

Non-coding RNAs-related polymorphisms in pediatric acute lymphoblastic leukemia susceptibility

Running title: MicroRNA-related SNPs in ALL

Angela Gutierrez-Camino¹, Elixabet Lopez-Lopez ¹, Idoia Martin-Guerrero ¹, Maria A. Piñan², Purificacion Garcia-Miguel ³, Jose Sanchez-Toledo ⁴, Ana Carbone Bañeres⁵, Javier Uriz⁶, Aurora Navajas⁷, Africa Garcia-Orad¹.

AGC and ELL contributed equally to this work.

¹Department of Genetics, Physical Anthropology and Animal Physiology, University of the Basque Country (UPV/EHU), Leioa, Spain

²Department of Hematology and Hemotherapy, University Hospital Cruces, Bilbao, Spain

³Service of Pediatric Oncohematology, University Hospital La Paz, Madrid, Spain

⁴Service of Pediatric Oncology and Hematology, University Hospital Vall d' Hebron, VHIR, Barcelona, Spain

⁵Unit of Pediatric Oncohematology, University Hospital Miguel Servet, Zaragoza, Spain

⁶Unit of Pediatric Oncohematology, University Hospital Donostia, San Sebastian, Spain

⁷Unit of Pediatric Hematology/Oncology, University Hospital Cruces, Bilbao, Spain

Corresponding Author: Africa Garcia-Orad. Department of Genetics, Physical Anthropology and Animal Physiology, Faculty of Medicine and Dentistry-University of the Basque Country, Barrio Sarriena s/n, 48940 Leioa, Spain. Phone: international +34.946012909. Fax: international +34.946013400. E-mail: africa.garciaorad@ehu.es

Financial Support: AGC was supported by a predoctoral grant from the Gangoiti Barrera Foundation, Bilbao, Spain. ELL was supported by a predoctoral grant of the Basque Government and "Fellowship for recent Doctors until their integration in postdoctoral

programs” by the Investigation Vice-rector’s office of the University of Basque Country (UPV/EHU). This project was supported by Spanish Thematic Network of Cooperative Research in Cancer (RTIC) (RD/06/0020/0048), Basque Government (IT661-13, S-PE12UN060) and UPV/EHU (UFI11/35). Support by SGIker (University of the Basque Country, UPV/EHU) is gratefully acknowledged.

Conflict of interest disclosure statement: The authors have no conflicts of interest to disclose.

Category of study: Clinical Research

Abstract

Background: Evidence for an inherited genetic risk for pediatric acute lymphoblastic leukemia has been provided in several studies. Most of them focused on coding regions. However, those regions represent only 1.5% of the entire genome. In ALL, it has been suggested that the expression of microRNAs is dysregulated, which suggests that they may have a role in ALL risk. Changes in miRNA function may occur through SNPs. Therefore, the aim of this study was to evaluate whether polymorphisms in pre-miRNAs, and/or miRNA processing genes, contribute to a predisposition for childhood ALL.

Methods: In this study, we analyzed 118 SNPs in pre-miRNAs and miRNA processing genes in 213 B-ALL patients and 387 controls.

Results: We found 11 SNPs significantly associated with ALL susceptibility. These included three SNPs present in miRNA genes (*miR-612*, *miR-499*, and *miR-449b*) and eight SNPs present in six miRNA biogenesis pathway genes (*TNRC6B*, *DROSHA*, *DGCR8*, *EIF2C1*, *CNOT1*, and *CNOT6*). Among the 118 SNPs analyzed, rs12803915 in mir-612 and rs3746444 in mir-499 exhibited a more significant association, with a p value <0.01.

Conclusion: The results of this study indicate that SNP rs12803915 located in pre-mir-612, and SNP rs3746444 located in pre-mir-499, may represent novel markers of B-ALL susceptibility.

Introduction

Acute lymphoblastic leukemia (ALL) is the most common pediatric hematological malignancy in developed countries. Its etiology is believed to be multifactorial, with both environmental and genetic risk factors being relevant ¹. Recently, several studies have provided evidence for an inherited genetic risk for pediatric ALL ^{2,3}. Most of these studies focused on the coding regions of these genetic components. However, this represents only ~1.5% of the entire genome, and non-coding regions of the genome have also been shown to mediate regulatory functions. For example, microRNAs (miRNAs) are a class of small non-coding RNA molecules that regulate gene expression at the post-transcriptional level by binding to the 3' untranslated region (UTR) of a target gene ⁴. This can lead to an inhibition of translation or enhanced degradation of a target mRNA (Figure 1). Primary double-stranded miRNA transcripts (pri-miRNA) are processed in the nucleus by microprocessor machinery which includes *DROSHA* RNase and the double-stranded RNA binding protein, DGCR8. A hairpin precursor miRNA molecule of 70-100 nucleotides (pre-miRNA) is then produced, and its translocation into the cytoplasm is facilitated by RAN GTPase and Exportin 5 (*XPO5*). In the cytoplasm, pre-miRNAs are further processed by a protein complex that includes *DICER1*, *TRBP*, *EIF2C1*, *EIF2C2*, *GEMIN3*, and *GEMIN4*, resulting in the production of mature miRNAs ⁴. It has been predicted that there are more than 1000 miRNA genes in the human genome ⁵ and approximately 30% of human genes are regulated by miRNAs.

