Genetic variant rs2301721 in the *HOXA7* gene is associated with leukemia risk in Jammu region of India

Amrita Bhat [a], Deepak Abrol [b], Gh. Rasool Bhat [a], Ruchi Shah [a], Sonali Verma [a], Divya Bakshi [a], Bhanu Sharma [a], Supinder Singh [c], Rakesh Raina [d], Rakesh Kumar [a] *

[a] Shri Mata Vaishno Devi University, Katra
[b] Department of Radiotherapy, GMC, Kathua, J&K
[c] Department of Medicine, ASCOMS, Sidhra.
[d] Siddhi Vinayak Dental and Health Care Centre, Rupnagar, Jammu, India

*Corresponding author e-mail: kumar.rakesh@smvdu.ac.in

ABSTRACT

Background: Leukemia is a heterogeneous disorder, characterized by high proliferation of white blood cells. Various genetic studies have tried to assess the role of SNPs contributing to the development of leukemia. The HOX is implicated in normal as well as abnormal hematopoiesis. The role of the present variant rs2301721 has been previously studied in various population groups; however, the role of this variant in Leukemia in the Jammu region/population is unclear.

Aim: In the present study, we investigated the genetic variant rs2301721 in the HOXA7 gene in leukemia patients from the Jammu region of India.

Methods: The variant was genotyped by using Sanger sequencing in 180 individuals (90 leukemic cases and 90 healthy controls). The association of SNP with the disease was evaluated by using logistic regression.

Results: It was observed that the variant rs2301721 showed significant association with Leukemia risk in Jammu population The allelic OR of variant rs2301721 was 1.54 (1.01 – 2.34) with p-value = 0.004 and under H.W.E it was 0.838 when corrected for age, gender, BMI, smoking and alcohol.

Conclusion: The present study concludes that the variant rs2301721 in the HOXA7 gene acts as a risk factor in the development of leukemia in the study population.

KEYWORDS: Homeobox proteins (HOXA7); Linkage Disequilibrium (LD); Jammu region; SNP

Citation: Bhat et al. Genetic variant rs2301721 in the *HOXA7* gene is associated with Leukemia risk in Jammu region of India. Polymorphism 2020; 5: 68-76.

INTRODUCTION

Leukemia consists of a group of heterogeneous malignancies in which immature and dysfunctional hematopoietic progenitors proliferate and accumulate in the bone marrow (Alharbi, Pettengell, Pandha, & Morgan, 2013). In the hematopoietic stem cells, a disruption of the cellular processes, including proliferation, differentiation, and cell growth ultimately causes Leukemia ("Leukemia," 2017). In the United States, leukemia is ranked 8th for being the most commonly diagnosed cancer with approximately 61,789 new cases and 23,100 deaths in 2019. According to the Population-Based Cancer Registry of India, males are more affected than females by the ratio 2:1 (Asthana, 2018). The incidence of leukemia in Jammu has shot up rapidly from the previous decade (Chokkalingam & Buffler, 2008). One of India's highly diverse and predominately inbred population groups, the population of Jammu and Kashmir (J&K) region, remained unexplored for blood disorders. Leukemia is multifactorial in origin, which can be caused by both genetic as well as non-genetic factors (Sinnett, 2000). Various genetic alterations, including genetic polymorphisms resulting in gain-of-function mutation in proto-oncogenes as well as inactivating tumor suppressor genes can lead to the progression of leukemia (Münger, 2002). Genome-wide association studies (GWAS) have advanced our understanding of susceptibility to leukemia; however, much of the heritable risk remains unidentified. This factor has been increasingly documented as an inherited genetic factor that can contribute to the progression of leukemia. GWAS data indicate that much of the heritable risk of leukemia is ascribable to common genetic variations which need to be discovered. Recently, GWAS has found the association of variant rs2301721 of HOXA7 with depression disease in Russian population (Shadrina, Bondarenko, & Slominsky, 2018) and colon cancer in Korean population (Kan et al., 2007). The genetic variant rs2301721 of HOXA7 has cytogenetic localization to 7p15.2 with the 1st exonic position and encodes a DNA-binding transcription factor which may regulate gene expression, morphogenesis, and differentiation in hematopoietic stem cells (Inoue et al., 2013). The present study aimed at exploring the association of variant rs2301721 of HOXA7 with the risk of Leukemia in Jammu population.

