T cell antigen discovery using signaling and antigen presenting bifunctional receptors (SABRs)

Alok Joglekar (✉ alokvj@caltech.edu)
  Joglekar Lab

Michael Leonard
John Jeppson
Margaret Swift
David Baltimore

Abstract

CD8+ T cells recognize and eliminate tumors in an antigen-specific manner. Despite progress in characterizing the antitumor T cell repertoire and function, identifying their target antigens remains a challenge. Here, we describe the use of chimeric receptors called Signaling and Antigen-presenting Bifunctional Receptors \(SABRs) in a novel cell-based platform for T Cell Receptor \(TCR) antigen discovery. SABRs present an extracellular peptide-MHC complex and induce intracellular signaling via a TCR-like signal upon binding with a cognate TCR. We devised a strategy for antigen discovery using SABR libraries to screen thousands of antigenic epitopes. We validated this platform by identifying the targets recognized by public TCRs of known specificities. Moreover, we extended this approach for personalized neoantigen discovery. The antigen discovery platform reported here will provide a scalable and versatile way to develop novel targets for immunotherapy.

Introduction

Here, we sought to develop novel antigen discovery techniques to address the unmet need. We present a cell-based platform for T cell antigen discovery. In this study, we describe chimeric receptors called Signaling and Antigen-presenting Bifunctional Receptors \(SABRs). By virtue of genetically linking the peptide epitope with MHC, SABRs can be used to present a defined antigen and to report its successful recognition by a TCR. Therefore, we asked if SABR libraries presenting a large number of epitopes can be used to screen successful TCR-pMHC interactions. An orphan TCR may be obtained by single cell sequencing on tumor infiltrating lymphocytes isolated from a patient sample. We designed a strategy to construct and use SABR libraries for T cell antigen discovery of 'orphan' antitumor TCRs with unknown antigens. First, a list of target epitopes was generated from an existing database or from tumor exome data followed by prediction of MHC binding. The protein sequences of the target epitopes were backtranslated to generate oligonucleotide sequences. Overhangs of sequences with 15bp overlap with the SABR vector were added to the oligonucleotide sequences corresponding to the epitopes. The entire list of oligonucleotides for the library was synthesized using pooled synthesis. The pooled library was then amplified and cloned using ligation-free cloning using the 15bp overhangs into the SABR vector plasmid. The SABR libraries were packaged into lentiviral vectors and used to transduce NFAT-GFP-Jurkat cells. NFAT-GFP-Jurkat cells expressing the SABR library were co-cultured with Jurkat cells expressing an 'orphan' TCR. GFP+CD69+ NFAT-GFP-Jurkat cells were sorted using fluorescence activated cell sorting \(FACS), followed by genomic DNA extraction. The epitope portion of the SABRs was amplified and subjected to high throughput sequencing. The sequencing reads were aligned with the SABR vector backbone using Burrows-Wheeler alignment. Aligned reads were translated to reveal the epitope. The number of reads corresponding to each epitope was counted and reported in a list. A minimum of three replicates of the co-incubation assay were performed. For each replicate, a numerical rank was given to each epitope based on descending order of the number of reads. The rank from three replicates for each assays was averaged and reported as 'Average Rank'. The top ranked epitopes were putative antigens for

that TCR and are subsequently validated by constructing individual SABRs presenting each of the epitopes and measuring GFP expression in co-culture assays.

Reagents

Jurkat Cells, Clone E6-1 ATCC, Manassas, VA TIB¬-152 NFAT-GFP-Jurkat Cells Provided by Arthur Weiss and Yvonne Chen N/A K562-A2.1+ cells ATCC, Manassas, VA CCL-243 GXR-B27+ cells Provided by Bruce D. Walker N/A Primary T Cells UCLA, CFAR Virology Core N/A HEK-293T Cells ATCC, Manassas, VA CRL-3216 RPMI 1640, 1X with L-glutamine Corning™, Corning, NY 10-040-CV 10% Fetal Bovine Serum Corning™, Corning, NY 35-015-CV Penicillin-Streptomycin Solution, 100X Corning™, Corning, NY 30-002-CI G-418 Sulfate Corning™, Corning, NY 30-234-CI Immunocult™ CD3/28 T Cell Activator StemCell Technologies™, Vancouver, Canada 10991 Human IL-2 IS, premium grade MACS Miltenyi Biotec, Bergisch Gladbach, Germany 130-097-746 All indicated peptides Synthesized by Pierce Thermo Fisher N/A DMEM, 1X with L-Glutamine, 4.5g/L Glucose and Sodium Pyruvate Corning™, Corning, NY 10-013-CV Clontech In-Fusion® HD Cloning Kit Clontech, Takara Bio USA, Mountain View, CA 639650 BsmBI New England BioLabs®, Inc, Ipswich, MA R0580S Clontech Stellar™ Competent Cells Clontech, Takara Bio USA, Mountain View, CA 636763 Zyppy™ Plasmid Miniprep Kit Zymo Research, Irvine, CA D4036 Carbenicillin \(disodium) Gold Biotechnology®, Saint Louis, MO C-103-SL10 NucleoBond® Xtra Maxi EF Kit Clontech, Takara Bio USA, Mountain View, CA 740424.50 TransIT®-293 Transfection Reagent Mirus® Bio, LLC Madison, WI MIR 2704 RetroNectin Takara Bio Inc, CA T100B Gibco™ Opti-MEM™ I Reduced Serum Medium Life Technologies, Thermo Fisher Scientific, Waltham, MA 31985-062 Millex®-HV Syringe Filter Unit, 0.45 μm, PVDF, 33 mm, gamma sterilized EMD Millipore, Burlington, MA SLHV033RS MACSQuant® Analyzer 10 MACS Miltenyi Biotec, Bergisch Gladbach, Germany 130-096-343 APC/Cy7 anti-human CD69 Clone: FN50 BioLegend®, San Diego, CA 310913 BD FACSort™ Becton Dickinson, Franklin Lakes, NJ CFSE Cell Division Tracker Kit BioLegend®, San Diego, CA 423801 Invitrogen™ PureLink™ Genomic DNA Mini Kit Invitrogen™ Life Technologies, Thermo Fisher Scientific, Waltham, MA K182001 KOD DNA Polymerase EMD Millipore, Burlington, MA 71085 Macherey-Nagel NucleoSpin® Gel and PCR Purification Kit Clontech, Takara Bio USA, Mountain View, CA 740609.250 TreeStar FlowJo® Flow Cytometric Data Analysis Software v10 FlowJo, LLC Ashland, Oregon N/A GraphPad Prism v7 GraphPad, San Diego, California N/A The plasmids for HLA-A**0201-SABR backbone \(pCCLc-MND-A0201-SABR-Backbone, ID 119050), HLA-B**2705-SABR backbone \(pCCLc-MND-B2705-SABR-Backbone, ID 119051), A2-Mart1-SABR \(pCCLc-MND-A0201-Mart1-SABR, ID 119052), and B27-KK10-SABR \(pCCLc-MND-B2705-KK10-SABR, ID 119053) are available through Addgene Inc.

