Risk assessment of FLT3 and PAX5 variants in B-acute lymphoblastic leukemia: a case - control study in a Pakistani cohort

Ammara Khalid Corresp., 1, Sara Aslam 1, Mehboob Ahmed 1, Shahida Hasnain 1, Aimen Aslam 2

1 Department of Microbiology & Molecular Genetics, University of the Punjab, Lahore, Pakistan
2 Department of statistics and actuarial science, University of the Punjab, Lahore, Pakistan

Corresponding Author: Ammara Khalid
Email address: amara.mmg@pu.edu.pk

AIMS: B-cell acute lymphoblastic leukemia (B-ALL) is amongst most prevalent cancers of children in Pakistan. Genetic variations in FLT3 are associated with auto-phosphorylation of kinase domain that leads to increased proliferation of blast cells. Paired box family of transcription factor (PAX5) plays a critical role in commitment and differentiation of B-cells. Variations in PAX5 are associated with the risk of B-ALL. We aimed to analyze the association of FLT3 and PAX5 polymorphisms with B cell leukemia in Pakistani cohort.

METHODS: We collected 155 B-ALL subject and 155 control blood samples. For analysis, genotyping was done by tetra ARMS-PCR. SPSS was used to check the association of demographic factors of SNPs present in the population with the risk of B-ALL.

RESULTS: Risk allele frequency A at locus 13q12.2 (rs35958982, FLT3) was conspicuous and showed positive association (OR= 2.30, CI=1.20-4.50, P=0.005) but genotype frequency (OR=3.67, CI=0.75-18.10, P=0.088) failed to show any association with the disease. At locus 9p13.2 (rs3780135, PAX5), the risk allele frequency was significantly more in B-ALL subjects than ancestral allele frequency (OR=2.17, CI=1.37-3.43, P=0.000). Genotype frequency analysis of rs3780135 polymorphism exhibited the protective effect (OR=0.55, CI=0.72-1.83, P=0.029). At locus 13q12.2 (rs12430881, FLT3), the minor allele frequency G (OR=1.15, CI=1.37-3.43, P=0.043) and genotype frequency (OR=2.52, P=0.006) reached significance as showed p<0.05.

CONCLUSION: In present study, a strong risk of B-cell acute lymphoblastic leukemia was associated with rs35958982 and rs12430881 polymorphisms. However, rs3780135 polymorphism showed the protective effect. Additionally, other demographic factors like family history, smoking and consanguinity were also found to be important in risk assessment. We anticipate that the information from genetic variations in this study can aid in therapeutic approach in the future.
Risk assessment of \textit{FLT3} and \textit{PAX5} variants in B-acute lymphoblastic leukeima: a case-control study in a Pakistani cohort

Ammara Khalid$^1$, Sara Aslam$^2$, Dr. Mehboob Ahmed$^3$, Dr. Shahida Hasnain$^4$, Aimen Aslam$^5$

$^1$Department of Microbiology & Molecular Genetics, University of the Punjab, Lahore, Pakistan.
$^2$Department of Microbiology & Molecular Genetics, University of the Punjab, Lahore, Pakistan.
$^3$Department of Microbiology & Molecular Genetics, University of the Punjab, Lahore, Pakistan.
$^4$Department of Microbiology & Molecular Genetics, University of the Punjab, Lahore, Pakistan.
$^5$Department of statistics and actuarial science, University of the Punjab, Lahore, Pakistan.

Corresponding Author:
Ammara Khalid$^1$

Department of Microbiology & Molecular Genetics, University of the Punjab, Lahore, Pakistan.

Email address: amara.mmg@pu.edu.pk

Abstract

AIMS: B-cell acute lymphoblastic leukemia (B-ALL) is amongst most prevalent cancers of children in Pakistan. Genetic variations in \textit{FLT3} are associated with auto-phosphorylation of kinase domain that leads to increased proliferation of blast cells. Paired box family of transcription factor (\textit{PAX5}) plays a critical role in commitment and differentiation of B-cells. Variations in \textit{PAX5} are associated with the risk of B-ALL. We aimed to analyze the association of \textit{FLT3} and \textit{PAX5} polymorphisms with B cell leukemia in Pakistani cohort.

METHODS: We collected 155 B-ALL subject and 155 control blood samples. For analysis, genotyping was done by tetra ARMS-PCR. SPSS was used to check the association of demographic factors of SNPs present in the population with the risk of B-ALL.

RESULTS: Risk allele frequency A at locus 13q12.2 (rs35958982, \textit{FLT3}) was conspicuous and showed positive association (OR= 2.30, CI=1.20-4.50, \(P=0.005\)) but genotype frequency (OR=3.67, CI=0.75-18.10, \(P=0.088\)) failed to show any association with the disease. At locus 9p13.2 (rs3780135, \textit{PAX5}), the risk allele frequency was significantly more in B-ALL subjects than ancestral allele frequency (OR=2.17, CI=1.37-3.43, \(P=0.000\)). Genotype frequency analysis of rs3780135 polymorphism exhibited the protective effect (OR=0.55, CI=0.72-1.83, \(P=0.029\)). At locus 13q12.2 (rs12430881, \textit{FLT3}), the minor allele frequency G (OR=1.15, CI=1.37-3.43, \(P=0.043\)) and genotype frequency (OR=2.52, \(P=0.006\)) reached significance as showed \(P<0.05\).
CONCLUSION: In present study, a strong risk of B-cell acute lymphoblastic leukemia was associated with rs35958982 and rs12430881 polymorphisms. However, rs3780135 polymorphism showed the protective effect. Additionally, other demographic factors like family history, smoking and consanguinity were also found to be important in risk assessment. We anticipate that the information from genetic variations in this study can aid in therapeutic approach in the future.