MATERIAL AND METHODS

Ethics statement

The study was approved by the Institutional Ethics Review Board (IERB) of Shri Mata Vaishno Devi University (SMVDU) under notification number (SMVDU/IERB/16/41). All the details have been recorded in a predesigned proforma and written informed consent was obtained from each participant before conducting the study. All experimental protocols were conducted according to the guidelines and regulations set by IERB, SMVDU.

Sample collection and DNA isolation

A total of 180 subjects were recruited for the study, out of which 90 were cases (leukemic patients) and 90 were healthy controls. All the cases were histopathologically confirmed. The genomic DNA was isolated from the blood samples using a Qiagen DNA Isolation kit (Catalogue No. 51206). Agarose gel electrophoresis was used to analyze the quality of the genomic DNA and quantification was performed using UV spectrophotometer (Nanodrop).

Exome sequencing

The exome sequencing for 05 cases was done on Illumina NGS Platform. The alignment against the hg GrCh37 genome was carried out using BWA and the variant calling was done using GATK pipeline. The alignment against the hg19 GrCh37 genome was carried out using BWA, and the variant calling was done using GATK pipeline. The variants were identified using GATK tool and annotated using ANNOVAR. The databases used for study include 1000 genome, DBNSFP, dbSNP, GWAS and ClinVar.

After filtration of variants against various filters by the use of Exome Capture Kits targeted region.

Sanger Sequencing

Polymerase chain reaction (PCR) was done to amplify PAX5 using Mastercycler® pro, Eppendorf AG, Hamburg, Germany. Overall the target region was amplified using the PCR program with 95ºC for 1 minute, 95ºC for 1 min, 60ºC for 45 seconds with 36 cycles, 72ºC for 1 min, 72ºC for 2min, 4ºC infinity (∞) shown in Table 1 and sequencing was performed by Aggrigenome Pvt. Limited using Sequence Scanner Software (ABI 3730 XL DNA Analyser) and Chromas v2.6.6 (Technelysium Pty Ltd, South Brisbane, Australia).

Statistical Analysis

Statistical Analysis of the data was done by using SPSS software (v.20; Chicago, IL). Chi-square ($\chi 2$) was performed, and genotype frequencies were tested for total Hardy–Weinberg equilibrium. Binary Logistic Regression was also used to estimate OR at 95% CI, and the respective level of significance was estimated as p-value. Moreover, various interactions of genes were analyzed by the string software tool.

RESULT

The total data was generated 45.63 GB with more than 100X sequencing coverage. HOXA7 variant rs2301721 was frequently observed in all samples. Detail of gene is given in Table 2 and Figure 1

HOXA7 was not been reported in leukemia from Jammu Kashmir population of North India. As the variant of HOXA7 was shortlisted from WES, it was imperative to further validate in a larger number of samples and the same was done by sanger sequencing. The main aim was to explore the association of variant rs2301721 of HOXA7 with the Leukemia risk.

This case-control association study included 180 samples (90 leukemic cases and 90 healthy controls). The clinical characteristic distributions of the cases and the controls are given in Table 2. In Table 1, 60% of Leukemia patients were males, and 40% were females, while in controls, 70% were males, and 30% were females. The mean ages were 40.51(±14.67) for leukemia patients and 50.76 (±13.30) years for controls. BMI in cases was 21.2 (±6.10) and 24.9 (±5.3) in Controls.

Table: 1 Showing the details of forwarding and reversed primers of variant rs2301721 of HOXA7.

Gene	Zygosity	SNP ID	cDNA Change	Protein change	Exon No.	Function	Type of Alteration
HOXA7	Heterozygosity	rs2301721	g.28624294 G>A	p.A18T	01	Mis-sense variant	Non synonymous SNV

Table 2: Clinical Characteristics of Cases and Controls.