Equipment

2100 Bioanalyzer Instrument Agilent, Santa Clara, CA G2939BA Ilumina® HiSeq 2500 Sequencing System Illumina® San Diego, CA SY−401−2501 BD FACSort™ Becton Dickinson, Franklin Lakes, NJ MACSQuant® Analyzer 10 MACS Miltenyi Biotec,

Procedure

Generate pooled oligonucleotide library \(Timing: 30 min) 1. Starting with a list of desired peptides to express in a SABR library, generate back-translated nucleotide sequences with the IDT Codon Optimization Tool. Select "Amino Acids" as the input sequence type, "gBlocks Gene Fragments" as the Product Type, "Homo sapiens \(human)" as the Organism, and enter a tab-delimited list of peptides on the "Bulk Entry" page. 2. Append nucleotide sequences for cloning into a BsmBI digested backbone to the back-translated epitope sequences "CAGGAGGGCTCGGCA" + + "GGATGCGGAGGGTCC" 3. Append random nucleotides to the end of the epitope oligo for normalization + "GGCGGCCGCTCATCTCGTGCGACATCAAGCTATACTCTAATATAGCATTCCGTTCGAGTATAGCAG" 4. Submit the first 100 bp of each nucleotide sequence with appended random nucleotides for synthesis =Left\(, 100). Pause point: The synthesized pooled library can be stored at -20°C indefinitely **Cloning Oligonucleotides in SABR vectors** The plasmids for HLA-A**0201-SABR backbone \(pCCLc-MND-A0201-SABR-Backbone, ID 119050), HLA-B**2705-SABR backbone \(pCCLc-MND-B2705-SABR-Backbone, ID 119051), A2-Mart1-SABR \(pCCLc-MND-A0201-Mart1-SABR, ID 119052), and B27-KK10-SABR \(pCCLc-MND-B2705-KK10-SABR, ID 119053) are available through Addgene Inc. A schematic of the cloning procedure is shown in Fig 2 **Day 1 \(Timing: 1-2 hours)** Digest backbone plasmid with restriction enzyme specific to overhangs selected above 5. First backbone digest: Prepare 1200 ul of digestion buffer \(120 ul of 10X NEB 3.1 Buffer, 48 ul of NEB BsmBI Enzyme, 72 ul of Addgene backbone plasmid at 1 ug/ul, 960 ul of Nuclease-free water), aliquot 100 ul across 12 PCR strips, and incubate for 6 hours at 55°C, then heat inactivate enzyme at 80°C for 20 minutes. Pause point: Store at 4°C for up to one month or at -20°C indefinitely 6. Pool all digestion reactions in a 1.5 mL eppendorf tube. Purify with NucleoSpin Gel and PCR purification kit, using one column per 240 ul of digested plasmid \(5 in total), and elute in 20 ul Elution Buffer \(supplied) per column. Pool elutions in a 1.5 mL eppendorf tube. About 50% recovery of digested plasmid expected \(36 ug) in 100 ul is about 300 ng/ul final concentration. 7. Second backbone digest: Prepare 600 ul of second digestion buffer \(60 ul of 10X NEB 3.1 Buffer, 24 ul of NEB BsmBI Enzyme, 100 ul of digested Addgene backbone from Step 2, 416 ul of Nuclease-free water), aliquot 100 ul across 6 PCR strips, and incubate for 6 hours at 55°C, then heat inactivate enzyme at 80°C for 20 minutes. Pause point: Store at 4°C for up to one month or at -20°C indefinitely **Day 2 \(Timing: 3-4 hours)** Prepare insert and vector, ligate, and transform 8. Gel purify backbone: Pour a 0.9% agarose gel \(mix 1.08 g agarose and 120 mL TAE buffer, heat to above 95°C until fully dissolved, cool to 60°C, add 10 ul gel green, mix and cast in a gel tray with wide combs until solidified). Pool all second digest reactions in a 1.5 mL eppendorf tube \(600 ul). Add 120 ul 6X Purple Loading Dye to digest, mix, and load mixture on gel at 65 ul per well \(11 total). Load 10 ul 2-log Ladder in a remaining empty well. Run gel at 120 V for 1 hour, or until dye front has run 75% of the gel. Briefly image the gel in a gel imager at 365 nm UV wavelength. Cut out the larger of the two expected bands with a razor, and place each excised band in a 2 mL eppendorf tube. Follow instructions for NucleoSpin Gel and PCR purification kit, using one column per 4 wells of excised plasmid \(3 in total), and elute in 20 ul Elution Buffer \(supplied) per column. Pool elutions in a 1.5 mL eppendorf tube. About 30% recovery of digested plasmid expected \(11 ug) in 60 ul is about 180 ng/ul final concentration. Note: The smaller band corresponds to the 2kb stuffer sequence,

