Genome-wide haplotype association study identify the FGFR2 gene as a risk gene for Acute Myeloid Leukemia

Hongchao Lv[1],*, Mingming Zhang[1],*, Zhenwei Shang[1],*, Jin Li[1], Shanshan Zhang[2], Duan Lian[1], Ruijie Zhang[1]

[1]College of Bioinformatics Science and Technology, Harbin Medical University, Harbin, China

[2]Hospital of Harbin Turbine Company Limited, Harbin Electric Corporation, Harbin, China

*Joint first authors

Correspondence to: Ruijie Zhang, **email:** zhangruijie_2015@163.com

ABSTRACT

Acute myeloid leukemia (AML) is a cancer of the myeloid line of blood cells, and generally considered to be caused by environment and genetic factors. In this study, we combined a genome-wide haplotype association study (GWHAS) and gene prioritization strategy to mine AML-related genetic affect factors and understand its pathogenesis. A total of 175 AML patients were downloaded from the public GEO database (GSE32462) and 218 matched Caucasian controls were from the HapMap Project. We first identified the linkage disequilibrium (LD) blocks and performed a GWHAS to scan AML-related haplotypes. Then we mapped these haplotypes to the corresponding genes as candidate. And finally, we prioritized all the AML candidate genes based on the similarity with 38 known AML susceptibility genes. The results showed that 1754 haplotypes were significant associated with AML (P<1E-5) and mapped to 591 candidate genes. After prioritizing all 591 AML candidate genes, we obtained four genes ranking at the front as AML risk genes: RUNX1, JAK1, PDGFRA, and FGFR2. Among them, RUNX1, JAK1 and PDGFRA had been confirmed as AML risk genes. In particular, we found that the gene FGFR2 was a novel AML susceptibility gene with a haplotype TT (rs7090018 and rs2912759) showed significant association with AML (P-value = 7.07E-06). In a word, our findings might provide a new perspective to understand the pathogenesis of AML.

INTRODUCTION

Acute myeloid leukemia (AML) is one of the cancer in myeloid line of blood cells, characterized by clonal expansion of myeloid progenitors in the bone marrow and peripheral blood [1]. AML was the most common acute leukemia, and increased with the population ages [2]. In the USA alone, its incidence was 3.5 cases per 100,000 people, and approximately 20,000 patients were diagnosed with AML and 10,000 AML-related deaths occurred in 2015 [3]. AML with t(8;21)(q22;q22) [RUNX1-RUNX1T1] and inv(16)(p13.1q22) or t(16;16) (p13.1;q22) [CBFB-MYH11], which the aberrant fusion genes leaded to impaired differentiation of hematopoietic progenitors, commonly referred to as core-binding factor (CBF)–CAML. It accounted for approximately 15% of AML patients, and was considered to have relatively good prognosis compared to other leukemia subtypes [4]. In recent years, many studies had revealed novel mutations, epigenetic changes, and/or aberrant expression levels of protein-coding and noncoding genes involved in leukemogenesis. These results could help us to understand the genetic basis of the disease and refine the risk assessment in CBF-AML. Furthermore, they would serve as targets for novel therapeutic approaches [4, 5]. With the development of high-resolution genome-wide scanning technologies, we could detect more AML-related SNPs, copy number alterations (CNAs) and other regions in AML and other myeloid neoplasms.

Kühn MW, et al shared a genome-wide CBF-AML SNP-array data set with pediatric and adult CBF-AML patients (GSE32462). Based on the mutation analyses and CNAs analyses, they identified multiple activating mutations in KIT, FLT3, JAK2, NRAS, and KRAS, and

revealed that in some cases, more than one gene within the RAS/Kinase signaling pathway were activated [4]. As we know, a sequence of linkage disequilibrium (LD) alleles on a particular chromosome region might make up of a haplotype block, which would contain genetic information of several SNPs with marginal effects. Comparing the results about single-marker analysis, several previous studies confirmed that haplotype analysis results could increase power for detecting associations with disease and were robustness for mapping disease genes [6, 7]. Thus we performed a genome-wide haplotype association analysis study (GWHAS) to identify potentially novel AML risk genes. The International HapMap Project is an important resource in the human genome research and provides a variety of adjacent populations as a reference. In view of there were not any normal individuals in GSE32462, we constructed a control samples set from the HapMap Project to effectively use the precious genome-wide SNPs data set. Furthermore, according to the related genes in a disease or biological process generally shared similar characteristics in multiple omics data resources [8, 9], a gene prioritization strategy was adopted to assess the similarity between AML candidate genes and known AML susceptibility genes.

RESULTS

The results of genome-wide haplotype association study

In this study, a total of 739,981 autosomal SNPs were included, which shared in both GPL6801 platform and the HapMap database. After the quality control (detail see methods), our study obtained a total of 734,624 eligible autosomal SNPs. Then a genome-wide haplotype association study was performed, which contained 393 samples (175 AML patients and 218 matched healthy individuals). We identified 118,057 haplotype blocks, which contained total 519,865 haplotypes. For each haplotype, we carried out a chi-square test to assess the relationship with AML risk. In the end, 1,754 haplotypes located on 1,089 block regions showed significant correlation with AML (P< 1E-5) (Supplementary Table 2).

