RESEARCH ARTICLE

WILEY

Characterisation of the genomic landscape of CRLF2-rearranged acute lymphoblastic leukemia

Lisa J. Russell¹ | Lisa Jones¹ | Amir Enshaei¹ | Stefano Tonin¹ | Sarra L. Ryan¹ | Jeyanthy Eswaran¹ | Sirintra Nakjang² | Elli Papaemmanuil^{3,4} | Jose M. C. Tubio⁴ | Adele K. Fielding⁵ | Ajay Vora⁶ | Peter J. Campbell⁴ | Anthony V. Moorman^{1†} | Christine J. Harrison^{1†}

Correspondence

Dr Lisa J Russell, Wolfson Childhood Cancer Research Centre. Northern Institute for Cancer Research, Newcastle University, Level 6, Herschel Building, Brewery Lane, Newcastle upon Tyne, NE1 7RU. Email: lisa.russell@newcastle.ac.uk

Funding information

Support by: The Kay Kendall Leukaemia Fund, Leuka, European Haematology Association and Bloodwise (formerly Leukaemia and Lymphoma Research)

Abstract

Deregulated expression of the type I cytokine receptor, CRLF2, is observed in 5-15% of precursor B-cell acute lymphoblastic leukaemia (B-ALL). We aimed to determine the clinical and genetic landscape of those with IGH-CRLF2 or P2RY8-CRLF2 (CRLF2-r) using multiple genomic approaches. Clinical and demographic features of CRLF2-r patients were characteristic of B-ALL. Patients with IGH-CRLF2 were older (14 y vs. 4 y, P < .001), while the incidence of CRLF2-r among Down syndrome patients was high (50/161, 31%). CRLF2-r co-occurred with primary chromosomal rearrangements but the majority (111/161, 69%) had B-other ALL. Copy number alteration (CNA) profiles were similar to B-other ALL, although CRLF2-r patients harbored higher frequencies of IKZF1 (60/138, 43% vs. 77/1351, 24%) and BTG1 deletions (20/138, 15% vs. 3/1351, 1%). There were significant differences in CNA profiles between IGH-CRLF2 and P2RY8-CRLF2 patients: IKZF1 (25/35, 71% vs. 36/108, 33%, P < .001), BTG1 (11/35, 31% vs. 10/108, 9%, P = .004), and ADD3 deletions (9/19, 47% vs. 5/38, 13%, P = .008). A novel gene fusion, USP9X-DDX3X, was discovered in 10/54 (19%) of patients. Pathway analysis of the mutational profile revealed novel involvement for focal adhesion. Although the functional relevance of many of these abnormalities are unknown, they likely activate additional pathways, which may represent novel therapeutic targets.

1 | INTRODUCTION

Acute lymphoblastic leukaemia (ALL) is defined by primary chromosomal abnormalities that drive disease progression, with impact on prognosis and treatment stratification. One quarter of patients, known as B-other ALL, lack a known primary abnormality. A group of B-other patients, known as Ph-like/BCR-ABL1-like, constitute 10-15% of B-ALL. Although they lack the BCR-ABL1 fusion, their gene expression profile is similar to

BCR-ABL1 positive ALL.^{2,3} They are characterised by high expression of the type I cytokine receptor, cytokine receptor-like factor 2 (CRLF2), the presence of tyrosine kinase fusion genes and mutations of genes within the JAK/STAT and RAS signaling pathways.⁴ Deregulated expression of CRLF2 (CRLF2-d) is observed in 27-50% of patients with BCR-ABL1-like disease (5-15% of B-ALL). 1,4-6 Deregulation occurs via three genomic rearrangements (CRLF2-r): a cryptic reciprocal translocation with the immunoglobulin heavy chain locus (IGH); an interstitial deletion within

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

¹Leukaemia Research Cytogenetics Group, Northern Institute for Cancer Research, Newcastle University, Newcastle-upon-Tvne, UK

²Bioinformatics Support Unit, Newcastle University, Newcastle-upon-Tyne, UK

³Department of Epidemiology-Biostatistics, Memorial Sloan Kettering Cancer Center, USA

⁴Cancer Genome Project, Wellcome Trust Sanger Institute, Hinxton, UK

⁵Research Department of Haemaoloty, UCL Cancer Institute, London, UK

⁶Department of Haematology, Sheffield Children's Hospital, Sheffield, UK

[†]These authors contributed equally to this study.

the pseudoautosomal region (PAR1) of chromosomes X and Y (P2RY8-CRLF2); rare but recurrent CRLF2 mutations. All three CRLF2-r result in overexpression of CRLF2 mRNA and protein; however, alone they are insufficient to cause overt leukaemia.^{5,7} Interestingly, studies identifying CRLF2-d patients by mRNA and protein expression have shown that some patients do not harbor one of the three known genomic rearrangements.⁸ The cause of this overexpression is currently unknown. The incidence of CRLF2-r is high in patients with Down syndrome ALL (DS-ALL) (>50%) and intrachromosomal amplification of chromosome 21 (iAMP21) (25%), 5,7,9-11 with other established cytogenetic abnormalities rarely associated. However, data indicating whether CRLF2-r is a primary or secondary event are scarce. 12 It is well documented that deletions of genes involved in B-cell differentiation and cell cycle control are recurrent in these patients. 5,13-15 Mutations of the Janus kinase family, in particular JAK2, and mutations of IL7R, are also recurrently observed and together result in IL-3 independent growth of mouse BaF3 cells. 11,16 Recently, mutations affecting other kinase genes have also been reported in CRLF2-d ALL. 4,17 Both BCR-ABL1-like and CRLF2-d ALL have been associated with poor outcome and increased risk of relapse.^{2,3,10} Although MRD-directed treatment intensification improves survival, ¹⁸ outcome for CRLF2-d ALL remains less favorable than for patients with good risk cytogenetics. Hence, novel therapeutic strategies are required to improve survival and quality of life. In this study, we have explored the clinical and genetic landscape of patients with known CRLF2 rearrangements. The aims of this study were: (1) to identify the clinical and genomic differences that may exist between patients with IGH or P2RY8 driven overexpression of CRLF2; (2) to undertake pathway analysis of whole genome (WGS) and whole exome sequencing (WES) data to highlight additional pathways that may co-operate with rearrangements of CRLF2.