expected to run at 2,000 bases. The larger band is the digested plasmid. Critical Step: Do not use 302nm wavelength for visualizing the gel. Caution: If using Ethidium bromide to visualize gels: Ethidium bromide is a carcinogen. Dispose of it appropriately. 9. Amplify library oligos: Dilute or resuspend single-stranded oligo pool to 10 ng/ul with Elution Buffer \(supplied in the PCR purification kit). Prepare 600 ul library amplification mix \(300 ul of 2X KOD master mix, 24 ul of 10 uM Oligo-Amplify-Fwd, 24 ul of 10 uM Oligo-Amplify-Rev, 2.4 ul of 10 ng/ul single-stranded oligo pool, and 250 ul of Nuclease-free water), aliquot mix to 24 wells of a PCR plate or strips, and amplify according to the following program: Initial Denaturation at 95°C for 2 minutes, 6 cycles of 95°C for 20 seconds \(denature), 61°C for 10 seconds \(anneal), and 70°C for 15 seconds \(extension), then a final extension at 70°C for 1 minute, and store at 4°C for up to a month or at -20°C indefinitely. Troubleshooting: If the oligonucleotides do not amplify, perform the reaction using a gradient of annealing temperatures from 50-70°C. Cycle number can be increased to 10-15 at the optimal temperature. Verify on gel to make sure alternate products do not exist, which can occur for high cycle numbers after primers are exhausted. 10. Pool all amplification reactions in a 1.5 mL Eppendorf tube. PCR purify with NucleoSpin Gel and PCR purification kit, using one column, and elute in 17 ul Elution Buffer \(supplied). Expect above 10 ng/ul of double-stranded oligos. 11. Ligation: Dilute linearized vector to 100 ng/ul. Dilute double-stranded oligo pool to 12.5 ng/ul. In PCR strips, prepare 10 ul of ligation mix \(2 ul of 5X In-Fusion mix, 2 ul of linearized vector at 100 ng/ul, 2 ul of double-stranded oligo pool at 12.5 ng/ul, 4 ul of Nuclease-free water) and 10 ul of no-insert mix \(2 ul of linearized vector and 8 ul of Nuclease-free water). Incubate at 50°C for 15 minutes. Pause point: Store mixes at 4°C for up to one month or at -20°C indefinitely. 12. Dilute both ligation mix and no-insert mix to 24 ul with Nuclease-free water. Using 12 tubes \(50 ul each) of NEB 5-alpha high efficiency competent cells, transform 2 ul of ligation mix per tube, according to manufacturer instructions. Pool cells diluted in SOC after recovery. Perform four serial dilutions on an aliquot of pooled cells and plate equal volumes of each dilution on LB + Carbenicillin plates to calculate transformation efficiency. Inoculate 1.5 L of LB + Carbenicillin liquid media with the remainder of the cell mixture, split across 3 maxi prep flasks. Incubate plates at 37°C for 12 hours or greater, and liquid media in a shaking incubator 37°C for 12 hours or greater. Note: Use cells with a transformation efficiency at 1-3 x 10^9 cfu/ug or above. Critical Step: Calculate the number of transformation reactions to be used beforehand to ensure library coverage. We recommend transforming 2 ul of ligation mix per 1000 epitopes in the library. **Day 3 \(Timing: 1 hour)** 13. Calculate library coverage from the amount of transformed cells pre dilution, total amount of pooled cell mixture, and number of unique oligos in the library. Plasmid coverage = number of transformants / number of oligos Critical step: Ensure atleast 100X coverage for the library. Troubleshooting: If there is not sufficient coverage, perform more ligation/transformation reactions. If ligation fails completely, i.e. no colonies are obtained, optimize Insert:Vector ratio. 14. Maxi prep the 1.5 L of liquid media across three maxi columns. Pool maxi preps in a 2 mL eppendorf tube. Pause point: Store plasmids at 4°C for up to one month or at -20°C indefinitely. **Packaging Lentiviral Vectors** **Caution: All the steps in this subsection are to be performed at BSL2 with lentiviral vector precautions.** **Day 1 \(Timing: 30 minutes)** 15. Prepare D10 media \(55 mL of FBS and 5.5 mL Penicillin Streptomycin added to a 500 mL bottle of DMEM), warm D10 to room temperature. 16. Coat a 10 cm Tissue-culture treated dish with poly-L-Lysine. Add 10 ml of poly-L-Lysine to the plate and swirl to ensure the entire surface is covered. Aspirate

