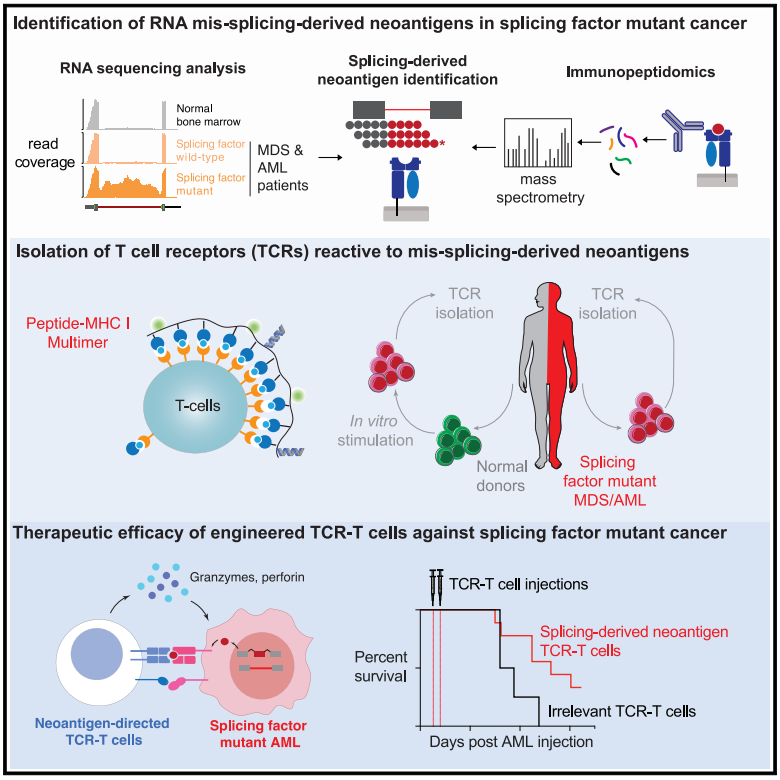
Cell



# Mis-splicing-derived neoantigens and cognate TCRs in splicing factor mutant leukemias

## Graphical abstract




## Highlights

- Oncogenic splicing factor mutations generate recurrent RNA-splicing-derived neoantigens
- T cells reactive to splicing-derived neoantigens are dysfunctional in leukemia patients
- Panels of TCRs targeting splicing-derived neoantigens can be generated
- Adoptive transfer of TCR-engineered T cells mediates specific tumor control *in vivo*

## Authors

Won Jun Kim, Edie I. Crosse, Emma De Neef, ..., Christopher A. Klebanoff, Robert K. Bradley, Omar Abdel-Wahab

## Correspondence

klebanoc@mskcc.org (C.A.K.), rbradley@fredhutch.org (R.K.B.), abdelwao@mskcc.org (O.A.-W.)

## In brief

Common oncogenic RNA splicing factor mutations in myeloid malignancies generate recurrent splicing-derived neoantigens, which can be selectively targeted using a T cell receptor-based approach.

Kim et al., 2025, Cell *188*, 1–19
June 26, 2025 
https://doi.org/10.1016/j.cell.2025.03.047

CellPress

Please cite this article in press as: Kim et al., Mis-splicing-derived neoantigens and cognate TCRs in splicing factor mutant leukemias, Cell (2025), https://doi.org/10.1016/j.cell.2025.03.047

Cell

CellPress
OPEN ACCESS

Article

# Mis-splicing-derived neoantigens and cognate TCRs in splicing factor mutant leukemias

Won Jun Kim,[1,14] Edie I. Crosse,[2,14] Emma De Neef,[2,3] Inaki Etxeberria,[4] Erich Y. Sabio,[1] Eric Wang,[5] Jan Philipp Bewersdorf,[1] Kuan-Ting Lin,[6] Sydney X. Lu,[7] Andrea Belleville,[2] Nina Fox,[1] Cynthia Castro,[1] Pu Zhang,[1] Takeshi Fujino,[1] Jennifer Lewis,[1] Jahan Rahman,[1] Beatrice Zhang,[1] Jacob H. Winick,[1] Alexander M. Lewis,[1] Robert F. Stanley,[1] Susan DeWolf,[1] Brigita Meškauskaitė Urben,[8] Meril Takizawa,[8] Tobias Krause,[8] Henrik Molina,[9] Ronan Chaligne,[8] Priya Koppikar,[10] Jeffrey Molldrem,[10] Mathieu Gigoux,[11] Taha Merghoub,[11] Anthony Daniyan,[1] Smita S. Chandran,[4] Benjamin D. Greenbaum,[12] Christopher A. Klebanoff,[4,13,*] Robert K. Bradley,[2,3,*] and Omar Abdel-Wahab[1,15,*]

[1]Molecular Pharmacology Program, Sloan Kettering Institute, Memorial Sloan Kettering Cancer Center (MSK), New York, NY, USA
[2]Public Health Sciences and Basic Sciences Divisions, Fred Hutchinson Cancer Center, Seattle, WA, USA
[3]Department of Genome Sciences, University of Washington, Seattle, WA, USA
[4]Human Oncology and Pathogenesis Program, MSK, New York, NY, USA
[5]The Jackson Laboratory for Genomic Medicine, Farmington, CT, USA
[6]Codify Therapeutics, Cambridge, MA, USA
[7]Department of Medicine, Division of Hematology, Stanford University, Palo Alto, CA, USA
[8]Single-cell Analytics Innovation Laboratory, MSK, New York, NY, USA
[9]Proteomics Resource Center, Rockefeller University, New York, NY, USA
[10]Department of Hematopoietic Biology and Malignancy, University of Texas MD Anderson Cancer Center, Houston, TX, USA
[11]Swim Across America and Ludwig Collaborative Laboratory, Department of Pharmacology, Sandra and Edward Meyer Cancer Center, Weill Cornell Medical Center, New York, NY, USA
[12]Computational Oncology, Department of Epidemiology and Biostatistics, MSK, New York, NY, USA
[13]Parker Institute for Cancer Immunotherapy, New York, NY, USA
[14]These authors contributed equally
[15]Lead contact
*Correspondence: klebanoc@mskcc.org (C.A.K.), rbradley@fredhutch.org (R.K.B.), abdelwao@mskcc.org (O.A.-W.)
https://doi.org/10.1016/j.cell.2025.03.047

## SUMMARY

Mutations in RNA splicing factors are prevalent across cancers and generate recurrently mis-spliced mRNA isoforms. Here, we identified a series of bona fide neoantigens translated from highly stereotyped splicing alterations promoted by neomorphic, leukemia-associated somatic splicing machinery mutations. We utilized feature-barcoded peptide-major histocompatibility complex (MHC) dextramers to isolate neoantigen-reactive T cell receptors (TCRs) from healthy donors, patients with active myeloid malignancy, and following curative allogeneic stem cell transplant. Neoantigen-reactive $CD8^+$ T cells were present in the blood of patients with active cancer and had a distinct phenotype from virus-reactive T cells with evidence of impaired cytotoxic function. T cells engineered with TCRs recognizing SRSF2 mutant-induced neoantigens arising from mis-splicing events in *CLK3* and *RHOT2* resulted in specific recognition and cytotoxicity of SRSF2-mutant leukemia. These data identify recurrent RNA mis-splicing events as sources of actionable public neoantigens in myeloid leukemias and provide proof of concept for genetically redirecting T cells to recognize these targets.

## INTRODUCTION

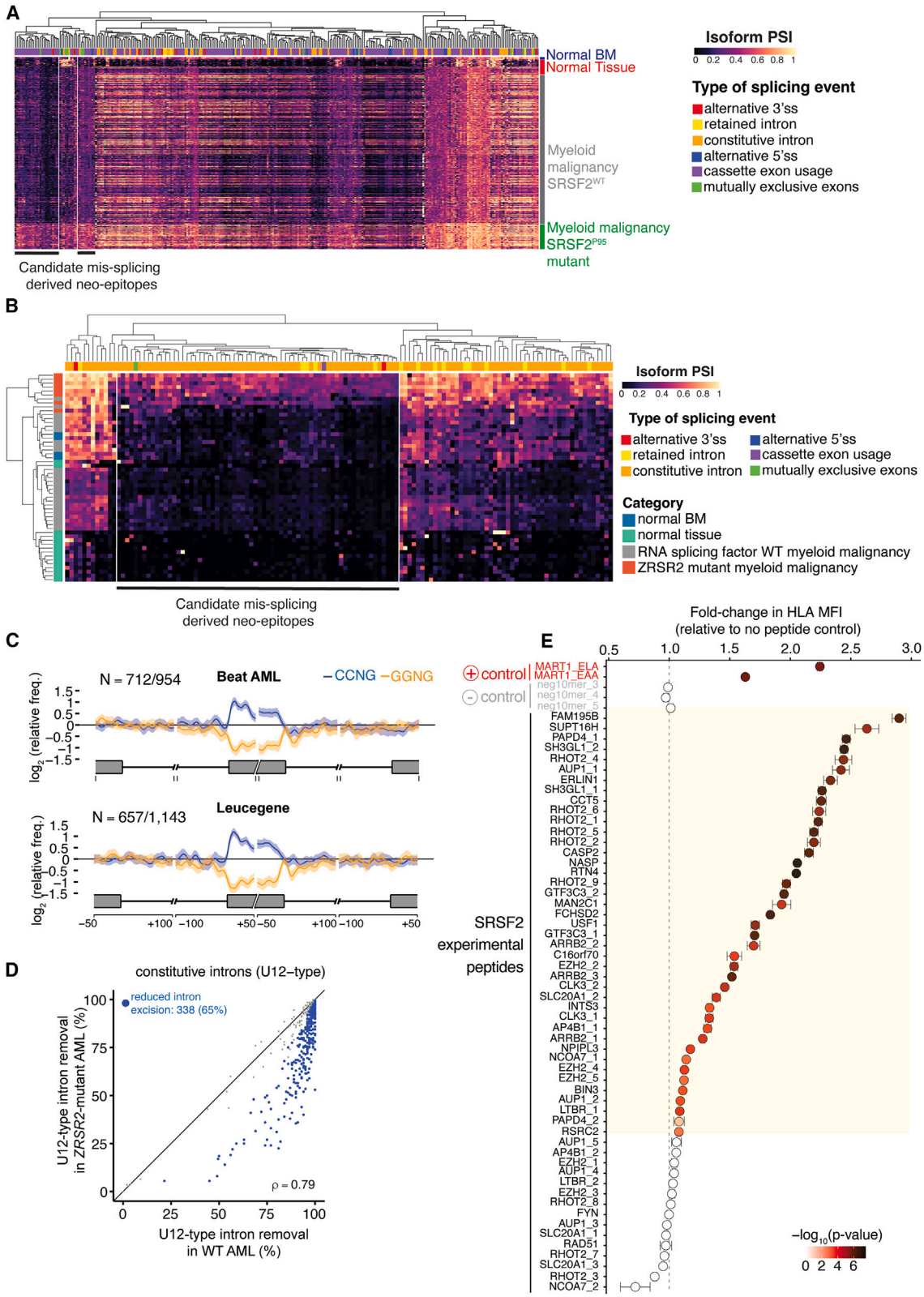
Following five decades with few approved therapies for acute myeloid leukemia (AML) and high-risk myelodysplastic syndromes (MDS), the last 5 years have brought remarkable progress with multiple new US Federal Drug Administration-approved therapies. Despite these advances, the 5-year survival for most AML patients is less than 20%,[1] and there are few effective therapies for high-risk MDS.

One major challenge in treatment of MDS and AML has been the lack of effective immunotherapies. Still, the promise for T cell-based immunotherapeutic approaches for treating myeloid leukemias is clear, given the curative potential of allogeneic hematopoietic cell transplantation (allo-HCT) for AML and MDS patients, arising from the graft-versus-leukemia (GVL) activity of donor T cells.

To date, chimeric antigen receptor (CAR) T cell approaches for AML have nearly entirely relied on antigens which are shared

Please cite this article in press as: Kim et al., Mis-splicing-derived neoantigens and cognate TCRs in splicing factor mutant leukemias, Cell (2025), https://doi.org/10.1016/j.cell.2025.03.047






*(legend on next page)*

Please cite this article in press as: Kim et al., Mis-splicing-derived neoantigens and cognate TCRs in splicing factor mutant leukemias, Cell (2025), https://doi.org/10.1016/j.cell.2025.03.047



across myeloid neoplasms and normal hematopoietic stem cells, risking unacceptable on-target but off-tumor toxicities. Given these challenges, identifying additional sources of leukemia-specific neoantigens shared across patients and applicable to multiple human leukocyte antigen (HLA) alleles could represent a major therapeutic advance. To this end, we hypothesized that cancer-associated mutations in genes encoding RNA splicing factors, which create stereotypical neomorphic changes in splicing consistently across patients,[2,3] may yield novel HLA class-I (HLA-I) displayed peptides.

Mutations in the RNA splicing factor genes *SF3B1*, *SRSF2*, *U2AF1*, and *ZRSR2* are seen in 50%–70% of MDS[4–6] and are also prevalent in AML and the related aggressive myeloid blood cancer chronic myelomonocytic leukemia (CMML). These mutations skew usage of alternative RNA splicing events observed in healthy cells but also create novel RNA isoforms, which are reproducibly observed across patients with the same mutations.[7–10] As such, aberrant mRNA splicing events have the potential to generate mis-splicing-derived neoantigens that are "public," or shared across patients with the same splicing factor mutation.

A prior study identified putative aberrant RNA-splicing-derived neoantigens in *SF3B1*-mutant uveal melanoma[11]; however, validation of such antigens as immunotherapeutic targets through T cell receptor (TCR) gene transfer was not performed. Moreover, no studies have evaluated mis-splicing-derived neoantigens in myeloid malignancies, where mutations affecting splicing factors occur at the highest frequency. Here, we sought to identify mis-splicing-derived neoantigens in myeloid leukemias with mutations in the RNA splicing factor genes *SRSF2* and *ZRSR2*. We focused on these two factors given their high mutation frequencies, associations with adverse outcomes, and unique mechanistic roles in RNA splicing. SRSF2 mutations in particular are especially common across myeloid neoplasms, as they occur in 30%–50% patients with CMML,[12] 17%–25% of patients with MDS,[13] 18%–20% of patients with AML over the age of 60,[14,15] and 6%–8% of younger patients with AML (≤60 years).[15]

Here, we identified a series of neoantigens derived from shared mis-splicing events observed across myeloid leukemia patients with SRSF2 or ZRSR2 mutations. These empirically validated neoantigens were then used to construct HLA-I dextramers, which allowed us to discover rare, circulating neoantigen-reactive $CD8^+$ T cells in patients, evaluate their transcriptional characteristics, and isolate their TCRs. Finally, a panel of TCRs reactive against *SRSF2* mutation-induced mis-splicing-derived neoantigens redirected primary $CD8^+$ T cells for specific recognition and cytotoxicity of *SRSF2*-mutant leukemia cells. This study thereby catalogs endogenously produced neoantigens and their cognate TCRs in patients with splicing factor mutant leukemia and proposes neoantigen-based TCR-T cell therapies as novel immunotherapeutic strategies in this population.

## RESULTS

### Identification of putative RNA mis-splicing-derived neoantigens in *SRSF2*- or *ZRSR2*-mutant myeloid leukemias

Mutations in the splicing factors *SF3B1*, *SRSF2*, *U2AF1*, and *ZRSR2* cause highly sequence-specific changes to splicing consistent across cancers and cell types which are uniformly observed across patients with mutations in the same gene.[7–10] We therefore hypothesized that a subset of RNA mis-splicing events created by mutations in RNA splicing factors could give rise to novel HLA-I presented peptides that are immunogenic.

To test this hypothesis, we performed large-scale RNA sequencing (RNA-seq) analyses to identify recurrently mis-spliced isoforms in myeloid neoplasm patient samples from five patient cohorts with *SRSF2* mutations ($n = 107$), *ZRSR2* mutations ($n = 33$), or no mutations in *SF3B1*, *SRSF2*, *U2AF1*, or *ZRSR2* (referred to as wild-type or WT; $n = 837$; Figure S1A; Table S1).[7,10,16] We identified mis-spliced isoforms consistently produced across multiple patient cohorts and minimally expressed in healthy bone marrow (BM), peripheral blood mononuclear cells (PBMCs), and 14 normal tissues. Mis-spliced isoforms produced in *SRSF2*-mutant MDS and AML affected diverse alternative splicing events (differentially spliced cassette exons, alternative 5′ or 3′ splice sites, mutually exclusive exons, and retention of constitutive or alternative introns), which could be broadly categorized into two groups (Figure 1A). The first and larger group encompasses mis-splicing events that enhance production of isoforms that are normally produced in healthy tissues at a lower level. The second, smaller group is

**Figure 1. Prediction of RNA mis-splicing-derived neoantigens and confirmation of HLA-I binding in splicing factor mutant myeloid leukemias**

(A) Heatmap illustrating isoform percent spliced-in (PSI) values for splicing events differentially spliced between *SRSF2* P95-mutant ($n = 54$) and splicing factor wild-type myeloid malignancies ($n = 325$) across the Beat AML cohort,[17] Body Map 2.0 panel of healthy tissues ($n = 14$), and healthy bone marrow ($n = 4$[18]). Each row represents a patient transcriptome, and columns represent isoforms. Groups of isoforms exhibiting specificity for *SRSF2*-mutant cells are below heatmap.

(B) As in (A), but for *ZRSR2*-mutant malignancies ($n = 33$).

(C) Top: relative frequency of CCNG and GGNG (N = any nucleotide) motifs in cassette exons promoted versus repressed by *SRSF2* mutations from (A). Bottom: identical analysis but for the Leucegene AML cohort.[19] Shading, 95% confidence interval by bootstrapping. Schematic illustrates a metagene containing the differentially spliced cassette exon. Horizontal axis, genomic coordinates defined with respect to 5′ and 3′ splice sites where 0 is the splice site. Vertical axis, relative frequency of the motifs over genomic loci containing cassette exons promoted versus repressed by *SRSF2* mutations (log scale).

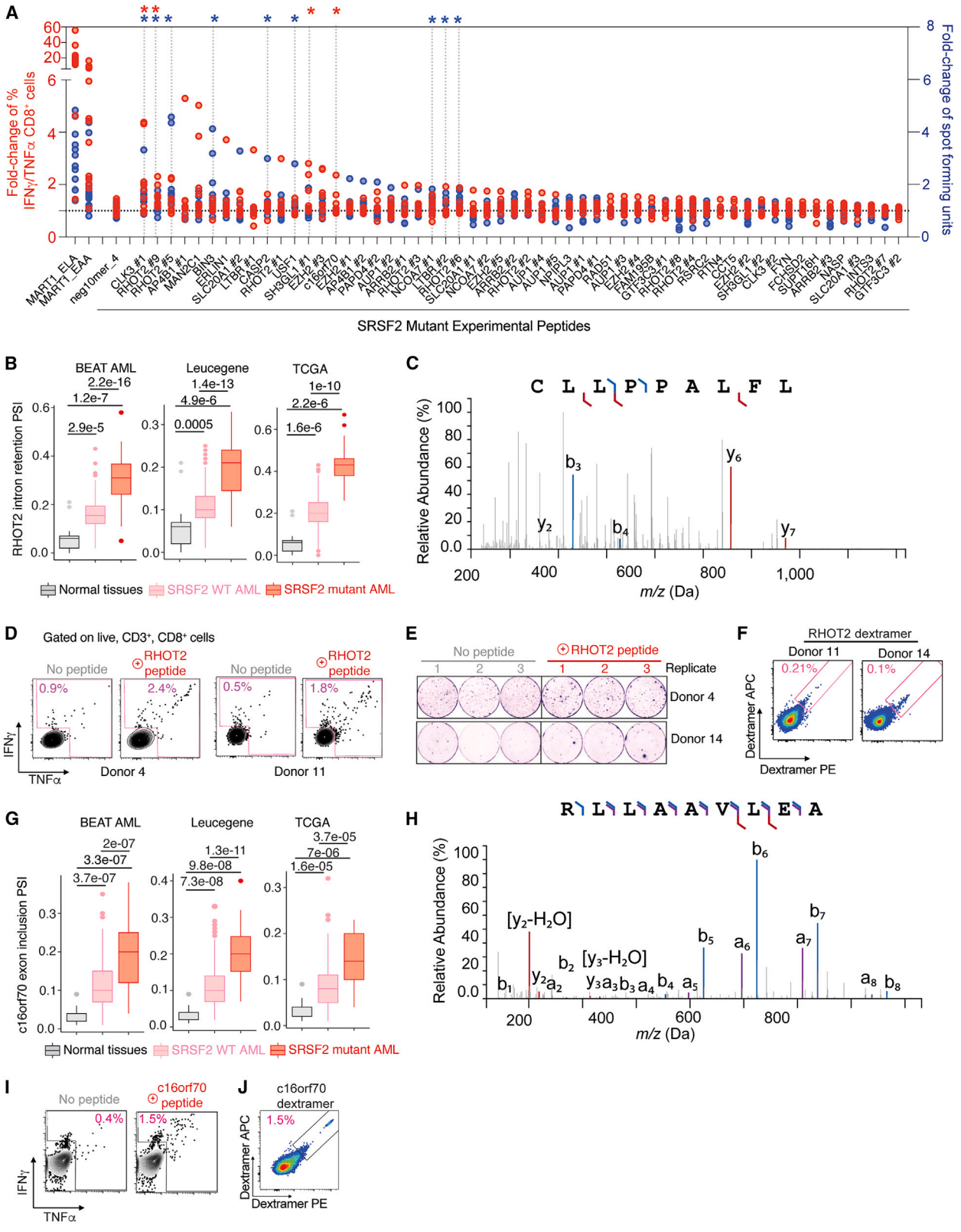
(D) Quantification of differential splicing of minor (U12-type) introns from (B). Each point corresponds to a single intron and illustrates percentage of mRNA in which the intron is spliced out. Blue dots, introns with significantly increased retention in *ZRSR2*-mutant versus WT cells, defined as an absolute change in retention of $\geq 10\%$ or absolute log fold-change of $\geq 2$ with associated $p \leq 0.05$.

(E) Median fluorescence intensity (MFI) values of HLA-ABC surface expression from T2 HLA-A2 stabilization experiments. Yellow shaded area, peptides with significant HLA-I binding. Mean ± SEM. $p$ values, unpaired Student's t test. Red text, HLA-A*02:01 MART-1 peptide positive controls ("ELA": ELAGIGILTV; "EAA": EAAGIGILTV). Negative control 10-mer peptides are predicted non-binders to HLA-A*02:01.

See also Figure S1.

Please cite this article in press as: Kim et al., Mis-splicing-derived neoantigens and cognate TCRs in splicing factor mutant leukemias, Cell (2025), https://doi.org/10.1016/j.cell.2025.03.047






*(legend on next page)*

Please cite this article in press as: Kim et al., Mis-splicing-derived neoantigens and cognate TCRs in splicing factor mutant leukemias, Cell (2025), https://doi.org/10.1016/j.cell.2025.03.047





characterized by mis-splicing events that result in production of aberrant isoforms absent from healthy tissues and therefore represent candidates for neoepitope production. *ZRSR2* mutations similarly gave rise to diverse mis-spliced isoforms, including strong representation of intron retention events and a large number of mis-spliced isoforms specific to *ZRSR2*-mutant cells, and not observed in healthy tissues (Figure 1B). Cassette exons included in *SRSF2*-mutant cells exhibited a strong enrichment for C-rich exonic splicing enhancer sequences (Figure 1C), while G-rich exons were predominantly skipped (as described previously[7]). *ZRSR2*-mutant transcriptomes exhibited a striking retention of U12-type introns (Figure 1D), consistent with ZRSR2's known role as a component of the minor (or U12-type) spliceosome.[10]

We next ranked mis-spliced isoforms based on the following criteria: (1) significant and consistent mis-splicing across patient cohorts, (2) minimal mis-splicing in healthy tissues, (3) high mis-spliced isoform expression (taking both mis-splicing rate and gene-expression levels into account) in tumor and low expression in healthy tissues, (4) encode epitopes with predicted high-affinity binding to HLA-I which uniquely arise from the mis-spliced isoform. We translated each tumor-specific aberrant RNA isoform *in silico*, split each into 8- to 12-mer peptides, limited to those that arose uniquely from their parent mis-spliced isoforms, and predicted high-affinity binders to HLA-A*02:01 with NetMHCpan 4.0.[20] We further prioritized resulting candidates based on their parent isoforms' mis-splicing patterns (Figure S1A; Tables S2 and S3; STAR Methods). In total, we selected 56 candidate mis-splicing-derived neoantigens created by SRSF2 mutations and 19 by ZRSR2 mutations for further studies (Tables S2 and S3).

## Validation of immunogenicity

We next tested whether candidate neoepitopes stably bind HLA-A*02:01 *in vitro* by performing a T2 HLA-A2 shift assay[21] (Figure 1E). We evaluated 56 candidate SRSF2-mutant and 12 ZRSR2-mutant mis-splicing-derived peptides over a peptide concentration range of 0–100 μg/mL (Figures 1E and S1B). As positive controls, we used the well-studied HLA-A*02:01-restricted Melan-A/MART1 (26–35) peptide (EAAGIGILTV) and its high-affinity heteroclitic variant (ELAGIGILTV).[22] As negative controls, we used three distinct 10-mer peptides predicted not to bind HLA-A*02:01. Relative to controls, 42 out of 56 SRSF2 mutant-induced peptides (75%) and 9 out of 12 ZRSR2 mutant-induced peptides (75%) exhibited significant HLA-I binding (Figures 1E and S1B).

To evaluate the immunogenicity of the candidate peptides as well as to isolate antigen-reactive CD8$^+$ T cells with high-avidity cognate TCRs, we performed *in vitro* sensitization of CD8$^+$ T cells from HLA-A*02:01$^+$ healthy donors using our predicted neoantigens.[23,24] We stimulated bulk PBMCs from 14 unique HLA-A*02:01 healthy donors with candidate neopeptides to induce CD8$^+$ T cell priming (Figure S2A). After 1 week, we re-stimulated the expanded T cells with the same peptide and measured interferon γ (IFN-γ)/tumor necrosis factor alpha (TNF-α) production using intracellular cytokine staining (ICS) and IFN-γ ELISpot (enzyme linked immunosorbent spot). A number of SRSF2-mutant-induced peptides were immunogenic across multiple donors with statistical significance, which include those derived from mis-spliced transcripts encoding CLK3, RHOT2, and c16orf70 (Figure 2A).

