

## MYELOID NEOPLASIA

# Recurrent noncoding somatic and germline *WT1* variants converge to disrupt MYB binding in acute promyelocytic leukemia

Huan Song,<sup>1,\*</sup> Yabin Liu,<sup>1,2,\*</sup> Yun Tan,<sup>1</sup> Yi Zhang,<sup>1</sup> Wen Jin,<sup>1,3</sup> Li Chen,<sup>1</sup> Shishuang Wu,<sup>1</sup> Jinsong Yan,<sup>4</sup> Junmin Li,<sup>1</sup> Zhu Chen,<sup>1,2</sup> Saijuan Chen,<sup>1,2</sup> and Kankan Wang<sup>1-3</sup>

<sup>1</sup>Shanghai Institute of Hematology, State Key Laboratory of Medical Genomics, National Research Center for Translational Medicine at Shanghai, Ruijin Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai, China; <sup>2</sup>School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai, China; <sup>3</sup>Sino-French Research Center for Life Sciences and Genomics, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China; and <sup>4</sup>Department of Hematology, the Second Hospital of Dalian Medical University, Dalian, China

## KEY POINTS

- Noncoding mutations occur preferentially in active enhancers in APL.
- Recurrent noncoding somatic mutations and germline variants converge to decrease *WT1* expression by disrupting MYB binding.

**Genetic alternations can occur at noncoding regions, but how they contribute to cancer pathogenesis is poorly understood. Here, we established a mutational landscape of cis-regulatory regions (CREs) in acute promyelocytic leukemia (APL) based on whole-genome sequencing analysis of paired tumor and germline samples from 24 patients and epigenetic profiling of 16 patients. Mutations occurring in CREs occur preferentially in active enhancers bound by the complex of master transcription factors in APL. Among significantly enriched mutated CREs, we found a recurrently mutated region located within the third intron of *WT1*, an essential regulator of normal and malignant hematopoiesis. Focusing on noncoding mutations within this *WT1* intron, an analysis on 169 APL patients revealed that somatic mutations were clustered into a focal hotspot region, including one site identified as a germline polymorphism contributing to APL risk. Significantly decreased *WT1* expression was observed in APL patients bearing somatic and/or germline noncoding *WT1* variants. Furthermore, biallelic *WT1* inactivation was**

**recurrently found in APL patients with noncoding *WT1* variants, which resulted in the complete loss of *WT1*. The high incidence of biallelic inactivation suggested the tumor suppressor activity of *WT1* in APL. Mechanistically, noncoding *WT1* variants disrupted MYB binding on chromatin and suppressed the enhancer activity and *WT1* expression through destroying the chromatin looping formation. Our study highlights the important role of noncoding variants in the leukemogenesis of APL.**

## Introduction

Although noncoding regions account for >98% of the human genome, genetic alternations occurring at noncoding regions remain poorly understood in cancer pathogenesis. This is largely owing to the challenging interpretation of noncoding variants and the limited understanding of their functional consequences. Emerging evidence has revealed that most noncoding somatic mutations are passenger events without functional consequence.<sup>1,2</sup> Therefore, it is essential to distinguish functional noncoding mutations from these passenger mutations.

Noncoding regions that contain gene regulatory elements are recognized as critical determinants of gene expression, namely cis-regulatory elements (CREs).<sup>3-5</sup> These regulatory noncoding regions are bound by transcriptional regulators, such as transcription factors and coregulators. The dynamic interaction/dissociation between transcriptional regulators and CREs

determines the transcriptional activity and chromatin conformations of nearby genes.<sup>6</sup> Furthermore, recent studies have shown that pathogenic noncoding variants are often located in CRE regions and dysregulate gene expression through redepositing or reshaping cis-regulatory elements to promote malignant transformation.<sup>7-9</sup> Thus, the identification of cancer-associated noncoding variants based on CREs can expand our understanding of the epigenetic control of malignancies.

Acute promyelocytic leukemia (APL) is pathologically characterized by the accumulation of promyelocytic blast cells in bone marrow and peripheral blood and cytogenetically depicted by a typical t(15;17) translocation in most patients. Mounting evidence has demonstrated that the PML/RAR $\alpha$  fusion protein, generated by t(15;17), mainly impairs neutrophil differentiation.<sup>10</sup> However, accumulating data suggest PML/RAR $\alpha$  alone is not sufficient to cause APL, and multiple genetic events are required for the development

of APL.<sup>11-13</sup> However, current knowledge on genetic events contributing to APL development is singularly restricted to mutations in coding regions, with the most recurrently found in *FLT3*, *WT1*, and *RAS* (*N-RAS* and *K-RAS*).<sup>14-16</sup> There is unmet need for systematic discovery of noncoding mutations/variants in the APL genome, advancing our understanding of APL pathogenesis.

The Wilms tumor 1 (*WT1*) gene encodes a zinc-finger transcription factor (TF), an important regulator of normal and malignant hematopoiesis. The role of *WT1* is dynamic and context dependent. During normal hematopoiesis, *WT1* in early undifferentiated cells is essential for the survival or the quiescence of uncommitted cells, but in committed cells, *WT1* is critical for differentiation.<sup>17</sup> In leukemogenesis, the investigation of *WT1* is of a particular challenge because both overexpression and loss-of-function coding mutations are found in patients. On the one hand, *WT1* expression has been reported to be of prognostic importance, with higher levels associated with relapse and poor overall survival.<sup>18</sup> On the other hand, coding *WT1* mutations are predominantly nonsense and frameshift, resulting in loss-of-function truncated proteins. *WT1* mutations are mutually exclusive with *TET2* and *IDH1/2* mutations, establishing an epigenetic function of *WT1*.<sup>19,20</sup> Furthermore, mice studies have shown that *Wt1* haploinsufficiency can enhance the self-renewal capacity of hematopoietic stem and progenitor cells and cooperate with *FLT3*-internal tandem duplication to induce leukemogenesis.<sup>21</sup> In APL, *WT1* is one of the most frequently mutated genes, and the prevalence of coding mutations is even higher than that in other hematological malignancies.<sup>14</sup> These coding alterations show a high incidence in relapsed APL patients.<sup>16</sup>

Here, we systematically analyzed noncoding mutations that mapped to regulatory regions using whole-genome sequencing (WGS) and regulome profiling in primary APL samples. Among mutated CREs, we found a frequently mutated focal region within the third intronic enhancer of *WT1*. Further analysis demonstrated that these noncoding somatic and/or germline *WT1* variants diminished the MYB binding of the *WT1* intronic enhancer and the enhancer-promoter interactions to decrease *WT1* expression in APL cells.

## Methods

### Patient samples

This study was approved by the institutional review board from Ruijin Hospital affiliated to Shanghai Jiao Tong University School of Medicine. The written informed consent was obtained from all participants. This study was also reviewed by the Chinese Ministry of Science and Technology for the Review and Approval of Human Genetic Resources (approval no. 2021BAT4751). We included 169 APL patients between 18 and 77 years of age. Tumor samples were obtained from bone marrow collected at diagnosis. For paired normal samples, 149 of 169 were extracted from peripheral blood or bone marrow of the remission stage from the same patient. For those suffering from early death (20/169), paired normal samples were extracted from their hair follicles.

### Calling somatic mutations

Somatic single nucleotide variants (SNVs) and small insertions and deletions (INDELs) were called using GATK Mutect2,<sup>22</sup>

Strelka,<sup>23</sup> VarScan2 2.4.1,<sup>24</sup> and VarDict 1.5.1.<sup>25</sup> For SNVs, SomaticSniper 1.0.5.0<sup>26</sup> and MuSE 1.0<sup>27</sup> were also used. Each variant caller was run using default parameters or according to those recommended by the authors where available. Furthermore, somatic SNVs and INDELs identified by  $\geq 2$  somatic mutation calling tools were retained for subsequent analysis. All variants were annotated by Annovar.<sup>28</sup>

### ChIP-seq peak calling

Peaks from chromatin immunoprecipitation sequencing (ChIP-seq) data were called using MACS2.<sup>29</sup> A *q* value threshold of 0.01 was used to call peaks for transcription factor ChIP-seq data, and a *P* value cutoff of 1e-9 was used for histone modification ChIP-seq data.

