Novel biological subtypes and clinically important genetic aberrations (druggable lesions, prognostic factors) have been described in B-other acute lymphoblastic leukemia (ALL) during the last decade; however, due to a lack of studies on unselected cohorts, their population frequency and mutual associations still have to be established. We studied 110 consecutively diagnosed and uniformly treated childhood B-other patients using single nucleotide polymorphism arrays and whole exome/transcriptome sequencing. The frequency of DUX4-rearranged, BCR-ABL1-like, ZNF384-rearranged, ETV6-RUNX1-like, iAMP21 and MEF2D-rearranged subtypes was 27%, 15%, 5%, 5%, 4%, and 2%, respectively; 43% of cases were not classified into any of these subtypes (B-rest). We found worse early response to treatment in DUX4-rearranged leukemia and a strong association of ZNF384-rearranged leukemia with B-myeloid immunophenotype. Of the druggable lesions, JAK/STAT-class and RAS/RAF/MAPK-class aberrations were found in 21% and 43% of patients, respectively; an ABL-class aberration was found in one patient. A recently described negative prognostic factor, IKZF1plus, was found in 14% of patients and was enriched in (but not exclusive for) BCR-ABL1-like subtype. PAX5 fusions (including 4 novel), intragenic amplifications and P80R mutations were mutually exclusive and only occurred in the B-rest subset, altogether accounting for 20% of the B-other group. PAX5 P80R was associated with a specific gene expression signature, potentially defining a novel leukemia subtype. Our study shows unbiased European population-based frequencies of novel ALL subtypes, recurrent (cyto)genetic aberrations and their mutual associations. This study also strengthens and widens the current knowledge of B-other ALL and provides an objective basis for optimization of current genetic diagnostics.

Introduction

Knowledge about the genetics of pediatric B-cell precursor (BCP) acute lymphoblastic leukemia (ALL) has grown substantially in the past decade due to the boom in genome-wide technologies. Some of the most important novel findings concern B-other ALL, a heterogeneous group of BCP-ALL with previously unknown genetic backgrounds, usually defined by the absence of all routinely assessed classifying aberrations. Several subtypes have been described within B-other ALL based on the presence of unique (presumably primary) genetic aberrations [iAMP21 ALL1, DUX4 rearranged (r) ALL2, ZNF384r ALL3,5,6, MEF2Dr ALL3,6,7] or gene expression signatures (BCR-ABL1-like ALL8-10 and ETV6-RUNX1-like ALL11-13). Although the classification of BCP-ALL has been substantially refined, it is complicated by a partial overlap of some of these subtypes and/or their imperfect definitions. It has been shown that a significant proportion of iAMP21 ALL have...
the BCR-ABL1-like phenotype or may coincide with the ETV6-RUNX1-like phenotype. Moreover, it has been demonstrated that the definition of BCR-ABL1-like ALL varies significantly among studies; this likely results from the limited specificity of BCR-ABL1-associated gene expression signature. Thus, not only a complex genetic and gene expression characterization but also a hierarchical approach is needed in order to classify BCP-ALL into non-overlapping subsets.

In parallel to aberrations that define subtypes, a wide range of additional genetic aberrations has been identified. Among them, two groups of aberrations have already become therapeutically relevant either as novel prognostic factors or as therapy targets. The first group is represented by IKZF1 gene deletion (IKZF1del) and IKZF1 promoter deletion. IKZF1del was repeatedly shown to confer a higher risk of relapse and is used in therapy algorithm in some contemporary treatment protocols. The second group of therapeutically relevant aberrations largely comprises aberrations affecting genes encoding kinases, cytokine receptors and signaling regulators, resulting in activation of kinase signaling. These aberrations occur especially, but not exclusively, in BCR-ABL1-like ALL. A wide range of kinase aberrations has already been described, and novel ones still continue to be reported. These aberrations vary by type (gene fusions/juxtapositions, copy number changes, mutations) and frequency; CRLF2, NRAS or KRAS lesions are relatively common and occur in various ALL subtypes, while, for example, fusions involving ABL1, ABL2, EPOR, PDGFRB or JAK2 are less frequent and occur predominantly in BCR-ABL1-like ALL. Two functional classes of aberrations, ABL1-class and JAK/STAT-class, are currently used for targeted ALL therapy in some ongoing clinical trials. The ongoing translation of these genetic findings into treatment is bringing new diagnostic demands and a dilemma as to which genetic methods should be incorporated into diagnostic algorithms and how this should be achieved. In order to explore this, well-defined consecutive cohorts need to be analyzed to assess the real incidence of various ALL subtypes and genetic aberrations. Unfortunately, so far, such data have been very scarce, and virtually all the published cohorts were in some way selected, enriched, biased, or not analyzed completely.

