blood advances

American Society of Hematology
2021 L Street NW, Suite 900,
Washington, DC 20036
Phone: 202-776-0544 | Fax 202-776-0545
bloodadvances@hematology.org

Diagnostic utility of whole genome sequencing in adults with B-other acute lymphoblastic leukemia

Tracking no: ADV-2022-008992R1

Daniel Leongamornlert (Wellcome Sanger Institute, United Kingdom) Jesús Gutiérrez-Abril (MSKCC, United States) Soo Wah Lee (UCL Cancer Institute, United Kingdom) Emilio Barretta (Ilumina Centre, United Kingdom) Tom Creasey (Newcastle University, United Kingdom) Gunes Gundem (Memorial Sloan Kettering Cancer Center, United States) Max Levine (Memorial Sloan Kettering, United States) Juan Arango Ossa (Memorial Sloan Kettering Cancer Center, United States) Konstantinos Liosis (Memorial Sloan Kettering Cancer Center, United States) Juan Medina-Martinez (Memorial Sloan Kettering Cancer Center, United States) Krisztina Zuborne Alapi (Barts Health NHS Trust, United Kingdom) Amy Kirkwood (UCL, United Kingdom) Laura Clifton-Hadley (University College London, United Kingdom) Pip Patrick (University College London, United Kingdom) David Jones (Wellcome Sanger Institute, United Kingdom) Laura O'Neill (Wellcome Trust Sanger Institute, United Kingdom) Adam Butler (The Wellcome Trust Sanger Institute, United Kingdom) Christine Harrison (Newcastle University, United Kingdom) Peter Campbell (Wellcome Trust Sanger Institute, United Kingdom) Bela Patel (Queen Mary University of London,, United Kingdom) Anthony Moorman (Newcastle University, United Kingdom) Adele Fielding (UCL Cancer Institute, United Kingdom) Elli Papaemmanuil (MSKCC, United Kingdom)

Abstract:
Genomic profiling at diagnosis of B-cell precursor Acute Lymphoblastic Leukemia (BCP-ALL) in adults is used to guide disease classification, risk stratification and treatment decisions. Patients for which diagnostic screening fails to identify disease defining or risk stratifying lesions are classified as B-other ALL. We screened a cohort of 652 BCP-ALL cases enrolled in UKALL14 to identify and perform whole genome sequencing (WGS) on paired tumor-normal samples. For 52 B-other patients we compared WGS findings to data from clinical and research cytogenetics. WGS identifies a cancer associated event in 51/52 cases, this includes an established subtype defining genetic alteration in 5/52 that were previously missed by standard-of-care genetics. Of the 47 true B-other ALL we identified a recurrent driver in 87% (41). Complex karyotype by cytogenetics emerges as a heterogeneous group, including distinct genetic alterations associated with either favorable (DUX4-r) or poor outcomes (MEF2D-r, IGK::BCL2). For a subset of 31 cases, we integrate findings from RNA-sequencing (RNA-seq) analysis to include fusion gene detection, and classification by gene expression. Compared to RNA-seq, WGS was sufficient to detect and resolve recurrent genetic subtypes, however RNA-seq can provide orthogonal validation of findings. In conclusion, we demonstrate that WGS can identify clinically relevant genetic abnormalities missed by standard-of-care testing and identify leukemia driver events in virtually all cases of B-other ALL.

Conflict of interest: COI declared - see note

COI notes: G.G. is a consultant at Isabl Inc. M.F.L. is an employee, equity holder in Isabl Inc. J.S.M.M. is a founder, equity holder in Isabl Inc. A.A.K has received honoraria from Kite. L.C.H on behalf of CRUK and UCL CTC has received research funding from Astra Zeneca, GSK, Pfizer, MSD, BMS, Amgen, Millennium Takeda. P.C. is a founder in FL86 inc. B.P. is on the advisory board for Pfizer. A.V.M has received honoraria from Amgen. A.K.F. has served as a consultant for Amgen, Pfizer and Novartis. E.P. is a founder, equity holder and holds fiduciary role in Isabl Inc. and is an equity holder in TenSixteen Bio.

Preprint server: No;

Author contributions and disclosures: Conception and Design: P.C., B.P., A.V.M., A.K.F. and E.P. Financial support: C.J.H., P.C., B.P., A.V.M., A.K.F. and E.P. Provision of study patients: S.L., K.Z.A., A.A.K., L.C.H., P.P. and A.K.F. Collection and assembly of data: D.L., S.L., E.B., T.C., K.Z.A., A.A.K., L.C.H., P.P., A.V.M. and A.K.F. Laboratory processing of biospecimens and sequencing: S.L. and L.O. Processed, analyzed and interpreted data: D.L., J.G.A., G.G., M.F.L., J.E.A.O., K.L., J.S.M.M., D.J., A.P.B., P.C., A.V.M. and E.P. Wrote the manuscript: D.L., A.V.M., A.K.F. and E.P. Final approval of manuscript: All authors.

Non-author contributions and disclosures: No;

Agreement to Share Publication-Related Data and Data Sharing Statement: All sequencing data has been deposited in EGA accession numbers EGAS00001002474 & EGAS00001003428.

Clinical trial registration information (if any):

Diagnostic utility of whole genome sequencing in adults with B-other acute lymphoblastic leukemia

Daniel Leongamornlert(1), Jesús Gutiérrez-Abril(2), SooWah Lee(3), Emilio Barretta(4), Thomas Creasey(4), Gunes Gundem (2), Max F. Levine(2), Juan E. Arango-Ossa(2), Konstantinos Liosis(2), Juan S Medina-Martinez(2), Krisztina Zuborne-Alapi (3), Amy A. Kirkwood(5), Laura Clifton-Hadley(5) Pip Patrick (5), David Jones(1), Laura O'Neil(1), Adam P. Butler(1), Christine J. Harrison(4), Peter Campbell(1), Bela Patel (6), Anthony V Moorman*(4), Adele K. Fielding*(3), Elli Papaemmanuil*(2)

(1) Cancer, Ageing and Somatic Mutation (CASM), Wellcome Sanger Institute, Hinxton, UK

(2) Department of Epidemiology & Biostatistics, Memorial Sloan Kettering Cancer Center, New York, New York, USA

(3) Department of Haematology, UCL Cancer Institute, London, UK

(4) Leukaemia Research Cytogenetics Group, Translational and Clinical Research Institute, Newcastle University, Newcastle upon Tyne, UK

(5) Cancer Research UK & UCL Cancer Trials Centre, UCL Cancer Institute, University College London, UK

(6) Department of Haemato-Oncology, Barts Cancer Institute, London, Queen Mary University of London UK

Statement of equal author contribution

A.V.M., AKF and E.P. contributed equally to this study.

Keywords: B-cell precursor acute lymphoblastic leukemia (BCP-ALL), B-other, whole genome sequencing, genetics, genomics,

Presented in abstract form at the 64th American Society of Hematology Annual Meeting in December 2022.

All sequencing data has been deposited in EGA accession numbers EGAS00001002474 & EGAS00001003428. Contact the corresponding author for other forms of data sharing: papaemme@mskcc.org.

Abstract

Genomic profiling at diagnosis of B-cell precursor Acute Lymphoblastic Leukemia (BCP-ALL) in adults is used to guide disease classification, risk stratification and treatment decisions. Patients for which diagnostic screening fails to identify disease defining or risk stratifying lesions are classified as B-other ALL. We screened a cohort of 652 BCP-ALL cases enrolled in UKALL14 to identify and perform whole genome sequencing (WGS) on paired tumor-normal samples. For 52 B-other patients we compared WGS findings to data from clinical and research cytogenetics. WGS identifies a cancer associated event in 51/52 cases, this includes an established subtype defining genetic alteration in 5/52 that were previously missed by standard-of-care genetics. Of the 47 true B-other ALL we identified a recurrent driver in 87% (41). Complex karyotype by cytogenetics emerges as a heterogeneous group, including distinct genetic alterations associated with either favorable (*DUX4*-r) or poor outcomes (*MEF2D*-r, *IGK::BCL2*). For a subset of 31 cases, we integrate findings from RNA-sequencing (RNA-seq) analysis to include fusion gene detection, and classification by gene expression. Compared to RNA-seq, WGS was sufficient to detect and resolve recurrent genetic subtypes, however RNA-seq can provide orthogonal validation of findings. In conclusion, we demonstrate that WGS can identify clinically relevant genetic

abnormalities missed by standard-of-care testing and identify leukemia driver events in virtually all cases of B-other ALL.

Key Points

1. Whole genome sequencing stratifies 88% of B-other ALL to an established genetic subtype that was not possible to detect by cytogenetics
2. Complex karyotype B-ALL emerges as a heterogeneous group of genetic subtypes including *MEF2D*-r, *DUX4*-r and *IGK::BCL2*

Subjects: Article, Lymphoid Neoplasia, Adult B-other ALL

Topics: Whole genome sequencing

Introduction

At diagnosis, genetic classification of adult B-cell precursor Acute Lymphoblastic Leukemia (BCP-ALL) is informed by the detection of structural chromosome alterations and altered ploidy states, typically assessed by cytogenetic analysis, fluorescence in situ hybridization (FISH) or reverse transcriptase PCR (RT-PCR). These diagnostic strategies inform treatment decisions through the identification of targetable lesions such as *BCR::ABL1*[1], or by assigning patients with high-risk genetic abnormalities such as *KMT2A* fusion, low hypodiploidy or complex karyotype to more intensive therapies such as allogeneic stem cell transplant[2,3].

