**Al-Rafidain J Med Sci. 2025;9(1):134-142.**
**DOI:** https://doi.org/10.54133/ajms.v9i1.2112

AJMS

Online ISSN (2789-3219)

**Research Article**

# Comprehensive Molecular Profiling of the 3′UTR Region of the *CEBPA* Gene in Iraqi Patients with Acute Myeloid Leukemia Reveals Novel Regulatory Variants

Noor Hussein Ali*, Nuha Joseph Kandala
*Department of Biotechnology, College of Science, University of Baghdad, Baghdad, Iraq*
Received: 28 May 2025; Revised: 12 July 2025; Accepted: 17 July 2025

## ABSTRACT

***Background***: Acute myeloid leukemia (AML) is a genetically heterogeneous leukemia characterized by abnormal myeloid blast accumulation, disrupting normal hematopoiesis and leading to rapid progression. ***Objective***: To investigate SNPs within the 3'UTR of the CCAAT/enhancer-binding protein alpha (*CEBPA*) gene and its association with AML in Iraqi patients. ***Methods***: The study was carried out on 120 AML patients classified into newly diagnosed, induction chemotherapy, and consolidation chemotherapy stages (40 each), and 40 individuals as a control group. Genomic DNA was extracted from AML patients and controls, followed by PCR amplification and Sanger sequencing of the 3'UTR region of the *CEBPA* gene. The AML patients were characterized by age, sex, FMS-like tyrosine kinase 3 internal tandem duplication (*FLT3-ITD*), Nucleophosmin1 (*NPM1*) mutations, the French-American-British classification (FAB), and the World Health Organization (WHO). ***Results***: The results revealed significant age differences among AML subgroups and notable hematological abnormalities, including reduced hemoglobin and platelet levels. According to the WHO classification, *PML-RARA* emerged as the most frequent fusion transcript. Based on FAB classification, M3 was the most common, followed by M4 and M0. The *NPM1* mutations were more common than *FLT3-ITD*. The sequencing of the CEBPA 3′UTR region identified 83 variants, including 46 novels, 14 new forms of known SNPs, and 23 registered SNPs, reflecting substantial regulatory heterogeneity in this non-coding region. ***Conclusions***: The *CEBPA* 3′UTR mutations reveal considerable genetic diversity among Iraqi AML patients, suggesting a potential regulatory role.

***Keywords***: AML patient, *CEBPA* gene, Novel SNPs, Sanger sequencing.

**تحليل جزيئي شامل لمنطقة 3′UTR من جين *CEBPA* لدى مرضى سرطان الدم النقياني الحاد في العراق يكشف عن طفرات تنظيمية جديدة**

**الخلاصة**

**الخلفية:** ابيضاض الدم النقياني الحاد (AML) مرض غير متجانس وراثيًا يتميز بتراكم غير طبيعي للأورام النقوية، مما يُعطل عملية تكوين الدم الطبيعية ويؤدي إلى تطور سريع للمرض. **الهدف:** دراسة تعدد أشكال النوكليوتيدات المفردة (SNPs) ضمن منطقة 3′UTR لجين CCAAT/بروتين ألفا الرابط للمعزز (*CEBPA*) وارتباطها بابيضاض الدم النقياني الحاد لدى المرضى العراقيين. **الطرائق:** أُجريت الدراسة على 120 مريضًا مصابًا ب AML مُصنّفين إلى (40 مريضًا مُشخّصًا حديثًا، و40 في مرحلة العلاج الكيميائي التحريضي، و40 في مرحلة العلاج الكيميائي التعزيزي)، و40 فردًا كمجموعة تحكم. استُخلص الحمض النووي من المرضى ومجموعة التحكم، تبعه التضخيم بتقنية تفاعل البوليميراز المتسلسل وتسلسل سانجر لمنطقة 3′UTR من جين *CEBPA*. تم توصيف مرضى AML حسب العمر، والجنس، والتضاعف الداخلي للتيروزين كيناز 3 الشبيه بمتلازمة FMS (*FLT3-ITD*)، وطفرات النوكليوفوسمين 1 (*NPM1*)، والتصنيف الفرنسي-الأمريكي-البريطاني، ومنظمة الصحة العالمية. **النتائج:** كشفت الدراسة عن فروقات عمرية ملحوظة بين المجموعات الفرعية لمرضى AML، بالإضافة إلى اضطرابات دموية بارزة شملت انخفاض مستويات الهيموغلوبين والصفائح الدموية. ووفقًا لتصنيف منظمة الصحة العالمية، برز اندماج *PML-RARA* كأكثر التحولات الجينية شيوعًا. أما حسب تصنيف FAB، فكان النمط M3 هو الأكثر شيوعًا، يليه M4 ثم M0. كما وُجد أن طفرات *NPM1* كانت أكثر شيوعًا من طفرات *FLT3-ITD*. وقد أظهر تسلسل منطقة 3′UTR من جين *CEBPA* وجود 83 تغيرًا وراثيًا، شملت 46 طفرة جديدة، و14 شكلًا جديدًا من SNPs معروفة، و23 SNP مسجلًا سابقًا، مما يعكس درجة عالية من التنوع الجيني التنظيمي في هذه المنطقة غير المُشفّرة. **الاستنتاجات:** تكشف الطفرات عن تنوع جيني كبير بين مرضى AML الحاد، مما يشير إلى دور تنظيمي محتمل.