We further interrogated these immunogenic peptides and their related mis-splicing events. One example is RHOT2 peptide #5 (CLLPPALFL), derived from an inclusion of intron 5 in *RHOT2* (Figure S2B). This mis-splicing event is recurrently upregulated across SRSF2-mutant AML patients, compared with SRSF2 WT AML patients and normal human tissues across three cohorts of AML RNA-seq datasets (Figure 2B). We validated endogenous presentation of this peptide by HLA-immunoprecipitation liquid chromatography-tandem mass spectrometry (HLA-IP LC-MS/MS)[26] on HLA-A*02:01-transduced K562 cells containing a knockin SRSF2$^{P95H/WT}$ mutation (Figures 2C and S2C). As noted above, this RHOT2 neoantigen was immunogenic across multiple donors (Figures 2A, 2D, and 2E). To unequivocally prove the *in vitro* priming of antigen-reactive CD8$^+$ T cells using our immunogenicity assay, we generated dual fluorescence-labeled dextramers for RHOT2 peptide #5, which are composed of up to 20 peptide/HLA-I (p/HLA-I) complexes assembled on a dextran scaffold. We stained the *in vitro* primed T cells with dextramers and identified dextramer double-positive CD8$^+$ T cells reactive to RHOT2 peptide #5 (Figure 2F).

Another example of immunogenic peptides is c16orf70 peptide (RLLAAVLEA). SRSF2 mutations promote inclusion of a

**Figure 2. Immunogenicity of candidate SRSF2 mutant-induced neoantigens and identification of neoantigen-reactive T cell**

(A) Left y axis: fold-change in percent of IFN-γ$^+$ and/or TNF-α$^+$ CD8$^+$ T cells upon candidate peptide re-stimulation (normalized to DMSO), tested across $n$ = 14 healthy donor PBMCs. Right y axis: fold-change in IFN-γ ELISpot spot forming units upon re-stimulation with the test peptide (versus DMSO). Each dot represents mean value across $n$ = 3 technical replicates for a single healthy donor. Red text, HLA-A*02:01 MART-1 positive control peptides. Negative control peptide is a predicted nonbinder to HLA-A*02:01. $p$ values calculated for each candidate peptide (versus neg10mer_4 peptide) using Wilcoxon signed-rank test (*$p$.adj $\leq$ 0.05). Red and blue asterisks: peptides with statistical significance from ICS and IFN-γ ELISpot, respectively.

(B) Boxplot quantifications of PSI values of the RHOT2 intron retention event in the BeatAML,[17] Leucegene,[19] and AML TCGA[25] cohorts. Normal tissue controls from Body Map 2.0. Horizontal line, median; colored box, interquartile range; whiskers, minimum and maximum values within 1.5× of the interquartile range; dots outside whiskers, outliers. $p$ values, one-sided Mann-Whitney U test.

(C) MS2 spectra of RHOT2 peptide #5 eluted from HLA-A*02:01$^+$ SRSF2$^{P95H/WT}$ K562 cells. Peaks represent b ions (blue) and y ions (red).

(D) ICS FACS plots from (A) demonstrating IFN-γ$^+$ and/or TNF-α$^+$ CD8$^+$ T cells upon restimulation with RHOT2 peptide #5 or DMSO.

(E) IFN-γ ELISpot images from PBMCs upon restimulation with RHOT2 peptide #5 or DMSO.

(F) FACS plots of RHOT2 peptide #5-primed CD8$^+$ T cells from two donors, stained with dual color RHOT2 peptide #5-HLA-A*02:01 dextramers.

(G) As in (B) but for the exon inclusion event in *c16orf70*.

(H) As in (C) but for c16orf70 peptide.

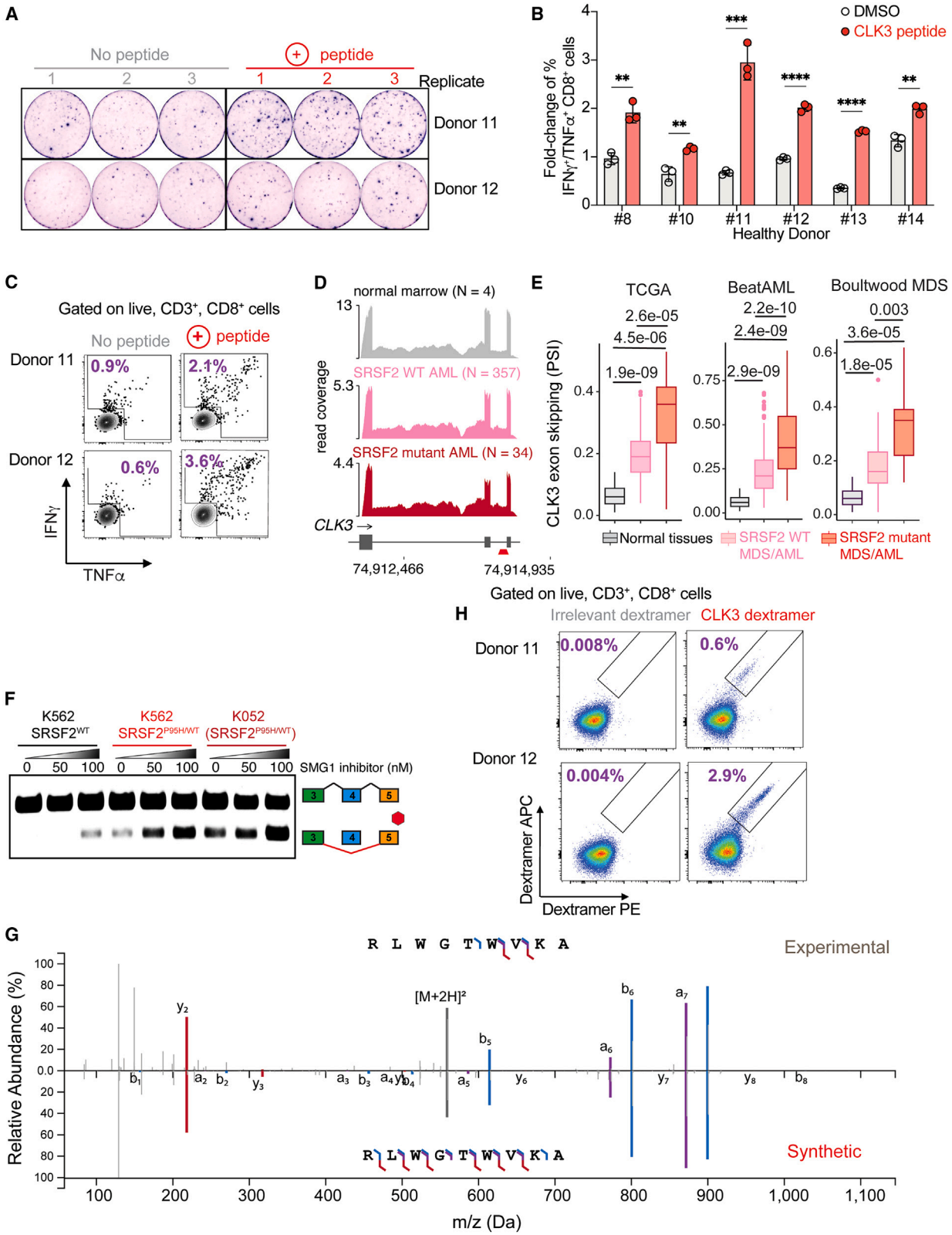
(I) As in (D) but for c16orf70 peptide.

(J) As in (F) but for c16orf70 peptide.

See also Figure S2.

Please cite this article in press as: Kim et al., Mis-splicing-derived neoantigens and cognate TCRs in splicing factor mutant leukemias, Cell (2025), https://doi.org/10.1016/j.cell.2025.03.047






*(legend on next page)*

Please cite this article in press as: Kim et al., Mis-splicing-derived neoantigens and cognate TCRs in splicing factor mutant leukemias, Cell (2025), https://doi.org/10.1016/j.cell.2025.03.047



poison exon in *c16orf70*, and the resulting truncated protein product gives rise to the c16orf70 neoantigen (Figure S2D). We confirmed the endogenous production and presentation of the peptide in SRSF2$^{P95H/WT}$-mutant cells and validated the immunogenicity of this peptide across multiple donors and presence of neoantigen-reactive CD8$^+$ T cells in *in vitro* primed T cells (Figures 2G–2J).

For 12 ZRSR2 mutant-induced peptides, two peptides were immunogenic across multiple donors with statistical significance (Figures S2E–S2I). These neoantigens are translated from minor introns in *ATG3* and *MYO1F* (Figures S2F and S2H). Dual-color p/HLA-I dextramer staining validated the successful *in vitro* priming of CD8$^+$ T cells reactive to ATG3 and MYO1F peptides (Figures S2J and S2K).

To understand how broadly applicable our proposed neoantigen-based therapies are for patients with MDS, CMML, and AML, we analyzed RNA-seq data from cohorts of patients with each of these diseases to quantify expression of SRSF2 or ZRSR2 mutant-induced mis-spliced mRNA encoding each neoantigen (Figures S3A and S3B). These data demonstrate that expression of most of the neoantigens are consistent across patients with MDS, CMML, and AML bearing mutations in SRSF2 or ZRSR2.

## SRSF2 mutations induce *CLK3* exon 4 skipping and give rise to a novel antigen

One peptide (RLWGTWVKA) derived from SRSF2-mutation-induced aberrant splicing of *CLK3* transcript was the most frequently antigenic peptide across donors (Figures 2A, 3A–3C, and S3C). Given the robust capacity of this peptide to elicit T cell responses, we investigated the related mis-splicing event in more detail.

SRSF2 mutations promote skipping of exon 4 in *CLK3* and the resulting transcript is a predicted substrate for nonsense-mediated mRNA decay (NMD) (Figures 3D and S3D). This aberrant CLK3 isoform is significantly upregulated in SRSF2 mutant leukemia patients in RNA-seq data across three AML cohorts and patients with MDS (Figures 3D, 3E, and S3E). The splicing alteration was validated by RT-qPCR in isogenic K562 cells with heterozygous knockin of the SRSF2$^{P95H/WT}$ mutation as well as in KO52 cells with a naturally occurring SRSF2$^{P95H/WT}$ mutation (Figure 3F). Consistent with prediction that the *CLK3* mis-spliced transcript is an NMD substrate, NMD inhibition (using an inhibitor of SMG1[28,29]), led to a dose-dependent increase in the mis-spliced isoform in SRSF2-mutant cells (Figure 3F). SMG1 inhibition also subtly increased *CLK3* mis-spliced transcript in SRSF2 WT K562 cells but to a much lesser degree compared with SRSF2 mutant cells.

Prior studies suggested that NMD transcripts have a higher rate of RNA turnover accompanied by rapid decay of nascent peptides by the proteosome that may drive frequent production of HLA-I peptides.[30] To confirm that the *CLK3* mis-spliced isoform gives rise to the predicted CLK3 neoantigen endogenously, we performed HLA-IP LC-MS/MS on HLA-A*02:01$^+$ SRSF2$^{P95H/WT}$-mutant K562 cells as well as KO52 cells. Eluted peptides exhibited the expected sequence preferences at anchor residues (Figures S3F and S3G). In both cell lines, we identified the eluted peptide matching MS/MS spectra with that of synthetic CLK3 peptide (Figures 3G and S4A).

## Characterization of TCRs recognizing CLK3 neoantigen

To isolate CLK3 neoantigen-reactive TCRs, we synthesized dual-color CLK3 p/HLA-I dextramers and identified CLK3 neoantigen-reactive CD8$^+$ T cells from two healthy donors (Figure 3H). We sorted dextramer$^+$CD8$^+$ populations and performed single-cell RNA- and TCR-seq (Figure 4A). Expression of cytotoxic markers (*GZMB*, *CCL5*, *NKG7*, and *PRF1*) indicated productive T cell activation and priming (Figures S4B–S4D). TCR sequencing revealed 11 distinct TCR clonotypes across two donors (Figures 4B and S4E).

We next sought to functionally characterize the CLK3 neoantigen-reactive TCRs. Polyclonal human primary CD8$^+$ T cells were transduced with individual TCR panel members. T cells expressing each of the candidate TCRs bound the CLK3 p/HLA-I dextramers, establishing their specificity for the CLK3 neoantigen (Figure 4C). As a negative control, we confirmed the absence of dextramer staining on CD8$^+$ T cells transduced with an irrelevant CMV (Cytomegalovirus) TCR. We next evaluated the reactivity of each TCR to CLK3 neoantigen by co-culturing TCR-T cells with AML cells (1) loaded with increasing concentrations of CLK3 peptide or (2) electroporated with *CLK3* full length or exon 4 skipped mRNA (Figure 4D). We utilized HLA-A*02:01-expressing K562 cells co-transduced with CD80 and CD83 to mimic an artificial antigen-presenting cell (aAPC) system[31,32] (Figure S4F). All CLK3 neoantigen-reactive TCR clonotypes elicited robust, dose-dependent production of IFN-$\gamma$ and TNF-$\alpha$

**Figure 3. Discovery and validation of SRSF2 mutant-induced CLK3 neoantigen**

(A) IFN-$\gamma$ ELISpot images of two healthy donor PBMCs *in vitro* primed and restimulated with CLK3 peptide or DMSO.

(B) Fold-change in percent of IFN-$\gamma^+$ and/or TNF-$\alpha^+$ CD8$^+$ T cells restimulated with CLK3 peptide or DMSO, across six healthy donors. $n = 3$ technical replicates. Mean ± SEM. $p$ values, unpaired Student's t test (**$p$.adj $\leq$ 0.01, ***$p$.adj $\leq$ 0.001, ****$p$.adj $\leq$ 0.0001).

(C) ICS FACS plots in two donors from (B).

(D) CLK3 exon 4 skipping RNA-seq coverage plots in SRSF2$^{P95}$-mutant ($n = 34$) and splicing factor wild-type ($n = 357$) patients from the Leucegene cohort[19] and normal bone marrow samples ($n = 4$).[18]

(E) Quantification of PSI of CLK3 exon 4 skipping event in AML TCGA,[25] BeatAML,[17] and Boultwood MDS[27] datasets. Normal tissue controls from Body Map 2.0. Horizontal line, median; colored box, interquartile range; whiskers, minimum and maximum values within 1.5× of the interquartile range; dots outside whiskers, outliers. $p$ values, one-sided Mann-Whitney U test.

(F) RT-qPCR of CLK3 exon 4 skipping event in isogenic K562 cells and SRSF2$^{P95H}$-mutant KO52 cells treated with SMG1 inhibitor.

(G) MS2 spectra of CLK3 peptide eluted from HLA-A*02:01$^+$ SRSF2$^{P95H/WT}$ knockin K562 cells (top) and synthetic peptide (bottom). Peaks represent $b$ ions (blue) and $y$ ions (red).

(H) FACS plots of CLK3 peptide-primed CD8$^+$ T cells from two donors stained with dual-color CLK3 peptide-HLA-A*02:01 dextramers.

See also Figure S3.

Please cite this article in press as: Kim et al., Mis-splicing-derived neoantigens and cognate TCRs in splicing factor mutant leukemias, Cell (2025), https://doi.org/10.1016/j.cell.2025.03.047



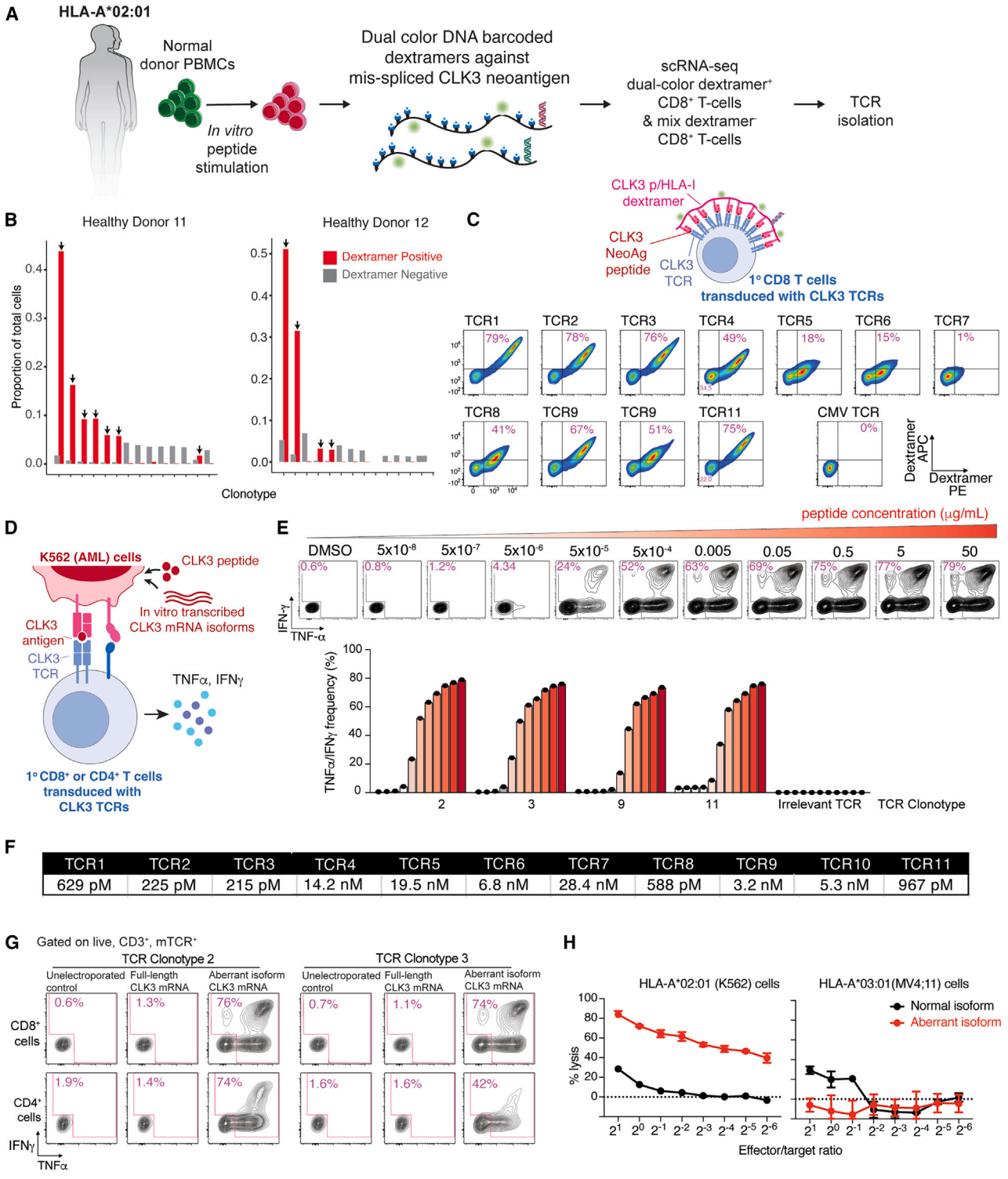



**Figure 4. Discovery and characterization of TCRs recognizing the CLK3 neoantigen**

(A) Schema of assay to isolate CLK3 neoantigen-reactive TCRs from HLA-A*02:01$^+$ healthy donors.

(B) Proportion of total cells with the indicated TCR clonotype in dextramer$^+$ (red bars) versus dextramer$^-$ (gray bars) cells. Arrows, TCR clonotypes enriched in dextramer$^+$ cells.

(C) FACS plots of CLK3 peptide-HLA-A*02:01 dextramer staining of CD8$^+$ T cells transduced with the isolated TCRs. Percentages denote dextramer-double-positive cells.

*(legend continued on next page)*

Please cite this article in press as: Kim et al., Mis-splicing-derived neoantigens and cognate TCRs in splicing factor mutant leukemias, Cell (2025), https://doi.org/10.1016/j.cell.2025.03.047





by TCR-T cells when co-cultured with peptide-loaded K562 cells (Figure 4E). By contrast, $CD8^+$ T cells transduced with control CMV TCR did not react to the CLK3 peptide. We calculated the half-maximal effective concentration ($EC_{50}$) peptide concentration of each TCR clonotype, and found two TCR clonotypes (TCR2 and TCR3) to be especially potent with $EC_{50}$ values in the picomolar range (Figures 4F and S4G). CD8 co-receptor independence of a TCR, which can be measured by TCR activity in $CD4^+$ cells, often correlates with TCR affinity and anti-tumor efficacy.[33] We measured robust IFN-$\gamma$ and TNF-$\alpha$ production in both $CD8^+$ and $CD4^+$ T cells transduced with candidate CLK3 TCRs following co-culture with K562 cells electroporated with *CLK3* mis-spliced mRNA (Figure 4G). These data indicate that the CLK3 neoantigen-reactive TCRs we identified are potently reactive to their cognate neoantigen.

Finally, we evaluated whether CLK3 neoantigen-reactive TCRs can redirect primary $CD8^+$ T cells for specific cytotoxicity of leukemic cells overexpressing *CLK3* mis-spliced transcript. $CD8^+$ CLK3 TCR-T cells specifically lysed HLA-A*02:01-expressing K562 cells electroporated with *CLK3* mis-spliced mRNA and did not lyse cells electroporated with *CLK3* WT mRNA (Figure 4H). In contrast, the same TCR-T cells did not lyse HLA-A*02:01-negative MV4;11 AML cells electroporated with either *CLK3* mRNA isoform, confirming HLA restriction of these TCRs.

## CLK3 neoantigen-reactive TCR-T cells specifically lyse SRSF2 mutant AML cells

We next asked whether endogenous SRSF2 mutations can lead to production and HLA presentation of the CLK3 neoantigen at a sufficient level for recognition by CLK3 neoantigen-reactive TCR-T cells. We engineered KO52 cells, established from an HLA-A*02:01$^-$ AML patient with an endogenous $SRSF2^{P95H/WT}$ mutation,[7] to express HLA-A*02:01 (Figure S5A). These cells were co-cultured with $CD8^+$ T cells transduced with each of the most potent CLK3 neoantigen-reactive TCRs. All TCRs conferred cytolytic activity to $CD8^+$ T cells (Figures 5A and S5B). No cytolytic activity was seen against KO52 cells without HLA-A*02:01 (Figure S5B). CLK3 TCR-T cells upregulated canonical T cell activation markers CD69 and 4-1BB only when co-cultured with HLA-A*02:01$^+$ KO52 target cells, with a concomitant increase in frequencies of activated caspase-3/7$^+$ and/or 7-AAD$^+$ populations in target cells (Figure 5B).

As an orthogonal experiment, we introduced firefly luciferase into KO52 cells and co-cultured with CLK3 neoantigen-reactive TCR-T cells across varying effector-to-target (E:T) ratios. We again observed specific lysis of HLA-A*02:01$^+$ KO52 cells by CLK3 TCR-T cells (Figure 5C). The same TCR-T cells failed to recognize SRSF2-WT AML cells (THP-1 and MV4;11) that endogenously or exogenously express HLA-A*02:01, respectively (Figures 5C and S5A). To confirm that these tumor lines expressed sufficient HLA-A*02:01 for immune recognition, we pulsed CLK3 peptide onto THP-1 and MV4;11 cells, and this rendered them susceptible to potent killing by the TCR-T cells (Figure S5C).

In addition to KO52 cells, we tested whether CLK3 neoantigen-reactive TCR-T cells are reactive to primary cells from patients with SRSF2 mutant AML. We co-cultured CLK3 TCR-T cells with BM mononuclear cells (BM MNCs) or PBMCs from HLA-A*02:01$^+$ SRSF2 mutant AML patients (Table S4) and observed a significantly increased frequency of CD69$^+$ and 4-1BB$^+$ population in CLK3 TCR-T cells (compared with CMV TCR-T cells) (Figures 5D, 5E, and S5D). As negative controls, we performed the same experiment on primary cells from HLA-A*02:01$^+$ SRSF2 WT AML patients or normal donors, and we did not observe significant activation of CLK3 TCR-T cells. We therefore conclude that SRSF2 mutant AML cells endogenously process and present a sufficient amount of the CLK3 neoantigen to allow their recognition and lysis by CLK3 neoantigen-reactive TCR-T cells.

