Supplementary Methods

Sequencing Method

An optimised and standardised protocol was followed by all testing centres. All samples underwent a buffer exchange using a Zymo Research DNA & Concentrator kit (Zymo Research, Tustin, Ca, USA). Whole genome shotgun library preparation was performed from 100 ng HMW or 200 ng FFPE gDNA using the KAPA Hyperplus kit (Roche Sequencing Solutions, Pleasanton, CA, US). Points in the protocol which were optimised included, fragmentation conditions of 37°C for 20 min, inclusion of a dual size selection step to enable retention of library molecules between 250-450 bp and 6 cycles of pre-hybridization PCR cycles. Including the negative and positive controls, 16 and 24 libraries were pooled for the proficiency and validation runs, respectively, followed by hybrid capture selected using the custom designed SeqCap EZ Choice Library (Roche Sequencing Solutions, Pleasanton, CA, US). For hybridization a total of 1 µg of pooled libraries were utilised which equated to 71.4 ng and 45.5 ng for each proficiency and validation sample library respectively. To remove bias during hybridisation arising from variability in DNA integrity, HMW and FFPE samples were pooled and hybridised on separate runs. Eleven cycles of post-hybridization PCR were performed prior to final enriched pooled libraries being sequenced on the NextSeq 500 (Illumina, Cambridge, UK) using a 150bp paired-end strategy performed on a 300-cycle NextSeq 500/550 Mid Output Kit (Illumina, Cambridge, UK). An in-silico analysis was performed of different read lengths (75 and 100bp pair-end reads) and the reporting of rearrangements using the EuroClonality-NDC assay was found to be remarkably stable throughout the cohort of 270 samples (manuscript in preparation).

ARResT/Interrogate pipeline

For both rearrangements and Structural Variant (SV) detection reads were paired-end joined using flash2, and for those without overlapping sequence, we concatenated the R1 and R2 reads into one sequence using a 12-N spacer[1]. No quality trimming was performed to be able to evaluate duplicated fragments downstream. Identical reads were deduplicated irrespective of orientation.

Bioinformatic Pipeline - Rearrangement Detection

Reads were aligned to IG/TCR germline genes and V gene recombination signal sequences (RSS) with minimap2[2]. A subset of reads that aligned to the IG/TCR germline genes and not RSS motifs were extracted using using seqtk subseq and retained for further analysis. Firstly, reads with the same junction or informative junction fragments were grouped and deduplicated based on strict clustering using the vsearch tool with a 98% cut off[3]. Clustered reads were then used to create contigs using the kmer based assembler, tadpole, from the BBMap program to maximise the quality and length of the

available sequence information. Secondly, the contigs were then used for a second round of finding rearrangements to finalise and improve the immunogenetic annotation. Finally, the rearrangement contig sequences were aligned to the genome [bwa mem, samtools] to detect false positive events (e.g. if the complete contig aligned to a contiguous genomic region, even with alignment breaks or other artefacts) or in fact confirm them (if the annotations of the aligned genomic coordinates matched the rearranged genes). A simplified version of the rearrangement results which formed part of the integrated report used for the study is shown below.

sample	in FASTQs[a]	Unique[b]	Usable[c]	locus	Class[d]	Event[e]	Frag[f]	% in locus[g]	% in class[h]
Ex_1	7544400	170529	1502	TRB	R VJ:Vb-(Db)-Jb	V4-2 J2-7 CASSQDGVNEQYF	212	31.6	32.1
Ex_1	7544400	170529	1502	TRB	R VJ:Vb-(Db)-Jb	V30 J2-7 CAWSAYREAL#F	182	27.2	27.5
Ex_1	7544400	170529	1502	TRB	R VJ:Vb-(Db)-Jb	V20-1 J1-1 CSARWVNTEAFF	165	24.6	25
Ex_1	7544400	170529	1502	TRG	R VJ:Vg-Jg	V2 J2 CATWDYKKLF	225	57	57
Ex_1	7544400	170529	1502	TRG	R VJ:Vg-Jg	V3 J2 CATWDRPGG#F	113	28.6	28.6

[a] Number of reads within the fastq file
[b] Number of IG/TCR reads following deduplication
[c] Number of reads which contained a rearrangement junction and were classed as usable
[d] Description of the class of the identified rearrangement
[e] Detail of the gene involved in the rearrangement and the subsequent CDR3 sequence
[f] Number of fragments which had the same number
[g] Percentage in locus was calculated by determining the number of fragments for the specific rearrangement detailed and dividing it by the total number of fragments for rearrangements identified for that specific loci
[h] Percentage in class was calculated by determining the number of fragments for the specific rearrangement and dividing it by the total number of fragments for the specific class of rearrangement within that loci.

Bioinformatic Pipeline - Translocation Detection

We created target sequences from the panel's BED file and the human genome, padded by 150bp in both directions [bedtools] to avoid missed alignments[4]. Reads were aligned to these target sequences [minimap2, samtools], and alignments kept if mapQ > 0[2,5]. Alignments needed to be soft-clipped (and the soft-clip of a minimum length of 30bp) and span two IG/TCR gene segments, with a minimum identity of 98% (or 90% in case of potential IG/TCR rearrangements). All potential breakpoints were calculated. Similar to rearrangements, reads with the same (or close-neighbouring) breakpoint were

grouped and deduplicated based on strict clustering [vsearch 98%], then used to create contigs [bbmap's tadpole] to maximise the quality and length of the available sequence information. The translocation contig sequences were aligned to the genome [minimap2, samtools] and the process described above for breakpoint detection is repeated. This allowed the identification of all translocation partners. Again, similar to rearrangements, a function of this alignment to the genome was to detect false positive events. The breakpoint coordinates were also used to identify reads supporting germline configurations of the partner loci, towards an approximation of event coverage. Often, a detected translocation was in fact a rearrangement, in which case the event was annotated as such and tagged as SV-R. A simplified version of the translocation results which formed part of the integrated report used for the study is shown below.

sample	in FASTQs[a]	unique[b]	usable[c]	locus	class[d]	event[e]	Frag[f]	% in locus[g]	% in class[h]
Ex_2	7519784	113011	1399	misc	SV transloc	IGHJ6 / CCND1 14:105,863,241- / 11:69,623,982-	219	99.9	99.9
Ex_2	7519784	113011	1399	misc	SV transloc	CCND1 / IGHD3-9 11:69,623,989- / 14:105,904,683-	116	99.9	99.9

[a] Number of reads within the fastq file
[b] Number of IG/TCR reads following deduplication
[c] Number of reads which contained a rearrangement junction was a rearrangement parameter
[d] Description of the structural variant, which for translocations were labelled 'SV transloc'
[e] Detail of the genes involved in the translocations and chromosomal coordinates of the breakpoint
[f] Number of fragments which had the same number
[g] Percentage in locus was a rearrangement parameter and was set as an arbitrary figure of 99.9 for translocations
[h] Percentage in class was a rearrangement parameter and was set as an arbitrary figure of 99.9 for translocations

Bioinformatic Pipeline - Copy Number Variation Detection

FASTQ files were aligned to the human reference genome (GRCh38/hg38) using Burrow-Wheeler Aligner (bwa, v0.7.17) following which Samtools (v4.0.12.0) was used for sorting, merging and filtering the bam files generated by bwa[5,6]. GATK4's (v 4.0.12) picard was used to sort, mark and remove the duplicated reads[7]. The aligned bam files for all the samples were then normalized using TSS (total sum scaling) using Bioconductor package CNVPanelizer[8]. The normalized samples with a coverage of at least 250X were used as group of reference samples against all the tumour samples. An average value of all the normal samples for each region in bed file was then calculated. Thereafter, the copy number variation is calculated as a ratio between individual tumour sample and mean value of normal samples for each region included in the bed file. Samples with a high level of aneuploidy (defined by $\geq$ 40 regions showing amplification or deletion across all regions on the bed file) were excluded (8 out of 280 samples [2.8%]) from analysis. CNVPanelizer used a threshold of 0.8 and 1.2 to classify a region

as either exhibiting either loss or gain of DNA copy number respectively. Analysis of IG/TCR rearrangements using CNA data was performed as above with verification of results using manual curation.

Bioinformatic Pipeline - Somatic Mutation Detection

All sequencing reads were aligned to human reference sequence (GRCh38), using the BWA (v0.7.15) MEM algorithm to generate SAM files[6]. SAMtools (v1.9) was used to convert the SAM files to BAM files, to remove reads with low mapping quality (MQ < 30) and to merge files from the same cell[5]. Picard tools (v2.6.0)[3] was used to sort the BAM files by chromosome coordinates, to mark/remove PCR duplicates, and to calculate the sequencing metrics. Variant calls were obtained using Vardict (v1.6) with tumour sample only mode[9]. Two sets of normal sample datasets were generated using the DNA materials from HMW and FFPE tissue types which were termed the pool of normal (PON) samples. To generate the initial high-confidence set of variant calls, any variants were removed with any of following conditions: (i) If the variant allele frequency was below 4%, (ii) were not within the coding sequence or splice site, (iii) the median variant allele frequency (VAF) is ≥ 20% in the HMW or FFPE PON samples at the same chromosomal coordinate, (iv) if ≥ 4 samples of the HMW PON samples or ≥ 2 samples of the FFPE PON samples had a VAF > 0 at the same chromosomal coordinate. The R Bioconductor package, Maftools, was used to create oncoplots detailing the most frequently mutated genes by disease entity in the validation cohort[10]. Using the Horizon cell line blends, with VAFs ranging from 4-30%, which were sequenced in triplicate we determined the LOD for SNV and indel calling to be 4% using our previously published methodology[11].

Supplementary References

1. Magoc T, Salzberg SL. FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics*. 2011;27(21):2957-2963.
2. Li H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*. 2018;34(18):3094-3100.
3. Rognes T, Flouri T, Nichols B, Quince C, Mahe F. VSEARCH: a versatile open source tool for metagenomics. *PeerJ*. 2016;4:e2584.
4. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*. 2010;26(6):841-842.
5. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25(16):2078-2079.
6. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25(14):1754-1760.
7. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010;20(9):1297-1303.
8. Oliveira C, Wolf T. CNVPanelizer: Reliable CNV detection in targeted sequencing applications. Vol. R package version 1.18.0.

https://bioconductor.org/packages/release/bioc/vignettes/CNVPanelizer/inst/doc/CNVPanelizer.pdf: Bioconductor; 2019.
9. Lai Z, Markovets A, Ahdesmaki M, et al. VarDict: a novel and versatile variant caller for next-generation sequencing in cancer research. *Nucleic Acids Res*. 2016;44(11):e108.
10. Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res*. 2018;28(11):1747-1756.
11. Izquierdo E, Yuan L, George S, et al. Development of a targeted sequencing approach to identify prognostic, predictive and diagnostic markers in paediatric solid tumours. *Oncotarget*. 2017;8(67):112036-112050.

