azacytidine-resistant clones were subsequently tested for sensitivity to azacytidine in the presence and absence of an ABCB1 inhibitor such as verapamil. Azacytidine-resistant clones which are relatively resistant to single-agent azacytidine were re-sensitised to azacytidine in the presence of verapamil (Figure 5.8 and Table 5.2). Interestingly, the azacytidine sensitivity of all the *TET2* monoallelic azacytidine resistant clones were brought down to similar levels as that of *TET2* monoallelic parental cells. However, the effect size of sensitisation was relatively lower in *TET2* biallelic azacytidine resistant clones. Irrespective of the extent of sensitisation, there was a clear increase in azacytidine sensitivity of all the azacytidine resistant clones in the presence of verapamil. These data provide further evidence that downregulation of ABCB1 has a role in azacytidine sensitivity of *TET2* null cells.

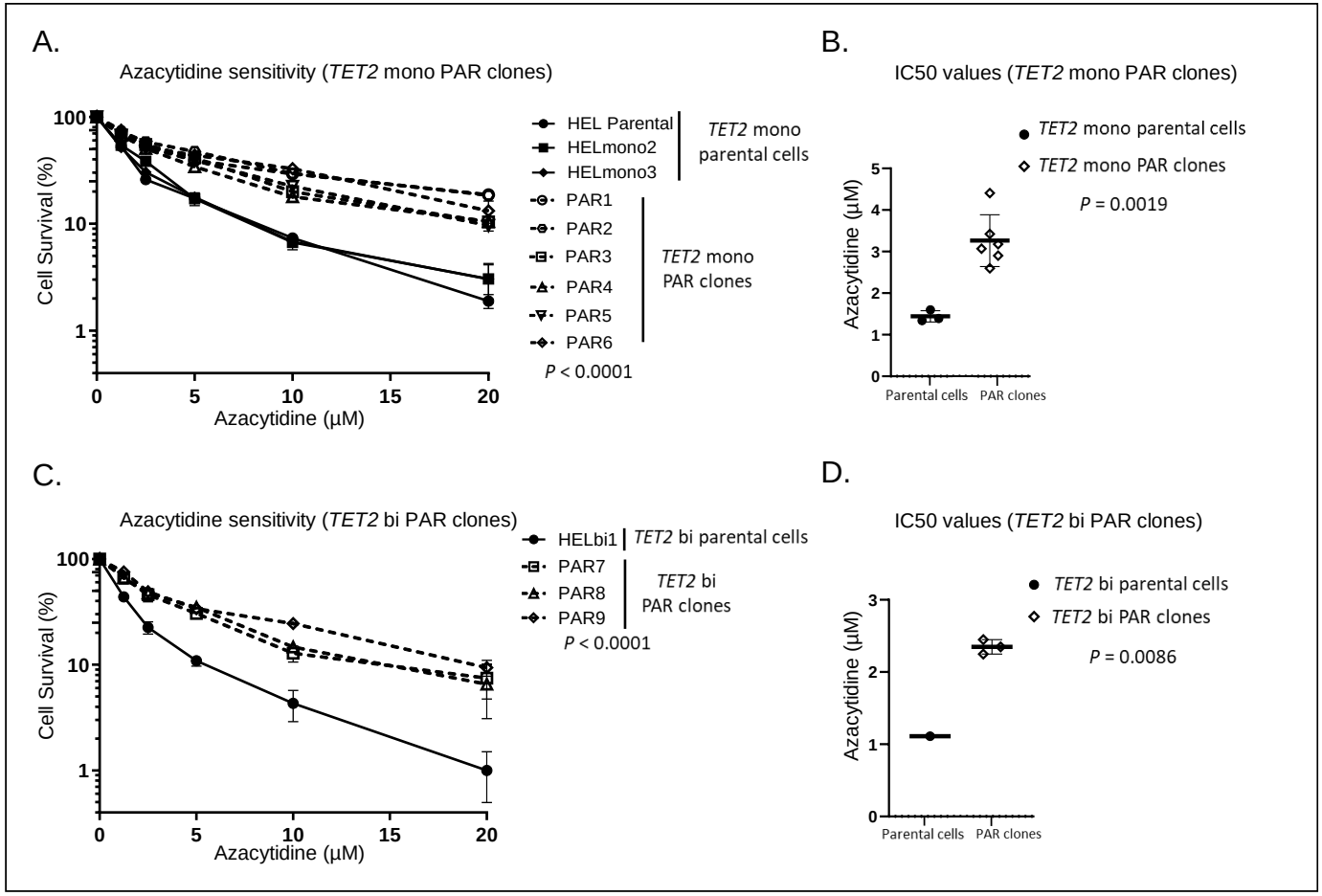

Figure 5.6. Growth inhibition confirms the azacytidine resistance of PAR clones. Cells were exposed to escalating doses (0-20 µM) of azacytidine and the cell viability was measured 96 hours post treatment. (A) PAR clones developed from *TET2* monoallelic mutant cells. PAR1 and PAR2 are developed from HEL parental cells, PAR3, PAR4 and PAR5 are developed from HELmono2 cells and PAR6 is developed from HELmono3 cells. The P values (ANOVA) based on standard deviation from mean of three independent replicates (B) IC50 values of *TET2* monoallelic azacytidine-resistant clones were significantly higher than *TET2* monoallelic parental cells. P values (student's *t*-test) compare the mean IC50 values of *TET2* mono parental cells and azacytidine resistant clones (C) PAR clones such as PAR7, PAR8 and PAR9 developed from *TET2* biallelic mutant HELbi1. The P values (ANOVA) based on standard deviation from mean of three independent replicates (D) IC50 values of *TET2* biallelic mutant azacytidine-resistant clones were significantly higher than the IC50 of HELbi1. P values (student's *t*-test) compare the mean IC50 values of *TET2* mono parental cells and azacytidine resistant clones. All the developed clones had comparatively reduced sensitivity to azacytidine compared to their parental counterparts. Open symbols represent the PAR clones whereas closed symbols represent the cells from which the PAR clones were developed.

PAR, putative azacytidine resistant clones

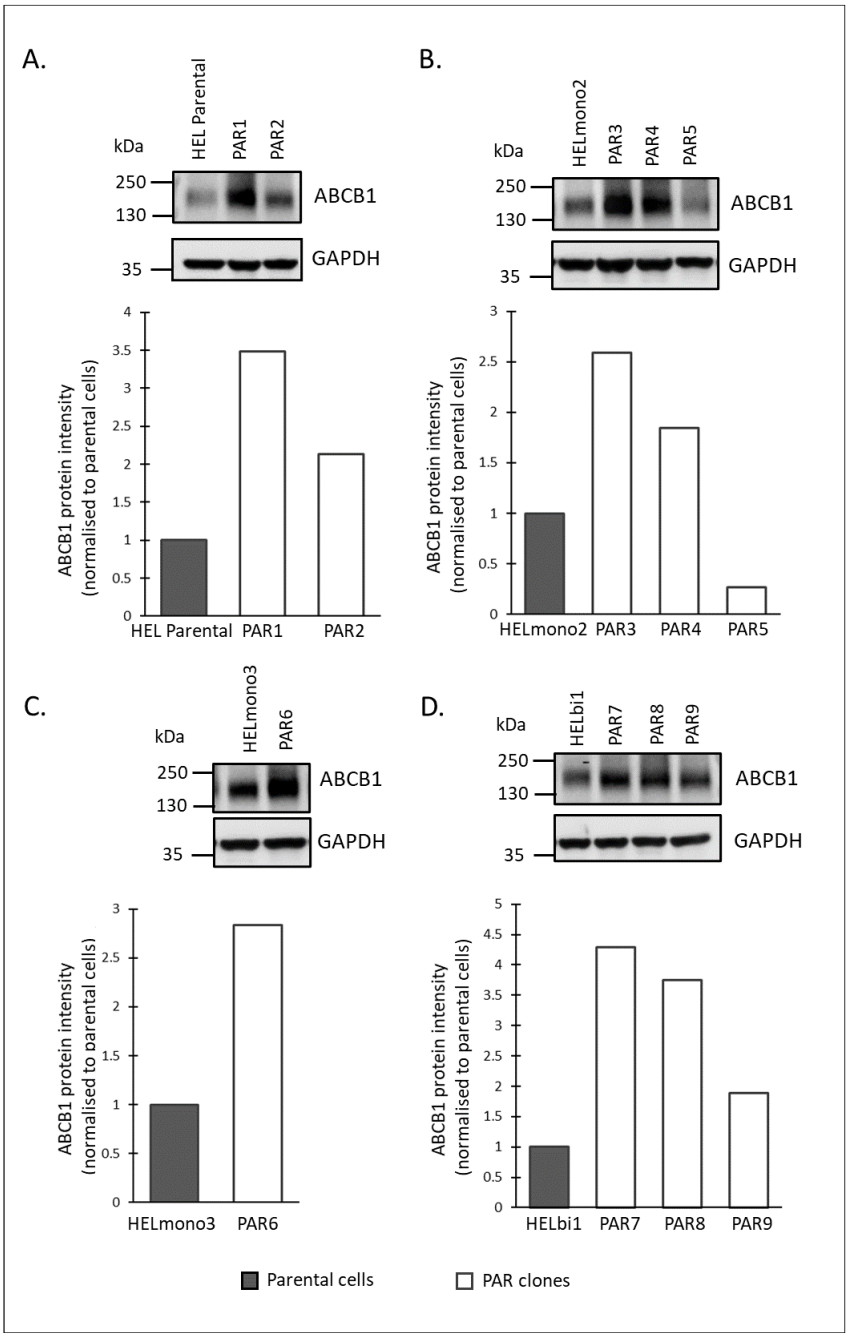

Figure 5.7. ABCB1 protein expression in azacytidine resistant clones. Western blot showing ABCB1 protein expression in aza-resistant clones (A) PAR1 and PAR2 developed from HEL parental cells (B) PAR3, PAR4 and PAR5 developed from HELmono2 cells (C) PAR6 developed from HELmono3 cells and (D) PAR7, PAR8 and PAR9 developed from HELbi1 cells. ABCB1 protein levels of azacytidine-resistant clones were quantified and normalised to their respective parental counterparts. In the bar plot, dark grey denotes parental cells and white denotes PAR clones. ABCB1 was upregulated at protein levels in all the azacytidine-resistant cells (except PAR5) when compared to their respective parental cells.

PAR, Putative azacytidine resistant

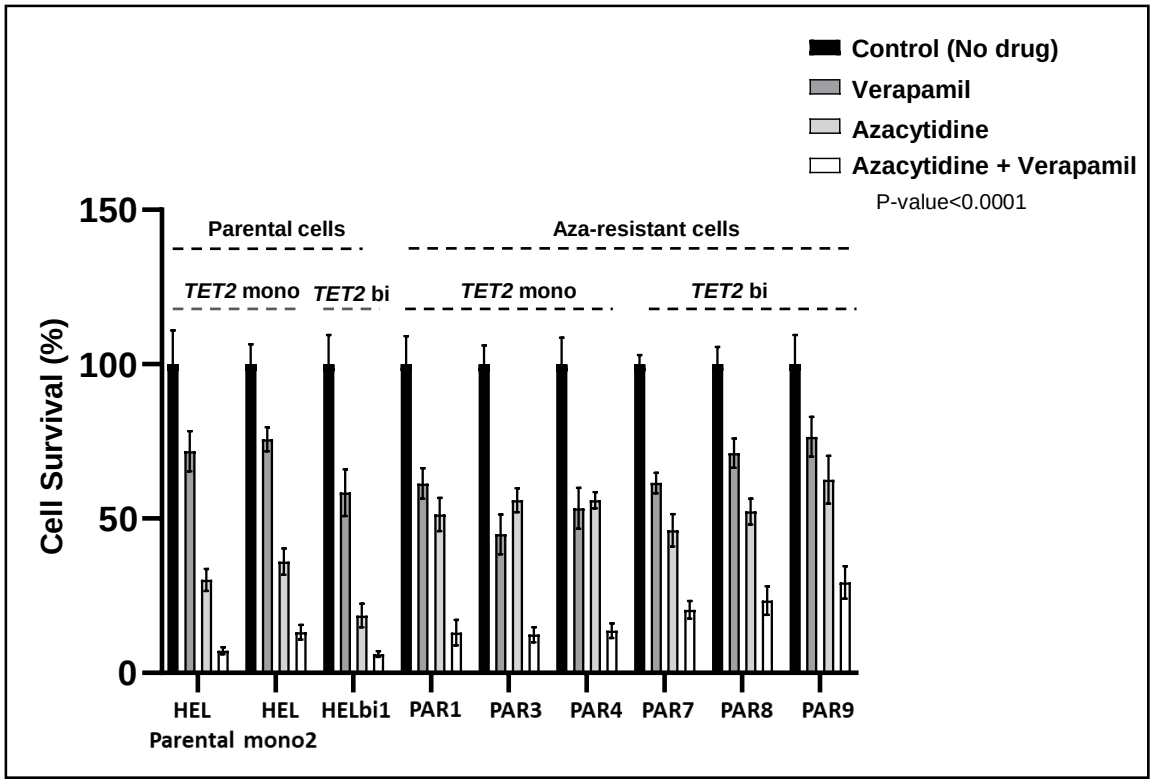

Figure 5.8. Effect of ABCB1 inhibition on azacytidine response of aza-resistant clones. A growth inhibition assay was performed where *TET2* monoallelic aza-resistant clones PAR1, PAR3 and PAR4 and *TET2* biallelic aza-resistant clones PAR7, PAR8 and PAR9 along with their respective parental cells were treated with either single-agent verapamil (50 µM) (dark grey bars) or single-agent azacytidine (5 µM) (light grey bars) or a combination of azacytidine (5 µM) and verapamil (50 µM) (white bars). The control wells were left untreated (black bars). Cell viability was measured after 96 hours of drug exposure and cell survival percentage was calculated based on the cell viability of respective control wells. All the aza-resistant clones were re-sensitised to azacytidine in the presence of ABCB1 inhibitor, verapamil. The summary *P* value is based on the results of three independent replicates. Individual *P* values are provided in Table 5.2.

PAR, Putative azacytidine resistant

Comparison criteria	P_{adj} value	Comparison criteria	P_{adj} value
HEL Parental		**PAR3**	
Control vs. Verapamil	<0.0001	Control vs. Verapamil	<0.0001
Control vs. Azacytidine	<0.0001	Control vs. Azacytidine	<0.0001
Control vs. Azacytidine+Verapamil	<0.0001	Control vs. Azacytidine+Verapamil	<0.0001
Verapamil vs. Azacytidine	<0.0001	Verapamil vs. Azacytidine	0.9616
Verapamil vs. Azacytidine+Verapamil	<0.0001	Verapamil vs. Azacytidine+Verapamil	<0.0001
Azacytidine vs. Azacytidine+Verapamil	<0.0001	Azacytidine vs. Azacytidine+Verapamil	<0.0001
HELmono2		**PAR7**	
Control vs. Verapamil	<0.0001	Control vs. Verapamil	<0.0001
Control vs. Azacytidine	<0.0001	Control vs. Azacytidine	<0.0001
Control vs. Azacytidine+Verapamil	<0.0001	Control vs. Azacytidine+Verapamil	<0.0001
Verapamil vs. Azacytidine	<0.0001	Verapamil vs. Azacytidine	<0.0001
Verapamil vs. Azacytidine+Verapamil	<0.0001	Verapamil vs. Azacytidine+Verapamil	<0.0001
Azacytidine vs. Azacytidine+Verapamil	<0.0001	Azacytidine vs. Azacytidine+Verapamil	<0.0001
HELbi1		**PAR8**	
Control vs. Verapamil	<0.0001	Control vs. Verapamil	<0.0001
Control vs. Azacytidine	<0.0001	Control vs. Azacytidine	<0.0001
Control vs. Azacytidine+Verapamil	<0.0001	Control vs. Azacytidine+Verapamil	<0.0001
Verapamil vs. Azacytidine	<0.0001	Verapamil vs. Azacytidine	<0.0001
Verapamil vs. Azacytidine+Verapamil	<0.0001	Verapamil vs. Azacytidine+Verapamil	<0.0001
Azacytidine vs. Azacytidine+Verapamil	0.0012	Azacytidine vs. Azacytidine+Verapamil	<0.0001
PAR1		**PAR9**	
Control vs. Verapamil	<0.0001	Control vs. Verapamil	<0.0001
Control vs. Azacytidine	<0.0001	Control vs. Azacytidine	<0.0001
Control vs. Azacytidine+Verapamil	<0.0001	Control vs. Azacytidine+Verapamil	<0.0001
Verapamil vs. Azacytidine	0.0149	Verapamil vs. Azacytidine	0.0005
Verapamil vs. Azacytidine+Verapamil	<0.0001	Verapamil vs. Azacytidine+Verapamil	<0.0001
Azacytidine vs. Azacytidine+Verapamil	<0.0001	Azacytidine vs. Azacytidine+Verapamil	<0.0001
PAR2			
Control vs. Verapamil	<0.0001		
Control vs. Azacytidine	<0.0001		
Control vs. Azacytidine+Verapamil	<0.0001		
Verapamil vs. Azacytidine	0.0058		
Verapamil vs. Azacytidine+Verapamil	<0.0001		
Azacytidine vs. Azacytidine+Verapamil	<0.0001		

Table 5.2. Multiple comparisons test on ABCB1 functional study 2. The cytotoxic response corresponding to any two experimental conditions (comparison criteria) were compared and the statistical significance was estimated using multiple comparisons test in two-way ANOVA (GraphPad Prism). The table shows the individual *P* values resulting from respective comparisons.

P_{adj}, adjusted *P* value; PAR, Putative azacytidine resistant

5.4. Discussion

ABCB1 is one of the most characterised ABC transporters, which plays a pivotal role in the cellular efflux of chemotherapeutics, thus conferring cells with chemoresistance to various cytotoxic agents (Robey et al. 2018). As a consequence, upregulation of ABCB1 is a poor prognosis factor in AML and is correlated with shorter event-free survival (Corrêa et al. 2014; Boyer et al. 2019). Function of this multidrug resistance protein as a critical determinant of treatment response is consistent with data presented here that downregulation of ABCB1 transcript in *TET2* biallelic mutated cells is a predictor of sensitivity to azacytidine.

Consistent with the transcript profile, ABCB1 was downregulated at protein level in *TET2* biallelic mutated cells compared to their isogenic *TET2* monoallelic mutated cells. One of the mechanisms of ABCB1 regulation that can be affected by TET2 loss of function is promoter methylation. Hypermethylation of the *ABCB1* promoter along with histone 3 deacetylation is considered as a double repressed state for *ABCB1* which is generally released prior to ABCB1 expression (Chen et al. 2012; David et al. 2004). Moreover, *ABCB1* promoter hypermethylation is observed in approximately 4% of AML (Olesen et al. 2005). Although the mechanism responsible for ABCB1 downregulation in *TET2* null cells is unclear, it is unlikely that this phenotype is a direct impact of hypermethylation driven by *TET2* biallelic mutation because *ABCB1* was not detected as a significantly differentially methylated probe during Illumina 450K methylation array analysis. Consistently, silencing of ABCB1 did not correlate with the methylation status of *ABCB1* in haematopoietic cell lines as well as AML patient samples (Fryxell et al. 1999; Vasconcelos et al. 2021). As well as promoter methylation, there are other mechanisms known to regulate the expression of ABCB1 including gene rearrangements that brings *ABCB1* under the influence of strong transcriptionally active promoters. For example, this mechanism has been reported in patients with refractory ALL who have upregulated ABCB1 expression (Huff et al. 2006; Huff et al. 2005; Robey et al. 2018). Also, mutant TP53 has been reported to enhance *ABCB1* promoter activity whereas wild-type TP53 represses ABCB1 expression (Chen et al. 2012; Chin et al. 1992; Sampath et al. 2001; Scotto 2003; Thottassery et al. 1997). Other regulatory mechanisms include transcriptional regulation by CEBPβ which upregulates ABCB1 transcription activity (Chen et al. 2012).

We demonstrated that inhibition of ABCB1 using either verapamil or tariquidar increased the sensitivity of both *TET2* monoallelic and biallelic mutated HEL cells to azacytidine. Of note, the combinatorial effect of either of these ABCB1 inhibitors was synergistic with azacytidine (denoted by synergy score > 10) in *TET2* monoallelic mutants, whereas the interaction was

additive (denoted by synergy score ≤ 10) in *TET2* biallelic mutants. Moreover, the synergy scores correlated well with ABCB1 protein expression. Specifically, although inhibition of ABCB1 increased the sensitivity of all the cells to azacytidine, the impact was much more pronounced in *TET2* monoallelic mutants which had higher ABCB1 protein expression compared to *TET2* biallelic mutants. Consistently, the effect of ABCB1 inhibition along with induction therapy (cytarabine + daunorubicin) on overall survival rate was maximum on patients with moderate or high ABCB1 protein level expression compared to patients with low or no ABCB1 expression (List et al. 2001), suggesting that the impact of ABCB1 inhibition positively correlates with ABCB1 protein expression.

Inhibition of ABCB1 leads to increased intracellular accumulation of azacytidine in SKM1 (AML) cells, suggesting that ABCB1 is involved in azacytidine efflux (Lainey et al. 2012; Lainey et al. 2013). Treatment of AML cell lines with a combination of azacytidine and erlotinib, which antagonizes ABCB1, demonstrated synergistic cytotoxicity in cell lines such as SKM1, MOLM-13, HL-60 and MV4-11 (Lainey et al. 2013). Interestingly, the synergy was lost when azacytidine was substituted by its mechanistic twin decitabine, whereas synergy was retained even when erlotinib was substituted by gefitinib which are both ABCB1 inhibitors (Lainey et al. 2013). These observations suggest that the combinatorial effect was a response that is specific for azacytidine (and not decitabine) and ABCB1 inhibitor as the synergy was lost when azacytidine was substituted with decitabine. In addition, treatment with either gefitinib or erlotinib invariably increased the intracellular accumulation of azacytidine further supporting the involvement of ABCB1 in azacytidine efflux (Lainey et al. 2013).

As a further functional study, we generated azacytidine resistant clones through long-term exposure of cells to high dose of azacytidine. Of the nine azacytidine resistant clones generated, six were developed from *TET2* monoallelic mutants and three from *TET2* biallelic mutants, and all were relatively resistant to azacytidine compared to their parental counterparts. Interestingly, ABCB1 was upregulated in eight out of nine azacytidine-resistant clones, ranging from approximately 1.8 to 4.3-fold increase in ABCB1 protein levels. However, PAR5 had markedly reduced ABCB1 protein expression compared to its respective parental cells (HELmono2) albeit being resistant to azacytidine and this suggests the potential existence of other possible mechanisms of azacytidine resistance. Messingerova and colleagues also reported upregulation of ABCB1 protein in azacytidine resistant clones developed from SKM1 and MOLM-13 AML cell lines (Messingerova et al. 2015). Inhibition of ABCB1 using verapamil re-sensitised the azacytidine resistant clones to azacytidine. In *TET2* monoallelic azacytidine resistant clones, the azacytidine sensitivity was restored to levels comparable with those in *TET2* monoallelic

mutated parental cells (HEL parental and HELmono2), whereas such a trend was absent in *TET2* biallelic mutant aza-resistant clones.

Numerous clinical trials have investigated the combination treatment of ABCB1 inhibitors and chemotherapeutics. Inhibition of ABCB1 restored sensitivity to chemotherapeutics such as cytarabine and daunorubicin and significantly improved the overall survival rate and event-free survival rate in patients with high-risk AML, despite having no effect on complete remission rate (List et al. 2001). In contrast, other clinical trials that tested ABCB1 inhibitor in combination with chemotherapeutics failed to demonstrate a significant effect on survival rate (Cripe et al. 2010; Binkhathlan et al. 2013). Treatment failure was primarily attributed to differences in baseline ABCB1 expression among the different patients which led to variability in response to ABCB1 inhibitors (Binkhathlan et al. 2013). Another reason was increased toxicity due to inhibition of ABCB1 in normal tissues and off-target activity against other ABC transporters (Binkhathlan et al. 2013; Cripe et al. 2010; Robey et al. 2018). However, the lack of success in clinical trials including ABCB1 inhibitors should not undermine the importance of ABCB1 as a determinant of chemoresistance in AML.

As such, our results highlight the role of ABCB1 in azacytidine resistance and suggest that the downregulation of ABCB1 in *TET2* biallelic mutated cells could play a role in the hypersensitivity of null cells to azacytidine, via a mechanism mediated by reduced azacytidine efflux and hence increased intracellular accumulation of the drug.

5.4.1. Summary of Chapter 5

The results of chapter 5 can be summarised as follows:

(i) ABCB1 was downregulated at the transcript and protein level in *TET2* biallelic mutated cells compared to *TET2* monoallelic mutated cells.

(ii) Inhibition of ABCB1 using verapamil or tariquidar increased the sensitivity of both *TET2* monoallelic and biallelic mutated cells to azacytidine.

(iii) The synergy score when treated with a combination of ABCB1 inhibitor and azacytidine was significantly higher in *TET2* monoallelic mutants compared to *TET2* biallelic mutants, which correlates with the ABCB1 protein levels in these cells. Specifically, there was greater synergy in *TET2* monoallelic mutants that have higher ABCB1 expression, consistent with greater potential forABCB1 inhibition.

(iv) ABCB1 was upregulated in eight out of nine azacytidine resistant clones developed through long-term exposure to azacytidine.

(v) Inhibition of ABCB1 using verapamil re-sensitized the azacytidine resistant clones to azacytidine.

Chapter 6. Effect of TET2 null phenotype on sensitivity of AML cell lines to azacytidine

6.1. Introduction

6.1.1. Modelling biallelic TET2 mutation in SKM1 cells

AML is a clinically heterogenous disease with a broad mutational landscape caused by molecular alterations such as *FLT3*, *NPM1* and *TET2* mutations as well as chromosomal alterations such as t(8;21)(q22;q22), t(15;17)(q22;q21) and inv(16)(p13q22) (Weissmann et al. 2012). Data presented in this thesis has demonstrated that *TET2* biallelic mutation confers hypersensitivity to AML cells. However, these findings are based solely on data derived from a single AML cell line (HEL). As such, it is important to test the azacytidine sensitivity in other AML cell models with *TET2* biallelic mutation.

Replication of our HEL-based cellular model requires another independent leukemic cell line which developed in the background of a *TET2* monoallelic mutation. Another requirement is the absence of reported mutations in other DNA methylation related genes such as *IDH1*, *IDH2*, *WT1*, *DNMT3A* or *DNMT3B* which phenocopy *TET2* mutation. The SKM1 AML cell line is the closest match to these criteria as it is reported to possess a *TET2* monoallelic mutation (c.4253_4254insTT, p.1419fsX30) in exon 10 (Cluzeau et al. 2014; Tate et al. 2018).

(i) Brief history of the SKM1 AML cell line

The SKM1 AML cell line was established from the peripheral blood of a 76-year-old Japanese male with acute monoblastic leukaemia that evolved from prior MDS (Nakagawa et al. 1995; Nakagawa et al. 1993; Nakagawa et al. 1991). The patient was initially diagnosed with refractory anaemia with excess blasts and at this stage there were multiple *RAS* mutations (five point mutations in *N-RAS* and *K-RAS* genes) and absence of any chromosomal abnormalities (Nakagawa et al. 1995). Over a period of 7 months, the patient underwent multiple treatment cycles of low dose cytarabine until chemoresistance developed and he was diagnosed with AML. Thereafter the patient was treated with a combination of cytarabine, daunorubicin, 6-mercaptopurine, prednisolone and mitoxantrone (Nakagawa et al. 1991). However, at 17 months from initial diagnosis, the patient died due to the progression of the disease into acute monoblastic leukaemia. The cell line was derived at this progressive stage of the disease and was characterised to have lost the previously identified *RAS* mutations but have acquired multiple chromosomal alterations (Nakagawa et al. 1991; Nakagawa et al. 1993). The karyotype of this cell line is characterised as follows:

43, XY, del(2)(p11), del(7)(q3?3q3?6), del(9)(q12), t(10;?21)(q10;q10), ?add(11)(q2?3), -12, -14, -17, -19, -20x2, -21, +4mars[19] (Cluzeau et al. 2014)

In addition to the *TET2* monoallelic mutation, SKM1 cells are reported to possess an *ASXL1* homozygous mutation (c.1773C>A,p.YS91X) but are wild-type for *IDH1*, *IDH2*, *WT1*, *DNMT3B* and *DNMT3A* (Cluzeau et al. 2014; Tate et al. 2018). Coexistence of TET2 and ASXL1 mutation is a common event in AML and MDS (Cluzeau et al. 2014). Given the characteristics of SKM1, including the TET2 deletion, this cell line was selected for TET2 targeting in order to generate a null cell clone. It was decided to use the same CRISPR-CAS9 lentiviral vector (Figure 3.1) to generate biallelic TET2 mutation in SKM1 cells such that the experimental conditions of the replicate are as close as the original HEL-based model system.

6.1.2. TET2 null phenotype in AML

In addition to *TET2* biallelic mutation, there are various other mutations that potentially impact TET2 functionality, and which could phenocopy TET2 loss, including mutations in methylation related genes such as *IDH1*, *IDH2* and *WT1*. *IDH1/IDH2* mutations disrupt TET2 function via the accumulation of 2-hydroxyglutarate, an inhibitor of TET2 (Figueroa et al. 2010). Similarly, *WT1* mutated cells have reduced 5-hmC levels as seen in *TET2* mutated AML and are phenotypically similar to *TET2* mutant AML (Rampal et al. 2014).

Given the hypersensitivity of *TET2* biallelic mutated HEL AML cells to azacytidine, it is important to determine whether TET2 expression levels correlate with azacytidine sensitivity in other AML cell lines. This part of the study will be performed using well established AML cell lines, which possess a broad mutational spectrum and variable TET2 protein expression. The genotypic status of *TET2* and other methylation related genes in the panel of 10 AML cell lines that are used in this chapter are summarised (Table 6.1).

Cell line	Genotypic status of							Reference
	TET2	***DNMT1***	***DNMT3A***	***DNMT3B***	***IDH1***	***IDH2***	***WT1***	
AML2	wild-type	---	mutated (homo/hemi)	---	---	---	---	(Tiacci et al. 2012; Jing et al. 2020)
AML3	wild-type	---	mutated (mono)	---	wild-type	wild-type	---	(Tiacci et al. 2012; Chen et al. 2018)
HEL	mutated (mono)	---	---	---	---	---	---	(Abdel-Wahab et al. 2009)
HL-60	wild-type	---	wild-type	---	---	---	wild-type	(Lyu et al. 2017; Tiacci et al. 2012)
Kasumi-1	wild-type	---	wild-type	---	---	---	wild-type	(Lyu et al. 2017; Chen et al. 2018)
MV4-11	---	---	wild-type	---	---	---	---	(Guryanova et al. 2016)
NB4	wild-type	---	wild-type	---	wild-type	---	wild-type	(Lyu et al. 2017; Boutzen et al. 2016; Chen et al. 2018)
SKM1	mutated (mono)	---	wild-type	---	wild-type	wild-type	---	(Cluzeau et al. 2014)
THP1	wild-type	---	wild-type	---	---	---	wild-type	(Morinishi et al. 2020; Koya et al. 2016; Sinha et al. 2015)
U937	wild-type (Partially methylated)	---	wild-type	---	---	---	mutated	(Chen et al. 2015; Tiacci et al. 2012)

Table 6.1. Genotypic status of *TET2* and other methylation related genes in AML cell lines.
---, No reported mutations; hemi, hemizygous; homo, homozygous; mono, monoallelic.

6.2. Aims of Chapter 6

The main objective of this chapter is to understand the role of a TET2 null phenotype or a downregulation of TET2 protein expression as a determinant of sensitivity to azacytidine in independent AML cell lines. The specific experimental aims of this chapter are:

- To generate biallelic *TET2* mutated SKM1 cells using CRISPR-CAS9 gene editing and to establish an isogenic cell model system based on SKM1 cells similar to that of HEL cells.
- To test the role of mutant allele dosage on azacytidine sensitivity of *TET2* monoallelic and biallelic mutated SKM1 cells.
- To characterise the TET2 protein levels in a panel of 10 AML cell lines namely AML2, AML3, HEL, HL-60, Kasumi-1, MV4-11, NB4, SKM1, THP1 and U937.
- To test the azacytidine sensitivity of the different AML cell lines and to determine if there is significant correlation between TET2 phenotype (protein expression) and azacytidine sensitivity (IC50 and IC90).

6.3. Materials and methods

The experimental methods that are specific to chapter 3 are outlined in this section.

6.3.1. CRISPR-CAS9 TET2 targeting in SKM1 cells

The lentivirus one vector system with the gRNA targeting exon 6 of *TET2* (Section **Error! R eference source not found.**was re-purchased (Merck Life Science UK Limited). All the procedures, as described in Section 2.4, were replicated invariably in SKM1 cells, unless otherwise stated. For the two initial replicates the target MOI was 15 and 30 respectively and as there was no evidence of successful transduction, the MOI was increased to 75. Lentiviral particles were introduced in SKM1 cells through spinfection protocol (Section 3.3.1) and the same volume of RF10 media was added in the vehicle-treated control cells, following which all the cells were incubated at 37°C and 5% CO_2.

Cells that had grown adequately were sequentially expanded to higher volumes from a 96-well plate to a 25 cm^3 T25 flask. At this stage, the CRISPR-treated and the control cells were tested for GFP expression using an Attune Nxt flow cytometer (Thermo Fisher Scientific). The forward scatter and side scatter voltages were set at 100 V and 320 V respectively. Gating was performed to include potentially viable cells based on the forward and side scatter values. Fluorescence of the gated population was measured at wavelength 530/30 nm by selecting the GFP option under the blue laser (BL1) channel, which was set at 280 V. All data analysis was performed using FCSExpress 7 flow cytometry software (De Novo). The overlay option in the software was used to precisely compare the GFP expression profile of CRISPR-treated and the respective vehicle-treated control cells.

Following determination of lack of substantial shift in GFP expression in CRISPR-treated cells as compared to control cells, as an alternative approach, an aliquot of CRISPR-treated cells was allowed to grow in the presence of 1 µg/ml of puromycin. Cells were monitored daily with an aim to isolate the puromycin-resistant population of cells. As there was no evidence of successful transduction, the gRNA target region of *TET2* in SKM1 cells was sanger sequenced as per the protocol (Section 3.3.2). The target sequence in SKM1 cells were compared with that of HEL parental cells using Finch TV software, to check for the presence of any major mutations in the target region, that can potentially hamper the binding of CRISPR gRNA.

6.4. Results

6.4.1. Generation of TET2 biallelic mutated cells SKM1 cells

SKM1 cells were treated with CRISPR-CAS9 lentivirus construct consisting of gRNA targeting exon 6 of *TET2*, which is the same construct previously used to generate biallelic *TET2* mutation in HEL cells (Figure 3.1). Three independent repeats of the lentiviral transduction were conducted at increasing multiplicities of infections (MOI) 15, 30 and 75.

Successfully transduced cells are predicted to express GFP and a puromycin resistance cassette. Therefore, the CRISPR-treated cells and vehicle-treated control cells were tested for GFP expression using flow cytometry post-transduction. Despite increasing the MOI, the lentivirus treated cells did not show a substantial shift in GFP peak compared to the control cells during flow cytometry analysis Figure 6.1). As a further confirmation cell cultures were supplemented with puromycin (1ug/ml) with an aim to isolate successfully transduced cells, if any existed. However, cells did not survive in puromycin supplemented media, further confirming that the cells are not transduced with the CRISPR lentivirus.

Furthermore, the CRISPR gRNA targeting region in exon 6 of *TET2* in SKM1 cells was sanger sequenced to confirm whether the region has any indels or mutations that might have potentially hampered the gRNA binding. Of note, the gRNA targeting region in SKM1 was intact and identical to the expected wild-type sequence (Figure 6.2) (Appendix A).

