Supplementary File

Article: HOW TO COMBINE MULTIPLE TOOLS FOR THE GENETIC DIAGNOSIS WORK-UP OF PEDIATRIC B-CELL ACUTE LYMPHOBLASTIC LEUKEMIA

Gloria Hidalgo-Gómez[1-2], Bárbara Tazón-Vega[1], Carlos Palacio[1], Silvia Saumell[1], Noemi Martínez-Morgado[1], Víctor Navarro[3], Laura Murillo[4], Pablo Velasco[4], Thais Murciano[4], Cristina Díaz de Heredia[4], Francesc Bosch[1], Gemma Armengol[2], Margarita Ortega[1]

Author affiliations:
[1]Hematology Service, Vall d'Hebron Hospital Universitari, Experimental Hematology, Vall d'Hebron Institute of Oncology (VHIO), Hospital Universitari Vall d'Hebron, Vall d'Hebron Barcelona Hospital Campus, Passeig Vall d'Hebron 119-129, 08035 Barcelona, Spain.
[2]Unit of Biological Anthropology, Department of Animal Biology, Plant Biology, and Ecology, Faculty of Biosciences, Universitat Autònoma de Barcelona, 08193 Barcelona, Catalonia, Spain.
[3]Oncology Data Science (ODysSey) Group, Vall d'Hebron Institute of Oncology (VHIO), Vall d'Hebron Barcelona Hospital Campus, Passeig Vall d'Hebron 119-129, 08035 Barcelona, Spain.
[4]Pediatric Oncology and Hematology Department, Hospital Universitari Vall d'Hebron, Vall d'Hebron Barcelona Hospital Campus, Passeig Vall d'Hebron 119-129, 08035 Barcelona, Spain.

Corresponding author:
Margarita Ortega
Vall d'Hebron Hospital Universitari | Hematology Service
Passeig Vall d'Hebron 119-129, 08035 Barcelona, Spain
Tel: +34 93 274 6205
Email: margarita.ortega@vallhebron.cat

Table of contents of supplementary patients and methods

List of supplementary figures

References for supplementary information

Samples and patients

Samples were collected from bone marrow aspirates at the time of the patient's diagnosis and/or relapse. All patients provided informed consent in accordance with protocols approved by Institutional Ethics Committee of Vall d'Hebron University Hospital and in compliance with the Declaration of Helsinki.

Morphology and immunophenotype

Bone marrow (BM) aspirate smears were stained with May–Grünwald–Giemsa according to standard laboratory protocols for cytochemical stains, including myeloperoxidase and Periodic Acid-Schiff (PAS). The diagnosis of ALL was based on the morphologic features of the cells, namely the presence of more than 20% cells showing lymphoblast morphology.

Immunophenotyping was performed by flow cytometry on BM samples according to standardized procedures. Antibodies used included anti-CD45, CD34, CD123, CD10, CD19, CD20, CD22, CD9, CD24, CD25, CD38, CD73, CD15, NG2, CD66c (KORSA), CD33, CD13, cit IgM, citMPO, nucTdT and citCD3 (from Beckman Coulter, CA, USA), CD304 (from Mitenyi Biotec, Germany).

Cytogenetic analysis

G-banded chromosomal analyses were performed on BM samples using standard techniques. Karyotypes were described according to the International System for Cytogenetic Nomenclature (ISCN 2020)[1].

Fluorescent in situ hybridization (FISH) assay was carried out on interphase nuclei using commercial probes for *BCR::ABL1*, *KMT2A*, *ETV6::RUNX1, TCF3, CRLF2, ABL1, ABL2, PDGFRB, JAK2, IGH* (MetaSystems Probes, Germany), *EPOR, MEF2D, ZNF384* and *PAX5* (Empire Genomics, NY, USA), according to the manufacturer's protocol.

RT-PCR

The fusion transcript *BCR::ABL1* was analyzed from 1µg of total RNA by RT-PCR using primers and procedures previously described[2]. A delta-delta Ct analysis was performed normalizing through *ABL1* as the control gene. This technique is used diagnostically to confirm the presence of the BCR::ABL1 transcript and, if positive, is also employed for follow-up.

