1 **Antigen identification and high-throughput interaction mapping by reprogramming**
2 **viral entry**
3
4 Connor S. Dobson[1,2], Anna N. Reich[1,2], Stephanie Gaglione[2,3], Blake E. Smith[2,4], Ellen J.
5 Kim[1,2], Jiayi Dong[1,2], Larance Ronsard[5], Vintus Okonkwo[5], Daniel Lingwood[5], Michael
6 Dougan[4,6], Stephanie K. Dougan[4,7], Michael E. Birnbaum[1,2,5,*]
7
8 [1] Department of Biological Engineering, Massachusetts Institute of Technology,
9 Cambridge, MA, USA.
10 [2] Koch Institute for Integrative Cancer Research, Cambridge, MA, USA.
11 [3] Department of Chemical Engineering, Massachusetts Institute of Technology,
12 Cambridge, MA, USA.
13 [4] Program in Immunology, Harvard Medical School, Boston, MA, USA.
14 [5] Ragon Institute of MIT, MGH, and Harvard, Cambridge, MA, USA.
15 [6] Division of Gastroenterology, Department of Medicine, Massachusetts General Hospital
16 and Harvard Medical School, Boston, MA, USA
17 [7] Department of Cancer Immunology and Virology, Dana-Farber Cancer Institute, Boston,
18 MA, USA
19 *e-mail: mbirnb@mit.edu
20
21 **Abstract**
22 Deciphering immune recognition is critical for understanding a broad range of diseases
23 and for the development of effective vaccines and immunotherapies. Efforts to do so are
24 limited by a lack of technologies capable of simultaneously capturing the complexity of
25 adaptive immune receptor repertoires and the landscape of potential antigens. To
26 address this, we present RAPTR (**R**eceptor-**A**ntigen **P**airing by **T**argeted **R**etroviruses).
27 RAPTR combines viral pseudotyping and molecular engineering approaches to enable
28 one-pot library on library interaction screens by displaying antigens on the surface of
29 lentiviruses and encoding their identity in the viral genome. Antigen-specific viral infection
30 of cells allows readout of both antigen and receptor identities via single-cell sequencing.
31 The resulting system is modular, scalable, and compatible with any cell type, making it
32 easily implemented. These techniques provide a suite of new tools for targeted viral entry,
33 molecular engineering, and interaction screens with broad potential applications.

34   The adaptive immune system recognizes and responds to the diverse pathogens

35   humans encounter throughout their lives. B cells recognize extracellular antigens via the

36   B cell receptor (BCR), leading to the production of secreted antibodies to inhibit and

37   eliminate pathogens. T cells use T cell receptors (TCRs) to recognize short peptide

38   fragments presented by major histocompatibility complex proteins (pMHCs), enabling

39   antigen-specific coordination of the immune response and elimination of malignant or

40   infected cells. This remarkable ability to sense, respond to, and remember threats is key

41   to any successful adaptive immune response, and accordingly is the cornerstone of

42   successful vaccines and immunotherapies.

43   The importance of T cell-mediated immunity has led to the development of a series

44   of approaches dedicated to identifying antigens or pMHC-TCR pairs[1]. Conventional T cell

45   assays, such as ELISPOT or intracellular cytokine staining[2], provide direct readouts of T

46   cell function. While these approaches can be highly multiplexed, they do not readily lend

47   themselves to TCR sequencing and are limited in their ability to identify single reactive

48   antigens. More recently, T-Scan has broadened the antigenic scope of functional assays

49   to genome scale[3], but cannot provide paired receptor information without panning of pre-

50   determined TCRs or iterated steps of target antigen identification followed by sorting of

51   reactive cells for TCR sequencing. Other recent cell-based reporter assays convert

52   pMHCs into the recognition domain of immune signaling complexes[4,5]. These approaches

53   can identify successful TCR-pMHC interactions for single TCRs from antigen libraries on

54   the scale of 10^3-10^4, but require substantial library redundancy, are limited in their ability

55   to multiplex TCRs, and may require multiple rounds of screening. Requiring a discrete

56   experiment per TCR imposes a significant scalability constraint; each individual has a

57 TCR repertoire of ~10^7 unique clones[6,7], and very little overlap is expected between

58 individual repertoires even for people who share common MHC alleles.

59   Recombinant protein-based screening can broaden either the number of T cell

60 clones or the number of antigens that are feasible to screen. Yeast display of pMHCs

61 enables screening of ~10^8 unique pMHC antigens[8,9], but can only examine a limited

62 number of recombinantly expressed TCRs at a time, and can require significant

63 optimization for each MHC allele. Conversely, recent advances in barcoded pMHC

64 multimers enable screens on the order of 10^3 antigens in bulk[10,11], or hundreds while

65 maintaining receptor-antigen pairing[12,13]. While such analyses can be performed on

66 polyclonal T cells, they are inherently bottlenecked by several technical limitations: (i) the

67 need to manually assemble individual barcoded multimers; (ii) the ability to correctly

68 identify interactions in large pools of multimers; and (iii) the relatively small set of MHC

69 molecules that have been recombinantly expressed successfully.

70   Identifying antigenic targets of B cells poses similar limitations. Recent approaches

71 have shown that antigen-BCR pairs can be identified via the oligonucleotide tagging of

72 recombinantly expressed proteins[14,15], which is inherently limited in scale. Thus, our

73 understanding of antigen recognition by both arms of the adaptive immune system is

74 currently constrained by significant hurdles to experimental scale.

75   Deciphering the full complexity of immune recognition requires the ability to screen

76 for interaction pairs while incorporating diversity of both antigen receptors and their

77 targets at the same time. This represents a broader experimental challenge for interaction

78 screens commonly known as "library vs library screening." The most well-established

79 approach to this problem is yeast two-hybrid[16], in which intracellular protein pairs are used

80   to drive expression of a reporter gene. Other previously established approaches include

81   automated individual ELISA screens of recombinantly expressed protein pairs[17–19], as

82   well as mass-spectrometry-based identification of interacting pairs[20,21]. There have been

83   several recent efforts at designing novel systems, including those based on yeast mating

84   or spatial colocalization of DNA-barcoded molecules[22,23]. However, these approaches

85   can be labor-intensive, are often not suited for complex, extracellular protein complexes

86   such as immune receptors, and may be inefficient at low (i.e. micromolar) affinities.

87   Together, these limitations have thus far rendered such approaches unsuitable for

88   applications such as identifying TCR-pMHC pairs.

89   To overcome the above limitations, we have developed a technique that combines

90   lentiviral surface display[24–26] with a versatile pseudotyping strategy and viral genome

91   engineering to enable one-pot library vs library screening (**Figure 1**). We demonstrate

92   that our pseudotyping strategy – an engineered fusogen termed VSVGmut co-expressed

93   with a targeting moiety – is general and versatile for both receptor and ligand usage. We

94   leverage these abilities to present Receptor-Antigen Pairing by Targeted Retroviruses

95   (RAPTR), which matches receptors with their cognate antigens based on specific

96   infection of receptor-expressing cells by antigen-displaying viruses. Putative hits can be

97   read out by bulk or single-cell sequencing, enabling screens of single or polyclonal

98   receptors. We demonstrate the feasibility of this approach for both T and B cell receptors,

99   including a library on library screen comprising 96 pMHC antigens and receptors enriched

100   from a library of >450,000 TCRs (thousands of potential interactions).

101   **Results**

102   VSVGmut pseudotyped lentiviruses enable modular tropism

103   Developing a scalable pipeline for antigen-receptor screening presents several

104   requirements: (i) a straightforward ability to track both receptor and antigen sequences;

105   (ii) mammalian expression systems to maximize the ability to express complex, multimeric

106   proteins; and (iii) the ability to generate selection reagents without the need to

107   recombinantly express and characterize each protein for every experiment. To address

108   these needs, we turned to lentiviruses, which have decades of precedent for molecular

109   manipulations. They can be created at large scale, are already known to infect T cells,

110   and, upon successful infection, leave a permanent record of infection due to their

111   integration into the host genome.

112   Lentiviruses have been pseudotyped via a number of different strategies to enable

113   their use as biotechnology tools and gene therapy vectors[27]. Due to its robustness and

114   efficient infection of many cell types, vesicular stomatitis virus G protein (VSVG, referred

115   to as VSVGwt in this manuscript) is the most common pseudotype for laboratory studies

116   and cell manufacturing for clinical applications[28]. More recently, approaches to enable

117   cell type-specific targeting, via co-expressing receptor-blinded versions of Sindbis virus[29–

118   [31] or paramyxovirus envelope proteins[32–37] with targeting moieties have been described.

119   While these approaches show promise, they have been reported to have strict limitations

120   on targeting ligand and receptor binding topology due to their mechanism of entry.

121   Since these factors constrain the generality of an interaction screening system, we

122   developed an alternative strategy based upon VSVG. We used the recently described

123   crystal structure of VSVG in complex with its native receptor[38], the low-density lipoprotein

124   receptor (LDLR), to engineer VSVGmut, which incorporates the K47Q and R354A

125 mutations reported to ablate affinity for the LDLR family of receptors (**Figure 1A**). To

126 retarget VSVGmut pseudotyped viruses, we co-expressed a variety of surface-bound

127 molecules during viral production (**Figure 1B**). To benchmark against established

128 systems[33], we used IL-13 as a viral targeting ligand. After testing several surface

129 architectures (**Supplementary Figure 1**), we found that while many distinct constructs

130 conferred specific infection of IL-13Rα1-expressing cells, the dimerized, surface-tethered

131 IL-13 yielded the most efficient infection while retaining excellent specificity (**Figure 1C**).

132 We also found that VSVGmut pseudotyped viruses displaying an anti-CD19 single-chain

133 antibody fragment (scFv) efficiently infected CD19$^+$ B cell lines, but not CD19$^-$ T cell lines

134 (**Figure 1D**).

135 Next, we sought to exploit the unique modularity of the VSVGmut system to

136 develop combinatorial targeting strategies. While display of the anti-CD3 Fab UCHT1

137 yielded only modest infection of Jurkat T cell lines, CD80 mediated robust infection

138 (**Supplementary Figure 2A**). Although the infection rate for the combinatorial strategy

139 was not better than for CD80 alone, we reasoned that the ability to incorporate multiple

140 signals during infection could be useful for engineering primary cells by providing user-

141 defined phenotypic inputs. Therefore, we used these viruses to infect primary CD8 T cells

142 and measured infection and CD25 upregulation as a marker of activation (**Figure 1E**).

