Received: 8 October 2024 | Accepted: 14 March 2025

DOI: 10.1111/bjh.20056

BJHaem
BRITISH JOURNAL OF HAEMATOLOGY

ORIGINAL PAPER

Biology and Translational Science

RNA-sequencing: A reliable tool to unveil transcriptional landscape of paediatric B-other acute lymphoblastic leukaemia

Clara Vicente-Garcés[1,2] | Guerau Fernández[3,4] | Elena Esperanza-Cebollada[1] | Mercè Richarte-Franqués[1] | Alba Crespo-Carrasco[1] | Sara Montesdeoca[1,5] | Ignacio Isola[1,5] | Edurne Sarrate[1,5] | Esther Cuatrecasas[1,3] | Susana Rives[1,4,6] | José Luis Dapena[1,6] | Mireia Camós[1,4,5] | Nerea Vega-García[1,5]

[1]Developmental Tumors Biology Group, Leukaemia and Paediatric Haematology Disorders, Pediatric Cancer Center Barcelona (PCCB), Institut de Recerca Hospital Sant Joan de Déu (IRSJD), Esplugues de Llobregat, Barcelona, Spain

[2]Facultat de Medicina i Ciències de la Salut, Universitat de Barcelona (UB), Barcelona, Spain

[3]Department of Genetic and Molecular Medicine-IPER, Hospital Sant Joan de Déu, Institut de Recerca Sant Joan de Déu, Barcelona, Spain

[4]Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Instituto de Salud Carlos III, Madrid, Spain

[5]Hematology Laboratory, Hospital Sant Joan de Déu Barcelona, Esplugues de Llobregat, Barcelona, Spain

[6]Leukemia and Lymphoma Department. Pediatric Cancer Center Barcelona (PCCB), Hospital Sant Joan de Déu Barcelona, Barcelona, Spain

Correspondence
Nerea Vega-García, Hematology Laboratory, Hospital Sant Joan de Déu, Passeig Sant Joan de Déu 2, Esplugues de Llobregat, Barcelona 08950, Spain.
Email: nerea.vega@sjd.es

Funding information
Instituto de Salud Carlos III, Grant/Award Number: PI21/00213

Summary

B-cell precursor acute lymphoblastic leukaemia (BCP-ALL) comprises multiple subtypes characterized by different genetic alterations. With the use of current standard-of-care tests used in clinical practice, 20%–30% of the cases may not be classified into the main genetic subtypes and additional approaches are needed. These patients are grouped in the heterogeneous category B-other ALL. Transcriptome sequencing (RNA-seq) has allowed the identification of novel fusion genes and gene expression profiles that define new molecular subtypes. We present RNA-seq results integrated, in a real-world scenario, with clinical routine diagnostic data to identify new biomarkers and reclassify a cohort of 60 B-other ALL patients in the newly described genetic subtypes. Overall, 49 rearrangements were identified, including 32 different fusion genes in 41 B-other patients (68%). Moreover, we reported six novel rearrangements (*IGK::PAX5*, *PAX5::IL1RAPL1*, *ETV6::KRT78*, *IGH::HIC1*, *IGH::MIR100HG* and *NKAIN4::PNPLA7*). The integration of RNA-seq results with standard-of-care data allowed us to classify 72% of the patients (43/60) in 11 different subtypes, being *DUX4* rearranged and *PAX5*alt the most represented subtypes. In summary, RNA-seq is a reliable tool for the identification of new emerging genetic subtypes contributing to a better genetic risk stratification of BCP-ALL paediatric patients on the path towards a more personalized medicine.

KEYWORDS
gene expression profile, paediatric acute lymphoblastic leukaemia, RNA-seq

INTRODUCTION

B-cell precursor acute lymphoblastic leukaemia (BCP-ALL) is the most frequent childhood cancer,[1] characterized by heterogeneous biology comprising different genetic alterations.[2–4] These features are crucial for classifying patients into different subgroups and guiding treatment protocols. Standard diagnostic methods used in clinical practice, including karyotyping, fluorescent in situ hybridization (FISH) and reverse transcription (quantitative) polymerase chain reaction

[Correction added on 14 May 2025, after first online publication: The subcategory has been changed.]

(RT-qPCR), identify well-established leukaemia subtypes: high hyperdiploidy (51–67 chromosomes), hypodiploidy (<44 chromosomes), *ETV6::RUNX1*, *BCR::ABL1*, *TCF3::PBX1*, *TCF3::HLF*, *KMT2A* rearrangements and intrachromosomal amplification of chromosome 21 (iAMP21), accounting for 70% of BCP-ALL patients. However, with conventional methods, we fail to classify into the main recurrent genetic categories around 20%–30% of cases, currently named B-other ALL, and additional or more directed approaches are needed to precisely identify all molecular alterations.[5–8]

Next-generation sequencing (NGS) has improved leukaemia diagnosis by revealing genetic alterations defining new subgroups such as *DUX4*, *PAX5*, *MEF2D*, *ZNF384* rearrangements or *CDX2* and *UBTF::ATXN7L3* alterations, among others.[3,9–14] Particularly, RNA-sequencing (RNA-seq) has further identified patients with highly distinctive gene expression profiles (GEP) such as *BCR::ABL1-like* and *ETV6::RUNX1-like*[15,16] or *PAX5*-P80R and *IKZF1*-N159Y single mutations.[17] Some of these discoveries have been incorporated into the latest World Health Organization (WHO) classification of paediatric tumours[18] and the International Consensus Classification (ICC) of Myeloid Neoplasms and Acute Leukaemia,[19] highlighting the need to incorporate genome-wide methodologies in BCP-ALL diagnosis.

We present our experience integrating RNA-seq in diagnostics routine, in a B-other ALL cohort, to reclassify those patients in the newly established leukaemia subgroups.

MATERIALS AND METHODS

Patients cohort

We have studied 83 B-other ALL patients diagnosed between 2009 and 2023 at a single institution, treated with SHOP-2005 and SEHOP-PETHEMA 2013 protocols. B-other ALL diagnosis was achieved by discarding the presence of *ETV6::RUNX1*, *TCF3::PBX1*, *TCF3::HLF*, *BCR::ABL1*, *KMT2A* rearrangements, high hyperdiploidy (51–67 chromosomes), hypodiploidy (<44 chromosomes) and iAMP21 using karyotype, FISH, RT-qPCR, multiplex ligation-dependent probe amplification (MLPA) and DNA Index by flow cytometry. High-quality RNA from bone marrow or peripheral blood was available for RNA-seq in 60 of the 83 patients.

This study strictly followed the ethical standards and the Declaration of Helsinki and obtained a favourable report from our Ethics Committee (PIC-215-20). All samples were stored in our legally recognized biobank 'Biobanco Hospital Sant Joan de Déu (B.000059 ISCIII)' after obtaining the corresponding written informed consent from patients or their legal guardians.

RNA-seq library preparation and sequencing

RNA was extracted using Direct-zol™ RNA MiniPrep (Zymo Research, CA, USA). Quality and concentration were assessed by TapeStation (Agilent, Santa Clara, CA, USA) and Qubit 4.0 fluorometer (Thermo Fisher Scientific, MA, USA). Samples with RIN ≥6 were included in the study. As assessed by flow cytometry, all samples presented more than 60% leukaemic cells.

Libraries were prepared using the Illumina® TruSeq™ stranded mRNA kit following the manufacturer's instructions with an initial input of 500–700 ng of total RNA. Library quality was assessed by TapeStation D1000 reagents and quantified using the Qubit 4.0 fluorometer. Sequencing was performed on a NextSeq 550 using a v2 HighOutput 150 cycle kit 2 × 75 bp (Illumina, San Diego, CA, USA). Generated data were submitted to the European Genome-phenome Archive (EGA) (EGAD50000000981).

