Supplemental Methods

Figure S1. Schematic of the CAR19 transposon cassette. The **pVAX1PB CAR19h28TM41BBz** transposon plasmid contains the following functional DNA sequences making up the transposon cassette: the **hEF1alpha promoter** (GenBank HG530137.1, FJ716125.1)[1,2], the **CAR19h28TM41BBz chimeric antigen receptor** and the **SV40 polyadenylation signal**[3-5]. In addition, the functional sequences are flanked by **chicken β-globin cHS4 insulators** intended to prevent the transgene altering expression of surrounding genes or CAR transgene silencing[6-9], which are in-turn flanked by the **transposon inverted repeats** (GenBank HQ888845.1, EU257621.1, KF658273.1, AB779767.1, AB779766.1) [10-12] [13,14]. IR= inverted repeat, Ins= Insulator, EF1α= human elongation factor-1 alpha promoter, CAR19h28TM41BBz= CD19-specific CAR transgene, SV40= simian virus late polyadenylation *signal.*

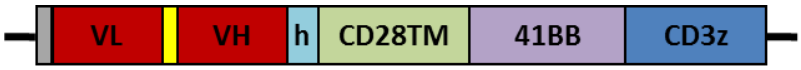

Figure S2. Schematic of CAR19h28TM41BBz Transgene. The **CAR19h28TM41BBz transgene** is a second generation CAR consisting of the CD8 alpha chain leader peptide (amino acids 1-21, UniProtKB reference P01782), a single chain variable fragment (scFv) derived from the CD19-specific antibody, FMC63 [15,16] containing a flexible (G4S)3 linker [17,18]. The scFv is followed by a second (G4S)3 linker and an extracellular spacer consisting of the hinge domain of the human IgG1 heavy chain (amino acids 99-110, UniProtKB reference P01857) with a C to P

amino acid substitution in the hinge domain at position 103 of the native protein. The spacer is followed by the transmembrane domain of the CD28 molecule (amino acids 153-179, UniProtKB reference P10747) and then the intracellular domains of 4-1BB (CD137- amino acids 214-255, UniProtKB reference Q07011) and CD3zeta (amino acids 52-164, UniProtKB reference P230963). Grey bar= CD8 alpha leader peptide sequence; VL= FMC63 antibody light chain variable region sequence; VH= FMC63 antibody heavy chain variable region sequence; h= IgG1 hinge; CD28TM= transmembrane domain of the CD28 co-stimulatory molecule; 41BB= intracellular domain of 4-1BB (CD137), CD3z= intracellular domain of CD3 zeta chain.

Annotated Sequence of the pVAX1PB CAR19h28TM41BBz Plasmid

5' IR	**201-139 (c)**
Insulator	377-608
EF1alpha Promoter	622-1167
CAR19	1194-2657
SV40 polyA	2779-2909
Insulator	3388-3158
3' IR	**3663-3697**
Kanamycin R	4271-5065
pUC ori	5364-6037

Non-coding

GACTCTTCGCGGCGCGCCTCGTTCATTCACGTTTTTGAACCCGTGGAGGACGGGCAG
ACTCGCGGTGCAAATGTGTTTTACAGCGTGATGGAGCAGATGAAGATGCTCGACAC
GCTGCAGAACACGCAGCTAGATTAA**CCCTAGAAAGATAATCATATTGTGACGTAC**
GTTAAAGATAATCATGTGTAAAATTGACGCATGTGTTTTATCGGTCTGTATATCGA
GGTTTATTTATTAATTTGAATAGATATTAAGTTTTATTATATTTACACTTACATACTA
ATAATAAATTCAACAAACAATTTATTTATGTTTATTTATTTATTAAAAAAAACAAAA
ACTCAAAATTTCTTCTATAAAGTAACAAAACTTTTATGAGGGACAGCCCCCCCCCAA
AGCCCCCAGGGATGTAATTACGTCCCTCCCCCGCTAGGGGGCAGCAGCGAGCCGCC
CGGGGCTCCGCTCCGGTCCGGCGCTCCCCCCGCATCCCCGAGCCGGCAGCGTGCGG
GGACAGCCCGGGCACGGGGAAGGTGGCACGGGATCGCTTTCCTCTGAACGCTTCTC
GCTGCTCTTTGAGCCTGCAGACACCTGGGGGGATACGGGGAAAAGGCCTCCACGGC
CAAGGATCTGCGATCGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCCACA
GTCCCCGAGAAGTTGGGGGGAGGGGTCGGCAATTGAACGGGTGCCTAGAGAAGGTG
GCGCGGGGTAAACTGGGAAAGTGATGTCGTGTACTGGCTCCGCCTTTTTCCCGAGGG
TGGGGGAGAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTCGCAACGG
GTTTGCCGCCAGAACACAGCTGAAGCTTCGAGGGGCTCGCATCTCTCCTTCACGCGC
CCGCCGCCCTACCTGAGGCCGCCATCCACGCCGGTTGAGTCGCGTTCTGCCGCCTCC
CGCCTGTGGTGCCTCCTGAACTGCGTCCGCCGTCTAGGTAAGTTTAAAGCTCAGGTC
GAGACCGGGCCTTTGTCCGGCGCTCCCTTGGAGCCTACCTAGACTCAGCCGGCTCTC
CACGCTTTGCCTGACCCTGCTTGCTCAACTCTACGTCTTTGTTTCGTTTTCTGTTCTGC
GCCGTTACAGATCCAAGCTGTGACCGGCGCCTACTCTAGAGCTAGCGAATTCGAATG
GCCATGGCCTTACCAGTGACCGCCTTGCTCCTGCCGCTGGCCTTGCTGCTCCACGCC
GCCAGGCCGGACATCCAGATGACACAGACTACATCCTCCCTGTCTGCCTCTCTGGGA