143 We observed that the anti-CD3/CD80 combination viruses both infected and activated

144 cells, but viruses displaying CD80 alone infected but did not activate the cells. To our

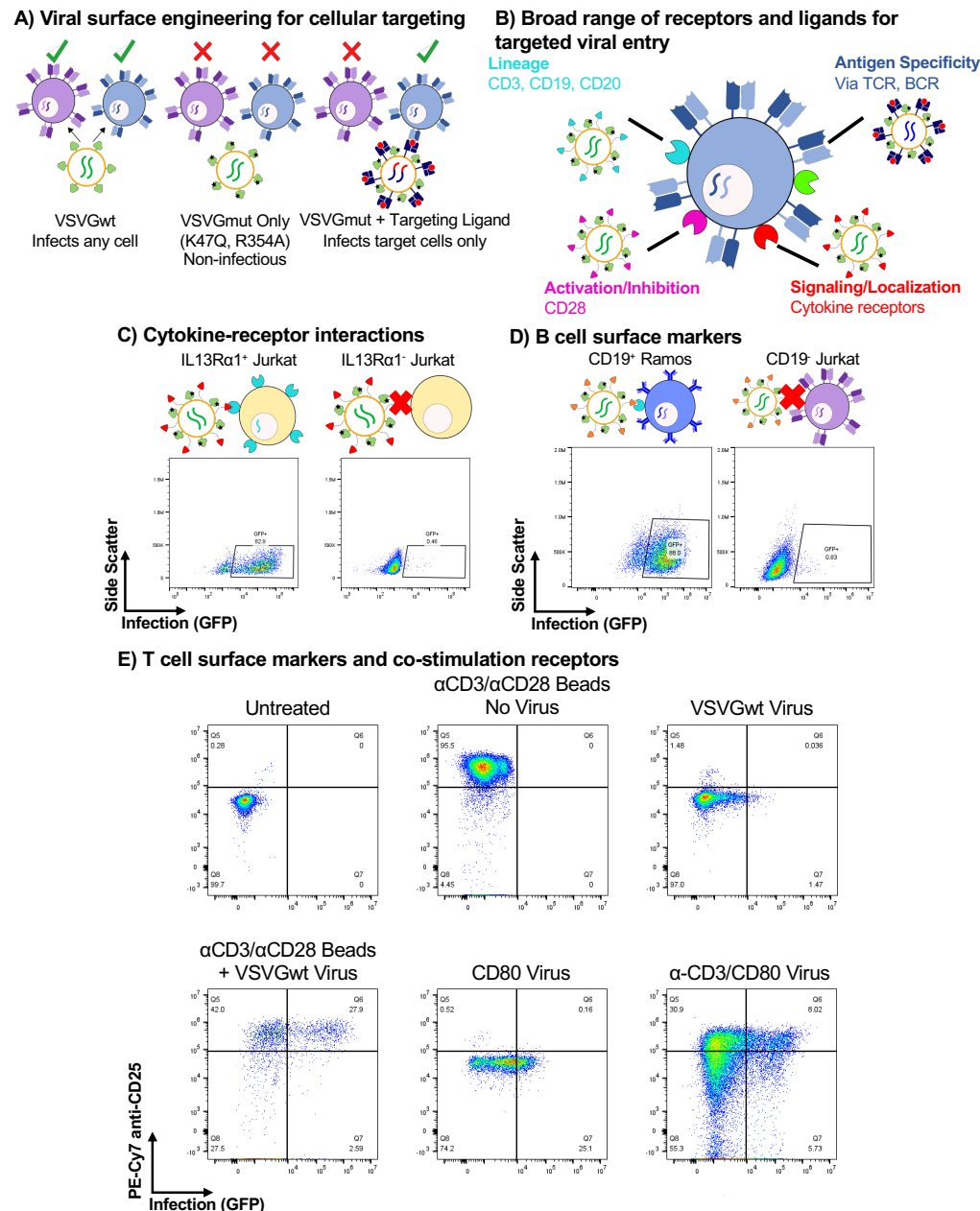145 knowledge, this is the first example of a combinatorial viral targeting system reported to

146 date.

Figure 1 | The VSVGmut pseudotyping system is modular and utilizes a broad range of receptors and ligands

A, VSVGmut (VSVG K47Q, R354A) pseudotyping system enables specific targeting by coexpression of receptor-blinded VSV-G and a modular targeting ligand. **B,** VSVGmut is compatible with a broad range of receptors and targeting ligands, including immune cell surface markers, signaling and antigen receptors. **C,** Specific infection of IL-13Rα1-expressing Jurkat cells (left) as compared to parental cell line (IL-13Rα1⁻, right) via VSVGmut lentiviruses displaying surface tethered IL-13; data representative of 2 biologically independent experiments. **D,** Specific infection of CD19⁺ Ramos cells (left), but not CD19⁻ Jurkat cells (right) by VSVGmut pseudotyped viruses displaying an anti-CD19 scFv; data are representative of 2 biologically independent experiments. **E,** Infection of Primary CD8 T cells (MOI = 1) via lentiviruses displaying anti-CD3 Fab and/or the costimulatory receptor CD80, or VSVGwt viruses with or without anti-CD3/anti-CD28 magnetic beads; data were collected at day 4 post-infection and are representative of three biological replicates.

147   Antigen-specific T cell targeting

148   Following the observation that anti-CD3 viruses yielded infection of T cells via a

149   component of the TCR complex, we next sought to determine whether we could achieve

150   antigen-specific cell entry via the TCR itself. One potential challenge to repurposing the

151   TCR-pMHC interaction for viral entry is its affinity, which is typically 1-50 μM[39]. We

152   therefore generated cell lines expressing a range of previously characterized affinity

153   variants for the 1G4 TCR[40,41], which recognizes a peptide derived from the NY-ESO-1

154   cancer testis antigen (SLLMWITQV) presented in the context of HLA-A*02:01, with

155   reported affinities ranging from picomolar to micromolar (**Figure 2A**). To display these

156   molecules on the virus surface, we expressed them as single-chain trimers[42], which are

157   comprised of covalently linked peptide, beta-2-microglobulin, and MHC. We found that

158   pMHC-displaying lentiviruses are able to efficiently infect T cells in an antigen-dependent

159   manner. We observed similar infection efficiency across the tested affinity range, with

160   only a modest reduction for the lowest-affinity variant (**Figure 2B**). To ensure that TCR-

161   mediated infection was generalizable, we displayed several individual pMHCs as single-

162   chain trimers alongside VSVGmut and used them to infect either Jurkat cells expressing

163   its parental TCR (off-target) or on-target J76 cell lines expressing their cognate TCRs

164   (**Figure 2C**). We observed specific infection across 3 different TCR-pMHC pairs with

165   minimal background infection. As VSVG enables cell entry via an endocytic route[43], and

166   because endocytosis is one of the earliest downstream events following T cell

167   activation[44–46], we next sought to determine whether viral infection was accompanied by

168   TCR signaling. We measured CD69 expression following viral infection (**Figure 2D**), and
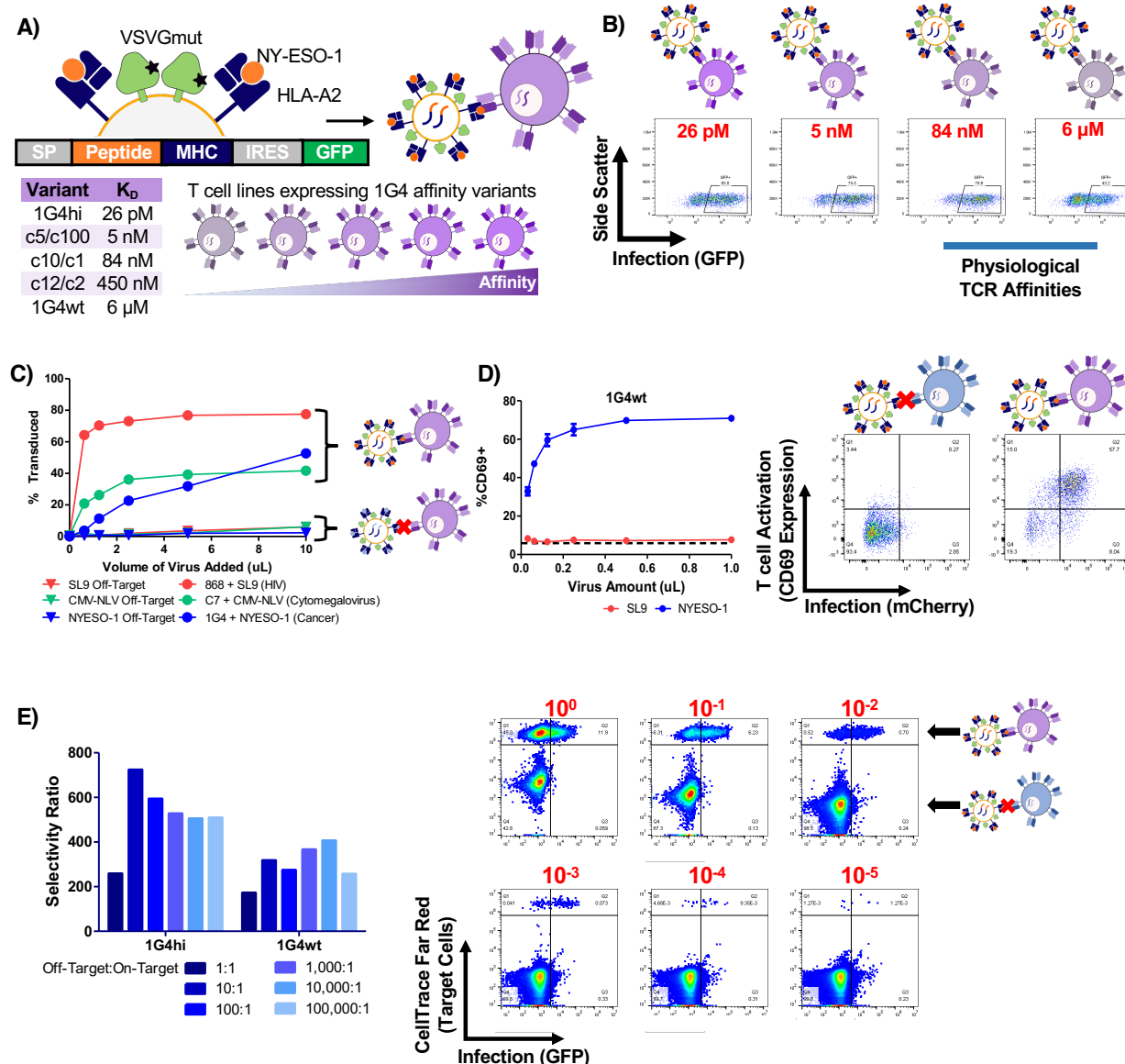
Figure 2 | TCR-mediated infection is sensitive, specific, and induces signaling
A, Schematic of the 1G4 TCR-HLA-A2/NY-ESO-1 system used for TCR-pMHC characterization. **B,** Representative infection data for various 1G4 TCR affinity variants by HLA-A2/NY-ESO-1 displaying VSVGmut lentiviruses. **C,** infection of J76 cells via specific TCR-pMHC interactions via pMHC-displaying VSVGmut lentiviruses for three independent TCR-pMHC pairs (HLA-A2/SL9-displaying viruses infecting 868 TCR-expressing J76 cells; HLA-A2/NLV-displaying viruses infecting C7 TCR-expressing J76 cells; HLA-A2-NY-ESO-1-displaying viruses infecting 1G4 TCR-expressing J76 cells). **D,** Upregulation of CD69 on J76 cells transduced with the 1G4wt TCR during viral entry demonstrates that viral entry induces TCR signaling; data shown represent mean + SD across three biological replicates, dashed line represents CD69 expression in untreated J76-1G4wt cells. **E,** Selectivity ratio of on target to off target infection of J76 cells expressing 1G4 TCR variants mixed at indicated ratios with off-target Jurkat cells; representative of three independent experiments, flow plots correspond to 1G4wt TCR-expressing J76 cells with target cell frequencies indicated in red.