Mapping and prioritizing the AML candidate genes

According to the chromosome location information, 1,754 significant haplotypes were mapped to 591 genes, and regarded them as AML-related candidate genes. Meanwhile, we retrieved 38 known AML susceptibility genes from OMIM and GAD database, which had been confirmed to associate with AML risk. After that, we analyzed the similarity on the structure and function between two gene sets, and total 42 features were included, such as Ensembl Est, Gene Ontology, KEGG,

Blast, HPRD, and so on (See methods for more detailed). We performed all above analyses by the Endeavour software. At last, four genes, RUNX1, JAK1, PDGFRA, and FGFR2, had lower P-value (P<0.05), and regarded as high-risk AML genes (Table 1). All detailed results for candidate genes prioritization could be seen in Supplementary Table 3.

The analysis of RUNX1 gene

The runt related transcription factor 1 (RUNX1) gene, also named as AML1, PEBP2, located on chromosome 21q22.3, and the physical location was: 36,160,098bp (start) – 36,421,595bp (end). The RUNX1 gene ranked the first, and also was the only gene existed in both training and candidate gene sets. From the detail information on Figure 1, we could see a LD block located on the gene with 2 SNPs: rs8133478 and rs762247, and a haplotype GC showed significant association with AML (P-value = 1.66E-6). Based on the EM algorithm to estimate the frequency of haplotype, it in patients was 5.95E-2, whereas in control individuals was 2.57E-3. Obviously, the haplotype GC could be inferred as a risk factor for AML with an OR=23.14. Generally, as a member of the RUNX gene family, it appeared frequently in the development of all hematopoietic cell types. The RUNX1 gene was involved in the t(8;21) translocation in acute and chronic myeloid leukemia, and could produce oncogenic transformation to AML [10]. It was also demonstrated associated with other types of leukemia [11, 12], such as frequently mutated in pediatric B progenitor acute lymphoblastic leukemia (ALL) in Pui (1995)'s research [13].

The analysis of JAK1 gene

The janus kinase 1 (JAK1) gene located from 65,298,906bp (start) to 65,432,187bp (end) on chromosome 1p32.3-p31.3. The detailed information could be seen on the Figure 2. In this gene, we found a haplotype CAGGG showed significant association with AML (chi-square test p value is 1.83E-12), which located on a block that consisted of 5 SNPs: rs3790541, rs11208534, rs310202, rs310201, and rs310199. As the frequency of the haplotype CAGGG in patients was 0.118, and in control individuals was 2.57E-3, and the odds ratio was 45.8, we could infer that the haplotype CAGGG was an AML risk factor, and the individuals carried it might increase the genetic risk of AML. The oncogenic JAK1 gene encoded a cytoplasmic tyrosine kinase and involved in lymphoid cell precursor proliferation, survival, and differentiation [14]. Dysregulation of JAK-STAT pathway had been found as key events in a variety of hematological malignancies [15]. Additionally, it has been reported that somatic

Table 1: The AML related genes for prioritization according to reliability verification

Ensembl gene ID	Gene symbol	Gene description	location	P-value
ENSG00000159216	RUNX1	runt related transcription factor 1	21q22.3	0.0016
ENSG00000162434	JAK1	janus kinase 1	1p32.3-p31.3	0.0318
ENSG00000134853	PDGFRA	platelet derived growth factor receptor alpha	4q12	0.043
ENSG00000066468	FGFR2	fibroblast growth factor receptor 2	10q26	0.0494

Compared with the known AML genes, the candidate genes were ranked based on P-values of global prioritization. We listed the top four candidate genes with P-values lower than 0.05 in Table 1.

