Title: CBFB-MYH11 fusion neoantigen enables T cell recognition and killing of acute myeloid leukemia

Authors: Melinda A Biernacki, Kimberly A Foster, Kyle B Woodward, Michael E Coon, Carrie Cummings, Tanya M Cunningham, Robson G Dossa, Michelle Brault, Jamie Stokke, Tayla M Olsen, Kelda Gardner, Elihu Estey, Soheil Meshinchi, Anthony Rongvaux, Marie Bleakley

Supplementary Materials

Supplementary Materials and Methods

Table S1. Additional predicted HLA-B*40:01-binding peptides.

Table S2. Sex and HLA-typing of healthy donor PBMC used in immunogenicity screening experiments.

Table S3. HLA-binding predictions for HLA-B*40:01 binding of peptides with sequential alanine substitutions at each position of REEMEVHEL epitope.

Figure S1. CBFB-MYH11/B*40:01 is reproducibly immunogenic in three additional healthy donors.

Figure S2. Gating strategy used for tetramer staining of clones.

Figure S3. REEMEVHEL is presented, but less efficiently, on HLA-B*40:02.

Figure S4. Generation of epitope-bearing AML cell lines.

Figure S5. *CBFB-MYH11* transcript expression significantly decreases after treatment with CBFB-MYH11/B*40:01-specific T cells in vivo.

Figure S6. Persistent AML in xenografts continues to express HLA and present neoantigen.

Figure S7. CBFB-MYH11/B*40:01-specific T cells control circulating AML in a short-term in vivo experiment.

Figure S8. CBFB-MYH11/B*40:01 specific clones do not recognize similar human peptides.

Figure S9. $CD8^+$ T cells isolated from healthy donors by CBFB-MYH11/B*40:01 pHLA tetramer enrichment and sorting are functional.

Figure S10. Additional evaluation of TCR constructs.

Supplementary Materials and Methods

Class I HLA typing of normal donors and cell lines

Genomic DNA was isolated from cell lines, donor PBMC or primary AML (QIAamp DNA Blood Kit: Qiagen) and used for HLA typing by PCR (Allset Gold Low-Resolution ABC Kit: One Lambda) or by next-generation sequencing (NGS) using the ScisGo HLA v6 typing kit (Scisco Genetics Inc., Seattle, WA). Briefly, the NGS employs an amplicon-based 2-stage PCR, followed by sample pooling and sequencing using a MiSeq v2 PE500 (Illumina, San Diego, CA).

HLA binding predictions

HLA binding predictions were made using the Immune Epitope Database (IEDB) analysis resource (http://tools.iedb.org/main/) Consensus tool (72), which combines predictions from NetMHC (4.0) (73-75), SMM (87), and Comblib (77). Predictions were also made using netMHCpan 4.0 (76, 78, 79). All predictions were made on 11/1/2018 using the most recent

versions of the algorithms. The 20 HLA class I molecules evaluated for predicted binding were HLA-A*01:01, -A*02:01, -A*03:01, -A*11:01, -A*24:02, -B*07:02, -B*08:01, -B*15:01, -B*35:01, -B*40:01, -B*44:02, -B*44:03, -C*03:03, -C*03:04, -C*04:01, -C*05:01, -C*06:02, -C*07:01, -C*07:02, and -C*12:03.

Cell lines

Epstein-Barr virus (EBV) transformed lymphoblastoid cell lines (B-LCL) were prepared and maintained in RPMI 1640, 10% fetal calf serum and 1% penicillin/streptomycin (LCL medium), as described (35). Lenti-X 293T cells (Clontech) used in LV production were maintained in DMEM (Invitrogen) supplemented with 10% FCS, 25mM HEPES, 2mM L-glutamine, 1% penicillin/streptomycin, and detached for passage using 0.05% trypsin-EDTA (Invitrogen). T cells were maintained in RPMI 1640, 10% human serum, 1% penicillin/streptomycin, 3mM L-glutamine, and 50μM β-mercaptoethanol (CTL medium). NB-4 and ME-1 AML cell lines were purchased from Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ) and maintained in LCL medium with 20% FCS. NB-4 was confirmed to be genotypically positive for HLA-B*40:01 by HLA typing and was transduced with the *CBFB-MYH11* fusion sequence and an RQR8 tag (42, 88) for selection. ME-1 expresses the type A *CBFB-MYH11* fusion sequence (89), and was transduced with the restricting HLA-B*40:01 allele and an RQR8 tag.

*Generation of CBFB-MYH11$^+$ B*40:01$^+$ AML cell lines*

NB-4 cells were transduced with the *CBFB-MYH11* fusion sequence, specifically the

published nucleotide sequence encoding the type A fusion (16, 90). An epitope-based marker gene, *RQR8* (42, 88), was included upstream of the fusion in the construct for tracking and selection of transduced cells. To ensure coordinated gene expression, the transgene components were separated by 2A elements from the porcine teschovirus (P2A). The transgenes were codon-optimized and synthesized by GeneArt (Life Technologies), cloned into the LV vector pRRLSIN.cPPT.MSCV.WPRE by restriction digestion and ligation, and confirmed by Sanger sequencing.

ME-1 cells were transduced to express the restricting HLA-B*40:01 molecule. The sequence for *HLA-B*40:01:02* was obtained from IPD-IMGT/HLA (https://www.ebi.ac.uk/cgi-bin/ipd/imgt/hla/get_allele.cgi?B*40:01:02:01) and synthesized as a synthetic nucleotide block (GeneArt, Life Technologies), transferred into a Topo vector (LifeTechnologies), and then cloned into the LV vector pRRLSIN.cPPT.MSCV.WPRE with upstream RQR8 tag as described above.

LV transduction of AML cell lines and T cells

LentiX-293T cells (Clontech) were transfected with the LV backbone plasmids along with PAX2 and VSVg packaging plasmids using the CalPhos transfection system (Clontech) per manufacturer's protocol. Virus particles were harvested after 48 hours and filtered through a 0.45 μm filter prior to use. AML cell lines were transduced by the addition of LV supernatant and 1μg/mL of polybrene followed by 90 minutes of centrifugation at 800xg and returned to the incubator at 37°C, 5% CO_2. Transduction efficiency of AML cell lines and $CD4^+$ T cells was assessed by flow cytometry 72 hours after spin inoculation based on staining for the CD34

epitope (Q component of RQR8) or CD20 epitope (R component of RQR8). Transduced cell lines were then enriched to >95% purity by flow sorting, expanded, and cryopreserved until use.

^{51}Chromium-release cytotoxicity assays (CRA)

Cytotoxicity was measured in short-term (4-hour) assays using ^{51}Cr-labeled target cells. Briefly, target cells were labeled with ^{51}Cr overnight (cell lines) or for 6 hours (primary leukemia) at 37°C and 5% CO_2. Effector cells were added to labeled target cells, and incubated for 4 hours. After co-incubation, supernatant was harvested for γ-counting. Spontaneous release (SR) in counts per minute (CPM) was determined from target cells incubated with media alone. Maximal release (MR) was determined from target cells incubated with detergent. The calculation for percent specific lysis by effector cells (T cell clones or TCR-transduced T cells) in experimental wells was performed using the standard formula (91):

$$\%\ lysis = ((average\ CPM_{EXPERIMENTAL})-(average\ CPM_{SR}))/ ((average\ CPM_{MR})-(average\ CPM_{SR}))$$

Targets used included autologous and allogeneic LCL with and without peptides in varying concentrations. For primary AML targets, leukemia cells were thawed, washed, suspended in LCL medium supplemented with 500 U/mL interferon-γ, and incubated for 24 hours at 37°C prior to CRA initiation. All targets were washed to remove excess ^{51}Cr and residual peptide or interferon-γ before co-culture initiation. For peptide titration assays, unconstrained 4-parameter logistic curves were fitted to the CRA data using the R package nplr (92) and used to determine IC50.

