Original Article

Comprehensive genetic analyses of childhood acute leukemia in Iraq using next-generation sequencing

Lika'a Fasih Y. Al-Kzayer[1], Raghad M. Saeed[2], Hasanein Habeeb Ghali[2,3], Miyuki Tanaka[1], Mazin F. Al-Jadiry[2,3], Safa A. Faraj[2,4], Salma A. Al-Hadad[2,3], Hussam M. Salih Al Abdullah[5], Athar A. Majeed[5], Ali Omer Qadir[6], Dana Ahmed Abdullah[7,8], Kani Dlawar Noori[7], Zheyan Mohammed Hama[7], Abdulrahman A. Muhsin[9], Adnan Anwer Al-Doski[10], Yasir S. Al-Agele[11], Abduladheem H. Malallah[11], Khalid S. Al-Badrani[12], Asmaa M. A. Khaleel[12], Minoru Kamata[13], Motoharu Hamada[14], Seiji Kojima[14], Yozo Nakazawa[1], Yusuke Okuno[15,16]

[1]Department of Pediatrics, Shinshu University School of Medicine, Matsumoto, Japan; [2]Department of Pediatric Oncology, Children Welfare Teaching Hospital, Baghdad Medical City, Baghdad, Iraq; [3]Department of Pediatrics, College of Medicine, Baghdad University, Baghdad Medical City, Baghdad, Iraq; [4]Department of Pediatrics, College of Medicine, Wasit University, Al Kut, Iraq; [5]Department of Pediatric Oncology, Basra Children's Specialty Hospital, Basra, Iraq; [6]Department of Pediatric Hematology and Oncology, Hiwa Cancer Hospital, Sulaymaniyah, Iraq; [7]Department of Hematopathology, Hiwa Cancer Hospital, Sulaymaniyah, Iraq; [8]Department of Pathology, College of Medicine, University of Sulaimani, Sulaymaniyah, Iraq; [9]Department of Medical Laboratories, College of Health Science, University of Duhok, Duhok, Iraq; [10]Department of Hematopathology, College of Medicine, University of Duhok, Duhok, Iraq; [11]Department of Pediatric Oncology, Ibn Al-Atheer Hospital for Children, Mosul, Iraq; [12]Department of Hematopathology, Ibn Al-Atheer Hospital for Children, Mosul, Iraq; [13]Japan Chernobyl Foundation (JCF) NPO, Matsumoto, Japan; [14]Department of Pediatrics, Nagoya University Graduate School of Medicine, Nagoya, Japan; [15]Department of Virology, Nagoya City University Graduate School of Medical Sciences, Nagoya, Japan; [16]Medical Genomics Centre, Nagoya University Hospital, Nagoya, Japan

Contributions: (I) Conception and design: LFY Al-Kzayer, Y Okuno, Y Nakazawa, S Kojima, M Tanaka, M Kamata, SA Al-Hadad; (II) Administrative support: LFY Al-Kzayer, Y Okuno; (III) Provision of study materials or patients: RM Saeed, HH Ghali, MF Al-Jadiry, SA Faraj, SA Al-Hadad, HMS Al Abdullah, AA Majeed, AO Qadir, DA Abdullah, KD Noori, ZM Hama, AA Muhsin, AA Al-Doski, YS Al-Agele, AH Malallah, KS Al-Badrani, AMA Khaleel; (IV) Collection and assembly of data: LFY Al-Kzayer, Y Okuno, M Hamada; (V) Data analysis and interpretation: Y Okuno, M Hamada, LFY Al-Kzayer; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Lika'a Fasih Y. Al-Kzayer, MD, PhD. Department of Pediatrics, Shinshu University School of Medicine, 3-1-1, Asahi, Matsumoto 390-8621, Japan. Email: alkzayerlikaa@yahoo.co.uk; Yusuke Okuno, MD, PhD. Department of Virology, Nagoya City University Graduate School of Medical Sciences, Nagoya, Japan. Email: yusukeo@med.nagoya-cu.ac.jp.

Background: Molecular analyses in hematological malignancies provide insights about genetic makeup. Probable etiological factors in leukemogenesis could also be disclosed. Since genetic analyses are still primitive in Iraq, a country of repeated wars, we conceived of performing next-generation sequencing (NGS), to disclose the genomic landscape of acute lymphoblastic leukemia (ALL) and acute myeloid leukemia (AML) among a cohort of Iraqi children.

Methods: Dried blood samples were collected from Iraqi children with ALL (n=55), or AML (n=11), and transferred to Japan where NGS was done. Whole-exome, whole-genome, and targeted gene sequencings were performed.

Results: Somatic point mutations and the copy number variations among Iraqi children with acute leukemia were comparable with those in other countries, and cytosine-to-thymine nucleotide alterations were dominant. Strikingly, *TCF3-PBX1* was the most recurrent fusion gene (22.4%) in B-cell precursor ALL (B-ALL), and acute promyelocytic leukemia (AML-M3) was subtyped in 5 AML cases. Additionally, a high frequency of *RAS* signaling pathway mutations was detected in children with B-ALL (38.8%), along with 3 AML cases that carried oncogenic *RAS*.

Conclusions: Apart from disclosing the high frequency of *TCF3-PBX1*, NGS confirmed our previous finding of recurrent *RAS* mutations in Iraqi childhood acute leukemia. Our results suggest that the biology of

Iraqi childhood acute leukemia is in part characteristic, where the war-aftermath environment or geography might play a role.

Keywords: Acute lymphoblastic leukemia (ALL); acute myeloid leukemia (AML); next-generation sequencing (NGS); Iraq; RAS

Submitted Oct 11, 2022. Accepted for publication Mar 15, 2023. Published online Apr 20, 2023.
doi: 10.21037/tp-22-512
View this article at: https://dx.doi.org/10.21037/tp-22-512

Introduction

Childhood acute leukemia has heterogeneous biological and multifactorial etiology mechanisms linked with genetic susceptibility factors and subsequently acquired somatic mutations (1-3). Differences in the incidences, risk factors, and survival of pediatric acute leukemia along with the different frequencies of molecular markers have been reported across various countries (4,5). Such differences could be attributed to the interaction between genomic drivers, which are associated with race and ethnicity, and environmental factors (2,3,6-8).

Over the last four decades of wars and their aftermath in Iraq, the health system underwent a serious regression. No genetic analysis is yet available for diagnosing pediatric acute lymphoblastic leukemia (ALL) and acute myeloid leukemia (AML) in Iraq. Limited diagnostic facilities have a negative impact on disease understanding, management, and consequently its outcome. Meanwhile, childhood leukemia rates doubled over 15 years (1993–2007) according to a study from Basra in Southern Iraq, and the trend was deemed significant when compared to neighboring countries like Kuwait and Oman, as well as the United States (9). Notably, Basra was the nearest spot that experienced repeated gulf wars and was exposed to repeated bombing and by-products of the petroleum fires.

Despite the improved 5-year survival rate of pediatric ALL outcome of more than 90% in developed countries, Iraqi pediatric oncologists struggle to achieve around 70% (10,11). The international collaboration from Japan was, therefore, established aiming to scale up the diagnosis of Iraqi children with acute leukemia by performing molecular analysis using the dried blood spot (DBS) samples; concomitantly, Italy had settled a telemedicine program in the main pediatric oncology center in Baghdad to support in protocol guidance for acute leukemia cases. Through our collaboration studies, the prevalence of *RAS* mutations was previously noted to be higher among Iraqi childhood ALL and AML than in other countries (12,13). Likewise, acute promyelocytic leukemia (APL) was unusually frequent among Iraqi children with AML in our study (14) and in a report by the Italian team (15).

Next-generation sequencing (NGS) technology with its characteristic high throughput and high sensitivity and specificity provides a good platform for acute leukemia diagnosis and research to improve the understanding of molecular alterations in patients with such diseases. Thus, it aids in refining their treatment plans accordingly. NGS when compared to conventional genetic sequencing, has several advantages such as comprehensive genomic coverage, higher capacity with sample multiplexing, and the ability to sequence hundreds to thousands of genes or gene regions simultaneously (16).

In this new collaborative study, NGS was utilized for the first time for illustrating the landscape of genetic mutations

Highlight box

Key findings

- Among Iraqi children with acute leukemia, we disclosed a high frequency of *TCF3-PBX1* in ALL, and a frequent AML-M3 subtype, along with recurrent RAS mutations in ALL/AML.

What is known, and what is new?

- Molecular analyses in hematological malignancies provide insights about genetic makeup. Probable etiological factors in leukemogenesis could be disclosed.
- For the first time in Iraq, NGS was performed to disclose the molecular landscape of a cohort of childhood ALL/AML from Iraq in Japan using dried blood spot samples. Our results suggest that the biology of Iraqi childhood acute leukemia is, in part, characteristic, where the war-aftermath environment or geography might play a role.

What is the implication, and what should change now?

- Understanding the biology of acute leukemia in Iraq could help doctors there in modifying the management protocols or arranging the plan for required and applicable analysis in their locations for achieving better results.

in a series of Iraqi children with ALL and AML, and the DBS-extracted DNA was used for the NGS analysis. We aimed to perform a more comprehensive genetic analysis using NGS and to assess the possible differences in the biology of pediatric acute leukemia in Iraq in association with genetic or non-genetic factors with the consideration of environmental factors. We present this article in accordance with the MDAR reporting checklist (available at https://tp.amegroups.com/article/view/10.21037/tp-22-512/rc).

Methods

General information about the study

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The research work was approved by the Ethical Committee of Shinshu University School of Medicine (No. 622/2020), Nagoya University Graduate School of Medicine (No. 18185/2020), and by the Ministry of Health in Iraq (No. 2553/2018). All methods were carried out in accordance with relevant guidelines and regulations and a written informed consent was obtained from all subjects and/or their legal guardian(s).

Five main pediatric oncology centers from Iraq have participated in this study, including Children Welfare Teaching Hospital (CWTH) in Baghdad (the major referral center for childhood cancers in the country), Basra Children's Specialty Hospital (BCSH) in Basra, Ibn Al-Atheer Hospital for Children (IAH) in Mosul, Hiwa Cancer Hospital (HCH) in Sulaymaniyah, and Jin Pediatric Hematology-Oncology Centre (JPHOC) in Duhok. CWTH, BCSH, and IAH are in Arab provinces, whereas HCH and JPHOC are in Kurdistan, the area inhabited mostly by the Kurdish ethnicity in the north of Iraq. HCH in Sulaymaniyah is the only oncology center in Iraq, which is equipped with hematopoietic stem cell transplantation unit under the supervision of an Italian team, and they are performing the minimal residual disease detection using flow cytometry. Patients in the mentioned centers were treated according to the United Kingdom-Medical Research Council (UK-MRC) protocols for pediatric acute leukemia, including the modified UKALL 2011 for ALL, and AML-MRC15 for AML, described elsewhere (10,17).

Sample collection

In the form of DBS, paired bone marrow (BM) samples were collected at diagnosis (day 0, tumor status), and at (day 30 or 60, remission status), from Iraqi patients aged ≤16 years, who were newly diagnosed with ALL or AML, from June 2016 to December 2019. Providing that no molecular data are available in Iraq upon diagnosis and the samples were received sequentially within the first few weeks of diagnosis, selection bias are not expected. In total, 101 cases were recruited from Iraq, 53 from CWTH, 25 from BCSH, 15 from HCH, and 8 from JPHOC. However, 66 (55 ALL and 11 AML) cases who had paired BM samples (at diagnosis and remission) were eligible for NGS, including 36, 17, 8, and 5, from CWTH, BCSH, HCH, and JPHOC, respectively. The remaining 35 cases were either missing one sample (n=15), died before reaching remission (n=9), insufficient in terms of DNA concentration (n=5), transferred to be treated outside Iraq (n=3), or abandoned therapy (n=3).

Flinders Technology Associates (FTA) paper processing

A few drops of blood from BM aspirate at initial diagnosis and at remission status were applied to the FTA classic card's filter paper (Cat No. WB120205, GE Healthcare, Buckinghamshire, UK Limited) (18) at the five Iraqi hospitals. After the blood spots were dried for 1 hour at room temperature, the FTA card was kept in a special FTA envelope in a refrigerator for up to several weeks and was then transported by airplane to Japan. Two mm disks (eight disks) were punched out from the DBS on FTA cards using a sterile hole puncher (Harris Micro-Punch, Shunderson Communications Inc., Ottawa, Canada). For the matched remission status samples especially those with hypoplastic BM samples, more DBS disks (up to 40) were consumed to increase the DNA yield. DNA was extracted from the DBS from the samples of FTA cards and was purified using the QIA amp DNA Blood Mini Kit (Cat No. 56304, Qiagen, Ltd., Tokyo, Japan) as per the manufacturer's instructions. After the extraction, DNA was measured using Qubit® 2.0 Fluorometer (Thermo Fisher Scientific, Life Technologies, MA, USA) as per the manufacturer's instructions.

Whole-exome sequencing (WES)

NGS analyses have been performed essentially as described (19). Briefly, WES libraries starting from 50–200 ng of DNA have been prepared using a SureSelect Human All Exon V5 bait and SureSelect Reagents (Agilent, Santa Clara, CA, USA) as per the manufacturer's

instructions. The libraries were run on a HiSeq X next-generation sequencer (Illumina, San Diego, CA, USA), with a 2×150-bp paired end-reads option. The sequence reads were aligned to the hg19 reference genome using the Burrows-Wheeler Aligner (http://bio-bwa.sourceforge.net/) with default parameters and a "–mem" option. Polymerase chain reaction (PCR) duplicates were removed from constructed BAM files using the Picard tools (https://broadinstitute.github.io/picard/).

To identify somatic point mutations, paired tumor-normal data were analyzed using VarScan2. We then called candidate variants in the coding region that have variant allele frequencies (VAF) of >0.1 (in tumor) and <0.05 (in normal), 10 or more reads with the variant, and minor allele frequencies (MAF) of <0.001 in single nucleotide polymorphism (SNP) databases (ESP6500, 1000 genomes, ExAC, and Kaviar). A candidate variant was considered as an artifact and was filtered out if the identical variant was present in 12 irrelevant germline samples with an average VAF of >0.01. The variants were then annotated using ANNOVAR (https://annovar.openbioinformatics.org/).

In total, 50 candidates were randomly selected, and PCR-based deep sequencing was performed. Briefly, a *Not*I-tagged PCR primers (having 5'-AAGCGGCCGC-3' tag on their 5'- side) were designed to cover 100–200 bp regions including candidates. PCR products were digested using *Not*I (New England Biolabs, Ipswich, MA, USA) and concatenated using T4 DNA Ligase (TaKaRa Bio, Otsu, Japan). The concatemers were fragmented to an average length of 400 bp by Covaris M220 (Covaris, Wobam, MA, USA) and were prepared for sequencing using an NEBNext Ultra DNA Prep Kit for Illumina (New England Biolabs) as per the manufacturer's instructions. BAM files have been assembled, and VAF of candidates has been measured. The candidate is considered present if the VAF in tumor was three times more than that in normal. As a result, (48/50, 96%) were confirmed to be present as somatic mutations.

To identify germline variants, variants with VAF >0.25 in normal data have been picked up using VarScan2. The variants were annotated using ANNOVAR. Genetic diagnoses were considered only when germline variants fulfilled the criteria of "pathogenic" or "likely pathogenic" as provided by the American College of Medical Genetics guideline, as described (20). The zygosity of a variant was considered homozygous when the VAF of the variant exceeded 0.85.

To identify copy number alterations (CNAs), a read count of an exon in a tumor sample was normalized for the total coverage of the sample and was compared with 12 irrelevant germline samples. The exon was considered a candidate of CNAs if the standard deviation of the tumor sample's read count was >3. If three or more continuous exons are the candidates, the exons are considered affected by amplifications or deletions.

Run of homozygosity (ROH) was identified by detecting a run of homozygous common (>1% MAF in SNP databases) SNPs in normal samples. ROH was classified by simply counting the number of continuous SNPs (10–50 SNPs; short ROH, and >50 SNPs; long ROH). A total of 60 germline samples of Japanese origin were used as control.

Targeted gene sequencing (TGS)

A custom SureSelect bait was designed targeting the whole gene body of 31 genes associated with fusion genes or deletions in B-ALL (Table S1). To detect chromosomal structural variations (SVs), soft-clipped bases were realigned to the hg19 using BLAT (https://hgwdev.gi.ucsc.edu/~kent/src/). A candidate SV supported by five or more reads with the identical breakpoint was visually interrogated using the Integrative Genomics Viewer (https://software.broadinstitute.org/software/igv/).

Whole-genome sequencing (WGS)

A total of 13 cases were analyzed using WGS. WGS libraries were prepared starting from 50–100 ng of DNA using an NEBNext Ultra II DNA Prep Kit for Illumina (New England Biolabs), according to the manufacturer's instructions. Somatic and germline variants and SVs were detected using the approaches used in WES and TGS. A copy number estimate of 10 kb bin was made simply from the number of reads within the bin divided by the mean coverage of the sample.

Co-occurrence simulation

The probability of co-occurrence of two kinds of genetic alterations was calculated using a Monte-Carlo simulation approach, based on the number of total patients, the number of patients with either of the mutations, and the number of patients with both mutations, as described (21).

