




Integrated genetic and epigenetic analysis revealed heterogeneity of acute lymphoblastic leukemia in Down syndrome

Yasuo Kubota¹  | Kumiko Uryu¹ | Tatsuya Ito² | Masafumi Seki¹ | Tomoko Kawai³ | Tomoya Isobe¹ | Tadayuki Kumagai⁴ | Tsutomu Toki² | Kenichi Yoshida⁵ | Hiromichi Suzuki⁵ | Keisuke Kataoka⁵  | Yuichi Shiraishi⁶ | Kenichi Chiba⁶ | Hiroko Tanaka⁷ | Kentaro Ohki⁸ | Nobutaka Kiyokawa⁸ | Jiro Kagawa⁴ | Satoru Miyano⁷ | Akira Oka¹ | Yasuhide Hayashi⁹ | Seishi Ogawa⁵ | Kiminori Terui² | Atsushi Sato¹⁰ | Kenichiro Hata³ | Etsuro Ito² | Junko Takita¹ 

¹Department of Pediatrics, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

²Department of Pediatrics, Hirosaki University Graduate School of Medicine, Hirosaki, Japan

³Department of Maternal-Fetal Biology, National Research Institute for Child Health and Development, Tokyo, Japan

⁴Department of Pediatrics, Fujieda Municipal General Hospital, Fujieda, Japan

⁵Department of Pathology and Tumor Biology, Graduate School of Medicine, Kyoto University, Kyoto, Japan

⁶Section of Genome Analysis Platform, Center for Cancer Genomic and Advanced Therapeutics, National Cancer Center, Tokyo, Japan

⁷Laboratory of DNA Information Analysis, Human Genome Center, Institute of Medical Science, University of Tokyo, Tokyo, Japan

⁸Department of Pediatric Hematology and Oncology Research, National Research Institute for Child Health and Development, Setagaya-ku, Japan

⁹Institute of Physiology and Medicine, Jobu University, Takasaki, Japan

¹⁰Department of Hematology and Oncology, Miyagi Children's Hospital, Sendai, Japan

Correspondence

Junko Takita, Department of Pediatrics, The University of Tokyo, Tokyo, Japan.
Email: jtakita-ty@umin.ac.jp

Abstract

Children with Down syndrome (DS) are at a 20-fold increased risk for acute lymphoblastic leukemia (ALL). Compared to children with ALL and no DS (non-DS-ALL), those with DS and ALL (DS-ALL) harbor uncommon genetic alterations, suggesting DS-ALL could have distinct biological features. Recent studies have implicated several genes on chromosome 21 in DS-ALL, but the precise mechanisms predisposing children with DS to ALL remain unknown. Our integrated genetic/epigenetic analysis revealed that DS-ALL was highly heterogeneous with many subtypes. Although each subtype had genetic/epigenetic profiles similar to those found in non-DS-ALL, the subtype distribution differed significantly between groups. The Philadelphia chromosome-like subtype, a high-risk B-cell lineage variant relatively rare among the entire

Abbreviations: ALL, acute lymphoblastic leukemia; AML, acute myeloid leukemia; BCP, B-cell precursor; DS, Down syndrome; HeH, high hyperdiploid; JACLS, Japan Association of Childhood Leukemia Study; MLPA, multiplex ligation-dependent probe amplification; Ph, Philadelphia chromosome; ROSE, recognition of outliers by sampling ends; SNP, single nucleotide polymorphism.

Yasuo Kubota, Kumiko Uryu, and Tatsuya Ito contributed equally to this work.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2019 The Authors. *Cancer Science* published by John Wiley & Sons Australia, Ltd on behalf of Japanese Cancer Association.

pediatric ALL population, was the most common form in DS-ALL. Hypermethylation of *RUNX1* on chromosome 21 was also found in DS-ALL, but not non-DS-ALL. *RUNX1* is essential for differentiation of blood cells, especially B cells; thus, hypermethylation of the *RUNX1* promoter in B-cell precursors might be associated with increased incidence of B-cell precursor ALL in DS patients.

KEYWORDS

acute lymphoblastic leukemia, children, Down syndrome, epigenetic analysis, genetic analysis

1 | INTRODUCTION

Children with DS harbor one extra copy of chromosome 21. Among other congenital deficits and disease susceptibilities, children with DS are at a 20-fold increased risk for ALL.¹ Compared to ALL in children without DS (non-DS-ALL), children with DS and ALL (DS-ALL) show uncommon genetic alterations,² such as mutations in *JAK2*, *NRAS*, and *KRAS* as well as *CRLF2* overexpression, suggesting that certain tumorigenic or other disease-related processes could be unique to DS-ALL. Consistent with distinct ALL pathomechanisms in DS, DS-ALL patients have generally worse prognosis than non-DS-ALL patients. Therefore, it is critical to identify factors contributing to enhance ALL propensity and severity in DS. Genes on chromosome 21 are obvious candidates. Indeed, recent studies have implicated *HMGN1*³ and *DYRK1A*⁴ overexpression in DS-ALL.

To better understand DS-ALL molecular pathogenesis, we carried out an integrated genetic/epigenetic analysis. Expression and methylation analyses revealed high subtype heterogeneity in DS-ALL. Although each subtype had expression and methylation profiles similar to corresponding non-DS-ALL subtypes, subtype frequency distribution differed significantly between patient groups. The hypermethylation of *RUNX1* on chromosome 21 was unique to DS-ALL, which could explain the increased incidence of BCP-type ALL in DS.

2 | MATERIALS AND METHODS

2.1 | Patients and materials

Forty-three patients with DS-ALL (Table S1), 38 patients with non-DS-ALL (Table S2), and 18 DS patients without ALL or AML (Table S3) were enrolled in this retrospective study. The 43 DS-ALL patients included one large cohort (n = 31) from the JACLS as well as patients from various hospitals in Japan. The 38 non-DS-ALL patients were from the University of Tokyo Hospital. Peripheral blood and bone marrow samples were collected after written informed consent was obtained from legal guardians according to protocols approved by the Human Genome, Gene Analysis Research Ethics Committee of the University of Tokyo, and other participating institutions. All protocols conformed to the tenets of the Declaration of Helsinki.

2.2 | Next-generation sequencing

Next-generation sequencing was carried out using the Illumina HiSeq 2000, 2500, or MiSeq platform (Illumina) with a standard 100-bp paired-end read protocol according to the manufacturer's instructions.⁵

2.3 | RNA sequencing

High-quality RNA samples isolated by Agilent TapeStation (Agilent) were available from 25 DS-ALL patients (RNA integrity number equivalent greater than 5.5). These samples were used to prepare libraries for RNA sequencing using the NEBNext Ultra RNA library prep kit for the Illumina platform (New England BioLabs). Fusion transcripts were detected by Genomon version 2.5.0 and filtered by excluding fusions mapping to repetitive regions. Normalized count data obtained by the variance-stabilizing transformation function of the R package DESeq2 were used for clustering analysis. Prior to clustering analysis, the data were filtered to remove gene sets deemed unrelated to disease, such as those from sex-determining regions of X and Y chromosomes and genes (Table S4) from contaminating normal erythrocytes. For clustering analysis, we used Ward's hierarchical clustering method and included 250 of the top 1% of differentially expressed genes (Table S5) extracted using DESeq2. For comparing DS-ALL and non-DS-ALL samples, we also used the open dataset of expression⁶ (Table S6).