Supplementary Tables

Table S1 Description of reactive lesions

Sample ID	Diagnosis
GMS1795	Chronic tonsillitis
GMS1796	Chronic tonsillitis
GMS1797	Chronic tonsillitis
GMS1798	Chronic tonsillitis
GMS1799	Reactive lymph node, NOS
GMS1801	Chronic tonsillitis
GMS1802	Chronic tonsillitis
GMS1803	Chronic tonsillitis
GMS1804	Chronic tonsillitis
GMS1805	Chronic tonsillitis
GMS1806	Chronic tonsillitis
GMS1807	Chronic tonsillitis
GMS1808	Chronic tonsillitis
GMS1809	Chronic tonsillitis
GMS1810	Chronic tonsillitis
GMS1811	Chronic tonsillitis
GMS1812	Chronic tonsillitis
GMS1813	Reactive lymph node, NOS
GMS1814	Chronic tonsillitis
GMS1815	Chronic tonsillitis
GMS1816	Chronic tonsillitis

Table S2 Description of cell lines used in the study

Cell Line	B / T cell	Disease Subtype	ATCC	DSMZ	ExPASy
ARR	T	T-ALL	No	No	CVCL_WV44
CA 46	B	BL	CRL-1648	ACC 73	CVCL_1101
CML-T1	T	T-ALL	No	ACC 7	CVCL_1126
H929	B	MM	CRL-9068	ACC 163	CVCL_1600
HAL-01	B	B-ALL	No	ACC 610	CVCL_1242
KCA	B	B-ALL	No	No	No
Loucy	T	T-ALL	CRL-2629	ACC 394	CVCL_1380

P12-Ichikawa	T	T-ALL	No	ACC 34	CVCL_1630
K-T1	T	T-ALL	No	No	CVCL_F690
KARPAS299	T	ALCL	No	No	CVCL_1324
Namalwa	B	BL	CRL-1432	ACC 24	CVCL_0067
REH	B	B-ALL	CRL-8286	ACC 22	CVCL_1650
RS4	B	B-ALL	CRL-1873	ACC 508	CVCL_0093
TOM-1	B	B-ALL	No	ACC 578	CVCL_1895

Table S3 Description of samples with Southern blot data used in study

Sample ID	Disease Subtype
DE-9	B-ALL
DE-6	CLL
DE-7	CLL
ES-1	CLL
ES-2	CLL
FR-3	CLL
GBN-10	CLL
GBS-2	CLL
GBS-3	CLL
NL-10	CLL
NL-11	CLL
ES-5	DLBCL
FR-6	DLBCL
GBN-4	DLBCL
GBS-16	DLBCL
NL-6	DLBCL
PT-11	DLBCL
PT-13	DLBCL
ES-7	FCL
FR-8	FCL
GBN-3	FCL
NL-4	FCL
PT-7	FCL
PT-8	FCL
ES-4	MCL
GBS-1	MCL
GBS-17	MCL
NL-15	MCL
PT-10	MCL
NL-17	MF
ES-8	MM_PCL
NL-12	MM_PCL
ES-14	Normal Peripheral Blood
PT-2	Normal Peripheral Blood
FR-10	PTCL
GBN-1	PTCL
GBN-5	PTCL
NL-18	PTCL
PT-3	PTCL
DE-5	Reactive lymph node
GBN-9	Reactive lymph node
PT-1	Reactive lymph node
ES-11	Reactive tonsil

GBS-11	Reactive tonsil
NL-7	Reactive tonsil
DE-1	T-ALL
DE-10	T-ALL
FR-12	T-ALL
ES-9	T-LBL
GBS-15	T-PLL

Table S4.1 List of IGH Genes on Panel

IG Locus	Chromosome	V Gene	D Gene	J Gene
IGH	chr14	IGHV6-1	IGHD1-26	IGHJ1
IGH	chr14	IGHV1-2	IGHD6-25	IGHJ2
IGH	chr14	IGHV1-3	IGHD5-24	IGHJ3
IGH	chr14	IGHV4-4	IGHD4-23	IGHJ4
IGH	chr14	IGHV7-4-1	IGHD3-22	IGHJ5
IGH	chr14	IGHV2-5	IGHD2-21	IGHJ6
IGH	chr14	IGHV3-7	IGHD1-20	
IGH	chr14	IGHV1-8	IGHD6-19	
IGH	chr14	IGHV3-9	IGHD5-18	
IGH	chr14	IGHV3-11	IGHD4-17	
IGH	chr14	IGHV3-13	IGHD3-16	
IGH	chr14	IGHV3-15	IGHD2-15	
IGH	chr14	IGHV3-16	IGHD1-14	
IGH	chr14	IGHV1-18	IGHD6-13	
IGH	chr14	IGHV3-20	IGHD5-12	
IGH	chr14	IGHV3-21	IGHD4-11	
IGH	chr14	IGHV3-23	IGHD3-10	
IGH	chr14	IGHV1-24	IGHD3-9	
IGH	chr14	IGHV3-25	IGHD2-8	
IGH	chr14	IGHV2-26	IGHD1-7	
IGH	chr14	IGHV4-28	IGHD6-6	
IGH	chr14	IGHV3-30	IGHD5-5	
IGH	chr14	IGHV4-31	IGHD4-4	
IGH	chr14	IGHV3-33	IGHD3-3	
IGH	chr14	IGHV4-34	IGHD2-2	
IGH	chr14	IGHV3-35	IGHD1-1	
IGH	chr14	IGHV3-38		
IGH	chr14	IGHV4-39		
IGH	chr14	IGHV3-43		
IGH	chr14	IGHV1-45		
IGH	chr14	IGHV1-46		
IGH	chr14	IGHV3-48		
IGH	chr14	IGHV3-49		
IGH	chr14	IGHV5-51		
IGH	chr14	IGHV3-53		

IGH	chr14	IGHV1-58		
IGH	chr14	IGHV4-59		
IGH	chr14	IGHV4-61		
IGH	chr14	IGHV3-64		
IGH	chr14	IGHV3-66		
IGH	chr14	IGHV1-69		
IGH	chr14	IGHV2-70		
IGH	chr14	IGHV3-72		
IGH	chr14	IGHV3-73		
IGH	chr14	IGHV3-74		
IGH	chr14	IGHV7-81		

Table S4.2 List of IGK Genes on Panel

IG Locus	Chromosome	V Gene	J Gene	Misc
IGK	chr2	IGKV4-1	IGKJ1	Kde
IGK	chr2	IGKV5-2	IGKJ2	IntronRSS
IGK	chr2	IGKV1-5	IGKJ3	
IGK	chr2	IGKV1-6	IGKJ4	
IGK	chr2	IGKV3-7	IGKJ5	
IGK	chr2	IGKV1-8		
IGK	chr2	IGKV1-9		
IGK	chr2	IGKV3-11		
IGK	chr2	IGKV1-12		
IGK	chr2	IGKV1-13		
IGK	chr2	IGKV3-15		
IGK	chr2	IGKV1-16		
IGK	chr2	IGKV1-17		
IGK	chr2	IGKV3-20		
IGK	chr2	IGKV6-21		
IGK	chr2	IGKV2-24		
IGK	chr2	IGKV1-27		
IGK	chr2	IGKV2-28		
IGK	chr2	IGKV2-29		
IGK	chr2	IGKV2-30		
IGK	chr2	IGKV1-33		
IGK	chr2	IGKV1-37		
IGK	chr2	IGKV1-39		
IGK	chr2	IGKV2-40		

Table S4.3 List of IGL Genes on Panel

IG Locus	Chromosome	V Gene	J Gene
IGL	chr22	IGLV4-69	IGLJ1
IGL	chr22	IGLV8-61	IGLJ2
IGL	chr22	IGLV4-60	IGLJ3
IGL	chr22	IGLV6-57	IGLJ4
IGL	chr22	IGLV11-55	IGLJ5
IGL	chr22	IGLV10-54	IGLJ6
IGL	chr22	IGLVpreB	IGLJ7

IGL	chr22	IGLV5-52	
IGL	chr22	IGLV1-51	
IGL	chr22	IGLV1-50	
IGL	chr22	IGLV9-49	
IGL	chr22	IGLV5-48	
IGL	chr22	IGLV1-47	
IGL	chr22	IGLV7-46	
IGL	chr22	IGLV5-45	
IGL	chr22	IGLV1-44	
IGL	chr22	IGLV7-43	
IGL	chr22	IGLV1-41	
IGL	chr22	IGLV1-40	
IGL	chr22	IGLV5-37	
IGL	chr22	IGLV1-36	
IGL	chr22	IGLV7-35	
IGL	chr22	IGLV2-34	
IGL	chr22	IGLV2-33	
IGL	chr22	IGLV3-32	
IGL	chr22	IGLV3-27	
IGL	chr22	IGLV3-26	
IGL	chr22	IGLV3-25	
IGL	chr22	IGLV2-23	
IGL	chr22	IGLV3-22	
IGL	chr22	IGLV3-21	
IGL	chr22	IGLV3-19	
IGL	chr22	IGLV2-18	
IGL	chr22	IGLV3-16	
IGL	chr22	IGLV2-14	
IGL	chr22	IGLV3-12	
IGL	chr22	IGLV2-11	
IGL	chr22	IGLV3-10	
IGL	chr22	IGLV3-9	
IGL	chr22	IGLV2-8	
IGL	chr22	IGLV4-3	
IGL	chr22	IGLV3-1	

Table S4.4 List of TRA Genes on Panel

TCR Locus	Chromosome	V Gene	J Gene
TRA	chr14	TRAV1-1	TRAJ61
TRA	chr14	TRAV1-2	TRAJ59
TRA	chr14	TRAV2	TRAJ58
TRA	chr14	TRAV3	TRAJ57
TRA	chr14	TRAV4	TRAJ56
TRA	chr14	TRAV5	TRAJ54
TRA	chr14	TRAV6	TRAJ53
TRA	chr14	TRAV7	TRAJ52
TRA	chr14	TRAV8-1	TRAJ50

TRA	chr14	TRAV9-1	TRAJ49
TRA	chr14	TRAV10	TRAJ48
TRA	chr14	TRAV12-1	TRAJ47
TRA	chr14	TRAV8-2	TRAJ46
TRA	chr14	TRAV8-3	TRAJ45
TRA	chr14	TRAV13-1	TRAJ44
TRA	chr14	TRAV12-2	TRAJ43
TRA	chr14	TRAV8-4	TRAJ42
TRA	chr14	TRAV13-2	TRAJ41
TRA	chr14	TRAV14/DV4	TRAJ40
TRA	chr14	TRAV9-2	TRAJ39
TRA	chr14	TRAV12-3	TRAJ38
TRA	chr14	TRAV8-6	TRAJ37
TRA	chr14	TRAV16	TRAJ36
TRA	chr14	TRAV17	TRAJ35
TRA	chr14	TRAV18	TRAJ34
TRA	chr14	TRAV19	TRAJ33
TRA	chr14	TRAV20	TRAJ32
TRA	chr14	TRAV21	TRAJ31
TRA	chr14	TRAV22	TRAJ30
TRA	chr14	TRAV23DV6	TRAJ29
TRA	chr14	TRDV1	TRAJ28
TRA	chr14	TRAV24	TRAJ27
TRA	chr14	TRAV25	TRAJ26
TRA	chr14	TRAV26-1	TRAJ25
TRA	chr14	TRAV8-7	TRAJ24
TRA	chr14	TRAV27	TRAJ23
TRA	chr14	TRAV29DV5	TRAJ22
TRA	chr14	TRAV30	TRAJ21
TRA	chr14	TRAV26-2	TRAJ20
TRA	chr14	TRAV34	TRAJ19
TRA	chr14	TRAV35	TRAJ18
TRA	chr14	TRAV36DV7	TRAJ17
TRA	chr14	TRAV38-1	TRAJ16
TRA	chr14	TRAV38-2DV8	TRAJ15
TRA	chr14	TRAV39	TRAJ14
TRA	chr14	TRAV40	TRAJ13
TRA	chr14	TRAV41	TRAJ12
TRA	chr14		TRAJ11
TRA	chr14		TRAJ10
TRA	chr14		TRAJ9
TRA	chr14		TRAJ8
TRA	chr14		TRAJ7
TRA	chr14		TRAJ6
TRA	chr14		TRAJ5

