JCI The Journal of Clinical Investigation

Crossreactive public TCR sequences undergo positive selection in the human thymic repertoire

Mohsen Khosravi-Maharlooei, … , Todd M. Brusko, Megan Sykes

J Clin Invest. 2019. https://doi.org/10.1172/JCI124358.

Research In-Press Preview **Autoimmunity** **Immunology**

We investigated human T-cell repertoire formation using high throughput TCRβ CDR3 sequencing in immunodeficient mice receiving human hematopoietic stem cells (HSCs) and human thymus grafts. Replicate humanized mice generated diverse and highly divergent repertoires. Repertoire narrowing and increased CDR3β sharing was observed during thymocyte selection. While hydrophobicity analysis implicated self-peptides in positive selection of the overall repertoire, positive selection favored shorter shared sequences that had reduced hydrophobicity at positions 6 and 7 of CDR3βs, suggesting weaker interactions with self-peptides than unshared sequences, possibly allowing escape from negative selection. Sharing was similar between autologous and allogeneic thymi and occurred between different cell subsets. Shared sequences were enriched for allo-crossreactive CDR3βs and for Type 1 diabetes-associated autoreactive CDR3βs. Single-cell TCR-sequencing showed increased sharing of CDR3αs compared to CDR3βs between mice. Our data collectively implicate preferential positive selection for shared human CDR3βs that are highly cross-reactive. While previous studies suggested a role for recombination bias in producing “public” sequences in mice, our study is the first to demonstrate a role for thymic selection. Our results implicate positive selection for promiscuous TCRβ sequences that likely evade negative selection, due to their low affinity for self-ligands, in the abundance of “public” human TCRβ sequences.

Title: Crossreactive public TCR sequences undergo positive selection in the human thymic repertoire

Authors: Mohsen Khosravi-Maharlooei[1,5,*], Aleksandar Obradovic[1,*], Aditya Misra[1], Keshav Motwani[6], Markus Holzl[1,5], Howard R. Seay[6], Susan DeWolf[1], Grace Nauman[1,5], Nichole Danzl [1,5], Haowei Li[1,5], Siu-hong Ho[1], Robert Winchester[2], Yufeng Shen[3], Todd M. Brusko [6], Megan Sykes[1,4,5]

Affiliations: [1] Columbia Center for Translational Immunology, Department of Medicine, Columbia University Medical Center, New York, NY, 10032, USA;

[2] Division of Rheumatology, Department of Medicine, Columbia University Medical Center, New York, NY, 10032, USA;

[3] Center for Computational Biology and Bioinformatics, Columbia University Medical Center, New York, USA, 10032, USA;

[4] Department of Surgery, Columbia University Medical Center, Columbia University, New York, NY, 10032, USA;

[5] Department of Microbiology & Immunology, Columbia University Medical Center, Columbia University, New York, NY, 10032, USA;

[6] Department of Pathology, Immunology and Laboratory Medicine, University of Florida, Gainesville, FL, 32610, USA.

***** These authors equally contributed to this work.

Declaration of interests

The authors have declared that no conflict of interest exists.

Corresponding Author:

Megan Sykes, MD, Columbia Center for Translational Immunology

650 West 168th Street, Black Building, 1512, (Mailbox 127)

New York, NY 10032

Tel: 212.304.5696 Fax: 646.426.0019 megan.sykes@columbia.edu

Abstract

We investigated human T-cell repertoire formation using high throughput TCRβ CDR3 sequencing in immunodeficient mice receiving human hematopoietic stem cells (HSCs) and human thymus grafts. Replicate humanized mice generated diverse and highly divergent repertoires. Repertoire narrowing and increased CDR3β sharing was observed during thymocyte selection. While hydrophobicity analysis implicated self-peptides in positive selection of the overall repertoire, positive selection favored shorter shared sequences that had reduced hydrophobicity at positions 6 and 7 of CDR3βs, suggesting weaker interactions with self-peptides than unshared sequences, possibly allowing escape from negative selection. Sharing was similar between autologous and allogeneic thymi and occurred between different cell subsets. Shared sequences were enriched for allo-crossreactive CDR3βs and for Type 1 diabetes-associated autoreactive CDR3βs. Single-cell TCR-sequencing showed increased sharing of CDR3αs compared to CDR3βs between mice. Our data collectively implicate preferential positive selection for shared human CDR3βs that are highly cross-reactive. While previous studies suggested a role for recombination bias in producing “public” sequences in mice, our study is the first to demonstrate a role for thymic selection. Our results implicate positive selection for promiscuous TCRβ sequences that likely evade negative selection, due to their low affinity for self-ligands, in the abundance of “public” human TCRβ sequences.

Introduction

A functional TCR repertoire requires great diversity to recognize a wide range of pathogens. Extrapolations of early sequencing results on a few hundred TCRs led to an estimate of about 10^6 different TCR β chains in human blood, each pairing, on average, with at least 25 different α chains [1]. Recent studies have used next-generation TCR sequencing to capture the diversity of TCR repertoires. The first such study estimated that normal human blood contains 3-4 × 10^6 unique TCR β chains [2], while a later study provided a minimal estimate of 100 × 10^6 unique TCRβ sequences in naïve $CD4^+$ and $CD8^+$ T-cell repertoires of young adults [3]. Another study estimated that there are 40-70 × 10^6 unique TCR β sequences and 60-100 × 10^6 TCRα sequences in human pediatric thymi [4]. However, a surprising degree of repertoire overlap has been observed between individuals in studies of peripheral T cells [5]. *In silico* modeling suggested a role for recombination bias in generating shared sequences in mice [6]. In a human system, we have investigated the role of thymic selection in the generation of shared sequences.

Diversity of the TCR repertoire is formed at different levels. First, different Vβ, Dβ and Jβ genes recombine to generate TCR β chains. Similarly, different Vα and Jα genes recombine to generate TCR α chains [7]. Exonuclease removal of exposed residues and addition of random nucleotides at the junctional sites mediated by the enzyme terminal deoxynucleotidyl transferase (TdT) further diversifies the TCR repertoire [8]. Finally, α and β chains combine to form functional TCRs. Studies in mice have shown that thymocytes undergo different selection processes that shape the TCR repertoire after formation of functional TCRs [9,10]. This repertoire is further altered in the periphery by expansion and deletion of certain clones [11]. Our knowledge of the formation of the human TCR repertoire is limited by access to human thymus samples and the inability to manipulate variables *in vivo*. Humanized mice generated by transplanting immunodeficient mice with human thymus and hematopoietic stem cells provide a unique opportunity to investigate the factors involved in the formation of a human TCR repertoire. The few studies in the literature that have used humanized mice to investigate the human

thymus T cell repertoire[12–15] are limited to V-J gene usage and CDR3 length distribution analysis (spectratyping). Here, we performed high-throughput and single cell TCR immunosequencing of human thymocyte subsets generated in human thymi and periphery of replicate humanized mice to determine whether or not exposure to the same antigens for positive and negative selection in thymus tissue from the same human donor would lead to the formation of similar TCR repertoires in different individuals. We also used this model to investigate the impact of positive and negative selection on the human TCR repertoire and to examine the factors that determine V-J pairing. While initial repertoire formation was highly stochastic and differed markedly between biological replicates, we demonstrate preferential selection of a set of “public” sequences that can be selected by disparate HLA alleles. Comparison with known cross-alloreactive and type 1 diabetes (T1D)-associated autoreactive TCRs support the interpretation that these are highly cross-reactive TCRs that are preferentially selected and shared between different individuals.

Results

Kinetics of human cell development in humanized mice and the histology of grafted thymi

To study the formation of the human thymic and peripheral TCR repertoire, three batches of mice were generated. As shown in Figure 1A, the first experiment consisted of three mice (1autoA, 1autoB and 1autoC) that were generated with the same fetal liver HSCs and autologous fetal thymus. Therefore, they had the same genetic background and selection took place in the same thymus (Figure 1A). The second batch consisted of six mice that were generated with the same fetal liver HSCs (different from Experiment 1). Mice designated “2autoA, 2autoB and 2autoC” received an autologous fetal thymus, while the mice designated “2alloA, 2alloB and 2alloC” received an allogeneic fetal thymus so that thymic selection occurred in a different thymus while the thymocyte genetic backgrounds were the same as that of the other three mice (Figure 1B). Experiment 3 consisted of two mice (2autoA and 2autoB) that were thymectomized and transplanted with the same fetal liver HSCs and autologous fetal thymus. Peripheral CD4 and CD8 cells were also sequenced and analyzed in these mice, in addition to thymic SP-CD4 and SP-CD8 cells. The mice in the first experiment were euthanized at 14 weeks post-transplantation, while the mice in the second and third experiments were euthanized at 20 and 22 weeks post-transplantation, respectively. Figure S1A shows the gross appearance of the spleen, lymph nodes and the grafted thymus under the kidney capsule of a representative humanized mouse at the time of harvest. Figure S1B shows hematoxylin and eosin (H&E) staining of a representative grafted thymus and a thymus from a 13-year-old child. Cortical (hypercellular) and medullary (hypocellular) areas and Hassall's corpuscles in the medullary areas are noticeable in the H&E stains. Immunofluorescent staining of a representative grafted thymus and a thymus from a 13-year-old child stained for HLA-DR, cytokeratin (CK) 8 and CK14 is shown in Figure S1C. HLA-DR^{+} cells that are not stained for CKs are HSC-derived antigen presenting cells (APCs) that are mainly concentrated in medullary areas. We further characterized the APCs in grafted thymi by FCM. B cells (CD19^{+}),

monocytes ($CD14^+$) and dendritic cells ($CD11c^+$) collectively constituted about 30% of the double negative ($CD4^-$ $CD8^-$) cells in grafted thymi (Fig S1D).

The kinetics of the peripheral appearance of human immune cells ($hCD45^+$), B cells ($CD19^+$) and T cells ($CD3^+$), as well as the T cell naïve/memory phenotype are shown in Figure S1E-H. The majority of T cells in peripheral blood at weeks 14-16 were naïve.

Our method for constructing humanized mice included several measures to eliminate pre-existing thymocytes and their progeny from the transplanted fetal thymic tissue. These measures include freezing and thawing the thymus tissues as described [16], pipetting up and down to physically release thymocytes and injecting 2 weekly doses of a depleting anti-CD2 antibody as described [16]. To assess the role of cells carried in the thymic tissue in producing peripheral and intrathymic T cell populations in this model, we generated a batch of mice with allogeneic fetal HSCs and thymus tissue. The fetal thymic cells were HLA-$A3^-$ while the fetal HSCs were HLA-$A3^+$. At 24 weeks post-transplantation, we euthanized the animals and evaluated the origin of T cells in grafted thymi and peripheral lymphoid tissues. Approximately 3% of DP and SP-CD8 thymocytes and 2% of SP-CD4 cells were thymus graft-derived (HLA-$A3^-$) (Figure S1I). Approximately 0.5% of CD4 and CD8 cells in the spleen were thymus graft-derived (Figure S1J). Therefore, the majority of T cells in the grafted thymi and spleens of these animals were derived from the HSCs that were given intravenously.

Effect of selection on diversity

The cell counts of grafted thymi in addition to the sorted cell numbers are summarized in Table S1. For each sample, we obtained template counts, clonality scores and unique clone counts at the nucleotide level (for both productive rearrangements, and non-productive rearrangements that include frame shifts or premature stop codons) and the amino acid level. These are shown in Table S1. Template counts for $CD69^-$ DP cells were lower than expected from the number of cells, likely reflecting the rearrangement of TCRβ after acquisition of the DP phenotype in a significant fraction of cells [17].

Clonality (a normalized measure of inverse diversity based on CDR3β sequences) in all thymic samples was very low, demonstrating production and selection of a highly diverse repertoire in the human thymus grafts. Clonality scores are typically much higher for both CD4 and CD8 T cells in human peripheral blood, most markedly for CD8 cells, presumably reflecting antigen-driven expansions [18]. Accordingly, clonality of peripheral CD4 and CD8 cells was markedly higher than that of thymic SP-CD4 and SP-CD8 cells in Experiment 3 (Figure 1F). Although only some differences achieved statistical significance, all thymocyte subsets (CD8 SP, CD4 SP non-Tregs, CD4 Tregs) showed increased clonality scores for amino acid compared to nucleotide sequences (Figures 1D and 1E and Table S1). Collectively, these results show the effect of selection on narrowing the TCR repertoire, since selection is applied to TCR protein and multiple productive nucleotide sequences can produce the same peptide sequence.

In Experiment 2, in which $CD69^-$ and $CD69^+$ DP thymocytes were sequenced in addition to the three SP subsets, each animal demonstrated very low clonality scores for the $CD69^-$ DP population. With the exception of one animal (2alloB), we were unable to detect a positive selection-induced increase in clonality in the $CD69^-$ to $CD69^+$ transition. However, comparison of $CD69^-$ DP populations and all 3 SP subsets (SP-CD4 non-Treg [referred to henceforth as SP-CD4 for simplicity], SP-CD8 and CD4 Tregs [referred to henceforth as Tregs for simplicity]) revealed an increase in amino acid sequence clonality (Figure 1E and Table S1). Collectively, these data demonstrate narrowing of the T cell repertoire due to thymic selection.

Compared to the original fetal thymus, clonality scores were lower in the grafted thymi for SP-CD8 and Treg cell populations (Figure 1G), demonstrating greater diversity of the thymocytes generated from human HSCs used to construct humanized mice than in the original fetal thymus, which had a gestational age of 17 weeks, when thymic development and generation of a fully diversified repertoire is not complete. It has been previously reported that the TCR repertoire of mouse neonates (day 1 after birth) is much narrower than that of adult mice, due to lack of random nucleotide insertions in the CDR3s

[19]. As shown in Figure 1H, TdT is not expressed in DP cells of fetal human thymus. However, it is expressed in DP cells of post-natal thymi as well as grafted human thymus in humanized mice, thus explaining the greater diversity of TCR repertoires in grafted human thymi in our study compared to the fetal human thymus.