169   observed robust CD69 upregulation during TCR-mediated viral entry, but not during off-

170   target combinations. Moreover, TCR-mediated entry was inhibited by dasatinib, a TCR

171   signaling inhibitor[47], while TCR-independent infection via VSVGwt was unaffected

172   (**Supplementary Figure 2B**). Thus, we concluded that pMHC-targeted viruses integrate

173   both binding and signaling as a means of infection.

174   We next sought to determine the limitations of our system by characterizing its

175   sensitivity and specificity using the 1G4-NY-ESO-1 system. To determine the ability of

176   our approach to discern on-target interactions in complex mixtures, we labeled 1G4-

177   expressing cells with a cell tracking dye, mixed them at varying ratios with unlabeled

178   Jurkats, and infected the mixture with NY-ESO-1/HLA-A2 displaying viruses (**Figure 2E**).

179   We calculated a selectivity ratio of on-target infection relative to off-target infection, and

180   observed on-target selectivity greater than 200:1, down to target cell frequencies of 1x

181   10^{-5}. Thus, we concluded that our system can detect on-target interactions in the

182   micromolar affinity range, even in complex mixtures containing hundreds of thousands of

183   non-target T cells.

184

185   Development of viral packaging strategy that maintains genotype-phenotype linkage

186   To fully enable interaction screening at library scale, we next turned to the

187   challenge of producing the lentiviral libraries. While lentiviral screens have been widely

188   used for functional genomics, standard lentiviral packaging techniques pose inherent

189   limitations that are particularly relevant to interaction screens. The key challenge stems

190   from the fact that at least hundreds of plasmids containing library elements are mixed in

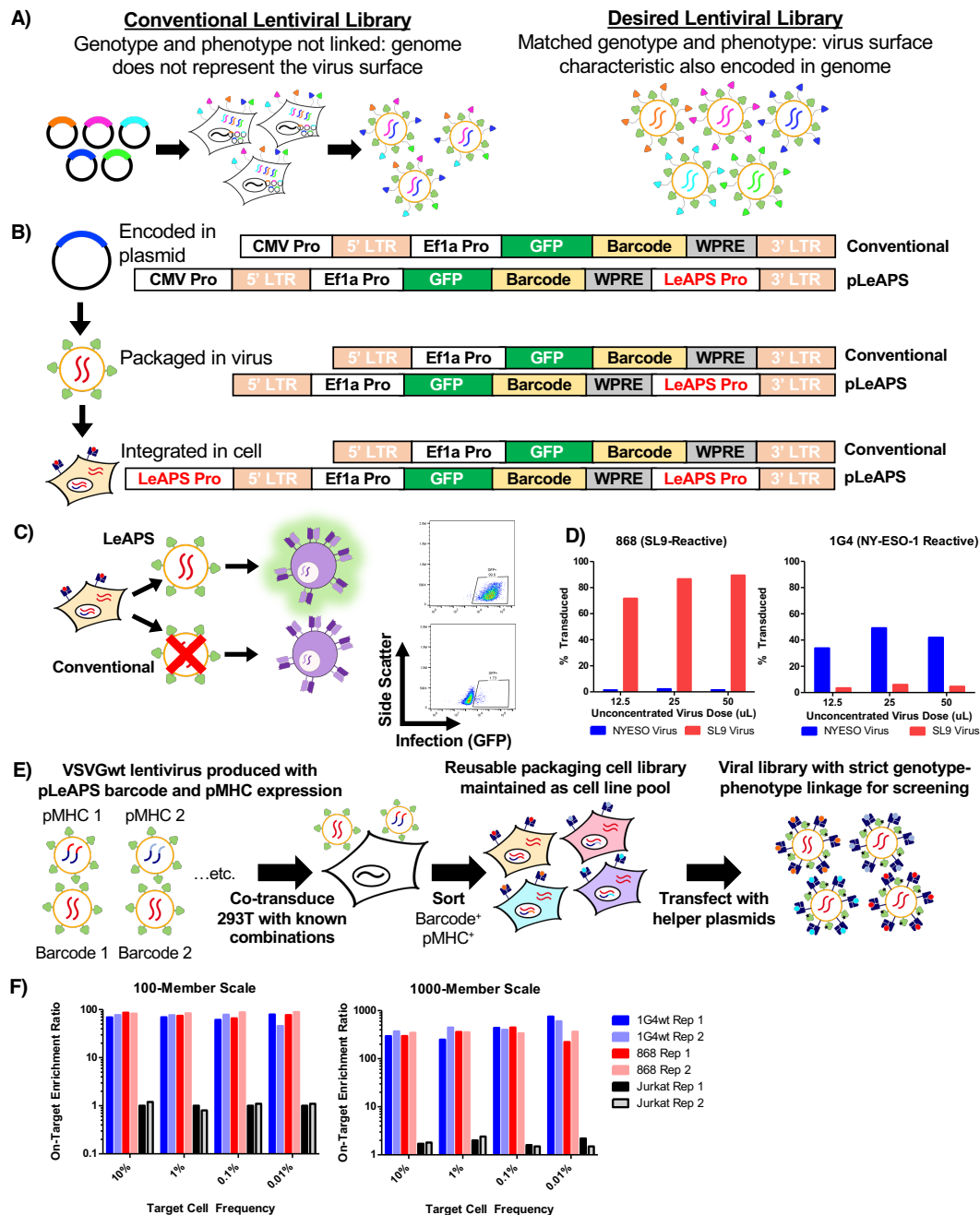191   each transfected packaging cell (**Figure 3A**). In functional genomics, this leads to well-

Figure 3 | LeAPS-based lentiviral packaging enables scalable viral library construction
A, Schematic overview of the limitation of conventional lentiviral packaging strategies for surface displayed lentiviral libraries, in which a mixture of displayed surface proteins disrupts the linkage between viral pseudotype and genomic identity. **B,** Schematic overview of the mechanism of promoter translocation used in the LeAPS system, in which a promoter sequence placed internally in the lentiviral genome is placed 5' of the 5' LTR during the viral lifecycle, thus allowing the integrated vector to produce viral genomes at high copy number. If cells are infected with a single lentiviral vector, each infected cell is ensured to produce viruses that match viral surface phenotype and genotype. **C,** Packaging cells transduced with viruses containing LeAPS-encoding genomes and then transfected with lentiviral helper plasmids produce high titer virus (top), while those transduced with viruses containing standard genomes produce minimal functional

virus (bottom). **D,** LeAPS-produced viruses maintain TCR specificity in two independent TCR-pMHC systems, demonstrating functional targeted viruses can be created. **E,** Schematic overview of the implementation of LeAPS for viral library assembly. 293T cells are co-transduced with known combinations of LeAPS-fluorophore-barcode viruses and pMHC expression viruses, all cells are pooled and sorted to generate a library packaging line that generates virus following helper plasmid transfection. **F,** On-target enrichment of libraries using defined mixtures of A2-NY-ESO-1 and A2-SL9 packaging cells to represent 100- or 1,000-variant libraries.

192   described intermolecular recombination of library elements[48,49], which is a significant

193   source of noise in combinatorial experiments. To date, solutions to this problem have

194   included plasmid dilution (accompanied by a 100x reduction in viral titer)[49], or simply

195   restricting libraries to sizes suitable for arrayed screens. This problem is compounded in

196   our approach, as multiple plasmids entering the same packaging cell compromises the

197   link between viral genotype and surface phenotype.

198   To obviate this issue, we exploited a detail of HIV-1 replication that results in the

199   copying of sequences between the polypurine tract and the 3' LTR to the 5' end of the

200   genome during reverse transcription and integration[50]. There have been several recent

201   reports exploiting this phenomenon to enable genomic screens via copying of CRISPR

202   gRNA cassettes[51,52]. Here, we insert a promoter to enable re-packaging of the viral

203   genome following infection of a cell[53] (**Figure 3B**). Conventional 3rd generation lentiviral

204   packaging approaches are inherently self-inactivating in part due to an inability to

205   transcribe packageable viral genomes following infection; i.e. the LTRs, which are weak

206   promoters, are truncated, and there is no promoter upstream of the viral genome following

207   successful integration. Our approach, which we term **Le**ntiviruses **A**ctivated by **P**romoter

208   **S**huffling (LeAPS), enables generation of lentiviruses from previously transduced cells

209   upon the re-introduction of helper plasmids. As a result, we can use lentiviral transduction

210   to ensure a single library member per packaging cell without incorporating additional viral

211   elements into the genome. We first validated that viruses produced with the LeAPS

212   strategy yielded high-titer virus (>10^6 TU/mL unconcentrated), whereas conventional

213   constructs (without a LeAPS promoter) yielded very little following transduction (**Figure**

214   **3C**). By stably expressing pMHC cassettes in LeAPS-transduced packaging cells, we

215   verified that viruses produced in this manner maintain antigen-specific targeting capability

216   (**Figure 3D**).

217   To develop a scalable library assembly strategy, we co-transduced 293T cells with

218   two VSVGwt-pseudotyped viruses: one that drives surface expression of a known pMHC

219   using a conventional lentivirus, and another that uses the LeAPS system to deliver a

220   repackageable genome driving expression of a fluorescent protein and a defined barcode

221   (**Figure 3E**). Following initial transduction, packaging cells for each library member were

222   pooled and sorted for expression of pMHC and barcode in a single sorting step. This

223   yields a library packaging line that can be used to generate the same lentiviral library

224   indefinitely via single transfections with only helper plasmids. By contrast, conventional

225   approaches require arrayed transfection to produce each library member for each

226   experiment[54], which poses a substantial impediment to experimental throughput.

227   To validate the utility of our approach, we employed a two-component system of

228   the HIV SL9 antigen (SLYNTVATL) paired with a LeAPS-mCherry cassette and the NY-

229   ESO-1 antigen paired with a LeAPS GFP cassette. By mixing the packaging cells for each

230   at different ratios, we assessed the feasibility of pMHC libraries at 100-member and

231   1,000-member scales (**Supplementary Figure 3**). We then used flow cytometry to

232   quantify signal to noise across different frequencies of on-target cells (**Figure 3F**). Across

233   both library sizes, and extending from target cell frequencies of 10% down to 0.01%, we

234   observed robust enrichment of on-target interactions, but minimal enrichment of off-target

235   interactions (i.e. no enrichment of either component on Jurkat cells expressing an

236   irrelevant TCR).