GACAGAGTCACCATCAGTTGCAGGGCAAGTCAGGACATTAGTAAATATTTAAATTG
GTATCAGCAGAAACCAGATGGAACTGTTAAACTCCTGATCTACCATACATCAAGATT
ACACTCAGGAGTCCCATCAAGGTTCAGTGGCAGTGGGTCTGGAACAGATTATTCTCT
CACCATTAGCAACCTGGAGCAAGAAGATATTGCCACTTACTTTTGCCAACAGGGTAA
TACGCTTCCGTACACGTTCGGAGGGGGGACTAAGTTGGAAATAACACGGGCTGATG
CTGCACCAACTGTATCCATCTTCCCACCATCCAGTAATGGTGGCGGTGGCTCGGGCG
GTGGTGGGTCGGGTGGCGGCGGATCTGAGGTGAAACTGCAGGAGTCAGGACCTGGC
CTGGTGGCGCCCTCACAGAGCCTGTCCGTCACATGCACTGTCTCAGGGGTCTCATTA
CCCGACTATGGTGTAAGCTGGATTCGCCAGCCTCCACGAAAGGGTCTGGAGTGGCTG
GGAGTAATATGGGGTAGTGAAACCACATACTATAATTCAGCTCTCAAATCCAGACTG
ACCATCATCAAGGACAACTCCAAGAGCCAAGTTTTCTTAAAAATGAACAGTCTGCA
AACTGATGACACAGCCATTTACTACTGTGCCAAACATTATTACTACGGTGGTAGCTA
TGCTATGGACTACTGGGGTCAAGGAACCTCAGTCACCGTCTCCTCAGGTGGAGGCGG
GTCTGGGGGCGGAGGTTCAGGCGGGGGTGGTTCCGAGCCCAAATCTCCTGACAAAA
CTCACACATGCCCATTTTGGGTGCTGGTGGTGGTTGGTGGAGTCCTGGCTTGCTATA
GCTTGCTAGTAACAGTGGCCTTTATTATTTTCTGGGTGAAACGGGGCAGAAAGAAAC
TCCTGTATATATTCAAACAACCATTTATGAGACCAGTACAAACTACTCAAGAGGAAG
ATGGCTGTAGCTGCCGATTTCCAGAAGAAGAAGAAGGAGGATGTGAACTGAGAGTG
AAGTTCAGCAGGAGCGCAGACGCCCCCGCGTACCAGCAGGGCCAGAACCAGCTCTA
TAACGAGCTCAATCTAGGACGAAGAGAGGAGTACGATGTTTTGGACAAGAGACGTG
GCCGGGACCCTGAGATGGGGGGAAAGCCGAGAAGGAAGAACCCTCAGGAAGGCCT
GTACAATGAACTGCAGAAAGATAAGATGGCGGAGGCCTACAGTGAGATTGGGATGA
AAGGCGAGCGCCGGAGGGGCAAGGGGCACGATGGCCTTTACCAGGGTCTCAGTACA
GCCACCAAGGACACCTACGACGCCCTTCACATGCAGGCCCTGCCCCCTCGCTAAGTC
GACAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTAT
GTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCGTTAA
CTAAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAAT
TTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCA
ATGTATCTTATCATGTCTGGAATTGACTCAAATGATGTCAATTAGTCTATCAGAAGC
TCATCTGGTCTCCCTTCCGGGGGACAAGACATCCCTGTTTAATATTTAAACAGCAGT
GTTCCCAAACTGGGTTCTTATATCCCTTGCTCTGGTCAACCAGGTTGCAGGGTTTCCT
GTCCTCACAGGAACGAAGTCCCTAAAGAAACAGTGGCAGCCAGGTTTAGCCCCGGA
ATTGACTGGATTCCTTTTTTAGGGCCCATTGGTATGGCTTTTCCCCGTATCCCCCCA
GGTGTCTGCAGGCTCAAAGAGCAGCGAGAAGCGTTCAGAGGAAAGCGATCCCGTGC
CACCTTCCCCGTGCCCGGGCTGTCCCCGCACGCTGCCGGCTCGGGGATGCGGGGGG
AGCGCCGGACCGGAGCGGAGCCCCGGGCGGCTCGCTGCTGCCCCCTAGCGGGGGAG
GGACGTAATTACATCCCTGGGGGCTTTGGGGGGGGGCTGTCCCTGATATCTATAACA
AGAAAATATATATATAATAAGTTATCACGTAAGTAGAACATGAAATAACAATATAA
TTATCGTATGAGTTAAATCTTAAAAGTCACGTAAAAGATAATCATGCGTCATTTTGA
CTCACGCGGTCGTTATAGTTCAAAATCAGTGACACTTACCGCATTGACAAGCACGCC
TCACGGGAGCTCCAAGCGGCGACTGAGATGTCCTAAATGCACAGCGACGGATTCGC
GCTATTTAGAAAGAGAGAGCAATATTTCAAGAATG**CATGCGTCAATTTTACGCAG**
ACTATCTTTCTAGGGTTAATCTAGCTGCATCAGGATCATATCGTCGGGTCTTTTTTC
CGGCTCAGTCATCGCCCAAGCTGGCGCTATCTGGGCATCGGGGAGGAAGAAGCCCG
TGCCTTTTCCCGCGAGGTTGAAGCGGCATGGAAAGAGTTTGCCGAGGATGACTGCTG
CTGCATTGACGTTGAGCGAAAACGCACGTTTACCATGATGATTCGGGAAGGTGTGGC

CATGCACGCCTTTAACGGTGAACTGTTCGTTCAGGCCACCTGGGATACCAGTTCGTC
GCGGCTTTTCCGGACACAGTTCCGGATGGTCAGCCCGAAGCGCATCAGCAACCCGA
ACAATACCGGCGACAGCCGGAACTGCCGTGCCGGTGTGCAGATTAATGACAGCGGT
GCGGCGCTGGGATATTACGTCAGCGAGGACGGGTATCCTGGCTGGATGCCGCAGAA
ATGGACATGGATACCCCGTGAGTTACCCGGCGGGCGCGCGCCCTGCAAAGTAAACT
GGATGGCTTTCTTGCCGCCAAGGATCTGATGGCGCAGGGGATCAAGCTCTGATCAAG
AGACAGGATGAGGATCGTTTCGCATGATTGAACAAGATGGATTGCACGCAGGTTCT
CCGGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACAGACAATCGG
CTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGTTCTTTTTGTC
AAGACCGACCTGTCCGGTGCCCTGAATGAACTGCAAGACGAGGCAGCGCGGCTATC
GTGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTGTCACTGAAGC
GGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTC
ACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATA
CGCTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGA
GCACGTACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGGACGAAGAGCA
TCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGAGCATGCCCGACG
GCGAGGATCTCGTCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGGAAA
ATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATC
AGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCT
GACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCT
ATCGCCTTCTTGACGAGTTCTTCTGAATTATTAACGCTTACAATTTCCTGATGCGGTA
TTTTCTCCTTACGCATCTGTGCGGTATTTCACACCGCATCAGGTGGCACTTTTCGGGG
AAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCG
CTCATGAGACAATAACCCTGATAAATGCTTCAATAATAGCACGTGCTAAAACTTCAT
TTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATC
CCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGA
TCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCAC
CGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGG
TAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCCGTAGT
TAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCC
TGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAA
GACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACA
CAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCT
ATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGC
GGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGT
ATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATG
CTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGT
TCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTT

Detailed CAR T-cell Production Protocol

This study was conducted in accordance with the Declaration of Helsinki. Donor peripheral blood mononuclear cells (PBMCs) were isolated from whole blood by Ficoll gradient centrifugation and rested in AIM-V (Thermo-Fisher) for 2-24 hours. Rested cells were washed in phosphate buffered saline and suspended in Neon Resuspension Buffer T at a concentration of $4x10^7$/ml with 20 ug/ml piggyBac transposase mRNA and 50 ug/ml of the CAR transposon plasmid. Up to $50x10^7$ cells were electroporated in 100ul aliquots using the Neon transfection system (Thermo-Fisher) with 1 pulse at 2400v for 20ms. Transduced cells were rested in media overnight, harvested, counted, suspended in AIM-V supplemented with 10% autologous serum (complete AIM-V) at $1x10^6$/ml and transferred to a G-Rex10 culture flask (Wilson Wolf Manufacturing). Irradiated (30 Gy) donor PBMC feeder cells were added at an effector:feeder ratio of 1:2 and supplemented with interleukin-15 at a final concentration of 200 IU/ml. Interleukin-15 was replenished every 2-3 days and irradiated PBMCs every 7 days. CAR T cell manufacturing was completed within 15 days of electroporation for both products, and products were cryopreserved at 1×10^7 cells/ml in 70% saline (Baxter), 20% Albumex20 (CSL Behring) and 10% dimethyl sulfoxide (WAK-Chemie Medical).

Flow cytometry detection of CAR T-cells

Cells were extracted from biopsy specimens using the Miltenyi gentleMACS Dissociator (Miltenyi Biotec, Bergisch Gladbach, Germany). Mononuclear cells were isolated from the peripheral blood and ascites by ficoll gradient centrifugation. Live cells were identified by exclusion of 7AAD (eBioscience, Thermo Fisher) and stained using fluorophore conjugated antibodies (BD Biosciences Franklin Lakes, NJ, Table S1). The FMC63-specific anti-idiotype

antibody (a gift from Prof Laurence Cooper, MD Anderson[19]) was conjugated using the Alexa Fluor 647 labelling kit (Invitrogen, Thermo Fisher Scientific, Waltham, MA). Immunophenotypic analysis of malignant CAR T-cells was performed on FACS Canto or Fortessa flow cytometers (BD Biosciences) and analyzed using FCS Express 6 (De Novo Software, Pasadena, CA).

Abseq immunophenotypic analysis

Aliquots of one million cells were incubated with human Fc Block (BD Pharmingen) followed by a cocktail of antibody-oligonucleotide conjugates (AbO) (BD Biosciences) (Table S1) for 45 minutes. Malignant cells were loaded directly onto a 10X Chromium. CAR T-cell product was sorted for viable $CD3^{+}CAR19^{+}$ cells on a BD FACSAria IIIu (BD Biosciences) and loaded onto a 10X Chromium platform.

Library preparation was performed using Chromium Single Cell 3' Reagents Kits v2 (10X Genomics, PN-120267) and v3 (10X Genomics, PN-1000092). An AbO-specific PCR1 primer was used to amplify the AbSeq library which were separated following cDNA amplification by SPRIselect size selection (Beckman Coulter). Sample index PCRs were performed according to manufacturers' instructions, final libraries were pooled and sequenced together on NextSeq500 or NovaSeq6000 platforms (Illumina). Library quality control was performed using a combination of agarose gel electrophoresis, Quant-iT PicoGreen (Invitrogen) and LabChip GX (Perkin Elmer).

The Cellranger pipeline was applied to obtain the protein expression (AbSeq) matrix. This resulted in a total of 3173 cells for the product and 4362 cells for the tumor sample. All plots

with AbSeq expression values show the result of a centered-log-ratio normalization applied to the raw counts.

Phosphoflow

$5x10^5$ rested cells were fixed and permeabilized with PhosFlow Lyse/Fix Buffer and PhosFlow Perm Buffer III (BD Biosciences). Cells were stained with anti-CD3-APC, anti-CD4-PECy7, anti-CD8-Pacific Blue, anti-CD3zeta(pY142)-PE, anti-AKT(pS473)-AF488 and anti-ZAP70(pY319)-AF488 (BD Biosciences, Table S1). Data were collected on the LSRFortessa flow cytometer (BD Biosciences) and analyzed using FCS Express 6. To control for autofluorescence, an FMO sample was stained with surface antibodies.

Flow cytometry cell sorting

CD4+, CD8+ and CAR+ T-cells were selected from healthy donor peripheral blood, the patient's CAR T-cell product, enriched peripheral blood CAR T-cells and malignant CAR T-cell tumor by flow cytometry sorting on the FACSAria III sorter (BD Biosciences). Untransduced T-cells from the peripheral blood of healthy donors were identified as lineage negative (CD19neg, CD56neg, CD14neg) and CD4+ or CD8+. CAR+ T-cells were identified as CD4+CAR+ or CD8+CAR+.