Here, we present a genetic study of all pediatric B-other ALL cases consecutively diagnosed and uniformly treated in the Czech Republic over the last seven years. We performed this study to facilitate the use of optimal treatment strategies by extending and refining diagnostic algorithms, and to assess the feasibility and the benefits of introducing genome-wide technologies into routine diagnostics. A further study aim was to determine unbiased population frequencies of known ALL subtypes derived from B-other

<table>
<thead>
<tr>
<th>Table 1. Early response to treatment of B-other acute lymphoblastic leukemia patients stratified by DUX4r.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total B-other (n=110)</strong></td>
</tr>
<tr>
<td>Prednisone response</td>
</tr>
<tr>
<td>Good</td>
</tr>
<tr>
<td>Poor</td>
</tr>
<tr>
<td>FC MRD d15&lt;sup&gt;6&lt;/sup&gt;</td>
</tr>
<tr>
<td>&lt;10%</td>
</tr>
<tr>
<td>≥10%</td>
</tr>
<tr>
<td>NA</td>
</tr>
<tr>
<td>PCR MRD&lt;sup&gt;6&lt;/sup&gt;</td>
</tr>
<tr>
<td>TP1 + TP2 neg.</td>
</tr>
<tr>
<td>Other</td>
</tr>
<tr>
<td>TP1 ±10-3 + TP2 pos. or TP2 ≥10-3</td>
</tr>
<tr>
<td>NA</td>
</tr>
<tr>
<td>Final risk group&lt;sup&gt;7&lt;/sup&gt;</td>
</tr>
<tr>
<td>SR</td>
</tr>
<tr>
<td>MR</td>
</tr>
<tr>
<td>HR</td>
</tr>
<tr>
<td>NA</td>
</tr>
</tbody>
</table>

<sup>1</sup>Good: <1000 leukemia blood blasts/μL on treatment day (d) 8; poor: ≥1000/μL. Minimal residual disease (MRD) at d15 of treatment measured by flow cytometry (FC): MRD measured by polymerase chain reaction (PCR): n: number; TP: time point; TP1: treatment day 33; TP2: treatment day 78. Standard risk (SR), medium risk (MR), high risk (HR); for definitions see Online Supplementary Methods. P-value by Fisher exact test. <sup>2</sup>Patient died before TP2 or no MRD target; neg.: negative; pos.: positive (any positive value); NA: not applicable/not available; Unknown values are not considered for percentage and statistics.
ALL and genetic aberrations, their mutual associations, and associations with demographic and clinical parameters.

Methods

Patients

The present study included 110 children (aged 1-18 years) diagnosed with B-other ALL [BCP-ALL negative for ETV6-RUNX1, BCR-ABL1, TCF3-PBX1, KMT2A-rearrangements, hyperdiploidy (>50 chromosomes), and hypodiploidy (<44 chromosomes)] between December 2010 and December 2017 treated according to the AIEOP-BFM ALL 2009 trial (clinicaltrials.gov identifier: 01117441). In total, 410 BCP-ALL and/or BCP-myeloid mixed phenotype acute leukemia (MPAL) cases were diagnosed and treated according to this protocol within the period of study (Online Supplementary Table S1). The study was approved by the local Institutional Review Board. Informed consent was obtained in accordance with the Declaration of Helsinki.

