ELUCIDATING THE GENOMIC AND TRANSCRIPTOMIC PROFILES CONTRIBUTING TO THE OVERALL SURVIVAL OF ACUTE MYELOID LEUKAEMIA PATIENTS WITH NORMAL KARYOTYPE IN MALAYSIA

A ANGELI A/P AMBAYYA @ AMPIAH

UNIVERSITI SAINS MALAYSIA

2024

ELUCIDATING THE GENOMIC AND TRANSCRIPTOMIC PROFILES CONTRIBUTING TO THE OVERALL SURVIVAL OF ACUTE MYELOID LEUKAEMIA PATIENTS WITH NORMAL KARYOTYPE IN MALAYSIA

by

A ANGELI A/P AMBAYYA @ AMPIAH

Thesis submitted in fulfilment of the requirements
for the degree of
Doctor of Philosophy

August 2024

ACKNOWLEDGEMENTS

First and foremost, all praise and glory to my Lord and Saviour, Jesus Christ, who strengthened me to complete this research. Only by His grace have I made it this far. Heartfelt gratitude to my main supervisor, Prof. Dr Rosline Hassan, for her constant support and motherly love and for being there from day one. I sincerely thank my co-supervisors, Assoc. Prof. Dr Sarina Sulong and Dr Rozaimi Razali. I am profoundly grateful for the bioinformatics mentoring by Dr Rozaimi. I am so blessed to have such an incredible team of supervisors. I appreciate the endless support and favour of Dr Jameela Sathar, Dr Mandy Yap, and Dr Veena Selvaratnam. Thanks to the Department of Haematology, Hospital Ampang and Universiti Sains Malaysia (USM), my fellow post-graduate students, especially Nor Hayati Ismail and all my subjects who willingly donated their blood for this project. Thanks to the Ministry of Health (MOH) Malaysia for granting me the PhD scholarship. My heartfelt gratitude goes to Ts. Mohd Zuhaimi Yacob, who had been a brother and a good listener. Special thanks to my JRM team and Rev. John Ganapathy for constantly enveloping me in prayers. My deepest gratitude goes to my family, especially my mum, for her prayers and love. She played an integral role in my accomplishments. This thesis was written in memory of my eldest brother, Edward James, who passed away in 2021. He was always proud of his little sister, and I am sure he rejoiced in Heaven with my accomplishments. Through it all, I finally made it to the finish line, indeed, by the grace of God. Words are not good enough to thank everyone who participated in this journey. Again, thank you. Every good and perfect gift is from above!

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF SYMBOLS

^{0}C	Celsius
Δ	Delta
g	Gram
g/dl	Gram per decilitre
mL	Millilitre
µL	Microlitre
mg	Milligram
ng	Nanogram
ng/µL	Nanogram per microlitre
%	Percent
U	Unit

LIST OF ABBREVIATIONS

ACMG	American College of Medical Genetics and Genomics
allo-SCT	Allogenic-SCT
AML-NK	Acute myeloid leukaemia-normal karyotype
ANC	Absolute neutrophil count
AMP	Association for Molecular Pathology
ASCO	American Society for Clinical Pathology
AUC	The area under the ROC curve
auto-SCT	Autologous-SCT
BP	Biological process
bp	base pair
CADD	Combined annotation-dependent depletion
DA 3+7	Daunorubicin and cytarabine (also known as Ara C) (3+7)
DTU	Differential transcript usage
EFS	Event-free survival
CAP	College of American Pathologists
cDNA	Complementary DNA
CKB Boost	Clinical knowledgebase by the Jackson's Laboratory
Clinvar	Clinical genome resource
Clinvarminer	Genome database
CNV	Copy number variation
COSMIC	Catalogue of somatic mutations in cancer
CR1	First complete remission was described as patients who have attained remission following the completion of chemotherapy cycles (induction and consolidation). RX is synonym to CR1 in this thesis.

CT	Cycle threshold value
dbSNP	The single nucleotide polymorphism database
DEG	Differentially expressed genes
DNA	Deoxyribonucleic acid
ds	Double-stranded
DX	Refers to patient sample collected at presentation
ELN	European Leukaemia Network
EveClass	Evolutionary model of variant effect score
FC	Fold change
FDR	False discovery rate
FFPM	Fusion fragments per million
F-GT	Genotype parameter
GO	Gene ontology
GOF	Gain of function
GOI	Gene of interest
GSEA	Gene set enrichment analysis
HA	Hospital Ampang
HG	Housekeeping genes
HIS	Hospital information system
HRM	High-resolution melting analysis
HUSM	Hospital Universiti Sains Malaysia
IL	Interleukin
InDels	Insertion-deletion
lfc	log fold change
IQR	Interquartile range
KM	Kaplan-Meier
KEGG	Kyoto Encyclopedia of Genes and Genomes

LDAS	Large double anchor support
lincRNA	Long intergenic non-coding RNA
lncRNA	Long noncoding RNA
LOF	Loss of function is predicted
MF	Molecular function
MRD	Minimal residual disease
mRNA	Messenger RNA
ncRNA	Non-coding RNA
NES	Normalised enrichment score
NGS	Next-generation sequencing
NMD	Nonsense-mediated mRNA decay
OR	Odds ratio
ORA	Over-representation analysis
PA	Pathway analysis
padj	Adjusted p-value
PCR	Polymerase chain reaction
PE	Paired-end
Plt	Platelet
PON	Panel of normal
MREC	Medical Research Ethics Committee
Real-Time qRT-PCR	Real-time quantitative reverse transcription PCR
RIN	RNA integrity number
RNA	Ribonucleic acid
ROC	Receiver operating characteristic
RT-PCR	Reverse transcription PCR
SCT	Stem cell transplant

SD	Standard deviation
SEER	Surveillance, Epidemiology, and End Results
SIFT	Sorting intolerant from tolerant
SNP	Single nucleotide polymorphism
SNV	Single nucleotide variant
SPSS	Statistical Package for the Social Sciences
TSG	Tumour suppressor gene
TWBC	Total white blood cell
VAF	Variant allele frequency
VEP	Variant effect predictor
WES	Whole exome sequencing
WHO	World Health Organization
WGS	Whole genome sequencing
wt	Wildtype

LIST OF APPENDICES

PROFIL GENOMIK DAN TRANSKRIPTOMIK YANG MENYUMBANG KEPADA SURVIVAL PESAKIT MYELOID LEUKEMIA AKUT DI MALAYSIA

ABSTRAK

Leukemia mieloid akut-kariotip normal (AML-NK) terdiri daripada hampir separuh daripada subjenis AML dan diketahui mempamerkan heterogeniti klinikal dari segi kelangsungan hidup keseluruhan (OS). Namun sehingga kini, landskap genomik AML-NK lengkap yang menyumbang kepada OS masih belum terungkai. Oleh itu, kajian ini merungkai profil genomik dan transkritom pesakit AML-NK yang berkait dengan OS. Sebanyak 51 sampel AML-NK diagnosis baru (DX) ,14 pasang sampel yang remisi lengkap pertama (CR1) dan 14 orang normal yang sihat telah diuji untuk penjujukan transkritom mendalam menggunakan NovaSeq 6000 (200M bacaan setiap sampel, bacaan akhir berpasangan). Lapan pasang DNA DX dan CR1 telah dijalankan penjujukan DNA disasar berpandukan panel Archer HGC VariantPlex Myeloid menggunakan NovaSeq 6000 (kedalaman 3M untuk sampel DX dan 30M untuk sampel CR1 secara bacaan akhir berpasangan). Penjujukan transkritom menghasilkan penemuan profil gen yang diekspresikan secara berbeza (DEG) dan laluan fungsional terperinci dalam beberapa analisis subkumpulan yang merangkumi perbandingan pesakit AML-NK dengan kumpulan normal yang sihat (5126 DEGs), pasangan DX dan CR1 (5621 DEGs), genotip *FLT3*/*NPM1* (750 DEGs), dan kelangsungan hidup keseluruhan (OS) yang kurang serta yang melebihi lima tahun (211 DEGs). Analisis gen gabungan mendedahkan 27 jenis gen gabungan (n=68), termasuk *LATS2*-*SAP18* dan *HOXA3*-*HOXA9* yang novel, yang mempamerkan kaitan prognostik terhadap