Statistics

Statistical analyses were performed using SPSS program

Table 1 Clinical characteristics of 66 Iraqi children with ALL and AML

Acute leukemia type	Variable	Number of patients (%)
ALL (n=55)	Sex	
	Male	36 (65.5)
	Female	19 (34.5)
	Age (years)	
	1–<5	29 (52.7)
	5–<10	19 (34.6)
	≥10	7 (12.7)
	WBC ($\times10^9$/L)	
	<20	27 (49.1)
	20–<50	9 (16.4)
	≥50	19 (34.5)
	ALL subtypes	
	B-ALL	49 (89.1)
	T-ALL	6 (10.9)
AML (n=11)	Sex	
	Male	8 (72.7)
	Female	3 (27.3)
	Age (years)	
	1–<5	5 (45.5)
	5–<10	1 (9.0)
	≥10	5 (45.5)
	AML subtypes	
	M2	4 (36.3)
	M3 (APL)	5 (45.5)
	M5	1 (9.1)
	M6	1 (9.1)

ALL, acute lymphoblastic leukemia; AML, acute myeloid leukemia; WBC, white blood cell; APL, acute promyelocytic leukemia.

v. 28 (SPSS, IBM Corporation, Armonk, NY, USA). The unpaired Student's *t*-test was used in determining the significance of differences between two independent groups, and the Mann-Whitney U-test was used for data that were not normally distributed. Chi-square test or Fisher's exact test was used to compare the frequencies of genetic mutations between our cases and those from other countries. Statistical significance was defined as a P value of <0.05.

Results

Study cohort and design

This study included 49, 6, and 11 cases of B-cell precursor ALL (B-ALL), T-cell precursor ALL (T-ALL), and AML, respectively (*Table 1*). The median age among B-ALL was 4.2 (1–13) years, with a male to female ratio (M/F) of 1.7, meanwhile, the median age among T-ALL cases was 9.3 (3.5–12.8) years, and 5/6 of them were males. The median white blood cell (WBC) count in B-ALL and T-ALL was 16.4 (2.4–181) $\times10^9$/L and 280.5 (4.2–700) $\times10^9$/L, respectively. The average age and WBC were significantly higher in T-ALL compared to B-ALL, with P values of 0.007 and <0.001, respectively. The 3-year event-free survival (EFS) was 70.9%, and the 3-year overall survival (OS) was 74.5%. In AML, a higher frequency of APL or French-American-British (FAB) AML-M3 morphology (5/11, 45%) was observed, followed by FAB-M2 (4/11, 36%), and one case of M5 and M6. A case of AML-M2 (UPN49) was found to be secondary to chemotherapy or therapy-related AML (s-AML), for a previously cured germ cell tumor.

Either WES (53 cases), or WGS (13 cases) for each patient was performed. Paired tumor (BM specimen at diagnosis) and germline (BM at remission) samples were analyzed for each patient to identify somatic and germline mutations. Additionally, for patients with B-ALL analyzed by WES (40 cases), targeted sequencing was performed to identify fusion genes (Table S1).

Quality assessment of DBS-derived DNA

Since using DBS-derived DNA was unusual for NGS, the performance of our analysis was checked. As a result, an average of 81.0× and 30.3× coverage was obtained in WES and WGS analyses, respectively. The coverage resulted in 97.7% and 97.3% of the coding region covered by 10 or more unique reads, suggesting that DBS-derived DNA can be utilized for NGS.

Somatic mutations

A comprehensive detection of point mutations, small insertions/deletions (indels), copy number variants, and chromosomal SVs was performed in 66 patients with acute

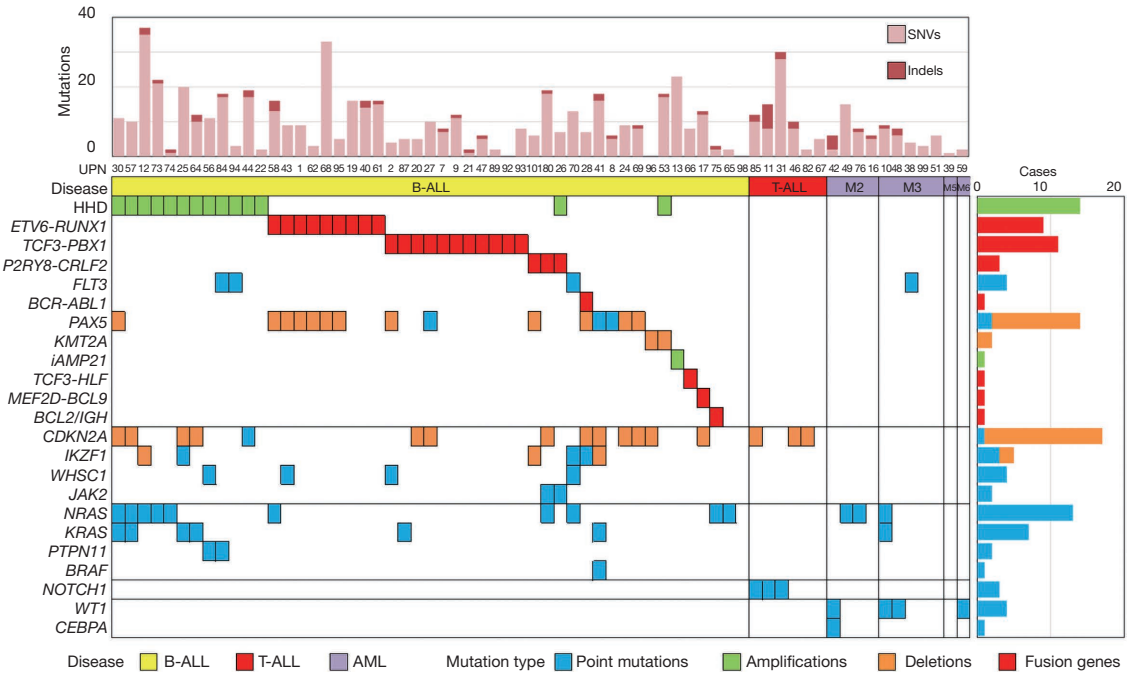

Figure 1 Mutational landscape of childhood acute leukemia in Iraq. Mutational landscape of B-ALL; n=49, T-ALL; n=6, and AML (M2, M3, M5, and M6); n=11. Each column indicates a patient, while each row indicates the kind of mutation. Boxes indicate mutations (blue, point mutations; green, chromosomal amplifications; orange, deletions; red, fusion genes). The bar chart on the top indicates the number of mutations in the coding region. The bars on the right indicate the number of cases with the indicated mutations. SNVs, single nucleotide variants; indels, insertions and deletions; UPN, unique patient number; HHD, high hyperdiploidy; iAmp21, intrachromosomal amplification of chromosome 21; B-ALL, B-cell precursor acute lymphoblastic leukemia; T-ALL, T-cell precursor ALL; AML, acute myeloid leukemia.

leukemia in Iraq (*Figure 1*, Tables S2,S3).

At least one driver mutation in 48 (95%) cases with B-ALL was identified, and accordingly B-ALL cases were classified. In 21 (42%) cases, 2 major subsets of B-ALL, including, high hyperdiploid (HHD) (>50 chromosomes), and *ETV6-RUNX1*, were identified representing 12, and 9 cases, respectively. The pattern of chromosomal amplification in HHD cases and the positions of chromosomal recombination in *ETV6-RUNX1* cases were found to be typical (*Figure 2A*,*2B*). Surprisingly, *TCF3-PBX1*, which usually constitutes 3–5% of a B-ALL cohort, explained 11 (22.4%) cases. This observation was not caused by cross-contamination of samples, because each patient carried a unique chromosomal breakpoint (*Figure 2C*). Moreover, the fusion gene was frequently associated with the amplification and deletion of chromosomes 1 and 19, respectively (*Figure 2D*). Four Ph-like ALL cases (3 with *P2RY8-CRLF2* and a case with *FLT3*-tyrosine kinase domain mutation) were identified (22), in addition to a Ph-ALL case with *BCR-ABL1*. Two of the 3 cases with *P2RY8-CRLF2* carried concomitant *JAK2* p.Arg683Gly point mutations. Other classifications were *PAX5* alterations (four cases) (22), *KMT2A* (*MLL*) deletion (two cases), intrachromosomal amplification of chromosome 21 (iAMP21), one case (*Figure 2E*), *TCF3-HLF* fusion (one), *MEF2D-BCL9* fusion (one), and *BCL2-IGH* fusion (one case).

Two cases (UPN65 and UPN98) did not carry any mutation associated with B-ALL classification and thus were classified with B-other ALL. The contamination of tumor in the germline sample or the scarcity of tumor cells in the tumor specimen was considered to explain the absence of classification in these two patients (Table S4).

Some mutations showed co-occurrence within a patient. Both HHD and RAS pathway mutations (*NRAS*, *KRAS*, *PTPN11*, and *BRAF* mutations) were detected in nine patients (P=0.0058) (Table S5). Also, 6/9 patients with *ETV6-RUNX1* carried *PAX5* mutations (P=0.0066).

In patients with T-ALL or AML, several mutations that

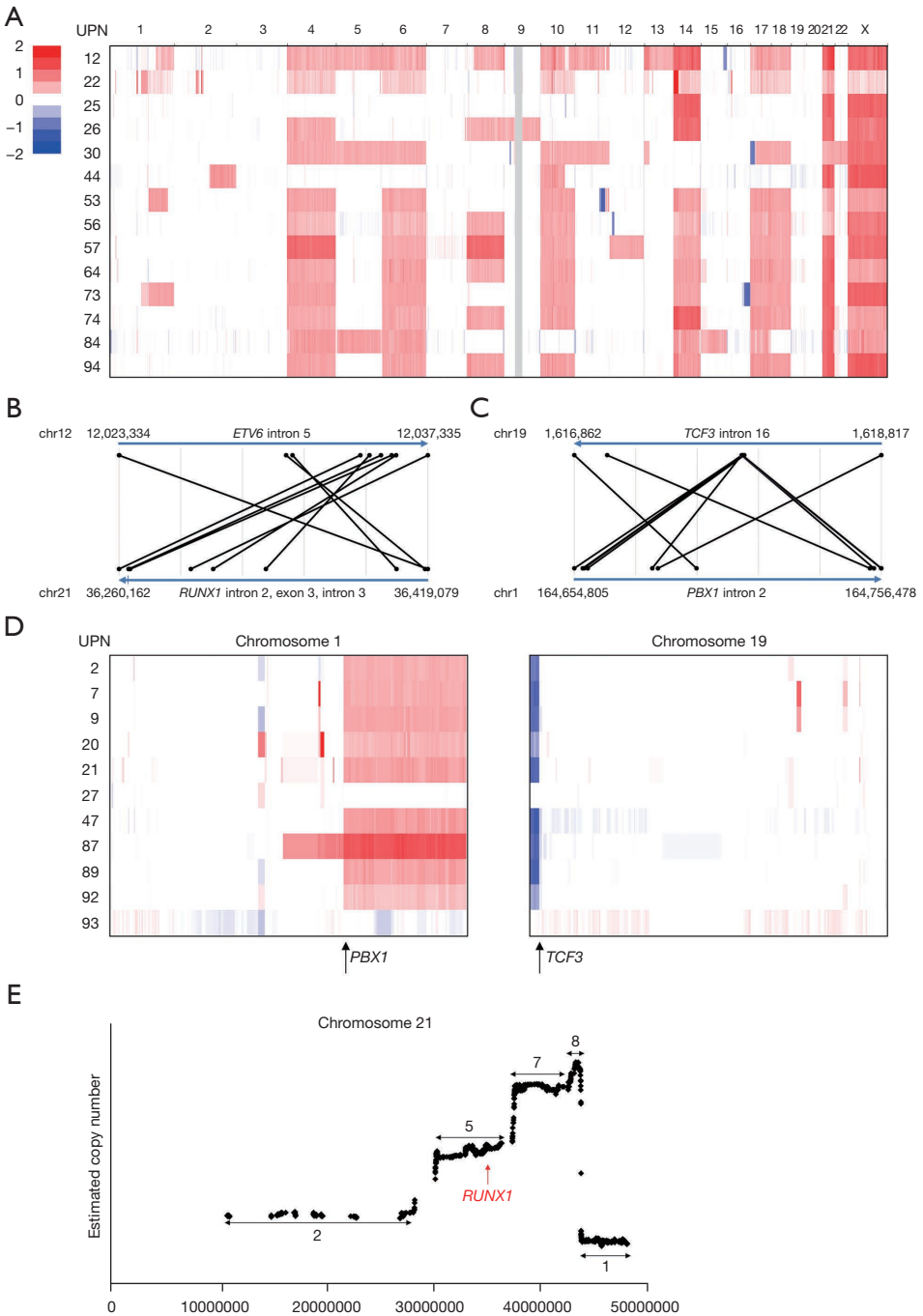

Figure 2 Copy number aberrations and fusion genes in B-ALL in Iraq. (A) Chromosomal copy number alterations in B-ALL patients with HHD. Numbers on the top indicate chromosome numbers. Red and blue indicate amplified and deleted regions, respectively, while grey indicates regions where the copy number could not be determined. (B) Chromosomal breakpoints of B-ALL patients with *ETV6-RUNX1*. Each black line indicates a patient's breakpoints. Numbers indicate the hg19 genomic coordinate. (C) Chromosomal breakpoints of *TCF3-PBX1*. (D) Chromosomal amplification and deletion in chromosomes 1 and 19 in patients with *TCF3-PBX1*, using the same color codes as in (A). The positions of *TCF3* and *PBX1* are indicated by arrows. (E) Intrachromosomal amplification of 21 identified in UPN13. The X- and Y-axes indicate the genomic coordinate and the estimated copy number, respectively. Numbers and arrows also indicate the estimated copy numbers. UPN, unique patient number; B-ALL, B-cell precursor acute lymphoblastic leukemia; HHD, high hyperdiploidy.

are characteristic of these diseases have been identified. T-ALL carried *NOTCH1*, *PTEN*, *ETV6*, *IL7R*, *RUNX1*, *RPL10*, and *SUZ12* mutations and *CDKN2A* deletions. AML carried *WT1*, *CEBPA*, *FLT3*, *MYC*, *KRAS*, and *NRAS* mutations. Because fusion gene detection was not performed in patients who had these diseases and were analyzed by WES, characteristic fusion genes were mostly not identified in these patients; but at least, *PML-RARA* was detected by WGS in UPN99 with AML-M3 and thus confirmed the PCR result of that patient in Iraq.

The number of somatic point mutations in the coding region was 0–37 (9.9 on average) and was considered comparable with similar diseases in other countries. B-ALL and T-ALL were noted to significantly differ in terms of the number of indels (0.81 *vs.* 2.66 on average, P=0.002), while the small number of T-ALL cases defies its interpretation. The number of indels looked high in several patients (including UPN11 who carried seven indels); however, the number was not statistically significant.

The type of nucleotide alterations of somatic mutations was biased toward C-to-T transitions (40%), suggesting that most somatic point mutations were acquired because of cell division (*Figure 3A*). Indels accounted for 7.5% of the somatic mutations.

The RAS pathway mutations were present in (16/49, 32.7%) of B-ALL cases and (3/11, 27.3%) of AML cases. Thus, in line with our previous reports, *RAS* mutations are prevalent among Iraqi children with acute leukemia compared with that of other countries (23-28). Three patients carried 2 *RAS* pathway mutations in ALL (*NRAS* and *KRAS* in both UPN30 and UPN57, *KRAS* and *BRAF* in UPN41), and one AML case (UPN10) had *NRAS* and *KRAS* mutations.

Germline variations

Germline mutations were analyzed; however, those related to the known inherited diseases were not identified. We also could not point out any pathogenic variants associated with leukemia predisposition. Meanwhile, several drug metabolism-associated SNPs were disclosed for 6-mercaptopurine (6-MP) and methotrexate (MTX) (29-31) (*Figure 3B*). Two SNPs (*ITPA* rs1127354 and *NUDT15* rs116855232) associated with 6-MP toxicity were present with MAF of 0.038, and 0.015, respectively. *MTHFR* rs1801131 (MTHFR-A or c.1298A>C) and rs1801133 (MTHFR-C or c.677C>T), which affect MTX metabolism, were frequent (MAF of 0.409 and 0.265, respectively). As a result, (23/66, 34.8%) cases of our cohort were affected by 2 or more MTHFR-A/C risk alleles. *SLCO1B1*, another MTX catalyzer, carried several drug metabolism-associated SNPs including rs11045819 (MAF =0.144) and rs4149056 (MAF =0.197). Additionally, it may be notable that *SLCO1B1* was also affected by rare nonsense mutations including rs7158941 (p.Arg580Ter, three patients, MAF =0.023) and p.Trp171Ter (one patient).

ROH was frequently observed in the germline of children with acute leukemia in Iraq, possibly because of consanguinity (*Figure 3C*). In total, (29/66, 43.9%) patients carried at least 1 ROH that had a length of >10 Mb. However, a significant accumulation of ROH could not be identified. At the very least, the lengths of ROH were significantly longer compared with those of Japanese samples [Iraq: 0–392,289,752 bp (70,539,478 bp on average); Japan: 0–148,869,110 bp (8,397,369 bp on average), $P=2.6\times10^{-6}$].

Genetic findings and clinical presentations

Several fusion genes were associated with clinical parameters; 9 patients with *ETV6-RUNX1* fusion gene had a median age of 4 (3–4.75) years, with M/F ratio of 3.5, and a median WBC of 11.7 (4.5–72) $\times10^9$/L (*Table 2*). The average WBC associated with *ETV6-RUNX1* cases was lower than those B-ALL cases without *ETV6-RUNX1*; however, it was of no significance (22.5 *vs.* 42) $\times10^9$/L, respectively (P=0.237). Eleven patients with *TCF3-PBX1* fusion gene had a median age of 5.7 (2–12) years, M/F ratio of 4.5, and a median WBC of 52.5 (4.6–152) $\times10^9$/L. The average WBC in patients with *TCF3-PBX1* was significantly higher than those *TCF3-PBX1*-negative B-ALL cases (63.4 *vs.* 31.2) $\times10^9$/L, respectively (P=0.033), whereas the average number of somatic mutations per patient associated with *TCF3-PBX1* was significantly lower than those B-ALL cases without *TCF3-PBX1* (5.6 *vs.* 11.9), respectively (P=0.023).