Analysis of ERG gene deletion and P2RY8-CRLF2 by polymerase chain reaction

Presence of the ERG gene deletion (ERGdel) was analyzed by DNA-based multiplex polymerase chain reaction (PCR). Two primers corresponding to additional centromeric breakpoint sites were added: 5’-GCCGCTACTTGTTGTTCCCAAGAA-3’ and 5’-CTATCCTGACCTTGCTGCGACG-3’. Presence of the P2RY8-CRLF2 transcript was analyzed by real time (RT)-PCR and quantified by quantitative real time (qRT)-PCR as described previously with replacement of the reverse primer for qRT-PCR (5’-AGCCTC CCCA GAAAGACG-3’).

Single-nucleotide polymorphism array

Single-nucleotide polymorphism array (SNPa) was performed in 103 out of 110 B-other patients. Copy number aberrations (CNA) and regions of uniparental disomy (UPD) were identified using HumanOmni Express BeadChip (Illumina, USA) or CytoScan HD arrays (Affymetrix, USA).

Whole-transcriptome and whole-exome sequencing

Whole-transcriptome (RNA-sequencing) and whole-exome (WES) sequencing were performed using diagnostic samples in 109 out of 110 and 69 out of 110 B-other patients, respectively, as described previously. In one patient without material from initial diagnosis, the relapse sample was used for subtype classification. Sequencing libraries were prepared from DNA and total RNA using Agilent SureSelectXT HumanAllExon V4, V5 or V6 and Agilent SureSelect mRNA Strand Specific kits, respectively, according to the manufacturer’s instructions (Agilent Technologies Inc., Santa Clara, CA, USA). High-throughput sequencing was performed on HiSeq2500 (1x50, 2x50 or 2x100 bp) or NextSeq500 (2x75 bp) using TruSeq Rapid SBS and PE Cluster kits and High Output Kit (Illumina).

Selected data metrics are shown in Online Supplementary Tables S2 and S3.

Analysis of fusion transcripts and IGH, DUX4 and CRLF2 rearrangements

RNA-sequencing data were analyzed using TopHat and deFuse. In-frame fusion transcripts and those out-of-frame fusion transcripts that disrupted known ALL-associated genes and/or correlated with boundaries of CNAs found by SNPa were selected for further verification. Fusions verified by RT-PCR as leukemia-spe-
cific (absent in healthy mononuclear blood cells) are reported. To analyze the presence of IGH-DUX4 and IGH-CRLF2 fusion transcripts that may not be revealed by TopHat/deFuse, reads mapped to fusion partner genes were analyzed again manually.

Analysis of genomic variants
Read pairs were aligned to hg19 using BWA35 (WES) and STAR36 (RNA-sequencing) and further processed by Picard tools (http://broadinstitute.github.io/picard/). Variant calling was performed using VarScan37 and Samtools (http://samtools.sourceforge.net/). To distinguish somatic and germline variants, WES results from ALL diagnosis were compared to the remission sample.

Gene expression profiling, hierarchical clustering analysis, DUX4 gene expression analysis
Genome-wide gene expression analysis was performed using RNA-sequencing data of the 110 B-other study patients plus two BCR-ABL1-positive and nine ETV6-RUNX1-positive patients. Alignment and counting were performed as described previously.38 Data normalization and hierarchical clustering analysis (HCA) based on the expression of the most variably expressed genes or of genes belonging to particular gene sets were performed using R package Deseq239 (vst normalization, ward.D method and Euclidean distance linkage for HCA). The list of gene sets and included genes used for HCA are shown in Online Supplementary Tables S4 and S5.

To assess DUX4 gene expression, reads mapped to the DUX4 reference were counted as described previously.3 To correct for uneven sequencing depth, counts were normalized using size factors computed by Deseq2.

Flow cytometry
Routine flow cytometry was used for the diagnosis of BCP-ALL and mixed-phenotype acute leukemia (MPAL) using criteria of the World Health Organization and/or of the European Group for Immunophenotyping of Leukemia. CRLF2 expression was assessed using PE anti-human TSLPR Antibody (clone 1D3; BioLegend, USA). In general, a 50% (10%) cut-off was used to assign strong (weak) positivity. A small CRLF2-positive subclone (<10%), distinct from the major negative population and clearly distinguishable from the technical background, was reported in one patient.

For more details of the methods used see the Online Supplementary Methods.