mutations in JAK1 occurred in individuals with acute lymphoblastic leukemia (ALL) and the dysregulated JAK1 function effected on ALL, particularly of T cell origin [14, 16]. Furthermore, Xiang Z–s report firstly demonstrated somatic JAK1 mutations in AML and suggested that JAK1 mutations might have function as disease-modifying mutations in AML pathogenesis [17]. According to the association analysis about the regions of strong LD, our study further revealed that the JAK1 gene might be an AML-related risk gene.
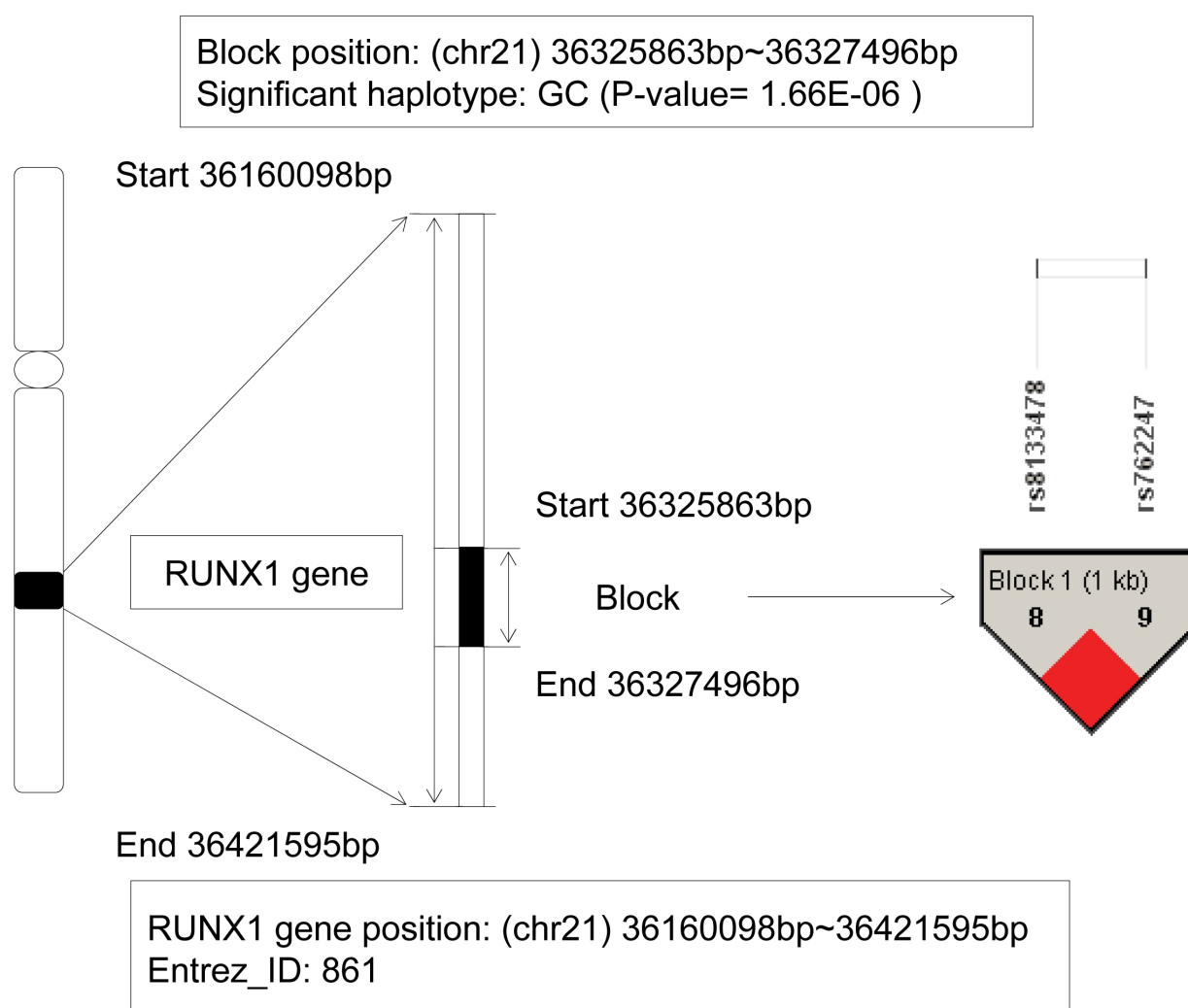

Figure 1: The haplotype analysis result of RUNX1 gene.

The analysis of PDGFRA gene

From our GWHAS, a LD block was found on the platelet derived growth factor receptor alpha (PDGFRA) gene with a physical location from 55,063,358bp to 55,102,425bp. Among six haplotypes, two haplotypes GGGCTC and GGGCTT showed significant association with AML (P-value = 6.35E-08, and 3.34E-18) (see Figure 3). For the haplotype GGGCTC, it can be inferred as a protective factor with an OR=0.284, and however, the haplotype GGGCTT showed a risk effect with OR=841.67. Studies revealed that this gene involved in organ development, wound healing, and tumor progression. Mutations in the PDGFRA gene had been implicated in idiopathic hypereosinophilic syndrome [18], somatic and familial gastrointestinal stromal tumors [19], brain tumor [20] and a variety of other cancers. Hiwatari M's study revealed [21] novel missense mutations in the tyrosine kinase domain of the PDGFRA gene in childhood acute myeloid leukemia with t(8;21)(q22;q22) or inv(16) (p13q22). The results suggested that PDGFRA mutations might be implicated in oncogenic mechanisms in AML. Moreover, recent studies showed if the eosinophilia-associated AML patients presented the FIP1L1-PDGFRA fusion gene, they should be as excellent candidates for treatment with tyrosine kinase inhibitors [22]. In our study, the relationship between PDGFRA gene and AML risk was detected again.