TRA	chr14		TRAJ4
TRA	chr14		TRAJ3
TRA	chr14		TRAJ2
TRA	chr14		TRAJ1

Table S4.5 List of TRD Genes on Panel

TCR Locus	Chromosome	V Gene	D Gene	J Gene
TRD	chr14	TRDV2	TRDD1	TRDJ1
TRD	chr14	TRDV3	TRDD2	TRDJ4
TRD	chr14		TRDD3	TRDJ2
TRD	chr14			TRDJ3

Table S4.6 List of TRB Genes on Panel

TCR Locus	Chromosome	V Gene	D Gene	J Gene
TRB	chr7	TRBV2	TRBD1	TRBJ1-1
TRB	chr7	TRBV3-1	TRBD2	TRBJ1-2
TRB	chr7	TRBV4-1		TRBJ1-3
TRB	chr7	TRBV5-1		TRBJ1-4
TRB	chr7	TRBV6-1		TRBJ1-5
TRB	chr7	TRBV7-1		TRBJ1-6
TRB	chr7	TRBV4-2		TRBJ2-1
TRB	chr7	TRBV6-4		TRBJ2-2
TRB	chr7	TRBV7-3		TRBJ2-3
TRB	chr7	TRBV5-3		TRBJ2-4
TRB	chr7	TRBV9		TRBJ2-5
TRB	chr7	TRBV10-1		TRBJ2-6
TRB	chr7	TRBV11-1		TRBJ2-7
TRB	chr7	TRBV10-2		
TRB	chr7	TRBV6-5		
TRB	chr7	TRBV7-4		
TRB	chr7	TRBV5-4		
TRB	chr7	TRBV6-6		
TRB	chr7	TRBV5-5		
TRB	chr7	TRBV5-6		
TRB	chr7	TRBV6-7		
TRB	chr7	TRBV7-6		
TRB	chr7	TRBV6-8		
TRB	chr7	TRBV19		
TRB	chr7	TRBV20-1		
TRB	chr7	TRBV23-1		
TRB	chr7	TRBV24-1		
TRB	chr7	TRBV25-1		
TRB	chr7	TRBV27		
TRB	chr7	TRBV28		
TRB	chr7	TRBV29-1		

TRB	chr7	TRBV30		

Table S4.7 List of TRG Genes on Panel

TCR Locus	Chromosome	V Gene	J Gene
TRG	chr7	TRGV11	TRGJ2
TRG	chr7	TRGVB	TRGJP2
TRG	chr7	TRGV10	TRGJ1
TRG	chr7	TRGV9	TRGJP
TRG	chr7	TRGVA	TRGJP1
TRG	chr7	TRGV8	
TRG	chr7	TRGV7	
TRG	chr7	TRGV6	
TRG	chr7	TRGV5	
TRG	chr7	TRGV4	
TRG	chr7	TRGV3	
TRG	chr7	TRGV2	
TRG	chr7	TRGV1	

Table S4.8 List of Genes Assessed for Mutation on Panel

Chromosome	Region Name	Coding Sequence / Hotspot
chr1	ID3	Entire Coding Sequence
chr1	ARID1A	Entire Coding Sequence
chr1	JAK1	Exons 12-15
chr1	NRAS	Exons 2-4
chr1	NOTCH2	Exon 34
chr2	DNMT3A	Exons 8-23
chr2	XPO1	Exons 15-16
chr2	CXCR4	Entire Coding Sequence
chr2	SF3B1	Exons 12-17
chr2	IDH1	Exon 4
chr3	MYD88	Exon 5
chr3	RHOA	Entire Coding Sequence
chr3	PIK3CA	Exons 2, 5 10 & 21
chr4	KIT	Exons 9, 11, 13, 17 & 18
chr4	TET2	Entire Coding Sequence
chr4	FBXW7	Entire Coding Sequence
chr4	FAT1	Entire Coding Sequence
chr5	IL7R	Exon 6
chr6	HIST1H1C	Entire Coding Sequence
chr6	HIST1H1E	Entire Coding Sequence
chr6	HIST1H1D	Entire Coding Sequence
chr6	HIST1H1B	Entire Coding Sequence
chr6	CCND3	Exon 5
chr6	NFKBIE	Exon 1
chr6	TNFAIP3	Entire Coding Sequence
chr7	CARD11	Exons 2-16
chr7	IKZF1	Entire Coding Sequence
chr7	POT1	Entire Coding Sequence

chr7	BRAF	Exons 11 & 15
chr7	EZH2	Entire Coding Sequence
chr8	MYC	Entire Coding Sequence
chr9	JAK2	Exons 12 & 14
chr9	CDKN2A	Entire Coding Sequence
chr9	CDKN2B	Entire Coding Sequence
chr9	PAX5	Entire Coding Sequence
chr9	ABL1	Exons 4-9
chr9	NOTCH1	Exons 26, 27 & 34
chr9	TRAF2	Entire Coding Sequence
chr10	EGR2	Mutation hotspot in Exon 2
chr10	PTEN	Entire Coding Sequence
chr10	NT5C2	Exons 9-15
chr11	WT1	Entire Coding Sequence
chr11	BIRC3	Exons 8 & 9
chr11	ATM	Entire Coding Sequence
chr11	CBL	Exons 7-9
chr12	KRAS	Exons 2, 3 & 4
chr12	MLL2	Entire Coding Sequence
chr12	BTG1	Entire Coding Sequence
chr14	TRAF3	Entire Coding Sequence
chr15	MAP2K1	Exons 1-8
chr15	IDH2	Exons 4 & 7
chr16	CREBBP	Entire Coding Sequence
chr16	SOCS1	Entire Coding Sequence
chr16	PLCG2	Exons 19, 20 & 24
chr17	TP53	Entire Coding Sequence
chr17	STAT5B	Exons 13-19
chr17	STAT3	Exons 10-16 & 20-21
chr17	MAP3K14	Entire Coding Sequence
chr17	CD79B	Exons 4, 5 & 6
chr18	BCL2	Entire Coding Sequence
chr19	TCF3	Entire Coding Sequence
chr19	KLF2	Entire Coding Sequence
chr19	JAK3	Exons 12-14
chr19	CD79A	Exons 4 & 5
chr20	ASXL1	Mutation hotspot in last exon
chr20	SAMHD1	Entire Coding Sequence
chr21	RUNX1	Entire Coding Sequence
chr21	ERG	Entire Coding Sequence
chr22	MAPK1	Entire Coding Sequence
chr22	EP300	Entire Coding Sequence
chrX	BTK	Entire Coding Sequence
chrX	PHF6	Entire Coding Sequence

Table S4.9 List of Genes or Chromosome Regions Assessed for Copy Number Variation on Panel

Chromosome	Region Name	Parameter
chr3	PIK3CA	Copy Number
chr7	IKZF1	Copy Number
chr8	MYC	Copy Number
chr8	8qTEL	Copy Number
chr9	9pTEL	Copy Number
chr9	CDKN2A	Copy Number
chr9	CDKN2B	Copy Number
chr9	PAX5	Copy Number
chr9	9pCEN	Copy Number
chr11	11qCEN	Copy Number
chr11	CCND1	Copy Number
chr11	ATM	Copy Number
chr11	11qTEL	Copy Number
chr12	12pTEL	Copy Number
chr12	ET6	Copy Number
chr12	12pCEN	Copy Number
chr12	12qCEN	Copy Number
chr12	BTG1	Copy Number
chr12	12qTEL	Copy Number
chr13	13qCEN	Copy Number
chr13	RB1	Copy Number
chr13	13qTEL	Copy Number
chr17	17pTEL	Copy Number
chr7	7pTEL	Copy Number
chr17	TP53	Copy Number
chr17	17pCEN	Copy Number

Table S4.10 List of Additional Regions Assessed for Translocations on Panel

Chromosome	Region Name	Parameter	Additional Description
chr1	STIL	Translocation	N/A
chr2	ALK	Translocation	N/A
chr3	BCL6	Translocation	N/A
chr11	BCL1	Translocation	N/A
chr11	BIRC3	Translocation	N/A
chr11	KMT2A	Translocation	N/A
chr14	IGHA2	Translocation	switch region
chr14	IGHE	Translocation	switch region
chr14	IGHG4	Translocation	switch region
chr14	IGHG2	Translocation	switch region
chr14	IGHA1	Translocation	switch region
chr14	IGHG1	Translocation	switch region
chr14	IGHG3	Translocation	switch region
chr18	BCL2	Translocation	N/A
chrX	CRLF2	Translocation	N/A

Table S5. Regions covered on smMIP-based NGS panel performed at Radboud University Medical Center

Gene	RefSeq	Region of interest
AKT1	NM_005163.2	codon 17
BIRC3	NM_182962.2	exon 7 to 10
BRAF	NM_004333.4	codon 582-615
BTK	NM_000061.2	codon 481
CDKN2A	(P14) NM_058195.1	>95% of coding sequence and splice sites (-5/+5)
CDKN2A	(P16) NM_000077.3	>95% of coding sequence and splice sites (-5/+5)
CTNNB1	NM_001904.3	codon 19-48
CXCR4	NM_001008540.1	codon 281-357
CYSLTR2	NM_001308471.2	codon 129
EGFR	NM_005228.3	codon 434-499, 688-823, 849-875
EIF1AX	NM_001412.4	codon 1-16
ERBB2	NM_004448.3	codon 770-785
EZH2	NM_004456.4	codon 471-502, 618-645, 679-704
GNA11	NM_002067.4	codon 183 and 209
GNAQ	NM_002072.3	codon 183 and 209
GNAS	NM_000516.4	codon 201 and 227
H3F3A	NM_002107.4	codon 28 and 35
H3F3B	NM_005324.4	codon 37
HRAS	NM_005343.2	codon 12, 13, 59 and 61
IDH1	NM_005896.3	codon 132
IDH2	NM_002168.3	codon 140 and 172
JAK2	NM_004972.3	codon 617
KIT	NM_000222.2	codon 412-513, 550-591, 628-713,799-828
KRAS	NM_004985.4	codon 12, 13, 59, 61, 117 and 146
MPL	NM_005373.2	codon 515
MYD88	NM_002468.4	codon 169-280
NRAS	NM_002524.4	codon 12, 13, 59, 61, 117 and 146
PDGFRA	NM_006206.4	codon 552-596, 632-667, 814-848
PIK3CA	NM_006218.2	codon 520-554, 1020-1069
PIM1	NM_002648.3	exon 1 to 6
PLCB4	NM_000933.3	codon 630
PLCG2	NM_002661.4	codon 665, 707-708, 845-846, 993, 1140-1144
SF3B1	NM_012433.2	codon 603-671, 694-727, 833-906
TP53	NM_000546.5	>95% of coding sequence and splice sites (-5/+5)