Several drug metabolism-associated SNPs were found to be correlated with adverse effects of chemotherapy (*Table 3*); UPN93, who carried three *MTHFR* risk alleles (heterozygous MTHFR-A and homozygous MTHFR-C), had experienced frequent interruptions in chemotherapy protocol owing to neutropenia or elevated liver function test. Likewise, UPN25, who carried heterozygous NUDT15 and MTHFR-C risk alleles, had recurrent febrile neutropenia and abnormal liver function test, resulting in relapse, and eventually died. Additionally, UPN87, who carried homozygous MTHFR-C, suffered from frequent neutropenia. Finally, UPN80 with heterozygous

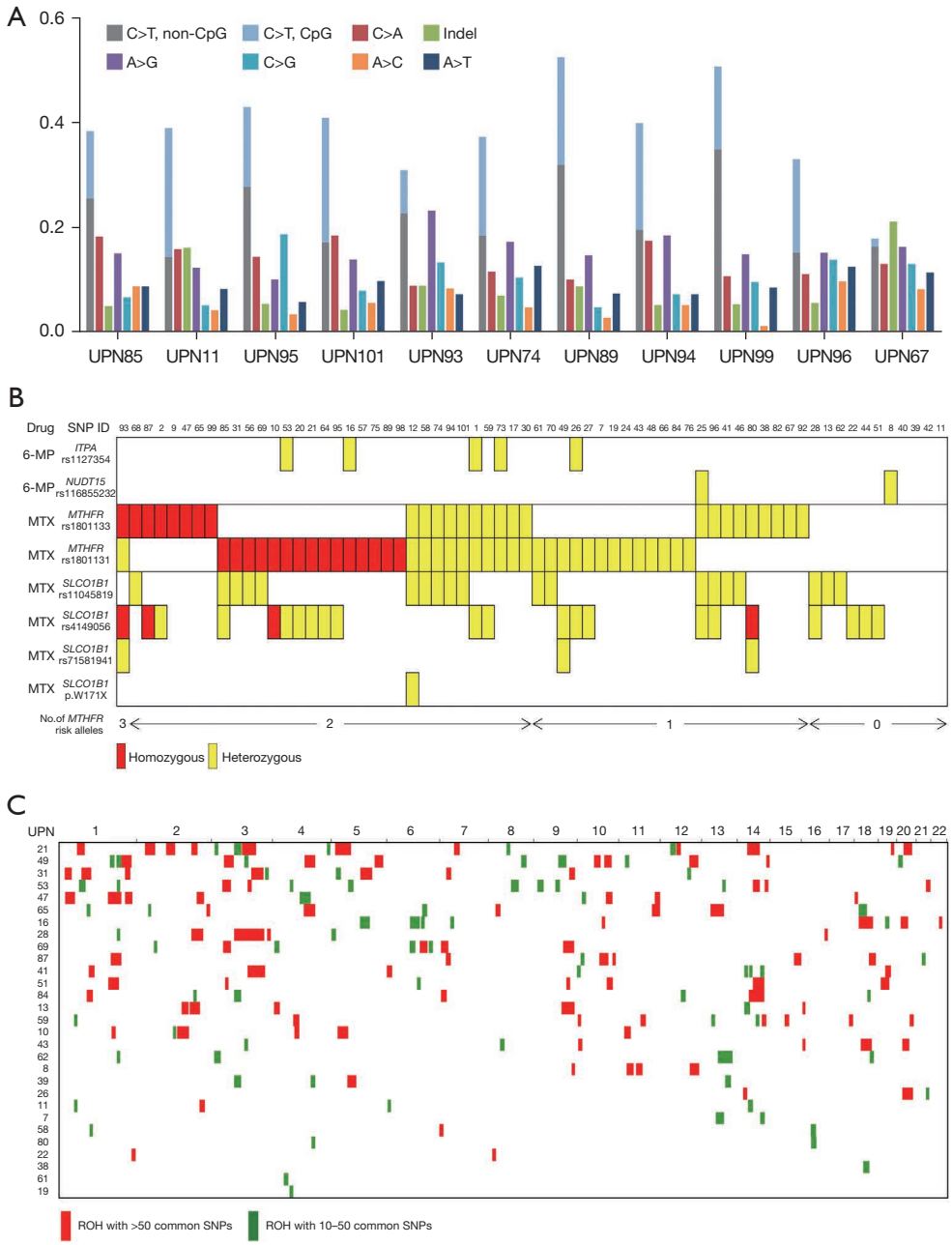

Figure 3 Other genetic findings. (A) Nucleotide alteration patterns of somatic mutations identified using whole-genome sequencing. The result of 11 patients who carried >50 somatic point mutations is presented, sorted in the descending order of total number of mutations. C-to-T transitions were separated into those in the CpG context and those in the non-CpG context. (B) Germline variants associated with drug metabolism. Each column and row indicate a patient and a SNP, respectively. Red and yellow indicate homozygous and heterozygous variants, respectively. The sum of the MTHFR alleles for each patient is also indicated. (C) ROH map. Red and green indicate ROH with >50 and 10–50 common SNPs, respectively. Only patients who carried one or more ROH regions in whole-exome sequencing are visualized. SNP, single nucleotide polymorphism; 6-MP, 6-mercaptopurine; MTX, methotrexate; UPN, unique patient number; ROH, run of homozygosity.

Table 2 Clinical criteria, important genetic aberrations results, and the outcome of 55 Iraqi childhood ALL cases

UPN/sex	Modified UKALL-2011 clinical-based risk factors classification						Genetic aberrations		Notes and outcome****		
	Age (years)	Initial WBC $\times 10^9$/L	Initial CSF (CNS status)*	Pre-phase steroid response**	Post-induction BMA status	Protocol regimen	ALL classification	Potentially deleterious somatic point mutation	Relapse, site/time in weeks	Death, time of death in weeks/cause	Cured, in continuous CR and others
1/M	4.20	34.07	–	Good	CR by Morph***	A	*ETV6-RUNX1*		CNS/54		Follow-up
19/M	3.50	11.6	–	Good	CR by Morph	A	*ETV6-RUNX1*				Follow-up
43/M	4.10	22.4	–	Good	CR by Morph	A	*ETV6-RUNX1*	*WHSC1*			On therapy
58/M	4.00	5.1	–	Good	CR by Morph	A	*ETV6-RUNX1*	*NRAS*			On therapy
61/F	3.00	11.4	–	Good	CR by Morph	A	*ETV6-RUNX1*				Follow-up
62/M	3.00	29.8	–	NU	Flow-MRD: 0.005%	A	*ETV6-RUNX1*				On therapy
68/M	4.00	4.5	–	NU	Flow-MRD: 0.001%	A	*ETV6-RUNX1*				On therapy
95/F	4.75	11.7	–	Good	CR by Morph	A	*ETV6-RUNX1*				On therapy
40/F	3.75	72	–	Good	CR by Morph	B	*ETV6-RUNX1*	*U2AF1*			Follow-up
44/F	6.10	5.2	–	Good	CR by Morph	A	HHD	*CDKN2A, CDKN2A, CHD4*			Follow-up
64/F	4.00	7.4	–	NU	Flow-MRD: 0.003%	A	HHD	*KRAS*			Follow-up
56/F	6.00	2.4	–	Good	CR by Morph	A	HHD	*PTPN11, WHSC1*			Follow-up
57/M	2.50	11.1	–	Good	CR by Morph	A	HHD	*KRAS, NRAS*			On therapy
73/M	4.70	5.94	–	Good	CR by Morph	A	HHD	*NRAS, IKZF3*			On therapy
74/F	6.50	8.32	–	Good	CR by Morph	A	HHD	*NRAS*			Follow-up
84/F	5.10	16.4	–	NU	CR by Morph	A	HHD	*PTPN11, FLT3, ARID5B*			Follow-up
94/F	2.75	12.63	–	Good	CR by Morph	A	HHD	*FLT3*			On therapy
22/F	1.10	9.6	–	Good	CR by Morph	A	HHD			84/infection	
25/F	5.60	5.7	–	Good	CR by Morph	A	HHD	*IKZF1, KRAS*	BM/111	114/PD	
30/M	7.00	71.5	CNS3	Good	CR by Morph	B	HHD	*KRAS, NRAS*	BM/166	170/PD	
12/M	10.70	15.8	–	Poor	CR by Morph	B	HHD	*NRAS, KMT2D, KMT2D*			Follow-up
2/M	5.70	33.7	–	Good	CR by Morph	A	*TCF3-PBX1*	*WHSC1*			Follow-up
21/F	7.60	16.6	–	Good	CR by Morph	A	*TCF3-PBX1*	*SETD2*			Abandon/12 weeks

Table 2 (*continued*)

Table 2 (*continued*)

UPN/sex	Modified UKALL-2011 clinical-based risk factors classification						Genetic aberrations		Notes and outcome****		
	Age (years)	Initial WBC $\times 10^9$/L	Initial CSF (CNS status)*	Pre-phase steroid response**	Post-induction BMA status	Protocol regimen	ALL classification	Potentially deleterious somatic point mutation	Relapse, site/time in weeks	Death, time of death in weeks/cause	Cured, in continuous CR and others
92/M	2.00	4.6	–	Good	CR by Morph	A	*TCF3-PBX1*		CNS/37		On palliative therapy
7/M	7.00	40.7	–	Poor	CR by Morph	B	*TCF3-PBX1*				Follow-up
87/M	12.00	6.21	–	Good	CR by Morph	B	*TCF3-PBX1*	*KRAS*			On therapy
9/M	4.20	134	–	Good	Not in CR by Morph	C	*TCF3-PBX1*				Follow-up
89/F	10.00	65.75	–	Good	CR by Morph	B	*TCF3-PBX1*				Follow-up
93/M	9.90	52.5	–	Poor	CR by Morph	B	*TCF3-PBX1*				On therapy
20/M	3.40	125	–	Good	CR by Morph	B	*TCF3-PBX1*				Follow-up
27/M	2.20	152	–	Good	CR by Morph	B	*TCF3-PBX1*	*PAX5*			Follow-up
47/M	3.00	66.14	–	Good	CR by Morph	B	*TCF3-PBX1*	*IKZF3*	CNS/63	77/infection	
24/M	2.20	14.6	–	Good	CR by Morph	A	*PAX5alt*	*TP53*			On therapy
69/M	3.90	89.19	–	Good	CR by Morph	B	*PAX5alt*				On therapy
41/F	13.00	11.8	–	Good	CR by Morph	B	*PAX5alt*	*KRAS, BRAF, PAX5, XBP1*		45/infection	
26/M	2.00	14.4	–	Good	CR by Morph	A	*P2RY8-CRLF2*	*JAK2*			On therapy
80/M	3.00	47.4	–	NU	Flow-MRD: 0.001%	A	*P2RY8-CRLF2*	*NRAS, JAK2*			On therapy
101/M	4.20	85.07	–	Good	CR by Morph	B	*P2RY8-CRLF2*				On therapy
70/M	7.40	46.5	–	Good	CR by Morph	A	*Ph-like (FLT3)*	*NRAS, IKZF1, FLT3, WHSC1*	BM + CNS/61	92/PD	
53/F	2.90	3.8	–	Good	CR by Morph	A	*del(11)(q23)*				Follow-up
96/M	3.25	23	–	Poor	CR by Morph	B	*del(11)(q23)*				On therapy
28/M	6.80	144	CNS3	Good	CR by Morph	B	*BCR-ABL1*	*IKZF1*			Follow-up
75/F	7.75	3.7	–	Good	CR by Morph	A	*BCL2-IGH*	*NRAS*	CNS/111	112/PD	
66/F	13.00	4	–	NU	Flow-MRD: 0.009%	B	*TCF3-HLF*		BM/101	102/PD	
13/M	3.60	58.3	–	Good	CR by Morph	B	*iAMP21*	*BCORL1, CSF3R*	BM/118	149/PD	

Table 2 (*continued*)

Table 2 (*continued*)

UPN/sex	Modified UKALL-2011 clinical-based risk factors classification						Genetic aberrations		Notes and outcome****		
	Age (years)	Initial WBC $\times 10^9$/L	Initial CSF (CNS status)*	Pre-phase steroid response**	Post-induction BMA status	Protocol regimen	ALL classification	Potentially deleterious somatic point mutation	Relapse, site/time in weeks	Death, time of death in weeks/cause	Cured, in continuous CR and others
17/M	7.40	30.15	–	Good	CR by Morph	A	*MEF2D-BCL9*	*ARID1A*	BM + CNS/47	113/PD	
8/F	7.11	181	CNS3	Good	CR by Morph	B	B-other ALL	*PAX5*			Follow-up
65/M	1.00	51	–	NU	Flow-MRD: 0.02%	B	B-other ALL	*NRAS*			On therapy
98/M	8.00	3.16	–	Poor	Not in CR by Morph	C	B-other ALL			24/PD	
11/M	9.20	73	–	Poor	CR by Morph	B	T-ALL	*NOTCH1, ETV6, IL7R*			Follow-up
46/M	3.50	563	–	Poor	CR by Morph	B	T-ALL	*PTEN*		43/infection	
67/F	8.00	4.2	–	NU	Flow-MRD: 0.001%	B	T-ALL				Follow-up
82/M	10.60	700	CNS3	NU	CR by Morph	B	T-ALL		CNS/45	53/PD	
85/M	12.80	111	–	Good	CR by Morph	B	T-ALL	*NOTCH1, RPL10*			Follow-up
31/M	9.40	450	–	Good	CR by Morph	B	T-ALL	*NOTCH1, RUNX1, SUZ12, SUZ12*			On therapy

*, CNS status, CNS3 defined as CSF of >5 WBC/μL and cytospin positive for blasts; **, a seven-day pre-phase steroid response defined as a drop in peripheral blasts count of <1.0×10^9/L at day 8; ***, CR by Morph, BM blasts <5% with normal hematopoietic recovery; ****, no patient had testicular involvement at initial diagnosis, and no case in our cohort had Down syndrome. ALL, acute lymphoblastic leukemia; UPN, unique patient number; WBC, white blood cells; CSF, cerebrospinal fluid; CNS, central nervous system; BMA, bone marrow aspirate; CR, complete remission; M, male; Morph, morphology; F, female; NU, not used; Flow-MRD, flow cytometry based-minimal residual disease; HHD, high hyperdiploidy; BM, bone marrow; PD, progressive disease.

Table 3 Drug metabolism-associated SNPs of 55 Iraqi childhood ALL cases and related clinical notes

UPN/ sex	Age (years)	Drug metabolism-associated SNPs								Clinical notes of drug toxicity
		Methotrexate						6-mercaptopurine		
		SLCO1B1				*MTHFR*		*ITPA*	*NUDT15*	
		rs4149056	*rs11045819*	*rs71581941*	*p.W171X*	*rs1801131*	*rs1801133*	*rs1127354*	*rs116855232*	
1/M	4.20	het	–	–	–	het	het	het	–	
19/M	3.50	–	–	–	–	het	–	–	–	
43/M	4.10	–	–	–	–	het	–	–	–	
58/M	4.00	–	het	–	–	het	het	–	–	
61/F	3.00	–	het	–	–	het	–	–	–	
62/M	3.00	–	het	–	–	–	–	–	–	
68/M	4.00	–	het	–	–	–	hom	–	–	
95/F	4.75	het	–	–	–	hom	–	–	–	
40/F	3.75	–	–	–	–	–	–	–	–	
44/F	6.10	het	–	–	–	–	–	–	–	
64/F	4.00	het	–	–	–	hom	–	–	–	MTX toxicity
56/F	6.00	–	het	–	–	hom	–	–	–	
57/M	2.50	–	–	–	–	hom	–	–	–	
73/M	4.70	–	–	–	–	het	het	het	–	
74/F	6.50	–	het	–	–	het	het	–	–	
84/F	5.10	–	–	–	–	het	–	–	–	
94/F	2.75	–	het	–	–	het	het	–	–	
22/F	1.10	het	–	–	–	–	–	–	–	
25/F	5.60	het	het	–	–	–	het	–	het	Frequent FN, abnormal LFT and jaundice
30/M	7.00	–	–	–	–	het	het	–	–	
12/M	10.70	–	het	–	het	het	het	–	–	
2/M	5.70	het	–	–	–	–	hom	–	–	
21/F	7.60	het	–	–	–	hom	–	–	–	
92/M	2.00	–	–	–	–	–	het	–	–	
7/M	7.00	–	–	–	–	het	–	–	–	
87/M	12.00	hom	–	–	–	–	hom	–	–	Recurrent neutropenia
9/M	4.20	–	–	–	–	–	hom	–	–	
89/F	10.00	–	–	–	–	hom	–	–	–	
93/M	9.90	hom	–	het	–	het	hom	–	–	Recurrent neutropenia/and elevated LFT

Table 3 (*continued*)

Table 3 (*continued*)

UPN/ sex	Age (years)	Drug metabolism-associated SNPs								Clinical notes of drug toxicity
		Methotrexate						6-mercaptopurine		
		SLCO1B1				*MTHFR*		*ITPA*	*NUDT15*	
		rs4149056	*rs11045819*	*rs71581941*	*p.W171X*	*rs1801131*	*rs1801133*	*rs1127354*	*rs116855232*	
20/M	3.40	het	–	–	–	hom	–	–	–	
27/M	2.20	het	–	–	–	het	–	–	–	
47/M	3.00	–	–	–	–	–	hom	–	–	
24/M	2.20	–	–	–	–	het	–	–	–	
69/M	3.90	–	het	–	–	hom	–	–	–	
41/F	13.00	–	het	–	–	–	het	–	–	
26/M	2.00	het	–	–	–	het	–	het	–	
80/M	3.00	hom	–	het	–	–	het	–	–	MTX toxicity
101/M	4.20	–	het	–	–	het	het	–	–	
70/M	7.40	–	het	–	–	het	–	–	–	
53/F	2.90	het	–	–	–	hom	–	het	–	
96/M	3.25	het	het	–	–	–	het	–	–	
28/M	6.80	het	het	–	–	–	–	–	–	
75/F	7.75	–	–	–	–	hom	–	–	–	
66/F	13.00	–	–	–	–	het	–	–	–	MTX toxicity? Frequent neutropenia
13/M	3.60	–	het	–	–	–	–	–	–	
17/M	7.40	–	–	–	–	het	het	–	–	
8/F	7.11	–	–	–	–	–	–	–	het	
65/M	1.00	–	–	–	–	–	hom	–	–	
98/M	8.00	–	–	–	–	hom	–	–	–	
11/M	9.20	–	–	–	–	–	–	–	–	
46/M	3.50	–	het	–	–	–	het	–	–	
67/F	8.00	–	–	–	–	–	het	–	–	
82/M	10.60	–	–	–	–	–	het	–	–	
85/M	12.80	het	het	–	–	hom	–	–	–	
31/M	9.40	–	het	–	–	hom	–	–	–	

SNPs, single nucleotide polymorphism; ALL, acute lymphoblastic leukemia; UPN, unique patient number; M, male; het, heterozygous; F, female; hom, homozygous; MTX, methotrexate; FN, febrile neutropenia; LFT, liver function test.