The analysis of FGFR2 gene

The fibroblast growth factor receptor 2 (FGFR2) gene located on chromosome 10q26 with the physical location: 123,237,844bp (start) – 123,357,972bp (end). From Figure 4, a LD block contained 2 SNPs: rs7090018 and rs2912759, located on this gene, and a haplotype TT showed significant association with AML (P-value = 7.07E-06). When we implemented the haplotype frequency estimation via the EM algorithm, the frequencies of the haplotype TT were 0.0463 in patients and only zero in control individuals, respectively. Although the allele frequency in patients showed relatively low, the haplotype TT still revealed a risk effect compared to the lower control group. In recent years, the association between mutations in FGFR2 gene and various cancers had been found, such as breast cancer [23], endometrial cancer [24], oral squamous cell carcinoma [25], gastric cancer [26], and so on. Further, fibroblast growth factor receptors
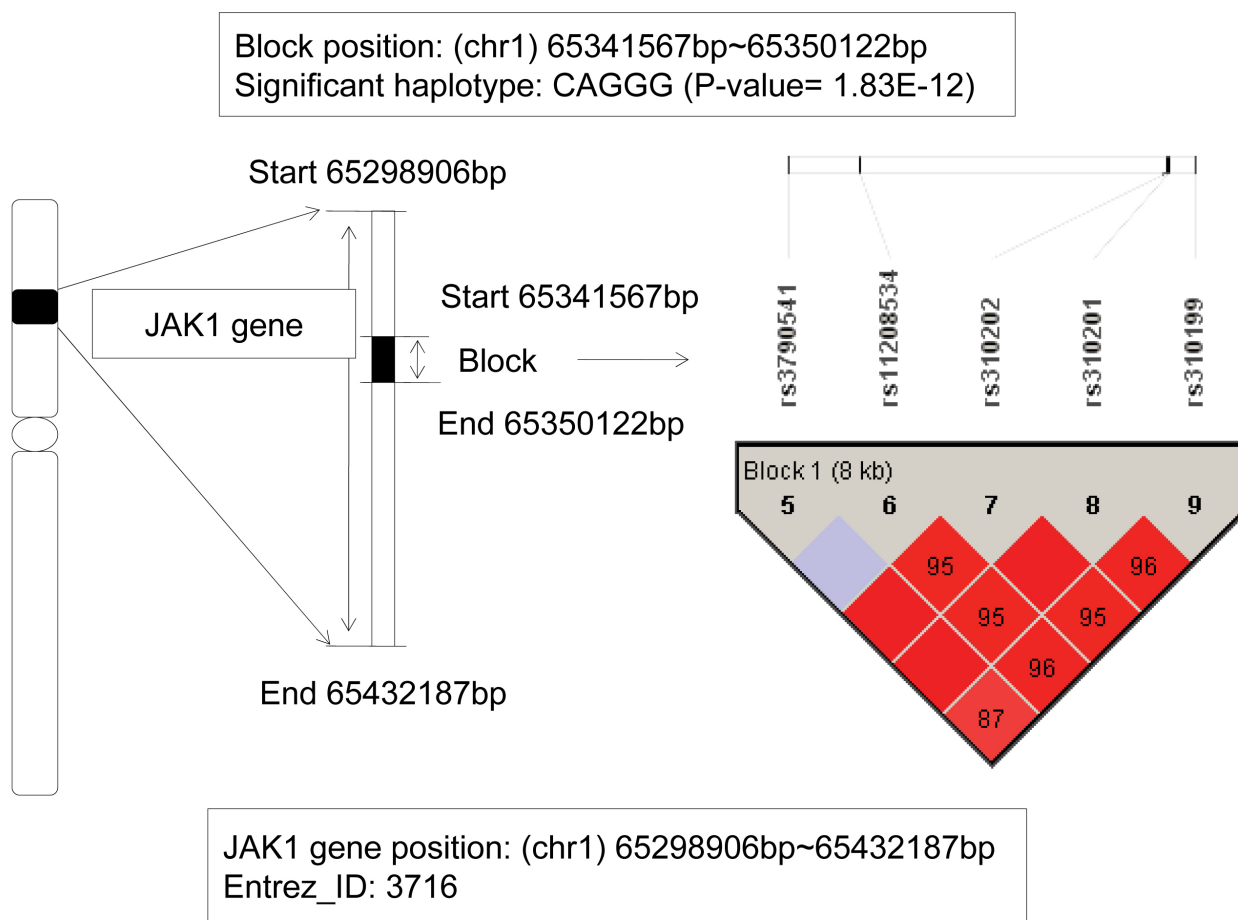

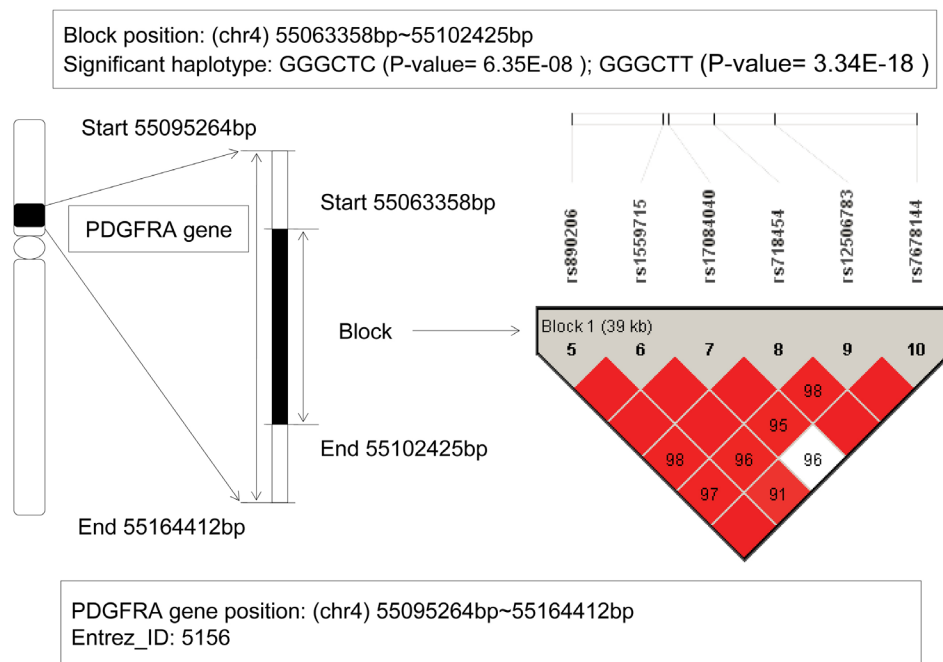Figure 2: The haplotype analysis result of JAK1 gene.