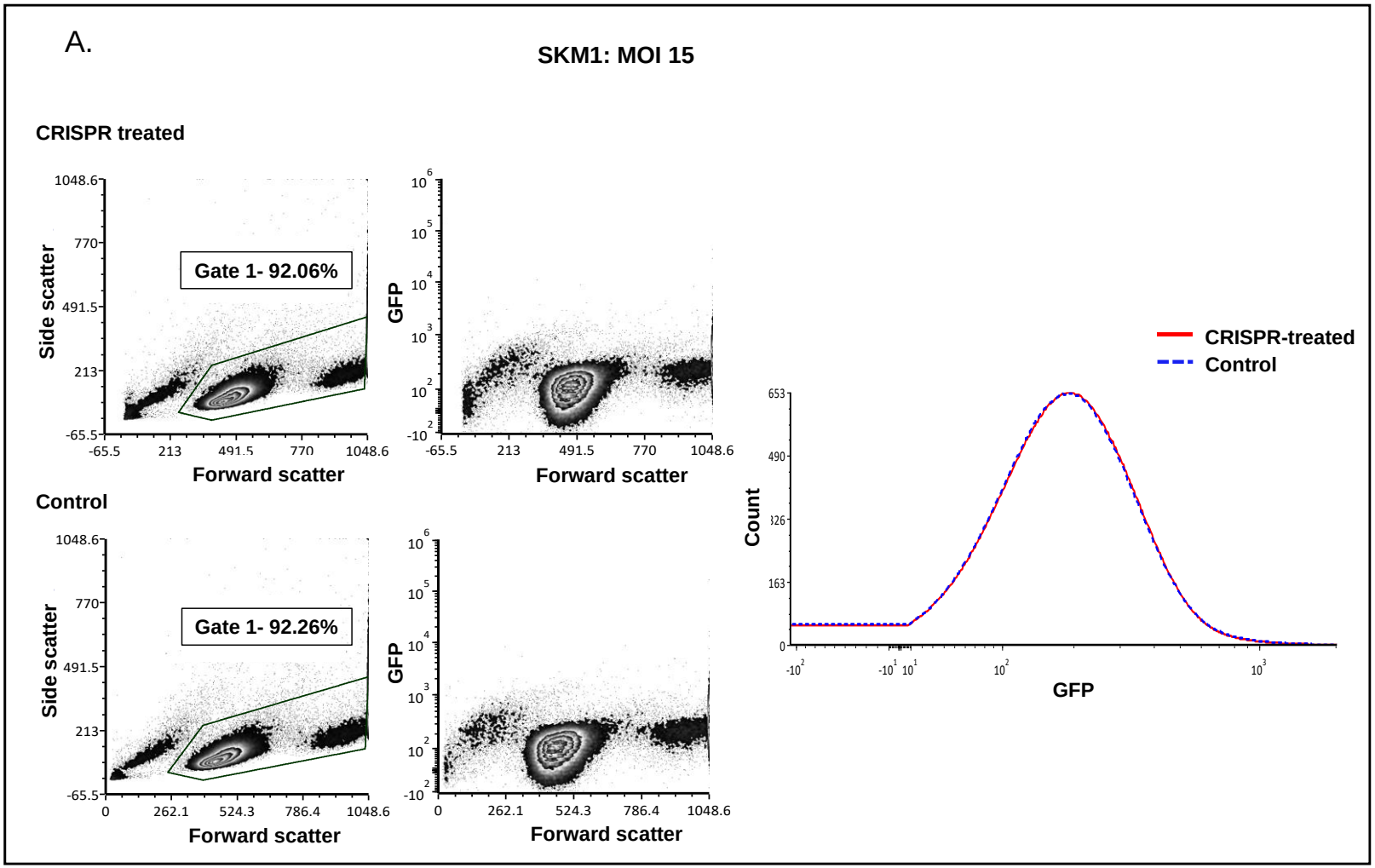

Figure 6.1. Flow cytometric assessment of GFP expression after CRISPR-lentivirus treatment (continued on next page).

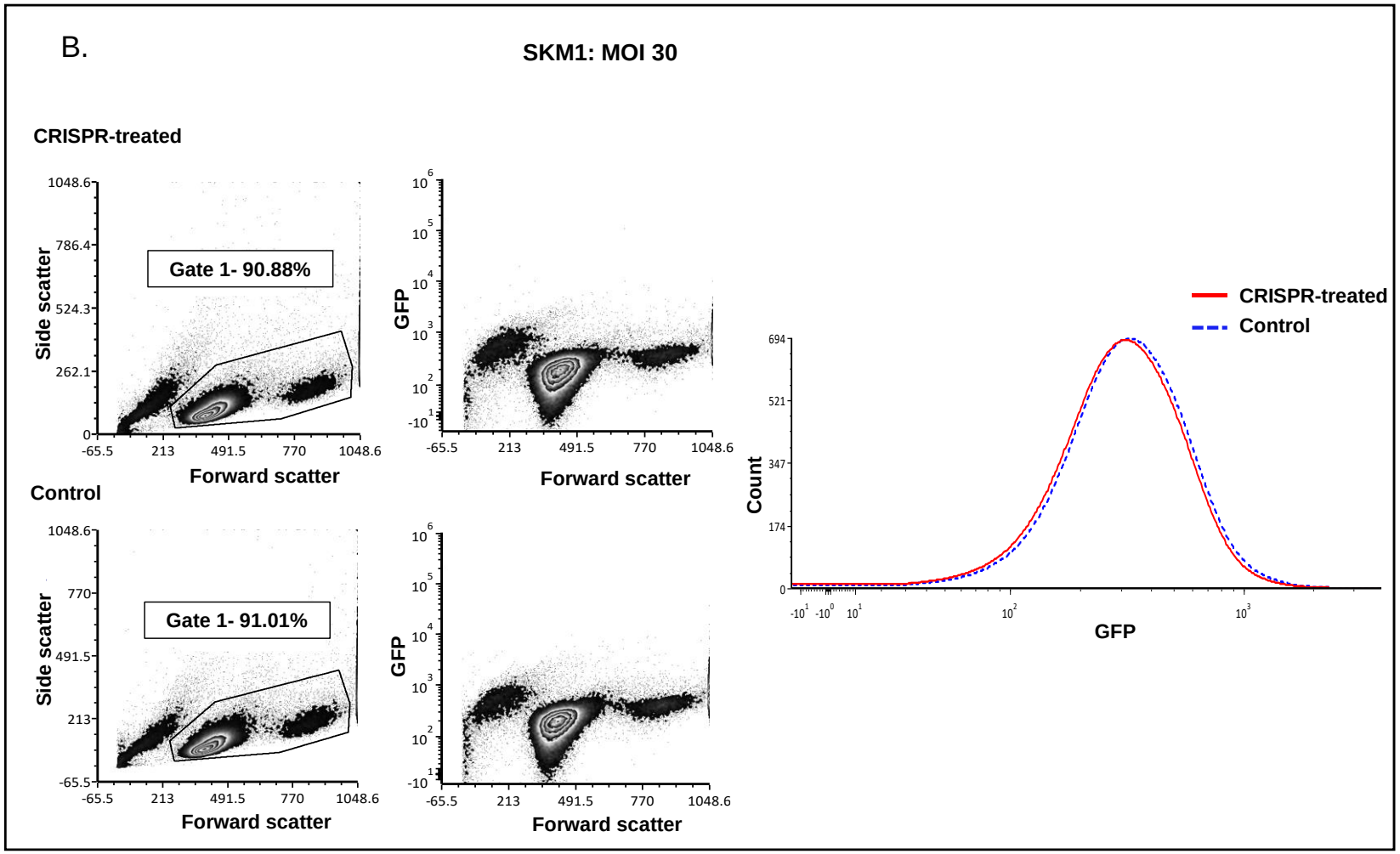

Figure 6.1. Flow cytometric assessment of GFP expression after CRISPR-lentivirus treatment (continued on next page).

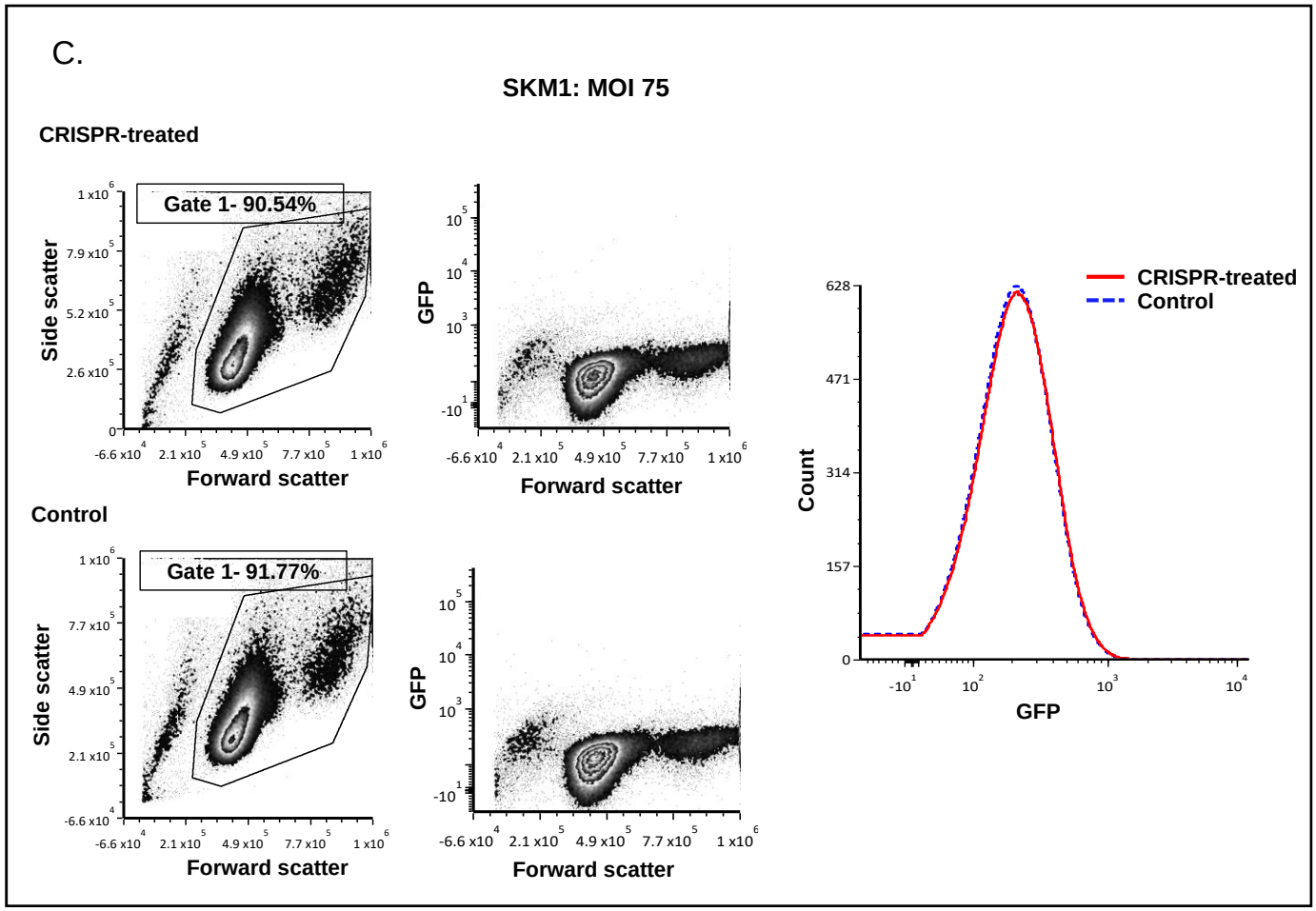

Figure 6.1. Flow cytometric assessment of GFP expression after CRISPR-lentivirus treatment (continued on next page).

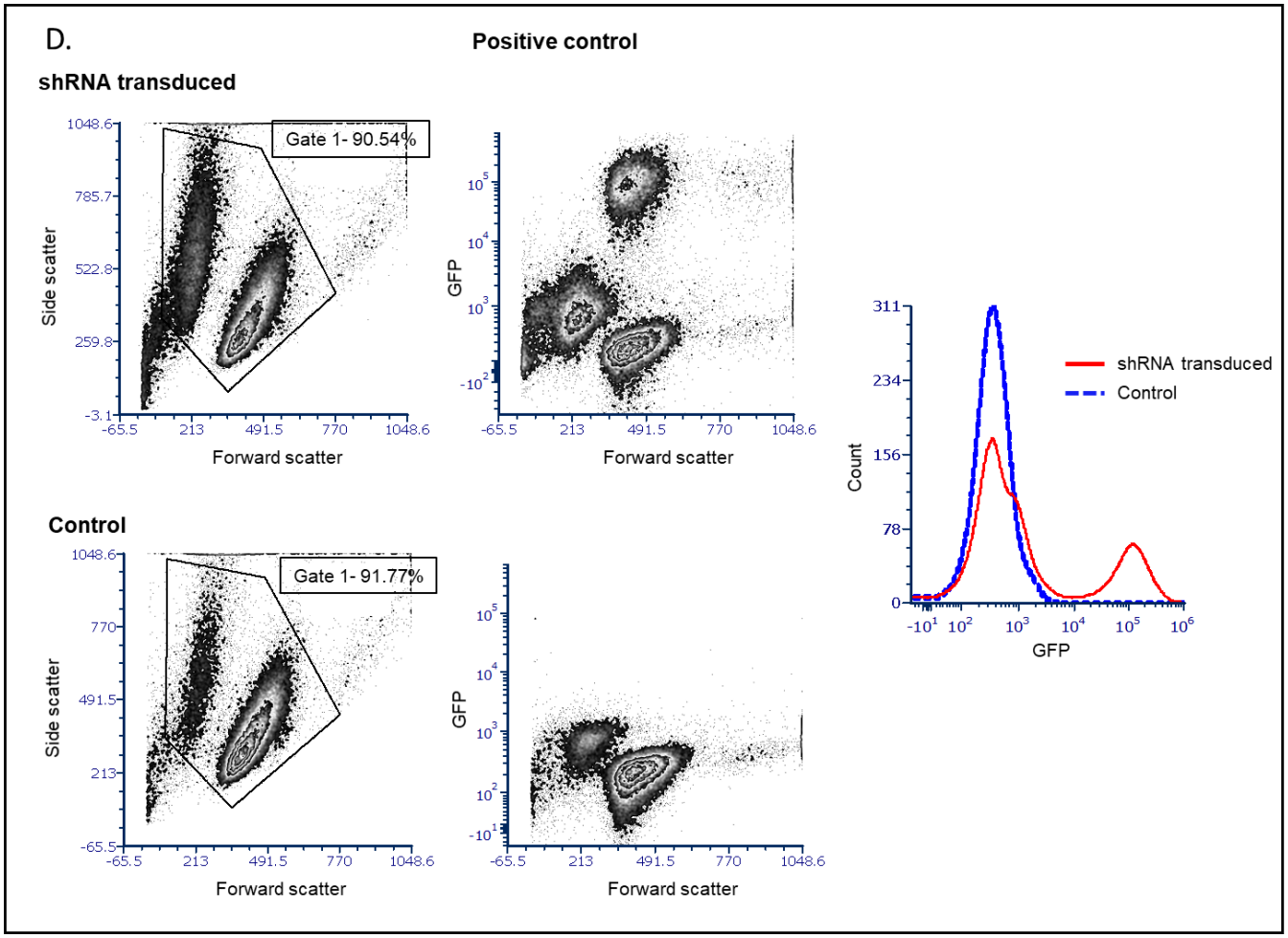

Figure 6.1. Flow cytometric assessment of GFP expression after CRISPR-lentivirus treatment (continued from previous page). 10000 cells which were transduced with an MOI of (A) 15, (B) 30 and (C) 75 and the respective control cells which were vehicle-treated (bottom panel) (D) Kasumi-1 cell line stably transduced with lentiviral vector pLenti.TTMPVIV-(N) expressing shNTC transcript and venus fluorescently labelled protein was used as a positive control in the experiment (cells were provided by Asmida Isa, Newcastle University). Cells were assessed for GFP expression using an Attune NxT Flow Cytometer. The dot plots on the left-hand side represents the forward and side scatter profiles which are used to identify viable cells (Gate 1). The percentage of cells included in the gated population are shown in respective boxes. The dot plots shown in the middle represent the GFP fluorescence of the gated cells. The histogram overlay compares the number of cells expressing GFP at increasing intensities in CRISPR-treated and control population of cells. There was no evidence of successful transduction as there was no substantial shift in GFP peak in either of the three MOIs showing that the GFP expression profile was similar in CRISPR-treated and control cells. The positive control cells demonstrated GFP positivity compared to the respective non-transduced control cells.

GFP, Green Fluorescence Protein, shNTC, non-targeting control shRNA

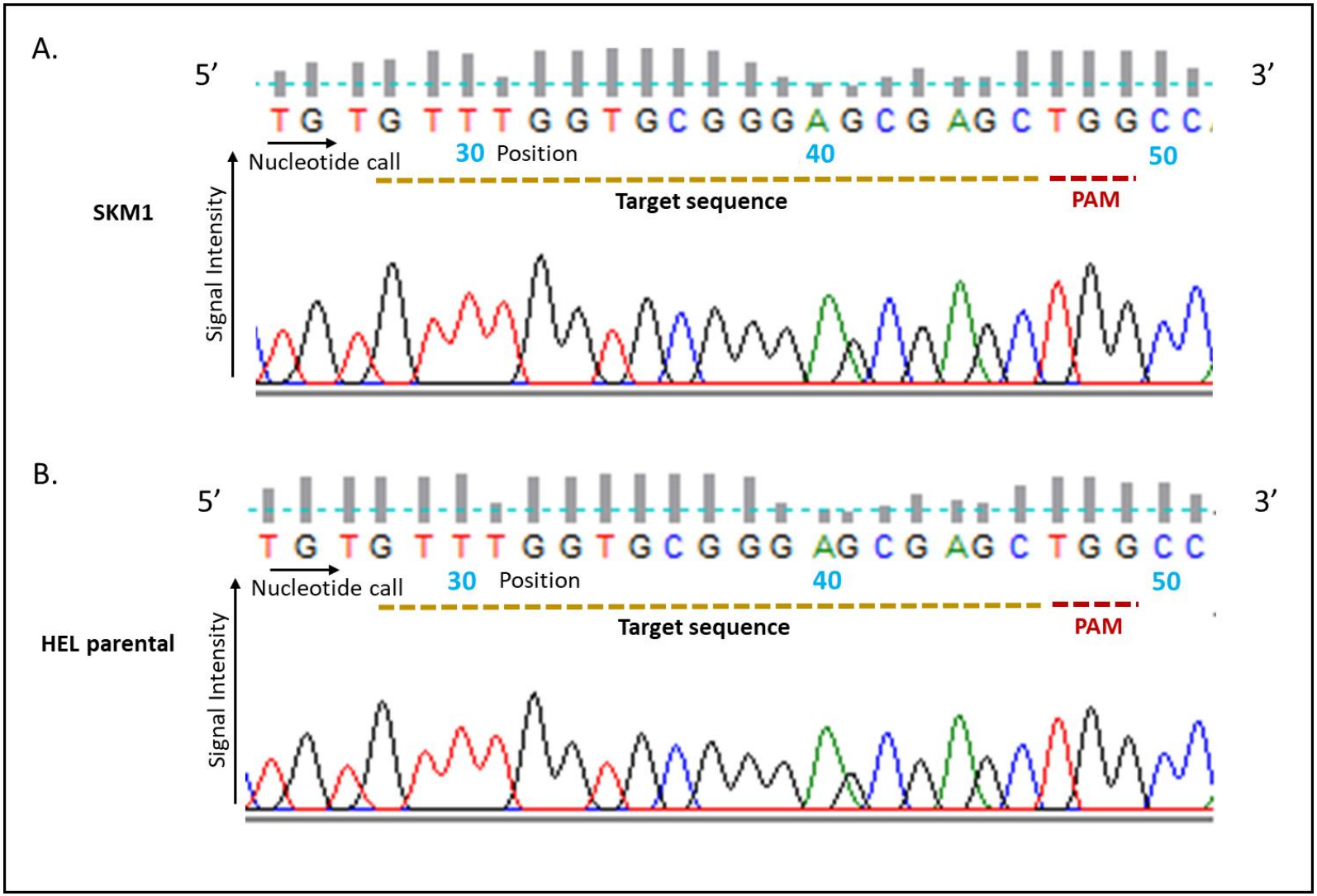

Figure 6.2. Comparison of the CRISPR gRNA binding site sequences of SKM1 and HEL parental cells. Forward sequencing (5'-3') electropherogram of (A) SKM1 and (B) HEL parental cells showing wild-type sequences in *TET2* exon 6, confirming that the CRISPR gRNA binding site in SKM1 cells is intact. The respective sequence signal intensity, nucleotide call, base position numbers (blue), CRISPR-target region (gold dotted lines) and PAM sequence (red dotted lines) are shown.

CRISPR, clustered regularly interspaced short palindromic repeats; gRNA, guide RNA; PAM, protospacer adjacent motif

6.4.2. Effect of TET2 protein expression on azacytidine sensitivity in AML cells

(i) Characterisation of TET2 phenotype in AML cell lines

As an alternative approach to CRISPR-based *TET2* knockout in SKM1 cells, it was decided to phenotypically characterise the baseline TET2 protein levels in a panel of 10 AML cell lines, namely AML2, AML3, HEL, HL-60, Kasumi, NB4, MV4-11, SKM1, THP1 and U937 (Figure 6.3). THP1, AML2, Kasumi and HL-60 AML cell lines had relatively high TET2 expression compared to other AML cell lines. Intriguingly, SKM1 cells which are reported to be *TET2* monoallelic mutated (c.4253_4254insTT, p.1419fsX30) had complete absence of TET2 protein despite the presumed presence of an intact *TET2* allele.

(ii) Correlation of TET2 protein levels and azacytidine sensitivity in AML cells

A panel of 10 AML cell lines were treated with increasing doses of azacytidine and cell survival was measured as a function of dose. Consistent with an association between TET2 protein levels and response to azacytidine, THP1 cells had the highest TET2 protein and were the least azacytidine sensitive, whereas SKM1 cells, which were completely null for TET2 protein, was the most azacytidine sensitive (Figure 6.4A). Of note, there was a statistically significantly correlation between TET2 protein levels and azacytidine IC50 (R^2=0.88, P<0.0001) and IC90 (R^2=0.77, P=0.0008) (Figure 6.4 B-C).

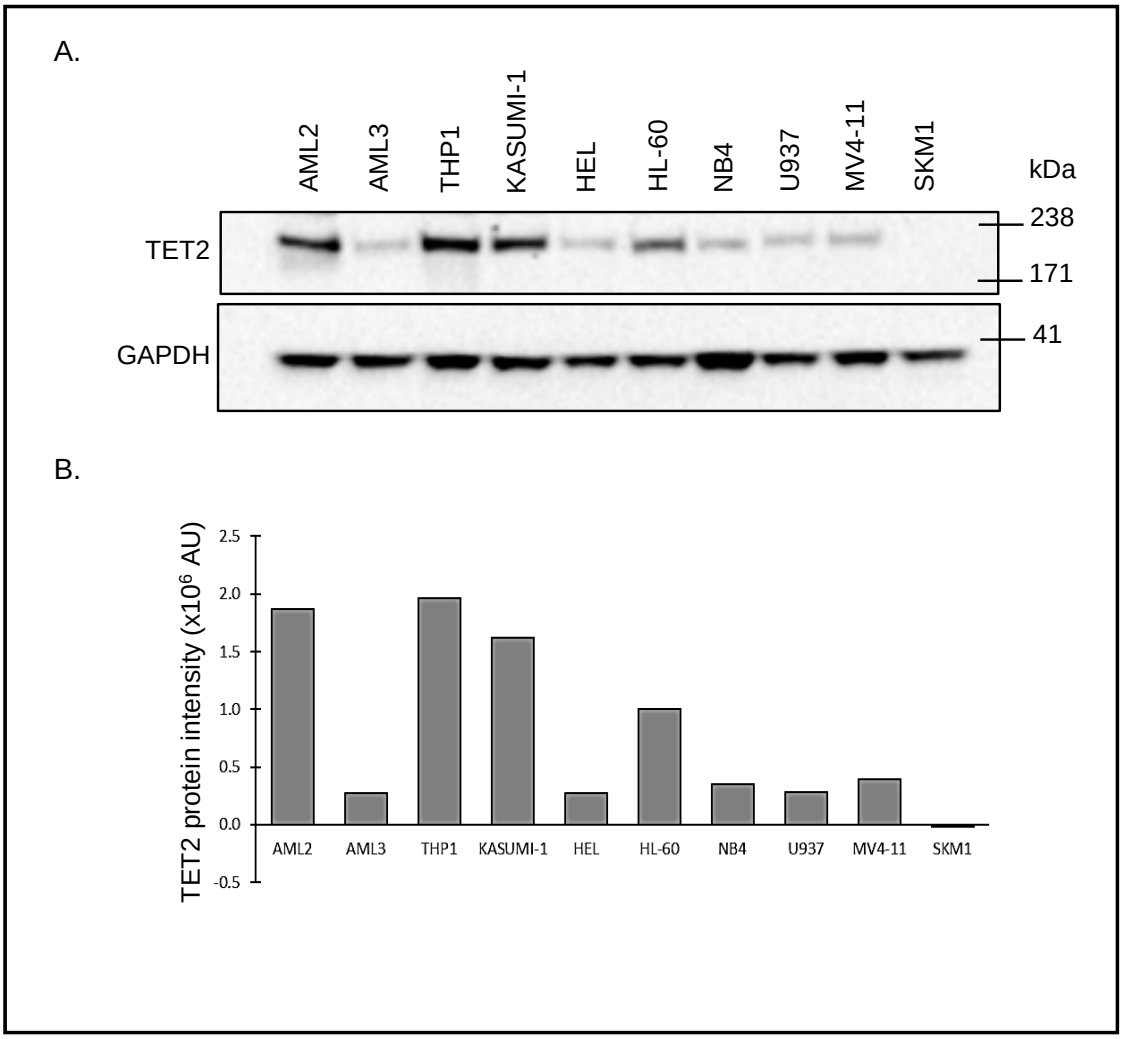

Figure 6.3. TET2 protein levels in a panel of 10 AML cell lines. A. Western blot showing TET2 protein expression profile in a panel of 10 AML cell lines. GAPDH was used as a loading control. B. TET2 protein expression was quantified in each cell line. AU represents arbitrary units as measured by the Fuji LAS-300 Image Analyzer System. SKM1 cells are completely null for TET2. THP1, AML2, Kasumi-1 and HL-60 have higher TET2 expression compared to other AML cell lines.

AU, arbitrary units; kDa, kilodaltons

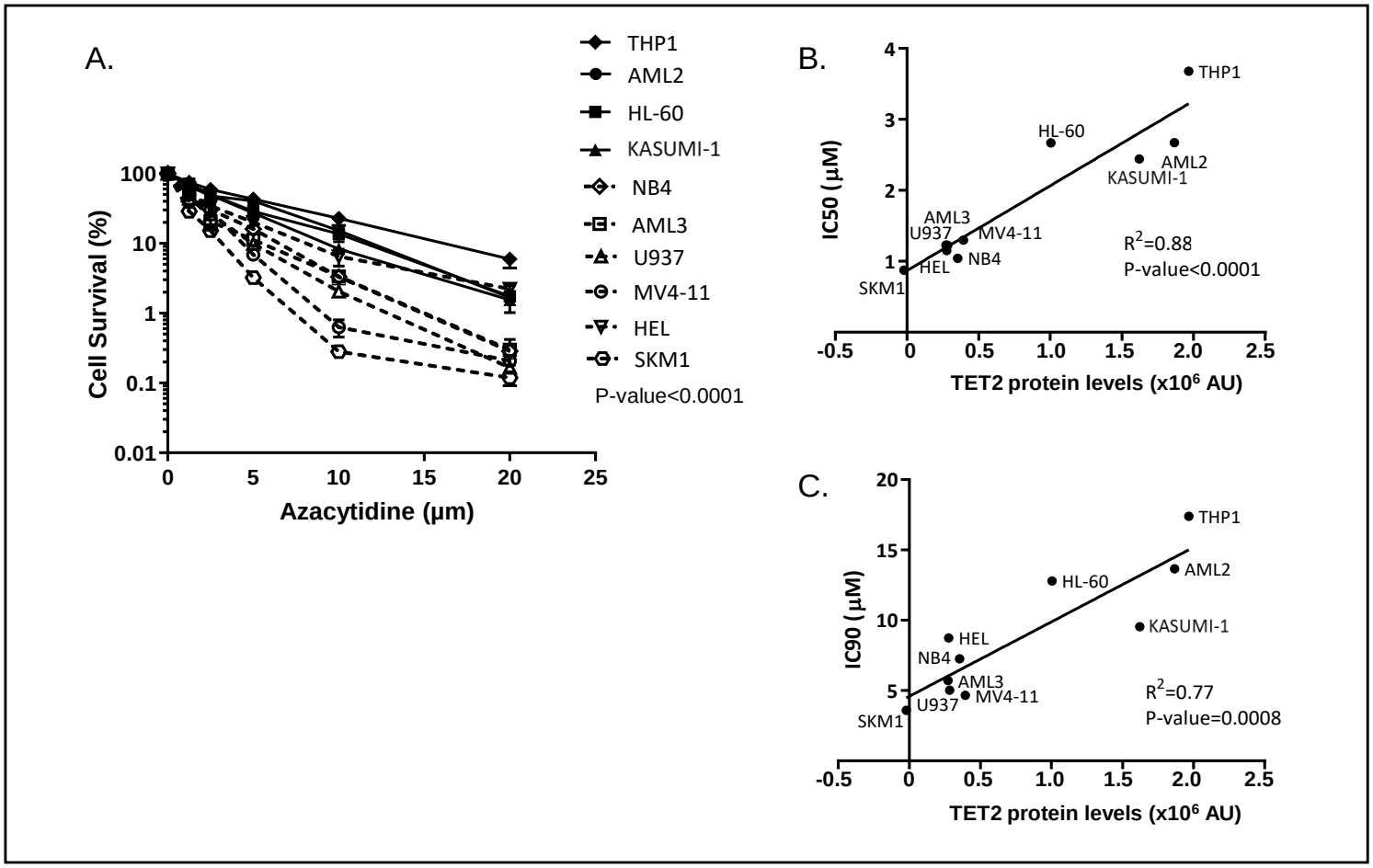

Figure 6.4. Correlation of TET2 protein expression and azacytidine sensitivity. (A) Growth inhibition profile AML cell lines in response to increasing doses (0-20 µM) of azacytidine at 96 hours post treatment. Closed symbols and solid lines show the cell lines with relatively high TET2 protein expression (THP1, AML2, HL-60 and Kasumi-1) and open symbols and dotted lines show the cell lines with relatively low TET2 protein expression (NB4, AML3, HEL, U937, MV4-11, SKM1). P-values are calculated by two-way ANOVA. (B) TET2 protein expression was quantified in each cell line and plotted against azacytidine IC50 (B) and IC90 (C) values. AU represents arbitrary units as measured by the Fuji LAS-300 Image Analyzer System. The R^2 values denote the respective correlation scores.

6.5. Discussion

Characterisation of the mutational landscape in AML paved the way for the development of targeted therapeutics such as the *FLT3* and *IDH1/IDH2* inhibitors (Döhner et al. 2021). Our results demonstrate that biallelic *TET2* mutation confers hypersensitivity to azacytidine and emphasise the importance of considering *TET2* mutant allele dosage in determining the treatment strategy for AML.

As most of our results are solely based on a single cell line (HEL) and its clones, we decided to replicate the key experiments on an independent cell line (SKM1), which is reported to be *TET2* monoallelic mutated (c.4253_4254insTT, p.1419fsX30) (Cluzeau et al. 2014). As such, it was decided to exploit the same CRISPR-CAS9 lentivirus vector that was previously used to target exon 6 of *TET2* in HEL cells. However, despite increasing the target MOI to 75, there was no evidence of successful lentivirus transduction following puromycin selection or flow cytometry. Nevertheless, it must be noted that the increased target MOI of 75 was still lower than the target MOI of 100, which was previously used for HEL cells. However, we were limited by the available volume of lentivirus and the optimal SKM1 cell density requirement and therefore it was not possible to increase the MOI any further.

Generally, the factors contributing to failure of CRISPR-CAS9 lentiviral transduction are low titre values, the gRNA design, inefficient selective markers, target cells that are difficult to transfect and overgrowth of wild-type cells which outcompete CRISPR edited cells (Jin et al. 2020). When attempting to target *TET2* in SKM1 cells, all the experimental conditions (except the target MOI) were identical to those used to target *TET2* in HEL cells. There are previous records of successful CRISPR-targeting in SKM1 cells, including studies that have used electroporation to introduce targeting vectors and spinfection (Hoerster et al. 2021; Yang, Liu, et al. 2017), similar to our protocol, suggesting that SKM1 cells are amenable to CRISPR-targeting.

Furthermore, the CRISPR gRNA binding site in SKM1 cells was sequenced and confirmed to be wild-type demonstrating that targeting failure was not due to mutations in the guide binding site. Taken together, the possible reasons for failure to generate *TET2* targeted SKM1 cell clones might be low transduction efficiency due to an insufficient target MOI, low expression of selection markers or potential outgrowth of wild-type cells which outcompete the CRISPR-targeted pool of cells.

The primary objective of CRISPR-targeting the SKM1 cells was to target the retained wild-type *TET2* allele and therefore rendering the cells completely null for TET2 protein expression. Intriguingly, western blot confirmed that the parental SKM1 cells are completely null for TET2 protein expression, despite retaining a wild-type *TET2* allele. Of note, SKM1 cells are heterozygous for *TET2* mutation and there are no copy number alterations affecting the *TET2* locus 4q24 (Cluzeau et al. 2014), suggesting that silencing of retained wild-type *TET2* allele is a possible explanation for the complete lack of protein expression. Regardless, the complete lack of TET2 protein expression in SKM1 AML cells renders CRISPR-targeting of the wild-type *TET2* allele redundant.

Characterisation of TET2 protein expression levels in a panel of 10 AML cell lines revealed a significant correlation with azacytidine sensitivity (in terms of IC50 and IC90 values). THP1 cells that have the highest TET2 expression was the least azacytidine sensitive, whereas SKM1 cells that are completely null for TET2 were the most azacytidine sensitive. Despite retaining a wild-type *TET2* allele, the TET2 null phenotype in SKM1 cells appears to be a key determinant of azacytidine sensitivity. As such, unlike biallelic *TET2* mutation which leads to a complete loss of TET2 protein, monoallelic *TET2* mutation does not necessarily correlate with protein expression.

In addition, the *WT1*-mutated and *TET2*-partially methylated U937 AML cells that are otherwise genetically wild-type for *TET2* (Chen et al. 2015), expressed very low levels of TET2 protein and were concomitantly highly sensitive to azacytidine compared to most other AML cell lines. Likewise, U937 cells are reported to be highly sensitive to decitabine (the functional analogue of azacytidine) compared to KG-1 cells which express higher TET2 protein levels (Chen et al. 2015). THP1, HL-60, AML2 and Kasumi-1 AML cell lines, which all express higher TET2 levels, were all relatively resistant to azacytidine compared to AML cell lines with lower TET2 expression.

Taken together, the strong correlation between TET2 phenotype (protein expression) and azacytidine sensitivity highlights the predictive value of TET2 expression profiling in determining azacytidine response. Overall, the results from this chapter argue in favour of characterising *TET2* mutant allele dosage and TET2 protein expression to identify patients who might benefit from azacytidine based treatment.

6.5.1. Summary of Chapter 6

The results of chapter 6 can be summarised as follows:

(i) SKM1 cells did not show evidence of successful lentiviral transduction or CRISPR-targeting. As such, a *TET2* biallelic mutated isogenic model system could not be generated in SKM1 cells.

