THE AETIOLOGY OF PATHOGENIC MUTATIONS IN ACUTE MYELOID LEUKAEMIA

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

Mohd Fadly Md Ahid

Northern Institute for Cancer Research,
Faculty of Medical Sciences,
Newcastle University

October 2018

ABSTRACT

Relapse in acute myeloid leukaemia (AML) is driven by additional cooperating somatic mutations that were acquired (or selected for) after chemotherapy, where the relapsing clone evolves from a cell carrying leukaemia-initiating genetic lesions, including gene fusions such as *RUNX1/ETO*. Some fusion genes are thought to confer a mutator phenotype that predisposes cells to the acquisition of cooperating mutations. As such, it is important to understand how chemotherapy and fusion gene expression contribute to mutagenesis in relapsing AML.

The effect of chemotherapeutic agents on the mutation frequency was determined at the thymidine kinase (TK) and hypoxanthine-guanine phosphoribosyltransferase (HPRT) loci in a fully controlled cell model system using TK6 human lymphoblastoid cell lines. Both daunorubicin and cytarabine were mutagenic to DNA at the *TK* and *HPRT* loci. Comparison of the daunorubicin-treated and vehicle-treated mutational spectra at the *HPRT* coding region revealed a significant increase in large deletions in daunorubicin-treated cells relative to vehicle-treated cells, demonstrating daunorubicin as a powerful mutagen, inducing almost exclusively gene deletions presumably via strand break induction.

The effect of RUNX1/ETO fusion gene expression on mutation frequency was also determined both spontaneously and after chemotherapy treatment using a cell line transduced with the full-length *RUNX1/ETO* fusion gene. RUNX1/ETO significantly increases spontaneous mutation frequency at both *TK* and *HPRT* loci; however, there is no strong evidence that RUNX1/ETO fusion gene expression sensitises cells to chemotherapy-induced mutation. Interrogation of the spontaneous *HPRT* base substitution spectrum revealed that RUNX1/ETO significantly increases T:A > G:C transversions *in vitro*, and particularly at the central base position of $^{5'}$ApTpA$^{3'}$ / $^{5'}$TpApT$^{3'}$ sequences, although this observation was not evident in primary t(8;21) AML.

A flow cytometric method was established to evaluate the effect of RUNX1/ETO fusion gene expression on mutation frequency at specific base residues in an AML relevant gene. RUNX1/ETO significantly increase the type A exon 12 *NPM1* mutation frequency over time, suggesting this fusion protein predisposes cells to the acquisition of somatic mutation at loci relevant to leukaemogenesis.

Relapsed AML is associated with the acquisition of additional somatic mutations which are thought to drive phenotypic adaptability driving clonal selection during treatment and evolution of leukaemic clones. Analysis of high-throughput exome

sequencing of a small cohort of matched presentation and relapse AML samples revealed clonal evolution patterns in AML, characterised by the continuous acquisition of additional somatic mutations during disease progression and provided important insight on the clonal origins of relapsed AML.

Data generated from these studies demonstrate the *in vitro* mutagenicity of chemotherapeutic agents used in AML remission induction treatment. Furthermore, RUNX1/ETO fusion gene expression increases spontaneous mutation rate, including mutation at the *NPM1* locus. However, there is insufficient evidence from the induced mutation spectrum to discern the underlying mechanism of RUNX1/ETO-driven mutagenesis. Taken together, these data inform on the aetiology of somatic mutations in AML.

ACKNOWLEDGEMENTS

Firstly, I would like to express my deepest gratitude to my supervisors, Professor James Allan and Dr Sarah Fordham for their valuable guidance and support throughout the project. I would also like to thank the present and past members of molecular carcinogenesis group, Dr Helen Marr, Dr Thahira Rahman, Dr Mohammed Nahari, Dr Claire Elstob, Emmanouela Niki Soura, Nurettin Ayvali, Catherine Park and Devi Suchitra Devi for their help in the laboratory and 'shared moment' during my time in the institute. I would like to thank Dr Yaobo Xu and Dr Wei-Yu Lin for their help in NGS data analysis. I would like to acknowledge Dr Luise Hartmann (German Cancer Research Centre) for providing next generation sequencing data of 56 diagnostic AML t(8;21) samples. I would also like to acknowledge collaborators from Institute for Medical Research, Malaysia, Dr Zubaidah Zakaria and Dr Yuslina Mat Yusoff for providing the matched presentation and relapse AML samples as well as the funding for the whole exome sequencing study.

My deepest gratitude to my parents, Md Ahid Bakar and Zainab Abdullah, whom I love and respect. Without them, maybe I would not be the way I am now. My special thanks to my one and only beloved wife, Siti Rahayu Sabtu for her encouragement, support and patience in taking care of our two wonderful UK-born children, Fateh and Aisyah during this period of time. There is no way I can repay your kindness and sacrifice since we came to the UK.

Finally, I would like to thank Ministry of Health Malaysia for their sponsorship of the PhD study.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

6-TG	6-thioguanine
6-TGMP	6-thioguanosine monophosphate
8-OxoG	8-oxoguanine
AKAPs	PKA anchoring proteins
ALL	Acute lymphoblastic leukaemia
AML	Acute myeloid leukaemia
ANOVA	Analysis of variance
AP	Apurinic site
APL	Acute promyelocytic leukaemia
Ara-C	Cytarabine
Ara-CTP	Ara-C triphosphate
ASXL1	Additional sex-comb like 1
BCA	Bicinchoninic acid
BER	Base excision repair
BGI	Beijing genome institute
BSA	Bovine serum albumin
C/EBPα	CCAAT/enhancer binding protein alpha
CBF	Core-binding factor
CBFβ	Core-binding factor beta
CBP	p300/CREB-binding protein
CDKs	Cyclin dependent kinases
CE	Cloning efficiency
CHAT	Cytidine hypoxanthine aminopterin thymidine
CHIP	Clonal haematopoiesis of indeterminate potential
CLL	Chronic lymphocytic leukaemia
CML	Chronic myeloid leukaemia
CMML	Chronic myelomonocytic leukaemia
CN-AML	Cytogenetically-normal acute myeloid leukaemia
COSMIC	Catalogue of somatic mutations in cancer
CR	Complete remission
ddH_2O	Deionised distilled water
del(9q)	Chromosome 9q deletion

DMSO	Dimethyl sulphoxide
DSBs	Double strand breaks
ENU	Ethylnitrosourea
FAB	French-American-British
FBS	Fetal bovine serum
FCCF	Flow cytometry core facility
FLT3	FMS-related tyrosine kinase 3
GM-CSF	Granulocyte-macrophage colony-stimulating factor
GO	Gemtuzumab ozogamicin
GTPase	Guanosine triphosphatase
HDAC	Histone deacetylases
HHR	Hydrophobic amino acid heptad repeat
HL	Hodgkin lymphoma
HNPCC	Hereditary non-polyposis colorectal cancer
HPRT	Hypoxanthine-guanine phosphoribosyltransferase
HPSCs	Haematopoietic stem/progenitor cells
HR	Homologous recombination
H-RAS	Harvey sarcoma virus
Indels	Insertions/deletions
ITD	Internal tandem duplication
K-RAS	Kirsten sarcoma virus
LOH	Loss of heterozygosity
MDS	Myelodysplastic syndrome
MF	Mutation frequency
MFI	Meidan fluorescence intensity
MMR	Mismatch repair
MNU	N-methyl-N-nitrosourea
MPN	Myeloproliferative neoplasia
MRD	Minimal residual disease
MYND	Myeloid-Nervy-DEAF1
NADPH	Nicotinamide adenine dinucleotide phosphate
NCoR	Nuclear receptor corepressor
NER	Nucleotide excision repair
NGS	Next generation sequencing
NHEJ	Non-homologous end joining

NHL	Non-Hodgkin lymphoma
NHR (1, 2, 3, 4)	Nervy-homology regions
NPM1	Nucleophosmin 1
N-RAS	Neuroblastoma RAS viral (v-ras) oncogene homolog
OGG1	8-oxoguanine glycosylase
PAH	Polycyclic aromatic hydrocarbons
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PKA RIIα	Type II cyclic AMP-dependent protein kinase
Pol β	Polymerase β
Pol λ	Polymerase λ
PVDF	Polyvinylidenefluoride
qRT-PCR	Quantitative real-time PCR
RARα	Retinoic acid receptor alpha
RF10	RPMI, FBS 10%
RHD	Runt homology domain
ROS	Reactive oxygen species
RTK	Receptor tyrosine kinase
RUNX1	RUNT-related transcription factor 1
sAML	Secondary AML
SCF	Stem cell factor
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEM	Standard error of the mean
SMART	Somatic mutation and recombination test
SMRT	Thyroid hormone receptors
SNPs	Single nucleotide polymorphisms
SNVs	Single nucleotide variants
SSB	Single strand break
t-AML	Therapy-related acute myeloid leukaemia
TAFs	TATA binding protein-associated factors
TET2	Ten-eleven translocation 2
TFT	trifluorothymidine
THC	Thymidine hypoxanthine cytidine

TK	Thymidine kinase
TKD	Tyrosine kinase domain
Topo I	Topoisomerase I
Topo II	Topoisomerase II
UK	United Kingdom
UV	Ultraviolet light
VAF	Variant allele frequency
WES	Whole exome sequencing
WHO	World Health Organisation
WT	Wild type
α-KG	α-ketoglutarate

Chapter 1. Introduction

1.1. Acute Myeloid Leukaemia

Acute myeloid leukaemia (AML) is a malignant disorder of the blood and bone marrow which arises as a result of haematopoietic failure or de-regulation of haematopoiesis, the process by which the cellular components of blood are formed (Figure 1.1). It is a heterogeneous clonal disorder characterised by the proliferation and accumulation of immature myeloid progenitor cells (blasts) in bone marrow and blood, which occurs when the blasts lose the ability to differentiate normally and fail to respond to normal regulators of proliferation. Patients with AML are usually taken ill over a short timeframe and experience infection, bleeding or organ infiltration, and succumb to the disease within weeks or months if untreated.

1.1.1. AML incidence and epidemiology

In the United Kingdom (UK), AML is a relatively rare cancer, with an incidence of approximately 2,480 cases annually and median age at presentation of 71.6 years (Figure 1.2) (https://www.hmrn.org). Recent incidence data demonstrates a rate of diagnosis of 2 – 3 per 100, 000 annually in children, increasing to 15 per 100, 000 in older adults (Burnett and Venditti, 2011). As such, AML can be considered a disease related with older age, although it can present at any age. AML risk is slightly higher in males than females with three men affected for every two women. AML is the most common type of acute leukaemia in adults that accounts for about 6.3% of total haematological malignancies, and has the lowest survival rate of all the leukaemias (Deschler and Lübbert, 2006). The high median age of diagnosis has important implications for treatment strategies which are associated with chemoresistance and comorbidity that limits treatment options.

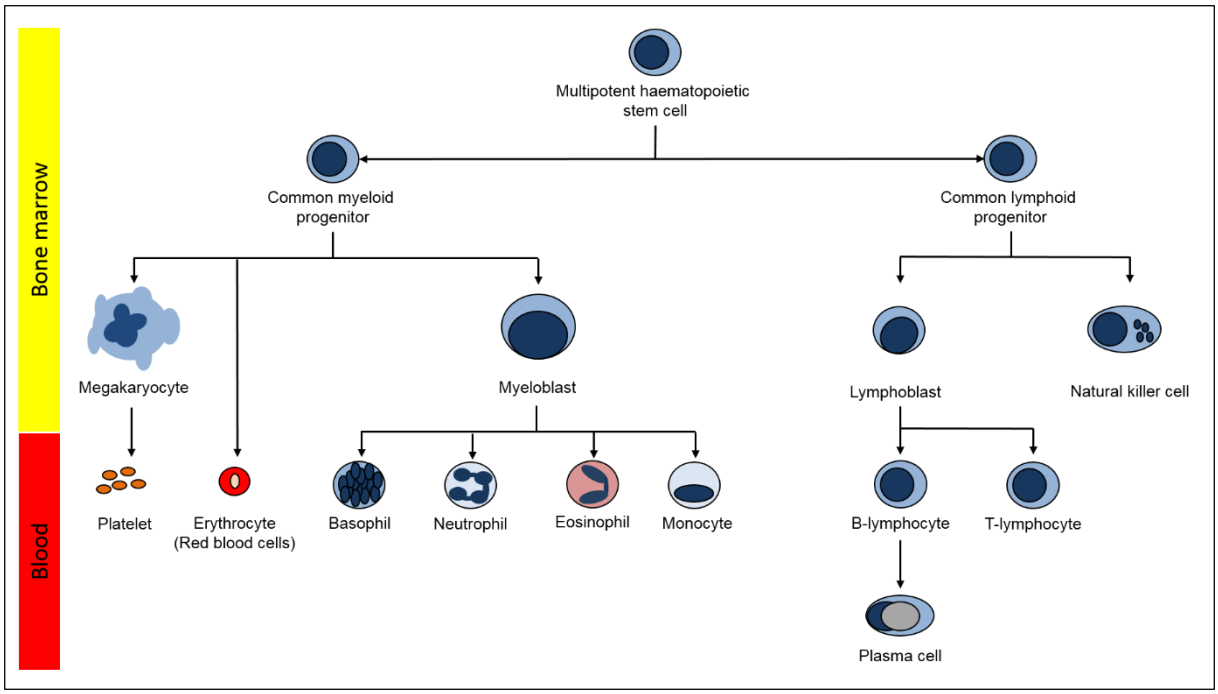


Figure 1.1. Haematopoiesis in human.

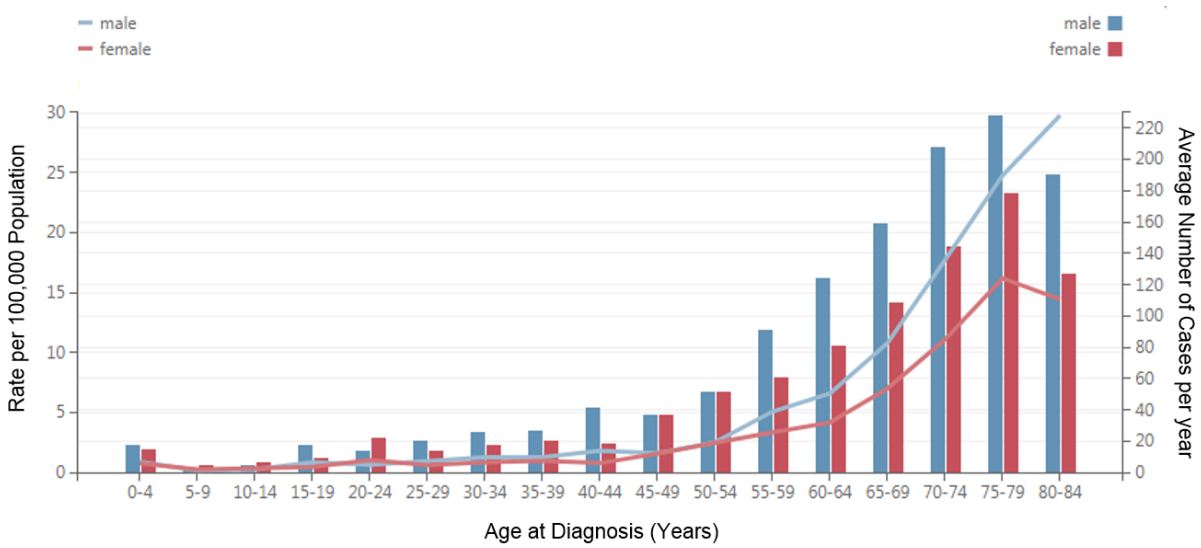


Figure 1.2. Incidence of AML in the UK in different age and gender (https://www.hmrn.org; 15 September 2018).

1.1.2. AML classification

The French-American-British (FAB) classification system was introduced in 1976 for classifying AML (Bennett *et al.*, 1976), and noted that many cases of AML are associated with recurring genetic abnormalities that affect cellular pathways of myeloid maturation and proliferation. The FAB classification provides a well-defined structure, based on morphological, immunological and clinical features, by which to classify disease into prognostically-relevant groups. Under the FAB classification system, AML is classified into 8 subtypes depending on the specific cell lineage(s) affected, defined by morphological appearance of the myeloblasts populating the leukaemic bone marrow (Table 1.1).

However, it was found that the correlation between morphology and genetic findings predicts the prognosis and biologic properties of the leukaemia more accurately than morphology alone. Thus, although the FAB classification recognises the morphologic heterogeneity of AML, it does not always reflect the genetic or clinical diversity of the disease. A revised classification devised under the auspices of the World Health Organization (WHO) recognizes accumulating knowledge of the cytogenetic and molecular characteristics (Table 1.2) (Arber *et al.*, 2016). In addition to morphological classification, detection of certain genetic mutations (cytogenetic and molecular genetic) can also delineate specific AML subtypes. Furthermore, many commonly detected genetic mutations are prognostically relevant (Table 1.3) (Döhner *et al.*, 2017) and, in some cases, can direct therapeutic choice. Specific chromosome aberrations and their molecular counterparts have been included in the WHO classification of haematologic malignancies, and together with morphology, immunophenotype and clinical features are used to define distinct disease entities for proper classification and prognostication (Vardiman *et al.*, 2002; Betz and Hess, 2010).

FAB Subtype	Description	Morphology
M0	Acute myeloblastic leukaemia, minimally differentiated	Immature myeloblasts without definite myeloid differentiation
M1	Acute myeloblastic leukaemia without maturation	Immature myeloblasts, Auer rods may be present
M2	Acute myeloblastic leukaemia with maturation	Immature myeloblasts (more mature than M1), Auer rods often present
M3	Acute promyelocytic leukaemia (APL)	Hypergranularpromyelocytes, bundles of Auer rods
$M4^{eo}$	Acute myelomonocytic leukaemia	Mixture of abnormal monocytes and myeloblasts/promyelocytes (eosinophil precursors in $M4^{eo}$)
$M5^{a}$ $M5^{b}$	Acute monocytic leukaemia	$M5^{a}$ 80% immature monoblasts $M5^{b}$ 20% differentiated monocytes
M6	Acute erythroleukaemia	Abnormal multinucleated erythroblasts, characterised by ringed sideroblasts
M7	Acute megakaryoblastic leukaemia	Megakaryoblasts, often myelofibrils, increased bone marrow reticulin

Table 1.1. Overview of the French-American-British (FAB) classification system of AML (Bennett *et al.*, 1976).

Acute myeloid leukaemia with recurrent genetic abnormalities
AML with t(8;21)(q22;q22); *RUNX1/ETO*
AML with inv(16)(p13q22) or t(16;16)(p13;q22); *CBFβ/MYH11*
APML with t(15;17)(q22;q12); *PML/RARα*
AML with t(9;11)(p21.3;q23.3); *MLLT3/KMT2A (MLL)*
AML with t(6;9)(p23;q34); *DEK/NUP214*
AML with inv(3)(q21q26.2) or t(3;3)(q21;q26.2); *GATA2/MECOM*
AML (megakaryoblastic) with t(1;22)(p13.3;q13.3); *DEK/NUP214*
AML with mutated *NPM*1
AML with biallelic mutations of *CEBPA*
Provisional entity: AML with *BCR/ABL1*
Provisional entity: AML with mutated *RUNX1*
Acute myeloid leukaemia with myelodysplasia changes
Therapy-related myeloid neoplasms
Acute myeloid leukaemia, not otherwise categorized
Classify as:
AML with minimally differentiated
AML without maturation
AML with maturation
Acute myelomonocytic leukaemia
Acute monoblastic/acute monocytic leukaemia
Acute erythroid leukaemia (erythroid/myeloid and pure erythroleukaemia)
Acute megakaryoblastic leukaemia
Acute basophilic leukaemia
Acute panmyelosis with myelofibrosis
Myeloid sarcoma
Myeloid proliferations related to Down syndrome
Transient abnormal myelopoiesis (TAM)
Myeloid leukaemia associated with Down syndrome

Table 1.2. WHO classification of AML (Arber *et al.*, 2016).

Prognosis	Genetic abnormality
Favourable	t(8;21)(q22;q22); *RUNX1/ETO*
	t(15;17)(q22;q21); *PML/RARα*
	inv(16)(p13;q22) or t(16;16)(p13;q22); *CBFβ/MYH11*
	Mutated *NPM1* without *FLT3*-ITD (normal karyotype)
	Biallelic mutated *CEBPA* (normal karyotype)
Intermediate I	Mutated NPM1 and FLT3-ITD (normal karyotype)
	Wild-type NPM1 and FLT3-ITD (normal karyotype)
	Wild-type NPM1 without FLT3-ITD (normal karyotype)
Intermediate II	t(9;11)(p22;q23); *MLLT3/KMT2A*
	Cytogenetic abnormalities not classified as favourable or adverse
Adverse	inv(3)(q21q26) or t(3;3)(q21;q26); *GATA2/MECOM* (*EVI1*)
	t(6;9)(p23;q34); *DEK/NUP214*
	t(v;11)(v;q23); *KMT2A* rearranged
	-5 or del(5q); -7; -17/abnormality (17p)
	Complex karyotype, monosomal karyotype
	Wild-type *NPM1* and *FLT3*-ITD^{high}
	Mutated *RUNX1*
	Mutated *ASXL1*
	Mutated *TP53*

Table 1.3. Prognostic classification of AML based on genetic abnormalities (Döhner *et al.*, 2017).

1.2. Genetic Mutations in AML Pathogenesis

The acquisition of somatic mutations underpins cancer development, which can either be caused by large-scale mutations (e.g. chromosomal translocations, inversions and numerical aberrations) or small-scale mutations (e.g. point mutations and microdeletions). Cytogenetic analysis has contributed to the identification of gross chromosomal abnormalities that play essential roles in AML pathogenesis, and remains the first tier of screening for AML classification. Abnormal karyotype with a recurrent chromosomal aberration constitutes approximately 55% of *de novo* AML cases, while the remaining 45% have a cytogenetically normal karyotype (CN-AML) (Mrózek *et al.*, 2004; Sanderson *et al.*, 2006a). Clinical management of AML patients with a normal karyotype was previously a major challenge, but which is now improving with the identification of distinct sub-groups characterised by somatic base substitutions.

Recent advances in molecular diagnostic techniques have facilitated in the discovery of nucleotide-level mutations that undetectable by standard cytogenetics (Betz and Hess, 2010). AML has fewer mutations than many other adult cancer with an average of 13 somatic driver mutations. In contrast, some solid tumours such as breast, lung or pancreatic cancers can have hundreds of mutations (Cancer Genome Atlas Research, 2013). The most frequent mutations found in CN-AML are *FLT3* internal tandem duplications (ITDs), *FLT3* tyrosine kinase domain mutations, *MLL* partial tandem duplications, and a range of base-substitution mutations, short insertion or deletion mutations within the coding region of the *NPM1*, *CEBPA*, *NRAS*, *WT1* and other genes. Among these mutations, *NPM1*, *CEBPA*, and *FLT3*-ITD mutations now have sufficiently well-established prognostic significance that testing is recommended for CN-AML patients in order to stratify patients for appropriate treatment (Renneville *et al.*, 2008; Döhner *et al.*, 2009)

As well as major cytogenetic abnormalities, sub-cytogenetic somatic mutations also constitute key events in AML pathogenesis and the affected genes have significant roles in key pathways involves in the regulation of haematopoietic cell survival, proliferation or differentiation. The two-hit model of leukaemogenesis has served as a basis to understand how genetic mutations drive AML (Kelly and Gilliland, 2002). According to this model, mutations that frequently detected in AML can be grouped into two classes, class I mutation and class II mutation, which cooperate and cause leukaemic transformation of a myeloid progenitor. This model is supported by

the observation in which a single mutation alone is often insufficient to drive AML transformation. AML-associated chromosomal translocations such as core-binding factor (CBF) translocations [t(8;21), inv(16)] and translocation involving retinoic acid receptor alpha, RARα [t(15;17)] can be found in a proportion of healthy individuals (Quina Ana *et al.*, 2000; Basecke *et al.*, 2002; Mori *et al.*, 2002; Song *et al.*, 2011). It was demonstrated in murine models that these chromosomal translocations have a significant effect on cells by conferring increased self-renewal capacity and reduced differentiation; however, none of these mutations are sufficient to induce malignant transformation (Downing, 2003). Collectively, these findings demonstrate a requirement for both classes of mutation in the same cell for AML transformation.

In addition, genomic studies have revealed the presence of mutations in genes associated with epigenetic modification (*ASXL1*, *DNMT3A*, *IDH1/2*, *TET2*, *EZH2* and *MLL*) in a significant proportion of AML patients (Shih *et al.*, 2012). Disruption of epigenetic processes including DNA methylation and histone modification can lead to altered function in genes involve in key cellular pathways such as DNA repair, RAS signalling, cell cycle and apoptosis, which contribute to malignant transformation (Kanwal and Gupta, 2012). Furthermore, a genome-wide DNA methylation profiling study has demonstrated that specific methylation profiles are associated with specific AML subtypes (Figueroa *et al.*, 2010b), which identified oncogenic cooperativity between somatic alterations in epigenetic regulators and known AML driver mutations. Therefore, an additional class of mutation was defined, termed epigenetic modifiers (Conway O'Brien *et al.*, 2014).

1.2.1. Class I mutations

The class I mutations confer a proliferative and survival advantage to cells and frequently target key components of kinase signalling pathways. These mutations generally occur late and are associated with AML progression (Kelly and Gilliland, 2002). The genes in this group are:

FLT3

The *FMS-related tyrosine kinase 3* (*FLT3*) gene, located on chromosome 13q12 encodes a membrane-bound receptor tyrosine kinase (RTK) involved in the regulation of proliferation, differentiation and apoptosis of haematopoietic cell progenitors (Stirewalt and Radich, 2003). Mutation in *FLT3* is one of the most frequent mutations

observed in AML, found in 25 – 45% of all AML patients (Renneville *et al.*, 2008). ITDs, which result from the duplication and tandem insertion of a variably sized (3–400 nucleotides) fragment of the gene, are the most common mutations in *FLT3* and occur in a wide range of cytogenetic subsets including 20 – 30% of CN-AML cases (Bacher *et al.*, 2007). Other mutations detected in *FLT3* include missense point mutations and small insertion/deletions within the tyrosine kinase domain (*FLT3*-TKD). These mutations result in uncontrolled proliferation and decreasing apoptosis of leukaemic blasts (Gilliland and Griffin, 2002).

RAS

RAS oncogenes encode a family of guanine nucleotide-binding proteins that regulate intracellular signal transduction on binding to a variety of membrane receptors, including *c-KIT* and *FLT3*, and play an important role in proliferation, differentiation and apoptosis (Reuter *et al.*, 2000). H-RAS (Harvey sarcoma virus), K-RAS (Kirsten sarcoma virus), and N-RAS (neuroblastoma RAS viral oncogene homolog) are the most frequently mutated oncogenes in human cancer (Renneville *et al.*, 2008). In AML, point mutations in N-RAS and K-RAS are found in approximately 10 – 15% and 5% of the cases, respectively, with N-RAS codon 12 mutation being the most frequent (43%), followed by N-RAS codon 13 (21%) and K-RAS codon 12 (21%) (Pedersen-Bjergaard *et al.*, 2008; Renneville *et al.*, 2008). These mutations confer constitutive activation of the RAS protein by abrogating guanosine triphosphatase (GTPase) activity that affect cellular proliferation, differentiation and apoptotic processes (Beaupre and Kurzrock, 1999).

KIT

The *KIT* proto-oncogene, located on chromosome 4q12 encodes a transmembrane glycoprotein, which is a member of the type III RTK family, and whose ligand is stem cell factor (SCF). SCF binding promotes c-KIT dimerization and transphosphorylation that leads to activation of downstream signalling pathways involved in proliferation, differentiation, migration and survival, particularly of haematopoietic stem cells (Blume-Jensen and Hunter, 2001). Ligand-independent activation of *c-KIT* can be caused by different types of mutations which lead to deregulation of intracellular signal-transduction pathways mediating development and multicellular communication. *c-KIT* mutations have been identified with high frequency in CBF AML and are mainly found in AML FAB subtype M1, M4, M4eo and particularly in M2, where nearly 70% of

patients have mutated *c-KIT* (Beghini *et al.*, 2000; Goemans *et al.*, 2005). *c-KIT* mutation is a common co-operating event is t(8;21) AML, detected in up to 50% of the cases (Beghini *et al.*, 2000; Care Rory *et al.*, 2003; Goemans *et al.*, 2005; Kohl *et al.*, 2005; Wang *et al.*, 2005; Reikvam *et al.*, 2011).

1.2.2. Class II mutations

In contrast to class1 mutations, class II mutations primarily lead to impaired myeloid differentiation during haematopoiesis by affecting genes involved in transcriptional regulation. These mutations occur early during leukaemogenesis and are stable throughout the disease course and have been proposed to be founder (initiating) mutations (Kelly and Gilliland, 2002). The genes in this group are:

NPM1

The *nucleophosmin 1* (*NPM1*) gene, located on chromosome 5q35 encodes a multifunctional phosphoprotein that shuttles between nuclear compartments and the cytoplasm. Among other functions, one role of *NPM1* thought to be important for leukaemogenesis is its involvement in regulation of ARF and p53 tumour suppressor function (Falini *et al.*, 2009). Insertion/deletion in exon 12 of *NPM1* is the most common mutation in AML, reported in approximately 35% of primary AML cases (Falini *et al.*, 2005). There are different types of *NPM1* exon 12 insertions. Type A mutation (c.860_863insTCTG) are the most frequent *NPM1* mutation in AML, representing 69% of *NPM1* mutations. Type B (c.862_863insCATG) and Type D (c.863_864insCCTG) mutations are also relatively common, representing 11% and 8% of *NPM1* mutations respectively, while other mutations are rare, accounting for <1% of *NPM1* mutations (Alpermann *et al.*, 2016; Kumar *et al.*, 2018). In addition, chromosomal aberrations involving *NPM1* are also a common observation in AML and other haematological malignancies, such as non-Hodgkin lymphoma, acute promyelocytic leukaemia (APL) and myelodysplastic syndrome (MDS) (Falini *et al.*, 2007). *NPM1* mutations are now considered the most common genetic lesion in AML, occurring in approximately 30% and 60% of adult *de novo* cases and CN-AML cases, respectively (Betz and Hess, 2010; Webersinke *et al.*, 2014). Mutation of *NPM1* results in aberrant localisation of the protein in the cytoplasm (Mariano *et al.*, 2006) which could promote tumour growth by inactivation of tumour suppressor p53/ARF pathway, and is thought to be an early event in leukaemogenesis (Thiede *et al.*, 2006; Cheng *et al.*, 2010; Vassiliou *et al.*,

2011). However, there is recent data suggesting that some *NPM1* mutations occur late and are the final transforming event in secondary AML (Schnittger *et al.*, 2014). Although this mutation is associated with a good response to induction therapy and a favourable prognosis, the favourable impact of *NPM1* mutation is highly dependent on *FLT3*-ITD status (Betz and Hess, 2010).

C/EBPα

The *CCAAT/enhancer binding protein alpha* (*C/EBPα*) gene, located on chromosome 19q13.1, encodes a transcription factor that plays a crucial role during differentiation of various cell types. In haematopoiesis, *C/EBPα* plays a key role in early stages of granulocyte development and differentiation from haematopoetic precursors (Tenen, 2003; Nerlov, 2004). Mutations of *C/EBPα* are reported in 15 – 20% CN-AML (Pedersen-Bjergaard *et al.*, 2007), particularly in patients with AML M1 and M2 FAB morphologic subtypes (Pabst and Mueller, 2009; Dufour *et al.*, 2010). In general, *C/EBPα* mutations span the whole protein coding region of the gene and 2 types of mutation are predominantly observed: N-terminal frameshift mutations which result in a truncated form of the CEBPA protein; and C-terminal insertions which result in deficient DNA binding (Pabst *et al.*, 2001). The majority of patients have biallelic *C/EBPα* mutation with many having an N-terminal frameshift and a C-terminal insertion or deletion on different alleles (Pabst and Mueller, 2007). Mutations resulting in loss of C/EBPα function are thought to promote leukaemogenesis by blocking granulocytic differentiation (Pabst *et al.*, 2001). In addition, it was demonstrated that mutation of *C/EBPα* is associated with the upregulation of genes involved in erythroid differentiation, and downregulation of genes involved in driving myeloid differentiation and proliferation (Marcucci *et al.*, 2008). Like mutations in *NPM1*, mutations of *C/EBPα* are also associated with a good response to induction therapy and a favourable prognosis, but the favourable impact is highly dependent on *FLT3*-ITD status.

RARα

The *retinoic acid receptor alpha* (*RARα*) gene, located on chromosome 17q21 encodes a nuclear receptor in humans that regulates transcription in a ligand-dependent manner. Translocations between this locus and several other loci have been associated with acute promyelocytic leukemia (APL), AML of subtype M3. The most frequent translocation is t(15,17)(q21;q22), which fuses the *RARα* gene with the *PML* gene (Kakizuka *et al.*, 1991). The resulting chimeric protein suppresses transcription

of a number of genes containing retinoic acid response elements, leading to blocked myeloid differentiation and promoting cellular proliferation and survival (Pandolfi, 2001).

RUNX1

The human *RUNT-related transcription factor 1* (*RUNX1*, also called *AML1)* gene, located on chromosome 21q22 encodes for one of the two subunits forming the human core binding factor (CBF) (Miyoshi et al., 1991). *RUNX1* is one of the genes most frequently deregulated in leukaemia mainly through chromosomal translocations, point mutations and amplifications. In AML, the RUNX1/ETO fusion oncoprotein arises from a balanced translocation between chromosomes 8 and 21 represents the most common type of mutation affecting the *RUNX1* gene AML and is reported in 10 – 15% of adult AML cases (Kantarjian, 2015). t(8;21) is a defining feature of the FAB M2 AML subtype (Renneville *et al.*, 2008). This fusion oncoprotein cause deregulation of haematopoietic differentiation by targeting a number of haematopoietic transcription factors including PU.1, GATA-1 and C/EBPα (Petrie and Zelent, 2007). The mechanism by which RUNX1/ETO contribute to the AML leukaemogenesis will be further described in Section 1.3.

CBFβ

The *core binding factor beta* (*CBFβ*) gene, located on chromosome 16q22 encodes the β subunit of a heterodimeric core-binding transcription factor belonging to the PEBP2/CBF transcription factor family which functions as a master regulator of numerous genes specific to haematopoiesis (e.g. *RUNX1*) and osteogenesis (e.g. *RUNX2*) (Liu *et al.*, 1993). Mutations in this gene are commonly associated with an inversion on chromosome 16 (inv16), giving rise to a fusion with the *MYH11* gene at 16p13 (Sinha *et al.*, 2015). The resulting CBFβ/MYH11 fusion oncoprotein, has a similar function to the RUNX1/ETO fusion protein in that is can repress transcription leading to impaired haematopoietic differentiation (Reilly John, 2004). Given that both fusion products have similar pathogenic mechanisms they are grouped together and termed CBF leukaemias.

1.2.3. Class III mutations

Class III mutations represent a recently discovered group of mutations that affect proteins involved in epigenetic regulation which can lead to malignant cellular transformation by disrupting key cellular pathways such as DNA repair, RAS signalling, cell cycle and apoptosis (Conway O'Brien *et al.*, 2014). The genes mutated in this group are:

ASXL1

The *additional sex-comb like 1* (*ASXL1*) gene, located on chromosome 20q11.21 encodes for a chromatin-binding protein that has repressive or activating effects on transcription through modification of histone methylation in haematopoietic cells (Chi *et al.*, 2010; Abdel-Wahab *et al.*, 2012). Somatic mutations in the *ASXL1* gene have been reported in several types of myeloid malignancies including MDS, myeloproliferative neoplasia (MPN), chronic myelomonocytic leukaemia (CMML) and AML (Pratcorona *et al.*, 2012). In AML, *ASXL1* mutation occur in approximately 10.8% of *de novo* cases (Chou *et al.*, 2010), with most mutations located in exon 12 of the gene resulting in protein truncation and loss of the PHD domain, an important feature for physical interaction with chromatin (Boultwood *et al.*, 2010; Chou *et al.*, 2010). *ASXL1* mutations in leukaemia are associated with loss of ASXL1 function (Abdel-Wahab *et al.*, 2012) and deregulated expression of several genes involved in haematopoiesis (Davies *et al.*, 2013). In addition, loss of ASXL1 function has been shown to impair granulocyte differentiation in human $CD34^+$ progenitor cells (Davies *et al.*, 2013) and contribute to the development of myelodysplastic syndrome-like disease in mice (Wang *et al.*, 2014).

IDH1/2

The *IDH1/2* genes encode for isocitrate dehydrogenase, a metabolic enzyme that catalyse the conversion of isocitrate to α-ketoglutarate (α-KG) to produce reduced nicotinamide adenine dinucleotide phosphate (NADPH) from $NADP^+$, which is essential for many cellular processes including epigenetic regulation (Clark *et al.*, 2016; Medeiros *et al.*, 2016). Recurring mutations in *IDH* genes are reported in approximately 20% of adult AML and 5% of adult MDS (Figueroa *et al.*, 2010a; Abdel-Wahab and Levine, 2013), with *IDH2* mutations being more common than *IDH1* mutations (Döhner *et al.*, 2015). Both *IDH1* and *IDH2* mutations are frequently

associated with substitution of arginine residues; R132C or R132H for IDH1 and R140Q or R172K for *IDH2* (Molenaar *et al.*, 2015). Expression of mutant IDH1/2 is associated with DNA and histone hypermethylation, altered gene expression and blocked differentiation of hematopoietic progenitor cells (Clark *et al.*, 2016; Medeiros *et al.*, 2016).

TET2

The ten-eleven translocation 2 (*TET2*) gene, located on chromosome 4q24 encodes for tet methylcytosine dioxygenase 2 which plays a key role in DNA demethylation by catalysing the conversion of the modified DNA base methylcytosine to 5-hydroxymethylcytosine (Pastor *et al.*, 2013). Somatic *TET2* mutations are frequently observed in myeloid malignancies including MDS, MPN, CMML, AML and secondary AML (sAML) (Ko *et al.*, 2010). In AML, mutations in *TET2* are reported in 7 – 27% of cases (Weissmann *et al.*, 2011; Gaidzik *et al.*, 2012), and are predominantly associated with CN-AML (Weissmann *et al.*, 2011). In a murine model loss of *TET2* leads to DNA hypermethylation and downregulation of several tumour suppressor genes in haematopoietic cells, conferring a proliferative and survival advantage contributing to the development of leukaemia (Rasmussen *et al.*, 2015).

1.3. RUNX1/ETO Confers a Mutator Phenotype

Cancer is driven by somatic mutations acquired during disease progression. The high number of mutations observed in many cancers led to the mutator phenotype hypothesis, which suggests that an early initiating mutational event in many cancers gives rise to an increase in spontaneous mutation rate (Loeb, 1991). This initiating mutational event acts as a driver for the acquisition of additional mutations which eventually lead to transformation of normal cells to the malignant phenotype. This theory is supported by the observation that constitutional mutations in DNA mismatch repair (MMR) genes confer susceptibility to mutations and promote genomic instability, causing malignant transformation in hereditary non-polyposis colorectal cancer (HNPCC) (Eshleman and Markowitz, 1996). In leukaemia, such a mutator phenotype was identified in chronic myeloid leukaemia (CML) where the BCR/ABL fusion protein promotes the acquisition of point mutations in cells through inhibition of the mismatch repair (MMR) pathway, driving genomic instability and point mutations in tumour

suppressor genes including *p53* and *Rb*, eventually leading to malignant progression of the disease (Stoklosa *et al.*, 2008).

A recent study has demonstrated that expression of the RUNX1/ETO fusion protein confers a mutator phenotype by predisposing cells to the acquisition of somatic mutations, both spontaneously and after exposure to genotoxic chemotherapy, with downregulation of OGG1 expression hypothesised as one possible mechanism (Forster *et al.*, 2015). The *RUNX1/ETO* fusion gene, a product of t(8;21) translocation, is one of the most common genetic lesion observed in AML (Klaus *et al.*, 2004; Sanderson *et al.*, 2006b; Grimwade *et al.*, 2010). It is observed in approximately 40% of FAB M2 AML cases (Peterson *et al.*, 2007) and is defined as a distinct entity according to WHO classification (Arber *et al.*, 2016), having specific clinical and biological characteristics compared to other AML subtype. The resulting fusion oncoprotein was demonstrated to block myeloid differentiation and downregulate several DNA repair proteins, leading to genomic instability and elevated DNA damage which are crucial in AML pathogenesis (Alcalay *et al.*, 2003).

1.3.1. RUNX1/ETO structure and interactions

The t(8;21) translocation fuses the *RUNX1* gene (also called *AML1*) on chromosome 21 with the *ETO* gene (also called *RUNX1T1*) on chromosome 8 to form the *RUNX1/ETO* fusion gene. The expressed RUNX1/ETO fusion protein consists of 752 amino acids comprising the N-terminal portion of RUNX1, with the intact Runt homology domain (RHD) and almost the entire ETO protein, containing four intact nervy-homology regions (NHRs) (Figure 1.3). Numerous *RUNX1/ETO* transcript isoforms are reported in t(8;21) AML patients, which encode truncated proteins shorter than the full-length RUNX1/ETO protein due to alternative splicing (Tighe and Calabi, 1994; van de Locht *et al.*, 1994). One transcript variant of particular interest widely detected in AML is *RUNX1/ETO9a*, characterised by an additional 155bp from exon 9 of ETO (Yan *et al.*, 2006). The *RUNX1/ETO9a* transcript results in the production of a C-terminal truncated RUNX1/ETO protein of 572 amino acids (Yan *et al.*, 2006). Similar to other truncated isoforms of RUNX1/ETO, RUNX1/ETO9a lacks the NHR3 and NHR4 domains which are important for interaction with transcriptional repressors like NCoR and SMRT (Lutterbach *et al.*, 1998; Wang *et al.*, 1998; Zhang *et al.*, 2001). Despite that, RUNX1/ETO9a rapidly induces leukaemia compared to full-length RUNX1/ETO in a murine model (Yan *et al.*, 2006), which demonstrated that disruption

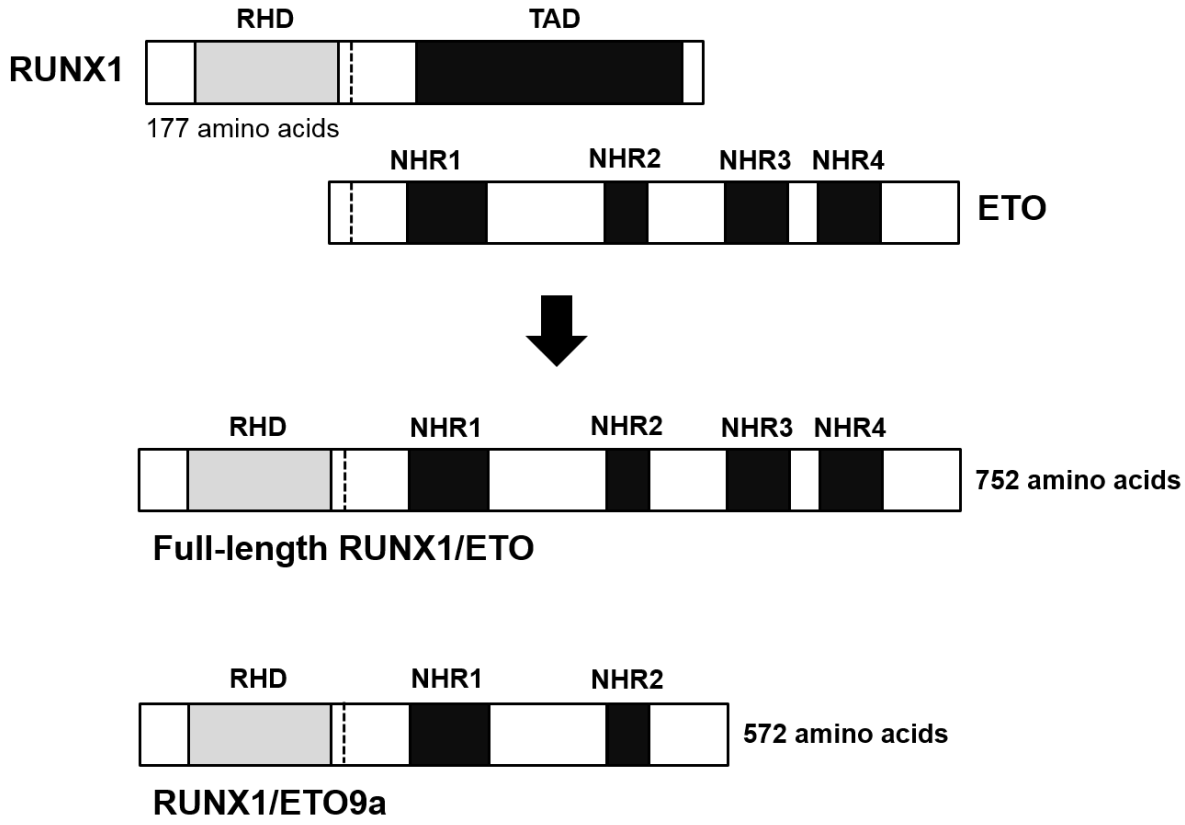


Figure 1.3. Structures of the full-length RUNX1/ETO protein and truncated RUNX1/ETO9a protein.

The RUNX1/ETO full-length fusion protein is comprised of 177 amino acids from the N-terminal of RUNX1 (including a RHD) and almost the whole ETO protein containing four NHRs, produce a 752 amino acid form of RUNX1/ETO. The RUNX1/ETO9a fusion protein comprises the same 177 amino acids of RUNX1 but lacks the NHR3 and NHR4 domains from ETO, giving rise to a truncated protein of 572 amino acids.

of the NHR3 and NHR4 domains via deletion or mutation of full-length RUNX1/ETO enhances leukaemia development (Ahn *et al.*, 2008). However, the exact mechanism by which RUNX1/ETO9a drives leukaemogenesis remains unclear and further studies are needed to delineate the role of this truncated fusion protein, especially because it is frequently co-expressed with full length RUNX1/ETO protein in primary t(8;21) AML (Ommen Hans *et al.*, 2010).

RUNX1 belong to heterodimeric transcription factors family known as CBFs. Each CBF constitutes a non-binding CBFβ subunit and one of three DNA-binding CBFα subunit (RUNX1, RUNX2 and RUNX3) (Ogawa *et al.*, 1993; Levanon *et al.*, 1994). The RHD of RUNX1 mediates DNA binding of RUNX1 with CBFβ subunit to interact with other transcription factors (Figure 1.4a). A wide variety of genes including *IL3*, *GM-CSF*, *MPO* and *P14^{ARF}* are bound by the RUNX1/CBFβ complex, where transcriptional activity is generally up-regulated. Numerous RUNX1/CBFβ target genes are involved in the regulation of haematopoiesis (Meyers *et al.*, 1993). In addition to interacting with CBFβ, RUNX1 can also interact with various other transcription factors and transcriptional co-regulators including ETS1, PU.1, C/EBPα, p300, mSin3a, GATA1, and FLI1 by coordinating their transcriptional activity via its RHD (Gunther and Graves, 1994; Zhang *et al.*, 1996; Petrovick *et al.*, 1998; Gu *et al.*, 2000; Imai *et al.*, 2004; Xu *et al.*, 2006; Huang *et al.*, 2009). These studies show that the function of RUNX1 depends on its interaction partners to regulate target genes in an organised manner.

The four NHRs of ETO, all of which are retained in the RUNX1/ETO fusion protein, mediate interaction with other proteins. Briefly, NHR1 is homologous to several TATA binding protein-associated factors (TAFs) and mediates interactions with E proteins such as inhibition of transcriptional activation to recruit co-activator molecules such as p300/CREB-binding protein (CBP) (Zhang *et al.*, 2004). NHR2 contains a hydrophobic amino acid heptad repeat (HHR) and mediates oligomerization of ETO and its family members (MTG8, MTG16, and MTGR1), as well as interaction with co-repressor proteins such as nuclear receptor corepressor (NCoR), thyroid hormone receptors (SMRT), mSin3a, and histone deacetylases (HDAC) (Kitabayashi *et al.*, 1998; Zhang *et al.*, 2001; Liu *et al.*, 2006). NHR3 contains a coiled-coil structure, which facilitates binding to the regulatory subunit of type II cyclic AMP-dependent protein kinase (PKA RIIα), mediating interaction with PKA anchoring proteins (AKAPs) (Fukuyama *et al.*, 2001). NHR4, a myeloid-Nervy-DEAF1 (MYND) homology domain, contains a zinc chelating structure and mediates interaction with co-repressor

NCoR/SMRT complexes (Lutterbach *et al.*, 1998; Wang *et al.*, 1998; Zhang *et al.*, 2001). Despite their significant involvement in protein-protein interaction, these domains lack DNA-binding ability as compared to RUNX1. In a polymerase chain reaction (PCR)-based DNA-binding site screen ETO was not able to bind DNA, unlike RUNX1 which preferentially bound to DNA sequences with duplicated RUNX1 consensus sites (Okumura *et al.*, 2008).

Based on interaction of ETO with co-repressors it was suggested that the N-terminal portion of RUNX1 in RUNX1/ETO binds to DNA (RUNX1 target gene promoter) via RHD to repress the transcription of RUNX1 target genes (Wildonger and Mann, 2005). The ETO portion of RUNX1/ETO recruits NCoR/SMRT/SIN3/HDAC complexes, leads to decreased histone acetylation and chromatin accessibility at RUNX1 targets (Zhang *et al.*, 2001; Liu *et al.*, 2006; Lin *et al.*, 2017). Taken together, these data provide evidence that RUNX1/ETO drives repression of RUNX1 target genes and inhibits expression of haematopoietic differentiation genes (Figure 1.4b).

1.3.2. RUNX1/ETO cooperating mutations and additional cytogenetic abnormalities

There is a growing body of evidence demonstrating that RUNX1/ETO alone is insufficient for leukaemogenesis, as it requires additional co-operating mutations for progression to AML. Transgenic mice with expression of full-length RUNX1/ETO did not result in the development of leukaemia and remain healthy during their life spans (Yuan *et al.*, 2001). However, following treatment with a strong mutagen (ethylnitrosourea (ENU)), 55% of RUNX1/ETO-expressing transgenic mice developed AML and the remaining transgenic mice and all of the wild-type mice developed acute lymphoblastic leukaemia (ALL). These studies demonstrate a requirement for additional co-operating mutations to facilitate transformation to AML in transgenic mice. In support of this conclusion, t(8;21)(q22;q22) has been reported in non-leukaemic haematopoietic stem cells (HPSCs) of AML patients, potentially representing ancestral cells prior to the acquisition of co-operating mutations (Miyamoto *et al.*, 2000).

Collaboration between different classes of mutations is necessary for the malignant transformation of hematopoietic progenitor cells (Section 1.2). The *RUNX1/ETO* fusion gene is generally recognised as a class II mutation, which affects genes involved in transcriptional regulation and leads to impaired myeloid

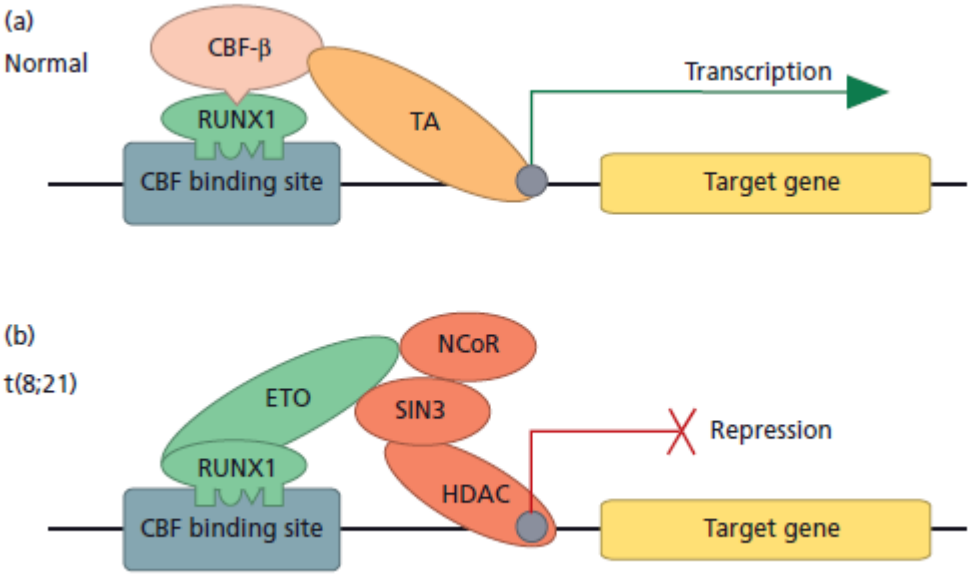


Figure 1.4. Transcriptional repression of RUNX1 target genes by RUNX1/ETO (Vassiliou and Green, 2010).

(a) The normal CBF dimer (composed of RUNX1 and CBFβ) binds specific DNA sequences through the RUNX1 subunit and recruits transcriptional activators (TA) that upregulate the expression of many genes important for haematopoietic cell differentiation. (b) In leukaemic cells carrying the RUNX1/ETO fusion gene, ETO recruits a co-repressor complex comprising NcoR, SIN3 and HDAC to repress transcription of RUNX1 target genes which may lead to inhibition of the expression of haematopoietic differentiation genes.

differentiation during haematopoiesis (Kelly and Gilliland, 2002). Although numerous co-operating mutations have been reported in t(8;21) AML (Table 1.4) (Reikvam *et al.*, 2011), only mutations in certain categories of genes such as type III RTK family were shown to be 'co-operating hits' responsible for driving leukaemic transformation in t(8;21)/RUNX1/ETO AML (Yuan *et al.*, 2001; Higuchi *et al.*, 2002). One example is *FLT3* gene, which encodes a membrane-bound RTK is involved in the regulation of proliferation, differentiation and apoptosis of haematopoietic cell progenitors (Stirewalt and Radich, 2003). Activating mutations of *FLT3* such as *FLT3*-ITD cooperate with RUNX1/ETO to induce leukaemia in a murine bone marrow transplantation model (Schessl *et al.*, 2005).

Mutations in the *KIT* gene represent the most common co-operating event associated with t(8;21) AML, and are detected in up to 50% of the cases (Beghini *et al.*, 2000; Care Rory *et al.*, 2003; Goemans *et al.*, 2005; Kohl *et al.*, 2005; Wang *et al.*, 2005). This gene encodes a transmembrane glycoprotein, and whose ligand is SCF. SCF binding promotes KIT dimerization and transphosphorylation that leads to activation of downstream signalling pathways involved in proliferation, differentiation, migration and survival, particularly of haematopoietic stem cells (Blume-Jensen and Hunter, 2001). Ligand-independent activation of KIT caused by mutations in the gene leads to deregulation of intracellular signal-transduction pathways (Blume-Jensen and Hunter, 2001). Different *KIT* mutations have been detected in t(8;21) AML, with most mutations predominantly occurring in the TKD2 of the gene (Wang *et al.*, 2005; Paschka *et al.*, 2006; Shimada *et al.*, 2006). The most common KIT mutation in TKD2 is N822K with a frequency of 39% (Reikvam *et al.*, 2011), followed by three variants of D816 (D816H, D816V and D816Y) with frequencies between 10 – 35% (Paschka *et al.*, 2006; Reikvam *et al.*, 2011). Oncogenic cooperativity between activating KIT mutation and RUNX1/ETO has been demonstrated in murine models where the N822K mutation was shown to cooperate with the full-length RUNX1/ETO in the induction of a transplantable AML in mice (Wang *et al.*, 2011).

RAS mutations represent another frequent co-operating event in t(8;21), found in approximately 17% of t(8;21) AML cases, with majority of the mutations occurred at codon 12 of *N-RAS* (Krauth *et al.*, 2014). *RAS* is a family of small GTPase proteins that regulate intracellular signal transduction from a variety of membrane receptors, including c-KIT and FLT3 (Reuter *et al.*, 2000). Activating RAS mutations abrogate GTPase activity which affect cellular proliferation, differentiation and apoptotic processes (Beaupre and Kurzrock, 1999). Co-expression of N-RAS G12D mutation

Abnormality	Frequency in t(8;21) AML (%)[a]
Chromosomal abnormalities:	
-X in female patients	30 – 40
-Y in male patients	50 – 60
Del(9q)	13 – 20[a,b]
Trisomy 8	4 – 8[a,b]
Complex karyotype	9 – 23
Gene mutations:	
FLT3-ITD	5
FLT3 D853	3 – 7
KIT mutations	25 – 50
Specifically:	
KIT N288K	39
KIT D816 (D816H, D816V, D816Y)	10 – 35[a,c]
RAS G12D	17[b]
JAK2 V617F	6 – 8

Table 1.4. Genetic abnormalities commonly detected in combination with t(8;21).

[a] Frequency obtained from (Reikvam *et al.*, 2011) unless otherwise stated.

[b] Frequency obtained from (Krauth *et al.*, 2014).

[c] Frequency obtained from (Paschka *et al.*, 2006).

and RUNX1/ETO in human haematopoietic cells increases colony forming and replating ability, which reflects the proliferative advantage conferred by N-RAS G12D, thus promoting progression towards transformation in RUNX1/ETO-expressing cells (Chou *et al.*, 2011).

In addition to gene mutations, a variety of chromosomal abnormalities are observed concurrently with t(8;21). Loss of a sex chromosome represents the most frequent cytogenetic abnormality, with 30 – 40% of *de novo* t(8;21) AML cases having monosomy X in females and 50 – 60% having loss of the Y chromosome in males (Reikvam *et al.*, 2011). The loss of sex chromosomes is associated with aging; however, a higher incidence observed in younger patients with t(8;21) AML (Mertens *et al.*, 1993) indicates possible importance of this abnormality in t(8;21) AML leukaemogenesis. One gene of particular interest is *CSF2RA*, located on the located on the pseudo-autosomal region of the sex chromosomes, encodes the granulocyte-macrophage colony-stimulating factor (GM-CSF) receptor α subunit that functions as a cytokine to promote myeloid proliferation, differentiation and cell survival (Burgess and Metcalf, 1980). A high penetrance of AML was observed in irradiated mice when transplanted with RUNX1/ETO-expressing HSCs that lack GM-CSF (Matsuura *et al.*, 2012), demonstrating potential oncogenic cooperativity between RUNX1/ETO and disrupted GM-CSF signalling via loss of CSF2RA expression in the development of t(8;21) AML.

Other cytogenetic abnormalities recurrently observed in t(8;21) AML include chromosome 9q deletion (del(9q)) and trisomy 8, with incidence of 13 – 20% and 4 – 8%, respectively (Reikvam *et al.*, 2011; Krauth *et al.*, 2014). Oncogenic cooperation between del(9q) and RUNX1/ETO in AML leukaemogenesis is associated with loss of transcriptional co-repressors TLE1 and TLE4 on chromosome 9 that cause increased cell proliferation in RUNX1/ETO-expressing cells (Dayyani *et al.*, 2008). In contrast, the contribution of trisomy 8 to leukaemogenesis in t(8;21) AML is unclear, since it commonly presents as a sole cytogenetic abnormality in AML (Paulsson and Johansson, 2007).

Collectively, these observations indicate that genetic abnormalities affecting transcription factors (class I mutations) and mutations affecting genes involved in signal transduction (class II mutations) represent two classes of the most frequent somatic events in t(8;21) AML, and both are required for malignant transformation. However, the fact that other cytogenetic abnormalities found concurrently with t(8;21) in AML cannot be disregarded as their role in AML leukaemogenesis is still unknown.

1.3.3. Regulation of DNA repair by RUNX1/ETO and susceptibility to mutations

DNA is subjected to a constant barrage of stress from both normal cellular metabolic processes (endogenous stress) and environmental factors (exogenous stress) that can ultimately lead to DNA damage. Endogenous sources of DNA damage include hydrolysis, oxidation and alkylation while exogenous sources of DNA damage include ultraviolet (UV) light, ionising radiation and genotoxic chemicals (Marnett and Plastaras, 2001; Norbury and Hickson, 2001; Fry *et al.*, 2005). DNA damage produced by reactive oxygen species (ROS), a natural by-product of normal cellular oxidative metabolism, is the most frequently occurring damage in cancer cells (Liou and Storz, 2010). ROS are reactive chemical species containing oxygen that include superoxide, hydrogen peroxide, hydroxyl radicals and singlet oxygen. The production of ROS is also influenced by exogenous factors such as exposure to UV and ionising radiation which can increase the level of intracellular ROS (Liou and Storz, 2010). Uncontrolled production of ROS lead to oxidative stress in cells and can cause several types of DNA damage, including oxidized bases and single- and double-strand breaks (Wiseman and Halliwell, 1996; Maynard *et al.*, 2009). Cells are equipped with a number of different DNA repair pathways that attempt to remove damage and restore the structure of DNA, by either repairing the damage or inducing cell death if the damage is irreparable. The base excision repair (BER) pathway is the primary mechanism for the removal of damaged bases caused by oxidation, deamination and alkylation (Wilson and Bohr, 2007), whereas the nucleotide excision repair (NER) pathway is responsible for the removal of bulky DNA adducts caused by UV radiation, mutagenic chemicals or chemotherapeutic drugs (Leibeling *et al.*, 2006). The MMR pathway detects and removes mismatched DNA bases, including those arising due to DNA replication errors, and inserts the correct complementary base into the daughter DNA strand (Jiricny, 2006). Each of these DNA repair pathways excise a damaged region and insert new DNA bases to fill the gap. Furthermore, there are two major pathways for repairing double strand breaks (DSBs); non-homologous end joining (NHEJ) and homologous recombination (HR) (Lieber, 2010). However, in cancer cells, many of the components of DNA repair pathways are dysregulated or mutated, and such defect can result in low efficiency or low fidelity of repair and predispose to mutations.

Evidence suggests that RUNX1/ETO promotes mutagenesis through downregulation of genes involved in several DNA repair pathways (Alcalay *et al.*, 2003;

Krejci *et al.*, 2008). The impairment of DNA repair pathways and the accumulation of DNA damage in RUNX1/ETO-expressing cells is theorised to promote an elevated mutation rate and increase the chances of acquiring subsequent genetic alterations. It is suggested that the mechanisms by which RUNX1/ETO downregulate expression is through direct binding of RUNX1/ETO to conserved RUNX1 binding sites in regulatory regions of DNA repair genes (Krejci *et al.*, 2008).

1.3.3.1. Inactivation of p53 Pathway

Activation of the p53 pathway is a key cellular response to DNA damage and its inactivation plays an important role in the development of many cancers (Kirsch and Kastan, 1998; Amundson *et al.*, 1999). It was demonstrated *in vitro* that p53 is upregulated after exposure to ROS-inducing agents such as hydrogen peroxide and 2-methoxyestradiol, which indicates importance of this pathway in regulating response to DNA damage (Kitamura *et al.*, 1999; Siebert *et al.*, 2011). p53 is a nuclear transcription factor, encoded by *TP53* gene on chromosome 17p13, which modulates cellular responses to DNA damage through the transcriptional activation of genes involved in cell cycle control, DNA repair and apoptosis (Vousden and Lu, 2002; Harris and Levine, 2005).

Tight regulation of p53 function is critical for normal cell growth and development. In normal cells, the expression of p53 protein is controlled through interaction with the Mdm2 protein. Mdm2 forms a complex with p53 that stabilises the endogenous levels of p53 through the proteasomal degradation pathway mediated by Mdm2 E3 ubiquitin ligase (Haupt *et al.*, 1997; Kubbutat *et al.*, 1997). Upon DNA damage, p53 become phosphorylated through post-translational modifications, preventing the formation of the p53/Mdm2 complex resulting in increased endogenous p53 levels (Harris and Levine, 2005). As a transcription factor, p53 mainly exerts its function through transcriptional regulation of its target genes. One of the important genes regulated by p53 is *WAF1*. Activated p53 binds to the promoter region of *WAF1* resulting in increased expression of the encoded protein, p21. p21 then binds to CDK4/Cyclin D complex and inhibits transcription of genes essential for cell cycle progression from G1 phase to S phase, activating cell cycle arrest at G1/S checkpoint. p53-mediated cell cycle arrest allows cells to repair damaged DNA. Once DNA is repaired, p53 degrades spontaneously and cells re-enter the normal cell cycle. In contrast, if the damage is too serious and irreparable, p53 exerts its function by causing

permanent arrest/senescence or activating pro-apoptotic genes to induce apoptosis. Therefore, dysfunction of p53 causes improper binding to target genes, and as a consequence, p21 protein is not transactivated and is not available to act as the "stop signal" for cell division. In this scenario, cell cycle progresses with damaged DNA resulting in fixation to mutation, genomic instability and malignant transformation.

Loss of p53 function is can be caused by mutation in the *TP53* gene itself or by defects in the p53 signalling pathway. Mutation in *TP53* is one of the most frequent genetic alterations in human cancers, found in about 50% of lung, oesophageal, colorectal, breast and ovarian cancers (Hollstein *et al.*, 1991; Olivier *et al.*, 2004). However, *TP53* is rarely mutated in AML, and occurs in only 2 – 3% of all AML cases (Fenaux *et al.*, 1992; Haferlach *et al.*, 2008). Banker and colleagues reported no TP53 mutations in in a group of 28 t(8;21) AML cases (Banker *et al.*, 1998). Given the low rate of *TP53* mutation in AML, including t(8;21) AML, it is thought that p53 is often inactivated by alternative mechanisms.

Numerous mechanisms have been proposed to interfere with p53 signalling in t(8;21) AML, including up-regulation of the *BCL2* gene (Klampfer *et al.*, 1996), silencing of the *p15* and *p16* gene loci (Wong *et al.*, 2000) and direct transcriptional repression of $p14^{ARF}$ (Linggi *et al.*, 2002; Hiebert *et al.*, 2003). The BCL2 protein encoded by is a founding member of Bcl-2 family and constitutes an important regulator of apoptosis by either inducing apoptosis (pro-apoptotic) or inhibiting apoptosis (anti-apoptotic) in a context dependent manner (Kelly and Strasser, 2011). A consensus RUNX1 DNA binding sequence ($TG^{T}/_{C}GGT$) was identified in the 5' regulatory region of the *BCL2* gene, which demonstrated binding affinity for RUNX1/ETO leading to direct upregulation of BCL2 expression (Klampfer *et al.*, 1996) delayed apoptosis in several cell types (Kelly and Strasser, 2011). The p15 and p16 proteins are recognised tumour suppressor proteins in both solid tumours and haematological malignancies (Takeuchi *et al.*, 1995; Herman *et al.*, 1996; Ng *et al.*, 1997). These proteins function as cyclin-dependent kinase inhibitor which form complexes with CDK4 or CDK6 and prevent the activation of cyclin dependent kinases (CDKs) by cyclin D, which subsequently inhibit cell cycle progression from G1 phase to S phase (Takeuchi *et al.*, 1995). Concomitant *p15* and *p16* methylation is reported in M2 and M4 AML subtypes (Wong *et al.*, 2000), and is associated with transcriptional loss of gene expression (Herman *et al.*, 1996; Wong *et al.*, 1998) leading to activation of the CDK/Cyclin D complex and failure of the G1/S checkpoint. Another important component of the p53 pathway is $p14^{ARF}$, which acts as a mediator of p53 by indirectly activating p53 through inhibition of Mdm2 E3

ubiquitin ligase activity (Jing *et al.*, 2003). However, evidence demonstrates that $p14^{ARF}$ is able to bind p53 directly if MDM2 is absent (Sherr, 1998; Lin and Lowe, 2001). The human *$p14^{ARF}$* gene contains eight consensus RUNX1 binding sites in its promoter region, which makes it a direct transcriptional target of RUNX1/ETO (Linggi *et al.*, 2002). Experimental evidence has confirmed reduced endogenous levels of $p14^{ARF}$ expression in cells expressing RUNX1/ETO (Linggi *et al.*, 2002). Loss of $p14^{ARF}$ leads to increased levels of MDM2 and reduced p53 levels, and therefore impairs p53-mediated cell cycle arrest and/or apoptosis.

Collectively, these observations indicate that the mechanisms by which RUNX1/ETO impairs p53 signalling is complex and may involve deregulated expression of various genes. Regardless of the mechanism, disruption of p53 function could lead to inappropriate progression of cell cycle with damaged DNA in the genome, which is hypothesised to be one mechanism responsible in driving mutagenesis in t(8;21) AML.

1.3.3.2. Downregulation of OGG1 by RUNX1/ETO

DNA glycosylases are a group of enzymes that play an essential role in the initiating step of BER in response to a wide range of DNA damage caused by oxidation, deamination and alkylation (Jacobs and Schär, 2012). 8-oxoguanine glycosylase (OGG1) encoded by the *OGG1* gene, is a primary DNA glycosylase enzyme responsible for the excision of 8-oxoguanine (8-oxoG), a common DNA base lesion that occurs as a result of exposure to ROS (Ba and Boldogh, 2018). Indeed, previous studies have demonstrated increased OGG1 expression in response to increased ROS levels (Ishchenko *et al.*, 2003) and oxidative DNA damage (Jankowska *et al.*, 2007). 8-oxoG is a pro-mutagenic DNA lesion which can cause mispairing of guanine with adenine, resulting in G:C to T:A transversion mutations if left unrepaired (Cheng *et al.*, 1992; Boiteux and Radicella, 2000; Suzuki *et al.*, 2010). OGG1 has the ability to directly remove 8-oxoG from DNA and protect against the mutagenic effects of 8-oxoG. Briefly, 8-oxoG lesions are principally detected by OGG1, which activates the BER pathway by catalysing the excision of 8-oxoG from the DNA strand, creating an apurinic (AP) site. OGG1 is also known as a bifunctional glycosylase, due to its ability to create a 3’ incision of the phospodiester backbone of DNA without the need for an independent AP endonuclease (Boiteux and Radicella, 2000; Shinmura and Yokota, 2001). Following the 3’ incision of the phospodiester backbone a 5’ incision is made by

AP exonuclease. Re-synthesis of the resultant DNA gap is mediated by polymerase β and subsequently sealed by DNA ligase I to complete the repair process.

Mutation or deletion of the *OGG1* gene is associated with impaired OGG1 DNA glycosylase activity that can lead to the accumulation of 8-oxoG and elevated spontaneous mutation in cells, which is predicted to contribute to cancer in humans (Shinmura and Yokota, 2001). Disruption of *Ogg1* by partial deletion of the coding sequence in yeast increases spontaneous G:C to T:A transversions in cells (Thomas *et al.*, 1997). In Ogg1-deficient mice, accumulation of 8-oxoG was associated with the development of lung adenoma/carcinoma (Sakumi *et al.*, 2003). Moreover, Ogg1-deficient mice were also susceptible to 8-oxoG formation after ultraviolet B (UVB) exposure, resulting in skin carcinogenesis (Kunisada *et al.*, 2005). In addition, it was shown that substitution of arginine to glutamine at codon 229 (R229Q) caused by point mutation in the human *OGG1* gene results in functional loss of OGG1 enzymatic activity that leads to accumulation of 8-oxoG lesions and activation of apoptosis in the myeloid leukaemia cell line KG-1 (Hyun *et al.*, 2000).

The *OGG1* gene is highly polymorphic in human populations and somatically mutated in some cancers (Shinmura and Yokota, 2001), which is thought to contribute to the loss of OGG1 function. Single nucleotide polymorphisms (SNPs) and somatic mutations in *OGG1* have been associated with increased cancer risk (Boiteux and Radicella, 2000). One of the most common *OGG1* polymorphisms is a C > G transversion at nucleotide position 1245 in *OGG1* exon 7, resulting in a serine to cysteine alteration at codon 326 (S326C; ref SNP ID: rs1052134) (Simonelli *et al.*, 2013) which has been associated with impaired OGG1 DNA glycosylase activity (Kershaw and Hodges, 2012). This polymorphism can be found in 40 – 60% of Asian and 13 – 38% of Caucasian individuals (Morreall *et al.*, 2015), and has been associated with increased risk of stomach, lung, prostate, nasopharyngeal, cervical and oesophageal cancers (Goode *et al.*, 2002; Weiss *et al.*, 2004). In haematological malignancies, rs1052134 is reported with increased risk of childhood ALL (Stanczyk *et al.*, 2011). However, there is no strong evidence that rs1052134 is associated with AML risk.

OGG1 is downregulated in human $CD34^+$ haematopoietic progenitor cells expressing *RUNX1/ETO* when assessed by quantitative real-time PCR (qRT-PCR) (Krejci *et al.*, 2008). In addition, gene expression data from a cohort of AML patient database showed that *OGG1* was significantly downregulated in patient samples with t(8;21) as compared to other sub-types of AML (Liddiard *et al.*, 2010). In a more recent

study, reduced OGG1 protein levels were observed in TK6 RUNX1/ETO clones and *OGG1* transcript levels were found to be inversely proportional to the expression of RUNX1/ETO (Forster *et al.*, 2015). Moreover, siRNA-mediated depletion of RUNX1/ETO resulted in an increase in *OGG1* transcript levels and a concomitant increase in OGG1 protein (Forster *et al.*, 2015). Taken together, these data suggest that OGG1 is downregulated in RUNX1/ETO-expressing cells. Moreover, the RUNX1/ETO fusion protein directly binds to the *OGG1* gene promoter and negatively regulate its transcription (Ptasinska *et al.*, 2012; Forster *et al.*, 2015), providing a plausible mechanism by which RUNX1/ETO downregulates OGG1 expression and drives mutagenesis in t(8;21) AML patients.

1.4. Treatment Strategies for Acute Myeloid Leukaemia

The aim of AML treatment is to restore normal haematopoiesis. To achieve this, the predominant abnormal cells, which are immature leukaemic blasts, need to be eliminated in order to allow normal cells to re-populate the marrow. Significant progress has been made in AML treatment including the use of alternative nucleoside analogues and targeted therapies in specific contexts. However, standard remission induction chemotherapeutic regimens consisting of cytarabine (ara-C) and an anthracycline, remain the main treatment for AML in most patients.