MLPA analysis

MLPA reactions were performed using the SALSA MLPA probemix P335 ALL-IKZF1 (version C2) and P327 iAMP21-ERG (version B2) (MRC-Holland, Amsterdam, Netherlands) according to the manufacturer's instructions. DNA samples from three healthy donors served as controls. For MLPA studies, a minimum infiltration of ≥30% blasts in bone marrow and high-quality DNA was required. The P335 probemix contains probes for *IKZF1, CDKN2A/B, PAX5, EBF1, ETV6, BTG1, RB1,* as well as genes from the X/Y PAR1

region (*CRLF2, CSF2RA, IL3RA and P2RY8*). The P327 probemix contains probes for RUNX1 and ERG genes. MLPA amplification products were analyzed on an ABI 3130XL Genetic Analyzer (Applied Biosystems), using the Genescan 500LIZ internal size standard (Applied Biosystems). MLPA analysis was performed on the Coffalyser.NET Software (MCR-Holland). CNA were defined, following Schwab et al.[3], as: biallelic deletion (<0.25), monoallelic deletion (≥0.25 to <0.75), normal (≥0.75 to <1.3) and gain/amplification (≥1.3).

Targeted next generation sequencing (t-NGS)

t-NGS was performed using Oncomine Childhood Cancer Research Assay (ThermoFisher Scientific, Life Technologies Corporation, CA, USA), which includes 203 genes involved in childhood cancer and allows the analysis of point mutations and fusions. Libraries were automatically prepared using the Ion Chef and sequenced on the S5 sequencer (ThermoFisher Scientific, Life Technologies Corporation, CA, USA), with an average coverage of 2500x. Variant alignment and detection were performed in Ion Reporter with the human reference genome (hg19). Variants with >1% allelic frequency were considered and manually reviewed with the Integrated Genome Viewer and classified according to the VARSOME23 software. Fusions were considered when transcript levels were >0.1%. Data from the literature and the Catalogue Of Somatic Mutations In Cancer (COSMIC) database were used for the interpretation of the somatic nature of the fusions and mutant calls. Single-nucleotide polymorphisms (SNPs) listed in the SNP database (dbSNP) and the Genome 10K Project were excluded.

Optical Genome Mapping

Optical Genome Mapping (OGM) requires ultra-high molecular weight (UHMW) DNA. It was extracted from ~$1.5x10^6$ cryopreserved cells (from bone marrow or peripheral blood) according to the manufacturer's instructions (Bionano Genomics, San Diego, USA). DNA was labeled using a sequence-specific DLE-1 (Direct Labeling Enzyme), that attaches a green fluorophore to a specific 6bp sequence, present around 15 times per 100 kb in the human genome. The labeled DNA was loaded onto a Saphyr chip and scanned on the Saphyr instrument (Bionano Genomics, San Diego, USA). Changes in the patterning or spacing of the labels were identified by software solutions to accurately detect all classes of SVs. For each sample, an effective genome coverage of about 300x is achieved, with a theoretical mean variant allele frequency sensitivity of 5%.

Summary of genetic diagnosis techniques

Karyotyping allowed us to detect aneuploidies, including hyperdiploidy (HeH, characterized by 51–67 chromosomes) and hypodiploidy (low hypodiploidy is defined as 30–39 chromosomes while near haploidy includes 24–29 chromosomes). Recurrent translocations such as t(9;22)(q34;q11.2), t(1;19)(q23;p13), and 11q23 rearrangements were also identified through this method.

FISH detected specific fusions based on the probes used, focusing on rearrangements like *KMT2A*, *ETV6::RUNX1*, *ZNF384*, *CRLF2*, among others. For instance, iAMP21 was diagnosed using the *ETV6::RUNX1* FISH probe. Initial suspicion arose from observing clustered signals in the interphase nucleus study, and the diagnosis was confirmed by detecting tandem duplicated *RUNX1* signals on a single abnormal chromosome 21 in metaphase. Furthermore, *BCR::ABL1* fusions could be identified using either FISH or RT-PCR, both of which are well-suited for detecting this fusion with high specificity.

t-NGS provided high sensitivity for detecting specific fusions and mutations, offering critical molecular insights. OGM can detect a wide range of structural variants, including balanced translocations, inversions, and large insertions or deletions, offering an even broader genomic perspective. In this study, it was used as a final step to identify genetic alterations in cases that remained unclassified after applying the previous diagnostic methods. Additionally, MLPA was applied to identify copy number alterations (CNAs) commonly associated with B-ALL, such as deletions of *ETV6*, *PAX5*, *IKZF1*, and *CDKN2A/B*.