(ii) SKM1 cells are completely null for TET2 protein, despite retaining a wild-type *TET2* allele. Consistent with a complete lack of TET2 protein expression, SKM1 cells are acutely sensitive to azacytidine. Therefore, *TET2* monoallelic mutation is not necessarily a good indicator of TET2 protein levels or response to azacytidine.

(iii) There was a statistically significantly correlation between TET2 protein levels and azacytidine sensitivity (in terms of IC50 and IC90 values (in a panel of 10 AML cell lines.

(iv) A TET2 null phenotype sensitises AML cells to azacytidine and therefore TET2 expression profiling and *TET2* mutant allele dosage can be used to predict response to azacytidine.

Chapter 7. Concluding Discussion

7.1. Mutational complexity of AML and the need for precision medicine

AML is a genetically and clinically heterogenous disease characterised by a plethora of different mutations. Moreover, more than one mutation concurrently occurs in the same patient revealing the concerted interaction of multiple mutations in the pathogenesis of AML (Hou et al. 2020; Grimwade et al. 2016). Despite the heterogeneity in mutational profile, the treatment for AML primarily depends on Ara-C and anthracycline-based induction therapy and a significant number of patients are either not fit for this intensive chemotherapy regimen or eventually experience relapse or refractoriness (Hou et al. 2020).

The recent advances in sequencing technologies have paved the way for understanding the enormous mutational complexity of AML, refining risk stratification and prognostication, and development of novel targeted therapies. For instance, *FLT3* mutations are among the most frequent mutations in AML, which leads to constitutional activation of *FLT3* and its downstream pathways (Döhner et al. 2021). As such, clinical development of therapeutics targeting FLT3 such as midostaurin, sorafenib and other FLT3 inhibitors has been fundamental to the evolution of precision medicine in AML. Among the other molecular alterations in AML, mutations in DNA methylation related genes account for 44% of all gene mutations (Coombs, Tallman, et al. 2016). In particular, *TET2* is somatically mutated in up to 25% of AML, with approximately 75% monoallelic mutations and 25% biallelic mutations (Ahn et al. 2015). *TET2* mutations in AML are associated with poor outcome (Patel et al. 2012). Given the role of *TET2* as a CHIP-driver gene (age-associated event) (Section 1.4.3. (ii)), it is not surprising that *TET2* mutations in AML are associated with older age (p-value<0.001) (Metzeler, Maharry, Radmacher, Mrózek, et al. 2011).

Generally, elderly AML cases over 65 years are considered not fit for Ara-C and anthracycline-based induction therapy due to increased toxicity, co-morbidities, organ impairment and other treatment-related complications such as severe infections (Palmieri et al. 2020). As a consequence, the 5-year relative survival for patients aged above 70 years is as low as ~1% (HMRN(2010-2016)). Therefore, there is an urgent clinical need to deliver treatments which are personalised rather than a “one-size-fits-all” approach, and for that it is essential to identify therapeutically exploitable genomic vulnerabilities which can be targeted using existing or newly developed agents. To this end, our data identifies *TET2* biallelic mutation as a therapeutically exploitable genomic vulnerability and presents a novel therapeutic strategy based on hypomethylating agents to target *TET2*-mutated AML.

7.2. Development of azacytidine-based therapeutic strategy for *TET2*-mutated AML and understanding the role of *TET2* mutant allele dosage on azacytidine sensitivity

The index patient described in section 1.6 presented with a biallelic *TET2* mutation-positive AML that was chemoresistant to standard remission induction therapy while being acutely sensitive to single agent azacytidine, which led to long-term morphological remission. This suggests the potential of single-agent azacytidine treatment in inducing CR in AML harboring *TET2* biallelic mutation, including in patients with disease refractory to conventional induction therapy and in patients of adverse risk group who are not fit for chemotherapy due to their age or poor performance status.

Our data derived from an isogenic cell model system strongly supports the azacytidine hypersensitivity of *TET2* biallelic mutated AML compared to *TET2* monoallelic mutated cells. *TET2* biallelic mutated cells had a lower proliferation rate and clonogenicity in the presence of azacytidine compared to *TET2* monoallelic mutated cells. Consistently, *TET2* mutations have been positively correlated with better response following treatment with hypomethylating agents in MDS (Itzykson et al. 2011; Bejar et al. 2014; Lin et al. 2017; Cedena et al. 2017; Patel et al. 2017), although some other studies did not replicate these findings (Döhner et al. 2018). However, none of these studies investigated the role of *TET2* mutant allele dosage on response to hypomethylating agents.

TET2 mutant allele dosage as such did not have a significant impact on baseline proliferation rate in liquid culture or clonogenicity of cells in soft agar. In contrast, murine *Tet2*$^{-/-}$ cells displayed increased proliferation and clonogenicity compared to *Tet2*$^{+/-}$ cells (Li et al. 2011). *Tet2*$^{-/-}$ mice had significantly shorter latency to the development of AML and reduced overall survival compared to *Tet2*$^{+/-}$ mice (Li et al. 2011; Shih et al. 2015). Murine *Tet2*$^{-/-}$ cells outcompete *Tet2* wild-type cells at a much faster rate compared to *Tet2*$^{+/-}$ cells, showing a mutant allele dose dependent increase in self-renewal potential (Moran-Crusio et al. 2011). Taken together, *TET2* mutant allele dosage has an important role in leukaemia development and response to azacytidine.

TET2 biallelic mutated cells were relatively more sensitive to decitabine compared to *TET2* monoallelic mutated cells, however the effect size was much weaker compared to that seen with azacytidine. Furthermore, *TET2* mutant allele dosage did not have a significant impact on response to other commonly used chemotherapeutic agents, such as cytarabine, daunorubicin, fludarabine, idarubicin and calicheamicin (the conjugated drug in GO). *TET2* mutations are

associated with poor response to standard chemotherapy and reduced overall survival in AML (Metzeler, Maharry, Radmacher, Mrózek, et al. 2011). The chemoresistance showed by the index patient (section 1.6) in response to standard chemotherapy based on Ara-C and anthracycline, further supports the inferior response of *TET2* biallelic mutated AML to standard remission induction chemotherapy.

As such, our data demonstrates that *TET2* mutant allele dosage has a pronounced effect on response to azacytidine in AML and specifically that *TET2* biallelic mutated cells are hypersensitive to azacytidine compared to *TET2* monoallelic mutated cells.

7.3. Unveiling the mechanism behind the azacytidine sensitivity of *TET2* null cells

Given the analogous function of TET2 and azacytidine in demethylation of DNA, it was initially speculated that the global hypermethylation phenotype driven by *TET2* biallelic mutation was responsible for azacytidine sensitivity. To this end, the methylation and transcript profiles of *TET2* monoallelic and biallelic mutated cells were interrogated using Illumina 450K methylation array and RNAseq, respectively. Although we noted a significant increase in the number of hypermethylated probes in *TET2* biallelic mutated cells compared to *TET2* monoallelic mutated cells, the effect size was relatively modest due to the presence of numerous hypomethylated CpG probes and concomitant upregulation of multiple transcripts in *TET2* biallelic mutated cells. Therefore, it seems unlikely that the modest genome-wide hypermethylation phenotype has a major role in sensitizing *TET2* biallelic mutated cells to azacytidine. Consistently, baseline DNA methylation level did not correlate with better response to decitabine in MDS (Shen et al. 2010). Taken together, the accumulating evidence suggests the role of specific genes or pathways in rendering *TET2* biallelic mutated cells hypersensitive to azacytidine.

We identified multiple genes that are differentially expressed in *TET2* biallelic mutated cells compared to *TET2* monoallelic mutated cells and clinically relevant pathways associated with the differentially expressed genes. The ribosomal pathway was among the top enriched pathways associated with the differentially expressed transcripts and the majority of the pathway components were downregulated at transcript and protein levels in *TET2* biallelic mutated cells compared to *TET2* monoallelic mutated cells. Consistently, the total cellular protein concentration which is a cumulative measure of cellular protein synthesis and hence ribosomal pathway functionality, was significantly lower in *TET2* biallelic mutated cells compared to *TET2* monoallelic mutated cells. Of note, 80-90% of azacytidine is incorporated

into RNA, alters RNA methylation and inhibits protein synthesis (Saliba et al. 2021; Schaefer et al. 2009; Lee et al. 1976; Lu et al. 1980; Stresemann et al. 2008; Christman 2002).

Given that accumulating evidence demonstrates the discordance between pre-treatment DNA methylation levels and response to azacytidine in AML (Shen et al. 2010; Treppendahl et al. 2014), it has been hypothesized that changes in RNA metabolism might influence azacytidine action (Saliba et al. 2021). To this end, Cheng and colleagues demonstrated an interaction between RNA methyltransferases, RNA splicing components and RNA Pol II in modulating azacytidine response (section 4.4) (Cheng et al. 2018). Moreover, upregulation of ribosomal pathway and intensive protein synthesis have been associated with poor response to azacytidine (Monika Belickova et al. 2016). As such, given that protein synthesis is already downregulated in *TET2* biallelic mutated cells and it can be speculated that any residual protein synthesis activity in these cells is essential for survival and hence *TET2* biallelic mutated cells might be hypersensitive to any further inhibition in protein synthesis mediated by azacytidine. In addition, although azacytidine and decitabine are structurally very similar, azacytidine incorporates into RNA whereas decitabine incorporates into DNA, and hence decitabine is generally considered as a more potent DNA methylation inhibitor (Christman 2002). Even then, the effect size of the hypersensitivity response of *TET2* biallelic mutated cells was much stronger against azacytidine compared to decitabine, which suggests a potential role of RNA-related pathways as determinants of sensitivity.

Our investigations also identified significant downregulation of ABCB1 in *TET2* biallelic mutated cells at the transcript and protein level. ABCB1 is an important azacytidine efflux transporter which is involved in pumping azacytidine out of the cell thus leading to a reduction in cellular azacytidine sensitivity. Inhibition of ABCB1 has been reported to increase intracellular accumulation of azacytidine and hence exert a synergistic effect with azacytidine on cell death (Lainey et al. 2012; Lainey et al. 2013). Consistently, we demonstrated increased azacytidine sensitivity in the presence of either of the ABCB1 inhibitors — tariquidar and verapamil. Moreover, drug synergy between azacytidine and both ABCB1 inhibitors was significantly higher for *TET2* monoallelic mutants compared to *TET2* biallelic mutants. As such, the extent of synergy is directly correlated with ABCB1 protein levels. In addition, azacytidine resistant cells (developed through long-term exposure of cells to azacytidine) had significant upregulation of ABCB1 protein expression compared to their respective parental azacytidine sensitive counterparts, and inhibition of ABCB1 in the azacytidine resistant clones resensitized these cells to azacytidine. Taken together, these data demonstrate strong association between ABCB1 expression and azacytidine response and suggest a role for

ABCB1 downregulation in *TET2* biallelic mutated cells as a potential mechanism for azacytidine sensitivity in these cells.

Unlike many other mutations that exert only a limited impact by impairing their own function or the function of direct downstream targets, *TET2* mutations have a global impact by affecting genome-wide methylation status leading to impairment of a plethora of different pathways with varied functions. Similarly, azacytidine is a drug that has a global genome-wide impact affecting multiple genes and pathways. As such, it is possible that multiple genes and pathways might have a concerted role in sensitizing *TET2* biallelic mutated cells to azacytidine. Here we identified two important determinants of azacytidine sensitivity in *TET2* biallelic mutated cells namely downregulation of ABCB1 and ribosomal pathway.

7.4. The effect of TET2 null phenotype on azacytidine sensitivity

Mutations in other methylation-related genes that operate in the TET2 hydroxymethylation pathway phenocopy *TET2* loss, including *IDH1*, *IDH2* and *WT1* mutation. *IDH1/IDH2* mutations occur in approximately 15-30% of AML and are always heterozygous, suggesting an oncogenic gain of function phenotype where mutation in one allele is sufficient for malignant transformation (Medeiros et al. 2017; Figueroa et al. 2010; Schnittger et al. 2010). IDH1/IDH2 catalyse the oxidative decarboxylation of isocitrate into α-KG, whereas mutant *IDH1/IDH2* generates 2-HG, which is a structural analogue of α-KG and inhibits α-KG dependent enzymes such as TET2 (Figure 1.8) (Figueroa et al. 2010; Medeiros et al. 2017). Consistently, *IDH1/IDH2* mutated AML acquire a hypermethylation phenotype compared to wild-type *IDH1/IDH2* AML and TET2-mediated demethylation of 5-mC is significantly reduced in the presence of mutant *IDH1* compared to wild-type *IDH1* (Figueroa et al. 2010). In contrast to *IDH1/IDH2* mutations, *WT1* mutation occurs in approximately 6-15% of AML and can be either monoallelic or biallelic (Rampal et al. 2016; Hou et al. 2010). WT1 physically interacts with and recruits TET2 to the target site prior to demethylation (Wang, Xiao, et al. 2015) and *WT1* loss-of-function mutations confer a cellular global hypermethylated phenotype which impairs hematopoietic differentiation similar to that observed in *TET2/IDH1/IDH2* mutated AML (Rampal et al. 2014). Taken together, mutations in *IDH1/IDH2/WT1* inhibits TET2 activity and therefore phenocopy TET2 loss in AML. Therefore, it can be speculated that mutations in these genes might potentially confer sensitivity to azacytidine in AML, though this needs to be fully deciphered. Consistently, a retrospective analysis involving 42 AML patients receiving hypomethylating therapy (either decitabine or azacytidine) demonstrated a

significantly higher response rate in patients with *IDH1/IDH2* mutation compared to those without the mutation (Emadi et al. 2015; Rampal et al. 2016).

Unlike *TET2* biallelic mutation that renders cells null for TET2 protein, *TET2* monoallelic mutation need not necessarily predict the TET2 protein levels. For instance, SKM1 cells which are *TET2* monoallelic mutated (Cluzeau et al. 2014) were completely null for TET2 protein whereas U937 cells which are *WT1*-mutated and genetically wild-type for *TET2* (Chen et al. 2015) had very low TET2 protein expression. Consistent with TET2 protein levels, both SKM1 and U937 cells were highly sensitive to azacytidine. Furthermore, we noted a remarkable correlation between TET2 protein levels and azacytidine sensitivity across a panel of 10 AML cell lines. Therefore, in addition to *TET2* biallelic mutation, TET2 protein expression can also be an important determinant of azacytidine sensitivity.

7.5. Summary of Findings

In summary, this project has demonstrated that *TET2* mutant allele dosage in AML is an important indicator of azacytidine response and *TET2* biallelic mutation results in hypersensitivity to azacytidine. Based on our results, downregulation of ABCB1 can be an important mechanism behind the sensitivity of TET2 null cells to azacytidine. In addition, it can be speculated that downregulation of ribosomal pathway has a potential role in sensitising TET2 biallelic mutants to azacytidine (Figure 7.1). Furthermore, the significant correlation between TET2 protein levels and azacytidine sensitivity demonstrates the importance of TET2 protein expression profiling to identify patients with AML likely to be sensitive to azacytidine, irrespective of *TET2* mutation status.

Comprehensive mutation profiling using sequence and copy number analysis might be required to completely characterise complex mutations affecting *TET2* locus and thereby to predict the likely outcome after azacytidine treatment. Screening for TET2 status is particularly important in elderly patients with AML, where the disease is likely a consequence of *TET2*-driven clonal haematopoiesis and the probability of *TET2* biallelic mutation or TET2 null phenotype is higher (Hirsch et al. 2018). Multiple clinical studies have demonstrated excellent responses in terms of improved CR rates and reduced mortality in elderly patients treated with azacytidine (Almeida et al. 2017; Pollyea et al. 2013; Wen et al. 2020), although the *TET2* mutation status in these patients were not investigated. Therefore, larger clinical studies are warranted to confirm the impact of *TET2* mutant allele dosage on response to azacytidine in AML patients of all age groups.

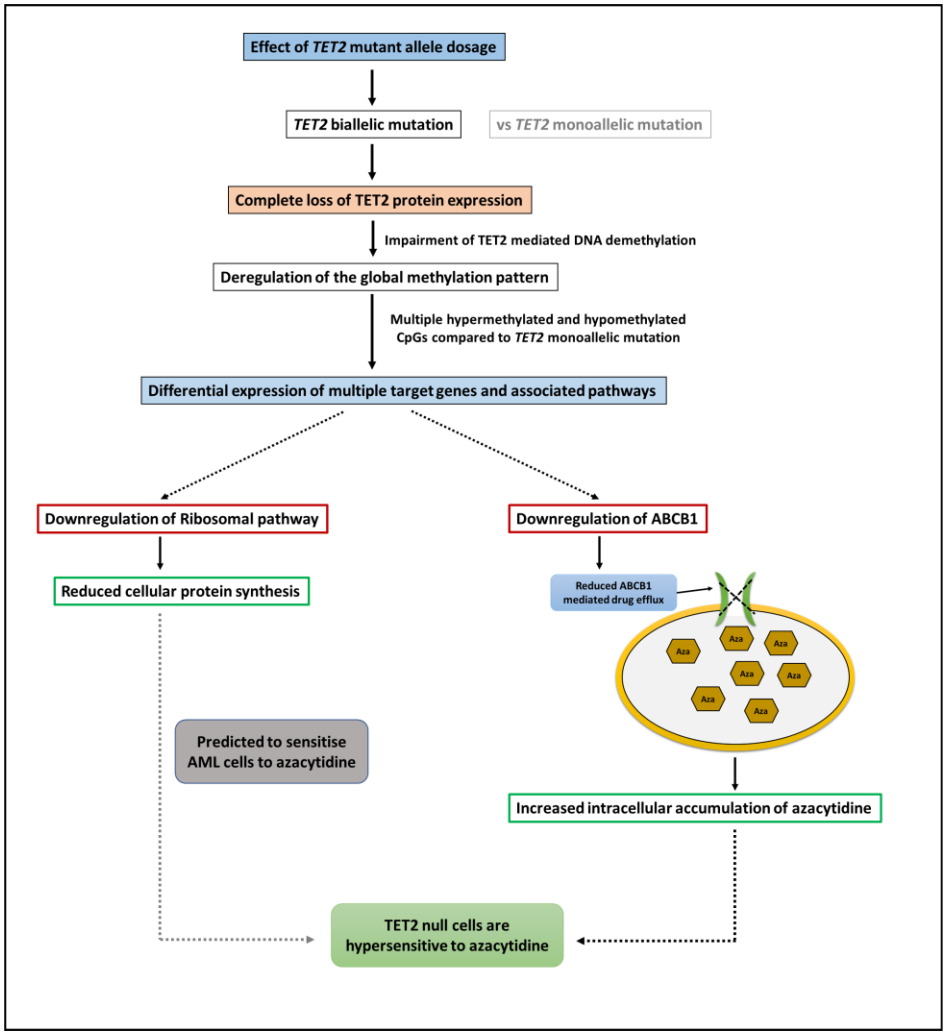

Figure 7.1. Summary of major findings of this project. The effect of *TET2* mutant allele dosage on differential response to AML therapeutics was investigated using isogenic *TET2* monoallelic and biallelic mutated HEL cells. Our data demonstrate that *TET2* biallelic mutated cells are hypersensitive to azacytidine. While investigating the mechanism behind the azacytidine sensitivity, it was noted that *TET2* biallelic mutation resulted in complete loss of TET2 protein expression in HEL cells, which led to deregulation of global methylation pattern and acquirement of a global hypermethylation phenotype, consistent with the demethylation function of TET2. This resulted in differential expression of multiple target genes and associated pathways. ABCB1, which is involved in azacytidine efflux is downregulated in *TET2* biallelic mutated cells compared to *TET2* monoallelic mutated counterparts, which likely resulted in an increased intracellular accumulation of azacytidine due to reduced ABCB1 mediated drug efflux. Our data demonstrates that downregulation of ABCB1 is an important mechanism behind hypersensitivity of *TET2* null cells to azacytidine. In addition, downregulation of ribosomal pathway which is already predicted to have a role in azacytidine sensitivity might be another mechanism behind the azacytidine sensitivity of *TET2* null cells.

Aza, azacytidine

As such, our data strongly argues in favour of considering *TET2* mutant allele dosage and TET2 protein expression for identifying patients who might benefit from azacytidine therapy. Finally, this study represents an important step in the development of precision medicine for patients with *TET2* mutated AML, who currently do not generally respond favourably to conventional chemotherapy.

7.6. Future directions

Our data opens up many possibilities for further investigations in this field, as follows:

7.6.1. Investigate the response of TET2 null cells to 2-hydroxyglutarate

Normally, IDH1/IDH2 (wild-type) catalyses the formation of α-KG, which acts as a co-factor necessary for TET2 enzyme activity (Figure 7.2A). Mutations in *IDH1/IDH2* results in the increased production of 2-HG, which acts as an inhibitor of TET2. Pagliuca reported synthetic lethality of *TET2* inhibition in *IDH1/IDH2* mutant cell lines, due to the phenotypic redundancy of *TET2* and *IDH1/IDH2* (Pagliuca et al. 2020). This highlights the relevance of testing whether the reverse also holds true – whether *IDH1/IDH2* loss-of-function is synthetically lethal in *TET2* null cells. To this end, we noted that *TET2* mutant allele dosage did not affect sensitivity to *IDH1/IDH2* inhibitors such as ivosidenib and enasidenib. This is likely because inhibition of IDH1/IDH2 are reported to reduce 2-HG levels whereas *IDH1/IDH2* mutation results in an increase in 2-HG (Liu et al. 2019) and therefore, inhibition of IDH1/IDH2 and *IDH1/IDH2* mutations are phenotypically different with respect to the levels of 2-HG (Figure 7.2B-C). As such, to completely phenocopy *IDH1/IDH2* mutations, it would be better to supplement the TET2 null cells directly with 2-HG which is predicted to result in synthetic lethality in TET2 null cells (Figure 7.2). Consistently, Guan and colleagues reported synthetic lethality in *TET2* deficient cells when treated with *IDH1/IDH2* mutant-derived-2HG in myeloid neoplasms. (Guan et al. 2021).

7.6.2. Investigate the impact of IDH1/IDH2/WT1 mutations on azacytidine response

In Chapter 6, it was noted that the *WT1*-mutated U937 cells were highly sensitive to azacytidine consistent with a significant correlation between TET2 null phenotype and azacytidine sensitivity. To this end, it would be interesting to test whether mutations in genes that are part of TET2 hydroxymethylation pathway such as *IDH1/IDH2/WT1* mutations that phenocopy *TET2* null cells also renders cells sensitivity to azacytidine. Future investigations based on

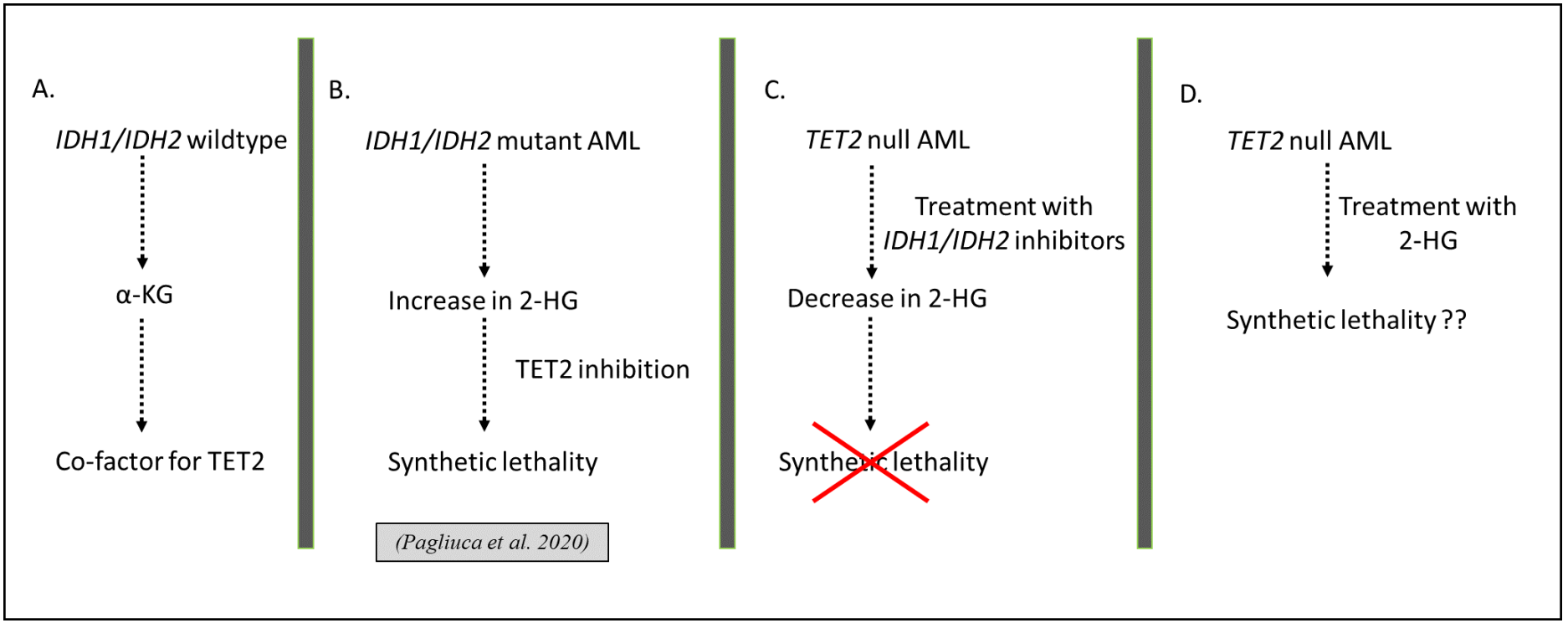

Figure 7.2. Synthetic lethality due to phenotypic redundancy of IDH1/IDH2 and TET2. A. In wild-type HSCs, IDH1/IDH2 catalyses the formation of α-KG, which is an essential co-factor of TET2 activity. B. Mutations in IDH1/IDH2 leads to the increased production of 2-HG (inhibitor of TET2) instead of α-KG. Further inhibition of TET2 in *IDH1/IDH2* mutant AML is reported to induce synthetic lethality due to phenotypic redundancy of IDH1/IDH2 and TET2 (Pagliuca et al. 2020). C. Treatment with IDH1/IDH2 inhibitors are reported to decrease the 2-HG levels (unlike IDH1/IDH2 mutations) and therefore did not induce synthetic lethality in TET2 null cells. D. To completely phenocopy *IDH1/IDH2* mutations it might better to treat the TET2 null cells directly with 2-HG and to test for synthetic lethality.

α-KG, α-Ketoglutarate; 2-HG, 2-Hydroxyglutarate

isogenic cell models with *IDH1/IDH2/WT1* monoallelic and biallelic mutations are warranted to understand the impact of these mutations on azacytidine response and whether azacytidine-based treatment regimen would prove beneficial in these AML patient groups.

7.6.3. Validation of the impact of TET2 mutant allele dosage on azacytidine sensitivity in other AML cell models or using xenograft models

The investigations reported in this thesis were carried out on a single AML cell line (HEL). As such, it is necessary to validate the results on an independent cell line and preferably a cell line derived from an AML with monoallelic *TET2* mutation. Although, SKM1 cells are *TET2* monoallelic mutated (Cluzeau et al. 2014) they are completely null for TET2 protein (Chapter 6) suggesting that the retained wild-type allele is silenced. As such, knocking out the retained wild-type allele is predicted to have no effect on TET2 protein expression. Given this, there is a need to identify another cell line with a monoallelic *TET2* mutation and use this as the basis to generate a second isogenic model system to validate the impact of *TET2* mutant allele dosage on azacytidine sensitivity. As an alternative approach, it is also possible to knock out *TET2* in an AML cell line that is wild-type for *TET2* such as THP1 or KG-1. This would generate a model that is not seen in AML patients and hence might not be very representative. Patient-derived xenograft models can also be exploited to validate the effect of *TET2* mutant allele dosage on azacytidine sensitivity. This model is expected to better recapitulate the mutational complexity of human AML.

7.6.4. Clinical trial to investigate the impact of TET2 mutant allele dosage on azacytidine sensitivity

This study has identified clinically relevant findings with the potential to transform the treatment outcome of *TET2* mutated AML, which is currently associated with poor outcome (Patel et al. 2012). As such, these findings will need to be validated in a clinical setting before the therapeutic strategy can be translated into routine clinical practice for the treatment of AML. Therefore, large clinical trials are warranted in order to identify and stratify AML patients of all age groups based on their *TET2* mutant allele dosage (and TET2 protein expression), and to determine their response to single-agent azacytidine. This study would also benefit from an analysis of *IDH1/IDH2/WT1* mutation status as a determinant of azacytidine response. Successful completion of a large-scale clinical trial has the potential to identify subgroups of patients who will benefit from azacytidine-based therapy.

Appendices

Appendix A: Sanger sequencing results of HEL parental cells and *TET2* monoallelic biallelic mutated HEL cell clones and SKM1 cells (continued in next page).

HEL Parental cells (5'-3')

GGGGGCAAACATGAAAAGAACTACTGTGTTTGGTGCGGGAGCGAGCTGGCCACACCTGT
GAGGCTGCAGTGATTGTGATTCTCATCCTGGTGTGGGAAGGAATCCCGCTGTCTCTGGCT
GACAAACTCTACTCGGAGCTTACCGAGACGCTGAGGAAATACGGCACGCTCACCAATCG
CCGGTGTGCCTTGAATGAAGAGTAAGTGAAGCCCAGGGCCTCTCCCCTCTTTGCGGCCAC
TGATAGGAAAGCCCAATCTTTGGTTGAAAGGAAGAGAGTTCAGCGTGCACTTTTACATTT
ATAAAATGGGCATCAAAATGCCTGTTTGGCAGTCATGCGATAAGAAGTTGTATTTGCTAA
TGTGAATAACTTGAGATGATTTCATTATCTGAATTGTACAGTTTAGCCATTAATTAGGAGC
AGTCAGAGTGTCTGTAACCACATGGCCTCAGTTATACCATAAACTTGAAATTGTTTATGT
GCTCACATGCTACAAGTGAACGGGCTCCTAAAAAAT

HELmono1 cells (5'-3')

GGGGGCAAACATGAAAAGAACTACTGTGTTTGGTGCGGGAGCGAGCTGGCCACACCTGT
GAGGCTGCAGTGATTGTGATTCTCATCCTGGTGTGGGAAGGAATCCCGCTGTCTCTGGCT
GACAAACTCTACTCGGAGCTTACCGAGACGCTGAGGAAATACGGCACGCTCACCAATCG
CCGGTGTGCCTTGAATGAAGAGTAAGTGAAGCCCAGGGCCTCTCCCCTCTTTGCGGCCAC
TGATAGGAAAGCCCAATCTTTGGTTGAAAGGAAGAGAGTTCAGCGTGCACTTTTACATTT
ATAAAATGGGCATCAAAATGCCTGTTTGGCAGTCATGCGATAAGAAGTTGTATTTGCTAA
TGTGAATAACTTGAGATGATTTCATTATCTGAATTGTACAGTTTAGCCATTAATTAGGAGC
AGTCAGAGTGTCTGTAACCACATGGCCTCAGTTATACCATAAACTTGAAATTGTTTATGT
GCTCACATGCTACATGAACCGGGCTCCTAAA

HELmono2 cells (5'-3')

GGCAAACATGAAAGAACTACTGTGTTTGGTGCGGGAGCGAGCTGGCCACACCTGTGAGG
CTGCAGTGATTGTGATTCTCATCCTGGTGTGGGAAGGAATCCCGCTGTCTCTGGCTGACA
AACTCTACTCGGAGCTTACCGAGACGCTGAGGAAATACGGCACGCTCACCAATCGCCGGT
GTGCCTTGAATGAAGAGTAAGTGAAGCCCAGGGCCTCTCCCCTCTTTGCGGCCACTGATA
GGAAAGCCCAATCTTTGGTTGAAAGGAAGAGAGTTCAGCGTGCACTTTTACATTTATAAA
ATGGGCATCAAAATGCCTGTTTGGCAGTCATGCGATAAGAAGTTGTATTTGCTAATGTGA
ATAACTTGAGATGATTTCATTATCTGAATTGTACAGTTTAGCCATTAATTAGGAGCAGTCA
GAGTGTCTGTAACCACATGGCCTCAGTTATACCATAAACTTGAAATTGTTTATGTGCTCAC
ATGCTACATGAAACGGGCTTCCTAAAA

HELmono3 cells (5'-3')

GGGGGCCAAACATGAAAAGAACTACTGTGTTTGGTGCGGGAGCGAGCTGGCCACACCTG
TGAGGCTGCAGTGATTGTGATTCTCATCCTGGTGTGGGAAGGAATCCCGCTGTCTCTGGC
TGACAAACTCTACTCGGAGCTTACCGAGACGCTGAGGAAATACGGCACGCTCACCAATC
GCCGGTGTGCCTTGAATGAAGAGTAAGTGAAGCCCAGGGCCTCTCCCCTCTTTGCGGCCA
CTGATAGGAAAGCCCAATCTTTGGTTGAAAGGAAGAGAGTTCAGCGTGCACTTTTACATT
TATAAAATGGGCATCAAAATGCCTGTTTGGCAGTCATGCGATAAGAAGTTGTATTTGCTA
ATGTGAATAACTTGAGATGATTTCATTATCTGAATTGTACAGTTTAGCCATTAATTAGGAG
CAGTCAGAGTGTCTGTAACCACATGGCCTCAGTTATACCATAAACTTGAAATTGTTTATGT
GCTCACATGCTACAAGTGACGGCTCCTACCA

Appendix A: Sanger sequencing results of HEL parental cells, *TET2* monoallelic biallelic mutated HEL cell clones and SKM1 cells (continued from previous page).