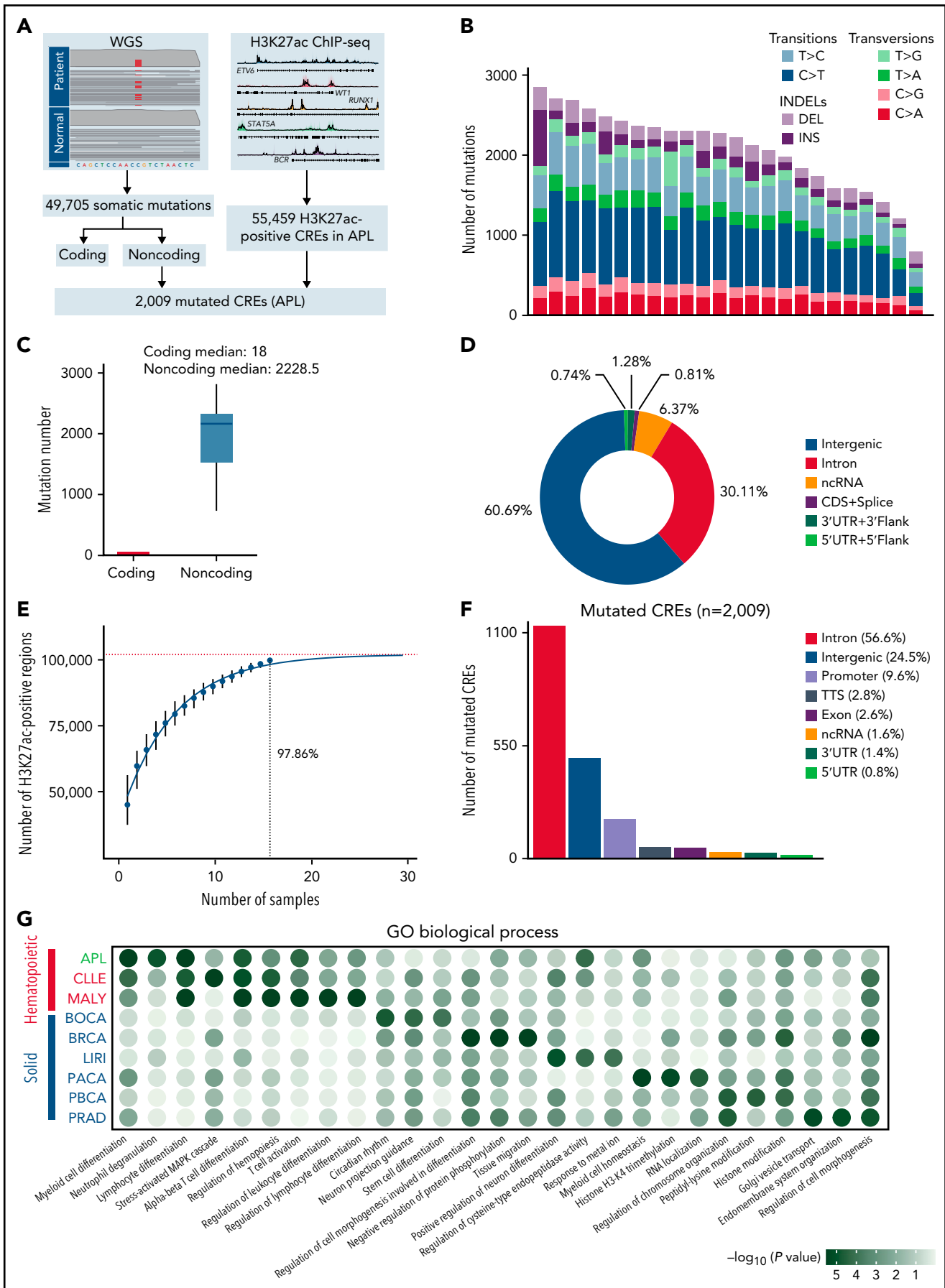
## Results

### Comprehensive analysis identifies mutated noncoding regulatory regions in APL

CREs represent a group of functional sequences in the noncoding genome.<sup>30</sup> To systematically identify APL-associated somatic noncoding mutations, we generated a mutational landscape of CREs by performing WGS and H3K27ac ChIP-seq on APL patients (Figure 1A, supplemental Table 1). First, we performed WGS analysis on paired tumor and normal samples from 24 APL patients (supplemental Table 2), identifying 49 705 somatic alterations with an average of 2071 alterations per patient (supplemental Table 3). Somatic SNVs were predominantly transitions (83%) (Figure 1B). As expected, the number of noncoding mutations outnumbered coding mutations (Figure 1C), with as high as 99% found in noncoding regions (Figure 1D).<sup>31</sup> Next, we defined APL-associated CREs by performing H3K27ac ChIP-seq on 16 APL patients (10 also had WGS; supplemental Table 4). Our modeling analysis suggested that the number of identified CREs had reached  $\sim 98\%$  saturation with a sample size of 16 (Figure 1E; details in supplemental Methods). A total of 55 459 CREs were identified (supplemental Table 5). Finally, we obtained 2009 APL-associated mutated CREs by restricting those CREs with somatic mutations (supplemental Table 6). These mutated CREs were mainly located in intronic regions (56.6%), intergenic regions (24.5%), and promoter and 5'UTR regions (10.4%) (Figure 1F). To assess the functional relevance of CRE mutations, we performed functional enrichment analysis on mutated CREs in APL vs other types of hematopoietic malignancies and solid tumors (details in supplemental Methods; supplemental Tables 7 and 8). We found that GO terms related to myeloid leukemia, such as myeloid differentiation and neutrophil degranulation, were exclusively enriched in mutated CREs in APL, suggesting that these mutated CREs contributed to myeloid dysfunction (Figure 1G). Moreover, the hematopoietic-specific GO terms, such as those related to the regulation of hematopoiesis and immunity, were consistently enriched in hematopoietic malignancies (Figure 1G). The data suggested that somatic noncoding mutations on CREs were biologically meaningful in APL development.

### Somatic noncoding mutations prefer to accumulate on chromatin bound by master transcription factors essential for APL

It is recognized that CREs may act as the regulatory platform enabling combinatorial regulations involving master transcription



**Figure 1.**

factors and coregulators.<sup>32,33</sup> We then explored whether somatic noncoding mutations on CREs may affect the forming of the transcriptional regulatory complex in APL, and more specifically, which transcription factors are affected and how. Master transcription factors usually form the core transcriptional regulatory circuitry (CRC) (ie, constructing the tightly integrated networks to establish and maintain lineage-specific transcriptional programs and forming the interconnected feed-forward regulatory loops).<sup>34-36</sup> To predict master transcription factors in APL, we constructed the CRC models using a CRC mapper algorithm<sup>35</sup> based on the super-enhancer profiles of these 16 APL patients (details in supplemental Methods). As shown in Figure 2A, the master transcription factors captured in the CRC modeling were PML/RAR $\alpha$ , SPI1 (PU.1), MYB, RUNX1, ETV6, IRF1, GFI1, KLF13, MEF2D, and NFE2. As expected, PML/RAR $\alpha$ , the initiating oncogenic fusion protein in APL, contributed to form the master transcription factor-mediated complex in APL.<sup>10,37-39</sup> PU.1, MYB, GFI1, RUNX1, and ETV6 are well-known transcriptional regulators for myeloid development. Their disruption has been reported to contribute to the malignant transformation of myeloid leukemia<sup>40-44</sup>, and NFE2, KLF13, and MEF2D are involved in the hematopoiesis, such as erythropoiesis and lymphopoiesis.<sup>45-47</sup> To prove that mutated CREs are targeted by the above master transcription factors, we performed genome-wide occupancy analysis of MYB, GFI1, and PU.1 (ChIP-seq experiments, detailed in supplemental Methods) and the remaining TFs (based on the published ChIP-seq data).<sup>31,48-50</sup> As shown in Figure 2B and supplemental Table 9, the binding regions for these master transcription factors were significantly enriched within the mutated CREs. Comparing mutated CREs vs nonmutated CREs, the motif enrichment analysis revealed the enrichment for DNA recognition motifs recognized by these transcription factors (Figure 2C), validating the findings obtained from ChIP-seq data (Figure 2B). These data indicated that somatic noncoding mutations preferred to accumulate on chromatin regions bound by master transcription factors essential for APL.

Next, to determine how somatic noncoding mutations affected these APL-associated master transcription factor complex, we evaluated the mutation burden on binding sites for each transcription factor. We calculated the mutation frequency over binding regions centered on DNA recognition motifs and the surrounding area. As shown in Figure 2D, master transcription factors bound on mutated CREs displayed 3 major patterns. First, APL-associated somatic noncoding mutations were significantly enriched at the binding sites of MYB, PU.1, and IRF1 (Z score = 8.12, 6.07, and 3.89, respectively: the left panel), all of which are known pioneer factors. These pioneer factors have been reported to be capable of accessing, opening silent or

compacted chromatin, and recruiting downstream transcription factors to establish lineage-specific transcriptional programs.<sup>51-53</sup> Second, a moderate enrichment was found on the 50- to 100-bp flanking regions of the binding sites of ETV6 and GFI1 (Z score = 3.10 and 2.56, respectively: the middle panel), suggesting their close interactions with the above pioneer factors. Third, no enrichment was observed for the remaining master transcription factors (ie, PML/RAR $\alpha$ , RUNX1, MEF2D, KLF13, and NFE2 [the right panel]). It suggested that this group of transcription factors might act as transcriptional coregulators without directly binding to DNA, consistent with the previous findings from our group and others.<sup>39,54,55</sup> Together, these results indicated that somatic noncoding mutations tended to target the motifs of IRF1, PU.1, and MYB and subsequently influenced the master transcription factor complex in APL.

