Science
AAAS

Supplementary Materials for

Pediatric patients with acute lymphoblastic leukemia generate abundant and functional neoantigen-specific CD8$^+$ T cell responses

Anthony E. Zamora, Jeremy Chase Crawford, E. Kaitlynn Allen, Xi-zhi J. Guo, Jesse Bakke, Robert A. Carter, Hossam A. Abdelsamed, Ardiana Moustaki, Yongjin Li, Ti-Cheng Chang, Walid Awad, Mari H. Dallas, Charles G. Mullighan, James R. Downing, Terrence L. Geiger, Taosheng Chen, Douglas R. Green, Benjamin A. Youngblood, Jinghui Zhang, Paul G. Thomas*

*Correspondence to: paul.thomas@stjude.org

This PDF file includes:

Other Supplementary Material for this manuscript includes the following:

Materials and Methods

Patient Samples, Sequencing, and Variant Detection

The use of human tissues was approved by the institutional review board (IRB #00000029 – Expedited Protocol XPD-13-098) of St. Jude Children's Research Hospital. Tumor samples were obtained from diagnostic bone marrow biopsies, whereas matched germline samples were obtained either from peripheral blood, bone marrow, or adjacent normal tissue. Bone marrow mononuclear cells (BMMCs) were isolated from bone marrow biopsy samples by density gradient centrifugation using Lymphoprep (StemCell Technologies) under sterile conditions in a biological safety cabinet. After removal of the mononuclear layer, cells were counted, adjusted to 1-2 x10^7 cells/mL in 1-1.5 mL of cryopreservative media, and underwent program-controlled freezing in a Planer Kryo 560-16 freezer. Whole-genome sequencing was performed for each tissue type on all patients, and in some cases (N = 4 in this cohort) additional whole-exome sequencing (WES) was performed. mRNA sequencing was performed on tumor tissues only. The sequencing, alignment against the human reference genome, and the identification and validation of somatic non-synonymous mutations was described in detail elsewhere *(71,72)*. Somatic gene fusions were identified with mRNA sequences using the "CICERO" algorithm as performed elsewhere *(72)*.

Neoepitope prediction

Class I HLAs were inferred for each patient from diagnostic paired-end mRNA-Seq data using OptiType *(73)* configured with the IBM CPLEX optimizer and at 5 enumerations, and FPKM values were obtained using gencode annotations with featureCounts *(74)* and edgeR *(75)*. Using these predicted HLAs, all tumor-specific nonsynonymous mutations and genomic fusions were

computationally screened for antigenicity using the same tools as a pipeline described in detail elsewhere *(36)* and locally adapted for this study. For each mutation with a known reading frame, a sliding-window approach was used to generate all unique peptide sequences of lengths 8-15aa that contained the tumor-specific amino acid at any position. Unique peptides were similarly generated for genomic fusions, instead sliding the two amino acids at the fusion junction along the peptide sequence window. The binding affinity between each of these peptides and each of the patient's HLA class I alleles was modeled using NetMHCcons *(76)*, which was chosen due to this algorithm's high sensitivity and specificity across multiple HLA alleles *(76,77)*. Based on criteria used in similar studies *(30, 78)*, we classified peptides as either predicted strong binders (IC_{50} < 150 nM), intermediate binders (150 nM ≥ IC_{50} ≤ 500 nM), or non-binders (IC_{50} > 500 nM), and only predicted binders were considered further. For each putative neoepitope derived from a nonsynonymous mutation, the predicted binding affinity of the respective parent peptide (i.e., the germline peptide sequence without the tumor-specific amino acid) was also modeled for each of the patient's HLA alleles.

Flow Cytometry

To determine the phenotype of tumor infiltrating lymphocytes, we performed flow cytometric analysis on $CD8^+$ T-cell subsets (Naïve: $CD3^+CD8^+CCR7^+CD45RO^-$, Tcm: $CD3^+CD8^+CCR7^+CD45RO^+$, Teff/emra $CD3^+CD8^+CCR7^-CD45RO^-$ and Tem: $CD3^+CD8^+CCR7^-CD45RO^+$) from thawed bone marrow samples obtained from patients with B-ALL. BMMCs were initially stained with Live/Dead Aqua (Life Technologies) per manufacturer's instructions. Briefly, cells were centrifuged, washed, and resuspended in 1 mL of PBS to which 1 µL of reconstituted dye was added. Cells were then incubated for 20-30 minutes

at room temperature, protected from light. Following incubation, the cells were washed twice with 1 mL of staining buffer (1X DPBS, 2% FBS, 1mM EDTA, and 0.02% sodium azide). Cells were then stained for surface markers, as described above, in sorting buffer for 20 minutes at room temperature using the following human monoclonal antibodies (mAbs): APC/Cy7- CD3 (BioLegend 344818, Clone SK7), PB-CD8a (Biolegend 301033, Clone RPA-T8), FITC-CCR7 (BioLegend 353216, Clone G043H7), APC-CD45RO (BioLegend 304210, Clone UCHL1). Following surface staining, cells were washed twice, resuspended in 300 µL of staining buffer, and then analyzed using a custom-configured BD Fortessa using FACSDiva software (Becton-Dickinson). Data were analyzed using FlowJo software (TreeStar).

To determine the phenotype of tumor cells, we performed flow cytometric analysis on $CD19^+$ B cells from thawed bone marrow samples obtained from patients with B-ALL. BMMCs were initially stained with Live/Dead Aqua (Life Technologies) as described above. Cells were also stained for various surface markers, using the following human mAbs: CD3 (APC/Fire750-conjugated, BioLegend 344840, clone SK7; used as dump gate), CD8a (BV785-conjugated, BioLegend 301046, clone RPA-T8; used as dump gate), CD19 (PerCP/Cy5.5-conjugated, BioLegend 302230, clone HIB19), HLA-A2 (FITC-conjugated (BioLegend 343303) or BV421-conjugated (BioLegend 343325), clone BB7.2), HLA-A30, A31 (biotin-conjugated, One Lambda), CD273 (APC-conjugated, BioLegend 329608, clone 24F.10C12), CD274 (PE-conjugated, BioLegend 329706, clone 29E.2A3), Galectin-9 (FITC-conjugated, BioLegend 348912, clone 9M1-3). Following surface staining, cells were washed twice, resuspended in 300 µL of staining buffer, and then analyzed using a custom-configured BD Fortessa using FACSDiva software (Becton-Dickinson). Data were analyzed using FlowJo software (TreeStar).

Peptide synthesis

Peptides corresponding to patient-specific neoantigens were synthesized at the Peptide Synthesis Facility at St. Jude Children's Research Hospital, GenScript, or the Immune Monitoring Core at Fred Hutchinson Cancer Research Center. All peptides were synthesized as free acids and were more than 80% pure. For functional assays, sterile stock solutions of 1 mM 15mer peptides were prepared from lyophilized peptides reconstituted in the manufacturer's recommended solvent (ultrapure water or dimethyl sulfoxide) and stored at -20^{o}C. On the day of the experiment, stock solutions were diluted in complete RPMI media (RPMI 1640 supplemented with 10% FBS, 100 U/mL penicillin, 100 μg/mL streptomycin, 2mM glutamine, 1mM sodium pyruvate, 25 mM Hepes, and 1X non-essential amino acids) to a final concentration ranging from 1μM to 10 pM. For tetramer binding assays, 9-10mer peptides were synthesized by the Immune Monitoring Core at Fred Hutchinson Cancer Research Center and used to synthesize monomers and tetramers corresponding to patient-specific neoepitopes.

Generation of lentiviral-transduced K562 aAPCs

gBlock gene fragments, encoding HLA-A*02:01, HLA-A*30:02, HLA-B*15:03, HLA-B*18:01, or HLA-B*53:01 were obtained from Integrated DNA Technologies (IDT). Each HLA gBlock was cloned into the pLVX-EF1a-IRES-Puro bicistronic lentiviral expression vector (Clonetech). Lentivirus was generated by transfection of pLVX lentivirus vector containing an HLA insert, psPAX2 packaging plasmid, and pMD2.G envelope plasmid into 293T packaging cell line. Viral supernatant was harvested and filtered 24 and 48 hours after transfection. K562 cells (American Type Culture Collection (ATCC)) were transduced and selected with 2 μg/ml puromycin for one week in Iscove's Modified Dulbecco's Medium (IMDM; ATCC) containing 10% FBS. Routine

assays for expression of single HLA class I molecules via flow cytometric analysis and for mycoplasma (Lonza) were conducted.

Artificial APC (aAPC) functional assays using exogenous peptides

BMMCs were thawed, resuspended to 2 x 10^6 cells/mL in complete RPMI media, and rested in an incubator (37ºC, 5% CO_2) for 2 hours with DNAse I (100 U/mL; Worthington Biochemicals). The BMMCs were examined for viability by trypan blue exclusion and adjusted to 5 x10^7 cells/mL in separation buffer (PBS with 2% FBS and 1mM EDTA). $CD8^+$ T cells were enriched by negative selection (>90% purity) using the EasySep Human $CD8^+$ T Cell Isolation Kit (StemCell Technologies) according to the manufacturer's instructions. Enriched $CD8^+$ T cells (1-2 x 10^5) were co-cultured with aAPCs (1-2 x 10^5) expressing a single patient-specific HLA molecule that were pulsed with 1 μM of the appropriate peptide (either patient-specific 15mer neoantigen or 15mer irrelevant peptide), and 1 μg/mL each of anti-human CD28/CD49d (BD Biosciences) for 1 hour in an incubator (37ºC, 5% CO_2). An unstimulated (CD28/CD49d) and positive control (SEB, 100 μg/mL or 1X eBioscience Cell Stimulation Cocktail, PMA/ionomycin) were included in each assay. Following the one-hour incubation, monensin (Golgistop, 0.7 μL/mL; BD Biosciences) and brefeldin A (GolgiPlug, 1 μl/mL; BD Biosciences) were added to the cell cultures and the cells were placed back into the incubator (37ºC, 5% CO_2) for an additional 5 hours. Following incubation, the cells were washed twice with PBS, and stained with live/dead fixable violet amine reactive dye (Invitrogen Corporation) according to the manufacturer's instructions. The cells were then washed with staining buffer and stained with a cocktail of surface antibodies. The cells were washed with staining buffer, and then fixed and permeabilized using the Cytofix/Cytoperm kit (BD Biosciences) according to the manufacturer's

instructions. Following fixation and permeabilization, the cells were washed twice with 1X Perm/Wash buffer and stained with anti-IFNγ (FITC-conjugated, BD Biosciences, clone 25723.11 or AF647-conjugated, BioLegend, clone 4S.B3) and anti-TNFα (APC-conjugated, BD Biosciences, clone 6401.1111 or BV605-conjugated, BioLegend, clone MAb11) or corresponding isotype control antibodies. Cytokine-responsive cells were determined by subtracting the amount of IFNγ and TNFα produced by $CD8^+$ T cells stimulated with an irrelevant peptide from the amount produced when stimulating with a peptide corresponding to patient-specific tumor neoantigens and subsequently normalizing to the response obtained using a polyclonal source of stimulation (SEB and/or PMA/Ionomycin).

Autologous tumor functional assays using exogenous peptides

BMMCs were thawed, resuspended to 2 x 10^6 cells/mL in complete RPMI media, and rested in an incubator (37ºC, 5% CO_2) for 2-6 hours with DNAse I (100 U/mL; Worthington Biochemicals). The BMMCs were examined for viability by trypan blue exclusion and adjusted to 1-2 x 10^6 cells/mL. BMMCs (1-2 x10^6) were incubated with 1 μg/mL each of anti-human CD28/CD49d (BD Biosciences), 1 μM to 10 pM of the appropriate peptide (either patient-specific 15mer neoantigen or 15mer irrelevant peptide; see Table S6), and CD107a (PE-conjugated or BV650-conjugated, BioLegend, clone H4A3) and CD107b (PE-conjugated, BioLegend, clone H4B4) for 1 hour in an incubator (37ºC, 5% CO_2). An unstimulated (CD28/CD49d) and positive control (Staphylococcus enterotoxin B (SEB), 100 μg/mL or 1X eBioscience Cell Stimulation Cocktail, phorbol 12-myristate 13-acetate (PMA) and ionomycin) were included in each assay. Following the one-hour incubation, monensin (Golgistop, 0.7 μL/mL; BD Biosciences) and brefeldin A (GolgiPlug, 1 μl/mL; BD Biosciences) were added to

the cell cultures and the cells were placed back into the incubator (37°C, 5% CO_2) for an additional 5 hours. Following incubation, the cells were washed twice with PBS, and stained with live/dead fixable violet amine reactive dye (Invitrogen Corporation) according to the manufacturer's instructions. The cells were then washed with staining buffer (PBS containing 1% FBS and 0.02% Sodium Azide) and stained with a cocktail of surface antibodies. The cells were washed with staining buffer, and then fixed and permeabilized using the Cytofix/Cytoperm kit (BD Biosciences) according to the manufacturer's instructions. Following fixation and permeabilization, the cells were washed twice with 1X Perm/Wash buffer and stained with anti-IFNγ (FITC-conjugated, BD Biosciences, clone 25723.11 or AF647-conjugated, BioLegend, clone 4S.B3) and anti-TNFα (APC-conjugated, BD Biosciences, clone 6401.1111 or BV605-conjugated, BioLegend, clone MAb11) or corresponding isotype control antibodies. Cytokine-responsive and degranulating cells were determined by subtracting the amount of IFNγ, TNFα, and CD107a/b produced by $CD8^+$ T cells stimulated with an irrelevant peptide from the amount produced when stimulating with a peptide corresponding to patient-specific tumor neoantigens and subsequently normalizing to the response obtained using a polyclonal source of stimulation (SEB and/or PMA/Ionomycin).

Generation of tandem minigene (TMG) constructs, and in-vitro-transcribed (IVT) RNA

Minigenes were generated for each non-synonymous somatic mutation (and corresponding wild-type sequence) to encode the mutated (or wild-type) amino acid flanked bilaterally with 12 wild-type amino acids, resulting in a 25 amino acid peptide. TMG constructs were synthesized from multiple minigenes (2-7 per plasmid) arranged in tandem for a given patient. Each TMG construct was codon optimized, synthesized, and cloned into a pcDNA3.1 expression vector

using available EcoRI and NotI cut sites (GenScript). TMG plasmids were linearized with the restriction enzyme SmaI (NEB), and each linearized plasmid was used as a template for in-vitro-transcribed RNA using the mMESSAGE mMACHINE T7 Ultra Transcription Kit (Thermo Fisher Scientific) according to the manufacturer's instructions. RNA was purified using the MEGAclear Transcription Clean-Up Kit (Thermo Fisher Scientific) according to the manufacturer's instructions. Subsequently, RNA concentration was determined using a NanoDrop spectrophotometer and stored at -80^{o}C until later use.