In the past few years, it was suggested that miRNAs in ALL are dysregulated. For example, in the study of Zhang et al., up to 171 miRNAs have been found to be differentially expressed between ALL patients and normal donors. These results suggest that dysregulation of these miRNAs may be associated with an increased risk for ALL. Changes in miRNAs function have the potential to affect the expression of a large number of genes, including genes involved in the origin and evolution of pediatric ALL ^{6,7}. Changes in miRNA function may occur through

genetic variations⁴. For example, single nucleotide polymorphisms (SNPs) present in genes involved in miRNA processing can affect levels of miRNA expression, while SNPs in miRNA genes can affect miRNA biogenesis and function. There have been several polymorphisms found to be associated with other malignancies, and a recent pilot study has found an association between rs2910164 in mir-146a and ALL risk⁸. However, despite accumulating evidence that inherited genetic variation can contribute to a predisposition for pediatric ALL, the suggested role of miRNAs in the development of this disease, as well as the role of miRNA related polymorphisms in cancer risk, SNPs in miRNA genes and miRNA processing genes have not been extensively studied in association with pediatric ALL risk.

Therefore, the aim of this study was to evaluate whether polymorphisms in pre-miRNAs, and/or miRNA processing genes, contribute to a predisposition for childhood ALL.

Results

Genotyping Results

A total of 213 patients with B-ALL and 387 unrelated healthy controls were available for genotyping. Successful genotyping was achieved for 550 DNA samples (91.67%). Of the SNPs, 106/118 (89.83%) were genotyped satisfactorily. Failed genotyping was due to an absence of PCR amplification, insufficient intensity for cluster separation, poor cluster definition, or an inability to define clusters. The average genotyping rate for all SNPs was 98.12%. Furthermore, of the 106 SNPs genotyped, 14 were not in Hardy Weinberg equilibrium (HWE) in the population of 387 healthy controls, and therefore, were not considered for further analysis. In total, 26 SNPs were excluded from the association study (Supplemental Table S1 (online)), leaving 92 SNPs available for association studies.

Analysis of association

To investigate if genetic variation influences the risk of ALL, the 92 polymorphisms successfully genotyped were compared between cases and controls. As shown in Tables 1 and 2, statistically significant associations ($p < 0.05$) were observed for 11 polymorphisms present in miRNA-related genes. Of these, three were located in pre-miRNAs (Table 1) and eight were located in miRNA processing genes (Table 2).

Among the SNPs located in pre-miRNA genes, the SNP that was the most significantly associated with ALL risk was SNP rs12803915 in pre-mature *mir-612* (Figure 2A). The A allele for this SNP was found to be protective (Odd ratio (OR) 0.61; 95% CI: 0.42–0.88; $p = 0.007$) in the log-additive (GG vs. GA vs. AA) genetic model. The second most significant association involved SNP rs3746444 in *mir-499* (Figure 2B). Moreover, this SNP is located in the seed region of mature miR-499-3p. The G allele of this SNP was found to be protective (OR 0.67, 95% CI: 0.49–0.91; $p = 0.009$) in the log-additive (AA vs. AG vs. GG) genetic model. An association with rs10061133 in *mir-449b* was also identified (Table 1).

These two top SNPs, rs12803915 and rs3746444, were also studied in association with B-hyperdiploid ALL subtype (Supplemental table S2 and S3 (online)) and chromosomal translocations (data not shown) and no significant differences were found.

In the genes of the miRNA processing components, the most significant SNP was rs139919, a SNP located in the *TNRC6B* gene. In the recessive genetic model, the variant homozygous genotype, CC, of this SNP was associated with a 5.1-fold increase in ALL risk (95% CI: 1.31–19.11; $p = 0.011$). Associations involving rs9606248 and rs1640299 in *DGCR8*, rs11866002 in *CNOT1*, rs6877400 in *CNOT6*, rs10035440 in *DROSHA*, and rs636832 and rs595961 in *EIF2C1* were also identified (Table 2).

All the SNPs remained significantly associated with ALL risk after multivariate logistic regression to account for the possible confounding effect of sex. These SNPs did not reach the significant value when Bonferroni correction was applied.