Fusion gene detection

Pair-end reads quality control was assessed using FastQCv.0.11.5 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/), followed by the removal of low-quality reads and adapters using cutadapt v1.13. The remaining reads were aligned to the reference human genome GRCh38 using the STAR aligner.[20] Fusion calling was conducted using the Fusion InPipe algorithm.[21] Overall, we selected those fusions called by three or more algorithms, and we applied different filters, as described,[21] to discard artefacts and false-positive variants. Due to the complexity of *DUX4* rearrangements these rearrangements were kept regardless of the number of callers detecting the fusion. If no fusion candidates were found, a second manual inspection was done, focusing on fusion events involving genes linked to leukaemia that might have been missed or not detected by at least three algorithms. The final candidates were visually verified by reviewing supporting reads in the New Genome Browser (NGB) data viewer.

The identified fusion genes were validated by RT-qPCR after specific oligonucleotides and probes were designed (Table S1), except in those cases, where more than one breakpoint was identified, the rearrangement included intragenic regions or no sample was available.

Gene expression profile analysis

Read counts per gene were calculated using gencode. v34lift37 annotation by the HTSeq package v1.99.2.[22] Gene expression levels were normalized by the DESeq2 v1.42,[23] and batch effect introduced by different library preparation, gender and sequencing strategies was corrected using limma v.3.58.1 R package.[24]

A cohort of patients from St. Jude Children's Research Hospital (SJCRH) (EGAD00001004461 and EGAD00001004463)[17] was used to generate a t-distributed Stochastic Neighbour Embedding (t-SNE) plot and visualize the different cluster subtypes. With the SJCRH cohort

as a baseline, we individually interrogated our patients to assign them to a specific subtype based on the patient's GEP.

The GEP defined different groups, including *BCR::ABL1/BCR::ABL1-like*, *DUX4*, *ETV6::RUNX1/ETV6::RUNX1-like*, hyperdiploidy, *iAMP21*, *KMT2A*, *MEF2D*, *NUTM1*, *PAX5alt*, *PAX5*-P80R, *TCF3::PBX1*, *ZEB2/CEBPE* and *ZNF384*. The *PAX5alt* group included patients with rearrangements, focal amplifications and mutations different from P80R.

Reclassification of B-other ALL patients and genetic risk stratification

Patients were assigned to one of the subtypes proposed by the WHO[18] and the ICC[19] classifications when the identified fusion gene and the GEP were concordant. For discordant results, available data from conventional methods (including karyotype, FISH, MLPA, flow cytometry), targeted NGS (using AmpliSeq™ for Illumina® Childhood Cancer Panel) and ALLCatchR,[25] were integrated to finally reclassify the B-other ALL patients into a specific subtype. Patients were classified when two or more techniques provided evidence to support this new classification (Figure 1). Finally, genetic risk was assigned based on updated literature.[12,19,26]

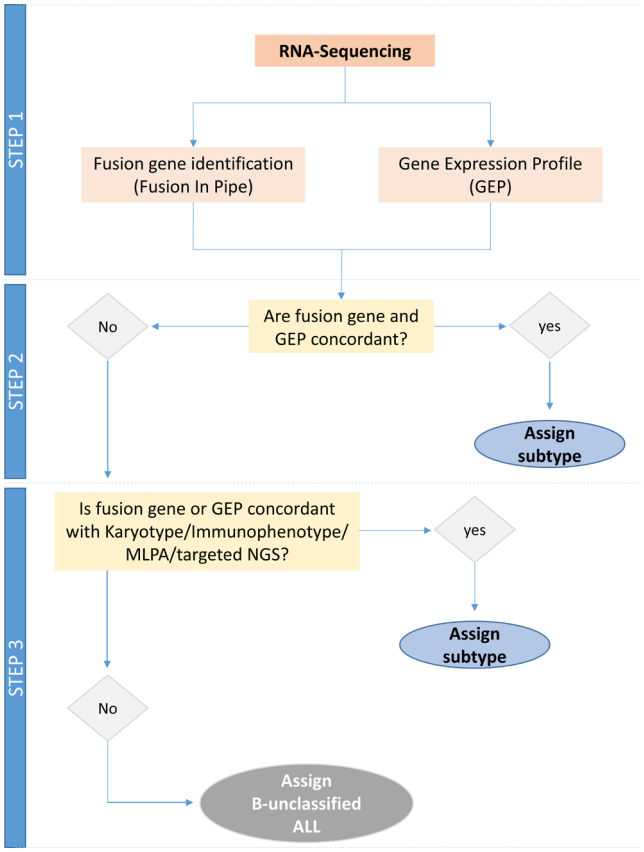

FIGURE 1 Flowchart with the steps to assign patients into different BCP-ALL subtypes. During the first step, results from RNA-sequencing were obtained and analysed separately. The second step integrated results from fusion calling and gene expression profile (GEP) and, finally, step 3 integrated RNA-seq results with standard-of-care diagnostic tests (karyotype, flow cytometry, MLPA), as well as targeted NGS and ALLCatchR. BCP-ALL, B-cell precursor acute lymphoblastic leukaemia; MLPA, multiplex ligation-dependent probe amplification; NGS, next-generation sequencing.

RESULTS

RNA-seq allows identifying new fusion genes

We identified 49 rearrangements, including 32 different fusion genes, in 41 patients (68%). Thirty-six fusion genes were validated by RT-qPCR, and all but two cases were confirmed. Detailed rearrangement information (including distinction between in-frame and out-of-frame fusions) is included in Figure 2 and Table S2. Most of the rearrangements (44/49, 90%) belonged to the emerging genetic subtypes (*ABL2*, *CRLF2*, *DUX4*, *ETV6*, *CEBP*, *IGH*, *MEF2D*, *NUTM1*, *PAX5* or *ZNF384* rearrangements). *PAX5* was the most frequently rearranged gene, present in 11 patients (11/41, 27%, including eight in-frame and three out-of-frame rearrangements) with eight different gene partners (*NOL4L*, *PAN3*, *PML*, *ETV6*, *FOXP1*, *IGK*, *IL1RAPL1*, and *ZCCHC7*), followed by *DUX4* (9/41, 22%, five CDS-truncated, four out-of-frame), with *IGH* as a partner. Other rearrangements were less represented: *CRLF2* ($n = 6$, 15%, one CDS-truncated, three UTR, two intergenic), *ETV6* ($n = 5$, 12%, three in-frame, two out-of-frame), *CEBP* family ($n = 3$, 7%, one CDS-truncated, one out-of-frame, one UTR), *IGH* ($n = 3$, 7%, one in-frame, one CDS-truncated, one UTR), *MEF2D* ($n = 2$, 5%, in-frame), *ZNF384* ($n = 2$, 5%, in-frame) and one patient each with *ABL*-class (in-frame), *NUTM1* (in-frame) and *TCF3* (out-of-frame) rearrangements. Among them, two patients presented *ETV6::RUNX1* and *TCF3::PBX1* fusions that were missed by conventional methodologies as they presented rare breakpoints not routinely checked. Additionally, we identified five fusion genes that do not belong to any of the newly established subtypes (5/48, 10%) including *TFG::ADGRG7*, *ZEB2::CXCR4*, *CDKN2A::TRAC*, *SPG11::PATL2* and *NKAIN4::PNPLA7* (3/5 in frame) (Figure 2; Table S2).

Overall, we identified six new rearrangements; *PAX5* fused with two novel partners resulting in out-of-frame products (*IGK::PAX5*, *PAX5::IL1RAPL1*), *ETV6::KRT78*, *IGH::HIC1*, *IGH::MIR100HG* and *NKAIN4::PNPLA7*.

