CDKN2A and CDKN2B Gene Variants in Acute Lymphoblastic Leukemia in Tunisian Population

Sana Mahjoub1, Rabeb Ghali1, Vera Chaeib1, Bechir Aichou2, Fatma Megdich1, Malek Souayed1, Faouzi Janhanni1, Abdelaziz Soukri3, Wassim Y Alfawi4 and Touhami Mahjoub1,4

1Laboratory of Human Genome and Multifactorial Diseases (LR12ES07), Faculty of Pharmacy of Monastir, Haifa Regaieg University of Monastir
2Hematology Clinical departments CHU Farhat Hached Sousse
3Professor Laboratoire de Physiopathologie, Génétique moléculaire et Biotechnologie PGMBessor Casablanca Maroc
4Faculty of Sciences, El-Manar University, Tunis, Tunisia

Corresponding author: Touhami Mahjoub, Laboratory of Human Genome and Multifactorial Diseases (LR12ES07), Faculty of Pharmacy of Monastir, University of Monastir, Tunis, Tel: +0021698246153; E-mail: touhamimahjoub@gmail.com

Received date: February 18, 2018; Accepted date: March 08, 2018; Published date: March 15, 2018

Copyright: © 2018 Mahjoub T. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Introduction

Acute lymphoblastic leukemia (ALL) is an aggressive lymphoid cancer, frequent in younger than in adult population with 25 new children diagnosed with ALL every year in Tunisia [1,2]. ALL constitutes 25% of cancer that affect people before their twenties; with 2 at 5 years peak age. Current therapy allows the cure of approximately 60%-80% nowadays of (young) ALL patients and the five-year survival is estimated at 60% [2-5]. ALL is characterized by infiltration of monoclonal immature cells medullar and extra medullar areas [6-8]. The physiopathology of ALL is multifactorial, and includes interaction between modifiable (environmental), and non-modifiable factors, in particular genetic factors [9,10]. However, the genetic etiology dominates in younger subject, because children are less exposed to the environmental factors compared to adults [10,11]. In these regards, acquired somatic mutations contributing to ALL were reported [12]. These act by increasing the proliferation and survival of progenitor cells and impaired further cellular differentiation [12]. Though these and related alterations participate in the diagnosis and prognosis they remain insufficient to physio-pathological processes [12]. The introduction of novel diagnostic tools, such as high resolution karyotyping or CGH Array and genome wide molecular analysis may reveal additional somatic mutations not captured by standard cytogenetic analysis [7,13].

![CDKN2A/CDKN2B gene localised in chr 9; CDKN2A encode two suppressor proteins p14ARF (exon 1a and exon 2) and p16INK4A (exon 1b, exon2 and exon 3), CDKN2B encode one protein p15.](image)

Figure 1: CDKN2A/CDKN2B gene localised in chr 9; CDKN2A encode two suppressor proteins p14ARF (exon 1a and exon 2) and p16INK4A (exon 1b, exon 2 and exon 3), CDKN2B encode one protein p 15.

Earlier studied identified 18 single nucleotide polymorphisms (SNP) to be associated with ALL of which only 12 SNP had moderate odds ratio (OR) of 1.43-3.6, which included zinc finger protein subfamily 1A gene [11]. Subsequently, genome-wide associated scans (GWAS) confirmed an association between ALL and polymorphic variants in cyclin-dependent kinases-inhibitor (CDKN) 2A and CDKN2B genes in some, but not in other populations [7,14]. CDKN2A/B is important in control of cell cycle [15]. Protein CDKN2A and CDKN2B are encoded by CDKN2A and CDKN2B adjacent genes, localized on chromosome 9p21 (Figure 1) are frequently deleted, mutated and promoter methylated in some cases of cancer [16-18]. In response to these alterations, cancer cells become unable to control proliferation, and thus cell cycle regulation [19,20]. CDKN2A and CDKN2B genes were reported to be inactivated in hematologic malignancies, such as B-ALL (21%) and T-ALL (50%) patients [21]. Furthermore, heterozygous and homozygous deletions are more frequent than hyper methylation and mutations in exons 1 and 2 of CDKN2A in childhood ALL [21].