Alternative methods, such as RT-PCR, are useful for detecting fusion transcripts, while SNP-arrays allow for detailed profiling of aneuploidies and copy number variations. The combination of these techniques ensures comprehensive genetic characterization, highlighting the importance of an integrative diagnostic approach.

Stratification risk of SEHOP-PETHEMA 2013 treatment guidelines

SEHOP-PETEHEMA 2013 protocol[4] incorporates risk-adapted treatment regimens based on factors such as age, white blood cell count (WBC), extramedullary infiltration (specifically CNS or testes involvement), immunophenotype, minimal residual disease (MRD), and cytogenetics.

For standard-risk group, patient must meet all of the following criteria: age >1 and <10 years, leukocytes $<20x10^9$/L at diagnosis, non-T immunophenotype, absence of CNS and/or testicular infiltration, cytogenetics with high hyperdiploidy (51-67 chromosomes) and *ETV6::RUNX1* positive, and not presenting t(1;19) or *KMT2A* rearrangement. Additionally, they must have <1,000 blasts/mm^3 in peripheral blood on day +8 of induction, <5% blasts and <0.1% MRD in bone marrow on day +15 of induction.

The existence of any of the following criteria determines inclusion in the high-risk group: t(4;11)/*KMT2A::AFF1*, hypodiploidy (<44 chromosomes or DNA index <0.81), presence of >1,000 blasts/mm^3 in peripheral blood on day +8 of Induction, > 25% blasts and > 10% MRD in bone marrow on day +15 of induction, MRD > 1% on day +33 of induction or MRD >0.1% before consolidation in bone marrow.

Those patients who do not meet the standard-risk or high-risk criteria will be included in the intermediate-risk group.

Statistical methods

Overall survival (OS) was defined as time to death, censoring at date of last contact. Event-free survival (EFS) was defined as time from randomization to relapse, second tumour or death, censoring at date of last contact. All survival rates are quoted at 5 years. Kaplan-Meier methods were used to estimate survival rates. Other comparisons were performed using chi-square, Fisher's exact, and Mann–Whitney tests, as appropriate.

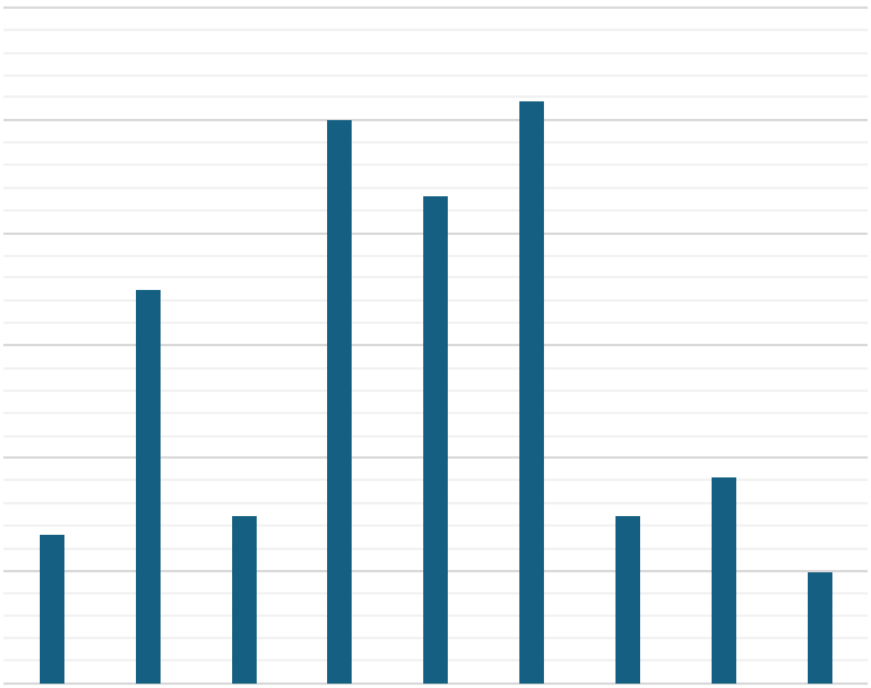
Analyses were conducted using IBM SPSS version 25.0 (IBM Corp., Armonk, NY, USA) and R (www.r-project.org) software.

Data availability statement

Due to concerns regarding patient privacy, the datasets employed in this study are not accessible to the public; however, upon reasonable request to the corresponding, they may be made accessible. Please contact Gloria Hidalgo (gloria.hidalgo@vallhebron.cat) and Margarita Ortega (margarita.ortega@vallhebron.cat).