* ***Corresponding author***: Noor H. Ali, Department of Biotechnology, College of Science, University of Baghdad, Baghdad, Iraq; Email: noor.ali@sc.uobaghdad.edu.iq
***Article citation***: Ali NH, Kandala NJ. Comprehensive Molecular Profiling of the 3′UTR Region of the *CEBPA* Gene in Iraqi Patients with Acute Myeloid Leukemia Reveals Novel Regulatory Variants. *Al-Rafidain J Med Sci*. 2025;9(1):134-142. doi: https://doi.org/10.54133/ajms.v9i1.2112


## INTRODUCTION

Acute myeloid leukemia (AML) is a genetically diverse hematologic malignancy marked by clonal proliferation of myeloid precursors in the bone marrow, leading to impaired hematopoiesis and bone marrow failure [1]. Its molecular pathogenesis involves mutations in transcription factors, tumor suppressor genes, and signaling pathways regulating differentiation and proliferation [2,3]. AML represents ~80% of adult acute leukemia cases, and its prognosis depends largely on underlying cytogenetic and molecular features [4]. The

World Health Organization (WHO) classification of AML incorporates genetic alterations to refine diagnosis and prognosis, highlighting the significance of specific mutations and polymorphisms in disease progression [5,6]. While the French-American-British (FAB) classification categorizes AML into subtypes (M0-M7) based on blast morphology, molecular studies have identified recurrent genetic mutations [7]. Genetic mutations play a crucial role in AML pathogenesis, with recurrent alterations detected in key genes involved in hematopoietic differentiation, proliferation, and apoptosis. Among these, mutations in FMS-like tyrosine kinase 3 (*FLT3*), 30-35%; Nucleophosmin 1 (*NPM1*), 25-30%; and CCAAT/enhancer-binding protein alpha (*CEBPA*), 10-15%, are frequently observed [8,9]. The *CEBPA* gene, located at chromosome 19q13.1, is a GC-rich (77%) single-exon gene encoding a 358-amino-acid transcription factor critical for myeloid differentiation [10]. It exists in two isoforms: full-length p42 and truncated p30, both derived from the same gene but translated from 2 distinct AUG translational start sites [11]. This protein is a transcription factor, meaning it binds to specific DNA sequences to help regulate the expression of certain genes [12]. Mutations in this gene can disrupt normal cellular differentiation and promote leukemogenesis. In AML, *CEBPA* mutations are considered a significant marker for prognosis, and their presence can influence treatment response and survival outcomes. Some studies have linked *CEBPA* mutations with a better prognosis, especially in patients with isolated mutations, but their clinical significance is often context-dependent, requiring consideration of co-occurring mutations like *FLT3* and *NPM* [13,14]. The 3′ untranslated region (3′UTR) of *CEBPA* plays a vital role in regulating gene expression through control of mRNA stability and translation. Single Nucleotide Polymorphism (SNP) within this region disrupts miRNA binding or alters RNA structure, potentially affecting *CEBPA* function [15]. This study aims to investigate the presence of SNPs within the 3'UTR of *CEBPA* and assess their possible association with AML in Iraqi patients. Since the genetic variations in this gene have not been previously studied in Iraq, this research aims to fill that knowledge gap.

## METHODS

### *Study design and sampling*

In this case-control study, one hundred twenty blood samples were collected from AML patients aged 14-80 years at the National Center for Hematology in Baghdad, Iraq. The samples were evenly distributed between 60 males and 60 females. The patients were classified into three groups: 40 newly diagnosed (NDX) cases, 40 instances undergoing induction chemotherapy (ICT), and 40 cases in the consolidation chemotherapy stage (CCT). Additionally, 40 control samples (20 males and 20 females) from healthy individuals aged 14-80 years were obtained from the National Blood Bank in Baghdad, Iraq.

### *Sample collection*

This study's sample collection period extended from November 11, 2023, to July 10, 2024. AML diagnosis and classification were performed according to the FAB system (M0–M7) and the WHO guidelines for promyelocytic leukemia-retinoic acid receptor alpha (*PML-RARA*), Runt-related transcription factor 1–RUNX1 partner transcriptional co-repressor 1 (*RUNX1-RUNX1T1*), and core-binding factor subunit beta–myosin heavy chain 11 (*CBFB-MYH11*), which included clinical, morphological, biochemical, cytogenetic, and gene mutation analyses. Clinical data, including age, sex, white blood cell (WBC) count, platelet count (PLT), disease status, and the presence of *FLT3-ITD* and *NPM1* mutations, were obtained from patient records. Patients with comorbidities such as diabetes, infectious diseases, Down syndrome, and other chronic illnesses were excluded from the study.

### *Molecular analysis and variant detection*

Peripheral blood samples were collected in EDTA tubes and processed for DNA extraction using the FavorPrep™ Genomic DNA Kit (FAVORGEN, Taiwan) following the manufacturer's instructions. DNA purity was assessed via NanoDrop spectrophotometry (Thermo Fisher Scientific, USA). The primers used in this investigation were designed based on their reference sequences from the National Center for Biotechnology Information (NCBI) database and obtained from Macrogen Company. The forward primer sequence was 5′-CGGCAACTCTAGTATTTAGGA-3′, and the reverse primer was 5′-ATACAAGTGTTGATATCGGCT-3′. This study specifically designed these primers to amplify a 706 bp fragment within the 3′UTR of the *CEBPA* gene. All PCR procedures were completed following the manufacturer's instructions. A final volume of 25 µl reaction contained 12.5 µl of OneTaq NEB® master mix (NEB®, USA), 4 µl of DNA sample, 1 µl of 10 pmol/µl from each primer, and 6.5 µl of nuclease-free water. Amplification was performed using the following thermal cycling conditions: an initial denaturation step at 95 °C for 5 minutes, followed by 35 cycles of denaturation at 95 °C for 30 seconds, annealing at 48 °C for 30 seconds, and extension at 72 °C for 30 seconds. A final extension was carried out at 72°C for 5 minutes. The amplified products were visualized by 1.5% agarose gel electrophoresis. The annealing temperature that resulted in a clear and sharp band was selected as optimal for the target gene. The PCR products were subsequently dispatched to Macrogen, Korea, for sequencing via the Sanger method. The

resulting sequencing data, provided in FASTA format, were analyzed using Geneious Prime software. Sequence alignment was performed against the reference sequence (RefSeq) of the *CEBPA* gene, accession number 1050, to accurately identify and characterize potential SNPs and mutations within the gene.