HELbi1 cells (5'-3')

GAGTCTACCGTGAGAGAGCTACTGTGTTTGGTGCGGGAGCTGGGACACCAGAGATGCTG
CAGTGAGTGGGATTCCCATCCTGTGGTGGGAAGGAGTCCGGCTGTCTCTGGCTGATGTGG
TCTACTCTCACCTTACCGATACGCTAAAGAATCAACGCACACTCAACAATCGCCCGTATG
CCTTGAATGAAGATCACCTGACGCCCAGGGCCTCTCCCCTCTTTGCAGCCACTGATAAGA
AACTCCAATCTTTGGTTGAAAGGAGAAAAGTTCACCGTGTCTTTTTACATTTATAAAATG
GGCATCAAAATGCTTGTTTGTTTATCATGCGATAATAAATTGGATTGGCTAATGTGAATA
ACTTGAGATGATTTCTTTATCTGAGTTGTAAACTTTAGACGATATCTATAACCAATCTGAA
TGTCTGAAACCTCATGTACTAACATACACAATGACCTTGAACTTGTGTATGTGCTCACATG
CTACAATTGACATTGCCTATGTGCTCACATGCTACAAGTGACGGCTCCTAA

HELbi2 cells (5'-3')

GCGTCTACGCGTGAGAGAGCTACTGTGTTTGGTGCGGGAGCTGGGAGCCTAGTGATGCTG
CAGTGAGTGGGAGTCCCATCCTGGGGTGCTAAGGAGTCCGGCTGTCTCTGCCTGACGTGC
TCTACTCTCACCTTACCTCTGCGCTAAAGAATCAACGCACGCTCTACCGTCGCCCGTATGC
CTTGAAGGACGATCACCTGACGCCCAGGTGCTCTCCCCTCTTTGCGGCCGATGATAAGGG
ACTCCAATCTCTGGTTGAAAGGAGAAAAGTTCGCCGTGTCTTTGTACATTTATAAAAAGA
GTTATCAAAATGCCTGTTTGTTTATCATGCGATAATAAATTGGATTGGCTAATGTGAATGA
CTTGAGATGATTTCTTTATCTGAGTTGAAAACTTTAGATGATAATTATAACCAATCTGAAT
GTCTGAAACCTCATGTACTAACATACACAATGACCTTGAACTTGTGTATGTGCTCACATG
CTACACTGAACAGCGCCTATGTGCTCACATGCTACAATGAACGGCTCCTA

HELbi3 cells (5'-3')

AACTACTGTGTTTGGTGCGGGAGCTGGCCACACCTGTGAGGCTGCAGTGATTGTGATTCT
CATCCTGGTGTGGRAAGGAATCCCGCTGTCTCTGGCTGACAAACTCTACTCGGAGCTTAC
MGAGACGCTGAGGAAATACGGCACGCTCACCAMTCGCCGGTGTGCCTTGAATGAAGAGT
ARRTGAAKCCCMSGGCCTCTCCCCTCTTTGCGGCCACTGATAGGAAASCCCAATCTTTGGT
TGAAAGGRAGAGAGTTCAGCGTGCACTTTTACATTTATAAAATGGGCATCRAAATGSCTG
TTTGGCAKYCATGCGAYAAKAAGTTGTATTTGCTAATGTGAATAACTTGAGATGATTTCW
TTATCTGAATTGTACAGTTTAGCCATTAATTAGGAGCAGTCAGAGTGTCTGTAACCACAT
GGCCTCAGTTATACCATAAACTTGAAAKTGKTTATGTGCTCACATGCKACAATGACCGGC
TCC

SKM1 cells (5'-3')

GGGGGCTAACATGAAAGAAGCTACTGTGTTTGGTGCGGGAGCGAGCTGGCCACACCTGT
GAGGCTGCAGTGATTGTGATTCTCATCCTGGTGTGGGAAGGAATCCCGCTGTCTCTGGCT
GACAAACTCTACTCGGAGCTTACCGAGACGCTGAGGAAATACGGCACGCTCACCAATCG
CCGGTGTGCCTTGAATGAAGAGTAAGTGAAGCCCAGGGCCTCTCCCCTCTTTGCGGCCAC
TGATAGGAAAGCCCAATCTTTGGTTGAAAGGAAGAGAGTTCAGCGTGCACTTTTACATTT
ATAAAATGGGCATCAAAATGCCTGTTTGGCAGTCATGCGATAAGAAGTTGTATTTGCTAA
TGTGAATAACTTGAGATGATTTCATTATCTGAATTGTACAGTTTAGCCATTAATTAGGAGC
AGTCAGAGTGTCTGTAACCACATGGCCTCAGTTATACCATAAACTTGAAATTGTTTATGT
GCTCACATGCTACATGAACCGGGCTCCTAAAAAAT

Appendix B: Significantly differentially expressed genes (Padj < 0.05 and |Log2FC| ≥ 0.3) in HEL cell clones with biallelic TET2 mutations compared to parental clones with monoallelic TET2 mutation identified (continued in next page).

Gene ID	P_{adj}	log2FC		Gene ID	P_{adj}	log2FC
LIMCH1	6.31E-09	-9.8468		C3	0.0132914	-2.4344
EPHA7	2.12E-15	-8.4818		B4GALNT3	0.011724	-2.4206
LOC100507377	0.0214126	-7.9974		ITGAL	0.0171647	-2.3851
POMC	9.69E-11	-7.9757		ARRDC4	2.31E-12	-2.3805
ALDH1L1	0.0011572	-6.7135		ARHGAP20	0.0336812	-2.3582
GABRB2	0.0324817	-6.5139		OSM	0.00069	-2.3320
TMPRSS4	0.0254611	-6.4465		ACSL1	4.71E-10	-2.3063
PLA2G3	0.0302844	-5.7966		CABP4	0.040943	-2.2866
LONRF2	0.0179756	-5.7897		DHDH	1.05E-05	-2.2606
HLA-DRA	0.0222096	-5.7454		TMEM154	0.0006301	-2.2456
KIAA0930	0.0338257	-5.5195		RNU5D-1	0.0038921	-2.2453
CETP	1.59E-06	-5.1217		IL12RB1	0.002749	-2.2132
ASB2	0.0028428	-5.0083		CACNA1D	0.0396629	-2.1849
MUC19	1.00E-06	-4.0412		RPL3L	0.04242	-2.1834
FCGR1A	0.021559	-3.6180		WNT11	0.0083745	-2.1782
GABRA6	0.006725	-3.5104		DIRAS1	0.0434361	-2.1734
EFR3B	0.0241546	-3.4940		FOS	0.0073787	-2.1705
DUSP2	0.0404843	-3.4552		CXCR4	8.89E-10	-2.1689
EDDM13	0.0015062	-3.4135		CD69	1.97E-05	-2.1354
BCL2	0.0391088	-3.3898		ADAMTS4	0.0307332	-2.1334
KLHDC7B	0.0091497	-3.3536		SLC9A3	0.04184	-2.1119
CXCR5	0.0043528	-3.3379		GADD45B	0.001505	-2.0935
CXCL8	0.0429074	-3.3220		TMEM40	0.0067087	-2.0735
RAET1E	0.0038476	-3.3139		LINC01033	0.0064317	-2.0724
UBBP4	0.0203554	-3.2134		IL18RAP	0.0086981	-2.0543
GLDC	0.0263893	-3.1704		FRGCA	0.0190321	-2.0538
RNVU1-7	0.0089546	-3.1414		ITGB2	0.0043281	-2.0512
PDGFB	0.0004895	-3.0379		IL1RL1	0.0330801	-1.9505
CLEC9A	0.0025758	-2.9795		TNFRSF4	0.0108378	-1.8938
SLC13A3	0.0163214	-2.9602		RNVU1-6	0.017224	-1.8913
TPSB2	0.0044395	-2.9566		RYR2	0.0003848	-1.8732
SEPTIN1	0.0008786	-2.9490		CCDC88B	0.0408906	-1.8518
RNU5F-1	0.0006953	-2.8908		B4GALNT1	0.0426431	-1.8404
LRRC8E	0.0061218	-2.8722		FSCN1	0.0171647	-1.8278
CD80	0.0283118	-2.8457		TMOD1	0.0206749	-1.7927
CD52	0.0009906	-2.6970		PTGER1	0.0097536	-1.7808
CKM	0.0012109	-2.6833		FCHO1	0.0033996	-1.7805
SLC5A10	0.0419417	-2.6814		PRG2	0.0206749	-1.7673
LINC01270	0.0477769	-2.6720		SLC7A7	0.0307755	-1.7669
B3GNT5	0.0131813	-2.6358		RNU1-1	0.0307332	-1.7635
FGR	0.0330326	-2.6336		P2RX7	0.0408514	-1.7396
MIR3150BHG	0.0470821	-2.5706		PRRG4	1.59E-06	-1.7013
PTPRR	0.0007792	-2.4896		FOSB	0.0497518	-1.6677
LTF	0.0338257	-2.4859		UPB1	0.0353189	-1.6517
DYSF	0.0336401	-2.4705		KLF11	0.0011353	-1.6442

Appendix B: Significantly differentially expressed genes (P_{adj} < 0.05 and |Log2FC| ≥ 0.3) in HEL cell clones with biallelic *TET2* mutations compared to parental clones with monoallelic *TET2* mutation identified (continued in next page).

Gene ID	P_{adj}	log2FC		Gene ID	P_{adj}	log2FC
HSD3B7	0.0424008	-1.6220		RNU5A-1	1.03E-08	-1.0024
AP1S2	2.75E-06	-1.5882		ST8SIA4	0.0084267	-0.9903
LGMN	0.0278768	-1.5731		SNORA73B	5.61E-05	-0.9863
NCR3	0.0207176	-1.5730		RYR3	3.59E-05	-0.9841
RNU4ATAC	0.0158496	-1.5541		TMEM107	0.0004199	-0.9594
SNORD3B-2	0.0403285	-1.5517		HLA-B	0.0073229	-0.9570
LAT2	0.0008436	-1.5149		KLF10	0.0214524	-0.9416
LOC100506544	0.040943	-1.5112		TMED7	0.0046881	-0.9378
TNFRSF25	0.0401502	-1.4920		CNKSR2	0.0141949	-0.9335
ZG16B	0.0067465	-1.4628		MICAL1	0.0456586	-0.9283
ANTXR2	2.55E-05	-1.4561		IDH1	0.0063113	-0.9255
APOL1	0.0002456	-1.4195		CDC6	0.0030937	-0.9239
TMEM150B	0.0162304	-1.4084		VAT1	0.0022107	-0.9176
ARHGAP45	0.0099369	-1.3915		RAB37	0.0052704	-0.9096
PDGFD	0.0320038	-1.3600		SYPL1	0.0283015	-0.9088
IL1RAP	0.0004199	-1.3527		EEF1E1	0.0083461	-0.9040
RNVU1-18	0.0394566	-1.3521		RHBDD2	0.0073787	-0.8969
ID2	3.60E-05	-1.3483		LGALS1	0.0296427	-0.8916
ALDH2	0.0002171	-1.3406		SLC35E1	0.0313809	-0.8895
NCALD	0.0061271	-1.3396		APOBEC3B	0.0290631	-0.8817
KCNAB2	1.04E-05	-1.3148		MRAP2	0.0413116	-0.8688
ATP8B4	0.021559	-1.3066		ERP29	0.0171647	-0.8626
RHOB	0.0045946	-1.2795		ITPRIP	0.0010148	-0.8619
GNS	4.12E-16	-1.2550		DENND2C	0.0246773	-0.8615
PTGS2	0.0024814	-1.2535		ARHGAP15	0.0214126	-0.8611
BSDC1	0.0454163	-1.2522		SRGN	0.017842	-0.8573
SMIM30	0.0170448	-1.2433		OGFRL1	0.0313184	-0.8495
DPEP2	0.0342147	-1.2413		HERPUD1	0.0367631	-0.8388
PARVG	0.043576	-1.2061		TWSG1	7.93E-05	-0.8353
TATDN3	0.0273097	-1.2052		PSMB8-AS1	0.0108731	-0.8327
DUSP1	0.0099656	-1.2043		ASNS	0.0003234	-0.8165
FAM83G	0.0163214	-1.1991		SNORA53	0.0091518	-0.8137
DUSP6	6.04E-07	-1.1748		P2RX1	5.20E-05	-0.8117
DNAJB9	0.0006468	-1.1456		IFRD1	1.42E-05	-0.8068
DERA	1.88E-05	-1.1452		ESYT2	8.45E-06	-0.8001
PRSS57	0.0273771	-1.1158		HLA-C	0.0111237	-0.7985
CALB2	0.0131813	-1.0960		TFEC	0.0243107	-0.7979
SDSL	0.0463286	-1.0912		SNHG32	0.0451561	-0.7939
HTATIP2	0.0171647	-1.0883		NAB2	0.0327336	-0.7936
EPS8L1	0.0377624	-1.0824		SPX	0.0010048	-0.7930
SNORA74B	0.0381615	-1.0711		SEL1L3	0.037515	-0.7922
IKZF2	0.0045946	-1.0646		TTC39B	0.0037975	-0.7891
CTH	8.21E-07	-1.0461		TSLP	0.0347729	-0.7871
GPR85	2.06E-05	-1.0115		STK17A	0.0030479	-0.7859
ZNF22	0.0111237	-1.0114		CHAC2	0.0021827	-0.7825

Appendix B: Significantly differentially expressed genes (P_{adj} < 0.05 and |Log_2FC| ≥ 0.3) in HEL cell clones with biallelic *TET2* mutations compared to parental clones with monoallelic *TET2* mutation identified (continued in next page).

Gene ID	P_{adj}	log2FC		Gene ID	P_{adj}	log2FC
PDE4D	0.0003776	-0.7807		KIAA1324L	0.0041503	-0.6466
PIGK	0.0176479	-0.7769		ASF1A	0.0162006	-0.6329
XBP1	0.0464647	-0.7724		SSBP2	0.0020614	-0.6324
PMAIP1	0.031282	-0.7719		MCL1	0.0449597	-0.6313
CYCS	0.0201477	-0.7714		GRAMD2B	0.043576	-0.6262
MIS12	0.0407952	-0.7701		NXT2	0.0424008	-0.6136
CTBS	0.0011594	-0.7696		FIS1	0.0136999	-0.6113
SPCS3	0.0003236	-0.7669		TM6SF1	1.54E-05	-0.6072
TMSB4X	0.0022107	-0.7627		ACTR3	0.0122162	-0.6064
H3F3B	0.0273097	-0.7592		CAP1	0.0199543	-0.6011
LINC01814	0.0225029	-0.7543		STK38L	0.0102836	-0.5993
BTG1	0.0190794	-0.7541		H2AFZ	0.0388841	-0.5986
SNORA26	0.0392106	-0.7514		PPIP5K2	0.0013212	-0.5985
RNY1	0.0398187	-0.7508		RPS6KA5	0.0252092	-0.5943
RNU4-1	0.0245154	-0.7378		SNORA74A	0.03426	-0.5910
IGF2BP3	0.0029183	-0.7310		RNU5B-1	0.0002992	-0.5900
PRUNE2	0.0211409	-0.7303		HSPA5	0.0464834	-0.5900
RNU6ATAC	0.014186	-0.7288		PNPLA8	0.0407952	-0.5887
TMEM173	0.0449911	-0.7276		ZNF266	0.0220295	-0.5867
TMPO	0.0346387	-0.7240		NBAS	0.0055466	-0.5819
CNST	0.0338257	-0.7213		CHMP3	0.0381357	-0.5812
TRAF3IP2	0.0338625	-0.7204		NR3C1	0.0416743	-0.5810
DOCK2	0.0081649	-0.7021		VPS29	0.0005133	-0.5763
MCM7	0.000223	-0.7009		TWF1	0.0136095	-0.5740
LSM8	0.0011572	-0.7004		NEK7	0.0479261	-0.5736
RNU4-2	0.0045946	-0.6956		GPR158	0.027657	-0.5700
ZNF277	0.0023607	-0.6935		ARHGEF6	0.0477456	-0.5691
SAT1	0.0449597	-0.6928		PPP1R35	0.021559	-0.5667
ING3	0.0243651	-0.6904		CDK2	0.0228704	-0.5628
ADCK2	0.0283176	-0.6903		RPL22	0.000104	-0.5617
GLUL	0.0250118	-0.6867		HIGD1A	0.0334269	-0.5611
RBMS1	0.0014022	-0.6852		VRK1	0.037515	-0.5589
PSMC2	0.0081521	-0.6791		SRSF7	0.021559	-0.5588
CALU	4.18E-05	-0.6791		NARS	0.0416743	-0.5583
SBDSP1	0.03426	-0.6737		SLC38A2	0.0185449	-0.5565
UBE2J1	0.0394566	-0.6680		TMEM38B	0.0055641	-0.5548
SBDS	0.014186	-0.6630		ITGB1	0.0254611	-0.5505
SPIRE1	0.014186	-0.6625		CDKN2AIP	0.0204293	-0.5490
ABCB1	0.0008096	-0.6589		CDC42	0.0456224	-0.5466
BROX	0.0188823	-0.6531		CEMIP2	0.007294	-0.5426
ARFGAP3	0.0490086	-0.6528		UGCG	0.0120669	-0.5423
HSPA13	0.0345871	-0.6525		TMED5	0.0037232	-0.5411
COL4A6	0.0307332	-0.6518		CAPZA1	0.0011806	-0.5405
PYROXD1	0.0295009	-0.6494		DCLRE1A	0.0028259	-0.5391

Appendix B: Significantly differentially expressed genes (P_{adj} < 0.05 and |Log2FC| ≥ 0.3) in HEL cell clones with biallelic *TET2* mutations compared to parental clones with monoallelic *TET2* mutation identified (continued in next page).

Gene ID	P_{adj}	log2FC		Gene ID	P_{adj}	log2FC
CDADC1	0.0471611	-0.5363		CRY1	0.0383284	-0.4371
VAMP7	0.010518	-0.5331		TMEM41B	0.0401502	-0.4351
GSTK1	0.0319484	-0.5304		HNRNPH1	0.0033224	-0.4301
SCYL2	0.0011346	-0.5296		SMC2	0.0404843	-0.4299
MANEA	0.0080397	-0.5285		ARL6IP1	0.021559	-0.4244
RSL24D1	0.0259861	-0.5283		MTF2	0.0113783	-0.4223
C1D	0.022754	-0.5254		ACSL3	0.0319484	-0.4190
CALCOCO2	0.0466057	-0.5216		HAT1	0.0277667	-0.4144
SNORA38	0.0444447	-0.5189		ACAD8	0.0492494	-0.4138
RRM2B	0.037246	-0.5142		SLC39A10	0.0417604	-0.4133
ORAI2	0.0097847	-0.5139		SLK	0.0491242	-0.4044
TP53RK	0.0203554	-0.5138		ANXA7	0.0339977	-0.3866
ZCCHC10	0.0307952	-0.5137		ZNF562	0.0416743	-0.3840
CLINT1	0.0052282	-0.5135		BCAT2	0.037515	0.3795
MRPL32	0.0131747	-0.5119		ZBTB39	0.0355799	0.4113
ABI1	0.0347791	-0.5090		FAM168A	0.0165351	0.4377
MTHFD2	0.0170448	-0.5085		SERPINH1	0.0405748	0.4589
PDIA4	0.0122162	-0.5060		BSG	0.0190794	0.4609
GGCT	0.0131813	-0.5030		SSH2	0.0160116	0.4653
KIAA1143	0.0167563	-0.5022		URB1	0.0486285	0.4731
MTURN	0.0329497	-0.5006		DIS3L2	0.0220592	0.4839
ARF5	0.0078284	-0.4980		MAML1	0.0298407	0.4842
WSB1	0.0408514	-0.4957		DHX34	0.0327494	0.4875
CHPT1	0.0039025	-0.4889		ZNF574	0.0438398	0.4897
SNORA2C	0.0478033	-0.4843		GRWD1	0.008837	0.4943
SUMF2	0.0061178	-0.4833		PLCB2	0.0111078	0.4949
WDR36	0.0065331	-0.4793		ETV6	0.0102305	0.4988
ERAP1	0.0054274	-0.4788		DAP	0.026751	0.5030
MYCBP	0.0305968	-0.4775		SMG9	0.0231993	0.5055
MLH1	0.0073787	-0.4768		POLR3H	0.0494573	0.5066
TBK1	0.0081278	-0.4731		DLG4	0.0416743	0.5167
RFC2	0.0377813	-0.4695		ATF7IP2	0.0416743	0.5188
SRSF10	0.0239416	-0.4685		LTB4R	0.0451566	0.5209
ASL	0.0283379	-0.4638		PDE1B	0.0025748	0.5333
VAV3	0.0300691	-0.4622		PPARGC1B	0.0401365	0.5333
SND1	0.0075503	-0.4599		RIPOR2	0.0426431	0.5469
CHPF2	0.0163011	-0.4555		INTS3	0.0088464	0.5500
ATP6V0A2	0.0228964	-0.4553		VPS37B	0.0394981	0.5533
NAPG	0.0356801	-0.4469		PDXDC1	0.0469199	0.5557
MKRN1	0.0355707	-0.4422		BICD1	0.0481645	0.5572
ELOVL5	0.0167698	-0.4414		ZNF589	0.0464647	0.5595
FAM76B	0.0311155	-0.4405		ZNF749	0.0080397	0.5739
CYREN	0.0394981	-0.4379		ABL2	0.0340878	0.5809
CCT6A	0.0057997	-0.4377		C20orf27	0.0376705	0.5827

Appendix B: Significantly differentially expressed genes (P_{adj} < 0.05 and |Log_2FC| ≥ 0.3) in HEL cell clones with biallelic *TET2* mutations compared to parental clones with monoallelic *TET2* mutation identified (continued in next page).

Gene ID	P_{adj}	log2FC		Gene ID	P_{adj}	log2FC
CABIN1	0.0231993	0.5836		ZNF696	0.0148909	0.7419
TMPPE	0.0225658	0.5873		TNS1	0.027657	0.7457
AK4	0.022754	0.5989		ANKDD1A	0.0497518	0.7458
SERPING1	0.0108731	0.5996		ADAMTSL4	0.0220321	0.7480
RDH13	0.0250118	0.6021		COTL1	0.0062792	0.7511
MRTFA	0.0394566	0.6021		FARSA	0.0024814	0.7515
PSMG4	0.0065236	0.6023		ABCC1	0.0471373	0.7526
CPNE1	0.0352562	0.6092		LENG8	0.022815	0.7552
MED12	0.0001972	0.6093		TET3	0.033638	0.7553
MAVS	0.0393189	0.6100		ENG	0.0355707	0.7566
GP6	0.0317255	0.6128		CAVIN1	0.0002731	0.7583
CEP164	0.033638	0.6162		PEAR1	2.06E-05	0.7607
ZXDC	0.0111237	0.6210		ANGEL1	0.0026738	0.7679
CCDC69	0.033565	0.6336		PTPN14	0.0001385	0.7742
WDR4	0.0111237	0.6339		DOK6	0.0007792	0.7748
MECP2	0.0449597	0.6373		LINC01089	0.0228704	0.7765
MEF2D	0.0313184	0.6431		FGD5	0.0381357	0.7796
ILF3-DT	0.0437856	0.6437		CIC	0.0457126	0.7902
FAM78A	0.0248673	0.6468		ZMYM3	0.0265514	0.7929
TDRD3	0.0279602	0.6488		DDHD1	0.0058708	0.7946
INKA2	0.0228704	0.6520		EGLN1	0.0408514	0.7951
NBEAL2	0.0453506	0.6521		RHCE	0.0194573	0.7960
AGPAT4	0.0108731	0.6581		PRRC2A	0.037246	0.7969
SPHK2	0.0063113	0.6617		RIOX2	0.010518	0.8021
ADCY7	0.0407952	0.6694		PROSER2	0.0033961	0.8043
TRIM16	0.0191504	0.6743		ABHD2	0.0158168	0.8046
PNKD	0.0318547	0.6770		ST3GAL2	0.0081278	0.8085
CDC42BPB	0.0320038	0.6804		ICAM5	0.0032453	0.8138
ZNF592	0.0407952	0.6819		OPHN1	0.014186	0.8162
MPL	0.0163517	0.6882		POLR2A	0.0496329	0.8179
PTOV1	0.0093674	0.6915		AGO1	0.0338257	0.8259
GABRE	0.0183977	0.6924		FAM171A1	0.0449597	0.8281
NRROS	0.0005695	0.6925		FAXDC2	0.040943	0.8294
MYLK	0.0095849	0.6941		GYS1	0.0451177	0.8316
DIDO1	0.02278	0.6966		ONECUT2	0.0407952	0.8319
TMPRSS3	0.0424008	0.6997		LRRC28	6.38E-05	0.8336
XXYLT1	0.0307332	0.7032		TMEM63B	0.0381357	0.8344
EHD3	0.01226	0.7057		CHST2	5.96E-06	0.8408
IGF2BP1	0.000873	0.7069		NKAPP1	0.0348494	0.8420
ACP6	0.0006301	0.7184		RUBCN	0.0426875	0.8486
TNK2	0.0179756	0.7230		TSPAN14	0.0377813	0.8494
ZYX	0.0334269	0.7245		SIK3	0.0091497	0.8522
PDE4A	0.0206194	0.7300		PLAGL2	0.0171647	0.8618
KCNC4	0.0321911	0.7323		SUFU	0.0122128	0.8650

Appendix B: Significantly differentially expressed genes (P_{adj} < 0.05 and |Log$_2$FC| ≥ 0.3) in HEL cell clones with biallelic *TET2* mutations compared to parental clones with monoallelic *TET2* mutation identified (continued in next page).

Gene ID	P_{adj}	log2FC		Gene ID	P_{adj}	log2FC
RBM38	0.0176479	0.8651		GPR173	0.0415703	1.0550
CFAP44	0.0030056	0.8663		TGFB1	0.0283015	1.0604
DOT1L	0.0263135	0.8705		SEMA4B	0.0417604	1.0630
SMARCAL1	0.0058899	0.8749		BCORL1	0.0214524	1.0670
D2HGDH	0.0243107	0.8815		TSPOAP1-AS1	1.03E-07	1.0685
STX18-AS1	0.0319484	0.8858		IKZF4	0.0144898	1.0691
FBRS	0.0009114	0.8908		ST3GAL4	0.0338257	1.0760
RUNX1	0.0469199	0.9015		ATP11A	0.0254611	1.0760
MFHAS1	0.0449597	0.9069		KIAA0513	2.20E-06	1.1000
SEMA6C	0.0214126	0.9084		JUP	1.52E-05	1.1055
IGSF10	1.09E-05	0.9124		SELENBP1	0.0298284	1.1056
MED25	0.0283379	0.9291		DSC2	2.93E-06	1.1057
BCL9	0.0250118	0.9302		KCNJ14	0.0245154	1.1081
FAM189B	0.0171647	0.9379		B3GNT7	0.0401633	1.1142
ASIC4	0.0348494	0.9452		ADAM20	0.0265441	1.1178
TNRC6C	0.0051177	0.9465		SLC25A30	1.18E-07	1.1186
HCFC1	0.0381357	0.9467		ZHX2	0.0024104	1.1216
SLC27A3	0.0032735	0.9497		EXOC6B	0.0332642	1.1272
TRIM58	0.0464647	0.9561		APBB1IP	0.0401502	1.1327
POGK	1.17E-05	0.9580		FKBP11	0.0032379	1.1466
HSPA1B	0.0003605	0.9682		GRIK4	0.0308186	1.1480
ARHGAP42	0.0045946	0.9747		KLHL31	0.0086981	1.1563
ZMIZ1	0.0085611	0.9748		THEMIS2	5.95E-05	1.1579
STX3	0.0019807	0.9771		ATP7B	0.0193175	1.1631
SRCAP	0.0398315	0.9842		SOGA1	2.64E-07	1.1647
CYP2S1	0.0063113	0.9868		SUGCT	0.0111237	1.1705
MBD6	0.0433919	0.9870		GPR135	0.0176479	1.1738
LRRN2	0.0498196	0.9897		PAQR8	0.0099369	1.1839
CCDC85C	0.0214202	0.9939		NHSL1	0.0017016	1.1914
FASN	0.0108731	0.9980		KMT2D	0.0063113	1.1943
IGSF3	5.59E-05	0.9991		PKIG	0.0102631	1.1985
PRKACA	0.0328737	1.0130		BCL9L	0.0484534	1.2166
HDAC7	0.0171647	1.0131		SSH1	0.0136219	1.2415
CLDN10	0.0417604	1.0139		YPEL2	2.69E-08	1.2461
EMCN	0.0104464	1.0184		TTC8	0.0338257	1.2509
SOX12	0.0263893	1.0193		NFIX	0.0111237	1.2589
GP1BA	0.0344666	1.0236		ABCB6	0.0214202	1.2590
DANT2	0.0355707	1.0285		VASH1	1.17E-05	1.2606
EPHX2	0.0277541	1.0309		LYST	0.0137311	1.2662
SMAD6	2.26E-05	1.0328		THBS4	0.0203315	1.2682
SHANK3	0.0021357	1.0362		IQSEC1	0.005962	1.2798
SYNGAP1	0.021559	1.0429		LIPC	0.0006301	1.2844
SAMD4B	0.0190794	1.0500		KDM6B	0.0167101	1.2991
CD84	0.000238	1.0500		CACNB4	0.00999	1.3031

Appendix B: Significantly differentially expressed genes (P_{adj} < 0.05 and |Log_2FC| ≥ 0.3) in HEL cell clones with biallelic *TET2* mutations compared to parental clones with monoallelic *TET2* mutation identified (continued in next page).

Gene ID	P_{adj}	log2FC		Gene ID	P_{adj}	log2FC
BCL7A	0.0048025	1.3113		PTK7	0.0003605	1.7490
ZFPM1	0.0188069	1.3351		CRYBG3	2.13E-05	1.7590
ARHGEF11	4.35E-09	1.3369		WDR49	0.0091213	1.7630
RNF139-AS1	0.0189259	1.3390		FAM131B	0.003285	1.7928
DUSP4	0.0001999	1.3539		ALS2CL	0.0014022	1.7979
FAT4	0.0329902	1.3555		NPHP1	0.0204293	1.8311
C2CD2	6.27E-11	1.3643		HOXB13	0.0007792	1.8324
HOMER2	0.0270251	1.3659		RBMS2	0.0010401	1.8568
CR1L	0.0139747	1.3671		SOX4	0.0165148	1.8590
APOE	1.54E-06	1.3842		PPM1L	0.0204293	1.9069
TUBB4A	0.0003425	1.3847		FOXL2NB	4.67E-06	1.9124
TMEM108	0.006715	1.3956		WNT5A	1.72E-05	1.9210
ZFHX2	0.0075135	1.3973		MAGI1-IT1	0.0321911	1.9546
GVQW3	0.0338257	1.4008		AKAP12	0.000104	1.9628
RNF24	0.0005357	1.4013		TANC2	1.59E-06	1.9804
LY6G6F	0.0321911	1.4102		ADCY4	0.0068753	1.9808
FPR3	0.0006836	1.4120		HAGLR	0.0394566	1.9882
PM20D2	0.0180861	1.4178		IGFBP4	1.98E-12	1.9945
VANGL2	8.49E-06	1.4185		NCAM2	0.0024163	2.0151
ARHGAP12	0.0077582	1.4203		NAV1	0.0009114	2.0202
ANTXR1	0.0090572	1.4415		GATM	0.0291534	2.0536
ADAM11	0.0089995	1.4594		ASIC1	2.01E-08	2.0668
ABCG2	0.0313613	1.4671		FRAS1	1.84E-12	2.0753
SPTBN4	0.0358628	1.4833		MMP24	0.0414246	2.0893
ZNF385A	0.0038882	1.4900		RAP1GAP2	3.63E-17	2.0950
SCARB2	0.0208077	1.5165		KREMEN1	0.001161	2.1108
SLCO2B1	0.0398187	1.5492		ZNF790-AS1	0.0180474	2.1671
SNPH	0.0261493	1.5715		KCNN3	3.41E-07	2.1785
RIPOR1	3.61E-05	1.5733		TBX2-AS1	0.0158168	2.1791
KIAA1522	2.75E-13	1.5874		GBP2	0.0214524	2.2239
LINC01133	0.0217647	1.5884		PLPP3	3.43E-06	2.2284
RHBDL3	0.0010473	1.6024		TRIM9	0.0073787	2.3038
NINL	0.0199543	1.6079		SELENOP	1.23E-05	2.3320
BNC2	0.0004188	1.6155		KCNQ4	0.0004632	2.3325
MEX3A	0.0006652	1.6183		ITGA8	0.035275	2.3334
ALPK3	0.0003425	1.6264		TBX2	0.0147582	2.3475
IFNGR2	1.05E-05	1.6530		LAMA5	2.23E-05	2.3632
NTN4	0.0383456	1.6784		IRF6	0.005676	2.3684
NLGN3	0.0180801	1.6935		SMAD3	0.0172344	2.3696
FABP6	0.01623	1.7078		GRID1	3.80E-06	2.4004
PDE4B	0.0214126	1.7175		PRSS12	0.0345871	2.4076
ZNF572	0.0021428	1.7254		PCDH7	0.0407952	2.4160
TGFB2	0.0171647	1.7358		PPP2R3A	0.0212543	2.4470
PROX1	2.26E-05	1.7439		ACE	0.033565	2.4535

Appendix B: Significantly differentially expressed genes (P_{adj} < 0.05 and |Log$_2$FC| ≥ 0.3) in HEL cell clones with biallelic *TET2* mutations compared to parental clones with monoallelic *TET2* mutation identified (continued from previous page).

Gene ID	P_{adj}	log2FC		Gene ID	P_{adj}	log2FC
EPB41L4A	0.0048817	2.4782		FOXO1	1.01E-06	4.1565
IGFBP5	8.68E-10	2.4958		SAMD4A	0.0093649	4.1659
CCNO	0.0013494	2.5159		DCT	8.86E-11	4.1749
GLP1R	0.0433755	2.5259		WDR88	0.0211532	4.1985
GLIS3	0.0006036	2.5568		FLG-AS1	0.0147582	4.2103
ANO2	0.0107628	2.5603		AKT3	8.86E-11	4.2365
VWF	7.55E-06	2.5605		EPHB1	0.0011346	4.3259
HSPA6	3.55E-05	2.5763		CCDC50	2.01E-07	4.3534
UNC5B	0.0078668	2.6003		MOBP	0.0171261	4.3552
DENND5B	5.34E-07	2.6146		SEMA3E	0.024977	4.5860
ESYT3	0.0002752	2.6487		TRMT9B	0.0196939	4.8280
SIGLEC11	0.0006079	2.6660		ADGRL3	0.0003871	4.9549
ZNF385B	0.0338257	2.7510		RNF180	2.22E-09	4.9912
RNASE1	3.74E-06	2.7588		TENT5A	0.0415703	5.1065
TMEM132E	7.55E-06	2.7765		PAGE1	0.0419417	5.1466
CSF1	2.22E-09	2.7977		ADGRB3	5.47E-07	5.1475
NELL2	8.94E-12	2.8196		SCD5	0.0005695	5.1565
TMEM132A	5.39E-05	2.8203		ZNF704	0.0035634	5.7404
TMEM150C	3.46E-05	2.8700		LOC101928307	4.58E-06	5.9022
SHROOM4	0.0045336	2.8877		ARFGEF3	1.26E-31	6.6169
MGAT4A	0.0259753	3.0270		MARK1	2.43E-05	6.7876
HBA2	0.0459902	3.0341		ZEB1	3.30E-55	6.7877
ZFHX3	0.0001242	3.0533		LRATD1	0.0206199	7.0385
FRMD6	0.0001615	3.1504		CNTN1	8.89E-10	7.0644
ARMC4	0.0001994	3.1532		LRIG1	8.21E-07	7.3216
KCTD12	0.0060617	3.1914		SH2D4A	0.0081521	7.6960
MAP1B	1.06E-05	3.2174		ZEB1-AS1	1.33E-11	7.8430
ZNF568	0.0062792	3.3173		MYO3A	0.0062792	8.3124
ANXA6	0.0032992	3.3314		POU6F2	0.0013974	8.8588
SIGLEC16	3.37E-05	3.3502		ATRNL1	3.48E-30	9.3264
DSP	1.83E-07	3.4430				
LINC01470	0.0048025	3.4830				
SAMD12	9.83E-05	3.5283				
ZNF829	0.0070635	3.5349				
EIF4E3	0.0204293	3.5564				
EPHB4	2.98E-14	3.6207				
PLPP1	0.0006301	3.7286				
MPP7	3.43E-06	3.7419				
KCNS3	0.0024163	3.9531				
PTPRF	3.40E-07	3.9561				
FSTL1	0.0081521	3.9782				
BCHE	3.15E-06	4.0111				
SATB1	0.006725	4.0405				
PERP	0.000342	4.1314				

Appendix C: Gene ontology component analysis of significantly differentially expressed genes in HEL cell clones with biallelic TET2 mutations compared to parental clones with monoallelic TET2 mutation (continued in next page).