Supplementary figures

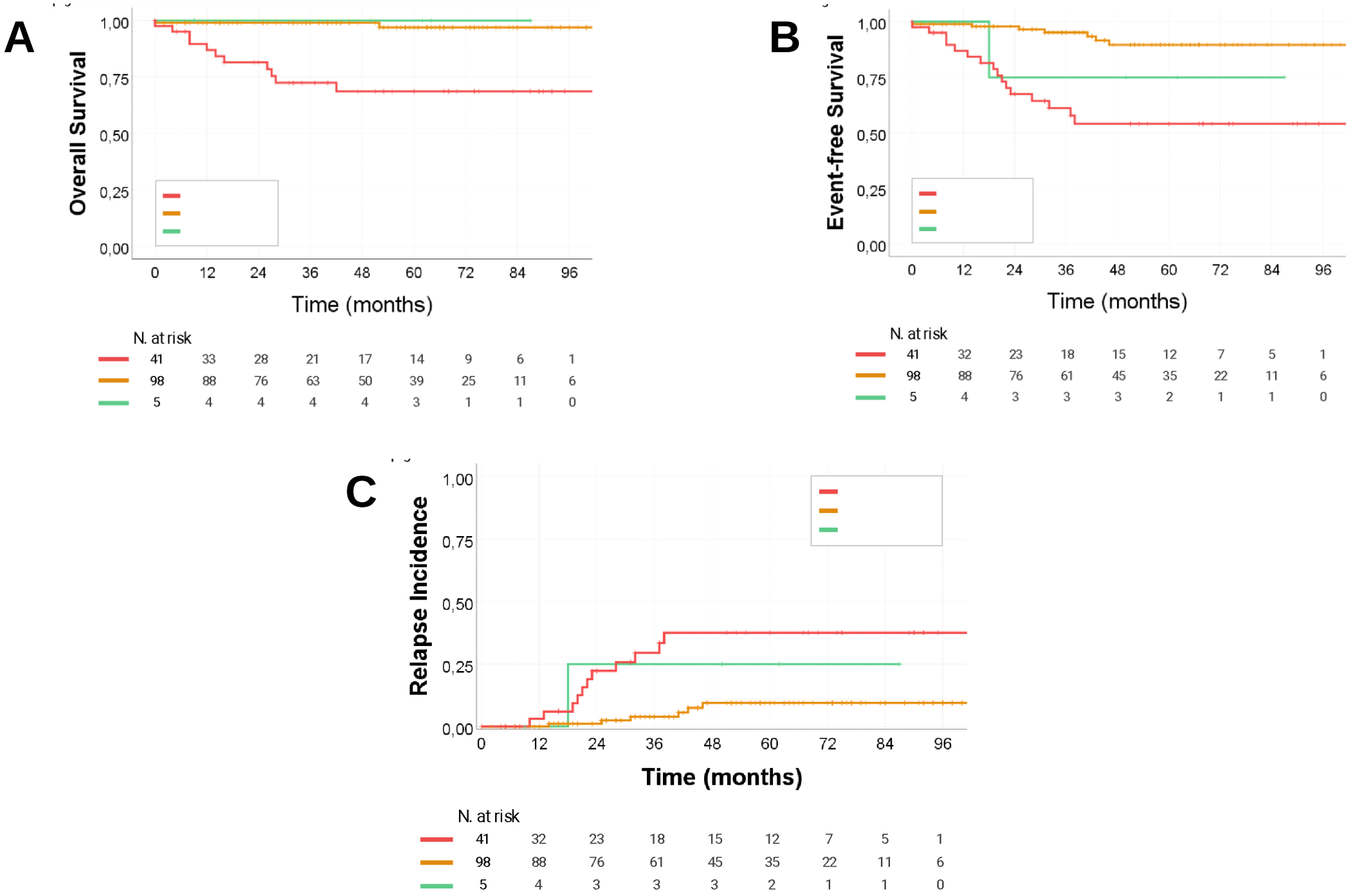
Figure S1. Bar chart and table showing the frequency and co-occurrence of copy number alterations (CNA) of the genes studied by P335 MLPA our pediatric B-ALL cohort. Significant (p<0.05) concurrent deletions are marked in light blue.