Discussion

In this study, 11 SNPs were found to be significantly associated with ALL susceptibility. These included three SNPs present in miRNA genes (*mir-612*, *mir-499*, and *mir-449b*) and eight SNPs present in six miRNA biogenesis pathway genes (*TNRC6B*, *DROSHA*, *DGCR8*, *EIF2C1*, *CNOT1*, and *CNOT6*). Among them, rs12803915 in *mir-612* and rs3746444 in *mir-499* exhibited a more significant association, with a p value < 0.01. In spite of not reaching a significant p value after the restrictive Bonferroni correction, our results point to a putative role of these SNPs in ALL susceptibility, which could be of lower penetrance.

SNP rs12803915 located in the premature region of *mir-612* showed the strongest association with ALL risk, with the A allele being protective (p= 0.007). To our knowledge, this is the first report of this SNP being associated with cancer risk. It has been suggested that a SNP in the premature region of a miRNA could alter its secondary structure and inhibit or enhance pri-microRNA processing⁹. Therefore, SNPs in miRNAs could lead to dysregulation of miRNA expression⁴. Accordingly, it has been shown that SNP rs12803915 significantly decreases mature miR-612 levels, and this may represent a mechanism by which cancer risk is increased¹⁰. One of the potential targets of *mir-612* is IKZF2¹¹, a member of the Ikaros family of zinc-finger proteins. This protein is a hematopoietic-specific transcription factor involved in the regulation of lymphocyte development and other members of this family have been associated with ALL susceptibility^{2,3}. Therefore, the presence of SNP rs12803915 in *pre-mir-612* may

contribute to an increased susceptibility to ALL based on its capacity to affect expression of mir-612, as well as its downstream targets, such as IKZF2.

The second most significant association identified in this study involved SNP rs3746444 ($p=0.009$) located in *pre-mir-499*. In this case, the G allele was associated with a lower risk for ALL. This SNP has recently been implicated in the etiology of several types of cancer with controversial results. The results of the present study are consistent with those of three previous studies, where the rs3746444 GG genotype was shown to be associated with a decreased risk of cancer¹²⁻¹⁴. By contrast in six other studies^{5,15-19}, the GG genotype was found to be associated with a higher risk of cancer development. There have also been studies that have found no association between this SNP and cancer risk^{8,20-24}. The inconsistency of these results may be due to differences in the carcinogenic mechanisms of different cancers, as well as differences in genetic backgrounds²⁵. Correspondingly, a meta-analysis observed that a tendency for reduced cancer risk was associated with the *mir-499* rs3746444 GG genotype in Caucasian populations²⁶.

This SNP, rs3746444, is located in *pre-mir-499*, in the premature sequences of mir-499-5p and mir-499-3p, and in the seed sequence of mir-499-3p. Therefore, it could have a double effect. On the one hand, a SNP located in a pre-miRNA region has the potential to impact the processing of pre-miRNAs into mature miRNAs¹⁴. Correspondingly, different genotypes of rs3746444 in *pre-mir-499* have been associated with dysregulated expression of mir-499-5p in a colorectal cancer model⁵. These data are of particular interest considering that mir-499-5p seems to be upregulated in ALL⁶. On the other hand, SNP rs3746444 is also present in the seed region of a mature miRNA-499-3p, and could be essential for the accurate recognition of target mRNA sequences. Potential targets of mir-499-3p¹¹ include FOXO1A (a transcription factor which is dysregulated in B-ALL²⁷), MS4A1 or CD20 (a B-lymphocyte surface molecule which

plays a role in the development and differentiation of B-cells into plasma cells), and PBX1 (pre-B-cell leukemia homeobox 1, which is a dysregulated transcription factor in ALL)²⁸. Therefore, alterations in the sequence of mir-499-5p and mir-499-3p which affect expression of these miRNAs, and/or the binding of mir-499-3p to target mRNAs, may have functional consequences for ALL.

Among the SNPs located in pre-miRNA genes, other interesting result was that the rs10061133 G allele in *mir-449b* was associated with a decreased risk of ALL ($p = 0.012$). To our knowledge, this is the first time that this polymorphism has been associated with cancer risk. Based on its location, this polymorphism is also present in the mature sequence of the miRNA; therefore, it could affect the strength of miRNA-mRNA binding, as well as miRNA levels. This miRNA has also been found to be upregulated in endometrial cancer and bladder cancer^{29,30}.

However, in this larger population, we have not replicated the association previously found in a small pilot study between rs2910164 in mir-146a and ALL risk⁸.