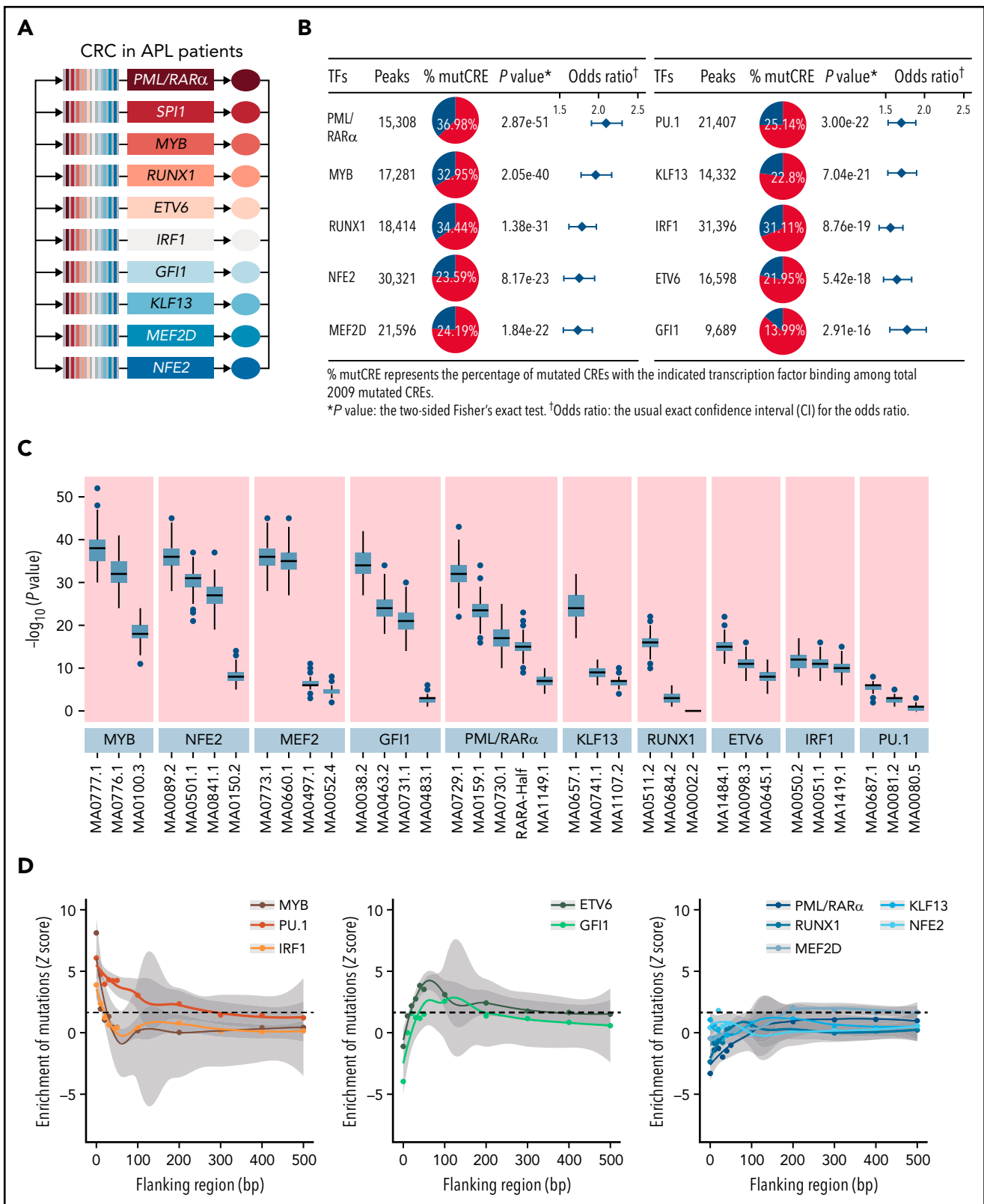
### Recurrent noncoding *WT1* variants are found at both somatic and germline levels

It is well-recognized that functional somatic noncoding mutations should be observed recurrently and may exert impacts on expression and phenotype.<sup>56-58</sup> Therefore, we further performed RNA-seq on APL patients with WGS data (supplemental Table 10) and searched for functional mutated CREs, considering (1) the recurrence occurring within a small region (ie, 50 bp), (2) the transcriptional outcome of the CRE mutation by comparing the expression levels of the respective target gene in mutated and nonmutated patients, and (3) the phenotypic relevance (details in supplemental Methods). Among mutated CREs, we identified 38 functional regions harboring somatic noncoding mutations, which affected the expression levels of leukemia-related genes, such as *WT1*, *ETS1*, *PLCG2*, *PRKCB*, *PDGFA*, and *CD247* (Figure 3A; supplemental Figure 1; supplemental Table 11).

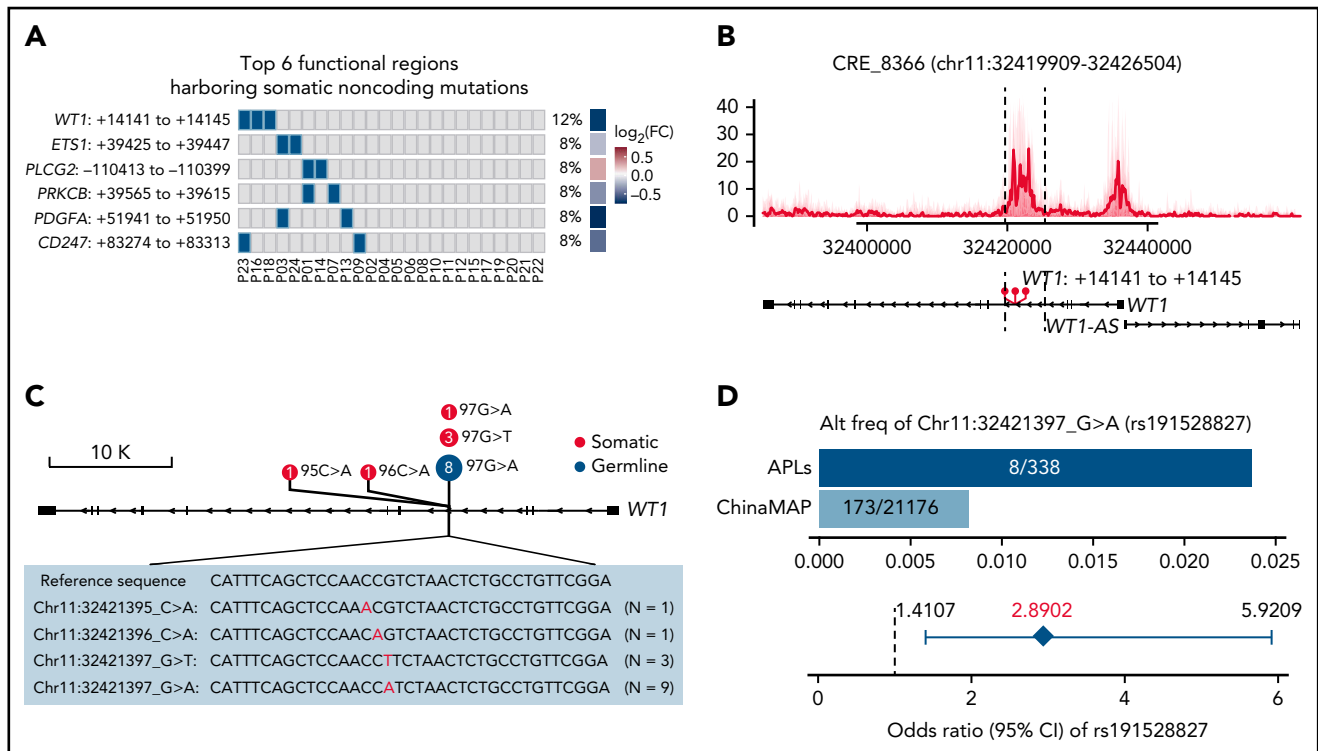
Given the critical and complicated role of *WT1* in leukemia development, we sought to investigate in more detail the recurrent region harboring somatic noncoding mutations located in the *WT1* third intron (Figure 3B; supplemental Figure 2; supplemental Table 12). Sanger sequencing validation in 169 APL patients (including 24 WGS cases) revealed 6 cases with somatic noncoding mutations (Figure 3C; supplemental Figure 3; supplemental Table 13). These mutations on the *WT1* CRE were located within a very focal region (chr11:32421395-97) (ie, chr11:32421395 C>A [n = 1], chr11:32421396 C>A [n = 1], chr11:32421397 G>T [n = 3], and chr11:32421397 G>A [n = 1]) (Figure 3C).

Of note, chr11:32421397 G>A, a germline polymorphism (rs191528827), was also present in 8 of 169 APL patients (4.73%) (Figure 3C; supplemental Figures 2 and 3; supplemental Tables 12

**Figure 1. Identification of mutated noncoding regulatory regions in APL by WGS and H3K27ac ChIP-seq data.** (A) Schematic diagram for analyzing mutated CREs in APL. (B) The number of somatic single-nucleotide variants and short insertions/deletions in each patient. Different colors indicate different types of mutations. (C) Repartition of the 49705 somatic mutations identified across the coding regions and noncoding regions. Each box plot represents the median, interquartile range, and minimum and maximum quartile of the mutation number. (D) Genomic localization of somatic mutations annotated using RefSeq hg38. (E) Saturation analysis for H3K27ac-positive regions identified from ChIP-seq across 16 APL samples. Individual points represent median peaks per sample added, and error bars represent standard deviations from the mean. (F) Genomic distribution of the mutated CREs over exons, promoter (-1 kb to +100 bp of the transcription start site), 3' UTR, 5' UTR, noncoding RNA (ncRNA), transcription termination site (TTS) (-100 bp to +1 kb of the TTS position), intron, and intergenic regions. (G) Enriched gene ontology (GO) terms within genes regulated by mutated CREs in APL compared with other types of hematopoietic malignancies and solid cancers. Other hematopoietic malignancies include chronic lymphocytic leukemia (CLLE) and malignant lymphoma (MALY). Solid cancers include bone cancer (BOCA), breast cancer (BRCA), liver cancer (LIRI), pancreatic cancer (PACA), pediatric brain cancer (PBCA), and prostate adenocarcinoma (PRAD). The variant call format files of WGS data for other cancer types were downloaded from the Pan-Cancer Analysis of Whole Genomes project, and the H3K27ac ChIP-seq data of these cancer types were downloaded from the Gene Expression Omnibus database. The bubble color indicates the P value. CDS, coding sequence.



**Figure 2. Somatic mutations in CREs prefer to accumulate on chromatin regions bound by master transcription factors essential for APL.** (A) Master transcription factors identified by the CRC modeling using H3K27ac ChIP-seq data of 16 APL samples. (B) Enrichment of binding regions for master transcription factors within mutated CREs. The enrichment level of binding regions for the indicated transcription factor in mutated CREs was analyzed by the Fisher's exact test. (C) Enrichment analysis of DNA recognition motifs for the identified master transcription factors within mutated CREs compared with nonmutated CREs. The position weight matrices of master transcription factors and their paralogs for motif enrichment analysis were downloaded from JASPAR (<http://jaspar.genereg.net>). The top 3 enriched motifs (JASPAR ID) for the indicated transcription factor family were listed (eg, MA0100.3, MA0776.1, and MA0777.1 for MYB, MYBL1, and MYBL2, respectively). Box plot represents the median, interquartile range, and minimum and maximum quartile from motif enrichment analysis of 200 random subsampling from background regions. (D) Mutation enrichment analysis for master transcription factors over binding regions centered on DNA recognition motifs and the surrounding area. Z scores are computed based on a permutation test, and the dashed line indicates a *P* value < .05 significance threshold. mutCRE, mutated CRE.