1.4.1. Standard remission induction chemotherapeutic regimes in AML treatment

The nucleoside analogue ara-C in combination with and anthracycline has formed the backbone of intensive AML remission induction chemotherapy for the past 40 years (Saultz and Garzon, 2016). AML treatment includes at least one course of intensive remission induction chemotherapy, followed by an additional course of intensive consolidation therapy (Figure 1.5). The first phase of the treatment aims to improve marrow function by inducing complete remission (CR), a condition in which the return of normal blood count with blast constituting less than 5% of the marrow population, while the second phase of the treatment aims to prolong the CR. The possibility of relapse declines to less than 10% if the patient has been in remission for 3 years (Creutzig and Kaspers, 2004). In AML, the induction period is very intensive and patients do have significant treatment related mortality and morbidities. Such

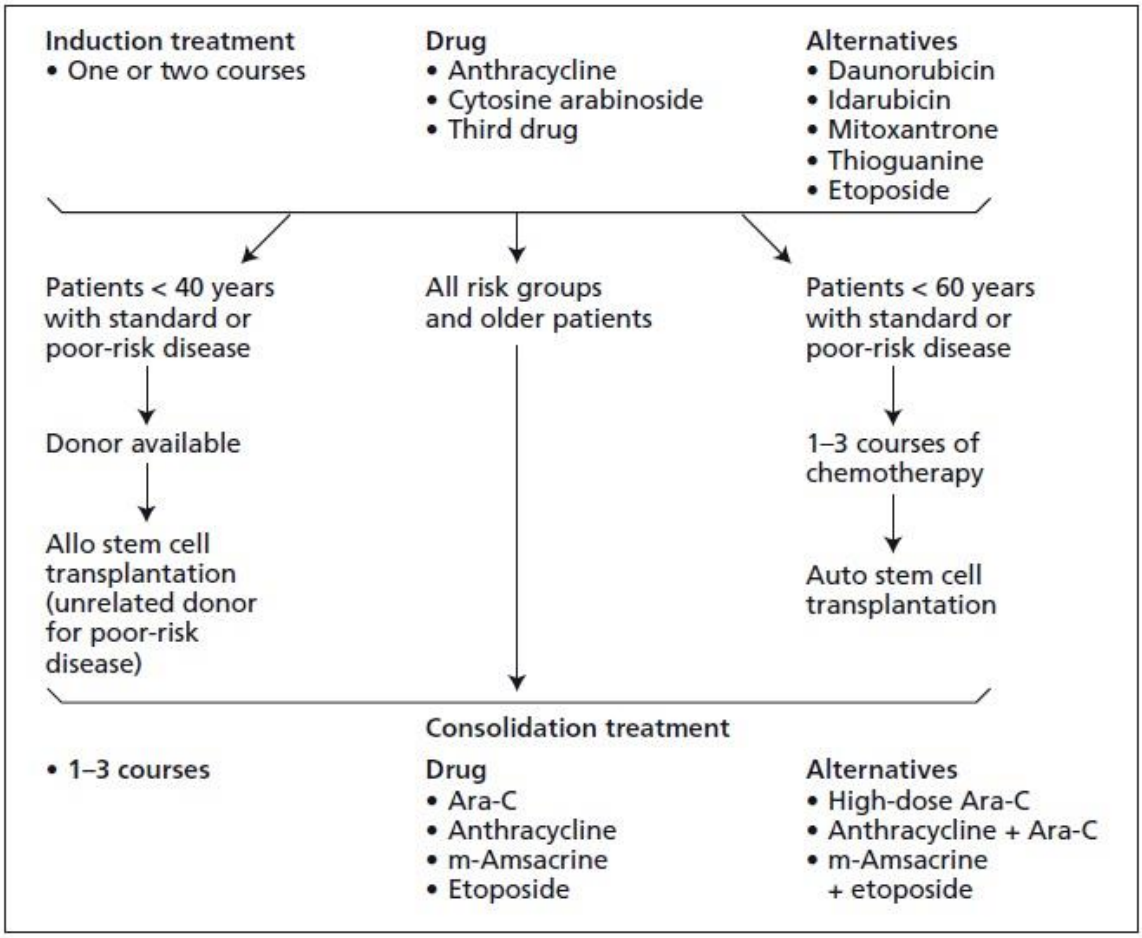


Figure 1.5. Treatment options in acute myeloid leukaemia (Burnett and Venditti, 2011).

treatment is necessary because AML could only be brought under control with intensive chemotherapy. Ara-C is usually administered in combination with daunorubicin (or in some cases idarubicin). Other agents sometimes used in combination with ara-C and daunorubicin include the base analogue 6-thioguanine and topoisomerase inhibitors such as etoposide (Burnett and Venditti, 2011).

Overall, approximately 60 – 80% of adult AML patients will achieve CR following standard remission induction chemotherapy (Shipley and Butera, 2009), with a higher CR rate reported in younger patients (< 60 years) compared to older patients (Saultz and Garzon, 2016). The low CR rate in older patients is partly due to an inability to tolerate aggressive therapeutic regimes. In contrast to AML at presentation, the treatment of relapsed or refractory AML remains a major problem and more difficult to treat, with reported CR rates after the first relapse to be as low as 28% (Larson, 2007), where disease is often associated with a chemoresistant phenotype. Relapsed AML often warrants the use of high-dose chemotherapy or alternative treatments such as targeted therapies.

1.4.2. Alternative nucleoside analogues

Efforts to improve the CR rate have also included the use of alternative nucleoside analogue such as clofarabine and fludarabine, in combination with high-dose ara-C. Both clofarabine and fludarabine share a similar mechanism of action to ara-C, with incorporation of nucleoside analogues into replicating DNA leading to termination of chain elongation and inhibition of DNA synthesis (see Section 1.7.1).

Clofarabine is a next generation nucleoside analogue used for the treatment of ALL in children and young people up to the age of 21 who failed initial treatment. It is also being tested in elderly AML patients who are unable to tolerate intensive chemotherapy and patients with relapsed or refractory AML. Previous studies have evaluated the use of clofarabine in combination with high-dose ara-C in relapsed/refractory AML patients (Faderl *et al.*, 2005) and with low-dose ara-C in elderly AML patients (Faderl *et al.*, 2008), with some evidence of improved CR rates and minimal cytotoxicity.

Fludarabine is an adenosine nucleoside analogue commonly used in the treatment of chronic lymphocytic leukaemia (CLL). It is currently being tested in combination with ara-C in patients with relapsed or refractory AML. The use of fludarabine in combination with high-dose ara-C, granulocyte colony-stimulating factor

and idarubicin significantly increased CR rate and reduced the risk of relapse, particularly in patients with favourable and intermediate risk karyotype (Table 1.3) (Burnett *et al.*, 2013).

1.4.3. Targeted therapies

Increased understanding of the pathogenesis of AML has fostered the development of targeted therapies that utilise the molecules present on the leukaemic cells as a mediator for the chemotherapy drug. The primary aim of a targeted approach is to preserve normal cell function and minimise adverse side effect while maintaining anti-leukaemic efficacy and improving the quality of life of AML patients.

The immunotoxin gemtuzumab ozogamicin (GO), commercially known as Mylotarg®, was the first approved targeted therapy as a single agent for the treatment of relapse in adult AML patients over the age of 60 who are unfit for further intensive cytotoxic chemotherapy (Stasi *et al.*, 2008). GO consists of a monoclonal antibody conjugated to a highly potent anti-tumour antibiotic calicheamicin. This antibody binds to CD33, which is highly expressed on the surface of most AML cells, leading to internalisation of the chemotherapy drug and induction of cytotoxic DNA DSBs (Stasi *et al.*, 2008). GO has also shown remarkable activity in APL, possibly due to the high expression of CD33 on APL blasts (Estey *et al.*, 2002; Lo-Coco *et al.*, 2004), showing promising efficacy for the treatment of other AML patient subgroups. Recently, GO has also been approved for the treatment of adults with newly diagnosed $CD33^+$ AML and for patients aged ≥2 years with $CD33^+$ AML who have experienced a relapse or who have not responded to initial treatment (Appelbaum and Bernstein, 2017).

The advent of next generation sequencing (NGS) technologies has facilitated the discovery of sub-cytogenetic recurrent genetic mutations responsible for the development and progression of AML, which has subsequently driven the development of targeted therapies against mutant driver proteins. Examples of targeted therapies include FLT3 and IDH inhibitors, that are used either as a single agents or in combination with standard chemotherapy drugs to treat newly diagnosed AML patients with *FLT3* and *IDH1/2* gene mutations (Stein and Tallman, 2016). In addition to targeted therapies, the development of novel therapies that targeting biological processes essential to AML cell survival, such as cell cycle inhibitors, have shown promising early results in clinical trials either as single agents or in combination with standard remission induction agents (Saygin and Carraway, 2017).

1.5. Therapy-related Acute Myeloid Leukaemia

Therapy-related acute myeloid leukaemia (t-AML) is recognised as a specific form of AML that occurred as a late complication following cytotoxic therapy (chemotherapy and/or radiotherapy) for a primary cancer, such as Hodgkin lymphoma (HL), non-Hodgkin lymphoma (NHL), ALL, multiple myeloma, breast cancer, ovarian cancer and prostate cancer, or for non-malignant condition in some cases (Bhatia, 2013). The latency period between diagnosis of the primary cancer and development of t-AML varies, which can range between several months to several years and is associated with cumulative dose, dose intensity and type of previous cytotoxic therapy (Godley and Larson, 2008). Two distinct subtypes of t-AML have been recognised, depending on the causative therapeutic exposures: an alkylating agents/radiation-related type and a topoisomerase II (topo II) inhibitor-related type (Vardiman *et al.*, 2002). Abnormal cytogenetics are the frequent observation in t-AML patients, associated with increased prevalence of adverse-risk karyotype (Pedersen-Bjergaard *et al.*, 2008). Patients with alkylating agent-related t-AML have a high incidence of abnormalities involving chromosome 5 (-5/del(5q)) and 7 (-7/del(7q)) (Smith *et al.*, 2003). In contrast, t-AML following chemotherapy with topo II inhibitors is frequently associated with balanced translocations involving chromosome bands 11q23 or 21q22 (Olney *et al.*, 2002). However, the classification of t-AML are no longer subcategorised since most patients received treatment with both alkylating agents and topoisomerase II inhibitors in recent years (Arber *et al.*, 2016). Currently, t-AML account for approximately 10% of all AML cases in the UK (www.hmrn.org). t-AML patients is associated with low CR rate, to be as low as 28% (Larson, 2007).

1.6. Clonal Evolution in Acute Myeloid Leukaemia

Despite the ability of standard regimes to induce remission in 60 – 80% of AML patients at diagnosis, the majority of patients develop relapse within 2 years (Shipley and Butera, 2009). Relapse remains difficult to treat and most patients with AML die from progressive disease after relapse. One reason for this is that relapse often occurs because the leukaemia cells have become resistant to drug treatment, which is associated with clonal evolution of the leukaemic cells.

AML demonstrates clonal heterogeneity, where clones and cells compete with each other for resources and space within the bone marrow microenvironment. Competition generates natural selection that favour clones that grow faster, can disperse and invade. This model of clonal selection and evolution is thought to operate in most human cancers and is similar in principle to Darwin's theory of natural selection, in which heritable traits that help organisms survive and reproduce become more common in a population over time. The genetic heterogeneity of cancer cells is shaped by clonal evolution, a multistep process by which random mutations are acquired during disease progression, resulting in a genetically diverse cell population that is then the subject of natural selection (Nowell, 1976; Merlo *et al.*, 2006; Burrell *et al.*, 2013). The advent of NGS technologies has allowed for a more comprehensive analysis of somatic variants in individual cancers, including the characterisation of very low frequency clones, which has facilitated better understanding of clonal evolution in cancer (Figure 1.6). AML genomes and HSPCs from normal individuals generally contain hundreds of random background mutations which accumulate over the time (Ley *et al.*, 2008; Ding *et al.*, 2012; Welch *et al.*, 2012). Most of these mutations, which can be termed "passenger mutations", are thought to be benign and irrelevant for AML pathogenesis. In addition to non-functional passenger mutations, initiating leukaemia-driver mutations have also been reported and are associated with clonal haematopoiesis of indeterminate potential (CHIP) (Steensma *et al.*, 2015). Specifically, CHIP refers to a common aging-related genetic event in which HSPCs or other early blood cell progenitors acquire somatic mutations, primarily in genes associated with myeloid neoplasm (e.g. *DNMT3A*, *TET2*, *ASXL1*) that drive clonal expansion and results in a genetically distinct subclone, without any apparent sign of haematological malignancy (Steensma *et al.*, 2015). It is suggested that the presence of CHIP increases the risk of subsequent haematological malignancy when acquire an additional advantageous mutation (cooperating mutation) that accelerate the growth of cells which eventually transform to a clone that results in a clinically diagnosed AML phenotype. This is supported by genetic analysis of normal HSPCs and AML blasts from the same AML patients, where normal HSPCs contain pathogenic mutations in *DNMT3A* gene, but without the *NPM1* mutations found in association with the same *DNMT3A* mutations in AML blasts (Shlush *et al.*, 2014). In addition to CHIP mutations,

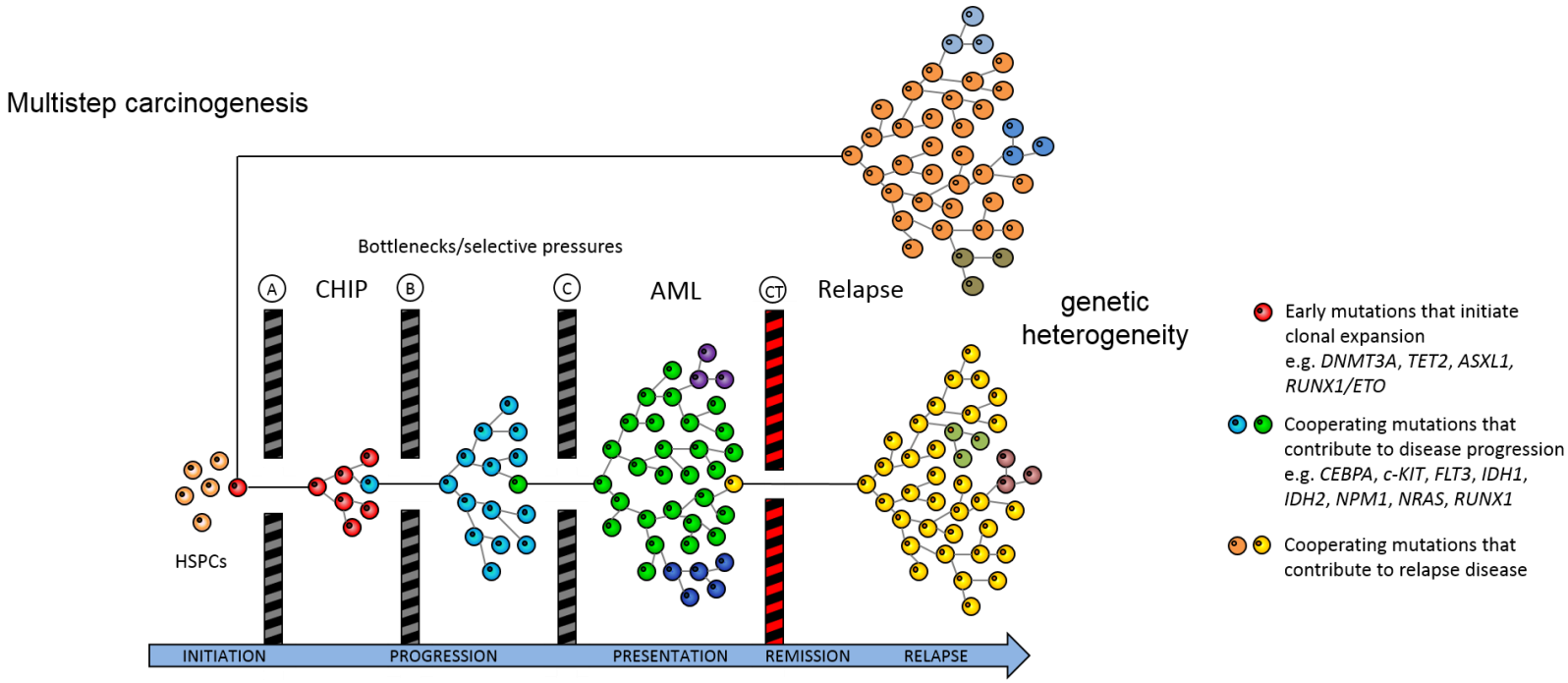


Figure 1.6. Clonal evolution of leukaemia.

Leukaemia, like all cancers, is an evolutionary process that is driven by stepwise, sequential acquisition of somatic mutations and subclonal selection. Hematopoietic stem or progenitor cells (HPSCs) commonly acquire mutations throughout the human lifespan, with many of these are passenger mutations. However, acquisition of certain mutations can confer a survival advantage to the cells and allow clonal expansion without any apparent disease feature, a condition called clonal haematopoiesis of indeterminate potential (CHIP), until they acquire a cooperating mutation for progression to AML. Although remission is achievable after chemotherapy (CT), relapse may occur due to re-population of the dominant presentation clone (green), or as a result of subsequent clonal evolution from either the presentation clone (yellow) or from an ancestral clone (red). In a small subset of cases, the relapse clone is entirely genetically distinct from the presentation clone and may represent a second independent leukaemia or a therapy-related leukaemia. This genetic heterogeneity poses a difficult challenge in the management of relapsed AML.

chromosomal translocations are recurrently found in 6 to 15% of cases (Schoch and Haferlach, 2002) and are thought to play an essential role in the early stages of disease pathogenesis (Döhner and Döhner, 2008; Chen *et al.*, 2010). Many AML-associated chromosomal translocations, including t(15;17) and t(8;21), are found in non-leukaemic cells in a small number of healthy individuals (Quina Ana *et al.*, 2000; Mori *et al.*, 2002), which represent early genetic events in AML pathogenesis. In addition, expression of RUNX1/ETO in Drosophila also causes an expansion of blood cell lineage progenitors (Osman *et al.*, 2009). Like CHIP mutations, these chromosomal translocations also require additional cooperating mutations to induce AML. Taken together, these evidences suggest that cells harbouring these mutations may serve as pre-leukaemic stage until one of the cell acquire an additional cooperating mutation for progression to AML.

Genomic instability is a prominent source of genetic diversity in leukaemic clones and is thought to drive phenotypic adaptability which may contribute to clonal selection during treatment and ultimately result in the evolution of relapse clones. High throughput sequencing of matched presentation and relapse AML samples have revealed two major patterns of AML relapse evolution (Ding *et al.*, 2012; Welch *et al.*, 2012; Garg *et al.*, 2015; Sood *et al.*, 2015). Firstly, a leukaemic clone at disease presentation acquires additional mutations and evolves into the relapse clone after the chemotherapy, and secondly, the pre-leukaemic ancestral clone from the founding clone acquires additional mutations and evolves into a relapse clone via a semi-independent pathway (Figure 1.6). In the former model, the relapse clone is a direct evolutionary descendent of the presentation clone. In the latter model, the relapse clone shares ancestry with the presentation clone, but is not a direct descendent. This is supported by a finding of *RUNX1/ETO* fusion gene in HPSCs of AML patients with long-term remission (Miyamoto *et al.*, 2000), which suggests that the fusion gene persists in a long lived pre-leukaemic ancestral clone that survive chemotherapy. Importantly, both patterns of relapse are defined by the acquisition of additional novel somatic mutations.

The fact that AML relapse is associated with the addition of new mutations and clonal evolution, which is shaped, in part by the chemotherapy, warrant further investigation on the mutagenicity of chemotherapy agents used in AML remission induction treatment. A recent study has showed that ara-C, which is commonly used in AML remission induction treatment, preferentially induces mutation at the central base position of $^{5'}$TpGpA$^{3'}$/$^{5'}$TpCpA$^{3'}$ sequences in the AML genome. These mutations

occur at a significantly higher frequency in relapsed AML patients after exposure to cytarabine-containing regimens (Fordham *et al.*, 2015) compared to chemotherapy-naïve AML, suggesting the existence of a “mutational fingerprint” induced by chemotherapy that can be identified in relapsed AML.

1.7. Chemotherapy-induced DNA Damage

The current treatment of many cancers including AML is still largely based on the use of DNA-damaging chemotherapy agents. Although direct evidence is limited, chemotherapy used to treat AML is predicted to induce mutations detectable in relapsed disease. Chemotherapy contributes to relapse in cancer patients by damaging DNA and generating new mutations that allow tumour cells to evolve and become resistant to treatment. This is supported in a study that genotoxic chemotherapy such as FLT3 inhibitors might predispose to the development of new mutations in a cellular model (Leung *et al.*, 2013). In addition, the type of DNA damage caused by each chemotherapeutic agent can be different, and is determined by the specific mechanism of action of each drug.

1.7.1. Mechanism of ara-C in inducing DNA damage and mutation

Ara-C is a pyrimidine nucleoside analogue which resembles the structure of deoxycytidine in DNA, distinguished by a hydroxyl group in the β-D-configuration on the 2’-carbon of the deoxyribose ring (Figure 1.7). The cytotoxicity of ara-C is primarily mediated by its rapid conversion into ara-C triphosphate (ara-CTP) and incorporation into replicating DNA during cell cycle, resulting in chain termination (Graham and Whitmore, 1970; Zahn *et al.*, 1972). Briefly, the penetration of ara-C into cells is through specific membrane-bound nucleoside transporters (Matsuda and Sasaki, 2005). Once inside cells, ara-C is converted into ara-CTP through a series of phosphorylation events mediated by deoxycytidine kinase. Ara-CTP then competes with CTP for incorporation into replicating DNA during S phase of cell cycle. Following incorporation, ara-CTP induces significant conformational perturbations in the DNA backbone, which subsequently inhibit DNA polymerase binding, leading to induced stalled replication and inhibition of chain elongation (Figure 1.8) (Gao *et al.*, 1991; Schweitzer *et al.*, 1994). This process triggers the intra-S phase DNA damage checkpoint which results in cell cycle arrest (Shi *et al.*, 2001) that allows the removal of ara-C lesion by DNA

repair pathways, particularly BER pathway (Wilson and Bohr, 2007). However, experimental evidence suggests that the exonuclease activities associated with replicative DNA polymerases are inefficient at removing ara-C lesions from the DNA strand (Huang *et al.*, 1991; Perrino *et al.*, 1999).

Another potential mechanism of ara-C-mediated cytotoxicity is incorporation into repair patches during DNA repair. Ara-CTP is an efficient substrate for both polymerase λ (pol λ) and polymerase β (pol β), which are responsible for re-synthesis of DNA during BER pathway (Garcia-Diaz *et al.*, 2010; Prakasha Gowda *et al.*, 2010). Incorporation of fraudulent pyrimidine nucleoside bases of ara-C into DNA by polymerases during BER can decrease the ligation rate of DNA ligases, subsequently leading to to stalled replication and fork collapse (Ewald *et al.*, 2008; Prakasha Gowda *et al.*, 2010).

The mechanisms by which ara-C induces cytotoxicity may also be enhanced through inhibition of the topoisomerase I (topo I) enzyme. In normal conditions, topo I binds to DNA and catalyses the relaxation of supercoiled DNA distal to replication forks during DNA replication by inducing transient single strand break (SSB) and subsequent rejoining (Roca, 1995). Incorporation of ara-C into replicating DNA increases the concentration of topo I-DNA covalent complexes and decrease the re-ligation rate giving rise to unstable protein:DNA complexes that result in the generation of DSBs (Chrencik *et al.*, 2003). If unrepaired, DNA DSBs can induce apoptosis and cell death.

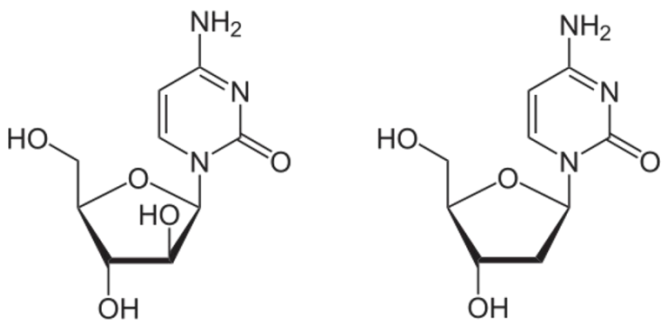


Figure 1.7. Structure of ara-C and deoxycytidine.

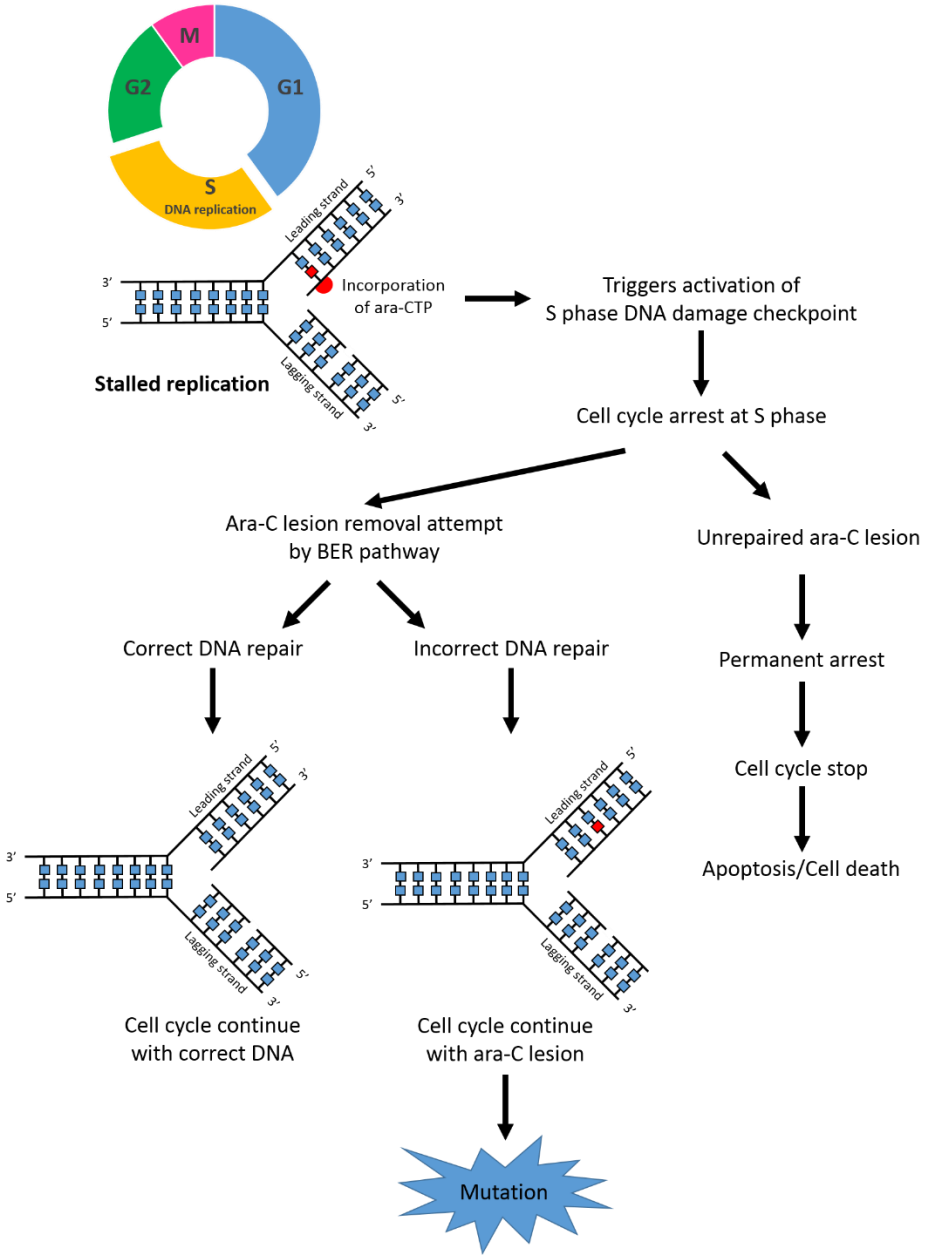


Figure 1.8. Mechanism of action by which ara-C induces DNA damage and mutation.

Incorporation of ara-CTP into replicating DNA during S phase induces conformational perturbations at the site of incorporation, leading to stalled replication and activation of the intra-S phase DNA damage checkpoint. If the damaged base cannot be repaired, permanent arrest can occur which eventually leads to mutation, apoptosis or cell death. Inefficient removal of incorporated ara-C lesions by exonucleases during BER (Huang *et al.*, 1991; Perrino *et al.*, 1999) may also increase rates of mutation or induce cell death.

1.7.2. Mechanism of daunorubicin in inducing DNA damage and mutation

Daunorubicin belongs to the anthracycline class of drugs, which are widely used in the treatment of numerous cancers, including AML (Minotti *et al.*, 2004). Daunorubicin is composed of four tetracyclic ring with an adjacent quinone-hydroquinone group (aglycone moiety) and a sugar (aminosugar 42oiety) attached by a glycosidic linkage (Figure 1.9).

Daunorubicin can intercalate between DNA base pairs and inhibit DNA replication and transcription (Minotti *et al.*, 2004). However, the primary mechanism by which daunorubicin induces cytotoxicity to cancer cells is via the inhibition of the DNA topo II enzyme (Gewirtz, 1999). DNA topoisomerases become targets for several classes of anti-cancer drugs including daunorubicin due to their key function in catalysing the relaxation of supercoiled DNA during DNA replication and transcription, mediated through the controlled transient generation and re-ligation of DSBs (Fortune and Osheroff, 2000). During DNA replication and transcription DNA becomes overwound distal to replication forks (Figure 1.10). Topo II binds to DNA and excises the phosphodiester backbone of both DNA strands to allow the DNA to be unwound temporarily in order to release the torsional strain before being re-ligated (Champoux, 2001; Wang, 2002). However, intercalation of daunorubicin into DNA stabilises topo II-DNA complexes, preventing re-ligation of the break and inhibiting the progression of replication forks (Hande, 2008). Failure to repair DNA DSBs at sites of topo II-DNA complex formation can lead to chromosomal breakage, which can ultimately induce mutation, apoptosis and cell death.

There are two major pathways for repairing DSBs; NHEJ and HR (Lieber, 2010). HR is generally considered to be a high-fidelity repair pathway whereas evidence demonstrates that NHEJ has lower fidelity. Specifically, NHEJ frequently induces insertions or deletions (indels) at the site of breakpoint repair. Although HR-mediated repair is considered to be an error-free repair process, it can also lead to mutations such as base pair substitutions, indels, and complex chromosome rearrangements (Rodgers and McVey, 2016). Gene translocations represent one form of DSBs misrepair that arise from inhibition of DNA topo II enzyme (Richardson and Jasin, 2000). Several partner genes involved in leukaemis-specific translocations are proposed to occur as a result of misrepair of DSBs induced by daunorubicin-mediated inhibition of topo II. For example, fusion of *AF4* and *AF9* in the t(9;11) translocation, *PML* and *RARA* in the t(15;17) translocation and *RUNX1* and *ETO* in the t(8;21)

translocation are all thought to be induced by exposure to topo II poisons such as daunorubicin. Furthermore, translocations involving *MLL* at chromosome 11q23 (Lovett *et al.*, 2001; Zhang *et al.*, 2002; Whitmarsh *et al.*, 2003; Mistry *et al.*, 2005) are also common in cancers that arise following exposure to topo II posions.

In addition to inhibiting ligation of topo II-cleaved DNA complexes, daunorubicin can also redox cycle resulting in the formation of free radicals, such as reactive oxygen species (ROS), that damage DNA, proteins and cell membranes (Powis, 1989). Specifically, the quinone moiety of daunorubicin can undergo enzymatic reduction forming semiquinone and oxygen radicals (Bachur *et al.*, 1978; Bates and Winterbourn, 1982; Muindi *et al.*, 1984). However, the contribution of free radical generation to the cytotoxicity of the anthracyclines remains unclear, as there is no strong evidence demonstrating free radical generation at clinically relevant concentrations of anthracyclines and at the oxygen tension in cancer cells (Gewirtz, 1999).

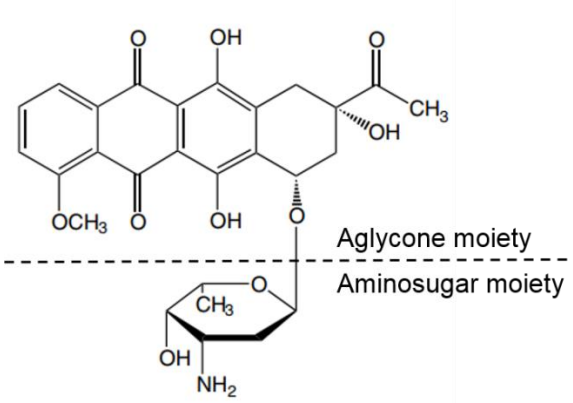


Figure 1.9. Structure of daunorubicin.

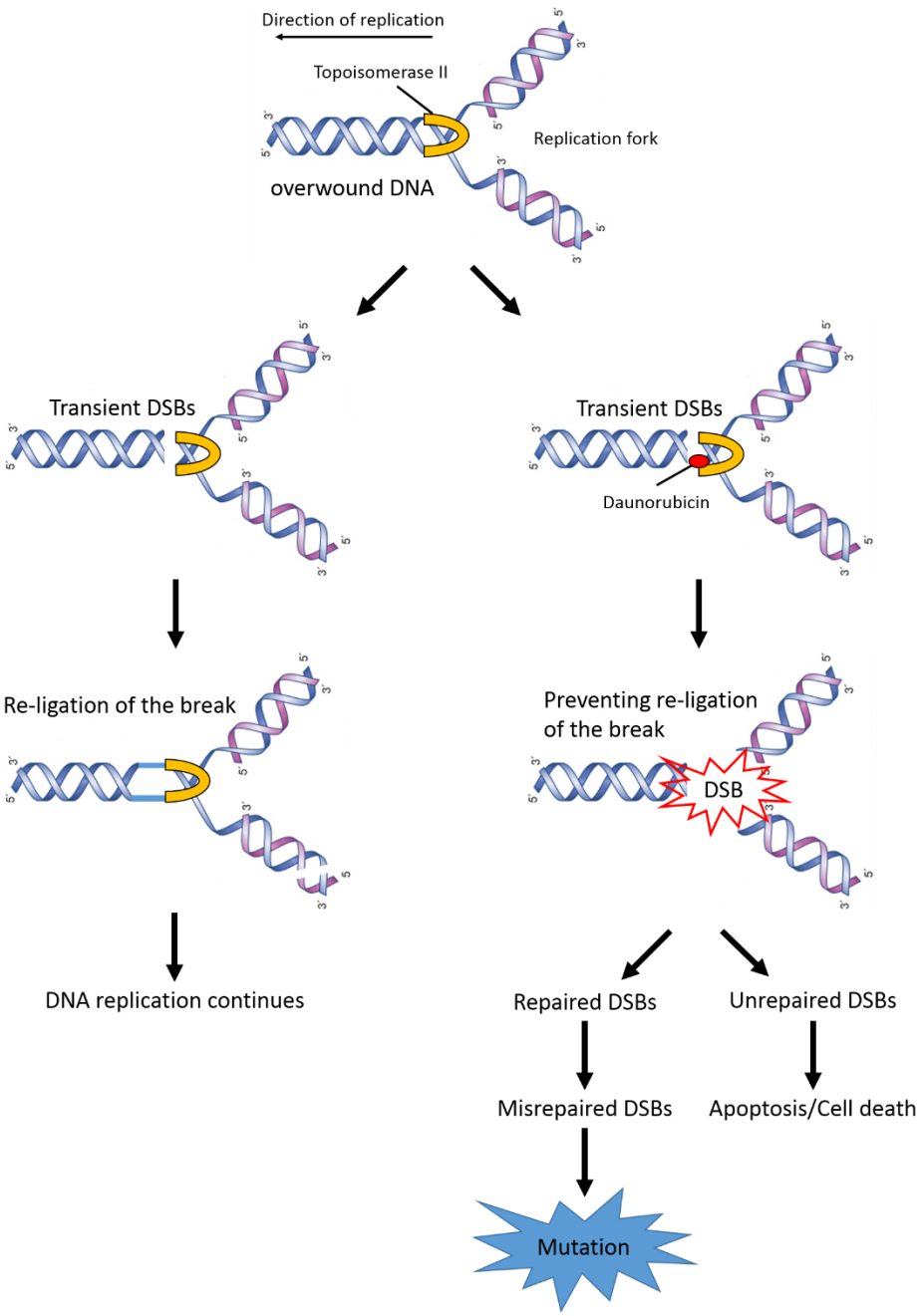


Figure 1.10. Mechanism of daunorubicin-induced DNA damage and mutation. DNA topoisomerase II (topo II) is responsible for relaxing the torsional strain in overwound DNA through controlled transient generation of double strand breaks (DSBs) and re-ligation of the breaks. Intercalation of daunorubicin between DNA base pairs stabilises topo II-cleaved DNA covalent complexes, preventing re-ligation of the break, thus inhibiting the progression of replication forks. If the damage cannot be repaired, the induced DSBs can ultimately lead to mutation, apoptosis or cell death.

1.8. Aims of Project

Although remission is achievable in AML patients, relapse remains a major clinical problem and most patients with AML die from progressive disease after relapse. It is thought that relapse is driven by additional cooperating somatic mutations that were acquired (or selected) after chemotherapy, which drive phenotypic adaptability and clonal selection during treatment. The mechanisms by which novel mutations are acquired at relapse remain unclear, although there is strong evidence that some are chemotherapy-induced. Despite considerable efficacy in the treatment of cancer, many chemotherapeutic agents are also mutagenic and could contribute to the aetiology of relapse-driver mutations.

Deep sequencing has identified two patterns of AML relapse evolution, where in both patterns, the relapsing clone evolves from a cell carrying leukaemia-initiating genetic lesions, including gene fusions such as *RUNX1/ETO*. Moreover, evidence suggests that the *RUNX1/ETO* fusion gene confers a mutator phenotype which predisposes cells to the acquisition of additional cooperating mutations through downregulation of DNA repair. Therefore, understanding how chemotherapy and fusion gene expression cause mutations in cells that survive remission induction is fundamental in preventing the acquisition of relapse-driver mutations and developing approaches to reduce AML relapse risk.

The primary objective of this project was to investigate the contribution of remission induction chemotherapy and fusion gene expression to the aetiology of pathogenic mutations in AML. The strategies for achieving this were:

- To determine the effect of chemotherapeutic agents, particularly ara-C and daunorubicin on mutation frequency *in vitro* using a cell model system.
- To determine whether RUNX1/ETO fusion gene expression affects mutagenesis *in vitro*, both spontaneously and after exposure to ara-C and daunorubicin.
- To investigate whether RUNX1/ETO fusion gene expression predisposes to mutation at specific sequence of the genome.
- To determine whether RUNX1/ETO fusion gene expression predisposes cells to mutations in leukaemia relevant genes, particularly *NPM1* gene.

- To perform high-throughput exome sequencing of a cohort of matched presentation and relapse AML samples, in order to investigate the clonal architecture of AML and clonal evolution during AML treatment.

Chapter 2. Materials and Methods

2.1. Chemicals and Reagents

All chemicals and reagents used were of Analar grade and purchased from Sigma Aldrich Co. Ltd. (Dorset, UK) unless otherwise stated. Phosphate buffered saline (PBS) was prepared from PBS tablets (Life Technologies, Paisley, UK) dissolved in deionised distilled water (ddH_2O) and autoclaved prior to use for sterility. Preparation of all additional reagents for use in specific experiments are detailed in relevant sections.

2.2. Cell Lines

TK6 B lymphoblastoid cells were a kind gift from Professor William Thilly (MIT, MA, USA). TK6 RE1 and TK6 RE8 cell lines were engineered to express RUNX1/ETO by Dr Vicky Forster (Northern Institute for Cancer Research (NICR), Newcastle, UK). These cell lines were used throughout the project to determine the effect of chemotherapy and the *RUNX1/ETO* fusion gene expression on the acquisition of somatic mutations. However, TK6 and its derivative cell lines are B lymphoblastoid cells; therefore, consideration needs to be given to expression of recombination-activating genes (RAGs) and activation-induced cytidine deaminase (AID) in relation to the types of mutations seen. RAG expression is important for V(D)J genes recombination during early B-cell development and AID drives somatic hypermutation via deamination of cytosine to uracil in order to produce antibody diversity during late stage B cell development. Both of these pro-mutagenic processes are known to inappropriately target non-immunoglobulin genes and induce somatic mutation in other parts of the genome (Jones and Gellert, 2004; Larson and Maizels, 2004; Teng and Papavasiliou, 2007).

Specific features of TK6 and other cell lines used in this project are described in Table 2.1. All cell lines were defrosted from liquid nitrogen stocks held in the Northern Institute for Cancer Research (NICR).

2.3. General Cell Culture Methods

2.3.1. Routine cell culture

Cell culture RPMI-1640 medium (Roswell Park Memorial Institute 1640) and reagents were purchased from Sigma Aldrich Co. Ltd. (Dorset, UK). Fetal bovine serum (FBS) was purchased from Gibco (ThermoFisher Scientific) and cell culture plasticware was from Corning-Costar (VWR International Ltd., Leicesterchire, UK). All cell culture work was performed in a class II microbiological safety cabinet ((BIOMAT-2), Medical Air Technology., Oldham, UK)) using aseptic technique. All cell lines were maintained as suspension cultures in sterile 25cm^2 cell culture flasks and passaged as required (Table 2.1) using 10 ml RF10 medium [RPMI-1640 medium supplemented with 10% (v/v) FBS and 50 µg/ml penicillin/streptomycin]. Cultures were incubated at 37°C in a humidified 5% CO_2 incubator (Hareus Equipment Ltd., Essex, UK). Experiments were performed on exponentially growing cells with fewer than 25 passages. Testing for mycoplasma was performed by E. C. Matheson (Northern Institute for Cancer Research (NICR), Newcastle, UK) at 3 month intervals using a MycoAlert kit (Lonza Biologics, Slough, UK).

2.3.2. Cell counting and determination of dell density

A 10µl aliquot of cells in suspension was mixed with 10µl 0.4% solution of trypan blue (Life Technologies, Paisley, UK) in a 1:1 ratio and 10µl of the mixture was loaded onto a Neubauer haemocytometer (VWR International Ltd.) and counted using established techniques. An average from at least two separate counting for each cell sample were taken to determine the cell concentration.

2.3.3. Cryopreservation of cell stocks

Cell stocks from all cell lines were prepared and stored long term in liquid nitrogen. An appropriate volume of cell suspension containing 5 x 10^6 cells was dispensed into a sterile 15ml BD Falcon™ centrifuge tube (BD Biosciences, Oxford, UK) and centrifuged at 300 x g for 5 minutes (min). Supernatant was discarded and the cell pellet was resuspended 1ml freezing medium [FBS supplemented with 10% (v/v) dimethyl sulphoxide (DMSO)]. The cell suspension was then transferred to a

sterile 1.8ml polypropylene cryovial (Invitrogen Life Technologies) and frozen slowly inside a Mr. Frosty™ Freezing Container (Thermo Fisher Scientific) by incubation at -80°C overnight before being transferred to liquid nitrogen.

2.3.4. Resuscitation of frozen cell stocks

Frozen cell aliquots were quickly thawed at 37°C and immediately transferred to 5ml of pre-warmed RF10 medium in a sterile 15ml BD Falcon™ centrifuge tube followed by centrifugation at 300 x g for 5 min. Supernatant was discarded and the cell pellet was resuspended in 5ml RF10 medium. Cell suspensions were then transferred to sterile 25cm^2 cell culture flasks and monitored daily for growth. Once cell populations had recovered, suspensions were maintained as in Table 2.1 in preparation for use in experiments.

Cell Line	Morphology	Optimum Density (x10^6 cells/ml)	No. of Days between Passage	Split
TK6	Lymphoblastoid, non-leukaemic	0.2 – 1.0	2	1:10
TK6 RE1	Lymphoblastoid, *RUNX1/ETO*-transduced, clone 1	0.2 – 1.0	2	1:10
TK6 RE8	Lymphoblastoid, *RUNX1/ETO*-transduced, clone 8	0.2 – 1.0	2	1:10
Kasumi-1	Myeloid leukaemia with t(8;21)	0.3 – 3.0	3	1:10
AML-3	Myeloid leukaemia with *NPM1* mutation (type A) and the *DNMT3A* R882C mutation	0.5 – 2.0	3	1:8
THP-1	Monocytic leukaemia with *RAS* G12D mutation	0.2 – 0.8	3	1:8

Table 2.1. Cell lines used in this project and their routine passage schedule.

2.3.5. Preparation of cell pellets

Aliquot of cell suspension containing 5 x 10^6 cells were dispensed into 15ml BD Falcon™ centrifuge tubes and centrifuged at 300 x g for 5 min. Supernatant was discarded and cell pellets were washed using 3ml cold sterile PBS. Suspension was centrifuged again and cell pellets were resuspended 1ml cold sterile PBS. Suspension was transferred to a sterile 1.5ml Eppendorf tube and centrifuged at 300 x g for 5 min. Supernatant was discarded and cell pellets were stored at -80°C until required.

2.4. Western Immunoblotting

Western blotting was used to quantify RUNX1/ETO fusion protein expression in *RUNX1/ETO*-transduced cell lines relative to their respective parental cells (Chapters 3 and 5), and also to determine the specificity of mutant NPM1-specific antibodies (Chapter 5). Western immunoblotting comprises extraction of proteins, electrophoretic separation of proteins according to size on a polyacrylamide gel followed by immobilisation onto a membrane, and subsequent detection of target proteins of interest using labelled antibodies. The number of antibody molecules that bind to the target protein is proportional to the amount of protein present, therefore protein expression levels can be semi-quantitatively compared between different cell lines.

2.4.1. Preparation of protein extract

A frozen cell pellet containing 5 x 10^6 cells (described in Section 2.3.5) was thawed and resuspended in 150µl sodium dodecyl sulphate (SDS) sample buffer [62.5M Tris-HCl pH 6.8, 2% (w/v) SDS, 20% (v/v) glycerol]. The cells were then disrupted by passing through a 21-gauge needle (attached to a 1ml syringe) several times. The cell lysate was heated at 100°C for 5 min in a heat block followed by centrifugation at 14,000 x g for 10 min to pellet cell debris. The supernatant was transferred to a fresh sterile Eppendorf for quantification of protein concentration using the Pierce BCA assay as described in Section 2.4.2.

After determination of protein concentration, 250µg of protein was transferred to a fresh 1.5ml Eppendorf tube and made up to 200µl using SDS sample buffer followed by the addition of 25µl 0.2% bromophenol blue aqueous solution (final concentration 0.01% v/v) and 25µl β-mercaptoethanol (final concentration 20% v/v),

resulting in a protein concentration of 1mg/ml. Aliquots of 50µl were stored at -20°C until required.

2.4.2. Quantification of protein concentration by Pierce BCA assay

Pierce® BCA Protein Assay Kit (Fisher Scientific UK Ltd., Leicestershire, UK) was used for the quantification of protein concentration in the extract. The principle of this assay is based on colorimetric detection following reduction of copper cations (Cu^{+2} to Cu^{+1}) by protein with bicinchoninic acid (BCA) in an alkaline medium, results in the generation of a purple coloured complex. The absorbance of this coloured complex can be measured at 562nm and increases linearly with increasing protein concentration, thus allowed for an estimation of protein concentration to be made. The assay was performed according to manufacturer's protocols which are described as follows:

A fifty microliter aliquot of diluted extract was made by adding 5µl of neat extract to 45µl dH_2O (1:10 dilution) to give an approximate protein concentration within the detection range of the kit (0.2 – 1.2 mg/ml). A set of standard dilutions of bovine serum albumin (BSA) was prepared according to Table 2.2. For each standard and diluted extract, 10µl was transferred into four wells of a 96-well flat-bottomed plate (VWR International Ltd.) starting from the second column of the plate and 10µl of dH_2O was added to each eight wells in the first column of the plate as blanks. An appropriate volume of BCA working reagent was prepared (Reagent A + Regent B in a 50:1 ratio) and 190µl was added into each well and the contents of the plate were gently mixed on a plate shaker, followed by incubation at 37°C for 30 min.

Following incubation, the plate was placed in a Spectromax® 250 Microplate Spectrophotometer System (Molecular Devices Corporation, Crawley, UK) for absorbance reading at 562nm. Protein concentrations were calculated via comparison of a corrected mean absorbance value of the diluted extract to a standard curve generated by the set of BSA standards (0.2 – 1.2 mg/ml). The concentration of the neat protein extract was determined by multiplying the measured value by 10 (to account for dilution factor).

Volume of BSA stock solution (µl)	Volume of dH_2O (µl)	Final concentration of standard (mg/ml)
240	160	1.2
200	200	1.0
160	240	0.8
120	280	0.6
80	320	0.4
40	360	0.2

Table 2.2. Preparation of BSA standard solution.

2.4.3. SDS PAGE and electrophoretic transfer

Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) was performed for separation of proteins in the extract according to size. Mini-PROTEAN® TGX™ 4-20% Tris-Glycine precast gels (Bio-Rad) were placed in a mini PROTEAN Tetra system filled with SS electrode buffer [41.2mM Tris-HCl pH 6.8, 192 mM glycerol, 0.1% (w/v) SDS]. In the first lane of each gel, 10µl of PageRuler™ Prestained Protein Ladder (ThermoFisher Scientific) was loaded. For each sample, 20µl of cytosolic extract (20µg of protein) was loaded per well. Electrophoresis was performed using a constant voltage of 150V for approximately 1 hour.

Following separation by electrophoresis, proteins were transferred from the gels onto a polyvinylidenefluoride (PVDF) membrane (Bio-Rad Laboratories Ltd., Hertfordshire, UK) via another round of electrophoresis. Briefly, gels were removed from their casing and assembled on the cassette accordingly to Figure 2. PVDF membranes were soaked in 100% methanol for 20 seconds (sec) and placed in transfer buffer [10mM CAPS-NaOH pH 11, 10% (v/v) methanol] before use. Gels were removed from their casing transferred onto the membrane before being sandwiched between two 3mm Whatman® chromatography papers (VWR International Ltd.) and two transfer sponges, all of which had been pre-soaked in transfer buffer. Cassettes were placed in a Mini Trans Blot Electrophoretic Transfer Cell (Bio-Rad Laboratories Ltd.) filled with transfer buffer and electrophoresis was performed overnight at a constant voltage of 100V using a magnetic stirrer to maintain ion distribution within the buffer and with an ice block to prevent extreme increases in temperature.

2.4.4. Antibody detection and visualisation of bound proteins

PVDF membranes with bound proteins were removed from transferred cassettes and placed in 50ml BD Falcon™ tubes containing 20ml of 5% blocking solution [TBS/Tween [0.01M Tris-HCl pH 7.5, 0.1M NaCl, 0.05% (v/v) Tween-20] and 5% (w/v) dried non-fat skimmed milk powder]. The membranes were constantly agitated on a roller mixer at room temperature for one hour to block non-specific antibody binding sites. After the blocking, membranes were divided into appropriate sections depending on the size of protein of interest (as estimated using the protein ladder). Each section of the membrane was transferred to a separate 50ml BD Falcon™ tube containing 5ml of primary antibody solution (prepared according to

dilutions specified in Table 2.3). All membranes were incubated overnight at 4°C on a roller mixer (with constant gentle agitation).

Following exposure to appropriate primary antibodies, membranes were transferred to a fresh 50ml BD Falcon™ tube and washed three times using 5ml TBS/Tween [0.01M Tris-HCl pH 7.5, 0.1M NaCl, 0.05% (v/v) Tween-20] on a roller mixer at room temperature for 10 min each wash. Wash steps were performed to ensure removal of any unbound primary antibody. Membranes were then transferred to fresh 50ml BD Falcon™ tubes containing appropriate secondary antibody conjugated to horseradish peroxidase (prepared according to dilution specified in Table 2.3) and incubated at room temperature for 1 hour on a roller mixer (with constant gentle agitation). After exposure to secondary antibodies, membranes were washed four times in the same way as above to remove any unbound antibody.

Bound antibodies were detected using Amersham™ ECL™ Prime Western Blotting Detection Reagent (GE Healthcare Ltd., Buckinghamshire, UK) according to the manufacturer's protocol. Chemiluminescence from bound antibodies was visualised by exposing the membrane to Carestream Kodak BioMax light film (Sigma Aldrich Co. Ltd., Dorset, UK) for between 5 sec and 10 min, depending on the strength of the signal. Films were developed using a Mediphot 937 X-Ray Filmprocessor (Colenta Lobortechnik, Austria).

2.4.5. Quantification of protein bands

The developed film was placed in a LAS-3000 Luminescent Image Reader V2.2 (Fujifilm) to capture the image of the protein bands by using white light transilluminator. The captured protein bands were then quantified in relative to control sample using AIDA Image Analyser V3.28 software (Elysia-Ratest, Germany).

Antibody		Host / Type	Supplier	Dilution
Primary	AML1 (RUNX1)	Rabbit, Polyclonal	Merck Millipore #PC285	1 : 40
	NPM1	Rabbit, Polyclonal	Invitrogen #PA5-12446	1 : 1000
	NPM1 (mutant)	Rabbit, Polyclonal	Invitrogen #PA1-46356	1 : 1000
	RAS	Rabbit Monoclonal	Cell Signalling Tech. #8955s	1 : 1000
	RAS (G12D mutant)	Rabbit monoclonal	Cell Signalling Tech. #14429	1 : 1000
	GAPDH	Rabbit, Polyclonal	Santa Cruz Biotech. #SC25778	1 : 400
Secondary	Anti-rabbit	Goat, Polyclonal	Dako #P0448	1 : 2000

Table 2.3. Antibodies used in western immunoblotting.

2.5. Cytotoxicity Assay

Cytotoxicity assays were performed to assess drug response in TK6 and its derivative subclones, in order to determine the efficacy of chemotherapeutic agents on the survival and proliferation of cells and to identify appropriate doses for evaluating the effect of chemotherapeutic agents on mutation frequency (MF) (Chapter 3).

2.5.1. Preparation of chemotherapeutic agents

All chemotherapeutic agents used for in *in vitro* cytotoxicity and mutation assays were purchased from Sigma-Aldrich Co. Ltd. (Dorset, UK). Stock solutions were prepared according to Table 2.4 using sterile ddH_2O and appropriate working solutions were prepared by dilution of stock solutions in sterile ddH_2O before use in experiments.

2.5.2. 96-well clonogenic assay

Clonogenic assays were performed to assess the ability of TK6 cells to grow and survive at a very low density (i.e. 10^1 cells/ml) following exposure to appropriate doses of daunorubicin.

An accurate cell count was performed to and a cell suspension at a density of 1 x 10^5 cells/ml in 10ml RF10 medium was generated for each cell line. Following this, a 10-fold serial dilution was carried out using RF10 medium to derive cell suspensions at 1 x 10^3 cells/ml, 1 x 10^2 cells/ml and 1 x 10^1 cells/ml (each in a volume of 30ml). An appropriate volume of daunorubicin working solution (depending on the final concentration of daunorubicin) was added to each cell suspension in a 50ml BD Falcon™ tube. Drug exposed cells (200 µls) were then transferred to each well of a 96-well culture plate (VWR International Ltd.). The lowest doses were applied to cell suspensions at a density of 1x10^1 cells/ml (2 cells/well), mid-range doses were applied to cell suspensions at a density of 1x10^2 cells/ml (20 cells/well) and the highest doses were applied to cell suspensions at a density of 1x10^3 cells/ml (200 cells/well). A vehicle-only control was also prepared by adding an appropriate volume of solvent to 30ml of 1x10^1 cells/ml cell suspension and 200µl was transferred to each well of a plate. The design of the plating for each cell line is illustrated in Figure 2.1. All plates were incubated at 37°C/5% CO2 for approximately 20 days.

Following 20 days growth period, the number of negative wells (wells in which no viable colony had grown) was scored per plate at each dose including vehicle-only control. Calculation of cloning efficiency (CE) was performed using the following formula (Furth *et al.*, 1981) :

$$Cloning\ efficiency\ (CE) = \frac{-\ln(No.negative\ wells \div total\ no.of\ wells)}{No.of\ cells\ per\ well}$$

All calculations were performed using Microsoft Excel. The survival fraction at each dose was determined by calculating the mean CE of the dosed cells as a percentage of the mean CE of the vehicle-treated cells and each experiment was repeated three times. The dose-response curve was generated using mean survival fractions and standard error of the mean (SEM) of the data from three independent experiments using GraphPad Prism 6 software (San Diego, USA).

2.5.3. Dose finding assay

A dose-finding assay was carried out to identify appropriate doses inducing relevant cytotoxicity following 16 hours exposure to escalating doses of chemotherapeutic agents prior to use in mutation assays.

An accurate cell count was performed to generate a cell suspension at a density of 1 x 10^5 cells/ml in 10ml RF10 medium. An appropriate volume of working concentration drug solution was added to each flask and incubated at 37°C/5% CO2 for 16 hours. Following incubation, the cell suspension was transferred to 15ml BD Falcon™ centrifuge tube and centrifuged at 300 x g for 5 min. Supernatant was discarded and cells were washed twice using 3 ml sterile PBS and resuspended in fresh RF10 medium for cytotoxicity assessment.

An accurate cell count was performed to determine the volume of cell suspension required to give 5 x 10^5 cells/ml. An appropriate volume of cell suspension was dispensed into a sterile BD Falcon™ centrifuge tube and centrifuged at 300 x g for 5 min. Supernatant was discarded and the cell pellet was resuspended in 5ml RF10 medium to give a cell suspension at 1 x 10^5 cells/ml. A 10-fold serial dilution was performed using RF10 medium to derive cell suspensions at 1 x 10^2 cells/ml and 1 x 10^1 cells/ml (in a volume of 30ml). For each cell density a 96-well culture plate was prepared and incubated at 37°C/5% CO_2 for approximately 20 days.

Chemotherapeutic Agent	Molecular Weight	Solvent	Stock Solution (mM)
Daunorubicin hydrochloride	563.98	Sterile ddH_2O	5
Cytarabine hydrochloride (*1-β-D-ararbinofuranosylcytosine*)	279.68	Sterile ddH_2O	1

Table 2.4. Preparation of stock solutions of chemotherapeutic agents.

Daunorubicin dose / Cell density	Vehicle-treated (0nM)	2nM	4nM	6nM	8nM	10nM
$1x10^1$ cells/ml (2 cells/well)	●	●	●			
$1x10^2$ cells/ml (20 cells/well)		●	●	●	●	
$1x10^3$ cells/ml (200 cells/well)			●	●	●	●

Figure 2.1. The design of the 96-well plating for clonogenic assay.

Following a 20 days growth period, the survival fraction at each dose was determined according to Section 2.5.2. Each experiment was performed on three independent occasions. A dose-response curve was generated using mean survival fractions and SEM of the data from three independent experiments (using GraphPad Prism 6 software).

2.6. *In Vitro* Mutation Assay

In vitro mutation assays were used to evaluate the effect of chemotherapeutic agents and RUNX1/ETO status on MF in TK6 cells using thymidine kinase (*TK*) and hypoxanthine-guanine phosphoribosyltransferase (*HPRT*) as reporter genes (Chapter 3). The workflow for setup of the assays is summarised in Figure 2.2 and specific details of each step are described in Sections 2.6.1, 2.6.2, 2.6.3 and 2.6.4.

2.6.1. Removal of spontaneous mutants using CHAT medium

Prior to performing mutation assays, pre-existing spontaneous *TK* mutants ($TK^{+/+}$/$TK^{+/-}$) and *HPRT* mutants ($HPRT^{+/-}$) in a cell population were purged by culture in CHAT (cytidine, hypoxanthine, aminopterin, thymidine) medium. Inhibition of purine nucleotide synthesis by aminopterin requires cells to rely on exogenous DNA precursors in the medium (hypoxanthine and thymidine) to evade DNA synthesis blockage (salvage pathways). TK and HPRT mutants unable to survive due to an inability to incorporate DNA precursors from the culture medium.

For each cell line, 5 x 10^5 cells were centrifuged at 300 x g for 5 min, resuspended in 25ml CHAT medium [RF10 medium supplemented with 10μM 2-deoxycytidine, 17.5μM thymidine, 200μM hypoxanthine and 0.2μM aminopterin] and transferred to a 75cm^2 cell culture flask. Cultures were incubated at 37°C/5% CO_2 for 48 hours.

Following 48 hours incubation, cells were centrifuged at 300 x g for 5 min and the supernatant discarded. The cells were washed twice with 3ml sterile PBS and resuspended in 25ml THC medium [the same as CHAT medium but without aminopterin] and transferred to a fresh 75cm^2 culture flask. Cultures in THC medium were incubated at 37°C/5% CO_2 until cell populations recovered as indicated by exponential growth. The cultures were maintained daily by gradually replacing THC

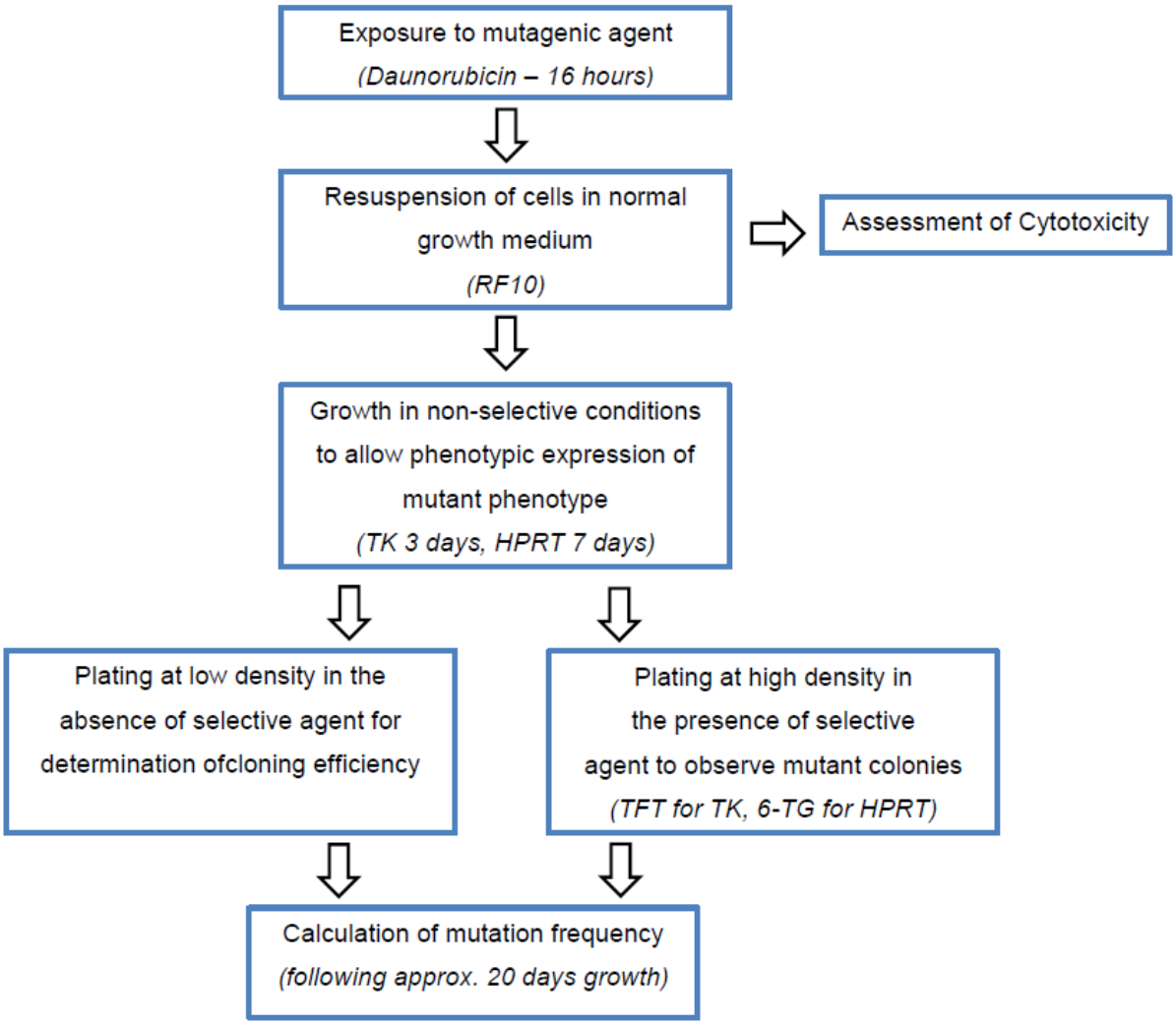


Figure 2.2. General workflow of *in vitro* mutation assay (*TK* and *HPRT*).

medium with normal RF10 medium. Mutation assays were commenced immediately once normal exponential cell growth had been achieved.

2.6.2. Drug exposure

A 50ml aliquot of cell suspension at a density of 1 x 10^5 cells/ml for each cell line (TK6 and TK6 RE8) was prepared using RF10 medium. The cell suspension was transferred to a sterile 50ml BD Falcon™ centrifuge tube and centrifuged at 300 x g for 5 min. Supernatant was discarded and the cell pellet resuspended in 50ml RF10 medium supplemented with an appropriate volume of drug or solvent and transferred to 175cm^2 cell culture flasks. Each cell line was divided into four 50ml cultures in 175 cm^2 flasks for drug dosing. Three flasks from each cell line were designated as drug-dosed flasks and one flask as a solvent control flask (vehicle-treated control). Cultures were incubated at 37°C/5% CO_2 for 16 hours and each experiment was performed on three independent occasions.

Following a 16 hour incubation, the dosed cell suspension was transferred to a sterile 50ml BD Falcon™ centrifuge tube and centrifuged at 300 x g for 5 min. Supernatant was discarded and cells were washed twice using sterile PBS. An aliquot of 5 x 10^5 cells from each dosed cell suspension was transferred to a sterile 50ml BD Falcon™ tube for cytotoxicity assessment as described in Section 2.5.3. The remaining cell suspensions were transferred to fresh 175cm^2 culture flasks and incubated at 37°C/5% CO_2. Cell suspensions were maintained daily at a density between 2 x 10^5 and 1 x 10^6 cells/ml (in a volume of 50ml) using RF10 medium.

2.6.3. Selection of TK mutants

TK mutants were selected based on resistance to trifluorothymidine (TFT) after 3 days in culture; the time required for expression of the *TK* mutant phenotype (Liber and Thilly, 1982). A 96-well non-treated plate was prepared to enable calculation of CE in the absence of selection and six 96-well TFT-treated plates were prepared for *TK* mutant selection.

For each drug dosed and vehicle control cell suspension, an accurate cell count was performed and used to establish a cell suspension at a density of 1 x 10^5 cells/ml (in a volume of 155ml) using RF10 medium. The remaining cell suspension was returned to the incubator for subsequent *HPRT* mutant selection. In order to determine

cloning efficiency in the absence of TFT selection, a 5ml aliquot from this suspension was removed and a 10-fold serial dilution was performed using RF10 medium to derive a cell suspension at 1 x 10^1 cells/ml (in a volume of 30ml). The cell suspension was then transferred to a 96-well culture plate (200µl/well with average 2 cells/well). For the TFT-treated plates, the remainder of the cell suspension (150ml) was supplemented with 37.5µl of a 4mg/ml TFT solution (final concentration 1µg/ml) and transferred to six 96-well culture plates (200µl/well with average 20,000 cells/well). All plates were incubated at 37°C/5% CO2 for approximately 20 days.

2.6.4. Selection of HPRT mutants

HPRT mutants were selected based on resistance to 6-thioguanine (6-TG) after 7 days in culture; the time required for expression of the *HPRT* mutant phenotype (Liber and Thilly, 1982). Similar to the *TK* mutant selection protocol, a 96-well plate was prepared to enable calculation of CE in the absence of 6-TG selection and six 96-well 6-TG treated plates were prepared for *HPRT* mutants selection.

For each drug dosed and vehicle control cell suspension, an accurate cell count was performed and used to establish a cell suspension at density of 1 x 10^5 cells/ml (in a volume of 155ml) using RF10 medium. In order to determine cloning efficiency in the absence of 6-TG selection, a 5ml aliquot from this suspension was removed and a 10-fold serial dilution was performed using RF10 medium to derive a cell suspension at 1 x 10^1 cells/ml (in a volume of 30ml). The cell suspension was then transferred to a 96-well culture plate (200µl/well with average 2 cells/well). For the 6-TG-treated plates, the remainder of the cell suspension (150ml) was supplemented with 120µl of a 2.5mg/ml 6-TG solution (final concentration 2µg/ml) and transferred to six 96-well culture plates (200µl/well with average 20,000 cells/well). All plates were incubated at 37°C/5% CO_2 for approximately 20 days.

2.6.5. Calculation of mutation frequency

Following a 20 day growth period for mutant colonies, the number of positive wells (wells in which mutant colonies had grown) was scored per plate from the six TFT- and 6-TG-treated plates for each dose. For calculation of CE in the absence of selection, the number of negative wells (wells in which no cell growth had occurred) was scored and the CE was determined according to the formula in Section 2.5.2.

MF at each dose was calculated according to Liber and Thilly (Liber and Thilly, 1982) using the following formula:

$$Mutation\ frequency\ (MF) = \frac{Total\ no.\ of\ mutant\ colonies}{CE\ x\ Total\ no.\ of\ cellls\ plated}$$

All calculations were performed using Microsoft Excel. The survival fraction at each dose was determined as described in Section 2.5.2 and each experiment was repeated three times. The cytotoxicity-MF exponential curve was generated using mean and SEM values from three independent experiments (using GraphPad Prism 6 software).

2.6.6. Statistical analysis

One-way and two-way analysis of variance (ANOVA) was used to assess the effects of drug dose and RUNX1/ETO status on MF. A p-value of < 0.05 was considered statistically significant in all cases. All statistical analyses were performed using GraphPad Prism 6 software (San Diego, USA).

2.7. Molecular Analysis of *HPRT* Mutants

To further investigate the mutagenicity of RUNX1/ETO, both spontaneously and after chemotherapy treatment (Chapter 3), a number of spontaneous and daunorubicin-induced *HPRT* mutants were subject to molecular analysis in order to generate a daunorubicin-induced and RUNX1/ETO-induced mutational spectrum. *HPRT* cDNA was used for the mutation screening as it can specifically reveal mutations in the *HPRT* coding region which might give rise to dysfunctional protein (missense and nonsense mutations). In addition, mutation screening using cDNA can be achieved via a single amplicon. In contrast, mutation screening using genomic DNA would require amplification of each exon independently or the use of multiplex PCR. Both of these approaches would require optimisation and would be time consuming and laborious compared to the PCR of the *HPRT* cDNA.

2.7.1. Generation of spontaneous and daunorubicin-induced mutant cell populations

A small batch of CHAT-treated cell population was performed every time to minimise the number of duplicate mutations derived from the same parental flask which likely represent clonal expansions. The CHAT-treated parental population was divided into twenty independent populations in 25cm^2 cell culture flasks (5ml/flask). A summary of the workflow used to generate spontaneous and daunorubicin-induced *HPRT* mutants is shown in Figure 2.3.

For the assessment of spontaneous mutants, the cell suspension were passaged every 2 days for 21 days to allow for the acquisition of spontaneous mutations. Following this acquisition period, the cell suspension from each flask was transferred to a 96-well culture plate under 6-TG selection to select for spontaneous *HPRT* mutant colonies. On the other hand, the cells were exposed to the highest dose of daunorubicin for 16 hours and transferred to a 96-well plate under selection of 6-TG to generate daunorubicin-induced *HPRT* mutant colonies.

Following 20 days of growth, one mutant colony from each plate was randomly selected and transferred to a well of a sterile 24-well cell culture plate containing 1ml RF10 medium supplemented with 2µg/ml 6-TG. Plates were incubated at 37°C/5% CO_2 and assessed daily for cell growth (indicated by a change in colour of culture medium and confirmed microscopically). Once sufficient cell numbers were generated (after approximately 7 days in culture), as estimated microscopically, each population was washed and cell pellets were prepared according to Section 2.3.5 (without performing cell counts). The cell pellets were stored at -80°C until required for DNA/RNA extraction (Section 2.7.2).

2.7.2. DNA/RNA extraction and quantitation

Frozen cell pellets of *HPRT* mutants were thawed and resuspended in 350µl RLT Plus buffer. DNA and RNA extraction was performed using the AllPrep DNA/RNA Mini Kit (Qiagen) according to the manufacturer's protocol. Briefly, DNA and RNA extraction was achieved through binding of DNA/RNA to a QIAamp® silica gel membrane (in spin-column format) after which contaminants were removed by washing using provided buffers. The DNA and RNA was eluted in respective elution

buffer (DNA: 50µl Buffer EB; RNA: 30µl RNase-free water). The eluted DNA and RNA was stored at -20°C and -80°C respectively until required.

Quantitation of RNA was performed using a NanoDrop® ND-1000 spectrophotometer (Thermo Scientific, DE, USA) which measures the absorbance of UV light at 260nm passing through a 1.5µl aliquot of extracted RNA, and performs the necessary calculations according to the Beer Lambert Law to provide an estimate of RNA concentration (in ng/µl). Following quantitation, an appropriate dilution was performed using nuclease-free water to obtain 15µl of RNA at a concentration of 200ng/µl for *HPRT* cDNA preparation.

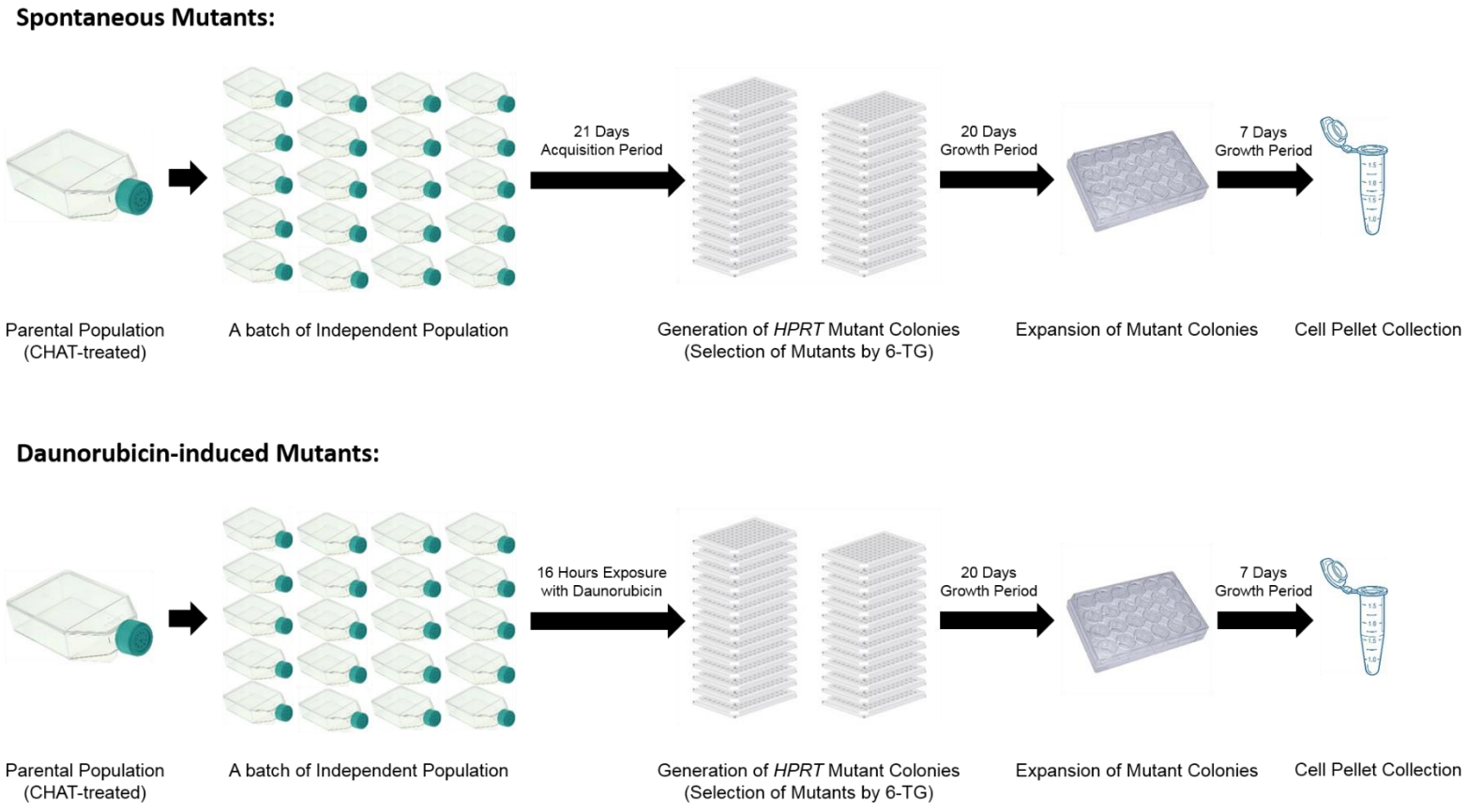


Figure 2.3. Summarised workflow for the generation of spontaneous and daunorubicin-induced *HPRT* mutants.