Results
Frequency of the B-other-derived acute lymphoblastic leukemia subtypes and associations with clinical parameters
To classify B-other patients into iAMP21, ZNF384r, MEF2Dr, DUX4r, ETV6-RUNX1-like and BCR-ABL1-like subtypes, frequencies of specific PAX5 aberrations which were mutually exclusive with these subtypes are shown. Aberrations which are not mutually exclusive with the established subtypes [e.g. CRLF2r, dic(9;20)] are not shown. Infants and BCR-ABL1-positive patients were treated according to different protocols and are not included. Chr: chromosomes. PAX5 fusions do not include ZDHC7-PAX5.
of BCR-ABL1-like and ETV6-RUNX1-like subtypes with respective genetic groups, we amended the cohort with two BCR-ABL1-positive and nine ETV6-RUNX1-positive ALLs. Eighteen patients had the BCR-ABL1-like phenotype according to HCA; two of them had iAMP21 and were not included in BCR-ABL1-like ALL for the subsequent analyses, and the remaining 16 patients were classified as BCR-ABL1-like ALL. Five patients were classified as ETV6-RUNX1-like ALL. Unlike DUX4r and ETV6-RUNX1-positive/-like ALL, BCR-ABL1-like ALL did not co-cluster together in unsupervised HCA, confirming that the gene expression profile of BCR-ABL1(-like) ALL is less unique or specific (Figure 1D).

Overall, the most common B-other ALL subtype was DUX4r ALL (27%), followed by BCR-ABL1-like (15%), ZNF384r (5%), ETV6-RUNX1-like (5%), iAMP21 (4%), and MEF2Dr ALL (2%); the largest proportion of cases ("B-rest"; n=47, 43%) did not belong to any of these subtypes (Figure 2).

Compared to remaining BCP-ALLs, B-others were more frequently males, older, had higher white blood cell (WBC) count, and worse early response to therapy (Online Supplementary Table S1). Analysis of demographic and clinical parameters in the two largest B-others-derived subtypes, DUX4r, BCR-ABL1-like, and B-rest subgroup showed a significantly higher proportion of females in BCR-ABL1-like compared to non-BCR-ABL1-like ALL, while there were no significant differences in age or WBC count, and worse early response to therapy (Online Supplementary Table S1). Importantly, we found a significantly worse early response to treatment assessed by several criteria in the DUX4r ALL compared to non-DUX4r ALL, resulting in a higher proportion of high-risk patients among the DUX4r ALL (Table 1).

In total, 7 out of 110 B-other ALLs had an MPAL of BCR-ABL1-like ALL. Unlike BCR-ABL1 (27%), followed by BCR-ABL1-like (15%), ZNF384r (5%), ETV6-RUNX1-like (5%), iAMP21 (4%), and MEF2Dr ALL (2%); the largest proportion of cases ("B-rest"; n=47, 43%) did not belong to any of these subtypes (Figure 2).

Compared to remaining BCP-ALLs, B-others were more frequently males, older, had higher white blood cell (WBC) count, and worse early response to therapy (Online Supplementary Table S1). Analysis of demographic and clinical parameters in the two largest B-others-derived subtypes, DUX4r, BCR-ABL1-like, and B-rest subgroup showed a significantly higher proportion of females in BCR-ABL1-like compared to non-BCR-ABL1-like ALL, while there were no significant differences in age or WBC count, and worse early response to therapy (Online Supplementary Table S1). Importantly, we found a significantly worse early response to treatment assessed by several criteria in the DUX4r ALL compared to non-DUX4r ALL, resulting in a higher proportion of high-risk patients among the DUX4r ALL (Table 1).
Figure 3. Graphical overview of clinical characteristics, the analyses performed and the genetic findings in 110 B-other acute lymphoblastic leukemia (ALL) patients grouped by ALL subtype. Patients are ordered according to the ALL subtype; demographic and clinical data are shown above, while genetic data are displayed below the ALL subtype track. Recurrently affected genes are arranged according to functional categories. Epigen.: epigenetic regulators/modifiers; MPAL: BCP-myeloid mixed phenotype acute leukemia; Dg WBC: initial white blood cell count (cells/μL); y: years; FC: flow cytometry; SR: standard risk; MR: medium risk; HR: high risk; Neg: negative; Pos: positive; ncounts: normalized counts. For definitions of good and poor prednisone response, polymerase chain reaction (PCR) minimal residual disease (MRD) risk, and final risk group stratification see Table 1, Online Supplementary Methods and Online Supplementary Table S6.
immunophenotype and these were significantly enriched in ZNF384 r ALL [MPALs comprised 67% (4 out of 6) of ZNF384 r vs. 3% (3 out of 104) of non-ZNF384 r ALLs; P<0.0001].

