



A Mechanism for TCR Sharing between T Cell Subsets and Individuals Revealed by Pyrosequencing

This information is current as of August 5, 2022.

Vanessa Venturi, Máire F. Quigley, Hui Yee Greenaway, Pauline C. Ng, Zachary S. Ende, Tina McIntosh, Tedi E. Asher, Jorge R. Almeida, Samuel Levy, David A. Price, Miles P. Davenport and Daniel C. Douek

J Immunol 2011; 186:4285-4294; Prepublished online 7 March 2011;
doi: 10.4049/jimmunol.1003898
<http://www.jimmunol.org/content/186/7/4285>

Supplementary Material <http://www.jimmunol.org/content/suppl/2011/03/02/jimmunol.1003898.DC1>

References This article **cites 42 articles**, 21 of which you can access for free at:
<http://www.jimmunol.org/content/186/7/4285.full#ref-list-1>

Why *The JI*? Submit online.

- **Rapid Reviews! 30 days*** from submission to initial decision
- **No Triage!** Every submission reviewed by practicing scientists
- **Fast Publication!** 4 weeks from acceptance to publication

**average*

Subscription Information about subscribing to *The Journal of Immunology* is online at:
<http://jimmunol.org/subscription>

Permissions Submit copyright permission requests at:
<http://www.aai.org/About/Publications/JI/copyright.html>

Email Alerts Receive free email-alerts when new articles cite this article. Sign up at:
<http://jimmunol.org/alerts>

A Mechanism for TCR Sharing between T Cell Subsets and Individuals Revealed by Pyrosequencing

Vanessa Venturi,^{*,1} Máire F. Quigley,^{†,1} Hui Yee Greenaway,^{*,1} Pauline C. Ng,^{‡,§}
Zachary S. Ende,[†] Tina McIntosh,[§] Tedi E. Asher,[†] Jorge R. Almeida,[†] Samuel Levy,[¶]
David A. Price,^{†,||} Miles P. Davenport,[#] and Daniel C. Douek[†]

The human naive T cell repertoire is the repository of a vast array of TCRs. However, the factors that shape their hierarchical distribution and relationship with the memory repertoire remain poorly understood. In this study, we used polychromatic flow cytometry to isolate highly pure memory and naive CD8⁺ T cells, stringently defined with multiple phenotypic markers, and used deep sequencing to characterize corresponding portions of their respective TCR repertoires from four individuals. The extent of interindividual TCR sharing and the overlap between the memory and naive compartments within individuals were determined by TCR clonotype frequencies, such that higher-frequency clonotypes were more commonly shared between compartments and individuals. TCR clonotype frequencies were, in turn, predicted by the efficiency of their production during *V(D)J* recombination. Thus, convergent recombination shapes the TCR repertoire of the memory and naive T cell pools, as well as their interrelationship within and between individuals. *The Journal of Immunology*, 2011, 186: 4285–4294.

Although the importance of T cells in the control of infectious agents throughout the lifetime of an individual is now well established, the influence of variation within the repertoires of Ag-specific TCRs has only more recently become appreciated (1–3). Furthermore, the relative impact of the different events involved in T cell development on shaping the peripheral T cell repertoire remains poorly understood. The processes of germline *V*, *J*, and *D* (for β -chains) gene recombination and junctional diversification by nucleotide deletion and addition generate a vast array of $>10^{18}$ unique TCR $\alpha\beta$ sequences in humans; in turn, these nascent TCRs undergo thymic selection

and peripheral expansion to populate the naive T cell pool, which is estimated to contain $\sim 2.5 \times 10^7$ unique TCRs (4–6). Yet, because our current understanding of T cell repertoire composition derives predominantly from studies of rather restricted epitope-specific responses, it remains unclear whether the individual clonotypes (T cell populations defined by their expressed TCRs) that make up such highly diverse naive T cell pools are equally represented to produce a featureless “repertoire landscape” or whether the distribution of clonotype sizes is quantitatively varied. If the latter is the case, the question arises as to what are the mechanisms that underlie the hierarchical features of clonotype distribution and whether any such patterns are preserved and transmitted to the memory T cell pools.