UKALL14 is a UK National Cancer Research Institute Adult ALL group study in which patients were stratified according to the Moorman cytogenetic risk classification[2]. The outcome for the primary, clinical randomized question has recently been published[4]. The study also aimed to better understand the relationship between the B-ALL cancer genome, clinical phenotype, and therapeutic response. Genetic profiling of 652 BCP-ALL UKALL14 participants by cytogenetics, FISH and MLPA confirmed the high-risk nature of *KMT2A*-r, low hypodiploid and complex karyotype and identified JAK-STAT abnormalities as a new high-risk genetic subgroup[4]. In unison, these profiling approaches assigned 70% cases to an established genetic subgroup leaving 30% of cases in the undefined and clinically heterogeneous B-other subgroup.

Recent comprehensive profiling approaches by whole transcriptome sequencing have been instrumental in the discovery of disease defining alterations in ALL[5–9]. In the current study, we deploy retrospective whole genome and transcriptome sequencing to characterize the genetic landscape of B-other adult ALL. We demonstrate that comprehensive genome profiling allows for the detection of all lesions reported by standard of care (SoC) profiling and critically resolve genetic subtypes for the majority of BCP-ALL patients where SoC failed to deliver a definitive diagnosis.

Methods

Patients and sample selection

Patients were treated on a randomized trial for adults (25-65 years) with newly diagnosed ALL (UKALL14, ISRCTN66541317, NCT01085617). Trial participation and correlative research was supported by written informed consent. The study received institutional review board approval. To identify patients with B-other ALL, we performed central review of genetic data collected through standard of care testing (cytogenetics, FISH, RT-PCR) as well as research profiling including FISH (*CRLF2*, *JAK2*, *ABL1*, *ABL2* & *PDGFRB*) and MLPA (SALSA MLPA Probemix P335, MRC Holland)[10]. A total of 58 cases satisfied criteria for whole genome sequencing (WGS): (a) absence of genomic drivers (Supplemental Figure 1, Supplemental Table 1); (b) availability of a diagnostic, pre-treatment DNA; and (c) availability of germline control DNA (MRD <1% or buccal swab). Diagnostic RNA was available for 33 of the 58 cases (Supplemental Table 2).

Sequencing and Bioinformatic analysis

WGS was performed (2x 150bp) to a target read depth of 60x and 30x for tumor and normal samples respectively. RNA-seq was performed using oligo dT pulldown to a target coverage of 50 million reads (2x 75bp).

WGS and RNA-seq were aligned to GRCh37d5 using BWA-MEM (Supplemental Table 3) and STAR (version 2.5.0c, Supplemental Table 4). WGS variant calling proceeded with 57/58 sequence complete pairs to determine; somatic single nucleotide variants, insertion/deletions, copy number aberrations and structural variants[11–15]. Variants were annotated with OncoKb, to determine putative role in cancer pathogenesis[16].

Data from 33 RNA-seq cases was analyzed to classify ALL subtypes using the consensus of any two of the ALLSorts[17], ALLSpice[18] and ALLCatchR (https://github.com/ThomasBeder/ALLCatchR) classifiers. Sample clustering with the City of Hope (COH) B-ALL reference gene expression data set from the MD-ALL R package (https://github.com/gu-lab20/MD-ALL) was used to finalize subtype classification. RNA fusions were detected using FusionCatcher[19] and CICERO[20]. RNA mutations in PAX5 (R38H/C, P80R & R140L) and ZEB2 (H1038R), were manually checked using IGV[21]. Data integration, analysis execution and visualization were conducted using the Isabl platform[22].

To detect enhancer hijacking events associated with the *IGH* locus, we used gGnome (https://github.com/mskilab/gGnome) to construct a graph representation of the SVs detected in the tumor sample. For each graph the gGnome "proximity" function was used to "walk" from the Eμ *IGH* super enhancer locus to genes in GENCODE v29[23,24]. Candidate IGH walks <100 kbp were annotated using the Atlas of Genetics and Cytogenetics in Oncology and Haematology and Mitelman Database of Chromosome Aberrations and Gene Fusions in Cancer[25,26].

For DUX4r detection, we used the GRIDSS[27] SV caller, which allows the identification of single breakends (SVs unambiguously anchored at only one locus). For each single breakend identified, BLAT was used to map the unplaced sequence to hg19[28]. Single breakend sequences which mapped to chrUn_gl000228 (an unplaced hg19 contig with DZ4Z repeats containing *DUX4*) were manually reviewed in IGV[21]. For further details please refer to Supplementary methods.

Mutational signatures

For single base substitutions (SBS), mutation signature analysis was performed using the R/Bioconductor "MutationalPatterns" (v3.5.6) package[29] using a seven signature SBS mutational catalog as reference (Supplemental Figure 3)[30]. RAG mediated deletions were annotated by motif analysis using MEME[31,32].

Statistical analysis

R version 3.6.3 (2020-02-29) was used; Mann-Whitney U-test was used for between groups comparison of continuous variables using the "wilcox.test" function[33]. A robust linear model was used to calculate the per year SNV burden using the "rlm" function from the MASS R package[34] and 95% confidence intervals with "confint.default" function[33].

Data Sharing Statement

All sequencing data has been deposited in EGA accession numbers EGAS00001002474 & EGAS00001003428.

Results

Comprehensive profiling by WGS, characterized arm-level and focal copy number alterations (CNAs), structural variants (SV), acquired somatic and germline mutations mapping at least one aberrant somatic lesion in 52/57 B-other ALL samples in our cohort (Figure 1a). The remaining five samples were excluded from subsequent analysis owing to low mutation burden (<100, Supplemental Figure 4) (Supplemental Figure 5), suggestive of low tumor burden. Notably, all five low purity samples were erroneously considered to have normal karyotypes by cytogenetics (Supplemental Table 6). This resulted in a set of 52 cases of which 31 also had RNA-seq (Supplemental Table 7).

Arm-level CNAs were the most common class of alteration (n=139, median 1), followed by focal deletions (<10Mb) in tumor suppressor genes (n=106, median 2) and acquired somatic mutations in cancer genes (n=43, median 1). The majority of cases (81% 42/52) had focal deletions, 60% (31/52) had acquired somatic mutation targeting established cancer genes, 58% (30/52) had recurrent gene rearrangements and 57% (28/52) harbored arm-level CNAs (Figure 1a).

Genomic Classification of B-other cases by WGS

WGS analysis of the 52 B-other cases identified 5 (10%) cases with an established WHO[2016] genetic subtype that was missed by SoC testing (Supplemental Table 8). Three high-risk (two low hypodiploid and a near haploid) cases and two standard risk (high hyperdiploid and *TCF3::PBX1*) cases. Cytogenetic analysis for these 5 cases either failed (n=2) or produced a normal karyotype (n=3). The median inferred tumor purity for these cases was 46% (range 33%-96%) compared to 89% (range 13%-99%) for the remaining patients (n=47). Of note was the near-triploidy case (PD37195) with a germline TP53 (p.N239D, 17:7577566 T>C, VAF in normal tissue = 0.29, tumor VAF = 0.87 with chr17 cnLOH) mutation (Supplemental Table 8).

Among the remaining 47 B-other cases, WGS analysis identified abnormalities enabling 41 (87%) cases to be classified into the following subtypes: *DUX4* rearrangements (*DUX4*-r, n=8), PAX5alt (n=7), PAX5 P80R (n=4), *ZNF384* rearrangements (*ZNF384*-r, n=5), *ZEB2*/*CEBP* (n=4), *MEF2D* rearrangements (*MEF2D*-r, n=3), *UBTF::ATXN7L3* fusions (n=2), *IDH1/2* mutations (n=2), Ph-like JAK-STAT abnormalities (n=2), along with single cases of *IGK::BCL2*, *IGH::CEBPA*, *IGH::ID4* and *IGH::MIR125B1* (Figure 1b, Supplemental Figure 6, Supplemental Table 8). With the exception of *IGH* rearrangements and ZEB2 H1038R, these genetic lesions define non-overlapping subgroups. Deletions targeting *CDKN2A*/*B*, *PAX5*, *IKZF1* are prevalent across genetic subtypes. The same is true for mutations targeting RAS/MAPK signaling genes (*NRAS*, *KRAS*, *FLT3*, *PTPN11*, *GNB1*, *EGFR*) (Supplemental Table 8, Figure 1b).