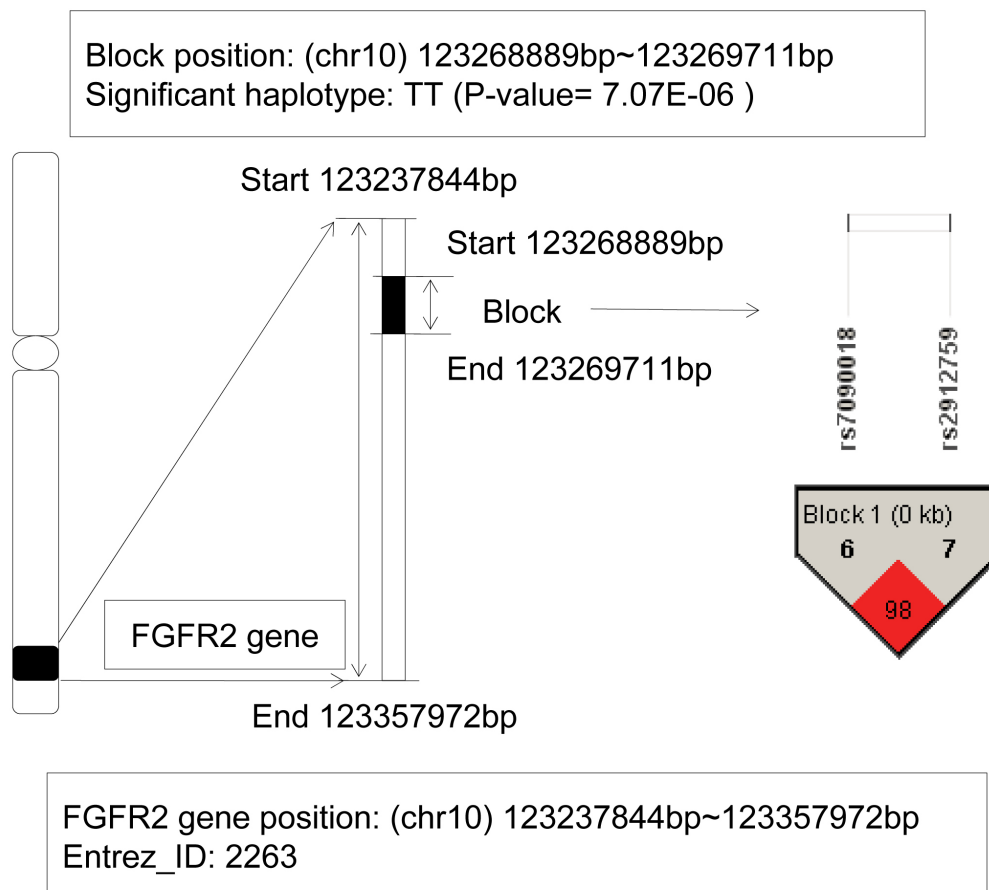Figure 3: The haplotype analysis result of PDGFRA gene.

Figure 4: The haplotype analysis result of FGFR2 gene.

(FGFRs) genes had been shown to be translocated in multiple myeloma (MM) and myeloproliferative disorder (MPD) [27, 28]. Jang JH, et al's study also found a splice variant of FGFR2 gene in human leukemia HL-60 cells, and revealed that it was associated with AKT and MAPK pathway activation [29]. All of these indicated its oncogenic potential. In additional, Tanizaki J, et al's study identified a novel FGFR2 extracellular domain insertion mutation and demonstrated that they were both oncogenic and sensitive to inhibition by FGFR kinase inhibitors [30]. From the perspective of linkage disequilibrium, our study for the first time revealed the relationship between the FGFR2 gene and AML risk.

DISCUSSION

Acute myeloid leukemia was recognized as a complex disease of hematopoietic stem cell disorders. It was classified into several categories based on underlying genetic alterations to facilitate diagnosis and prognosis by the World Health Organization (WHO). Among them, CBF-AML was characterized by the presence of distinct cytogenetic abnormalities [31]. The genetic alterations discovered in CBF-AML would help us to understand the process of leukemogenesis and also serve as targets for novel therapeutic approaches [32]. Although its pathogenesis mechanisms remained unclear, new molecular technologies had allowed for in-depth molecular analyses of AML patients and revealed some molecular aberrations such as novel mutations, epigenetic changes, and so on [5]. Moreover, the LD region on a gene could be regarded as a sensitive indicator of the population genetic forces [33]. Based on the genotyping arrays dataset GSE32462, we firstly performed a GWHAS to identify the AML related haplotypes. It covered the marginal effects of SNPs in a haplotype block and might have higher efficiency to mined AML susceptibility genes [6]. Here, a total of 1,089 block regions were caught, and mapped to 591 AML candidate genes. Then based on multi-omics data resources, we prioritized and identified 4 AML risk genes: RUNX1, JAK1, PDGFRA, and FGFR2. Among of them, RUNX1 ranked the first and was an oncogene for AML; JAK1 and PDGFRA gene also were confirmed association with AML. Through retrieving the published literature, our study revealed that FGFR2 gene, which had been involved in various cancers, was association with AML risk for the first time.

Recent years, the development of high-resolution genome-wide scanning technologies had given us abundant genetic variation resources for mining disease pathogenesis. Although still a lot of work need to be done, we believe that our work is effective and feasible. The study about conserved linkage disequilibrium chromosome regions can provide a new perspective for subsequent understanding of the pathogenesis about AML.

MATERIALS AND METHODS

Datasets and Quality control

In this study, we downloaded all the original genotypes data of Caucasian CBF-AML patients from GEO database (accession GSE32462) under the chip platform GPL6801 (Affymetrix SNP 6.0 array), which contained 930,662 SNP markers. The purpose of our study was to identify the genetic risk factors of acute myeloid leukemia and avoid the effect of acquired somatic mutation. Therefore we selected the germline DNA of 175 pediatric and adult CBF-AML patients obtained from bone marrow or blood at remission. Considering the International HapMap Project could provide a variety of geographically neighboring reference populations, we compared the similarity of allele frequencies between HapMap Caucasian populations and GSE32462, and selected the highest correlation the HapMap CEU and TSI populations as a reference. Finally, 218 unrelated Caucasian control individuals were downloaded from phase II+III of the International HapMap Project (CEU, 116; TSI, 102, ftp://ftp.ncbi.nih.gov/hapmap) [34, 35]. We took them as normal control samples. There were 739,981 autosomal SNPs shared in both GPL6801 platform and the HapMap database. Then we performed a quality control (QC) to select data according to the following criteria: Hardy–Weinberg equilibrium (HWE) P > 0.001, minor allele frequency (MAF) > 0.001, and percentage of successfully genotyping for the marker >75%. Finally, a total of 734,624 autosomal SNPs passed the QC and were used in the subsequent analyses.

