Association of GATA3 Rs3824662 Gene Polymorphism with Response to Induction Therapy in Acute Lymphoblastic Leukemia
Anna Muhammed Salih 1, Jaffar Nouri Jaffar Alalsaidissa 2
1 Ibn Al-Bitar Center for Cardiac Surgery, Baghdad, Iraq
2 Department of Pathology and Forensic Medicine, College of Medicine, University of Baghdad, Baghdad, Iraq
Corresponding Author: Anna Muhammed Salih, Email: annamuhammed485@gmail.com, Mobile: 07814957600

ABSTRACT
Background: The malignant transformation and multiplication of lymphoid precursor cells in the bone marrow, blood, and extramedullary tissues is known as acute lymphoblastic leukemia. ALL susceptibility has been linked to the GATA3 SNP rs3824662. Numerous studies have found a link between the rs3824662 risk allele and a worse prognosis and relapse.

Aim: To evaluate the frequency of GATA3 gene single nucleotide polymorphism (SNP rs3824662) in newly diagnosed Iraqi patients with acute lymphoblastic leukemia and its correlation to response to induction therapy.

Patients and methods: A cross-sectional study was conducted on 53 patients (33 adult and 20 pediatric patients) with newly diagnosed acute lymphoblastic leukemia. Sanger sequencing polymerase chain reaction-based techniques have been employed by us to enable the detection of GATA3 rs3824662 gene polymorphism.

Results: For the adult group genotype frequency of GATA3 rs3824662 wild type (CC) was 19 (57.6%) of patients, and heterozygous state (CA) genotype was detected in 14 (42.4%) of patients while the homozygous state (AA) genotype was not detected in the adult group of this study. In the pediatric group of ALL patients wild type (CC) was detected in 11 (55%) of patients, the heterozygous state (CA) genotype was present in 6 (30%) of patients, while the homozygous state (AA) genotype was detected in 3 (15%) of patients. There was no significant association between the GATA3 rs3824662 genotype and response to induction therapy, p-value = 0.54.

Conclusion: The GATA3 rs3824662 AA genotype and A allele could be risk factors for childhood and adult acute lymphoblastic leukemia. There was a nonsignificant association between the GATA3 rs3824662 genotype and response to induction therapy.

Keywords: GATA3 rs3824662, Acute lymphoblastic leukemia, Single nucleotide polymorphism.

INTRODUCTION
Acute lymphoblastic leukemia (ALL) is a clonal hemopoietic stem cell condition that is defined by the growth of blasts with little to no differentiation, which are typically B-cell or T-cell precursors. Assessing disease prognosis and choosing current and better treatments require an understanding of blast lineage

It was characterized by somatically acquired genomic alterations, but research over the past ten years has revealed that genetic polymorphisms that are inherited (germline) have a significant role in the variability in susceptibility between patients to ALL, drug action, and side effects of treatment. Germline variants strongly linked to ALL susceptibility have been discovered using genome-wide association studies (GWAS), which has led to a new understanding of the driving forces behind leukemogenesis and evidence of the intricate correlation between inherited and acquired genetic variants in ALL.

The GATA3 gene is situated at location p14, close to the end of the short arm of chromosome 10. The GATA3 SNP rs3824662 corresponds to 10p14 and is located in intron 3 of the transcription factor. A transcription factor required for early T-cell development and lymphoid cell lineage commitment is encoded by GATA3.

GATA3 has been referred to as a vital regulator of T-helper 2 (Th2) cell development in matured CD4+ T cells. The thymus's functions in the CD4 versus CD8 lineage option and at the β- selection checkpoint during T-cell development are best understood. Early B-cell commitment requires transcriptional regulation of GATA3, a crucial early regulator of innate lymphoid cells.

GATA3, which has been demonstrated to be expressed in multipotent hematopoietic stem cells, regulates the equilibrium between self-renewal and differentiation in hematopoietic stem cells (HSCs).

MATERIALS AND METHODS
Patients:
A cross-sectional study was done on 53 patients (33 adult and 20 pediatric patients) who are newly diagnosed with acute lymphoblastic leukemia. The collection period was from December 2021 to May 2022. The patients were seen and followed up in the Hematology Unit at Baghdad Medical City and Central child's hospital.

For each patient, a questionnaire form was done including the patient's general information, past and recent medical history, and results of blood and flow cytometry tests.

Bone marrow and peripheral blood were used to monitor patients following the induction therapy to see their response to treatment.
Inclusion criteria:
1. Newly diagnosed ALL patients who had not received any treatment for ALL before
2. For the adult group, the only patient who had received the Hyper-CVAD regimen which includes the drugs cyclophosphamide, vincristine sulfate, doxorubicin hydrochloride (Adriamycin), and dexamethasone was included in this study.

