

# Deciphering Membrane-Protein Interactions and High-Throughput Antigen Identification with Cell Doublets

Yuqian Wang, Zhe Wang, Juan Yang, Xiaobo Lei, Yisu Liu, Luke Frankiw, Jianwei Wang,\* and Guideng Li\*

Deciphering cellular interactions is essential to both understand the mechanisms underlying a broad range of human diseases, but also to manipulate therapies targeting these diseases. Here, the formation of cell doublets resulting from specific membrane ligand-receptor interactions is discovered. Based on this phenomenon, the study developed DoubletSeeker, a novel high-throughput method for the reliable identification of ligand-receptor interactions. The study shows that DoubletSeeker can accurately identify T cell receptor (TCR)-antigen interactions with high sensitivity and specificity. Notably, DoubletSeeker effectively captured paired TCR-peptide major histocompatibility complex (pMHC) information during a highly complex library-on-library screening and successfully identified three mutant TCRs that specifically recognize the MART-1 epitope. In turn, DoubletSeeker can act as an antigen discovery platform that allows for the development of novel immunotherapy targets, making it valuable for investigating fundamental tumor immunology.

## 1. Introduction

Intercellular communication is essential for the cell-fate decisions that drive both physiological and pathological processes in a wide range of fields, from tissue homeostasis to development to immunology.<sup>[1]</sup> Thus, decoding the interactions that drive intercellular communication is extremely important to understand the underlying physiology and pathology, as well as to facilitate the design of potential drug targets. With respect to the latter, the complex interactions between peptide major histocompatibility complex (pMHC) ligands and T cell receptors (TCRs) are of particular interest due to the implications they have for the field of cancer immunotherapy. With the advance of high-throughput microscopy and single-cell-based molecular analysis, a variety of methods have been developed that allow for the investigation

of membrane protein-protein interactions (PPIs), such as PIC-seq,<sup>[2]</sup> PUP-IT,<sup>[3]</sup> and BioID.<sup>[4]</sup> However, it remains unclear whether these methods are sensitive enough to capture the complex interactions between pMHC ligands and TCRs.

TCR repertoire diversity, especially in the complementarity-determining regions (CDR1, CDR2, and CDR3), is generated by a process of somatic recombination. Accurate interactions between these CDRs and the pMHC complex account for specific TCR-pMHC recognition. The CDR1 and CDR2 loops interact mainly with the MHC molecules, while the hypervariable CDR3 loop directly contacts the antigen peptides.<sup>[5]</sup> The complicated epitope recognition of TCRs<sup>[6]</sup> and the enormous diversity of TCR repertoire in each individual<sup>[7]</sup> demand methods that allow screening the TCR-pMHC interactions in a high-throughput manner. Thus, some high-throughput approaches have been recently established for mapping TCR-antigen interactions, including DNA-barcoded pMHC tetramer,<sup>[8]</sup> yeast display platform,<sup>[9]</sup> cell-based approaches<sup>[10]</sup> (e.g., trogocytosis<sup>[11]</sup>) and engineered retrovirus-based systems (ENTER and RAPTR).<sup>[12,13]</sup> These approaches enable the identification of TCRs that recognize given antigens or the antigen recognized by a given TCR. However, most of the current methods are still constrained by extensive antigenic peptide synthesis, poor sensitivity to low-affinity antigens, inability to perform large-scale “library-on-library” screening, and labor intensity.

Doublets (a pair of two cells) are routinely observed in flow cytometry or single-cell analysis, but are often considered

Y. Wang, Z. Wang, J. Yang, Y. Liu, G. Li  
National Key Laboratory of Immunity and Inflammation  
Suzhou Institute of Systems Medicine  
Chinese Academy of Medical Sciences & Peking Union Medical College  
Suzhou, Jiangsu 215123, China  
E-mail: [lgd@ism.cams.cn](mailto:lgd@ism.cams.cn)

Y. Wang, Z. Wang, J. Yang, Y. Liu, G. Li  
Key Laboratory of Synthetic Biology Regulatory Element  
Suzhou Institute of Systems Medicine  
Chinese Academy of Medical Sciences & Peking Union Medical College  
Suzhou, Jiangsu 215123, China

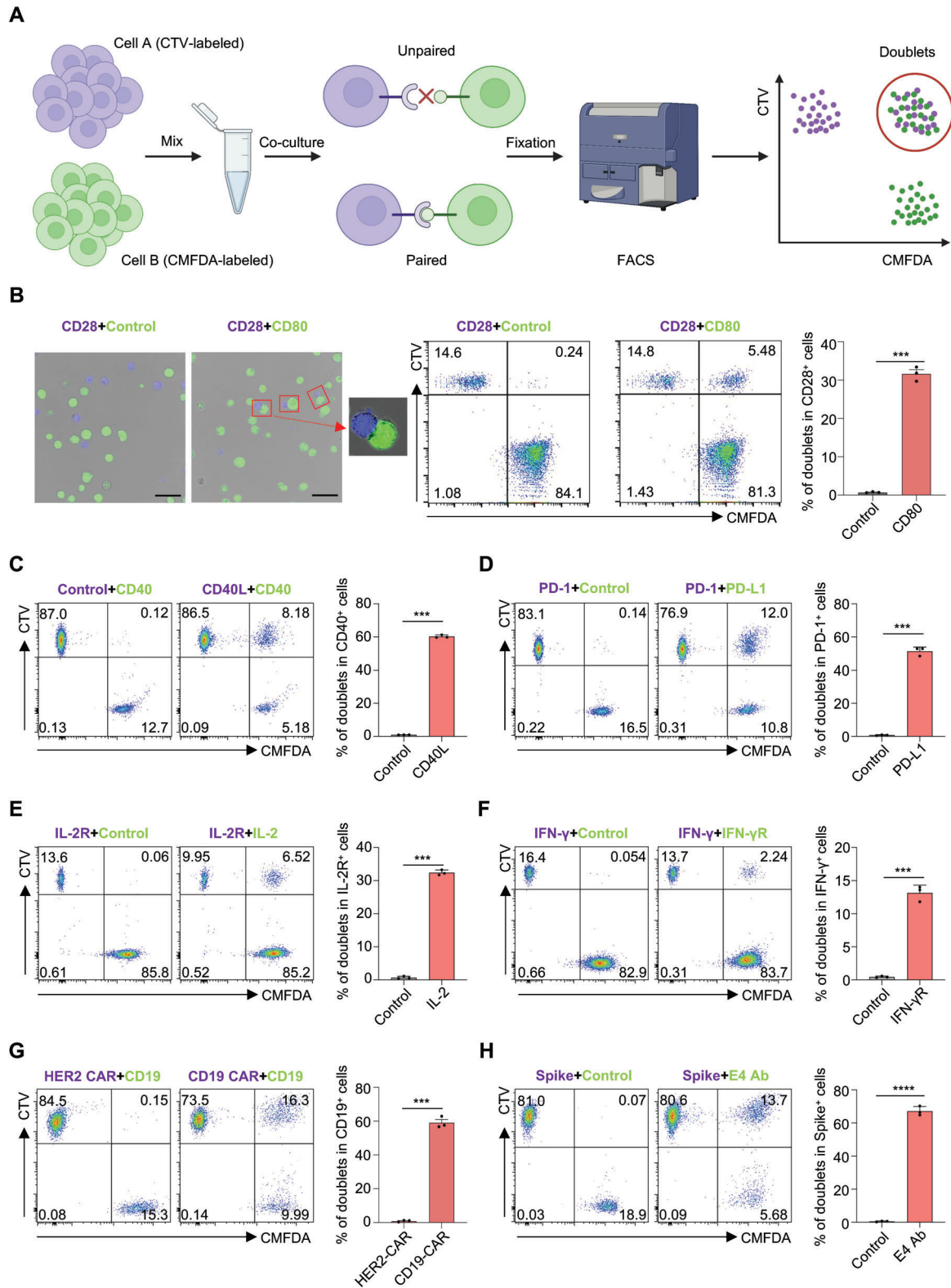
X. Lei, J. Wang  
NHC Key Laboratory of Systems Biology of Pathogens  
Institute of Pathogen Biology  
Chinese Academy of Medical Sciences and Peking Union Medical College  
Beijing 100730, China  
E-mail: [wangjianwei@cams.cn](mailto:wangjianwei@cams.cn)

L. Frankiw  
Department of Pediatrics  
Boston Children's Hospital  
Boston, MA 02115, USA

 The ORCID identification number(s) for the author(s) of this article can be found under <https://doi.org/10.1002/advs.202305750>

© 2024 The Authors. Advanced Science published by Wiley-VCH GmbH. This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

DOI: [10.1002/advs.202305750](https://doi.org/10.1002/advs.202305750)



**Figure 1.** DoubletSeeker allows for the detection of a broad range of interactions between various ligand-receptor types by specifically targeting cell doublets. A) Schema of the DoubletSeeker method. Cell A, expressing the ligand of interest, and cell B, expressing a receptor, were labeled with distinct fluorescence CellTrace dyes. Subsequently, the labeled cells were co-incubated, and the interaction between the ligand and receptor on the cells led to the formation of cell doublets. In this diagram, cell A was labeled with CTV, while cell B was labeled with CMFDA. The FACS technique was utilized to capture the double fluorescent doublets. B) Detection of doublets formation by confocal microscopy (left) and by flow cytometry (right) after CD28-Jurkat

technological artifacts and are thus discarded, or ignored before gathering data.<sup>[14]</sup> However, doublets are not simply technological artifacts of the flow cytometry process but rather represent accurate physical cell–cell contact.<sup>[15]</sup> Due to physical cell–cell interactions, doublets were commonly observed in many different cell types, including T cells with dendritic cells (DCs) and tumor cells.<sup>[16]</sup> The specificity of doublets formation between the T cells and antigen-presenting cells (APCs) is likely dictated by the immunological synapses created via the TCRs' recognition of pMHCs.<sup>[17]</sup> In the present study, we utilized the phenomenon of doublet formation to develop a new method, DoubletSeeker, to facilitate the library-on-library screening of the ligand-receptor interactions. DoubletSeeker combines doublets formation and single-doublet sequencing technology, allowing high-throughput capture of ligand-receptor pairing information. We showed that DoubletSeeker enables the characterization of a variety of membrane protein-protein interactions, including interactions between cytokine receptors and membrane-anchored ligands, immune costimulatory or immunosuppressive molecules and their respective ligands, antibodies, and antigens, as well as TCRs and pMHCs. We further verified the high sensitivity and specificity of DoubletSeeker in identifying T cells specific for a given antigen and in discovering antigens for a given TCR. Finally, we have successfully demonstrated the capability of DoubletSeeker to capture pairs of TCR-pMHC through a comprehensive library-on-library screen.