Role of stochastic rearrangement and selection in TCR repertoire formation

To obtain an understanding of the impact of stochastic TCR rearrangement vs. background genetics on the TCR repertoire, we compared repertoires generated under the same conditions from the same progenitor pool across identical as well as allogeneic, extensively HLA-mismatched (Table S2) thymi by measuring the Jenson-Shannon Divergence (JSD) and the number of shared CDR3β TCR sequences, as shared CDR3β fraction quantifies sharing of unique sequences and JSD additionally accounts for frequency of shared sequences. In Experiment 1, even though all three mice received the same HSCs and thymus from the same human fetal donor, their TCR repertoires at the level of CDR3β were highly divergent at both the nucleotide and amino acid levels (Figure 2A). In Experiment 2, in which six mice received the same HSCs, with mice 2autoA, 2autoB and 2autoC receiving autologous thymus and mice 2alloA, 2alloB and 2alloC receiving allogeneic thymus tissues, there was a similarly high divergence among all thymi in different cell populations (Figure 2B). Furthermore, there was no difference in divergence between pairs of mice whose T cells developed in the same thymus vs those whose T cells developed in the allogeneic thymi (Table S3). In addition, observed divergence between mice for both amino acid and nucleotide repertoires was significantly higher than baseline generated from repeated under-sampling of identical repertoires for all thymic subpopulations, both by JSD (Figures S3A) and by shared CDR3β fraction (Figures S3B). All of these findings emphasize the highly stochastic nature of TCR repertoire formation at the level of CDR3β.

In all experiments, and in every thymocyte and peripheral subset, divergence was lower at the amino acid level compared to the nucleotide level and the JSD decreased for selected (CD69+ DP and SP populations) compared to unselected (CD69- DP) populations at the amino acid but not the nucleotide

level (Figures 2A-C). This finding suggested that, despite the stochastic nature of repertoire formation, thymic selection in identical thymi results in selection of some shared sequences between individuals.

We compared the fraction of CDR3βs that were shared between every possible pair of mice for each thymocyte population. As shown in Figure 3, the fraction of shared CDR3βs between paired mice in all three experiments was less than 4% of all thymic TCRβs of each mouse. The proportion of shared CDR3βs was highest at the amino acid level and also was higher at the productive/nucleotide level compared to the non-productive/nucleotide level (Fig 3A-C, Table S4). In addition, the proportion of shared CDR3βs increased significantly during transition from the $CD69^-$ DP to the $CD69^+$ DP stage, indicating positive selection for these shared CDR3βs. The proportion of shared CDR3βs further increased for the sorted mature ($CD3^{high}CD5^{high}$) SP CD4 and CD8 cell populations, which have completed positive selection and partially undergone negative selection, compared to the positively selected $CD69^+$ DP population (Figure 3B). CDR3β sharing was even higher in peripheral CD4 and CD8 subsets compared to thymic SP-CD4 and SP-CD8 samples in Experiment 3 (Figure 3C), possibly due to repertoire narrowing via completion of negative selection and post-thymic selection and expansion of certain clones. Pairwise divergence analyses by JSD yielded similar results, with significant decreases in DP $CD69^+$ compared to DP $CD69^-$ cells and further decreases in the mature SP cell populations (SP-CD4, SP-CD8) (Figure 2B) and peripheral populations (p. CD4, P. CD8) (Fig 2C). Together, these findings demonstrate that both positive and negative selection of human thymocytes increase CDR3β overlap between individual T cell repertoires. However, the highest proportion of sequences shared between two replicate thymic repertoires found in the SP-CD4 subset accounted for only 3.5% of the repertoire.

As another readout of the effect of selection, we compared AA sequence convergence of nucleotide sequences for shared and unshared CDR3βs. For each pair of mice in Experiment 2, we measured the number of unique CDR3β nucleotide sequences corresponding to each amino acid sequence shared between the same cell population in both mice (shared) compared to the number of unique nucleotide

sequences corresponding to each amino acid sequence present in at least one of the mice but not shared between both (unshared). While the average nucleotide-per-amino-acid sequence ratio was close to 1 for unshared sequences, it was significantly higher for the shared CDR3βs in all populations, indicating preferential selection of shared amino acid sequences (Figure 3G, Table S5). Within the population of shared CDR3βs, this ratio was significantly higher in DP CD69^{+} cells compared to DP CD69^{-} cells, and in SP cell populations (except Tregs) compared to DP CD69^{+} cells, indicating selection for the shared sequences (Table S6).

The JSD between different mice was lower among the 100 most frequent sequences compared to all sequences for each of the five selected and non-selected cell populations, indicating greater overlap among the more abundant sequences (Figure 2D). Consistently, the fraction of CDR3B sequences overlapping between animals was greater among the 100 top sequences compared to the entire population (Figure 3D). While this finding may reflect the greater likelihood of detecting abundant sequences in general, it is also consistent with the possibility that the shared CDR3βs are preferentially selected.

Surprisingly, for the five different selected and non-selected cell populations, the proportion of shared CDR3βs was not different between mice with allogeneic vs autologous thymi (Figures 3E). Furthermore, there was no dramatic increase in shared CDR3βs among mice within an experiment compared to those between experiments, despite the different genetic background of HSCs and thymi used to generate the T cells in each experiment (Figure 3F). In addition, different cell subsets in allogeneic and autologous thymi in Experiment 2 had similar divergences compared to the original autologous fetal thymus (Figure 2E). Tables S7 and S8 show the numbers of unique and shared CDR3βs between the three mice that were generated from the same thymus and HSCs in Experiment 1 and the six mice with allogeneic and autologous thymi in Experiment 2, respectively. Table S9 shows the total number of shared and non-shared CDR3βs for each cell population in each experiment at both amino acid and nucleotide levels. Consistent with results described above, the number of overlapping

CDR3βs increased as selection progressed and there were dramatically greater numbers of overlapping CDR3βs at the amino acid sequence level compared to the nucleotide sequence level.

In order to address variable template counts across samples, we validated the results of repertoire divergence analysis by randomly subsampling each sample to the same template count and then repeating the analysis. With three subsamples of 1,000 templates each, we observed the same trends as in whole-sample comparisons by shared CDR3β fraction (Figure S3C). Namely, we observed consistently increased sharing at the amino acid level compared to nucleotide level and increased sharing in DP CD69^{+} samples compared to DP CD69^{-} samples, and from DP to SP samples. Therefore, our results are stable across random subsamples of the data, regardless of variable sample size (Figure S3C).

Shared CDR3βs have shorter length due to fewer N insertions than unique CDR3βs, often use different V genes

Further characterization of the CDR3βs that were shared between any two thymi vs those that were detected in only one thymus ("unique" sequences) revealed that the shared CDR3βs were significantly shorter than the unique CDR3βs. The shared CDR3βs had an average length of around 40 nucleotides, while the unique CDR3βs had an average CDR3β length of around 44 nucleotides (Figure 4A). Number of inserted nucleotides at V-D and D-J junctions was significantly lower for the shared CDR3βs compared to the unshared CDR3βs (Figure 4B). As the number of V and J nucleotide deletions were slightly higher in unshared CDR3βs (Figures 4C and 4D), the shorter length of shared CDR3βs was thus attributable to the lower number of nucleotide insertions in these sequences. The shorter length of shared CDR3βs did not simply reflect the fact that they tended to be relatively abundant, as the average CDR3 length of the 1000 most abundant CDR3βs overall was significantly greater than that of the 1000 least abundant CDR3βs (Figure 4E). However, each animal demonstrated CDR3β shortening as selection progressed from the CD69+ DP to the SP stage among the 1000 most abundant but not among the least abundant sequences (Figure 4E, Table S10), indicating a selective preference for

shorter shared CDR3βs. Shortening of CDR3βs continued further in the transition from thymic SP to peripheral CD4 and CD8 cells (Figure 4F).

We characterized the V and J gene usage of shared CDR3βs among SP CD4 cells. Only about 20-25% of CDR3βs shared between different thymi used the same V gene, while almost all shared CDR3βs used the same J gene. The ratio of shared CDR3βs that used the same V-J pair and hence the same TCRβ chain is therefore 20-25% (Figure 4G). These ratios were not different comparing allogeneic vs. autologous thymi

TCR β chain overlap between different cell subsets in individual human thymus grafts

To better understand the selection of shared CDR3β sequences, we compared CDR3β sequences of SP and DP T cell populations within each mouse in Experiment 2. As shown in Figure 5A, there was overlap in CDR3β sequences between SP thymic populations of each mouse, especially among the 100 most frequent sequences. Among the sequences with identical CDR3β in different mature thymocyte subsets, approximately 60% used the same V gene within individual mice (Figure 5B), while about 40% used different V genes. Almost all of the shared CDR3βs were associated with the same J gene, so about 60% of TCRs with shared CDR3βs used the same V-J pair and shared the entire TCRβ chain (Figure 5B). As the HLAs that SP CD4 and SP CD8 cells are selected on are different, these results suggest that cross-reactive TCR β chains can be selected on different MHCs. CDR3βs that were shared in both SP CD4 and SP CD8 cells of each mouse in Experiment 2 had an average nucleotide-per-amino-acid sequence ratio of about 2, while the CDR3βs that were not shared had an average ratio close to 1, pointing to preferential selection of the shared CDR3βs in both T cell subsets (Table S11).

Shared CDR3βs are more likely to be cross-reactive than unshared sequences

The data above suggested that shared CDR3 sequences might be highly cross-reactive against disparate specificities. To address this possibility, we compared the repertoires of shared and unshared

CDR3βs from SP cell populations in both experiments to a list of cross-reactive CDR3βs defined by greater than two-fold frequency expansion in mixed lymphocyte reactions of a human peripheral blood sample against two different allogeneic donors sharing no HLA alleles. Among 100,112 and 29,033 alloreactive CDR3β sequences, 1,019 sequences expanded to both stimulators and were therefore identified as cross-reactive. Fisher's Exact test revealed a highly significant increase in the rate at which shared compared to unshared CDR3β sequences from Experiments 1, 2 and 3 were cross-reactive against two different sets of alloantigens (Figure 5C). Conversely, a highly significant increase was observed in the odds of allo-crossreactive sequences compared to alloreactive but non-cross-reactive sequences being shared between mice in Experiments 1 and 2 (Figure 5C). P-values by Fisher's Exact Test for the odds ratio of cross-reactivity in shared vs unshared sequences as well as for the odds ratio of sharing in cross-reactives vs allo-non-crossreactives are listed in Table S12. These data demonstrate that shared CDR3 sequences are more cross-reactive than unshared sequences.

Selection of autoreactive TCRs

In view of the evidence for cross-reactivity of shared sequences selected between disparate thymi and subsets, we hypothesized that shared sequences might be enriched for autoreactivity. We interrogated a previously-described list of 1,655 Type 1 diabetes (T1D)-associated autoreactive CDR3βs [20], along with some newer unique CDR3β amino acid sequences (total 2,208 sequences) associated with T1D, largely from peripheral blood, but also found in pancreas, LN and spleen of T1D donors from the network for Pancreatic Organ donors with Diabetes (nPOD) program [21]. These sequences were derived from a number of assays including sequencing of T cells following FACS-proliferation of dye-labeled responding T cells harvested following culture with autoantigens [22], direct MHC tetramer isolation of autoreactive T cells [22–25] or following the isolation and examination of peptide reactivities from islet infiltrating T cells [26]. T1D reactivity for these sequences was defined as reactivity to islet antigens such as GAD65 and insulin as described [21].

Comparison of these autoreactive TCRs to the TCR repertoires of grafted thymi in Experiment 2 revealed a significant increase in both cumulative frequency and clone fraction of T1D-associated sequences in SP-CD8 compared to DP CD69$^-$ populations (Figures 5E). Remarkably, the odds that a CDR3β shared between SP subsets of any two mice in Experiments 1, 2 or 3 was T1D-reactive was highly significantly greater than that for non-shared CDR3βs (Figure 5D), suggesting that shared CDR3s were enriched for autoreactivity. The P-values for the odds of T1D-reactivity in shared vs unshared sequences are listed in Table S12.

CDR3α and TCR sharing from single cell sequencing

To determine the extent to which CDR3β sharing was associated with sharing of the entire TCR, including the α chain, we performed single-cell TCR sequencing of thymic SP-CD4 cells from the same mice bulk-sequenced in Experiment 2 (except mouse 2autoA, due to technical failure). Comparing each pair of mice, the level of CDR3α sharing was significantly higher than CDR3β sharing (Figure 6A). However, the level of sharing for paired CDR3α-CDR3β was near-zero, and significantly lower than for either TCR chain on its own (Figure 6A), showing that the TCRs were almost always different among clones with a shared CDR3 α or β sequence. Consistent with the findings from bulk sequencing, the levels of shared CDR3s were not different between mice with allogeneic vs autologous thymi, either for TCRα, TCRβ or paired TCRα+TCRβ (Figure 6B). The number of unique CDRαs, CDR3βs and paired CDR3α-CDR3βs, the fraction of cells with a β chain that have at least one paired α chain or two paired α chains and the fraction of cells with an α chain that have a paired β chain is shown in Table S13.

Sub-sequence features are conserved in shared CDR3βs

Methods from Greiff et al. [27] which successfully distinguished between public and private antibody repertoires were applied to this dataset to see whether sub-sequence-level features can distinguish between shared and unshared sequences. This method uses normalized gapped k-mer (two sub-sequences of length k, separated by a gap of up to m amino acids) count as an input to a support vector machine (SVM) to see whether shared/unshared status can be predicted. Optimal parameters

determined by Greiff et al (k = 1, m = 1 and cost = 100) were used for SVM analysis and 10-fold cross validation was performed to assess performance of the classifier, using balanced accuracy (mean of sensitivity and specificity) as a performance metric. This was repeated on 100 length matched shared and unshared sequence datasets generated as described above. As shown in Figure S4A, these features can be used to predict shared or unshared status of sequences with a median balanced accuracy of ~62-78% for all cell subsets, where 50% would be equivalent to a random classifier. Frequency of gapped k-mers in shared sequences plotted against the frequency in unshared sequences further supported the hypothesis that there are sub-sequence features that are conserved in shared sequences (Figure S4B). We also found a notable enrichment in the "CASSL" motif at the 5' end of shared CDR3βs relative to unshared sequences, even in the unselected CD69- DP population (Figure S5), though that motif was highly represented in both shared and unshared sequences (Figure S5).