237

238   Receptor-Antigen Pairing by Targeted Retroviruses

239   With all the necessary tools for RAPTR validated, we assembled a library of 96

240   pMHC constructs consisting of known seroprevalent viral or vaccine antigens from the

241   IEDB[55], and further filtered for binding to HLA-A2 using NetMHC4.1[56] (**Table 1**). We co-

242   transduced each into 293T cells alongside a LeAPS vector encoding a barcode, pooled

243   the resulting cells proportionally, and sorted for cells expressing both barcode (via

244   mCherry) and pMHC (via GFP). We verified that antigen and barcode expression in our

245   packaging cell library was stable after at least 9 passages post-sort (**Supplementary**

246   **Figure 4A**). We next used the library to infect a cell line expressing the C7 TCR[57], which

247   is known to recognize the CMV peptide NLVPMVATV (known as NLV). Without any cell

248   sorting step required, we sequenced genomic DNA from transduced cells to enumerate

249   barcode frequencies (**Figure 4A**). We observed a large enrichment of NLV barcode in

250   transduced cells relative to the packaging cell line (**Figure 4B**), with NLV representing

251   15% of all reads in transduced cells and no notable enrichment of other sequences. We

252   further validated this approach in C7 replicates (**Figure 4C** and **Supplementary Figure**

253   **4B**), and in a separate screen of the JM22 TCR [57], which recognizes the GL9 peptide

254   from influenza A (**Figure 4D** and **Supplementary Figure 4C**). In all cases, we observed

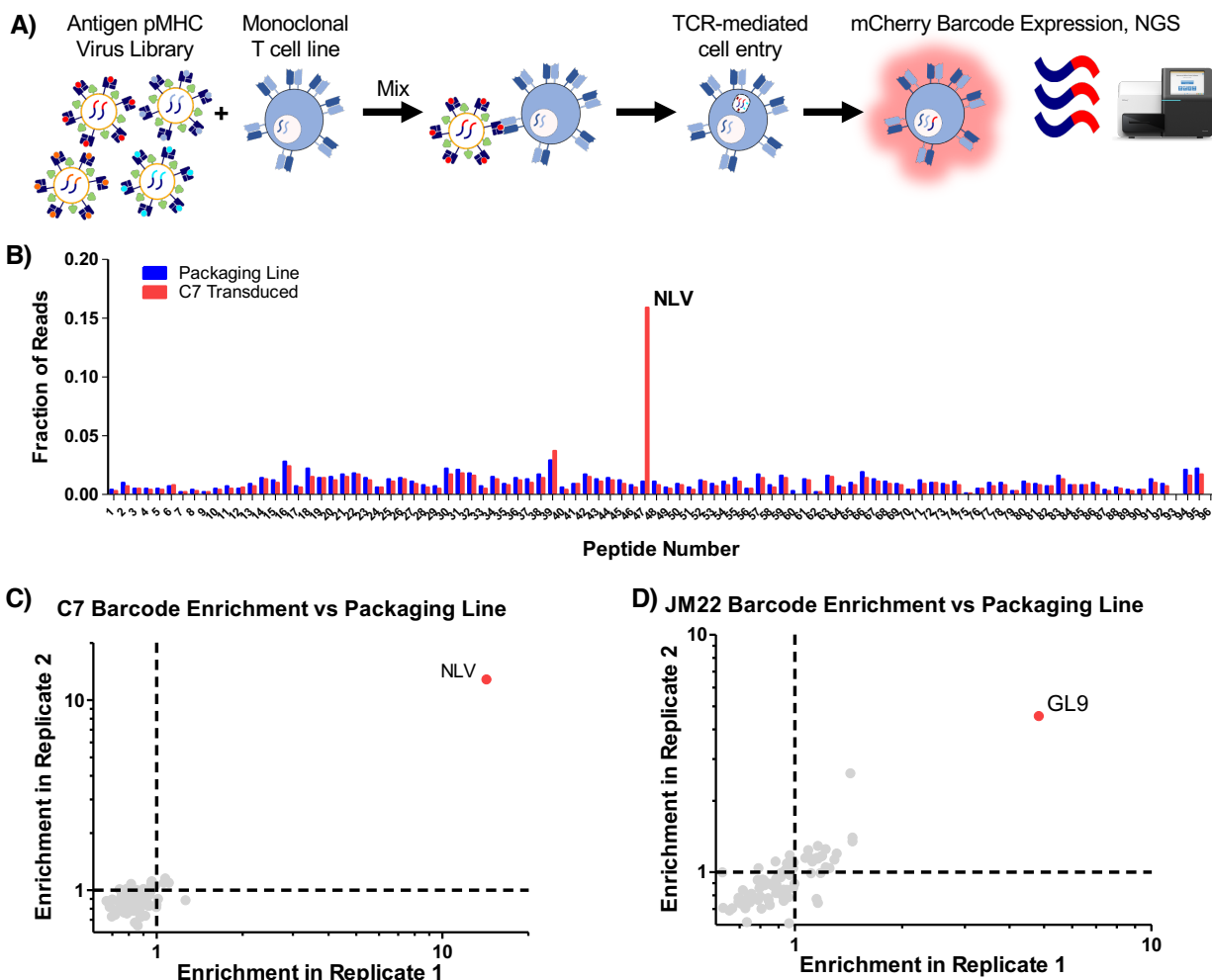255   robust enrichment of on-target, but not off-target, barcodes.

Figure 4 | RAPTR enables facile antigen identification for T cell receptors
A, Workflow for RAPTR on monoclonal cell lines. **B,** Read fraction of barcodes following infection of C7 TCR-transduced J76 cells. **C,** Comparison of antigen enrichment upon library infection of C7 TCR-transduced J76 cells across two additional replicates. **D,** Comparison of antigen enrichment upon library infection of JM22 TCR-transduced J76 cells across two replicates.

256   RAPTR can be adapted to BCRs

257   Many of the same constraints that limit TCR-pMHC interaction screens are also

258   shared by efforts to identify B cell antigens at scale. Most approaches rely on barcoded

259   multimers for very few targets, and the largest antigen library reported to date is 9

260   antigens[15]. As BCRs are also rapidly endocytosed upon antigen engagement[58,59], we

261   reasoned that we could use stabilized viral surface antigens or known immunogens as

262   targeting ligands for viral infection. We thus generated a "hybrid" pseudotype co-

263   expressing either the surface-tethered receptor binding domain (RBD) or the full-length,

264   pre-fusion stabilized spike (S2P) protein from SARS-CoV-2[60] alongside VSVGmut

265   (**Figure 5A**). We used each of these to achieve efficient, antigen-specific infection of a

266   Ramos cell line expressing CR3022, a B cell receptor clone that is cross-reactive for the

267   RBDs of SARS-CoV-1 and 2[61,62] (**Figure 5B**). Similarly, hybrid pseudotypes incorporating

268   HIV env CD4 binding site (CD4bs) constructs from diverse clades[63] each infected cell

269   lines expressing the bNab VRC01[64], and a known affinity-reducing mutation (D368R)

270   dramatically reduced, but did not fully abrogate, infection[65].

271   To determine the sensitivity for BCR-mediated cell entry, we mixed Ramos cells

272   expressing either the CR3022 BCR (IgM$^+$) or no BCR (IgM-negative), and tracked on-

273   target cells via surface IgM expression (**Figure 5C**). Using RBD-displaying viruses, we

274   observed specific infection at target cell frequencies at least as rare as 1×10^{-5}, but could

275   not calculate a meaningful selectivity ratio due to a lack of off-target infection. Next, to

276   assess the potential feasible library size, we mixed on-target RBD-displaying viruses

277   encoding GFP with off-target CD4bs-displaying viruses encoding mCherry. and used

278   them to infect CR3022 BCR-expressing Ramos cells (**Figure 5D**). Keeping the total virus

279   amount constant, we were able to robustly detect on-target interactions with minimal off-

280   target infection in mixtures as dilute as 1:1,000 on-target viruses. We observed similar

281   results when displaying SARS-CoV-2 S2P instead of RBD (**Supplementary Figure 5**).

282   These results gave us confidence in the feasibility of extracting meaningful signal from

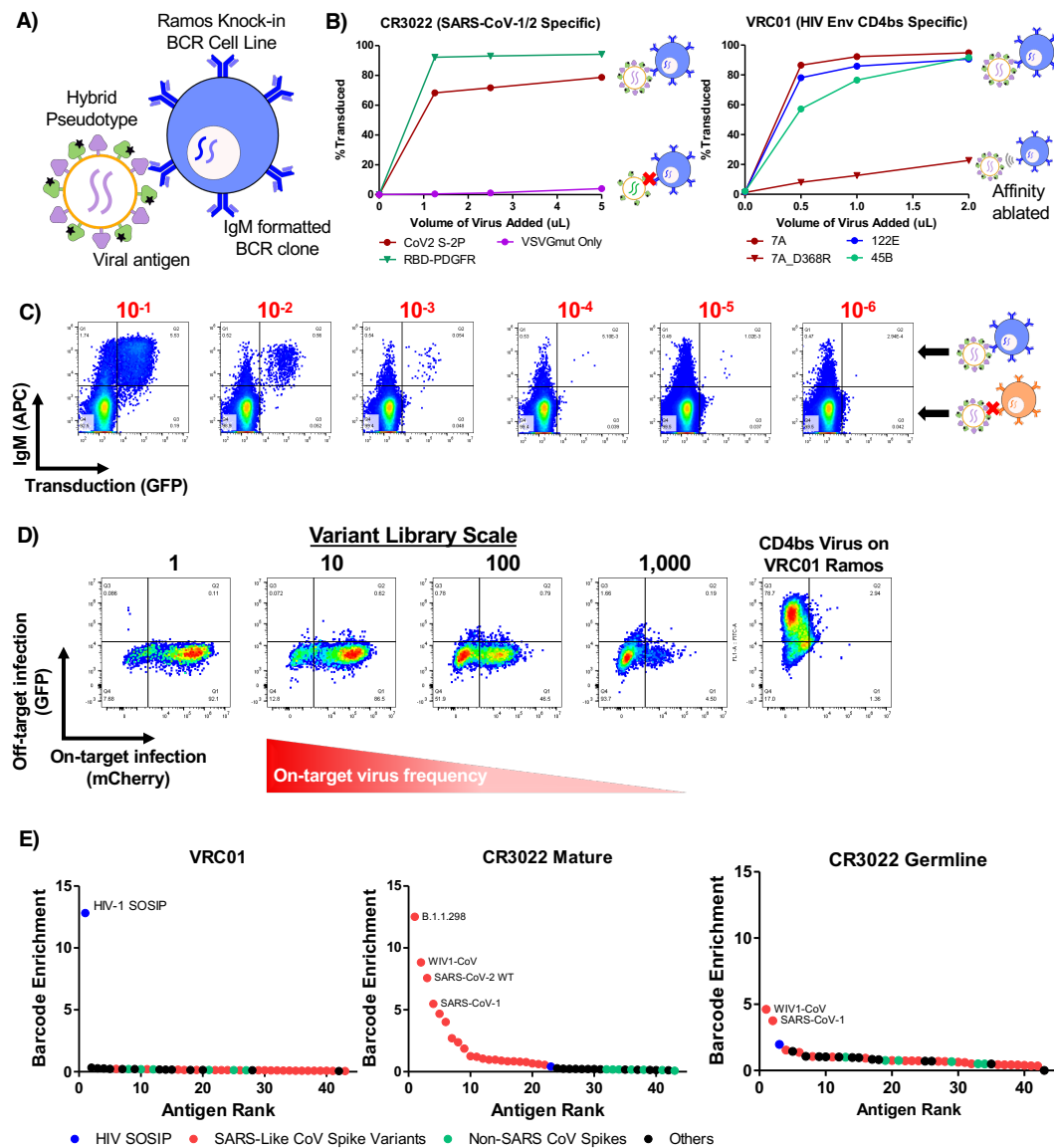283   interaction screens using lentiviral infection via the BCR.