2 | MATERIALS AND METHODS

2.1 | Patient samples

We identified 172 patients with *CRLF2*-r ALL by fluorescence in situ hybridisation (FISH) and multiplex ligation-dependent probe amplification (MLPA) from the following trials: UKALL97/99 (n=68), UKALL2003 (n=75), UKALLXI (n=6), and UKALLXII (n=26) (Supporting Information Table 1). The patients included in this study were identified by screening for *CRLF2*-r^{6,14,19-21} and were representative of the trial. Demographic and clinical details were collected by the Clinical Trial Service Unit (CTSU, Oxford University, UK). Each contributing centre obtained relevant ethical approval. Informed consent was obtained in accordance with the Declaration of Helsinki. Diagnostic immunophenotypes were collected centrally with review of original reports for 50 patients (Supporting Information Table 2).

2.2 | Cytogenetics and fluorescence in situ hybridisation

For details on genetic testing, see Supporting Information Figure 1. Karyotype data were collected from UK cytogenetics laboratories (Supporting Information Table 1). Cytogenetic analysis and FISH were carried out on the same diagnostic patient samples. The involvement of *IGH* was determined using the LSI IGH Dual Color Break-Apart Rearrangement Probe (Abbott Molecular, Green Oaks, IL). Additional FISH probes used in this study are shown in Supporting Information Table 3. Five control slides of fixed cells from normal individuals were hybridised with all probe combinations (Supporting Information Table 3) to determine the cut-off percentages for false positive results (\pm 3 \times standard deviations). A minimum of 100 nuclei were scored by eye for each FISH test by two independent analysts. When combining three or more florophores, capture and scoring was carried out using an automated Olympus BX-61 8-bay stage florescence microscope. Images were analysed using the Cyto-Vision 7.1 SPOT counting system (Leica Microsystems, Gateshead, UK).

2.3 | Multiplex ligation-dependent probe amplification

Copy number alterations (CNA) were investigated (n=154) by MLPA using the SALSA MLPA kit P335 (MRC Holland, Amsterdam, Netherlands) according to manufacturer's guidelines as previously reported.¹⁵

2.4 | Affymetrix genome-wide human SNP6.0 array

Sufficient material was available from 26 (representitive for cohort) *CRLF2*-r patients (15 with matched germ-line sample) for Affymetrix Genome-Wide Human SNP6.0 analysis, performed by AROS Applied Biotechnology A/S (Aarhus, Denmark). Copy number alterations were analysed using Genotyping Console (Affymetrix software) with additional manual curation (GEO accession number GSE83272).

2.5 | Low depth paired-end and whole exome sequencing

Sufficient diagnostic DNA was available from 11 *CRLF2*-r patients (representitive of cohort and all patients have SNP analysis completed), which was prepared for library construction (300–500 bp), flow cell preparation and cluster formation using the Illumina no-PCR library protocol (Illumina Inc, San Diego, CA), and 50 bp reads were performed using the Illumina Genome Analyzer IIx instrument following the manufactures guidelines. Structural variants were selected as previously reported.²³

Sufficient diagnostic and germ-line DNA from the same 11 patients was prepared for Illumina paired-end sequencing with subsequent exome enrichment using the Agilent SureSelect Human All Exon 50 MB kit (Agilent Technologies LTD, Berkshire, UK). Guidelines for the Illumina Genome Analyzer IIx instrument were followed for flow-cell preparation, cluster generation and paired-end sequencing of 75 bp reads.^{23,24} Whole

[†]These authors contributed equally to this study.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

^{© 2017} Wiley Periodicals, Inc.

genome and exome sequencing data are available using EGA accession numbers EGAD00001002007 and EGAD00001002008, respectively.

2.6 | Structural variant detection and validation

Sequencing reads were mapped to the human genome (hg19) with a minimum and mean fold coverage of 5.61 and 8.46, respectively. The algorithm BRASS²⁵ was used to identify groups of ≥2 discordantly mapped (distance or orientation) paired-end reads. Rearrangements observed in 5 or more paired-end reads were validated (excluding those where both reads were in the same intron or outside the coding regions) by conventional PCR and Sanger sequencing using diagnostic and remission DNA. If a rearrangement failed to validate using two sets of primers, repeated twice over a temperature gradient, it was regarded as a fail. Breakpoints within immunoglobulin loci were not validated. MLPA and SNP data were also used to validate CNA. A total of 476 SV were identified from WGS of 11 patients with *CRLF2*-r ALL (average 43 SV/patient, range 22–90), including 416 intrachromosomal and 57 interchromosomal rearrangements.

2.7 | Mutation and in/del detection and validation

A total of 458 SNVs and insertion/deletions (in/dels) were identified from the same 11 patients. A minimum sequencing coverage of 30-fold was required for each sample. Default setting of BWA²⁴ and CaVEMan was used to align the reads and detect somatically acquired single nucleotide substitutions as previously reported²³. The algorithm PIN-DEL²⁶ was used to detect in/dels.²³ Conventional PCR was used to validate the substitutions, with subsequent 454 pyrosequencing for confirmation.²³ Conventional PCR validated the in/dels, with subsequent sequencing using the ABI terminator Cycle Sequencing Kit (Applied Biosystems) to confirm each in/del.²³

2.8 | Targeted JAK mutation screening

Primers were used to amplify exon 14 of the *JAK2* and *JAK1* genes only²⁷ and Sanger sequencing confirmed the presence of the mutation. As far as we are aware *CRLF2* mutations have not been reported in patients with *IGH-CRLF2* or *P2RY8-CRLF2*, therefore screening for these mutations was not carried out.

2.9 | Statistical analysis

The distribution of categorical variables was examined using Fisher's exact test. Due to the investigative nature of this analysis, we did not apply stringent multiple comparisons adjustment (all tests were conducted at the 5% significance level). All analyses were performed using Intercooled Stata 14.1 for Windows (Stata Corporation, College Station, TX).

2.10 | Pathway analysis

Pathway mapping was performed on a defective gene list identified for each sample. A gene was considered to be defective if: (1) it contained

at least one nonsilent mutation, (2) was located at the break point junction of the identified structural variants, and (3) was located within a defective region (detected either from SNP6 or WGS) of no more than 1 MB in size. Gene lists from the defective regions were retrieved from Human reference genome hg19 via biomart²⁸ and defective gene lists were mapped to KEGG pathways²⁹ using R/Bioconductor package KEGGREST (Tenenbaum D. KEGGREST: Client-side REST access to KEGG. R package version 1.11.0). Pathway enrichment analysis was determined using hypergeometric test. The test is based on the probability of observing x number of genes from a given pathway as being defective, given a process of sampling without replacement of all protein-coding genes (20805) annotated in the human genome (GRCh37.p13, INSDC Assembly GCA_000001405.14, Feb 2009). This gene list is considered as a representative gene set that are mutated among CRLF2-r cases. The basic assumption for the analysis is that mutations among CRLF2-r cases arise randomly, so the enrichment test is used as a statistical method to test whether there is any selective pressure that makes mutations occur in particular pathways more often than expect by chance. Enrichment tests were performed with the defective gene lists from all samples as well as for each subtype separately.