Flow cytometry-based cytotoxicity assays

Cytotoxicity of clones against AML cell lines was measured in co-culture up to 120 hours by assessing survival of targets using flow cytometry. Briefly, effector and targets cells were plated in a 1:1 E:T ratio (5 x 10^5 each) in six replicate wells in LCL medium in a 96 well plate, and co-cultured at 37°C in 5% CO_2. Wells with only effectors and only targets were included as controls. At various timepoints, the cultures were centrifuged and the cell pellets stained with antibodies against CD33, CD8, DAPI (0.002 µg/mL; Sigma-Aldrich), and, in some assays, CD34 and/or CD20 mAb to distinguish target cells expressing transgene constructs. Fluorescent CountBright counting beads (Invitrogen) were used to calculate absolute numbers of live (DAPI negative) target cells. Percent survival was calculated as (absolute number of live targets with effector)/(absolute number of live targets without effector)*100.

In cytotoxic degranulation (CD107a) assays, effector T cells and stimulator cells (AML cell lines and primary AML) were washed and plated in a 1:1 or 1:2 E:T ratio in LCL medium with GolgiStop transport inhibitor (BD Biosciences) and PE-conjugated anti-CD107a mAb. Primary AML cells were pre-incubated in interferon-γ and then washed as described above. Effectors and targets were co-incubated for 5 hours at 37°C. Cells were then washed and stained with CBFB-MYH11/B*40:01 tetramer as well as mAb against CD8, CD33, CD4, CD56, and DAPI.

Reporting on animal studies

Mice used in experiments were exclusively *M-CSF*$^{h/h}$ *IL-3/GM-CSF*$^{h/h}$ *SIRPA*$^{h/h}$ *TPO*$^{h/h}$ *RAG2*$^{-/-}$ *IL2RG*$^{-/-}$ genotype, MISTRG strain, sourced from Dr.

Rongvaux's laboratory (81, 93), backcrosses 129 x Balb/c N2-N3. Animals were both males and females, randomized across groups. Mice were injected with AML at day 1-3 after birth and received T cell injections at 13 (**Figure 5**) or 21 (**Figure S7**) weeks of age.

Terminal tissue harvesting

Marrow cells were harvested by flushing bilateral femurs and tibias from each animal. Spleen cells in the second experiment were harvested by mechanical dissociation. Harvested cells were passed through a 70 μm filter and subjected to ACK lysis followed by washing in RPMI and PBS. Viable cell counts from tissues were determined by standard Trypan blue staining and counting. In the second murine experiment, values for total white blood cell count per μL on terminal blood samples were obtained using a Heska Element HT5 Veterinary Hematology Analyzer.

Flow cytometry evaluation of MISTRG AML xenografts

Peripheral blood was collected from mice at timepoints by retro-orbital bleed. Red blood cells were lysed using ACK lysing buffer (Gibco). Cells were washed and centrifuged and the cell pellets were then stained with antibodies against murine CD45 (mCD45), human CD45 (hCD45), human CD33, human CD3, and DAPI. Gating strategy is shown in **Figure 5B**. The percent human cells was calculated from viable single cells as %hCD45 ÷ (%mCD45+%hCD45) x 100.

RNA Extraction and rt-QPCR

RNA was extracted from mouse peripheral blood using the RNeasy Mini Kit (QIAGEN) and quantitated on a Nanodrop 1000 spectrophotometer (ThermoFisher). 500 ng of RNA was reverse-transcribed into cDNA using SuperScript IV VILO Master Mix (ThermoFisher) following the manufacturer's protocol (20 µL reactions). The subsequent cDNA volume was increased to 50 µL with molecular grade water after transcription. All QPCR reactions were run in duplicate on a CFX 384 Real Time System (Bio-Rad). *CBFB-MYH11* typeA QPCR was performed using the Ipsogen *CBFB-MYH11* type A Kit (QIAGEN) following the manufacturer's protocol. Total reaction volumes were scaled back to 20 µL and 5 µL of cDNA was used in each reaction. Mouse CD45 rt-QPCR was performed using the *Ptprc* QPCR assay from Integrated DNA Technologies (Assay ID #Mm.PT.58.7583849). The assay was resuspended at a 20x concentration and the QPCR reactions were run using TaqMan Universal PCR Master Mix (ThermoFisher). 5 µL of cDNA were used in a total 20 µL reaction volume. rt-QPCR was performed only on samples from which RNA yield was adequate to perform both CBFB-MYH11 and *Ptprc* assays. Normalized *CBFB-MYH11* expression ($2^{(-\Delta Cq)}$) was calculated by first averaging Cq values of technical replicates, then calculating for each sample ΔCq = (average $Cq_{CBFB\text{-}MYH11}$-average $Cq_{murine\ CD45}$). For calculation purposes, the Cq value was set to the limit of detection (*i.e.,* 39) for samples with undetectable *CBFB-MYH11* transcript.

Statistics

Two-sample t-tests were performed to compare % survival of AML cell lines at each timepoint to % survival at baseline (0 hours). Unpaired t-tests with Welch's correction were performed to compare degranulation (CD107a staining) of T cell clones in response to AML samples with and without the fusion or restricting HLA, and on lysis of AML samples with and

without the fusion or restricting HLA by T cell clones. All t-tests were two-tailed. Repeated measures 2-way analysis of variance (ANOVA) testing was performed to compare human $CD33^+$ cells detected by flow cytometry in the peripheral blood of xenotransplanted animals treated with CBFB-MYH11/B*40:01-specific or irrelevant T cells. Unpaired parametric T tests were performed to compare *CBFB-MYH11* transcript expression normalized to murine CD45 transcript from rt-QPCR data ($2^{(-\Delta Cq)}$) in the peripheral blood of xenotransplanted animals treated with CBFB-MYH11/B*40:01-specific or irrelevant T cells. Correlation between different measurements of bone marrow disease burden was determined by calculation of Pearson correlation coefficient. For rt-QPCR data, statistical testing was performed on log-transformed data. A p-value <0.05 was considered significant. All statistical tests were performed using GraphPad Prism for Mac OS X, version 7.0d.

Identification of similar peptides

The CBFB-MYH11/B*40:01 epitope REEMEVHEL was evaluated for similarity to peptides from wild-type human proteins by performing a protein BLAST (blastp, https://blast.ncbi.nlm.nih.gov/Blast.cgi?) search for the original amino acid sequence and with X substituted in positions 1, 4, and 8 alone and in combinations against the non-redundant database of Homo sapiens sequences using default search parameters, including automatic adjustment of parameters for short input sequences.

Known microbial T cell epitopes were identified by searching IEDB (http://www.iedb.org/), restricting to all human infectious disease, class I-restricted targets with positive immune assays. REEMEVHEL was then used as the query sequence and aligned to the identified known microbial epitopes using protein BLAST with a BLOSUN62 matrix and default

parameters (gap existence cost 11, gap extension cost 1, conditional compositional score matrix for compositional adjustment) and automatic adjustment of parameters for short input sequences. Sequences were compared visually using WebLogo (94).

Tetramer enrichment of CBFB-MYH11-specific T cells

CBFB-MYH11-specific T cells were enriched from PBMC from healthy donors by tetramer enrichment performed using methods adapted from Moon and co-workers (86). Briefly, cryopreserved PBMC from healthy donors or individuals with active or treated AML were thawed and rested in CTL medium with 0.25 units/mL benzonase (EMD Millipore) and 1 ng/mL IL-15, at 37°C for 2 hours. Cells were then pre-incubated in 500 μL 1X PBS/5% FBS with 50 nM dasatinib and incubated for 15 minutes at 37°C to enhance TCR expression (95). Cells were then washed to removed dasatinib, and a single-cell suspension was prepared in 200 μl of 1X PBS/5% FBS. APC-conjugated CBFB-MYH11/B*40:01 tetramer was added to a concentration of 30 nM and incubated at room temperature for 15 minutes, followed by a wash in 15 ml of ice-cold MACS buffer (PBS + 1% FBS and 2 mM EDTA). Tetramer-stained cells were then resuspended to a volume of 200 μL of MACS buffer, mixed with 25 μL anti-APC–conjugated immunomagnetic microbeads (Miltenyi Biotec), and incubated for 15 minutes on ice. The cells were then resuspended in 4 mL of MACS buffer and passed over a magnetized LS column (Miltenyi Biotec). After sample was run, the column was washed with 4 mL of MACS buffer twice and removed from the magnetic field. Labeled cells were eluted by flushing the unmagnetized column twice with 5 mL of MACS buffer. After centrifugation, cell pellets from the enriched and column flow-through fractions were resuspended in PBS/5% FBS. Cell suspensions were incubated with PE-conjugated CBFB-MYH11/B*40:01 tetramer for 30

minutes on ice, then with surface antibodies for 15 minutes on ice. Cells were washed with MACS buffer twice and resuspended in MACS buffer with DAPI prior to sorting. Surface antibodies, used in various combinations, were as follows: anti-CD8, anti-CD45RA, anti-CD45RO, anti-CCR7, anti-CD27, anti-CD28. BV421-conjugated anti-CD4, -CD14, -CD16, -CD123, and -CD117 were also included on a single channel with DAPI (dump). Gating strategy to identify tetramer-positive antigen-specific $CD8^+$ T cells was as follows: single cells, BV421/DAPI negative, $CD8^+$, tetramer double-positive (gating strategy, **Figure S6A**).