Table S6.1 Underperforming regions of interest

Chromosome	Start	Stop	Target[a]	Classification	GC Content[b]	HMW[c] (n=26)	FFPE[c] (n=21)
chr14	106088022	106088322	IGHV1-8	Rearrangement	0.52	53.8	71.4
chr11	132697125	132697225	11qTEL_rs1940150	reference SNP ID	0.30	100.0	81.0
chr12	43448951	43449051	12qCEN_rs2220865	reference SNP ID	0.34	100.0	90.5
chr12	131030078	131030178	12qTEL_rs1195889	reference SNP ID	0.56	57.7	100.0
chr12	131033171	131033271	12qTEL_rs7960677	reference SNP ID	0.60	100.0	90.5
chr13	23741443	23741543	13qCEN_rs36116586	reference SNP ID	0.38	100.0	71.4
chr17	20211440	20211540	17pCEN_rs2703812	reference SNP ID	0.57	100.0	90.5
chr17	19014895	19014995	17pCEN_rs28526726	reference SNP ID	0.65	92.3	76.2
chr17	714749	714849	17pTEL_rs2474694	reference SNP ID	0.72	100.0	95.2
chr7	1318876	1318976	7pTEL_rs10227104	reference SNP ID	0.65	69.2	66.7
chr7	1268215	1268315	7pTEL_rs12535561	reference SNP ID	0.64	100.0	100.0
chr7	1260115	1260215	7pTEL_rs36022572	reference SNP ID	0.71	80.8	81.0
chr7	1298433	1298533	7pTEL_rs4724908	reference SNP ID	0.61	92.3	57.1
chr8	140556338	140556438	8qTEL_rs2271736	reference SNP ID	0.66	53.8	52.4
chr8	140542050	140542150	8qTEL_rs2977475	reference SNP ID	0.63	100.0	66.7
chr8	140563209	140563309	8qTEL_rs2977490	reference SNP ID	0.54	96.2	95.2
chr12	11893674	11893774	ET6CN_rs2855750	reference SNP ID	0.26	100.0	100.0
chr9	36943480	36943580	PAX5CN_rs10123881	reference SNP ID	0.41	100.0	100.0
chr9	37025371	37025471	PAX5CN_rs3758178	reference SNP ID	0.65	100.0	100.0
chr13	48317650	48317750	RB1CN_rs3825417	reference SNP ID	0.72	96.2	61.9
chr13	48457646	48457746	RB1CN_rs9568042	reference SNP ID	0.72	100.0	95.2
chr13	48461409	48461509	RB1CN_rs9568043	reference SNP ID	0.33	100.0	100.0
chr19	17830098	17830228	JAK3_23	Somatic Mutation	0.69	100.0	95.2
chr22	21867311	21867520	MAPK1_1	Somatic Mutation	0.76	100.0	76.2
chr19	1622044	1622233	TCF3_10	Somatic Mutation	0.72	80.8	81.0
chr19	1621103	1621203	TCF3_12	Somatic Mutation	0.71	84.6	76.2
chr19	1650158	1650258	TCF3_2	Somatic Mutation	0.64	100.0	90.5
chr19	1622303	1622425	TCF3_9	Somatic Mutation	0.67	100.0	95.2
chr9	136886480	136886581	TRAF2_1	Somatic Mutation	0.80	100.0	100.0
chr9	136886711	136886814	TRAF2_2	Somatic Mutation	0.70	100.0	76.2
chrX	1212556	1218022	CRLF2_MBR	Translocation	0.53	50.0	71.4

[a] Regions were classed as underperforming if coverage was >2 standard deviations (SD) below the mean in ≥ 50% samples.

[b] GC content, expressed as a range between 0-1, of the region captured by the specified probe

[c] Percentage of samples with a log transformed coverage value >2 standard deviations (SD) of the mean .

Table S6.2 Sequencing performance

	Proficiency[a] (n = 98)	Validation (All)[b] (n = 308)	Validation (HMW)[c] (n = 220)	Validation (FFPE)[d] (n = 88)
Total Reads - Average	17,594,145	12,711,040	13,282,842	11,281,534
Total Reads - SD	6,914,318	3,619,403	3,305,593	3,978,340
Unique PF Reads - Average	10,977,284	7,261,695	7,700,370	6,165,008
Unique PF Reads - SD	4,337,890	2,585,797	2,625,606	2,130,213
% PCR Duplicates - Average	37.3	45.7	45.2	47.0
% PCR Duplicates - SD	13.4	10.4	10.5	10.0
Coverage Depth - Average	1628.2	1037.9	1067.0	965.3
Coverage Depth - SD	627.1	422.7	417.4	429.7

[a] Each of the seven sites performed a proficiency run comprising 14 cell lines (7 x 14 = 98)

[b] Each of the seven sites performed 2 validations runs comprising 22 samples (14 x 22 = 308)

[c] High molecular weight (HMW) samples which comprised 220 of the 308 samples in the validation run were analysed separately

[d] Formalin fixed paraffin embedded (FFPE) samples which comprised 88 of the 308 samples in the validation run were analysed separately

Table S7. Limit of Detection

NGS Site	Cell Line[a]	Cell Lineage	Rearrangements[b]				Translocations[c]			
			10%	5%	2.5%	1.25%	10%	5%	2.5%	1.25%
Site 1	RS4:11	B-cell	4/4	4/4	4/4	1/4	1/1	1/1	1/1	1/1
Site 2	KCA	B-cell	4/4	4/4	2/4	2/4	N/A	N/A	N/A	N/A
Site 3	Loucy	T-cell	6/6	6/6	6/6	0/6	N/A	N/A	N/A	N/A
Site 4	KT1	T-cell	4/4	4/4	3/4	2/4	N/A	N/A	N/A	N/A
Site 5	H929	B-cell	4/4	2/4	0/4	1/4	1/1	1/1	0/1	0/1
Site 6	Ichikawa	T-cell	5/5	5/5	5/5	0/5	1/1	1/1	0/1	0/1
Site 7	REH	T-cell	7/7	7/7	6/7	5/7	1/1	0/1	0/1	0/1
Expected Events			34	34	34	34	4	4	4	4
Observed Events			34	32	26	11	4	3	1	1
Percentage Found			100.0	94.1	76.5	32.4	100	75	25	25

[a] Cell line detailed was diluted with gDNA from a different cell line at mass ratios of 10%, 5.0%, 2.5% and 1.25%.

[b] The number of observed (numerator) and expected (denominator) rearrangements for each dilution of the cell line is shown.

[c] The number of observed (numerator) and expected (denominator) translocations for each dilution of the cell line is shown.

Table S8. Summary of cases with clonal IGH rearrangement by EuroClonality-NDC assay but not by BIOMED-2 PCR

Disease Entity[a]	BIOMED-2 PCR Result								EuroClonality-NDC Result			
	IGH[b]					IGK[b]		IGL[b]	IGH[b]	IGK[b]	IGL[b]	Translocation
		A	B	C	D	A	B					
	LD	FR1	FR2	FR3	D-J	V-J	Kde	V-J				
AITL	-	NC	NC	NC	-	C	NC	-	C	C	N	None
ALCL	-	P	P	P	-	P	P	-	C	N	N	ALK-NPM1
ALCL	-	P	P	P	-	P	P	-	C	C	N	ALK-NPM1
B-ALL	-	P	P	P	NC	NC	C	-	C	C	C	IGH-BLID
DLBCL	-	P	P	P	-	C	P	-	C	C	N	None
DLBCL	-	P	P	P	-	P	P	-	C	C	N	IGH-BCL6
FL	-	P	P	P	-	C	P	-	C	C	N	None
FL	-	P	P	P	-	C	P	-	C	C	C	IGH-BCL2
FL	-	P	P	P	-	C	P	-	C	C	C	IGH-BCL2
MCL	-	P	P	P	-	C	P	-	C	C	N	IGH-CCND1
PCM	-	-	P	P	-	-	-	-	C	C	N	IGH-CCND1

[a]Disease entity: AITL, angioimmunoblastic T-cell lymphoma; ALCL, anaplastic large-cell lymphoma; B-ALL, B-cell Acute Lymphoblastic Leukemia; DLBCL, diffuse large B-cell lymphoma; FL, follicular lymphoma; MCL, mantle cell lymphoma; PCM, plasma cell myeloma.

[b]Rearrangement patterns of IG loci: C, clonal; NC, not clonal; P, Polyclonal; "-", clonality data not provided by original laboratory.

Table S9. Summary of cases with clonal IGK rearrangement by EuroClonality-NDC assay but not by BIOMED-2 PCR

Disease Entity[a]	BIOMED-2 PCR Result								EuroClonality-NDC Result				
	IGH[b]					IGK[b]		IGL[b]	IGH[b]	IGK[b]		IGL[b]	Translocation
		A	B	C	D	A	B						
	LD	FR1	FR2	FR3	D-J	V-J	Kde	V-J		Vκ-Jκ	Vκ-Kde		
ALCL	-	P	P	P	-	P	P	-	C	NC	C	NC	ALK-NPM1
B-ALL	-	C	C	C	P	P	P	-	C	C	NC	NC	IGH-MYC
BL	-	C	C	C	-	P	P	-	C	C	N	NC	IGH-MYC
BL	-	NC	NC	C	C	NC	NC	-	C	C	NC	NC	IGH-MYC
BL	-	P	C	C	-	P	P	-	C	NC	C	C	IGH-MYC
DLBCL	-	P	C	P	N	P	P	-	C	C	NC	NC	IGH-BCL2
DLBCL	-	P	P	P	-	P	P	-	C	C	NC	NC	IGH-BCL6
FL	-	C	P	P	-	P	P	-	C	C	C	C	IGH-BCL2
FL	-	P	P	P	-	P	P	-	NC	C	NC	NC	IGH-BCL2
FL	-	P	P	P	-	P	P	-	NC	NC	C	NC	None
FL	-	P	C	C	-	P	P	-	C	C	C	NC	IGH-BCL2
FL	-	P	P	P	-	P	P	-	NC	C	C	NC	IGH-BCL2
MALT	-	C	C	P	P	P	NC	-	C	C	NC	NC	None
MALT	-	C	C	C	P	P	P	-	C	C	NC	NC	None
MCL	-	C	C	C	P	P	P	-	C	C	NC	NC	IGH-CCND1
MCL	-	P	C	C	-	P	P	-	C	C	NC	NC	IGH-CCND1
MCL	-	C	C	C	-	P	P	-	C	C	C	C	IGH-CCND1
PCM	-	C	C	P	C	P	P	-	C	C	NC	NC	None
PCM	-	P	C	P	C	P	P	-	C	C	NC	NC	None
PCM	-	C	C	-	P	-	P	-	C	C	NC	C	None
PCM	-	-	C	P	P	P	P	-	C	C	NC	NC	IGH-CCND1
MZL	-	C	C	C	-	P	P	-	C	C	C	NC	None
MZL	-	P	P	P	C	P	P	-	C	NC	C	NC	MALT1-BIRC3
MZL	-	C	C	C	C	P	P	-	C	C	NC	NC	None
MZL	C	-	-	-	NC	NC	NC	C	C	C	NC	NC	None

[a]Disease entity: ALCL, anaplastic large-cell lymphoma; B-ALL, B-cell Acute Lymphoblastic Leukemia; BL, Burkitt lymphoma; DLBCL, diffuse large B-cell lymphoma; FL, follicular lymphoma; MALT, Mucosa-Associated Lymphoid Tissue; MCL, mantle cell lymphoma; PCM, plasma cell myeloma, MZL, Marginal zone lymphoma.
[b]Rearrangement patterns of IG loci: C, clonal; NC, not clonal; P, Polyclonal; "-", clonality data not provided by original laboratory.

Table S10. IGL cases per disease entity

Disease Entity[a]	Total No.	Lambda RR	% samples with lambda RR's
BL	17	7	41.2
CLL	37	17	45.9
DLBCL	21	7	33.3
FL	34	12	35.3
MALT	3	0	0.0
MCL	34	17	50.0
PCM	24	8	33.3
MZL & SMZL	7	2	28.6
B-ALL	20	6	30
Total	197	76	38.6

[a]Disease entity: BL, burkitt lymphoma; CLL, chronic lymphocytic leukemia; DLBCL, diffuse large B-cell lymphoma; FL, follicular lymphoma; MALT, Mucosa-Associated Lymphoid Tissue; MCL, mantle cell lymphoma; PCM, plasma cell myeloma, MZL, Marginal zone lymphoma; B-ALL, B-cell Acute Lymphoblastic Leukemia.

