Postnatal origin of the chromosomal gains in older patients with high hyperdiploid acute lymphoblastic leukemia

High hyperdiploidy (HeH), characterized by nonrandom chromosomal gains resulting in 51-67 chromosomes, is the most common genetic subtype of childhood B-cell precursor (BCP) acute lymphoblastic leukemia (ALL). HeH ALL typically occurs in children aged 2-4 years and overall corresponds to 25% of cases in the pediatric (<18 years) population, but, for reasons unknown, becomes less frequent with increasing age; in adult ALL it is relatively rare. Several studies have shown clear evidence of HeH frequently arising already before birth in pediatric cases.[1-9] However, there are very little data on whether this also holds true for HeH arising in older children, adolescents and adults.

We previously utilized somatic single nucleotide variants (SNV) to study the "age" of the trisomies, i.e., whether their origin was recent or not in terms of clonal evolution, in 16 cases of pediatric HeH ALL.[10] This was done by analyzing the mutant allele frequencies (MAF) of all somatic SNV in trisomic chromosomes to investigate whether they were present in one of two or two of three chromosomal homologues. Whereas the former could arise either before or after the chromosome became trisomic (B/ATRI SNV), the latter must have arisen before the trisomy was formed and become duplicated with the homologue (BTRI SNV) (Figure 1A). We found that BTRI SNV constituted only a small fraction of the SNV in most cases, suggesting that the chromosomal gains arose very early during clonal evolution and before the cell had acquired many passenger mutations, in line with a prenatal origin. However, the two adolescent patients in the study had much higher frequencies of BTRI SNV, indicating that older patients with HeH may differ from younger patients in this regard. Here, we have addressed this possibility further.

We ascertained whole genome sequencing (WGS) data from HeH cases from local biobanks (not included in Paulsson *et al.*[10] and selected based on material being available except for ten cases selected based on age >16 years) (N=31), from the Therapeutically Applicable Research to Generate Effective Treatments (TARGET) Initiative (https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000464.v21.p8) (N=33), and from the St. Jude Cloud[11-14] (N=127). HeH was defined as 51-67 chromosomes without concurrent *ETV6::RUNX1, TCF3::PBX1, BCR::ABL1*, or *KMT2A* rearrangements. In addition, four cases with concurrent *BCR::ABL1* and HeH were investigated (*Online Supplementary Table S1*). Analyses of BTRI data for 40 cases have previously been published.[15] Informed consent was obtained according to the Declaration of Helsinki and the study was approved by the Swedish Ethical Review Authority. All somatic SNV in clonal trisomies were ascertained and categorized as BTRI (MAF=~0.67) or B/ATRI SNV (MAF=~0.33) (Figure 1A), excluding somatic SNV with MAF<0.15. Cases with large subclones for trisomic chromosomes, making the distinction between BTRI and B/ATRI unclear (N=18), with less than 65% leukemic cells based on MAF (N=15), or with <100 (N=26) or >1,000 (N=10) somatic SNV in trisomic chromosomes, were excluded. The final HeH cohort included 117 pediatric (1-17 years at diagnosis; median age 4 years) and five adult cases; 99 were aged 1-9 years and 23 ≥10 years at diagnosis (25 from local biobanks, 25 from TARGET, and 72 from St. Jude Cloud) (*Online Supplementary Table S1*). There were no differences in age (P=0.27; Mann-Whitney two-sided test) or sex distribution (P=0.81; Fisher's exact test) in cases excluded (36% of initial cases) and those retained. Mann-Whitney two-sided test, Spearman correlation, and Fisher's exact test were used as indicated below and P values <0.05 were considered significant. All statistical analyses were done using R x64 v4.4.1 (https://www.r-project.org/).

The median numbers of somatic, BTRI and B/ATRI SNV were 277.5 (range, 101-842), eight (range, 0-79), and 265 (range, 98-825), respectively, in HeH cases. To adjust for possible differences related to read depths between the cohorts and to modal number, all subsequent analyses were done on the proportion of BTRI/all SNV in trisomic chromosomes, since this should not be affected by the total number of SNV.