NGS of TCR genes

Survey-level sequencing of the variable V-J or V-D-J regions of the *TRA* and *TRB* genes (Adaptive Biotechnologies, Seattle, WA) was performed on genomic DNA extracted from ~$6x10^4$ T cells from CBFB-MYH11-specific T cell clones. These regions encode the complementarity-determining region 3 (CDR3) of the TCR α and β chains, *i.e.* the hypervariable amino acid sequences responsible for contact with the cognate peptide. The coverage per sample was >10X.

RACE-PCR sequencing of TCR transcripts

RNA was extracted from each CBFB-MYH11/B*40:01-specific T cell clone. Full TCR regions were identified using 5′ first-strand complementary DNA (cDNA) amplification and rapid amplification of cDNA ends polymerase chain reaction (RACE-PCR) using a SMARTer RACE cDNA Amplification Kit (Clontech Laboratories). Briefly, cDNA was synthesized from RNA using 5′ CDS Primer A, SMARTer IIA oligo, and SMARTScribe Reverse Transcriptase.

The cDNA was then used to perform a RACE-PCR reaction using Phusion High-Fidelity DNA Polymerase and gene-specific primers for the TCR α (hTCR_Calpha-R 5′-CAGCCGCAGCGTCATGAGCAGATTA-3′) or TCR β chain (hTCR_Cb1-R 5′-CCACTTCCAGGGCTGCCTTCAGAAATC-3′ and hTCR_Cb2-R 5′-TGGGATGGTTTTGGAGCTAGCCTCTGG-3′). RACE-PCR products were purified and sequenced to identify TCR α and β chains. TCR variable, diversity, and joining regions were defined using IMGT/V-QUEST software.

TCR construction

TRA V-J–encoding sequences and *TRB* V-D-J–encoding sequences were fused to sequences encoding human constant *TRA* and *TRB* chains, respectively, and were confirmed by PCR using a forward primer from the 5' end of the appropriate V region and reverse primers from the *TRA* or *TRB* constant region(s) followed by Sanger sequencing. Complementary cysteine residues at positions 48 (Thr to Cys) and 57 (Ser to Cys) were incorporated introduced into TRA and TRB constant domains to increase exogenous TCR pairing and decrease mispairing with endogenous TCR. TCR chains were separated by codon-diversified P2A elements to ensure coordinated expression. Transgenes were codon-optimized to enhance expression prior to synthesis as Gene Art (Life Techologies) and transfer into pRRT (42).

Flow cytometry monoclonal antibodies and instruments

Flow cytometry was performed on a 5-laser (355 nm, 405 nm, 488 nm, 552 or 532 nm and 628 or 640 nm) Fortessa X50 or Symphony instrument (BD). Cell sorting was performed on

a 3-laser (405 nm, 488 nm, 633 nm) or 5-laser (355 nm, 408 nm, 488 nm, 561 nm, 635 nm) Aria II device (BD). All data was analyzed with FlowJo software (Tree Star). Fluorochrome-conjugated pHLA tetramers were produced in-house by the Fred Hutchinson Cancer Research Center Immune Monitoring Lab. All mAb used for flow cytometry were mouse anti-human, except for rituximab, which is an engineered chimeric anti-human mAb, and anti-murine CD45, which was produced in a rat host.

Flow cytometry monoclonal antibodies

Marker	*Clone*	*Fluor*	*Host*	*Specificity (anti-)*	*Item number*	*Manufacturer*
CD107a	H4A3	PE	Mouse	Human	555801	BD
CD14	M5E2	FITC	Mouse	Human	555397	BD
CD197 (CCR7)	3D12	PE-Cy7	Mouse	Human	557648	BD
CD20	Rituximab	none	Chimeric	Human	NA	Genentech
CD27	L128	BV650	Mouse	Human	563228	BD
CD28	CD28.2	BUV737	Mouse	Human	564438	BD
CD3	UCHT1	BV786	Mouse	Human	565491	BD
CD33	WM53	BB515	Mouse	Human	564588	BD
CD34	QBenD10	APC	Mouse	Human	FAB7227A	R&D
CD4	RPA-T4	BV421	Mouse	Human	562424	BD
CD4	RPA-T4	FITC	Mouse	Human	555346	BD
CD45	30-F11	BV421	Rat	Mouse	103134	Biolegend

CD45	2D1	APC-H7	Mouse	Human	560178	BD
CD45RA	HI100	PerCP-Cy 5.5	Mouse	Human	563429	BD
CD45RO	UCHL1	BV786	Mouse	Human	564290	BD
CD56	B159	PE	Mouse	Human	555516	BD
CD8	RPA-T8	BUV395	Mouse	Human	563795	BD
CD8	SK1	APC-H7	Mouse	Human	560179	BD
CD8	RPA-T8	APC	Mouse	Human	555369	BD
CD8	RPA-T8	PE	Mouse	Human	555367	BD
CD80	L307.4	PE	Mouse	Human	340294	BD
CD83	HB15e	PE	Mouse	Human	556855	BD
CD86	2331 (FUN-1)	FITC	Mouse	Human	555657	BD
Human IgG	HP6017	APC	Mouse	Human	409306	Biolegend

Supplementary Tables

Table S1. Additional predicted HLA-B*40:01-binding peptides. Underlined text indicates typical residue at anchor positions for HLA-B*40:01 binding.

Parent gene	Sequence	Length	Predicted HLA restriction
MLL-AFF1	N<u>E</u>VHCVEEI<u>L</u>	10	B*40:01
NPM1 type A/D	Q<u>E</u>AIQDLC<u>L</u>	9	B*40:01
CBFA2T3-GLIS2	S<u>E</u>DFQPLRY<u>L</u>	10	B*40:01
MLL-MLLT1	K<u>E</u>KCTVQVR<u>L</u>	10	B*40:01
MLL-MLLT4	G<u>E</u>DLEFHG<u>V</u>	9	B*40:01

Table S2. Sex and HLA-typing of healthy donor PBMC used in immunogenicity screening experiments.

ID	Sex	HLA-A		HLA-B		HLA-C	
Donor 1 (D1)	Male	A*02:01	A*02:01	B*40:01	B*07:02	C*03:04	C*07:02
Donor 2 (D2)	Female	A*02:01	A*02:01	B*40:01	B*44:02	C*03:04	C*05:01
Donor 3 (D3)	Female	A*02:01	A*01:01	B*40:01	B*07:02	C*03:04	C*07:02
Donor 4 (D4)	Female	A*02:01	A*03:01	B*40:01	B*35:01	C*03:04	C*04:01
Donor 5 (D5)	Female	A*02:01	A*11:01	B*40:01	B*56:01	C*03:04	C*01:02
Donor 6 (D6)	Female	A*02:01	A*02:01	B*40:01	B*57:01	C*03:04	C*06:02

Table S3. HLA-binding predictions for HLA-B*40:01 binding of peptides with sequential alanine substitutions at each position of REEMEVHEL epitope.

