haematologica
Journal of The Ferrata Storti Foundation

Molecular characterisation and clinical outcome of B-cell precursor acute lymphoblastic leukemia with IG-MYC rearrangement

by Simon Bomken, Amir Enshaei, Edward C Schwalbe, Aneta Mikulasova, Yunfeng Dai, Masood Zaka, Kent T.M. Fung, Matthew Bashton, Huezin Lim, Lisa Jones, Nefeli Karataraki, Emily Winterman, Cody Ashby, Andishe Attarbaschi, Yves Bertrand, Jutta Bradtke, Barbara Buldini, G. A. Amos Burke, Giovanni Cazzaniga, Gudrun Göhring, Hesta A. de Groot-Kruseman, Claudia Haferlach, Luca Lo Nigro, Mayur Parihar, Adriana Plesa, Emma Seaford, Edwin Sonneveld, Sabine Strehl, Vincent H.J. van der Velden, Vikki Rand, Stephen P. Hunger, Christine J. Harrison, Chris M. Bacon, Frederik W. van Delft, Mignon L. Loh, John Moppett, Josef Vormoor, Brian A. Walker, Anthony V. Moorman, and Lisa J. Russell

Received: January 6, 2022.
Accepted: March 31, 2022.

Citation: Simon Bomken, Amir Enshaei, Edward C Schwalbe, Aneta Mikulasova, Yunfeng Dai, Masood Zaka, Kent T.M. Fung, Matthew Bashton, Huezin Lim, Lisa Jones, Nefeli Karataraki, Emily Winterman, Cody Ashby, Andishe Attarbaschi, Yves Bertrand, Jutta Bradtke, Barbara Buldini, G. A. Amos Burke, Giovanni Cazzaniga, Gudrun Göhring, Hesta A. de Groot-Kruseman, Claudia Haferlach, Luca Lo Nigro, Mayur Parihar, Adriana Plesa, Emma Seaford, Edwin Sonneveld, Sabine Strehl, Vincent H.J. van der Velden, Vikki Rand, Stephen P. Hunger, Christine J. Harrison, Chris M. Bacon, Frederik W van Delft, Mignon L. Loh, John Moppett, Josef Vormoor, Brian A. Walker, Anthony V. Moorman, and Lisa J. Russell. Molecular characterisation and clinical outcome of B-cell precursor acute lymphoblastic leukemia with IG-MYC rearrangement. Haematologica. 2022 Apr 28. doi: 10.3324/haematol.2021.280557. [Epub ahead of print]

Molecular characterisation and clinical outcome of B-cell precursor acute lymphoblastic leukaemia with *IG-MYC* rearrangement

Simon Bomken[1,2]*, Amir Enshaei[1], Edward C Schwalbe[3], Aneta Mikulasova[4], Yunfeng Dai[5], Masood Zaka[6,7], Kent TM Fung[1], Matthew Bashton[8], Huezin Lim[1], Lisa Jones[1], Nefeli Karataraki[1], Emily Winterman[1], Cody Ashby[9], Andishe Attarbaschi[10], Yves Bertrand[11], Jutta Bradtke[12], Barbara Buldini[13], G. A. Amos Burke[14], Giovanni Cazzaniga[15,16], Gudrun Göhring[17], Hesta A de Groot-Kruseman[18,19], Claudia Haferlach[20], Luca Lo Nigro[21], Mayur Parihar[22], Adriana Plesa[23], Emma Seaford[24], Edwin Sonneveld[19], Sabine Strehl[25], Vincent HJ van der Velden[26], Vikki Rand[6,7], Stephen P Hunger[27], Christine J Harrison[1], Chris M Bacon[1,2], Frederik W van Delft[1,2], Mignon L Loh[28], John Moppett[24], Josef Vormoor[1,19], Brian A Walker[29], Anthony V Moorman[1], Lisa J Russell[1]*

[1]Wolfson Childhood Cancer Centre, Translational and Clinical Research Institute, Newcastle University, Newcastle upon Tyne, United Kingdom

[2]The Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK

[3] Department of Applied Sciences, Northumbria University, Newcastle upon Tyne, UK

[4]Biosciences Institute, Newcastle University, Newcastle upon Tyne, UK

[5]Department of Biostatistics, Colleges of Medicine, Public Health and Health Professions, University of Florida, Gainesville, Florida, USA

[6]School of Health and Life Sciences, Teesside University, Middlesbrough, UK

[7]National Horizons Centre, Teesside University, Darlington, UK

[8]The Hub for Biotechnology in the Built Environment, Faculty of Health and Life Sciences, Northumbria University, Newcastle upon Tyne, UK

[9]Department of Biomedical Informatics / Cancer Institute, University of Arkansas for Medical Sciences, Little Rock, Arkansas, USA

[10]St Anna Children's Hospital, Medical University of Vienna, Vienna, Austria

[11]Department of Institute of Hematology Oncology Pediatric (IHOP), Hospices Civils de Lyon, Lyon, France

[12]Institute of Pathology, Department Cytogenetics, University Hospital Giessen and Marburg, Giessen

[13]Maternal and Child Health Department, Padua University. Italy

[14]Department of Paediatric Haematology, Oncology, and Palliative Care, Cambridge University Hospitals NHS Foundation Trust, Addenbrooke's Hospital, Cambridge, UK

[15]School of Medicine and Surgery, University of Milano-Bicocca, Monza, Italy

[16]Centro Ricerca Tettamanti, University of Milano-Bicocca, Monza, Italy

[17]Department of Human Genetics, Hannover Medical School, Hannover, Germany

[18]Dutch Childhood Oncology Group (DCOG), Utrecht, The Netherlands

[19]Princess Máxima Center for Pediatric Oncology, Utrecht, The Netherlands

[20]MLL Munich Leukemia Laboratory, Munich, Germany

[21]Head of Cytogenetic-Cytofluorimetric-Molecular Biology Laboratory, Center of Pediatric Hematology Oncology, Azienda Policlinico "G. Rodolico - San Marco", Catania, Italy

[22]Department of Cytogenetics and Laboratory Haematology, Tata Medical Centre, Kolkata, India

[23]Hematology and Flow cytometry Laboratory, Lyon Sud University Hospital, Hospices Civils de Lyon, Lyon, France

[24]Department of Paediatric Oncology, Bristol Royal Hospital for Children, Bristol, UK

[25]St. Anna Children's Cancer Research Institute, Vienna, Austria

[26]Department of Immunology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands

[27]Department of Pediatrics and the Center for Childhood Cancer Research, Children's Hospital of Philadelphia and the Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA

[28]Department of Pediatrics, Benioff Children's Hospital and the Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, California, USA

[29]Melvin and Bren Simon Comprehensive Cancer Center, Division of Hematology Oncology, Indiana University, Indianapolis, Indiana, USA

*These authors contributed equally to this work

Running head: *IG-MYC*-r B-cell precursor ALL

Corresponding authors: Dr. Lisa J Russell and Dr. Simon Bomken, Wolfson Childhood Cancer Research Centre, Level 6, Herschel Building, Brewery Lane, Newcastle upon Tyne, NE1 7RU, UK.

Tel: +44 (0)191 2082235

Email: lisa.russell@newcastle.ac.uk, s.n.bomken@newcastle.ac.uk

Competing Interests

The authors declare no competing financial interests.

Authorship Contributions

Contribution: SB, LJR, FWvD and AVM conceived and designed the study; SB, YD, AA, YB, JB, BB, GAAB, GC, GG, HAdG, CH, LLN, MP, AP, ESe, ESo, SS, VVHvdV, VR, SPH, CJH, FWvD, MLL,

JV, AVM, LJR provided clinical data and contributed patient samples; SB, ECS, AM, MZ, KTMF, MB, HL, LJ, NK, EW, CA, CMB, BAW, AVM, and LJR performed experiments and analysed experimental data; AE and YD collected and analysed outcome data; SB, LJR and AVM wrote the manuscript with additional contributions from MLL, JM and JV; all co-authors contributed to manuscript review and editing and approved the final manuscript for submission.

Data availability

All previously publicly available datasets used are referred to in the Methods section and in Supplementary Table 3. Methylation data are available from the GEO repository under accession number GSE174248. RNA, whole exome and targeted sequencing data are available from the EGA repository under EGA study ID EGAS00001005111.

Acknowledgements

We would like to thank patients and their families for donating their samples for research and each of the institutional/national BioBanks for providing those samples for this study, including the Blood Cancer UK Childhood Leukaemia Cell Bank. We thank Adele Fielding for help in identifying UK patients.

Word counts
Abstract: 185
Main text: 3890
Tables: 1
Figures: 7
Supplementary data: 1 main file, 1 Excel document (containing 5 Supplementary Tables)

Abstract

Rarely, immunophenotypically immature B-cell precursor acute lymphoblastic leukaemia (BCP-ALL) carries an immunoglobulin-*MYC* rearrangement (*IG-MYC*-r). This can result in diagnostic confusion with Burkitt lymphoma/leukaemia and use of unproven individualised treatment schedules. Here we contrast the molecular characteristics of these conditions and investigate historic clinical outcome data.

We identified 90 cases registered on a national BCP-ALL clinical trial/registry. Where present, diagnostic material underwent cytogenetic, exome, methylome and transcriptome analysis. Outcome was analysed to define 3-year event free survival (EFS) and overall survival (OS).

IG-MYC-r was identified in diverse cytogenetic backgrounds, co-existing with either: established BCP-ALL specific abnormalities (high hyperdiploidy n=3, *KMT2A*-rearrangement n=6, iAMP21 n=1, *BCR-ABL* n=1); *BCL2*/*BCL6*-rearrangements (n=15); or, most commonly, as the only defining feature (n=64). Within this final group, precursor-like V(D)J breakpoints predominated (8/9) and *KRAS* mutations were common (5/11). DNA methylation identified a cluster of V(D)J rearranged cases, clearly distinct from Burkitt leukaemia/lymphoma. Children with *IG-MYC*-r within that subgroup had 3-year EFS of 47% and OS of 60%, representing a high-risk BCP-ALL. To develop effective management strategies this patient group must be allowed access to contemporary, minimal residual disease adapted, prospective clinical trial protocols.

Introduction

Immunoglobulin (IG) driven overexpression of the oncogene *MYC* is the genetic hallmark of mature, germinal centre derived Burkitt lymphoma (BL). However, *IG-MYC* translocations are observed in other mature B-cell malignancies including 5-14% of diffuse large B-cell lymphoma, high-grade B-cell lymphomas with *BCL2* and/or *BCL6*-rearrangements, and multiple myeloma[1-5]. High *MYC* expression is driven by juxtaposition to powerful *IG* super-enhancers, most commonly the heavy chain locus resulting from the translocation, t(8;14)(q24;q32), but alternatively kappa or lambda light chain loci in the translocations t(2;8)(p11;q24) and t(8;22)(q24;q11), respectively.

Much less commonly, *IG-MYC* rearrangements (*IG-MYC*-r) have been identified in lymphoid malignancies expressing a surface immunoglobulin (sIg) negative, immature B-cell precursor (BCP) immunophenotype[6-13]. Several recent acute lymphoblastic leukaemia (ALL) clinical trials have excluded these patients to avoid the risk of mistreating mature B-NHL. However, a large RNA sequencing study recently identified a distinct group of 18/1988 (0.9%) BCP-ALL cases characterised by *IG* translocations with either *BCL2*, *BCL6* and/or *MYC*[14], whilst a Nordic population-based study estimated the frequency of *IG-MYC-r* in childhood BCP-ALL at 0.6%[15].

Despite the rarity, knowing whether to diagnose and treat as an acute lymphoblastic leukaemia or Burkitt leukaemia/lymphoma represents an important question, with the relevant therapies being markedly different. A recent molecular study of 12 patients

demonstrated genetic/epigenetic similarities to BCP-ALL[16]. In contrast, a retrospective clinical study of 14 cases registered with the German BFM-NHL group encouraged treatment according to a mature B-NHL protocol[17]. Clearly there is no current consensus, with many reported patients receiving hybrid protocols outside the clinical trial setting[17,18].

Here, we have studied the molecular and clinical characteristics of a cohort of 90 BCP-ALL cases, enrolled in large national clinical trials/registries. Combining cytogenetic characteristics, somatic mutations, DNA methylation and gene expression analysis, we addressed the question of which disease precursor B cell *IG-MYC*-r leukaemias represent - BCP-ALL or BL. We further sought to determine the prognostic implications of the presence of an *IG-MYC*-r. We found *IG-MYC*-r to be present in diverse cytogenetic backgrounds and defined three sub-groups within our cohort: cases with *IG-MYC-r* and a cytogenetic abnormality recurrently seen in BCP-ALL; cases with *IG-MYC*-r and *BCL2/BCL6* rearrangements; and the majority of cases, with *IG-MYC-r* as the defining cytogenetic abnormality. We demonstrate that *IG-MYC*-r can be either the candidate driver event or a secondary cytogenetic feature. We provide further evidence that the majority are distinct from BL and instead represent a subtype of BCP-ALL with a high risk of early relapse. To improve the care of these patients, clinical outcome should now be assessed in the setting of prospective minimal residual disease (MRD) adapted clinical trials.

Methods

Sample cohort

Patients were identified by the UK Leukaemia Research Cytogenetics Group (LRCG), US Children's Oncology Group (COG), Dutch Children's Oncology Group (DCOG) or International BFM study group members. Full patient characteristics, immunophenotype and additional methodological detail are included in the Supplementary Data, Supplementary Table 1 and 2 and Supplementary Figure 1. The primary inclusion criteria were enrolment/registration on BCP-ALL clinical trial/registry and *IG-MYC*-r detected by karyotyping or FISH as a translocation involving an immunoglobulin locus (*IG*) and *MYC* (*IGH-MYC*-r, t(8;14)(q24;q32); *IGK-MYC*-r, t(2;8)(p11;q24); *IGL-MYC*-r, t(8;22)(q24;q11)). We also identified two patients with T-cell receptor (*TCR*) translocations to the *MYC* locus. Non-trial cases were identified from local/national registries based on a local diagnostic immunophenotype of B-cell precursor malignancy. Children were defined as those patients under the age of 18 years at diagnosis. Informed consent and institutional review board approval was obtained at each collaborating centre. This study was performed in accordance with the Declaration of Helsinki.