Evidence suggesting a role for self-peptides in human thymocyte selection

In pre-selection murine thymocytes, TCRβ CDR3 interfacial hydrophobicity at position 6 and position 7 (P6 and P7), the residues that interface with the peptide/MHC, correlated with the ability to be activated by self-peptide/MHC [28]. Stadinski et al. developed a self-reactivity index based on hydrophobicity of amino acids at CDR3β P6 and P7 and showed that this index correlates well with increased and decreased self-reactivity during positive and negative selection, respectively. We performed a similar analysis on human thymocyte and peripheral T cell subsets from Experiment 2 and 3, focusing on CDR3β lengths of greatest frequency in all thymocyte subsets (13 to 16 amino acids). For each mouse, P6 and P7 amino acid frequencies were analyzed in the thymic (DPCD69^{-}, DPCD69^{+}, SP-CD4^{+}, SP-CD8^{+} and Treg) and peripheral (P. CD4 and P. CD8) cell populations and frequencies were normalized within each cell population. For each animal, a fold-change in frequency of P6 and P7 amino acid residues between cell populations was recorded and these values were averaged across the mice within each grafted thymus group (i.e., Experiment 2 allogeneic vs. autologous thymus and Experiment

3). For Experiment 2 samples, we compared SP-CD4/SP-CD8/Tregs against DP CD69$^-$ thymocytes to evaluate the entire thymic selection process, DP CD69$^+$ thymocytes against DP CD69$^-$ thymocytes, and SP-CD4/SP-CD8/Tregs against DP CD69$^+$ thymocytes. For Experiment 3 samples, we compared P. CD4 and P. CD8 against SP-CD4 and SP-CD8, respectively. As shown for P6 in Figures 7A and S6A and P7 in Figures 7B and S6B, a trend to enrichment of hydrophobic amino acids (as defined in a reference table [29]) at both positions was observed as thymic selection progresses. Results of Spearman's non-parametric rank test for the mice in Experiments 2 and 3 are shown in Figure 7A and B. Statistically significant correlations between the fold changes of the amino acid residue at P6 (Figure 7A) or P7 (Figure 7B) and its hydrophobicity were observed during selection from the DP CD69$^-$ stage to the mature SP CD4, CD8 and Treg populations. Both autologous and allogeneic thymi show a similar trend of increasing hydrophobicity at P6/P7 as selection progresses (Figure S6). Positive selection from the CD69$^-$ to the CD69$^+$ DP stage was associated with significantly increased P6 hydrophobicity and overall selection from CD69$^-$ DP to both CD4 and CD8 SP populations was associated with significantly increased hydrophobicity at P6 and P7 and at P6 for the CD69- to Treg transition. The CD69$^+$ DP to SP transition was associated with significantly increased hydrophobicity only for SP-CD8 cells at P6 and SP-CD8 and Tregs at P7. Overall, increased hydrophobicity from DP CD69$^-$ to SP cells was more pronounced compared to transition from DP CD69$^+$ to SP cells. This trend was stopped or reversed in the transition from SP-CD4 and SP-CD8 cells to peripheral CD4 and CD8 cells, both at P6 and P7 (Figures 7A, 7B and S6). In sum, our data demonstrate an increase in hydrophobic amino acid usage at P6 and P7 in association with selection of human thymocytes (more associated with positive selection) and arrest or reversal of this trend in the transition from SP thymocytes to peripheral T cells, possibly in association with completion of negative selection.

Shared sequences might escape negative selection

To analyze differential usage of amino acids at each position as defined by the international ImMunoGeneTics (IMGT), Fisher's exact test was performed for all sequences in each of the 100 length-matched datasets of shared and unshared sequences. Differentially used amino acids were

plotted if Benjamini-Hochberg-adjusted p-value was less than 0.05 for Fisher's exact test to ensure that differences were significant (Figure 7C). Only amino acids showing up in at least 75 out of the 100 downsamples are annotated. There was a significant enrichment for the neutral amino acid G and hydrophilic amino acids (e.g. Q) and a significant decrease in hydrophobic amino acids (e.g. W) at P6 and P7 (equal to positions 109 and 110 on the plots, respectively) in shared sequences among CD69+ DP cells, most thymic SP populations and peripheral populations (Figure 7C). This pattern was not seen for shared sequences among CD69- DP cells. The reduced hydrophobicity at P6/P7 in shared sequences among selected but not unselected populations suggests that selected shared sequences may have weaker interactions with self-peptides than unshared sequences and that this may allow them to escape negative selection.

V and J gene usage

As shown in Figure S7, the pattern of V and J gene usage for the SP-CD4 cell population was very similar between the mice with autologous vs allogeneic thymus in Experiment 2, with no significant differences in V and J gene usage, arguing against a major role for selection in determining V and J gene usage. Overall, a similar pattern of V and J gene usage for SP-CD4 cells was detected when comparing mice in Experiments 1, 2 and 3, which received different HSCs as well as different thymi (Figure S7). There was also a similar pattern of V and J gene usage between thymic SP-CD4 and peripheral CD4 cells in Experiment 3 mice (Figure S7). Similar V and J gene usage patterns were also observed across all cell populations for the mice in Experiment 2 (Figure S8). Few statistically significant differences are shown in Figures S7 and S8. Thus, only small effects of genetic background and/or thymic selection were detected on the overall pattern of V and J gene usage. These findings were confirmed by single cell TCR sequencing data, which show a similar pattern of V and J gene usage for both α and β chains comparing SP-CD4 cells in mice with allogeneic vs autologous thymi in Experiment 2 (Figure S9).

Figure S10A shows plots of VJ usage for CD4 repertoires of mice 1autoA, 1autoB and 1autoC, which received the same thymus and HSCs. These representative plots show no disproportionately favored V-J pairing in the repertoire. Strong correlation was seen between observed VJ usage and VJ usage expected from stochastic combination of V genes with J genes according to the background frequency of each V and J (Figure S10B). Observed and expected VJ distributions were compared by Mann-Whitney U-Test, failing to reject the null hypothesis that VJ pairing is stochastic.

Discussion

We have demonstrated the impact of positive and negative selection in a human thymus on the human TCR repertoire. We have observed very high diversity among human thymocytes at all stages of development. This diversity is much greater than that for peripheral blood human T cells that include memory cell populations [18], consistent with studies demonstrating that TCR repertoires of human naïve T cells are much more diverse than those of memory T cells [3]. The observed lower diversity at the amino acid compared to the nucleotide sequence level, in productive compared to non-productive sequences, and in selected SP cell populations (SP-CD4, SP-CD8 and Tregs) compared to non-selected DP thymocytes demonstrates that thymic selection narrows the human TCR repertoire at both the DP and the SP stages. We demonstrated that diversity further decreases in peripheral CD4 and CD8 cells in humanized mice.

The diversity of the SP-CD4 and SP-CD8 TCR repertoires in autologous and allogeneic grafted thymi was greater than that of the original fetal thymus used for construction of the humanized mice, reflecting the immaturity of the repertoire in the 17 gestational-week fetus used, and confirming that the thymocytes developing in our studies arose de novo from engrafted stem cells due to the success of our procedures for purging the graft of pre-existing thymocytes prior to implantation. The TCR repertoire of mouse neonates is much narrower than that of adult mice until day 4-5 after birth, when the TdT enzyme is activated [19]. Limited studies indicated that $CD45^+$ cells enter the human fetal thymus at 8 weeks of gestation and all steps of TCR development are detectable by week 16 [30], but TdT was

reportedly undetectable in fetal human thymus [31], consistent with our observation of very low TdT expression in DP cells of such thymi. In contrast, we detected much higher TdT levels in postnatal human thymi and similar levels in grafted human thymi, thus explaining the greater TCR diversity in grafted human thymi compared to fetal thymus donors. While grafted and human post-natal thymi showed very similar structures and cell populations, we cannot rule out the possibility that undetected structural differences might influence selection.

High divergence of TCRβ repertoires in thymocytes generated in identical thymi from the same HSCs at the same time is consistent with the stochastic nature of TCR repertoire formation and may contribute to the incomplete penetrance of genetically-controlled autoimmune diseases in identical twins [32]. Increased CDR3β sharing among productive compared to non-productive TCR rearrangements, and in selected vs unselected thymocyte populations along with high convergence of nucleotide sequences at the amino acid level demonstrate that thymic selection favors these shared TCRβ CDR3s. While many of these shared CDR3βs used different V genes but were associated with identical J genes, 20-25% also had identical V genes, indicating that the same entire β chain was used. Single cell analysis revealed that TCRs with the same β chain are almost always paired with different α chains.

Our observation of shared sequences between different donors and thymi is consistent with studies in mice showing the existence of MHC-independent public, abundant CDR3 sequences with convergent recombination, also with fewer N insertions than average, among peripheral T cells [6]. These authors performed numerical simulations of TCR rearrangements to demonstrate that biases in TCR recombination could only partially explain the observation. Our study is the first to specifically analyze the impact of selection on shared human thymocyte sequences, and demonstrates that thymic selection is a significant factor driving their abundance among mature T cells. Previous next-generation sequencing studies of the peripheral blood TCR repertoire of monozygotic twins revealed increased shared CDR3βs among highly-abundant clonotypes and similar overall overlap between the repertoires

of monozygotic twins and unrelated individuals [33], consistent with our observation of similar sharing between mice with the same HSCs but different thymi and between mice with different thymi and HSCs. The increased TCRβ sharing observed by Qi et al in peripheral blood of twins compared to unrelated individuals was more pronounced in memory than in naive T cells [34], and may be explicable by post-thymic selection events rather than thymic selection.

Several possibilities could explain our finding of similar TCR sharing among thymocytes developing in autologous vs allogeneic thymi. First, the allogeneic and autologous thymi may share common peptide motifs that contribute to the selection of shared CDR3βs in the context of different HLAs, as previously suggested [5]. According to a model based on sequencing of adult peripheral blood CD8 T cells, CDR3β sharing was observed at a >500-fold greater rate than expected from stochastic rearrangements and was independent of HLA sharing [5]. Consistent with our observation of shared sequences among thymocyte subsets, different peripheral blood T cells subsets revealed significant CDR3β sharing [35]. While peripheral blood T cell studies may reflect post-thymic antigen-driven expansion, this had not occurred in the thymic grafts in our studies.

We favor a second possible explanation for inter-individual and inter-T cell subset CDR3β sharing, namely that the shared CDR3βs are more highly cross-reactive than non-shared CDR3βs, and can therefore be positively selected by diverse HLA/peptide complexes. In all three experiments, we observed increased overlap of more abundant CDR3βs. This may reflect the greater likelihood of detecting more abundant sequences in a given sample size. However, our observation of increased CDR3β sharing following positive selection of DP thymocytes suggests that they are preferentially selected. We also observed an increase in P6-P7 hydrophobicity during positive selection, demonstrating the role of HLA-peptide complexes in this process. CDR1 and CDR2 interact prominently with MHC to influence CD4 and CD8 T cell development, whereas the CDR3β predominantly interacts with peptides bound to MHC [36], which differ for recognition by SP-CD4 and SP-CD8 cells. Thus, the hypothesis that most of the shared CDR3βs are cross-reactive seems most probable. TCRs have

flexible CDR3 loops [37] and can bind to different peptide-MHC complexes through different mechanisms, including binding with an altered angle or register [38]. For example, a human, preproinsulin-reactive CD8 T cell clone (1E6) is capable of recognizing more than 1 million peptides in the context of a single MHC I [39], including microbial peptides as well as self peptides [40]. Such cross-reactivity is necessary for the limited number of TCRs in the body to recognize more than 10^{15} different peptide-MHC complexes [40]. In mice, TCRs that survive thymic selection are enriched for cross-reactive TCRs compared to those that do not. Interestingly, cross-reactive TCRs were frequently reactive to both MHC I and II [41]. Furthermore, in mice genetically engineered to express human TCR genes with either a single allele of human MHCII or a single mouse MHCII, a surprisingly high number of CD4 TCRs was shared between the two groups [42]. The finding that single MHCs from different species are able to select shared TCRs provides support for our hypothesis about selection of cross-reactive TCRs.

The shared sequences in our study had shorter CDR3β length than non-shared sequences, consistent with previous studies [5,33], as a result of lower numbers of N insertions. Since more abundant CDR3βs were, on average, longer than the less frequent ones, even in the pre-selection DP CD69$^-$ population, the increased sharing among shorter sequences does not simply indicate an inherent bias for β-selection of shorter CDR3β's. We found that the average length of more abundant CDR3βs decreases during thymic selection, consistent with previous reports in human and mouse [43–46]. Shorter CDR3βs may be positively selected more easily due to increased low affinity cross-reactions with diverse MHC-peptide complexes. TCRs in mice deficient for TdT, which mediates N insertion in CDR3s, have a shorter average CDR3 length [8] and demonstrate increased inter-animal sharing of TCR sequences with increased cross-reactivity for different peptides compared to wild-type mice [47]. We observed increased usage of neutral and hydrophilic amino acids (G and Q specifically) among shared sequences at P6 and P7 (positions 109 and 110 in Figure 7B) compared to unshared sequences following the transition from DP CD69$^-$ to DP CD69$^+$. This difference was maintained in the single-positive populations, despite the overall increase in hydrophobicity with selection at these positions in the total repertoire. In view of murine studies indicating that hydrophobicity at P6 and P7 of CDR3β

correlates with reactivity to self peptide/MHC complexes [28], these data suggest that shared sequences may be preferentially positively selected and survive negative selection because of their low affinity for self peptide/MHC. Consistently, shared sequences in mice had reduced affinity for self peptide/MHC [27]. Thus, while the low affinity of shared sequences for self peptide/HLA is sufficient to promote positive selection, the low level of hydrophobicity at the peptide binding site may allow them to preferentially evade negative selection in the thymus. The observed average CDR3β shortening as negative selection progresses (CD69+ DP to SP transition) among the most abundant but not among the least abundant sequences is consistent with the avoidance of negative selection by these shorter, more commonly shared, CDR3βs. While one SP thymocyte subset (CD8 SPs) showed an unexpected increase in P6/P7 hydrophobicity during the transition from the CD69+ DP phase, more consistent increases were observed with overall selection between the CD69- DP and SP stage for each subset, suggesting that the major increase may reflect positive selection. The CD69+ DP population likely had already undergone considerable negative selection, diluting the increased hydrophobicity associated with positive selection, and negative selection likely continued into the SP phase, making comparisons between these populations complex to interpret.