ID	Description	P_{adj}
GO:0000244	spliceosomal tri-snRNP complex assembly	0.0003564
GO:0000353	formation of quadruple SL/U4/U5/U6 snRNP	0.0014256
GO:0000365	mRNA trans splicing, via spliceosome	0.0014256
GO:0045291	mRNA trans splicing, SL addition	0.0014256
GO:0001503	ossification	0.0014886
GO:0003279	cardiac septum development	0.005555
GO:0003281	ventricular septum development	0.0058147
GO:0003231	cardiac ventricle development	0.0058147
GO:0060411	cardiac septum morphogenesis	0.0058147
GO:0060412	ventricular septum morphogenesis	0.0083832
GO:0060326	cell chemotaxis	0.0086431
GO:0003205	cardiac chamber development	0.0086431
GO:0030595	leukocyte chemotaxis	0.0091921
GO:0003206	cardiac chamber morphogenesis	0.0118888
GO:0061162	establishment of monopolar cell polarity	0.0126127
GO:0000387	spliceosomal snRNP assembly	0.0127196
GO:0060415	muscle tissue morphogenesis	0.0135387
GO:1902895	positive regulation of pri-miRNA transcription by RNA polymerase II	0.0135387
GO:0061339	establishment or maintenance of monopolar cell polarity	0.0135387
GO:0030279	negative regulation of ossification	0.0139035
GO:0099173	postsynapse organization	0.0227007
GO:0048644	muscle organ morphogenesis	0.0227007
GO:0035088	establishment or maintenance of apical/basal cell polarity	0.0249714
GO:0061245	establishment or maintenance of bipolar cell polarity	0.0249714
GO:0050890	cognition	0.0254047
GO:0050921	positive regulation of chemotaxis	0.0254047
GO:1901342	regulation of vasculature development	0.0254047
GO:0060840	artery development	0.0254047
GO:0000395	mRNA 5'-splice site recognition	0.0254047
GO:0002699	positive regulation of immune effector process	0.0254047
GO:0022612	gland morphogenesis	0.0256173
GO:0002705	positive regulation of leukocyte mediated immunity	0.0268494
GO:0030593	neutrophil chemotaxis	0.0268494
GO:0034330	cell junction organization	0.0314656
GO:0043113	receptor clustering	0.0314656
GO:0048844	artery morphogenesis	0.033326
GO:0035089	establishment of apical/basal cell polarity	0.034083
GO:0055008	cardiac muscle tissue morphogenesis	0.034083
GO:0007163	establishment or maintenance of cell polarity	0.0342637
GO:0045765	regulation of angiogenesis	0.0388315
GO:0031667	response to nutrient levels	0.0388315
GO:1902893	regulation of pri-miRNA transcription by RNA polymerase II	0.0388315
GO:0007612	learning	0.0388315
GO:0003007	heart morphogenesis	0.0388315
GO:0002888	positive regulation of myeloid leukocyte mediated immunity	0.0388315
GO:0010876	lipid localization	0.0388315

Appendix C: Gene ontology component analysis of significantly differentially expressed genes in HEL cell clones with biallelic *TET2* mutations compared to parental clones with monoallelic *TET2* mutation (continued from previous page).

ID	Description	P_{adj}
GO:1904018	positive regulation of vasculature development	0.0404208
GO:0046467	membrane lipid biosynthetic process	0.0404208
GO:0034329	cell junction assembly	0.0404208
GO:0045216	cell-cell junction organization	0.0405342
GO:0014074	response to purine-containing compound	0.0405342
GO:0003209	cardiac atrium morphogenesis	0.0405342
GO:0045197	establishment or maintenance of epithelial cell apical/basal polarity	0.0405342
GO:1903522	regulation of blood circulation	0.0405342
GO:0060537	muscle tissue development	0.0405342
GO:0006376	mRNA splice site selection	0.0405342
GO:0030278	regulation of ossification	0.0405342
GO:0009991	response to extracellular stimulus	0.0405342
GO:0060562	epithelial tube morphogenesis	0.0405342
GO:0046683	response to organophosphorus	0.0407276
GO:0072578	neurotransmitter-gated ion channel clustering	0.0415456
GO:0032103	positive regulation of response to external stimulus	0.0421347
GO:0071621	granulocyte chemotaxis	0.0426451
GO:0051056	regulation of small GTPase mediated signal transduction	0.0426451
GO:0046058	cAMP metabolic process	0.0430189
GO:1904951	positive regulation of establishment of protein localization	0.0436303
GO:0042692	muscle cell differentiation	0.0458284
GO:0035904	aorta development	0.0466924
GO:0071496	cellular response to external stimulus	0.0470128
GO:0032970	regulation of actin filament-based process	0.047886
GO:1990266	neutrophil migration	0.047886
GO:0032754	positive regulation of interleukin-5 production	0.047886
GO:0071872	cellular response to epinephrine stimulus	0.047886
GO:0001819	positive regulation of cytokine production	0.047886
GO:0051924	regulation of calcium ion transport	0.047886
GO:0009267	cellular response to starvation	0.0493682

Appendix D: Gene ontology component analysis of significantly differentially expressed genes in HEL cell clones with biallelic TET2 mutations compared to parental clones with monoallelic TET2 mutation excluding HELmono1 cells (continued in next page).

ID	Description	P_{adj}
GO:0022626	cytosolic ribosome	1.20E-23
GO:0044391	ribosomal subunit	1.60E-20
GO:0005840	ribosome	1.26E-16
GO:0022625	cytosolic large ribosomal subunit	1.06E-15
GO:0005925	focal adhesion	7.50E-15
GO:0030055	cell-substrate junction	3.29E-14
GO:0015934	large ribosomal subunit	7.63E-12
GO:0015935	small ribosomal subunit	6.59E-10
GO:0022627	cytosolic small ribosomal subunit	8.44E-10
GO:0005844	polysome	4.42E-09
GO:0042788	polysomal ribosome	4.28E-07
GO:0140534	endoplasmic reticulum protein-containing complex	9.38E-06
GO:0042470	melanosome	4.22E-05
GO:0048770	pigment granule	4.22E-05
GO:0060205	cytoplasmic vesicle lumen	7.98E-05
GO:0034774	secretory granule lumen	9E-05
GO:0031983	vesicle lumen	9.42E-05
GO:0000502	proteasome complex	9.53E-05
GO:0031252	cell leading edge	0.000137
GO:0101002	ficolin-1-rich granule	0.000147
GO:0030667	secretory granule membrane	0.000166
GO:1904813	ficolin-1-rich granule lumen	0.000192
GO:0071013	catalytic step 2 spliceosome	0.000192
GO:1905369	endopeptidase complex	0.000192
GO:1905368	peptidase complex	0.00024
GO:0005765	lysosomal membrane	0.000453
GO:0098852	lytic vacuole membrane	0.000453
GO:0030666	endocytic vesicle membrane	0.000458
GO:0036464	cytoplasmic ribonucleoprotein granule	0.001
GO:0030139	endocytic vesicle	0.001
GO:0005774	vacuolar membrane	0.001018
GO:0035770	ribonucleoprotein granule	0.001044
GO:0010494	cytoplasmic stress granule	0.001044
GO:0030670	phagocytic vesicle membrane	0.001044
GO:0022624	proteasome accessory complex	0.001129
GO:0030532	small nuclear ribonucleoprotein complex	0.001433
GO:0030863	cortical cytoskeleton	0.001478
GO:0005790	smooth endoplasmic reticulum	0.001733
GO:0031227	intrinsic component of endoplasmic reticulum membrane	0.001825
GO:0005681	spliceosomal complex	0.001825
GO:0032587	ruffle membrane	0.002429
GO:0030027	lamellipodium	0.002553
GO:1904949	ATPase complex	0.002751
GO:0045335	phagocytic vesicle	0.003373
GO:0030662	coated vesicle membrane	0.004707

Appendix D: Gene ontology component analysis of significantly differentially expressed genes in HEL cell clones with biallelic *TET2* mutations compared to parental clones with monoallelic *TET2* mutation excluding HELmono1 cells (continued in next page).

ID	Description	P_{adj}
GO:0120114	Sm-like protein family complex	0.005095
GO:0070603	SWI/SNF superfamily-type complex	0.005139
GO:0005667	transcription regulator complex	0.006917
GO:0000932	P-body	0.007135
GO:0005788	endoplasmic reticulum lumen	0.007138
GO:0034663	endoplasmic reticulum chaperone complex	0.007765
GO:0005743	mitochondrial inner membrane	0.008077
GO:0097525	spliceosomal snRNP complex	0.008077
GO:0001726	ruffle	0.008671
GO:0031300	intrinsic component of organelle membrane	0.008671
GO:0016469	proton-transporting two-sector ATPase complex	0.009765
GO:0030176	integral component of endoplasmic reticulum membrane	0.009765
GO:0030867	rough endoplasmic reticulum membrane	0.009765
GO:0005838	proteasome regulatory particle	0.010736
GO:0090575	RNA polymerase II transcription regulator complex	0.012022
GO:0071556	integral component of lumenal side of endoplasmic reticulum membrane	0.012434
GO:0098553	lumenal side of endoplasmic reticulum membrane	0.012434
GO:0032279	asymmetric synapse	0.013063
GO:0031838	haptoglobin-hemoglobin complex	0.014285
GO:0005741	mitochondrial outer membrane	0.014431
GO:0099572	postsynaptic specialization	0.015572
GO:0005839	proteasome core complex	0.015572
GO:0098984	neuron to neuron synapse	0.015572
GO:0030118	clathrin coat	0.015572
GO:0030864	cortical actin cytoskeleton	0.017005
GO:0014069	postsynaptic density	0.017249
GO:0046540	U4/U6 x U5 tri-snRNP complex	0.018053
GO:0005938	cell cortex	0.018587
GO:0005684	U2-type spliceosomal complex	0.019
GO:0042581	specific granule	0.01917
GO:0000118	histone deacetylase complex	0.01917
GO:0098798	mitochondrial protein-containing complex	0.020042
GO:0097526	spliceosomal tri-snRNP complex	0.022344
GO:0005833	hemoglobin complex	0.023039
GO:0031209	SCAR complex	0.023039
GO:0031256	leading edge membrane	0.023991
GO:0031301	integral component of organelle membrane	0.023991
GO:0005793	endoplasmic reticulum-Golgi intermediate compartment	0.023991
GO:0016363	nuclear matrix	0.024742
GO:0031968	organelle outer membrane	0.024742
GO:0030135	coated vesicle	0.024877
GO:1902562	H4 histone acetyltransferase complex	0.026117
GO:0030665	clathrin-coated vesicle membrane	0.026804
GO:0035579	specific granule membrane	0.02768
GO:0098576	lumenal side of membrane	0.028499

Appendix D: Gene ontology component analysis of significantly differentially expressed genes in HEL cell clones with biallelic *TET2* mutations compared to parental clones with monoallelic *TET2* mutation excluding HELmono1 cells (continued from previous page).

ID	Description	P_{adj}
GO:0019867	outer membrane	0.028499
GO:0002102	podosome	0.029547
GO:0033176	proton-transporting V-type ATPase complex	0.03477
GO:0008250	oligosaccharyltransferase complex	0.035711
GO:0005682	U5 snRNP	0.039742
GO:0071682	endocytic vesicle lumen	0.039742
GO:0005635	nuclear envelope	0.041391
GO:0031902	late endosome membrane	0.041434
GO:0030136	clathrin-coated vesicle	0.044117
GO:0016471	vacuolar proton-transporting V-type ATPase complex	0.044477
GO:0005791	rough endoplasmic reticulum	0.044653
GO:0030125	clathrin vesicle coat	0.044653
GO:1990204	oxidoreductase complex	0.045196
GO:0000922	spindle pole	0.04767
GO:0030120	vesicle coat	0.048354

References

Abbaspour Babaei, M., B. Kamalidehghan, M. Saleem, H. Z. Huri, and F. Ahmadipour. 2016. 'Receptor tyrosine kinase (c-Kit) inhibitors: a potential therapeutic target in cancer cells', *Drug design, development and therapy*, 10: 2443-59.

Abdel-Wahab, O., M. Adli, L. M. LaFave, J. Gao, T. Hricik, A. H. Shih, S. Pandey, J. P. Patel, Y. R. Chung, R. Koche, F. Perna, X. Zhao, J. E. Taylor, C. Y. Park, M. Carroll, A. Melnick, S. D. Nimer, J. D. Jaffe, I. Aifantis, B. E. Bernstein, and R. L. Levine. 2012. 'ASXL1 mutations promote myeloid transformation through loss of PRC2-mediated gene repression', *Cancer Cell*, 22: 180-93.

Abdel-Wahab, O., A. Mullally, C. Hedvat, G. Garcia-Manero, J. Patel, M. Wadleigh, S. Malinge, J. Yao, O. Kilpivaara, R. Bhat, K. Huberman, S. Thomas, I. Dolgalev, A. Heguy, E. Paietta, M. M. Le Beau, M. Beran, M. S. Tallman, B. L. Ebert, H. M. Kantarjian, R. M. Stone, D. G. Gilliland, J. D. Crispino, and R. L. Levine. 2009. 'Genetic characterization of TET1, TET2, and TET3 alterations in myeloid malignancies', *Blood*, 114: 144-47.

Abdel-Wahab, O., A. Tefferi, and R. L. Levine. 2012. 'Role of TET2 and ASXL1 mutations in the pathogenesis of myeloproliferative neoplasms', *Hematol Oncol Clin North Am*, 26: 1053-64.

Achour, M., X. Jacq, P. Ronde, M. Alhosin, C. Charlot, T. Chataigneau, M. Jeanblanc, M. Macaluso, A. Giordano, A. D. Hughes, V. B. Schini-Kerth, and C. Bronner. 2008. 'The interaction of the SRA domain of ICBP90 with a novel domain of DNMT1 is involved in the regulation of VEGF gene expression', *Oncogene*, 27: 2187-97.

Ahn, J. S., H. J. Kim, Y. K. Kim, S. H. Jung, D. H. Yang, J. J. Lee, I. K. Lee, N. Y. Kim, M. D. Minden, C. W. Jung, J. H. Jang, H. J. Kim, J. H. Moon, S. K. Sohn, J. H. Won, S. H. Kim, N. Kim, K. Yoshida, S. Ogawa, and D. D. Kim. 2015. 'Adverse prognostic effect of homozygous TET2 mutation on the relapse risk of acute myeloid leukemia in patients of normal karyotype', *Haematologica*, 100: e351-3.

Almeida, A., A. R. Ferreira, M. J. Costa, S. Silva, K. Alnajjar, I. Bogalho, F. Pierdomenico, S. Esteves, M. Alpoim, G. Braz, E. Cortesão, and R. Pinto. 2017. 'Clinical outcomes of AML patients treated with Azacitidine in Portugal: A retrospective multicenter study', *Leukemia Research Reports*, 7: 6-10.

Almosailleakh, M., and J. Schwaller. 2019. 'Murine Models of Acute Myeloid Leukaemia', *Int J Mol Sci*, 20.

Altman, J. K., A. E. Perl, J. E. Cortes, M. J. Levis, C. C. Smith, M. R. Litzow, M. R. Baer, D. F. Claxton, H. P. Erba, S. C. Gill, S. L. Goldberg, J. G. Jurcic, R. A. Larson, C. Lui, E. K. Ritchie, B. Sargent, G. J. Schiller, A. I. Spira, S. A. Strickland, R. Tibes, C. Ustun,

E. S. Wang, R. K. Stuart, C. D. Baldus, C. Rollig, A. Neubauer, G. Martinelli, and E. Bahceci. 2015. 'Antileukemic Activity and Tolerability of ASP2215 80mg and Greater in FLT3 Mutation-Positive Subjects with Relapsed or Refractory Acute Myeloid Leukemia: Results from a Phase 1/2, Open-Label, Dose-Escalation/Dose-Response Study', *Blood*, 126: 321-21.

An, J., A. Rao, and M. Ko. 2017. 'TET family dioxygenases and DNA demethylation in stem cells and cancers', *Experimental & Molecular Medicine*, 49: e323-e23.

Aran, D., G. Toperoff, M. Rosenberg, and A. Hellman. 2011. 'Replication timing-related and gene body-specific methylation of active human genes', *Hum Mol Genet*, 20: 670-80.

Arber, D. A., A. Orazi, R. Hasserjian, J. Thiele, M. J. Borowitz, M. M. Le Beau, C. D. Bloomfield, M. Cazzola, and J. W. Vardiman. 2016. 'The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia', *Blood*, 127: 2391-405.

Aref, S., S. El Sharawy, M. Sabry, E. Azmy, and D. A. Raouf. 2014. 'Prognostic relevance of Wilms tumor 1 (WT1) gene Exon 7 mutations in-patient with cytogenetically normal acute myeloid leukemia', *Indian J Hematol Blood Transfus*, 30: 226-30.

Aryee, M. J., A. E. Jaffe, H. Corrada-Bravo, C. Ladd-Acosta, A. P. Feinberg, K. D. Hansen, and R. A. Irizarry. 2014. 'Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays', *Bioinformatics*, 30: 1363-9.

Asada, S., T. Fujino, S. Goyama, and T. Kitamura. 2019. 'The role of ASXL1 in hematopoiesis and myeloid malignancies', *Cellular and Molecular Life Sciences*, 76: 2511-23.

Awada, H., Y. Nagata, A. Goyal, M. F. Asad, B. Patel, C. M. Hirsch, T. Kuzmanovic, Y. Guan, B. P. Przychodzen, M. Aly, V. Adema, W. Shen, L. Williams, A. Nazha, M. E. Abazeed, M. A. Sekeres, T. Radivoyevitch, T. Haferlach, B. K. Jha, V. Visconte, and J. P. Maciejewski. 2019. 'Invariant phenotype and molecular association of biallelic TET2 mutant myeloid neoplasia', *Blood Advances*, 3: 339-49.

Bacher, U., T. Haferlach, C. Schoch, W. Kern, and S. Schnittger. 2006. 'Implications of NRAS mutations in AML: a study of 2502 patients', *Blood*, 107: 3847-53.

Bakshi, S. R., M. M. Brahmbhatt, P. J. Trivedi, E. N. Dalal, D. M. Patel, S. S. Purani, S. N. Shukla, P. M. Shah, and P. S. Patel. 2012. 'Trisomy 8 in leukemia: A GCRI experience', *Indian J Hum Genet*, 18: 106-8.

Ball, B. J., M. Hsu, S. M. Devlin, C. Famulare, S. F. Cai, A. Dunbar, Z. D. Epstein-Peterson, K. Menghrajani, J. L. Glass, J. Taylor, A. D. Viny, A. D. Goldberg, M. S. Tallman, and E. M. Stein. 2019. 'RAS Mutations Are Independently Associated with Decreased

Overall Survival and Event-Free Survival in Patients with AML Receiving Induction Chemotherapy', *Blood*, 134: 18-18.

Ball, M. P., J. B. Li, Y. Gao, J. H. Lee, E. M. LeProust, I. H. Park, B. Xie, G. Q. Daley, and G. M. Church. 2009. 'Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells', *Nat Biotechnol*, 27: 361-8.

Barbosa, K., S. Li, P. D. Adams, and A. J. Deshpande. 2019. 'The role of TP53 in acute myeloid leukemia: Challenges and opportunities', *Genes, Chromosomes and Cancer*, 58: 875-88.

Bauer, C., K. Göbel, N. Nagaraj, C. Colantuoni, M. Wang, U. Müller, E. Kremmer, A. Rottach, and H. Leonhardt. 2015. 'Phosphorylation of TET proteins is regulated via O-GlcNAcylation by the O-linked N-acetylglucosamine transferase (OGT)', *J Biol Chem*, 290: 4801-12.

Baylin, S. B., and P. A. Jones. 2011. 'A decade of exploring the cancer epigenome - biological and translational implications', *Nat Rev Cancer*, 11: 726-34.

Bejar, R., A. Lord, K. Stevenson, M. Bar-Natan, A. Pérez-Ladaga, J. Zaneveld, H. Wang, B. Caughey, P. Stojanov, G. Getz, G. Garcia-Manero, H. Kantarjian, R. Chen, R. M. Stone, D. Neuberg, D. P. Steensma, and B. L. Ebert. 2014. 'TET2 mutations predict response to hypomethylating agents in myelodysplastic syndrome patients', *Blood*, 124: 2705-12.

Belitsky, G., T. Fetisov, K. Kirsanov, E. Lesovaya, O. Vlasova, and M. Yakubovskaya. 2020. 'Therapy-related acute myeloid leukemia and its prevention', *American journal of blood research*, 10: 416-33.

Bertani, G., F. Guolo, V. Mancini, R. Greco, P. Minetto, M. Passannante, R. M. Lemoli, and R. Cairoli. 2020. 'Therapy-Related AML (t-AML), a Heterogeneous Disease: Multicenter Analysis on Biological and Clinical Differences between Cases Following Breast Cancer and Lymphoma Treatment', *Blood*, 136: 31-31.

Bhatia, S. 2013. 'Therapy-related myelodysplasia and acute myeloid leukemia', *Seminars in oncology*, 40: 666-75.

Bhutani, N., D. M. Burns, and H. M. Blau. 2011. 'DNA demethylation dynamics', *Cell*, 146: 866-72.

Bibikova, M., B. Barnes, C. Tsan, V. Ho, B. Klotzle, J. M. Le, D. Delano, L. Zhang, G. P. Schroth, K. L. Gunderson, J. B. Fan, and R. Shen. 2011. 'High density DNA methylation array with single CpG site resolution', *Genomics*, 98: 288-95.

Binkhathlan, Z., and A. Lavasanifar. 2013. 'P-glycoprotein inhibition as a therapeutic approach for overcoming multidrug resistance in cancer: current status and future perspectives', *Curr Cancer Drug Targets*, 13: 326-46.

Bocker, M. T., F. Tuorto, G. Raddatz, T. Musch, F.-C. Yang, M. Xu, F. Lyko, and A. Breiling. 2012. 'Hydroxylation of 5-methylcytosine by TET2 maintains the active state of the mammalian HOXA cluster', *Nat Commun*, 3: 818.

Borthakur, G., and H. Kantarjian. 2021. 'Core binding factor acute myelogenous leukemia-2021 treatment algorithm', *Blood Cancer J*, 11: 114.

Bose, P., P. Vachhani, and J. E. Cortes. 2017. 'Treatment of Relapsed/Refractory Acute Myeloid Leukemia', *Current Treatment Options in Oncology*, 18: 17.

Bostick, M., J. K. Kim, P. O. Esteve, A. Clark, S. Pradhan, and S. E. Jacobsen. 2007. 'UHRF1 plays a role in maintaining DNA methylation in mammalian cells', *Science*, 317: 1760-4.

Boutzen, H., E. Saland, C. Larrue, F. de Toni, L. Gales, F. A. Castelli, M. Cathebas, S. Zaghdoudi, L. Stuani, T. Kaoma, R. Riscal, G. Yang, P. Hirsch, M. David, V. De Mas-Mansat, E. Delabesse, L. Vallar, F. Delhommeau, I. Jouanin, O. Ouerfelli, L. Le Cam, L. K. Linares, C. Junot, J.-C. Portais, F. Vergez, C. Récher, and J.-E. Sarry. 2016. 'Isocitrate dehydrogenase 1 mutations prime the all-trans retinoic acid myeloid differentiation pathway in acute myeloid leukemia', *The Journal of experimental medicine*, 213: 483-97.

Bowen, D. T., M. E. Frew, R. Hills, R. E. Gale, K. Wheatley, M. J. Groves, S. E. Langabeer, P. D. Kottaridis, A. V. Moorman, A. K. Burnett, and D. C. Linch. 2005. 'RAS mutation in acute myeloid leukemia is associated with distinct cytogenetic subgroups but does not influence outcome in patients younger than 60 years', *Blood*, 106: 2113-9.

Bowman, R. L., and R. L. Levine. 2017. 'TET2 in Normal and Malignant Hematopoiesis', *Cold Spring Harb Perspect Med*, 7.

Boyer, K. 1996. 'The Adult Leukemias - Part 1: Acute Myeloid Leukemia.', *The Internet Journal of Academic Physician Assistants.*, 1.

Boyer, T., F. Gonzales, A. Barthélémy, A. Marceau-Renaut, P. Peyrouze, S. Guihard, P. Lepelley, A. Plesa, O. Nibourel, C. Delattre, M. Wetterwald, N. Pottier, I. Plantier, S. Botton, H. Dombret, C. Berthon, C. Preudhomme, C. Roumier, and M. Cheok. 2019. 'Clinical Significance of ABCB1 in Acute Myeloid Leukemia: A Comprehensive Study', *Cancers (Basel)*, 11.

Breen, E. C., A. R. Rezai, K. Nakajima, G. N. Beall, R. T. Mitsuyasu, T. Hirano, T. Kishimoto, and O. Martinez-Maza. 1990. 'Infection with HIV is associated with elevated IL-6 levels and production', *J Immunol*, 144: 480-4.

Brown, D., S. Kogan, E. Lagasse, I. Weissman, M. Alcalay, P. G. Pelicci, S. Atwater, and J. M. Bishop. 1997. 'A PMLRARalpha transgene initiates murine acute promyelocytic leukemia', *Proc Natl Acad Sci U S A*, 94: 2551-6.

Burnett, A. K. 2018. 'Treatment of Older Patients With Newly Diagnosed AML Unfit for Traditional Therapy', *Clin Lymphoma Myeloma Leuk*, 18: 553-57.

Burnett, A. K., N. H. Russell, R. K. Hills, A. E. Hunter, L. Kjeldsen, J. Yin, B. E. S. Gibson, K. Wheatley, and D. Milligan. 2013. 'Optimization of Chemotherapy for Younger Patients With Acute Myeloid Leukemia: Results of the Medical Research Council AML15 Trial', *Journal of Clinical Oncology*, 31: 3360-68.

Busque, L., M. Buscarlet, L. Mollica, and R. L. Levine. 2018. 'Concise Review: Age-Related Clonal Hematopoiesis: Stem Cells Tempting the Devil', *Stem Cells*, 36: 1287-94.

Cacemiro, M. d. C., J. G. Cominal, R. Tognon, N. d. S. Nunes, B. P. Simões, L. L. d. Figueiredo-Pontes, L. F. B. Catto, F. Traina, E. X. Souto, F. A. Zambuzi, F. G. Frantz, and F. A. d. Castro. 2018. 'Philadelphia-negative myeloproliferative neoplasms as disorders marked by cytokine modulation', *Hematology, transfusion and cell therapy*, 40: 120-31.

Care, R. S., P. J. Valk, A. C. Goodeve, F. M. Abu-Duhier, W. M. Geertsma-Kleinekoort, G. A. Wilson, M. A. Gari, I. R. Peake, B. Löwenberg, and J. T. Reilly. 2003. 'Incidence and prognosis of c-KIT and FLT3 mutations in core binding factor (CBF) acute myeloid leukaemias', *Br J Haematol*, 121: 775-7.

Cedena, M. T., I. Rapado, A. Santos-Lozano, R. Ayala, E. Onecha, M. Abaigar, E. Such, F. Ramos, J. Cervera, M. Díez-Campelo, G. Sanz, J. H. Rivas, A. Lucía, and J. Martínez-López. 2017. 'Mutations in the DNA methylation pathway and number of driver mutations predict response to azacitidine in myelodysplastic syndromes', *Oncotarget*, 8: 106948-61.

Chang, W. 2016. 'shiny: Web Application Framework for R. R package version 1.7.1'. https://shiny.rstudio.com/.

Chen, D., M. Christopher, N. M. Helton, I. Ferguson, T. J. Ley, and D. H. Spencer. 2018. 'DNMT3A(R882)-associated hypomethylation patterns are maintained in primary AML xenografts, but not in the DNMT3A(R882C) OCI-AML3 leukemia cell line', *Blood Cancer J*, 8: 38-38.

Chen, K. G., and B. I. Sikic. 2012. 'Molecular pathways: regulation and therapeutic implications of multidrug resistance', *Clin Cancer Res*, 18: 1863-9.

Chen, Q., Y. Chen, C. Bian, R. Fujiki, and X. Yu. 2013. 'TET2 promotes histone O-GlcNAcylation during gene transcription', *Nature*, 493: 561-64.

Chen, Q., J. He, X. Guo, J. Chen, L. Yang, X. Lin, Y. Li, E. Zhang, Y. Yang, W. Wu, D. He, H. Huang, S. Liu, and Z. Cai. 2015. 'TET2 Downregulation Leads to AML Cells Sensitive to Decitabine', *Blood*, 126: 3643-43.

Cheng, J., S. Guo, S. Chen, Stephen J. Mastriano, C. Liu, Ana C. D'Alessio, E. Hysolli, Y. Guo, H. Yao, Cynthia M. Megyola, D. Li, J. Liu, W. Pan, Christine A. Roden, X.-L. Zhou, K. Heydari, J. Chen, I.-H. Park, Y. Ding, Y. Zhang, and J. Lu. 2013. 'An Extensive Network of TET2-Targeting MicroRNAs Regulates Malignant Hematopoiesis', *Cell Reports*, 5: 471-81.

Cheng, J. X., L. Chen, Y. Li, A. Cloe, M. Yue, J. Wei, K. A. Watanabe, J. M. Shammo, J. Anastasi, Q. J. Shen, R. A. Larson, C. He, M. M. Le Beau, and J. W. Vardiman. 2018. 'RNA cytosine methylation and methyltransferases mediate chromatin organization and 5-azacytidine response and resistance in leukaemia', *Nat Commun*, 9: 1163.

Chin, K. V., K. Ueda, I. Pastan, and M. M. Gottesman. 1992. 'Modulation of activity of the promoter of the human MDR1 gene by Ras and p53', *Science*, 255: 459-62.

Choi, S. M., R. Dewar, P. W. Burke, and L. Shao. 2018. 'Partial tandem duplication of KMT2A (MLL) may predict a subset of myelodysplastic syndrome with unique characteristics and poor outcome', *Haematologica*, 103: e131-e34.

Chong, T., E. Ahearn, and L. Cimmino. 2019. 'Reprogramming the Epigenome With Vitamin C', *Frontiers in Cell and Developmental Biology*, 7.

Chou, W. C., S. C. Chou, C. Y. Liu, C. Y. Chen, H. A. Hou, Y. Y. Kuo, M. C. Lee, B. S. Ko, J. L. Tang, M. Yao, W. Tsay, S. J. Wu, S. Y. Huang, S. C. Hsu, Y. C. Chen, Y. C. Chang, Y. Y. Kuo, K. T. Kuo, F. Y. Lee, M. C. Liu, C. W. Liu, M. H. Tseng, C. F. Huang, and H. F. Tien. 2011. 'TET2 mutation is an unfavorable prognostic factor in acute myeloid leukemia patients with intermediate-risk cytogenetics', *Blood*, 118: 3803-10.

Choudhuri, S., and C. D. Klaassen. 2006. 'Structure, Function, Expression, Genomic Organization, and Single Nucleotide Polymorphisms of Human ABCB1 (MDR1), ABCC (MRP), and ABCG2 (BCRP) Efflux Transporters', *International Journal of Toxicology*, 25: 231-59.

Christiansen, D. H., M. K. Andersen, F. Desta, and J. Pedersen-Bjergaard. 2005. 'Mutations of genes in the receptor tyrosine kinase (RTK)/RAS-BRAF signal transduction pathway in therapy-related myelodysplasia and acute myeloid leukemia', *Leukemia*, 19: 2232-40.

Christman, J. K. 2002. '5-Azacytidine and 5-aza-2′-deoxycytidine as inhibitors of DNA methylation: mechanistic studies and their implications for cancer therapy', *Oncogene*, 21: 5483-95.

Cimmino, L., I. Dolgalev, Y. Wang, A. Yoshimi, G. H. Martin, J. Wang, V. Ng, B. Xia, M. T. Witkowski, M. Mitchell-Flack, I. Grillo, S. Bakogianni, D. Ndiaye-Lobry, M. T. Martín, M. Guillamot, R. S. Banh, M. Xu, M. E. Figueroa, R. A. Dickins, O. Abdel-Wahab, C. Y. Park, A. Tsirigos, B. G. Neel, and I. Aifantis. 2017. 'Restoration of TET2 Function Blocks Aberrant Self-Renewal and Leukemia Progression', *Cell*, 170: 1079-95.e20.

Cloos, J., B. F. Goemans, C. J. Hess, J. W. van Oostveen, Q. Waisfisz, S. Corthals, D. de Lange, N. Boeckx, K. Hählen, D. Reinhardt, U. Creutzig, G. J. Schuurhuis, M. Zwaan Ch, and G. J. Kaspers. 2006. 'Stability and prognostic influence of FLT3 mutations in paired initial and relapsed AML samples', *Leukemia*, 20: 1217-20.

Cloutier, A., L. Shkreta, J. Toutant, M. Durand, P. Thibault, and B. Chabot. 2018. 'hnRNP A1/A2 and Sam68 collaborate with SRSF10 to control the alternative splicing response to oxaliplatin-mediated DNA damage', *Scientific reports*, 8: 2206-06.

Cluzeau, T., A. Dubois, A. Jacquel, F. Luciano, A. Renneville, C. Preudhomme, J. M. Karsenti, N. Mounier, P. Rohrlich, S. Raynaud, B. Mari, G. Robert, and P. Auberger. 2014. 'Phenotypic and genotypic characterization of azacitidine-sensitive and resistant SKM1 myeloid cell lines', *Oncotarget*, 5: 4384-91.

Coombs, C. C., D. A. Sallman, S. M. Devlin, S. Dixit, A. Mohanty, K. Knapp, N. H. Al Ali, J. E. Lancet, A. F. List, R. S. Komrokji, E. Padron, M. E. Arcila, V. M. Klimek, M. R. van den Brink, M. S. Tallman, R. L. Levine, R. K. Rampal, and F. Rapaport. 2016. 'Mutational correlates of response to hypomethylating agent therapy in acute myeloid leukemia', *Haematologica*, 101: e457-e60.

Coombs, C. C., M. S. Tallman, and R. L. Levine. 2016. 'Molecular therapy for acute myeloid leukaemia', *Nat Rev Clin Oncol*, 13: 305-18.

Cooper, G. 2000. 'Translation of mRNA.' in, *The Cell: A Molecular Approach* (Sinauer Associates: Sunderland (MA)).

Corey, S. J., M. D. Minden, D. L. Barber, H. Kantarjian, J. C. Y. Wang, and A. D. Schimmer. 2007. 'Myelodysplastic syndromes: the complexity of stem-cell diseases', *Nature Reviews Cancer*, 7: 118-29.

Corrêa, S., E. Abdelhay, P. Paschka, V. I. Gaidzik, R. Hassan, R. Claus, R. F. Schlenk, K. Döhner, and L. Bullinger. 2014. 'ABCB1 Expression in Acute Myeloid Leukemia (AML): A Possible Predictive Value for Treatment Resistance?', *Blood*, 124: 3618-18.

Cortes, J. E., A. E. Perl, H. Dombret, S. Kayser, B. r. Steffen, P. Rousselot, G. Martinelli, E. H. Estey, A. K. Burnett, G. Gammon, D. Trone, E. Leo, and M. J. Levis. 2012. 'Final Results of a Phase 2 Open-Label, Monotherapy Efficacy and Safety Study of Quizartinib (AC220) in Patients ≥ 60 Years of Age with FLT3 ITD Positive or Negative Relapsed/Refractory Acute Myeloid Leukemia', *Blood*, 120: 48-48.

Cripe, L. D., H. Uno, E. M. Paietta, M. R. Litzow, R. P. Ketterling, J. M. Bennett, J. M. Rowe, H. M. Lazarus, S. Luger, and M. S. Tallman. 2010. 'Zosuquidar, a novel modulator of P-glycoprotein, does not improve the outcome of older patients with newly diagnosed acute myeloid leukemia: a randomized, placebo-controlled trial of the Eastern Cooperative Oncology Group 3999', *Blood*, 116: 4077-85.

Crouch, S. P., R. Kozlowski, K. J. Slater, and J. Fletcher. 1993. 'The use of ATP bioluminescence as a measure of cell proliferation and cytotoxicity', *J Immunol Methods*, 160: 81-8.

Damaraju, V. L., D. Mowles, S. Yao, A. Ng, J. D. Young, C. E. Cass, and Z. Tong. 2012. 'Role of human nucleoside transporters in the uptake and cytotoxicity of azacitidine and decitabine', *Nucleosides Nucleotides Nucleic Acids*, 31: 236-55.

Damiani, D., M. Tiribelli, E. Calistri, A. Geromin, A. Chiarvesio, A. Michelutti, M. Cavallin, and R. Fanin. 2006. 'The prognostic value of P-glycoprotein (ABCB) and breast cancer resistance protein (ABCG2) in adults with de novo acute myeloid leukemia with normal karyotype', *Haematologica*, 91: 825-8.

Daver, N., R. F. Schlenk, N. H. Russell, and M. J. Levis. 2019. 'Targeting FLT3 mutations in AML: review of current knowledge and evidence', *Leukemia*, 33: 299-312.