pesakit dengan OS kurang lima tahun. Sebanyak 74,604 varian somatik yang merbahaya merangkumi 88.3% varian nukleotida tunggal (SNV) dan 11.7% pemadaman sisipan anjakan bingkai (InDels) ditemui dalam 51 pesakit AML-NK. Terdapat terapi bersasar untuk lima SNVs manakala InDels yang dikesan dalam gen *NPM1, DNMT3A* dan *FLT3* ketara secara prognostik. Penjujukan DNA disasar mendedahkan 208 penemuan (122 dalam sampel DX dan 86 dalam sampel CR1), yang membawa kepada pengenalpastian biomarker yang sesuai untuk pemantauan baki minima sel leukaemia. Puncak penemuan DEG adalah penjanaan model pemarkahan prognostik berdasarkan penemuan OS untuk perbandingan kurang serta yang lima tahun ke atas. Enam gen yang disregulasi secara menaik dengan ketara dalam (*FHL1, SOCS2, IL17RC, STAT4, INHBA* dan *TNFSF8*) terlibat dalam laluan isyarat JAK-STAT serta interaksi reseptor sitokin-sitokin yang dimasukkan dalam model pemarkahan prognostik mendedahkan bahawa skor gen adalah penanda prognostik bebas dalam pesakit AML-NK dalam kohort ini. Seterusnya, penemuan ini digambarkan secara *oncoprint* untuk mencerminkan bagaimana penemuan genomik dalam kajian ini menambah baik stratifikasi risiko pesakit untuk ramalan hasil dan terapi sasaran yang berpotensi. Oleh itu, kajian pelbagai aspek ini telah memberikan pandangan baharu tentang profil genomik dan penerangan tentang hasil rawatan klinikal yang heterogen dalam kalangan pesakit AML-NK.

ELUCIDATING THE GENOMIC AND TRANSCRIPTOMIC PROFILES CONTRIBUTING TO THE OVERALL SURVIVAL OF ACUTE MYELOID LEUKAEMIA PATIENTS WITH NORMAL KARYOTYPE IN MALAYSIA

ABSTRACT

Acute myeloid leukaemia-normal karyotype (AML-NK) comprises almost half of the AML subtypes and exhibits heterogeneity in overall survival (OS). Hitherto, the genomic and transcriptomic profiles of AML-NK that contribute to OS remain veiled. Therefore, this study aimed to elucidate the genomic and transcriptomic profiles of AML—NK patients predisposed to their overall survival. Deep transcriptome sequencing using NovaSeq 6000 (200M reads per sample, paired-end) was performed on 51 AML-NK patients at diagnosis (DX), 14 paired first remission (CR1) samples and 12 healthy controls. Targeted DNA sequencing (Archer HGC VariantPlex Myeloid panel) was performed using NovaSeq 6000 on eight paired DX and CR1 with sequencing depths of 3M and 30M, paired-end, respectively. The transcriptome sequencing yielded discoveries of differentially expressed genes (DEGs) profiles and functionally enriched pathways in several subgroup analyses as follows: AML-NK patients with the healthy normal groups (5126 DEGs), paired DX and CR1 (5621 DEGs), *FLT3/NPM1* genotypes (750 DEGs), and OS of below and above five years (211 DEGs). There were 27 types of fusion genes (n=68) discovered in this study, of which several novel recurrent fusion genes, including *LATS2-SAP18* and *HOXA3-HOXA9*, exhibited prognostic relevance in patients with OS below five years. A total of 74,604 deleterious variants (88.3% single nucleotide variants (SNVs) and 11.7% insertion-deletions (InDels)) were identified in the 51 AML-NK patients. Five SNVs

were candidates for targeted therapies, whereas prognostically significant frameshift InDels were detected in the *NPM1, DNMT3A* and *FLT3* genes. The targeted DNA sequencing included eight AML-NK samples and revealed 208 findings (122 in the DX samples and 86 in the CR1 samples), leading to the identification of suitable biomarkers for minimal residual disease monitoring. The highlight of this study was a prognostic scoring model developed based on the findings of the OS below and above five-year comparison. Six significantly upregulated genes in the (*FHL1, SOCS2, IL17RC, STAT4, INHBA* and *TNFSF8*) in the JAK-STAT signalling pathway and cytokine-cytokine receptor interaction were included in the prognostic scoring model that revealed that the gene scores were an independent prognostic marker in the AML-NK patients in this cohort. Ultimately, the findings were depicted in an oncoprint that reflected how the genomic discoveries in this study improvised the patient's risk stratification for outcome predictions and potential targeted therapies. Hence, this multifaceted study has provided new insights into the genomic and transcriptomic profiles of AML-NK patients and shed light on their heterogeneous clinical outcomes.

CHAPTER 1

INTRODUCTION

1.1 Study background

Cancer is a myriad of disorders that cause rapid abnormal proliferation of cells encompassing malignant tumours and neoplasms, as defined by the World Health Organisation (WHO). Abnormal cancer cells can invade adjacent parts of the body and metastasise to other organs, which ultimately causes death. There are five main categories of cancers according to the Surveillance, Epidemiology, and End Results (SEER) Program (United States of America) and Cancer Research (United Kingdom) as follows: i) carcinoma, ii) sarcoma, iii) leukaemia, iv) myeloma and lymphoma, and v) brain and spinal cord cancers (SEER, 2023a).

Carcinoma is defined as cancer that begins in the skin or the tissue line or covers internal organs, while sarcoma includes cancer that initiates at the connective or supportive tissues. Leukaemia involves blood cells that are formed in the bone marrow, whereas lymphoma and myeloma involve cancers that affect the immune system. Brain and spinal cord cancers involve central nervous system cancers (Cancer Research UK, 2023; SEER, 2023).

According to the WHO Cancer statistics, cancer is the leading cause of death worldwide, with an estimated 19.3 million new cases and almost 10 million deaths in 2020 (Sung et al., 2021). The WHO expects the annual cancer incidence to increase to 20 million annually by 2025 (Sung et al., 2021; Ferlay et al., 2023). According to the Malaysian National Cancer Registry (MNCR) 2012-2016 report, there were 115,238 (44.7% males and 55.3% females) new cancer cases registered between 2012-2016 in this country. The age-standardised incidence rate (ASR) per 100,000 population for

males was 86.1, whereas for females was 101.6. According to the MCNR 2012-2016 report, breast cancer was the most common among females, while colorectal cancer was the most common in males (National Cancer Registry, 2019). Leukaemia was the sixth commonest form of cancer in Malaysia among all residents between the year 2012-2016, as illustrated in Figure 1.1.