Eighty-five per cent of patients ($n = 35/41$) presented a single rearrangement. However, additional fusion events were found in six cases: Three patients harboured *DUX4::IGH* and additionally *PAX5::ZCCHC7*, *IGK::PAX5* or *ZEB2::CXCR4*, one patient presented *P2RY8::CRLF2* and *IGH::CEBPB* fusions, one patient had *IGH::MIR100HG* and *ETV6::BCL2L14* and one patient presented four concomitant fusion genes: *CDKN2A::TRAC*, *P2RY8::CRLF2*, *PAX5::IL1RAPL1* and *SPG11::PATL2* (Table S2).

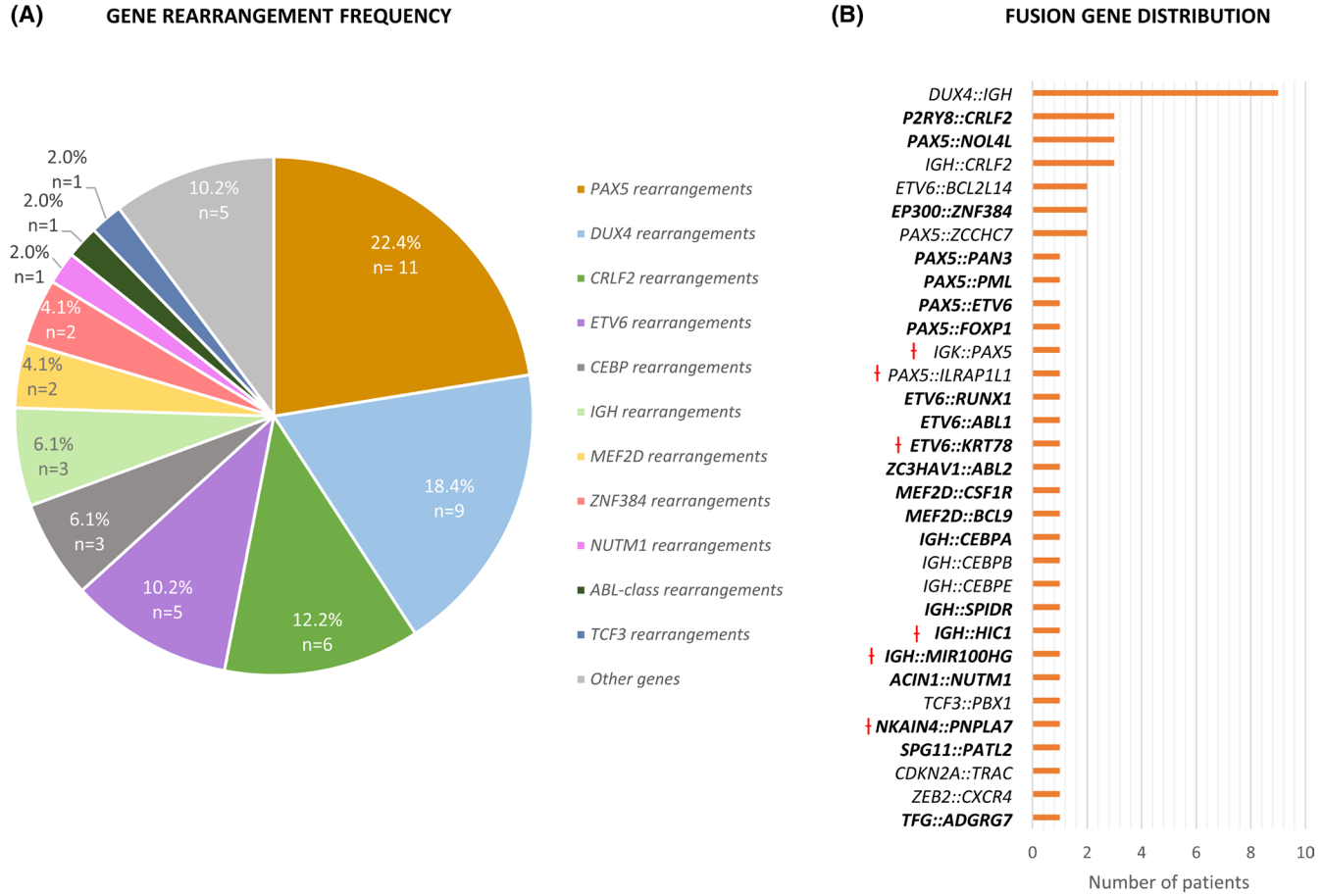

FIGURE 2 Frequency of the gene rearrangements and distribution of the fusion genes identified by RNA-seq. (A) Frequency of the different rearrangements identified. Rearrangements were grouped according to the new emerging genetic subtypes. (B) List of the different rearrangements and the number of patients harbouring them. Red '†' indicates not previously reported rearrangements. In-frame fusions are indicated in bold. The percentages are shown considering the number of total rearrangements ($n = 48$) and not the number of patients. Similarly, the number of fusion genes is not equivalent to the number of patients with a rearrangement identified ($n = 41$), as some of them presented more than one alteration. RNA-seq, RNA-sequencing.

Gene expression profile and its integration with RNA-seq fusion calling

Considering the transcriptional profile, we found a highly distinct GEP corresponding to one of the defined molecular subtypes in 45/60 (75%) B-other ALL patients (Figure 3). Eleven patients were assigned to *PAX5alt*, nine patients to *DUX4*, seven patients to *BCR::ABL1-like*, seven patients to hyperdiploidy, three patients to *ETV6::RUNX1-like*, three patients to *PAX5*-P80R, two patients to *ZNF384*, one patient to *ETV6::RUNX1*, one patient to *MEF2D* and one patient to *NUTM1* subtype. The remaining 15 patients (15/60, 25%) did not match any subtypes and remained unclassified according to their GEP.

Of the 45 patients to whom we assigned a GEP, 29 harboured a fusion gene. We confirmed the subgroup by a matching GEP and fusion gene in 26 of them (26/29, 90%). The fusion gene and the specific GEP for these patients are shown in Table S3. The three discordant patients presented a hyperdiploid GEP but harboured the *IGH::SPIDR*, *PAX5::ZCCHC7* or *P2RY8::CRLF2* and *IGH::CEBPB* rearrangements (Figure 4). The six patients with multiple fusion genes were analysed to determine which alteration potentially correlated better with the GEP. Two patients with *DUX4* rearrangements grouped in the *DUX4* GEP-defined group, while the third patient with a *DUX4* rearrangement (*DUX4::IGH* and *ZEB2::CXCR4*) was placed near this subgroup, between *DUX4* and *ZNF384* clusters. In contrast, for the other three patients with more than one fusion gene, the GEP was inconclusive.

The remaining 16 patients were grouped in a specific BCP-ALL subtype according solely to their GEP (16/45, 36%): Four patients belonged to the *PAX5alt*, three patients to the *PAX5*-P80R subgroup, four patients were clustered in the hyperdiploid group, two patients in the *BCR::ABL1-like*, two patients in the *DUX4* and one patient in the *ETV6::RUNX1-like* group (Figure 4).

Regarding the 15 patients who stood unclassified by the GEP, 12 presented a rearrangement (Figure 4). However, they did not cluster with the expected subtype.