Preparation of autologous tumor cells and CD8^{+}TILs for TMG and IVT functional assays

Autologous tumor cells (CD19^{+} B cells) were positively selected (>90% purity) from patient's bone marrow samples using the EasySep Human CD19^{+} Positive Selection Kit II (StemCell Technologies) according to the manufacturer's instructions. Enriched CD19^{+} tumor cells were cultured overnight in complete RPMI media, supplemented with multimeric CD40L (1000 ng/mL; AdipoGen Life Sciences), IL-4 (10 ng/mL), IL-21 (10 ng/mL), and BAFF (10 ng/mL) (all cytokines from GenScript). Following positive selection of CD19^{+} tumor cells, the flow through was used to enrich for CD8^{+} T cells by negative selection (>85% purity) using the EasySep Human CD8^{+} T Cell Isolation Kit (StemCell Technologies) according to the manufacturer's instructions. Enriched CD8^{+} T cells were cultured overnight in GT-T551 T cell culture media (Takara Bio), supplemented with recombinant human IL-2 (300 IU/mL; PeproTech).

TMG plasmid DNA transfection of aAPCs

aAPCs expressing a single patient-specific HLA molecule were transfected with corresponding mutant or wild-type (parent) TMG plasmid DNA using the Neon Transfection System (Thermo Fisher Scientific) according to the manufacturer's instructions. Briefly, aAPCs were resuspended in the kit's Resuspension Buffer R at a concentration of 1 x 10^7 cells/mL. The aAPCs were mixed with TMG plasmid DNA (1 μg plasmid DNA per 1 x 10^6 cells), aspirated into the 100 μL Neon Tip, and electroporated with 1,450 V for 10 ms and three pulses in the Neon device. Electroporated cells were transferred into 6-well plates containing 2 mL of IMDM media supplemented with 10% FBS, and incubated (37°C, 5% CO_2) overnight for use in TMG co-culture functional assays. Transfection efficiencies were routinely between 70-90%, on the basis of green fluorescent protein (GFP) expression 24 hours after transfection of aAPCs with control pcDNA3.1-GFP plasmid DNA (GenScript).

TMG RNA transfection of autologous tumor cells

Positively selected autologous tumor cells ($CD19^+$ B cells) were transfected with either mutant or wild-type (parent) TMG RNA using the Lipofectamine MessengerMAX (Thermo Fisher Scientific) according to the manufacturer's instructions. Briefly, 10 μg of the in-vitro-transcribed plasmid RNA was complexed to diluted Lipofectamine MessengerMAX Reagent in Opti-MEM reduced serum media. After a brief 5 minute incubation at room temperature, the RNA-lipid complex was added to 1 x 10^6 tumor cells in 6-well plates, and incubated (37°C, 5% CO_2) overnight for use in TMG co-culture functional assays. Transfection efficiencies were routinely between 40-60%, as assessed by GFP expression at 24 hours after transfection of tumor cells with a control GFP-encoding RNA (TriLink BioTechnologies) or with in-vitro-transcribed RNA generated from the pcDNA3.1-GFP plasmid (GenScript).

TMG co-culture functional assays

Enriched $CD8^+$ T cells from patient bone marrow samples or from healthy donor PBMCs were co-cultured with mutant or wild-type (parent) TMG-electroporated target cells at a ratio of 1:2 aAPCs or 1:5 tumor cells in complete RPMI media, respectively. Additionally, mock-TMG cells (target cells electroporated or cultured with lipofectamine reagents in the absence of plasmid DNA or RNA, respectively) or $CD8^+$ T cells alone, with or without PMA/Ionomycin, were used as controls. Co-cultured cells were incubated (37°C, 5% CO_2) for two hours, after which monensin (Golgistop, 0.7 µL/mL; BD Biosciences) and brefeldin A (GolgiPlug, 1 µl/mL; BD Biosciences) were added to the cell cultures and the cells were placed back into the incubator (37°C, 5% CO_2) for an additional 8-10 hours. Following incubation, the cells were washed twice with PBS, and stained with live/dead fixable violet amine reactive dye (Invitrogen Corporation) according to the manufacturer's instructions. The cells were then washed with staining buffer and stained with a cocktail of surface antibodies. The cells were washed with staining buffer, and then fixed and permeabilized using the Cytofix/Cytoperm kit (BD Biosciences) according to the manufacturer's instructions. Following fixation and permeabilization, the cells were washed twice with 1X Perm/Wash buffer and stained with anti-IFNγ (AF647-conjugated, BioLegend, clone 4S.B3) and anti-TNFα (BV605-conjugated, BioLegend, clone MAb11) antibodies. Cytokine-responsive $CD8^+$ T cells were determined by subtracting the amount of IFNγ or TNFα produced by $CD8^+$ T cells co-cultured with mock-TMG cells from the amount produced when co-culturing with mutant or wild-type (parent) TMG transfected cells and subsequently normalizing to the response by enriched $CD8^+$ T cells stimulated with PMA/Ionomycin.

TIL expansion and generation of tumor-reactive CD8$^+$ T cell lines

Tumor-reactive CD8$^+$ T cells from two diagnostic patient (ERG009 and ETV078) specimens were generated by co-culturing 2-3 x 10^4 sorted CD8$^+$ T cells with 8 x 10^4 autologous CD19$^+$ tumor cells in GT-T551 T cell culture media (Takara Bio), supplemented with 10% human AB serum (Gemini Bio-Products), recombinant human IL-2 (6,000 IU/mL, PeproTech,) and CD3 stimulation (30 ng/mL anti-CD3 antibody, clone OKT3, Miltenyi Biotec; or Dynabeads Human T-Activator CD3/CD28, Thermo Fisher) for a total of 21 days. A negative control (non-neoantigen expressing target) for each patient was included, in parallel, by co-culturing 2-3 x 10^4 sorted CD8$^+$ T cells with 8 x 10^4 aAPCs expressing patient-specific HLAs using the same culturing conditions and reagents outlined above. Tumor reactivity was determined by flow cytometry after 21 days of co-culturing by staining samples with Live/Dead Aqua (Life Technologies), an antibody against CD8a (BV785-conjugated, BioLegend 301046, clone RPA-T8), and with pre-titrated mutant, wild-type, or irrelevant APC- and/or PE-conjugated tetramers as outlined below.

Tetramer-binding assay

Tetramers corresponding to patient-specific neoepitopes (9-10mers binding patient-specific HLA class I molecules; Table S4) were synthesized at the Immune Monitoring Core at Fred Hutchinson Cancer Research Center. Table S4 lists all of the 10mer-containing tetramers for the neoepitopes for ERG009 that successfully refolded. For several cases, we had greater success producing tetramers containing 9mers due to improper folding when using peptides of longer length, thus we focused on the 9mer containing tetramers for subsequent patient samples. For tetramer-binding assays, BMMCs were thawed, resuspended to 2 x 10^6 cells/mL in complete

RPMI media, and rested in an incubator (37°C, 5% CO_2) for 2 hours with DNAse I (100 U/mL; Worthington Biochemicals). After 2 washes with PBS, 2-4 x 10^6 cells were incubated (37°C, 5% CO_2) for 30 mins in PBS containing 50 nM dasatinib. Subsequently, cells were stained with pre-titrated mutant (cancer neoantigen bound to patient-specific HLA), wild-type (parent self-peptide bound to patient-specific HLA), and/or irrelevant (nonself-antigen bound to patient-specific HLA) APC- or PE-conjugated tetramers for 20 mins in the incubator (37°C, 5% CO_2). After 2 washes with PBS, anti-PE (BioLegend, clone PE001) or anti-APC (BioLegend, clone APC003) primary unconjugated antibodies were added to the cells. After 2 washes with staining buffer, cells were stained with a selection of the following cell surface monoclonal antibodies (mAbs): CD3 (APC/Fire750-conjugated, BioLegend 344840, clone SK7), CD4 (PerCP/Cy5.5-conjugated, BioLegend 317428, clone OKT4), CD14 (PerCP/Cy5.5-conjugated, BioLegend 301824, or PE/Cy7-cojugated, BD Biosciences 557742, clone M5E2), CD19 (PerCP/Cy5.5-conjugated, BioLegend 302230, clone HIB19), CD8a (BV785-conjugated, BioLegend 301046, clone RPA-T8), CD45RO (BV605-conjugated, BioLegend 304238, clone UCHL1), CCR7 (FITC-conjugated, BioLegend 353216, clone G043H7), PD-1 (PE- or PE/Cy7-conjugated, BioLegend 329906/329918, clone EH12.2H7), TIM-3 (SuperBright436-conjugated, eBioscience 62-3109-42, or PE/Cy7-conjugated, BioLegend 345014, clone F38-2E2), and/or CD45RA (BV421-conjugated, BioLegend 304130, clone HI100). The cells were washed twice with PBS, and stained with live/dead fixable violet amine reactive dye (Invitrogen Corporation) according to the manufacturer's instructions. The cells were then washed with staining buffer and analyzed on a custom-configured BD Fortessa using FACSDiva software (Becton-Dickinson). Data were analyzed using FlowJo software (TreeStar). The distinction of neoepitope-specific $CD8^+$ T cells was restricted to those tetramer-positive $CD8^+$ T cells that bound patient-specific neoepitope

(mutant) tetramers with a 5-fold greater frequency than $CD8^+$ T cells bound to either the parent (self-peptide containing) tetramer or to an irrelevant tetramer containing a patient associated HLA.

Generation of retroviral-transduced SUP-T1 cells

αβ TCR sequences were amplified from tetramer-sorted (HLA-A*30:02 PLCD3($_{311\text{-}319}$) mutant tetramer) single cells as described previously *(79)*. gBlock gene fragments encoding the clonal TCRαβ were obtained from Integrated DNA Technologies (IDT). Each TCRαβ gBlock was cloned into the MSCV-IRES-Thy1.1 DEST retroviral expression vector (Addgene). Recombinant MSCV-IRES-Thy1.1 plasmids with full length TCRαβ inserts were isolated in small scale using a NucleoSpin Plasmid kit (Clontech) and in large scale, for transfection, using a Plasmid Midi kit (Qiagen) per the manufacturer's instructions. TCRαβ and retroviral packaging vectors (5 μg of each) were co-transfected into Phoenix GP cell line (ATCC® CRL-3215) using Lipofectamine 3000 (ThermoFisher) per the manufacturer's instructions. Supernatant was collected 24 and 48 hours after transfection and centrifuged at 2000g for 2 hours after overnight incubation with retrovirus concentrator (Retro-X, Clontech) at 4°C. The retrovirus pellets were resuspended in complete RPMI medium and stored at -20°C. SUP-T1 cells (ATCC CRL-1942) were cultured in the retrovirus-containing c-RPMI medium at 37°C and 5% CO_2 for 72 hours. Routine assays for TCRαβ expression were performed using flow cytometric analysis, and

neoepitope specificity of the TCRαβ-transduced SUP-T1 cells were verified using neoepitope-specific tetramers.

Fluidigm gene expression methods

Single-cell gene expression experiments were performed using Fluidigm's DynamicArray microfluidic chips (Fluidigm) following manufacturer's instructions (Two-Step Single-cell gene expression using EvaGreen Supermix on Biomark HD System). Single cells were sorted from three patient samples using the Sony SY3200 into individual wells of 96-well plates that had been preloaded with 5 µL consisting of 1.2 µL 5X SuperScript VILO Reaction Mix (Thermo Fisher), 0.3 µL SUPERase-In (20 U/µL; Thermo Fisher), 0.25 µL 10% NP40 (Thermo Fisher), and nuclease free water (Teknova). Each plate contained two empty wells for use as non-template controls (NTC). After successful sorting, the plates were incubated at 65°C for 90 seconds, and chilled on ice for 5 minutes for RNA denaturation. The remaining components of the RT reaction were added to each well, including 0.15 µL 10X SuperScript Enzyme Mix (Thermo Fisher), 0.12 µL T4 Gene 32 Protein (NEB), and nuclease free water (Teknova). Thermal cycling conditions for the reverse transcription of single-cell mRNA into cDNA was 25°C for 5 min, 50°C for 30 min, 55°C for 25 min, 60°C for 5 min, 70°C for 10 min, and then hold at 4°C. After reverse transcription, the cDNA was amplified using 1.5 µL of a pool of the designed primers (STA mix; Fluidigm), 7.5 µL TaqMan PreAmp Master Mix (Thermo Fisher), and 0.075 µL 0.5M EDTA (Thermo Fisher). The thermal cycling conditions were 95°C for 10 min, 20 cycles of 96° for 5 sec and 60°C for 4 min, and then hold at 4°C. Amplified cDNA was then treated with Exonuclease I (NEB) to remove any unincorporated primers by adding 6 µL of 4 U/µL Exonuclease I (20 U/µL Exonuclease I diluted with water and 10X Exonuclease I

Reaction buffer). Thermal cycling conditions were 37°C for 30 min, 80°C for 15 min, and then hold at 4°C. The cDNA was then diluted 5-fold for loading onto the chip.

The sample mix to be loaded onto the chip included 2.5 µL 2X Sso Fast EvaGreen Supermix with low ROX (Bio-Rad), 0.25 µL 20S DNA Binding Dye Sample Loading Reagent (Fluidigm), and 2.25 µL of the prepared cDNA. The assay mix loaded onto the chip included 2.5 µL 2X Assay Loading Reagent (Fluidigm), 1X DNA Suspension Buffer (Teknova), and 100 µM of each forward and reverse primer mix (Fluidigm). Each chip was primed immediately before loading the 5 µL of each sample mix and 5 µL of each assay mix onto the chip. The chips were then loaded using the IFC Controller HX (Fluidigm) and transferred to the Biomark HD real-time PCR reader to analyze using the GE Fast 96x96 PCR+Melt v2 protocol.

Analysis and visualization of single-cell expression data

Ct values were recovered from the BioMark HD, and data were then analyzed using the Fludigm Real-Time PCR Analysis software. The quality threshold was set to 0.65, and a linear derivative was used for baseline correction. Cells were removed from further analysis if they had low or absent ACTB and GAPDH expression (n=35). As positive controls, PBMC samples from healthy donors were sorted exactly as patient samples ($CCR7^{-}CD45RO^{+}$ and $CCR7^{+}CD45RO^{-}$) in bulk (at least 100 cells per well). Fluidigm reactions with Tm that did not match what was observed in the positive controls were excluded from analysis. Genes were removed if amplification was found in the non-template control wells (n=3), if there were several melting temperature curves (n=1), if median Ct values were <6 (n=1), or if there was no expression in

any reaction (n=1). A list of the 88 DeltaGene assays used in this study with primer sequences is provided in Table S5.