Our single-cell TCR sequencing data showed even greater TCRα sharing than TCRβ sharing between animals and also demonstrated that almost all TCRs with shared α or β chains do not form the same αβ pairs. This result implies that the selective pressure for shared CDR3βs is applied separately to the β chain. Since we did not perform bulk analysis of CDR3α at various stages of thymocyte selection, we do not know whether the high level of CDR3α sharing reflects generational bias or selective pressure. Consistent with our results, greater TCRα than TCRβ sharing has been reported among healthy humans, while no difference was found in diversity of CDR3α compared to CDR3β [48]. Structural analyses have shown that both CDR3α and CDR3β interact with peptides and contribute to recognition of peptide/MHC complexes [49]. However, a predominance of charged and polar amino acids in CDR3α compared to CDR3β may impact peptide interactions [50].

Comparison of shared and unshared sequences in our study with known autoreactive and cross-reactive TCRs supports our hypothesis that shared CDR3s are cross-reactive. Cross-reactive CDR3βs identified through sequencing of healthy donor T cells responding to disparate alloantigens were strikingly enriched for shared compared to unshared sequences identified in thymi in both experiments. Moreover, the frequency of known T1D-associated CDR3βs was increased during positive selection in the grafted thymi and these sequences were highly significantly increased among shared CDR3s. These data suggest that autoreactive T1D-associated sequences are enriched for cross-reactive TCRs that can be selected by different thymi with different MHCs

Despite the very divergent CDR3β TCR repertoires, V and J gene was overall very similar among different mice generated with different HSCs and different thymi. There is controversy in the literature about the role of genetic background and thymic selection in determining V and J gene usage [33,51–54]. Our analysis of thymocyte subsets generated from different donors demonstrates only a minor role for thymic selection in determining overall V and J gene usage. Intrinsic genomic factors such as promoter strength or differential structural features are likely the main determinants of V and J gene usage in TCRs. However, the small number of HSC donors in our study does not permit detection of more subtle effects of genetic background and different HLA alleles on V and J gene usage, as has been reported [55,56]. A report of increased similarity in V and J gene usage of peripheral naive T cells between identical twins than non-twins but overall similar V and J usage frequencies of any two donors, regardless of relatedness [56] is consistent with a subtler effect of HLA alleles in selecting for certain V and J genes. In one mouse study [42], selection of T cells on a single murine or human MHC II molecule resulted in some differences in V and J gene usage. However, the proportion of SP thymocytes in these mice was significantly lower compared to that in normal B6 mice, suggesting inefficient selection on a single MHC allele, which could potentially skew the repertoire toward using certain V and J genes. Our study shows the thymic MHC does not markedly influence human V and J gene when there is a full complement of HLA molecules.

Interestingly, sequences shared between mice were enriched across all cell populations for the 5' "CASSL" motif in CDR3β within the V-region of CDR3β. This motif is found in the 3' tail of TRB5-01, which is consistently the most highly-expressed V gene across all samples, at roughly 12%. Though it is enriched in shared sequences, the motif has high representation in both shared and unshared sequences sets. Published structural analysis of TCR binding to MHC-peptide complex for a clone with this motif shows no direct interaction of the motif with peptide-MHC. Our observation of enrichment for this motif in shared sequences of pre-selection $CD69^-$ DP thymocytes suggests a more structural role, such as increased binding to pre-TCRα, rather than a role in self-antigen-driven selection. After successful rearrangement of the TCR-β chain, developing thymocytes undergo an estimated six to seven rounds of cell division before independent rearrangements of TCRα [57,58], which could enrich such sequences in preselection thymocytes. Increased stability of TCR tertiary structure or increased binding to TCRα [59] could also explain this enrichment in shared sequences of preselection thymocytes. Our detection of fewer unique CDR3α than CDR3β sequences in the single cell analysis was surprising in light of the ability of early DP thymocytes to expand during beta selection and the ability of individual T cell clones to rearrange only one β chain and two α chains. This result likely reflects reduced efficiency in detecting α chains than β chains with the sequencing platform used, as less than 100% of cells with β chains sequenced had a sequenced α chain.

Besides V and J gene usage, some reports indicate that the frequency of particular TCR Vβ-TCR Jβ recombinations in human lymphocytes is controlled genetically [60]. By comparing the observed VJ frequency distributions in productive and nonproductive repertoires to VJ frequency distribution expected from a stochastic combination of V genes with J genes according to their background frequency in thymic TCR repertoires, our results show that VJ pairing is only determined by the amount of expression of each gene, with no evidence for preferential V-J recombinations.

In conclusion, our data indicate that human thymus in humanized mice selects a very diverse TCR repertoire. Formation of the human TCR repertoire is largely stochastic and can be almost totally

divergent in thymi of animals with identical hematopoietic stem cells, thymus, genetic background and environment. However, we show that thymic selection increases the overlap between human TCR repertoires and recognition of self HLA-peptide plays a role in human thymocyte selection. The overlap of CDR3βs in disparate thymi and different thymocyte subsets, shorter CDR3β lengths of shared sequences, direct evidence for cross-alloreactivity and autoreactivity of shared sequences, and analysis of amino acid usage in these CDR3 sequences are consistent with the interpretation that shared sequences are preferentially positively selected due to high cross-reactivity and evade negative selection due to low affinity for self peptides. Thus, we have used the humanized mouse model in our studies to obtain novel insights into the factors determining human T cell repertoire formation.

Methods

Generation of humanized mice

Three, six and two humanized mice were generated by combined human fetal thymic and hematopoietic cell transplantation in thymectomized immunodeficient mice for Experiments 1, 2 and 3, respectively (Figure 1A-C; details in supplemental materials).

FACS sorting of different subsets of grafted thymus and peripheral cells

At weeks 14, 20 and 22 after thymus transplantation, mice from Experiments 1, 2 and 3, respectively, were euthanized. Grafted thymi (for mice in all experiments) and spleen and lymph nodes (LNs) (only mice in Experiment 3) were harvested and the thymocytes and pooled spleen and LN cells were isolated and FACS sorted (Figure S2; see supplemental materials for details of FACS sorting). HLA typing of fetal tissues used to generate humanized mice in all three experiments are shown in Table S2.

DNA isolation and high throughput CDR3β TCR sequencing

Genomic DNA was isolated from sorted cell populations using the Qiagen DNeasy Blood and Tissue Kit. DNA was frozen at −20°C and shipped on dry ice to Adaptive Biotechnologies for high-throughput TCRB CDR3 sequencing. The TCR sequencing data were retrieved from Adaptive's ImmunoSEQ software. The Adaptive raw sequencing data of all samples are available at https://github.com/Aleksobrad/Humanized-Mouse-Data.

Single cell TCR sequencing

Single cell TCR sequencing was performed using the 10X Genomics platform as detailed in the supplemental materials.

Computational and statistical analysis

These methods are described in supplemental materials.

Cross-reactive TCR list

Peripheral blood mononuclear cells (PBMCs) from a human subject were stained with carboxyfluorescein succinimidyl ester (CFSE) and co-cultured separately with two fully HLA-mismatched irradiated PBMCs, as stimulator cells, for six days. Dividing CD4 and CD8 T cells (CFSE-low) and unstimulated CD4 and CD8 T cells from the same donor were FACS sorted and after DNA isolation their TCRβ was sequenced as described above. Alloreactive sequences were defined as CDR3β amino acid sequences that were expanded at least two-fold by frequency from unstimulated to stimulated samples, above a minimum frequency in the stimulated samples of $1*10^{-5}$, as described previously [18,61]. Alloreactivity was determined separately for CD4 and CD8 samples, and alloreactive CDR3βs responding to both stimulators were considered cross-reactive, since they responded to different stimulators with no MHC sharing. We compared the repertoires of T cell populations of grafted thymi in humanized mice to the list of cross-reactive TCRs to investigate the selection of cross-reactive TCRs in human thymus, calculating the odds ratio of allo-cross-reactivity in shared vs unshared clones as well as the odds ratio of shared clone status in allo-cross-reactive vs allo-non-cross-reactive clones, with significance assessed by Fisher's exact test. Lists of the CDR3 sequences determined to be cross-reactive and allo-non-cross-reactive are available as .csv files at https://github.com/Aleksobrad/Humanized-Mouse-Data along with the raw Adaptive sequencing data from healthy control MLRs of the same responders against two fully-HLA-mismatched stimulators.

T1D-reactive TCR list

Our T1D-reactive TCR data set contains 2208 unique CDR3β amino acid sequences associated with Type 1 diabetes derived from peripheral blood, pancreas, LN and spleen of T1D donors from the network for Pancreatic Organ donors with Diabetes (nPOD) program [20]. These sequences were derived from a number assays including sequencing of T cells following FACS-proliferation of dye-labeled responding T cells harvested in response to culture with autoantigens [21], direct MHC tetramer isolation of autoreactive T cells [21–24] , or in certain situations following the isolation and examination of peptide reactivities from islet infiltrating T cells [25]. T1D reactivity for these sequences was defined as reactivity to islet antigens such as GAD65 and insulin as described [26]. A single prominent T1D-reactive nucleotide

sequence [encoding CASSSFWGSDTGELFF TCRBV11-02 TCRBJ02-02] present in bulk-sequencing data but missing from single-cell data was removed from Experiment 2 due to suspected contamination from a vector with the same sequence that was present in the lab. We compared the repertoires of T cell populations of grafted thymi in humanized mice to this list of T1D-reactive sequences to investigate the selection of T1D-reactive TCRs in human thymus, calculating the odds ratio of T1D-reactivity among shared vs unshared sequences, with significance assessed by Fisher's exact test. The .csv files with the CDR3 sequences of T1D-reactive sequences are available at https://github.com/Aleksobrad/Humanized-Mouse-Data.

Statistics.

For comparisons between different cell populations considering clonality, JSD and fraction of shared sequences, paired t-tests with Bonferonni multiple-testing correction were used. For studies that involved calculation of odds ratio, Fisher's exact test was performed. The Spearman correlation coefficient R value and p-value from the nonparametric Spearman correlation test were reported where fold changes in the relative amino acid frequencies are plotted against amino acid hydrophobicity based on Gibbs Free Energy. Unpaired t-test with Bonferroni correction for multiple testing was performed to compare the Vβ and Jβ gene usages between different groups for each gene. For comparing the observed and expected frequencies of VJ gene pairs, Mann-Whitney U-Tests were performed with the null hypothesis that there is no difference between the distribution of observed frequencies and the distribution of frequencies expected from random VJ combination. In all statistical tests, corrected P-values <0.05 were considered significant.

Study approval.

Protocols involving the use of human tissues and animals were approved by the Institutional Review Board and the Institutional Animal Care and Use Committee of Columbia University (New York, NY), and all of the experiments were performed in accordance with the protocols.

Author Contributions

Conceptualization, M.K.M. and M.S.; Methodology, M.K.M., A.O. and M.S.; Software/Analysis, A.O., K.M., H.R.S., and A.M.; Validation, R.W., Y.S., T.M.B. and M.S.; Investigation, M.K.M., A.O., M.H., A.M., G.N., N.D., H.L., and S.H.; Writing – Original Draft, M.K.M., A.O., A.M. and M.S.; Writing – Review & Editing, M.K.M., A.O., M.H., R.W., Y.S., H.R.S., T.M.B. and M.S.; Funding Acquisition, M.S.; Supervision, M.S.

Acknowledgements

Research reported in this publication was supported by the following grants: P01 AI04589716 and R01DK103585 (M.S.), P01 AI42288 and DK106191 (T.M.B.). Funding was provided by the Human Islet Research Network (HIRN) Opportunity Pool Fund, RRID:SCR_014393; https://hirnetwork.org ; U01 DK104162 to M.S. and T.M.B. This research was performed using resources and/or funding provided by the NIDDK-supported Human Islet Research Network (HIRN, RRID: SCR_014393; https://hirnetwork.org; CMAI UC4 DK104207 to M.S. and DK104194 to T.M.B). Research was performed in the CCTI Flow Cytometry Core, supported in part by the Office of the Director, National Institutes of Health under awards S10OD020056, S10RR027050, P30CA013696, 5P30DK063608 and R01DK106436. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. MKM was supported by a Friedman Award from the University of British Columbia (Canada) and also an American Diabetes Association (ADA) Postdoctoral Fellowship. We thank Drs. Arup Chakraborty and Peter Sims for helpful comments on the manuscript and Ms. Nicole Casio for assistance with the submission.

References

1. Arstila TP, et al. A direct estimate of the human alphabeta T cell receptor diversity. *Science*. 1999;286(5441):958-961.

2. Robins HS, et al. Comprehensive assessment of T-cell receptor β-chain diversity in αβ T cells. *Blood*. 2009;114(19):4099-4107.

3. Qi Q, et al. Diversity and clonal selection in the human T-cell repertoire. *Proc Natl Acad Sci U S A*. 2014;111(36):13139-13144.

4. Vanhanen R, et al. T cell receptor diversity in the human thymus. *Mol Immunol*. 2016;76:116-122.

5. Robins HS, et al. Overlap and effective size of the human CD8+ T-cell receptor repertoire. *Sci Transl Med*. 2010; 2(47):47ra64.

6. Madi A, et al. T-cell receptor repertoires share a restricted set of public and abundant CDR3 sequences that are associated with self-related immunity. *Genome Res*. 2014;24(10):1603-1612.

7. Clambey ET, et al. Molecules in medicine mini review: the αβ T cell receptor. *J Mol Med*. 2014;92(7):735-741.

8. Cabaniols JP, et al. Most alpha/beta T cell receptor diversity is due to terminal deoxynucleotidyl transferase. *J Exp Med*. 2001;194(9):1385-1390.