Conventional cytogenetics and fluorescence *in situ* hybridisation

Fixed cells were available for 65 *MYC*-r patients (Supplementary Table 1). *IGH*, *IGK*, *IGL*, *MYC*, *BCL2* and *BCL6* fluorescence *in situ* hybridisation (FISH) were performed in all cases with sufficient available cells, in that order of priority. For labelling, capture and scoring

methods see Supplementary Methods.

Multiplex Ligation-dependent Probe Amplification

Copy number alterations were determined by Multiplex Ligation-dependent Probe Amplification (MLPA) using the SALSA MLPA kit P335 (MRC Holland, Amsterdam, Netherlands) in 27 patients[19]. The kit includes probes to *IKZF1*, *CDKN2A/B*, *PAX5*, *EBF1*, *ETV6*, *BTG1*, *RB1* and *CSF2RA/IL3RA/CRLF2*.

Targeted sequencing

DNA from 17 patients was prepared and sequenced using a targeted next generation approach to identify translocation breakpoints at *IG* and *MYC* loci, as described previously[4]. BWA-MEM (v0.7.12) and genome assembly GRCh37 was used for sequence mapping. Chromosomal rearrangements were called using Manta (v.0.29.6)[4,20] and manually inspected in Integrative Genomics Viewer (Broad Institute).

Whole exome sequencing

DNA libraries of 15 diagnostic and 2 relapse samples were generated using the TWIST Human Core Exome kit (Twist Biosciences, San Francisco, CA, USA) and paired end sequencing performed on a NovaSeq (Illumina, San Diego, CA, USA) at the Newcastle University Genomics Core Facility. Data were analysed using the Genome Analysis Toolkit (GATK) including variant caller MuTect2[21-23] (Broad Institute, Cambridge, MA, USA).

Illumina Infinium MethylationEPIC BeadChip array

Methylation data were generated for 18 diagnostic samples at Eurofins Genomics (Ebersberg, Germany) using HumanMethylation Epic arrays (Illumina) and standard manufacturer's protocols. To assess relationships between *IG-MYC*-r leukaemia and other disease entities, we selected the 1404 CpG loci previously reported to be related to B-cell maturation[16] and performed tSNE visualisation of the combined sample set (Supplementary Table 3), as previously reported[24].

RNA-sequencing

RNA-sequencing was performed for six patient samples at Eurofins Genomics (Ebersburg, Germany) using TruSeq Stranded mRNA Library Prep Kit (Illumina). Data were quantified using Kallisto[25] and analysed using DEseq2[26]. These data were compared with publicly available data (EGAS00001001795)[27]. The top 5% of the genes with the highest Median Absolute Deviation (MAD) were selected. The resulting count matrix was used for t-SNE analysis[28].

Outcome analysis

Event free survival (EFS) was defined as time from diagnosis to relapse, second tumour, or death, censoring at last contact. Overall survival (OS) was defined as time from diagnosis to death, censoring at last contact. For patients without data on relapse or second tumour, death was assumed to be the first event. Survival rates were calculated and compared using Kaplan-Meier methods, log-rank tests, and Cox regression models (univariate and

multivariate analyses). Other comparisons were performed using the χ^2 or Fisher's exact test. All tests were conducted at the 5% significance level. All outcome analyses were performed using Intercooled Stata 15.0.

Results

Demographic and clinical characterisation

We collected 90 cases of BCP-ALL diagnosed by local haematologists between August 1989 and July 2019 and subsequently identified as having an *IG-MYC*-r by karyotype or FISH (Supplementary Figure 1, Supplementary Table 1). The majority were registered at diagnosis on a BCP-ALL clinical trial/registry via UKALL (n=32), COG (n=32) or DCOG (n=5) groups. Eighteen additional cases were identified by international BFM study group centres, and eight of these were enrolled on ALL trials. Three further patients in whom *IG-MYC*-r was only demonstrated at relapse were not included in the survival analysis. Patients received a wide variety of trial and clinician determined treatments.

Amongst 87 cases identified at presentation, the median age at diagnosis was 10 years (range 0-81 years) (Table 1). Fifty-two (60%) patients were male, 35 (40%) female. The median presenting white blood cell count was $14x10^9$/L (range 1.6-1200, n=77 patients). Diagnostic bone marrow blast percentage was available for 64 patients ranging from 30%-100% and median of 86.5%. CNS disease was present in 15/64 (23%) cases.

Immunophenotype data were available for 33 diagnostic cases and 1 relapse case

(Supplementary Table 2). TdT was positive in 22/29 (76%) cases. Of TdT negative cases, 7/7 (100%) were slg negative, confirming a BCP immunophenotype.

Cytogenetic characterization identifies both heterogeneity and distinct genomic features

Karyotype was available for 89/90 patients (86/87 at diagnosis) and was abnormal in all cases. Three patients had high hyperdiploidy (HeH, 51-65 chromosomes), a favourable risk feature and eight had high-risk genetics (*KMT2A* rearrangements n=6, iAMP21 n=1, *BCR-ABL1* n=1) (Table 1, Supplementary Table 1). The remaining 75/86 (87%) cases were classified as intermediate cytogenetic risk at diagnosis. Fourteen patients (n=5 children, n=9 adults) had a co-existing *BCL2* rearrangement (*BCL2*-r), one of whom also had a co-existing *BCL6* rearrangement (Supplementary Table 1). In addition, one adult patient had only a *BCL6* rearrangement co-existent with their *IGH-MYC*-r, identified by FISH. In the remaining 64 patients the *IG-MYC*-r was the defining cytogenetic feature.

Karyotyping and/or FISH showed that 47/90 patients (52%) had rearrangement of the *IGH* locus, 39 (43%) the *IGL* locus and three (3%) the *IGK* locus (Figure 1A). Two (2%) patients had *TCR-MYC* rearrangements. One patient had both *IGH-MYC* and *IGL-MYC* rearrangements.

We estimated the percentage of cells carrying *IG-MYC*-r in 53 patients with quantitative FISH data and found that, irrespective of partner gene, *IG-MYC*-r can be either clonal or sub-clonal (Figure 1B, Supplementary Figure 2). Additional evidence for the sub-clonal nature of some *IG-MYC* rearrangements was derived from 3/5 *KMT2A*-r cases (22901, 30611, 17659) in which karyotype and FISH data demonstrated the presence of *KMT2A*-rearranged but

MYC-germline cells, implying the *KMT2A*-rearranged transformation preceded the *MYC*-rearrangement (Figure 1C, Supplementary Table 1). Furthermore, it was possible to reconstruct the clonal evolution of *KMT2A*-r case 17659 from the presentation and relapse metaphase data (Figure 1D, Supplementary Table 1), showing that the relapse was derived from a clone carrying the cooperating *IGL-MYC* rearrangement. Patients with other established ALL-specific cytogenetic abnormalities (HeH, iAMP21) harboured *MYC*-r within the same clone (Supplementary Table 1). Together these data show that, for a proportion of cases at least, *IG-MYC* rearrangements are secondary chromosomal events. Nevertheless, when compared with *MYC*-germline cases of BCP-ALL, a high level of *MYC* expression was identified, irrespective of *IG* partner or degree of clonality (Figure 1E).

In nineteen patients, *IG-MYC*-r was the sole abnormality (Supplementary Table 1). Amongst the remaining cases, abnormalities of chromosome 1 were the most frequently observed aneuploidy/structural gains (28/89, 31%) (Figure 1F, Figure 2, Supplementary Table 1). In the majority (25/28, 89%), 1q abnormalities were shown to be present within the same clone as the *IG-MYC*-r. One specific feature we identified, which has not been recurrently described within either BL[29-32] or BCP-ALL[33], was isochromosome of the long arm of chromosome 1, i(1)(q10). This was observed in nine patients and was associated with *IGH-MYC* in eight.

We further analysed copy number alterations (CNAs) in nine of the most commonly deleted genes/regions in ALL (Figure 3) for twenty seven patients with sufficient DNA. Amongst the fifteen cases with *IG-MYC*-r as the defining cytogenetic abnormality, 11/15 (73%) showed no deletion within the genes/regions analysed. This is remarkably high and even more so when

specifically considering paediatric patients amongst whom 9/10 (90%) had no deletions at diagnosis. This contrasts with just 26% of childhood ALL classified as B-other, the subgroup to which all 10 patients would be assigned[19], implying an important driver role for *IG-MYC*-r in this subgroup of patients.

Finally, in order to investigate the *IG* and *MYC* breakpoints at high resolution, we used a targeted sequencing approach[4] and were able to resolve exact breakpoints for 15/17 patients analysed (Figure 3, Figure 4, Supplementary Table 4). *MYC* breakpoints differed according to *IG* partner locus. *IGH-MYC* breakpoints (n=9) were detected 5' of *MYC*, within the 5' UTR or within intron 1. One case had an additional breakpoint within intron 2 (Figure 4A). Each resulted in *MYC* being juxtaposed to the *IGH* super-enhancers on chromosome 14. In contrast, all *IGL-MYC* breakpoints (n=5) resulted in the translocation of the *IGL* super-enhancer 3' (telomeric) of *MYC* (Figure 4A). In patients with *IGH-MYC*-r, the majority (7/9, 78%) had breakpoints located with the V(D)J gene segments. Notably, just two patients had constant region breakpoints, more commonly associated with germinal centre class switch recombination activity (Figure 4B). One of these cases also carried an *IGH-BCL2*-r (30279). In addition, this analysis identified three patients (26683, 27424 and 29785) to have cryptic *IGH-DUX4* rearrangements[27,34-36], following which they were considered within the group of cases with an established ALL-specific cytogenetic abnormality (Supplementary Table 1).

Together these analyses show significant variability in the cytogenetic characteristics of BCP-ALL with *IG-MYC*-r. In some cases the *IG-MYC* recombination is clearly a secondary chromosomal event but is still capable of imparting a potentially important biological effect.

In other cases it is accompanied by additional *IG* translocations, notably to *BCL2* and/or *BCL6*. However, in the majority of this cohort, *IG-MYC*-r occurred either in isolation or in association with additional cytogenetic features not typical of either BCP-ALL or BL. We therefore sought to investigate the relationship between these disease entities further.

IG-MYC-r BCP-ALLs display diverse genetic, epigenetic and transcriptomic features

In order to better characterise the molecular landscape of *IG-MYC*-r BCP-ALL we analysed exome sequencing data from 15 diagnostic and two relapse samples, interrogating a panel of genes recurrently mutated in BCP-ALL or BL (Figure 3)[14,36-38]. In keeping with a recent analysis of exome data from five similar cases[16], we found non-synonymous variants in *KRAS* in 5/15 (33%) patients at diagnosis. Mutations were also recurrently identified in genes mutated in BCP-ALL including, *FAT4*, *ASMTL* and *IKZF1* observed in 6/15 (40%), 3/15 (20%) and 3/15 (20%) patients, respectively. Additional mutations were observed in *ANK3*, *FAT1*, *FAT2* (2/15, 13%) and *PDGFRA* (1/15, 7%). However, the cohort also included two cases (4352 and 30279) which harboured BL hotspot mutations in *ID3* (L64F) and *TCF3* (N554K) (Supplementary Figure 3), respectively, and additional characteristic BL mutations in *DDX3X*, *SMARCA4*, *CCND3*. Interestingly, case 30279 also had *IGH-BCL2* and *BCL6* rearrangements as have recently been described in a cohort of high grade, molecular Burkitt lymphomas[39].

Wagener *et al* were also able to analyse the methylome of two cases, neither of which carried an *IG-BCL2/BCL6*-r, and found these to be distinct from that of BL, instead clustering with BCP-ALL cases[16]. Here, we integrated 18 of our cases with published DNA methylation

data of BCP-ALL[40-42] and BL[16,43] cases and identified four clusters (Figure 5A). The biggest, in which no *IG-BCL2*/*BCL6*-r or ALL-specific cytogenetic abnormality was identified (Cluster A), contained seven cases (four with *IGH-MYC*-r, two with *IGL-MYC*-r and one with *TRA-MYC*-r) and appeared most closely associated to BCP-ALL with the *TCF3-PBX1* fusion gene. Distinct from cluster A, three patients clustered with cell lines and primary samples from BL patients (Cluster B). These three patients carried either BL-like mutations in *ID3*/*TCF* and *IGH* constant region breakpoints (30279 and 4352) and/or *IG-BCL2*-r (30279 and 25729). Cluster C contained three cases identified by targeted *IG*-sequencing to have *IGH-DUX4*. Patient 30611 had a *KMT2A*-r in 40% of nuclei and clustered with other *KMT2A*-r samples.

Furthermore, transcriptome analysis was conducted for six patients (five with methylation data) for whom RNA was available (Figure 5B&C). Patients with *IG-MYC*-r and no *IG-BCL2*/*BCL6*-r or ALL-specific cytogenetic abnormality again associate, although here they were less close to BCP-ALL samples with *TCF3-PBX1*. Two *IGH-DUX4* patients (27424 and 29785) were confirmed to cluster with published *IGH-DUX4* samples, despite substantially different *MYC*-r frequencies of 16% and 90% respectively, and *KMT2A*-r patient 30611 associated with published *KMT2A*-r patients.

Overall these analyses show that the four cases with an ALL-specific cytogenetic abnormality clustered with the relevant leukaemic subgroup. In our methylome analysis, cases with either an additional *BCL2*/*BCL6*-r, *IGH* constant region break or a mutational profile in keeping with BL, cluster with BL. However, cases without *BCL2*/*BCL6*-r or a BL-like profile form a distinct cluster, strongly indicating they are a subtype of BCP-ALL.