Ct values of filtered single-cell expression data were converted to expression thresholds (Et) and analysed using the MAST statistical framework *(70)* in R, which utilizes hurdle models designed specifically for single-cell expression data that take into account both the proportion of cells expressing a gene and the expression of the gene. Cells with a cellular detection rate in the 1% and 99% quantiles were filtered from further analysis, and missing data points were considered as zero-values unless otherwise indicated.

To facilitate simultaneous analysis of transcript and protein expression, transcript data were combined with cell-specific protein expression data obtained from indexed flow cytometry by centering each measure on its mean and scaling to unit variance. Initial investigations of the data were conducted via Principal Components Analysis (PCA), using the two principal components that accounted for the greatest variation in the data to search for patterns related to patient identity and tetramer specificity. We then conducted cut-tree hierarchical clustering with Ward's minimum variance method on the first two principal components in order to further investigate potential expression similarities among groups of cells, visually confirming the consistent separation of clusters by overlaying cluster membership on the PCA plot. We employed a sequential approach whereby we expanded the number of clusters iteratively to search for the largest value for which expression differences among putative clusters most closely resembled the published expression profiles of known $CD8^+$ T cell subsets. Whereas two clusters broadly recapitulated differences typically observed between exhausted and non-exhausted $CD8^+$ T cells, three clusters revealed an effector subset and a dysfunctional subset within the non-exhausted cluster, resulting in three broad functional groups supported by the literature: exhausted cells,

effectors, and dysfunctional cells. Additional iterations of dendrogram cutting revealed putative subdivisions of various clusters, but we were unable to find literature supporting the biological relevance of these potential subpopulations. Heatmaps were generated using the R package NMF *(80)* with missing data points excluded, and rather than using the center-scaled data, the raw protein expression data were re-scaled to the full range of the Et data in order to visualize the heatmap using an absolute metric.

Methylation

Genomic DNA (100 ng) was subjected to bisulfite (BS) treatment using the EZ DNA methylation Gold kit (Zymo Research), which converts all unmethylated cytosines to uracils while protecting methylated cytosines from the deamination reaction. To perform loci-specific methylation analysis, bisulfite-modified DNA was PCR amplified with locus-specific primers for *IFNγ*, *TCF7*, and *TBX21*. The PCR product was gel extracted using the Zymoclean Gel DNA Recovery Kit (Zymo Research) and ligated overnight at 4^{o}C to the pGEMT easy vector using T4 DNA ligase following the manufacturer's protocol (pGEM T vector system-I, Promega). XL-Gold ultracompetent bacteria (Stratagene) were transformed with the ligated DNA and spread on LB agar plates containing Ampicillin, X-Gal, and IPTG and incubated for 16 hrs at 37^{o}C. White colonies were selected and inoculated into 1ml of 2x LB broth, and incubated overnight at 37^{o}C in a shaking incubator at 220 rpm. Plasmid DNA was isolated from the cultured bacteria using the Directpret 96 Miniprep Kit (Qiagen), and the genomic insert was Sanger sequenced.

Sequencing results were analyzed using the QUMA online server to assess the extent of methylation at the selected loci *(81)*.

Code and Data availability

Single-cell gene expression data have been submitted to the Gene Expression Omnibus (GSE130670) and single-cell protein data is provided in Table S8. All other sequencing data can be accessed via European Genome-Phenome Archive (www.ebi.ac.uk/ega/search/site/PCGP) and St Jude Pediatric Cancer (PeCan) Data Portal (https://pecan.stjude.org/home). The EGA numbers corresponding to the data in this manuscript are EGAD00001002654 and EGAS00001000514.

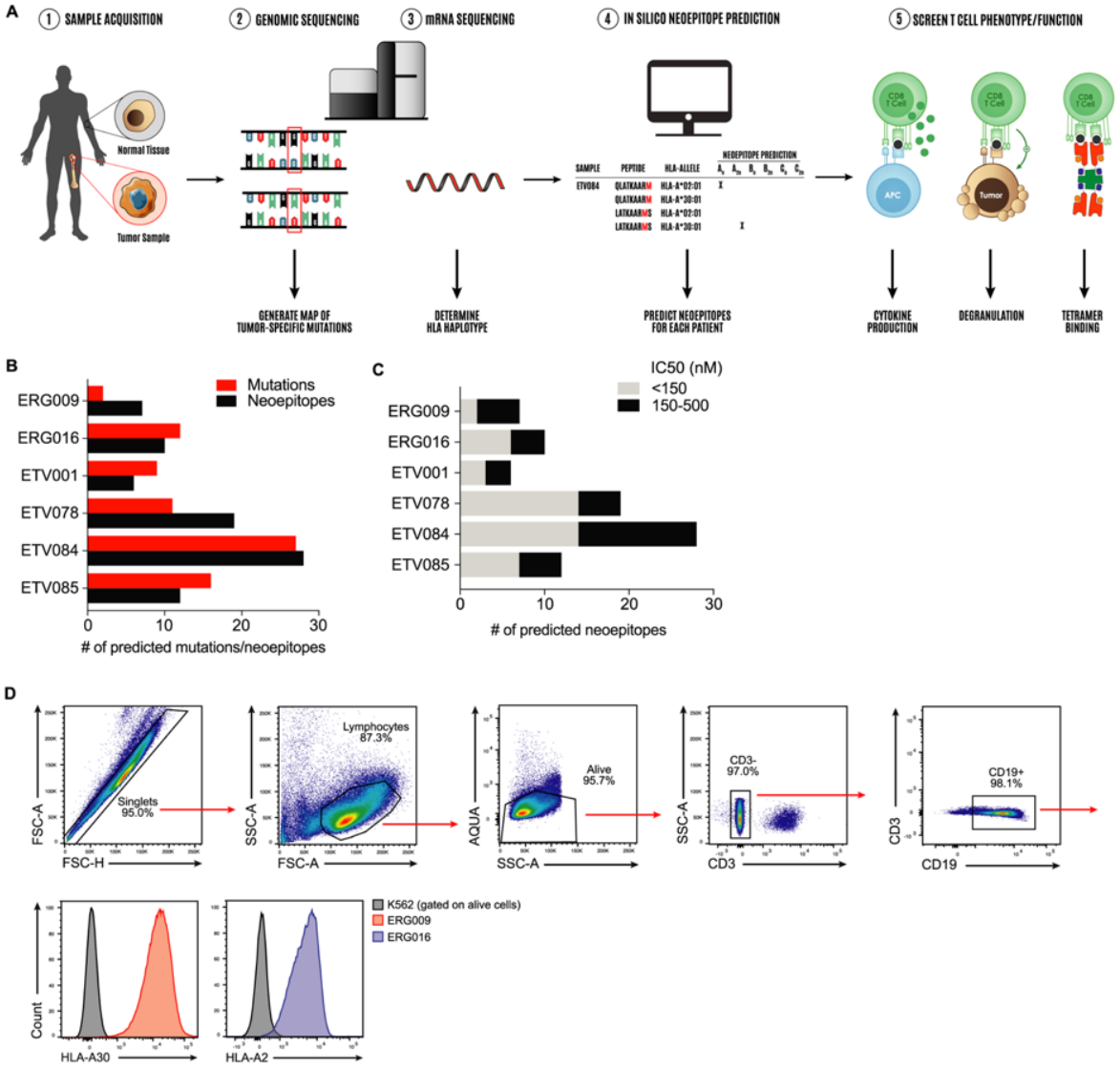

Fig. S1. Experimental pipeline used to identify cancer neoepitopes. (A) (Step 1) Tumor material was analyzed against matched germline tissue to identify tumor-specific mutations using WES and/or WGS (Step 2). (Step 3) Expression of HLA class I genes was assessed from paired-end mRNA sequencing from tumor biopsies using OptiType. (Step 4) Neoepitope prediction was performed using NetMHCcons to determine the binding affinity between mutant peptides and patient-specific HLA class I alleles. (Step 5) Putative neoantigens and neoepitopes were experimentally validated for their ability to induce tumor - specific $CD8^+$ T cell responses by tetramer binding, cytokine production, and degranulation. (B) Number of somatic mutations (red) and putative neoeptiopes (black) from patients with B-ALL generated from WES and WGS data. (C) Number of putative neoepitopes with HLA binding affinity <150 nM (gray) and 150 to 500 nM (black) from patients with B-ALL. (D) Flow cytometric analysis of BMMCs depicting the frequency of $CD19^+$ tumor cells expressing HLA molecules for ERG009 (red histogram, HLA-A30) and ERG016 (blue histogram, HLA-A2).

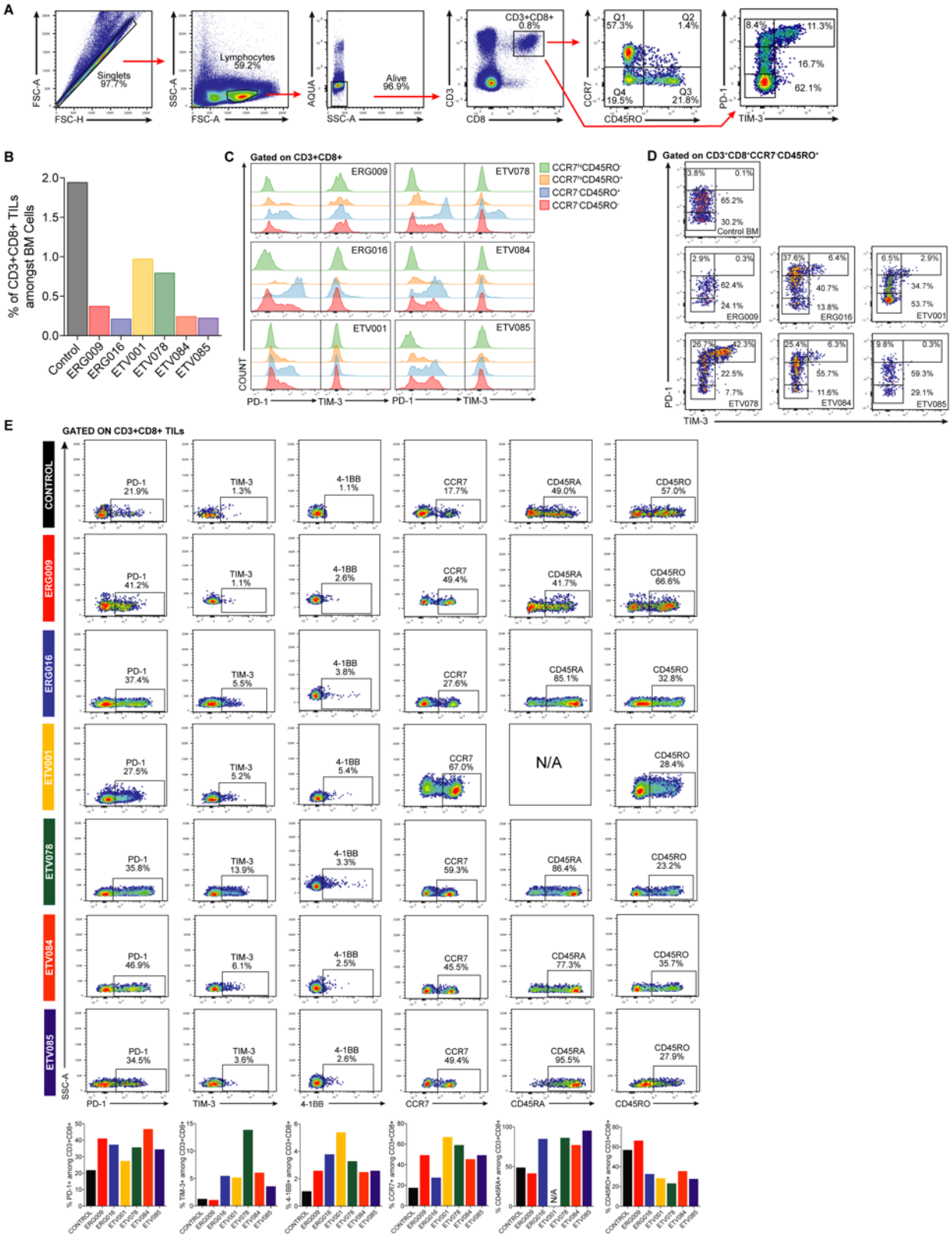
A
FSC-A
FSC-H
Singlets
97.7%
SSC-A
FSC-A
Lymphocytes
59.2%
AQUA
SSC-A
Alive
96.9%
CD3
CD8
CD3+CD8+
0.8%
CCR7
CD45RO
Q1
57.3%
Q2
1.4%
Q4
19.5%
Q3
21.8%
PD-1
TIM-3
8.4%
11.3%
16.7%
62.1%
B
% of CD3+CD8+ TILs amongst BM Cells
2.0
1.5
1.0
0.5
0.0
Control
ERG009
ERG016
ETV001
ETV078
ETV084
ETV085
C
Gated on CD3+CD8+
ERG009
ETV078
ERG016
ETV084
ETV001
ETV085
CCR7hiCD45RO-
CCR7hiCD45RO+
CCR7-CD45RO+
CCR7-CD45RO-
COUNT
PD-1
TIM-3
D
Gated on CD3+CD8+CCR7-CD45RO+
3.8%
0.1%
65.2%
30.2%
Control BM
2.9%
0.3%
62.4%
24.1%
ERG009
37.6%
6.4%
40.7%
13.8%
ERG016
6.5%
2.9%
34.7%
53.7%
ETV001
26.7%
42.3%
22.5%
7.7%
ETV078
25.4%
6.3%
55.7%
11.6%
ETV084
9.8%
0.3%
59.3%
29.1%
ETV085
PD-1
TIM-3
E
GATED ON CD3+CD8+ TILs
CONTROL
ERG009
ERG016
ETV001
ETV078
ETV084
ETV085
PD-1
21.9%
TIM-3
1.3%
4-1BB
1.1%
CCR7
17.7%
CD45RA
49.0%
CD45RO
57.0%
PD-1
41.2%
TIM-3
1.1%
4-1BB
2.6%
CCR7
49.4%
CD45RA
41.7%
CD45RO
66.6%
PD-1
37.4%
TIM-3
5.5%
4-1BB
3.8%
CCR7
27.6%
CD45RA
85.1%
CD45RO
32.8%
PD-1
27.5%
TIM-3
5.2%
4-1BB
5.4%
CCR7
67.0%
N/A
CD45RO
28.4%
PD-1
35.8%
TIM-3
13.9%
4-1BB
3.3%
CCR7
59.3%
CD45RA
86.4%
CD45RO
23.2%
PD-1
46.9%
TIM-3
6.1%
4-1BB
2.5%
CCR7
45.5%
CD45RA
77.3%
CD45RO
35.7%
PD-1
34.5%
TIM-3
3.6%
4-1BB
2.6%
CCR7
49.4%
CD45RA
95.5%
CD45RO
27.9%
SSC-A
PD-1
TIM-3
4-1BB
CCR7
CD45RA
CD45RO
% PD-1+ among CD3+CD8+
% TIM-3+ among CD3+CD8+
% 4-1BB+ among CD3+CD8+
% CCR7+ among CD3+CD8+
% CD45RA+ among CD3+CD8+
N/A
% CD45RO+ among CD3+CD8+
CONTROL
ERG009
ERG016
ETV001
ETV078
ETV084
ETV085

Fig. S2. Phenotypic characterization of $CD8^+$ tumor infiltrating lymphocytes from human bone marrow samples. (A) Flow cytometric analysis of one representative patient depicting the gating strategy used to characterize the differentiation status (CCR7 and CD45RO expression), PD-1 expression, and TIM-3 expression. (B) Frequency of $CD8^+$ T cells among total bone marrow cells from a healthy donor (control) and patient bone marrow samples. (C) Overlaid histogram plots show expression of PD-1 and TIM-3 in different $CD3^+CD8^+$ TIL subsets. $CD8^+$ T cell subsets were defined as follows: TN cells, $CD3^+CD8^+CCR7^+CD45RO^-$; TCM cells, $CD3^+CD8^+CCR7^+CD45RO^+$; TEM cells, $CD3^+CD8^+CCR7^-CD45RO^+$; TEFF/EMRA cells, $CD3^+CD8^+CCR7^-CD45RO^-$. (D) Flow cytometric analysis of PD-1 and TIM-3 on TEM $CD8^+$ lymphocytes from a healthy donor (control) or patient bone marrow samples. (E) Flow dot plots depicting the expression of PD-1, TIM-3, 4-1BB, CCR7, CD45RA, and CD45RO on $CD3^+CD8^+$ TILs from healthy donor (control) and patient bone marrow samples. Summary frequency graphs (bottom) depict the differential expression of each marker between healthy donor and patient samples. Bone marrow from a single patient (ETV001) was not stained for marker CD45RA.