the remaining poly-L-Lysine. Note: Poly-L-Lysine can be reused multiple times. 17. From a 15 cm plate of HEK-293T cells, aspirate supernatant. 18. Add 6 ml of Trypsin and ensure that it covers the entire surface. 19. Incubate plate at 37°C for 2-3 min until the cells detach from the surface. 20. Add 15 ml of D10 and wash all the cells off the surface using a 10-ml serological pipet. 21. Form a single cell suspension by repeatedly pipetting up and down the cells. 22. Count the cells using a hemocytometer and adjust the concentration to 0.5e6 cells/ml with D10. 23. Add 10 ml of cell suspension to the poly-L-Lysine coated plate. Shake the plate forward-backward and side-by-side to ensure even distribution of the cells. 24. Incubate the plate overnight in a CO2 incubator at 37°C, 5% CO2, and 95% relative humidity. **Day 2 \ (Timing: 1.5 hours)** 25. Warm OPTI-MEM and D10 to room temperature. 26. Prepare a mixture of OPTI-MEM with TransIT-293. Use 1ml of OPTI-MEM and 100 ul of TransIT-293 per 10 cm plate. Mix by vortexing. Incubate at room temperature for 20 minutes. 27. Prepare a mixture of lentiviral vector shuttle and packaging plasmids. Use 5 ug of the lentiviral shuttle plasmid \(e.g. pCCLc-MND-A0201-MART1-SABR or a TCR-Lentiviral vector), 5 ug of the pCMV-RD8.9 plasmid, and 1 ug of the pMDG-VSVG plasmid. 28. After the 20 minute incubation, combine both mixtures and mix well by pipetting up and down gently. Critical Step: ensure that the incubation is atleast 15 min, but no more than 30 min. Ensure that the mixing is performed gently. Do not vortex. 29. Incubate at room temperature for 20 minutes. 30. During this incubation, aspirate the medium from the plated HEK-293T cells from Day 1. Replace the medium with 10 ml of fresh D10. 31. After the incubation in step 29, add the mixture dropwise to the plated HEK-293T cells. Critical Step: ensure that the incubation is atleast 15 min, but no more than 30 min. Do not disturb the cell monolayer. 32. Incubate the plate for 3 days in a CO2 incubator at 37°C, 5% CO2, and 95% relative humidity. **Day 5 \(Timing: 1 hour)** 33. Remove the plate from the incubator. Note: the cell-free medium from this plate will be the lentiviral vector to be used for transduction. 34. Harvest the medium with a 10 ml serological pipet without disturbing the cell monolayer. Filter the medium through a 0.45 micron PVDF filter. Critical step: Use only 0.45 micron PVDF filters. Using 0.22 micron or using other polymers will result in reduction of viral titers. Do not mix the cell-free medium while harvesting. 35. Aliquot the filtered vector in 1.5ml microcentrifuge tubes and store at -80°C until further use. Pause point: Store filtered virus at -80°C indefinitely. Troubleshooting: Viral titers may be negatively affected by plating a higher or lower cell number per plate, by prolonging or shortening the incubation beyond the recommended time, by filtering through the wrong filter, by using high passage number HEK-293T cells, by using plasmids that have degraded, by using incorrect amounts and ratios of the plasmids. **Transducing NFAT-GFP-Jurkat or Jurkat Cells** **Caution: All the steps in this subsection are to be performed at BSL2 with lentiviral vector precautions.** **Day 1 \(Timing: 30 minutes)** 36. Prepare R10 media. Warm to room temperature. 37. Count NFAT-GFP-Jurkat cells or Jurkat cells with FACS or hemocytometer. Dilute to 1e6 cells/ml in R10. 38. Thaw the viral vector from Step 35 at room temperature or 37°C. Mix by vortexing briefly. 39. In 12 well Tissue-culture treated plates, distribute 500 ul of cell suspension from Step 37. 40. Add 500 ul of the thawed viral vector. Mix well by pipetting up and down. 41. Incubate the plate for 3 days in a CO2 incubator at 37°C, 5% CO2, and 95% relative humidity. 42. If using NFAT-GFP-Jurkat cells, add 1 ml of R10+ 2mg/ml G-418 at 48 hours post-transduction. Troubleshooting: Transduction efficiency may be increased by using fewer cells or by using higher amount of the viral vector. **Co-culture Assay** **Caution: All the steps in this subsection are to be