Amino acid sequence	Substituted position	Predicted HLA	IEDB ANN (nM)	IEDB SMM (nM)	netMHCpan 4.0 (nM)
REEMEVHEL	NA	B*40:01	8	24.5	25.5
AEEMEVHEL	1	B*40:01	18	42.8	36.5
RAEMEVHEL	2	B*40:01	9105	5246.9	17928.0
REAMEVHEL	3	B*40:01	16	32.1	13.8
REEAEVHEL	4	B*40:01	18	27.0	22.4
REEMAVHEL	5	B*40:01	6	12.5	13.0
REEMEAHEL	6	B*40:01	9	14.7	27.3
REEMEVAEL	7	B*40:01	19	27.2	21.7
REEMEVHAL	8	B*40:01	8	17.2	16.2
REEMEVHEA	9	B*40:01	2395	1103.8	2251.7

Supplementary Figures

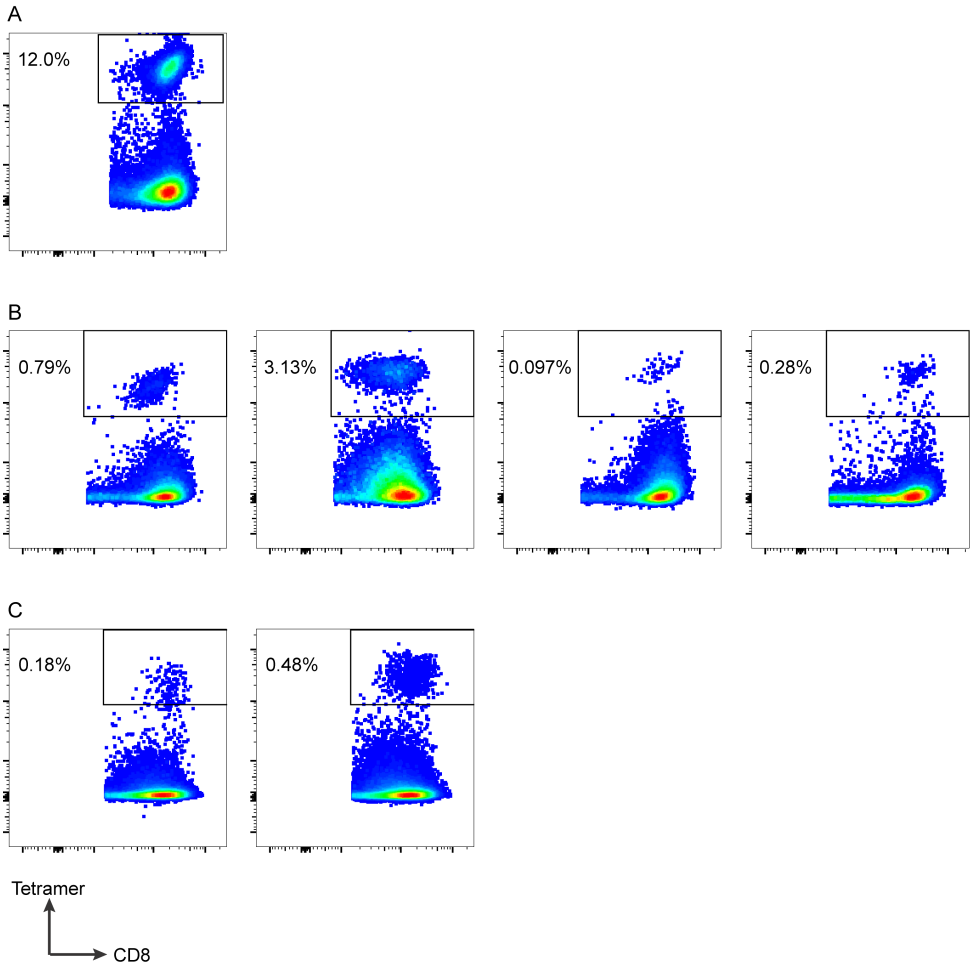


Figure S1. CBFB-MYH11/B*40:01 is reproducibly immunogenic in three additional healthy donors. $CD8^+$ T cells were isolated from three additional HLA-$B*40:01^+$ healthy donors and stimulated with autologous DC pulsed with CBFB-MYH11/B*40:01 peptide. On day 11-13 of co-culture, expansion of CBFB-MYH11/B*40:01-specific T cells was detected by staining with pHLA tetramer. Representative flow plots from CBFB-MYH11/B*40:01-specific lines from each donor (one experiment) are shown: (**A**) donor 4; (**B**) donor 5; (**C**) donor 6.

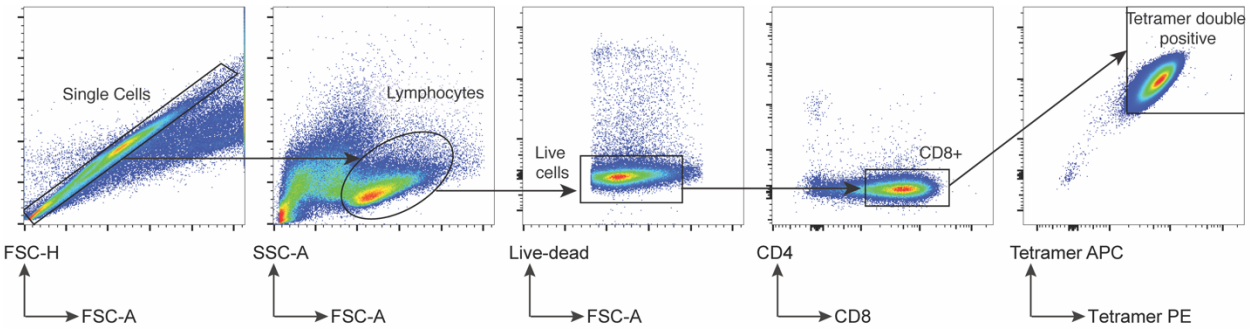


Figure S2. Gating strategy used for tetramer staining of clones shown in **Figure 2D**.

A

Amino acid sequence	Length	Predicted HLA	IEDB ANN (nM)	IEDB SMM (nM)	netMHCpan 4.0 (nM)
REEMEVHEL	9	B*40:01	16.4	24.5	25.5
REEMEVHEL	9	B*40:02	35.2	138	53.5
HREEMEVHEL	10	B*40:02	286.5	4580.3	1161.6
REEMEVHELE	10	B*40:02	438.8	216.7	1763.5

B

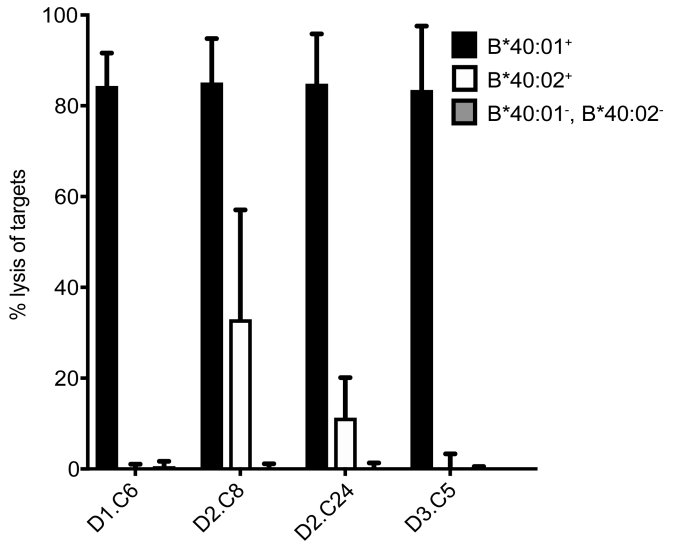


Figure S3. REEMEVHEL binds, but somewhat less efficiently, to HLA-B*40:02. **(A**) Three peptides spanning the junction of the CBFB-MYH11 type A fusion protein with a high probability of binding to HLA-B*40:02 were identified using three HLA-binding prediction algorithms. HLA-binding predictions for REEMEVHEL and HLA-B*40:01 are included on the first line for comparison. (**B**) High-avidity REEMEVHEL-specific clones D1.C6, D2.C8, D2.C24, and D3.C5 were tested for lysis of *HLA-B*40:01*$^{+}$ LCL, *HLA-B*40:02*$^{+}$ LCL, and *HLA-B*40:01*$^{-}$/ *B*40:02*$^{-}$ LCL pulsed with 1000 ng/mL REEMEVHEL peptide in a 4-hour CRA. Mean and SD of 3 technical replicate experiments are shown.