9. Klein L, et al. Positive and negative selection of the T cell repertoire: what thymocytes see (and don't see). *Nat Rev Immunol*. 2014;14(6):377-391.

10. Wiegers GJ, et al. Shaping the T-cell repertoire: a matter of life and death. *Immunol Cell Biol*. 2011;89(1):33-39.

11. Allen S, et al. Shaping the T-cell repertoire in the periphery. *Immunol Cell Biol*. 2011;89(1):60-69.

12. Pham HP, et al. Half of the T-cell repertoire combinatorial diversity is genetically determined in humans and humanized mice. *Eur J Immunol*. 2012;42(3):760-770.

13. Marodon G, et al. High diversity of the immune repertoire in humanized NOD.SCID.gamma c-/- mice. *Eur J Immunol*. 2009;39(8):2136-2145.

14. Vandekerckhove BA, et al. Thymic Selection of the Human T Cell Receptor V Beta Repertoire in SC1D-hu Mice. *J Exp Med*. 1992;176(6):1619-24.

15. Shimizu I, et al. Comparison of human T cell repertoire generated in xenogeneic porcine and human thymus grafts. *Transplantation*. 2008;86:601-610.

16. Kalscheuer H, et al. A Model for Personalized in Vivo Analysis of Human Immune Responsiveness. *Sci Transl Med*. 2012;4(125):125ra30.

17. Mitchell JL, et al. Ikaros, Helios, and Aiolos protein levels increase in human thymocytes after β selection. *Immunol Res*. 2016;64(2):565-575.

18. Morris H, et al. Tracking donor-reactive T cells: Evidence for clonal deletion in tolerant kidney transplant patients. *Sci Transl Med*. 2015;7(272):272ra10.

19. Bogue M, et al. A special repertoire of alpha:beta T cells in neonatal mice. *EMBO J*. 1991;10(12):3647-3654.

20. Jacobsen LM, et al. T Cell Receptor Profiling in Type 1 Diabetes. *Curr Diab Rep*. 2017;17(11):118.

21. Eugster A, et al. High Diversity in the TCR Repertoire of GAD65 Autoantigen-Specific Human CD4 $^{+}$ T Cells. *J Immunol*. 2015;194(6):2531-2538.

22. Yang J, et al. Antigen-Specific T Cell Analysis Reveals That Active Immune Responses to β Cell Antigens Are Focused on a Unique Set of Epitopes. *J Immunol*. 2017;199(1):91-96.

23. Michels AW, et al. Islet-derived CD4 T cells targeting proinsulin in human autoimmune diabetes. *Diabetes*. 2017;66(3):722-734.

24. Bulek AM, et al. Structural basis of human β -cell killing by CD8+ T cells in Type 1 diabetes. *Nat Immunol*. 2012;13(3):283-289.

25. Kent SC, et al. Expanded T cells from pancreatic lymph nodes of type 1 diabetic subjects recognize an insulin epitope. *Nature*. 2005;435(7039):224-228.

26. Seay HR, et al. Tissue distribution and clonal diversity of the T and B cell repertoire in type 1 diabetes. *JCI Insight*. 2016;1(20):e88242.

27. Greiff V, et al. Learning the High-Dimensional Immunogenomic Features That Predict Public and Private Antibody Repertoires. *J Immunol*. 2017;199(8):2985-2997.

28. Stadinski BD, et al. Hydrophobic CDR3 residues promote the development of self-reactive T cells. *Nat Immunol*. 2016;17(8):946-55.

29. White SH, Wimley WC. MEMBRANE PROTEIN FOLDING AND STABILITY: Physical Principles. *Annu Rev Biophys Biomol Struct*. 1999;28(1):319-365.

30. Haynes BF, Heinly CS. Early human T cell development: analysis of the human thymus at the time of initial entry of hematopoietic stem cells into the fetal thymic microenvironment. *J Exp Med*. 1995;181(4):1445-1458.

31. Deibel MR, et al. Expression of terminal deoxynucleotidyl transferase in human thymus during ontogeny and development. *J Immunol*. 1983;131(1):195-200.

32. Redondo MJ, et al. Heterogeneity of Type I diabetes: Analysis of monozygotic twins in Great Britain and the United States. *Diabetologia*. 2001;44(3):354-362.

33. Zvyagin I V., et al. Distinctive properties of identical twins' TCR repertoires revealed by high-throughput sequencing. *Proc Natl Acad Sci*. 2014;111:5980-5985.

34. Qi Q, et al. Diversification of the antigen-specific T cell receptor repertoire after varicella zoster vaccination. *Sci Transl Med*. 2016;8(332):332ra46.

35. Wang C, et al. High throughput sequencing reveals a complex pattern of dynamic

interrelationships among human T cell subsets. *Proc Natl Acad Sci*. 2010;107(4):1518-1523.

36. Petrova G, et al. Cross-reactivity of T cells and its role in the immune system. *Crit Rev Immunol*. 2012;32(4):349-372.

37. Reiser J, et al. CDR3 loop flexibility contributes to the degeneracy of TCR recognition. *Nat Immunol*. 2003;4(3):241-247.

38. Sewell AK. Why must T cells be cross-reactive ? *Nat Publ Gr*. 2012;12(9):669-677.

39. Wooldridge L, et al. A single autoimmune T cell receptor recognizes more than a million different peptides. *J Biol Chem*. 2012;287(2):1168-1177.

40. Cole DK, et al. Hotspot autoimmune T cell receptor binding underlies pathogen and insulin peptide cross-reactivity. *J Clin Invest*. 2016;126(6):2191-2204.

41. McDonald BD, et al. Crossreactive αβ T Cell Receptors Are the Predominant Targets of Thymocyte Negative Selection. *Immunity*. 2015;43(5):859-869.

42. Chen X, et al. Human TCR-MHC coevolution after divergence from mice includes increased nontemplate-encoded CDR3 diversity. *J Exp Med*. 2017;214(11):3417-3433.

43. Nishio J, et al. Development of TCRB CDR3 length repertoire of human T lymphocytes. *Int Immunol*. 2004;16(3):423-431.

44. Yassai M, Gorski J. Thymocyte maturation: selection for in-frame TCR alpha-chain rearrangement is followed by selection for shorter TCR beta-chain complementarity-determining region 3. *J Immunol*. 2000;165(7):3706-3712.

45. Matsutani T, et al. Shortening of complementarity determining region 3 of the T cell receptor α chain during thymocyte development. *Mol Immunol*. 2011;48(4):623-629.

46. Matsutani T, et al. Comparison of CDR3 length among thymocyte subpopulations: Impacts of MHC and BV segment on the CDR3 shortening. *Mol Immunol*. 2007;44(9):2378-2387.

47. Gavin MA, Bevan MJ. Increased peptide promiscuity provides a rationale for the lack of N

regions in the neonatal T cell repertoire. *Immunity*. 1995;3(6):793-800.

48. Kitaura K, et al. A new high-throughput sequencing method for determining diversity and similarity of T cell receptor (TCR) α and β repertoires and identifying potential new invariant TCR α chains. *BMC Immunol*. 2016; 17(1):38.

49. Adams JJ, et al. T cell receptor signaling is limited by docking geometry to peptide-major histocompatibility complex. *Immunity*. 2011;35(5):681-93.

50. Moss PAH, Bell JI. Sequence analysis of the human αβ T-cell receptor CDR3 region. *Immunogenetics*. 1995;42(1):10-8.

51. Hawes GE, et al. Differential usage of T cell receptor V gene segments in CD4+ and CD8+ subsets of T lymphocytes in monozygotic twins. *J Immunol*. 1993;150:2033-2045.

52. Malhotra U, et al. Variability in T cell receptor V beta gene usage in human peripheral blood lymphocytes. Studies of identical twins, siblings, and insulin-dependent diabetes mellitus patients. *J Immunol*. 1992;149(5):1802-1808.

53. Kohsaka H, et al. The expressed T cell receptor V gene repertoire of rheumatoid arthritis monozygotic twins: rapid analysis by anchored polymerase chain reaction and enzyme-linked immunosorbent assay. *Eur J Immunol*. 1993;23(8):1895-1901.

54. Reinhardt C, Melms A. Skewed TCRV beta repertoire in human thymus persists after thymic emigration: influence of genomic imposition, thymic maturation and environmental challenge on human TCRV beta usage in vivo. *Immunobiology*. 1998;199(1):74-86.

55. Sharon E, et al. Genetic variation in MHC proteins is associated with T cell receptor expression biases. *Nat Genet*. 2016;48(9):995-1002.

56. Rubelt F, et al. Individual heritable differences result in unique cell lymphocyte receptor repertoires of naïve and antigen-experienced cells. *Nat Commun*. 2016;7:11112.

57. Hamrouni A, et al. T cell receptor gene rearrangement lineage analysis reveals clues for the

origin of highly restricted antigen-specific repertoires. *J Exp Med*. 2003;197(5):601-614.

58. Dudley EC, et al. T cell receptor beta chain gene rearrangement and selection during thymocyte development in adult mice. *Immunity*. 1994;1(2):83-93.

59. Bartok I, et al. T cell receptor CDR3 loops influence alphabeta pairing. *Mol Immunol*. 2010;47(7-8):1613-1618.

60. Nanki T, et al. Genetic control of T cell receptor BJ gene expression in peripheral lymphocytes of normal and rheumatoid arthritis monozygotic twins. *J Clin Invest*. 1996;98(7):1594-1601.

61. DeWolf, S, et al. Quantifying size and diversity of the human T cell alloresponse. *JCI Insight*. 2018;3(15).

Figures

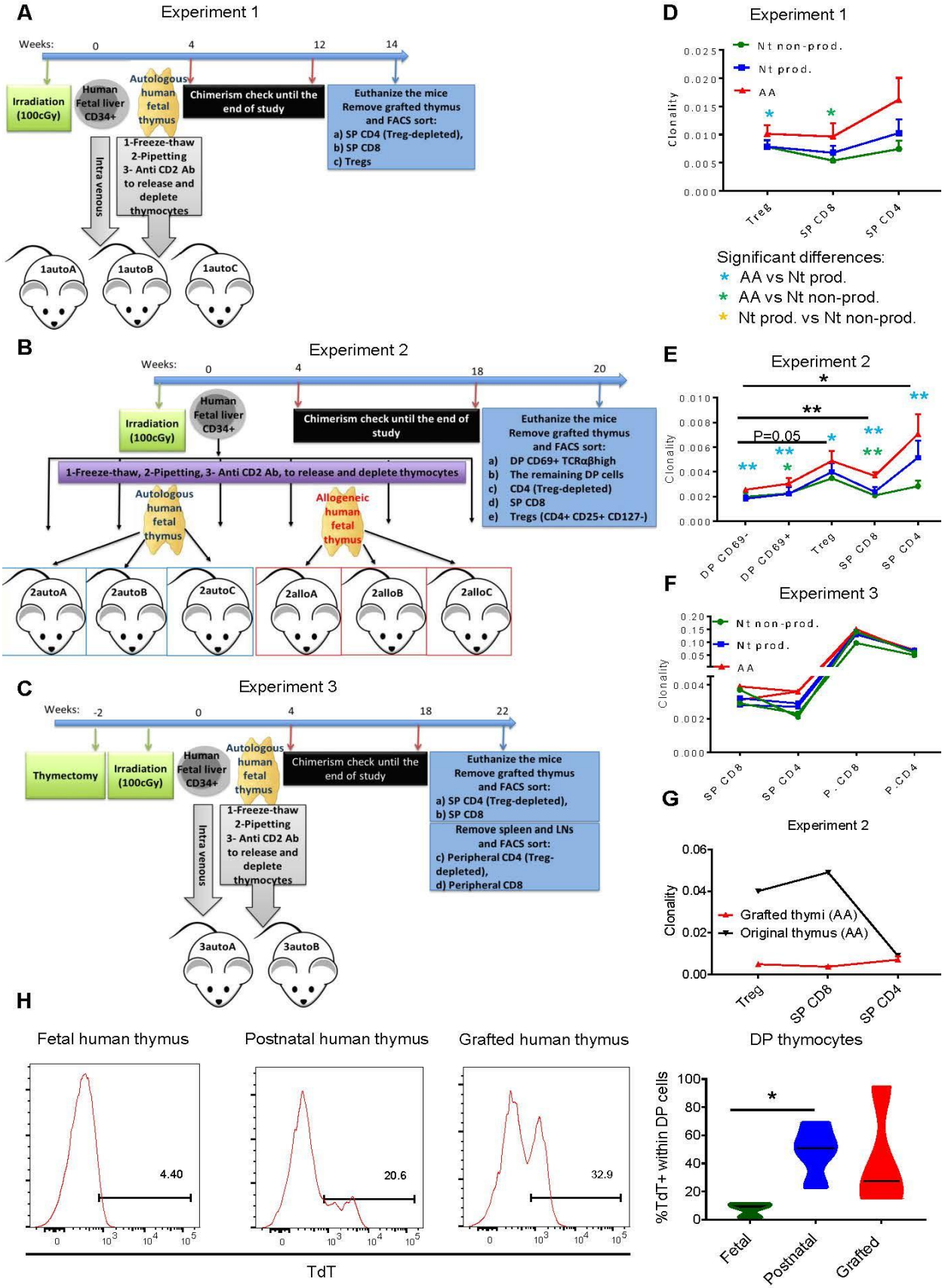
A
Experiment 1
Weeks: 0 4 12 14
Irradiation (100cGy)
Human Fetal liver CD34+
Autologous human fetal thymus
Chimerism check until the end of study
Euthanize the mice
Remove grafted thymus and FACS sort:
a) SP CD4 (Treg-depleted),
b) SP CD8
c) Tregs
Intra venous
1-Freeze-thaw
2-Pipetting
3- Anti CD2 Ab to release and deplete thymocytes
1autoA
1autoB
1autoC
B
Experiment 2
Weeks: 0 4 18 20
Irradiation (100cGy)
Human Fetal liver CD34+
Chimerism check until the end of study
Euthanize the mice
Remove grafted thymus and FACS sort:
a) DP CD69+ TCRαβhigh
b) The remaining DP cells
c) CD4 (Treg-depleted)
d) SP CD8
e) Tregs (CD4+ CD25+ CD127-)
1-Freeze-thaw, 2-Pipetting, 3- Anti CD2 Ab, to release and deplete thymocytes
Autologous human fetal thymus
Allogeneic human fetal thymus
2autoA
2autoB
2autoC
2alloA
2alloB
2alloC
C
Experiment 3
Weeks: -2 0 4 18 22
Thymectomy
Irradiation (100cGy)
Human Fetal liver CD34+
Autologous human fetal thymus
Chimerism check until the end of study
Euthanize the mice
Remove grafted thymus and FACS sort:
a) SP CD4 (Treg-depleted),
b) SP CD8
Remove spleen and LNs and FACS sort:
c) Peripheral CD4 (Treg-depleted),
d) Peripheral CD8
Intra venous
1-Freeze-thaw
2-Pipetting
3- Anti CD2 Ab to release and deplete thymocytes
3autoA
3autoB
D
Experiment 1
Nt non-prod.
Nt prod.
AA
Clonality
Treg
SP CD8
SP CD4
Significant differences:
AA vs Nt prod.
AA vs Nt non-prod.
Nt prod. vs Nt non-prod.
E
Experiment 2
P=0.05
Clonality
DP CD69-
DP CD69+
Treg
SP CD8
SP CD4
F
Experiment 3
Nt non-prod.
Nt prod.
AA
Clonality
SP CD8
SP CD4
P.CD8
P.CD4
G
Experiment 2
Grafted thymi (AA)
Original thymus (AA)
Clonality
Treg
SP CD8
SP CD4
H
Fetal human thymus
Postnatal human thymus
Grafted human thymus
4.40
20.6
32.9
TdT
DP thymocytes
%TdT+ within DP cells
Fetal
Postnatal
Grafted