Outcome of patients with *IG-MYC* BCP-ALL

Our genomic analysis identified three groups of *IG-MYC*-r cases, defined by: (1) concomitant *BCL2*/*BCL6*-r; (2) co-occurrence of an established ALL-specific cytogenetic abnormality; (3) *IG-MYC*-r in the absence of another defining cytogenetic abnormality (Figure 6). That analysis demonstrated at least a proportion of *IG-MYC*/*IG-BCL2*/*BCL6*-r "double/triple hit" cases to be biologically distinct from BCP-ALL, instead showing overlap with molecular BL as has recently been described[39]. In keeping with published case reports[44,45], we identified a markedly inferior EFS (HR 3.10, 95% CI 1.42-6.76, p=0.004) and OS (HR 3.78, 95% CI 1.63-8.79, p=0.002) in comparison to *BCL2* wildtype cases. We therefore excluded these cases from subsequent analyses.

For those cases with an ALL-specific cytogenetic abnormality, the *IG-MYC*-r was often sub-clonal and DNA methylation/transcriptomic analysis confirmed they continued to cluster according to the underlying cytogenetic subgrouping. We believe these cases present less of a diagnostic challenge and should continue to be eligible for inclusion in ALL clinical trials.

We therefore further investigated the clinical course of the diagnostically more challenging patients where *IG-MYC*-r represents the only recurrent cytogenetic feature. Amongst 59 cases enrolled on an ALL clinical trial/registry, follow-up data were available from 55 patients. However, only 20 patients (15 children and 5 adults) received the full protocol on trial (Supplementary Figure 1). The remaining patients received unknown off-trial

treatment, which was at the clinical team's discretion. As a result, these off-trial patients had markedly shorter follow-up available, particularly with regard to relapse events (Figure 7). Within the small sub-cohort suitable for analysis, we did not identify a survival difference between children treated on an ALL trial versus those taken off an ALL trial with 3-year EFS (47% (21-69) v 67% (37-85), p= 0.13) and OS (60% (32-80) v 80% (50-93), p=0.15). Although there were long-term survivors among trial and non-trial patients, the predominant feature among both children and adults was a high frequency of early relapse, consistent with high-risk disease (Figure 7A&B). Insufficient MRD was available to allow analysis.

Twenty six children demonstrated a molecular feature identified within our study (*IG-MYC*-r clonality >50%, *KRAS* mutation, abnormal chromosome 1 or i(1)(q10) (Supplementary Figure 4). Univariate analysis did not identify any of these to be associated with EFS or OS (Supplementary Table 5).

Discussion

Although rare, the identification of an *IG-MYC*-r in the karyotype of patients diagnosed with BCP-ALL challenges diagnostic and clinical decision making, with many case reports/series demonstrating the lack of uniform management, often involving hybrid ALL/B-NHL treatment[17,18,45]. Furthermore, given the poor outcome of non-lymphoblastic/Burkitt leukaemia treated with ALL therapy[46,47], several ALL clinical trials have excluded patients with *IG-MYC*-r, restricting the development of effective protocols. To address this clinical need we have collected the largest published cohort of *IG-MYC*-r BCP-ALL cases, from the UK, Europe and the US, studying their molecular biology and clinical outcome.

The principal molecular distinction required is from Burkitt leukaemia/lymphoma, which shares the hallmark cytogenetic feature of an *IG-MYC*-r. However, the majority (approximately 80%) of BL cases have an *IGH-MYC*-r, while the remainder have *IGL* or *IGK* rearrangements in similar proportions[46]. In contrast, we identified a much higher frequency of light chain rearrangements and particularly *IGL-MYC*-r. A substantial proportion of BCP-ALL cases harboured subclonal *IGH* or *IGL* rearrangements whereas in BL *IG-MYC* rearrangement is considered a founder oncogenic event and is therefore clonal. We also identified recurrent i(1)(q10). Whilst gains of 1q are the most common abnormality in BL, i(1)(q10) has not been previously reported in that disease[29-32]. Equally, we identified only two non-*IG-MYC*-r patients with i(1)(q10) within the Leukaemia Research Cytogenetics Group (LRCG) database of >12,000 cases of BCP-ALL and whilst a recently published cohort

of ALL had 2 cases with i(1)(q10) amongst 18 identified as harboring a *BCL2*/*BCL6*/*MYC* rearrangement, only 2 other cases with i(1)(q10) were seen amongst the remaining 1970 cases[14]. Interestingly, isolated cases of immunophenotypically immature BCP-ALL with *IG-MYC*-r and i(1)(q10) have been reported previously[11,16]. Whilst these data support a specific role for i(1)(q10) in BCP-ALL with *IG-MYC*-r +/- *IG-BCL2*-r, the biological implications of this specific finding are unknown. Our mutational analysis of cases confirmed recurrent *KRAS* mutations within this group[14,16] and identified additional recurrent mutations in BCP-ALL related genes *IKZF1* and *ASMTL*.

Targeted sequencing of the *IGH* locus identified breakpoints predominantly within the V(D)J gene segments. This contrasts with germinal centre derived malignancies, including BL, in which just 9-19% of cases carry a V(D)J gene segment breakpoint with rearrangements instead predominantly affecting the constant/switch regions[45-49]. Whilst our DNA methylation/transcriptome analyses were able to cluster cases with their underlying cytogenetic aberration (*KMT2A*-r or *IGH-DUX4*) when present, other *IG-MYC*-r cases with V(D)J gene segment breakpoints formed a separate cluster distinct from BL. In just two cases we identified *IGH* constant region breakpoints, co-occurring with the only examples in our study of typical BL related mutations in *ID3*/*TCF3*, *DDX3X*, *SMARCA4* and *CCND3*[36-38]. Furthermore, one of these two cases also had an *IG-BCL2*-r. DNA methylation analysis clustered all three cases displaying either *IGH* constant region breakpoints and/or *IG-BCL2*-r with BL cases/cell lines, in keeping with the recently described occurrence of *BCL2*-r in adult molecular-BL[39]. However, these findings contrast with a subgroup of *BCL2*/*BCL6*/*MYC*

rearranged ALL which were recently described to cluster together by RNAseq analysis[14]. Whether this relates to alternative experimental approaches/comparators, immunophenotypic inclusion criteria or in fact further demonstrates heterogeneity within this patient group remains unclear. However, the recent publication of the clinical outcome of this subgroup further supports the dismal prognosis associated with *BCL2*/*BCL6*-r ALL[50]. Whether *BCL2*/*BCL6*-r identifies a different disease from the majority of cases of *IG-MYC*-r BCP-ALL remains to be determined but neither published data nor the data presented here support the treatment of these patients with standard BCP-ALL therapy (Figure 7C).

Given the inconsistent treatment protocols and poor follow-up associated with patients not enrolled on a clinical trial, we sought to define patients suitable for ALL therapy for whom inclusion in prospective clinical trials would offer standardised risk-stratified, response-adapted therapy with analysis of long-term follow-up data. The inclusion of cases with BCP-ALL and an ALL-specific cytogenetic abnormality is, we believe, uncontentious. Nonetheless, the presence of a *MYC*-r in these cases is bound to generate questions and therefore clinical trial inclusion/exclusion criteria need to be definitive. More challenging are cases in which *IG-MYC*-r is the only defining cytogenetic abnormality. Historically, patients' outcome has been very poor, but predominantly the very few data which exist predate MRD analysis and response adaptation. With small patient numbers in both the group remaining on an ALL trial and those removed from trial to receive therapy for which the details are not available, we did not observe a significant difference in EFS or OS. Whilst there was a non-significant trend towards improved outcome for those removed from trial protocol therapy, it remains

unclear what the impact of heterogeneous clinician determined therapies was. These are likely to have included ALL or BL protocols but also hybrid treatments delivering intensified early "BL-type" therapy with subsequent addition of ALL maintenance chemotherapy[17,18]. These limited data underscore the need for a prospective analysis of outcome using a standardized protocol.

The rarity of these cases makes it unfeasible to conduct a dedicated trial solely for these patients. Currently they are being lost to investigation and receive unproven therapy resulting in an outcome significantly inferior to standard ALL or BL. We believe that the underlying biology of these *IG-MYC*-r patients indicate that they are part of the ALL spectrum. To improve our understanding of this BCP-ALL subtype and patient outcome, we strongly suggest that they be included in clinical trials able to prospectively measure early treatment response, supporting informed decision making on whether to continue a standard treatment strategy or not. Those with very poor response to initial therapy may be considered for alternative therapies, similar to the approach being taken with acute leukaemia of ambiguous lineage by some study groups[51,52]. Within that setting, use of disease agnostic approaches such as antibody or advanced cellular therapies, may be logical.

In summary, we have demonstrated that surface Ig negative BCP-ALL with *IG-MYC-r* are a high-risk subgroup of BCP-ALL. Optimal treatment can only be determined by removing barriers to treating patients with uniform therapy and collection of detailed long-term clinical follow-up data. We believe that clinical trial protocols should not exclude these patients, instead providing access to contemporary, risk-stratified, response-adapted

therapy.

References

1. Klapper W, Stoecklein H, Zeynalova S, et al. Structural aberrations affecting the MYC locus indicate a poor prognosis independent of clinical risk factors in diffuse large B-cell lymphomas treated within randomized trials of the German High-Grade Non-Hodgkin's Lymphoma Study Group (DSHNHL). Leukemia. 2008;22(12):2226-2229.

2. Savage KJ, Johnson NA, Ben-Neriah S, et al. *MYC* gene rearrangements are associated with a poor prognosis in diffuse large B-cell lymphoma patients treated with R-CHOP chemotherapy. Blood. 2009;114(17):3533-3537.

3. Ott G, Rosenwald A, Campo E. Understanding MYC-driven aggressive B-cell lymphomas: pathogenesis and classification. Blood. 2013;122(24):3884-3891.

4. Mikulasova A, Ashby C, Tytarenko RG, et al. Microhomology-mediated end joining drives complex rearrangements and overexpression of MYC and PVT1 in multiple myeloma. Haematologica. 2020;105(4):1055-1066.

5. Swerdlow SH, Campo E, Harris NL, et al. WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues (ed Revised 4th Edition): International Agency for Research on Cancer (IARC). 2017

6. Navid F, Mosijczuk AD, Head DR, et al. Acute lymphoblastic leukemia with the (8;14)(q24;q32) translocation and FAB L3 morphology associated with a B-precursor immunophenotype: the Pediatric Oncology Group experience. Leukemia. 1999;13(1):135-141.

7. Loh ML, Samson Y, Motte E, et al. Translocation (2;8)(p12;q24) associated with a cryptic

t(12;21)(p13;q22) TEL/AML1 gene rearrangement in a child with acute lymphoblastic leukemia. Cancer Genet Cytogenet. 2000;122(2):79-82.

8. Reid MM, Drewery C, Windebank KP. Surface immunoglobulin-negative acute lymphoblastic leukaemia with predominant L1 morphology, occasional L3 cells and t(8;22). Br J Haematol. 2003;122(5):693.

9. Gupta AA, Grant R, Shago M, Abdelhaleem M. Occurrence of t(8;22)(q24.1;q11.2) involving the MYC locus in a case of pediatric acute lymphoblastic leukemia with a precursor B cell immunophenotype. J Pediatr Hematol Oncol. 2004;26(8):532-534.

10. Meeker ND, Cherry AM, Bangs CD, Kimble Frazer J. A pediatric B lineage leukemia with coincident MYC and MLL translocations. J Pediatr Hematol Oncol. 2011;33(2):158-160.

11. Roug AS, Wendtland P, Bendix K, Kjeldsen E. Supernumerary isochromosome 1, idic(1)(p12), leading to tetrasomy 1q in Burkitt lymphoma. Cytogenet Genome Res. 2014;142(1):7-13.

12. Sato Y, Kurosawa H, Fukushima K, Okuya M, Arisaka O. Burkitt-Type Acute Lymphoblastic Leukemia With Precursor B-Cell Immunophenotype and Partial Tetrasomy of 1q: A Case Report. Medicine (Baltimore). 2016;95(10):e2904.

13. Meznarich J, Miles R, Paxton CN, Afify Z. Pediatric B-Cell Lymphoma With Lymphoblastic Morphology, TdT Expression, MYC Rearrangement, and Features Overlapping With Burkitt Lymphoma. Pediatr Blood Cancer. 2016;63(5):938-940.

14. Gu Z, Churchman ML, Roberts KG, et al. PAX5-driven subtypes of B-progenitor acute lymphoblastic leukemia. Nat Genet. 2019;51(2):296-307.

15. Forestier E, Johansson B, Borgström G, Kerndrup G, Johansson J, Heim S. Cytogenetic findings in a population-based series of 787 childhood acute lymphoblastic leukemias from the Nordic countries. The NOPHO Leukemia Cytogenetic Study Group. Eur J Haematol. 2000;64(3):194-200.

16. Wagener R, Lopez C, Kleinheinz K, et al. IG-MYC (+) neoplasms with precursor B-cell phenotype are molecularly distinct from Burkitt lymphomas. Blood. 2018;132(21):2280-2285.

17. Herbrueggen H, Mueller S, Rohde J, et al. Treatment and outcome of IG-MYC(+) neoplasms with precursor B-cell phenotype in childhood and adolescence. Leukemia. 2020;34(3):942-946.

18. Zhang C, Amos Burke GAA. Pediatric precursor B-cell acute lymphoblastic leukemia with MYC 8q24 translocation - how to treat? Leuk Lymphoma. 2018;59(8):1807-1813.

19. Schwab CJ, Chilton L, Morrison H, et al. Genes commonly deleted in childhood B-cell precursor acute lymphoblastic leukemia: association with cytogenetics and clinical features. Haematologica. 2013;98(7):1081-1088.

20. Chen X, Schulz-Trieglaff O, Shaw R, et al. Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. Bioinformatics. 2016;32(8):1220-1222.

21. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics. 2010;26(5):589-595.

22. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010;20(9):1297-1303.

23. Cibulskis K, Lawrence MS, Carter SL, et al. Sensitive detection of somatic point mutations in

impure and heterogeneous cancer samples. Nat Biotechnol. 2013;31(3):213-219.

24. Sharma T, Schwalbe EC, Williamson D, et al. Second-generation molecular subgrouping of medulloblastoma: an international meta-analysis of Group 3 and Group 4 subtypes. Acta Neuropathol. 2019;138(2):309-326.

25. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. Nat Biotechnol. 2016;34(5):525-527.

26. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.

27. Lilljebjorn H, Henningsson R, Hyrenius-Wittsten A, et al. Identification of ETV6-RUNX1-like and DUX4-rearranged subtypes in paediatric B-cell precursor acute lymphoblastic leukaemia. Nat Commun. 2016;7:11790.

28. van der Maaten L, Hinton G. Visualising data using t-SNE. J Mach Learn Res. 2008;9(86):2579-2605.

29. Newman AM, Zaka M, Zhou P, et al. Genomic abnormalities of *TP53* define distinct risk groups of paediatric B-cell non-Hodgkin lymphoma. Leukemia. 2022;36(3):781-789.

30. Schiffman JD, Lorimer PD, Rodic V, et al. Genome wide copy number analysis of paediatric Burkitt lymphoma using formalin-fixed tissues reveals a subset with gain of chromosome 13q and corresponding miRNA over expression. Br J Haematol. 2011;155(4):477-486.

31. Scholtysik R, Kreuz M, Klapper W, et al. Detection of genomic aberrations in molecularly defined Burkitt's lymphoma by array-based, high resolution, single nucleotide polymorphism

analysis. Haematologica. 2010;95(12):2047-2055.

32. Havelange V, Pepermans X, Ameye G, et al. Genetic differences between paediatric and adult Burkitt lymphomas. Br J Haematol. 2016;173(1):137-144.

33. Moorman AV, Ensor HM, Richards SM, et al. Prognostic effect of chromosomal abnormalities in childhood B-cell precursor acute lymphoblastic leukaemia: results from the UK Medical Research Council ALL97/99 randomised trial. Lancet Oncol. 2010;11(5):429-438.

34. Yasuda T, Tsuzuki S, Kawazu M, et al. Recurrent DUX4 fusions in B cell acute lymphoblastic leukemia of adolescents and young adults. Nat Genet. 2016;48(5):569-574.

35. Liu YF, Wang BY, Zhang WN, et al. Genomic Profiling of Adult and Pediatric B-cell Acute Lymphoblastic Leukemia. EBioMedicine. 2016;8:173-183.

36. Zhang J, McCastlain K, Yoshihara H, et al. Deregulation of DUX4 and ERG in acute lymphoblastic leukemia. Nat Genet. 2016;48(12):1481-1489.

36. Love C, Sun Z, Jima D, et al. The genetic landscape of mutations in Burkitt lymphoma. Nat Genet. 2012;44(12):1321-1325.

37. Schmitz R, Young RM, Ceribelli M, et al. Burkitt lymphoma pathogenesis and therapeutic targets from structural and functional genomics. Nature. 2012;490(7418):116-120.

38. Richter J, Schlesner M, Hoffmann S, et al. Recurrent mutation of the ID3 gene in Burkitt lymphoma identified by integrated genome, exome and transcriptome sequencing. Nat Genet. 2012;44(12):1316-1320.

39. Bouska A, Bi C, Lone W, et al. Adult high-grade B-cell lymphoma with Burkitt lymphoma

signature: genomic features and potential therapeutic targets. Blood. 2017;130(16):1819-1831.

40. Nordlund J, Bäcklin CL, Wahlberg P, et al. Genome-wide signatures of differential DNA methylation in pediatric acute lymphoblastic leukemia. Genome Biol. 2013;14(9):r105.

41. Gabriel AS, Lafta FM, Schwalbe EC, et al. Epigenetic landscape correlates with genetic subtype but does not predict outcome in childhood acute lymphoblastic leukemia. Epigenetics. 2015;10(8):717-726.

42. Bergmann AK, Castellano G, Alten J, et al. DNA methylation profiling of pediatric B-cell lymphoblastic leukemia with KMT2A rearrangement identifies hypomethylation at enhancer sites. Pediatr Blood Cancer. 2017;64(3):e26251.

43. Hernandez-Vargas H, Gruffat H, Cros MP, et al. Viral driven epigenetic events alter the expression of cancer-related genes in Epstein-Barr-virus naturally infected Burkitt lymphoma cell lines. Sci Rep. 2017;7(1):5852.

44. Liu W, Hu S, Konopleva M, et al. De Novo MYC and BCL2 Double-hit B-Cell Precursor Acute Lymphoblastic Leukemia (BCP-ALL) in Pediatric and Young Adult Patients Associated With Poor Prognosis. Pediatr Hematol Oncol. 2015;32(8):535-547.

45. Sakaguchi K, Imamura T, Ishimaru S, et al. Nationwide study of pediatric B-cell precursor acute lymphoblastic leukemia with chromosome 8q24/MYC rearrangement in Japan. Pediatr Blood Cancer. 2020;67(7):e28341.

46. Anderson JR, Wilson JF, Jenkin DT, et al. Childhood non-Hodgkin's lymphoma. The results of a randomized therapeutic trial comparing a 4-drug regimen (COMP) with a 10-drug regimen (LSA2-

L2). N Engl J Med. 1983;308(10):559-565.

47. Moorman AV, Harrison CJ, Buck GAN, et al. Karyotype is an independent prognostic factor in adult acute lymphoblastic leukemia (ALL): analysis of cytogenetic data from patients treated on the Medical Research Council (MRC) UKALLXII/Eastern Cooperative Oncology Group (ECOG) 2993 trial. Blood. 2007;109(8):3189-3197.

45. Grande BM, Gerhard DS, Jiang A, et al. Genome-wide discovery of somatic coding and non-coding mutations in pediatric endemic and sporadic Burkitt lymphoma. Blood. 2019;133(12):1313-1324.

46. Gelmann EP, Psallidopoulos MC, Papas TS, Dalla Favera R. Identification of reciprocal translocation sites within the c-myc oncogene and immunoglobulin mu locus in a Burkitt lymphoma. Nature. 1983;306(5945):799-803.

47. Busch K, Keller T, Fuchs U, et al. Identification of two distinct *MYC* breakpoint clusters and their association with various *IGH* breakpoint regions in the t(8;14) translocations in sporadic Burkitt-lymphoma. Leukemia. 2007;21(8):1739-1751.

48. López C, Kleinheinz K, Aukema SM, et al. Genomic and transcriptomic changes complement each other in the pathogenesis of sporadic Burkitt lymphoma. Nat Commun. 2019;10(1):1459.

49. Bendig S, Walter W, Meggendorfer M, et al. Whole genome sequencing demonstrates substantial pathophysiological differences of *MYC* rearrangements in patients with plasma cell myeloma and B-cell lymphoma. Leuk Lymphoma. 2021;62(14):3420-3429.

50. Paietta E, Roberts KG, Wang V, et al. Molecular classification improves risk assessment in

adult BCR-ABL1-negative B-ALL. Blood. 2021;138(11):948-958.

51. Oberley MJ, Raikar SS, Wertheim GB, et al. Significance of minimal residual disease in pediatric mixed phenotype acute leukemia: a multicenter cohort study. Leukemia. 2020;34(7):1741-1750.

52. Orgel E, Alexander TB, Wood BL, et al. Mixed-phenotype acute leukemia: A cohort and consensus research strategy from the Children's Oncology Group Acute Leukemia of Ambiguous Lineage Task Force. Cancer. 2020;126(3):593-601.

Tables

		Overall	Children	Adults
		87	66	21
Sex	Female	35 (40)	26 (39)	9 (43)
	Male	52 (60)	40 (61)	12 (57)
Age (years)	Median	10	6	52
	Range	0-81	0-17	20-81
WCC	>50 x 10^9/L	13/74 (18)	10/57 (18)	3/17 (18)
	Median	14	14	19
	Range	1.6-1200	1.6-1200	2-131
CNS status	Pos	15 (23)	12 (22)	3 (30)
	Neg	49 (77)	42 (78)	7 (70)
	NA	23	12	11
Cytogenetic risk group	GR	3 (3)	1 (1)	2 (10)
	IR	75 (86)	57 (86)	18 (86)
	HR	8 (9)	7 (11)	1 (5)
	NA	1	1	0

Table 1. Clinical characteristics of patients presenting at diagnosis with BCP-ALL carrying an *IG-MYC*-r. WCC, white blood cell count. CNS, central nervous system. Pos, positive. Neg, negative. GR, good risk. IR, intermediate risk. HR, high risk. NA, not available. Numbers in parenthesis are a percentage.

Figure Legends

Figure 1. Cytogenetic characterization of *IG-MYC*-r patients.

(A) Karyotype reveals the distribution of immunoglobulin chain involvement in *MYC* translocations (FISH data were used for one case who presented with a normal karyotype). +*BCL2/6*-r, concomitant *BCL2/BCL6*-rearrangement. +ALL-r, concomitant established ALL rearrangement. *IG-MYC*, *IG-MYC* rearrangement as the defining cytogenetic abnormality. (B) Percentage of nuclei carrying *IG-MYC*-r grouped by immunoglobulin chain involvement. Blue dots, *IG-MYC*-r alone. Green dots, *IG-MYC*-r and established ALL rearrangement. Red dots, *IG-MYC*-r and *BCL2*/*BCL6*-r. Dotted and dashed lines, interquartile ranges. (C) Percentage rearrangement of MYC (blue dots) and KMT2A (green dots) using FISH. (D) Evolution of t(4;11) and subsequent gain of t(8;22) in case 17659. Percentages at diagnosis represent the proportion of metaphases seen with each abnormality (left panel). Representative chromosomes taken from diagnostic and relapse metaphases (right panel). Grey arrows mark the portion of chromosome 22 translocated to chromosome 8. (E) RNA sequencing comparing expression of MYC amongst (labelled) cases from *IG-MYC*-r cohort with other BCP-ALL cases in house. TPM, transcripts per million reads. (F) Chromosomal abnormalities observed for chromosome 1 in the karyotypes of patients with *IG-MYC*-r. Red line - patients with concomitant *BCL2*-r.

Figure 2. Ideogram of cytogenetic rearrangements reported in the karyotypes of 70 patients with *IG-MYC*-r.

Figure 3. Oncoplot depicting the incidence of selected genomic features of *IG-MYC*-r patients with material available for genomic studies. D - diagnosis; R - relapse; dark shaded square, positive result; light shaded square, negative result; white square, not tested; T, 1q translocation.

Figure 4. Targeted *IG* and *MYC* sequencing identifies heterogeneous breakpoints

(A) Distribution of breakpoints within the *MYC* locus. (B) Distribution of breakpoints within the *IGH* locus. Each line shows the breakpoints for patients with translocations involving individual genomic loci. Frequency distribution defines a region of increased frequency of breaks (peach shaded area). Upper panels provide an expanded view of the breakpoint hotspots. Each dot represents an individual breakpoint. Blue dots, *IG-MYC*-r alone. Green dots, *IG-MYC*-r and established ALL rearrangement. Red dots, *IG-MYC*-r and *BCL2*/*BCL6*-r.

Figure 5. Methylation and transcriptome data identify five clusters within the *IG-MYC*-r cohort.

(A) Consensus clustering of methylation data. Common *t*-SNE visualisation of *IG-MYC*-r patients combined with publicly available data for subtypes of B-ALL (GSE49031, GSE69229, GSE76585) and BL patient (GSE114210) and cell line samples (GSE92378). Each dot represents an individual patient.

(B) Consensus clustering of RNA-seq data. Common *t*-SNE visualisation of *IG-MYC*-r patients combined with publicly available RNA-seq data for subtypes of B-ALL (EGAS00001001795). Each dot represents an individual patient. (C) Patient IDs assigned to clusters. *, Methylation and RNA-seq data, **RNA-seq only.

Figure 6. Molecular characteristics of *IG-MYC*-r ALL cases. Summary of three groups defined within the molecular analysis. Cases with a co-existing ALL-specific cytogenetic abnormality are shown in green in the left panel. Cases with an *IG-MYC*-rearrangement (*IG-MYC*-r) as the only recurrent cytogenetic feature are shown in blue in the centre panel and cases with a co-existing rearrangement of *BCL2* and/or *BCL6* are shown in red in the right panel. Each donut plot shows the proportion of cases studied which have a specific molecular feature. The total number of cases in each study is shown within each donut. The upper donut plots show the proportion with variable *V(D)J* gene segment breakpoints (dark segments) and constant segment breakpoints (light segments). The middle donut plots show the proportion of cases with either Burkitt-like *TCF3*/*ID3* mutations (light segments), *KRAS* mutations (medium segments) or neither such mutation (dark segments). The lower donut plots show the proportion of cases within each of three DNA methylation clusters, either Burkitt-like cluster B (light segments), the *IG-MYC*-r cluster A (medium segments) or one of the other ALL specific clusters (dark segments). *The single case in the centre panel harbouring both a constant region breakpoint and an *ID3* mutation (case 4352) clustered with Burkitt lymphoma/leukaemia cases within our DNA methylation analysis.

Figure 7. Swimmer plots displaying event free survival and overall survival. (A) Children and (B) adults with *IG-MYC*-r but no other ALL-specific cytogenetic abnormality. Suggested approach for the assignment of patients to a treatment strategy (C). + For patients without data on relapse or second tumour, death was assumed to be the first event.