## 2. Results

### 2.1. Establishing DoubletSeeker to Capture Membrane Protein–Protein Interactions

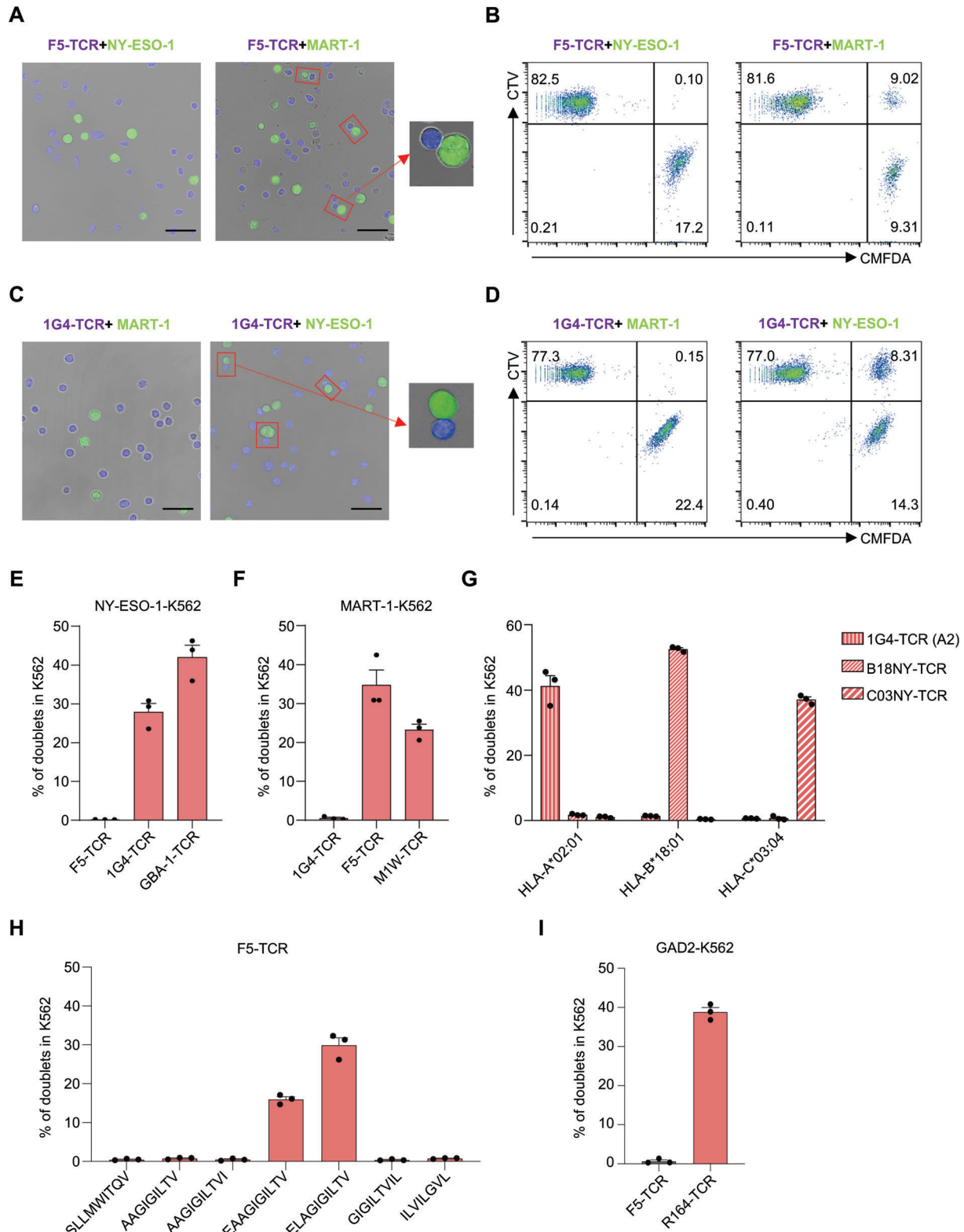
Here, we sought to develop a method, termed DoubletSeeker, that can simultaneously capture interacting cell pairs and uses sequencing technology to decode the information of both the ligand and receptor in these cell pairs (Figure 1A). We first established cell lines stably expressing genes encoding either receptors or the respective ligand, including Jurkat cells expressing CD28 and K562 cells expressing CD80. To minimize the potential signal-to-noise ratio of the doublet, we compared the fluorescent intensity of fluorescent proteins and tracer dyes (CellTrace Violet-CTV and CellTrace Green-CMFDA) by flow cytometry. We found that the tracer dyes-labeled cells are much brighter than the cells expressing fluorescent proteins and the tracer dyes do not leak from cells (Figure S1A,B, Supporting Information), making the tracer dyes suitable for long-term cell labeling and doublet tracking. Indeed, tracer dye-labeled cells displayed less non-specific formation of doublets (Figure S1C,D, Supporting Information). To directly test the formation of cell doublets upon specific ligand-receptor interaction, we co-incubated tracer dyes-

labeled CD28-Jurkat cells with control K562 cells or CD80-K562 cells and examined the cell doublets (CTV<sup>+</sup>CMFDA<sup>+</sup>). Using both microscopy and flow cytometry, we observed significant amounts of cell doublets formed upon the incubation of CD28-Jurkat cells with cognate CD80-K562 cells, but not with control K562 cells (Figure 1B), suggesting cell doublets are sufficiently stable for flow cytometry tracing. We next tested if cell doublets could be formed among different ligand-receptor interactions. We observed the formation of doublets through various ligand-receptor interactions, including interactions between immune-associated proteins and their receptors (CD40L/CD40 and programmed death 1(PD1)/programmed death ligand 1(PD-L1) and CD28/CD86), glycosylphosphatidylinositol(GPI)-modified membrane-anchored cytokines and their receptors (interleukin-2(IL-2)/interleukin-2 receptor(IL-2R) and interferon- $\gamma$ (IFN- $\gamma$ )/interferon- $\gamma$  receptor(IFN- $\gamma$ R)), viral protein and their receptors (the Spike protein of severe acute respiratory syndrome coronavirus 2(SARS-CoV-2) and its different receptors, angiotensin converting enzyme(ACE2),<sup>[18]</sup> asialoglycoprotein receptor 1(ASGR1),<sup>[19]</sup> neuropilin(NRP1),<sup>[20]</sup> kringle containing transmembrane protein 1(KREMEN1),<sup>[19]</sup> tyrosine-protein kinase receptor(AXL),<sup>[21]</sup> C-type lectin domain family 10 member A(CLEC10A),<sup>[19]</sup> dipeptidyl peptidase-4(DPP4),<sup>[22]</sup> low-density lipoprotein receptor(LDLR),<sup>[23]</sup> and so on<sup>[24]</sup> (Figure 1C–F; Figure S2A–G, Supporting Information). In addition, the interaction of membrane proteins with antibodies can also be captured by doublet formation, including chimeric antigen receptor (CAR) and their target membrane proteins (CD19-CAR/CD19, epidermal growth factor receptor(EGFR)-CAR/EGFR, human epidermal growth factor receptor(HER2)-CAR/HER2), and the Spike protein of SARS-CoV-2 and its neutralizing antibody (E4) (Figure 1G,H; Figure S2H, Supporting Information). Importantly, the introduction of recombinant E4 antibody into the coculture medium effectively blocks the formation of doublets between Spike protein and its receptor ACE2 (Figure S2I, Supporting Information). Thus, the DoubletSeeker method can be used to capture various membrane protein interactions.

### 2.2. Identification of Interactions between TCR-pMHC by DoubletSeeker

The binding strength between TCR and pMHC is generally very low.<sup>[25]</sup> Thus, we examined if the interaction of TCR and pMHC would be sufficiently stable for FACS tracing. By co-incubation of Jurkat cells expressing F5-TCR with K562 cells expressing its cognate peptide HLA-A\*02:01/MART-1<sub>26-35(A27L)</sub>, or K562 cells expressing HLA-A\*02:01/NY-ESO-1<sub>157-165(C165V)</sub> as single-chain trimers (SCTs),<sup>[26]</sup> we observed antigen-specific forming of cell doublets of F5-TCR-Jurkat cells with MART-1-K562 cells

cells (CTV-labeled, blue) co-incubated with either control K562 or CD80-K562 cells (CMFDA-labeled, green). Red rectangle boxes mark the presence of doublets. Scale bars, 50  $\mu$ m. Representative flow cytometry plots for doublet identification are presented in the middle. Bar plots presented to the right show the percentage of doublets for CD28-Jurkat cells. (C to H) Formation of doublets was tested across a broad range of receptors and targeting ligands, including C) CD40/CD40L, D) PD1/PD-L1, E) IL-2R/IL-2, F) IFN- $\gamma$ R/IFN- $\gamma$ , G) CD19-CAR/CD19, and the Spike protein of SARS-CoV2 with its specific antibody E4-Ab (H). Representative flow cytometry plots for doublet formation are presented on the left. Bar plots on the right show the percentage of doublets within the indicated ligand or receptor-expressing cells. CTV: CellTrace Violet, CMFDA: CellTrace CMFDA. Data are represented as the means  $\pm$  SEM.  $n = 3$ . \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , and \*\*\*\* $p < 0.0001$  (unpaired two-tailed Student's t-test). Data are representative of three independent experiments (B–H).



**Figure 2.** DoubletSeeker serves as a powerful tool for the identification of interactions between the TCRs and pMHCs. A–D) Doublets formed by co-cubation of CTV-labeled F5-TCR-Jurkat (A,B) or 1G4-TCR-Jurkat (C,D) cells with CMFDA-labeled NY-ESO-1-K562 or MART-1-K562 cells (TCR-expressing cells: SCT-expressing cells = 5:1) were analyzed by confocal microscopy (A,C) and flow cytometry (B,D). F5-TCR is paired with MART-1-SCT, and 1G4-TCR is paired with NY-ESO-1-SCT. Red rectangle boxes mark the presence of doublets. Scale bars, 50  $\mu$ m. E) Percentage of doublets in NY-ESO-1-K562 cells after co-incubation with their cognate 1G4-TCR-Jurkat, GBA-1-TCR-Jurkat or noncognate F5-TCR-Jurkat cells (TCR-expressing cells: SCT-expressing

but not with NY-ESO-1-K562 cells (Figure 2A). Notably, these TCR-pMHC interacting cell doublets were very stable and can be acquired by flow cytometry (Figure 2B). Similarly, the co-incubation of 1G4-TCR-Jurkat cells with cognate NY-ESO-1-K562 cells also led to the stable and specific formation of cell doublets (Figure 2C,D). Coreceptor CD8 acts to stabilize the binding of TCR to the pMHC complex.<sup>[27]</sup> Indeed, we observed significantly decreased percentages of cell doublets when CD8 was not expressed in F5-TCR-Jurkats or 1G4-TCR-Jurkat cells (Figure S3A,B, Supporting Information).

Next, we examined the formation of doublets in SCT-K562 cells upon co-incubation with TCR-Jurkat cells at different ratios or for different time-periods. We observed a gradual increase in the percentage of cell doublets in MART-1- or NY-ESO-1-K562 cells as the proportions of their cognate F5- or 1G4-TCR-Jurkat cells increased (Figure S3C,D, Supporting Information). Furthermore, after 30 min of co-incubation, approximately 40–50% of MART-1-K562 and NY-ESO-1-K562 cells formed doublets with their cognate TCR-Jurkat cells. However, with extended co-incubation beyond 30 min, the formation of doublets gradually decreased (Figure S3E,F, Supporting Information). In addition, the percentage of nonspecific doublets remained  $\approx$ 1% and showed minimal changes across different conditions, further demonstrating that the formation of doublets was TCR-pMHC interaction specific.