3 | RESULTS

3.1 | The clinical and demographic features of *CRLF2*-r patients are characteristic of B-ALL

All patients with *CRLF2*-r were B-ALL (n=172): *IGH-CRLF2* (n=47) and *P2RY8-CRLF2* (n=125), median age 5 years (range 1–60 y), male predominance (59%), 70% had WBC $<50 \times 10^6$ /l (Supporting Information Table 4; Figure 1A). The age distribution was similar to B-ALL overall, with the majority <10 years (71%). 20 Median age (4 y vs. 14 y, P<.001), WBC (25% vs. 43% with WBC $>50 \times 10^6$ /l, P=.016) and percentage NCI high risk (38% vs. 77%, P<.001) differed significantly between *P2RY8-CRLF2* and *IGH-CRLF2*, respectively (Supporting Information Table 4; Figure 1A). We confirmed the high frequency of DS-ALL among *CRLF2*-r ALL (31%, 50/161) (DS status not available for 11 patients), with more DS-ALL patients harboring the *P2RY8-CRLF2* fusion (*P2RY8-CRLF2*, n=41 vs. *IGH-CRLF2* n=9, 35% vs. 20%, respectively, P=.087) (Supporting Information Table 4; Figure 1B).

Cytogenetic analysis was successful in 160 *CRLF2*-r patients. Normal karyotype was observed in 51 (36%) patients. The majority, including those with DS-ALL, were classified as B-other ALL (Figure 1B,C); however, primary, established chromosomal abnormalities co-occurred in 22% (Supporting Information Table 1): *ETV6-RUNX1* (n=3), *BCR-ABL1* (n=5), high hyperdiploidy (n=12), and iAMP21 (n=17). All but one of these patients had *P2RY8-CRLF2* (n=36) (Figure 1B). No patients had t(1;19)(q23;p13), *KMT2A* (*MLL*) rearrangements or *ABL*-class fusions involving *ABL1*, *ABL2*, *PDGFRB*, and *CSF1R*. 4,30,31

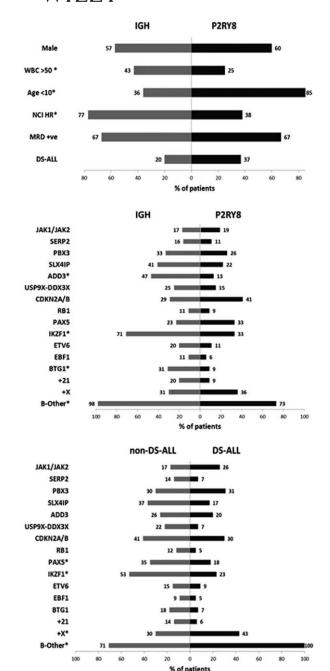


FIGURE 1 Comparative histograms of recurrent genetic abnormalities in patients with CRLF2-r ALL. (A) Comparison of clinical data between patients with IGH-CRLF2 (gray bars) and P2RY8-CRLF2 (black bars). (B) Comparison of genetic data between patients with IGH-CRLF2 (gray bars) and P2RY8-CRLF2 (black bars). (C) Comparison between CRLF2-r DS-ALL (black bars) and CRLF2-r non DS-ALL patients (gray bars). * denotes those abnormalities where the incidence between the two groups is statistically significant

3.2 | Additional chromosomal abnormalities in *CRLF2*-r patients

Recurrent somatic structural and numerical aberrations were present in 109 patients (Supporting Information Figure 2), including gains of chromosomes X (37/83, 45%), 21 (13/83, 16%), 17 (7/83, 8%), and 9 (4/83,

5%), in patients with and without DS. The gain of chromosome X was significantly enriched in *CRLF2*-r patients, particularly among DS-ALL, when compared to a cohort of B-other patients, where only 5 of 1019 (5%) patients had gain of X (Supporting Information Table 5; Figure 1C).

Copy number alterations (CNA) were detected among patients with CRLF2-r ALL by MLPA (n=154) and SNP6.0 arrays (n=26) (Table 1; Supporting Information Table 1 and 6-7; Figure 1B). Deletions of IKZF1, CDKN2A/B, PAX5, and BTG1 were present in 43% (n=60), 38% (n=52), 30% (n=41), and 14% (n=20) of cases, respectively. Deletions of IKZF1 and BTG1 occurred at higher incidences than seen in Bother ALL at 23% and 2%, respectively. Deletions of IKZF1 (71% vs. 33% P<.001) and BTG1 (31% vs. 9%, P=.004) were more frequent in IGH-CRLF2 than P2RY8-CRLF2 patients, respectively (Figure 1B). There was a lower incidence of PAX5 and IKZF1 deletions in patients with DS-associated CRLF2-r ALL compared to those without DS (18% vs. 35% P<.047 and 23% vs. 53%, P=.003, respectively) (Figure 1C, Supporting Information Table 5). In fact, CRLF2-r DS patients were more likely to have none of these deletions (48% vs. 15%, P<.001), with 89% (39/44) having fewer than three gene deletions.

Other recurrent CNA observed from SNP6.0 arrays included: the histone cluster at 6p22.2 (n=7), VPREB1 (n=6), ADD3 (n=5), BTLA (n=4), SLX4IP (n=3), SERP2 and TSC22D1 (n=3), and PBX3 (n=2). Publically available SNP6.0 data³² showed these deletions to be present in other ALL subtypes: ADD3, n=9 (4.69%), SLX4IP, n=9 (4.69%), BTLA, n=13 (6.77%). To verify their frequency in CRLF2-r ALL, we screened additional samples by FISH, as fixed cells were the most abundant source of patient material, that also allowed detection of low level populations. However, the resolution of FISH restricted accurate detection of deletions <40 kb. Combined FISH and SNP6.0 data generated incidences of; ADD3 (14/57, 25%), SLX4IP (13/44, 30%), SERP2 and TSC22D1 (7/55, 13%), and PBX3 (16/56, 29%), with no difference in the prevalence between DS and non-DS CRLF2-r ALL. IGH-CRLF2 patients had a higher incidence of ADD3 deletions (46% vs. 13%, IGH vs. P2RY8, P=.008) (Supporting Information Table 6; Figure 1B).