2.4 | Validation of fusion genes detected by RNA sequencing

Novel in-frame fusion transcripts were validated by RT-PCR followed by Sanger sequencing (Table S7).

2.5 | Detection of the Ph-like signature

We defined a ROSE gene set (Table S8), including the top 25 genes in clusters R1-R8.⁷ Ward's hierarchical clustering method was then used for clustering analysis of the ROSE gene set. As reported,⁷ samples in cluster R8 showed a signature similar to Philadelphia chromosome-positive ALL and were thus labeled Ph-like. To detect other samples with a Ph-like signature, we undertook hierarchical

clustering of 25 DS-ALL (Table S1) and 118 non-DS-ALL samples (Tables S2 and S6) based on genes (Table S9) significantly upregulated (adjusted $P < .0001$) in samples previously labeled Ph-like in R8.

2.6 | DNA methylation analysis

Comprehensive DNA methylation analysis was carried out on 35 DS-ALL and 24 non-DS-ALL samples using the Infinium HumanMethylationEPIC BeadChip (Illumina) for the JACLS samples or Infinium HumanMethylation 450K BeadChip (Illumina) for samples from other sources, all according to the manufacturer's protocols. Beta values were corrected for probe design bias using a beta-mixture quantile normalization method⁸ and converted to M values.⁹ Next, the R package *pcaMethods* bioconductor was used to impute incomplete M values, which were then converted to beta values. Imputed beta values were later used for further analyses. For the analysis of DNA methylation, the open dataset of DNA methylation profiling in pediatric ALL determined by the Infinium HumanMethylation 450K BeadChip¹⁰ was also included. For clustering analysis, we used unsupervised consensus clustering with 8000 probes using Ward's method. Cluster stability was determined by consensus clustering with 1000 iterations using the R package *ConsensusClusterPlus*. Statistical significance was assessed using Wilcoxon's rank-sum test, and values were corrected by employing the Benjamini-Hochberg method.

2.7 | Bisulfite conversion of *RUNX1*

To confirm *RUNX1* methylation in DS patients without ALL or AML, bisulfite sequencing was carried out using nested primers (forward, 5'-tcttgaaagaagaacagacca-3'; reverse, 5'-agtaaattctagcattactcagggga-3'). Genomic DNA (500 ng) was bisulfite-modified using the EpiTect Plus DNA bisulfite kit (Qiagen) according to the manufacturer's instructions.

2.8 | Targeted capture sequencing

We undertook mutation analysis of genes (*JAK2*, *SH2B3*, *KRAS*, *NRAS*, *PTPN11*, *FLT3*, *BRAF*, *NF1*, *KIT*, *IKZF1*, *PAX5*, and *IKZF1*) reported as mutually exclusive oncogenic drivers² or transcription factors in DS-ALL ($n = 43$). The samples from the JACLS ($n = 31$) were captured and sequenced using SureSelect (Agilent), HaloPlex (Agilent), Nextera (Illumina), or SeqCap (Roche NimbleGen). Polymerase chain reaction-based targeted deep sequencing was undertaken on other samples ($n = 12$).

2.9 | Copy number analysis

DNA samples extracted from peripheral blood or bone marrow were processed using MLPA for the JACLS samples and SNP array analysis for the other samples. The MLPA was carried out using the SALSA MLPA probemix P335-B1 ALL-*IKZF1* kit according to the manufacturer's instructions (MRCHolland). The SNP array

analysis was undertaken using the Affymetrix GeneChip 250K Nsp or CytoScan HD (Affymetrix), according to the manufacturer's protocol.

2.10 | Mutations in non-DS-ALL samples

Among the 118 samples used for RNA sequencing, DNA was available for 42 paired normal samples, which were evaluated using whole-exome sequencing. The mutations in the remaining 76 samples, for which paired normal DNA was not available, were called by the RNA sequencing results.

3 | RESULTS

3.1 | Expression analysis

Initial next-generation sequencing analysis of 25 DS-ALL samples (Table S1) identified 19 fusions (Table S10), including 15 frequently reported in pediatric BCP-ALL, as well as 4 novel fusions. Among common fusions previously identified in ALL, *P2RY8-CRLF2* was detected in 9 samples; *IGH-CRLF2* was not detected in our cohort. Among novel fusions, *PDGFA-TTYH3* was detected in 2 samples. Two DS-ALL samples with $t(14;19)(q32;q13)$ ¹¹ and $t(8;14)(q11;q32)$ ¹² showed high *CEBPA* and *CEBPD* expression levels, respectively (Figure S1). These samples were assumed to harbor *IGH-CEBPA* and *IGH-CEBPD* fusions, respectively, because both karyotypes involved *IGH* at 14q32. These fusions are rare in non-DS-ALL, but could be more common in DS-ALL.^{12,13} To characterize the unique expression profiles of DS-ALL, we applied hierarchical clustering analysis (Figure 1 and Table S5) including the 25 DS-ALL and 118 non-DS-ALL samples (expression cohort). These non-DS-ALL samples included several subtypes, such as the *ETV6-RUNX1* fusion and HeH (Tables S2 and S6). We also identified 8 samples with *PAX5* abnormalities in the absence of other genetic alterations, which we defined as "PAX5-altered." Unsupervised hierarchical clustering of the entire ALL sample population yielded 6 clusters (E1-E6), and DS-ALL samples fell into 4 clusters, with significant enrichment in E6 (Table S11). Cluster E3 was heterogeneous and contained several ALL subtypes, including *MLL*-rearranged, *TCF3-HLF*, *IGH-DUX4*, *MEF2D*-rearranged, and *PAX5*-altered. The *PAX5*-altered samples fell into clusters E3 and E6. All DS-ALL samples with *ETV6-RUNX1* fell into cluster E4, which also included all non-DS-ALL samples with the *ETV6-RUNX1* fusion. Additionally, E4 included one DS-ALL sample without *ETV6-RUNX1*. Six DS-ALL fell into cluster E5, which also included most non-DS-ALL samples with HeH. Among these 6 DS-ALL cases, only 1 had HeH. Cluster E6 was characterized by the presence of *BCR-ABL1* fusions, Ph-like expression profiles, and *CRLF2* fusions.

To detect the Ph-like signature, we clustered the expression cohort by gene sets using the ROSE method⁷ (Figure S2 and Table S8), which revealed that 7 DS-ALL samples had the Ph-like signature. Altogether, expression analysis revealed that DS-ALL samples were highly heterogeneous, but individual subtypes showed