The remaining 6 (13%) B-other cases could not be classified into a recognized B-other subtype (Supplemental Table 9). Case PD37191 harbored a *MYO18A::FGFR1* fusion (t(8;17)(p11;q23)), which most likely maps to the WHO[2016] "Myeloid/lymphoid neoplasms with FGFR1 rearrangement" classification[35]. Two cases (PD40822, PD40836) had clonal frameshifts in genes associated with T-ALL (PTPN2 p.A108fs*5 and WT1 p.V371fs*14 respectively). PD40803 had a somatic ETV6 p.R399C previously reported as germline risk[36] with a concurrent complex SV within *KDM6A* (Supplemental Table 10) and a single case harbored a clonal mutation typically associated with myeloid disease ASXL1 p.R417* (PD40826). Lastly, WGS failed to identify a putative driver in one case (PD40831), with an estimated tumor purity of 28% (Supplemental Table 9).

Comparison of WGS and RNA-seq subtype allocation

To evaluate the diagnostic utility of WGS to RNA-seq, we compared diagnostic findings in 31 cases with both WGS and RNA-seq data (Supplemental Table 7). WGS analysis identified a disease defining genetic alteration in all 31(Figure 1c, Supplemental Table 8). Of the 31 cases with RNA-seq data, five were classified as "low" RNA sample quality: three cases (PD37187, PD43260 & PD43262) with evidence of

cross contamination (Supplemental Table 7) and two cases (PD37188 & PD40803) identified as low purity which impairs gene expression-based classification (Supplemental Figure 7a, Supplemental Table 11). Estimation of % blast count by RNA and WGS correlated well, with the expected exception of the high hyperdiploid case (Supplemental Figure 7b).

In 19/31 cases the underlying lesion contained a genomic rearrangement. Evaluation of RNA-seq classification by gene expression and fusion detection (Supplemental Table 11 & 12), correctly classified 19 cases (16 by both, 1 fusion only and 2 classifier only, Supplemental Table 13). Fusion detection missed two *IGH::DUX4* rearrangements. In a single case with *P2RY8::CRLF2* (PD40808); RNA classification was split between Ph-like (ALLCatchR & ALLSorts) and iAMP21 (COH reference data set). In this case WGS analysis only identified a single chr21 gain rather than a high-level amplification[37] which was consistent with the diagnostic karyotype (50,XY,+21,+22 inc/46,XY).

In three cases (PD37187, PD43260 & PD43262) additional RNA fusions were found not concordant with the consensus RNA classifier subtype (Supplemental Table 12), these were the same samples identified by Somalier with cross contamination (Supplemental Table 7).

Of the 13 cases, defined by coding gene mutations (4x ZEB2 H1038R, 3x PAX5 P80R, 2x PAX5alt, 2x IDH2 R140Q, ETV6 R399C and WT1 frameshift), RNA classification assigned subtypes to 12/13 cases leaving the ETV6 R399C case (PD40803) unclassified by both WGS and RNA. Of the 12 classified cases 9 were concordant between WGS and RNA derived subtypes (Supplemental Table 13). Of the three discordant cases, two had IDH2 R140Q mutations (PD40809, PD40814) and one had a frameshift WT1 mutation (PD40836). All three cases were classified as PAX5alt by 3/4 RNA classifiers, in comparison 3 cases with bona fide PAX5 alterations were classified as PAX5alt by 4/4 classifiers (Figure 1c, Supplemental Table 11 & 13). Lastly, the B-other case with ETV6 R399C had an additional *KDM6A::CGA* rearrangement identified by RNA-seq fusion detection, not predicted by WGS fusion calling.

Resolution of cytogenetic findings

Cytogenetics analysis of the cohort (n=52) revealed either a normal (n=19), non-specific abnormal (n=23), complex karyotype (n=8) or failed (n=2) karyotype. These cytogenetic categories represent highly heterogeneous genetic subgroups (Figure 2a). WGS analysis identified a driver event in 100% (8/8) of cases with a complex karyotype, classifying them into one of six different genomic subtypes: *DUX4*-r, *MEF2D*-r, *ZEB2*/*CEBP*, *UBTF::ATXN7L3*, *IGK::BCL2* and *IGH::ID4* (Figure 2b). There was no correlation between complex karyotype and genomic subtype; however, 2/3 *MEF2D*-r cases with cytogenetics had a complex karyotype and 3/8 cases had a TP53 mutation with 17p loss (2x *DUX4*-r and *IGK::BCL2*), Supplemental Table 8, Supplemental Figure 8). The same is true for cases with normal, non-specific abnormal events or failed karyotype where we identify diverse genetic lesions to include *DUX4*-r, *ZNF384*-r, *ZEB2*/*CEBP* amongst others. Notably, 56% (9/16) of cases with normal karyotype had either a cytogenetically cryptic genomic rearrangement (7/16) or an altered ploidy class (2/16).

Demographic, clinical and genetic features of B-other genomic subtypes

The majority of *DUX4*-r cases harbored the canonical *IGH::DUX4* rearrangement (7/8) but a single case had *ERG::DUX4* (PD40815) (Supplemental Table 8). Only a single concurrent intragenic *ERG* deletion was identified (PD40800a, 21:39773785-39875948), which has been reported in up to two-thirds of pediatric *DUX4*-r ALL[38] (Supplemental Figure 9). Interestingly, 6/8 (75%) *DUX4*-r were female compared with the study cohort 17/44 (39%). Unlike prior observations in both pediatric and adult ALL suggesting *DUX4*-r was associated with favorable outcome[9,39], while only 1/8 were MRD positive at the end of induction and 5/8 relapsed (Supplemental Figure 10).

We identify 11 cases with *PAX5* abnormalities mapping to PAX5alt (n=7) subtypes and PAX5 P80R (n=4). Notably, 10/11 cases showed evidence of biallelic targeting: by either two mutations (n=4) or a single mutation/fusion and concomitant LOH (n=6) (Supplemental Table 8, Supplemental Figure 11).

Both PAX5alt fusions had corresponding karyotypic events: *ETV6::PAX5* was detected by cytogenetics as dic(9;12)(p13;p13) and *FOXP1::PAX5* presented as add(9)(p13). The other five PAX5alt contained; three cases with R38 and R140 biallelic mutations, one case harboring an intragenic *PAX5* amplification (Supplemental Figure 12a) and one with *PAX5* deletion and a protein truncating mutation (p.M335fs*68) (Supplemental Figure 12b). We also detected four PAX5 P80R; three with arm-level LOH and one with a secondary PAX5 R38H mutation (Supplemental Figure 13). Whilst monoallelic *PAX5* deletions are observed across genomic subgroups, both PAX5alt and PAX5 P80R are demarcated by a dependency on bi-allelic targeting (Supplemental Table 8, Supplemental Figure 6).

Five cases had a *ZNF384* fusion including *EP300::ZNF384* (n=4) and *AKAP8::ZNF384* (n=1). We have recently reported that *ZNF384* patients are typically younger and have good outcomes[10]. As expected, the same is true for patients in this subset. 4/5 (80%) were ≤40 years compared to 18/47 (40%) for the rest of the cohort, only 1/5 were MRD positive at the end of induction and 2/5 relapsed; (Supplemental Figure 10)[10].

WGS analysis identified 4 cases belonging to the *ZEB2*/*CEBP* group: two cases with ZEB2 H1038R and concurrent *IGH::CEBPB* rearrangements and two cases with just ZEB2 H1038R. A fifth ZEB2 H1038R mutation was detected in an *IGH::DUX4* case (PD40800, VAF=∼20) but RNA-seq classified this sample as *DUX4*-r, and the remaining 4 ZEB2 H1038R were all classified as *ZEB2*/*CEBP* (Supplemental Table 11 & 13).

Of the three MEF2D-r cases identified: two cases were complex karyotype (PD40816, *MEF2D::BCL9*; PD40832, *MEF2D::HNRNPUL1*) and exhibited copy number oscillation involving multiple chromosomes (chr1,4,9,11 and chr6,9,13 respectively), indicative of chromothripsis (Figure 3a-b). The remaining case with *MEF2D::BCL9* (PD40810a) failed cytogenetics and had a relatively quiet genome (Figure 3c). However, the *MEF2D::BCL9* rearrangement was a consequence of a highly complex localized structural variant involving an interstitial jump into an intergenic region (Figure 3d). Notably, all three *MEF2D*-r

cases exhibited a copy gain of *MEF2D*, which confounded initial detection by FISH (Supplemental Figure 14). All three *MEF2D*-r cases were classified as high-risk on UKALL14: two on the basis of complex karyotype and a single case on age. In keeping with this high-risk status, all three *MEF2D*-r cases relapsed (Supplemental Figure 10).

We had previously screened patients in this cohort for ABL-class fusions and JAK-STAT abnormalities by FISH and MLPA (Supplemental Figure 1). However, WGS revealed an additional two cases with Ph-like JAK-STAT abnormalities. The first had *IGH::EPOR* which is not detectable by FISH or MLPA while the second had *P2RY8::CRLF2* fusion, which was validated by FISH and MLPA[10].