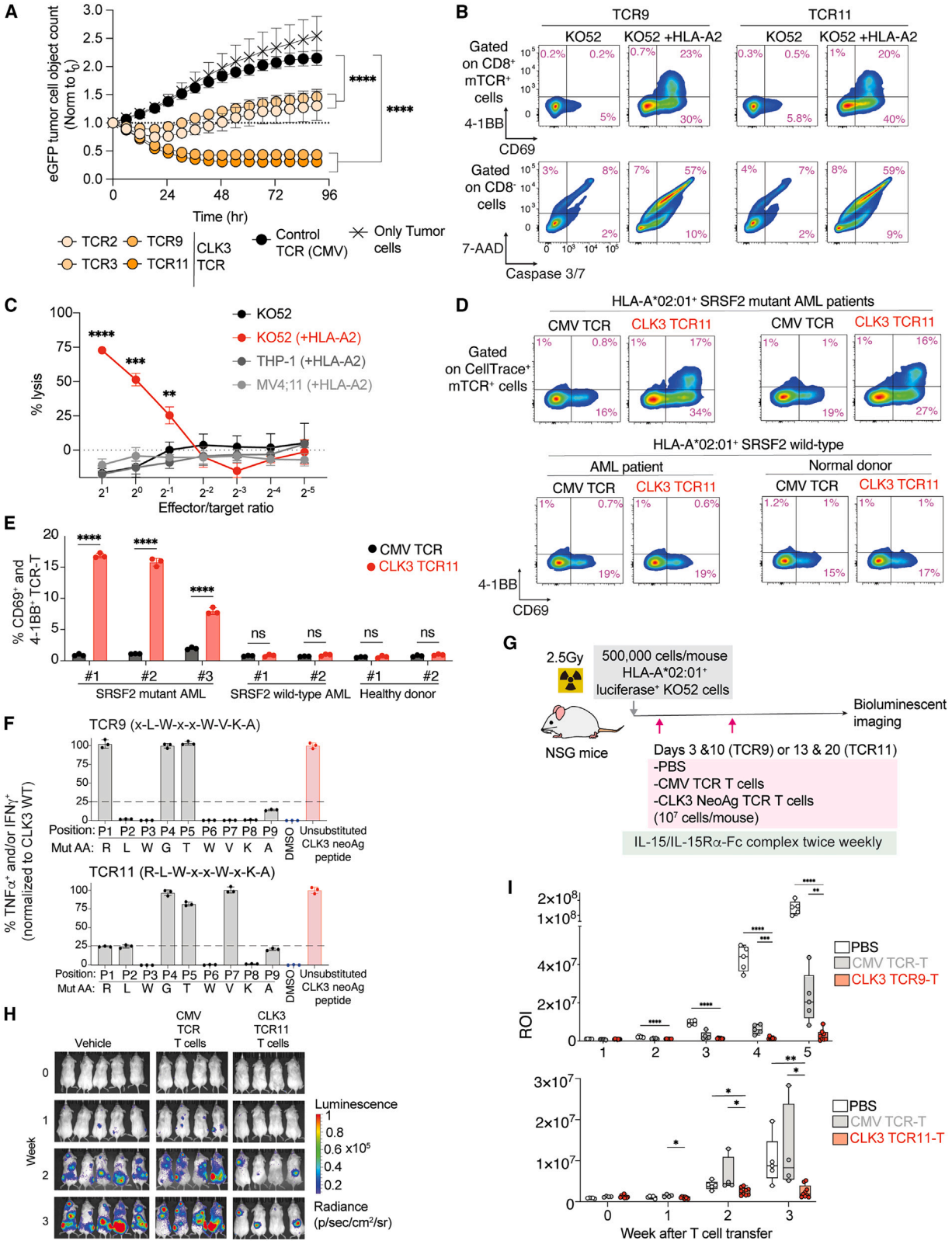
We next assessed the specificity of the CLK3 neoantigen-reactive TCRs and potential off-target peptides using alanine scanning. We generated a set of 9-mer peptides where each of the amino acid residues of the CLK3 neoantigen (RLWGTWVKA) was substituted with alanine (or glycine for the amino acid #9 as it is already an alanine). We pulsed HLA-A*02:01$^+$ K562 cells with these peptides individually, co-cultured with $CD8^+$ T cells transduced with each of the four CLK3 TCRs and assessed TNF-$\alpha$ and IFN-$\gamma$ production in TCR-T cells. We defined peptide positions resulting in TNF-$\alpha$ and/or IFN-$\gamma$ release >25% relative to the native CLK3 peptide as being permissive. These data established a peptide recognition motif for each TCR (Figures 5F and S5E). We then utilized ScanProsite[34] to survey the human proteome for these motifs. Importantly, for TCR3, TCR9, and TCR11, the CLK3 neoantigen is the only sequence recognized. For TCR2, potential off-target peptides derived from the proteins BOC (cLWraWsKq), DOC-1 (kLWipWmKs), GVINP1 (qLWhhWcKk), CD360 (pLWrlWkKi), and TEX50 (yLWkkWkKh) were identified. We synthesized these peptides and confirmed that none of them resulted in activation of $CD8^+$ T cells transduced with TCR2 (Figure S5F).

Finally, we tested whether the anti-tumor activity of CLK3 neoantigen-reactive TCR-T cells translates into meaningful therapeutic efficacy *in vivo*. We generated a xenograft model of SRSF2 mutant AML by intravenously injecting firefly luciferase-labeled HLA-A*02:01$^+$ KO52 cells into sublethally irradiated

(D) Schema of experiment testing antigen reactivity of isolated TCRs. T cells transduced with each TCR were co-cultured with K562 cells loaded with CLK3 peptide or electroporated with *CLK3* mis-spliced mRNA.
(E) Top: IFN-$\gamma$ and TNF-$\alpha$ ICS FACS plots of $CD8^+$ T cells transduced with CLK3 TCR clonotype 3 and exposed to K562 cells loaded with increasing concentration of CLK3 peptide. Bottom: frequency of IFN-$\gamma^+$ and/or TNF-$\alpha^+$ $CD8^+$ T cells expressing four distinct CLK3 TCRs or CMV TCR.
(F) Half-maximal effective concentration ($EC_{50}$) of the CLK3 peptide for each TCR based on TNF-$\alpha$ and IFN-$\gamma$ ICS from (E).
(G) IFN-$\gamma$ and TNF-$\alpha$ ICS FACS plots of $CD8^+$ or $CD4^+$ T cells transduced with CLK3 TCR clonotype 2 or 3 and exposed to K562 cells electroporated with *CLK3* wild-type or mis-spliced mRNA.
(H) Percent lysis of HLA-A*02:01$^+$ K562 cells (left) or HLA-A*03:01$^+$ MV4;11 cells (right) expressing CLK3 mRNA and exposed to $CD8^+$ T cells transduced with CLK3 TCR clonotype 3. Mean ± SEM.
See also Figure S4.

Please cite this article in press as: Kim et al., Mis-splicing-derived neoantigens and cognate TCRs in splicing factor mutant leukemias, Cell (2025), https://doi.org/10.1016/j.cell.2025.03.047






(legend on next page)

Please cite this article in press as: Kim et al., Mis-splicing-derived neoantigens and cognate TCRs in splicing factor mutant leukemias, Cell (2025), https://doi.org/10.1016/j.cell.2025.03.047



NSG (NOD SCID gamma) mice (Figure 5G). In the first experiment, we tested the therapeutic efficacy of CLK3 TCR9 by randomizing tumor-bearing mice to one of the following treatments: (1) PBS ($n$ = 5), (2) CMV-reactive TCR-T cells as a specificity control ($n$ = 5), and (3) CLK3 TCR9-T cells ($n$ = 8) (Figure 5G). All mice received an assigned treatment on days 3 and 10 post-tumor engraftment in addition to an extended half-life variant of interleukin (IL)-15.[35] Mice treated with CLK3 TCR9-T cells had a significantly lower tumor burden compared with those that received PBS or CMV TCR-T controls (Figures 5I and S5G). None of the mice experienced significant weight loss (Figure S5H).

In a second experiment, we tested the therapeutic efficacy of another CLK3 neoantigen-reactive TCR (TCR11) in a larger tumor burden model. Here, we allowed more time for the tumor to engraft in mice, confirmed tumor engraftment prior to treatment, and randomized mice to one of the following treatments on days 13 and 20 post-tumor injections: (1) PBS ($n$ = 5), (2) CMV-reactive TCR-T cells ($n$ = 4), and (3) CLK3 neoantigen-reactive TCR11-T cells ($n$ = 9) (Figure 5G). Mice treated with CLK3 TCR11-T cells had a significantly lower tumor burden and lived significantly longer compared with the controls (Figures 5H, 5I, and S5I).

## Isolation of neoantigen-reactive CD8$^+$ T cells in myeloid leukemia patients

The discovery of immunogenic neoantigens in splicing factor mutant leukemias raised the question of how such malignancies persist despite potential neoantigen-reactive immune responses. To establish whether mis-splicing-derived neoantigens can elicit productive CD8$^+$ T cell responses in AML patients, we generated a panel of dextramers for the predicted neoantigens. We labeled each dextramer with a unique DNA barcode to detect, quantify, and characterize neoantigen-reactive CD8$^+$ T cells. By performing single-cell RNA- and TCR-seq on dextramer-positive patient CD8$^+$ T cells without *in vitro* manipulation, we interrogated gene-expression profiles of neoantigen-reactive CD8$^+$ T cells in patients *in vivo* and captured TCR sequence of the individual clones (Figure 6A). We generated separate pools of dextramers for HLA-A*02:01$^+$ SRSF2 mutant or ZRSR2 mutant patients. The SRSF2 mutant dextramer pool consisted of one positive control dextramer against an HLA-A*02:01 CMV epitope, three negative control dextramers (1 HLA-matched and 2 HLA-mismatched), and 46 dextramers against SRSF2 mutation-induced neopeptides (Figure 6A; Table S5). The ZRSR2 mutant dextramer pool consisted of two positive control dextramers against HLA-A*02:01 CMV and EBV (Epstein-Barr virus) epitopes, three negative controls, and 12 dextramers against ZRSR2 mutation-induced neopeptides (Table S5).

We stained bulk T cells from patient PBMCs with a dextramer pool. We sorted live CD3$^+$CD8$^+$ dextramer$^+$ cells and followed by single-cell RNA-, TCR-, and dextramer feature barcode sequencing (Figure S6A). For samples where dextramer$^+$ cell numbers were limiting, live CD3$^+$ CD8$^+$ dextramer$^-$ cells were stained with cell hashing antibodies and spiked into dextramer$^+$ populations. Using this approach, we performed single-cell profiling of CD8$^+$ T cells from nine samples across five HLA-A*02:01$^+$ SRSF2 mutant leukemia patients (Figures 6B and S6B; Table S4) yielding 75,343 T cells, of which 46,030 were dextramer$^+$. Neoantigen-reactive TCRs accounted for 56% of sequenced TCRs while viral-reactive were 4.5%. Neoantigen-reactive TCRs were present across naive and memory T cell clusters while viral-specific clones exclusively had a memory profile (green Figure 6B). Based on differentially expressed genes, canonical immune markers, and curated gene signatures,[36,37] cells were categorized into naive and memory subsets with unsupervised clustering revealing 12 distinct CD8$^+$ T cell clusters (Figures 6B and S6C).

Prior studies in patients with solid tumors have identified differences in cell states between viral-reactive and tumor antigen-reactive CD8$^+$ T cells.[36,38] Neoantigen-reactive T cells were most highly concentrated within cluster 3 marked by pronounced reduction in expression of nuclear factor κB (NF-κB) signaling components (Figures 6B–6E). Scoring against reference gene signatures from datasets of human solid TILs (tumor infiltrating lymphocytes)[36,37] showed that neoantigen-reactive T cells consistently exhibited increased signatures of

**Figure 5. Anti-tumor efficacy of CLK3 neoantigen-reactive TCR-T cells against SRSF2 mutant AML cells**

(A) Normalized number of HLA-A*02:01$^+$eGFP$^+$ KO52 cells over time in co-culture with CD8$^+$ T cells expressing indicated CLK3 TCRs or CMV TCR. E:T ratio of 4:1. Mean ± SEM. $p$ values, one-way ANOVA (****$p$.adj ≤ 0.0001).

(B) 24-h co-culture of CD8$^+$ CLK3 TCR9− or TCR11-T cells with KO52 cells at 1:1 E:T ratio. TCR-T cell activation assessed by CD69 and 4-1BB expression in CD8$^+$mTCR$^+$ (mouse TCR) cells, and AML cell killing by activated caspase 3/7 and 7-AAD expression in CD8$^-$ cells.

(C) Percent lysis of firefly luciferase-transduced leukemia cells co-cultured with CD8$^+$ CLK3 TCR9-T cells across E:T ratios. AML cell lines used include SRSF2 mutant KO52 cells with or without HLA-A2 and HLA-A2$^+$ SRSF2 wild-type THP-1 and MV4;11 cells. Mean ± SEM. $p$ values, unpaired Student's t test (**$p$.adj ≤ 0.01, ***$p$.adj ≤ 0.001, ****$p$.adj ≤ 0.0001).

(D) CD69 and 4-1BB expression in CLK3 or CMV TCR-T cells (labeled with CellTrace dye), upon co-culture with BM MNCs or PBMCs from HLA-A*02:01$^+$ SRSF2 mutant AML patients at 1:1 E:T ratio for 24 h. Primary cells from HLA-A*02:01$^+$ SRSF2 wild-type patients or normal donors were negative controls.

(E) Frequency of CD69$^+$ and 4-1BB$^+$ TCR-T cells upon co-culture with HLA-A*02:01$^+$ patient samples. Mean + SEM. $p$ values, unpaired Student's t test (****$p$.adj ≤ 0.0001).

(F) Identification of peptide recognition motif for CLK3 TCR9 (top) and TCR11 (bottom) using alanine scanning. Frequency of TNF-α$^+$ and/or IFN-γ$^+$ TCR-T cells upon 1:1 E:T co-culture with peptide-loaded K562 cells (1 μg/mL). % maximum response relative to unsubstituted CLK3 peptide (y axis). $n$ = 3 technical replicates.

(G) Schema of *in vivo* experiments to test therapeutic efficacy of TCR-T cells expressing CLK3 TCR9 or TCR11.

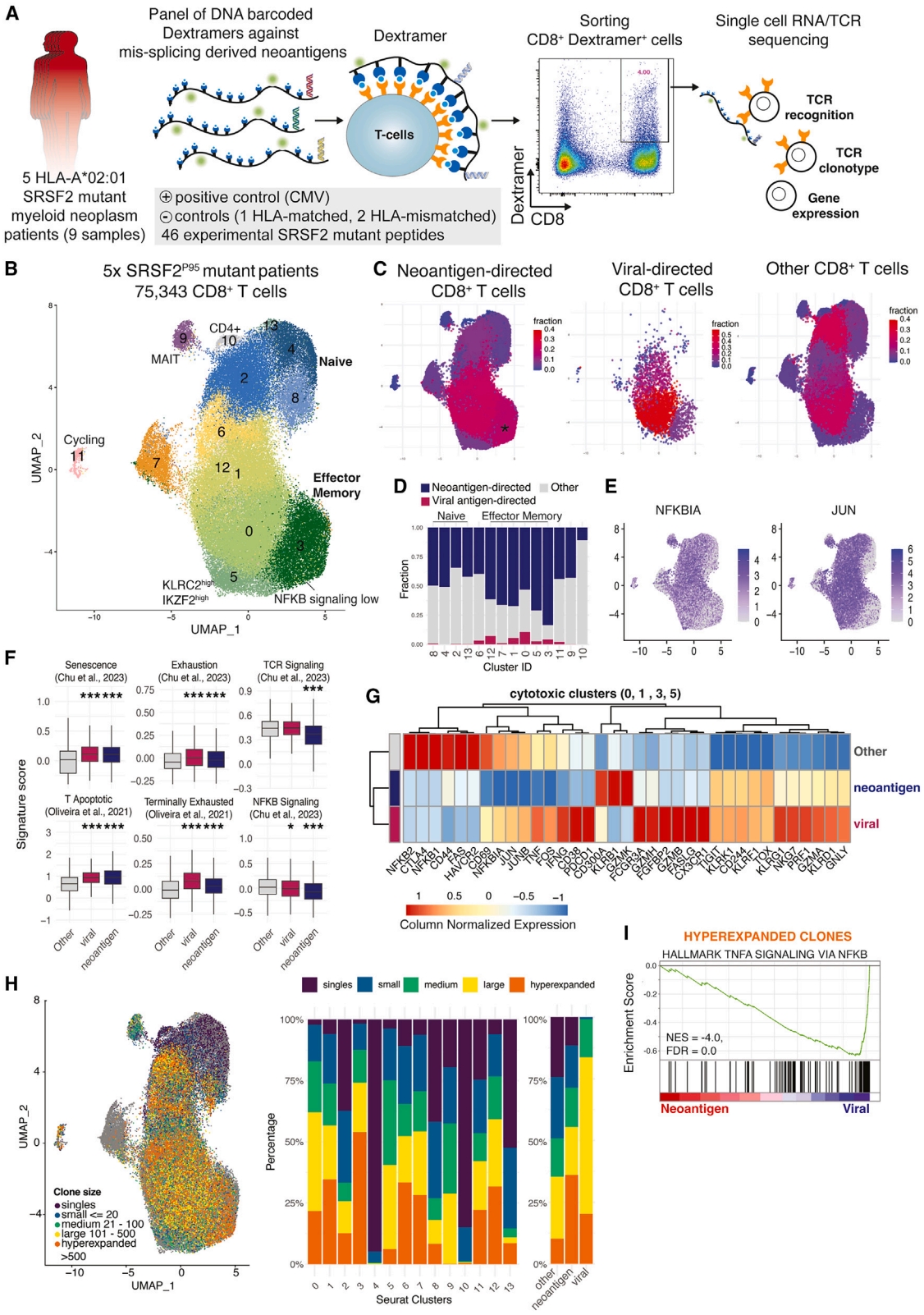
(H) BLI images of animals treated with CLK3 TCR11-T cells, or either PBS or CMV TCR-T cells.

(I) Box and whisker plots of region of interest (ROI) BLI signals in mice from (G). Bar, median; box edges, first and third quartile values; and whisker edges, minimum and maximum values. $p$ value, unpaired Student's t test (*$p$.adj ≤ 0.05, **$p$.adj ≤ 0.01, ***$p$.adj ≤ 0.001, ****$p$.adj ≤ 0.0001).

See also Figure S5.

Please cite this article in press as: Kim et al., Mis-splicing-derived neoantigens and cognate TCRs in splicing factor mutant leukemias, Cell (2025), https://doi.org/10.1016/j.cell.2025.03.047






(legend on next page)

Please cite this article in press as: Kim et al., Mis-splicing-derived neoantigens and cognate TCRs in splicing factor mutant leukemias, Cell (2025), https://doi.org/10.1016/j.cell.2025.03.047





senescence, exhaustion, and apoptosis with reduced TCR and NF-κB signaling (Figures 6F and S6D). Focusing on genes differentially expressed within effector memory CD8 T cells (clusters 0, 1, 3, and 5) revealed marked downregulation of TNF-α via NF-κB signaling genes (*NFKBIA*, *JUN*, *JUNB*, *TNF*, and *FOS*) in neoantigen-reactive T cells (Figure 6G). Gene set enrichment analysis (GSEA) further highlighted diminished activity in TNF-α via NF-κB pathways for neoantigen-reactive cells against the other two groups (Figures S6E and S6F).

We analyzed peripheral blood neoantigen-reactive CD8$^+$ T cells in a separate cohort of four HLA-A*02:01$^+$ ZRSR2 mutant myeloid leukemia patients using the ZRSR2 dextramer pool (Figures S6G–S6I; Table S4). This analysis also revealed enrichment of neoantigen-reactive CD8$^+$ T cells with an effector memory profile and reduced expression of TNF-α via NF-κB pathway genes compared with CD8$^+$ T cells of undefined specificity (Figures S6G–S6I).

We also evaluated the relationship between T cell phenotype and TCR clonal abundance among 10,976 unique clonotypes (Figure 6H). As expected, highly expanded clones resided within the memory compartment. Although CMV-reactive TCRs comprised predominantly large clones, large and hyperexpanded neoantigen-reactive TCRs clones were also highly abundant. GSEA demonstrated significant loss of TNF-α versus NF-κB signaling in hyperexpanded clones in neoantigen-reactive compared with CMV-reactive TCRs (Figure 6I). Taken together, these results revealed that neoantigen-reactive CD8$^+$ T cells are present in the peripheral blood of AML patients, are clonally expanded, and yet have a distinct phenotype from virus-reactive T cells with transcriptional evidence of impaired cytotoxicity.

## Characterizing neoantigen-reactive CD8$^+$ T cells post-curative allogeneic transplant

For patients with high-risk myeloid leukemia, allo-HCT remains the most established curative therapy. In the allo-HCT setting, donor alloreactive T cells recognizing host minor histocompatibility antigens contribute to the GVL effect. We hypothesized that CD8$^+$ T cells reactive to SRSF2 mutation-induced neoantigens may be identifiable in this therapeutic setting.

To investigate this hypothesis, we performed dextramer-based single-cell profiling of pre- and post-transplant PBMCs from a 54-year-old patient with SRSF2$^{R94dup}$-mutant MDS with excessive blasts-1 (patient 2). Pre-transplant sample was collected in the setting of persistent disease following two cycles of decitabine while post-transplant sample was collected 15 months post curative allo-HCT from an HLA-matched donor (Table S4). Post-transplant T cells primarily fell within the memory T cell cluster characterized by high expression of NKG7, PRF1, and CCL5 (cluster 0, Figures 7A, 7B, and S7A). GSEA of differentially expressed genes pre- versus post-transplant within memory CD8$^+$ T cell clusters (0, 1, and 3) revealed enrichment of multiple inflammatory and proliferation pathways in post-transplant T cells, indicative of T cell activation (Figure S7B). Post-transplant effector T cells showed increased expression of hallmark T cell activation and cytotoxicity genes (*PRF1*, *KLRF1*, *GZMB*, and *GNLY*) (Figure S7C). There was no sharing of neoantigen-reactive TCRs pre- and post-transplant, suggesting the emergence of likely donor-derived tumor-reactive TCR clonotypes post-transplant (Figure S7D). Interestingly, when we compared clone size of TCR clonotypes between two time points, hyperexpanded clonotypes (>500 cells) were only present post-transplant (Figure 7C). We profiled another patient (patient 1, Table S4) with SRSR2$^{P95H/WT}$-mutant CMML prior to and 1-year-post curative allo-HCT (Figures S7E–S7G). Once again, post-transplant T cells were most abundant within effector memory CD8$^+$ T cell clusters, and GSEA revealed upregulated inflammatory (TNF-α via NF-κB) and proliferation (mitotic spindle, mTORC1 [mammalian target of rapamycin complex 1]) pathways compared with pre-transplant T cells or other neoantigen-reactive T cells in the same cluster (Figure S7H). TCR clonotypes were completely non-overlapping pre- versus post-transplant, and hyperexpanded TCR clones were only seen post-transplant (Figures S7I and S7J).

We hypothesized that abundant and phenotypically activated neoantigen-reactive TCRs post-transplant might be driven by mis-splicing-derived neoantigens presented by splicing mutant leukemic cells. We found expansion of one TCR clonotype (CAIRGGDSFLFNQPQHF) in the post-transplant sample of patient 2, which was predicted to bind dextramers created with a peptide (CLLPPALFL) derived from an SRSF2 mutant-induced mis-splicing event in *RHOT2* (Figures 7D and 7E). We previously validated the immunogenicity of this peptide (Figures 2A–2F).

To validate the antigen reactivity of the RHOT2 neoantigen-reactive TCR, we transduced human primary T cells with this TCR and confirmed binding to the RHOT2 #5 peptide by staining with RHOT2 #5 p/HLA-I dextramers (Figure S7K). We then co-cultured the TCR-T cells with RHOT2 peptide-pulsed K562 aAPCs (Figures S2C and S4F) and measured dose-dependent production of IFN-γ and TNF-α in CD8$^+$ and CD4$^+$ T cells ($EC_{50}$

**Figure 6. Characterization of neoantigen-reactive CD8$^+$ T cells in splicing factor mutant leukemia patients**

(A) Schema of experiment to evaluate antigen-reactive CD8$^+$ T cells in leukemia patients using 50 distinct DNA-barcoded p/HLA-I-dextramers.

(B) UMAP of CD8$^+$ T cells from five HLA-A*02:01$^+$ SRSF2 mutant patients. Clusters denoted by colors and labeled with inferred cell states.

(C) As in (B) but indicating fraction of total cells within each cluster for neoantigen-directed CD8$^+$ T cells, viral-antigen-directed CD8$^+$ T cells, and CD8$^+$ T cells without dextramer barcode specificity. Asterisk, NF-κB signaling low effector memory CD8$^+$ T cell cluster populated by neoantigen-directed CD8$^+$ T cells.

(D) Fraction of CD8$^+$ T cells in each cell state cluster from (B) based on antigen reactivity.

(E) NFKBIA and JUN expression on UMAP from (B) and reduced expression in NF-κB signaling in effector memory CD8$^+$ T cell cluster 3.

(F) Box and whisker plots of scores for gene signatures[36,37] for antigen-directed CD8$^+$ T cell types indicated in (C). *p* values, t test, and Bonferroni correction (*$p$.adj < 1 × 10$^{-3}$, **$p$.adj < 1 × 10$^{-4}$, ***$p$.adj < 1 × 10$^{-5}$).

(G) Heatmap of selected differentially expressed genes within effector memory CD8$^+$ T cell clusters 0, 1, 3, and 5 from (B).

(H) Left: same UMAP as in (B) but instead T cells marked on basis of TCR clone frequency defined through scTCR-seq. Right: proportional representation of clone sizes within each cluster and type of antigen-directed CD8$^+$ T cell.

(I) Gene set enrichment analysis of differentially expressed genes in hyperexpanded TCR clonotypes in viral-antigen versus neoantigen-directed CD8$^+$ T cells from (H).

See also Figure S6.

Please cite this article in press as: Kim et al., Mis-splicing-derived neoantigens and cognate TCRs in splicing factor mutant leukemias, Cell (2025), https://doi.org/10.1016/j.cell.2025.03.047



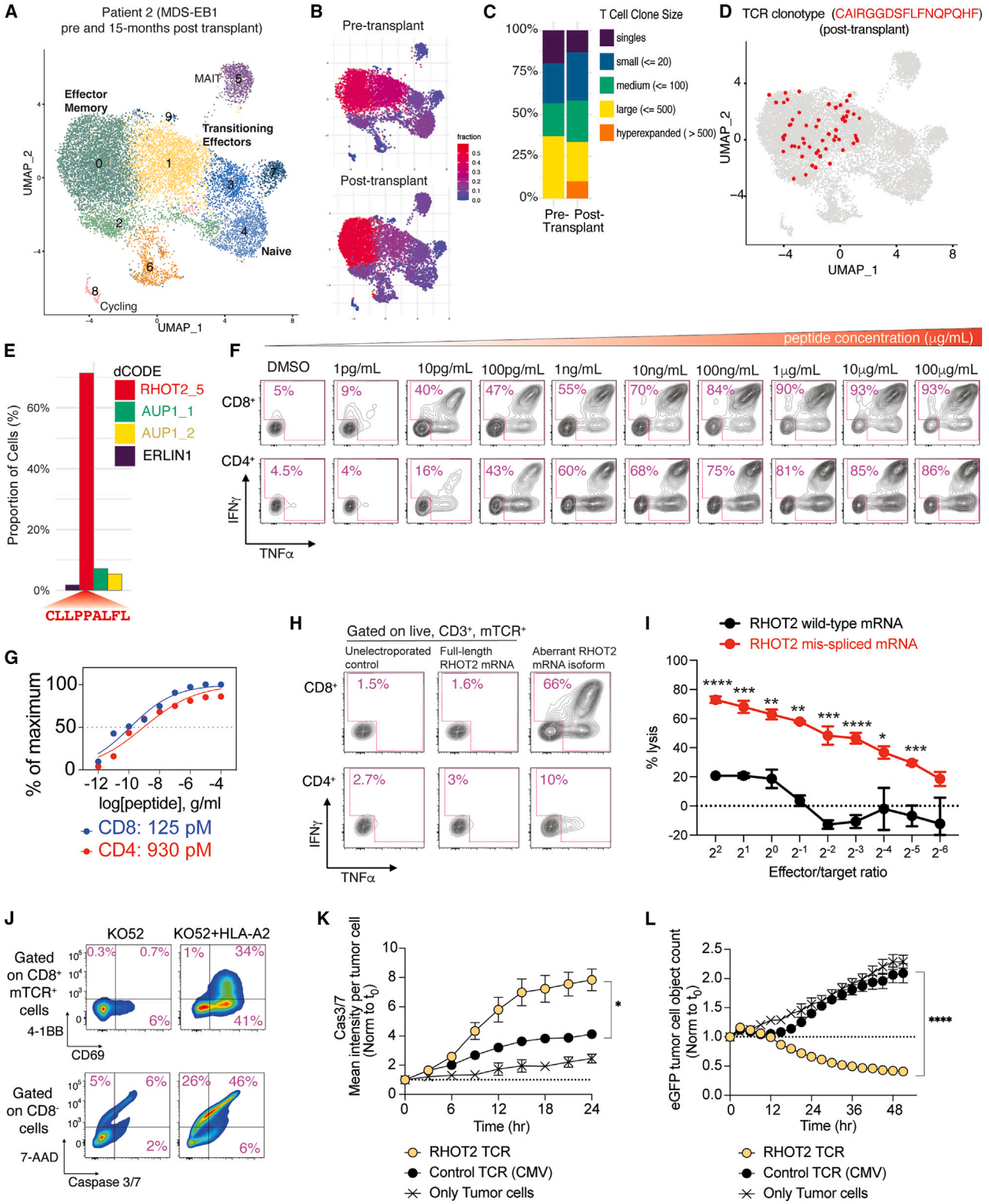



**Figure 7. Neoantigen-reactive T cells and TCRs pre and post HLA-matched allogeneic transplantation in splicing factor mutant leukemia patients**

(A) UMAP of CD8+ T cells pre- and post- transplant for patient 2 (see Table S4).