Introduction

According to the Punjab cancer registry report, acute lymphoblastic leukemia (ALL) is a predominant malignancy of children and it makes up most prevalent cancer in Punjab, Pakistan. The worldwide incidence rate is 1-4.75 per 100,000 people. In Pakistan ALL contributes to 17.9% among all cancers. It is characterized by mutation in blast cells in hematopoietic stem cells, spleen, neurons, gonads, lymph nodes, and hepatic cells (Portell et al., 2013). Although, B-ALL is very common in children but it may also occur in adult populace (Forero et al., 2013). Several demographic parameters like gender, age, family history and biological factors also play an important role in the prevalence of disease. Other factors like exposure to UV, radiations, lifestyle may also act as risk factors (Levine et al., 2016; Acharya et al., 2018). Mutation in certain genes involved in different processes like apoptosis, proliferation and differentiation of B-cells may also cause B-ALL. These genetic alterations largely affect the prediction and therapeutic approach used for medication and therapy of ALL (Tasian and Hunger, 2017).

FMS-like tyrosine kinase (FLT3) belongs to class III receptor tyrosine kinase (RTK) family. Structurally, FLT3 consist of an extracellular domain at the amino terminus. This domain comprises of immunoglobulin-like transmembrane region and intracellular juxta-membrane domain (JMD). At carboxyl terminus, there are two kinase domains, separated by a kinase insert region (Gilliland and Griffin, 2002). FLT3 is expressed in normal human bone marrow especially in CD34+ hematopoietic stem, brain (Çakmak-Görür et al., 2019) and gonads (Matthews et al., 1991; Small et al., 1994) and encodes 1000 amino acid protein in humans. In the hematopoietic tissues, Binding of FL with its receptor causes auto-phosphorylation of tyrosine residues present in the kinase domain and stimulates growth of progenitor cells in the marrow and blood (Marhäll et al., 2018). This results in downstream activation of signaling pathways that are involved in regulation of cell cycle or apoptosis, including (PI3K), caspase-9 and Ras/Raf pathways and
causes multiplied proliferation of cells, reduced cell apoptosis, and inhibition of B-cell differentiation (Zhang and Broxmeyer, 2000).

In hematologic malignancy, 70% to 100% increased expression of FLT3 in acute myeloid leukemia (AML) and acute lymphoblastic leukemia (ALL) is reported previously (Brown et al., 2005; Griffith et al., 2016). Rosnet and colleagues reported 3 of 5 ALL subjects with increased expression of FLT3 in leukemia blasts (Rosnet et al., 1996). Another study showed that up regulation of FLT3 gene is a potential risk factor of leukemia (Cheng et al., 2018).

B-cell-specific activator protein (PAX5) encodes transcription factors that are the member of a paired box domain. PAX5 plays imperative role in the commitment of B-cell lineage from blast cells as it controls the differentiation of a pro-B cell to pre-B cells (Fuxa and Skok, 2007; Lang et al., 2007). In pre-pro-B cells the immunoglobin gene rearrangement starts and matures into pro-B cells. Expression of PAX5 gene initiates from pro-B stage and terminates at pre-B stage. In late B-lymphoposis, PAX5 maintains the function of mature B-cells (Shahjahani et al., 2015).

In B-cell malignancies, PAX5 act as an oncogene. Down-regulation of PAX5 halts B-cells and reverts B-cell precursors (BCPs) to progenitors (pro B-cell stage) (Schebesta et al., 2007; Carotta and Nutt, 2008). Conversely, uncontrolled proliferation of the B-cells leads to the abnormal expression of PAX5 in precursor cells and inhibit T-cell proliferation (Souabni et al., 2007). It is reported that in childhood acute lymphoblastic leukemia (ALL), translocations and mutation in PAX5 are more prevalent (Bousquet et al., 2007; Nebral et al., 2009; Santoro et al., 2009; Iacobucci and Mullighan, 2017). Alternative splicing of PAX5 in exon 7 to exon 9 results into five isoforms. These isoforms are more expressed in primary B-cell lymphoma tissues and cancerous cell lines (Zwollo et al., 1997; Arseneau et al., 2009).

Previous studies showed that the presence of single nucleotide polymorphisms (SNPs) in genome maybe risk causing or protective for the disease and it may also alters the pharmacokinetic and pharmacodynamics properties of drugs (Kumanayake, 2013; Pui, 2015; Tasian and Hunger, 2017). We selected two non-synonymous SNPs including 13q12.2 (rs35958982, FLT3), 557 (Val > Ile) at position Chr13:28034336 (GRCh38.p12) and 9p13.2 (rs3780135, PAX5), 293 (Thr >Ile) at position Chr9:36840626 (GRCh38.p12). A synonymous SNP 13q12.2 (rs12430881, FLT3), (A>G) at position Chr13:28020665 (GRCh38.p12) was also selected. The change in amino acid sequence due to non-synonomous SNP alters the protein structure implicating its expression and function. Current study is designed to evaluate the role of FLT3 and PAX5 genes in B-cell
lymphoblastic leukemia. For this purpose, a case control analysis was conducted to evaluate the polymorphic association of rs35958982, rs3780135 and rs12430881 with B-cell acute lymphoblastic leukemia (B-ALL) incidence.