Four cases harbored newly described subtypes. Two cases (PD40820, PD40824) with *UBTF::ATXN7L3* fusion resulting from a sub-microscopic interstitial deletion at 17q21.31. Both cases also harbored the 13q12.2 deletion which hijacks the *PAN3* enhancer to drive over-expression of *CDX2* (Supplemental Figure 15)[40,41]. Two further cases (PD40809, PD40814) harbored clonal IDH2 p.R140Q mutations (Supplemental Figure 16), a subtype recently described by Yasuda et al[41].

Lastly, 13 cases involved the hijacking of an immunoglobulin gene enhancer (*IGH* x 12 and *IGK* x1). In addition to the aforementioned *DUX4*-r (n=7) and *ZEB2*/*CEBP* (n=2) cases, WGS analysis revealed four additional cases of *IGK::BCL2*, *IGH::CEBPA*, *IGH::ID4* and *IGH::MIR125B1*. The patient with *IGK::BCL2* (PD40797) was classified as “high risk” (age >40yrs, complex karyotype) and although achieving remission post induction, responded poorly to therapy dying shortly after diagnosis (< 200 days). The detection of this abnormality at initial diagnosis may have prompted a re-evaluation of the diagnosis.

Detection of DUX4 rearrangements

Using our custom workflow, eight rearrangements targeting the *DUX4* locus were identified (Supplemental Table 14). RNA-seq data were available for 7/8 cases, allowing the comparison of WGS against RNA fusion and RNA classification (Figure 4a). RNA classification identified all 7 cases, but RNA fusion gene analysis failed to detect two *IGH::DUX4* cases PD37188a and PD40804a. WGS showed

evidence that in both cases the 5' IGH locus translocation mapped to telomeric repeats while the 3' IGH translocation mapped to *DUX4* loci. Therefore, we can infer that in these two cases the *DUX4* locus along with proximal telomeric repeats was inserted into the IGH locus in an inverted orientation which confounds detection by RNA fusion calling (Figure 4b). The *ERG::DUX4* (PD40815a) case also showed expression of an alternative exon 6 (Supplemental Figure 17) as previously described[38]. WGS analysis and *DUX4* gene expression analysis but not RNA-fusion analysis has high specificity for the detection of *DUX4*-r. Implementation of WGS enables the detection of all *DUX4* rearrangements and provides resolution on both the structure of the rearrangement as well as partner genes.

Detection of *IGH* enhancer rearrangements

Using the graph based *IGH* enhancer hijack calling workflow we reclassify eight SV calls as candidate *IGH* events (Supplemental Table 15). Partner genes included recurrent events in *DUX4*, *EPOR*, *CEBPA*, *CEBPB* and the less common *MIR125B1* and *ID4*. Seven *IGH* rearrangements were directly validated by FISH or RNA-seq fusion analysis and a single case (PD43259) had indirect evidence from FISH and karyotype showing a 14q32 deletion correlating to the loss incurred due to an unbalanced translocation to form *IGH::ID4* (Supplemental Table 15, Figure 4c). Last, speaking to the sensitivity of this approach, this workflow identified an *IGH* rearrangement that was validated both by RNA-seq and FISH in a case that failed the purity criteria (20%) for WGS analysis (PD40837, *IGH::CEBPA*) (Supplemental Table 15).

Detection of focal deletions

Focal deletions in *EBF1*, *IKZF1*, *CDKN2A*, *CDKN2B*, *PAX5*, *ETV6*, *BTG1* & *RB1* are strongly associated with ALL pathogenesis. For 48 cases, high quality MLPA and WGS CNA data was available for these eight gene targets[10]. MLPA detected 82 deletions in 32 cases (median deletion count 3 range 1-4). WGS analysis was 100% concordant, by detecting all 82 deletions characterized by MLPA. (Figure 5). Importantly, WGS analyses detected a further 21 deletions in 15 cases missed by MLPA, the vast

majority can be attributed to sample purity/ploidy (n=14) or a subclonal event (n=5), with the final two events not identified by MLPA due to probe placement (Figure 5, Supplemental Table 16).

Patterns of genomic instability in BCP- ALL

Assessment of the genome-wide patterns of mutations present in each leukemia genome allows for the characterization of putative biological processes or environmental exposures that result in genomic instability and contribute to leukemic transformation[31,42].

Structural variant analysis identified B-cell specific processes such as RAG mediated deletions[31] and observed complex SV events such as chromothripsis[43]. RAG median deletion was identified in all cases with a median burden of 7 events (range 1-49). A higher burden of RAG mediated deletions was observed in *DUX4*-r and *ZNF384*-r subtypes (Figure 6a). The highest RAG deletion burden was identified in a single *IGH::MIR125B1* case (PD37197) with 49/69 deletions attributable to RAG activity (Supplemental Figure 18). The subtype with the lowest contribution for RAG mediated deletions was *MEF2D*-r, this correlates with the previous observation that RAG1 is downregulated in *MEF2D*-r[44]. We also observed outlier high telomere length in *MEF2D*-r (Figure 6a).

Analysis of single base substitution (SBS) patterns for established mutation signatures revealed that the majority of SNVs mapped to SBS Blood signature (Figure 6b,c). SBS Blood is a clock-like signature operative in hematopoietic stem cells enriched in C>T mutations[30,45]. SBS Blood was detected in 51/52 samples in our cohort and demonstrated a linear relationship with age explaining ~16 mutations per year of life (Supplemental Figure 19). The only exception was an *IGH::DUX4* case with homozygous loss of *MSH6* that was defined by hypermutation (PD40812 with 77,183 SNVs) enriched in SBS1, a phenotype previously observed in hypermutator cases[46,47]. Although the sample mutational profile was atypical for SBS1 (Cosine similarity = 0.941) and a much better fit was the thio-dMMR signature (Cosine similarity = 0.996) (Supplemental Figure 20a,b). A signature seen in relapsed pediatric hypermutator ALL cases and recently attributed to MMR deficient ALL treated with thiopurine[48,49]. The patient had a

history of Crohn's disease, where thiopurines are a common therapy. SBS9 is a signature dominated by T>G and T>C mutations attributed to replication errors by polymerase η during somatic hypermutation in lymphoid cells[50]. In our cohort, SBS9 is observed in four cases; both low hypodiploid cases, an *IGH::CEBPB* and the *IGK::BCL2* case with the highest mutation burden (42%) which we would expect given association of this lesion with mature B-cell neoplasms. Lastly, SBS18, a signature dominated by C>A mutations and attributed to reactive oxygen species, was enriched in *MEF2D*-r, *PAX5* and *IDH1/2* subtypes.

Discussion

The diagnostic work up of adult ALL does not incorporate WGS in the vast majority of centers worldwide. Here, we assess the utility of WGS profiling in 52 adult ALL cases, that in the absence of informative biomarkers by standard of care were classified as B-other.

WGS assigned 88% (46/52) of cases called B-other to an established genetic subtype of ALL, with ~20% (10/46) of subtypes being assigned solely via novel WGS workflows developed in this study. This included five cases with WHO[2016] subtypes crucially, 3/5 cases harbored high-risk genetic events that would have changed their UKALL14 risk group and post-induction treatment in the absence of other risk factors. Among the remaining B-other cases, 87% (41/47) were assigned to one of the newly described genetic subtypes of ALL[9,39,41]. Our recent studies have proposed that several of these subtypes are linked to a good (e.g. *ZNF384*-r) or poor (e.g. JAK-STAT) outcome[10,39].

Our findings demonstrate that unless cytogenetics reveals an established genetic rearrangement or ploidy subtype, a designation of failed, normal or complex karyotype frequently misses subtype defining events that can readily be picked up by WGS, as also demonstrated from our related studies in childhood ALL[51,52]. For example, the presence of a complex karyotype, defined as ≥5 chromosomal abnormalities, has been associated with a poor prognosis[2,10]. WGS identified a driver event in all eight cases, revealing a heterogeneous spectrum of drivers, linked both to favorable (*DUX4*-r) and poor

(*MEF2D*-r and *UBTF::ATXN7L3*) outcomes. This suggests that complex karyotype is not a robust classification.

WGS also identified events which could also alter clinical management. One *IGH::DUX4* case had hypermutation caused by homozygous loss of *MSH6,* a candidate for checkpoint inhibition. A second case had homozygous loss of CD58 and concurrent LOH of *HLA-B* which would likely confer immune escape[53,54] and reduce efficacy of CAR T-cell therapy[55,56]. Lastly, a near-triploidy case harbored a germline *TP53* mutation, which has implications for carrier screening.

Integration of RNA-seq for fusion gene detection and gene expression classification enables concomitant validation of WGS based findings using an orthogonal assay. This is of particular importance for transcriptomically defined subtypes such as Ph-like or PAX5alt where WGS on its own would only be able to evidence previously defined recurrent DNA lesions for these phenocopy subtypes. RNA-seq is also able to directly identify fusions caused by complex SVs which confound standard WGS fusion calling.

As this study was retrospective, we used remission samples to source germline DNA. To avoid tumor contamination, we selected samples with MRD negative results or MRD <1%. Therefore, the observed genomic subtype frequencies may not reflect the true distribution in adult B-ALL. In this study, we focused on cases that did not harbor informative clinical biomarkers, classified as B-other. We did not formally assess the performance of WGS and RNA-seq against standard of care for the patients where an ALL subtype was previously determined.