MTHFR-C and UPN64 with homozygous MTHFR-A had been complaining from MTX toxicity, with delay in their treatment progress.

Discussion

The use of FTA cards for conventional molecular analysis including PCR and Sanger sequencing for Iraqi pediatric acute leukemia was previously reported (12-14,18). Whereas in this study, and for the first time NGS was utilized for illustrating the landscape of genetic mutations in a series of Iraqi children with acute leukemia.

Our results disclosed apparent differences in some genetic aberrations, including the unusually high frequency of *TCF3-PBX1* fusion gene in ALL (22.4%) and the prevalent APL in AML (45.5%), along with the high frequency of *RAS* signaling pathway mutations in both ALL (38.8%) and AML (36.4%). Less frequent, however, still comparable results were detected with *ETV6-RUNX1* (18.4%), *PAX5alt* (6.1%), and Ph-like ALL (8.2%) compared to those in the developed world. While HHD (24.5%), *BCR-ABL1* (2%), *iAMP21* (2%), *KMT2A* (*MLL*) deletion (2%), *MEF2D-BCL9* (2%), and *TCF3-HLF* (2%), were similar to those in other studies (19,22,32).

Risk stratification of our cohort according to the clinical characteristics set in the modified UKALL-11 protocol (10) assigned (28/55, 51%) as good risk group eligible for regimen-A treatment plan. Although a total of (32/55, 58.2%) cases carried the favorable risk according to genetic subsets made of *ETV6-RUNX1*, HHD, and *TCF3-PBX1*, the stepwise risk refinements, by combining the data, had recognized only (20/55, 36.4%), with the favorable prognostic criteria. Among them, (18/20, 90%) were in continuous complete remission, whether finished or still under treatment, albeit 1 died from infection before completing the therapy. As a result, the 3-year EFS was 70.9%, and the 3-year OS was 74.5%.

One of the striking observations of this study is the unprecedented high frequency of B-ALL cases possessing *TCF3-PBX1* fusion gene associated with the translocation t(1;19)(q23;p13), (11/49, 22.4%). Abundance of *TCF3-PBX1* in the current cohort is significantly higher than several studies from different ethnicities and countries, ranging from 3% to 7.2% (33-37), including our previous report of (11/264, 4.2%) in pediatric ALL in Iraq (18). Arguably, there might be an underestimation of the frequency of *TCF3-PBX1* using the DBS-derived RNA in our previous study compared to DNA. Also, in the previous study the cases were not defined whether of B or T-ALL subtype. Notably, although the number of cases in this study is fewer, our current results are supported both by the chromosomal breakpoints and copy number changes of chromosomes (1 and 19). In fact, our frequency was significantly higher than neighboring Arab countries; Saudi Arabia of 3.4% (38), and Palestine of 7.3% (39), as well as than the Middle Eastern countries of 6.2% (40). Remarkably, *TCF3-PBX1* incidence in this report is higher even than that of the African-American B-ALL cases of 16.3% and the Mexican of 14.6% (4,33).

Among AML cases, APL subtype was recurrent (5/11, 45.5%) representing about half of our AML cohort. Of note, frequencies of (9/26, 35%) and (24/134, 18%) of APL among Iraqi children with AML were reported by Testi *et al.*, and Al-Kzayer *et al.* (based on molecular diagnosis), respectively (14,15). Interestingly, APL seems to be a prevalent AML subtype in Iraq, and records from adolescents and adults in locally published study over 5-year period in a single center at Sulaymaniyah province by Tawfiq *et al.* (41) showed that APL represented 25.5% of the total AML cases. Indeed, compared to nearby Middle Eastern countries, our frequency is yet higher than that in Saudi Arabia, Israel, Oman, and Iran, which is 3.4%, 8%, 13%, and 16%, respectively (5). Our incidence was also higher than Japan (9%) and other international registries including the United States (5–10%) and Switzerland (2%) (5).

In agreement with our previous work (12,13), in B-ALL, the most frequent somatic mutations were those in the RAS signaling pathway made up of 10 *NRAS*, 6 *KRAS*, 2 *PTPN11*, and 1 *BRAF*, which were detected in (16/49, 32.7%), including 3 with double mutations. Compared to literature, the overall somatic *RAS* signaling pathway mutations of around 39% in Iraqi children with ALL are among the highest reported frequencies. Our incidence was comparable to that reported by Case *et al.* (42) with overall mutations of (26/86, 30%) of childhood ALL cases, provided that *FLT3* mutations were excluded from their results. Moreover, our frequency was higher than those reported by Liang *et al.* (24), with overall *RAS* mutations of (122/530, 23%) of B-ALL Taiwanese pediatric cohort (P=0.16), and by Zhang *et al.* (23), with overall mutations of (24/114, 21.1%) in 23 Chinese children with B-ALL (P=0.1). *RAS* mutations were reported in less frequency of 15–20% in previous studies among childhood ALL (25-28). Wiemels *et al.* elucidated that *RAS* mutation frequency among Hispanics was > twice compared to non-Hispanic whites, of 28%, and 13%, respectively, in their cohort, and that HHD-associated *RAS* mutations were 30%; while, in

our series, the latter was 75% (28).

Numerous researchers had investigated the role of environmental exposure to chemicals, including hydrocarbons, and the risk factors behind childhood acute leukemia. Moreover, *RAS* oncogene was linked to hydrocarbons and other environmental insults. However, whether such association is causal in fact or not remains unclear (7,12,25,28,43).

Iraq was exposed to environmental and chemical hazards that carried potential health risks during repeated wars. Furthermore, the chaotic situation that characterized Iraq, as a consequence of repeated wars, and the damaged infrastructures had resulted in an ongoing process of undifferentiated water and air pollution, with a negative impact on several health aspects in Iraq, including cancer (9,11,12,25,44).

Although racial, ethnic, and geographic differences in the frequency of molecular markers of childhood ALL are widely of concern, the distinct biological difference in the genetics of Iraqi childhood acute leukemia, which varies even with the surrounding countries and sometimes in a significant manner, despite the ethnic similarity, may emphasize the concept of the environmental impact, especially considering Iraq has a war zone environment.

Conclusions

Apart from disclosing the high frequency of *TCF3-PBX1*, NGS confirmed our previous finding of recurrent *RAS* mutations in Iraqi childhood acute leukemia. Our results suggest that the biology of Iraqi childhood acute leukemia is in part characteristic, where the war-aftermath environment or geography might play a role. Given the environmental differences in respect to the above complicated status of Iraq, our findings maybe of special interest to encourage more studies enrolling more ALL/AML cases from Iraq to focus on this issue.

Acknowledgments

We thank Mrs. Sadako Kamiya and the non-governmental organization "Japan Chernobyl Foundation" (JCF) for supporting this study. Furthermore, we thank all collaborators from different hospitals in Iraq who have made this work possible.

Funding: This work was supported by a grant from the "Japan Society for the Promotion of Science" (JSPS) for the Grants-in-Aid for Scientific Research (KAKENHI) [18K07295 to LFY Al-Kzayer and Y Okuno], through Shinshu University School of Medicine, Nagano, Matsumoto, Japan. In addition to the grant from "TERUMO LIFE SCIENCE FOUNDATION" to [LFY Al-Kzayer and Y Okuno], for our overseas research in 2019, in Japan. We are thankful to the above agencies for their generous grants.

Footnote

Reporting Checklist: The authors have completed the MDAR reporting checklist. Available at https://tp.amegroups.com/article/view/10.21037/tp-22-512/rc

Data Sharing Statement: Available at https://tp.amegroups.com/article/view/10.21037/tp-22-512/dss

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://tp.amegroups.com/article/view/10.21037/tp-22-512/coif). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The research work was approved by the Ethical Committee of Shinshu University School of Medicine (No. 622/2020), Nagoya University Graduate School of Medicine (No. 18185/2020), and by the Ministry of Health in Iraq (No. 2553/2018). All methods were carried out in accordance with relevant guidelines and regulations and informed consent was obtained from all subjects and/or their legal guardian(s).

References

1. Greaves M. Author Correction: A causal mechanism for

childhood acute lymphoblastic leukaemia. Nat Rev Cancer 2018;18:526.
2. Buffler PA, Kwan ML, Reynolds P, et al. Environmental and genetic risk factors for childhood leukemia: appraising the evidence. Cancer Invest 2005;23:60-75.
3. Hein D, Borkhardt A, Fischer U. Insights into the prenatal origin of childhood acute lymphoblastic leukemia. Cancer Metastasis Rev 2020;39:161-71.
4. Jiménez-Morales S, Miranda-Peralta E, Saldaña-Alvarez Y, et al. BCR-ABL, ETV6-RUNX1 and E2A-PBX1: prevalence of the most common acute lymphoblastic leukemia fusion genes in Mexican patients. Leuk Res 2008;32:1518-22.
5. Zhang L, Samad A, Pombo-de-Oliveira MS, et al. Global characteristics of childhood acute promyelocytic leukemia. Blood Rev 2015;29:101-25.
6. Rodriguez-Galindo C, Friedrich P, Alcasabas P, et al. Toward the Cure of All Children With Cancer Through Collaborative Efforts: Pediatric Oncology As a Global Challenge. J Clin Oncol 2015;33:3065-73.
7. Lim JY, Bhatia S, Robison LL, et al. Genomics of racial and ethnic disparities in childhood acute lymphoblastic leukemia. Cancer 2014;120:955-62.
8. Chen W, Liu D, Wang G, et al. Screening diagnostic markers for acute myeloid leukemia based on bioinformatics analysis. Transl Cancer Res 2022;11:1722-9.
9. Hagopian A, Lafta R, Hassan J, et al. Trends in childhood leukemia in Basrah, Iraq, 1993-2007. Am J Public Health 2010;100:1081-7.
10. Al-Hadad SA, Al-Jadiry MF, Ghali HH, et al. Treatment of childhood acute lymphoblastic leukemia in Iraq: a 17-year experience from a single center. Leuk Lymphoma 2021;62:3430-9.
11. Al-Hadad SA, Al-Jadiry MF, Al-Darraji AF, et al. Reality of pediatric cancer in Iraq. J Pediatr Hematol Oncol 2011;33 Suppl 2:S154-6.
12. Al-Kzayer LF, Sakashita K, Al-Jadiry MF, et al. Analysis of KRAS and NRAS Gene Mutations in Arab Asian Children With Acute Leukemia: High Frequency of RAS Mutations in Acute Lymphoblastic Leukemia. Pediatr Blood Cancer 2015;62:2157-61.
13. Al-Kzayer LF, Sakashita K, Al-Jadiry MF, et al. Frequent coexistence of RAS mutations in RUNX1-mutated acute myeloid leukemia in Arab Asian children. Pediatr Blood Cancer 2014;61:1980-5.
14. Al-Kzayer LF, Uyen le TN, Al-Jadiry MF, et al. Analysis of class I and II aberrations in Iraqi childhood acute myeloid leukemia using filter paper cards. Ann Hematol 2014;93:949-55.
15. Testi AM, Al-Hadad SA, Al-Jadiry MF, et al. Impact of international collaboration on the prognosis of childhood acute promyelocytic leukemia in Iraq. Haematologica 2006;91:509-12.
16. Zhang HH, Wang HS, Qian XW, et al. Ras pathway mutation feature in the same individuals at diagnosis and relapse of childhood acute lymphoblastic leukemia. Transl Pediatr 2020;9:4-12.
17. Burnett AK, Russell NH, Hills RK, et al. Optimization of chemotherapy for younger patients with acute myeloid leukemia: results of the medical research council AML15 trial. J Clin Oncol 2013;31:3360-8.
18. Al-Kzayer LF, Sakashita K, Matsuda K, et al. Genetic evaluation of childhood acute lymphoblastic leukemia in Iraq using FTA cards. Pediatr Blood Cancer 2012;59:461-7.
19. Suzuki K, Okuno Y, Kawashima N, et al. MEF2D-BCL9 Fusion Gene Is Associated With High-Risk Acute B-Cell Precursor Lymphoblastic Leukemia in Adolescents. J Clin Oncol 2016;34:3451-9.
20. Muramatsu H, Okuno Y, Yoshida K, et al. Clinical utility of next-generation sequencing for inherited bone marrow failure syndromes. Genet Med 2017;19:796-802.
21. Murakami N, Okuno Y, Yoshida K, et al. Integrated molecular profiling of juvenile myelomonocytic leukemia. Blood 2018;131:1576-86.
22. Ratti S, Lonetti A, Follo MY, et al. B-ALL Complexity: Is Targeted Therapy Still A Valuable Approach for Pediatric Patients? Cancers (Basel) 2020;12:3498.
23. Zhang H, Wang H, Qian X, et al. Genetic mutational analysis of pediatric acute lymphoblastic leukemia from a single center in China using exon sequencing. BMC Cancer 2020;20:211.
24. Liang DC, Chen SH, Liu HC, et al. Mutational status of NRAS, KRAS, and PTPN11 genes is associated with genetic/cytogenetic features in children with B-precursor acute lymphoblastic leukemia. Pediatr Blood Cancer 2018;65:10.1002/pbc.26786.
25. Shu XO, Perentesis JP, Wen W, et al. Parental exposure to medications and hydrocarbons and ras mutations in children with acute lymphoblastic leukemia: a report from the Children's Oncology Group. Cancer Epidemiol Biomarkers Prev 2004;13:1230-5.
26. Perentesis JP, Bhatia S, Boyle E, et al. RAS oncogene mutations and outcome of therapy for childhood acute lymphoblastic leukemia. Leukemia 2004;18:685-92.
27. Yamamoto T, Isomura M, Xu Y, et al. PTPN11, RAS and FLT3 mutations in childhood acute lymphoblastic

leukemia. Leuk Res 2006;30:1085-9.
28. Wiemels JL, Zhang Y, Chang J, et al. RAS mutation is associated with hyperdiploidy and parental characteristics in pediatric acute lymphoblastic leukemia. Leukemia 2005;19:415-9.
29. Zhou H, Li L, Yang P, et al. Optimal predictor for 6-mercaptopurine intolerance in Chinese children with acute lymphoblastic leukemia: NUDT15, TPMT, or ITPA genetic variants? BMC Cancer 2018;18:516.
30. Fukushima H, Fukushima T, Sakai A, et al. Polymorphisms of MTHFR Associated with Higher Relapse/Death Ratio and Delayed Weekly MTX Administration in Pediatric Lymphoid Malignancies. Leuk Res Treatment 2013;2013:238528.
31. Treviño LR, Shimasaki N, Yang W, et al. Germline genetic variation in an organic anion transporter polypeptide associated with methotrexate pharmacokinetics and clinical effects. J Clin Oncol 2009;27:5972-8.
32. Inaba H, Mullighan CG. Pediatric acute lymphoblastic leukemia. Haematologica 2020;105:2524-39.
33. Pui CH, Sandlund JT, Pei D, et al. Results of therapy for acute lymphoblastic leukemia in black and white children. JAMA 2003;290:2001-7.
34. Lin A, Cheng FWT, Chiang AKS, et al. Excellent outcome of acute lymphoblastic leukaemia with TCF3-PBX1 rearrangement in Hong Kong. Pediatr Blood Cancer 2018;65:e27346.
35. Yen HJ, Chen SH, Chang TY, et al. Pediatric acute lymphoblastic leukemia with t(1;19)/TCF3-PBX1 in Taiwan. Pediatr Blood Cancer 2017;64:10.1002/pbc.26557.
36. Wang Y, Xue YJ, Lu AD, et al. Long-Term Results of the Risk-Stratified Treatment of TCF3-PBX1-Positive Pediatric Acute Lymphoblastic Leukemia in China. Clin Lymphoma Myeloma Leuk 2021;21:e137-44.
37. Asai D, Imamura T, Yamashita Y, et al. Outcome of TCF3-PBX1 positive pediatric acute lymphoblastic leukemia patients in Japan: a collaborative study of Japan Association of Childhood Leukemia Study (JACLS) and Children's Cancer and Leukemia Study Group (CCLSG). Cancer Med 2014;3:623-31.
38. Ahmed AM, Al-Trabolsi H, Bayoumy M, et al. Improved Outcomes of Childhood Acute Lymphoblastic Leukemia: A Retrospective Single Center Study in Saudi Arabia. Asian Pac J Cancer Prev 2019;20:3391-8.
39. Shawahna R, Mosleh S, Odeh Y, et al. Clinical characteristics and outcomes of patients with pediatric acute lymphoblastic leukemia after induction of chemotherapy: a pilot descriptive correlational study from Palestine. BMC Res Notes 2021;14:259.
40. Al-Mulla NA, Chandra P, Khattab M, et al. Childhood acute lymphoblastic leukemia in the Middle East and neighboring countries: a prospective multi-institutional international collaborative study (CALLME1) by the Middle East Childhood Cancer Alliance (MECCA). Pediatr Blood Cancer 2014;61:1403-10.
41. Tawfiq SA, Yassin AK, AlGetta HA, et al. Acute myeloblastic leukemia: Important clinical and epidemiological facts from Hiwa Hospital in Sulaimaniyah, Iraq. Iraqi J Hematol 2019;8:69-73.
42. 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;68:6803-9.
43. North M, Shuga J, Fromowitz M, et al. Modulation of Ras signaling alters the toxicity of hydroquinone, a benzene metabolite and component of cigarette smoke. BMC Cancer 2014;14:6.
44. Busby C, Hamdan M, Ariabi E. Cancer, infant mortality and birth sex-ratio in Fallujah, Iraq 2005-2009. Int J Environ Res Public Health 2010;7:2828-37.