Characteristics	Cases	Controls	p-value
Age *	40.7 (±14.5)	49.1(±15.30)	<0.01
Gender (in %)	F=40 M=60	F=30 M=70	-
BMI**	21.2 (±6.10)	24.9 (±5.3)	<0.01
Histopathology (Leukemia Types) Acute Lymphoblastic Leukemia (ALL) Acute Myeloid Leukemia (AML) Chronic Myeloid Leukemia (CML) Chronic Lymphoid Leukemia (CLL)	 0.17 0.23 0.50 0.10		
Smoker	0.63	0.22	-
Non-Smoker	0.37	0.78	-
Alcoholic	0.53	0.17	-
Non-Alcoholic	0.47	0.83	-

*age in years and **BMI in kg/m2

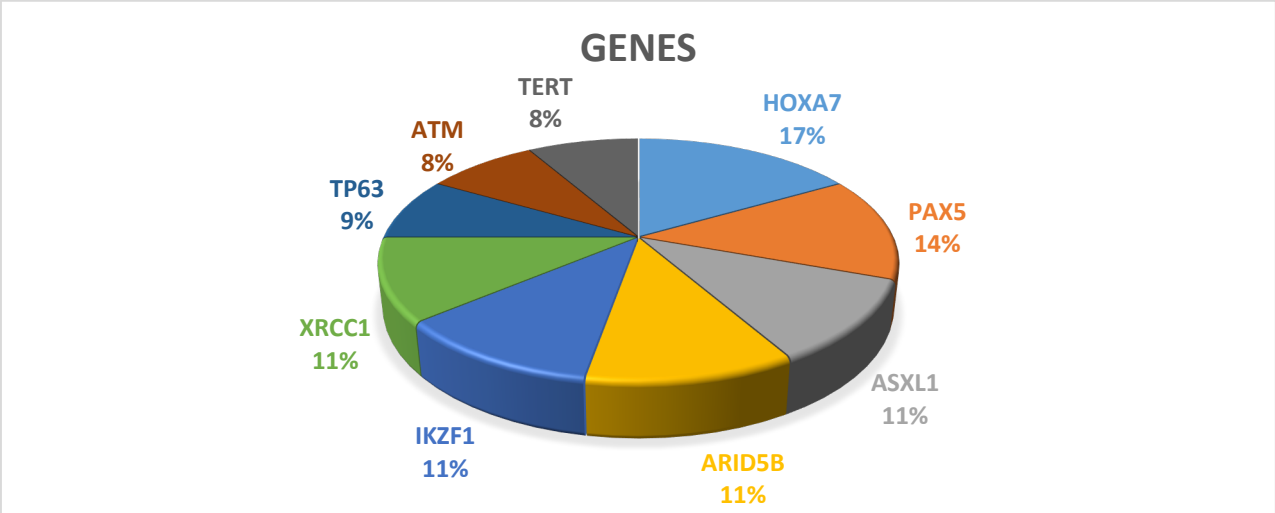

Figure 1: Showing the percentage frequency of genes targeted in exome sequencing, where HOXA7 had a higher percentage (17%) frequency among all genes.

The C>T genetic variant (rs10899750) in the HOXA7 gene results in an amino acid change from Alanine to tryptophan. An allelic frequency distribution is shown in Figure 2 and the chromatogram of HOXA7 shows homozygous wild (CC), heterozygous (CT) and Homozygous mutant (TT) (Figure 3).