Figure 1

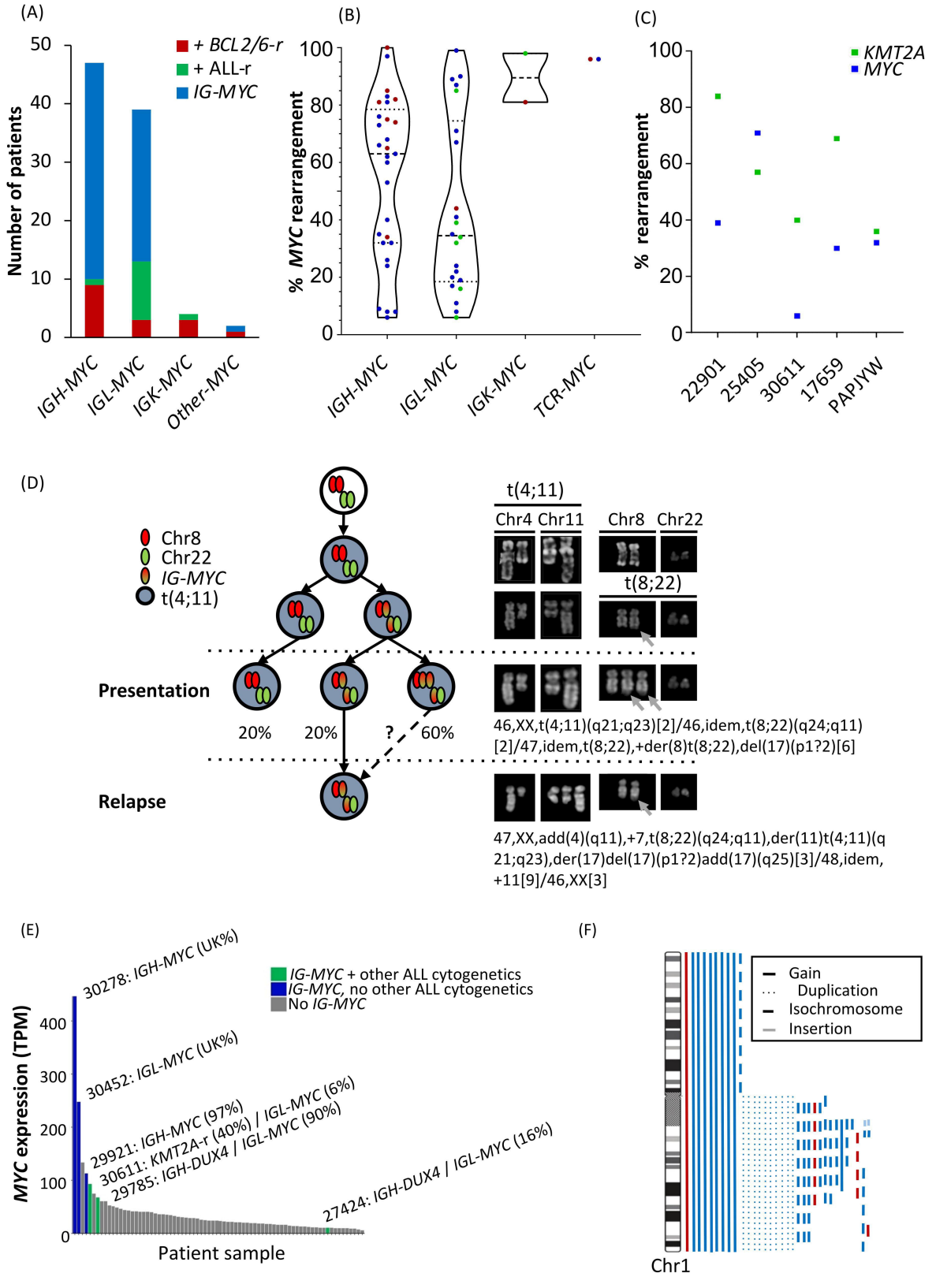
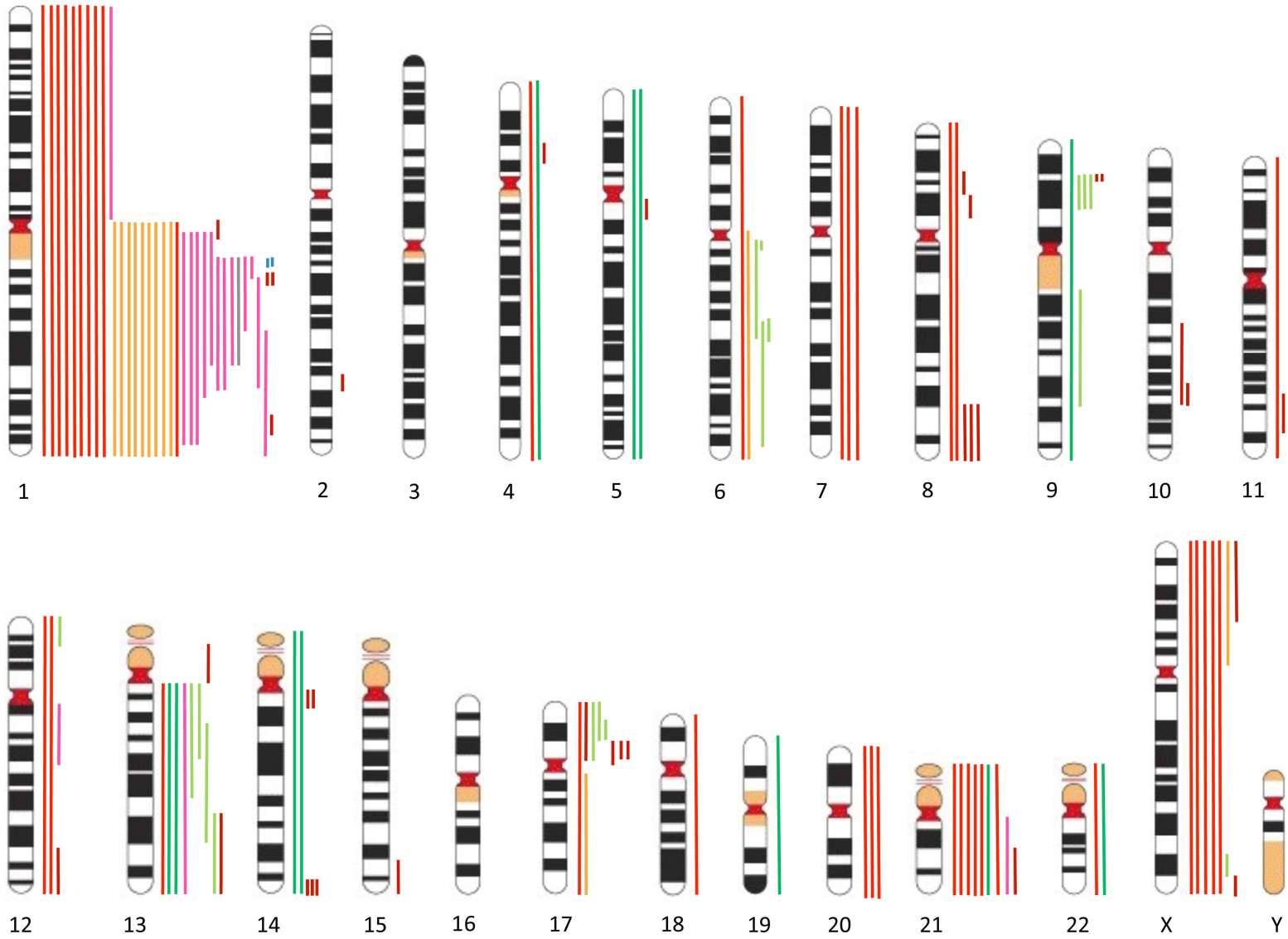

(A)
Number of patients
+ BCL2/6-r
+ ALL-r
IG-MYC
IGH-MYC
IGL-MYC
IGK-MYC
Other-MYC
(B)
% MYC rearrangement
IGH-MYC
IGL-MYC
IGK-MYC
TCR-MYC
(C)
% rearrangement
KMT2A
MYC
22901
25405
30611
17659
PAPJYW
(D)
Chr8
Chr22
IG-MYC
t(4;11)
Presentation
Relapse
20%
20%
?
60%
t(4;11)
Chr4 Chr11
Chr8 Chr22
t(8;22)
46,XX,t(4;11)(q21;q23)[2]/46,idem,t(8;22)(q24;q11)[2]/47,idem,t(8;22),+der(8)t(8;22),del(17)(p1?2)[6]
47,XX,add(4)(q11),+7,t(8;22)(q24;q11),der(11)t(4;11)(q21;q23),der(17)del(17)(p1?2)add(17)(q25)[3]/48,idem,+11[9]/46,XX[3]
(E)
MYC expression (TPM)
IG-MYC + other ALL cytogenetics
IG-MYC, no other ALL cytogenetics
No IG-MYC
30278: IGH-MYC (UK%)
30452: IGL-MYC (UK%)
29921: IGH-MYC (97%)
30611: KMT2A-r (40%) / IGL-MYC (6%)
29785: IGH-DUX4 / IGL-MYC (90%)
27424: IGH-DUX4 / IGL-MYC (16%)
Patient sample
(F)
Gain
Duplication
Isochromosome
Insertion
Chr1

Figure 2



Whole chromosome gain Additional material
Whole chromosome loss
Deletion of material
Duplication of material
Insertion of material
Isochromosome/ isodicentric Inversion of material
Triplication of material

Solid line – confident aberration reported. Dotted line – query aberration present.
Dotted and dashed line – constitutional abnormality

Figure 3

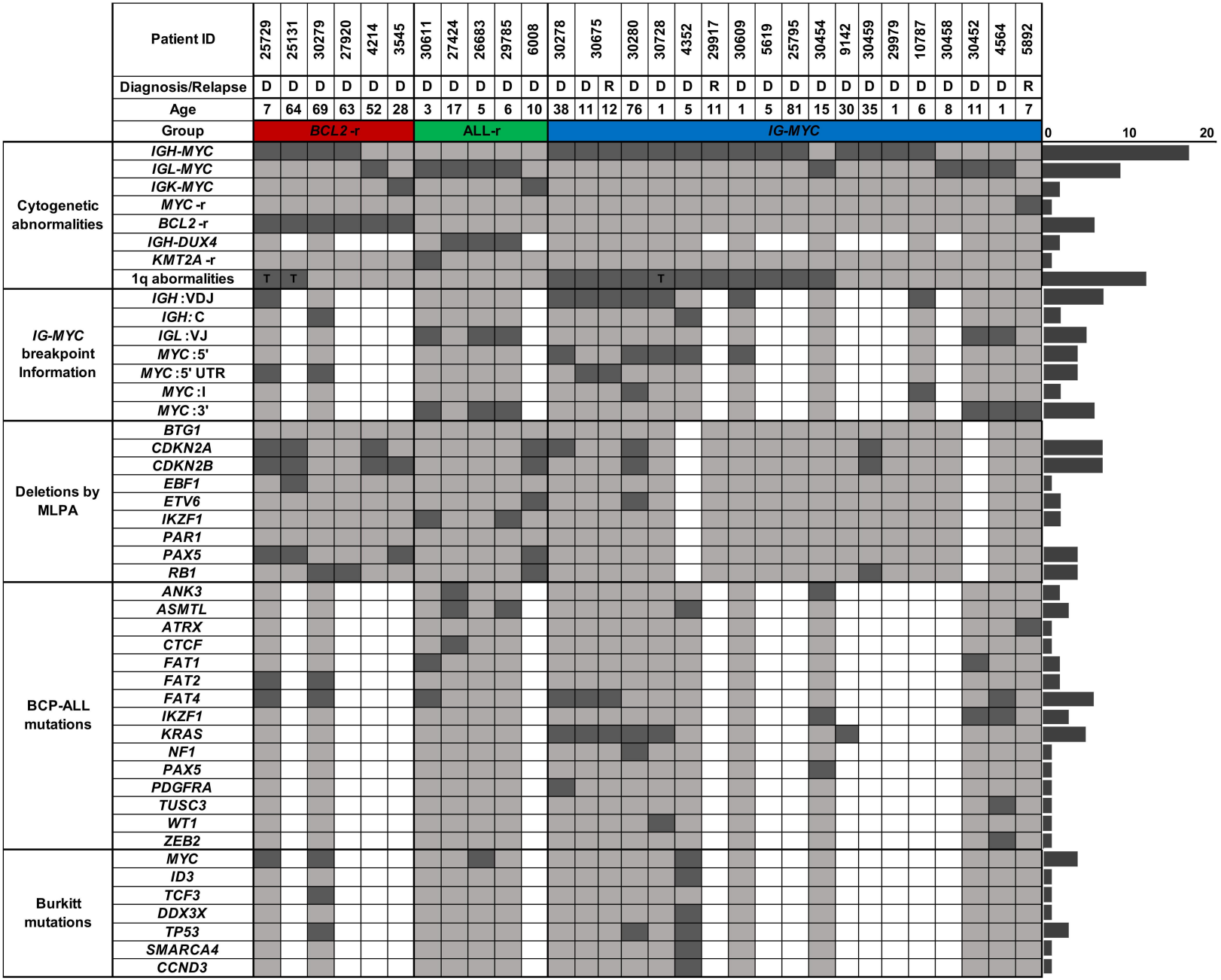

Patient ID
25729 25131 30279 27920 4214 3545 30611 27424 26683 29785 6008 30278 30675 30280 30728 4352 29917 30609 5619 25795 30454 9142 30459 29979 10787 30458 30452 4564 5892
Diagnosis/Relapse
D D D D D D D D D D D D D R D D D R D D D D D D D D D D R
Age
7 64 69 63 52 28 3 17 5 6 10 38 11 12 76 1 5 11 1 5 81 15 30 35 1 6 8 11 1 7
Group
BCL2-r
ALL-r
IG-MYC
0 10 20
Cytogenetic abnormalities
IGH-MYC
IGL-MYC
IGK-MYC
MYC-r
BCL2-r
IGH-DUX4
KMT2A-r
1q abormalities
IG-MYC breakpoint Information
IGH:VDJ
IGH:C
IGL:VJ
MYC:5'
MYC:5' UTR
MYC:I
MYC:3'
Deletions by MLPA
BTG1
CDKN2A
CDKN2B
EBF1
ETV6
IKZF1
PAR1
PAX5
RB1
BCP-ALL mutations
ANK3
ASMTL
ATRX
CTCF
FAT1
FAT2
FAT4
IKZF1
KRAS
NF1
PAX5
PDGFRA
TUSC3
WT1
ZEB2
Burkitt mutations
MYC
ID3
TCF3
DDX3X
TP53
SMARCA4
CCND3