2.7.3. PCR of HPRT cDNA

2.7.3.1. Reverse transcription of RNA into cDNA

HPRT cDNA was prepared using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Warrington, UK), which uses the random primer method for initiating reverse transcriptase (RT)-mediated cDNA synthesis.

An appropriate volume of RT mastermix was prepared according to the manufacturer's protocol as shown in Table 2.5 and 10µl of the mastermix was dispensed into an appropriate number of sterile 0.2ml PCR tubes (Eppendorf). For each sample, 10µl of 200ng/µl total RNA was added. A blank reaction (negative control) was also prepared using nuclease-free dH_2O in place of RNA, to ensure no contamination of reagents. The following single cycle was performed using a GeneAmp® PCR System 9700 (Applied Biosystems):

Step 1:	25°C	10 min
Step 2:	37°C	120 min
Step 3:	85°C	5 sec
Step 4:	4°C	Hold

1 cycle.

Hypothetically, this protocol will yield 20µl of cDNA at a concentration of 100ng/µl by assuming 100% efficiency of RT reaction.

2.7.3.2. Amplification of HPRT cDNA

PCR of *HPRT* cDNA was performed using a set of primer that amplify the exonic sequence of *HPRT* (Meng *et al.*, 2002), as follows:

Forward: $^{5'}$CCTGAGCAGTCAGCCCGCGC$^{3'}$
Reverse: $^{5'}$CAATAGGACTCCAGATGTTT$^{3'}$

Primers were manufactured by Sigma-Aldrich, reconstituted to 100pmol/µl using TE buffer upon receipt and stored at -20°C. Working stocks of primers at a concentration of 10pmol/µl were prepared using nuclease-free dH_20 and stored at 4°C.

2X ReddyMix PCR Master Mix containing 0.625 u Taq DNA Polymerase, 75mM Tris-HCL, 20mM $(NH_4)_2SO_4$, 1.5mM $MgCl_2$ and 0.2mM each dNTP was obtained from ThermoFisher Scientific. This pre-mix of PCR components enables quick PCR preparation and reduces risk of contamination due to a reduced number of pipetting steps required for PCR set-up.

An appropriate volume of PCR master mix with primers was prepared according to Table 2.6 for each PCR reaction and 15µl of the master mix was dispensed into an appropriate number of sterile 0.2ml PCR tubes (Eppendorf). For each sample, 2µl of cDNA (200ng) was added and made up to a total volume of 25µl using nuclease-free dH_2O. A blank reaction (negative control) was also prepared using nuclease-free dH_2O in place of cDNA to ensure no contamination of reagents. Thermal cycle conditions were as follows:

Initial denaturation	95°C	2 min	
Denaturation	95°C	25 sec	36 cycles
Annealing	55°C	3 sec	
Extension	72°C	65 sec	
Final extension	72°C	5 min	

2.7.4. Agarose gel electrophoresis

A 1% agarose solution was prepared by melting 1g of Ultrapure™ agarose in 100ml 1X TBE buffer using a 650W microwave on medium power. Following cooling to 'hand heat', 10µl of GelGreen Nucleic Acid Stain (10,000X in water) was added. Gel solutions were mixed thoroughly, poured into assembled large gel cassettes with combs and allowed to set at room temperature.

Gel cassettes were placed in Sub-Cell® GT Agarose Gel Electrophoresis Systems filled with 1X TBE buffer. To allow estimation of the size of PCR products, 10µl of Quick-Load® 100bp DNA Ladder was loaded into the first well of each gel. For each reaction, 5µl of PCR product was loaded into each well of the gel. A negative and positive control was loaded into the second and third well, respectively. Amplicons derived from *HPRT* mutant colonies were loaded into the remaining wells and electrophoresis was performed at 100V for approximately 1 hour.

Following electrophoresis, the gel was removed from the cassette and placed in a Gel Doc™ XR (Bio-Rad Laboratories Ltd.) system and the Quantity One® V4.5.0

Reagent	Volume per reaction (µl)	Final concentration
RT Buffer (10X)	2.0	2X
dNTP mix (100mM)	0.8	8mM
RT Random Primers (10X)	2.0	2X
Multiscribe Reverse Transcriptase (50U/µl)	1.0	5U/µl
RNase inhibitor (20U/µl)	1.0	2U/µl
Nuclease-free dH_2O	3.2	
Total volume	**10.0**	

Table 2.5. Preparation of reverse-transcription master mix.

Component	Volume per reaction (µl)	Final concentration
2X ReddyMix PCR Master Mix	12.5	1X
Forward primer (10µM)	1.25	0.5µM
Reverse primer (10µM)	1.25	0.5µM
Template cDNA*	2.0	200ng
Nuclease-free dH_2O*	8.0	
Total volume	**25.0**	

Table 2.6. Preparation of PCR master mix.

*Separately added into each PCR tube containing 15µl of PCR master mix with primers.

software was used to capture an image of the gel by using UV light for visualisation of PCR products. The presence of a DNA band in the positive control and absence of a DNA band in the negative control indicated successful PCR. Amplicons derived from *HPRT* mutants were assessed by size comparison with the positive control sample (as estimated using the DNA ladder) and PCR reactions where there was an absence of a DNA band were not further investigated for Sanger sequencing.

2.7.5. Clean-up of PCR products for Sanger sequencing

Prior to Sanger sequencing, clean-up of PCR products (mutants with DNA band) was performed using the QIAquick® PCR Purification kit (Qiagen) according to the manufacturer's protocol. Briefly, buffer PBI was mixed with PCR products in a 5:1 ratio, followed by the addition of 10µl 3M sodium acetate (pH 5) to ensure correct pH of the mixture (indicated by the yellow colour of the mixture). The mixture was then transferred to a spin column, centrifuged at 15,000 x g for 1 min and the flow-through was discarded. A wash step was performed by adding 750µl of Buffer PE to the spin column and centrifuged at 15,000 x g for 1 min. The flow-through was discarded and the spin column was again centrifuged for an additional 1 min to dry the spin column membrane. The spin column was then transferred to a sterile 1.5ml Eppendorf tube and 30µl of Buffer EB was added directly to the centre of the spin column membrane to elute the DNA. The spin column was incubated at room temperature for 5 min, followed by centrifugation at 15,000 x g for 1 min for the collection of purified PCR products.

Quantitation of purified PCR products was performed as described in Section 2.7.2 and the concentration was adjusted to 10ng/µl using nuclease-free dH_2O. The purified PCR products were Sanger sequenced by Source BioScience LifeSciences (Nottingham, UK) using appropriate primers (section 2.7.3.2).

2.7.6. Sequence analysis

Sanger sequence data was received in FASTA format, viewed using EditSeq (DNASTAR Inc., WI, USA) and compared to human *HPRT* exon sequences (NCBI Reference Sequence: NM_000194) using Multiple Sequence Alignment tools (https://www.ebi.ac.uk/Tools/msa/muscle). Both forward and reverse sequences were analysed to ensure accurate coverage of entire exons. In addition, sequence traces

(chromatogram format) were viewed using FinchTV (Geospiza Inc., WA, USA) in order to assess quality of sequence reads.

2.7.7. Classification of spontaneous and daunorubicin-induced mutation

HPRT mutants were grouped into five classes; partial gene duplications, base substitutions, small deletion/insertions, internal exon deletions and terminal deletions (5' or 3' terminal deletions). The number of mutants from each class was used to estimate the proportion of each class as a fraction of the total *HPRT* mutants analysed and mutation spectra were generated using Microsoft Excel.

2.7.8. Statistical analysis

A Chi-squared test was performed to determine if there were any significant difference between the spontaneous and daunorubicin-induced *HPRT* mutational spectra, and also to determine whether RUNX1/ETO status significant affected mutation induction. The Fisher's exact test was used to determine whether the frequency of specific classes of mutants differed significantly between spontaneous and daunorubicin-induced and also between RUNX1/ETO-negative and RUNX1/ETO-positive cells, both spontaneously and after exposure to daunorubicin. A p-value of < 0.05 was considered statistically significant in all cases. All statistical analyses were performed using GraphPad Prism 6 software.

2.7.9. Interrogation of base substitution spectrum

Base substitution spectrum was further investigated as a function of local sequence context to identify a mutational signature of base substitutions induced by RUNX1/ETO (Chapter 4). Mutations at the central base in each possible trinucleotide sequence (32 combinations) were examined to identify sequences over-represented in RUNX1/ETO transduced cells compared to cells negative for RUNX1/ETO expression. In addition, gene specific sequencing data from 56 diagnostic AML t(8;21) cases (Hartmann *et al.*, 2016) were also interrogated to provide an *in vivo* dataset for the base substitution pattern in RUNX1/ETO-positive AML. Collectively, these data were used to determine whether RUNX1/ETO preferentially induces mutation at specific sequences in the genome. For this analysis genes interrogated include

ZBTB7A, *ASXL2*, *NRAS*, *KIT*, *FLT3*, *ASXL1*, *RAD21*, *TET2*, *SMC3*, *DNMT3A*, *JAK2*, *CBL*, *FAT1*, *SMC1A*, *WT1*, *KRAS*, *PTPTN11*, *IDH2*, *GATA2* and *JAK3*.

Heatmaps displaying the number of base substitution mutations at the central position in each trinucleotide sequence were generated using Microsoft Excel. The Fisher's exact test was used to determine whether the frequency of specific classes of mutation differed significantly between RUNX1/ETO-negative and RUNX1/ETO-positive cells. A p-value of < 0.05 was considered statistically significant in all cases. All statistical analyses were performed using GraphPad Prism 6 software.

2.8. Flow Cytometry

Flow cytometry was used to evaluate the effect of RUNX1/ETO expression on MF in AML relevant genes using mutant-specific antibody (Chapter 5). Parental TK6 and its derivative RUNX1/ETO-positive cell clones (TK6 RE1 and TK6 RE8) were used as a model to determine whether the RUNX1/ETO fusion oncoprotein predisposes cells to mutation in genes relevant to leukaemogenesis. All antibodies used in flow cytometry experiments and their details are shown in Table 2.7. Preparation of all additional reagents for use in specific experiments are detailed in sections 2.8.2 and 2.8.3.

2.8.1. Generation of single clonal population of TK6 and TK6 RUNX1/ETO

TK6 have a high single cell cloning efficiency in liquid culture of approximately 70% when plated in a 96-well culture plate. This allows for the generation of single subclones with a range of RUNX1/ETO expression.

For each clone, an accurate cell count using was performed to establish a cell suspension at a concentration of 1×10^5 cells/ml using RF10 medium. A 10-fold serial dilution was carried out using RF10 medium to give a cell concentration of 10 cells/ml, followed by a1 in 2 dilution to give a final cell concentration of 5 cells/ml. 200µl of suspension was transferred to each well of sterile 96-well culture plate (average 1 cell/well) using a multichannel pipette. Ten plates were prepared for each clone and incubated at 37°C/5% CO_2. Following 20 days growth period, plates were checked under microscope for colony formation and any which had visible colonies were transferred to 800µl fresh RF-10 media in a 48 well plate for expansion, and finally transferred to a $25cm^3$ sterile cell culture flasks prior to flow cytometry.

2.8.2. Cell preparation, fixing and permeabilisation

For each cell clone, an accurate cell count was performed to get 3 x 10^6 cells. The appropriate volume of cell suspension was dispensed into a sterile 15ml BD Falcon™ centrifuge tube, centrifuged at 400 x g for 5 min and supernatant was discarded. Cells were washed twice with 1ml of 1% BSA/PBS by centrifugation at 400 x g for 5 min and the supernatant was discarded each time. A hundred microliter of ice cold absolute ethanol was added to the cells, gently mixed and incubated for 15 min on ice. Cells were then washed twice with 1ml of 1% BSA/PBS. A hundred microliter of 1X FACS permeabilising solution II (0.5 mL; Becton Dickinson, USA) was added to the cell suspension, vortexed and incubated for 10 min on ice. Cells were subsequently washed twice with 1ml of 1% BSA/PBS and cell pellets were collected by centrifugation at 400 x g for 5 min.

2.8.3. Primary and secondary antibodies staining

Primary and secondary antibodies were prepared from stocks by dilution with 3% BSA/PBS (Table 2.7). A hundred microliters of NPM1 mutant-specific antibody (final concentration 1µg/ml) and rabbit IgG isotype control (final concentration 2.5µg/ml) was added to the appropriate cells, vortexed and incubated at room temperature for 1 hour. Cells were then washed twice with 1ml of 1% BSA/PBS by centrifugation at 400 x g for 5 min and the supernatant were discarded each time. A hundred microliters of conjugated secondary goat anti-rabbit antibody (final concentration 2µg/ml) was added to the cells, vortexed and incubated at room temperature for 30 min in dark. Samples were then washed twice with 1ml of 1% BSA/PBS and the cell pellets were resuspended in 500µl of ice cold 3% BSA/PBS.

Antibody	Type	Supplier	Host / Isotype	Concentration	Conjugate	Dilution
NPM1 (mutant) Polyclonal	Primary	Invitrogen #PA1-46356	Rabbit / IgG	1 mg/ml	Unconjugated	1 : 100
Rabbit IgG Isotype Control	Primary	Invitrogen #02-6102	Rabbit / IgG	5 mg/ml	Unconjugated	1 : 200
Goat anti-Rabbit IgG (H+L) Cross-Adsorbed Polyclonal	Secondary	Invitrogen #31212	Goat / IgG	2 mg/ml	Alexa Fluor 488 conjugated	1 : 100

Table 2.7. Antibodies used in flow cytometry.

2.8.4. Detection of mutant cells

The labelled samples were analysed on a FACSCalibur flow cytometer (Becton Dickinson, USA) using Cell Quest™ Pro (Becton Dickinson, USA) and FCS Express 6 Flow cytometry softwares (De Novo, California, USA). The target cell populations were gated based on the forward and side scatter signal intensity to exclude the non-target populations including dead cells and cell debris. Positive control (AML-3) and isotype control (Rabbit IgG Isotype) samples were simultaneously analysed alongside parental TK6 and RUNX1/ETO-transduced cells (TK6 RE1 and TK6 RE8) to optimise instrument settings (adjust voltages of fluorescence detectors to remove background fluorescence and adjust compensation settings for overlap of fluorescent spectra). Following optimisation, NPM1 mutant positivity was determined by analysing the Alexa Fluor 488 fluorescence of the gated population in a histogram format with a minimum of 10,000 events being acquired in this gate. A fluorescence intensity value between 10^1 and 10^2 was defined as "positive" by overlapping the fluorescence spectra of the samples simultaneously to exclude the negative events and non-specific binding of the secondary antibody (Chapter 5, Section 5.2.2). Based on the acquired criteria for positivity, the validity of the assay was determined using flow sorting and Sanger sequence analysis of positive events (Section 2.8.5, 2.8.6 and 2.8.7).

2.8.5. Cell sorting of NPM1 mutant population

The flow cytometry assay was investigated to confirm that positive events were cells mutated for *NPM1* exon 12. The target population was sorted using an ASTRIOS EQ Cell Sorter (Beckman Coulter, London, UK) in the Flow Cytometry Core Facility (FCCF) (School of Medical Sciences, Newcastle University). Fifty million of labelled TK6 RE8 clone (clone with high expression of RUNX1/ETO) and mixed population of 5% OCI-AML3/95% TK6 RE8 cells were prepared according to the sample preparation protocol (Sections 2.8.2 and 2.8.3) and diluted in 3% BSA/PBS to give a final cell concentration of 2 x 10^7 cells/ml as recommended by the FCCF. Sorted cells were collected into 100 µl of 3% BSA/PBS in a 1.5 ml Eppendorf tube and stored at -20°C until required.

2.8.6. Whole genome amplification of the sorted cells

Genomic DNA of sorted cells was extracted using a REPLI-g® Single Cell kit (Qiagen) according to the manufacturer's protocol. This kit utilises a long-range DNA polymerase capable of amplifying up to 100 kilobases (kb) from small number of cells (1 to <1000 bacterial or tumour cells) at a constant temperature (30°C) without dissociating from the genomic template.

Quantification of DNA was performed as described in Section 2.7.2 and the concentration was adjusted to 100ng/µl using nuclease-free dH_2O for use in PCR reactions. Remaining DNA was stored at -20°C until required.

2.8.7. Amplification of NPM1 exon 12 for sequencing of NPM1 mutation

PCR of *NPM1* DNA was performed using primers that target exon 12 of *NPM1* gene (NCBI Reference Sequence: NG_016018.1) as follows:

Forward: $^{5'}$AACTCTCTGGTGGTAGAATGAAA$^{3'}$
Reverse: $^{5'}$TGAGAACTTTCCCTACCGTGT$^{3'}$

Primers were manufactured by Sigma-Aldrich, reconstituted to 100 pmol/µl using TE buffer upon receipt and stored at -20°C. Working stocks of primers at a concentration of 10 pmol/µl were prepared using nuclease-free dH_20 and stored at 4°C.

Reaction conditions and setup for PCR amplification was prepared according to Section 2.7.3.2. A control was also prepared using nuclease-free dH_2O in place of DNA, to ensure no contamination of reagents. Thermal cycling conditions were as follows:

Initial denaturation	95°C	2 min	
Denaturation	95°C	25 sec	36 cycles
Annealing	55°C	3 sec	
Extension	72°C	65 sec	
Final extension	72°C	5 min	

To confirm successful amplification of *NPM1* exon 12 DNA, an aliquot of each reaction (5µl) was assessed using agarose gel electrophoresis according to Section 2.7.4. Remaining PCR product was prepared and Sanger sequenced by Source BioScience LifeSciences (Nottingham, UK) (Section 2.7.5). Sequence analysis was performed as described in Section 2.7.6.

2.8.8. Calculation of mutation frequency

Following confirmation of NPM1 mutant positivity, subsequent experiments were performed every 2 weeks to measure the *NPM1* MF in RUNX1/ETO-positive subclones (TK6 RE1 and TK6 RE8) relative to parental control TK6 subclones. The frequency of *NPM1* exon 12 positive cells in each cell population was calculated according to the following formula:

$$Mutation\ frequency\ (MF) = \frac{No.of\ positive\ event}{No.of\ gated\ event}$$

All calculations were performed using Microsoft Excel. The individual value, mean and SEM of ten replicates were plotted on X-Y scatter graph (using GraphPad Prism 6 software). One-way ANOVA and paired *t*-test were used to assess the effects of RUNX1/ETO status on MF. A p-value of < 0.05 was considered statistically significant in all cases. All statistical analyses were performed using GraphPad Prism 6 software (San Diego, USA).

2.9. Whole Exome Sequencing

Whole exome sequencing (WES) was performed on matched presentation and relapsed AML samples in order to investigate clonal evolution of relapsed AML (Chapter 6). Exome sequencing provides high throughput sequencing of the protein-coding genes in a genome (~23 000 genes) which constitute about 2% of the human genome, by using biotinylated oligonucleotide probes to selectively hybridize to target regions. This allow identification of genetic variants across the region that contains ~85% of known disease-related variants (van Dijk *et al.*, 2014) and can be used to compare genetic variants in presentation and relapse samples for the identification of relapse-specific mutations responsible for AML disease progression.

2.9.1. Recruitment of patients

AML samples were collected within the National Referral Centre for Bone Marrow Cytogenetics and Molecular Genetics, Institute for Medical Research (IMR, Kuala Lumpur, Malaysia) who are collaborating on this study. Ethical approval had been acquired form Medical Research and Ethics Committee (MREC), Ministry of Health, Malaysia. Written informed consent was obtained from each patient for the use of their DNA samples in this study as well as for any scientific publication derived from this study. A cohort of 11 matched presentation and relapsed AML cases were identified. Patients achieved clinical remission after treatment with combination chemotherapy that included ara-C and an anthracycline (daunorubicin or idarubicin), followed by consolidation chemotherapy that included high-dose of ara-C. The median age was 33 years (range: 10 – 51 years). Detailed information of the recruited patients will be further described in Chapter 6.

2.9.2. Preparation of DNA from primary AML samples

All DNA extraction and preparation were performed by IMR. RNA from all primary patient samples was extracted using QIAamp DNA Blood Mini Kit (Qiagen) according to manufacturer's protocol. Quantitation of the extracted DNA was carried out using a Nanodrop ND-1000 spectrophotometer and the required concentration and volume of RNA for exome sequencing was prepared according to sequencing service provider instruction.

2.9.3. Exome sequencing

Sample preparation, library construction, sequencing and bioinformatics analysis were performed offsite by the Beijing Genome Institute (BGI) (Hong Kong, China). SureSelect Human All Exon V5 + UTR kit (Agilent Technologies, Santa Clara, USA) was used for enrichment of exon sequence including untranslated regions (UTR). Samples were sequenced to give an average read depth (reads per base) of 100X, supporting the identification of minor leukaemic cell clones. Paired-end sequencing was performed using Illumina HiSeq 4000 platform and analysis of raw data was processed according to BGI analysis pipeline to map reads against the hg19

build of the human genome and identify single nucleotide variants (SNVs) and insertion/deletions (indels).

2.9.4. Data analysis

Variant calling files (.VCF) were received from the BGI. Analysis of genetic variants was performed using Ingenuity Variant Analysis™ software (Qiagen Bioinformatics) to compare the somatic mutation profile of matched AML samples at presentation and relapse in order to identify somatic mutations that drive disease progression.

A series of filters were applied to select variants most likely to impact on gene function or biological processes known to be implicated in disease progression. Variants with a read depth < 10 were excluded from the analysis as were variants with an allele frequency (VAF) ≥ 5% in the healthy population [databases used include the 1000 Genomes Project, ExAC, gnomAD, NHLBI ESP exomes; minor allele frequency (MAF) < 1%)] .

Identification of putative AML driver mutations were determined by identification of variants in established AML genes. Specifically, the Catalogue of Somatic Mutations in Cancer (COSMIC) database was interrogated to identify the most frequently recurrently mutated genes in AML (Table 2.8).

No.	Gene	Mutated samples	Samples tested	Frequency (%)
1	NPM1	6209	19236	32.28
2	FLT3	16141	68481	23.57
3	DNMT3A	2525	13085	19.3
4	NRAS	1157	8594	13.46
5	TET2	928	7274	12.76
6	RET	104	1042	9.98
7	CEBPA	1351	13823	9.77
8	IDH2	1489	15414	9.66
9	TLE4	98	1041	9.41
10	ATM	102	1116	9.14
11	WT1	563	6347	8.87
12	TP53	382	4454	8.58
13	MAP2K2	88	1042	8.45
14	ASXL1	671	7965	8.42
15	KIT	789	9601	8.22
16	PBRM1	82	1042	7.87
17	RUNX1	685	8806	7.78
18	TEK	77	1041	7.4
19	IDH1	1139	17340	6.57
20	SRSF2	337	5149	6.54
21	DNM2	65	1041	6.24
22	PTPN11	379	6248	6.07
23	NF1	149	2583	5.77
24	GATA2	284	5150	5.51
25	BCOR	243	4464	5.44
26	KRAS	343	6319	5.43
27	MUC12	51	947	5.39
28	GATA1	176	3614	4.87
29	KMT2D	120	2534	4.74
30	RAD21	188	4142	4.54

Table 2.8. Top 30 genes frequently mutated in AML (https://cancer.sanger.ac.uk/cosmic).

Chapter 3. Evaluating the Mutagenicity of Chemotherapy Agents and RUNX1/ETO Expression in a Cell Model System

3.1. Introduction

Cancer is the ultimate genetic disease characterised by the accumulation of somatic mutations within the genome of a somatic cell acquired during disease progression. Although somatic mutations are known to cause cancer, the biological processes generating these mutations are poorly understood. In normal circumstances somatic mutations can be acquired spontaneously due to replication errors and defective DNA repair. However, the acquisition of somatic mutations can be influenced by additional sources of DNA damage including exposure to exogenous agents such as UV radiation (Brash *et al.*, 1991; Pierceall William *et al.*, 1991; Weihrauch *et al.*, 2002; Brash, 2015) and endogenous cellular processes such as energy metabolism and lipid peroxidation, which generate reactive genotoxic chemicals (e.g. ROS) (Pluskota-Karwatka, 2008). These DNA damaging agents have the potential to be mutagenic to DNA and cause genomic instability if the damage is not detected and repaired, or if apoptosis is not initiated when repair cannot successfully occur.

Despite considerable efficacy in the treatment of cancer, many chemotherapy agents are also mutagenic and some have already been shown to induce mutations detectable after therapy (Hunter *et al.*, 2006). Ara-C in combination with an anthracycline such as daunorubicin are the two most widely used chemotherapeutic agents in AML remission induction treatment. Ara-C is a nucleoside analogue incorporated into replicating DNA at internucleotide positions leading to inhibition of chain extension (Chapter 1, Section 1.7.1). The mutagenicity of ara-C was previously described although the effect was relatively weak compared to the mutagenic potency of N-methyl-N-nitrosourea (MNU) (Fordham *et al.*, 2015). The cytotoxicity of daunorubicin is primarily mediated by the inhibition of topo II, preventing the DNA double helix from being re-ligated, thereby stopping the process of replication (see Chapter 1, Section 1.7.2), which may eventually lead to elevated DNA strand breaks. As well as inducing strand breaks via the inhibition of DNA topo II, daunorubicin is a redox cycling agent which can form a complex with iron to generate ROS (Davies and Doroshow, 1986; Juchau *et al.*, 1986). This process leads to oxidative stress and causes DNA base damage such as 8-oxoG lesions, which can result in base mispairing with adenine and generation of G:C to T:A substitution mutations in the genome (Cheng *et al.*, 1992). Indeed, evidence has showed that daunorubicin and related anthracyclines can induce loss of heterozygosity (LOH) or genetic rearrangements

such as deletion and chromosomal translocation, demonstrating the mutagenicity of this group of chemotherapeutic agent (Anderson *et al.*, 1991; Lehmann *et al.*, 2004).

In addition, genomic instability caused by fusion genes may also contribute to increase somatic mutations in AML. Evidence suggests that initiating lesions such as *RUNX1/ETO* may promote mutagenesis (Krejci *et al.*, 2008; Araten *et al.*, 2013) and ectopic expression of RUNX1/ETO has been observed to cause downregulation of several DNA repair proteins (*BRCA2*, *OGG1* and *ATM*) and failure to repair DNA damage (Alcalay *et al.*, 2003). A recent study has shown that RUNX1/ETO increases genomic instability and mutagenesis by predisposing cells to the acquisition of mutations, both spontaneously and particularly after treatment with genotoxic agents (Forster *et al.*, 2015). *RUNX1/ETO* is the most frequent fusion gene found in AML and is the product of the t(8;21) chromosomal translocation (Klaus *et al.*, 2004; Sanderson *et al.*, 2006b; Grimwade *et al.*, 2010). t(8;21) AML is prominent in adolescence but with intensive chemotherapy a 60% survival rate can be achieved (Shipley and Butera, 2009). Despite this success, relapse of t(8;21) AML remains a major clinical problem. In both models of relapsed AML (see Chapter 1, Section 1.6), the relapsing clone evolves from a cell carrying leukaemia-initiating genetic lesions, including gene fusions, suggesting the persistence of pre-leukaemic fusion-positive clones susceptible to chemotherapy-induced mutation and which could drive relapse. As such, it is important to determine whether the expression of gene fusions affects the frequency and pattern of chemotherapy-induced mutation.

The mutagenic potential of DNA damaging agents can be evaluated by a number of *in vitro* and *in vivo* assays using a variety of reporter genes (e.g. *HPRT*, *APRT*, *XPRT*, *TK*, *PIGA*) to detect mutations induced by these agents (Aaron *et al.*, 1994). In mammalian cells, *TK* and *HPRT* are the most commonly used reporter genes for *in vitro* mutation studies. The *in vitro TK* and *HPRT* mutation assays work by treating cells with the agent of interest followed by exposure to a selective agent that results in death of non-mutant cells. The HPRT enzyme, encoded by the *HPRT* gene, is an important component of the purine salvage pathway. This enzyme is also required for conversion of 6-TG to its active DNA precursor, 6-thioguanosine monophosphate (6-TGMP) which inhibits DNA synthesis and is cytotoxic to cells. Due to loss of HPRT enzymatic function, mutant cells are resistant to the cytotoxic effect of 6-TG while cells with functional HPRT mutation are selectively killed by 6-TG. The *TK* assay works on the same principle in which TFT is used to select cells with *TK* mutation. The cells that survive and form colonies after the selection are *TK* mutant resulting from either a

spontaneous mutation or from an induced mutation caused by an exogenous DNA damaging agent. The frequency of induced mutants, relative to control untreated cells, can be used as a measure of the mutagenicity of the agent under investigation. In addition, the *TK* and *HPRT* assays detect different spectra of genetic events. The autosomal location of *TK* gene (chromosome 17) allows for detection of larger mutations, such as large deletions, which is unlikely to be detected at the *HPRT* locus on X chromosome (Liber and Thilly, 1982).

A previous study has successfully demonstrated the capability of this assay in determining the mutagenicity of cytarabine in mismatch repair proficient and deficient cell lines (Fordham *et al.*, 2015). Using a similar approach, in this study, the effect of *RUNX1/ETO* fusion gene on mutation frequency (MF) was determined at the *TK* and *HPRT* loci, both spontaneously and after treatment with chemotherapy agents in a cell model system using immortalised TK6 human lymphoblastoid cells. Importantly, these assays allow mutation in surviving cells to be measured simultaneously with cytotoxicity, such that the mutagenicity/cytotoxicity ratio can be calculated. In addition, *HPRT* mutant clones were generated (following selection of 6-TG resistant cell clones) and subjected to molecular analysis, allowed for a comparison of *HPRT* mutation spectra between chemotherapy-treated and vehicle-treated cells in order to determine the spectra of gene deletion mutants and base substitution mutants.

3.1.1. Aims of Chapter 3

The objective of this investigation is to identify the contribution of remission induction chemotherapy and expression of the RUNX1/ETO fusion gene to the aetiology of pathogenic mutations in AML using *in vitro* reporter gene assays.

Specifically, the experimental aims were as follows:

- Determine chemotherapy dose-dependent mutagenicity in a cell model system.
- Determine the effect of *RUNX1/ETO* fusion gene expression on spontaneous mutation frequency.
- Determine the effect of *RUNX1/ETO* fusion gene expression on chemotherapy-induced mutation frequency.
- Determine the spectrum of mutations at the *HPRT* locus, both spontaneously and after exposure to daunorubicin.

3.2. Results

3.2.1. Expression levels of the RUNX1/ETO fusion protein determined by western blotting

A TK6 cell clone with stable expression of full-length RUNX1/ETO (TK6 RE8) was previously generated via lentiviral transduction (see Chapter 2, Section 2.2). The clone was re-assessed for RUNX1/ETO expression, in order to confirm the expected phenotype. RUNX1/ETO expression was observed in TK6 RE8 cells line which has approximately 94% of RUNX1/ETO expression relative to Kasumi-1 but was absent in parental TK6 (Figure 3.1). These data confirm that the TK6 RE8 cells have the expected phenotype based on genotype and are suitable for use in subsequent experiments investigating the influence of *RUNX1/ETO* fusion gene on the acquisition of mutations, both spontaneously and after chemotherapy treatments.

3.2.2. Cytotoxicity in TK6 and TK6 RUNX1/ETO cells following treatment with daunorubicin (chronic drug exposure)

Cytotoxicity in response to treatment with daunorubicin was assessed in TK6 *RUNX1/ETO* cells relative to parental TK6 cells using the 96-well clonogenic assay. There was a progressive drop in cloning efficiency up to 6nM daunorubicin (Figure 3.2), demonstrating a significant dose response suggesting that daunorubicin is cytotoxic to TK6 cells. TK6 *RUNX1/ETO* cells were significantly resistant to cell killing mediated by daunorubicin compared to parental TK6 cells ($P = 0.01$; two-way ANOVA). This daunorubicin resistant phenotype might be partly caused by a slower growth rate of TK6 *RUNX1/ETO* cells compared to parental TK6 cells (Figure 3.3). These data also confirmed that TK6 and its derivative subclones have a high cloning efficiency (50 – 74%) as demonstrated by their ability to survive and proliferate at low density (2 cells per well) in liquid media (data not shown). Collectively, these data were used to identify a suitable starting dose of daunorubicin to be used in mutation assays investigating the effect of daunorubicin exposure on MF.

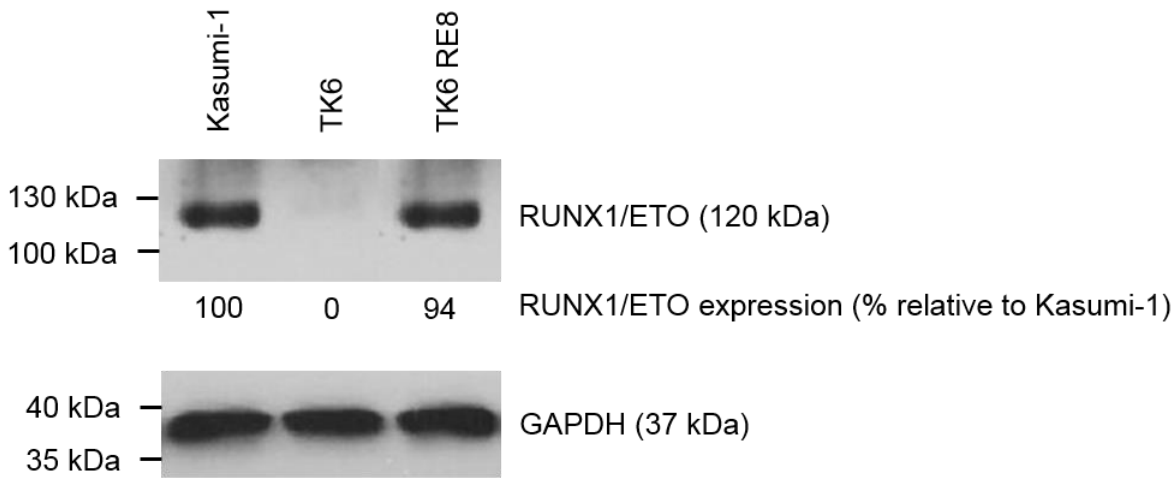


Figure 3.1. RUNX1/ETO protein expression in TK6 cell lines.

Western immunoblotting was performed to determine protein expression of RUNX1/ETO in parental TK6 and TK6 RE8 cells relative to Kasumi-1 cells. GAPDH was used as a loading control for all blots. Each blot is representative of three independent experiments.

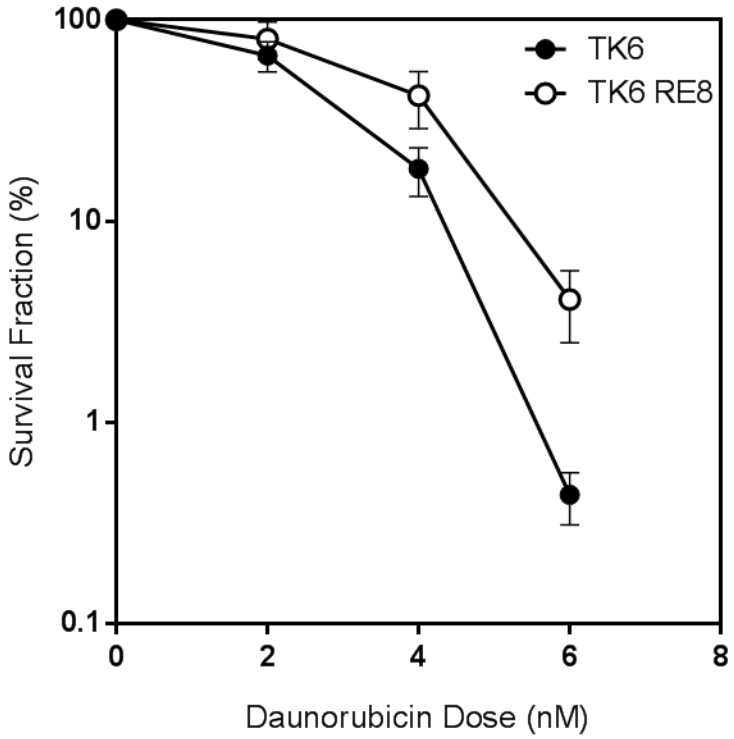


Figure 3.2. Cytotoxicity in response to daunorubicin in TK6 and TK6 RUNX1/ETO cells.

Cell survival in response to daunorubicin in parental TK6 cells (black circles) and TK6 *RUNX1/ETO* cells (throughout circles) determined via clonogenic assay in 96- well plates using limiting dilution. Data is presented as the number of viable colonies (positive wells) from dosed cell suspensions as a percentage of the number of viable colonies from vehicle-only treated cell suspensions. Data represents the mean and standard deviation (error bars) of three independent experiments.

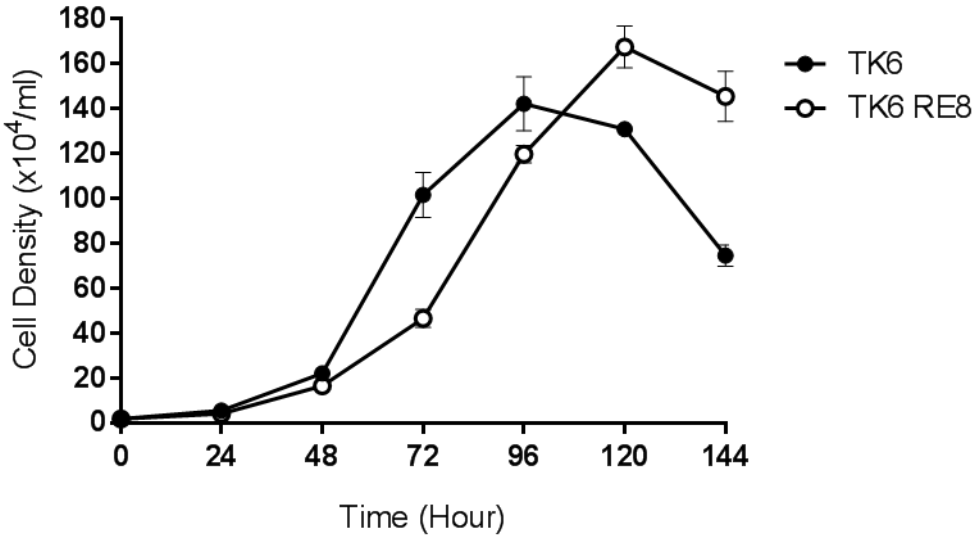


Figure 3.3. Cell proliferation in TK6 and TK6 RUNX1/ETO cell lines.

Cell proliferation was compared in parental TK6 cells (black circles) and TK6 *RUNX1/ETO* cells (throughout circles). Data is presented as the number of cells counted every 24 hours. Data represents the mean and standard deviation (error bars) of three independent experiments.

3.2.3. Cytotoxicity in TK6 and TK6 RUNX1/ETO cells following treatment with daunorubicin and ara-C (acute drug exposure)

In order to identify a suitable starting dose for daunorubicin and ara-C in each cell line, dose finding assays were performed by treating cell lines with escalating doses of daunorubicin and ara-C for 16 hours. This assay was adapted from the 96-well clonogenic assay with the aim of identifying doses which result in approximately 60 – 95% cytotoxicity to be used in *in vitro* mutation assays.

A progressive drop in colony forming efficiency in response to escalating doses of daunorubicin and ara-C demonstrated a significant dose response (Figure 3.4). A concentration of 15nM, which induces 70 – 80% cytotoxicity, was selected as a starting dose for daunorubicin mutagenicity assays. Likewise, a concentration of 1µM, which induces 70 – 80% cytotoxicity, was selected as a starting dose for ara-C mutagenicity assays.

3.2.4. Effect of daunorubicin exposure on mutation frequency at TK and HPRT loci in TK6 and TK6 RUNX1/ETO

To assess the mutagenicity of daunorubicin, cultures of parental TK6 and its *RUNX1/ETO*-positive subclone TK6 RE8, were first purged of any pre-existing spontaneous *TK* and *HPRT* mutants (see Chapter 2, Section 2.6.1), then exposed to increasingly doses of daunorubicin for 16 hours. MF was subsequently determined at the *TK* and *HPRT* loci in the surviving cells.

Exposure of TK6 and TK6 *RUNX1/ETO* cells to daunorubicin induced a significant dose-dependent increase in MF at both *TK* and *HPRT* loci, with the effect more prominent at the *TK* locus (Table 3.1 and Figure 3.5). Moreover, RUNX1/ETO significantly increase MF relative to RUNX1/ETO-negative cells in response to escalating doses of daunorubicin, although it was statistically not significant at the *HPRT* locus (Table 3.1). The mutagenicity of daunorubicin was further evaluated based on cytotoxicity induced by each individual dose of daunorubicin (Figure 3.6). An exponential increase in MF demonstrated that exposure to doses that induced higher levels of cytotoxicity increase the MF at both loci in both cell lines. A comparison of the daunorubicin-induced MF at doses giving near equivalent cytotoxicity was also performed to further determine whether RUNX1/ETO influences the mutagenic potency of daunorubicin relative to RUNX1/ETO-negative cells. At doses inducing

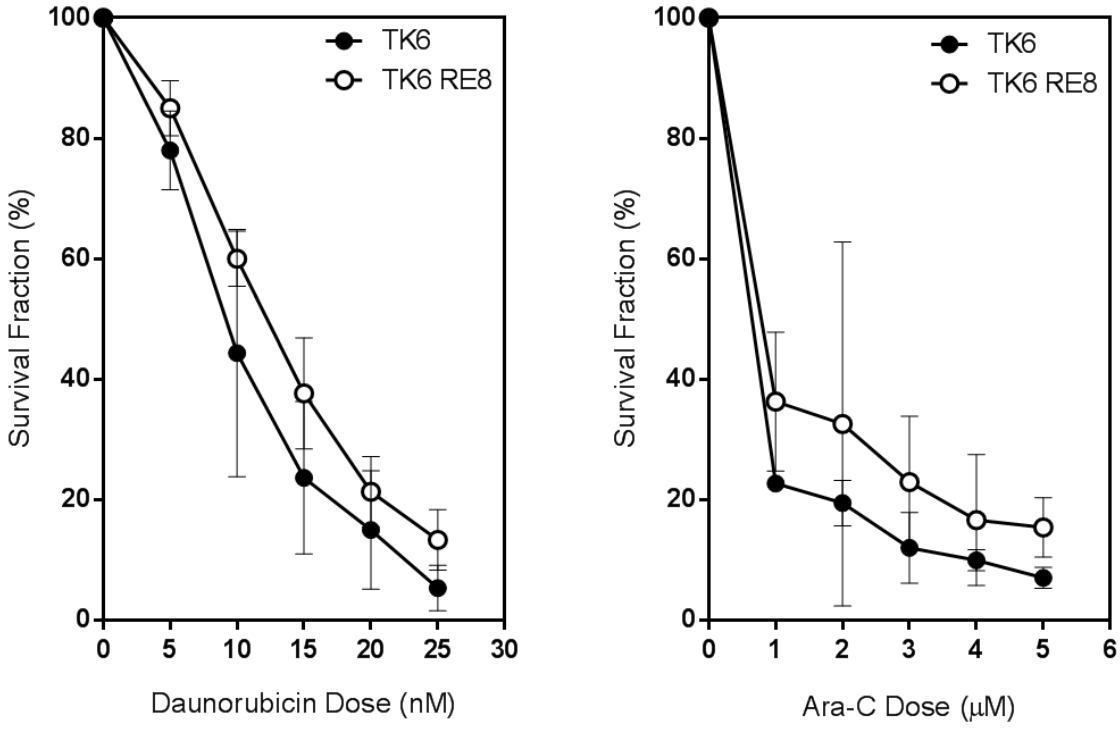


Figure 3.4. Cytotoxicity in response to daunorubicin and ara-C in TK6 and TK6 RUNX1/ETO cells.

Cell survival following 16 hours exposure to daunorubicin and ara-C in parental TK6 cells (black circles) and TK6 *RUNX1/ETO* cells (throughout circles) via clonogenic assay in 96 well plates using limiting dilution. Data is presented as the number of viable colonies (positive wells) from dosed cell suspensions as a percentage of the number of viable colonies from vehicle-only treated cell suspensions. Data represents the mean and standard deviation (error bars) of three independent experiments.

Daunorubicin (nM)	*TK* MF ($x10^{-6}$)		*HPRT* MF ($x10^{-6}$)	
	TK6	TK6 *RUNX1/ETO*	TK6	TK6 *RUNX1/ETO*
0	8.2 (0.0)	13.0 (0.0)	2.9 (0.0)	4.0 (0.0)
15	57.8 (49.6)	55.5 (42.5)	26.3 (23.4)	18.7 (14.7)
25	59.2 (51.0)	68.3 (55.3)	38.5 (35.6)	33.2 (29.2)
30	60.9 (52.7)	74.3 (61.3)	39.6 (36.7)	40.7 (36.7)
One-way ANOVA[a]	< 0.0001	< 0.0001	< 0.0001	< 0.0001
Two-way ANOVA[b]	0.0090		0.1696	

Table 3.1. Effect of daunorubicin exposure on mutation frequency at *TK* and *HPRT* loci.

Figures represent the observed MF at each drug concentration. Data represent the mean of three independent experiments. Numbers in parentheses represent the induced MF (x 10^{-6}) at each dose, calculated by subtraction of the spontaneous MF (0nM daunorubicin; vehicle-treated). These figures are displayed graphically in Figure 3.5.

[a] Figures represent the P value determined using a one-way analysis of variance (ANOVA) test for effect of daunorubicin dose on MF.

[b] Figures represent the P value determined using a two-way ANOVA test for effect of RUNX1/ETO status on MF.

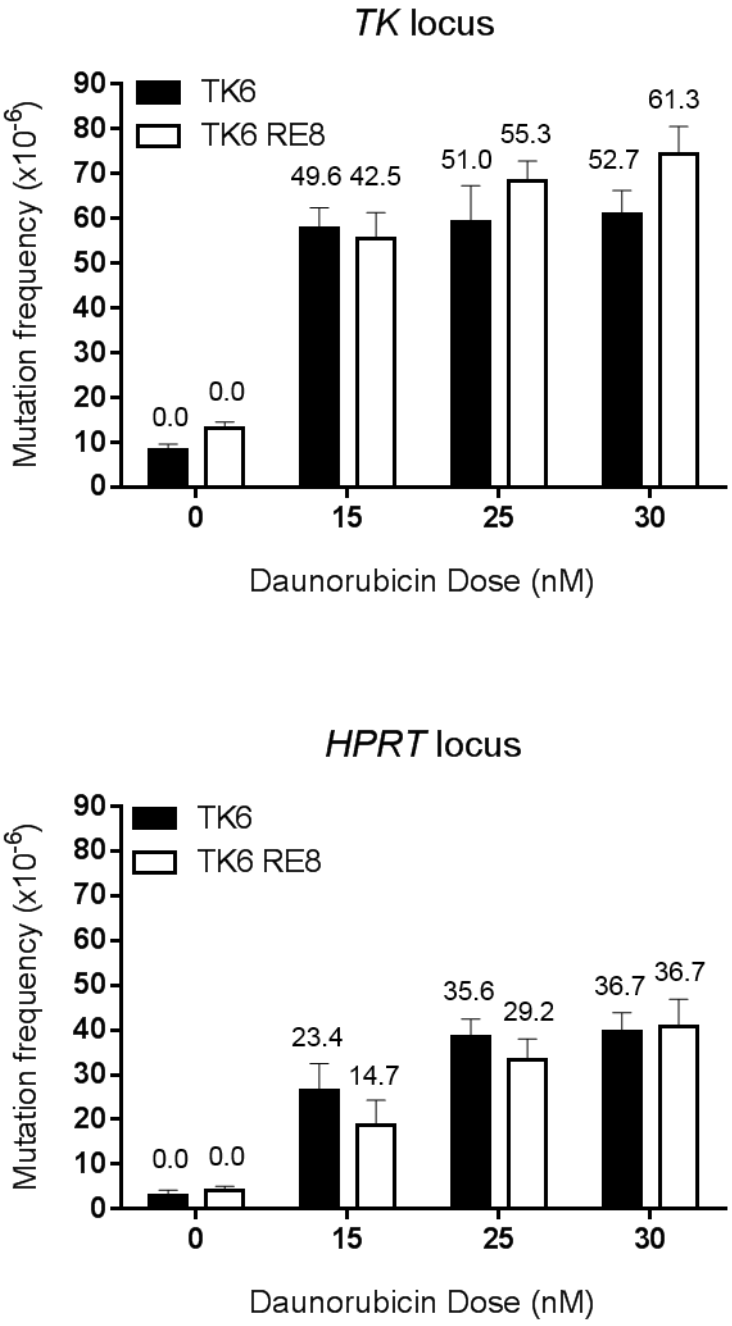


Figure 3.5. Effect of daunorubicin exposure on mutation frequency at the *TK* and *HPRT* loci.

MF at the *TK* and *HPRT* loci was assessed in parental TK6 cells (black bars) and TK6 *RUNX1/ETO* cells (throughout bars) following a 16 hour exposure to daunorubicin. Data represents the mean and standard deviation (error bars) of three independent experiments. Numbers above bars represent the induced MF (x 10^{-6}) at each dose, calculated by subtraction of the spontaneous MF (0nM daunorubicin; vehicle-treated).

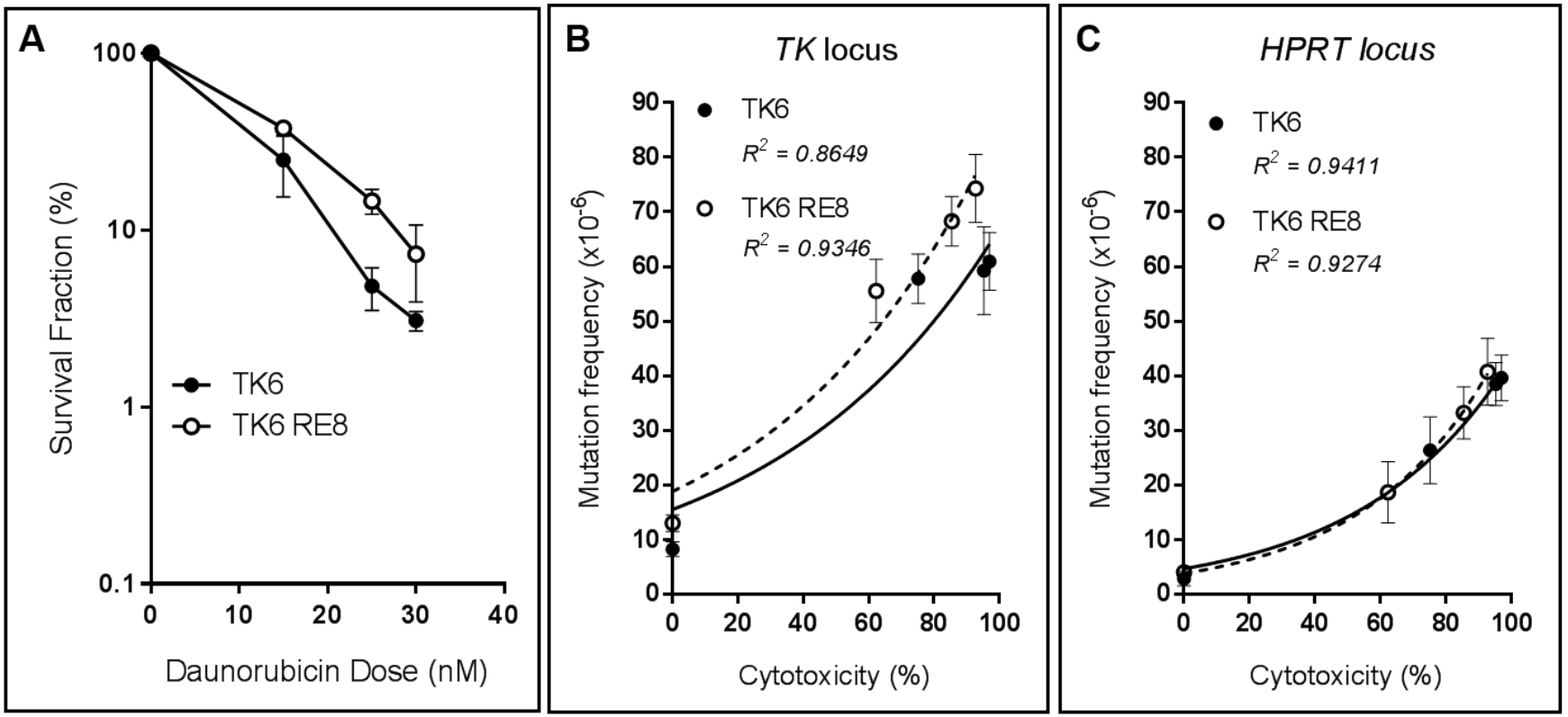


Figure 3.6. Effect of daunorubicin exposure on mutation frequency at the *TK* and *HPRT* loci as a function of cytotoxicity.
MF at the *TK* and *HPRT* loci was assessed in parental TK6 cells and TK6 *RUNX1/ETO* cells following 16 hours exposure to daunorubicin. Data represents the mean and standard deviation (error bars) of three independent experiments and is presented as MF observed in control vehicle-treated cells (0% cytotoxicity) and at drug doses causing up to 95% cytotoxicity. Exponential curves are displayed for each dataset. The associated R^2 values represent the square of the Pearson product moment correlation coefficient, indicating the 'goodness of fit' of the data.

approximately 95% cytotoxicity, there was no significant difference in induced MF between parental TK6 and TK6 *RUNX1/ETO*-positive cells despite a slight increase in MF at both loci (*TK*: *P* = 0.0857, *HPRT*: *P* = 0.8840; paired *t*-test) (Figure 3.7).

3.2.5. Effect of ara-C exposure on mutation frequency at the TK and HPRT loci in TK6 and TK6 RUNX1/ETO cells

The mutagenicity of ara-C was evaluated in the same way as daunorubicin, with higher doses of ara-C were used to induce the required cytotoxicity. Similar to the response observed for daunorubicin, exposure of TK6 and TK6 *RUNX1/ETO* cells to ara-C caused a significant dose-dependent increase in MF at both *TK* and *HPRT* loci, with the effect being more prominent at the *TK* locus (Table 3.2 and Figure 3.8). However, there was no significant increase in MF at both loci in RUNX1/ETO-positive cells relative to RUNX1/ETO-negative cells in response to escalating doses of ara-C (Table 3.2), which might be caused by unequal cytotoxicity induced by each individual ara-C dose. Despite that, an exponential increase in MF demonstrated that exposure to doses that induced higher levels of cytotoxicity increased the MF at both loci in both cell lines (Figure 3.9). In addition, comparison of ara-C induced MF at equitoxic doses (80% cytotoxicity) showed significant increase in TK6 *RUNX1/ETO* cells relative to parental TK6 cells at the *HPRT* locus, although it was not statistically significant at the TK locus (*TK*: *P* = 0.5970, *HPRT*: *P* = 0.0424; paired *t*-test) (Figure 3.10).

3.2.6. Effect of RUNX1/ETO expression on spontaneous mutation frequency at the TK and HPRT Loci

The effect of *RUNX1/ETO* fusion gene expression on the acquisition of mutations were further evaluated by comparison of MF in vehicle-treated control populations (Figure 3.11). In daunorubicin vehicle-treated cell populations, a significant increase in MF was observed at the *TK* locus in TK6 *RUNX1/ETO* cells relative to parental TK6 cells (*P* = 0.0109; paired *t*-test). Consistently, in ara-C vehicle-treated cells, a significant increase in MF was also observed at the *TK* locus in TK6 *RUNX1/ETO* cells relative to parental TK6 cells (*P* = 0.0469; paired *t*-test). Similar statistical testing performed at the *HPRT* locus however, showed no significant differences in spontaneous MF comparing RUNX1/ETO-positive and parental cell populations, although a modest increase in MF was observed in TK6 *RUNX1/ETO*

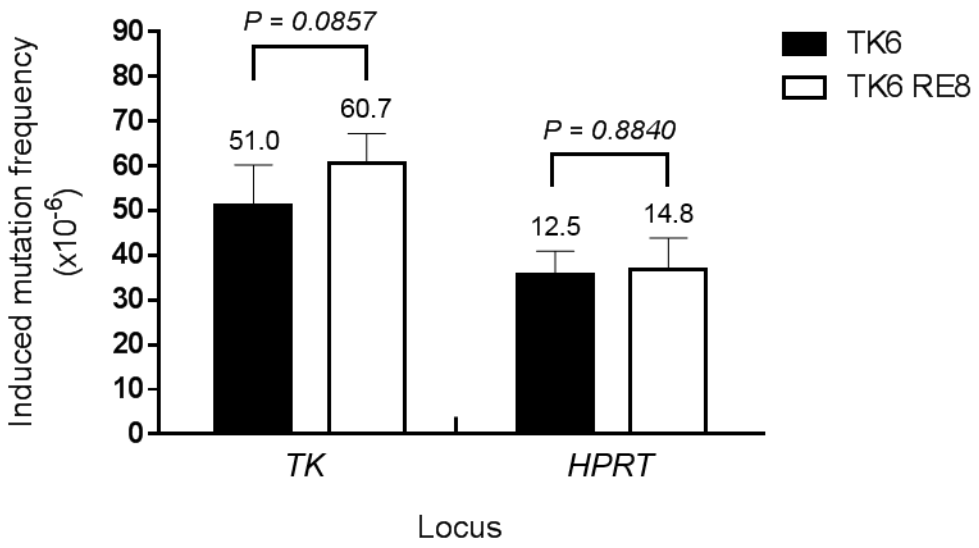


Figure 3.7. Daunorubicin-induced mutation frequency at the *TK* and *HPRT* loci at 95% cytotoxicity.

Induced MF at the *TK* and *HPRT* loci was assessed in parental TK6 cells (black bars) and TK6 *RUNX1/ETO* cells (throughout bars) at daunorubicin doses cytotoxic to approximately 95% of cells. Data represents the mean and standard deviation (error bars) of three independent experiments. Induced MF (numbers shown above bars) were calculated by subtraction of the spontaneous MF (0nM daunorubicin). P-value was calculated by paired *t*-test comparing the induced MF between cell lines at each locus.

Ara-C (μM)	*TK* MF		*HPRT* MF	
	TK6	TK6 *RUNX1/ETO*	TK6	TK6 *RUNX1/ETO*
0	9.5 (0.0)	14.5 (0.0)	2.6 (0.0)	3.8 (0.0)
1	38.4 (28.9)	36.7 (22.2)	11.5 (8.9)	5.8 (2.0)
3	46.6 (37.1)	43.7 (29.2)	15.1 (12.5)	13.7 (9.9)
5	52.9 (43.4)	53.5 (39.0)	22.0 (19.4)	18.6 (14.8)
One-way ANOVA[a]	< 0.0001	< 0.0001	0.0014	0.0014
Two-way ANOVA[b]	0.9171		0.2608	

Table 3.2. Effect of ara-C exposure on mutation frequency at the *TK* and *HPRT* loci.

Figures represent the observed MF at each drug concentration. Data represents the mean of three independent experiments. Numbers in parentheses represent the induced MF (x 10^{-6}) at each dose, calculated by subtraction of the spontaneous MF (0μM ara-C; vehicle-treated). These figures are displayed graphically in Figure 6.1.

[a] Figures represent the P value determined using a one-way analysis of variance (ANOVA) test for effect of ara-C dose on MF.

[b] Figures represent the P value determined using a two-way ANOVA test for effect of RUNX1/ETO status on induced MF.

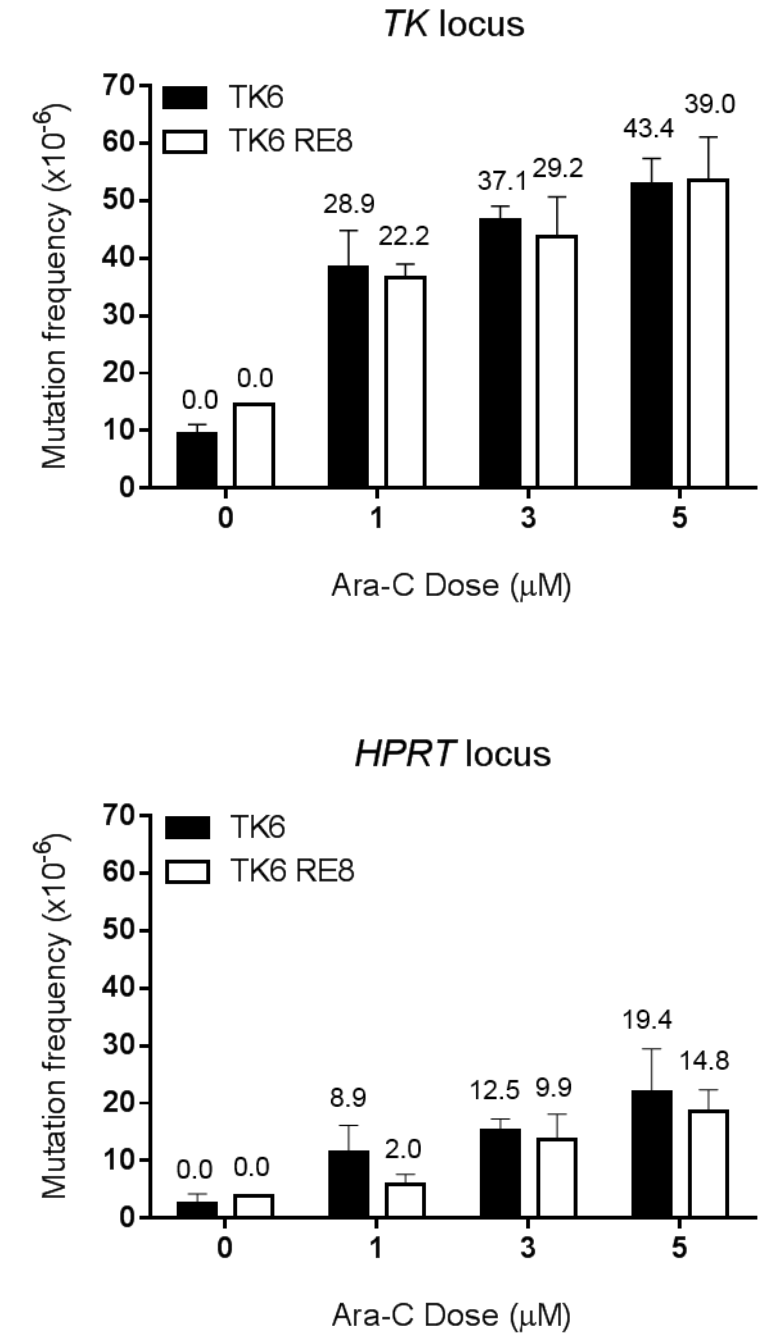


Figure 3.8. Effect of ara-C exposure on mutation frequency at the *TK* and *HPRT* loci.

MF at the *TK* and *HPRT* loci was assessed in parental TK6 cells (black bars) and TK6 *RUNX1/ETO* cells (throughout bars) following 16 hour exposure to ara-C. Data represents the mean and standard deviation (error bars) of three independent experiments. Numbers above bars represent the induced MF (x 10^{-6}) at each dose, calculated by subtraction of the spontaneous MF (0μM ara-C; vehicle-treated).

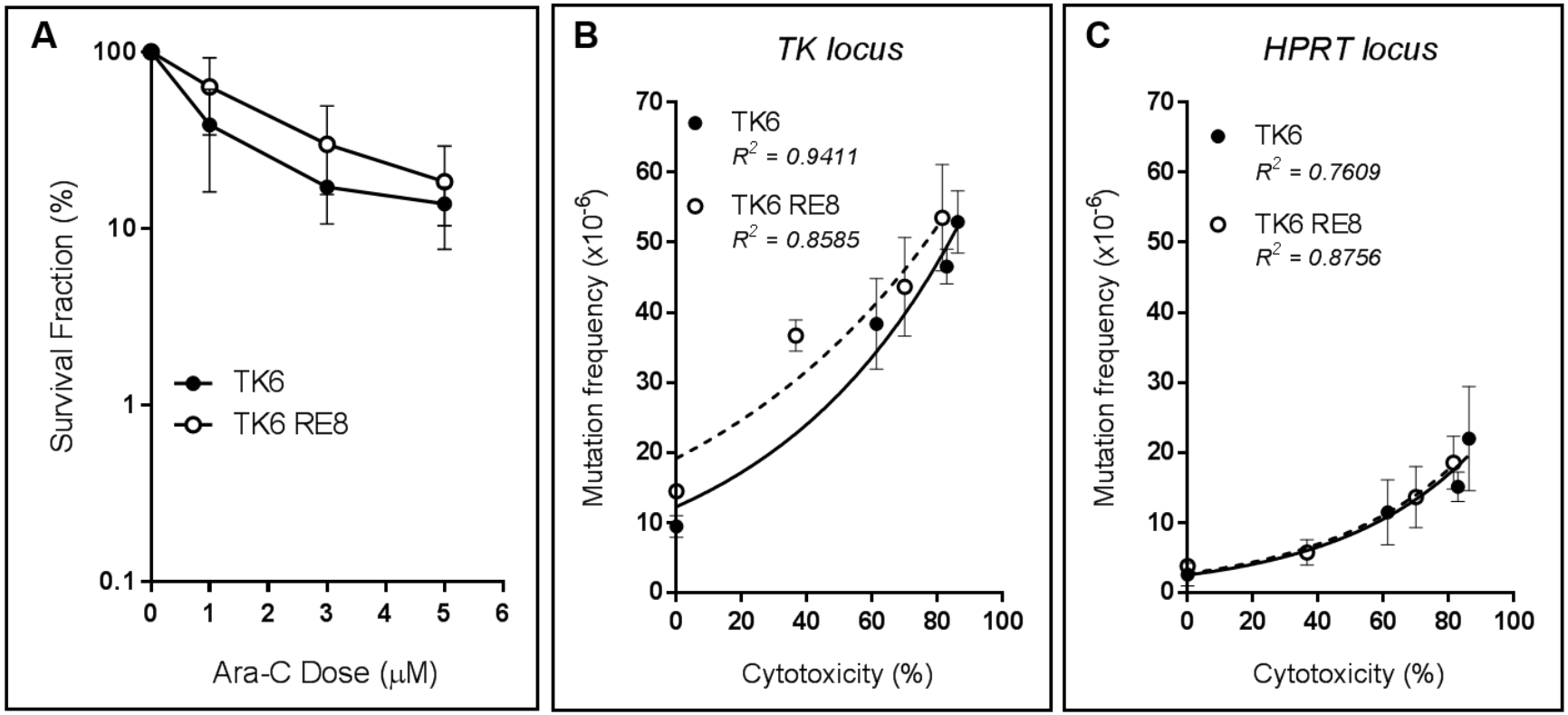


Figure 3.9. Effect of ara-C exposure on mutation frequency at the *TK* and *HPRT* loci as a function of cytotoxicity.

MF at the *TK* and *HPRT* loci was assessed in parental TK6 and TK6 *RUNX1/ETO* cells following 16 hours exposure to ara-C. Data represents the mean and standard deviation (error bars) of three independent experiments and is presented as MF observed in control vehicle-treated cells (0% cytotoxicity) and at drug doses causing up to 90% cytotoxicity. Exponential curves are displayed for each dataset. The associated R^2 values represent the square of the Pearson product moment correlation coefficient, indicating the 'goodness of fit' of the data.

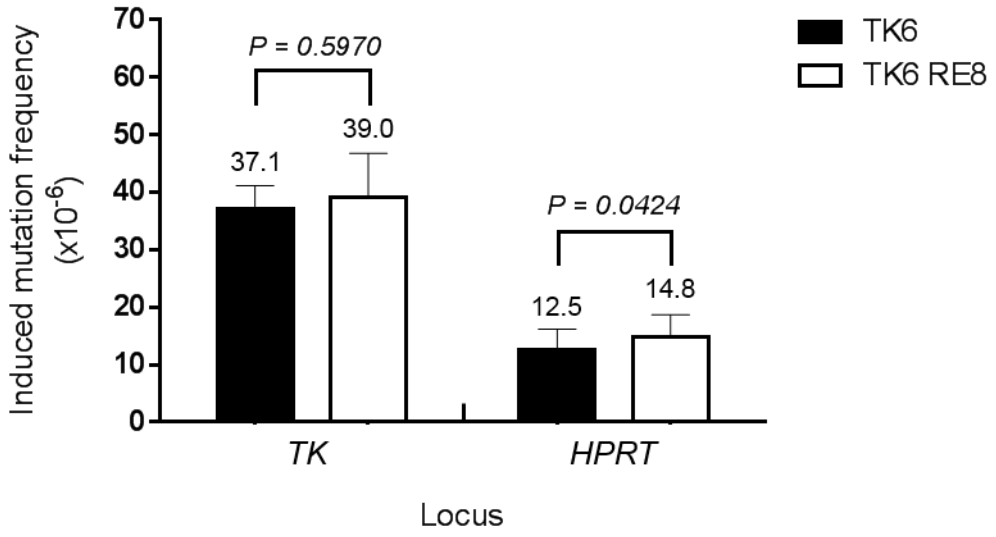


Figure 3.10. Ara-C-induced mutation frequency at the TK and HPRT loci at 80% cytotoxicity.

Induced MF at the *TK* and *HPRT* loci was assessed in parental TK6 cells (black bars) and TK6 *RUNX1/ETO* cells (throughout bars) at ara-C doses inducing approximately 80% cytotoxicity. Data represents the mean and standard deviation (error bars) of three independent experiments. Induced MF (numbers shown above bars) were calculated by subtraction of the spontaneous MF (0µM ara-C; vehicle-treated). P-value was calculated by paired *t*-test comparing the induced MF between cell lines at each locus.

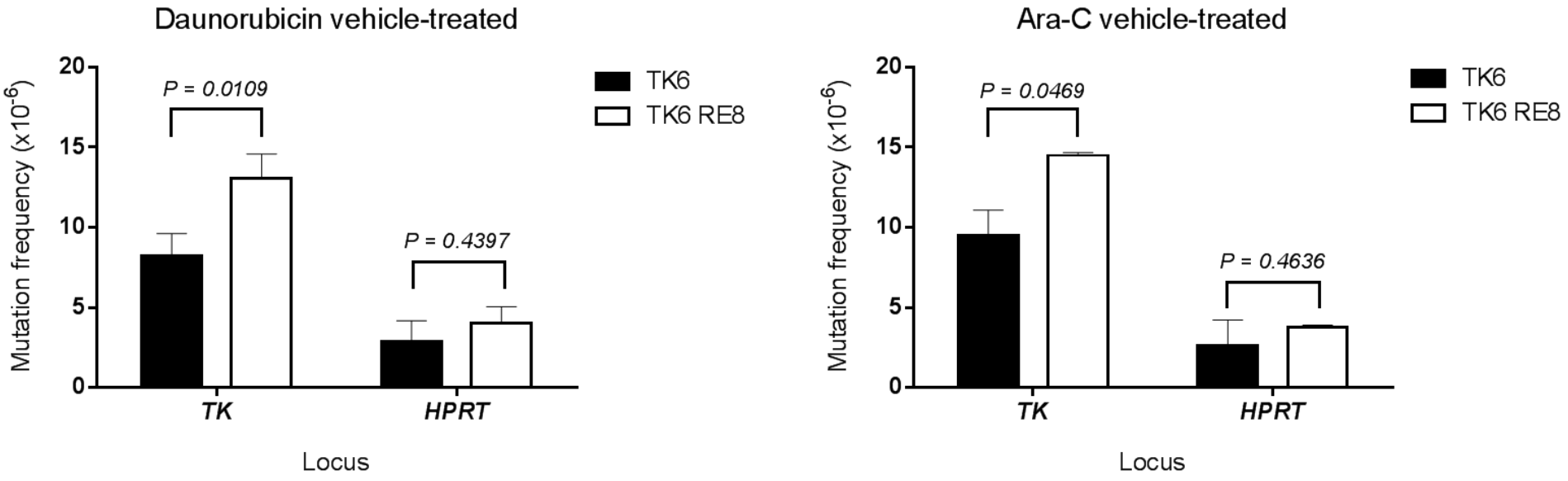


Figure 3.11. Spontaneous mutation frequency at the *TK* and *HPRT* loci.

MF at the *TK* and *HPRT* loci was assessed in parental TK6 cells (black bars) and TK6 *RUNX1/ETO* cells (throughout bars) following 16 hour exposure to vehicle control of daunorubicin and ara-C. Numbers above bars represent the MF value (x 10^{-6}) and data represents the mean and standard deviation (error bars) of three independent experiments. P-value was calculated by paired *t*-test comparing the MF between cell lines at each locus.

cells compared to parental TK6. Taken together, these data demonstrate that *RUNX1/ETO* expression increases spontaneous MF at the *TK* locus, with some evidence of higher mutation at the *HPRT* locus.

3.2.7. Molecular analysis of mutations at the HPRT locus in TK6 and TK6 RUNX1/ETO cells

To further investigate the mutagenicity of daunorubicin and the influence of the *RUNX1/ETO* fusion gene on the acquisition of mutations, in terms of the structural types and their frequency, a number of individual *HPRT* mutants generated from TK6 and TK6 *RUNX1/ETO* cells were subject to molecular analysis in order to generate both spontaneously-occurring and daunorubicin-treated mutational spectra.