For clinical implementation, detailed laboratory analytical and clinical validity studies are warranted to delineate standardization metrics for diagnostic assay deployment and optimal source of normal DNA in order to determine assays performance against standard of care molecular diagnostic assays across the spectrum of B-ALL subtypes.

Acknowledgements

The authors thank all the participating sites, local investigators, and research teams for their ongoing participation in the study, together with patients who took part in this trial as well as their families. We acknowledge the input of all the scientists and technicians working in the adult ALL MRD laboratory based at UCL. The UKALL14 trial was coordinated by the CRUK & UCL Cancer Trials Centre and funded by CRUK (C27995/A9609, A.K.F.). This study was supported by research grants from Cancer Research UK (A21019, A.V.M. and A.K.F.) and Blood Cancer UK (15036, A.V.M. and C.J.H.). We thank the member laboratories of the UK Cancer Cytogenetic Group for cytogenetic data and material. This research benefited from core institutional funding to the Wellcome Sanger Institute. E.P. is a Josie Robertson Investigator and is supported by the MDS foundation and is a Damon-Runyon Rachleff Innovator Award recipient.

Authorship

Contributions

Conception and Design: P.C., B.P., A.V.M., A.K.F. and E.P.

Financial support: C.J.H., P.C., B.P., A.V.M., A.K.F. and E.P.

Provision of study patients: S.L., K.Z.A., A.A.K., L.C.H., P.P. and A.K.F.

Collection and assembly of data: D.L., S.L., E.B., T.C., K.Z.A., A.A.K., L.C.H., P.P., A.V.M. and A.K.F.

Laboratory processing of biospecimens and sequencing: S.L. and L.O.

Processed, analyzed, and interpreted data: D.L., J.G.A., G.G., M.F.L., J.E.A.O., K.L., J.S.M.M., D.J., A.P.B., P.C., A.V.M. and E.P.

Wrote the manuscript: D.L., A.V.M., A.K.F. and E.P.

Final approval of manuscript: All authors.

Disclosure of conflict of interests

G.G. is a consultant at Isabl Inc.

M.F.L. is an employee, equity holder in Isabl Inc.

J.S.M.M. is a founder, equity holder in Isabl Inc.

A.A.K has received honoraria from Kite.

L.C.H on behalf of CRUK and UCL CTC has received research funding from Astra Zeneca, GSK, Pfizer, MSD, BMS, Amgen, Millennium Takeda.

P.J.C. is a co-founder, stock-holder and consultant for FL86 Inc

B.P. is on the advisory board for Pfizer.

A.V.M has received honoraria from Amgen.

A.K.F. has served as a consultant for Amgen, Pfizer and Novartis.

E.P. is a founder, equity holder and holds fiduciary role in Isabl Inc. and is an equity holder in TenSixteen Bio.

Current affiliations

The current affiliation for S. L. is St. George's University Hospital, NHS Foundation Trust and the current affiliation for K. Z-A. is Barts Health NHS Trust, The Royal London Hospital.

References

1. Tanasi I, Ba I, Sirvent N, et al. Efficacy of tyrosine kinase inhibitors in Ph-like acute lymphoblastic leukemia harboring ABL-class rearrangements. *Blood*. 2019;134(16):1351–1355.
2. 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.
3. Marks DI, Moorman AV, Chilton L, et al. The clinical characteristics, therapy and outcome of 85 adults with acute lymphoblastic leukemia and t(4;11)(q21;q23)/MLL-AFF1 prospectively treated in the UKALLXII/ECOG2993 trial. *Haematologica*. 2013;98(6):945–952.
4. Marks DI, Kirkwood AA, Rowntree CJ, et al. Addition of four doses of rituximab to standard induction chemotherapy in adult patients with precursor B-cell acute lymphoblastic leukaemia (UKALL14): a phase 3, multicentre, randomised controlled trial. *Lancet Haematol*. 2022;9(4):e262–e275.
5. 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.
6. Lilljebjörn 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.
7. Gu Z, Churchman M, Roberts K, et al. Genomic analyses identify recurrent MEF2D fusions in acute lymphoblastic leukaemia. *Nat. Commun.* 2016;7:13331.
8. Li J-F, Dai Y-T, Lilljebjörn H, et al. Transcriptional landscape of B cell precursor acute lymphoblastic leukemia based on an international study of 1,223 cases. *Proc. Natl. Acad. Sci. U. S. A.* 2018;115(50):E11711–E11720.
9. Gu Z, Churchman ML, Roberts KG, et al. PAX5-driven subtypes of B-progenitor acute lymphoblastic leukemia. *Nat. Genet.* 2019;51(2):296–307.
10. Moorman AV, Barretta E, Butler ER, et al. Prognostic impact of chromosomal abnormalities and copy number alterations in adult B-cell precursor acute lymphoblastic leukaemia: a UKALL14 study. *Leukemia*. 2022;36(3):625–636.
11. Ye K, Schulz MH, Long Q, Apweiler R, Ning Z. Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. *Bioinformatics*. 2009;25(21):2865–2871.
12. Nik-Zainal S, Van Loo P, Wedge DC, et al. The life history of 21 breast cancers. *Cell*. 2012;149(5):994–1007.
13. Nik-Zainal S, Davies H, Staaf J, et al. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature*. 2016;534(7605):47–54.
14. Jones D, Raine KM, Davies H, et al. cgpCaVEManWrapper: Simple Execution of CaVEMan in Order to Detect Somatic Single Nucleotide Variants in NGS Data. *Curr. Protoc. Bioinformatics*. 2016;56:15.10.1–15.10.18.
15. Raine KM, Van Loo P, Wedge DC, et al. ascatNgs: Identifying Somatically Acquired Copy-Number Alterations from Whole-Genome Sequencing Data. *Curr. Protoc. Bioinformatics*. 2016;56:15.9.1–15.9.17.
16. Chakravarty D, Gao J, Phillips SM, et al. OncoKB: A Precision Oncology Knowledge Base. *JCO Precis Oncol*. 2017;2017.:
17. Schmidt B, Brown LM, Ryland GL, et al. ALLSorts: a RNA-Seq classifier for B-Cell Acute Lymphoblastic Leukemia. *bioRxiv*. 2021;2021.08.01.454393.
18. Mäkinen V-P, Rehn J, Breen J, Yeung D, White DL. Multi-Cohort Transcriptomic Subtyping of B-Cell Acute Lymphoblastic Leukemia. *Int. J. Mol. Sci.* 2022;23(9.):
19. Nicorici D, Şatalan M, Edgren H, et al. FusionCatcher – a tool for finding somatic fusion genes in paired-end RNA-sequencing data. *bioRxiv*. 2014;011650.
20. Tian L, Li Y, Edmonson MN, et al. CICERO: a versatile method for detecting complex and diverse driver fusions using cancer RNA sequencing data. *Genome Biol.* 2020;21(1):126.
21. Robinson JT, Thorvaldsdóttir H, Winckler W, et al. Integrative genomics viewer. *Nat. Biotechnol.* 2011;29(1):24–26.
22. Medina-Martínez JS, Arango-Ossa JE, Levine MF, et al. Isabl Platform, a digital biobank for processing multimodal patient data. *BMC Bioinformatics*. 2020;21(1):549.
23. Hnisz D, Abraham BJ, Lee TI, et al. Super-enhancers in the control of cell identity and disease. *Cell*.