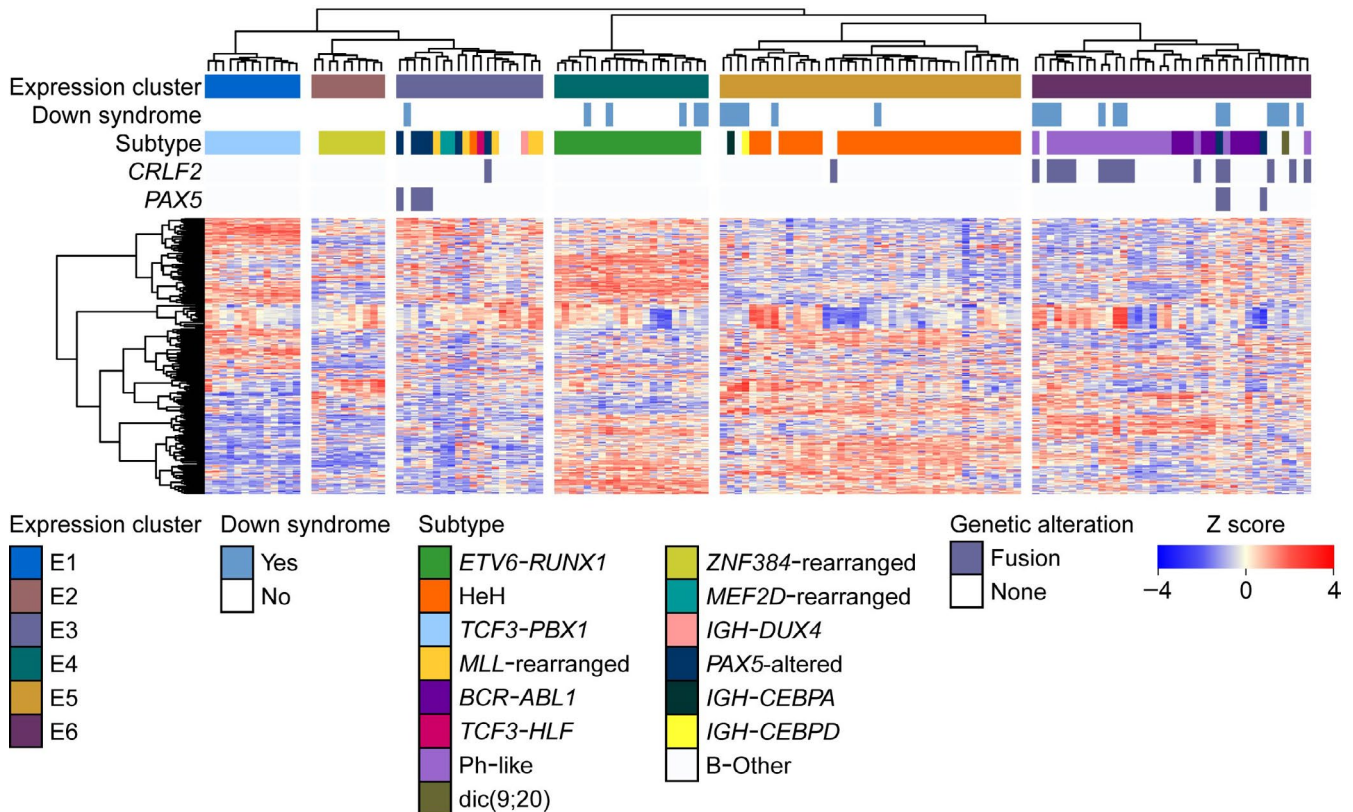


FIGURE 1 Gene expression clusters in 143 samples of B-cell precursor acute lymphoblastic leukemia (BCP-ALL) (25 samples from children with Down syndrome and ALL [DS-ALL] and 118 non-DS-ALL samples) on hierarchical clustering. Hierarchical clustering reveals BCP-ALL samples are grouped into 6 clusters (E1-E6). In those 6 clusters, DS-ALL samples are clustered into 4 clusters. Expression clusters, DS-ALL samples, subtypes, and genetic aberrations are shown by colors as indicated. DS-ALL and non-DS-ALL are clustered into the same cluster corresponding to each biological subtype. Philadelphia chromosome (Ph)-like signature is confirmed by recognition of outliers by sampling ends (ROSE) gene set clustering. HeH, high hyperdiploid

expression patterns similar to corresponding non-DS-ALL subtypes. Alternatively, the frequency distribution differed between groups.

Although it is reasonable to speculate that genes on chromosome 21 could be associated with DS-ALL development, surprisingly no genes on chromosome 21, such as *HMGN1* or *DYRK1A*, showed significantly higher expression in DS-ALL (Table S12).

3.2 | DNA methylation analysis

We then compared DNA methylation status between DS-ALL and non-DS-ALL samples. This methylation cohort included 59 samples of our study cohort (methylation cohort, 35 DS-ALL and 24 non-DS-ALL samples) and 664 samples (14 DS-ALL samples and 650 non-DS-ALL samples) from the open dataset of pediatric ALL.¹⁰ In methylation-based clustering analysis, DS-ALL samples fell into 5 clusters, M1-M5 (Figures 2 and S3, Table S13). Clusters M1 and M2 contained mainly HeH, whereas clusters M3 and M5 were highly heterogeneous. Two-step clustering further divided M3 into 5 (Figure S4) and M5 into 6 (Figure S5) additional clusters. Conversely, cluster M4 included the majority of samples with *ETV6-RUNX1* as well as the two DS-ALL samples without *ETV6-RUNX1*. In previous

DNA methylation profiling,¹⁰ the *ETV6-RUNX1* cluster also included *ETV6-RUNX1*-like samples, suggesting that the DS-ALL samples without *ETV6-RUNX1* in cluster M4 might have an *ETV6-RUNX1*-like signature.

In accordance with expression analysis, DS-ALL samples were clustered in subtypes similar to those of the non-DS-ALL samples by methylation analysis. However, the methylation levels of several genes, including *RUNX1* and *KDM2B*, were higher in the DS-ALL samples than non-DS siblings or mothers.¹⁴ Although methylation clustering revealed no distinct DS-ALL subtypes, the direct comparison of methylation status revealed that the P1 promoter region¹⁵ of *RUNX1* was hypermethylated in DS-ALL, but not in non-DS-ALL (Figure 3A and Table S14). Methylation of *RUNX1* promoters was higher in all DS-ALL subtypes compared to corresponding non-DS-ALL subtypes except for *ETV6-RUNX1*, although the sample size of several subtypes, such as Ph-like or *PAX5*-altered, were small in non-DS-ALL (Figure S6). Next, we used the most significant probe (cg22698744) to compare ALL samples with CD19⁺ B cells and CD19⁻ B cells from fetal bone marrow samples,¹⁶ adult bone marrow samples,¹⁷ and bone marrow samples of DS-ALL patients in remission. The methylation of CD19⁺ B cells from bone marrow samples of DS-ALL patients in remission was higher than that of CD19⁺