**Figure 3. Recurrent noncoding somatic and germline variants are located in the third intron of WT1.** (A) Top 6 functional regions harboring somatic noncoding mutations identified by the integrative analysis. The abscissa represents each patient, the ordinate represents candidate loci containing noncoding mutations for the indicated gene, the blue represents the sample with the specified mutation, and the gray represents the sample that does not contain the specified mutation. Also included is the log<sub>2</sub> fold-change of associated genes by comparing the expression levels of respective target genes in mutated and nonmutated samples. (B) Detailed annotation of the recurrent somatic noncoding mutation-containing locus in the third intron of WT1. The H3K27ac peaks of each sample are shown as a transparent area plot. The thick line represents the median profile from all samples. The height of peaks from each sample was determined by the H3K27ac read density of the corresponding sample within the region. (C) Lollipop plot showing the distribution and classes of noncoding variants in the third intron of WT1. The red point represents the somatic mutation, and the blue point represents the germline mutation. The size of the point represents the number of a given mutation type. (D) Association of rs191528827 with APL risk. Upper panel: the bar plot shows the alternative allele frequency in 169 APL patients of this study and a nontumor cohort in the ChinaMAP. Lower panel: the diamond-shaped point represents the odds ratio, and the error bar represents the 95% confidence interval of the odds ratio. CI, confidence interval; FC, fold change.

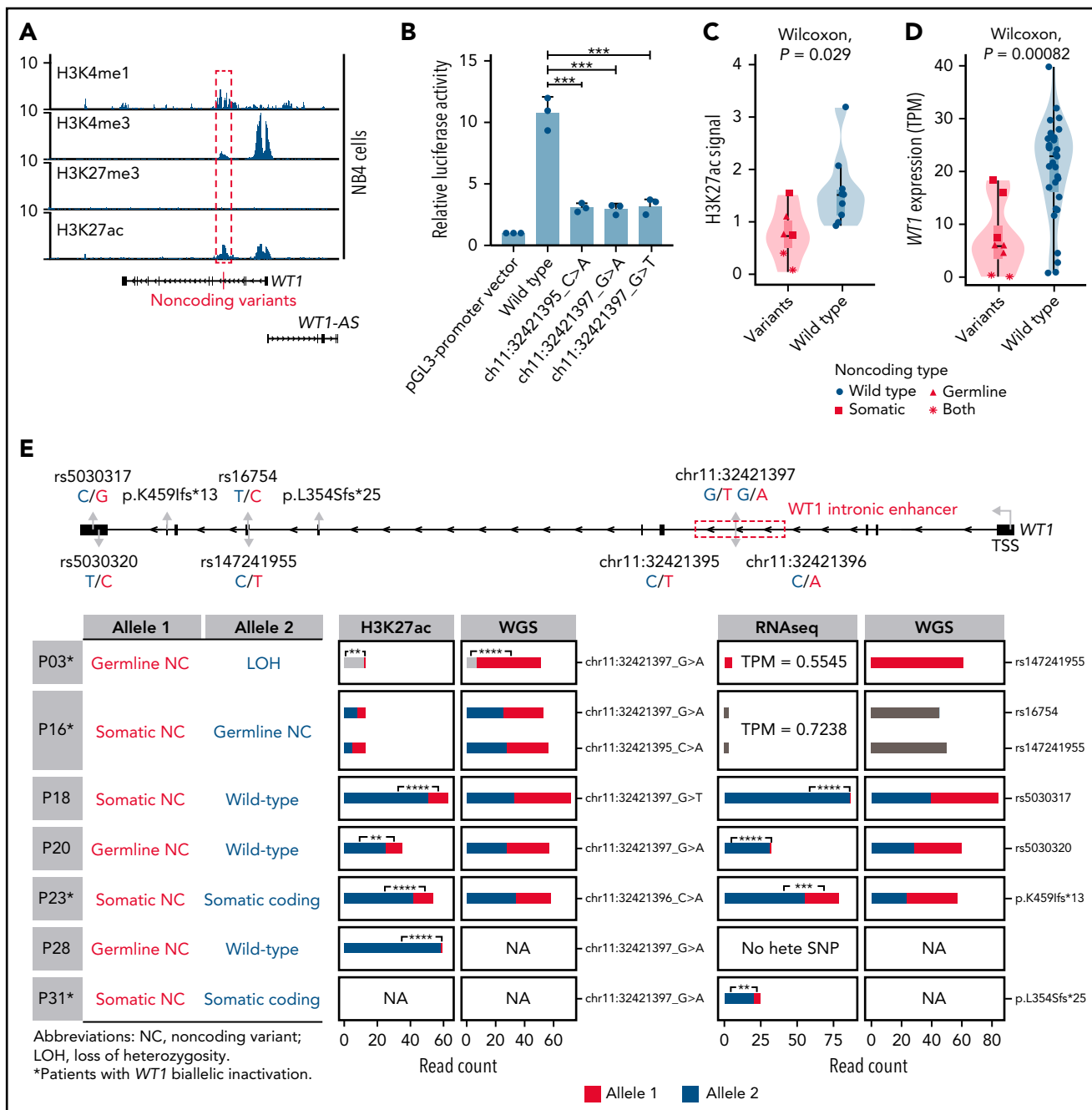
and 13). The allele frequency of this single nucleotide polymorphism (SNP) is 0.82% in the Chinese population (the China Metabolic Analytics Project, ChinaMAP) (supplemental Table 14).<sup>59</sup> We then asked whether this germline variant is a risk-associated SNP in APL. Statistical analysis revealed a significant excess of this site in APL patients (odds ratio, 2.8902; 95% confidence interval, 1.4107-5.9209; Fisher's exact test *P* value, .0091) (Figure 3D), suggesting that chr11:32421397 G>A was a newly identified risk variant for APL.

### Noncoding WT1 somatic/germline variants lead to allele-specific loss of H3K27ac binding and allele-specific repression of WT1 expression

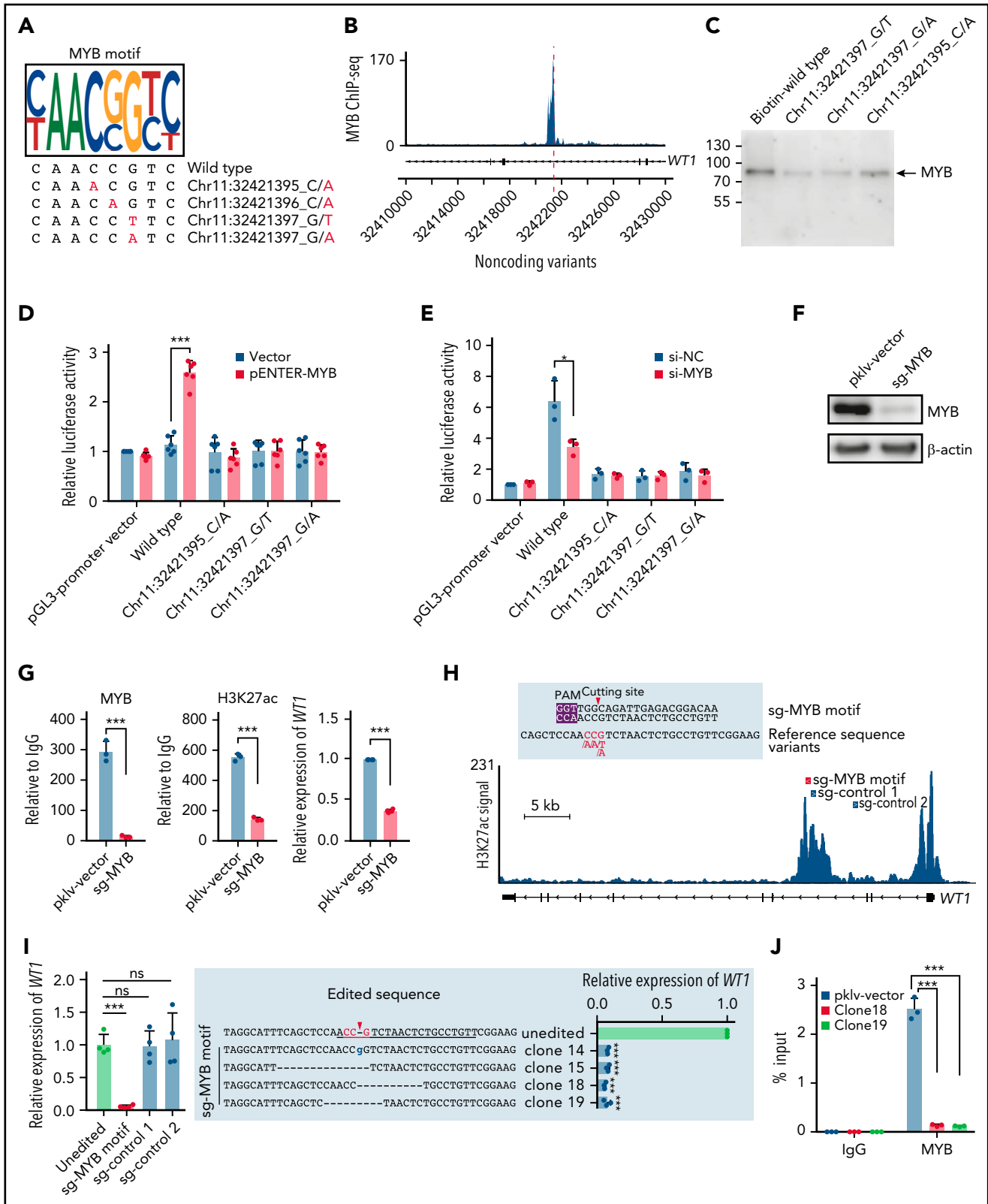
To investigate the epigenetic and regulatory impacts of noncoding WT1 somatic/germline variants on WT1 expression, we first analyzed the chromatin state of the surrounding region for H3K4me1, H3K4me3, H3K27ac, and H3K27me3 in hematopoietic cell lines. This region, marked by a high abundance of H3K4me1 and H3K27ac signals, was indeed an active enhancer in hematopoietic cell lines without these alterations, including NB4, K562, and Kasumi-1 cells (Figure 4A; supplemental Figure 4). Furthermore, luciferase reporter assays confirmed the enhancer activity of this region, and mutations mimicking

noncoding WT1 somatic/germline variants observed in APL patients significantly impaired the enhancer activity (Figure 4B).