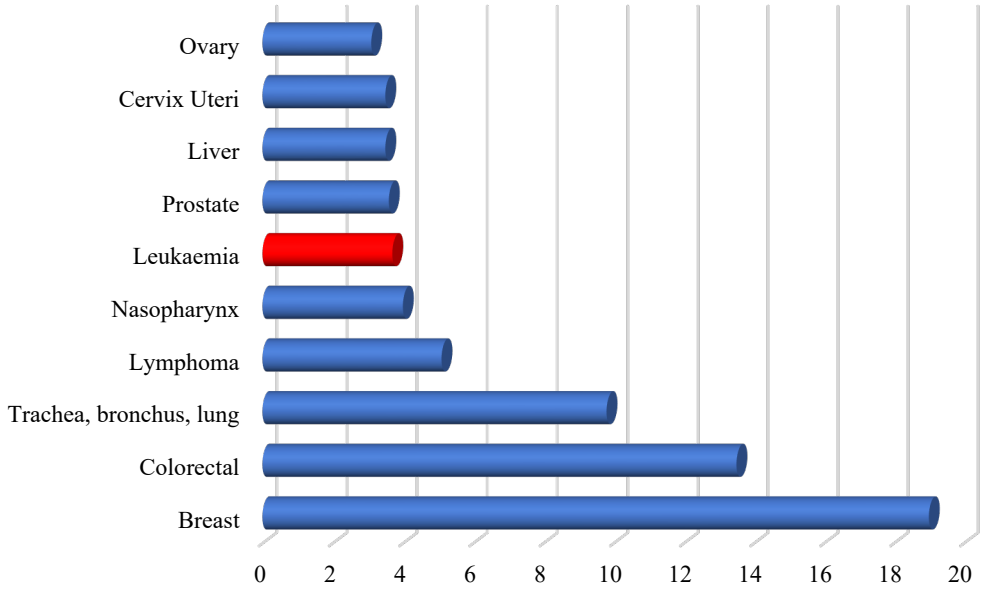

Figure 1.1 Top ten cancers in Malaysia based on MCNR report, 2012-2016 (National Cancer Registry, 2019).

A total number of 4,273 cases of leukaemia were registered between 2012 and 2016, of which about 55.2% (n= 2,359/4273) were myeloid leukaemia. Approximately 54.2% (n= 1,279) of the myeloid leukaemia patients were males, and 45.8% (n=1,080) were females. The incidence of myeloid leukaemia per 100,000 population increased with age, especially after 40 years of age in both genders, as shown in Figure 1.2 and Figure 1.3 (National Cancer Registry, 2019).

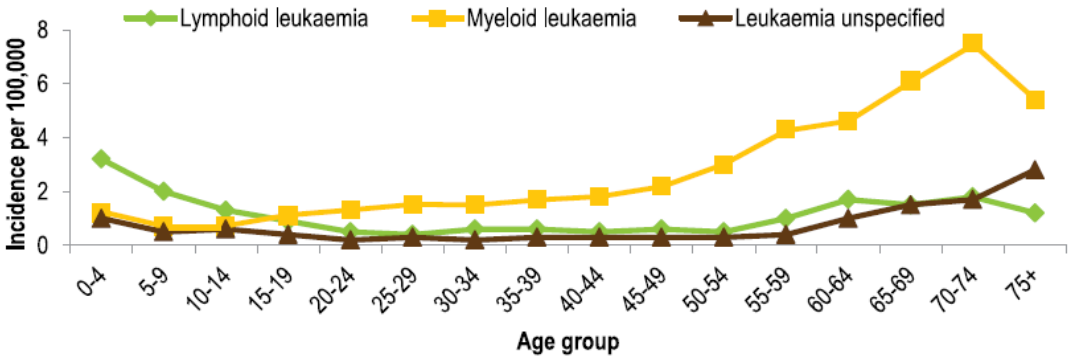

Figure 1.2 Age-specific incidence rates of leukaemia in males between 2012 and 2016.
The figure was adapted from the National Cancer Registry report(National Cancer Registry, 2019).

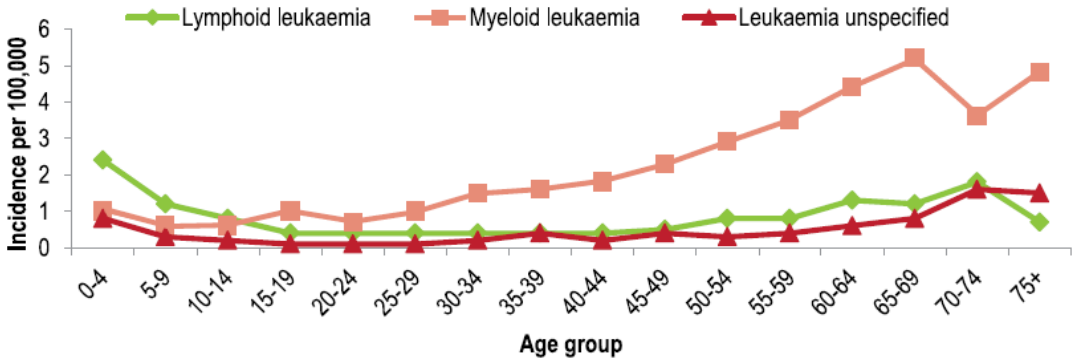

Figure 1.3 Age-specific incidence rates of leukaemia in females between 2012 and 2016.
The figure was adapted from the National Cancer Registry Report(National Cancer Registry, 2019).

Acute myeloid leukaemia (AML) is the most common in the adult population and is characterised by clonal expansion of immature myeloid cells in the blood and bone marrow, affecting erythropoiesis and increasing the risk of bone marrow failure (Bain & Béné, 2019; Medeiros et al., 2019; Naymagon et al., 2019). The recent 5th WHO classification of AML has enhanced the grouping into two main categories. The first category includes AML with defining genetic abnormalities encompassing fusion genes and somatic mutations. In the second category, AML, defined by differentiation,

considers the maturation stages of the leukaemic cells for classification (Khoury *et al.*, 2022).

Figure 1.4 depicts the timelines in the AML development that began with the French-American-British (FAB) classification of AML in 1976 based on the cell morphology to genomic-based classifications over the recent years (Bennett et al., 1976). In Malaysia, the current practice for AML diagnosis relies on incorporating morphologic, flow cytometry immunophenotyping and a few genetic markers, including the mutational status of *FLT3* and *NPM1* genes as routine diagnostic workups.