The median BTRI/all SNV was 3.1% (range, 0-23%; *Online Supplementary Table S1*), indicating that the time period after the HeH arose was always longer than the time period before, in line with previous findings.[10] No sex differences were seen (Mann-Whitney two-sided test; P=0.0531). However, whereas the BTRI/all SNV percentages appeared to be constant in patients aged 1-9 years, they increased linearly in those ≥10 years at diagnosis (Spearman correlation coefficient=0.64, P=0.00092; Figure 1B, C). The latter cases also had a significantly higher percentage of BTRI SNV (median 2.6% in patients 1-9 years *vs*. 9.8% in patients ≥10 years; Mann-Whitney two-sided test, $P=5.26\times10^{-9}$; Figure 1D). That cases <10 years generally had the same proportion of BTRI SNV agrees well with a prenatal origin for the majority of these because their expected number of BTRI SNV would then be constant. In contrast, that cases ≥10 years show an increasing number of BTRI SNV by age suggests that the chromosomal gains generally arise postnatally at different time points in these cases, with the number of BTRI SNV being higher in older patients since the chromosomes that become duplicated will have acquired more passenger mutations with increasing

age. Thus, our results show that the chromosomal gains in HeH ALL arising in older patients - outside the "age peak" in early childhood – may not arise *in utero*.

Support for the hypothesis that pediatric HeH ALL may have a prenatal origin has previously been obtained by detection of clonotypic *IGH* rearrangements in Guthrie cards, observations of monozygotic twins with concurrent HeH ALL, and detection of trisomies in saved cord blood cells from children who later developed leukemia.[1-9] Whereas the latter two of these methods have provided unequivocal evidence of the HeH itself being present before birth in some cases, studies of clonotypic IGH rearrangements without analysis of chromosome 14 copy number can only provide evidence for the presence of a (pre)leukemic clone, not HeH itself. In total, we found 39 reported HeH cases where prenatal origin was investigated and the patient age was given (Table 1);[1-9] both positive (N=29) and negative (N=10) cases had a median age of 2 years at diagnosis. Two of three HeH cases older than 10 years at diagnosis had evidence of a preleukemic clone - albeit not conclusive for HeH since chromosomal copy numbers were not addressed - at birth.[7] This agrees with our findings because the proportion of BTRI/all SNV varied also in patients ≥10 years, with some having low levels of BTRI SNV and thereby possibly a prenatal origin of the chromosomal gains.

A second question arising from our findings is whether chromosomal gains arising postnatally in older patients are the primary event or if they occur as a secondary aberration to another leukemia-initiating lesion. We have previously shown that in BCP ALL with concurrent *BCR::ABL1*-fusion and HeH, the chromosomal gains are most likely secondary to the fusion gene.[16] Thus, such cases can serve as a model