A total of 154 and 292 independent *HPRT* mutants were generated from TK6 and TK6 *RUNX1/ETO* cells. Likewise, a total of 56 and 50 independent *HPRT* mutants were generated from parental TK6 cells and TK6 *RUNX1/ETO* cells, respectively, following a dose of daunorubicin that induced 95% cytotoxicity. RNA was extracted from each individual *HPRT* mutant for reverse-transcription PCR (RT-PCR), PCR and cDNA sequencing for characterisation of mutations.

3.2.7.1. Analysis of mutations in the HPRT coding region

HPRT mutants characterised by the absence of a cDNA PCR product (Figure 3.12) were defined as having terminal deletions (either 5' or 3') and were grouped by the structure of large deletion (Table 3.3). For parental TK6 cells, a total of 41 (73%) daunorubicin-treated mutants, as well as 56 (36%) spontaneously-occurring mutants were found to harbour terminal deletions at the *HPRT* locus. Likewise, 28 (53%) daunorubicin-treated *HPRT* mutants and 105 (36%) spontaneously-occurring mutants carry a terminal deletion at this locus for TK6 *RUNX1/ETO HPRT* mutant population. However, no further investigation was performed to distinguish terminal deletions from complete gene deletions.

Internal exon deletions were defined as the absence of one or more consecutive exons in a gene, as evidenced by the presence of a cDNA PCR product, but which is shorter than full-length *HPRT* cDNA (737 bp) (Figure 3.12). For parental TK6 *HPRT* mutant population, a total of 14 (25%) daunorubicin-treated mutants, as well as 54 (35%) spontaneously-occurring mutants were found to harbour internal exon deletions

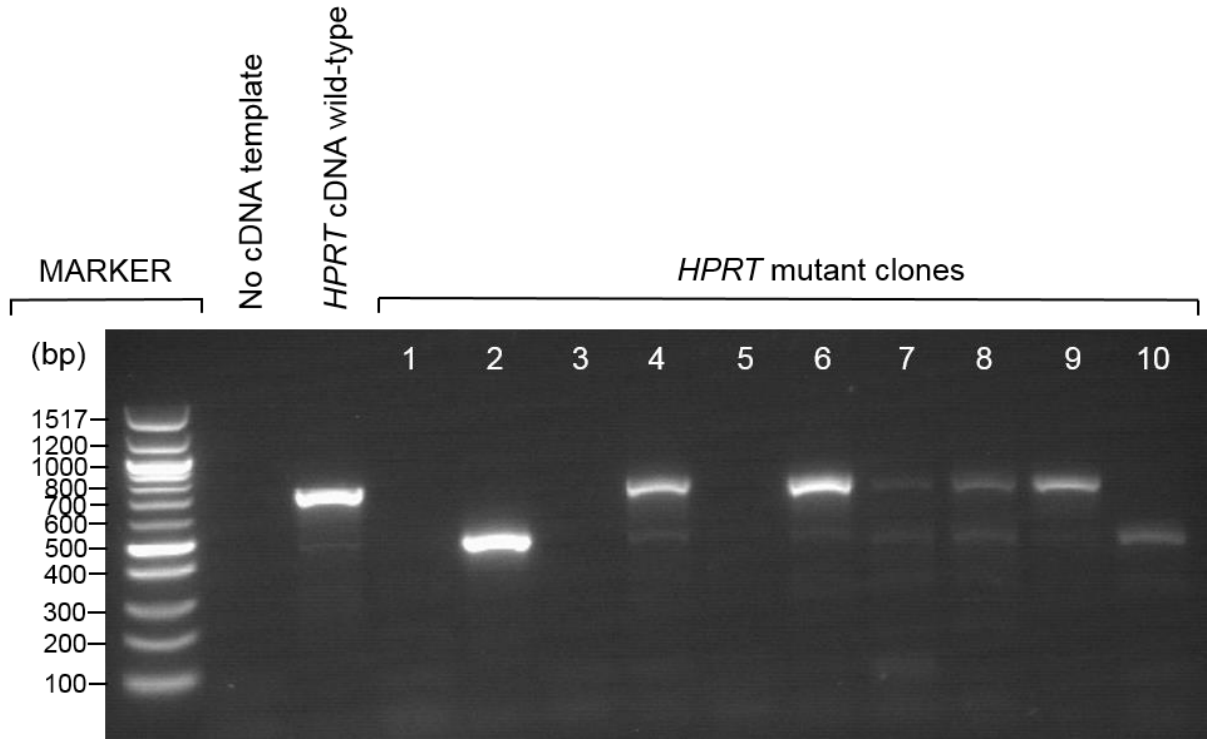


Figure 3.12. Large deletion in HPRT coding region.

The gel shows examples of PCR profiles from mutant with terminal deletions (lane 1, 3, 5), as evidenced by the absence of a cDNA product, and from mutants with internal exon deletion (lane 2, 10), as evidenced by shorter than full-length *HPRT* cDNA wild-type sequence (737 bp). 'No cDNA template' was used as negative control and *HPRT* cDNA wild-type (untreated parental TK6 cells) was used as positive control for PCR amplification.

Deleted Region	Spontaneous		Daunorubicin-treated	
	TK6	TK6 *RUNX1/ETO*	TK6	TK6 *RUNX1/ETO*
5' / 3' terminal deletions				
Total	56	105	41	28
Internal exon deletions				
1	6			
1 – 3		1		
2	4	7	2	2
2 – 3	5	28	3	9
2 – 4	1	1		
2 – 6		4		1
2 – 8	1			
3	2		2	
4	16	14		1
4 – 5	1	5		2
4 – 6		4		
3 – 8			1	
5	3	20	1	3
5 – 6	1	1	1	
6	4	15	1	2
6 – 8			1	
7	5	5		
7 – 8	4	4	2	1
8	1	3		
Total	54	112	14	21
Total Large Deletion	**110**	**217**	**55**	**49**

Table 3.3. Large deletions observed in spontaneous and daunorubicin-treated *HPRT* mutants.

Figures represent the number of deletions observed in both spontaneously-occurring and daunorubicin-treated *HPRT* mutant populations. Internal exon deletions were further classified according to the extent of the deleted region and affected exons.

at the *HPRT* locus. Likewise, 21 (42%) daunorubicin-treated *HPRT* mutants and 112 (38%) spontaneously-occurring mutants carried an internal *HPRT* exon deletion derived from TK6 *RUNX1/ETO HPRT* cells. cDNA sequencing was employed to determine the extent of the deleted region. However, no notable differences in the structural type of internal exon deletions were observed between daunorubicin-treated and spontaneously-occurring *HPRT* mutants, with both *HPRT* mutant populations have a high proportion of deletions encompassing exons 2 – 3.

For *HPRT* mutants with full length *HPRT* cDNA (737 bp), cDNA sequencing was used to identify small deletions/insertions and base substitutions (Figure 3.13). Of these, a single mutant clone carrying a base substitution was observed in parental TK6 and TK6 *RUNX1/ETO HPRT* populations treated with daunorubicin, which included a transition mutation (A > G) in parental TK6 and a transversion mutation (T > A) in *TK6 RUNX1/ETO* (Table 3.4). In contrast, a total of 21 (14%) and 49 (17%) base substitutions were observed in spontaneously-occurring parental TK6 and TK6 *RUNX1/ETO* HPRT mutant populations, respectively. Moreover, there was more T > G / A > C, T > C / A > G and C > A / G > T base substitutions derived from TK6 *RUNX1/ETO* cells compared to parental TK6. The comprehensive analysis of base substitution mutation will be described in the next chapter (Chapter 4). In addition, 23 (15%) and 26 (9%) spontaneous *HPRT* mutants derived from parental TK6 and TK6 *RUNX1/ETO HPRT* were small deletions/insertions (Table 3.5). This class of mutation was not observed in either parental TK6 or TK6 *RUNX1/ETO* daunorubicin-treated cell populations.

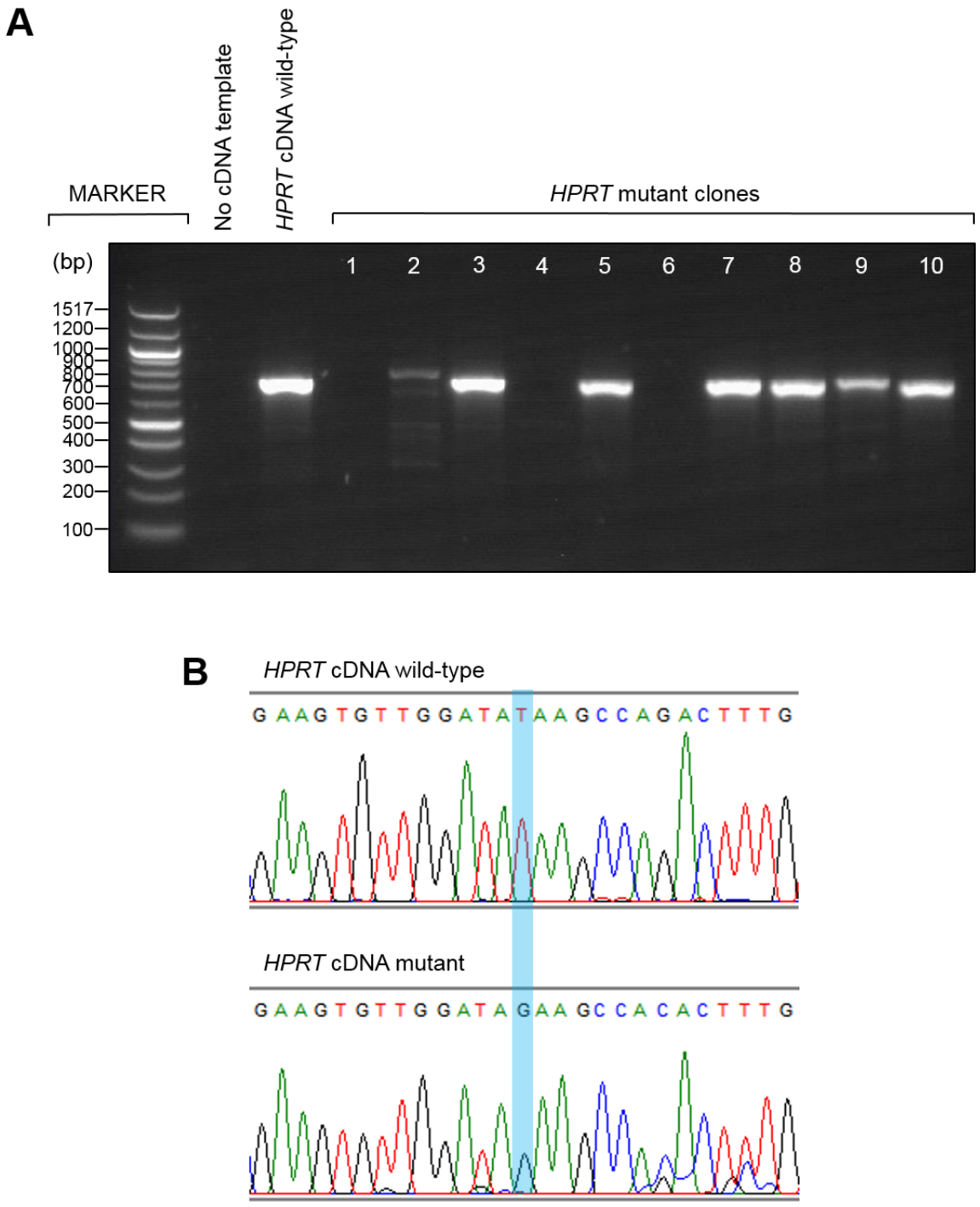


Figure 3.13. Small aberration in *HPRT* coding region.

A. The gel shows example of PCR profiles from mutant with base substitutions or small deletions/insertions (lane 3, 5, 7, 8, 9, 10), as evidenced by similar size with full length of *HPRT* cDNA wild-type sequence (737 bp). 'No cDNA template' was used as negative control and *HPRT* cDNA wild-type (untreated parental TK6 cells) was used as positive control for PCR amplification.

B. Sanger sequencing of cDNA from one *HPRT* mutant clone (lane 3) revealed a T > G base substitution at position E7:522 of the *HPRT* cDNA.

Position	Sequence change	Target sequence	Spontaneous		Dauno-treated	
			TK6	TK6 RE	TK6	TK6 RE
Transitions						
E2:48	T > C	AGGTTAT		1		
E2:118	G > A	CATGGAC	1			
E2:121	C > T	GGACTAA		1		
E2:122	T > C	GACTAAT		1		
E3:145	C > T	CGTCTTG		1		
E3:151	C > T	GCTCGAG		1		
E3:281	C > T	TTCCTAT		1		
E4:325	C > T	GACCAGT	1	2		
E4:355	G > A	GGTGGAG		2		
E5:395	T > C	TGATTGT	1			
E5:401	A > G	TGGAAGT		4		
E6:409	A > G	ATAATTG		2		
E6:416	C > T	ACACTGG		1		
E6:419	G > A	CTGGCAA	1	1		
E6:425	C > T	AAACAAT	1			
E6:440	T > C	TGCTTTC		1		
E6:454	C > T	AGGCAGT	1			
E7:508	C > T	CCACGAA	2			
E7:530	A > G	CAGACTG	1			
E8:533	T > C	TAGTTGT	1			
E8:539	G > A	TTGGATT		1		
E8:569	G > A	TAGGATA	1			
E8:598	A > G	TTCAGGG			1	
E9:610	C > T	TAGCATG	1	2		

Table 3.4. Base substitutions observed in spontaneous and daunorubicin-treated *HPRT* mutants (continued on next page).

Position	Sequence change	Target sequence	Spontaneous		Dauno-treated	
			TK6	TK6 RE	TK6	TK6 RE
Transversions						
E1:1	A > T	GTTATGG	1	1		
E1:2	T > A	TTATGGC	1			
E2:82	T > G	CATTATG		1		
E2:107	T > G	TATTCCT	1			
E2:109	A > T	TTTATTC		1		
E2:110	T > A	TTATTCC				1
E2:122	T > G	GACTAAT		1		
E2:124	A > T	CTAATTA		1		
E2:130	G > T	ATGGACA	1			
E3:136	A > C	GACTGAA		3		
E3:148	G > C	CTTGCTC		1		
E3:194	T > G	CCCTCTG		1		
E3:197	G > C	TCTGTGT		1		
E3:211	G > C	GGGGGCT		1		
E3:216	T > G	AATAGT		1		
E3:229	G > T	GCTGACC		1		
E3:233	T > A	ACCTGCT	1			
E3:238	G > T	CTGGATT		1		
E3:292	G > T	GTAGATT		1		
E4:354	T > G	TGGTGGA		1		
E5:388	G > T	AATGTCT		2		
E5:397	G > T	ATTGTGG		1		
E6:407	T > G	ATATAAT		1		
E6:409	A > T	ATAATTG	1			

Table 3.4. Base substitutions observed in spontaneous and daunorubicin-treated *HPRT* mutants (continued from previous page).

Position	Sequence change	Target sequence	Spontaneous		Dauno-treated	
			TK6	TK6 RE	TK6	TK6 RE
E6:410	T > A	TAA<u>T</u>TGA		1		
E6:481	G > C	GTC<u>G</u>CAA	1			
E6:482	C > A	TCG<u>C</u>AAG	1			
E7:522	T > G	ATA<u>T</u>AAG		1		
E8:533	T > G	TAG<u>T</u>TGT		1		
E8:585	T > G	ATT<u>A</u>TAG		2		
E8:606	G > C	TTT<u>G</u>AAT	1			
E9:610	C > A	TAG<u>C</u>ATG		1		
Total			**21**	**49**	**1**	**1**

Table 3.4. Base substitutions observed in spontaneous and daunorubicin-treated *HPRT* mutants.

Figures represent the number of mutations observed in both spontaneously-occurring and daunorubicin-treated *HPRT* mutant populations. Each mutation is identified by the exon in which it occurred and is numbered according to the cDNA position. Within the target sequence the affected base is underlined.

Affected Region	Spontaneous		Daunorubicin-treated	
	TK6	TK6 *RUNX1/ETO*	TK6	TK6 *RUNX1/ETO*
Small deletion / insertion				
E1:18	1			
E1:27		1		
E2:34	1			
E2:36	2			
E2:37	1			
E2:40	1	1		
E2:70-86	1			
E2:93		1		
E2:114-124		1		
E2:115		1		
E3:140	4			
E3:144		1		
E3:148	1			
E3:158-171	1			
E3:179	1			
E3:204-231		1		
E3:208-225		1		
E3:215-228	1			
E3:231	1			
E3:239		1		
E3:275-294	1			
E3:289-290	1			
E4:352		1		

Table 3.5. Small deletion/insertion observed in spontaneous and daunorubicin-treated *HPRT* mutants (continued to next page).

Affected Region	Spontaneous		Daunorubicin-treated	
	TK6	TK6 *RUNX1/ETO*	TK6	TK6 *RUNX1/ETO*
Small deletion / insertion				
E5:392		1		
E5:399	2			
E5:402		1		
E6:425		1		
E6:436-444		1		
E6:440		1		
E6:441-443	1			
E6:445		1		
E6:471-476		1		
E7:513-525	1			
E8:536-540		1		
E8:583-606		1		
E8:569-578		1		
E8:607		1		
E9:610-626		2		
E9:615		1		
E9:615-634		1		
E9:633-634		1		
E9:645	1			
Total	**23**	**26**	**0**	**0**

Table 3.5. Small deletion/insertion observed in spontaneous and daunorubicin-treated *HPRT* mutants (continued from previous page).

Figures represent the number of mutations observed in both spontaneously-occurring and daunorubicin-treated *HPRT* mutant populations. Small deletions/insertions were further classified according to the extent of the affected region and exons.

3.2.7.2. Comparison of spontaneous and daunorubicin-treated HPRT mutational spectra

The spectrum of mutations at the *HPRT* locus derived from daunorubicin-treated cells was compared to that derived from vehicle-treated cells (spontaneously-occurring mutation) (Figure 3.14). Specifically, the proportion of large deletions, comprising terminal and internal exon deletions was significantly higher in daunorubicin-treated cells (55/56; 98%) relative to vehicle-treated cells (110/154; 71%) in parental TK6 mutant population ($P < 0.0001$; two-tailed Fisher's exact test). Similarly, in TK6 *RUNX1/ETO*-expressing cells, the proportion of large deletions was also significantly higher following daunorubicin treatment (49/50; 98%) relative to vehicle treatment (217/292; 74%) ($P < 0.0001$; two-tailed Fisher's exact test). Strikingly, only 1 base substitution event (TK6: 1/56; 2%, TK6 RUNX1/ETO: 1/50; 2%) was observed in each daunorubicin-treated mutant population compared to 21 (14%) and 49 (17%) base substitution events observed in vehicle-treated parental TK6 and TK6 *RUNX1/ETO* cells, respectively. Taken together, this demonstrates that the spectrum of daunorubicin-treated *HPRT* mutant is clearly distinct from the spontaneously-occurring mutational spectrum.

Daunorubicin-treated mutant populations were compared to investigate the influence of the *RUNX1/ETO* fusion gene on the daunorubicin-induced mutational spectrum. A two-tailed Fisher's exact test demonstrated no significant difference in frequency for each class of mutation between both *HPRT* mutant populations (data not shown), although a higher proportion of internal exon deletion was observed in TK6 *RUNX1/ETO*-positive cells (42%) compared to parental TK6 cells (25%). In addition, similar statistical testing performed for each class of mutation between both vehicle-treated *HPRT* mutant populations also did not reach statistical significance (data not shown). Collectively, these data demonstrated that the *RUNX1/ETO* fusion gene does not significantly affect the frequency of each class of *HPRT* mutation, either spontaneously or after daunorubicin treatment.

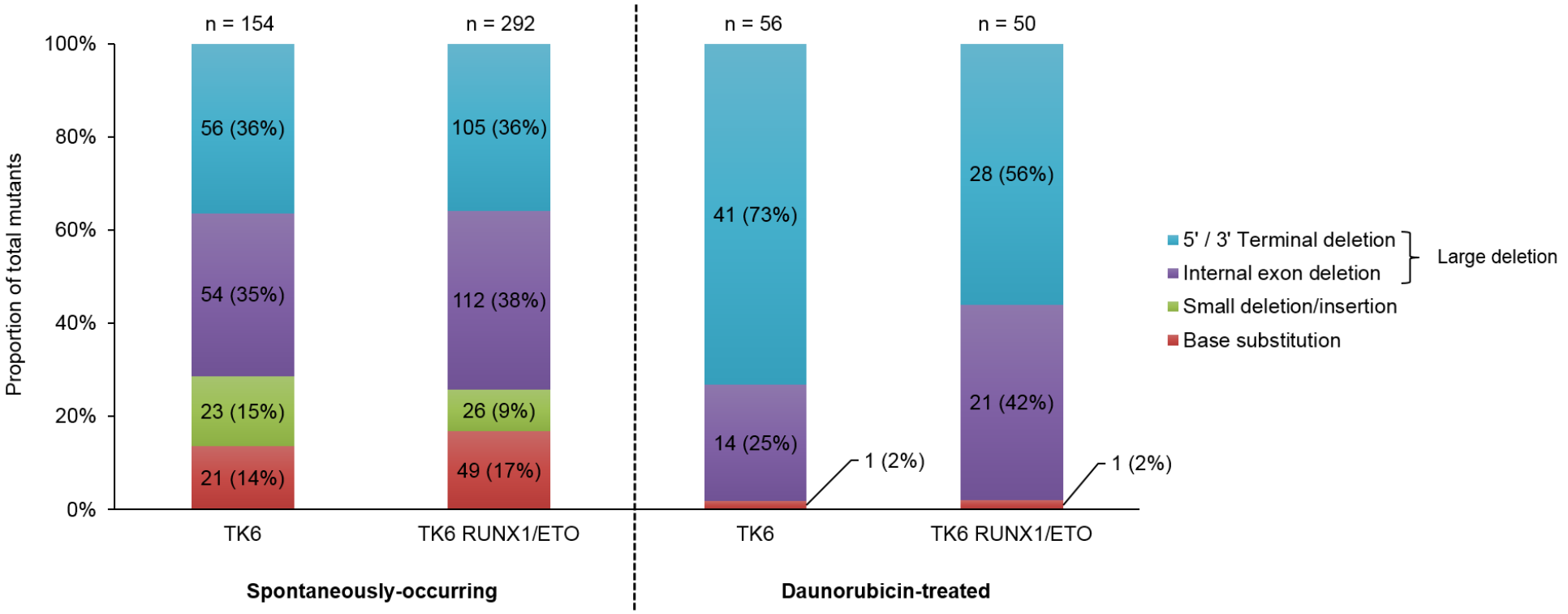


Figure 3.14. Mutational spectra of spontaneously-occurring and daunorubicin-treated mutants derived from parental TK6 and TK6 RUNX1/ETO-positive cells.

Spectra of mutations observed at the *HPRT* locus following 16 hour exposure to 30nM daunorubicin (inducing ~95% cytotoxicity) or vehicle control in TK6 and TK6 *RUNX1/ETO* cells. Coloured sections represent different structural classes of mutation. Numbers in parentheses indicate the percentage of mutants in each class. Total number of mutant analysed (n) is shown above the bars.

3.3. Discussion

The mechanisms by which novel mutations are acquired at relapse remain unclear, although there is strong evidence that some are chemotherapy-induced (Fordham *et al.*, 2015). The carcinogenicity of many chemotherapy agents is partly dependent on their ability to induce pro-mutagenic DNA damage. However, there is lack of data investigating the mutagenic potential of chemotherapy agents despite widespread use in the treatment of AML. In addition, emerging evidence has demonstrated the influence of the *RUNX1/ETO* fusion gene on the acquisition of mutations, both spontaneously and after exposure to genotoxic agents (Forster *et al.*, 2015), suggesting a major contribution of this fusion oncogene to the induction of genomic instability and mutagenesis. The main purpose in this study therefore, is to determine whether RUNX1/ETO expression sensitises cells to mutations at the *TK* and *HPRT* loci, both spontaneously and after exposure to chemotherapeutic agents.

The *TK* and *HPRT* mutation assays used in this study are clonogenic assays, a type of *in vitro* cell survival assay based on the ability of a single cell to grow into a colony. For this reason, TK6 human lymphoblastoid cells were selected as a cell model to investigate the mutagenicity of chemotherapy agents due to a high cloning efficiency as demonstrated by their ability to survive and proliferate at low density (2 cells per well) in liquid media. TK6 is a B lymphoblastoid line that is easy to maintain and has a low somatic mutation rate (~ 0.3%) (Sie *et al.*, 2009), which make it the preferred choice for *in vitro* mutation studies. Moreover, prior to initiation of the assay, pre-existing *TK* and *HPRT* mutant cells can be purged from the cell population to minimise the intrinsic mutation frequency at these loci, allowing for an accurate measurement of the induced mutation frequency as a function of time and exposure, such as oncoprotein expression or chemotherapy.

TK6 cells are heterozygous at the *TK* locus and hemizygous at the *HPRT* locus (refs). Therefore, mutagenic mechanisms that involve homologous interaction such as gene conversion cannot occur at the X-linked *HPRT* locus. Similarly, large multi-locus deletions are unlikely to occur at the *HPRT* locus since large deletions may span adjacent genes essential for cell survival. Therefore, the mutagenic effect of chemotherapy agents might be underestimated by the *HPRT* mutation assay if those agents induce homologous interaction or large deletion. However, performing both *TK* and *HPRT* mutation assays simultaneously using the same cell population may

compensate to some extent for the limitations of each individual locus as a reporter gene, thus providing accurate information on the mutagenicity of that agents.

As expected, a statistically significant dose-dependent increase in MF was observed at both *TK* and *HPRT* loci in both TK6 and TK6 *RUNX1/ETO* cells following 16 hour exposure to escalating doses of daunorubicin and ara-C, suggesting that these chemotherapy agents are mutagenic at the doses investigated. As a nucleoside analogue which is similar to human cytosine deoxyribose (deoxycytidine), ara-C can be incorporated into replicating DNA at internucleotide positions leading to inhibition of chain extension (Graham and Whitmore, 1970; Zahn *et al.*, 1972). The induced chain termination gives rise to stalled replication forks, which can in turn be converted to DSBs (Arnaudeau *et al.*, 2000; Arnaudeau *et al.*, 2001). This could result in large deletions if two DSBS are in close proximity due to major loss of DNA sequence. There are two repair mechanisms for DSBs; however, HR is thought to be the dominant mechanism for DSBs at stalled replication forks (Arnaudeau *et al.*, 2001; Lundin *et al.*, 2002; Rothkamm *et al.*, 2003), which raises the possibility that ara-C could also induce gene conversion type mutations. Collectively, large deletions and gene conversion mutations could give rise to LOH which may partly explain the higher ara-C-induced MF observed at the *TK* locus relative to the *HPRT* locus. This finding is consistent with data demonstrating a greater increase in MF at the *TK* locus than at the *HPRT* locus after exposure to ara-C (Fordham *et al.*, 2015). In addition, studies investigating Zidovudine (3'-Azido-2',3'-dideoxythymidine), a nucleoside analogue which also causes DNA chain termination also demonstrated a higher MF at the *TK* locus than at the *HPRT* locus (Sussman *et al.*, 1999; Meng *et al.*, 2000).

The cytotoxicity of daunorubicin is primarily mediated by the inhibition of topo II enzyme (Nielsen *et al.*, 1996; Arcamone *et al.*, 1997). In normal condition, topo II binds to DNA to catalyse the relaxation of supercoiled DNA distal to replication fork during DNA replication by inducing transient double strand break and rejoining both strands of the double helix (see Section 1.7.2, Chapter 1). Intercalation of daunorubicin between DNA base pairs stabilises the topo II-cleaved DNA covalent complex, preventing the DNA double helix from being resealed and thereby inhibiting replication (see Chapter 1, Section 1.7.2). Many anti-cancer drugs including anthracyclines that target topo II are reported to increase the concentration of topo II-DNA covalent complexes and decrease the re-ligation rate which eventually collapse to generate a DNA DSB (Fortune and Osheroff, 2000) and which can then induce chromosomal breakage and cell death. Indeed, evidence has shown that daunorubicin and related

anthracyclines are direct inducers of DNA strand breaks (Blasiak *et al.*, 2002; Yang *et al.*, 2015). As such, induced DNA damage, responsible for daunorubicin cytotoxicity, may cause genomic rearrangements (translocations, inversions or deletions) as a consequence of attempts to repair the strand break (Felix Carolyn, 2001). NHEJ is a a major mechanism for the repair of double strand breaks in DNA; however, data from wing Somatic Mutation and Recombination Test (SMART) in *Drosophila melanogaster* demonstrated a high frequency of HR after exposure to daunorubicin, indicating daunorubicin as a direct inducer of HR (Lehmann *et al.*, 2004). Although HR-mediated repair is considered to be a relatively error-free repair process, it can also lead to LOH or genomic rearrangements (Bishop and Schiestl, 2001; Rodgers and McVey, 2016). In addition, a study has demonstrated that other anthracycline drug, doxorubicin, also induces a substantial increase in the frequency of large deletions *in vitro* (Anderson *et al.*, 1991). Therefore, it is suggested that DNA topo II-mediated chromosomal breakage caused by daunorubicin may partly explains the acquisition of large-scale mutations, as demonstrated by higher induced MF at the *TK* locus compared to induced MF at *HPRT* locus at equivalent doses of daunorubicin.

TK6 cells transduced with full length RUNX1/ETO (TK6 *RUNX1/ETO*) were used to investigate the influence of RUNX1/ETO expression on the acquisition of mutations, both spontaneously and after chemotherapy treatment. Evidences suggest that RUNX1/ETO suppresses cellular DNA repair activity to promote mutagenesis and facilitate the acquisition of secondary mutations (Alcalay *et al.*, 2003; Krejci *et al.*, 2008; Araten *et al.*, 2013). One important DNA repair protein implicated in t(8;21) AML pathogenesis is OGG1, which is a key component in BER pathway that initiates repair of oxidised lesions, including 8-oxoG (Alcalay *et al.*, 2003; Liddiard *et al.*, 2010). OGG1 protein expression is downregulated in RUNX1/ETO expressing cells and the depletion of RUNX1/ETO via RNA interference increased *OGG1* transcript, indicating *OGG1* as a likely RUNX1/ETO target gene (Forster *et al.*, 2015). Subsequent studies confirmed that the RUNX1/ETO fusion protein binds theOGG1 promoter and negatively regulates transcription (Forster et al, 2015), which provides a plausible mechanism of mutagenesis driven by the RUNX1/ETO fusion protein. In addition, a study in Drosophila showed that RUNX1/ETO expression increased the levels of intracellular ROS (Sinenko *et al.*, 2010), which may induce elevated DNA damage, making cells that express RUNX1/ETO more susceptible to additional genetic mutations. This is consistent with evidence in this study which showed that RUNX1/ETO sensitises cells to the acquisition of mutations in the absence of an exogenous exposure, as

demonstrated by an increased spontaneous mutation frequency at both the *TK* and *HPRT* loci relative to RUNX1/ETO-negative cells. However, there was no strong evidence that RUNX1/ETO expression renders cells more susceptible to mutagenesis after exposure to equitoxic doses of daunorubicin; a result that is inconsistent with previously reported findings (Forster *et al.*, 2015). This inconsistent finding might be due to the very high mutation burden in cells after exposure to highly cytotoxic doses, thus mitigating any increase on mutation due to RUNX1/ETO expression.

The influence of RUNX1/ETO expression on daunorubicin mutagenicity was further investigated to determine the frequency of each class of mutation (i.e. large deletion, small deletion/insertion, base substitution) induced by this agent at the *HPRT* locus. Two *HPRT* mutant spectrums were generated, for RUNX1/ETO-negative and RUNX1/ETO-positive cells, to determine the effect of fusion oncoprotein expression on spontaneous mutation and chemotherapy-induced mutation. The daunorubicin-induced mutational spectrum for both RUNX1/ETO-negative and RUNX1/ETO-positive *HPRT* mutant populations are dominated by mutants originating from DNA strand breakage events (large deletions), consistent with the inhibitory effect of topo II enzyme as a major mechanism of this agent. In fact, an almost complete lack of base substitution mutants (base substitution; 2%) suggests that daunorubicin is efficient at inducing mutations derived from DNA strand breaks but that mutations arising from base mispairing are rare. Although a slightly higher proportion of *HPRT* internal exon deletions was observed in daunorubicin-treated TK6 *RUNX1/ETO* cells compared to daunorubicin-treated parental TK6 cells, further characterisation of these deletion revealed no notable differences in the structural type of internal exon deletions. Taken together, these data provide strong evidence that daunorubicin almost exclusively induces mutations derived from DNA strand breaks, likely resulting from the inhibition of topo II.

The influence of RUNX1/ETO on the acquisition of mutations was further evaluated by interrogating the spontaneous mutational spectrum to determine whether fusion protein expression is association with acquisition of a particular mutation type. In contrast to the daunorubicin-induced mutational spectrum, which is dominated by deletions, the frequency for each class of mutation observed in TK6 cells is similar with the reported frequency for spontaneous TK6 *HPRT*-mutations in other studies, comprising 45% large deletions, 30% base substitutions, 15% small deletions and 10% compound alterations (Giver Cynthia *et al.*, 1993), although the proportion of large deletion identified in the current study was slightly higher. The base substitution

mutation spectrum analysis is the most anticipated in this study since many mutations in AML relevant genes carry this type of mutation in primary AML. All base substitutions observed in both *HPRT* mutant populations were single base substitutions. In the absence of an exogenous exposure, such as daunorubicin, base substitutions occurred spontaneously at a moderate frequency, although the frequency was slightly lower than the 30% frequency reported previously (Giver Cynthia *et al.*, 1993). One of the assumptions prior to analysis of spontaneously-occurring mutants was that RUNX1/ETO expression would increase the frequency of base substitutions, given the fact that downregulation of OGG1 is predicted to lead to an increase of 8-oxoG and cause elevated G:C to T:A mutation. However, there was no strong evidence that *RUNX1/ETO* expression resulted in a significant increase in base substitution mutations specifically, despite an overall increase in mutation frequency associated with expression of the fusion oncoprotein. However, RUNX1/ETO expression was associated with an increase in T > G / A > C, T > C / A > G and C > A / G > T *HPRT* mutations, which might suggest the preference DNA base/basepair for RUNX1/ETO mutagenicity. Given the heterogeneity in base substitution identified in this study, a thorough analysis of the base substitution spectrum will be described in Chapter 4, in order to identify the mutational pattern associated with RUNX1/ETO expression.

In conclusion, *in vitro* evidence suggests that daunorubicin and ara-C are mutagenic. However, there was no strong evidence that RUNX1/ETO significantly affect the chemotherapy-induced mutation frequency after exposure to daunorubicin and ara-C. This observation is on contrast to published data showing that RUNX1/ETO increases mutagenicity at the the *PIGA* locus in TK6 cells (Forster *et al.*, 2015). One possible explanation for this is that the *PIGA* locus is more sensitive than either TK or HPRT to those types mutations induced by either daunorubicin or ara-C. Indeed, this is almost certainly the case for the *TK* locus, which is recognised as a poor reporter gene for large deletions. Furthermore, the PIGA locus is generally more mutable than either the *TK* or *HPRT* loci and could be generally a more sensitive reporter gene for mutation. Despite a low number of characterised *HPRT* mutants, daunorubicin exclusively induces deletions derived from DNA strand break. However, a more extensive molecular analysis of spontaneous *HPRT* mutants is required to expand the current dataset so that the types of mutation induced by RUNX1/ETO can be fully delineated. Moreover, the data from this study may form the basis for future studies to investigate the mutagenic potential of other chemotherapy agents used in AML, as well

as the influence of *RUNX1/ETO* fusion gene on the acquisition of mutations, both spontaneously and after chemotherapy treatment.

3.3.1. Summary of chapter

In summary, these investigations have demonstrated that:

- Exposure to daunorubicin and ara-C results in a significant dose-dependent increase in mutation frequency at both *TK* and *HPRT* loci in both TK6 and TK6 *RUNX1/ETO* cells, suggesting both chemotherapy agents are mutagenic to DNA *in vitro*.
- RUNX1/ETO increases spontaneous mutation frequency at both *TK* and *HPRT* loci relative to RUNX1/ETO-negative cells, although it was not statistically significant at the *HPRT* locus.
- There was no strong evidence that RUNX1/ETO significantly affect the chemotherapy-induced mutation frequency following exposure to daunorubicin and ara-C.
- Large deletions account for most of daunorubicin-treated mutations, consistent with the inhibitory effect of topo II enzyme which resulted in DNA strand breaks.

Chapter 4. Identifying RUNX1/ETO Mutational Signature in Acute Myeloid Leukaemia

4.1. Introduction

Cancer is a genetic disease characterised by the accumulation of mutations in the somatic genome acquired during disease progression. Although somatic mutations are known to cause cancer, the biological processes generating these mutations is poorly understood. It is thought that the collective somatic mutations observed in a cancer genome are the consequence of several mutational/mutagenic processes that have been operative throughout the lifetime of a patient (Helleday *et al.*, 2014). Each mutational process comprises both a DNA damage component (endogenous or exogenous DNA damage) and a DNA repair or replicative component (defective DNA repair pathways) which leave a characteristic imprint on the cancer genome that reflect those processes. This characteristic imprint, also known as 'mutational signature' is represented by specific mutations in the genome that can include base substitutions, insertions and deletions or structural variations (genomic rearrangements). Previous studies have shown that a certain type of DNA damage has its own preference for specific nucleotides. The high frequency of G:C > T:A transversion found in smoking-associated lung cancer has been attributed to the mutagenic action of polycyclic aromatic hydrocarbons (PAH) compounds in cigarette smoke (Pfeifer and Hainaut, 2003) and single C:G > T:A transition and tandem double CC:GG > TT:AA transition are associated with UV DNA damage in skin cancer (Brash *et al.*, 1991; Pierceall William *et al.*, 1991; Weihrauch *et al.*, 2002; Brash, 2015).

However, it would be naïve to define the cause of cancer by just a single mutational process since cancer aetiology is often multifactorial. The numerous processes operating during cancer development suggests that cancers genomes contain two or more mutational signatures. Mutational signatures in human cancers are based on six classes of base substitutions: C > A, C > G, C > T, T > A, T > C and T > G (all substitutions are referred to by the pyrimidine of the mutated Watson–Crick base pair) which can be further classified by the triplet sequence in which they are positioned with the mutated base residing at the central base position of 96 possible trinucleotide sequences (Figure 4.1) (Nik-Zainal *et al.*, 2012; Alexandrov *et al.*, 2013a; Alexandrov *et al.*, 2013b; Alexandrov *et al.*, 2015). This 96 substitution classification is particularly useful for distinguishing mutational signatures that cause the same substitutions but in different sequence contexts. To date, 30 mutational signatures have been identified across 40 cancer types with at least two mutational signatures found in each cancer type and some cancer types with more than 10 mutational

signatures (Figure 4.2) (https://cancer.sanger.ac.uk/cosmic/signatures). Moreover, some of the mutational signatures are found in multiple cancer types, while others are confined to a single cancer type. The identification of these mutational signatures has provided essential information on the aetiology of mutagenesis although the causes of many of these remain unknown. Examples of mutational signatures that have been attributed to possible aetiology are shown in Figure 4.1.

AML is a complex disease comprises of different subtypes, but contains relatively few somatic mutations compared to other cancers (Cancer Genome Atlas Research, 2013). Two mutational signatures have been found in AML (Figure 4.2). The first signature (signature 1) is characterised by an increase of C > T mutations and is associated with endogenous mutational process initiated by spontaneous deamination of 5-methylcytosine, while the second signature (signature 5) is dominated by C > T and T>C mutations, but yet has unknown aetiology (https://cancer.sanger.ac.uk/cosmic/signatures). However, it is thought that other biological processes may also influence the mutational pattern in AML. A recent study has showed that ara-C which is mainly used in AML remission induction treatment preferentially induces mutation at central base position of $^{5'}$TpGpA$^{3'}$/$^{5'}$TpCpA$^{3'}$ sequences in the genome which are significantly higher in relapsed AML patients after exposure to ara-C-containing regimens (Fordham *et al.*, 2015). The finding from this study demonstrated that specific sequences in the genome were susceptible to additional exogenous mutagenic exposures by chemotherapy agents which might contribute to the aetiology of relapse disease. It is important to explore patterns of mutations induced by specific mutagens since various endogenous and exogenous sources of DNA damage are involved in AML pathogenesis. RUNX1/ETO is one of the important initiating genetic lesions in AML and was demonstrated to increase mutation frequency both spontaneously and after exposure to genotoxic chemotherapeutic agents (Forster *et al.*, 2015). This finding is further supported by the data presented in Chapter 3 of this thesis demonstrating elevated mutagenicity in RUNX1/ETO-expressing cells. Using the same approaches as Fordham and colleagues (Fordham *et al.*, 2015), the base substitution data described in previous chapter (see Table 3.4, Chapter 3) was further investigated in order to identify a spontaneous mutational signature associated with RUNX1/ETO expression in order to better understand the mutational processes contributing to disease progression.

4.1.1. Aims of Chapter 4

The objective of this investigation is to identify a somatic mutational signature associated with RUNX1/ETO expression and to determine whether this oncoprotein predisposes to mutation at specific sequences in the genome.

Specifically, the experimental aims were as follows:

- Compare the mutational pattern between RUNX1/ETO-positive and RUNX1/ETO-negative clones at the *HPRT* locus.
- Compare the RUNX1/ETO-induced mutational pattern observed *in vitro* using the cell model system and that seen in primary AML.
- Identify nucleotide sequences susceptible to RUNX1/ETO-induced mutation.
- Identify “leukaemia genes” susceptible to RUNX1/ETO-induced mutation in AML.

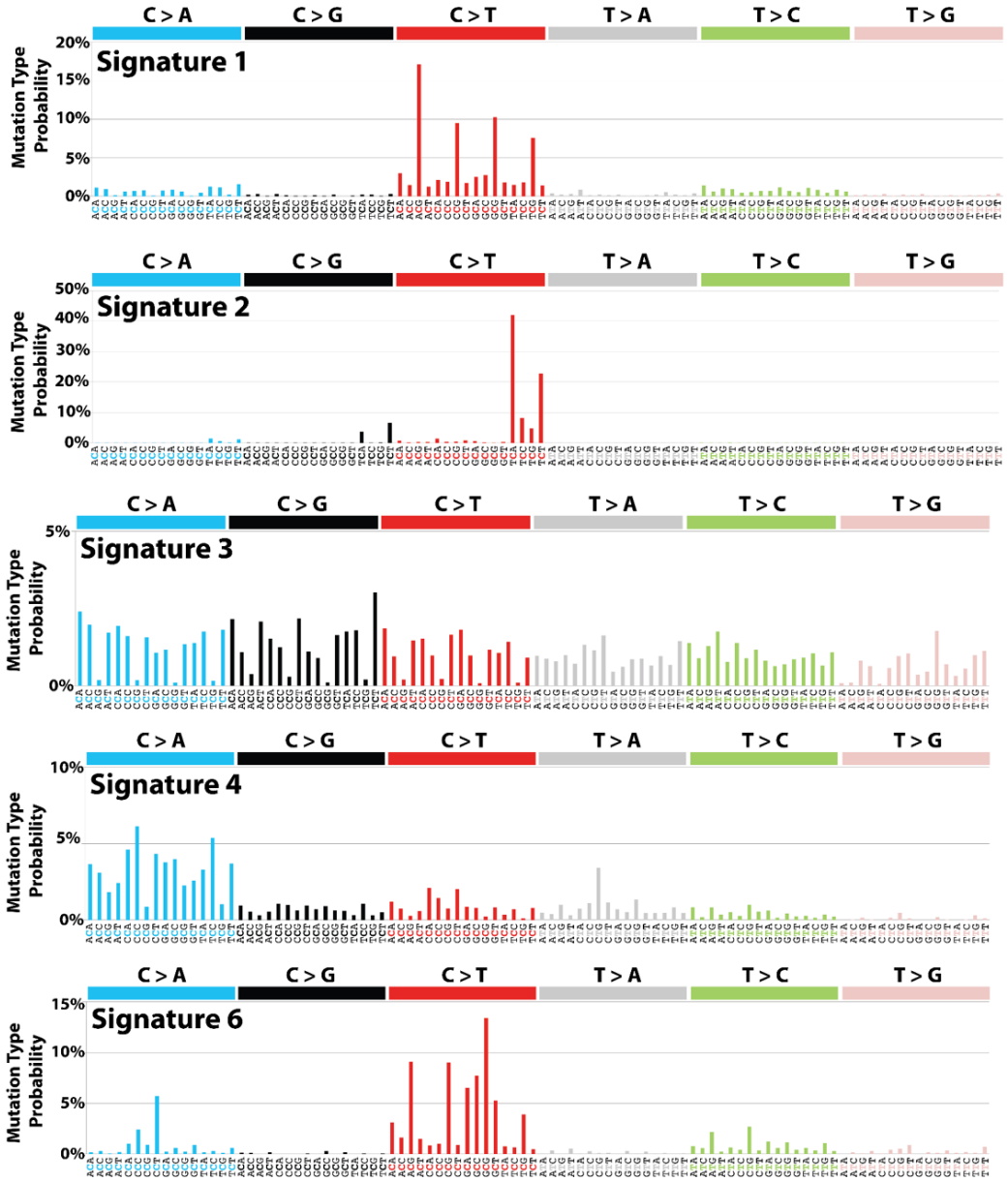


Figure 4.1. Example of mutational signatures found in human cancers that have been attributed to proposed aetiology (https://cancer.sanger.ac.uk/cosmic/signatures).

The signature is displayed according to the 96 substitution classification defined by the substitution class and sequence context immediately 5' and 3' to the mutated base.

Signature 1: spontaneous deamination of 5-methylcytosine.

Signature 2: activity of the AID/APOBEC family of cytidine deaminases.

Signature 3: failure of DNA double-strand break repair by homologous recombination.

Signature 4: exposed to tobacco carcinogens (e.g., benzo[a]pyrene).

Signature 6: defective DNA mismatch repair.

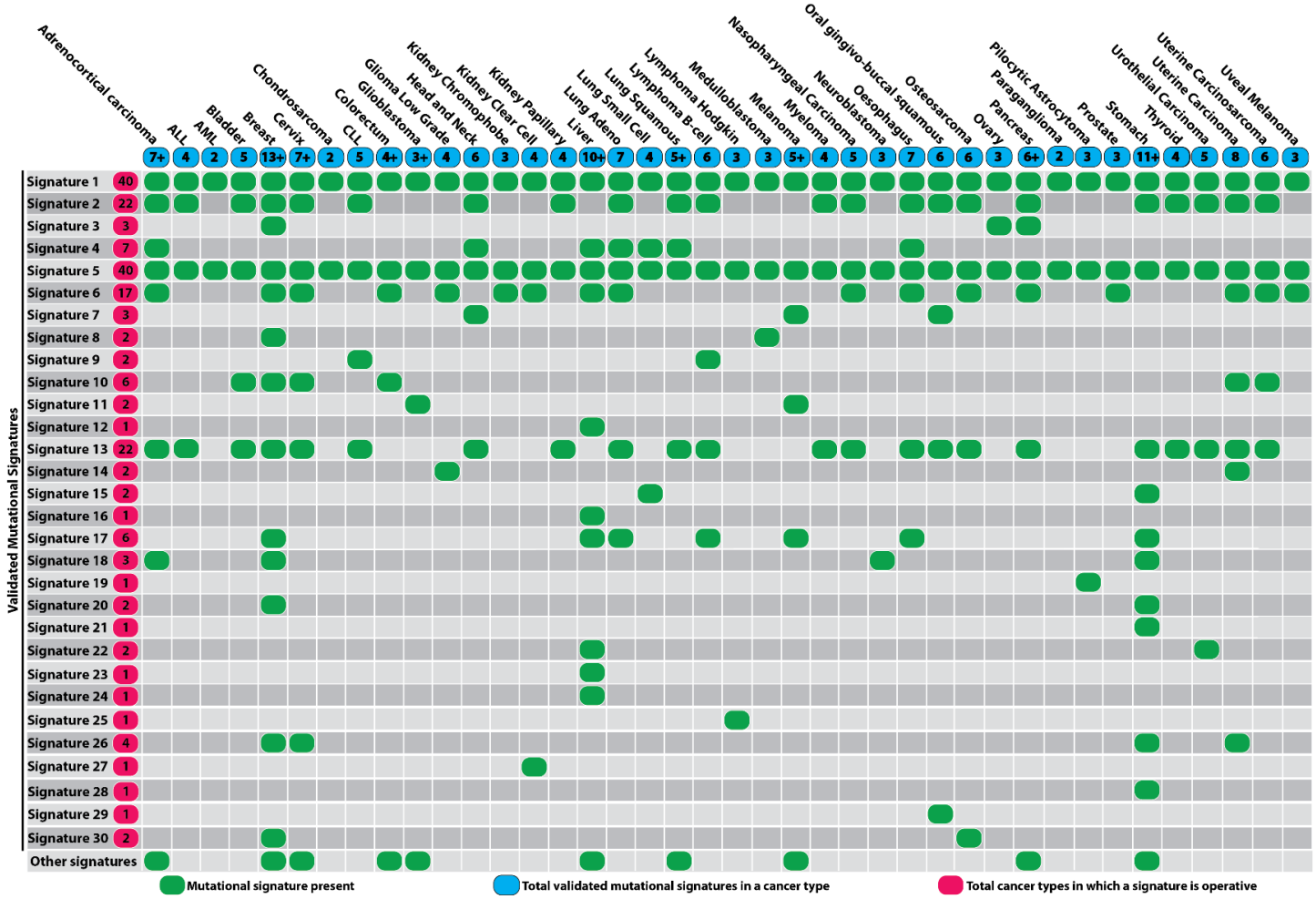


Figure 4.2. The presence of mutational signatures identified across 40 human cancer types (https://cancer.sanger.ac.uk/cosmic/signatures).

Two mutational signatures have been associated with AML (signature 1 and signature 5).

4.2. Results

4.2.1. Base substitution in TK6 and TK6 RUNX1/ETO at the HPRT locus

The single base substitutions identified in TK6 and TK6 RUNX1/ETO cells (see Table 3.4, Chapter 3) were classified into six classes of mutation: C > A / G > T, C > G / G > C, C > T / G > A, T > A / A > T, T > C / A > G and T > G / A > C. In total, 21 and 49 independent single base substitutions were derived from *HPRT* mutant clones of TK6 and TK6 RUNX1/ETO cells, respectively. In proportion, the frequency of T > G / A > C, T > C / A > G and C > A / G > T base substitutions was higher in TK6 RUNX1/ETO compared to parental TK6; however, none of these were statistically significant (Figure 4.3). Regarding type of mutation, in proportion, transversion mutations occurred at a slightly higher frequency in TK6 RUNX1/ETO compared to parental TK6 [55% (27/49) vs 43% (9/21)], but this was also not statistically significant (*P* = 0.44; two-tailed Fisher's exact test) (Figure 4.4).

4.2.2. Mutational pattern between TK6 and TK6 RUNX1/ETO at HPRT locus

The base substitution mutation data were further analysed in sequence context to identify any mutational pattern associated with RUNX1/ETO expression. The six classes of base substitution were grouped according to mutation at the central base position of 16 different possible trinucleotide sequences, result in 96 possible trinucleotide sequences for all 6 classes of mutation (Figure 4.5). Regarding T > G / A > C base substitutions, 5 of 13 (39%) occurred at $^{5'}$ApTpA$^{3'}$/$^{5'}$TpApT$^{3'}$ sequences in RUNX1/ETO clones, compared to 0 of 1 (0%) in parental TK6 cells (*P* = 1.00; two-tailed Fisher's exact test).

The 96 possible trinucleotide sequences were then combined and used to generate a heatmap in order to identify the affected trinucleotide sequences regardless of type of base substitution (Figure 4.6). Strikingly, 7 of 49 (14.3%) base substitutions from RUNX1/ETO-positive clones affected the central T:A base-pair in $^{5'}$ApTpA$^{3'}$/$^{5'}$TpApT$^{3'}$ sequences compared to 1 of 21 (4.8%) in RUNX1/ETO-negative cells; however, this was not statistically significant (*P* = 0.42; two-tailed Fisher's exact test). In addition, mutation frequency at the other 31 possible trinucleotide sequences in RUNX1/ETO-positive mutational spectrum were also not substantially different compared to fusion-negative controls.

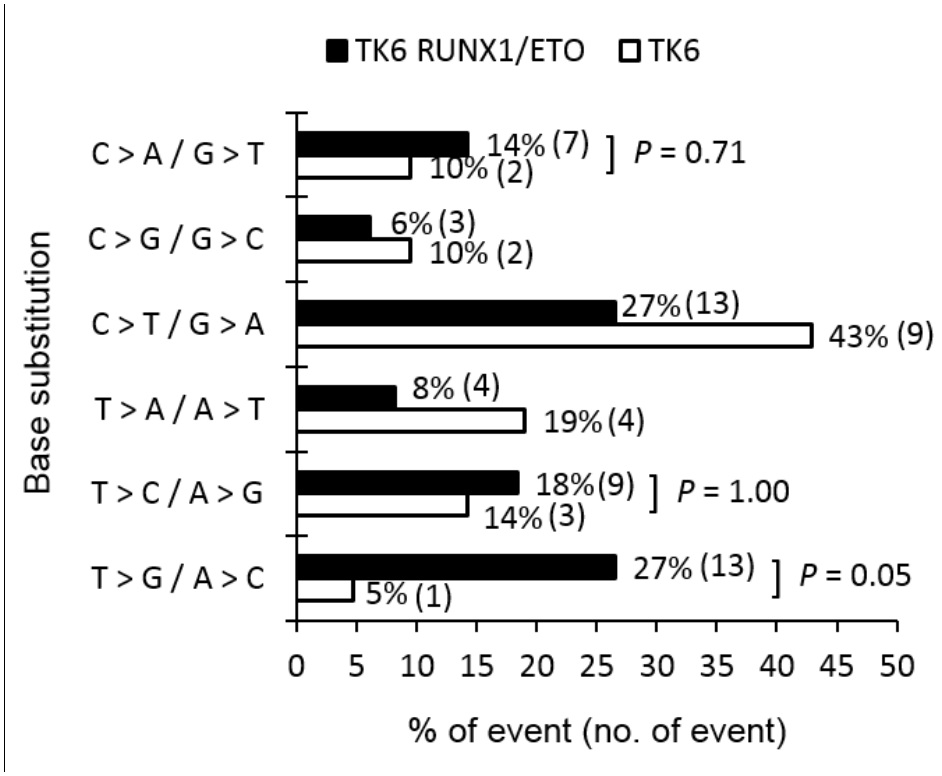


Figure 4.3. Number of single base substitution identified in TK6 and TK6 RUNX1/ETO at *HPRT* locus.

Single base substitutions were classified into six classes of mutation. The number of events (and proportion) from the total number of mutations observed in TK6 and TK6 RUNX1/ETO clones is shown. Figure shown in parenthesis is the number of mutations identified in each class of mutation for TK6 and TK6 RUNX1/ETO cells. P-value was calculated by two-tailed Fisher's exact test comparing the number of mutation between TK6 and TK6 RUNX1/ETO cells.

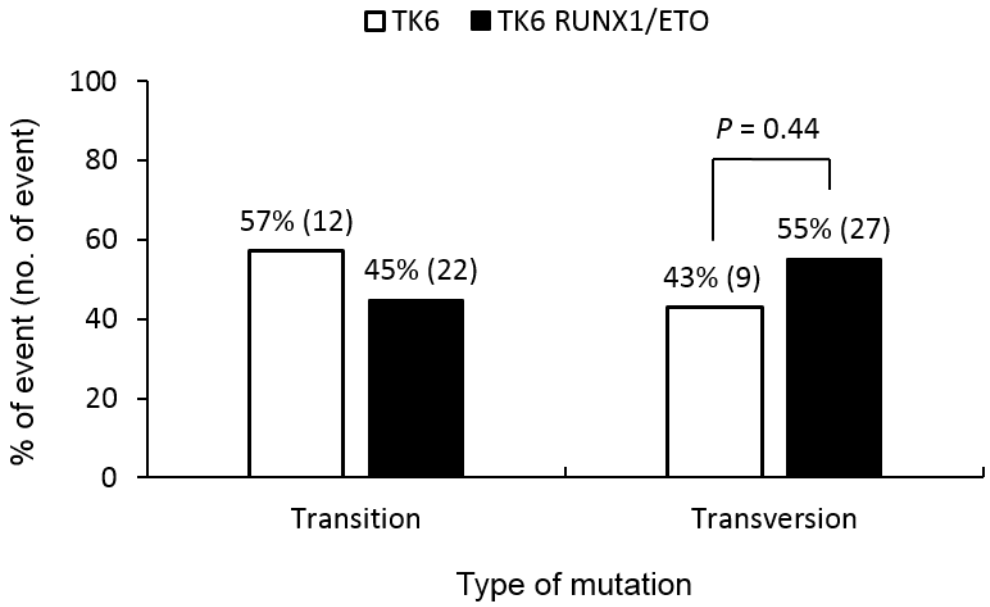


Figure 4.4. Type of base substitution mutation.

Single base substitutions were classified according to type of mutation (transition or transversion). The number of events (and proportion) from the total number of mutations observed in TK6 and TK6 RUNX1/ETO clones is shown. Figure shown in parenthesis is the number of mutations identified in each class of mutation for TK6 and TK6 RUNX1/ETO cells. P-value was calculated by two-tailed Fisher's exact test comparing the number of mutation between TK6 and TK6 RUNX1/ETO cells.

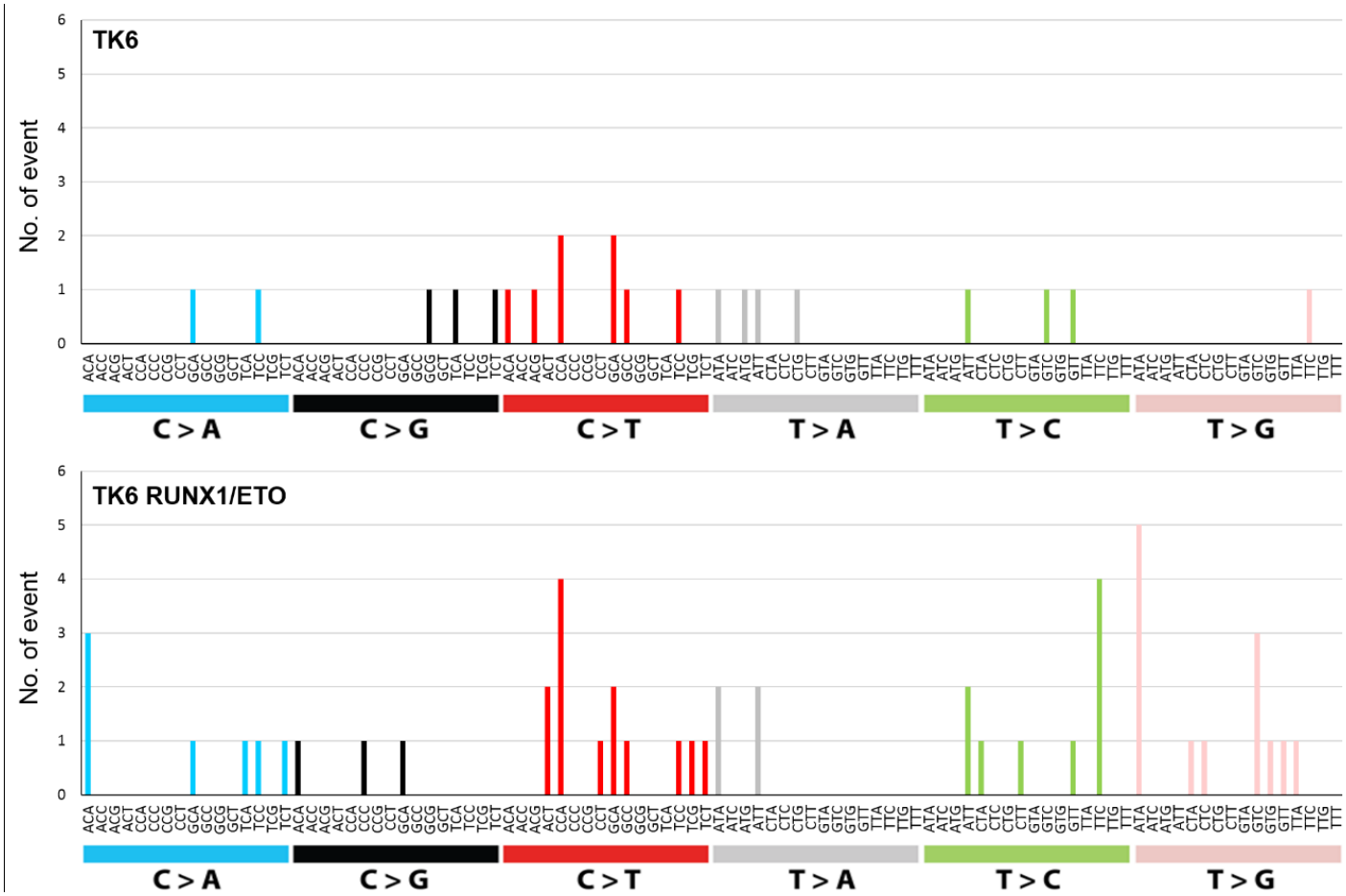


Figure 4.5. Comparison of *HPRT* base substitution mutational pattern between TK6 and TK6 RUNX1/ETO.

The mutation at central base position is displayed according to the 96 substitution classification defined by the substitution class and sequence context immediately 5' and 3' to the mutated base (trinucleotide sequence). Number of event represents number of mutation at the central base position of each trinucleotide sequence.

TK6

5' base		A				C				G				T			
Central base	A	0	1	0	2	0	0	1	1	1	1	0	0	0	0	0	1
	C	1	0	2	0	2	0	0	0	3	1	1	0	1	2	0	0
	G	0	0	0	0	0	1	0	2	2	1	0	0	1	3	2	1
	T	1	0	1	2	0	0	1	0	0	1	0	1	0	1	0	0
3' base		A	C	G	T	A	C	G	T	A	C	G	T	A	C	G	T

TK6 RUNX1/ETO

5' base		A				C				G				T			
Central base	A	0	2	1	4	0	1	0	0	4	3	1	0	1	0	2	7
	C	4	0	0	2	4	1	0	1	4	1	0	0	1	2	1	2
	G	2	0	1	2	1	0	0	0	2	1	1	0	1	4	4	4
	T	7	0	0	4	2	1	0	1	0	3	1	2	1	4	0	0
3' base		A	C	G	T	A	C	G	T	A	C	G	T	A	C	G	T

0 7

Figure 4.6. Mutation at central base in each possible trinucleotide.

The sequence context of *HPRT* base substitution mutations in TK6 and TK6 RUNX1/ETO clones. Heatmaps display the number of base substitution mutations at the central base position of each trinucleotide sequence.

4.2.3. Mutational pattern identified in next generation sequencing data of primary RUNX1/ETO AML patients

To identify the base substitution mutational associated with RUNX1/ETO in primary AML cells, interrogation of available targeted amplicon sequencing data from 56 diagnostic AML t(8;21) samples (Hartmann *et al.*, 2016) was performed. A total of 86 single base substitutions in 19 leukaemia relevant genes were analysed (Table 4.1). Of these, 56% (48/86) were transitions and 44% (38/86) were transversions. In term of type of base substitution, C > T / G > A transitions occurred at a higher frequency (38/ 86, 44.2%) than the other five classes of base substitution (Figure 4.7). Regarding gene context, there were 16 of 19 genes affected by C > T / G > A transitions, with *ZBTB7A* (7/38) and *NRAS* (6/38) were the genes mostly frequently affected by this mutation type (Table 4.1).

Using the same approach as used to analyse RUNX1/ETO-induced mutations at the *HPRT* locus *in vitro* (Section 4.2.2), base substitutions were grouped according to mutation at the central base position of 16 possible trinucleotide sequences for each class of single base substitution (Figure 4.8). Of the most frequent class of mutations were C > T / G > A transitions, which commonly occurred at the central base position of $^{5'}$CpCpG$^{3'}$ / $^{5'}$CpGpG$^{3'}$ (10/38, 26.3%) and $^{5'}$GpCpG$^{3'}$ / $^{5'}$CpGpC$^{3'}$ (8/38, 21.1%) sequences. Other notable trinucleotide sequences affected by C > T / G > A transitions at central base position were $^{5'}$ApCpC$^{3'}$/$^{5'}$GpGpT$^{3'}$ (5/38, 13.2%) and $^{5'}$TpCpG$^{3'}$/$^{5'}$CpGpA$^{3'}$ (4/38, 10.5%). Genes that were found affected specifically by C > T / G > A transitions at C:G basepairs in $^{5'}$CpCpG$^{3'}$ / $^{5'}$CpGpG$^{3'}$ sequences were *ZBTB7A*, *DNMT3A*, *SMC1A*, *PTPN11* and *IDH2* whereas *ZBTB7A*, *ASXL1*, *TET2*, *DNMT3A*, *FAT1* and *WT1* were affected by C > T / G > A transitions at C:G base pair in $^{5'}$GpCpG$^{3'}$ / $^{5'}$CpGpC$^{3'}$ sequences (Figure 4.8). In contrast, only the *NRAS* gene was affected by C > T / G > A transitions at C:G basepairs in $^{5'}$ApCpC$^{3'}$ / $^{5'}$GpGpT$^{3'}$.

Regardless of the type of base substitution, $^{5'}$CpCpG$^{3'}$ / $^{5'}$CpGpG$^{3'}$ sequences were the most commonly mutated (11/86, 12.8%), followed by $^{5'}$GpCpG$^{3'}$ / $^{5'}$CpGpC$^{3'}$ sequences as the second most affected (8/86, 9.3%) (Figure 4.9). Other notable affected sequences were $^{5'}$ApCpC$^{3'}$/$^{5'}$GpGpT$^{3'}$ (7/86, 8.1%), $^{5'}$ApCpA$^{3'}$/$^{5'}$TpGpT$^{3'}$ (6/86, 7.0%), $^{5'}$GpTpC$^{3'}$ / $^{5'}$GpApC$^{3'}$ (6/86, 7.0%) and 5'TpCpT3' / 5'ApGpA3' (6/86, 7.0%). Interestingly, it was observed that most of C > T / G > A transitions occurred at the cytosine (C) bases of $^{5'}$CpG$^{3'}$ dinucleotide sequences (23/38, 60.5%), with the most affected gene was *ZBTB7A* (7/23, 30.4%) (Figure 4.8) (Table 4.2).

Gene	Transition		Total no. of transition	Transversion				Total no. of transversion	Total no. of mutation
	C > T / G > A	T > C / A > G		C > A / G > T	C > G / G > C	T > A / A > T	T > G / A > C		
ZBTB7A	7		7	2			3	5	12
NRAS	6	2	8	3				3	11
KIT		1	1		2	4	3	9	10
FLT3		1	1	1	3	2	2	8	9
TET2	4	1	5	1	1		1	3	8
DNMT3A	3		3		1	1		2	5
JAK2	1	2	3	2				2	5
FAT1	3		3		1			1	4
ASXL2	1		1	1			1	2	3
SMC3	1		1	1			1	2	3
IDH2	2		2					0	2
JAK3	1	1	2					0	2
KRAS	2		2					0	2
PTPN11	2		2					0	2
RAD21	1		1	1				1	2
SMC1A	2		2					0	2
WT1	1	1	2					0	2
ASXL1	1		1					0	1
CBL		1	1					0	1
Total no. of mutation	38	10	**48**	12	8	7	11	**38**	**86**

Table 4.1. Single base substitution mutations identified in t(8;21) AML at 19 leukaemia relevant genes.

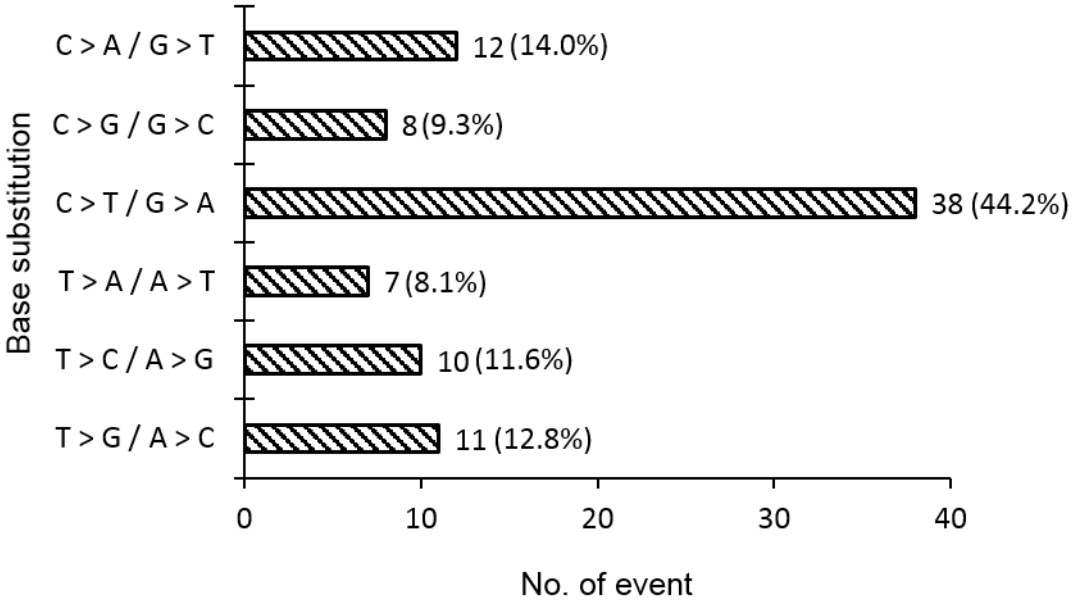


Figure 4.7. Number of single base substitution mutations identified in t(8;21) AML at 19 leukaemia relevant genes.

Single base substitutions were classified into six types of mutation. Number of event represents the number of mutations identified in t(8;21) AML samples. Figure shown in parenthesis is the percentage of mutations from the total number of mutations identified in t(8;21) AML patients.

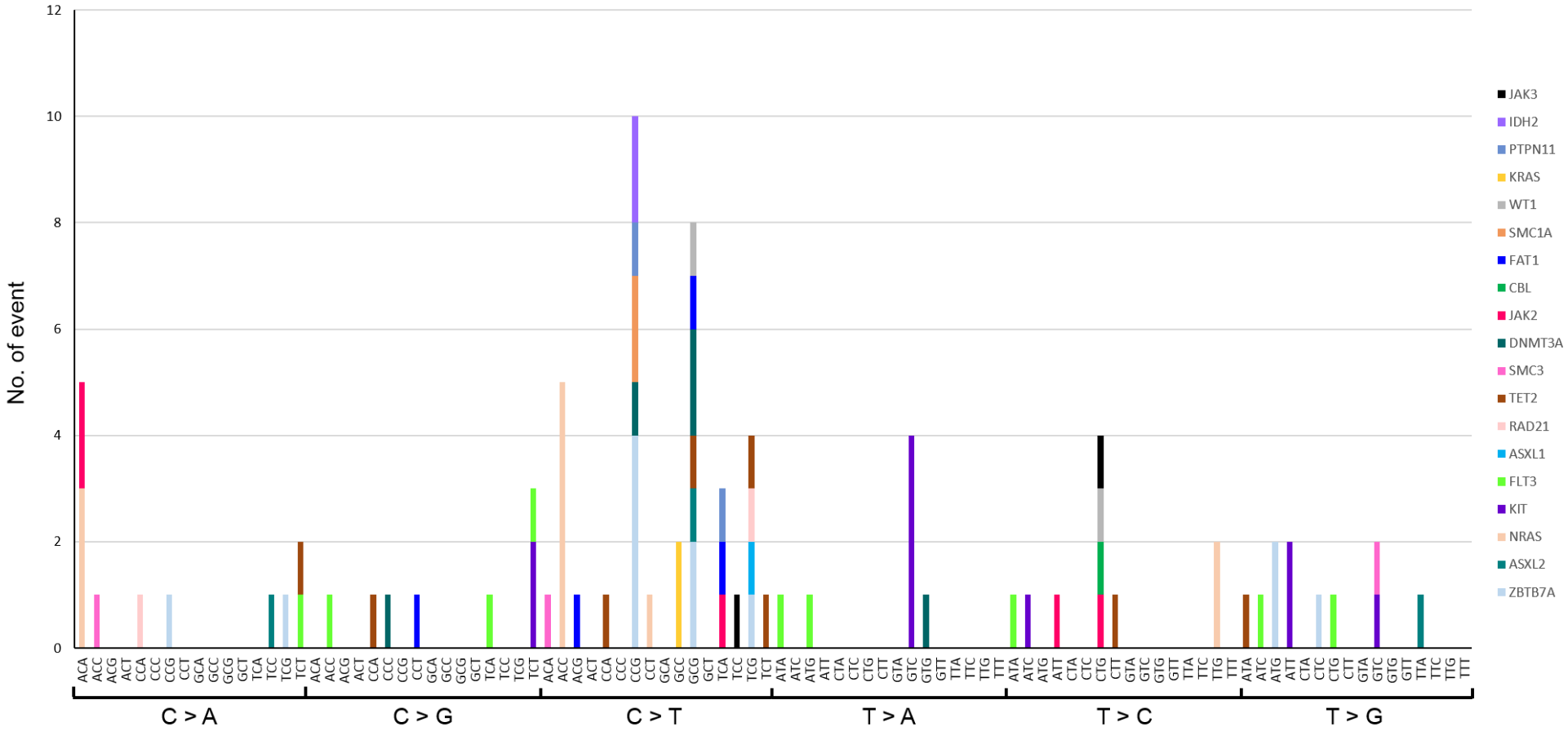


Figure 4.8. Base substitution mutation pattern identified in t(8;21) AML patients at 19 leukaemia relevant genes.

The mutation at central base position is displayed according to the 96 substitution classification defined by the substitution class and sequence context immediately 5' and 3' to the mutated base (trinucleotide sequence). Number of event represents number of mutation at the central base position of each trinucleotide sequence.

5‘ base		A				C				G				T			
Central base	A	0	0	1	3	2	1	5	3	0	6	1	2	1	0	0	3
	C	6	7	1	0	3	1	11	2	0	2	8	0	4	2	5	6
	G	6	0	2	0	5	8	11	1	2	2	1	7	4	0	3	6
	T	3	2	3	3	0	1	5	1	0	6	1	0	1	0	2	0
3‘ base		A	C	G	T	A	C	G	T	A	C	G	T	A	C	G	T

Figure 4.9. Mutation at the central base in each possible trinucleotide based on data derived from t(8;21) AML leukaemias and mutation from 18 leukaemia relevant genes.

The sequence context of base substitution mutations in t(8;21) AML patients. Heatmap displays the number of base substitution mutations at the central position.