Figure 4

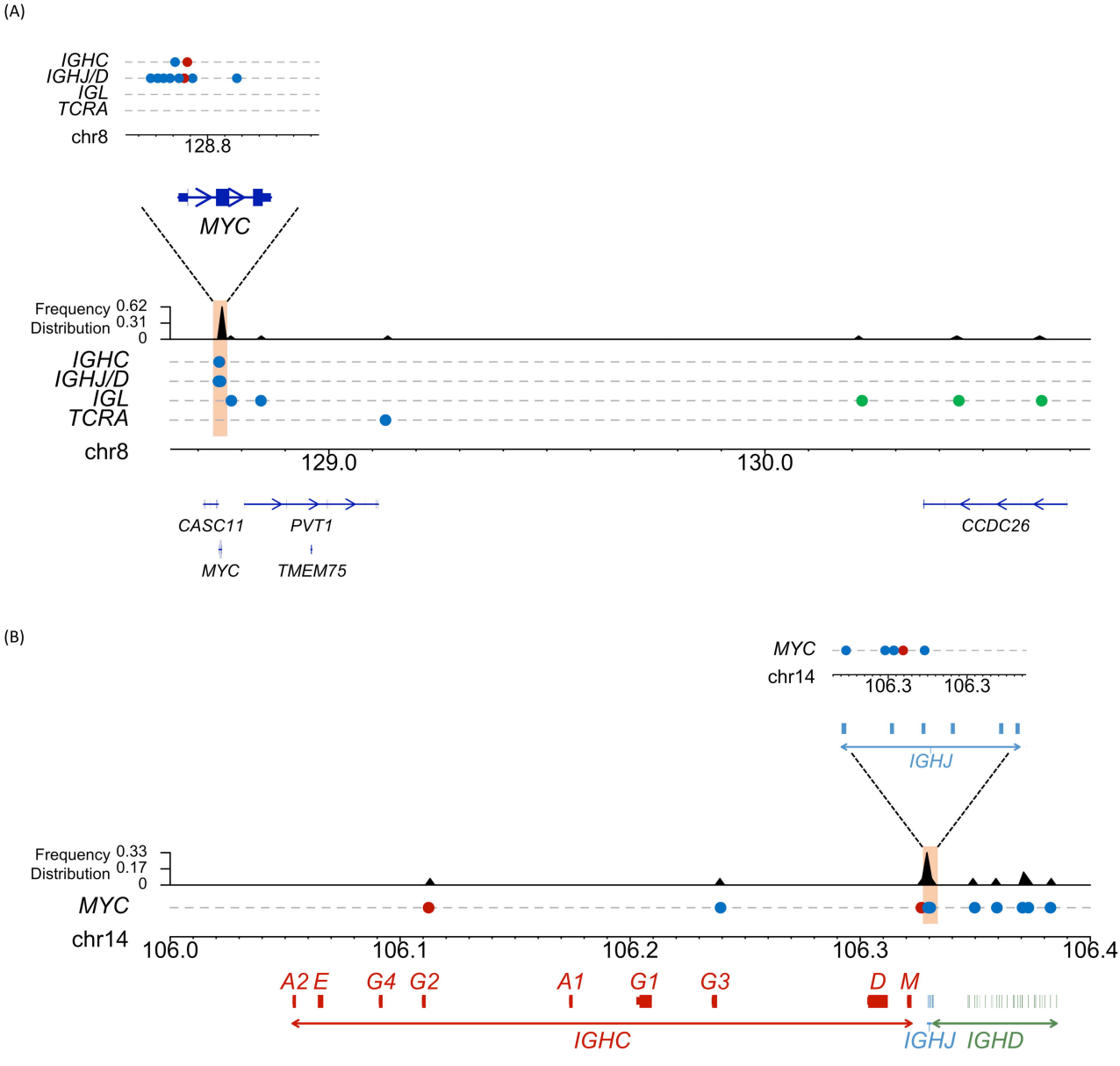
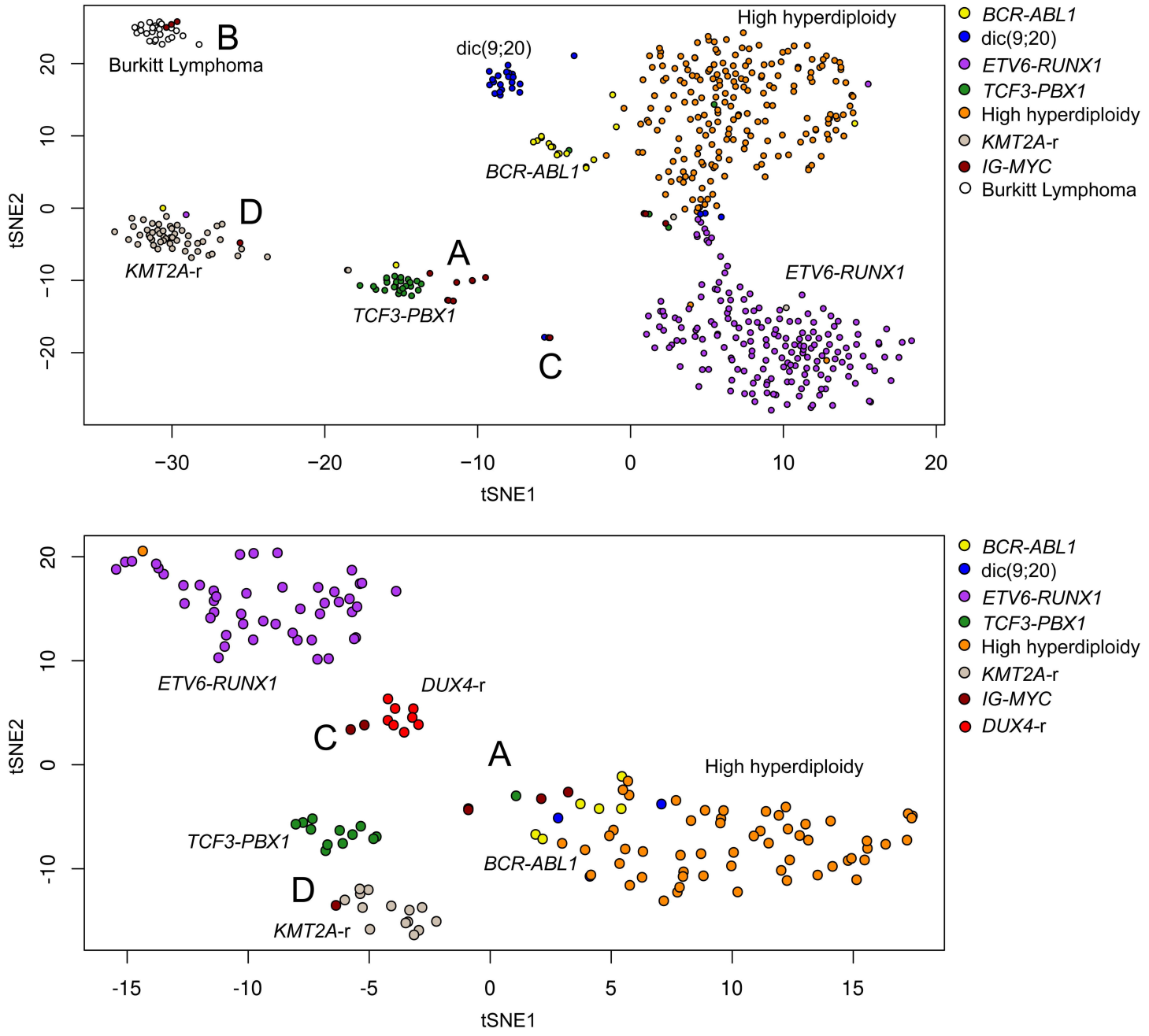

(A)
IGHC
IGHJ/D
IGL
TCRA
chr8
128.8
MYC
Frequency 0.62
Distribution 0.31
0
IGHC
IGHJ/D
IGL
TCRA
chr8
129.0
130.0
CASC11
PVT1
CCDC26
MYC
TMEM75
(B)
MYC
chr14
106.3
106.3
IGHJ
Frequency 0.33
Distribution 0.17
0
MYC
chr14
106.0
106.1
106.2
106.3
106.4
A2 E
G4 G2
A1
G1
G3
D M
IGHC
IGHJ IGHD

Figure 5



	A	B	C	D
Cluster ID	*IG-MYC*	Burkitt-like	*IGH-DUX4*	*KMT2A*-r
IGH	30728*	25729	29785*	30611*
	30675	30279	27424*	
	30280	4352	26683	
	30278			
	29921**			
IGL	4564			
	30452*			
OTHER	5892			

Figure 6.

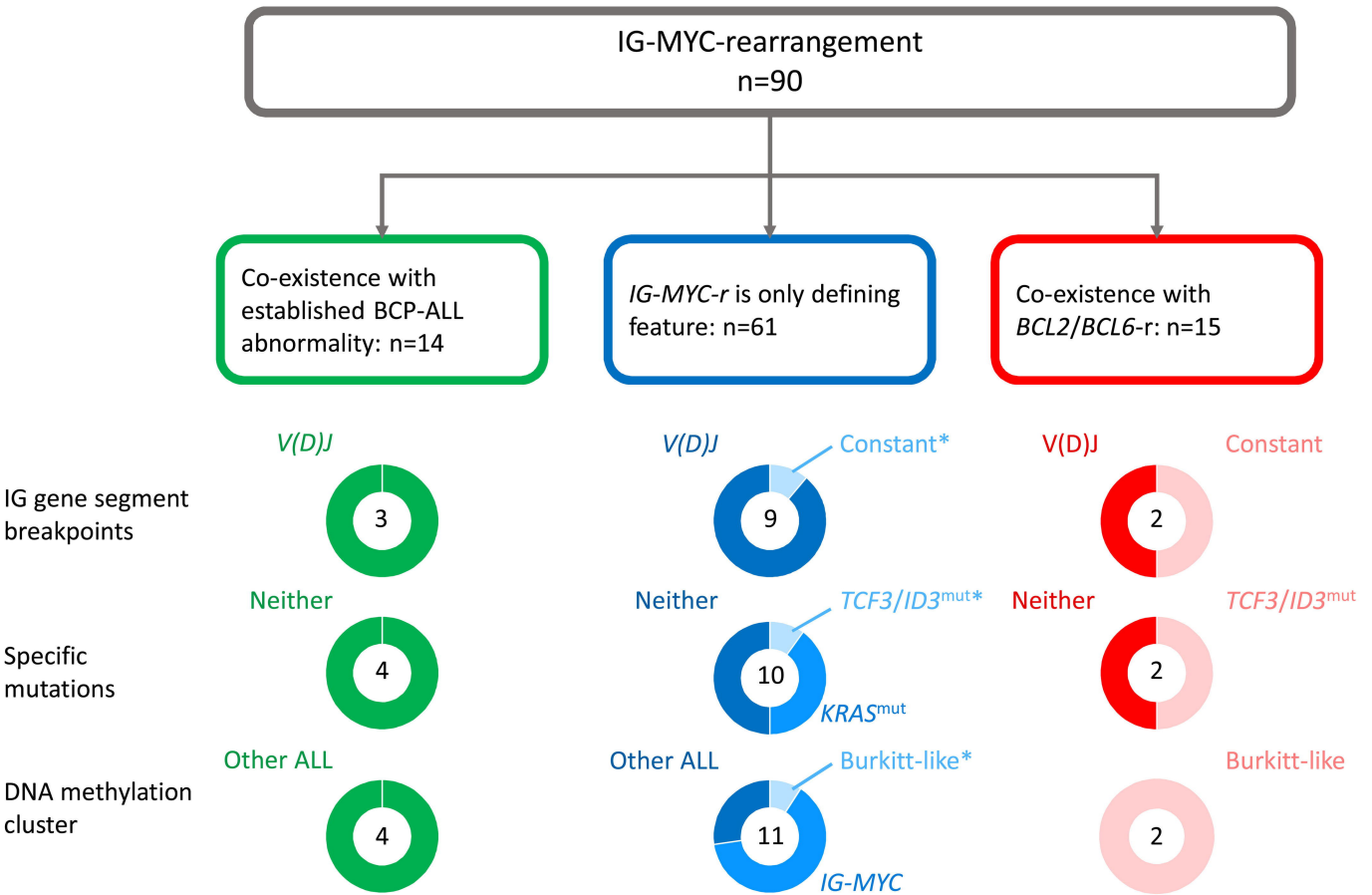
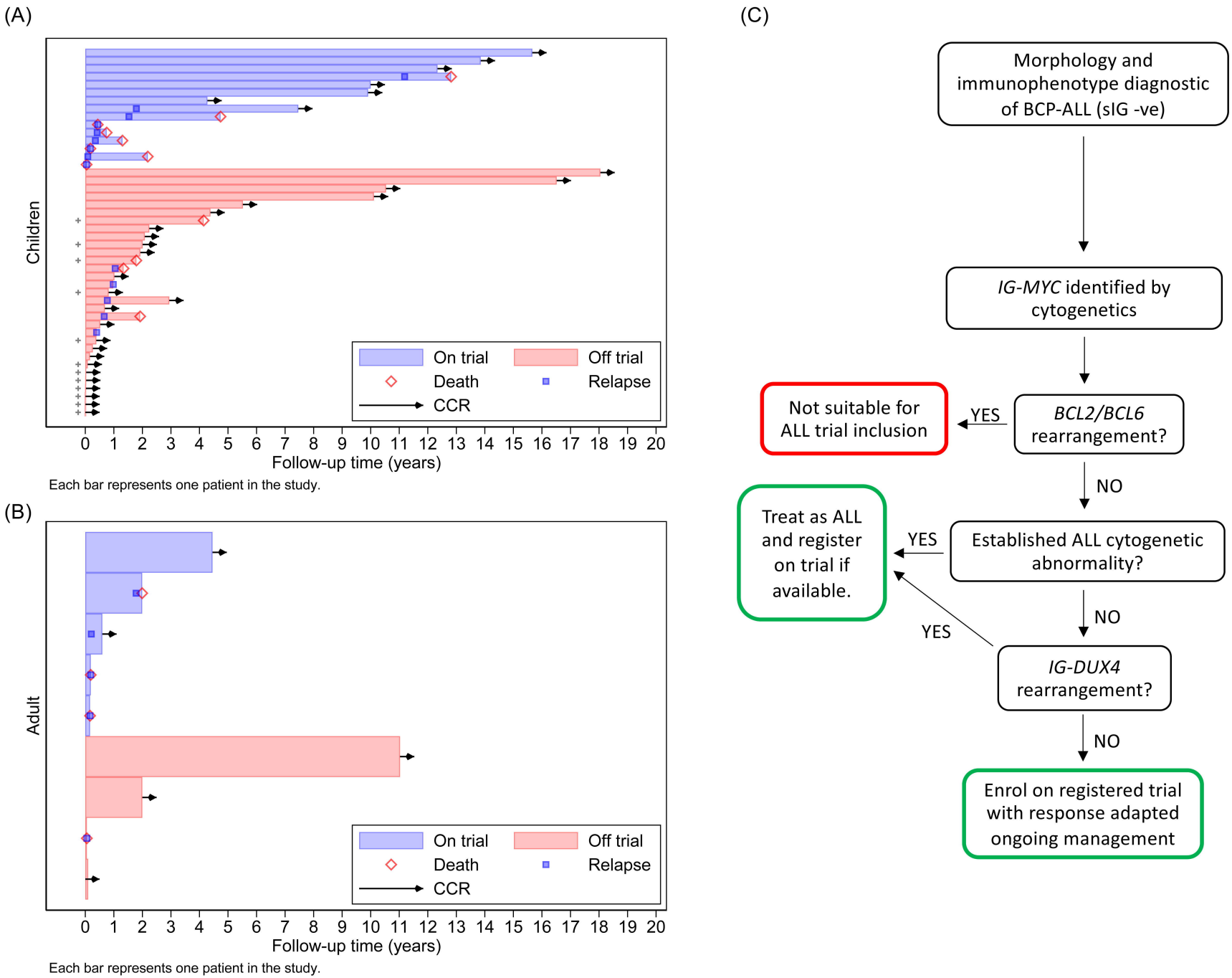