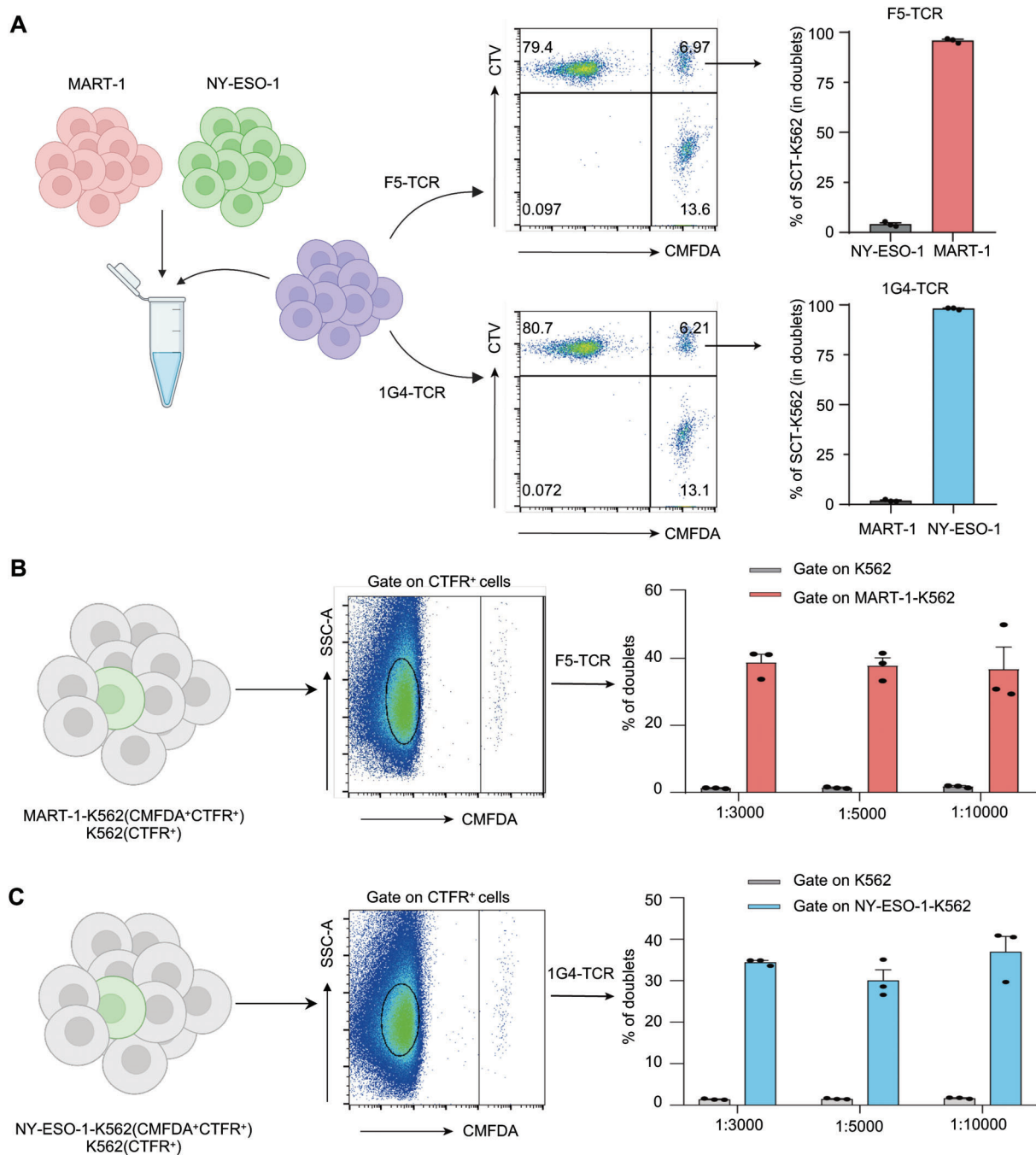
In addition to F5- and 1G4-TCR, we also observed specific doublets formation for five additional TCR-pMHC cognate pairings, including a novel A2-restricted NY-ESO-1-specific TCR (GBA-1-TCR), a low-affinity A2-restricted MART-1-specific TCR (M1W-TCR), and two HLA-B\*18:01- and C\*03:04-restricted NY-ESO-1-specific TCRs (Figure 2E–G). Of note, even low-affinity M1W-TCR displayed a substantial percentage of cell doublets, suggesting the DoubletSeeker method would be able to enrich lower affinity TCR-pMHC interactions. We next evaluated the sensitivity of DoubletSeeker in differentiating the recognition of peptide variants by a TCR. By incubating F5-TCR-Jurkat cells with A2-K562 cells that express different MART-1 peptide variants at the same level, we found that there were more cell doublets formed in the co-incubation of F5-TCR-Jurkat cells with K562 cells presenting heteroclitic MART-1 peptide (ELAGIGILTV) than that with K562 cells presenting native low avidity peptide (EAAGIGILTV), suggesting the extent of doublets formation correlates with the avidity of the TCR-pMHC interaction (Figure 2H). Of note, such correlation is absent in receptor-ligand interactions and antibody-antigen interactions (Figure S3G, Supporting Information). By co-culturing K562-ACE2 cells and Jurkat-Spike cells with varying levels of Spike protein expression, we observed a substantial increase in the percentage of cell doublet formation when target protein was expressed at higher levels on the cells, revealing an association between doublets formation and protein ex-

pression levels (Figure S3H,I, Supporting Information). Therefore, we hypothesize that optimizing TCR or peptide presentation levels could enhance the applicability of DoubletSeeker in primary cells. To investigate the impact of increased antigen presentation on doublet formation, we treated B16F10 and B16F10-OVA tumor cells with IFN- $\gamma$ , which is known to upregulate MHC class I expression and antigen presentation. We observed a significant increase in doublet formation (14.3%) during coculture with OVA<sub>SIINFEKL</sub>-specific primary OT-I T cells, in comparison to untreated B16F10-OVA cells (4.02%) and control B16F10 cells (1.15%). Additionally, we found that  $\approx$ 50% of OVA<sub>SIINFEKL</sub>-SCT overexpressing B16F10 (B16F10-OVA SCT) cells formed doublets with primary OT-I T cells, further supporting the notion that high levels of surface antigen expression increase doublet formation (Figure S3J,K, Supporting Information). In addition to MHC-I restricted TCRs, we also tested the ability of DoubletSeeker to detect MHC-II restricted TCR-antigen interactions. We found that Jurkat cells expressing MHC-II restricted R164-TCR can also form stable doublets with K562 cells presenting the cognate GAD2 peptide (Figure 2I). These findings demonstrate that DoubletSeeker is capable of capturing specific interactions between TCRs and pMHCs.

### 2.3. Specificity and Sensitivity of DoubletSeeker for Resolving Cognate Antigens

We next assessed DoubletSeeker's ability to distinguish antigen-expressing target cells from cells expressing noncognate antigens for a given TCR. We co-incubated CMFDA-labeled NY-ESO-1-K562 cells and MART-1-K562 cells (in which target cells were additionally labeled with CellTrace Far-red-CTFR) in a 1:1 ratio with CTV-labeled Jurkat cells expressing F5- or 1G4-TCR respectively. Target MART-1-K562 cells for F5-TCR were substantially enriched (98%) in the cell doublets when the NY-ESO-1- and MART-1-K562 cell mixture was co-incubated with F5-TCR-Jurkat cells (Figure 3A, upper panel). Similarly, the enrichment of NY-ESO-1-K562 cells was observed when the mixture was incubated with 1G4-Jurkat cells (Figure 3A, lower panel). Thus, DoubletSeeker can specifically enrich cognate TCR-pMHC cell pairs, allowing the efficient identification of on-target cells from noncognate epitope-presenting cells. To further validate the sensitivity of DoubletSeeker in capturing cognate target cells, we conducted a proof-of-concept library screening. We mixed on-target MART-1-K562 (CTFR<sup>+</sup>CMFDA<sup>+</sup>) and control K562 cells (CTFR<sup>+</sup>CMFDA<sup>-</sup>) at different ratios, and then cocultured this mixture with CTV-labeled F5-TCR-Jurkat cells specific to MART-1. We then quantified the percentage of cell doublets formed by MART-1-K562 and control K562 cells with F5-TCR-Jurkat cells (CTFR<sup>+</sup>CMFDA<sup>+</sup>CTV<sup>+</sup> versus CTFR<sup>+</sup>CMFDA<sup>-</sup>CTV<sup>+</sup>). A

cells = 5:1) was analyzed by flow cytometry. F) Percentage of doublets in MART-1-K562 cells, following co-incubation with Jurkat cells expressing the cognate F5-TCR or M1W-TCR (TCR-expressing cells: SCT-expressing cells = 5:1), was analyzed by flow cytometry. G) Antigen-specific doublet formation between K562 expressing HLA-A\*02:01, HLA-B\*18:01 or HLA-C\*03:04 restricted NY-ESO-1 epitopes and Jurkat cells expressing their cognate TCRs (TCR-expressing cells: SCT-expressing cells = 5:1). H) Comparison of doublets formation in K562 cells expressing diverse HLA-A\*02:01-restricted MART-1 peptide variants after co-incubation with F5-TCR-Jurkat cells. I) Percentage of doublets in K562 cells expressing MHC-II restricted GAD2 antigen after co-incubation with cognate R164-TCR-Jurkat or noncognate F5-TCR-Jurkat cells (TCR-expressing cells: SCT-expressing cells = 5:1) was analyzed by flow cytometry. CTV: CellTrace Violet, CMFDA: CellTrace CMFDA. Data are represented as the means  $\pm$  SEM.  $N = 3$ . Data are representative of three independent experiments (A–I).



**Figure 3.** DoubletSeeker demonstrates remarkable specificity and sensitivity in identifying cognate antigens for a given TCR. A) Schematic and representative flow cytometry plots for a 1:1 mixture of MART-1-K562 and NY-ESO-1-K562 cells (CMFDA-labeled) co-incubated with either CTV-labeled F5-TCR-Jurkat cells (top) or 1G4-TCR-Jurkat cells (bottom) (TCR-expressing cells: SCT-expressing cells = 5:1). K562 cells expressing the cognate antigens were additionally labeled with CTFR. The percentage quantification of each SCT-K562 cell in the doublet population is presented on the right. B,C) Schematic and representative flow cytometry plots for the mixture containing cognate B) MART-1-K562 or C) NY-ESO-1-K562 cells with noncognate K562 cells at different ratios (1:3000, 1:5000, 1:10 000) co-incubated with either F5-TCR-Jurkat cells (B) or 1G4-TCR-Jurkat cells (C) (TCR-expressing cells: SCT-expressing cells = 1:1). The percentage quantification of doublets in each K562 cells are presented on the right. MART-1-K562 and NY-ESO-1-K562 cells were labeled with CMFDA and CTFR, whereas K562 cells were labeled with CTFR. CTV: CellTrace Violet, CMFDA: CellTrace CMFDA, CTFR: CellTrace Far-red. Data are represented as the means  $\pm$  SEM.  $n = 3$ . Data are representative of three independent experiments.

similar experiment was also performed using NY-ESO-1-K562 cells and control K562 cells in combination with 1G4-TCR-Jurkat cells. We found that, compared with control K562 cells, a remarkably higher percentage of target antigen-presenting K562 cells ( $\approx 40\%$ ) formed doublets with their cognate TCR-Jurkat cells, even when the target cells were at a ratio as low as 1:10 000 (Figure 3B,C; Figure S4, Supporting Information). In addition, DoubletSeeker also effectively isolates various interacting cells, including CD28/CD80, CD28/CD86, CD19-CAR/CD19, and CD40/CD40L interactions (Figure S5, Supporting Information). These results demonstrated the high specificity and sensitivity of DoubletSeeker for resolving cognate antigens for a given TCR, as well as a broad range of membrane protein interactions.