Figure 5 | RAPTR enables profiling of B cell reactivity via infection of BCR knock-in cells
A, Overview of BCR-based targeting via the VSVGmut hybrid pseudotype in which a viral protein is co-expressed with VSVGmut. **B,** Targeting of Ramos cells expressing the SARS-CoV S protein specific CR3022 BCR via VSVGmut lentiviruses displaying either SARS-CoV-2 spike (2P) or surface-bound SARS-CoV-2 S protein RBD, or VRC01 BCR via env constructs representing multiple clades of HIV; representative data from 2 independent experiments. **C,** Infection of CR3022 BCR-expressing Ramos cells mixed into IgM-negative Ramos cells at target cell frequencies indicated in red, demonstrating selectivity in a mixed population of cells. **D,** Infection of CR3022 BCR-expressing Ramos cells with mixtures of SARS-CoV-2 RBD (on-target, expressing mCherry) and HIV env (off-target, expressing GFP) viruses; variant library size indicates the ratio of off-target to on-target virus present. Infection of VRC01-expressing Ramos cells with HIV env hybrid pseudotyped viruses (right) demonstrates that the viral particles remain functional. **E,** Enrichment of antigen barcodes following viral glycoprotein library infection of Ramos cells expressing BCRs for VRC01, mature CR3022, or the germline-reverted version CR3022 demonstrating selectivity even for antibody sequences before affinity maturation.

284   Next, we sought to establish the capabilities of RAPTR for BCR profiling in a true

285   library setting. We assembled a library composed of 21 prefusion-stabilized SARS-CoV-

286   2 spike variants and 22 additional vial surface glycoproteins or immunogens from diverse

287   sources as controls to enable proof of concept experiments (**Table 2**, **Supplementary**

288   **Table 2**). When we used this library to infect cells expressing the VRC01 BCR, we

289   observed clear enrichment of only HIV-1 SOSIP[66] (**Figure 5E**), highlighting the potential

290   for direct receptor deorphanization. When we infected cells expressing the CR3022 BCR,

291   we observed clear enrichment of SARS-CoV spikes, including multiple SARS-CoV-2

292   variants. Notably, a version of CR3022 that excluded the mutations acquired via somatic

293   hypermutation yielded enrichment of only SARS-CoV-1 spike and the closely related

294   WIV-1 spike, highlighting the ability of RAPTR to distinguish BCR cross-reactivity even

295   for naïve BCRs that have yet to undergo affinity maturation.

296   RAPTR scales to library-on-library screens

297   To fully realize the potential of RAPTR for antigen identification, we sought to apply

298   it to library vs library screening by using our 96-member pMHC viral library to pan a

299   previously reported library of >450,000 TCRs[67]. To simplify the first attempt, we pre-

300   enriched for potentially reactive cells using tetramers for CMV, EBV, and Flu (CEF)

301   antigens presented by HLA-A2, resulting in a polyclonal pool of TCRs with greater

302   prevalence of these specificities, similar to the routine process of pre-expanding cells

303   (**Figure 6A**). Following transduction and FACS to isolate transduced cells, we used bulk

304   sequencing to identify which antigens were enriched in aggregate. In line with tetramer

305   staining of the untransduced cells, we observed strong enrichment of Flu GL9, and more

306   modest enrichment of EBV GLC (**Figure 6B**, **Supplementary Figure 6**).
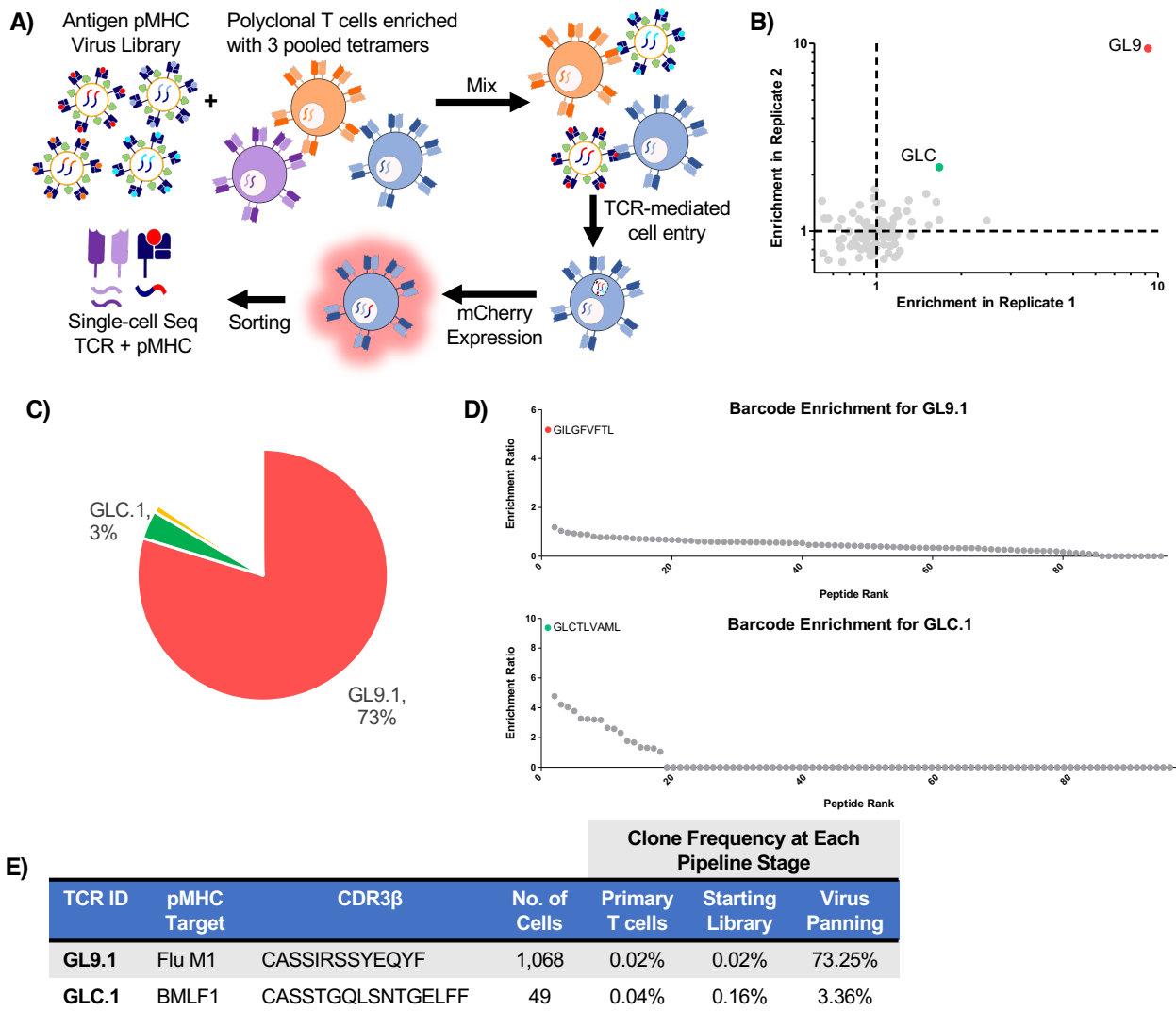
Figure 6 | RAPTR scales to polyclonal antigen identification in a single, scalable pipeline
A, Schematic overview of the library on library RAPTR paradigm for TCR antigen identification. **B,** Bulk sequencing of barcodes following infection of enriched CSS-930 TCR library with viral antigen library shows enrichment of immunodominant EBV and Influenza antigens in two independent infections. **C,** Clonal frequencies of enriched TCRs in cells analyzed by single cell sequencing; white space represents TCRs found in < 2 cells. **D,** Enrichment of antigen barcodes in the top two most abundant clones. **E,** Detailed tracking of TCR clones with matched antigens, including CDR3 sequence, number of cells analyzed, and clonal frequencies at each stage of the pipeline.

307   With this validation in hand, we performed single-cell RNA sequencing using the

308   10x Genomics Chromium 5' chemistry with V(D)J enrichment. We included a targeted

309   primer at the reverse transcription step and a custom amplification protocol to enable

310   recovery of our pMHC barcodes, akin to ECCITEseq and other similar methods[68,69]. In

311 this experiment, we recovered 1,458 cells, with TCR clonotypes dominated by a few

312 clones (**Figures 6C** and **6E**). We grouped cells of common clonotype together and

313 analyzed pMHC barcode expression to determine enrichment relative to packaging library

314 frequency. The most prevalent TCR clone, GL9.1, demonstrated strong enrichment of

315 GL9 pMHC signal (**Figure 6D**). Upon literature search, we noted that GL9.1 represents a

316 well-characterized public clonotype known to recognize GL9 peptide presented by HLA-

317 A2[70].

318   The next most abundant clone, GLC.1, was identified in a previous screen of this

319 same TCR library by recombinant HLA-A2 GLC dextramer[67]. Our screen also identified

320 GLC as the target pMHC (**Figure 6D**). Notably, this was the only clone previously

321 identified by GLC dextramer panning that was validated by monoclonal binding and

322 functional tests. Our approach did not identify any of the previously reported false

323 positives as hits, providing further evidence that RAPTR is an efficient means of

324 identifying T cell antigens that integrates both binding and signaling.

325

326 **Discussion**

327   RAPTR is a high-throughput platform for directly linking immune receptors with

328 their cognate antigens. It is based on the combination of several conceptual and

329 technological advances that enable (i) versatile and efficient targeted viral entry via the

330 VSVGmut pseudotyping system; (ii) a method for scalable, reproducible lentiviral library

331 packaging (LeAPS); and (iii) the use of viral entry as a means of screening for interactions.

332 We demonstrated a system for efficient, antigen-specific infection via both the TCR and

333 the BCR that integrates both binding and target cell activation. We then exploited this

334   mode of infection to match receptors with target antigens in complex mixtures and in a

335   library vs library format.

336   We demonstrated that the VSVGmut pseudotyping system is efficient, specific,

337   and uniquely modular by targeting a variety of cell surface proteins, including cytokine

338   receptors, costimulatory molecules, lineage markers, and both T and B cell receptors. We

339   report here as many different targeting approaches for VSVGmut as have been reported

340   for any other single pseudotyping system[27], including the first true antigen-specific

341   infection of lymphocytes and the first combinatorial targeting approach. In addition, the

342   VSVGmut system produces viruses at higher average titers than other systems while

343   maintaining a high degree of specificity. Whereas most pseudotyping systems have been

344   limited to the use of small, stable, high-affinity targeting ligands (e.g. scFvs or DARPins),

345   VSVGmut is uniquely versatile in ligand usage due to the modular design it permits. We

346   report targeting even at micromolar affinities, whereas the efficiency of infection by

347   paramyxoviral pseudotypes was reported to dramatically decrease in the low nanomolar

348   range[71].

