THE CLINICAL SIGNIFICANCE OF *NPM1*, *FLT3* AND *CEBPA* GENE MUTATIONS IN ADULT PATIENTS WITH ACUTE MYELOID LEUKAEMIA

IZAZ UL HAQ

UNIVERSITI SAINS MALAYSIA

2020

THE CLINICAL SIGNIFICANCE OF *NPM1*, *FLT3* AND *CEBPA* GENE MUTATIONS IN ADULT PATIENTS WITH ACUTE MYELOID LEUKAEMIA

by

IZAZ UL HAQ

Thesis submitted in the fulfilment of the requirements
for the degree of
Master of Science

November 2020

In the Name of Allāh, the Most Gracious, the Most Merciful

ACKNOWLEDGEMENT

First and foremost, I would like to thank Allah Almighty for giving me the strength, knowledge, ability, and opportunity to undertake this research study and persevere and complete it satisfactorily. Without his blessings, this achievement would not have been possible.

Special appreciation goes to my supervisor, Dr Abdul Rahim Hussein, for his supervision and constant support. His invaluable help of constructive comments and suggestions throughout the experimental and thesis works have contributed to this research's success. He has given me all the freedom to pursue my research while silently and non-obtrusively ensuring that I stay on course and do not deviate from the core of my research. Without his able guidance, this thesis would not have been possible, and I shall eternally be grateful to him for his assistance. I have incredible respect to offer the heartiest appreciation to my worthy co-supervisor Associate Prof Dr Muhammad Farid Johan, Associate Prof Dr Badrul Hisham Yahaya, their guidance, motivating directions, caring conduct, and encouraging me throughout my research work.

Expressions don’t come straightforward for me to say the feeling of my modest commitment towards my sweetest Parents who dependably raised their hands for my victory and gave me an air to attain this objective. I pay my exceptional respects to my brothers and dear sisters for their otherworldly impulse, well wishes, and genuine requests to Allah for my victory.I shall be missing something if I don’t extend my admiration and appreciation to all my friends, especially to my brother Abdulwali sister Nur Azirah and Dr Siti Asma, and my other colleague during my laboratory work they provide me with their excellent co-operation.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF ABBREVIATIONS

AML	Acute myeloid leukaemia
NPM1	nucleophosmin
$NPM1^{+}$	nucleophosmin
CBF	Core binding factor
°C	Degree Celsius
CSGE	Conformation sensitive gel electrophoresis
PCR	Polymerase chain reaction
Plt	platelet
CN	cytogenetically normal
CA	cytogenetical abnormal
NLS	nuclear localisation signal
DNA	deoxyribonucleic acid
Del	deletion
WT	wild type
WHO	world health organisation
WBC	white blood cell
EDTA	ethylenediamine tetra acetic acid
Nt	nucleotide
µM	micromolar
µg	microgram
RBC	red blood cell
µl	microlitre
Sec	second
SNP	single nucleotide polymorphism
RTK	receptor tyrosine kinase
Ng	nanogram

NoLS	nucleoli localization signal
NRAS	neuroblastoma RAS
TM	transmembrane
TEMED	tetramethylethylenediamine
TKI	tyrosine kinase inhibitor
U	unit
TET2	tet methyl cytosine dioxygenase 2
Tm	melting temperature
NES	nuclear export signal
MDS	myelodysplasia syndrome
M	molar
Hb	haemoglobin
H	hour
Bp	Base pairs
CEBPA	CCAAT enhancer-binding protein alpha
NPM1-	FLT3-ITD non-mutant
FLT3-ITD-	FLT3-ITD non-mutant
FLT3-ITD+	FLT3-ITD mutant
Rpm	revolutions per minute
SSCP	single-strand conformational polymorphism
FLT3	Fms-like tyrosine kinase-3
APML	Acute promyelocytic leukaemia
FAB	French-American-British
BM	Bone marrow
TKD	Tyrosine kinase domain
ITD	internal tandem duplication
ddH2O	Double distilled water
PLT	Platelets
NA	Not available

PB	Peripheral blood
NK	Normal karyotype
Abn	Abnormalities
DBD	DNA binding domain
LZD	leucine zipper domain

LIST OF APPENDICES

KEPENTINGAN KLINIKAL MUTASI GENE *NPM1, FLT3* DAN *CEBPA* PADA PESAKIT LEUKEMIA MIELOID AKUT DEWASA

ABSTRAK

Leukemia myeloid akut (AML) adalah kanser hematologi biologi kompleks dan klinikal heterogen dan peningkatan insiden di kalangan orang tua. Umur, sitogenetik dan mutasi genetik tetap menjadi faktor prognostik penting untuk hasil rawatan. Tujuan utama kajian ini adalah untuk menentukan kelaziman dan kepentingan klinikal mutasi *NPM1*, *FLT3* dan *CEBPA* AML di Semenanjung Malaysia.Tiga gen mutasi AML yang biasa dan penting iaitu *FLT3* (exon 14-15, exon 20), *NPM1* (exon 12), dan *CEBPA* (exon 1) ditentukan oleh polymerase chain reaction (PCR) dan elektroforesis gel sensitif konformasi (CSGE) pada 47 pesakit dewasa yang baru didiagnosis dengan AML. Fisher-Exact Test digunakan untuk menganalisis perbezaan signifikan AML dengan mutan *FLT3*-ITD, *NPM1*, dan *CEBPA* dalam hubungannya dengan parameter klinikal. Sampel sumsum tulang (BM) dan sampel darah periferal (PB) dianalisis mengikut ujian yang diperlukan dan data dikumpulkan. Kelangsungan hidup kumulatif (OS) satu tahun dari pelbagai mutasi gen gabungan dianalisis dengan analisis keluk Kaplan-Meier. Dalam kajian ini, mutasi gen *NPM1* 9 (19.2%) mempunyai prevalensi tertinggi, diikuti oleh mutasi *CEBPA* 7 (14.9%) dan *FLT3*-ITD 5 (10.6%) tetapi tidak ada mutasi yang dikesan pada *FLT3*-TKD. Sedangkan kelaziman gabungan Mutasi mengesan $NPM1^{+}/FLT3^{-}/CEBPA^{-}$ (17%), $NPM1^{+}/FLT3^{+}/CEBPA^{-}$ (2.12%), $NPM1^{-}/FLT3^{-/}CEBPA^{+}$ (10.68%), $NPM1^{-}/FLT3^{+}/CEBPA^{+}$ (4.25%), $NPM1^{-}/FLT3^{+}/CEBPA^{-}$(4.25%) dan $NPM1^{-}/FLT3^{-}/CEBPA^{-}$ *(61.7.%)*. Tidak ada pesakit yang dikesan dengan ketiga-tiga mutasi ($NPM1^{+}/FLT3^{+}/CEBPA^{+}$). Walau bagaimanapun, FLT3-ITD mutasi gen telah