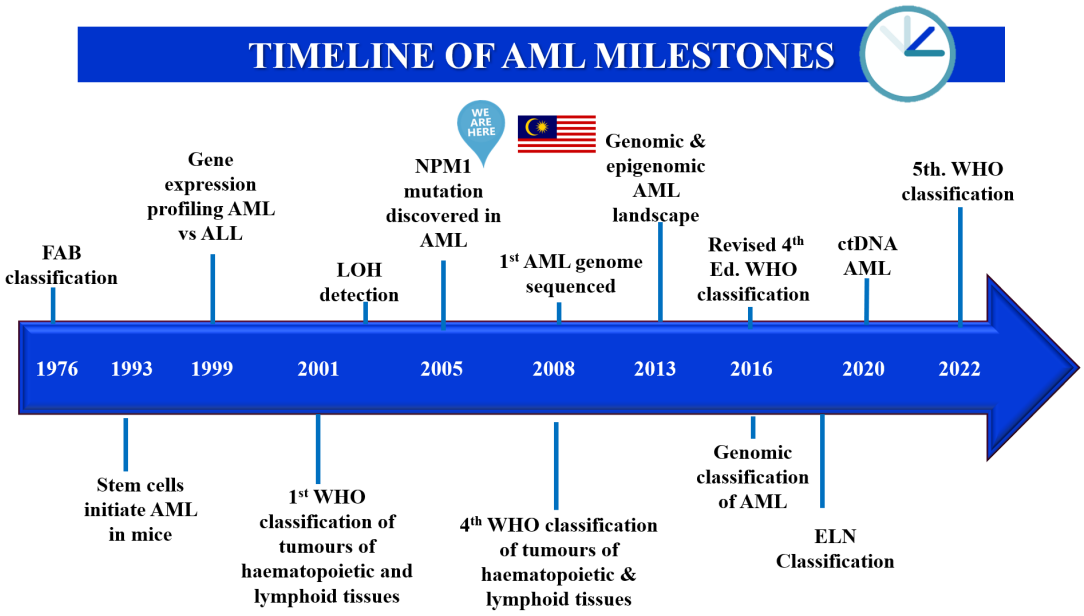

Figure 1.4 Timeline of AML milestones.

The earliest classification of AML was based on the FAB classification. Over the last decades, advancements in sequencing-based technologies have enabled genomic-based classification and prognostication of AML. In 2005, NPM1 mutation was discovered in AML (Falini et al., 2005)., and some laboratories in Malaysia

performed this mutation analysis as part of their routine diagnostic workup for AML patients.

Several studies have revealed that mutations in *FLT3*, *CEBPA* and *NPM1* genes were useful in predicting the outcome in AML-NK patients (van Waalwijk et al., 2003; Dickson et al., 2016; Konstandin et al., 2018; Abbas et al., 2019; Carbonell et al., 2022). However, the incidences of *FLT3*- internal tandem duplication (ITD) and *NPM1* mutations in Malaysian AML patients were lower (Abdullah et al., 2011; Yunus et al., 2015; Roszymah et al., 2016).

Over the last three decades, cytogenetic analysis has remained an indispensable tool in AML as recurrent chromosomal aberrations have been utilised to diagnose and prognosticate this heterogeneous disorder (Mrózek et al., 2004; Rowley, 2008). However, nearly half of the AML patients have a normal karyotype (AML-NK), lacking structural abnormalities. AML-NK is diagnosed based on conventional cytogenetics analysis of at least 20 normal bone marrow metaphases (Mawad & Estey, 2012a; Döhner, Andrew H Wei, et al., 2022). In a study among Malaysian AML patients, 49% (n=294/601) presented with a normal karyotype (Ambayya et al., 2021), consistent with the findings at different geographic locations worldwide (Gmidène et al., 2012; Amare et al., 2016; Ait Boujmia et al., 2021).

Although AML-NK patients are categorised as intermediate prognoses, the outcomes of the patients are rather heterogeneous (Döhner et al., 2017; Khoury et al., 2022; Bouligny et al., 2023). Studies have reported that the 5-year survival rates of AML-NK ranged between 24-42% (Marcucci et al., 2005; Nimer, 2008; Stölzel et al., 2016; Samra et al., 2020). Intriguingly, the complexities of AML-NK are more than what appears on the surface, involving interplays between the karyotypes, pathobiology, and clinical expressions.

Genomic assays enable the characterisation of aberrations at genomic, transcriptomic, and epigenetics levels, especially when studying the defects of AML patients. Genomic methods provide insight into targetable genomic alteration by precision medicine and reveal the complexities of gene regulations in AML patients (Berger & Mardis, 2018; Hayashi et al., 2020). Gene regulations are controlled by interactions between transcription factors and their target genes and extension interactions in different pathways in humans (MacNeil & Walhout, 2011). In AML patients, the deregulated genes disrupt the regulatory networks involving the molecular interplay between the DNA, RNA and proteins that govern the AML cells and maintain the leukaemic state via these interactions(Jothi et al., 2009; Wrighton, 2019; Kanellou, Georgakopoulos-Soares & Zaravinos, 2023).

Despite advancements in various technologies, including high-density comparative genomic hybridisation (single nucleotide polymorphism array) and next-generation sequencing (NGS), studies have revealed that a substantial number of AML patients carry zero or very few mutations in the AML-associated driver genes (Patel *et al.*, 2011; Shen *et al.*, 2011; The Cancer Genome Atlas Research Network, 2013). Moreover, the average number of mutations per genome in AML patients is lower than in other cancer types (Welch *et al.*, 2012; The Cancer Genome Atlas Research Network, 2013; Garg *et al.*, 2015; Mat Yusoff *et al.*, 2021)

The AML-NK genomic and transcriptome landscape is yet to be explored in Malaysia. It is paramount to elucidate the cryptic aberrations in AML-NK patients, which will yield crucial insights into the disease pathobiology. This study aims to uncover the underlying genomic profiles and regulatory networks contributing to the clinical outcome of AML-NK patients in Malaysia that are pragmatic in the diagnosis and prognosis of this clinically heterogeneous disease.

1.2 Justification of the study

It is crucial to look beyond the existing WHO subgroups of genomic categories of AML as improvisation can be made in the classification of AML-NK patients for diagnosis and outcome prediction; this is particularly important in the AML-NK cases where provisional new disease entities could be discovered in this study, which may influence treatment decisions for better clinical outcomes.

Based on the increasing use of "omics" technologies in research and clinical services, the current list of AML-relevant genes will likely be expanded, and routine testing for single gene mutations will likely be replaced by microarray and NGS-based panel diagnostics. Although some organisations are carrying out studies, these findings are not incorporated into an accessible knowledge database in Malaysia, which will be invaluable in advancing individualised treatment approaches.

Moreover, there is no publication on DNA and transcriptome profiling of AML-NK patients in Malaysia, which could discover new disease entities and biologic clusters for risk classification and outcome predictions. Although some studies were conducted on a smaller scale of AML patients, no association with patient clinical findings and survival analysis has been carried out (Osman *et al.*, 2020; Mat Yusoff *et al.*, 2021).

Therefore, this study adds an essential facet by integrating targeted DNA and deep transcriptome sequencing to uncover the AML-NK genomic profiles and exemplify regulatory networks that predict the patient's outcome. Ultimately, elucidating genomic biomarkers and regulatory networks could shed some light on the pathogenesis of AML-NK and the patient's clinical outcome, which could pave the way for translational research on developing targeted therapies and precision medicine.

1.3 Research objectives

1.3.1 General objective

To uncover the underlying genomic profiles and regulatory networks contributing to the clinical outcome of acute myeloid leukaemia-normal karyotype (AML-NK) patients in Malaysia.

1.3.2 Specific objectives

The specific objectives of this study are:

1. To discover the genomic profiles of AML-NK patients using targeted DNA and transcriptome sequencing.
2. To explore the differentially expressed gene (DEG) profiles and to interrogate regulatory networks in:
    i. AML-NK patient samples at presentation (DX) (n=51) versus the healthy control group (n=12).
    ii. AML-NK DX (n=14) sample and their paired first complete remission (CR1) sample (n=14).
    iii. AML-NK patient subgroups based on the genotypes of *FLT3* and *NPM1* mutation groups.
    iv. AML-NK is based on their survival status (overall survival (OS) below and above 5 years).
3. To determine the fusion genes discovered in the AML-NK using two pipelines and their relevance in patient survival.
4. To explore somatic variants, including single nucleotide variants (SNVs) and insertion and deletions (InDels) and their clinical relevance in AML-NK patients.