Exclusion criteria:
1. Patients who have previously received chemotherapy or radiation therapy.
2. Patients with Down syndrome.
3. Patients with Burkitt lymphoma.
4. Patients with acute lymphoblastic leukemia test positive for the Philadelphia chromosome.

Therapy and Follow-Up
Following induction therapy, patients were monitored using peripheral blood and bone marrow to see how well they were responding to treatment.

Adult participants in this study got the Hyper-CVAD regimen, which contains the medications cyclophosphamide, vincristine sulfate, doxorubicin hydrochloride (Adriamycin), and dexamethasone.

All Pediatric patients received UKALL-2003 and the assessment of induction therapy for pediatric patients was done on 28 days of chemotherapy.

Patients were categorized as follows:
1. Complete remission (9): less than 5% of bone marrow blasts, no extramedullary illness (such as CNS or soft tissue illness), and absolute neutrophil count (ANC) of at least 1000/μL; platelets greater than 100,000/μL.
2. Incomplete remission: Patients who did not attain total hematological remission.
3. Death: who passed away before the initial induction therapy was finished.

Methods
Blood sampling
Three milliliters of fresh peripheral blood was drawn from the eligible patients into EDTA tubes. DNA extraction using the Promega relia prep kit for DNA extraction (Lot No.0000493063) then DNA was directly frozen at −20°C in the teaching laboratories of Medical City until the time of sequencing.

Data analysis identified sequencing differences between samples of a particular gene using generous software following amplification, all the tests were done in the National Centre of Teaching Laboratories/Genetic department/Medical City and in the Advance Scientific (ASCO) learning center.

Standard sequencing
ABI3730XL automated DNA sequencers were used to perform Sanger sequencing on PCR products provided to Macrogen Corporation - Korea. The results were emailed to us, and we used specialized software to analyze them. Figure (1).

Figure (1): DNA sequencing for of rs3824662 shows wild (C/C=C), homogenous (A/A=A) and heterogenous (C/A=M) polymorphism.
Standard sequencing

This technique depends on DNA polymerase's selective inclusion of chain-terminating dideoxynucleotides during in vitro DNA replication. Single-stranded DNA templates, DNA primers, DNA polymerases, regular deoxynucleotide triphosphates (dNTPs), and modified di-deoxynucleotide triphosphates (ddNTPs), which end DNA strand elongation, are required for this.

DNA polymerase ceases expanding DNA when a modified ddNTP is present because these chain-terminating nucleotides lack the 3'-OH group required to produce a phosphodiester link between two nucleotides. The ddNTPs could be fluorescently or radioactively tagged for automated sequencing devices to detect them (10).

The DNA sample is split into four separate sequencing procedures, each of which includes DNA polymerase and all four of the basic deoxynucleotides (dATP, dGTP, dCTP, and dTTP).

Only one of the four dideoxynucleotides chain terminators—ddATP, ddGTP, ddCTP, or ddTTP—that have been labeled with fluorescent dyes and emit light at distinct wavelengths is added to each process; all additional nucleotides are ordinary nucleotides. Cycles of template DNA extension from the attached primer are followed by heat denature and capillary electrophoresis size (or length) separation of the resultant DNA fragments. The data are then generated as fluorescent peak trace chromatograms after detecting and recording the fluorescence dye (10). (Figure 2)

Figure (2): DNA sequencing using the standard (sanger) method.
Ethics consent:  
The study's approval was given by the ethical council of the College of Medicine at the University of Baghdad. All of the patients who participated in this study provided their informed consent.

Statistical analysis  
The IBM SPSS (Statistical Package for Social Science) program, version 26 (IBM SPSS, Inc., Chicago, IL), was used for the statistical analysis. Frequency and percentage were used to display qualitative data. Quantitative information was displayed as median and range or mean and standard deviation. The Kolmogrov-Smirnov test was utilized to determine the normality of the quantitative data.

Cross tabulation and chi-square were applied to analyze associations between categorical variables. We compared the medians of continuous variables using the Mann-Whitney test (a non-parametric version of the independent samples t-test) and the Kruskal-Wallis test (a non-parametric equivalent of the one-way ANOVA test). The p-value was calculated using exact testing. In all statistical analyses, a p-value of less than 0.05 was considered significant.

RESULTS  
The current study included 53 patients with newly diagnosed acute lymphoblastic leukemia, (33 adult and 20 pediatric patients).