Trinucleotide Sequence	No. of event	Genes
$^{5'}$ApCpA$^{3'}$ / $^{5'}$TpGpT$^{3'}$	1	*SMC3* (1)
$^{5'}$ApCpC$^{3'}$ / $^{5'}$GpGpT$^{3'}$	5	*NRAS* (5)
$^{5'}$ApCpG$^{3'}$ / $^{5'}$CpGpT$^{3'}$	1	*FAT1* (1)
$^{5'}$ApCpT$^{3'}$ / $^{5'}$ApGpT$^{3'}$	0	-
$^{5'}$CpCpA$^{3'}$ / $^{5'}$TpGpG$^{3'}$	1	*TET2* (1)
$^{5'}$CpCpC$^{3'}$ / $^{5'}$GpGpG$^{3'}$	0	-
$^{5'}$CpCpG$^{3'}$ / $^{5'}$CpGpG$^{3'}$	10	*ZBTB7A* (4), *DNMT3A* (1), *SMC1A* (2), *PTPN11* (1), *IDH2* (2)
$^{5'}$CpCpT$^{3'}$ / $^{5'}$ApGpG$^{3'}$	1	*NRAS* (1)
$^{5'}$GpCpA$^{3'}$ / $^{5'}$TpGpC$^{3'}$	0	-
$^{5'}$GpCpC$^{3'}$ / $^{5'}$GpGpC$^{3'}$	2	*KRAS* (2)
$^{5'}$GpCpG$^{3'}$ / $^{5'}$CpGpC$^{3'}$	8	*ZBTB7A* (2), *ASXL1* (1), *TET2* (1), *DNMT3A* (2), *FAT1* (1), *WT1* (1)
$^{5'}$GpCpT$^{3'}$ / $^{5'}$ApGpC$^{3'}$	0	-
$^{5'}$TpCpA$^{3'}$ / $^{5'}$TpGpA$^{3'}$	3	*JAK2* (1), *FAT1* (1), *PTPN11* (1)
$^{5'}$TpCpC$^{3'}$ / $^{5'}$GpGpA$^{3'}$	1	*JAK3* (1)
$^{5'}$TpCpG$^{3'}$ / $^{5'}$CpGpA$^{3'}$	4	*ZBTB7A* (1), *ASXL1* (1), *RAD21* (1), *TET2* (1)
$^{5'}$TpCpT$^{3'}$ / $^{5'}$ApGpA$^{3'}$	1	*TET2* (1)
Total no. of event	**38**	

Table 4.2. Trinucleotide sequences in leukaemia relevant genes affected by C > T / G > A transitions.

Figure shown in parenthesis represents number of C > T / G > A events at specific trinucelotide sequence for each gene. Genes that were affected by mutation at C base of $^{5'}$CpG$^{3'}$ dinucleotide sequences are higlighted in grey.

4.3. Discussion

Single base substitutions constitute the most frequent type of human gene mutation and are a leading cause of genetic disease including cancer. The analysis of base substitution has previously been identified as a useful approach to demonstrate how the mutation spectra can be specific to tumour type and the underlying mutational processes (Helleday *et al.*, 2014). Other types of mutation, such as insertions and deletions (indels) also contribute to disease pathogenesis, but these are technically difficult to assay due to size local sequence, which is often repetitive. The base substitutions included in the analysis presented in this chapter were independently derived from unique *HPRT* mutant clones with just one mutation derived from each CHAT-treated cell population. As such, the base substitution analysis includes only independent mutations after correction for duplicate mutations which likely represent clonal expansions.

Single base substitutions identified in the *HPRT* gene sequence were classified into six classes of mutation based on the affected basepair and whether a transition or transversion mutation. Furthermore, because it is not possible to discern in which strand a mutation was initiated, mutations are grouped by complementary sequence. On the basis of type of mutation, there was more transversions than transitions in RUNX1/ETO-positive cells. In the human genome, transitions are generated at higher spontaneous frequency than transversions somatically, with the latter generally having more deleterious effects on the amino acid sequence (Rubin and Green, 2009). The data generated in this project suggest that an increase in T > G / A > C mutations is largely responsible for the elevated mutation rate in RUNX1/ETO-positive cells. The increased of T > G / A > C mutations suggests that DNA repair pathways especially those repairing damaged or mismatched DNA bases such as BER and MMR might be compromised. A previous study has shown that OGG1 DNA glycosylase, a key component in BER pathway, was downregulated in cells expressing RUNX1/ETO (Forster *et al.*, 2015). However, impaired activity of OGG1 is associated with G:C > T:A transversions, which was not evident at the *HPRT* locus in this study. Downregulation of other key DNA glycosylase in BER pathway such as NEIL1 was also shown to result in base substitution mutations, as characterised by a higher rate of spontaneous A:T > T:A transversions at the *HPRT* locus in Chinese hamster lung fibroblast V79 cells (Maiti *et al.*, 2008); but, this was also not evident at the *HPRT* locus in this study. Although no strong evidence of OGG1 and NEIL1 downregulation in RUNX1/ETO-

expressing cells from this study, investigation on association between RUNX1/ETO and other DNA repair gene expression especially those involved in BER and MMR pathways is warranted, given the potential of these pathways in causing base substitution mutations when disrupted.

In terms of sequence context, the central T:A base pair in $^{5'}$ApTpA$^{3'}$/$^{5'}$TpApT$^{3'}$ sequences were most affected by T > G / A > C transversions. However, there is no previous evidence of $^{5'}$ApTpA$^{3'}$/$^{5'}$TpApT$^{3'}$ susceptibility to mutations in the literature, indicating unknown aetiology. Moreover, the mutational pattern observed in RUNX1/ETO-positive cells did not resemble any of the recognised mutational signatures previously identified in human cancer (https://cancer.sanger.ac.uk/cosmic/signatures). Mutational signature 9 which is predominantly composed of T > G mutation at the central base of $^{5'}$TpTpT$^{3'}$, $^{5'}$CpTpT$^{3'}$, $^{5'}$TpTpA$^{3'}$ and $^{5'}$ApTpA$^{3'}$ trinucleotides sequences is the closest resemblance (Figure 4.10), but caution must be exercised given the limited dataset generated in this study. Mutational signature 9 has been observed in CLL and malignant B-cell lymphoma and is associated with the activity of activation induced deaminase (AID) during somatic hypermutation (https://cancer.sanger.ac.uk/cosmic/signatures); however, no evidence of this mutational signature was previously reported in AML.

Interrogation of available targeted amplicon sequencing data from 56 diagnostic AML t(8;21) samples (Hartmann *et al.*, 2016) revealed that mutation in primary t(8;21) AML was dominated by C > T / G > A transitions, which occurred at a substantially higher frequency than the other five classes of base substitution. Interestingly, it was found that mutation at C bases of $^{5'}$CpG$^{3'}$ dinucleotide sequences dominated the primary t(8;21) AML mutational pattern. CpG dinucleotides are susceptible to spontaneous mutation via spontaneous deamination of 5-methylcytosine to thymine, giving rise to a T:G mispair (Holliday and Grigg, 1993; Wiebauer *et al.*, 1993). T:G mismatches are normally corrected by specific DNA glycosylase during base excision repair (BER) pathway; however, misrepair often occurred as this DNA lesion is difficult to be recognised which can result in the generation of C:G > T:A transition mutations at the next round of DNA replication (Kow, 2002). The CpG dinucleotides are recognised as a mutation hotspot for some cancer types, especially colorectal cancer where there is a predominance of C > T transitions (Sjöblom *et al.*, 2006). The observation of most C > T transition mutation at CpG sequences in a number of genes, including *ZBTB7A*, *DNMT3A*, *SMC1A*, *PTPN11*, *IDH2*, *ASXL1*, *TET2*, *FAT1*, *WT1* and

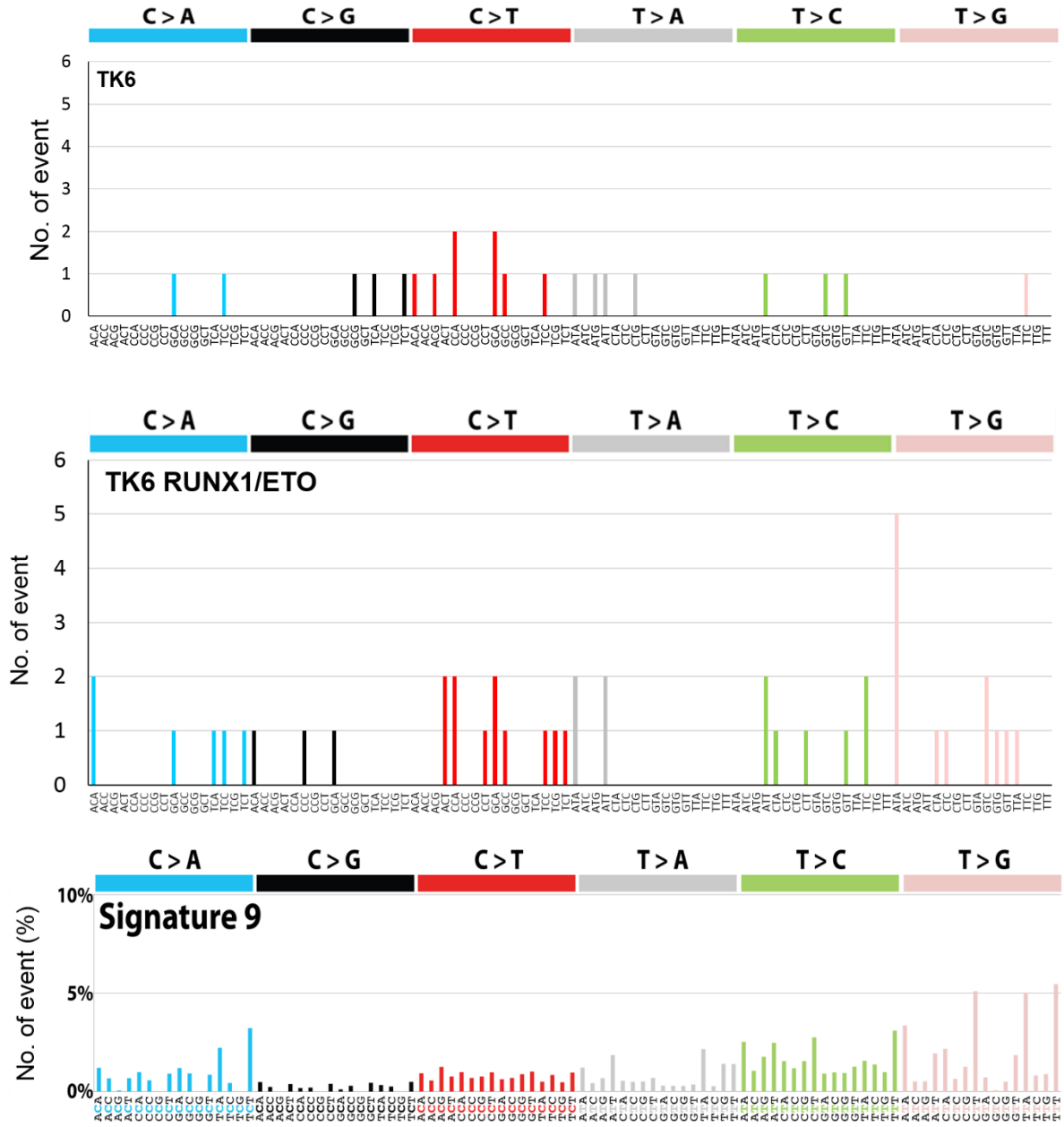


Figure 4.10. Comparison of the *HPRT* mutational signature derived from RUNX1/ETO-positive cells *in vitro* and Signature 9.

Mutation at central base position is defined by the substitution class and sequence context immediately 5' and 3' to the mutated base.

RAD21, in t(8;21) primary AML, suggesting deamination of 5-methylcytosine as a likely source of mutagenesis in this AML sybtype.

Collectively, the mutational pattern observed in t(8;21) primary AML was largely resembles mutational signature 1 (Figure 4.11) (https://cancer.sanger.ac.uk/cosmic/signatures). This mutational signature has been found in all cancer types and is associated with endogenous mutational process initiated by spontaneous deamination of 5-methylcytosine. Increase deamination of cytosine was observed in cells after exposure to DNA damaging agents such as UV, ionizing radiation, sodium bisulfite and intercalating agents such as echinomycin (Chen and Shaw, 1993; Moyer *et al.*, 1993; Pfeifer *et al.*, 2005; d'Ischia *et al.*, 2011). Therefore, it is suggested the mechanisms that increase mutations at CpG sequence in primary t(8;21) AML are caused by the hypermutability of methylated CpG which is vulnerable to mutagenic effect of various endogenous and exogenous mutagens.

In general, the mutational pattern observed in t(8;21) primary AML was different to the mutational pattern observed *in vitro* in RUNX1/ETO-positive cells and there are three possible reasons that could be responsible for this observation. Firstly, it is important to note that the analysis of base substitutions from *in vitro* study were limited to mutations detected solely in *HPRT* cDNA sequence. The *HPRT* coding sequence contains only about 2.3% of CpG dinucleotide sequences, which may not be sufficient to identify a susceptibility to mutations driven by 5-methylcytosine deamination. Secondly, the cell model used for *in vitro* studies is lymphoblastic in origin and could be susceptible to the acquisition of mutations via RAG expression and other B-cell specific mutational processes. This could explain why the *in vitro HPRT* base substitution mutational pattern is consistent with mutational signature 9, more commonly found in lymphoid cancers, instead of other mutational signatures found in myeloid leukaemia. Thirdly, there is a clear selection bias on two levels driving mutagenesis in t(8;21) primary AML: firstly, analysis was limited to selected leukaemia relevant genes which may not be representative of the entire genome; secondly, the mutation spectrum in any human disease, including AML, is driven by functionality rather than local base sequence. For example, *KIT*, *FLT3* and *NRAS* represent the most frequently mutated genes in t(8;21) AML (Goemans *et al.*, 2005; Kohl *et al.*, 2005; Wang *et al.*, 2005; Krauth *et al.*, 2014), and mutations in these genes have been shown to cooperate with RUNX1/ETO to induce AML in murine models (Schessl *et al.*, 2005; Chou *et al.*, 2011; Wang *et al.*, 2011). As such, the mutations observed in t(8;21) primary AML are positively selected because they affect protein function or expression.

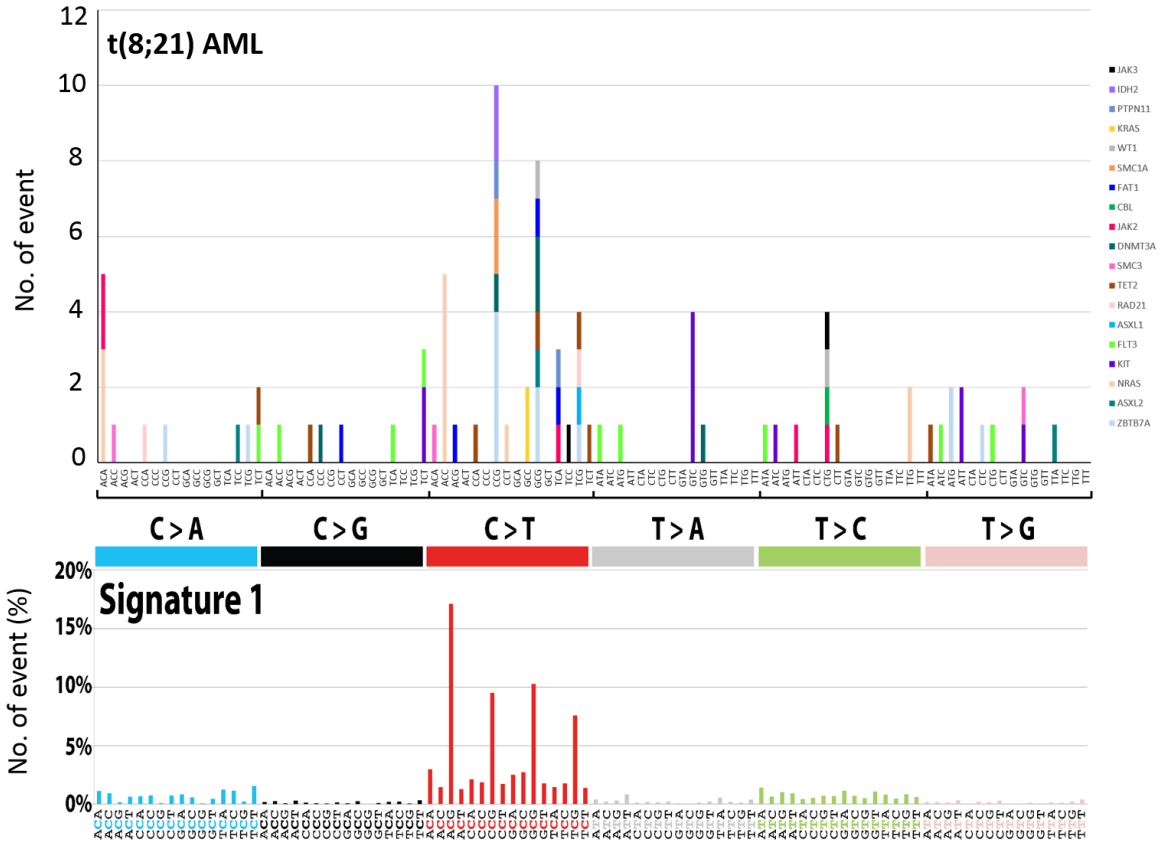


Figure 4.11. Comparison of mutational pattern between t(8;21) primary AML and Signature 1.

The mutation at central base position is displayed by the substitution class and sequence context immediately 5' and 3' to the mutated base.

Despite a lack of supporting evidence from t(8;21) primary AML, $^{5'}$ApTpA$^{3'}$/$^{5'}$TpApT$^{3'}$ sequences appeared to be a putative hotspot for RUNX1/ETO-induced mutation *in vitro*. However, given the low number of characterised *HPRT* mutants, no definite conclusion can be made. Therefore, more *HPRT* mutants need to be generated to expand the current dataset and improve statistical power to discern a RUNX1/ETO-induced mutational signature.

4.3.1. Summary of chapter

In summary, these investigations have demonstrated that:

- RUNX1/ETO potentially increases T:A > G:C transversion at *HPRT* locus *in vitro*.
- RUNX1/ETO preferentially induced mutations at central base position of $^{5'}$ApTpA$^{3'}$ / $^{5'}$TpApT$^{3'}$ sequences at *HPRT* locus *in vitro*.
- The base substitution in t(8;21) primary AML is dominated by C:G > T:A transition.
- $^{5'}$CpG$^{3'}$ dinucleotides sequence were susceptible to C > T transition possibly caused by spontaneous deamination of 5-methylcytosine.

Chapter 5. Development of Novel Flow Cytometric Method to detect Mutated Proteins in TK6 RUNX1/ETO

5.1. Introduction

Evidence suggests that the *RUNX1/ETO* fusion gene resulting from the t(8;21) translocation is a leukaemia initiating lesion, but which is insufficient to cause AML as a sole lesion and requires additional co-operating mutation to drive disease progression. For example, c-*KIT* mutations are reported in nearly half of AML patients with t(8;21) (Goemans *et al.*, 2005; Kohl *et al.*, 2005; Wang *et al.*, 2005) identifying these as a key trigger for leukaemogenesis in combination with RUNX1/ETO (Wang *et al.*, 2011). However, the mechanisms driving the acquisition of cooperating mutations remain unclear, although there is evidence that RUNX1/ETO promotes mutagenesis (Krejci *et al.*, 2008; Araten *et al.*, 2013; Forster *et al.*, 2015). In addition, ectopic expression of RUNX1/ETO downregulates several DNA repair proteins (BRCA2, OGG1 and ATM) and induces elevated DNA damage (Alcalay *et al.*, 2003), suggesting that RUNX1/ETO induces a pro-mutagenic phenotype predisposing cells to the acquisition of somatic mutations. Much of these data have been acquired using reporter gene assays, such as HPRT and TK (Forster *et al.*, 2015). As such, it would be useful to investigate the effect of RUNX1/ETO on mutation frequency in genes with an established role in AML pathogenesis, such as *FLT3* and *NPM1*. For the purposes of this experiment, the effect of RUNX1/ETO on mutagenesis was investigated at the *NPM1* locus, based exclusively on the availability of a high quality antibody specific for mutated NPM1 and will not reported affinity against wild-type NPM1 protein. However, it is noted that *NPM1* mutation is reported at very low frequency in t(8;21) AML (Thiede *et al.*, 2006) and no oncogenic cooperativity have been demonstrated between these lesions in AML leukaemogenesis.

The *NPM1* gene encodes for a protein involved in several cellular processes, including centrosome duplication, protein chaperoning, cell growth and cell proliferation (Box *et al.*, 2016). Mutations in this gene are very important diagnostic and prognostic markers in AML, and in combination with mutations in other genes such as *FLT3, CEBPA* and *DNMT3A*, are prognostic in AML. In 2016, a revised World Health Organisation (WHO) classification of myeloid neoplasms and acute leukaemia changed *NPM1* mutations status from a provisional entity to a distinct entity (Arber *et al.*, 2016) confirming the importance of dysregulated nucleophosmin in AML pathogenesis.

Insertion/deletion in exon 12 of *NPM1* is the most common mutation in AML, reported in approximately 35% of primary AML cases (Falini *et al.*, 2005) and is thought

to be an early event in AML leukaemogenesis (Thiede *et al.*, 2006; Cheng *et al.*, 2010; Vassiliou *et al.*, 2011). This mutation causes a frameshift in translation and alters the amino acid sequence at the C-terminus, resulting in a lengthening of the protein. NPM1 is a nucleocytoplasmic shuttling protein, but which primarily resides in the nucleolus. Mutation causes disruption of normal nucleocytoplasmic shuttling by changing the nuclear localisation signal and shifting the balance towards nuclear export. Accumulation of NPM1 in the cytoplasm promotes tumour growth by inactivation of p53/ARF tumour suppressor function (Falini *et al.*, 2009).

There are different types of *NPM1* exon 12 insertion mutations. Type A mutation are characterised by a duplication and insertion of 4 DNA bases (c.860_863insTCTG), and are the most frequent NPM1 mutation in AML, representing 69% of *NPM1* mutations. Type B (c.862_863insCATG) and Type D (c.863_864insCCTG) mutations are also relatively common, representing 11% and 8% of *NPM1* mutations respectively, while other mutations are rare, accounting for <1% of *NPM1* mutations (Alpermann *et al.*, 2016; Kumar *et al.*, 2018). *NPM1* mutations are significantly more common in adult patients, occurring up to 60% of patients with a normal karyotype, and are less frequent in children, reported in approximately 25% of patients with normal karyotype (Rau and Brown, 2009; Betz and Hess, 2010). Although predominantly found in AML with normal karyotype, *NPM1* mutations are also reported in AML with recurrent genetic abnormalities and minor chromosomal abnormalities (Haferlach *et al.*, 2009).

Although *NPM1* mutation is associated with a good response to induction therapy and a favourable prognosis, the favourable impact of *NPM1* mutation is highly dependent on *FLT3*-ITD status (Betz and Hess, 2010). In the absence of *FLT3*-ITD, *NPM1* mutations are associated with a favourable or better risk prognosis in patients with normal karyotype. In contrast, *FLT3-ITD* and *DNMT3A* mutations are associated with poor outcome regardless of *NPM1* status. As such, *FLT3* mutations dominate over *NPM1* mutation in terms of prognostic impact, highlighting the need to identify co-occurring somatic mutations in order to discern the prognostic impact of NPM1 mutations.

Somatic mutation screening is currently performed using high throughput sequencing, typically targeted or exome sequencing. However, immunophenotyping using mutant-specific antibody by flow cytometry could potentially replace routine sequencing techniques to detect mutated cells with *NPM1* mutation, which also has the advantage of yielding accurate quantitative information at the phenotype level in

newly diagnosed AML patients. Indeed, flow cytometry is now routine practice in clinical laboratories for several markers essential to the diagnosis, prognosis and monitoring of disease. For example, flow cytometry is commonly used in the clinical setting for the assessment of DNA ploidy, cell cycle analysis and the measurement of oxidative burst in granulocytes and reticulocytes (RBCs and platelets). In principle, flow cytometry uses fluorescently-labelled antibodies against specific proteins to identify and measure the target population of cells in a heterogeneous sample such as blood, bone marrow or lymphatic material. One of the major advantages of flow cytometry is the ability to detect and quantify small numbers of target cells in a vast excess of non-target cells. For example, minimal residual disease (MRD) monitoring by flow cytometry is an example of immunophenotyping able to detect residual tumour cells at low frequency, providing data on treatment effectiveness. However, the quality of flow cytometry data is determined by the quality of the primary and secondary antibodies, and their specificity for their target. Given the availability of a high quality antibody raised against mutant NPM1, we sought to use this to develop a flow cytometric method for measuring mutagenesis at the NPM1 locus in RUNX1/ETO positive cells.

5.1.1. Aims of Chapter 5

The purpose of the work described in this chapter was to design and establish a reliable flow cytometric method to evaluate the effect of RUNX1/ETO expression on mutation frequency in AML relevant genes. Specifically, the primary aim was to determine whether the RUNX1/ETO fusion gene predisposes cells to mutations in leukaemia relevant genes.

The experimental aims were as follows:

- Determine specificity of NPM1 mutant-specific antibody via western immunoblotting for expression of mutant protein.
- Validate the assay by flow-sorting putative NPM1 mutated cells via Sanger sequencing to confirm the mutation.
- Optimise flow cytometric assay for the identification of NPM1-positive cells.
- Measure the change in frequency of NPM1 mutated cells in RUNX1/ETO-positive cell populations cultured over an extended time period.

5.2. Results

5.2.1. Assessment of NPM1 mutant antibody specificity by western immunoblotting

The specificity of commercially available NPM1 mutant-specific polyclonal antibody (#PA1-46356, Lot No.: SE2380401, ThermoFisher Scientific) raised against an epitope (within residues 250 – 294) in the C-terminal region of human mutant nucleophosmin was determined via western immunoblotting (Figure 5.1). A panel of AML cell lines and B lymphoblastoid cell lines (TK6, TK6 RE1 and TK6 RE8) used as models in this study were assessed for both NPM1 wild type (WT) and mutant protein status. The OCI-AML3 cell line, which harbours a confirmed Type A exon 12 *NPM1* mutation (c.860_863insTCTG; p.W288fs) (Quentmeier *et al.*, 2005; Alpermann *et al.*, 2016) was used as a positive control. A band corresponding to the expected size of NPM1 wild type protein was observed in all cell lines, although there was heterogeneity in expression levels (Figure 5.1). In contrast, putative NPM1 mutant protein was observed exclusively in AML-3 cells (Figure 5.1). This western blot profile validates the specificity of NPM1 mutant-specific antibody for the detection of cells with *NPM1* exon 12 mutation.

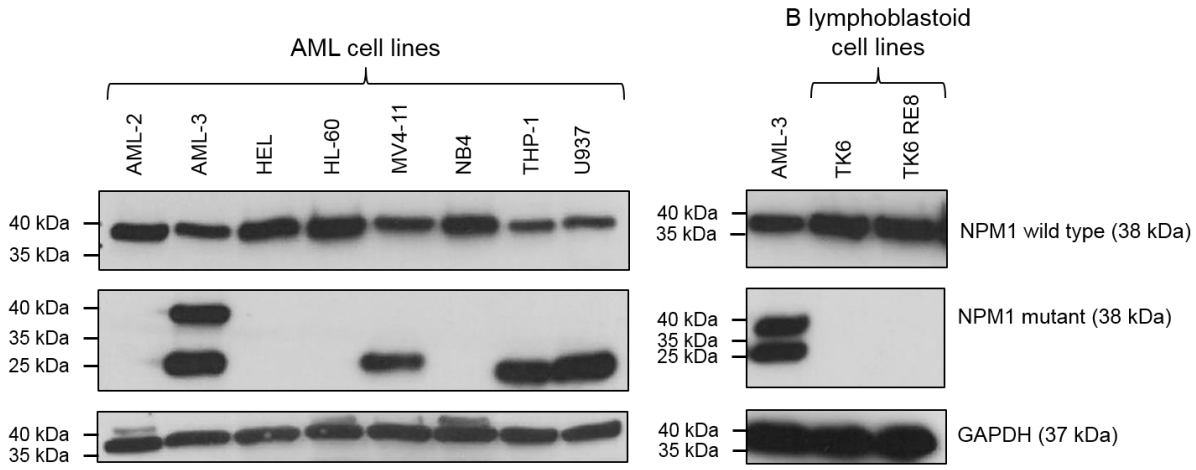


Figure 5.1. NPM1 protein expression in AML and TK6 cell lines.

Western immunoblotting was used to determine the specificity of NPM1 mutant-specific polyclonal antibody (#PA1-46356, Lot No.: SE2380401, ThermoFisher Scientific) relative to an antibody raised against wild type NPM1 (#PA5-12447, Lot No.: SH090717BE, ThermoFisher Scientific). The AML-3 cell line was included as a positive control for NPM1 mutant protein. GAPDH was used as a loading control for all blots.

5.2.2. Optimisation of flow cytometric method for the detection of NPM1-mutated cells

A series of initial antibody staining conditions were tested to identify conditions with minimal non-specific binding of the mutant-specific antibody against NPM1-WT cells, whilst maximising sensitivity of detection for NPM1-mutant cells. AML-3 cells are approximately the same cell size and have a similar granularity as TK6 cells (Figure 5.2), facilitating optimisation of a flow cytometric assay for the detection of NPM1-mutant TK6 cells. Individual fluorescent intensity profiles for each antibody staining condition were determined for AML-3, TK6 and TK6 RE8 (Figure 5.3). There was a substantial increase of approximately 10-fold in fluorescence intensity for AML-3 cells stained with mutant-specific antibody (positive control) relative to cells stained with IgG isotype (negative control), while cells stained with secondary antibody only display far less fluorescent intensity relative to the positive control (Figure 5.3A). No clear discrimination was observed between cells stained with mutant-specific antibody and isotype control for both TK6 (Figure 5.3B) and TK6 RE8 (Figure 5.3C), in contrast to clear discrimination observed for AML-3 cells. This data demonstrate the specificity of the primary antibody to detect cells with *NPM1* exon 12 mutation using flow cytometry.

The assay sensitivity was then assessed by mixing NPM1-mutant cells (AML-3) with NPM1-WT cells (TK6 RE8) to give a mixed population that was 5% positive for NPM1-mutant cells (Figure 5.4). The frequency of positive events was measured at 3.85% in this mixed sample (Figure 5.4B) and was approximate to the expected frequency of 5%. This AML-3/TK6 RE8 cell mixture showed clear discrimination between the two peaks, with the positive NPM1 mutant population displays 10-fold higher in fluorescent intensity relative to negative NPM1 mutant population. The median fluorescence intensity (MFI) of the NPM1 peak for the AML-3/TK6 RE8 cell mixture was determined and divided by the MFI of the NPM1 peak for the simultaneously stained WT control sample to generate the normalised NPM1 MFI value (Figure 5.4C). A normalised NPM1 MFI value of 6.0 and above was determined to correspond with NPM1-mutant positivity. This data demonstrate the sensitivity of the assay to detect NPM1 mutated event at low frequency.

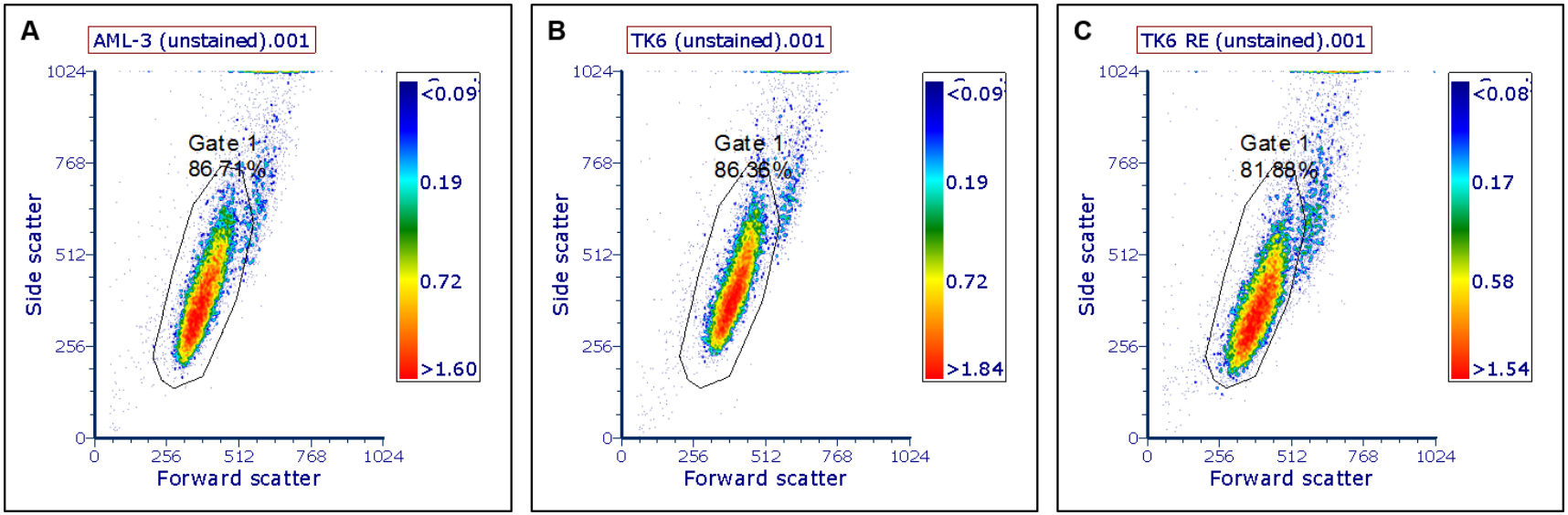


Figure 5.2. Density plot for AML-3, TK6 and TK6 RE8 cells.

(A) AML-3, (B) TK6 and (C) TK6 transduced with RUNX1/ETO (TK6 RE8). Coloured area represent the relative cell frequency of each population. Cells were gated based on forward scatter (size) and side scatter (granularity) to identify the cell population (Gate 1) and exclude debris and cell doublets. All cell populations were captured using the same gate. 1 x 10^4 events were recorded per sample.

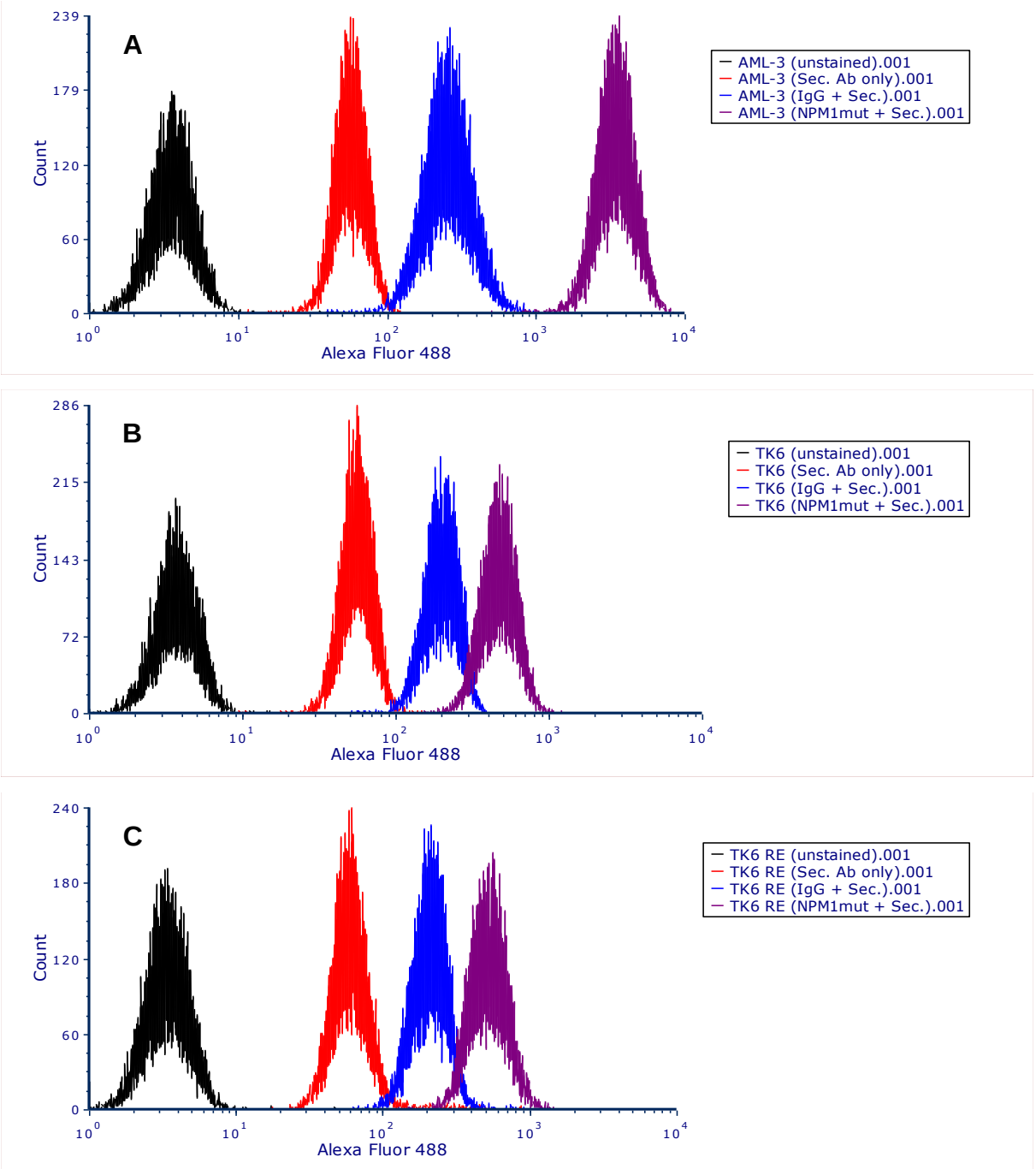
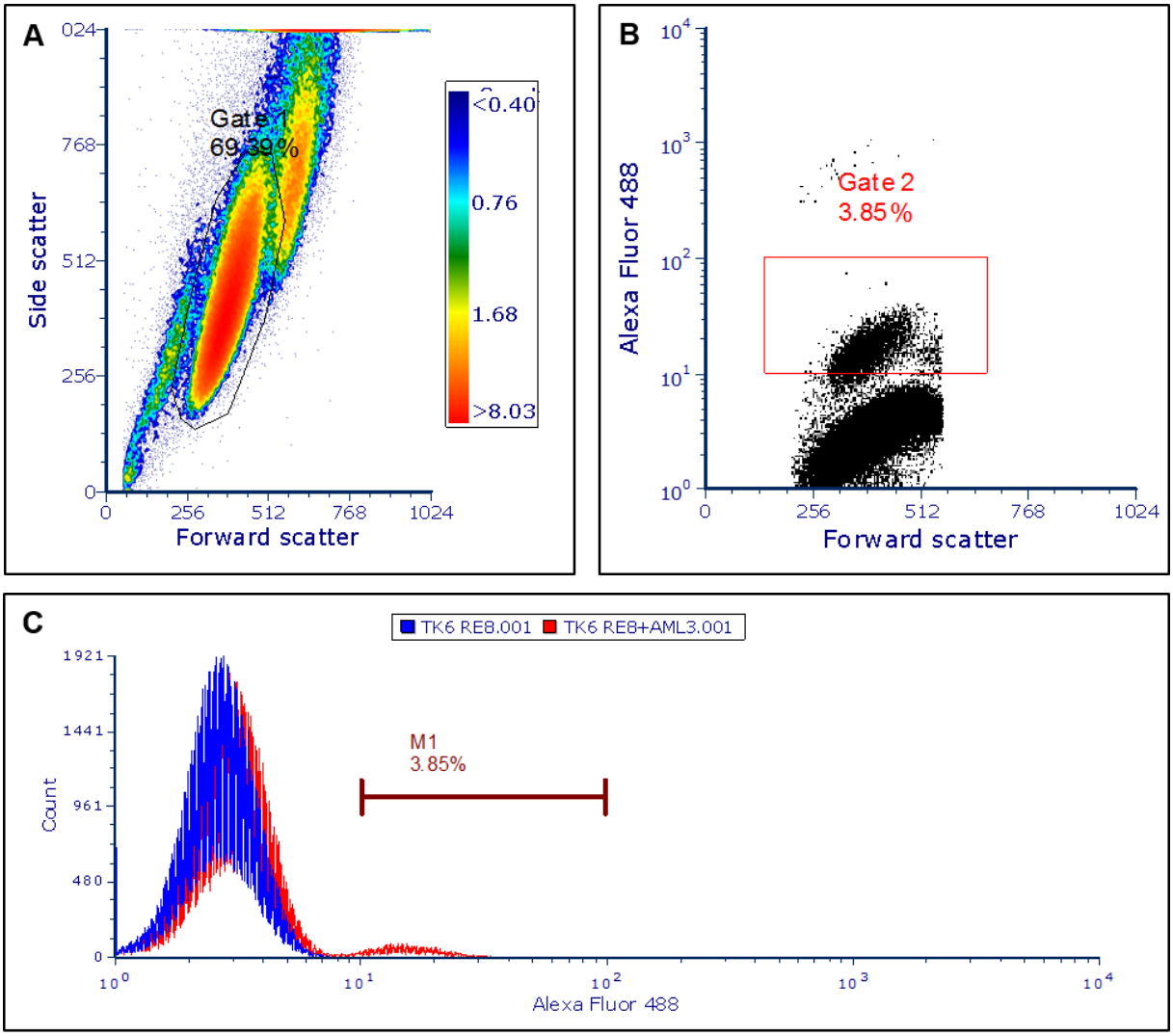


Figure 5.3. Fluorescence intensity profiles for AML-3, TK6 and TK6 RUNX1/ETO cells stained with primary antibody, secondary antibody or IgG isotype.

(A) AML-3, (B) TK6, and (C) TK6 transduced with RUNX1/ETO (TK6 RE8). Histogram displays the fluorescence intensity profiles for each cell line by comparison of the negative unstained control cells stained with primary antibody, secondary antibody or IgG isotype negative control. Clear discrimination was observed between AML-3 cells stained with primary antibody (NMP1 mutant polyclonal antibody) and AML-3 cells stained with IgG isotype (internal negative or isotype control), demonstrating specificity of the primary antibody to detect cells with mutated *NPM1* using flow cytometry. 1 x 10^4 events were recorded per sample.



Histogram	Parameter	Marker	MFI	Normalised MFI
TK6 RE8	Alexa Fluor 488	None	2.64	1.00
5% AML-3/TK6 RE8	Alexa Fluor 488	M1	16.20	6.14

Figure 5.4. Sensitivity of detection for NPM1-mutant cells in a majority population of NPM1 WT cells.

(A) Density plot showing size and granularity of the gated population, (B) Dot plot showing clear separation of positive and negative cell populations, and (C) Histogram subtraction of TK6 RE cells (NPM1 WT) population from the TK6 RE/AML-3 (5% NPM1-mutant) cell population. The NPM1 mutant positive population was generated by mixing AML-3 cells (positive control) with TK6 RE8 cells to give a population that was 5% positive for NPM-1 mutant cells. Positive events were defined as those having normalised MFI of more than 6.0 relative to the negative population.

5.2.3. Validation of NPM1 mutations in flow-sorted cells by Sanger sequencing

5.2.3.1. Flow-sorting of positive events for NPM1 exon 12 mutation

To confirm the NPM1 mutation status of positive events, cells were sorted and collected for DNA extraction. A mixed population of 5% AML-3/95% TK6 RE8 cells was used to determine gating parameters for selection of positive events (Figure 5.5). A population of positive events was gated and sorted, representing 2.83% of the total mixed cell population. In contrast, using the same gating parameters, only 0.07% of positive events were sorted from the pure TK6 RE8 cell population, broadly consistent with the low frequency of *NPM1* exon 12 mutations expected in this population.

5.2.3.2. PCR and Sanger sequencing analysis of NPM1 exon 12 mutation status

PCR of DNA extracted from flow-sorted cells was performed using primers targeting *NPM1* exon 12 predicted to give an amplicon of 269bp (Figure 5.6). Extracted DNA was either amplified directly or following whole genome amplification using the REPLI-g Single Cell Kit. PCR products for DNA extracted from REPLI-g Single Cell Kit produced stronger bands when compared to those amplified using DNA extracted directly from QIAamp DNA micro kit that had not been whole genome amplified.

Sequencing of the PCR product (DNA extracted from REPLI-g Single Cell Kit) confirm a heterozygous 4 bases insertion (TCTG sequence) in cells flow-sorted from the AML-3/TK6 RE8 cells (Figure 5.7). This heterozygous insertion was also present in flow-sorted TK6 RE8 cells. These findings validate the assay sensitivity to identify NPM1 mutated events at low frequency in a cell population using commercially available NPM1 mutant-specific polyclonal antibody.

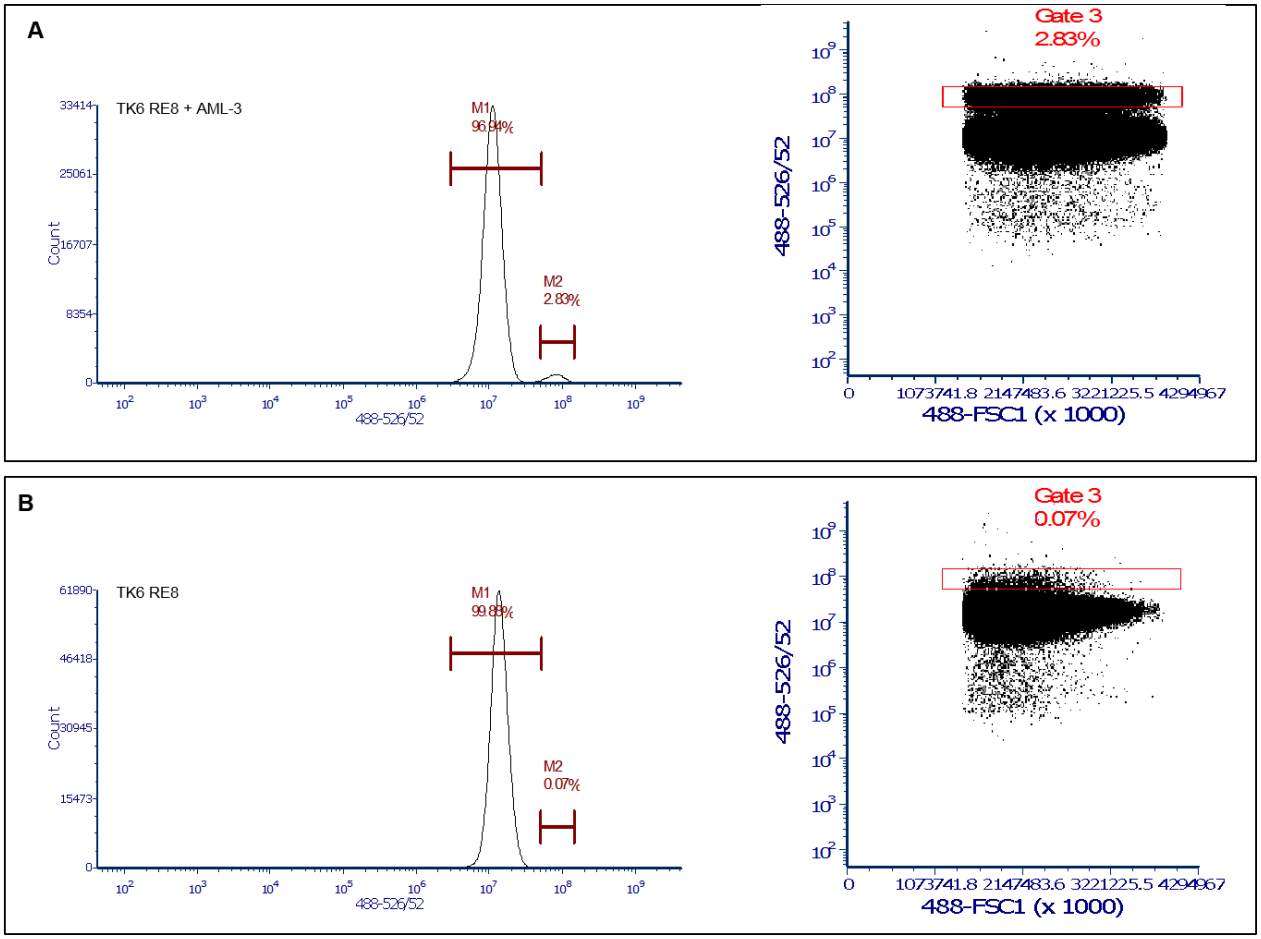


Figure 5.5. Fluorescence-activated cell sorting (FACs) of the positive *NPM1* exon 12 mutation events.

(A) 5% AML-3/TK6 RE8 mixed sample and (B) TK6 RE8. 5% AML-3/TK6 RE8 mixed sample was used as a control to identify the positivity of *NPM1* exon 12 mutation. Positive events in TK6 RE8 population were gated and sorted based on the positivity of AML-3 cells in mixed sample.

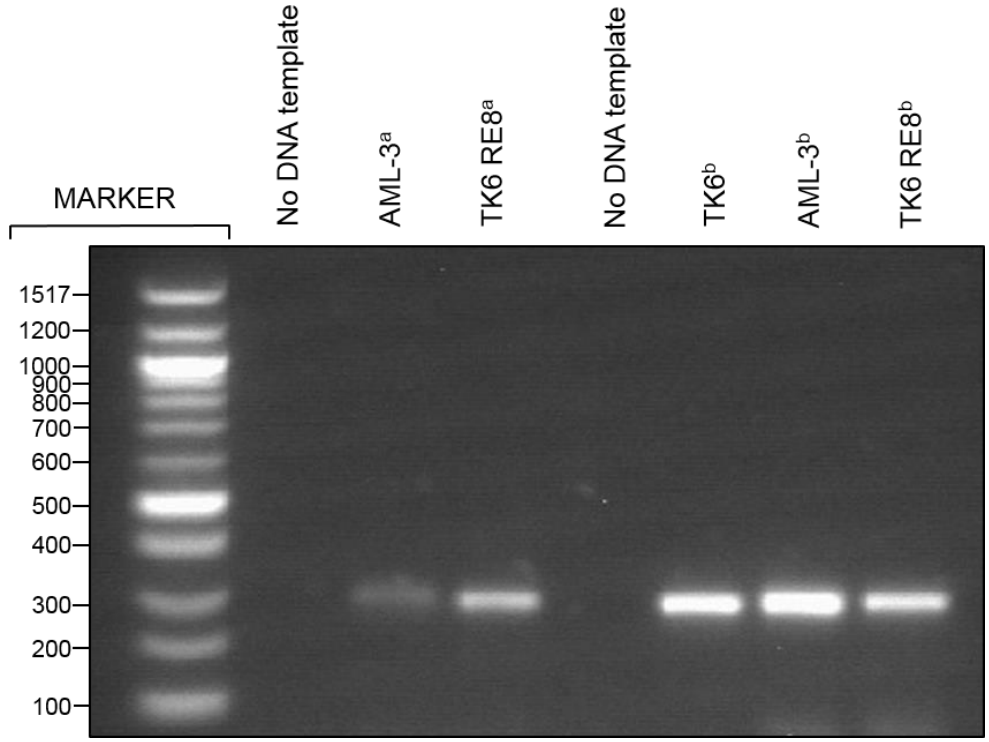


Figure 5.6. PCR amplification of the *NPM1* exon 12.

The gel shows the PCR profile of the flow-sorted cells representing the targeted C-terminal region of *NPM1* exon 12. 'No DNA template' was used as negative control for PCR amplification.

[a]DNA extracted via QIAamp DNA micro kit (Qiagen).

[b]DNA extracted via REPLI-g Single Cell Kit (Qiagen).

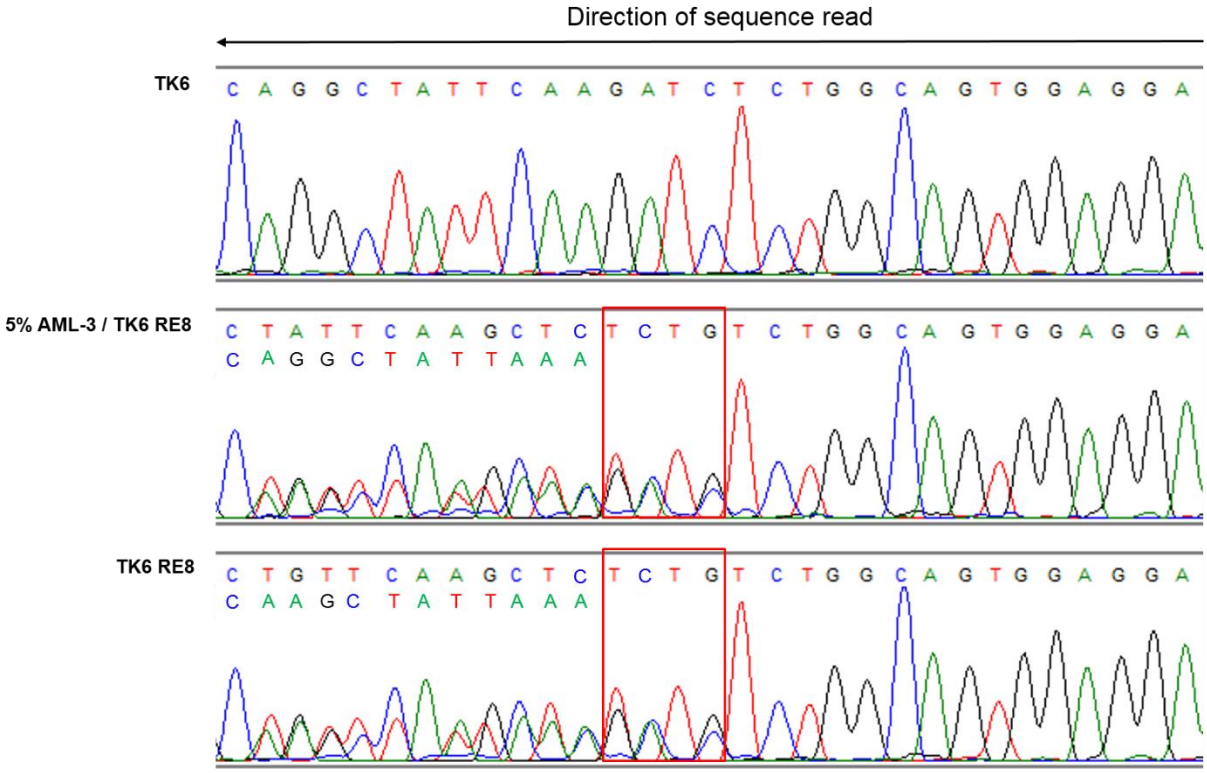


Figure 5.7. Sequencing of the targeted *NPM1* exon 12 region.

Sequencing of DNA from the flow-sorted cells confirmed a heterozygous 4 bases duplication in the positive events from the AML-3 / TK6 RE8 mixed population and TK6 RE8 population. The duplication of a TCTG sequence is highlighted in a red box. Unsorted TK6 cells were included as a negative control for *NPM1* mutation. Amplified *NPM1* exon 12 region was sequenced for each cell population; however, only the portion of the positive strand (reverse read) containing the mutation is shown. Underlying sequence in a chromatogram trace characterised by two peaks and starting from the base adjacent to the site of insertion is also shown.

5.2.4. Effect of RUNX1/ETO expression on the acquisition of somatic mutations

To investigate whether RUNX1/ETO expression predisposes cells to the acquisition of somatic mutations in the *NPM1* gene 10 independent subclones each of parental TK6 and RUNX1/ETO-transduced clones (TK6 RE1 and TK6 RE8) were generated using limiting dilution (see Chapter 2, Section 2.8.1). These subclones were assessed for the acquisition of NPM1 mutant cells over an extended culture to determine the effect of RUNX1/ETO expression on mutagenesis in exon 12 of *NPM1*.

5.2.4.1. Expression levels of RUNX1/ETO in TK6 and TK6 RE subclones determined by western immunoblotting

RUNX1/ETO expression was assessed in TK6, TK6 RE1 and TK6 RE8 subclones relative to Kasumi-1 cells via western immunoblotting (Figure 5.8). RUNX1/ETO protein was expressed in all RUNX1/ETO subclones, derived from TK6 RE1 and TK6 RE8 clones. Quantification of the protein bands revealed RUNX1/ETO subclones with a range of expression relative to that found in Kasumi-1 cells. The majority of TK6 RE1 subclones demonstrated lower levels of RUNX1/ETO expression compared to Kasumi-1 cells, although some TK6 RE1 subclones (TK6 RE1_8, 9, 10) had RUNX1/ETO expression levels similar to Kasumi-1. The majority of TK6 RE8 subclones had RUNX1/ETO expression at levels similar to Kasumi-1, although there were some with lower RUNX1/ETO expression. (TK6 RE8_7, 8, 9). RUNX1/ETO expression was not observed in TK6 parental cells, as expected (Figure 5.8). All clones were cultured long-term and used to determine the effect of RUNX1/ETO expression on the acquisition of exon 12 *NPM1* mutations.

5.2.4.2. Effect of RUNX1/ETO on NPM1 exon 12 mutation frequency

The TK6, TK6 RE1 and TK6 RE8 subclones were grouped into 3 groups based on the RUNX1/ETO expression level relative to Kasumi-1 cells: 1) No RE expression (0%), 2) Low RE expression (10 – 50%) and 3) High RE expression (51 – 100%) (Table 5.1). Using a normalised NPM1 MFI value of more than 6.0 relative to the negative population to correspond with NPM1-mutant positivity (Section 5.2.2), MF of *NPM1* exon 12 mutation was measured in these subclones via flow cytometry for every 2 weeks, starting at week 7 to week 22 post-cloning. A significant increase in mutation

frequency was observed in all clones over the course of the acquisition period (P < 0.0001; one-way ANOVA) (Table 5.1) (Figure 5.9), although there was some fluctuation in mean mutation frequency in each of the groups over the time course of the experiment. The effect of RUNX1/ETO fusion expression on the acquisition of *NPM1* exon 12 mutation was also evaluated by comparing the *NPM1* exon 12 MF between subclones with RUNX1/ETO expression and subclones with no RUNX1/ETO expression at each timepoint. For subclones with low RUNX1/ETO expression, a significant increase in MF was observed only at week 22 relative to subclones with no RUNX1/ETO expression (P < 0.05; paired *t*-test), while for subclones with high RUNX1/ETO expression, a significant increase in MF was observed at week 11, 20 and 22 relative to subclones with no RUNX1/ETO expression (week 11: P < 0.05; week 20: P < 0.05; week 22: P < 0.0001; paired *t*-test). In addition, no significant increase in MF was observed in subclones with high RUNX1/ETO expression relative to subclones with low RUNX1/ETO expression at each timepoint except at week 20 (P < 0.05; paired *t*-test). Despite that, overall MF demonstrated subclones with RUNX1/ETO expression had significantly higher *NPM1* exon 12 mutation frequency compared to subclones with no RUNX1/ETO expression (low RE vs no RE, P = 0.01; high RE vs no RE, P < 0.01; paired *t*-test) (Figure 5.10). Furthermore, subclones with high RUNX1/ETO fusion expression had a higher mean *NPM1* exon 12 mutation frequency than subclones with low RUNX1/ETO fusion expression (P < 0.05; paired *t*-test), indicating the significant effect of RUNX1/ETO expression level. Taken together, these data provide evidence that RUNX1/ETO predisposes cells to the acquisition of somatic mutations in the *NPM1* gene, although it was not strongly evident at each timepoint.

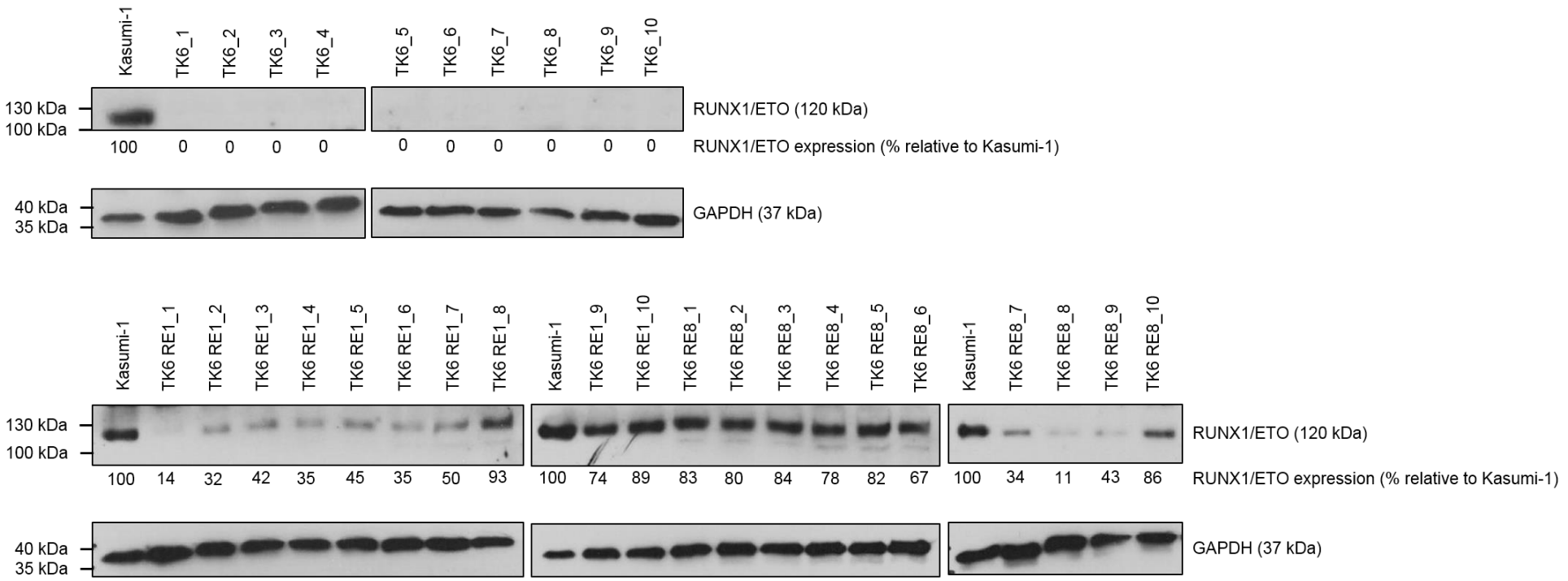


Figure 5.8. RUNX1/ETO protein levels in TK6 and its derivative subclones.

Western immunoblotting was performed to assess RUNX1/ETO fusion protein expression in parental TK6 subclones and TK6 RE subclones. Ten subclones for each cell population were generated by cloning in soft agar and expansion in culture. Kasumi-1 cells were included as an RE positive control. A band of approximately 120 kDa equivalent to the expected size of RUNX1/ETO is present in RUNX1/ETO-transduced sub clones, with the majority of TK6 RE1 subclones having lower expression of RUNX1/ETO compared to TK6 RE8 subclones. GAPDH was used as a loading control for all blots.

Time (Week)	Mean MF ($x10^{-5}$)			Paired *t*-test[b]		
	No RE expression	Low RE expression	High RE expression	Low RE vs No RE	High RE vs No RE	High RE vs Low RE
7	1.81	2.15	1.86	0.4213	0.8834	0.4111
9	1.85	2.06	2.43	0.4669	0.2811	0.4623
11	2.16	3.06	3.62	0.0626	0.0021	0.1939
13	3.19	3.71	3.98	0.0718	0.0930	0.6360
15	2.99	3.35	4.00	0.4502	0.0939	0.2482
18	4.64	5.23	5.34	0.3949	0.1562	0.8235
20	5.00	5.32	6.91	0.5137	0.0428	0.0149
22	5.63	7.28	8.12	0.0174	< 0.0001	0.2167
One-way ANOVA[a]	< 0.0001	< 0.0001	< 0.0001			

Table 5.1. Effect of RUNX1/ETO expression on *NPM1* mutation frequency over time.

Data represent the mean MF from ten independent subclones for each original clone (no RE expression, low RE expression and high RE expression) measured every 2 weeks, starting at week 7 to week 22 post-cloning (Appendix E). These figures are displayed graphically in Figure 5.9.

[a] Figures represent the P value determined using a one-way analysis of variance (ANOVA) test for effect of time (week) on MF.

[b] Figures represent the P value determined using a paired *t*-test for effect of RUNX1/ETO expression on MF at each timepoint.

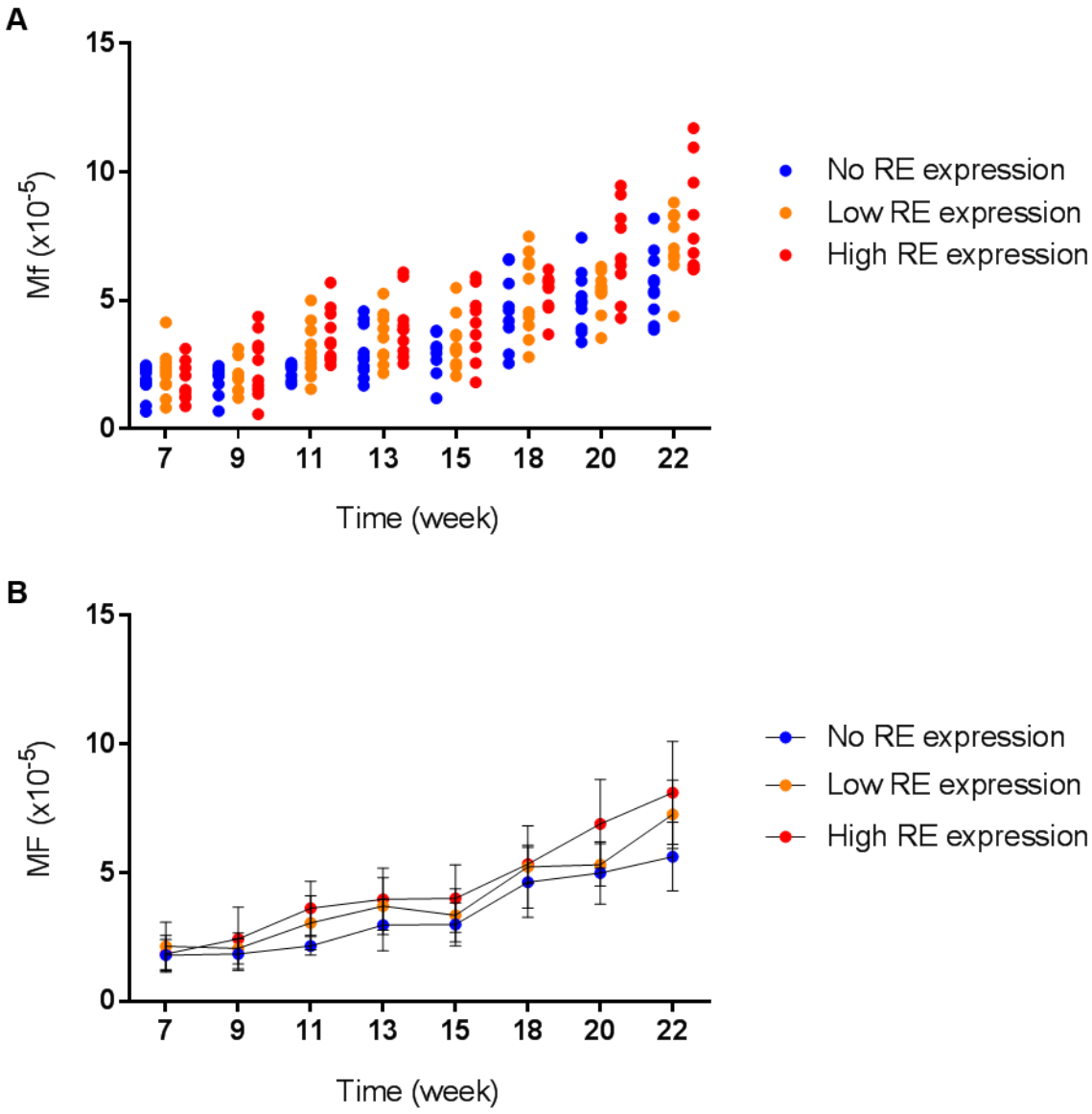


Figure 5.9. Effect of RUNX1/ETO expression on *NPM1* mutation frequency over time.

MF of *NPM1* exon 12 mutation was measured in no RE expression subclones (blue dots), low RE expression subclones (orange dots) and high RE expression subclones (red dots) every 2 weeks, starting at week 7 to week 22 post-cloning. (A) Dot plot represents data for each subclone. (B) Mean and standard deviation (error bars) for all subclones from each original clone at each timepoint. Data is a representative of 10 subclones for each original clone.

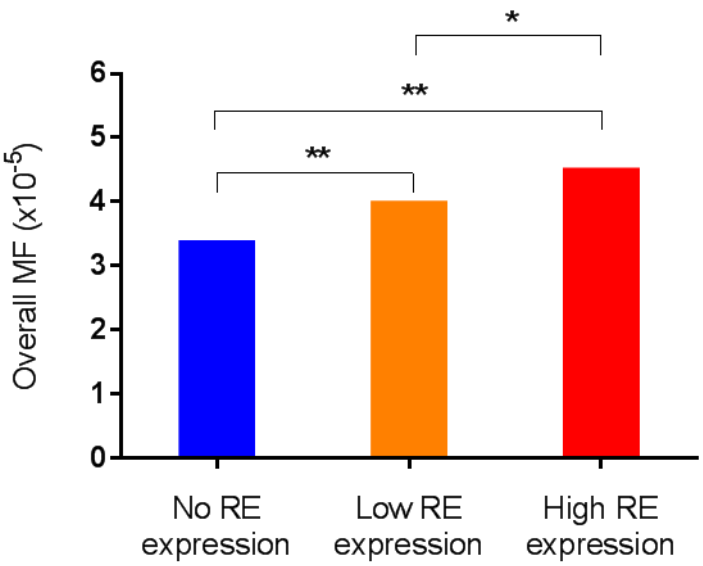


Figure 5.10. Effect of RUNX1/ETO expression on *NPM1* mutation frequency. Overall mean MF of *NPM1* exon 12 mutation was measured in no RE expression subclones (blue bar), low RE expression subclones (orange bar) and high RE expression subclones (red bar) for a total of 22 weeks. P-value was calculated by paired *t*-test comparing the overall MF between groups. $P < 0.01$ (**), $P < 0.05$ (*).

5.3. Discussion

The effect of RUNX1/ETO in the acquisition of somatic mutations in AML genes is poorly documented although numerous studies have identified the co-occurrence of mutations with the *RUNX1/ETO* fusion gene in patients with t(8;21)-positive AML (Hartmann *et al.*, 2016; Patel *et al.*, 2017). *c-Kit* mutation is the most common recurrent mutation found in t(8;21) AML, reported in nearly half of cases (Goemans *et al.*, 2005; Kohl *et al.*, 2005; Wang *et al.*, 2005) and has been described to cooperate with RUNX1/ETO in models of AML (Wang *et al.*, 2011; Wichmann *et al.*, 2015). has been identified as a cooperating mutation in this disease (Wang *et al.*, 2011). Other recurrent mutations in t(8;21) AML include *NRAS*, *FLT3*, *ASXL2* and *ZBTB7A* (Reikvam *et al.*, 2011; Hartmann *et al.*, 2016). In addition, other mutations were also found in t(8;21) AML such as *DNMT3A*, *JAK2* and *CBL* at low frequency (Hartmann *et al.*, 2016) of which can be classified as secondary “passenger” mutations as a result of continuous somatic mutation acquisition. Although many of these secondary mutations seem to have no importance in RUNX1/ETO AML leukaemogenesis, a contribution to disease progression cannot be excluded. In this study, we have evaluated the effect of RUNX1/ETO expression on *NPM1* exon 12 mutation frequency in RUNX1/ETO negative cells and two RUNX1/ETO positive cell populations with different levels of RUNX1/ETO expression.

A cell line harbouring a confirmed type A exon 12 *NPM1* mutation was used in conjunction with a commercially available NPM1 mutant specific polyclonal antibody to develop a flow cytometric method for the detection of *NPM1* mutation. Detection of NPM1 mutant cells in a mixed population of primarily NPM1 WT cells requires a very sensitive technique. As such, flow cytometry was selected as the method of choice based on the ability to screen large numbers of cells quickly and efficiently at the single cell level. In contrast, western immunoblotting, for example, is not sensitive enough to detect mutation at the single cell level. However, western immunoblotting was essential to conform the suitability of this antibody for detection of NPM1 mutated event in flow cytometry. During development of the flow cytometric method it was essential to investigate non-specific binding and the frequency of false-positive events. Assay development included analysis of unstained cell, cells stained with conjugated secondary antibody only, cells stained with isotype control and cells stained with mutant antibody; controls which are standard practice in evaluating the specificity of antibody for flow cytometry. The positive cell population was observed to have a higher

fluorescent intensity than cells stained with the IgG isotype control and cells stained with conjugated secondary antibody alone, as demonstrated by clear separation between the three peaks from the histogram data of AML-3 cells (Figure 5.3A). In contrast, due to the low frequency of *NPM1* mutation in TK6 cells, the frequency of positive events was very low and there was no clear separation in term of fluorescent intensity, with most of the cells staining negative and consistent with the IgG control (Figure 5.3B and 5.3C). A mixed population of 5% AML-3/95% TK6 cells was used to validate the method. This was aided by the fact that AML-3 and TK6 cells have similar characteristics (size and granularities), which allowed for a direct comparison based solely on fluorescence intensity. The frequency of positive events in the mixed cell population was estimated to be 3.85% with positivity defined as MFI of 6.0 and above relative to the putative negative population. Sanger sequencing was subsequently used to confirm that the population of cells deemed positive were in fact almost exclusively positive for *NPM1* exon 12 mutation (a heterozygous TCTG insertion), as evidenced by a virtually identical profile of sorted positive cells and AML-3 positive control cells. These data confirmed the sensitivity of the flow cytometry assay for the detection of *NPM1* exon 12 mutated cells in a mixed population.