Table S1- Abseq and fluorophore conjugated antibodies used for phenotypic analysis, phosphoflow and flow cytometry sorting- all from BD Bioscience

Target	Clone	Barcode ID (precommercial in brackets)	Catalogue number
CD127	HIL-7R-M21	AHS0028	940012
CD14	MϕP9	AHS0037	940005
CD161	DX12	AHS0002	940070
CD183	1C6/CXCR3	AHS0031	940030
CD185 (CXCR5)	RF8B2	AHS0039	940042
CD19	HIB19	(v2a_Hs_0030)	
CD194	1G1	AHS0038	940047
CD196 (CCR6)	11A9	AHS0034	940033
CD197 (CCR7)	150503	(v2a_Hs_007)	
CD25	M-A251	(v2a_Hs_0026)	
CD27	M-T271	AHS0025	940018
CD28	CD28.2	AHS0024	940017
CD3	SK7	AHS0033	940000
CD38	HIT2	AHS0022	940013
CD4	SK3	AHS0032	940001
CD45RA	HI100	AHS0009 (v2a_Hs_0029)	940011
CD45RO	UCHL1	AHS0036	940022
CD8	RPA-T8	AHS0027	940003
CD95	DX2	AHS0023	940037
HLA-DR	G46-6	(v2a_Hs_0035)	
PD1 (CD279)	MIH4	(v2a_Hs_0014)	
CD10	HI10a	AHS0051	940045
CD11b	M1/70	AHS0005	940008
CD11c	B-ly6	AHS0056	940024
CD137	4B4-1	AHS0003	940055
CD141	1A4	AHS0083	940079
CD154	TRAP1	AHS0077	940053
CD16	3G8	AHS0053	940006
CD20	2H7	AHS0008	940016
CD206	19.2	AHS0072	940068
CD21	B-ly4	AHS0074	940048
CD274 (B7-H1)	MIH1	AHS0004	940035
CD40	5C3	AHS0117	940049
CD56	NCAM16.2	AHS0019	940007
CD80 (B7-1)	L307.4	AHS0046	
CD86 (B7-2)	2331 (FUN-1)	AHS0057	940025
IgD	IA6-2	AHS0058	940026

IgG	G18-145	AHS0059	940027
LAG-3 (CD223)	T47-530	AHS0018	940080
TIM-3 (CD366)	7D3	AHS0016	940066
CD19	SJ25C1	AHS0030	940004
CD25	2A3	AHS0026	940009
CD197 (CCR7)	3D12	AHS0007	940014
PD1 (CD279)	EH12.1	AHS0014	940015
Target	**Clone**	**Fluorophore**	**Catalogue number**
CD2	RPA-2.10	FITC	555326
CD3	UCHT1	BV480	566105
CD3	UCHT1	Pacific Blue	558117
CD3	UCHT1	BV510	563109
CD3	SK7	PECy7	341091
CD4	SK3	APC	340443
CD4	SK3	PE-Cy7	557852
CD4	RPA-T4	BUV395	564724
CD8	RPA-T8	APC-R700	565165
CD8	RPA-T8	Pacific Blue	558207
CD8	SK1	FITC	347313
CD8	SK1	PE	340046
CD14	M5E2	PE-Cy7	557742
CD19	4G7	PerCP	347544
CD19	HIB19	PE-Cy5	555414
CD56	NCAM16.2	FITC	340410
CD3ζ-pY142	K25-407.69	PE	558448
ZAP70-pY319	17A/P-ZAP70	AF488	557818
SRC-pY418	K98-37	PE	560094
Akt-pS473	M89-61	AF488	560404

T-Cell receptor clonality analysis

Multiplex PCR was performed using BIOMED-2 primers targeting the V and J regions of the T-cell receptor gamma (TCRG) with the Invivoscribe Identiclone TCRG Gene Clonality Assay, followed by fragment analysis by capillary electrophoresis.

T-cell receptor beta (TRB) loci deep amplicon sequencing was performed using LymphoTrack TRB (Invivoscribe, San Diego, CA). Sequence assembly from FASTQs, annotation and error

correction was performed by MiXCR[20] with diversity assessment, VDJ family usage analysis performed using VDJtools (ver 1.1.9).[21]

Hybridization-based next generation sequencing analysis

Extracted DNA was analysed for the presence of mutations in a panel of genes commonly mutated in haematological malignancies as previously described.[22] Indexed libraries were sequenced on an Illumina NextSeq (paired-end 75 bp reads). After base calling and de-multiplexing, a Seqliner-framework analysis pipeline was used to align reads to the human reference genome (GRCh37 assembly) using BWA-MEM, followed by marking of duplicate reads, base quality score recalibration, local indel realignment and variant calling using GATK Haplotype Caller (https://software.broadinstitute.org/gatk/). Aligned sequence data was processed through a dedicated bioinformatics pipeline which included variant calling with GATK4/Mutect2 (https://software.broadinstitute.org/gatk/) in order to improve detection of low level acquired variants. Genomic copy number analysis was performed using on and off target reads from this hybridization-based NGS panel.[22,23]

Transgene copy number analysis

Droplet digital PCR was conducted on genomic DNA[22] using a custom Bio-Rad PrimerPCR ddPCR primer/probe specific for the CAR gene at the junction of CD28 transmembrane and 4-1BB intracellular sequences and the Bio-Rad Human RPP30. Copy number assay primer/probe (Bio-Rad, Hercules, CA) CAR and RPP30 gene copies per microliter were absolutely quantified utilising QuantaSoft Software v1.7.4 (Bio-Rad).

CAR copies per cell (CCPC) were calculated using the equation:

CCPC = 2 x (CAR copies/ul)/(RPP30 copies/ul)

Whole genome sequencing analysis

Sequencing libraries were prepared from 100 ng of genomic DNA using the NEBNext Ultra II DNA Library Prep Kit for Illumina (New England Biolabs, Ipswich, MA, USA) with a double-sided SPRI bead size selection and 6 cycles of PCR. Whole genome sequencing was performed by the Australian Genome Research Facility (AGRF) on the NovaSeq 6000 platform (Illumina, San Diego CA, USA) with a 2 × 150 bp read length using NovaSeq S4 300-cycle reagent kit. Each sample was sequenced to a minimum of 30-fold base coverage (30X) per sample with a total output of a minimum 100GB. Primary data analysis was performed in real time by the NovaSeq Control Software v1.6.0 and Real Time Analysis (RTA) v3.4.4. Demultiplexed fastq files were generated using Illumina's bcl2fastq software (version 2.20.0.422).