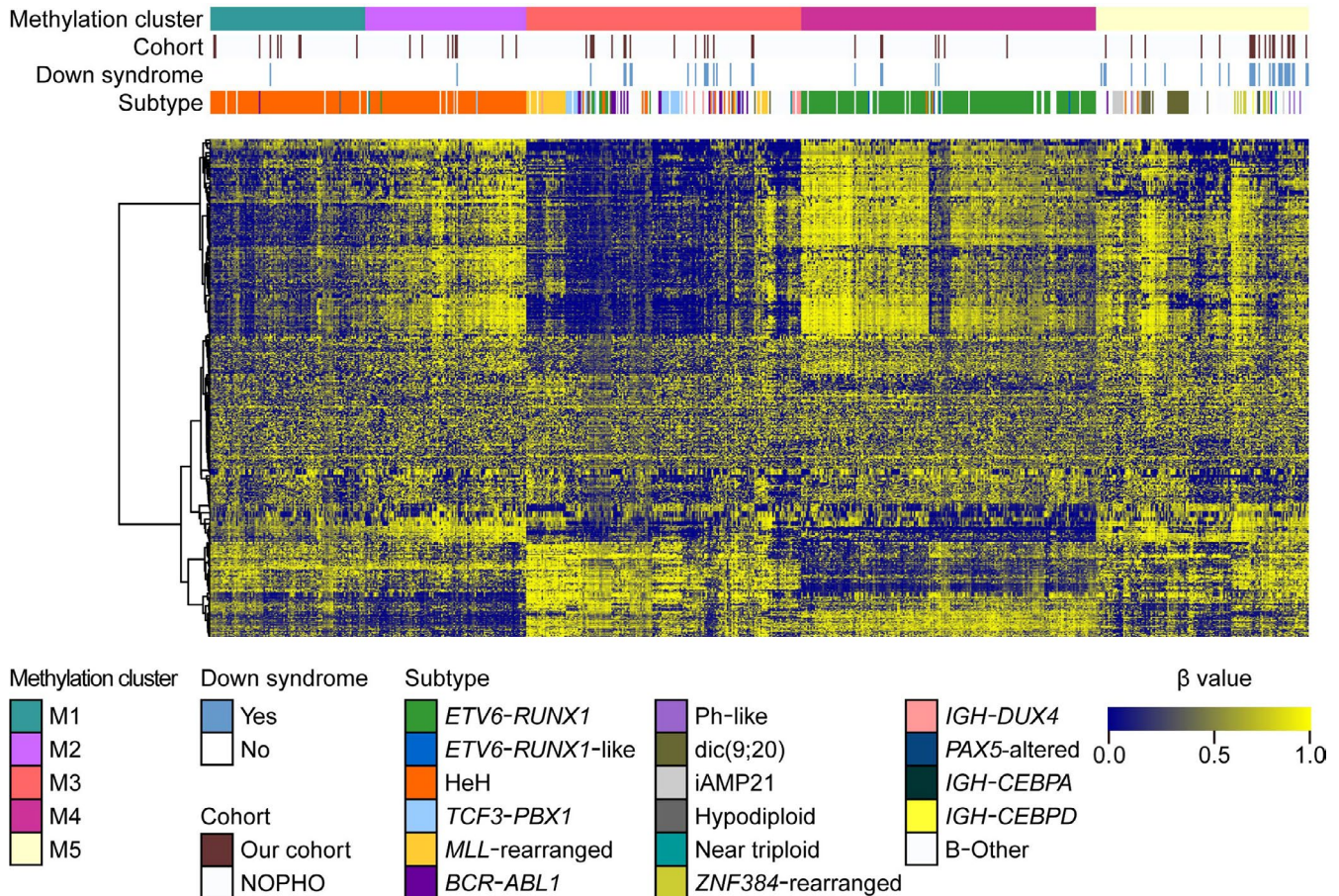


FIGURE 2 DNA methylation clusters in 723 B-cell precursor acute lymphoblastic leukemia (BCP-ALL) samples on consensus unsupervised clustering. Consensus unsupervised clustering suggests 723 BCP-ALL samples are grouped into 5 clusters (M1-M5). Clusters, cohorts, Down syndrome (DS)-ALL samples, and subtypes are shown by colors as indicated. Methylation clusters M3 and M5 are heterogeneous clusters, which are reclustered into 5 and 6 clusters, respectively. As shown in the expression analysis, DS-ALL and non-DS-ALL are clustered into the same cluster corresponding to each biological subtype. HeH, high hyperdiploid; NOPHO, Nordic Society for Pediatric Hematology and Oncology; Ph, Philadelphia chromosome

B cells from fetal and adult ALL bone marrow samples, although the sample size was small, suggesting that *RUNX1* promoter hypermethylation might be a unique characteristic of DS-ALL (Figure 3B). However, *RUNX1* promoters were also hypermethylated in DS patients without ALL (Figure S7). *RUNX1* promoter methylation was also higher in adults than fetuses.

Thus, consistent with a previous report,¹⁴ the *RUNX1* promoter appears congenitally hypermethylated in DS. The *RUNX1* isoform transcribed from the P1 promoter is expressed predominantly in hematopoietic cells.¹⁸ As *RUNX1* is essential for the differentiation of B cells, hypermethylation of the *RUNX1* promoter could be associated with increased BCP-ALL incidence in DS.

3.3 | Mutation and copy number analysis

To investigate the relationships among expression, methylation, and genetic status, we undertook mutational (Table S15) and copy number analyses (Tables S16 and S17) on 43 DS-ALL samples previously analyzed for expression (Figure 4A) or DNA methylation (Figure 4B).

In addition, we included the mutational analysis results (Tables S18 and S19) of 118 non-DS-ALL samples (expression cohort). Six of 25 samples with *CRLF2* fusions harbored *JAK2* mutations. As all DS-ALL samples with *JAK2* mutations and *CRLF2* fusions in the expression analysis had Ph-like signatures similar to non-DS-ALL,^{19,20} all 6 DS-ALL samples in the mutation cohort were labeled as Ph-like. In cluster E5, one non-DS-ALL sample with a *JAK2* mutation and *CRLF2* fusion was not labeled as Ph-like.

To detect other samples with Ph-like signatures, we carried out a hierarchical clustering of expression cohort samples (Tables S1, S2, and S6), based on the significantly highly expressed (adjusted $P < .0001$) genes (Table S9) in cluster R8 (Figure S8). This analysis revealed 3 additional samples, 2 DS-ALL and 1 non-DS-ALL, with similar gene expression profiles. The 3 B-other samples, 2 DS-ALL and 1 non-DS-ALL, in this cluster had Ph-like signatures. In contrast, several subtypes in the non-DS-ALL group showed mutations in *RAS* pathway genes,²¹ which are common drivers in pediatric BCP-ALL.

Copy number analysis revealed that one DS-ALL sample in cluster E3 had a known focal amplification of chromosome 9 involving