Furthermore, H3K27ac signals around this enhancer region exhibited remarkably lower signals in patients with noncoding somatic/germline WT1 variants (*n* = 6) than in those without such variants (*n* = 10) (Figure 4C). Accordingly, patients with these variants showed a significantly lower expression of WT1 (Figure 4D). Notably, 2 mutant cases harboring 2 noncoding mutations/variants (P03 and P16) showed the extremely low H3K27ac signals and WT1 expression (Figure 4C-E; supplemental Figure 5). P16 harbored a somatic mutation and a germline variant on 2 alleles, respectively. P03 contained the same germline variant in a homozygous state due to copy-neutral loss of heterozygosity. We further analyzed H3K27ac ChIP-seq data in 4 cases that were heterozygous for noncoding WT1 variants (Figure 4E; P18, P20, P23, and P28). The allele with noncoding WT1 variants exhibited significantly lower H3K27ac ChIP-seq reads compared with the wild-type allele, showing the evidence of allelic imbalance in noncoding WT1 variants (Figure 4E; supplemental Table 15). Moreover, RNA-seq reads from patients (P18, P20, P23, and P31) with heterozygous exonic SNPs or exonic mutations revealed an allelic imbalance in WT1 expression (Figure 4E; supplemental Table 16). The data



**Figure 4. Noncoding *WT1* variants lead to allele-specific downregulation of *WT1*, and biallelic *WT1* inactivation is recurrently found in APL patients.** (A) ChIP-seq tracks of H3K4me1, H3K4me3, H3K27me3, and H3K27ac occupancy in the *WT1* locus in NB4 cells without noncoding *WT1* variants. The site for noncoding *WT1* variants is marked with a red line. (B) Impairment of the enhancer activity by noncoding *WT1* variants. Wild-type or mutated intronic sequences were cloned into the pGL3-promoter plasmid, and the luciferase reporter assay was performed in NB4 cells (without noncoding *WT1* variants). Data are plotted as means plus or minus standard deviation ( $n = 3$ ).  $***P < .001$ . (C) Decreased H3K27ac signals on the intronic region of *WT1* in noncoding *WT1* variants containing APL patients compared with patients without such variants. The  $P$  value was calculated by the Wilcoxon test. (D) Significant downregulation of *WT1* expression in APL patients with noncoding *WT1* mutations/variants compared with patients without these alterations. Triangle represents the case with the somatic mutation. Quadrilateral represents the case with the germline variant. Star represents the case harboring both somatic mutation and germline variant. Circle represents the case without the noncoding mutation/variant. (E) Allelic imbalance of H3K27ac ChIP-seq signals and *WT1* expression in noncoding *WT1* variants for APL patients. Upper panel: diagram of the *WT1* locus. Left panel: the mutation status in each allele of *WT1* in APL samples with noncoding *WT1* variants. Middle panel: allelic distribution of H3K27ac ChIP-seq reads with and without the indicated noncoding variant. Right panel: allelic distribution of *WT1* expression based on RNA-seq reads mapped to the indicated heterozygous exonic SNPs or exonic mutations. The red and blue colors represent different alleles, and gray represents the indistinguishable alleles. One-tailed binomial test (expected probability 0.5) was performed: \* $P < .05$ ; \*\* $P < .01$ ; \*\*\* $P < .001$ ; \*\*\*\* $P < .0001$ . LOH, loss of heterozygosity; NA, not available; NC, noncoding variant; No hete SNP, no heterozygous exonic SNP.



**Figure 5. Noncoding variants in the intronic enhancer of WT1 disrupt MYB binding and MYB-mediated transcriptional activation of WT1.** (A) Noncoding WT1 variants disrupted the binding site of MYB. The binding motif of MYB was determined by the MYB-specific ChIP-seq data for NB4 cells (detailed in supplemental Methods). The lower panel shows patient-derived mutation sequences with mutated nucleotides labeled in red. (B) ChIP-seq tracks showing MYB binding to the WT1 intronic enhancer in NB4 cells (without noncoding WT1 variants). (C) DNA pull-down assays with anti-MYB antibodies and cell lysates from HEK293T cells stably expressing MYB. (D-E) The dual-luciferase assays on the enhancer activity of nonmutated and mutated constructs with or without MYB expression in 293T cells (D) and NB4 cells (E). The firefly luciferase activity was normalized to the renilla luciferase and presented as the ratio relative to the pGL3-SV40 promoter vector. (F) The knockout efficiency of sg-MYB was tested by western blot. (G) MYB knockout significantly reduced the H3K27ac signals of the WT1 enhancer and expression of WT1. MYB binding



suggested that the noncoding *WT1* variants resulted in the mutant allele-specific loss of H3K27ac binding around this enhancer region and subsequent allelic-specific *WT1* repression.

We also provided evidence supporting biallelic inactivation of *WT1* expression. Among 11 patients suitable for detecting allele-specific expression, 4 (36%) showed biallelic inactivation, including the above 2 cases with 2 noncoding variants (P03 and P16) and additional 2 patients positive for both noncoding variants and coding *WT1* mutations (P23 and P31) (Figure 4E; supplemental Table 13). P03 and P16 (with 2 noncoding mutations/variants on each allele, respectively) exhibited extremely low *WT1* expression on both alleles. P23 and P31 (with noncoding and coding mutations on each allele, respectively) showed allele-specific expression of the coding mutations (p.K459Ifs in P23 and p.L354Sfs in P31). In these 2 cases, the inactivation of 1 allele of the *WT1* gene was due to somatic noncoding *WT1* mutations, supported by the absence of reads without the exonic mutations (Figure 4E). The inactivation of the other allele was caused by coding *WT1* mutations, which are frameshift insertions encoding a truncated protein. Together, the biallelic inactivation of *WT1* expression was recurrently found in APL patients, strongly suggesting that *WT1* functions as a tumor suppressor gene.

### Impaired MYB binding caused by noncoding *WT1* variants leads to *WT1* downregulation

Next, we explored the potential mechanism of how noncoding *WT1* variants repressed *WT1* expression. Motif scanning of this region showed that the wild-type sequence perfectly matched the MYB motif, whereas the mutated nucleotides altered the core positions of the MYB motif (Figure 5A). MYB binding on this enhancer region was confirmed in cell lines without noncoding *WT1* mutations (Figure 5B; supplemental Figure 6). Furthermore, we performed a series of experiments to evaluate the effects of impaired MYB binding. First, we conducted DNA pulldown assays and found that MYB binding was significantly reduced in the presence of probes with noncoding *WT1* variants (Figure 5C). Second, the luciferase activity of the *WT1* enhancer was significantly increased upon MYB overexpression in HEK293T cells, a cell line without endogenous MYB expression (Figure 5D), suggesting the *WT1* enhancer was activated by MYB. Third, the MYB-dependent enhancer activity was significantly impaired when the MYB overexpression plasmid was cotransfected with the luciferase constructs with noncoding *WT1* variants (Figure 5D). Knockdown of endogenous MYB expression in NB4 cells confirmed the above findings (Figure 5E). Forth, MYB knock-out significantly decreased the H3K27ac signals on the *WT1*

enhancer and downregulated *WT1* expression in NB4 cells (Figure 5F-G). Finally, we employed the CRISPR/Cas9 technology to introduce small INDELS into the MYB motif of the *WT1* enhancer in NB4 cells (Figure 5H). We found that the MYB motif-mutated clones showed a remarkably lower expression of *WT1* as compared with controls (including unedited and control sgRNA-transduced cells) (Figure 5I; supplemental Figure 7). This suggested that these clones could imitate patient samples with the noncoding *WT1* mutations/variants. We then performed the MYB-specific ChIP-qPCR assays at the *WT1* enhancer region and found that these engineered NB4 clones showed the depletion of MYB binding compared with controls (Figure 5J). Together, these results indicated the noncoding mutations/variants on the *WT1* enhancer decreased *WT1* expression by destroying MYB-mediated transcriptional activation in APL cells.