Cite this article as: Al-Kzayer LFY, Saeed RM, Ghali HH, Tanaka M, Al-Jadiry MF, Faraj SA, Al-Hadad SA, Al Abdullah HMS, Majeed AA, Qadir AO, Abdullah DA, Noori KD, Hama ZM, Muhsin AA, Al-Doski AA, Al-Agele YS, Malallah AH, Al-Badrani KS, Khaleel AMA, Kamata M, Hamada M, Kojima S, Nakazawa Y, Okuno Y. Comprehensive genetic analyses of childhood acute leukemia in Iraq using next-generation sequencing. Transl Pediatr 2023;12(5):827-844. doi: 10.21037/tp-22-512

Supplementary

Table S1 Target-captured region for fusion gene detection

Gene	Target region*	Design remarks
ABL1	chr9:130713831-130887725	Whole gene
ABL2	chr1:179099277-179229734	Whole gene
BCL2	chr18:63123296-63320178	Whole gene
BCL9	chr1:147541362-147626269	Whole gene
BCR	chr22:23179654-23318087	Whole gene
CDKN2A	chr9:21967702-21995351	Whole gene
CDKN2B	chr9:22002853-22009413	Whole gene
CRLF2	chrX:1187499-1212800	Whole gene
CSF1R	chr5:150053241-150113422	Whole gene
DUX4	chr4:190173724-190185992	Whole gene
EBF1	chr5:158695865-159099830	Whole gene
EPOR	chr19:11377155-11384392	Whole gene
ERG	chr21:38367211-38661830	Whole gene
ETV6	chr12:11649804-11895452	Whole gene
FLT3	chr13:28003224-28100642	Whole gene
HNRNPUL1	chr19:41262426-41307742	Whole gene
IKZF1	chr7:50304033-50405151	Whole gene
IL7R	chr5:35852645-35879653	Whole gene
JAK2	chr9:4984983-5128233	Whole gene
KMT2A	chr11:118436440-118526882	Whole gene
MEF2D	chr1:156463671-156500892	Whole gene
MYC	chr8:127732884-127741484	Whole gene + 2.5 kb upstream
NUTM1	chr15:34343265-34357787	Whole gene
P2RY8	chrX:1462522-1537194	Whole gene
PAX5	chr9:36833225-37034529	Whole gene
PBX1	chr1:164555534-164899346	Whole gene
PDGFRB	chr5:150113787-150155922	Whole gene
RUNX1	chr21:34787751-35049348	Whole gene
TCF3	chr19:1609240-1652655	Whole gene
TP53	chr17:7661729-7687600	Whole gene
ZNF384	chr12:6666427-6689622	Whole gene

*, hg38 coordinate.

https://dx.doi.org/10.21037/tp-22-512

Table S2 Somatic point mutations identified on the exome

UPN	Gene	Reference	Nucleotide change	Effect	Amino acid change	VAF
1	*AGRN*	NM_001305275	c.1177+3G>C	splice site	(exon 6)	0.17
1	*FAM110B*	NM_147189	c.772C>T	missense	p.R258W	0.34
1	*GLIS3*	NM_152629	c.1926C>T	silent	p.T642T	0.39
1	*MAGI1*	NM_001033057	c.2819C>T	missense	p.T940M	0.33
1	*MAP3K10*	NM_002446	c.1971C>T	silent	p.S657S	0.16
1	*SERPINI1*	NM_001122752	c.125A>G	missense	p.E42G	0.44
1	*TRDN*	NM_001251987	c.1285C>T	nonsense	p.R429*	0.42
1	*ZNF148*	NM_021964	c.1979G>A	missense	p.R660Q	0.14
1	*ZNF552*	NM_024762	c.129G>T	silent	p.T43T	0.37
2	*EYS*	NM_001142800	c.2883C>T	silent	p.P961P	0.26
2	*SLC35E2*	NM_182838	c.784A>G	missense	p.M262V	0.59
2	**WHSC1*	NM_001042424	c.3295G>A	missense	p.E1099K	0.14
2	*ZNF506*	NM_001145404	c.677G>A	missense	p.R226K	0.16
7	*AIPL1*	NM_001033055	c.201C>T	silent	p.F67F	0.44
7	*DBH*	NM_000787	c.958_959insGGGGTCC	frameshift	S325Rfs*254	0.40
7	*IGFN1*	NM_001164586	c.1767C>T	silent	p.D589D	0.35
7	*MYO15A*	NM_016239	c.7520C>G	missense	p.P2507R	0.54
7	*OVCH1*	NM_183378	c.1593G>T	missense	p.L531F	0.25
7	*RNF150*	NM_020724	c.718G>T	missense	p.A240S	0.22
7	*SMC5*	NM_015110	c.340G>A	missense	p.V114M	0.51
7	*ZNF273*	NM_021148	c.1644C>T	silent	p.D548D	0.12
8	*CASC5*	NM_144508	c.628G>A	missense	p.E210K	0.44
8	*CFAP74*	NM_001304360	c.3451G>A	missense	p.A1151T	0.44
8	*KAT6B*	NM_001256468	c.3252C>T	silent	p.T1084T	0.42
8	*MCPH1*	NM_024596	c.1875_1876insGG	frameshift	p.F627Afs*12	0.47
8	**PAX5*	NM_016734	c.A397C	missense	p.S133R	0.10
8	*XKR4*	NM_052898	c.976G>A	missense	p.V326I	0.39
9	*BICC1*	NM_001080512	c.2124C>T	silent	p.A708A	0.21
9	*CSNK1A1*	NM_001271742	c.25G>A	missense	p.E9K	0.46
9	*ELF1*	NM_001145353	c.142dupT	frameshift	p.Y48Lfs*12	0.24
9	*FBXL18*	NM_024963	c.742C>T	missense	p.R248W	0.40
9	*GATB*	NM_004564	c.1651C>T	silent	p.L551L	0.41
9	*LRFN1*	NM_020862	c.245G>A	missense	p.R82H	0.54
9	*OLFML2B*	NM_001297713	c.969C>T	silent	p.S323S	0.93
9	*RPH3A*	NM_014954	c.1759A>G	missense	p.K587E	0.16
9	*SLC6A1*	NM_003042	c.846C>T	silent	p.S282S	0.11
9	*SMIM24*	NM_001136503	c.132C>A	silent	p.I44I	0.32
9	*TCTN2*	NM_001143850	c.1350C>T	silent	p.N450N	0.19
9	*UGT2B4*	NM_001297615	c.495A>G	silent	p.K165K	0.41
10	**KRAS*	NM_004985	c.35G>T	missense	p.G12V	0.06
10	*MZF1*	NM_003422	c.1987C>T	missense	p.R663W	0.11
10	*NLGN3*	NM_001166660	c.1377G>A	silent	p.S459S	0.47
10	*NOTCH2*	NM_024408	c.2546_2547delAA	frameshift	p.K849Rfs*6	0.38
10	**NRAS*	NM_002524	c.34G>A	missense	p.G12S	0.09
10	*RASSF9*	NM_005447	c.380G>A	missense	p.R127Q	0.12
10	*RBPMS*	NM_001008710	c.111T>C	silent	p.P37P	0.26
10	*REV1*	NM_001037872	c.1583C>A	missense	p.A528D	0.38
10	**WT1*	NM_000378	c.1091C>A	nonsense	p.S364*	0.39
11	*AKIP1*	NM_001206647	c.-6-5C>G	splice site	(exon 2)	0.52
11	*CACNA1B*	NM_000718	c.159G>A	silent	p.A53A	0.23
11	*CEND1*	NM_016564	c.421G>T	missense	p.G141C	0.42
11	*DHX34*	NM_014681	c.3410_3411insCT	frameshift	p.H1138Sfs*19	0.53
11	**ETV6*	NM_001987	c.771dupC	frameshift	p.R259Pfs*41	0.39
11	**IL7R*	NM_002185	c.760_761insAAA	in-frame	p.A254delinsET	0.42
11	*KANK2*	NM_015493	c.362A>G	missense	p.N121S	0.59
11	*MDGA1*	NM_153487	c.741C>T	silent	p.N247N	0.62
11	*NLGN3*	NM_018977	c.501C>T	silent	p.D167D	0.31
11	*NLRC5*	NM_032206	c.1210_1211insCC	frameshift	p. V405Rfs*32	0.66
11	**NOTCH1*	NM_017617	c.4719_4720insGGT	in-frame	p.L1574delinsGL	0.17
11	*OGFR*	NM_007346	c.1468_1469insCT	frameshift	p.H490Pfs*225	0.29
11	*PCDH7*	NM_001173523	c.2833C>A	missense	p.Q945K	0.44
11	*SATB1*	NM_001131010	c.454_455insAAGATAACCGGA	in-frame	p.T152delinsKDNRT	0.44
11	*ZSCAN5A*	NM_024303	c.1389C>T	silent	p.S463S	0.52
12	*ABCA6*	NM_080284	c.1529C>T	missense	p.T510M	0.23
12	*ABHD4*	NM_022060	c.579C>T	silent	p.A193A	0.31
12	*ADAMTS2*	NM_014244	c.2140G>A	missense	p.V714M	0.29
12	*AFF1*	NM_001313960	c.735A>C	missense	p.K245N	0.37
12	*ALPI*	NM_001631	c.1383C>T	silent	p.R461R	0.37
12	*BAIAP2L1*	NM_018842	c.276+4C>T	splice site	(exon 4)	0.41
12	*CCT5*	NM_012073	c.873+1G>C	splice site	(exon 6)	0.23

Table S2 (*continued*)

https://dx.doi.org/10.21037/tp-22-512

Table S2 (*continued*)

UPN	Gene	Reference	Nucleotide change	Effect	Amino acid change	VAF
12	*CNGB1*	NM_001286130	c.3584C>T	missense	p.P1195L	0.33
12	*CYSLTR2*	NM_001308471	c.269C>T	missense	p.T90M	0.29
12	*EIF2AK1*	NM_001134335	c.1231C>T	missense	p.P411S	0.42
12	*ENDOG*	NM_004435	c.576C>T	silent	p.N192N	0.38
12	*FARP1*	NM_001001715	c.354G>A	silent	p.A118A	0.19
12	*GATA3*	NM_001002295	c.520G>A	missense	p.G174S	0.37
12	*GLI3*	NM_000168	c.2718C>T	silent	p.S906S	0.38
12	*GPAM*	NM_001244949	c.952C>T	missense	p.R318C	0.18
12	*GRM6*	NM_000843	c.1780G>A	missense	p.V594M	0.32
12	*IRX1*	NM_024337	c.1194C>T	silent	p.H398H	0.21
12	**KMT2D*	NM_003482	c.12022_12036del	in-frame	p.4008_4012del	0.32
12	**KMT2D*	NM_003482	c.12015_12018delAAGA	frameshift	p.L4006Nfs*15	0.44
12	*LOC389199*	NM_203423	c.205G>A	missense	p.G69R	0.27
12	*MMAA*	NM_172250	c.1206C>T	silent	p.S402S	0.35
12	*MUC16*	NM_024690	c.23043G>A	silent	p.V7681V	0.37
12	*MYH4*	NM_017533	c.3301G>A	missense	p.E1101K	0.40
12	*MYH7*	NM_000257	c.1401C>A	silent	p.I467I	0.23
12	**NRAS*	NM_002524	c.38G>A	missense	p.G13D	0.45
12	*OR2L2*	NM_001004686	c.778C>T	missense	p.R260C	0.28
12	*OR5D14*	NM_001004735	c.808C>T	missense	p.R270W	0.28
12	*OVCH1*	NM_183378	c.2220G>A	silent	p.G740G	0.40
12	*RUNX1T1*	NM_175636	c.1389C>T	silent	p.D463D	0.21
12	*SLC45A4*	NM_001080431	c.234C>T	silent	p.G78G	0.27
12	*SYPL2*	NM_001040709	c.55-4G>A	splice site	(exon 2)	0.51
12	*TENM2*	NM_001080428	c.3278C>T	missense	p.T1093M	0.19
12	*TENM4*	NM_001098816	c.4682G>A	missense	p.R1561Q	0.33
12	*TSNARE1*	NM_001291931	c.92C>T	missense	p.T31I	0.28
12	*TTN*	NM_003319	c.48275G>A	missense	p.R16092Q	0.42
12	*VPS18*	NM_020857	c.1176A>G	silent	p.Q392Q	0.41
12	*ZNF503*	NM_032772	c.1704C>T	silent	p.A568A	0.34
13	*APBA3*	NM_004886	c.869C>G	missense	p.A290G	0.39
13	**BCORL1*	NM_001184772	c.2896G>T	nonsense	p.E966*	0.10
13	*CHSY1*	NM_014918	c.1650T>G	missense	p.F550L	0.20
13	**CSF3R*	NM_000760	c.1853C>T	missense	p.T618I	0.16
13	*CYFIP1*	NM_001033028	c.1882C>G	missense	p.L628V	0.35
13	*DFFB*	NM_001282669	c.431C>T	missense	p.A144V	0.48
13	*DLG2*	NM_001142702	c.752C>T	missense	p.A251V	0.30
13	*EPHA2*	NM_004431	c.1855A>T	missense	p.I619F	0.38
13	*FOXO3*	NM_001455	c.700T>G	missense	p.W234G	0.44
13	*FREM1*	NM_144966	c.1309T>A	missense	p.F437I	0.11
13	*GALNT9*	NM_021808	c.380G>T	missense	p.C127F	0.52
13	*GYS2*	NM_021957	c.719A>T	missense	p.H240L	0.41
13	*HAS3*	NM_001199280	c.121C>T	missense	p.H41Y	0.13
13	*IGF2BP1*	NM_001160423	c.600C>T	silent	p.A200A	0.33
13	*PRDM5*	NM_018699	c.723T>C	silent	p.S241S	0.24
13	*SCN1A*	NM_001165963	c.5612T>G	missense	p.F1871C	0.16
13	*SP7*	NM_152860	c.1253C>T	missense	p.A418V	0.38
13	*SPATA19*	NM_001291992	c.14C>A	missense	p.T5K	0.41
13	*ST18*	NM_014682	c.2371G>A	missense	p.G791R	0.38
13	*TNC*	NM_002160	c.324C>T	silent	p.R108R	0.25
13	*UROD*	NM_000374	c.994C>T	missense	p.R332C	0.35
13	*ZEB2*	NM_001171653	c.3041A>G	missense	p.H1014R	0.26
13	*ZNF536*	NM_014717	c.2288C>G	missense	p.S763C	0.26
16	*C3orf70*	NM_001025266	c.606G>A	silent	p.S202S	0.45
16	*COL18A1*	NM_030582	c.2445C>T	silent	p.P815P	0.48
16	*DAZAP1*	NM_018959	c.1216C>T	nonsense	p.R406*	0.54
16	*PIWIL1*	NM_001190971	c.1028T>A	nonsense	p.L343*	0.34
16	*PNPLA5*	NM_001177675	c.282C>T	silent	p.N94N	0.51
16	*SIN3B*	NM_015260	c.2711_2712insG	frameshift	p.D904Efs*32	0.44
17	*ARHGEF26*	NM_001251962	c.328C>T	missense	p.R110W	0.16
17	**ARID1A*	NM_006015	c.4906delC	frameshift	p.R1636Gfs*18	0.45
17	*CWH43*	NM_001286791	c.789C>T	silent	p.F263F	0.38
17	*ERN2*	NM_001308220	c.654G>A	silent	p.T218T	0.45
17	*FAT2*	NM_001447	c.7808C>T	missense	p.P2603L	0.13
17	*GRIK1*	NM_000830	c.280C>T	missense	p.R94W	0.53
17	*HOXA13*	NM_000522	c.1032T>C	silent	p.N344N	0.47
17	*HPGDS*	NM_014485	c.301T>C	missense	p.C101R	0.12
17	*MAP2*	NM_002374	c.2713C>T	nonsense	p.R905*	0.43
17	*MXRA5*	NM_015419	c.444C>T	silent	p.N148N	0.35
17	*MYBPC1*	NM_001254722	c.1197A>G	silent	p.K399K	0.37

Table S2 (*continued*)

https://dx.doi.org/10.21037/tp-22-512

Table S2 (*continued*)