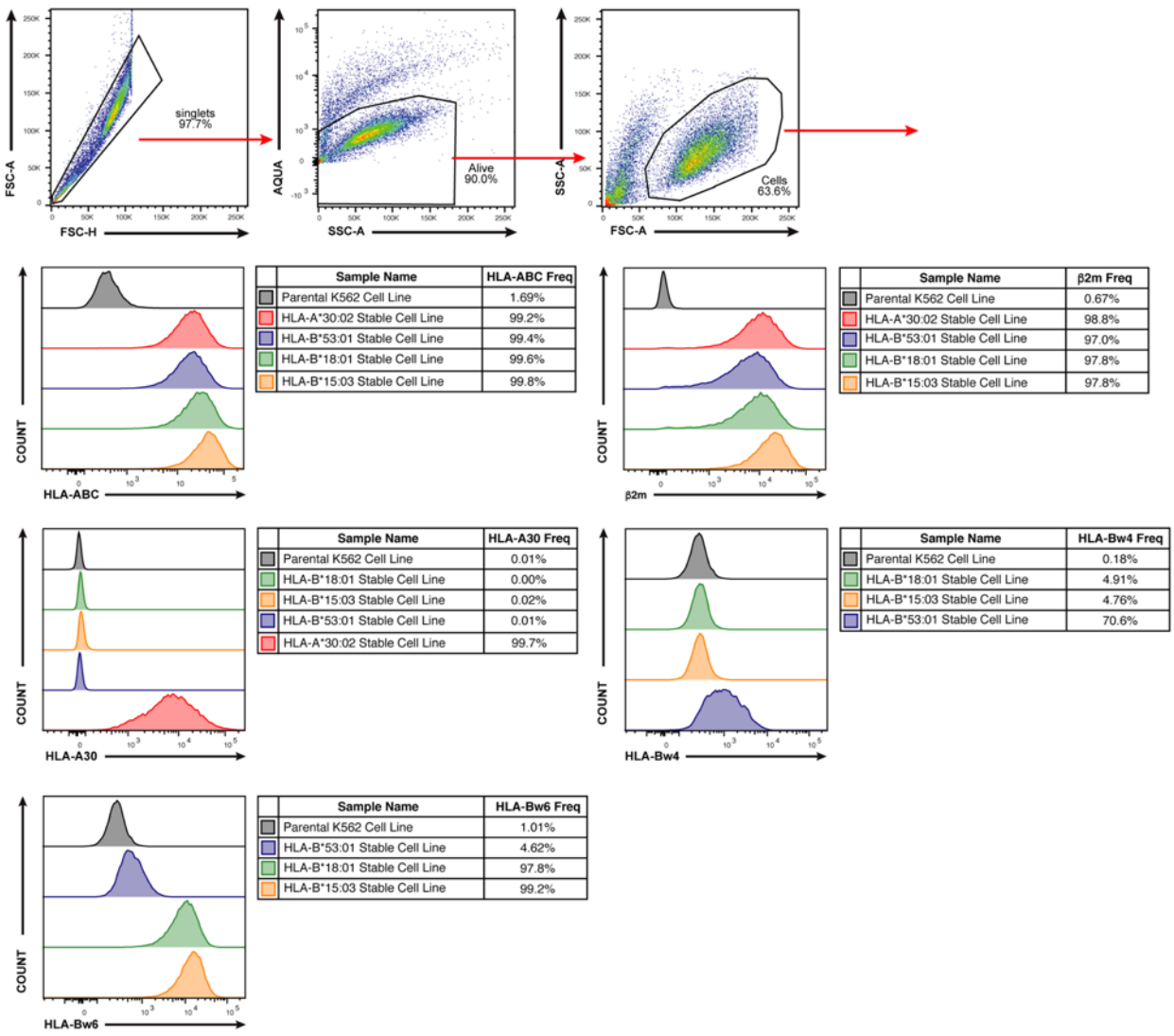

Fig. S3. Generation of aAPCs expressing single patient-specific HLA class I molecules.

Representative gating strategy used to identify HLA expression patterns among single HLA transduced K562 cells using monoclonal antibodies against b2-microglobulin, pan HLA-A,B,C molecules, HLA-A30, HLA-Bw4, or HLA-Bw6. Overlaid histogram plots show expression of HLA-A*30:02, HLA-B*15:03, HLA-B*18:01, and HLA-B*53:01 between the parental K562 cell line, and single HLA class I expressing stable cell lines (aAPCs).

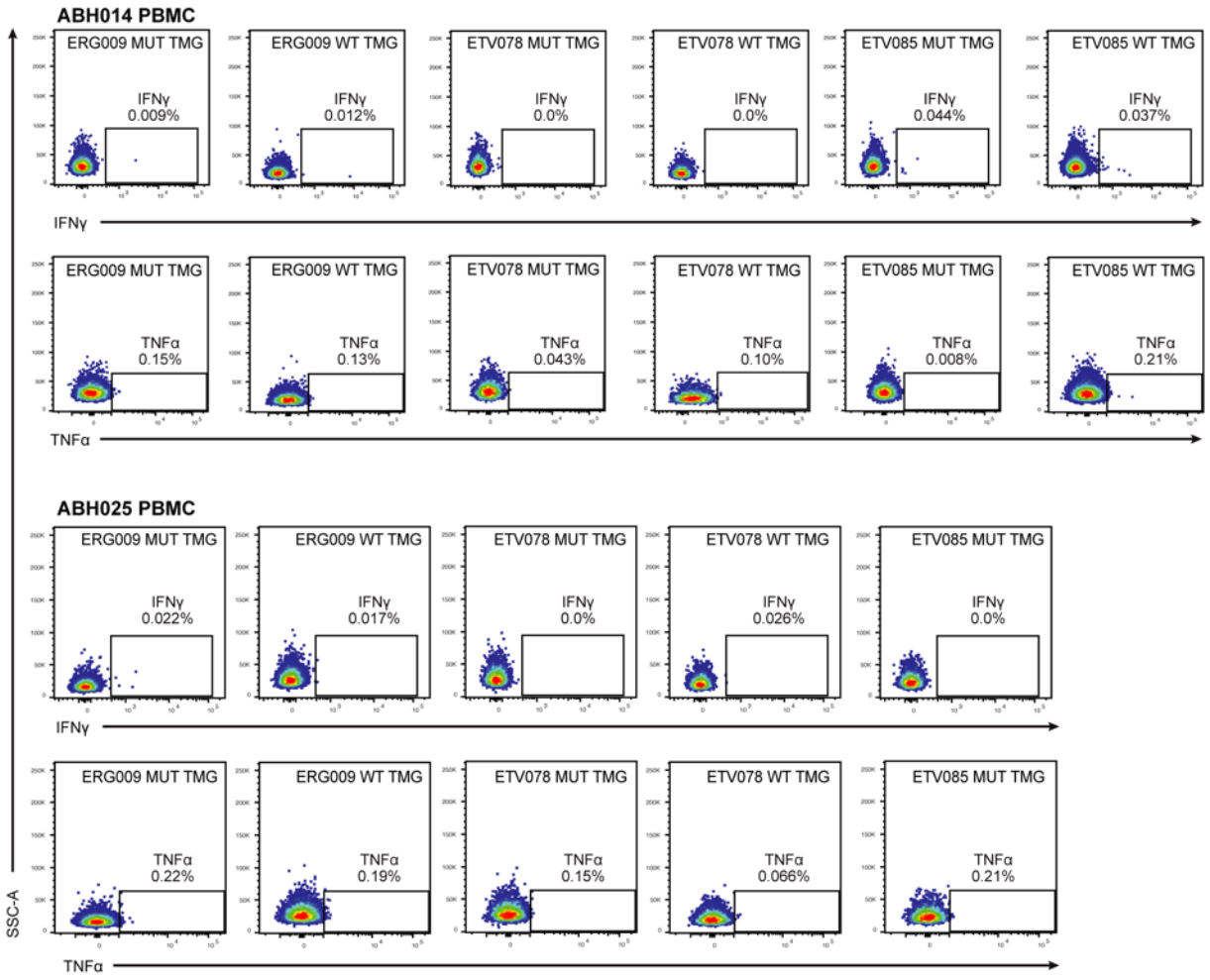

Fig. S4. Healthy donors exhibit negligible responses against endogenous neoantigens. Enriched $CD8^+$ T cells from healthy donor PBMCs (ABH014 PBMC and ABH025 PBMC) were cultured overnight in GT-T551 T cell culture media supplemented with rhIL-2 (300 IU/mL). Following overnight incubation, enriched $CD8^+$ T cells were co-cultured with RNA-transfected tumor cells from three patients at a 1:5 effector to target ratio in complete RPMI media for a total of 12 hours. Co-cultured $CD8^+$ T cells were subsequently subjected to intracellular flow cytometric analysis, and the frequency cytokine producing cells was determined. Flow dot plots depict the frequency IFNγ- and TNFα-positive $CD8^+$ T cells.


ERG009
ETV078
ETV001
ERG016
ETV084
ETV085
CD8
TETRAMER
Mutant Tet
Parent Tet
Irrelevant Tet
HLA-A*30:01 PLCD3 MUT
PLCD3(312-320)MUT 7.90%
HLA-A*30:02 PLCD3 MUT
PLCD3(311-319)MUT 11.9%
HLA-B*18:01 PLCD3 MUT
PLCD3(308-316)MUT 0.6%
HLA-B*18:01 EEF1A2 MUT
EEF1A2(104-142)MUT 1.10%
HLA-A*30:01 PLCD3 MUT
PLCD3(311-319)MUT 5.87%
HLA-B*18:01 PLCD3 MUT
PLCD3(308-317)MUT 0.5%
HLA-A*30:02 PLCD3 MUT
PLCD3(310-318)MUT 1.6%
HLA-A*30:01 PLCD3 PAR
PLCD3(312-320)PAR 1.64%
HLA-A*30:02 PLCD3 PAR
PLCD3(311-319)PAR 1.82%
HLA-A*02:01 ETV6/RUNX1
Irrelevant Tet 0.43%
HLA-A*03:01 GPR139
Irrelevant Tet 0.25%
HLA-A*24:02 CD101
Irrelevant Tet 0.06%
HLA-A*24:02 FAM157B
Irrelevant Tet 0.01%
HLA-A*24:02 TMEM104
Irrelevant Tet 0.0%
HLA-A*24:02 ACE2
Irrelevant Tet 0.0%
HLA-B*15:01 IPO13
Irrelevant Tet 0.05%
HLA-B*35:01 ANKS1A
Irrelevant Tet 0.03%
HLA-B*35:01 ASXL3
Irrelevant Tet 0.94%
HLA-B*35:01 ATP13A4
Irrelevant Tet 0.04%
HLA-B*35:01 DTX3L
Irrelevant Tet 0.01%
HLA-B*35:01 SLC6A7
Irrelevant Tet 0.0%
HLA-B*35:01 ETV6/RUNX1
Irrelevant Tet 0.0%
HLA-C*07:02 MCTP1
Irrelevant Tet 0.0%
HLA-A*03:01 PCNXL2 MUT
PCNXL2(1176-1184)MUT 0.34%
HLA-A*03:01 GPR139 MUT
GPR139(297-305)MUT 5.64%
HLA-A*03:01 GPR139 MUT
GPR139(293-301)MUT 8.38%
HLA-A*03:01 GPR139 PAR
GPR139(297-305)PAR 0.18%
HLA-A*03:01 GPR139 PAR
GPR139(293-301)PAR 0.83%
HLA-A*03:01 CMV IE1
Irrelevant Tet 0.01%
HLA-A*02:01 ETV6/RUNX1
Irrelevant Tet 0.71%
HLA-A*24:02 FAM157B
Irrelevant Tet 0.13%
HLA-A*24:02 TMEM104
Irrelevant Tet 0.12%
HLA-A*24:02 ACE2
Irrelevant Tet 0.65%
HLA-A*24:02 CD101
Irrelevant Tet 0.05%
HLA-A*30:02 PLCD3
Irrelevant Tet 0.08%
HLA-B*35:01 ASXL3
Irrelevant Tet 0.31%
HLA-B*35:01 DTX3L
Irrelevant Tet 0.0%
HLA-C*07:02 MCTP1
Irrelevant Tet 0.50%
HLA-A*02:01 SPATA5
SPATA5(807-815) Tet 5.53%
HLA-A*02:01 ETV6/RUNX1
ETV6/RUNX1 Tet 13.4%
HLA-A*02:01 VWA3A
Irrelevant Tet 0.80%
HLA-A*03:01 GPR139
Irrelevant Tet 0.01%
HLA-A*03:01 PCNXL2
Irrelevant Tet 0.0%
HLA-A*30:01 PLCD3
Irrelevant Tet 0.0%
HLA-A*30:02 PLCD3
Irrelevant Tet 0.0%
HLA-A*24:02 ACE2
ACE2(614-622) Tet 1.10%
HLA-C*07:02 MCTP1
MCTP1(779-787) Tet 0.28%
HLA-A*24:02 TMEM104
Irrelevant Tet 0.23%
HLA-A*24:02 EBV 5 MIX
Irrelevant Tet 0.1%
HLA-A*02:01 ETV6/RUNX1
Irrelevant Tet 0.18%
HLA-B*35:01 ETV6/RUNX1
Irrelevant Tet 0.0%
HLA-B*15:01 ETV6/RUNX1
Irrelevant Tet 0.06%
HLA-A*30:02 PLCD3
Irrelevant Tet 0.10%
HLA-A*30:01 PLCD3
Irrelevant Tet 0.0%
HLA-A*03:01 GPR139
Irrelevant Tet 0.07%
HLA-B*35:01 ANKS1A
ANKS1A(428-436) Tet 0.90%
HLA-B*35:01 ASXL3
ASXL3(1714-1722) Tet 1.22%
HLA-B*35:01 ATP13A4
ATP13A4(107-115) Tet 1.32%
HLA-B*35:01 SLC6A7
SLC6A7(57-65) Tet 1.56%
HLA-B*35:01 ETV6/RUNX1
ETV6/RUNX1 Tet 2.2%
HLA-B*35:01 DTX3L
DTX3L(139-147) Tet 3.6%
HLA-B*15:01 IPO13
Irrelevant Tet 0.22%
HLA-A*24:02 FAM157B
FAM157B(155-163) Tet 1.1%
HLA-A*24:02 TMEM104
TMEM104(276-284) Tet 2.2%
HLA-B*15:01 ETV6/RUNX1
ETV6/RUNX1 Tet 0.05%
HLA-B*15:01 IPO13
IPO13(267-275) Tet 0.2%
HLA-B*15:01 CD101
CD101(884-892) Tet 1.9%
HLA-A*24:02 ACE2
Irrelevant Tet 0.18%