When we analyzed 72 polymorphisms present in miRNA biogenesis pathway, eight of these were found to be significantly associated with ALL risk. Polymorphism rs139919 in *TNRC6B* was the SNP most highly associated with ALL susceptibility among the miRNA processing genes, with the CC genotype associated with an increase in ALL risk ($p = 0.011$). *TNRC6B* encodes an RNA interference (RNAi) machinery component which contributes to the RISC complex, and is crucial for miRNA-dependent translational repression or degradation of target messenger RNAs³¹. Although this SNP has not previously been analyzed, other genetic variants in *TNRC6B* have been associated with prostate cancer risk^{31,32}. It has also been suggested that alterations in expression of *TNRC6B* are due to genetic variations that may affect mRNA levels that are normally regulated by *TNRC6B*, thereby affecting carcinogenesis³².

We have also found other SNPs located in genes which produce proteins that contribute to the RISC complex associated with ALL susceptibility. These included two SNPs in *EIF2C1* (rs595961 and rs636832), one SNP in *CNOT1* (rs11866002), and one SNP in *CNOT6* (rs6877400). The SNPs rs595961 and rs636832 in *EIF2C1* have previously been associated with renal cell carcinoma risk in males³³ and lung cancer³⁴, respectively. However, to our knowledge, this is the first study to identify an association between the SNPs, rs11866002 in *CNOT 1* and rs6877400 in *CNOT6*, and cancer risk. In addition, it is hypothesized that rs595961, rs11866002, and rs6877400 have putative roles in transcription and/or regulation of splicing events, thereby affecting gene expression. SNPs that affect expression levels of these proteins may have deleterious effects on miRNA-mRNA interactions, and may affect cancer development and progression. Correspondingly, *EIF2C1* is frequently lost in human cancers such as Wilms tumor, neuroblastoma, and carcinomas of the breast, liver, and colon³⁵.

Among the genes of biogenesis machinery, when we studied genes that contribute to the processing of pri-miRNAs to pre-miRNAs, we identified an association between SNPs and ALL risk. These SNPs included rs10035440 in *DROSHA*, and rs9606248 and rs1640299 in *DGCR8*. Interestingly, these three SNPs have putative roles in transcriptional regulation, and may affect the expression levels of *DROSHA* or *DGCR8*. As a result, levels of miRNAs could be affected, leading to an increased ALL risk. Consistent with this hypothesis, a differentially expressed miRNA profile and *DROSHA* gene expression have been observed in relation to another SNP, rs640831, being present in lung tissue³⁶. Furthermore, increased expression of *DROSHA* and *DGCR8* has been shown to dysregulate miRNAs present in the pleomorphic adenomas of the salivary gland³⁷. In contrast, decreased expression of *DROSHA* and *DGCR8* have been shown to accelerate cellular transformation and tumorigenesis³⁸. Surprisingly, the association of these polymorphisms with the risk to develop other tumors has previously been analyzed^{25,33,39} and

no significant association was found. However, they had not been analyzed in ALL patients until now.

In conclusion, the results of this study indicate that SNP rs12803915 located in *pre-mir-612*, and SNP rs3746444 located in *pre-mir-499*, may represent novel markers of B-ALL susceptibility. It would be of great interest to confirm these results in different cohorts of patients. To our knowledge, this is the first extensive study to report miRNA-related SNPs associated with ALL risk.

Materials and methods

Study Participants

A total of 213 children (1-15 years) of European origin all diagnosed with precursor B-cell ALL (B-ALL) in the Pediatric Oncology Units of five Spanish hospitals (Hospital Cruces; Hospital Donostia; Hospital Vall d'Hebrón, Hospital La Paz, and Hospital Miguel Servet) were enrolled in this study. These patients were the entire incident population diagnosed and treated in the participating centers between 1995 and 2011. In addition, 387 Spanish healthy individuals of European origin with no previous history of cancer from the collection C.0001171 registered in The Institute of Health Carlos III (ISCIII) were enrolled as controls (Table 3). Patients were classified with B-hyperdiploid ALL if the DNA index was > 1.16 and/or the karyotype had more than 50 chromosomes. For 51 cases out of 213 B-ALL patients, cytogenetic data were not available. Informed consent was obtained from all participants, or from their parents prior to sample collection. The study was approved by The Ethic Committee for Clinical Research (CEIC-E) and conducted in accordance with the Declaration of Helsinki.

Selection of genes and polymorphisms

Twenty-one genes involved in miRNA biogenesis and processing, as determined based on available literature and the Patrocles (<http://www.patrocles.org/>) (University of Liege, Liège, Belgium) database, were selected for analysis. For each gene, all of the SNPs with potential functional effects were examined using F-SNP (<http://compbio.cs.queensu.ca/F-SNP/>) (Queen's University, Kingston, Canada), Fast-SNP (<http://fastsnp.ibms.sinica.edu.tw>)(Academia Sinica, Taipei, Taiwan), polymirTS (<http://compbio.uthsc.edu/miRSNP/>)(University of Tennessee Health Science Center, Memphis, TN) and Patrocles databases. Functional effects were considered to be those that resulted in amino acid changes and/or alternate splicing, those that were located in the promoter region of putative transcription factor binding sites, or those that disrupted or created miRNAs binding interactions. SNPs previously included in association studies were also examined. The final selection of SNPs was made based on those having a minor allele frequency greater than 5% ($MAF \geq 0.05$) in European/Caucasian populations.