5. To identify potential biomarkers and develop a prognostic scoring model to predict the AML-NK patient's survival.
6. To integrate the targeted DNA and transcriptome sequencing findings (DEG, fusion genes, and somatic variants) and their association with AML-NK patients' survival.

CHAPTER 2

LITERATURE REVIEW

2.1 Leukaemia overview

Leukaemia encompasses a broad spectrum of complex blood-related malignancies that begin in the bone marrow. Leukaemia is characterised by the abnormal proliferation of leukemic cells that accumulate in the bone marrow, blood and/or lymphatic tissues, disrupting blood cells' normal function and production. Leukaemia is classified based on its progression rate, which includes acute and chronic leukaemia.

Acute leukaemia develops rapidly and presents with more severe symptoms, and the leukemic cells are immature progenitor cells. Chronic leukaemia progresses slowly with milder symptoms, and the neoplastic cells are more mature than acute leukaemia (S.H. Campo, 2008; Chennamadhavuni *et al.*, 2024). Leukaemia is categorised by the malignant cell lineage being either myeloid or lymphoid. However, some leukemia cases have no lineage-specific phenotypes or admixture of different lineages of leukemic cells (Lernoux *et al.*, 2020; Chennamadhavuni *et al.*, 2024).

Among the features of leukaemia include fatigue, weakness, malaise, abnormal bleeding, excessive bruising, abnormal weight loss, bone and/or joint pain, infection and fever, abdominal discomfort and usually present with enlarged spleen, lymph nodes and liver (Iqbal, 2012) The current state-of-the-art leukaemia diagnosis integrates morphology, cytochemistry, immunophenotyping, cytogenetics and molecular genetics to allow comprehensive characterisation of patients for optimal diagnosis and management (Haferlach & Schmidts, 2020). Despite tremendous

development in diagnosis technologies, the heterogeneity of this disease is yet to be characterised thoroughly.

2.2 Acute myeloid leukaemia

Acute myeloid leukaemia is the most common leukaemia among adults, with increasing incidence with age. Higher incidences of AML in the White population (4.1 per 100,000 people) were reported compared to the Asian population (3.2 per 100,000 people) (Howlader et al., 2009; SEER, 2023). In Malaysia, the incidence of AML displayed an increase in incidence rates with age in both males and females. According to the Malaysia National Cancer Registry Report, the incidence of myeloid leukaemia in those above 40 to 60 years ranged between 1.8 and 7.5 (National Cancer Registry, 2019). In the largest and most comprehensive study on AML genetic profiles (n=854 patients, 443 male and 411 female), the median age at presentation was 45, ranging between 12-93 years (Ambayya *et al.*, 2021).

2.2.1 Classification of AML

2.2.1(a) French American British (FAB) classification system

The landmark categorisation of AML developed in the 1970s by a collaboration among haematologists is the French-American-British (FAB) classification system, revised in 1985. The FAB classification was considered the landmark classification system as this is the first systematic classification of AML based on the leukaemia cell differentiation stages. The FAB classification relies on the morphology and cytochemistry of the leukemic cells in the bone marrow aspirates and peripheral blood samples. Based on the level of leukemic blast differentiation, AML patients were designated into eight subtypes in the FAB classification system, as illustrated in Table 2.1 (Bennett *et al.*, 1976, 1985).

Table 2.1 French American British (FAB) classification of AML.

Subtype	Description
M0	• Minimal differentiation – predominantly stem cells or unidentified cell type.
M1	• Myeloblastic leukaemia without maturation – predominantly immature white blood cells.
M2	• Myeloblastic leukaemia with maturation – partial differentiation.
M3	• Hypergranular promyelocytic leukaemia – predominantly promyelocytes.
M4	• Myelomonocytic leukaemia – a combination of myeloblasts and monoblasts, with each component constituting>20% of the blasts in the bone marrow. • M4 – acute myelomonocytic leukaemia). • M4Eo – acute myelomonocytic leukaemia with eosinophilia).
M5	• Monocytic leukaemia – predominantly monocytes. • M5a – acute monocytic leukaemia without differentiation (monoblastic). • M5b – acute monocytic leukaemia with differentiation (promonocytic).
M6	• Erythroleukaemia Di Guglielmo's disease – predominantly immature red and white blood cells.
M7	• Megakaryoblastic leukaemia.

The table was adapted from Bennett *et al*. (1967, 1985) (Bennett et al., 1976; 1985).

The FAB system facilitated a standardised framework for AML diagnosis and reporting and enabled communication and comparison of findings across different clinical settings. Nevertheless, with the advancement in other laboratory techniques, including cytochemistry, flow cytometry, immunophenotyping, molecular genetics and genomics, there were restrictions in applying the FAB classification system as it relied heavily on morphological classification. Hence, with limited cytogenetic and molecular features, the heterogeneous genetic and biological complexities underpinning the AML subtypes were inadequately captured.

2.2.1(b) World Health Organisation (WHO) classification system

The World Health Organisation (WHO) introduced another classification in 2001 which incorporated morphological, cytochemical, immunophenotypic, and cytogenetics findings which revolutionised a paradigm shift as emphasis was given to genetics findings in the diagnosis of leukaemia(Vardiman et al., 2002). The WHO 2001 classification was revised in 2008 as genetic mutations associated with AML patients' prognosis, particularly those who presented with normal karyotypes, were

discovered (Swerdlow *et al.*, 2008). As an example, AML with mutated *NPM1* or *CEBPA* has been added as provisional entities in the revised WHO 2008 classification as frequently seen in normal karyotype AML patients (Swerdlow *et al.*, 2008).

Advances in sequencing technologies in the last two decades have led to discoveries of biomarkers that improved the diagnosis and prognosis of AML patients. Hence, the WHO classification was revised in 2016 by incorporating clinically relevant genetic biomarkers (Arber *et al.*, 2016). The most recent classification of AML was established in 2022, where the classification of AML was reenvisioned to highlight breakthrough findings that improved the understanding and management of this disease (Khoury *et al.*, 2022).

Firstly, the AML was regrouped into two main categories: AML with defining genetic abnormalities, and AML characterised by differentiation. The AML with defining genetic abnormalities as follows (Khoury et al., 2022):

- Acute promyelocytic leukaemia with *PML-RARA* fusion
- AML with *RUNX1-RUNX1T1* fusion
- AML with *CBFB-MYH11* fusion
- AML with *DEK-NUP214* fusion
- AML with *RBM15-MRTFA* fusion
- AML with *BCR-ABL1* fusion
- AML with *KMT2A* rearrangement
- AML with *MECOM* rearrangement
- AML with *NUP98* rearrangement
- AML with *CEBPA* mutation
- AML, myelodysplasia-related
- AML with other defined genetic alterations

Changes were made to what was formerly known as AML, with myelodysplasia-related changes to AML-myelodysplasia-related changes (AML-MR). The crucial changes in this category include the removal of sole morphology-based diagnosis, including defining cytogenetic findings and introducing a mutation-based definition consisting of eight genes listed below. The diagnosis of AML-MR requires the presence of one or more cytogenetic abnormalities and/or a history of myelodysplastic syndrome (MDS) or MDS/myeloproliferative neoplasms (MPN) (Khoury *et al.*, 2022). AML-MR that are defined by cytogenetic abnormalities include the following:

- Complex karyotype (≥3 abnormalities)
- 5q deletion or loss of 5q due to unbalanced translocation
- Monosomy 7, 7q deletion, or loss of 7q due to unbalanced translocation
- 11q deletion
- 12p deletion or loss of 12p due to unbalanced translocation
- Monosomy 13 or 13q deletion
- 17p deletion or loss of 17p due to unbalanced translocation
- Isochromosome 17q
- idic(X)(q13)

AML-MR is also classified by the presence of defining somatic mutations in the *ASXL1*, *BCOR*, *EZH2*, *SF3B1*, *SRSF2*, *STAG2*, *U2AF1* and *ZRSR2* genes.