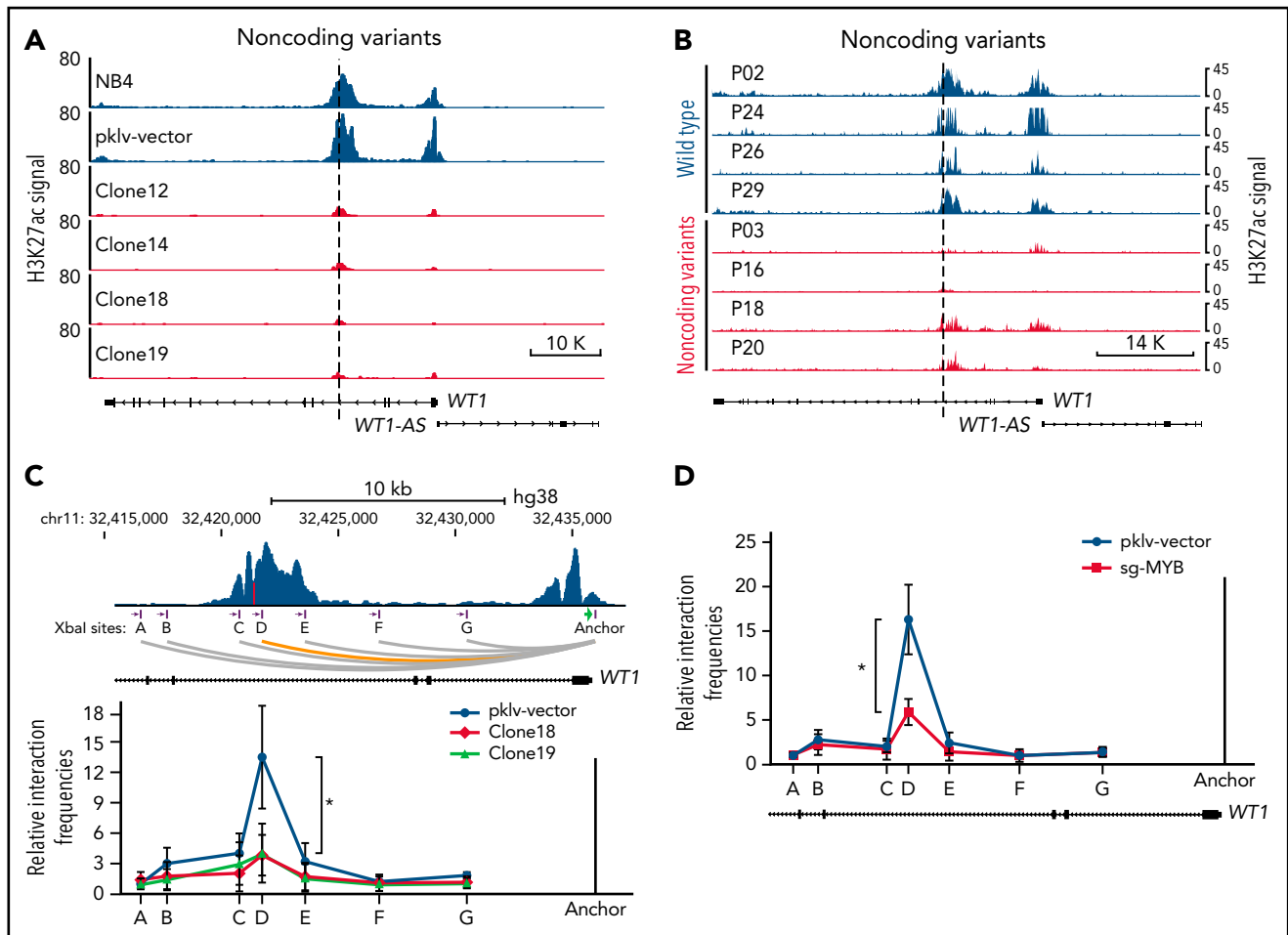
### Noncoding *WT1* variants disrupt the MYB-mediated enhancer-promoter interactions

Interestingly, we found that the H3K27ac signals were also dramatically decreased at the promoter of *WT1*, along with the decrease in H3K27ac signals at the intronic enhancer in MYB motif-mutated clones (Figure 6A). A similar finding was observed in APL patients with noncoding *WT1* variants but not in patients without such variants (Figure 6B). These results implicated that noncoding *WT1* variants might affect *WT1* expression by disrupting long-range chromatin conformations (the interaction between promoters and enhancers) in APL cells. To validate this assumption, we performed chromatin conformation capture (3C) experiments to assess the long-range interaction between the promoter and enhancer at the *WT1* locus in the MYB motif-mutated clones and the parental cells. As shown in Figure 6C, the intronic enhancer of *WT1* showed strong interactions with the *WT1* promoter in NB4 cells, but when the intronic site was mutated, the enhancer-promoter interactions were lost. As the noncoding *WT1* variants altered the MYB motif (Figure 5A), we further knocked out MYB expression by CRISPR/Cas9 and found the enhancer-promoter interactions on the *WT1* locus were dramatically decreased after MYB knockout (Figure 6D). Together, these results demonstrated that noncoding *WT1* variants interfered with MYB binding to the enhancer and that this interference disrupted the enhancer-promoter interactions on the *WT1* genomic locus and subsequently decreased *WT1* expression.

## Discussion

Emerging studies have revealed the importance of noncoding mutations in cancer development, and now, a major challenge is the functional interpretation of disease-associated noncoding

**Figure 5 (continued)** signals (left), H3K27ac signals (middle), and relative expression of *WT1* (right) were determined. (H) The schematic diagram of the single guide RNA (sgRNA) target sites in the H3K27ac ChIP-seq track for NB4 cells. The red box indicates the sgRNA targeting the MYB motif, and the blue boxes indicate control sgRNAs targeting regions surrounding the MYB motif within intron 3. The protospacer adjacent motif (PAM) sequence is shown in the purple frame. The red arrowhead indicates the expected cleavage site. (I) Clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated protein 9 (Cas9)-mediated targeted mutagenesis of the MYB motif within the third intronic region of *WT1* inhibited *WT1* expression. The left panel shows the relative *WT1* expression in mutated clones targeted by the indicated sgRNAs and the unedited clones. CRISPR/Cas9-edited sequences were validated by Sanger sequencing. Data are normalized against the mean expression level of the unedited clones and are plotted as means plus or minus SD (n = 4). The middle panel shows the edited genomic sequences from CRISPR/Cas9-edited single-cell clones. The right panel shows the mRNA levels of *WT1* expression in parental and CRISPR/Cas9-edited clones. Data are represented as the fold change relative to the expression of the parental cells and are plotted as means plus or minus SD (n = 3). (J) Disruption of the MYB motif resulted in the loss of MYB binding at the *WT1* enhancer region. ChIP-quantitative polymerase chain reaction (qPCR) assays were performed to detect MYB binding using an antibody specific for MYB in the representative MYB motif-mutated clones and the parental cells. Data are calculated as the percentage of input and are plotted as means plus or minus SD (n = 3). The statistical significance was determined using a 2-tailed Student t test. \*P < .05; \*\*P < .01; \*\*\*P < .001. IgG, immunoglobulin G; ns, not significant; si-NC, negative control small interfering RNA (siRNA); si-MYB, siRNA targeting MYB; pklv-vector, sgRNA empty vector control; sg-MYB, sgRNA targeting MYB; SD, standard deviation.



**Figure 6. Alterations in the recurrently mutated site of the *WT1* enhancer disrupt the MYB-mediated enhancer-promoter interactions.** (A-B) ChIP-seq tracks of H3K27ac signals at the *WT1* locus in the MYB motif-mutated clones and the parental cells (A) and APL patients with or without noncoding *WT1* variants (B). (C) The relative interaction frequencies between the anchor region (the *WT1* promoter) and distal sites (purple bars) in the MYB motif-mutated clones and the parental cells. The relative interaction frequencies were determined by chromatin conformation capture (3C)-qPCR and normalized to the control region (region A). The upper panel is the schematic diagram showing the H3K27ac signals at the *WT1* locus and the design of the 3C assay. *Xba*I restriction sites are indicated by purple blocks, 3C primers are indicated by purple arrows, and the anchor is shown by a green arrow. \* $P < .05$  for comparison between the MYB motif-mutated clones and the parental cells. (D) The relative interaction frequencies analyzed by 3C-qPCR in NB4 cells with or without MYB knockout. \* $P < .05$  for comparison between MYB knockout cells and control cells. pklv-vector, the sgRNA empty vector control; sg-MYB, sgRNA targeting MYB.

mutations. In this study, we made an important contribution, reporting systematic identification and functional characterization of noncoding regulatory mutations based on the combination of WGS data and the CRE catalog from primary APL patients. We identified >2000 H3K27ac-positive CREs containing somatic mutations. Our data analysis revealed that somatic noncoding mutations preferred to accumulate on chromatin bound by master transcription factors essential for hematopoiesis and APL leukemogenesis. Furthermore, we elucidated the mechanisms by which noncoding mutations within the *WT1* intronic enhancer disrupted the binding of MYB on chromatin and thereby suppressed the enhancer activity and gene expression of *WT1* by destroying the chromatin looping formation in APL. Our work provides novel insights into the role of noncoding mutations in the pathogenesis of APL.

Our comprehensive investigations of noncoding *WT1* mutations have addressed several long-standing questions in the pathogenesis of APL. First, we found that the third intron of *WT1* was recurrently mutated in APL patients. Previous studies have

shown a high incidence of coding *WT1* mutations in APL compared with non-APL AML.<sup>14</sup> However, these studies may underestimate the true incidence of *WT1* deregulation in APL; our data showed that ~7.69% of APL contained mutations/variants in the intronic enhancer. Second, biallelic inactivation of *WT1* caused by noncoding and/or coding mutations supports *WT1* acting as a tumor suppressor in APL because 2 inactivating hits are considered the hallmark of tumor suppressor genes.<sup>60,61</sup> Third, we found that one of the noncoding *WT1* mutations also existed as a germline polymorphism (rs191528827) in the Chinese population, and the G to A variant contributed to the disease susceptibility in APL patients. This SNP represents the first description of a clinically relevant noncoding germline variant in APL. This may explain a relatively higher frequency of APL in China, but the actual incidence rate of APL in China is unknown because population-based registries do not distinguish APL from other AML subtypes. This is the first report demonstrating that noncoding *WT1* variants have a role in the susceptibility of APL. Admittedly, the sample size is small, and further confirmation in a larger cohort is needed.

Our study suggested a functional relevance of noncoding *WT1* variants in APL leukemogenesis, and the functional consequences of these variants can be inferred from previous studies. For example, *Wt1* haploinsufficiency can enhance the self-renewal capacity of hematopoietic stem and progenitor cells and cooperate with *FLT3*-internal tandem duplication to induce AML with full penetrance.<sup>21</sup> The growth advantage conferred by the inactivation of *WT1* has been reported in CD34<sup>+</sup> cells expressing PML/RAR $\alpha$ , providing convincing evidence for the tumor suppressor activity of *WT1* in APL.<sup>62</sup> As noncoding *WT1* variants in this intronic enhancer led to decreased *WT1* expression, it is tempting to speculate that the resulting consequences could contribute to the growth advantage of hematopoietic progenitors.