Considering that miRNAs can regulate a wide range of genes and that the number of polymorphisms in miRNAs was affordable, we selected all the known SNPs at the moment of the selection with a $MAF > 0.01$ in European/Caucasoid populations, using Patrocles and Ensembl (<http://www.ensembl.org/>)(Wellcome Trust Genome Campus, Cambridge, UK) databases and literature review.

Genotyping

Genomic DNA was extracted with the phenol-chloroform method previously described⁴⁰ from remission (containing less than 5% blast cells) peripheral blood, bone marrow slides or granulocytes isolated with Ficoll-Plaque PLUS (GE Healthcare Life Sciences, Piscataway, NJ), or from saliva samples using Oragene DNA kit (DNA Genotek, Inc., Ottawa, Canada) according to the manufacturer's instructions.

SNP genotyping was performed using TaqMan OpenArray Genotyping technology (Applied Biosystems, Life Technologies, Carlsbad, CA) according to the published Applied Biosystems protocol. Initially, 131 SNPs were considered for analysis. After considering compatibility with the Taqman OpenArray platform, 118 SNPs were included in a Taqman OpenArray Plate (Applied Biosystems), and these included 72 SNPs present in 21 genes involved in miRNA biogenesis and 46 SNPs present in 42 pre-miRNA genes (Supplemental Table S4 and S5 (online))

Data were analyzed using Taqman Genotyper software (Applied Biosystems) for genotype clustering and genotype calling. Duplicate samples were genotyped across the plates.

Statistical analysis

Statistical analyses were performed using R software (version v2.14.1) (Institute for Statistics and Mathematics, Wien, Austria). To identify any deviation in HWE for the healthy controls (n = 387), a χ^2 test was used. The association between genetic polymorphisms in B-ALL patients and controls was also evaluated using the χ^2 or Fisher's exact test. Fisher's exact test was used if a genotype class had less than five individuals. We also tested the association considering genetic characteristics (hyperdiploid subtype and chromosomal translocations). The effect sizes of the associations were estimated by the OR from univariate logistic regression and multivariate logistic regression to account for the possible confounding effect of sex.

The most significant test among codominant (major allele homozygotes vs. heterozygotes and major allele homozygotes vs. minor allele homozygotes), dominant (major allele homozygotes vs. heterozygotes+minor allele homozygotes), recessive (major allele homozygotes+heterozygotes vs. minor allele homozygotes), and additive (doses-dependent

effect: major allele homozygotes vs. heterozygotes vs. minor allele homozygotes) genetic models was selected. In all cases, the significance level was set at 5%. The results were adjusted for multiple comparisons using the Bonferroni correction.

References

1. Ellinghaus E, Stanulla M, Richter G, et al. Identification of germline susceptibility loci in ETV6-RUNX1-rearranged childhood acute lymphoblastic leukemia. *Leukemia* 2012;26:902-9.
2. Treviño LR, Yang W, French D, et al. Germline genomic variants associated with childhood acute lymphoblastic leukemia. *Nat Genet* 2009;41:1001-5.
3. Papaemmanuil E, Hosking FJ, Vijayakrishnan J, et al. Loci on 7p12.2, 10q21.2 and 14q11.2 are associated with risk of childhood acute lymphoblastic leukemia. *Nat Genet* 2009;41:1006-10.
4. Ryan BM, Robles AI, Harris CC. Genetic variation in microRNA networks: the implications for cancer research. *Nat Rev Cancer* 2010;10:389-402.
5. Vinci S, Gelmini S, Mancini I, et al. Genetic and epigenetic factors in regulation of microRNA in colorectal cancers. *Methods* 2013;59:138-46.
6. Zhang H, Yang JH, Zheng YS, et al. Genome-wide analysis of small RNA and novel MicroRNA discovery in human acute lymphoblastic leukemia based on extensive sequencing approach. *PLoS One* 2009;4:e6849.
7. Schotte D, Akbari Moqadam F, Lange-Turenhout EA, et al. Discovery of new microRNAs by small RNAome deep sequencing in childhood acute lymphoblastic leukemia. *Leukemia* 2011;25:1389-99.
8. Hasani SS, Hashemi M, Eskandari-Nasab E, Naderi M, Omrani M, Sheybani-Nasab M. A functional polymorphism in the miR-146a gene is associated with the risk of childhood acute lymphoblastic leukemia: a preliminary report. *Tumour Biol* 2013.
9. Salzman DW, Weidhaas JB. SNPing cancer in the bud: microRNA and microRNA-target site polymorphisms as diagnostic and prognostic biomarkers in cancer. *Pharmacol Ther* 2013;137:55-63.