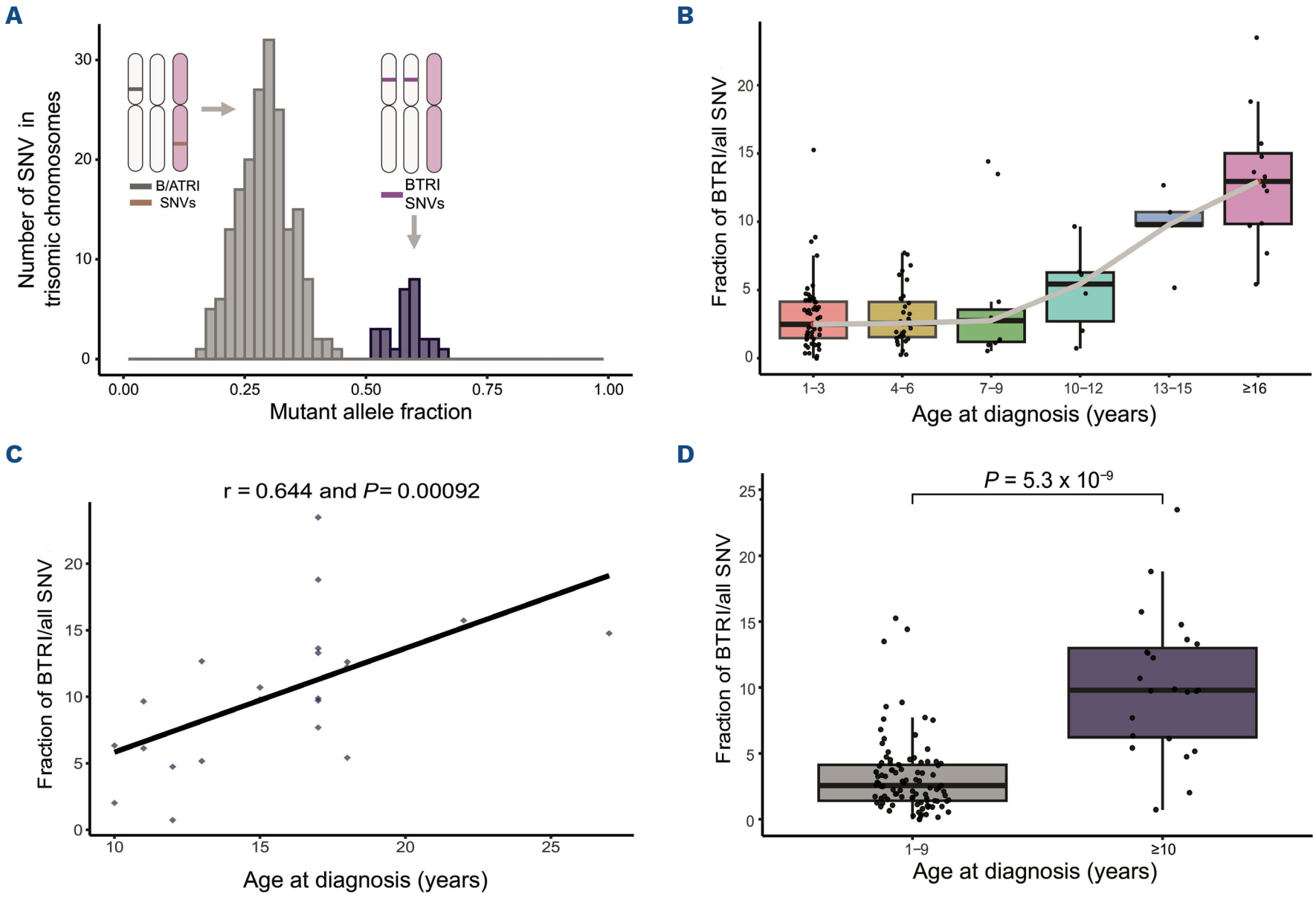

Figure 1. The fraction of BTRI mutations, arising before the chromosomal gains, is higher in older patients with high hyperdiploid acute lymphoblastic leukemia. (A) Example of somatic single nucleotide variants (SNV) in trisomic chromosomes from case L1, showing 2 distinct peaks corresponding to B/ATRI and BTRI SNV, respectively. (B) The fraction of BTRI/all SNV in relation to age is constant in cases 1-9 years old at diagnosis but increases with age in cases 10 years or older at diagnosis. The center of the boxplot is the median and lower/upper hinges correspond to the first/third quartiles; whiskers are 1.5-times the interquartile range and the individual data points are shown as black points. The median value is shown as red points for each group. (C) Spearman correlation shows that the fraction of BTRI/all SNV increases linearly with age in patients older than 10 years at diagnosis. (D) Boxplot of the fraction of BTRI/all SNV in patients 1-9 and ≥10 years old at diagnosis, respectively, showing a higher fraction in the older age group. The center of the boxplot is the median and lower/upper hinges correspond to the first/third quartiles; whiskers are 1.5-times the interquartile range and the individual data points are shown as black points.

Table 1. High hyperdiploid acute lymphoblastic leukemia reported in the literature investigated for prenatal origin and with age information available.

