UNIVERSITY OF CALIFORNIA SAN DIEGO

The Impact of Dual TCRα Expression on Formation of the T Cell Repertoire Mediating Protection from Viral Infection

A thesis submitted in partial satisfaction of the requirements
for the degree Master of Science

in

Biology

by

Letitia Yang

Committee in charge:

Professor Gerald P. Morris, Chair
Professor Li-Fan Lu, Co-Chair
Professor Elina I. Zuniga

2021

The thesis of Letitia Yang is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2021

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

Chapter 1, in full, is the reprint of the material as it appears in Proceedings of the National Academy of Sciences 2020. Yang, Letitia; Jama, Burhan; Wang, Huawei; Labarta-Bajo, Lara; Zúñiga, Elina I.; Morris, Gerald P., PNAS, 2020. The thesis author was a co-author of this paper.

Chapter 2, in part, is currently being prepared for submission for publication of the material. Yang, Letitia; Morris, Gerald P. The thesis author was the primary investigator and author of this material.

Appendix, in full, is the reprint of the material as it appears in Proceedings of the National Academy of Sciences 2020. Yang, Letitia; Jama, Burhan; Wang, Huawei; Labarta-Bajo, Lara; Zúñiga, Elina I.; Morris, Gerald P., PNAS, 2020. The thesis author was a co-author of this paper.

ABSTRACT OF THE THESIS

The Impact of Dual TCRα Expression on Formation of the T Cell Repertoire Mediating Protection from Viral Infection

By

Letitia Yang

Master of Science in Biology

University of California San Diego, 2021

Professor Gerald P. Morris, Chair
Professor Li-Fan Lu, Co-Chair

The T cell receptor (TCR) plays a crucial role in T cell development, response, and homeostasis. A small population of T cells naturally express two TCRα clonotypes (dual TCRα cells). The study of dual TCRα cells had been limited by the lack of tools available to definitively identify them. A transgenic B6.TCRA-GFP/RFP mouse model linking fluorophores to TCRα constant region to enable identification of dual TCRα cells was developed by our lab. In Chapter

1, I hypothesized that dual TCRα cells influence immune response against foreign antigens. LCMV-Armstrong infection model revealed that dual TCRα cells expanded significantly from pre-infection baseline of ~18% to 40-50% of virus-specific T cells at 8 dpi. Dual TCR frequency at 28 dpi remained higher than pre-infection among LCMV-specific $CD4^+$ effector memory and central memory subsets. These findings demonstrate that dual TCRα expression influences protective antiviral responses.

In Chapter 2, I hypothesized that dual TCRα cells contain a distinct repertoire of receptors. Bulk sequencing revealed that peripheral dual TCR repertoire contained unique clonotypes not found in the single TCR repertoire. Relatively low correlation, similarity, and overlap were observed between the single and dual TCR clones. The $CD4^+$ dual TCRα repertoire was enriched for clones associated with autoantigen epitopes. These results support my hypothesis that the dual TCR repertoire contains unique clonotypes and suggest that dual TCRα cells may have increased potential for autoreactivity. Altogether, these findings demonstrate that dual TCR cells play a critical role in mediating antiviral immunity and T cell repertoire formation.

GENERAL INTRODUCTION

T cells recognize peptide antigens presented by self major histocompatibility molecules (MHC) via highly specific T cell receptors (TCR) (1). The DNA sequences encoding these receptors are not germline-encoded, but rather generated in somatic cells via non-homologous recombination of TCR gene segments during thymocyte development. This enables the production of a vastly diverse repertoire (in mice and humans) capable of recognizing a diverse array of antigens. T cell recognition of antigens presented by MHC is crucial for activation of T cell response against their specific foreign antigen (1, 62).

During development, immature thymocytes undergoing thymopoiesis begin as $CD4^-CD8^-$ double-negative (DN) cells, where recombination of variable (V), diversity (D), and joining (J) TCRβ gene segments occur to generate a functional TCRβ chain (63). Successful production of TCRβ induces thymocyte proliferation and differentiation into $CD4^+CD8^+$ double-positive (DP) cells, where TCRα V-J recombination occurs to generate a TCRα chain. The ability of the newly generated TCR to interact with self-MHC is tested via positive selection, which retains cells bearing a TCR capable of low affinity interaction with a limited repertoire of self-peptides presented by self-MHC. Failure to engage in such interaction results in death by neglect, while engagement with sufficient affinity allows thymocytes to mature as $CD4^+$ or $CD8^+$ single-positive (SP) T cells (63). In negative selection, thymocytes bearing TCRs that bind to self-peptide-self-MHC complexes with high affinity are eliminated to avoid generation of potentially autoreactive T cells capable of responding to self-antigens (63). The interaction between the TCR and self-peptide-self-MHC complex is therefore crucial for ensuring the generation of functional TCRs and elimination of autoreactive T cells.

Unlike TCRβ rearrangement, both TCRα loci remain in an open chromatin state in DP thymocytes until they receive a positively-selecting signal. This allows for multiple chances to produce a functional TCRα (64). The absence of allelic exclusion in TCRα rearrangement enables the possibility for two functional TCRs with distinct clonotypes to be expressed on a T cell during normal thymic development, which gives rise to the generation of dual TCRα cells (2, 3, 65). Dual TCRα expression has previously been shown to increase auto- and alloreactivity by promoting positive selection and rescuing autoreactive cells from negative selection (8, 11, 14, 15, 17). Furthermore, studies on the effects of dual TCR expression on shaping the T cell repertoire have suggested that dual TCR cells may contain unique antigen specificities, and that the presence of a secondary TCR may expand the peripheral T cell repertoire (8, 13-18). Despite evidence of the contributions of dual TCRα cells in protective immune response, their specific role in memory generation and T cell repertoire formation required further investigation.

Previous studies on dual TCR function relied on genetic elimination of dual TCRα cells from the T cell repertoire by knocking out one chromosomal copy of TRAC (encoding the constant region of the TCRα protein) (13, 17). However, this method may not accurately reflect dual TCRα function under normal physiology since the polyclonal nature of the T cell repertoire could compensate for the absence of dual TCRα cells. Pairwise labeling with monoclonal antibodies (mAbs) has been used to study dual TCRα cells in intact T cell repertoires, though this approach is limited by the reagents available, covering only approximately 13-15% of mouse and human TCRα repertoires (2). Single-cell RNA sequencing has been used to identify cells with two TCR gene rearrangements, but this technique damages the cell, which prevents mechanistic investigation of dual TCRα cells (20, 21). To overcome these limitations, the B6.TCRA-GFP/RFP mouse model was developed by linking GFP and RFP fluorophores to the TCRα constant region.

I hypothesized that this system may enable a more accurate determination of how dual TCRα expression shapes T cell repertoires and antigen-specific memory responses during acute lymphocytic choriomeningitis virus (LCMV) infection.

CHAPTER 1: TCRα REPORTER MICE REVEAL CONTRIBUTION OF DUAL TCRα EXPRESSION TO T CELL REPERTOIRE AND FUNCTION

Proceedings of the National Academy of Sciences of the United States of America, 2020

1.1: Introduction

The T cell receptor (TCR) clonotype present on a T cell determines reactivity to specific peptide–major histocompatibility complex (pMHC) ligands, which in turn directs development, function, and homeostasis (1). Thus, T cell identity and function are intrinsically linked to TCR clonotype. Conventional T cells bear a single TCR clonotype formed as a heterodimer of TCRα and TCRβ proteins. However, a subset of T cells expresses two functional TCRαβ receptors (2–4). Co-expression of two TCRs results from incomplete allelic exclusion of TCRα and TCRβ gene loci during thymopoiesis (4–7). A prevailing view of dual TCR expression as a by-product of TCR gene rearrangement posits that it affects only a small (1 to 10%) subset of T cells and does not have significant functional consequence. However, a growing body of evidence indicates that dual TCRα-expressing T cells contain a unique repertoire of TCRα clonotypes (8) and that these cells may have distinct potential to respond to ligands such as auto- or alloantigens (8–18). Despite evidence of the involvement of dual TCRα cells in pathogenic responses including autoimmunity and graft-versus-host disease, they remain understudied due to limitations in definitively identifying and isolating these cells.

Transgenic TCR systems have demonstrated the potential for dual TCR co-expression to enable emergence of pathologic dual receptor cells bearing clonotypes that would otherwise be eliminated during thymic selection (9, 11, 14, 15, 17). However, the effect of dual TCR co-expression on naturally developing T cell repertoires has been more difficult to evaluate. Genetic

elimination of dual TCRα expression, commonly by knockout of one chromosomal copy of TRAC, has exhibited heterogenous effects on T cell development and function (8, 10, 16, 19). Studies based on genetic elimination of dual TCRα cells may not accurately reflect the role of dual receptor cells in normal physiology, as the breadth of the TCR repertoire could compensate for loss of specific subsets. Evaluation of dual TCRα cells in intact T cell repertoires has relied on pairwise labeling with monoclonal antibodies (mAbs) against TCRVα (2, 4). This approach is critically limited by a paucity of reagents (covering only ~13% of mouse and human TCRVα), requiring extrapolation and potentially biased estimation of dual TCRα cells. Single-cell RNA sequencing can provide unbiased evaluation of cells with two TCR gene rearrangements, though analysis pipelines often filter multiple TCR sequences from individual cells (20–22), leading to underestimation of dual TCR cell frequencies. Furthermore, presence of in-frame TCR gene rearrangements may not necessarily translate into functional protein (23). Finally, single-cell sequencing is a terminal event for the cells studied and precludes further functional testing, hindering mechanistic investigation.

To overcome these limitations, we generated transgenic B6.TCRA-GFP/RFP (green fluorescent protein/red fluorescent protein) mice with TCRα protein generated from one chromosomal copy of TRAC labeled with GFP and TCRα protein generated from the other chromosomal copy labeled with RFP. Using this system, single and dual TCRα cells are unequivocally identifiable. This system defines dual TCRα cells as a much larger component of the naive T cell repertoire than previously appreciated. Importantly, this system enables the discovery that dual TCRα cells have a selective contribution to protective immune response and subsequent memory formation during viral infection.

1.2: Results

Generation of Fluorophore-Tagged TCRα Reporter Mice.

Most (>90%) dual TCR cells result from production of two functional TCRα proteins (2, 3), each paring with the cell's single TCRβ. To overcome the limitations of existing approaches for identification of single- and dual-TCR cells, we used CRISPR-Cas9–mediated gene editing (24) to insert genes encoding enhanced GFP (GFP) and tdTomato (RFP) at the 3′ end of exon 3 of TRAC [exon 4 is not translated (25)] to generate two transgenic mouse lines (Appendix, Figure S1A). Reporter genes were attached to the C-terminal end of TCRCα by an 18-amino-acid flexible linker protein to colocalize reporters to the TCRα while avoiding interference with TCR trafficking or function (26). Double-stranded DNA plasmid containing reporter gene constructs flanked by 1-kb homology arms of genomic TRAC sequence, CRISPR RNA, guide RNA, and Cas9 protein were injected by micropipette into ~200 C57BL/6 (B6) embryos. Flow cytometry of peripheral blood leukocytes identified 1/11 pups expressed TCRα-GFP and 1/26 pups expressed TCRα-RFP. Founder mice with individual reporters were bred to generate B6.TCRA-GFP and B6.TCRA-RFP lines homozygous for TCRα reporters.

The insertion site of the GFP and RFP reporter genes was confirmed by sequencing of the flanking DNA regions. Expression of GFP and RFP in transgenic mice is specific for CD4+ and CD8+ T cells (Appendix, Figure S1B). Confocal microscopy confirmed that GFP and RFP reporters localize to the T cell membrane (Appendix, Figure S1C). GFP and RFP reporters are first detectable at low levels in double-positive $CD4^+CD8^+$ (DP) thymocytes and increase in expression throughout maturation to single-positive $CD4^+CD8^-$ (CD4SP) and $CD4^-CD8^+$ (CD8SP) thymocytes (Appendix, Figure S1 D–F), consistent with increasing TCRα expression during

thymopoiesis. B6.TCRA-GFP and B6.TCRA-RFP mice had normal thymic development, with equivalent generation of CD4SP and CD8SP thymocytes at frequencies similar to wild-type B6 mice (Appendix, Figure S1G). Reporter mice had normal frequencies of $CD4^+$ and $CD8^+$ T cells in the spleen (Appendix, Figure S1H). $CD4^+$ and $CD8^+$ T cells from reporter mice proliferated equivalently to T cells from wildtype B6 mice following in vitro stimulation with plate-bound anti-CD3 and anti-CD28 (Appendix, Figure S1 I and J), indicating that GFP and RFP do not interfere with TCR function.

B6.TCRA-GFP/RFP Mice Reveal High Frequency of Dual TCR Expression.

B6.TCRA-GFP and B6.TCRA-RFP mice enable evaluation of the entire repertoire of TCRα expression by flow cytometry, removing the limitation of anti-TCRVα mAb labeling. Interbreeding TCRα reporter transgenic lines to produce mice with one TRAC chromosomal copy containing the GFP reporter and the other containing the RFP reporter (B6.TCRA-GFP/RFP) provided the ability to distinguish T cells expressing a single GFP or RFP TCRα from those co-expressing two TCRα chains. Flow cytometry of splenocytes from B6.TCRA-GFP/RFP mice clearly identified populations of $CD4^+$ and $CD8^+$ T cells co-expressing GFP- and RFP-labeled TCRα (Figure 1.1 A). The frequency of GFP^+RFP^+ dual receptor cells was significantly higher than established consensus estimates of 1 to 10% (2–4), with 16.7 ± 1.3% of $CD4^+$ and 16.9 ± 0.9% (mean ± SD) of $CD8^+$ T cells co-expressing dual TCRα (Figure 1.1 B). Confocal microscopy demonstrated colocalization of TCRα-GFP and TCRα-RFP at the cell membrane in GFP^+RFP^+ cells (Figure 1.1 C). This suggests that the high frequency of dual GFP and RFP expression measured by flow cytometry is not a result of nonproductive TCRα protein not expressed on the cell surface but represents true co-expression. The frequency of dual-GFP^+RFP^+ T cells identified

by confocal microscopy was 16.0 ± 1.0% (Figure 1.1 D), similar to that observed by flow cytometry. However, it is possible that cells appearing as single-TCRα cells have low levels of a second TCRα protein below the threshold for detection by these methods, suggesting that the frequencies of dual TCRα cells detected here represent minimums with a possibility for cells expressing very low levels of second TCRα proteins.

Cell-surface expression of TCR proteins is a dynamic process, which can be both negatively and positively regulated (27). Studies in Jurkat cells have demonstrated that cell-surface levels of co-expressed TCRs can be differentially regulated during T cell stimulation (28), and nonengaged TCRαβ proteins can be actively recruited to the immune synapse at the cell surface (29). To examine the potential for single GFP^+ or RFP^+ cells to harbor second TCRα proteins that could be up-regulated to the cell surface, we isolated GFP^+RFP^- and GFP^-RFP^+ T cells from B6.TCRA-GFP/RFP mice by FACS (fluorescence-activated cell sorting) and examined TCRα co-expression following 5-d stimulation with anti-CD3/anti-CD28 mAbs (Appendix, Figure S2 A and B). In vitro stimulation resulted in 3.3 ± 1.0% of GFP^+RFP^- and 1.2 ± 0.5% of GFP^-RFP^+ cells expressing a second TCRα detectable by flow cytometry. We do not attribute changes in reporter expression to induced transcription/translation of nonfunctional TRAV-J rearrangements, as the fluorophore reporters are linked to TCRCα as single-chain proteins and nonproductive rearrangements would likely result in the protein's being out of frame for translation. We also do not attribute dual TCRα expression to trogocytosis (30), as in vitro stimulation of cocultured T cells from B6.TCRA-GFP and B6.TCRA-RFP mice did not produce dual GFP^+RFP^+ cells (Appendix, Figure S2 C and D). These data support that our dual-transgenic system robustly identifies single- and dual-TCRα T cells.

Expression of Dual TCRα Is Associated with Positive Selection of DP Thymocytes.

Co-expression of dual TCRα results from a lack of TCRα allelic exclusion in DP thymocytes. It has been presumed that ~33% of thymocytes have two in-frame TRAV-J rearrangements capable of producing functional TCRα protein, though frequencies of mature thymocytes co-expressing two TCRα proteins measurable by mAb labeling is significantly lower at 5 to 7% (23). Similar to results from identification of dual TCRα peripheral T cells (Figure 1.1), flow cytometry analysis of thymocytes from B6.TCRA-GFP/RFP mice demonstrates significantly higher frequencies of dual TCRα CD4SP (18.0 ± 1.4%) and CD8SP (19.5 ± 2.0%) (Figure 1.2 A and B) than previously estimated.

The ability to recombine TRAV-J on both chromosomal loci and express dual TCRα has been associated with facilitating positive selection and promoting production of mature CD4SP and CD8SP (8, 19). However, these results are from $TRAC^{+/-}$ systems, which may not directly reflect the effect of dual TCRα expression; decreased positive selection and thymocyte maturation associated with loss of dual-chromosome TRAV-J recombination may reflect an inability to generate a single productive TCRα rather than an increased ability of dual TCRα cells to successfully undergo positive selection. To investigate this, we measured the frequency of dual TCRα expression in $TCR^{low}CD69^{+}$, $TCR^{high}CD69^{+}$, and $TCR^{high}CD69^{low}$ stages of DP thymocytes selection (GFP and RFP reporter expression in $TCR^{low}CD69^{-}$ DP thymocytes was beneath the threshold for reliable detection by flow cytometry; Appendix, Figure S1 D–F) (Figure 1.2 C and D and Appendix, Figure S3A). Compared to preselection $TCR^{low}CD69^{+}$ cells (14.0 ± 5.6%), dual TCRα frequency was significantly increased among $TCR^{high}CD69^{+}$ (22.6 ± 6.1%, $P < 0.05$) and $TCR^{high}CD69^{low}$ DP thymocytes successfully undergoing positive selection (18.3 ± 8.5%, $P < 0.05$). These data are consistent with previous findings and suggest that dual TCRα protein expression,

rather than just the ability to recombine two sets of TRAV-J gene segments, promotes positive selection.

The positive effect of dual TCRα co-expression on positive selection contrasts with a paradoxical observation that dual TCRα expression impairs agonist selection of regulatory FoxP3^{+}CD4^{+} T cells (Treg) (19). Agonist selection of thymocytes, essential for thymic generation of Tregs, invariant natural killer T (NKT) cells, and CD8αα^{+} intraepithelial T cells, results from high-affinity interactions with self-pMHC ligands (31). Thymic Tregs in B6.TCRA-GFP/RFP mice demonstrated frequencies of dual TCRα expression similar to CD4SP thymocytes (Figure 1.2 E and F). Thymic NKT cells had similar frequencies of dual TCRα expression (Appendix, Figure S3 B and C). Consistent with the lack of observable effect of dual TCRα expression on agonist selection, we did not observe increased expression of CD5, a surrogate marker for TCR signal strength in response to positively selecting self-pMHC (32–34), on dual TCRα CD4SP or CD8SP (Appendix, Figure S3 D and E). These data suggest that while dual TCRα expression promotes recognition of self-pMHC during positive selection it may not result in a generally increased affinity/avidity for selecting pMHC ligands.

Dual TCR T Cells Have Unchanged Total TCR Expression and Function but Increased CD5.

Previous descriptions of TCR allelic inclusion have indicated that co-expression of two TCR clonotypes does not increase total antigen receptor expression at the cell surface (2, 35). This likely results from TCR complex stoichiometry regulating cell-surface expression via limited availability of CD3 proteins (36). Measurement of CD3 by flow cytometry as an assessment of total TCR expression confirmed that GFP^{+}RFP^{+} dual TCRα cells expressed similar amounts of CD3 as compared to single TCR cells (Figure 1.3 A and B). Similarly, co-expression of two TCR

clonotypes had no effect on T cell response to non-specific TCR stimulation, as graded doses of anti-CD3/anti-CD28 mAbs or staphylococcal enterotoxin B (SEB) resulted in equivalent in vitro proliferation of single and dual TCRα cells (Appendix, Figure S4). These data confirm previous observations that co-expression of two TCR clonotypes does not confer a general advantage for TCR stimulation (14, 18, 35).

TCR is required not only for recognition of foreign antigens but also for recognition of self-pMHC ligands for homeostatic maintenance of T cells (1). The affinity of this interaction can be estimated as proportionate to the cell-surface expression of CD5 (34). We have previously demonstrated that dual TCRα cells identified by anti-TCRVα mAb co-labeling express higher levels of CD5 than single TCRα cells. This correlates with increased dual TCRα T cell potential for acute lymphopenia-induced proliferation (35), a function dependent on TCR interaction with self-pMHC ligands (37, 38). Similar to our previous investigations, dual TCRα cells from B6.TCRA-GFP/RFP mice demonstrated increased expression of CD5 on the cell surface as compared to single TCRα $CD4^+$ and $CD8^+$ cells (Figure 1.3 C and D). This effect was more pronounced in $CD4^+$ dual TCRα cells, which had an average 31.6% higher expression of CD5 than single TCRα cells from the same mouse. Comparatively, $CD8^+$ dual TCRα cells demonstrated an average 12.4% increase in CD5 expression. This contrasts to the absence of any difference in CD5 expression by dual TCRα CD4SP or CD8SP (Appendix, Fig. S3 D and E). However, CD5 expression and its effects on TCR signaling can be changed in response to self-pMHC–induced tonic signaling in the periphery (39), providing a plausible explanation for this apparent discrepancy.

Increased expression of CD5 by dual TCRα T cells in the periphery suggests that they may have an increased reactivity with self-antigens. However, dual TCRα co-expression was not

observed at increased frequencies among T cell subsets associated with reactivity against self-ligands including splenic Tregs (Appendix, Figure S5 A and B), NKT cells (Appendix, Figure S5 C and D), intestinal $CD4^+$ and $CD8^+$ T cells (Appendix, Figure S5 G and H), intestinal Tregs (Appendix, Figure S5 I and J), and $CD8\alpha\alpha^+$ T cells (Appendix, Figure S5 K and L). The frequency of splenic T cells from immunologically naive mice expressing high levels of CD44, a marker of prior activation in response to antigen, was also similar between single and dual TCR $CD4^+$ and $CD8^+$ T cells from immunologically naive B6.TCRA-GFP/RFP mice (Figure 1.3 E and F). Similarly, dual TCRα cells were not observed at increased frequencies among $CD44^{high}CD62L^-$ effector memory (EM) or $CD44^{low}CD62L^+$ central memory (CM) cells in immunologically naive mice (Appendix, Figure S5 E and F). However, dual TCRα T cells in the spleen were more frequently positive for the early activation marker CD69 ($CD4^+$ 10.0 ± 2.7% dual TCRα vs. 5.8 ± 1.6% single TCRα, $P < 0.005$; $CD8^+$ 2.2 ± 1.1% dual TCRα vs. 1.3 ± 0.4% single TCRα, $P < 0.005$) (Figure 1.3 G and H). Concurrently, a higher proportion of dual TCRα $CD4^+$ cells had an anergy-associated $CD73^{high}FR4^{high}$ phenotype (40) (6.6 ± 2.3% dual TCRα vs. 3.2 ± 1.3% single TCRα, $P < 0.005$) (Figure 1.3 I and J). Together these data support a model where dual TCRα expression may promote increased reactivity against self-pMHC antigens, but the biologic effect is limited at an immunologic steady state via up-regulation of the negative regulator of TCR signaling CD5 and induction of anergy.

Dual TCR Expression Promotes Protective Immune Response to Lymphocytic Choriomeningitis Virus Infection.

Identification of dual TCRα populations significantly larger than previous estimates suggests that there may be unappreciated effects of dual receptor expression on immune responses.

To examine the role of dual TCRα cells in a protective immune response, B6.TCRA-GFP/RFP mice were infected with murine lymphocytic choriomeningitis virus (LCMV) Armstrong strain. Infected mice were killed 8 d after infection to evaluate the acute immune response (41). Dual TCRα cells were increased in the spleen and lymph nodes for total dual TCRα $CD4^+$ cells (23.8 ± 3.9%, $P < 0.001$) and $CD8^+$ cells (23.4 ± 5.3%, $P < 0.005$) compared to immunologically naive mice (Figure 1.4 A). This indicates a potential advantage for dual TCRα expression during an immune response, though antigen specificity in the broad population is undefined.

LCMV-specific T cells were identified from spleen and lymph node cells by flow cytometry for binding I-A^b:$GP_{66–77}$ and H2-D^b:$GP_{33–41}$ tetramers (Figure 1.4 B and C). While labeling with these tetramers does not exhaustively identify all LCMV-specific T cells, it enables examination of defined LCMV-specific T cells. Focusing on the LCMV-specific response revealed more dramatic expansion of virus-specific dual TCR cells during infection. LCMV $tetramer^+$ $CD4^+$ and $CD8^+$ cells contained significantly higher frequencies of dual TCRα cells ($CD4^+$ 48.1 ± 11.1%, $P < 0.005$; $CD8^+$ 36.5 ± 8.7%, $P < 0.005$) than tetramer-negative populations within individual mice (Figure 1.4 D and E). Dual TCRα cell populations in tetramer-negative cells were increased as compared to uninfected mice ($CD4^+$ 22.4 ± 3.0%, $P < 0.05$; $CD8^+$ 25.0 ± 4.8%, $P < 0.05$), similar to the observed effect in total T cells (Figure 1.4 A). This likely reflects a broad response against multiple LCMV antigens. The dramatic specific expansion of LCMV-specific dual TCRα cells in $tetramer^+$ populations demonstrates a selective benefit of dual TCRα expression to the protective immune response.

Within LCMV tetramer-specific T cell populations, dual TCRα expression directed differing phenotypes for $CD4^+$ and $CD8^+$ T cells. LCMV-specific $CD4^+$ dual TCRα cells had higher frequencies of $CD11a^+CD49d^+$ cells, associated with an activated T cell response during

acute LCMV infection (42), compared to single TCRα cells from the same mouse (Figure 1.4 F and G). In LCMV-specific $CD8^+$ cells, dual TCRα expression also affected effector cell differentiation, though in an opposite manner than for $CD4^+$ cells (Figure 1.4 H and I), as dual TCRα expression was associated with lower frequencies of KLRG-1^{high}Ly6c$^+$ terminal effector $CD8^+$ cells (43). These data suggest that while dual TCRα expression may promote responses in both $CD4^+$ and $CD8^+$ T cells, dual TCRα expression may differentially affect the nature of those responses.

Binding of pMHC tetramers is considered indicative of reactivity against a specific antigen. However, tetramer can have nonspecific binding with T cells that do not have measurable functional response against the antigen or, conversely, can miss antigen-specific T cells with low affinity for antigen (44, 45). To address this, functional responses against LCMV were assessed by measurement of cytokine production after ex vivo antigen-specific stimulation of splenocytes from infected mice. Similar to the tetramer data, $CD4^+$ dual TCRα cells from infected mice demonstrated a preferential functional response. Stimulation of splenocytes with LCMV $GP_{66–77}$ peptide resulted in higher frequencies of dual TCRα cells producing interferon gamma (IFNγ) and tumor necrosis factor alpha (TNFα) compared to single TCRα cells in the same culture (Figure 1.5 A and B). Responding single and dual TCRα cells produced equivalent levels of cytokine on a per-cell basis (Appendix, Figure S6A). Dual TCRα expression had similar effects on cytokine production by $CD8^+$ cells in response to LCMV $GP_{33–41}$ stimulation. Dual TCRα cells had higher frequencies of cells producing IFNγ and TNFα but not granzyme B in response to antigenic stimulation (Figure 1.5 C and D). Responding single and dual TCRα $CD8^+$ cells produced equivalent amounts of cytokines (Appendix, Figure S6B). These data are consistent with the LCMV tetramer data and support that dual TCRα expression favors functional participation in the

protective antiviral immune response. The data also indicate that dual TCRα cells do not have an increased functional capacity but rather an increased ability to respond to antigen.

Dual TCR Expression Differentially Affects Persistence and Memory of CD4$^+$ and CD8$^+$ T Cell Immune Responses.

In the acute phase of the immune response against LCMV, dual TCRα expression promoted participation in the immune response for both CD4$^+$ and CD8$^+$ T cells. We evaluated whether dual TCRα expression persistently affected the antiviral immune response by examining dual TCRα T cells 28 d after infection with LCMV Armstrong (Figure 1.6A). Dual TCRα CD4$^+$ T cells (20.0 ± 1.9%, $P < 0.001$) remained increased compared to immunologically naive mice. Conversely, dual TCRα CD8$^+$ T cells (16.8 ± 2.9%) returned to baseline frequencies. This pattern was consistent when examining LCMV-specific T cells identified by flow cytometry for binding I-A^b:GP$_{66–77}$ and H2-D^b:GP$_{33–41}$ tetramers (Figure 1.6 B and C). At 28 d after infection, LCMV-specific dual TCRα cells remained significantly increased among LCMV tetramer$^+$ CD4$^+$ T cells (33.8 ± 11.7%, $P < 0.005$) as compared to tetramer-negative cells, while CD8$^+$ cells did not demonstrate the same persistence of LCMV tetramer$^+$ dual TCRα cells (13.5 ± 5.3%).

LCMV-specific dual TCRα cells among both CD4$^+$ and CD8$^+$ T cells predominantly (78.4 ± 16.7% and 89.3 ± 7.6%) had an EM phenotype 28 d after infection (Figure 1.6 D). Examining EM and CM cells after resolution of LCMV infection revealed that dual TCRα cells were markedly increased in both total and I-A^b:GP$_{66–77}$–specific EM (total 26.7 ± 1.6%, Tet$^+$ 32.6 ± 12.6%) and CM (total 25.7 ± 11.0%, Tet$^+$ 41.6 ± 24.6%) CD4$^+$ populations (Figure 1.6 E). A similar effect was not seen for CD8$^+$ EM and CM populations, where dual TCRα cells were found at frequencies similar to pre-infection (total EM 18.0 ± 3.8%, Tet$^+$ EM 12.5 ± 3.7%, total CM 15.1

± 3.4%, Tet$^+$ CM 18.1 ± 9.9%). While dual TCRα expression did not seem to provide a benefit to the persistence of LCMV-specific CD8$^+$ cells or their differentiation into memory cells, the difference between dual TCRα cell frequency among total CD4$^+$ T cells, which returned to pre-infection frequencies after the resolution of infection, and the increased frequency of dual TCRα cells among total CD4$^+$ EM and CM cells indicates that dual TCRα expression affects the long-term fate of CD4$^+$ T cell function during protective immune response and may promote CD4$^+$ memory formation.

1.3: Discussion

These data uncover unappreciated roles for dual TCRα expression in the function of the immune system. The physiologic impact of dual TCRα expression has been debated, as genetic elimination of dual TCRα co-expression (as in TCRC$\alpha^{+/-}$ mice) does not broadly change the composition of peripheral T cell numbers or subsets and does not completely eliminate reactivity against any antigen tested. This singularly reductionist view represents an unachievable standard, as the breadth of the TCR repertoire (46) and the flexibility with which TCRs interact with pMHC ligands (47) are likely to preclude absolute "present or absent" effects of changes to the TCR repertoire. Our approach enabled unambiguous identification of single and dual TCRα cells from the intact T cell repertoire, defining the dual TCRα cell subset as ~16% of peripheral T cells in immunologically naive adult mice (Figure 1.1). This is significantly higher than traditional estimates of 1 to 10% (2, 10) and in line with our recent description of ~18% of T cells from healthy adult humans having two in-frame TRAV-J gene rearrangements identified by single-cell RNA sequencing (22).

Generation of T cells co-expressing two TCRα proteins results from the simultaneous rearrangement of both TCRα loci in DP thymocytes. Allelic inclusion of TCRα has been demonstrated to facilitate positive selection (8, 19), though it is undefined whether this depends on generation of two TCRα proteins or if it reflects an increased efficiency of having twice the opportunity to generate a TCRα protein capable of pairing with TCRβ and responding to the limited repertoire of selecting self-pMHC ligands in the thymic cortex. Our data demonstrating that thymocytes expressing two TCRα proteins are significantly increased among TCR^{high} post-selection DP thymocytes (Figure 1.2) suggest that the benefit of TCRα allelic inclusion is associated with co-expression of two TCRα proteins, rather than an increased efficiency of

generating a single productive TCRα. Transgenic TCR systems have demonstrated that dual TCR co-expression can have either potentiating or inhibitory effects on TCR signaling in response to agonist ligand (48–50). Our current data do not indicate that dual TCRα expression affects the strength of the signal with positively selecting ligands, as expression of CD5 is unaffected by dual TCRα expression on mature thymocytes (Appendix, Figure S3). However, CD5 is only a surrogate marker and future studies examining TCR signaling pathways and more discriminatory measures of TCR signal strength such as Nur77 (51) will be important to more thoroughly evaluate this. In similar form, dual TCRα expression does not seem to promote agonist selection (Figure 1.2 and Appendix, Figure S3). Data from a $TRAC^{+/-}$ model have demonstrated that TCRα co-expression impairs agonist selection of Tregs (19). It was hypothesized that this occurred via alteration of TCR signal strength in response to positively selecting pMHC. Future studies using our model to dissect how TCRα co-expression affects TCR signaling during thymocyte development and whether this imprints future functional ability on dual TCRα cells will be important for understanding this process.

The robust ability of our model to identify dual TCRα cells enabled discovery of their function during a physiologic immune response. LCMV Armstrong infection experiments revealed that dual TCRα cells were unexpectedly prominent contributors to the antiviral immune response, as measured by recognizing I-A^b:$GP_{66–77}$ and H2-D^b:$GP_{33–41}$ tetramers, as well as by ex vivo stimulation of cells with LCMV $GP_{66–77}$ and $GP_{33–41}$ antigens (Figure 1.4 and 1.5). Expansion of dual TCRα cells during acute infection also extended outside of cells reactive to these antigens (Figure 1.4). It remains to be determined whether this represents selective dual TCRα responses against other LCMV antigens or if dual TCRα cells are expanding via off-target or bystander activation. This difference could have important consequences related to epitope spreading and

heterologous immunity. Our work demonstrating the alloreactive potential of dual TCRα cells would suggest that this risk could extend to heterologous immune responses against alloantigens, which could present significant risk for transplant rejection or graft-versus-host disease.

The mechanism through which dual TCRα cells are selectively increased in number and function during an immune response is unclear. An intuitive notion that co-expression of two TCR clonotypes could broaden the antigenic reactivity of a given T cell would suggest that dual TCRα cells are expanded in response to infection simply because they have a second opportunity to recognize a viral antigen. However, we have previously demonstrated that dual receptor cells contain a unique subset of TCRs (8). These are presumably TCR clonotypes that would otherwise be removed during thymic selection but instead emerge as co-expressed with TCRs capable of independently mediating positive selection or downregulating or "masking" the cross-reactive or autoreactive TCR from negative selection (9, 11, 14, 15, 17). This implies that dual TCRα cells potentially harbor highly cross-reactive or self-reactive TCRs (52, 53). The broad expansion of dual TCRα cells not specific for LCMV $GP_{66–77}$ and $GP_{33–41}$ antigens (Figure 1.4) may be partly a result of this type of cross-reactivity, in addition to reactivity against other LCMV epitopes. Further studies to interrogate the biochemistry of ligand recognition by dual TCR cells participating in immune responses are necessary.

The data here present evidence that co-expression of dual TCRs can have qualitatively different effects for $CD4^+$ and $CD8^+$ T cells during an immune response. Both $CD4^+$ and $CD8^+$ dual TCRα cells expanded robustly during the acute immune response (Figure 1.4), but $CD4^+$ dual TCRα cells remained significantly increased among LCMV-specific cells at 28 d after infection, while LCMV-specific $CD8^+$ cells returned to baseline frequencies (Figure 1.6). Both LCMV-specific $CD8^+$ and $CD4^+$ T cells were predominantly EM cells, with a significant contribution of

CM cells, though only in CD4+ T cells were the post-infection EM and CM compartments enriched for dual TCRα cells (Figure 1.6). This difference may reflect the different kinetics of expansion and contraction of virus-specific cells, with a delayed contraction of CD4+ LCMV-specific cells compared to CD8+ cells (41). However, the quality of TCR–ligand interaction affects T cell effector function, determination of the responding T cell repertoire, and potential for memory formation (54, 55). In LCMV infection, CD4+ cell function has been positively correlated with affinity for viral antigens (56), and the results presented here could reflect an increased function of dual TCRα cells after infection driven by higher affinity for viral antigens. Again, biochemical studies of ligand recognition by dual TCRα cells are necessary to evaluate this possibility.

An alternate hypothesis for the effect of dual TCRα expression on T cell function during acute infection and selective promotion of memory cells after the resolution of infection is that this may result from an increased potential for reactivity against self-pMHC ligands, rather than differential affinity for viral antigens. Results from other models indicate that reactivity of TCR for self-pMHC ligands, including as evidenced by increased expression of CD5, correlates with the formation and persistence of protective CD4+ T cell responses and memory formation (39, 57–59). Here (Figure 1.3), as well as in prior studies (35), we have identified that dual TCRα cells express higher levels of CD5 than single TCRα cells and that this difference is more pronounced in CD4+ dual TCRα cells. We have previously confirmed the functional impact of this in acute lymphopenia-induced proliferation, a recognized effect of affinity for self-pMHC (35). This difference in peripheral T cells is in contrast to a lack of observable difference in CD5 expression between single and dual TCRα CD4SP and CD8SP (Appendix, Figure S3). However, CD5 expression can change in response to self-pMHC–induced tonic signaling in the periphery (39). We propose this difference may be related to a more narrow window of permissive reactivity

against self-pMHC in the thymus compared to the periphery. It may also result from a wider range of self-antigens present in peripheral tissues as compared to the thymus. Mechanisms potentially underlying how dual TCR co-expression affects reactivity against self-ligands, including not only the potential for autoreactivity but also coagonism and TCR signal attenuation by CD5, are essential questions for continued study.

1.4: Figures

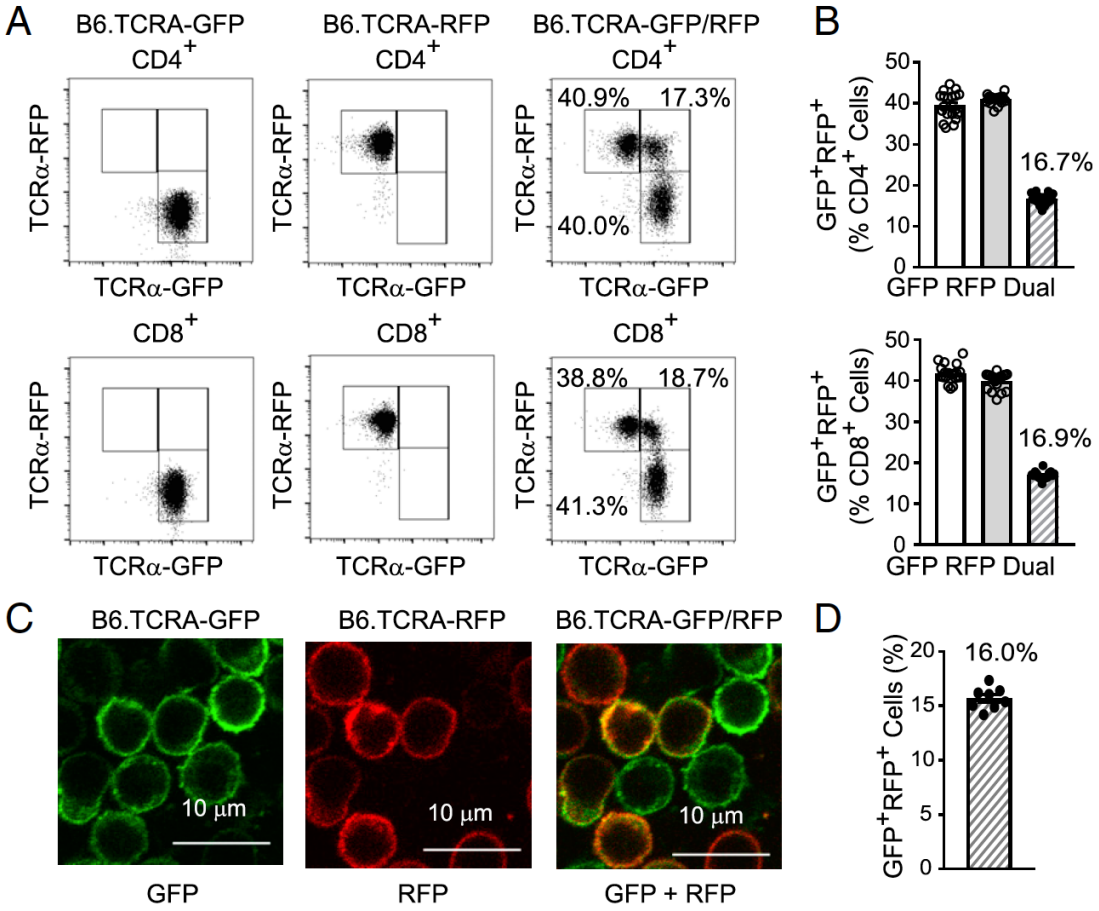

Figure 1.1: B6.TCRA-GFP/RFP mice identify high frequency of dual TCRα cells. (A) Identification of GFP$^+$, RFP$^+$, and dual GFP$^+$RFP$^+$ cells from adult B6.TCRA-GFP, B6.TCRA-RFP, and B6.TCRA-GFP/RFP mice. Representative samples shown. (B) Quantification of TCRA-GFP$^+$, TCRA-RFP$^+$, and dual TCRA-GFP$^+$RFP$^+$ CD4$^+$ and CD8$^+$ T cells. Dots represent 19 individual mice from six independent experiments, mean ± SEM. (C) Confocal microscopy of T cells from B6.TCRA-GFP/RFP mice. Images shown for GFP and RFP channels and merged channels for a representative single field at 600× magnification. (D) Quantification of TCRA-GFP$^+$, TCRARFP$^+$, and TCRA-GFP$^+$RFP$^+$ T cell-enriched splenocytes identified by manual counting of 10 confocal microscopy image fields containing 300 to 1,000 cells per image field per sample. Dots represent eight individual mice from three independent experiments, mean ± SEM.

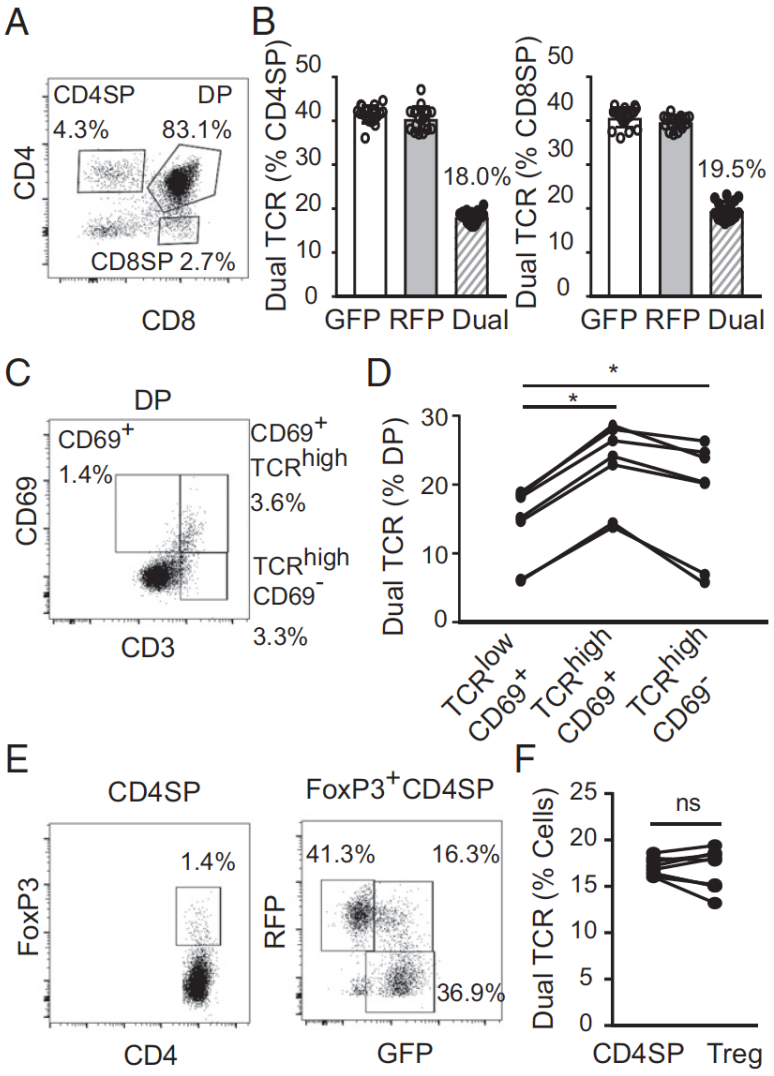

Figure 1.2: Dual TCRα expression arises during positive selection of thymocytes. (A) CD4$^+$CD8$^+$ DP, CD4$^+$CD8$^-$ single-positive (CD4SP), and CD4$^-$CD8$^+$ single-positive (CD8SP) thymocytes were identified from B6.TCRA-GFP/RFP mice by flow cytometry. Representative sample shown. (B) Quantification of dual TCRα GFP$^+$RFP$^+$ CD4SP and CD8SP. Dots represent 16 individual mice from six independent experiments, mean ± SEM. (C) TCR^{low}CD69$^+$, TCR^{high}CD69$^+$, and TCR^{high}CD69^{low} maturation stages of DP thymocytes were identified by flow cytometry. Representative sample shown. (D) Quantification of dual TCRα GFP$^+$RFP$^+$ DP thymocytes related to developmental stage. Linked dots represent seven individual mice from three independent experiments. Data compared by Wilcoxon matched-pairs rank-sign test. (E) Thymic CD4$^+$FoxP3$^+$ Tregs were identified among CD4SP by flow cytometry. Representative sample shown. (F) Quantification of dual TCRα GFP$^+$RFP$^+$ CD4SP Tregs. Dots represent nine individual mice from three independent experiments, mean ± SEM. Data compared by Wilcoxon matched-pairs rank-sign test. *P <0.05; ns = not statistically significant.

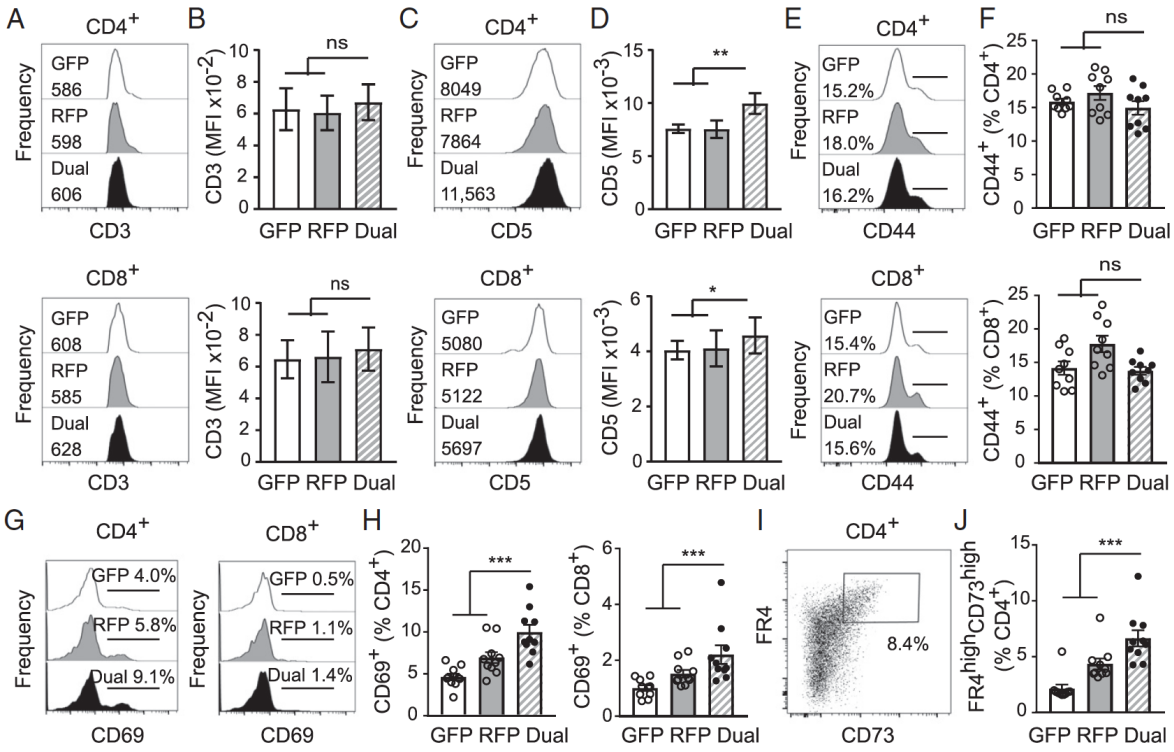

Figure 1.3: Dual TCR cells have equivalent total TCR but increased CD5. CD4+ and CD8+ splenic T cells from TCRA-GFP+, TCRA-RFP+, and TCRA-GFP+RFP+ mice were examined by flow cytometry. (A) Comparison of CD3 expression by GFP+, RFP+, and GFP+RFP+ cells with MFI from single representative sample shown. (B) Aggregate data for CD3 expression of nine mice from three independent experiments, mean ± SD. Data compared using paired Student's t test. (C) Comparison of CD5 expression on GFP+, RFP+, and GFP+RFP+ cells with MFI from single representative sample shown. (D) Aggregate data for CD5 expression of nine mice from three independent experiments, mean ± SD. Data compared using paired Student's t test. (E) Comparison of CD44 expression by GFP+, RFP+, and GFP+RFP+ cells with percentages of CD44+ cells from single representative sample shown. (F) Aggregate frequencies of CD44+ cells. Dots represent nine individual mice from three independent experiments, mean ± SD. Data compared using Wilcoxon matched-pairs rank-sign test. (G) Comparison of CD69 expression by GFP+, RFP+, and GFP+RFP+ cells with percentages of CD69+ cells from single representative sample shown. (H) Aggregate frequencies of CD69+ cells. Dots represent nine individual mice from three independent experiments, mean ± SEM. Data compared using Wilcoxon matched-pairs rank-sign test. (I) Identification of $CD73^{high}FR4^{high}$ CD4+ T cells as a marker for anergic phenotype. Representative sample shown, with single TCRα GFP+ or RFP+ cells in gray and dual TCRα GFP+RFP+ cells in black. (J) Dots represent nine individual mice from three independent experiments, mean ± SEM. Data compared using Wilcoxon matched-pairs rank-sign test. *P < 0.05, **P < 0.01, ***P < 0.001; ns = not statistically significant.

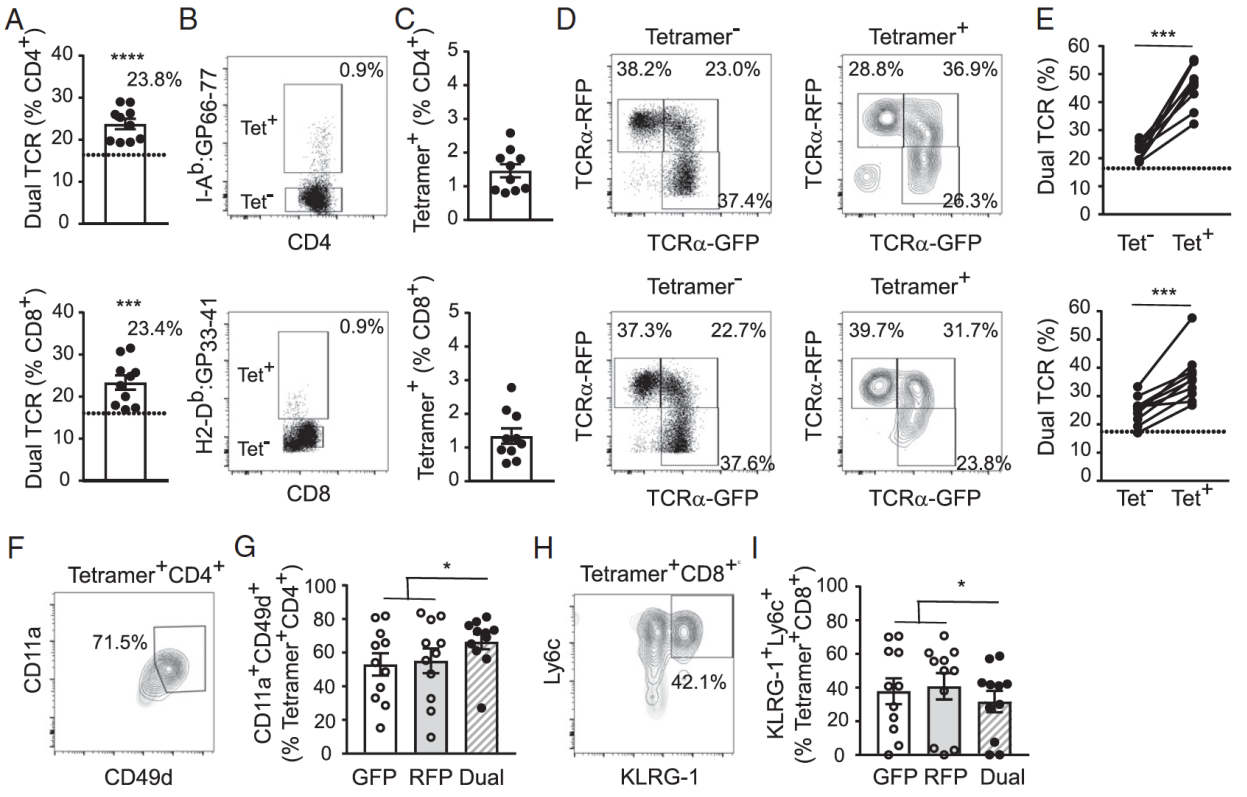

Figure 1.4: Dual TCR cells are selectively increased in response to LCMV infection. All data represent 11 individual mice from three independent experiments. (A) B6.TCRA-GFP/RFP mice were infected i.v. with 2 Å~ 105 plaque-forming units LCMV Armstrong and dual TCRα cells were measured by flow cytometry of splenocytes and lymph node cells 8 d after infection. Dots represent individual mice with mean ± SEM. Data compared to immunologically naive mice (dotted line) nonparametrically by Mann–Whitney U test. (B) LCMV-specific T cells were identified by flow cytometry for binding I-A^b:$GP_{66–77}$ and H2-D^b:$GP_{33–41}$ tetramers. Representative samples shown. (C) Quantification of LCMV-tetramer–binding cells in LCMV infected mice. Dots represent individual mice with mean ± SEM. (D) Flow cytometry to identify dual TCRα $CD4^+$ and $CD8^+$ T cells from tetramer $(Tet)^+$ and $Tet^–$ cells. Representative sample shown. (E) Comparison of dual TCRα T cells among $Tet^–$ and Tet^+ $CD4^+$ and $CD8^+$ T cells from LCMV-infected mice. Linked dots represent $Tet^–$ and Tet^+ cells from individual mice. Dotted line represents mean dual TCRα cell frequencies from naive mice. Comparison of Tet^+ and $Tet^–$ dual TCRα frequencies within individual samples performed nonparametrically using Wilcoxon matched-pairs rank-sign test. (F) Activated $CD4^+$ T cells present 8 d after LCMV infection were identified from Tet^+ cells by flow cytometry for expression of CD11a and CD49d. Gray lines are total $CD4^+$ cells and dark lines are gated on Tet^+ cells. Representative sample shown. (G) Quantification of $CD4^+Tet^+CD11a^+CD49d^+$ activated T cells. Dots represent individual mice with mean ± SEM. Data compared within individual samples nonparametrically using Wilcoxon matched-pairs rank-sign test. (H) Effector $CD8^+$ T cells present 8 d after LCMV infection were identified from Tet^+ cells by flow cytometry for expression of Ly6c and KLRG-1. Gray lines are total $CD8^+$ cells and dark lines are gated on Tet^+ cells. Representative sample shown. (I) Quantification of $CD8^+Tet^+KLRG\text{-}1^+Ly6c^+$ effector T cells. Dots represent individual mice with mean ± SEM. Data compared within individual samples nonparametrically using Wilcoxon matched-pairs rank-sign test. *P < 0.05, ***P < 0.005, ****P < 0.001.

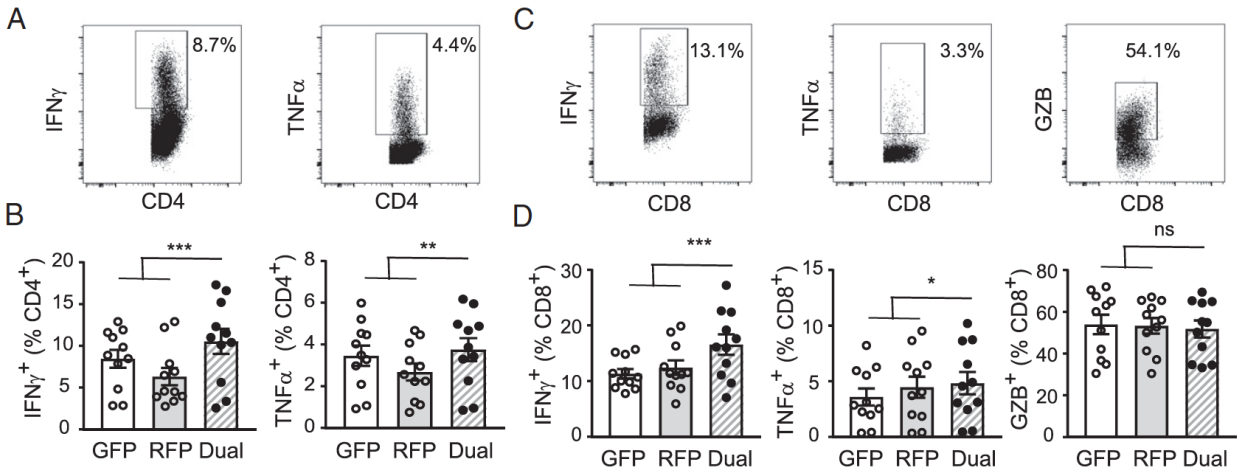

Figure 1.5: Dual TCR cells have increased functional response to LCMV. All data represent 11 individual mice from three independent experiments. (A) Cytokine production of $CD4^+$ T cells from LCMV-infected mice was assessed 8d post-infection by intracellular flow cytometry after ex vivo stimulation of splenocytes with LCMV $GP_{66–77}$ peptide. Representative sample shown. (B) Quantification of IFNγ- and TNFα-producing cells for TCRA-GFP^+, TCRA-RFP^+, and TCRA-GFP^+RFP^+ cells. Dots represent individual mice with mean ± SEM. Data compared within individual samples nonparametrically using Wilcoxon matched-pairs rank-sign test. (C) Cytokine production of $CD8^+$ T cells from LCMV-infected mice was assessed by intracellular flow cytometry after ex vivo stimulation of splenocytes with LCMV $GP_{33–41}$ peptide. Representative sample shown. (D) Quantification of IFNγ-, TNFα-, and granzyme B (GZB)-producing cells for TCRA-GFP^+, TCRA-RFP^+, and TCRA-GFP^+RFP^+ cells. Dots represent individual mice with mean ± SEM. Data compared within individual samples nonparametrically using Wilcoxon matched-pairs rank-sign test. *$P < 0.05$, **$P < 0.01$, ***$P < 0.005$; ns = not statistically significant.

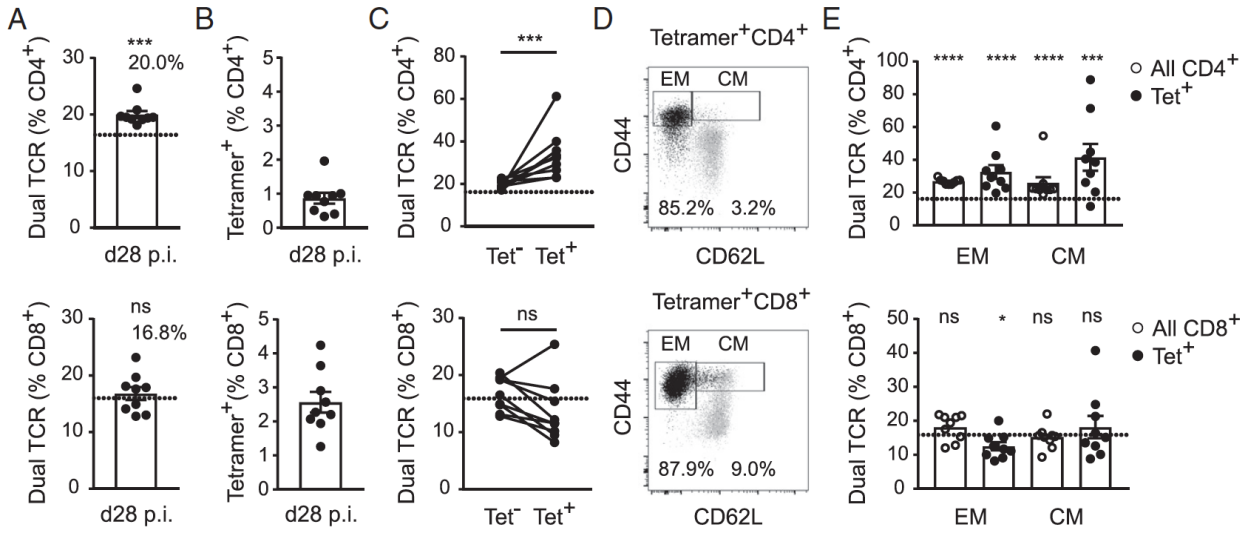

Figure 1.6: Dual TCR expression affects persistence of LCMV-specific cells after infection. All data represent nine individual mice from three independent experiments. (A) Dual TCR cell response assessed by flow cytometry of splenocytes and lymph node cells 28 d after infection of B6.TCRA-GFP/RFP mice with 2 x 10^5 plaque-forming units LCMV Armstrong. Dots represent individual mice with mean ± SEM. Data compared to immunologically naive mice (dotted line) nonparametrically by Mann–Whitney U test. (B) Quantification of LCMV I-A^b:$GP_{66–77}$ and H2-D^b:$GP_{33–41}$ tetramer (Tet)-binding cells. Dots represent individual mice with mean ± SEM. (C) Comparison of dual TCRα T cells among $Tet^–$ and Tet^+ $CD4^+$ and $CD8^+$ T cells. Linked dots represent $Tet^–$ and Tet^+ cells from individual mice. Dotted line represents mean dual TCRα cell frequencies from naive mice. Comparison of Tet^+ and $Tet^–$ dual TCRα frequencies within individual samples performed nonparametrically using Wilcoxon matched-pairs rank-sign test. (D) LCMV-specific EM and CM cells identified from $CD4^+$ and $CD8^+Tet^+$ cells by flow cytometry. Gray lines are total $CD4^+$ or $CD8^+$ cells and dark lines are gated on Tet^+ cells. Representative sample shown. (E) Quantification of dual TCR cells in EM and CM subsets for total (open circles) or Tet^+ (closed circles) T cells. Dots represent individual mice with mean ± SEM. Groups compared to dual TCR cell frequencies from naive mice (dotted line) nonparametrically using Mann–Whitney U test. *P < 0.05, ***P < 0.005, ****P < 0.001; ns = not statistically significant.

1.5: Materials and Methods

Mice.

C57BL/6 mice were originally purchased from Charles River Laboratories. Mice with fluorescent reporters linked to the TCRCα protein were generated by CRISPR-Cas9–mediated recombination (Appendix, Supplemental Methods). B6.TCRA-GFP/RFP mice were generated by interbreeding B6.TCRA-GFP and B6.TCRA-RFP mice. Both male and female mice were used in this study. Age- and sex-matched mice of 6 to 8 wk of age were used for all studies. All experimental mice were bred and housed in specific-pathogen-free conditions in a 12-h light/dark cycle with ad libitum access to chow and water at the University of California San Diego (UCSD). All breeding and experiments were performed according to UCSD Institutional Animal Care and Use Committee–approved protocols and under the supervision of the UCSD
Animal Care Program.

Flow Cytometry.

Cells were incubated with Zombie yellow (Biolegend) viability dye prior to labeling with antibodies as indicated. Samples were run in batches containing control and experimental samples with color and fluorescence-minus-one (FMO) controls. Where applicable, cells from single-transgenic B6.TCRA-GFP and B6.TCRA-RFP mice were used as FMO controls for gating single and dual TCR cells from B6.TCRA-GFP/RFP mice. Flow cytometry analyses were performed using FACSCanto or LSR II instruments (BD Biosciences) with FACSDiva software. Data were analyzed using FlowJo v10 software.

Confocal Microscopy.

T cells were isolated by negative paramagnetic bead selection and directly plated in Nunc Lab-Tek chambered coverglass. Confocal imaging was performed using Olympus FV1000 confocal microscope (Olympus America) at 600× magnification. GFP (excitation: 488 nm, filter: 500 nm to 530 nm) and RFP (excitation: 543 nm, filter: 555 nm to 655 nm) channels were scanned in sequential-frame manner to exclude overlapping. Ten independent fields containing 300 to 1,000 cells per field were taken for each mouse. Images were processed and analyzed using ImageJ Fiji software (60).

In Vitro T Cell Proliferation.

Peripheral T cells were isolated from spleens using negative paramagnetic bead selection and labeled with Tag-it Violet tracking dye (Biolegend). T cells were cultured in vitro in RPMI 1640 supplemented with 10% fetal bovine serum, 2 mM Glutamax, and 0.5 μg/mL gentamicin (Thermo Scientific) in 96-well flat-bottomed plates. Cells were stimulated with graded doses of plate-bound anti-CD3 (2C.11)/anti-CD28 (37.51) mAbs and SEB as indicated and cultured for 5 d at 37 °C 6% CO2. Proliferation was measured by flow cytometry and proliferation index was calculated as the total number of cell divisions divided by calculated number of divided cells (61).

LCMV Infection.

LCMV Armstrong strain was propagated on BHK cells and quantified by plaque assay performed on Vero cells. Vero cell monolayers were infected with 500 mL serially diluted viral stock and incubated for 60 min at 37 °C, 5% CO2 with gentle shaking. Agarose overlay was applied to infected cells and incubated at 37 °C, 5% CO2 for 6 d, after which cells were fixed with

formaldehyde and stained with crystal violet for 5 min at room temperature for plaque enumeration. Mice were infected by intravenous (i.v.) injection of 2×10^5 plaque forming units of murine LCMV Armstrong strain. Mice were killed 8 or 28 d after infection, and spleen and lymph nodes were collected for analysis by flow cytometry. LCMV-specific T cells were identified by flow cytometry for pMHC-tetramer labeling for $GP_{33–41}$ (KAVYNFATM):H2-D^b and $GP_{66–77}$ (DIYKGVYQFKSV):I-A^b (NIH Tetramer Core). Tetramer labeling was performed by 30′ incubation of cells (labeled with viability dye) with 1:100 dilution of tetramer at room temperature. Cells were subsequently labeled with mAb, including antibodies against B220, F4/80, CD11c, and CD11b to exclude nonspecific tetramer binding.

Statistical Analyses.

Data were analyzed using Prism 6 software. Data from individual mice were analyzed nonparametrically using Mann–Whitney U test. Analysis of intersample groups was performed nonparametrically using Wilcoxon matched-pairs rank-sign test. Proliferation indices were compared between groups across graded stimulation using two-way ANOVA with Tukey's multiple comparison test. Mean fluorescence intensity (MFI) values from intersample groups were compared using paired Student's *t* test for mean values. Two-tailed *P* values ≤ 0.05 were considered statistically significant for all analyses.

Data Availability.

All study data are included in the paper and Appendix.

Acknowledgements.

We thank Wei Feng (UCSD) and Ella Kothari and Jun Zhao (UCSD Transgenic Mouse Shared Resource) for assistance in CRISPRCas9 blastocyst microinjection, Cody Fine and Jesus Olivera (UCSD Human Embryonic Stem Cell Core Facility) for assistance in flow cytometry cell sorting, Jack Bui (UCSD) for use of flow cytometry equipment, and Paul M. Allen (Washington University) and Wan-Lin Lo (UCSF) for helpful discussion and comments. This research was supported by NIH R56 AI137150 (G.P.M.), R01 AI151293 (G.P.M.), R01 AI081923 (E.I.Z.), and an American Society of Hematology Bridge grant (G.P.M.). The UCSD Human Embryonic Stem Cell Core Facility is supported by California Institute for Regenerative Medicine Major Facilities Grant FA1-00607.

Chapter 1, in full, is the reprint of the material as it appears in Proceedings of the National Academy of Sciences 2020. Yang, Letitia; Jama, Burhan; Wang, Huawei; Labarta-Bajo, Lara; Zúñiga, Elina I.; Morris, Gerald P, PNAS, 2020. The thesis author was a co-author of this paper.

CHAPTER 2: BULK TCR SEQUENCING UNCOVERS UNIQUE DUAL TCRα REPERTOIRE ENRICHED WITH POTENTIALLY AUTOREACTIVE CLONOTYPES

2.1: Introduction

Dual TCRα expression has been hypothesized to allow thymocytes expressing autoreactive TCRs to avoid negative selection. This is thought to occur due to the decreased overall surface expression of both TCRs, which leads to diminished signal transduction from the autoreactive TCR, thereby allowing the thymocyte to survive (2, 10-12). The experiments shown in Chapter 1 demonstrated that dual TCRα cells have increased CD5 expression, a marker for TCR signal strength in response to self-peptide-self-MHC (33), suggesting that they may have increased reactivity for self-antigens compared to single TCRα cells (Figure 1.3 C, D). The same phenomenon has been observed using anti-TCRVα mAb co-labeling for T cells in an acute lymphopenia-induced proliferation model (35). Furthermore, previous TCRα repertoire studies using TCRα$^{+/-}$ mice have shown that dual TCRα cells contain unique TCR clonotypes that would otherwise have been negatively selected in single TCRα cells during thymic development (8). These results led us to postulate that the increased ability for dual TCRα cells to escape negative selection changes the stringency of thymic selection, which may result in dual TCRα cells having a unique repertoire containing TCRs reactive to self-antigens.

TCR sequencing technology relies on the identity of V and J segments along with the complementarity-determining region 3 (CDR3) to quantify T cell clonalities. Somatic recombination of V(D)J genes, imprecise recombination junctions, and combinatorial pairing give rise to an immensely diverse repertoire of approximately 2 x 10^6 different TCRαβ clonotypes in mice and 2.5 x 10^7 in humans (68, 69). CDR3 also encodes a portion of the ligand-binding site on

the TCR, making it essential for the interaction between the TCR and peptide-MHC complex (69, 70). Since it is highly unlikely that two T cells derived from different clonal lineages will share the same V(D)J segments and CDR3 sequences, these properties are widely used for TCR repertoire analysis.

The two complementary sequencing approaches currently available for studying the TCR repertoire are bulk and single-cell sequencing. Single-cell methods reveal information about paired TCR sequences that can be used to study T cell ancestry and specificity. However, single-cell sequencing covers a limited number of cells and is therefore less suitable for studying TCR diversity in larger cohorts (21, 71). On the other hand, bulk sequencing of combined immune cell populations does not provide information on TCR pairs but is more advantageous and cost-effective for our purpose of comparing distinct repertoires on a larger scale (21, 71).

Using this approach, I hypothesized that expression of a secondary TCR may alter the efficiency of negative selection, leading to the generation of a distinct peripheral dual TCRα repertoire containing TCRs that may be autoreactive. To compare the single and dual TCR clones, R correlation coefficient, Jaccard index, and Morisita-Horn index were used to assess correlation, similarity and overlap between clonotypic frequencies of the repertoires, respectively. Our bulk sequencing data showed that the peripheral dual TCRα repertoire contained unique TCR clonotypes, and that the $CD4^+$ dual TCRα repertoire was enriched with clones associated with autoantigen and pathogen epitopes. These findings demonstrate the potential for dual TCR cells to harbor reactivities against distinct antigens and highlight a potentially significant role of dual TCRα cells in the development of autoimmune diseases.

2.2: Results

Insertion of Fluorescent Reporter Genes Does Not Create Oligoclonal Bias in TCR Repertoire of B6.TCRA-GFP and B6.TCRA-RFP Mice.

To effectively compare single and dual TCRα clonotypes, bulk sequencing of TCRα from wildtype B6, B6.TCRA-GFP and B6.TCRA-RFP mice was first performed to ensure that the fluorescent reporter genes do not alter the repertoire. Splenocytes from B6, B6.TCRA-GFP and B6.TCRA-RFP mice were extracted, and T cells were isolated by fluorescence activated cell sorting (FACS). RNA was extracted from sorted cells and TCR profiling was performed using a 5'-RACE-based approach, which relies on the template-switching effect and allows for all TCR variants in the sample to be captured while minimizing bias caused by the target sequences (21, 72). Indexed libraries were then sequenced by Illumina sequencing technology (HiSeq 4000).

To investigate whether insertion of the fluorescent reporter genes would influence development of the T cell repertoire in our reporter mouse models, we analyzed the repertoires of wildtype B6 mice and our reporter mice. Comparative analysis of their TCRα repertoires was performed using the normalized Shannon-Weiner index and Inverse Simpson index to measure clonal diversity while accounting for evenness of the samples (73). Evenness was crucial for ensuring that the target repertoires were not skewed, as uneven distribution of individual TCR clones would create bias in the comparative results. If a significant difference in evenness was observed, it would be more effective to focus on the most abundant clones rather than assessing them as a population.

According to the Shannon-Weiner and Inverse Simpson indices, there was no significant difference in clonal diversity between the B6 control and the reporter mice for both $CD4^+$ and

$CD8^+$ T cell populations (Figure 2.1 A, 2.1 B). Consistent with this finding, no difference in TRAV and TRAJ usage was found between the control and reporter mice (Figure 2.1 C, 2.1 D). To quantify the number of shared amino acid sequences among the single and dual TCRα repertoire, the amino acid length distribution and characteristics of CDR3α was measured. The CDR3α chains for B6 control and reporter mice ranged between 8 and 18 amino acids in length for both $CD4^+$ and $CD8^+$ T cell populations (Figure 2.1 E). For all three models, the CDR3α chains were high in polarity and strength, but relatively low in charge and hydropathy (Figure 2.1 F). These results indicate that the TCRα repertoires from B6 control and reporter mice have similar germline features and CDR3 biophysical characteristics.

To compare clonotypic frequencies between the mouse models, the R coefficient was used to measure correlation, Jaccard index was used to measure similarity, and Morisita-horn index was used to measure overlap between the TCR clones. Analysis of TCR clonotypic frequencies revealed strong correlation ($R > 0.79$), high similarity (Jaccard > 0.22), and high overlap (Morisita-Horn > 0.76) between the control and the reporter mice, which indicate that the repertoires of the reporter mice are diverse, conserved, and functionally similar to the B6 control (Figure 2.1 G). Altogether, these results demonstrate that insertion of the GFP or RFP reporter genes in B6.TCRA-GFP and B6.TCRA-RFP mice does not skew the T cell repertoire, and that our reporter system can be used as a model for unbiased repertoire analysis.

Peripheral Dual TCRα Repertoire Contains Clonotypes Distinct from Single TCRα Repertoire.

Next, the dual TCRα repertoire was examined using the B6.TCRA-GFP/RFP model. Splenocytes were extracted from B6.TCRA-GFP/RFP mice and $CD4^+$ or $CD8^+$ T cells expressing

single or dual TCRα were isolated by FACS (Figure 2.4 A). RNA was purified and bulk TCR sequencing was performed. There was no significant difference in clonal diversity between the single and dual TCRα repertoire according to the Shannon-Weiner and Inverse Simpson indices, which indicates that the compared repertoires were equally diverse without oligoclonal bias (Figure 2.2 A, 2.2 B). No significant difference in TRAV/J and TRBV/J usage was observed between $CD4^+$ or $CD8^+$ single and dual TCRα cells (Figure 2.3 A, 2.3 B). Both CDR3α and CDR3β lengths ranged from 8 to 18 amino acids for single and dual TCRα cells (Figure 2.3 C, 2.3 D). Similar to data from the control mice, CDR3α chains from single and dual TCRα clones were high in polarity and strength, but relatively low in charge and hydropathy (Figure 2.1 F, 2.3 E). CDR3β chains from single and dual TCRα clones were also high in polarity and strength, but much lower in charge and hydropathy compared to CDR3α (Figure 2.3 F). These findings indicate that the single and dual TCRα repertoires have broadly similar features, including V/J gene segment usage, CDR3 lengths and amino acid characteristics.

Analysis of the TCRα clonotypic frequencies revealed lower correlation and overlap between the single and dual TCRα clonotypes compared to frequencies from the control and reporter mice (Figure 2.4 B, 2.4 C, 2.1 E). The Jaccard index revealed a more dramatic difference. Compared to data from the control and reporter mice, the Jaccard values between the dual and single TCRα clonotypes were lower by 10-fold (Jaccard < 0.04), indicating that the two populations were non-overlapping (Figure 2.1 E, 2.4 B, 2.4 C). These results demonstrate that the peripheral dual TCRα repertoire contains unique CDR3 sequences that gave rise to clones distinct from the single TCRα repertoire.

CD4$^+$ Dual TCRα Repertoire is Enriched for Clonotypes Associated with Autoantigen and Pathogen Epitopes.

Based on our finding that the dual TCRα repertoire contained unique clonotypes, I hypothesized that a high frequency of these clones may be autoreactive. To test this hypothesis, the frequency of single versus dual TCRα clonotypes specific for autoantigen, tumor antigen, and pathogen epitopes was examined using a database of TCR sequences associated with defined ligands (McPAS-TCR) (74). Tumor antigen epitopes were included in this study because dual TCR cells have been thought to mediate anti-tumor immunity through the recognition of modified self-antigens (75). Additionally, a broad expansion of dual TCR cells was observed in response to acute LCMV infection in Chapter 1, so TCR specificity for pathogen epitope was also assessed to investigate whether dual TCR reactivity to non-specific foreign antigens contributed to the increased expansion.

Analysis of CD4$^+$ single and dual TCRα clonotypes revealed a 0.3% increase in the proportion of CD4$^+$ dual TCRα clonotypes specific for pathogen epitopes compared to single TCRα clones (Figure 2.5 A top panel). A significant increase of approximately 1.3% in the frequency of dual TCRα clones specific for autoantigen epitopes was also observed (Figure 2.5 A top panel). Notably, the relative increase in frequency of autoantigen-specific clones is nearly five times the frequency of pathogen-specific clones, which highlights the autoreactive potential of the dual TCRα repertoire. No significant difference in specificity was found for CD8$^+$ TCRα, CD4$^+$ TCRβ, or CD8$^+$ TCRβ clonotypes (Figure 2.5 A bottom panel, 2.5 B), but this was likely due to the fact that only a small percentage of clonotypes was functionally annotated. Although the current database of annotated TCR sequences is limited, these findings are consistent with previous

speculation that the dual TCRα repertoire contains a higher frequency of autoreactive TCR clones compared to the single TCRα repertoire.

2.3: Discussion

Dual TCR cells expressing two distinct TCR clonotypes comprise a small population of the normal T cell repertoire. The simultaneous rearrangement of two TCRα loci has been shown to enhance positive selection and mask the secondary TCR from negative selection, altering the stringency of thymic selection and leading to the generation of distinct repertoires (2, 8, 10-12). Chapter 2 presents evidence that the peripheral $CD4^+$ and $CD8^+$ dual TCRα repertoire contained unique clonotypes with CDR3 sequences that are distinct from the single TCRα repertoire (Figure 2.4). The demonstration that dual TCR cells contain a unique TCR repertoire not found in conventional single TCR cells suggests that they may be specific to distinct ligands.

Our results showed a higher frequency of $CD4^+$ dual TCRα cells associated with autoantigen epitopes compared to single TCRα cells, indicating that dual TCRα expression is linked to increased reactivity against self-ligands. This finding was consistent with previous work showing increased ability for dual TCRα cells to respond to self-peptide-MHC tetramers (8). Additionally, this increased response to self-ligands may explain our observation showing a selective persistence of dual TCRα memory cells after the resolution of infection, as cells with high affinity for self-ligands have been demonstrated to favor memory formation (76). The same phenomenon was not observed within the $CD8^+$ dual TCRα subset, though this was likely due to the limited number of functionally annotated $CD8^+$ TCR sequences currently available. As a result, further studies utilizing larger sample sizes will be necessary to determine the epitope specificities of the $CD8^+$ dual TCR repertoire. Altogether, these findings suggest that dual TCRα cells may contribute to the development of autoimmune diseases.

Unlike single-cell sequencing, bulk sequencing methods do not provide information regarding TCR α/β pairs, which have been shown to influence MHC and antigen specificities of a

TCR repertoire (77, 78). Although our results showed that the $CD4^+$ dual TCRα repertoire encompassed a high frequency of TCRα capable of recognizing autoantigen epitopes, the peptide specificity of each TCR pair remains unclear because the α and β chains were assessed separately in this study (Figure 2.5). Therefore, optimizing single-cell approaches for dual TCRα profiling will be useful for defining TCR α/β pairs and performing studies such as single-cell sequencing of barcoded peptide-MHC multimers. The use of barcoded multimers would allow for single-cell isolation of autoantigen- or pathogen-specific TCR sequences, which would be crucial for elucidating the ligand biochemistry of dual TCR cells (79). Further studies on the mechanisms underlying how dual TCRα cells influence self-reactivity, as well as their potential for autoreactivity are necessary to define their role in autoimmunity. Understanding their peptide specificities will further provide valuable insight towards the mechanisms underlying their increased affinity for self-ligands and their potential applications for treating autoimmune diseases.

2.4: Figures

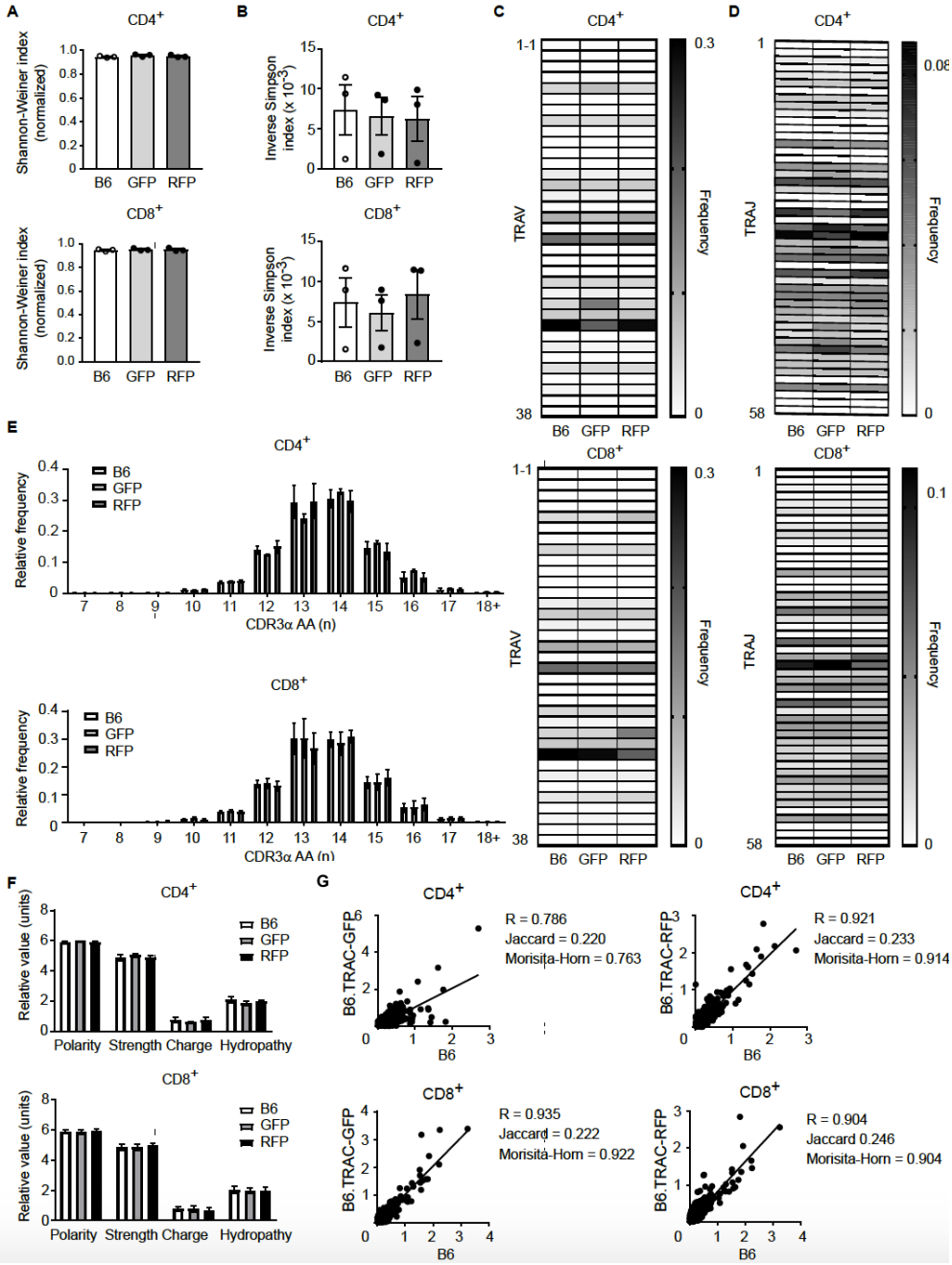

Figure 2.1: Insertion of fluorescent reporter gene does not alter the TCR repertoire in B6.TCRA-GFP and B6.TCRA-RFP mice. *(A, B)* Quantification of TCR diversity in adult wildtype B6, B6.TCRA-GFP and B6.TCRA-RFP mice. Dots represent 9 individual mice. *(C, D)* Heat-map showing frequency of TRAV and TRAJ usage from CD4+ and CD8+ T cells in adult wildtype B6, B6.TCRA-GFP and B6.TCRA-RFP mice. *(E, F)* Length distribution and amino acid properties of CDR3 from CD4+ and CD8+ T cells in adult wildtype B6, B6.TCRA-GFP and B6.TCRA-RFP mice. *(G)* Comparison of clonal frequencies in adult B6 wildtype and B6.TCRA-GFP or B6.TCRA-RFP mice. R represents correlation coefficient, Jaccard index denotes similarity, and Morisita-Horn index shows overlap between two samples.

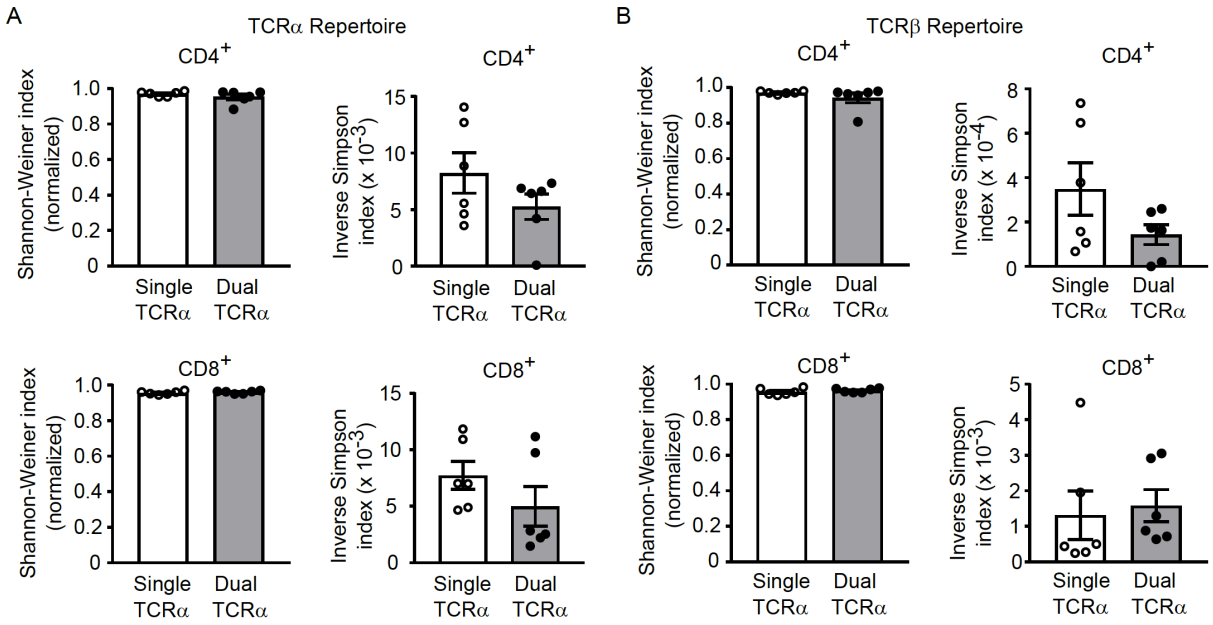

Figure 2.2: The single and dual TCR repertoires were equally diverse without oligoclonal bias. *(A, B)* Quantification of clonal diversity in the TCRα and TCRβ repertoire from adult B6.TCRA-GFP/RFP mice. Dots represent 6 individual mice. Normalized Shannon-Weiner index and Inverse Simpson index indicate species diversity, which encompasses richness and evenness of each sample.

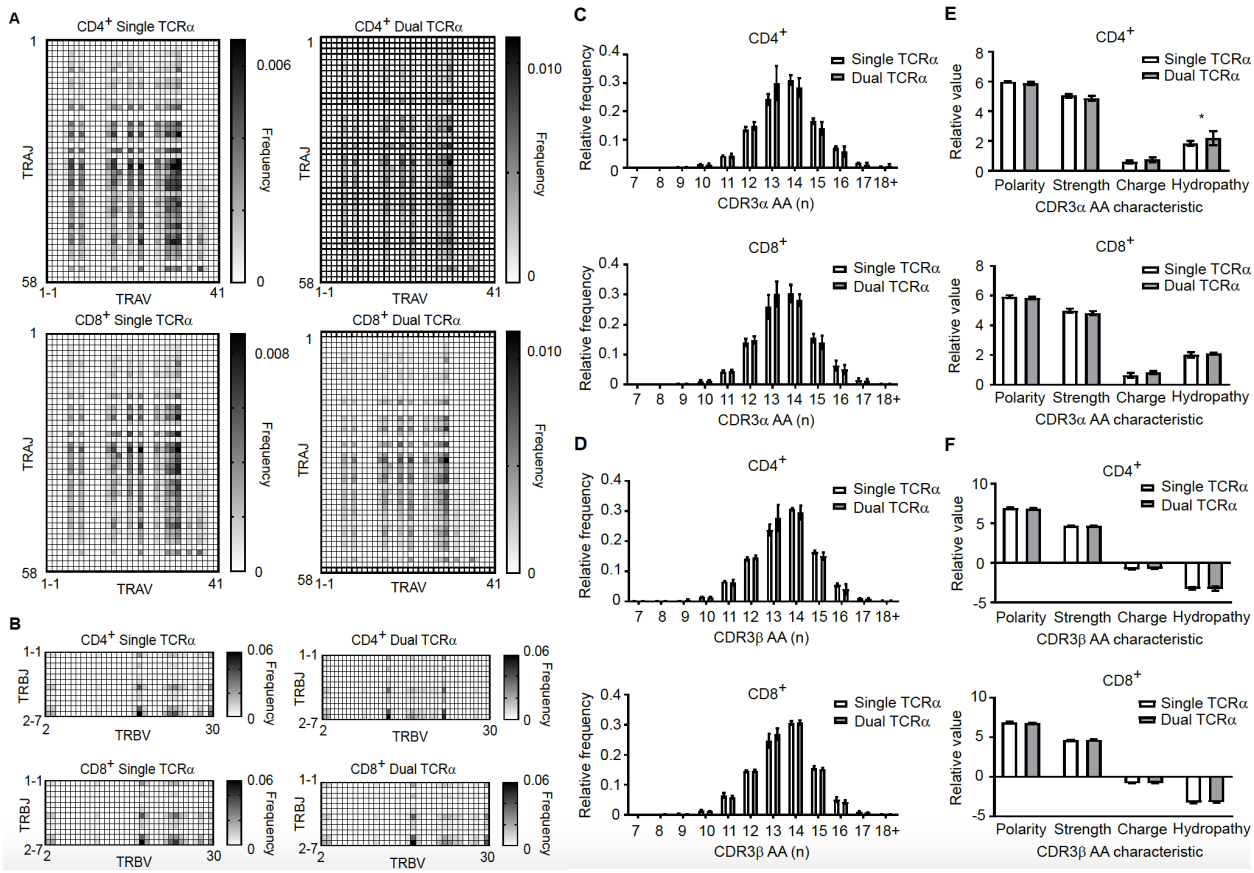

Figure 2.3: Single and dual TCRα repertoires have similar V/J segment usage and CDR3 characteristics. (*A, B*) Heatmap showing TRAV and TRBV usage, respectively, in CD4⁺ and CD8⁺ single and dual TCRα cells. (*C, D*) Length distribution of CDR3α and CDR3β amino acids, respectively, from CD4⁺ and CD8⁺ T cells in adult B6.TCRA-GFP/RFP mice. (*E, F*) CDR3α and CDR3β amino acid characteristics, respectively, from CD4⁺ and CD8⁺ T cells in adult B6.TCRA-GFP/RFP mice. $*P < 0.05$.

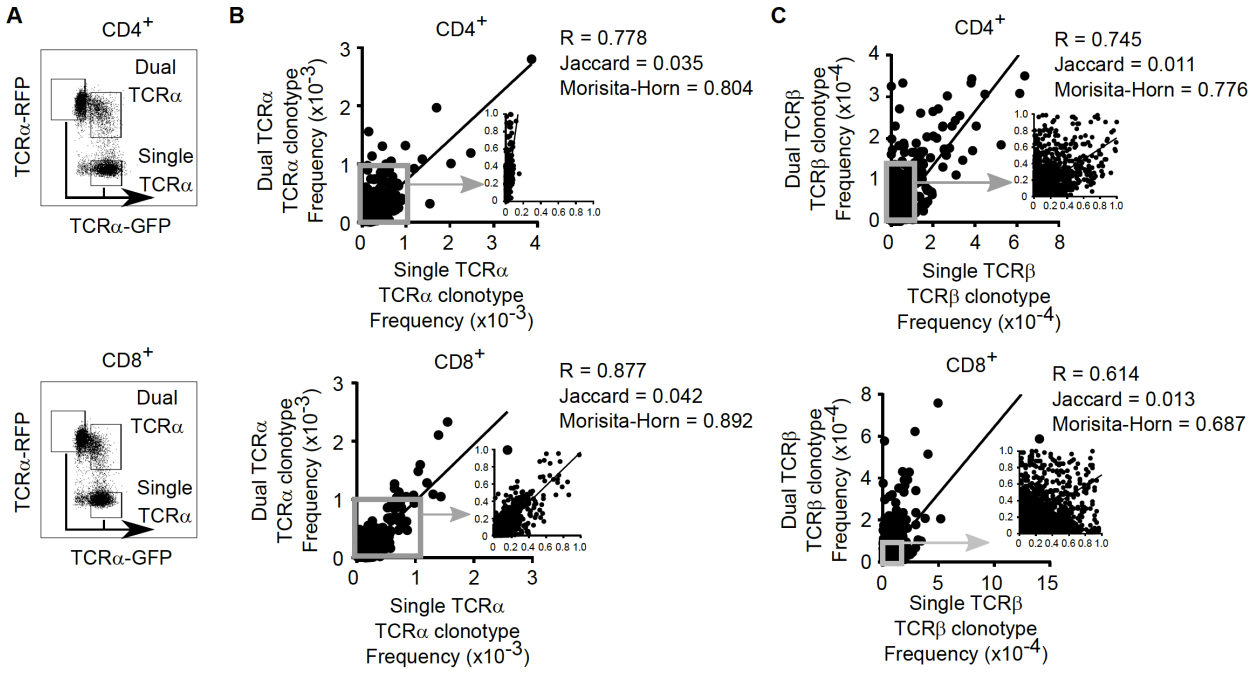

Figure 2.4: The peripheral dual TCRα repertoire contains clonotypes distinct from the single TCRα repertoire. *(A)* Representative flow cytometry plot to identify CD4+ and CD8+ single versus dual TCRα cells. *(B, C)* Comparison of frequencies between single and dual TCRα and TCRβ clonotypes from adult B6.TCRA-GFP/RFP mice. Smaller subplots (to the right) are enlargements of low frequency clones. R represents correlation coefficient; Jaccard index denotes similarity between two samples; Morisita-Horn index shows overlap between two samples.

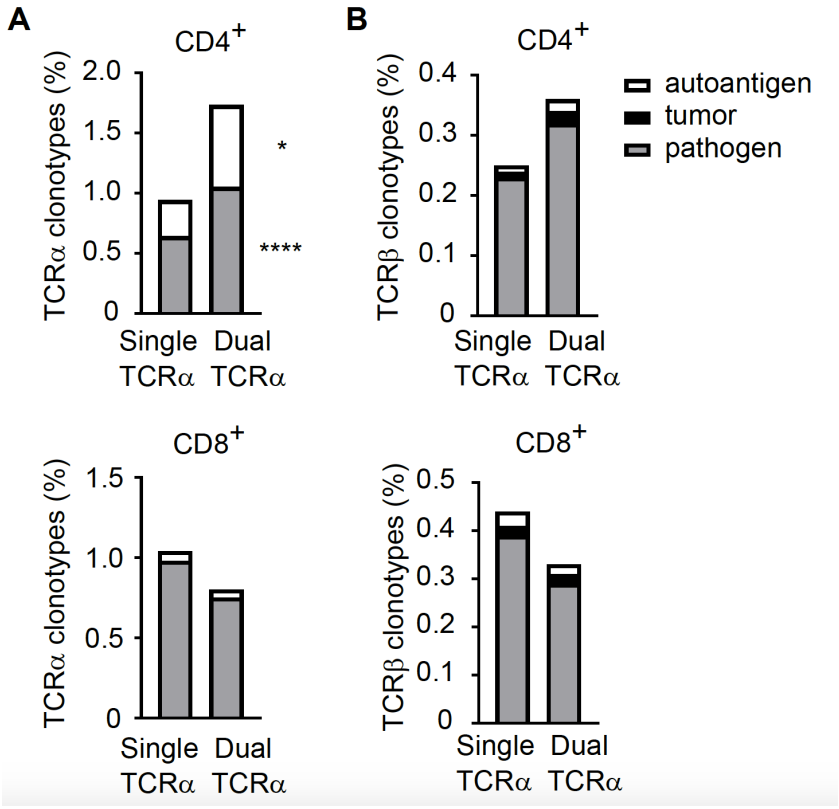

Figure 2.5: CD4+ dual TCRα repertoire may contain a high frequency of clonotypes associated with autoantigen and pathogen epitopes. (*A, B*) Frequency of single versus dual TCRα and TCRβ clonotypes, respectively, associated with autoantigen, tumor, or pathogen epitopes according to the McPAS-TCR database. * $P < 0.05$; **** $P < 0.00005$.

2.5: Materials and Methods

Fluorescence Activated Cell sorting.

T cells were isolated from mouse splenocytes using MojoSort CD3 T cell isolation kit (Biolegend #480031). T cells were labeled with Zombie Yellow (Biolegend #423104) viability dye prior to staining with anti-CD4 (Biolegend #100526) and anti-CD8 (Biolegend #100734) antibodies. Cell sorting was performed at the UCSD Human Embryonic Stem Cell Core Facility.

RNA Isolation and TCR Sequencing.

RNA was purified from sorted cells using the RNEasy mini prep kit (Qiagen #74104). TCR profiling was performed using SMARTer TCRα/β profiling kit (Takara #634403). Sample QC, library preparation, and sequencing on Illumina HiSeq 4000 was conducted by the UCSD IGM Genomics Center.

Sequencing Data Analysis.

Input sequence data was processed using MiXCR to build alignments, assemble clonotypes, and conduct frequency-based error correction. TCR repertoire was extracted and VDJtools was used to analyze repertoire diversity, overlaps, CDR3 properties, and V/J usage. Computational association of TCR sequences with their epitopes was performed using the McPAS-TCR database.

Statistical Analysis.

Student's *t* test was used to analyze statistical data. $P \leq 0.05$ was considered statistically significant for all analyses.

Acknowledgements.

Chapter 2, in part, is currently being prepared for submission for publication of the material. Yang, Letitia; Morris, Gerald P. The thesis author was the primary investigator and author of this material.

APPENDIX

Chapter 1 Supplemental Methods:

Generation of B6.TCRA-GFP and B6.TCRA-RFP mice.

Mice with fluorescent reporters linked to the TCRCα protein were generated by CRISPR/Cas9-mediated recombination of DNA encoding GSG-linked (GSGSGSGSGSGSGSQPVAT) eGFP (GFP) and tdTomato (RFP) at the 3' end of Exon 3 of TRAC (CRISPR guide site CTGAGGCTGTGGTCCAG/TTG(AGG) with "/" indicating insertion site and PAM site in parentheses) (Figure S1A). DNA encoding eGFP and tdTomato flouorophores was cloned into a pUCIDT plasmid containing ~1 kb of synthesized DNA sequence of the 3' end of TRAC Exon 3, with DNA sequence including the GSG linker at the insertion site, and ~1 kb of TRAC intronic sequence downstream of the insertion site. This generated a targeting plasmid containing an ~800 bp GSG-fluorophore insertion sequence flanked by ~1 kb homology arms. CRISPR mixture containing Cas9 protein (1000 nM), AltR-modified guideRNA (6 μM), and AltR-modified tracRNA (6 μM) was mixed in TE buffer pH 7.5 and incubated at room temperature for 5 min. Targeting plasmid (50 nM) was added and incubated for 5 min at room temperature. Mixture was then centrifuged 10,000 rpm for 1 min, and the supernatant was collected and put on ice until injection. CRISPR/Cas9/DNA mixture was injected into the cytoplasm of d 0.5 C57BL/6 (B6) embryos via micropipette. Injected embryos were implanted into pseudo pregnant dams. Pups were screened for presence of transgene by flow cytometry of peripheral blood lymphocytes, and insertion sequence was confirmed by forward sequencing of TRAC and reverse sequencing of the 3' end of the fluorophore reporter. Transgenic mice were bred to make produce transgenic homozygous B6.TCRA-GFP and B6.TCRA-RFP mice.

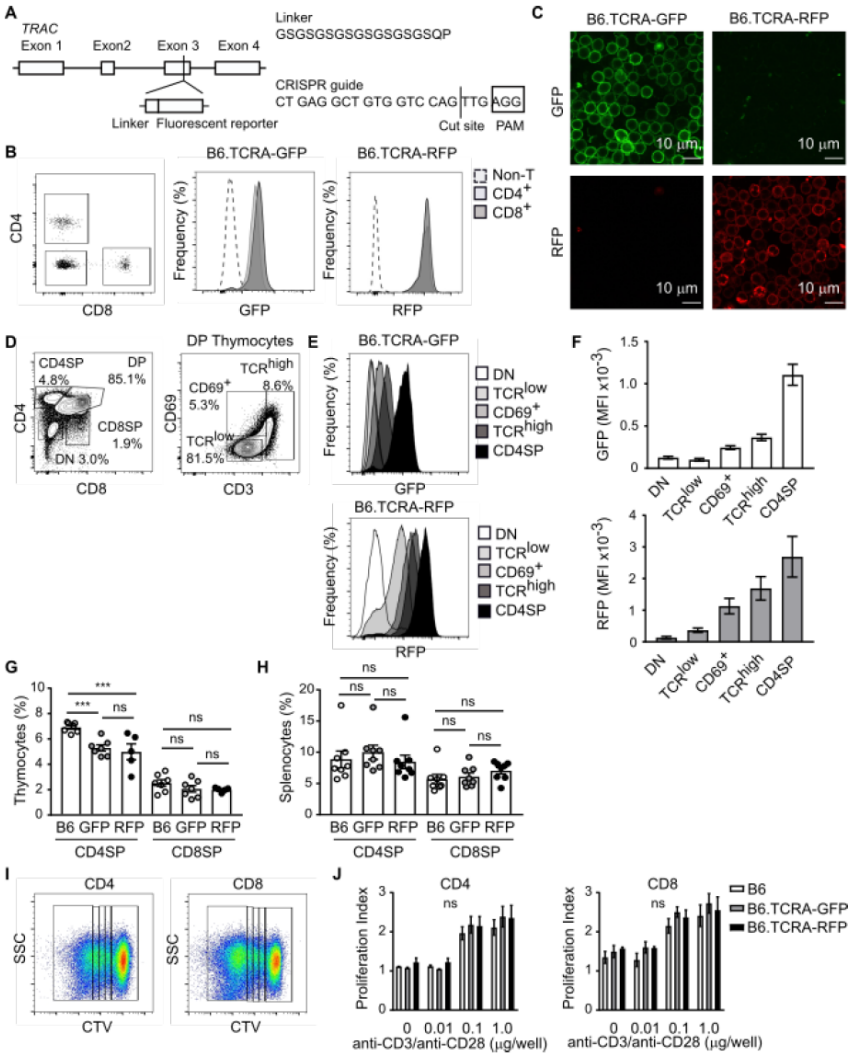

Figure S1.1: Generation of B6.TCRA-GFP and B6.TCRA-RFP mice. (*A*) Schematic for insertion of DNA encoding 18 amino acid GSG linker and eGFP or tdTomato reporters into the 3' end of Exon 3 of *TRAC* by CRISPR/Cas9-mediated gene editing. GSG linker amino acid sequence and CRISPR guide RNA sequences shown. (*B*) Representative flow cytometry of splenocytes from adult B6.TCRA-GFP and B6.TCRA-RFP mice. Expression of GFP and RFP were measured for $CD4^+$ T cells, $CD8^+$ T cells, and $CD4^-CD8^-$ non-T cells. (*C*) Confocal microscopy of T cells from B6.TCRA-GFP and B6.TCRA-RFP mice. Images shown for GFP and RFP channels for individual images at 600X magnification. Representative samples. (*D*) Examination of thymocyte development in TCRA-reporter mice by flow cytometry. Representative B6.TCRA-RFP sample shown. (*E*) Expression of GFP and RFP reporters evaluated by mean fluorescence intensity (MFI) among thymocyte developmental stages. Representative samples shown. (*F*) Quantification of GFP and RFP reporter expression from 5-6 individual mice in each group from 3 independent experiments. Data shown as mean ± SD. (*G*) Quantification of CD4SP and CD8SP thymocytes. Dots represent 5-7 individual mice per group from 3 independent experiments with mean ± SEM. Data compared non-parametrically using Mann-Whitney test. (*H*) Flow cytometry of splenocytes from B6, B6.TCRA-GFP, and B6.TCRA-RFP mice gating on $CD4^+$ and $CD8^+$ T cells. Dots represent 8 individual mice per group from 3 independent experiments with mean ± SEM. Data compared non-parametrically using Mann-Whitney test. (*I*) *In vitro* proliferation of $CD4^+$ and $CD8^+$ T cells measured 5 d after stimulation with plate-bound anti-CD3/anti-CD28 mAbs was measured by dilution of cell-tracker violet (CTV). Representative B6. TCRA-RFP sample shown. (*J*) Proliferation in response to graded doses of anti-CD3/CD28 stimulation was assessed as Proliferation Index. Data are shown mean ± SD of aggregate data from 3 independent experiments. Data analyzed using 2-way ANOVA with Tukey's multiple comparison test. ***$P < 0.005$, ns = not statistically significant.

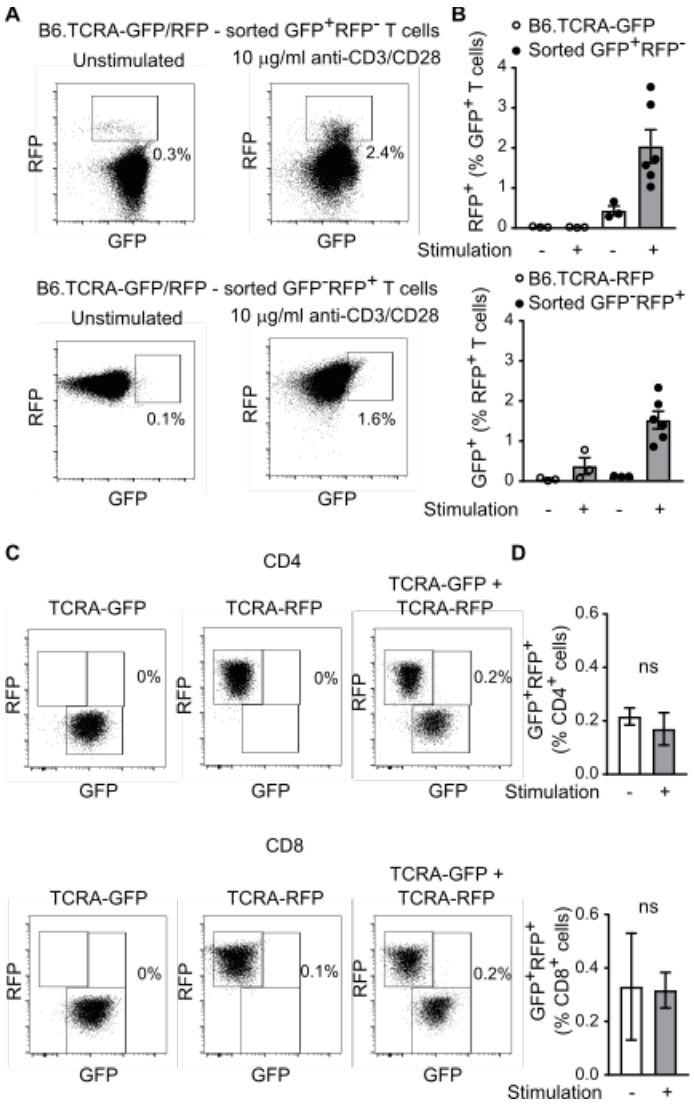

Figure S1.2: Dual TCRα expression is stable on B6.TCRA-GFP/RFP T cells. (*A*) GFP$^+$RFP$^-$ and GFP$^-$RFP$^+$ T cells were isolated by FACS from B6.TCRA-GFP/RFP mice and stimulated *in vitro* with 1.0 μg/ml anti-CD3/CD28 mAbs to observe potential changes in TCRα expression. Expression of GFP and RFP reporters was measured in live CD4$^+$ and CD8$^+$ T cells on d 5 by flow cytometry. Representative samples shown. (*B*) Quantification of dual TCRα GFP$^+$RFP$^+$ cells detectable after 5 d *in vitro* stimulation of isolated GFP$^+$RFP$^-$ and GFP$^-$RFP$^+$ T cells. Dots represent 3-6 individual samples per group from 3 independent experiments with mean ± SEM. (*C*) Dual GFP and RFP expression resulting from trogocytosis was tested by co-culturing a 1:1 ratio of paramagnetic bead-enriched T cells from B6.TCRA-GFP and B6.TCRA-RFP mice in the presence of stimulation with 1.0 μg/ml anti-CD3/CD28 mAbs and assessing reporter expression by flow cytometry at d 5. Representative samples shown. (*D*) Quantification of dual TCRα GFP$^+$RFP$^+$ cells detectable after 5 d *in vitro* co-culture of enriched T cells from B6.TCRA-GFP and B6.TCRA-RFP mice. Data shown as mean ± SEM of 3-6 individual samples per group from 3 independent experiments.

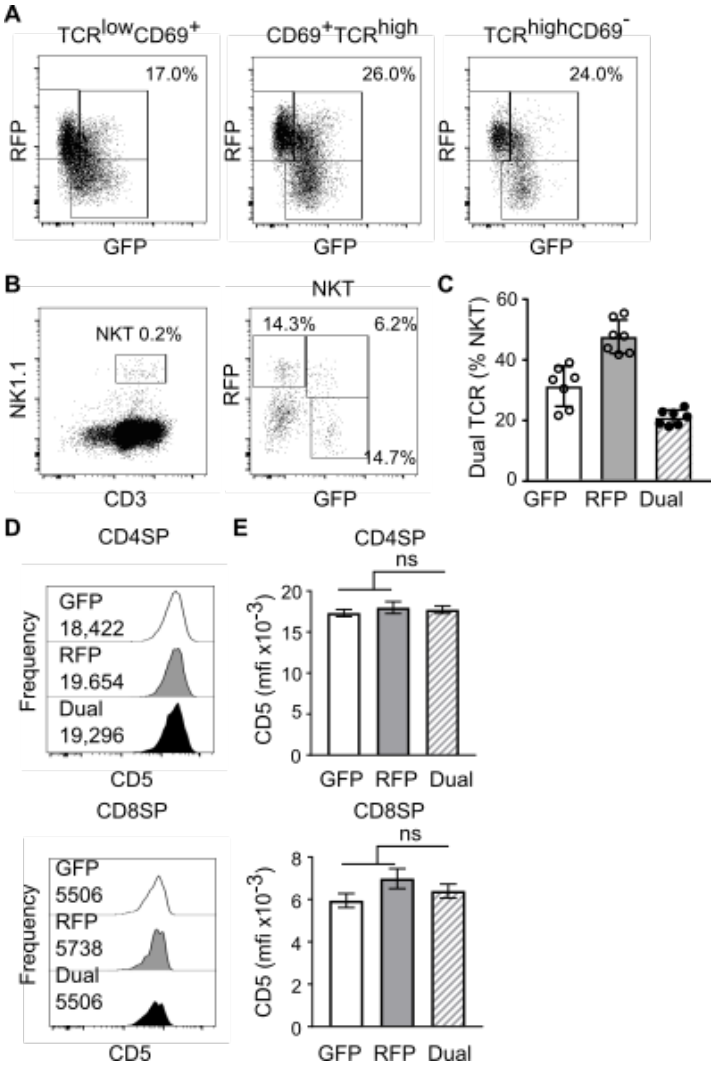

Figure S1.3: Evaluation of dual TCRα expression in thymocytes. (A) Dual TCRα expression by $TCR^{low}CD69^{+}$, $TCR^{high}CD69^{+}$, and $TCR^{high}CD69^{low}$ DP thymocytes from B6.TCRA-GFP/RFP mice as identified in Figure 1.2C by flow cytometry. Representative sample shown. (B) Identification of NKT thymocytes from B6.TCRA-GFP/RFP mice by flow cytometry. Representative sample shown. (C) Quantification of GFP^{+}, RFP^{+}, and $GFP^{+}RFP^{+}$ NKT thymocytes. Dots represent 7 individual mice from 3 independent experiments. (D) Comparison of CD5 expression on GFP^{+}, RFP^{+}, and $GFP^{+}RFP^{+}$ CD4SP and CD8SP cells with MFI from single representative sample shown. (E) Aggregate data for CD5 expression of 7 mice from 3 independent experiments, mean ± SD. Data compared using paired Student's t-test.

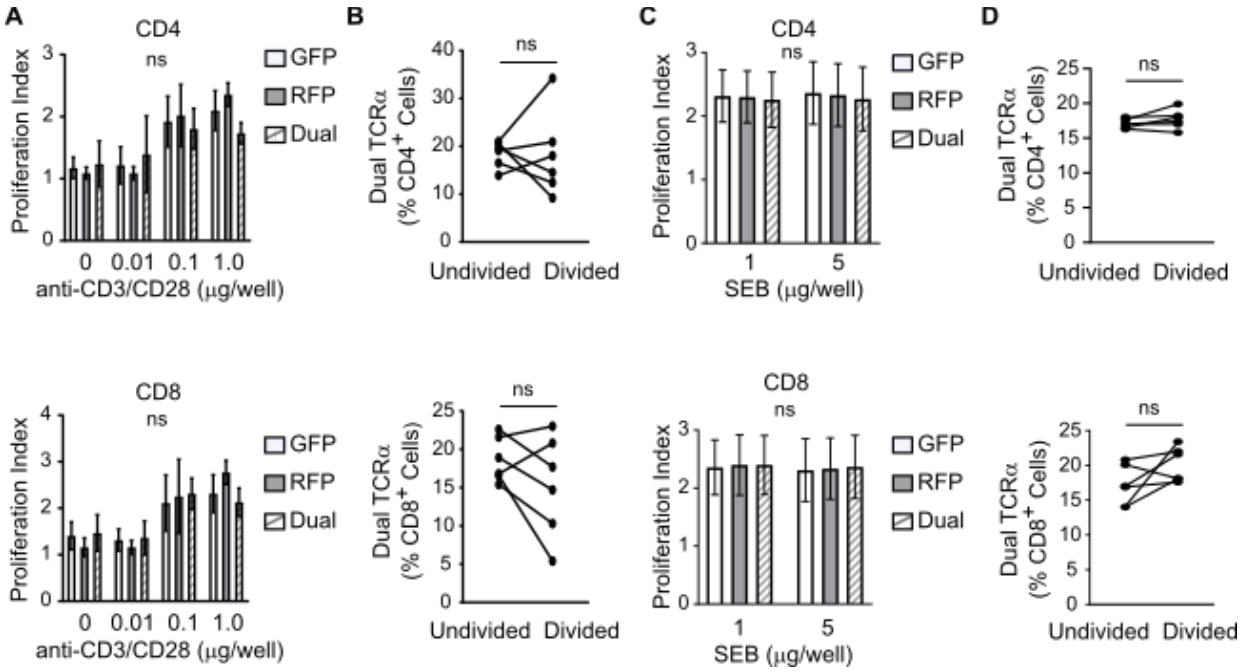

Figure S1.4: Dual TCR expression does not affect proliferative response to TCR stimulation. (*A*) Proliferation of paramagnetic bead-enriched T cells isolated from B6.TCRA-GFP/RFP splenocytes in response to 5 d *in vitro* stimulation with graded doses of anti-CD3/CD28 stimulation was measured by flow cytometry for dilution of CTV dye. TCRA-GFP$^+$, TCRA-RFP$^+$, and TCRA-GFP$^+$RFP$^+$ CD4$^+$ and CD8$^+$ T cells were gated and proliferation assessed as Proliferation Index. Data are mean ± SD of aggregate data from 3 independent experiments. Data analyzed using 2-way ANOVA with Tukey's multiple comparison test. (B) The frequency of dual TCRα GFP$^+$RFP$^+$ cells compared between undivided and divided cells from cultures stimulated with anti-CD3/CD28. Linked dots represent populations within the same sample from 3 replicate experiments. Data analyzed by Wilcoxan matched-pairs rank-sign test. (*C*) Proliferation of paramagnetic bead-enriched T cells isolated from B6.TCRA-GFP/RFP splenocytes in response to 5 d *in vitro* stimulation with graded doses of SEB was measured by flow cytometry for dilution of CTV dye. TCRA-GFP$^+$, TCRA-RFP$^+$, and TCRA-GFP$^+$RFP$^+$ CD4$^+$ and CD8$^+$ T cells were gated and proliferation assessed as Proliferation Index. Data are mean ± SD of aggregate data from 3 independent experiments. Data analyzed using 2-way ANOVA with Tukey's multiple comparison test. (D) The frequency of dual TCRα GFP$^+$RFP$^+$ cells compared between undivided and divided cells from cultures stimulated with SEB. Linked dots represent populations within the same sample from 3 replicate experiments. Data analyzed by Wilcoxan matched-pairs rank-sign test.

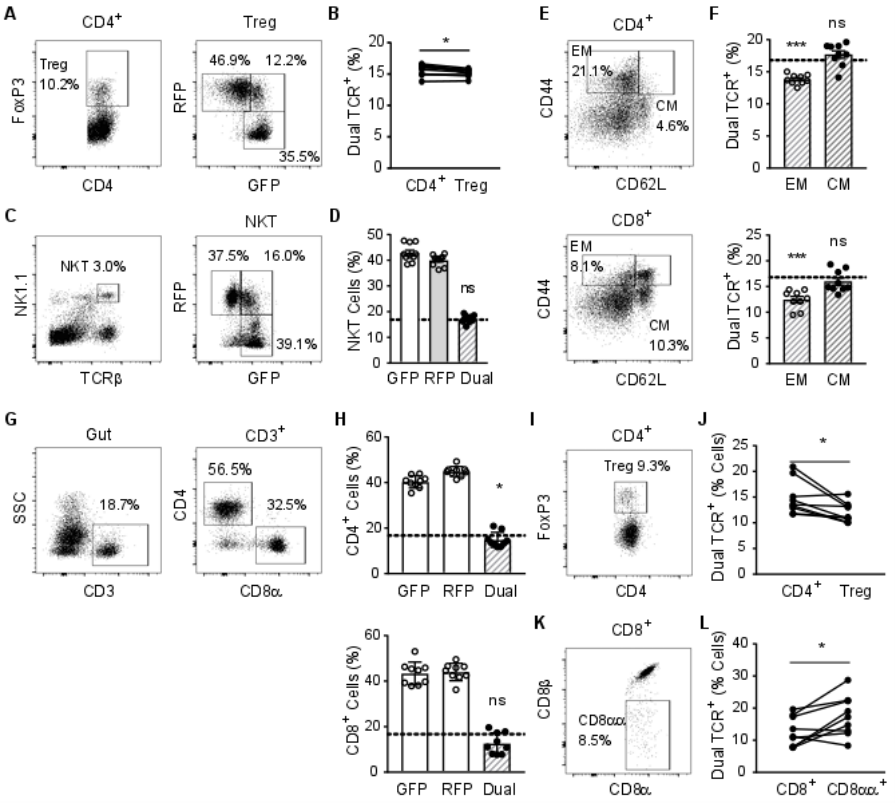

Figure S1.5: Dual TCRα expression has limited effects on T cell subsets in immunologically naive mice. (A) $CD4^+FoxP3^+$ Tregs identified from B6.TCRA-GFP/RFP splenocytes by flow cytometry. Representative sample shown. (B) Quantification of dual TCRα Tregs. Linked dots represent frequency of dual TCRα $CD4^+$ splenocytes and splenic Tregs from 9 individual mice from 3 independent experiments. Data analyzed by Wilcoxan matched-pairs rank-sign test. (C) NKT cells identified from B6.TCRA-GFP/RFP splenocytes by flow cytometry. Representative sample shown. (D) Quantification of GFP^+, RFP^+, and GFP^+RFP^+ NKT cells from splenocytes of 10 individual mice from 3 independent experiments. Dotted line represents frequency of dual TCRα T cells in the spleen. Dual TCRα NKT cell frequency compared to dual TCRα T cell frequency by Mann-Whitney test. (E) $CD44^{high}CD62L^-$ effector memory (EM) and $CD44^{high}CD62L^+$ central memory (CM) cells identified from $CD4^+$ and $CD8^+Tet^+$ splenocytes by flow cytometry. Representative sample shown. (F) Quantification of dual TCRα cells in EM (open circles) and CM (closed circle) subsets. Dots represent 9 individual mice from 3 independent experiments with mean ± SEM. Groups compared to total splenic dual TCRα $CD4^+$ and $CD8^+$ T cells (dotted line) using Mann-Whitney test. (G) Gut-associated T lymphocytes were isolated by enzymatic digestion and physical dissociation of intestines from B6.TCRA-GFP/RFP mice and examined by flow cytometry. Representative sample shown. (H) Quantification of GFP^+, RFP^+, and GFP^+RFP^+ $CD4^+$ and $CD8^+$ T cells from gut-associated T cells of 9 individual mice from 3 independent experiments. Dotted lines represent the average frequency of dual TCRα $CD4^+$ and $CD8^+$ T cells in the spleen. Gut-associated dual TCRα cell frequencies compared to splenic dual TCRα T cell frequencies by Mann-Whitney test. (I) Identification of gut-associated Tregs. Representative sample shown. (J) Quantification of dual TCRα gut-associated Tregs. Linked dots represent frequency of gut-associated dual TCRα $CD4^+$ cells and Tregs from 9 individual mice from 3 independent experiments. Data analyzed by Wilcoxan matched-pairs rank-sign test. (K) Identification of gut CD8αα T cells by flow cytometry. Representative sample shown. (L) Quantification of dual TCRα gut-associated CD8αα T cells. Linked dots represent frequency of gut-associated dual TCRα $CD8^+$ cells and CD8αα T cells from 9 individual mice from 3 independent experiments. Data analyzed by Wilcoxan matched-pairs rank-sign test. $^*P < 0.05$, $^{***}P < 0.005$, ns = not statistically significant.

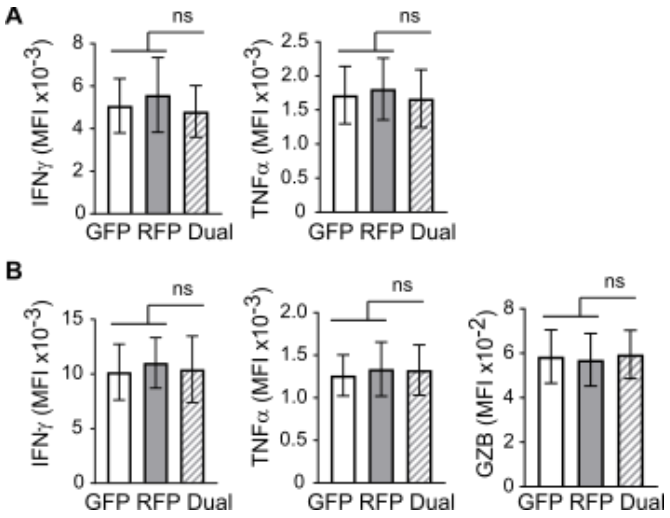

Figure S1.6: Single- and dual-TCR cells produce equivalent amounts of cytokine in response to antigen-specific stimulation. Cytokine production of T cells from LCMV-infected mice was assessed by intracellular flow cytometry after ex vivo stimulation of (A) $CD4^+$ and (B) $CD8^+$ splenocytes with LCMV GP_{66-77} and LCMV GP_{33-41} peptide. Cytokine production measured as MFI, mean ± SD for 11 mice from 3 independent experiments. Data compared using paired Student's t-test, ns = not statistically significant.

Acknowledgements.

Appendix, in full, is the reprint of the material as it appears in Proceedings of the National Academy of Sciences 2020. Yang, Letitia; Jama, Burhan; Wang, Huawei; Labarta-Bajo, Lara; Zúñiga, Elina I.; Morris, Gerald P., PNAS, 2020. The thesis author was a co-author of this paper.

REFERENCES

1. G. P. Morris, P. M. Allen, How the TCR balances sensitivity and specificity for the recognition of self and pathogens. *Nat. Immunol.* 13, 121–128 (2012).

2. E. Padovan, G. Casorati, P. Dellabona, S. Meyer, M. Brockhaus, A. Lanzavecchia, Expression of two T cell receptor α chains: dual receptor T cells. *Science* 262, 422–424 (1993).

3. E. Padovan, C. Giachino, M. Cella, S. Valitutti, O. Acuto, A. Lanzavecchia, Normal T lymphocytes can express two different T cell receptor β chains: Implications for the mechanism of allelic exclusion. *J. Exp. Med.* 181, 1587–1591 (1995).

4. S. M. Alam, I. N. Crispe, N. R. Gascoigne, Allelic exclusion of mouse T cell receptor α chains occurs at the time of thymocyte TCR up-regulation. *Immunity* 3, 449–458 (1995).

5. J. L. Casanova, P. Romero, C. Widmann, P. Kourilsky, J. L. Maryanski, T cell receptor genes in a series of class I major histocompatibility complex-restricted cytotoxic T lymphocyte clones specific for a Plasmodium berghei nonapeptide: Implications for T cell allelic exclusion and antigen-specific repertoire. *J. Exp. Med.* 174, 1371–1383 (1991).

6. P. Borgulya, H. Kishi, Y. Uematsu, H. von Boehmer, Exclusion and inclusion of α and β T cell receptor alleles. *Cell* 69, 529–537 (1992).

7. H. T. Petrie, F. Livak, D. G. Schatz, A. Strasser, I. N. Crispe, K. Shortman, Multiple rearrangements in T cell receptor alpha chain genes maximize the production of useful thymocytes. *J. Exp. Med.* 178, 615–622 (1993).

8. P. P. Ni, B. Solomon, C. S. Hsieh, P. M. Allen, G. P. Morris, The ability to rearrange dual TCRs enhances positive selection, leading to increased Allo- and Autoreactive T cell repertoires. *J. Immunol.* 193, 1778–1786 (2014).

9. F. Hardardottir, J. L. Baron, C. A. Janeway Jr, T cells with two functional antigen-specific receptors. *Proc. Natl. Acad. Sci. U.S.A.* 92, 354–358 (1995).

10. J. I. Elliott, D. M. Altmann, Dual T cell receptor α chain T cells in autoimmunity. *J. Exp. Med.* 182, 953–959 (1995).

11. T. Zal, S. Weiss, A. Mellor, B. Stockinger, Expression of a second receptor rescues self-specific T cells from thymic deletion and allows activation of autoreactive effector function. *Proc. Natl. Acad. Sci. U.S.A.* 93, 9102–9107 (1996).

12. A. Sarukhan, C. Garcia, A. Lanoue, H. von Boehmer, Allelic inclusion of T cell receptor α genes poses an autoimmune hazard due to low-level expression of autospecific receptors. *Immunity* 8, 563–570 (1998).

13. X. He, C. A. Janeway Jr, M. Levine, E. Robinson, P. Preston-Hurlburt, C. Viret, K. Bottomly, Dual receptor T cells extend the immune repertoire for foreign antigens. *Nat. Immunol*. 3, 127–134 (2002).

14. G. P. Morris, P. M. Allen, Cutting edge: Highly alloreactive dual TCR T cells play a dominant role in graft-versus-host disease. *J. Immunol*. 182, 6639–6643 (2009).

15. Q. Ji, A. Perchellet, J. M. Goverman, Viral infection triggers central nervous system autoimmunity via activation of CD8+ T cells expressing dual TCRs. *Nat. Immunol*. 11, 628–634 (2010).

16. E. Kekäläinen, A. Hänninen, M. Maksimow, T. P. Arstila, T cells expressing two different T cell receptors form a heterogeneous population containing autoreactive clones. *Mol. Immunol*. 48, 211–218 (2010).

17. J. L. Auger, S. Haasken, E. M. Steinert, B. A. Binstadt, Incomplete TCR-β allelic exclusion accelerates spontaneous autoimmune arthritis in K/BxN TCR transgenic mice. *Eur. J. Immunol*. 42, 2354–2362 (2012).

18. G. P. Morris, G. L. Uy, D. Donermeyer, J. F. Dipersio, P. M. Allen, Dual receptor T cells mediate pathologic alloreactivity in patients with acute graft-versus-host disease. *Sci. Transl. Med*. 5, 188ra74 (2013).

19. N. J. Schuldt, J. L. Auger, J. A. Spanier, T. Martinov, E. R. Breed, B. T. Fife, K. A. Hogquist, B. A. Binstadt, Cutting edge: Dual TCRα expression poses an autoimmune hazard by limiting regulatory T cell generation. *J. Immunol*. 199, 33–38 (2017).

20. D. A. Bolotin, M. Shugay, I. Z. Mamedov, E. V. Putintseva, M. A. Turchaninova, I. V. Zvyagin, O. V. Britanova, D. M. Chudakov, MiTCR: Software for T-cell receptor sequencing data analysis. *Nat. Methods* 10, 813–814 (2013).

21. A. Han, J. Glanville, L. Hansmann, M. M. Davis, Linking T-cell receptor sequence to functional phenotype at the single-cell level. *Nat. Biotechnol*. 32, 684–692 (2014).

22. A. Balakrishnan, N. Gloude, R. Sasik, E. D. Ball, G. P. Morris, Proinflammatory dual receptor T cells in chronic graft-versus-host disease. *Biol. Blood Marrow Transplant*. 23, 1852–1860 (2017).

23. S. M. Alam, N. R. Gascoigne, Posttranslational regulation of TCR Valpha allelic exclusion during T cell differentiation. *J. Immunol*. 160, 3883–3890 (1998).

24. H. Yang, H. Wang, R. Jaenisch, Generating genetically modified mice using CRISPR/Cas-mediated genome engineering. *Nat. Protoc*. 9, 1956–1968 (2014).

25. B. F. Koop, R. K. Wilson, K. Wang, B. Vernooij, D. Zaller, C. L. Kuo, D. Seto, M. Toda, L. Hood, Organization, structure, and function of 95 kb of DNA spanning the murine T-cell receptor C α/C δ region. *Genomics* 13, 1209–1230 (1992).

26. R. S. Friedman, P. Beemiller, C. M. Sorensen, J. Jacobelli, M. F. Krummel, Real-time analysis of T cell receptors in naive cells in vitro and in vivo reveals flexibility in synapse and signaling dynamics. *J. Exp. Med.* 207, 2733–2749 (2010).

27. A. Alcover, B. Alarcon, V. Di Bartolo, Cell biology of T cell receptor expression and regulation. *Ann. Rev. Immunol.* 36, 103–125 (2018).

28. A. Monjas, A. Alcover, B. Alarcón, Engaged and bystander T cell receptors are downmodulated by different endocytotic pathways. *J. Biol. Chem.* 279, 55376–55384 (2004).

29. E. Fernandez-Arenas, E. Calleja, N. Martínez-Martín, S. I. Gharbi, R. Navajas, N. García-Medel, P. Penela, A. Alcamí, F. Mayor Jr, J. P. Albar, B. Alarcón, β-Arrestin-1 mediates the TCR-triggered re-routing of distal receptors to the immunological synapse by a PKC-mediated mechanism. *EMBO J.* 33, 559–577 (2014).

30. J. F. Huang, Y. Yang, H. Sepulveda, W. Shi, I. Hwang, P. A. Peterson, M. R. Jackson, J. Sprent, Z. Cai, TCR-Mediated internalization of peptide-MHC complexes acquired by T cells. *Science* 286, 952–954 (1999).

31. G. L. Stritesky, S. C. Jameson, K. A. Hogquist, Selection of self-reactive T cells in the thymus. *Annu. Rev. Immunol.* 30, 95–114 (2012).

32. A. Tarakhovsky, S. B. Kanner, J. Hombach, J. A. Ledbetter, W. Müller, N. Killeen, K. Rajewsky, A role for CD5 in TCR-mediated signal transduction and thymocyte selection. *Science* 269, 535–537 (1995).

33. H. S. Azzam, A. Grinberg, K. Lui, H. Shen, E. W. Shores, P. E. Love, CD5 expression is developmentally regulated by T cell receptor (TCR) signals and TCR avidity. *J. Exp. Med.* 188, 2301–2311 (1998).

34. C. Sinclair, M. Saini, I. Schim van der Loeff, S. Sakaguchi, B. Seddon, The long-term survival potential of mature T lymphocytes is programmed during development in the thymus. *Sci. Signal.* 4, ra77 (2011).

35. A. Balakrishnan, B. Jama, G. P. Morris, Endogenous co-expression of two T cell receptors promotes lymphopenia-induced proliferation via increased affinity for self-antigen. *J. Leukoc. Biol.* 104, 1097–1104 (2018).

36. J. A. Punt, J. L. Roberts, K. P. Kearse, A. Singer, Stoichiometry of the T cell antigen receptor (TCR) complex: Each TCR/CD3 complex contains one TCR α, one TCR β, and two CD3 e chains. *J. Exp. Med.* 180, 587–593 (1994).

37. B. Ernst, D. S. Lee, J. M. Chang, J. Sprent, C. D. Surh, The peptide ligands mediating positive selection in the thymus control T cell survival and homeostatic proliferation in the periphery. *Immunity* 11, 173–181 (1999).

38. A. W. Goldrath, M. J. Bevan, Low-affinity ligands for the TCR drive proliferation of mature CD8+ T cells in lymphopenic hosts. *Immunity* 11, 183–190 (1999).

39. S. P. Persaud, C. R. Parker, W. L. Lo, K. S. Weber, P. M. Allen, Intrinsic CD4+ T cell sensitivity and response to a pathogen are set and sustained by avidity for thymic and peripheral complexes of self peptide and MHC. *Nat. Immunol.* 15, 266–274 (2014).

40. L. A. Kalekar, S. E. Schmiel, S. L. Nandiwada, W. Y. Lam, L. O. Barsness, N. Zhang, G. L. Stritesky, D. Malhotra, K. E. Pauken, J. L. Linehan, M. G. O'Sullivan, B. T. Fife, K. A. Hogquist, M. K. Jenkins, D. L. Mueller, CD4(+) T cell anergy prevents autoimmunity and generates regulatory T cell precursors. *Nat. Immunol.* 17, 304–314 (2016).

41. D. Homann, L. Teyton, M. B. Oldstone, Differential regulation of antiviral T-cell immunity results in stable $CD8^+$ but declining $CD4^+$ T-cell memory. *Nat. Med.* 7, 913–919 (2001).

42. D. S. McDermott, S. M. Varga, Quantifying antigen-specific CD4 T cells during a viral infection: CD4 T cell responses are larger than we think. *J. Immunol.* 187, 5568–5576 (2011).

43. S. M. Kaech, J. T. Tan, E. J. Wherry, B. T. Konieczny, C. D. Surh, R. Ahmed, Selective expression of the interleukin 7 receptor identifies effector CD8 T cells that give rise to long-lived memory cells. *Nat. Immunol.* 4, 1191–1198 (2003).

44. J. J. Sabatino, Jr, J. Huang, C. Zhu, B. D. Evavold, High prevalence of low affinity peptide-MHC II tetramer-negative effectors during polyclonal $CD4^+$ T cell responses. *J. Exp. Med.* 208, 81–90 (2011).

45. C. Rius, M. Attaf, K. Tungatt, V. Bianchi, M. Legut, A. Bovay, M. Donia, P. T. Straten, M. Peakman, I. M. Svane, S. Ott, T. Connor, B. Szomolay, G. Dolton, A. K. Sewell, Peptide-MHC class I tetramers can fail to detect relevant functional T cell clonotypes and underestimate antigen-reactive T cell populations. *J. Immunol.* 200, 2263–2279 (2018).

46. H. S. Robins, P. V. Campregher, S. K. Srivastava, A. Wacher, C. J. Turtle, O. Kahsai, S. R. Riddell, E. H. Warren , C. S. Carlson, Comprehensive assessment of T-cell receptor β-chain diversity in alphabeta T cells. *Blood* 114, 4099–4107 (2009).

47. L. A. Colf, A. J. Bankovich, N. A. Hanick, N. A. Bowerman, L. L. Jones, D. M. Kranz, K. C. Garcia, How a single T cell receptor recognizes both self and foreign MHC. *Cell* 129, 135–146 (2007).

48. J. M. Robertson, B. D. Evavold, Cutting edge: Dueling TCRs: Peptide antagonism of $CD4^+$ T cells with dual antigen specificities. *J. Immunol.* 163, 1750–1754 (1999).

49. B. N. Dittel, I. Stefanova, R. N. Germain, C. A. Janeway Jr, Cross-antagonism of a T cell clone expressing two distinct T cell receptors. *Immunity* 11, 289–298 (1999).

50. R. M. Teague, P. D. Greenberg, C. Fowler, M. Z. Huang, X. Tan, J. Morimoto, M. L. Dossett, E. S. Huseby, C. Ohlén, Peripheral CD8+ T cell tolerance to self-proteins is regulated proximally at the T cell receptor. *Immunity* 28, 662–674 (2008).

51. A. E. Moran, K. L. Holzapfel, Y. Xing, N. R. Cunningham, J. S. Maltzman, J. Punt, K. A. Hogquist, T cell receptor signal strength in Treg and iNKT cell development demonstrated by a novel fluorescent reporter mouse. *J. Exp. Med.* 208, 1279–1289 (2011).

52. E. S. Huseby, F. Crawford, J. White, J. Kappler, P. Marrack, Negative selection imparts peptide specificity to the mature T cell repertoire. *Proc. Natl. Acad. Sci. U.S.A.* 100, 11565–11570 (2003).

53. B. D. McDonald, J. J. Bunker, S. A. Erickson, M. Oh-Hora, A. Bendelac, Crossreactive αβ T cell receptors are the predominant targets of thymocyte negative selection. *Immunity* 43, 859–869 (2015).

54. E. Corse, R. A. Gottschalk, J. P. Allison, Strength of TCR-peptide/MHC interactions and in vivo T cell responses. *J. Immunol.* 186, 5039–5045 (2011).

55. N. J. Tubo, A. J. Pagán, J. J. Taylor, R. W. Nelson, J. L. Linehan, J. M. Ertelt, E. S. Huseby, S. S. Way, M. K. Jenkins, Single naive CD4+ T cells from a diverse repertoire produce different effector cell types during infection. *Cell* 153, 785–796 (2013).

56. M. A. Williams, E. V. Ravkov, M. J. Bevan, Rapid culling of the CD4+ T cell repertoire in the transition from effector to memory. *Immunity* 28, 533–545 (2008).

57. I. Stefanová, J. R. Dorfman, R. N. Germain, Self-recognition promotes the foreign antigen sensitivity of naive T lymphocytes. *Nature* 420, 429–434 (2002).

58. J. N. Mandl, J. P. Monteiro, N. Vrisekoop, R. N. Germain, T cell-positive selection uses self-ligand binding strength to optimize repertoire recognition of foreign antigens. *Immunity* 38, 263–274 (2013).

59. R. B. Fulton, S. E. Hamilton, Y. Xing, J. A. Best, A. W. Goldrath, K. A. Hogquist, S. C. Jameson, The TCR's sensitivity to self peptide-MHC dictates the ability of naive CD8(+) T cells to respond to foreign antigens. *Nat. Immunol.* 16, 107–117 (2015).

60. J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, A. Cardona, Fiji: An open-source platform for biological-image analysis. *Nat. Methods* 9, 676–682 (2012).

61. M. Roederer, Interpretation of cellular proliferation data: Avoid the panglossian. *Cytometry* A 79, 95–101 (2011).

62. W. L. Lo, P. M. Allen, Self-awareness: how self-peptide/MHC complexes are essential in the development of T cells. *Mol Immunol.* 55(2), 186-189 (2013).

63. L. Klein, B. Kyewski, P. M. Allen, K. A. Hogquist, Positive and negative selection of the T cell repertoire: what thymocytes see (and don't see). *Nat Rev Immunol* 14**,** 377–391 (2014).

64. R. Tripathi, A. Jackson, M. S. Krangel, A Change in the Structure of Vβ Chromatin Associated with TCR β Allelic Exclusion. *J Immunol*, 168(5), 2316-2324 (2002).

65. D. Balomenos, R. S. Balderas, K. P. Mulvany, J. Kaye, D. H. Kono, A. N. Theofilopoulos, Incomplete T cell receptor V beta allelic exclusion and dual V beta-expressing cells. *J Immunol*. 155 (7), 3308-3312 (1995).

66. E. S. Huseby, F. Crawford, J. White, T. Vass, D. Becker, C. Pinilla, P. Marrack, J. W. Kappler, How the T cell repertoire becomes peptide and MHC specific. *Cell*. 122, 247–260 (2005).

67. A. Casrouge, E. Beaudoing, S. Dalle, C. Pannetier, J. Kanellopoulos, P. Kourilsky, Size Estimate of the αβ TCR Repertoire of Naive Mouse Splenocytes. *J Immunol* 164 (11), 5782-5787 (2000).

68. T. P. Arstila, A. Casrouge, V. Baron, J. Even, J. Kanellopoulos, P. Kourilsky, A direct estimate of the human αβ T cell receptor diversity. *Science* 286, 958–961 (1999).

69. A. Farmanbar, R Kneller, S. Firouzi, RNA sequencing identifies clonal structure of T-cell repertoires in patients with adult T-cell leukemia/lymphoma. *npj Genom. Med*. 4, 10 (2019)

70. N. Felix, P. Allen, Specificity of T-cell alloreactivity. *Nat Rev Immunol* 7**,** 942–953 (2007).

71. E. Rosati, C. M. Dowds, E. Liaskou, E. Henriksen, T. H. Karlsen, A. Franke, Overview of methodologies for T-cell receptor repertoire analysis. *BMC biotechnology*, 17(1), 61 (2017).

72. M. Matz, D. Shagin, E. Bogdanova, O. Britanova, S. Lukyanov, L. Diatchenko, A. Chenchik, Amplification of cDNA ends based on template-switching effect and step-out PCR. *Nucleic Acids Research*, 27 (6), 1558–1560 (1999).

73. F.S. Roberts, Measurement of Biodiversity: Richness and Evenness. In: Kaper H., Roberts F. (eds) Mathematics of Planet Earth. *Mathematics of Planet Earth*, vol 5. Springer, Cham (2019).

74. N. Tickotsky, T. Sagiv, J. Prilusky, E. Shifrut, N. Friedman, McPAS-TCR: a manually curated catalogue of pathology-associated T cell receptor sequences. *Bioinformatics*, 33(18):2924-2929 (2017).

75. E. Alspach, D. M. Lussier, A. P. Miceli, I. Kizhvatov, M. DuPage, A. M. Luoma, W. Meng, C. F. Lichti, E. Esaulova, A. N. Vomund, D. Runci, J. P. Ward, M. M. Gubin, R. F. V. Medrano, C. D. Arthur, J. M. White, K. C. F. Sheehan, A. Chen, K. W.

Wucherpfennig, T. Jacks, E. R. Unanue, M. N. Artyomov, R. D. Schreiber, MHC-II neoantigens shape tumour immunity and response to immunotherapy. *Nature*, 574(7780):696-701 (2019).

76. J. Sprent, C. Surh, Normal T cell homeostasis: the conversion of naive cells into memory-phenotype cells. *Nat Immunol* 12**,** 478–484 (2011).

77. J. A. Carter, J. B. Preall, K. Grigaityte, S. J. Goldfless, E. Jeffery, A. W. Briggs, F. Vigneault, G. S. Atwal, Single T Cell Sequencing Demonstrates the Functional Role of $\alpha\beta$ TCR Pairing in Cell Lineage and Antigen Specificity. *Front Immunol*. 10:1516 (2019).

78. B. D. Stadinski, P. Trenh, B. Duke, P. G. Huseby, G. Li, L. J. Stern, E. S. Huseby, Effect of CDR3 sequences and distal V gene residues in regulating TCR–MHC contacts and ligand specificity. *J Immunol*. 192 (12) 6071-6082 (2014).

79. S. A. Overall, J. S. Toor, S. Hao, M. Yarmarkovich, S. M. O'Rourke, G. I. Morozov, S. Nguyen, A. S. Japp, N. Gonzalez, D. Moschidi, M. R. Betts, J. M. Maris, P. Smibert, N. G. Sgourakis, High throughput pMHC-I tetramer library production using chaperone-mediated peptide exchange. *Nat Commun*. 11**,** 1909 (2020).