*(legend continued on next page)*

Please cite this article in press as: Kim et al., Mis-splicing-derived neoantigens and cognate TCRs in splicing factor mutant leukemias, Cell (2025), https://doi.org/10.1016/j.cell.2025.03.047





125 and 930 pM, respectively) (Figures 7F and 7G). In addition, we co-cultured RHOT2 TCR-T cells with K562 aAPCs electroporated with *RHOT2* WT or mis-spliced mRNA and confirmed RHOT2 neoantigen-induced T cell activation and cytotoxicity (Figures 7H and 7I). Moreover, we co-cultured RHOT2 TCR-T cells with SRSF2 mutant KO52 cells. RHOT2 TCR-T cells upregulated surface CD69 and 4-1BB only when co-cultured with KO52 cells expressing HLA-A*02:01, while simultaneously inducing apoptosis of KO52 cells (Figures 7J and 7K). Lastly, RHOT2 TCR-T cells potently suppressed growth of KO52 cells *in vitro* (Figure 7L). Overall, these data provide evidence of an endogenous T cell response to mis-splicing-derived neoantigens in myeloid leukemia patients in the post-transplant setting, with donor-derived T cell clones potentially contributing to long-term anti-leukemic T cell immunity.

## DISCUSSION

Here, we present a series of splicing-factor-mutation-induced neoantigens and cognate TCRs which can selectively recognize and lyse spliceosomal mutant AML cells. Recently, a number of studies documented neoantigens arising from RNA mis-splicing events that involve non-protein coding regions such as long non-coding RNAs,[39] introns,[40] circular RNAs,[41] and transposable elements.[42] While some of these neoantigens have been described as recurrent, the molecular basis for why these neoantigens would be shared between patients is elusive. By contrast, the molecular mechanisms of how mutations in RNA splicing factors lead to highly recurrent mis-splicing events across patients are now increasingly understood.[7–10] The recurrent nature of mis-splicing occurs due to the defined roles that each splicing factor plays in RNA splicing and the sequence-specific impact of these mutations on RNA recognition. Thus, many neoantigens translated from these mis-splicing events are shared across patients. Moreover, RNA splicing factor mutations lead to numerous mis-splicing events which can generate hundreds to even thousands of potential mis-splicing-derived neoepitopes. The sheer number of neoepitopes created by these mutations increases the likelihood that neoantigens can be presented on a wide diversity of HLA alleles.

Our studies characterizing neoantigen-reactive CD8$^+$ T cells provide new insights into the immune dysregulation that occurs in MDS/AML patients. We found that CD8$^+$ T cells reactive to mis-splicing-derived neoantigens exist in the peripheral blood of patients with spliceosomal mutant leukemias. However, these T cells have defective NF-κB proinflammatory pathways, which might explain why these cells are unable to mount effective antigen-reactive immune responses in patients. Single-cell profiling of matched pre- and post-transplant samples suggested that allogeneic donor T cells have gene-expression profiles of activation and cytotoxicity compared with the patient's endogenous neoantigen-reactive T cells. Moreover, we discovered donor-derived CD8$^+$ T cells against RHOT2 neoantigen in one patient. Future studies are warranted to test the therapeutic efficacy and safety of this TCR, along with CLK3 neoantigen-reactive TCRs, against leukemias with splicing factor mutations. Of note, two recent clinical trials of TCR-edited T cells in solid cancers have created and infused gene-edited T cells with TCRs against up to three distinct epitopes into the same patient.[43,44] Such approaches could be readily applied for myeloid leukemia patients with splicing factor mutations given the multitude of public mis-splicing-derived neoantigens created by these mutations.

While our studies focused on patients with myeloid neoplasms, it is important to note that mutations in splicing factors are recurrent in the pre-disease settings of clonal hematopoiesis (CH) and clonal cytopenia of undetermined significance (CCUS). Importantly, mutations in RNA splicing factors when present in CH confer higher risk of transformation to overt myeloid leukemia.[45,46] It would be very intriguing to utilize a similar dextramer-based approach to study the characteristics of such neoantigen-reactive T cells in the earliest stages of pre-malignant disease. Such studies may motivate neoantigen vaccination approaches in splicing factor mutant CH and CCUS using collections of mis-splicing-derived peptides or the mRNAs encoding them in the future.

### Limitations of the study

While we focused on the most common HLA class-I allele in our patient cohorts (i.e., HLA-A*02:01), future studies are needed to identify neoantigens presentable on other common HLA-I alleles as well as their cognate TCRs. On a similar note, our therapeutic

(B) As in (A) but indicating fraction of total cells within each cluster for pre- and post-transplant samples.
(C) Proportional representation of TCR clone sizes within pre- versus post-transplant antigen-directed CD8$^+$ T cells.
(D) As in (A) but indicating a CDR3β TCR clonotype specific to RHOT2 peptide #5. This TCR clonotype largely maps to post-transplant effector memory CD8$^+$ T cells.
(E) Proportion of dextramer barcodes among cells with the TCR clonotype indicated in (D).
(F) IFN-γ and TNF-α ICS FACS plots of CD8$^+$ and CD4$^+$ T cells transduced with RHOT2 TCR from (D) and exposed to K562 cells pulsed with increasing concentrations of RHOT2 peptide.
(G) Percent maximum T cell activation across peptide concentrations from (F). $EC_{50}$ concentrations below graph.
(H) IFN-γ and TNF-α ICS FACS plots of CD8$^+$ or CD4$^+$ T cells transduced with RHOT2 TCR and exposed to K562 cells electroporated with *RHOT2* wild-type or mis-spliced mRNA.
(I) Percent lysis of HLA-A*02:01 K562 cells expressing *RHOT2* mRNA isoforms and exposed to CD8$^+$ RHOT2 TCR-T cells across E:T ratios. Mean ± SEM. *p* values, unpaired Student's t test (*$p$.adj ≤ 0.05, **$p$.adj ≤ 0.01, ***$p$.adj ≤ 0.001, ****$p$.adj ≤ 0.0001).
(J) 24-h co-culture of RHOT2 TCR-T cells with KO52 cells with or without HLA-A*02:01 at 1:1 E:T ratio. TCR-T cell activation assessed by CD69 and 4-1BB expression in CD8$^+$mTCR$^+$ cells, and AML cell killing by activated caspase 3/7 and 7-AAD expression in CD8$^-$ cells.
(K) Time course detection of HLA-A*02:01$^+$ KO52 cell apoptosis upon co-culture with CD8$^+$ RHOT2 TCR-T cells at 1:1 E:T ratio, using caspase 3/7-based eGFP probe. Mean ± SEM. *p* values, one-way ANOVA (*$p$.adj ≤ 0.05).
(L) Normalized number of HLA-A*02:01$^+$eGFP$^+$ KO52 cells in co-culture with CD8$^+$ T cells expressing RHOT2 or CMV TCR over time. E:T ratio of 4:1. Mean ± SEM. *p* values, ordinary one-way ANOVA (****$p$.adj ≤ 0.0001).
See also Figure S7.

Please cite this article in press as: Kim et al., Mis-splicing-derived neoantigens and cognate TCRs in splicing factor mutant leukemias, Cell (2025), https://doi.org/10.1016/j.cell.2025.03.047



approach is theoretically achievable through targeting HLA-class-II-restricted neoantigens, which is especially relevant in neoantigen-based vaccination.

In our *in vivo* experiments using immune-deficient mice, we focused on early time points after T cell transfer to evaluate the therapeutic efficacy of neoantigen-reactive TCR-T cells due to the potential for graft-versus-host-disease, which could confound interpretation of overall survival.[47,48] Future *in vivo* studies in HLA-matched humanized mice are needed, particularly as we optimize neoantigen-reactive TCR constructs to improve TCR-T cell activation and persistence.[49,50] Finally, while we propose neoantigen-based vaccination as a viable therapeutic strategy in patients with myeloid malignancies, we note that endogenous neoantigen-reactive $CD8^+$ T cells in patients with active disease are dysfunctional. It is therefore unclear whether patients with active MDS or AML could respond to neoantigen vaccination. As such, careful consideration of when to perform vaccination would be critical, and future studies to better understand the molecular mechanisms of immune dysfunction in patients with MDS or AML will provide critical insight on how to maximize the therapeutic potential of our proposed therapies.

## RESOURCE AVAILABILITY

### Lead contact

Further information and requests for resources and reagents should be directed to the lead contact, Omar Abdel-Wahab (abdelwao@mskcc.org).

### Materials availability

All materials from this study are included in this article and its supplemental information files. Any material that can be shared will be released via a material transfer agreement for non-commercial usage.

### Data and code availability

- Data are available from the corresponding authors upon reasonable request. Single-cell RNA-seq data have been deposited to GEO under accession ID GSE268157.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

## ACKNOWLEDGMENTS

W.J.K. was supported by a Medical Scientist Training Program grant from the NIH (T32GM152349) to the Weill Cornell/Rockefeller/Sloan Kettering Tri-institutional MD-PhD Program, NIH/NHLBI (1F30HL172602-01), and American Society of Hematology. E.I.C. is a Damon Runyon Cancer Research Foundation fellow. O.A.-W. and R.K.B. were supported by the Edward P. Evans Foundation, NIH/NCI (R01 CA251138 and R01 CA283364), and NIH/NHLBI (R01 HL128239), and the Leukemia & Lymphoma Society. O.A.-W. was supported by NIH/NCI (R01 CA242020 and P50 CA254838-01). P.K., J.M., and O.A.-W. were supported by Break Through Cancer. O.A.-W. is supported by the Neil S. Hirsch Foundation. R.K.B. was supported by NIH/NHLBI (R01 HL151651), Mark Foundation for Cancer Research, and Paul G. Allen Frontiers Group (8023-20). R.K.B. is a Scholar of the Leukemia & Lymphoma Society (1344-18) and holds the McIlwain Family Endowed Chair in Data Science. C.A.K. was supported by R37 CA259177, R01 CA269733, R01 CA286507, and P50CA217694, the Metropoulos Family Foundation, and the Parker Institute for Cancer Immunotherapy. We acknowledge FHCC's Scientific Computing Infrastructure (ORIP S10 OD028685), NCI Cancer Center Support Grant (CCSG; P30 CA08748), Cycle for Survival, and Marie-Josée and Henry R. Kravis, Center for Molecular Oncology.

## AUTHOR CONTRIBUTIONS

Conceptualization, W.J.K., E.I.C., E.D.N., I.E., C.A.K., R.K.B., and O.A.-W.; RNA-seq analyses, E.I.C., E.D.N., A.B., B.Z., B.G., and R.K.B.; mass spectrometry, E.D.N., I.E., E.Y.S., H.M., and R.K.B.; primary human cell experiments, W.J.K., I.E., J.P.B., S.X.L., N.F., C.C., P.Z., T.F., J.L., J.H.W., A.M.L., R.F.S., S.D., B.M.U., M.T., T.K., R.C., P.K., J.M., M.G., T.M., A.D., and C.A.K.; bioinformatics, data analyses, and visualization, E.I.C., E.D.N., E.Y.S., K.-T.L., A.B., J.R., B.Z., H.M., B.D.G., and R.K.B.; resources, J.B., S.D., R.K.B., and O.A.-W.; writing – review and editing, all authors; supervision, C.A.K., R.K.B., O.A.-W.; project administration, W.J.K., E.I.C., R.K.B., and O.A.-W.; funding acquisition, C.A.K., R.K.B., and O.A.-W.

## DECLARATION OF INTERESTS

R.K.B. and O.A.-W. are founders and scientific advisors of Codify Therapeutics; they hold equity and receive research funding from this company. K.T.-L. is an employee of Codify Therapeutics. W.J.K., E.I.C., R.K.B., and O.A.-W. are inventors on a patent related to this study. C.A.K. and I.E. are inventors on TCR patents unrelated to this study and are recipients of licensing revenue shared according to MSK institutional policies. C.A.K. has consulted for or is on scientific advisory boards for Achilles Therapeutics, Affini-T Therapeutics, Aleta BioTherapeutics, Bellicum Pharmaceuticals, Bristol Myers Squibb, Catamaran Bio, Cell Design Labs, Decheng Capital, G1 Therapeutics, Klus Pharma, Obsidian Therapeutics, PACT Pharma, Roche/Genentech, and Royalty Pharma. C.A.K. is a scientific co-founder and equity holder in Affini-T Therapeutics. R.K.B. is a founder and scientific advisor of Synthesize Bio and holds equity in this company. O.A.-W. has served as a consultant for Amphista Therapeutics and MagnetBio and is on scientific advisory boards of Envisagenics Inc. and Harmonic Discovery Inc.; O.A.-W. received research funding from Nurix Therapeutics, Minovia Therapeutics, and LOXO Oncology unrelated to this study.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Primary cells and cell lines
  - Patient PBMC and BM MNC samples
  - Animals
- METHOD DETAILS
  - Plasmids and peptides
  - Lentiviral production and AML cell line transduction
  - T2 HLA-A2 shift assay
  - *In vitro* immunogenicity screening
  - RT-PCR
  - Western blotting
  - Dual fluorophore-conjugated dextramer staining of human primary T cells
  - Dextramer-based single-cell profiling of patient PBMCs
  - Library preparation and sequencing for dextramer-guided single cell RNA and TCR sequencing
  - TCR transduction
  - Tumor-immune co-culture experiments
  - Alanine scanning
  - Adoptive cell transfer xenograft model
  - HLA-I immunoprecipitation, peptide purification, and mass spectrometry
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - RNA-seq data analysis for differential RNA splicing
  - Identification of mis-splicing-derived candidate neopeptides
  - Dextramer-guided single cell RNA- and TCR-sequencing
  - Single cell data pre-processing, dimension reduction and clustering
  - Cell hashing demux and dextramer background thresholding
  - TCR clonotype assignment and clonal analysis
  - Dextramer / TCR pair identification

## SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cell.2025.03.047.

Received: May 29, 2024
Revised: February 6, 2025
Accepted: March 28, 2025
Published: April 23, 2025

## REFERENCES

1. Daver, N., Wei, A.H., Pollyea, D.A., Fathi, A.T., Vyas, P., and DiNardo, C.D. (2020). New directions for emerging therapies in acute myeloid leukemia: the next chapter. Blood Cancer J. *10*, 107. https://doi.org/10.1038/s41408-020-00376-1.
2. Stanley, R.F., and Abdel-Wahab, O. (2022). Dysregulation and therapeutic targeting of RNA splicing in cancer. Nat. Cancer *3*, 536–546. https://doi.org/10.1038/s43018-022-00384-z.
3. Bradley, R.K., and Anczuków, O. (2023). RNA splicing dysregulation and the hallmarks of cancer. Nat. Rev. Cancer *23*, 135–155. https://doi.org/10.1038/s41568-022-00541-7.
4. Yoshida, K., Sanada, M., Shiraishi, Y., Nowak, D., Nagata, Y., Yamamoto, R., Sato, Y., Sato-Otsubo, A., Kon, A., Nagasaki, M., et al. (2011). Frequent pathway mutations of splicing machinery in myelodysplasia. Nature *478*, 64–69. https://doi.org/10.1038/nature10496.
5. Papaemmanuil, E., Cazzola, M., Boultwood, J., Malcovati, L., Vyas, P., Bowen, D., Pellagatti, A., Wainscoat, J.S., Hellstrom-Lindberg, E., Gambacorti-Passerini, C., et al. (2011). Somatic SF3B1 mutation in myelodysplasia with ring sideroblasts. N. Engl. J. Med. *365*, 1384–1395. https://doi.org/10.1056/NEJMoa1103283.
6. Graubert, T.A., Shen, D., Ding, L., Okeyo-Owuor, T., Lunn, C.L., Shao, J., Krysiak, K., Harris, C.C., Koboldt, D.C., Larson, D.E., et al. (2011). Recurrent mutations in the U2AF1 splicing factor in myelodysplastic syndromes. Nat. Genet. *44*, 53–57. https://doi.org/10.1038/ng.1031.
7. Kim, E., Ilagan, J.O., Liang, Y., Daubner, G.M., Lee, S.C.W., Ramakrishnan, A., Li, Y., Chung, Y.R., Micol, J.B., Murphy, M.E., et al. (2015). SRSF2 Mutations Contribute to Myelodysplasia by Mutant-Specific Effects on Exon Recognition. Cancer Cell *27*, 617–630. https://doi.org/10.1016/j.ccell.2015.04.006.
8. Ilagan, J.O., Ramakrishnan, A., Hayes, B., Murphy, M.E., Zebari, A.S., Bradley, P., and Bradley, R.K. (2015). U2AF1 mutations alter splice site recognition in hematological malignancies. Genome Res. *25*, 14–26. https://doi.org/10.1101/gr.181016.114.
9. Darman, R.B., Seiler, M., Agrawal, A.A., Lim, K.H., Peng, S., Aird, D., Bailey, S.L., Bhavsar, E.B., Chan, B., Colla, S., et al. (2015). Cancer-Associated SF3B1 Hotspot Mutations Induce Cryptic 3′ Splice Site Selection through Use of a Different Branch Point. Cell Rep. *13*, 1033–1045. https://doi.org/10.1016/j.celrep.2015.09.053.
10. Inoue, D., Polaski, J.T., Taylor, J., Castel, P., Chen, S., Kobayashi, S., Hogg, S.J., Hayashi, Y., Pineda, J.M.B., El Marabti, E., et al. (2021). Minor intron retention drives clonal hematopoietic disorders and diverse cancer predisposition. Nat. Genet. *53*, 707–718. https://doi.org/10.1038/s41588-021-00828-9.
11. Bigot, J., Lalanne, A.I., Lucibello, F., Gueguen, P., Houy, A., Dayot, S., Ganier, O., Gilet, J., Tosello, J., Nemati, F., et al. (2021). Splicing Patterns in SF3B1-Mutated Uveal Melanoma Generate Shared Immunogenic Tumor-Specific Neoepitopes. Cancer Discov. *11*, 1938–1951. https://doi.org/10.1158/2159-8290.CD-20-0555.
12. Patnaik, M.M., and Tefferi, A. (2024). Chronic myelomonocytic leukemia: 2024 update on diagnosis, risk stratification and management. Am. J. Hematol. *99*, 1142–1165. https://doi.org/10.1002/ajh.27271.
13. Bernard, E., Tuechler, H., Greenberg, P.L., Hasserjian, R.P., Arango Ossa, J.E., Nannya, Y., Devlin, S.M., Creignou, M., Pinel, P., Monnier, L., et al. (2022). Molecular International Prognostic Scoring System for Myelodysplastic Syndromes. NEJM Evid. *1*, EVIDoa2200008. https://doi.org/10.1056/EVIDoa2200008.
14. Jahn, E., Saadati, M., Fenaux, P., Gobbi, M., Roboz, G.J., Bullinger, L., Lutsik, P., Riedel, A., Plass, C., Jahn, N., et al. (2023). Clinical impact of the genomic landscape and leukemogenic trajectories in non-intensively treated elderly acute myeloid leukemia patients. Leukemia *37*, 2187–2196. https://doi.org/10.1038/s41375-023-01999-6.
15. Bottomly, D., Long, N., Schultz, A.R., Kurtz, S.E., Tognon, C.E., Johnson, K., Abel, M., Agarwal, A., Avaylon, S., Benton, E., et al. (2022). Integrative analysis of drug response and clinical outcome in acute myeloid leukemia. Cancer Cell *40*, 850–864.e9. https://doi.org/10.1016/j.ccell.2022.07.002.
16. Inoue, D., Chew, G.L., Liu, B., Michel, B.C., Pangallo, J., D'Avino, A.R., Hitchman, T., North, K., Lee, S.C.W., Bitner, L., et al. (2019). Spliceosomal disruption of the non-canonical BAF complex in cancer. Nature *574*, 432–436. https://doi.org/10.1038/s41586-019-1646-9.
17. Tyner, J.W., Tognon, C.E., Bottomly, D., Wilmot, B., Kurtz, S.E., Savage, S.L., Long, N., Schultz, A.R., Traer, E., Abel, M., et al. (2018). Functional genomic landscape of acute myeloid leukaemia. Nature *562*, 526–531. https://doi.org/10.1038/s41586-018-0623-z.
18. Madan, V., Kanojia, D., Li, J., Okamoto, R., Sato-Otsubo, A., Kohlmann, A., Sanada, M., Grossmann, V., Sundaresan, J., Shiraishi, Y., et al. (2015). Aberrant splicing of U12-type introns is the hallmark of ZRSR2 mutant myelodysplastic syndrome. Nat. Commun. *6*, 6042. https://doi.org/10.1038/ncomms7042.
19. Lavallée, V.P., Baccelli, I., Krosl, J., Wilhelm, B., Barabé, F., Gendron, P., Boucher, G., Lemieux, S., Marinier, A., Meloche, S., et al. (2015). The transcriptomic landscape and directed chemical interrogation of MLL-rearranged acute myeloid leukemias. Nat. Genet. *47*, 1030–1037. https://doi.org/10.1038/ng.3371.
20. Jurtz, V., Paul, S., Andreatta, M., Marcatili, P., Peters, B., and Nielsen, M. (2017). NetMHCpan-4.0: Improved Peptide-MHC Class I Interaction Predictions Integrating Eluted Ligand and Peptide Binding Affinity Data. J. Immunol. *199*, 3360–3368. https://doi.org/10.4049/jimmunol.1700893.
21. Hosken, N.A., and Bevan, M.J. (1990). Defective presentation of endogenous antigen by a cell line expressing class I molecules. Science *248*, 367–370. https://doi.org/10.1126/science.2326647.
22. Loftus, D.J., Squarcina, P., Nielsen, M.B., Geisler, C., Castelli, C., Odum, N., Appella, E., Parmiani, G., and Rivoltini, L. (1998). Peptides derived from self-proteins as partial agonists and antagonists of human CD8+ T-cell clones reactive to melanoma/melanocyte epitope MART1(27-35). Cancer Res. *58*, 2433–2439.
23. Roudko, V., Bozkus, C.C., Orfanelli, T., McClain, C.B., Carr, C., O'Donnell, T., Chakraborty, L., Samstein, R., Huang, K.L., Blank, S.V., et al. (2020). Shared Immunogenic Poly-Epitope Frameshift Mutations in Microsatellite Unstable Tumors. Cell *183*, 1634–1649.e17. https://doi.org/10.1016/j.cell.2020.11.004.
24. Cimen Bozkus, C., Blazquez, A.B., Enokida, T., and Bhardwaj, N. (2021). A T-cell-based immunogenicity protocol for evaluating human antigen-specific responses. Star Protoc. *2*, 100758. https://doi.org/10.1016/j.xpro.2021.100758.
25. Ley, T.J., Miller, C., Ding, L., Raphael, B.J., Mungall, A.J., Robertson, A.G., Hoadley, K., Triche, T.J., Laird, P.W., et al.; Cancer; Genome; Atlas; Research Network (2013). Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N. Engl. J. Med. *368*, 2059–2074. https://doi.org/10.1056/NEJMoa1301689.
26. Abelin, J.G., Keskin, D.B., Sarkizova, S., Hartigan, C.R., Zhang, W., Sidney, J., Stevens, J., Lane, W., Zhang, G.L., Eisenhaure, T.M., et al. (2017). Mass Spectrometry Profiling of HLA-Associated Peptidomes in Mono-allelic Cells Enables More Accurate Epitope Prediction. Immunity *46*, 315–326. https://doi.org/10.1016/j.immuni.2017.02.007.
27. Pellagatti, A., Armstrong, R.N., Steeples, V., Sharma, E., Repapi, E., Singh, S., Sanchi, A., Radujkovic, A., Horn, P., Dolatshad, H., et al. (2018). Impact of spliceosome mutations on RNA splicing in myelodysplasia:

dysregulated genes/pathways and clinical associations. Blood *132*, 1225–1240. https://doi.org/10.1182/blood-2018-04-843771.