#### 2.4. Identification of Cognate TCR Antigens by DoubletSeeker

With the key characteristics of DoubletSeeker validated, we next asked whether DoubletSeeker can be used to screen the candidate antigen for a given TCR in a more extensive library. We designed a strategy to utilize DoubletSeeker for the acquisition of FACS-sorted cell doublets, which were then subjected to Next-generation sequencing (NGS) analysis for the identification of target epitopes (Figure 4A). We incubated CTV-labeled F5-TCR-Jurkat cells with CMFDA-labeled K562 cells expressing an A2-restricted SCT cDNA library consisting of all known HLA-A\*02:01-restricted epitopes, from the Immune Epitope Database (IEDB). FACS-sorted (CTV<sup>+</sup>CMFDA<sup>+</sup>) cell doublets were then subjected to NGS to identify the enriched epitopes. After one round of doublet selection, the top four enriched epitopes that share similar sequences were known cognate ligands for the F5-TCR (Figure 4B; Figure S6A, Supporting Information). These data demonstrate that DoubletSeeker can enrich cognate antigens and decode the cross-reactive ligands with a similar sequence for a given TCR. We further tested the versatility of DoubletSeeker for neoantigen discovery using a neoantigen-specific tumor-reactivity Neo-TCR that was isolated from a subject with metastatic melanoma. We used the Neo-TCR as a surrogate for a tumor-reactive orphan TCR and conducted co-incubation experiments with CTV-labeled Jurkat cells expressing Neo-TCR and CMFDA-labeled K562 cells transduced with a neoepitope SCT cDNA library containing 3251 predicted A2-restricted neoepitopes (8 to 12 amino acids in length) corresponding to nonsynonymous mutations found in the tumor. Subsequent NGS sequencing and the ranking result of enriched epitopes showed that the top two hits were epitopes derived from USP-7 with a nonsynonymous D798Y mutation (YLYHRVDVI and YLYHRVDVIF) (Figure 4C; Figure S6B, Supporting Information). These results demonstrate that the DoubletSeeker method is an efficient and effective method for the deorphanization of TCRs.

#### 2.5. Identification of Antigen-Specific TCRs by DoubletSeeker

Next, we explored if the DoubletSeeker method could enable the capture of TCRs that specifically recognize a given antigen. When CTV-labeled F5-TCR-Jurkat cells and 1G4-TCR-Jurkat cells were co-incubated with CMFDA-labeled MART-1- or NY-ESO-1-K562 cells, significant formation of cell doublets was ob-

served in the presence of their respective target antigens. Conversely, no such interaction was observed when co-incubating the TCR-Jurkat cells with K562 cells expressing noncognate antigens (Figure S7A,B, Supporting Information). This data indicates a specific interaction between TCR-expressing Jurkat cells and K562 cells expressing their cognate antigens. To further assess the sensitivity of DoubletSeeker to capture antigen-specific TCRs, we mixed F5- or 1G4-TCR-Jurkat (CTV<sup>+</sup>CTFR<sup>+</sup>) cells with control Jurkat cells (CTV<sup>-</sup>CTFR<sup>+</sup>) at different ratios and then co-incubated them with MART-1-K562 cells and NY-ESO-1-K562 cells (CMFDA<sup>+</sup>), respectively. We observed a significantly higher proportion of cell doublets (CTV<sup>+</sup>CTFR<sup>+</sup>CMFDA<sup>+</sup>) formed by F5-TCR-Jurkat and 1G4-TCR-Jurkat cells ( $\approx 40\%$ ) with targeted antigen-presenting K562 cells compared to their control Jurkat counterparts, even when the ratio of antigen-specific Jurkat cells was as low as 1:10 000 (Figure 5A,B; and Figure S7C, Supporting Information).

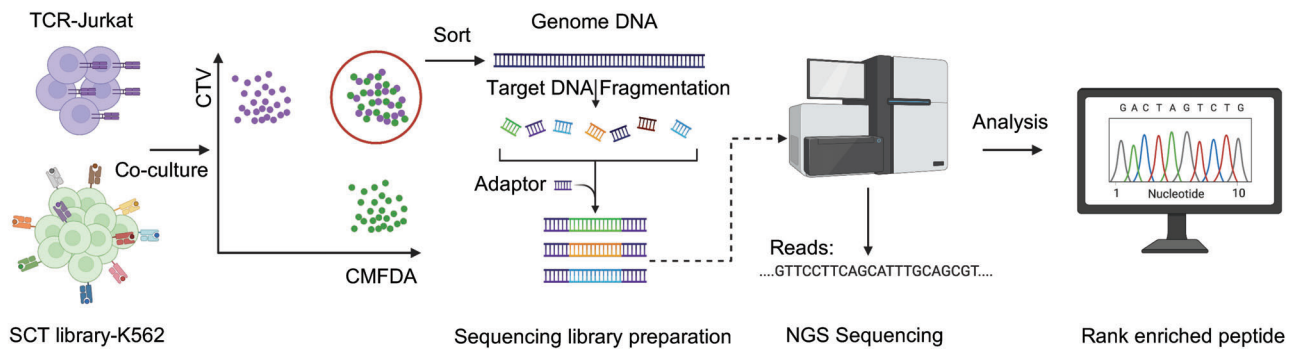
We further generated a comprehensive mutant M1W-TCR library by performing saturation mutagenesis on the six amino acids (AA) in the CDR3 $\beta$  region of M1W TCR $\beta$ , which pairs with the identical  $\alpha$  chain of the M1W-TCR (Figure 5C). The M1W-TCR with its corresponding mutation library (mixed with 0.1% WT M1W-TCR virus) was transduced into CD8-expressed J76 cells (referred to as M1W-TCR pools-J76), an endogenous TCR $\alpha$  and TCR $\beta$  deficient Jurkat cell, at a multiplicity of infection (MOI) of 4 to achieve a high infection rate (Figure S7D, Supporting Information). Subsequently, TCR-positive J76 cells were isolated and co-incubated with MART-1-K562 cells, followed by FACS isolation of doublets and sequencing to identify the target TCRs. The top TCR enriched in the doublets was the wild-type (WT) M1W-TCR that recognizes the MART-1<sub>ELAGIGILT</sub> epitope (Figure 5D,E). The other top enriched mutant TCRs exhibited much fewer reads in the doublets sequencing data and further validation experiments confirmed that none of these TCRs show strong interaction with the MART-1<sub>ELAGIGILT</sub> epitope (Figure S8, Supporting Information). These results indicate that DoubletSeeker is capable of enriching antigen-specific TCRs with high specificity and sensitivity.

#### 2.6. The Library-on-Library Screening Capability of DoubletSeeker

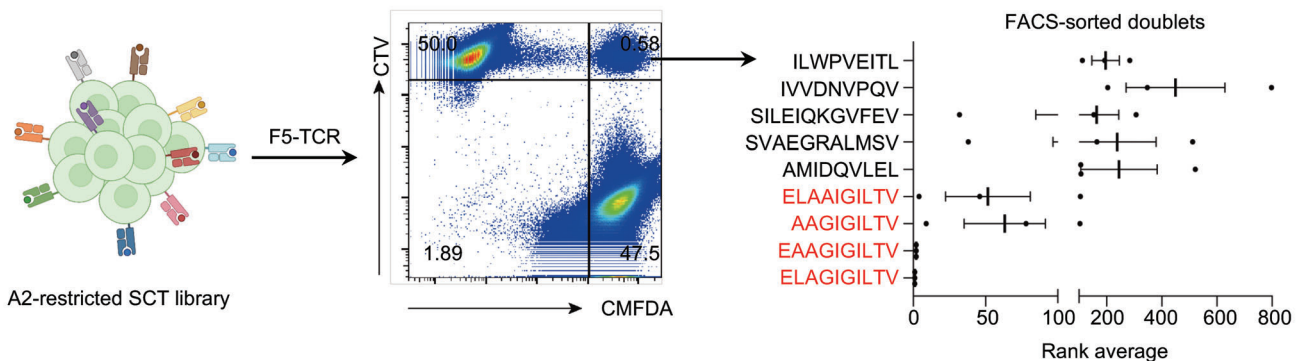
To evaluate the library-on-library screening capability of DoubletSeeker, we mixed the TCR-expressing Jurkat cells (F5-TCR, 1G4-TCR and Neo-TCR) with their targeted SCT-expressing K562 cells (MART-1, NY-ESO-1 and USP-7) at a 1:1 ratio. Each population of TCR-Jurkat cells and SCT-K562 cells were labeled with distinct fluorescent markers (Figure S9A, Supporting Information). We found a significant increase in the formation of cell doublets was observed in the paired TCR-SCT expressing cells compared to the unpaired cells (Figure S9B, Supporting Information). This finding indicates the potential of DoubletSeeker in facilitating the identification of TCR-pMHC interactions within complex cellular mixtures.

Next, we devised a single doublet-capturing strategy to validate the accuracy of DoubletSeeker in library-on-library screening, as illustrated in Figure 6A. The A2-restricted SCT of the K562 library, as mentioned in Figure 4B, was co-incubated with a mixture of F5-, 1G4-, GBA-1-, and Neo-TCR-Jurkat cells (a mixture