IG-MYC-rearrangement
n=90
Co-existence with established BCP-ALL abnormality: n=14
IG-MYC-r is only defining feature: n=61
Co-existence with BCL2/BCL6-r: n=15
IG gene segment breakpoints
Specific mutations
DNA methylation cluster
V(D)J
3
Neither
4
Other ALL
4
V(D)J
Constant*
9
Neither
TCF3/ID3mut*
10
KRASmut
Other ALL
Burkitt-like*
11
IG-MYC
V(D)J
Constant
2
Neither
TCF3/ID3mut
2
Burkitt-like
2

Figure 7


(A)
Children
On trial
Off trial
Death
Relapse
CCR
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Follow-up time (years)
Each bar represents one patient in the study.
(B)
Adult
On trial
Off trial
Death
Relapse
CCR
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
Follow-up time (years)
Each bar represents one patient in the study.
(C)
Morphology and immunophenotype diagnostic of BCP-ALL (sIG -ve)
IG-MYC identified by cytogenetics
BCL2/BCL6 rearrangement?
YES
Not suitable for ALL trial inclusion
NO
Established ALL cytogenetic abnormality?
YES
Treat as ALL and register on trial if available.
NO
IG-DUX4 rearrangement?
YES
NO
Enrol on registered trial with response adapted ongoing management

Molecular characterisation and clinical outcome of B-cell precursor acute lymphoblastic leukaemia with IG-MYC rearrangement

Simon Bomken[1,2], Amir Enshaei[1], Edward C Schwalbe[3], Aneta Mikulasova[4], Yunfeng Dai[5], Masood Zaka[6,7], Kent TM Fung[1], Matthew Bashton[8], Huezin Lim[1], Lisa Jones[1], Nefeli Karataraki[1], Emily Winterman[1], Cody Ashby[9], Andishe Attarbaschi[10], Yves Bertrand[11], Jutte Bradtke[12], Barbara Buldini[13], GA Amos Burke[14], Giovanni Cazzaniga[15,16], Gudrun Göhring[17], Hesta A de Groot-Kruseman[18,19], Claudia Haferlach[20], Luca Lo Nigro[21], Mayur Parihar[22], Adriana Plesa[23], Emma Seaford[24], Edwin Sonneveld[19], Sabine Strehl[25], Vincent HJ van der Velden[26], Vikki Rand[6,7], Stephen P Hunger[27], Christine J Harrison[1], Chris M Bacon[1,2], Frederik W van Delft[1,2], Mignon L Loh[28], John Moppett[24], Josef Vormoor[1,19], Brian A Walker[29], Anthony V Moorman[1], Lisa J Russell[1]

[1]Wolfson Childhood Cancer Centre, Translational and Clinical Research Institute, Newcastle University, Newcastle upon Tyne, United Kingdom
[2]The Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK
[3] Department of Applied Sciences, Northumbria University, Newcastle upon Tyne, UK
[4]Biosciences Institute, Newcastle University, Newcastle upon Tyne, UK
[5]Department of Biostatistics, Colleges of Medicine, Public Health and Health Professions, University of Florida, Gainesville, Florida, USA
[6]School of Health and Life Sciences, Teesside University, Middlesbrough, UK
[7]National Horizons Centre, Teesside University, Darlington, UK
[8]The Hub for Biotechnology in the Built Environment, Faculty of Health and Life Sciences, Northumbria University, Newcastle upon Tyne, UK
[9]Department of Biomedical Informatics / Cancer Institute, University of Arkansas for Medical Sciences, Little Rock, Arkansas, USA
[10]St Anna Children's Hospital, Medical University of Vienna, Vienna, Austria
[11]Department of Institute of Hematology Oncology Pediatric (IHOP), Hospices Civils de Lyon, Lyon, France
[12]Institute of Pathology, Department Cytogenetics, University Hospital Giessen and Marburg, Giessen
[13]University of Padova, Padova, Veneto, Italy

[14]Department of Paediatric Haematology, Oncology, and Palliative Care, Cambridge University Hospitals NHS Foundation Trust, Addenbrooke's Hospital, Cambridge, UK
[15]School of Medicine and Surgery, University of Milano-Bicocca, Milan, Italy
[16]Centro Ricerca Tettamanti, University of Milano-Bicocca, Milan, Italy
[17]Department of Human Genetics, Hannover Medical School, Hannover, Germany
[18]Dutch Childhood Oncology Group (DCOG), Utrecht, The Netherlands
[19]Princess Máxima Center for Pediatric Oncology, Utrecht, The Netherlands
[20]MLL Munich Leukemia Laboratory, Munich, Germany
[21]Head of Cytogenetic-Cytofluorimetric-Molecular Biology Laboratory, Center of Pediatric Hematology Oncology, Azienda Policlinico "G. Rodolico - San Marco," Catania, Italy
[22]Department of Cytogenetics and Laboratory Haematology, Tata Medical Centre, Kolkata, India
[23]Hematology and Flow cytometry Laboratory, Lyon Sud University Hospital, Hospices Civils de Lyon, Lyon, France
[24]Department of Paediatric Oncology, Bristol Royal Hospital for Children, Bristol, UK
[25]St. Anna Children's Cancer Research Institute, Vienna, Austria
[26]Department of Immunology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands
[27]Department of Pediatrics and the Center for Childhood Cancer Research, Children's Hospital of Philadelphia and the Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA
[28]Department of Pediatrics, Benioff Children's Hospital and the Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, California, USA
[29]Melvin and Bren Simon Comprehensive Cancer Center, Division of Hematology Oncology, Indiana University, Indianapolis, Indiana, USA

Contents

Supplementary Methods

Conventional cytogenetics and fluorescence *in situ* hybridisation

Briefly, *IGH (*CytoCell (Oxford Gene Technologies, Begbroke, Oxfordshire, UK) IGH breakapart, Vysis (Abbott Laboratories. Abbott Park, Illinois, U.S.A) LSI IGH Dual Color, Break Apart Rearrangement Probe)*, MYC* (CytoCell MYC breakapart, V*ysis MYC* Dual Color, Break Apart Rearrangement Probe) *, BCL2* (CytoCell BCL2 breakapart) and *BCL6* (CytoCell BCL6 breakapart) fluorescence *in situ* hybridisation (FISH) probes were mixed 1:1 with hybridisation buffer and denatured at 75°C for five minutes followed by hybridisation at 37°C overnight. Slides were washed for two-minutes in 0.02% SSC with 0.003% NP40 at 72°C followed by two minute room temperature incubation in 0.1% SSC. Slides were mounted with 10µl DAPI (Vector laboratories, California, USA). Manual visualisation and scoring were performed using an Olympus BX-61 fluorescence microscope with a x100 oil objective (Leica Microsystems, Gateshead, UK). Where possible, more than 100 interphase nuclei were scored for each FISH test by two independent analysts. A cut-off threshold of >5% was used for all probes to allow for interference and obscuring of signals (false positives). This cut-off level was established by counting the number of abnormal signals generated when the FISH probes were hybridised to normal cells.

Whole exome sequencing

Data analysis

Quality reports were generated using Java base tool FastQC (Babraham Bioinformatics, Cambridge, UK) and read pairs were mapped to human genome reference (hg19) (ftp://ftp.broadinstitute.org/bundle/hg19/)), using default settings from BWA-MEM (Burrows-Wheel Aligner, 0.7.17-r1188 version). Broad Institute best practices PICARD (2.19.0 version, https://broadinstitute.github.io/picard/) tools were applied for the pre-processing (merging the read groups, coordinate sorting, marking duplicates and correction of technical biases) of aligned bam files.

SNVs and Indels in tumour samples were identified using an in-house NGS data analysis pipeline. A separate analysis was run for the generation of a panel of normal (PON) from high quality 61 constitutional/normal exome samples generated using the same exome preparation methods, software and sequencing technologies to capture recurrent technical artefacts in order to improve the results of the somatic variant calling. Initial genomics variants for SNVs were identified using tumour only algorithm of MuTect2 caller of GATK (Genome Analysis Tool Kit, version 3.8), followed by extensive review of corresponding genomic positions in the PON database built in the previous step. Variant call sets were filtered after calculating the new quality scores using the VQSR (variant quality score recalibration) method. All putative somatic variants were annotated using Ensembl Variant Effect Predictor (VEP, version 90). Candidate variants

coding for synonymous consequences and allelic frequency > 1% in gnomAD (version 2.1) were discarded from further analysis.

Due to the absence of paired constitutional DNA samples, we took a targeted approach to identify common BCP-ALL and BL mutations. Genes included in the BCP-ALL panel were *ANK3, ARID1A, ASMTL, ASXL3, ATM, ATRX, BIRC6, CBL, CBLB, CHD4, CREBBP, CSF3R, CTCF, EBF1, EP300, ERG, ETV6, FAT1, FAT2, FAT4, FBXW7, FLT3, IKZF1, IKZF2, IKZF3, IL7R, JAK1, JAK2, JAK3, KAT6B, KDM6A, KIT, KRAS, KMT2A, KMT2B, KMT2C, NCL, NF1, NOTCH2, NRAS, PAX5, PDGFRA, PTPN11, RUNX1, RUNX2, SEMA7A, SF3B1, TET1, TET2, TP53, TUSC3, WHSC1, WT1* and *ZEB2*. Genes included in the BL panel were *FOXO1, MYC, ID3, TP53, DDX3X, TCF3, SMARCA4, GNA13, CCND3, ARID1A, RET, RHOA, PIK3R1, KMT2C, KMT2D and PTEN*.

Illumina Infinium MethylationEPIC BeadChip array

Briefly, 250ng dsDNA from 17 patients was hybridised to the Illumina Infinium MethylationEPIC BeadChip Array (Illumina, San Diego, CA, USA) by Eurofins Genomics (Ebersberg, Germany). DNA was treated with sodium bisulfite using the EZ-96 DNA methylation kit (Zymo Research, Irvine, CA, USA). The bisulfite-treated DNA was analysed using the manual Illumina Infinium HD.

Methylation assay according to manufacturer's guidelines. The BeadChip was scanned with both red (Cy5) and green (Cy3) laser on an iScan instrument with the iScan Control software (Illumina) determining intensity values for each bead type. The data was analysed using Illumina's Genomestudio software. Data were quality controlled and normalised as previously reported[1]. We augmented this data with additional publicly available methylation array data (Supplementary Table 3).

RNA-sequencing

RNA integrity was assessed using the Bioanalyser with RIN scores >6 being included. 500ng of RNA was used to prepare RNA strand-specific libraries using TruSeq Stranded mRNA Library Prep Kit (Illumina) according to the manufacturer's instructions. For sequencing, pooled libraries were loaded on the cBot (Illumina) and cluster generation was performed using manufacturer's instructions. Paired-end sequencing using 100bp read length was performed on a HiSeq2500 machine (HiSeq Control Software 2.2.58) using HiSeq Flow Cell v4 and TruSeq SBS Kit v4. For processing of raw data RTA version 1.18.64 and CASAVA 1.8.4 was used to generate FASTQ-files. Data were quantified by kallisto[2]. The estimated counts were transformed using voom from the limma R package[3] and batch effect was corrected by permuted SVA[4].