Genome-wide haplotype association analysis

On the basis of the autosomal SNPs data of 393 case-control samples, we performed a GWHAS to identify AML-related haplotypes. First, we used the Four Gamete Tests (FGT) method to assess the linkage disequilibrium (LD) blocks [36]. Then the haplotype phasing was estimated using the Maximum Likelihood Estimation (MLE) and the frequencies of the haplotypes were estimated using the Expectation Maximization (EM) algorithm. Finally, a chi-square test was performed, and if the P-value less than 1E-5, the haplotype was captured and considered as an AML-related candidate. All above analyses were done using the HAPLOVIEW software [37].

Mapping AML candidate genes

We mapped all the AML-related candidate haplotypes to genes based on the chromosome location information. All the genes location information was queried from the NCBI gene database. If a gene shared the same chromosome fragment with at least one AML

candidate haplotype, it would be regarded as an AML candidate gene.

Prioritizing candidate genes to find AML risk genes

In some ways, the genes involved in the defined biological processes or diseases often shared some related characteristics, such as the similarity for sequence, functional annotation, expression and regulation information [9, 38]. Therefore, it could be considered that the AML potential risk genes should share some genetic features with the known AML genes either directly or indirectly. In order to identify the AML risk genes, we adopted a gene prioritization strategy proposed by Aerts S, et al [9]. We performed it in three simple steps. First of all, we used the known disease genes as a training set, and established the training models based on the biological process of interest. Each sub-model corresponds to one selected data source. For a given data source, the scores of each candidate gene were computed using the associated sub-model. Then we ranked the candidate genes in each selected data source. Obviously, the most promising candidate genes would be at the front. At last, we integrated all the candidate genes rankings correspond to the given data sources and used an order statistics to obtain a single global ranking for mining the optimal disease risk genes. Based on the above strategies, we firstly collected 38 AML known genes as training gene set, in which 34 genes were from the Online Mendelian Inheritance in Man (OMIM) database, the other 4 AML known genes were verified by at least three published studies in genetic association database (GAD) database (Supplementary Table 1; All the 3 supplementary tables could be found in the website: http://www.bioapp.org/research/AMLhaplotype). Then up to 42 available data sources were used to train features and construct models, which generally contained the following categories: gene and protein function, chemical information, bio-molecular pathways, phenotypic information, interaction networks, expression profiles, expression ontologies, sequence features [8]. For instance, the 'Gene and protein function' category included resources such as Gene Ontology, Pfam, UniProt and InterPro. We could obtain the ranking of the candidate genes associated with each given data source. Then we fused all of these rankings from the separate data sources into a single ranking and obtained a global prioritization using order statistics [39]. The global order statistic Q-value was modeled by the Gamma-distribution, and then an approximate P-value was calculated according to the cumulative distribution function. The P-value represented the significance of this combination of rankings. If one candidate gene had a smaller P-value, it was regarded as more similar to the known AML genes and more likely to be considered as an AML risk gene [40]. In this study, the comprehensive prioritization results

of candidate genes could be obtained from the online software Endeavour (https://endeavour.esat.kuleuven.be/Default.aspx) [8].

CONFLICTS OF INTEREST

The authors declare that they have no conflict of interest.

REFERENCES

1. Saultz JN, Garzon R. Acute Myeloid Leukemia: A Concise Review. J Clin Med. 2016; 5. doi: 10.3390/jcm5030033E33 [pii] jcm5030033 [pii].

2. Kavianpour M, Ahmadzadeh A, Shahrabi S, Saki N. Significance of oncogenes and tumor suppressor genes in AML prognosis. Tumour Biol. 2016. doi: 10.1007/s13277-016-5067-110.1007/s13277-016-5067-1 [pii].

3. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2015. CA Cancer J Clin. 2015; 65: 5-29. doi: 10.3322/caac.21254.

4. Kuhn MW, Radtke I, Bullinger L, Goorha S, Cheng J, Edelmann J, Gohlke J, Su X, Paschka P, Pounds S, Krauter J, Ganser A, Quessar A, et al. High-resolution genomic profiling of adult and pediatric core-binding factor acute myeloid leukemia reveals new recurrent genomic alterations. Blood. 2012; 119: e67-75. doi: 10.1182/blood-2011-09-380444 blood-2011-09-380444 [pii].

5. Khaled S, Al Malki M, Marcucci G. Acute Myeloid Leukemia: Biologic, Prognostic, and Therapeutic Insights. Oncology (Williston Park). 2016; 30: 318-29. doi: 216453 [pii].

6. Morris RW, Kaplan NL. On the advantage of haplotype analysis in the presence of multiple disease susceptibility alleles. Genet Epidemiol. 2002; 23: 221-33. doi: 10.1002/gepi.10200.