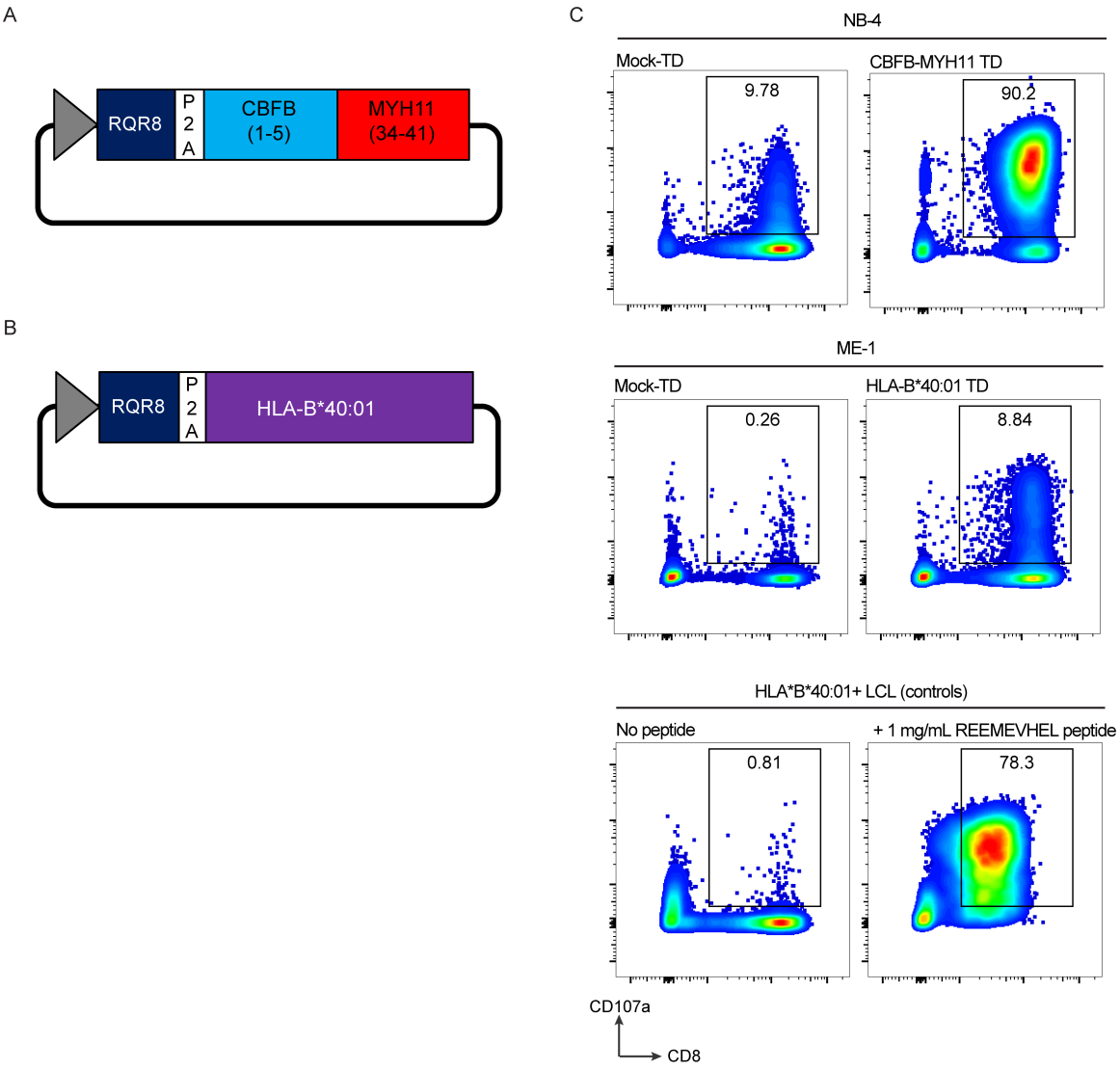


Figure S4. Generation of epitope-bearing AML cell lines. (**A**) Schematic of the lentiviral construct used to transduce HLA-B*40:01^{+} NB-4 cells with the full-length *CBFB-MYH11* type A fusion. (**B**) Schematic of the lentiviral construct used to transduce *CBFB-MYH11*$^{+}$ ME-1 cells with HLA-B*40:01. (**C**) Representative flow plots from CD107a assay (technical duplicates) showing degranulation of high-avidity REEMEVHEL-specific T cell clone D2.C24 in response to AML cell lines (NB-4, top row; ME-1, middle row) transduced to express the *CBFB-MYH11* or *HLA-B*40:01* transgenes, respectively (right panels), but not the untransduced cell lines (left

panels). *HLA-B*40:01*$^{+}$ LCL without and with REEMEHVEL peptide are included as negative and positive controls, respectively (bottom row).

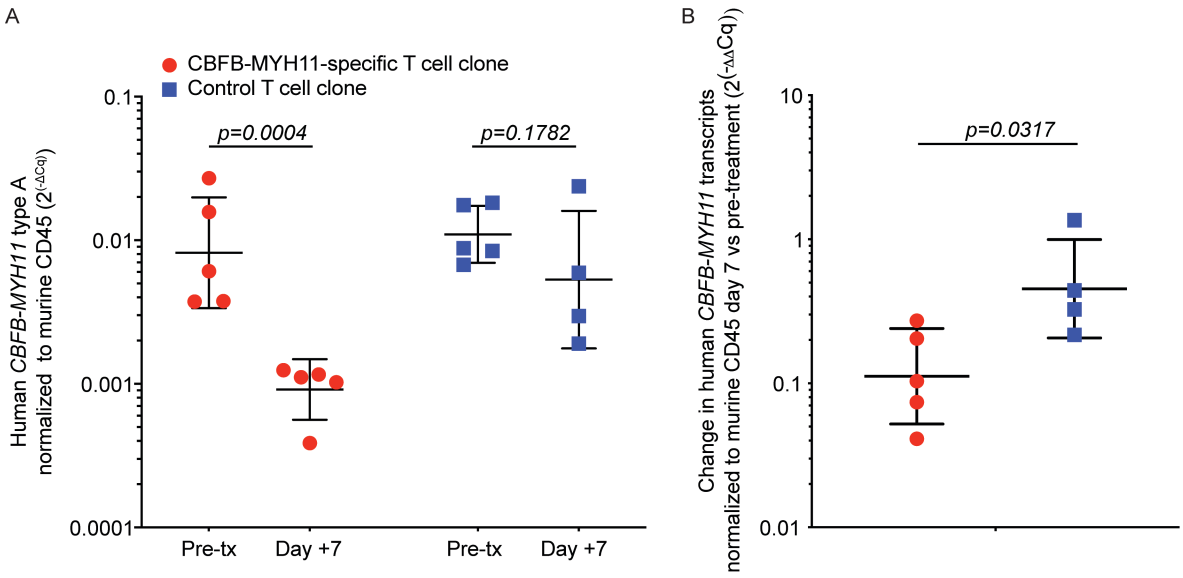


Figure S5. *CBFB-MYH11* transcript expression significantly decreases after treatment with CBFB-MYH11/B*40:01-specific T cells in vivo. (**A**) Expression of human CBFB-MYH11 type A transcript normalized to murine CD45 was assessed before and 7 days after administration of CBFB-MYH11/B40:01-specific (red squares) or control (blue circles) T cell clones. (**B**) The change in relative *CBFB-MYH11* transcript expression in peripheral blood between pre-treatment samples and day 7 post-T cells samples as calculated as $2^{(-\Delta\Delta Cq)}$. For all groups and timepoints, n=5, except for control T cell-treated mice on day 7 (n=4) due to poor RNA yield from one sample. Statistics were calculated using unpaired parametric t-tests. Geometric means and SD are shown.