Fig. S5. Neoepitope tetramers are patient-specific and exhibit negligible nonspecific binding. Bone marrow cells from each patient were independently stained using an antibody cocktail and either an irrelevant (either nonself-antigen bound to patient-specific HLA or mutated neoantigen from a different patient), parent (wild-type self-peptide bound to patient-specific HLA), or mutant (cancer neoantigen bound to patient-specific HLA) tetramers. Flow dot plots depict the frequency of $CD8^+$ TILs binding either a parent (blue shaded dot plots), irrelevant (gray shaded dot plots), or mutant (dot plots without shading) tetramer across all patients. Above each dot plot is a label indicating the specific tetramer (HLA allele and peptide) used for staining.

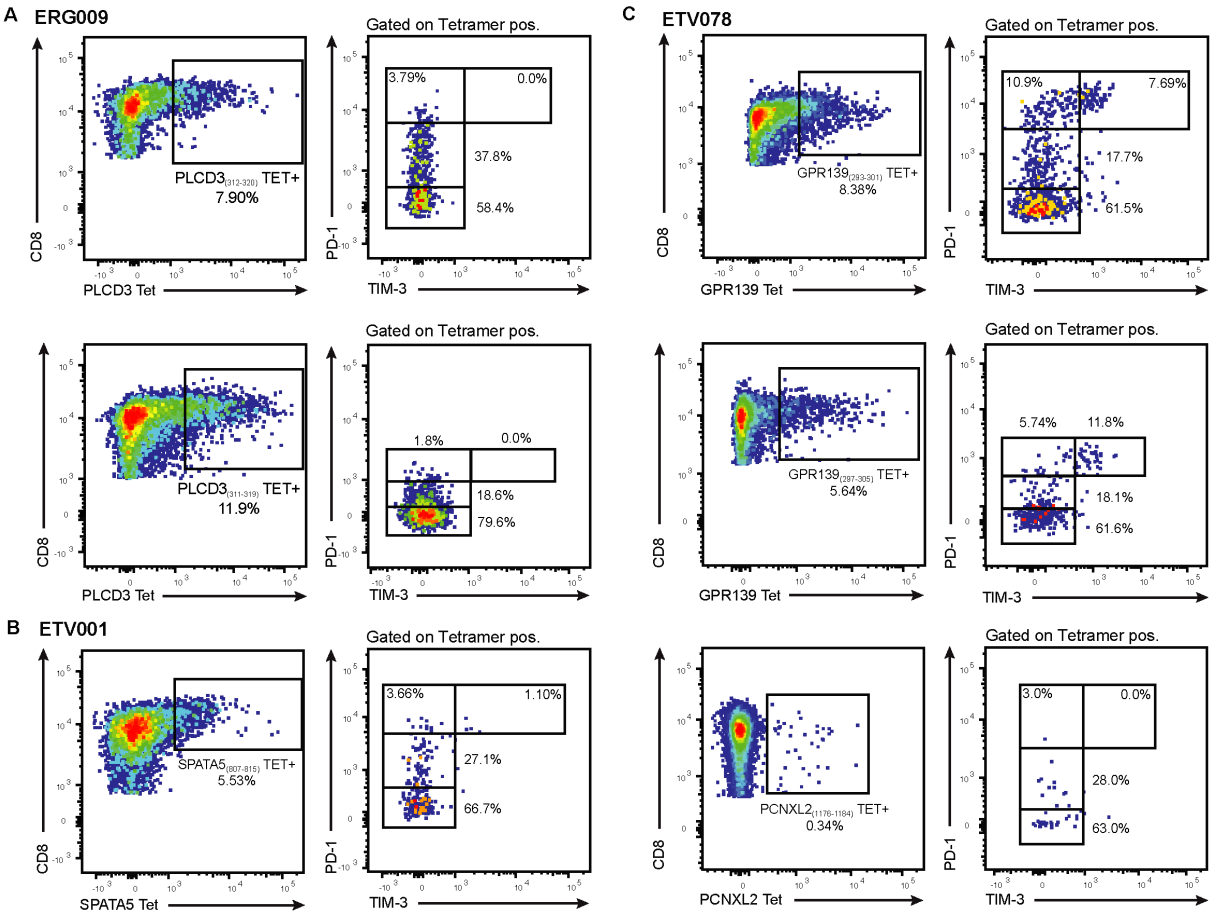

Fig. S6. Phenotypic characterization of neoepitope-specific CD8$^+$ TILs. PD-1^{low}, PD-1^{mid}, PD-1^{high}, and PD-1^{high}/TIM-3^{high} expression on tetramer-positive CD3$^+$CD8$^+$ TILs from patients ERG009 (A), ETV001 (B), and ETV078 (C) determined by flow cytometric analysis.

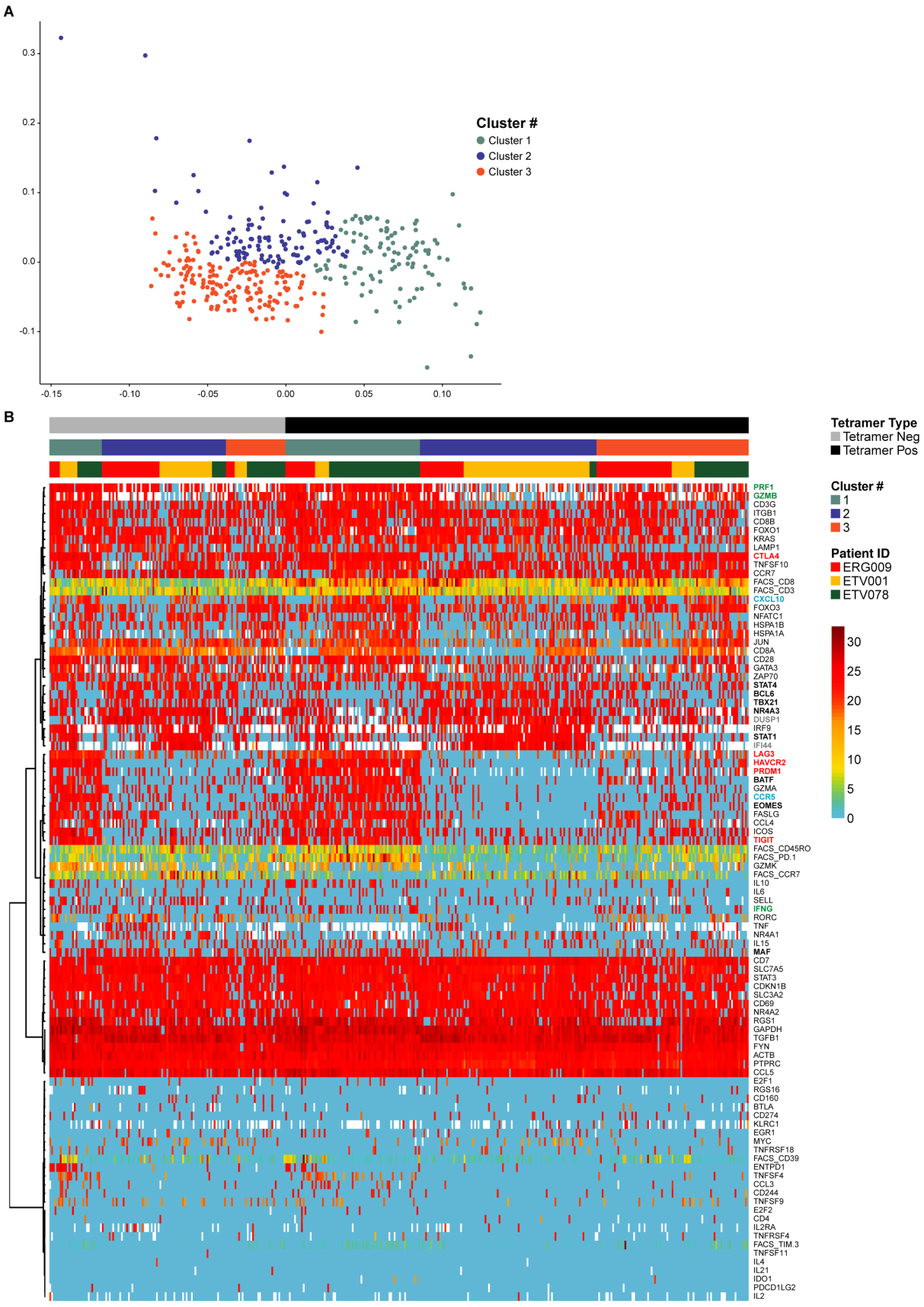
A
0.3
0.2
0.1
0.0
-0.1
-0.15
-0.10
-0.05
0.00
0.05
0.10
Cluster #
Cluster 1
Cluster 2
Cluster 3
B
Tetramer Type
Tetramer Neg
Tetramer Pos
Cluster #
1
2
3
Patient ID
ERG009
ETV001
ETV078
30
25
20
15
10
5
0
PRF1
GZMB
CD3G
ITGB1
CD8B
FOXO1
KRAS
LAMP1
CTLA4
TNFSF10
CCR7
FACS_CD8
FACS_CD3
CXCL10
FOXO3
NFATC1
HSPA1B
HSPA1A
JUN
CD8A
CD28
GATA3
ZAP70
STAT4
BCL6
TBX21
NR4A3
DUSP1
IRF9
STAT1
IFI44
LAG3
HAVCR2
PRDM1
BATF
GZMA
CCR5
EOMES
FASLG
CCL4
ICOS
TIGIT
FACS_CD45RO
FACS_PD.1
GZMK
FACS_CCR7
IL10
IL6
SELL
IFNG
RORC
TNF
NR4A1
IL15
MAF
CD7
SLC7A5
STAT3
CDKN1B
SLC3A2
CD69
NR4A2
RGS1
GAPDH
TGFB1
FYN
ACTB
PTPRC
CCL5
E2F1
RGS16
CD160
BTLA
CD274
KLRC1
EGR1
MYC
TNFRSF18
FACS_CD39
ENTPD1
TNFSF4
CCL3
CD244
TNFSF9
E2F2
CD4
IL2RA
TNFRSF4
FACS_TIM.3
TNFSF11
IL4
IL21
IDO1
PDCD1LG2
IL2

Fig. S7. Patient-specific transcriptional profiles. (A) Projection of principal components from single-cell gene-expression data derived from tetramer-negative and tetramer-positive sorted $CD45RO^{+}CD8^{+}$ TIL subsets from patients ERG009, ETV001, and ETV078. Each circle represents an individual cell and each color represents the corresponding cluster number. (B) Heatmap visualizing unscaled expression of genes (transcript expression threshold values; Et) and scaled surface protein data (MFI) for sorted single-cell tetramer-negative and tetramer-positive $CD45RO^{+}CD8^{+}$ TILs. Top margin color bars represent, from top to bottom, tetramer type, hierarchical cluster number (clusters 1-3), and patient IDs (ERG009, ETV001, and ETV078), respectively. Bolded gene name colors represent: transcription factors (black), inhibitory receptors (red), functional molecules (green), chemokine/chemokine receptors (blue), and transcriptional regulators (gray).

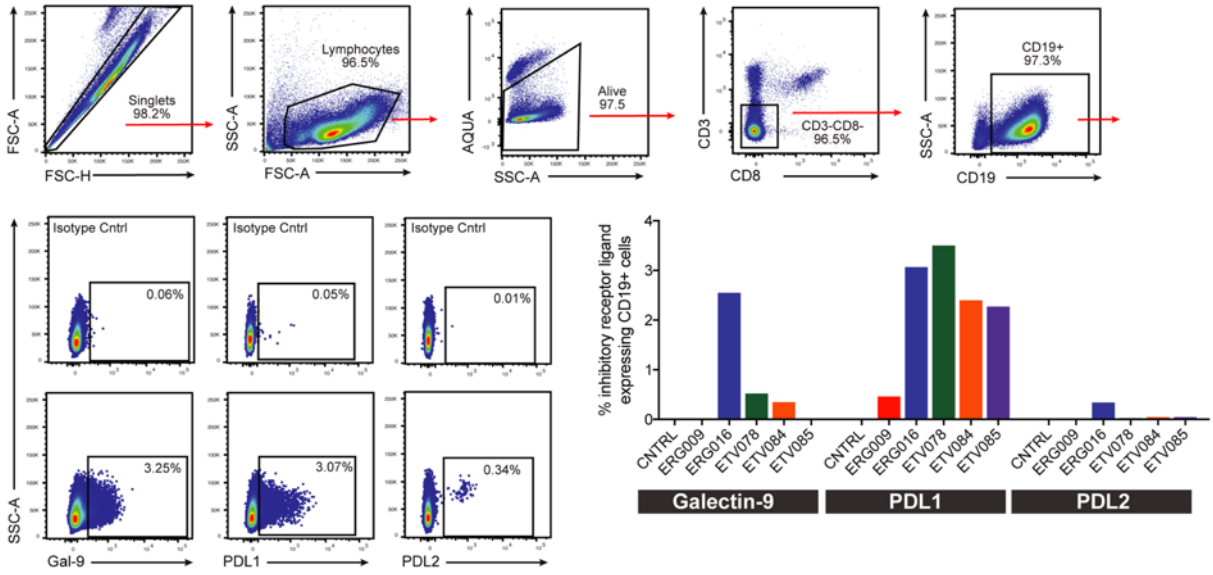

Fig. S8. Phenotypic characterization of patient-specific CD19^{+} tumor cells. Galectin-9, PDL1, and PDL2 co-inhibitory receptor ligand expression on CD19^{+} tumor cells from ERG009, ERG016, ETV078, ETV084, and ETV085 or CD19^{+} BMMCs from control samples determined by flow cytometric analysis.