TK6 and RUNX1/ETO transduced subclones with a range of RUNX1/ETO expression (TK6 RE1 and TK6 RE8) were used as a model to investigate the effect of RUNX1/ETO expression on *NPM1* exon 12 mutation frequency. Unlike a reporter gene assay which takes advantage of a selective agent to reduce the level of spontaneous *TK* and *HPRT* mutants, there is no mechanism to purge a population of pre-existing cells with *NPM1* mutation. Alternatively, single cell cloning can be used to ensure that a cell population was WT at a given point in time; specifically, when seeded using limiting dilution. Therefore, 10 subclones for each TK6 cell population (parental TK6, TK6 RE1 and TK6 RE8) were generated and used to monitor the acquisition of *NPM1* mutation over time. Ten independent replicates for each TK6 population was deemed sufficient to support statistical analysis. It is noteworthy that all TK6 subclones, irrespective of the population from which they were derived, predominantly stained negative by flow cytometry, indicating that the initial cell of each clone must have therefore been *NPM1* WT. However, the acquisition of spontaneous mutations was observed in all TK6 subclones as demonstrated by the significant increase at every time point within the acquisition period regardless of RUNX1/ETO expression status (Table 5.1), suggesting that cells acquired mutations over the time as a result of continuous cell division. Furthermore, overall analysis revealed that the frequency of

NPM1 mutation was significantly higher in RUNX1/ETO-positive subclones compared to RUNX1/ETO-negative subclones (Figure 5.10). The level of RUNX1/ETO expression was also observed to influence the induced mutation frequency. TK6 subclones displaying high levels of RUNX1/ETO expression had a significant higher *NPM1* mutation frequency compared to TK6 subclones that have lower RUNX1/ETO expression (Figure 5.10), although it was not statistically significant at most of the timepoint (Table 5.1). Nevertheless, these data provide evidence that RUNX1/ETO predisposes cells to the acquisition of somatic mutations in the *NPM1* gene, although it was not strongly evident at each timepoint.

Despite the significant effect of RUNX1/ETO expression on the acquisition of somatic mutation in *NPM1* exon 12, *NPM1* mutations are commonly found in patients with normal karyotype (Rau and Brown, 2009), although a very low frequency of *NPM1* mutation was reported in t(8;21) AML (Thiede *et al.*, 2006). Regardless of rare reports of *NPM1* and *RUNX1/ETO* concurrence, these mutations are generally considered to be mutually exclusive. Furthermore, *RUNX1/ETO* and *NPM1* mutations are classified as distinct entities by the WHO classification of myeloid neoplasms and acute leukaemia (Arber *et al.*, 2016). Like *RUNX1/ETO*, *NPM1* mutation alone is insufficient for leukaemogenesis and requires cooperating events to drive transformation (Grisendi *et al.*, 2005; Sportoletti *et al.*, 2008; Cheng *et al.*, 2010; Sportoletti *et al.*, 2013). It has been suggested that mutations in receptor tyrosine kinase genes such as *c-Kit*, *FLT3* and *RAS* involved in signal transduction pathways, are required for leukaemia development in RUNX1/ETO AML (Jiao *et al.*, 2009). In contrast, *NPM1* mutations are often reported with mutation in genes involved in the regulation of DNA methylation (*DNMT3A, TET2, IDH1, IDH2*), RNA splicing (*SRSF2, SF3B1*), the cohesin complex (*RAD21, SMC1A, SMC3, STAG2*) and cell signalling pathways (*FLT3, NRAS,* and *PTPN11*) (Patel *et al.*, 2017). Both mutations have been well characterised as initiating mutation in AML (Cheng *et al.*, 2010; Vassiliou *et al.*, 2011; Wang *et al.*, 2011; Wichmann *et al.*, 2015) although there are some reports suggesting that *NPM1* mutation can occur late and might be the final transforming event AML pathogenesis, and particularly secondary AML (Schnittger *et al.*, 2014; Patel *et al.*, 2017).

The data from this study suggests that RUNX1/ETO promotes mutagenesis by predisposing cells to the acquisition of somatic mutations in the *NPM1* gene. Although *NPM1* mutations are rarely found to co-exist with the *RUNX1/ETO* fusion oncoprotein, these data provide proof of principle that RUNX1/ETO drives mutagenesis at loci relevant in the pathogenesis of AML. Previous studies have shown that RUNX1/ETO

alone is sufficient to increase genomic instability by downregulating DNA-repair proteins (Alcalay *et al.*, 2003; Krejci *et al.*, 2008; Forster *et al.*, 2015) (see Chapter 1, Section 1.3.3). This will lead to the increase of mutation frequency in gene such as *NPM1* as a result of mutagenic effect of *RUNX1/ETO*.

5.3.1. Summary of chapter

In summary, these investigations have demonstrated that:

- Flow cytometry is sufficiently sensitive to identify and measure NPM1-mutated cells at low very frequency in a predominantly wild-type cell population.
- *NPM1* exon 12 mutation frequency is significantly increased over time, suggesting continuous spontaneous somatic mutation acquisition.
- Mutation frequency is positively associated with RUNX1/ETO expression level.

Chapter 6. High Throughput Exome Sequencing to Investigate the Clonal Evolution of Relapsed AML

6.1. Introduction

Cytogenetic analysis has been used for more than three decades to define the molecular pathogenesis of AML, and remains as the first tier screening for AML classification. Recurrent chromosomal structural variations such as t(8;21), inv(16), t(15;17), del(5) and del(7) are established diagnostic and prognostic markers suggesting that acquired genomic abnormalities play an essential role in leukaemogenesis (Betz and Hess, 2010). However, nearly half of AML cases have a normal karyotype lacking recurrent structural abnormalities.

Somatic point mutations affecting numerous genes have been described in AML at presentation, many of which are pathogenic and prognostic, including *RUNX1 (*Tang *et al., 2009)*, *NPM1 (*Becker *et al., 2010)*, *NRAS/KRAS* (Janssen *et al.,* 1987) and *CEBPA* (Green *et al.*, 2010). High throughput sequencing technology has intensified the search for somatic mutations in AML, leading to the discovery of novel mutations in established pathways, such as *RAS (Tyner et al., 2009)*, and identified new pathways through whole genome-based approaches (Ley *et al.*, 2008). This approach has also been applied to the study of relapsed leukaemia, successfully identifying relapse-specific coding sequence mutations affecting *ETV6* and *MYO18B* (Ding *et al.*, 2012).

The accumulation of somatic mutations in cancer cells as a result of continuous acquisition of mutations during progression of the disease suggests cancer is an evolutionary process. Most mutations at disease presentation in AML are thought to be acquired after initiating genetic lesions such as t(8;21) and t(15;17), although a small number of pre-existing mutations was already present before cells acquire these advantageous initiating mutations. There is some evidences that initiating events promote mutagenesis predisposing cells to the acquisition of additional somatic mutations (Alcalay *et al.*, 2003; Krejci *et al.*, 2008; Araten *et al.*, 2013; Forster *et al.*, 2015) which may contribute to genomic instability in a cancer genome. Genomic instability can increase the mutation rate in the cancer genome through many different mechanisms which play an important role in cancer evolution.

Chemotherapy using cytotoxic agents during remission induction treatment may eradicate cancer cells with dominant clones, however most of AML patients experience disease relapse which is difficult to treat. The mechanisms driving relapse evolution remain unclear although there is evidence that relapse is driven by novel mutations acquired after the chemotherapy. High throughput sequencing of matched

presentation and relapse AML samples have revealed two major patterns of AML relapse evolution (Ding *et al.*, 2012; Welch *et al.*, 2012; Garg *et al.*, 2015; Sood *et al.*, 2015). Firstly, a leukaemic clone at disease presentation acquires additional mutations and evolves into the relapse clone after the chemotherapy and secondly, the pre-leukaemic clone from the founding clone acquires additional mutations and evolves into a relapse clone. Importantly, both patterns of relapse are defined by the acquisition of additional somatic mutations.

Therefore, it is thought that there are two major underlying mechanisms driving mutagenesis in relapsing AML. Firstly, the genomic instability of the leukaemic clone is a major cause of mutagenesis and contributes to genetic heterogeneity in AML. Secondly, chemotherapy agents used in induction chemotherapy contribute to relapsed AML since most chemotherapy agents are genotoxic which may lead mutation. Collectively, both mechanisms are predicted to be responsible for the aetiology of relapse-driver mutations.

In this study, whole exome sequencing of matched presentation and relapse AML samples was performed to map changes in mutational landscape between presentation and relapse AML in order to determine the clonal origins of relapsed AML and clonal evolution during treatment of AML in order to better understanding disease evolution and heterogeneity in AML.

6.1.1. Aims of Chapter 6

The objective of the work described in this chapter was to compare genetic variants at initial AML diagnosis to those present at relapse in a cohort of Malaysian AML cases in order to identify novel mutations at relapse (relapse-specific mutations) and to identify low-frequency mutations at presentation that are selected for by chemotherapy important in the clonal evolution of relapsed AML.

Specifically, the experimental aims were as follows:

- Identify a cohort of matched presentation and relapse AML cases for whole exome sequencing.
- Identify somatic variants at presentation and relapse through several analysis filters using Ingenuity Variant Analysis™ software (Qiagen Bioinformatics) in order to select variants with putative functional impact on disease.
- Interrogate filtered variants for their frequencies and deleterious effect in reported and published AML cases using the Catalogue of Somatic Mutations in Cancer (COSMIC) database.
- Compare somatic variants detected at presentation and relapse using variant allele fraction (VAF) score to map the clonal evolution of relapsed AML in each patient.

6.2. Results

6.2.1. Patients and clinical data

A cohort of 11 matched presentation and relapsed AML cases with different cytogenetics and molecular profiles were selected from diagnostic samples referred to National Referral Centre for Bone Marrow Cytogenetics and Molecular Genetics, Institute for Medical Research, Kuala Lumpur, Malaysia (Table 6.1). All samples have insufficient metaphase spread for conclusive karyotyping (fewer than 20 cells and some samples were not sent for cytogenetics) except 1 relapse sample (patient 10). Multiplex RT-PCR using HemaVision®-28N kit (DNA Diagnostic, Risskov, Denmark) was performed for all samples to screen for 28 leukaemia causing translocations including common translocations in AML such as t(8;21)(q22;q22), t(15;17)(q24;q21) and inv(16)(p13;q22). This qualitative test uses reverse transcription of RNA to cDNA followed by multiplex nested polymerase chain reaction and agarose gel electrophoresis to identify chromosomes, genes and exons at the breakpoint in fusion genes. In addition, some samples were further screened for AML prognostic markers including gene mutations in *NPM1*, *KIT* and *FLT3* via targeted PCR using specific primers targeted to regions of interest.

6.2.2. Identification of somatic variants

Paired-end whole exome sequencing was performed for each sample to identify somatic variants. By enriching for exonic sequence, high quality whole exome sequencing reads with average coverage of 93.8% of exome (range: 77.3 – 96.5%) and average sequencing depth/read depth of 151X (range: 15 – 434X) was achieved.

Somatic variants including nonsynonymous single nucleotide variant (SNV, both missense and nonsense) and small insertion/deletion (indel, both in-frame and frameshift), as well as splice site variants were identified (Table 6.2) following several layers of filters using Ingenuity Variant Analysis software (Chapter 2.9.4) and after exclusion of variants with benign and uncertain significance. Two mutations at presentation, *NPM1* p.W288fs (Patient 1), *FLT3*-ITD (Patient 1) and one mutation at relapse, *NPM1* p.W288fs (Patient 9) detected from initial RT-PCR screening were not detected by exome sequencing. In total, 40 and 48 variants were identified in

Patient	Status	Date at Referral	Age at Referral (Years)	Blast %	Cytogenetics	Molecular Diagnostic (28 translocations)	AML Prognostic Marker
1	P	15/05/2014	26	55	46, XX [4]	Negative	*FLT3*-ITD positive, *NPM1* (MutA)
	R	28/10/2014		30	ND	Negative	*FLT3*-ITD negative, *NPM1* (MutA)
2	P	25/01/2014	10	> 80	ND	t(9;11)(p22;q23)	ND
	R	05/03/2014		> 80	ND	t(9;11)(p22;q23)	ND
3	P	30/05/2012	10	ND	ND	t(15;17)(q24;q21)	ND
	R	23/04/2013		> 90	ND	t(15;17)(q24;q21)	ND
4	P	18/04/2013	17	> 80	46, XX [11]	Negative	ND
	R	21/12/2015		60	ND	Negative	*FLT3*-ITD negative, *NPM1* negative
5	P	02/10/2015	43	91	ND	Negative	*FLT3*-ITD negative, *NPM1* negative
	R	13/06/2016		94	ND	Negative	*FLT3*-ITD negative, *NPM1* negative
6	P	20/01/2015	35	78	ND	t(8;21)(q22;q22)	*FLT3*-ITD negative, *cKIT* negative
	R	20/05/2016		54	ND	t(8;21)(q22;q22)	*FLT3*-ITD negative, *cKIT* negative
7	P	15/05/2014	36	> 70	ND	t(8;21)(q22;q22)	*FLT3*-ITD negative, *cKIT* negative
	R	01/12/2015		18	46, XY [4]	t(8;21)(q22;q22)	*FLT3*-ITD negative, *cKIT* negative
8	P	16/06/2014	40	11	ND	Negative	*FLT3*-ITD negative, *NPM1* negative
	R	04/08/2015		70	ND	inv(16)(p13;q22)	*FLT3*-ITD negative, *NPM1* negative
9	P	20/06/2014	51	54	ND	Negative	*FLT3*-ITD negative, *NPM1* (MutA)
	R	18/12/2015		20	ND	Negative	*FLT3*-ITD negative, *NPM1* (MutA)
10	P	22/01/2016	27	21	46, XY [5]	inv(16)(p13;q22)	ND
	R	07/02/2017		60	46, XY, inv(16)(p13;q22) [22]	inv(16)(p13;q22)	*FLT3*-ITD negative, *cKIT* negative
11	P	21/01/2016	33	61	46, XX [2]	Negative	*FLT3*-ITD negative, *NPM1* negative
	R	17/03/2017		91	46, XX [2]	Negative	ND

Table 6.1. Selected features of matched presentation and relapsed AML patients included in this study.

P – Initial disease presentation, R – relapse disease

ND – not determined

Patient	Gene	Reference Allele	Variant Allele	Variation Type	Gene Region	AA Change	Presentation		Relapse		Frequency in AML as reported in COSMIC (%)
							Read Depth	VAF (%)	Read Depth	VAF (%)	
1	PTPRQ	GGGT		deletion	exonic	c.6452_6453+2delGGGT	-	-	103	17.5	0.0 (0/0)
1	SELE	C	T	SNV	splice site	c.902-1G>A	113	55.8	111	55.9	0.0 (0/0)
3	LIFR	C	A	SNV	splice site	c.992-1G>T	-	-	68	16.2	0.0 (0/0)
3	PTPRQ	GGGT		deletion	exonic	c.6452_6453+2delGGGT;	111	19.8	-	-	0.0 (0/0)
3	VPS13B	C		frameshift deletion	exonic	c.5334delC; p.R1779fs*2	252	48.0	222	57.2	0.0 (0/5)
4	ADAM9		TTTT	frameshift insertion	exonic	c.1883_1884insTTTT; p.C629fs*3	43	18.6	-	-	0.0 (0/0)
4	APOB	G	A	nonsynonymous SNV	exonic	c.10579C>T; p.R3527W	278	45.0	247	40.9	0.0 (0/11)
4	DNAH5	T		deletion	splice site	c.2578-2delA	-	-	17	17.7	0.0 (0/8)
4	NRAS	C	T	nonsynonymous SNV	exonic	c.35G>A; p.G12D	390	14.4	434	33.4	29.13 (328/1126)
5	CEBPA		CTG	in-frame insertion	exonic	c.934_936dupCAG; p.Q312dup	283	87.6	268	47.0	2.08 (28/1346)
5	CFTR	G	A	nonsynonymous SNV	exonic	c.1865G>A; p.G622D	103	50.5	94	56.4	0.0 (0/2)
5	MUTYH	C	A	stop gain SNV	exonic	c.1435G>T; p.E479*	152	47.4	126	48.4	0.0 (0/0)

Table 6.2. Somatic variants identified from whole exome sequencing (continued on next page).

Patient	Gene	Reference Allele	Variant Allele	Variation Type	Gene Region	AA Change	Presentation		Relapse		Frequency in AML as reported in COSMIC (%)
							Read Depth	VAF (%)	Read Depth	VAF (%)	
6	ASXL1		G	frameshift insertion	exonic	c.1934dupG; p.G646fs*12	-	-	126	40.5	25.56 (171/669)
6	CACNA2D4	G	T	stop gain SNV	exonic	c.2637C>A; p.C879*	86	37.2	54	63.0	0.0 (0/1)
6	IL12RB1	G	A	stop gain SNV	exonic	c.1681C>T; p.R561*	111	47.8	71	94.4	0.0 (0/0)
6	NFIX		T	frameshift insertion	exonic	c.1168_1169insT; p.P390fs*41	139	20.1	-	-	0.0 (0/11)
6	RASA2	TG		frameshift deletion	exonic	c.2202_2203delTG; p.G735*	170	38.2	128	40.6	0.0 (0/0)
6	TET2		AGAGG	frameshift insertion	exonic	c.1016_1017insGAGGA; p.N339fs*10	246	65.5	184	88.6	0.0 (0/921)
6	TP53	ACCACTACTC		deletion	exonic	c.783-6_786delGAGTAGTGGT	-	-	35	80.0	0.0 (0/379)
6	VPS13B		TTTTTTT	insertion	exonic	c.9406-1_9406insTTTTTTTTTT	15	33.3	-	-	0.0 (0/5)
7	CCDC8	CTCT		frameshift deletion	exonic	c.1283_1286delAGAG p.Q428fs*50	122	43.4	129	36.4	0.0 (0/0)
7	CFTR	G	A	nonsynonymous SNV	exonic	c.1865G>A; p.G622D	86	45.4	64	48.4	0.0 (0/2)
7	FKTN	G	A	SNV	splice site	c.910+1G>A	132	54.6	114	47.4	0.0 (0/3)
7	PKP2	C	A	SNV	splice site	c.2300-1G>T	-	-	39	15.4	0.0 (0/0)
7	USH1C	A	T	start loss SNV	exonic	c.25T>A; p.M1K	92	55.4	95	47.4	0.0 (0/1)

Table 6.2. Somatic variants identified from whole exome sequencing (continued from previous page).

Patient	Gene	Reference Allele	Variant Allele	Variation Type	Gene Region	AA Change	Presentation		Relapse		Frequency in AML as reported in COSMIC (%)
							Read Depth	VAF (%)	Read Depth	VAF (%)	
8	ESRRB	C	T	nonsynonymous SNV	exonic	c.1144C>T; p.R382C	202	49.5	236	48.7	0.0 (0/0)
8	FLT3	G	A	nonsynonymous SNV	exonic	c.1577C>T; p.T526M	419	48.5	385	48.8	0.0 (0/16119)
8	HBB	AAAG		frameshift deletion	exonic	c.126_129delCTTT; p.F42fs*19	241	50.2	263	39.5	0.0 (0/0)
8	LFNG		GATG	frameshift insertion	exonic	c.163_166dupGATG; p.E56fs*2	104	50.0	116	35.3	0.0 (0/1)
8	MAP3K20	A	T	SNV	splice site	c.416-2A>T	-	-	66	15.2	0.0 (0/0)
8	MAP3K20	G	T	SNV	splice site	c.416-1G>T	-	-	66	15.2	0.0 (0/0)
9	ESRRB	C	T	nonsynonymous SNV	exonic	c.1144C>T; p.R382C	218	48.6	216	50.5	0.0 (0/0)
9	EYS	C	A	stop gain SNV	exonic	c.8170G>T; p.E2724*	155	57.4	153	54.3	0.0 (0/0)
9	GUCY2D	C	T	nonsynonymous SNV	exonic	c.164C>T; p.T55M	75	66.7	64	53.1	0.0 (0/0)
9	IDH2	C	T	nonsynonymous SNV	exonic	c.419G>A; p.R140Q	124	38.7	115	26.1	42.89 (636/1483)
9	NPM1		TCTG	frameshift insertion	exonic	c.860_863dupTCTG; p.W288fs	-	-	17	41.2	36.28 (2252/6207)
9	SLC12A1	C	A	stop gain SNV	exonic	c.1584C>A; p.Y528*	204	55.4	146	38.4	0.0 (0/2)
9	WRN	AGGG		deletion	splice site	c.840-2_841delAGGG;	-	-	31	19.4	0.0 (0/0)
9	WRN		TTTT	frameshift insertion	exonic	c.844_845insTTTT; p.S282fs*24	-	-	31	19.4	0.0 (0/0)

Table 6.2. Somatic variants identified from whole exome sequencing (continued from previous page).

Patient	Gene	Reference Allele	Variant Allele	Variation Type	Gene Region	AA Change	Presentation		Relapse		Frequency in AML as reported in COSMIC (%)
							Read Depth	VAF (%)	Read Depth	VAF (%)	
10	AIMP1	C	T	stop gain SNV	exonic	c.115C>T; p.Q63*	58	60.3	42	47.6	0.0 (0/0)
10	KIT	ACG		in-frame deletion	exonic	c.1255_1257delGAC; p.D419del	-	-	113	37.2	4.58 (35/764)
10	TTN	C	A	SNV	splice site	c.32554+1G>T	127	40.9	120	50.8	0.0 (0/0)
11	ATM	G	T	nonsynonymous SNV	splice site	c.497-1G>T	-	-	18	33.3	0.0 (0/102)
11	ATRX	G	A	stop gain SNV	exonic	c.7219C>T p.R2407*	-	-	120	44.2	14.29 (2/14)
11	CEBPA		TTGGCCTT	frameshift insertion	exonic	c.929_930insAAGGCCAA p.Q346fs*10	303	42.9	349	49.6	0.0 (0/1346)
11	CEBPA	CTCCA		frameshift deletion	exonic	c.923_927delTGGAG; p.V343fs*11	329	43.5	350	47.4	0.0 (0/1346)
11	MPST	C	A	stop gain SNV	exonic	c.52G>T; p.E18*	184	51.1	193	49.7	0.0 (0/0)
11	RYR2	C	T	nonsynonymous SNV	exonic	c.13739C>T; p.T4580M	-	-	29	51.7	0.0 (0/5)
11	TBXAS1	C	T	stop gain SNV	exonic	c.997C>T; p.R333*	148	41.2	150	52.0	0.0 (0/2)
11	VWF	G	A	nonsynonymous SNV	exonic	c.4195C>T; p.R1399C	161	40.4	154	37.7	0.0 (0/1)

Table 6.2. Somatic variants identified from whole exome sequencing (continued from previous page).

presentation and relapse samples respectively (Table 6.3) with an average of 3.6 variants per patient at presentation (range 0 – 9) and 4.4 variants per patient at relapse (range 0 – 8). In terms of type of variant, SNVs were more common than indels at both presentation and relapse with 18 indels and 22 SNVs identified at presentation and 19 indels and 29 SNVs identified at relapse.

A total of 34 shared variants occurred at both presentation and relapse (Figure 6.1). Moreover, 4 variants were presentation-specific, and 14 variants were relapse-specific. Additional somatic variants were acquired in six relapse samples (Patient 1, 7, 8, 9, 10 and 11) which gained between 1 and 3 new variants when compared to their corresponding presentation samples. Four patients had the same number of variants at both presentation and relapse samples (Patient 3, 4, 5 and 6), of which 3 patients (Patient 3, 4 and 6) acquired and lost equivalent numbers of variants.

6.2.3. Evaluation of somatic variants at presentation and relapse

The filtered somatic variants were further evaluated to define the oncogenic effect of the variants in determining their potential importance as AML driver mutations via interrogation of the COSMIC database. The mutated genes were ranked according to frequency of genes implicated in AML as reported in the COSMIC database (Figure 6.2) regardless of type and position of the variants in genes, as either classical types of variants, or previously reported and/or recurrent hotspots variants, or novel variants.

Somatic variants in recurrently mutated genes in AML (*NPM1*, *FLT3*, *NRAS*, *TET2*, *ATM*, *CEBPA*, *IDH2*, *TP53*, *ASXL1* and *KIT*) were identified in 8 patients (Patient 1, 4, 5, 6, 8, 9, 10 and 11) with four of these variants newly acquired at relapse. Six of these variants are classic AML driver and recurrent hotspot mutations whereas variants such as *TET2* p.N339fs*10 (Patient 6), *TP53* c.783-6_786del (Patient 6), *FLT3* p.T526M (Patient 8), *ATM* c.497-1G>T (Patient 11), *CEBPA* p.Q346fs*10 (Patient 11) and *CEBPA* p.V343fs*11 (Patient 11) were not previously reported in AML (Table 6.2). In addition to somatic variants in recurrently mutated genes, there were 14 mutated genes that have been previously reported in a small number of AML cases in COSMIC (less than 2%) and an additional 18 mutations in genes not previously reported in AML (Figure 6.2).

Patient	Presentation			Relapse		
	Indel	SNV	Total no. of variants	Indel	SNV	Total no. of variants
1	2	1	3	3	1	4
2	0	0	0	0	0	0
3	2	0	2	1	1	2
4	1	2	3	1	2	3
5	1	2	3	1	2	3
6	4	2	6	4	2	6
7	1	3	4	1	4	5
8	2	2	4	2	4	6
9	3	5	8	3	5	8
10	0	2	2	1	2	3
11	2	3	5	2	6	8
Total no. of variants	18	22	**40**	19	29	**48**

Table 6.3. Number of somatic variants at presentation and relapse.

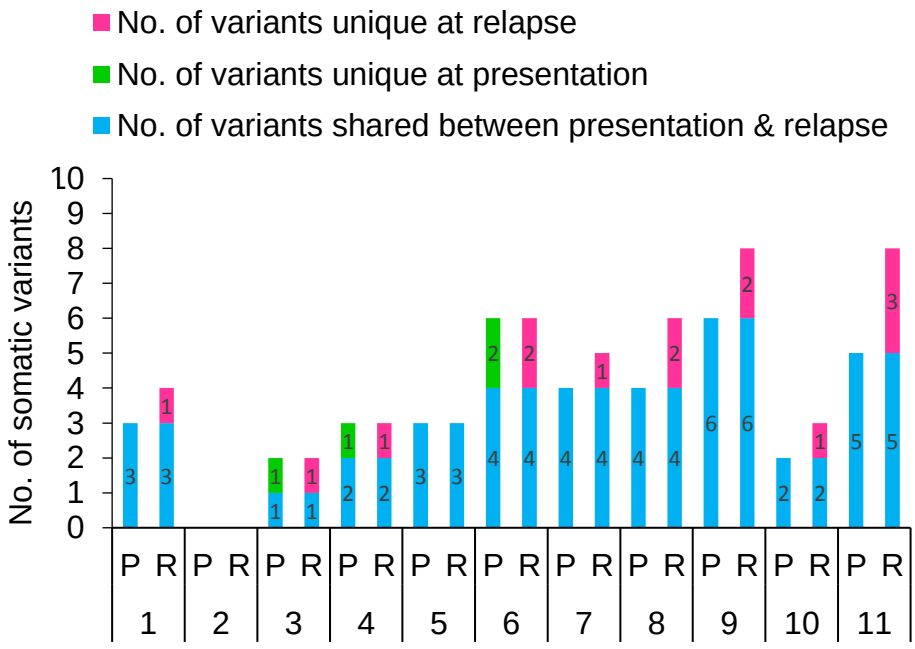


Figure 6.1. Total number of somatic variants specific to presentation or relapse and somatic variants shared at both disease stages.

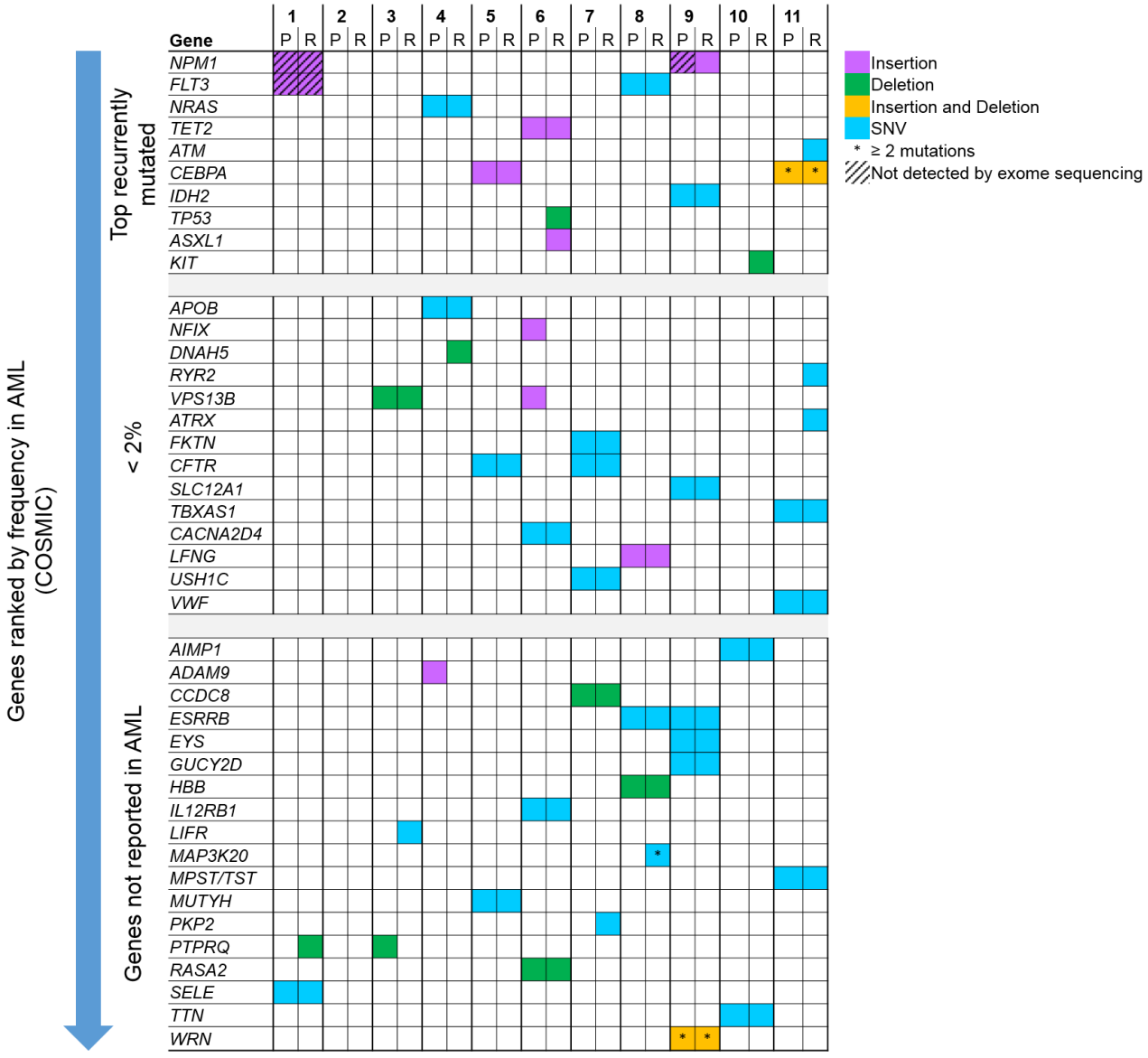


Figure 6.2. Mutational spectrum of somatic variants at presentation and relapse characterised according to frequency in AML as reported in COSMIC.

The mutated genes identified in each patient were ranked according to frequency of genes implicated in AML as reported in the COSMIC database.

To further analyse identified mutations, the mutated genes were clustered into four groups based on their putative effect that may relevant for leukaemic transformation according to existing model of leukaemogenesis (Kelly and Gilliland, 2002; Conway O'Brien *et al.*, 2014): (1) mutations affecting genes that contribute to cell proliferation (class I mutations); (2) mutations affecting genes involved in myeloid differentiation (class II mutations); (3) mutations affecting genes involved in epigenetic modification and (4) mutations affecting genes that have as yet unknown function in leukaemogenesis (Figure 6.3). All patients have mutations in at least one gene that either affect cell proliferation, are involved in myeloid differentiation or involved in epigenetic modification at presentation, with 4 patients (Patient 1, 5, 6 and 9) having both class I and II mutations at presentation. At relapse, acquisition of both classes of mutations was observed in 3 patients (3, 10 and 11) that lack complementary class of mutation at presentation. In contrast, the remaining 28 mutated genes identified in this cohort have yet unknown functions in leukaemogenesis.

6.2.4. Correlation of presentation VAF and relapse

VAF was used to track each somatic variant from presentation to relapse to map clones that persisted, those that resolved from presentation to relapse as well as novel mutations that emerged at relapse in order to delineate the clonal architecture of leukaemic cells and to identify driver mutations that promote clonal expansion in relapse AML (Figure 6.4). However, caution must be exercised given that the VAF will vary according to the blast count, therefore minor differences in VAF between diagnosis and relapse samples cannot be considered as significant.

NRAS p.G12D (Patient 4), *CEBPA* p.Q312dup (Patient 5), *TET2* p.N339fs*10 (Patient 6), *FLT3* p.T526M (Patient 8), *IDH2* p.R140Q (Patient 9), *CEBPA* p.Q346fs*10 (Patient 11) and *CEBPA* p.V343fs*11 (Patient 11) were among mutations in known AML driver genes that were retained after chemotherapy with VAF of 14.4%, 87.6%, 65.5%, 48.5%, 38.7%, 42.9% and 43.5%, respectively, in the presentation samples and 33.4%, 47.0%, 88.6%, 48.8%, 26.1%, 49.6% and 47.4% in the relapse samples. On the other hand, *ASXL1* p.G646fs*12 (Patient 6), *TP53* (Patient 6), *KIT* p.D419del (Patient 10) and *ATM* c.497-1G>T (Patient 11) were mutations in known AML driver genes acquired at relapse with VAF of 40.5%, 80.0%, 37.2% and 33.3% respectively. VAF score for two mutations at both presentation and relapse in Patient 1 (*NPM1* p.W288fs and *FLT3*-ITD) and one mutation at presentation in Patient 9 (*NPM1*

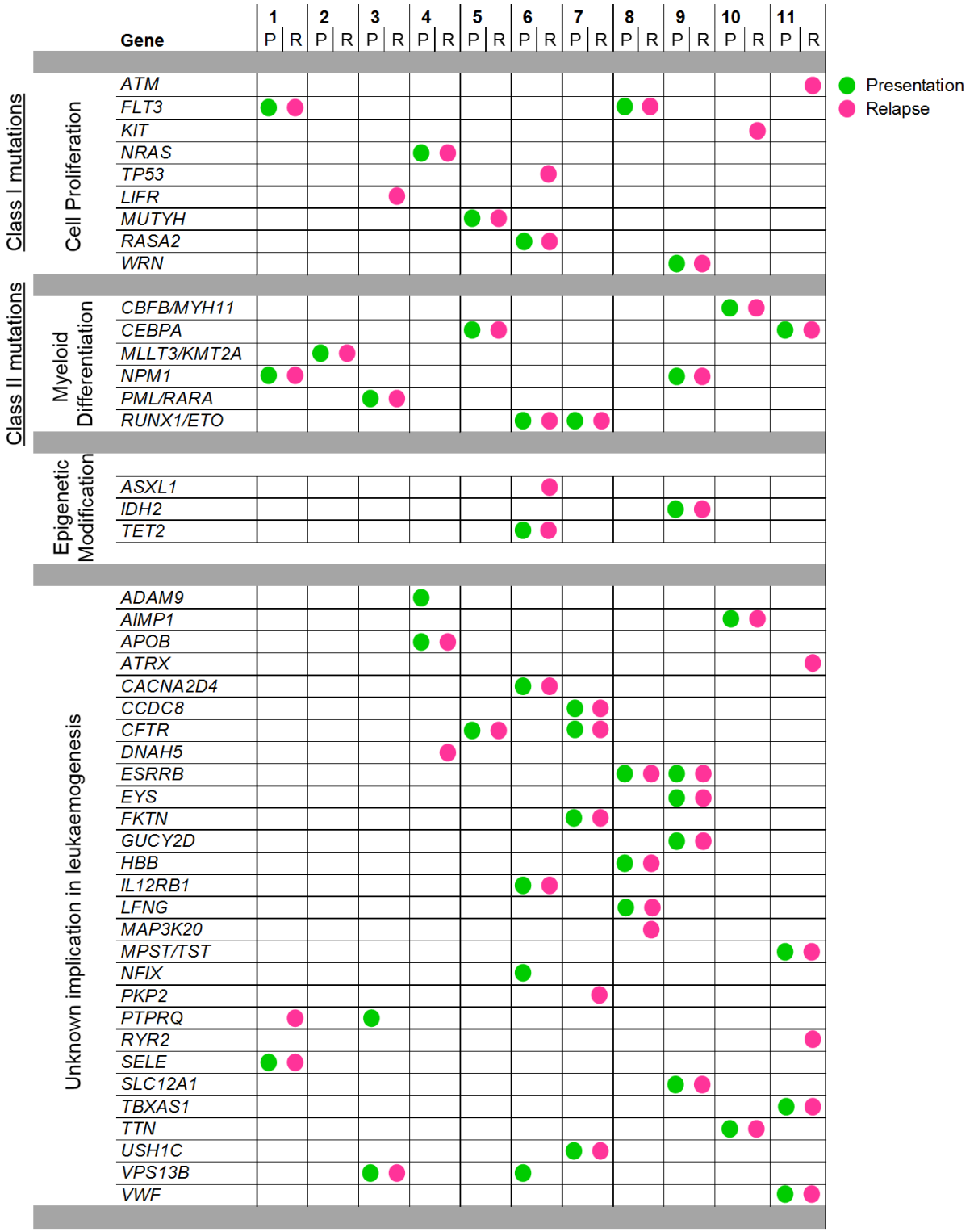


Figure 6.3. Relationship between somatic variants at presentation and relapse samples based on their known or predicted role in leukaemogenesis (Kelly and Gilliland, 2002; Conway O'Brien *et al.*, 2014).

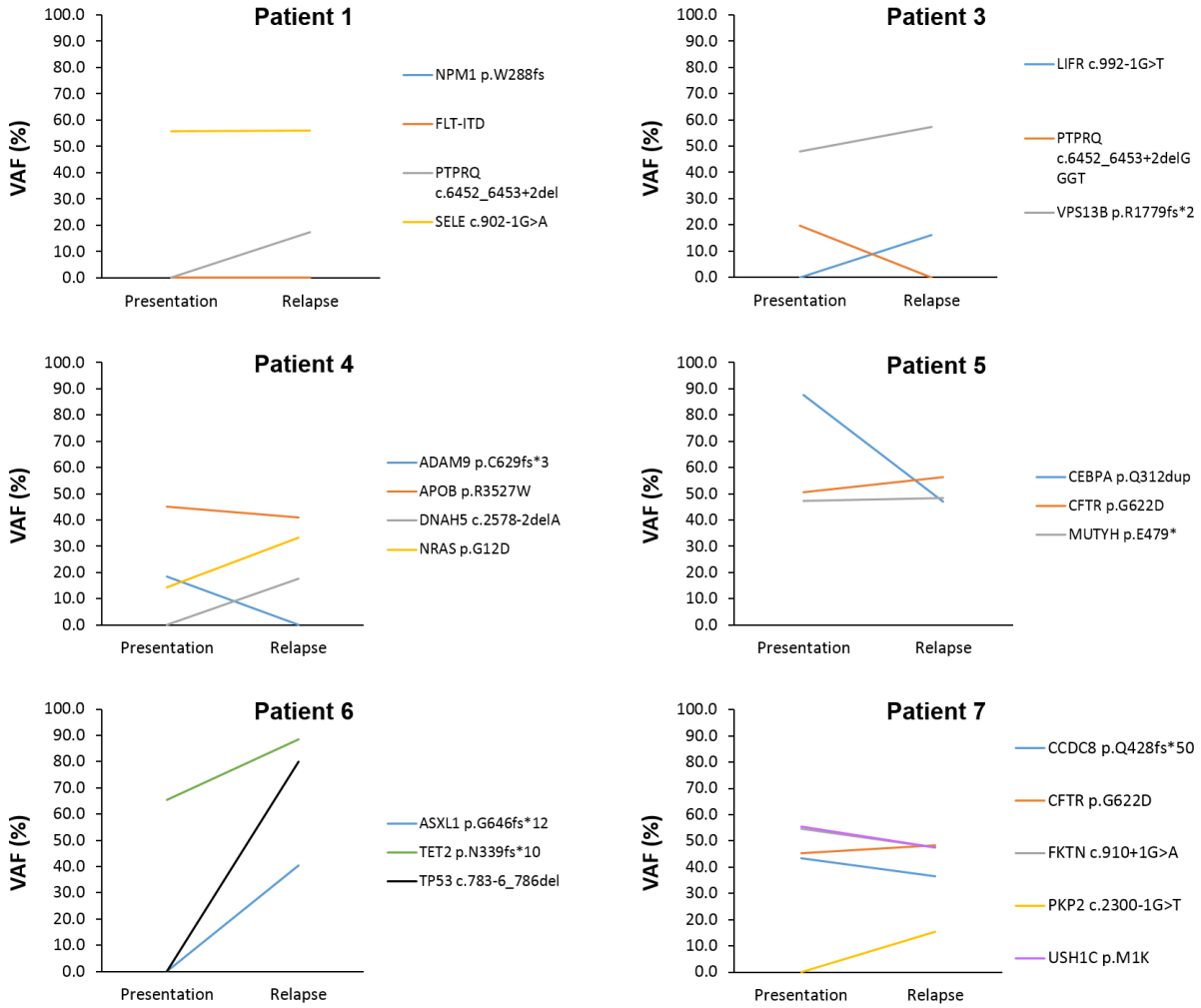


Figure 6.4. Variant allele fraction (VAF) score for each somatic variants identified from disease presentation to relapse in individual patient (continued on next page).

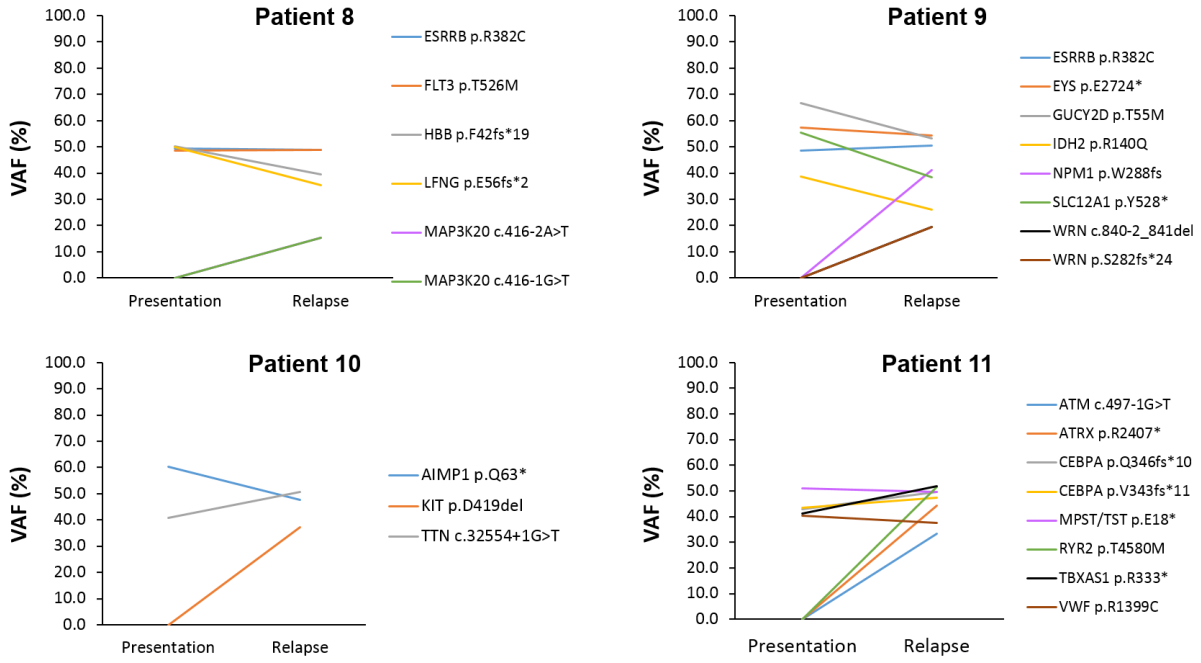


Figure 6.4. Variant allele fraction (VAF) score for somatic variants in each individual AML patient (continued from previous page).

VAF score for two mutations at both presentation and relapse in Patient 1 (*NPM1* p.W288fs and *FLT3*-ITD) and one mutation at presentation in Patient 9 (*NPM1* p.W288fs) were not available (not detected by exome sequencing). Each line represent one variant.

p.W288fs) detected by molecular diagnostic screening, were not available as they were not detectable by exome sequencing. In addition, mutations in other genes acquired at relapse include *PTPRQ* c.6452-6453+2del (Patient 1), *LIFR* c.992-1G>T (Patient 3), *DNAH5* c.2578-2delA (Patient 4), *PKP2* c.2300-1G>T (Patient 7), *MAP3K20* c.416-2A>T (Patient 8), *MAP3K20* c.416-1G>T (Patient 8), *WRN* c.840-2_841delAGGG (Patient 9), *WRN* p.S282fs*24 (Patient 9), *ATRX* p.R2407* (Patient 11) and *RYR2* p.T4580M (Patient 11) with a range of 15.2% – 80.0% VAF score.

Variants were further clustered into four groups to define the effect of chemotherapy on the prevalence of somatic variants at relapse specifically in known AML driver genes based on comparison between VAF score at presentation and relapse (Figure 6.5). Except for *NPM1* p.W288fs and *FLT3*-ITD mutations in which no VAF scores were available (not detected by exome sequencing), *FLT3* p.T526M, *CEBPA* p.Q346fs*10 and *CEBPA* p.V343fs*10 mutations were found to be persist after treatment by having similar values at presentation and relapse (less than 10% difference). *NRAS* p.G12D and *TET2* p.N339fs*10 mutations were predicted to confer relative resistance to standard combination chemotherapy treatment based on their increased VAF (greater than 10% difference) at relapse, whereas *CEBPA* p.Q312dup mutation was predicted to confer sensitivity to chemotherapy based on a reduced VAF score (greater than 10% difference) at relapse. Four mutations including *ASXL1* p.G646fs*12, *TP53* c.783-6_786del, *KIT* p.D419del and *ATM* c.497-1G>T are predicted to be induced by or selected by chemotherapy since these were not detected at presentation.

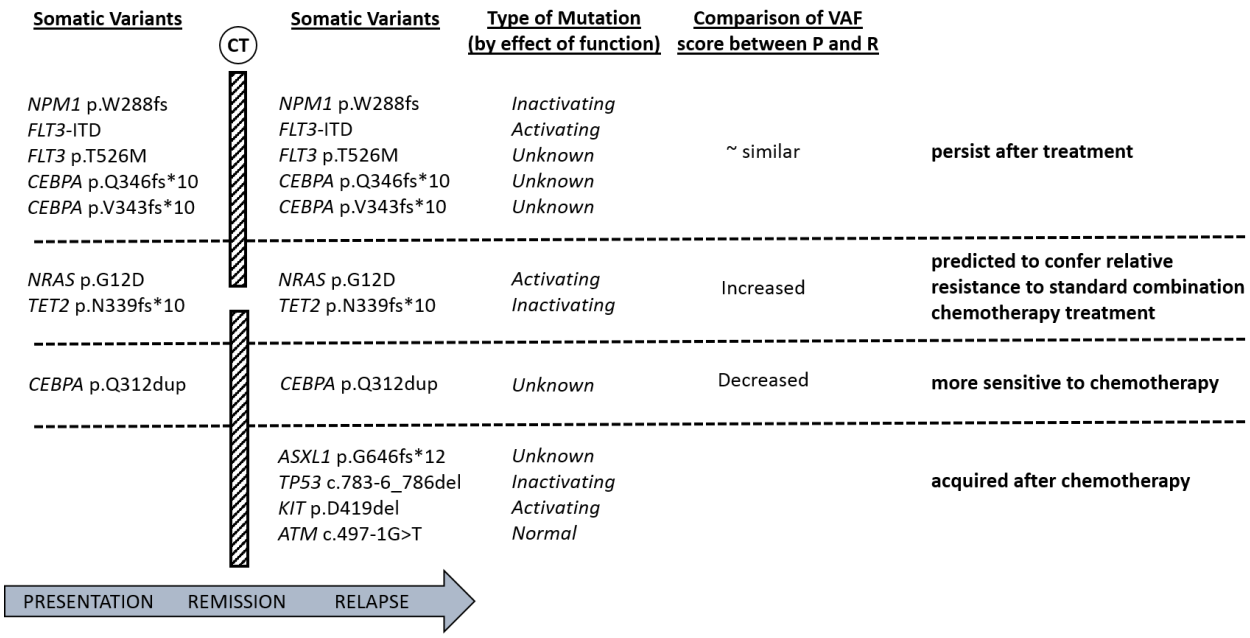


Figure 6.5. Effect of chemotherapy on the prevalence of somatic variants in known AML driver genes based on comparison of VAF score at presentation and relapse.

6.3. Discussion

Genetic heterogeneity is a major concern in cancer especially AML as it drives phenotypic adaptability which may contribute to clonal selection during treatment. One of the important causes of genetic heterogeneity is genomic instability. This instability leads to an increased mutation rate which can shape the evolution of the cancer genome through many mechanisms. Moreover, the diversity of genetic events in a cancer genome poses a significant challenge to identifying the oncogenic effect of specific genetic abnormalities. The advent of next generation sequencing (NGS) technology has allowed for more comprehensive analysis of somatic variants which has facilitated understanding of AML initiation and progression (Ding *et al.*, 2012; Welch *et al.*, 2012; Cancer Genome Atlas Research, 2013; Garg *et al.*, 2015; Sood *et al.*, 2015; Farrar *et al.*, 2016; Patel *et al.*, 2017).

In this study, high-throughput paired-end exome sequencing of 11 matched presentation and relapse AML samples was performed which enabled comparison of mutational landscape to better understand cancer evolution from presentation to relapse. The advantage of paired-end sequencing allowed both ends of a DNA fragment to be sequenced thus generating high-quality and alignable sequence data to facilitate detection of genomic variants with high confidence. However, one caveat with this approach is that the genome coverage was very low (the human exome represents ~1.2% of the total genome) and mutations in non-coding sequence were not identified. Despite that, with high read depth achieved in this study as a result of exonic sequence enrichment, identification of somatic variants present in minor leukaemic cell clones both at presentation and relapse was achieved.

Despite the sensitivity of NGS to identify low frequency mutations, this approach depends on good quality DNA with a high proportion of the leukaemic DNA (purity) in the sample to be sequenced. As such, poor quality sequencing or low purity can lead to false negative results, as evidenced by the failure to detect known *FLT3*-ITD and *NPM1* pW288fs (TCTG insertion) by exome sequencing in patients 1 and 9 that had previously been confirmed using targeted PCR. These mutations might have been present at very low frequency at presentation and were not detected via NGS due to low read depth. Unfortunately, further investigation of these mutations could not be performed due to the unavailability of raw sequencing data (.bam files) for these samples.

6.3.1. Somatic variants at presentation and relapse

SNVs and indels are the most abundant type of genetic variation in the human genome with ratios of indels to SNVs were 0.19 – 0.22 in WGS and 0.14 in WES (Mills *et al.*, 2011; Shigemizu *et al.*, 2013). Higher ratio of indels to SNVs were identified in this study with 0.82 at presentation and 0.66 at relapse. However, these data is incomparable since the former data was generated solely based on the prevalence of all genetic variation in the human genome whereas SNVs and indels identified in this study were strictly limited to somatic variants that may have biological implications in AML pathogenesis. Nevertheless, accurate identification of pathogenic SNVs and indels is one of the most important challenges in cancer genome analysis and requires detailed information of their impact on disease development. Ingenuity Variant Analysis software is equipped with linked databases that provides comprehensive information of each variant which can be used to discern functionality and a role in leukaemogenesis.

The average number of somatic variants (SNVs and indels) identified in this study were fewer compared to average number of mutation found in previous somatic mutation studies (Cancer Genome Atlas Research, 2013; Garg *et al.*, 2015; Sood *et al.*, 2015; Farrar *et al.*, 2016; Hartmann *et al.*, 2016). However, some previous studies had greater coverage of the genome (whole genome sequencing) and larger sample sizes. Despite this, the number of somatic mutations in AML is relatively low with an average of 13 mutations per genome as opposed to other cancers occurring in adults which often have hundreds of somatic mutations, especially solid tumours such as breast, lung or pancreas (Cancer Genome Atlas Research, 2013).

6.3.2. Coexisting variants within individual patients

The two-hit model of leukaemogenesis has served as a basis to understand how mutations in certain genes drive AML (Kelly and Gilliland, 2002). Based on this model, mutations were divided according to two fundamental characteristics of cancer cells, including mutations that confer cellular proliferation (class I mutations) and mutations that impair myeloid differentiation (class II mutations). The model predicts that both classes of mutation are required for AML transformation. However, genomic studies have revealed the presence of mutations in genes associated with epigenetic modification (*ASXL1*, *DNMT3A*, *IDH1/2*, *TET2*, *EZH2* and *MLL*) in a significant

proportion of AML patients (Shih *et al.*, 2012). Disruption of epigenetic processes including DNA methylation and histone modification can lead to altered function in genes involve in key cellular pathways such as DNA repair, RAS signalling, cell cycle and apoptosis which may cause malignant cellular transformation in cancer (Kanwal and Gupta, 2012). In addition, a genome-wide DNA methylation profiling study has demonstrated that specific methylation profiles are associated with specific AML subtypes (Figueroa *et al.*, 2010b), which identified oncogenic cooperativity between somatic alterations in epigenetic regulators and known AML driver mutations.

However, determining oncogenic cooperativity between mutations is difficult as it requires prior knowledge of the function of the genes as well as the impact of the mutated genes on the translated protein. Therefore, given the abundant number of mutated genes with uncertain significance identified in this cohort, analysis was limited to genes most recurrently mutated in AML as reported in COSMIC and genes not previously implicated in AML pathogenesis that have important functions potentially relevant for leukaemogenesis.

The mutational spectrum exhibited by each patient was totally different which demonstrates the heterogeneous nature and complexity of the genomic landscape in AML. However, despite this variation, more than half of patients (7 out of 11 patients) had mutations in at least one known AML driver genes that is reported as recurrently mutated in AML in COSMIC. *NPM1*, *FLT3*, *NRAS*, *CEBPA*, *TET2* and *IDH2* were the mutated genes identified in these patients at presentation which are well reported for their importance and contribution to AML pathogenesis. Co-occurrence of class I and II mutations was observed in three patients with normal karyotype at presentation (Patient 1, 5 and 9). *FLT3*-ITD and *NPM1* p.W288fs (TCTG insertion) mutations harboured by Patient 1 are inarguably defined as AML driver mutations since numerous studies has showed the co-occurrence of these mutations in AML patients (Krönke *et al.*, 2013; Garg *et al.*, 2015; Alpermann *et al.*, 2016; Patel *et al.*, 2017). In patient 5, *CEBPA* mutation appeared to be a driver mutation as it is recurrently mutated in AML patients (Pabst *et al.*, 2001; Fasan *et al.*, 2013) and plays a key role in AML initiation as demonstrated in murine models (Kirstetter *et al.*, 2008; Bereshchenko *et al.*, 2009). Although no additional mutations in known AML driver genes were detected in this patient, mutation in *MUTYH* gene appeared as a candidate mutation to cooperate with *CEBPA* mutation at presentation since its involvement in oxidative DNA damage repair (Banda *et al.*, 2017; Dumanski *et al.*, 2017) compared to another accompanying mutation in *CFTR* gene which was commonly reported in patients with

cystic fibrosis (Ortiz *et al.*, 2017; Furgeri *et al.*, 2018). *WRN* gene was recognised as tumour-suppressor gene and evidence suggests a role in promoting oncogenic proliferation (Agrelo *et al.*, 2006; Opresko *et al.*, 2007) which may likely contribute to AML transformation in patient 9 in addition to *NPM1* and *IDH2* mutations.

The mutational landscape observed in two patients with t(8;21) AML (Patient 6 and 7) provides further evidence for the complexity of AML, despite these leukaemias sharing a common driver lesion. It is well described that *RUNX1/ETO* fusion gene (a product of t(8;21)) alone is insufficient for leukaemogenesis, as it requires additional co-operating mutations for progression to AML. *KIT* and *FLT3* mutations are commonly found in t(8;21) patients (Boissel *et al.*, 2006; Santos *et al.*, 2011) and their oncogenic cooperativity for leukaemogenesis have been well established in murine models (Schessl *et al.*, 2005; Wichmann *et al.*, 2015). However, both patients lack these mutations which suggests that other mutations may also cooperate with t(8;21) to drive leukaemogenesis. *TET2* mutation is commonly associated with cytogenetically normal AML and concurrently observed with mutations in *NPM1*, *FLT3*, *JAK2*, *RUNX1*, *CEBPA*, *CBL* and *KRAS* (Weissmann *et al.*, 2011). Targeted amplicon sequencing of 46 leukaemia relevant genes in 56 t(8;21) patients has revealed *TET2* mutation in 12% of AML (Hartmann *et al.*, 2016). Furthermore, TET2 is predicted to act as a tumour-suppressor (Delhommeau *et al.*, 2008; Mohamedali *et al.*, 2009; Smith *et al.*, 2010; Mercher *et al.*, 2012) which might cooperates with the *RUNX1/ETO* fusion gene to cause AML phenotype in patient 6. Moreover, an interesting candidate gene, *RASA2* was identified in this patient as mutation in this gene is thought to have oncogenic effects due to its important role in suppressing RAS function (RAS signalling pathway) which control cell proliferation and differentiation (Arafeh *et al.*, 2015).

Collectively, these findings emphasise the necessity of further investigation of mutation cooperativity, especially for mutations in genes that does not belong to either class I or class II mutations. Indeed, it is particularly important to identify mutation cooperativity in patients that lack known recurrent combinations of mutations (Patient 3, 4, 8 and 11) as this might identify novel mechanisms of AML leukaemogenesis.

6.3.3. Clonal evolution from presentation to relapse

One major advantage of NGS approaches is that it can be used to quantitate the proportion of variant reads for any given mutation, also known as the variant allele fraction (VAF), which indicates the percentage of tumour cells that harbour a specific

mutation assuming a relatively pure leukaemic sample. Using this VAF score, each mutation at presentation can be tracked to map clones that persisted, those that resolved from presentation to relapse as well as novel mutations that emerged at relapse in order to delineate clonal evolution from presentation to relapse.

A relationship between the presentation and relapse samples was clearly observed in all patients studied. In most cases, mutations detected at presentation were also present at relapse with additional mutations (epitomised by patient 11; Figure 6.6A), except for one case (Patient 5) which had no additional mutation acquired at relapse. Despite no evidence of additional mutations at relapse in patient 5, the VAF score for *CEBPA* p.Q312dup was 40% higher than VAF scores of other two accompanying mutations which had rather similar VAF score at presentation, suggesting two major clones at presentation as illustrated in Figure 6.6B. The clone with *CEBPA* p.Q312dup, *CFTR* p.G622D, *MUTYH* p.E479* mutations persisted at relapse, suggesting relative resistance to chemotherapy, whereas the second clone with a *CEBPA* p.Q312dup mutation failed to survived chemotherapy.

The characterisation and study of acquired cytogenetic abnormalities, especially common translocations, has facilitated a better understanding of AML leukaemogenesis. The prevailing evidence identifies chromosomal translocations such as t(8;21), inv(16) and t(15;17) as early driver mutations in AML based on the fact that they are recurrently found in 6 to 15% of AML patients (Schoch and Haferlach, 2002) and promote mutagenesis that give rise to secondary mutations (Alcalay *et al.*, 2003; Krejci *et al.*, 2008; Araten *et al.*, 2013; Forster *et al.*, 2015). These chromosomal translocations are assumed to play an essential role in the early stages of leukaemogenesis as demonstrated in murine models which showed their cooperativity with other mutations to cause clonal expansion or block differentiation (Schessl *et al.*, 2005; Wichmann *et al.*, 2015). In addition, other studies have identified these AML-associated chromosomal translocations in non-leukaemic cells in a small number of healthy individuals which further supports their occurrence as early genetic events (Quina Ana *et al.*, 2000; Mori *et al.*, 2002). Given this evidence, characterising these genetic lesions in terms of order of acquisition would facilitate in understanding their role in leukaemogenesis. In this regard, there is evidence to suggest that the order of mutation acquisition can affect disease progression in AML (Ortmann *et al.*, 2013; Ortmann *et al.*, 2015; Sun *et al.*, 2016). In patients with chromosomal translocations, there are two possible patterns of clonal evolution suggested. Firstly, the leukaemic clone at presentation acquired additional somatic mutations and evolve into relapse

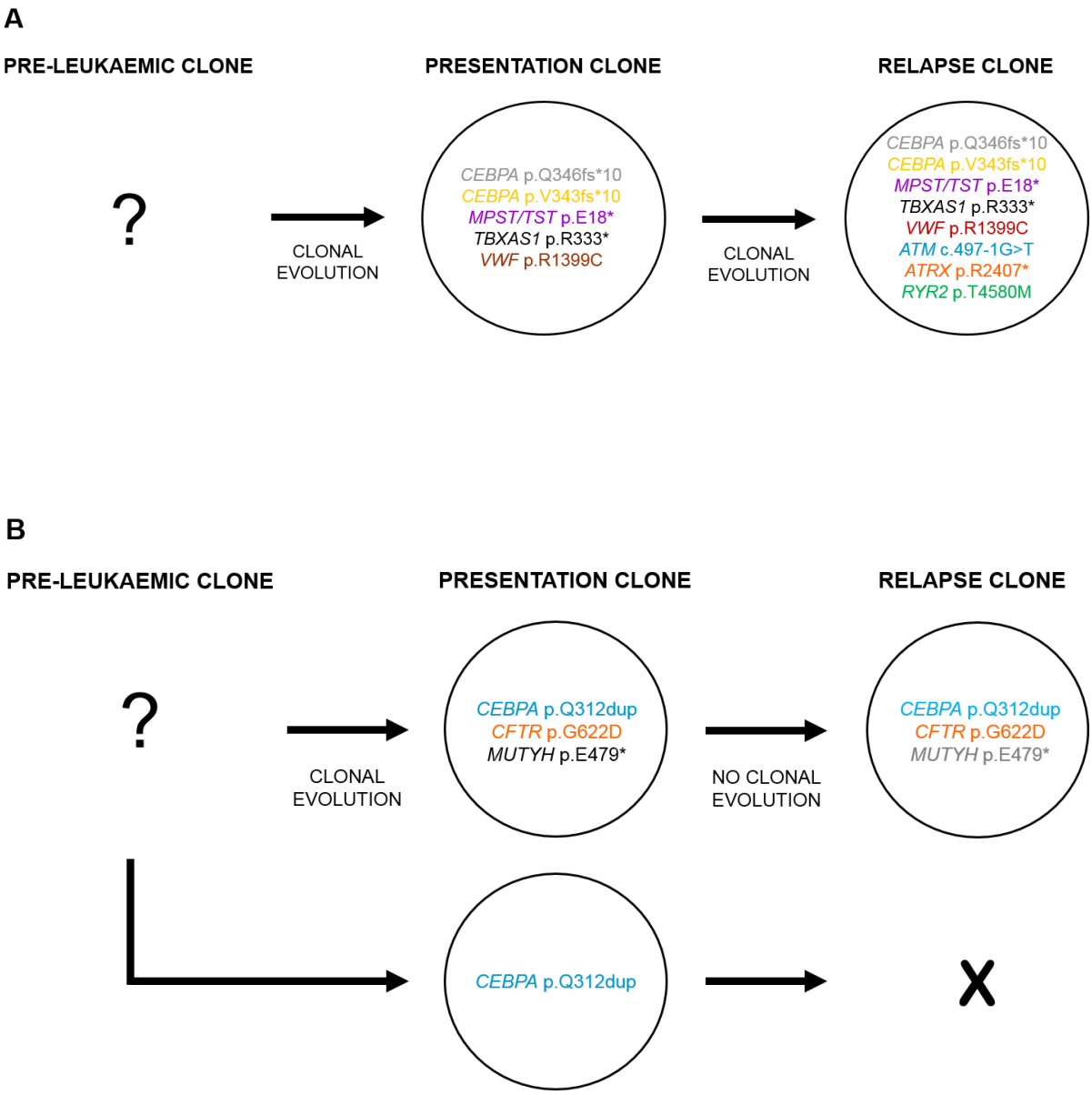


Figure 6.6. Clonal evolution of AML (continued on next page).

clone. Secondly, rather than evolving directly from the major presentation clone, the relapse clone is derived from an ancestral, pre-leukaemic clone. This is illustrated in the case of patient 6 (Figure 6.6C) where two major clones at relapse have been identified in which one additional mutation was acquired by the presentation clone and pre-leukaemic clone respectively after chemotherapy.

Collectively, the findings from this study suggests two possible mechanisms of relapse: (1) a subclone at disease presentation survived chemotherapy and re-emerged after acquired additional mutations; (2) a pre-leukaemic clone survived chemotherapy and acquired additional mutations during remission and gave rise to relapse. Both mechanisms showed that relapse is driven by the acquisition of additional somatic mutations. These findings are consistent with other previous studies investigating clonal evolution of AML (Ding *et al.*, 2012; Welch *et al.*, 2012; Garg *et al.*, 2015; Sood *et al.*, 2015; Greif *et al.*, 2018). However, in some cases, relapse might be driven by the same set of mutations acquired at presentation, suggesting these mutations confer selective advantage during AML treatment, resulting in clonal expansion and eventually leading to relapse. Sequence analysis of a remission sample is therefore essential to determine the VAF of these mutations at remission in order to confidently identify the pattern of clonal evolution.

Data from this study suggest that further investigations of the clonal origins of relapse in AML using high-throughput sequencing approaches are warranted in order to further understand the relationship between presentation and relapse AML. Critically, such studies should focus on the use of more samples sequenced using whole genome approaches to a greater depth. Ultimately, this approach will allow for the identification of relapse-driver mutations that could be exploited for the development of synthetic lethal interactions and novel therapies to improve outcomes and reduce death from disease.

C

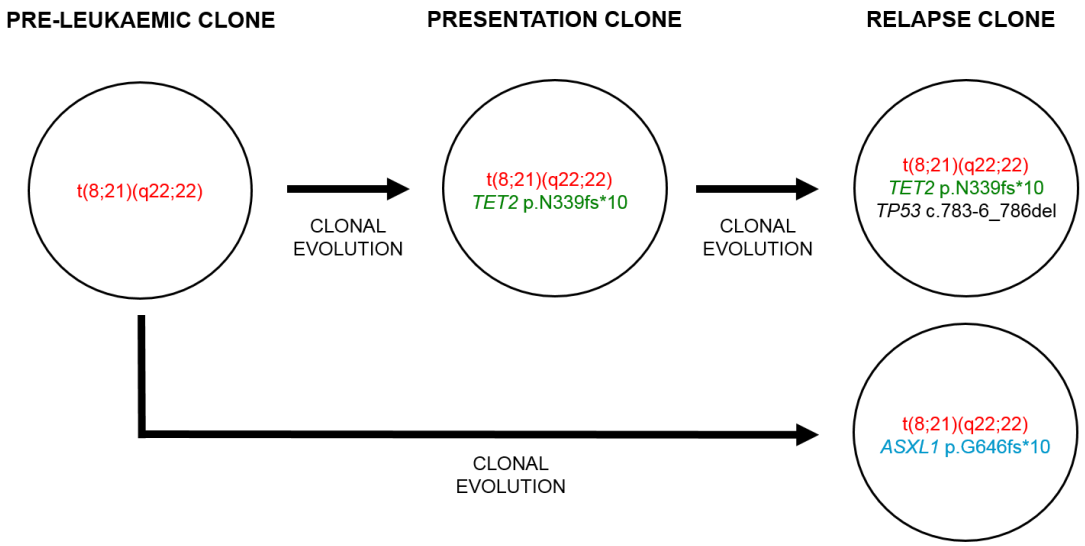


Figure 6.6. Clonal evolution of AML (continued from previous page).

A. Evolution from the major clone detected at disease presentation (patient 11).

B. Evolution from a pre-leukaemic clone detectable at disease presentation but resolved at disease relapse (patient 5).

C. Evolution from a pre-leukaemic clone detectable at disease presentation (patient 6).

6.3.4. Summary of chapter

In summary, these investigations have demonstrated that:

- High-throughput sequencing is a powerful approach that can provide in-depth analysis of the genome to identify somatic mutations in AML.
- In many cases, at least two driver mutations are required for AML transformation at presentation.
- Clonal evolution of relapsed AML is frequently characterised by the acquisition of additional somatic mutations after chemotherapy.
- Larger studies are required to fully delineate the heterogeneity of AML and disease evolution from presentation to relapse.

Chapter 7. Concluding Discussion

7.1. The Mutagenicity of Chemotherapeutic Agents used in AML Treatment

Despite significant progress in AML treatment, which includes alternative nucleoside analogues and targeted therapies, the current treatment of AML remains primarily dependent of ara-C and anthracycline containing chemotherapy, which function by inducing DNA damage and inhibiting DNA replication. Although these DNA damaging agents are efficacious in the treatment of AML, the majority of the patients experience relapse, leading to morbidity and mortality. The primary mechanism that contributes to disease recurrence or relapse is the acquired resistance of cancer cells to chemotherapy (Cheung-Ong *et al.*, 2013). Resistance to chemotherapy can arise in various ways, including mutations in genes involved in pathways that promote drug metabolism and degradation (Housman *et al.*, 2014). Emerging evidence demonstrates that the majority of genotoxic chemotherapeutic agents are mutagenic (Silva *et al.*, 2005; Pendleton *et al.*, 2014; Fordham *et al.*, 2015; Szikriszt *et al.*, 2016), demonstrating their ability to induce pro-mutagenic DNA damage which may give rise to mutations after chemotherapy.

The data from this study adds to a growing body of literature demonstrating that both ara-C and daunorubicin are mutagenic, as evidenced by increasing mutation frequency at doses that induce significant cytotoxicity (Chapter 3). In addition, the type of mutation induced by these chemotherapeutic agents is determined by their specific mechanism of action. Although the ara-C-induced mutational spectrum was not investigated in this project, a previous study has demonstrated that ara-C is a base substitution mutagen which preferentially induces mutation at G:C basepairs of $^{5'}$TpGpA$^{3'}$/$^{5'}$TpCpA$^{3'}$ sequences, consistent with its function as a cytosine analogue (Fordham *et al.*, 2015). This finding was supported by the interrogation of somatic base substitutions in presentation and relapsed AML samples, which demonstrated significant increase of mutations at the central base position of $^{5'}$TpGpA$^{3'}$/$^{5'}$TpCpA$^{3'}$ in relapse disease after treatment with ara-C-containing regimes (Fordham *et al.*, 2015). In contrast to ara-C, daunorubicin induces almost exclusively large deletions, consistent with its primary function as topo II enzyme inhibitor which resulted in DNA strand breaks (Chapter 3). Indeed, daunorubicin and related anthracyclines have been identified as direct inducers of DNA strand breaks (Blasiak *et al.*, 2002; Yang *et al.*, 2015) which can results in LOH or genetic rearrangements such as deletion and chromosomal translocation (Lehmann *et al.*, 2004). In addition, balanced translocations involving chromosome bands 11q23 (*MLL*) or 21q22 (*RUNX1*) observed

in t-AML have been associated with prior exposure to anthracyclines/topo II inhibitor-containing therapy (Pedersen-Bjergaard and Philip, 1991; Godley and Larson, 2008; Bhatia, 2013), demonstrating their ability to induce mutational events that drive leukaemogenesis. Taken together, these observations suggest that the use of pro-mutagenic agents during AML treatment contributes to genomic instability and causes mutations in key genes that drive relapsed AML.

7.2. Expression of the RUNX1/ETO Fusion Gene Promotes a Mutator Phenotype

Fusion genes are one of the most frequent types of somatic DNA alteration in cancer and have been recognised as important drivers in various human cancers, particularly haematological malignancies (Mertens *et al.*, 2015; Gao *et al.*, 2018). Gene fusion can be generated through structural chromosomal rearrangements such as translocations, deletions and inversions, which involves the co-localisation of coding or regulatory DNA sequences between previously separated genes. The identification of recurrent chromosomal translocations and their corresponding fusion genes in leukaemia has provided important insight into the underlying mechanisms involved in leukaemogenesis. The first fusion gene identified, *BCR/ABL1* (encoded by the Philadelphia (Ph) chromosome), results from a balanced translocation between chromosome 9 and 22 (t(9;22)(q34;q11) and was identified in CML (Nowell, 1960; Rowley, 1973; Klein *et al.*, 1982). The *BCR/ABL* fusion gene occurs in more than 95% of CML cases (Harrison, 2001; Frazer *et al.*, 2007; Bennour *et al.*, 2016). The *BCR/ABL1* fusion gene exerts its oncogenic influence by encoding a constitutively active tyrosine kinase that disrupt downstream signalling pathways, causing increased cell proliferation, blocked differentiation and evasion of apoptosis (Cilloni and Saglio, 2012; Hantschel, 2012; Sinclair *et al.*, 2013). The identification of the BCR/ABL fusion led to the first approved fusion-targeted drug (Imatinib) for newly diagnosed CML patients, which inhibits constitutive tyrosine kinase activity resulting in blocked transmission of proliferative signals to the nucleus and induced cell apoptosis (Schindler *et al.*, 2000; Cohen *et al.*, 2002). A considerable number of fusion genes now had been identified in leukaemia and many have served as important diagnostic and prognostic markers as they show specificity for distinct types of leukaemia. Many leukaemia-specific fusion genes represent therapeutic target for promising new treatments that could significantly improve patient outcomes.

In AML, t(15;17)(q22;q12), inv(16)(p13;q22) or t(16;16)(p13;q22), t(8;21)(q22;q22) and t(9;11)(p22;q23) are the most common chromosomal abnormalities that resulted in the formation of fusion genes in AML (Wang *et al.*, 2017). Of these, the *RUNX1/ETO* fusion gene is the most extensively studied fusion gene in AML, with a large body of literature investigating the molecular and cellular basis of this type of t(8;21) AML (Ferrara and Del Vecchio, 2002; Reikvam *et al.*, 2011; Solh *et al.*, 2014). Although patients with t(8;21) have relatively favourable prognosis compared to other types of AML, the presence of additional genetic abnormalities can impact on outcome. In addition, substantial numbers of t(8;21) AML patients eventually relapse, which is thought to be associated with the genetic heterogeneity of the leukaemic cells. It is well described that RUNX1/ETO expression alone is insufficient for leukaemogenesis and additional cooperating mutations are required (Yuan *et al.*, 2001; Higuchi *et al.*, 2002; Schessl *et al.*, 2005; Chou *et al.*, 2011; Wang *et al.*, 2011). However, the mechanism by which RUNX1/ETO driving the acquisition of additional cooperating mutations remains unclear, although evidence suggests that RUNX1/ETO promotes a mutator phenotype through downregulation of genes involved in several DNA repair pathways (Alcalay *et al.*, 2003; Krejci *et al.*, 2008; Forster *et al.*, 2015).