### *Ethical considerations*

The study protocol was approved by the ethical committee of the Biotechnology Department, College of Science, University of Baghdad, according to the reference number CSEC/1123/0070.

### Statistical analysis

Statistical analysis was performed using several statistical packages for Windows, SPSS version 20 (IBM Corp., Armonk, N.Y., USA). The probability was calculated using a Mann-Whitney or Kruskal-Wallis test for non-parametric data. Chi-square was used to calculate the probability for categorical data. Differences were considered significant at $P \leq 0.05$. Bioinformatics tools were used in this study. Geneious Prime was employed to analyze Sanger sequencing data and its alignment with NCBI reference sequences.

## RESULTS

The median age of participants in the control group was 34.5 years (range: 14–80), while it was 56 years (range: 14–80) in the newly diagnosed AML group, 32 years (range: 14–70) in the induction group, and 34.5 years (range: 18–61) in the consolidation group. Age was analyzed as a continuous variable and categorized into defined age groups. The Kruskal-Wallis test showed a statistically significant difference in age distribution among the groups ($p$= 0.0003), as shown in Table 1.

**Table 1:** Distribution of study groups according to age (n= 40 in each group)

| Groups | Age (year) | Age group (year) | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | 10-19 | 20-29 | 30-39 | 40-49 | 50-59 | 60-69 | ≥ 70 |
| Control | 34.5(14-80) | 4(10) | 9(22.5) | 11(27.5) | 8(20) | 6(15) | 2(5) | 0(0.0) |
| Newly diagnosed | 56(14-80) | 5(12.5) | 5 (12.5) | 2(5) | 4(10) | 13(32.5) | 6(15) | 5(12.5) |
| Induction | 32 (14-70) | 3(7.5) | 5(12.5) | 5(12.5) | 6(15) | 8(20) | 7(17.5) | 6(15) |
| Consolidation | 34.5(18-61) | 7(17.5) | 11(27.5) | 6(15) | 5(12.5) | 8(20) | 2(5) | 1(2.5) |
| *p*-value | 0.0003 | | | | | | | |

Data are presented as median and range for continuous variables (age) and frequency and percentage for age groups. The Kruskal-Wallis test was used, and a $p$-value < 0.05 was considered statistically significant.

Regarding sex distribution, no statistically significant difference was observed among the study groups ($p$ = 0.361), indicating that the distribution of males and females was comparable between groups at the statistical level, as illustrated in Table 2.

**Table 2:** Distribution of study groups according to sex (n=40 in each group)

| Groups | Sex | |
|---|---|---|
| | Male | Female |
| Control | 20(50) | 20(50) |
| Newly diagnosed | 24(60) | 16(40) |
| Induction | 16(40) | 24(60) |
| Consolidation | 20(50) | 20(50) |
| *p*-value | 0.361 | |

Data are presented as numbers and percentages. Chi-square ($\chi^2$) test was used; $p$< 0.05 was considered statistically significant.

The clinical characteristics of the study participants were assessed based on hematological parameters, including WBC count, Hb concentration, and PLT count, as shown in Table 3.

**Table 3:** Comparison of hematological parameters between AML patients and the control group

| Groups | Median (IQR) | | |
|---|---|---|---|
| | WBC count (×10$^9$\L) | Hb (g/dl) | PLT count (×10$^9$/L) |
| Control (n= 40) | 7.19(5.6-8.4) | 13.7(12.5-14.5) | 252(212.2-293.5) |
| Patients (n= 120) | 5.7(2.2–30.3) | 8.07(7–9.6) | 74.5(28.2-108) |
| *p*-value | < 0.001 | < 0.001 | < 0.001 |

Data are presented as median and interquartile range (IQR). Mann-Whitney U test was used; $p$< 0.05 was considered statistically significant.

The AML patient group demonstrated significantly altered values compared to the control group. The median WBC count in patients was lower than that of the control group. Hemoglobin levels showed a marked reduction in patients compared to controls. Likewise, platelet counts were substantially decreased in patients relative to controls. These differences were all statistically significant ($p < 0.001$). Based on the FAB classification system, six AML subtypes were identified among the patients: M0 6 (5%), M1 1 (0.83%), M3 20 (16.67%), M4 7 (5.83%),

M5 2 (1.67%), and M7 2 (1.67%). In many cases, 82 (68.33%) were categorized as unclassified (UC). M3 was the most common, followed by M4 and M0. In contrast, M1, M5, and M7 were observed at lower frequencies. Among the WHO classification-related genetic abnormalities, the *PML-RARA* fusion was detected in 12 out of 120 AML patients (10%). The *RUNX1-RUNX1T1* fusion gene was observed in 3 patients (2.5%), while the *CBFB-MYH11* rearrangement was found in only one patient (0.38%). Most patients (86.67%) did not exhibit these three WHO-defining mutations. Regarding molecular mutation, *FLT3-ITD* mutations were identified in 29 cases (24.2%), while *NPM1* mutations were detected in 42 cases (35.0%). In contrast, 91 patients (75.8%) were negative for *FLT3-ITD* mutations, and 78 patients (65.0%) did not carry *NPM1* mutations. The PCR reactions were conducted on control and AML patient samples to amplify the 3'UTR of the *CEBPA* gene. The primer set *CEBPA*(P) successfully amplified the target region in all tested samples, as illustrated in Figure 1.