performed at BSL2 with lentiviral vector precautions.** **Day 1 \(Timing: 1 hour)** 43. Prepare R10 media \(55 mL of FBS and 5.5 mL Penicillin Streptomycin added to a 500 mL bottle of RPMI), warm R10 to room temperature. 44. Count transduced target Jurkat cells from step 43 and transduced effector NFAT-GFP cells from step 43 on FACS or hemocytometer. 45. For each pair of targets and effectors, pool 1.5e6 effector cells and 1.5e6 target cells per replicate in 50 mL conical tube. Bring all tubes to equal volume with R10. 46. Centrifuge tubes at 500 x g for 5 minutes at 4°C. Aspirate supernatant. Resuspend with 6 mL R10 per replicate. Distribute resuspended cell mixture across wells of a 6 well plate, aliquoting 6 mL of cell mixture per well. Critical step: Do not add G-418 to the medium at this step. 47. Incubate cells for 8 hours, or up to 14 hours if performing overnight, in a CO2 incubator at 37°C, 5% CO2, and 95% relative humidity. Troubleshooting: Assay efficiency can be negatively affected if the NFAT-GFP-Jurkat cells are high passage, or are not split regularly. Prolonged experiment setup at room temperature may reduce the efficiency. **Day 2 \(Timing: 2-4 hours)** 48. Prepare MACS buffer \(10 mL of FBS added to a 500 mL bottle of PBS). Keep MACS buffer on ice. 49. Harvest cells by pipetting each well up and down until mixed, pool replicates in a 50 mL conical tube. Bring all tubes to equal volume with MACS buffer. Keep everything on ice. 50. Stain cells: Prepare staining solution \(2.5 ul of CD69 antibody added to 250 ul MACS Buffer per reaction). Centrifuge tubes at 800 x g for 5 minutes at 4°C. Aspirate supernatant. Resuspend with 250 ul staining solution per tube. Incubate at 4°C in the dark for 20 minutes. Add 5 mL MACS Buffer to each tube to wash staining solution, briefly vortex. Centrifuge tubes at 800 x g for 5 minutes at 4°C. Aspirate supernatant. Resuspend with 1 mL MACS Buffer. 51. Filter cells by pipetting through filter cap of 5 mL tube. Keep tubes on ice until sort, no longer than two hours after staining. 52. Verify on FACS: Transfer 5 ul of stained cells from each tube to a 96 well FACS plate, bring each well to 100 ul with MACS Buffer. Select lasers and gates for lymphocytes, GFP, and antibody color. Confirm that a double-positive GFP+CD69+ population exists. 53. Sort cells at a sorting facility according to their protocols. Sort GFP+CD69+ cells in MACS buffer. **Illumina Sequencing** **Day 1 \(Timing: 2-4 hours)** 54. Extract genomic DNA from sorted cells in Step 53 using the PureLink Genomic DNA kit, following manufacturer's instructions. Elute in 30 ul of elution buffer. Pause point: Store genomic DNA at 4°C for up to a month or at -20°C indefinitely. 55. Amplicon PCR: Prepare 200 ul Illumina amplification mix \(100 ul of 2X KOD master mix, 6 ul of 10 uM TruSeq-Univ-SCTfixed-F, 6 ul of 1 uM TruSeq-Read2-SCTfixed-R, 6 ul of 10 uM Truseq-Adapter-Index, 15 ul of extracted genomic DNA, and 58 ul of Nuclease-free water), using the appropriate index primer per sample, and aliquot mix to 8 wells of a PCR plate or strips, 25 ul per well. Amplify according to the following program: Initial Denaturation at 95°C for 2 minutes, 35 cycles of 95°C for 20 seconds \(denature), 66°C for 10 seconds \(anneal), and 70°C for 15 seconds \(extension), then a final extension at 70°C for 2 minutes. Pause point: Store PCR product at 4°C for up to a month or at -20°C indefinitely. 56. Pool all 8 reactions per index in a 1.5 mL eppendorf tube. Purify reactions with NucleoSpin Gel and PCR purification kit, using one column per index, and elute in 15 ul Elution Buffer \(supplied) per column. 57. Verify amplicon on gel or BioAnalyzer 1. Pour a 2% agarose gel \(mix 2.4 g agarose and 120 mL TAE buffer, heat to above 95°C until fully dissolved, cool to 60°C, add 10 ul gel green, mix and cast in a gel tray with wide combs until solidified). Combine 1 ul purified amplicon and 1 ul 6X loading dye, then load mixture on gel. Run gel at 120 V for 20 minutes, or until dye front has run 30% of the gel. Image the gel in a gel imager at 302 nm UV wavelength. Expected size of amplicon is 200

bp. Note: Amplicon might be very faint depending on the number of cells sorted. When amplifying very few cells, off-target amplification may occur. Analysis will ignore those products, but be sure to account for them while loading the flowcell. Caution: If using Ethidium bromide to visualize gels: Ethidium bromide is a carcinogen. Dispose of it appropriately. 2. Run on BioAnalyzer following manufacturer's instructions. In brief: Load 9 ul gel-dye on to chip, add 5 ul Nano Marker to each sample well, load sample and ladder into appropriate wells, vortex chip, place chip in BioAnalyzer, and start analysis. 58. Perfom Illumina sequencing on HiSeq2500 at a sequencing facility according to their instructions. Obtain the final data as FASTQ files. **Analysis \(Timing: 2 hours)** 59. Convert and de-multiplex basecalls with Illumina bcl2fastq2 software 1. Create sample sheet with Illumina Experiment Manager to define indexes 2. Configure bclToFastq: configureBclToFastq.pl --input-dir --output-dir --sample-sheet /SampleSheet.csv Note: This script will generate a makefile 3. Run bclToFastq: make 60. Generate bwa index 1. Download .fasta sequence for backbone plasmid from addgene 2. Replace the 2kb BsmBI stuffer sequence with a string of random nucleotides the length of the average genetically encoded epitope \(36 random bp for 12 AA) 3. Build the bwa index: bwa index -p ref backbone.fasta 61. Align each sample to plasmid backbone with bwa bwa mem -B 1 ref input.fastq > output.sam Note: the flag "-B 1" reduces the mismatch penalty to account for the variable epitope sequences 62. Trim alignment to only epitope region For each read: target_sequence_start = epitope_start_position - read_start target_sequence_stop = epitope_stop_position - read_start + 1 target_sequence = read_sequence\[target_sequence_start:target_sequence_stop] 63. Translate to peptide For each codon: amino_acid_sequence += codon_table\[codon] 64. Count library peptide hits For each amino_acid_sequence: if amino_acid_sequence in library_sequences_table: library_sequences_table\[amino_acid_sequence] += 1 65. Export the read counts to .xls or .csv files. 66. Assign each epitope in a sorted sample a rank according to the number of reads corresponding to the epitope. 67. Average the ranks for replicates for a given sample to calculate 'Average Rank'. 68. Plot Average Rank for the sort for a given TCR against the Mock-sort in scatter plots with each sort on either axis. Troubleshooting: If there are no standout hits, it could mean that the TCR under investigation does not have its target peptide in the library or does not recognize the MHC used in SABRs. 69. Validate the top hits using traditional in vitro assays.