3.3 | A novel fusion between USP9X and DDX3X

An interstitial deletion of chromosome sub-band Xp11.4, fusing *USP9X* to *DDX3X* (Supporting Information Figure 3A), was observed in 2/26 cases by SNP6.0 arrays and 7/24 cases by FISH, including one DS patient, giving an incidence of 19% (Supporting Information Figure 3B). No cases were identified among the publically available SNP6.0 data. FISH revealed that the fusion was present in both minor and major clones (8-91% of nuclei) in both male (n = 4) and female patients (n = 5). All male patients had either an additional copy of normal chromosome X or derived X involved in the *CRLF2*-r. It was identified in patients with both *IGH* and *P2RY8* involvement (4 vs. 5, respectively). Real-time PCR confirmed that the fusion was in-frame and was expressed at the mRNA level (Supporting Information Figure 3C). Sanger sequencing verified fusion of *USP9X* exon 31 to exon 2 of *DDX3X* (Supporting Information Figure 3D). The DNA breakpoint sequence from three patients consistently fell within intron 31 of

TABLE 1 Focal aberrations identified by SNP, paired-end and exome sequencing

Patient ID	B-cell differentiation	Cell cycle	Kinase	RAS	Cell adhesion	TP53	Other interesting genes
P2RY8-CRLF2							
9534		BTG1, CDKN2AB	P2RY8-CRLF2, JAK2, IL7R	NRAS	LAMA1	CDKN2AB	NFATC4, NCOA3
11538	ETV6, IKZF1, PAX5	CDKN2AB, FOXN3	P2RY8-CRLF2, MAST4, STK38L, RPS6KA5	NRAS	ITGBL1, USH2A	CDKN2AB	CRIPAK, KDM4A, SKIL, CCDC88C, SMEK1, FBLN5, CYP4A11
11706	ETV6		P2RY8-CRLF2, CKMT1A, SPHKAP, NBEA, PTPRT		PTPRT		CREBBP, GLI1
20638		CDKN2AB	P2RY8-CRLF2, ACVRL1, IL7R		ITGA7, ANK2	CDKN2AB	MLLT3, TUSC1, TOPORS, DDX58, APTX
20753	IKZF1, PAX5	CDKN2A	P2RY8-CRLF2, MAPK10		USH2A	CDKN2A	PPP2R3B, ZEDB1, IL9R
21819	IKZF1, PAX5, VPREB1	CDKN2AB, MCC	P2RY8-CRLF2, INSRR	NF1		CDKN2AB	DDX6, CASZ1
21991	ETV6, VPREB1	CDKN2AB, CCND3	P2RY8-CRLF2		CD38, NPNT	CDKN2AB	ARID5A, SEMA4C, BTLA, TBL1XR1, FHIT, PAXIP1
IGH-CRLF2							
11543	EBF1, IKZF1	BTG1	IGH-CRLF2, JAK2, TTBK1				NOX4, CHORDC1
19599	IKZF1, VPREB1	CDK13	IGH-CRLF2, JAK2, CDK13		DST		TOP3A, CREB5, TRL4, CHN2, MPLKIP, INHBA, GLI3, TSC22D1
21245	EBF1, ETV6, IKZF1, PAX5, VPREB1	RB1, CDK6	IGH-CRLF2, JAK2, ERBB4		CTNNA2	RB1	ABRA, BTLA, AFF1, MEF2C, ADD3, RUNX1
21470	ETV6, IKZF1, VPREB1		IGH-CRLF2, CRLF2, PTPN11, RHBDL2	PTPN11	ITGB7, LAMC2, COL3A1, CTTNBP2		MIR181B1, MIR181A1, CADM1

 $\it USP9X$ (covering a 3.5 kb region) and intron 1 of $\it DDX3X$ (covering a 2.1 kb region).

3.4 | Somatically acquired structural variants are rare, while the incidence of kinase and *JAK* mutations is high in *CRLF2*-r ALL

The somatic nature of 137 structural variants (SV) (6 tandem duplication-, 104 deletion-, 14 inversion-, and 13 translocation-type rearrangements) were validated, providing an average of 12.7 SV per patient, irrespective of DS status (Figure 2; Supporting Information Table 8; Supporting Information Figure 4). No novel recurrent rearrangements or fusion genes typical of B-other ALL were identified in these 11 patients. There was no difference in the number of SV between patients with *IGH-* and *P2RY8-CRLF2* and also between patients with and without DS.

WES from the same 11 patients detected 218 mutations, 187 point mutations, and 31 in/del events. Among the point mutations, 122 were predicted as possibly disease causing by PolyPhen-2 and

Mutation Taster (Figure 3; Supporting Information Table 9; Supporting Information Figure 4). The average number of mutations was 19.8 (point mutations alone: 17) per patient, with no difference between *IGH*- and *P2RY8-CRLF2* (P=.85) or those with or without DS (P=.41). In the majority of patients, at least half of the mutations were at variant allele frequency (VAF) higher than 30%; only one patient (#11706) had a higher number of mutations at a VAF of <20%. Recurrent mutations were identified in *CACNA1D*, *IKZF1*, *JAK2*, *IL7RA*, *NRAS*, and *USH2A* (n=2 each). Ten patients had missense or nonsense point mutations in genes with protein kinase functions (Table 1), with six patients harboring either a mutation or CNA in two or more kinase genes.

Targeted screening for JAK2 (n=75) and JAK1 (n=36) mutations identified an incidence of 35% and 14%, respectively. There was no difference in the number of JAK1 or JAK2 mutations between IGH- and P2RY8-CRLF2 (17% vs. 19%, P=.46) or DS and non-DS patients (JAK2 Ex14, 44% vs. 35%, P=.613; JAK1 Ex14, 17% vs. 11%, P=1; Supporting Information Table 5; Figure 1B,C). Due to selective screening of only exon 14 in the majority of patients, these numbers may represent an underestimate of the true incidence of JAK1 and JAK2 mutations in CRLF2-r ALL.