The aim of the current study is to evaluate the association between the risk of ALL in Tunisian children and CDKN2A and rs3731217 (intron1) gene variant which creates two overlapping cis-acting intronic splice enhancer motifs (CCCAAG and CAGTAC) which may regulate alternative splicing, and CDKN2A rs3734249 (exon2) missense SNP associated with alanine-to-threonine replacement [19,20,22,23]. In addition, the association of CDKN2B rs6626463 in ANRIL, regulates CDKN2B expression by disrupting a transcription factor binding site for CEFPB with ALL will also be investigated [24,25].

Subjects and Methods

Subjects

The retrospective study group includes 122 children with ALL, recruited in the period January 2013 to May 217, with mean age of 7.6 ± 4.5 yr. ALL was diagnosed and classified according to morphological and immunophenotypic characterization of blast cells in the bone marrow (OMS 2016). Of these patients, 106 were diagnostic with B-ALL and 16 with T-ALL. In addition, 91 children with mean age 7.9 ± 5.0 yr served as controls, and were recruited from pediatric general service for routine checkup, and were matched to cases according to self-declared ethnic origin. Blood samples were taken from all participants in EDTA-containing tube for total genomic DNA extraction. The guardians of patients and controls were required to sign a consent form before inclusion in the study, which was approved by the Ethic committee of CHU Farhat Hached (Sousse, Tunisia).
SNP Genotyping

Total genomic DNA was isolated from peripheral blood leukocytes by the salting-out method. Genotyping of rs3731217, rs3731249 (CDKN2A) and rs662463 (CDKN2B) was done using TaqMan® SNP Genotyping Assay, and specific primer pairs for each SNP. A standard 10 μl PCR reaction consisted of 1X TaqMan® Genotyping Master Mix (Applied Biosystems), 1X SNP Genotyping Assay Mix and 20 ng DNA.

Statistical analysis

Statistical analysis was performed on SPSS v. 23.0 (SPSS Inc., Chicago, IL). Data were expressed as percentages of total (categorical variables), or mean ± SD (continuous variables). Student’s t-test was used to determine differences in means, and Pearson χ² test was used to assess inter–group significance. Genotypes were tested for departures from Hardy–Weinberg equilibrium (HWE) in the control population using Haplovie 4.2 (www.broadinstitute.org/haplovie). All analyses were conducted under additive genetic model. Pairwise linkage disequilibrium (LD) values were calculated with Haplovie 4.2 and haplotype reconstruction was performed by the expectation maximization method. Logistic regression analysis was performed in order to determine the odds ratios (OR) and 95% confidence intervals (95%CI) associated ALL after controlling for age and gender as covariates, taking control subjects as the reference group. Statistical significance was set at P<0.05.

Results

Study subjects

The characteristics of ALL patients and control subjects are shown in Table 1.

Table 1: Characteristics of Study Participants.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Cases (122)</th>
<th>Controls (142)</th>
<th>P1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age²</td>
<td>7.6 ± 4.5</td>
<td>7.9 ± 5.0</td>
<td>0.687</td>
</tr>
<tr>
<td>≤ 10 years of age³</td>
<td>38 (30.9)</td>
<td>43 (30.3)</td>
<td>1.000</td>
</tr>
<tr>
<td>Gender (M:F)</td>
<td>73:49</td>
<td>79:63</td>
<td>0.533</td>
</tr>
<tr>
<td>WBC²</td>
<td>532,855 ± 4,883,861</td>
<td>13582 ± 18,724</td>
<td>0.206</td>
</tr>
<tr>
<td>WBC: 20,000 – 100,000³</td>
<td>23 (18.7)</td>
<td>11 (7.7)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>WBC &gt;100,000³</td>
<td>33 (26.8)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>RBC²</td>
<td>10.9 ± 2.0</td>
<td>8.1 ± 2.9</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Platelets²</td>
<td>92,881 ± 105,428</td>
<td>352,316 ± 155,183</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Positive family history (%)</td>
<td>93.9</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Immunophenotype: LAL-B³</td>
<td>107 (87.0)</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>LAL-T³</td>
<td>16 (13.0)</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

1. Student t-test for continuous variables, chi-square analysis for categorical variables.
2. Mean ± SD
3. Number (percent total)

No significant inter-group differences were recorded for age (P=0.687), or gender (P=0.533). All hematological indices were significantly different between ALL patients and control subjects.