Anticipated Results

To demonstrate proof-of-concept of the approach detailed in Fig 1, we asked if SABR libraries can identify the cognate antigens of known public TCRs in an unbiased manner. We constructed a SABR library encoding 12,055 epitopes presented on HLA-A**0201 \(A2-SABR-library) consisting of all known HLA-A**0201-restricted epitopes from the Immune Epitope Database24 \(IEDB). We first interrogated if the A2-SABR-library allows identification of the cognate antigen for two TCRs with known specificities \(F5, which recognizes EAAGIGILTV, and SL9, which recognizes SLYNTVATL). We transduced NFAT-GFP-Jurkat cells with the A2-SABR-library, and incubated them with Jurkat cells expressing F5 or SL9 TCRs. After 10 hours of co-culture, we sorted GFP+CD69+ cells by FACS \(Fig 3a), extracted genomic DNA, sequenced the epitopes, and calculated average ranks for each epitope as described in Fig 41 First, we plotted the

average ranks of each of the epitopes from the SL9 sort against those from F5 sort \(Fig 3b). Six epitopes formed a distinct cluster by their rank in the F5 sort. In the SL9 sort, the top ranks were outliers, but did not form a separate cluster. The top six epitopes in the F5 sort were analogs of EAAGIGILTV, indicating successful identification of its antigen \(Fig 3c). Six out of the top ten epitopes from the SL9 sort were analogs of SLYNTAVATL \(Fig 3d). Epitopes enriched for their corresponding TCRs were not enriched in mismatched TCRs \(Fig 3c and d). The average fold-enrichment of the top hits from the F5- and SL9-sorts over the Mock-sort was 296 and 70 respectively. The noise observed in the SL9 sort is possibly due to the higher number of analogs of the SLYNTVATL peptide. The A2-SABR library contains 22 analogs EAAGIGILTV and 60 analogs of SLYNTVATL. A higher number of recognized analogs would lower the Average Rank for each of the epitopes because of competition among the epitopes. We compared the ranks of all the analogs of EAAGIGILTV and SLYNTVATL in the sorts. Six out of twenty-two EAAGIGILTV analogs were identified in the F5 sort \(Fig 3e), whereas nine out of sixty SLYNTVATL analogs were identified in the SL9 sort \(Fig 3f). The lack of identification of all the analogs is presumably due to reduced cross-reactivity of the F5 or SL9 TCRs towards them. Indeed, analogs SLYNTIATL \(V6I) and SLFNTVATL \(Y3F) are documented escape mutations in the SLYNTVATL epitope. We validated the top six hits from the F5 sort by in vitro cytotoxicity assays. We observed that all six analogs of the Mart1 peptide were specifically recognized by the F5 TCR, leading to induction of cytotoxicity \(Fig 3g). Nevertheless, these experiments showed that a SABR library approach could identify the cognate antigen of a TCR by screening thousands of epitopes.