UPN	Gene	Reference	Nucleotide change	Effect	Amino acid change	VAF
17	*PER2*	NM_022817	c.433G>A	missense	p.V145M	0.47
17	*SORCS3*	NM_014978	c.983G>A	missense	p.R328Q	0.42
19	*AGRN*	NM_198576	c.2796C>T	silent	p.N932N	0.27
19	*APOA4*	NM_000482	c.386G>A	missense	p.R129Q	0.15
19	*ASB7*	NM_024708	c.784C>T	nonsense	p.R262*	0.31
19	*C11orf85*	NM_001037225	c.93G>C	missense	p.K31N	0.41
19	*CCDC33*	NM_182791	c.753G>C	missense	p.L251F	0.42
19	*CCDC91*	NM_018318	c.925-4G>C	splice site	(exon 10)	0.58
19	*IL18*	NM_001243211	c.523G>C	missense	p.E175Q	0.58
19	*INTU*	NM_015693	c.1276G>C	missense	p.E426Q	0.50
19	*LGI2*	NM_018176	c.414-4C>G	splice site	(exon 5)	0.40
19	*LOXL1*	NM_005576	c.1158G>C	missense	p.Q386H	0.39
19	*MYO16*	NM_001198950	c.3486C>G	silent	p.L1162L	0.40
19	*PPL*	NM_002705	c.4669G>C	missense	p.E1557Q	0.37
19	*PRSS23*	NM_001293179	c.658C>T	nonsense	p.Q220*	0.32
19	*UBA2*	NM_005499	c.349G>A	missense	p.D117N	0.26
19	*ZNF526*	NM_001314033	c.492T>C	silent	p.L164L	0.47
19	*ZNF880*	NM_001145434	c.80C>T	missense	p.A27V	0.29
20	*BOD1*	NM_001159651	c.315G>A	silent	p.T105T	0.11
20	*BOD1*	NM_001159651	c.343C>T	silent	p.L115L	0.11
20	*BOD1*	NM_001159651	c.330G>A	silent	p.Q110Q	0.11
20	*PCDHB11*	NM_018931	c.1487T>C	missense	p.L496P	0.15
20	*PTPRJ*	NM_001098503	c.722A>G	missense	p.E241G	0.33
21	*GAL3ST1*	NM_004861	c.837C>T	silent	p.N279N	0.57
21	**SETD2*	NM_014159	c.1717_1720delTTCT	frameshift	p.F573Vfs*5	0.18
22	*C4B*	NM_001002029	c.3214C>T	missense	p.R1072W	0.15
22	*POTEE*	NM_001083538	c.2738A>C	missense	p.K913T	0.13
24	*ALDOB*	NM_000035	c.385G>T	missense	p.D129Y	0.44
24	*ANKRD2*	NM_001129981	c.888C>T	silent	p.H296H	0.65
24	*EXTL3*	NM_001440	c.839G>A	missense	p.R280H	0.38
24	*HMCN1*	NM_031935	c.11378G>A	missense	p.R3793H	0.43
24	*KIF5C*	NM_004522	c.566C>A	missense	p.A189E	0.35
24	*RGAG4*	NM_001024455	c.1462T>C	missense	p.Y488H	0.96
24	*RYR2*	NM_001035	c.14137G>A	missense	p.V4713I	0.31
24	*SEMA4F*	NM_001271661	c.374G>A	missense	p.R125Q	0.52
24	**TP53*	NM_001126115	c.460G>A	missense	p.E154K	0.36
25	*ADAMTS12*	NM_030955	c.4046C>T	missense	p.A1349V	0.23
25	*CCSER1*	NM_001145065	c.2297G>A	missense	p.R766H	0.45
25	*CYR61*	NM_001554	c.213C>A	missense	p.D71E	0.44
25	*DGKD*	NM_003648	c.1977G>A	silent	p.P659P	0.98
25	*DNAH8*	NM_001206927	c.9080G>A	missense	p.R3027Q	0.31
25	*ERC2*	NM_015576	c.1376C>T	missense	p.T459I	0.66
25	*FAM120C*	NM_017848	c.2770G>A	missense	p.V924I	0.31
25	*FSIP2*	NM_173651	c.3637G>A	missense	p.V1213I	0.35
25	**IKZF1*	NM_001291840	c.830C>T	missense	p.S277L	0.46
25	*KCNA10*	NM_005549	c.724C>T	missense	p.R242W	0.48
25	**KRAS*	NM_004985	c.38G>A	missense	p.G13D	0.52
25	*KRTAP5-3*	NM_001012708	c.528C>T	silent	p.C176C	0.22
25	*MST1L*	NM_001271733	c.1377T>C	silent	p.C459C	0.34
25	*NBPF1*	NM_017940	c.2667-3delC	splice site	(exon 25)	0.14
25	*PDLIM5*	NM_001256428	c.75G>A	silent	p.S25S	0.45
25	*PITX1*	NM_002653	c.434G>A	missense	p.R145H	0.51
25	*SEC61A2*	NM_001142628	c.298A>T	missense	p.I100F	0.34
25	*SNAP25*	NM_003081	c.13G>A	missense	p.A5T	0.20
25	*TBC1D30*	NM_015279	c.223G>A	missense	p.D75N	0.52
25	*USP17L20*	NM_001256861	c.860C>T	missense	p.T287I	0.12
26	**JAK2*	NM_004972	c.2047A>G	missense	p.R683G	0.23
26	*KAT6A*	NM_001305878	c.368G>A	missense	p.R123H	0.15
26	*PTGES2*	NM_025072	c.560C>T	missense	p.T187I	0.73
26	*RFPL4A*	NM_001145014	c.388C>T	nonsense	p.Q130*	0.15
26	*SRL*	NM_001098814	c.75T>C	silent	p.D25D	0.42
26	*TEKT4*	NM_001286559	c.456A>G	silent	p.K152K	0.15
26	*TPM1*	NM_001018008	c.144C>T	silent	p.D48D	0.33
27	*CHMP4C*	NM_152284	c.344C>T	missense	p.A115V	0.30
27	*CMYA5*	NM_153610	c.895G>A	missense	p.V299I	0.56
27	*LAMB1*	NM_002291	c.936C>A	nonsense	p.C312*	0.48
27	*MUC4*	NM_018406	c.11523T>G	silent	p.L3841L	0.16
27	*NEIL1*	NM_001256552	c.876G>A	silent	p.P292P	0.50
27	*NKD1*	NM_033119	c.835G>A	missense	p.V279M	0.29
27	*OBSCN*	NM_001271223	c.14091G>A	silent	p.V4697V	0.42

Table S2 (*continued*)

https://dx.doi.org/10.21037/tp-22-512

Table S2 (*continued*)

UPN	Gene	Reference	Nucleotide change	Effect	Amino acid change	VAF
27	**PAX5*	NM_016734	c.C101G	missense	p.P34R	0.34
27	*SALL2*	NM_005407	c.2606C>T	missense	p.P869L	0.27
27	*TANC2*	NM_025185	c.4863C>T	silent	p.A1621A	0.33
28	*CLCN6*	NM_001256959	c.818G>A	missense	p.R273H	0.22
28	*HYDIN*	NM_001198542	c.2350T>C	missense	p.S784P	0.16
28	**IKZF1*	NM_001220768	c.265G>T	nonsense	p.G89*	0.77
28	*LUZP1*	NM_001142546	c.2401T>C	missense	p.S801P	0.44
28	*MGAM*	NM_004668	c.1841C>T	missense	p.T614I	0.39
28	*PGGT1B*	NM_005023	c.1025C>T	missense	p.P342L	0.42
28	*WASF3*	NM_006646	c.691G>A	missense	p.E231K	0.49
30	*DDX11*	NM_001257144	c.1221C>A	missense	p.S407R	0.11
30	*DLK1*	NM_003836	c.712G>A	missense	p.E238K	0.46
30	*FARP1*	NM_001286839	c.2093G>A	missense	p.R698Q	0.28
30	*FMO1*	NM_001282692	c.724C>T	missense	p.R242C	0.10
30	**KRAS*	NM_004985	c.38G>A	missense	p.G13D	0.21
30	*NARS*	NM_004539	c.279G>T	missense	p.K93N	0.34
30	*NEBL*	NM_006393	c.430G>C	missense	p.E144Q	0.25
30	**NRAS*	NM_002524	c.35G>C	missense	p.G12A	0.07
30	*RPTOR*	NM_001163034	c.420C>T	silent	p.N140N	0.18
30	*SDK1*	NM_152744	c.1345C>T	missense	p.R449C	0.25
30	*ZNF208*	NM_007153	c.3480G>T	missense	p.K1160N	0.37
31	*ACAN*	NM_001135	c.823C>T	missense	p.R275W	0.61
31	*ADGRV1*	NM_032119	c.9314G>A	missense	p.R3105Q	0.45
31	*ATP12A*	NM_001185085	c.563G>A	missense	p.R188Q	0.36
31	*BIN3*	NM_018688	c.260C>T	missense	p.T87M	0.10
31	*C6*	NM_000065	c.2222C>T	missense	p.P741L	0.44
31	*CYP4F22*	NM_173483	c.889G>T	missense	p.A297S	0.42
31	*DNAH3*	NM_017539	c.1912T>C	missense	p.F638L	0.44
31	*DSE*	NM_001080976	c.449C>T	missense	p.P150L	0.45
31	*FBN1*	NM_000138	c.3026C>T	missense	p.P1009L	0.49
31	*FLNC*	NM_001127487	c.6708C>A	silent	p.G2236G	0.41
31	*HEATR1*	NM_018072	c.5259C>T	silent	p.S1753S	0.46
31	*IFNL2*	NM_172138	c.359T>G	missense	p.V120G	0.29
31	*KCNA10*	NM_005549	c.640G>A	missense	p.A214T	0.33
31	*KMT5A*	NM_020382	c.42_47delGGCGGC	in-frame	p.14_16del	1.00
31	*LAMA1*	NM_005559	c.4723G>A	missense	p.V1575I	0.53
31	*MB21D1*	NM_138441	c.162C>T	silent	p.A54A	0.44
31	*NOL4L*	NM_080616	c.180G>A	silent	p.T60T	0.48
31	**NOTCH1*	NM_017617	c.7205_7206insGGGCGCTT	frameshift	p.I2402Mfs*23	0.39
31	*NPFFR2*	NM_004885	c.55G>A	missense	p.V19I	0.38
31	*NUP205*	NM_015135	c.4600C>T	missense	p.R1534C	0.43
31	*NWD1*	NM_001007525	c.865C>A	missense	p.Q289K	0.47
31	*OTX2*	NM_001270525	c.500C>T	missense	p.P167L	0.59
31	*OXGR1*	NM_080818	c.803G>A	missense	p.R268H	0.40
31	*POFUT2*	NM_015227	c.301C>T	missense	p.R101W	0.43
31	*ROR1*	NM_001083592	c.554G>T	missense	p.R185L	0.50
31	**RUNX1*	NM_001001890	c.415C>T	nonsense	p.R139*	0.53
31	**SUZ12*	NM_015355	c.856C>T	nonsense	p.R286*	0.43
31	**SUZ12*	NM_015355	c.758G>C	missense	p.R253T	0.44
31	*TLL1*	NM_001204760	c.1085C>A	missense	p.S362Y	0.49
31	*TMEM132D*	NM_133448	c.242C>A	nonsense	p.S81*	0.53
38	*ANKFN1*	NM_153228	c.1170T>C	silent	p.G390G	0.23
38	**FLT3*	NM_004119	c.2503G>T	missense	p.D835Y	0.25
38	**MYC*	NM_002467	c.218C>A	missense	p.T73N	0.27
38	*TECTA*	NM_005422	c.1332C>T	silent	p.Y444Y	0.16
39	*DYNAP*	NM_001307955	c.218T>C	missense	p.M73T	0.19
40	*CHRM2*	NM_001006629	c.1070A>T	missense	p.K357M	0.32
40	*CNBP*	NM_001127192	c.237C>T	silent	p.C79C	0.43
40	*ERG*	NM_001136155	c.278_279insTGCGGG	in-frame	p.A93delinsAAG	0.30
40	*FAM229A*	NM_001167676	c.282-1insATTTCCCCA	splice site	(exon 3)	0.42
40	*FAM57A*	NM_024792	c.660C>G	missense	p.F220L	0.39
40	*GCLC*	NM_001498	c.447-1G>C	splice site	(exon 4)	0.55
40	*IRF5*	NM_001098629	c.490C>T	nonsense	p.Q164*	0.54
40	*KIF2B*	NM_032559	c.338C>T	missense	p.T113M	0.45
40	*KLHL18*	NM_025010	c.394C>T	nonsense	p.R132*	0.15
40	*PABPN1*	NM_004643	c.866G>A	missense	p.R289Q	0.49
40	*PAK7*	NM_177990	c.1463G>A	missense	p.R488Q	0.32
40	*TNPO3*	NM_001191028	c.1870-1G>T	splice site	(exon 16)	0.34
40	*TNPO3*	NM_001191028	c.138C>A	silent	p.I46I	0.38
40	**U2AF1*	NM_001025203	c.101C>T	missense	p.S34F	0.37

Table S2 (*continued*)

https://dx.doi.org/10.21037/tp-22-512

Table S2 (*continued*)

UPN	Gene	Reference	Nucleotide change	Effect	Amino acid change	VAF
40	*ZNF462*	NM_021224	c.676C>T	missense	p.R226C	0.48
40	*ZNF770*	NM_014106	c.2024_2031delACTTTAAA	frameshift	p.H675Rfs*18	0.26
41	**BRAF*	NM_004333	c.1803A>T	missense	p.K601N	0.33
41	*CCDC120*	NM_001163321	c.1640G>A	missense	p.R547H	0.26
41	*CCDC88C*	NM_001080414	c.4808G>A	missense	p.S1603N	0.55
41	*FAM47A*	NM_203408	c.1494T>A	silent	p.T498T	0.17
41	*GNB1*	NM_001282539	c.239T>A	missense	p.I80N	0.24
41	*HS3ST6*	NM_001009606	c.860C>A	missense	p.P287H	0.25
41	**KRAS*	NM_004985	c.436G>A	missense	p.A146T	0.16
41	*MTOR*	NM_004958	c.4377_4378insTCC	in-frame	p.L1460delinsSL	0.25
41	*MUC4*	NM_018406	c.10400C>A	missense	p.T3467K	0.10
41	*MUC4*	NM_018406	c.10387T>A	missense	p.S3463T	0.11
41	*MYO9A*	NM_006901	c.455G>A	missense	p.C152Y	0.15
41	**PAX5*	NM_016734	c.T404C	missense	p.I135T	0.80
41	*PLEKHG2*	NM_022835	c.1099G>A	missense	p.V367M	0.35
41	*RAI2*	NM_021785	c.328G>A	missense	p.A110T	0.39
41	*TBX22*	NM_001303475	c.48G>A	silent	p.K16K	0.45
41	*TDRD9*	NM_153046	c.936T>A	silent	p.I312I	0.39
41	**XBP1*	NM_005080	c.581dupT	frameshift	p.L194Ffs*190	0.38
41	*XIRP2*	NM_001199144	c.9224G>A	missense	p.R3075H	0.41
42	**CEBPA*	NM_001285829	c.579_580insCAG	in-frame	p.K194delinsQK	0.41
42	**CEBPA*	NM_004364	c.78delC	frameshift	p.S27Afs*133	0.47
42	*CREB5*	NM_001011666	c.417G>A	silent	p.P139P	0.42
42	*TAF1L*	NM_153809	c.1975C>G	missense	p.L659V	0.41
42	**WT1*	NM_001198552	c.458_459insGTACGGTCGGC	frameshift	p.S154Yfs*70	0.27
42	**WT1*	NM_001198552	c.539dupA	frameshift	p.M181Dfs*9	0.40
43	*ACTR8*	NM_022899	c.937G>C	missense	p.D313H	0.31
43	*CCDC40*	NM_001243342	c.2895A>G	silent	p.A965A	0.12
43	*COBLL1*	NM_001278461	c.1860T>C	silent	p.H620H	0.30
43	*ESYT3*	NM_031913	c.2201C>G	missense	p.S734C	0.37
43	*EYA1*	NM_172059	c.1302C>T	silent	p.A434A	0.33
43	*HAPLN4*	NM_023002	c.1069C>T	missense	p.R357W	0.29
43	*KCNB2*	NM_004770	c.155C>T	missense	p.T52M	0.31
43	*SLTM*	NM_001013843	c.400G>A	missense	p.E134K	0.57
43	**WHSC1*	NM_001042424	c.3295G>A	missense	p.E1099K	0.41
44	*ACACB*	NM_001093	c.3105G>A	silent	p.P1035P	0.53
44	*AKNAD1*	NM_152763	c.866C>T	missense	p.S289F	0.28
44	*ARNT2*	NM_014862	c.379G>A	missense	p.A127T	0.41
44	*CDH12*	NM_001317227	c.2068C>T	missense	p.R690C	0.52
44	*CDH2*	NM_001308176	c.1625C>T	missense	p.A542V	0.54
44	**CDKN2A*	NM_000077	c.181_182insCGG	in-frame	p.E61delinsAE	0.48
44	**CDKN2A*	NM_000077	c.172_173insT	frameshift	p.R58Lfs*62	0.54
44	**CHD4*	NM_001297553	c.3259G>A	missense	p.E1087K	0.53
44	*CTNND2*	NM_001288716	c.1792C>T	nonsense	p.R598*	0.41
44	*CYFIP1*	NM_014608	c.351T>A	silent	p.P117P	0.50
44	*PCDH11X*	NM_001168362	c.3833A>C	missense	p.D1278A	0.27
44	*PCDH15*	NM_001142765	c.3320T>C	missense	p.V1107A	0.27
44	*RAPH1*	NM_203365	c.329G>A	missense	p.R110H	0.29
44	*STRA6*	NM_001142617	c.1963C>T	missense	p.R655C	0.40
44	*TMEM11*	NM_003876	c.240C>T	silent	p.C80C	0.42
44	*TTN*	NM_133378	c.27338A>G	missense	p.H9113R	0.26
44	*UNC80*	NM_032504	c.3333C>A	missense	p.D1111E	0.36
44	*USP54*	NM_152586	c.602G>A	missense	p.R201Q	0.25
44	*ZNF275*	NM_001080485	c.637G>T	nonsense	p.E213*	0.22
46	*C17orf74*	NM_175734	c.573G>A	silent	p.L191L	0.41
46	*CPXM2*	NM_198148	c.1697G>A	missense	p.R566Q	0.45
46	*CRAMP1*	NM_020825	c.2627G>A	missense	p.R876Q	0.33
46	*DNAH8*	NM_001206927	c.8324G>A	missense	p.G2775E	0.42
46	*EGFR*	NM_005228	c.1939G>A	missense	p.A647T	0.39
46	*GDNF*	NM_000514	c.100delG	frameshift	p. E34Kfs*14	0.51
46	*HRNR*	NM_001009931	c.6330C>T	silent	p.H2110H	0.28
46	*NELL1*	NM_001288713	c.75C>T	silent	p.P25P	0.35
46	*PDE2A*	NM_001243784	c.2553-5C>T	splice site	(exon 31)	0.67
46	**PTEN*	NM_000314	c.697_699delinsTA	frameshift	p.R233Yfs*23	0.54
47	*ADAM21*	NM_003813	c.1252G>T	nonsense	p.E418*	0.32
47	**IKZF3*	NM_001257408	c.162_163insGAATA	frameshift	p. D55Efs*2	0.23
47	*KCNH5*	NM_139318	c.605C>T	missense	p.T202M	0.61
47	*KCTD4*	NM_198404	c.569C>T	missense	p.S190L	0.48
47	*OR5M10*	NM_001004741	c.516T>G	silent	p.L172L	0.44
47	*STEAP3*	NM_138637	c.1272G>A	silent	p.P424P	0.45