Association studies

Summarizes the association between CDKN2Ars3731249, CDKN2Brs662463SNPs and ALL in case-control subjects Table 2.

Table 2: Association between CDKN2A, CDKN2B SNPs and ALL.

<table>
<thead>
<tr>
<th>Locus</th>
<th>SNP</th>
<th>Position1</th>
<th>Alleles</th>
<th>Cases2</th>
<th>Controls2</th>
<th>HWE</th>
<th>χ²</th>
<th>P 3</th>
<th>OR (95% CI)</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDKN2A</td>
<td>rs3731249</td>
<td>21970917</td>
<td>C:T</td>
<td>7 (0.03)</td>
<td>10 (0.04)</td>
<td>0.16</td>
<td>0.01</td>
<td>0.92</td>
<td>0.95 (0.35 – 2.53)</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td>rs3731217</td>
<td>21984662</td>
<td>G:T</td>
<td>24 (0.14)</td>
<td>28 (0.13)</td>
<td>0.21</td>
<td>0.01</td>
<td>0.92</td>
<td>1.03 (0.57 – 1.85)</td>
<td>0.28</td>
</tr>
<tr>
<td>CDKN2B</td>
<td>rs662463</td>
<td>22030439</td>
<td>C:T</td>
<td>41 (0.20)</td>
<td>38 (0.18)</td>
<td>0.51</td>
<td>0.32</td>
<td>0.57</td>
<td>1.15 (0.71 – 1.88)</td>
<td>0.36</td>
</tr>
</tbody>
</table>

MAF, Minor allele frequency; HWE, Hardy-Weinberg Equilibrium.
1. Location on chromosome based on dbSNP build 125.
2. Minor allele (frequency).
The genotypes of CDKN2A rs3731249 (P=0.16), rs3731217 (P=0.21), and CDKN2B rs662463 (P=0.51) were in HWE among control population. Minor allele frequency (MAF) of CDKN2A rs3731249 (P=0.92), rs3731217 (P=0.92), and CDKN2B rs662463 (P=0.57) were not significantly different between ALL cases and control subjects, even before applying the Bonferroni correction method for multiple testing. Setting homozygous major allele as reference (OR=1.00), results from Table 3 demonstrated lack of association of CDKN2A rs3731249 (P=0.54), rs3731217 (P=0.06), and CDKN2B rs662463 (P=0.47) genotypes with ALL under the additive model, as well as dominant or recessive models (data not shown).

### Table 2: Distribution of CDKN2A and CDKN2B alleles in cancer cases and control subjects.

<table>
<thead>
<tr>
<th>Gene</th>
<th>SNP</th>
<th>Genotype</th>
<th>Cases</th>
<th>Control</th>
<th>P</th>
<th>aOR1 (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDKN2 A</td>
<td>rs3731249</td>
<td>C / C</td>
<td>94 (0.93)</td>
<td>128 (0.93)</td>
<td>0.54</td>
<td>1.00 (Reference)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C / T</td>
<td>7 (0.07)</td>
<td>8 (0.06)</td>
<td>1.19</td>
<td>3.40 (0.42 - NA)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T / T</td>
<td>0 (0.00)</td>
<td>1 (0.01)</td>
<td>0.00</td>
<td>(0.00 – NA)</td>
</tr>
<tr>
<td>CDKN2 A</td>
<td>rs3731217</td>
<td>G / G</td>
<td>66 (0.76)</td>
<td>76 (0.73)</td>
<td>0.06</td>
<td>1.00 (Reference)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G / A</td>
<td>18 (0.21)</td>
<td>28 (0.27)</td>
<td>0.74</td>
<td>1.46 (0.38 - 1.46)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A / A</td>
<td>3 (0.03)</td>
<td>0 (0.00)</td>
<td>NA</td>
<td>(0.00 – NA)</td>
</tr>
<tr>
<td>CDKN2 B</td>
<td>rs662463</td>
<td>C / C</td>
<td>65(0.64)</td>
<td>69(0.66)</td>
<td>0.47</td>
<td>1.00 (Reference)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C / T</td>
<td>31(0.31)</td>
<td>34 (0.32)</td>
<td>0.97</td>
<td>1.75 (0.53 - 1.75)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T / T</td>
<td>5 (0.05)</td>
<td>2 (0.02)</td>
<td>2.65</td>
<td>(0.5014-16)</td>
</tr>
</tbody>
</table>