dikaitkan dengan ketara dengan peratusan letupan yang lebih tinggi (nilai-*p* = 0.045) dan tuduhan WBC tinggi (nilai-*p* = 0.005). Dan *CEBPA* mutasi gen telah dikaitkan dengan ketara dengan kiraan tinggi platelet (nilai-*p* = 0.001), dan klasifikasi FAB (nilai-*p* = 0,044). Ekspresi tertinggi antigen CD33 adalah (93.6%). Selepas itu (89.4%) ekspresi antigen MPO, CD13, dan CD117 diikuti oleh CD64 (68.1%), HLA-DR (59.6%), CD34 (57.4%) dan CD56 (37.8%). Walau bagaimanapun, ekspresi antigen CD7 dicatatkan paling rendah (13%). Keabnormalan kromosom didapati 44.7%, kariotip normal 42.6%, dan hasil kariotip 12.7% tidak terdapat pada pesakit AML. Kelangsungan hidup keseluruhan kumpulan mutasi gabungan setelah satu tahun diagnosis adalah $NPM1^{-}/FLT3^{+}/CEBPA^{-}$ (100%), $NPM1^{-}/FLT3^{-}/CEBPA^{+}$ (80%), $NPM1^{-}/FLT3^{-}/CEBPA^{-}$ (48.1%), $NPM1^{+}/FLT3^{-}/CEBPA^{-}$ (50%). Walau bagaimanapun, pesakit dengan $NPM1^{-}/FLT3^{+}/\ CEBPA^{+}$ dan $NPM1^{+}/FLT3^{+}/CEBPA^{-}$ tidak terselamat setelah satu tahun didiagnosis. Kajian menunjukkan bahawa mutasi gen *NPM1* 9 (19.2%) mempunyai prevalensi tertinggi, diikuti oleh mutasi *CEBPA* 7 (14.9%) dan *FLT3*-ITD 5 (10.6%). Mutasi gen *FLT3*-ITD berkorelasi secara signifikan dengan peratusan letupan yang lebih tinggi (nilai *p* = 0.045) dan jumlah WBC yang tinggi (nilai *p* = 0.005). Mutasi gen *CEBPA* berkorelasi secara signifikan dengan jumlah platelet yang lebih tinggi (nilai *p* = 0.001), dan klasifikasi FAB (nilai *p* = 0,044). Sedangkan mutasi gen *NPM1* tidak berkorelasi secara signifikan dengan parameter klinikal. Dalam kajian ini, ekspresi antigen tertinggi adalah CD33 (93.6%) sementara MPO, CD13, dan CD117 adalah (89.4%). Tiga mutasi gen gabungan pertama dengan kelangsungan hidup terpanjang adalah $NPM1^{-}/FLT3^{+}/CEBPA^{-}$ (1.0%), $NPM1^{-}/FLT3^{-}/CEBPA^{+}$ (0.8%) dan $NPM1^{+}/FLT3^{-}/CEBPA^{-}$ (0.5%).

THE CLINICAL SIGNIFICANCE OF *NPM1, FLT3* AND *CEBPA* GENE MUTATIONS IN ADULT PATIENTS WITH ACUTE MYELOID LEUKAEMIA

ABSTRACT

Acute myeloid leukaemia (AML) is a biologically complex and clinically heterogeneous haematological cancer, and its incidence increases in the elderly. Age, cytogenetics and genetic mutations remain important prognostic factors for treatment outcome. The main aim of the study was to determine the prevalence and clinical significance of *NPM1, FLT3* and *CEBPA* mutations of AML in the Peninsula Malaysia. Three common and important gene mutations in AML *FLT3* (exon 14-15, exon 20), *NPM1* (exon 12), and *CEBPA* (exon 1) were determined by polymerase chain reaction and conformation-sensitive gel electrophoresis (CSGE) in 47 newly diagnosed adult patients with AML. Fisher-Exact Test was used to analyse the significant difference of AML with *FLT3-ITD, NPM1*, and *CEBPA* mutants in association with the clinical parameters. The bone marrow (BM) aspirates and peripheral blood (PB) samples were analysed according to the tests required, and the data were collected. The one-year cumulative overall survival (OS) of various combined genes mutations was analysed by Kaplan-Meier curve analysis. In this study, *NPM1* 9(19.2%) gene mutations had the highest prevalence, followed by *CEBPA* 7(14.9%) and *FLT3*-ITD 5(10.6%) mutations but there was no mutation detected in *FLT3*-TKD. whereas the prevalence of Combined mutations detected $NPM1^{+}/FLT3^{-}/CEBPA^{-}$ (17%), $NPM1^{+}/FLT3^{+}/CEBPA^{-}$ (2.12%), $NPM1^{-}/FLT3^{-/}CEBPA^{+}$ (10.68%), $NPM1^{-}/FLT3^{+}/CEBPA^{+}$ (4.25%), $NPM1^{-}/FLT3^{+}/CEBPA^{-}$(4.25%) and $NPM1^{-}/FLT3^{-}/CEBPA^{-}$ *(61.7.%)*. None of these patients detected with all three mutations

(*NPM1*$^{+}$/*FLT3*$^{+}$/*CEBPA*$^{+}$). However, *FLT3*-ITD gene mutation were significantly correlated with higher blast percentage (*p*-value= 0.045) and high WBC counts (*p*-value= 0.005). And *CEBPA* gene mutation was significantly correlated with higher platelet count (*p*-value = 0.001), and FAB classification (*p*-value = 0.044). The highest expression of antigen CD33 was (93.6%). After that (89.4%) expression of antigens MPO, CD13, and CD117 follow by CD64 (68.1%), HLA-DR (59.6%), CD34 (57.4%) and CD56 (37.8%). However, the antigen CD7 expression was recorded lowest (13%). The cytogenetic abnormality was found 44.7%, normal karyotype was 42.6%, and 12.7% cytogenetic data were not available in AML patients. The overall survival of the combined mutational groups after one year of diagnosis was *NPM1*$^{-}$/*FLT3*$^{+}$/*CEBPA*$^{-}$ (100%), *NPM1*$^{-}$/*FLT3*$^{-}$/*CEBPA*$^{+}$ (80%), *NPM1*$^{-}$/*FLT3*$^{-}$/*CEBPA*$^{-}$ (48.1%), *NPM1*$^{+}$/*FLT3*$^{-}$/*CEBPA*$^{-}$ (50%). However, those patients with *NPM1*$^{-}$/*FLT3*$^{+}$/*CEBPA*$^{+}$ and *NPM1*$^{+}$/*FLT3*$^{+}$/*CEBPA*$^{-}$ were not survived after one year of diagnosis. The study revealed *NPM1* 9(19.2%) gene mutations had the highest prevalence, followed by *CEBPA* 7(14.9%) and *FLT3*-ITD 5(10.6%) mutations. *FLT3*-ITD gene mutation were significantly correlated with higher blast percentage (*p*-value= 0.045) and high WBC counts (*p*-value= 0.005). *CEBPA* gene mutation was significantly correlated with higher platelet count (*p*-value = 0.001), and FAB classification (*p*-value = 0.044). Whereas *NPM1* gene mutation was not significantly correlated with any clinical parameters. In this study, the highest expression of antigens was CD33 (93.6%) while MPO, CD13, and CD117 was (89.4%). The first three combined gene mutations with longest overall survival were *NPM1*$^{-}$/*FLT3*$^{+}$/*CEBPA*$^{-}$ (1.0%), *NPM1*$^{-}$/*FLT3*$^{-}$/*CEBPA*$^{+}$ (0.8%) and *NPM1*$^{+}$/*FLT3*$^{-}$/*CEBPA*$^{-}$ (0.5%).