Supplementary Table Legends

Supplementary Table 1. Supplementary Table 1. Patient cohort demographic, clinical and cytogenetic characteristics

Supplementary Table 2. Immunophenotype data collected from patients with IG-MYC-r

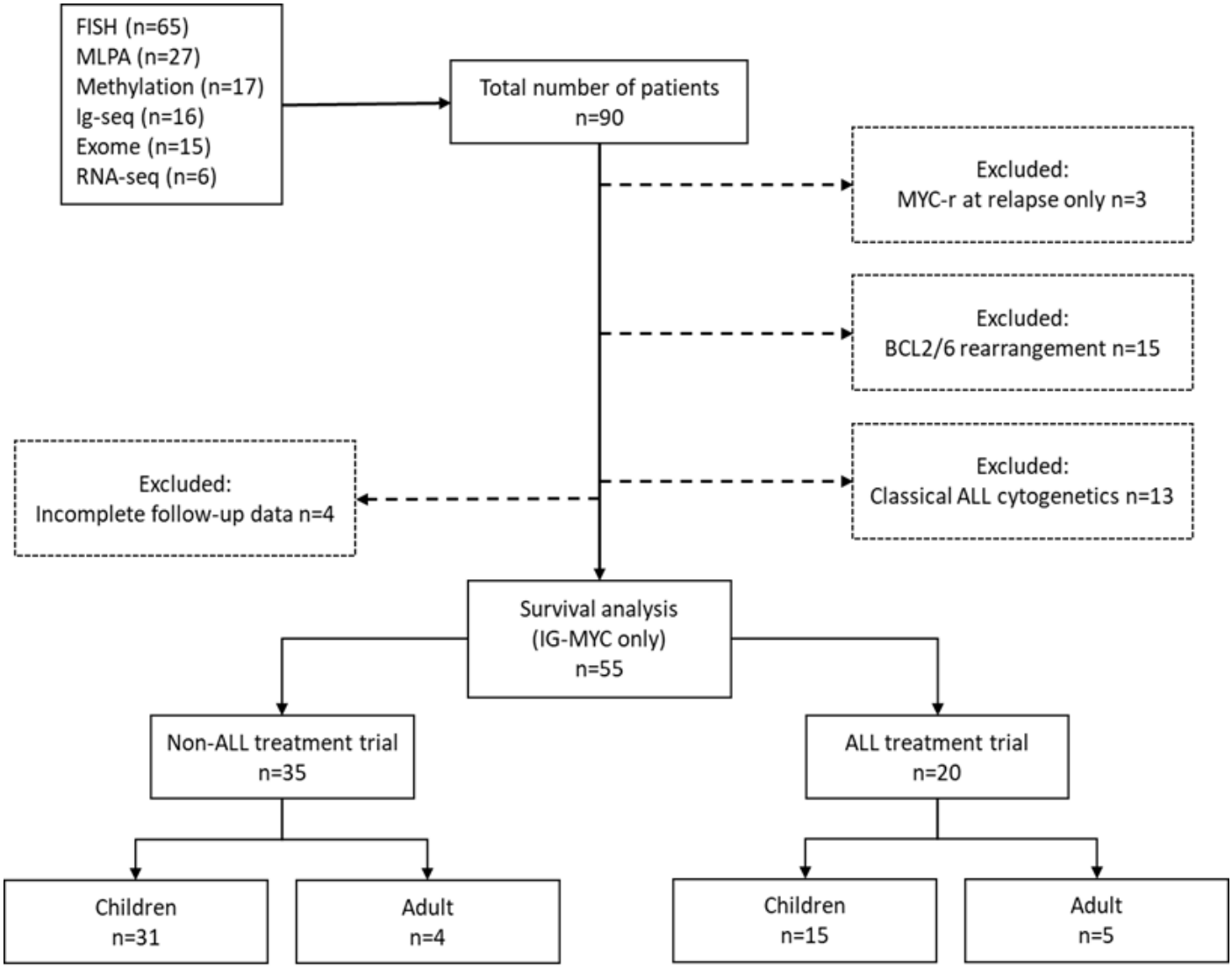
Supplementary Table 3. Publicly available methylation array datasets used in our analysis

Supplementary Table 4. Detail of translocations involving immunoglobulin and/or MYC genes detected by targeted sequencing

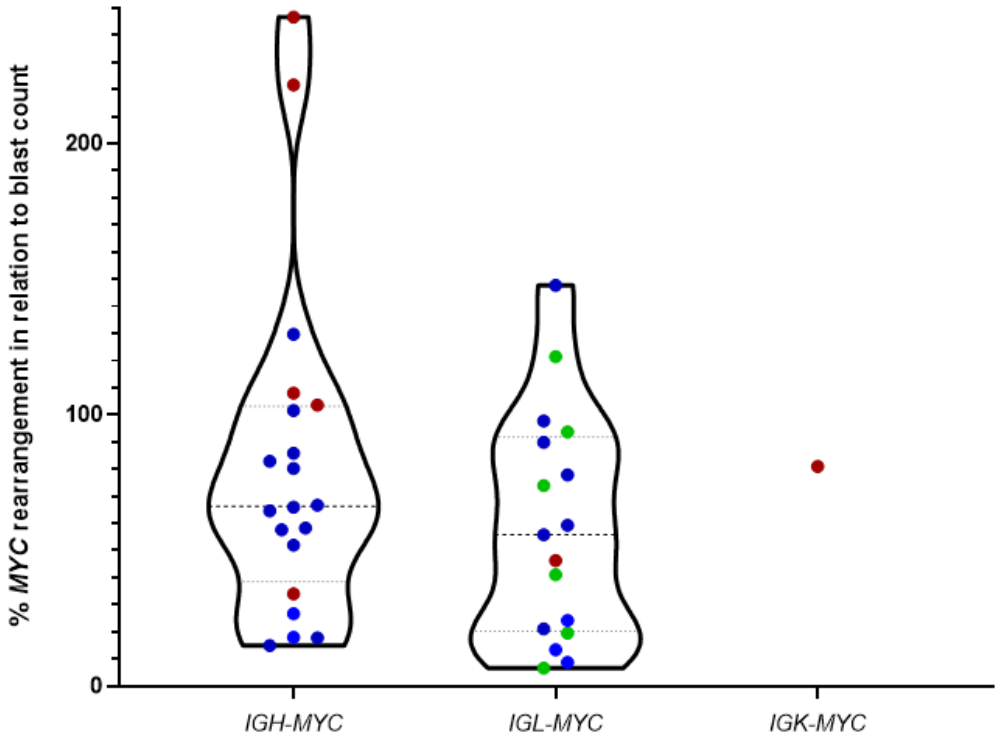
Supplementary Table 5. Univariate analysis of the key molecular features identified in children and adults with IG-MYC (excluding those with BCL2/6-r and established ALL-specific rearrangements).

Supplemental tables provided as excel files

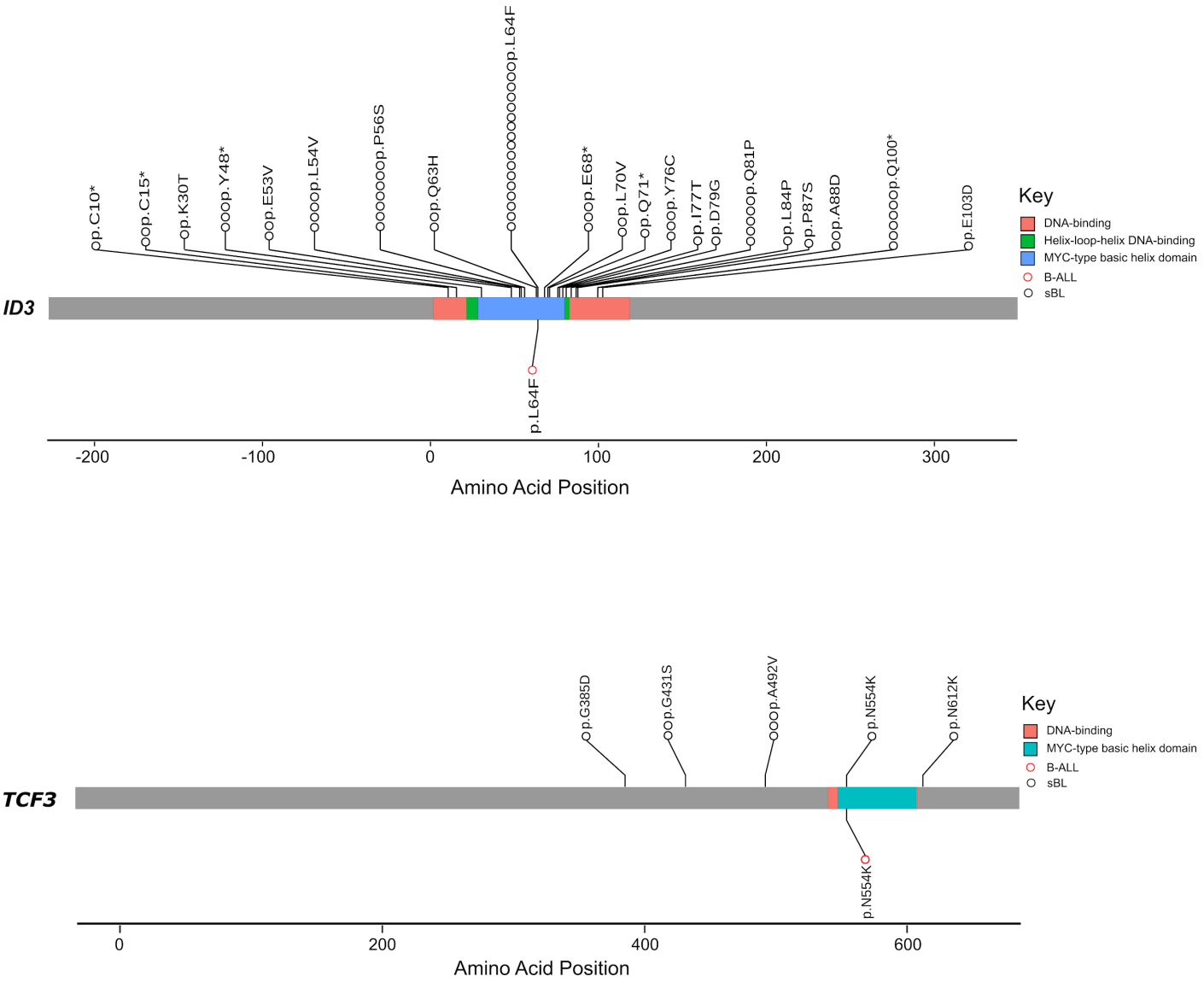
Supplementary Figures



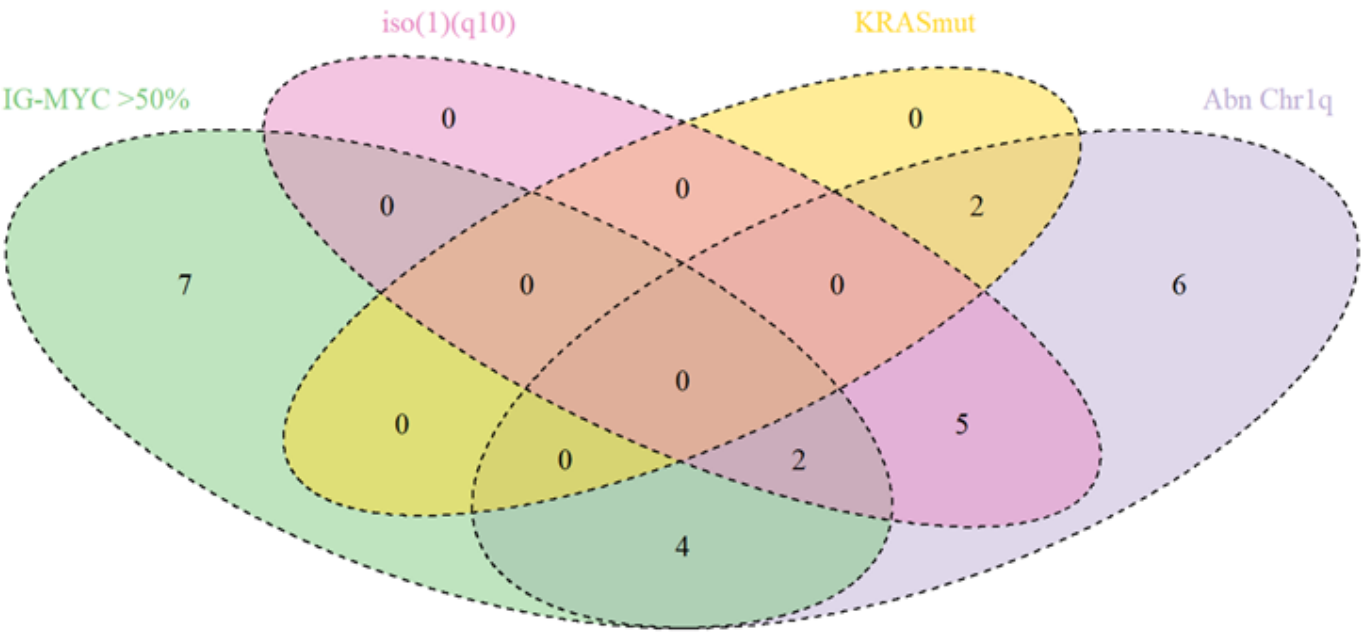
Supplementary Figure 1. Consort diagram for the genetic characterisation and survival analysis of patients with BCP-ALL and *IG-MYC*-r.



Supplementary Figure 2. Percentage of blasts carrying *IG-MYC*-r grouped by immunoglobulin chain involvement. Red dots - patients with concomitant BCL2/BCL6-r. Green dots - patients with concomitant ALL-specific abnormality. Blue dots - patient with *IG-MYC*-r as the defining cytogenetic abnormality. Dotted and dashed lines, interquartile ranges.



Supplementary Figure 3. Lolliplots for common BL mutations. ID3 (upper panel) and TCF3 (lower panel) in Burkitt lymphoma samples (black circles, Newman *et al*[5]) and BCP-ALL samples from this study (red circles).



Supplementary Figure 4. Venn diagram of the key molecular features identified within children with *IG-MYC*-r BCP-ALL.

Supplementary References

1. Schwalbe EC, Lindsey JC, Nakjang S, et al. Novel molecular subgroups for clinical classification and outcome prediction in childhood medulloblastoma: a cohort study. Lancet Oncol. 2017;18(7):958-971.

2. Bray, N.L., Pimentel, H., Melsted, P., and Pachter, L. Near-optimal probabilistic RNA-seq quantification. Nature Biotechnology. 2016;34(5):525-527.

3. Law, C.W., Chen, Y., Shi, W., and Smyth, G.K. Voom: precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biology. 2014;15:R29.

4. Parker, H.S., Leek, J.T., Favorov, A.V., et al. Preserving biological heterogeneity with a permuted surrogate variable analysis for genomics batch correction. Bioinformatics. 2014;30(19):2757-2763.

5. Newman AM, Zaka M, Zhou P, et al. Genomic abnormalities of *TP53* define distinct risk groups of paediatric B-cell non-Hodgkin lymphoma. Leukemia. 2022;36(3):781-789.