Figures
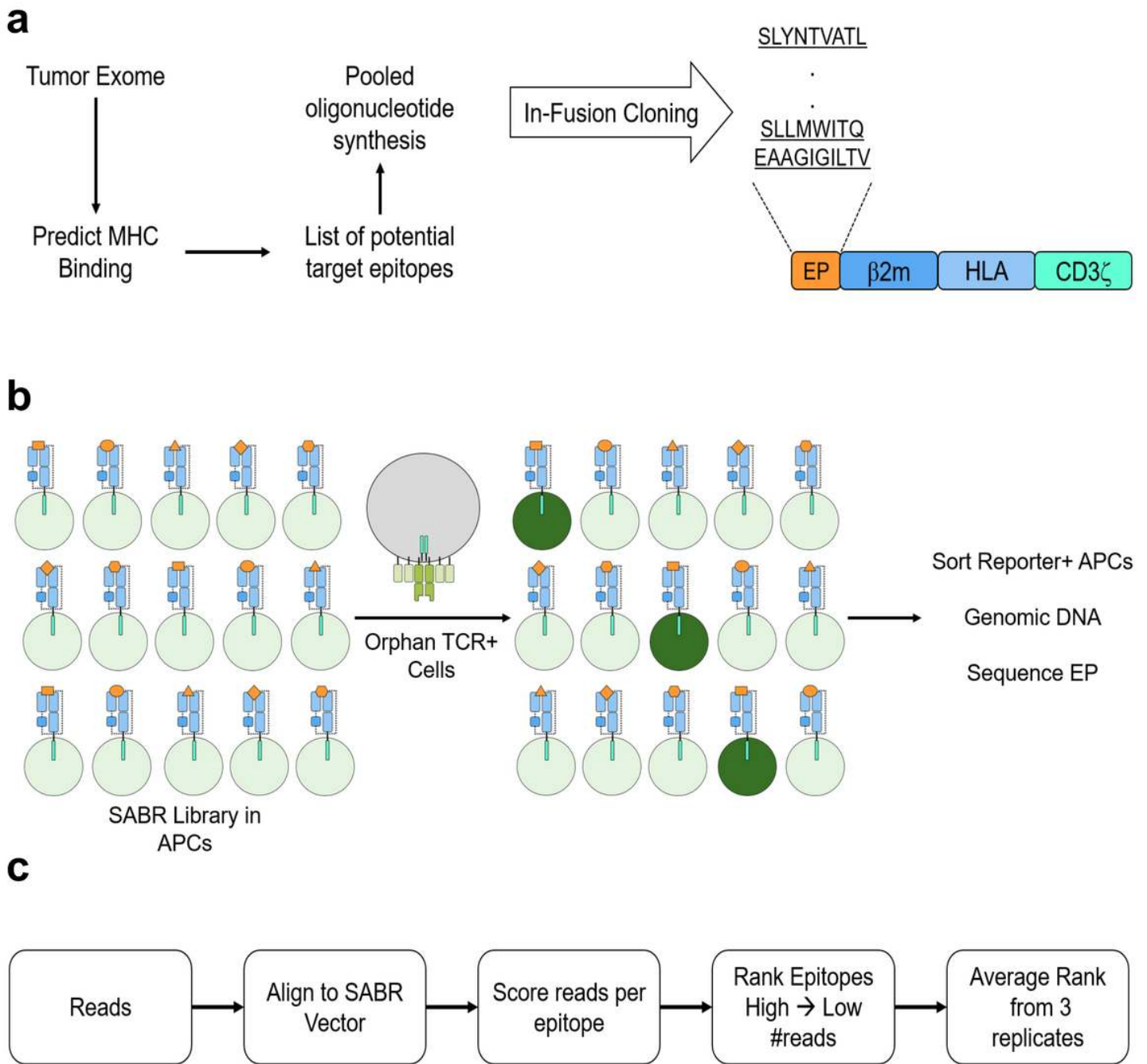
Figure 1

Overview of SABR-mediated antigen discovery a. Schematic showing the pipeline to construct custom SABR libraries. EP – epitope. The left panel shows the procedure to obtain and synthesize a list of epitopes. The right panel shows the schematic of SABR library. b. Schematic showing co-culture experiment to select cells from SABR library that are recognized by an orphan TCR. Left panel shows a SABR library presenting numerous unique epitopes. The middle panel shows cells APCs showing reporter expression induced by SABRs presenting the cognate epitope for the orphan TCR. Right panel shown processing of the selected cells. c. Flowchart showing the computational analysis pipeline.
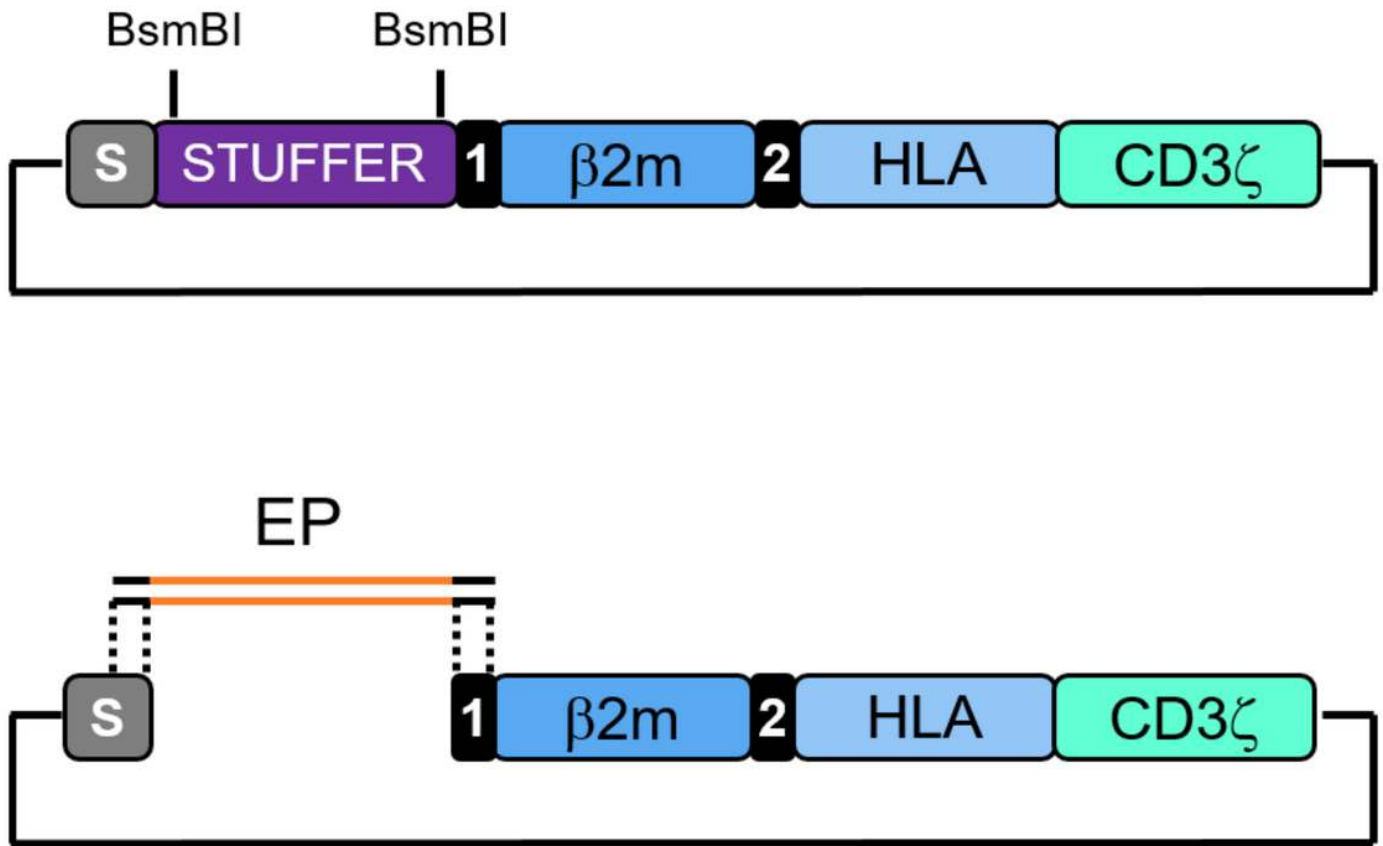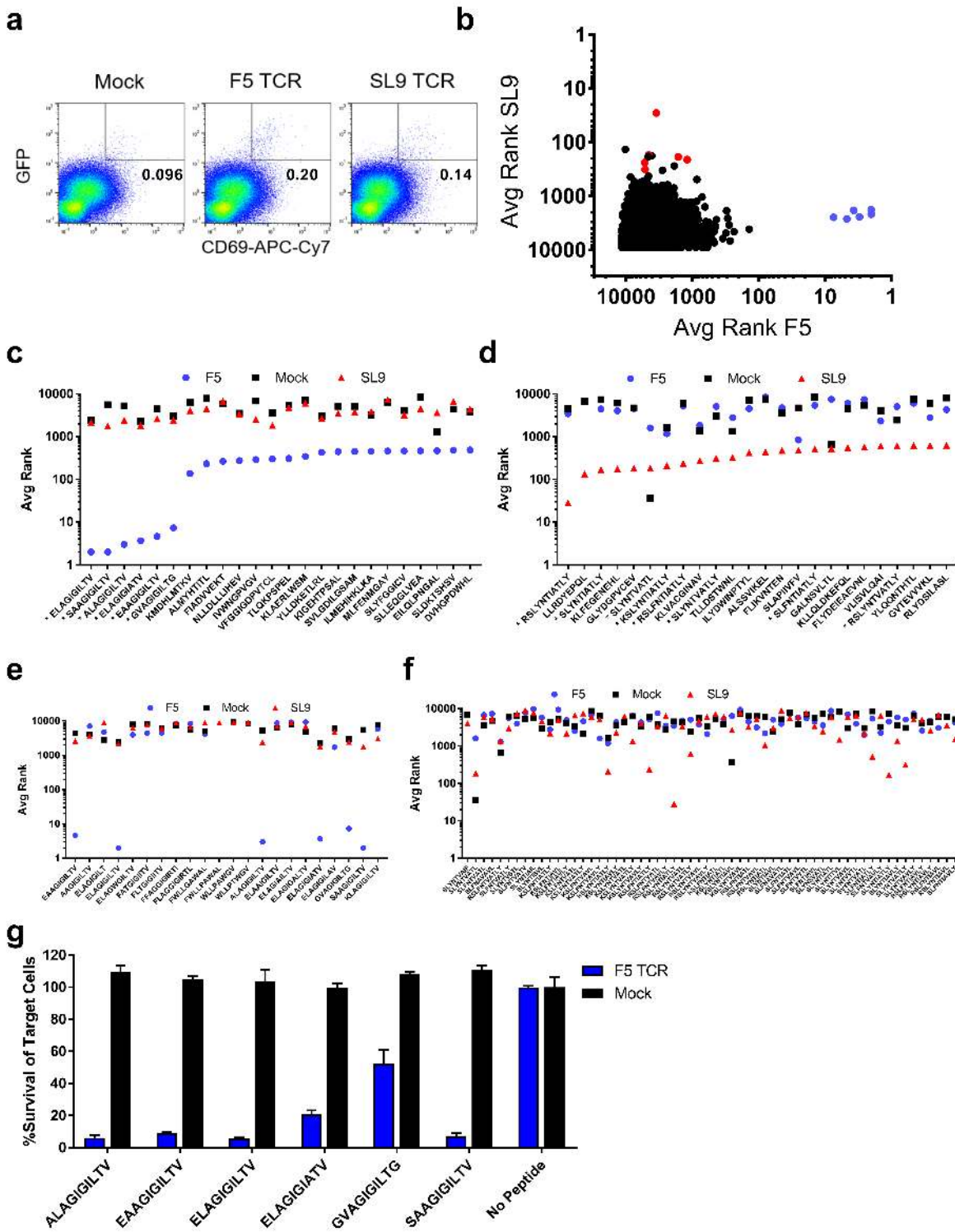
Figure 2