349   Applications of RAPTR

350   We have demonstrated the utility of RAPTR for identifying antigens recognized by

351   TCRs and BCRs. For T cell reactivity, our approach could be directly applied to rapidly

352   deorphanize TCRs from existing single-cell studies, which routinely pair TCR sequences

353   with gene expression for cancer, autoimmunity, or infectious disease. Existing

354   techniques, including yeast display, SABRs, and T-Scan, require individual screens to

355   pair each TCR with its cognate antigen, including at least one cell sorting step per TCR.

356   In contrast, RAPTR can be performed on pools of TCRs in a single step, with only one

357   cell sorting step required (and no sorting required for single-TCR experiments). While

358   barcoded pMHC multimers can be built in to single-cell workflows, large antigen pools are

359   inherently difficult to produce and validate, and require equivalent effort for each new

360   batch. These factors have limited their use to only a few specialized labs thus far. RAPTR

361   library production is straightforward and far more scalable in comparison, requiring only

362   a simple transfection per batch following initial library assembly. In principle, these

363   reagents could be made and distributed at large scale, either via sharing the packaging

364   cell line or by leveraging current industrial infrastructure for production of lentiviruses.

365   For BCRs, we envision applying pools of protein variants, such as the one

366   presented here, to large libraries of BCRs. Such an approach would greatly accelerate

367   isolation and optimization of cross-reactive mAbs to a wide range of target classes.

368   Current techniques require the recombinant expression of individual mAbs followed by

369   ELISA for each target of interest, which represents a major bottleneck to experimental

370   scale and scope[72]. Using RAPTR with BCR knock-in cells, which could themselves be

371   prepared as pooled lentiviral libraries[62], could alleviate this bottleneck. For example, even

372   scales smaller than the one we presented would be sufficient to represent every subtype

373   of influenza hemagglutinin and neuraminidase, multiple representatives of each of the

374   major clades of HIV-1, and more. Larger libraries on the scale of 1000s of variants, whose

375   feasibility we demonstrated, could enable detailed antigenic site or epitope mapping for

376   many mAbs at once, a process currently limited to structural studies that are difficult to

377   scale.

378   Finally, the demonstrated modularity of the system opens the door to many

379   potential applications. Since our approach only requires 293T cell surface expression of

380   the targeting molecule, we hypothesize that RAPTR can be readily adapted to other class

381   I MHC alleles, class II MHC alleles, as well as non-human systems for vaccine and

382   immunotherapy development. This compares favorably to recombinant MHC expression

383   or yeast display, each of which can require allele-specific optimizations and/or mutations

384   to ensure proper folding[8,73,74]. However, our system is not restricted to immunological

385   applications. In principle, it is applicable to any receptor-ligand system in which one

386   binding partner can be displayed on a lentivirus and the other can be expressed as an

387   endocytic receptor on cells. This may enable large-scale interactome and co-evolution

388   studies that are not readily achievable by existing techniques. Furthermore, the VSVGmut

389   pseudotyping system may be useful as a new platform for the development of efficient,

390   specific gene therapies.

391   <u>Limitations</u>

392   As with any technology, RAPTR has several limitations. Our proof of concept

393   studies were performed on ~100-antigen scale, and while our control experiments

394   indicate this could be readily increased to 1,000s, these scales will still require pre-

395   selection of antigen targets for analysis. Future work will focus on further streamlining

396   library assembly and increasing transduction efficiency toward larger-scale assays. In

397   addition, the single-cell sequencing step must be performed in sufficient depth to facilitate

398   acceptable signal to noise determinations. While we chose to use a commercially

399   available platform to ensure broad applicability, recent and future advances in the scale

400   of single-cell analysis[75] will improve the utility of RAPTR regardless of platform. Like other

401   genetically encoded techniques, RAPTR also requires prior selection of MHC

402   haplotype(s) to generate libraries. Finally, the antigen identification studies we presented

403 were on cell lines with receptors knocked in, rather than primary cells. While the

404 decreasing cost and turnaround time of gene synthesis will make such resources

405 increasingly available, future studies can apply RAPTR directly to primary cells.

406 Altogether, RAPTR is a versatile platform for high-throughput interaction screens

407 that uniquely incorporates diversity of both receptors and ligands in a single assay. The

408 resulting tools will enable detailed studies of antigen recognition for understanding and

409 engineering the adaptive immune response, as well as broader interactome studies.

410

411 **Methods**

412 *Media and Cells*

413 HEK293T cells (ATCC CRL-11268) were cultured in DMEM (ATCC) supplemented

414 with 10% fetal bovine serum (FBS, Atlanta Biologics) and penicillin-streptomycin

415 (Pen/Strep, Gibco).

416 Jurkat (ATCC TIB-152), J76 cells, and Ramos cells (ATCC CRL-1596) were

417 cultured in RPMI-1640 (ATCC) supplemented with 10% FBS and Pen/Strep. J76 cells[76]

418 (PMID 29707134) were a gift from Mirjam Heemskerk and Mark Davis.

419 CSS-930 TCR library cells were a gift from David Johnson[67], and were cultured in

420 RPMI-1640 (ThermoFisher) supplemented with 10% FBS, Pen/Strep, 1x non-essential

421 amino acids (ThermoFisher), 2mM GlutaMAX (ThermoFisher), and 1mM sodium pyruvate

422 (ThermoFisher).

423 *Plasmid Construction*

424 The plasmid pHIV-EGFP was gifted by Bryan Welm & Zena Werb (Addgene

425 plasmid #21373) and pMD2.G and psPAX2 were gifted by Didier Trono (Addgene plasmid

426   #12259 and #12260). pLentiCRISPR v2 was a gift from Feng Zhang (Addgene plasmid #

427   52961). IL-13 (Uniprot ID P35225) residues 35-146 were cloned into the pHIV backbone

428   containing the Igκ leader peptide, the PDGFR stalk and transmembrane domain (Uniprot

429   P09619 residues 449-497) with extracellular linkers listed in Supplementary Figure 1B.

430   The anti-CD19 scFv FMC63 and was cloned into the IgG4 hinge-PDGFR display format

431   in the pMD2 backbone. The anti-CD3 Fab UCHT1 was cloned into the PDGFR stalk only

432   format. Human CD80 (Uniprot P33681, residues 1-273) was cloned into the pMD2

433   backbone. pMHC single-chain trimers[42] were cloned into either the pMD2 backbone for

434   individual infections or the pHIV backbone for library construction. To generate the

435   pLeAPS backbone, the CMV core promoter was cloned into the pLenti backbone between

436   the polypurine tract and the 3' LTR, analogous to previous work[52]. For individual

437   infections, SARS-CoV-2 RBD or HIV env CD4bs constructs were cloned into the pMD2

438   backbone on the PDGFR stalk only display architecture. Prefusion-stabilized SARS-CoV-

439   2 spike (2P) was cloned into the pMD2 backbone for individual infections. For library

440   assembly, viral constructs described in **Supplementary Table 2** were cloned into the

441   pLenti backbone, C-terminally fused to GFP via a P2A motif. For pLeAPS barcodes,

442   mCherry with an 8-nt degenerate sequence was cloned into the pLeAPS backbone

443   downstream of the Ef1α core promoter.

444   *Transfection for Lentiviral Production*

445   Lentiviruses were prepared by transient transfection of HEK293T cells with linear

446   25 kDa polyethylenimine (PEI, Santa Cruz Biotechnology) at a 3:1 mass ratio of PEI to

447   DNA. Briefly, DNA and PEI were diluted in Opti-MEM (ThermoFisher) and mixed to form

448   complexes. Complex formation was allowed to proceed for 15 minutes at room

449 temperature before dropwise addition to cells. Media was changed to complete DMEM +

450 25 mM HEPES after 3-6 hours.

451 For individually targeted or VSVGwt viruses, plasmid mass ratios were 5.6:3:3:1

452 for transfer plasmid to psPAX2.1 to targeting plasmid (when used) to fusogen plasmid

453 (either VSVGmut or VSVGwt). Targeting plasmids contain expression cassettes for

454 virally-displayed ligands in the pMD2 backbone. Total plasmid amounts are indicated in

455 the table below. For LeAPS-based virus production, packaging cells were transfected with

456 a 3:1 mass ratio of psPAX2.1:fusogen.

457

Production Scale (per well)	Total Plasmid Amount (µg)		
	Non-Targeted	Targeted	LeAPS-based
6-well plate	2.4	3.15	1
10 cm dish	24	31.5	10
T225	72	94.5	30

458

459 *Viral Purification*

460 Unconcentrated viruses were filtered (0.45µm PES) and used directly. If needed,

461 they were stored at 4°C for up to 2 weeks or at -80°C indefinitely. Concentrated viruses

462 were filtered (0.45µm PES) and concentrated 200x by ultracentrifugation for 90 minutes

463 at 100,000xg at 4°C. The supernatant was discarded and viral pellets were resuspended

464 in Opti-MEM overnight at 4°C.

465

466 *Single lentiviral infections*

467 Individual infections were carried out with the indicated amounts of virus and cells,

468 in the presence of 8 µg/mL of either polybrene or diethylaminoethyl (DEAE)-dextran

469   (Sigma-Aldrich). After 24h, an additional 1x volume of media was added. Cells were

470   analyzed by flow cytometry 48-72h post-infection for Jurkat or J76 infections, and 24-48h

471   post-infection for Ramos cell infections. If cells were only assessed for viral infection, cells

472   were washed once in FACS buffer (PBS + 0.1% BSA and 1 mM EDTA) before analysis

473   on an Accuri C6 or Cytoflex S flow cytometer. In experiments where an additional cell

474   lineage or activation marker such as CD25 or CD69 were assessed, cells were washed

475   once in FACS buffer, stained for 10 minutes in FACS buffer containing a marker-specific

476   antibody, then washed 2x with FACS buffer before analysis for infection (GFP or

477   mCherry) via flow cytometry.

478   *Human Primary T cell activation and transduction*

479   Peripheral blood mononuclear cells from healthy donors were purified from leukopaks

480   purchased from Stem Cell Technologies using Ficoll-Paque PLUS (GE Healthcare)

481   density gradient centrifugation with SepMate tubes (Stem Cell Technologies) per

482   manufacturer instructions. Primary CD8$^+$ T cells were isolated using EasySep Human

483   CD8$^+$ T Cell Enrichment Kits (Stem Cell Technologies) and cultured in RPMI-1640

484   (ATCC) supplemented with 10% fetal bovine serum, 100 U/ml penicillin-streptomycin

485   (Corning), and 30 IU/ml recombinant human IL-2 (R&D Systems). Prior to transduction

486   with VSVGwt viruses, T cells were activated using a 1:1 ratio of DynaBeads Human T-

487   Activator CD3/CD28 (Thermo Fisher) for 24 hours, after which 8 μg/mL of polybrene

488   (Santa Cruz Biotechnology) and concentrated lentivirus were added to culture at a

489   multiplicity of infection of 1. For targeted viruses, the same protocol was used, but

490   DynaBeads were omitted.

491   *Antigen receptor cell line generation*

492   Lentiviral TCR cassettes were formatted as TCRβ-P2A-TCRα and cloned into the

493   pHIV backbone. TCR KO J76 cells were transduced as described above and sorted

494   based on TCR expression to establish monoclonal cell lines. Ramos BCR cells were

495   established according to the published protocol[62].