David, G. L., S. Yegnasubramanian, A. Kumar, V. L. Marchi, A. M. De Marzo, X. Lin, and W. G. Nelson. 2004. 'MDR1 promoter hypermethylation in MCF-7 human breast cancer cells: changes in chromatin structure induced by treatment with 5-Aza-cytidine', *Cancer Biol Ther*, 3: 540-8.

De Braekeleer, E., N. Douet-Guilbert, and M. De Braekeleer. 2014. 'RARA fusion genes in acute promyelocytic leukemia: a review', *Expert Review of Hematology*, 7: 347-57.

De Keersmaecker, K., Z. K. Atak, N. Li, C. Vicente, S. Patchett, T. Girardi, V. Gianfelici, E. Geerdens, E. Clappier, M. Porcu, I. Lahortiga, R. Lucà, J. Yan, G. Hulselmans, H. Vranckx, R. Vandepoel, B. Sweron, K. Jacobs, N. Mentens, I. Wlodarska, B. Cauwelier, J. Cloos, J. Soulier, A. Uyttebroeck, C. Bagni, B. A. Hassan, P. Vandenberghe, A. W. Johnson, S. Aerts, and J. Cools. 2013. 'Exome sequencing identifies mutation in CNOT3 and ribosomal genes RPL5 and RPL10 in T-cell acute lymphoblastic leukemia', *Nat Genet*, 45: 186-90.

De Kouchkovsky, I., and M. Abdul-Hay. 2016. "Acute myeloid leukemia: a comprehensive review and 2016 update", *Blood Cancer J*, 6: e441.

de Noronha, T. R., M. Mitne-Neto, and M. d. L. Chauffaille. 2019. 'JAK2-mutated acute myeloid leukemia: comparison of next-generation sequencing (NGS) and single nucleotide polymorphism array (SNPa) findings between two cases', *Autopsy & case reports*, 9: e2018084-e84.

de Thé, H., and Z. Chen. 2010. 'Acute promyelocytic leukaemia: novel insights into the mechanisms of cure', *Nature Reviews Cancer*, 10: 775-83.

de Thé, H., J. Zhu, R. Nasr, J. Ablain, and V. Lallemand-Breittenbach. 2015. 'PML/RARA as the Master Driver of APL Pathogenesis and Therapy Response.' in Michael Andreeff (ed.), *Targeted Therapy of Acute Myeloid Leukemia* (Springer New York: New York, NY).

Deplus, R., B. Delatte, M. K. Schwinn, M. Defrance, J. Méndez, N. Murphy, M. A. Dawson, M. Volkmar, P. Putmans, E. Calonne, A. H. Shih, R. L. Levine, O. Bernard, T. Mercher, E. Solary, M. Urh, D. L. Daniels, and F. Fuks. 2013. 'TET2 and TET3 regulate GlcNAcylation and H3K4 methylation through OGT and SET1/COMPASS', *The EMBO journal*, 32: 645-55.

Derissen, E. J., J. H. Beijnen, and J. H. Schellens. 2013. 'Concise drug review: azacitidine and decitabine', *Oncologist*, 18: 619-24.

DiNardo, C. D., and J. E. Cortes. 2016. 'Mutations in AML: prognostic and therapeutic implications', *Hematology Am Soc Hematol Educ Program*, 2016: 348-55.

DiNardo, C. D., K. P. Patel, G. Garcia-Manero, R. Luthra, S. Pierce, G. Borthakur, E. Jabbour, T. Kadia, N. Pemmaraju, M. Konopleva, S. Faderl, J. Cortes, H. M. Kantarjian, and F. Ravandi. 2014. 'Lack of association of IDH1, IDH2 and DNMT3A mutations with outcome in older patients with acute myeloid leukemia treated with hypomethylating agents', *Leuk Lymphoma*, 55: 1925-9.

DiNardo, C. D., K. Pratz, V. Pullarkat, B. A. Jonas, M. Arellano, P. S. Becker, O. Frankfurt, M. Konopleva, A. H. Wei, H. M. Kantarjian, T. Xu, W. J. Hong, B. Chyla, J. Potluri, D. A. Pollyea, and A. Letai. 2019. 'Venetoclax combined with decitabine or azacitidine in treatment-naive, elderly patients with acute myeloid leukemia', *Blood*, 133: 7-17.

DiNardo, C. D., F. Ravandi, S. Agresta, M. Konopleva, K. Takahashi, T. Kadia, M. Routbort, K. P. Patel, M. Brandt, S. Pierce, G. Garcia-Manero, J. Cortes, and H. Kantarjian. 2015. 'Characteristics, clinical outcome, and prognostic significance of IDH mutations in AML', *Am J Hematol*, 90: 732-6.

Döhner, H., A. Dolnik, L. Tang, J. F. Seymour, M. D. Minden, R. M. Stone, T. B. Del Castillo, H. K. Al-Ali, V. Santini, P. Vyas, C. L. Beach, K. J. MacBeth, B. S. Skikne, S. Songer, N. Tu, L. Bullinger, and H. Dombret. 2018. 'Cytogenetics and gene mutations influence survival in older patients with acute myeloid leukemia treated with azacitidine or conventional care', *Leukemia*, 32: 2546-57.

Döhner, H., E. Estey, D. Grimwade, S. Amadori, F. R. Appelbaum, T. Büchner, H. Dombret, B. L. Ebert, P. Fenaux, R. A. Larson, R. L. Levine, F. Lo-Coco, T. Naoe, D. Niederwieser, G. J. Ossenkoppele, M. Sanz, J. Sierra, M. S. Tallman, H.-F. Tien, A. H. Wei, B. Löwenberg, and C. D. Bloomfield. 2017. 'Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel', *Blood*, 129: 424-47.

Dohner, H., E. H. Estey, S. Amadori, F. R. Appelbaum, T. Buchner, A. K. Burnett, H. Dombret, P. Fenaux, D. Grimwade, R. A. Larson, F. Lo-Coco, T. Naoe, D. Niederwieser, G. J. Ossenkoppele, M. A. Sanz, J. Sierra, M. S. Tallman, B. Lowenberg, and C. D. Bloomfield. 2010. 'Diagnosis and management of acute myeloid leukemia in adults: recommendations from an international expert panel, on behalf of the European LeukemiaNet', *Blood*, 115: 453-74.

Döhner, H., A. H. Wei, and B. Löwenberg. 2021. 'Towards precision medicine for AML', *Nature Reviews Clinical Oncology*, 18: 577-90.

Dombret, H., and C. Gardin. 2016. 'An update of current treatments for adult acute myeloid leukemia', *Blood*, 127: 53.

Dombret, H., J. F. Seymour, A. Butrym, A. Wierzbowska, D. Selleslag, J. H. Jang, R. Kumar, J. Cavenagh, A. C. Schuh, A. Candoni, C. Récher, I. Sandhu, T. Bernal del Castillo, H. K. Al-Ali, G. Martinelli, J. Falantes, R. Noppeney, R. M. Stone, M. D. Minden, H. McIntyre, S. Songer, L. M. Lucy, C. L. Beach, and H. Döhner. 2015. 'International phase 3 study of azacitidine vs conventional care regimens in older patients with newly diagnosed AML with >30% blasts', *Blood*, 126: 291-9.

Donati, G., S. Bertoni, E. Brighenti, M. Vici, D. Treré, S. Volarevic, L. Montanaro, and M. Derenzini. 2011. 'The balance between rRNA and ribosomal protein synthesis up- and downregulates the tumour suppressor p53 in mammalian cells', *Oncogene*, 30: 3274-88.

Dotson, J. L., and Y. Lebowicz. 2021. 'Myelodysplastic Syndrome.' in, *StatPearls* (StatPearls Publishing LLC.: Treasure Island (FL)).

Du, P., X. Zhang, C.-C. Huang, N. Jafari, W. A. Kibbe, L. Hou, and S. M. Lin. 2010. 'Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis', *BMC Bioinformatics*, 11: 587.

Du, Q., Z. Wang, and V. L. Schramm. 2016. 'Human DNMT1 transition state structure', *Proceedings of the National Academy of Sciences*, 113: 2916.

Duchmann, M., and R. Itzykson. 2019. 'Clinical update on hypomethylating agents', *International Journal of Hematology*, 110: 161-69.

Duchmann, M., F. F. Yalniz, A. Sanna, D. Sallman, C. C. Coombs, A. Renneville, O. Kosmider, T. Braun, U. Platzbecker, L. Willems, L. Adès, M. Fontenay, R. Rampal, E. Padron, N. Droin, C. Preudhomme, V. Santini, M. M. Patnaik, P. Fenaux, E. Solary, and R. Itzykson. 2018. 'Prognostic Role of Gene Mutations in Chronic Myelomonocytic Leukemia Patients Treated With Hypomethylating Agents', *EBioMedicine*, 31: 174-81.

Dupont, C., D. R. Armant, and C. A. Brenner. 2009. 'Epigenetics: Definition, Mechanisms and Clinical Perspective', *Seminars in reproductive medicine*, 27: 351-57.

Edgar, R., M. Domrachev, and A. E. Lash. 2002. 'Gene Expression Omnibus: NCBI gene expression and hybridization array data repository', *Nucleic Acids Res*, 30: 207-10.

Egger, G., G. Liang, A. Aparicio, and P. A. Jones. 2004. 'Epigenetics in human disease and prospects for epigenetic therapy', *Nature*, 429: 457.

Emadi, A., R. Faramand, B. Carter-Cooper, S. Tolu, L. A. Ford, R. G. Lapidus, M. Wetzler, E. S. Wang, A. Etemadi, and E. A. Griffiths. 2015. 'Presence of isocitrate dehydrogenase mutations may predict clinical response to hypomethylating agents in patients with acute myeloid leukemia', *Am J Hematol*, 90: E77-9.

Estes, M. L., Y. Ozawa, A. H. Williams, A. F. List, and L. Sokol. 2005. 'Homozygous JAK2 V617P Gain-of-Function Mutation Is Responsible for Constitutive JAK2/STAT5 Activation and Proliferation of HEL.92.1.7. Cell Line', *Blood*, 106: 4377-77.

Estey, E. H. 2014. 'Acute myeloid leukemia: 2014 update on risk-stratification and management', *Am J Hematol*, 89: 1063-81.

Etheridge, S. L., M. E. Roh, M. E. Cosgrove, V. Sangkhae, N. E. Fox, J. Chen, J. A. López, K. Kaushansky, and I. S. Hitchcock. 2014. 'JAK2V₆₁₇F-positive endothelial cells contribute to clotting abnormalities in myeloproliferative neoplasms', *Proceedings of the National Academy of Sciences*, 111: 2295.

Falconi, G., E. Fabiani, A. Piciocchi, M. Criscuolo, L. Fianchi, E. L. Lindfors Rossi, C. Finelli, E. Cerqui, T. Ottone, A. Molteni, M. Parma, S. Santarone, A. Candoni, S. Sica, G. Leone, F. Lo-Coco, and M. T. Voso. 2019. 'Somatic mutations as markers of outcome after azacitidine and allogeneic stem cell transplantation in higher-risk myelodysplastic syndromes', *Leukemia*, 33: 785-90.

Falini, B., I. Nicoletti, M. F. Martelli, and C. Mecucci. 2007. 'Acute myeloid leukemia carrying cytoplasmic/mutated nucleophosmin (NPMc+ AML): biologic and clinical features', *Blood*, 109: 874-85.

Federici, L., and B. Falini. 2013. 'Nucleophosmin mutations in acute myeloid leukemia: a tale of protein unfolding and mislocalization', *Protein science : a publication of the Protein Society*, 22: 545-56.

Fenaux, P., G. J. Mufti, E. Hellstrom-Lindberg, V. Santini, C. Finelli, A. Giagounidis, R. Schoch, N. Gattermann, G. Sanz, A. List, S. D. Gore, J. F. Seymour, J. M. Bennett, J. Byrd, J. Backstrom, L. Zimmerman, D. McKenzie, C. Beach, and L. R. Silverman. 2009. 'Efficacy of azacitidine compared with that of conventional care regimens in the treatment of higher-risk myelodysplastic syndromes: a randomised, open-label, phase III study', *Lancet Oncol*, 10: 223-32.

Fenaux, P., G. J. Mufti, E. Hellström-Lindberg, V. Santini, N. Gattermann, U. Germing, G. Sanz, A. F. List, S. Gore, J. F. Seymour, H. Dombret, J. Backstrom, L. Zimmerman, D. McKenzie, C. L. Beach, and L. R. Silverman. 2010. 'Azacitidine prolongs overall survival compared with conventional care regimens in elderly patients with low bone marrow blast count acute myeloid leukemia', *J Clin Oncol*, 28: 562-9.

Feng, Y., X. Li, K. Cassady, Z. Zou, and X. Zhang. 2019. 'TET2 Function in Hematopoietic Malignancies, Immune Regulation, and DNA Repair', *Frontiers in oncology*, 9: 210-10.

Fernandez-Mercado, M., B. H. Yip, A. Pellagatti, C. Davies, M. J. Larrayoz, T. Kondo, C. Perez, S. Killick, E. J. McDonald, M. D. Odero, X. Agirre, F. Prosper, M. J. Calasanz, J. S. Wainscoat, and J. Boultwood. 2012. 'Mutation patterns of 16 genes in primary and secondary acute myeloid leukemia (AML) with normal cytogenetics', *PLoS One*, 7: e42334.

Figueroa, M. E., O. Abdel-Wahab, C. Lu, P. S. Ward, J. Patel, A. Shih, Y. Li, N. Bhagwat, A. Vasanthakumar, H. F. Fernandez, M. S. Tallman, Z. Sun, K. Wolniak, J. K. Peeters, W. Liu, S. E. Choe, V. R. Fantin, E. Paietta, B. Lowenberg, J. D. Licht, L. A. Godley, R. Delwel, P. J. Valk, C. B. Thompson, R. L. Levine, and A. Melnick. 2010. 'Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation', *Cancer Cell*, 18: 553-67.

Filion, G. J., S. Zhenilo, S. Salozhin, D. Yamada, E. Prokhortchouk, and P. A. Defossez. 2006. 'A family of human zinc finger proteins that bind methylated DNA and repress transcription', *Mol Cell Biol*, 26: 169-81.

Flotho, C., R. Claus, C. Batz, M. Schneider, I. Sandrock, S. Ihde, C. Plass, C. M. Niemeyer, and M. Lübbert. 2009. 'The DNA methyltransferase inhibitors azacitidine, decitabine

and zebularine exert differential effects on cancer gene expression in acute myeloid leukemia cells', *Leukemia*, 23: 1019-28.

Fryxell, K. B., S. B. McGee, D. K. Simoneaux, C. L. Willman, and M. M. Cornwell. 1999. 'Methylation analysis of the human multidrug resistance 1 gene in normal and leukemic hematopoietic cells', *Leukemia*, 13: 910-17.

Furth, P. A., L. St Onge, H. Böger, P. Gruss, M. Gossen, A. Kistner, H. Bujard, and L. Hennighausen. 1994. 'Temporal control of gene expression in transgenic mice by a tetracycline-responsive promoter', *Proceedings of the National Academy of Sciences*, 91: 9302-06.

Gaidzik, V. I., P. Paschka, D. Spath, M. Habdank, C. H. Kohne, U. Germing, M. von Lilienfeld-Toal, G. Held, H. A. Horst, D. Haase, M. Bentz, K. Gotze, H. Dohner, R. F. Schlenk, L. Bullinger, and K. Dohner. 2012. 'TET2 mutations in acute myeloid leukemia (AML): results from a comprehensive genetic and clinical analysis of the AML study group', *J Clin Oncol*, 30: 1350-7.

Garcia-Manero, G., G. Roboz, K. Walsh, H. Kantarjian, E. Ritchie, P. Kropf, C. O'Connell, R. Tibes, S. Lunin, T. Rosenblat, K. Yee, W. Stock, E. Griffiths, J. Mace, N. Podoltsev, J. Berdeja, E. Jabbour, J. J. Issa, Y. Hao, H. N. Keer, M. Azab, and M. R. Savona. 2019. 'Guadecitabine (SGI-110) in patients with intermediate or high-risk myelodysplastic syndromes: phase 2 results from a multicentre, open-label, randomised, phase 1/2 trial', *Lancet Haematol*, 6: e317-e27.

García, M. G., A. Carella, R. G. Urdinguio, G. F. Bayón, V. Lopez, J. R. Tejedor, M. I. Sierra, E. García-Toraño, P. Santamarina, R. F. Perez, C. Mangas, A. Astudillo, M. D. Corte-Torres, I. Sáenz-de-Santa-María, M.-D. Chiara, A. F. Fernández, and M. F. Fraga. 2018. 'Epigenetic dysregulation of TET2 in human glioblastoma', *Oncotarget*, 9: 25922-34.

Gardin, C., and H. Dombret. 2017. 'Hypomethylating Agents as a Therapy for AML', *Current Hematologic Malignancy Reports*, 12: 1-10.

Genovese, G., A. K. Kahler, R. E. Handsaker, J. Lindberg, S. A. Rose, S. F. Bakhoum, K. Chambert, E. Mick, B. M. Neale, M. Fromer, S. M. Purcell, O. Svantesson, M. Landen, M. Hoglund, S. Lehmann, S. B. Gabriel, J. L. Moran, E. S. Lander, P. F. Sullivan, P. Sklar, H. Gronberg, C. M. Hultman, and S. A. McCarroll. 2014. 'Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence', *N Engl J Med*, 371: 2477-87.

Gilliland, D. G., and J. D. Griffin. 2002. 'The roles of FLT3 in hematopoiesis and leukemia', *Blood*, 100: 1532-42.

Giuseppe, L., P. Livio, B.-Y. Dina, and V. Maria Teresa. 2007. 'Therapy-related leukemia and myelodysplasia: susceptibility and incidence', *Haematologica*, 92: 1389-98.

Godley, L. A., and R. A. Larson. 2008. 'Therapy-related myeloid leukemia', *Seminars in oncology*, 35: 418-29.

Gore, S. D., P. Fenaux, V. Santini, J. M. Bennett, L. R. Silverman, J. F. Seymour, E. Hellström-Lindberg, A. S. Swern, C. L. Beach, and A. F. List. 2013. 'A multivariate analysis of the relationship between response and survival among patients with higher-risk myelodysplastic syndromes treated within azacitidine or conventional care regimens in the randomized AZA-001 trial', *Haematologica*, 98: 1067-72.

Gossen, M., S. Freundlieb, G. Bender, G. Müller, W. Hillen, and H. Bujard. 1995. 'Transcriptional Activation by Tetracyclines in Mammalian Cells', *Science*, 268: 1766-69.

Granfeldt Østgård, L. S., B. C. Medeiros, H. Sengeløv, M. Nørgaard, M. K. Andersen, I. H. Dufva, L. S. Friis, E. Kjeldsen, C. W. Marcher, B. Preiss, M. Severinsen, and J. M. Nørgaard. 2015. 'Epidemiology and Clinical Significance of Secondary and Therapy-Related Acute Myeloid Leukemia: A National Population-Based Cohort Study', *Journal of Clinical Oncology*, 33: 3641-49.

Griffiths, E. A., and S. D. Gore. 2008. 'DNA methyltransferase inhibitors: Class effect or unique agents?', *Leukemia & Lymphoma*, 49: 650-51.

Grimwade, D., A. Ivey, and B. J. Huntly. 2016. 'Molecular landscape of acute myeloid leukemia in younger adults and its clinical relevance', *Blood*, 127: 29-41.

Grisolano, J. L., R. L. Wesselschmidt, P. G. Pelicci, and T. J. Ley. 1997. 'Altered myeloid development and acute leukemia in transgenic mice expressing PML-RAR alpha under control of cathepsin G regulatory sequences', *Blood*, 89: 376-87.

Grundler, R., C. Miething, C. Thiede, C. Peschel, and J. Duyster. 2005. 'FLT3-ITD and tyrosine kinase domain mutants induce 2 distinct phenotypes in a murine bone marrow transplantation model', *Blood*, 105: 4792-9.

Guan, Y., A. D. Tiwari, J. G. Phillips, M. Hasipek, D. R. Grabowski, S. Pagliuca, P. Gopal, C. M. Kerr, V. Adema, T. Radivoyevitch, Y. Parker, D. J. Lindner, M. Meggendorfer, M. Abazeed, M. A. Sekeres, O. Y. Mian, T. Haferlach, J. P. Maciejewski, and B. K. Jha. 2021. 'A Therapeutic Strategy for Preferential Targeting of TET2-Mutant and TET Dioxygenase–Deficient Cells in Myeloid Neoplasms', *Blood Cancer Discovery*, 2: 146.

Guenova, M., G. Balatzenko, and G. Mihaylov. 2013. 'Therapy-Related Acute Myeloid Leukemias', IntechOpen, Accessed 10-01-2022. https://www.intechopen.com/chapters/44517.

Guo, J. U., D. K. Ma, H. Mo, M. P. Ball, M. H. Jang, M. A. Bonaguidi, J. A. Balazer, H. L. Eaves, B. Xie, E. Ford, K. Zhang, G. L. Ming, Y. Gao, and H. Song. 2011. 'Neuronal activity modifies the DNA methylation landscape in the adult brain', *Nat Neurosci*, 14: 1345-51.

Guryanova, O. A., K. Shank, B. Spitzer, L. Luciani, R. P. Koche, F. E. Garrett-Bakelman, C. Ganzel, B. H. Durham, A. Mohanty, G. Hoermann, S. A. Rivera, A. G. Chramiec, E. Pronier, L. Bastian, M. D. Keller, D. Tovbin, E. Loizou, A. R. Weinstein, A. R. Gonzalez, Y. K. Lieu, J. M. Rowe, F. Pastore, A. S. McKenney, A. V. Krivtsov, W. R. Sperr, J. R. Cross, C. E. Mason, M. S. Tallman, M. E. Arcila, O. Abdel-Wahab, S. A. Armstrong, S. Kubicek, P. B. Staber, M. Gönen, E. M. Paietta, A. M. Melnick, S. D. Nimer, S. Mukherjee, and R. L. Levine. 2016. 'DNMT3A mutations promote anthracycline resistance in acute myeloid leukemia via impaired nucleosome remodeling', *Nature medicine*, 22: 1488-95.

Gwynn, B., K. Lueders, M. S. Sands, and E. H. Birkenmeier. 1998. 'Intracisternal A-particle element transposition into the murine beta-glucuronidase gene correlates with loss of enzyme activity: a new model for beta-glucuronidase deficiency in the C3H mouse', *Mol Cell Biol*, 18: 6474-81.

He, F. 2011. 'BCA (Bicinchoninic Acid) Protein Assay', *Bio-protocol*, 1: e44.

He, L. Z., C. Tribioli, R. Rivi, D. Peruzzi, P. G. Pelicci, V. Soares, G. Cattoretti, and P. P. Pandolfi. 1997. 'Acute leukemia with promyelocytic features in PML/RARalpha transgenic mice', *Proc Natl Acad Sci U S A*, 94: 5302-7.

He, P. F., J. D. Zhou, D. M. Yao, J. C. Ma, X. M. Wen, Z. H. Zhang, X. Y. Lian, Z. J. Xu, J. Qian, and J. Lin. 2017. 'Efficacy and safety of decitabine in treatment of elderly patients with acute myeloid leukemia: A systematic review and meta-analysis', *Oncotarget*, 8: 41498-507.

Heath, E. M., S. M. Chan, M. D. Minden, T. Murphy, L. I. Shlush, and A. D. Schimmer. 2017. 'Biological and clinical consequences of NPM1 mutations in AML', *Leukemia*, 31: 798-807.

Hellman, A., and A. Chess. 2007. 'Gene body-specific methylation on the active X chromosome', *Science*, 315: 1141-3.

Hemsing, A. L., R. Hovland, G. Tsykunova, and H. Reikvam. 2019. 'Trisomy 8 in acute myeloid leukemia', *Expert Review of Hematology*, 12: 947-58.

Higuchi, M., D. O'Brien, P. Kumaravelu, N. Lenny, E. J. Yeoh, and J. R. Downing. 2002. 'Expression of a conditional AML1-ETO oncogene bypasses embryonic lethality and establishes a murine model of human t(8;21) acute myeloid leukemia', *Cancer Cell*, 1: 63-74.

Hills, R. K., S. Castaigne, F. R. Appelbaum, J. Delaunay, S. Petersdorf, M. Othus, E. H. Estey, H. Dombret, S. Chevret, N. Ifrah, J.-Y. Cahn, C. Récher, L. Chilton, A. V. Moorman, and A. K. Burnett. 2014. 'Addition of gemtuzumab ozogamicin to induction chemotherapy in adult patients with acute myeloid leukaemia: a meta-analysis of individual patient data from randomised controlled trials', *The Lancet. Oncology*, 15: 986-96.

Hirsch, C. M., A. Nazha, K. Kneen, M. E. Abazeed, M. Meggendorfer, B. P. Przychodzen, N. Nadarajah, V. Adema, Y. Nagata, A. Goyal, H. Awada, M. F. Asad, V. Visconte, Y. Guan, M. A. Sekeres, R. Olinski, B. K. Jha, T. LaFramboise, T. Radivoyevitch, T. Haferlach, and J. P. Maciejewski. 2018. 'Consequences of mutant TET2 on clonality and subclonal hierarchy', *Leukemia*, 32: 1751-61.

HMRN(2010-2016). ' Haematological Malignancy Research Network (HMRN) '.

Hoerster, K., M. Uhrberg, C. Wiek, P. A. Horn, H. Hanenberg, and S. Heinrichs. 2021. 'HLA Class I Knockout Converts Allogeneic Primary NK Cells Into Suitable Effectors for “Off-the-Shelf” Immunotherapy', *Frontiers in Immunology*, 11.

Hollenbach, P. W., A. N. Nguyen, H. Brady, M. Williams, Y. Ning, N. Richard, L. Krushel, S. L. Aukerman, C. Heise, and K. J. MacBeth. 2010. 'A comparison of azacitidine and decitabine activities in acute myeloid leukemia cell lines', *PLoS One*, 5: e9001-e01.

Hou, H.-A., T.-C. Huang, L.-I. Lin, C.-Y. Liu, C.-Y. Chen, W.-C. Chou, J.-L. Tang, M.-H. Tseng, C.-F. Huang, Y.-C. Chiang, F.-Y. Lee, M.-C. Liu, M. Yao, S.-Y. Huang, B.-S. Ko, S.-C. Hsu, S.-J. Wu, W. Tsay, Y.-C. Chen, and H.-F. Tien. 2010. 'WT1 mutation in 470 adult patients with acute myeloid leukemia: stability during disease evolution and implication of its incorporation into a survival scoring system', *Blood*, 115: 5222-31.

Hou, H.-A., and H.-F. Tien. 2020. 'Genomic landscape in acute myeloid leukemia and its implications in risk classification and targeted therapies', *Journal of Biomedical Science*, 27: 81.

Hu, L., Z. Li, J. Cheng, Q. Rao, W. Gong, M. Liu, Y. G. Shi, J. Zhu, P. Wang, and Y. Xu. 2013. 'Crystal structure of TET2-DNA complex: insight into TET-mediated 5mC oxidation', *Cell*, 155: 1545-55.

Huff, L. M., J. S. Lee, R. W. Robey, and T. Fojo. 2006. 'Characterization of gene rearrangements leading to activation of MDR-1', *J Biol Chem*, 281: 36501-9.

Huff, L. M., Z. Wang, A. Iglesias, T. Fojo, and J. S. Lee. 2005. 'Aberrant transcription from an unrelated promoter can result in MDR-1 expression following drug selection in vitro and in relapsed lymphoma samples', *Cancer Res*, 65: 11694-703.

Hunter, A. M., and D. A. Sallman. 2019. 'Current status and new treatment approaches in TP53 mutated AML', *Best Practice & Research Clinical Haematology*, 32: 134-44.

Ianevski, A., A. K. Giri, and T. Aittokallio. 2020. 'SynergyFinder 2.0: visual analytics of multi-drug combination synergies', *Nucleic Acids Research*, 48: W488-W93.

Inoue, S., F. Lemonnier, and T. W. Mak. 2016. 'Roles of IDH1/2 and TET2 mutations in myeloid disorders', *International Journal of Hematology*, 103: 627-33.

Ito, S., L. Shen, Q. Dai, S. C. Wu, L. B. Collins, J. A. Swenberg, C. He, and Y. Zhang. 2011. 'Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine', *Science*, 333: 1300-3.

Itzykson, R., O. Kosmider, T. Cluzeau, V. Mansat-De Mas, F. Dreyfus, O. Beyne-Rauzy, B. Quesnel, N. Vey, V. Gelsi-Boyer, S. Raynaud, C. Preudhomme, L. Adès, P. Fenaux, M. Fontenay, and M. on behalf of the Groupe Francophone des. 2011. 'Impact of TET2 mutations on response rate to azacitidine in myelodysplastic syndromes and low blast count acute myeloid leukemias', *Leukemia*, 25: 1147-52.

Jankowska, A. M., H. Szpurka, R. V. Tiu, H. Makishima, M. Afable, J. Huh, C. L. O'Keefe, R. Ganetzky, M. A. McDevitt, and J. P. Maciejewski. 2009. 'Loss of heterozygosity 4q24 and TET2 mutations associated with myelodysplastic/myeloproliferative neoplasms', *Blood*, 113: 6403-10.

Jimenez, J. J., R. S. Chale, A. C. Abad, and A. V. Schally. 2020. 'Acute promyelocytic leukemia (APL): a review of the literature', *Oncotarget*, 11.

Jin, J., Y. Xu, L. Huo, L. Ma, A. W. Scott, M. P. Pizzi, Y. Li, Y. Wang, X. Yao, S. Song, and J. A. Ajani. 2020. 'An improved strategy for CRISPR/Cas9 gene knockout and subsequent wildtype and mutant gene rescue', *PLoS One*, 15: e0228910-e10.

Jing, C.-B., C. Fu, N. Prutsch, M. Wang, S. He, and A. T. Look. 2020. 'Synthetic lethal targeting of TET2-mutant hematopoietic stem and progenitor cells (HSPCs) with TOP1-targeted drugs and PARP1 inhibitors', *Leukemia*, 34: 2992-3006.

Jolliffe, I. T., and J. Cadima. 2016. 'Principal component analysis: a review and recent developments', *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374: 20150202.

Kaelin, William G., Jr., and Steven L. McKnight. 'Influence of Metabolism on Epigenetics and Disease', *Cell*, 153: 56-69.

Kang, J., N. Brajanovski, K. T. Chan, J. Xuan, R. B. Pearson, and E. Sanij. 2021. 'Ribosomal proteins and human diseases: molecular mechanisms and targeted therapy', *Signal Transduction and Targeted Therapy*, 6: 323.

Kantarjian, H. M., X. G. Thomas, A. Dmoszynska, A. Wierzbowska, G. Mazur, J. Mayer, J.-P. Gau, W.-C. Chou, R. Buckstein, J. Cermak, C.-Y. Kuo, A. Oriol, F. Ravandi, S. Faderl, J. Delaunay, D. Lysák, M. Minden, and C. Arthur. 2012. 'Multicenter, randomized, open-label, phase III trial of decitabine versus patient choice, with physician advice, of either supportive care or low-dose cytarabine for the treatment of older patients with newly diagnosed acute myeloid leukemia', *J Clin Oncol*, 30: 2670-77.

Karamyshev, A. L., and Z. N. Karamysheva. 2018. 'Lost in Translation: Ribosome-Associated mRNA and Protein Quality Controls', *Frontiers in Genetics*, 9.

Kawashima, N., Y. Ishikawa, J. H. Kim, Y. Ushijima, A. Akashi, Y. Yamaguchi, H. Hattori, M. Nakashima, S. Ikeno, R. Kihara, T. Nishiyama, T. Morishita, K. Watamoto, Y. Ozawa, K. Kitamura, and H. Kiyoi. 2022. 'Comparison of clonal architecture between primary and immunodeficient mouse-engrafted acute myeloid leukemia cells', *Nat Commun*, 13: 1624.

Kelly, L. M., and D. G. Gilliland. 2002. 'Genetics of myeloid leukemias', *Annu Rev Genomics Hum Genet*, 3: 179-98.

Khatodia, S., K. Bhatotia, N. Passricha, S. M. P. Khurana, and N. Tuteja. 2016. 'The CRISPR/Cas Genome-Editing Tool: Application in Improvement of Crops', *Frontiers in Plant Science*, 7.

Kihara, R., Y. Nagata, H. Kiyoi, T. Kato, E. Yamamoto, K. Suzuki, F. Chen, N. Asou, S. Ohtake, S. Miyawaki, Y. Miyazaki, T. Sakura, Y. Ozawa, N. Usui, H. Kanamori, T. Kiguchi, K. Imai, N. Uike, F. Kimura, K. Kitamura, C. Nakaseko, M. Onizuka, A. Takeshita, F. Ishida, H. Suzushima, Y. Kato, H. Miwa, Y. Shiraishi, K. Chiba, H. Tanaka, S. Miyano, S. Ogawa, and T. Naoe. 2014. 'Comprehensive analysis of genetic alterations and their prognostic impacts in adult acute myeloid leukemia patients', *Leukemia*, 28: 1586-95.

Kim, H. G., J. LeGrand, C. S. Swindle, H. J. Nick, R. A. Oster, D. Chen, S. Purohit-Ghelani, C. V. Cotta, R. Ko, L. Gartland, V. Reddy, S. W. Hiebert, A. D. Friedman, and C. A. Klug. 2017. 'The assembly competence domain is essential for inv(16)-associated acute myeloid leukemia', *Leukemia*, 31: 2267-71.

Kim, J., R. Y. Haddad, and E. Atallah. 2012. 'Myeloproliferative Neoplasms', *Disease-a-Month*, 58: 177-94.

Kimura, H., and K. Shiota. 2003. 'Methyl-CpG-binding protein, MeCP2, is a target molecule for maintenance DNA methyltransferase, Dnmt1', *J Biol Chem*, 278: 4806-12.

Kisseleva, T., S. Bhattacharya, J. Braunstein, and C. W. Schindler. 2002. 'Signaling through the JAK/STAT pathway, recent advances and future challenges', *Gene*, 285: 1-24.

Ko, M., J. An, H. S. Bandukwala, L. Chavez, T. Aijo, W. A. Pastor, M. F. Segal, H. Li, K. P. Koh, H. Lahdesmaki, P. G. Hogan, L. Aravind, and A. Rao. 2013. 'Modulation of TET2 expression and 5-methylcytosine oxidation by the CXXC domain protein IDAX', *Nature*, 497: 122-6.

Ko, M., Y. Huang, A. M. Jankowska, U. J. Pape, M. Tahiliani, H. S. Bandukwala, J. An, E. D. Lamperti, K. P. Koh, R. Ganetzky, X. S. Liu, L. Aravind, S. Agarwal, J. P. Maciejewski, and A. Rao. 2010. 'Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2', *Nature*, 468: 839-43.

Kosmider, O., E. Delabesse, V. M. de Mas, P. Cornillet-Lefebvre, O. Blanchet, A. Delmer, C. Recher, S. Raynaud, D. Bouscary, F. Viguié, C. Lacombe, O. A. Bernard, N. Ifrah, F. Dreyfus, and M. Fontenay. 2011. 'TET2 mutations in secondary acute myeloid leukemias: a French retrospective study', *Haematologica*, 96: 1059-63.

Koya, J., K. Kataoka, T. Sato, M. Bando, Y. Kato, T. Tsuruta-Kishino, H. Kobayashi, K. Narukawa, H. Miyoshi, K. Shirahige, and M. Kurokawa. 2016. 'DNMT3A R882 mutants interact with polycomb proteins to block haematopoietic stem and leukaemic cell differentiation', *Nat Commun*, 7: 10924-24.

Kristensen, L. S., and L. L. Hansen. 2009. 'PCR-Based Methods for Detecting Single-Locus DNA Methylation Biomarkers in Cancer Diagnostics, Prognostics, and Response to Treatment', *Clinical Chemistry*, 55: 1471-83.