CHAPTER 1

INTRODUCTION

1.1 Acute myeloid leukaemia

1.1.1 Definition

Acute myeloid leukaemia (AML) is a heterogeneous group of malignant clonal disorders originating from early bone marrow haemopoietic progenitors, called myeloblasts, characterised by the proliferation of poorly differentiated blast cells in the bone marrow, blood and other tissue which interfere with normal haematopoiesis (Heerema-McKenney & Arber, 2009).

1.1.2 Clinical presentation of AML

Clinical AML manifestation bone marrow failure is the result of the accumulation of immature myeloid blast cells leads to a decrease in average blood cell production. AML patients would appear with symptoms such as fever, weight loss, lethargy, and breathlessness caused by anaemia, bleeding caused by thrombocytopenia and infections, especially of the skin, chest, and mouth caused by neutropenia. Leukaemic infiltration can also arise in any organ, but most commonly occurs in liver, spleen, gums, and skin.

1.1.3 Diagnosis of AML

The diagnosis of AML is based on the appearance of the leukaemic blast cells in the bone marrow (BM) and/or peripheral blood (PB) examined under light microscopy using a different cytochemical stains cytomorphology. The World Health Organization (WHO) criteria for AML diagnosis is ≥20% of blast cells in bone marrow or blood. Patients with the following cytogenetic abnormalities t(8;21)(q22;q22), inv(16)(p13.1q22), t(16;16)(p13.1;q22) or t(15;17)(q22;q12) may be diagnosed as AML though the blast count is below 20% (Vardiman *et al.*, 2009). Forty to fifty per cent of AML patients have normal karyotype (AML-NK) at diagnosis (Gulley, Shea & Fedoriw, 2010).

1.1.4 AML Classification

There are two main classifications of AML. Initially, the AML classification system has been established by the French-American-British (FAB) classification, and the other one is the World Health Organization (WHO) classification system.

The classification French-American-British (FAB) system was first proposed in 1976 (Bennett *et al.*, 1976), and AML is classified into eight subtypes, M0-M7 based on the type of cell from which the leukaemia originated and its degree of maturity (Table 1.1). This classification system was initially based on the leukaemic cell morphology and cytochemical stains and immunological analysis. (Lowenberg, B & JR, 1999). However, the FAB classification is still commonly used.

The AML classification of the World Health Organization (WHO) classification is a new system that integrates cytogenetics, molecular genetics, immunological markers, and morphology (Table 1.2). Furthermore, the provisional entities AML with genetic mutations namely, mutated *NPM1* and AML with mutated *CEBPA* have being presented in 2008 revision (Vardiman *et al.*, 2009), Though AML with *RUNX1* and AML with *BCR-ABL1* mutated were familiarised as part of revision in 2016 (Arber *et al.*, 2016) (Table 1.2), making the classification of the WHO more detailed and clinically more significant.

A new type is given in the revised 2016 WHO classification of hematopoietic tumours: "myeloid neoplasms with germline predisposition." These entities are infrequent but are also now underdiagnosed and underreported. WHO comprises three types of myeloid neoplasms with germline predisposition: neoplasms with preexisting platelet disorders, neoplasms without the preexisting disorder or organ dysfunction, and neoplasms with other organ dysfunction Table 1.3 (Geyer, 2019; Baptista *et al.*, 2017).

1.1.5 Prognosis of AML

Cytogenetics is one of the most important independent prognostic factors in AML. Cytogenetically, AML can be divided into two main groups, one with normal karyotype (NK), which is the larger group (55%), and the other one with an abnormal karyotype. Within the normal cytogenetics, further important subgroup definitions are possible based on the mutation status of genes such as *NPM1* (Nucleophosmin), *FLT3* (FMS-like tyrosine kinase 3), and *CEBPA* (CCAAT/enhancer-binding protein alpha), which are common in AML. Based on cytogenetics and molecular genetics, patients are

classified into one of three risk groups: favourable, intermediate, and adverse (Table 1.4) (Döhner *et al.*, 2017).

Table 1.1 French-American-British (FAB) classification system of AML

FAB subtype	**Description**
M0	Acute myeloblastic leukaemia minimal maturation
M1	Acute myeloblastic leukaemia without maturation
M2	Acute myeloblastic leukaemia with maturation
M3	Acute promyelocytic leukaemia
M4	Acute promyelocytic leukaemia
M5	Acute monoblastic leukaemia
M6	Acute erythroid leukaemia
M7	Acute megakaryoblastic leukaemia

Table 1.2 WHO classification of acute myeloid leukaemias (AML)

1. AML with recurrent genetic abnormalities
AML with inv(16)(p13.1q22) or t(16;16)(p13.1;q22); *CBFB-MYH11*
AML with t(9;11)(p22;q23); *MLLT3-MLL*
AML with t(8;21)(q22;q22); *RUNX1-RUNX1T1*
AML with inv(3)(q21q26) or t(3;3)(q21;q26.2); *RPN1-EVI1*
APL with t(15;17)(q22;q12); *PML-RARA*
AML with t(6;9)(p23;q34); *DEK-NUP214*
AML (megakaryoblastic) with t(1;22)(p13;q13); *RBM15-MKL1*
AML with mutated *CEBPA*
AML with mutated *NPM1*
AML with BCR-ABL1 (provisional entity)
AML with mutated *RUNX1* (provisional entity)
2. AML with myelodysplasia-related changes
3. Therapy-related myeloid neoplasms
4. AML, not otherwise specified (NOS)
AML with maturation
AML without maturation
AML with minimal differentiation
Acute myelomonocytic leukaemia
Acute monoblastic and monocytic leukaemia
Pure erythroid leukaemia
Acute erythroid leukaemia
Acute basophilic leukaemia
Acute megakaryoblastic leukaemia
Acute panmyelosis with myelofibrosis
5. Myeloid sarcoma
6. Myeloid proliferations related to Down syndrome:
Transient abnormal myelopoiesis
Myeloid leukaemia associated with Down syndrome