	EBF1	IKZF1	JAK2	CDKN2A/B	PAX5	ETV6	BTG1	RB1	PAR1
EBF1	8								
IKZF1	4	21							
JAK2	1	3	9						
CDKN2A/B	1	8	9	30					
PAX5	1	7	8	15	26				
ETV6	5	7	4	9	8	31			
BTG1	1	4	1	3	4	5	9		
RB1	4	4	2	4	4	4	2	11	
PAR1	1	4	1	3	1	2	2	1	6

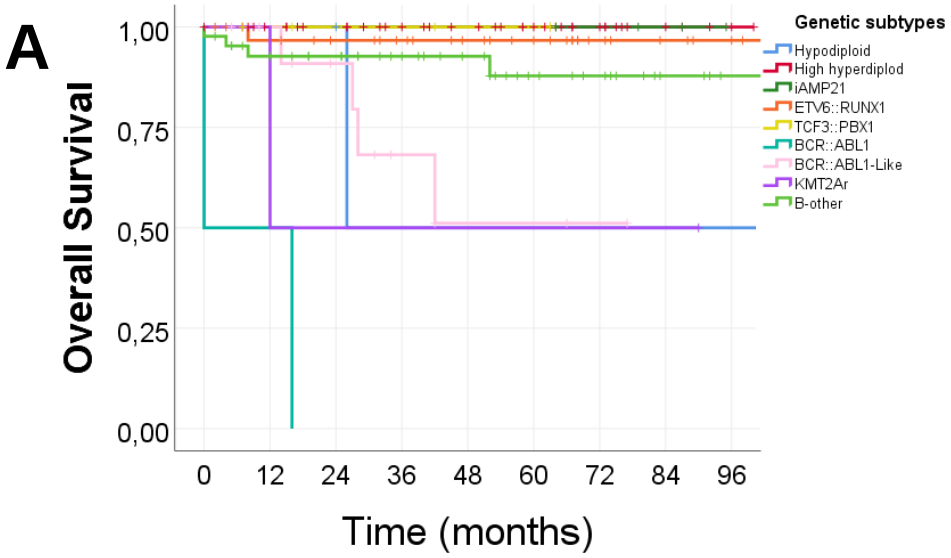
Supplementary figures

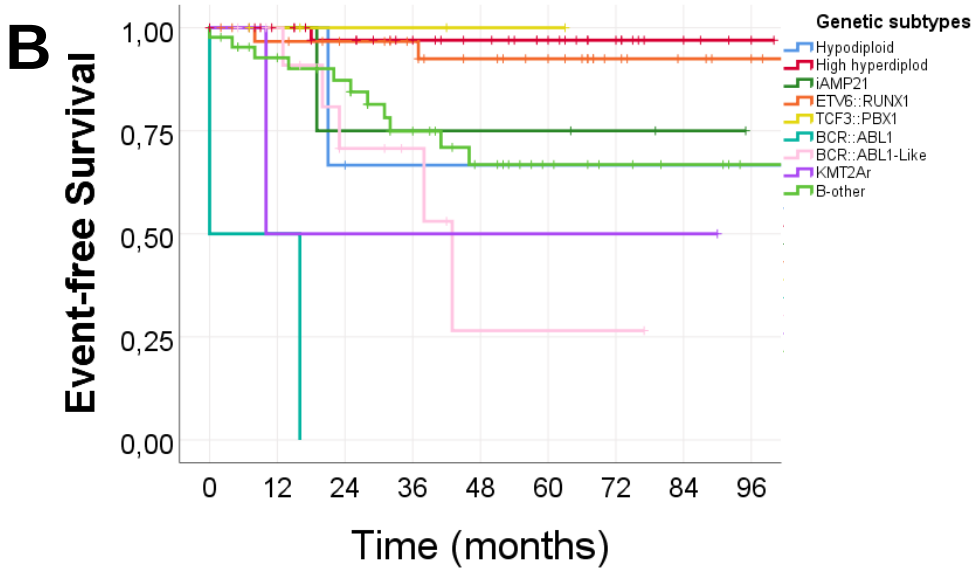
Figure S2. Outcomes according to SEHOP-PETHEMA 2013 Protocol risk classification for our pediatric B-ALL cohort. (A–C) Kaplan–Meier survival curves showing the overall survival, event-free survival and relapse incidence according to SEHOP-PETHEMA 2013 risk classification.