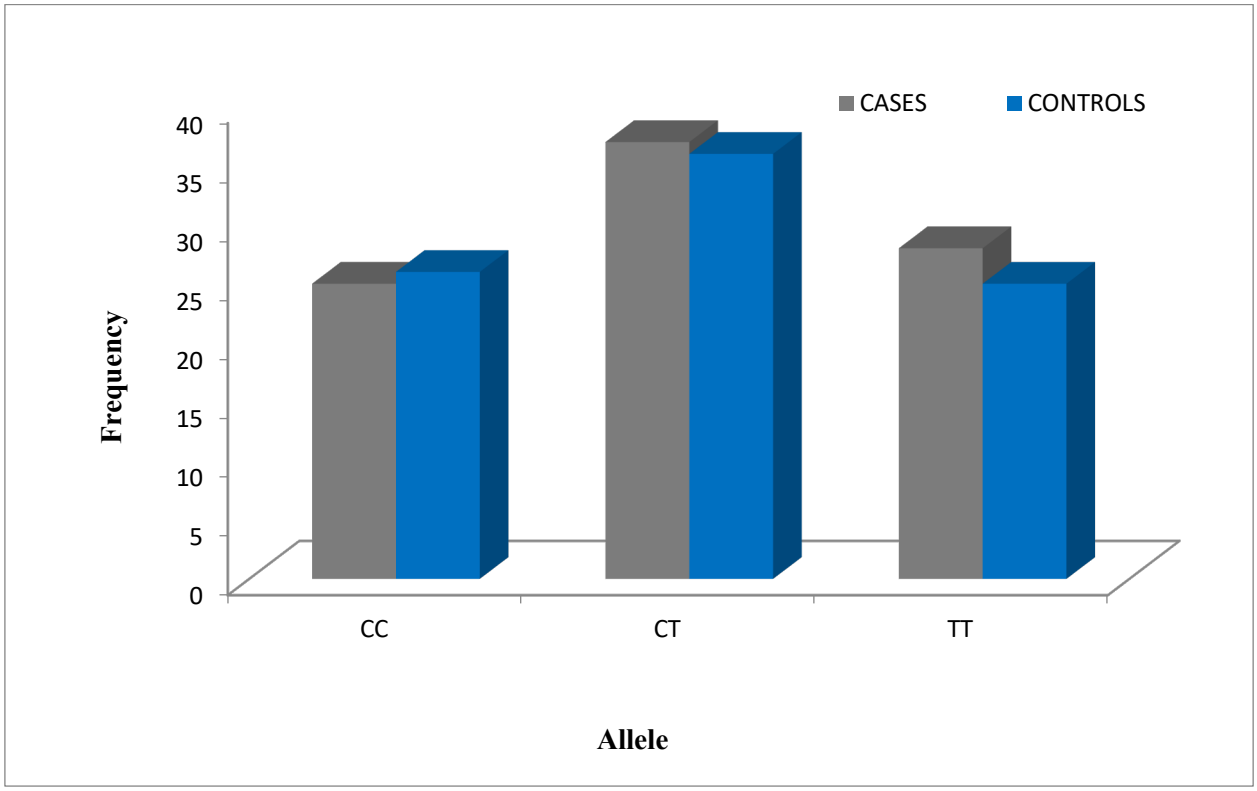

Figure 2: Showing Allelic frequency distribution of variant rs2301721 of HOXA7

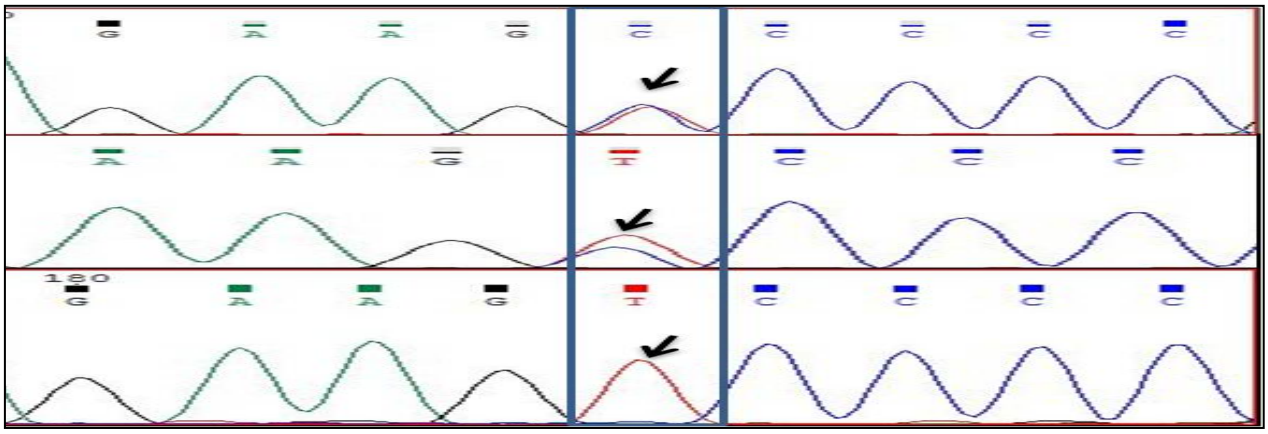

Figure 3: Chromatogram of HOXA7 with variant rs2301721 having missense mutation C >T. Showing A) homozygous wild (CC) B) heterozygous (CT) and C) Homozygous mutant (TT).

The allelic OR of variant rs2301721 was 1.54 (1.01 – 2.34) with p-value = 0.004 and under H.W.E it was 0.838. To observe the maximum effect of allele 'T' dominant model was applied. The OR observed was 2.34 (1.09-3.21, at 95% CI) with p= 0.002 as shown in Table 3 corrected for age, gender, and BMI. Thus, it was observed that the variant rs2301721 of HOXA7 shows a significant association with the risk of leukemia in the population of Jammu region of North India.