Table S11. IGL assessment in proficiency cell lines across 7 testing sites

Sample	Site 1	Site 2	Site 3	Site 4	Site 5	Site 6	Site 7	Tan et al., 2018 (PMID: 30285677)	BIOMED-2 (IGL) PCR
KCA	V1-44>J2=J3 V6-57>J3	V1-44>J2=J3 V6-57>J3	V1-44>J2=J3 V6-57>J3	V1-44>J2=J3 V6-57>J3	V1-44>J2=J3 V6-57>J3	V1-44>J2=J3 V6-57>J3	V1-44>J2=J3 V6-57>J3	Not Tested	C (x2)
Loucy	No RR	No RR	No RR	No RR	No RR	No RR	No RR	Not Clonal	P
K-T1	No RR	No RR	No RR	No RR	No RR	No RR	No RR	Not Tested	P
H929	No RR	No RR	No RR	No RR	No RR	No RR	No RR	Not Clonal	P
REH	V3-21>J3 V2-8>J2=J3	V3-21>J3 V2-8>J2=J3	V3-21>J3 V2-8>J2=J3	V3-21>J3 V2-8>J2=J3	V3-21>J3 V2-8>J2=J3	V3-21>J3 V2-8>J2=J3	V3-21>J3 V2-8>J2=J3	Not Detected	C
RS4;11	V4-3>J3	V4-3>J3	V4-3>J3	V4-3>J3	V4-3>J3	V4-3>J3	V4-3>J3	V4-3>J3 V11-55>J7	C (x2)
P12-Ichikawa	No RR	No RR	No RR	No RR	No RR	No RR	No RR	Not Clonal	P
HAL-01	No RR	No RR	No RR	No RR	No RR	No RR	No RR	Not Tested	P
Karpas299	No RR	No RR	No RR	No RR	No RR	No RR	No RR	Not Clonal	P
CA-46	No RR	No RR	No RR	No RR	No RR	No RR	No RR	Not Clonal	P
CML-T1	No RR	No RR	No RR	No RR	No RR	No RR	No RR	Not Clonal	P
TOM-1	No RR	No RR	No RR	No RR	No RR	No RR	No RR	Not Tested	P
ARR	No RR	No RR	No RR	No RR	No RR	No RR	No RR	Not Tested	P
Namalwa	V4-60>J3	V4-60>J3	V4-60>J3	V4-60>J3	V4-60>J3	V4-60>J3	V4-60>J3	V4-60>J3	C

Table S12. Cases exhibiting a discrepancy between originating lab PCR data and EuroClonality-NDC assay results for TCR rearrangement

TCR Loci[a]	Disease Entity	BIOMED-2 PCR Result							EuroClonality-NDC Result				
		TRB[c] A V-J	TRB[c] B V-J	TRB[c] C D-J	TRG[c] A V-J	TRG[c] B V-J	TRD[c] V-J, D-J	TRA[c] V-J	TRB[c]	TRG[c]	TRD[c]	TRA[c]	Translocation
TRB	AITL	P	P	P	C	C	-	-	C	C	NC	C	None
TRB	ALCL	P	P	P	P	P	-	-	C	C	C	NC	None
TRG	AITL	P	P	P	P	P	-	-	NC	C	NC	C	None
TRG	ALCL	C	P	P	P	P	-	-	C	C	C	C	ALK
TRG	ALCL	P	P	P	P	P	-	-	C	C	C	NC	ALK
TRG	ALCL	P	P	C	P	P	-	-	C	C	C	NC	ALK
TRG	ALCL	P	P	C	P	P	-	-	C	C	NC	C	ALK
TRD	B-ALL	-	-	-	C	NC	NC	-	NC	C	C	C	None
TRD	B-ALL	-	-	-	NC	C	NC	-	C	C	C	NC	None
TRD	B-ALL	-	-	-	NC	NC	NC	-	NC	NC	C	C	None
TRD	T-ALL	C	NC	C	NC	C	NC	-	C	C	C	C	TAL1-STIL
TRA	B-ALL	C	NC	NC	NC	C	NC	NC	C	C	NC	C	IGH-CRLF2
TRA	B-ALL	NC	NC	NC	NC	C	NC	NC	NC	C	NC	C	IGH-CRLF2
TRA	B-ALL	-	-	-	C	NC	NC	NC	NC	C	NC	C	IGH-CRLF2
TRA	B-ALL	-	-	-	C	C	NC	NC	C	C	NC	C	IGH-CRLF2
TRA	B-ALL	-	-	-	NC	C	NC	NC	C	C	NC	C	IGH-CRLF2
TRA	T-ALL	C	P	C	C	C	C	P	C	C	C	C	None
TRA	T-ALL	C	P	C	C	P	C	P	C	C	C	C	TRD-TLX1
TRA	T-ALL	R	P	C	C	P	C	P	C	C	C	C	TRB-TLX1
TRA	T-ALL	P	C	C	C	C	C	P	C	C	C	C	None

[a]Loci where discrepancy between EuroClonality-NDC assay and originating lab PCR data occurred.

[b] Disease entity: AITL, Angioimmunoblastic T-cell lymphoma; ALCL, Anaplastic large-cell lymphoma; B-ALL, B-cell Acute Lymphoblastic Leukemia; T-ALL, T Acute Lymphoblastic leukemia.

[c]Rearrangement patterns of TCR loci were described as C, clonal; NC, not clonal; P, Polyclonal; R, Restricted; "-", clonality data not provided by original laboratory.

Table S13. Comparison of the EuroClonality-NDC assay with Southern blot results in 50 samples from the original BIOMED-2 paper

Locus	Total Tested	Benchmark - Clonal		Benchmark – Not Clonal	
		NGS - Clonal	NGS – Not Clonal	NGS - Clonal	NGS – Not Clonal
IGH	50	32	1	3	14
IGK	50	28	2	10	10
IGL	50	9	1	4	36
TRA	0	0	0	0	0
TRB	50	13	2	7	28
TRD	50	8	1	2	39
TRG	49	10	0	12	27

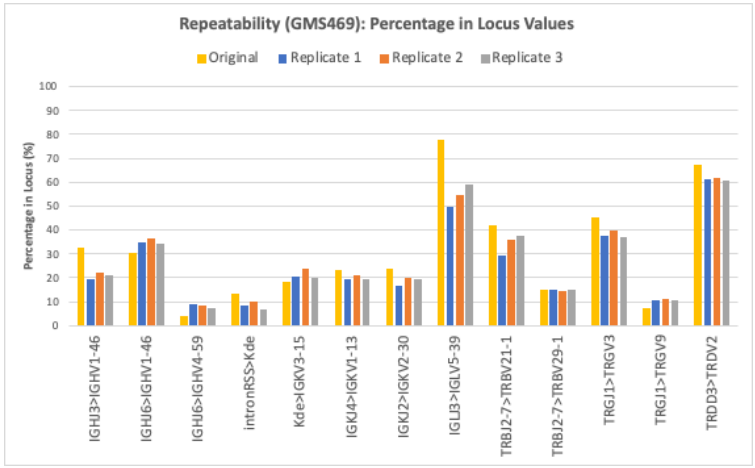
Table S14. Combination of four parameters for clonality detection

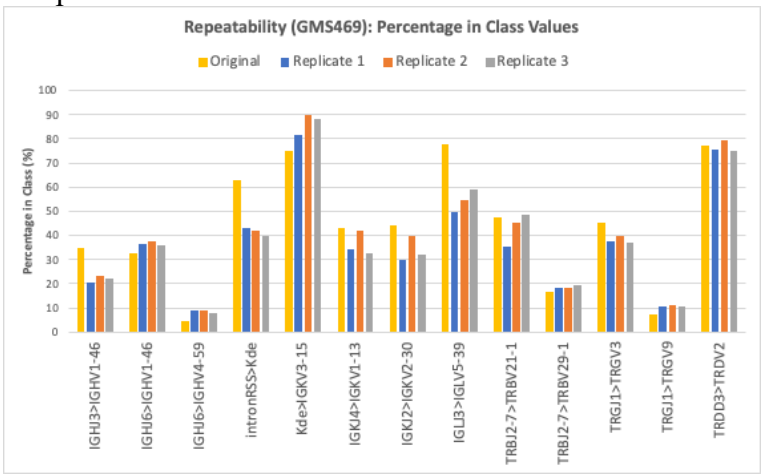
Disease Entity[a]	IG/TCR Rearrangements							Translocation	Copy Number	Mutation					
	IGH	IGK	IGL	TRB	TRG	TRD	TRA			Chr Location	Gene	VAF (%)	Annotation	HGVS.c	HGVS.p
B-ALL	N	N	N	C	C	C	C	IGH-SLC25A12	IGH, TRB, TRG	N	N	N	N	N	N
BL	N	N	N	N	N	N	N	N	IGH	N	N	N	N	N	N
BL	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
MCL	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
FL	N	N	N	N	N	N	N	IGH-BCL2	IGH, IGK	6:137879333	TNFAIP3	19.1	stop gained	c.1888G>T	p.Glu630*
										12:49024679	KMT2D	18.0	stop gained	c.15951T>A	p.Tyr5317*
										16:3749667	CREBBP	17.0	stop gained	c.3796C>T	p.Gln1266*
FL	N	N	N	N	N	N	N	IGH-BCL2	N	N	N	N	N	N	N
T-ALL	N	N	N	N	N	N	N	N	N	X:134425206	PHF6	33.5	missense	c.974A>G	p.Tyr325Cys
										11:108256229	ATM	33.0	missense	c.2139A>C	p.Glu713Asp
										9:136504937	NOTCH1	30.7	missense	c.4754T>C	p.Leu1585Pro
										21:34880619	RUNX1	30.3	frameshift	c.445_446insGG	p.Ala149fs
										1:64846682	JAK1	13.0	missense	c.1954T>A	p.Tyr652Asn
ALCL	N	N	N	N	N	N	N	ALK-RANBP2	N	N	N	N	N	N	N

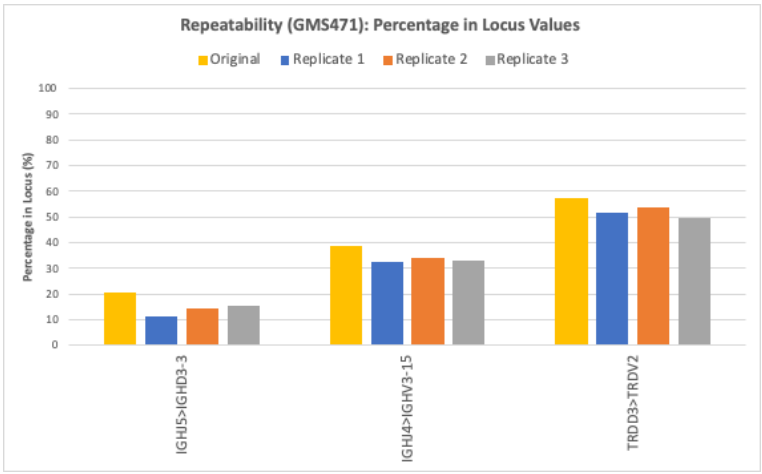
[a]Disease entity: B-ALL, B-cell Acute Lymphoblastic Leukemia; BL, burkitt lymphoma; MCL, mantle cell lymphoma; FL, follicular lymphoma; T-ALL, T Acute Lymphoblastic leukemia; ALCL Anaplastic large-cell lymphoma.

Supplementary Figures

Figure S1. Assessment of intra- and inter-run replicates using two B-ALL (GMS469 and GMS471) samples. Both samples were included in the validation cohort (labelled 'Original') and were also ran in triplicate (labelled Replicate 1-3) during the intra-run precision assessment. The 'percentage in locus' and 'percentage in class' values were compared across all runs.