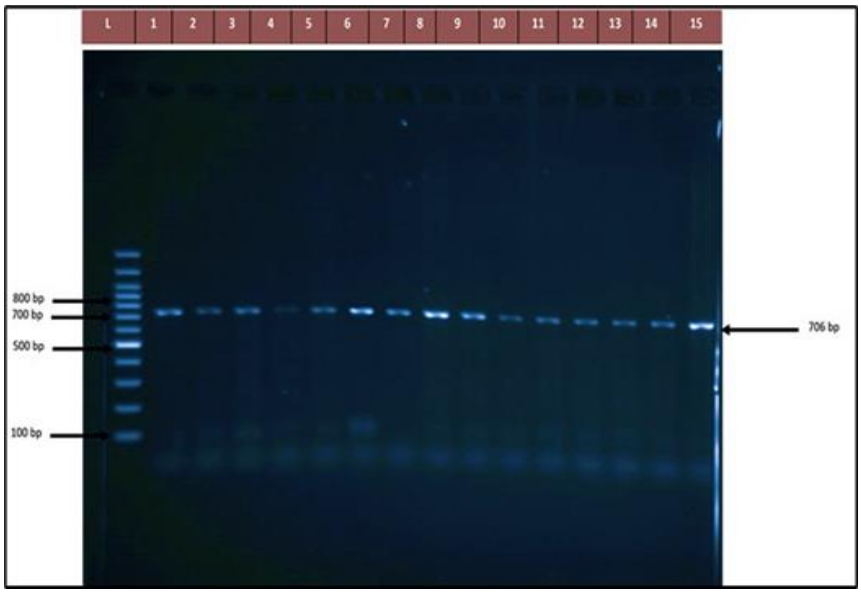


**Figure 1:** Agarose gel electrophoresis of the PCR product of the *CEBPA* fragment with a molecular base of 706 bp for the patient and the control. L: 100 -1500bp DNA ladder,1-8 were patients,9-15 were controls. Fragments were fractionated by electrophoresis on a 1.5 % agarose gel and visualized by Safe Red dye.

The sequencing results focused on variants in the 3' UTR, which may be associated with AML. Identifying 3' UTR variants is significant, as they may affect gene expression regulation and contribute to AML pathogenesis. Among the study cohort, genomic variants (including SNPs and mutations) were detected in 54 out of 66 AML patients, compared to only 7 out of 40 individuals in the control group. This difference was statistically significant ($p$= 0.0019), suggesting a potential association between the presence of genetic alterations and AML, as shown in Table 4.

**Table 4:** Comparison of the presence of genomic variants (SNPs/mutations) between AML patients and controls

| Characteristics | | Number of subjects | | *p*-value |
|---|---|---|---|---|
| | | Control | Patients | |
| Genomic variant (SNP/mutation) | Yes | 7 | 54 | 0.0019 |
| | No | 33 | 66 | |

Data are presented as frequencies (number of subjects). Chi-square test was used; $p< 0.05$ was considered statistically significant.

Subsequently, 83 distinct genetic variants were identified within the 3′ untranslated region (3′UTR) of the *CEBPA* gene across all study groups, including newly diagnosed patients, those receiving induction or consolidation therapy, and the control sample. These variants comprised single-nucleotide variants (SNVs), short insertions, deletions, and complex indels. Among the detected variants, 46 were classified as novel, not previously reported in public databases such as dbSNP or ClinVar, as summarized in Table 5.