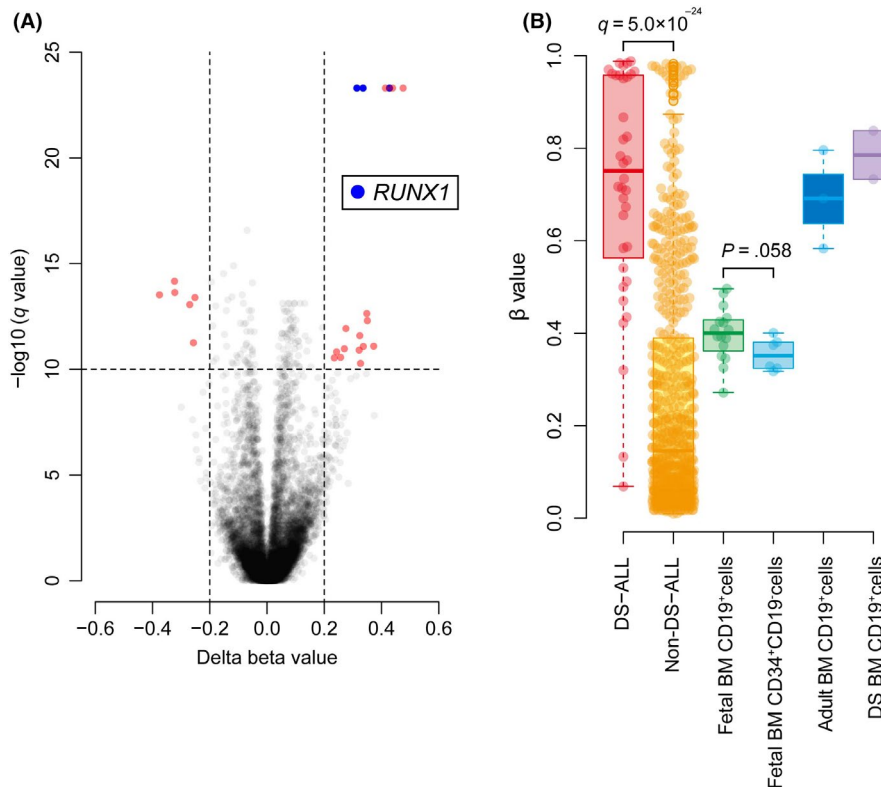


FIGURE 3 Methylation of *RUNX1* promoter. A, Volcano plot comparing significant delta beta values between samples from children with Down syndrome and acute lymphoblastic leukemia (DS-ALL) and non-DS-ALL samples. Significant probes showing the delta beta value greater than .2 or less than -.2, and $-\log_{10}(q \text{ value})$ greater than 10 are colored red. Among these probes, probes of promoter regions of *RUNX1* are colored blue and significantly highly methylated in DS-ALL. q Values are calculated using the Wilcoxon rank-sum test adjusted by the Benjamini-Hochberg correction. B, Box plot comparing beta values of the most significant probe of *RUNX1* promoter (cg22698744) in ALL samples ($n = 723$), $CD19^+$ cells of fetuses ($n = 16$), $CD34^+ CD19^-$ cells of fetuses ($n = 6$) and $CD19^+$ cells of adults ($n = 3$), and $CD19^+$ cells of DS-ALL samples in remission ($n = 2$). In normal samples, $CD19^+$ cells of adults and DS-ALL samples in remission are highly methylated compared to fetal cells. q Values are calculated using the Wilcoxon rank-sum test adjusted by the Benjamini-Hochberg correction. BM, bone marrow

PAX5 exons 2-5²² (Figure S9). The expression levels of *PAX5* exons 2-5 were higher than other exons (Figure S10), suggesting the expression of structurally aberrant *PAX5* protein or a loss of function. In addition, the expression profile of this particular sample (DS-15) was similar to other *PAX5*-altered samples from the expression analysis and so was categorized as *PAX5*-altered.

With the exception of *PAX5* deletion, *PAX5* status has not been evaluated previously in DS-ALL. Our results raised the possibility that miscellaneous aberrations, such as amplifications or fusions, as well as deletions, might occur in DS-ALL. Furthermore, several of the *PAX5*-altered DS-ALL samples, including those with *PAX5* amplification, were clustered in E3. Although various *PAX5* alterations, including fusions, amplifications, or mutations, were detected in BCP-ALL, the expression profiles of these samples were similar regardless of *PAX5* alterations without *PAX5* deletions.²³ In our cohort, our defined *PAX5*-altered samples were also clustered into the single cluster, E3, although our sample size was small and our analysis might not be precise. Because *PAX5* deleted samples were not necessarily clustered into E3 and were detected in 4 clusters, *PAX5* deletion would have different pathogenicity from other *PAX5* alterations. *PAX5* deletions decrease the expression

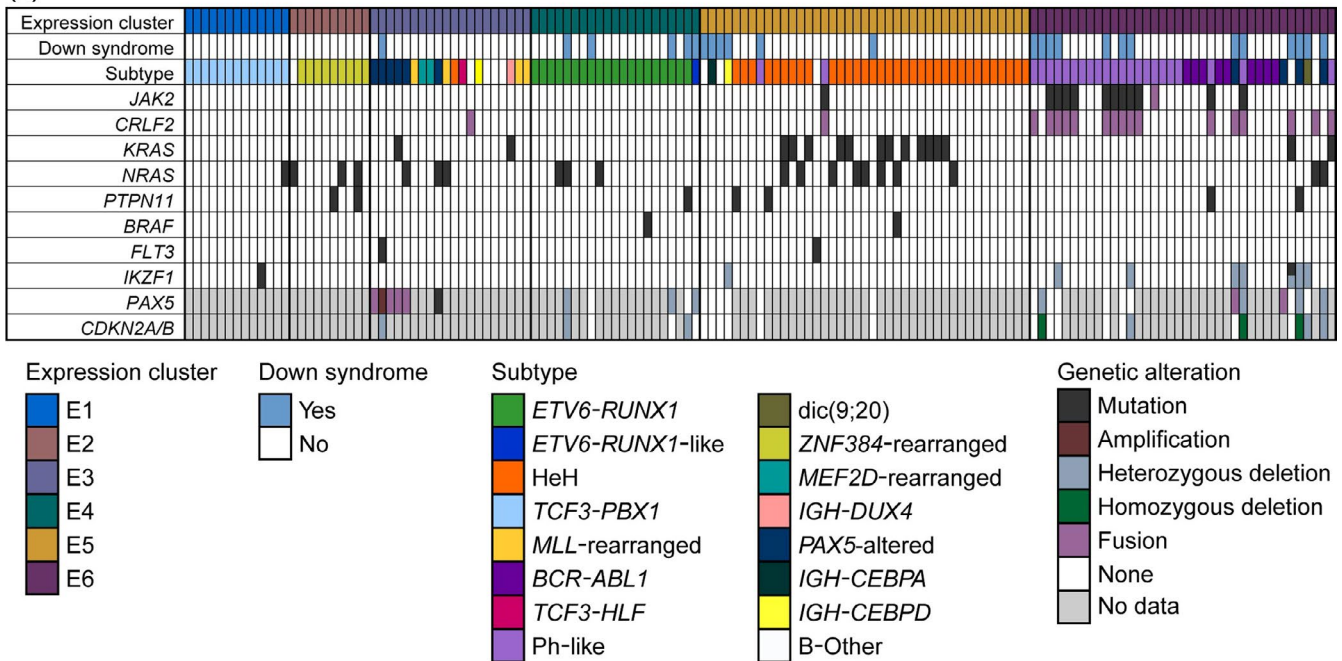
of *PAX5*, however, other *PAX5* alterations might express aberrant *PAX5* protein.

One DS-ALL sample (DS-13), without the *ETV6-RUNX1* fusion in cluster E4, had homozygous deletions of *ETV6* (Figure S11), suggesting an *ETV6-RUNX1*-like signature.²⁴ Cluster M4 included 2 DS-ALL samples (DS-10 and DS-13) without the *ETV6-RUNX1* fusion. Copy number analysis revealed one more sample with an *ETV6-RUNX1*-like signature (Figure S12). Although DS-10 was not available due to poor RNA quality, copy number analysis showed that DS-10 harbored deletions of several genes, including *ETV6* and *ARPP21*,²⁵ suggesting that DS-10 was also the *ETV6-RUNX1*-like signature.