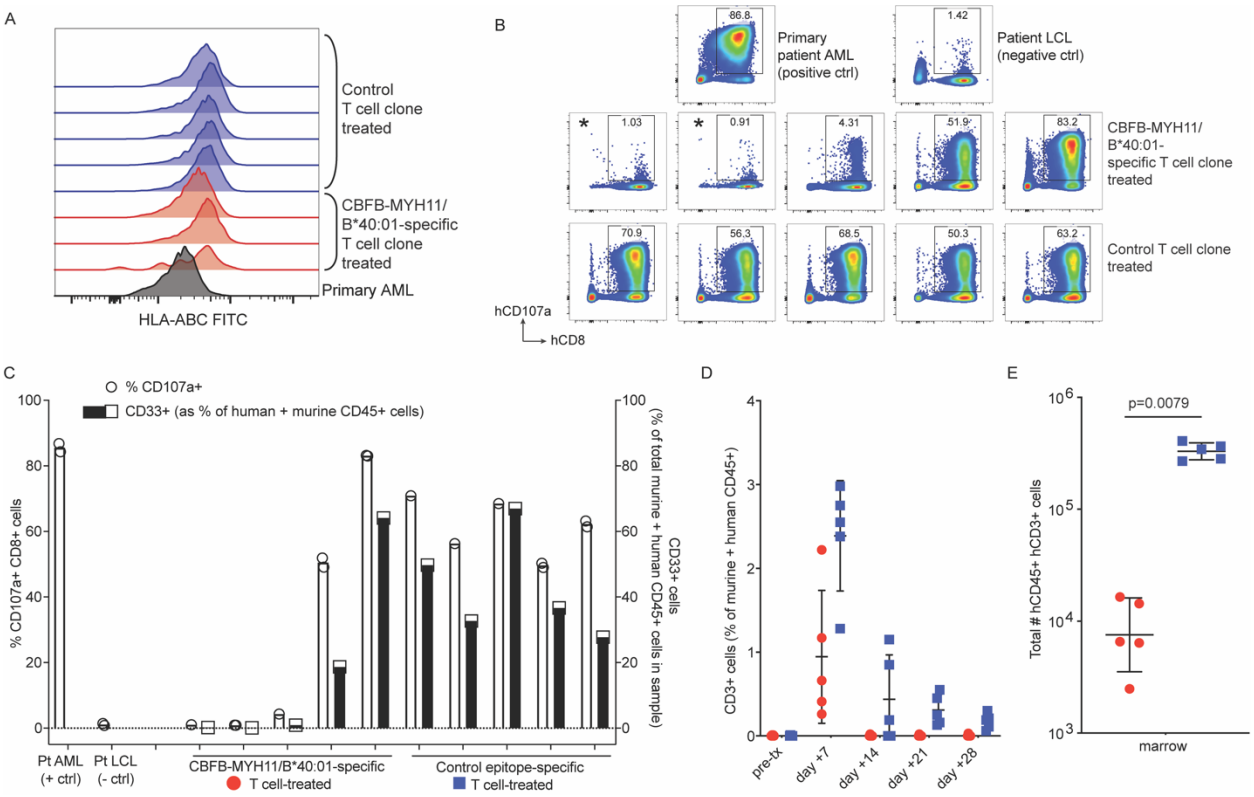


Figure S6. Persistent AML in xenografts continues to express HLA and present neoantigen. (**A**) HLA-class I staining of human CD45$^+$ cells (AML) from bone marrow of mice after treatment with control epitope-specific T cells (blue) or CBFB-MYH11/B*40:01-specific T cells (red). Only samples with >1% human CD45$^+$ in the marrow were included. The original primary AML sample (black) is included for comparison. (**B, C**) Presentation of the CBFB-MYH11/B*40:01 neoantigen by residual AML in xenografts was assessed by testing whether bone marrow samples from xenografts could induce degranulation of the CBFB-MYH11/B*40:01-specific T cell clone D2.C24 as measured by CD107a staining. (**B**), Representative flow plots (technical duplicates) from CD107a assay showing CD107a staining of D2.C24 CD8$^+$ cell clone in response to different stimulators. Top row, controls: left, primary AML2 sample (positive control), right, LCL from AML2 patient (negative control). Middle row, marrow cells from

CBFB-MYH11/B*40:01-specific T cell-treated mice as stimulators. Bottom row, marrow cells from control T cell-treated mice as stimulators. * indicates marrow samples with no detectable AML by flow or rt-QPCR. (**C**) Comparison of % $CD107a^+$ $CD8^+$ T cells (left y-axis, open bars with open circles) and human $CD33^+$ cells in murine marrow samples (right y-axis, black bars with open squares), indicating that T cell antigen recognition (as evidence by degranulation) and disease burden roughly correspond. (**D**) Circulating human $CD3^+$ T cells by flow cytometry in mice treated with CBFB-MYH11/B40:01-specific (red squares) or control (blue circles) T cell clones. (**E**) Absolute number of human $CD3^+$ cells in terminal bone marrow of mice treated with CBFB-MYH11/B40:01-specific (red squares) or control (blue circles). Statistics were calculated using unpaired parametric t-tests. In (**D**) and (**E**), mean and SD are shown.