7. Akey J, Jin L, Xiong M. Haplotypes vs single marker linkage disequilibrium tests: what do we gain? Eur J Hum Genet. 2001; 9: 291-300. doi: 10.1038/sj.ejhg.5200619.

8. Tranchevent LC, Barriot R, Yu S, Van Vooren S, Van Loo P, Coessens B, De Moor B, Aerts S, Moreau Y. ENDEAVOUR update: a web resource for gene prioritization in multiple species. Nucleic Acids Res. 2008; 36: W377-84. doi: 10.1093/nar/gkn325 gkn325 [pii].

9. Aerts S, Lambrechts D, Maity S, Van Loo P, Coessens B, De Smet F, Tranchevent LC, De Moor B, Marynen P, Hassan B, Carmeliet P, Moreau Y. Gene prioritization through genomic data fusion. Nat Biotechnol. 2006; 24: 537-44. doi: nbt1203 [pii] 10.1038/nbt1203.

10. Cohen MM, Jr. Perspectives on RUNX genes: an update. Am J Med Genet A. 2009; 149A: 2629-46. doi: 10.1002/ajmg.a.33021.

11. Nucifora G, Rowley JD. AML1 and the 8;21 and 3;21 translocations in acute and chronic myeloid leukemia. Blood. 1995; 86: 1-14. doi:

12. Okuda T, van Deursen J, Hiebert SW, Grosveld G, Downing JR. AML1, the target of multiple chromosomal translocations in human leukemia, is essential for normal fetal liver hematopoiesis. Cell. 1996; 84: 321-30. doi: S0092-8674(00)80986-1 [pii].

13. P ui CH. Childhood leukemi a s . N Engl J Med. 1995; 332: 1618-30. doi: 10.1056/NEJM199506153322407.

14. Flex E, Petrangeli V, Stella L, Chiaretti S, Hornakova T, Knoops L, Ariola C, Fodale V, Clappier E, Paoloni F, Martinelli S, Fragale A, Sanchez M, et al. Somatically acquired JAK1 mutations in adult acute lymphoblastic leukemia. J Exp Med. 2008; 205: 751-8. doi: 10.1084/jem.20072182 jem.20072182 [pii].

15. Furqan M, Mukhi N, Lee B, Liu D. Dysregulation of JAK-STAT pathway in hematological malignancies and JAK inhibitors for clinical application. Biomark Res. 2013; 1: 5. doi: 10.1186/2050-7771-1-52050-7771-1-5 [pii].

16. Losdyck E, Hornakova T, Springuel L, Degryse S, Gielen O, Cools J, Constantinescu SN, Flex E, Tartaglia M, Renauld JC, Knoops L. Distinct Acute Lymphoblastic Leukemia (ALL)-associated Janus Kinase 3 (JAK3) Mutants Exhibit Different Cytokine-Receptor Requirements and JAK Inhibitor Specificities. J Biol Chem. 2015; 290: 29022-34. doi: 10.1074/jbc.M115.670224 M115.670224 [pii].

17. Xiang Z, Zhao Y, Mitaksov V, Fremont DH, Kasai Y, Molitoris A, Ries RE, Miner TL, McLellan MD, DiPersio JF, Link DC, Payton JE, Graubert TA, et al. Identification of somatic JAK1 mutations in patients with acute myeloid leukemia. Blood. 2008; 111: 4809-12. doi: blood-2007-05-090308 [pii] 10.1182/blood-2007-05-090308.

18. Cools J, DeAngelo DJ, Gotlib J, Stover EH, Legare RD, Cortes J, Kutok J, Clark J, Galinsky I, Griffin JD, Cross NC, Tefferi A, Malone J, et al. A tyrosine kinase created by fusion of the PDGFRA and FIP1L1 genes as a therapeutic target of imatinib in idiopathic hypereosinophilic syndrome. N Engl J Med. 2003; 348: 1201-14. doi: 10.1056/NEJMoa025217348/13/1201 [pii].

19. Chompret A, Kannengiesser C, Barrois M, Terrier P, Dahan P, Tursz T, Lenoir GM, Bressac-De Paillerets B. PDGFRA germline mutation in a family with multiple cases of gastrointestinal stromal tumor. Gastroenterology. 2004; 126: 318-21. doi: S0016508503017748 [pii].

20. De Bustos C, Smits A, Stromberg B, Collins VP, Nister M, Afink G. A PDGFRA promoter polymorphism, which disrupts the binding of ZNF148, is associated with primitive neuroectodermal tumours and ependymomas. J Med Genet. 2005; 42: 31-7. doi: 42/1/31 [pii] 10.1136/jmg.2004.024034.

21. Hiwatari M, Taki T, Tsuchida M, Hanada R, Hongo T, Sako M, Hayashi Y. Novel missense mutations in the tyrosine kinase domain of the platelet-derived growth factor receptor alpha(PDGFRA) gene in childhood acute myeloid leukemia with t(8;21)(q22;q22) or inv(16)(p13q22). Leukemia. 2005; 19: 476-7. doi: 2403638 [pii] 10.1038/sj.leu.2403638.

22. Metzgeroth G, Walz C, Score J, Siebert R, Schnittger S, Haferlach C, Popp H, Haferlach T, Erben P, Mix J, Muller MC, Beneke H, Muller L, et al. Recurrent finding of the FIP1L1-PDGFRA fusion gene in eosinophilia-associated acute myeloid leukemia and lymphoblastic T-cell lymphoma. Leukemia. 2007; 21: 1183-8. doi: 2404662 [pii] 10.1038/sj.leu.2404662.