Figure 1. Experimental design and clonality scores. A, B and C: Construction of humanized mice for Experiments 1, 2 and 3. Cell populations were sorted for sequencing at 14, 20 and 22 weeks post-transplantation, respectively. D, E and F: clonality scores for cell populations of Experiments 1, 2 and 3 at the nucleotide/nonproductive (Nt. non-prod.), Nt/productive (Nt. prod.) and amino acid (AA) levels (means ±SEM except for Experiment 3, which shows individual animals). Paired t-tests compare clonality of each sequence set within each cell population. Paired t-tests with Bonferonni multiple-testing correction were performed to compare different cell populations in Experiment 2. * 0.01<p-value<0.05, **0.001<p-value<0.01, ***p-value<0.001. G: AA clonality scores of grafted thymi and the original autologous thymus in Experiment 2. H: expression of TdT in DP thymocytes of fetal (n=3, gestational ages of 17, 20 and 21 weeks), postnatal (n=4, age 4 months, 6 months, 13 years and 17 years) and grafted human thymi in humanized mice (n=3, at 18, 26 and 33 weeks post-transplantation).

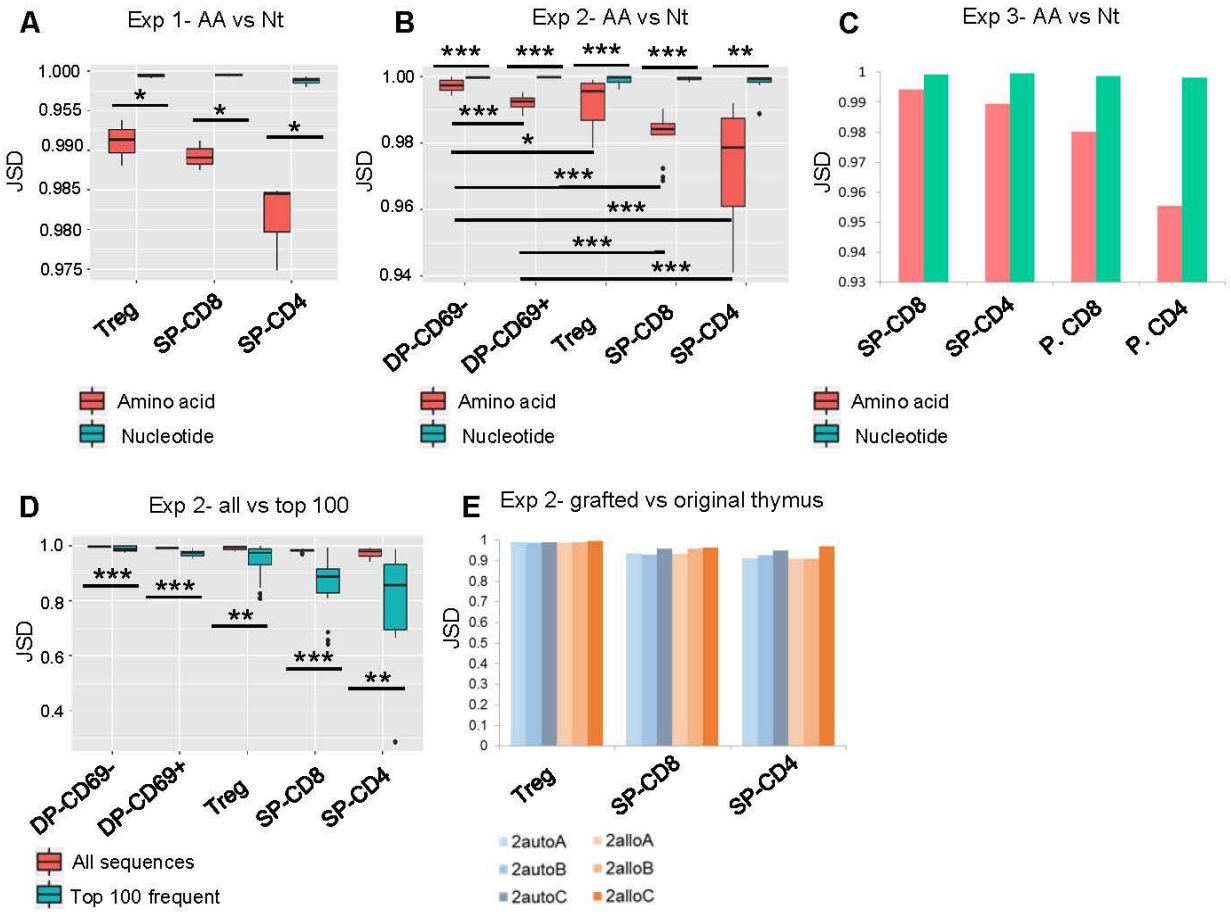

Figure 2. Repertoire divergence between animals in each experiment. A, B and C: Jensen-Shannon Divergence (JSD) scores at nucleotide (Nt) and amino acid (AA) levels for each cell population in Experiments 1 (n=3 comparisons), 2 (n=15 comparisons), and 3 (n=1 comparison), respectively. JSDs for each possible pair of mice were calculated and presented in box and whisker plots, showing median, range, and interquartile range, as well as outliers (Except Experiment 3, for which we only show one comparison per cell subset, because there were only two mice). D: AA JSD comparing each pair of mice in Experiment 2 for different cell populations. E: AA JSDs across different cell populations in Experiment 2 for all sequences vs the 100 most frequent sequences. F: AA JSD scores for TCR repertoires of different cell populations from grafted thymi of the 6 mice in Experiment 2 compared to the original autologous fetal thymus. Paired t-tests with Bonferonni multiple-testing correction were performed for all comparisons. * 0.01<p-value<0.05, **0.001<p-value<0.01, ***p-value<0.001.

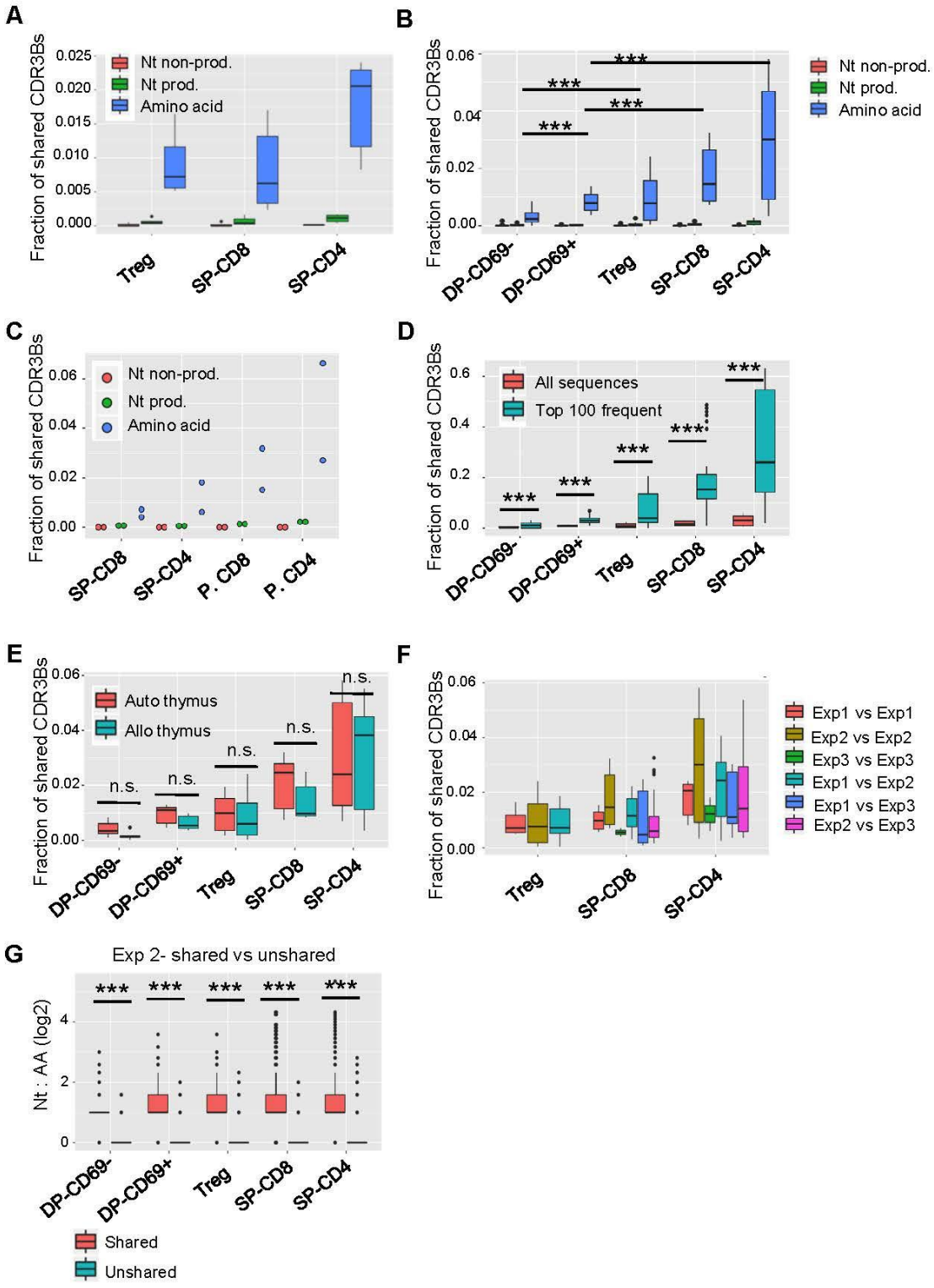

Figure 3. Proportion of shared CDR3βs between animals and experiments. A, B and C: box and whisker plots (dot plot for Experiment 3 due to lower sample size) comparing proportions of shared CDR3βs between each asymmetric mouse pair in Experiment 1 (n=6 comparisons), 2 (n=30

comparisons) and 3 (n=2 comparisons) for each cell population at the Nt/nonproductive, Nt/productive and amino acid (AA) levels. D: box and whisker plot distributions of the proportion of shared CDR3βs comparing all vs the top 100 sequences by frequency in Experiment 2. E: comparisons in both directions between each pair of mice depending on whether the mice received the same (auto thymus, n=12 comparisons) or a different thymus (allo thymus, n=18 comparisons). F: distributions of the proportion of shared CDR3βs between each pair of mice within and across experiments. G: ratio of unique CDR3β nucleotide sequences per amino acid sequence (Nt/AA ratio) in shared vs unshared sequences for each pair of mice in Experiment 2. The tables show the mean Nt/AA ratio for each subset and p-values comparing different subsets. Box and whisker plots show median, range, interquartile range and outliers. Paired t-tests with Bonferonni multiple-testing correction were performed. * 0.01<p-value<0.05, **0.001<p-value<0.01, ***p-value<0.001.