### Table 3: Genotype frequencies of CDKN2A and CDKN2B variants in cases and controls.

### Correlation studies

We next examined the correlation between CDKN2A and CDKN2B and ALL-associated hematological and biochemical indices. Spearman correlation calculation demonstrated correlation between CDKN2A rs3731249 and LAL/B/LALT phenotype (r=0.244, P=0.014), and between CDKN2B rs662463 and the liver function tests ALAT (r=-0.372, P=0.036), ASAT (r=-0.415, P=0.018), and LDH (r=-0.379, P=0.043) (Table 4). Lack of correlation was noted between CDKN2ArS3731217 and the biochemical and hematological examined in ALL patients Table 4.

### Table 4: Correlation of CDKN2A and CDKN2B variants with clinical features.

### Haplotype analysis

Three-locus [rs3731249, rs3731217 (CDKN2A) and rs662463 (CDKN2B)] haplotype analysis demonstrated that the majority of haplotype diversity was captured by 5 haplotypes in controls (96.1%) and cases (100.0%), which comprised CGC, CGT, CTC, TGC, and TGT haplotypes. Taking the common CGC haplotype as reference (OR=1.00), univariate and multivariate analysis demonstrated lack of association of any of the identified haplotypes with ALL, even before correcting for multiple testing Table 5.

### Table 5: CDKN2Aand CDKN2B haplotype distribution in cases and controls.
Discussion

This is the first study to analyze the association of CDKN2A and CDKN2B variants with ALL in North-African Arab-speaking Tunisians. ALL is a common childhood cancer. The genetic alterations increase cell proliferation, prolonged their survival, and/or impair the differentiation of the lymphoid hematopoietic progenitors [26,27]. Genetic alteration of CDKN2A/2B was reported as a risk factor in many type of cancer including ALL [9]. Migliorini G [7] and Lopez-Lopez E [14] have shown that the alterations of the CDKN2A–CDKN2B locus was associated with increased leukemia risk in some population but not in other [14,28]. In this study, we replicated the results from previous studies which evaluated the association of the risk ALL with rs3731217 and rs3731249 (CDKN2A) and rs662463 (CDKN2B). Our data showed lack of association of the tested CDKN2Aand CDKN2B variants with ALL, irrespective of the genetic model used (codominant, dominant or recessive). The five common haplotypes identified were also not associated with ALL.

The rs3731217 could affect gene expression by different mechanisms. It creates two overlapping cis-acting intronic splice enhancer motifs, CCCAGG and CAGTAC that may regulate alternative splicing of CDKN2A [24]. Its association with ALL is controversial, highlighted by the independent association of rs3731217 with ALL in Polish, Hispanic, and Thai populations [29-31]. The results of two GWAS of the scale study, and Latvian population reported comparable frequency of rs3731217 in ALL patients and controls comparable to our results, and with a negative association with childhood ALL [32,33]. On the other hand, studies on German, French and Canadian populations reported strong association of rs3731217 and ALL [32,34]. Collectively, this suggests that rs3731217 constitutes a European ALL-susceptibility locus [23,30]. In European population the risk remains statistically significant after adjustment for multiple testing. This SNP is associated with Pediatric B-cell precursor acute lymphoblastic leukemia by alteration of mRNA stability of the two tumor suppressors proteins p16 and p14 ARF [24].

GWAS identified rs3731249 as risk factor [24]. This SNP is localized in exon 2 of CDKN2A gene and encoded two proteins: p16INK4A, a negative regulator cyclin-dependant kinase, and p14ARF, a p53 activator Figure 2 [31].

This SNP is a mis-sense variant, associated with C-to-T substitution resulting in alanine-to-threonine replacement [25]. This substitution induces a missense change for p16INK4A open reading frame, reduce the capacity of tumor suppressor CDK4 and CDK6 that favor the B cell proliferation [23]. In p14ARF, the rs3731249 is in the 3'UTR region, where the T allele creates a miRNA binding site for miR-132-5p and miR-4642 [35]. These miRNAs could modify p14ARF expression, and then increased its function as cyclin inhibitor. Therefore, T allele of C/T SNP inCDKN2A caused B-ALL through its effect on the function of both p16INK4A andp14ARF.