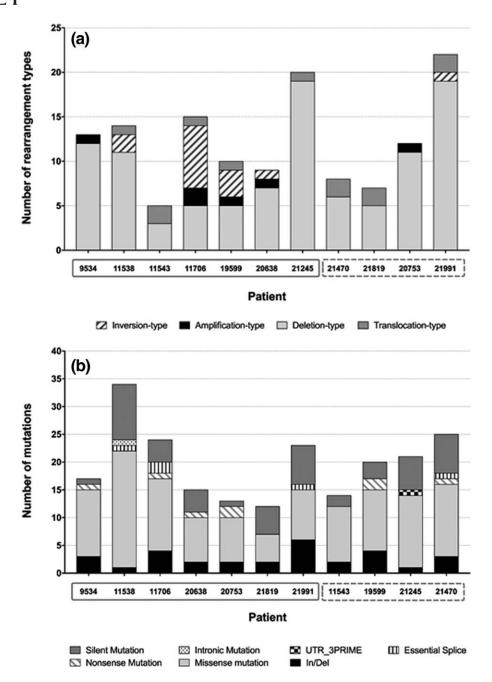


FIGURE 2 Acquired mutations and structural variations in *CRLF2*-r ALL. (A) Histogram representing the distribution of structural variants (SV) in 11 patients with *CRLF2*-r ALL (x-axis, solid gray box depicts patients with *P2RY8-CRLF2* and the dashed gray box, patients with *IGH-CRLF2*), detected by paired-end sequencing and validated by FISH, MLPA, SNP arrays and/or PCR and subsequent Sanger sequencing for breakpoint locations. Each patient is represented by a bar with the height of each bar being representative of the number of aberrations present in that patients sample. Different shades/patterns denote the different types of SV that were detected. (B) Histogram representing the distribution of coding mutations in same 11 patients with *CRLF2*-r ALL (x-axis, solid gray box depicts patients with *P2RY8-CRLF2* and the dashed gray box, *IGH-CRLF2* patients), detected by whole exome sequencing. Each patient is represented by a bar with the height of each bar being representative of the number of aberrations present in that patients sample. Different shades/patterns denote the different types of coding mutations that were detected

3.5 | Involvement of cell adhesion mutations in *CRLF2*-r patients

WES data identified \geq 1 somatic point mutation in a gene(s) involved in cell adhesion in eight patients including; NPNT, ITGB7, COL3A1, ITGA7,

ITGBL1, DST, LAMC2, PTPRT, MAPK10, CCND3, CTNNA2, CTNND2, and LAMA1 in single patients and USH2A in two patients (Table 1).

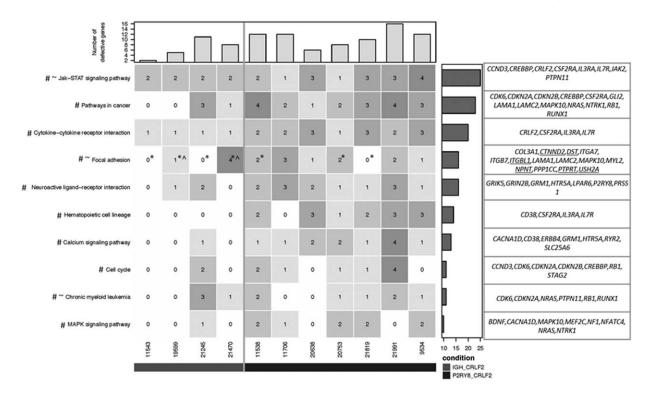


FIGURE 3 Heatmap showing the frequency distribution of defective genes. The top ten pathways in which defective genes are most frequently identified across 11 patients with *CRLF2*-r ALL are shown. The number in each cell represents numbers of defective genes identified for particular pathways and patients, ranging from 0 (white) to 4 (dark gray). Total numbers of defective genes for each patient are summarized in the bar chart on the top panel of the heatmap. Note that genes which are members of two or more pathways were counted only once. Bar chart on the right panel summarises the total number of defective genes identified across the 11 patients for a given pathway (KEGG pathways, with additional manually annotated genes underlined). # significantly enriched (hypergeometric test *P*-value < .05) in patients with *P2RY8-CRLF2*; ~ significantly enriched in patients with *IGH-CRLF2*; * CNA of *IKZF1*; ^ *IKZF1* point mutation

3.6 | Pathway analysis reveals deregulation in known and novel pathways

Pathway analysis revealed those pathways likely to be deregulated due to the presence of mutations (not including silent mutations), in/dels and focal SV <1 Mb in size. The top ten pathways in which defective genes were most frequently identified are depicted in Figure 3. Unsurprisingly, *JAK-STAT* involvement was seen in all patients. Other pathways included those involved in cancer, cytokine-cytokine interaction, haematopoietic cell lineage and MAPK signaling. Interestingly, the focal adhesion pathway was also recurrently involved.

4 | DISCUSSION

In this study investigating the genomic landscape of *CRLF2*-r ALL, we have confirmed the high incidence of *CRLF2*-r in DS-ALL, demonstrated its co-existence with other primary chromosomal rearrangements and enrichment of specific chromosomal gains and deletions of *IKZF1*, *BTG1*, *ADD3*, *SERP2*, *TSC22D1*, *SLX4IP*, and *PBX3*. There were significant differences in CNA profiles between *P2RY8-CRLF2* and *IGH-CRLF2* patients with an increased incidence of *IKZF1*, *BTG1*, and *ADD3* deletions and a higher age at diagnosis being observed in the latter. These disparities correlate with reported differences in outcome, where for example older age and *IKZF1* status drive the inferior prognosis

observed in adult IGH-CRLF2.14 A fifth of CRLF2-r patients harbor a primary chromosomal abnormality with all but one being P2RY8-CRLF2, suggesting a co-operating role for CRLF2 deregulation. Interestingly, the sole patient with IGH-CRLF2 and a primary abnormality had the BCR-ABL1 fusion in a separate clone.²² Collectively, these data do not imply that IGH and P2RY8 are driving distinct subgroups but rather that CRLF2-d may play a dual role as both a primary and co-operating driver in ALL, with the latter being more prominent among P2RY8-CRLF2 patients. This model is akin the BCR-ABL1 fusion, which although predominatly a primary driver, has also been reported as a secondary abnormality^{22,33}. Independent pathway analysis of the aberration profile of IGH-CRLF2 and P2RY8-CRLF2 patients found enrichment of all top ten pathways in the P2RY8-CRLF2 patients and only a subset of pathways in IGH-CRLF2 patients. Previous studies have reported the involvement of MAPK signaling in CRLF2-r ALL.34 This pathway was not significantly enriched in our IGH-CRLF2 patients; however, this observation was based on only four patients and thus requires further validation. WES identified a frequency of 20 mutations per patient, similar to iAMP21-ALL,35 but higher than other subgroups, ^{23,36-38}. The high VAF of kinase mutations suggested that they were clonal and were likely acquired early in disease development. Mutations in ERBB4, TTBK1, and STK38L occurred within the catalytic domain with potential for constitutive activation. These mutations may activate STAT5 (ERBB4), deregulate alternate pathways including, NFKB and chromosome alignment pathways (*TTBK1*), or regulate the Hippo signaling pathway (*STK38L*) by binding *MOB* kinase activators, which have tumour suppressor roles. Taken together, these data imply that *CRLF2*-r may define a distinct subgroup of B-other ALL.