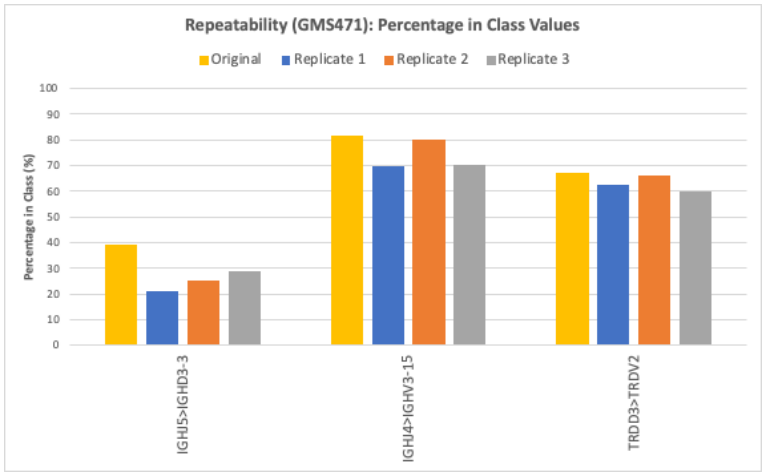


Figure S2. Association between ddPCR result and the number of fragments identified assigned to a specific rearrangement by the EuroClonality-NDC assay. The number of fragments identified by the EuroClonality-NDC assay containing the clonal rearrangement were displayed on a box and whisker plot and grouped by ddPCR classification (specific amplification defined as the clonal abundance of the rearrangement being >1% [sa]; minor amplification defined as the clonal abundance of the rearrangement being <1% [minor]; the specific signal was lower than the polyclonal control [ua]; no PCR product was evident [neg]).

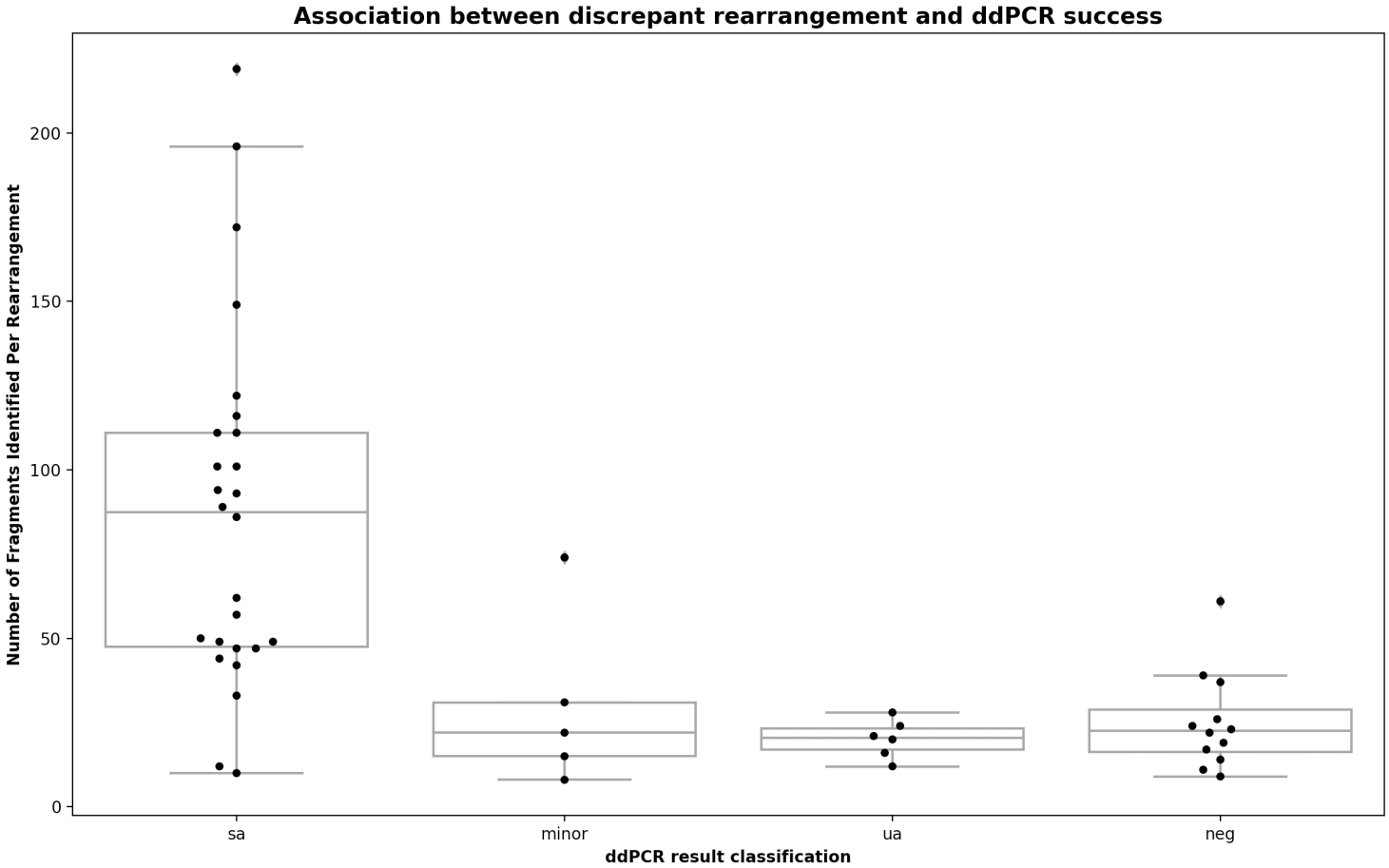


Figure S3. Confirmation of IGH rearrangements detected by the EuroClonality-NDC assay, but not by the EuroClonality/BIOMED-2 PCR, using the EuroClonality-NDC CNA pipeline. The upper panel displays the result of the rearrangement pipeline (ARResT/Interrogate). The lower panel is a graphical representation of the CNA coverage depth ratio calculated between an individual tumour sample and the mean value of normal samples for each region included in the bed file. CNA ratios <0.8 and >1.2 are coloured in blue and red respectively.

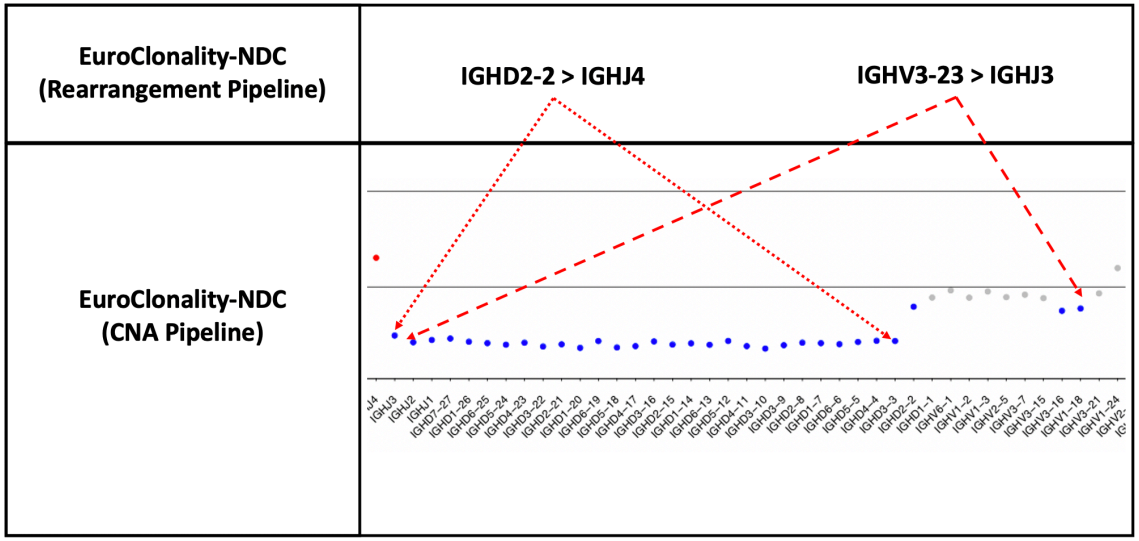


Figure S4. Confirmation of an IGK rearrangement detected by the EuroClonality-NDC assay, but not by the EuroClonality/BIOMED-2 PCR, using the EuroClonality-NDC CNA pipeline. The upper panel displays the result of the rearrangement pipeline (ARResT/Interrogate). The lower panel is a graphical representation of the CNA coverage depth ratio calculated between an individual tumour sample and the mean value of normal samples for each region included in the bed file. CNA ratios <0.8 and >1.2 are coloured in blue and red respectively.

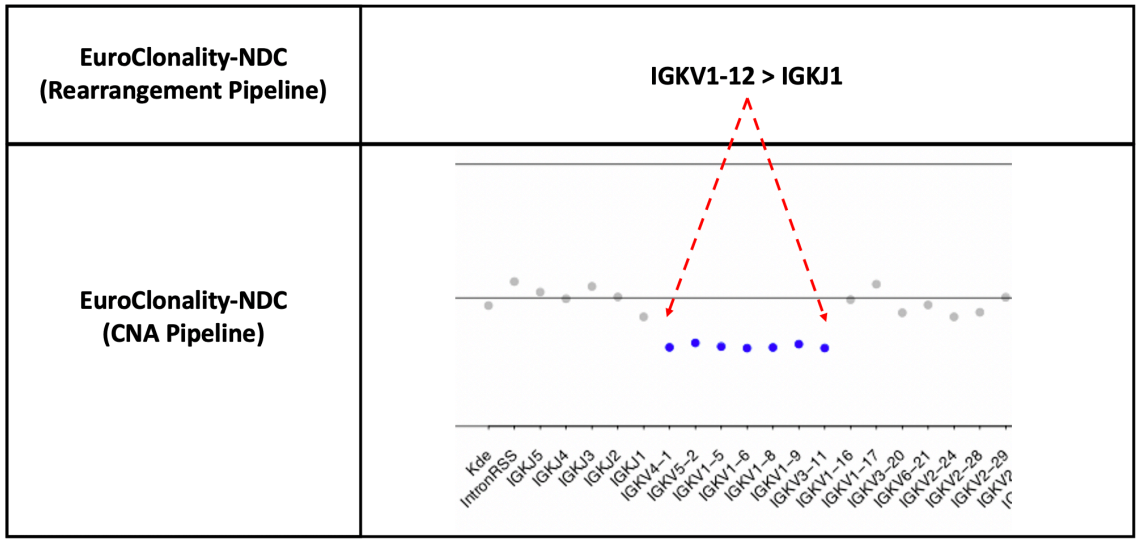


Figure S5. Confirmation of an IGL rearrangement detected by the EuroClonality-NDC assay, but not by the EuroClonality/BIOMED-2 PCR, using the EuroClonality-NDC CNA pipeline. The upper panel displays the result of the rearrangement pipeline (ARResT/Interrogate). The lower panel is a graphical representation of the CNA coverage depth ratio calculated between an individual tumour sample and the mean value of normal samples for each region included in the bed file. CNA ratios <0.8 and >1.2 are coloured in blue and red respectively.