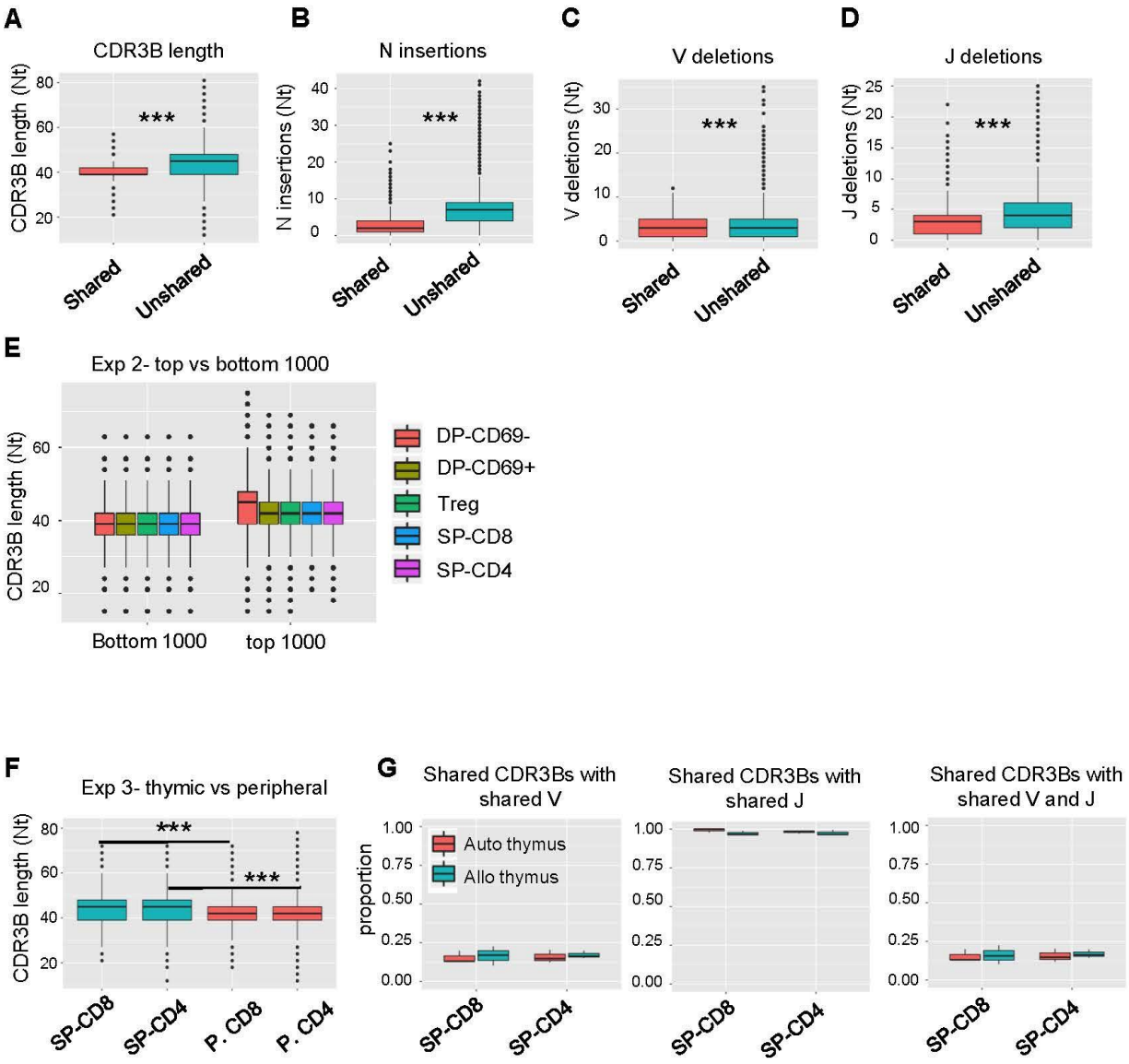

Figure 4. Characteristics of shared vs. unshared CDR3βs. A, Nucleotide length distribution of shared vs unshared SP-CD4 CDR3βs in Experiment 2. B, Number of non-template nucleotide insertions at V-D plus D-J junctions for shared vs unshared SP CD4 CDR3βs. C and D: Number of nucleotides that are deleted from the 3' end of V genes and the 5' end of J genes at V-D and D-J junctions of SP-CD4 CDR3βs, respectively. E: distribution of combined (all 6 mice in Experiment 2) CDR3β length for the 1000 most frequent CDR3βs and the 1000 CDR3βs with lowest frequencies across different thymic cell populations. The table shows p-values comparing different cell subsets (unpaired t-test). F: Nucleotide CDR3β length of all thymic and peripheral T cell subsets in Experiment 3. G: Proportion of shared CDR3βs (amino acid level) using the same Vβ gene, Jβ gene and Vβ-Jβ pair for SP-CD4 and

SP-CD8 T cell populations comparing mice with the same (auto) vs allogeneic thymus in Experiment 2. Paired t-tests with Bonferonni multiple-testing correction were performed. * 0.01<p-value<0.05, **0.001<p-value<0.01, ***p-value<0.001. Box and whisker plots show median, range, interquartile range and outliers.

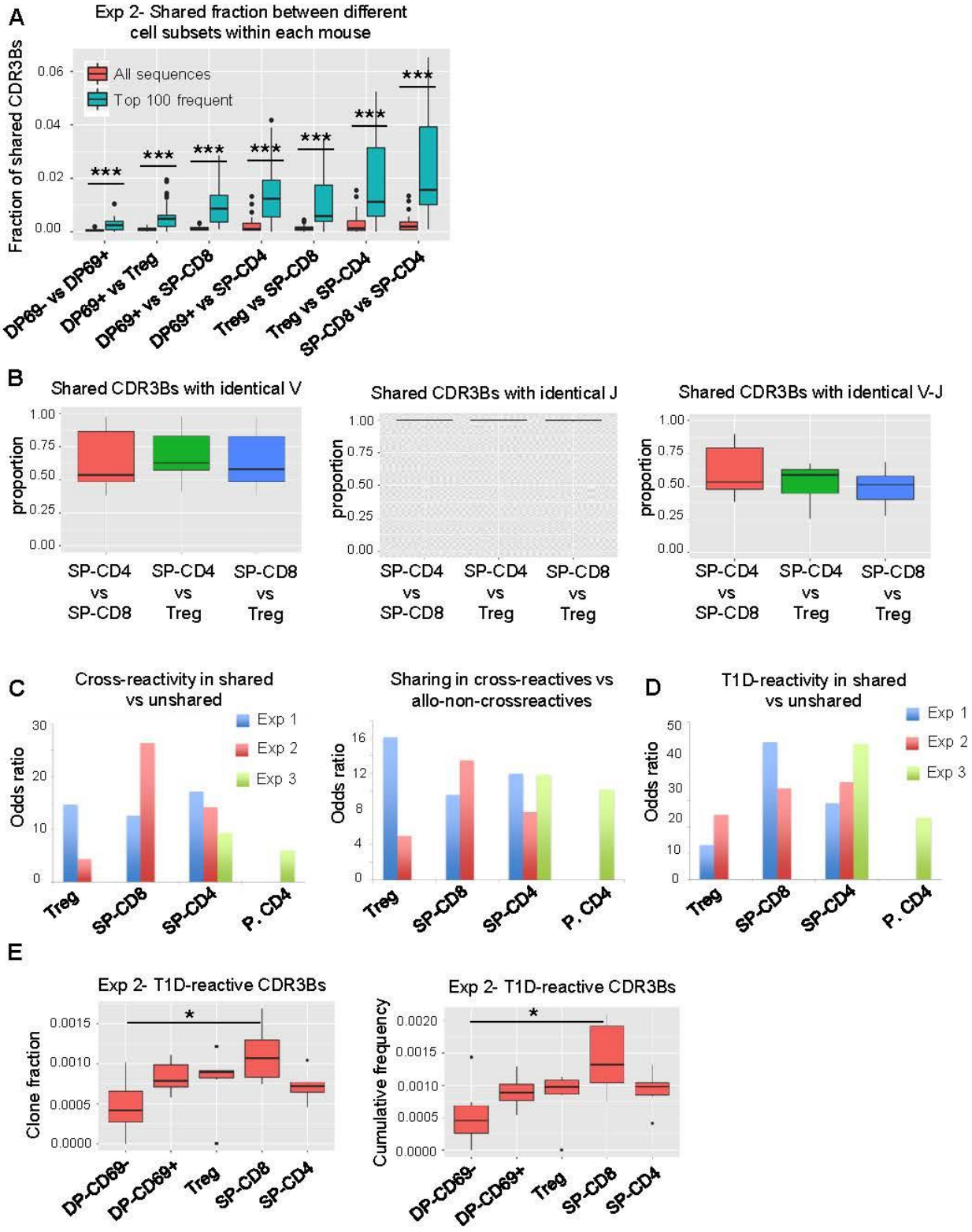

Figure 5. Overlap between different cell subsets and enrichment for cross-reactive/autoreactive CDR3βs among shared sequences. A: Proportions of shared CDR3βs between paired cell populations within each thymus graft in Experiment 2 (AA level) among all vs the 100 most frequent

CDR3βs (n=6). Potentially ambiguous sequences present in more than one cell population were not removed from this analysis. The table shows the average number of unique nucleotide sequences per amino acid sequence for shared vs unshared CDR3βs between SP-CD4 and SP-CD8 cells. B: Proportion of shared CDR3βs with a shared Vβ gene, Jβ gene and Vβ-Jβ pair, comparing each pair of SP cell populations within each mouse in Experiment 2. C and D: Odds ratios of cross-reactivity in shared vs unshared sequences, sharing in cross-reactive vs allo-non-crossreactive sequences, and T1D-reactivity in shared vs unshared sequences for Experiments 1, 2 and 3. E: P-values for the results in panels C and D (Fisher's exact test). F: Clone fraction and cumulative frequency of T1D-reactive CDR3βs in different cell subsets of Experiment 2. * 0.01<p-value<0.05, **0.001<p-value<0.01, ***p-value<0.001. Box and whisker plots show median, range, interquartile range and outliers.

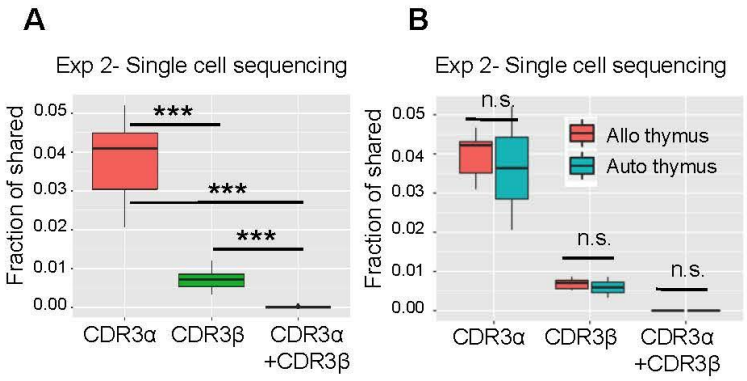

Figure 6. Fraction of shared CDR3αs, CDR3βs and paired CDR3α-CDR3βs revealed by single cell T cell sequencing. A: Fraction of shared CDR3αs, CDR3βs and paired CDR3α-CDR3βs for SP-CD4 cells between each pair of mice in Experiment 2 (except 2autoA) at the AA level (comparisons in both directions, n=20 comparisons). B: Comparisons in both directions between each pair of mice depending on whether the mice received the same (auto thymus, n=8 comparisons) or a different thymus (allo thymus, n=12 comparisons). C: Number of unique CDRαs, CDR3βs and paired CDR3α-CDR3βs, the fraction of cells with a β chain that have at least one paired α chain or two paired α chains and the fraction cells with an α chain that have a paired β chain for SP-CD4 cells in all five mice in Experiment 2. Box and whisker plots show median, range, interquartile range and outliers.

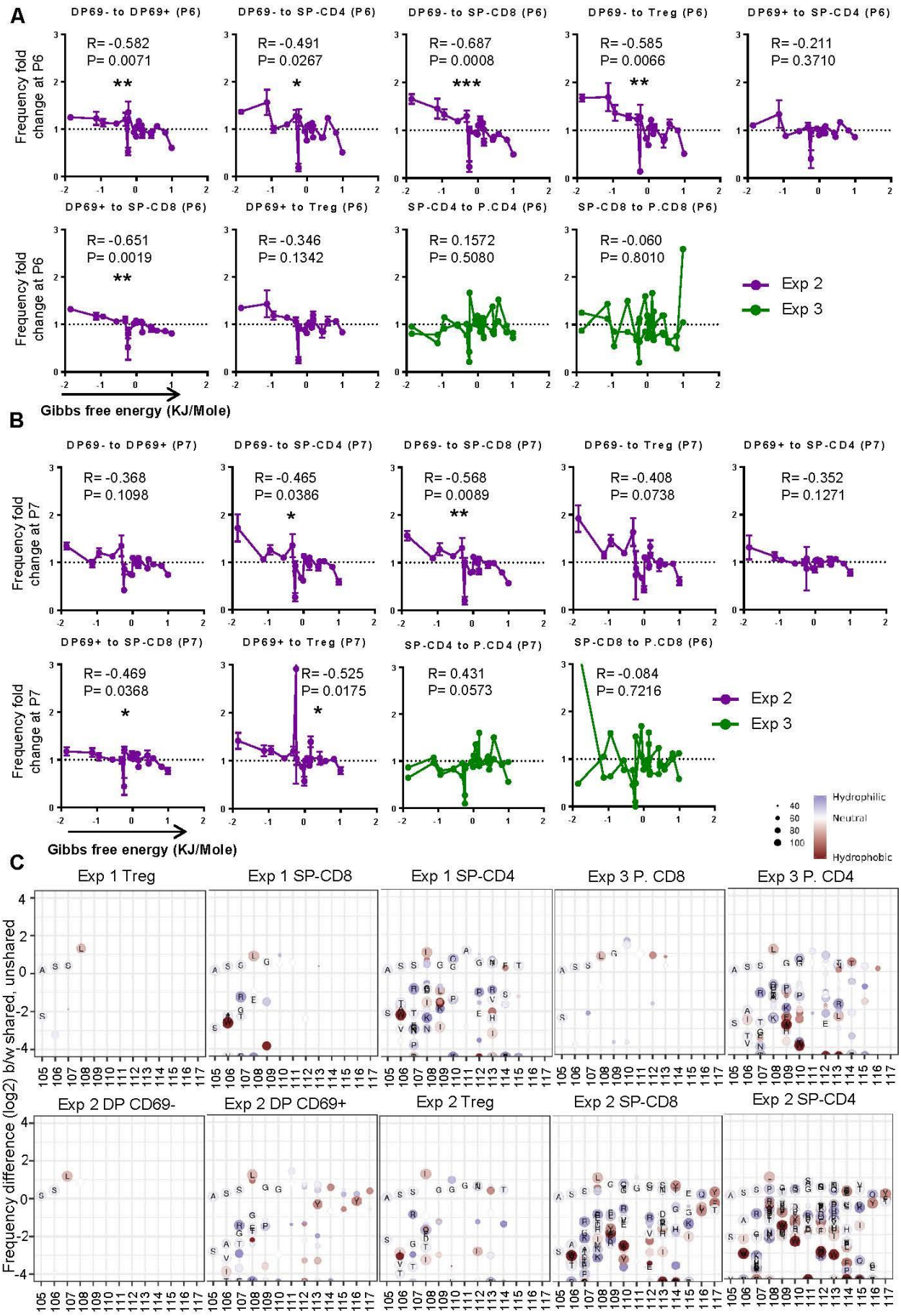
A
DP69- to DP69+ (P6)
R= -0.582
P= 0.0071
**
DP69- to SP-CD4 (P6)
R= -0.491
P= 0.0267
*
DP69- to SP-CD8 (P6)
R= -0.687
P= 0.0008
***
DP69- to Treg (P6)
R= -0.585
P= 0.0066
**
DP69+ to SP-CD4 (P6)
R= -0.211
P= 0.3710
DP69+ to SP-CD8 (P6)
R= -0.651
P= 0.0019
**
DP69+ to Treg (P6)
R= -0.346
P= 0.1342
SP-CD4 to P.CD4 (P6)
R= 0.1572
P= 0.5080
SP-CD8 to P.CD8 (P6)
R= -0.060
P= 0.8010
Frequency fold change at P6
Exp 2
Exp 3
Gibbs free energy (KJ/Mole)
B
DP69- to DP69+ (P7)
R= -0.368
P= 0.1098
DP69- to SP-CD4 (P7)
R= -0.465
P= 0.0386
*
DP69- to SP-CD8 (P7)
R= -0.568
P= 0.0089
**
DP69- to Treg (P7)
R= -0.408
P= 0.0738
DP69+ to SP-CD4 (P7)
R= -0.352
P= 0.1271
DP69+ to SP-CD8 (P7)
R= -0.469
P= 0.0368
*
DP69+ to Treg (P7)
R= -0.525
P= 0.0175
*
SP-CD4 to P.CD4 (P7)
R= 0.431
P= 0.0573
SP-CD8 to P.CD8 (P6)
R= -0.084
P= 0.7216
Frequency fold change at P7
Exp 2
Exp 3
Gibbs free energy (KJ/Mole)
40
60
80
100
Hydrophilic
Neutral
Hydrophobic
C
Exp 1 Treg
Exp 1 SP-CD8
Exp 1 SP-CD4
Exp 3 P. CD8
Exp 3 P. CD4
Exp 2 DP CD69-
Exp 2 DP CD69+
Exp 2 Treg
Exp 2 SP-CD8
Exp 2 SP-CD4
Frequency difference (log2) b/w shared, unshared