10. Kim HK, Prokunina-Olsson L, Chanock SJ. Common genetic variants in miR-1206 (8q24.2) and miR-612 (11q13.3) affect biogenesis of mature miRNA forms. *PLoS One* 2012;7:e47454.
11. Friedman RC, Farh KK, Burge CB, Bartel DP. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res* 2009;19:92-105.
12. Kim WH, Min KT, Jeon YJ, et al. Association study of microRNA polymorphisms with hepatocellular carcinoma in Korean population. *Gene* 2012;504:92-7.
13. Liu Z, Li G, Wei S, et al. Genetic variants in selected pre-microRNA genes and the risk of squamous cell carcinoma of the head and neck. *Cancer* 2010;116:4753-60.
14. Ahn DH, Rah H, Choi YK, et al. Association of the miR-146aC>G, miR-149T>C, miR-196a2T>C, and miR-499A>G polymorphisms with gastric cancer risk and survival in the Korean population. *Mol Carcinog* 2012.
15. Xiang Y, Fan S, Cao J, Huang S, Zhang LP. Association of the microRNA-499 variants with susceptibility to hepatocellular carcinoma in a Chinese population. *Mol Biol Rep* 2012;39:7019-23.
16. Zhou B, Wang K, Wang Y, et al. Common genetic polymorphisms in pre-microRNAs and risk of cervical squamous cell carcinoma. *Mol Carcinog* 2011;50:499-505.
17. Hu Z, Chen J, Tian T, et al. Genetic variants of miRNA sequences and non-small cell lung cancer survival. *J Clin Invest* 2008;118:2600-8.
18. Alshatwi AA, Shafi G, Hasan TN, et al. Differential expression profile and genetic variants of microRNAs sequences in breast cancer patients. *PLoS One* 2012;7:e30049.
19. George GP, Gangwar R, Mandal RK, Sankhwar SN, Mittal RD. Genetic variation in microRNA genes and prostate cancer risk in North Indian population. *Mol Biol Rep* 2011;38:1609-15.

20. Akkiz H, Bayram S, Bekar A, Akgöllü E, Üsküdar O. Genetic variation in the microRNA-499 gene and hepatocellular carcinoma risk in a Turkish population: lack of any association in a case-control study. *Asian Pac J Cancer Prev* 2011;12:3107-12.
21. Catucci I, Yang R, Verderio P, et al. Evaluation of SNPs in miR-146a, miR196a2 and miR-499 as low-penetrance alleles in German and Italian familial breast cancer cases. *Hum Mutat* 2010;31:E1052-7.
22. Min KT, Kim JW, Jeon YJ, et al. Association of the miR-146aC>G, 149C>T, 196a2C>T, and 499A>G polymorphisms with colorectal cancer in the Korean population. *Mol Carcinog* 2012;51 Suppl 1:E65-73.
23. Okubo M, Tahara T, Shibata T, et al. Association between common genetic variants in pre-microRNAs and gastric cancer risk in Japanese population. *Helicobacter* 2010;15:524-31.
24. Tian T, Shu Y, Chen J, et al. A functional genetic variant in microRNA-196a2 is associated with increased susceptibility of lung cancer in Chinese. *Cancer Epidemiol Biomarkers Prev* 2009;18:1183-7.
25. Yang H, Dinney CP, Ye Y, Zhu Y, Grossman HB, Wu X. Evaluation of genetic variants in microRNA-related genes and risk of bladder cancer. *Cancer Res* 2008;68:2530-7.
26. Qiu MT, Hu JW, Ding XX, et al. Hsa-miR-499 rs3746444 polymorphism contributes to cancer risk: a meta-analysis of 12 studies. *PLoS One* 2012;7:e50887.
27. Andersson A, Ritz C, Lindgren D, et al. Microarray-based classification of a consecutive series of 121 childhood acute leukemias: prediction of leukemic and genetic subtype as well as of minimal residual disease status. *Leukemia* 2007;21:1198-203.
28. Li Z, Zhang W, Wu M, et al. Gene expression-based classification and regulatory networks of pediatric acute lymphoblastic leukemia. *Blood* 2009;114:4486-93.
29. Chung TK, Lau TS, Cheung TH, et al. Dysregulation of microRNA-204 mediates migration and invasion of endometrial cancer by regulating FOXC1. *Int J Cancer* 2012;130:1036-45.