4 | DISCUSSION

Although incidence of a Ph-like signature was high in DS-ALL, especially in children younger than 10 years old, the genetic profile of DS-ALL was highly heterogeneous (Figure 5). Although the expression and DNA methylation analyses revealed many similarities between DS-ALL and non-DS-ALL, the subtype profile of DS-ALL was distinct, with lower frequencies of *ETV6-RUNX1* and HeH subtypes²⁶ and a high incidence of a Ph-like signature, especially

(A)



(B)

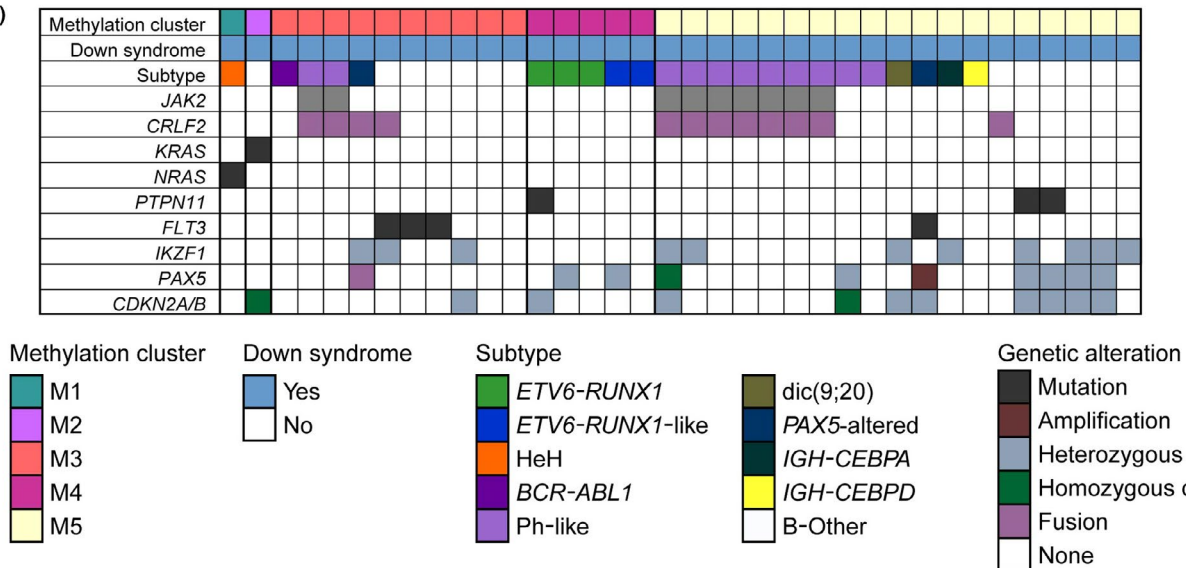
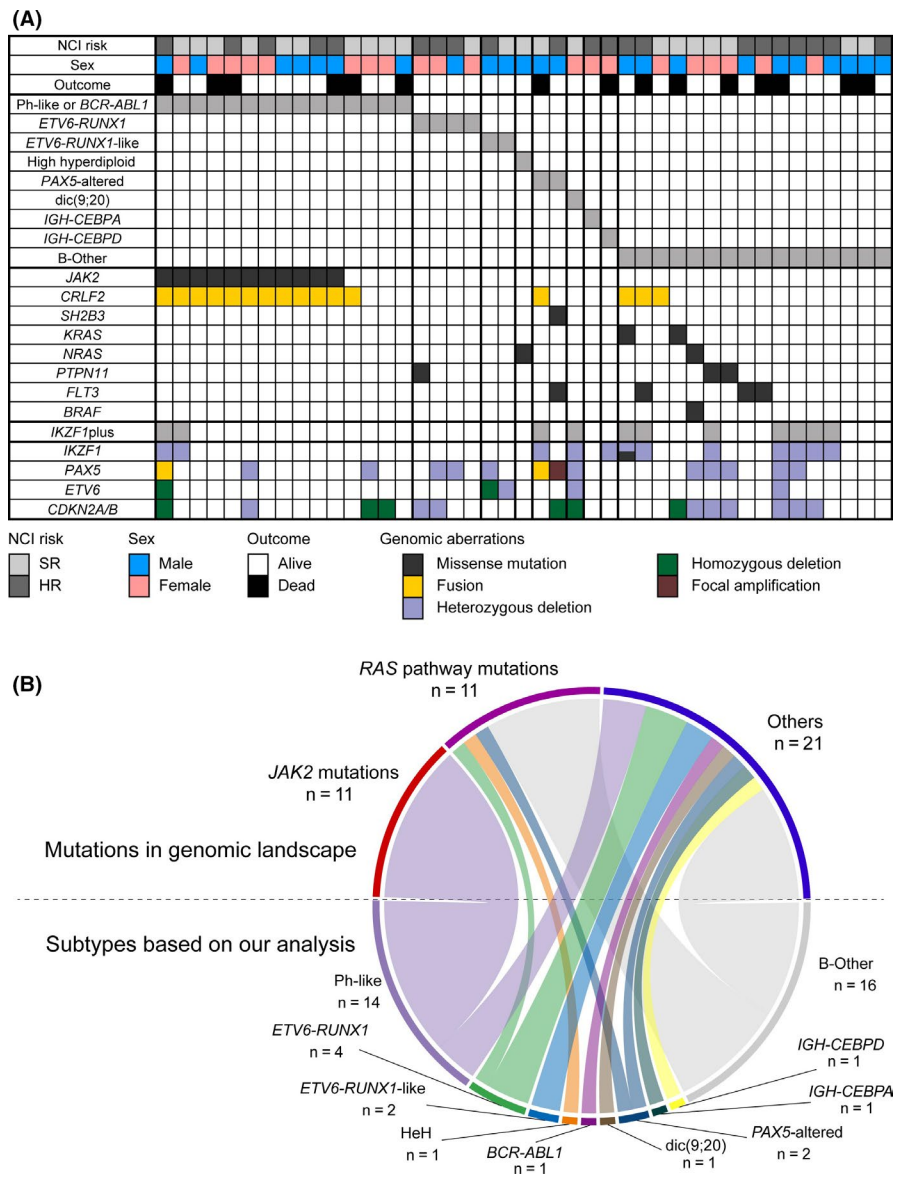


FIGURE 4 Relations of gene expression, DNA methylation, and genomic status. A, Mutational and copy number analysis in expression clusters (E1-E6). Clusters, samples from children with Down syndrome and acute lymphoblastic leukemia (DS-ALL), subtypes, and genetic aberrations are shown by different colors as indicated. Copy number analysis of *IKZF1* and *PAX5* are represented in only DS-ALL samples. DS-ALL sample in E3 had *PAX5* amplification. Samples with *PAX5* alteration are clustered into E3 or E6. DS-ALL sample without *ETV6-RUNX1* has deletion of *ETV6*, implicating *ETV6-RUNX1*-like signature. All samples with *JAK2* mutation and *CRLF2* fusion, including DS-ALL samples, reveal Philadelphia chromosome (Ph)-like signature. In contrast, mutations of *RAS* pathway genes are detected in several subtypes. B, Results of DNA methylation analysis are combined with results of mutational and copy number analyses. Copy number analyses of *IKZF1* and *PAX5* are represented in only DS-ALL samples. Samples with Ph-like signature are divided into M3 and M5 clusters. Ph-like samples in M3 and M5 are clustered in *BCR-ABL1* and *iAMP21* cluster, respectively. HeH, high hyperdiploid

in children younger than 10 years, compared to non-DS-ALL. There were also discrepancies in cluster designation based on differential expression and DNA methylation. Differences in gene expression could stem from mutations, the presence of chimeric

genes, and (or) DNA methylation. Previous studies reported that the frequencies of *MLL*-rearranged, *BCR-ABL1*, *ETV6-RUNX1*, and HeH subtypes were low in DS-ALL; however, considering that ALL risk is 20-fold higher in DS,¹ the incidence rates of these subtypes