AML that lacks defining genetic abnormalities is categorised as AML defined by differentiation as listed in Table 2.2. AML classification based on leukemic cell differentiation provides a long-standing practical approach for prognostications and therapeutic implications (Khoury *et al.*, 2022).

Table 2.2 AML defined by differentiation category.

Type	Diagnostic criteria*
AML with minimal differentiation	• Blasts are negative (<3%) for MPO and SBB by cytochemistry • Expression of two or more myeloid-associated antigens, such as CD13, CD33, and CD117
AML without maturation	• ≥3% blasts positive for MPO (by immunophenotyping or cytochemistry) or SBB and negative for NSE by cytochemistry • Maturing cells of the granulocytic lineage constitute <10% of the nucleated bone marrow cells • Expression of two or more myeloid-associated antigens, such as MPO, CD13, CD33, and CD117
AML with maturation	• ≥3% blasts positive for MPO (by immunophenotyping or cytochemistry) or SBB by cytochemistry • Maturing cells of the granulocytic lineage constitute ≥10% of the nucleated bone marrow cells • Monocyte lineage cells constitute < 20% of bone marrow cells • Expression of two or more myeloid-associated antigens, such as MPO, CD13, CD33, and CD117
Acute basophilic leukaemia	• Blasts and immature/mature basophils with metachromasia on toluidine blue staining • Blasts are negative for cytochemical MPO, SBB, and NSE. • No expression of strong CD117 equivalent (to exclude mast cell leukaemia)
Acute myelomonocytic leukaemia	• ≥20% monocytes and their precursors • ≥20% maturing granulocytic cells • ≥3% of blasts positive for MPO (by immunophenotyping or cytochemistry)
Acute monocytic leukaemia	• ≥80% monocytes and/or their precursors (monoblasts and/or promonocytes) • <20% maturing granulocytic cells • Blasts and promonocytes expressing at least two monocytic markers including CD11c, CD14, CD36 and CD64, or NSE positivity on cytochemistry
Acute erythroid leukaemia	• ≥30% immature erythroid cells (proerythroblasts) • Bone marrow with erythroid predominance, usually ≥80% of cellularity
Acute megakaryoblastic leukaemia	• Blasts express at least one or more of the platelet glycoproteins: CD41 (glycoprotein llb), CD61 (glycoprotein IIIa), or CD42b (glycoprotein lb)b

BM: bone marrow, MPO: myeloperoxidase, NSE: nonspecific esterase, PB: peripheral blood, SBB: Sudan Black B.

*Shared diagnostic criteria include:

- ≥20% blasts in bone marrow and/or blood (except for acute erythroid leukaemia).
- Criteria for AML types with defined genetic alterations are not met.
- Criteria for mixed-phenotype acute leukaemia are unmet (relevant for AML with minimal differentiation).
- Not fulfilling diagnostic criteria for myeloid neoplasm post-cytotoxic therapy.
- No prior history of myeloproliferative neoplasm.

The table was adapted from Khoury *et al*. (2022) (Khoury *et al.*, 2022).

Studies have revealed diversified genetic profiles of acute myeloid leukaemia in different geographical locations. The prevalence of cytogenetic abnormalities and genetic aberrations varied in different studies across the globe, suggesting gender, age

groups, ethnicity, and environmental factors are interrelated factors predisposing to the types of acute myeloid leukaemia reported (Moorman *et al.*, 2002; Marcucci, Mrózek & Bloomfield, 2005b; Schlenk, Döhner, Krauter, Fröhling, Corbacioglu, Bullinger, Habdank, Späth, Morgan, Benner, Schlegelberger, Heil, Ganser, Döhner & Group, 2008; Mawad & Estey, 2012a; P S Kadam Amare *et al.*, 2016; kaltoum Ait boujmia *et al.*, 2021).

2.2.1(c) ELN-based risk stratifications

The European LeukaemiaNet (ELN) international expert panel published the first recommendation for diagnosing and managing AML in 2010, widely adopted in clinical practice, trials, and regulatory agencies (Döhner *et al.*, 2010). As considerable progress has been made in the elucidation of AML pathogenesis, development of diagnostic tests, and introduction of novel therapies, in parallel with the WHO classification of myeloid neoplasm and acute leukaemia, revisions of ELN guidelines for AML have been made in 2017 and 2022 (Döhner et al., 2017; 2022). The management of patients diagnosed with *PML- RARA* t(15;17 translocation) fusion is described in the ELN special report on acute promyelocytic leukemia (Sanz et al., 2019).

2.2.1(d) ELN 2022 genetic risk classification at diagnosis

The ELN 2022 has incorporated new data that emerged from data of intensively treated AML patients that prompted the need to adjust the risk classifications. One of the most significant changes to the ELN 2022 genetic risk classification is that the *FLT3*-ITD allelic ratio is no longer considered. Allelic ratios refer to the number of mutant alleles compared to the number of wildtype alleles in AML. Next, AML with

FLT3-ITD (without adverse-risk genetic aberrations) is now considered an intermediate risk irrespective of allelic ratio or concurrent *NPM1* mutations, as there were issues in standardising the *FLT3*-ITD allelic measurement assays and improved outcomes of *FLT3*-ITD AML patients with the incorporation of Midostaurin-based therapy (Döhner *et al.*, 2017).

Table 2.3 outlines the risk categories and genetic abnormalities based on ELN 2022. The ELN 2022 updated the risk stratification schema based on the 5th edition of the WHO classification of haematolymphoid tumours: myeloid and histiocytic/dendritic neoplasms depicted in the Sankey plot (Döhner, Andrew H. Wei, *et al.*, 2022) (Figure 2.1).

Table 2.3 Risk categories and genetic abnormalities as defined by ELN 2022.

Risk category	Genetic abnormality
Favourable	• t(8;21)(q22;q22.1)/*RUNX1-RUNX1T1*†,‡ • inv(16)(p13.1q22) or t(16;16)(p13.1;q22)/*CBFB-MYH11*†,‡ • Mutated *NPM1*†, §without *FLT3*-ITD • bZIP in-frame mutated CEBPA‖
Intermediate	• Mutated *NPM1*†,§with *FLT3*-ITD • Wild-type *NPM1* with *FLT3*-ITD (without adverse-risk genetic lesions) • t(9;11)(p21.3;q23.3)/*MLLT3-KMT2A*† • Cytogenetic and/or molecular abnormalities not classified as favourable or adverse.
Adverse	• t(6;9)(p23.3;q34.1)/*DEK-NUP214* • t(v;11q23.3)/KMT2A-rearranged# • t(9;22)(q34.1;q11.2)/*BCR-ABL1* • t(8;16)(p11.2;p13.3)/*KAT6A-CREBBP* • inv(3)(q21.3q26.2) or t(3;3)(q21.3;q26.2)/*GATA2*, *MECOM(EVI1)* • t(3q26.2;v)/*MECOM*(*EVI1*)-rearranged • −5 or del(5q); −7; −17/abn(17p) • Complex karyotype, **monosomal karyotype†† • Mutated *ASXL1, BCOR, EZH2, RUNX1, SF3B1, SRSF2, STAG2, U2AF1,* and/or *ZRSR2*‡‡ • Mutated *TP53*

Frequencies, response rates and outcome measures should be reported by risk category and, if sufficient numbers are available, by specific genetic lesions indicated.