*Adapted from (Arber *et al.*, 2016)

Table 1.3 Classification of myeloid neoplasms with germline predisposition

Myeloid neoplasms with germline predisposition without a preexisting disorder or organ dysfunction
Acute myeloid leukaemia with germline *CCAAT*/enhancer-binding protein-A mutation
Myeloid neoplasm with germline *DDX41* mutation
Myeloid neoplasms with germline predisposition and preexisting platelet disorders
Myeloid neoplasms with germline *RUNX1* mutation
Myeloid neoplasms with germline *ANKRD26* mutation
Myeloid neoplasms with germline *ETV6* mutation
Myeloid neoplasms with germline predisposition and other organ dysfunction
Myeloid neoplasms with germline *GATA2* mutation
Myeloid neoplasms with germline predisposition with BM failure syndromes
Myeloid neoplasms with germline predisposition with telomere biology disorders
Juvenile myelomonocytic leukaemia associated with neurofibromatosis, Noonan syndrome, or Noonan syndrome-like disorders
Myeloid neoplasms associated with Down syndrome

*Adapted from (Baptista *et al.*, 2017)

Table 1.4 Classification of cytogenetic risk groups (Döhner *et al.*, 2017)

Genetic group	Subsets
Favorable	t(8;21)(q22;q22.1); *RUNX1-RUNX1T1* inv(16)(p13.1q22); *CBFB-MYH11* Mutated *NPM1* without *FLT3*-ITD (normal karyotype) Biallelic mutated *CEBPA* (normal karyotype)
Intermediate	Mutated *NPM1* and *FLT3*-ITD Wild-type *NPM1* without *FLT3*-ITD t(9;11)(p21.3;q23.3); *MLLT3-KMT2A* Cytogenetic abnormality not classified as favourable or adverse
Adverse	t(6;9)(p23;q34.1); *DEK-NUP214* t(v;11q23.3); *KMT2A* rearranged t(9;22)(q34.1;q11.2); *BCR-ABL1* inv(3)(q21.3q26.2); *GATA2,MECOM(EVI1)* -5 or del(5q); -7; -17/abn(17p) Complex karyotype and monosomal karyotype Wild-type *NPM1* and *FLT3*-ITD Mutated *ASXL1* Mutated *TP53*

Abn: Abnormalities, *CEBPA*: CCAAT/enhancer-binding protein alpha, del: deletion, *FLT3* ITD, Fms-related tyrosine kinase 3 internal tandem duplications; *MLLT3-MLL*, mixed-lineage leukemia; *NPM1*, nucleophosmin 1. *RUNX1:* Runt-related transcription factor 1

1.2 AML Pathogenesis

It is generally recognised that AML is the result of a multistep process (Sawyers, Denny & Witte, 1991; Fröhling *et al.*, 2005). According to this idea, the transformed haemopoietic progenitor cells initially acquire a single genetic lesion resulting in only slight variation in cell survival, proliferation, or maturational abilities as compared to their normal counterparts. The additional genetic lesion may result in full leukaemic transformation (Smith *et al.*, 2004; Grisolano *et al.*, 2003; Kelly & Gilliland, 2002a). Mutational analysis of occasional AML has confirmed that many patients have more than one recurring genetic abnormality. Numerous genes have been observed to be frequently mutated in the leukaemic cells of AML patients. The combination of a range variety of chromosomal abnormalities and heterogeneity basis found in this disease. Therefore, AML is not as a result of a single genetic alteration, but it requires at least two genetic alterations to develop into AML. For many years the two-hit hypothesis was used as a model of leukemogenesis. These genetic events underlying AML pathogenesis can be divided into two complementary main classes, I (proliferative) and II (blocking mutation) (figure 1.1). Which divide mutations and other genetic alterations. These different types of genetic mutations collaborate in leukaemogenesis (Kelly & Gilliland, 2002a).

Class I mutations incorporate the mutations that activate signal transduction pathways, leading to increased proliferation and existence of haemopoietic progenitor cells. This group include genes such as *RAS, c-KIT, FLT3,* etc (Schlenk *et al.*, 2008). Class II mutation, consists of mutations that affect transcription factors and impaired haematopoietic differentiation such as *CEBPA* (CCAAT/enhancer-binding protein α

gene), *MLL* (mixed-lineage leukaemia gene) and the *NPM1* (Nucleophosmin gene) (Kelly & Gilliland, 2002a).

However, recent studies discovered genetic mutations that do not belong to these two classes, which were categorised with new class mutations (class III) associated with epigenetic regulation (figure 1.2). Class III genes are encoding epigenetic modifiers, including *EZH2*, *DNMT3A*, *TET2*, *ASXL1*, *IDH1*, and *IDH2*, which would also contribute a significant role in AML pathogenesis (Chen, Shen & Chen, 2013; Shen *et al.*, 2011; Dombret, 2011). These mutations are affecting epigenetic regulators which are not considered as Class I and Class II, and this indicates that the "two-hit model" is no longer adequate (Shih *et al.*, 2012).

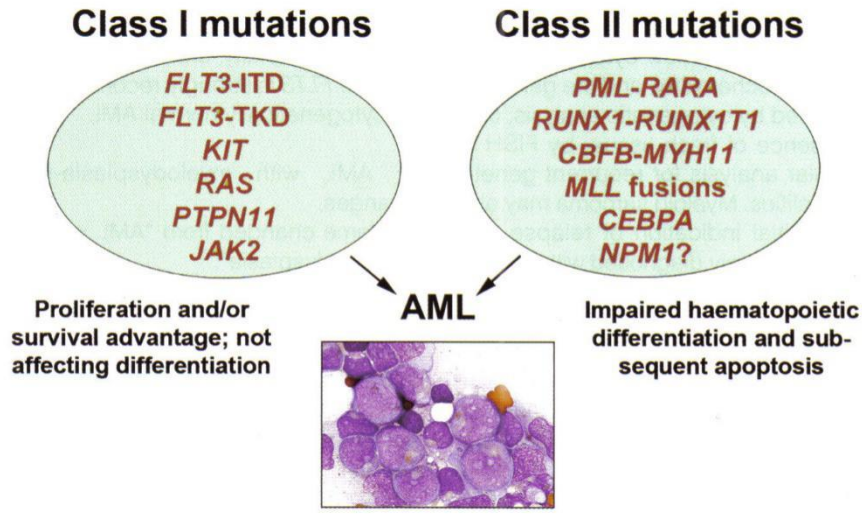

Figure 1.1 Mutations in AML pathogenesis (Kelly & Gilliland, 2002a).