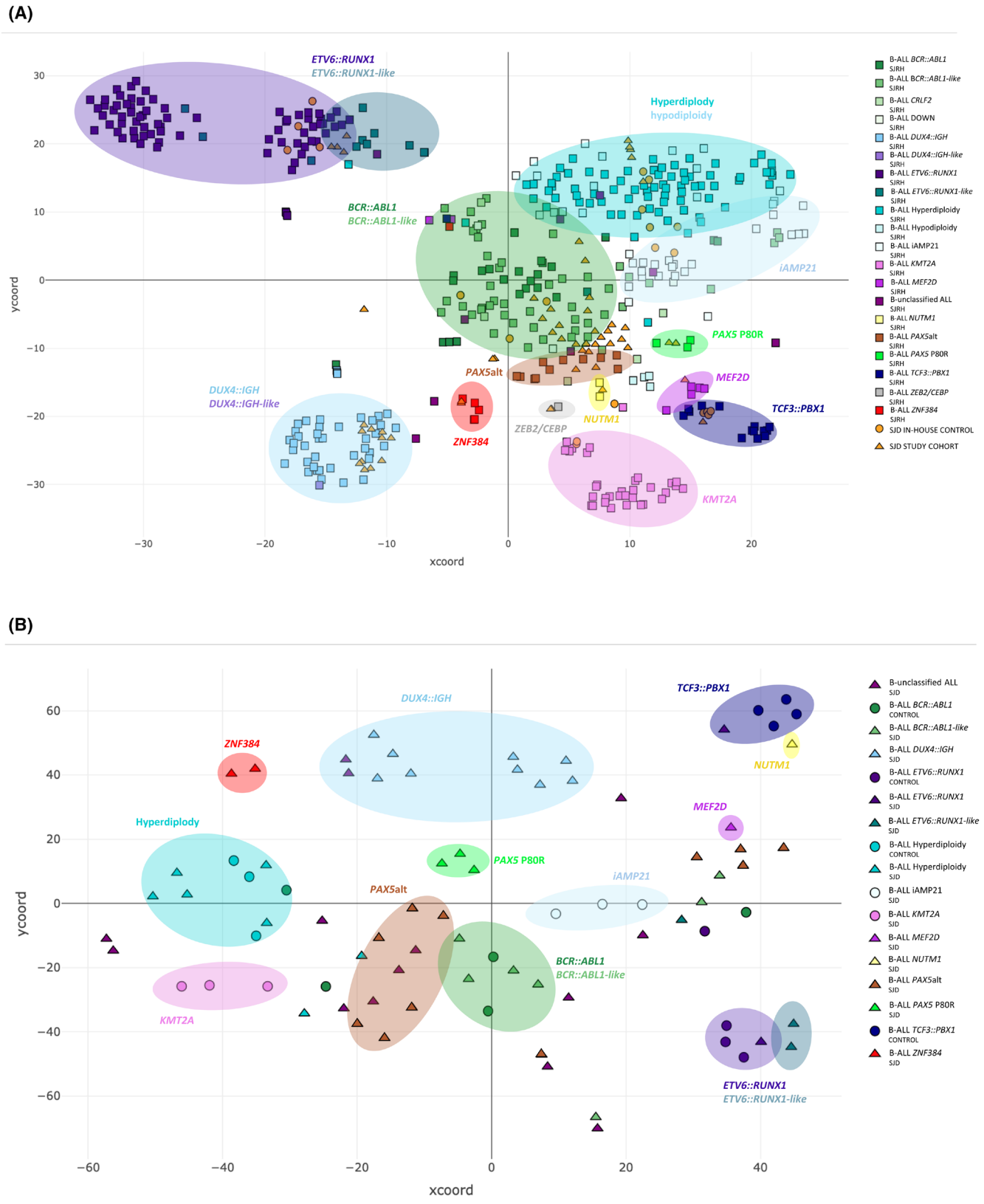

FIGURE 3 t-Distributed stochastic neighbour embedding (t-SNE) plot of BCP-ALL subtypes according to their gene expression profile (GEP). Each dot represents a patient colour-coded according to their subtype (Squares = St Jude Research Hospital, triangles = Study cohort, circles = in-house controls). Coloured circles highlight specific categories. (A) Matrix created using a validated cohort from St Jude Hospital. The 1000 most differential expressed genes were used to differentiate the groups. Twenty-one in-house controls with hyperdiploidy, *ETV6::RUNX1*, *BCR::ABL1*, *TCF3::PBX1* and iAMP21 were included to validate the matrix. Our study cohort (triangles) was distributed in the clusters according to their GEP. (B) Study B-other ALL cohort ($n = 60$) and in-house controls ($n = 22$) distributed according to their GEP. ALL, acute lymphoblastic leukaemia; BCP-ALL, B-cell precursor acute lymphoblastic leukaemia.

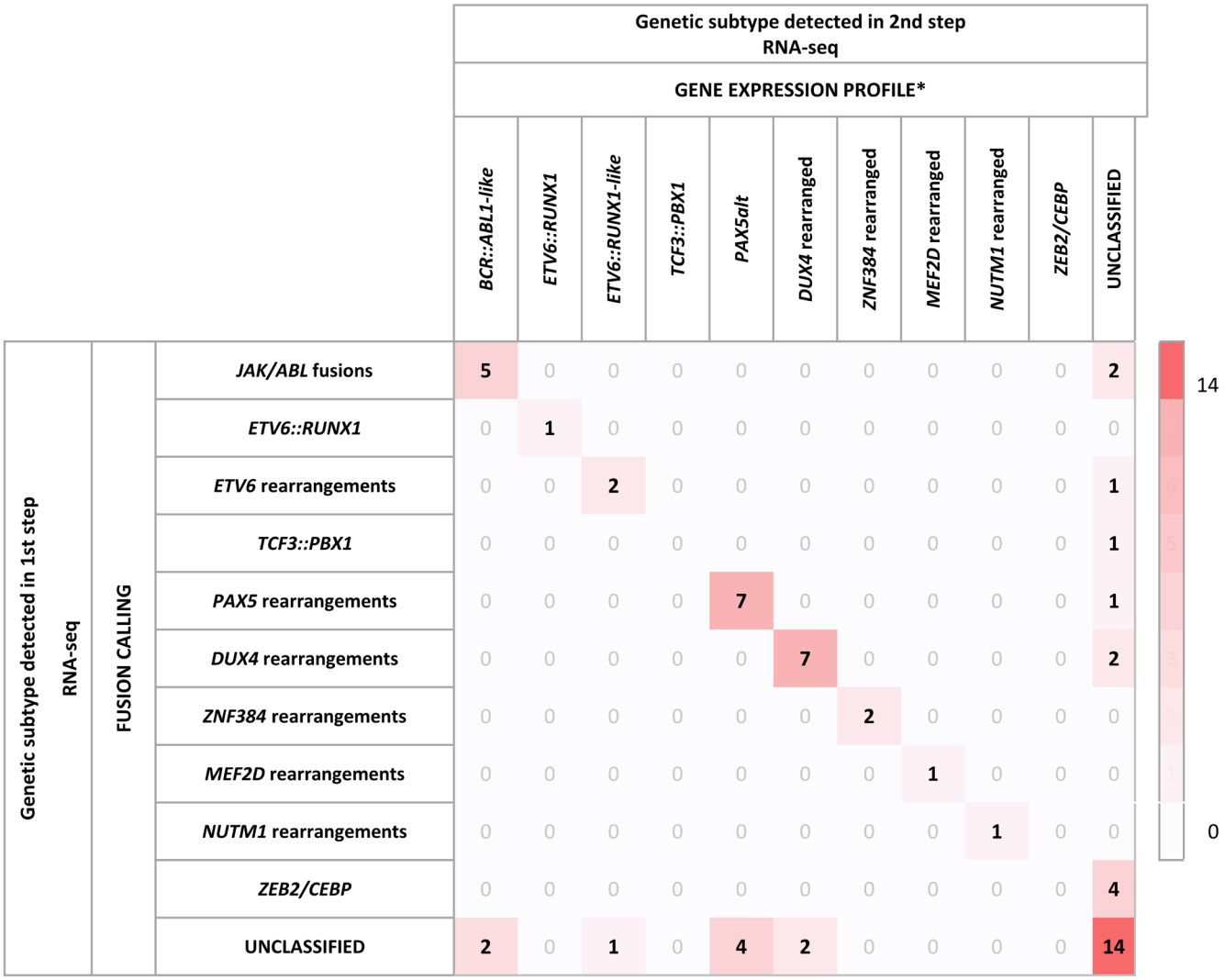

FIGURE 4 Concordance between fusion genes and the gene expression profile identified. The confusion matrix relates the subtype according to the fusion gene identified with the GEP assigned. Diagonal values show the number of concordant patients. **IGH* rearrangements do not have a specific GEP, and *PAX5*-P80R mutation and hyperdiploid cases do not have a specific fusion gene; therefore, they were included in the unclassified group. GEP, gene expression profiling.