28. Gopalsamy, A., Bennett, E.M., Shi, M., Zhang, W.G., Bard, J., and Yu, K. (2012). Identification of pyrimidine derivatives as hSMG-1 inhibitors. Bioorg. Med. Chem. Lett. *22*, 6636–6641. https://doi.org/10.1016/j.bmcl.2012.08.107.
29. Langer, L.M., Bonneau, F., Gat, Y., and Conti, E. (2021). Cryo-EM reconstructions of inhibitor-bound SMG1 kinase reveal an autoinhibitory state dependent on SMG8. eLife *10*, e72353. https://doi.org/10.7554/eLife.72353.
30. Apcher, S., Daskalogianni, C., Lejeune, F., Manoury, B., Imhoos, G., Heslop, L., and Fåhraeus, R. (2011). Major source of antigenic peptides for the MHC class I pathway is produced during the pioneer round of mRNA translation. Proc. Natl. Acad. Sci. USA *108*, 11572–11577. https://doi.org/10.1073/pnas.1104104108.
31. Hirano, N., Butler, M.O., Xia, Z., Ansén, S., von Bergwelt-Baildon, M.S., Neuberg, D., Freeman, G.J., and Nadler, L.M. (2006). Engagement of CD83 ligand induces prolonged expansion of CD8+ T cells and preferential enrichment for antigen specificity. Blood *107*, 1528–1536. https://doi.org/10.1182/blood-2005-05-2073.
32. Butler, M.O., Lee, J.S., Ansén, S., Neuberg, D., Hodi, F.S., Murray, A.P., Drury, L., Berezovskaya, A., Mulligan, R.C., Nadler, L.M., and Hirano, N. (2007). Long-lived antitumor CD8+ lymphocytes for adoptive therapy generated using an artificial antigen-presenting cell. Clin. Cancer Res. *13*, 1857–1867. https://doi.org/10.1158/1078-0432.CCR-06-1905.
33. Chandran, S.S., Ma, J., Klatt, M.G., Dündar, F., Bandlamudi, C., Razavi, P., Wen, H.Y., Weigelt, B., Zumbo, P., Fu, S.N., et al. (2022). Immunogenicity and therapeutic targeting of a public neoantigen derived from mutated PIK3CA. Nat. Med. *28*, 946–957. https://doi.org/10.1038/s41591-022-01786-3.
34. de Castro, E., Sigrist, C.J.A., Gattiker, A., Bulliard, V., Langendijk-Genevaux, P.S., Gasteiger, E., Bairoch, A., and Hulo, N. (2006). ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. Nucleic Acids Res. *34*, W362–W365. https://doi.org/10.1093/nar/gkl124.
35. Gattinoni, L., Finkelstein, S.E., Klebanoff, C.A., Antony, P.A., Palmer, D.C., Spiess, P.J., Hwang, L.N., Yu, Z., Wrzesinski, C., Heimann, D.M., et al. (2005). Removal of homeostatic cytokine sinks by lymphodepletion enhances the efficacy of adoptively transferred tumor-specific CD8+ T cells. J. Exp. Med. *202*, 907–912. https://doi.org/10.1084/jem.20050732.
36. Oliveira, G., Stromhaug, K., Klaeger, S., Kula, T., Frederick, D.T., Le, P.M., Forman, J., Huang, T., Li, S., Zhang, W., et al. (2021). Phenotype, specificity and avidity of antitumour CD8+ T cells in melanoma. Nature *596*, 119–125. https://doi.org/10.1038/s41586-021-03704-y.
37. Chu, Y., Dai, E., Li, Y., Han, G., Pei, G., Ingram, D.R., Thakkar, K., Qin, J.J., Dang, M., Le, X., et al. (2023). Pan-cancer T cell atlas links a cellular stress response state to immunotherapy resistance. Nat. Med. *29*, 1550–1562. https://doi.org/10.1038/s41591-023-02371-y.
38. Caushi, J.X., Zhang, J., Ji, Z., Vaghasia, A., Zhang, B., Hsiue, E.H.C., Mog, B.J., Hou, W., Justesen, S., Blosser, R., et al. (2021). Transcriptional programs of neoantigen-specific TIL in anti-PD-1-treated lung cancers. Nature *596*, 126–132. https://doi.org/10.1038/s41586-021-03752-4.
39. Barczak, W., Carr, S.M., Liu, G., Munro, S., Nicastri, A., Lee, L.N., Hutchings, C., Ternette, N., Klenerman, P., Kanapin, A., et al. (2023). Long non-coding RNA-derived peptides are immunogenic and drive a potent anti-tumour response. Nat. Commun. *14*, 1078. https://doi.org/10.1038/s41467-023-36826-0.
40. Kahles, A., Lehmann, K.V., Toussaint, N.C., Huser, M., Stark, S.G., Sachsenberg, T., Stegle, O., Kohlbacher, O., Sander, C., and Cancer Genome Atlas Research Network, and Ratsch, G. (2018). Comprehensive Analysis of Alternative Splicing Across Tumors from 8,705 Patients. Cancer Cell *34*, 211–224.e6. https://doi.org/10.1016/j.ccell.2018.07.001.
41. Huang, D., Zhu, X., Ye, S., Zhang, J., Liao, J., Zhang, N., Zeng, X., Wang, J., Yang, B., Zhang, Y., et al. (2024). Tumour circular RNAs elicit anti-tumour immunity by encoding cryptic peptides. Nature *625*, 593–602. https://doi.org/10.1038/s41586-023-06834-7.
42. Merlotti, A., Sadacca, B., Arribas, Y.A., Ngoma, M., Burbage, M., Goudot, C., Houy, A., Rocañín-Arjó, A., Lalanne, A., Seguin-Givelet, A., et al. (2023). Noncanonical splicing junctions between exons and transposable elements represent a source of immunogenic recurrent neo-antigens in patients with lung cancer. Sci. Immunol. *8*, eabm6359. https://doi.org/10.1126/sciimmunol.abm6359.
43. Foy, S.P., Jacoby, K., Bota, D.A., Hunter, T., Pan, Z., Stawiski, E., Ma, Y., Lu, W., Peng, S., Wang, C.L., et al. (2023). Non-viral precision T cell receptor replacement for personalized cell therapy. Nature *615*, 687–696. https://doi.org/10.1038/s41586-022-05531-1.
44. Parkhurst, M., Goff, S.L., Lowery, F.J., Beyer, R.K., Halas, H., Robbins, P.F., Prickett, T.D., Gartner, J.J., Sindiri, S., Krishna, S., et al. (2024). Adoptive transfer of personalized neoantigen-reactive TCR-transduced T cells in metastatic colorectal cancer: phase 2 trial interim results. Nat. Med. *30*, 2586–2595. https://doi.org/10.1038/s41591-024-03109-0.
45. Abelson, S., Collord, G., Ng, S.W.K., Weissbrod, O., Mendelson Cohen, N., Niemeyer, E., Barda, N., Zuzarte, P.C., Heisler, L., Sundaravadanam, Y., et al. (2018). Prediction of acute myeloid leukaemia risk in healthy individuals. Nature *559*, 400–404. https://doi.org/10.1038/s41586-018-0317-6.
46. Desai, P., Mencia-Trinchant, N., Savenkov, O., Simon, M.S., Cheang, G., Lee, S., Samuel, M., Ritchie, E.K., Guzman, M.L., Ballman, K.V., et al. (2018). Somatic mutations precede acute myeloid leukemia years before diagnosis. Nat. Med. *24*, 1015–1023. https://doi.org/10.1038/s41591-018-0081-z.
47. Ito, R., Katano, I., Kawai, K., Yagoto, M., Takahashi, T., Ka, Y., Ogura, T., Takahashi, R., and Ito, M. (2017). A Novel Xenogeneic Graft-Versus-Host Disease Model for Investigating the Pathological Role of Human CD4+ or CD8+ T Cells Using Immunodeficient NOG Mice. Am. J. Transplant. *17*, 1216–1228. https://doi.org/10.1111/ajt.14116.
48. King, M.A., Covassin, L., Brehm, M.A., Racki, W., Pearson, T., Leif, J., Laning, J., Fodor, W., Foreman, O., Burzenski, L., et al. (2009). Human peripheral blood leucocyte non-obese diabetic-severe combined immunodeficiency interleukin-2 receptor gamma chain gene mouse model of xenogeneic graft-versus-host-like disease and the role of host major histocompatibility complex. Clin. Exp. Immunol. *157*, 104–118. https://doi.org/10.1111/j.1365-2249.2009.03933.x.
49. Drakes, D.J., Rafiq, S., Purdon, T.J., Lopez, A.V., Chandran, S.S., Klebanoff, C.A., and Brentjens, R.J. (2020). Optimization of T-cell Receptor-Modified T Cells for Cancer Therapy. Cancer Immunol. Res. *8*, 743–755. https://doi.org/10.1158/2326-6066.CIR-19-0910.
50. Yi, F., Cohen, T., Zimmerman, N., Dündar, F., Zumbo, P., Eltilib, R., Brophy, E.J., Arkin, H., Feucht, J., Gormally, M.V., et al. (2024). CAR-engineered lymphocyte persistence is governed by a FAS ligand/FAS autoregulatory circuit. Preprint at bioRxiv. https://doi.org/10.1101/2024.02.26.582108.
51. Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L.D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., et al. (2019). Welcome to the Tidyverse. J. Open Source Softw. *4*, 1686. https://doi.org/10.21105/joss.01686.
52. Kolde, R. (2018). pheatmap: Pretty Heatmaps. R package version 1.0.12. https://github.com/raivokolde/pheatmap.
53. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842. https://doi.org/10.1093/bioinformatics/btq033.
54. Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. *10*, R25. https://doi.org/10.1186/gb-2009-10-3-r25.
55. Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal. *17*, 10–12. https://doi.org/10.14806/ej.17.1.200.

56. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). 1000 Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352.

57. Leinonen, R., Sugawara, H., and Shumway, M.; International Nucleotide Sequence Database Collaboration (2011). The sequence read archive. Nucleic Acids Res., D19–D21. https://doi.org/10.1093/nar/gkq1019.

58. Li, B., and Dewey, C.N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics *12*, 323. https://doi.org/10.1186/1471-2105-12-323.

59. Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics *25*, 1105–1111. https://doi.org/10.1093/bioinformatics/btp120.

60. Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat. Protoc. *7*, 562–578. Erratum in Oct 2014: Nat. Protoc. *9*, 2513. https://doi.org/10.1038/nprot.2012.016.

61. Katz, Y., Wang, E.T., Airoldi, E.M., and Burge, C.B. (2010). Analysis and design of RNA sequencing experiments for identifying isoform regulation. Nat. Methods *7*, 1009–1015. https://doi.org/10.1038/nmeth.1528.

62. Lawrence, M., Huber, W., Pagès, H., Aboyoun, P., Carlson, M., Gentleman, R., Morgan, M.T., and Carey, V.J. (2013). Software for computing and annotating genomic ranges. PLoS Comput Biol. *9*, e1003118. https://doi.org/10.1371/journal.pcbi.1003118.

63. Bembom, O., and Ivanek, R. (2024). seqLogo: Sequence logos for DNA sequence alignments. R package. version 1.72.0.

64. O'Donnell, T.J., Rubinsteyn, A., and Laserson, U. (2020). MHCflurry 2.0: Improved Pan-Allele Prediction of MHC Class I-Presented Peptides by Incorporating Antigen Processing. Cell Syst. *11*, 42–48.e7. https://doi.org/10.1016/j.cels.2020.06.010.

65. Seiler, M., Yoshimi, A., Darman, R., Chan, B., Keaney, G., Thomas, M., Agrawal, A.A., Caleb, B., Csibi, A., Sean, E., et al. (2018). H3B-8800, an orally available small-molecule splicing modulator, induces lethality in spliceosome-mutant cancers. Nat. Med. *24*, 497–504. https://doi.org/10.1038/nm.4493.

66. Durham, B.H., Getta, B., Dietrich, S., Taylor, J., Won, H., Bogenberger, J.M., Scott, S., Kim, E., Chung, Y.R., Chung, S.S., et al. (2017). Genomic analysis of hairy cell leukemia identifies novel recurrent genetic alterations. Blood *130*, 1644–1648. https://doi.org/10.1182/blood-2017-01-765107.

67. Ishihama, Y., Rappsilber, J., and Mann, M. (2006). Modular stop and go extraction tips with stacked disks for parallel and multidimensional Peptide fractionation in proteomics. J. Proteome Res. *5*, 988–994. https://doi.org/10.1021/pr050385q.

68. Hao, Y., Stuart, T., Kowalski, M.H., Choudhary, S., Hoffman, P., Hartman, A., Srivastava, A., Molla, G., Madad, S., Fernandez-Granda, C., and Satija, R. (2024). Dictionary learning for integrative, multimodal and scalable single-cell analysis. Nat. Biotechnol. *42*, 293–304. https://doi.org/10.1038/s41587-023-01767-y.

69. Butler, A., Hoffman, P., Smibert, P., Papalexi, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat. Biotechnol. *36*, 411–420. https://doi.org/10.1038/nbt.4096.

70. Borcherding, N., Bormann, N.L., and Kraus, G. (2020). scRepertoire: An R-based toolkit for single-cell immune receptor analysis. F1000Res *9*, 47. https://doi.org/10.12688/f1000research.22139.2.

Please cite this article in press as: Kim et al., Mis-splicing-derived neoantigens and cognate TCRs in splicing factor mutant leukemias, Cell (2025), https://doi.org/10.1016/j.cell.2025.03.047



## STAR★METHODS

### KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| **Antibodies** | | |
| FITC mouse anti-human CD3, clone: SK7 | BD Biosciences | Cat#349201; RRID: AB_400405 |
| APC mouse anti-human CD8, clone: SK1 | BD Biosciences | Cat#340659; RRID: AB_400529 |
| PE/Cy7 mouse anti-human CD8, clone: SK1 | BD Biosciences | Cat#335805; RRID: AB_399980 |
| FITC mouse anti-human CD3, clone: SK7 | BioLegend | Cat#344803; RRID: AB_2043992 |
| APC mouse anti-human CD8, clone: RPA-T8 | BioLegend | Cat#301049; RRID: AB_314132 |
| Brilliant Violet 785 mouse anti-human CD4, clone: RPA-T4 | BioLegend | Cat#300554; RRID: AB_2564382 |
| PE mouse anti-human IFN-g, clone: B27 | BioLegend | Cat#506507; RRID: AB_315440 |
| PE/Cy7 mouse anti-human TNF-a, clone: Mab11 | BioLegend | Cat#502930; RRID: AB_2204079 |
| BV785 mouse anti-human CD69, clone: FN50 | BioLegend | Cat#310932; RRID: AB_2563696 |
| APC mouse anti-human CD137 (4-1BB), clone: 4B4-1 | BioLegend | Cat#309810; RRID: AB_830672 |
| Brilliant Violet 421 mouse anti-human CD137 (4-1BB), clone: 4B4-1 | BioLegend | Cat#309820; RRID: AB_2563830 |
| PE mouse anti-human HLA-A2, clone: BB7.2 | BioLegend | Cat#343306; RRID: AB_1877227 |
| Brilliant Violet 605 Armenian hamster anti-mouse TCR β, clone: H57–597 | BioLegend | Cat#109241; RRID: AB_2629563 |
| Brilliant Violet 785 Armenian hamster anti-mouse TCR β, clone: H57-597 | BioLegend | Cat#109249; RRID: AB_2810347 |
| TotalSeq-C0251 anti-human hashtag 1 antibody, clone: LNH-94; 2M2 | BioLegend | Cat#394661; RRID: AB_2801031 |
| Rabbit polyclonal anti-CLK3 | Cell Signaling Technology | Cat#3256S; RRID: AB_2082274 |
| Rabbit polyclonal anti-MIRO2 | Proteintech | Cat#11237-1-AP; RRID: AB_2179539 |
| Rabbit polyclonal anti-C16orf70 | Proteintech | Cat#20602-1-AP; RRID: AB_10694825 |
| Mouse monoclonal anti-β-actin | Sigma-Aldrich | Cat#A1978; RRID: AB_476692 |
| Purified mouse anti-human CD28, clone: CD28.2 | BD Biosciences | Cat#555726; RRID: AB_396069 |
| Purified mouse anti-human CD49d, clone: 9F10 | BD Biosciences | Cat#555502; RRID: AB_395892 |
| FITC mouse anti-human HLA-ABC, clone: G46-2.6 | BD Biosciences | Cat#555552; RRID: AB_395935 |
| *InVivo*MAb anti-human MHC class I (HLA-A, HLA-B, HLA-C), clone: W6/32 | BioXCell | Cat#BE0079; RRID: AB_1107730 |
| Purified mouse anti-human HLA-A2 monoclonal antibody (BB7.2) | ThermoFisher Scientific | Cat#MA5-16586; RRID: AB_2538089 |
| IRDye 800CW donkey anti-rabbit IgG secondary antibody | LI-COR | Cat#926-32213; RRID: AB_621848 |
| IRDye 680RD donkey anti-mouse IgG secondary antibody | LI-COR | Cat#926-68072; RRID: AB_10953628 |
| ImmunoCult human CD3/CD28 T cell activator | STEMCELL Technologies | Cat#10991; RRID: AB_2827806 |
| dCODE dextramer (10x) | Immudex | This paper; Table S5 |
| **Bacterial and virus strains** | | |
| One Shot Stbl3 chemically competent *E. coli* | ThermoFisher Scientific | Cat#C737303 |
| **Biological samples** | | |
| Human peripheral blood mononuclear cells, frozen | STEMCELL Technologies | Cat#70025 |
| Human bone marrow mononuclear cells, frozen | STEMCELL Technologies | Cat#70001.2 |
| Human peripheral blood $CD8^+$ T Cells, frozen | STEMCELL Technologies | Cat#200-0164 |

(*Continued on next page*)

Please cite this article in press as: Kim et al., Mis-splicing-derived neoantigens and cognate TCRs in splicing factor mutant leukemias, Cell (2025), https://doi.org/10.1016/j.cell.2025.03.047





*Continued*

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| **Chemicals, peptides, and recombinant proteins** | | |
| hSMG-1 inhibitor 11e | MedChemExpress | HY-124760 |
| Custom peptides | GenScript | This paper; Tables S2 and S3 |
| Recombinant human GM-CSF | PeproTech | Cat#300-03 |
| Recombinant human IL-4 | PeproTech | Cat#200-04 |
| Recombinant human Flt3-ligand | PeproTech | Cat#300-19 |
| Recombinant human IL-1β | PeproTech | Cat#200-01B |
| Recombinant human IL-2 | PeproTech | Cat#200-02 |
| Recombinant human IL-7 | PeproTech | Cat#200-07 |
| Recombinant human IL-15 | PeproTech | Cat#200-15 |
| Recombinant human IL-15 receptor alpha Fc | PeproTech | Cat#200-15RA |
| R848 (Resiquimod) | InvivoGen | Cat#tlrl-r848-10 |
| LPS-SM Ultrapure | InvivoGen | Cat#tlrl-smlps |
| DNase I solution (1 mg/mL) | STEMCELL Technologies | Cat#100-0762 |
| XbaI | New England Biolabs | Cat#R0145L |
| RNase-free DNase set | Qiagen | Cat#79254 |
| Beta 2 microglobulin | Bio-Rad | Cat#PHP135 |
| RetroNectin recombinant human fibronectin fragment | Takara | Cat#T100B |
| **Critical commercial assays** | | |
| Human IFN-γ ELISpot kit | R&D Systems | XEL285 |
| Bright-Glo luciferase assay system | Promega | E2620 |
| CellTiter-Glo luminescent cell viability assay | Promega | G7573 |
| CellEvent caspase-3/7 green flow cytometry assay kit | Invitrogen | C10427 |
| CellEvent caspase-3/7 detection reagents | Invitrogen | C10432 |
| CellTrace blue cell proliferation kit, for flow cytometry | Invitrogen | C34568 |
| Chromium Next GEM Single Cell 5′ Kit v2 | 10x Genomics | PN-1000263 |
| Chromium 5′ Feature Barcode Kit | 10x Genomics | PN-1000541 |
| **Deposited data** | | |
| Dextramer-based single-cell sequencing data for this paper | NCBI Gene Expression Omnibus | GEO: GSE268157 |
| **Experimental models: Cell lines** | | |
| 293T | ATCC | CRL-3216 |
| 293GP | Chandran et al.[33] | N/A |
| K-562 | ATCC | CCL-243 |
| MV-4-11 | ATCC | CRL-9591 |
| T2 | ATCC | CRL-1992 |
| KO52 | AcceGen | ABC-TC0536 |
| **Oligonucleotides** | | |
| GGACTACTATGGACCTTCACGT | This paper | CLK3 RT-PCR forward primer |
| CCAGGTTCCCCACAATCTCA | This paper | CLK3 RT-PCR reverse primer |
| **Recombinant DNA** | | |
| HLA-A*02:01 cDNA | PMID: 35704596 | N/A |
| RD114 | Chandran et al.[33] | N/A |
| pMSGV-RV | Chandran et al.[33] | N/A |
| pENTR223 CD80 | DNASU | Clone ID: 515155 |
| pENTR223 CD83 | DNASU | Clone ID: 506972 |
| pLX303 | Addgene | Cat#25897 |
| VSV.G | Addgene | Cat#14888 |

(*Continued on next page*)

*Continued*

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| psPAX2 | Addgene | Cat#12260 |
| Lenti-luciferase-P2A-Neo | Addgene | Cat#105621 |
| pLX303 CD80 | This paper | N/A |
| pLX303 CD83 | This paper | N/A |
| pcDNA3.1(+) CLK3 full-length cDNA | This paper | N/A |
| pcDNA3.1(+) CLK3 mis-spliced cDNA | This paper | N/A |
| pcDNA3.1(+) RHOT2 full-length cDNA | This paper | N/A |
| pcDNA3.1(+) RHOT2 mis-spliced cDNA | This paper | N/A |
| pcDNA3.1(+) c16orf70 full-length cDNA | This paper | N/A |
| pcDNA3.1(+) c16orf70 mis-spliced cDNA | This paper | N/A |
| **Software and algorithms** | | |
| CellRanger v7.0.1 | PMID: 28091601 | https://www.10xgenomics.com/support/software/cell-ranger/latest |
| tidyverse | Wickham et al.[51] (https://doi.org/10.21105/joss.01686) | Bioconductor |
| Seurat | PMID: 37231261, PMID: 29608179 | https://github.com/satijalab/seurat |
| Harmony | PMID: 31740819 | https://portals.broadinstitute.org/harmony/ |
| Gene Set Enrichment Analysis | PMID: 16199517 | https://www.gsea-msigdb.org/gsea/index.jsp |
| scRepertoire | PMID: 32789006 | https://github.com/ncborcherding/scRepertoire |
| Pheatmap | R Kolde[52] (Kolde, Raivo. "Pheatmap: pretty heatmaps." R package *version* 1.2 (2019): 726.) | https://github.com/raivokolde/pheatmap |
| FlowJo V8.7 | TreeStar (BD Biosciences) | https://www.flowjo.com/ |
| Prism 9.0 | GraphPad | https://www.graphpad.com |
| BEDTools | Quinlan and Hall[53] | https://github.com/arq5x/bedtools2 |
| Bowtie | Langmead et al.[54] | https://bowtie-bio.sourceforge.net/bowtie2/index.shtml |
| Cutadapt | Martin[55] | https://cutadapt.readthedocs.io/en/stable/installation.html |
| Picard (v2.8.2) | Broad Institute | https://broadinstitute.github.io/picard/ |
| Python (v2.7.2 or later) | Python Software | https://www.python.org/downloads/ |
| SAMtools | Li et al.[56] | https://samtools.sourceforge.net/ |
| SRA Toolkit | Leinonen et al.[57] | https://github.com/ncbi/sra-tools |
| RSEM v1.2.4 | Li and Dewey[58] | https://deweylab.github.io/RSEM/ |
| TopHat v2.0.8b | Trapnell et al.[59,60] | https://ccb.jhu.edu/software/tophat/index.shtml |
| MISO v2.0 | Katz et al.[61] | http://hollywood.mit.edu/burgelab/miso/ |
| tidyverse | Wickham et al.[51] (https://doi.org/10.21105/joss.01686) | Bioconductor |
| GenomicRanges | Lawrence et al.[62] | Bioconductor |
| seqLogo | Bembom and Ivanek[63] | Bioconductor |
| NetMHCpan4.0 | Jurtz et al.[20] | https://services.healthtech.dtu.dk/services/NetMHCpan-4.0/ |
| MHCFlurry2.0 | O'Donnell et al.[64] | https://github.com/openvax/mhcflurry |
| R version 4.2.2 | R Core Team | https://www.R-project.org/ |
| **Other** | | |
| High-Capacity RNA-to-cDNA kit | ThermoFisher Scientific | Cat#4387406 |
| UltraComp eBead plus compensation beads | ThermoFisher Scientific | Cat#01-3333-42 |
| ArC amine reactive compensation bead kit | ThermoFisher Scientific | Cat#A10346 |

(*Continued on next page*)

Please cite this article in press as: Kim et al., Mis-splicing-derived neoantigens and cognate TCRs in splicing factor mutant leukemias, Cell (2025), https://doi.org/10.1016/j.cell.2025.03.047





*Continued*

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| LIVE/DEAD fixable blue dead cell stain kit | ThermoFisher Scientific | Cat#L34962 |
| NuPAGE LDS sample buffer (4X) | ThermoFisher Scientific | Cat#NP0008 |
| NuPAGE sample reducing agent (10X) | ThermoFisher Scientific | Cat#NP0009 |
| NuPAGE Bis-Tris mini protein gels, 4-12%, 1.0-1.5 mm | ThermoFisher Scientific | Cat#NP0322BOX |
| SYBR safe DNA gel stain | ThermoFisher Scientific | Cat#S33102 |
| Nunc poly-D-lysine-coated multidishes | ThermoFisher Scientific | Cat#152035 |
| Lipofectamine 3000 transfection reagent | ThermoFisher Scientific | Cat#L3000015 |
| UltraPure herring sperm DNA solution | ThermoFisher Scientific | Cat#15634017 |
| Pierce BCA protein assay kits | ThermoFisher Scientific | Cat#PI23225 |
| Neon transfection system 100 μ L kit | ThermoFisher Scientific | Cat#MPK10025 |
| Pierce C18 spin tips & columns | ThermoFisher Scientific | Cat#87784 |
| Gateway LR clonase II enzyme mix | ThermoFisher Scientific | Cat#11791020 |
| RNeasy Mini Kit | Qiagen | Cat#74106 |
| QIAquick PCR purification kit | Qiagen | Cat#28104 |
| Multiplex PCR kit | Qiagen | Cat#206143 |
| GolgiPlug (containing Brefeldin A) | BD Biosciences | Cat#555029 |
| GolgiStop (containing Monensin) | BD Biosciences | Cat#554724 |
| BD Cytofix/Cytoperm fixation/permeabilization kit | BD Biosciences | Cat#554714 |
| X-VIVO 15 hematopoietic serum-free culture media | Lonza | Cat#BW04-418Q |
| GemCell U.S. origin human serum | GeminiBio | Cat#100-512-100 |
| RIPA buffer (10X) | Cell Signaling Technology | Cat#9806 |
| cOmplete, mini protease inhibitor cocktail | Roche | Cat#11836153001 |
| PhosSTOP | Roche | Cat#4906845001 |
| Immun-Blot PVDF membrane, precut, 25 x 28 cm | Bio-Rad | Cat#1620255 |
| Intercept (TBS) blocking buffer | LI-COR | Cat#927-60001 |
| GammaBind Plus Sepharose and antibody purification resin | Cytiva | Cat#17088601 |
| HiScribe T7 ARCA mRNA kit (with tailing) | New England Biolabs | Cat#E2060S |
| IVISbrite RediJect D-luciferin | Revvity Healthy Sciences, Inc | Cat#50-209-9345 |
| EasySep human T cell isolation kit | STEMCELL Technologies | Cat#17951 |
| EasySep human CD8$^+$ T cell isolation kit | STEMCELL Technologies | Cat#17953 |

## EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

### Primary cells and cell lines

Peripheral blood mononuclear cells (PBMC; catalog #70025) and bone marrow mononuclear cells (BM MNC; catalog #70001.2) from HLA-A*02:01$^+$ healthy donors were purchased from STEMCELL Technologies (catalog #70025). Isolated peripheral blood CD8$^+$ T cells from healthy donors were also purchased from STEMCELL Technologies (#200-0164). Bulk T cells or CD8$^+$ T cells were also isolated with EasySep human T or CD8$^+$ T cell isolation kits (STEMCELL Technologies, catalog #17951 or 17953, respectively). Human primary T cells were cultured in RPMI medium supplemented with 10% human serum (GeminiBio, catalog #100-512-100), GlutaMAX (ThermoFisher Scientific, catalog #35050061), and 10 mM HEPES. T2 cell line was purchased from ATCC (catalog #CRL-1992) and cultured in IMDM medium with 20% FCS. K562 cell line was purchased from ATCC (catalog #CCL-243) and cultured in IMDM medium with 10% FCS. Isogenic K562 cell line with knockin SRSF2$^{WT/P95H}$ mutation in the endogenous *SRSF2* locus were generated as previously described.[65] KO52 AML cell line was purchased from AcceGen (catalog #ABC-TC0536) and cultured in RPMI medium with 10% FCS. 293T cell line was purchased from ATCC (catalog #CRL-3216) and 293GP cell line was purchased from Takara Bio (catalog #631458), both of which were cultured in DMEM medium with 10% FCS. All media were supplemented with 10 IU/mL penicillin-streptomycin.