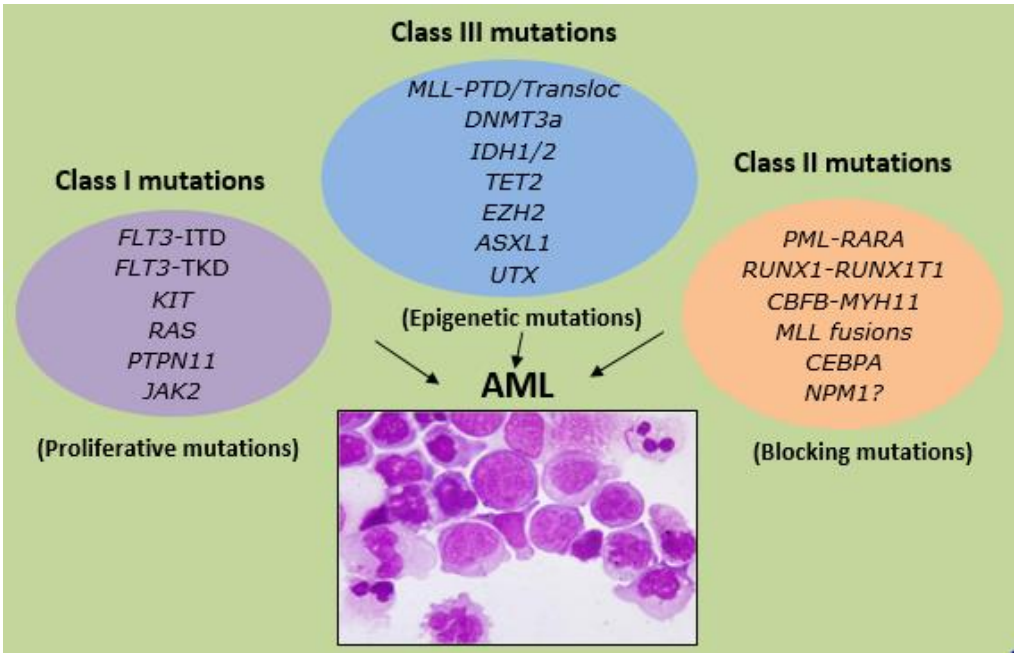

Figure 1.2 Mutations in AML pathogenesis with a new class.

1.3 Target gene

1.3.1 *FLT3* gene

FMS-like tyrosine kinase 3 (*FLT3*) is a receptor tyrosine kinase; it is also known as stem cell tyrosine kinase-1 (STK1) or fetal liver tyrosine kinase-2 (FLK-2). Human *FLT3* gene is located on chromosome 13q12, which contains 24 exons and 993 amino acid (GenBank Accession No: NM_004119.2) (Abu-duhier *et al.*, 2001a). It encodes for the *FLT3* receptor protein belonging to the class III receptor tyrosine kinase. The class III receptor tyrosine kinase family comprises of the platelet-derived growth factor receptors (PDGFR), *FLT3*, and *KIT*. Class III receptor tyrosine kinases (RTKs) share sequence homology and have a similar overall structure, with five immunoglobin-like domains in extracellular with a single TM (transmembrane domain), a JM (juxtamembrane domain), two intracellular tyrosine kinase domains (TK1 and TK2) separated by a kinase insert domain (KI) and a C-terminal domain (Figure 1.3) (Abu-Duhier *et al.*, 2000). Class I mutations incorporating the mutations that activate signal transduction pathways, leading to increased proliferation and existence of haemopoietic progenitor cells. (Kelly *et al.*, 2002). Typically, wild type FLT3 is expressed on regular hematopoietic stem cells. It is involved in receptor dimerisation and autophosphorylation activation of downstream signal transduction pathways that play a vital function in proliferation, differentiation, survival and cell growth of hematopoietic progenitors and creating a connection with the RAS pathway which is also involved in leukaemogenesis (Grafone *et al.*, 2012).

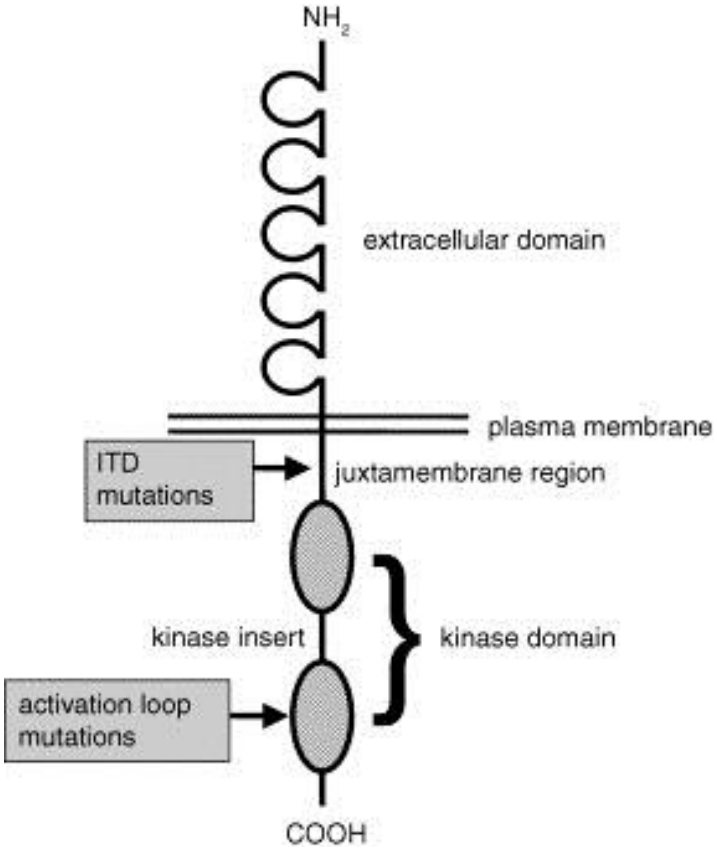

Figure 1.3 The schematic description of the functional domains of the *FLT3* gene (Adapted from (Small, 2006)).