RNA-seq data need integration with other methodologies to precisely reclassify B-other ALL patients

By integrating the biological data from standard-of-care (SoC) methods, targeted NGS, ALLCatchR and RNA-seq results, we could classify 43/60 (72%) B-other ALL patients (Figure S1). A detailed description of the integrated data is summarized in Figure 5 and Tables S4–S7. Altogether, *DUX4* ($n=11$) and *PAX5* ($n=9$) were the most common subtypes followed by *BCR::ABL1-like* ($n=7$), *ZEB2/CEBP* ($n=4$), *ETV6::RUNX1-like* ($n=3$), *PAX5*-P80R ($n=3$), *ZNF384* rearrangements ($n=2$), *MEF2D* rearrangements ($n=1$), *NUTM1* fusions ($n=1$), *ETV6::RUNX1* ($n=1$) and *TCF3::PBX1* ($n=1$). Patients without driver genetic lesions or GEP and any other supporting information were finally classified as B-unclassified ($n=17$), representing 28% (17/60) of the B-other ALL patients and 7% (17/255) of the total BCP-ALL cohort (Figure 6).

DUX4 group

We identified 11 *DUX4* cases: seven with both the rearrangement and the GEP, two with the rearrangement and two with the GEP (Figure 5). The two patients with an inconclusive GEP and the two patients without a fusion gene were finally classified as *DUX4* due to other supporting alterations such as *ERG* deletion and strong CD371 cell surface expression (Table S6).[27] Although half of the cases passed all the filters, five were detected in a secondary manual analysis. Regarding the GEP, nine patients clustered together in this molecular subtype profile, with seven showing the rearrangement. Despite

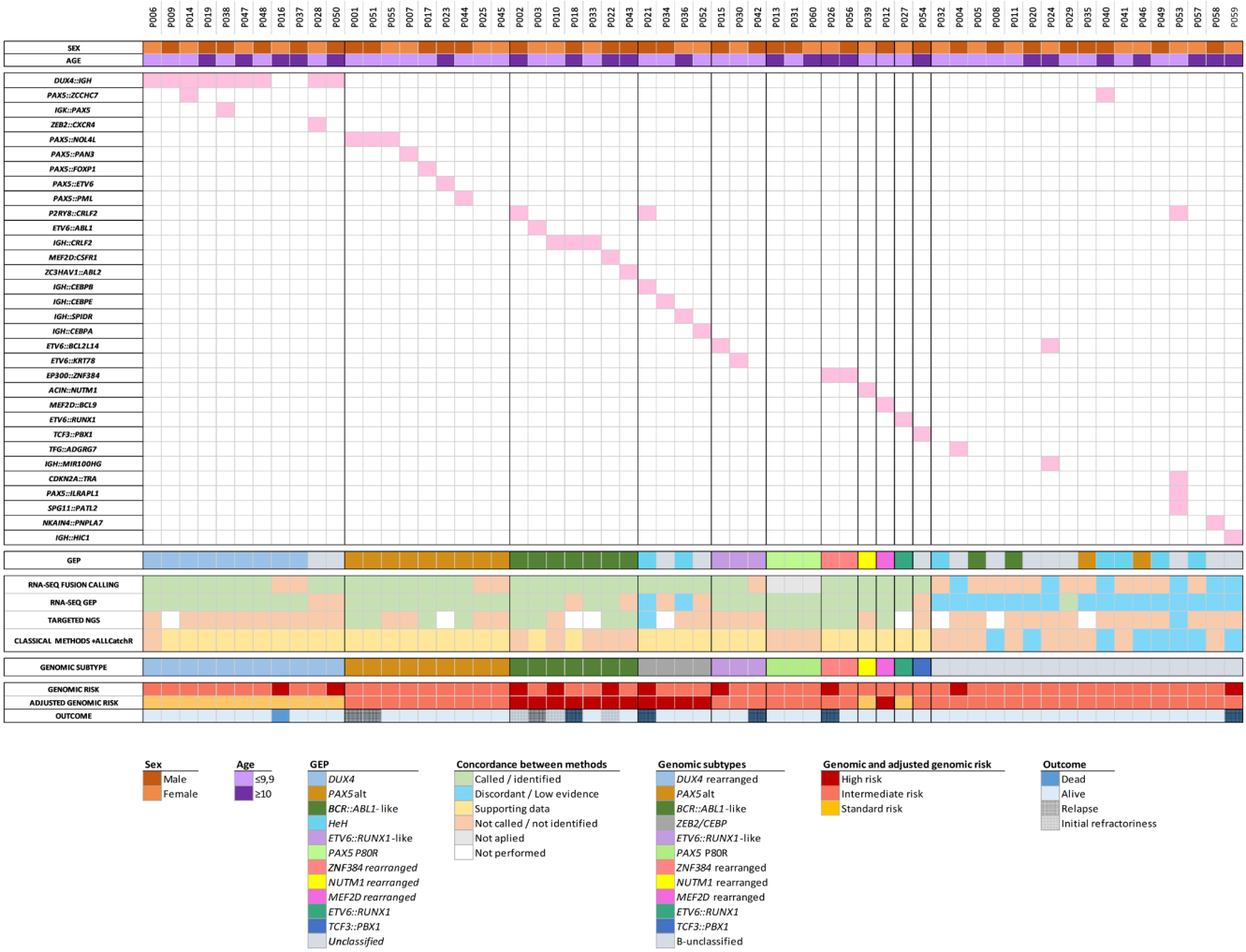

FIGURE 5 Integrated representation of clinical and molecular data in leukaemia. To classify a patient into a subtype, we assessed whether the patient had two strong evidence (green) or one strong evidence and supporting data (green and yellow) from different methodologies. 'Classical methods' included karyotype, FISH, MLPA, flow cytometry and RT-qPCR. Discordant/low evidence: Alteration or GEP not concordant with the subtype assigned or discrepant between them. *Supporting data*: Identified alterations that by themselves do not have enough evidence to classify the patient into a subgroup. Supporting data and discordant data can be consulted in Table S5. FISH, fluorescent in situ hybridization; GEP, gene expression profiling; MLPA, multiplex ligation-dependent probe amplification; RT-qPCR, reverse transcription (quantitative) polymerase chain reaction.

clustering together, two *DUX4* subgroups were observed (Figure 3), as previously described by Li et al.,[28] based on the presence or absence of *ERG* deletion.

PAX5 group

We classified 12 patients in the *PAX5* group: nine patients as *PAX5alt* and three as *PAX5*-P80R. Seven patients harboured a *PAX5* rearrangement with five different partner genes (*NOL4L* [$n=3$], *ETV6*, *PML*, *FOXP1* and *PAN3*, each one case). Most of these rearrangements could be suspected based on the karyotype and supported by *PAX5* deletions observed by MLPA. In all cases, the GEP was concordant (Table S3). Additionally, the GEP classified two more patients in this group without detected fusion, mutation or CNV (Figure 5) and identified three patients with a GEP corresponding to the *PAX5*-P80R subtype, confirmed by targeted NGS (Figure 5; Table S5).

Despite presenting *PAX5* rearrangements (*IGK::PAX5*, *PAX5::ZCCHC7*, *PAX5::IL1RAPL1*), three patients were finally classified into other subgroups (*DUX4* [$n=2$] and B-unclassified, respectively) due to the presence of other fusion genes, other GEP assignation or insufficient supporting evidence.

Other subtypes

Within the less represented groups, *BCR::ABL1-like*, *ZEB2/CEBP* and *ETV6::RUNX1-like* patients accounted for a higher proportion of cases.