**Materials & Methods**

**Study subjects**

The present study was conducted at the University of Punjab, Pakistan and granted ethical approval to carry out the study within its facilities (Ethical Application Ref: sbs/222/18). Blood samples were collected during the period of January 2017 to February 2017 from Children's Hospital, Lahore, Pakistan. Study population comprised of 155 cases and 155 controls younger than 15 years of age. The diagnostic criteria for B-ALL cases include B-cell positive markers (CD19, CD10, CD22, and CD20) confirmed by flow cytometry analysis. Cases with relapsed and newly diagnosed B-ALL were also included. All 310 subjects recruited were consented to participate in this study after filling the questionnaire. The subjects with any other type of leukemia, blood infectious disease, and B-ALL subjects older than 15 years of age were excluded from the study. Family history with cancer, parental consanguinity (first and second degree relatives) and smoking status (>100 cigarettes in lifetime) were gathered by questionnaire interviewed.

**Genotyping:**

Venous blood samples of cases and controls were collected in EDTA vials. DNA extraction was done using Sambrook 2001 organic protocol. The genes and SNPs associated with B-ALL were screened using DisGeNET platform (Queralt-Rosinach *et al.*, 2016) and were verified by dbSNP database (Sherry *et al.*, 2001). Presence of the selected SNPs in Pakistani population was confirmed by Ensembl genome browser (Frankish *et al.*, 2017). In order to identify the SNPs, tetra arm primers were designed using Primer1 software (Ye *et al.*, 2001) as shown in table 1. Tetra arms PCR was done using advanced primus 96 (PeqLab) thermal cycler (table 2). PCR products were further analyzed by gel electrophoresis (fig. 1, 2, and 3).

**Statistical analysis:**

Statistical studies were performed using IBM SPSS 23. Chi-square test was conducted to compare categorical data. Allele and genotype association between SNPs and B-ALL were calculated by computing odds ratio (OR) and 95% confidence interval (CI). The Bonferroni
corrections were applied for all multiple tests. A logistic regression model was used to adjust
different B-ALL risk factors. The probability level accepted for significance was $P < 0.05$.

**Results**

Family history of cancer and parental consanguinity showed significant association with B-ALL
while, there was no association with the smoker parents. Subjects with a family history of any
type of cancer showed a high risk of having B-ALL (OR=15.42, $P=0.000$). Previous studies
showed smoking as a risk factor for cancer but our cohort displayed a contradictory results as no
significant association was found in B-ALL subjects (OR=0.85, $P=0.580$). In the present study,
more B-ALL subjects were product of parental consanguinity and showed highly significant
association with the risk of B-ALL (OR=1.87, $P=0.050$) (Table: 3).

Our data showed that none of the subjects and their parents was exposed to radiations.
Furthermore, eighteen patients had liver hepatomegaly sized $11.7 \pm 3.3\text{mm}$, nine cases had
nephropathy of right kidney $9.47 \pm 2.9\text{mm}$ and left kidney $10.01 \pm 2.3\text{mm}$. Symptoms like night
sweating, dizziness, abdominal pain, vomiting, bruises, pallor, enlarged lymph nodes, cough
with blood, loose stools, jaundice, pedal edema, pain, dehydration, hepatosplenomegaly, atypical
blast cells mild abdominal ascites, low leukocytes and thrombocytopenia were also recorded.

In our cohort, rs35958982 encoding isoleucine form of codon frequency in subjects was 13.4%
and 5.6% in controls. Moreover, statistical analysis showed positive association of allele
frequency (OR=2.30, CI=1.20-4.50, $P=0.005$) and no association of genotype frequency
(OR=3.67, CI=0.75-18.10, $P=0.088$) with the disease. Another polymorphism rs3780135, a
minor allele frequency in subjects was 47.1% and 32.58% in controls showing positive
association with B-ALL (OR=2.17, CI=1.37-3.43, $P=0.000$). The genotype frequency showed
protective effect with B-ALL (OR=0.55, CI=0.72-1.83, $P=0.029$). The SNP rs12430881 allele
frequency (OR=1.15, CI=1.37-3.43, $P=0.043$) and genotype frequency (OR=2.52, CI=1.28-4.95,
$P=0.006$) showed strong association with the disease as shown in table: 4. After applying
Bonferroni correction, SNPs rs35958982, rs3780135 and rs12430881 remained statistically
significant and showed $P$-value 0.030, 0.010 and 0.002 respectively.

Multivariate regression analysis was performed after adjusting the baseline for conventional B-
ALL risk factors such as family history, smoking and parental consanguinity. As shown in table
4, the multivariate analysis indicated that outcome of heterozygous genotype GA in SNPs
rs35958982, rs3780135 and rs12430881 had significant association with B-ALL and showed
odds ratio (OR=1.13, CI=0.41-3.08), (OR=1.19, CI=0.52-2.73) and (OR=1.09, CI=0.48-2.69) respectively. Additionally, risk genotypes in SNPs rs35958982 (AA) and rs12430881 (GG) showed positive association with disease having an odds ratio (OD=1.30, CI=0.37-5.08) and (OD=1.03, CI=0.50-2.68) respectively. However, the risk genotype (AA) of SNP rs3780135 displayed no association with B-ALL after adjusting for environmental factors. Stratification analysis of environmental factors showed smoking as major risk factor in both heterozygous and risk genotype of SNP rs3780135 and rs12430881 whereas, parental consanguinity act as risk factor only in heterozygous genotype of SNP rs3780135 as shown in table 5.