FLT3 mutations have been observed 25% in patients with acute myeloid leukaemia (AML). The FLT3 mutation has been categorised into two main types. The type one mutation is *FLT3*/ITD (internal tandem duplications) in juxtamembrane (JM) domain within exon 14 and 15, and type two mutation is point mutation or in the tyrosine kinase domain (*FLT3*/TKD) (Figure 1.3) (Kiyoi *et al.*, 1997; Chauhan *et al.*, 2011). *FLT3-ITD* is similarly well-known as an FLT3-LM (length mutation), and the first time was describe by (Nakao *et al.*, 1996). The length of the ITD varies from 3 to over 400 base pair nucleotides. This type of mutation is the most common mutation in the *FLT3* gene (Stirewalt & Radich, 2003; Schnittger *et al.*, 2002). Internal tandem duplications (*FLT3*-ITD) have been detected in 20-25% in the JM of adult patients with AML (Fröhling *et al.*, 2002; Thiede *et al.*, 2002) and 15 per cent in paediatrics patients (Meshinchi *et al.*, 2008). Even though the second most occurring mutation in *FLT3* is point mutations in exon 20 at codon D835 (aspartate) within the tyrosine kinase domain activation loop (*FLT3-TKD*). Typically, *FLT3-TKD* mutation is less found in AML, and the previous study represents 7% of adults and 3% of children in AML (Abu-Duhier *et al.*, 2001b; Yamamoto *et al.*, 2001; Thiede *et al.*, 2002).

In AML patients, *FLT3*-ITD mutation is associated with poor prognosis, including higher relapse rates from complete remission and decrease OS (Small, 2006; Kottaridis *et al.*, 2001; Fröhling *et al.*, 2002). However, the *FLT3*-TKD prognosis in AML is still indistinguishable (Whitman *et al.*, 2010; Fröhling *et al.*, 2002; Thiede *et al.*, 2002; Abu-Duhier *et al.*, 2001b).

1.3.2 *NPM1* gene

The gene *NPM1* (nucleophosmin 1) is found in human beings in chromosome 5q35 and contains 12 exons. It encodes for the nucleophosmin (NPM) protein, which contains 294 amino acids. *NPM1* is also known as numatrin, B23. The *NPM1* protein is a multifunctional nucleocytoplasmic shuttling protein that is localised primarily in the nucleolus (Cordell *et al.*, 1999). It has several functions that carry ribosome components for ribosome biogenesis to the cytoplasm (Szebeni & Olson, 1999). Operation of centrosomal duplication throughout cell division (Okuda *et al.*, 2000), and association with the tumour suppressor gene (p53 and p19Arf) to operation apoptosis and proliferation of cells (Colombo *et al.*, 2002).

NPM1's function is based on its nucleocytoplasmic shuttle cycle, which is regulated by three central areas, such as NES (nuclear export signal), NLS (nuclear localisation signal), and NoLS (nucleoli localisation signal) (Figure 1.6). NES is essential for the transport of NPM protein from the nucleoplasm to the cytoplasm. While NLS and NoLS are essential for localisation of the NPM protein in nucleoplasm and nucleolus (Falini *et al.*, 2007; Chen, Rassidakis & Medeiros, 2006).

NPM1 gene mutation in adult AML patients is accounted for approximately 26-35% (Falini *et al.*, 2005; Chetsada, Wanna & Chirayu, 2008) and *NPM1* mutations occur in about 40-60% of adult CN-AML cases (Nafea *et al.*, 2011; Chetsada, Wanna & Chirayu, 2008). These mutations frequently occurred in older age groups than in younger adults. However, in pediatric AML, the incidences are lower 8% (Braoudaki *et al.*, 2010). The WHO 2016 accepts AML as a distinct entity with mutated *NPM1* (Arber *et al.*, 2016). Exon 12 shows these mutations, encoding the NPM protein C-terminus and commonly heterozygous (Figure 1.5). Several *NPM1* gene mutations were identified in AML at nucleotide positions 956 through 971 (Chen, Rassidakis & Medeiros, 2006; Zhu, Ma & Liu, 2010). All these mutations are either deletions or insertions that cause the frameshift mutation at the C-terminus region in the *NPM1* gene, and these mutations stop the production of protein in the cytoplasm.

In exon 12 of the *NPM1* gene, more than 50 different types of mutations have been reported. These mutations characterise either simple insertions of 4bp or more complex deletion and insertions (Figure1.4) (Schnittger *et al.*, 2009). Mutation A is the most common among all types of mutations and accounts for 75-85% of these mutations. Mutation A is the duplication of a TCTG tetranucleotide at position 956 to 959 (Falini *et al.*, 2005). Mutations B and D, which represent 10 per cent and 5 per cent of all *NPM1* mutations in CN-AML, individually. Relatively other types of mutations are unusual (Falini *et al.*, 2007).

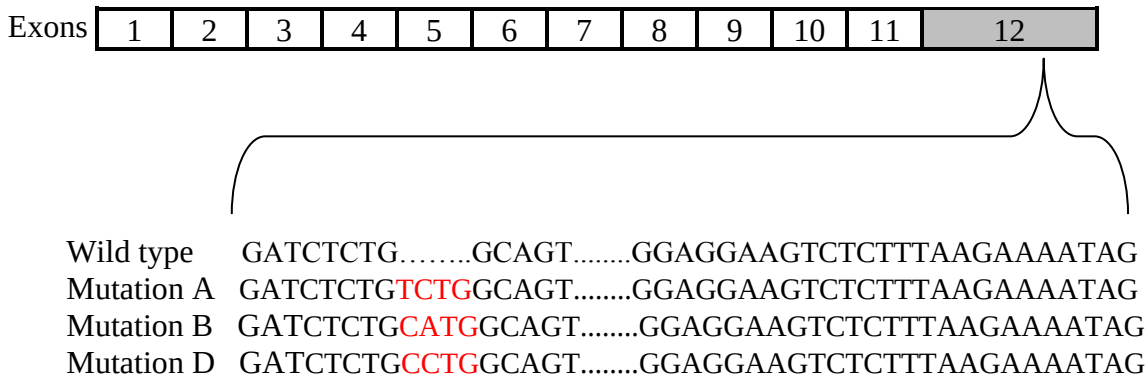

Figure 1.4 The common mutation of NPM1 type A, B, and D in exon 12.