BCR::ABL1-like related fusion genes were identified in seven patients. Rearrangements of *JAK–STAT* pathway genes were identified in four patients: *IGH::CRLF2* ($n=3$) and *P2RY8::CRLF2* ($n=1$), and three patients had fusion genes involving *ABL*-class genes (*ABL1*, *ABL2* and *CSF1R*) (Table S5).

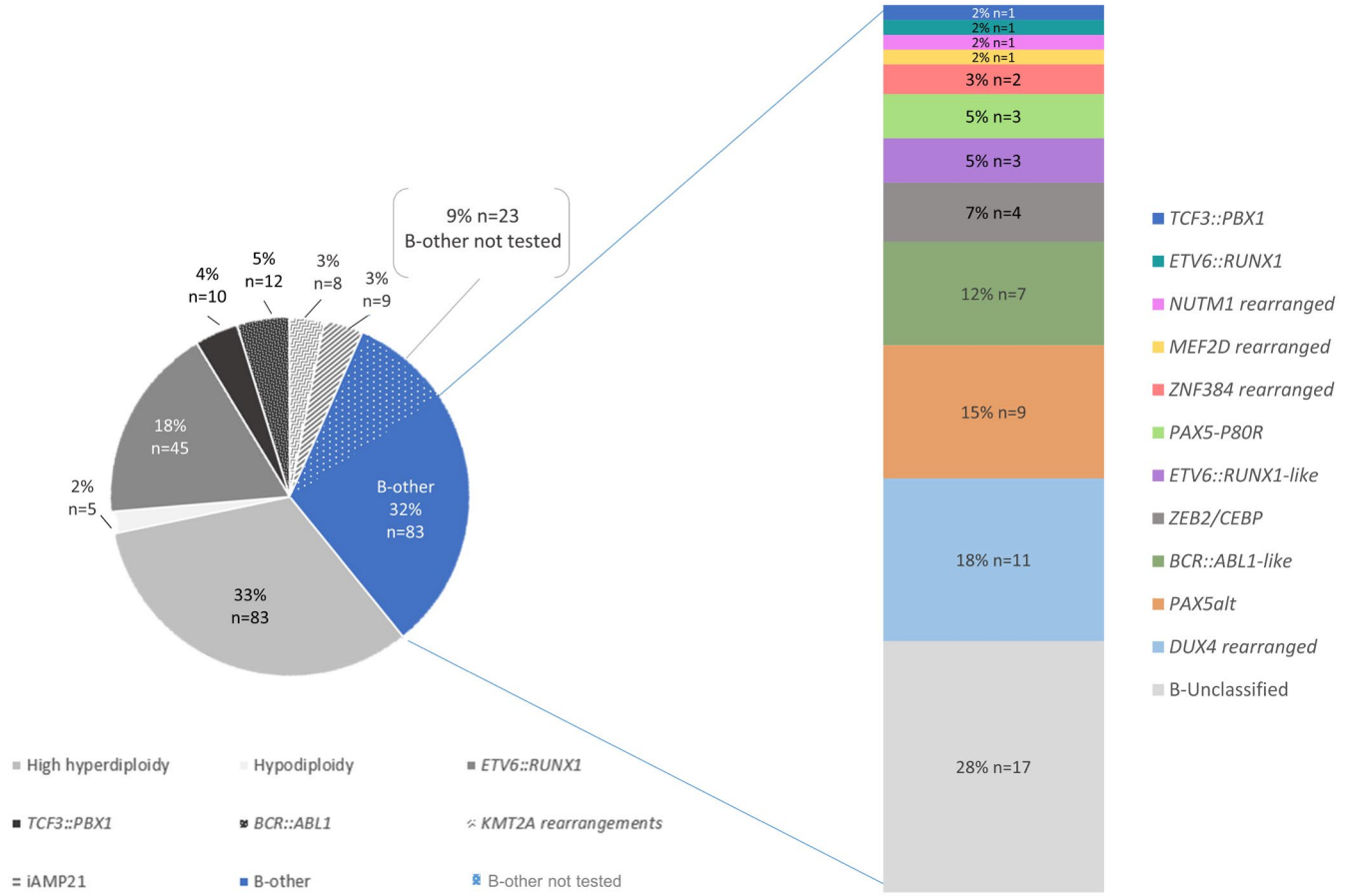

FIGURE 6 Reclassification of the B-other ALL patients after the integration of RNA-seq results and conventional methodologies. The pie chart shows the BCP-ALL patients classification before applying the RNA-seq. The B-other cohort represented 32% (83 patients). Twenty-three of them (23/255, 9%) did not have available samples for the study. Of the 60 patients analysed by RNA-seq, we reclassified $n = 43$ (72%) into one of the genomic subtypes represented in the bar graph at the right part of the figure. ALL, acute lymphoblastic leukaemia; BCP-ALL, B-cell precursor acute lymphoblastic leukaemia; RNA-seq, RNA-sequencing.

The *ZEB2/CEBP* subgroup ($n = 4$) resulted in a very heterogeneous group in which each patient presented *IGH* rearranged with a different partner (*CEBPA*, *CEBPB*, *CEBPE* and *SPIDR*), and the GEP was heterogeneous.

In the *ETV6::RUNX1-like* group ($n = 3$), we identified two patients harbouring different rearrangements: *ETV6::KRT78* and *ETV6::BCL2L14*; one more patient was included in the group due to the assigned GEP and the presence of a deletion in this gene as supporting evidence.

Genetic risk stratification

RNA-seq enabled the identification of genomic abnormalities, facilitating the refinement of patient stratification into distinct risk categories in 25/60 patients (42%). According to the new classification, 13 patients (13/60, 22%) were reclassified as low-risk associated genomic group, including those patients presenting *DUX4*, *NUTM1* and *ETV6::RUNX1* rearrangements. Twelve patients (12/60, 20%) harbouring *CEBP*, *MEF2D* and *BCR::ABL1-like* rearrangements were allocated to the group of genomic abnormalities associated with a higher risk. The remaining 35 patients, including B-unclassified patients among others, remained in the intermediate risk.

DISCUSSION

RNA-seq has allowed the identification of novel alterations hitherto cryptic by conventional methods.[3,17] In our study, we integrated RNA-seq into clinical routine to reclassify B-other ALL patients, combining novel insights with classical methodologies. We evidenced the RNA-seq utility in identifying new alterations, classifying 72% of our B-other cohort into recent genetic subtypes, significantly reducing the B-unclassified group from 32% to 7% of the total BCP-ALL cohort, consistent with recent studies.[29] The identification of *DUX4* and *IGH* rearrangements is challenging, and these subtypes may be underrepresented in our cohort. However, the most represented subgroups were *DUX4* and *PAX5*, followed by *BCR::ABL1-like*, with frequencies similar to those previously reported.[10,11,17,29,30]

We identified 32 different rearrangements involving genes previously described in leukaemia. Notably, we reported for

the first time *IGK::PAX5*, *PAX5::IL1RAPL1*, *ETV6::KRT78*, *IGH::HIC1*, *IGH::MIR100HG* and *NKAIN4::PNPLA7*. These genes have been observed rearranged in other haematological neoplasia and other tumour types,[31–35] suggesting a role in malignant capacity. Functional consequences of fusions vary by reading frame: in-frame fusions preserve the coding sequence and may produce functional chimeric proteins driving oncogenesis. Conversely, out-of-frame fusions have been described as contributors through mechanisms of haploinsufficiency or dominant-negative effects. Functional studies should be performed to better understand the leukaemogenic role of the out-of-frame fusions.[36] Additionally, we identified a patient with an *IGH::SPIDR* rearrangement exhibiting a similar GEP to the *ZEB2/CEBP* subgroup. Considering the *IGH* oncogene-activating mechanisms, the breakpoint within exon 19 of *SPIDR* would argue against the upregulation of the entire gene and suggest instead an *SPIDR* disruption and an upregulation of *CEBPD* via the juxtaposed IGH regulatory elements. However, we observed high expression of both *SPIDR* and *CEBPD* genes, which could explain the similar GEP. Only one case has been previously reported, which was also classified in this subgroup.[37]

New emerging subtypes are sometimes described as phenocopies of established subtypes. These cases, together with *DUX4* or *IGH* rearrangements, can be missed when considering solely fusion calls. In this sense, based on GEP, we identified five patients without rearrangements but with a highly distinctive GEP (two *PAX5alt*, two *DUX4*, and one *ETV6::RUNX1-like*) with secondary supporting alterations that allowed their classification. This highlights the importance of integrating RNA-seq-derived GEPs into diagnostic workflows.