**Table 5:** Novel genomic variants (SNPs/mutations) identified in the 3′UTR of the *CEBPA* gene among AML patients

| Genomic variance (SNP/mutation) | Allele variance | Variation type | Genomic location | Number of patients with SNP/mutation | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | Control | Newly diagnosed | Induction | Consolidation | Total |
| Novel, 33300944 | C>T | SNV | chr19: 33300944 | 1 | 0 | 1 | 2 | 4 |
| Novel, 33300396 | G>C | SNV | chr19: 33300396 | 0 | 0 | 2 | 2 | 4 |
| Novel, 33300386 | T>G | SNV | chr19: 33300386 | 0 | 2 | 1 | 1 | 4 |
| Novel, 33300973 | A>T | SNV | chr19: 33300973 | 0 | 1 | 2 | 0 | 3 |
| Novel, 33300408 | A>C | SNV | chr19: 33300408 | 0 | 2 | 1 | 0 | 3 |
| Novel, 33301012 | A>T | SNV | chr19: 33301012 | 0 | 2 | 0 | 0 | 2 |
| Novel, 33300992 | A>T | SNV | chr19: 33300992 | 0 | 1 | 1 | 0 | 2 |
| Novel, 33300989 | Insertion, T | Insertion | chr19: 33300989 | 0 | 0 | 0 | 2 | 2 |
| Novel, 33300386 | C>A | SNV | chr19: 33300387 | 0 | 0 | 2 | 0 | 2 |
| Novel, 33300989 | A>T | SNV | chr19: 33300989 | 0 | 0 | 0 | 1 | 1 |
| Novel, 33300980 | A>T | SNV | chr19: 33300980 | 0 | 0 | 1 | 0 | 1 |
| Novel, 33300974 | A>T | SNV | chr19: 33300974 | 0 | 0 | 1 | 0 | 1 |
| Novel, 33300972 | A>T | SNV | chr19: 33300972 | 0 | 0 | 1 | 0 | 1 |
| Novel, 33300961 | A>T | SNV | chr19: 33300961 | 0 | 0 | 1 | 0 | 1 |
| Novel, 33300954 | C>T | SNV | chr19: 33300954 | 0 | 0 | 1 | 0 | 1 |
| Novel, 33300948 | C>T | SNV | chr19: 33300948 | 0 | 0 | 0 | 1 | 1 |
| Novel, 33300947 | C>T | SNV | chr19: 33300947 | 0 | 0 | 0 | 1 | 1 |
| Novel, 33300931 | C>T | SNV | chr19: 33300931 | 0 | 0 | 0 | 1 | 1 |
| Novel, 33300931 | C>G | SNV | chr19: 33300931 | 0 | 0 | 1 | 0 | 1 |
| Novel, 33300928 | Insertion, T | Insertion | chr19: 33300928 | 0 | 0 | 0 | 1 | 1 |
| Novel, 33300917 within rs1016346090 | C>G | SNV | chr19: 33300917 | 0 | 0 | 1 | 0 | 1 |
| Novel, 33300903 | C>A | SNV | chr19: 33300903 | 0 | 0 | 0 | 1 | 1 |
| Novel, 33300899 | C>T | SNV | chr19: 33300899 | 0 | 0 | 0 | 1 | 1 |
| Novel, 33300897 | G>T | SNV | chr19: 33300897 | 0 | 0 | 1 | 0 | 1 |
| Novel, 33300854 | A>T | SNV | chr19: 33300854 | 0 | 0 | 0 | 1 | 1 |
| Novel, 33300783 | Insertion, A | Insertion | chr19: 33300783 | 0 | 0 | 0 | 1 | 1 |
| Novel, 33300780 | A>G | SNV | chr19: 33300780 | 0 | 0 | 1 | 0 | 1 |
| Novel, 33300655 | T>C | SNV | chr19: 33300655 | 0 | 0 | 1 | 0 | 1 |
| Novel, 33300643 | C>T | SNV | chr19: 33300643 | 0 | 0 | 0 | 1 | 1 |
| Novel, 33300629 | C>G | SNV | chr19: 33300629 | 0 | 0 | 1 | 0 | 1 |
| Novel, 33300623 | G>C | SNV | chr19: 33300623 | 0 | 0 | 0 | 1 | 1 |
| Novel, 33300613 | G>T | SNV | chr19: 33300613 | 0 | 0 | 1 | 0 | 1 |
| Novel, 33300609 | C>T | SNV | chr19: 33300609 | 1 | 0 | 0 | 0 | 1 |
| Novel, 33300600 | G>C | SNV | chr19: 33300600 | 0 | 0 | 1 | 0 | 1 |
| Novel, 33300575 | A>T | SNV | chr19: 33300575 | 0 | 0 | 1 | 0 | 1 |
| Novel, 33300558 | C>A | SNV | chr19: 33300558 | 1 | 0 | 0 | 0 | 1 |
| Novel, 33300558 | C>T | SNV | chr19: 33300558 | 0 | 0 | 0 | 1 | 1 |
| Novel, 33300558 | C>G | SNV | chr19: 33300558 | 0 | 0 | 1 | 0 | 1 |
| Novel, 33300533 | C>G | SNV | chr19: 33300533 | 1 | 0 | 0 | 0 | 1 |
| Novel, 33300406 | T>A | Insertion | chr19: 33300406 | 0 | 0 | 1 | 0 | 1 |
| Novel, 33300393 | C>G | SNV | chr19: 33300393 | 0 | 0 | 1 | 0 | 1 |
| Novel, 33300387 | G>T | SNV | chr19: 33300387 | 0 | 0 | 1 | 0 | 1 |
| Novel, 33300386 | G>T | SNV | chr19: 33300386 | 0 | 0 | 1 | 0 | 1 |
| Novel, 33300386 | A>C | SNV | chr19: 33300386 | 0 | 0 | 1 | 0 | 1 |
| Novel, 33300381 | insertion of A | insertion | chr19: 33300381 | 0 | 0 | 1 | 0 | 1 |
| Novel, 33300969 | G>T | SNV | chr19: 33300969 | 0 | 0 | 1 | 0 | 1 |

Variants were identified and annotated using Geneious Prime

These novel variants were individually distributed across the cohort, with most variants identified in different single individuals without recurrence. They included various types of substitutions, such as transitions and transversions (e.g., C>T, G>A, A>C), as well as insertions and deletions. An additional 14 variants, presented in Table 6, were recognized as new forms of previously reported SNPs. These were located at known SNP positions but involved different nucleotide substitutions than those recorded in reference databases. For example, a substitution of G>A was observed at rs1967123357 instead of the registered G>T. Each new form was observed within the dataset in no more than two cases.

**Table 6:** New forms of reported SNPs identified in the 3′UTR of the *CEBPA* gene among AML patients

| Genomic variance (SNP/mutation) | Allele variance | Variation type | Genomic location | Number of patients with (SNP/mutation) | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | Control | Newly diagnosed | Induction | Consolidation | Total |
| New form of rs1967123357, G >T instead of G >A | G>A | SNV | chr19:33300385 | 0 | 2 | 5 | 0 | 7 |
| New form of rs2145257064, G>T instead of G >A | G>A | SNV | chr19:33300908 | 0 | 0 | 1 | 2 | 3 |
| deletion of T at position 33301025 of rs1303853358 | T>C | SNV | chr19:33301025 | 0 | 0 | 0 | 2 | 2 |
| C > T in rs1016346090 instead of an Indel | delGCCACGGGCCTGCT(C)TCCT | Indel | chr19:33300902-33300929 | 0 | 0 | 1 | 0 | 1 |
| New form of rs957209037, C >T instead of C > G | C>G | SNV | chr19:33300923 | 0 | 0 | 0 | 1 | 1 |
| New form of rs2513326585, A > T instead of A > C | A>C | SNV | chr19:33300973 | 0 | 1 | 0 | 0 | 1 |
| New form of rs2513325909, C >T | C>A | SNV | chr19:33300618 | 1 | 0 | 0 | 0 | 1 |
| New form of rs2513325854, C >T | C>G | SNV | chr19:33300568 | 1 | 0 | 0 | 0 | 1 |
| New form of rs1967143107, G >T instead of G >A | G>A | SNV | chr19:33300958 | 0 | 0 | 1 | 0 | 1 |
| New form of rs1470470253, G >T | G>A | SNV | chr19:33300723 | 0 | 0 | 0 | 1 | 1 |
| New form of rs1417869963, C >G instead of C >T | C>T | SNV | chr19:33300640 | 0 | 0 | 1 | 0 | 1 |
| New form of rs1190219619, G>T instead of G >C | G>C | SNV | chr19:33300862 | 0 | 0 | 0 | 1 | 1 |
| New form of rs1005626115 | C>T | SNV | chr19:33300393 | 0 | 0 | 1 | 0 | 1 |
| New form of rs2513326459, G >T | G>A / G>C | SNV | chr19:3330091 | 0 | 0 | 0 | 1 | 1 |

Variants were identified and annotated using Geneious Prime.