For the adult group, out of 33 patients, 19 were males (57.6%) and 14 were females (42.4%) with a male-to-female ratio was 1.3. Their age ranged between 15 and 70 years with a median age of 32.0 years. Immunological subtyping for adult patients revealed that 81.8% of them were B-ALL and 18.2% were T-AL.

For the pediatric group, out of 20 patients, 15 were males (75%) and 5 were females (25%) with a male-to-female ratio was 3. Their age ranged from 1 to 14 years with a median age of 5 years. Immunological subtyping for pediatric patients revealed that 75% were B-ALL and 25% were T-ALL. The hematological parameters of adult patients were arranged in Table (1) and the hematological parameters of pediatrics were arranged in Table (2).

Table (1): The hematological parameters of ALL adult patients

<table>
<thead>
<tr>
<th>Hematological parameter</th>
<th>Hemoglobin (gm/dL)</th>
<th>Total WBC count (10⁹/L)</th>
<th>Platelet count (10⁹/L)</th>
<th>Blast cells percentage (PB)</th>
<th>Blast cells percentage (BM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median</td>
<td>7.5000</td>
<td>17.4000</td>
<td>37.0000</td>
<td>55.00</td>
<td>80.00</td>
</tr>
<tr>
<td>Range</td>
<td>3.7 -12.5</td>
<td>0.8 -300</td>
<td>13 - 180</td>
<td>1 - 97</td>
<td>50-97</td>
</tr>
</tbody>
</table>

Table (2): The hematological parameters of ALL pediatric patients

<table>
<thead>
<tr>
<th>Hematological parameter</th>
<th>Hemoglobin (gm/dL)</th>
<th>Total WBC count (10⁹/L)</th>
<th>Platelet count (10⁹/L)</th>
<th>Blast cells percentage (PB)</th>
<th>Blast cells percentage (BM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median</td>
<td>9.0000</td>
<td>12.2500</td>
<td>35.5000</td>
<td>68.00</td>
<td>92.00</td>
</tr>
<tr>
<td>Range</td>
<td>4.0 – 11.20</td>
<td>1.20- 305.0</td>
<td>7.9-236.0</td>
<td>0-93</td>
<td>70-99</td>
</tr>
</tbody>
</table>
In the adult group of this study, we found that the frequency of GATA3 rs3824662 wild type (CC) was 19 (57.6%) of patients, and the heterozygous state (CA) genotype was present in 14 (42.4%) of patients while homozygous state (AA) genotype was not detected in them.

In the pediatric group, the frequency of GATA3 rs3824662 wild type (CC) genotype was 11 (55%) of patients, the heterozygous state (CA) genotype was present in 6 (30%) of patients, and the homozygous state (AA) genotype was detected in 3 (15%) of them.

Following induction therapy, bone marrow and peripheral blood analyses were performed on all patients to see how they responded to the treatment. In the adult group (21/33, 63.6%) of patients had complete remission, (5/33, 15.2%) had incomplete remission, (4/33, 12.1%) died before they completed their therapy, and (3/33, 9.1%) lost from follow-up. In the pediatric group, there was complete remission in all patients (20/20, 100%).

There was no significant association between rs3824662 genotype and response to induction therapy, \( p = 0.54 \) (Table 3).

### Table (3): The correlation between rs3824662 genotype and response to induction therapy in the adult group.

<table>
<thead>
<tr>
<th>SNP rs3824662 genotype</th>
<th>Response to induction therapy</th>
<th>Total</th>
<th>( p )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Complete remission</td>
<td>Incomplete remission</td>
<td>Death</td>
</tr>
<tr>
<td>Wild type (CC)</td>
<td>12</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>75.0%</td>
<td>18.8%</td>
<td>6.3%</td>
</tr>
<tr>
<td>Heterozygous (CA)</td>
<td>9</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>64.3%</td>
<td>14.3%</td>
<td>21.4%</td>
</tr>
<tr>
<td>Total</td>
<td>21</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>70.0%</td>
<td>16.7%</td>
<td>13.3%</td>
</tr>
</tbody>
</table>

Chi – square test : \( p = 0.54 \) (N.S)

In both groups, there was a non-significant association between rs3824662 genotype and immunophenotype or gender with a \( p \)-value > 0.05.

There was a non-significant association between the median of all the hematological parameters (Hemoglobin, Total WBC count, Platelet count Blast cells percentage (PB), Blast cells percentage (BM), and rs 3838662 genotypes, \( p \)-value>0.05 in both groups, which was summarized in Table 4 for adult group and Table 5 for the pediatric group.