FIGURE 5 Genetic landscape of Down syndrome and acute lymphoblastic leukemia (DS-ALL). A, Genetic landscape of DS-ALL combined with clinical information. The NCI criteria (standard risk [SR] or high risk [HR]), sex, and outcomes (alive or dead) together with subtypes, the affected genes, and the types of genomic aberrations are shown by colors as indicated. *IKZF1*plus was defined as *IKZF1* deletions co-occurring with deletions in *CDKN2A*, *CDKN2B*, *PAX5*, or *PAR1* in the absence of *ERG* deletion. We detected Ph-like, *ETV6*-*RUNX1*-like, *PAX5* alteration, *IGH*-*CEBPA* fusion, and *IGH*-*CEBPD* fusion in addition to known subtypes such as *ETV6*-*RUNX1*, high hyperdiploid, *dic(9;20)*, *TCF3*-*PBX1*, and *BCR*-*ABL1*. Samples are ordered by biological subtypes. B, The relation between mutational status and biological subtypes in DS-ALL samples. The upper half of this figure shows already known genetic alterations in DS-ALL; the lower half shows subtypes of ALL detected by our analysis. All samples with *JAK2* mutations and *CRLF2* fusions have the Philadelphia chromosome (Ph)-like signature. In contrast, several subtypes of DS-ALL have mutations of *RAS* pathway genes



were predicted to be 5- to 8-fold higher in DS compared to healthy children. Alternatively, the incidence of the Ph-like subtype was predicted to be 70-fold higher. Indeed, Fisher's exact test revealed a significantly higher Ph-like subtype frequency in DS-ALL ($P = .0056$) compared to non-DS-ALL. Generally, the Ph-like samples were divided into 3 classes: (i) ABL-class rearrangements involving *ABL1*, *ABL2*, *CSF1R*, and *PDGFRB*; (ii) JAK/STAT pathway alterations of *CRLF2*, *JAK2*, *EPOR*, *IL7R*, and *SH2B3*; and (iii) other rare kinase fusions involving *NTRK3*, *DGKH*, or *FGFR1*. The Ph-like signatures with JAK/STAT pathway alterations accounted for approximately 50% of all Ph-like signatures in the non-DS-ALL group. However, most DS-ALL samples with the Ph-like subtype (all except 2) in our cohort had *JAK2* mutations and *CRLF2* fusions, whereas none with Ph-like signatures had fusions of *JAK2*, *EPOR*, or *ABL*-class kinases. No studies to date have reported a Ph-like signature in DS-ALL cases harboring fusions without *CRLF2* fusions. The high incidence of a Ph-like signature in DS-ALL could be

due to the high frequency of *CRLF2* fusions in DS patients. Indeed, *CRLF2* fusion preceded *JAK2* mutation frequency in ALL with a Ph-like signature.

Our cohort also included DS-ALL samples with the *IKZF1* G158S mutation, the amplification of *PAX5* exons 2-5, and an *ETV6*-*RUNX1*-like signature, all of which were recently identified in other cohorts. In total, the B-other subtype accounted for approximately one-third of the cohort samples. Half of the B-other subtype samples in the DS-ALL cohort harbored mutations in *RAS* pathway genes. Mutations in *RAS* pathway genes are frequently detected in non-DS-ALL subtypes, especially the HeH subtype,²⁷ and thus might not be specific for DS-ALL. One other reason for the high proportion of the B-other subtype in DS-ALL could be the unique *IGH*-*CEBPA* or *IGH*-*CEBPD* subtypes revealed by our analysis.

Previous reports revealed DS patients have a genetic predisposition to ALL due to the presence of extra copies of genes such as *HMG1* and *DYRK1A*. In addition, as our analysis revealed,

hypermethylation of *RUNX1* promoter is also a potential predisposing factor. *RUNX1* is the differentiation factor from common lymphoid progenitors to pre-pro-B cells and hypermethylation of the *RUNX1* promoter is predicted to decrease *RUNX1* expression²⁸ in common lymphoid progenitors. Furthermore, chromosomal instabilities²⁹ caused by aneuploidy in DS patients might be associated with ALL development. These genetic and epigenetic factors could act synergistically to promote DS-ALL. Several genes, including *RUNX1*, were also shown to be hypermethylated in DS patients with myeloid leukemia or transient myeloproliferative disorder compared to DS noncancerous fetal liver mononuclear cells.³⁰ However, our comparison of DS-ALL and non-DS-ALL samples revealed similar expression levels of *RUNX1*, *HMGN1*, and *DYRK1A* as well as other chromosome 21 genes between the 2 groups, suggesting that the roles of these genes could be limited to the development of ALL and not afterwards. In addition, hypermethylation of the *RUNX1* promoter was not detected in several subtypes, including *ETV6-RUNX1*, some Ph-like, and some B-other samples, suggesting that these subtypes have unique molecular pathomechanisms independent of *RUNX1*.

In our cohort, DS-ALL prognosis was worse than non-DS-ALL prognosis (Figure S13A,B), possibly due to the greater sensitivity of DS-ALL patients to chemotherapy side-effects and higher treatment-related mortality rate. Additionally, unique high-risk subtypes and common high-risk subtypes with higher frequencies influence the general risk classification of DS-ALL. However, it is difficult to accurately estimate outcomes of these various subtypes due to the small individual sample sizes. Based on previous clinical investigations, the outcomes of *ETV6-RUNX1*, *ETV6-RUNX1*-like, and hyperdiploid may be excellent. Compared to Ph-like non-DS-ALL, the outcome of the Ph-like DS-ALL was fair (Figure S13C,D). Because outcomes of B-other DS-ALL cases are also still worse (Figure S13E,F), further identification of B-other DS-ALL subtypes and associated risk classifications are needed. In addition to personalized supportive care for DS-ALL patients, precise risk classification based on each subtype identified by molecular profiling is a promising treatment approach.

The major limitation of this study is the small DS-ALL sample size, which precluded direct comparisons of many subtypes between groups and limited whole-genome sequencing. Thus, our findings must be verified by a prospective study involving a large cohort.

In conclusion, we confirmed that DS-ALL is highly heterogeneous with many subtypes. Although subtype expression and methylation profiles were similar to non-DS-ALL, subtype frequencies differed significantly between groups. The Ph-like subtype was the most common in DS-ALL, accounting for approximately one-third of all samples. Hypermethylation of the *RUNX1* promoter was also a unique characteristic of DS-ALL, and so could explain the increased ALL incidence in DS.

DATA AVAILABILITY

All sequencing data supporting these findings, including RNA sequencing, DNA methylation arrays, and SNP arrays, were deposited in the DNA Data Bank of Japan and are accessible through

accession number JGAS0000000147. All other relevant data are available from the corresponding authors upon reasonable request.

ACKNOWLEDGMENTS

We gratefully acknowledge the TCGA Consortium and all its members for making their invaluable data publically available. The results published here are in whole or part based on data generated by the Therapeutically Applicable Research to Generate Effective Treatment (TARGET) initiative managed by the NCI. Information about TARGET can be found at <http://ocg.cancer.gov/programs/target>. We are also grateful to M. Matsumura, K. Yin, and F. Saito for their excellent technical assistance. We also wish to express our appreciation to H. Goto (Kanagawa Children's Medical Center), K. Koh (Saitama Children's Medical Center), R. Hanada (Saitama Children's Medical Center), T. Imamura (Kyoto Prefectural University of Medicine), K. Kato (Ibaraki Children's Hospital), and J. Kamiyama (Tottori University) for collecting samples.