23. Xi J, Su Y, Beeghly Fadiel A, Lin Y, Su FX, Jia WH, Tang LY, Ren ZF. Association of physical activity and polymorphisms in FGFR2 and DNA methylation related genes with breast cancer risk. Cancer Epidemiol. 2014; 38: 708-14. doi: 10.1016/j.canep.2014.09.0 02S1877-7821(14)00152-0 [pii].

24. Lin YC, Er TK, Yeh KT, Hung CH, Chang JG. Rapid Identification of FGFR2 Gene Mutations in Taiwanese Patients With Endometrial Cancer Using High-resolution Melting Analysis. Appl Immunohistochem Mol Morphol. 2015; 23: 532-7. doi: 10.1097/PAI.0000000000000114.

25. Nayak S, Goel MM, Makker A, Bhatia V, Chandra S, Kumar S, Agarwal SP. Fibroblast Growth Factor (FGF-2) and Its Receptors FGFR-2 and FGFR-3 May Be Putative Biomarkers of Malignant Transformation of Potentially Malignant Oral Lesions into Oral Squamous Cell Carcinoma. PLoS One. 2015; 10: e0138801. doi: 10.1371/journal.pone.0138801 PONE-D-15-17251 [pii].

26. Jang JH, Shin KH, Park JG. Mutations in fibroblast growth factor receptor 2 and fibroblast growth factor receptor 3 genes associated with human gastric and colorectal cancers. Cancer Res. 2001; 61: 3541-3. doi:

27. Chesi M, Nardini E, Brents LA, Schrock E, Ried T, Kuehl WM, Bergsagel PL. Frequent translocation t(4;14)(p16.3;q32.3) in multiple myeloma is associated with increased expression and activating mutations of fibroblast growth factor receptor 3. Nat Genet. 1997; 16: 260-4. doi: 10.1038/ng0797-260.

28. Ollendorff V, Guasch G, Isnardon D, Galindo R, Birnbaum D, Pebusque MJ. Characterization of FIM-FGFR1, the fusion product of the myeloproliferative disorder-associated t(8;13) translocation. J Biol Chem. 1999; 274: 26922-30. doi:

29. Jang JH, Chung CP. A novel splice variant of fibroblast growth factor receptor 2 in human leukemia HL-60

cells. Blood Cells Mol Dis. 2002; 29: 133-7. doi: S1079979602905486 [pii].

30. Tanizaki J, Ercan D, Capelletti M, Dodge M, Xu C, Bahcall M, Tricker EM, Butaney M, Calles A, Sholl LM, Hammerman PS, Oxnard GR, Wong KK, et al. Identification of Oncogenic and Drug-Sensitizing Mutations in the Extracellular Domain of FGFR2. Cancer Res. 2015; 75: 3139-46. doi: 10.1158/0008-5472.CAN-14-3771 0008-5472.CAN-14-3771 [pii].

31. Paschka P. Core binding factor acute myeloid leukemia. Semin Oncol. 2008; 35: 410-7. doi: 10.1053/j.seminoncol.2008.04.011 S0093-7754(08)00121-8 [pii].

32. Paschka P, Dohner K. Core-binding factor acute myeloid leukemia: can we improve on HiDAC consolidation? Hematology Am Soc Hematol Educ Program. 2013; 2013: 209-19. doi: 10.1182/asheducation-2013.1.209 2013/1/209 [pii].

33. Slatkin M. Linkage disequilibrium—understanding the evolutionary past and mapping the medical future. Nat Rev Genet. 2008; 9: 477-85. doi: 10.1038/nrg2361 nrg2361 [pii].

34. Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, Belmont JW, Boudreau A, Hardenbol P, Leal SM, Pasternak S, Wheeler DA, Willis TD, et al. A second generation human haplotype map of over 3.1 million SNPs. Nature. 2007; 449: 851-61. doi: nature06258 [pii] 10.1038/nature06258.

35. Altshuler DM, Gibbs RA, Peltonen L, Dermitzakis E, Schaffner SF, Yu F, Bonnen PE, de Bakker PI, Deloukas P, Gabriel SB, Gwilliam R, Hunt S, Inouye M, et al. Integrating common and rare genetic variation in diverse human populations. Nature. 2010; 467: 52-8. doi: 10.1038/nature09298 nature09298 [pii].

36. Wang N, Akey JM, Zhang K, Chakraborty R, Jin L. Distribution of recombination crossovers and the origin of haplotype blocks: the interplay of population history, recombination, and mutation. Am J Hum Genet. 2002; 71: 1227-34. doi: S0002-9297(07)60418-2 [pii] 10.1086/344398.

37. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics. 2005; 21: 263-5. doi: 10.1093/bioinformatics/bth457 bth457 [pii].

38. Turner FS, Clutterbuck DR, Semple CA. POCUS: mining genomic sequence annotation to predict disease genes. Genome Biol. 2003; 4: R75. doi: 10.1186/gb-2003-4-11-r75 gb-2003-4-11-r75 [pii].

39. Stuart JM, Segal E, Koller D, Kim SK. A gene-coexpression network for global discovery of conserved genetic modules. Science. 2003; 302: 249-55. doi: 10.1126/science.1087447 1087447 [pii].

40. Goodridge DM, Sloan JA, LeDoyen YM, McKenzie JA, Knight WE, Gayari M. Risk-assessment scores, prevention strategies, and the incidence of pressure ulcers among the elderly in four Canadian health-care facilities. Can J Nurs Res. 1998; 30: 23-44. doi: