# CRISPR-targeted display of functional T cell receptors enables engineering of enhanced specificity and prediction of cross-reactivity

Rodrigo Vazquez-Lombardi[1], Johanna S. Jung, Florian Bieberich, Edo Kapetanovic, Erik Aznauryan, Cédric R. Weber and Sai T. Reddy[1]

Department of Biosystems Science and Engineering, ETH Zurich

Mattenstrasse 26, 4058 Basel, Switzerland

[1]To whom correspondence should be addressed

rodrigva@ethz.ch, sai.reddy@ethz.ch

## ABSTRACT

**T cell receptor (TCR) gene therapy is a promising cell therapy approach for the treatment of cancer. However, most naturally occurring TCRs display low affinities to their peptide-MHC targets, and engineering of TCRs for enhanced affinity is complicated by the risk of introducing cross-reactivity and the poor correlation between affinity and function. Here we report the establishment of the TCR-accepting T cell (TnT) platform through five sequential CRISPR-Cas9 genome editing steps of a human T cell line, and demonstrate its application for functional engineering of TCRs and prediction of cross-reactivity. Using the TnT platform, we profile the mutational landscapes of tumor-specific TCRs at high-throughput to reveal a substantial discordance between antigen binding and antigen-induced signaling. Furthermore, we combine CRISPR-targeting, functional selection and deep sequencing to screen TCR mutagenesis libraries and identify variants with enhanced recognition of the cancer-testis antigen MAGE-A3. Finally, functional cross-reactivity profiling using TnT cells was able to accurately predict off-targets and identify engineered TCRs with exquisite specificity to MAGE-A3. Thus, the TnT platform represents a valuable technology for the engineering of TCRs with enhanced functional and safety profiles.**

## INTRODUCTION

T cell receptor (TCR) gene therapy is a promising cell therapy approach against cancer that relies on viral transfer of genes encoding the α and β chains of tumor-reactive TCRs into autologous T cells, followed by their expansion and re-infusion into patients[1]. Unlike chimeric antigen receptor (CAR)-T cell therapies, which target surface tumor antigens, TCR-redirected T cells recognize processed intracellular tumor antigen peptides displayed by major histocompatibility complex (MHC) / human leukocyte antigen (HLA) molecules. This key feature vastly increases the number of potential antigen targets and is thought to allow for more efficient infiltration into solid tumors[2], a known limitation of CAR-T cells[3]. As such, TCR gene therapy has been identified as an effective approach for driving durable responses against multiple cancers[2,4–12]. Despite its significant promise, the discovery, engineering and selection of optimal therapeutic TCRs remains a time-consuming process complicated by the low affinities of TCRs to their peptide-MHC targets and the inherent risk of TCR cross-reactivity[6,13,14].

Different from antibodies, which typically recognize a single epitope, TCRs are able to recognize multiple peptide antigens presented by MHC[15]. For example, naturally-occurring TCRs have been shown to display a wide range of specificity profiles[16], in some cases having the potential to recognize up to a million different peptides[17]. Therefore, cross-reactivity is vital to measure and subsequently avoid in order to engineer safe TCR gene therapies. In addition to specificity, TCR affinity and potency (i.e., functional avidity) are crucial parameters to consider when selecting optimal TCRs for gene therapy applications. Notably, TCR specificity, affinity and function are not necessarily correlated with each other. For example, while TCRs with high affinities to antigen (1-5 μM) tend to display high potency *in vitro*, TCRs with lower affinities (5-100 μM) often display a poor correlation between affinity and function[18,19]. In addition, engineered TCRs with supraphysiological affinities (< 1 μM) may display sub-optimal therapeutic activity due to increased T cell dysfunction[20,21], inability to undergo serial TCR triggering[22], and potential for reactivity against the presenting HLA molecule[23,24].

Due to central and peripheral negative selection processes, naturally occurring high-affinity TCRs targeting self-antigens are extremely rare and their discovery requires extensive screening[25–27]. To overcome this, directed evolution and protein engineering methods that rely on display technologies (e.g., phage and yeast display) have been employed[28,29]. These require reformatting of TCRs into synthetic single-chain fragments, which are then screened on the surface of phage or yeast for binding towards peptide-MHC multimers. In one notable example, phage display was used to increase the affinity of a TCR targeting the cancer-testis antigen MAGE-A3. When applied as a TCR gene therapy, unexpected cross-reactivity towards a self-antigen expressed by beating cardiomyocytes cells was observed, which ultimately resulted in treatment-induced patient deaths in a clinical trial[14,30]. In light of this outcome, the development of TCR display platforms enabling simultaneous TCR engineering and detection of cross-reactivity on the basis of function (i.e., antigen-induced signaling) would be highly desirable. While a number of TCR engineering methods have been developed in mammalian cell lines[25,31–34] and primary T cells[35,36], these have only reported selections based on antigen binding. In addition, their use of viral transduction or plasmid transfection for TCR reconstitution is associated with limitations including random integration, constitutive expression and possible expression of different TCRs by a single cell.

Several technologies allowing the assessment of TCR specificity have been reported recently, highlighting a strong interest in the development and safety screening of TCRs for gene therapy. Affinity-based methods include the use of barcoded peptide-MHC multimer libraries for profiling TCRs displayed on primary T cells[16,37,38] and the assessment of soluble TCR interaction with peptide-MHC libraries displayed on the surface of yeast[18,39–41]. A potential limitation of such methods, however, is that they are unable to detect very low-affinity interactions between TCR and peptide-MHC that are nevertheless functional[42]. Epitope screening platforms based on cellular function include those relying on the display of peptide libraries by cells expressing chimeric MHC receptors[43,44], cells undergoing trogocytosis[45], or cells harboring reporters of exogenous granzyme activity[46,47]. Another method for the functional assessment of TCR cross-reactivity involves measuring T cell reactivity to single amino acid mutants of the target peptide, which has been recently utilized to profile the cross-reactivity of phage display-engineered TCRs reformatted for expression in primary T cells[48,49] or a murine cell line[31]. While all of the above methods focus on specificity screening of TCRs, they have not been reported to directly enable TCR engineering.

Here we report the development and application of the TCR-accepting T cell (TnT) platform for the functional display, engineering and cross-reactivity profiling of TCRs. Reconstitution of Jurkat-derived TnT cells with transgenic TCRs was targeted by CRISPR-Cas9 to the endogenous TCRβ genomic region, specifically to the recombined complementary determining region 3 β (CDR3β) sequence, thus providing a monoallelic target that ensured a single integration event in every cell and physiological expression of transgenic TCRs. We use this approach to characterize > 30 individual TCRs and perform functional screening of 437 single amino acid variants and ~150,000 combinatorial variants. We dissect the mutational landscapes of two TCRs to reveal contrasting patterns and substantial discordance between antigen binding and antigen-induced signalling. Furthermore, through a combination of positive and negative functional selection steps coupled with deep sequencing, we identify combinatorial TCR variants with enhanced recognition of the MAGE-A3 cancer-testis antigen and lacking cross-reactivity to a known off-target peptide. Finally, we use peptide scanning to evaluate the cross-reactivity of resulting engineered TCRs displayed on TnT cells, leading to the identification of TCR variants targeting MAGE-A3 antigen with exquisite specificity and minimal cross-reactivity, which thus represent therapeutic candidates for TCR gene therapy.

## RESULTS

### Development of a TCR-accepting T cell (TnT) platform for the functional display of transgenic TCRs

The development of the TnT platform required extensive, multistep CRISPR-Cas9 genome editing. Using the human Jurkat E6-1 leukemia T cell line as a starting point, several genomic components were sequentially edited in order to facilitate the introduction and functional screening of TCRs of different specificities (Fig. 1a and Supplementary Fig. 1). Each genome editing step consisted of transfection with a gene-targeting guide RNA (gRNA) and, when required, a homology-directed repair (HDR) template encoding desired transgenes. This was followed by single-cell fluorescence-activated cell sorting (FACS), cell expansion and clone validation by flow cytometry and Sanger sequencing. In the first step, we equipped the Jurkat T cell line with constitutive Cas9 and human CD8 expression by CRISPR-Cas9 HDR targeting the CCR5 safe harbor locus[50] (Supplementary Figs. 2a-e). This was performed in order to simplify and increase genome editing efficiency[51], and to allow screening of CD8+ T cell-derived TCRs recognizing MHC class I-restricted peptides. In line with this, in the second step we knocked out the endogenous Jurkat CD4

co-receptor by CRISPR-Cas9 non-homologous end joining (NHEJ) (Supplementary Figs. 3a-c). In the third step, we introduced an NFAT-GFP construct into the AAVS1 safe harbor locus through CRISPR-Cas9 HDR, which provides a fluorescence reporter of TCR signaling and activation (Supplementary Figs. 4a-e). Notably, our design incorporated a promoter-less mRuby cassette that acted as a PPP1R12C gene-trap[52] and served to identify successfully edited cells. In the fourth step, we targeted the endogenous Jurkat TCRα chain for knockout through CRISPR-Cas9 NHEJ, leading to the generation of a cell line with abolished surface expression of the TCR-CD3 complex (Supplementary Figs. 5a-f). In addition to eliminating the possibility of transgenic TCR chains mispairing with the endogenous Jurkat TCRα chain, this approach allows us to use restoration of CD3 surface expression as a selectable marker for successful integration of transgenic TCRs (Figs. 1a-c). In the final step, we knocked out expression of the Fas cell surface death receptor (Fas) by CRISPR-Cas9 NHEJ in order to provide our platform with resistance to activation-induced cell death (AICD) (Supplementary Figs. 6a-c). This resulting cell line constitutively expresses Cas9, human CD8 and mRuby, harbors an NFAT-GFP reporter of TCR signaling, and lacks expression of CD4, endogenous TCR and Fas, and thus represents the TnT platform used throughout the rest of this study.

We next reconstituted TnT cells with transgenic TCRs in a manner that would allow for monoallelic, homogenous and physiological TCR expression. To this end, we used CRISPR-Cas9 HDR to target integration of TCR transgenes to the recombined Jurkat TCRβ locus. First, we confirmed that the Jurkat T cell line expresses a single TCRβ chain by performing template-switching RT-PCR and Sanger sequencing (Supplementary Figs. 7a-b). We then designed and validated a gRNA molecule targeting the CDR3β of the Jurkat TCRβ chain (Supplementary Fig. 7c). Since the CDR3β sequence arises from the allele-independent recombination of V-, D- and J-genes, it provides a genomic target that is both highly specific and monoallelic. Having identified a suitable gRNA, we then proceeded to design TCRαβ HDR templates for targeted TCR reconstitution in TnT cells. We selected three previously discovered TCRs recognizing HLA-A*0201-restricted tumor-associated antigens: $TCR_{1G4}$ with specificity to NY-ESO-1, and $TCR_{DMF4}$ and $TCR_{DMF5}$ with specificity to MART-1[53] (Supplementary Table 1). HDR templates contained sequences encoding TCRα variable (VJα) and constant (TRAC) domains, a self-processing T2A peptide and a TCRβ variable (VDJβ) domain, flanked by ~ 900 bp homology arms mapping to the recombined Jurkat TCRβ locus (Fig. 1b). The lack of a constant TCRβ domain (TRBC) in the designed constructs made splicing with endogenous Jurkat TRBC exons a requirement for transgenic TCR expression. This feature allowed us to detect targeted genomic integration based on restored surface expression of CD3 following transfection of TnT cells with CDR3β gRNA and designed HDR templates (PCR product) (Fig. 1c). Furthermore, our strategy ensured that cells displaying restored CD3 expression underwent knockout of the endogenous Jurkat TCRβ chain, as integration of transgenic TCRs relied on the introduction of a dsDNA break at the targeted CDR3β genomic region. Although our initial experiments yielded HDR efficiencies of 1-2%, the development of an enhanced transfection protocol led to HDR efficiencies in the range of 5-20% (Fig. 3c and Supplementary Fig. 8). Targeted TCR reconstitution for the generation of TnT-TCR cells was further validated by detecting binding to cognate peptide-MHC dextramer using flow cytometry (Fig. 1d), PCR amplification of the Jurkat TCR genomic locus (Fig. 1e) and RT-PCR using reverse primers annealing to endogenous Jurkat TRBC sequences (Supplementary Figs. 9a and 9b). We thus successfully generated TnT-$TCR_{1G4}$, TnT-$TCR_{DMF4}$ and TnT-$TCR_{DMF5}$ cell lines that were then subjected to further functional validation.

In order to assess antigen-induced signaling in TnT-TCR cells, we performed co-culture experiments using the HLA-A*0201-positive T2 cell line[54]. Co-culture of TnT-$TCR_{1G4}$ cells with unpulsed (no peptide) T2 cells yielded no detectable expression of NFAT-GFP, while co-culture with T2 cells pulsed with NY-ESO-$1_{157-165}$ cognate peptide induced robust expression of the NFAT-GFP reporter (Fig. 1f). To further assess the specificity of our platform, we performed co-cultures of TnT, TnT-$TCR_{1G4}$, TnT-$TCR_{DMF4}$ and TnT-$TCR_{DMF5}$ cells with T2 cells pulsed with NY-ESO-$1_{157-165}$ peptide, MART-$1_{26-35(2L)}$ peptide or no peptide. We found that NFAT-GFP expression was fully restricted to correct TCR-peptide pairings, with no detectable NFAT-GFP expression across negative controls (Fig. 1g).

We next compared the binding and functional avidities of TnT-$TCR_{1G4}$, TnT-$TCR_{DMF4}$ and TnT-$TCR_{DMF5}$ cells (Fig. 1h-i). TnT-$TCR_{DMF5}$ cells displayed the highest binding avidity to their target peptide-MHC (EC50 = 7 pM), followed by TnT-$TCR_{1G4}$ (EC50 = 57 pM) and TnT-$TCR_{DMF4}$ (EC50 = 267 pM), with picomolar EC50 values reflecting the multivalent nature of surface-displayed TCRs and peptide-MHC dextramers (10-20 peptide-MHC copies per molecule). We evaluated the functional avidity of TnT-TCR cells following co-culture with T2 cells pulsed with serial dilutions of cognate peptide, which revealed a dose-dependent response in terms of NFAT-GFP-expressing cells (Fig. 1i). TnT-$TCR_{DMF5}$ displayed the highest functional avidity (EC50 = 6 nM), followed by TnT-$TCR_{DMF4}$ (EC50 = 11 nM) and TnT-$TCR_{1G4}$ (EC50 = 90 nM). In contrast to their 39-fold difference in binding to peptide-MHC dextramer, the difference between TnT-$TCR_{DMF4}$ and TnT-$TCR_{DMF5}$ in terms of functional avidity was only 2-fold. This indicated that a high binding avidity is not a requirement for antigen-induced expression of NFAT-GFP in TnT-TCR cells. Finally, additional co-culture experiments confirmed TnT-TCR resistance to AICD (Supplementary Fig. 6d) and physiological down-regulation of surface TCR-CD3 expression with increasing amounts of presented antigen[56] (Supplementary Fig. 6e).

**Deep mutational scanning reveals the expression, binding and functional landscapes of TCRs**

In contrast to previous TCR engineering methods[28,29], the TnT platform allows us to assess TCRs across multiple parameters which include their surface expression in complex with CD3, binding to peptide-MHC multimers and signaling in response to antigen-presentation. We selected the MART-1-specific $TCR_{DMF4}$ and the MAGE-A3-specific $TCR_{A3}$ for comprehensive profiling using the TnT platform (Fig. 2a). $TCR_{A3}$ is a low-avidity TCR isolated from a melanoma patient treated with a viral vaccine encoding the HLA-A*0101-restricted MAGE-$A3_{168-176}$ peptide[57,58] (Supplementary Table 1). Co-culture experiments with the MAGE-A3-positive EJM myeloma cell line revealed that the avidity of $TCR_{A3}$ was not sufficient to induce NFAT-GFP expression in TnT-$TCR_{A3}$ cells. However, we detected increased surface expression of the early T cell activation marker CD69 following co-culture, which provided us with a complementary and more sensitive functional readout (Supplementary Fig. 10c).

Having characterized parental $TCR_{DMF4}$ (Fig. 1g-i) and $TCR_{A3}$, we aimed to determine their mutational landscapes across the multiple parameters provided by TnT cells. To this end, we performed deep mutational scanning (DMS) of CDR3β as this TCR region is typically enriched for direct contacts to peptide antigen rather than to MHC[55]. To generate DMS libraries we introduced NNK degenerate codons at each

CDR3β position using plasmid nicking saturation mutagenesis[59] (Supplementary Figs. 11a-c). Plasmid DMS libraries were then used to generate TCRαβ HDR templates (PCR products), which were transfected into TnT cells alongside CDR3β gRNA (Supplementary Fig. 11b). TnT cells displaying restored TCR-CD3 surface expression were isolated by FACS (SEL 1), expanded and re-sorted based on binding to cognate peptide-MHC dextramer (SEL 2A) or activation following co-culture with cells displaying target peptides (SEL 2B) (Fig. 2b and Supplementary Figs. 12a-c). Due to the low avidity of the parental $TCR_{A3}$, functional selections of $TCR_{A3}$ DMS libraries were based on CD69 expression, instead of NFAT-GFP. Following FACS, deep sequencing was performed to identify TCR variants that were enriched across selection steps.

Sequence enrichment analysis was performed relative to observed variant frequencies in plasmid DMS libraries of $TCR_{A3}$ (Fig. 2c) and $TCR_{DMF4}$ (Fig. 2d). We found that wild-type TCR sequences (Figs. 2c and 2d, boxed residues) were enriched in most selections, which likely reflected the depletion of unfavorable variants. The only exception occurred in $TCR_{A3}$ SEL 2B (signaling), in which wild-type $TCR_{A3}$ was depleted relative to the original plasmid library. This was in agreement with the low levels of antigen-induced signaling found in TnT-$TCR_{A3}$ cells (Supplementary Fig. 12c), and pointed towards the occurrence of $TCR_{A3}$ single mutants with substantial increases in function relative to wild-type $TCR_{A3}$. In terms of TCR-CD3 expression (SEL 1), we detected sequencing reads originating from all but five $TCR_{A3}$ variants, indicating that the vast majority of $TCR_{A3}$ mutants were capable of cell surface expression in complex with CD3 (Fig. 2c). This was in stark contrast to the $TCR_{DMF4}$ surface expression landscape, which was heavily restricted (Fig. 2d). As expected, the mutational landscapes for peptide-MHC binding (SEL 2A) and antigen-induced signaling (SEL 2B) were more restricted than those seen for TCR-CD3 surface expression (Figs. 2c and 2d). In the case of $TCR_{A3}$, while a similar number of enriched variants was observed in binding (SEL 2A) and signaling (SEL 2B) selections, there was little correlation between the deep sequencing enrichment levels of individual variants in the two fractions (Fig. 2c and Supplementary Fig. 13a). Some $TCR_{A3}$ variants, however, did display high levels of sequence enrichment for both parameters (e.g., S3G, P4L, M6L and A7V) (Supplementary Fig. 13a). In the case of $TCR_{DMF4}$, there was a modest correlation between enrichment levels of individual variants in binding (SEL 2A) and signaling (SEL 2B) fractions (Supplementary Fig. 13b). However, a very limited number of $TCR_{DMF4}$ variants displayed enrichment levels that were higher than wild-type $TCR_{DMF4}$ for either binding or signaling. Notably, the number of enriched $TCR_{DMF4}$ variants in the signaling fraction (SEL 2B) was considerably higher than in the binding fraction (SEL 2A). This indicated that antigen-induced signaling may occur in the absence of detectable peptide-MHC binding in some of these $TCR_{DMF4}$ variants, particularly those with substitutions in CDR3β positions 2-4 (Fig. 2d).

In order to further characterize the discordance observed between antigen binding and antigen-induced signaling, we reconstituted TnT cells with variants showing enrichment above wild-type TCR levels. For $TCR_{A3}$, we selected nine variants (S3G, P4F, P4L, P4M, P4T, N5C, M6L, A7V, Q10Y) and included the high-affinity $TCR_{a3a}$ as a positive control (see next section for more information on $TCR_{a3a}$). For $TCR_{DMF4}$, we selected six variants (I2H, I2L, E4F, E4Q, E4P, E4V) and included the high-affinity $TCR_{DMF5}$ as a positive control. All selected TCR variants were successfully expressed on the surface of TnT cells and were assessed for binding with cognate peptide-MHC dextramer (Fig. 2e, g top row) and antigen-induced signaling (Fig. 2e, g bottom row). Notably, the majority of $TCR_{A3}$ variants displayed enhancements in binding

and signaling relative to wild-type $TCR_{A3}$. Consistent with sequence enrichment data, variants S3G, P4L, M6L and A7V displayed the largest increases in terms of both binding and signaling (Fig. 2f). For $TCR_{DMF4}$ variants, only modest improvements were observed relative to wild-type (Fig. 2i), namely enhanced binding of variants I2L, E4V and E4F to MART-1 peptide-MHC dextramer (Fig. 2 h). Surprisingly, $TCR_{DMF4}$ variants I2H and I2L showed similar levels of antigen-induced signaling despite I2H showing no detectable binding to MART-1 peptide-MHC dextramer (Figs. 2g and 2h), further highlighting the discordance between binding and signaling in certain TCRs.

**Functional selection of combinatorial CDR3β libraries identifies TCRs with enhanced target specificity and reactivity**

Next, we aimed to engineer the low-avidity $TCR_{A3}$ for enhanced reactivity to the cancer-testis antigen MAGE-A3, as it represents an attractive target that is widely expressed in several types of epithelial tumors[60]. In a previous effort, $TCR_{A3}$ was engineered by phage display to generate $TCR_{a3a}$, a variant with >200-fold higher affinity to MAGE-$A3_{168-176}$ peptide-MHC[14,30] (Supplementary Fig. 10a and Supplementary Table 1). However, use of $TCR_{a3a}$ in a gene therapy clinical trial resulted in two treatment-induced patient deaths[14]. Retrospective analysis uncovered that $TCR_{a3a}$ cross-reactivity to a peptide derived from the protein titin, which is highly expressed in beating cardiomyocytes, led to fatal cardiac toxicity[30]. Notably, this occurred despite the titin-derived peptide (ESDPIVAQY) having four amino acid differences relative to MAGE-$A3_{168-176}$ (EVDPIGHLY). Consistent with previous findings, we found that TnT-$TCR_{A3}$ cells showed low but detectable binding and TnT-$TCR_{a3a}$ showed high binding to MAGE-A3 peptide-MHC dextramer, while only TnT-$TCR_{a3a}$ displayed binding to titin peptide-MHC dextramer (Supplementary Fig. 10b).

We sought to generate and screen combinatorial CDR3β libraries using the TnT platform in order to identify TCR variants with enhanced antigen specificity and reactivity. Combinatorial library design was based on data obtained from $TCR_{A3}$ DMS (Fig. 2c), as this allowed us to maximize both library functionality and size for screening in mammalian cells[51]. By utilizing a sequence space optimization algorithm[51], we designed degenerate codons that mimic the $TCR_{A3}$ single mutant frequencies found in DMS binding and signaling selections (Supplementary Fig. 14), leading to a combinatorial library diversity of $2.6 \times 10^5$ (Fig. 3a). The designed $TCR_{A3}$ combinatorial library was generated by overlap-extension PCR, and cloned into a $TCR_{A3}$-encoding plasmid for bacterial transformation (Supplementary Fig. 15a). We estimated that the physical plasmid library contained as many as $1.5 \times 10^5$ $TCR_{A3}$ variants based on the number of bacterial transformants and the proportion of unique clones after deep sequencing (Supplementary Fig. 15b). The constructed $TCR_{A3}$ plasmid library was used to generate HDR templates, which were then integrated into TnT cells by CRISPR-Cas9 HDR.

The selection strategy for the TnT-$TCR_{A3}$ library consisted of a combination of positive and negative selection steps based on binding to peptide-MHC dextramers and signaling in response to cell-displayed antigens, with deep sequencing of TCR regions performed after every selection step (Fig. 3b). In the first selection step, TnT cells displaying restored CD3 expression were isolated by FACS and expanded (SEL 1, Fig. 3c). Next, expanded cells from SEL 1 were co-cultured overnight with MAGE-A3-positive EJM cells and co-stained with MAGE-A3 and titin peptide-MHC dextramers. Different to wild-type TnT-$TCR_{A3}$ cells, a fraction of TnT-$TCR_{A3}$ library cells displayed robust NFAT-GFP expression (Fig. 3d). Analysis of the NFAT-

GFP-positive population revealed that ~ 20% of cells that bound to MAGE-A3 also recognized the titin peptide-MHC dextramer (Fig. 3d). In order to exclude these cross-reactive variants, we isolated cells that were NFAT-GFP-positive, MAGE-A3 peptide-MHC dextramer-positive and titin peptide-MHC dextramer-negative (SEL 2). A final functional selection step was performed by co-culturing expanded cells from SEL 2 with Colo 205 cells pulsed with either MAGE-A3 or titin peptides. As expected, we observed a substantial fraction of TnT-TCR cells with NFAT-GFP expression after co-culture with MAGE-A3-pulsed Colo 205 cells (Fig. 3e). Interestingly, we also found that a subset of TnT-TCR cells still expressed NFAT-GFP in response to titin stimulation despite the exclusion of titin peptide-MHC binders in SEL 2. Analysis of the sequence space landscape revealed a number of notable trends, including the preference of alanine codons at position CDR3β-1 for TCR:CD3 expression (SEL 1), and the increased frequency of specific residues at CDR3β-6 and CDR3β-10 in MAGE-A3-based selections (SEL 2 and SEL 3A). Most notably, we observed a substantial increase in the frequency of glutamate residues at CDR3β-6 in SEL 3B (titin-induced signaling), highlighting this substitution as a potential determinant of titin cross-reactivity.

We next analyzed deep sequencing data for clone enrichment (Supplementary Table 2). We identified a total of 195 unique clones displaying greater than 2-fold enrichment in SEL 3A (MAGE-A3 signaling) and that were either absent or not enriched in SEL 3B (titin signaling). We scored these clones based on their frequency and enrichment levels and identified the top 29 candidate $TCR_{A3}$ variants, of which we selected 14 for further characterization (Fig. 3g and Supplementary Table 3). To this end, we reconstituted TnT cells with wild-type $TCR_{A3}$, $TCR_{a3a}$ and selected $TCR_{A3}$ variants via CRISPR-Cas9 HDR and assessed them for CD3 expression, binding to MAGE-A3 and titin peptide-MHC dextramers. All selected $TCR_{A3}$ variants were successfully displayed on the surface of TnT cells and showed enhanced MAGE-A3 dextramer binding relative to wild-type TnT-$TCR_{A3}$ cells (Fig 3h). Importantly, binding to titin peptide-MHC dextramer was undetectable in most of the selected $TCR_{A3}$ variants, while readily detected in TnT cells displaying the phage display-engineered $TCR_{a3a}$ (Fig 3i).

**The TnT platform enables TCR cross-reactivity profiling and accurate prediction of off-target antigens**

Since engineering TCRs for enhanced reactivity carries the risk of introducing unwanted specificities, we decided to apply the TnT platform for the screening of TCR cross-reactivity. As a proof of concept, we profiled the cross-reactivity of the phage display-engineered $TCR_{a3a}$[30], which has titin as a known off-target, by performing deep mutational scanning (DMS) of the MAGE-A3 target peptide. Accordingly, we designed a synthetic peptide DMS library containing every possible single amino acid mutant of $MAGE\text{-}A3_{168\text{-}176}$ (EVDPIGHLY) (n = 171). Each library peptide was then individually pulsed on Colo 205 cells (HLA-A*0101-positive, MAGE-A3-negative) and co-cultured with TnT-$TCR_{a3a}$ cells (n = 171 co-cultures). Activation of TnT-$TCR_{a3a}$ cells was assessed by NFAT-GFP expression, which was corrected for background and normalized to the response induced by Colo 205 cells pulsed with wild-type MAGE-A3 peptide (Fig. 4a). In agreement with a previous glycine-serine scan[30], we found that most mutations at peptide-MHC positions 1, 3, 4, 5 and 9 of the target MAGE-A3 peptide resulted in substantially reduced TnT-$TCR_{a3a}$ activation. By contrast, several peptides with mutations at positions 2, 6, 7 and 8 induced strong activation of TnT-$TCR_{a3a}$ cells.

In order to predict potential off-targets, we generated peptide sequence space motifs of allowed substitutions at discrete thresholds of TnT-TCR$_{a3a}$ activation, and used them to interrogate the UniProtKB database[48,49]. As expected, the number of unique peptide hits resulting from these searches decreased with increasing activation thresholds (Fig. 4b), as this reflected the use of more restricted peptide motifs (Supplementary Table 4). The only human peptide returned at the highest activation threshold (100%) was indeed the target MAGE-A3$_{168\text{-}176}$ (EVDPIGHLY) (Fig. 4b). Remarkably, the human peptide returned at the second highest activation threshold was the titin peptide ESDPIVAQY, which was identified despite the fact that the most similar peptides in the DMS library had three amino acid mismatches. The human peptide returned at the third highest activation threshold was MAGE-A6$_{168\text{-}176}$ (EVDPIGHVY), another known target of TCR$_{a3a}$[30]. We also identified the cancer-testis antigen MAGE-B18$_{166\text{-}174}$ (EVDPIRHYY), which has been previously shown to activate TCR$_{a3a}$[30].

We next selected a subset of predicted off-target peptides with potential clinical relevance for experimental validation. These were peptides returned across several activation thresholds and included ten human and two viral peptides (Supplementary Table 5). We also included the bacterial *C. difficile* EKDPIKENY peptide, which was predicted in a previous study to be a potential off-target of TCR$_{a3a}$[30], but was not predicted in our high-resolution DMS analysis as inclusion of lysine at position 2 resulted in negligible TnT-TCR$_{a3a}$ activation (Fig. 4a). Peptides were pulsed on Colo 205 cells and co-cultured with TnT-TCR$_{a3a}$ cells overnight, followed by assessment of activation based on both NFAT-GFP and CD69 expression. MAGE-A3 induced strong expression of NFAT-GFP and CD69, while MAGE-A6 and MAGE-B18 also induced substantial but lower activation (Fig. 4c). Since NFAT-GFP expression was low in some co-cultures, quantification of CD69$^{high}$ TnT-TCR$_{a3a}$ cells allowed us to identify peptides inducing responses significantly higher than background (CMV negative control peptide) (Fig. 4d). We found that titin significantly activated TnT-TCR$_{a3a}$ cells, however this response was still considerably lower than that induced by MAGE-A3, which is consistent with the reported 50-fold difference in the binding affinity of TCR$_{a3a}$ towards titin peptide-MHC[30]. We also found that human peptides derived from the proteins ANR16, CD166 and MRCKA isoform 2; and a peptide derived from a Kaposi Sarcoma-associated herpesvirus exonuclease (AN_HHV8P) induced significant activation of TnT-TCR$_{a3a}$ cells (Fig. 4d), and thus represent potentially novel and clinically relevant off-targets of TCR$_{a3a}$.

Having profiled TCR$_{a3a}$ cross-reactivity through target peptide DMS, we wondered if any of the resulting off-targets were capable of activating our engineered TCR$_{A3}$ combinatorial variants (Fig. 3g-i). To assess this possibility, we co-cultured TnT cells expressing wild-type TCR$_{A3}$, TCR$_{a3a}$ or selected TCR$_{A3}$ variants with Colo 205 cells pulsed with activating off-target peptides (excluding MAGE-A6 and MAGE-B18), as well as MAGE-A3, titin and CMV control peptides (Fig. 4e). We found that all TnT-TCR$_{A3}$ variants displayed a stronger response to MAGE-A3 than wild-type TnT-TCR$_{A3}$, which showed negligible activation. Interestingly, we found that TnT -TCR$_{A3\text{-}27}$ and TnT-TCR$_{A3\text{-}28}$ were activated by titin. While the low response observed in TnT-TCR$_{A3\text{-}28}$ cells might be explained by its residual level of binding to titin peptide-MHC dextramer (Fig. 3i), the robust response in TnT-TCR$_{A3\text{-}27}$ cells was highly unexpected considering its minimal binding to titin peptide-MHC dextramer (Fig. 3i). In contrast, TnT-TCR$_{A3\text{-}09}$ showed higher binding to titin peptide-MHC dextramer than TnT-TCR$_{A3\text{-}27}$ cells (Fig 3i) but were not activated in response to titin-induced signaling (Fig 5e and Supplementary Fig. 15). These results provide further examples of the discordance between TCR

antigen binding and signaling, and emphasize the importance of functional screening to better assess TCR specificity and cross-reactivity. Another interesting finding was that negative selection against titin recognition (Figs. 3b and 3c) did not impede the emergence of additional cross-reactivity to other peptides in some of the engineered $TCR_{A3}$ variants. This was clearly the case for variants $TCR_{A3\text{-}09}$ (cross-reactive to ANR16 and MRCKA) and $TCR_{A3\text{-}12}$ (cross-reactive to CD166 and MRCKA), which provide examples of increased cross-reactivity following TCR engineering (Fig. 4e). Most importantly, we found that variants $TCR_{A3\text{-}03}$, $TCR_{A3\text{-}04}$, $TCR_{A3\text{-}05}$, $TCR_{A3\text{-}08}$ and $TCR_{A3\text{-}10}$ showed low or negligible activation in response to all tested peptides (Fig. 4e, highlighted in blue, and Supplementary Fig. 17).

**Engineered TCRs expressed in primary human T cells display high target reactivity with minimal cross-reactivity**

As a final validation step, we profiled the activities of selected engineered TnT-$TCR_{A3}$ variants in primary human CD8+ T cells. To achieve this, we adapted a CRISPR-Cas9-based method for dual knockout of endogenous TCRα and TCRβ chains and simultaneous reconstitution with transgenic TCRs targeted to the TRAC locus[56,61,62] (Fig. 5a). The lack of a complete TRAC region in the designed TCRβα HDR templates made splicing with endogenous TRAC exons a requirement for transgenic TCR surface expression. In order to confirm targeted TCR reconstitution, we first validated this approach using $TCR_{a3a}$. Accordingly, RT-PCR using a forward TRBV5-1 (present in $TCR_{a3a}$) primer and reverse endogenous TRAC primer yielded a 1.6 kb product only in T cells transfected with both gRNA and HDR template (Fig. 5b). Sanger sequencing of the resulting product revealed correct integration of $TCR_{a3a}$ sequences. By contrast, a control RT-PCR amplification of TCRα chains utilizing the TRAV21 gene yielded products of expected size for test and control transfections (Fig. 5b). Flow cytometry confirmed CD3 expression and binding to MAGE-A3 peptide-MHC dextramer in T cells reconstituted with $TCR_{a3a}$, and a greater than 95% endogenous TCR knockout efficiency (Fig. 5c).

We next generated primary human CD8+ T cells expressing wild-type $TCR_{A3}$, $TCR_{a3a}$, selected $TCR_{A3}$ variants or no transgenic TCR (TCR KO). In order to verify transgenic TCR function, we first co-cultured transfected T cells with MAGE-A3-positive EJM or MAGE-A3-negative Colo205 cells, and assessed their activation by means of IFN-γ ELISpot (Figs. 5d and 5e). No significant differences in MAGE-A3-induced activation were observed in T cells transfected with wild-type $TCR_{A3}$, $TCR_{A3\text{-}04}$, or no TCR (Fig. 5e). Notably, while wild-type $TCR_{A3}$ showed a minimal response to EJM cells, the complete lack of response observed in T cells transfected with $TCR_{A3\text{-}04}$ indicated that this variant was not successfully expressed by primary T cells. In contrast, we found that T cells expressing $TCR_{a3a}$, $TCR_{A3\text{-}03}$, $TCR_{A3\text{-}05}$, $TCR_{A3\text{-}08}$ and $TCR_{A3\text{-}10}$ were strongly and significantly activated after co-culture with EJM cells (Figs. 5d and 5e).

We next performed a series of co-culture experiments using peptide-pulsed Colo 205 cells (HLA-A*0101-positive, MAGE-A3-negative) coupled with IFN-γ ELISpot readouts. Consistent with our results using TnT-TCR cells, we observed $TCR_{a3a}$ cross-reactivity to titin, while titin-induced responses in our engineered $TCR_{A3}$ variants were either absent or of lower magnitude (Fig. 5f and Supplementary Fig. 18). Notably, primary T cells expressing $TCR_{a3a}$ showed a significant titin-induced IFN-γ response relative to background, with a magnitude that was 62% of their observed MAGE-A3 response (Fig. 5f). We also observed a low but

significant titin-induced IFN-γ response in primary T cells expressing $TCR_{A3\text{-}08}$ (Fig. 5f). Importantly, we found that primary T cells expressing variants $TCR_{A3\text{-}03}$, $TCR_{A3\text{-}05}$ or $TCR_{A3\text{-}10}$ displayed a significant MAGE-A3-induced IFN-γ response but undetectable activation following co-culture with titin-pulsed Colo 205 cells (Fig. 5f). Furthermore, primary T cells expressing wild-type $TCR_{A3}$ showed a similarly negligible titin-induced response in this assay (Fig. 5f). Assessment of TCR cross-reactivity against off-target peptides previously identified for $TCR_{a3a}$ revealed that primary T cells expressing $TCR_{A3\text{-}03}$ or $TCR_{A3\text{-}08}$ had significant IFN-γ responses to peptides derived from CD166, MRCKA isoform 2 and AN_HHV8 ($TCR_{A3\text{-}08}$ only), and were thus excluded from further characterization (Supplementary Fig. 19). By contrast, primary T cells expressing $TCR_{A3\text{-}05}$ or $TCR_{A3\text{-}10}$ displayed high levels of specificity for the MAGE-A3 target peptide, with no other peptide inducing detectable responses.

Given the remarkable specificity of variants $TCR_{A3\text{-}05}$ and $TCR_{A3\text{-}10}$, we next decided to assess their cross-reactivity profiles using the TnT platform and DMS of the target MAGE-$A3_{168\text{-}176}$ peptide (EVDPIGHLY). We assessed activation of TnT-TCR cells by flow cytometric detection of CD69 expression, as we previously validated it as a more sensitive readout for cross-reactivity detection (Fig. 4c). Similar to TnT-$TCR_{a3a}$ peptide DMS (Fig. 4a), most mutations at positions 1, 3, 4, 5 and 9 were detrimental for TnT-$TCR_{A3\text{-}05}$ and TnT-$TCR_{A3\text{-}010}$ activation (Figs. 5g-h). Remarkably, we found that several peptides with mutations at positions 6 and 7, which were mostly activating in TnT-$TCR_{a3a}$ cells, led to substantially reduced activation in TnT-$TCR_{A3\text{-}10}$ and TnT-$TCR_{A3\text{-}05}$ cells. Of note, we found that the presence of valine at peptide position 6, a substitution present in the titin off-target peptide (ESDPIVAQY), drastically reduced responses in both TnT-$TCR_{A3\text{-}05}$ and $TCR_{A3\text{-}10}$ (4-8% of the wild-type MAGE-A3 response), which rationalizes the lack of cross-reactivity of these engineered $TCR_{A3}$ variants to titin (Fig. 5f). As a way of comparison, the same peptide induced a response of 91% relative to wild-type MAGE-A3 in TnT-$TCR_{a3a}$ cells (Fig. 4a).

Querying of the UniProtKB database with motifs derived from peptide DMS data of TnT-$TCR_{A3\text{-}05}$ and TnT-$TCR_{A3\text{-}10}$ revealed substantial reductions in the number of predicted off-targets relative to TnT-$TCR_{a3a}$, both in terms of human and non-human sequences (Fig. 5i and Supplementary Fig. 20). Most notably, the MAGE-A3 target peptide became the only returned human hit for $TCR_{A3\text{-}05}$ and $TCR_{A3\text{-}10}$ at much lower activation thresholds (26% and 44%, respectively) than $TCR_{a3a}$ (89%), further highlighting their reduced tolerance to peptide substitutions. In a similar manner, additional human hits for $TCR_{A3\text{-}05}$ and $TCR_{A3\text{-}10}$ were returned at considerably lower activation thresholds relative to $TCR_{a3a}$ (Fig. 5j).

## DISCUSSION

Here we describe the development of the TnT platform, an extensively CRISPR-edited human T cell line that supports the functional display, engineering and cross-reactivity profiling of TCRs. TnT cells harbor fully-defined genomic changes that facilitate the display and functional engineering of transgenic TCRs at high-throughput. As such, the TnT platform provides important advantages over previous TCR engineering methods that rely exclusively on affinity-based readouts, especially in light of the poor correlation that exists between TCR affinity and function[18,19]. Furthermore, TCR reconstitution by CRISPR-Cas9 targeting of the endogenous TCRβ locus offers several advantages over plasmid transfection[33] or viral transduction[25,31,32,34–36] such as homogenous and physiological expression of TCRs, and occurrence of a single integration event

per cell ensured by targeting of the monoallelic Jurkat CDR3β genomic sequence. Furthermore, optimization of our CRISPR-Cas9 HDR protocol led to TCR integration efficiencies of up to 20%, which allowed for the screening of large mutagenesis libraries and selection of variants with enhanced specificity and function.

The unique features of the TnT platform allowed us to apply CDR3β DMS to comprehensively profile two tumor-reactive TCRs ($TCR_{A3}$ and $TCR_{DMF4}$) for their patterns of expression, antigen binding and antigen-induced signaling. Deep sequencing of original libraries and FACS selections was a crucial component of our TCR engineering pipeline (both in DMS and combinatorial selections) as it allowed us to accurately determine the enrichment of specific TCR variants across selections. Remarkably, we identified several variants that were enriched for antigen binding but not for antigen-induced signaling, and vice-versa, which emphasized the discordance between TCR binding and function. For example, two $TCR_{DMF4}$ variants had nearly identical signaling capacity despite one of them displaying undetectable binding to the MART-1 peptide-MHC dextramer. This is in agreement with previous reports of peptide-MHC multimers failing to detect fully functional TCRs[42]. Crucially, DMS enrichment data allowed us to tailor the design of a combinatorial $TCR_{A3}$ library in order to maximize the number of productive variants.

We utilized the TnT platform for engineering the MAGE-A3-specific $TCR_{A3}$, a naturally occurring low-avidity TCR that has been previously engineered by phage display, which resulted in affinity-enhanced variants with cross-reactivity to titin[14,30,63]. To this end, we designed a DMS-based combinatorial library of CDR3β $TCR_{A3}$ and used the TnT platform to perform selections based on binding and signaling in response to both MAGE-A3 and titin antigens. Throughout the selection process, we observed the emergence of $TCR_{A3}$ variants with enhanced binding and signaling in response to MAGE-A3. Interestingly, we found that a considerable proportion of library members also displayed binding and signaling in response to titin, highlighting the high propensity of $TCR_{A3}$ mutants to develop titin cross-reactivity. By using deep sequencing enrichment data, we could identify $TCR_{A3}$ variants with increased binding and function in response to MAGE-A3 that neither bound or responded to titin. Surprisingly, we found that some variants (e.g., $TCR_{A3\text{-}27}$ and $TCR_{A3\text{-}28}$) displayed robust signaling in response to titin, despite showing low or undetectable binding to titin-MHC dextramer. This finding further emphasizes the importance of functional screening of engineered TCRs, as screening approaches based on binding alone may fail to identify cross-reactive TCRs[16,40]. Remarkably, $TCR_{A3\text{-}05}$, which showed the highest specificity to MAGE-A3 throughout our experiments in both TnT cells and primary T cells, was the variant with the highest score for predicted MAGE-A3 specificity based on deep sequencing enrichment data (Supplementary Table 3). Taken together, these findings highlight the power of TnT functional selections coupled with deep sequencing for engineering highly specific TCRs.

The TnT platform allows us to both engineer TCRs and assess their cross-reactivity potential without the need of reformatting TCRs for cellular display[31,48,49] or soluble expression[39,40]. By performing DMS of the target $MAGE\text{-}A3_{168\text{-}176}$ peptide, we predicted and validated known and potentially novel off-targets for $TCR_{a3a}$. Despite its four amino acid difference to MAGE-A3, the titin ESDPIVAQY peptide was identified at the second highest activation threshold. Interestingly, we also identified a peptide derived from MRCKA isoform 2 (UniProtKB, Q5VT25-2) as a potential $TCR_{a3a}$ off-target, and confirmed its ability to activate TnT-

$TCR_{a3a}$ cells *in vitro*. As MRCKA is highly expressed in the heart[64], it is possible that recognition of this peptide may have also contributed to the cardiac toxicity elicited by $TCR_{a3a}$[14]. We observed that a number of our engineered $TCR_{A3}$ variants were activated by predicted $TCR_{a3a}$ off-targets. As these TnT-$TCR_{A3}$ variants showed weaker binding to MAGE-A3 peptide-MHC than TnT-$TCR_{a3a}$ ($TCR_{a3a}$ $K_D$ ~ 2.3 – 6.6 μM)[23,30], they constitute examples of cross-reactivity emergence in engineered TCRs within the physiological affinity range (1 – 100 μM)[19].

Assessment of primary T cell activation by means of IFN-γ ELISpot assays showed a higher sensitivity than CD69 expression in TnT-TCR cells following co-culture, which may reflect higher levels of co-receptors in primary CD8+ T cells relative to TnT cells. Despite this difference, we were able to identify variants that displayed remarkable specificity for MAGE-A3, namely $TCR_{A3\text{-}05}$ and $TCR_{A3\text{-}10}$, in both TnT-TCR and primary T cell assays. Functional cross-reactivity profiling of these variants using the TnT platform revealed a large reduction in the number of predicted off-targets compared to $TCR_{a3a}$. While reduced cross-reactivity was particularly evident for titin, it was also observed for peptides that were not negatively selected for (e.g., MAGE-A6, CD166 and MRCKA isoform 2), thus highlighting $TCR_{A3\text{-}05}$ and $TCR_{A3\text{-}10}$ as intrinsically specific variants. This is consistent with the recent identification of naturally-occurring TCRs displaying a wide range of specificity levels for the same target[16]. Thus, the enhanced function and specificity of $TCR_{A3\text{-}05}$ and $TCR_{A3\text{-}10}$ make them promising candidates for use in TCR gene therapies targeting MAGE-A3-positive tumors in HLA-A*0101-positive patients. Thus, the ability of the TnT platform to support both TCR engineering and the accurate prediction of cross-reactivity on the basis of function makes it a promising technology for the development of therapeutic TCRs with improved efficacy and safety profiles.

# METHODS

## Cell lines and cell culture

The Jurkat leukemia E6-1 T cell line was obtained from the American Type Culture Collection (ATCC) (#TIB152); the T2 hybrid cell line (#ACC598) and the EJM multiple myeloma cell line (#ACC560) were obtained from the German Collection of Cell Culture and microorganisms (DSMZ); and the Colo 205 colon adenocarcinoma cell line (#87061208) was obtained from the European Collection of Authenticated Cell Cultures (ECACC). Jurkat cells, engineered TnT cells and Colo 205 cells were cultured in ATCC-modified RPMI-1640 (Thermo Fisher, #A1049101), T2 cells were cultured in RPMI-1640 (Thermo Fisher, #11875093), and EJM cells were cultured in IMDM (Thermo Fisher, #12440053). All media were supplemented with 10% FBS, 50 U $ml^{-1}$ penicillin and 50 µg $ml^{-1}$ streptomycin. Detachment of EJM and Colo 205 adherent cell lines for passaging was performed using the TrypLE reagent (Thermo Fisher, #12605010). All cell lines were cultured at 37 °C, 5% $CO_2$ in a humidified atmosphere.

## Polymerase chain reaction (PCR)

PCRs for cloning, generation of HDR templates, genotyping of mammalian cells, generation of TCR libraries and generation of amplicons for deep sequencing were performed using the KAPA HiFi PCR kit with GC buffer (Roche Diagnostics, #07958846001) and custom designed primers (Supplementary Table 6). Annealing temperatures (x) were optimized for each reaction by gradient PCR and cycling conditions were as follows: 95°C for 3 min; 35 cycles of 98°C for 20 s, x°C for 15 s, 72°C for 30 s per kb; final extension 72°C for 1 min per kb. PCRs for genotyping of bacterial colonies after transformation were performed using the KAPA2G Fast ReadyMix kit (Sigma Aldrich, #KK5102) with custom designed primers and the following cycling conditions: 95°C for 3 min; 35 cycles of 95°C for 15 s, 60°C for 15 s, 72°C for 15 s per kb; final extension 72°C for 1 min per kb.

## Cloning and generation of HDR templates

DNA for gene-encoding regions and homology regions were generated by gene synthesis (Twist Bioscience) or PCR and introduced into desired plasmid backbones via restriction cloning (Supplementary File 1). The following plasmids were used as backbones: pX458 (Addgene, #48138), AAVS1_Puro_Tet3G_3xFLAG_Twin_Strep (Addgene, #92099), pGL4.30 (Promega, #E8481) and pTwist Amp High Copy (Twist Bioscience). Targeted knock-in of Cas9/GFP into the CCR5 locus was performed utilizing circular plasmid DNA as the HDR template. HDR templates for all other targeted knock-in experiments were provided as linear double-stranded DNA (dsDNA) generated by PCR. Prior to transfection, PCR products were column-purified using the QIAquick PCR Purification Kit (Qiagen, #28106). For targeted TCR reconstitution of TnT cells, homology arms flanking the recombined Jurkat TCRβ VDJ locus were designed and cloned in pTwist (Twist Bioscience), resulting in pJurTCRB. TCRαβ cassettes encoding transgenic TCRs were generated by gene synthesis (Twist Bioscience) and cloned into pJurTCRB using naturally-occurring XbaI and BsaI restriction sites present within the homology arms. Next, HDR templates were generated by PCR using primer pair RVL-127/128 and PCR products purified prior to transfection. For targeted TCR reconstitution of primary human CD8+ T cells, TCRβα cassettes lacking TRAC exons 2-3 and flanked by homology arms mapping to TRAC exon 1[56] were designed and cloned in

pTwist (Twist Bioscience). HDR templates were generated by PCR using primer pair RVL-164/165 and PCR products purified prior to transfection.

**CRISPR-Cas9 genome editing**

Transfection of TnT cells and Jurkat-derived cell lines was performed by electroporation using the 4D-Nucleofector device (Lonza) and the SE cell line kit (Lonza, #V4XC-1024). The day before transfection, cells were seeded at $2.5x10^5$ cells/mL and cultured for 24 h. Prior to transfection, gRNA molecules (Supplementary Table 7) were assembled by mixing 4 µl of custom Alt-R crRNA (200 µM, IDT) with 4 µL of Alt-R tracrRNA (200 µM, IDT, #1072534), incubating the mix at 95°C for 5 min and cooling it to room temperature. For transfection of Cas9-negative cell lines, 2 µL of assembled gRNA molecules were mixed with 2 µL of recombinant SpCas9 (61 µM, IDT, #1081059) and incubated for > 10 min at room temperature to generate Cas9 RNP complexes. Immediately prior to transfection, cells were washed twice in PBS and $1x10^6$ cells were re-suspended in 100 µL of SE buffer. 1.5 µg of HDR template and 7 µL of assembled gRNA (or 4 µL of Cas9 RNP complexes) were added to the cell suspension, mixed and transferred into a 1 mL electroporation cuvette. Cells were electroporated using program CK116, topped-up with 1 mL of complete media and rested for 10 min prior to transfer into a 12-well plate. Alt-R HDR enhancer (IDT, #1081073) was added at a 30 µM final concentration and removed after 16 h of culture by centrifugation. HDR efficiency was assessed by flow cytometry on day 5 post-transfection. For transfections at the 20 µL scale (Lonza, #V4XC-1032), cell numbers and reagent volumes were reduced 5-fold.

**Flow cytometry and fluorescence-activated cell sorting (FACS)**

Flow cytometric analysis of cell lines and primary T cells was performed according to standard protocols. The following antibodies were purchased from BioLegend and used at 1 µg $ml^{-1}$ in flow cytometry buffer (PBS, 2% FBS, 2 mM EDTA): PE-Cy7-conjugated or APC-conjugated anti-human CD3e (clone UCHT1, #300420 or #300458), APC-conjugated anti-human CD4 (clone RPA-T4, #300552), PE-conjugated anti-human CD8a (clone HIT8a, #300908), PE-Cy7-conjugated anti-human CD19 (clone HIB19, #302216), APC-conjugated anti-human CD69 (clone FN50, #310910), APC-conjugated anti-human Fas (clone DX2, #305611) and PE-conjugated anti-human TCR α/β (clone IP26, #306707). DAPI viability dye (Thermo Fisher, #62248) was added to antibody cocktails at a final concentration of 1 µg $ml^{-1}$. Cells were washed once in flow cytometry buffer prior to staining, stained for 20 min on ice and washed twice in flow cytometry buffer before analysis using BD LSRFortessa or Beckman-Coulter CytoFLEX flow cytometers. Blocking of Fc receptors in T2 cells was performed prior to staining using the TruStain FcX reagent (BioLegend, #422301). Staining with peptide-MHC dextramers was performed for 10 min at room temperature (RT), followed by addition of 2X antibody cocktails (2 ug $ml^{-1}$ antibodies, 2 ug $ml^{-1}$ DAPI) and incubation for 20 min on ice. The following peptide-MHC dextramers were commercially obtained from Immudex: NY-ESO-$1_{157-165}$ (SLLMWITQC, HLA-A*0201, #WB2696-APC); MART-$1_{26-35(27L)}$ (ELAGIGILTV, HLA-A*0201, #WB2162-APC); MAGE-$A3_{168-176}$ (EVDPIGHLY, HLA-A*0101, #WA3249-PE) and titin$_{24,337-24,345}$ (ESDPIVAQY, HLA-A*0101, custom-made, APC-conjugated). Peptide-MHC dextramers were used at a 3.2 nM final concentration (i.e., 1:10 dilution) for staining, unless indicated otherwise in figure legends. FACS was performed using BD FACSAria III or BD FACSAria Fusion instruments. Single-cell sorts were collected

in 96-well flat-bottom plates containing conditioned media and clones were cultured for 2-3 weeks prior to characterization.

**Genotyping of cell lines and transfectants**

Genomic DNA was extracted from $2x10^5$ cells by resuspension in 100 µL of QuickExtract solution (Lucigen, #0905T), incubation at 65°C for 6 min, vortexing for 15 s and incubation at 98°C for 2 min. 5 µL of genomic DNA extract were then used as templates for 25 µL PCR reactions. For genotyping by two-step reverse transcription PCR (RT-PCR), RNA from $1x10^5$ cells was extracted using the TRIZol reagent (Invitrogen, # 15596018) and column-purified using the PureLink RNA Mini kit (Invitrogen, #12183025). For reverse transcription, 100 pmol of oligo dT, 10 nmol of each dNTP, 5 µL RNA and sufficient nuclease-free water for a final 14 µl volume were mixed, incubated at 65°C for 5 min and chilled on ice for 5 min. This was followed by addition of 4 µL of 5X RT buffer, 40 units of RiboLock RNAse inhibitor (Thermo Fisher, #EO0381) and 200 units of Maxima H-minus reverse transcriptase (Thermo Fisher, #EP0751) and mixing. In some experiments, 40 pmol of template-switching oligonucleotide (TSO, Supplementary Table 6) was added for labelling of first-strand cDNA 3' ends[65]. Reverse transcription was performed at 50°C for 30 min, followed by inactivation at 85°C for 5 min. 5 µl of the resulting cDNA-containing reverse transcription mixes were then used as templates for 25 µL PCR reactions.

**Peptides and peptide pulse**

Peptides and peptide libraries were generated by custom peptide synthesis (Genscript), re-suspended at 10 mg $ml^{-1}$ in DMSO and placed at -80°C for prolonged storage. For peptide pulsing, T2 cells or Colo 205 cells were harvested and washed twice in serum-free RPMI 1640 (SF-RPMI). Peptides were diluted to 10 µg $ml^{-1}$ in SF-RPMI (or to concentrations indicated in figure legends) and the solution was used to re-suspend cells at $1x10^6$ cells $ml^{-1}$. Cells were incubated for 90 min at 37°C, 5% CO2, washed once with SF-RPMI, re-suspended in complete media and added to co-culture wells (see section below).

**TnT stimulation and co-culture assays**

For clone screening and assessment of AICD, TnT cells and Jurkat-derived cell lines were stimulated overnight with either 10 µg $ml^{-1}$ plate-bound anti-human CD3e antibody (clone OKT3, BioLegend, #317326) or 1X eBioscience Cell Stimulation Cocktail (81 nM PMA, 1.34 µM ionomycin; Thermo Fisher, #00497093). For co-culture experiments, TnT-TCR cells at ~ $1x10^6$ cells $ml^{-1}$ density were harvested, pelleted by centrifugation and re-suspended in fresh complete media at $1x10^6$ cells $ml^{-1}$. $1x10^5$ TnT-TCR cells (100 µL) were seeded in wells of a V-bottom 96-well plate. Antigen-expressing cells (EJM) or peptide-pulsed cells (T2, Colo 205) were adjusted to $1x10^6$ cells $ml^{-1}$ in complete media and $5x10^4$ cells (50 µL) added to each well. Anti-human CD28 antibody (clone CD28.2, BioLegend, #302933) was added at a final concentration of 1 µg $ml^{-1}$ for co-stimulation of all samples (including negative controls) and plates were incubated overnight at 37°C, 5% $CO_2$. The next day, expression of NFAT-GFP and CD69 in TnT-TCR cells was assessed by flow cytometry. Flow cytometric discrimination between TnT-TCR cells and Colo 205 cells (or EJM cells) was based on side scatter area (SSC-A) and mRuby expression, while discrimination between TnT-TCR cells (CD19-negative) and T2 cells (CD19-positive) was based on CD19 expression.

**Generation of deep mutational scanning (DMS) libraries**

DMS libraries of the CDR3β regions of $TCR_{A3}$ and $TCR_{DMF4}$ were generated by plasmid nicking mutagenesis as described in Wrenbeck et al. 2016[59]. The protocol relies on the presence of a single BbvCI restriction site for sequential targeting with Nt.BbvCI and Nb.BbvCI nickases, digestion of wild-type plasmid and plasmid re-synthesis using mutagenic oligonucleotides. A plus-strand BbvCI restriction site was introduced into the pJurTCRB-$TCR_{A3}$ plasmid by means of PCR and blunt-end ligation, while the endogenous minus-strand BbvCI site present in the TRBV10-3 gene of pJutTCRB-$TCR_{DMF4}$ was targeted. The order of BbvCI nickase digestion was adjusted for each plasmid so that the plus DNA strand was digested first. Mutagenic oligonucleotides were designed using the QuikChange Primer Design online tool (Agilent) and assessed for the presence of secondary structures using the Oligo Evaluator online tool (Sigma-Aldrich) (Supplementary Table 6). Oligonucleotides showing strong potential for forming secondary structures were manually modified to reduce this propensity. After nicking mutagenesis, mutated plasmids were transformed into 100 µL of chemically-competent *E. coli* DH5α cells (NEB, #C2987H) and plated on ampicillin (100 µg $ml^{-1}$) LB agar in Nunc BioAssay dishes (Sigma-Aldrich, #D4803). Serial dilutions of transformed cells were plated separately to quantify bacterial transformants. Plasmid libraries were purified from bacterial transformants using the QIAprep Spin Miniprep kit (Qiagen, #27106). HDR templates were generated from plasmid libraries by PCR using primer pair RVL-127/128 and column-purified prior to transfection.

**DMS library screening and selections**

DMS library HDR templates and CDR3B gRNA were used to transfect $1x10^6$ TnT cells. In $TCR_{A3}$ DMS selections, cells with restored CD3 surface expression and no binding to control titin peptide-MHC dextramer were isolated by FACS on day 8 post-transfection (SEL 1). Sorted cells were expanded for 13 days and either stained with MAGE-A3 peptide-MHC dextramer or co-cultured overnight with MAGE-A3-positive EJM cells. Dextramer-positive cells (SEL 2A) and activated $CD69^{high}$ cells (SEL 2B) were then isolated by FACS. In $TCR_{DMF4}$ DMS selections, cells with restored CD3 surface expression and no binding to control NY-ESO-1 peptide-MHC dextramer were isolated by FACS on day 8 post-transfection (SEL 1). Sorted cells were expanded for 13 days and either stained with MART-1 peptide-MHC dextramer or co-cultured overnight with MART-1 peptide--pulsed T2 cells. Dextramer-positive cells (SEL 2A) and activated NFAT-GFP-positive cells (SEL 2B) were then isolated by FACS.

**Generation of combinatorial $TCR_{A3}$ libraries**

Degenerate codons reflecting the combined CDR3β amino acid frequencies observed in $TCR_{A3}$ DMS binding and signaling selections (SEL2A+2B) were determined as previously described[51]. The library resulting from two iterations of our algorithm was modified to include VNB codons at CDR3β positions 4 and 6. For library construction, ssDNA oligonucleotides containing a 28 nt complementary overlap were designed and purchased as custom ultramers (IDT, Supplementary Table 6). The forward ultramer encoded exclusively wild-type $TCR_{A3}$ codons, while the reverse ultramer contained the reverse complement of both wild-type and library degenerate codons. 200 pmol of each ultramer were mixed and subjected to single-cycle PCR using the following conditions: 95°C for 3 min, 98°C for 20 s, 70°C for 15 s, 72°C for 10 min. The resulting 270 bp dsDNA product was gel-purified (Zymogen, #D4002) and 8 ng were utilized as template for a 200 µL PCR reaction using external primers with the following cycling conditions: 95°C for 3 min; 25 cycles

of 98°C for 20 s, 62°C for 15 s, 72°C for 15 s; final extension 72°C for 30 s. The PCR product was column-purified, digested with KpnI and BsaI restriction enzymes, and re-purified. In parallel, the pJurTCRB-$TCR_{A3}$ plasmid was digested with KpnI and BsaI, de-phoshorylated (CIP, NEB, #M0290) and gel-purified. Digested PCR product (112.5 ng) and plasmid (750 ng) were ligated in a 75 µl reaction containing 1X T4 PNK buffer, 1 mM ATP and 3 units of T4 DNA ligase for 2 h at RT (all from NEB). Next, the ligation mix was transformed into 750 µL of chemically-competent *E. coli* DH5α cells (NEB, #C2987I) and plated on ampicillin LB agar in Nunc BioAssay dishes. Quantification of bacterial transformants, purification of plasmid library and generation of HDR templates was performed as described for DMS libraries.

**Combinatorial $TCR_{A3}$ library screening and selections**

Combinatorial library HDR templates (20 µg) and CDR3B gRNA (10 nmol) were used to transfect $1x10^8$ TnT cells using the 4D-Nucleofector LV unit (Lonza, #AAF-1002L). TnT cells with restored CD3 surface expression were bulk-sorted (SEL 1) on day 6 post transfection. SEL 1 cells were expanded for 6 days prior to overnight co-culture with MAGE-A3-positive EJM cells followed by co-staining with MAGE-A3 and titin peptide-MHC dextramers. After co-culture, NFAT-GFP-positive cells displaying positive MAGE-A3 peptide-MHC binding and negative titin peptide-MHC binding were bulk-sorted (SEL 2) and expanded in culture for 12 days. SEL 2 cells were co-cultured overnight with either peptide MAGE-A3-pulsed (MAGE-A3) or titin-pulsed Colo 205 cells. Activated NFAT-GFP-positive cells from MAGE-A3 (SEL 3A) and titin (SEL 3B) co-cultures were bulk-sorted for RNA extraction, RT-PCR and deep sequencing.

**Deep sequencing and analysis of TCR libraries**

TCR amplicons for deep sequencing of plasmid libraries were generated by PCR using primer pair RVL-144/154, while TCR amplicons for deep sequencing of TnT-TCR selections were generated by two-step RT-PCR using primer pair RVL-144/145. In both cases, PCR was limited to 25 cycles. TCR amplicons were column-purified and deep-sequenced using the Amplicon-EZ service (Genewiz), which includes adaptor/index ligation and paired-end Illumina sequencing (250 cycles) followed by delivery of 50,000 assembled reads per sample with unique sequence identification and abundance analysis. For DMS plasmid libraries and selections, unique sequences with less than ten sequencing reads were excluded from enrichment analysis, as every library member had sequencing reads above this threshold. Sequence enrichment of unique DMS variants was determined by dividing their observed frequencies in SEL 1 (TCR-CD3 expression), SEL 2A (binding) and SEL 2B (signaling) over their plasmid DMS library frequencies, and heatmaps were generated using the GraphPad Prism software. For the $TCR_{A3}$ combinatorial plasmid library and selections, unique clone frequency data was filtered to remove clones containing insertions, deletions or mutations outside CDR3β. Filtered data was used to generate sequence logos weighted on amino acid frequencies at specific CDR3β positions using R packages ggseqlogo[66] and ggplot2. The frequencies of specific $TCR_{A3}$ variants across selections were identified by merging unique clone datasets using a custom Python script. Sequence enrichment of unique $TCR_{A3}$ combinatorial variants was determined by dividing their observed frequencies in SEL 2 (MAGE-A3-induced activation and binding), SEL 3A (MAGE-A3-induced activation) and SEL 3B (titin-induced activation) over their SEL 1 (TCR-CD3 expression) frequencies.

**Peptide DMS and assessment of TCR cross-reactivity**

A DMS library of the target MAGE-A3$_{168-176}$ EVDPIGHLY peptide was designed and generated by custom peptide synthesis (Genscript). Each library member (n = 171) was individually pulsed at a 50 ug ml$^{-1}$ concentration on Colo 205 cells for co-culture with TnT-TCR cells (n = 171 co-cultures). Co-cultures with MAGE-A3-pulsed (n = 3), titin-pulsed (n = 3), CMV-pulsed (n = 6) peptides and unpulsed (n = 6) Colo 205 cells were included as controls. After overnight co-culture, TnT-TCR activation was assessed by NFAT-GFP and CD69 expression by means of flow cytometry. The mean background activation observed in CMV peptide controls (i.e., VTEHDTLLY peptide pulse) was subtracted from all samples and their responses normalized to the mean MAGE-A3 response level. Normalized data was used to generate heatmaps (GraphPad Prism), weighted sequence logos (ggseqplot, ggplot2 in R) and peptide sequence motifs of allowed substitutions at discrete activation thresholds (Bioconductor package Biostrings in R). Peptide sequence motifs (Supplementary Table 4) were then used to query the UniProtKB database (including splice variants) with the ScanProsite online tool. The output of these searches was processed using the Biostrings package in order to compute the number of unique peptide hits.

**Primary T cell culture and genome editing**

Human peripheral blood mononuclear cells were purchased from Stemcell Technologies (#70025) and CD8+ T cells isolated using the EasySep Human CD8+ T Cell Isolation kit (Stemcell Technologies, #17953). Primary human CD8+ T cells were cultured for up to 24 days in ATCC-modified RPMI (Thermo Fisher, #A1049101) supplemented with 10% FBS, 10 mM non-essential amino acids, 50 µM 2-mercaptoethanol, 50 U ml$^{-1}$ penicillin, 50 µg ml$^{-1}$ streptomycin and freshly added 20 ng ml$^{-1}$ recombinant human IL-2, (Peprotech, #200-02). T cells were activated with anti-CD3/anti-CD28 tetrameric antibody complexes (Stemcell Technologies, #10971) on days 1 and 13 of culture and expanded every 3-4 days. Transfection of primary T cells with Cas9 RNP complexes and TCRβα HDR templates was performed 3-4 days following activation using the 4D-Nucleofector and a 20 uL format P3 Primary Cell kit (Lonza, V4XP-3032). Briefly, 1x10$^6$ primary CD8+ T cells were transfected with 1 µg of HDR template, 1 µl of TRAC Cas9 RNP complex and 1 µl of TRBC1/2 Cas9 RNP complex using the EO115 electroporation program (Cas9 RNP complexes = 50 µM gRNA, 30.5 µM recombinant SpCas9). For RT-PCR validation of TCR reconstitution, RNA was extracted from 1x10$^6$ T cells, quantified using a Nanodrop instrument, and 40 ng RNA used as input for reverse transcription. 2 µL of reverse transcription mixes were then utilized as templates for 25 µL PCR reactions.

**Co-culture of primary T cells and IFN-γ ELISpot**

IFN-γ ELISpot assays were performed using the Human IFN-γ ELISpot Pair (BD, #551873), 96-well ELISpot plates (Millipore, #MSIPS4W10), Avidin-HRP (Biolegend, #405103) and precipitating TMB substrate (Mabtech, #3651-10). Wells were activated with 15%(v/v) ethanol for 30 s, washed twice with PBS and coated with 5 µg ml$^{-1}$ capture antibody (in PBS) at 4°C overnight (or up to a week). On the day of co-culture (i.e, day 5 post-transfection), wells were washed twice with PBS and blocked with primary T cell media lacking IL-2 (RP10-TC) for > 2 h at 37°C. In parallel, TCR-reconstituted primary CD8+ T cells were rested in the absence IL-2 for 6 h. After resting, T cells were washed and re-suspended in fresh RP10-TC media. A 100 µL volume of cell suspensions containing 5x10$^4$ to 4x10$^5$ T cells was then transferred into blocked

743 ELISpot wells, as specified in figure legends. Next, $1.5x10^4$ antigen-expressing (EJM) or peptide-pulsed
744 (Colo 205) cells were added into wells in a 50 µL volume of RP10-TC media. Anti-CD28 monoclonal antibody
745 was added into every well at a 1 µg $ml^{-1}$ final concentration and plates were incubated for 20 h at 37°C, 5%
746 $CO_2$. Following co-culture, cells were removed and wells washed three times with wash buffer (0.01%(v/v)
747 Tween 20 in PBS). Detection antibody was then added at 2 µg $ml^{-1}$ in dilution buffer (0.5%(v/v) BSA in PBS)
748 followed by 2 h incubation at RT. After incubation, wells were washed three times with wash buffer and
749 1:2000 avidin-HRP (in dilution buffer) added for 45 min at RT. Wells were washed three times with wash
750 buffer and once in PBS, followed by development with precipitating TMB substrate for 3-10 min at RT.
751 Development was stopped by washing with deionized water and plates were dried for > 24 h in the dark
752 prior to analysis using an AID ELR08 ELISpot reader (Autoimmun Diagnostika).
753

754 ## AUTHOR CONTRIBUTIONS

755 R.V.-L. and S.T.R. designed the study; R.V.-L., J.J. and S.T.R. contributed to experimental design; R.V.-
756 L., J.J., F.B., E.A. and E.K. performed experiments; R.V.-L., J.J., F.B., and C.R.W. analyzed data; R.V.-L.
757 and S.T.R. wrote the manuscript with input from all authors.
758

759 ## ACKNOWLEDGMENTS

760 We thank Dr. Joseph Taft for his advice on the plasmid nicking mutagenesis method, and Darya Palianina
761 for her assistance with ELISpot assays. We acknowledge the ETH Zurich D-BSSE Single Cell Unit for their
762 assistance with FACS. This study is supported by funding from the Personalized Health and Related
763 Technologies Postdoctoral Fellowship (to R.V.-L), the NCCR Molecular Systems Engineering (to S.T.R.)
764 and Helmut Horten Stiftung (to S.T.R.).
765

766 ## COMPETING INTERESTS

767 ETH Zurich has filed for patent protection on the technology described herein, and R.V.-L., J.J. and S.T.R.
768 are named as co-inventors on this patent (European Patent Application: EP19202970.0).
769

770 ## DATA AVAILABILITY

771 The raw FASTQ files from deep sequencing that support the findings of this study will be deposited in the
772 Sequence Read Archive (SRA) with the primary accession code(s) <code(s)
773 (https://www.ncbi.nlm.nih.gov/sra)>. Additional data that support the findings of this study are available from
774 the corresponding author upon reasonable request.
775

776

777

778

**Figure 1_Vazquez-Lombardi et al.**

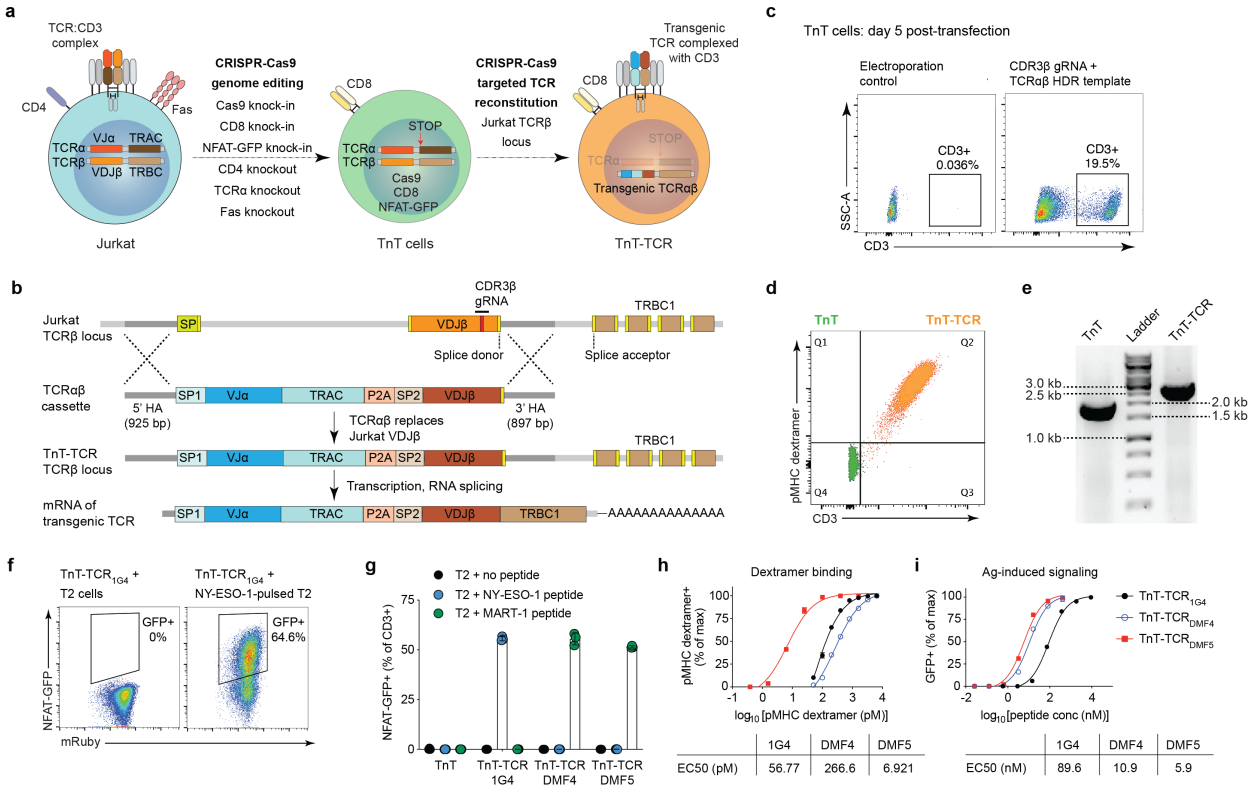



**Figure 1. The TCR-accepting T cell (TnT) platform supports targeted reconstitution and functional display of transgenic TCRs. a,** The Jurkat E6-1 cell line was subjected to sequential CRISPR-Cas9 genome editing steps in order to generate the TnT platform. TnT cells constitutively express Cas9 and human CD8, harbor an NFAT-GFP reporter of TCR signaling, and lack expression of CD4, Fas and endogenous TCR. Reconstitution of TnT cells with transgenic TCRs via CRISPR-Cas9 HDR results in TnT-TCR cells with restored surface expression of the TCR:CD3 complex. **b**, TCR reconstitution of TnT cells targeted to the Jurkat TCRβ locus. An HDR template encoding transgenic TCRαβ chains is integrated into TnT cells via Cas9 and a gRNA targeting the endogenous Jurkat CDR3β sequence. Transgenic TCR expression is dependent on correct RNA splicing with endogenous Jurkat TRBC1. **c,** Flow cytometric assessment of CD3 restoration in TnT cells after their targeted reconstitution with $TCR_{1G4}$ (specific for NY-ESO-$1_{157-165}$ peptide). **d**, Representative flow cytometry plot showing peptide-MHC dextramer binding to TnT cells (no TCR expression) and MART-1-specific TnT-$TCR_{DMF5}$ cells. **e,** Validation of targeted TCR reconstitution in TnT-$TCR_{DMF5}$ cells by genomic PCR of the Jurkat VDJβ region. **f,** Representative flow cytometry dot plots displaying NFAT-driven GFP expression in TnT-$TCR_{1G4}$ cells, but not TnT cells, after overnight co-culture with T2 cells pulsed with NY-ESO-$1_{157-164}$ peptide. **g**, NFAT-GFP expression in TnT and TnT-TCR cells after overnight co-culture with T2 cells pulsed with NY-ESO-$1_{157-164}$, MART-$1_{26-35(27L)}$ or no peptide. **h**, Serially-diluted target peptide-MHC dextramers were used to assess the binding avidities of TnT-$TCR_{1G4}$, TnT-$TCR_{DMF4}$ and TnT-$TCR_{DMF5}$ (n = 3). Peptide-MHC dextramer concentrations resulting in half-maximal proportions of dextramer positive cells (EC50) were derived from non-linear least squares fits. **i,** Normalized NFAT-GFP expression in TnT-TCR cells after overnight co-culture with T2 cells pulsed with serially-diluted cognate peptide (n = 2). Peptide pulse concentrations resulting in half-maximal proportions of NFAT-GFP+ cells (EC50) were derived from non-linear least squares fits. Data are displayed as mean ± SD.

**Figure 2_Vazquez-Lombardi et al.**

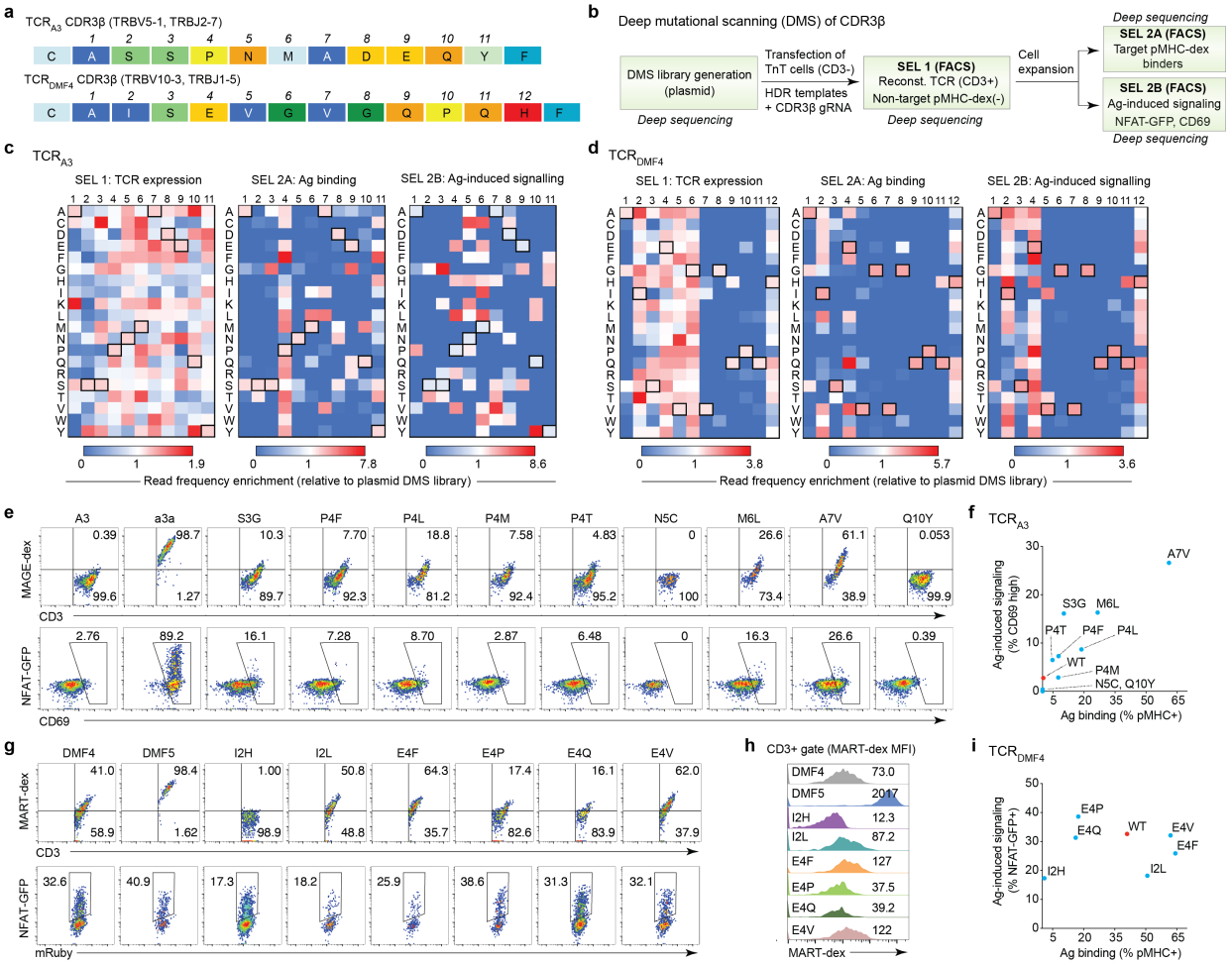



**Figure 2. Deep mutational scanning of CDR3β reveals the sequence landscape for TCR expression, binding and signaling. a,** The CDR3β sequences of $TCR_{A3}$ and $TCR_{DMF4}$ were subjected to deep mutational scanning (DMS). **b**, DMS libraries were generated by means of plasmid nicking saturation mutagenesis and integrated into TnT cells by Cas9 HDR. Reconstituted cells were selected by FACS based on CD3 surface expression, target peptide-MHC binding and antigen-induced signalling (co-culture with antigen-presenting cells). Deep sequencing of CDR3β sequences was performed at every selection step. **c**, **d**, Heatmaps displaying the enrichment of sequencing reads from $TCR_{A3}$ and $TCR_{DMF4}$ variants across selections relative to their occurrence in the starting plasmid DMS libraries. Outlined boxes represent wild-type CDR3β residues. **e-f,** Flow cytometric assessment of antigen binding and antigen-induced signalling in TnT cells reconstituted with $TCR_{A3}$, $TCR_{a3a}$ and selected $TCR_{A3}$ variants (CD3+ gate). **e**, The levels of binding to MAGE-A3 peptide-MHC dextramer (top row) and activation after overnight co-culture with MAGE-A3+ EJM cells (bottom row, proportion $CD69^{high}$ TnT-TCR) are displayed. **f**, Graph displays the levels of antigen binding and antigen-induced signaling in $TCR_{A3}$ single mutants relative to wild-type $TCR_{A3}$. **g-i**, Flow cytometric assessment of antigen binding and antigen-induced signalling in TnT cells reconstituted with $TCR_{DMF4}$, $TCR_{DMF5}$ and selected $TCR_{DMF4}$ variants (CD3+ gate). **g**, The levels of binding to MART-1 peptide-MHC dextramer (top row) and activation after overnight co-culture with MART-1-pulsed T2 cells (bottom row, proportion NFAT-GFP+ TnT-TCR) are displayed. **i**, Histograms displaying the levels of MART-1 peptide-MHC dextramer bound by TnT-TCR cells. **h**, Graph displays the levels of antigen binding and antigen-induced signaling in $TCR_{DMF4}$ single mutants relative to wild-type $TCR_{DMF4}$.

**Figure 3_Vazquez-Lombardi et al.**

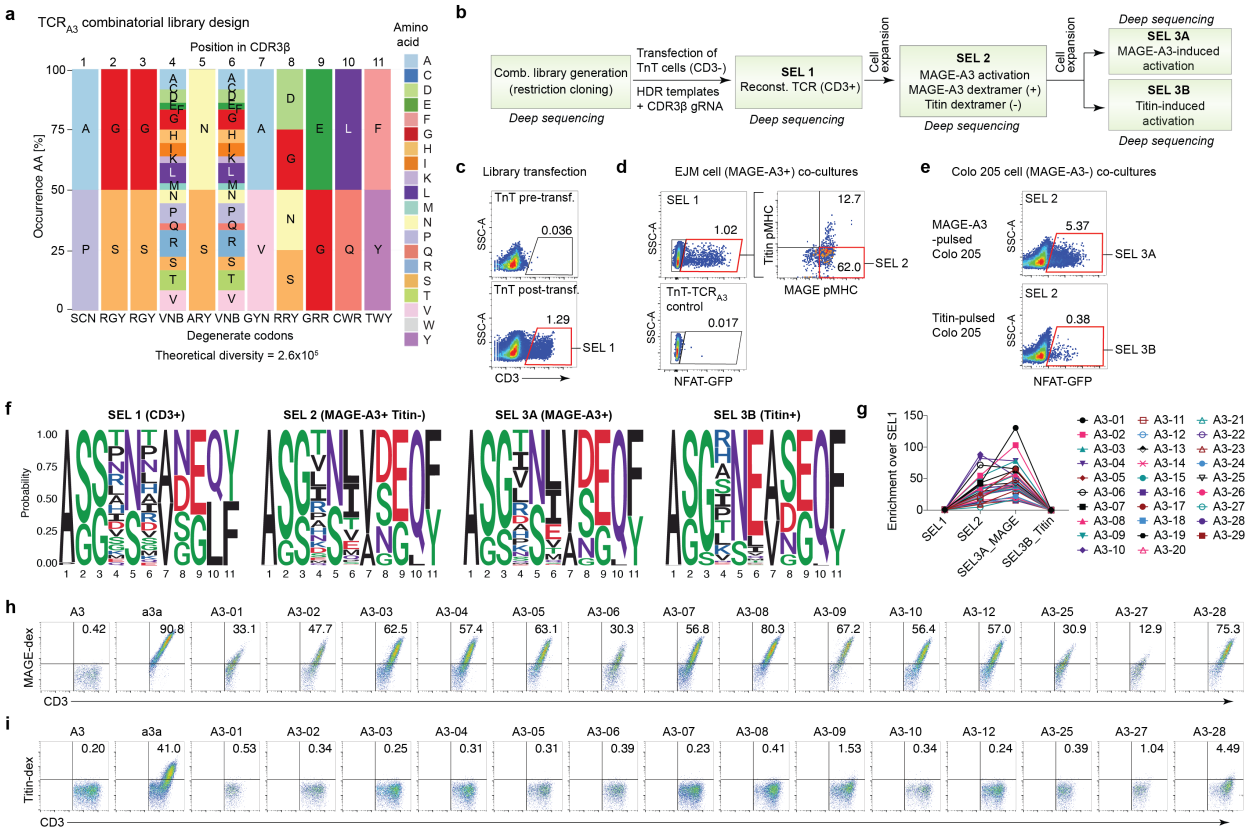



**Figure 3. Rational design and functional screening of combinatorial TCR mutagenesis libraries in TnT cells for selection of enhanced target specificity. a,** Data obtained from $TCR_{A3}$ CDR3β DMS was utilized as input to design a combinatorial mutagenesis library that has a theoretical diversity of 2.6 x $10^5$ variants. Degenerate codons were designed to recapitulate the amino acid frequencies observed in DMS selections based on MAGE-A3 peptide-MHC dextramer binding and MAGE-A3-induced signalling. **b**, Strategy for the selection of $TCR_{A3}$ combinatorial mutagenesis variants with enhanced recognition of the MAGE-$A3_{168-176}$ peptide (EVDPIGHLY), while avoiding cross-reactivity to the known titin$_{24,337-24,345}$ off-target peptide (ESDPIVAQY). Deep sequencing of CDR3β sequences was performed at every selection step. **c**, Flow cytometry plots show selection of $TCR_{A3}$ variants from the combinatorial library that are capable of surface expression in TnT cells. TnT cells with restored CD3 surface expression were bulk-sorted (SEL 1). **d**, Flow cytometry plots showing TnT-TCR cells (from SEL 1) with positive MAGE-A3-induced signaling (NFAT-GFP+), positive MAGE-A3 peptide-MHC binding and negative titin peptide-MHC binding; cells were bulk-sorted (SEL 2) and expanded in culture. **e,** Expanded SEL 2 cells were co-cultured overnight with either MAGE-A3 peptide-pulsed or titin peptide-pulsed Colo 205 cells and bulk-sorted for activation by NFAT-GFP+ expression (SEL 3A and SEL 3B). **f**, Amino acid sequence logos showing the frequency of specific residues at each CDR3β position across selections (logos weighted on unique clone frequencies). **g**, $TCR_{A3}$ variants with predicted high specificity for MAGE-A3 and their enrichment across selections based on deep sequencing data. **h-i**, Flow cytometry plots of TnT cells reconstituted with $TCR_{A3}$, $TCR_{a3a}$ and selected $TCR_{A3}$ combinatorial variants. Cells were assessed for CD3 expression, MAGE-A3 peptide-MHC binding (**h**) and titin peptide-MHC binding (**i**). Cells in the CD3+ gate are shown. Degenerate nucleotide symbols: R = A, G; Y = C, T; S = G, C; W = A, T; K = G, T; M = A, C; B = C, G, T; D = A, G, T; H = A, C, T; V = A, C, G; N = any base.

**Figure 4_Vazquez-Lombardi et al.**

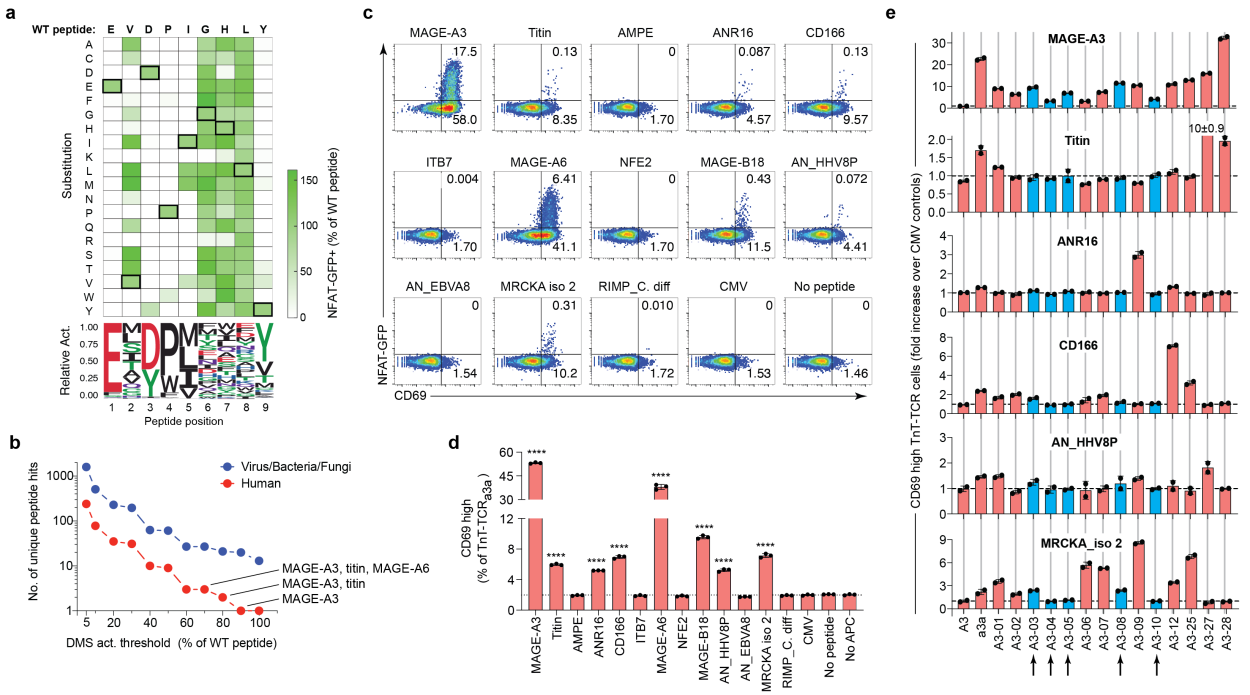
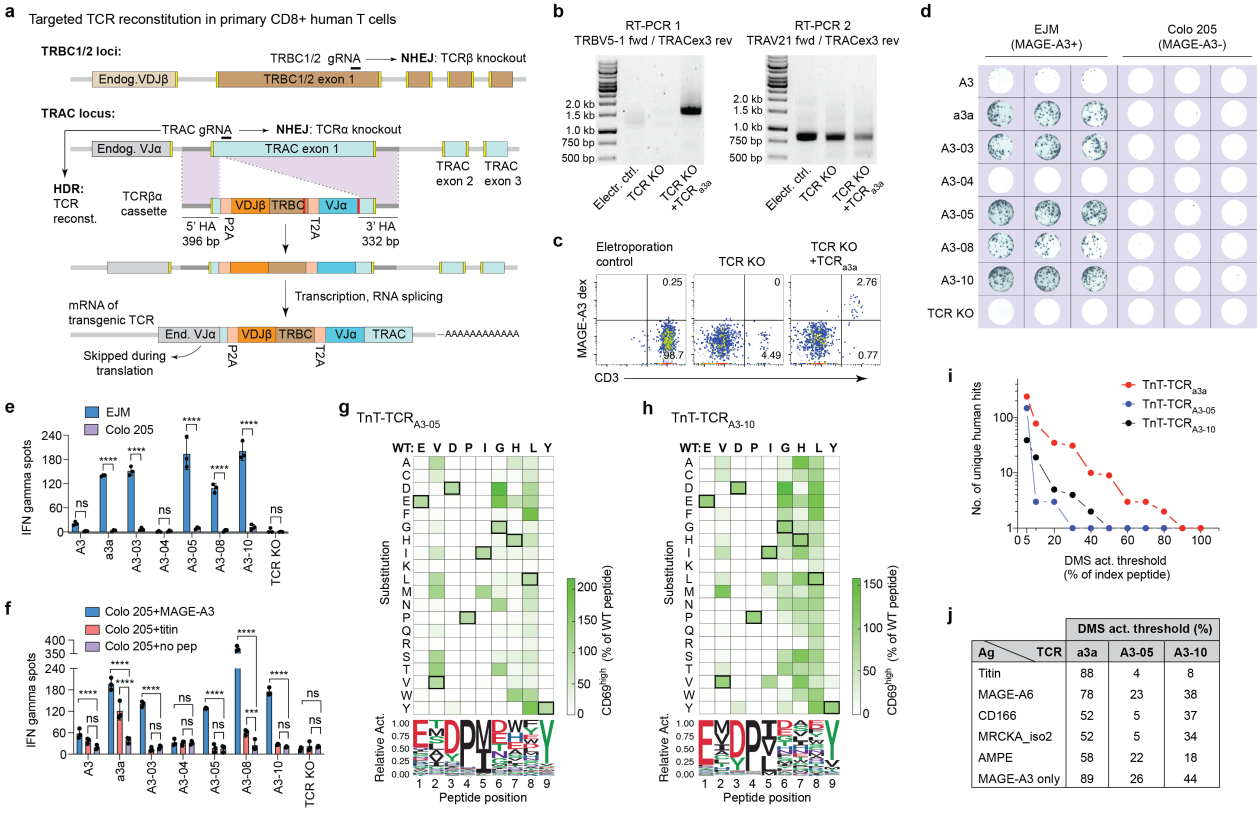



**Figure 4. The TnT platform enables TCR cross-reactivity profiling and prediction of off-target peptides. . a,** The cross-reactivity profile of TnT-TCR$_{a3a}$ cells was assessed using single mutant variants of the wild-type MAGE-A3 peptide (peptide DMS library), which were pulsed on Colo 205 cells for individual co-culture assays (n = 171). Heatmap shows the proportion of NFAT-GFP+ TnT-TCR$_{a3a}$ cells after co-culture, as determined by flow cytometry. Data are normalized to the response induced by the MAGE-A3 wild-type peptide (boxed residues). Sequence logo shows the relative activity of peptide DMS library members carrying mutations at the same peptide position. **b**, The sequences of peptides mediating 5, 10, 20, 30, 40, 50, 60, 70, 80, 90 and 100 percent activation relative to the MAGE-A3 wild-type peptide were utilized to generate motifs to query the UniProtKB database. Dot plot displays the number of human (red) and non-human (blue) unique peptide hits resulting from these searches. The first (MAGE-A3), second (titin) and third (MAGE-A6) highest predicted activating peptides are highlighted. **c**, Flow cytometry of CD69 and NFAT-GFP expression following co-culture with peptide-pulsed Colo 205 cells shows cross-reactivity of TnT-TCR$_{a3a}$ cells against a subset of predicted off-target peptides. The CMV peptide (HLA-A*0101-restricted, VTEHDTLLY) is included as a negative control. **d**, Bar graph shows repeat of the experiment in (**c**) performed in triplicate. The percentages of CD69$^{high}$ TnT-TCR$_{a3a}$ cells after co-culture with peptide-pulsed Colo 205 cells were utilized to assess reactivity to each peptide. Dotted line reflects the mean of CMV-pulsed controls (Y = 1.98). Asterisks indicate significant differences to CMV controls as determined by two-way ANOVA with Bonferroni post-hoc test for multiple comparisons. **e**, Assessment of selected TnT-TCR$_{A3}$ CDR3β variants for cross-reactivity. TnT cells expressing TCR$_{A3}$, TCR$_{a3a}$ and selected TnT$_{A3}$ variants were co-cultured with Colo 205 cells pulsed with a subset of predicted off-target peptides. The percentages of CD69$^{high}$ TnT-TCR cells were determined by flow cytometry and normalized to their respective CMV backgrounds (n = 2). Selected TnT-TCR$_{A3}$ variants displaying favorable cross-reactivity profiles are highlighted in blue. Data are displayed as mean ± SD. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$, ns = not significant. See Supplementary Table 5 for a full list of peptides and their sequences.

**Figure 5_Vazquez-Lombardi et al.**




| Ag \ TCR | DMS act. threshold (%) a3a | A3-05 | A3-10 |
|---|---|---|---|
| Titin | 88 | 4 | 8 |
| MAGE-A6 | 78 | 23 | 38 |
| CD166 | 52 | 5 | 37 |
| MRCKA_iso2 | 52 | 5 | 34 |
| AMPE | 58 | 22 | 18 |
| MAGE-A3 only | 89 | 26 | 44 |

**Figure 5. The TnT platform and primary T cells validate the engineering of TCRs with enhanced target reactivity and minimal cross-reactivity. a**, Schematic shows TCR reconstitution in primary human CD8+ T cells is targeted to the TRAC locus by Cas9 HDR. T cells were transfected with TRAC and TRBC1/2 gRNAs in complex with recombinant Cas9 for dual knockout of endogenous TCRα and TCRβ chains. TCRβα cassettes consisted of a P2A self-processing peptide, followed by VDJβ and TRBC domains, a T2A self-processing peptide and a VJα domain. Regions highlighted in red were re-coded in order to prevent targeting of HDR templates. Transgenic TCR expression is dependent on correct RNA splicing with endogenous TRAC exons 2 and 3. **b**, Validation of targeted $TCR_{a3a}$ integration by means of RT-PCR with a forward primer annealing to the transgenic TRBV5-1 gene and a reverse primer annealing to the endogenous TRAC exon 3. Electroporation control and dual TCR knockout samples display no PCR product, while samples derived from $TCR_{a3a}$ transfectants display a 1.6 kb product consistent with targeted TCR integration and correct splicing with endogenous TRAC. A second RT-PCR reaction amplifying both endogenous and transgenic TRAV21-TRAC sequences was included as a positive control. **c,** Flow cytometric assessment of CD3 expression and MAGE-A3 peptide-MHC dextramer binding in primary CD8+ T cells reconstituted with $TCR_{a3a}$. **d**, IFN-γ ELISPOT assays for assessment of primary CD8+ T cells reconstituted with selected TCRs following co-culture with MAGE-A3-positive EJM cells or MAGE-A3-negative Colo 205 cells (n = 3; $5x10^4$ T cells per well). **e,** Quantification of IFN-γ ELISpot data in (**d**). Asterisks indicate significant differences as determined by two-way ANOVA with Bonferroni post-hoc test for multiple comparisons. **f**, Quantification of IFN-γ ELISpot data from primary CD8+ T cells reconstituted with selected TCRs following overnight co-culture with Colo 205 cells pulsed with $MAGE\text{-}A3_{168\text{-}176}$, $titin_{24,337\text{-}24,345}$, or no peptide (n = 3; $4x10^5$ T cells per well). Asterisks indicate significant differences to non-pulsed Colo 205 controls as determined by two-way ANOVA with Bonferroni post hoc test for multiple comparisons. **g**, **h**, The cross-reactivity profiles of $TnT\text{-}TCR_{A3\text{-}05}$ and $TnT\text{-}TCR_{A3\text{-}10}$ cells were assessed by individual co-cultures with Colo 205 cells pulsed with single mutant variants (DMS library) of the wild-type $MAGE\text{-}A3_{168\text{-}176}$ peptide (n = 171 variants). Heatmaps represent the proportions of $CD69^{high}$ TnT-TCR cells after overnight co-culture, as determined by flow cytometry. Data are normalized to the response induced by wild-type peptide (boxed residues). Sequence logos show the relative activity of peptide DMS library members carrying mutations at the same peptide position. **i**, The sequences of peptides mediating 5, 10, 20, 30, 40, 50, 60, 70, 80, 90 and 100 percent activation relative to wild-type peptide were utilized to generate motifs to query the UniProtKB database. Dot plots display the number of unique human activating peptides resulting from these searches. The $TnT\text{-}TCR_{a3a}$ cross-reactivity profile is displayed for comparison purposes. **j**, Comparison of the highest activation thresholds at which peptide DMS motifs derived from $TnT\text{-}TCR_{a3a}$, $TnT\text{-}TCR_{A3\text{-}05}$ and $TnT\text{-}TCR_{A3\text{-}10}$ return off-target peptides: titin, MAGE-A6, CD166, MRCKA isoform 2 or glutamyl aminopeptidase (AMPE). The lowest activation threshold at which MAGE-A3 peptide becomes the only predicted activating antigen is also displayed. Data are displayed as mean ± SD; * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$, ns = not significant.

## References


1. Schmitt, T. M., Ragnarsson, G. B. & Greenberg, P. D. T cell receptor gene therapy for cancer. *Hum. Gene Ther.* **20**, 1240–1248 (2009).
2. Garber, K. Driving T-cell immunotherapy to solid tumors. *Nat. Biotechnol.* **36**, 215–219 (2018).
3. Newick, K., Moon, E. & Albelda, S. M. Chimeric antigen receptor T-cell therapy for solid tumors. *Molecular Therapy - Oncolytics* vol. 3 16006 (2016).
4. Zhang, J. & Wang, L. The Emerging World of TCR-T Cell Trials Against Cancer: A Systematic Review. *Technology in Cancer Research & Treatment* vol. 18 153303381983106 (2019).
5. Rapoport, A. P. *et al.* NY-ESO-1-specific TCR-engineered T cells mediate sustained antigen-specific antitumor effects in myeloma. *Nat. Med.* **21**, 914–921 (2015).
6. Morgan, R. A. *et al.* Cancer Regression and Neurological Toxicity Following Anti-MAGE-A3 TCR Gene Therapy. *Journal of Immunotherapy* vol. 36 133–151 (2013).
7. Johnson, L. A. *et al.* Gene therapy with human and mouse T-cell receptors mediates cancer regression and targets normal tissues expressing cognate antigen. *Blood* **114**, 535–546 (2009).
8. Morgan, R. A. *et al.* Cancer regression in patients after transfer of genetically engineered lymphocytes. *Science* **314**, 126–129 (2006).
9. Robbins, P. F. *et al.* A pilot trial using lymphocytes genetically engineered with an NY-ESO-1-reactive T-cell receptor: long-term follow-up and correlates with response. *Clin. Cancer Res.* **21**, 1019–1027 (2015).
10. Parkhurst, M. R. *et al.* T cells targeting carcinoembryonic antigen can mediate regression of metastatic colorectal cancer but induce severe transient colitis. *Mol. Ther.* **19**, 620–626 (2011).
11. Klebanoff, C. A., Rosenberg, S. A. & Restifo, N. P. Prospects for gene-engineered T cell immunotherapy for solid cancers. *Nat. Med.* **22**, 26–36 (2016).
12. Strønen, E. *et al.* Targeting of cancer neoantigens with donor-derived T cell receptor repertoires. *Science* **352**, 1337–1341 (2016).
13. Uttenthal, B. J., Chua, I., Morris, E. C. & Stauss, H. J. Challenges in T cell receptor gene therapy. *The Journal of Gene Medicine* vol. 14 386–399 (2012).
14. Linette, G. P. *et al.* Cardiovascular toxicity and titin cross-reactivity of affinity-enhanced T cells in myeloma and melanoma. *Blood* **122**, 863–871 (2013).
15. Sewell, A. K. Why must T cells be cross-reactive? *Nat. Rev. Immunol.* **12**, 669–677 (2012).
16. Bentzen, A. K. *et al.* T cell receptor fingerprinting enables in-depth characterization of the interactions governing recognition of peptide-MHC complexes. *Nat. Biotechnol.* (2018) doi:10.1038/nbt.4303.
17. Wooldridge, L. *et al.* A single autoimmune T cell receptor recognizes more than a million different peptides. *J. Biol. Chem.* **287**, 1168–1177 (2012).
18. Sibener, L. V. *et al.* Isolation of a Structural Mechanism for Uncoupling T Cell Receptor Signaling from Peptide-MHC Binding. *Cell* **174**, 672–687.e27 (2018).
19. Hebeisen, M. *et al.* Identifying Individual T Cell Receptors of Optimal Avidity for Tumor Antigens. *Front. Immunol.* **6**, 582 (2015).
20. Presotto, D. *et al.* Fine-Tuning of Optimal TCR Signaling in Tumor-Redirected CD8 T Cells by Distinct TCR Affinity-Mediated Mechanisms. *Front. Immunol.* **8**, 1564 (2017).
21. Hebeisen, M. *et al.* SHP-1 phosphatase activity counteracts increased T cell receptor affinity. *J. Clin. Invest.* **123**, 1044–1056 (2013).
22. Thomas, S. *et al.* Human T cells expressing affinity-matured TCR display accelerated responses but fail to recognize low density of MHC-peptide antigen. *Blood* vol. 118 319–329 (2011).
23. Tan, M. P. *et al.* T cell receptor binding affinity governs the functional profile of cancer-specific CD8+ T cells. *Clin. Exp. Immunol.* **180**, 255–270 (2015).
24. Duong, M. N., Erdes, E., Hebeisen, M. & Rufer, N. Chronic TCR-MHC (self)-interactions limit the functional potential of TCR affinity-increased CD8 T lymphocytes. *J Immunother Cancer* **7**, 284 (2019).
25. Spindler, M. J. *et al.* Massively parallel interrogation and mining of natively paired human TCRαβ repertoires. *Nat. Biotechnol.* **38**, 609–619 (2020).
26. Johnson, L. A. *et al.* Gene transfer of tumor-reactive TCR confers both high avidity and tumor reactivity to nonreactive peripheral blood mononuclear cells and tumor-infiltrating lymphocytes. *J. Immunol.* **177**, 6548–6559 (2006).
27. Chapuis, A. G. *et al.* T cell receptor gene therapy targeting WT1 prevents acute myeloid leukemia relapse post-transplant. *Nat. Med.* **25**, 1064–1072 (2019).
28. Kieke, M. C. *et al.* Selection of functional T cell receptor mutants from a yeast surface-display library. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 5651–5656 (1999).
29. Li, Y. *et al.* Directed evolution of human T-cell receptors with picomolar affinities by phage display. *Nat. Biotechnol.* **23**, 349–354 (2005).
30. Cameron, B. J. *et al.* Identification of a Titin-derived HLA-A1-presented peptide as a cross-reactive target for

engineered MAGE A3-directed T cells. *Sci. Transl. Med.* **5**, 197ra103 (2013).
31. Karapetyan, A. R. *et al.* TCR Fingerprinting and Off-Target Peptide Identification. *Front. Immunol.* **10**, 2501 (2019).
32. Malecek, K. *et al.* Engineering improved T cell receptors using an alanine-scan guided T cell display selection system. *J. Immunol. Methods* **392**, 1–11 (2013).
33. Wagner, E. K. *et al.* Human cytomegalovirus-specific T-cell receptor engineered for high affinity and soluble expression using mammalian cell display. *J. Biol. Chem.* **294**, 5790–5804 (2019).
34. Chervin, A. S., Aggen, D. H., Raseman, J. M. & Kranz, D. M. Engineering higher affinity T cell receptors using a T cell display system. *Journal of Immunological Methods* vol. 339 175–184 (2008).
35. Kessels, H. W., van Den Boom, M. D., Spits, H., Hooijberg, E. & Schumacher, T. N. Changing T cell specificity by retroviral T cell receptor display. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 14578–14583 (2000).
36. Schmitt, T. M. *et al.* Generation of higher affinity T cell receptors by antigen-driven differentiation of progenitor T cells in vitro. *Nat. Biotechnol.* **35**, 1188–1195 (2017).
37. Bentzen, A. K. *et al.* Large-scale detection of antigen-specific T cells using peptide-MHC-I multimers labeled with DNA barcodes. *Nat. Biotechnol.* **34**, 1037–1045 (2016).
38. Zhang, S.-Q. *et al.* High-throughput determination of the antigen specificities of T cell receptors in single cells. *Nature Biotechnology* vol. 36 1156–1159 (2018).
39. Gee, M. H. *et al.* Antigen Identification for Orphan T Cell Receptors Expressed on Tumor-Infiltrating Lymphocytes. *Cell* **172**, 549–563.e16 (2018).
40. Gee, M. H., Yang, X. & Christopher Garcia, K. Facile method for screening clinical T cell receptors for off-target peptide-HLA reactivity. doi:10.1101/472480.
41. Birnbaum, M. E. *et al.* Deconstructing the peptide-MHC specificity of T cell recognition. *Cell* **157**, 1073–1087 (2014).
42. Rius, C. *et al.* Peptide–MHC Class I Tetramers Can Fail To Detect Relevant Functional T Cell Clonotypes and Underestimate Antigen-Reactive T Cell Populations. *The Journal of Immunology* vol. 200 2263–2279 (2018).
43. Joglekar, A. V. *et al.* T cell antigen discovery via signaling and antigen-presenting bifunctional receptors. *Nat. Methods* **16**, 191–198 (2019).
44. Kisielow, J., Obermair, F.-J. & Kopf, M. Deciphering CD4 T cell specificity using novel MHC-TCR chimeric receptors. *Nat. Immunol.* **20**, 652–662 (2019).
45. Li, G. *et al.* T cell antigen discovery via trogocytosis. *Nat. Methods* **16**, 183–190 (2019).
46. Sharma, G., Rive, C. M. & Holt, R. A. Rapid selection and identification of functional CD8 T cell epitopes from large peptide-coding libraries. *Nat. Commun.* **10**, 4553 (2019).
47. Kula, T. *et al.* T-Scan: A Genome-wide Method for the Systematic Discovery of T Cell Epitopes. *Cell* **178**, 1016–1028.e13 (2019).
48. Border, E. C., Sanderson, J. P., Weissensteiner, T., Gerry, A. B. & Pumphrey, N. J. Affinity-enhanced T-cell receptors for adoptive T-cell therapy targeting MAGE-A10: strategy for selection of an optimal candidate. *Oncoimmunology* **8**, e1532759 (2019).
49. Sanderson, J. P. *et al.* Preclinical evaluation of an affinity-enhanced MAGE-A4-specific T-cell receptor for adoptive T-cell therapy. *Oncoimmunology* **9**, 1682381 (2020).
50. Sather, B. D. *et al.* Efficient modification of CCR5 in primary human hematopoietic cells using a megaTAL nuclease and AAV donor template. *Sci. Transl. Med.* **7**, 307ra156 (2015).
51. Mason, D. M. *et al.* High-throughput antibody engineering in mammalian cells by CRISPR/Cas9-mediated homology-directed mutagenesis. *Nucleic Acids Res.* **46**, 7436–7449 (2018).
52. Hockemeyer, D. *et al.* Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. *Nat. Biotechnol.* **27**, 851–857 (2009).
53. Robbins, P. F. *et al.* Single and dual amino acid substitutions in TCR CDRs can enhance antigen-specific T cell functions. *J. Immunol.* **180**, 6116–6131 (2008).
54. Luft, T. *et al.* Exogenous peptides presented by transporter associated with antigen processing (TAP)-deficient and TAP-competent cells: intracellular loading and kinetics of presentation. *J. Immunol.* **167**, 2529–2537 (2001).
55. Borbulevych, O. Y., Santhanagopolan, S. M., Hossain, M. & Baker, B. M. TCRs used in cancer gene therapy cross-react with MART-1/Melan-A tumor antigens via distinct mechanisms. *J. Immunol.* **187**, 2453–2463 (2011).
56. Schober, K. *et al.* Orthotopic replacement of T-cell receptor α- and β-chains with preservation of near-physiological T-cell function. *Nat Biomed Eng* **3**, 974–984 (2019).
57. Karanikas, V. *et al.* Monoclonal anti-MAGE-3 CTL responses in melanoma patients displaying tumor regression after vaccination with a recombinant canarypox virus. *J. Immunol.* **171**, 4898–4904 (2003).
58. Connerotte, T. *et al.* Functions of Anti-MAGE T-cells induced in melanoma patients under different vaccination modalities. *Cancer Res.* **68**, 3931–3940 (2008).
59. Wrenbeck, E. E. *et al.* Plasmid-based one-pot saturation mutagenesis. *Nat. Methods* **13**, 928–930 (2016).
60. Weon, J. L. & Potts, P. R. The MAGE protein family and cancer. *Current Opinion in Cell Biology* vol. 37 1–8 (2015).

61. Roth, T. L. *et al.* Reprogramming human T cell function and specificity with non-viral genome targeting. *Nature* **559**, 405–409 (2018).
62. Ren, J. *et al.* A versatile system for rapid multiplex genome-edited CAR T cell generation. *Oncotarget* **8**, 17002–17011 (2017).
63. Raman, M. C. C. *et al.* Direct molecular mimicry enables off-target cardiovascular toxicity by an enhanced affinity TCR designed for cancer immunotherapy. *Sci. Rep.* **6**, 18851 (2016).
64. Zhao, Y. *et al.* Cloning and chromosomal location of a novel member of the myotonic dystrophy family of protein kinases. *J. Biol. Chem.* **272**, 10013–10020 (1997).
65. Picelli, S. *et al.* Full-length RNA-seq from single cells using Smart-seq2. *Nat. Protoc.* **9**, 171–181 (2014).
66. Wagih, O. ggseqlogo: a versatile R package for drawing sequence logos. *Bioinformatics* vol. 33 3645–3647 (2017).