The remaining 23 variants corresponded to registered SNPs listed in dbSNP, as shown in Table 7. The most frequent among them was rs41367646, observed in seven patients, mainly from the induction group. This SNP exhibited both C>G and C>T substitutions. Other recurring variants included rs707656 and rs10499969, each observed in up to five cases. The distribution of these registered SNPs varied among the AML subgroups. More than 50 of the 83 variants were detected in individual patients only once, emphasizing this cohort's predominance of non-recurrent, unique events. Despite being located in a non-coding region, this variant profile demonstrated substantial molecular diversity within the analyzed segment of *CEBPA*. Table 6 summarizes sequence alterations that occur at known SNP positions but differ from the reference alleles reported in public databases. Table 7 describes previously registered SNPs reported in dbSNP, detected in the 3′UTR of the *CEBPA* gene.

## DISCUSSION

The present study aimed to investigate the frequency and clinical relevance of *CEBPA* mutations in Iraqi AML patients. The findings provide new insights into the genetic heterogeneity of the *CEBPA* gene and its potential association with disease progression. A significant variation in age distribution was observed among study groups, with younger median ages in the induction and consolidation groups compared to the newly diagnosed and control groups. This reflects clinical practice, where younger patients are more likely to receive intensive chemotherapy. The findings are consistent with international data showing AML incidence increases with age [16,17], while younger patients are often selected for aggressive treatment [18]. In contrast, sex distribution did not differ significantly across the study groups . Equal male-to-female ratios were observed in the control and consolidation groups, while slight variations appeared in the newly diagnosed and induction groups.

xxxxxxxxxxx

**Table 7:** Registered SNPs identified in the 3′UTR of the *CEBPA* gene among patients with AML

| Genomic variance (SNP/mutation) | Allele variance | Variation type | Genomic location | Number of patients with SNP/mutation | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | Control | Newly diagnosed | Induction | Consolidation | Total |
| rs41367646 | C>G / C>T | SNV | chr19:33300977 | 0 | 5 | 4 | 0 | 9 |
| rs707656 | G>C | SNV | chr19:33300623 | 2 | 1 | 1 | 1 | 5 |
| rs1049969 | T>C | SNV | chr19:33300901 | 2 | 1 | 1 | 0 | 4 |
| rs941652275 | delCCC/ delC/ dupC | Indel | chr19:33300529-33300534 | 0 | 2 | 0 | 1 | 3 |
| rs867364161 | G>A | SNV | chr19:33300446 | 1 | 0 | 0 | 0 | 1 |
| rs771978125 | C>T | SNV | chr19:33300576 | 1 | 0 | 0 | 0 | 1 |
| rs549322947 | G>A / G>C | SNV | chr19:33300630 | 1 | 0 | 0 | 0 | 1 |
| rs2513325587 | T>G | SNV | chr19:33300384 | 0 | 0 | 0 | 1 | 1 |
| rs2145257028 | T>C | SNV | chr19:33300894 | 0 | 0 | 1 | 0 | 1 |
| rs1967142985 | C>A / C>T | SNV | chr19:33300956 | 0 | 1 | 0 | 0 | 1 |
| rs1967135055 | A>G | SNV | chr19:33300788 | 0 | 0 | 1 | 0 | 1 |
| rs1600019277 | T>G | SNV | chr19:33300722 | 0 | 0 | 1 | 0 | 1 |
| rs1490729652 | C>A | SNV | chr19:33300967 | 0 | 1 | 0 | 0 | 1 |
| rs1469326844 | G>A | SNV | chr19:33300388 | 0 | 1 | 0 | 0 | 1 |
| rs1460208064, G >A | G>A / G>T | SNV | chr19:33300416 | 0 | 0 | 0 | 1 | 1 |
| rs1443902201 | G>C | SNV | chr19:33300857 | 0 | 1 | 0 | 0 | 1 |
| rs1428971730 | G>C | SNV | chr19:33300957 | 0 | 1 | 0 | 0 | 1 |
| rs1417869963 | C>T | SNV | chr19:33300640 | 1 | 0 | 0 | 0 | 1 |
| rs1380779275 | C>A / C>T | SNV | chr19:33301007 | 1 | 0 | 0 | 0 | 1 |
| rs1247396010 | C>T | SNV | chr19:33300898 | 0 | 0 | 0 | 1 | 1 |
| rs1240079452 | A>C / A>G | SNV | chr19:33300645 | 1 | 0 | 0 | 0 | 1 |
| rs1184457934 | G>T | SNV | chr19:33300963 | 0 | 1 | 0 | 0 | 1 |
| rs1005626115 | C>T | SNV | chr19:33300393 | 0 | 0 | 1 | 0 | 1 |

Variants were identified and annotated using Geneious Prime.