The association of rs3731249 with ALL remains controversy [23]. It has high frequency of risk in African-American children than other populations. Gutierrez-Camino reported a positive association between rs3731249 and ALL risk in Spanish population. Also this variant is associated with B-ALL subtypes, including B-hyperdiploid ALL, and ETV6-RUNX1 ALL [25].

German (n=1155) and U.K. (n=824) GWAS of B-ALL patients demonstrated that Ala148Thr variant is a factor risk of B-ALL with calculated OR of 2.46 (95% CI=1.84-3.28), and 2.48 (95% CI=1.77-3.48) [31]. It is noteworthy that rs3731249 is implicated in other cancers, such as colorectal, lung and melanoma [36,37]. We have not established an association between CDKN2A Ala148Thr variant and ALL risk in our patient cohort children. MAF of Tunisian controls and case patients (0.04 vs. 0.03) were comparable to those of German (0.03 vs 0.08) and UK (0.03 vs 0.05) populations [31].

We also examined the association between rs662463 CDKN2B variant and ALL. This variant is localized in the long non-coding RNA encoded in the chromosome 9p21 region, ANRIL [31,38]. ANRIL regulates the expression of CDKN2A/B tumor suppressors, and thus control the cell proliferation, apoptosis, senescence and aging [39,40]. The presence of rs662463 in ANRIL alters the binding site for CEBPB transcription factor, frequently mutated in BCP-ALL, increased the BALL risk by diminishing the level of p 15 tumor suppressor [24]. This variant was demonstrated to be a B-ALL risk factor African-Americans and populations of European descent [24]. The CDKN2B rs662463 variant is in almost complete LD with rs2811712 (r²= 1.00), which is a
B-ALL risk locus in European-Americans and African-Americans [24]. While our results did not establish a correlation between rs662463 and ALL risk, there was correlation between CDKN2B rs662463 and liver function tests ALAT, ASAT, and LDH. This prompts the speculation as to possible diagnostic capacity of this variant, given that increased liver enzyme is likely due to hepatic injury from leukemic infiltrates or treatment toxicity [41]. Our data also established correlation between CDKN2A rs3731249 and LAL-B/LAL-T phenotype, with CDKN2A/2B polymorphisms being more correlated with LAL-T than LAL-B [42-47].

Conclusion

Our data don't show a correlation between the three SNPs, although, the OR is nearly of the significativity for the SNPs. The controversies of the results are probably associated to the variation of the ethnicity that may be influence the mechanism physiopathology of ALL, the still of patient live, age and patient cohort. Our study has several strengths namely, which the first analysis of three SNPs in Arabic speaking populations, and that revealed the correlation between CDKN2Brs 662463 and the liver function. However some limitations of this case – control study relate that the cohort and the power are insufficiently, because the ALL prevalence is weak in Tunisia.

Authors’ Roles

Sana Mahjoub: genotyping and paper writing.
Rabbeb Ghali: Paper writing.
Verra Chaeib: DNA extraction.
Fatma Megdiche: patients recruitement
Malmek Souayed: genotyping technical
Faouzi Janhani: Project funding.
Abdelaziz Soukri: Head Moroccan of research project
Wassim Youssef Almawi: Statistic study.
Touhami Mahjoub

Highlights

This work is the first study associated three CDKN2A/2B genotype and haplotype in Arabic speakin population. It is one of few study replicated those rs with ALL in African populations Sperman correction reveled the correlation between CDKN2B rs 662463 and the liver function and between CDKN2A rs3731249 and LAL-B/LAL-T phenotype.

Acknowledgments

We wish to express our gratitude to all the patients and controls

Formatting of Funding Sources

The work was funded by the Laboratory of Human Genome and Multifactorial Diseases (LR12ES07), Faculty of Pharmacy of Monastir, University of Monastir and collaboration project Tunisia- Maroccan.

Conflict of Interest

None

Disclosure Statement

The authors have nothing to disclose

References