2013;155(4):934–947.
24. Frankish A, Diekhans M, Ferreira A-M, et al. GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res.* 2019;47(D1):D766–D773.
25. Huret J-L, Ahmad M, Arsaban M, et al. Atlas of genetics and cytogenetics in oncology and haematology in 2013. *Nucleic Acids Res.* 2013;41(Database issue):D920–4.
26. Mitelman F JBAMF (eds). Mitelman Database of Chromosome Aberrations and Gene Fusions in Cancer (2022).
27. Cameron DL, Schröder J, Penington JS, et al. GRIDSS: sensitive and specific genomic rearrangement detection using positional de Bruijn graph assembly. *Genome Res.* 2017;27(12):2050–2060.
28. James Kent W. BLAT—The BLAST-Like Alignment Tool. *Genome Res.* 2002;12(4):656–664.
29. Blokzijl F, Janssen R, van Boxtel R, Cuppen E. MutationalPatterns: comprehensive genome-wide analysis of mutational processes. *Genome Med.* 2018;10(1):33.
30. Machado HE, Mitchell E, Øbro NF, et al. Diverse mutational landscapes in human lymphocytes. *Nature*. 2022;
31. Papaemmanuil E, Rapado I, Li Y, et al. RAG-mediated recombination is the predominant driver of oncogenic rearrangement in ETV6-RUNX1 acute lymphoblastic leukemia. *Nature Genetics*. 2014;46(2):116–125.
32. Bailey TL, Johnson J, Grant CE, Noble WS. The MEME Suite. *Nucleic Acids Res.* 2015;43(W1):W39–49.
33. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2020.
34. Venables WN, Ripley BD. Modern Applied Statistics with S. *Statistics and Computing*. 2002;
35. Shomali W, Gotlib J. World Health Organization-defined eosinophilic disorders: 2019 update on diagnosis, risk stratification, and management. *Am. J. Hematol.* 2019;94(10):1149–1167.
36. Zhang MY, Churpek JE, Keel SB, et al. Germline ETV6 mutations in familial thrombocytopenia and hematologic malignancy. *Nat. Genet.* 2015;47(2):180–185.
37. Li Y, Schwab C, Ryan S, et al. Constitutional and somatic rearrangement of chromosome 21 in acute lymphoblastic leukaemia. *Nature*. 2014;508(7494):98–102.
38. Zhang J, McCastlain K, Yoshihara H, et al. Deregulation of DUX4 and ERG in acute lymphoblastic leukemia. *Nat. Genet.* 2016;48(12):1481–1489.
39. 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.
40. Kimura S, Montefiori L, Iacobucci I, et al. Enhancer retargeting of CDX2 and UBTF::ATXN7L3 define a subtype of high-risk B-progenitor acute lymphoblastic leukemia. *Blood*. 2022;139(24):3519–3531.
41. Yasuda T, Sanada M, Kawazu M, et al. Two novel high-risk adult B-cell acute lymphoblastic leukemia subtypes with high expression of CDX2 and IDH1/2 mutations. *Blood*. 2022;139(12):1850–1862.
42. Alexandrov LB, Nik-Zainal S, Wedge DC, et al. Signatures of mutational processes in human cancer. *Nature*. 2013;500(7463):415–421.
43. Stephens PJ, Greenman CD, Fu B, et al. Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell*. 2011;144(1):27–40.
44. Liu Y-F, Wang B-Y, Zhang W-N, et al. Genomic Profiling of Adult and Pediatric B-cell Acute Lymphoblastic Leukemia. *EBioMedicine*. 2016;8:173–183.
45. Lee-Six H, Øbro NF, Shepherd MS, et al. Population dynamics of normal human blood inferred from somatic mutations. *Nature*. 2018;561(7724):473–478.
46. Sanders MA, Chew E, Flensburg C, et al. MBD4 guards against methylation damage and germ line deficiency predisposes to clonal hematopoiesis and early-onset AML. *Blood*. 2018;132(14):1526–1534.
47. Meier B, Volkova NV, Hong Y, et al. Mutational signatures of DNA mismatch repair deficiency in *C. elegans* and human cancers. *Genome Research*. 2018;28(5):666–675.
48. Waanders E, Gu Z, Dobson SM, et al. Mutational landscape and patterns of clonal evolution in relapsed pediatric acute lymphoblastic leukemia. *Blood Cancer Discov*. 2020;1(1):96–111.
49. Yang F, Brady SW, Tang C, et al. Chemotherapy and mismatch repair deficiency cooperate to fuel TP53 mutagenesis and ALL relapse. *Nature Cancer*. 2021;2(8):819–834.
50. Puente XS, Pinyol M, Quesada V, et al. Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. *Nature*. 2011;475(7354):101–105.
51. Schwab C, Cranston RE, Ryan SL, et al. Integrative genomic analysis of childhood acute lymphoblastic leukaemia lacking a genetic biomarker in the UKALL2003 clinical trial. *Leukemia*. 2022;
52. Ryan SL, Peden JF, Kingsbury Z, et al. Whole genome sequencing provides comprehensive genetic testing in childhood B-cell acute lymphoblastic leukaemia. *Leukemia*. 2023;1–11.

53. Frangieh CJ, Melms JC, Thakore PI, et al. Multimodal pooled Perturb-CITE-seq screens in patient models define mechanisms of cancer immune evasion. *Nat. Genet.* 2021;53(3):332–341.
54. Ho P, Melms JC, Rogava M, et al. The CD58:CD2 axis is co-regulated with PD-L1 via CMTM6 and governs anti-tumor immunity. *bioRxiv*. 2022;2022.03.21.485049.
55. Yan X, Chen D, Ma X, et al. CD58 loss in tumor cells confers functional impairment of CAR T cells. *Blood Adv*. 2022;
56. Majzner RG, Frank MJ, Mount C, et al. CD58 Aberrations Limit Durable Responses to CD19 CAR in Large B Cell Lymphoma Patients Treated with Axicabtagene Ciloleucel but Can be Overcome through Novel CAR Engineering. *Blood*. 2020;136(Supplement 1):53–54.

Tables

No main tables

Figure Legends

Figure 1: Comprehensive genome profiling utilizing WGS

a) Heatmap describing the presence of genomic markers by WGS in 52 cases, with samples in columns and markers in rows; 1. Focal deletions (<10Mb) in tumor suppressor genes (TSG), 2. Small mutations (SNV/INDEL) in cancer genes, 3. Recurrent gene rearrangements seen in ALL, 4. Arm-level CNA (asterisk denotes samples with sub-optimal copy number profiles), 5. Outlier single base substitution (SBS) mutational signature profile, 6. Germline variant and 7. Outlier tumor mutation burden (TMB). b) Oncoplot of true B-other cases (n=47) showing “Subtype defining lesions'' and recurrent “Deletion events'' and genes recurrently targeted by “Small mutations'' (SNV/INDEL) in OncoKB™ or putative drivers identified in the “Other” category. c) Comparison of WGS defined subtypes to RNA expression subtype classification and RNA fusion calling in 31 cases with both DNA and RNA available.

Figure 2: Resolution of cytogenetic findings

Resolution of cytogenetic finding in 52 cases: a) Barplot of standard of care cytogenetic classification. b) Somatic event profile per cytogenetic group describing occurrence of; “Driver”, “arm-level CNAs” and “coding co-mutations” (y-axis denotes group-wise frequency).

Figure 3: Complexity of MEF2D genomes in B-ALL

Genomic landscape of MEF2D rearranged B-ALL: a-b) Circos plots showing the high SV burden of two MEFDr cases with chromothripsis. The circos plot shows four concentric panels, the outermost ring shows the chromosome ideogram. The second ring shows the intermutation distance for all SNVs color-coded by the pyrimidine partner of the mutated base (C>A blue, C>G black, C>T red, T>A grey, T<C green and T>G pink). The third ring shows small insertions (dark green) and deletions (dark red). The fourth ring shows copy number changes gains (green) and losses (red). Intra and inter chromosomal SVs are shown by arcs describing translocations (black), inversions (blue), deletions (red) and duplications (green). Circo plot c) and locus plot of 1q21 d) showing the "quiet" global SV burden and complex causative local SV event in an *MEF2D::BCL9* case. In the locus plot d); the upper panel describes the SV events as arcs between two breakpoint loci, with the color denoting the type of SV, the middle panel shows the absolute copy number and the lower panel shows the two causative SVs (inversion and deletion) overlaying the transcript structure of *BCL9* and *MEF2D*.

Figure 4: Identification of *DUX4* and *IGH@* rearranged by WGS

a) Upset plot detailing *DUX4* rearranged event identification in 7 cases with both WGS and RNA data. b) Graphical representation of a complex *IGH::DUX4* locus configuration detected by WGS only in two cases (PD37188 & PD40804) which confounds RNA fusion calling. Upper and lower loci represent the wild type (WT) *IGH* locus and *DUX4* loci (4q and 10q). The middle represents the *IGH::DUX4* loci, where an inverted insertion of a *DUX4* loci including the proximal telomere into the IGH locus confounds RNA fusion calling. c) Heatmap of *IGH@* events identified by WGS and correlative evidence from RNA fusion calling, RNA expression classification and cytogenetics (karyotype and FISH).

Figure 5: MLPA and WGS copy number comparison

Tile plot comparing deletions across MLPA (P335) target genes called by MLPA and WGS (ASCAT & BRASS). For each case, events recovered by MLPA and WGS are shown as paired columns. Events missed by MLPA are marked by a gray tile; additionally, two events missed due to MLPA probe placement are marked by an asterisk.

Figure 6: Mutational patterns

a) Dotplot showing RAG burden (total count and % of all deletions) and telomere length for each subtype. b) Plot showing total SNV burden (upper panel) and normalized SBS signature exposure (lower panel). c) 96 context mutational profiles of SBS1, SBS9, SBS18 and SBS blood.