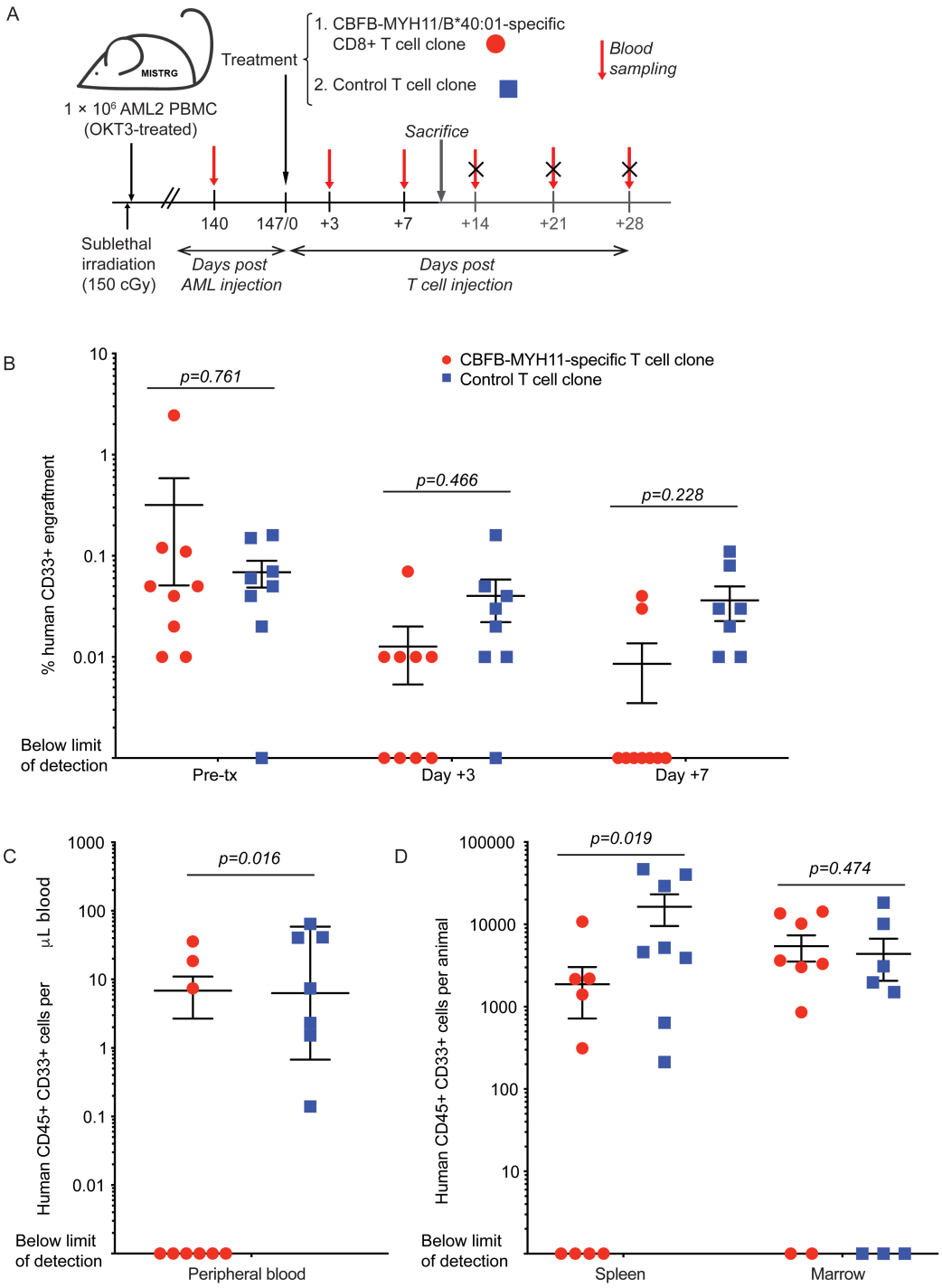


Figure S7. CBFB-MYH11/B*40:01-specific T cells control circulating AML in a short-term in vivo experiment. (**A**) Newborn, preconditioned MISTRG mice were injected intrahepatically with 1x10^6 OKT3-treated PBMC from an HLA-B*40:01$^+$ patient with active CBFB-MYH11$^+$

AML (AML1, 51% blasts), allowed to engraft with AML over 5 months, then treated intravenously with T cells either from the high-avidity CBFB-MYH11/B*40:01-specific D2.C24 clone or an irrelevant control clone specific for an HLA-B*07:02-presented epitope, an allele for which this AML was genotypically negative. Mice were monitored for peripheral disease burden by serial sampling of peripheral blood (PB). The planned endpoint was one month after T cell injection but was terminated prematurely due to COVID-19 (gray arrow, gray lines). (**B**) Summary of PB disease burden as evaluated by flow cytometry in mice treated with CBFB-MYH11/B40:01-specific (red squares, n=9) or control (blue circles, n=8) T cell clones. Statistics were calculated using repeated measures 2-way analysis of variance (ANOVA) (**C, D**) AML burden at sacrifice on day 10. Statistics were calculated using unpaired parametric t-tests. (**B**). Absolute numbers of human $CD33^+$ cells in terminal peripheral blood. (**C**), Total numbers of human $CD33^+$ cells in terminal spleen (left) and bone marrow (right). Mean and SEM are shown.

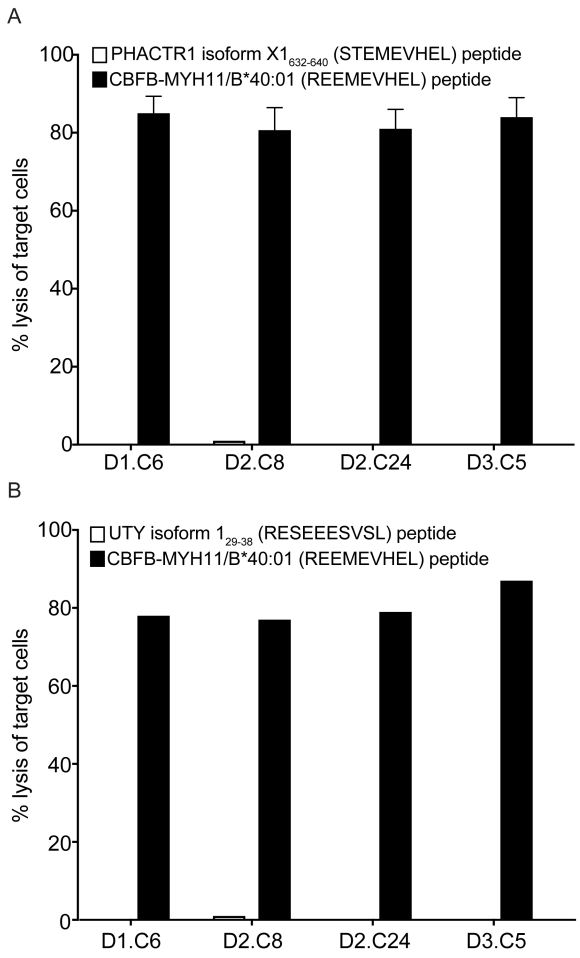


Figure S8. CBFB-MYH11/B*40:01 specific clones do not recognize similar human peptides. CBFB-MYH11/B*40:01-specific clones D1.C6, D2.C8, D2.C24, and D3.C5 were tested in 4-hour CRA for lysis of *HLA-B*40:01*$^{+}$ LCL pulsed with 1 µg/mL of either (**A**) STEMEVHEL peptide (PHACTR1 isoform X1$_{632\text{-}640}$; open bars) in three technical replicate experiments or (**B**) RESEEESVSL peptide (UTY isoform 1$_{29\text{-}38}$; open bars) in one experiment. Lysis of *HLA-B*40:01*$^{+}$ LCL pulsed with 1 µg/mL REEMEVHEL peptide is included in both (**A**) and (**B**) as positive control (solid bars).