The proposed mutator phenotype theory in cancer suggests that the spontaneous mutation rate in normal cells is insufficient to cause large number of mutations observed in cancer cells, and other events are required to increase the spontaneous mutation rate and drive genomic instability (Loeb, 1991). The first direct evidence supporting this theory was demonstrated in cells from patients with HNPCC, which exhibit genomic instability in due to mutations in DNA MMR genes (Eshleman and Markowitz, 1996). In leukaemia, the *BCR/ABL1* fusion gene induces a mutator phenotype through inhibition of the MMR pathway, driving genomic instability and point mutations in tumour suppressor genes including *p53* and *Rb*, eventually leading to malignant progression of the disease (Stoklosa *et al.*, 2008). In addition, a recent study has demonstrated that expression of the RUNX1/ETO fusion protein predisposes cells to the acquisition of somatic mutations, both spontaneously and after exposure to genotoxic chemotherapy, with downregulation of OGG1 expression hypothesised as one possible mechanism (Forster *et al.*, 2015), providing evidence that RUNX1/ETO induces a mutator phenotype. Collectively, these studies showed that a mutator phenotype is associated with mutations in key genes and pathways such as DNA repair, which may promote an elevated mutation rate and increase the chances of acquiring subsequent genetic alterations.

As an extension from the Forster study (Forster *et al.*, 2015), the present study was undertaken to evaluate the influence of the *RUNX1/ETO* fusion gene on the mutagenesis, both spontaneously and after exposure to genotoxic agents. It was demonstrated that RUNX1/ETO significantly increases spontaneous mutation frequency at two independent loci (*TK* and *HPRT*) *in vitro*, despite no strong evidence that RUNX1/ETO expression renders cells more susceptible to mutagenesis after exposure to chemotherapeutic agents (Chapter 3). RUNX1/ETO fusion protein has been recognised as a leukaemia-initiating transcription factor that interferes with normal RUNX1 function, which includes deregulation of genes involved in haematopoiesis (Lam and Zhang, 2012) and repression of genes involved in several DNA repair pathways (Alcalay *et al.*, 2003; Wildonger and Mann, 2005; Krejci *et al.*, 2008; Forster *et al.*, 2015). One gene of particular interest is *OGG1*, which encodes a key DNA glycosylase involved in the excision of 8-oxoG lesions during base excision repair (Ba and Boldogh, 2018). Although the association of RUNX1/ETO and OGG1 was not investigated in this project, previous studies have provided strong evidence that OGG1 is downregulated in cells expressing RUNX1/ETO (Krejci *et al.*, 2008; Forster *et al.*, 2015). Moreover, *OGG1* gene is a direct transcriptional target of RUNX1/ETO fusion protein, which is negatively regulated by RUNX1/ETO (Ptasinska *et al.*, 2012; Forster *et al.*, 2015), providing a plausible mechanism by which RUNX1/ETO downregulates OGG1 expression and drives mutagenesis in t(8;21) AML. Impaired OGG1 DNA glycosylase activity can lead to the accumulation 8-oxoG lesion, giving rise to G:C > T:A transversion mutations (Suzuki *et al.*, 2010). Interrogation of the spontaneous base substitution mutational spectrum at the *HPRT* locus *in vitro* however showed no significant increase of G:C > T:A transversion in cells expressing RUNX1/ETO (Chapter 4). Instead, an increase of T:A > G:C transversions was observed, particularly at the central base position of $^{5'}$ApTpA$^{3'}$ / $^{5'}$TpApT$^{3'}$ trinucleotide sequences, although it was not statistically significant. However, the aetiology of T:A > G:C transversions is currently unknown and no mutational signature that exhibit the dominance of T:A > G:C transversions, particularly at $^{5'}$ApTpA$^{3'}$ / $^{5'}$TpApT$^{3'}$ sequences was previously reported in human cancer (https://cancer.sanger.ac.uk/cosmic/signatures). Moreover, there is lack of evidence for both G:C > T:A and T:A > G:C transversions in t(8;21) primary AML, which is dominated by C:G > T:A transitions. Interestingly, most of C:G > T:A transitions in primary t(8;21) AML occurred at the C base of $^{5'}$CpG$^{3'}$ dinucleotides sequences, which is associated with spontaneous deamination of 5-methylcytosine (Holliday and Grigg,

1993; Wiebauer *et al.*, 1993). The apparently contradictory findings between data from cell based reporter assays and and primary t(8;21) AML is likely influenced by the potential at specific DNA bases to affect protein function. Specifically, *HPRT* is considered a good reporter gene because mutation at over 60% of bases in the coding sequence affect protein function and can be detected using the reporter assay. In contrast, mutation in primary AML in genes such as *NPM1* and *FLT3* is observed at a relatively small number of bases, which affect protein function when mutated. As such, it is likely that mutation at the majority of bases in these and other leukaemia-specific genes is not observed because it is not functional. In this regard, many leukaemia-associated genes, such as *FLT3*, do not represent appropriate reporter genes with which to identify mutational signatures.

The data from this project has demonstrated that RUNX1/ETO expression predisposes to acquisition of the type A exon 12 *NPM1* mutation (c.863_864insTCTG), providing direct evidence that this fusion protein predisposes cells to the acquisition of somatic mutation at loci relevant to leukaemogenesis (Chapter 5). Although *NPM1* mutation is commonly associated with normal karyotype, it is reported at very low frequency in t(8;21) AML (Thiede *et al.*, 2006). Specifically, two NPM1 base substitutions have been reported in a patient with t(8;21), which were a T > G (W290G) transversion and a T > C (S293P) transition (Braoudaki *et al.*, 2010), with none of these mutations were previously reported in other AML subtype. Taken together, these data suggest that RUNX1/ETO drives genomic instability and mutagenesis at other loci in the genome, including leukaemia-relevant genes such as *NPM1*. It is well described that mutations in certain categories of genes, such as type III RTK family, co-operate with RUNX1/ETO to drive t(8;21) AML (Yuan *et al.*, 2001; Higuchi *et al.*, 2002). For example, mutations in *KIT* gene are found in up to 50% of t(8;21) AML cases, but are comparatively rare in other AML subtypes (Goemans *et al.*, 2005; Kohl *et al.*, 2005; Wang *et al.*, 2005). KIT mutation co-operates with RUNX1/ETO to drive AML in mice (Wang *et al.*, 2011). The most common *KIT* mutations reported in AML are N822K, D816H, D816V and D816Y, which are caused by T > G, G > C, A > T and G > T base substitutions, respectively (Care Rory *et al.*, 2003; Beghini *et al.*, 2004; Wang *et al.*, 2005; Roumier *et al.*, 2006). Of these, only KIT D816Y mutation is possibly caused by direct downregulation of OGG1 by RUNX1/ETO. It is hypothesised that downregulation of other genes involved in the BER pathway might responsible for other base substitution mutations in *KIT*. A study has shown that downregulation of the NEIL1 DNA glycosylase in Chinese hamster lung fibroblast V79 cells results in a higher rate

of spontaneous A > T base substitutions, causing amino acid change from aspartic asid (D) to valine (V) at the *HPRT* locus (Maiti *et al.*, 2008), which might explain the acquisition of D816V in *KIT* gene. Moreover, majority of the mutations occurred at A:T basepairs, indicating that NEIL1 has a preference for repair of oxidatively damaged A:T basepairs (Maiti *et al.*, 2008). However, there is no data on NEIL1 expression in t(8;21) AML. Hence, further studies are needed to confirm the downregulation of NEIL1 in RUNX1/ETO-expressing cells in order to fully understand the mechanisms by which RUNX1/ETO predisposes the acquisition of base substitutions in *KIT* as well as in other common cooperating mutation hotspot of RUNX1/ETO such as NRAS G12D (G > A).

7.3. Clonal Evolution and Clonal Heterogeneity in AML

Genomic diversity within single cancers has long be recognised. Some of the first evidence of genetic heterogeneity in leukaemia was reported by Anderson *et al*., (2010), who demonstrated low frequency of genetically distinct cell populations in propagating cells at leukaemia diagnosis. This clonal heterogeneity represents a key factor contributing to chemotherapy resistance by providing a reservoir of cells with the potential for clonal expansion leading to relapse. The advent of NGS technology has allowed for a more comprehensive analysis of cancer evolution and provided a better understanding of cancer initiation and progression. The exome sequencing data of matched presentation and relapsed AML samples in this project (Chapter 6) has provided insight on clonal evolution in AML, along with data from other studies (Ding *et al.*, 2012; Welch *et al.*, 2012; Garg *et al.*, 2015; Sood *et al.*, 2015; Greif *et al.*, 2018). Consistent with these studies, two major patterns of relapsed AML were identified from this project, with relapse evolving either from the major leukaemic clone at presentation, or from an ancestral, pre-leukaemic clone. Critically, both patterns of relapse are defined by the acquisition of additional somatic mutations. The fact that relapse is driven by additional acquired mutations acquired suggests that chemotherapeutic agents used in AML remission induction treatment either induces pro-mutagenic DNA damage or selects for pre-existing cells with mutations. However, in some cases, the dominant clone at disease presentation persists at relapse without the addition of novel somatic mutations, suggesting the presence of non-genetic alterations, such as epigenetic alterations, that might confer chemotherapy resistance. Taken together, the findings from this study suggest that the relapse-initiating clones may pre-exist before therapy, which harbour or acquire mutations that confer selective

advantage during chemotherapy, resulting in clonal expansion and eventually leading to relapse.

7.4. Summary of Findings

In summary, this project has demonstrated that chemotherapeutic agents used in AML remission induction treatment are mutagenic to DNA *in vitro*. Additionally, the influence of RUNX1/ETO fusion gene expression on mutagenesis has been demonstrated, with data suggesting that RUNX1/ETO confers a mutator phenotype by predisposing cells to the acquisition of additional mutations through downregulation of genes involved in DNA repair. Comparison of the mutational landscape between presentation and relapse AML samples has provided important insight into the clonal origins of relapsed AML, although additional studies are required to fully delineate the clonal evolution of this complex disease. Nevertheless, these data could prove valuable in the development of novel targeted therapies specifically tailored to individual patients with certain AML subgroups, thus limiting the use of genotoxic pro-mutagenic chemotherapy.

7.5. Limitations of the Study

Despite several major findings obtained from this study, there are few limitations that need to be addressed:

- The use of lymphoblastoid cells is perhaps not the best model to study myeloid specific mutations (Chapter 3 and 4). Specifically, consideration needs to be given to the expression of recombination-activating genes (RAGs) and other mechanisms intrinsic to B cells that have evolved to drive mutation at immunoglobulin loci, and how these might affect mutation at other sites in the genome, as described in Section 2.2.

- Ideally, antibodies raised against specific mutant proteins such as KIT D816V and RAS G12D should be used to study the effect of *RUNX1/ETO* fusion gene expression on the acquisition of mutations because these genes are relevant to t(8;21) AML pathogenesis (Chapter 5). In contrast, *NPM1* mutation is a very rare occurrence in t(8;21) AML and is therefore less relevant as mutational

target in RUNX1/ETO expressing cells. The selection of genes for study in this model was primarily determined by the availability of high specificity antibodies for mutant proteins. In addition, there is a concern that the increasing level of *NPM1* mutation over time which might represent a false-positive result as the intrinsic *NPM1* mutation is assumed to be very rare in normal cells especially in B cell lines. However, the data from the study demonstrated a reasonable increase in spontaneous *NPM1* mutation frequency over time as a result of continuous cell division in long term culture.

- The ability to quantify VAF is one of the major advantages of NGS technology that can provide information of the percentage of tumour cells that harbour a specific mutation (Chapter 6). However, VAF will vary according to the blast cell count and a high proportion of leukaemic DNA in the sample to be sequenced is essential to give accurate results. Therefore, a reasonable cut-off for major difference in VAF between diagnosis and relapse samples need to be determined for the changes to be considered significant, as applied in this study. In addition, strict selection criteria for the sample to be sequenced can be implemented for future studies to exclude samples with low blast cell count and samples contaminated with normal cells.

7.6. Future Directions

A number of further investigations are anticipated following the findings of the present study.

7.6.1. In vitro mutagenicity of other chemotherapeutic agents used in AML treatment

Exposure to daunorubicin and ara-C results in a significant dose-dependent increase in mutation frequency at both the *TK* and *HPRT* loci in TK6 cells, providing evidence on the *in vitro* mutagenicity of both chemotherapeutic agents (Chapter 3). In addition, mutation screening at the *HPRT* coding region in TK6 cells revealed that daunorubicin can induce a spectrum of mutations which are unique compared to spontaneous background mutation spectrum. Thus, the data from this study may form the basis for future studies to investigate the mutagenic potential of other

chemotherapeutic agents used in AML in order to discern their likely contribution to the aetiology of relapse driver mutations.

7.6.2. RUNX1/ETO-induced spontaneous mutation spectrum

Interrogation of the spontaneous *HPRT* base substitution spectrum revealed that RUNX1/ETO can induces unique mutation at the central base position of $^{5'}$ApTpA$^{3'}$ / $^{5'}$TpApT$^{3'}$ trinucleotide sequences, but this mutation was not evident in primary t(8;21) AML (Chapter 4). Therefore, a much larger number of independent spontaneous RUNX1/ETO-induced *HPRT* mutants need to be generated to expand the current dataset and improve statistical power to discern a RUNX1/ETO-induced mutation signature. Alternatively, whole genome sequencing of RUNX/ETO-positive cells can be carried out to identify all mutations in genes other than the leukaemia-relevant genes.

7.6.3. Effect of RUNX1/ETO gene expression on mutation frequency at specific loci relevant to AML pathogenesis

A reliable flow cytometric method was established to evaluate the effect of RUNX1/ETO fusion gene expression on mutation frequency at specific base residue in NPM1 (Chapter 5). It is now planned to perform similar mutation screening at other loci relevant to t(8;21) AML pathogenesis using commercially available antibodies raised against specific mutant proteins such as Kit D816V and RAS G12D.

7.6.4. Relationship between RUNX1/ETO and DNA repair gene expression

As most RUNX1/ETO cooperating mutations are base substitutions, an investigation on the association between RUNX1/ETO and DNA repair gene expression is warranted, with a particular focus on pathways that protect against base substitution mutation, such as BER and MMR. Such studies can be performed via quantification of transcript level of the candidate genes by real-time PCR (qPCR) or RNA-sequencing to also identify alternative transcripts that could affect repair function. Furthermore, ChIP sequencing can be used to identify DNA repair genes that are transcriptional targets of the RUNX1/ETO oncoprotein.

7.6.5. High-throughput sequencing of matched presentation and relapsed AML

Data from high-throughput exome sequencing of a small cohort of matched presentation and relapse AML samples has provided evidence of clonal evolution during disease progression (Chapter 6). Future studies should focus on the analysis of additional cytogenetically defined cases, including t(8;21), using whole genome approaches to a greater depth in order to fully delineate the heterogeneity of AML and disease evolution from presentation to relapse and fully elucidate the origins of relapse-initiating clones.

Appendices

Appendix A: Cytotoxicity in response to daunorubicin in TK6 and TK6 RUNX1/ETO cells.

The survival fraction at each dose was determined by calculating the mean CE of the dosed cells as a percentage of the mean CE of the vehicle-treated cells and each experiment was repeated three times.

Daunorubicin dose (nM)	TK6				TK6 RE8			
	Survival fraction (%)			Mean (%)	Survival fraction (%)			Mean (%)
	Exp 1	Exp 2	Exp 3		Exp 1	Exp 2	Exp 3	
0 (vehicle-treated)	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
2	60.12	59.90	79.60	66.54	61.40	94.22	86.09	80.57
4	22.88	18.91	12.96	18.25	30.81	56.64	38.84	41.98
6	0.30	0.55	0.46	0.44	2.25	5.15	4.86	4.09

Appendix B: Effect of daunorubicin exposure on mutation frequency at *TK* and *HPRT* loci in TK6 and TK6 RUNX1/ETO.

MF at the *TK* and *HPRT* loci was assessed in parental TK6 cells and TK6 *RUNX1/ETO* cells following a 16 hour exposure to daunorubicin. Each experiment was repeated three times.

Daunorubicin (nM)	*TK*							
	TK6				TK6 RUNX1/ETO			
	Exp 1	Exp 2	Exp 3	Mean MF ($x10^{-6}$)	Exp 1	Exp 2	Exp 3	Mean MF ($x10^{-6}$)
0	8.12	6.9	9.66	8.2	11.9	12.4	14.8	13.0
15	61.9	58.5	52.9	57.8	54.2	61.9	50.5	55.5
25	64.8	62.9	50.0	59.2	63.4	72.2	69.3	68.3
30	66.9	56.9	60.0	60.9	79.5	76.0	67.4	74.3

Daunorubicin (nM)	*HPRT*							
	TK6				TK6 RUNX1/ETO			
	Exp 1	Exp 2	Exp 3	Mean MF ($x10^{-6}$)	Exp 1	Exp 2	Exp 3	Mean MF ($x10^{-6}$)
0	1.9	2.3	4.4	2.9	5.2	3.4	3.4	4.0
15	30.6	29.1	19.3	26.3	15.6	25.1	15.4	18.7
25	41.6	39.9	34.1	38.5	27.8	36.9	35.0	33.2
30	43.9	39.6	35.5	39.6	33.7	44.7	43.8	40.7

Appendix C: Effect of ara-C exposure on mutation frequency at *TK* and *HPRT* loci in TK6 and TK6 RUNX1/ETO

MF at the *TK* and *HPRT* loci was assessed in parental TK6 cells and TK6 *RUNX1/ETO* cells following a 16 hour exposure to ara-C. Each experiment was repeated two times. 'x' denotes experimental replicate not done.

Ara-C (µM)	*TK*							
	TK6				TK6 RUNX1/ETO			
	Exp 1	Exp 2	Exp 3	Mean MF ($x10^{-6}$)	Exp 1	Exp 2	Exp 3	Mean MF ($x10^{-6}$)
0	8.4	10.6	x	9.5	14.4	14.6	x	14.5
1	42.9	33.8	x	38.4	38.3	35.1	x	36.7
3	48.3	44.8	x	46.6	48.6	38.7	x	43.7
5	56.0	49.8	x	52.9	58.9	48.1	x	53.5

Ara-C (µM)	*HPRT*							
	TK6				TK6 RUNX1/ETO			
	Exp 1	Exp 2	Exp 3	Mean MF ($x10^{-6}$)	Exp 1	Exp 2	Exp 3	Mean MF ($x10^{-6}$)
0	3.7	1.5	x	2.6	3.9	3.7	x	3.8
1	8.2	14.8	x	11.5	7.1	4.5	x	5.8
3	13.6	16.6	x	15.1	10.6	16.7	x	13.7
5	16.7	27.3	x	22.0	15.9	21.2	x	18.6

Appendix D: Sequence of *HPRT* cDNA (NCBI Reference Sequence: NM_000194).

GGCGGGGCCTGCTTCTCCTCAGCTTCAGGCGGCTGCGACGAGCCCTCAGGCGAACCTCTCGGCTTTCCCGCGC
GGCGCCGCCTCTTGCTGCGCCTCCGCCTCCTCCTCTGCTCCGCCACCGGCTTCCTCCTCCTGAGCAGCAGCCCGC
GCGCCGGCCGGCTCCGTTATGGCGACCCGCAGCCCTGGCGTCGTGATTAGTGATGATGAACCAGGTTATGACC
TTGATTTATTTTGCATACCTAATCATTATGCTGAGGATTTGGAAAGGGTGTTTATTCCTCATGGACTAATTATGG
ACAGGACTGAACGTCTTGCTCGAGATGTGATGAAGGAGATGGGAGGCCATCACATTGTAGCCCTCTGTGTGCT
CAAGGGGGGCTATAAATTCTTTGCTGACCTGCTGGATTACATCAAAGCACTGAATAGAAATAGTGATAGATCC
ATTCCTATGACTGTAGATTTTATCAGACTGAAGAGCTATTGTAATGACCAGTCAACAGGGGACATAAAAGTAAT
TGGTGGAGATGATCTCTCAACTTTAACTGGAAAGAATGTCTTGATTGTGGAAGATATAATTGACACTGGCAAA
ACAATGCAGACTTTGCTTTCCTTGGTCAGGCAGTATAATCCAAAGATGGTCAAGGTCGCAAGCTTGCTGGTGA
AAAGGACCCCACGAAGTGTTGGATATAAGCCAGACTTTGTTGGATTTGAAATTCCAGACAAGTTTGTTGTAGG
ATATGCCCTTGACTATAATGAATACTTCAGGGATTTGAATCATGTTTGTGTCATTAGTGAAACTGGAAAAGCAA
AATACAAAGCCTAAGATGAGAGTTCAAGTTGAGTTTGGAAACATCTGGAGTCCTATTGACATCGCCAGTAAAA
TTATCAATGTTCTAGTTCTGTGGCCATCTGCTTAGTAGAGCTTTTTGCATGTATCTTCTAAGAATTTTATCTGTTT
TGTACTTTAGAAATGTCAGTTGCTGCATTCCTAAACTGTTTATTTGCACTATGAGCCTATAGACTATCAGTTCCCT
TTGGGCGGATTGTTGTTTAACTTGTAAATGAAAAAATTCTCTTAAACCACAGCACTATTGAGTGAAACATTGAA
CTCATATCTGTAAGAAATAAAGAGAAGATATATTAGTTTTTTAATTGGTATTTTAATTTTTATATATGCAGGAAA
GAATAGAAGTGATTGAATATTGTTAATTATACCACCGTGTGTTAGAAAAGTAAGAAGCAGTCAATTTTCACATC
AAAGACAGCATCTAAGAAGTTTTGTTCTGTCCTGGAATTATTTTAGTAGTGTTTCAGTAATGTTGACTGTATTTT
CCAACTTGTTCAAATTATTACCAGTGAATCTTTGTCAGCAGTTCCCTTTTAAATGCAAATCAATAAATTCCCAAA
AATTTAAAAAAAAAAAAAAAAAAAA

Different coloured letters represent different exon of *HPRT* (exon 1 – 9).

Start codon is highlighted in light blue.

Forward primers are highlighted in green.

Reverse primers are highlighted in yellow.

Appendix E: Effect of RUNX1/ETO expression on *NPM1* mutation frequency over time (continued of next page).

Week	No RE expression				Low RE expression				High RE expression			
	Clone	Positive event	Gated event	MF ($x10^{-5}$)	Clone	Positive event	Gated event	MF ($x10^{-5}$)	Clone	Positive event	Gated event	MF ($x10^{-5}$)
7	1	6	350204	1.71	1	3	364383	0.82	1	4	330681	1.21
	2	2	304717	0.66	2	15	362736	4.14	2	8	338807	2.36
	3	5	230834	2.17	3	6	344971	1.74	3	11	353997	3.11
	4	6	344495	1.74	4	4	348021	1.15	4	7	338075	2.07
	5	6	310541	1.93	5	6	350092	1.71	5	5	328952	1.52
	6	8	348034	2.30	6	7	336006	2.08	6	5	398978	1.25
	7	11	445475	2.47	7	8	358257	2.23	7	9	338075	2.66
	8	3	333974	8.98	8	8	334301	2.39	8	7	341107	2.05
	9	7	294954	2.37	9	10	366220	2.73	9	3	339519	0.88
	10	6	330538	1.82	10	9	352337	2.55	10	5	346527	1.44
9	1	7	298812	2.34	1	4	333888	1.20	1	10	309892	3.23
	2	7	303499	2.31	2	6	398370	1.51	2	8	299325	2.67
	3	4	310888	1.29	3	6	300554	2.00	3	4	296842	1.35
	4	2	292176	6.85	4	7	326267	2.15	4	12	304630	3.94
	5	7	285463	2.45	5	6	311426	1.93	5	5	298305	1.68
	6	5	287081	1.74	6	5	337695	1.48	6	9	289189	3.11
	7	6	284843	2.11	7	7	322538	2.17	7	2	348788	5.73
	8	6	292781	2.05	8	9	288761	3.12	8	6	316825	1.89
	9	7	307843	2.27	9	9	313934	2.87	9	5	327654	1.53
	10	4	311012	1.29	10	7	318292	2.20	10	13	297746	4.37
11	1	8	341066	2.35	1	11	335360	3.28	1	17	298610	5.69
	2	6	347807	1.73	2	15	299735	5.00	2	15	317815	4.72
	3	6	344732	1.74	3	8	289410	2.76	3	12	363393	3.30
	4	8	320327	2.50	4	8	311793	2.57	4	9	316623	2.84
	5	8	310843	2.57	5	7	344427	2.03	5	13	329718	3.94

Appendix E: Effect of RUNX1/ETO expression on *NPM1* mutation frequency over time (continued of next page).

Week	No RE expression				Low RE expression				High RE expression			
	Clone	Positive event	Gated event	MF ($x10^{-5}$)	Clone	Positive event	Gated event	MF ($x10^{-5}$)	Clone	Positive event	Gated event	MF ($x10^{-5}$)
11	6	6	324241	1.85	6	9	303318	2.97	6	9	325869	2.76
	7	8	328703	2.43	7	7	293969	2.38	7	14	312942	4.47
	8	8	312070	2.56	8	5	323979	1.54	8	9	363760	2.47
	9	8	387374	2.07	9	12	313108	3.83	9	10	297726	3.36
	10	6	331766	1.81	10	13	308105	4.22	10	8	298210	2.68
13	1	11	407178	2.70	1	11	386466	2.85	1	16	413232	3.87
	2	10	337389	2.96	2	8	370393	2.16	2	22	371260	5.93
	3	6	305792	1.96	3	13	366749	3.54	3	10	395011	2.53
	4	8	349667	2.29	4	20	379940	5.26	4	22	360374	6.10
	5	9	372216	2.42	5	14	358878	3.90	5	11	363153	3.03
	6	6	358848	1.67	6	11	378992	2.90	6	14	363426	3.85
	7	15	351643	4.27	7	16	374996	4.27	7	17	422907	4.02
	8	15	327744	4.58	8	16	358565	4.46	8	12	430448	2.79
	9	9	322686	2.79	9	9	361026	2.49	9	17	401458	4.23
	10	13	318828	4.08	10	19	360878	5.26	10	13	378705	3.43
15	1	9	287685	3.13	1	12	332790	3.61	1	20	348763	5.73
	2	10	338996	2.95	2	7	339426	2.06	2	13	355045	3.66
	3	4	337841	1.18	3	10	325511	3.07	3	14	380378	3.68
	4	13	340043	3.82	4	11	349011	3.15	4	17	353870	4.80
	5	9	336717	2.67	5	8	316022	2.53	5	20	337855	5.92
	6	12	318583	3.77	6	8	330095	2.42	6	6	333631	1.80
	7	12	311133	3.86	7	12	327508	3.66	7	8	312825	2.56
	8	7	324698	2.16	8	19	345937	5.49	8	13	314827	4.13
	9	10	312154	3.20	9	18	398135	4.52	9	10	315013	3.17
	10	11	344704	3.19	10	11	367228	3.00	10	15	326951	4.59

Appendix E: Effect of RUNX1/ETO expression on *NPM1* mutation frequency over time (continued of next page).

Week	No RE expression				Low RE expression				High RE expression			
	Clone	Positive event	Gated event	MF ($x10^{-5}$)	Clone	Positive event	Gated event	MF ($x10^{-5}$)	Clone	Positive event	Gated event	MF ($x10^{-5}$)
18	1	21	319573	6.57	1	18	308130	5.84	1	18	324954	5.54
	2	18	318644	5.65	2	23	333136	6.90	2	18	328741	5.48
	3	8	315535	2.54	3	23	307127	7.49	3	15	318388	4.71
	4	9	309959	2.90	4	22	338473	6.50	4	18	290089	6.20
	5	16	346933	4.61	5	11	317899	3.46	5	19	325623	5.83
	6	13	309164	4.20	6	7	250469	2.79	6	12	326856	3.67
	7	16	336416	4.76	7	14	307903	4.55	7	12	249676	4.81
	8	21	317874	6.61	8	16	250172	6.40	8	18	312548	5.76
	9	15	324607	4.62E	9	7	160056	4.37	9	15	263229	5.70
	10	12	304760	3.94	10	10	248454	4.02	10	18	313451	5.74
20	1	12	318933	3.76	1	18	338183	5.32	1	27	285419	9.46
	2	16	321044	4.98	2	21	340392	6.17	2	24	307006	7.82
	3	15	306725	4.89	3	20	316990	6.31	3	30	328896	9.12
	4	13	332907	3.90	4	19	330064	5.76	4	19	298312	6.37
	5	16	309415	5.17	5	17	323468	5.26	5	29	353932	8.19
	6	19	312786	6.07	6	15	339759	4.41	6	21	327920	6.40
	7	18	312615	5.76	7	18	327902	5.49	7	13	301587	4.31
	8	23	309054	7.44	8	18	323864	5.56	8	15	315106	4.76
	9	10	297810	3.36	9	18	334025	5.39	9	21	316569	6.63
	10	14	300481	4.66	10	12	340186	3.53	10	21	347717	6.04
22	1	27	329623	8.19	1	24	289968	8.28	1	35	299240	1.17
	2	12	300051	4.00	2	27	324988	8.31	2	21	337030	6.23
	3	18	315816	5.70	3	21	310745	6.76	3	10	146062	6.85
	4	18	310723	5.79	4	29	329132	8.81	4	23	239753	9.59
	5	17	322712	5.27	5	14	319618	4.38	5	25	337503	7.41

Appendix E: Effect of RUNX1/ETO expression on *NPM1* mutation frequency over time (continued from previous page).

Week	No RE expression				Low RE expression				High RE expression			
	Clone	Positive event	Gated event	MF ($x10^{-5}$)	Clone	Positive event	Gated event	MF ($x10^{-5}$)	Clone	Positive event	Gated event	MF ($x10^{-5}$)
22	6	21	320773	6.55	6	26	330786	7.86	6	35	319203	1.10
	7	15	322585	4.65	7	20	313897	6.37	7	25	331562	7.54
	8	18	336046	5.36	8	21	315378	6.66	8	20	313312	6.38
	9	23	330868	6.95	9	23	327640	7.02	9	26	312200	8.33
	10	12	310718	3.86	10	28	336254	8.33	10	19	306382	6.20

Detection of NPM1-mutant positive event was performed every 2 weeks, starting at week 7 to week 22 post-cloning using flow cytometry. The frequency (MF) of *NPM1* exon 12 mutation in no RE expression subclones, low RE expression subclones and high RE expression subclones was calculated by dividing the no. of positive event with the no. of gated event.

References

Aaron, C.S., Bolcsfoldi, G., Glatt, H.R., Moore, M., Nishi, Y., Stankowski, L., Theiss, J. and Thompson, E. (1994) 'Mammalian cell gene mutation assays working group report', *Mutation Research/Environmental Mutagenesis and Related Subjects*, 312(3), pp. 235-239.

Abdel-Wahab, O., Adli, M., LaFave, Lindsay M., Gao, J., Hricik, T., Shih, Alan H., Pandey, S., Patel, Jay P., Chung, Young R., Koche, R., Perna, F., Zhao, X., Taylor, Jordan E., Park, Christopher Y., Carroll, M., Melnick, A., Nimer, Stephen D., Jaffe, Jacob D., Aifantis, I., Bernstein, Bradley E. and Levine, Ross L. (2012) 'ASXL1 Mutations Promote Myeloid Transformation through Loss of PRC2-Mediated Gene Repression', *Cancer Cell*, 22(2), pp. 180-193.

Abdel-Wahab, O. and Levine, R.L. (2013) 'Mutations in epigenetic modifiers in the pathogenesis and therapy of acute myeloid leukemia', *Blood*, 121(18), p. 3563.

Agrelo, R., Cheng, W.-H., Setien, F., Ropero, S., Espada, J., Fraga, M.F., Herranz, M., Paz, M.F., Sanchez-Cespedes, M., Artiga, M.J., Guerrero, D., Castells, A., von Kobbe, C., Bohr, V.A. and Esteller, M. (2006) 'Epigenetic inactivation of the premature aging Werner syndrome gene in human cancer', *Proceedings of the National Academy of Sciences of the United States of America*, 103(23), pp. 8822-8827.

Ahn, E.-Y., Yan, M., Malakhova, O.A., Lo, M.-C., Boyapati, A., Ommen, H.B., Hines, R., Hokland, P. and Zhang, D.-E. (2008) 'Disruption of the NHR4 domain structure in AML1-ETO abrogates SON binding and promotes leukemogenesis', *Proceedings of the National Academy of Sciences of the United States of America*, 105(44), pp. 17103-17108.

Alcalay, M., Meani, N., Gelmetti, V., Fantozzi, A., Fagioli, M., Orleth, A., Riganelli, D., Sebastiani, C., Cappelli, E., Casciari, C., Sciurpi, M.T., Mariano, A.R., Minardi, S.P., Luzi, L., Muller, H., Di Fiore, P.P., Frosina, G. and Pelicci, P.G. (2003) 'Acute myeloid leukemia fusion proteins deregulate genes involved in stem cell maintenance and DNA repair', *The Journal of Clinical Investigation*, 112(11), pp. 1751-1761.

Alexandrov, L.B., Nik-Zainal, S., Siu, H.C., Leung, S.Y. and Stratton, M.R. (2015) 'A mutational signature in gastric cancer suggests therapeutic strategies', *Nat Commun*, 6.

Alexandrov, L.B., Nik-Zainal, S., Wedge, D.C., Aparicio, S.A.J.R., Behjati, S., Biankin, A.V., Bignell, G.R., Bolli, N., Borg, A., Borresen-Dale, A.-L., Boyault, S., Burkhardt, B., Butler, A.P., Caldas, C., Davies, H.R., Desmedt, C., Eils, R., Eyfjord, J.E., Foekens, J.A., Greaves, M., Hosoda, F., Hutter, B., Ilicic, T., Imbeaud, S., Imielinsk, M., Jager, N., Jones, D.T.W., Jones, D., Knappskog, S., Kool, M., Lakhani, S.R., Lopez-Otin, C., Martin, S., Munshi, N.C., Nakamura, H., Northcott, P.A., Pajic, M., Papaemmanuil, E., Paradiso, A., Pearson, J.V., Puente, X.S., Raine, K., Ramakrishna, M., Richardson, A.L., Richter, J., Rosenstiel, P., Schlesner, M., Schumacher, T.N., Span, P.N., Teague, J.W., Totoki, Y., Tutt, A.N.J., Valdes-Mas, R., van Buuren, M.M., van /'t Veer, L., Vincent-Salomon, A., Waddell, N., Yates, L.R., Australian Pancreatic Cancer Genome, I., Consortium, I.B.C., Consortium, I.M.-S., PedBrain, I., Zucman-Rossi, J., Andrew Futreal, P., McDermott, U., Lichter, P., Meyerson, M., Grimmond, S.M., Siebert, R., Campo, E., Shibata, T., Pfister, S.M., Campbell, P.J. and Stratton, M.R. (2013a) 'Signatures of mutational processes in human cancer', *Nature*, 500(7463), pp. 415-421.

Alexandrov, Ludmil B., Nik-Zainal, S., Wedge, David C., Campbell, Peter J. and Stratton, Michael R. (2013b) 'Deciphering Signatures of Mutational Processes Operative in Human Cancer', *Cell Reports*, 3(1), pp. 246-259.

Alpermann, T., Schnittger, S., Eder, C., Dicker, F., Meggendorfer, M., Kern, W., Schmid, C., Aul, C., Staib, P., Wendtner, C.-M., Schmitz, N., Haferlach, C. and Haferlach, T. (2016) 'Molecular subtypes of NPM1 mutations have different clinical profiles, specific patterns of accompanying molecular mutations and varying outcomes in intermediate risk acute myeloid leukemia', *Haematologica*, 101(2), p. e55.

Amundson, S.A., Myers, T.G. and Fornace Jr, A.J. (1999) 'Roles for p53 in growth arrest and apoptosis: putting on the brakes after genotoxic stress', *Oncogene*, 17, p. 3287.

Anderson, R.D., Veigl, M.L., Baxter, J. and Sedwick, W.D. (1991) 'DNA Sequence Specificity of Doxorubicin-induced Mutational Damage in Escherichia coli', *Cancer Research*, 51(15), p. 3930.

Appelbaum, F.R. and Bernstein, I.D. (2017) 'Gemtuzumab ozogamicin for acute myeloid leukemia', *Blood*, 130(22), p. 2373.

Arafeh, R., Qutob, N., Emmanuel, R., Keren-Paz, A., Madore, J., Elkahloun, A., Wilmott, J.S., Gartner, J.J., Di Pizio, A., Winograd-Katz, S., Sindiri, S., Rotkopf, R., Dutton-Regester, K., Johansson, P., Pritchard, A., Waddell, N., Hill, V.K., Lin, J.C., Hevroni, Y., Rosenberg, S.A., Khan, J., Ben-Dor, S., Niv, M.Y., Ulitsky, I., Mann, G.J., Scolyer, R.A., Hayward, N.K. and Samuels, Y. (2015) 'Recurrent inactivating RASA2 mutations in melanoma', *Nature genetics*, 47(12), pp. 1408-1410.

Araten, D.J., Krejci, O., DiTata, K., Wunderlich, M., Sanders, K.J., Zamechek, L. and Mulloy, J.C. (2013) 'The rate of spontaneous mutations in human myeloid cells', *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 749(1–2), pp. 49-57.

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

Arcamone, F., Animati, F., Capranico, G., Lombardi, P., Pratesi, G., Manzini, S., Supino, R. and Zunino, F. (1997) 'New developments in antitumor anthracyclines', *Pharmacology & Therapeutics*, 76(1), pp. 117-124.

Arnaudeau, C., Lundin, C. and Helleday, T. (2001) 'DNA double-strand breaks associated with replication forks are predominantly repaired by homologous recombination involving an exchange mechanism in mammalian cells11Edited by J. Karn', *Journal of Molecular Biology*, 307(5), pp. 1235-1245.

Arnaudeau, C., Tenorio Miranda, E., Jenssen, D. and Helleday, T. (2000) 'Inhibition of DNA synthesis is a potent mechanism by which cytostatic drugs induce homologous recombination in mammalian cells', *Mutation Research/DNA Repair*, 461(3), pp. 221-228.

Ba, X. and Boldogh, I. (2018) '8-Oxoguanine DNA glycosylase 1: Beyond repair of the oxidatively modified base lesions', *Redox Biology*, 14, pp. 669-678.

Bacher, U., Haferlach, C., Kern, W., Haferlach, T. and Schnittger, S. (2007) 'Prognostic relevance of FLT3-TKD mutations in AML: the combination matters—an analysis of 3082 patients', *Blood*, 111(5), pp. 2527-2537.

Bachur, N.R., Gordon, S.L. and Gee, M.V. (1978) 'A General Mechanism for Microsomal Activation of Quinone Anticancer Agents to Free Radicals', *Cancer Research*, 38(6), p. 1745.

Banda, D.M., Nuñez, N.N., Burnside, M.A., Bradshaw, K.M. and David, S.S. (2017) 'Repair of 8-oxoG:A mismatches by the MUTYH glycosylase: Mechanism, metals and medicine', *Free Radical Biology and Medicine*, 107, pp. 202-215.

Banker, D.E., Radich, J., Becker, A., Kerkof, K., Norwood, T., Willman, C. and Appelbaum, F.R. (1998) 'The t(8;21) translocation is not consistently associated with high Bcl-2 expression in de novo acute myeloid leukemias of adults', *Clinical Cancer Research*, 4(12), p. 3051.

Basecke, J., Cepek, L., Mannhalter, C., Krauter, J., Hildenhagen, S., Brittinger, G., Trumper, L. and Griesinger, F. (2002) 'Transcription of AML1/ETO in bone marrow and cord blood of individuals without acute myelogenous leukemia', *Blood*, 100(6), p. 2267.
Bates, D.A. and Winterbourn, C.C. (1982) 'Deoxyribose breakdown by the adriamycin semiquinone and H2O2: evidence for hydroxyl radical participation', *FEBS Letters*, 145(1), pp. 137-142.

Beaupre, D.M. and Kurzrock, R. (1999) 'RAS and Leukemia: From Basic Mechanisms to Gene-Directed Therapy', *Journal of Clinical Oncology*, 17(3), pp. 1071-1071.

Becker, H., Marcucci, G., Maharry, K., Radmacher, M.D., Mrózek, K., Margeson, D., Whitman, S.P., Wu, Y.-Z., Schwind, S., Paschka, P., Powell, B.L., Carter, T.H., Kolitz, J.E., Wetzler, M., Carroll, A.J., Baer, M.R., Caligiuri, M.A., Larson, R.A. and Bloomfield, C.D. (2010) 'Favorable Prognostic Impact of NPM1 Mutations in Older Patients With Cytogenetically Normal De Novo Acute Myeloid Leukemia and Associated Gene- and MicroRNA-Expression Signatures: A Cancer and Leukemia Group B Study', *Journal of Clinical Oncology*, 28(4), pp. 596-604.

Beghini, A., Peterlongo, P., Ripamonti, C.B., Larizza, L., Cairoli, R., Morra, E. and Mecucci, C. (2000) 'C-kit mutations in core binding factor leukemias', *Blood*, 95(2), pp. 726-728.

Beghini, A., Ripamonti, C.B., Cairoli, R., Cazzaniga, G., Colapietro, P., Elice, F., Nadali, G., Grillo, G., Haas, O.A., Biondi, A., Morra, E. and Larizza, L. (2004) 'KIT activating mutations: incidence in adult and pediatric acute myeloid leukemia, and identification of an internal tandem duplication', *Haematologica*, 89(8), p. 920.

Bennett, J.M., Catovsky, D., Daniel, M.T., Flandrin, G., Galton, D.A., Gralnick, H.R. and Sultan, C. (1976) 'Proposals for the classification of the acute leukaemias. French-American-British (FAB) co-operative group', (0007-1048 (Print)).

Bennour, A., Saad, A. and Sennana, H. (2016) 'Chronic myeloid leukemia: Relevance of cytogenetic and molecular assays', *Critical Reviews in Oncology/Hematology*, 97, pp. 263-274.

Bereshchenko, O., Mancini, E., Moore, S., Bilbao, D., Månsson, R., Luc, S., Grover, A., Jacobsen, S.E.W., Bryder, D. and Nerlov, C. (2009) 'Hematopoietic Stem Cell Expansion Precedes the Generation of Committed Myeloid Leukemia-Initiating Cells in C/EBPα Mutant AML', *Cancer Cell*, 16(5), pp. 390-400.

Betz, B.L. and Hess, J.L. (2010) 'Acute Myeloid Leukemia Diagnosis in the 21st Century', *Archives of Pathology & Laboratory Medicine*, 134(10), pp. 1427-1433.

Bhatia, S. (2013) 'Therapy-related myelodysplasia and acute myeloid leukemia', *Seminars in oncology*, 40(6), p. 10.1053/j.seminoncol.2013.09.013.

Bishop, A.J.R. and Schiestl, R.H. (2001) 'Homologous recombination as a mechanism of carcinogenesis', *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer*, 1471(3), pp. M109-M121.

Blasiak, J., Gloc, E. and Warszawski, W. (2002) 'A comparison of the in vitro genotoxicity of anticancer drugs idarubicin and mitoxantrone', *Acta Biochim Pol*, 49(1), pp. 145-155.

Blume-Jensen, P. and Hunter, T. (2001) 'Oncogenic kinase signalling', *Nature*, 411(6835), pp. 355-365.

Boissel, N., Leroy, H., Brethon, B., Philippe, N., de Botton, S., Auvrignon, A., Raffoux, E., Leblanc, T., Thomas, X., Hermine, O., Quesnel, B., Baruchel, A., Leverger, G., Dombret, H. and Preudhomme, C. (2006) 'Incidence and prognostic impact of c-Kit, FLT3, and Ras gene mutations in core binding factor acute myeloid leukemia (CBF-AML)', *Leukemia*, 20, p. 965.

Boiteux, S. and Radicella, J.P. (2000) 'The Human OGG1 Gene: Structure, Functions, and Its Implication in the Process of Carcinogenesis', *Archives of Biochemistry and Biophysics*, 377(1), pp. 1-8.

Boultwood, J., Perry, J., Pellagatti, A., Fernandez-Mercado, M., Fernandez-Santamaria, C., Calasanz, M.J., Larrayoz, M.J., Garcia-Delgado, M., Giagounidis, A., Malcovati, L., Della Porta, M.G., Jädersten, M., Killick, S., Hellström-Lindberg, E., Cazzola, M. and Wainscoat, J.S. (2010) 'Frequent mutation of the polycomb-associated gene ASXL1 in the myelodysplastic syndromes and in acute myeloid leukemia', *Leukemia*, 24, p. 1062.

Box, J.K., Paquet, N., Adams, M.N., Boucher, D., Bolderson, E., O'Byrne, K.J. and Richard, D.J. (2016) 'Nucleophosmin: from structure and function to disease development', *BMC Molecular Biology*, 17(1), p. 19.

Braoudaki, M., Papathanassiou, C., Katsibardi, K., Tourkadoni, N., Karamolegou, K. and Tzortzatou-Stathopoulou, F. (2010) 'The frequency of NPM1 mutations in childhood acute myeloid leukemia', *Journal of Hematology & Oncology*, 3, pp. 41-41.
Brash, D.E. (2015) 'UV Signature Mutations', *Photochemistry and photobiology*, 91(1), pp. 15-26.

Brash, D.E., Rudolph, J.A., Simon, J.A., Lin, A., McKenna, G.J., Baden, H.P., Halperin, A.J. and Pontén, J. (1991) 'A role for sunlight in skin cancer: UV-induced p53 mutations in squamous cell carcinoma', *Proceedings of the National Academy of Sciences of the United States of America*, 88(22), pp. 10124-10128.

Burgess, A.W. and Metcalf, D. (1980) 'The nature and action of granulocyte-macrophage colony stimulating factors', *Blood*, 56(6), p. 947.

Burnett, A.K., Russell, N.H., Hills, R.K., Hunter, A.E., Kjeldsen, L., Yin, J., Gibson, B.E.S., Wheatley, K. and Milligan, D. (2013) 'Optimization of Chemotherapy for Younger Patients With Acute Myeloid Leukemia: Results of the Medical Research Council AML15 Trial', *Journal of Clinical Oncology*, 31(27), pp. 3360-3368.

Burnett, A.K. and Venditti, A. (2011) 'Acute myeloid leukaemia', in Hoffbrand, V.A., Catovsky, D., Tuddenham, E.G.D. and Green, A.R. (eds.) *Postgraduate Haematology*. 6th edn. Blackwell Publishing Ltd., p. 415.

Burrell, R.A., McGranahan, N., Bartek, J. and Swanton, C. (2013) 'The causes and consequences of genetic heterogeneity in cancer evolution', *Nature*, 501, p. 338.

Cancer Genome Atlas Research, N. (2013) 'Genomic and Epigenomic Landscapes of Adult De Novo Acute Myeloid Leukemia', *New England Journal of Medicine*, 368(22), pp. 2059-2074.

Care Rory, S., Valk Peter, J.M., Goodeve Anne, C., Abu-Duhier Faisel, M., Geertsma-Kleinekoort Wendy, M.C., Wilson Giu, A., Gari Mamdooh, A., Peake Ian, R., Löwenberg, B. and Reilly John, T. (2003) 'Incidence and prognosis of c-KIT and FLT3 mutations in core binding factor (CBF) acute myeloid leukaemias', *British Journal of Haematology*, 121(5), pp. 775-777.

Champoux, J.J. (2001) 'DNA Topoisomerases: Structure, Function, and Mechanism', *Annual Review of Biochemistry*, 70(1), pp. 369-413.

Chen, H. and Shaw, B.R. (1993) 'Kinetics of bisulfite-induced cytosine deamination in single-stranded DNA', *Biochemistry*, 32(14), pp. 3535-3539.

Chen, J., Odenike, O. and Rowley, J.D. (2010) 'Leukaemogenesis: more than mutant genes', *Nature Reviews Cancer*, 10, p. 23.

Cheng, K., Sportoletti, P., Ito, K., Clohessy, J.G., Teruya-Feldstein, J., Kutok, J.L. and Pandolfi, P.P. (2010) 'The cytoplasmic NPM mutant induces myeloproliferation in a transgenic mouse model', *Blood*, 115(16), p. 3341.

Cheng, K.C., Cahill, D.S., Kasai, H., Nishimura, S. and Loeb, L.A. (1992) '8-Hydroxyguanine, an abundant form of oxidative DNA damage, causes G----T and A----C substitutions', *Journal of Biological Chemistry*, 267(1), pp. 166-172.

Cheung-Ong, K., Giaever, G. and Nislow, C. (2013) 'DNA-Damaging Agents in Cancer Chemotherapy: Serendipity and Chemical Biology', *Chemistry & Biology*, 20(5), pp. 648-659.

Chi, P., Allis, C.D. and Wang, G.G. (2010) 'Covalent histone modifications — miswritten, misinterpreted and mis-erased in human cancers', *Nature Reviews Cancer*, 10, p. 457.

Chou, F.-S., Wunderlich, M., Griesinger, A. and Mulloy, J.C. (2011) 'N-Ras(G12D) induces features of stepwise transformation in preleukemic human umbilical cord blood cultures expressing the AML1-ETO fusion gene', *Blood*, 117(7), pp. 2237-2240.

Chou, W.-C., Huang, H.-H., Hou, H.-A., Chen, C.-Y., Tang, J.-L., Yao, M., Tsay, W., Ko, B.-S., Wu, S.-J., Huang, S.-Y., Hsu, S.-C., Chen, Y.-C., Huang, Y.-N., Chang, Y.-C., Lee, F.-Y., Liu, M.-C., Liu, C.-W., Tseng, M.-H., Huang, C.-F. and Tien, H.-F. (2010) 'Distinct clinical and biological features of de novo acute myeloid leukemia with additional sex comb-like 1 (ASXL1) mutations', *Blood*, 116(20), p. 4086.

Chrencik, J.E., Burgin, A.B., Pommier, Y., Stewart, L. and Redinbo, M.R. (2003) 'Structural Impact of the Leukemia Drug 1-β-d-Arabinofuranosylcytosine (Ara-C) on the Covalent Human Topoisomerase I-DNA Complex', *Journal of Biological Chemistry*, 278(14), pp. 12461-12466.

Cilloni, D. and Saglio, G. (2012) 'Molecular Pathways: BCR-ABL', *Clinical Cancer Research*, 18(4), p. 930.

Clark, O., Yen, K. and Mellinghoff, I.K. (2016) 'Molecular Pathways: Isocitrate Dehydrogenase Mutations in Cancer', *Clinical Cancer Research*, 22(8), p. 1837.

Cohen, M.H., Williams, G., Johnson, J.R., Duan, J., Gobburu, J., Rahman, A., Benson, K., Leighton, J., Kim, S.K., Wood, R., Rothmann, M., Chen, G., U, K.M., Staten, A.M. and Pazdur, R. (2002) 'Approval Summary for Imatinib Mesylate Capsules in the Treatment of Chronic Myelogenous Leukemia', *Clinical Cancer Research*, 8(5), p. 935. Conway O'Brien, E., Prideaux, S. and Chevassut, T. (2014) 'The Epigenetic Landscape of Acute Myeloid Leukemia', *Advances in Hematology*, 2014, p. 103175.

Creutzig, U. and Kaspers, G.J.L. (2004) 'Revised Recommendations of the International Working Group for Diagnosis, Standardization of Response Criteria, Treatment Outcomes, and Reporting Standards for Therapeutic Trials in Acute Myeloid Leukemia', *Journal of Clinical Oncology*, 22(16), pp. 3432-3433.

d'Ischia, M., Napolitano, A., Manini, P. and Panzella, L. (2011) 'Secondary Targets of Nitrite-Derived Reactive Nitrogen Species: Nitrosation/Nitration Pathways, Antioxidant Defense Mechanisms and Toxicological Implications', *Chemical Research in Toxicology*, 24(12), pp. 2071-2092.

Davies, C., Yip, B.H., Fernandez-Mercado, M., Woll, P.S., Agirre, X., Prosper, F., Jacobsen, S.E., Wainscoat, J.S., Pellagatti, A. and Boultwood, J. (2013) 'Silencing of ASXL1 impairs the granulomonocytic lineage potential of human CD34+ progenitor cells', *British Journal of Haematology*, 160(6), pp. 842-850.

Davies, K.J. and Doroshow, J.H. (1986) 'Redox cycling of anthracyclines by cardiac mitochondria. I. Anthracycline radical formation by NADH dehydrogenase', *Journal of Biological Chemistry*, 261(7), pp. 3060-7.

Dayyani, F., Wang, J., Yeh, J.-R.J., Ahn, E.-Y., Tobey, E., Zhang, D.-E., Bernstein, I.D., Peterson, R.T. and Sweetser, D.A. (2008) 'Loss of TLE1 & TLE4 from the del(9q) commonly deleted region in AML cooperates with AML1-ETO to affect myeloid cell proliferation and survival', *Blood*, 111(8), p. 4338.

Delhommeau, F., Dupont, S., James, C., Masse, A., le Couedic, J.P., Valle, V.D., Alberdi, A., Dessen, P., Fontenay, M., Casadevall, N., Soulier, J., Bernard, O. and Vainchenker, W. (2008) 'TET2 Is a Novel Tumor Suppressor Gene Inactivated in Myeloproliferative Neoplasms: Identification of a Pre-JAK2 V617F Event', *Blood*, 112(11), pp. lba-3.

Deschler, B. and Lübbert, M. (2006) 'Acute myeloid leukemia: Epidemiology and etiology', *Cancer*, 107(9), pp. 2099-2107.

Ding, L., Ley, T.J., Larson, D.E., Miller, C.A., Koboldt, D.C., Welch, J.S., Ritchey, J.K., Young, M.A., Lamprecht, T., McLellan, M.D., McMichael, J.F., Wallis, J.W., Lu, C., Shen, D., Harris, C.C., Dooling, D.J., Fulton, R.S., Fulton, L.L., Chen, K., Schmidt, H., Kalicki-Veizer, J., Magrini, V.J., Cook, L., McGrath, S.D., Vickery, T.L., Wendl, M.C., Heath, S., Watson, M.A., Link, D.C., Tomasson, M.H., Shannon, W.D., Payton, J.E., Kulkarni, S., Westervelt, P., Walter, M.J., Graubert, T.A., Mardis, E.R., Wilson, R.K. and DiPersio, J.F. (2012) 'Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing', *Nature*, 481(7382), pp. 506-510.

Döhner, H., Estey, E., Grimwade, D., Amadori, S., Appelbaum, F.R., Büchner, T., Dombret, H., Ebert, B.L., Fenaux, P., Larson, R.A., Levine, R.L., Lo-Coco, F., Naoe, T., Niederwieser, D., Ossenkoppele, G.J., Sanz, M., Sierra, J., Tallman, M.S., Tien, H.-F., Wei, A.H., Löwenberg, B. and Bloomfield, C.D. (2017) 'Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel', *Blood*, 129(4), pp. 424-447.

Döhner, H., Estey, E.H., Amadori, S., Appelbaum, F.R., Büchner, T., Burnett, A.K., Dombret, H., Fenaux, P., Grimwade, D., Larson, R.A., Lo-Coco, F., Naoe, T., Niederwieser, D., Ossenkoppele, G.J., Sanz, M.A., Sierra, J., Tallman, M.S., Löwenberg, B. and Bloomfield, C.D. (2009) 'Diagnosis and management of acute myeloid leukemia in adults: recommendations from an international expert panel, on behalf of the European LeukemiaNet', *Blood*, 115(3), pp. 453-474.

Döhner, H., Weisdorf, D.J. and Bloomfield, C.D. (2015) 'Acute Myeloid Leukemia', *New England Journal of Medicine*, 373(12), pp. 1136-1152.

Döhner, K. and Döhner, H. (2008) 'Molecular characterization of acute myeloid leukemia', *Haematologica*, 93(7), p. 976.

Downing, J.R. (2003) 'The core-binding factor leukemias: lessons learned from murine models', *Current Opinion in Genetics & Development*, 13(1), pp. 48-54.

Dufour, A., Schneider, F., Metzeler, K.H., Hoster, E., Schneider, S., Zellmeier, E., Benthaus, T., Sauerland, M.-C., Berdel, W.E., Büchner, T., Wörmann, B., Braess, J., Hiddemann, W., Bohlander, S.K. and Spiekermann, K. (2010) 'Acute Myeloid Leukemia With Biallelic CEBPA Gene Mutations and Normal Karyotype Represents a Distinct Genetic Entity Associated With a Favorable Clinical Outcome', *Journal of Clinical Oncology*, 28(4), pp. 570-577.

Dumanski, J.P., Rasi, C., Björklund, P., Davies, H., Ali, A.S., Grönberg, M., Welin, S., Sorbye, H., Grønbæk, H., Cunningham, J.L., Forsberg, L.A., Lind, L., Ingelsson, E., Stålberg, P., Hellman, P. and Tiensuu Janson, E. (2017) 'A MUTYH germline mutation is associated with small intestinal neuroendocrine tumors', *Endocrine-Related Cancer*, 24(8), pp. 427-443.

Eshleman, J.R. and Markowitz, S.D. (1996) 'Mismatch repair defects in human carcinogenesis', *Human Molecular Genetics*, 5, pp. 1489 - 1494.

Estey, E.H., Giles, F.J., Beran, M., Brien, S., Pierce, S.A., Faderl, S.H., Cortes, J.E. and Kantarjian, H.M. (2002) 'Experience with gemtuzumab ozogamycin (“mylotarg”) and all-trans retinoic acid in untreated acute promyelocytic leukemia', *Blood*, 99(11), p. 4222.

Ewald, B., Sampath, D. and Plunkett, W. (2008) 'Nucleoside analogs: molecular mechanisms signaling cell death', *Oncogene*, 27, p. 6522.

Faderl, S., Gandhi, V., Brien, S., Bonate, P., Cortes, J., Estey, E., Beran, M., Wierda, W., Garcia-Manero, G., Ferrajoli, A., Estrov, Z., Giles, F.J., Du, M., Kwari, M., Keating, M., Plunkett, W. and Kantarjian, H. (2005) 'Results of a phase 1-2 study of clofarabine in combination with cytarabine (ara-C) in relapsed and refractory acute leukemias', *Blood*, 105(3), p. 940.

Faderl, S., Ravandi, F., Huang, X., Garcia-Manero, G., Ferrajoli, A., Estrov, Z., Borthakur, G., Verstovsek, S., Thomas, D.A., Kwari, M. and Kantarjian, H.M. (2008) 'A randomized study of clofarabine versus clofarabine plus low-dose cytarabine as front-line therapy for patients aged 60 years and older with acute myeloid leukemia and high-risk myelodysplastic syndrome', *Blood*, 112(5), p. 1638.

Falini, B., Bolli, N., Liso, A., Martelli, M.P., Mannucci, R., Pileri, S. and Nicoletti, I. (2009) 'Altered nucleophosmin transport in acute myeloid leukaemia with mutated NPM1: molecular basis and clinical implications', *Leukemia*, 23(10), pp. 1731-1743.

Falini, B., Mecucci, C., Tiacci, E., Alcalay, M., Rosati, R., Pasqualucci, L., La Starza, R., Diverio, D., Colombo, E., Santucci, A., Bigerna, B., Pacini, R., Pucciarini, A., Liso, A., Vignetti, M., Fazi, P., Meani, N., Pettirossi, V., Saglio, G., Mandelli, F., Lo-Coco, F., Pelicci, P.-G. and Martelli, M.F. (2005) 'Cytoplasmic Nucleophosmin in Acute Myelogenous Leukemia with a Normal Karyotype', *New England Journal of Medicine*, 352(3), pp. 254-266.

Falini, B., Nicoletti, I., Bolli, N., Martelli, M.P., Liso, A., Gorello, P., Mandelli, F., Mecucci, C. and Martelli, M.F. (2007) 'Translocations and mutations involving the nucleophosmin (NPM1) gene in lymphomas and leukemias', *Haematologica*, 92(4), pp. 519-532.

Farrar, J.E., Schuback, H.L., Ries, R.E., Wai, D., Hampton, O.A., Trevino, L.R., Alonzo, T.A., Guidry Auvil, J.M., Davidsen, T.M., Gesuwan, P., Hermida, L., Muzny, D.M., Dewal, N., Rustagi, N., Lewis, L.R., Gamis, A.S., Wheeler, D.A., Smith, M.A., Gerhard, D.S. and Meshinchi, S. (2016) 'Genomic Profiling of Pediatric Acute Myeloid Leukemia Reveals a Changing Mutational Landscape from Disease Diagnosis to Relapse', *Cancer Research*, 76(8), p. 2197.

Fasan, A., Haferlach, C., Alpermann, T., Jeromin, S., Grossmann, V., Eder, C., Weissmann, S., Dicker, F., Kohlmann, A., Schindela, S., Kern, W., Haferlach, T. and Schnittger, S. (2013) 'The role of different genetic subtypes of CEBPA mutated AML', *Leukemia*, 28, p. 794.

Felix Carolyn, A. (2001) 'Leukemias related to treatment with DNA topoisomerase II inhibitors*', *Medical and Pediatric Oncology*, 36(5), pp. 525-535.

Fenaux, P., Preudhomme, C., Quiquandon, I., Jonveaux, P., Laï, J.L., Vanrumbeke, M., Loucheux-Lefebvre, M.H., Bauters, F., Berger, R. and Kerckaert, J.P. (1992) 'Mutations of the P53 gene in acute myeloid leukaemia', *British Journal of Haematology*, 80(2), pp. 178-183.

Ferrara, F. and Del Vecchio, L. (2002) 'Acute myeloid leukemia with t(8;21)/AML1/ETO: a distinct biological and clinical entity', *Haematologica*, 87(3), p. 306.

Figueroa, M.E., Abdel-Wahab, O., Lu, C., Ward, P.S., Patel, J., Shih, A., Li, Y., Bhagwat, N., Vasanthakumar, A., Fernandez, H.F., Tallman, M.S., Sun, Z., Wolniak, K., Peeters, J.K., Liu, W., Choe, S.E., Fantin, V.R., Paietta, E., Löwenberg, B., Licht, J.D., Godley, L.A., Delwel, R., Valk, P.J.M., Thompson, C.B., Levine, R.L. and Melnick, A. (2010a) 'Leukemic IDH1 and IDH2 Mutations Result in a Hypermethylation Phenotype, Disrupt TET2 Function, and Impair Hematopoietic Differentiation', *Cancer Cell*, 18(6), pp. 553-567.

Figueroa, M.E., Lugthart, S., Li, Y., Erpelinck-Verschueren, C., Deng, X., Christos, P.J., Schifano, E., Booth, J., van Putten, W., Skrabanek, L., Campagne, F., Mazumdar, M., Greally, J.M., Valk, P.J.M., Löwenberg, B., Delwel, R. and Melnick, A. (2010b) 'DNA Methylation Signatures Identify Biologically Distinct Subtypes in Acute Myeloid Leukemia', *Cancer Cell*, 17(1), pp. 13-27.

Fordham, S.E., Cole, M., Irving, J.A. and Allan, J.M. (2015) 'Cytarabine preferentially induces mutation at specific sequences in the genome which are identifiable in relapsed acute myeloid leukaemia', *Leukemia*.

Forster, V.J., Nahari, M.H., Martinez-Soria, N., Bradburn, A.K., Ptasinska, A., Assi, S.A., Fordham, S.E., McNeil, H., Bonifer, C., Heidenreich, O. and Allan, J.M. (2015) 'The leukemia-associated RUNX1/ETO oncoprotein confers a mutator phenotype', *Leukemia*.

Fortune, J.M. and Osheroff, N. (2000) 'Topoisomerase II as a target for anticancer drugs: When enzymes stop being nice', in *Progress in Nucleic Acid Research and Molecular Biology*. Academic Press, pp. 221-253.

Frazer, R., Irvine, A.E. and McMullin, M.F. (2007) 'Chronic Myeloid Leukaemia in The 21st Century', *The Ulster Medical Journal*, 76(1), pp. 8-17.

Fry, R.C., Begley, T.J. and Samson, L.D. (2005) 'GENOME-WIDE RESPONSES TO DNA-DAMAGING AGENTS', *Annual Review of Microbiology*, 59(1), pp. 357-377.

Fukuyama, T., Sueoka, E., Sugio, Y., Otsuka, T., Niho, Y., Akagi, K. and Kozu, T. (2001) 'MTG8 proto-oncoprotein interacts with the regulatory subunit of type II cyclic AMP-dependent protein kinase in lymphocytes', *Oncogene*, 20, p. 6225.

Furgeri, D.T., Marson, F.A.L., Correia, C.A.A., Ribeiro, J.D. and Bertuzzo, C.S. (2018) 'Cystic fibrosis transmembrane regulator haplotypes in households of patients with cystic fibrosis', *Gene*, 641, pp. 137-143.

Furth, E.E., Thilly, W.G., Penman, B.W., Liber, H.L. and Rand, W.M. (1981) 'Quantitative assay for mutation in diploid human lymphoblasts using microtiter plates', *Analytical Biochemistry*, 110(1), pp. 1-8.

Gaidzik, V.I., Paschka, P., Späth, D., Habdank, M., Köhne, C.-H., Germing, U., von Lilienfeld-Toal, M., Held, G., Horst, H.-A., Haase, D., Bentz, M., Götze, K., Döhner, H., Schlenk, R.F., Bullinger, L. and Döhner, K. (2012) 'TET2 Mutations in Acute Myeloid Leukemia (AML): Results From a Comprehensive Genetic and Clinical Analysis of the AML Study Group', *Journal of Clinical Oncology*, 30(12), pp. 1350-1357.

Gao, Q., Liang, W.-W., Foltz, S.M., Mutharasu, G., Jayasinghe, R.G., Cao, S., Liao, W.-W., Reynolds, S.M., Wyczalkowski, M.A., Yao, L., Yu, L., Sun, S.Q., Caesar-Johnson, S.J., Demchok, J.A., Felau, I., Kasapi, M., Ferguson, M.L., Hutter, C.M., Sofia, H.J., Tarnuzzer, R., Wang, Z., Yang, L., Zenklusen, J.C., Zhang, J., Chudamani, S., Liu, J., Lolla, L., Naresh, R., Pihl, T., Sun, Q., Wan, Y., Wu, Y., Cho, J., DeFreitas, T., Frazer, S., Gehlenborg, N., Getz, G., Heiman, D.I., Kim, J., Lawrence, M.S., Lin, P., Meier, S., Noble, M.S., Saksena, G., Voet, D., Zhang, H., Bernard, B., Chambwe, N., Dhankani, V., Knijnenburg, T., Kramer, R., Leinonen, K., Liu, Y., Miller, M., Reynolds, S., Shmulevich, I., Thorsson, V., Zhang, W., Akbani, R., Broom, B.M., Hegde, A.M., Ju, Z., Kanchi, R.S., Korkut, A., Li, J., Liang, H., Ling, S., Liu, W., Lu, Y., Mills, G.B., Ng, K.-S., Rao, A., Ryan, M., Wang, J., Weinstein, J.N., Zhang, J., Abeshouse, A., Armenia, J., Chakravarty, D., Chatila, W.K., de Bruijn, I., Gao, J., Gross, B.E., Heins, Z.J., Kundra, R., La, K., Ladanyi, M., Luna, A., Nissan, M.G., Ochoa, A., Phillips, S.M., Reznik, E., Sanchez-Vega, F., Sander, C., Schultz, N., Sheridan, R., Sumer, S.O., Sun, Y., Taylor, B.S., Wang, J., et al. (2018) 'Driver Fusions and Their Implications in the Development and Treatment of Human Cancers', *Cell Reports*, 23(1), pp. 227-238.e3.

Gao, Y.G., van der Marel, G.A., van Boom, J.H. and Wang, A.H. (1991) 'Molecular structure of a DNA decamber containing an anticancer nucleoside arabinosylcytosine: conformational perturbation by arabinosylcytosine in B-DNA', *Biochemistry*, 41(30), pp. 9922-9931.

Garcia-Diaz, M., Murray, M.S., Kunkel, T.A. and Chou, K.-m. (2010) 'Interaction between DNA Polymerase λ and Anticancer Nucleoside Analogs', *Journal of Biological Chemistry*, 285(22), pp. 16874-16879.

Garg, M., Nagata, Y., Kanojia, D., Mayakonda, A., Yoshida, K., Haridas Keloth, S., Zang, Z.J., Okuno, Y., Shiraishi, Y., Chiba, K., Tanaka, H., Miyano, S., Ding, L.-W., Alpermann, T., Sun, Q.-Y., Lin, D.-C., Chien, W., Madan, V., Liu, L.-Z., Tan, K.-T., Sampath, A., Venkatesan, S., Inokuchi, K., Wakita, S., Yamaguchi, H., Chng, W.J., Kham, S.-K.Y., Yeoh, A.E.-J., Sanada, M., Schiller, J., Kreuzer, K.-A., Kornblau, S.M., Kantarjian, H.M., Haferlach, T., Lill, M., Kuo, M.-C., Shih, L.-Y., Blau, I.-W., Blau, O., Yang, H., Ogawa, S. and Koeffler, H.P. (2015) 'Profiling of somatic mutations in acute

myeloid leukemia with FLT3-ITD at diagnosis and relapse', *Blood*, 126(22), pp. 2491-2501.

Gewirtz, D. (1999) 'A critical evaluation of the mechanisms of action proposed for the antitumor effects of the anthracycline antibiotics adriamycin and daunorubicin', *Biochemical Pharmacology*, 57(7), pp. 727-741.

Gilliland, D.G. and Griffin, J.D. (2002) 'The roles of FLT3 in hematopoiesis and leukemia', *Blood*, 100(5), pp. 1532-1542.

Giver Cynthia, R., Nelson Stephen, L. and Grosovsky Andrew, J. (1993) 'Spectrum of spontaneous HPRT mutations in TK6 human lymphoblasts', *Environmental and Molecular Mutagenesis*, 22(3), pp. 138-146.

Godley, L.A. and Larson, R.A. (2008) 'Therapy-related Myeloid Leukemia', *Seminars in oncology*, 35(4), pp. 418-429.

Goemans, B.F., Zwaan, C.M., Miller, M., Zimmermann, M., Harlow, A., Meshinchi, S., Loonen, A.H., Hahlen, K., Reinhardt, D., Creutzig, U., Kaspers, G.J.L. and Heinrich, M.C. (2005) 'Mutations in KIT and RAS are frequent events in pediatric core-binding factor acute myeloid leukemia', *Leukemia*, 19(9), pp. 1536-1542.

Goode, E.L., Ulrich, C.M. and Potter, J.D. (2002) 'Polymorphisms in DNA Repair Genes and Associations with Cancer Risk', *Cancer Epidemiology Biomarkers & Prevention*, 11(12), p. 1513.

Graham, F.L. and Whitmore, G.F. (1970) 'Studies in Mouse L-cells on the Incorporation of 1-β-D-Arabinofuranosylcytosine into DNA and on Inhibition of DNA Polymerase by 1-β-D-Arabinofuranosylcytosine 5′-Triphosphate', *Cancer Research*, 30(11), p. 2636.

Green, C.L., Koo, K.K., Hills, R.K., Burnett, A.K., Linch, D.C. and Gale, R.E. (2010) 'Prognostic Significance of CEBPA Mutations in a Large Cohort of Younger Adult Patients With Acute Myeloid Leukemia: Impact of Double CEBPA Mutations and the Interaction With FLT3 and NPM1 Mutations', *Journal of Clinical Oncology*, 28(16), pp. 2739-2747.

Greif, P.A., Hartmann, L., Vosberg, S., Stief, S.M., Mattes, R., Hellmann, I., Metzeler, K.H., Herold, T., Bamopoulos, S.A., Kerbs, P., Jurinovic, V., Schumacher, D., Pastore, F., Bräundl, K., Zellmeier, E., Ksienzyk, B., Konstandin, N.P., Schneider, S., Graf, A., Krebs, S., Blum, H., Neumann, M., Baldus, C.D., Bohlander, S.K., Wolf, S., Görlich, D., Berdel, W.E., Wörmann, B.J., Hiddemann, W. and Spiekermann, K. (2018) 'Evolution of Cytogenetically Normal Acute Myeloid Leukemia During Therapy and Relapse: An Exome Sequencing Study of 50 Patients', *Clinical Cancer Research*, 24(7), p. 1716.

Grimwade, D., Hills, R.K., Moorman, A.V., Walker, H., Chatters, S., Goldstone, A.H., Wheatley, K., Harrison, C.J. and Burnett, A.K. (2010) 'Refinement of cytogenetic classification in acute myeloid leukemia: determination of prognostic significance of rare recurring chromosomal abnormalities among 5876 younger adult patients treated in the United Kingdom Medical Research Council trials', *Blood*, 116(3), p. 354.

Grisendi, S., Bernardi, R., Rossi, M., Cheng, K., Khandker, L., Manova, K. and Pandolfi, P.P. (2005) 'Role of nucleophosmin in embryonic development and tumorigenesis', *Nature*, 437, p. 147.

Gu, T.-L., Goetz, T.L., Graves, B.J. and Speck, N.A. (2000) 'Auto-Inhibition and Partner Proteins, Core-Binding Factor β (CBFβ) and Ets-1, Modulate DNA Binding by CBFα2 (AML1)', *Molecular and Cellular Biology*, 20(1), pp. 91-103.

Gunther, C.V. and Graves, B.J. (1994) 'Identification of ETS domain proteins in murine T lymphocytes that interact with the Moloney murine leukemia virus enhancer', *Molecular and Cellular Biology*, 14(11), pp. 7569-7580.

Haferlach, C., Dicker, F., Herholz, H., Schnittger, S., Kern, W. and Haferlach, T. (2008) 'Mutations of the TP53 gene in acute myeloid leukemia are strongly associated with a complex aberrant karyotype', *Leukemia*, 22, p. 1539.

Haferlach, C., Mecucci, C., Schnittger, S., Kohlmann, A., Mancini, M., Cuneo, A., Testoni, N., Rege-Cambrin, G., Santucci, A., Vignetti, M., Fazi, P., Martelli, M.P., Haferlach, T. and Falini, B. (2009) 'AML with mutated NPM1 carrying a normal or aberrant karyotype show overlapping biologic, pathologic, immunophenotypic, and prognostic features', *Blood*, 114(14), p. 3024.

Hande, K.R. (2008) 'Topoisomerase II inhibitors', *Update on Cancer Therapeutics*, 3(1), pp. 13-26.

Hantschel, O. (2012) 'Structure, Regulation, Signaling, and Targeting of Abl Kinases in Cancer', *Genes & Cancer*, 3(5-6), pp. 436-446.

Harris, S.L. and Levine, A.J. (2005) 'The p53 pathway: positive and negative feedback loops', *Oncogene*, 24, p. 2899.

Harrison, C.J. (2001) 'Philadelphia Chromosome', in Brenner, S. and Miller, J.H. (eds.) *Encyclopedia of Genetics*. New York: Academic Press, pp. 1449-1450.