Please cite this article in press as: Kim et al., Mis-splicing-derived neoantigens and cognate TCRs in splicing factor mutant leukemias, Cell (2025), https://doi.org/10.1016/j.cell.2025.03.047



### Patient PBMC and BM MNC samples

Studies were approved by the Institutional Review Boards (IRB) of Memorial Sloan Kettering Cancer Center (MSKCC) and conducted in accordance with the Declaration of Helsinki protocol. Primary PBMC or BM MNC samples from de-identified patients with MDS, AML, CMML, or myeloproliferative neoplasms (MPNs) were utilized. Mutational genotyping of each sample was performed by the MSKCC IMPACT assay as previously described.[66] HLA typing was previously performed for the majority of patients; for those without such information, genomic DNA was extracted using DNeasy Blood & Tissue Kit (Qiagen, #69506) and sent for HistoGenetics HLA typing services. Patient samples were obtained as part of the IRB protocol #16-171, to which all subjected provided informed consent.

### Animals

All animal studies were performed in accordance with a MSKCC Institutional Animal Care and Use Committee-approved protocol (13-04-003). Four to six-week-old male and female NOD.Cg-*Prkdc*$^{scid}$ *Il2rg*$^{tm1Wjl}$/SzJ (NSG) mice were purchased from Jackson Laboratory and housed in pathogen-free conditions at the MSKCC vivarium.

## METHOD DETAILS

### Plasmids and peptides

HLA-A*02:01 cDNA plasmid was provided by Taha Merghoub. pENTR223 CD80 (clone ID: 515155) and pENTR223 CD83 (clone ID: 506972) were purchased from DNASU and cloned into pLX303 (Addgene, catalog #25897) using Gateway LR clonase II enzyme mix (ThermoFisher Scientific, catalog #11791020). VSV.G and psPAX2 for 293T transfection were purchased from Addgene (catalog #14888 and #12260, respectively). NLS-eGFP and NLS-mCherry sequences were synthesized and cloned into pMSGV-1 RV plasmid by Genscript. Full-length and mis-spliced isoform cDNA sequences of *CLK3*, *RHOT2*, and *c16orf70* were cloned into pcDNA3.1$^+$ vector by Genscript. These vectors were linearized using XbaI (New England Biolabs, catalog #R0145L) and purified using QIAquick PCR purification kit (Qiagen, catalog #28104). Linearized vectors were used as templates for *in vitro* transcription using HiScribe T7 ARCA mRNA kit (New England Biolabs, catalog #E2060S). Candidate TCR $\alpha/\beta$ chain sequences were cloned into pMSGV-1 RV plasmid as previously described[33] and the variable region of each chain was fused to a modified murine constant chain, and two chains were linked by a furinSGSGP2A linker. For *in vivo* experiments, pMSGV-1 plasmid encoding a TCR of interest and CD8$\beta$ and CD8$\alpha$ (downstream of the TCR, separated by a P2A sequence) was used. >98% purity peptides were synthesized by Genscript, with standard removal of trifluoroacetic acid (TFA) and replacement with chloride. Peptides were stored at -80° C in multiple aliquots to reduce the number of freeze-thaw cycles.

### Lentiviral production and AML cell line transduction

293T cells were plated in a tissue culture-treated 6-well plate at 2 million cells per well in DMEM medium with 10% FCS with antibiotics. The next day, medium was replaced with fresh DMEM medium with 10% FCS, 25 $\mu$M chloroquine diphosphate, and without antibiotics (3 ml per well). Cells were transfected with 5 $\mu$g target plasmid of interest, 2.5 $\mu$g VSV.G, and 3.75 $\mu$g psPAX2, and cells were incubated at 37 ° C for 6 hours. Medium was replaced with fresh DMEM medium with 10% FCS and without antibiotics (2 ml per well). After 48 hours, lentiviral supernatant was collected and filtered through 0.45 $\mu$m filters (Sigma-Aldrich, catalog #WHA9914-2504). 500,000 K562 cells in 1ml IMDM medium with 10% FCS, along with 1ml lentiviral supernatant and 7.5 $\mu$g/ml polybrene, were plated in a tissue culture-treated 12-well plate and centrifuged at 2,300 RPM for 90 minutes at 32 ° C. Spinoculation was repeated with fresh virus the second day.

### T2 HLA-A2 shift assay

T2 cells were washed twice with PBS, resuspended in IMDM medium without serum, and plated in 96-well U bottom plates (100,000 cells per well) with 1 $\mu$g/ml $\beta$2 microglobulin (Biorad, catalog #PHP135). A dose range of each experimental peptide was added in technical triplicates. Cells were incubated at 37° C for 18 hours, subsequently stained with HLA-ABC antibody (BD Biosciences, clone G46-2.6) and analyzed using BD LSR Fortessa Flow Cytometer.

### *In vitro* immunogenicity screening

Vials of cryopreserved PBMCs from HLA-A*02:01$^+$ healthy donors were obtained commercially from STEMCELL Technologies and stored in liquid nitrogen until use. Synthesized experimental peptides were tested for T cell immunogenicity as previously described.[24] On Day 0, a vial of PBMCs were thawed and cultured in X-VIVO 15 medium (Lonza, catalog #04-418Q) in 96-well U bottom plates (100,000 cells per well). On Day 1, PBMCs were supplemented with a cytokine cocktail of 1000 IU/ml GM-CSF (PeproTech, catalog #300-03), 500 IU/ml IL-4 (PeproTech, catalog #200-04), and 500 ng/ml Flt3-Ligand (PeproTech, catalog #200-19) for APC maturation. On Day 2, PBMCs were stimulated with experimental peptides individually at 1 $\mu$g/ml concentration, along with 10 $\mu$g/ml R848 (InvivoGen, catalog #tlrl-r848-5), 100 ng/ml LPS (InvivoGen, catalog #tlrl-smlps) and 10 ng/ml IL-1$\beta$ (PeproTech, catalog #200-01B). PBMCs were subsequently cultured in RPMI medium supplemented with 10% human serum, GlutaMAX, 10 mM HEPES, and a cytokine cocktail of 10 IU/ml IL-2 (PeproTech, catalog #200-02), 10 ng/ml IL-7 (PeproTech, catalog #200-07), and 10 ng/ml IL-15 (PeproTech, catalog #200-15) for 1 week to expand T cells. Medium was replaced with fresh

Please cite this article in press as: Kim et al., Mis-splicing-derived neoantigens and cognate TCRs in splicing factor mutant leukemias, Cell (2025), https://doi.org/10.1016/j.cell.2025.03.047



cytokine-containing RPMI medium on Day 3 and Day 5, and fresh cytokine-free medium on Day 8. PBMCs were harvested on Day 10, counted, and plated for downstream T cell activation experiments as follows.

#### *Intracellular cytokine staining (ICS) and flow cytometry*

Peptide-stimulated PBMCs were plated in new 96-well U bottom plates (200,000 cells per well) and re-stimulated with the same peptide (10 $\mu$g/ml) or DMSO, each in triplicates, along with 0.5 mg/ml each of CD28 and CD49d antibodies (BD Biosciences, catalog #555726 and 555502). PBMCs were incubated at 37° C for 1 hour, treated with 0.67 and 1 $\mu$l/ml protein transport inhibitors containing Monensin and Brefeldin A (BD Biosciences, catalog #554724 and 555029), and incubated at 37° C for another 8-12 hours. PBMCs were subsequently stained for viability and extracellular markers for 20 minutes, fixed and permeabilized using commercial buffers (BD Biosciences, catalog #554713), stained for intracellular cytokines for 30 minutes, and analyzed using BD LSR Fortessa Flow Cytometer.

#### *IFN-$\gamma$ ELISpot*

Human IFN-$\gamma$ ELISpot Kit was purchased from Bio-Techne (catalog #XEL285) and used for the experiment. Peptide-stimulated PBMCs were plated in an ELISpot plate (50,000 cells per well), along with 0.5 mg/ml each of CD28 and CD49d antibodies, and the cells were incubated at 37° C for 48 hours. The plate was subsequently processed as per manufacturer's instructions and imaged with the ImmunoSpot imager.

### RT-PCR

Total RNA was extracted using Qiagen RNeasy Mini Kit (catalog #74106) as per manufacturer's instructions, with an additional step of DNase treatment (Qiagen, catalog #79254). 1$\mu$g RNA was used for reverse transcription (RT) using High-Capacity RNA-to-cDNA Kit (ThermoFisher Scientific, catalog #4387406) with following RT protocol: 37° C for 1 hour, and 95° C for 5 minutes. cDNA product was diluted 1:1 with $H_2O$, and 1 $\mu$l was used as a template for PCR using the Qiagen Multiplex PCR kit (catalog #206143) in a total reaction volume of 25 $\mu$l, using the following PCR protocol: (1) 95° C for 15 minutes, and then (2) 35 cycles of 95° C for 30 seconds, 57° C for 90 seconds, and 72° C for 1 minute, followed by (3) 60° C for 30 minutes. Bands were visualized by staining with SYBR Safe DNA gel stain (ThermoFisher Scientific, catalog #S33102).

### Western blotting

Cells were harvested, washed in PBS twice, resuspended in 1X RIPA buffer (Cell Signaling Technology, catalog #9806S) with cOmplete Mini protease inhibitor cocktail (Roche, catalog #11836170001) and PhosSTOP phosphatase inhibitor (Roche, catalog #4906845001), vortexed vigorously and incubated on ice for 30 minutes. Lysed cells were spun down at 15,000 RPM for 10 minutes at 4° C, and protein supernatant was transferred to a new Eppendorf tube. Protein concentration was measured using Pierce BCA Protein Assay Kit (ThermoFisher Scientific, catalog #23225) and normalized with NuPAGE 4X LDS sample buffer (catalog #NP0008) and NuPAGE 10X sample reducing agent (catalog #NP0004). Normalized protein was boiled at 95° C for 5 minutes and loaded into each well of the NuPAGE protein gels (catalog #NP0322BOX) for gel electrophoresis. Protein was transferred to Immuno-Blot PVDF membrane (Biorad, catalog #1620255) at 100 V for 1 hour. Transferred membrane were blocked with Intercept blocking buffer (LI-COR, catalog #927-60001) at room temperature for 1 hour and incubated with primary antibodies at 4° C overnight. Membrane was subsequently incubated with LI-COR secondary antibodies (catalog #926-32213, 926-68072) at room temperature for 1 hour, and bands were visualized using Odyssey CLx Imager. Primary antibodies used are as follows: CLK3 (Cell Signaling Technology, catalog #3256S), RHOT2/MIRO2 (Proteintech, catalog #11237-1-AP), C16orf70 (Proteintech, catalog #20602-1-AP), and $\beta$-actin (Sigma-Aldrich, catalog #A1978).

### Dual fluorophore-conjugated dextramer staining of human primary T cells

HLA-A*02:01 dextramers bound to experimental peptides, feature-barcoded, and conjugated to PE or APC were purchased from Immudex. In PBS with 2% FCS and 0.1 g/L Herring sperm DNA (Invitrogen, catalog #15634017), 1-3 million PBMCs were stained with dual fluorophore-conjugated dextramers for 30 minutes at room temperature as per manufacturer's instructions, and subsequently with surface antibodies against CD3 and CD8 for additional 20 minutes at room temperature. PBMCs were washed twice, resuspended in PBS with 2% FCS containing DAPI, and analyzed using BD LSR Fortessa Flow Cytometer Fortessa or sorted using BD FACSAria or FACS Symphony S6. For primary T cells that were transduced with candidate TCR and hence clonal, 20,000-50,000 cells were stained with dextramers as per manufacturer's instructions.

### Dextramer-based single-cell profiling of patient PBMCs

Vial(s) of patient PBMCs was thawed and treated with 0.1 mg/ml DNase I (STEMCELL Technologies, catalog #100-0762) for 15 minutes at room temperature to dissociate cell clumps. For samples whose cell number was greater than 10 million, T cells were enriched by magnet-based negative selection using EasySep human T cell enrichment kit (STEMCELL Technologies, catalog #19051). PBMCs or enriched T cells were stained with a panel of feature-barcoded peptide-dextramers for 30 minutes at room temperature as per manufacturer's instructions, and subsequently with surface antibodies against CD3 and CD8 for additional 20 minutes at room temperature. Cells were washed twice, resuspended in PBS with 2% FCS containing DAPI, and sorted into (1) $DAPI^-CD3^+CD8^+dextramer^+$ and (2) $DAPI^-CD3^+CD8^+dextramer^-$ T cells using BD FACSAria or FACS Symphony S6. If the cell number

Please cite this article in press as: Kim et al., Mis-splicing-derived neoantigens and cognate TCRs in splicing factor mutant leukemias, Cell (2025), https://doi.org/10.1016/j.cell.2025.03.047



of population (1) was lower than 16,000, population (2) was stained with TotalSeq-C0251 hashtag antibody (BioLegend, catalog #394661) and doped into population (1) for downstream single-cell experiments.

### Library preparation and sequencing for dextramer-guided single cell RNA and TCR sequencing

The single-cell RNA-seq of FACS-sorted cell suspensions was performed on Chromium instrument (10X genomics) following the user guide manual for 5′ v2 chemistry. In brief, FACS-sorted cells were washed once with PBS containing 1% bovine serum albumin (BSA) and resuspended in PBS containing 1% BSA to a final concentration of 700–1,200 cells per μl. The viability of cells was above 80%, as confirmed with 0.2% (w/v) Trypan Blue staining (Countess II). Cells were captured in droplets. Following reverse transcription and cell barcoding in droplets, emulsions were broken and cDNA purified using Dynabeads MyOne SILANE followed by PCR amplification per manual instruction.

Approximately 10,000 cells were targeted for each sample. The scRNA-seq and scTCR-seq libraries were prepared using the 10x Single Cell Immune Profiling Solution Kit, according to the manufacturer's instructions. TotalSeq C (BioLegend, San Diego, USA) and Dextramer (Immudex, Fairfax, USA) signal were captured using and following instructions from 5' V2 Feature Barcoding Kit (1000256, 10x Genomics).

Briefly, amplified cDNA was used for both 5′ gene expression library construction and TCR enrichment. For gene expression library construction, amplified cDNA was fragmented and end-repaired, double-sided size-selected with SPRIselect beads, PCR-amplified with sample indexing primers (98°C for 45 s; 14–16 cycles of 98°C for 20 s, 54°C for 30 s, 72°C for 20 s; 72°C for 1 min), and double-sided size-selected with SPRIselect beads. For TCR library construction, TCR transcripts were enriched from 2 μl of amplified cDNA by PCR (primer sets 1 and 2: 98 °C for 45 s; 10 cycles of 98 °C for 20 s, 67 °C for 30 s, 72 °C for 1 min; 72 °C for 1 min). Following TCR enrichment, enriched PCR product was fragmented and end-repaired, size-selected with SPRIselect beads, PCR-amplified with sample-indexing primers (98 °C for 45 s; 9 cycles of 98 °C for 20 s, 54 °C for 30 s, 72 °C for 20 s; 72 °C for 1 min), and size-selected with SPRIselect beads. Final libraries (GEX, TCR and TSC / Dextramer) were sequenced on Illumina NovaSeq S4 platform (R1 – 26 cycles, i7 – 10 cycles, i5 – 10 cycles, R2 – 90 cycles).

### TCR transduction

293GP retroviral packaging cells were plated in poly-D-lysine-coated 6-well plates (ThermoFisher Scientific, catalog #152035), at 1 million cells per well in 2ml DMEM medium with 10% FCS. Cells were transfected with 2 μg candidate TCR pMSGV1 plasmid and 2 μg RD114 envelope plasmid using Lipofectamine 3000 transfection reagents (ThermoFisher Scientific, catalog #L3000015). Retroviral supernatant was harvested after 48 hours, loaded onto non tissue culture-treated 24-well plates coated with 20 μg/ml retronectin (Takara, catalog #T100B), and centrifuged at 2,300 RPM for 2 hours at 32° C. Primary $CD8^+$ or pan T cells from healthy donors were activated with ImmunoCult CD3/CD28 antibodies (STEMCELL Technologies, catalog #10991) and 300 IU/ml IL-2 for 48 hours. Activated T cells were plated at 250,000 cells per well with 300 IU/ml IL-2 and spun at 1,500 RPM for 10 minutes at 32° C. Cells were assessed for transduction efficiency by staining with mouse TCR antibodies (BioLegend, clone H57-597) at Day 4 post-transduction, and used for co-culture experiments at Day 5.

### Tumor-immune co-culture experiments

#### *Peptide titration experiments*

K562 cells transduced with HLA-A*02:01, CD80, and CD83 were used as artificial antigen-presenting cells (aAPCs) as previously described.[31,32] K562 aAPCs were plated in 96-well U bottom plates (100,000 cells per well), loaded with increasing concentrations of peptides, and incubated at 37° C for 1 hour. 100,000 $CD8^+$ or pan T cells transduced with a candidate TCR were added along with 0.67 and 1 μl/ml protein transport inhibitors containing Monensin and Brefeldin A, and cells were incubated at 37° C for additional 5 hours. Cells were subsequently stained for viability and extracellular markers for 20 minutes, fixed and permeabilized, stained for intracellular cytokines for 30 minutes, and analyzed using BD LSR Fortessa Flow Cytometer.

#### *Overexpression of mis-spliced mRNA isoform*

K562 aAPCs were electroporated with *in vitro* transcribed mRNA of wild-type or mis-spliced isoform of CLK3 or RHOT2 using Neon transfection system 100 μl kit (Invitrogen, catalog #MPK10025) with the following electroporation conditions: pulse voltage of 1450 V, width of 10 ms, and 3 pulses. Electroporated cells were rested in IMDM medium with 10% FCS without antibiotics at 37° C overnight. Subsequently, 100,000 electroporated K562 aAPC cells were co-cultured with 100,000 TCR-T cells in 96-well u-bottom plates along with protein transport inhibitors at 37° C for 5 hours. Cells were subsequently stained for intracellular cytokines and analyzed as described above. For tumor killing assays, 100,000 electroporated luciferase-expressing K562 aAPCs were plated in 96-well flat clear bottom tissue culture-treated plates (Corning, catalog #3610), and TCR-T cells were added across E:T ratios. After 18 hours of incubation at 37° C, Bright-Glo luciferase substrates (Promega, catalog #E2620) were added, and luminescence was measured in a plate reader. Percent lysis was calculated using the formula $100 \times (1 - (RLU_{sample})/(RLU_{target\ alone}))$.

#### *Cytotoxicity against SRSF2-mutant KO52 cells*

KO52 cells were lentivirally transduced with HLA-A*02:01 cDNA as described above. (1) For looking at T cell activation, 100,000 KO52 cells with or without HLA-A*02:01 were co-cultured with TCR-T cells at 1:1 E:T ratio in 96-well u-bottom plates. After 24 hours of incubation at 37° C, cell mixtures were stained with CD69 and 4-1BB antibodies (BioLegend, clones FN50 and 4B4-1, respectively) along with CD8 and mouse TCRβ antibodies (BioLegend, clones RPA-T8 and H57-597, respectively) at 4° C for 40 minutes. Cells

Please cite this article in press as: Kim et al., Mis-splicing-derived neoantigens and cognate TCRs in splicing factor mutant leukemias, Cell (2025), https://doi.org/10.1016/j.cell.2025.03.047



were washed with PBS +2% BSA twice and incubated with CellEvent caspase-3/7 and 7-AAD (Invitrogen, catalog #C10427) at 37° C for 60 and 5 minutes, respectively, as per manufacturer's instructions. Samples were analyzed using BD LSR Fortessa Flow Cytometer. (2) For luciferase-based tumor killing assay, luciferase-expressing leukemia cells were co-cultured with TCR-T cells across E:T ratios in 96-well flat clear bottom plates. After 18 hours of incubation 37° C, luciferase activity was measured to calculate percent lysis as described above. (3) For real-time live cell imaging, 100,000 KO52 cells expressing eGFP-NLS or mCherry-NLS were seeded in 100 μL DMEM media with 10% FCS in 96-well flat-bottom plates. The cell-coated plates were initially incubated at room temperature for 20 minutes before loading into the Incucyte® S3/SX1 G/R instrument (Sartorius) for baseline measurement. After baseline imaging, the plates were removed, T cells were added immediately in 100 μL DMEM media, and the plates were incubated at room temperature for 20 minutes before returning to the Incucyte® instrument for live imaging. For the caspase-3/7-based apoptosis detection, the caspase-3/7 green fluorogenic substrate (Invitrogen, catalog #C10432) was added to the wells together with the T cells as per manufacturer's instructions. Tumor cell killing over time was analyzed by normalizing each tumor cell count to the initial baseline measurement (t0) using the Incucyte® Basic Analysis Software. The time-course detection of tumor cell apoptosis using the caspase-3/7-based eGFP probe was analyzed by normalizing the eGFP fluorescence intensity within mCherry-positive tumor cells to the initial baseline measurement (t0) using the Incucyte® Basic Analysis Software.

#### *Co-culture with primary patient samples*

vials of primary BM MNC or PBMC patient samples were thawed and treated with 0.1 mg/ml DNase I (STEMCELL Technologies, catalog #100-0762) for 15 minutes at room temperature to dissociate cell clumps. 50,000 primary cells were co-cultured with TCR-T cells, which were labeled with CellTrace Blue dye (Invitrogen, catalog #C34568) as per manufacturer's instructions, at 1:1 E:T ratio in 96-well u-bottom plates. After 24-hour incubation at 37° C, cell mixtures were stained with CD69 and 4-1BB antibodies (along with mouse TCRβ antibody) at 4° C for 40 minutes and subsequently analyzed using BD LSR Fortessa Flow Cytometer.

### Alanine scanning

A set of 9-mer peptides, where each of the amino acid residues of the CLK3 neoantigen (RLWGTWVKA) was substituted with alanine (or glycine for the amino acid #9), were synthesized by Genscript. K562 aAPCs were plated in 96-well U bottom plates (100,000 cells per well), loaded with 1 μg/ml of each peptide, and incubated at 37° C for 1 hour. Subsequently, 100,000 $CD8^+$ T cells transduced with each of the CLK3 neoantigen-reactive TCRs were added along with protein transport inhibitors at 37° C for 5 hours. Cells were subsequently stained for intracellular cytokines and analyzed as described above. Once TCR recognition motifs were identified, ScanProsite was used to search all UniProtKB/Swiss-Prot database sequences, including splice variants, for proteins containing these recognition motifs. For TCR2, five potential off-target peptides predicted by ScanProsite were synthesized by Genscript and tested for CLK3 TCR-T cell activation as described above.

### Adoptive cell transfer xenograft model

500,000 HLA-A*02:01-expressing KO52 cells were systemically injected into sublethally irradiated (2.5 Gy) NSG mice via tail veins. KO52 cells were also lentivirally transduced with firefly luciferase to allow bioluminescence (BLI) imaging over time. On days 3 and 10 post-tumor injection (or days 13 and 20), mice were randomized to one of the following treatment groups via tail-vein injection: (1) PBS, (2) CMV-reactive TCR-T cells (10 million cells per injection), and (3) CLK3-reactive TCR-T cells (10 million cells per injection). After first TCR-T cell injection, all mice received a twice-weekly intraperitoneal injection of 1 μg of IL-15 pre-complexed with IL-15Rα-Fc (1:1 M) (pre-complexing done by 30-minute incubation at 37° C) for the duration of the transplant. Tumor burden in mice was measured via weekly BLI imaging using IVISbrite RediJect D-luciferin (Revvity Health Sciences Inc, catalog #50-209-9345).