### Table (4): Comparison between the median of the hematological parameters in the adult group and rs 3824662 genotype

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>SNP1: Wild type N = 19</th>
<th>SNP1: Heterozygous N = 14</th>
<th>Mann-Whitney U</th>
<th>( z ) Score</th>
<th>( P )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Median</td>
<td>Mean Ranks</td>
<td>Median</td>
<td>Mean Ranks</td>
<td>U</td>
</tr>
<tr>
<td>Hemoglobin (gm/dL)</td>
<td>7.0</td>
<td>16.26</td>
<td>7.55</td>
<td>18.00</td>
<td>119.00</td>
</tr>
<tr>
<td>Total WBC count (10^9/L)</td>
<td>9.50</td>
<td>14.63</td>
<td>32.900</td>
<td>20.21</td>
<td>88.00</td>
</tr>
<tr>
<td>Platelet count (10^9/L)</td>
<td>37.00</td>
<td>16.55</td>
<td>43.00</td>
<td>17.61</td>
<td>124.50</td>
</tr>
<tr>
<td>Blast cells percentage (PB)</td>
<td>50</td>
<td>14.24</td>
<td>65.50</td>
<td>20.75</td>
<td>80.50</td>
</tr>
<tr>
<td>Blast cells percentage (BM)</td>
<td>80.00</td>
<td>15.74</td>
<td>86.00</td>
<td>18.71</td>
<td>109.00</td>
</tr>
</tbody>
</table>

Mann-Whitney test
Table (5): Comparison between the median of the hematological parameters in the pediatric group and rs 3824662 genotype

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>SNP1: Wild type N = 11</th>
<th>SNP1: Heterozygous N = 6</th>
<th>SNP1: Homozygous N = 3</th>
<th>(H) (df)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Median</td>
<td>Mean</td>
<td>Ranks</td>
<td>Median</td>
<td>Mean</td>
</tr>
<tr>
<td>Hemoglobin (gm/dL)</td>
<td>8.50</td>
<td>9.23</td>
<td></td>
<td>9.25</td>
<td>11.08</td>
</tr>
<tr>
<td>Total WBC count (10^9/L)</td>
<td>11.30</td>
<td>10.27</td>
<td></td>
<td>15.655</td>
<td>11.00</td>
</tr>
<tr>
<td>Platelet count (10^9/L)</td>
<td>35.00</td>
<td>9.91</td>
<td></td>
<td>21.500</td>
<td>9.00</td>
</tr>
<tr>
<td>Blast cells percentage (PB)</td>
<td>30.00</td>
<td>9.45</td>
<td></td>
<td>77.50</td>
<td>13.42</td>
</tr>
<tr>
<td>Blast cells percentage (BM)</td>
<td>95.00</td>
<td>11.91</td>
<td></td>
<td>90.50</td>
<td>9.75</td>
</tr>
</tbody>
</table>

Kruskal-Wallis Test

Through the analysis of sanger sequencing, there were another three sites of GATA3 gene polymorphism. They are rs374641, rs386680, and rs371668.

**DISCUSSION**

In the present study, the median age of enrolled newly diagnosed adult acute lymphoblastic leukemia patients was 32 years which was similar to the median age of the Chinese study by Hao et al. (11) but higher than the median age of the Iraqi study made by Alwan (12) and Indian study by Jain et al. (13) and lower than other studies by Orvain et al. (14) and Liao et al. (15). This discrimination may be related to the geographic and biological factors, as well as the effect of the environmental pollution.

Most of the adult patients in this study were males with a male-to-female ratio of 1.3 and this was in agreement with other Iraqi studies by Hao et al. (11), Liao et al. (15), Mohammed et al. (16) and Liu et al. (17).

In the present study, the median of hemoglobin (gm/dL) and platelet (10^9/L) of the adult group was 7.5 gm/dL and 37 × 10^9 respectively and it was lower than the result reported by Yilmaz et al. (18), Mosaad et al. (19), Koller et al. (20).

Regarding the median of the total WBCs count of adult patients, it was higher than the results reported by Yilmaz et al. (18), Mosaad et al. (19), and Koller et al. (20) but lower than the Chinese study by Liu et al. (17).

In our study, the median of bone marrow blast in the adult group was 92 and it was close to the results of other studies by Yilmaz et al. (18) and Koller et al. (20).

In the present study, there was complete remission in 63.6 % of enrolled adult patients and this is lower than other studies by Mohammed et al. (16), and Yilmaz et al. (18).