Krivtsov, A. V., and S. A. Armstrong. 2007. 'MLL translocations, histone modifications and leukaemia stem-cell development', *Nature Reviews Cancer*, 7: 823-33.

Kroeze, L. I., M. G. Aslanyan, A. van Rooij, T. N. Koorenhof-Scheele, M. Massop, T. Carell, J. B. Boezeman, J.-P. Marie, C. J. M. Halkes, T. de Witte, G. Huls, S. Suciu, R. A. Wevers, B. A. van der Reijden, J. H. Jansen, E. L. G. on behalf of the, and Gimema. 2014. 'Characterization of acute myeloid leukemia based on levels of global hydroxymethylation', *Blood*, 124: 1110-18.

Kukurba, K. R., and S. B. Montgomery. 2015. 'RNA Sequencing and Analysis', *Cold Spring Harbor protocols*, 2015: 951-69.

Kuster, J. E., M. H. Guarnieri, J. G. Ault, L. Flaherty, and P. J. Swiatek. 1997. 'IAP insertion in the murine LamB3 gene results in junctional epidermolysis bullosa', *Mamm Genome*, 8: 673-81.

Ladines-Castro, W., G. Barragán-Ibañez, M. A. Luna-Pérez, A. Santoyo-Sánchez, J. Collazo-Jaloma, E. Mendoza-García, and C. O. Ramos-Peñafiel. 2016. 'Morphology of leukaemias', *Revista Médica del Hospital General de México*, 79: 107-13.

Lagunas-Rangel, F. A., V. Chávez-Valencia, M. Á. Gómez-Guijosa, and C. Cortes-Penagos. 2017. 'Acute Myeloid Leukemia-Genetic Alterations and Their Clinical Prognosis', *International journal of hematology-oncology and stem cell research*, 11: 328-39.

Lainey, E., M. Sébert, S. Thépot, M. Scoazec, C. Bouteloup, C. Leroy, S. De Botton, L. Galluzzi, P. Fenaux, and G. Kroemer. 2012. 'Erlotinib antagonizes ABC transporters in acute myeloid leukemia', *Cell cycle (Georgetown, Tex.)*, 11: 4079-92.

Lainey, E., A. Wolfromm, N. Marie, D. Enot, M. Scoazec, C. Bouteloup, C. Leroy, J. B. Micol, S. De Botton, L. Galluzzi, P. Fenaux, and G. Kroemer. 2013. 'Azacytidine and erlotinib exert synergistic effects against acute myeloid leukemia', *Oncogene*, 32: 4331-42.

Lam, K., and D.-E. Zhang. 2012. 'RUNX1 and RUNX1-ETO: roles in hematopoiesis and leukemogenesis', *Frontiers in bioscience (Landmark edition)*, 17: 1120-39.

Lee, T. T., and M. R. Karon. 1976. 'Inhibition of protein synthesis in 5-azacytidine-treated HeLa cells', *Biochem Pharmacol*, 25: 1737-42.

Lever, J., M. Krzywinski, and N. Altman. 2017. 'Principal component analysis', *Nature Methods*, 14: 641-42.

Levine, A. J. 1997. 'p53, the cellular gatekeeper for growth and division', *Cell*, 88: 323-31.

Ley, T. J., L. Ding, M. J. Walter, M. D. McLellan, T. Lamprecht, D. E. Larson, C. Kandoth, J. E. Payton, J. Baty, J. Welch, C. C. Harris, C. F. Lichti, R. R. Townsend, R. S. Fulton, D. J. Dooling, D. C. Koboldt, H. Schmidt, Q. Zhang, J. R. Osborne, L. Lin, M. O'Laughlin, J. F. McMichael, K. D. Delehaunty, S. D. McGrath, L. A. Fulton, V. J. Magrini, T. L. Vickery, J. Hundal, L. L. Cook, J. J. Conyers, G. W. Swift, J. P. Reed, P. A. Alldredge, T. Wylie, J. Walker, J. Kalicki, M. A. Watson, S. Heath, W. D. Shannon, N. Varghese, R. Nagarajan, P. Westervelt, M. H. Tomasson, D. C. Link, T. A. Graubert, J. F. DiPersio, E. R. Mardis, and R. K. Wilson. 2010. 'DNMT3A Mutations in Acute Myeloid Leukemia', *New England Journal of Medicine*, 363: 2424-33.

Ley, T. J., C. Miller, L. Ding, B. J. Raphael, A. J. Mungall, A. Robertson, K. Hoadley, T. J. Triche, Jr., P. W. Laird, J. D. Baty, L. L. Fulton, R. Fulton, S. E. Heath, J. Kalicki-Veizer, C. Kandoth, J. M. Klco, D. C. Koboldt, K. L. Kanchi, S. Kulkarni, T. L. Lamprecht, D. E. Larson, L. Lin, C. Lu, M. D. McLellan, J. F. McMichael, J. Payton, H. Schmidt, D. H. Spencer, M. H. Tomasson, J. W. Wallis, L. D. Wartman, M. A. Watson, J. Welch, M. C. Wendl, A. Ally, M. Balasundaram, I. Birol, Y. Butterfield, R. Chiu, A. Chu, E. Chuah, H. J. Chun, R. Corbett, N. Dhalla, R. Guin, A. He, C. Hirst, M.

Hirst, R. A. Holt, S. Jones, A. Karsan, D. Lee, H. I. Li, M. A. Marra, M. Mayo, R. A. Moore, K. Mungall, J. Parker, E. Pleasance, P. Plettner, J. Schein, D. Stoll, L. Swanson, A. Tam, N. Thiessen, R. Varhol, N. Wye, Y. Zhao, S. Gabriel, G. Getz, C. Sougnez, L. Zou, M. D. Leiserson, F. Vandin, H. T. Wu, F. Applebaum, S. B. Baylin, R. Akbani, B. M. Broom, K. Chen, T. C. Motter, K. Nguyen, J. N. Weinstein, N. Zhang, M. L. Ferguson, C. Adams, A. Black, J. Bowen, J. Gastier-Foster, T. Grossman, T. Lichtenberg, L. Wise, T. Davidsen, J. A. Demchok, K. R. Shaw, M. Sheth, H. J. Sofia, L. Yang, J. R. Downing, and G. Eley. 2013. 'Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia', *N Engl J Med*, 368: 2059-74.

Li, M., L. Jones, C. Gaillard, M. Binnewies, R. Ochoa, E. Garcia, V. Lam, G. Wei, W. Yang, C. Lobe, M. Hermiston, E. Passegué, and S. C. Kogan. 2013. 'Initially disadvantaged, TEL-AML1 cells expand and initiate leukemia in response to irradiation and cooperating mutations', *Leukemia*, 27: 1570-73.

Li, Z., X. Cai, C.-L. Cai, J. Wang, W. Zhang, B. E. Petersen, F.-C. Yang, and M. Xu. 2011. 'Deletion of Tet2 in mice leads to dysregulated hematopoietic stem cells and subsequent development of myeloid malignancies', *Blood*, 118: 4509-18.

Lian, C. G., Y. Xu, C. Ceol, F. Wu, A. Larson, K. Dresser, W. Xu, L. Tan, Y. Hu, Q. Zhan, C.-W. Lee, D. Hu, B. Q. Lian, S. Kleffel, Y. Yang, J. Neiswender, A. J. Khorasani, R. Fang, C. Lezcano, L. M. Duncan, R. A. Scolyer, J. F. Thompson, H. Kakavand, Y. Houvras, L. I. Zon, M. C. Mihm, Jr., U. B. Kaiser, T. Schatton, B. A. Woda, G. F. Murphy, and Y. G. Shi. 2012. 'Loss of 5-hydroxymethylcytosine is an epigenetic hallmark of melanoma', *Cell*, 150: 1135-46.

Liao, Y., G. K. Smyth, and W. Shi. 2019. 'The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads', *Nucleic Acids Research*, 47: e47-e47.

Lichtman, M. A. 2015. 'Clonal hematopoiesis: a "CHIP" off the old block', *Blood*, 126: 1-2.

Lin, Y., Z. Lin, K. Cheng, Z. Fang, Z. Li, Y. Luo, and B. Xu. 2017. 'Prognostic role of TET2 deficiency in myelodysplastic syndromes: A meta-analysis', *Oncotarget*, 8: 43295-305.

List, A. F., K. J. Kopecky, C. L. Willman, D. R. Head, D. L. Persons, M. L. Slovak, R. Dorr, C. Karanes, H. E. Hynes, J. H. Doroshow, M. Shurafa, and F. R. Appelbaum. 2001. 'Benefit of cyclosporine modulation of drug resistance in patients with poor-risk acute myeloid leukemia: a Southwest Oncology Group study', *Blood*, 98: 3212-20.

Liu, X., and Y. Gong. 2019. 'Isocitrate dehydrogenase inhibitors in acute myeloid leukemia', *Biomarker research*, 7: 22-22.

Lopes, E. C., E. Valls, M. E. Figueroa, A. Mazur, F. G. Meng, G. Chiosis, P. W. Laird, N. Schreiber-Agus, J. M. Greally, E. Prokhortchouk, and A. Melnick. 2008. 'Kaiso contributes to DNA methylation-dependent silencing of tumor suppressor genes in colon cancer cell lines', *Cancer Res*, 68: 7258-63.

Love, M. I., W. Huber, and S. Anders. 2014. 'Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2', *Genome Biol*, 15: 550.

Lu, L. J., and K. Randerath. 1980. 'Mechanism of 5-azacytidine-induced transfer RNA cytosine-5-methyltransferase deficiency', *Cancer Res*, 40: 2701-5.

Lubbert, M., and M. Minden. 2005. 'Decitabine in acute myeloid leukemia', *Semin Hematol*, 42: S38-42.

Lübbert, M., S. Suciu, L. Baila, B. H. Rüter, U. Platzbecker, A. Giagounidis, D. Selleslag, B. Labar, U. Germing, H. R. Salih, F. Beeldens, P. Muus, K. H. Pflüger, C. Coens, A. Hagemeijer, H. Eckart Schaefer, A. Ganser, C. Aul, T. de Witte, and P. W. Wijermans. 2011. 'Low-dose decitabine versus best supportive care in elderly patients with intermediate- or high-risk myelodysplastic syndrome (MDS) ineligible for intensive chemotherapy: final results of the randomized phase III study of the European Organisation for Research and Treatment of Cancer Leukemia Group and the German MDS Study Group', *J Clin Oncol*, 29: 1987-96.

Lv, L., Q. Wang, Y. Xu, L.-C. Tsao, T. Nakagawa, H. Guo, L. Su, and Y. Xiong. 2018. 'Vpr Targets TET2 for Degradation by CRL4VprBP E3 Ligase to Sustain IL-6 Expression and Enhance HIV-1 Replication', *Mol Cell*, 70: 961-70.e5.

Lv, L., J. Yu, and Z. Qi. 2020. 'Acute myeloid leukemia with inv(16)(p13.1q22) and deletion of the 5'MYH11/3'CBFB gene fusion: a report of two cases and literature review', *Molecular Cytogenetics*, 13: 4.

Lyu, Y., J. Lou, Y. Yang, J. Feng, Y. Hao, S. Huang, L. Yin, J. Xu, D. Huang, B. Ma, D. Zou, Y. Wang, Y. Zhang, B. Zhang, P. Chen, K. Yu, E. W. F. Lam, X. Wang, Q. Liu, J. Yan, and B. Jin. 2017. 'Dysfunction of the WT1-MEG3 signaling promotes AML leukemogenesis via p53-dependent and -independent pathways', *Leukemia*, 31: 2543-51.

Mackinnon, R. N., M. Wall, A. Zordan, S. Nutalapati, B. Mercer, J. Peverall, and L. J. Campbell. 2013. 'Genome organization and the role of centromeres in evolution of the erythroleukaemia cell line HEL', *Evolution, medicine, and public health*, 2013: 225-40.

Maksimovic, J., B. Phipson, and A. Oshlack. 2016. 'A cross-package Bioconductor workflow for analysing methylation array data', *F1000Res*, 5: 1281.

Malik, P., and A. F. Cashen. 2014. 'Decitabine in the treatment of acute myeloid leukemia in elderly patients', *Cancer management and research*, 6: 53-61.

Mannelli, F., V. Ponziani, S. Bencini, M. I. Bonetti, M. Benelli, I. Cutini, G. Gianfaldoni, B. Scappini, F. Pancani, M. Piccini, T. Rondelli, R. Caporale, A. M. Gelli, B. Peruzzi, M. Chiarini, E. Borlenghi, O. Spinelli, D. Giupponi, P. Zanghì, R. Bassan, A. Rambaldi, G. Rossi, and A. Bosi. 2017. 'CEBPA-double-mutated acute myeloid leukemia displays a unique phenotypic profile: a reliable screening method and insight into biological features', *Haematologica*, 102: 529-40.

Martelli, M. P., P. Sportoletti, E. Tiacci, M. F. Martelli, and B. Falini. 2013. 'Mutational landscape of AML with normal cytogenetics: Biological and clinical implications', *Blood Reviews*, 27: 13-22.

Martin, P., and T. Papayannopoulou. 1982. 'HEL cells: a new human erythroleukemia cell line with spontaneous and induced globin expression', *Science*, 216: 1233-5.

McCormick, S. R., M. J. McCormick, P. S. Grutkoski, G. S. Ducker, N. Banerji, R. R. Higgins, J. R. Mendiola, and J. J. Reinartz. 2010. 'FLT3 mutations at diagnosis and relapse in acute myeloid leukemia: cytogenetic and pathologic correlations, including cuplike blast morphology', *Arch Pathol Lab Med*, 134: 1143-51.

Medeiros, B. C., A. T. Fathi, C. D. DiNardo, D. A. Pollyea, S. M. Chan, and R. Swords. 2017. 'Isocitrate dehydrogenase mutations in myeloid malignancies', *Leukemia*, 31: 272-81.

Medinger, M., and J. R. Passweg. 2017. 'Acute myeloid leukaemia genomics', *Br J Haematol*, 179: 530-42.

Megías-Vericat, J. E., O. Ballesta-López, E. Barragán, and P. Montesinos. 2019. 'IDH1-mutated relapsed or refractory AML: current challenges and future prospects', *Blood Lymphat Cancer*, 9: 19-32.

Megías-Vericat, J. E., L. Rojas, M. J. Herrero, V. Bosó, P. Montesinos, F. Moscardó, J. L. Poveda, M. Á. Sanz, and S. F. Aliño. 2015. 'Influence of ABCB1 polymorphisms upon the effectiveness of standard treatment for acute myeloid leukemia: A systematic review and meta-analysis of observational studies', *The Pharmacogenomics Journal*, 15: 109-18.

Melamed, P., Y. Yosefzon, C. David, A. Tsukerman, and L. Pnueli. 2018. 'Tet Enzymes, Variants, and Differential Effects on Function', *Frontiers in Cell and Developmental Biology*, 6: 22.

Messingerova, L., D. Imrichova, H. Kavcova, K. Turakova, A. Breier, and Z. Sulova. 2015. 'Acute myeloid leukemia cells MOLM-13 and SKM-1 established for resistance by

azacytidine are crossresistant to P-glycoprotein substrates', *Toxicology in Vitro*, 29: 1405-15.

Metzeler, K. H., K. Maharry, M. D. Radmacher, K. Mrozek, D. Margeson, H. Becker, J. Curfman, K. B. Holland, S. Schwind, S. P. Whitman, Y. Z. Wu, W. Blum, B. L. Powell, T. H. Carter, M. Wetzler, J. O. Moore, J. E. Kolitz, M. R. Baer, A. J. Carroll, R. A. Larson, M. A. Caligiuri, G. Marcucci, and C. D. Bloomfield. 2011. 'TET2 mutations improve the new European LeukemiaNet risk classification of acute myeloid leukemia: a Cancer and Leukemia Group B study', *J Clin Oncol*, 29: 1373-81.

Metzeler, K. H., K. Maharry, M. D. Radmacher, K. Mrózek, D. Margeson, H. Becker, J. Curfman, K. B. Holland, S. Schwind, S. P. Whitman, Y. Z. Wu, W. Blum, B. L. Powell, T. H. Carter, M. Wetzler, J. O. Moore, J. E. Kolitz, M. R. Baer, A. J. Carroll, R. A. Larson, M. A. Caligiuri, G. Marcucci, and C. D. Bloomfield. 2011. 'TET2 mutations improve the new European LeukemiaNet risk classification of acute myeloid leukemia: a Cancer and Leukemia Group B study', *J Clin Oncol*, 29: 1373-81.

Meyer, S. C., and R. L. Levine. 2014. 'Translational implications of somatic genomics in acute myeloid leukaemia', *Lancet Oncol*, 15: e382-94.

Mhawech, P., and A. Saleem. 2001. 'Myelodysplastic syndrome: review of the cytogenetic and molecular data', *Critical Reviews in Oncology/Hematology*, 40: 229-38.

Michaud, E. J., M. J. van Vugt, S. J. Bultman, H. O. Sweet, M. T. Davisson, and R. P. Woychik. 1994. 'Differential expression of a new dominant agouti allele (Aiapy) is correlated with methylation state and is influenced by parental lineage', *Genes Dev*, 8: 1463-72.

Mitin, N., K. L. Rossman, and C. J. Der. 2005. 'Signaling interplay in Ras superfamily function', *Curr Biol*, 15: R563-74.

Mohammad, A. A. 2018. 'Myelodysplastic syndrome from theoretical review to clinical application view', *Oncology reviews*, 12: 397-97.

Momparler, R. L. 2013. 'Optimization of cytarabine (ARA-C) therapy for acute myeloid leukemia', *Experimental Hematology & Oncology*, 2: 20.

Monika Belickova, M., M. D. Merkerova, H. Votavova, J. Valka, J. Vesela, B. Pejsova, H. Hajkova, J. Klema, J. Cermak, and A. Jonasova. 2016. 'Up-regulation of ribosomal genes is associated with a poor response to azacitidine in myelodysplasia and related neoplasms', *International Journal of Hematology*, 104: 566-73.

Moore, L. D., T. Le, and G. Fan. 2013. 'DNA Methylation and Its Basic Function', *Neuropsychopharmacology*, 38: 23-38.

Moore, P. B., and T. A. Steitz. 2003. 'The Structural Basis of Large Ribosomal Subunit Function', *Annual review of biochemistry*, 72: 813-50.

Moran-Crusio, K., L. Reavie, A. Shih, O. Abdel-Wahab, D. Ndiaye-Lobry, C. Lobry, M. E. Figueroa, A. Vasanthakumar, J. Patel, X. Zhao, F. Perna, S. Pandey, J. Madzo, C. Song, Q. Dai, C. He, S. Ibrahim, M. Beran, J. Zavadil, S. D. Nimer, A. Melnick, L. A. Godley, I. Aifantis, and R. L. Levine. 2011. 'Tet2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation', *Cancer Cell*, 20: 11-24.

Morinishi, L., K. Kochanowski, R. L. Levine, L. F. Wu, and S. J. Altschuler. 2020. 'Loss of TET2 Affects Proliferation and Drug Sensitivity through Altered Dynamics of Cell-State Transitions', *Cell Systems*, 11: 86-94.e5.

Morita, K., F. Wang, K. Jahn, T. Hu, T. Tanaka, Y. Sasaki, J. Kuipers, S. Loghavi, S. A. Wang, Y. Yan, K. Furudate, J. Matthews, L. Little, C. Gumbs, J. Zhang, X. Song, E. Thompson, K. P. Patel, C. E. Bueso-Ramos, C. D. DiNardo, F. Ravandi, E. Jabbour, M. Andreeff, J. Cortes, K. Bhalla, G. Garcia-Manero, H. Kantarjian, M. Konopleva, D. Nakada, N. Navin, N. Beerenwinkel, P. A. Futreal, and K. Takahashi. 2020. 'Clonal evolution of acute myeloid leukemia revealed by high-throughput single-cell genomics', *Nat Commun*, 11: 5327.

Mrozek, K., G. Marcucci, D. Nicolet, K. S. Maharry, H. Becker, S. P. Whitman, K. H. Metzeler, S. Schwind, Y. Z. Wu, J. Kohlschmidt, M. J. Pettenati, N. A. Heerema, A. W. Block, S. R. Patil, M. R. Baer, J. E. Kolitz, J. O. Moore, A. J. Carroll, R. M. Stone, R. A. Larson, and C. D. Bloomfield. 2012. 'Prognostic significance of the European LeukemiaNet standardized system for reporting cytogenetic and molecular alterations in adults with acute myeloid leukemia', *J Clin Oncol*, 30: 4515-23.

Muller, C. I., B. Ruter, H. P. Koeffler, and M. Lubbert. 2006. 'DNA hypermethylation of myeloid cells, a novel therapeutic target in MDS and AML', *Curr Pharm Biotechnol*, 7: 315-21.

Murphy, T., and K. W. L. Yee. 2017. 'Cytarabine and daunorubicin for the treatment of acute myeloid leukemia', *Expert Opinion on Pharmacotherapy*, 18: 1765-80.

Nakagawa, T., L. Lv, M. Nakagawa, Y. Yu, C. Yu, A. C. D'Alessio, K. Nakayama, H. Y. Fan, X. Chen, and Y. Xiong. 2015. 'CRL4(VprBP) E3 ligase promotes monoubiquitylation and chromatin binding of TET dioxygenases', *Mol Cell*, 57: 247-60.

Nakagawa, T., and S. Matozaki. 1995. 'The SKM-1 leukemic cell line established from a patient with progression to myelomonocytic leukemia in myelodysplastic syndrome (MDS)-contribution to better understanding of MDS', *Leuk Lymphoma*, 17: 335-9.

Nakagawa, T., S. Matozaki, T. Murayama, R. Nishimura, M. Tsutsumi, R. Kawaguchi, Y. Yokoyama, K. Hikiji, T. Isobe, and K. Chihara. 1993. 'Establishment of a leukaemic

cell line from a patient with acquisition of chromosomal abnormalities during disease progression in myelodysplastic syndrome', *Br J Haematol*, 85: 469-76.

Nakagawa, T., S. Saitoh, S. Imoto, M. Itoh, M. Tsutsumi, K. Hikiji, Y. Nakao, and T. Fujita. 1991. 'Loss of multiple point mutations of RAS genes associated with acquisition of chromosomal abnormalities during disease progression in myelodysplastic syndrome ', *Br J Haematol*, 77: 250-52.

Nakajima, H., and H. Kunimoto. 2014. 'TET2 as an epigenetic master regulator for normal and malignant hematopoiesis', *Cancer Sci*, 105: 1093-9.

Nan, X., H. H. Ng, C. A. Johnson, C. D. Laherty, B. M. Turner, R. N. Eisenman, and A. Bird. 1998. 'Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex', *Nature*, 393: 386-9.

Nerlov, C. 2004. 'C/EBPα mutations in acute myeloid leukaemias', *Nature Reviews Cancer*, 4: 394-400.

Ng, H. H., Y. Zhang, B. Hendrich, C. A. Johnson, B. M. Turner, H. Erdjument-Bromage, P. Tempst, D. Reinberg, and A. Bird. 1999. 'MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex', *Nat Genet*, 23: 58-61.

Nibourel, O., O. Kosmider, M. Cheok, N. Boissel, A. Renneville, N. Philippe, H. Dombret, F. Dreyfus, B. Quesnel, S. Geffroy, S. Quentin, C. Roche-Lestienne, J. M. Cayuela, C. Roumier, P. Fenaux, W. Vainchenker, O. A. Bernard, J. Soulier, M. Fontenay, and C. Preudhomme. 2010. 'Incidence and prognostic value of TET2 alterations in de novo acute myeloid leukemia achieving complete remission', *Blood*, 116: 1132-5.

Nickerson, M. L., S. Das, K. M. Im, S. Turan, S. I. Berndt, H. Li, H. Lou, S. A. Brodie, J. N. Billaud, T. Zhang, A. J. Bouk, D. Butcher, Z. Wang, L. Sun, K. Misner, W. Tan, A. Esnakula, D. Esposito, W. Y. Huang, R. N. Hoover, M. A. Tucker, J. R. Keller, J. Boland, K. Brown, S. K. Anderson, L. E. Moore, W. B. Isaacs, S. J. Chanock, M. Yeager, M. Dean, and T. Andresson. 2017. 'TET2 binds the androgen receptor and loss is associated with prostate cancer', *Oncogene*, 36: 2172-83.

Nickerson, M. L., K. M. Im, K. J. Misner, W. Tan, H. Lou, B. Gold, D. W. Wells, H. C. Bravo, K. M. Fredrikson, T. T. Harkins, P. Milos, B. Zbar, W. M. Linehan, M. Yeager, T. Andresson, M. Dean, and G. S. Bova. 2013. 'Somatic alterations contributing to metastasis of a castration-resistant prostate cancer', *Hum Mutat*, 34: 1231-41.

O'Donohue, M. F., V. Choesmel, M. Faubladier, G. Fichant, and P. E. Gleizes. 2010. 'Functional dichotomy of ribosomal proteins during the synthesis of mammalian 40S ribosomal subunits', *J Cell Biol*, 190: 853-66.

Okano, M., D. W. Bell, D. A. Haber, and E. Li. 1999. 'DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development', *Cell*, 99: 247-57.

Olarte Carrillo, I., A. I. García Laguna, A. De la Cruz Rosas, C. O. Ramos Peñafiel, J. Collazo Jaloma, and A. Martínez Tovar. 2021. 'High expression levels and the C3435T SNP of the ABCB1 gene are associated with lower survival in adult patients with acute myeloblastic leukemia in Mexico City', *BMC Medical Genomics*, 14: 251.

Olesen, L. H., A. Aggerholm, B. L. Andersen, C. G. Nyvold, P. Guldberg, J. M. Nørgaard, and P. Hokland. 2005. 'Molecular typing of adult acute myeloid leukaemia: significance of translocations, tandem duplications, methylation, and selective gene expression profiling', *Br J Haematol*, 131: 457-67.

Ortolani, C. 2011. *Flow Cytometry of Hematological Malignancies* (Wiley-Blackwell).

Pagliuca, S., Y. Guan, A. Tiwari, D. Grabowski, C. Gurnari, C. Kerr, S. Kongkiatkamon, M. Hasipek, L. Terkawi, M. Zawit, H. Awada, V. Visconte, Y. Sauntharajah, J. Phillips, J. Maciejewski, and B. Jha. 2020. 'Inhibition of Critical DNA Dioxygenase Activity in IDH1/2 Mutant Myeloid Neoplasms', *Blood*, 136: 28-28.

Palii, S. S., B. O. Van Emburgh, U. T. Sankpal, K. D. Brown, and K. D. Robertson. 2008. 'DNA methylation inhibitor 5-Aza-2'-deoxycytidine induces reversible genome-wide DNA damage that is distinctly influenced by DNA methyltransferases 1 and 3B', *Mol Cell Biol*, 28: 752-71.

Palmieri, R., G. Paterno, E. De Bellis, L. Mercante, E. Buzzatti, F. Esposito, M. I. Del Principe, L. Maurillo, F. Buccisano, and A. Venditti. 2020. 'Therapeutic Choice in Older Patients with Acute Myeloid Leukemia: A Matter of Fitness', *Cancers*, 12: 120.

Papaemmanuil, E., M. Gerstung, L. Bullinger, V. I. Gaidzik, P. Paschka, N. D. Roberts, N. E. Potter, M. Heuser, F. Thol, N. Bolli, G. Gundem, P. Van Loo, I. Martincorena, P. Ganly, L. Mudie, S. McLaren, S. O'Meara, K. Raine, D. R. Jones, J. W. Teague, A. P. Butler, M. F. Greaves, A. Ganser, K. Döhner, R. F. Schlenk, H. Döhner, and P. J. Campbell. 2016. 'Genomic Classification and Prognosis in Acute Myeloid Leukemia', *New England Journal of Medicine*, 374: 2209-21.

Patel, B. J., B. Przychodzen, S. Thota, T. Radivoyevitch, V. Visconte, T. Kuzmanovic, M. Clemente, C. Hirsch, A. Morawski, R. Souaid, C. Saygin, A. Nazha, B. Demarest, T. LaFramboise, H. Sakaguchi, S. Kojima, H. E. Carraway, S. Ogawa, H. Makishima, M. A. Sekeres, and J. P. Maciejewski. 2017. 'Genomic determinants of chronic myelomonocytic leukemia', *Leukemia*, 31: 2815-23.

Patel, J. P., M. Gonen, M. E. Figueroa, H. Fernandez, Z. Sun, J. Racevskis, P. Van Vlierberghe, I. Dolgalev, S. Thomas, O. Aminova, K. Huberman, J. Cheng, A. Viale, N. D. Socci, A. Heguy, A. Cherry, G. Vance, R. R. Higgins, R. P. Ketterling, R. E. Gallagher, M. Litzow, M. R. van den Brink, H. M. Lazarus, J. M. Rowe, S. Luger, A. Ferrando, E. Paietta, M. S. Tallman, A. Melnick, O. Abdel-Wahab, and R. L. Levine. 2012. 'Prognostic relevance of integrated genetic profiling in acute myeloid leukemia', *N Engl J Med*, 366: 1079-89.

Paulsson, K., and B. Johansson. 2007. 'Trisomy 8 as the sole chromosomal aberration in acute myeloid leukemia and myelodysplastic syndromes', *Pathol Biol (Paris)*, 55: 37-48.

Pfeifer, G. P. 2018. 'Defining Driver DNA Methylation Changes in Human Cancer', *International Journal of Molecular Sciences*, 19: 1166.

Pollyea, D. A., J. Zehnder, S. Coutre, J. R. Gotlib, L. Gallegos, O. Abdel-Wahab, P. Greenberg, B. Zhang, M. Liedtke, C. Berube, R. Levine, B. S. Mitchell, and B. C. Medeiros. 2013. 'Sequential azacitidine plus lenalidomide combination for elderly patients with untreated acute myeloid leukemia', *Haematologica*, 98: 591-6.

Prada-Arismendy, J., J. C. Arroyave, and S. Röthlisberger. 2017. 'Molecular biomarkers in acute myeloid leukemia', *Blood Reviews*, 31: 63-76.

Pradhan, S., A. Bacolla, R. D. Wells, and R. J. Roberts. 1999. 'Recombinant human DNA (cytosine-5) methyltransferase. I. Expression, purification, and comparison of de novo and maintenance methylation', *J Biol Chem*, 274: 33002-10.

Prokhortchouk, A., B. Hendrich, H. Jorgensen, A. Ruzov, M. Wilm, G. Georgiev, A. Bird, and E. Prokhortchouk. 2001. 'The p120 catenin partner Kaiso is a DNA methylation-dependent transcriptional repressor', *Genes Dev*, 15: 1613-8.

Pronier, E., C. Almire, H. Mokrani, A. Vasanthakumar, A. Simon, B. da Costa Reis Monte Mor, A. Massé, J. P. Le Couédic, F. Pendino, B. Carbonne, J. Larghero, J. L. Ravanat, N. Casadevall, O. A. Bernard, N. Droin, E. Solary, L. A. Godley, W. Vainchenker, I. Plo, and F. Delhommeau. 2011. 'Inhibition of TET2-mediated conversion of 5-methylcytosine to 5-hydroxymethylcytosine disturbs erythroid and granulomonocytic differentiation of human hematopoietic progenitors', *Blood*, 118: 2551-5.

Provost, E., C. A. Weier, and S. D. Leach. 2013. 'Multiple ribosomal proteins are expressed at high levels in developing zebrafish endoderm and are required for normal exocrine pancreas development', *Zebrafish*, 10: 161-9.

Pulikkan, J. A., and L. H. Castilla. 2018. 'Preleukemia and Leukemia-Initiating Cell Activity in inv(16) Acute Myeloid Leukemia', *Frontiers in oncology*, 8: 129.

Qi, J., S. Singh, W. K. Hua, Q. Cai, S. W. Chao, L. Li, H. Liu, Y. Ho, T. McDonald, A. Lin, G. Marcucci, R. Bhatia, W. J. Huang, C. I. Chang, and Y. H. Kuo. 2015. 'HDAC8 Inhibition Specifically Targets Inv(16) Acute Myeloid Leukemic Stem Cells by Restoring p53 Acetylation', *Cell stem cell*, 17: 597-610.

Qin, T., J. Jelinek, J. Si, J. Shu, and J.-P. J. Issa. 2009. 'Mechanisms of resistance to 5-aza-2′-deoxycytidine in human cancer cell lines', *Blood*, 113: 659-67.

Quentmeier, H., R. A. F. MacLeod, M. Zaborski, and H. G. Drexler. 2006. 'JAK2 V617F tyrosine kinase mutation in cell lines derived from myeloproliferative disorders', *Leukemia*, 20: 471-76.

Quintás-Cardama, A., and J. Cortes. 2009. 'Molecular biology of bcr-abl1–positive chronic myeloid leukemia', *Blood*, 113: 1619-30.

Rampal, R., A. Alkalin, J. Madzo, A. Vasanthakumar, E. Pronier, J. Patel, Y. Li, J. Ahn, O. Abdel-Wahab, A. Shih, C. Lu, P. S. Ward, J. J. Tsai, T. Hricik, V. Tosello, J. E. Tallman, X. Zhao, D. Daniels, Q. Dai, L. Ciminio, I. Aifantis, C. He, F. Fuks, M. S. Tallman, A. Ferrando, S. Nimer, E. Paietta, C. B. Thompson, J. D. Licht, C. E. Mason, L. A. Godley, A. Melnick, M. E. Figueroa, and R. L. Levine. 2014. 'DNA hydroxymethylation profiling reveals that WT1 mutations result in loss of TET2 function in acute myeloid leukemia', *Cell Rep*, 9: 1841-55.

Rampal, R., and M. E. Figueroa. 2016. 'Wilms tumor 1 mutations in the pathogenesis of acute myeloid leukemia', *Haematologica*, 101: 672-79.

Ramsahoye, B. H., D. Biniszkiewicz, F. Lyko, V. Clark, A. P. Bird, and R. Jaenisch. 2000. 'Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a', *Proc Natl Acad Sci U S A*, 97: 5237-42.

Ran, F. A., P. D. Hsu, J. Wright, V. Agarwala, D. A. Scott, and F. Zhang. 2013. 'Genome engineering using the CRISPR-Cas9 system', *Nature Protocols*, 8: 2281-308.

Rao, S., S. Y. Lee, A. Gutierrez, J. Perrigoue, R. J. Thapa, Z. Tu, J. R. Jeffers, M. Rhodes, S. Anderson, T. Oravecz, S. P. Hunger, R. A. Timakhov, R. Zhang, S. Balachandran, G. P. Zambetti, J. R. Testa, A. T. Look, and D. L. Wiest. 2012. 'Inactivation of ribosomal protein L22 promotes transformation by induction of the stemness factor, Lin28B', *Blood*, 120: 3764-73.

Rasmussen, K. D., and K. Helin. 2016. 'Role of TET enzymes in DNA methylation, development, and cancer', *Genes Dev*, 30: 733-50.

Rasmussen, K. D., G. Jia, J. V. Johansen, M. T. Pedersen, N. Rapin, F. O. Bagger, B. T. Porse, O. A. Bernard, J. Christensen, and K. Helin. 2015. 'Loss of TET2 in hematopoietic cells

leads to DNA hypermethylation of active enhancers and induction of leukemogenesis', *Genes Dev*, 29: 910-22.

Ravandi, F., M. L. Alattar, M. R. Grunwald, M. A. Rudek, T. Rajkhowa, M. A. Richie, S. Pierce, N. Daver, G. Garcia-Manero, S. Faderl, A. Nazha, M. Konopleva, G. Borthakur, J. Burger, T. Kadia, S. Dellasala, M. Andreeff, J. Cortes, H. Kantarjian, and M. Levis. 2013. 'Phase 2 study of azacytidine plus sorafenib in patients with acute myeloid leukemia and FLT-3 internal tandem duplication mutation', *Blood*, 121: 4655-62.