30. Catto JW, Miah S, Owen HC, et al. Distinct microRNA alterations characterize high- and low-grade bladder cancer. *Cancer Res* 2009;69:8472-81.
31. Tao S, Wang Z, Feng J, et al. A genome-wide search for loci interacting with known prostate cancer risk-associated genetic variants. *Carcinogenesis* 2012;33:598-603.
32. Sun J, Zheng SL, Wiklund F, et al. Sequence variants at 22q13 are associated with prostate cancer risk. *Cancer Res* 2009;69:10-5.
33. Horikawa Y, Wood CG, Yang H, et al. Single nucleotide polymorphisms of microRNA machinery genes modify the risk of renal cell carcinoma. *Clin Cancer Res* 2008;14:7956-62.
34. Kim JS, Choi YY, Jin G, et al. Association of a common AGO1 variant with lung cancer risk: a two-stage case-control study. *Mol Carcinog* 2010;49:913-21.
35. Koesters R, Adams V, Betts D, et al. Human eukaryotic initiation factor EIF2C1 gene: cDNA sequence, genomic organization, localization to chromosomal bands 1p34-p35, and expression. *Genomics* 1999;61:210-8.
36. Rotunno M, Zhao Y, Bergen AW, et al. Inherited polymorphisms in the RNA-mediated interference machinery affect microRNA expression and lung cancer survival. *Br J Cancer* 2010;103:1870-4.
37. Zhang X, Cairns M, Rose B, et al. Alterations in miRNA processing and expression in pleomorphic adenomas of the salivary gland. *Int J Cancer* 2009;124:2855-63.
38. Kumar MS, Lu J, Mercer KL, Golub TR, Jacks T. Impaired microRNA processing enhances cellular transformation and tumorigenesis. *Nat Genet* 2007;39:673-7.
39. Sung H, Lee KM, Choi JY, et al. Common genetic polymorphisms of microRNA biogenesis pathway genes and risk of breast cancer: a case-control study in Korea. *Breast Cancer Res Treat* 2011;130:939-51.
40. Sambrook J, Russell DW. Preparation and Analysis of Eukaryotic Genomic DNA. *Molecular cloning: A laboratory manual*. 3rd edition ed. Cold Spring Harbor, USA: CSHL Press; 2001:4-12.

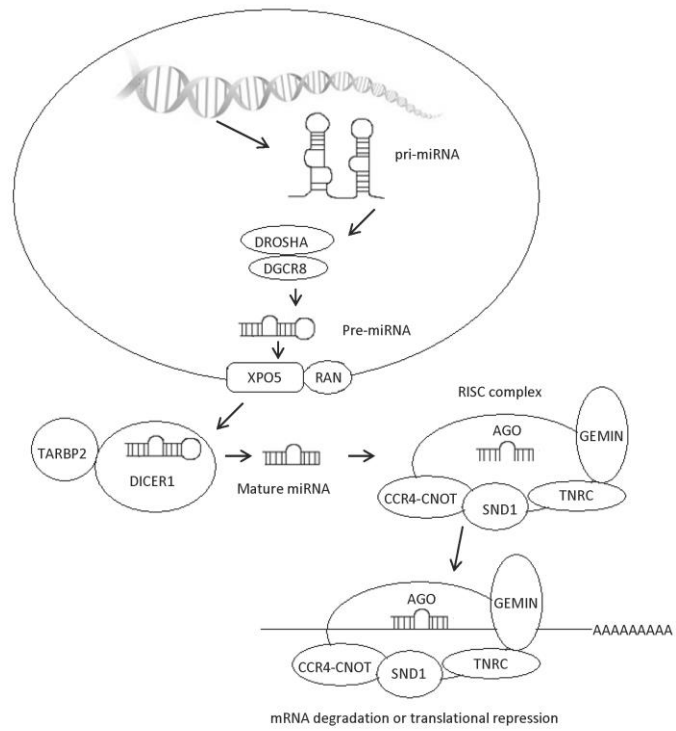
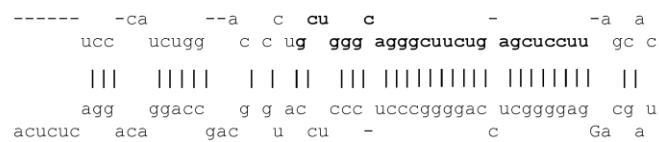
Figure legends**Figure 1.** MicroRNA biogenesis pathway.

Figure 2. Schematic diagram of the hairpin loop structure of the mir-612 and mir-499.

The sequence for mature miRNA is in bold. The polymorphisms sites are indicated in capital letters. (A) The A>G polymorphism is located in the loop of mir-612, this variation might alter the secondary structure (B) The G>A Polymorphism is located in the stem region opposite to the mature mir-499 sequence which results in a change from A:U pair to G:U mismatch in the stem structure of mir-499 precursor.