†Mainly based on results observed in intensively treated patients. Initial risk assignment may change during the treatment course based on the results from analyses of measurable residual disease.

‡Concurrent *KIT* and/or *FLT3* gene mutation does not alter risk categorisation.

§AML with *NPM1* mutation and adverse-risk cytogenetic abnormalities are categorised as adverse-risk.

‖Only in-frame mutations affecting the basic leucine zipper (bZIP) region of *CEBPA*, irrespective of whether they occur as monoallelic or biallelic mutations, have been associated with favourable outcomes.

¶The presence of t(9;11)(p21.3;q23.3) takes precedence over rare, concurrent adverse-risk gene mutations.

#Excluding *KMT2A* partial tandem duplication (PTD).

**Complex karyotype: ≥3 unrelated chromosome abnormalities without other class-defining recurring genetic abnormalities; excludes hyperdiploid karyotypes with three or more trisomies (or polysomies) without structural abnormalities.

††Monosomal karyotype: presence of two or more distinct monosomies (excluding loss of X or Y) or one single autosomal monosomy in combination with at least one structural chromosome abnormality (excluding core-binding factor AML).

These markers should not be used as adverse prognostic markers if they co-occur with favourable-risk AML subtypes.

[a]*TP53* mutation at a variant allele fraction of at least 10%, irrespective of the *TP53* allelic status (mono- or biallelic mutation); *TP53* mutations are significantly associated with AML with complex and monosomal karyotype.

The table was adapted from Döhner *et al*. (2022) (Döhner, Andrew H. Wei et al., 2022).

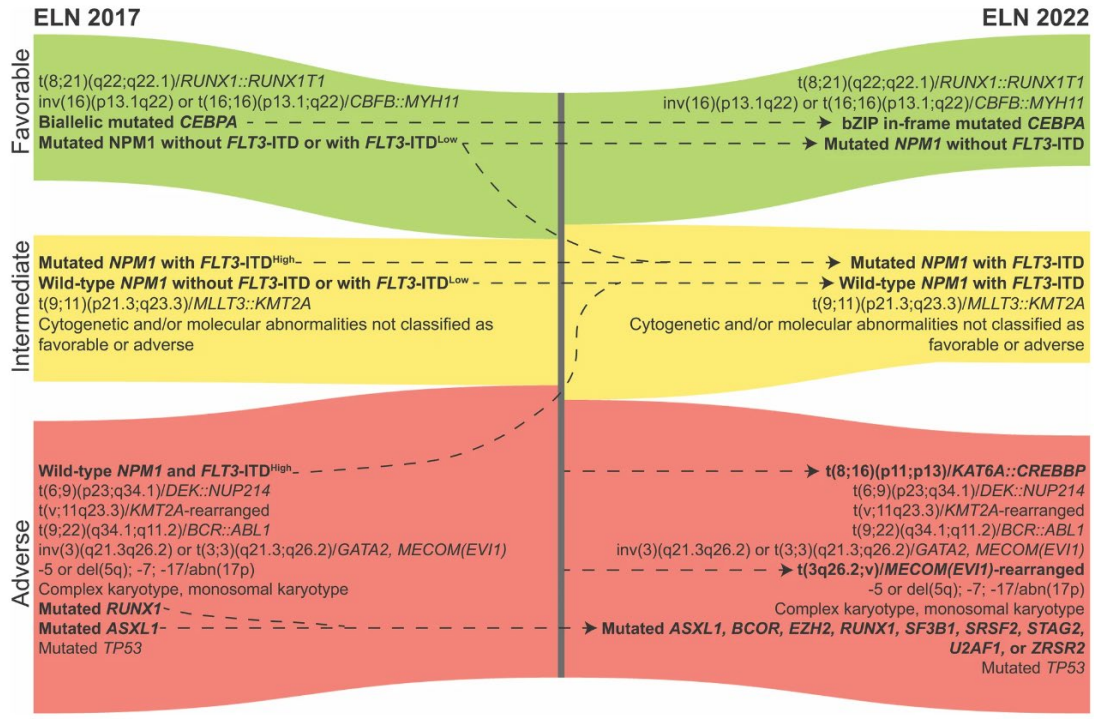

Figure 2.1 Sankey plots illustrating the changes in the ELN 2017 and 2022 risk stratification of AML (Bhansali et al., 2023).
The prognostic groups are coloured green (favourable), yellow (intermediate) and red (adverse). The dashed arrows show the changes in the risk stratification. bZIP refers to the basic leucine zipper domain.

2.2.2 Acute myeloid leukaemia-normal karyotype (AML-NK)

AML with a normal karyotype (AML-NK) is defined based on at least 20 bone marrow metaphases by conventional cytogenetics method (Mawad & Estey, 2012a; Döhner, Andrew H Wei, et al., 2022).Normal karyotypes were reported with varying frequencies between 25% to 70% in different geographical locations worldwide. Most studies reported that about 50% of the AML patients diagnosed presented with AML-NK, which was consistent with a study in Malaysia (Lazarevic et al., 2014; Amare et al., 2016; Ait Boujmia et al., 2021; Ambayya et al., 2021). The clinical presentation of AML-NK is similar to other subtypes of AML that include fatigue, weakness, malaise,

abdominal pain, fever, recurrent infections, and bleeding due to cytopenias, as well as enlarged spleen, lymph nodes and liver(Iqbal, 2012).

2.2.2(a) Genetic features of AML-NK

Although AML-NK presents with no chromosomal abnormalities by cytogenetic technique, studies have revealed a spectrum of somatic mutations in various genes associated with this disorder that affect the signalling pathways crucial for cell proliferation, maturation, and survival. Some frequently mutated genes in AML-NK include *NPM1*, associated with a favourable prognosis without other genetic lesions, including the *FLT3*-ITD mutation (Haferlach *et al.*, 2009; Abbas *et al.*, 2019).

The *FLT3*-ITD mutations occur in approximately 20 to 25% of AML-NK patients and are associated with adverse risks that include higher relapse rate, lower overall survival (OS) rate, poorer response to treatment and shorter disease-free survival (DFS) (Daver et al., 2021). Other relatively less common mutations include the *CEBPA* and *IDH1/2* mutations in about 10% of the AML-NK patients (Patel *et al.*, 2011; DiNardo & Cortes, 2016). Studies have revealed other genetic alterations in AML-NK patients, including epigenetic modifications, gene and microRNA dysregulation and fusion genes implicated in the AML-NK pathogenesis.

2.2.2(b) Treatment of AML and AML-NK

The general treatment approach for AML, including the AML-NK, has remained largely unchanged over the last 50 years. Most centres still rely on the standard induction regimen of daunorubicin and cytarabine (also known as Ara C) (DA 3+7), followed by a consolidation chemotherapy regimen to treat AML patients (Tang et al., 2021). Advancements in molecular genetics and genomics led to the

development of specific targeted therapy that targets mutant proteins and for in-depth monitoring of therapy response (minimal residual disease (MRD) markers) (Turkalj et al., 2023).