496

497   *Lentiviral infections in Mixed Cell Populations*

498   Cells were labeled with CellTrace dyes (ThermoFisher) according to the

499   manufacturer's instructions, counted, and mixed at the indicated ratios. After labeling,

500   viral infections were carried out as indicated above. After 48 hours, cells were washed in

501   FACS buffer and analyzed via flow cytometry to examine infection in CellTrace[+] vs

502   CellTrace[-] cells. After flow cytometry, the selectivity was calculated as follows:

503    $$\text{Selectivity Ratio} = \frac{\%\ \text{target cells transduced}}{\%\ \text{off-target cells transduced}} * (\text{frequency of target cells})$$

504

505   *Generation of LeAPS libraries*

506   To generate LeAPS production cell lines, 293T cells were transduced as outlined

507   above while being seeded at 20% confluency on either 6-well or 96-well plates

508   (ThermoFisher). For 6-well plates, 500 μL each of unconcentrated, VSVGwt-pseudotyped

509   LeAPS barcode and ligand expression viruses were used for infection. 50 μL of each virus

510   was used for 96-well plates.

511   For libraries, cells were transduced in duplicate. One of the duplicates was

512   assessed by flow cytometry to determine the proportion of cells transduced with both

513   ligand (GFP[+]) and LeAPS barcode (mCherry[+]). Cells from the remaining duplicate were

514   then pooled, with each library member normalized to include an equivalent number of

515 ligand-expressing (GFP⁺) and LeAPS barcode-transduced (mCherry⁺) cells for each

516 library member. This pool of cells was then sorted based on GFP and mCherry expression

517 to make library packaging cell pools. To produce lentiviruses, library packaging cells were

518 transfected as described above, but using only psPAX2.1 and pMD2-VSVGmut. For

519 library creation, both transfer plasmids and the pMD2 targeting plasmid were omitted, as

520 they were replaced by the LeAPS packaging line. After 48 hours, virus was then collected

521 and concentrated as described above.

522

523 *Library screening of monoclonal TCR lines*

524 10 µL of concentrated pMHC library virus was used to transduce one million TCR-

525 expressing J76 cells as outlined above. After infection was confirmed by flow cytometry

526 as described above, genomic DNA was isolated using the PureLink Genomic DNA kit

527 (ThermoFisher). Barcode inserts were then amplified via 25 cycles of PCR and submitted

528 for Amplicon-EZ analysis by Genewiz. Enrichment was calculated for each barcode as

529 the fraction of total barcode-containing reads divided by the barcode frequency in the

530 packaging cells.

531

532 *Library screening of monoclonal Ramos BCR lines*

533 500 µL of unconcentrated viral antigen library was used to transduce one million

534 BCR-expressing Ramos cells as outlined above. After infection was confirmed by flow

535 cytometry, genomic DNA was isolated using the PureLink Genomic DNA kit. Barcode

536 inserts were then amplified via 25 cycles of PCR and submitted for Amplicon-EZ analysis

537 by Genewiz or analyzed via a MiSeq v3 150x150nt PE Nano kit (MIT BioMicroCenter).

538   Enrichment was calculated for each barcode as the fraction of total barcode-containing

539   reads divided by the barcode frequency in the packaging cells.

540

541   *Tetramer enrichment of TCR libraries*

542   To enrich a polyclonal pool of cells from the CSS-930 TCR library for known

543   specificities, 50 million library cells were co-stained with an anti-TRBC antibody (clone

544   IP26) and a pool of HLA-A2 tetramers (NIH Tetramer Core) presenting the following

545   peptides: NLVPMVATV (NLV), GILGFVFTL (GL9), and GLCTLVAML (GLC), all at 1:800

546   dilutions. Sorting for double-positive cells was performed on a Sony MA900 FACS.

547

548   *Library vs library TCR-pMHC screen*

549   10 µL of concentrated pMHC library virus was used to transduce two million

550   tetramer-enriched cells. Cells were sorted for infection based on mCherry expression,

551   and then submitted for analysis using the 10x Genomics Chromium 5' v2 V(D)J kit with a

552   barcode construct-specific primer spiked in prior to droplet encapsulation (see

553   Supplementary Methods for further details). TCR amplicons were prepared and

554   sequenced according to the manufacturer's instructions. Following cDNA amplification,

555   mCherry barcodes were enriched via separate PCRs and sequenced on an Illumina

556   MiSeq (150x150nt paired end reads). CellRanger was used to assign TCR clone identities

557   for each cell. Cell barcodes were used to match TCRs with associated pMHC barcodes

558   counts.

559

560   **Data Availability**

561   The next-generation sequencing datasets generated during and/or analyzed during the

562   current study are available in the NCBI sequence read archive (accession numbers to be

563   assigned before publication). All data generated or analyzed during this study are

564   included in this published article, the sequence read archive, and its supplementary

565   information files.

566

Code Availability

567

568   A sample R workspace for extracting barcode counts based on CITEseq[68] will be posted

569   to Birnbaum lab code repository upon publication.

References

571
572   1. Joglekar, A. V. & Li, G. T cell antigen discovery. *Nat. Methods* **18**, 873–880 (2021).
573   2. McCutcheon, M. *et al.* A sensitive ELISPOT assay to detect low-frequency human T
574   lymphocytes. *J. Immunol. Methods* **210**, 149–166 (1997).
575   3. Kula, T. *et al.* T-Scan: A Genome-wide Method for the Systematic Discovery of T Cell
576   Epitopes. *Cell* **178**, 1016-1028.e13 (2019).
577   4. Joglekar, A. V. *et al.* T cell antigen discovery via signaling and antigen-presenting
578   bifunctional receptors. *Nat. Methods* **16**, 191–198 (2019).
579   5. Kisielow, J., Obermair, F.-J. & Kopf, M. Deciphering CD4+ T cell specificity using
580   novel MHC–TCR chimeric receptors. *Nat. Immunol.* **20**, 652–662 (2019).
581   6. Robins, H. S. *et al.* Overlap and effective size of the human CD8+ T cell receptor
582   repertoire. *Sci. Transl. Med.* **2**, 47ra64 (2010).
583   7. Arstila, T. P. *et al.* A direct estimate of the human alphabeta T cell receptor diversity.
584   *Science* **286**, 958–961 (1999).
585   8. Birnbaum, M. E. *et al.* Deconstructing the peptide-MHC specificity of T cell
586   recognition. *Cell* **157**, 1073–1087 (2014).
587   9. Gee, M. H. *et al.* Antigen Identification for Orphan T Cell Receptors Expressed on
588   Tumor-Infiltrating Lymphocytes. *Cell* **172**, 549-563.e16 (2018).
589   10. Bentzen, A. K. *et al.* Large-scale detection of antigen-specific T cells using peptide-
590   MHC-I multimers labeled with DNA barcodes. *Nat. Biotechnol.* **34**, 1037–1045 (2016).
591   11. Bentzen, A. K. *et al.* T cell receptor fingerprinting enables in-depth characterization
592   of the interactions governing recognition of peptide-MHC complexes. *Nat. Biotechnol.*
593   (2018) doi:10.1038/nbt.4303.
594   12. Overall, S. A. *et al.* High throughput pMHC-I tetramer library production using
595   chaperone-mediated peptide exchange. *Nat. Commun.* **11**, 1909 (2020).
596   13. Zhang, S.-Q. *et al.* High-throughput determination of the antigen specificities of T cell
597   receptors in single cells. *Nat. Biotechnol.* **36**, 1156–1159 (2018).

14. Shiakolas, A. R. *et al.* Efficient discovery of potently neutralizing SARS-CoV-2 antibodies using LIBRA-seq with ligand blocking. *bioRxivorg* 2021.06.02.446813 (2021) doi:10.1101/2021.06.02.446813.

15. Setliff, I. *et al.* High-throughput mapping of B cell receptor sequences to antigen specificity. *Cell* **179**, 1636-1646.e15 (2019).

16. Fields, S. & Song, O. A novel genetic system to detect protein-protein interactions. *Nature* **340**, 245–246 (1989).

17. Wojtowicz, W. M. *et al.* A human IgSF cell-surface interactome reveals a complex network of protein-protein interactions. *Cell* **182**, 1027-1043.e17 (2020).

18. Verschueren, E. *et al.* The immunoglobulin superfamily receptome defines cancer-relevant networks associated with clinical outcome. *Cell* **182**, 329-344.e19 (2020).

19. Özkan, E. *et al.* An extracellular interactome of immunoglobulin and LRR proteins reveals receptor-ligand networks. *Cell* **154**, 228–239 (2013).

20. Havugimana, P. C. *et al.* A census of human soluble protein complexes. *Cell* **150**, 1068–1081 (2012).

21. Babu, M. *et al.* Interaction landscape of membrane-protein complexes in Saccharomyces cerevisiae. *Nature* **489**, 585–589 (2012).

22. Younger, D., Berger, S., Baker, D. & Klavins, E. High-throughput characterization of protein-protein interactions by reprogramming yeast mating. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 12166–12171 (2017).

23. Gu, L. *et al.* Multiplex single-molecule interaction profiling of DNA-barcoded proteins. *Nature* **515**, 554–557 (2014).

24. Buchholz, C. J., Duerner, L. J., Funke, S. & Schneider, I. C. Retroviral display and high throughput screening. *Comb. Chem. High Throughput Screen.* **11**, 99–110 (2008).

25. Buchholz, C. J. *et al.* In vivo selection of protease cleavage sites from retrovirus display libraries. *Nat. Biotechnol.* **16**, 951–954 (1998).

26. Schneider, R. M. *et al.* Directed evolution of retroviruses activatable by tumour-associated matrix metalloproteases. *Gene Ther.* **10**, 1370–1380 (2003).

27. Frank, A. M. & Buchholz, C. J. Surface-Engineered Lentiviral Vectors for Selective Gene Transfer into Subtypes of Lymphocytes. *Molecular therapy. Methods & clinical development* **12**, 19–31 (2019).

28. Naldini, L. *et al.* In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science* **272**, 263–267 (1996).

29. Lei, Y., Joo, K.-I. & Wang, P. Engineering fusogenic molecules to achieve targeted transduction of enveloped lentiviral vectors. *J. Biol. Eng.* **3**, 8 (2009).

30. Yang, H., Joo, K.-I., Ziegler, L. & Wang, P. Cell type-specific targeting with surface-engineered lentiviral vectors co-displaying OKT3 antibody and fusogenic molecule. *Pharm. Res.* **26**, 1432–1445 (2009).

31. Yang, L., Bailey, L., Baltimore, D. & Wang, P. Targeting lentiviral vectors to specific cell types in vivo. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 11479–11484 (2006).