Figure 1. Global cohort overview

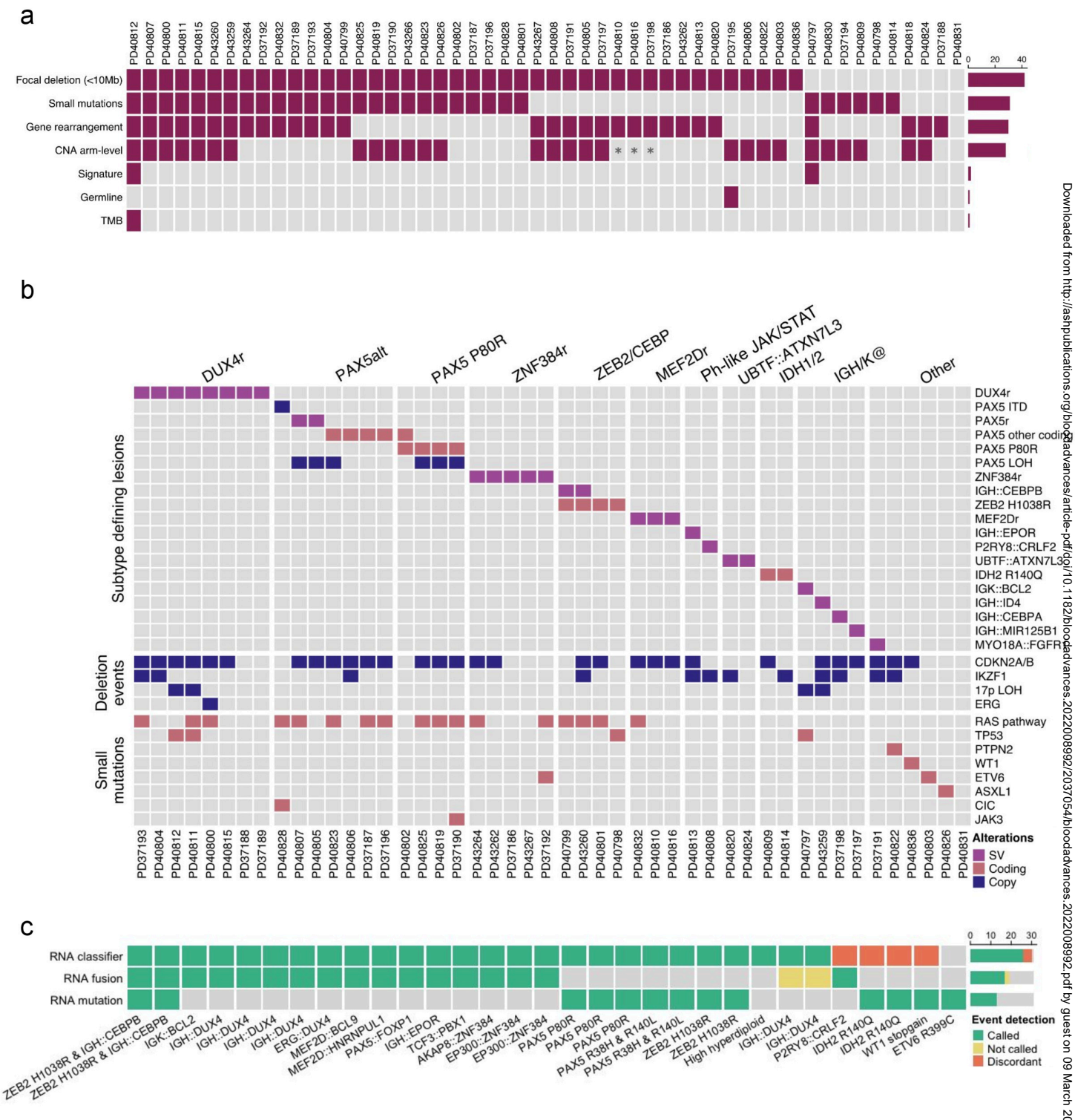

Figure 2: Resolution of cytogenetic findings

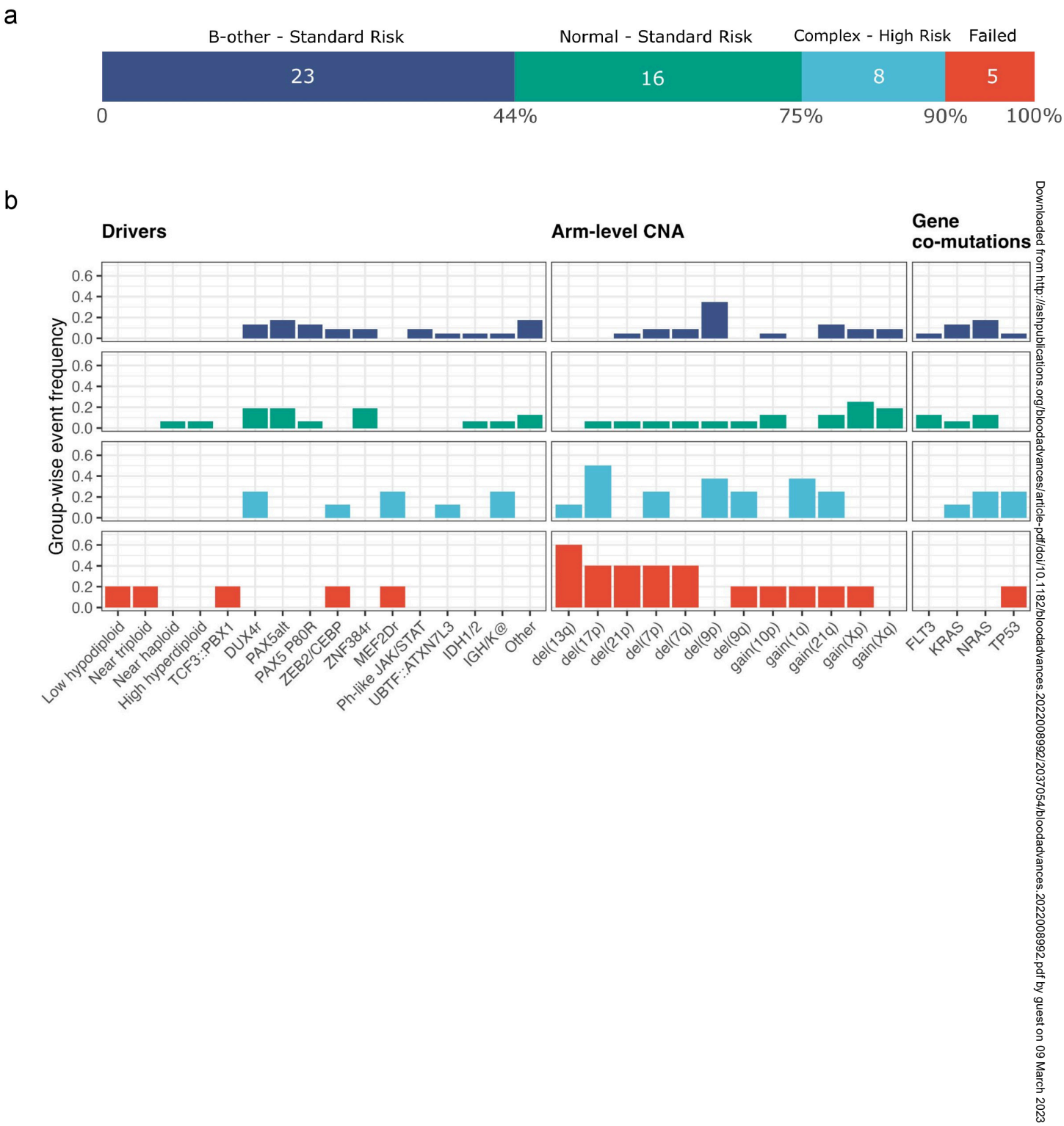

Figure 3. Complexity of MEF2D genomes in B-ALL.