NPM1 gene mutations occur more frequently in females than males and are strongly associated with older age, (Thiede *et al.*, 2006; Gale *et al.*, 2008) higher counts of WBC, blast cells, platelets and M4-M5 FAB subtypes (Thiede *et al.*, 2006; Boonthimat, Thongnoppakhun & Auewarakul, 2008; Chen, Rassidakis & Medeiros, 2006; Falini *et al.*, 2005). Previous studies showed that the absence of CD34 and CD133 hematopoietic stem cell markers characterise mutated *NPM1* in AML. The presence of CD33 and CD13 myeloid antigen markers (Döhner *et al.*, 2005). Mutations of *NPM1* are frequently associated with normal karyotype and *FLT3-ITD* mutations (Suzuki *et al.*, 2005; Schnittger *et al.*, 2005a; Döhner *et al.*, 2005; Falini *et al.*, 2005). whereas an *NPM1* mutation and no *FLT3-ITD* (*NPM1*$^{+}$/*FLT3*-ITD^{-}) is considered as a favourable factor for patient prognosis and *NPM1*$^{-}$/*FLT3*-ITD^{+} is associated with a poor prognosis, while AML with no mutation *NPM*$^{-}$/*FLT3*-ITD^{-} and *NPM1*$^{+}$/*FLT3*-ITD^{+} is related with intermediate prognosis (Thiede *et al.*, 2006; Döhner *et al.*, 2005; Gale *et al.*, 2008; Schnittger *et al.*, 2005a).

All the *NPM1* mutations lead to a distinct sequence in the C-terminus of the *NPM* protein. All the mutation shows at least one loss of the tryptophan residues (w) at amino acid positions 288 and 290 at the C-terminal domain, which leads to mutations in the frameshift and creates the last five common amino acid VSLRK residues. The C-terminal tryptophans 288 and 290 of the *NPM* protein are essential for the *NPM* nucleolar localisation. This alteration disrupts the *NPM* nucleolar localisation signal (NoLS), and They also produce an additional leucine-rich (NES) motif at the C-terminal. Both of them play essential roles in disturbing the normal nucleocytoplasmic transportation of the *NPM* protein, leading to abnormal cytoplasmic accumulation of the mutant NPM proteins (Figure 1.6) (Chen, Rassidakis & Medeiros, 2006; Falini *et al.*, 2007)

Falini *et al.* (2007) demonstrated that the cytoplasmic aggregation of NPM proteins is NES-dependent. The *NPM* protein of wild type includes only one NES motif at the amino acid position 92-104 at the N-terminus (Figure 1.5)(Falini *et al.*, 2006). It binds export receptors to the chromosomal region maintenance 1 (CRM1) and facilitates downstream nuclear export of wild NPM protein for normal functions, However, and the mutant *NPM* protein contains a new creation of the NES motif in the C – terminal domain, in which the accessibility of the *NPM* mutants to the CRM1 receptor can be reinforced. This would lead to more efficient nuclear export activity than nuclear import and result in an uncontrolled accumulation of NPM protein by cytoplasm (Falini *et al.*, 2007).

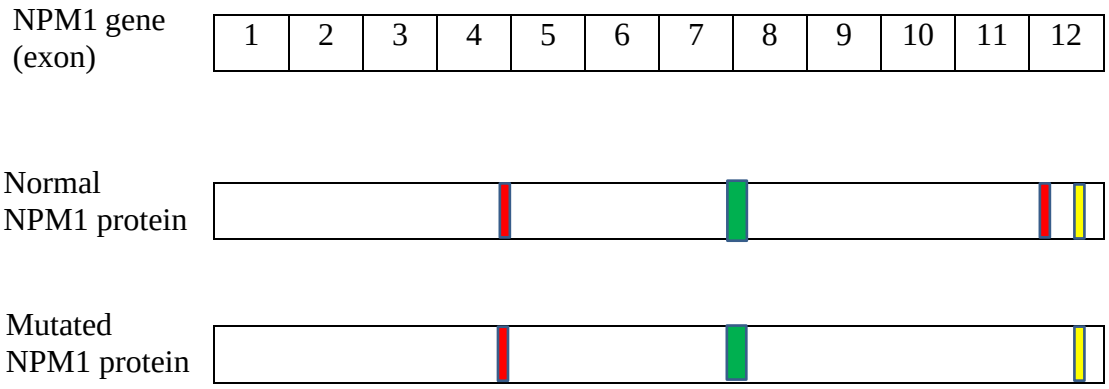

Figure 1.5 *NPM1* gene diagram, normal *NPM* protein and mutant *NPM* protein (Chen, Rassidakis & Medeiros, 2006)

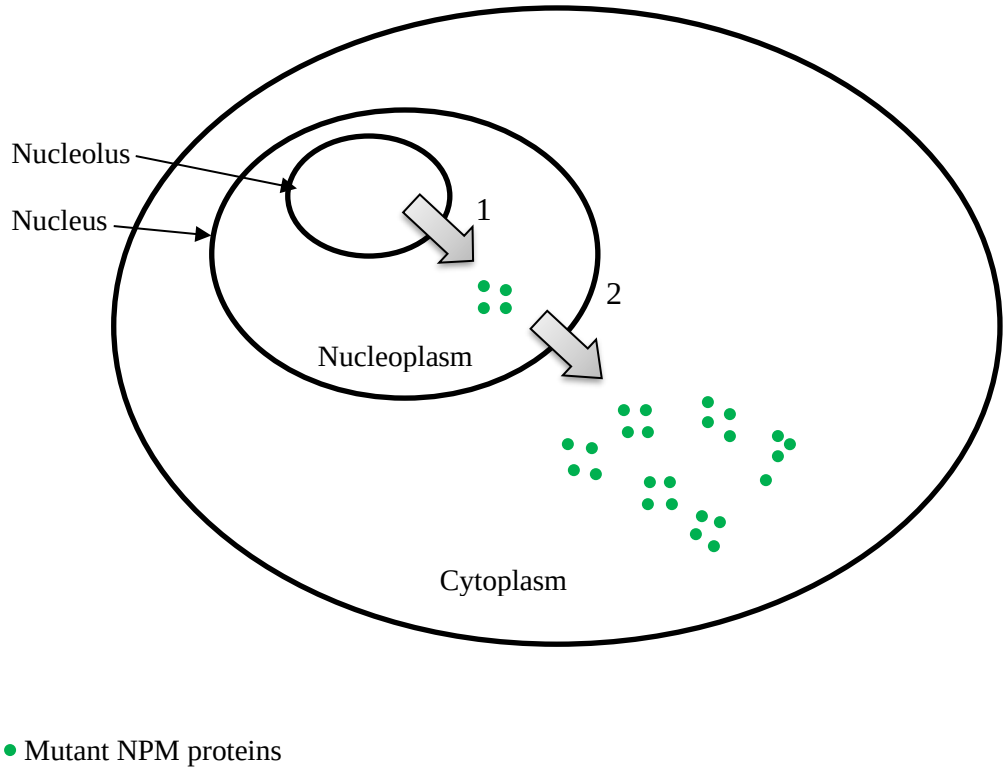

Figure 1.6 Diagram showing the accumulation of mutant *NPM* protein in the cytoplasm.