These findings align with global AML patterns showing a mild male predominance but no consistent sex-based differences in treatment response or prognosis [19]. Hematological analyses revealed significantly reduced hemoglobin and platelet levels compared to controls, indicating impaired bone marrow function and disease burden. While WBC counts varied widely, underscoring the disease's biological heterogeneity. These findings are consistent with previous studies that describe cytopenia as characteristic features of AML at diagnosis [20, 21]. Based on the FAB classification, M3 was the most frequent AML subtype, followed by M4 and M0, while M1, M5, and M7 were rare [22]. Many remained unclassified, possibly due to overlapping with molecularly defined subtypes or atypical morphology. M3 is strongly associated with the *PML-RARA* fusion and favorable response to ATRA/ATO therapy [23], whereas M4 may correlate with *CBFB-MYH11* rearrangements. Less frequent subtypes such as M0 and M7 often require immunophenotypic confirmation and may carry poorer prognoses [24]. According to the revised WHO classification, AML includes the following molecular subtypes: *RUNX1-RUNX1T1*, *CBFB-MYH11*, and *PML-RARA*. The results revealed that *PML-RARA* aberrations were the most frequently observed, whereas *CBFB-MYH11* and *RUNX1-RUNX1T1* exhibited the lowest frequencies. These findings align with previous studies, such as those reported by [25], which demonstrated that the *PML-RARA* fusion gene is the most prevalent molecular abnormality among AML patients compared to other fusion gene subtypes. Furthermore, the mutational profile revealed notable distributions of *FLT3* and *NPM1* alterations among AML patients. The higher frequency of the *NPM1* mutation in AML patients compared to *FLT3-ITD* can be attributed to several factors. *NPM1* mutations are commonly found in many AML cases, especially those with normal karyotypes. These mutations typically result in the mislocalization of the *NPM1* protein, which affects cellular functions but does not always lead to as aggressive disease progression as *FLT3-ITD* mutations [26]. When mutations occur in both *CEBPA* and *FLT3*, the negative impact of *FLT3* may outweigh the positive effects of *CEBPA*, creating an inverse relationship. In this case, *FLT3-ITD* mutations can cancel out the potential benefit of *CEBPA* mutations. Conversely, *NPM1* mutations in this gene are generally associated with favorable outcomes, especially in the absence of *FLT3* mutations. The relationship between *CEBPA* and *NPM1* is often positive [27]. The amplification of the *CEBPA* gene's 3′UTR region by PCR, the appearance of a single, distinct band at the expected size (706 bp) across all samples, confirms the primer specificity and integrity of the amplified region. The sequencing analysis of the *CEBPA* gene revealed extensive genetic variability across the AML subgroups and the control sample, despite the region being non-coding. The predominance of rare, non-recurrent variants, most of which were either novel or altered forms of known SNPs, highlights the potential for significant regulatory heterogeneity within this region. Such heterogeneity may affect post-

transcriptional gene regulation, including mRNA stability, localization, or miRNA binding, which are critical in modulating *CEBPA* expression during hematopoiesis and leukemogenesis [28]. The high proportion of novel variants detected in individual patients suggests patient-specific alterations that may reflect unique somatic changes or previously uncharacterized germline polymorphisms in the population. This is particularly relevant given the limited representation of Middle Eastern genomes in global databases like dbSNP or gnomAD, which may contribute to classifying population-specific polymorphisms as novel. Furthermore, observing new forms at known SNP loci suggests that even well-characterized positions may harbor unrecognized functional diversity. These altered substitutions may escape annotation by standard pipelines, despite occurring at regulatory hotspots. The recurrence of variants such as rs41367646 and rs707656 in multiple patients, particularly within specific AML subgroups, may indicate potential associations with disease stage or treatment response, warranting further functional investigation. These findings emphasize the potential regulatory role of non-coding regions in AML and support extending genetic analysis beyond coding sequences, particularly in key transcriptional regulators like *CEBPA* [29].

## Conclusion

This study highlights the complex genetic landscape of AML in Iraqi patients, focusing on *CEBPA* mutations and their clinical correlations. The detection of numerous rare and novel variants within the non-coding 3′UTR region of the *CEBPA* gene underscores the potential regulatory significance of these regions in leukemogenesis. The observed diversity, including population-specific polymorphisms and recurrent variants, indicates the need for expanded genomic databases to represent underrepresented populations better. Furthermore, the interplay between *CEBPA*, *NPM1*, and *FLT3* mutations reinforces the importance of integrative molecular profiling for accurate risk stratification and personalized therapy in AML. These findings support incorporating non-coding regions into routine molecular diagnostics and open new avenues for investigating regulatory mechanisms in hematologic malignancies.

### Conflict of interests

The authors declared no conflict of interest.

### Funding source

The authors did not receive any source of funds.

### Data sharing statement

Supplementary data can be shared with the corresponding author upon reasonable request.