While fusion calling and GEP analysis agreed in 26 patients, discrepancies arose in some cases. Secondary genetic lesions influencing the gene expression can alter the GEP, making it necessary to integrate other biological data to refine the classification. Although RNA-seq provides insights into the mutational landscape or CNVs, it is not considered the gold standard technique; thus, SoC diagnostic tests remain essential, providing fast, reliable data and additional information that RNA-seq might miss.[34] Hence, integrating SoC findings with RNA-seq enhances diagnostic precision.

Improved genomic characterization of B-other patients has identified new subtypes with prognosis[26] and therapeutic implications.[38–40] Based on the literature,[12,19,26] 25 patients (42%) of our cohort would have been re-categorized to a new genetic subtype associated with specific genetic risk. Among these, 12 out of 13 patients reclassified as low-risk were still alive in complete remission at the last follow-up, while 6 of the 12 reclassified into high-risk experienced refractoriness, relapse or death. However, risk stratification in current treatment protocols also depends on other clinic-biological factors.[41] Hence, incorporating these novel subtypes into therapeutic protocols is a gradual process, but recent classifications have begun to include some of them, underlining the importance of detecting these alterations for future treatment protocols.[18,19] Although targeted NGS has been gradually introduced into clinical practice with excellent results and reduced costs,[30] RNA-seq provides additional information about expression profiles unavailable through these approaches, supporting its utility in clinical practice; it not only uncovers cryptic alterations but gives a broad picture of the gene expression, allowing a more accurate classification of patients. However, the turnaround time (TAT) of RNA-seq is a crucial factor for its clinical implementation. Currently, results are available in 5–14 days, but TAT for SoC methodologies is shorter, and some clinical protocols require a faster diagnosis. A diagnostic algorithm could be followed, including the classical techniques to quickly classify the most clinically relevant subtypes needing a fast diagnosis, followed by a second step using RNA-seq for unclassified patients. With further standardization, pipeline improvements and consensus on RNA-seq analysis, a clinically acceptable TAT could be achieved. In this regard, emerging bioinformatics approaches, such as MD-ALL,[42] facilitate the integration of RNA-seq data and could accelerate patient classification.

In summary, we assessed the role of RNA-seq to refine the genomic classification of B-other ALL patients, showing subtype frequencies in a Spanish single-centre cohort. We proposed a model to integrate RNA-seq in real-world clinical practice, combining it with SoC methods to enhance the robustness of novel results. We also demonstrated RNA-seq's benefits as a reliable diagnostic tool, identifying alterations undetectable by SoC techniques. Incorporating RNA-seq into diagnostic algorithms can lead highly accurate B-ALL subtype classification, paving the way to the implementation of precision medicine and a more personalized management of patients.

AUTHOR CONTRIBUTIONS

C.V.-G., E.E.-C., M.C. and N.V.-G. designed the study, and M.C. and N.V.-G. supervised the project. S.R. and J.L.D. recruited patients. S.M., I.I., E.S. and E.C. performed the molecular diagnosis, flow cytometry, morphology and cytogenetics, respectively, and collected the data. C.V.-G. and E.E.-C. performed the RNA-seq libraries and the sequencing process; G.F. performed the bioinformatics analysis. C.V.-G. analysed the data; A.C.-C. and M.R.-F. performed the validation assays by RT-qPCR. C.V.-G. and N.V.-G. wrote the paper with the contribution of the rest of the authors, and all authors reviewed the manuscript and approved the submitted version.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the study participants and their families, the Obra Social from Hospital Sant Joan de Déu and many particular donors for their support in this project. The authors also wish to acknowledge 'Biobanc de l'Hospital Infantil Sant Joan de Déu per a la Investigació', integrated into the Spanish Biobank Network of ISCIII for the sample and data procurement employed in this work.

FUNDING INFORMATION

This study was supported by Instituto de Salud Carlos III grant (PI21/00213) and by associations of parents and families of children with cancer, in coordination with the Obra Social Hospital Sant Joan de Deu.

CONFLICT OF INTEREST STATEMENT

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. The authors have no competing interests.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in the European Genome-phenome Archive (EGA) reference dataset number EGAD50000000981.

ORCID

Clara Vicente-Garcés https://orcid.org/0000-0002-2556-043X
Elena Esperanza-Cebollada https://orcid.org/0000-0001-5451-0370
Mercè Richarte-Franqués https://orcid.org/0000-0001-6033-9880
Susana Rives https://orcid.org/0000-0002-5658-1831
Mireia Camós https://orcid.org/0000-0003-3658-7942
Nerea Vega-García https://orcid.org/0000-0002-3190-063X