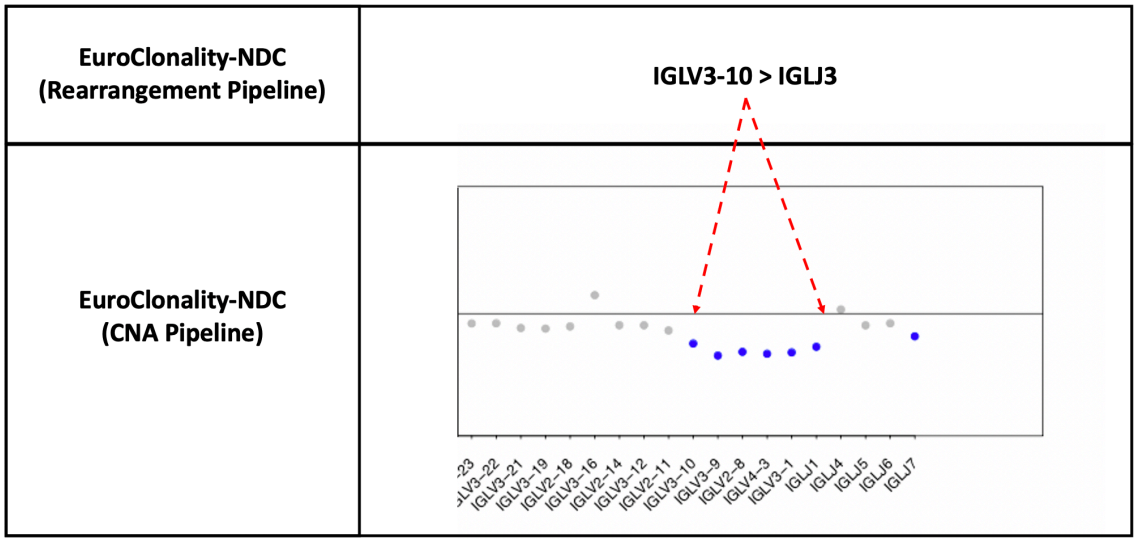


Figure S6. Oncoplot showing the most frequently mutated genes identified in B-cell Acute Lymphoblastic Leukemia (B-ALL) patients present within the validation cohort. Mutations were identified in 16 of the 20 B-ALL patients included within the final validation cohort of 270 samples. Each column represents one patient while each row represents the absence (grey) or presence (colour) of a mutation within the defined gene in each sample. Mutations are color-coded by mutation nomenclature using the key located at the bottom of the figure. Percentages indicate number of samples identified with a mutation of the defined gene. The top barplot has the number of mutations identified in each patient. The right-hand barplot provides the number of mutations in the defined gene which are coloured by mutation nomenclature.

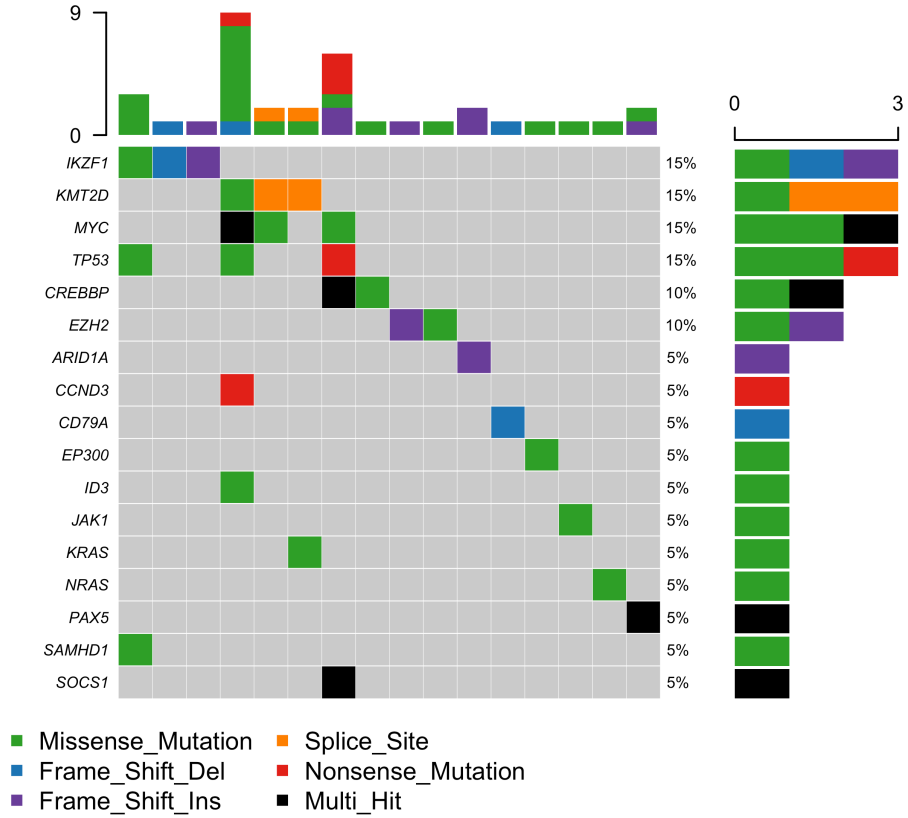


Figure S7. Oncoplot showing the most frequently mutated genes identified in Burkitt lymphoma (BL) patients present within the validation cohort. Mutations were identified in 13 of the 17 BL patients included within the final validation cohort of 270 samples. Each column represents one patient while each row represents the absence (grey) or presence (colour) of a mutation within the defined gene in each sample. Mutations are color-coded by mutation nomenclature using the key located at the bottom of the figure. Percentages indicate number of samples identified with a mutation of the defined gene. The top barplot has the number of mutations identified in each patient. The right-hand barplot provides the number of mutations in the defined gene which are coloured by mutation nomenclature.

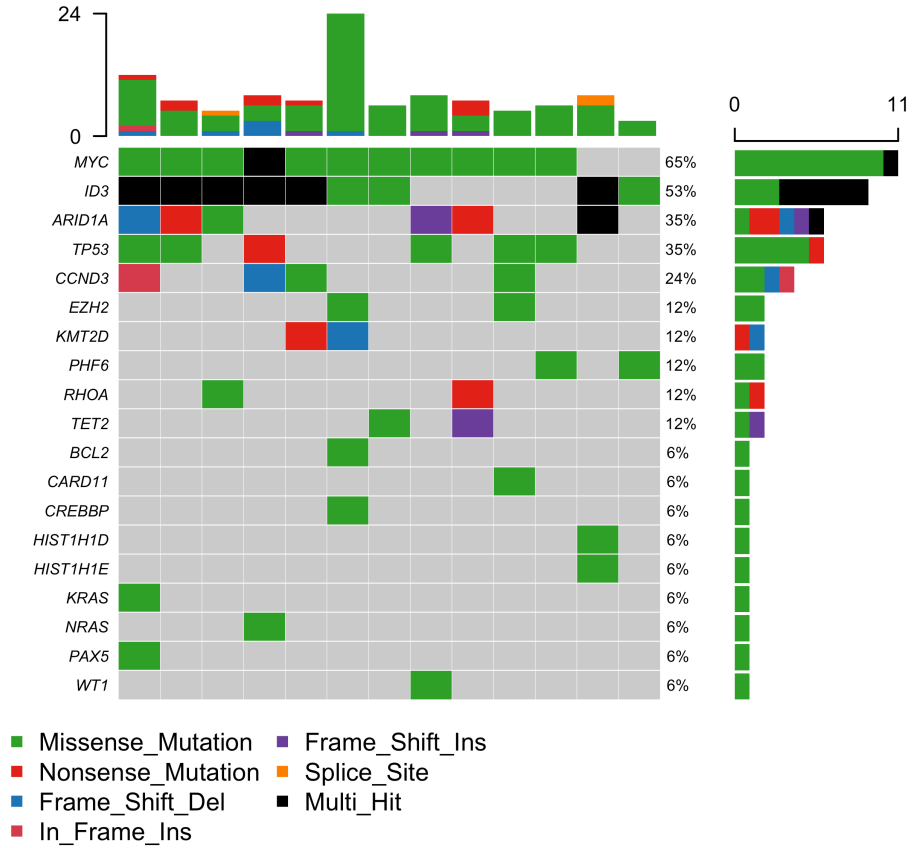


Figure S8. Oncoplot showing the most frequently mutated genes identified in Chronic lymphocytic leukemia (CLL) patients present within the validation cohort. Mutations were identified in 24 of the 37 CLL patients included within the final validation cohort of 270 samples. Each column represents one patient while each row represents the absence (grey) or presence (colour) of a mutation within the defined gene in each sample. Mutations are color-coded by mutation nomenclature using the key located at the bottom of the figure. Percentages indicate number of samples identified with a mutation of the defined gene. The top barplot has the number of mutations identified in each patient. The right-hand barplot provides the number of mutations in the defined gene which are coloured by mutation nomenclature.

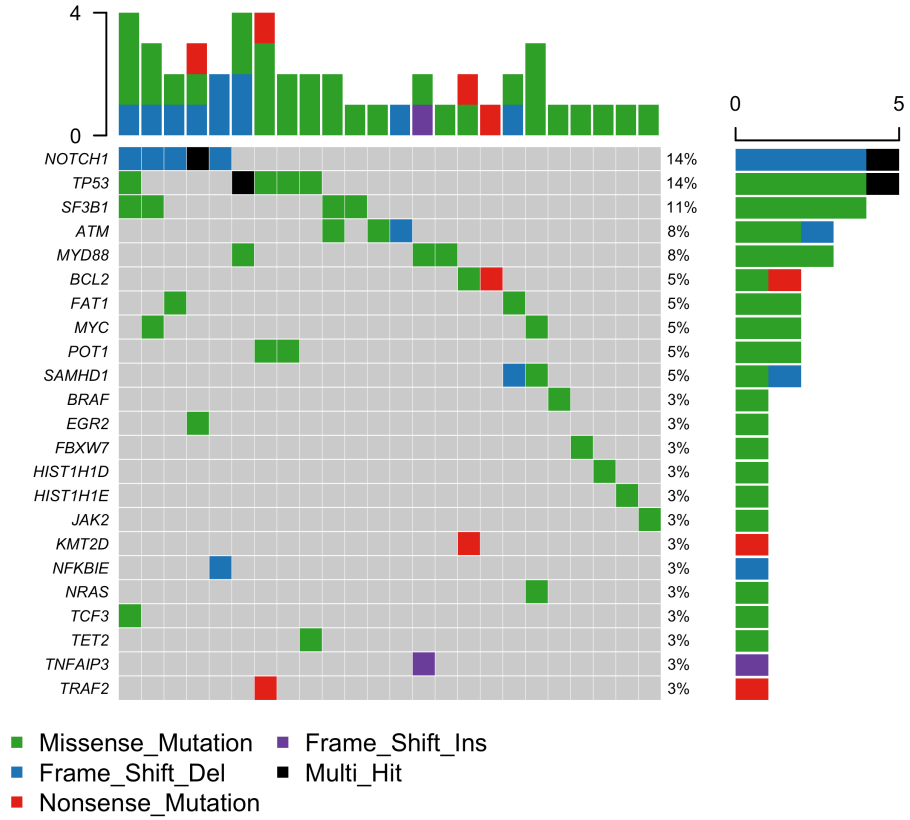


Figure S9. Oncoplot showing the most frequently mutated genes identified in Diffuse large B-cell lymphoma (DLBCL) patients present within the validation cohort. Mutations were identified in 19 of the 20 CLL patients included within the final validation cohort of 270 samples. Each column represents one patient while each row represents the absence (grey) or presence (colour) of a mutation within the defined gene in each sample. Mutations are color-coded by mutation nomenclature using the key located at the bottom of the figure. Percentages indicate number of samples identified with a mutation of the defined gene. The top barplot has the number of mutations identified in each patient. The right-hand barplot provides the number of mutations in the defined gene which are coloured by mutation nomenclature.