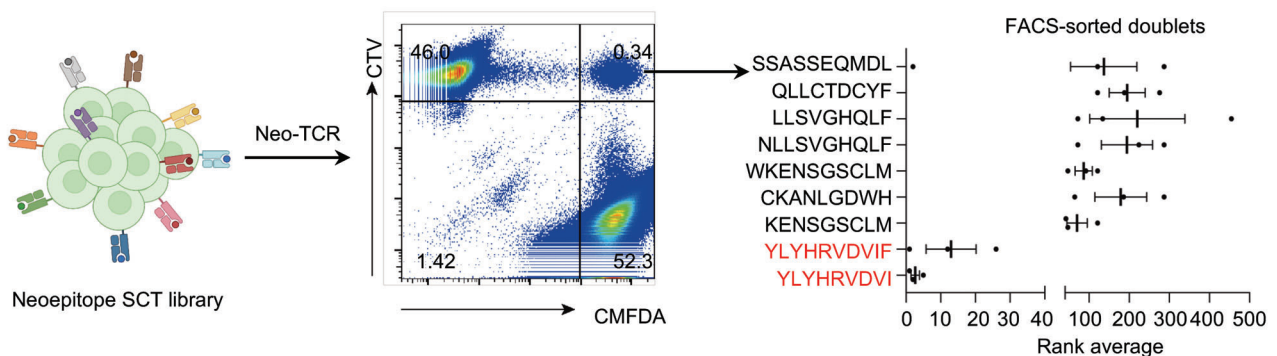
A



B

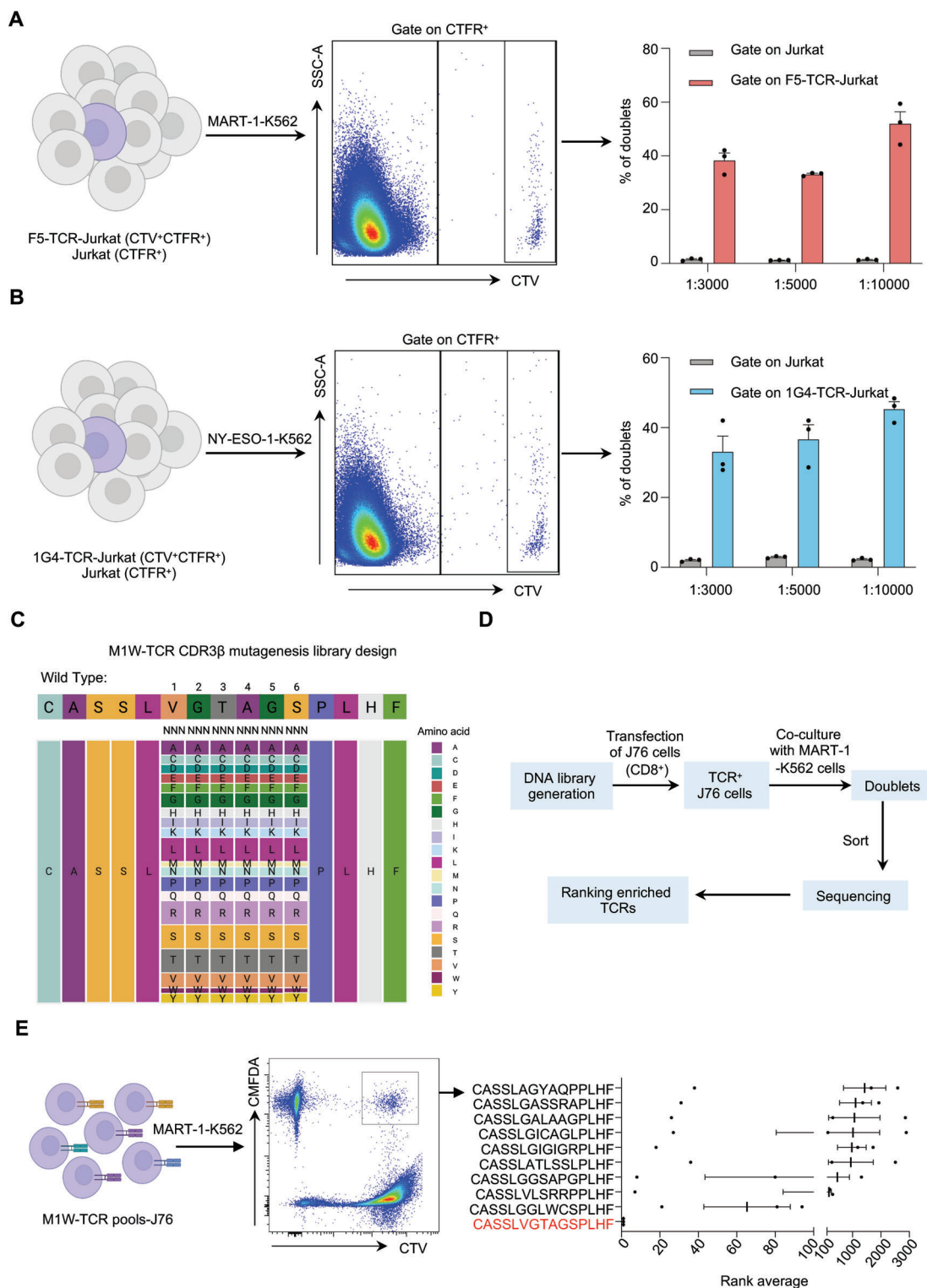


C



**Figure 4.** DoubletSeeker excels in identifying the cognate antigens for public F5-TCR and neoantigen-specific TCR. A) Schematic diagram of DoubletSeeker for screening TCR cognate epitopes from an SCT library. CTV-labelled TCR-expressing Jurkat cells co-incubated with CMFDA-labelled SCT library-expressing K562 cells at a 1:1 ratio. Subsequently, FACS-sorted (CTV<sup>+</sup>CMFDA<sup>+</sup>) cell doublets were then subjected to NGS to identify the enriched epitopes based on read counts. B,C) Schematic and representative flow cytometry plots for the identification of doublets in K562 cells expressing an A2-restricted SCT library containing 12055 epitopes (B) and a neoepitope SCT library containing 3251 unique neoepitopes (C) co-incubated with F5-TCR-Jurkat (B) or Neo-TCR-Jurkat (C) cells (TCR-expressing cells: SCT-expressing cells = 1:1), respectively. Identification of enriched epitopes after one round of selection by NGS is shown on the right. The rank average is determined by calculating the ranking of each peptide based on its abundance among all peptides, from three independent experiments. The SCT library-K562 cells were labeled with CMFDA, and the TCR-Jurkat cells were labeled with CTV. Data are represented as the means  $\pm$  SEM.  $n = 3$ .





**Figure 5.** DoubletSeeker enables identification of cognate TCRs for a given antigen. A, B) Schematic and representative flow cytometry plots for the mixture containing cognate A) F5-TCR-Jurkat or B) 1G4-TCR-Jurkat cells with control Jurkat cells at different ratios (1:3000, 1:5000, 1:10 000) co-incubated with either A) MART-1-K562 or B) NY-ESO-1-K562 cells labeled with CMFDA (TCR-expressing cells: SCT-expressing cells = 1:1). Bar plots that show the population of doublets in each Jurkat cell population are presented on the right. F5-TCR-Jurkat and 1G4-TCR-Jurkat cells were labeled with CTV and CTFR, whereas Jurkat cells were labeled with CTFR. C) Schematic for the design of the M1W-TCR CDR3 $\beta$  combinatorial mutagenesis library. D) Strategy for screening MART-1-specific TCRs from the M1W-TCR pools-J76 library using DoubletSeeker. E) Schematic and representative flow cytometry plots for the identification of doublets in CD8-J76 cells expressing the M1W-TCR and its corresponding mutation library (pre-mixed 0.1% WT M1W-TCR virus)

of 4 TCR-Jurkat cells). After isolating single doublet using FACS, we conducted single-cell PCR sequencing to obtain the TCR and antigen information for each doublet. Remarkably, the analysis of these sequenced doublets revealed that 93% of doublets containing well-described TCR-pMHC cognate pairings, including F5-TCR/MART-1-SCT (71.19%), 1G4-TCR/NY-ESO-1-SCT (20.34%) and GBA-1-TCR/NY-ESO-1-SCT (1.69%) (Figure 6B). Similarly, in the screening of the neopeptide SCT library with the 4 TCR-Jurkat cell mixture, the Neo-TCR/USP7-SCT cognate pairings were predominantly enriched, accounting for over 94% of the isolated doublets (Figure 6C).

To further evaluate the efficacy of DoubletSeeker in enriching TCR-pMHC interactions within a highly complex cellular mixture, we mixed the A2-restricted antigen library expressing K562 cells and the neopeptide library expressing K562 cells together. This combined antigen library was subsequently co-incubated with the low MOI transduced M1W-TCR pools-J76 cells, along with a small proportion of F5- and 1G4-TCR-Jurkat cells (Figure S9C, Supporting Information). After one round of doublets selection, our single-doublets sequencing data revealed that DoubletSeeker successfully enriched the well-described TCR-pMHC pairings, including F5-TCR/MART-1-SCT (8.82%), 1G4-TCR/NY-ESO-1-SCT (4.41%), and WT M1W-TCR/MART-1-SCT (69.12%) cognate TCR-pMHC interactions. In addition, we observed three mutant M1W-TCRs (T3N, S6K and S6T) were also enriched with MART-1<sub>ELAGIGILTV</sub> epitope, categorized as unknown TCR-pMHC pairings (Figure 6D). To verify the interaction of MART-1<sub>ELAGIGILTV</sub> epitope with the three mutant M1W-TCRs, DoubletSeeker was performed to assess the formation of doublets by co-incubating MART-1-K562 cells with J76 cells expressing M1W<sub>T3N</sub>, M1W<sub>S6K</sub>, or M1W<sub>S6T</sub>-TCR, respectively. As expected, there was an increased percentage of doublets in MART-1-K562 cells co-incubated with M1W<sub>T3N</sub>, M1W<sub>S6K</sub>, M1W<sub>S6T</sub>-TCR-Jurkat cells compared to noncognate NY-ESO-1-K562 cells, despite the difference in TCR expression (Figure 6E; Figure S9D, Supporting Information). Notably, the three mutated TCR cells showed a reduced level of doublet formation, suggesting that the three M1W mutant TCRs have a lower binding affinity to MART-1<sub>ELAGIGILTV</sub>. This was further validated by measuring the ability of the three mutant M1W-TCRs to induce human T cell cytotoxicity upon recognition of A2-expressing K562 cells pulsed with MART-1 peptide (ELAGIGILTV) (Figure 6F). Therefore, DoubletSeeker demonstrates the high efficiency in capturing the cognate TCR-pMHC pairs in a highly complex library-on-library screening.

### 3. Discussion

Cell-to-cell communications orchestrate complex biological processes like organismal development, tissue homeostasis, and immune responses. These communications rely on the interactions between ligands and their cell surface receptors. Deciphering these interactions is important for both fundamental research

and drug discovery efforts. In this study, we developed DoubletSeeker, a novel and high-throughput approach for identifying ligand-receptor interactions found between various cell surface proteins such as cytokine receptors, immune costimulatory molecules, viral proteins (i.e., the spike protein of SARS-CoV-2), antibodies, and TCRs. Through a series of proof-of-concept experiments and screening, we demonstrated the broad applicability of DoubletSeeker, including its potential capacity in screening drugs that target membrane protein interactions, such as blocking-antibody discovery.

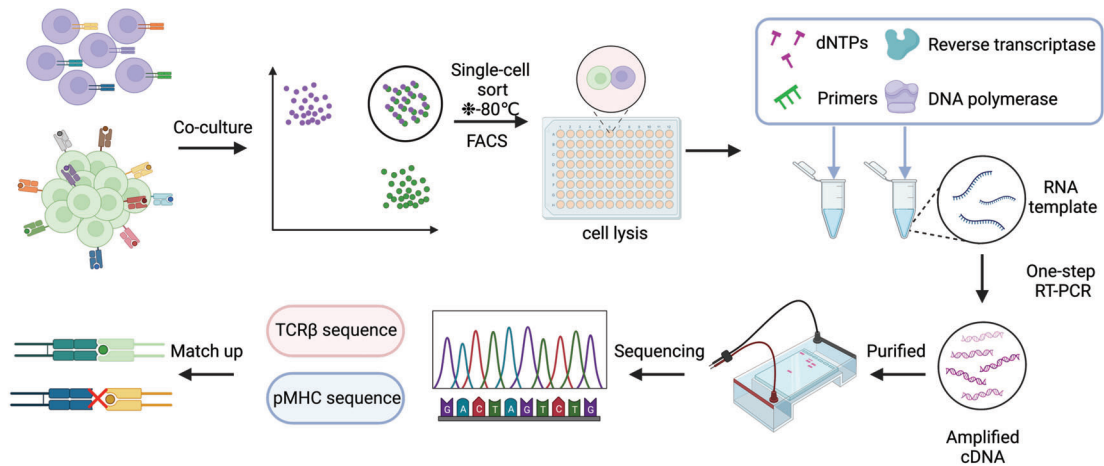
The user-friendly signature of DoubletSeeker makes it easily implementable in a basic biological laboratory. For TCR antigen discovery, DoubletSeeker enables T cell antigen identification with the ability to test a large number of antigens, screen multiple MHC alleles, perform double-blind screening, and overcome limitations of small antigen/TCR pools during “library-on-library” screening. A major limitation to the clinical application of the TCR-T-based strategy is the lack of efficient techniques for the discovery of the cognate antigens of a large number of orphan TCRs.<sup>[28]</sup> In recent decades, pMHC tetramers have been extensively used to identify antigen-specific T cells and facilitate high-throughput screening by coupling with DNA barcoding. However, the application of pMHC tetramers is constrained by the labor consumption of peptide synthesis, the inconvenient assembly of pMHC tetramers, and the requirement for pre-designed antigen information. Various cell-based methods,<sup>[29]</sup> such as T-Scan, TCR-MCR, signaling and antigen-presenting bifunctional receptors (SABRs) and trogocytosis,<sup>[11]</sup> have also been developed to decipher the antigen specificity of specific TCRs. However, these cell-based methods are still constrained by time-consuming multi-round screening, have low sensitivity for low-affinity antigens, and are unable to be used for “library-on-library” screening. DoubletSeeker overcomes these limitations and has the capacity to identify TCR-antigen interactions across various MHC alleles (MHC I and MHC II) and diverse TCR types, which would significantly enhance our understanding of antigen recognition and immune responses on a broader scale. Moreover, DoubletSeeker supports a wider range of application scenarios compared to pMHC tetramers and cell-based methods. It offers a bidirectional screening tool, facilitating the identification of both the TCR specific to a given antigen and the cognate antigen recognized by an orphan TCR.