DISCLOSURE

The authors declare no competing financial interests.

ORCID

Yasuo Kubota  <https://orcid.org/0000-0002-7909-5422>

Keisuke Kataoka  <https://orcid.org/0000-0002-8263-9902>

Junko Takita  <https://orcid.org/0000-0002-2452-6520>

REFERENCES

- Hasle H, Clemmensen IH, Mikkelsen M. Risks of leukaemia and solid tumours in individuals with Down's syndrome. *Lancet*. 2000;355:165-169.
- Nikolaev SI, Garieri M, Santoni F, et al. Frequent cases of RAS-mutated Down syndrome acute lymphoblastic leukaemia lack JAK2 mutations. *Nat Commun*. 2014;5:4654.
- Lane AA, Chapuy B, Lin CY, et al. Triplication of a 21q22 region contributes to B cell transformation through HMGN1 overexpression and loss of histone H3 Lys27 trimethylation. *Nat Genet*. 2014;46:618-623.
- Thompson BJ, Bhansali R, Diebold L, et al. *DYRK1A* controls the transition from proliferation to quiescence during lymphoid development by destabilizing Cyclin D3. *J Exp Med*. 2015;212:953-970.
- Sato Y, Yoshizato T, Shiraishi Y, et al. Integrated molecular analysis of clear-cell renal cell carcinoma. *Nat Genet*. 2013;45:860-867.
- Hirabayashi S, Ohki K, Nakabayashi K, et al. ZNF384-related fusion genes define a subgroup of childhood B-cell precursor acute lymphoblastic leukemia with a characteristic immunotype. *Haematologica*. 2017;102:118-129.
- Harvey RC, Mullighan CG, Wang X, et al. Identification of novel cluster groups in pediatric high-risk B-precursor acute lymphoblastic leukemia with gene expression profiling: correlation with genome-wide DNA copy number alterations, clinical characteristics, and outcome. *Blood*. 2010;116:4874-4884.

8. Teschendorff AE, Marabita F, Lechner M, et al. A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. *Bioinformatics*. 2013;29:189-196.
9. Du P, Zhang X, Huang CC, et al. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics*. 2010;11:587.
10. Nordlund J, Bäcklin CL, Wahlberg P, et al. Genome-wide signatures of differential DNA methylation in pediatric acute lymphoblastic leukemia. *Genome Biol*. 2013;14:r105.
11. Chapiro E, Russell L, Radford-Weiss I, et al. Overexpression of CEBPA resulting from the translocation t(14;19)(q32;q13) of human precursor B acute lymphoblastic leukemia. *Blood*. 2006;108:3560-3563.
12. Messinger YH, Higgins RR, Devidas M, Hunger SP, Carroll AJ, Heerema NA. Pediatric acute lymphoblastic leukemia with a t(8;14)(q11.2;q32): B-cell disease with a high proportion of Down syndrome: a Children's Oncology Group study. *Cancer Genet*. 2012;205:453-458.
13. Pui CH, Raimondi SC, Borowitz MJ, et al. Immunophenotypes and karyotypes of leukemic cells in children with Down syndrome and acute lymphoblastic leukemia. *J Clin Oncol*. 1993;11:1361-1367.
14. Bacalini MG, Gentilini D, Boattini A, et al. Identification of a DNA methylation signature in blood cells from persons with Down syndrome. *Aging (Albany NY)*. 2015;7:82-96.
15. Miyoshi H, Ohira M, Shimizu K, et al. Alternative splicing and genomic structure of the AML1 gene involved in acute myeloid leukemia. *Nucleic Acids Res*. 1995;23:2762-2769.
16. Lee ST, Xiao Y, Muench MO, et al. A global DNA methylation and gene expression analysis of early human B-cell development reveals a demethylation signature and transcription factor network. *Nucleic Acids Res*. 2012;40:11339-11351.
17. Ley TJ, Miller C, Ding L, et al. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N Engl J Med*. 2013;368:2059-2074.
18. Martinez M, Hinojosa M, Trombly D, et al. Transcriptional auto-regulation of RUNX1 P1 promoter. *PLoS ONE*. 2016;11:e0149119.
19. Roberts KG, Li Y, Payne-Turner D, et al. Targetable kinase-activating lesions in Ph-like acute lymphoblastic leukemia. *N Engl J Med*. 2014;371:1005-1015.
20. Wells J, Jain N, Konopleva M. Philadelphia chromosome-like acute lymphoblastic leukemia: progress in a new cancer subtype. *Clin Adv Hematol Oncol*. 2017;15:554-561.
21. Case M, Matheson E, Minto L, et al. Mutation of genes affecting the RAS pathway is common in childhood acute lymphoblastic leukemia. *Cancer Res*. 2008;68:6803-6809.
22. Ofverholm I, Tran AN, Heyman M, et al. Impact of IKZF1 deletions and PAX5 amplifications in pediatric B-cell precursor ALL treated according to NOPHO protocols. *Leukemia*. 2013;27:1936-1939.
23. Gu Z, Churchman ML, Roberts KG, et al. PAX5-driven subtypes of B-progenitor acute lymphoblastic leukemia. *Nat Genet*. 2019;51:296-307.
24. Lilljebjörn H, Henningson R, Hyrenius-Wittsten A, et al. Identification of ETV6-RUNX1-like and DUX4-rearranged subtypes in paediatric B-cell precursor acute lymphoblastic leukaemia. *Nat Commun*. 2016;7:11790.
25. Zaliava M, Kotrova M, Bresolin S, et al. ETV6/RUNX1-like acute lymphoblastic leukemia: a novel B-cell precursor leukemia subtype associated with the CD27/CD44 immunophenotype. *Genes Chromosom Cancer*. 2017;56:608-616.
26. Buitenkamp TD, Izraeli S, Zimmermann M, et al. Acute lymphoblastic leukemia in children with Down syndrome: a retrospective analysis from the Ponte di Legno study group. *Blood*. 2014;123:70-77.
27. Paulsson K, Horvat A, Strömbeck B, et al. Mutations of FLT3, NRAS, KRAS, and PTPN11 are frequent and possibly mutually exclusive in high hyperdiploid childhood acute lymphoblastic leukemia. *Genes Chromosom Cancer*. 2008;47:26-33.
28. Webber BR, Iacovino M, Choi SH, Tolar J, Kyba M, Blazar BR. DNA methylation of Runx1 regulatory regions correlates with transition from primitive to definitive hematopoietic potential in vitro and in vivo. *Blood*. 2013;122:2978-2986.
29. Hadi E, Sharony R, Goldberg-Bittman L, Biron-Shental T, Fejgin M, Amiel A. Telomere aggregates in trisomy 21 amniocytes. *Cancer Genet Cytogenet*. 2009;195:23-26.
30. Malinge S, Chlon T, Doré LC, et al. Development of acute megakaryoblastic leukemia in Down syndrome is associated with sequential epigenetic changes. *Blood*. 2013;122:e33-e43.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Kubota Y, Uryu K, Ito T, et al. Integrated genetic and epigenetic analysis revealed heterogeneity of acute lymphoblastic leukemia in Down syndrome. *Cancer Sci*. 2019;110:3358-3367. <https://doi.org/10.1111/cas.14160>