MYB can act as an oncogene in leukemogenesis.<sup>63-65</sup> It has been reported that somatic noncoding mutations can create MYB binding sites to cause MYB-mediated activation of LMO2 or TAL1 in T-ALL.<sup>66,67</sup> Moreover, increased MYB binding to specific MYB binding sites can cause a hijacked GATA2 enhancer driving *EVI1* overexpression in AML.<sup>68</sup> Independent from its role in activating oncogenes through creating or increasing MYB binding, our study expanded our understanding of the mechanisms of MYB in leukemogenesis. We found that noncoding somatic or germline variants in the *WT1* intron 3 destroyed MYB binding and thereby inhibited the expression of the tumor suppressor gene *WT1* in APL. An early work reported MYB-mediated *WT1* transactivation through a MYB binding site,<sup>69</sup> but our identified enhancer was located in a different position and showed highly abundant H3K27ac signals in hematopoietic cells (supplemental Figure 8). Furthermore, we revealed that noncoding *WT1* variants perturbed the MYB-mediated enhancer-promoter chromatin looping on the *WT1* locus. This is consistent with the recent findings on the role of noncoding variants facilitating chromatin looping.<sup>70-72</sup>

PML/RAR $\alpha$  is the initiating factor in APL. We indeed found that PML/RAR $\alpha$  was one of the master TFs for APL cells (Figure 2A). Of note, the less mutation burden was found on the PML/RAR $\alpha$  bound sites (Figure 2D). This was likely explained by the following factors: (1) At a more relaxed degree of DNA-binding specificity, PML/RAR $\alpha$  tends to bind the RARE half site.<sup>38</sup> (2) Instead of direct DNA binding, PML/RAR $\alpha$  can function as cofactors to be interacted with other hematopoietic TFs in mediating the formation of super-enhancers.<sup>39</sup>

This work opens additional important areas for future research. Considering the existence of *WT1* coding mutations in both APL and non-APL AML, it is deducible that noncoding *WT1* variants are also present in non-APL AML. Future studies will aim to pursue this issue using a large patient cohort. Furthermore, we identified an APL risk-associated SNP rs191528827. Our future work will explore the influence of rs191528827 on *WT1*

expression and cancer predisposition in healthy individuals in a population-based and long-term study.

## Acknowledgments

The computations were run on the  $\pi$  2.0 cluster supported by the Center for High Performance Computing at Shanghai Jiao Tong University.

This work was supported by the National Natural Science Foundation of China (81890994), the National Key R&D Program of China (2019YFA0905902), and the Innovative Research Team of High-Level Local Universities in Shanghai.

## Authorship

Contribution: K.W. designed the experiments; H.S., Y.T., and Y.Z. performed the experiments; Y.L. analyzed the data; L.C., J.Y., W.J., and S.W. collected clinical samples; K.W. supervised the study and analysis; L.C. and J.Y. followed up with the patients; K.W., Y.T., H.S., and Y.L. wrote the manuscript; Z.C. and S.C. gave conceptual advice; and all authors discussed the results and implications and reviewed the manuscript.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

ORCID profiles: H.S., 0000-0001-6811-7844; Y.L., 0000-0001-6248-1143; Y.T., 0000-0001-8450-0392; K.W., 0000-0001-7198-2134.

Correspondence: Kankan Wang, Shanghai Institute of Hematology, Ruijin Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, 197 Ruijin Er Rd, Shanghai 200025, China; e-mail: kankanwang@shsmu.edu.cn; and Yun Tan, Shanghai Institute of Hematology, Ruijin Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, 197 Ruijin Er Rd, Shanghai 200025, China; e-mail: ty12260@rjh.com.cn.

## Footnotes

Submitted 28 November 2021; accepted 24 May 2022; prepublished online on *Blood* First Edition 2 June 2022. DOI 10.1182/blood.2021014945.

\*H.S. and Y.L. contributed equally to this work.

Raw and processed WGS, ChIP-seq, and RNA-seq data have been deposited in the National Omics Data Encyclopedia (NODE) under the accession number OEP002856.

Send data sharing requests via e-mail to the corresponding author.

The online version of this article contains a data supplement.

There is a *Blood* Commentary on this article in this issue.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

## REFERENCES

- Corona RI, Seo JH, Lin X, et al. Non-coding somatic mutations converge on the PAX8 pathway in ovarian cancer. *Nat Commun*. 2020;11(1):2020.
- Mazrooei P, Kron KJ, Zhu Y, et al. Cistrome partitioning reveals convergence of somatic mutations and risk variants on master transcription regulators in primary prostate tumors. *Cancer Cell*. 2019;36(6):674-689.e6.
- Ramos-Rodríguez M, Raurell-Vila H, Colli ML, et al. The impact of proinflammatory cytokines on the  $\beta$ -cell regulatory landscape provides insights into the genetics of type 1 diabetes. *Nat Genet*. 2019;51(11):1588-1595.
- NandyMazumdar M, Yin S, Paranjapye A, et al. Looping of upstream cis-regulatory elements is required for CFTR expression in human airway epithelial cells. *Nucleic Acids Res*. 2020;48(7):3513-3524.
- Li J, Zhang T, Ramakrishnan A, et al. Dynamic changes in cis-regulatory occupancy by Six1 and its cooperative

- interactions with distinct cofactors drive lineage-specific gene expression programs during progressive differentiation of the auditory sensory epithelium. *Nucleic Acids Res.* 2020;48(6):2880-2896.
6. Oudelaar AM, Davies JOJ, Hanssen LLP, et al. Single-allele chromatin interactions identify regulatory hubs in dynamic compartmentalized domains. *Nat Genet.* 2018;50(12):1744-1751.
  7. Cherry TJ, Yang MG, Harmin DA, et al. Mapping the cis-regulatory architecture of the human retina reveals noncoding genetic variation in disease. *Proc Natl Acad Sci USA.* 2020;117(16):9001-9012.
  8. Li K, Zhang Y, Liu X, et al. Noncoding variants connect enhancer dysregulation with nuclear receptor signaling in hematopoietic malignancies. *Cancer Discov.* 2020;10(5):724-745.
  9. Akerman I, Maestro MA, De Franco E, et al. Neonatal diabetes mutations disrupt a chromatin pioneering function that activates the human insulin gene. *Cell Rep.* 2021;35(2):108981.
  10. de Thé H, Chen Z. Acute promyelocytic leukaemia: novel insights into the mechanisms of cure. *Nat Rev Cancer.* 2010;10(11):775-783.
  11. Grisolan JL, Wesselschmidt RL, Pelicci PG, Ley TJ. Altered myeloid development and acute leukemia in transgenic mice expressing PML-RAR alpha under control of cathepsin G regulatory sequences. *Blood.* 1997;89(2):376-387.
  12. Zimonjic DB, Pollock JL, Westervelt P, Popescu NC, Ley TJ. Acquired, nonrandom chromosomal abnormalities associated with the development of acute promyelocytic leukemia in transgenic mice. *Proc Natl Acad Sci USA.* 2000;97(24):13306-13311.
  13. Le Beau MM, Davis EM, Patel B, Phan VT, Sohal J, Kogan SC. Recurring chromosomal abnormalities in leukemia in PML-RARA transgenic mice identify cooperating events and genetic pathways to acute promyelocytic leukemia. *Blood.* 2003;102(3):1072-1074.
  14. Madan V, Shyamsunder P, Han L, et al. Comprehensive mutational analysis of primary and relapse acute promyelocytic leukemia [published correction appears in *Leukemia.* 2016;30(12):2430]. *Leukemia.* 2016;30(8):1672-1681.
  15. Lin X, Qiao N, Shen Y, et al. Integration of genomic and transcriptomic markers improves the prognosis prediction of acute promyelocytic leukemia. *Clin Cancer Res.* 2021;27(13):3683-3694.
  16. Lehmann-Che J, Bally C, Letouzé E, et al. Dual origin of relapses in retinoic-acid resistant acute promyelocytic leukemia. *Nat Commun.* 2018;9(1):2047.
  17. Huff V. Wilms' tumours: about tumour suppressor genes, an oncogene and a chameleon gene. *Nat Rev Cancer.* 2011;11(2):111-121.
  18. Nomdedéu JF, Hoyos M, Carricondo M, et al; CETLAM Group. Bone marrow WT1 levels at diagnosis, post-induction and post-intensification in adult de novo AML. *Leukemia.* 2013;27(11):2157-2164.
  19. Rampal R, Alkalin A, Madzo J, et al. DNA hydroxymethylation profiling reveals that WT1 mutations result in loss of TET2 function in acute myeloid leukemia. *Cell Rep.* 2014;9(5):1841-1855.
  20. Wang Y, Xiao M, Chen X, et al. WT1 recruits TET2 to regulate its target gene expression and suppress leukemia cell proliferation. *Mol Cell.* 2015;57(4):662-673.
  21. Pronier E, Bowman RL, Ahn J, et al. Genetic and epigenetic evolution as a contributor to WT1-mutant leukemogenesis. *Blood.* 2018;132(12):1265-1278.
  22. do Valle IF, Giampieri E, Simonetti G, et al. Optimized pipeline of MuTect and GATK tools to improve the detection of somatic single nucleotide polymorphisms in whole-exome sequencing data. *BMC Bioinformatics.* 2016;17(S12 Suppl 12):341.
  23. Saunders CT, Wong WS, Swamy S, Becq J, Murray LJ, Cheetham RK. Strelka: accurate somatic small-variant calling from sequenced tumor-normal sample pairs. *Bioinformatics.* 2012;28(14):1811-1817.
  24. Koboldt DC, Zhang Q, Larson DE, et al. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.* 2012;22(3):568-576.
  25. Lai Z, Markovets A, Ahdesmaki M, et al. VarDict: a novel and versatile variant caller for next-generation sequencing in cancer research. *Nucleic Acids Res.* 2016;44(11):e108.
  26. Larson DE, Harris CC, Chen K, et al. SomaticSniper: identification of somatic point mutations in whole genome sequencing data. *Bioinformatics.* 2012;28(3):311-317.
  27. Fan Y, Xi L, Hughes DS, et al. MuSE: accounting for tumor heterogeneity using a sample-specific error model improves sensitivity and specificity in mutation calling from sequencing data. *Genome Biol.* 2016;17(1):178.
  28. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 2010;38(16):e164.
  29. Zhang Y, Liu T, Meyer CA, et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* 2008;9(9):R137.
  30. Maston GA, Evans SK, Green MR. Transcriptional regulatory elements in the human genome. *Annu Rev Genomics Hum Genet.* 2006;7(1):29-59.
  31. ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature.* 2012;489(7414):57-74.
  32. Whyte WA, Orlando DA, Hnisz D, et al. Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell.* 2013;153(2):307-319.
  33. Yin Y, Morgunova E, Jolma A, et al. Impact of cytosine methylation on DNA binding specificities of human transcription factors. *Science.* 2017;356(6337):eaaj2239.
  34. Chen Y, Xu L, Lin RY, Müschen M, Koeffler HP. Core transcriptional regulatory circuitries in cancer. *Oncogene.* 2020;39(43):6633-6646.
  35. Saint-André V, Federation AJ, Lin CY, et al. Models of human core transcriptional regulatory circuitries. *Genome Res.* 2016;26(3):385-396.
  36. Ott CJ, Federation AJ, Schwartz LS, et al. Enhancer architecture and essential core regulatory circuitry of chronic lymphocytic leukemia. *Cancer Cell.* 2018;34(6):982-995.e7.
  37. Wang ZY, Chen Z. Acute promyelocytic leukemia: from highly fatal to highly curable. *Blood.* 2008;111(5):2505-2515.
  38. Wang K, Wang P, Shi J, et al. PML/RARalpha targets promoter regions containing PU.1 consensus and RARE half sites in acute promyelocytic leukemia. *Cancer Cell.* 2010;17(2):186-197.
  39. Tan Y, Wang X, Song H, et al. A PML/RARα direct target atlas redefines transcriptional deregulation in acute promyelocytic leukemia. *Blood.* 2021;137(11):1503-1516.
  40. McKenzie MD, Ghisi M, Oxley EP, et al. Interconversion between tumorigenic and differentiated states in acute myeloid leukemia. *Cell Stem Cell.* 2019;25(2):258-272.e9.
  41. Ramaswamy K, Forbes L, Minuesa G, et al. Peptidomimetic blockade of MYB in acute myeloid leukemia. *Nat Commun.* 2018;9(1):110.
  42. Hönes JM, Botezatu L, Helness A, et al. GF11 as a novel prognostic and therapeutic factor for AML/MDS. *Leukemia.* 2016;30(6):1237-1245.
  43. Sood R, Kamikubo Y, Liu P. Role of RUNX1 in hematological malignancies [published correction appears in *Blood.* 2018;131(3):373]. *Blood.* 2017;129(15):2070-2082.
  44. Barjesteh van Waalwijk van Doorn-Khosrovani S, Spensberger D, de Knecht Y, Tang M, Löwenberg B, Delwel R. Somatic heterozygous mutations in ETV6 (TEL) and frequent absence of ETV6 protein in acute myeloid leukemia. *Oncogene.* 2005;24(25):4129-4137.
  45. Gasiorek JJ, Blank V. Regulation and function of the NFE2 transcription factor in hematopoietic and non-hematopoietic cells. *Cell Mol Life Sci.* 2015;72(12):2323-2335.
  46. Wang X, Yang L, Wang YC, et al. Comparative analysis of cell lineage differentiation during hepatogenesis in humans and mice at the single-cell transcriptome level. *Cell Res.* 2020;30(12):1109-1126.
  47. Gu Z, Churchman M, Roberts K, et al. Genomic analyses identify recurrent MEF2D fusions in acute lymphoblastic leukaemia. *Nat Commun.* 2016;7:13331.