Genomic characterization of the study cohort

Genomic aberrations were analyzed by SNPa, WES, RNA-sequencing and PCR. On average, we found 9 CNAs/UPDs per patient by SNPa (range 0-35) and 14 mutations [single nucleotide variants (SNVs), insertions/deletions (indels)] by WES (range 1-65) (Online Supplementary Tables S7 and S8). In total, we identified 18 recurrently (in ≥2 patients) mutated genes by WES, and 43 recurrent chromosomal imbalances and CNAs subtracted to genes by SNPa; all had already been reported to occur in ALL. In addition to IGH-DUX4, we found 41 different fusion transcripts including 23 that were novel; 24 were in-frame (13 out of 24 novel), 9 contained a full coding sequence of one partner gene (3 out of 9 novel), and 8 were out-of-frame (7 out of 8 novel) (Online Supplementary Table S9). Subclonal deletions of ERG and P2RY8-CRLF2 fusions were detected by PCR in 11 and ten patients negative for these lesions by SNPa and RNA-sequencing, respectively. All recurrent findings are shown together with clinical parameters in Figure 3 and Online Supplementary Table S6.

Specific genetic features of the B-other-derived acute lymphoblastic leukemia subtypes

No significant differences were observed in numbers of CNAs/UPDs or SNVs/indels comparing individual ALL subtypes except for a significantly lower number of CNAs/UPDs in DUX4 r versus remaining ALL cases (4±0.5 vs. 10±0.9; P<0.0001) (Online Supplementary Figure S2). Apart from the MEF2D r and iAMP21 ALL, which were represented by a very low number of patients, each of the remaining subtypes was significantly enriched for specific genetic aberrations (Table 2). ERGdel occurred in 63% of DUX4 r ALL but not in non-DUX4 r ALL patients. Aberrations of TBLXR1 and DMD genes were also enriched in DUX4 r ALL. BCR-ABL1-like ALL was enriched for the CRLF2 fusions, IKZF1 del, both constitutional and somatic trisomy 21, and deletions of EBF1. Mutations in the KMT2D and SETD2 genes (encoding histone methyltransferases) were mutually exclusive and occurred in 5 out of 6 ZNF384 r ALL patients. In agreement with our previous study, all ETV6-RUNX1-like patients harbored an ETV6 aberration (5 out of 5 had a deletion and 2 out of 5 had a fusion). This subtype was also enriched for ARPP21 deletions, deletions of histone gene cluster on 6p22.2, and BTG1 aberrations.

Specific genetic features of the B-rest acute lymphoblastic leukemia subset

The B-rest subset was significantly enriched for the PAX5 fusions, PAX5 amplifications, PAX5 P80R mutation, CDKN2A/B deletions, dic(9;20), and TOX deletions compared to remaining B-other ALLs. It is worthy of note that the PAX5 fusions, amplifications and P80R mutations and TOX deletions occurred exclusively in the B-rest subset; 6 out of 8 TOX deletions occurred together with the PAX5 fusions (Table 2).
Table 3. PAX5-involving fusion genes.