1: The interruption of the nucleolar localisation signal (NoLS) as result Displacement of the mutant *NPM* proteins from the nucleolus to the nucleoplasm. 2: The addition of a nuclear export signal (NES) in the C-terminal of mutated *NPM* proteins as a result, the displacement of mutated *NPM* proteins from the nucleoplasm to the cytoplasm.

1.3.3 *CEBPA* gene

CEBPA or CCAAT/ enhancer-binding protein alpha (*C/EBPα*) gene is located on the lengthy arm of chromosome 19 at band q13.1. *CEBPA* is an intronless gene encoding a member of the transcription factor of the leucine zipper (*CEBPα*). It acts as a cell proliferation inhibitor and a tumour suppressant (Hendricks-Taylor *et al.*, 1992). Transcription factors play an essential role in the haemopoiesis process and are up or downregulated at specific points during the differentiation required to switch on or off their target genes. CCAAT enhancer-binding protein alpha (*C/EBPα*) is one such transcription factor with a pivotal role in myelopoiesis (Pabst *et al.*, 2001b).

The protein of *(C/EBPα)* contains several domains characteristic of this family, as shown in (figure 1.7). These include two domains of transactivation (TAD1 and TAD2) in the N-terminal part and the C-terminal region, and there is two domain DNA binding domain (DBD), leucine zipper domain (LZD). The roles of C/EBPα as a dimer, with LZD, mediated dimerisation with other *CEBP* family members. This domain contains an α-helical structure that interacts with the α-helix to form a coiled-coil structure in the binding partner. *C/EBPα* can form a homodimer or heterodimer with other members of the family, such as *C/EBPδ* and *C/EBPβ*, because of the conserved nature of the bZIP domain. Dimerisation is essential for DNA binding, which is facilitated by the DBD. It occurs at sequence-specific sites in the major groove of DNA, usually found in target gene promoters or enhancers. Transactivation domains can activate the transcription of target genes once they are bound to DNA. Two main isoforms of *C/EBPα* are present, which is full-length (p42) kDa protein and a truncated (p30) 30 kDa protein. The p30 isoforms are translated from an internal starting site in the mRNA, and the protein lacks at N-terminus the first 119 amino

acids, including the TAD1. The protein p30 has been found to have a lower potential for transcriptional activation than the p42 protein (Pabst *et al.*, 2001b).

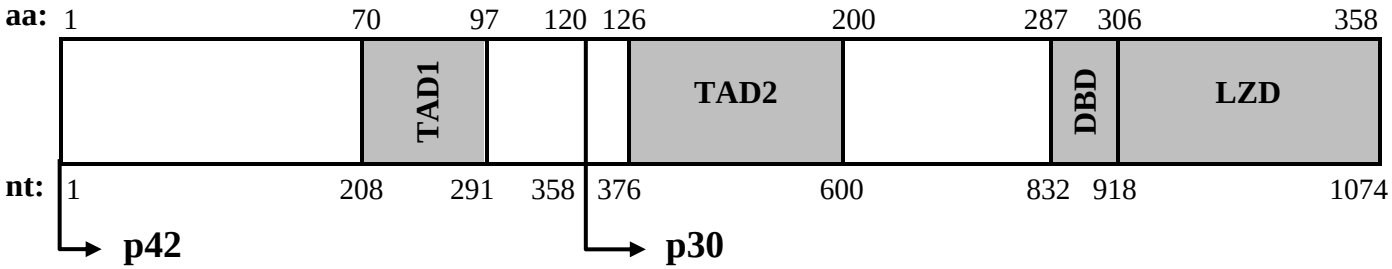

Figure 1.7 Functional domains of C/EBPα and Translational start sites for the two main isoforms p42 and p30, (aa) amino acid and (nt) nucleotide numbering are indicated (Adapted from (Mueller & Pabst, 2006)).

Mutation in *CEBPA* has been reported in approximately 6-20% of all AML patients (Pabst *et al.*, 2001b; Jennifer *et al.*, 2003; Frohling *et al.*, 2004; Pabst *et al.*, 2008; Said *et al.*, 2016; Ng *et al.*, 2018; Awad *et al.*, 2012). Mutations can occur across the entire coding region. In general, two common types of mutations have been identified. The first frameshift mutations are insertions or deletions in the N-terminus, resulting in premature termination of the (p42) full length 42- kDa. And increase isoform expression of p30, the isoform imbalance between p42 and p30, which interferes cell cycle arrest and differentiation. The second in-frame mutations, deletions or insertions in the DNA binding and leucine zipper domains at C-terminus, which are affecting to impair DNA-binding or homo- and heterodimerisation.

The mutation has been found across the entire *CEBPA* gene and can be either point mutations, small or large deletions or insertions. Most of the mutations are heterozygous, affecting only one allele. However, a small number of cases have been reported as homozygous (0.2%) (Frohling *et al.*, 2004; Jennifer *et al.*, 2003; Pabst *et al.*, 2009; Hou *et al.*, 2009). AML Patients with *CEBPA* mutations can be divided into two subgroups, which include single mutation *CEBPA* (*CEBPA*^{sm}) and double mutation *CEBPA* (*CEBPA*^{dm}). Single mutations of *CEBPA* (*CEBPA*^{sm}) may occur either in N-terminal or C-terminal of the gene, and double mutation of *CEBPA* (*CEBPA*^{dm}) involve both N-terminus or C-terminus of the gene (Ahn *et al.*, 2016; Li *et al.*, 2015; Wouters *et al.*, 2009).

It is well known that single nucleotide polymorphisms (SNPs) occur in the *CEBPA* gene. There was some most common polymorphism reported like 836T>G polymorphism has been detected (17%) in AML patients, which codes for the silent T230T. Moreover, there are numerous reports of rarer point mutations also predicted

to cause silent amino acid changes (Pabst *et al.*, 2001b; Frohling *et al.*, 2004; Preudhomme *et al.*, 2002). However, the most frequent mutation duplication of 6 base pairs in position (584_589dup) proline-rich region of TAD2 is expected to cause H195_P196dup (also termed HP196_197ins or P194_H195dup) (Frohling *et al.*, 2004; Fasan *et al.*, 2014). this was also found (39%) in healthy volunteers and AML patients (Lin *et al.*, 2005). The other study detected this change (8%) in the normal blood sample and (4%) in AML. It was found that H195_P196dup AML gene expression profiles did not cluster with those of other *CEBPA*-mutant cases (Wouters *et al.*, 2007). As a result, they conclude that HP196-197ins is a common polymorphism instead of mutation (Wouters *et al.*, 2007).