For the first time, we have implicated a role for *USP9X* and *DDX3X* in *CRLF2*-r ALL. This *USP9X-DDX3X* fusion removes the ubiquitin carboxyl-terminal hydrolase domain of *USP9X*, the promoter of *DDX3X*, multiple regulatory elements, MIR7641-2 and non-coding RNAs. The involvement of *USP9X* and *DDX3X* in cancer is known³⁹⁻⁴⁵ and another translocation involving *USP9X* has been reported in a *BCR-ABL1*-like ALL patient without *CRLF2*-r.⁴ In keeping with the emerging functions of *USP9X*, its overexpression has been reported in B-ALL, suggesting an oncogenic role.⁴⁶ Knockdown of USP9X sensitises both prednisolone sensitive and resistant cell lines to glucocorticoid (GC)-induced apoptosis, suggesting that reduced levels of USP9X may sensitise them to prednisolone treatment.⁴⁶

With the success of targeted approaches using tyrosine kinase inhibitors (TKI) in BCR-ABL1-positive disease, there is clear rational to apply similar targeted approaches to other ALL subtypes. CRLF2-r activate targetable pathways, including JAK-STAT, PI3K, and MAPK signaling, and many groups are now assessing the efficacy of inhibitors to these pathways. The first inhibitor to be tested in CRLF2-r ALL was the JAK1/2 inhibitor, Ruxolitinb.⁴⁷ While this study showed a response in vivo, greater effect was observed in JAK-activated, non-CRLF2-r ALL. Evidence of resistance to type I inhibitors in model systems of B-ALL is already driving the development of type II JAK inhibitors. 48 Targeting of mulitple pathways was an approach taken by Suryani et al, (2015) in assessing the JAK inhibitor, AZD1480, alone and in combination with the MEK inhibitor, Selumetinib.⁴⁹ While this study showed a strong anti-leukaemic effect in vitro, only modest effects were seen in vivo, an important consideration for future preclinical testing, which highlights the need to identify other common targets to provide further options for more patients. In addition to JAK-STAT, we have identified several deregulated pathways in CRLF2-r ALL, for example, focal adhesion signaling, one which has not previously been implicated in this subtype of ALL and was enriched in patients with both IGH- and P2RY8-CRLF2 ALL.

Mutations within genes controlling cell adhesion would seem to be more relevant in solid tumours, where such lesions would provide cells with a migratory and metastatic advantage. However, the presence of such mutations within the primary site may be advantageous, where small amounts of local movement can alter tumour growth and dispersal. In our study, eight patients (72%) had one or more focal adhesion gene mutations; an incidence much higher than reported in other ALL subtypes (ETV6-RUNX1, 12%²³, Ph-like non-CRLF2-r, 15%⁴, low hypodiploidy, 5%³⁷, high hyperdiploidy, 9.8%³⁸, KTM2A-rearranged infant ALL, 4.5%³⁶). Mutated genes were specifically located within the laminin-integrin-MAPK cell signaling axis. The laminins (LAMA1 and LAMC2) and their receptor integrins (ITGA7, ITIBL1, and ITGB7) play a central role in cell proliferation and tumour cell invasiveness through their critical role in basement membrane adhesion.⁵⁰ Although the functional consequence of these mutations in CRLF2-r ALL remains to

be determined, when laminin binding integrin signaling is altered, it activates downstream signaling kinases, such as FAK (focal adhesion kinase) and JUN (MAPK10)-FOS, which impacts on cell proliferation and migration in various cancers. ^{51,52} In a recent study, the presence of the Ikaros isoform 6 (IK6), *IKZF1* haploinsuficiency or mutations in *BCR-ABL1*-positive disease have been shown to increase expression of adhesion molecules and increased transcription of *FAK*, rendering these cells insensitive to tyrosine kinase inhibition. ⁵³ The application of inhibitors to FAK in patients with *BCR-ABL1*-positive ALL resulted in the same abrogation of adhesion and self renewal programmes. ⁵⁴ Interestingly, all 11 patients with sequencing data either showed mutation/ CNA of *IKZF1* or a mutation within a focal adhesion gene. Our data along with emerging literature in leukaemia highlight a role for the focal adhesion pathway in *CRLF2*-r ALL.

In summary, we show clear clinical and genomic differences between patients with IGH- and P2RY8-CRLF2. We describe these rearrangements in the presence of other established cytogenetic abnormalities, suggesting a secondary role for CRLF2-r in some patients, akin to BCR-ABL1 positive disease. However, due to low patient numbers, we were unable to ascertain any impact of P2RY8-CRLF2 on the prognosis of these patients. This co-occurrence needs to be further assessed in a larger patient cohort. The treatment of patients across four independent trails precluded meaningful survival analysis between patients IGHand P2RY8-CRLF2. It is clear from our data that CRLF2-r ALL is heterogeneous, requiring a combination of genetic abnormalities in functionally relevant genes to co-operate with deregulated expression of CRLF2. Although the functional relevance of some of the deletions and mutations presented in this descriptive study are currently unknown, there are pointers to activation of additional targetable pathways. There is a clear requirement for studies addressing the biological effect of these aberrations, which together with the identification of mutations in cell adhesion genes and a high incidence of IKZF1 deletions provide interesting targets for pre-clinical testing. Inhibitors to the focal adhesion pathway, as one example, may provide an insight into a new realistic therapeutic approach to improve outcome in CRLF2-r ALL.