Table S1. Patient ALL subtypes

PCGP ID	Sample Type	ALL Subtype	Comments	Rearrangement (fusion)	Prevalence (%)	Prognosis
SJETV001	Diagnostic	ETV-associated ALL	This is a form of ALL associated with a fusion between ETV-6 and AML-1; t(12;21)(p13;q22) translocation encoding ETV6-RUNX1	ETV6-RUNX1	15-25	Excellent prognosis
SJETV022	Diagnostic					
SJETV026	Diagnostic					
SJETV035	Diagnostic					
SJETV078	Diagnostic					
SJETV084	Diagnostic					
SJETV085	Diagnostic					
SJETV096	Diagnostic					
SJETV098	Diagnostic					
SJERG009	Diagnostic	ETS-Related Gene-Associated ALL	ALL with alterations in ERG; typically a distinct gene expression profile with the majority of patients having a focal ERG deletions	IGH-DUX4	~7	Typically favorable outcome
SJERG016	Diagnostic					

Table S2. FPKM values for HLA genes

	Gene			Tumor Milieu		
	HLA-A	HLA-B	HLA-C	% of Blasts (tumor cells)	% of Monocytic cells	% of T cells
SJERG009	206.764636	620.381172	411.862947	89.0%	1.0%	2.0%
SJERG016	111.185572	216.26332	132.588667	89.0%	2.0%	1.0%
SJETV001	306.695398	468.978469	385.35994	83.0%	2.0%	3.0%
SJETV078	466.678669	587.476285	703.273614	86.0%	1.0%	2.5%
SJETV084	129.479783	251.819536	234.751251	93.0%	3.5%	1.0%
SJETV085	205.053111	309.947074	326.763993	82.0%	1.0%	1.0%

Table S3. Patient-specific neoepitopes

PCGP ID	HLA Haplotype	Variant	Mutated Gene	Predicted High Affinity Neoepitopes	Corresponding HLA	Predicted Affinity (IC_{50})
SJETV001	A*02:01 A*33:03 B*58:01 B*27:05 C*03:02 C*02:02	MCTP1_A899V	Multiple C2 and Transmembrane Domain Containing 1	NILDEVVSF	HLA-C03:02	311.19
		SPATA5_S812C	Spermatogenesis Associated 5	**KLQFHCMPV**	**HLA-A02:01**	**8.29**
		ETV6-RUNX1	ETS Variant 6/Runt Related Transcription Factor 1	**RIAECILGM**	**HLA-A02:01**	**192.18**
				RIAECILGM	HLA-C03:02	80.28
		TECPR2_E1307A	Tectonin Beta-propeller Repeate Containing 2	TAWAHVPGL	HLA-C03:02	70.31
				WAHVPGLQA	HLA-C03:02	396.32
SJETV078	A*34:02 A*03:01 B*53:01 B*15:03 C*02:10 C*04:01	AHNAK_S5863F	Neuroblast Differentiation-assoicated Protein AHNAK Isoform 1	LASKKSRLF	HLA-B15:03	66.92
		BTBD16_V298M	BTB Domain Containing 16	QAIPTYETM	HLA-B15:03	60.38
				PTYETMMTF	HLA-B15:03	138.91
				TMMTFFKSF	HLA-B15:03	3.17
		GPR139_A298T	G Protein-couple Receptor 139	FRTMAATTL	HLA-B15:03	247.81
				RTMAATTLK	HLA-A34:02	41.19
				RTMAATTLK	**HLA-A03:01**	**21.03**
				ATTLKAFFK	**HLA-A03:01**	**93.08**
				ATTLKAFFK	HLA-A34:02	71.22
				MAATTLKAF	HLA-B53:01	121.34
				MAATTLKAF	HLA-B15:03	22.19
		ITPR1_V656L	Inositol 1,4,5-Triphosphate Receptor, Type 1	KLLLSRFEF	HLA-B15:03	34.77
		PCNXL2_F1180L	Pecanex-like 2 (Drosophila)	HLMWLERLY	HLA-A34:02	374.58
				HLMWLERLY	HLA-B15:03	232.24
				HLMWLERLY	**HLA-A03:01**	**390.37**
				LMWLERLYV	HLA-B15:03	141.95
		PRRC2B_R923W	Proline-Rich Coiled-Coil 2B	EPEWTPEPW	HLA-B53:01	51.89
		SCD_E240K	Stearoyl-CoA Desaturase	LVPWYFWGK	HLA-A34:02	338.02
				GKTFQNSVF	HLA-B15:03	28.78
SJETV084	A*02:02 A*02:02 B*42:01 B*35:01 C*17:01 C*04:01	ANKS1A_E434K	Ankyrin Repeat and SAM Domain-containing Protein 1A	FPLTASKVL	HLA-B42:01	12.78
				FPLTASKVL	**HLA-B35:01**	**20.69**
		ARHGAP12_S114L	Rho GTPase Activating Protein 12	KLPELLSFG	HLA-A02:02	207.3
		ASXL3_E1718K	Additional Sex Combs Like 3 (Drosophila)	SPMEKAISL	HLA-B42:01	21.25
				SPMEKAISL	**HLA-B35:01**	**96.68**
		ATP13A4_L109F	ATPase type 13A4; Probable Cation-Transporting ATPase 13A4	FGLTPDHPF	HLA-B35:01	153.95
				GLTPDHPFM	HLA-A02:02	279.13
				HPFMTDEEY	**HLA-B35:01**	**4.22**
				FMTDEEYII	HLA-A02:02	6.94
		CCDC108_R1124Q	Coiled-Coil Domain-Containing Protein 108	ITQKHLWRL	HLA-A02:02	476.88
		DTX3L_D141H	Deltex E3 Ubiquitin Ligase 3L	VTAHLNCNL	HLA-A02:02	319.56
				TAHLNCNLF	**HLA-B35:01**	**348.45**
		KIAA1715_R253Q	Lunapark, ER Junction Formation Factor	ILPREQGAL	HLA-A02:02	312.72
		KRT7_E263Q	Keratin 7	DLDGIIAQV	HLA-A02:02	253.23
		NNMT_L45V	Nicotinamide N-Methyltransferase	ILKHLLKNV	HLA-A02:02	153.12
				HLLKNVFKI	HLA-A02:02	126.7
		OR5H15_L45M	Olfactory Receptor Family 5 Subfamily H Member 15	TIMGNLGMI	HLA-A02:02	363.86
		PDS5B_D86H	Sister Chromatid Cohesion Protein PDS5 Homolog B	LVACCLAHI	HLA-A02:02	21.14
		RBCK1_L130V	RanBP-type and C3HC4-type Zinc Finger-Containing Protein 1 Isoform 2	VLSARNTSL	HLA-A02:02	15.03
		SEPT14_S82P	Septin 14	KPSHFYSNV	HLA-B42:01	74.97
		SERPINA6_E138Q	Serpin Family A Member 6	FLDGSLQLL	HLA-A02:02	3.93
		SLC6A7_R62C	Solute Carrier Family 6 Member 7	**LGNVWCFPY**	**HLA-B35:01**	**107.72**
		SULT4A1_T205M	Sulfotransferase Family 4A Member 1	MMVEQLARF	HLA-A02:02	165.16
				MMVEQLARF	HLA-B35:01	140.42
		USH2A_S1007L	Usher Syndrome 2A	HLLGALNET	HLA-A02:02	265.87
		ETV6-RUNX1	ETS Variant 6/Runt Related Transcription Factor 1	**MPIGRIAEC**	**HLA-B35:01**	**492.62**
				MPIGRIAEC	HLA-B42:01	228.5
				RIAECILGM	HLA-A02:02	93.08
SJETV085	A*24:02 A*68:02 B*53:01 B*15:01 C*03:03 C*04:01	AZI1_M681I	5-Azacytidine Induced 1; Centrosomal Protein of 131 kDa Isoform A	ELISATEKA	HLA-A68:02	104.28
		CD101_S889F	CD101 Molecule	**HLHCYRSSF**	**HLA-B15:01**	**28.93**
				CYRSSFTDF	HLA-A24:02	54.78
		FAM157B_L156F	Family With Sequence Similarity 157 Mmeber B	**RFLHDLHLL**	**HLA-A24:02**	**178.16**
		GRIN3A_A401T	Glutamate Receptor, Ionotropic, N-Methyl-D-Aspartate 3A	LVARAVTTA	HLA-A68:02	339.15
				RAVTTATMI	HLA-C03:03	59.73
		IPO13_I267M	Importin 13	**MSQPDAQRY**	**HLA-B15:01**	**267.31**
		SPIRE1_R332W	Protein Spire Homolog 1 Isoform B	MVNGDIPPW	HLA-B53:01	22.32
		TMEM104_V281I	Transmembrane Protein 104	**LFGVCIYSF**	**HLA-A24:02**	**110.68**
				FGVCIYSFM	HLA-A68:02	241.2
				FGVCIYSFM	HLA-C03:03	72.97
		ETV6-RUNX1	ETS Variant 6/Runt Related Transcription Factor 1	**RIAECILGM**	**HLA-B15:01**	**291.48**
SJERG009	A*30:02 A*30:01 B*18:01 B*13:02 C*06:02 C*05:01	PLCD3_D314H	Phospholipase C, Delta 3	**HELMTLHGF**	**HLA-B18:01**	**12.78**
				TLHGFMMYL	**HLA-A30:01**	**398.91**
				MTLHGFMMY	**HLA-A30:01**	**398.91**
				MTLHGFMMY	**HLA-A30:02**	**32.24**
				***HELMTLHGFM**	**HLA-B18:01**	**187.05**
				***LMTLHGFMMY**	**HLA-A30:02**	**127.39**
		EEF1A2_T142M	Eukaryotic Translation Elongation Factor 1 Alpha 2	**REHALLAYM**	**HLA-B18:01**	**412.08**
SJERG016	A*24:02 A*02:07 B*58:01 B*40:01 C*03:02 C*07:02	MCTP1_R783H	Multiple C2 and Transmembrane Domain-Containing Protein 1 Isoform L	FIRMKHCVM	HLA-C03:02	26.17
				IRMKHCVMV	**HLA-C07:02**	**342.84**
				RMKHCVMVL	HLA-C03:02	342.57
		PIGZ_P455S	Phosphatidylinositol Glycan Anchor Biosynthesis, Class Z; GPI Mannosyltransferase 4	TSTHYTLLF	HLA-C03:02	99.38
				TSTHYTLLF	HLA-B58:01	45.09
				STSTHYTLL	HLA-C03:02	284.78
		ACE2_D615Y	Angiotensin I Converting Enzyme 2	STDWSPYAY	HLA-C03:02	130.01
				AYQSIKVRI	**HLA-A24:02**	**91.58**
		FMOD_I117T	Fibromodulin	FQNNQTTSI	HLA-C03:02	27.49
		NUDT10_A62V	Nudix Hydrolase 10	EEEPGGAVV	HLA-B40:01	359.95

*denontes neoepitopes with 10mer amino acids
Bold denontes neoepitopes with corresponding tetramers that were generated and tested
For ETV6-RUNX1 neoepitopes, red color indicates the amino acids corresponding to ETV6 and the blue color indicates the amino acids corresponding to RUNX1