SABR backbone and schematic for cloning SABR vector constructs with a stuffer fragment showing BsmBI sites (top), and cloning strategy using double stranded oligonucleotides with encoding the epitope flanked by overlaps is shown

Figure 3

Example of expected outcome for two TCRs a. Sorting A2-SABR library cells based on reporter gene expression. Co-culture assays using 9 million library cells with 9 million TCR-transduced Jurkat cells were set up. At 10 hours post co-culture, cells were stained for CD69 and sorted using FACS. Representative flow cytometry plots from one replicate are shown. The rectangle in the top right corner of each flow plot shows the gate used for the sort. Frequency of cells in the sort gate is indicated as percentage. b. Average

ranks from F5 and SL9 sorts. Each dot represents the average rank for a unique epitope as calculated using the procedure described in Fig 4. Y-axis shows average rank in the SL9 sort, and X-axis shows the average rank in the F5 sort. Purple dots indicate EAAGIGILTV analogs and red dots indicate SLYNTVATL analogs. c. Average ranks for the top 24 hits from the F5 sort. Epitopes with asterisks indicate EAAGIGILTV analogs. d. Average ranks for the top 24 hits from the SL9 sort. Epitopes with asterisks indicate SLYNTVATL analogs. e. Average ranks for all the EAAGIGILTV analogs in the A2-SABR library. f. Average ranks for all the SLYNTVATL analogs in the A2-SABR library. In fig c-f, Mock-sort indicates an assay where the entire SABR library is co-cultured with Mock-transduced Jurkat cells, and the background of GFP+CD69+ is sorted. g. Validation of the top hits in the F5 sort by cytotoxicity assays. F5- or Mock-transduced primary T cells were incubated with CFSE-labeled K562 cells expressing HLA-A2.1 and pulsed with the indicated peptides. At 24 hours post-incubation, live K562 cells were counted by flow cytometry and normalized to the "no peptide sample" as described previously. The bars indicate means±s.d. for n=3.