Discussion

According to the previous studies, association of first and second degree family history of cancer signifies genetic and environmental risk factor for causing acute lymphoblastic leukemia. Our study also showed positive association of family history with B-ALL (OR=15.42, P=0.000). Earlier, parental smoking has also been associated with the prevalence of ALL but our study showed contrary results (OR=0.85, P=0.580) (Belson et al., 2007). Parental consanguinity is still practiced in Pakistan, which results in minor allele pool and contributes to the occurrence of disease. Our results are in accordance with (Steinberg and Steinfeld, 1960; Urtishak et al., 2016) which states that familial occurrence of leukemia exists (OR=1.87, P=0.050). Some studies found a correlation between parental exposure to radiation before conception, that may be due to their working environment (Shu et al., 2002). In our analysis, nor of patients neither of parents was ever exposed to radiations. Hepatomegaly and nephropathy are often seen in B-ALL subjects having chemotherapy. Malfunctioned leucocytes in the liver and kidney leads to enlargement of these organs (Rasool et al., 2015). Another study suggests that hepatomegaly and nephropathy may be the consequence of chemotherapeutic toxicity (Giamanco et al., 2016).

It is well established fact that cancer risk is influenced by numerous genetic variants having any risk or protective effect. The degree of penetrance of a certain genotype in the population and environmental factors is a major cause of cancer (Fletcher and Houlston, 2010). The information given by allelic and genotypic data of single nucleotide polymorphism in a population propose the possible genetic markers for cancer risk and predict possible targeted therapies (Griffith et al., 2016; Wu and Li, 2018).

In this study, SNP rs35958982 is a germline polymorphism present in transmembrane region of FLT3 gene. It is a non-synonymous variant which leads to the change in structure of the protein.
High throughput DNA sequence analysis has been done to check the frequency of rs35958982 with leukemiogenesis in drivers and passengers which showed no association with AML (Fröhling et al., 2007). Present study in contrast displayed the association of SNP rs35958982 with the disease (OR= 2.30, CI=1.20-4.50, P=0.005). Detailed analysis of genotype frequency in the population showed no association with B-ALL (OR=3.67, CI=0.75-18.10, P=0.088). This might be due to the fact that SNP rs35958982 is rare in acute lymphoblastic leukemia with low penetrance. Current study also depicts that individuals with risk allele A at locus 13q12.2 (rs12430881, FLT3) (OR=1.15, CI=1.37-3.43, P=0.0426) and genotype GG were more prone to B-ALL (OR=2.52, CI=1.28-4.95, P=0.006). It has been found that disruption of FLT3 gene due to the presence of mutation or SNP leads to deficiency of B-lymphoid progenitors suggesting its critical role in survival and proliferation of blast cells (Zriwil et al., 2018).

Bodian et al., studied allele frequency of paired box domain (PAX5) polymorphism rs3780135 in different populations i.e. African 34%, African European 49%, Central Asian 85%, East Asian 94%, European 95% and Hispanic 88% (Bodian et al., 2014). Pakistan lies in South East Asia having frequency of rs3780135 (47.1%) which is lower than previously reported in East Asian population. Firtina et al., found polymorphism rs3780135 in B-ALL subjects with increased mRNA expression of PAX5 suggesting the possible role of SNP with increased proliferation of blast cells (Firtina et al., 2012). In Pakistani population, minor allele frequency was significantly identified in B-ALL subjects (OR=2.17, CI=1.37-3.43, P=0.000). Heterozygous genotype GA (38.7%) was more frequently identified in our cohort than homozygous risk genotype AA (27.74%) which manifested significant difference in frequency (CI=0.72-1.83, P=0.029) and also showed protective effect (OR=0.55). PAX5 is involved in repression of T-cells, activation of B-cell proliferation from blast cell therefore, presence of any variant in this gene affects its pathway which may leads to increased expression of PAX5 and results into B-ALL (Firtina et al., 2012).

**Conclusions**

The findings of the present study significantly demonstrate that SNPs rs35958982 and rs12430881 correlates with the increase risk of B-ALL however, SNP rs3780135 has protective effect. The environmental risk factors of B-ALL including family history, parental consanguinity and smoking are found to have imperative role in progression of disease. Although the data is balanced but not robust and a small cohort of subjects limits the conclusion of manuscript. To
eliminate limitation and validate the results of present study, larger prospective studies need to be conducted in the same ethnic group. Furthermore various other demographic and environmental factors should also be considered and appraised for their association with B-ALL.

Acknowledgements

Authors would like to acknowledge The Children's Hospital and Institute of Child Health, Lahore, Pakistan for their assistance and support during blood sample collection.

References

Acharya, U. H., Halpern, A. B., Wu, Q., Voutsinas, J. M., Walter, R. B., Yun, S., Kanaan, M. and Estey, E. H. 2018. Impact of region of diagnosis, ethnicity, age, and gender on survival in acute myeloid leukemia (AML). *Journal of drug assessment*, 7:51-53.