Table S4. Tetramers used for specificity assays

Patient ID	Variant	Peptide	Peptide Length	HLA Allele	Patient-specific HLA?	Tetramer Type
ERG009	EEF1A2_T142M	REHALLAYM	9	HLA-B18:01	Yes	Mutant
	PLCD3_D314H	HELMTLHGF	9	HLA-B18:01	Yes	Mutant
	PLCD3_D314H	MTLHGFMMY	9	HLA-A30:01	Yes	Mutant
	PLCD3_D314H	TLHGFMMYL	9	HLA-A30:01	Yes	Mutant
	PLCD3_D314H	MTLHGFMMY	9	HLA-A30:02	Yes	Mutant
	PLCD3_D314H	HELMTLHGFM	10	HLA-B18:01	Yes	Mutant
	PLCD3_D314H	LMTLHGFMMY	10	HLA-A30:02	Yes	Mutant
	PLCD3_D314H	TLDGFMMYL	9	HLA-A30:01	Yes	Parent
	PLCD3_D314H	MTLDGFMMY	9	HLA-A30:02	Yes	Parent
	ETV6-RUNX1	RIAECILGM	9	HLA-A02:01	No	Irrelevant
	GPR139_A298T	ATTLKAFFK	9	HLA-A03:01	No	Irrelevant
	CD101_S889F	HLHCYRSSF	9	HLA-B15:01	No	Irrelevant
	FAM157B_L156F	RFLHDLHLL	9	HLA-A24:02	No	Irrelevant
	TMEM104_V281I	LFGVCIYSF	9	HLA-A24:02	No	Irrelevant
	ACE2_D615Y	AYQSIKVRI	9	HLA-A24:02	No	Irrelevant
	IPO13_I267M	MSQPDAQRY	9	HLA-B15:01	No	Irrelevant
	ANKS1A_E434K	FPLTASKVL	9	HLA-B35:01	No	Irrelevant
	ASXL3_E1718K	SPMEKAISL	9	HLA-B35:01	No	Irrelevant
	ATP13A4_L109F	HPFMTDEEY	9	HLA-B35:01	No	Irrelevant
	DTX3L_D141H	TAHLNCNLF	9	HLA-B35:01	No	Irrelevant
	SLC6A7_R62C	LGNVWCFPY	9	HLA-B35:01	No	Irrelevant
	ETV6-RUNX1	MPIGRIAEC	9	HLA-B35:01	No	Irrelevant
	MCTP1_R783H	IRMKHCVMV	9	HLA-C07:02	No	Irrelevant
ERG016	ACE2_D615Y	AYQSIKVRI	9	HLA-A24:02	Yes	Mutant
	MCTP1_R783H	IRMKHCVMV	9	HLA-C07:02	Yes	Mutant
	TMEM104_V281I	LFGVCIYSF	9	HLA-A24:02	Yes	Irrelevant
	EBV 5 Mix	5 peptide mix	9	HLA-A24:02	Yes	Irrelevant
	ETV6-RUNX1	RIAECILGM	9	HLA-A02:01	No	Irrelevant
	ETV6-RUNX1	MPIGRIAEC	9	HLA-B35:01	No	Irrelevant
	ETV6-RUNX1	RIAECILGM	9	HLA-B15:01	No	Irrelevant
	PLCD3_D314H	TLHGFMMYL	9	HLA-A30:01	No	Irrelevant
	PLCD3_D314H	MTLHGFMMY	9	HLA-A30:02	No	Irrelevant
	GPR139_A298T	ATTLKAFFK	9	HLA-A03:01	No	Irrelevant
ETV001	SPATA5_S812C	KLQFHCMPV	9	HLA-A02:01	Yes	Mutant
	ETV6-RUNX1	RIAECILGM	9	HLA-A02:01	Yes	Mutant
	VWA3A_V955I	RLFGTILES	9	HLA-A02:01	Yes	Irrelevant
	GPR139_A298T	ATTLKAFFK	9	HLA-A03:01	No	Irrelevant
	PCNXL2_F1180L	HLMWLERLY	9	HLA-A03:01	No	Irrelevant
	PLCD3_D314H	TLHGFMMYL	9	HLA-A30:01	No	Irrelevant
	PLCD3_D314H	MTLHGFMMY	9	HLA-A30:02	No	Irrelevant
ETV078	GPR139_A298T	RTMAATTLK	9	HLA-A03:01	Yes	Mutant
	GPR139_A298T	ATTLKAFFK	9	HLA-A03:01	Yes	Mutant
	PCNXL2_F1180L	HLMWLERLY	9	HLA-A03:01	Yes	Mutant
	GPR139_A298T	RTMAAATLK	9	HLA-A03:01	Yes	Parent
	GPR139_A298T	AATLKAFFK	9	HLA-A03:01	Yes	Parent
	CMV IE1	KLGGALQAK	9	HLA-A03:01	Yes	Irrelevant
	ETV6-RUNX1	RIAECILGM	9	HLA-A02:01	No	Irrelevant
	FAM157B_L156F	RFLHDLHLL	9	HLA-A24:02	No	Irrelevant
	TMEM104_V281I	LFGVCIYSF	9	HLA-A24:02	No	Irrelevant
	ACE2_D615Y	AYQSIKVRI	9	HLA-A24:02	No	Irrelevant
	CD101_S889F	HLHCYRSSF	9	HLA-B15:01	No	Irrelevant
	PLCD3_D314H	MTLHGFMMY	9	HLA-A30:02	No	Irrelevant
	ASXL3_E1718K	SPMEKAISL	9	HLA-B35:01	No	Irrelevant
	DTX3L_D141H	TAHLNCNLF	9	HLA-B35:01	No	Irrelevant
ETV084	ANKS1A_E434K	FPLTASKVL	9	HLA-B35:01	Yes	Mutant
	ASXL3_E1718K	SPMEKAISL	9	HLA-B35:01	Yes	Mutant
	ATP13A4_L109F	HPFMTDEEY	9	HLA-B35:01	Yes	Mutant
	DTX3L_D141H	TAHLNCNLF	9	HLA-B35:01	Yes	Mutant
	SLC6A7_R62C	LGNVWCFPY	9	HLA-B35:01	Yes	Mutant
	ETV6-RUNX1	MPIGRIAEC	9	HLA-B35:01	Yes	Mutant
	IPO13_I267M	MSQPDAQRY	9	HLA-B15:01	No	Irrelevant
ETV085	CD101_S889F	HLHCYRSSF	9	HLA-B15:01	Yes	Mutant
	FAM157B_L156F	RFLHDLHLL	9	HLA-A24:02	Yes	Mutant
	IPO13_I267M	MSQPDAQRY	9	HLA-B15:01	Yes	Mutant
	TMEM104_V281I	LFGVCIYSF	9	HLA-A24:02	Yes	Mutant
	ETV6-RUNX1	RIAECILGM	9	HLA-B15:01	Yes	Mutant
	ACE2_D615Y	AYQSIKVRI	9	HLA-A24:02	Yes	Irrelevant

Table S5. Fluidigm primer list

Target	Forward Primer Sequence	Reverse Primer Sequence
ACTB	GCCGTCTTCCCCTCCA	CTCGTCGCCCACATAGGAA
BATF	AGCAGTGACTCCAGCTTCA	CTCTTCTGGGCGGCAATAC
BCL6	GATGGAGCATGTTGTGGACAC	AGGAGGCTTGATGGCAGAAA
BTLA	TCCCATATCTGGACATCTGGAAC	CTCCTGCTAAGATGGAGTGTTCA
CCL3	GAGCAGCCAGTGCTCCAA	AGCTCCAGGTCGCTGACATA
CCL4	GTAGCTGCCTTCTGCTCTCC	TCTACCACAAAGTTGCGAGGAA
CCL5	CCCTCGCTGTCATCCTCA	GGGCAATGTAGGCAAAGCA
CCR5	TGAGACATCCGTTCCCCTACA	TGGCAGGGCTCCGATGTATA
CCR7	TGAGGTCACGGACGATTACA	CGTCCTTCTTGGAGCACAAA
CD160	CTCAGTTCAGGCTTCCTACA	TCTTTTGGCACAAGGCTTAC
CD244	AACCACAGCCCTTCCTTCAA	GAGCAGGGTTCTGGGCTTTA
CD274	ACCAGCCGCGCTTCTGT	TCAGCAAATGCCAGTAGGTCATGAAT
CD28	GTGGAGTCCTGGCTTGCTATA	GAGCCTGCTCCTCTTACTCC
CD3G	GCCCAATGACCAGCTCTA	TTCCTCCTCAACTGGTTTCC
CD4	AAAGTTGCATCAGGAAGTGAACC	CCCACACCTCACAGGTCAAA
CD69	TCACCCATGGAAGTGGTCAA	ACACACTTGTCAGACCCTGTA
CD7	GCCATCACGGAGGTCAATGT	AGCATCTGTGCCATCCTTGG
CD8A	CCATCATGTACTTCAGCCACTTCG	GCTGCGACGCGATGGT
CD8B	CCGGAAGACAGTGGCATCTA	CTGGGCAGTGGTGGGAA
CDKN1B	GCAATGCGCAGGAATAAGGAA	TTGGGGAACCGTCTGAAACA
CTLA4	CATGGACACGGGACTCTACA	AATCTGGGTTCCGTTGCCTA
CXCL10	GCTGTACCTGCATCAGCATTA	CTGGATTCAGACATCTCTTCTCAC
DUSP1	AGACATCAGCTCCTGGTTCA	CAGTGGACAAACACCCTTCC
E2F1	AGCTCATTGCCAAGAAGTCCAA	TCCTGGGTCAACCCCTCAA
E2F2	TTCAAGCACCTGACTGAGGAC	AGTTGCCAACAGCACGGATA
EGR1	AACCCTCAGGCGGACAC	CAGCACCTTCTCGTTGTTCA
ENTPD1	AGGTGCCTATGGCTGGATTAC	GTCCAAAGCTCCAAAGGTTTCC
EOMES	CTGTGGCAAAGCCGACAATA	CTCATCCAGTGGGAACCAGTA
FASLG	CTGAGGAAAGTGGCCCAT	ACAAAGTACAGCCCAGTTTCA
FOXO1	GGTGTCAGGCTGAGGGTTA	TTCTCTCAGTTCCTGCTGTCA
FOXO3	AGCTTCCCGTATACCACCAA	CCGAACACCGTGCTGTTAAA
FYN	GAGCCCATCTACATCGTCAC	TTCAGAGCTCTTCCTTCTCCA
GAPDH	GAACGGGAAGCTTGTCATCAA	ATCGCCCCACTTGATTTTGG
GATA3	CACGGTGCAGAGGTACCC	AGGGTAGGGATCCATGAAGCA
GZMA	CCTCCGAGGTGGAAGAGAC	GTGACCCCTCGGAAAACAC
GZMB	CTTCTCCAACGACATCATGCTAC	CTGGGCCTTGTTGCTAGGTA
GZMK	ATCCACAGTGGGTGCTGAC	AGAGTGTGCGCCTAAAACCA
HAVCR2	GGATCCAAATCCCAGGCATAA	CTTGGAAAGGCTGCAGTGAA
HSPA1A	AGGCTTCCCAGAGCGAAC	GAGAAGAGCTCGGTCCTTCC
HSPA1B	CAAGGCTTTCCAGAGCGAAC	AAGAGCTCAGTCCTTCGGAAC
ICOS	GCCAACTATTACTTCTGCAACC	GAACTTCAGCTGGCAACAAA
IDO1	GGATGCATCACCATGGCATA	TTGGCAGTAAGGAACAGCAA
IFI44	GGCTTTGGTGGGCACTAATA	TGCCATCTTTCCCGTCTCTA
IFNG	ACTGCCAGGACCCATATGTAA	GTTCCATTATCCGCTACATCTGAA
IL10	CCGTGGAGCAGGTGAAGAA	GTCAAACTCACTCATGGCTTTGTA
IL15	GTCCGGAGATGCAAGTATTCA	TCCTCACATTCTTTGCATCCA
IL2	ACCCAGGGACTTAATCAGCAA	GCATATTCACACATGAATGTTGTTTCA
IL21	CTGAATTTCTGCCAGCTCCA	TTGTTTCCTGTATTTGCTGACTTTA
IL2RA	TCCTGGGACAACCAATGTCA	GTCACTTGTTTCGTTGTGTTCC
IL4	CAGCTGATCCGATTCCTGAAA	GTTGGCTTCCTTCACAGGAC
IL6	AGAGCTGTGCAGATGAGTACAA	GTTGGGTCAGGGGTGGTTA
IRF9	CTCCAGCCATACTCCACAGAA	GGAGTCTGCTCCAGCAAGTA
ITGB1	GAATGTATACAAGCAGGGCCAAA	CGTGCAGAAGTAGGCATTCC
JUN	AAGAACTCGGACCTCCTCAC	TGGATTATCAGGCGCTCCA
KLRC1	AGCTTCTCAGGATTTTCAAGGGAA	GCTTCTCTGGAGCTGATGGTAA
KRAS	GGGAGGGCTTTCTTTGTGTA	ACTAGGACCATAGGTACATCTTCA
LAG3	TGGAGCCTTTGGCTTTCAC	GAGGGTGAATCCCTTGCTCTA
LAMP1	TCCAGGCTTTCAAGGTGGAA	CCACAGCGATGGGGATCA
MAF	TCGACGACCGCTTCTCC	ATCACCTCCTCCTTGCTGAC
MYC	CCTGGTGCTCCATGAGGA	CCTGCCTCTTTTCCACAGAAA
NFATC1	TCCTCTCCAACACCAAAGTCC	AGGATTCCGGCACAGTCAA
NR4A1	CCAGTTCTGCCGCTTCCA	GCTTGGGTTTTGAAGGTAGCC
NR4A2	TGGCTGTTGGGATGGTCAAA	TCTTCGGTTTCGAGGGCAAA
NR4A3	ACGTCGAAACCGATGTCAGTA	ACCTCTCCTCCCTTTCAGACTA
PDCD1LG2	GGAATTGCAGCTTCACCAGATA	CACATTGCTGCCATGCTCTA
PRDM1	CCTGGTACACACGGGAGAAAA	TTGAGATTGCTGGTGCTGCTA
PRF1	GTACAGCTTCAGCACTGACAC	CTGGGTGGAGGCGTTGAA
PTPRC	GTGGCTTAAACTCTTGGCATTT	GGGAAGGTGTTGGGCTTT
RGS1	TGCCCTGTAAAGCAGAAGAGATA	TTCTCGAGTGCGGAAGTCAA
RGS16	TCGAGTGGGGCAGTAAACAC	AGGTCGAACGACTCTCTCCA
RORC	CAAGACTCATCGCCAAAGCA	TTTCCACATGCTGGCTACAC
SELL	ACTATGGGCCCCAGTGTCA	GTGAGTACAGTCCATGGTACCC
SLC3A2	TGGCTGAGTGGCAAAATATCAC	GCTGAAGGTCGGAGGAGTTA
SLC7A5	TTCGGGGTCTGGTGGAAAA	CCTGCATGAGCTTCTGACAC
STAT1	ATGCTGGCACCAGAACGAA	GCTGGCACAATTGGGTTTCAA
STAT3	GGAAATAATGGTGAAGGTGCTGAAC	CCGAGGTCAACTCCATGTCAAA
STAT4	CAGTGCTGGAGGTAAAGGAA	AGAGGCAGATCTGTGTTTCAA
TBX21	GGGCGTCCAACAATGTGAC	CCGTCGTTCACCTCAACGATA
TGFB1	CTACTACGCCAAGGAGGTCAC	GCTGTGTGTACTCTGCTTGAAC
TIGIT	GTGGTGGTCGCGTTGACTA	TCCTGTCCAGCTGATTTTCTCC
TNF	CTTCTCGAACCCCGAGTGAC	ACTGGAGCTGCCCCTCA
TNFRSF18	GCTGCTGCCGCGATTA	GAATTCAGGCTGGACACACA
TNFRSF4	AACGACGTGGTCAGCTCCAA	CACAGCTGCTTCCGCTCAC
TNFSF10	AGAAGGAAGGGCTTCAGTGAC	CCTGGACCTCCATCATAGCC
TNFSF11	CTCAGCCTTTTGCTCATCTCAC	GGTACCAAGAGGACAGACTCAC
TNFSF4	CACATCGGTATCCTCGAATTCA	TGATTTCATCCTCCTTTTGGGAA
TNFSF9	CTTTCGCCCGACGATCC	ACCAGCTGCGCAAACA
ZAP70	AGTGTCCACCCGAACTGTAC	CTGCTCCACGGTCAGGAA