A modified human reference was generated by concatenating GRCh38 with the CAR19 sequence as an additional contig. Reads were aligned to the modified reference using Sentieon[24] with default settings. Structural Variants (SVs) were called using the SURVIVOR pipeline[25,26]. Alignments were processed through Individual SV callers Delly (v0.8.2)[27], Lumpy (v0.2.13)[28], and Manta (v1.6.0)[29]. The consensus among these callers was determined by SURVIVOR (v1.0.7) using the parameters "1000 2 1 0 0". Reads with low mapping quality (MQ < 4 per sample) were identified using samtools view and samtools depth and then clustered together using SURIVOR "bincov" with parameters 10 (window size) and 2 (minimum number of reads). SVs per sample were filtered using SURVIVOR "filter" to obtain SV calls outside of hard to map regions. The SV across the samples were merged using SURVIVOR "merge" with parameters "1000 1 1 0 0 0" to form the final multi-sample VCF file. The retained SVs were

annotated with annotatePeaks.pl from HOMER[30] using the hg38 build. CAR19 insertions in the malignancy detected by mu insertion analysis were confirmed by searching alignment files for paired end split reads between the CAR19 "contig" and regions of the autosomes, excluding calls that did not indicate a full-length insertion of CAR19.

Whole transcriptome analysis

Low input RNA-seq libraries were prepared using the SMARTer Stranded Total RNA-Seq kit v2 - Pico input mammalian (Takara Bio USA, Inc, CA) by AGRF and sequenced at 1 × 100 bp SE on an Illumina Novaseq 6000 SP 100 lane.

Library sequencing quality was determined using FastQC and FastQ0 Screen, Illumina adaptor sequence and reads of < 20 base pairs were trimmed using Trim Galore! (BabrahamBioinformatics: www.bioinformatics.babraham.ac.uk/). Reads were aligned to a custom genome (hg38 and CART-insert) using STAR[30] with Refseq annotations as a guide. Read counts data corresponding to Refseq transcript annotations were generated using HTSeq.[31] Alignments and transcriptional shadows were visualised using IGV (v2.72).[32] Genome position analyses were performed using genomic ranges[33] based on RefSeq annotations imported with 'rtracklayer'[34] and gene lists were annotated with KEGG pathways (adjusted p value <0.05) using the clusterProfiler package.[35]

All analyses were performed in the R Statistical Environment (http://www.Rproject.org) with tidyverse.[36] Counts data were background corrected and normalized for library size using edgeR[37] then transformed using voom[38] for differential expression analysis using LIMMA[39] (4-fold change, sdj.p <0.01).

In silico analysis of transcriptional readthrough mediated gene expression

In silico impact was estimated for genic insertion sites using RefSeq gene annotations. Briefly, where exons within a transcriptional shadow appeared overexpressed, if all coding exons were downstream, the ORF was considered intact and the proportion of reads arising due to insertion was estimated by taking the ratio of spliced reads relative to the proportion of spliced reads to the exon coverage. If insertion was within the coding region, the resultant sequence was verified for loss of frame using in silico translation and comparison to wildtype transcript sequence. Contribution of insertion to total gene signal was estimated as above.