Several “library-on-library” TCR antigen screening techniques have been established, such as the yeast display system-based YAMTAD, and engineered viral-based ENTER and RAPTR. The application of YAMTAD is hindered by the inherent drawbacks in yeast-display methods, including low stability of displayed proteins on the yeast surface and the potential for inaccurate antigen presentation. In contrast, engineered viral-based display systems address the limitations of the yeast display system but face challenges in preparing large-scale virus-displaying libraries, and their suitability for displaying protein complexes

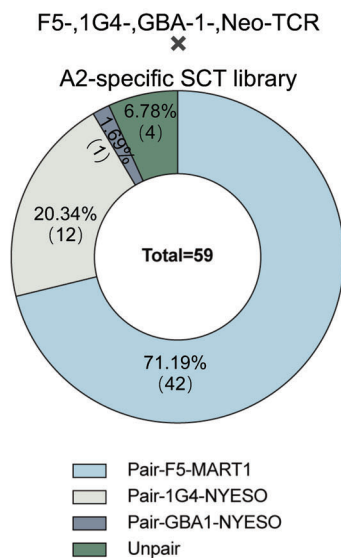
---

co-incubated with MART-1-K562 cells (TCR-expressing cells: SCT-expressing cells = 1:1). Identification of enriched epitopes after one round of selection by NGS is shown on the right. The rank average is determined by calculating the ranking of each peptide based on its abundance among all peptides, from three independent experiments. The M1W-TCR pools-J76 library was labeled with CTV, and the MART-1-K562 cells were labeled with CMFDA. CTV: CellTrace Violet, CMFDA: CellTrace CMFDA, CTRF: CellTrace Far-red. Data are represented as the means ± SEM. *n* = 3. Data are representative of three independent experiments (A,B,E).

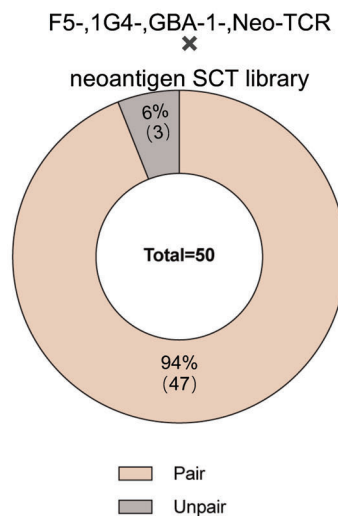
A



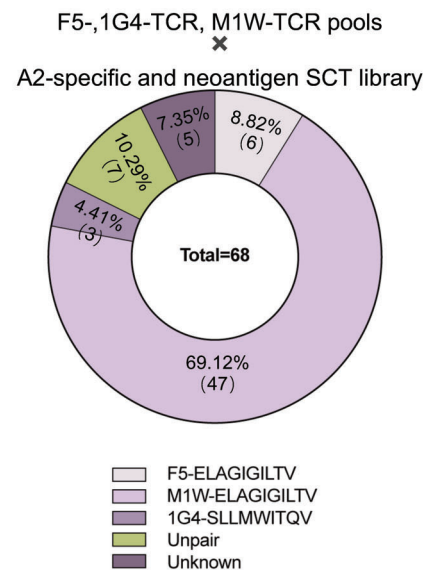
B



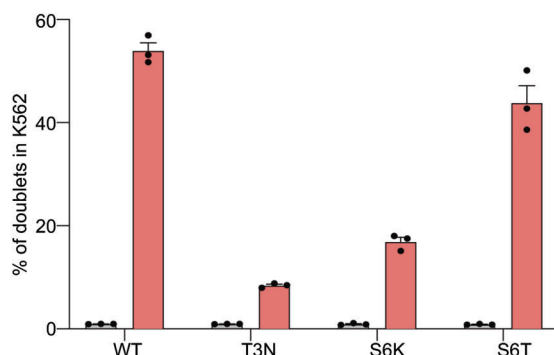
C



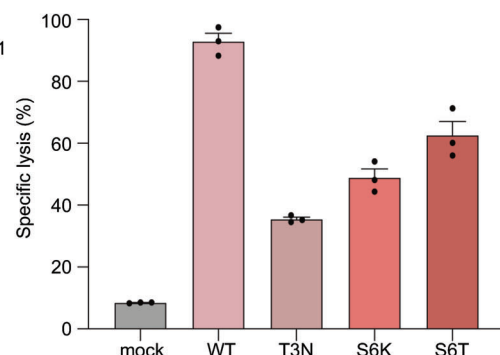
D



E



F



**Figure 6.** The library-on-library screening capability of DoubletSeeker. A) Schematic of DoubletSeeker for library-on-library screening of TCR-pMHC interactions. Single doublets are sorted into individual wells of 96-well plates, which are preloaded with cell lysis buffer, followed by freezing to lyse the doublet cells. Subsequently, the TCR and SCT information for each doublet is obtained through single-cell PCR sequencing. B,C) The proportion of TCR-pMHC pairs within the enriched doublets isolated from the co-incubation of A2-restricted SCT-K562 library (B) or neoepitope SCT-K562 library (C) with a mixture of F5-, 1G4-, GBA-1- and Neo-TCR-Jurkat cells (TCR-expressing cells: SCT-expressing cells = 1:1). D) Proportion of TCR-pMHC

is uncertain. The mammalian cell display system used in DoubletSeeker offers a convenient approach for generating large-scale antigen or TCR libraries while maintaining proper folding and modifications of membrane proteins. This endows DoubletSeeker to effectively enrich doublets that accurately represent genuine interactions. Furthermore, DoubletSeeker provides a reliable means of identifying TCR-antigen interactions through high-throughput “library-on-library” screening, offering distinct advantages in terms of accessibility and screening scale. By expanding the repertoire of TCR-antigen pairings, including the desired information on MHC II-restricted TCR/antigen interactions, DoubletSeeker would offer sufficient training data for studying T cell antigen-specific recognition through *in silico* methods.<sup>[28]</sup>

DoubletSeeker does have certain limitations. First, the time and conditions required for doublets formation may vary depending on the specific receptor-ligand interactions and the types of presenting cells involved. Thus, exploring and optimizing the detection conditions based on the specific experimental requirements is crucial. Second, the formation of doublets is primarily applicable to protein overexpression systems and may not be as effective for screening primary T cells or endogenously presented antigens. Because tumor cells naturally exhibit low levels of endogenously presented antigens on their surfaces, capturing interactions between these cells and primary T cells, which possess diverse T-cell receptor (TCR) repertoires, remains challenging for DoubletSeeker. As a result, there is a need for further improvement in the sensitivity of DoubletSeeker for primary cells. Nonetheless, our research suggests that, given its ability to identify OVA-specific primary OT-1 T cells, DoubletSeeker could serve as an alternative method for enriching antigen-specific T cells from peripheral blood.

Overall, DoubletSeeker represents a significant advancement in the field of intracellular communication by enabling the high-throughput capture of ligand-receptor interactions. It has the potential to revolutionize the study of immune cell communication and has the potential to greatly enhance our understanding of antigen recognition. DoubletSeeker facilitates comprehensive investigations into the adaptive immune response and as such, opens up avenues for the engineering of such responses against disease.

#### 4. Experimental Section

**Cells:** HEK293T (ATCC, CRL-3216), A549 (ATCC, CCL-185), and SK-BR-3 (ATCC, HTB-30) cells were cultured in Dulbecco’s modified Eagle’s medium (DMEM) (Gibco) supplemented with 10% (v/v) fetal bovine serum (FBS) (Gibco) and 1% (v/v) penicillin-streptomycin (P/S) (Gibco). Jurkat E6-1 (ATCC, TIB-152), K562 (ATCC, CCL-243), Raji (ATCC, CCL-86), and J76 cells were cultured in RPMI 1640 medium (Hyclone) supplemented with 10% (v/v) FBS, 1% (v/v) P/S, 0.1 M HEPES (Gibco), 1 mM sodium pyruvate (Gibco), 1% (v/v) non-essential amino acids (NEAA) (Gibco) and 50  $\mu$ M  $\beta$ -mercaptoethanol (Sigma–Aldrich). Primary human

peripheral blood mononuclear cells (PBMCs) from healthy donors were purchased from Sailybio, and cultured in RPMI 1640 medium supplemented with 5% (v/v) human serum (Gemini), 1% (v/v) P/S, 0.1 M HEPES, 1 mM sodium pyruvate, 1% (v/v) NEAA, 1% (v/v) glutaMAX (Gibco), and 50  $\mu$ M  $\beta$ -mercaptoethanol in the presence of recombinant human IL-2 (300 U/ml, Peprotech). All cells were cultured at 37 °C with 5% CO<sub>2</sub>.

**DNA Constructs:** The retroviral vector encoding TCR genes and the lentiviral vector encoding SCT genes were constructed as previously described.<sup>[10]</sup> Briefly, either human or murine TCRs were designed using the LNGFR-P2A-TCR $\alpha$ -F2A-TCR $\beta$  format and integrated into the MSGV retroviral vector. Peptide-MHC SCT composed of antigenic peptide (MART-1/HLA-A\*02:01, ELAGIGILTV; NY-ESO-1/HLA-A\*02:01, SLLMWITQV; NY-ESO-1/HLA-B\*18:01, LEFLAMPF; NY-ESO-1/HLA-C\*03:04, FATPMEAL; USP-7/HLA-A\*02:01, YLYHRVDVI; MART-1 variants/HLA-A\*02:01, AAGIGILTV, AAGIGILTVI, EAAGIGILTV, GIGILTVIL, and ILVILGVL),  $\beta$ 2-microglobulin, and HLA domains via flexible glycine-serine linkers were prepared with a disulfide trap modification expressed in a lentiviral vector encoding an enhanced green fluorescent protein (eGFP). In addition, the MHC-II restricted antigenic peptide (GAD2<sub>555-567</sub>, NFFRMVISNPAAT) was linked with HLA-DRB1 and HLA-DRA genes via flexible glycine-serine linkers and then subcloned into a lentiviral vector. The lentiviral or retroviral vectors encoding a blue fluorescent protein (BFP) or eGFP were used in the construction of various membrane proteins, including GPI-anchored cytokines, CD19-, HER2-, EGFR-CAR, CD8, CD40, CD40L.

**Cell Lines Construction:** Retroviruses and lentiviruses were produced in HEK-293T cells using polyethyleneimine (PEI) (Polysciences) as a transfection reagent, along with packaging plasmid vectors (pRD114 and pHIT60, or psPAX2 and pMD2.G) according to the manufacturer’s protocol. After 48 h of transfection, the viruses were filtered through 0.45  $\mu$ m syringe filters and stored at –80 °C until further use. The cells were spin-infected with the virus and supplemented with 10  $\mu$ g mL<sup>–1</sup> polybrene at 2500 r.p.m. at 30 °C for 90 min. After two days of infection, cells expressing the desired proteins were sorted by FACS.

**Primary T Cell Activation and Retroviral Transduction:** Anti-human CD3 (1  $\mu$ g mL<sup>–1</sup>, Biolegend) and anti-human CD28 (1  $\mu$ g mL<sup>–1</sup>, Biolegend) antibodies were used to activate PBMCs. After 48 h of activation, 1  $\times$  10<sup>6</sup> PBMCs per 24 well plates were spin-infected with the virus and fresh medium containing 300 U mL<sup>–1</sup> IL-2 and 1  $\mu$ g mL<sup>–1</sup> anti-human CD28 was added. LNGFR or mTCR $\beta$  were used to quantify the infection efficiency.

**A2-Restricted SCT and Neoantigen SCT Library Preparation:** The A2-restricted SCT cDNA library and the neoepitopes SCT cDNA library were generated as previously described.<sup>[10]</sup> Briefly, a pool of oligonucleotides encoding epitopes was synthesized and utilized as the template for PCR amplification. The PCR-amplified oligonucleotides were then inserted into a BsmBI-digested lentiviral vector encoding eGFP via flexible glycine-serine linkers along with  $\beta$ 2-microglobulin and HLA-A2 structural domains to form SCT by in-fusion cloning.