### HLA-I immunoprecipitation, peptide purification, and mass spectrometry

For experiments using HLA-A*02:01-expressing K562 cells, peptide/HLA-I complexes were isolated as previously described,[26] with a pan HLA class I antibody (BioXCell, clone W6/32, catalog #BE0079) non-covalently linked to GammaBind Plus Sepharose beads (Cytiva, catalog #17088601) co-incubated with soluble lysates overnight. For experiments using HLA-A*02:01-expressing KO52 cells (which also express other HLA class I alleles), HLA-A2 antibody was used (Invitrogen, clone BB7.2, catalog #MA5-16586). After washing with lysis buffer twice, 10mM Tris pH 8 twice, and dH2O twice, the peptides were desalted on C18 tips[67] (Pierce, catalog #87784) and eluted using a 20%–35%-50% acetonitrile stepwise gradient. Eluted fractions were dried using a SpeedVac and stored until mass spectrometry. Peptides were separated using a 12cm packed-in-emitter column and eluted using a gradient increasing from 2%B to 35% B in 70min or 2%B to 60% in 30min (A: 0.1% Formic acid; B: 80% ACN/0.1% Formic acid) delivered at 300nL/min and analyzed using a Fusion Lumos Orbitrap mass spectrometer operated in High/High hybrid DDA/PRM mode or dedicated PRM mode. DDA data was searched against a UniProt human database using PEAKS classic.

## QUANTIFICATION AND STATISTICAL ANALYSIS

### RNA-seq data analysis for differential RNA splicing

RNA-seq were aligned to a genome annotation (hg19/GRCh37) created by combining UCSC knownGene, Ensembl 71, and MISO v2.0[61] annotations. Reads were mapped using RSEM v1.2.4 -v2[58], modified to call Bowtie[54]. Reads that did not align during this step were then mapped using TopHat[60] v2.0.8b to the human genome and to a custom annotation of all possible splice junctions,

which was created using all gene-level pairwise combinations of annotation 5' and 3' splice sites. RSEM was used to generate gene expression estimates and normalized via the trimmed mean of M values (TMM) method. MISO v2.0 was used to quantify isoform expression levels. Events were considered differentially spliced if they contained at least 20 isoform-specific reads, a minimum absolute expression difference of 10%, and a p-value of less than 0.05 calculated using a two-sided Mann-Whitney U test.

### Identification of mis-splicing-derived candidate neopeptides

As described in the text, RNA-seq from five different patient datasets was used to identify potential neoantigens from patients with stereotyped splicing factor mutations in *SRSF2* or *ZRSR2*. The TCGA AML, Leucegene, and Beat AML cohorts were used for *SRSF2* analyses, comparing patients with ("mutant") and without ("wild-type") mutations in *SRSF2*. The Leucegene[19], Beat AML,[17] and two other datasets (Madan et al.[18], Inoue et al.[10]) were used for *ZRSR2* analyses. For each of these, the median splicing change (ΔPSI) and isoform expression was calculated across sample groups and used for further comparisons. The Body Map 2.0 dataset was used in combination with four healthy bone marrow samples from Madan et al. to determine healthy tissue expression and splicing patterns (grouped together as "normal"). Splicing events of interest were chosen to emphasize 1) novelty (lack of expression in normal tissues), 2) absolute expression (the degree of mis-splicing and parent gene expression observed in mutant samples), and 3) consistency (consistent mis-splicing in association with *SRSF2* or *ZRSR2* mutations across multiple datasets). This was accomplished by filtering mis-splicing events to include those that were mis-spliced in multiple datasets with a median ΔPSI of > 0.1 (note that we equivalently refer to this as 10% at times). Further emphasis was placed on absolute expression by calculating an isoform expression metric (psi * gene expression) and requiring a difference of at least 4.5 between the mis-spliced isoform and the wild-type isoform. The median normal tissue expression was required to be <= 0.1 or to have a much lower overall gene expression. In this case, the maximum psi of an isoform of interest in any healthy tissue was required to be <50% of the median PSI of that isoform across tumor samples with the relevant spliceosomal mutation, and the median expression of the parent gene in healthy tissues was required to be <50% of the median expression of the parent gene across tumor samples with the relevant spliceosomal mutation. All possible 8-12-mers were created from mis-spliced isoforms and were considered candidate neoantigens only if they were predicted to bind to HLA-A*02:01 by both NetMHCpan4.0[20] and MHCFlurry2.0[64] (ranks < 2 using both programs) and if the peptide sequences were specific (unique) to the mis-spliced isoform that was favored in patients with mutations in *SRSF2* or *ZRSR2*.

### Dextramer-guided single cell RNA- and TCR-sequencing

Single cell RNA-seq libraries were processed using Cellranger v7.0.1 count workflow to generate gene expression count matrices. Feature-barcoding libraries were processed to generate hashtag oligo (for demultiplexing) and dextramer count matrices using Cite-seq-count v1.4.5.

### Single cell data pre-processing, dimension reduction and clustering

Multimodal single cell data were loaded into Seurat[68,69] data structures and filtered to exclude cells with less than 500 transcriptome reads, less than 250 gene features and a mitochondrial gene ratio of more than 0.1. Data was then log normalized and 4000 variable features were identified. From the variable features, TRA/TRB genes were removed to remove variance associated with T cell clonotype. Data was then scaled and principal component analysis run using Seurat default settings. At this stage, for multi-sample analyses, samples were integrated using Harmony () method. Nearest neighbors, clustering and UMAP dimension reduction were then all performed using "PCA" reduction input for single samples and "harmony" reduction input for integrated samples. Population specific gene expression signatures were identified using a Wilcoxon sum-rank test, with a $\log_2$ fold change threshold of 0.1 and a minimum of 25% expression across all cells in that group. Seurat was used to assign module scores for gene expression signatures across different sub-populations. Significance of module score differences across groups was calculated with t-tests and Bonferroni correction. Gene Set Enrichment Analysis (GSEA) was carried out on ranked lists of genes using the GSEA software provided by the Broad Institute.[22]

### Cell hashing demux and dextramer background thresholding

Samples that had cell hashing were manually demultiplexed by plotting the density distribution of CITE-Seq reads and determining the threshold for positive cell identification through the analysis of the bimodal distribution peaks corresponding to negative and positive populations. Positive 'hashed' cells are dextramer negative and un-hashed cells are dextramer positive. From the dextramer positive population, density plots of dextramer barcode reads (dCODEs), $\log_2$+1 normalized, were plotted for each test dCODE against negative controls, to determine the background threshold for each dCODE.

### TCR clonotype assignment and clonal analysis

From the raw TCR annotated data generated by the CellRanger pipeline, clonotypes were filtered to include those that have a high confidence value (as called by Cell Ranger), have both TRA and TRB chains, include only sequences that are productive, and have a UMI threshold of at least 3. scRepertoire[70] was used to compare clonal diversities across different samples. Clonotypes were manually categorized as "singles" (1 cell), "small" (2 – 20 cells), "medium" (21-100 cells), "large" (101 – 500 cells) or "hyperexpanded" (> 500 cells).

Most plots were generated using ggplot2, heatmaps were generated using the pheatmap package.

Please cite this article in press as: Kim et al., Mis-splicing-derived neoantigens and cognate TCRs in splicing factor mutant leukemias, Cell (2025), https://doi.org/10.1016/j.cell.2025.03.047





### Dextramer / TCR pair identification

A z-scoring method of determining dextramer/TCR pairs that iterates over multiple background thresholds was developed. For each threshold, cells were grouped by clonotype and dCODEs, calculating the count of unique cells per dCODE|clonotype pairing. These counts per clonotype were then normalized, scaling them to per thousand cells to account for differences in clonotype size. The mean and standard deviation of these normalized counts was then computed for each dCODE across all clonotypes. The z-score for each dCODE|clonotype was then calculated by assessing the deviation of its normalized count from the mean, relative to the standard deviation, thus identifying dCODEs that were associated with a TCR clonotype higher than their own specific background pattern. dCODEs/clonotypes with a z-score greater than 1 and a representation in more than 10 cells in that clonotype were considered over-represented. When deciding final dCODE|clonotype hits to pursue, z-scores were prioritized if they were associated with a higher background threshold e.g. a z-score of 2 was weighted more heavily if found in cells above a background threshold of 6 than a threshold of 3.

$$Z = \frac{(X - \mu)}{\sigma}$$

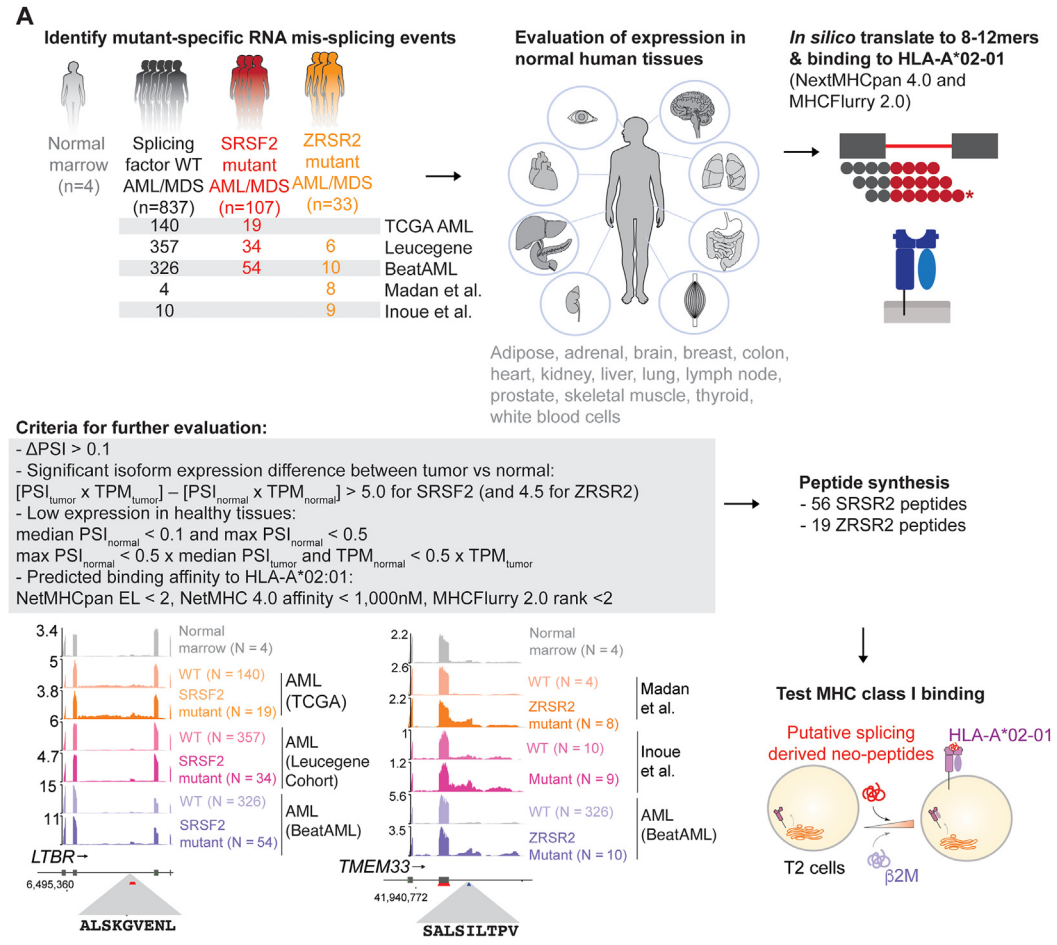
where: $Z$ represents the z-score of a dCODE|clonotype pair, $X$ is the normalized count of a specific dCODE per clonotype, $\mu$ is the mean of the normalized counts for that dCODE across all clonotypes, $\sigma$ is the standard deviation of the normalized counts for that dCODE across all clonotypes.

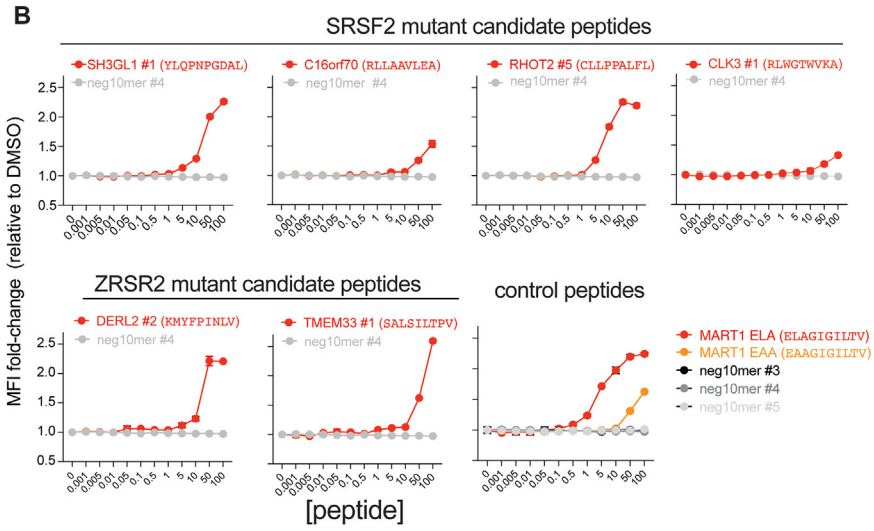
This method helps to overcome inherent challenges with analyzing dCODEs; a) each dCODE has its own background threshold, b) there is a tendency for patterns of co-expression of dCODEs for unknown reasons. Iterative looping over multiple threshold values identifies thresholds for each individual dCODE that appropriately distinguishes between background and positive signals, while z-scoring helps to distinguish specific clonotype binding from non-specific binding patterns. The specificity of this method was quantitatively validated through its performance metrics on positive control dCODEs.

For healthy samples stimulated with peptides, the datasets were analyzed as above but TCR clonotypes of interest were identified primarily by finding clonotypes with dominant expression in the cell hashing negative (i.e. dextramer positive) fraction.

CellPress
OPEN ACCESS

# Supplemental figures




(legend on next page)

**Figure S1. Workflow for prediction and characterization of RNA mis-splicing-derived neoantigens created by mutations in RNA splicing factors, related to Figure 1**

(A) Differential RNA splicing analyses were performed on patient samples from the AML TCGA,[25] Leucegene,[19] BeatAML,[17] Madan et al.,[18] and Inoue et al.[10] datasets of patients with myeloid neoplasms. The number of samples with the indicated mutations in *SRSF2* or *ZRSR2* or splicing factor wild-type (WT) is shown. Additionally, RNA-seq data from four normal adult bone marrows[18] and a collection of 14 normal human tissues from Body Map 2.0 were used. Disease- and *SRSF2*/*ZRSR2* mutation-specific RNA splicing events underwent *in silico* translation and splitting to 8-12-mer peptides, and binding to HLA-A*02:01 was predicted using NetMHCpan 4.0[20] and MHCFlurry 2.0.[64] The criteria noted in the gray box were used to prioritize among the available peptides and select a panel of candidate neoantigens for experimental study. The majority of selected candidates met all of the illustrated criteria. A minority of candidates met many, but not all, of the criteria and were selected based upon their particular biological interest (e.g., peptides arising from an *EZH2* poison exon of known functional importance in *SRSF2*-mutant MDS[7]). Predicted peptides were synthesized and experimentally tested for MHC class I binding in T2 cells. RNA-seq coverage plots illustrate exemplar differential splicing events across the indicated patient cohorts that give rise to candidate neoantigens of interest.

(B) HLA-ABC MFI of T2 cells treated with increasing doses of the indicated test SRSF2 or ZRSR2 mutation-induced derived peptides, or positive control MART1 or negative control 10-mer peptides.

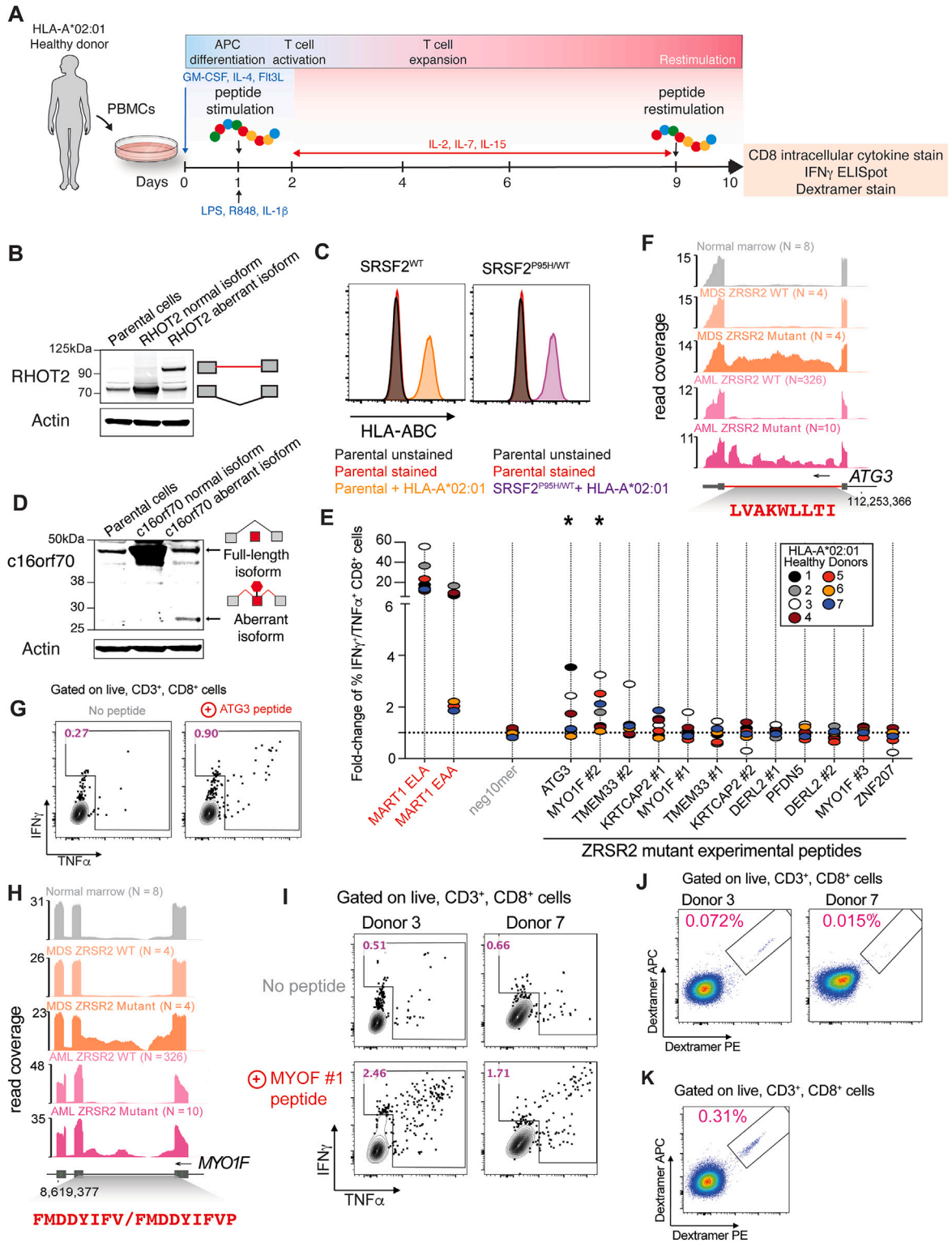



**Figure S2. Immunogenicity of candidate SRSF2 or ZRSR2 mutant-induced neoantigens and dextramer-based identification of neoantigen-reactive CD8+ T cells, related to Figure 2**

(A) Schema of *in vitro* immunogenicity assay to evaluate antigen-reactive T cell responses.[23,24] Healthy donor PBMCs were expanded *in vitro* following stimulation with candidate peptides along with adjuvants. Expanded T cells were then restimulated with the same peptide or DMSO.

*(legend continued on next page)*

(B) Western blot of K562 cells electroporated with *in vitro* transcribed *RHOT2* full-length annotated mRNA (middle column) or intron 5 retained isoform (third column). Left column is parental K562 cells.
(C) FACS histograms of HLA-ABC surface expression on SRSF2 WT or *SRSF2*$^{P95H/WT}$ heterozygous mutant knockin cells transduced with HLA-A*02:01 cDNA. Stained and unstained parental cells without HLA-A*02:01 cDNA shown as controls.
(D) As in (B) but for the *c16orf70* full-length isoform or the poison exon included isoform.
(E) Fold-change in percent of IFN-$\gamma^+$ and/or TNF-$\alpha^+$ CD8$^+$ T cells across healthy donor PBMCs ($n = 7$) upon candidate ZRSR2 peptide restimulation (normalized to DMSO). Red text, MART1 ELAGIGILTV ("ELA") and EAAGIGILTV ("EAA") positive control peptides. Neg10mer negative control peptide is a predicted nonbinder to HLA-A*02:01. $p$ values, Wilcoxon signed-rank test (*$p$.adj $\leq$ 0.05).
(F) RNA-seq coverage plots of a minor intron inclusion event in *ATG3* induced by ZRSR2 mutations. From top to bottom, data are from normal bone marrow, MDS patients from Madan et al.,[18] and MDS patients from Inoue et al.[10] $n$, number of patients per genotype.
(G) Representative ICS FACS plots of data from (E) demonstrating IFN-$\gamma^+$ and/or TNF$\alpha^+$ CD8$^+$ T cells restimulated with ATG3 peptide or DMSO.
(H) As in (F) but for a minor intron in *MYO1F*. This intron retention event is predicted to give rise to two distinct candidate neoepitopes with the amino acid sequences shown.
(I) As in (G) but for the MYO1F peptide (FMDDYIFV).
(J) FACS plots of dual-color ATG3 peptide-HLA-A*02:01 dextramer staining of CD8$^+$ T cells following priming with ATG3 peptide.
(K) As in (J) but using MYO1F peptide-HLA-A*02:01 dextramers.

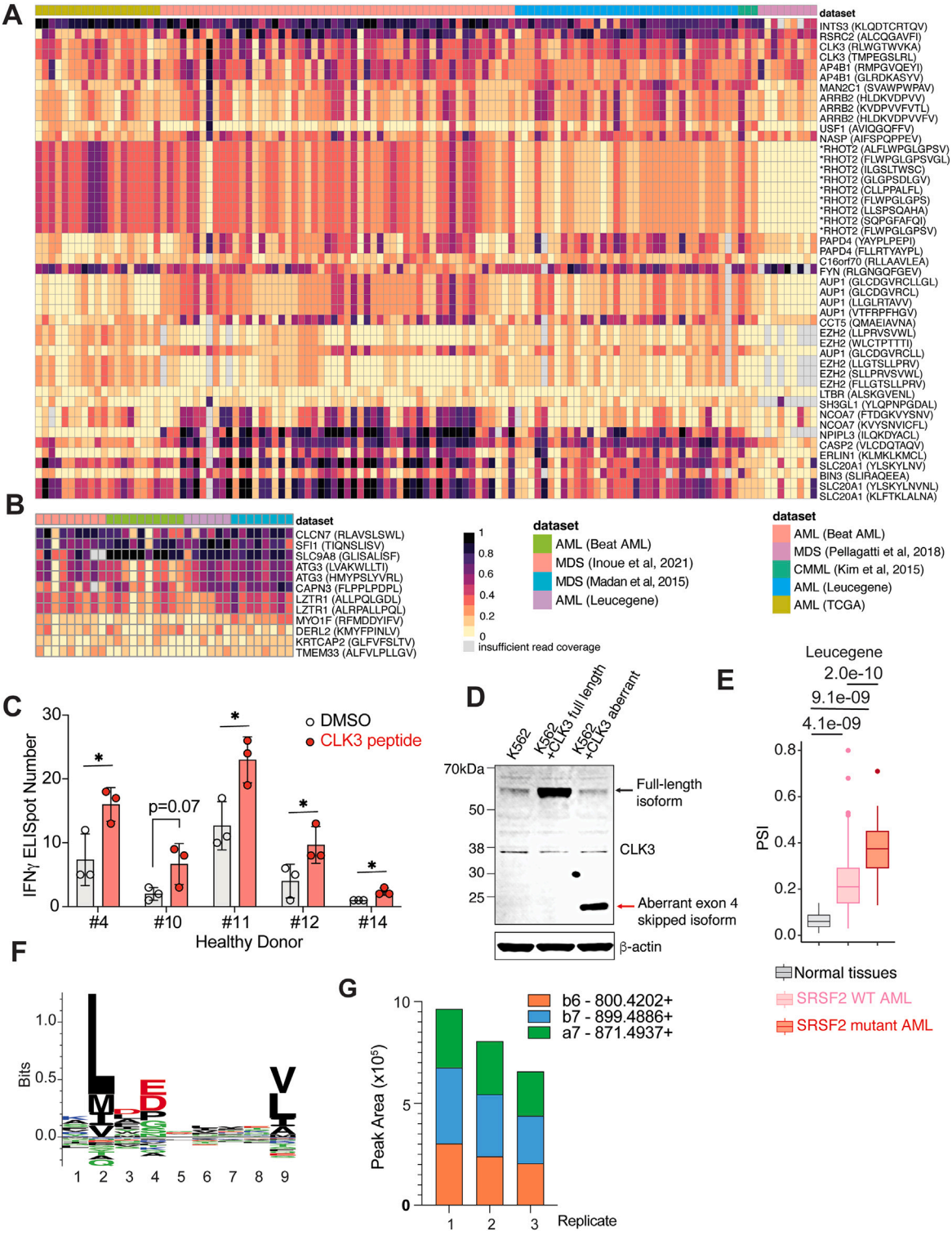



**Figure S3. Functional characterization of the CLK3 neoantigen, related to Figure 3**

(A) Heatmap illustrating the PSI values for SRSF2 mutant-induced mis-spliced isoforms encoding putative splicing-derived neoantigens across primary patient samples from the indicated cohorts. Each row represents a putative neoantigen that we experimentally studied via dextramers, indicated by the peptide sequence and the name of the mis-spliced gene. Each column is an individual patient with a hotspot SRSF2 mutation, categorized by dataset and disease type. "Insufficient read coverage" indicates samples with fewer than 5 informative RNA-seq reads, where an informative read distinguishes between different isoforms

*(legend continued on next page)*

of the splicing event of interest. Note that RNA-seq read coverage in the MDS dataset is relatively low coverage, which reduces the accuracy of quantification of isoform expression. Therefore, the RNA-seq data from MDS samples should be interpreted as providing evidence of some level of neoantigen-encoding isoform expression rather than viewed as accurate quantification of isoform expression levels.

(B) As in (A), but for patients with loss-of-function ZRSR2 mutations and putative ZRSR2 mutant-induced neoantigens that we studied via dextramers.