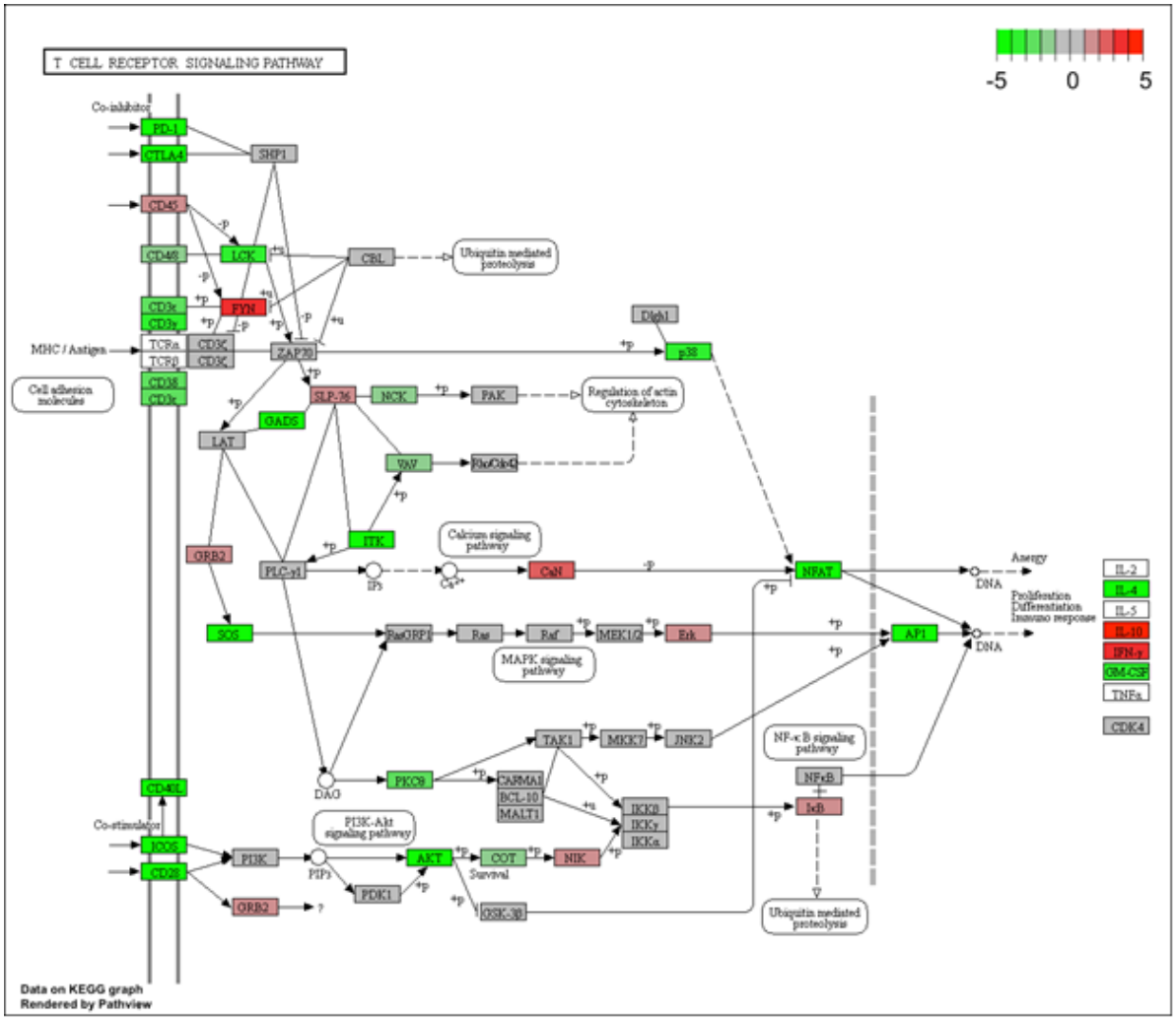

Figure S3. FYN and T-cell receptor activation pathways. KEGG pathway showing predominantly reduced levels of T-cell receptor associated genes and upregulation of FYN in the malignant CAR T-cells compared to untransduced healthy donor controls (green= down-regulated, red=up-regulated).

Table S2. Integration sites within the malignant CAR T-cells and their effect on expression of surrounding genes in Patient 8

	INSERT MAPPING			GENOMIC CONTEXT AND RNA SEQ		
Chr	Insert Site	Strand	Transcriptional shadow (kbp)	Gene same strand within shadow	Observed Log2FC (MalignancyCPM vs ProductCPM)	*in silico* evidence of functional product
chr1	8,3508,456	+	100	Non-coding RNA LOC107985043		Possible increase of uncertain significance
chr6	90,105,820	-	150	BACH2	-0.71	No
chr6	142,807,281	+	100			No
ChrX	65746639	-	30			No

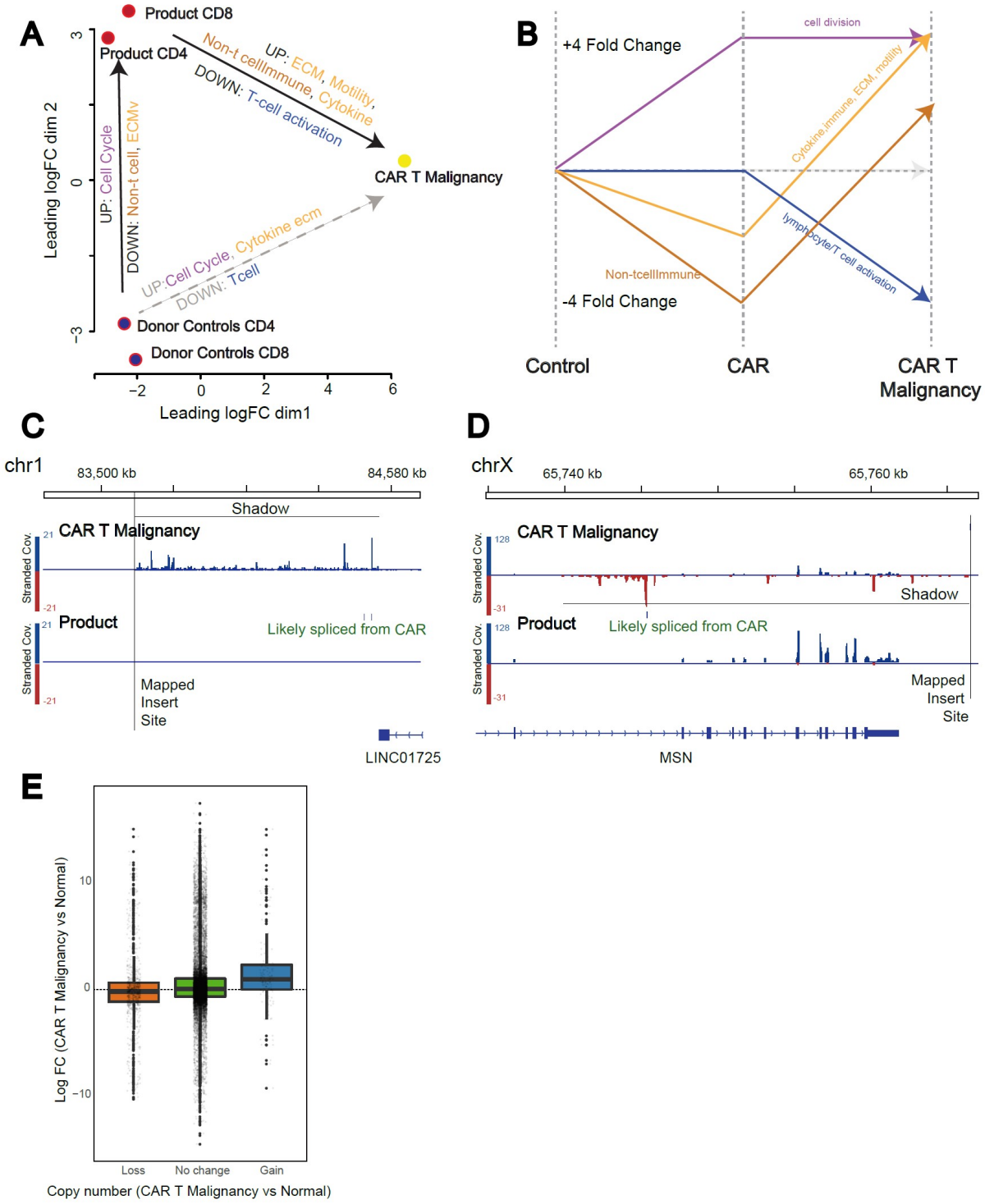

Figure S4. Transcription Analysis of Malignant CAR T-cells Compared to Non-Malignant CAR$^+$ and Untransduced T-cells related to Patient 8. (A) Multidimensional scaling of global

gene expression from malignant CAR T-cells (Malignancy), non-malignant $CD4^+$ and $CD8^+$ CAR T-cells from the product (Product) and untransduced healthy donor $CD4^+$ and $CD8^+$ T-cells (Donor Controls). (B) Line drawing showing 4-fold differential expression of clusters of genes in the malignant CAR T-cells (Malignancy) compared to $CD8^+$ T-cells from the CAR T-cell product (CAR) and untransduced related healthy donor T-cells (Control). (C) Example of transcriptional readthrough/shadow seen in the malignant CAR T-cells at the point of insertion proximal to the non-coding mRNA LOC107985043 producing increased positive strand expression (D) transcriptional readthrough at the point of insertion inverse to the MSN3 gene showing no increase in exonic expression. (E) Correlation between copy number variation and altered gene expression showing correlation between copy number gain and increased gene expression (blue box) in the malignant CAR T-cells (Malignancy) compared to CAR^+CD8^+ T-cells isolated from the product.