Table S2 (*continued*)

https://dx.doi.org/10.21037/tp-22-512

Table S2 (*continued*)

UPN	Gene	Reference	Nucleotide change	Effect	Amino acid change	VAF
48	*DMRTB1*	NM_033067	c.751_751delG	frameshift	p.V251Cfs*59	0.40
48	*FAT3*	NM_001008781	c.1670G>T	missense	p.R557L	0.35
48	*GIMD1*	NM_001195138	c.249G>A	silent	p.L83L	0.42
48	*LCE1E*	NM_178353	c.174C>A	silent	p.G58G	0.12
48	*LMNTD1*	NM_001145728	c.33G>C	silent	p.S11S	0.43
48	*MDGA2*	NM_001113498	c.1229C>T	missense	p.T410M	0.49
48	**WT1*	NM_001198552	c.134_144delGTGAGCAGCAG	frameshift	p.G45Vfs*32	0.74
48	*ZNFX1*	NM_021035	c.1461T>G	missense	p.N487K	0.10
49	*CORO2B*	NM_001190456	c.807G>C	silent	p.L269L	0.47
49	*CPAMD8*	NM_015692	c.3697C>T	missense	p.R1233C	0.17
49	*F11*	NM_000128	c.1679G>T	missense	p.C560F	0.35
49	*FAM188B*	NM_032222	c.1143G>A	silent	p.E381E	0.34
49	*FGFR1*	NM_001174066	c.1091G>A	missense	p.G364E	0.42
49	*NCAPH*	NM_001281712	c.1443G>A	silent	p.G481G	0.44
49	**NRAS*	NM_002524	c.38G>A	missense	p.G13D	0.21
49	*OR10K1*	NM_001004473	c.847C>T	missense	p.P283S	0.44
49	*OR11H2*	NM_001197287	c.827G>T	missense	p.S276I	0.19
49	*OSBPL11*	NM_022776	c.1984A>T	nonsense	p.R662*	0.42
49	*PMS2*	NM_000535	c.1146T>C	silent	p.G382G	0.18
49	*RASGRP2*	NM_001098670	c.1075G>A	missense	p.D359N	0.47
49	*SCN1A*	NM_001165963	c.1285C>T	nonsense	p.Q429*	0.13
49	*SLC7A7*	NM_001126105	c.248C>T	missense	p.S83F	0.18
49	*SPATA6*	NM_001286239	c.942G>A	silent	p.S314S	0.47
51	*AMDHD1*	NM_152435	c.798G>A	silent	p.P266P	0.44
51	*CSMD3*	NM_052900	c.6917A>G	missense	p.D2306G	0.37
51	*DOK3*	NM_001144876	c.406G>C	missense	p.G136R	0.41
51	*SMAD9*	NM_001127217	c.185C>T	missense	p.P62L	0.40
51	*UNC13C*	NM_001080534	c.3409G>A	missense	p.E1137K	0.35
51	*ZNF181*	NM_001029997	c.1352A>C	missense	p.H451P	0.15
53	*AASDH*	NM_001286668	c.974C>G	missense	p.A325G	0.26
53	*BLM*	NM_000057	c.872_873insTGA	in-frame	p.F291delinsFD	0.47
53	*CACNA1B*	NM_000718	c.2944C>T	missense	p.R982W	0.39
53	*CACNA1G*	NM_001256332	c.3317G>A	missense	p.R1106Q	0.40
53	*DSG3*	NM_001944	c.276C>T	silent	p.I92I	0.37
53	*GRIK4*	NM_001282470	c.1396C>T	missense	p.R466C	0.63
53	*GRIN2B*	NM_000834	c.3957G>A	silent	p.P1319P	0.43
53	*KCTD14*	NM_023930	c.89C>T	missense	p.T30M	0.43
53	*LOC100129307*	NM_001310140	c.717C>G	silent	p.V239V	0.25
53	*LY6D*	NM_003695	c.319G>A	missense	p.A107T	0.39
53	*MBTPS1*	NM_003791	c.569C>T	missense	p.P190L	0.13
53	*PRAMEF1*	NM_023013	c.558C>G	silent	p.V186V	0.16
53	*PRAMEF1*	NM_023013	c.560A>G	missense	p.N187S	0.16
53	*PTPRF*	NM_002840	c.1962C>T	silent	p.R654R	0.35
53	*QRFPR*	NM_198179	c.157G>A	missense	p.V53M	0.36
53	*SOX3*	NM_005634	c.527A>C	missense	p.D176A	0.29
53	*STX11*	NM_003764	c.338C>T	missense	p.A113V	0.24
53	*VSX2*	NM_182894	c.810C>T	silent	p.P270P	0.43
56	*FAM135A*	NM_001162529	c.923C>T	missense	p.A308V	0.25
56	*FGD1*	NM_004463	c.2697G>C	missense	p.W899C	0.20
56	*FGD2*	NM_173558	c.368T>C	missense	p.L123P	0.31
56	*LANCL3*	NM_001170331	c.1028T>C	missense	p.V343A	0.36
56	*MYO9B*	NM_001130065	c.5433G>A	silent	p.L1811L	0.46
56	*NETO1*	NM_001201465	c.620G>A	missense	p.R207Q	0.31
56	*OR12D2*	NM_013936	c.109G>T	missense	p.V37L	0.33
56	*PCDHGA12*	NM_003735	c.1437C>T	silent	p.P479P	0.44
56	**PTPN11*	NM_002834	c.218C>T	missense	p.T73I	0.53
56	**WHSC1*	NM_001042424	c.3448A>G	missense	p.T1150A	0.26
56	*ZNF41*	NM_007130	c.550C>A	missense	p.P184T	0.21
57	*HIST1H2AG*	NM_021064	c.72C>G	silent	p.L24L	0.28
57	**KRAS*	NM_004985	c.38G>A	missense	p.G13D	0.06
57	*MUC4*	NM_018406	c.10707T>G	silent	p.L3569L	0.10
57	*NLGN1*	NM_014932	c.1052G>A	missense	p.R351Q	0.43
57	**NRAS*	NM_002524	c.181C>A	missense	p.Q61K	0.34
57	*NUDT15*	NM_001304745	c.279T>C	silent	p.V93V	0.37
57	*OR8D1*	NM_001002917	c.379T>C	missense	p.C127R	0.30
57	*SERPINB11*	NM_001291279	c.269C>T	missense	p.S90L	0.30
57	*TBC1D30*	NM_015279	c.1920G>A	silent	p.P640P	0.28
57	*TNXB*	NM_019105	c.5495C>T	missense	p.P1832L	0.29
58	*ANKRD27*	NM_032139	c.3098C>T	missense	p.P1033L	0.40
58	*CWC27*	NM_001297645	c.1113G>A	silent	p.T371T	0.36

Table S2 (*continued*)

https://dx.doi.org/10.21037/tp-22-512

Table S2 (*continued*)

UPN	Gene	Reference	Nucleotide change	Effect	Amino acid change	VAF
58	*EAF2*	NM_018456	c.106+1delGTG	splice site	(exon 1)	0.22
58	*FER1L6*	NM_001039112	c.2718C>T	silent	p.D906D	0.24
58	*HTR3A*	NM_213621	c.1129G>A	missense	p.V377M	0.39
58	*MEF2A*	NM_001130928	c.1087dupC	frameshift	p.Q365Afs*20	0.46
58	*MTERF1*	NM_001301134	c.233A>T	missense	p.H78L	0.38
58	**NRAS*	NM_002524	c.38G>A	missense	p.G13D	0.33
58	*ORMDL1*	NM_001128150	c.344dupT	frameshift	p. Y116Lfs*5	0.35
58	*PLCG1*	NM_002660	c.2231_2232insCCGACC	in-frame	p.H744delinsHRP	0.17
58	*SETBP1*	NM_001130110	c.170C>T	missense	p.P57L	0.39
58	*SHANK2*	NM_133266	c.2099G>A	missense	p.R700Q	0.42
58	*TMEM132E*	NM_001304438	c.1257C>T	silent	p.G419G	0.50
58	*VGF*	NM_003378	c.1250A>G	missense	p.D417G	0.18
58	*XPR1*	NM_001135669	c.1769G>A	missense	p.R590H	0.20
58	*ZNF541*	NM_001277075	c.3158+1G>A	splice site	(exon 8)	0.36
59	*PCDH8*	NM_002590	c.2833C>A	missense	p.Q945K	0.12
59	**WT1*	NM_000378	c.1091C>A	nonsense	p.S364*	0.51
61	*DST*	NM_015548	c.9237A>G	silent	p.T3079T	0.25
61	*FAM81A*	NM_152450	c.1075C>T	nonsense	p.Q359*	0.25
61	*GRM8*	NM_000845	c.2186G>A	missense	p.R729Q	0.46
61	*HNRNPM*	NM_001297418	c.880_921del	in-frame	p.294_307del	0.23
61	*ITPRIP*	NM_001272012	c.166G>T	nonsense	p.E56*	0.30
61	*KRI1*	NM_023008	c.1600G>A	missense	p.V534M	0.16
61	*LOC100129697*	NM_001290330	c.912C>T	silent	p.H304H	0.12
61	*LOC100129697*	NM_001290330	c.915A>G	silent	p.R305R	0.13
61	*LY75-CD302*	NM_001198759	c.757G>C	missense	p.D253H	0.18
61	*MACC1*	NM_182762	c.636C>T	silent	p.V212V	0.20
61	*PRR32*	NM_001122716	c.403G>C	missense	p.G135R	0.48
61	*SALL1*	NM_001127892	c.384C>A	silent	p.A128A	0.30
61	*SPATA7*	NM_001040428	c.1423C>G	missense	p.Q475E	0.19
61	*TUBA3C*	NM_006001	c.727C>T	nonsense	p.R243*	0.37
61	*WNK2*	NM_001282394	c.2463G>A	silent	p.P821P	0.13
61	*ZMYM2*	NM_001190965	c.566C>A	missense	p.T189N	0.19
62	*C15orf59*	NM_001039614	c.62A>G	missense	p.E21G	0.11
62	*FAM83H*	NM_198488	c.1918G>A	missense	p.V640I	0.54
62	*RBBP6*	NM_006910	c.1918G>A	missense	p.E640K	0.17
64	*ABHD2*	NM_152924	c.207G>A	silent	p.P69P	0.45
64	*CLDN18*	NM_001002026	c.216G>T	silent	p.L72L	0.43
64	*HMBOX1*	NM_001135726	c.325C>T	missense	p.P109S	0.63
64	*IGSF22*	NM_173588	c.315C>T	silent	p.G105G	0.36
64	*KCNH8*	NM_144633	c.1125C>T	silent	p.Y375Y	0.39
64	**KRAS*	NM_004985	c.35G>T	missense	p.G12V	0.36
64	*LRIG3*	NM_001136051	c.1963G>A	missense	p.V655I	0.33
64	*LZTS1*	NM_021020	c.825C>T	silent	p.G275G	0.37
64	*NCOA1*	NM_003743	c.809_810insTAAAATCATC	frameshift	p.S274*	0.44
64	*PTCHD1*	NM_173495	c.1417G>A	missense	p.E473K	0.32
64	*QPCTL*	NM_001163377	c.709_710insTCC	in-frame	p.F237delinsFL	0.42
64	*RGPD3*	NM_001144013	c.3683C>T	missense	p.A1228V	0.42
65	**NRAS*	NM_002524	c.38G>A	missense	p.G13D	0.32
65	*RNF39*	NM_025236	c.1187T>C	missense	p.L396P	0.22
66	*ASTN1*	NM_001286164	c.3249C>T	silent	p.D1083D	0.29
66	*DAAM2*	NM_001201427	c.388G>T	missense	p.V130L	0.27
66	*EFEMP1*	NM_001039349	c.159C>T	silent	p.D53D	0.31
66	*KCNU1*	NM_001031836	c.39C>T	silent	p.D13D	0.34
66	*OR2M3*	NM_001004689	c.107C>T	missense	p.S36L	0.33
66	*PDZRN3*	NM_001303139	c.459C>T	silent	p.N153N	0.37
66	*SPIN3*	NM_001010862	c.48G>A	silent	p.T16T	0.35
66	*SPTAN1*	NM_001130438	c.2943G>A	silent	p.K981K	0.25
67	*ABCC8*	NM_000352	c.4504T>C	missense	p.F1502L	0.36
67	*ABCC8*	NM_000352	c.4525G>T	missense	p.A1509S	0.36
67	*MTCH2*	NM_001317232	c.710T>C	missense	p.V237A	0.37
67	*MTCH2*	NM_001317232	c.683T>C	missense	p.F228S	0.42
67	*ZNF732*	NM_001137608	c.1080C>G	silent	p.P360P	0.43
68	*ADCY5*	NM_001199642	c.102C>G	silent	p.L34L	0.21
68	*AP5Z1*	NM_014855	c.970-4G>A	splice site	(exon 9)	0.24
68	*CCNA1*	NM_001111045	c.79G>A	missense	p.G27R	0.25
68	*CD109*	NM_001159588	c.2493C>T	silent	p.I831I	0.12
68	*COL19A1*	NM_001858	c.2553C>T	silent	p.G851G	0.40
68	*COL6A3*	NM_057166	c.5203C>T	nonsense	p.R1735*	0.18
68	*DNAH7*	NM_018897	c.10602G>C	missense	p.Q3534H	0.21
68	*FRMPD4*	NM_014728	c.3714G>A	silent	p.P1238P	0.74

Table S2 (*continued*)

https://dx.doi.org/10.21037/tp-22-512

Table S2 (*continued*)