Original case N	Age at dx in years	Prenatal HeH	Method	Reference
1	1	Yes	Guthrie card-IGH no CN	Taub *et al.*[5]
T1	1	Yes	Twins and IGH	Maia *et al.*[2]
5	1	No	Guthrie card-IGH no CN	Maia *et al.*[6]
7	1	No	Guthrie card-IGH no CN	Maia *et al.*[6]
10	1	Yes	Guthrie card-IGH no CN	Maia *et al.*[6]
Detroit 1	1	Yes	Guthrie card-IGH no CN	Gruhn *et al.*[7]
T1A	1	Yes	Twins and IGH	Bateman *et al.*[3]
T1B	1	Yes	Twins and IGH	Bateman *et al.*[3]
4	2	Yes	Guthrie card-IGH no CN	Yagi *et al.*[4]
1	2	Yes	Guhtrie card-3 IGH	Panzer-Grümayer *et al.*[1]
3	2	Yes	Guthrie card-IGH no CN	Taub *et al.*[5]
6	2	Yes	Guthrie card-IGH no CN	Taub *et al.*[5]
T2	2	Yes	Twins and IGH	Maia *et al.*[2]
2	2	No	Guthrie card- IGH no CN	Maia *et al.*[6]
6	2	No	Guthrie card- IGH no CN	Maia *et al.*[6]
8	2	No	Guthrie card-IGH no CN	Maia *et al.*[6]
11	2	Yes	FISH on stored cord blood cells	Maia *et al.*[6]
Jena 10	2	Yes	Guthrie card-IGH no CN	Gruhn *et al.*[7]
Detroit 3	2	Yes	Guthrie card-IGH no CN	Gruhn *et al.*[7]
Detroit 4	2	Yes	Guthrie card-IGH no CN	Gruhn *et al.*[7]
T2A	2	Yes	Twins and IGH	Bateman *et al.*[3]
8	2	Yes	Guthrie card-IGH no CN	Kacanski *et al.*[8]
8	3	Yes	Guthrie card-IGH no CN	Taub *et al.*[5]
3	3	No	Guthrie card-IGH no CN	Maia *et al.*[6]
Jena 14	3	Yes	Guthrie card-IGH no CN	Gruhn *et al.*[7]
T2B	3	Yes	Twins and IGH	Bateman *et al.*[3]
1	3	Yes	FISH on stored cord blood cells	Zuna *et al.*[9]
1	4	No	Guthrie card-IGH no CN	Maia *et al.*[6]
9	4	No	Guthrie card-IGH no CN	Maia *et al.*[6]
12	5	Yes	Guthrie card-IGH no CN	Taub *et al.*[5]
Jena 20	5	Yes	Guthrie card-IGH no CN	Gruhn *et al.*[7]
Detroit 12	5	Yes	Guthrie card-IGH no CN	Gruhn *et al.*[7]
T3A	5	Yes	Twins and IGH	Bateman *et al.*[3]
Jena 26	8	No	Guthrie card-IGH no CN	Gruhn *et al.*[7]
14	9	Yes	Guthrie card-IGH no CN	Taub *et al.*[5]
Detroit 14	9	Yes	Guthrie card-IGH no CN	Gruhn *et al.*[7]
4	10	No	Guthrie card-IGH no CN	Maia *et al.*[6]
Jena 29	12	Yes	Guthrie card-IGH no CN	Gruhn *et al.*[7]
Jena 32	14	Yes	Guthrie card-IGH no CN	Gruhn *et al.*[7]

CN: copy number of chromosome 14; dx: diagnosis; FISH: fluorescence *in situ* hybridization; HeH: high hyperdiploidy; IGH: immunoglobulin heavy chain.

for the characteristics of HeH as a secondary aberration. We investigated four ALL with concurrent *BCR::ABL1* and HeH, finding very high BTRI/all SNV in all (median 22%; range 19-44%), including a 4-year-old case (*Online Supplementary Table S1*), in line with HeH being relatively new in terms of leukemic evolution. In fact, the proportion of BTRI SNV was higher than in all but one of the HeH cases without *BCR::ABL1* investigated. Thus, the analysis corroborated the view that HeH as a secondary event would indeed have high BTRI/all SNV.