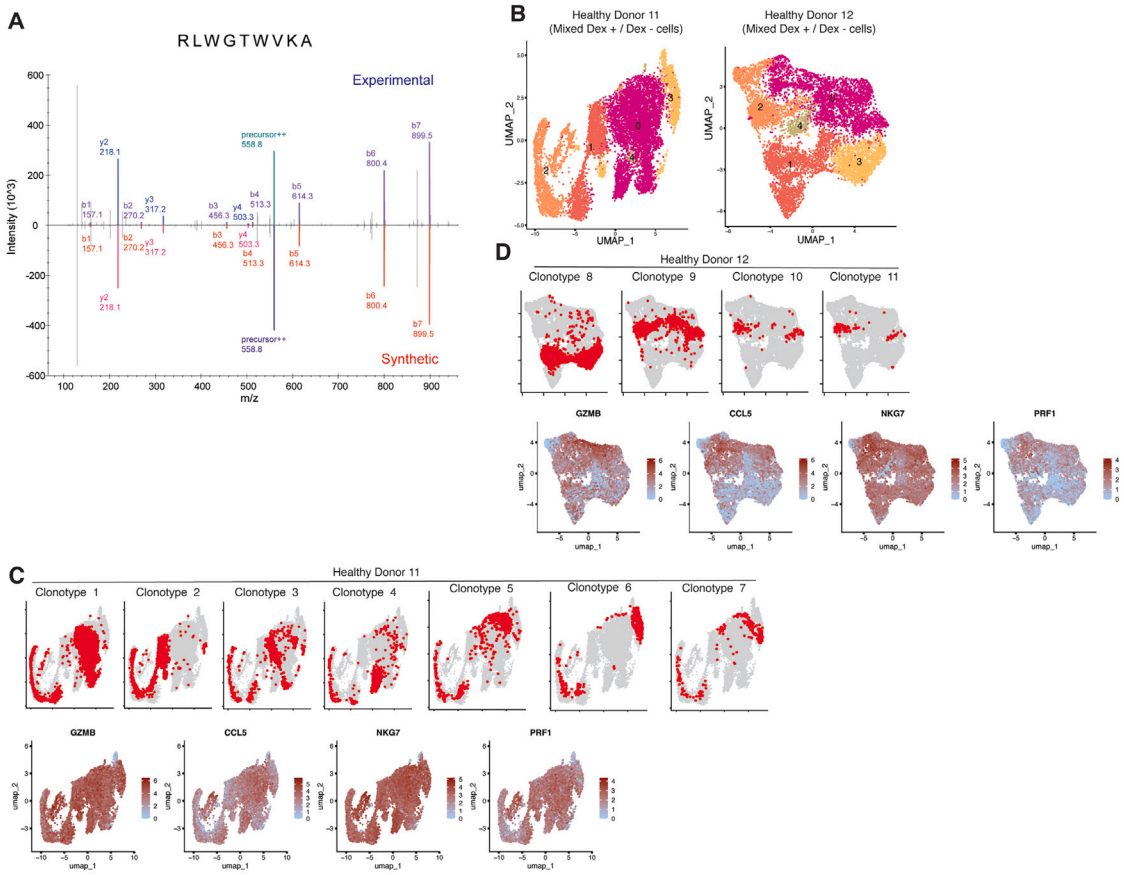
(C) IFN-$\gamma$ ELISpot spot numbers from healthy donor PBMCs *in vitro* primed and restimulated with the CLK3 peptide or DMSO. $n = 3$ technical replicates. Mean ± SEM. $p$ values, unpaired Student's t test (*$p$.adj $\leq 0.05$).

(D) Western blot of K562 cells electroporated with *in vitro* transcribed *CLK3* full-length annotated mRNA (middle column) or *CLK3* exon 4 skipped isoform (third column). Left column is parental K562 cells.

(E) Boxplot quantifications of PSI of the *CLK3* exon 4 skipping event in the Leucegene[19] cohort. Normal tissue controls are from Body Map 2.0. Horizontal line, median; colored box, interquartile range; whiskers, minimum and maximum values within 1.5× of interquartile range; dots outside whiskers, outliers. $p$ values, one-sided Mann-Whitney U test.

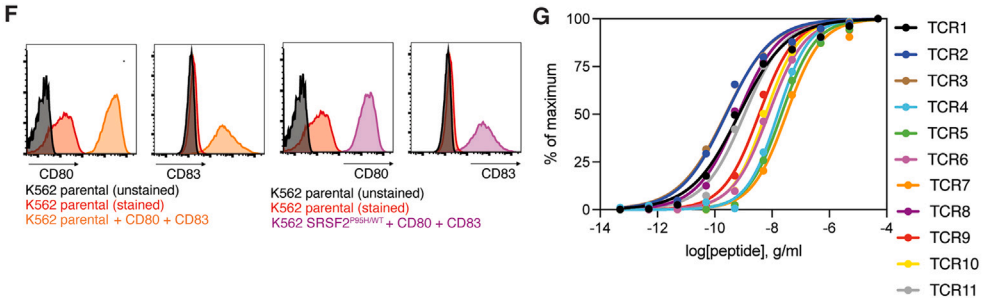
(F) Sequence logo for 9-mers identified from HLA-IP LC-MS/MS experiments using HLA-A*02:01$^{+}$ SRSF2$^{P95H/WT}$ K562 cells.

(G) Skyline analysis of HLA-I-bound peptides from the HLA-IP LC-MS/MS experiments from (F). Peak area intensity of fragment ions derived from the CLK3 peptide shown.




E

| TCR ID | TRAV | TRAJ | TRBV | TRBJ | TRBD | CDR3α/CDR3β | CDR3α length | CDR3β length |
|---|---|---|---|---|---|---|---|---|
| TCR 1 | TRAV26-2 | TRAJ39 | TRBV30 | TRBJ2-2 | N/A | CILRDPDAGNMLTF / CAWNWVNTGELFF | 14 | 13 |
| TCR 2 | TRAV38-2/ DV8 | TRAJ49 | TRBV30 | TRBJ2-3 | N/A | CASGAGNQFYF / CAWSFTSSTDTQYF | 11 | 14 |
| TCR 3 | TRAV19 | TRAJ36 | TRBV30 | TRBJ2-3 | TRBD1 | CALSEAGTFGANNLFF / CAWSVTGSTDTQYF | 16 | 14 |
| TCR 4 | TRAV29/ DV5 | TRAJ53 | TRBV13 | TRBJ2-1 | N/A | CAASDSGGSNYKLTF / CASSLVGGVLSYNEQFF | 15 | 17 |
| TCR 5 | TRAV19 | TRAJ26 | TRBV30 | TRBJ2-3 | N/A | CALSEADYGQNFVF / CAWSLAGGSDTQYF | 14 | 14 |
| TCR 6 | TRAV12-2 | TRAJ43 | TRBV30 | TRBJ1-1 | N/A | CAVFSGGDMRF / CAWNQQGGTEAFF | 11 | 13 |
| TCR 7 | TRAV12-3 | TRAJ49 | TRBV30 | TRBJ2-6 | N/A | CAMGENTGNQFYF / CASDSLSGANVLTF | 13 | 14 |
| TCR 8 | TRAV39 | TRAJ41 | TRBV30 | TRBJ2-3 | N/A | CAVDHSNSGYALNF / CAWSFLNTDTQYF | 14 | 13 |
| TCR 9 | TRAV13-1 | TRAJ26 | TRBV30 | TRBJ1-5 | TRBD1 | CAASYGQNFVF / CAWSSGAVGNQPQHF | 11 | 15 |
| TCR 10 | TRAV38-2/ DV8 | TRAJ43 | TRBV5-8 | TRBJ2-3 | N/A | CAYNNNNDMRF / CASSTGIASTDTQYF | 11 | 15 |
| TCR 11 | TRAV5 | TRAJ29 | TRBV30 | TRBJ2-3 | N/A | CAEGNTPLVF / CAVGGQKLGDTQYF | 10 | 14 |




*(legend on next page)*

**Figure S4. Identification of TCRs reactive to the CLK3 neoantigen, related to Figure 4**

(A) Mirror plot displaying the MS2 spectra of the CLK3 peptide (RLWGTWVKA) eluted from HLA-A*02:01$^{+}$ KO52 cells (top) and synthesized (bottom).

(B) Gene-expression UMAPs of the CLK3 peptide-primed CD8$^{+}$ T cells from healthy donors 11 (left) and 12 (right).

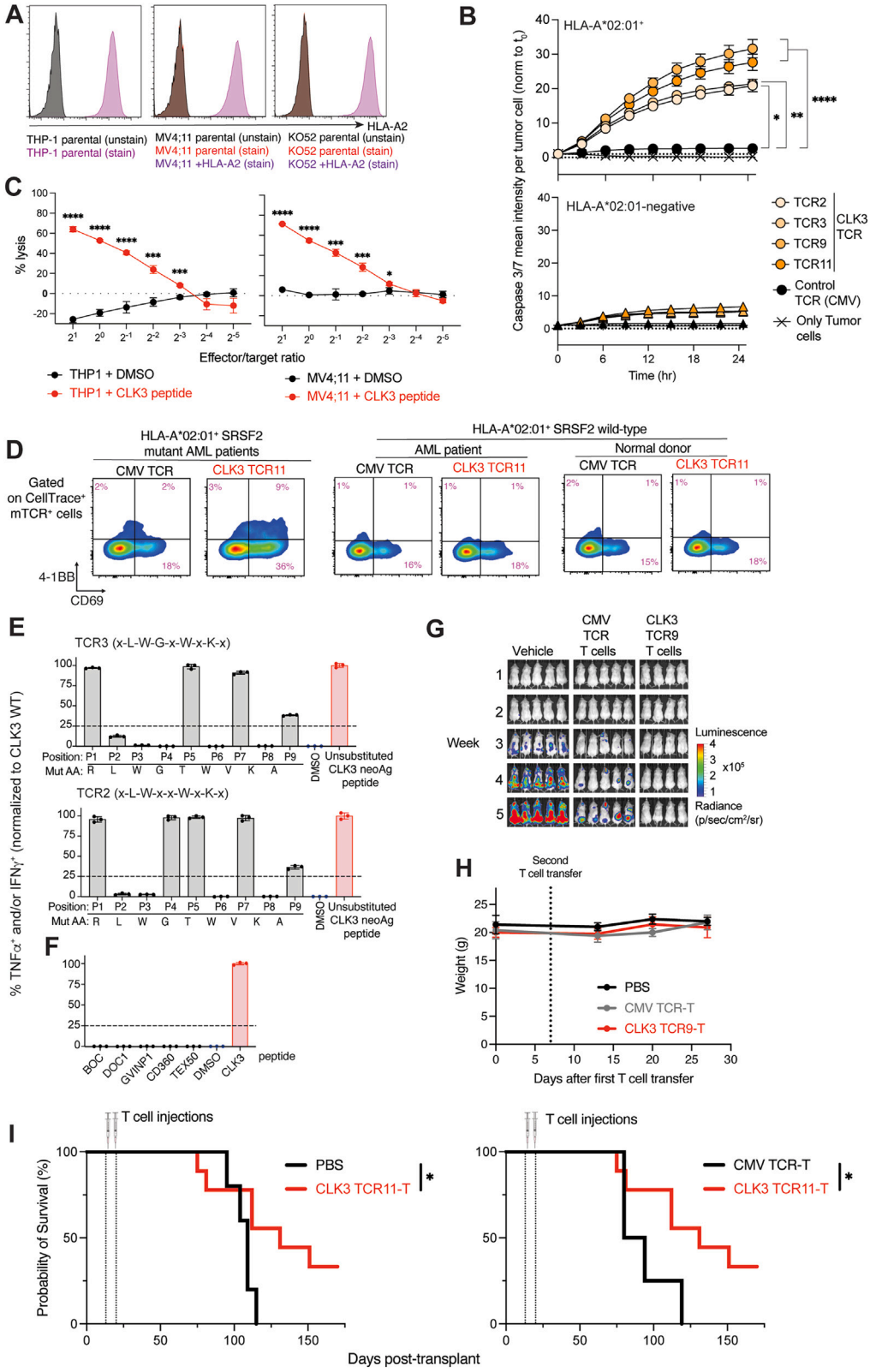
(C) UMAP mapping of TCR clonotypes 1–7 from healthy donor 11 (top) and cytotoxic marker expression in the CD8$^{+}$ T cells from (B) (bottom).

(D) As in (C) but for healthy donor 12 showing four TCR clonotypes.

(E) Table listing *TRAV*, *TRAJ*, *TRBV*, and *TRBJ* gene segment composition, CDR3 sequences, and CDR3 lengths for a panel of candidate TCRs reactive to the CLK3 neoantigen.

(F) FACS histograms of CD80 and CD83 surface expression on HLA-A*02:01$^{+}$ K562 parental or SRSF2$^{P95H/WT}$-mutant aAPCs. Stained and unstained parental cells without CD80/CD83 cDNA expression are shown as controls.

(G) Dose response curve of T cell activation across CLK3 peptide concentrations for 11 CLK3 neoantigen-reactive TCRs identified from (E).




*(legend on next page)*

**Figure S5. Anti-tumor efficacy CLK3 neoantigen-reactive TCR-T cells against SRSF2 mutant AML cells, related to Figure 5**

(A) FACS histograms of HLA-A2 surface expression in SRSF2 mutant KO52 and SRSF2 wild-type THP-1 (endogenously expresses HLA-A*02:01) and MV4;11 (transduced with HLA-A*02:01 cDNA).

(B) Time course detection of KO52 cell apoptosis using the caspase 3/7-based eGFP probe via live cell imaging. KO52 parental cells were co-cultured with $CD8^+$ T cells transduced with CLK3 neoantigen-reactive TCR clonotypes 2, 3, 9, or 11 (or CMV-reactive TCR) at 1:1 E:T ratio over time. Mean ± SEM. *p* values, ordinary one-way ANOVA (*$p$.adj $\leq$ 0.05, **$p$.adj $\leq$ 0.01, ****$p$.adj $\leq$ 0.0001).

(C) Percent lysis of HLA-$A2^+$ SRSF2 wild-type leukemia cells loaded with or without 1 μg/mL CLK3 peptide and co-cultured with $CD8^+$ CLK3 TCR9-T cells across E:T ratios. Mean ± SEM. *p* values, unpaired Student's t test (*$p$.adj $\leq$ 0.05, ***$p$.adj $\leq$ 0.001, ****$p$.adj $\leq$ 0.0001).

(D) FACS plots of CD69 and 4-1BB surface expression in CLK3 or CMV TCR-T cells, upon co-culture with BM MNCs or PBMCs from HLA-A*02:$01^+$ SRSF2 wild-type or mutant AML patients or a normal donor. FACS plots are gated on $CellTrace^+$, live, mouse $TCR^+$ T cells.

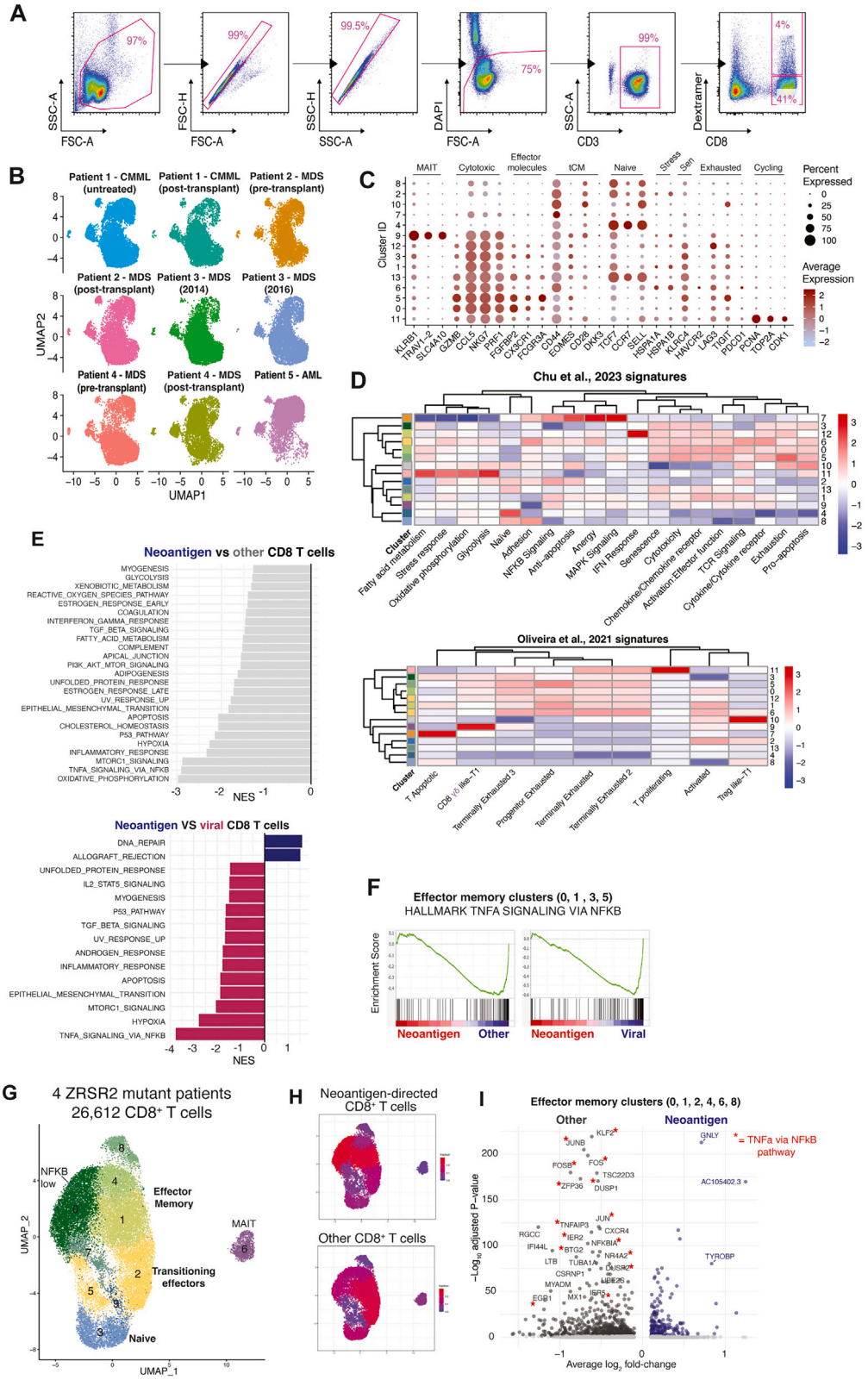
(E) Identification of peptide recognition motif for CLK3 TCR3 (top) and TCR2 (bottom) using alanine scanning. Frequency of TNF-$\alpha^+$ and/or IFN-$\gamma^+$ TCR-T cells upon 1:1 E:T co-culture with peptide-loaded K562 cells (1 μg/mL). % maximum response relative to unsubstituted CLK3 peptide (y axis). $n = 3$ technical replicates.

(F) Bar graph enumerating the frequency of TNF-$\alpha^+$ and/or IFN-$\gamma^+$ CLK3 TCR2-T cells upon co-culture with HLA-A*02:$01^+$ K562 cells loaded with potential off-target peptides derived from the proteins indicated on the x-axis. DMSO and native CLK3 peptide used as controls.

(G) Representative BLI images of animals treated with CLK3 TCR9-T cells or either of the controls as described in Figure 5G.

(H) Weight of tumor-bearing mice after two doses of $CD8^+$ CLK3 TCR9-T cells or either of the controls.

(I) Kaplan-Meier curves of the mice treated with CLK3 TCR11-T cells or either PBS (left) or CMV TCR-T cells (right). *p* values, log-rank test (*$p$.adj $\leq$ 0.05).




*(legend on next page)*

**Figure S6. Gene-expression analyses of neoantigen-reactive CD8$^+$ T cells from myeloid leukemia patients with mutations in SRSF2 or ZRSR2, related to Figure 6**

(A) FACS gating strategy for sorting dextramer-positive and -negative CD3$^+$CD8$^+$ cells from patient PBMCs. When included, dextramer-negative cells were stained with hashtag oligonucleotides before single-cell sequencing.

(B) UMAP from Figure 6B separated by patient sample.

(C) Bubble plots of marker gene expression used to define each CD8$^+$ T cell cluster in the gene-expression UMAP.

(D) Heatmap of gene signature scores across UMAP clusters in Figure 6B.

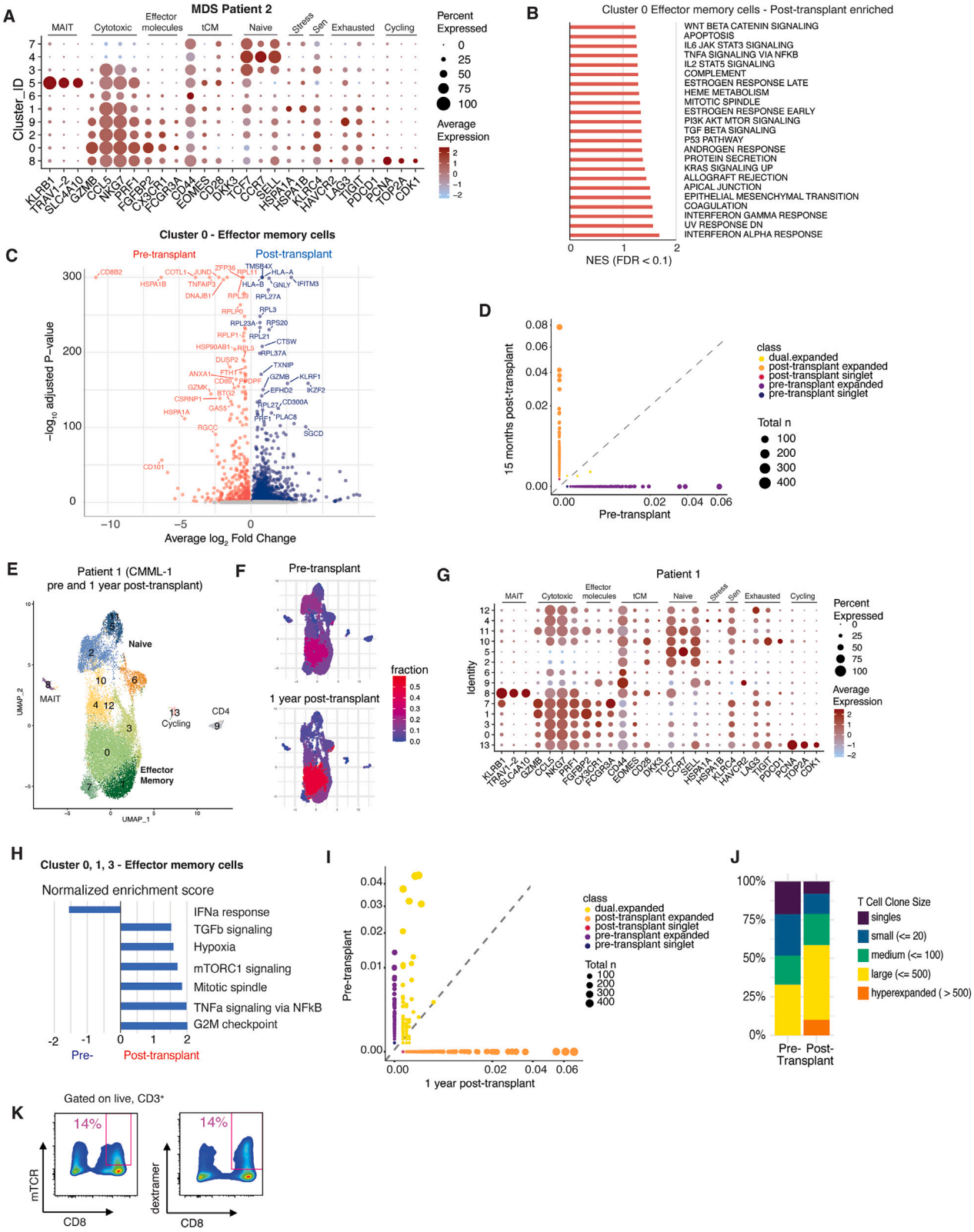
(E) Normalized enrichment score (NES) of hallmark gene sets differentially expressed in neoantigen-reactive CD8$^+$ T cells versus those cells without known antigen specificity (top) and CD8$^+$ T cells directed against CMV peptide (bottom) (FDR < 0.1).

(F) Gene set enrichment analysis of hallmark TNF-$\alpha$ signaling via NF-$\kappa$B pathway genes in neoantigen-reactive CD8$^+$ T cells versus those cells without known antigen specificity (left) and CD8$^+$ T cells reactive to CMV peptide (right).

(G) UMAP of CD8$^+$ T cells from four distinct HLA-A*02:01$^+$ ZRSR2 mutant myeloid leukemia patients. Clusters are denoted by colors and labeled with inferred cell states.

(H) As in (G) but indicating distribution of neoantigen-directed (top) and other (bottom) CD8$^+$ T cells.

(I) Volcano plot of differentially expressed genes in cytotoxic CD8$^+$ T cells in clusters 0, 1, 2, 4, 6, and 8 from neoantigen-directed versus other CD8$^+$ T cells from (G).




*(legend on next page)*

**Figure S7. Neoantigen-reactive T cells and TCRs pre and post HLA-matched allogeneic transplantation in splicing factor mutant leukemia patients, related to Figure 7**

(A) Bubble plots of marker gene expression used to define each $CD8^+$ T cell cluster in the gene-expression UMAP for patient 2.
(B) NES of differentially enriched pathways in pre- versus post-transplant effector memory $CD8^+$ T (tEff) cells from clusters 0, 1, and 3 for Patient 2.
(C) Volcano plot of differentially expressed genes in effector memory $CD8^+$ T cells in cluster 0 pre- versus post-transplant for patient 2.
(D) Comparison of TCR clonotype frequency in pre- versus post-transplant $CD8^+$ T cells for patient 2.
(E) UMAP of $CD8^+$ T cells pre- and post- transplant for patient 1 (see Table S4). Clusters are denoted by colors and labeled with inferred cell states.
(F) As in (E) but indicating fraction of total cells within each cluster for pre- and post-transplant samples.
(G) Bubble plots of marker gene expression used to define each $CD8^+$ T cell cluster in the gene-expression UMAP for patient 1.
(H) NES of differentially enriched pathways in pre- versus post-transplant effector memory $CD8^+$ T cells from clusters 0, 1, and 3 for patient 1.
(I) As in (D) but for patient 1.
(J) Bar plots for the frequency of T cell clone size in $CD8^+$ T cells from each time point in patient 1.
(K) FACS plots of primary human $CD8^+$ T cells transduced with putative RHOT2 neoantigen-reactive TCR and stained for mouse TCR constant region (mTCR) (left) or RHOT2 peptide #5-HLA-A*02:01 dextramers (right).