Arseneau, J. R., Laflamme, M., Lewis, S. M., Maïcas, E. and Ouellette, R. J. 2009. Multiple isoforms of PAX5 are expressed in both lymphomas and normal B-cells. *British journal of haematology*, 147:328-338.

Belson, M., Kingsley, B. and Holmes, A. 2007. Risk factors for acute leukemia in children: a review. *Environmental health perspectives*, 115:138.

Bodian, D. L., McCutcheon, J. N., Kothiyal, P., Huddleston, K. C., Iyer, R. K., Vockley, J. G. and Niederhuber, J. E. 2014. Germline variation in cancer-susceptibility genes in a healthy, ancestrally diverse cohort: implications for individual genome sequencing. *PloS one*, 9:e94554.

Bousquet, M., Broccardo, C., Quelen, C., Meggetto, F., Kuhlein, E., Delsol, G., Dastugue, N. and Brousset, P. 2007. A novel PAX5-ELN fusion protein identified in B-cell acute lymphoblastic leukemia acts as a dominant negative on wild-type PAX5. *Blood*, 109:3417-3423.

Brown, P., Levis, M., Shurtleff, S., Campana, D., Downing, J. and Small, D. 2005. FLT3 inhibition selectively kills childhood acute lymphoblastic leukemia cells with high levels of FLT3 expression. *Blood*, 105:812-820.

Çakmak-Görür, N., Radke, J., Rhein, S., Schumann, E., Willimsky, G., Heppner, F. L., Blankenstein, T. and Pezzutto, A. 2019. Intracellular expression of FLT3 in Purkinje cells: implications for adoptive T-cell therapies. *Leukemia*:1.

Carotta, S. and Nutt, S. L. 2008. Losing B cell identity. *Bioessays*, 30:203-207.
Cheng, J., Qu, L., Wang, J., Cheng, L. and Wang, Y. 2018. High expression of FLT3 is a risk factor in leukemia. Molecular medicine reports, 17:2885-2892.

Firtina, S., Sayitoglu, M., Hatirnaz, O., Erbilgin, Y., Oztunc, C., Cinar, S., Yildiz, I., Celkan, T., Anak, S. and Unuvar, A. 2012. Evaluation of PAX5 gene in the early stages of leukemic B cells in the childhood B cell acute lymphoblastic leukemia. Leukemia research, 36:87-92.

Fletcher, O. and Houlston, R. S. 2010. Architecture of inherited susceptibility to common cancer. Nature Reviews Cancer, 10:353.

Forero, R. M., Hernández, M. and Rivas, J. M. H. 2013. Genetics of Acute Lymphoblastic Leukemia. In: Guenova, M. and Balatzenko, G. [Eds.] Leukemia. InTech, Rijeka, pp. Ch. 01.

Frankish, A., Vullo, A., Zadissa, A., Yates, A., Thomann, A., Parker, A., Gall, A., Moore, B., Walts, B., Aken, B. L., Cummins, C., Girón, C. G., Ong, C. K., Sheppard, D., Staines, D., Murphy, D. N., Zerbinio, D. R., Ogeh, D., Perry, E., Haskell, E., Martin, F. J., Cunningham, F., Riat, H. S., Schuilenburg, H., Sparrow, H., Lavidas, I., Loveland, J. E., To, J. K., Mudge, J., Bhai, J., Taylor, K., Billis, K., Gil, L., Haggerty, L., Gordon, L., Amode, M. R., Ruffier, M., Patricio, M., Laird, M. R., Muffato, M., Nuhn, M., Kostadima, M., Langridge, N., Izuogu, O. G., Achuthan, P., Hunt, S. E., Janacek, S. H., Trevanion, S. J., Hourlier, T., Juettemann, T., Maurel, T., Newman, V., Akanni, W., McLaren, W., Liu, Z., Barrell, D. and Flicek, P. 2017. Ensembl 2018. Nucleic Acids Research, 46:D754-D761.

Fröhling, S., Scholl, C., Levine, R. L., Loriaux, M., Boggon, T. J., Bernard, O. A., Berger, R., Döhner, H., Döhner, K. and Ebert, B. L. 2007. Identification of driver and passenger mutations of FLT3 by high-throughput DNA sequence analysis and functional assessment of candidate alleles. Cancer cell, 12:501-513.

Fuxa, M. and Skok, J. A. 2007. Transcriptional regulation in early B cell development. Current opinion in immunology, 19:129-136.

Gardner, M. J. 1991. Father's occupational exposure to radiation and the raised level of childhood leukemia near the Sellafield nuclear plant. Environmental health perspectives, 94:5.
Giamanco, N. M., Cunningham, B. S., Klein, L. S., Parekh, D. S., Warwick, A. B. and Lieuw, K. 2016. Allopurinol Use During Maintenance Therapy for Acute Lymphoblastic Leukemia Avoids Mercaptopurine-related Hepatotoxicity. *Journal of pediatric hematology/oncology*, 38:147-151.

Gilliland, D. G. and Griffin, J. D. 2002. The roles of FLT3 in hematopoiesis and leukemia. *Blood*, 100:1532-1542.

Griffith, M., Griffith, O. L., Krysiak, K., Skidmore, Z. L., Christopher, M. J., Klco, J. M., Ramu, A., Lamprecht, T. L., Wagner, A. H. and Campbell, K. M. 2016. Comprehensive genomic analysis reveals FLT3 activation and a therapeutic strategy for a patient with relapsed adult B-lymphoblastic leukemia. *Experimental hematology*, 44:603-613.