Table S6. Peptides for functional assays

Patient ID	Sample Type	Mutation Type	Variant	Wild-type (parent) AA	Mutant AA	AA mutation position	Wild-type (parent) 15mer	Mutated 15mer
SJERG009	Diagnostic	SNV	EEF1A2_T142M	T	M	142	GQTREHALLAYTLGV	GQTREHALLAYMLGV
		SNV	PLCD3_D314H	D	H	314	KQHELMTLDGFMMYL	KQHELMTLHGFMMYL
SJERG016	Diagnostic	SNV	ACE2_D615Y	D	Y	615	GWSTDWSPYADQSIK	GWSTDWSPYAYQSIK
							SPYADQSIKVRISLK	SPYAYQSIKVRISLK
		SNV	FMOD_I117T	I	T	117	VYFFQNNQITSIQEG	VYFFQNNQTTSIQEG
		SNV	MCTP1_R783H	R	H	783	LRNFIRMKRCVMVLV	LRNFIRMKHCVMVLV
		SNV	NUDT10_A62V	A	V	62	EPEEEPGGAAVREVY	EPEEEPGGAVVREVY
		SNV	PIGZ_P455S	P	S	455	PVLPSTPTHYTLLFT	PVLPSTSTHYTLLFT
SJETV001	Diagnostic	SNV	MCTP1_A899V	A	V	899	VQNILDEVASFGERI	VQNILDEVVSFGERI
		SNV	SPATA5_S812C	S	C	812	IFKLQFHSMPVSNEV	IFKLQFHCMPVSNEV
		SNV	TECPR2_E1307A	E	A	1307	PVGTAWEHVPGLQAC	PVGTAWAHVPGLQAC
		Fusion	ETV6-RUNX1	-	-	-	-	AMPIGRIAECILGMN
SJETV078	Diagnostic	SNV	AHNAK_S5863F	S	F	5863	GVSLASKKSRLSSSS	GVSLASKKSRLFSSS
		SNV	BTBD16_V298M	V	M	298	IQAIPTYETVMTFFK	IQAIPTYETMMTFFK
							IPTYETVMTFFKSFP	IPTYETMMTFFKSFP
		SNV	GPR139_A298T	A	T	298	KRFRTMAAATLKAFF	KRFRTMAATTLKAFF
							FRTMAAATLKAFFKC	FRTMAATTLKAFFKC
		SNV	ITPR1_V656L	V	L	656	LIETKLVLSRFEFEG	LIETKLLLSRFEFEG
		SNV	PCNXL2_F1180L	F	L	1180	VAHLMWFERLYVWLQ	VAHLMWLERLYVWLQ
		SNV	PRRC2B_R923W	R	W	923	SSEPEWTPEPRSSSS	SSEPEWTPEPWSSSS
		SNV	SCD_E240K	E	K	240	PTLVPWYFWGETFQN	PTLVPWYFWGKTFQN
							YFWGETFQNSVFVAT	YFWGKTFQNSVFVAT
SJETV084	Diagnostic	SNV	ANKS1A_E434K	E	K	434	KYFPLTASEVLSMRP	KYFPLTASKVLSMRP
		SNV	ASXL3_E1718K	E	K	1718	STSSPMEEAISLATD	STSSPMEKAISLATD
		SNV	ATP13A4_L109F	L	F	109	PDHPLMTDEEYIINR	PDHPFMTDEEYIINR
		SNV	PDS5B_D86H	D	H	86	RLLVACCLADIFRIY	RLLVACCLAHIFRIY
		SNV	RBCK1_L130V	L	V	130	AYLYLLSARNTSLNP	AYLYVLSARNTSLNP
		Fusion	ETV6-RUNX1	-	-	-	-	AMPIGRIAECILGMN
SJETV085	Diagnostic	SNV	AZI1_M681I	M	I	681	KKLKELMSATEKARR	KKLKELISATEKARR
		SNV	CD101_S889F	S	F	889	RHLHCYRSSSTDFVL	RHLHCYRSSFTDFVL
		SNV	FAM157B_L156F	L	F	156	EFFRLLHDLHLLAFA	EFFRFLHDLHLLAFA
		SNV	GRIN3A_A401T	A	T	401	ELVARAVATATMIQP	ELVARAVTTATMIQP
		SNV	IPO13_I267M	I	M	267	AIVNAISQPDAQRYV	AIVNAMSQPDAQRYV
		SNV	SPIRE1_R332W	R	W	332	VMVNGDIPPRLKKSA	VMVNGDIPPWLKKSA
		SNV	TMEM104_V281I	V	I	281	RNLFGVCVYSFMCQH	RNLFGVCIYSFMCQH
		Fusion	ETV6-RUNX1	-	-	-	-	AMPIGRIAECILGMN

For ETV6-RUNX1, red color indicates the amino acids corresponding to ETV6 and the blue color indicates the amino acids corresponding to RUNX1

Table S7. Mutant allele frequencies (MAF) for sequenced mutations

Patient ID	Sample Type	Mutation Type	Variant	Mutation in Normal	Total in Normal	Mutation in Tumor	Total in Tumor	MAF	Blast Frequency in Tumor	MAF Relative To Blast Frequency
ERG009	DIAGNOSIS	SNV	EEF1A2_T142M	0	34	6	40	0.15	0.89	0.168539326
ERG009	DIAGNOSIS	SNV	PLCD3_D314H	1	24	14	41	0.341463	0.89	0.383666758
ERG016	DIAGNOSIS	SNV	ACE2_D615Y	0	42	35	63	0.555556	0.89	0.624219725
ERG016	DIAGNOSIS	SNV	APBA2_K495N	0	40	23	46	0.5	0.89	0.561797753
ERG016	DIAGNOSIS	SNV	OPRM1_R369C	0	32	24	60	0.4	0.89	0.449438202
ERG016	DIAGNOSIS	SNV	ST14_W826L	0	180	14	240	0.058333	0.89	0.065543071
ERG016	DIAGNOSIS	SNV	TBL1XR1_S461F	0	39	51	52	0.980769	0.89	1.1019879
ERG016	DIAGNOSIS	SNV	TMC6_Q240R	0	10	16	34	0.470588	0.89	0.528750826
ERG016	DIAGNOSIS	SNV	TMEM177_D267G	0	29	11	35	0.314286	0.89	0.353130016
ERG016	DIAGNOSIS	SNV	USP5_C451F	0	40	21	55	0.381818	0.89	0.429009193
ERG016	DIAGNOSIS	SNV	FMOD_I117T	0	29	23	43	0.534884	0.89	0.600992945
ERG016	DIAGNOSIS	SNV	MCTP1_R783H	1	40	36	73	0.493151	0.89	0.554101893
ERG016	DIAGNOSIS	SNV	NUDT10_A62V	0	10	10	18	0.555556	0.89	0.624219725
ERG016	DIAGNOSIS	SNV	PIGZ_P455S	0	113	12	130	0.092308	0.89	0.103716508
ETV001	DIAGNOSIS	SNV	MCTP1_A899V	0	47	5	30	0.166667	0.83	0.200803213
ETV001	DIAGNOSIS	SNV	SPATA5_S812C	0	44	6	40	0.15	0.83	0.180722892
ETV001	DIAGNOSIS	SNV	BLID_G9C	0	46	6	52	0.115385	0.83	0.139017609
ETV001	DIAGNOSIS	SNV	CDHR2_V1246I	0	27	27	45	0.6	0.83	0.722891566
ETV001	DIAGNOSIS	SNV	GTF3C3_H267D	0	55	8	49	0.163265	0.83	0.196705188
ETV001	DIAGNOSIS	SNV	ROCK1_A1064V	0	37	9	53	0.169811	0.83	0.204591953
ETV001	DIAGNOSIS	SNV	WDR17_E222K	0	35	8	49	0.163265	0.83	0.196705188
ETV001	DIAGNOSIS	SNV	WHSC1_E1099K	0	26	12	130	0.092308	0.83	0.111214087
ETV001	DIAGNOSIS	SNV	TECPR2_E1307A	0	22	2	9	0.222222	0.83	0.267737617
ETV001	DIAGNOSIS	Fusion	ETV6-RUNX1	NA	NA	NA	NA	NA	NA	NA
ETV078	DIAGNOSIS	SNV	AHNAK_S5863F	0	27	7	26	0.269231	0.86	0.313059034
ETV078	DIAGNOSIS	SNV	BTBD16_V298M	0	46	26	62	0.419355	0.86	0.487621905
ETV078	DIAGNOSIS	SNV	KCNK1_A308E	0	36	7	19	0.368421	0.86	0.428396573
ETV078	DIAGNOSIS	SNV	GPR139_A298T	0	37	10	26	0.384615	0.86	0.447227191
ETV078	DIAGNOSIS	SNV	KRAS_G12D	0	51	12	48	0.25	0.86	0.290697674
ETV078	DIAGNOSIS	SNV	ITPR1_V656L	0	47	14	39	0.358974	0.86	0.417412045
ETV078	DIAGNOSIS	SNV	PCNXL2_F1180L	0	35	30	51	0.588235	0.86	0.683994528
ETV078	DIAGNOSIS	SNV	PRRC2B_R923W	0	21	12	21	0.571429	0.86	0.664451827
ETV078	DIAGNOSIS	SNV	SLC24A2_A134V	0	34	17	42	0.404762	0.86	0.470653378
ETV078	DIAGNOSIS	SNV	WHSC1_E1099K	0	32	11	30	0.366667	0.86	0.426356589
ETV078	DIAGNOSIS	SNV	SCD_E240K	0	35	12	24	0.5	0.86	0.581395349
ETV078	DIAGNOSIS	Fusion	ETV6-RUNX1	NA	NA	NA	NA	NA	NA	NA
ETV084	DIAGNOSIS	SNV	ANKS1A_E434K	0	32	18	40	0.45	0.93	0.483870968
ETV084	DIAGNOSIS	SNV	ASXL3_E1718K	0	31	16	36	0.444444	0.93	0.477897252
ETV084	DIAGNOSIS	SNV	ATP13A4_L109F	0	32	33	51	0.647059	0.93	0.695762176
ETV084	DIAGNOSIS	SNV	PDS5B_D86H	0	30	16	40	0.4	0.93	0.430107527
ETV084	DIAGNOSIS	SNV	ARHGAP12_S114L	0	39	12	39	0.307692	0.93	0.330851944
ETV084	DIAGNOSIS	SNV	ARL15_E82Q	0	45	26	52	0.5	0.93	0.537634409
ETV084	DIAGNOSIS	SNV	ATP13A4_P105S	0	30	37	54	0.685185	0.93	0.736758264
ETV084	DIAGNOSIS	SNV	BRPF1_D344H	0	21	9	27	0.333333	0.93	0.358422939
ETV084	DIAGNOSIS	SNV	CC2D1A_Q695E	0	22	18	38	0.473684	0.93	0.509337861
ETV084	DIAGNOSIS	SNV	CCDC108_R1124Q	0	39	15	38	0.394737	0.93	0.424448217
ETV084	DIAGNOSIS	SNV	DTX3L_D141H	0	24	20	48	0.416667	0.93	0.448028674
ETV084	DIAGNOSIS	SNV	KIAA1715_R253Q	0	29	10	57	0.175439	0.93	0.188643652
ETV084	DIAGNOSIS	SNV	KRT7_E263Q	0	25	10	22	0.454545	0.93	0.488758553
ETV084	DIAGNOSIS	SNV	LZTS2_S451Y	0	36	11	20	0.55	0.93	0.591397849
ETV084	DIAGNOSIS	SNV	MORC4_R317K	0	32	24	59	0.40678	0.93	0.437397485
ETV084	DIAGNOSIS	SNV	NNMT_L45V	0	18	25	26	0.961538	0.93	1.033912324
ETV084	DIAGNOSIS	SNV	OR5H15_L45M	0	24	16	25	0.64	0.93	0.688172043
ETV084	DIAGNOSIS	SNV	PRL_E98Q	0	27	4	40	0.1	0.93	0.107526882
ETV084	DIAGNOSIS	SNV	SEPT14_S82P	0	25	31	65	0.476923	0.93	0.512820513
ETV084	DIAGNOSIS	SNV	SERPINA6_E138Q	0	30	11	33	0.333333	0.93	0.358422939
ETV084	DIAGNOSIS	SNV	SLC2A12_E21K	0	16	8	21	0.380952	0.93	0.409626216
ETV084	DIAGNOSIS	SNV	SLC6A7_R62C	0	27	19	41	0.463415	0.93	0.498295306
ETV084	DIAGNOSIS	SNV	SULT4A1_T205M	0	25	15	26	0.576923	0.93	0.620347395
ETV084	DIAGNOSIS	SNV	TMEM117_G377R	0	26	26	48	0.541667	0.93	0.582437276
ETV084	DIAGNOSIS	SNV	UGT2B28_S19R	0	11	26	26	1	0.93	1.075268817
ETV084	DIAGNOSIS	SNV	USH2A_S1007L	0	24	31	53	0.584906	0.93	0.628930818
ETV084	DIAGNOSIS	SNV	RBCK1_L130V	0	20	11	28	0.392857	0.93	0.422427035
ETV084	DIAGNOSIS	Fusion	ETV6-RUNX1	NA	NA	NA	NA	NA	NA	NA
ETV085	DIAGNOSIS	SNV	AZI1_M681I	0	33	24	42	0.571429	0.82	0.696864111
ETV085	DIAGNOSIS	SNV	CD101_S889F	0	36	34	54	0.62963	0.82	0.767841012
ETV085	DIAGNOSIS	SNV	FAM157B_L156F	0	28	12	29	0.413793	0.82	0.504625736
ETV085	DIAGNOSIS	SNV	GRIN3A_A401T	0	24	35	59	0.59322	0.82	0.723439438
ETV085	DIAGNOSIS	SNV	IPO13_I267M	0	30	24	60	0.4	0.82	0.487804878
ETV085	DIAGNOSIS	SNV	SPIRE1_R332W	0	43	28	56	0.5	0.82	0.609756098
ETV085	DIAGNOSIS	SNV	DCAF8L2_N405S	0	34	29	50	0.58	0.82	0.707317073
ETV085	DIAGNOSIS	SNV	FAM186A_R2174Q	0	34	31	71	0.43662	0.82	0.532463071
ETV085	DIAGNOSIS	SNV	KLF4_D462H	0	23	17	43	0.395349	0.82	0.482132728
ETV085	DIAGNOSIS	SNV	NEIL3_D413N	0	23	14	38	0.368421	0.82	0.449293967
ETV085	DIAGNOSIS	SNV	PDE3A_P472T	0	34	12	52	0.230769	0.82	0.281425891
ETV085	DIAGNOSIS	SNV	POLB_E203K	0	40	26	46	0.565217	0.82	0.689289502
ETV085	DIAGNOSIS	SNV	SHOX2_E3K	0	28	10	40	0.25	0.82	0.304878049
ETV085	DIAGNOSIS	SNV	TRIM64B_D246N	0	66	22	81	0.271605	0.82	0.331225534
ETV085	DIAGNOSIS	SNV	ZNF687_K588R	0	36	23	56	0.410714	0.82	0.50087108
ETV085	DIAGNOSIS	SNV	TMEM104_V281I	0	33	22	44	0.5	0.82	0.609756098
ETV085	DIAGNOSIS	Fusion	ETV6-RUNX1	NA	NA	NA	NA	NA	NA	NA

Table S8. Single-cell indexed FACS median fluorescence intensity (MFI)

(Table too large to embed in this Word document, so provided as a separate file)