32. Frecha, C. *et al.* Stable transduction of quiescent T cells without induction of cycle progression by a novel lentiviral vector pseudotyped with measles virus glycoproteins. *Blood* **112**, 4843–4852 (2008).

33. Ou, W. *et al.* Specific targeting of human interleukin (IL)-13 receptor α2-positive cells with lentiviral vectors displaying IL-13. *Hum. Gene Ther. Methods* **23**, 137–147 (2012).

34. Funke, S. *et al.* Targeted cell entry of lentiviral vectors. *Mol. Ther.* **16**, 1427–1436 (2008).

35. Bender, R. R. *et al.* Receptor-Targeted Nipah Virus Glycoproteins Improve Cell-Type Selective Gene Delivery and Reveal a Preference for Membrane-Proximal Cell Attachment. *PLoS Pathog.* **12**, e1005641 (2016).

36. Pfeiffer, A. *et al.* In vivo generation of human CD19-CAR T cells results in B-cell depletion and signs of cytokine release syndrome. *EMBO Mol. Med.* **10**, (2018).

37. Agarwal, S. *et al.* In vivo generation of CAR T cells selectively in human CD4+ lymphocytes. *Mol. Ther.* **28**, 1783–1794 (2020).

38. Nikolic, J. *et al.* Structural basis for the recognition of LDL-receptor family members by VSV glycoprotein. *Nat. Commun.* **9**, 1029 (2018).

39. Rossjohn, J. *et al.* T cell antigen receptor recognition of antigen-presenting molecules. *Annu. Rev. Immunol.* **33**, 169–200 (2015).

40. Zhao, Y. *et al.* High-affinity TCRs generated by phage display provide CD4+ T cells with the ability to recognize and kill tumor cell lines. *The Journal of Immunology* **179**, 5845–5854 (2007).

41. Li, Y. *et al.* Directed evolution of human T-cell receptors with picomolar affinities by phage display. *Nat. Biotechnol.* **23**, 349–354 (2005).

42. Yu, Y. Y. L., Netuschil, N., Lybarger, L., Connolly, J. M. & Hansen, T. H. Cutting edge: single-chain trimers of MHC class I molecules form stable structures that potently stimulate antigen-specific T cells and B cells. *The Journal of Immunology* **168**, 3145–3149 (2002).

43. Matlin, K. S., Reggio, H., Helenius, A. & Simons, K. Pathway of vesicular stomatitis virus entry leading to infection. *J. Mol. Biol.* **156**, 609–631 (1982).

44. Alcover, A. & Alarcón, B. Internalization and intracellular fate of TCR-CD3 complexes. *Crit. Rev. Immunol.* **20**, 325–346 (2000).

45. Valitutti, S., Müller, S., Salio, M. & Lanzavecchia, A. Degradation of T cell receptor (TCR)-CD3-zeta complexes after antigenic stimulation. *J. Exp. Med.* **185**, 1859–1864 (1997).

46. Liu, H., Rhodes, M., Wiest, D. L. & Vignali, D. A. On the dynamics of TCR:CD3 complex cell surface expression and downmodulation. *Immunity* **13**, 665–675 (2000).

47. Blake, S., Hughes, T. P., Mayrhofer, G. & Lyons, A. B. The Src/ABL kinase inhibitor dasatinib (BMS-354825) inhibits function of normal human T-lymphocytes in vitro. *Clin. Immunol.* **127**, 330–339 (2008).

48. Xie, S., Cooley, A., Armendariz, D., Zhou, P. & Hon, G. C. Frequent sgRNA-barcode recombination in single-cell perturbation assays. *PLoS One* **13**, e0198635 (2018).

49. Feldman, D., Singh, A., Garrity, A. J. & Blainey, P. C. Lentiviral co-packaging mitigates the effects of intermolecular recombination and multiple integrations in pooled genetic screens. *bioRxiv* (2018) doi:10.1101/262121.

50. Hu, W.-S. & Hughes, S. H. HIV-1 reverse transcription. *Cold Spring Harb. Perspect. Med.* **2**, a006882–a006882 (2012).

51. OhAinle, M. *et al.* A virus-packageable CRISPR screen identifies host factors mediating interferon inhibition of HIV. *Elife* **7**, (2018).

52. Datlinger, P. *et al.* Pooled CRISPR screening with single-cell transcriptome readout. *Nat. Methods* **14**, 297–301 (2017).

53. Broussau, S. *et al.* Inducible packaging cells for large-scale production of lentiviral vectors in serum-free suspension culture. *Mol. Ther.* **16**, 500–507 (2008).

54. Hill, A. J. *et al.* On the design of CRISPR-based single-cell molecular screens. *Nat. Methods* **15**, 271–274 (2018).

55. Vita, R. *et al.* The Immune Epitope Database (IEDB): 2018 update. *Nucleic Acids Res.* **47**, D339–D343 (2019).

56. Jurtz, V. *et al.* NetMHCpan-4.0: Improved peptide-MHC class I interaction predictions integrating eluted ligand and peptide binding affinity data. *J. Immunol.* **199**, 3360–3368 (2017).

57. Yang, X. *et al.* Structural basis for clonal diversity of the public T cell response to a dominant human Cytomegalovirus Epitope. *J. Biol. Chem.* **290**, 29106–29119 (2015).

58. Malhotra, S., Kovats, S., Zhang, W. & Coggeshall, K. M. B cell antigen receptor endocytosis and antigen presentation to T cells require Vav and dynamin. *J. Biol. Chem.* **284**, 24088–24097 (2009).

59. Dougan, S. K. *et al.* Antigen-specific B-cell receptor sensitizes B cells to infection by influenza virus. *Nature* **503**, 406–409 (2013).

60. Hsieh, C.-L. *et al.* Structure-based design of prefusion-stabilized SARS-CoV-2 spikes. *Science* **369**, 1501–1505 (2020).

61. Yuan, M. *et al.* A highly conserved cryptic epitope in the receptor binding domains of SARS-CoV-2 and SARS-CoV. *Science* **368**, 630–633 (2020).

62. Weaver, G. C. *et al.* In vitro reconstitution of B cell receptor-antigen interactions to evaluate potential vaccine candidates. *Nat. Protoc.* **11**, 193–213 (2016).

63. Zhou, T. *et al.* Structural repertoire of HIV-1-neutralizing antibodies targeting the CD4 supersite in 14 donors. *Cell* **161**, 1280–1292 (2015).

64. Wu, X. *et al.* Rational design of envelope identifies broadly neutralizing human monoclonal antibodies to HIV-1. *Science* **329**, 856–861 (2010).

65. Li, Y. *et al.* HIV-1 neutralizing antibodies display dual recognition of the primary and coreceptor binding sites and preferential binding to fully cleaved envelope glycoproteins. *J. Virol.* **86**, 11231–11241 (2012).

66. Sliepen, K. *et al.* Structure and immunogenicity of a stabilized HIV-1 envelope trimer based on a group-M consensus sequence. *Nat. Commun.* **10**, 2355 (2019).

67. Spindler, M. J. *et al.* Massively parallel interrogation and mining of natively paired human TCRαβ repertoires. *Nat. Biotechnol.* **38**, 609–619 (2020).

68. Mimitou, E. P. *et al.* Multiplexed detection of proteins, transcriptomes, clonotypes and CRISPR perturbations in single cells. *Nat. Methods* **16**, 409–412 (2019).

69. Replogle, J. M. *et al.* Combinatorial single-cell CRISPR screens by direct guide RNA capture and targeted sequencing. *Nat. Biotechnol.* **38**, 954–961 (2020).

70. Sant, S. *et al.* Single-cell approach to influenza-specific CD8+ T cell receptor repertoires across different age groups, tissues, and following influenza virus infection. *Front. Immunol.* **9**, 1453 (2018).

71. Hasegawa, K. *et al.* Affinity thresholds for membrane fusion triggering by viral glycoproteins. *J. Virol.* **81**, 13149–13157 (2007).

72. Utset, H. A., Guthmiller, J. J. & Wilson, P. C. Bridging the B cell gap: Novel technologies to study antigen-specific human B cell responses. *Vaccines (Basel)* **9**, (2021).

73. Adams, J. J. *et al.* T cell receptor signaling is limited by docking geometry to peptide-major histocompatibility complex. *Immunity* **35**, 681–693 (2011).

74. Altman, J. D. & Davis, M. M. MHC-peptide tetramers to visualize antigen-specific T cells. *Curr. Protoc. Immunol.* **115**, 17.3.1-17.3.44 (2016).

75. Datlinger, P. *et al.* Ultra-high throughput single-cell RNA sequencing by combinatorial fluidic indexing. *bioRxiv* (2019) doi:10.1101/2019.12.17.879304.

76. Rosskopf, S. *et al.* A Jurkat 76 based triple parameter reporter system to evaluate TCR functions and adoptive T cell strategies. *Oncotarget* **9**, 17608–17619 (2018).

Acknowledgments

767

Author contributions

769   Conception of project: C.S.D., M.E.B. Conducting experiments: C.S.D., A.R., S.G.,

770   B.E.S., E.J.K., J.D., L.R., V.O. Data analysis: C.S.D., A.R., S.G., B.E.S., E.J.K., J.D.

771   Supervision: D.L., M.D., S.K.D., M.E.B. Writing manuscript: C.S.D., M.E.B. Editing

772   manuscript: all authors.

773

Competing interests

775   The lentiviral targeting approach in this manuscript is the subject of US patent applications

776   with M.E.B., S.G., and C.S.D. as co-inventors. M.E.B. is a founder, consultant, and equity

777   holder of Viralogic Therapeutics and Abata Therapeutics. M.E.B. received research

778   funding from Pfizer, Inc. that partially funded this work. S.K.D. received unrelated

779   research funding from Novartis Pharmaceuticals, Eli Lilly and Company, and Bristol-

780   Myers Squibb, and is a founder, science advisory board member and equity holder in

781   Kojin. MD is a science advisory board member for Neoleukin.

782

783 **Table 1: Summary of HLA-A2-restricted viral antigens included in RAPTR T cell**

784 **library**

Antigen Source	# of Epitopes
Adenovirus 5	3
Herpes simplex virus 1 (HSV-1)	31
Herpes simplex virus 2 (HSV-2)	4
Varicella Zoster Virus (VZV)	4
Cytomegalovirus (CMV)	6
Coronavirus 229E	2
Coronavirus OC43	1
Epstein-Barr Virus (EBV)	26
Influenza A (IAV)	6
Measles	4
Mumps	6
Kaposi's sarcoma-associated virus (HHV-8)	3

785

786 **Table 2: Summary of viral surface glycoproteins for RAPTR B cell library**

Antigen Source	# of Variants
SARS-CoV-2 Spike	21
Other MERS- and SARS-like CoV Spike	4
Common Cold CoV Spike	4
SARS-CoV-2 E/M	2
Influenza	4
Seroprevalent or Childhood Vaccine Viruses (Positive Controls) Measles, RSV, CMV, EBV	5
Emerging or Uncommon in Healthy Donors (Negative Controls) HIV, Dengue, Nipah Virus	3

787