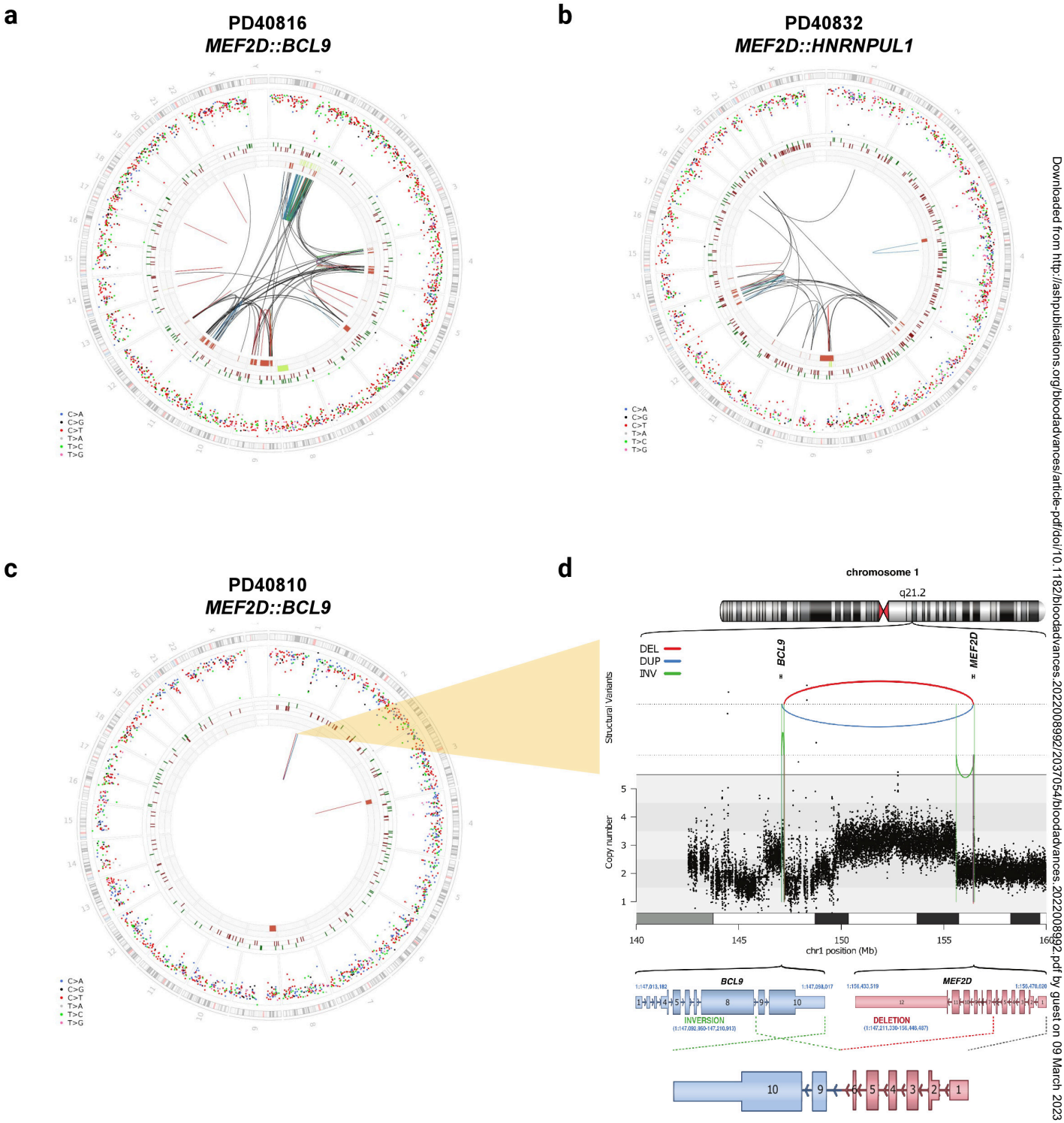

Figure 4. Identification of DUX4 and IGH@ rearranged by WGS

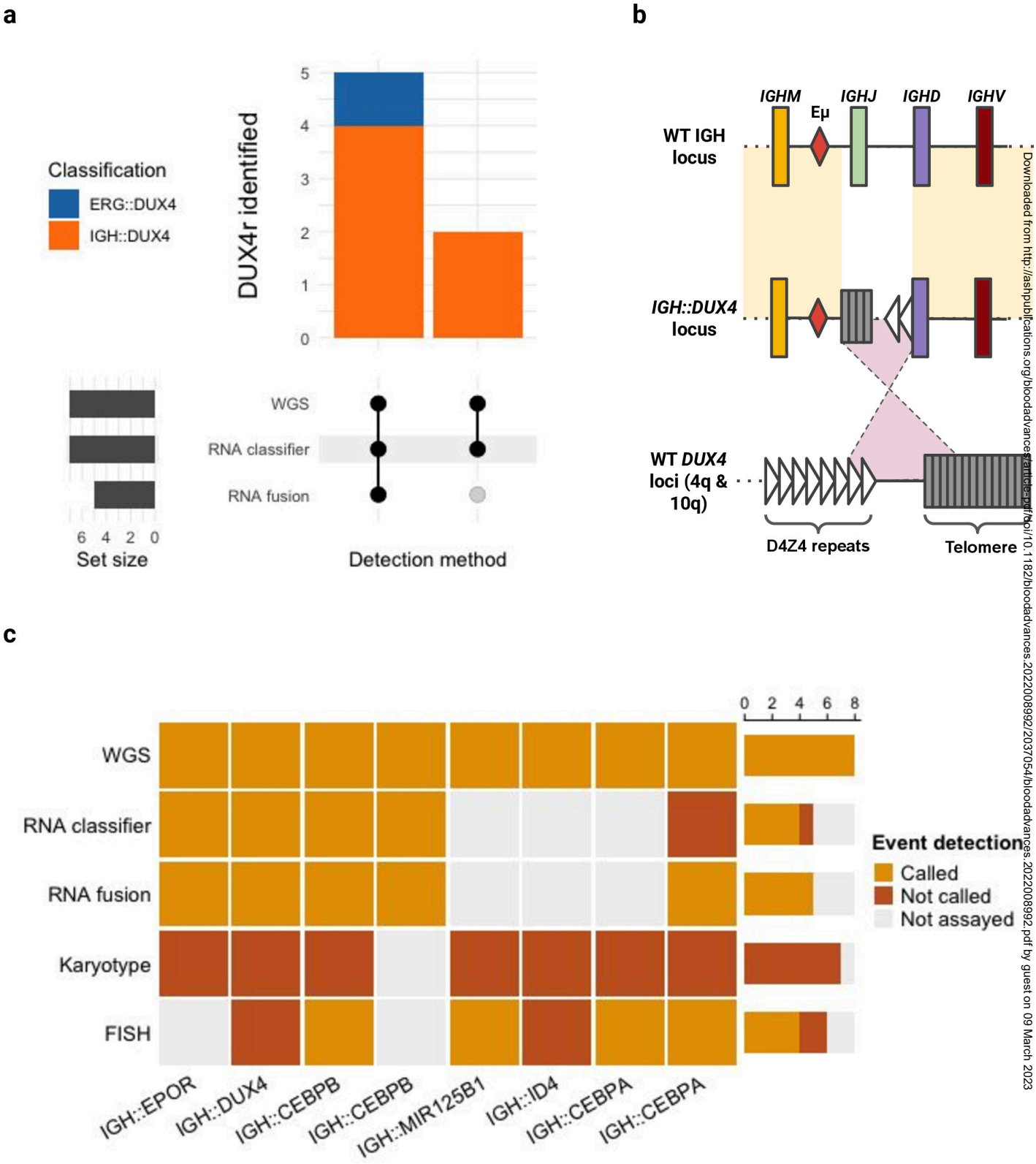

Figure 5: MLPA and WGS copy number comparison

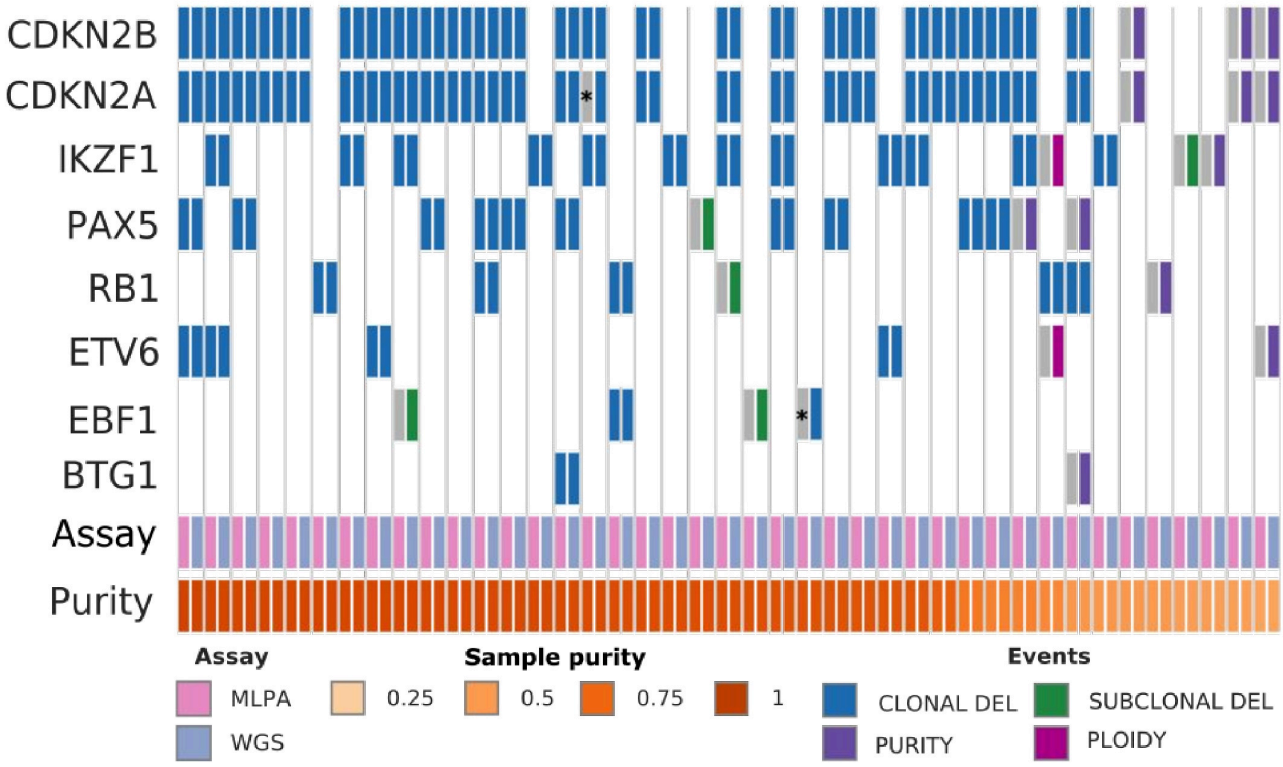

Figure 6: Mutational patterns