REFERENCES

1. World Health Organization. CureAll framework: WHO global initiative for childhood cancer. Increasing access, advancing quality, saving lives. Geneva: World Health Organization; 2021.
2. Iacobucci I, Mullighan CG. Genetic basis of acute lymphoblastic leukemia. J Clin Oncol. 2017;35(9):975–83. https://doi.org/10.1200/JCO.2016.70.7836
3. Brady SW, Roberts KG, Gu Z, Shi L, Pounds S, Pei D, et al. The genomic landscape of pediatric acute lymphoblastic leukemia. Nat Genet. 2022;54(9):1376–89.
4. Roberts KG, Mullighan CG. The biology of B-progenitor acute lymphoblastic leukemia. Cold Spring Harb Perspect Med. 2020;10(7):a034835.
5. Harrison CJ, Moorman AV, Schwab C, Iacobucci I, Mullighan C. Cytogenetics and molecular genetics. In: Vora, A (eds). Childhood acute lymphoblastic leukemia. Cham: Springer; 2017. p. 61–98. https://doi.org/10.1007/978-3-319-39708-5_4
6. Schwab C, Harrison CJ. Advances in B-cell precursor acute lymphoblastic leukemia genomics. Hemasphere. 2018;2(4):e53.
7. Iacobucci I, Kimura S, Mullighan CG. Biologic and therapeutic implications of genomic alterations in acute lymphoblastic leukemia. J Clin Med. 2021;10(17):3792.
8. Balciuniene J, Ning Y, Lazarus HM, Aikawa V, Sherpa S, Zhang Y, et al. Cancer cytogenetics in a genomics world: wedding the old with the new. Blood Rev. 2024;66:101209.
9. Inaba H, Mullighan CG. Pediatric acute lymphoblastic leukemia. Haematologica. 2020;105(11):2524–39.
10. Zaliova M, Stuchly J, Winkowska L, Musilova A, Fiser K, Slamova M, et al. Genomic landscape of pediatric B-other acute lymphoblastic leukemia in a consecutive European cohort. Haematologica. 2019;104(7):1396–406.
11. Ryan SL, Peden JF, Kingsbury Z, Schwab CJ, James T, Polonen P, et al. Whole genome sequencing provides comprehensive genetic testing in childhood B-cell acute lymphoblastic leukaemia. Leukemia. 2023;37(3):518–28.
12. Davis K, Sheikh T, Aggarwal N. Emerging molecular subtypes and therapies in acute lymphoblastic leukemia. Semin Diagn Pathol. 2023;40(3):202–15.
13. Kimura S, Montefiori L, Iacobucci I, Zhao Y, Gao Q, Paietta EM, et al. Enhancer retargeting of CDX2 and UBTF::ATXN7L3 define a subtype of high-risk B-progenitor acute lymphoblastic leukemia. Blood. 2022;139(24):3519–31. https://doi.org/10.1182/blood.2022015444
14. Pagliaro L, Chen SJ, Herranz D, Mecucci C, Harrison CJ, Mullighan CG, et al. Acute lymphoblastic leukaemia. Nat Rev Dis Primers. 2024;10(1):1–28.
15. Cario G, Leoni V, Conter V, Baruchel A, Schrappe M, Biondi A. BCR-ABL1-like acute lymphoblastic leukemia in childhood and targeted therapy. Haematologica. 2020;105(9):2200–4.
16. Zaliova M, Kotrova M, Bresolin S, Stuchly J, Stary J, Hrusak O, et al. ETV6/RUNX1-like acute lymphoblastic leukemia: a novel B-cell precursor leukemia subtype associated with the CD27/CD44 immunophenotype. Genes Chromosomes Cancer. 2017;56(8):608–16. https://doi.org/10.1002/gcc.22464
17. Gu Z, Churchman ML, Roberts KG, Moore I, Zhou X, Nakitandwe J, et al. PAX5-driven subtypes of B-progenitor acute lymphoblastic leukemia. Nat Genet. 2019;51(2):296–307.
18. Pfister SM, Reyes-Múgica M, Chan JKC, Hasle H, Lazar AJ, Rossi S, et al. A summary of the inaugural WHO classification of pediatric tumors: transitioning from the optical into the molecular era. Cancer Discov. 2022;12(2):331–55.
19. Duffield AS, Mullighan CG, Borowitz MJ. International Consensus Classification of acute lymphoblastic leukemia/lymphoma. Virchows Arch. 2023;482(1):11–26. https://doi.org/10.1007/s00428-022-03448-8
20. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. Sequence analysis STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29(1):15–21.
21. Vicente-Garcés C, Maynou J, Fernández G, Esperanza-Cebollada E, Torrebadell M, Català A, et al. Fusion InPipe, an integrative pipeline for gene fusion detection from RNA-seq data in acute pediatric leukemia. Front Mol Biosci. 2023;10:1141310.
22. Putri GH, Anders S, Pyl PT, Pimanda JE, Zanini F. Analysing high-throughput sequencing data in Python with HTSeq 2.0. Bioinformatics. 2022;38(10):2943–5. https://doi.org/10.1093/bioinformatics/btac166
23. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):1–21. https://doi.org/10.1186/s13059-014-0550-8
24. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43(7):e47. https://doi.org/10.1093/nar/gkv007
25. Beder T, Hansen BT, Hartmann AM, Zimmermann J, Amelunxen E, Wolgast N, et al. The gene expression classifier ALLCatchR identifies B-cell precursor ALL subtypes and underlying developmental trajectories across age. Hemasphere. 2023;7(9):E939.
26. Lejman M, Chałupnik A, Chilimoniuk Z, Dobosz M. Genetic biomarkers and their clinical implications in B-cell acute lymphoblastic leukemia in children. Int J Mol Sci. 2022;23(5):2755.
27. Schinnerl D, Mejstrikova E, Schumich A, Zaliova M, Fortschegger K, Nebral K, et al. CD371 cell surface expression: a unique feature of DUX4-rearranged acute lymphoblastic leukemia. Haematologica. 2019;104(8):e352–e355.
28. Li Z, Lee SHR, Chin WHN, Lu Y, Jiang N, Lim EH, et al. Distinct clinical characteristics of DUX4- and PAX5-altered childhood B-lymphoblastic leukemia. Blood Adv. 2021;5(23):5226–38.
29. Hu Z, Kovach AE, Yellapantula V, Ostrow D, Doan A, Ji J, et al. Transcriptome sequencing allows comprehensive genomic characterization of pediatric B-acute lymphoblastic leukemia in an academic clinical laboratory. J Mol Diagn. 2024;26(1):49–60.
30. Migita NA, Jotta PY, do Nascimento NP, Vasconcelos VS, Centoducatte GL, Massirer KB, et al. Classification and genetics of

pediatric B-other acute lymphoblastic leukemia by targeted RNA sequencing. Blood Adv. 2023;7(13):2957–71. https://doi.org/10.1182/bloodadvances.2022009179
31. Cleynen A, Szalat R, Kemal Samur M, Robiou Du Pont S, Buisson L, Boyle E, et al. Expressed fusion gene landscape and its impact in multiple myeloma. Nat Commun. 2017;8(1):1893.
32. Badar T, Johnson L, Trifilo K, Wang H, Kudlow BA, Padron E, et al. Detection of novel t(12;17)(p12;p13) in relapsed refractory acute myeloid leukemia by anchored multiplex PCR(AMP)-based next-generation sequencing. Appl Immunohistochem Mol Morphol. 2019;27(3):E28–E31.
33. Gao Q, Liang WW, Foltz SM, Mutharasu G, Jayasinghe RG, Cao S, et al. Driver fusions and their implications in the development and treatment of human cancers. Cell Rep. 2018;23(1):227–38.e3.
34. Hu X, Wang Q, Tang M, Barthel F, Amin S, Yoshihara K, et al. TumorFusions: an integrative resource for cancer-associated transcript fusions. Nucleic Acids Res. 2018;46:D1144–D1149. https://doi.org/10.1093/nar/gkx1018
35. Ghafouri-Fard S, Harsij A, Farahzadi H, Hussen BM, Taheri M, Mokhtari M. A concise review on the role of MIR100HG in human disorders. J Cell Mol Med. 2023;27(16):2278–89. https://doi.org/10.1111/jcmm.17875
36. Li Z, Qin F, Li H. Chimeric RNAs and their implications in cancer. Curr Opin Genet Dev. 2017;48:36–43.
37. Schwab C, Cranston RE, Ryan SL, Butler E, Winterman E, Hawking Z, et al. Integrative genomic analysis of childhood acute lymphoblastic leukaemia lacking a genetic biomarker in the UKALL2003 clinical trial. Leukemia. 2023;37(3):529–38. https://doi.org/10.1038/s41375-022-01799-4
38. Brivio E, Baruchel A, Beishuizen A, Bourquin JP, Brown PA, Cooper T, et al. Targeted inhibitors and antibody immunotherapies: novel therapies for paediatric leukaemia and lymphoma. Eur J Cancer. 2022;164:1–17.
39. Graff Z, Burke MJ, Gossai N. Novel therapies for pediatric acute lymphoblastic leukemia. Curr Opin Pediatr. 2024;36(1):64–70.
40. Iacobucci I, Roberts KG. Genetic alterations and therapeutic targeting of Philadelphia-like acute lymphoblastic leukemia. Genes (Basel). 2021;12(5):687. https://doi.org/10.3390/genes12050687
41. Malard F, Mohty M. Acute lymphoblastic leukemia. Lancet. 2020;395(10230):1146–62.
42. Hu Z, Jia Z, Liu J, Mao A, Han H, Gu Z. MD-ALL: an integrative platform for molecular diagnosis of B-acute lymphoblastic leukemia. Haematologica. 2024;109(6):1741–54.

SUPPORTING INFORMATION

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

How to cite this article: Vicente-Garcés C, Fernández G, Esperanza-Cebollada E, Richarte-Franqués M, Crespo-Carrasco A, Montesdeoca S, et al. RNA-sequencing: A reliable tool to unveil transcriptional landscape of paediatric B-other acute lymphoblastic leukaemia. Br J Haematol. 2025;206(5):1355–1365. https://doi.org/10.1111/bjh.20056