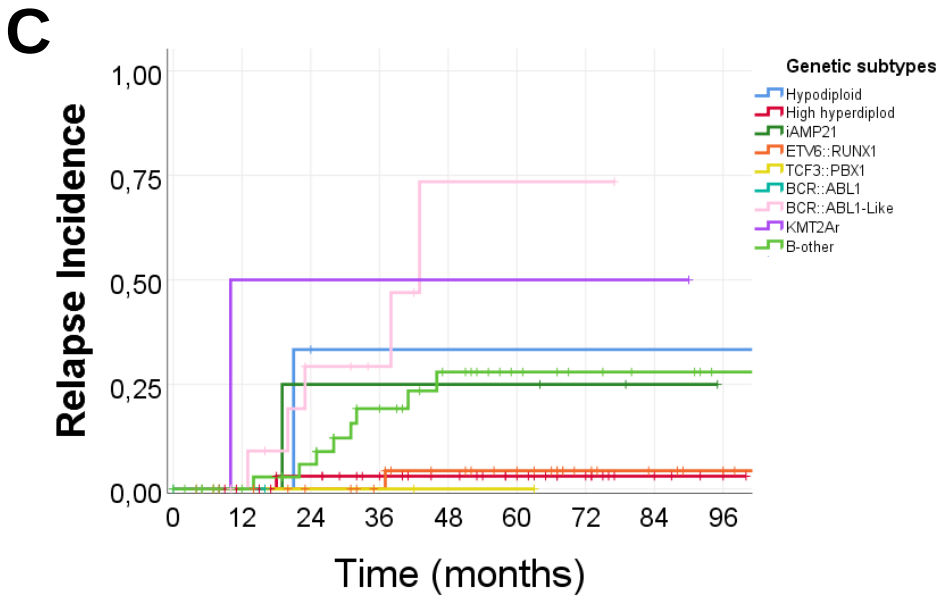


Supplementary figures

Figure S3. Outcomes according to the genetic defining subtype for our pediatric B-ALL cohort. (A–C) Kaplan–Meier survival curves showing the overall survival, event-free survival and relapse incidence according to genetic subtype.