ACKNOWLEDGMENTS

The authors would like to thank member laboratories of the United Kingdom Cancer Cytogenetic Group (UKCCG) for providing cytogenetic data and material. We are grateful to all the members of the NCRI Haematological Oncology Adult ALL Subgroup and the NCRI Childhood Cancer and Leukaemia Group (CCLG) Leukaemia Subgroup. Primary childhood leukaemia samples used in this study were provided by the Bloodwise Childhood Leukaemia Cell Bank working with the laboratory teams in the Bristol Genetics Laboratory, Southmead Hospital, Bristol: Molecular Biology Laboratory, Royal Hospital for Sick Children, Glasgow: Molecular Haematology Laboratory, Royal London Hospital, London: Molecular Genetics Service and Sheffield Children's Hospital, Sheffield. We also thank the Central England Haemato-Oncology Research Biobank and the Newcastle Haematology Biobank for providing material used in this study. Lisa J Russell has a John Goldman Fellowship from Leuka and Elli

Papaemmanuil has an EHA fellowship. Finally, we thank all the clinicians who entered patients into the trial and the patients and families who agreed to take part. Support by: The Kay Kendall Leukaemia Fund, Leuka, European Haematology Association and Bloodwise (formerly Leukaemia and Lymphoma Research)

REFERENCES

- [1] Moorman AV. The clinical relevance of chromosomal and genomic abnormalities in B-cell precursor acute lymphoblastic leukaemia. *Blood Rev.* 2012;26:123–135.
- [2] Den Boer ML, van Slegtenhorst M, De Menezes RX, et al. A subtype of childhood acute lymphoblastic leukaemia with poor treatment outcome: a genome-wide classification study. *Lancet Oncol.* 2009:10:125–134.
- [3] Mullighan CG, Su X, Zhang J, et al. Deletion of IKZF1 and prognosis in acute lymphoblastic leukemia. N Engl J Med. 2009;360:470–480.
- [4] Roberts KG, Li Y, Payne-Turner D, et al. Targetable kinaseactivating lesions in Ph-like acute lymphoblastic leukemia. N Engl J Med. 2014;371:1005–1015.
- [5] Russell LJ, Capasso M, Vater I, et al. Deregulated expression of cytokine receptor gene, CRLF2, is involved in lymphoid transformation in B-cell precursor acute lymphoblastic leukemia. *Blood.* 2009; 114:2688–2698.
- [6] Russell LJ, Enshaei A, Jones L, et al. IGH@ translocations are prevalent in teenagers and young adults with acute lymphoblastic leukemia and are associated with a poor outcome. J Clin Oncol. 2014;32: 1453–1462.
- [7] Mullighan CG, Collins-Underwood JR, Phillips LA, et al. Rearrangement of CRLF2 in B-progenitor- and Down syndrome-associated acute lymphoblastic leukemia. Nat Genet. 2009;41:1243–1246.
- [8] Yano M, Imamura T, Asai D, et al. An overall characterization of pediatric acute lymphoblastic leukemia with CRLF2 overexpression. *Genes Chromosomes Cancer.* 2014;53:815–823.
- [9] Chen I-M, Harvey RC, Mullighan CG, et al. Outcome modeling with CRLF2, IKZF1, JAK, and minimal residual disease in pediatric acute lymphoblastic leukemia: a Children's Oncology Group Study. *Blood*. 2012;119:3512–3522.
- [10] Harvey RC, Mullighan CG, Chen I-M, et al. Rearrangement of CRLF2 is associated with mutation of JAK kinases, alteration of IKZF1, Hispanic/Latino ethnicity, and a poor outcome in pediatric B-progenitor acute lymphoblastic leukemia. *Blood*. 2010;115:5312– 5321.
- [11] Haslam K, Kelly J, Morris T, et al. Screening for CRLF2 overexpression in adult acute lymphoblastic leukemia. Int J Lab Hematol. 2011; 33:e17-e19.
- [12] Morak M, Attarbaschi A, Fischer S, et al. Small sizes and indolent evolutionary dynamics challenge the potential role of P2RY8-CRLF2-harboring clones as main relapse-driving force in childhood ALL. Blood. 2012;120:5134–5142.
- [13] Buitenkamp TD, Pieters R, Gallimore NE, et al. Outcome in children with Down's syndrome and acute lymphoblastic leukemia: role of IKZF1 deletions and CRLF2 aberrations. *Leukemia*. 2012;26:2204– 2211.
- [14] Moorman AV, Schwab C, Ensor HM, et al. IGH@ translocations, CRLF2 deregulation, and microdeletions in adolescents and adults with acute lymphoblastic leukemia. J Clin Oncol. 2012;30:3100– 3108.
- [15] Schwab CJ, Jones LR, Morrison H, et al. Evaluation of multiplex ligation-dependent probe amplification as a method for the detec-

- tion of copy number abnormalities in B-cell precursor acute lymphoblastic leukemia. *Genes Chromosomes Cancer*. 2010;49:1104– 1113
- [16] Shochat C, Tal N, Bandapalli OR, et al. Gain-of-function mutations in interleukin-7 receptor- α (IL7R) in childhood acute lymphoblastic leukemias. *J Exp Med.* 2011;208:901–908.
- [17] Roberts Kathryn G, Morin Ryan D, Zhang J, et al. Genetic alterations activating kinase and cytokine receptor signaling in high-risk acute lymphoblastic leukemia. *Cancer Cell*. 2012;22:153–166.
- [18] Roberts KG, Pei D, Campana D, et al. Outcomes of children with BCR-ABL1-like acute lymphoblastic leukemia treated with riskdirected therapy based on the levels of minimal residual disease. J Clin Oncol. 2014;32:3012–3020.
- [19] Ensor HM, Schwab C, Russell LJ, et al. Demographic, clinical, and outcome features of children with acute lymphoblastic leukemia and CRLF2 deregulation: results from the MRC ALL97 clinical trial. *Blood*. 2011;117:2129–2136.
- [20] Moorman AV, Enshaei A, Schwab C, et al. A novel integrated cytogenetic and genomic classification refines risk stratification in pediatric acute lymphoblastic leukemia. *Blood*. 2014;124:1434–1444.
- [21] 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; 98:1081–1088.
- [22] Jeffries SJ, Jones L, Harrison CJ, Russell LJ. IGH@ translocations co-exist with other primary rearrangements in B-cell precursor acute lymphoblastic leukemia. *Haematologica*. 2014;99:1334–1342.
- [23] Papaemmanuil E, Rapado I, Li Y, et al. RAG-mediated recombination is the predominant driver of oncogenic rearrangement in ETV6-RUNX1 acute lymphoblastic leukemia. Nat Genet. 2014;46:116–125.
- [24] Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2010;26:589–595.
- [25] Nik-Zainal S, Alexandrov LB, Wedge DC, et al. Mutational processes molding the genomes of 21 breast cancers. Cell. 2012;149:979–993.
- [26] Raine KM, Hinton J, Butler AP, et al. cgpPindel: identifying somatically acquired insertion and deletion events from paired end sequencing. Curr Protoc Bioinformatics. 2015;52:11–12.
- [27] Kearney L, Gonzalez De Castro D, Yeung J, et al. Specific JAK2 mutation (JAK2R683) and multiple gene deletions in Down syndrome acute lymphoblastic leukemia. *Blood.* 2009;113:646–648.
- [28] Durinck S, Spellman PT, Birney E, Huber W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat Protoc.* 2009:4:1184–1191.
- [29] Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 2000;28:27–30.
- [30] Schwab C, Ryan SL, Chilton L, et al. EBF1-PDGFRB fusion in pediatric B-cell precursor acute lymphoblastic leukemia (BCP-ALL): genetic profile and clinical implications. *Blood*. 2016;127:2214–2218.
- [31] O'Connor D, Moorman AV, Wade R, Hancock J, Tan RM, Bartram J, Moppett J, Schwab C, Patrick K, Harrison CJ, Hough R, Goulden N, Vora A, Samarasinghe S. Use of minimal residual disease assessment to redefine induction failure in pediatric acute lymphoblastic leukemia. J Clin Oncol:Jco2016696278. 2017.
- [32] Mullighan CG, Goorha S, Radtke I, et al. Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature*. 2007; 446:758–764.
- [33] Miura I, Takatsu H, Yamaguchi A, et al. Standard Ph chromosome, t (9;22)(q34;q11), as an additional change in a patient with acute myelomonocytic leukemia (M4Eo) associated with inv(16)(p13q22). Am J Hematol. 1994;45:94–96.