**TCR Library Preparation:** To introduce random mutations in the CDR3 $\beta$  region of the TRBV gene in M1W-TCR, 5’-AATTTCCCCTGACTCTCGAGTCGCCAGCCCCAACCAGACCTCTGT ACTTCTGTGCCAG-CAGTTTANNNNNNNNNNNNNNNNNNCCCTCCACTTTGGGAACGGACCGCTCTACTGTGACAGAG-3’ was used as the template and then amplified with 5’-AATTCCCCTGATC-3’ as the forward primer and 5’-CTCTGTACAGTGAG-3’ as the reverse primer. After amplification, the PCR amplicons were digested with XhoI and MluI and then ligated into XhoI and MluI-digested M1W-TCR retroviral vector. The product of the ligation reaction was transformed into TG1 electrocompetent cells by the electroporation transformation method. Subsequently, all the transformed

pairs within the enriched doublets isolated from the co-incubation of the combined library of A2-restricted SCT-K562 and neoepitope SCT-K562 with M1W-TCR pools-J76 cells, in the presence of a small proportion of F5- and 1G4-TCR-Jurkat cells (TCR-expressing cells: SCT-expressing cells = 1:1). E) Interaction between MART-1-K562 or NY-ESO-1-K562 cells and J76 cells expressing M1W<sub>WT</sub>, M1W<sub>T3N</sub>, M1W<sub>S6K</sub>, or M1W<sub>S6T</sub>-TCR as detected by DoubletSeeker. F) Cytotoxicity of T cells expressing M1W<sub>WT</sub>, M1W<sub>T3N</sub>, M1W<sub>S6K</sub>, or M1W<sub>S6T</sub>-TCR against A2-expressing K562 cells pulsed with MART-1 peptide (ELAGIGILTV). The data are presented as percentage-specific lysis rates. Data are represented as the means  $\pm$  SEM. *n* = 3 (E and F).

bacteria were spread on LB agar plates. The plasmid midiprep kit (CW-BIO) was employed to isolate and purify the retroviral plasmid. In the case of M1W-TCR pool-J76 cells used in TCR library screening, the retroviral particle production was mixed with 0.1% WT M1W-TCR virus, and subsequently, CD8-J76 cells were infected with the virus at an MOI of 4. For M1W-TCR pool-J76 cells used in library-on-library screening, the CD8-J76 cells were infected with the virus at an MOI of 1 following retroviral particle production. Then, the cells were sorted to isolate the group that exhibited positive expression of the TCRs. These M1W-TCR pool-J76 libraries were then subjected to further screening and analysis.

**Doublet Formation Assay:** For doublets formation experiments, cells were labeled with indicated CellTrace dyes at 37 °C for 30 min as the experimental design. After the second washing, co-incubations of ligand-expressing cells and receptor-expressing cells were set up in 1.5 ml tubes at a ratio of 1:5 (30 000–50 000 cells in 75  $\mu$ L incubation buffer, PBS containing 0.5% BSA and 1 mM EDTA) and co-incubated for 30 min at 37 °C. Ratios for co-incubation experiments involving TCR-pMHC interactions are indicated in the text or figure legends. 500  $\mu$ L FACS buffer or 4% polyoxymethylene was then added to the co-incubation. The formation of doublets was then analyzed by flow cytometry. For the screening of A2-restricted SCT and neoantigen SCT libraries, co-incubations of SCT library-expressing cells and TCR-expressing cells were set up in 1.5 mL tubes at a ratio of 1:1 (300 000 cells in 200  $\mu$ L incubation buffer). After a 30 min incubation, 800  $\mu$ L of 4% polyoxymethylene was added to the mixture. For M1W-TCR pools-J76 library screening, co-incubations of MART-1-K562 and M1W-TCR pools-J76 library were set up in 1.5 mL tubes at a ratio of 1:1 (300 000 cells in 200  $\mu$ L incubation buffer). After a 30 min incubation, 800  $\mu$ L of FACS buffer was added to the mixture.

**Visualization of Doublets by Confocal Microscopy:** Cells labeled with either CTV or CMFDA were subjected to the doublet formation assay as previously described. The co-incubated cells were then plated on a confocal dish, and the formation of doublets was detected using a confocal microscope (TCS-SP6, Leica) with 405 and 488 lasers.

**Flow Cytometry and Sorting:** Ligand-expressing or receptor-expressing cells were stained with flow cytometry antibodies for 30 min at 4 °C, washed with FACS buffer twice, and then subjected to flow cytometry. The top 10% of cells that exhibited the highest expression of the target proteins were enriched for further analysis. For the cell co-incubation assay, the signal of different CellTrace dyes was measured by flow cytometry.

**PCR Amplification and Deep Sequencing:** Genomic DNA from the sorted SCT-K562 cells from A2-restricted SCT and neoantigen SCT library screening was extracted using the Nucleic acid extraction kit (CW-BIO), and then used as a template for barcoded PCR amplification, with the following program: 95 °C for 2 min; 35 cycles of 95 °C for 20 s, 63 °C for 10 s, 72 °C for 15 s; final extension of 72 °C for 1 min per kb. The primers used in PCR amplification were the following: SCT-Forward, 5'-AATGATACGGCGACCACCGAG ATCTACAATCCTTTCCCTACACGACGCTCTTCCGATCTGGCTGTGCTTGTGGCC-3'; SCT-Reverse, 5'-GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTCCACCA CCGCTACCTC-3'; and adaptor-index reverse primers. The amplified products were purified using DNA, RNA, and protein purification kits (NucleoSpin) according to the manufacturer's protocol, and then were sequenced using an Illumina HiSeq X Ten (Novogene).

**Single Doublets PCR:** Single doublets were sorted using FACS into ice-cold 96-well plates (Nest), with each well containing 20  $\mu$ L lysis buffer. The lysis buffer was supplemented with RNase inhibitor (Transgene) at a concentration of 1 U  $\mu$ L<sup>-1</sup>. Immediately after cell sorting, each plate was covered with an adhesive film (Beyotime), briefly spun down in a centrifuge, and then frozen at -80 °C for further analysis and storage. For the pre-amplification of CDR3 $\beta$  sequences, the primer mix included a 10  $\mu$ M concentration of primers designed to target the TRBV region (forward primer) and the TRBC region (reverse primer) for endogenous transcripts of the TCR $\beta$  chain. Similarly, for the pre-amplification of antigen sequences, the primer mix consisted of a 10  $\mu$ M concentration of primers targeting the signal peptide region (forward primer) and the  $\beta$ 2M region (reverse primer). To prepare the qRT-PCR reaction mixture, the primer mixtures were re-

spectively added into the OneStep qRT-PCR 5  $\times$  reaction mix (Qiagen), and DNase/RNase-free distilled water was added to adjust the final volume to 15  $\mu$ L<sup>-1</sup>. The plate containing the single doublets in the lysis mix was thawed on ice and briefly spun down in a centrifuge. Then, the contents were aliquoted to 10  $\mu$ L per reaction and added into the respective qRT-PCR reaction mix as described. This procedure enabled the amplification of both TCR and antigen sequences for each doublet. The reverse transcription step was performed using a thermal cycler (lid temperature 70 °C) for 30 min at 50 °C. Following this, PCR was performed with the following cycle conditions: 95 °C for 5 min; six cycles of 95 °C for 20 s, 60 °C for 60 s, and 70 °C for 15 s; final extension at 70 °C for 1 min, followed by a 4 °C hold. Nested PCRs for cloning target CDR3 $\beta$  region and antigen sequence were performed using the 2 $\times$  PrimeSTAR (Takara) and nest-designed primers. Cycling conditions were as follows: 95 °C for 2 min; 35 cycles of 95 °C for 20 s, 60 °C for 20 s, 72 °C for 15 s; final extension of 72 °C for 1 min.

**Peptide Loading in Antigen-Presenting Cells:** Lyophilized peptides (LifeTein) were redissolved in dimethylsulfoxide (Sigma-Aldrich) at a concentration of 10 mM. The peptides were then further diluted in water to achieve a final concentration of 100 nM. A2-K562-luciferase cells (50 000 total, 0.5  $\times$  10<sup>6</sup> cells mL<sup>-1</sup>) were pulsed with peptide dilution in a 96-well U-bottom plate and incubated for 2 h at 37 °C. After the incubation, 100  $\mu$ L of the medium was added to each well, and the plate was centrifuged for 5 min at 1500 r.p.m. The cells were subsequently washed once with 200  $\mu$ L of medium and resuspended in 100  $\mu$ L of medium in preparation for the cytotoxicity assay.

**Co-Culture of Primary Human T Cells and Cytotoxicity Assay:** Primary T cells expressing the TCR of interest and A2-K562-luciferase cells pulsed with the corresponding peptide were co-cultured in the wells of a 96-well U-bottom plate. The ratio of effector cells to target cells was maintained at 5:1. After 6 h of co-culture, the plate was centrifuged for 5 min at 1500 r.p.m. The cells were then resuspended in 50  $\mu$ L of PBS and frozen at -20 °C for 1 h to facilitate cell lysis. To assess T-cell mediated cytotoxicity, the intercellular luciferase activity of the surviving target cells was measured. Briefly, the luciferase substrate was added to each well at a 1:1 ratio. The plate was gently agitated in the dark to stabilize the luminescent signal. The luminescence of each well was then measured using a Multifunctional microporous plate reader (SpectraMax). The average values from triplicate wells were calculated, and the percent lysis was determined using the following equation: % specific lysis = 100  $\times$  (maximum fluorescence value - test fluorescence value) / maximum fluorescence value.

**Statistical Analysis:** Statistical analysis was performed using GraphPad Prism version 9.0. An unpaired t-test was used to assess the statistical significance. Data are represented as mean  $\pm$  SEM in bar blots. P values are as indicated in the figure legends.

## Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

## Acknowledgements

The authors would like to thank R. Zhu, X. Pei, Z. Peng, and X. Dai for their technical assistance. The authors would like to thank W. Zhu for performing the data analysis and H. Yang for providing recombinant E4 antibody and its corresponding expression plasmid. This work was supported by the National Natural Science Foundation of China (81972875 and 32270994), the Natural Science Foundation Outstanding Youth Fund of Jiangsu Province (BK20211505), the Non-profit Central Research Institute Fund of Chinese Academy of Medical Sciences (2021-RC310-014 and 2019PT310028), the CAMS Innovation Fund for Medical Sciences (CIFMS) (2023-I2M-QJ-019, 2021-I2M-1-047, 2021-I2M-1-061, and 2022-I2M-2-004), and the Gusu Innovation and Entrepreneurship Talent Program (ZXL2022503) to G.L. and the CAMS Innovation Fund for Medical

Sciences (CIFMS) (2022-I2M-1-021) to Y.L. and Science Fund for Creative Research Groups of the National Natural Science Foundation of China (82221004) to J.W. The authors thank the NCTIB Fund for R&D Platform for Cell and Gene Therapy and the Suzhou Municipal Key Laboratory (SZS2023005).

## Conflict of Interest

The authors declare no conflict of interest.

## Author Contributions

Y.W., Z.W., and J.Y. contributed equally to this work. G.L. conceived the idea. G.L., Z.W., and Y.W. designed the overall experiments, analyzed all the data, and wrote the manuscript. Y.W. and J.Y. performed the experiments. Y.L. analyzed the sequencing data. X.L., L.F., J.W., and G.L. reviewed the article. J.W. and G.L. supervised this research.

## Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## Keywords

doublets, membrane-protein interaction, pMHC, T cell antigen discovery, TCR

Received: August 16, 2023

Revised: January 2, 2024

Published online: February 11, 2024

- [1] E. Armingol, A. Officer, O. Harismendy, N. E. Lewis, *Nat. Rev. Genet.* **2021**, *22*, 71.
- [2] A. Giladi, M. Cohen, C. Medaglia, Y. Baran, B. Li, M. Zada, P. Bost, R. Blecher-Gonen, T.-M. Salame, J. U. Mayer, E. David, F. Ronchese, A. Tanay, I. Amit, *Nat. Biotechnol.* **2020**, *38*, 629.
- [3] Q. Liu, J. Zheng, W. Sun, Y. Huo, L. Zhang, P. Hao, H. Wang, M. Zhuang, *Nat. Methods* **2018**, *15*, 715.
- [4] K. J. Roux, D. I. Kim, M. Raida, B. Burke, *J. Cell Biol.* **2012**, *196*, 801.
- [5] a) A. J. Bankovich, K. C. Garcia, *Immunity* **2003**, *18*, 7; b) N. L. La Gruta, S. Gras, S. R. Daley, P. G. Thomas, J. Rossjohn, *Nat. Rev. Immunol.* **2018**, *18*, 467.
- [6] a) A. K. Sewell, *Nat. Rev. Immunol.* **2012**, *12*, 669; b) J. Nikolich-Zugich, M. K. Slifka, I. Messaoudi, *Nat. Rev. Immunol.* **2004**, *4*, 123.
- [7] a) H. S. Robins, S. K. Srivastava, P. V. Campregher, C. J. Turtle, J. Andriesen, S. R. Riddell, C. S. Carlson, E. H. Warren, *Sci. Transl. Med.* **2010**, *2*, 47ra64; b) T. P. Arstila, A. Casrouge, V. R. Baron, J. Even, J. Kanellopoulos, P. Kourilsky, *Science* **1999**, *286*, 958.
- [8] a) A. K. Bentzen, A. M. Marquard, R. Lyngaa, S. K. Saini, S. Ramskov, M. Donia, L. Such, A. J. S. Furness, N. Mcgranahan, R. Rosenthal, P. T. Straten, Z. Szallasi, I. M. Svane, C. Swanton, S. A. Quezada, S. N. Jakobsen, A. C. Eklund, S. R. Hadrup, *Nat. Biotechnol.* **2016**, *34*, 1037; b) S.-Q. Zhang, K.-Y. Ma, A. A. Schonnesen, M. Zhang, C. He, E. Sun, C. M. Williams, W. Jia, N. Jiang, *Nat. Biotechnol.* **2018**, *36*, 1156.
- [9] a) M. H. Gee, A. Han, S. M. Lofgren, J. F. Beausang, J. L. Mendoza, M. E. Birnbaum, M. T. Bethune, S. Fischer, X. Yang, R. Gomez-Eerland, D. B. Bingham, L. V. Sibener, R. A. Fernandes, A. Velasco, D. Baltimore, T. N. Schumacher, P. Khatri, S. R. Quake, M. M. Davis, K. C. Garcia, *Cell* **2018**, *172*, 549; b) M. E. Birnbaum, J. L. Mendoza, D. K. Sethi, S. Dong, J. Glanville, J. Dobbins, E. Özkan, M. M. Davis, K. W. Wucherpfennig, K. C. Garcia, *Cell* **2014**, *157*, 1073.
- [10] a) A. V. Joglekar, M. T. Leonard, J. D. Jeppson, M. Swift, G. Li, S. Wong, S. Peng, J. M. Zaretsky, J. R. Heath, A. Ribas, M. T. Bethune, D. Baltimore, *Nat. Methods* **2019**, *16*, 191; b) T. Kula, M. H. Dezfulian, C. I. Wang, N. S. Abdelfattah, Z. C. Hartman, K. W. Wucherpfennig, H. K. Lyster, S. J. Elledge, *Cell* **2019**, *178*, 1016; c) J. Kisielow, F.-J. Obermair, M. Kopf, *Nat. Immunol.* **2019**, *20*, 652; d) G. Sharma, C. M. Rive, R. A. Holt, *Nat. Commun.* **2019**, *10*, 4553; e) Z. Liu, J. P. Li, M. Chen, M. Wu, Y. Shi, W. Li, J. R. Teijaro, P. Wu, *Cell* **2020**, *183*, 1117.
- [11] G. Li, M. T. Bethune, S. Wong, A. V. Joglekar, M. T. Leonard, J. K. Wang, J. T. Kim, D. Cheng, S. Peng, J. M. Zaretsky, Y. Su, Y. Luo, J. R. Heath, A. Ribas, O. N. Witte, D. Baltimore, *Nat. Methods* **2019**, *16*, 183.
- [12] B. Yu, Q. Shi, J. A. Belk, K. E. Yost, K. R. Parker, R. Li, B. B. Liu, H. Huang, D. Lingwood, W. J. Greenleaf, M. M. Davis, A. T. Satpathy, H. Y. Chang, *Cell* **2022**, *185*, 4904.
- [13] C. S. Dobson, A. N. Reich, S. Gaglione, B. E. Smith, E. J. Kim, J. Dong, L. Ronsard, V. Okonkwo, D. Lingwood, M. Dougan, S. K. Dougan, M. E. Birnbaum, *Nat. Methods* **2022**, *19*, 449.
- [14] a) R. F. Kudernatsch, A. Letsch, H. Stachelscheid, H.-D. Volk, C. Scheibenbogen, *Cytometry, Part A* **2013**, *83A*, 173; b) J. G. Burel, M. Pomaznoy, C. S. Lindestam Arlehamn, G. Seumois, P. Vijayanand, A. Sette, B. Peters, *Cytometry A* **2020**, *97*, 1127.
- [15] J. G. Burel, M. Pomaznoy, C. S. Lindestam Arlehamn, D. Weiskopf, R. Da Silva Antunes, Y. Jung, M. Babor, V. Schulten, G. Seumois, J. A. Greenbaum, S. Premawansa, G. Premawansa, A. Wijewickrama, D. Vidanagama, B. Gunasena, R. Tippalagama, A. D. Desilva, R. H. Gilman, M. Saito, R. Taplitz, K. Ley, P. Vijayanand, A. Sette, B. Peters, *eLife* **2019**, *8*, e46045.
- [16] a) E. García-Guerrero, L. I. Sánchez-Abarca, E. Domingo, T. L. Ramos, J. A. Bejarano-García, J. A. Gonzalez-Campos, T. Caballero-Velázquez, J. A. Pérez-Simón, *Front. Immunol.* **2018**, *9*, 1971; b) A. Chudnovskiy, G. Pasqual, G. D. Victoria, *Curr. Opin. Immunol.* **2019**, *58*, 24; c) M. Cohen, A. Giladi, O. Barbo, P. Hamon, B. Li, M. Zada, A. Gurevich-Shapiro, C. G. Beccaria, E. David, B. B. Maier, M. Backup, I. Kamer, A. Deczkowska, J. Le Berichel, J. Bar, M. Iannacone, A. Tanay, M. Merad, I. Amit, *Nat. Cancer* **2022**, *3*, 303; d) B. Akkaya, O. Kamenyeva, J. Kabat, R. Kissinger, *Methods Mol. Biol.* **2021**, *2304*, 243.
- [17] a) M. L. Dustin, *Immunol. Rev.* **2008**, *221*, 77; b) D. J. Irvine, M. A. Purbhoo, M. Krogsgaard, M. M. Davis, *Nature* **2002**, *419*, 845; c) A. Grakoui, S. K. Bromley, C. Sumen, M. M. Davis, A. S. Shaw, P. M. Allen, M. L. Dustin, *Science* **1999**, *285*, 221.
- [18] R. Yan, Y. Zhang, Y. Li, L. Xia, Y. Guo, Q. Zhou, *Science* **2020**, *367*, 1444.
- [19] a) M. Hoffmann, S. Pöhlmann, *Cell Res.* **2022**, *32*, 1; b) Q. Lu, J. Liu, S. Zhao, M. F. Gomez Castro, M. Laurent-Rolle, J. Dong, X. Ran, P. Damani-Yokota, H. Tang, T. Karakousi, J. Son, M. E. Kaczmarek, Z. Zhang, S. T. Yeung, B. T. Mccune, R. E. Chen, F. Tang, X. Ren, X. Chen, J. C. C. Hsu, M. Teplova, B. Huang, H. Deng, Z. Long, T. Mudianto, S. Jin, P. Lin, J. Du, R. Zang, T. T. Su, et al., *Immunity* **2021**, *54*, 1304.
- [20] L. Cantuti-Castelvetri, R. Ojha, L. D. Pedro, M. Djannatian, J. Franz, S. Kuivanen, F. Van Der Meer, K. Kallio, T. Kaya, M. Anastasina, T. Smura, L. Levanov, L. Szirovicza, A. Tobi, H. Kallio-Kokko, P. Österlund, M. Joensuu, F. A. Meunier, S. J. Butcher, M. S. Winkler, B. Mollenhauer, A. Helenius, O. Gokce, T. Teesalu, J. Hepojoki, O. Vapalahti, C. Stadelmann, G. Balistreri, M. Simons, *Science* **2020**, *370*, 856.
- [21] S. Wang, Z. Qiu, Y. Hou, X. Deng, W. Xu, T. Zheng, P. Wu, S. Xie, W. Bian, C. Zhang, Z. Sun, K. Liu, C. Shan, A. Lin, S. Jiang, Y. Xie, Q. Zhou, L. Lu, J. Huang, X. Li, *Cell Res.* **2021**, *31*, 126.
- [22] Y. Li, Z. Zhang, L. Yang, X. Lian, Y. Xie, S. Li, S. Xin, P. Cao, J. Lu, *iScience* **2020**, *23*, 101160.
- [23] J. M. Condor Capcha, C. Iansen Irion, G. Lambert, A. Chahdi, P. Buchwald, D. Dykxhoorn, H. Farzaneh, S. Deo, S. Daunert, K. C. Young, L. A. Shehadeh, *Circulation* **2021**, *144*, A11197.
- [24] a) A. Leonardi, U. Rosani, P. Brun, *Ocul. Immunol. Inflamm.* **2020**, *28*, 735; b) S. M. Brockbank, J. Soden, R. Faba-Rodriguez, L. R. Ribeiro, C. Geh, H. Thomas, J. Delight, L. Coverley, W. M. Abbott, J. Freeth,

- bioRxiv* **2021**; c) N. Wanner, G. Andrieux, P. Badia-I-Mompel, C. Edler, S. Pfefferle, M. T. Lindenmeyer, C. Schmidt-Lauber, J. Czogalla, M. N. Wong, Y. Okabayashi, F. Braun, M. Lütgehetmann, E. Meister, S. Lu, M. L. M. Noriega, T. Günther, A. Grundhoff, N. Fischer, H. Bräuninger, D. Lindner, D. Westermann, F. Haas, K. Roedl, S. Kluge, M. M. Addo, S. Huber, A. W. Lohse, J. Reiser, B. Ondruschka, J. P. Sperhake, et al., *Nat. Metabolism* **2022**, 4, 310.
- [25] a) E. Corse, R. A. Gottschalk, J. P. Allison, *J. Immunol.* **2011**, 186, 5039; b) J. D. Stone, A. S. Chervin, D. M. Kranz, *Immunology* **2009**, 126, 165; c) P. D. Holler, P. O. Holman, E. V. Shusta, S. O'herrin, K. D. Wittrup, D. M. Kranz, *Proc. Natl. Acad. Sci. USA* **2000**, 97, 5387.
- [26] Y. Y. L. Yu, N. Netuschil, L. Lybarger, J. M. Connolly, T. H. Hansen, *J. Immunol.* **2002**, 168, 3145.
- [27] E. L. Reinherz, J.-H. Wang, *Trends Immunol.* **2015**, 36, 300.
- [28] D. Hudson, R. A. Fernandes, M. Basham, G. Ogg, H. Koohy, *Nat. Rev. Immunol.* **2023**, 23, 511.
- [29] A. V. Joglekar, G. Li, *Nat. Methods* **2021**, 18, 873.