<table>
<thead>
<tr>
<th>UPN</th>
<th>PAX5 exons (NM_016734)</th>
<th>Gene symbol</th>
<th>Reference</th>
<th>Exons</th>
<th>Locus</th>
<th>Fusion-related chromosomal imbalances</th>
</tr>
</thead>
<tbody>
<tr>
<td>1768</td>
<td>1-6</td>
<td>PML</td>
<td>NM_032338</td>
<td>3-9</td>
<td>15q24</td>
<td>None</td>
</tr>
<tr>
<td>1826</td>
<td>1-8</td>
<td>PROC</td>
<td>NM_000312</td>
<td>3-9</td>
<td>2q14</td>
<td>None</td>
</tr>
<tr>
<td>2188</td>
<td>1-5</td>
<td>NOAL</td>
<td>NM_080616</td>
<td>3-8</td>
<td>20q11.21</td>
<td>dic(9;20)</td>
</tr>
<tr>
<td>2319</td>
<td>1-5</td>
<td>CBFA2T3</td>
<td>NM_005187</td>
<td>2-12</td>
<td>16q24</td>
<td>deletion of 3' part of PAX5</td>
</tr>
<tr>
<td>2415</td>
<td>1-7</td>
<td>CBFA2T2</td>
<td>NM_005093</td>
<td>8-12</td>
<td>20q11.21</td>
<td>deletion of 3' part of PAX5 and deletion of 5' part of CBFA2T2</td>
</tr>
<tr>
<td>2544</td>
<td>1-5</td>
<td>CBFA2T3</td>
<td>NM_005187</td>
<td>4-12</td>
<td>16q24</td>
<td>None at PAX5 locus; CBFA2T3 locus unclear (multiple CNA across 16q)</td>
</tr>
<tr>
<td>2548</td>
<td>1-8</td>
<td>PIK3AP1</td>
<td>NM_152309</td>
<td>14-17</td>
<td>10q24</td>
<td>deletion of 5' part of PIK3AP1</td>
</tr>
<tr>
<td>2564</td>
<td>1-7</td>
<td>ESRRB</td>
<td>NM_004452</td>
<td>5-11</td>
<td>14q24</td>
<td>deletion of 3' part of PAX5 and deletion of 5' part of ESRRB</td>
</tr>
<tr>
<td>2578</td>
<td>1-6</td>
<td>MPRIP</td>
<td>NM_015134</td>
<td>11-23</td>
<td>17p11.2</td>
<td>deletion of 3' part of PAX5 and deletion of 5' part of MPRIP</td>
</tr>
<tr>
<td>2621</td>
<td>1-8</td>
<td>DNAJA1</td>
<td>NM_005359</td>
<td>4-9</td>
<td>9p13.3</td>
<td>deletion of 5' part of DNAJA1</td>
</tr>
<tr>
<td>2663</td>
<td>1-8</td>
<td>ARHGAP22</td>
<td>NM_021226</td>
<td>4-10</td>
<td>10q11.23</td>
<td>deletion of 3' part of PAX5</td>
</tr>
<tr>
<td>2701</td>
<td>1-5</td>
<td>ARHGAP22</td>
<td>NM_021226</td>
<td>5-10</td>
<td>10q11.23</td>
<td>deletion of 3' part of PAX5</td>
</tr>
<tr>
<td>2486</td>
<td>2-10</td>
<td>ZCCHC7</td>
<td>NM_032226</td>
<td>1-2</td>
<td>9p21.1</td>
<td>None</td>
</tr>
<tr>
<td>2596</td>
<td>2-10</td>
<td>ZCCHC7</td>
<td>NM_032226</td>
<td>1-2</td>
<td>9p21.1</td>
<td>None</td>
</tr>
</tbody>
</table>

PAX5 fusion partners newly identified in the present study are shown in bold.

In total, we found 11 different PAX5-involving in-frame fusions in 14 patients (Table 3). There were 2 types of fusions (Online Supplementary Figure S3). The more common type was comprised of the 5' part of PAX5 (exons 1-5/6/7/8) with the (partially) absent PAX5 transactivation domain (TAD), fused to various partner genes. This type was present in 12 patients, representing 11% of the total B-other cohort (Figure 2). The second type, represented only by the ZCCHC7-PAX5 fusion [generated by intrachromosomal inversion (Online Supplementary Figure S3)], involved the 3' part of PAX5 including both the DNA-binding domain and complete TAD (exons 2-10). We identified four novel PAX5 partner genes: ARHGAP22 (n=2), MPRIP, PIK3AP1, and PROC. To the best of our knowledge, except for PIK3AP1 (whose deletions were reported in adult ALL), none of these genes have previously been reported as recurrently affected in BCP-ALL. Interestingly, we found rearrangement of the second PIK3AP1 allele resulting in an in-frame PIK3AP1-WDR5 fusion in the PAX5-PIK3AP1-positive patient.