A miR-612



B miR-499

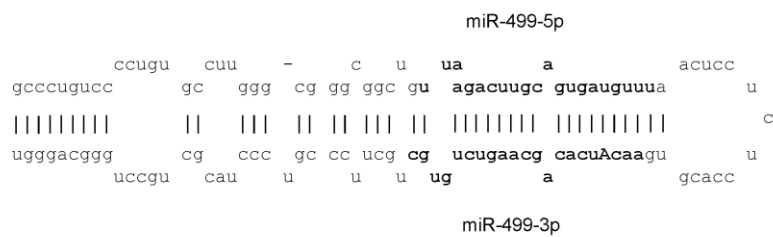


Table 1: Genotype frequencies of selected SNPs present in miRNA genes.

miRNA	SNP	Best fitting model	Genotype	Controls n (%)	Cases n (%)	OR (CI 95%)	p value	p value ^a
<i>mir-612</i>	rs12803915	Additive	GG	232 (68)	152 (78)	0.61 (0.42–0.88)	0.007	0.007
			GA	100 (29)	42 (21)			
			AA	11 (3)	2 (1)			
<i>mir-499</i>	rs3746444	Additive	AA	206 (59)	138 (69)	0.67 (0.49–0.91)	0.009	0.010
			AG	117 (41)	56 (28)			
			GG	24 (7)	6 (3)			
<i>mir-449b</i>	rs10061133	Dominant	AA	283 (82)	180 (90)	Reference	0.012	0.015
			GA/GG	63 (18)	21 (10)	0.52 (0.31 –0.89)		

^a calculated by multivariate logistic regressions to account for the possible confounding effect of sex

Table 2: Genotype frequencies of selected SNPs present in miRNA processing genes.

Gene	SNP	Best fitting model	Genotype	Control	Case	OR (CI 95%)	p value	p value ^a
				n (%)	n (%)			
<i>TNRC6B</i>	rs139919	Recessive	TT/CT	340 (99)	181 (96)	Reference	0.011	0.013
			CC	3 (1)	8 (4)	5.01 (1.31–19.11)		
<i>CNOT6</i>	rs6877400	Additive	TT	274 (78.)	173 (86)	0.58 (0.37–0.9)	0.011	0.011
			CT	68 (20)	26 (13)			
			CC	7 (3)	1 (1)			
<i>DGCR8</i>	rs9606248	Additive	AA	205 (59)	135 (68)	0.67 (0.48–0.92)	0.012	0.013
			AG	126 (36)	61 (31)			
			GG	17 (5)	3 (1)			
	rs1640299	Recessive	GG/GT	267 (76)	168 (84)	Reference	0.025	0.022
			TT	82 (24)	31 (16)	0.6 (0.38–0.95)		
<i>CNOT1</i>	rs11866002	Dominant	CC	134 (39)	97 (49)	Reference	0.017	0.017
			CT/TT	212 (61)	100 (51)	0.65 (0.46–0.93)		
<i>DROSHA</i>	rs10035440	Additive	TT	213 (63)	108 (55)	1.38 (1.04–1.83)	0.025	0.027
			CT	109 (32)	72 (36)			
			CC	17 (5)	18 (9)			
<i>EIF2C1</i>	rs636832	Recessive	GG/AG	344 (99)	194 (97)	Reference	0.031	0.032
			AA	3 (1)	7 (3)	4.14 (1.06–16.2)		
	rs595961	Recessive	AA/AG	329 (97)	182 (93)	Reference	0.046	0.042
			GG	10 (3)	13 (7)	2.35 (1.01–5.47)		

^acalculated by multivariate logistic regressions to account for the possible confounding effect

of sex

Table 3: Characteristics of the B-ALL patients and controls examined in this study.

	Patients	Controls
No of individuals (n)	213	387
Male (n, %)	124 (56)	199 (51)
Female (n, %)	95 (43)	187 (48)
Mean age \pm SE, years	5.7 \pm 3.5	51.2 \pm 7.7
Age at diagnosis		
1 to 9 years	173 (81)	-
\geq10 years	38 (18)	-
Leukocytes at diagnosis (WBC)		
<20 ($\times 10^9/l$)	94 (30)	-
20 to 200 ($\times 10^9/l$)	54 (25)	-
>200 ($\times 10^9/l$)	6 (3)	-
Hyperdiploid (n, %)	55 (26)	-
No-hyperdiploid (n, %)	107 (50)	-
Chromosomal Translocations		
TEL/AML1	28(13)	-
MLL	10 (5)	-
BCR/ABL	5 (2)	-
E2A/PBX1	3 (1)	-

SE: standard error.