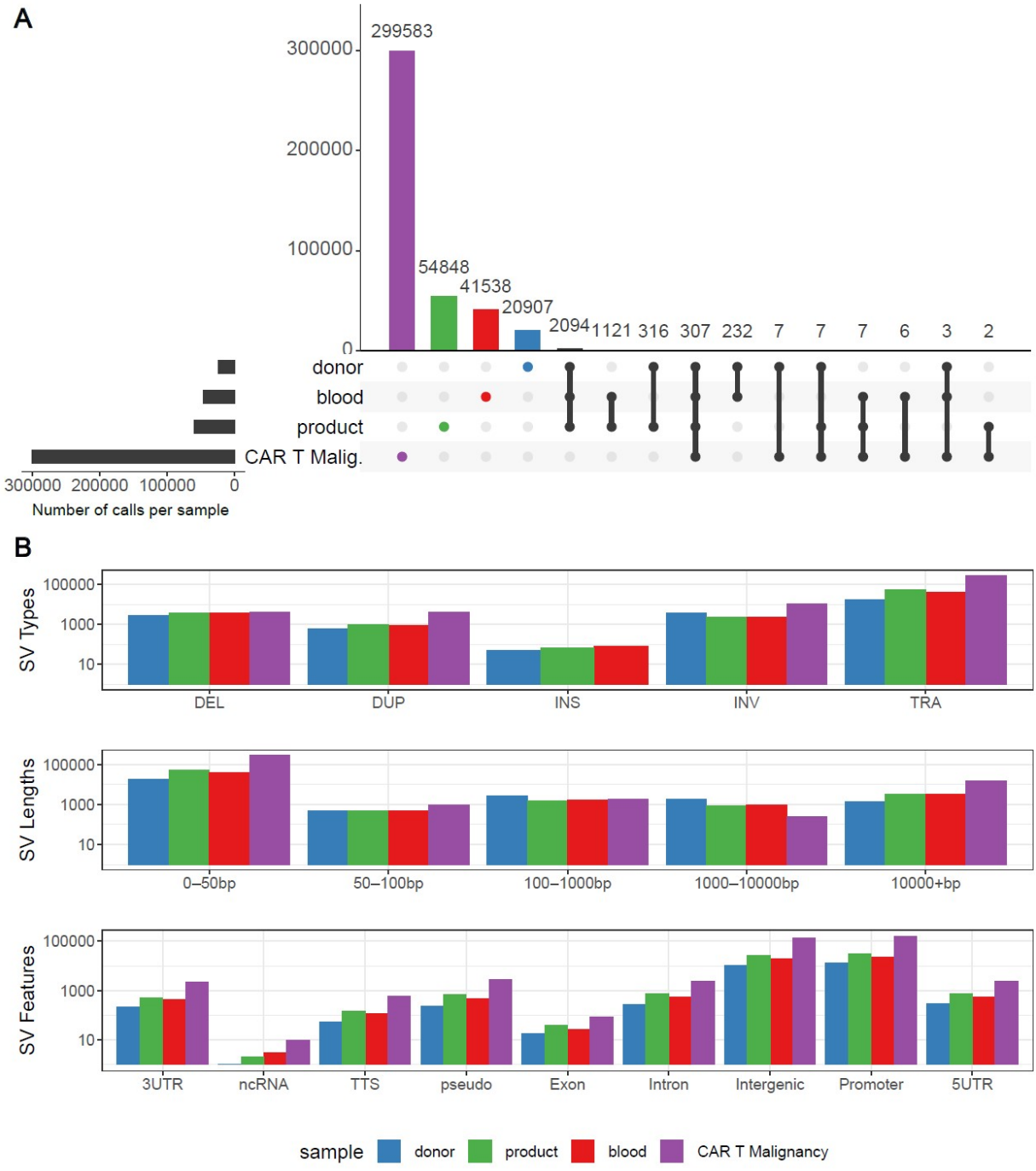

Figure S5. Structural variants across whole genome sequenced samples related to Patient 8. (A) UpSet plot showing the intersection of structural variants (SV) calls across DNA extracted from $CD8^+$ T-cells isolated from untransduced healthy donor T-cells (donor), $CD8^+$ CAR T-cells from the CAR T-cell product (product), expanded from Patient 8's peripheral blood (blood) and

from the malignancy (CAR T Malignancy). SVs detected uniquely within each sample are shaded in colors, maintained throughout the figure. On the bottom left is the number of calls per sample; intersections between samples (black) are shown by connecting bars in lower right plot. (B) Distribution of SVs called uniquely for each sample, including SV type (DEL=deletions; DUP=duplications; INS=insertions; INV=inversions; TRA=translocations; 3UTR= 3' untranslated regions; ncRNA= non-coding RNA; TTS= triplex target DNA sites; pseudo= pseudogenes; 5UTR= 5' untranslated region), the length of each SV, and the genomic annotation per SV.