Intragenic PAX5 amplification was found in six B-rest ALL patients, representing 5% of total B-other cohort (Figure 2). In five out of six patients, amplification encompassed exons 2-5 (n=4) or exon 5 (n=1), and thus corresponded to the recurrent type of amplification, PAX5Amp, which was recently characterized in detail; the remaining patient had atypical amplification encompassing exons 4-8 (Online Supplementary Figure S5). RNA-sequencing revealed aberrant PAX5 transcripts containing amplified exons in all six patients (Online Supplementary Figure S6). In line with the previous study, PAX5Amp was accompanied by the trisomy 5 in three out of five patients.

The PAX5 P80R mutation was found in five patients from the B-rest subset who co-clustered tightly together in unsupervised HCA (Figure 1D) and represented 5% of the total B-other cohort (Figure 2). The variant allele frequency (VAF) of PAX5 P80R ranged from 32% to 100%. The second PAX5 allele was deleted/lost in four out of five patients and likely inactivated in the remaining patient, with VAF 52% at the genomic but 98% at the mRNA level. We did not find any other genetic lesion which could explain co-clustering of these patients, which suggests that the specific expression signature was triggered by PAX5 P80R.

Unlike those with PAX5 P80R, patients with PAX5 fusions or PAX5Amp did not form clear clusters in unsupervised HCA; however, via an analysis of differential gene expression, we identified a gene set which distinguished patients with PAX5Amp from the remaining patients in supervised HCA (Online Supplementary Figure S7).

Two and three patients from the B-rest subset co-clustered together with ZNF384r and MEF2Dr ALL subtypes, respectively (Figure 1 and Online Supplementary Figure S7). We did not find any common genetic lesion in patients co-clustering with ZNF384r ALL; however, two out of three patients co-clustering with MEF2Dr ALL harbored the ZNF618-TERT fusion. The NUTM1-gene involving fusions were recently shown to occur recurrently in childhood ALL, outside the established ALL subtypes.

Frequency and distribution of therapeutically relevant aberrations

The IKZF1 pattern was found in 14 out of 103 B-other ALLs (BCR-ABL1-like, n=8; ETV6-RUNX1-like, n=2; B-rest ALLs, n=4) and was significantly enriched in the BCR-ABL1-like subtype (Table 2). It has been shown that IKZF1Amp has a negative prognostic impact only in patients with detectable MRD at the end of induction. All but two
**Discussion**

Although several large genomic studies have been published over the last few years, these were mostly based on non-consecutive cohorts of patients. Here, we present the first European population-based genomic study of pediatric B-other ALL based on a consecutive cohort of uniformly treated patients with a 100% inclusion rate. The only other consecutive study, the Swedish study by Liljebjoern et al., based on 54 B-other ALLs (considering the B-other definition used herein) with an inclusion rate of 69%, showed a similarly high proportion of the “B-rest” ALL (40–45%) and similarly infrequent ZNF384r and MEF2Dr ALL subgroups. Although the frequency of the remaining three subtypes differs between both studies (“ETV6-RUNX1-like ALL 5% vs. 11%, DUX4r ALL 27% vs. 15%, BCR-ABL1-like ALL 15% vs. 28%”), these differences are not statistically significant. Moreover, in the case of BCR-ABL1-like ALL, the difference might be influenced by heterogeneity in this subtype definition, which represents a broader phenomenon and reflects the low specificity of a BCR-ABL1-positive/-like gene expression signature.