- [34] Tasian SK, Doral MY, Borowitz MJ, et al. Aberrant STAT5 and Pl3K/ mTOR pathway signaling occurs in human CRLF2-rearranged B-precursor acute lymphoblastic leukemia. *Blood*. 2012;120:833–842.
- [35] Ryan SL, Matheson E, Grossmann V, Sinclair P, Bashton M, Schwab C, Towers W, Partington M, Elliott A, Minto L, Richardson S, Rahman T, Keavney B, Skinner R, Bown N, Haferlach T, Vandenberghe P, Haferlach C, Santibanez-Koref M, Moorman AV, Kohlmann A, Irving JA, Harrison CJ. The role of the RAS pathway in iAMP21-ALL. Leukemia. 2016;30:1824-1831.
- [36] Andersson AK, Ma J, Wang J, et al. The landscape of somatic mutations in infant MLL-rearranged acute lymphoblastic leukemias. Nat Genet. 2015;47:330–337.
- [37] Holmfeldt L, Wei L, Diaz-Flores E, et al. The genomic landscape of hypodiploid acute lymphoblastic leukemia. *Nat Genet*. 2013;45:242–252.
- [38] Paulsson K, Lilljebjorn H, Biloglav A, et al. The genomic landscape of high hyperdiploid childhood acute lymphoblastic leukemia. *Nat Genet*. 2015;47:672–676.
- [39] Brandimarte L, La Starza R, Gianfelici V, et al. DDX3X-MLLT10 fusion in adults with NOTCH1 positive T-cell acute lymphoblastic leukemia. *Haematologica*. 2014;99:64–66.
- [40] Cox JL, Wilder PJ, Wuebben EL, Ouellette MM, Hollingsworth MA, Rizzino A. Context-dependent function of the deubiquitinating enzyme USP9X in pancreatic ductal adenocarcinoma. *Cancer Biol Ther.* 2014;15:1042–1052.
- [41] Ojha J, Secreto CR, Rabe KG, et al. Identification of recurrent truncated DDX3X mutations in chronic lymphocytic leukaemia. Br J Haematol. 2015;169:445–448.
- [42] Oosterkamp H, Hijmans E, Brummelkamp T, et al. USP9X downregulation renders breast cancer cells resistant to tamoxifen. Cancer Res. 2014;74:3810–3820.
- [43] Robinson G, Parker M, Kranenburg TA, et al. Novel mutations target distinct subgroups of medulloblastoma. *Nature*. 2012;488:43–48.
- [44] Schmitz R, Young RM, Ceribelli M, et al. Burkitt lymphoma pathogenesis and therapeutic targets from structural and functional genomics. *Nature*. 2012;490:116–120.
- [45] Yan J, Zhong N, Liu G, et al. Usp9x- and Noxa-mediated Mcl-1 downregulation contributes to pemetrexed-induced apoptosis in human non-small-cell lung cancer cells. *Cell Death Dis*. 2014;5:e1316.
- [46] Zhou M, Wang T, Lai H, et al. Targeting of the deubiquitinase USP9X attenuates B-cell acute lymphoblastic leukemia cell survival and overcomes glucocorticoid resistance. Biochem Biophys Res Commun. 2015;459:333–339.

- [47] Maude SL, Tasian SK, Vincent T, et al. Targeting JAK1/2 and mTOR in murine xenograft models of Ph-like acute lymphoblastic leukemia. *Blood*. 2012;120:3510–3518.
- [48] Wu SC, Li LS, Kopp N, et al. Activity of the type II JAK2 inhibitor CHZ868 in B cell acute lymphoblastic leukemia. Cancer Cell. 2015; 28:29-41.
- [49] Suryani S, Bracken LS, Harvey RC, Sia KC, Carol H, Chen IM, Evans K, Dietrich PA, Roberts KG, Kurmasheva RT, Billups CA, Mullighan CG, Willman CL, Loh ML, Hunger SP, Houghton PJ, Smith MA, Lock RB. Evaluation of the in vitro and in vivo efficacy of the JAK inhibitor AZD1480 against JAK-mutated acute lymphoblastic leukemia. Mol Cancer Ther 2015;14:364–374.
- [50] Belkin AM, Stepp MA. Integrins as receptors for laminins. Microsc Res Tech. 2000;51:280–301.
- [51] Kim NG, Gumbiner BM. Adhesion to fibronectin regulates Hippo signaling via the FAK-Src-PI3K pathway. J Cell Biol. 2015;210:503– 515.
- [52] Tafolla E, Wang S, Wong B, Leong J, Kapila YL. JNK1 and JNK2 oppositely regulate p53 in signaling linked to apoptosis triggered by an altered fibronectin matrix: JNK links FAK and p53. J Biol Chem. 2005;280:19992–19999.
- [53] Churchman Michelle L, Low J, Qu C, et al. Efficacy of Retinoids in IKZF1-Mutated BCR-ABL1 Acute Lymphoblastic Leukemia. Cancer Cell. 2015;28:343–356.
- [54] Churchman ML, Jones L, Evans K, et al. Efficacy of Focal Adhesion Kinase Inhibition in Combination with Dasatinib in BCR-ABL1 Acute Lymphoblastic Leukemia. *Blood.* 2015;126:3766–3766.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

How to cite this article: Russell LJ, Jones L, Enshaei A, et al. Characterisation of the genomic landscape of *CRLF2*-rearranged acute lymphoblastic leukemia. *Genes Chromosomes Cancer*. 2017;56:363–372. doi:10.1002/gcc.22439.