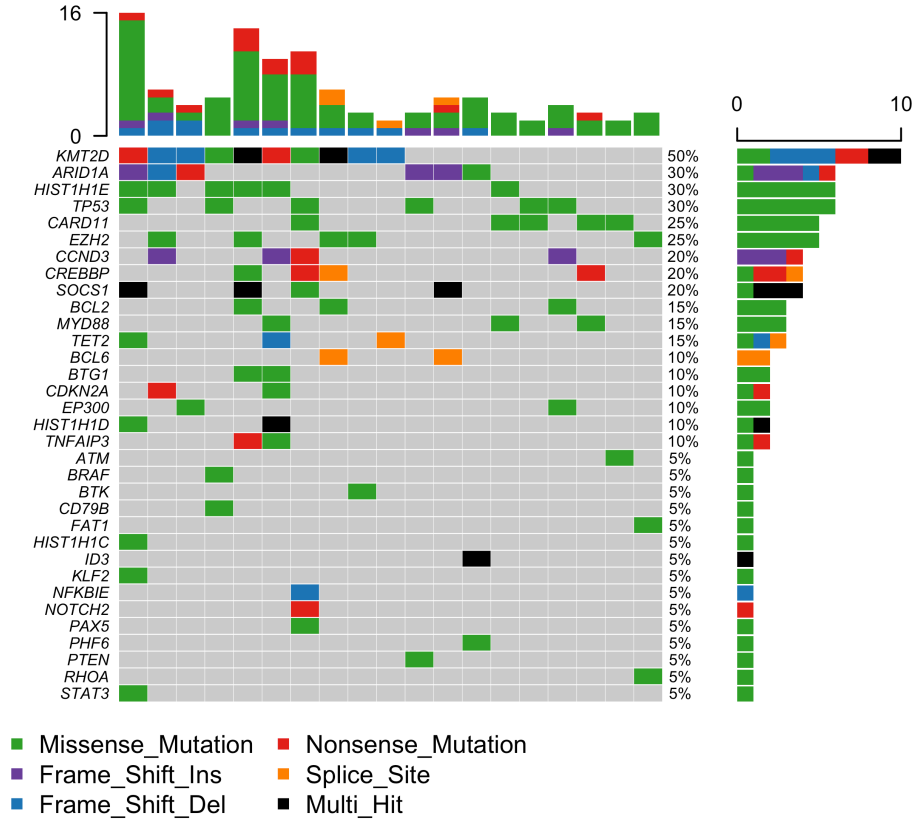


Figure S10. Oncoplot showing the most frequently mutated genes identified in Follicular lymphoma (FL) patients present within the validation cohort. Mutations were identified in 29 of the 34 FL patients included within the final validation cohort of 270 samples. Each column represents one patient while each row represents the absence (grey) or presence (colour) of a mutation within the defined gene in each sample. Mutations are color-coded by mutation nomenclature using the key located at the bottom of the figure. Percentages indicate number of samples identified with a mutation of the defined gene. The top barplot has the number of mutations identified in each patient. The right-hand barplot provides the number of mutations in the defined gene which are coloured by mutation nomenclature.

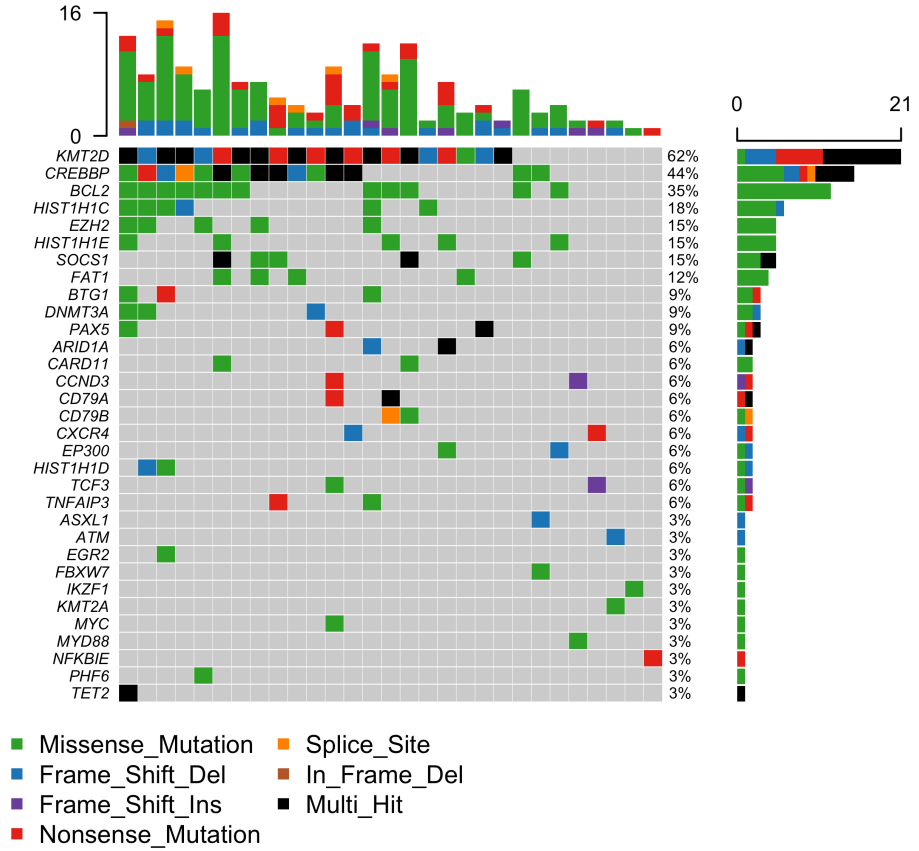


Figure S11. Oncoplot showing the most frequently mutated genes identified in Mantle Cell Lymphoma (MCL) patients present within the validation cohort. Mutations were identified in 29 of the 35 MCL patients included within the final validation cohort of 270 samples. Each column represents one patient while each row represents the absence (grey) or presence (colour) of a mutation within the defined gene in each sample. Mutations are color-coded by mutation nomenclature using the key located at the bottom of the figure. Percentages indicate number of samples identified with a mutation of the defined gene. The top barplot has the number of mutations identified in each patient. The right-hand barplot provides the number of mutations in the defined gene which are coloured by mutation nomenclature.

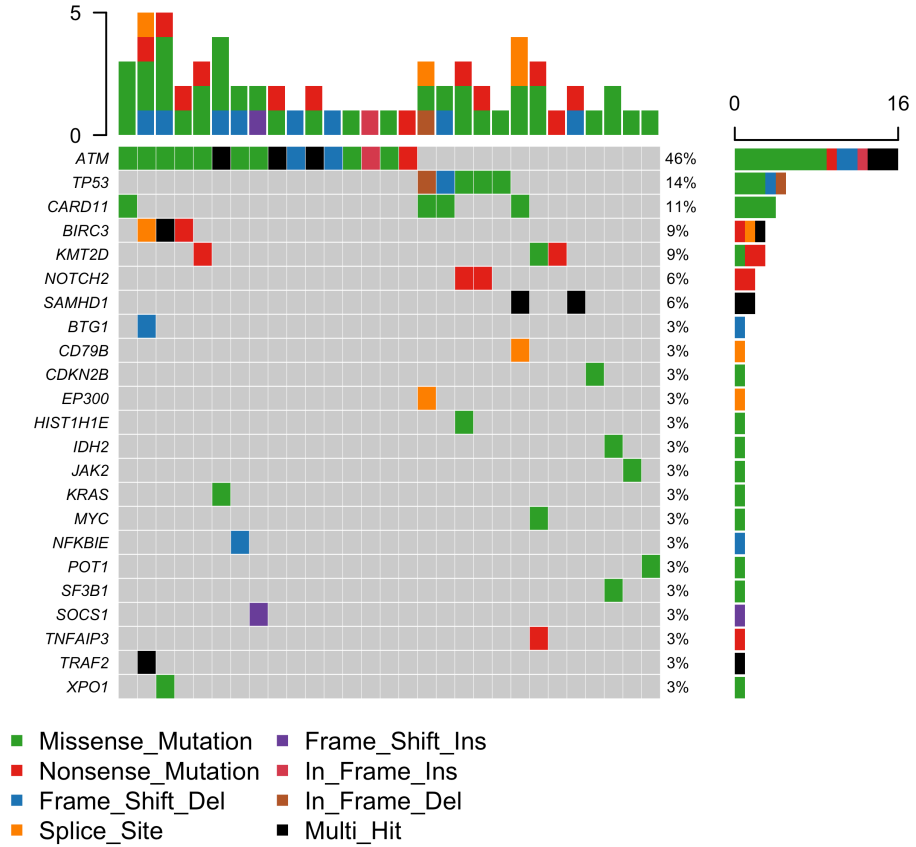


Figure S12. Oncoplot showing the most frequently mutated genes identified in Plasma Cell Myeloma (PCM) patients present within the validation cohort. Mutations were identified in 22 of the 24 T-ALL patients included within the final validation cohort of 270 samples. Each column represents one patient while each row represents the absence (grey) or presence (colour) of a mutation within the defined gene in each sample. Mutations are color-coded by mutation nomenclature using the key located at the bottom of the figure. Percentages indicate number of samples identified with a mutation of the defined gene. The top barplot has the number of mutations identified in each patient. The right-hand barplot provides the number of mutations in the defined gene which are coloured by mutation nomenclature.

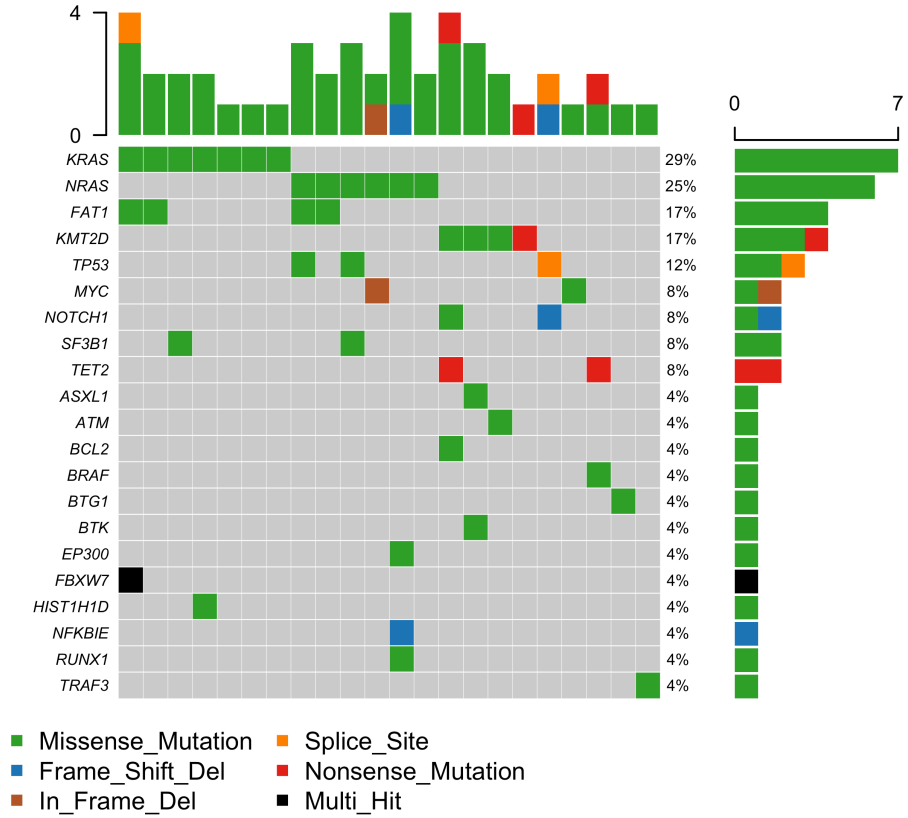


Figure S13. Oncoplot showing the most frequently mutated genes identified in T Acute Lymphoblastic Leukemia (T-ALL) patients present within the validation cohort. Mutations were identified in 37 of the 40 T-ALL patients included within the final validation cohort of 270 samples. Each column represents one patient while each row represents the absence (grey) or presence (colour) of a mutation within the defined gene in each sample. Mutations are color-coded by mutation nomenclature using the key located at the bottom of the figure. Percentages indicate number of samples identified with a mutation of the defined gene. The top barplot has the number of mutations identified in each patient. The right-hand barplot provides the number of mutations in the defined gene which are coloured by mutation nomenclature.