In this study, we discovered that the frequency of the wild GATA3 rs3824662 (CC) genotype was 57.6% of patients. The heterozygous (CA) genotype was detected in 42.4% of patients while the homozygous (AA) genotype was not detected in adult patients. These results were different from the results of American studies done by Koller et al. (20) and Perez et al. (21), which report a higher frequency of the A allele than our study. This difference in the frequency of the risk allele A may be due to a difference in sample size and may be due to ethnic variation; as there is a variation in susceptibility to ALL in diverse ancestry (22).

In the pediatric group of our study, the median age was 5 years and it was similar to the result of the Iraqi study by Al-Hadad et al. (23) and was in approximation with Chinese studies by Hao et al. (11), Liao et al. (15), and Liu et al. (17), but higher than the median age in a Sudanese study by Hussein et al. (24) and lower than Egyptian study by Ali et al. (25).

In the pediatric group, the male gender was observed more than females and this is in agreement with Iraqi studies by Al-Hadad et al. (23), Shalal et al. (26), Mustafa et al. (27), and Alani et al. (28) as well as other studies by Liu et al. (17), Mosaad et al. (19), Hussein et al. (24), and Ali et al. (25).

Bleeding tendency and splenomegaly were the most common presenting complaints in the pediatric group included and this following the Egyptian study by Mosaad et al. (19), but was different from the result of the Iraqi study by Shalal et al. (26) who reported that most frequent symptoms were fever and pallor and another Iraqi study by Alani (28) which show that swollen lymph node was the most presenting symptom.

In the pediatric group, the median of total WBCs count was lower than the result of the Iraqi study by Al-Hadad et al. (23), and the Egyptian study by Mosaad et al. (19) and higher than the Chinese studies by Hao et al. (11) and Liu et al. (17).

In the pediatric group of our study, there is complete remission of (100%) following the induction therapy and this was higher than the result reported by the Iraqi study by Al-Hadad et al. (23), American study by Pasquini et al. (29), Egyptian studies by Mosaad et al. (19), Ali (25) and Shibli et al (30) in approximation with
Sudanese study by Hussein et al. (24) and this disparity may result from different sample sizes.

In the pediatric group, we found that the frequency of GATA3 rs 3824662 wild (CC) genotype was 55% of patients. The Heterozygous (CA) genotype was detected in 30% of patients, while the homozygous (AA) genotype was detected in 15% of patients, and this was in approximation with the result of an Egyptian study by Mosaad et al. (19) and also in approximation with other study conducted by Madzio et al. (31) in Poland.

In this study we found that there was a non-significant association between GATA3 rs 3824662 genotype and response to induction therapy (p = 0.54) and this was similar to the result of the Egyptian study by Mosaad et al. (19) who found that the AA genotype was linked to a shortened disease-free survival DFS, a higher relapse rate, and a poor prognosis in pediatric ALL.

In this research non-significant association was found between GATA3 rs 3824662 genotype and immunophenotype or gender in both pediatric and adult groups and this was also similar to the result of the Egyptian study by Mosaad et al. (19).

Madjio et al. (31), Zhang et al. (32), and Li et al. (33) found that there was a significant statistical association between GATA3 rs 3824662 and minimal residual disease MRD following the completion of induction therapy. Zhang et al. (32) also found that with each extra copy of the risk allele, MRD positivity gradually increase. Madzio et al. (31) found that no matter the risk category or the protocol of therapy, the AA variant at rs3824662 was a significant factor impacting overall survival (OS), and AA homozygotes had a greater mortality rate than the CA and CC genotypes combined.

Jain et al. (13), Koller et al. (20), and Perez et al. (34) found that there was an association between the GATA3 rs3824662 genotype and Ph-like ALL, Additionally, they discovered that the rs3824662 risk allele was connected to variations in GATA3 expression as well as Ph-like ALL which is accompanied by somatic lesions (CRLF2 rearrangement, JAK gene mutation, and IKZF1 deletion).

Jain et al. (13), and Walsh et al. (35) found that there was a significant over-representation of this germline genetic variant in patients of Hispanic background comparable to Europeans and may explain their worse results, and discovered that rs3824662 contributes to an increase of 1.11 times in ALL in Hispanics compared to Europeans.

Migliorini et al. (3) and Walsh et al. (35) found that the rs3824662 risk allele was linked to a significantly lower rate of event-free survival and older age at diagnosis.

The analysis of sanger sequencing showed that there were another three sites of GATA3 gene polymorphism which are rs 374641, rs 386680, and rs 371668.

CONCLUSION

The GATA3 rs3824662 AA genotype and A allele could be risk factors for the emergence of childhood and adult acute lymphoblastic leukemia. There was no significant association between the rs3824662 genotype and response to induction therapy.

Funding: None.

Conflicts of interest: The authors affirm that they have no conflicts of interest.

REFERENCES