We attempted to find potential other primary events in cases with high BTRI/all SNV levels, but analysis of somatic WGS data for these cases did not reveal any unusual chromosomal events, structural rearrangements, targeted deletions, or coding mutations (*Online Supplementary Table S2*). Furthermore, the pattern of chromosomal gains, chromosomal modal numbers, and frequencies of the most common additional somatic events were similar (*Online Supplementary Table S2*). Thus, the only difference we can identify between HeH ALL with high BTRI/all SNV levels and the remaining HeH cases is that the former are, on average, older at diagnosis. Importantly though, some cases with high BTRI/all SNV were found in patients <10 at diagnosis. Thus, if there is an underlying primary event in these cases, it seems to be relatively age-independent although population-based cohorts would be needed to determine this.

In conclusion, we show that there is a marked difference in the percentage of BTRI/all SNV between young children and older children, adolescents, and adults in HeH ALL. We interpret this as most HeH ALL occurring in patients <10 years at diagnosis, constituting the majority of cases, have a prenatal origin of the chromosomal gains, in line with previous findings.[1-9] In contrast, HeH ALL occurring in older children, adolescents, and adults generally have a later, most likely postnatal, origin of the chromosomal gains. Notably, some cases in young children may also fall into this group, suggesting a hitherto unknown heterogeneity within HeH ALL. Considering that we have recently shown that the chromosomal gains occur simultaneously and very early in the leukemogenesis of HeH ALL,[15] we deem it likely that this means that also the ALL arises postnatally in these cases, although we cannot definitely exclude another, as yet hidden, primary event.

Authors

Minjun Yang,[1] Rebeqa Gunnarsson,[1] Nicolas Duployez,[2] Marketa Zaliova,[3,4] Jan Zuna,[3,4] Bertil Johansson[1,5] and Kajsa Paulsson[1]

[1]Department of Laboratory Medicine, Division of Clinical Genetics, Lund University, Lund, Sweden; [2]Laboratory of Hematology, Centre Hospitalier Universitaire (CHU) Lille, University of Lille, INSERM Unité 1277 Canther, Lille, France; [3]Department of Pediatric Hematology and Oncology, Second Faculty of Medicine, Charles University/University Hospital Motol, Prague, Czech Republic; [4]Childhood Leukemia Investigation Prague (CLIP), Prague, Czech Republic and [5]Department of Clinical Genetics, Pathology, and Molecular Diagnostics, Office for Medical Services, Laboratory Medicine, Region Skåne, Lund, Sweden

Correspondence:
K. PAULSON - kajsa.paulsson@med.lu.se

https://doi.org/10.3324/haematol.2023.284128

Received: August 21, 2023.
Accepted: February 8, 2024.
Early view: February 15, 2024.

Disclosures

No conflicts of interest to disclose.

Contributions

MY, RG, ND, MZ, JZ, BJ and KP performed research. MY and KP performed data analysis. KP supervised the study and wrote the manuscript with input from all authors.

Acknowledgments

The results published here are in part based upon data generated by the Therapeutically Applicable Research to Generate Effective Treatments (TARGET) Initiative (phs000464). This study makes use of data generated by the St. Jude Children's Research Hospital Genomes for Kids Study.[13] This study makes use of data generated by St. Jude Children's Research Hospital.[14]

Funding

This study was supported by grants from the Swedish Childhood Cancer Fund, grant numbers PR2020-0033 MY, TJ2020-0024 MY, PR2021-0005 BJ, and PR2021-0016 KP; the Swedish Cancer Fund, grant numbers 23 2694 Pj (to BJ) and 22 2062 Pj (to KP); governmental funding of clinical research within the National Health Service (to BJ and KP); National Institute for Cancer Research, grant number LX22NPO5102, funded by the European Union–Next Generation EU (to MZ and JZ), Czech Health Research Council, grant number NU21-03-00128 (to MZ and JZ); and the Swedish Research Council, grant numbers 2020-01164 (to BJ) and 2020-00997 (to KP).