Iacobucci, I. and Mullighan, C. G. 2017. Genetic basis of acute lymphoblastic leukemia. *Journal of Clinical Oncology*, 35:975.

Kumanayake, P. 2013. Genome-wide SNP discovery in associating with human diseases phenotypes. *Sri Lanka Journal of Bio-Medical Informatics*, 3.

Lang, D., Powell, S. K., Plummer, R. S., Young, K. P. and Ruggeri, B. A. 2007. PAX genes: roles in development, pathophysiology, and cancer. *Biochemical pharmacology*, 73:1-14.

Levine, P. H., Ajmera, K., O’Neill, B., Venkatesh, V., Garcia-Gonzalez, P. and Hoffman, H. J. 2016. Demographic factors related to young age at diagnosis of chronic myeloid leukemia in India. *Clinical Epidemiology and Global Health*, 4:188-192.

Marhäll, A., Heidel, F., Fischer, T. and Rönnstrand, L. 2018. Internal tandem duplication mutations in the tyrosine kinase domain of FLT3 display a higher oncogenic potential than the activation loop D835Y mutation. *Annals of hematology*, 97:773-780.

Matthews, W., Jordan, C. T., Wiegand, G. W., Pardoll, D. and Lemischka, I. R. 1991. A receptor tyrosine kinase specific to hematopoietic stem and progenitor cell-enriched populations. *Cell*, 65:1143-1152.

Nebral, K., Denk, D., Attarbaschi, A., König, M., Mann, G., Haas, O. A. and Strehl, S. 2009. Incidence and diversity of PAX5 fusion genes in childhood acute lymphoblastic leukemia. *Leukemia*, 23:134.

Portell, C. A., Wenzell, C. M. and Advani, A. S. 2013. Clinical and pharmacologic aspects of blinatumomab in the treatment of B-cell acute lymphoblastic leukemia. *Clinical pharmacology: advances and applications*, 5:5.
Pui, C.-H. 2015. Genomic and pharmacogenetic studies of childhood acute lymphoblastic leukemia. *Frontiers of medicine*, **9**:1-9.

Queralt-Rosinach, N., Pinero, J., Bravo, À., Sanz, F. and Furlong, L. I. 2016. DisGeNET-RDF: harnessing the innovative power of the Semantic Web to explore the genetic basis of diseases. *Bioinformatics*, **32**:2236-2238.

Rasool, M., Farooq, S., Malik, A., Shaukat, A., Manan, A., Asif, M., Sani, S., Qazi, M. H., Kamal, M. A. and Iqbal, Z. 2015. Assessment of circulating biochemical markers and antioxidative status in acute lymphoblastic leukemia (ALL) and acute myeloid leukemia (AML) patients. *Saudi journal of biological sciences*, **22**:106-111.

Rosnet, O., Bühring, H., Marchetto, S., Rappold, I., Lavagna, C., Sainty, D., Arnoulet, C., Chabannon, C., Kanz, L. and Hannum, C. 1996. Human FLT3/FLK2 receptor tyrosine kinase is expressed at the surface of normal and malignant hematopoietic cells. *Leukemia*, **10**:238-248.

Santoro, A., Bica, M. G., Dagnino, L., Agueli, C., Salemi, D., Cannella, S., Veltroni, M., Cetica, V., Giarin, E. and Fabbiano, F. 2009. Altered mRNA expression of PAX5 is a common event in acute lymphoblastic leukaemia. *British journal of haematology*, **146**:686-689.

Schebesta, A., McManus, S., Salvagiotto, G., Delogu, A., Busslinger, G. A. and Busslinger, M. 2007. Transcription factor Pax5 activates the chromatin of key genes involved in B cell signaling, adhesion, migration, and immune function. *Immunity*, **27**:49-63.

Shahjahani, M., Norozi, F., Ahmadzadeh, A., Shahrabi, S., Tavakoli, F., Asnafi, A. A. and Saki, N. 2015. The role of Pax5 in leukemia: diagnosis and prognosis significance. *Medical Oncology*, **32**:360.

Sherry, S. T., Ward, M.-H., Kholodov, M., Baker, J., Phan, L., Smigielski, E. M. and Sirotkin, K. 2001. dbSNP: the NCBI database of genetic variation. *Nucleic acids research*, **29**:308-311.

Shu, X. O., Potter, J. D., Linet, M. S., Severson, R. K., Han, D., Kersey, J. H., Neglia, J. P., Trigg, M. E. and Robison, L. L. 2002. Diagnostic X-rays and ultrasound exposure and risk of childhood acute lymphoblastic leukemia by immunophenotype. *Cancer Epidemiology and Prevention Biomarkers*, **11**:177-185.

Small, D., Levenstein, M., Kim, E., Carow, C., Amin, S., Rockwell, P., Witte, L., Burrow, C., Ratajczak, M. Z. and Gewirtz, A. M. 1994. STK-1, the human homolog of Flk-2/Flt-3,
selectively expressed in CD34+ human bone marrow cells and is involved in the proliferation of early progenitor/stem cells. Proceedings of the National Academy of Sciences, 91:459-463.

Souabni, A., Jochum, W. and Busslinger, M. 2007. Oncogenic role of Pax5 in the T-lymphoid lineage upon ectopic expression from the immunoglobulin heavy-chain locus. Blood, 109:281-289.