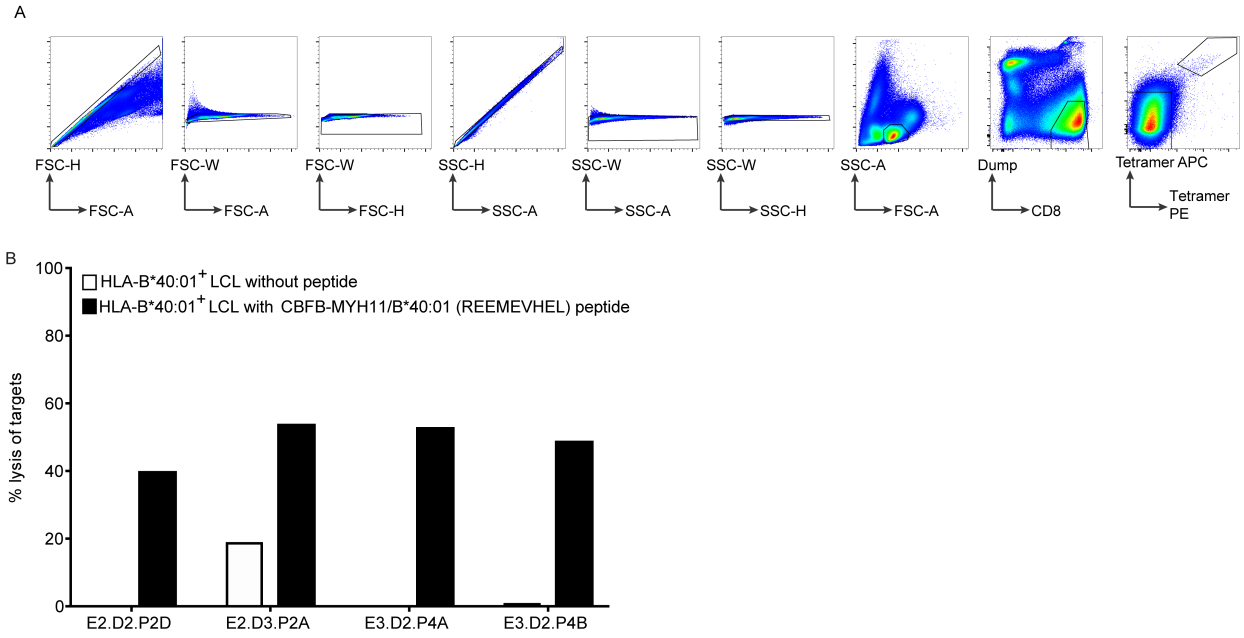
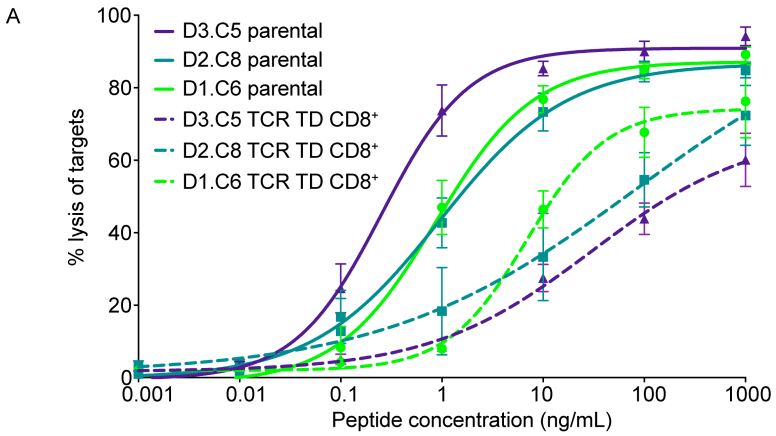
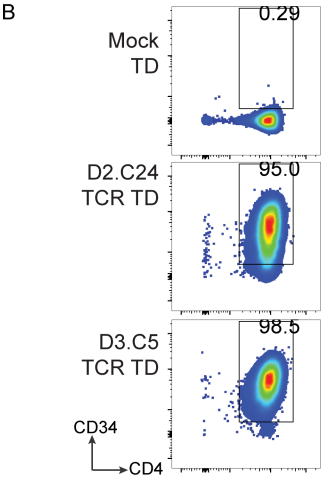
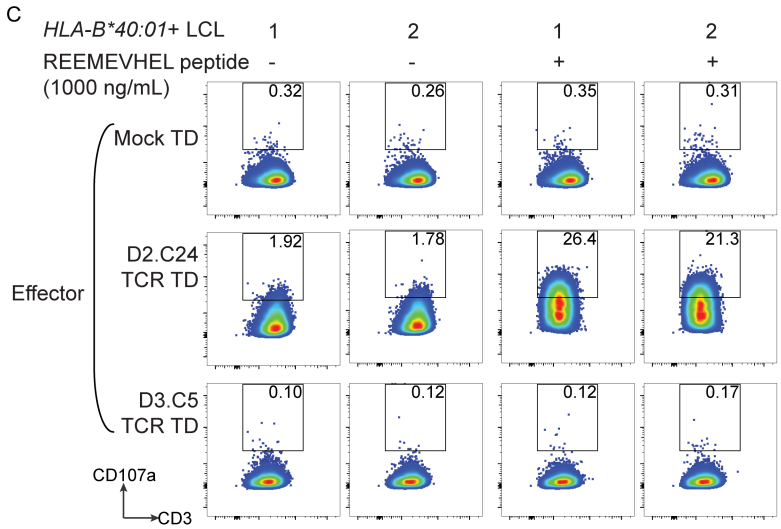
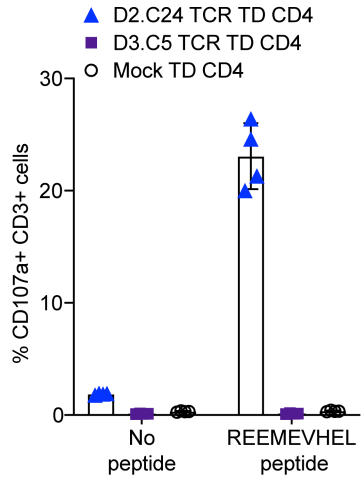


Figure S9. $CD8^+$ T cells isolated from healthy donors by CBFB-MYH11/B*40:01 pHLA tetramer enrichment and sorting are functional. (**A**) Gating strategy used for tetramer enrichment experiments using healthy donor PBMC. Dump channel is BV421 and encompasses cells staining positive for CD4, CD14, CD16, CD117, CD123, and for DAPI (i.e., nonviable). Tetramer refers to CBFB-MYH11/HLA-B*40:04 tetramer conjugated to APC (y-axis) or PE (x-axis). (**B**) $CD8^+$ APC $tetramer^+$ PE $tetramer^+$ cells were flow sorted from healthy donor PBMC after tetramer enrichment and expanded as polyclonal T cell lines (range of % $tetramer^+$ $CD8^+$ T cells in lines, 12.9-58.2%). T cell lines were used as effectors in CRA with HLA-B*40:01^+ LCL without peptide (open bars) or with 1 µg/mL REEMEVHEL peptide (solid bars). One experiment is shown.


A
D3.C5 parental
D2.C8 parental
D1.C6 parental
D3.C5 TCR TD CD8+
D2.C8 TCR TD CD8+
D1.C6 TCR TD CD8+
% lysis of targets
Peptide concentration (ng/mL)


B
Mock TD
0.29
D2.C24 TCR TD
95.0
D3.C5 TCR TD
98.5
CD34
CD4


C
HLA-B*40:01+ LCL
REEMEVHEL peptide (1000 ng/mL)
Mock TD
D2.C24 TCR TD
D3.C5 TCR TD
Effector
0.32
0.26
0.35
0.31
1.92
1.78
26.4
21.3
0.10
0.12
0.12
0.17
CD107a
CD3


D
D2.C24 TCR TD CD4
D3.C5 TCR TD CD4
Mock TD CD4
% CD107a+ CD3+ cells
No peptide
REEMEVHEL peptide

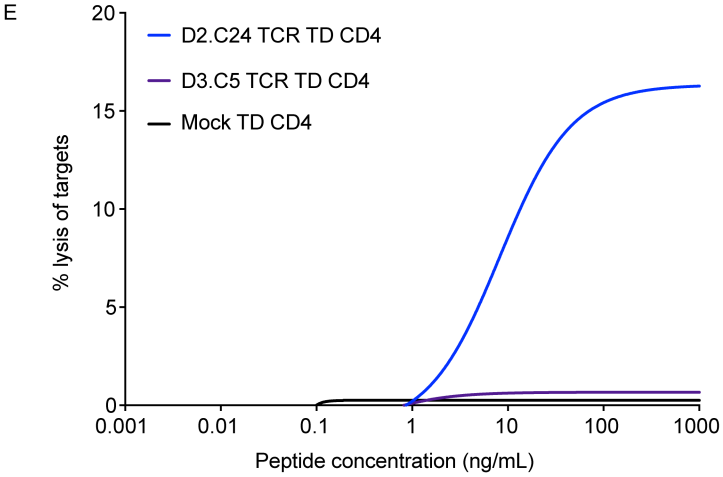

E
D2.C24 TCR TD CD4
D3.C5 TCR TD CD4
Mock TD CD4
% lysis of targets
Peptide concentration (ng/mL)

Figure S10. Additional evaluation of TCR constructs. (**A**) $CD8^+$ T cells transduced with each of the three constructs (D1.C6 TCR, D2.C8 TCR, D3.C5 TCR; broken lines) and corresponding parental clones (solid lines) were tested by CRA against *HLA-B*40:01*$^+$ LCL targets (n=3) pulsed with REEMEVHEL peptide at various concentrations. (**B-D**) TCRs from the two highest avidity $CD8^+$ T cell clones, D2.C24 and D3.C5, were transduced into $CD4^+$ T cells isolated from healthy volunteer donors, purified, and tested for transgene expression and CD8-independent recognition of REEMEVHEL peptide pulsed targets. Mean and SEM of technical triplicate experiments are shown. (**B**) Representative flow plots from evaluation of TCR transgene expression by CD34 staining, which detects an RQR8 selection marker in the transgene. All samples were >95% $CD4^+$ $CD34^+$. Representative flow plots (**C**) and summary data (**D**) from CD107a assay evaluating recognition of REEMEVHEL peptide by D2.C24 (blue triangles) or D3.C5 (purple squares) TCR-transduced or mock-transduced (open circles) $CD4^+$ T cells after co-culture with *HLA-B*40:01*$^+$ LCL cells (n=2) with or without the peptide (1000 ng/mL). Two technical replicates with each LCL and peptide condition were performed. Mean and SD are shown. (**E**) $CD4^+$ T cells transduced with D2.C24 or D3.C5 TCR or mock transduced were tested by CRA against *HLA-B*40:01*$^+$ LCL targets (n=1) pulsed with REEMEVHEL peptide at various concentrations.