48. Davis CA, Hitz BC, Sloan CA, et al. The encyclopedia of DNA elements (ENCODE): data portal update. *Nucleic Acids Res.* 2018; 46(D1):D794-D801.
49. Mandoli A, Singh AA, Prange KHM, et al. The hematopoietic transcription factors RUNX1 and ERG prevent AML1-ETO oncogene overexpression and onset of the apoptosis program in t(8;21) AMLs. *Cell Rep.* 2016;17(8):2087-2100.
50. McKeown MR, Corces MR, Eaton ML, et al. Superenhancer analysis defines novel epigenomic subtypes of Non-APL AML, including an RAR $\alpha$  dependency targetable by SY-1425, a potent and selective RAR $\alpha$  agonist. *Cancer Discov.* 2017;7(10):1136-1153.
51. Karwacz K, Miraldi ER, Pokrovskii M, et al. Critical role of IRF1 and BATF in forming chromatin landscape during type 1 regulatory cell differentiation. *Nat Immunol.* 2017;18(4):412-421.
52. Minderjahn J, Schmidt A, Fuchs A, et al. Mechanisms governing the pioneering and redistribution capabilities of the non-classical pioneer PU.1 [published correction appears in *Nat Commun.* 2020;11(1):1139]. *Nat Commun.* 2020;11(1):402.
53. Fuglerud BM, Lemma RB, Wanichawan P, Sundaram AYM, Eskeland R, Gabrielsen OS. A c-Myb mutant causes deregulated differentiation due to impaired histone binding and abrogated pioneer factor function. *Nucleic Acids Res.* 2017;45(13):7681-7696.
54. Neph S, Vierstra J, Stergachis AB, et al. An expansive human regulatory lexicon encoded in transcription factor footprints. *Nature.* 2012;489(7414):83-90.
55. Herglotz J, Unrau L, Hauschildt F, et al. Essential control of early B-cell development by Mef2 transcription factors. *Blood.* 2016; 127(5):572-581.
56. Huang FW, Hodis E, Xu MJ, Kryukov GV, Chin L, Garraway LA. Highly recurrent TERT promoter mutations in human melanoma. *Science.* 2013;339(6122):957-959.
57. Melton C, Reuter JA, Spacek DV, Snyder M. Recurrent somatic mutations in regulatory regions of human cancer genomes. *Nat Genet.* 2015;47(7):710-716.
58. Rheinbay E, Parasuraman P, Grimsby J, et al. Recurrent and functional regulatory mutations in breast cancer. *Nature.* 2017; 547(7661):55-60.
59. Cao Y, Li L, Xu M, et al; ChinaMAP Consortium. The ChinaMAP analytics of deep whole genome sequences in 10,588 individuals. *Cell Res.* 2020;30(9):717-731.
60. Knudson AG Jr. Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci USA.* 1971;68(4):820-823.
61. Berger AH, Knudson AG, Pandolfi PP. A continuum model for tumour suppression. *Nature.* 2011;476(7359):163-169.
62. Christopher MJ, Katerndahl CDS, LeBlanc HR, et al. Tumor suppressor function of WT1 in acute promyelocytic leukemia. *Haematologica.* 2022;107(1):342-346.
63. Nakano K, Uchimar K, Utsunomiya A, Yamaguchi K, Watanabe T. Dysregulation of c-Myb pathway by aberrant expression of proto-oncogene MYB provides the basis for malignancy in adult T-cell leukemia/lymphoma cells. *Clin Cancer Res.* 2016;22(23): 5915-5928.
64. Armenteros-Monterroso E, Zhao L, Gasparoli L, et al. The AAA+ATPase RUVBL2 is essential for the oncogenic function of c-MYB in acute myeloid leukemia. *Leukemia.* 2019;33(12):2817-2829.
65. Takao S, Forbes L, Uni M, et al. Convergent organization of aberrant MYB complex controls oncogenic gene expression in acute myeloid leukemia. *eLife.* 2021;10:e65905.
66. Mansour MR, Abraham BJ, Anders L, et al. Oncogene regulation. An oncogenic super-enhancer formed through somatic mutation of a noncoding intergenic element. *Science.* 2014;346(6215):1373-1377.
67. Rahman S, Magnussen M, León TE, et al. Activation of the LMO2 oncogene through a somatically acquired neomorphic promoter in T-cell acute lymphoblastic leukemia. *Blood.* 2017;129(24):3221-3226.
68. Smeenk L, Ottema S, Mulet-Lazaro R, et al. Selective requirement of MYB for oncogenic hyperactivation of a translocated enhancer in leukemia. *Cancer Discov.* 2021;11(11): 2868-2883.
69. Zhang X, Xing G, Fraizer GC, Saunders GF. Transactivation of an intronic hematopoietic-specific enhancer of the human Wilms' tumor 1 gene by GATA-1 and c-Myb. *J Biol Chem.* 1997;272(46):29272-29280.
70. Liu H, Leslie EJ, Carlson JC, et al. Identification of common non-coding variants at 1p22 that are functional for non-syndromic orofacial clefting. *Nat Commun.* 2017;8(1):14759.
71. Kikuchi M, Hara N, Hasegawa M, et al. Enhancer variants associated with Alzheimer's disease affect gene expression via chromatin looping. *BMC Med Genomics.* 2019;12(1):128.
72. Meng XH, Xiao HM, Deng HW. Combining artificial intelligence: deep learning with Hi-C data to predict the functional effects of non-coding variants. *Bioinformatics.* 2021; 37(10):1339-1344.

© 2022 by The American Society of Hematology. Licensed under Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0), permitting only noncommercial, nonderivative use with attribution. All other rights reserved.