Despite the limitations of our analysis of clinical parameters imposed by the small size of the individual subtypes, we do present two important observations. First, we show a strong association of ZNF384r with mixed, B-myeloid immunophenotype. Although this association had already been established, here we demonstrate that ZNF384r ALL represents the majority of all consecutively diagnosed non-infant pediatric B-other B-myeloid MPALS (4 out of 7, 57%). Second, we show a significant association of DUX4r ALL with worse early response to therapy. This ALL subtype is now often considered equivalent to ERGdel-positive ALL. However, an important proportion of DUX4r ALL lacks the ERGdel. We have shown previously that despite the association of ERGdel with worse early therapy response, this did not translate into worse outcome in the MRD-based FMA ALL 2000 protocol. Although two recent studies reported a favorable outcome of DUX4r ALL, its potential heterogeneity with respect to ERGdel has still not been studied. Thus, the outcome of DUX4r ALL in BMF studies, both in general and stratified by ERGdel, still remains to be addressed.

**Supplementary Information**

Additional information related to this article can be found in the supplementary information section, which includes detailed tables and figures to support the discussion and findings presented in the main text.
further specify the prognostic impact of $IKZF1^{del}$ and explain its absence in standard risk patients.\(^{17}\)

One of the most striking findings of our study is the low frequency of kinase/cytokine receptor gene fusions other than those involving $CRLF2$. Except for $CRLF2$ fusions and a single case with $ABL$-class fusion, no other kinase-activating fusions were detected among all B-other patients diagnosed in the Czech Republic over a 7-year period. The low frequency of non-$CRLF2$ fusions, supported also by the result of the Swedish study,\(^{7}\) seems to contrast with some American genomic studies;\(^{21,30}\) however, a direct comparison of frequencies between the European and American populations is not possible at the moment, as only the two European studies refer to consecutive, unselected cohorts. Considering the low costs and the excellent performance of flow cytometry in detecting $CRLF2^{2+}$-positive patients, it remains questionable whether any benefit can be gained from testing the $BCR-ABL1$-like signature in all B-other ALL patients (irrespective of their treatment response) in order to identify patients with a higher probability of having druggable kinase/cytokine receptor gene fusions. Instead, identification of targetable lesions (giving the choice of appropriate drug) can be performed in patients with poor response to treatment upfront and irrespective of potential presence/absence of $BCR-ABL1$-like phenotype.

In our study, we also aimed to assess the feasibility and benefit of incorporating genome-wide technologies into routine diagnostics. Thanks to centralization in a single laboratory, we successfully standardized all processes and minimized the time to integrate results of all analyses. Although we see some benefit in performing all three genome-wide analyses (SNPs, RNA-sequencing, WES) to obtain the most complex genomic profile, considering the cost-benefit ratio, we found the combination of SNPs and RNA-sequencing to be the most efficient. This combination enables the therapeutically most relevant aberrations ($IKZF1^{del}$ and kinase-activating lesions) to be detected. Moreover, RNA-sequencing enables patients to be classified into novel subtypes; this would not be possible with targeted approaches, which are methodologically and analytically less demanding and possibly cheaper. However, these advantages fade with the increasing number of aberrations that need to be tested and, in addition, undescribed/private, but druggable aberrations, cannot always be detected. Furthermore, the complex data gained through genome-wide methods are extremely valuable for retrospective discovery/validation analyses. Nevertheless, it should be noted that genome-wide methods also have additional limitations (e.g. lower sensitivity in fusion transcript detection than PCR), and these need to be considered and tested during the implementation process.

In conclusion, our study shows unbiased European population-based frequency of novel B-other ALL subtypes, recurrent (cyto)genetic aberrations as well as their mutual associations, some of which have not yet been reported. We believe that these findings not only help to strengthen and deepen the current knowledge of B-other ALL, but also provide an objective basis on which other groups/countries can decide how to optimize their current diagnostic routine.

Acknowledgments
We would like to thank to all centers of the Czech Paediatric Haematology Working Group (CPH).

Funding
This study was supported by the “Kapka nadeje” foundation, grants from the Czech Health Research Council (NV15-30626A), Czech Science Foundation (CJ15-06049Y) and Charles University (Primus/MED/28, UNCE 204012) and by the project (Ministry of Health, Czech Republic) for conceptual development of research organization 0064203 (University Hospital Motol, Prague, Czech Republic). Research infrastructure was supported by the Ministry of Education, Youth and Sports (NPÚ I no. LO1404 and LM2015094).

References
M. Zaliova et al.