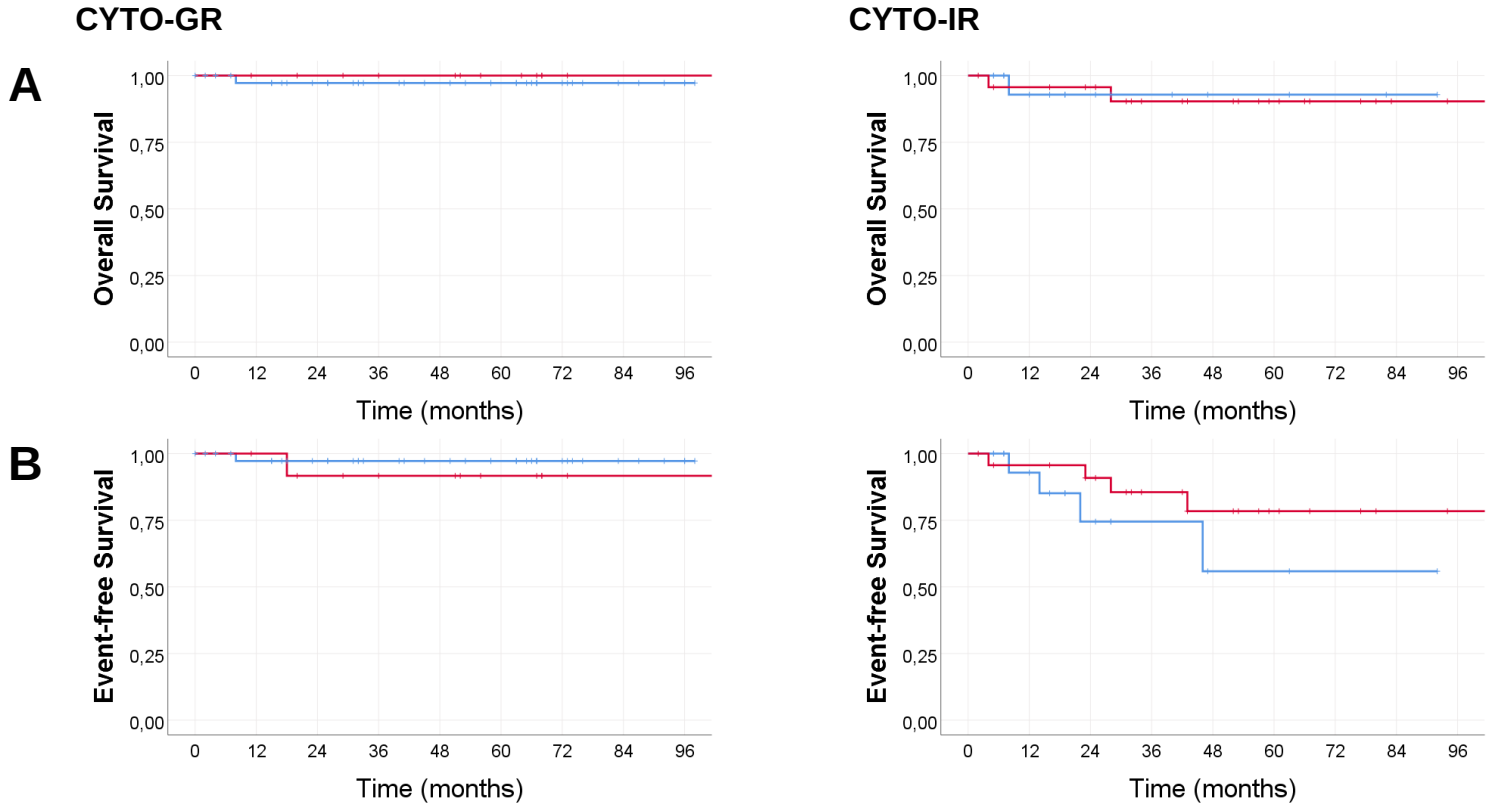






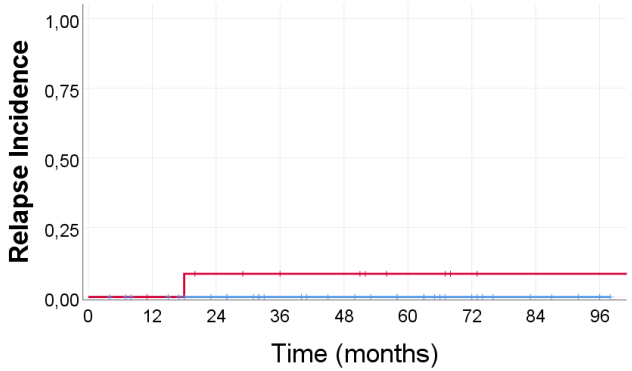
Supplementary figures

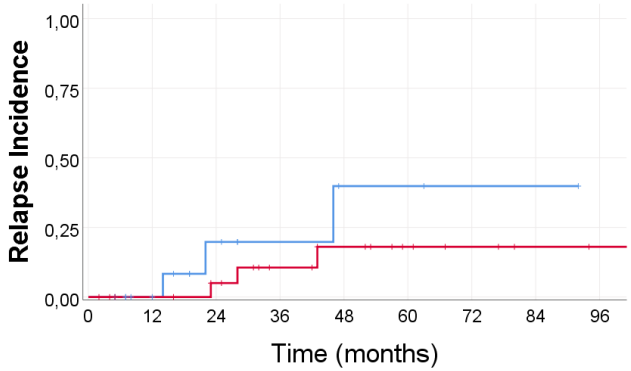
Figure S4. Outcomes according to the genetic risk classification plus the information of CNA profile for our pediatric B-ALL cohort. (A–C) Kaplan–Meier survival curves showing the overall survival, event-free survival and relapse incidence of the CYTO-GR and CYTO-IR by adding the CNA risk profile (CNA-GR: blue; CNA-PR: red). The results for the GEN-PR group are likely not shown due to a small sample size.