Hartmann, L., Dutta, S., Opatz, S., Vosberg, S., Reiter, K., Leubolt, G., Metzeler, K.H., Herold, T., Bamopoulos, S.A., Bräundl, K., Zellmeier, E., Ksienzyk, B., Konstandin, N.P., Schneider, S., Hopfner, K.-P., Graf, A., Krebs, S., Blum, H., Middeke, J.M., Stölzel, F., Thiede, C., Wolf, S., Bohlander, S.K., Preiss, C., Chen-Wichmann, L., Wichmann, C., Sauerland, M.C., Büchner, T., Berdel, W.E., Wörmann, B.J., Braess, J., Hiddemann, W., Spiekermann, K. and Greif, P.A. (2016) 'ZBTB7A mutations in acute myeloid leukaemia with t(8;21) translocation', *Nature Communications*, 7, p. 11733.

Haupt, Y., Maya, R., Kazaz, A. and Oren, M. (1997) 'Mdm2 promotes the rapid degradation of p53', *Nature*, 387, p. 296.

Helleday, T., Eshtad, S. and Nik-Zainal, S. (2014) 'Mechanisms underlying mutational signatures in human cancers', *Nat Rev Genet*, 15(9), pp. 585-598.

Herman, J.G., Jen, J., Merlo, A. and Baylin, S.B. (1996) 'Hypermethylation-associated Inactivation Indicates a Tumor Suppressor Role for INK4B', *Cancer Research*, 56(4), p. 722.

Hiebert, S.W., Reed-Inderbitzin, E.F., Amann, J., Irvin, B., Durst, K. and Linggi, B. (2003) 'The t(8;21) fusion protein contacts co-repressors and histone deacetylases to repress the transcription of the p14ARF tumor suppressor', *Blood Cells, Molecules, and Diseases*, 30(2), pp. 177-183.

Higuchi, M., O'Brien, D., Kumaravelu, P., Lenny, N., Yeoh, E.-J. and Downing, J.R. (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(1), pp. 63-74.

Holliday, R. and Grigg, G.W. (1993) 'DNA methylation and mutation', *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 285(1), pp. 61-67.

Hollstein, M., Sidransky, D., Vogelstein, B. and Harris, C.C. (1991) 'p53 mutations in human cancers', *Science*, 253(5015), p. 49.

Housman, G., Byler, S., Heerboth, S., Lapinska, K., Longacre, M., Snyder, N. and Sarkar, S. (2014) 'Drug Resistance in Cancer: An Overview', *Cancers*, 6(3), pp. 1769-1792.

Huang, H., Yu, M., Akie, T.E., Moran, T.B., Woo, A.J., Tu, N., Waldon, Z., Lin, Y.Y., Steen, H. and Cantor, A.B. (2009) 'Differentiation-Dependent Interactions between RUNX-1 and FLI-1 during Megakaryocyte Development', *Molecular and Cellular Biology*, 29(15), pp. 4103-4115.

Huang, P., Chubb, S., Hertel, L.W., Grindey, G.B. and Plunkett, W. (1991) 'Action of 2′,2′-Difluorodeoxycytidine on DNA Synthesis', *Cancer Research*, 22(51), pp. 6110-6117.

Hunter, C., Smith, R., Cahill, D.P., Stephens, P., Stevens, C., Teague, J., Greenman, C., Edkins, S., Bignell, G., Davies, H., Meara, S., Parker, A., Avis, T., Barthorpe, S., Brackenbury, L., Buck, G., Butler, A., Clements, J., Cole, J., Dicks, E., Forbes, S., Gorton, M., Gray, K., Halliday, K., Harrison, R., Hills, K., Hinton, J., Jenkinson, A., Jones, D., Kosmidou, V., Laman, R., Lugg, R., Menzies, A., Perry, J., Petty, R., Raine, K., Richardson, D., Shepherd, R., Small, A., Solomon, H., Tofts, C., Varian, J., West, S., Widaa, S., Yates, A., Easton, D.F., Riggins, G., Roy, J.E., Levine, K.K., Mueller, W., Batchelor, T.T., Louis, D.N., Stratton, M.R., Futreal, P.A. and Wooster, R. (2006) 'A Hypermutation Phenotype and Somatic Mutations in Recurrent Human Malignant Gliomas after Alkylator Chemotherapy', *Cancer Research*, 66(8), p. 3987.

Hyun, J.-W., Choi, J.-Y., Zeng, H.-H., Lee, Y.-S., Kim, H.-S., Yoon, S.-H. and Chung, M.-H. (2000) 'Leukemic cell line, KG-1 has a functional loss of hOGG1 enzyme due to a point mutation and 8-hydroxydeoxyguanosine can kill KG-1', *Oncogene*, 19, p. 4476.
Imai, Y., Kurokawa, M., Yamaguchi, Y., Izutsu, K., Nitta, E., Mitani, K., Satake, M., Noda, T., Ito, Y. and Hirai, H. (2004) 'The Corepressor mSin3A Regulates Phosphorylation-Induced Activation, Intranuclear Location, and Stability of AML1', *Molecular and Cellular Biology*, 24(3), pp. 1033-1043.

Ishchenko, A., Sinitsyna, O., Krysanova, O., Vasyunina, E.A., Saparbaev, M., Sidorkina, O. and Nevinsky, G.A. (2003) 'Age-dependent increase of 8-oxoguanine-, hypoxanthine-, and uracil- DNA glycosylase activities in liver extracts from OXYS rats

with inherited overgeneration of free radicals and Wistar rats', *Medical Science Monitor*, 1(9), pp. BR16-BR24.

Jacobs, A.L. and Schär, P. (2012) 'DNA glycosylases: in DNA repair and beyond', *Chromosoma*, 121(1), pp. 1-20.

Jankowska, A.M., Gondek, L.P., Szpurka, H., Nearman, Z.P., Tiu, R.V. and Maciejewski, J.P. (2007) 'Base excision repair dysfunction in a subgroup of patients with myelodysplastic syndrome', *Leukemia*, 22, p. 551.

Janssen, J.W., Steenvoorden, A.C., Lyons, J., Anger, B., Böhlke, J.U., Bos, J.L., Seliger, H. and Bartram, C.R. (1987) 'RAS gene mutations in acute and chronic myelocytic leukemias, chronic myeloproliferative disorders, and myelodysplastic syndromes', *Proceedings of the National Academy of Sciences of the United States of America*, 84(24), pp. 9228-9232.

Jiao, B., Wu, C.F., Liang, Y., Chen, H.M., Xiong, S.M., Chen, B., Shi, J.Y., Wang, Y.Y., Wang, J.H., Chen, Y., Li, J.M., Gu, L.J., Tang, J.Y., Shen, Z.X., Gu, B.W., Zhao, W.L., Chen, Z. and Chen, S.J. (2009) 'AML1-ETO9a is correlated with C-KIT overexpression/mutations and indicates poor disease outcome in t(8;21) acute myeloid leukemia-M2', *Leukemia*, 23, p. 1598.

Jing, X.L., Zhao, W.B. and Wang, Y.Y. (2003) 'ARF-Mdm2-p53', *Funct. Mater. (China)*, 34, p. 328.

Jiricny, J. (2006) 'The multifaceted mismatch-repair system', *Nature Reviews Molecular Cell Biology*, 7, p. 335.

Jones, J.M. and Gellert, M. (2004) 'The taming of a transposon: V(D)J recombination and the immune system', *Immunological Reviews*, 200(1), pp. 233-248.

Juchau, M.R., Fantel, A.G., Harris, C. and Beyer, B.K. (1986) 'The potential role of redox cycling as a mechanism for chemical teratogenesis', *Environmental Health Perspectives*, 70, pp. 131-136.

Kakizuka, A., Miller, W.H., Umesono, K., Warrell, R.P., Frankel, S.R., Murty, V.V.V.S., Dmitrovsky, E. and Evans, R.M. (1991) 'Chromosomal translocation t(15;17) in human acute promyelocytic leukemia fuses RARα with a novel putative transcription factor, PML', *Cell*, 66(4), pp. 663-674.

Kantarjian, H. (2015) 'Acute myeloid leukemia—Major progress over four decades and glimpses into the future', *American Journal of Hematology*, 91(1), pp. 131-145.

Kanwal, R. and Gupta, S. (2012) 'Epigenetic modifications in cancer', *Clinical genetics*, 81(4), pp. 303-311.

Kelly, L.M. and Gilliland, D.G. (2002) 'Genetics of Myeloid Leukaemias', *Annual Review of Genomics and Human Genetics*, 3(1), pp. 179-198.

Kelly, P.N. and Strasser, A. (2011) 'The role of Bcl-2 and its pro-survival relatives in tumourigenesis and cancer therapy', *Cell Death and Differentiation*, 18(9), pp. 1414-1424.

Kershaw, R.M. and Hodges, N.J. (2012) 'Repair of oxidative DNA damage is delayed in the Ser326Cys polymorphic variant of the base excision repair protein OGG1', *Mutagenesis*, 27(4), pp. 501-510.

Kirsch, D.G. and Kastan, M.B. (1998) 'Tumor-suppressor p53: implications for tumor development and prognosis', *Journal of Clinical Oncology*, 16(9), pp. 3158-3168.

Kirstetter, P., Schuster, M.B., Bereshchenko, O., Moore, S., Dvinge, H., Kurz, E., Theilgaard-Mönch, K., Månsson, R., Pedersen, T.Å., Pabst, T., Schrock, E., Porse, B.T., Jacobsen, S.E.W., Bertone, P., Tenen, D.G. and Nerlov, C. (2008) 'Modeling of C/EBPα Mutant Acute Myeloid Leukemia Reveals a Common Expression Signature of Committed Myeloid Leukemia-Initiating Cells', *Cancer Cell*, 13(4), pp. 299-310.

Kitabayashi, I., Ida, K., Morohoshi, F., Yokoyama, A., Mitsuhashi, N., Shimizu, K., Nomura, N., Hayashi, Y. and Ohki, M. (1998) 'The AML1-MTG8 Leukemic Fusion Protein Forms a Complex with a Novel Member of the MTG8(ETO/CDR) Family, MTGR1', *Molecular and Cellular Biology*, 18(2), pp. 846-858.

Kitamura, Y., Ota, T., Matsuoka, Y., Tooyama, I., Kimura, H., Shimohama, S., Nomura, Y., Gebicke-Haerter, P.J. and Taniguchi, T. (1999) 'Hydrogen peroxide-induced apoptosis mediated by p53 protein in glial cells', *Glia.*, 2(25), pp. 156-164.

Klampfer, L., Zhang, J., Zelenetz, A.O., Uchida, H. and Nimer, S.D. (1996) 'The AML1/ETO fusion protein activates transcription of BCL-2', *Proceedings of the National Academy of Sciences of the United States of America*, 93(24), pp. 14059-14064.

Klaus, M., Haferlach, T., Schnittger, S., Kern, W., Hiddemann, W. and Schoch, C. (2004) 'Cytogenetic profile in de novo acute myeloid leukemia with FAB subtypes M0, M1, and M2: a study based on 652 cases analyzed with morphology, cytogenetics, and fluorescence in situ hybridization', *Cancer Genetics and Cytogenetics*, 155(1), pp. 47-56.

Klein, A.d., Kessel, A.G.v., Grosveld, G., Bartram, C.R., Hagemeijer, A., Bootsma, D., Spurr, N.K., Heisterkamp, N., Groffen, J. and Stephenson, J.R. (1982) 'A cellular oncogene is translocated to the Philadelphia chromosome in chronic myelocytic leukaemia', *Nature*, 300, p. 765.

Ko, M., Huang, Y., Jankowska, A.M., Pape, U.J., Tahiliani, M., Bandukwala, H.S., An, J., Lamperti, E.D., Koh, K.P., Ganetzky, R., Liu, X.S., Aravind, L., Agarwal, S., Maciejewski, J.P. and Rao, A. (2010) 'Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2', *Nature*, 468(7325), pp. 839-843.

Kohl, T.M., Schnittger, S., Ellwart, J.W., Hiddemann, W. and Spiekermann, K. (2005) 'KIT exon 8 mutations associated with core-binding factor (CBF)–acute myeloid leukemia (AML) cause hyperactivation of the receptor in response to stem cell factor', *Blood*, 105(8), p. 3319.

Kow, Y.W. (2002) 'Repair of deaminated bases in DNA', *Free Radical Biology and Medicine*, 33(7), pp. 886-893.

Krauth, M.T., Eder, C., Alpermann, T., Bacher, U., Nadarajah, N., Kern, W., Haferlach, C., Haferlach, T. and Schnittger, S. (2014) 'High number of additional genetic lesions

in acute myeloid leukemia with t(8;21)/RUNX1-RUNX1T1: frequency and impact on clinical outcome', *Leukemia*, 28, p. 1449.

Krejci, O., Wunderlich, M., Geiger, H., Chou, F.-S., Schleimer, D., Jansen, M., Andreassen, P.R. and Mulloy, J.C. (2008) 'p53 signaling in response to increased DNA damage sensitizes AML1-ETO cells to stress-induced death', *Blood*, 111(4), p. 2190.
Krönke, J., Bullinger, L., Teleanu, V., Tschürtz, F., Gaidzik, V.I., Kühn, M.W.M., Rücker, F.G., Holzmann, K., Paschka, P., Kapp-Schwörer, S., Späth, D., Kindler, T., Schittenhelm, M., Krauter, J., Ganser, A., Göhring, G., Schlegelberger, B., Schlenk, R.F., Döhner, H. and Döhner, K. (2013) 'Clonal evolution in relapsed NPM1-mutated acute myeloid leukemia', *Blood*, 122(1), pp. 100-108.

Kubbutat, M.H.G., Jones, S.N. and Vousden, K.H. (1997) 'Regulation of p53 stability by Mdm2', *Nature*, 387, p. 299.

Kumar, D., Mehta, A., Panigrahi, M.K., Nath, S. and Saikia, K.K. (2018) 'NPM1 Mutation Analysis in Acute Myeloid Leukemia: Comparison of Three Techniques - Sanger Sequencing, Pyrosequencing, and Real-Time Polymerase Chain Reaction', *Turkish Journal of Hematology*, 35(1), pp. 49-53.

Kunisada, M., Sakumi, K., Tominaga, Y., Budiyanto, A., Ueda, M., Ichihashi, M., Nakabeppu, Y. and Nishigori, C. (2005) '8-Oxoguanine Formation Induced by Chronic UVB Exposure Makes Knockout Mice Susceptible to Skin Carcinogenesis', *Cancer Research*, 65(14), p. 6006.

Lam, K. and Zhang, D.-E. (2012) 'RUNX1 and RUNX1-ETO: roles in hematopoiesis and leukemogenesis', *Frontiers in bioscience : a journal and virtual library*, 17, pp. 1120-1139.

Larson, E.D. and Maizels, N. (2004) 'Transcription-coupled mutagenesis by the DNA deaminase AID', *Genome biology*, 5(3), pp. 211-211.

Larson, R.A. (2007) 'Etiology and Management of Therapy-Related Myeloid Leukemia', *ASH Education Program Book*, 2007(1), pp. 453-459.

Lehmann, M., Vilar, K.d.S.P., Franco, A., Reguly, M.L. and de Andrade, H.H.R. (2004) 'Activity of topoisomerase inhibitors daunorubicin, idarubicin, and aclarubicin in the Drosophila Somatic Mutation and Recombination Test', *Environmental and Molecular Mutagenesis*, 43(4), pp. 250-257.

Leibeling, D., Laspe, P. and Emmert, S. (2006) 'Nucleotide excision repair and cancer', *Journal of Molecular Histology*, 37(5), pp. 225-238.

Leung, A.Y.H., Man, C.H. and Kwong, Y.L. (2013) 'FLT3 inhibition: a moving and evolving target in acute myeloid leukaemia', *Leukemia*, 27(2), pp. 260-268.

Levanon, D., Negreanu, V., Bernstein, Y., Bar-Am, I., Avivi, L. and Groner, Y. (1994) 'AML1, AML2, and AML3, the Human Members of the runt domain Gene-Family: cDNA Structure, Expression, and Chromosomal Localization', *Genomics*, 23(2), pp. 425-432.
Ley, T.J., Mardis, E.R., Ding, L., Fulton, B., McLellan, M.D., Chen, K., Dooling, D., Dunford-Shore, B.H., McGrath, S., Hickenbotham, M., Cook, L., Abbott, R., Larson, D.E., Koboldt, D.C., Pohl, C., Smith, S., Hawkins, A., Abbott, S., Locke, D., Hillier, L.W.,

Miner, T., Fulton, L., Magrini, V., Wylie, T., Glasscock, J., Conyers, J., Sander, N., Shi, X., Osborne, J.R., Minx, P., Gordon, D., Chinwalla, A., Zhao, Y., Ries, R.E., Payton, J.E., Westervelt, P., Tomasson, M.H., Watson, M., Baty, J., Ivanovich, J., Heath, S., Shannon, W.D., Nagarajan, R., Walter, M.J., Link, D.C., Graubert, T.A., DiPersio, J.F. and Wilson, R.K. (2008) 'DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome', *Nature*, 456(7218), pp. 66-72.

Liber, H.L. and Thilly, W.G. (1982) 'Mutation assay at the thymidine kinase locus in diploid human lymphoblasts', *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 94(2), pp. 467-485.

Liddiard, K., Hills, R., Burnett, A.K., Darley, R.L. and Tonks, A. (2010) 'OGG1 is a novel prognostic indicator in acute myeloid leukaemia', *Oncogene*, 29, p. 2005.

Lieber, M.R. (2010) 'The Mechanism of Double-Strand DNA Break Repair by the Nonhomologous DNA End Joining Pathway', *Annual Review of Biochemistry*, (79), pp. 181-211.

Lin, A.W. and Lowe, S.W. (2001) 'Oncogenic ras activates the ARF-p53 pathway to suppress epithelial cell transformation', *Proceedings of the National Academy of Sciences of the United States of America*, 98(9), pp. 5025-5030.

Lin, S., Mulloy, J.C. and Goyama, S. (2017) 'RUNX1-ETO Leukemia', in Groner, Y., Ito, Y., Liu, P., Neil, J.C., Speck, N.A. and van Wijnen, A. (eds.) *RUNX Proteins in Development and Cancer*. Singapore: Springer Singapore, pp. 151-173.

Linggi, B., Müller-Tidow, C., van de Locht, L., Hu, M., Nip, J., Serve, H., Berdel, W.E., van der Reijden, B., Quelle, D.E., Rowley, J.D., Cleveland, J., Jansen, J.H., Pandolfi, P.P. and Hiebert, S.W. (2002) 'The t(8;21) fusion protein, AML1–ETO, specifically represses the transcription of the p14ARF tumor suppressor in acute myeloid leukemia', *Nature Medicine*, 8, p. 743.

Liou, G.-Y. and Storz, P. (2010) 'Reactive oxygen species in cancer', *Free radical research*, 44(5), p. 10.3109/10715761003667554.

Liu, P., Tarle, S.A., Hajra, A., Claxton, D.F., Marlton, P., Freedman, M., Siciliano, M.J. and Collins, F.S. (1993) 'Fusion between transcription factor CBF beta/PEBP2 beta and a myosin heavy chain in acute myeloid leukemia', *Science*, 261(5124), pp. 1041-1044.

Liu, Y., Cheney, M.D., Gaudet, J.J., Chruszcz, M., Lukasik, S.M., Sugiyama, D., Lary, J., Cole, J., Dauter, Z., Minor, W., Speck, N.A. and Bushweller, J.H. (2006) 'The tetramer structure of the Nervy homology two domain, NHR2, is critical for AML1/ETO's activity', *Cancer Cell*, 9(4), pp. 249-260.

Lo-Coco, F., Cimino, G., Breccia, M., Noguera, N.I., Diverio, D., Finolezzi, E., Pogliani, E.M., Di Bona, E., Micalizzi, C., Kropp, M., Venditti, A., Tafuri, A. and Mandelli, F. (2004) 'Gemtuzumab ozogamicin (Mylotarg) as a single agent for molecularly relapsed acute promyelocytic leukemia', *Blood*, 104(7), p. 1995.

Loeb, L.A. (1991) 'Mutator Phenotype May Be Required for Multistage Carcinogenesis', *Cancer Research*, 51(12), p. 3075.

Lovett, B.D., Lo Nigro, L., Rappaport, E.F., Blair, I.A., Osheroff, N., Zheng, N., Megonigal, M.D., Williams, W.R., Nowell, P.C. and Felix, C.A. (2001) 'Near-precise interchromosomal recombination and functional DNA topoisomerase II cleavage sites at MLL and AF-4 genomic breakpoints in treatment-related acute lymphoblastic leukemia with t(4;11) translocation', *Proceedings of the National Academy of Sciences of the United States of America*, 98(17), pp. 9802-9807.

Lundin, C., Erixon, K., Arnaudeau, C., Schultz, N., Jenssen, D., Meuth, M. and Helleday, T. (2002) 'Different Roles for Nonhomologous End Joining and Homologous Recombination following Replication Arrest in Mammalian Cells', *Molecular and Cellular Biology*, 22(16), pp. 5869-5878.

Lutterbach, B., Westendorf, J.J., Linggi, B., Patten, A., Moniwa, M., Davie, J.R., Huynh, K.D., Bardwell, V.J., Lavinsky, R.M., Rosenfeld, M.G., Glass, C., Seto, E. and Hiebert, S.W. (1998) 'ETO, a Target of t(8;21) in Acute Leukemia, Interacts with the N-CoR and mSin3 Corepressors', *Molecular and Cellular Biology*, 18(12), pp. 7176-7184.

Maiti, A.K., Boldogh, I., Spratt, H., Mitra, S. and Hazra, T.K. (2008) 'Mutator Phenotype of Mammalian Cells Due to Deficiency of NEIL1 DNA Glycosylase, An Oxidized Base-Specific Repair Enzyme', *DNA repair*, 7(8), pp. 1213-1220.

Marcucci, G., Maharry, K., Radmacher, M.D., Mrózek, K., Vukosavljevic, T., Paschka, P., Whitman, S.P., Langer, C., Baldus, C.D., Liu, C.-G., Ruppert, A.S., Powell, B.L., Carroll, A.J., Caligiuri, M.A., Kolitz, J.E., Larson, R.A. and Bloomfield, C.D. (2008) 'Prognostic Significance of, and Gene and MicroRNA Expression Signatures Associated With, CEBPA Mutations in Cytogenetically Normal Acute Myeloid Leukemia With High-Risk Molecular Features: A Cancer and Leukemia Group B Study', *Journal of Clinical Oncology*, 26(31), pp. 5078-5087.

Mariano, A.R., Colombo, E., Luzi, L., Martinelli, P., Volorio, S., Bernard, L., Meani, N., Bergomas, R., Alcalay, M. and Pelicci, P.G. (2006) 'Cytoplasmic localization of NPM in myeloid leukemias is dictated by gain-of-function mutations that create a functional nuclear export signal', *Oncogene*, 25(31), pp. 4376-4380.

Marnett, L.J. and Plastaras, J.P. (2001) 'Endogenous DNA damage and mutation', *Trends in Genetics*, 17(4), pp. 214-221.

Matsuda, A. and Sasaki, T. (2005) 'Antitumor activity of sugar-modified cytosine nucleosides', *Cancer Science*, 95(2), pp. 105-111.

Matsuura, S., Yan, M., Lo, M.-C., Ahn, E.-Y., Weng, S., Dangoor, D., Matin, M., Higashi, T., Feng, G.-S. and Zhang, D.-E. (2012) 'Negative effects of GM-CSF signaling in a murine model of t(8;21)–induced leukemia', *Blood*, 119(13), p. 3155.

Maynard, S., Schurman, S.H., Harboe, C., de Souza-Pinto, N.C. and Bohr, V.A. (2009) 'Base excision repair of oxidative DNA damage and association with cancer and aging', *Carcinogenesis*, 30(1), pp. 2-10.

Medeiros, B.C., Fathi, A.T., DiNardo, C.D., Pollyea, D.A., Chan, S.M. and Swords, R. (2016) 'Isocitrate dehydrogenase mutations in myeloid malignancies', *Leukemia*, 31, p. 272.

Meng, Q., Su, T., O'Neill, J.P. and Walker, V.E. (2002) 'Molecular analysis of mutations at the HPRT and TK loci of human lymphoblastoid cells after combined treatments with 3′-azido-3′-deoxythymidine and 2′,3′-dideoxyinosine†', *Environmental and Molecular Mutagenesis*, 39(4), pp. 282-295.

Meng, Q., Su, T., Olivero, O., Poirier, M., Shi, X., Ding, X. and Walker, V. (2000) 'Relationships between DNA incorporation, mutant frequency, and loss of heterozygosity at the TK locus in human lymphoblastoid cells exposed to 3'-azido-3'-deoxythymidine', *Toxicology Science*, 54(2), pp. 322-329.

Mercher, T., Quivoron, C., Couronné, L., Bastard, C., Vainchenker, W. and Bernard, O.A. (2012) 'TET2, a tumor suppressor in hematological disorders', *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer*, 1825(2), pp. 173-177.

Merlo, L.M.F., Pepper, J.W., Reid, B.J. and Maley, C.C. (2006) 'Cancer as an evolutionary and ecological process', *Nature Reviews Cancer*, 6, p. 924.

Mertens, F., Johansson, B., Fioretos, T. and Mitelman, F. (2015) 'The emerging complexity of gene fusions in cancer', *Nature Reviews Cancer*, 15, p. 371.

Mertens, F., Johansson, B. and Mitelman, F. (1993) 'Age- and gender-related heterogeneity of cancer chromosome aberrations', *Cancer Genetics and Cytogenetics*, 70(1), pp. 6-11.

Meyers, S., Downing, J.R. and Hiebert, S.W. (1993) 'Identification of AML-1 and the (8;21) translocation protein (AML-1/ETO) as sequence-specific DNA-binding proteins: the runt homology domain is required for DNA binding and protein-protein interactions', *Molecular and Cellular Biology*, 13(10), pp. 6336-6345.

Mills, R.E., Pittard, W.S., Mullaney, J.M., Farooq, U., Creasy, T.H., Mahurkar, A.A., Kemeza, D.M., Strassler, D.S., Ponting, C.P., Webber, C. and Devine, S.E. (2011) 'Natural genetic variation caused by small insertions and deletions in the human genome', *Genome Research*, 21(6), pp. 830-839.

Minotti, G., Menna, P., Salvatorelli, E., Cairo, G. and Gianni, L. (2004) 'Anthracyclines: Molecular Advances and Pharmacologic Developments in Antitumor Activity and Cardiotoxicity', *Pharmacological Reviews*, 56(2), p. 185.

Mistry, A.R., Felix, C.A., Whitmarsh, R.J., Mason, A., Reiter, A., Cassinat, B., Parry, A., Walz, C., Wiemels, J.L., Segal, M.R., Adès, L., Blair, I.A., Osheroff, N., Peniket, A.J., Lafage-Pochitaloff, M., Cross, N.C.P., Chomienne, C., Solomon, E., Fenaux, P. and Grimwade, D. (2005) 'DNA Topoisomerase II in Therapy-Related Acute Promyelocytic Leukemia', *New England Journal of Medicine*, 352(15), pp. 1529-1538.
Miyamoto, T., Weissman, I.L. and Akashi, K. (2000) 'AML1/ETO-expressing nonleukemic stem cells in acute myelogenous leukemia with 8;21 chromosomal translocation', *Proceedings of the National Academy of Sciences of the United States of America*, 97(13), pp. 7521-7526.

Mohamedali, A.M., Smith, A.E., Gaken, J., Lea, N.C., Mian, S.A., Westwood, N.B., Strupp, C., Gattermann, N., Germing, U. and Mufti, G.J. (2009) 'Novel TET2 Mutations Associated With UPD4q24 in Myelodysplastic Syndrome', *Journal of Clinical Oncology*, 27(24), pp. 4002-4006.

Molenaar, R.J., Thota, S., Nagata, Y., Patel, B., Clemente, M., Przychodzen, B., Hirsh, C., Viny, A.D., Hosano, N., Bleeker, F.E., Meggendorfer, M., Alpermann, T., Shiraishi, Y., Chiba, K., Tanaka, H., van Noorden, C.J.F., Radivoyevitch, T., Carraway, H.E., Makishima, H., Miyano, S., Sekeres, M.A., Ogawa, S., Haferlach, T. and Maciejewski, J.P. (2015) 'Clinical and biological implications of ancestral and non-ancestral IDH1 and IDH2 mutations in myeloid neoplasms', *Leukemia*, 29, p. 2134.

Mori, H., Colman, S.M., Xiao, Z., Ford, A.M., Healy, L.E., Donaldson, C., Hows, J.M., Navarrete, C. and Greaves, M. (2002) 'Chromosome translocations and covert leukemic clones are generated during normal fetal development', *Proceedings of the National Academy of Sciences of the United States of America*, 99(12), pp. 8242-8247.
Morreall, J., Limpose, K., Sheppard, C., Kow, Y.W., Werner, E. and Doetsch, P.W. (2015) 'Inactivation of a common OGG1 variant by TNF-alpha in mammalian cells', *DNA repair*, 26, pp. 15-22.

Moyer, R., Briley, D., Johnsen, A., Stewart, U. and Shaw, B.R. (1993) 'Echinomycin, a bis-intercalating agent, induces C→T mutations via cytosine deamination', *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 288(2), pp. 291-300.

Mrózek, K., Heerema, N.A. and Bloomfield, C.D. (2004) 'Cytogenetics in acute leukemia', *Blood Reviews*, 18(2), pp. 115-136.

Muindi, J.R.F., Sinha, B.K., Gianni, L. and Myers, C.E. (1984) 'Hydroxyl radical production and DNA damage induced by anthracycline-iron complex', *FEBS Letters*, 172(2), pp. 226-230.

Nerlov, C. (2004) 'C/EBP[alpha] mutations in acute myeloid leukaemias', *Nat Rev Cancer*, 4(5), pp. 394-400.

Ng, M.H.L., Chung, Y.F., Lo, K.W., Wickham, N.W.R., Lee, J.C.K. and Huang, D.P. (1997) 'Frequent Hypermethylation of p16 & p15 Genes in Multiple Myeloma', *Blood*, 89(7), p. 2500.

Nielsen, D., Maare, C. and Skovsgaard, T. (1996) 'Cellular resistance to anthracyclines', *General Pharmacology: The Vascular System*, 27(2), pp. 251-255.
Nik-Zainal, S., Alexandrov, Ludmil B., Wedge, David C., Van Loo, P., Greenman, Christopher D., Raine, K., Jones, D., Hinton, J., Marshall, J., Stebbings, Lucy A., Menzies, A., Martin, S., Leung, K., Chen, L., Leroy, C., Ramakrishna, M., Rance, R., Lau, King W., Mudie, Laura J., Varela, I., McBride, David J., Bignell, Graham R., Cooke, Susanna L., Shlien, A., Gamble, J., Whitmore, I., Maddison, M., Tarpey, Patrick S., Davies, Helen R., Papaemmanuil, E., Stephens, Philip J., McLaren, S., Butler, Adam P., Teague, Jon W., Jönsson, G., Garber, Judy E., Silver, D., Miron, P., Fatima, A., Boyault, S., Langerød, A., Tutt, A., Martens, John W., Aparicio, Samuel A., Borg, Å., Salomon, Anne V., Thomas, G., Børresen-Dale, A.-L., Richardson, Andrea L., Neuberger, Michael S., Futreal, P A., Campbell, Peter J., Stratton, Michael R. and the Breast Cancer Working Group of the International Cancer Genome, C. (2012) 'Mutational Processes Molding the Genomes of 21 Breast Cancers', *Cell*, 149(5-10), pp. 979-993.

Norbury, C.J. and Hickson, I.D. (2001) 'Cellular Responses to DNA Damage', *Annual Review of Pharmacology and Toxicology*, 41(1), pp. 367-401.

Nowell, P.C. (1960) 'A minute chromosome in human chronic granulocytic leukemia', *Science*, 132, pp. 497-501.

Nowell, P.C. (1976) 'The clonal evolution of tumor cell populations', *Science*, 194(4260), p. 23.

Ogawa, E., Maruyama, M., Kagoshima, H., Inuzuka, M., Lu, J., Satake, M., Shigesada, K. and Ito, Y. (1993) 'PEBP2/PEA2 represents a family of transcription factors homologous to the products of the Drosophila runt gene and the human AML1 gene', *Proceedings of the National Academy of Sciences of the United States of America*, 90(14), pp. 6859-6863.

Okumura, A.J., Peterson, L.F., Okumura, F., Boyapati, A. and Zhang, D.-E. (2008) 't(8;21)(q22;q22) fusion proteins preferentially bind to duplicated AML1/RUNX1 DNA-binding sequences to differentially regulate gene expression', *Blood*, 112(4), p. 1392.
Olivier, M., Hussain, S.P., Caron de Fromentel, C., Hainaut, P. and Harris, C. (2004) 'TP53 mutation spectra and load: a tool for generating hypotheses on the etiology of cancer', *IARC Science Publication*, 157, pp. 247-270.

Olney, H.J., Mitelman, F., Johansson, B., Mrózek, K., Berger, R. and Rowley, J.D. (2002) 'Unique balanced chromosome abnormalities in treatment-related myelodysplastic syndromes and acute myeloid leukemia: Report from an International Workshop†', *Genes, Chromosomes and Cancer*, 33(4), pp. 413-423.

Ommen Hans, B., Østergaard, M., Yan, M., Brændstrup, K., Zhang, D.-E. and Hokland, P. (2010) 'Persistent altered fusion transcript splicing identifies RUNX1-RUNX1T1+ AML patients likely to relapse', *European Journal of Haematology*, 84(2), pp. 128-132.
Opresko, P.L., Calvo, J.P. and von Kobbe, C. (2007) 'Role for the Werner syndrome protein in the promotion of tumor cell growth', *Mechanisms of Ageing and Development*, 128(7), pp. 423-436.

Ortiz, S.C., Aguirre, S.J., Flores, S., Maldonado, C., Mejía, J. and Salinas, L. (2017) 'Spectrum of CFTR gene mutations in Ecuadorian cystic fibrosis patients: the second report of the p.H609R mutation', *Molecular Genetics & Genomic Medicine*, 5(6), pp. 751-757.

Ortmann, C., Silber, Y., Baxter, J., Bellosillo, B., Doehner, K., Campbell, P.J. and Green, A.R. (2013) 'Order Matters: Sequence Of Mutation Acquisition In Myeloproliferative Neoplasms Impacts Disease Pathogenesis and Stem Cell Potency', *Blood*, 122(21), p. 2888.

Ortmann, C.A., Kent, D.G., Nangalia, J., Silber, Y., Wedge, D.C., Grinfeld, J., Baxter, E.J., Massie, C.E., Papaemmanuil, E., Menon, S., Godfrey, A.L., Dimitropoulou, D., Guglielmelli, P., Bellosillo, B., Besses, C., Döhner, K., Harrison, C.N., Vassiliou, G.S., Vannucchi, A., Campbell, P.J. and Green, A.R. (2015) 'Effect of Mutation Order on Myeloproliferative Neoplasms', *New England Journal of Medicine*, 372(7), pp. 601-612.
Osman, D., Gobert, V., Ponthan, F., Heidenreich, O., Haenlin, M. and Waltzer, L. (2009) 'A Drosophila model identifies calpains as modulators of the human leukemogenic fusion protein AML1-ETO', *Proceedings of the National Academy of Sciences of the United States of America*, 106(29), pp. 12043-12048.

Pabst, T. and Mueller, B.U. (2007) 'Transcriptional dysregulation during myeloid transformation in AML', *Oncogene*, 26(47), pp. 6829-6837.

Pabst, T. and Mueller, B.U. (2009) 'Complexity of CEBPA Dysregulation in Human Acute Myeloid Leukemia', *Clinical Cancer Research*, 15(17), pp. 5303-5307.

Pabst, T., Mueller, B.U., Zhang, P., Radomska, H.S., Narravula, S., Schnittger, S., Behre, G., Hiddemann, W. and Tenen, D.G. (2001) 'Dominant-negative mutations of CEBPA, encoding CCAAT/enhancer binding protein-α (C/EBPα), in acute myeloid leukemia', *Nature Genetics*, 27, p. 263.

Pandolfi, P.P. (2001) 'Oncogenes and tumor suppressors in the molecular pathogenesis of acute promyelocytic leukemia', *Human Molecular Genetics*, 10(7), pp. 769-775.

Paschka, P., Marcucci, G., Ruppert, A.S., Mrózek, K., Chen, H., Kittles, R.A., Vukosavljevic, T., Perrotti, D., Vardiman, J.W., Carroll, A.J., Kolitz, J.E., Larson, R.A. and Bloomfield, C.D. (2006) 'Adverse Prognostic Significance of KIT Mutations in Adult Acute Myeloid Leukemia With inv(16) and t(8;21): A Cancer and Leukemia Group B Study', *Journal of Clinical Oncology*, 24(24), pp. 3904-3911.

Pastor, W.A., Aravind, L. and Rao, A. (2013) 'TETonic shift: biological roles of TET proteins in DNA demethylation and transcription', *Nature Reviews Molecular Cell Biology*, 14, p. 341.

Patel, J.L., Schumacher, J.A., Frizzell, K., Sorrells, S., Shen, W., Clayton, A., Jattani, R. and Kelley, T.W. (2017) 'Coexisting and cooperating mutations in NPM1-mutated acute myeloid leukemia', *Leukemia Research*, 56, pp. 7-12.

Paulsson, K. and Johansson, B. (2007) 'Trisomy 8 as the sole chromosomal aberration in acute myeloid leukemia and myelodysplastic syndromes', *Pathologie Biologie*, 55(1), pp. 37-48.

Pedersen-Bjergaard, J., Andersen, M.K., Andersen, M.T. and Christiansen, D.H. (2008) 'Genetics of therapy-related myelodysplasia and acute myeloid leukemia', *Leukemia*, 22(2), pp. 240-248.

Pedersen-Bjergaard, J., Andersen, M.T. and Andersen, M.K. (2007) 'Genetic Pathways in the Pathogenesis of Therapy-Related Myelodysplasia and Acute Myeloid Leukemia', *ASH Education Program Book*, 2007(1), pp. 392-397.

Pedersen-Bjergaard, J. and Philip, P. (1991) 'Balanced translocations involving chromosome bands 11q23 and 21q22 are highly characteristic of myelodysplasia and leukemia following therapy with cytostatic agents targeting at DNA-topoisomerase II [letter]', *Blood*, 78(4), p. 1147.

Pendleton, M., Lindsey, R.H., Felix, C.A., Grimwade, D. and Osheroff, N. (2014) 'Topoisomerase II and leukemia', *Annals of the New York Academy of Sciences*, 1310(1), pp. 98-110.

Perrino, F.W., Mazur, D.J., Ward, H. and Harvey, S. (1999) 'Exonucleases and the incorporation of aranucleotides into DNA', *Cell Biochemistry and Biophysics*, 30(3), pp. 331-352.

Peterson, L.F., Boyapati, A., Ahn, E.-Y., Biggs, J.R., Okumura, A.J., Lo, M.-C., Yan, M. and Zhang, D.-E. (2007) 'Acute myeloid leukemia with the 8q22;21q22 translocation: secondary mutational events and alternative t(8;21) transcripts', *Blood*, 110(3), pp. 799-805.

Petrie, K. and Zelent, A. (2007) 'AML1/ETO, a promiscuous fusion oncoprotein', *Blood*, 109(10), pp. 4109-4110.

Petrovick, M.S., Hiebert, S.W., Friedman, A.D., Hetherington, C.J., Tenen, D.G. and Zhang, D.-E. (1998) 'Multiple Functional Domains of AML1: PU.1 and C/EBPα Synergize with Different Regions of AML1', *Molecular and Cellular Biology*, 18(7), pp. 3915-3925.

Pfeifer, G.P. and Hainaut, P. (2003) 'On the origin of G→T transversions in lung cancer', *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 526(1), pp. 39-43.

Pfeifer, G.P., You, Y.-H. and Besaratinia, A. (2005) 'Mutations induced by ultraviolet light', *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 571(1), pp. 19-31.

Pierceall William, E., Mukhopadhyay, T., Goldberg Leonard, H. and Ananthaswamy Honnavara, N. (1991) 'Mutations in the p53 tumor suppressor gene in human cutaneous squamous cell carcinomas', *Molecular Carcinogenesis*, 4(6), pp. 445-449.
Pluskota-Karwatka, D. (2008) 'Modifications of nucleosides by endogenous mutagens–DNA adducts arising from cellular processes', *Bioorganic Chemistry*, 36(4), pp. 198-213.

Powis, G. (1989) 'Free radical formation by antitumor quinones', *Free Radical Biology and Medicine*, 6(1), pp. 63-101.

Prakasha Gowda, A.S., Polizzi, J.M., Eckert, K.A. and Spratt, T.E. (2010) 'Incorporation of Gemcitabine and Cytarabine into DNA by DNA Polymerase β and Ligase III/XRCC1', *Biochemistry*, 49(23), pp. 4833-4840.

Pratcorona, M., Abbas, S., Sanders, M.A., Koenders, J.E., Kavelaars, F.G., Erpelinck-Verschueren, C.A.J., Zeilemakers, A., Löwenberg, B. and Valk, P.J.M. (2012) 'Acquired mutations in ASXL1 in acute myeloid leukemia: prevalence and prognostic value', *Haematologica*, 97(3), pp. 388-392.

Ptasinska, A., Assi, S.A., Mannari, D., James, S.R., Williamson, D., Dunne, J., Hoogenkamp, M., Wu, M., Care, M., McNeill, H., Cauchy, P., Cullen, M., Tooze, R.M., Tenen, D.G., Young, B.D., Cockerill, P.N., Westhead, D.R., Heidenreich, O. and Bonifer, C. (2012) 'Depletion of RUNX1/ETO in t(8;21) AML cells leads to genome-wide changes in chromatin structure and transcription factor binding', *Leukemia*, 26, p. 1829.

Quentmeier, H., Martelli, M.P., Dirks, W.G., Bolli, N., Liso, A., MacLeod, R.A.F., Nicoletti, I., Mannucci, R., Pucciarini, A., Bigerna, B., Martelli, M.F., Mecucci, C., Drexler, H.G. and Falini, B. (2005) 'Cell line OCI/AML3 bears exon-12 NPM gene mutation-A and cytoplasmic expression of nucleophosmin', *Leukemia*, 19, p. 1760.

Quina Ana, S., Gameiro, P., Sá da Costa, M., Telhada, M. and Parreira, L. (2000) 'PML-RARA fusion transcripts in irradiated and normal hematopoietic cells', *Genes, Chromosomes and Cancer*, 29(3), pp. 266-275.

Rasmussen, K.D., Jia, G., Johansen, J.V., Pedersen, M.T., Rapin, N., Bagger, F.O., Porse, B.T., Bernard, O.A., Christensen, J. and Helin, K. (2015) 'Loss of TET2 in hematopoietic cells leads to DNA hypermethylation of active enhancers and induction of leukemogenesis', *Genes & Development*, 29(9), pp. 910-922.

Rau, R. and Brown, P. (2009) 'Nucleophosmin (NPM1) Mutations in Adult and Childhood Acute Myeloid Leukemia: Towards Definition of a New Leukemia Entity', *Hematological oncology*, 27(4), pp. 171-181.

Reikvam, H., Hatfield, K.J., Kittang, A.O., Hovland, R. and Bruserud, Ø. (2011) 'Acute Myeloid Leukemia with the t(8;21) Translocation: Clinical Consequences and Biological Implications', *Journal of Biomedicine and Biotechnology*, 2011, p. 104631.
Reilly John, T. (2004) 'Pathogenesis of acute myeloid leukaemia and inv(16)(p13;q22): a paradigm for understanding leukaemogenesis?', *British Journal of Haematology*, 128(1), pp. 18-34.

Renneville, A., Roumier, C., Biggio, V., Nibourel, O., Boissel, N., Fenaux, P. and Preudhomme, C. (2008) 'Cooperating gene mutations in acute myeloid leukemia: a review of the literature', *Leukemia*, 22(5), pp. 915-931.

Reuter, C.W.M., Morgan, M.A. and Bergmann, L. (2000) 'Targeting the Ras signaling pathway: a rational, mechanism-based treatment for hematologic malignancies?', *Blood*, 96(5), p. 1655.

Richardson, C. and Jasin, M. (2000) 'Frequent chromosomal translocations induced by DNA double-strand breaks', *Nature*, 405, p. 697.

Roca, J. (1995) 'The mechanisms of DNA topoisomerases', *Trends in Biochemical Sciences*, 20(4), pp. 156-160.

Rodgers, K. and McVey, M. (2016) 'Error-prone repair of DNA double-strand breaks', *Journal of cellular physiology*, 231(1), pp. 15-24.

Rothkamm, K., Krüger, I., Thompson, L.H. and Löbrich, M. (2003) 'Pathways of DNA Double-Strand Break Repair during the Mammalian Cell Cycle', *Molecular and Cellular Biology*, 23(16), pp. 5706-5715.

Roumier, C., Lejeune-Dumoulin, S., Renneville, A., Goethgeluck, A.S., Philippe, N., Fenaux, P. and Preudhomme, C. (2006) 'Cooperation of activating Ras//rtk signal transduction pathway mutations and inactivating myeloid differentiation gene mutations in M0 AML: a study of 45 patients', *Leukemia*, 20(3), pp. 433-436.

Rowley, J.D. (1973) 'A New Consistent Chromosomal Abnormality in Chronic Myelogenous Leukaemia identified by Quinacrine Fluorescence and Giemsa Staining', *Nature*, 243, p. 290.

Rubin, A.F. and Green, P. (2009) 'Mutation patterns in cancer genomes', *Proceedings of the National Academy of Sciences*, 106(51), p. 21766.

Sakumi, K., Tominaga, Y., Furuichi, M., Xu, P., Tsuzuki, T., Sekiguchi, M. and Nakabeppu, Y. (2003) 'Ogg1 Knockout-associated Lung Tumorigenesis and Its Suppression by Mth1 Gene Disruption', *Cancer Research*, 63(5), p. 902.

Sanderson, R.N., Johnson, P.R.E., Moorman, A.V., Roman, E., Willett, E., Taylor, P.R., Proctor, S.J., Bown, N., Ogston, S. and Bowen, D.T. (2006a) 'Population-based demographic study of karyotypes in 1709 patients with adult Acute Myeloid Leukemia', *Leukemia*, 20(3), pp. 444-450.

Sanderson, R.N., Johnson, P.R.E., Moorman, A.V., Roman, E., Willett, E., Taylor, P.R., Proctor, S.J., Bown, N., Ogston, S. and Bowen, D.T. (2006b) 'Population-based demographic study of karyotypes in 1709 patients with adult Acute Myeloid Leukemia', *Leukemia*, 20, p. 444.

Santos, F.P.S., Jones, D., Qiao, W., Cortes, J.E., Ravandi, F., Estey, E.E., Verma, D., Kantarjian, H. and Borthakur, G. (2011) 'Prognostic value of FLT3 mutations among different cytogenetic subgroups in acute myeloid leukemia', *Cancer*, 117(10), pp. 2145-2155.

Saultz, J.N. and Garzon, R. (2016) 'Acute Myeloid Leukemia: A Concise Review', *Journal of Clinical Medicine*, 5(3), p. 33.

Saygin, C. and Carraway, H.E. (2017) 'Emerging therapies for acute myeloid leukemia', *Journal of Hematology & Oncology*, 10(1), p. 93.

Schessl, C., Rawat, V.P.S., Cusan, M., Deshpande, A., Kohl, T.M., Rosten, P.M., Spiekermann, K., Humphries, R.K., Schnittger, S., Kern, W., Hiddemann, W., Quintanilla-Martinez, L., Bohlander, S.K., Feuring-Buske, M. and Buske, C. (2005) 'The AML1-ETO fusion gene and the FLT3 length mutation collaborate in inducing acute leukemia in mice', *Journal of Clinical Investigation*, 115(8), pp. 2159-2168.

Schindler, T., Bornmann, W., Pellicena, P., Miller, W.T., Clarkson, B. and Kuriyan, J. (2000) 'Structural Mechanism for STI-571 Inhibition of Abelson Tyrosine Kinase', *Science*, 289(5486), p. 1938.

Schnittger, S., Haferlach, C., Nadarajah, N., Alpermann, T., Meggendorfer, M., Perglerová, K., Kern, W. and Haferlach, T. (2014) 'In AML Secondary to MDS NPM1 Mutations Are Late Events, Less Frequent, and Associated with a Different Pattern of Molecular Mutations Than in De Novo AML', *Blood*, 124(21), pp. 700-700.

Schoch, C. and Haferlach, T. (2002) 'Cytogenetics in acute myeloid leukemia', *Current Oncology Reports*, 4(5), pp. 390-397.

Schweitzer, B.I., Mikita, T., Kellogg, G.W., Gardner, K.H. and Beardsley, G.P. (1994) 'Solution Structure of a DNA Dodecamer Containing the Anti-Neoplastic Agent

Arabinosylcytosine: Combined Use of NMR, Restrained Molecular Dynamics, and Full Relaxation Matrix Refinement', *Biochemistry*, 33(38), pp. 11460-11475.

Sherr, C.J. (1998) 'Tumor surveillance via the ARF–p53 pathway', *Genes & Development*, 12(19), pp. 2984-2991.

Shi, Z., Azuma, A., Sampath, D., Li, Y.-X., Huang, P. and Plunkett, W. (2001) 'S-Phase Arrest by Nucleoside Analogues and Abrogation of Survival without Cell Cycle Progression by 7-Hydroxystaurosporine', *Cancer Research*, 61(3), p. 1065.

Shigemizu, D., Fujimoto, A., Akiyama, S., Abe, T., Nakano, K., Boroevich, K.A., Yamamoto, Y., Furuta, M., Kubo, M., Nakagawa, H. and Tsunoda, T. (2013) 'A practical method to detect SNVs and indels from whole genome and exome sequencing data', *Scientific Reports*, 3, p. 2161.

Shih, A.H., Abdel-Wahab, O., Patel, J.P. and Levine, R.L. (2012) 'The role of mutations in epigenetic regulators in myeloid malignancies', *Nature Reviews Cancer*, 12, p. 599.
Shimada, A., Taki, T., Tabuchi, K., Tawa, A., Horibe, K., Tsuchida, M., Hanada, R., Tsukimoto, I. and Hayashi, Y. (2006) 'KIT mutations, and not FLT3 internal tandem duplication, are strongly associated with a poor prognosis in pediatric acute myeloid leukemia with t(8;21): a study of the Japanese Childhood AML Cooperative Study Group', *Blood*, 107(5), p. 1806.

Shinmura, K. and Yokota, J. (2001) 'The OGG1 Gene Encodes a Repair Enzyme for Oxidatively Damaged DNA and Is Involved in Human Carcinogenesis', *Antioxidants & Redox Signaling*, 3(4), pp. 597-609.

Shipley, J.L. and Butera, J.N. (2009) 'Acute myelogenous leukemia', *Experimental Hematology*, 37(6), pp. 649-658.

Shlush, L.I., Zandi, S., Mitchell, A., Chen, W.C., Brandwein, J.M., Gupta, V., Kennedy, J.A., Schimmer, A.D., Schuh, A.C., Yee, K.W., McLeod, J.L., Doedens, M., Medeiros, J.J.F., Marke, R., Kim, H.J., Lee, K., McPherson, J.D., Hudson, T.J., Pan-Leukemia Gene Panel Consortium, T.H., Brown, A.M.K., Trinh, Q.M., Stein, L.D., Minden, M.D., Wang, J.C.Y. and Dick, J.E. (2014) 'Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia', *Nature*, 506(7488), pp. 328-333.

Sie, L., Loong, S. and Tan, E.K. (2009) 'Utility of lymphoblastoid cell lines', *Journal of Neuroscience Research*, 87(9), pp. 1953-1959.

Siebert, A.E., Sanchez, A.L., Dinda, S. and Moudgil, V.K. (2011) 'Effects of Estrogen Metabolite 2-Methoxyestradiol on Tumor Suppressor Protein p53 and Proliferation of Breast Cancer Cells', *Systems Biology in Reproductive Medicine*, 57(6), pp. 279-287.
Silva, M.J., Costa, P., Dias, A., Valente, M., Louro, H. and Boavida, M.G. (2005) 'Comparative analysis of the mutagenic activity of oxaliplatin and cisplatin in the Hprt gene of CHO cells', *Environmental and Molecular Mutagenesis*, 46(2), pp. 104-115.

Simonelli, V., Camerini, S., Mazzei, F., Van Loon, B., Allione, A., D'Errico, M., Barone, F., Minoprio, A., Ricceri, F., Guarrera, S., Russo, A., Dalhus, B., Crescenzi, M., Hübscher, U., Bjørås, M., Matullo, G. and Dogliotti, E. (2013) 'Genotype–phenotype analysis of S326C OGG1 polymorphism: a risk factor for oxidative pathologies', *Free Radical Biology and Medicine*, 63, pp. 401-409.

Sinclair, A., Latif, A.L. and Holyoake, T.L. (2013) 'Targeting survival pathways in chronic myeloid leukaemia stem cells', *British Journal of Pharmacology*, 169(8), pp. 1693-1707.

Sinenko, S.A., Hung, T., Moroz, T., Tran, Q.-M., Sidhu, S., Cheney, M.D., Speck, N.A. and Banerjee, U. (2010) 'Genetic manipulation of AML1-ETO–induced expansion of hematopoietic precursors in a Drosophila model', *Blood*, 116(22), pp. 4612-4620.

Sinha, C., Cunningham, L.C. and Liu, P.P. (2015) 'Core binding factor AML: New prognostic categories and therapeutic opportunities', *Seminars in hematology*, 52(3), pp. 215-222.

Sjöblom, T., Jones, S., Wood, L.D., Parsons, D.W., Lin, J., Barber, T.D., Mandelker, D., Leary, R.J., Ptak, J., Silliman, N., Szabo, S., Buckhaults, P., Farrell, C., Meeh, P., Markowitz, S.D., Willis, J., Dawson, D., Willson, J.K.V., Gazdar, A.F., Hartigan, J., Wu, L., Liu, C., Parmigiani, G., Park, B.H., Bachman, K.E., Papadopoulos, N., Vogelstein, B., Kinzler, K.W. and Velculescu, V.E. (2006) 'The Consensus Coding Sequences of Human Breast and Colorectal Cancers', *Science*, 314(5797), p. 268.

Smith, A.E., Mohamedali, A.M., Kulasekararaj, A., Lim, Z., Gäken, J., Lea, N.C., Przychodzen, B., Mian, S.A., Nasser, E.E., Shooter, C., Westwood, N.B., Strupp, C., Gattermann, N., Maciejewski, J.P., Germing, U. and Mufti, G.J. (2010) 'Next-generation sequencing of the TET2 gene in 355 MDS and CMML patients reveals low-abundance mutant clones with early origins, but indicates no definite prognostic value', *Blood*, 116(19), p. 3923.

Smith, S.M., Le Beau, M.M., Huo, D., Karrison, T., Sobecks, R.M., Anastasi, J., Vardiman, J.W., Rowley, J.D. and Larson, R.A. (2003) 'Clinical-cytogenetic associations in 306 patients with therapy-related myelodysplasia and myeloid leukemia: the University of Chicago series', *Blood*, 102(1), p. 43.

Solh, M., Yohe, S., Weisdorf, D. and Ustun, C. (2014) 'Core-binding factor acute myeloid leukemia: Heterogeneity, monitoring, and therapy', *American Journal of Hematology*, 89(12), pp. 1121-1131.

Song, J., Mercer, D., Hu, X., Liu, H. and Li, M.M. (2011) 'Common Leukemia- and Lymphoma-Associated Genetic Aberrations in Healthy Individuals', *The Journal of Molecular Diagnostics*, 13(2), pp. 213-219.

Sood, R., Hansen, N.F., Donovan, F.X., Carrington, B., Bucci, D., Maskeri, B., Young, A., Trivedi, N.S., Kohlschmidt, J., Stone, R.M., Caligiuri, M.A., Chandrasekharappa, S.C., Marcucci, G., Mullikin, J.C., Bloomfield, C.D. and Liu, P. (2015) 'Somatic mutational landscape of AML with inv(16) or t(8;21) identifies patterns of clonal evolution in relapse leukemia', *Leukemia*.

Sportoletti, P., Grisendi, S., Majid, S.M., Cheng, K., Clohessy, J.G., Viale, A., Teruya-Feldstein, J. and Pandolfi, P.P. (2008) 'Npm1 is a haploinsufficient suppressor of myeloid and lymphoid malignancies in the mouse', *Blood*, 111(7), p. 3859.

Sportoletti, P., Varasano, E., Rossi, R., Bereshchenko, O., Cecchini, D., Gionfriddo, I., Bolli, N., Tiacci, E., Intermesoli, T., Zanghì, P., Masciulli, A., Martelli, M.P., Falzetti, F.,

Martelli, M.F. and Falini, B. (2013) 'The human NPM1 mutation A perturbs megakaryopoiesis in a conditional mouse model', *Blood*, 121(17), p. 3447.

Stanczyk, M., Sliwinski, T., Cuchra, M., Zubowska, M., Bielecka-Kowalska, A., Kowalski, M., Szemraj, J., Mlynarski, W. and Majsterek, I. (2011) 'The association of polymorphisms in DNA base excision repair genes XRCC1, OGG1 and MUTYH with the risk of childhood acute lymphoblastic leukemia', *Molecular Biology Reports*, 38(1), pp. 445-451.

Stasi, R., Evangelista, M.L., Buccisano, F., Venditti, A. and Amadori, S. (2008) 'Gemtuzumab ozogamicin in the treatment of acute myeloid leukemia', *Cancer Treatment Reviews*, 34(1), pp. 49-60.

Steensma, D.P., Bejar, R., Jaiswal, S., Lindsley, R.C., Sekeres, M.A., Hasserjian, R.P. and Ebert, B.L. (2015) 'Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes', *Blood*, 126(1), p. 9.

Stein, E.M. and Tallman, M.S. (2016) 'Emerging therapeutic drugs for AML', *Blood*, 127(1), pp. 71-78.

Stirewalt, D.L. and Radich, J.P. (2003) 'The role of FLT3 in haematopoietic malignancies', *Nat Rev Cancer*, 3(9), pp. 650-665.

Stoklosa, T., Poplawski, T., Koptyra, M., Nieborowska-Skorska, M., Basak, G., Slupianek, A., Rayevskaya, M., Seferynska, I., Herrera, L., Blasiak, J. and Skorski, T. (2008) 'BCR/ABL Inhibits Mismatch Repair to Protect from Apoptosis and Induce Point Mutations', *Cancer Research*, 68(8), p. 2576.

Sun, Q.Y., Ding, L.W., Tan, K.T., Chien, W., Mayakonda, A., Lin, D.C., Loh, X.Y., Xiao, J.F., Meggendorfer, M., Alpermann, T., Garg, M., Lim, S.L., Madan, V., Hattori, N., Nagata, Y., Miyano, S., Yeoh, A.E.J., Hou, H.A., Jiang, Y.Y., Takao, S., Liu, L.Z., Tan, S.Z., Lill, M., Hayashi, M., Kinoshita, A., Kantarjian, H.M., Kornblau, S.M., Ogawa, S., Haferlach, T., Yang, H. and Koeffler, H.P. (2016) 'Ordering of mutations in acute myeloid leukemia with partial tandem duplication of MLL (MLL-PTD)', *Leukemia*, 31, p. 1.

Sussman, H.E., Olivero, O.A., Meng, Q., Pietras, S.M., Poirier, M.C., O'Neill, J.P., Finette, B.A., Bauer, M.J. and Walker, V.E. (1999) 'Genotoxicity of 3′-azido-3′-deoxythymidine in the human lymphoblastoid cell line, TK6: relationships between DNA incorporation, mutant frequency, and spectrum of deletion mutations in HPRT', *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 429(2), pp. 249-259.

Suzuki, T., Harashima, H. and Kamiya, H. (2010) 'Effects of base excision repair proteins on mutagenesis by 8-oxo-7,8-dihydroguanine (8-hydroxyguanine) paired with cytosine and adenine', *DNA Repair*, 9(5), pp. 542-550.

Szikriszt, B., Póti, Á., Pipek, O., Krzystanek, M., Kanu, N., Molnár, J., Ribli, D., Szeltner, Z., Tusnády, G.E., Csabai, I., Szallasi, Z., Swanton, C. and Szüts, D. (2016) 'A comprehensive survey of the mutagenic impact of common cancer cytotoxics', *Genome Biology*, 17(1), p. 99.

Takeuchi, S., Bartram, C.R., Seriu, T., Miller, C.W., Tobler, A., Janssen, J.W., Reiter, A., Ludwig, W.D., Zimmermann, M. and Schwaller, J. (1995) 'Analysis of a family of cyclin-dependent kinase inhibitors: p15/MTS2/INK4B, p16/MTS1/INK4A, and p18 genes in acute lymphoblastic leukemia of childhood', *Blood*, 86(2), p. 755.

Tang, J.-L., Hou, H.-A., Chen, C.-Y., Liu, C.-Y., Chou, W.-C., Tseng, M.-H., Huang, C.-F., Lee, F.-Y., Liu, M.-C., Yao, M., Huang, S.-Y., Ko, B.-S., Hsu, S.-C., Wu, S.-J., Tsay, W., Chen, Y.-C., Lin, L.-I. and Tien, H.-F. (2009) 'AML1/RUNX1 mutations in 470 adult patients with de novo acute myeloid leukemia: prognostic implication and interaction with other gene alterations', *Blood*, 114(26), pp. 5352-5361.

Tenen, D.G. (2003) 'Disruption of differentiation in human cancer: AML shows the way', *Nat Rev Cancer*, 3(2), pp. 89-101.

Teng, G. and Papavasiliou, F.N. (2007) 'Immunoglobulin Somatic Hypermutation', *Annual Review of Genetics*, 41(1), pp. 107-120.

Thiede, C., Koch, S., Creutzig, E., Steudel, C., Illmer, T., Schaich, M. and Ehninger, G. (2006) 'Prevalence and prognostic impact of NPM1 mutations in 1485 adult patients with acute myeloid leukemia (AML)', *Blood*, 107(10), pp. 4011-4020.

Thomas, D., Scot, A.D., Barbey, R., Padula, M. and Boiteux, S. (1997) 'Inactivation of OGG1 increases the incidence of G . C-->T . A transversions in Saccharomyces cerevisiae: evidence for endogenous oxidative damage to DNA in eukaryotic cells', *Molecular and General Genetics*, 2(254), pp. 171-178.

Tighe, J.E. and Calabi, F. (1994) 'Alternative, out-of-frame runt/MTG8 transcripts are encoded by the derivative (8) chromosome in the t(8;21) of acute myeloid leukemia M2', *Blood*, 84(7), p. 2115.

Tyner, J.W., Erickson, H., Deininger, M.W.N., Willis, S.G., Eide, C.A., Levine, R.L., Heinrich, M.C., Gattermann, N., Gilliland, D.G., Druker, B.J. and Loriaux, M.M. (2009) 'High-throughput sequencing screen reveals novel, transforming RAS mutations in myeloid leukemia patients', *Blood*, 113(8), pp. 1749-1755.

van de Locht, L.T., Smetsers, T.F., Wittebol, S., Raymakers, R.A. and Mensink, E.J. (1994) 'Molecular diversity in AML1/ETO fusion transcripts in patients with t(8;21) positive acute myeloid leukaemia', *Leukemia*, 10, pp. 1780 - 1784.

van Dijk, E.L., Auger, H., Jaszczyszyn, Y. and Thermes, C. (2014) 'Ten years of next-generation sequencing technology', *Trends in Genetics*, 30(9), pp. 418-426.

Vardiman, J.W., Harris, N.L. and Brunning, R.D. (2002) *The World Health Organization (WHO) classification of the myeloid neoplasms*.

Vassiliou, G.S., Cooper, J.L., Rad, R., Li, J., Rice, S., Uren, A., Rad, L., Ellis, P., Andrews, R., Banerjee, R., Grove, C., Wang, W., Liu, P., Wright, P., Arends, M. and Bradley, A. (2011) 'Mutant nucleophosmin and cooperating pathways drive leukemia initiation and progression in mice', *Nature Genetics*, 43, p. 470.

Vassiliou, G.S. and Green, A.R. (2010) 'The Molecular Basis of Leukaemia and Lymphoma', in *Postgraduate Haematology*. Wiley-Blackwell, pp. 380-394.

Vousden, K.H. and Lu, X. (2002) 'Live or let die: the cell response to p53', *Nature Reviews Cancer*, 2, p. 594.

Wang, J., Hoshino, T., Redner, R.L., Kajigaya, S. and Liu, J.M. (1998) 'ETO, fusion partner in t(8;21) acute myeloid leukemia, represses transcription by interaction with the human N-CoR/mSin3/HDAC1 complex', *Proceedings of the National Academy of Sciences of the United States of America*, 95(18), pp. 10860-10865.

Wang, J., Li, Z., He, Y., Pan, F., Chen, S., Rhodes, S., Nguyen, L., Yuan, J., Jiang, L., Yang, X., Weeks, O., Liu, Z., Zhou, J., Ni, H., Cai, C.-L., Xu, M. and Yang, F.-C. (2014) 'Loss of leads to myelodysplastic syndrome–like disease in mice', *Blood*, 123(4), p. 541.

Wang, J.C. (2002) 'Cellular roles of DNA topoisomerases: a molecular perspective', *Nature Reviews Molecular Cell Biology*, 3, p. 430.

Wang, Y.-Y., Zhao, L.-J., Wu, C.-F., Liu, P., Shi, L., Liang, Y., Xiong, S.-M., Mi, J.-Q., Chen, Z., Ren, R. and Chen, S.-J. (2011) 'C-KIT mutation cooperates with full-length AML1-ETO to induce acute myeloid leukemia in mice', *Proceedings of the National Academy of Sciences*, 108(6), pp. 2450-2455.

Wang, Y.-Y., Zhou, G.-B., Yin, T., Chen, B., Shi, J.-Y., Liang, W.-X., Jin, X.-L., You, J.-H., Yang, G., Shen, Z.-X., Chen, J., Xiong, S.-M., Chen, G.-Q., Xu, F., Liu, Y.-W., Chen, Z. and Chen, S.-J. (2005) 'AML1-ETO and C-KIT mutation/overexpression in t(8;21) leukemia: Implication in stepwise leukemogenesis and response to Gleevec', *Proceedings of the National Academy of Sciences of the United States of America*, 102(4), pp. 1104-1109.

Wang, Y., Wu, N., Liu, D. and Jin, Y. (2017) 'Recurrent Fusion Genes in Leukemia: An Attractive Target for Diagnosis and Treatment', *Current Genomics*, 18(5), pp. 378-384.
Webersinke, G., Kranewitter, W., Deutschbauer, S., Zach, O., Hasenschwandtner, S., Wiesinger, K., Erdel, M., Marschon, R., Böhm, A. and Tschurtschenthaler, G. (2014) 'Switch of the mutation type of the NPM1 gene in acute myeloid leukemia (AML): relapse or secondary AML?', *Blood Cancer Journal*, 4, p. e221.

Weihrauch, M., Bader, M., Lehnert, G., Wittekind, C., Tannapfel, A. and Wrbitzky, R. (2002) 'Carcinogen-specific mutation pattern in the p53 tumour suppressor gene in UV radiation-induced basal cell carcinoma', *International Archives of Occupational and Environmental Health*, 75(4), pp. 272-276.

Weiss, J.M., Goode, E.L., Ladiges, W.C. and Ulrich, C.M. (2004) 'Polymorphic variation in hOGG1 and risk of cancer: A review of the functional and epidemiologic literature', *Molecular Carcinogenesis*, 42(3), pp. 127-141.

Weissmann, S., Alpermann, T., Grossmann, V., Kowarsch, A., Nadarajah, N., Eder, C., Dicker, F., Fasan, A., Haferlach, C., Haferlach, T., Kern, W., Schnittger, S. and Kohlmann, A. (2011) 'Landscape of TET2 mutations in acute myeloid leukemia', *Leukemia*, 26, p. 934.

Welch, John S., Ley, Timothy J., Link, Daniel C., Miller, Christopher A., Larson, David E., Koboldt, Daniel C., Wartman, Lukas D., Lamprecht, Tamara L., Liu, F., Xia, J., Kandoth, C., Fulton, Robert S., McLellan, Michael D., Dooling, David J., Wallis,

John W., Chen, K., Harris, Christopher C., Schmidt, Heather K., Kalicki-Veizer, Joelle M., Lu, C., Zhang, Q., Lin, L., O'Laughlin, Michelle D., McMichael, Joshua F., Delehaunty, Kim D., Fulton, Lucinda A., Magrini, Vincent J., McGrath, Sean D., Demeter, Ryan T., Vickery, Tammi L., Hundal, J., Cook, Lisa L., Swift, Gary W., Reed, Jerry P., Alldredge, Patricia A., Wylie, Todd N., Walker, Jason R., Watson, Mark A., Heath, Sharon E., Shannon, William D., Varghese, N., Nagarajan, R., Payton, Jacqueline E., Baty, Jack D., Kulkarni, S., Klco, Jeffery M., Tomasson, Michael H., Westervelt, P., Walter, Matthew J., Graubert, Timothy A., DiPersio, John F., Ding, L., Mardis, Elaine R. and Wilson, Richard K. (2012) 'The Origin and Evolution of Mutations in Acute Myeloid Leukemia', *Cell*, 150(2), pp. 264-278.

Whitmarsh, R.J., Saginario, C., Zhuo, Y., Hilgenfeld, E., Rappaport, E.F., Megonigal, M.D., Carroll, M., Liu, M., Osheroff, N., Cheung, N.-K.V., Slater, D.J., Ried, T., Knutsen, T., Blair, I.A. and Felix, C.A. (2003) 'Reciprocal DNA topoisomerase II cleavage events at 5′-TATTA-3′ sequences in MLL and AF-9 create homologous single-stranded overhangs that anneal to form der(11) and der(9) genomic breakpoint junctions in treatment-related AML without further processing', *Oncogene*, 22, p. 8448.

Wichmann, C., Quagliano-Lo Coco, I., Yildiz, O., Chen-Wichmann, L., Weber, H., Syzonenko, T., Doring, C., Brendel, C., Ponnusamy, K., Kinner, A., Brandts, C., Henschler, R. and Grez, M. (2015) 'Activating c-KIT mutations confer oncogenic cooperativity and rescue RUNX1/ETO-induced DNA damage and apoptosis in human primary CD34+ hematopoietic progenitors', *Leukemia*, 29(2), pp. 279-289.

Wiebauer, K., Neddermann, P., Hughes, M. and Jiricny, J. (1993) 'The repair of 5-methylcytosine deamination damage', in Jost, J.-P. and Saluz, H.-P. (eds.) *DNA Methylation: Molecular Biology and Biological Significance*. Basel: Birkhäuser Basel, pp. 510-522.

Wildonger, J. and Mann, R.S. (2005) 'The t(8;21) translocation converts AML1 into a constitutive transcriptional repressor', *Development*, 132(10), p. 2263.

Wilson, D.M. and Bohr, V.A. (2007) 'The mechanics of base excision repair, and its relationship to aging and disease', *DNA Repair*, 6(4), pp. 544-559.

Wiseman, H. and Halliwell, B. (1996) 'Damage to DNA by reactive oxygen and nitrogen species: role in inflammatory disease and progression to cancer', *Biochemical Journal*, 313(Pt 1), pp. 17-29.

Wong, I.H.N., Ng, M.H.L., Huang, D.P. and Lee, J.C.K. (2000) 'Aberrant promoter methylation in adult and childhood acute leukemias of nearly all morphologic subtypes: potential prognostic implications', *Blood*, 95(6), p. 1942.

Wong, I.H.N., Ng, M.H.L., Lee, J.C.K., Chung, Y.F. and Huang, D.P. (1998) 'Transcriptional silencing of the p16 gene in human myeloma-derived cell lines by hypermethylation', *British Journal of Haematology*, 1(103), pp. 168-175.

Xu, G., Kanezaki, R., Toki, T., Watanabe, S., Takahashi, Y., Terui, K., Kitabayashi, I. and Ito, E. (2006) 'Physical association of the patient-specific GATA1 mutants with RUNX1 in acute megakaryoblastic leukemia accompanying Down syndrome', *Leukemia*, 20, p. 1002.

Yan, M., Kanbe, E., Peterson, L.F., Boyapati, A., Miao, Y., Wang, Y., Chen, I.M., Chen, Z., Rowley, J.D., Willman, C.L. and Zhang, D.-E. (2006) 'A previously unidentified alternatively spliced isoform of t(8;21) transcript promotes leukemogenesis', *Nature Medicine*, 12, p. 945.

Yang, F., Kemp, C.J. and Henikoff, S. (2015) 'Anthracyclines induce double-strand DNA breaks at active gene promoters', *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 773, pp. 9-15.

Yuan, Y., Zhou, L., Miyamoto, T., Iwasaki, H., Harakawa, N., Hetherington, C.J., Burel, S.A., Lagasse, E., Weissman, I.L., Akashi, K. and Zhang, D.-E. (2001) 'AML1-ETO expression is directly involved in the development of acute myeloid leukemia in the presence of additional mutations', *Proceedings of the National Academy of Sciences of the United States of America*, 98(18), pp. 10398-10403.

Zahn, R.K., Müller, W.E.G., Forster, W., Maidhof, A. and Beyer, R. (1972) 'Action of 1-β-d-Arabinofuranosylcytosine on mammalian tumor cells—1.: Incorporation into DNA', *European Journal of Cancer (1965)*, 8(4), pp. 391-396.

Zhang, D.E., Hetherington, C.J., Meyers, S., Rhoades, K.L., Larson, C.J., Chen, H.M., Hiebert, S.W. and Tenen, D.G. (1996) 'CCAAT enhancer-binding protein (C/EBP) and AML1 (CBF alpha2) synergistically activate the macrophage colony-stimulating factor receptor promoter', *Molecular and Cellular Biology*, 16(3), pp. 1231-1240.

Zhang, J., Hug, B.A., Huang, E.Y., Chen, C.W., Gelmetti, V., Maccarana, M., Minucci, S., Pelicci, P.G. and Lazar, M.A. (2001) 'Oligomerization of ETO Is Obligatory for Corepressor Interaction', *Molecular and Cellular Biology*, 21(1), pp. 156-163.

Zhang, J., Kalkum, M., Yamamura, S., Chait, B.T. and Roeder, R.G. (2004) 'E Protein Silencing by the Leukemogenic AML1-ETO Fusion Protein', *Science*, 305(5688), p. 1286.

Zhang, Y., Strissel, P., Strick, R., Chen, J., Nucifora, G., Le Beau, M.M., Larson, R.A. and Rowley, J.D. (2002) 'Genomic DNA breakpoints in RUNX1-ETO cluster with topoisomerase II DNA cleavage and DNase I hypersensitive sites in t(8;21) leukemia', *Proceedings of the National Academy of Sciences*, 99(5), p. 3070.