UPN	Gene	Reference	Nucleotide change	Effect	Amino acid change	VAF
68	*GCDH*	NM_000159	c.1138G>A	missense	p.D380N	0.19
68	*KATNA1*	NM_001204076	c.892G>C	missense	p.E298Q	0.21
68	*KIAA2022*	NM_001008537	c.1837G>C	missense	p.E613Q	0.40
68	*KLF1*	NM_006563	c.193G>A	missense	p.D65N	0.20
68	*KSR1*	NM_014238	c.1833C>T	silent	p.I611I	0.16
68	*MYCT1*	NM_025107	c.426C>G	silent	p.L142L	0.23
68	*NOL4*	NM_001198549	c.1022C>T	missense	p.S341F	0.17
68	*OR51I2*	NM_001004754	c.552G>A	missense	p.M184I	0.12
68	*PDS5B*	NM_015032	c.4151C>T	missense	p.P1384L	0.25
68	*PRKD1*	NM_002742	c.418G>C	missense	p.E140Q	0.21
68	*RUNX3*	NM_004350	c.879C>T	silent	p.S293S	0.30
68	*SI*	NM_001041	c.317G>C	missense	p.C106S	0.19
68	*SLITRK4*	NM_001184749	c.1133A>G	missense	p.N378S	0.18
68	*SORBS2*	NM_001145674	c.1122C>G	missense	p.I374M	0.13
68	*SPIRE1*	NM_001128627	c.1711T>G	missense	p.C571G	0.15
68	*STRADB*	NM_001206864	c.7C>G	missense	p.L3V	0.27
68	*SYCE2*	NM_001105578	c.273C>G	silent	p.L91L	0.15
68	*TINF2*	NM_001099274	c.1222G>C	missense	p.E408Q	0.19
68	*TTN*	NM_133379	c.13795G>A	missense	p.E4599K	0.21
68	*TTN*	NM_133379	c.13427G>C	missense	p.R4476T	0.21
68	*TTN*	NM_001256850	c.1100C>G	missense	p.S367C	0.24
68	*VPS13D*	NM_015378	c.2401G>A	missense	p.E801K	0.28
68	*VWF*	NM_000552	c.1616C>T	missense	p.S539F	0.11
68	*VWF*	NM_000552	c.1920C>A	silent	p.V640V	0.22
68	*ZNF709*	NM_152601	c.1249G>C	missense	p.E417Q	0.14
69	*BAIAP2L1*	NM_018842	c.666G>A	silent	p.L222L	0.46
69	*DCPS*	NM_014026	c.454C>T	nonsense	p.R152*	0.47
69	*DDX54*	NM_001111322	c.1283G>A	missense	p.R428H	0.43
69	*FAT3*	NM_001008781	c.12541C>T	missense	p.R4181C	0.50
69	*HIST2H2AB*	NM_175065	c.45C>A	silent	p.A15A	0.16
69	*RNF224*	NM_001190228	c.119G>A	missense	p.R40H	0.55
69	*SNHG32*	NM_001040438	c.163dupA	frameshift	p.N55Kfs*20	0.62
69	*SPEG*	NM_005876	c.578C>T	missense	p.T193M	0.57
69	*UBC*	NM_021009	c.632_633insAGGT	nonsense	p.Y211*	0.43
70	**FLT3*	NM_004119	c.1987A>C	missense	p.K663Q	0.25
70	*GABRB3*	NM_001191320	c.567T>G	silent	p.A189A	0.18
70	*GATAD2A*	NM_001300946	c.948G>C	silent	p.G316G	0.26
70	*IFI16*	NM_001206567	c.169C>T	nonsense	p.R57*	0.14
70	**IKZF1*	NM_001291840	c.424A>G	missense	p.N142D	0.17
70	*IWS1*	NM_017969	c.422G>A	missense	p.G141E	0.27
70	*KRT14*	NM_000526	c.978C>T	silent	p.S326S	0.34
70	*NCAM1*	NM_001076682	c.1810+1G>A	splice site	(exon 15)	0.12
70	**NRAS*	NM_002524	c.38G>A	missense	p.G13D	0.21
70	*OR13C8*	NM_001004483	c.762C>G	silent	p.T254T	0.19
70	*TJP1*	NM_003257	c.831C>T	silent	p.S277S	0.42
70	*UTY*	NM_001258265	c.233A>G	missense	p.Y78C	0.45
70	**WHSC1*	NM_001042424	c.3295G>A	missense	p.E1099K	0.13
73	*ALKBH7*	NM_032306	c.650C>T	missense	p.P217L	0.46
73	*ARHGEF33*	NM_001145451	c.1860C>T	silent	p.G620G	0.56
73	*C9orf170*	NM_001001709	c.340G>A	missense	p.V114M	0.55
73	*DHRS7C*	NM_001105571	c.454T>A	missense	p.F152I	0.24
73	*FABP1*	NM_001443	c.124G>A	missense	p.V42M	0.26
73	*FAM47A*	NM_203408	c.1616G>A	missense	p.R539Q	0.42
73	*HLA-DQA2*	NM_020056	c.208C>A	missense	p.Q70K	0.17
73	**IKZF3*	NM_001257408	c.96_99delCAAA	frameshift	p.K33Lfs*20	0.30
73	*ITGA2B*	NM_000419	c.1097G>A	missense	p.R366Q	0.39
73	*MUC17*	NM_001040105	c.3480T>C	silent	p.T1160T	0.40
73	*MUC4*	NM_018406	c.4530T>G	silent	p.P1510P	0.42
73	*NBEA*	NM_001204197	c.899G>A	missense	p.R300Q	0.27
73	*NOTCH3*	NM_000435	c.4884C>T	silent	p.D1628D	0.49
73	**NRAS*	NM_002524	c.38G>A	missense	p.G13D	0.43
73	*OR10Q1*	NM_001004471	c.290C>T	missense	p.S97L	0.45
73	*POM121L12*	NM_182595	c.475C>T	missense	p.R159C	0.45
73	*PRAMEF4*	NM_001009611	c.457G>A	missense	p.V153I	0.95
73	*PTPRT*	NM_007050	c.2679C>T	silent	p.Y893Y	0.32
73	*QRFPR*	NM_198179	c.732G>A	silent	p.K244K	0.23
73	*WNK3*	NM_020922	c.3812G>A	missense	p.R1271H	0.54
73	*WNT7A*	NM_004625	c.700G>A	missense	p.E234K	0.24
73	*ZNF804B*	NM_181646	c.1613C>T	missense	p.T538M	0.46
74	**NRAS*	NM_002524	c.38G>A	missense	p.G13D	0.37

Table S2 (*continued*)

https://dx.doi.org/10.21037/tp-22-512

Table S2 (*continued*)

UPN	Gene	Reference	Nucleotide change	Effect	Amino acid change	VAF
74	*VWA2*	NM_001272046	c.1369_1370insCC	frameshift	p.E458Pfs*22	0.39
75	*DYNAP*	NM_001307955	c.218T>C	missense	p.M73T	0.17
75	**NRAS*	NM_002524	c.183A>T	missense	p.Q61H	0.16
75	*POTEH*	NM_001136213	c.484_510del	in-frame	p.162_170del	0.40
76	*DHX38*	NM_014003	c.1051C>T	missense	p.R351W	0.48
76	*FAM84A*	NM_145175	c.414C>T	silent	p.P138P	0.44
76	*HFE*	NM_000410	c.173T>A	missense	p.F58Y	0.39
76	**NRAS*	NM_002524	c.35G>C	missense	p.G12A	0.20
76	*PGLYRP2*	NM_052890	c.697C>T	nonsense	p.R233*	0.32
76	*REPIN1*	NM_014374	c.969C>T	silent	p.A323A	0.67
76	*SPATA31E1*	NM_178828	c.3234G>A	silent	p.A1078A	0.36
76	*UBE2J1*	NM_016021	c.660delT	frameshift	p.A221Lfs*24	0.40
80	*ACSS3*	NM_024560	c.1783G>T	missense	p.G595C	0.45
80	*DUS3L*	NM_020175	c.849G>T	silent	p.G283G	0.38
80	*FAM205A*	NM_001141917	c.3797C>T	missense	p.T1266M	0.45
80	*FRMPD3*	NM_032428	c.2344C>A	missense	p.P782T	0.35
80	*GNAS*	NM_016592	c.258C>T	silent	p.H86H	0.45
80	**JAK2*	NM_004972	c.2047A>G	missense	p.R683G	0.34
80	*NLGN1*	NM_014932	c.1171G>T	nonsense	p.E391*	0.52
80	**NRAS*	NM_002524	c.201_202insGGAACC	in-frame	p.R68delinsGTR	0.33
80	*OR51E1*	NM_152430	c.813C>A	missense	p.D271E	0.39
80	*PALD1*	NM_014431	c.1687C>T	missense	p.R563W	0.38
80	*PCDHGA1*	NM_018912	c.1910C>T	missense	p.A637V	0.44
80	*PRSS54*	NM_001080492	c.63C>T	silent	p.L21L	0.51
80	*PSG9*	NM_001301707	c.549C>T	silent	p.N183N	0.44
80	*RBM45*	NM_152945	c.744G>A	silent	p.L248L	0.53
80	*TNR*	NM_003285	c.2935G>A	missense	p.E979K	0.39
80	*TRPM8*	NM_024080	c.2946G>A	silent	p.T982T	0.49
80	*TTN*	NM_003319	c.48729A>T	silent	p.P16243P	0.13
80	*ZMYM1*	NM_001289089	c.1437C>T	silent	p.H479H	0.46
80	*ZNF385D*	NM_024697	c.605G>A	missense	p.R202Q	0.36
82	*CTBP2*	NM_022802	c.2272G>T	nonsense	p.E758*	0.14
82	*SORL1*	NM_003105	c.5828C>T	missense	p.T1943M	0.20
84	*ACE*	NM_000789	c.1143G>A	silent	p.T381T	0.31
84	**ARID5B*	NM_032199	c.137dupG	frameshift	p. C46Wfs*29	0.41
84	*BRINP3*	NM_001317188	c.14C>A	missense	p.P5H	0.52
84	*C22orf29*	NM_024627	c.209G>A	missense	p.G70D	0.32
84	*CSMD3*	NM_052900	c.8558C>A	missense	p.T2853K	0.56
84	*EEF1A2*	NM_001958	c.912C>T	silent	p.P304P	0.41
84	**FLT3*	NM_004119	c.2039C>T	missense	p.A680V	0.48
84	*INADL*	NM_176877	c.1952G>A	missense	p.R651H	0.54
84	*KIR3DL3*	NM_153443	c.620C>T	missense	p.S207L	0.44
84	*LAS1L*	NM_001170650	c.317C>T	missense	p.P106L	0.32
84	*MPDZ*	NM_001261406	c.3382C>T	nonsense	p.R1128*	0.38
84	**PTPN11*	NM_002834	c.226G>A	missense	p.E76K	0.41
84	*RFC5*	NM_001206801	c.772G>A	missense	p.D258N	0.38
84	*RIPK4*	NM_020639	c.1855G>A	missense	p.V619M	0.24
84	*RPS6KA2*	NM_021135	c.1441T>G	missense	p.F481V	0.29
84	*SNX29*	NM_032167	c.2320G>A	missense	p.D774N	0.41
84	*UBE3C*	NM_014671	c.1552G>A	missense	p.E518K	0.50
84	*ZNF626*	NM_001076675	c.1046C>A	missense	p.A349D	0.13
85	*ADAMTS8*	NM_007037	c.1730C>T	missense	p.T577M	0.42
85	*ANKRD45*	NM_198493	c.311A>G	missense	p.N104S	0.54
85	*ASZ1*	NM_001301821	c.101C>T	missense	p.S34F	0.54
85	*CUL4A*	NM_001278513	c.1732-2A>G	splice site	(exon 19)	0.62
85	*FSIP2*	NM_173651	c.14275T>C	missense	p.S4759P	0.55
85	*KMT2D*	NM_003482	c.16599G>A	silent	p.R5533R	0.57
85	**NOTCH1*	NM_017617	c.7020dupC	frameshift	p.S2341Lfs*13	0.51
85	**RPL10*	NM_001256580	c.184C>A	missense	p.R62S	0.95
85	*TAL1*	NM_001290406	c.86G>A	missense	p.R29Q	0.48
85	*TBC1D10A*	NM_001204240	c.1449_1463del	in-frame	p.483_488delKDSAP	0.46
85	*THBS1*	NM_003246	c.2115C>T	silent	p.C705C	0.50
85	*TRMT5*	NM_020810	c.1118G>C	missense	p.G373A	0.51
87	*CHRM2*	NM_001006629	c.99C>T	silent	p.L33L	0.36
87	*DPYSL4*	NM_006426	c.747G>A	silent	p.P249P	0.47
87	*HBB*	NM_000518	c.76G>C	missense	p.G26R	0.11
87	*HBB*	NM_000518	c.84C>A	silent	p.A28A	0.14
87	**KRAS*	NM_004985	c.35G>T	missense	p.G12V	0.04
89	*CKAP5*	NM_001008938	c.2688G>A	silent	p.P896P	0.45
89	*GRHL3*	NM_001195010	c.1095C>T	silent	p.D365D	0.39

Table S2 (*continued*)

https://dx.doi.org/10.21037/tp-22-512

Table S2 (*continued*)

UPN	Gene	Reference	Nucleotide change	Effect	Amino acid change	VAF
93	*FBXL8*	NM_018378	c.937T>G	missense	p.S313A	0.96
93	*IRS1*	NM_005544	c.1021T>C	missense	p.S341P	0.67
93	*KRTAP10-2*	NM_198693	c.233C>T	missense	p.S78L	0.37
93	*MTUS1*	NM_001001924	c.2216A>T	missense	p.N739I	0.32
93	*SHISA7*	NM_001145176	c.568T>C	missense	p.C190R	0.63
93	*SRRM3*	NM_001291831	c.1161G>C	silent	p.R387R	1.00
93	*ZFHX4*	NM_024721	c.1870T>C	missense	p.S624P	0.53
93	*ZXDB*	NM_007157	c.732G>A	silent	p.A244A	1.00
94	**FLT3*	NM_004119	c.2503G>T	missense	p.D835Y	0.63
94	*MBLAC2*	NM_203406	c.568G>A	missense	p.V190I	0.59
94	*SKAP1*	NM_001075099	c.286G>A	missense	p.E96K	0.80
95	*APBB3*	NM_006051	c.1157A>G	missense	p.D386G	0.43
95	*DOCK5*	NM_024940	c.343C>T	missense	p.R115C	0.50
95	*DPH7*	NM_138778	c.741C>A	missense	p.S247R	0.47
95	*PIR*	NM_001018109	c.284C>T	missense	p.A95V	0.57
95	*ZHX1*	NM_001017926	c.581A>C	missense	p.K194T	0.48
99	*ATP11C*	NM_001010986	c.2788G>T	nonsense	p.E930*	0.88
99	*MYH7*	NM_000257	c.1324C>T	missense	p.R442C	0.37
99	*ROR1*	NM_005012	c.1387-3C>-	splice site	(exon 9)	0.51
101	*APOBEC3F*	NM_001006666	c.280G>A	missense	p.A94T	0.35
101	*CDC27*	NM_001114091	c.172T>C	missense	p.Y58H	0.56
101	*FAH*	NM_000137	c.782C>T	missense	p.P261L	0.50
101	*NHSL1*	NM_020464	c.2022G>T	missense	p.K674N	0.60
101	*ROBO3*	NM_022370	c.3412C>T	nonsense	p.R1138*	0.47
101	*UBE2D3*	NM_181893	c.406T>C	missense	p.Y136H	0.67

*, possible driver mutations. UPN, unique patient number; VAF, variant allele frequency.

Table S3 Fusion genes

UPN	Fusion gene	Breakpoint 1*	Breakpoint 2*
1	ETV6-RUNX1	chr12:12034688	chr21:36335734
19	ETV6-RUNX1	chr12:12037335	chr21:36297114
40	ETV6-RUNX1	chr12:12030912	chr21:36417670
43	ETV6-RUNX1	chr12:12034273	chr21:36260162
58	ETV6-RUNX1	chr12:12031200	chr21:36402962
61	ETV6-RUNX1	chr12:12023334	chr21:36419079
62	ETV6-RUNX1	chr12:12035898	chr21:36308613
68	ETV6-RUNX1	chr12:12035211	chr21:36265114
95	ETV6-RUNX1	chr12:12035696	chr21:36265894
2	TCF3-PBX1	chr1:164682489	chr19:1618817
7	TCF3-PBX1	chr1:164658024	chr19:1617928
9	TCF3-PBX1	chr1:164695245	chr19:1616862
20	TCF3-PBX1	chr1:164680606	chr19:1617927
21	TCF3-PBX1	chr1:164659227	chr19:1617944
27	TCF3-PBX1	chr1:164657580	chr19:1617931
47	TCF3-PBX1	chr1:164756478	chr19:1617932
87	TCF3-PBX1	chr1:164752679	chr19:1617071
89	TCF3-PBX1	chr1:164754008	chr19:1617926
92	TCF3-PBX1	chr1:164756367	chr19:1617926
93	TCF3-PBX1	chr1:164654805	chr19:1617931
26	P2RY8-CRLF2	chrX:1333754	chrX:1654735
80	P2RY8-CRLF2	chrX:1335073	chrX:1654734
101	P2RY8-CRLF2	chrX:1335077	chrX:1654914
28	BCR-ABL1	chr9:133678974	chr22:23577185
66	TCF3-HLF	chr17:53397769	chr19:1618340
17	MEF2D-BCL9	chr1:147095613	chr1:156445440
75	BCL2/IGH	chr14:106330842	chr18:60793550
99	PML-RARA	chr15:74326125	chr17:38494349

*, hg19 coordinate.

https://dx.doi.org/10.21037/tp-22-512

Table S4 B-ALL cases without known causative mutations

UPN	Method	Somatic point mutations on exome	Tumor content	Tumor contamination in germline samples	CNV	SV
65	WES	2	0.64	0.32	No significant finding	No significant finding
98	WGS	0	N/A	N/A	No significant finding	No significant finding

B-ALL, B-cell precursor acute lymphoblastic leukemia; UPN, unique patient number; CNV, copy number variations; SV, structural variations; WES, whole-exome sequencing + targeted sequencing; WGS, whole-genome sequencing; N/A, not available.

Table S5 B-ALL classification and concomitant somatic mutations including RAS signaling pathway mutations

Subtype	No. of patients	Percentages (%)	Average number of somatic mutations	RAS signaling pathway mutations	
				Number of cases	Number of mutated genes
HHD	12	24.5	14	9 [2]*	4 KRAS
					5 NRAS
					2 PTPN11
TCF3-PBX1	11	22.4	5.6	1	1 KRAS
ETV6-RUNX1	9	18.4	13.7	1	1 NRAS
PAX5alt	4	8.1	12	1 [1]*	1 KRAS
					1 BRAF
P2RY8-CRLF2	3	6.1	10.7	1	1 NRAS
del(11)(q23)	2	4.1	9	0	0
iAMP21	1	2	23	0	0
MEF2D-BCL9	1	2	13	0	0
BCR-ABL1	1	2	7	0	0
TCF3-HLF	1	2	8	0	0
Ph-like (FLT3)	1	2	13	1	1 NRAS
BCL2/IGH	1	2	3	1	1 NRAS
B-other-ALL	2	4.1	3	1	1 NRAS

*, number of cases with double mutations. B-ALL, B-cell precursor acute lymphoblastic leukemia; HHD, high hyperdiploidy.

https://dx.doi.org/10.21037/tp-22-512