## REFERENCES

1. De-Morgan A, Meggendorfer M, Haferlach C, Shlush L. Male predominance in AML is associated with specific preleukemic mutations. *Leukemia*. 2021;35(3):867–870. doi: 10.1038/s41375-020-0935-5.
2. Ahn JS, Kim HJ. FLT3 mutations in acute myeloid leukemia: a review focusing on clinically applicable drugs. *Blood Res*. 2022;57(S1):32–36. doi: 10.5045/br.2022.2022017.
3. Story BA. Tracing the transformation of hematopoietic stem cells into leukemic cells using single-cell multi-omics. ETH Zurich. PhD Thesis, 2021. doi:10.3929/ethz-b-000512056
4. Chen W, Wang H, Hu J. Incidence of myelosuppression in AML is higher compared with that in ALL. *Mol Clin Oncol*. 2024;21(6):95. doi: 10.3892/mco.2024.2793.
5. Bain BJ, Béné MC. Morphological and immunophenotypic clues to the WHO categories of acute myeloid leukaemia. *Acta Haematol*. 2019;141(4):232–244. doi: 10.1159/000496097.
6. Voso MT, De Bellis E, Ottone T. Diagnosis and classification of AML: WHO 2016. In: Acute Myeloid Leukemia. Springer; 2021. p. 23–54. doi: 10.1007/978-3-030-72676-8_2.
7. Ayub RH, Ibraheem AH. Evaluation of related flow cytometry parameters with FAB subtype of acute myeloid leukemia associated with histological assessment. *Samarra J Pure Appl Sci*. 2020;2(3):84–94. doi: 10.54153/sjpas.2020.v2i3.126.
8. Akabari R, Qin D, Hussaini M. Technological advances: CEBPA and FLT3 internal tandem duplication mutations can be reliably detected by next generation sequencing. *Genes (Basel)*. 2022;13(4):630. doi: 10.3390/genes13040630.
9. Ali AM, Salih GF. Molecular and clinical significance of FLT3, NPM1, DNMT3A and TP53 mutations in acute myeloid leukemia patients. *Mol Biol Rep*. 2023;50(10):8035–8048. doi: 10.1007/s11033-023-08680-2.
10. de la Torre E, Serrano J, Martínez-Cuadrón D, Torres L, Sargas C, Ayala R, et al. Validation of mutated CEBPA BZIP as a distinct prognosis entity in acute myeloid leukemia: a study by the Spanish PETHEMA registry. *Haematologica*. 2024;18;109(8):2682–2687. doi: 10.3324/haematol.2023.284601.
11. Rivera JC, Nuñez D, Millar E, Ramirez K, Chandía M, Aguayo C. Mutations in the bZip region of the CEBPA gene: A novel prognostic factor in patients with acute myeloid leukemia. *Int J Lab Hematol*. 2023. doi: 10.1111/ijlh.14157.
12. Tien FM, Hou HA. CEBPA mutations in acute myeloid leukemia: implications in risk stratification and treatment. *Int J Hematol*. 2024;1–7. doi: 10.1007/s12185-024-03625-3.
13. Bullinger L. CEBPA mutations in AML: site matters. *Blood*. 2022;139(1):6–7. doi: 10.1182/blood.2021013557.
14. Yener E, Ozcan K, Öngören Ş, Salihoğlu A, Demiröz A, Akı H. A clinicopathological experience in acute myeloid leukemia: Effects of clinical data and status of FLT3, CEBPA and NPM1 on prognosis. *Indian J Pathol Microbiol*. 2022;65(3):642–648. doi: 10.4103/ijpm.ijpm_1012_21.
15. Chan JJ, Tabatabaeian H, Tay Y. 3′ UTR heterogeneity and cancer progression. *Trends Cell Biol*. 2023;33(7):568–582. doi: 10.1016/j.tcb.2022.10.001.
16. Bazinet A, Kadia TM. Changing paradigms in the treatment of acute myeloid leukemia in older patients. *Clin Adv Hematol Oncol*. 2022;20(1):37–46. PMID: 35060961.
17. Döhner H, Wei AH, Löwenberg B. Towards precision medicine for AML. *Nat Rev Clin Oncol*. 2021;18(9):577–590. doi: 10.1038/s41571-021-00514-1.
18. de Leeuw DC, Ossenkoppele GJ, Janssen JJ. Older patients with acute myeloid leukemia deserve individualized treatment. *Curr Oncol Rep*. 2022;24(11):1387–1400. PMID: 35653050.

19. Döhner H, Estey E, Grimwade D, Amadori S, Appelbaum FR, Büchner T, et al. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood*. 2017;129(4):424–447. doi: 10.1182/blood-2016-08-733196.
20. Imad I, Ghaloub AN, Ozaslan M, Alwan AF. Comparison between WBCs, RBCs and HGHB Levels for Iraqi AML patients pre and post 3-and 7-treatment within four age groups segmented based on growth levels. *Indian J Public Health Res Dev*. 2019;10(10). doi: 10.5958/0976-5506.2019.03265.0.
21. Merdin A, Dal MS, Çakar MK, Yildiz J, Ulu BU, Batgi H, et al. Analysis of pre-chemotherapy WBC, PLT, monocyte, hemoglobin, and MPV levels in acute myeloid leukemia patients with WT1, FLT3, or NPM gene mutations. *Medicine*. 2020;99(14):e19627. doi: 10.1097/MD.0000000000019627.
22. Prakisya NPT, Liantoni F, Aristyagama Y, Hatta P. Classification of acute myeloid leukemia subtypes M1, M2 and M3 using K-Nearest Neighbor. *Int J Adv Sci Eng Inf Technol*. 2021;11(5):1847–1853. doi: 10.18517/ijaseit.11.5.13706.
23. Sanz MA, Fenaux P, Tallman MS, Estey EH, Löwenberg B, Naoe T, et al. Management of acute promyelocytic leukemia: updated recommendations from an expert panel of the European LeukemiaNet. *Blood*. 2019;133(15):1630–1643. doi: 10.1182/blood-2019-01-894980.
24. Wang T, Cui S, Lyu C, Wang Z, Li Z, Han C, et al. Molecular precision medicine: Multi-omics-based stratification model for acute myeloid leukemia. *Heliyon*. 2024;10(17):e123456. doi: 10.1016/j.heliyon.2024.e123456.
25. Voso MT, Ottone T, Lavorgna S, Venditti A, Maurillo L, Lo-Coco F, et al. MRD in AML: the role of new techniques. *Front Oncol*. 2019;9:655. doi: 10.3389/fonc.2019.00655.
26. Abbas HA, Ravandi F, Loghavi S, Patel KP, Borthakur G, Kadia TM, et al. NPM1 mutant variant allele frequency correlates with leukemia burden but does not provide prognostic information in NPM1-mutated AML. *Am J Hematol*. 2019;94(6):E158. doi: 10.1002/ajh.25456.
27. Zhao Y, Huang Y, Jiang L, Zhang Y, Liu F, Yan P, et al. Impact of different CEBPA mutations on therapeutic outcome in acute myeloid leukemia. *Ann Hematol*. 2024;103(9):3595–3604. doi: 10.1007/s00277-024-05323-w.
28. Wiśnik A, Jarych D, Krawiec K, Strzałka P, Potocka N, Czemerska M, et al. Role of microRNAs in acute myeloid leukemia. *Genes*. 2025;16(4):446. doi: 10.3390/genes16040446.
29. Mustafa MI, Mohammed ZO, Murshed NS, Elfadol NM, Abdelmoneim AH, Hassan MA. In silico genetics revealing 5 mutations in CEBPA gene associated with acute myeloid leukemia. *Cancer Inform*. 2019;18:1176935119870817. doi: 10.1177/1176935119870817.