Steinberg, A. G. and Steinfeld, J. L. 1960. The genetics of acute leukemia in children. Cancer, 13:985-999.

Tasian, S. K. and Hunger, S. P. 2017. Genomic characterization of paediatric acute lymphoblastic leukaemia: an opportunity for precision medicine therapeutics. British journal of haematology, 176:867-882.

Urtishak, K. A., Robinson, B. W., Rappaport, E. F., Sarezky, M. D., Biegel, J. A., Nichols, K. E., Wilmoth, D. M., Wang, L. S., Stern, J. W. and Felix, C. A. 2016. Unique Familial MLL (KMT2A)-Rearranged Precursor B-Cell Infant Acute Lymphoblastic Leukemia in Non-twin Siblings. Pediatric blood & cancer, 63:1175-1180.

Wu, C. and Li, W. 2018. Genomics and pharmacogenomics of pediatric acute lymphoblastic leukemia. Critical reviews in oncology/hematology.

Ye, S., Dhillon, S., Ke, X., Collins, A. R. and Day, I. N. 2001. An efficient procedure for genotyping single nucleotide polymorphisms. Nucleic acids research, 29:e88-e88.

Zhang, S. and Broxmeyer, H. E. 2000. Flt3 ligand induces tyrosine phosphorylation of gab1 and gab2 and their association with shp-2, grb2, and PI3 kinase. Biochemical and biophysical research communications, 277:195-199.

Zriwil, A., Böiers, C., Kristiansen, T. A., Wittmann, L., Yuan, J., Nerlov, C., Sitnicka, E. and Jacobsen, S. E. 2018. Direct role of FLT 3 in regulation of early lymphoid progenitors. British journal of haematology, 183:588-600.

Zwollo, P., Arrieta, H., Ede, K., Molinder, K., Desiderio, S. and Pollock, R. 1997. The Pax-5 gene is alternatively spliced during B-cell development. Journal of Biological Chemistry, 272:10160-10168.
Table 1 (on next page)

Tetra-ARMS primers
| Primer | Sequence (5’-3’) | Tm |
|--------|-----------------|----|
| rs35958982 | TGTGACAAATTAGCAGGGTTAACAC | 57.3 |
|         | CACAGAAGAGATCACAGAAGGAGTCT | 60.7 |
|         | GAAACTCCCCATTTGAGATCATATTCA | 56.0 |
|         | AGACAGAGACAAGCAGACATTCA | 58.4 |
| rs3780135 | CTCTTCCAGGCTCCCCCGAC | 59.2 |
|         | GGGCGGCAGCGCTATAAGAA | 59.5 |
|         | ACCCCAGCTCTAGATGGCGAAG | 56.6 |
|         | ATAGGTCATCAGTGGTGTC | 58.4 |
| rs12430881 | GTTTGCTCCTCCTTCATGGCA | 56.0 |
|         | GCCTCAGTGTCTCTTCGAATT | 56.3 |
|         | CCTTTTATCTTCACATCAGGCCT | 56.6 |
|         | CTAGTATGAGATGGGGTTTGCC | 58.4 |
Table 2 (on next page)

PCR program for SNPs
| PCR steps       | Temperature (°C) | Duration of steps (min) | No. of cycles |
|----------------|------------------|-------------------------|--------------|
| Initial duration | 92               | 5min                    |              |
| Denaturation    | 94               | 30sec                   |              |
| Annealing       |                  |                         |              |
| rs35958982      | 58.4°C           | 1min                    | 30-35        |
| rs3780135       | 56.6°C           |                         |              |
| rs12430881      | 58.8°C           |                         |              |
| Extension       | 72               | 1min                    |              |
| Final extension | 72               | 5min                    |              |
Table 3 (on next page)

Association of demographic factors with risk of B-ALL.

Significant values are shown in (*)
| Parameters                | Patients (%) | Control (%) | Odd Ratio | Chi-square | P value |
|--------------------------|--------------|-------------|-----------|------------|---------|
| Age (mean)               | 7.30         | 11.70       |           |            |         |
| A positive family history| 16.77        | 1.29        | 15.42     | 14.59      | 0.000*  |
| A negative family history| 83.22        | 98.70       |           |            |         |
| Smoking by parent        | 38.06        | 41.94       | 0.85      | 0.31       | 0.580   |
| No smoking parent        | 61.93        | 58.06       |           |            |         |
| Parental cousin marriage | 33.55        | 21.29       | 1.87      | 3.78       | 0.050*  |
| No cousin marriage       | 66.45        | 78.70       |           |            |         |
| Females                  | 56           | 72          | 0.65      | 3.41       | 0.070   |
| Males                    | 99           | 83          |           |            |         |
Table 4 (on next page)

Allele and genotype frequency

Adjusted ORs were obtained from logistic regression model with adjustment for family history, smoking and consanguinity.