Leukemia is divided into four subtypes like CML, AML, ALL, and CLL and in the present study, we performed stratified analysis as mentioned in Table 4. It was observed that Acute Lymphoblastic

Leukemia (ALL), Acute Myeloid Leukemia (AML), Chronic Myeloid Leukemia (CML) were found highly associated as compared to Chronic Lymphoid Leukemia (CLL).

Table 3: Showing the allele frequency distribution for variant rs2301721 of HOXA7.

SNPs /Gene	Allele frequency Cases (N=90)	Allele frequency Controls (N=90)	Allelic OR* (95% C.I.)	*p*-value	Total HWE	Dominant OR* (95% C.I.)	*p*-value
rs2301721 *HOXA7*	C=51.1 T=48.4	C=61.7 T=38.3	1.54 (1.0- 2.34)	0.004	0.838	2.34 (1.0 -3.21)	0.002

DISCUSSION

In this study, an attempt was made to explore the association of rs2301721 polymorphism in the ethnic population of the Jammu region. The same variant was previously examined in Mullerian duct abnormalities in Chinese women and the African population and it was not found to be associated with leukemia (Chen et al., 2014; Franceschini et al., 2013). The polymorphism rs2301721 has also been investigated in hypertension and in cardiovascular diseases (CVD) in African population (Franceschini et al., 2013) and was found to be associated. The same variant was explored in colorectal cancer in the American population (Kan et al., 2007). This genetic variant can also be explored in different ethnic populations in a variety of conditions, including differences in their allele frequency and in both the genetic and environmental backgrounds that interact with the variant.

HOXA7 is a homeodomain-containing transcription factor that plays an important role in hematopoietic stem cell expansion and is commonly deregulated in acute leukemias (Abramovich & Humphries, 2005). PcG proteins are evolutionarily conserved and known to maintain specific repressive states of homeotic (HOX) genes expression pattern within body segments of humans (Bracken, 2006). During every cell cycle transition, the PcG proteins maintain cellular identity by preserving chromatin states against chromatin disruption processes such as DNA replication and transcription. (Kundu et al., 2017). In mammals, PcG proteins incorporate two major functional complexes named Polycomb repressive complex PRC1 and PRC2 (Jo et al., 2011). PRC2 complexes are composed of four conserved components: suppressor of zeste (SUZ12), embryonic ectoderm development (EED), retinoblastoma binding proteins 46 and 48 (RbBP4 and RbBP7) and (EZH2) (Jørgensen et al., 2006). In addition to these core proteins, PRC2 complexes contain one of two methyltransferases with activity toward H3K27, Enhancer of Zeste Homolog 1 or 2 (EZH1 and EZH2) (Holoch & Margueron, 2017). Each of these four core subunits presents multiple orthologs, which incorporate dynamic patterns of the PRC1 complex depending on the differentiated status (Vidal & Starowicz, 2017). Through the ubiquitin E3 activity of its RING1 subunit, PRC1 mediates transcriptional repression by promoting the mono-ubiquitination of H2A at lysine (Cao, Tsukada, & Zhang, 2005). PRC1 can directly inhibit ATP-dependent nucleosome remodeling and block transcriptional initiation and may function in promoting chromatin compaction (Simon & Kingston, 2009). Prc1 complexes composed of CBX (CBX2, 4, 6, 7, or 8), HPF (HPF1, 2, or 3), PCGF, and the ubiquitin ligase RING1A or RING1B bind to H3K27me3 and mediate mono-ubiquitination of H2AK119 (Liss, 2016). The combined effects of PRC1 and PRC2 allow for compacting of chromatin and silencing of gene expression and is associated with the repression of HOX7 (Jäger et al., 2017) (Kim et al., 2018) (Visconte, Tiu, & Rogers, 2014) and lead to the progression of leukemia as shown in Figure 4.

HOXA7 has been shown to interact with many genes associated with many signaling pathways by using String tool software v 10.5 (Supplementary figure S1). Various HOX genes like HOXA7, HOXA5,

HOXA6 help in morphogenesis of hematopoietic stem cells; like-wise, HOXB7, HOXB6, HOXB4, HOXB5 play roles in stem cell differentiation and are upregulated in myeloid cells. Also, HOXC5, HOXC4, HOXC6 are expressed in both lymphoid as well as myeloid cells.