Supplementary figures

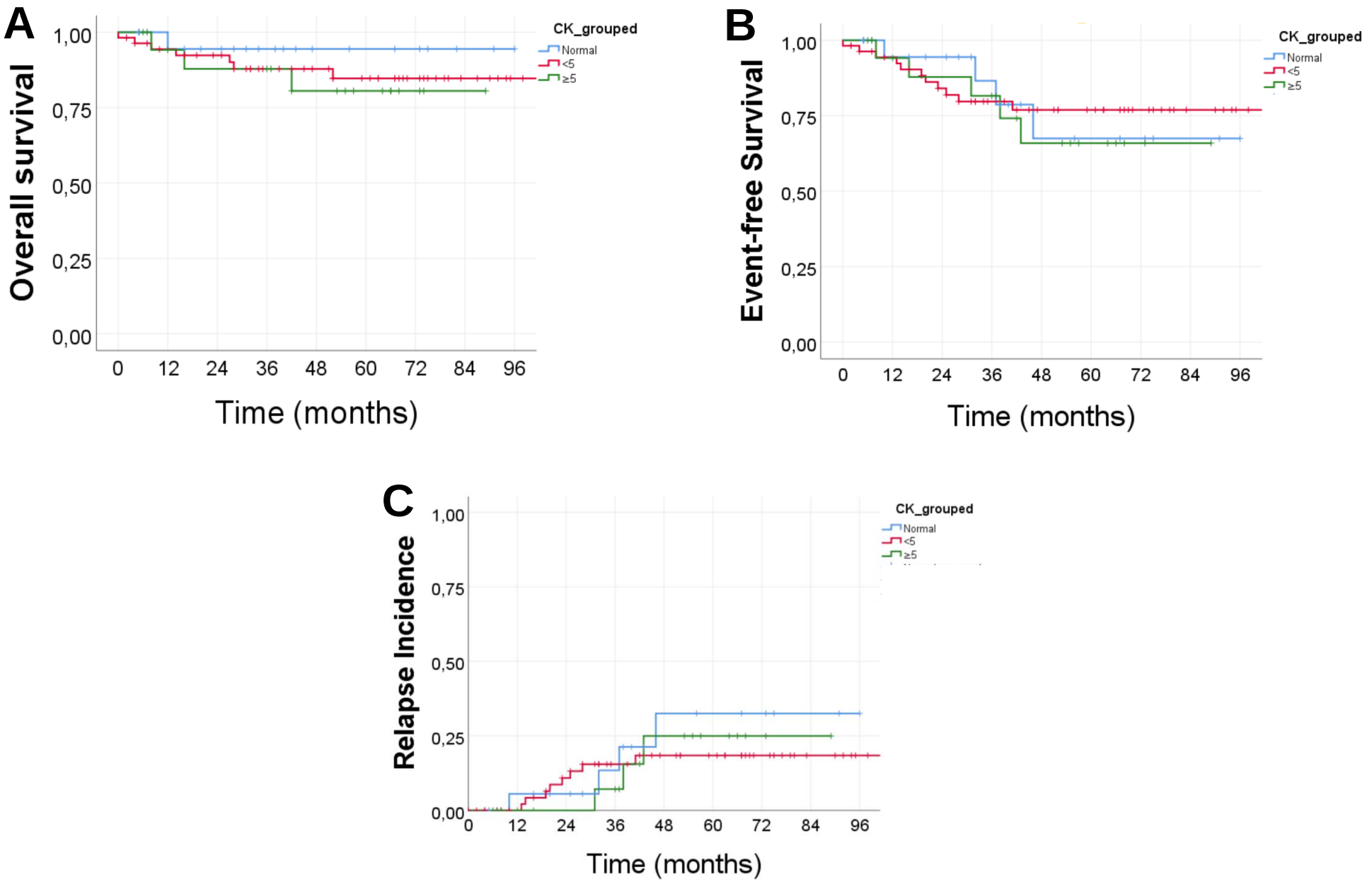
C





Supplementary figures

Figure S5. Outcomes according to normal (blue), altered (<5 alterations; red) or complex karyotype (≥5 alterations; green) for our pediatric B-ALL cohort. (A–C) Kaplan–Meier survival curves showing the overall survival, event-free survival and relapse incidence according to genetic subtype.



References

1. McGowan-Jordan J, Hastings RJ, Moore S (2020) ISCN 2020: An International System for Human Cytogenomic Nomenclature (2020). Kager

2. Van Dongen JJM, Macintyre EA, Gabert JA, et al (1999) Standardized RT-PCR analysis of fusion gene transcripts from chromosome aberrations in acute leukemia for detection of minimal residual disease. Report of the BIOMED-1 Concerted Action: Investigation of minimal residual disease in acute leukemia. Leukemia 13:1901–1928

3. Schwab CJ, Jones LR, Morrison H, et al (2010) Evaluation of multiplex ligation-dependent probe amplification as a method for the detection of copy number abnormalities in B-cell precursor acute lymphoblastic leukemia. Genes Chromosomes Cancer 49:1104–1113. https://doi.org/10.1002/gcc.20818

4. SEHOP, PETHEMA (2013) Protocolo para Leucemia Aguda Linfoblástica de la Sociedad Española de Hematología y Oncología Pediátricas (SEHOP) en colaboración con el grupo PETHEMA LAL/SEHOP-PETHEMA 2013