1Critical P value, 2Bonferroni corrected P value
| Gene/SNP   | Allele/Genotype | Controls (%) | Cases (%) | Crude OR (95%CI) | Adjusted OR (95%CI) | χ²^-value | ¹P-value | ²P-value |
|-----------|-----------------|--------------|----------|------------------|----------------------|-----------|---------|---------|
| Allele    |                 |              |          |                  |                      |           |         |         |
| A         | 5.60            | 13.40        | 2.30     | ---              |                      | 7.79      | 0.005*  | 0.002*  |
| G         | 94.40           | 86.60        |          | (1.20-4.50)      |                      |           |         |         |
| Genotype  |                 |              |          |                  |                      |           |         |         |
| rs35958982|                 |              |          |                  |                      |           |         |         |
| GG        | 90.70           | 79.60        |          | 1.00             |                      |           |         |         |
| GA        | 7.60            | 13.80        | 3.67     | (0.75-18.10)     | 1.13                 | 2.90      | 0.088   | 0.030*  |
| AA        | 1.85            | 6.50         |          | 1.30             | (0.37-5.08)          |           |         |         |
| Allele    |                 |              |          |                  |                      |           |         |         |
| A         | 32.58           | 47.10        | 2.17     | ---              |                      | 13.63     | 0.000*  | 0.000*  |
| G         | 67.42           | 52.90        |          | (1.37-3.43)      |                      |           |         |         |
| Genotype  |                 |              |          |                  |                      |           |         |         |
| rs3780135 |                 |              |          |                  |                      |           |         |         |
| GG        | 52.26           | 33.55        |          | 1.00             |                      |           |         |         |
| GA        | 30.32           | 38.70        | 0.55     | (0.39-0.95)      | 1.19                 | 4.72      | 0.029*  | 0.010*  |
| AA        | 17.42           | 27.74        |          | 0.97             | (0.26-1.42)          |           |         |         |
| Allele    |                 |              |          |                  |                      |           |         |         |
| G         | 22              | 29           | 1.15     | ---              |                      | 4.11      | 0.043*  | 0.014*  |
| A         | 78              | 71           | (0.72-1.83)| ---              |                      |           |         |         |
| rs12430881| Genotype        |              |          |                  |                      |           |         |         |
| AA        | 65.16           | 61.93        | 2.52     | 1.00             |                      |           |         |         |
| GA        | 25.80           | 18.06        |          | 1.09             | (0.48-2.69)          |           |         |         |
| GG | 9.03 | 20 | 1.03 (0.50-2.68) |
|----|------|----|-----------------|
| 2  |      |    |                 |
| 3  |      |    |                 |
| 4  |      |    |                 |
| 5  |      |    |                 |
| 6  |      |    |                 |
| 7  |      |    |                 |
| 8  |      |    |                 |
| 9  |      |    |                 |
| 10 |      |    |                 |
| 11 |      |    |                 |
| 12 |      |    |                 |
Table 5 (on next page)

Stratification analysis for association between genotypes and risk of B-ALL

ORs were obtained from logistic regression model with adjustment for family history, smoking and consanguinity.
| rs35958982 | OR (95% CI) | Family history status | Smoking status | Consanguinity status |
|---|---|---|---|---|
| | GG | GA | AA |
| | 1 | 0.92(0.18-4.56) | 0.50(0.06-4.15) |
| Yes | | | |
| No | | | |
| | 1 | 0.30(0.08-1.09) | 0.46(0.10-2.12) |
| Yes | | | |
| No | | | |
| | 1 | 0.86(0.24-3.10) | 0.94(0.20-4.37) |
| Yes | | | |
| No | | | |
| | 1 | 0.60(1.70-0.21) | 0.44(0.15-1.30) |
| Yes | | | |
| No | | | |
| | 1 | 1.08(0.46-2.50) | 1.3(0.59-2.90) |
| Yes | | | |
| No | | | |
| | 1 | 0.62(0.26-1.44) | 0.60(0.26-1.35) |
| Yes | | | |
| No | | | |
| | 1 | 0.58(0.16-2.15) | 1.80(0.65-4.98) |
| Yes | | | |
| No | | | |
| | 1 | 1.05(0.45-2.42) | 1.08(0.45-2.60) |
| Yes | | | |
| No | | | |
| | 1 | 1.15(0.50-2.65) | 0.49(0.18-1.35) |
| Yes | | | |
| No | | | |
| | 1 | 0.60(1.70-0.21) | 0.44(0.15-1.30) |
| Yes | | | |
| No | | | |
| | 1 | 1.08(0.46-2.50) | 1.3(0.59-2.90) |
| Yes | | | |
| No | | | |
| | 1 | 0.62(0.26-1.44) | 0.60(0.26-1.35) |
| Yes | | | |
| No | | | |
| | 1 | 0.58(0.16-2.15) | 1.80(0.65-4.98) |
| Yes | | | |
| No | | | |
| | 1 | 1.05(0.45-2.42) | 1.08(0.45-2.60) |
| Yes | | | |
| No | | | |
| | 1 | 1.15(0.50-2.65) | 0.49(0.18-1.35) |
| Yes | | | |
| No | | | |
Figure 1

SNP rs35958982

Well 1 indicates the DNA ladder (100bp), an amplicon (395bp) is outer band. Amplicon 272bp: allele ‘A’ and amplicon of 170bp allele ‘G’.
Figure 2

SNP rs3780135

Well 1 indicates the DNA ladder (50bp), an amplicon (331bp) is outer band. Amplicon 243bp: allele ‘A’ and amplicon of 128bp: allele ‘G’
Manuscript to be reviewed
Figure 3

SNP rs12430881

Well 1 indicates the DNA ladder (100bp), an amplicon (400bp) is outer band. Amplicon 165bp: allele G and amplicon of 280bp: allele A.