Data-sharing statement

The WGS data from Lund University have been deposited to the European Genome-phenome Archive (EGA) under accession number EGAS00001007052. This dataset is available under restricted access due to privacy concerns; access can be obtained for academic research by contacting the Data Access Committee via EGA. The WGS data generated by the Therapeutically Applicable Research to Generate Effective Treatments (TARGET) are available at https://portal.gdc.cancer.gov/projects.under accession code phs000464. WGS data for pediatric tumor samples used for analysis in this study were obtained from St. Jude Cloud.[14]

References

1. Panzer-Grümayer ER, Fasching K, Panzer S, et al. Nondisjunction of chromosomes leading to hyperdiploid childhood B-cell precursor acute lymphoblastic leukemia is an early event during leukemogenesis. Blood. 2002;100(1):347-349.
2. Maia AT, van der Velden VHJ, Harrison CJ, et al. Prenatal origin of hyperdiploid acute lymphoblastic leukemia in identical twins. Leukemia. 2003;17(11):2202-2206.
3. Bateman CM, Alpar D, Ford AM, et al. Evolutionary trajectories of hyperdiploid ALL in monozygotic twins. Leukemia. 2015;29(1):58-65.
4. Yagi T, Hibi S, Tabata Y, et al. Detection of clonotypic IGH and TCR rearrangements in the neonatal blood spots of infants and children with B-cell precursor acute lymphoblastic leukemia. Blood. 2000;96(1):264-268.
5. Taub JW, Konrad MA, Ge Y, et al. High frequency of leukemic clones in newborn screening blood samples of children with B-precursor acute lymphoblastic leukemia. Blood. 2002;99(8):2992-2996.
6. Maia AT, Tussiwand R, Cazzaniga G, et al. Identification of preleukemic precursors of hyperdiploid acute lymphoblastic leukemia in cord blood. Genes Chromosomes Cancer. 2004;40(1):38-43.
7. Gruhn B, Taub JW, Ge Y, et al. Prenatal origin of childhood acute lymphoblastic leukemia, association with birth weight and hyperdiploidy. Leukemia. 2008;22(9):1692-1697.
8. Kacanski N, Kolarovic J, Kostic T, et al. Presence of leukemic clone-specific immunoglobulin heavy chain rearrangements in neonatal blood spots of children with B-cell precursor acute lymphoblastic leukemia. Int J Lab Hematol. 2024;46(2):303-311.
9. Zuna J, Prouzova Z, Kalina T, et al. Backtracking of ALL to cord blood. Leuk Res. 2009;33(8):e107-108.
10. Paulsson K, Lilljebjörn H, Biloglav A, et al. The genomic landscape of high hyperdiploid childhood acute lymphoblastic leukemia. Nat Genet. 2015;47(6):672-676.
11. Downing JR, Wilson RK, Zhang J, et al. The pediatric cancer genome project. Nat Genet. 2012;44(6):619-622.
12. McLeod C, Gout AM, Zhou X, et al. St. Jude Cloud: a pediatric cancer genomic data-sharing ecosystem. Cancer Discov. 2021;11(5):1082-1099.
13. Newman S, Nakitandwe J, Kesserwan CA, et al. Genomes for kids: the scope of pathogenic mutations in pediatric cancer revealed by comprehensive DNA and RNA sequencing. Cancer Discov. 2021;11(12):3008-3027.
14. Rusch M, Nakitandwe J, Shurtleff S, et al. Clinical cancer genomic profiling by three-platform sequencing of whole genome, whole exome and transcriptome. Nat Commun. 2018;9(1):3962.
15. Woodward EL, Yang M, Moura-Castro LH, et al. Clonal origin and development of high hyperdiploidy in childhood acute lymphoblastic leukaemia. Nat Commun. 2023;14(1):1658.
16. Paulsson K, Harrison CJ, Andersen MK, et al. Distinct patterns of gained chromosomes in high hyperdiploid acute lymphoblastic leukemia with t(1;19)(q23;p13), t(9;22)(q34;q22) or MLL rearrangements. Leukemia. 2013;27(4):974-977.