Besides, this genetic variant has the putative regulatory effect (SNIPA online tool) as shown in Figure 5, thus polymorphism in any of the regions could affect the neighboring SNPs and disturb the interaction of genes. Linkage disequilibrium plots show the amount of correlation between a sentinel variant and its surrounding variants. The y-axis shows the correlation coefficient (r2); the x-axis shows the chromosomal position of each SNP. The plot symbol of each variant indicates its functional annotation. The plot symbol of each variant designates its functional observations (http://snipa.helmholtz-muenchen.de).

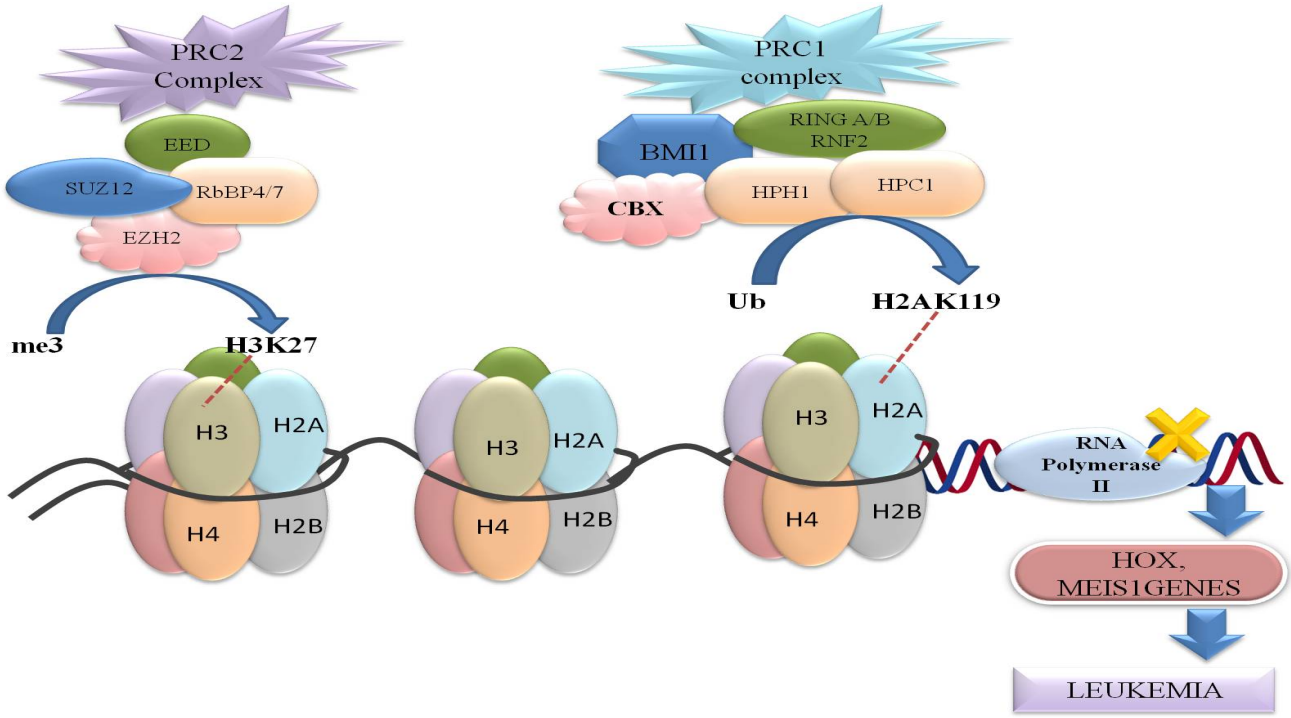

Figure 4: Showing the Pathway of HOXA7 related to leukemia.