Therapeutic targets in AML can be categorised into six categories, as depicted in Table 2.4. Figure 2.2 illustrates an overview of therapeutic targets and the targeted pathways in AML. The targeted molecules are present at the AML cell surface, within the cytoplasm or the nucleus of the mitochondria.

Table 2.4 Therapeutic targets in AML.

Type of target	**Targets**
Genetic targets	Fusion Oncoproteins, *FLT3, IDH1/2*, *p53, KIT*
Targets involved in apoptosis	*BCL2, MCL-1, BCL-XL, MDM2*
Signalling molecules or nuclear receptors	RAS pathway, *SYK, RARα*
Epigenetic regulators	*DOT1L*–Menin inhibition
Surface proteins and/or molecules involved in immune signalling	Smoothened, CD33, CD123, CD47, *SIRPα, AXL, TIM-3, CLL-1, PD-1/PD-L1,* and *CTLA-4* axes
Transcription factors and other molecules subject to targeted protein degradation	*EZH2, GSPT1*

The table was adapted from Tulkalj *et al*. (2023) (Turkalj et al., 2023).

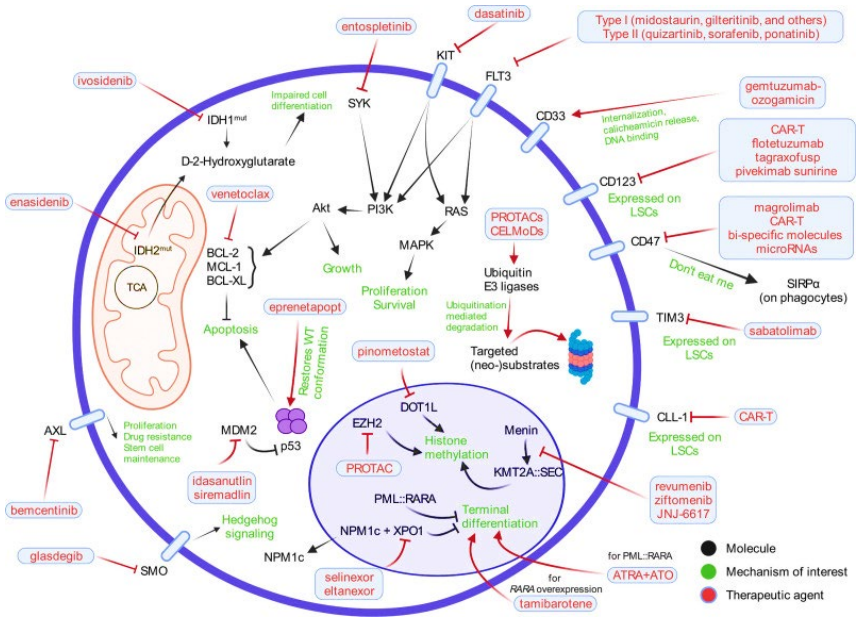

Figure 2.2 Schematic overview of therapeutic agents and targeted pathways in AML.
Therapeutic agents are labelled in red, and targeted molecules are marked in black. Mechanisms of interest (cellular or targeted molecule expression patterns) are labelled in green. The figure was adapted from Tulkalj *et al*. (2023) (Turkalj et al., 2023).

Although the majority of AML patients who have been treated with intensive chemotherapy achieve complete remission (CR), a substantial number of patients experience relapse if they were solely treated with intensive chemotherapy. Hence, stem cell transplant (SCT) will be the next potential curative modality in AML patients. Two main types of SCT are usually performed on AML patients: allogenic and autologous SCT (Suciu, 2003; Ferrara & Picardi, 2019; Chen *et al.*, 2022).

Allogenic-SCT (allo-SCT) is performed using stem cells derived from matched or partially matched healthy donors, while autologous-SCT (auto-SCT) is the patient's own stem cells. Allo-SCT is associated with a higher rate of side effects and mortality and risk of Graft Versus Host Disease (GVHD) in AML patients. Auto-SCT is more tolerated than allo-SCT, but there are risks of returning leukemic cells to the patients, and it is less commonly done in AML patients. Studies have shown that allo-SCT has superior outcomes compared to patients with auto-SCT when performed after

the first complete remission (CR1) was attained in AML patients (Suciu, 2003; Ferrara & Picardi, 2019; Chen *et al.*, 2022).

2.2.2(c) Prognostication and survival of AML-NK

The prognostication of AML-NK is challenging due to its inherent heterogeneity. Although AML-NK excludes high-risk associated chromosomal abnormalities, it does not necessarily stipulate good patient outcomes. Studies have reported that the 5-year survival rates of AML-NK ranged between 24.0-42.0% (Marcucci et al., 2005; Nimer, 2008; Stölzel et al., 2016; Samra et al., 2020). The OS of AML-NK patients are also influenced by the presence of genetic aberrations such as *FLT3*-ITD, *NPM1*, and *WT1* and clinical factors that include age at diagnosis, total white blood cell (TWBC) count at diagnosis and response to first-line chemotherapy(Liersch *et al.*, 2014; Alsulami *et al.*, 2021; Döhner, Andrew H. Wei, *et al.*, 2022).

The most extensive multicentred study in Malaysia on AML patients (non-M3) that included 1106 patients (82.2%. n=908 received intensive chemotherapy) revealed that the median OS was 15 months and event-free survival (EFS) was 12 months. The 3-year OS and EFS for the patients in this cohort were 32.9% and 28.5%, respectively. In the intermediate risk group (39.5%, n=437/1106), which includes the AML-NK patients, the median survival of patients without SCT was 10 months (3-year OS=23.0%) and in patients who have undergone SCT, the median survival was 53 months (3- year OS=56.6%). In their study, multivariate analysis revealed that the age of AML patients at diagnosis (>60 years), gender (male), TWBC >100 x 10^9/L, relapse (<12 months of treatment), refractory post-induction and the high-risk genetic group

as defined by ELN 2017 risk stratification were prognostic factors associated with dismal OS and EFS (Leong *et al.*, 2019).

2.2.3 AML in Malaysia

There were relatively few findings on AML cytogenetic and molecular findings in Malaysia. Two main studies with the largest cohort of AML patients include a study by Ambayya *et al*. (2021) comprising a total of 854 adult AML patients and the second study by Meng *et al*. (2013) that consisted of a total of 480 adults and paediatric *de novo* AML patients (Meng et al., 2013; Ambayya et al., 2021). Other studies were smaller-scale and focused on AML patients residing on the east coast of Peninsular Malaysia (Yunus *et al.*, 2015; Hamdan *et al.*, 2022).

2.2.3(a) Cytogenetic profiles of AML in Malaysia

In Malaysia, the cytogenetic profiles of adult AML patients, as described by Ambayya *et al*. (2021), are summarised in Table 2.5. In 601 patients with successful karyotypes, about 48.9% (n=294) of the AML patients presented with a normal karyotype. The most common translocations among the AML patients in this study were: t(15:17) seen in 14.3% (n=86/601), t(8;21) seen in 8.5% (n=51/601) and inv(16) seen in 5.0% (n=30/601). In terms of numerical changes in chromosomes in the study, del(5q/7q) was present in 5.3% (n=32/601), and trisomy 8 was seen in 3.16% (n=19/601) of the AML patients (Ambayya et al., 2021).