Figure 7. Interaction with self-peptides in selection of shared and unshared sequences. Fold changes (mean±SEM) in the relative AA frequencies vs hydrophobicity of the AA based on the Gibbs Free Energy at either Position 6 (A) or Position 7 (B) for transition from DP CD69^{-} to DP CD69^{+} cells and from there to SP cell subsets in Experiment 2 and also in transition from SP-CD8 and SP-CD4 to peripheral CD8 and CD4 cells for Experiment 3. Spearman correlation coefficient R values and p-values from the nonparametric Spearman correlation test are shown. Negative R values imply that as hydrophobicity increases, so does the fold change in the relative amino acid frequency across the two populations. P-value < 0.05 is shown with *. C: Differential abundance of each AA at each position in CDR3β, computed by randomly selecting a length-matched unshared sequence for each shared sequence. Shared sequences are those present in at least two mice and unshared sequences are unique to a single mouse. Only AAs producing a Benjamini-Hochberg-adjusted p-value <0.05 by Fisher's exact test are shown. The AAs plotted at frequencies 0 were preferentially used at that position in shared sequences, while those <0 were preferentially used in unshared sequences.

Supplemental Information

Supplemental Methods:

Generation of humanized mice

NOD-scid common cytokine gamma chain knockout (NOD.Cg-Prkdc^{scid} Il2rg^{tm1Wjl}/SzJ) (NSG) mice were obtained from the Jackson Laboratory and housed in a specific pathogen-free microisolator environment. Discarded human fetal thymus and liver tissues (gestational age 17 to 20 weeks) were obtained from Advanced Biosciences Resource. Fetal thymus fragments were cryopreserved in 10% dimethyl sulfoxide and 90% human AB serum (Atlanta Biologicals). In Experiment 1, three NSG mice were sublethaly irradiated (100cGy) and injected i.v with 2x10^5 human fetal liver (FL)-derived CD34$^+$ cells (referred to as hematopoietic stem cells, HSCs) (Figure 1A). Autologous human fetal thymus fragments measuring about 1 mm^3 were cryopreserved, thawed and transplanted under the kidney capsule of these mice, as described [16]. In Experiment 2, six mice received i.v injection of 2x10^5 human FL-derived CD34$^+$ HSCs from another donor. Three of these mice (mice 2autoA, 2autoB and 2autoC) received autologous human fetal thymus and the other three (mice 2alloA, 2alloB and 2alloC) received an allogeneic human fetal thymus transplant (Figure 1B). In Experiment 3, two NSG mice were thymectomized, sublethaly irradiated (100cGy) and injected i.v with 2x10^5 human fetal liver (FL)-derived CD34$^+$ HSCs from a different donor than those used in Experiments 1 and 2 (Figure 1C). To ensure that the transplanted donor thymus T cells were not able to persist, we froze and thawed the thymus tissues and also physically removed residual cells by repeated pipetting up and down before transplantation. To further deplete passenger thymocytes that might migrate to the periphery and limit allogeneic HSC engraftment, an anti-human CD2 antibody was injected to the mice in 2 weekly doses (400µg/mouse, i.p) as we have described [16]. For analysis of human reconstitution, mice were bled at regular intervals for FCM analysis of human T cells, B cells and monocytes and their naïve/memory state.

FACS sorting of different subsets of grafted thymus and peripheral cells

At weeks 14, 20 and 22 after thymus transplantation, mice from Experiments 1, 2 and 3, respectively, were euthanized. Grafted thymi (for mice in all experiments) and spleen and lymph nodes (LNs) (only mice in Experiment 3) were harvested and the thymocytes and pooled spleen and LN cells were isolated by physical force (crushing the thymus tissue between two slides and crushing the spleen and LNs through a 70µm cell strainer using a syringe plunger). After counting the total number of cells, they were stained with the following antibodies for FACS sorting: anti-human CD3 (PerCP-Cy5.5), anti-human CD5 (FITC), anti-human CD4 (PE-Cy7), anti-human CD8 (APC-Cy7), anti-human CD25 (PE) and anti-human CD127 (BV421). In Experiment 2, besides staining with these antibodies, a portion of cells were stained in a separate tube with the following markers: anti-human CD3 (PerCP-Cy5.5), anti-human CD4 (PE-Cy7), anti-human CD8 (APC-Cy7), anti-human CD69 (BV650) and anti-human TCRα/β (PE). In both experiments, after gating out the dead cells and doublets, Tregs, single positive (SP) CD8 cells and Treg-depleted SP CD4 cells were sorted within a $CD3^+$ $CD5^+$ gate. Tregs were sorted as $CD8^-$ $CD25^{high}$ $CD127^-$ $CD4^+$ cells (Figure S2A). In Experiment 2, the cells in the second tube were first gated out for doublets and dead cells. Within the population of CD4 and CD8 double positive (DP) cells, $CD69^+$ $TCR\alpha/\beta^{high}$ cells were sorted as positively-selected DP cells. The remaining DP cells were sorted as non-selected DP cells (Figure S2B). Thymic SP cells in Experiment 3 were sorted with the same panel as in Experiment 2. To sort peripheral (pooled spleen and LN) CD4 and CD8 cells in Experiment 3, after gating out the dead cells and doublets, CD4 and CD8 cells were sorted within a $CD3^+$ gate. Sorting was done using a BD Influx cell sorter. The purity of sorted cells was %90-%96 for different cell subsets (Figure S2C).

Single cell TCR sequencing

Single cell TCR sequencing was performed according to the manuals provided by the 10X Genomics company (Chromium Single Cell 5' Library & Gel Bead Kit, PN-1000006). Briefly, after sorting thymic SP-CD4 cells, 17,000 cells from each thymus graft were loaded into the chip along with partitioning oil,

the Gel beads and a master mix containing RT enzyme and poly-dt RT primers. The assembled chips were placed into the Chromium Controller, where the cells are mixed with the beads, master mix reagents and oil. Gel Beads-in-emulsion (GEMs) were generated, where all generated cDNAs shared a common 10x Barcode. After cDNA amplification and TCR locus target enrichment, enriched libraries were constructed. Each sample was indexed with a unique barcode for each well of the Chromium i7 Index Plate. After quantifying the amplified DNA using a Bioanalyser, the same amount of DNA from different samples was pooled and sequenced with an Illumina NextSeq machine. The output files were converted to FASTQ files using the Cell Ranger pipeline. The Loupe V(D)J Browser was used for preliminary analysis within each sample. Further analysis to compare different samples was done in R. The vloupe files of the single cell TCR sequences are available at https://github.com/Aleksobrad/Humanized-Mouse-Data.

Computational and statistical analysis

Adaptive ImmunoSeq performs PCR amplification, read sequencing, and mapping, with bias correction and internal controls. These analyses return tabulated read counts corresponding to unique clonal CDR3 DNA sequences across all samples, and including information on the CDR3 amino acid sequence and VJ usage of these clones. From this, we normalize read counts to frequency of clonal expression for each sample on the level of distinct CDR3 nucleotide sequence, distinct CDR3 amino acid sequence, and distinct V-J pair. This repertoire characterization process is done separately for read-count tables of productive clones and nonproductive clones, which are identified as being out of frame or including a stop codon.

For each sample, then, we generate clone frequency tables at the level of non-productive nucleotide sequence, productive nucleotide sequence, amino acid sequence, and VJ usage. Template counts, clonality scores, unique clone counts and entropies for each sample were calculated. Templates are cell count estimates for each clone, derived by Adaptive ImmunoSeq in their TCR-sequencing pipeline. Each unique TCR DNA-sequence in the repertoire (unique clone) may be represented by multiple

sequenced templates, with a greater number of templates indicating a higher-frequency clone. In every sample, clonality is calculated as an inverse measure of repertoire diversity, in order to ensure that repertoires are comparable. Clonality is entropy normalized for the number of clones N, where: $\forall i\ with\ frequency\ p_i, H_{obs} = \sum p_i \log_2 p_i\ , H_{max} = \sum \frac{1}{N} \log_2 \frac{1}{N} and\ clonality = 1 - H_{obs}/H_{max}$ such that clonality of 1 indicates a single dominant clone, and clonality of 0 indicates uniform distribution of clone frequencies. Our definition of clonality is based on CDR3β sequences and not the entire TCR β chain. Entropy (H) is a measure of diversity in a system, such that high-frequency clones in the repertoire decrease entropy and entropy for a sample is maximized if all clones are present at the same frequency. It is not a normalized metric and has no upper bound. Entropy is expected to be larger for larger samples, so only samples with similar numbers of unique clones can be compared in terms of relative diversity using entropy

We compared repertoires for the same cell population across mice using shared clone fraction, a non-symmetric measure such that the shared clone fraction of repertoire p compared to repertoire q is equal to the number of clonotypes present in both repertoires divided by the total number of unique clones defined in repertoire p. We alternately compared repertoires defined by their CDR3 non-productive nucleotide sequences, CDR3 productive nucleotide sequences, and CDR3 amino acid sequences for each thymic sub-population in both experiments. We also performed systematic comparison of repertoires using the Jensen-Shannon Divergence (JSD), which accounts for clone frequencies and scales for repertoire sizes. JSD is an information theory-based measure of the divergence of TCR repertoires. This is a symmetric value defined for any two repertoires p and q as: $JSD(p, q) = H_{obs}\left(0.5 * (p + q)\right) - 0.5 * (H_{obs}(p) + H_{obs}(q))$. JSD values range between 0 and 1, where 0 indicates identical repertoires, and 1 indicates complete divergence. For both shared clone fraction and JSD, we established a statistical baseline to distinguish any observed repertoire divergences across samples from divergence due to under-sampling of rare clones. This was done by ¼ sub-sampling (with replacement) of each repertoire 100 times, and computing mean and standard deviation of divergence by JSD and clone fraction when comparing all subsamples drawn from the same sample, thus

approximating divergence due to repertoire under-sampling and capturing any potential biases towards lower divergence across thymic sub-populations due to the presence of dominant high-frequency clones. All repertoire comparisons were validated for robustness to sample size differences by sub-sampling repertoires to the same low template count (2000 templates) three times each and repeating comparisons made between whole samples across the sub-samples.

We further plot the V and J gene frequencies across samples per cell population. Mann-Whitney U-tests are performed comparing the V and J distributions of different samples, as well as the observed distributions of combined VJ frequencies to the frequencies expected by stochastic pairing of 60 possible V genes with 13 possible J genes according to the background frequency of each V and J.

To identify correlations between amino acid use at P6 or P7 and hydrophobicity, the amino acid sequence data provided by Adaptive Immunoseq were tabulated for each of the five cell populations (DPCD69^{-}, DPCD69^{+}, SP CD4, SP CD8, SP Treg) in each animal and the amino acid and the corresponding relative frequency at P6 and at P7 was recorded for each of the CDR3β lengths. These frequencies were normalized such that the sum of all the amino acids within a given cell population and given CDR3β length in each mouse is one. These frequencies were subsequently chain-length matched, and the fold-change value was obtained as the ratio of the amino acid's relative frequency in cell population 2 to its relative frequency in cell population 1. The average fold change of the amino acid was determined as the numerical average of the fold changes across the mice.

To identify motifs at the sequence level comparing sequences shared between any two mice for a given cell population and sequences unique to a single mouse, a length-matched unshared sequences dataset of the same size as the shared sequences dataset was generated for each population by randomly selecting a sequence of the same length from the unshared sequence set for each sequence in the shared sequence set. Methods from Greiff et al. [27], which successfully distinguished between public and private antibody repertoires were applied to this dataset to identify subsequence level

features which can be used to distinguish between shared and unshared sequences. This method uses normalized gapped k-mer (two subsequences of length k, separated by a gap of up to m amino acids) count as an input to a support vector machine to predict shared/unshared status. SVM analysis was run using k = 1, m = 1 and cost = 100, and 10 fold cross-validation was performed to assess performance of the classifier, using balanced accuracy (mean of sensitivity and specificity) as a performance metric. This was repeated on 10 length-matched datasets generated as described above. To analyze differential usage of amino acids at each position as defined by IMGT, Fisher's exact test was performed for all sequences in one length matched dataset of shared and unshared sequences. Frequency differences of amino acid and position combinations were analyzed and plotted for all cases where $p < 0.05$ by Fisher's exact test.