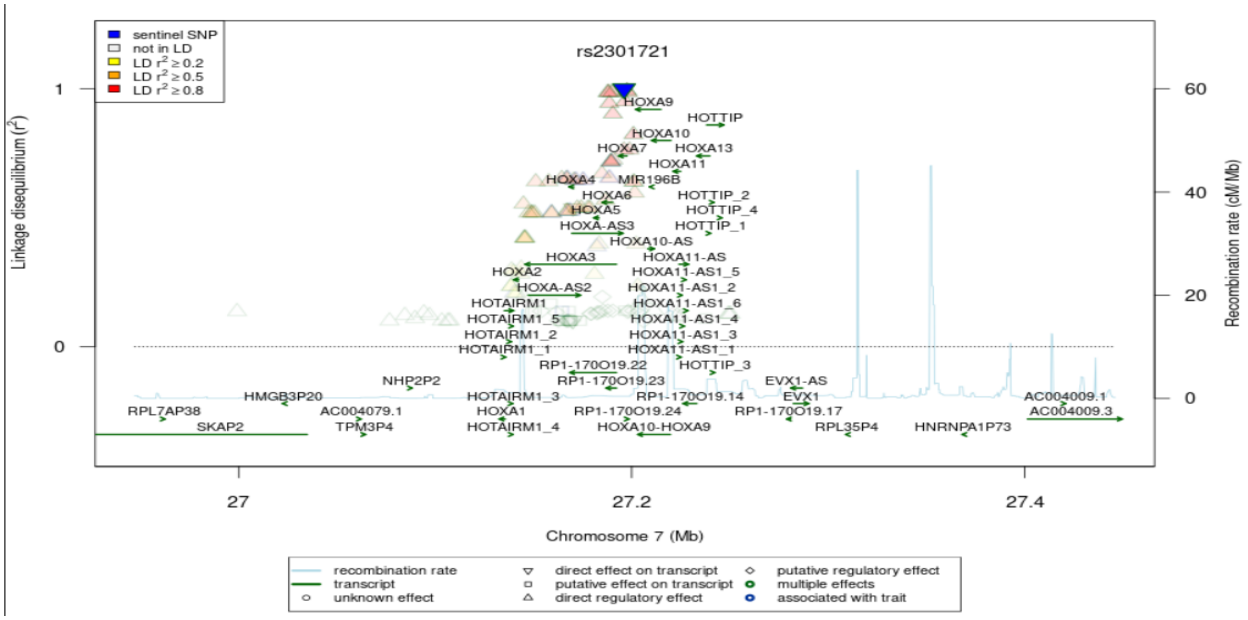

Figure 5: Linkage disequilibrium plot of the variant rs2301721 of HOXA7 with its functional annotation.

CONCLUSION

Our findings provide evidence that the variant rs2301721 of HOXA7 shows a significant association with leukemia in Northern India. Understanding the effects of this variant in different ethnic groups is crucial as they may confer varying risk to leukemia across populations and may act as biomarkers for preventive leukemia screening programs.

Acknowledgement

RK, and AB acknowledge the Department of Science & Technology, Government of India JK/SSTP/459.

Conflict of interest statement

The authors have declared to have no conflict of interest.

Authors' contributions

RK, DA and AB planned the work, AB carried out work all the experiment and wrote the manuscript GRB helped in statistical analysis, while SV, BS, DB helped in sampling processes. SS, RR, RS gave technical suggestions and RK and DA finally refined and approved the manuscript. All authors finally revised and approved the manuscript.

Source of Funding

Nil

Declaration of originality

The authors have declared that the data/text presented in this manuscript is original and no text, figure or data has been copied from any other source without appropriate citation.

Jurisdiction and maps

Polymorphism and Peer Publishers remain neutral to the jurisdictional claims, maps, boundaries and institutional affiliations shown or claimed in any of the articles published.

REFERENCES

Abramovich, C., & Humphries, R. K. (2005). Hox regulation of normal and leukemic hematopoietic stem cells. Current opinion in hematology, 12(3), 210-216.

Alharbi, R. A., Pettengell, R., Pandha, H. S., & Morgan, R. (2013). The role of HOX genes in normal hematopoiesis and acute leukemia. Leukemia, 27(5), 1000-1008.

Asthana, S., Labani, S., Mehrana, S., & Bakhshi, S. (2018). Incidence of childhood leukemia and lymphoma in India. Pediatric Hematology Oncology Journal.

Bracken, K. (2006). Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. Genes & development, 20(9), 1123-1136.

Cao, R., Tsukada, Y.-i., & Zhang, Y. (2005). Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing. Molecular cell, 20(6), 845-854.

Chen, X., Mu, Y., Li, C., Li, G., Zhao, H., Qin, Y., & Chen, Z.-J. (2014). Mutation screening of HOXA7 and HOXA9 genes in Chinese women with Müllerian duct abnormalities. Reproductive BioMedicine Online, 29(5), 595-599.

Chokkalingam, A. P., & Buffler, P. A. (2008). Genetic susceptibility to childhood leukaemia. Radiation protection dosimetry, 132(2), 119-129.