REFERENCES

1. Uetsuki T, Naito A, Nagata S, Kaziro Y. Isolation and characterization of the human chromosomal gene for polypeptide chain elongation factor-1 alpha. *J Biol Chem*. 1989;264(10):5791-5798.
2. Kim DW, Uetsuki T, Kaziro Y, Yamaguchi N, Sugano S. Use of the human elongation factor 1 alpha promoter as a versatile and efficient expression system. *Gene*. 1990;91(2):217-223.
3. Sweet BH, Hilleman MR. The vacuolating virus, S.V. 40. *Proc Soc Exp Biol Med*. 1960;105:420-427.
4. Fiers W, Contreras R, Haegemann G, et al. Complete nucleotide sequence of SV40 DNA. *Nature*. 1978;273(5658):113-120.
5. Hansen J, Mailand E, Swaminathan KK, Schreiber J, Angelici B, Benenson Y. Transplantation of prokaryotic two-component signaling pathways into mammalian cells. *Proc Natl Acad Sci U S A*. 2014;111(44):15705-15710.
6. Ramezani A, Hawley TS, Hawley RG. Performance- and safety-enhanced lentiviral vectors containing the human interferon-beta scaffold attachment region and the chicken beta-globin insulator. *Blood*. 2003;101(12):4717-4724.
7. Emery DW, Yannaki E, Tubb J, Nishino T, Li Q, Stamatoyannopoulos G. Development of virus vectors for gene therapy of beta chain hemoglobinopathies: flanking with a chromatin insulator reduces gamma-globin gene silencing in vivo. *Blood*. 2002;100(6):2012-2019.
8. Rivella S, Callegari JA, May C, Tan CW, Sadelain M. The cHS4 insulator increases the probability of retroviral expression at random chromosomal integration sites. *J Virol*. 2000;74(10):4679-4687.
9. Sharma N, Hollensen AK, Bak RO, Staunstrup NH, Schroder LD, Mikkelsen JG. The impact of cHS4 insulators on DNA transposon vector mobilization and silencing in retinal pigment epithelium cells. *PLoS One*. 2012;7(10):e48421.
10. Cary LC, Goebel M, Corsaro BG, Wang HG, Rosen E, Fraser MJ. Transposon mutagenesis of baculoviruses: analysis of Trichoplusia ni transposon IFP2 insertions within the FP-locus of nuclear polyhedrosis viruses. *Virology*. 1989;172(1):156-169.
11. Fraser MJ, Cary L, Boonvisudhi K, Wang HG. Assay for movement of Lepidopteran transposon IFP2 in insect cells using a baculovirus genome as a target DNA. *Virology*. 1995;211(2):397-407.
12. Fraser MJ, Ciszczon T, Elick T, Bauser C. Precise excision of TTAA-specific lepidopteran transposons piggyBac (IFP2) and tagalong (TFP3) from the baculovirus genome in cell lines from two species of Lepidoptera. *Insect Mol Biol*. 1996;5(2):141-151.
13. Gohl DM, Silies MA, Gao XJ, et al. A versatile in vivo system for directed dissection of gene expression patterns. *Nat Methods*. 2011;8(3):231-237.
14. Schuldiner O, Berdnik D, Levy JM, et al. piggyBac-based mosaic screen identifies a postmitotic function for cohesin in regulating developmental axon pruning. *Dev Cell*. 2008;14(2):227-238.
15. Zola H, MacArdle PJ, Bradford T, Weedon H, Yasui H, Kurosawa Y. Preparation and characterization of a chimeric CD19 monoclonal antibody. *Immunol Cell Biol*. 1991;69 (Pt 6):411-422.
16. Nicholson IC, Lenton KA, Little DJ, et al. Construction and characterisation of a functional CD19 specific single chain Fv fragment for immunotherapy of B lineage leukaemia and lymphoma. *Mol Immunol*. 1997;34(16-17):1157-1165.
17. Huston JS, Levinson D, Mudgett-Hunter M, et al. Protein engineering of antibody binding sites: recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in Escherichia coli. *Proc Natl Acad Sci U S A*. 1988;85(16):5879-5883.
18. Krebber A, Bornhauser S, Burmester J, et al. Reliable cloning of functional antibody variable domains from hybridomas and spleen cell repertoires employing a reengineered phage display system. *J Immunol Methods*. 1997;201(1):35-55.
19. Jena B, Maiti S, Huls H, et al. Chimeric antigen receptor (CAR)-specific monoclonal antibody to detect CD19-specific T cells in clinical trials. *PLoS One*. 2013;8(3):e57838.
20. Bolotin DA, Poslavsky S, Mitrophanov I, et al. MiXCR: software for comprehensive adaptive immunity profiling. *Nat Methods*. 2015;12(5):380-381.
21. Shugay M, Bagaev DV, Turchaninova MA, et al. VDJtools: Unifying Post-analysis of T Cell Receptor Repertoires. *PLoS Comput Biol*. 2015;11(11):e1004503.
22. Ryland GL, Jones K, Chin M, et al. Novel genomic findings in multiple myeloma identified through routine diagnostic sequencing. *J Clin Pathol*. 2018;71(10):895-899.

23. Markham JF, Yerneni S, Ryland GL, et al. CNspector: a web-based tool for visualisation and clinical diagnosis of copy number variation from next generation sequencing. *Sci Rep*. 2019;9(1):6426.
24. Freed D, Aldana R, Weber JA, Edwards JS. The Sentieon Genomics Tools - A fast and accurate solution to variant calling from next-generation sequence data. *bioRxiv*. 2017.
25. Jeffares DC, Jolly C, Hoti M, et al. Transient structural variations have strong effects on quantitative traits and reproductive isolation in fission yeast. *Nat Commun*. 2017;8:14061.
26. Dennenmoser S, Sedlazeck FJ, Schatz MC, Altmuller J, Zytnicki M, Nolte AW. Genome-wide patterns of transposon proliferation in an evolutionary young hybrid fish. *Mol Ecol*. 2019;28(6):1491-1505.
27. Rausch T, Zichner T, Schlattl A, Stutz AM, Benes V, Korbel JO. DELLY: structural variant discovery by integrated paired-end and split-read analysis. *Bioinformatics*. 2012;28(18):i333-i339.
28. Layer RM, Chiang C, Quinlan AR, Hall IM. LUMPY: a probabilistic framework for structural variant discovery. *Genome Biol*. 2014;15(6):R84.
29. Chen X, Schulz-Trieglaff O, Shaw R, et al. Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. *Bioinformatics*. 2016;32(8):1220-1222.
30. Heinz S, Benner C, Spann N, et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell*. 2010;38(4):576-589.