Rees, D. C., E. Johnson, and O. Lewinson. 2009. 'ABC transporters: the power to change', *Nature Reviews Molecular Cell Biology*, 10: 218-27.

Rego, E. M., L. Z. He, R. P. Warrell, Jr., Z. G. Wang, and P. P. Pandolfi. 2000. 'Retinoic acid (RA) and As2O3 treatment in transgenic models of acute promyelocytic leukemia (APL) unravel the distinct nature of the leukemogenic process induced by the PML-RARalpha and PLZF-RARalpha oncoproteins', *Proc Natl Acad Sci U S A*, 97: 10173-78.

Reikvam, H., K. J. Hatfield, A. O. Kittang, R. Hovland, and Ø. Bruserud. 2011. 'Acute Myeloid Leukemia with the t(8;21) Translocation: Clinical Consequences and Biological Implications', *Journal of Biomedicine and Biotechnology*, 2011: 104631.

Reindl, C., K. Bagrintseva, S. Vempati, S. Schnittger, J. W. Ellwart, K. Wenig, K. P. Hopfner, W. Hiddemann, and K. Spiekermann. 2006. 'Point mutations in the juxtamembrane domain of FLT3 define a new class of activating mutations in AML', *Blood*, 107: 3700-7.

Ren, R. 2005. 'Mechanisms of BCR–ABL in the pathogenesis of chronic myelogenous leukaemia', *Nature Reviews Cancer*, 5: 172-83.

Renneville, A., C. Roumier, V. Biggio, O. Nibourel, N. Boissel, P. Fenaux, and C. Preudhomme. 2008. 'Cooperating gene mutations in acute myeloid leukemia: a review of the literature', *Leukemia*, 22: 915-31.

Rhoades, K. L., C. J. Hetherington, N. Harakawa, D. A. Yergeau, L. Zhou, L. Q. Liu, M. T. Little, D. G. Tenen, and D. E. Zhang. 2000. 'Analysis of the role of AML1-ETO in leukemogenesis, using an inducible transgenic mouse model', *Blood*, 96: 2108-15.

Rieger, M. A., and T. Schroeder. 2012. 'Hematopoiesis', *Cold Spring Harbor perspectives in biology*, 4: a008250.

Ringnér, M. 2008. 'What is principal component analysis?', *Nat Biotechnol*, 26: 303-04.

Ritchie, M. E., B. Phipson, D. Wu, Y. Hu, C. W. Law, W. Shi, and G. K. Smyth. 2015. 'limma powers differential expression analyses for RNA-sequencing and microarray studies', *Nucleic Acids Research*, 43: e47-e47.

Rius, M., C. Stresemann, D. Keller, M. Brom, E. Schirrmacher, D. Keppler, and F. Lyko. 2009. 'Human concentrative nucleoside transporter 1-mediated uptake of 5-azacytidine enhances DNA demethylation', *Molecular Cancer Therapeutics*, 8: 225.

Robb, L. 2007. 'Cytokine receptors and hematopoietic differentiation', *Oncogene*, 26: 6715-23.

Robey, R. W., K. M. Pluchino, M. D. Hall, A. T. Fojo, S. E. Bates, and M. M. Gottesman. 2018. 'Revisiting the role of ABC transporters in multidrug-resistant cancer', *Nature Reviews Cancer*, 18: 452-64.

Robinson, M. D., D. J. McCarthy, and G. K. Smyth. 2010. 'edgeR: a Bioconductor package for differential expression analysis of digital gene expression data', *Bioinformatics*, 26: 139-40.

Roboz, G. J., H. M. Kantarjian, K. W. L. Yee, P. L. Kropf, C. L. O'Connell, E. A. Griffiths, W. Stock, N. G. Daver, E. Jabbour, E. K. Ritchie, K. J. Walsh, D. Rizzieri, S. D. Lunin, T. Curio, W. Chung, Y. Hao, J. N. Lowder, M. Azab, and J.-P. J. Issa. 2018. 'Dose, schedule, safety, and efficacy of guadecitabine in relapsed or refractory acute myeloid leukemia', *Cancer*, 124: 325-34.

Saha, A., S. Das, M. Moin, M. Dutta, A. Bakshi, M. S. Madhav, and P. B. Kirti. 2017. 'Genome-Wide Identification and Comprehensive Expression Profiling of Ribosomal Protein Small Subunit (RPS) Genes and their Comparative Analysis with the Large Subunit (RPL) Genes in Rice', *Front Plant Sci*, 8: 1553.

Saland, E., H. Boutzen, R. Castellano, L. Pouyet, E. Griessinger, C. Larrue, F. de Toni, S. Scotland, M. David, G. Danet-Desnoyers, F. Vergez, Y. Barreira, Y. Collette, C. Récher, and J. E. Sarry. 2015. 'A robust and rapid xenograft model to assess efficacy of chemotherapeutic agents for human acute myeloid leukemia', *Blood Cancer J*, 5: e297.

Saliba, A. N., A. J. John, and S. H. Kaufmann. 2021. 'Resistance to venetoclax and hypomethylating agents in acute myeloid leukemia', *Cancer Drug Resistance*, 4: 125-42.

Sampath, J., D. Sun, V. J. Kidd, J. Grenet, A. Gandhi, L. H. Shapiro, Q. Wang, G. P. Zambetti, and J. D. Schuetz. 2001. 'Mutant p53 cooperates with ETS and selectively up-regulates human MDR1 not MRP1', *J Biol Chem*, 276: 39359-67.

Samra, B., G. Richard-Carpentier, T. M. Kadia, F. Ravandi, N. Daver, C. D. DiNardo, G. C. Issa, P. Bose, M. Y. Konopleva, M. Yilmaz, M. Ohanian, G. Borthakur, G. Garcia-Manero, S. Pierce, J. E. Cortes, H. Kantarjian, and N. J. Short. 2020. 'Characteristics and outcomes of patients with therapy-related acute myeloid leukemia with normal karyotype', *Blood Cancer J*, 10: 47.

Sangle, N. A., and S. L. Perkins. 2011. 'Core-Binding Factor Acute Myeloid Leukemia', *Archives of Pathology & Laboratory Medicine*, 135: 1504-09.

Sarraf, S. A., and I. Stancheva. 2004. 'Methyl-CpG binding protein MBD1 couples histone H3 methylation at lysine 9 by SETDB1 to DNA replication and chromatin assembly', *Mol Cell*, 15: 595-605.

Saultz, J. N., and R. Garzon. 2016. 'Acute Myeloid Leukemia: A Concise Review', *Journal of Clinical Medicine*, 5: 33.

Saxonov, S., P. Berg, and D. L. Brutlag. 2006. 'A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters', *Proc Natl Acad Sci U S A*, 103: 1412-7.

Schaefer, M., S. Hagemann, K. Hanna, and F. Lyko. 2009. 'Azacytidine inhibits RNA methylation at DNMT2 target sites in human cancer cell lines', *Cancer Res*, 69: 8127-32.

Schermelleh, L., F. Spada, H. P. Easwaran, K. Zolghadr, J. B. Margot, M. C. Cardoso, and H. Leonhardt. 2005. 'Trapped in action: direct visualization of DNA methyltransferase activity in living cells', *Nat Methods*, 2: 751-6.

Schiffer CA, S. R. 2003. 'Morphologic Classification and Clinical and Laboratory Correlates.' in, *Holland-Frei Cancer Medicine, 6th Edition* (BC Decker: Hamilton (ON)).

Schneider, C., C. L. Will, O. V. Makarova, E. M. Makarov, and R. Lührmann. 2002. 'Human U4/U6.U5 and U4atac/U6atac.U5 tri-snRNPs exhibit similar protein compositions', *Mol Cell Biol*, 22: 3219-29.

Schnittger, S., F. Dicker, W. Kern, N. Wendland, J. Sundermann, T. Alpermann, C. Haferlach, and T. Haferlach. 2011. 'RUNX1 mutations are frequent in de novo AML with noncomplex karyotype and confer an unfavorable prognosis', *Blood*, 117: 2348-57.

Schnittger, S., C. Haferlach, M. Ulke, T. Alpermann, W. Kern, and T. Haferlach. 2010. 'IDH1 mutations are detected in 6.6% of 1414 AML patients and are associated with intermediate risk karyotype and unfavorable prognosis in adults younger than 60 years and unmutated NPM1 status', *Blood*, 116: 5486-96.

Schnittger, S., T. M. Kohl, T. Haferlach, W. Kern, W. Hiddemann, K. Spiekermann, and C. Schoch. 2006. 'KIT-D816 mutations in AML1-ETO-positive AML are associated with impaired event-free and overall survival', *Blood*, 107: 1791-99.

Schoch, C., A. Kohlmann, M. Dugas, W. Kern, W. Hiddemann, S. Schnittger, and T. Haferlach. 2005. 'Genomic gains and losses influence expression levels of genes located within the affected regions: a study on acute myeloid leukemias with trisomy 8, 11, or 13, monosomy 7, or deletion 5q', *Leukemia*, 19: 1224-28.

Schulz, W. A., C. Steinhoff, and A. R. Florl. 2006. 'Methylation of endogenous human retroelements in health and disease', *Curr Top Microbiol Immunol*, 310: 211-50.

Schwind, S., C. G. Edwards, D. Nicolet, K. Mrózek, K. Maharry, Y.-Z. Wu, P. Paschka, A.-K. Eisfeld, P. Hoellerbauer, H. Becker, K. H. Metzeler, J. Curfman, J. Kohlschmidt, T. W. Prior, J. E. Kolitz, W. Blum, M. J. Pettenati, P. Dal Cin, A. J. Carroll, M. A. Caligiuri, R. A. Larson, S. Volinia, G. Marcucci, C. D. Bloomfield, and O. Alliance for Clinical Trials in. 2013. 'inv(16)/t(16;16) acute myeloid leukemia with non-type A CBFB-MYH11 fusions associate with distinct clinical and genetic features and lack KIT mutations', *Blood*, 121: 385-91.

Scotto, K. W. 2003. 'Transcriptional regulation of ABC drug transporters', *Oncogene*, 22: 7496-511.

Seelan, R. S., P. Mukhopadhyay, M. M. Pisano, and R. M. Greene. 2018. 'Effects of 5-Aza-2'-deoxycytidine (decitabine) on gene expression', *Drug Metab Rev*, 50: 193-207.

Sharif, J., M. Muto, S. Takebayashi, I. Suetake, A. Iwamatsu, T. A. Endo, J. Shinga, Y. Mizutani-Koseki, T. Toyoda, K. Okamura, S. Tajima, K. Mitsuya, M. Okano, and H. Koseki. 2007. 'The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA', *Nature*, 450: 908-12.

Shen, L., H. Kantarjian, Y. Guo, E. Lin, J. Shan, X. Huang, D. Berry, S. Ahmed, W. Zhu, S. Pierce, Y. Kondo, Y. Oki, J. Jelinek, H. Saba, E. Estey, and J. P. Issa. 2010. 'DNA methylation predicts survival and response to therapy in patients with myelodysplastic syndromes', *J Clin Oncol*, 28: 605-13.

Shen, L., C.-X. Song, C. He, and Y. Zhang. 2014. 'Mechanism and function of oxidative reversal of DNA and RNA methylation', *Annual review of biochemistry*, 83: 585-614.

Shi, D.-Q., I. Ali, J. Tang, and W.-C. Yang. 2017. 'New Insights into 5hmC DNA Modification: Generation, Distribution and Function', *Frontiers in Genetics*, 8.

Shih, A. H., O. Abdel-Wahab, J. P. Patel, and R. L. Levine. 2012. 'The role of mutations in epigenetic regulators in myeloid malignancies', *Nat Rev Cancer*, 12: 599-612.

Shih, A. H., Y. Jiang, C. Meydan, K. Shank, S. Pandey, L. Barreyro, I. Antony-Debre, A. Viale, N. Socci, Y. Sun, A. Robertson, M. Cavatore, E. de Stanchina, T. Hricik, F. Rapaport, B. Woods, C. Wei, M. Hatlen, M. Baljevic, S. D. Nimer, M. Tallman, E. Paietta, L. Cimmino, I. Aifantis, U. Steidl, C. Mason, A. Melnick, and R. L. Levine. 2015. 'Mutational Cooperativity Linked to Combinatorial Epigenetic Gain of Function in Acute Myeloid Leukemia', *Cancer Cell*, 27: 502-15.

Shih, A. H., C. Meydan, K. Shank, F. E. Garrett-Bakelman, P. S. Ward, A. M. Intlekofer, A. Nazir, E. M. Stein, K. Knapp, J. Glass, J. Travins, K. Straley, C. Gliser, C. E. Mason,

K. Yen, C. B. Thompson, A. Melnick, and R. L. Levine. 2017. 'Combination Targeted Therapy to Disrupt Aberrant Oncogenic Signaling and Reverse Epigenetic Dysfunction in IDH2- and TET2-Mutant Acute Myeloid Leukemia', *Cancer Discov*, 7: 494-505.

Shih, L. Y., C. F. Huang, P. N. Wang, J. H. Wu, T. L. Lin, P. Dunn, and M. C. Kuo. 2004. 'Acquisition of FLT3 or N-ras mutations is frequently associated with progression of myelodysplastic syndrome to acute myeloid leukemia', *Leukemia*, 18: 466-75.

Shysh, A. C., L. T. Nguyen, M. Guo, M. Vaska, C. Naugler, and F. Rashid-Kolvear. 2017. 'The incidence of acute myeloid leukemia in Calgary, Alberta, Canada: a retrospective cohort study', *BMC Public Health*, 18: 94.

Sill, H., W. Olipitz, A. Zebisch, E. Schulz, and A. Wolfler. 2011. 'Therapy-related myeloid neoplasms: pathobiology and clinical characteristics', *Br J Pharmacol*, 162: 792-805.

Sinha, S., D. Thomas, L. Yu, A. J. Gentles, N. Jung, M. R. Corces-Zimmerman, S. M. Chan, A. Reinisch, A. P. Feinberg, D. L. Dill, and R. Majeti. 2015. 'Mutant WT1 is associated with DNA hypermethylation of PRC2 targets in AML and responds to EZH2 inhibition', *Blood*, 125: 316-26.

Smith, S. M., M. M. Le Beau, D. Huo, T. Karrison, R. M. Sobecks, J. Anastasi, J. W. Vardiman, J. D. Rowley, and R. A. Larson. 2003. 'Clinical-cytogenetic associations in 306 patients with therapy-related myelodysplasia and myeloid leukemia: the University of Chicago series', *Blood*, 102: 43-52.

Solary, E., O. A. Bernard, A. Tefferi, F. Fuks, and W. Vainchenker. 2014. 'The Ten-Eleven Translocation-2 (TET2) gene in hematopoiesis and hematopoietic diseases', *Leukemia*, 28: 485-96.

Song, S. J., K. Ito, U. Ala, L. Kats, K. Webster, S. M. Sun, M. Jongen-Lavrencic, K. Manova-Todorova, J. Teruya-Feldstein, D. E. Avigan, R. Delwel, and P. P. Pandolfi. 2013. 'The oncogenic microRNA miR-22 targets the TET2 tumor suppressor to promote hematopoietic stem cell self-renewal and transformation', *Cell stem cell*, 13: 87-101.

Speck, N. A., and D. G. Gilliland. 2002. 'Core-binding factors in haematopoiesis and leukaemia', *Nat Rev Cancer*, 2: 502-13.

Speck, N. A., and M. L. Iruela-Arispe. 2009. 'Conditional Cre/LoxP strategies for the study of hematopoietic stem cell formation', *Blood Cells Mol Dis*, 43: 6-11.

Sperling, A. S., C. J. Gibson, and B. L. Ebert. 2017. 'The genetics of myelodysplastic syndrome: from clonal haematopoiesis to secondary leukaemia', *Nature Reviews Cancer*, 17: 5-19.

Stahl, M., and M. S. Tallman. 2019. 'Acute promyelocytic leukemia (APL): remaining challenges towards a cure for all', *Leukemia & Lymphoma*, 60: 3107-15.

Stark, R., M. Grzelak, and J. Hadfield. 2019. 'RNA sequencing: the teenage years', *Nature Reviews Genetics*, 20: 631-56.

Steinbach, D., and O. Legrand. 2007. 'ABC transporters and drug resistance in leukemia: was P-gp nothing but the first head of the Hydra?', *Leukemia*, 21: 1172-76.

Stomper, J., J. C. Rotondo, G. Greve, and M. Lübbert. 2021. 'Hypomethylating agents (HMA) for the treatment of acute myeloid leukemia and myelodysplastic syndromes: mechanisms of resistance and novel HMA-based therapies', *Leukemia*, 35: 1873-89.

Stone, R. M., T. Fischer, R. Paquette, G. Schiller, C. A. Schiffer, G. Ehninger, J. Cortes, H. M. Kantarjian, D. J. DeAngelo, A. Huntsman-Labed, C. Dutreix, A. del Corral, and F. Giles. 2012. 'Phase IB study of the FLT3 kinase inhibitor midostaurin with chemotherapy in younger newly diagnosed adult patients with acute myeloid leukemia', *Leukemia*, 26: 2061-8.

Stresemann, C., and F. Lyko. 2008. 'Modes of action of the DNA methyltransferase inhibitors azacytidine and decitabine', *International Journal of Cancer*, 123: 8-13.

Sun, Q. Y., L. W. Ding, K. T. Tan, W. Chien, A. Mayakonda, D. C. Lin, X. Y. Loh, J. F. Xiao, M. Meggendorfer, T. Alpermann, M. Garg, S. L. Lim, V. Madan, N. Hattori, Y. Nagata, S. Miyano, A. E. Yeoh, H. A. Hou, Y. Y. Jiang, S. Takao, L. Z. Liu, S. Z. Tan, M. Lill, M. Hayashi, A. Kinoshita, H. M. Kantarjian, S. M. Kornblau, S. Ogawa, T. Haferlach, H. Yang, and H. P. Koeffler. 2017. 'Ordering of mutations in acute myeloid leukemia with partial tandem duplication of MLL (MLL-PTD)', *Leukemia*, 31: 1-10.

Syed, K., S. Naguib, Z.-J. Liu, L. Cimmino, and F.-C. Yang. 2020. 'Novel combinations to improve hematopoiesis in myelodysplastic syndrome', *Stem Cell Research & Therapy*, 11: 132.

Tahiliani, M., K. P. Koh, Y. Shen, W. A. Pastor, H. Bandukwala, Y. Brudno, S. Agarwal, L. M. Iyer, D. R. Liu, L. Aravind, and A. Rao. 2009. 'Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1', *Science*, 324: 930-35.

Takahashi, S. 2011. 'Current findings for recurring mutations in acute myeloid leukemia', *J Hematol Oncol*, 4: 36.

Tartaglia, M., S. Martinelli, I. Iavarone, G. Cazzaniga, M. Spinelli, E. Giarin, V. Petrangeli, C. Carta, R. Masetti, M. Arico, F. Locatelli, G. Basso, M. Sorcini, A. Pession, and A. Biondi. 2005. 'Somatic PTPN11 mutations in childhood acute myeloid leukaemia', *Br J Haematol*, 129: 333-9.

Tartaglia, M., C. M. Niemeyer, K. M. Shannon, and M. L. Loh. 2004. 'SHP-2 and myeloid malignancies', *Curr Opin Hematol*, 11: 44-50.

Tate, J. G., S. Bamford, H. C. Jubb, Z. Sondka, D. M. Beare, N. Bindal, H. Boutselakis, C. G. Cole, C. Creatore, E. Dawson, P. Fish, B. Harsha, C. Hathaway, S. C. Jupe, C. Y. Kok, K. Noble, L. Ponting, C. C. Ramshaw, C. E. Rye, H. E. Speedy, R. Stefancsik, S. L. Thompson, S. Wang, S. Ward, P. J. Campbell, and S. A. Forbes. 2018. 'COSMIC: the Catalogue Of Somatic Mutations In Cancer', *Nucleic Acids Research*, 47: D941-D47.

Tefferi, A., K. H. Lim, and R. Levine. 2009. 'Mutation in TET2 in myeloid cancers', *N Engl J Med*, 361: 1117; author reply 17-8.

Theocharides, A. P., A. Rongvaux, K. Fritsch, R. A. Flavell, and M. G. Manz. 2016. 'Humanized hemato-lymphoid system mice', *Haematologica*, 101: 5-19.

Thottassery, J. V., G. P. Zambetti, K. Arimori, E. G. Schuetz, and J. D. Schuetz. 1997. 'p53-dependent regulation of MDR1 gene expression causes selective resistance to chemotherapeutic agents', *Proc Natl Acad Sci U S A*, 94: 11037-42.

Tiacci, E., A. Spanhol-Rosseto, M. P. Martelli, L. Pasqualucci, H. Quentmeier, V. Grossmann, H. G. Drexler, and B. Falini. 2012. 'The NPM1 wild-type OCI-AML2 and the NPM1-mutated OCI-AML3 cell lines carry DNMT3A mutations', *Leukemia*, 26: 554-57.

To, K. K. W., M. Wu, C. W. S. Tong, and W. Yan. 2020. 'Chapter 2 - Drug transporters in the development of multidrug resistance in colorectal cancer.' in Chi Hin Cho and Tao Hu (eds.), *Drug Resistance in Colorectal Cancer: Molecular Mechanisms and Therapeutic Strategies* (Academic Press).

Tomita, A., H. Kiyoi, and T. Naoe. 2013. 'Mechanisms of action and resistance to all-trans retinoic acid (ATRA) and arsenic trioxide (As2O 3) in acute promyelocytic leukemia', *Int J Hematol*, 97: 717-25.

Traina, F., V. Visconte, P. Elson, A. Tabarroki, A. M. Jankowska, E. Hasrouni, Y. Sugimoto, H. Szpurka, H. Makishima, C. L. O'Keefe, M. A. Sekeres, A. S. Advani, M. Kalaycio, E. A. Copelan, Y. Sauntharajah, S. T. Olalla Saad, J. P. Maciejewski, and R. V. Tiu. 2014. 'Impact of molecular mutations on treatment response to DNMT inhibitors in myelodysplasia and related neoplasms', *Leukemia*, 28: 78-87.

Treppendahl, M. B., L. S. Kristensen, and K. Grønbæk. 2014. 'Predicting response to epigenetic therapy', *J Clin Invest*, 124: 47-55.

Ukai, H., H. Ishii-Oba, M. Ukai-Tadenuma, T. Ogiu, and H. Tsuji. 2003. 'Formation of an active form of the interleukin-2/15 receptor beta-chain by insertion of the intracisternal A particle in a radiation-induced mouse thymic lymphoma and its role in tumorigenesis', *Mol Carcinog*, 37: 110-9.

Ungureanu, D., J. Wu, T. Pekkala, Y. Niranjan, C. Young, O. N. Jensen, C.-F. Xu, T. A. Neubert, R. C. Skoda, S. R. Hubbard, and O. Silvennoinen. 2011. 'The pseudokinase

domain of JAK2 is a dual-specificity protein kinase that negatively regulates cytokine signaling', *Nature structural & molecular biology*, 18: 971-76.

Valencia, A., E. Masala, A. Rossi, A. Martino, A. Sanna, F. Buchi, F. Canzian, D. Cilloni, V. Gaidano, M. T. Voso, O. Kosmider, M. Fontenay, A. Gozzini, A. Bosi, and V. Santini. 2014. 'Expression of nucleoside-metabolizing enzymes in myelodysplastic syndromes and modulation of response to azacitidine', *Leukemia*, 28: 621-8.

Valinluck, V., H. H. Tsai, D. K. Rogstad, A. Burdzy, A. Bird, and L. C. Sowers. 2004. 'Oxidative damage to methyl-CpG sequences inhibits the binding of the methyl-CpG binding domain (MBD) of methyl-CpG binding protein 2 (MeCP2)', *Nucleic Acids Res*, 32: 4100-8.

Vardiman, J. W., N. L. Harris, and R. D. Brunning. 2002. 'The World Health Organization (WHO) classification of the myeloid neoplasms', *Blood*, 100: 2292-302.

Vasconcelos, F. C., P. S. de Souza, T. Hancio, F. C. C. de Faria, and R. C. Maia. 2021. 'Update on drug transporter proteins in acute myeloid leukemia: Pathological implication and clinical setting', *Critical Reviews in Oncology/Hematology*, 160: 103281.

Vella, P., A. Scelfo, S. Jammula, F. Chiacchiera, K. Williams, A. Cuomo, A. Roberto, J. Christensen, T. Bonaldi, K. Helin, and D. Pasini. 2013. 'Tet proteins connect the O-linked N-acetylglucosamine transferase Ogt to chromatin in embryonic stem cells', *Mol Cell*, 49: 645-56.

Venci, A., R. Mazza, O. Spinelli, L. Di Schiena, and D. Bettio. 2017. 'Acute promyelocytic leukemia with a cryptic insertion of RARA into PML on chromosome 15 due to uniparental isodisomy: A case report', *Oncol Lett*, 13: 4180-84.

Venney, D., A. Mohd-Sarip, and K. I. Mills. 2021. 'The Impact of Epigenetic Modifications in Myeloid Malignancies', *International Journal of Molecular Sciences*, 22: 5013.

Virtaneva, K., F. A. Wright, S. M. Tanner, B. Yuan, W. J. Lemon, M. A. Caligiuri, C. D. Bloomfield, A. de la Chapelle, and R. Krahe. 2001. 'Expression profiling reveals fundamental biological differences in acute myeloid leukemia with isolated trisomy 8 and normal cytogenetics', *Proceedings of the National Academy of Sciences*, 98: 1124.

Voelker, R., and C. Small. 2021. 'RNA-seqlopedia', Accessed 18-07-2021. https://rnaseq.uoregon.edu/#library-prep.

Walter, R. B., T. A. Gooley, V. H. J. van der Velden, M. R. Loken, J. J. M. van Dongen, D. A. Flowers, I. D. Bernstein, and F. R. Appelbaum. 2007. 'CD33 expression and P-glycoprotein–mediated drug efflux inversely correlate and predict clinical outcome in patients with acute myeloid leukemia treated with gemtuzumab ozogamicin monotherapy', *Blood*, 109: 4168-70.

Wang, K., and M. Bucan. 2014. 'Copy Number Variation Detection via High-Density SNP Genotyping', *Scientific Protocols*.

Wang, R., X. Gao, and L. Yu. 2019. 'The prognostic impact of tet oncogene family member 2 mutations in patients with acute myeloid leukemia: a systematic-review and meta-analysis', *BMC Cancer*, 19: 389-89.

Wang, W., S. Nag, X. Zhang, M. H. Wang, H. Wang, J. Zhou, and R. Zhang. 2015. 'Ribosomal proteins and human diseases: pathogenesis, molecular mechanisms, and therapeutic implications', *Med Res Rev*, 35: 225-85.

Wang, Y.-Y., G.-B. Zhou, T. Yin, B. Chen, J.-Y. Shi, W.-X. Liang, X.-L. Jin, J.-H. You, G. Yang, Z.-X. Shen, J. Chen, S.-M. Xiong, G.-Q. Chen, F. Xu, Y.-W. Liu, Z. Chen, and S.-J. Chen. 2005. 'AML1-ETO and C-KIT mutation/overexpression in t(8;21) leukemia: Implication in stepwise leukemogenesis and response to Gleevec', *Proc Natl Acad Sci U S A*, 102: 1104.

Wang, Y., M. Xiao, X. Chen, L. Chen, Y. Xu, L. Lv, P. Wang, H. Yang, S. Ma, H. Lin, B. Jiao, R. Ren, D. Ye, K. L. Guan, and Y. Xiong. 2015. 'WT1 recruits TET2 to regulate its target gene expression and suppress leukemia cell proliferation', *Mol Cell*, 57: 662-73.

Wang, Z., M. Gerstein, and M. Snyder. 2009. 'RNA-Seq: a revolutionary tool for transcriptomics', *Nat Rev Genet*, 10: 57-63.

Wang, Z., X. Wu, and Y. Wang. 2018. 'A framework for analyzing DNA methylation data from Illumina Infinium HumanMethylation450 BeadChip', *BMC Bioinformatics*, 19: 115.

Ward, P. S., J. Patel, D. R. Wise, O. Abdel-Wahab, B. D. Bennett, H. A. Coller, J. R. Cross, V. R. Fantin, C. V. Hedvat, A. E. Perl, J. D. Rabinowitz, M. Carroll, S. M. Su, K. A. Sharp, R. L. Levine, and C. B. Thompson. 2010. 'The Common Feature of Leukemia-Associated IDH1 and IDH2 Mutations Is a Neomorphic Enzyme Activity Converting α-Ketoglutarate to 2-Hydroxyglutarate', *Cancer Cell*, 17: 225-34.

Weissmann, S., T. Alpermann, V. Grossmann, A. Kowarsch, N. Nadarajah, C. Eder, F. Dicker, A. Fasan, C. Haferlach, T. Haferlach, W. Kern, S. Schnittger, and A. Kohlmann. 2012. 'Landscape of TET2 mutations in acute myeloid leukemia', *Leukemia*, 26: 934-42.

Wen, B., W. You, S. Yang, and X. Du. 2020. 'Indirect comparison of azacitidine and decitabine for the therapy of elderly patients with acute myeloid leukemia: a systematic review and network meta-analysis', *Experimental Hematology & Oncology*, 9: 3.

Wilhelm-Benartzi, C. S., D. C. Koestler, M. R. Karagas, J. M. Flanagan, B. C. Christensen, K. T. Kelsey, C. J. Marsit, E. A. Houseman, and R. Brown. 2013. 'Review of processing and analysis methods for DNA methylation array data', *British Journal of Cancer*, 109: 1394-402.

Wilkens, S. 2015. 'Structure and mechanism of ABC transporters', *F1000prime reports*, 7: 14-14.

Will, C. L., and R. Lührmann. 2011. 'Spliceosome structure and function', *Cold Spring Harbor perspectives in biology*, 3: a003707.

Winters, A. C., and K. M. Bernt. 2017. 'MLL-Rearranged Leukemias-An Update on Science and Clinical Approaches', *Front Pediatr*, 5: 4.

Wolf, S. J., M. Bachtiar, J. Wang, T. S. Sim, S. S. Chong, and C. G. L. Lee. 2011. 'An update on ABCB1 pharmacogenetics: insights from a 3D model into the location and evolutionary conservation of residues corresponding to SNPs associated with drug pharmacokinetics', *The Pharmacogenomics Journal*, 11: 315-25.

Wu, M., E. M. Rinchik, E. Wilkinson, and D. K. Johnson. 1997. 'Inherited somatic mosaicism caused by an intracisternal A particle insertion in the mouse tyrosinase gene', *Proc Natl Acad Sci U S A*, 94: 890-4.

Wunderlich, M., B. Mizukawa, F. S. Chou, C. Sexton, M. Shrestha, Y. Saunthararajah, and J. C. Mulloy. 2013. 'AML cells are differentially sensitive to chemotherapy treatment in a human xenograft model', *Blood*, 121: e90-7.

Xiaoguang, W., C. Zhangguo, K. M. Ameet, S. Alexa, R. Wenhua, P. Zenggang, and H. W. Jing. 2018. 'Chemotherapy-induced differential cell cycle arrest in B-cell lymphomas affects their sensitivity to Wee1 inhibition', *Haematologica*, 103: 466-76.

Xie, W., C. L. Barr, A. Kim, F. Yue, A. Y. Lee, J. Eubanks, E. L. Dempster, and B. Ren. 2012. 'Base-resolution analyses of sequence and parent-of-origin dependent DNA methylation in the mouse genome', *Cell*, 148: 816-31.

Yang, L., Y. N. Liu, N. Zhang, X. Y. Ding, W. Zhang, K. F. Shen, L. Huang, J. F. Zhou, S. Cui, Z. M. Zhu, Z. Hu, and M. Xiao. 2017. 'Novel impact of the DNMT3A R882H mutation on GSH metabolism in a K562 cell model established by TALENs', *Oncotarget*, 8: 30395-409.

Yang, L., H. Zhang, X. Yang, T. Lu, S. Ma, H. Cheng, K. Yen, and T. Cheng. 2021. 'Prognostic Prediction of Cytogenetically Normal Acute Myeloid Leukemia Based on a Gene Expression Model', *Frontiers in oncology*, 11.

Yang, X., and K. Qian. 2017. 'Protein O-GlcNAcylation: emerging mechanisms and functions', *Nature reviews. Molecular cell biology*, 18: 452-65.

Ye, J., G. Coulouris, I. Zaretskaya, I. Cutcutache, S. Rozen, and T. L. Madden. 2012. 'Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction', *BMC Bioinformatics*, 13: 134.

Yin, R., S.-Q. Mao, B. Zhao, Z. Chong, Y. Yang, C. Zhao, D. Zhang, H. Huang, J. Gao, Z. Li, Y. Jiao, C. Li, S. Liu, D. Wu, W. Gu, Y.-G. Yang, G.-L. Xu, and H. Wang. 2013. 'Ascorbic Acid Enhances Tet-Mediated 5-Methylcytosine Oxidation and Promotes DNA Demethylation in Mammals', *Journal of the American Chemical Society*, 135: 10396-403.

Yoon, H. G., D. W. Chan, A. B. Reynolds, J. Qin, and J. Wong. 2003. 'N-CoR mediates DNA methylation-dependent repression through a methyl CpG binding protein Kaiso', *Mol Cell*, 12: 723-34.

Yu, B., and D. Liu. 2019. 'Gemtuzumab ozogamicin and novel antibody-drug conjugates in clinical trials for acute myeloid leukemia', *Biomarker research*, 7: 24-24.

Yu, L., H. T. Kim, S. Kasar, P. Benien, W. Du, K. Hoang, A. Aw, B. Tesar, R. Improgo, S. Fernandes, S. Radhakrishnan, J. Klitgaard, C. Lee, G. Getz, S. R. Setlur, and J. R. Brown. 2017. 'Survival of Del17p CLL Depends on Genomic Complexity and Somatic Mutation', *Clin Cancer Res*, 23: 735-45.

Yu, M.-G., and H.-Y. Zheng. 2017. 'Acute Myeloid Leukemia: Advancements in Diagnosis and Treatment', *Chinese Medical Journal*, 130: 211-18.

Zarei, M., S. Lal, S. J. Parker, A. Nevler, A. Vaziri-Gohar, K. Dukleska, N. C. Mambelli-Lisboa, C. Moffat, F. F. Blanco, S. N. Chand, M. Jimbo, J. A. Cozzitorto, W. Jiang, C. J. Yeo, E. R. Londin, E. L. Seifert, C. M. Metallo, J. R. Brody, and J. M. Winter. 2017. 'Posttranscriptional Upregulation of IDH1 by HuR Establishes a Powerful Survival Phenotype in Pancreatic Cancer Cells', *Cancer Res*, 77: 4460-71.

Zawadzka, I., A. Jeleń, J. Pietrzak, M. Żebrowska-Nawrocka, K. Michalska, D. Szmajda-Krygier, M. Mirowski, M. Łochowski, J. Kozak, and E. Balcerczak. 2020. 'The impact of ABCB1 gene polymorphism and its expression on non-small-cell lung cancer development, progression and therapy – preliminary report', *Scientific reports*, 10: 6188.

Zhang, D. E., P. Zhang, N. D. Wang, C. J. Hetherington, G. J. Darlington, and D. G. Tenen. 1997. 'Absence of granulocyte colony-stimulating factor signaling and neutrophil development in CCAAT enhancer binding protein alpha-deficient mice', *Proc Natl Acad Sci U S A*, 94: 569-74.

Zhang, J., B. A. Hug, E. Y. Huang, C. W. Chen, V. Gelmetti, M. Maccarana, S. Minucci, P. G. Pelicci, and M. A. Lazar. 2001. 'Oligomerization of ETO is obligatory for corepressor interaction', *Mol Cell Biol*, 21: 156-63.

Zhu, J., and S. G. Emerson. 2002. 'Hematopoietic cytokines, transcription factors and lineage commitment', *Oncogene*, 21: 3295-313.