Franceschini, N., Fox, E., Zhang, Z., Edwards, T. L., Nalls, M. A., Sung, Y. J., . . . Adeyemo, A. (2013). Genome-wide association analysis of blood-pressure traits in African-ancestry individuals reveals common associated genes in African and non-African populations. The American Journal of Human Genetics, 93(3), 545-554.

Holoch, D., & Margueron, R. (2017). Mechanisms regulating PRC2 recruitment and enzymatic activity. Trends in biochemical sciences, 42(7), 531-542.

Inoue, D., Kitaura, J., Togami, K., Nishimura, K., Enomoto, Y., Uchida, T., . . . Izawa, K. (2013). Myelodysplastic syndromes are induced by histone methylation–altering ASXL1 mutations. The Journal of clinical investigation, 123(11), 4627-4640.

Jäger, D., Barth, T. F., Brüderlein, S., Scheuerle, A., Rinner, B., von Witzleben, A., . . . von Baer, A. (2017). HOXA7, HOXA9, and HOXA10 are differentially expressed in clival and sacral chordomas. Scientific reports, 7(1), 2032.

Jo, S., Lee, H., Kim, S., Hwang, E. M., Park, J.-Y., Kang, S. S., & Chung, H. (2011). Inhibition of PCGF2 enhances granulocytic differentiation of acute promyelocytic leukemia cell line HL-

Vidal, M., & Starowicz, K. (2017). Polycomb complexes PRC1 and their function in hematopoiesis. Experimental hematology, 48, 12-31.

Visconte, V., Tiu, R. V., & Rogers, H. J. (2014). Pathogenesis of myelodysplastic syndromes: an overview of molecular and non-molecular aspects of the disease. Blood research, 49(4), 216-227.

60 via induction of HOXA7. Biochemical and biophysical research communications, 416(1-2), 86-91.

Jørgensen, H. F., Giadrossi, S., Casanova, M., Endoh, M., Koseki, H., Brockdorff, N., & Fisher, A. G. (2006). Polycomb repressive complexes restrain the expression of lineage-specific regulators in embryonic stem cells. Cell Cycle, 5(13), 1411-1414.

Kan, T., Paun, B. C., Mori, Y., Sato, F., Jin, Z., Hamilton, J. P., . . . Olaru, A. V. (2007). Rarity of Somatic Mutation and Frequency of Normal Sequence Variation Detected in Sporadic Colon Adenocarcinoma Using High-Throughput cDNA Sequencing. Bioinformatics and biology insights, 1, 117793220700100001.

Kim, J., Lee, Y., Lu, X., Song, B., Fong, K.-W., Cao, Q., . . . Yu, J. (2018). Polycomb-and methylation-independent roles of EZH2 as a transcription activator. Cell reports, 25(10), 2808-2820. e2804.

Kundu, S., Ji, F., Sunwoo, H., Jain, G., Lee, J. T., Sadreyev, R. I., . . . Kingston, R. E. (2017). Polycomb repressive complex 1 generates discrete compacted domains that change during differentiation. Molecular cell, 65(3), 432-446. e435.

Leukemia. (2017). American Society of Hematology., 1.

Liss, A. (2016). Altered Chromatin Signaling in Cancer Chromatin Signaling and Diseases (pp. 329-346): Elsevier.

Münger, K. (2002). Disruption of oncogene/tumor suppressor networks during human carcinogenesis. Cancer investigation, 20(1), 71-81.

Shadrina, M., Bondarenko, E. A., & Slominsky, P. A. (2018). Genetics factors in major depression disease. Frontiers in psychiatry, 9, 334.

Simon, J. A., & Kingston, R. E. (2009). Mechanisms of polycomb gene silencing: knowns and unknowns. Nature reviews Molecular cell biology, 10(10), 697-708.

Sinnett, D., Krajinovic, M., & Labuda, D. (2000). Genetic susceptibility to childhood acute lymphoblastic leukemia. Leukemia & lymphoma, 38(5-6), 447-462.

Asthana. (2018). Incidence of childhood leukemia and lymphoma in India. *Pediatric Hematology Oncology Journal, 3*(4), 115-120.

Sinnett, e. a. (2000). Genetic susceptibility to childhood acute lymphoblastic leukemia. *Leukemia & lymphoma, 38*(5-6), 447-462.