The relationship between the memory and naive T cell repertoires has generally been examined within the fixed framework of epitope-specific T cells that survive an acute response to antigenic challenge and enter the memory pool. Yet, this approach should be considered in light of the dynamic relationship between these two populations, which involves continuous thymic output of new naive T cells, homeostatic maintenance of the peripheral repertoire, and recruitment of naive T cells to the memory pool through episodic and persistent antigenic stimulation (7, 8). Recruitment from the naive T cell pool seems to be highly efficient (9), such that even CD8⁺ T cell clonotypes with very low avidity for cognate Ag (10, 11) are mobilized from the naive repertoire. Therefore, one might expect that the clonotypic landscape of the memory T cell repertoire after initial recruitment should reflect the hierarchical distribution of available clonotypes within the naive T cell repertoire. Observations in mice have provided evidence for (12) and against (13) a role for thymic selection in shaping clonotypic output to generate a hierarchical distribution within the naive T cell pool, whereas other studies found only limited overlap between the repertoires of central and effector memory T cells in mice (14) and humans (15). However, small sample sizes and insufficient sequencing depth in some of these studies may have led to an underestimation of clonotypic frequencies and the extent of overlap.

Studies in genetically identical mice showed that almost a third of the naive peripheral TCR β repertoire overlaps between individuals

^{*}Computational Biology Group, Centre for Vascular Research, University of New South Wales, Sydney, New South Wales 2052, Australia; [†]Human Immunology Section, Vaccine Research Center, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892; [‡]Department of Computational and Mathematical Biology, Genome Institute of Singapore, Singapore; [§]J. Craig Venter Institute, Rockville, MD 20850; [¶]The Scripps Research Institute, La Jolla, CA 92037; ^{||}Department of Infection, Immunity, and Biochemistry, Cardiff University School of Medicine, Heath Park, Cardiff CF14 4XN, Wales, United Kingdom; and [#]Complex Systems in Biology Group, Centre for Vascular Research, University of New South Wales, Sydney, New South Wales 2052, Australia

¹V.V., M.F.Q., and H.Y.G. contributed equally to this work.

Received for publication November 30, 2010. Accepted for publication January 26, 2011.

This work was supported by the Intramural Research Program and the Office of AIDS Research of the National Institutes of Health, the Australian Research Council, and the Australian National Health and Medical Research Council. D.A.P. is a Medical Research Council (U.K.) senior clinical fellow, M.F.Q. is a Marie Curie International outgoing fellow, M.P.D. is an Australian National Health and Medical Research Council senior research fellow, and V.V. is an Australian Research Council future fellow.

Address correspondence and reprint requests to Dr. Daniel C. Douek, Dr. Miles P. Davenport, or Dr. David A. Price, Human Immunology Section, Vaccine Research Center, National Institute of Allergy and Infectious Diseases, National Institutes of Health, 40 Convent Drive, Room 3509, Bethesda, MD 20892 (D.C.D.), Complex Systems in Biology Group, Centre for Vascular Research, University of New South Wales, Sydney, New South Wales 2052, Australia (M.P.D.), or Department of Infection, Immunity, and Biochemistry, Cardiff University School of Medicine, Heath Park, Cardiff CF14 4XN, Wales, United Kingdom (D.A.P.). E-mail addresses: ddouek@mail.nih.gov (D.C.D.), m.davenport@unsw.edu.au (M.P.D.), or dprice1@mail.nih.gov (D.A.P.)

The online version of this article contains supplemental material.

Abbreviation used in this article: IMGT, ImMunoGeneTics.

(16). The extent to which total peripheral TCR repertoires overlap between individuals in outbred human populations is not known. However, the interindividual sharing of epitope-specific clonotypes that has been observed in a variety of immune responses in humans (as well as in mice and rhesus macaques) (17, 18) suggests that memory and naive human TCR repertoires are also not unique to an individual. Our previous studies of such epitope-specific CD8⁺ T cell repertoires suggested that a process of convergent recombination enables some TCR sequences to be produced more frequently by *V(D)J* gene recombination than others (19–21). Convergent recombination predicts that such frequently made clonotypes will have a greater likelihood of being shared between individuals within epitope-specific memory CD8⁺ T cell repertoires and that they will be present more frequently in the naive T cell repertoires within and among individuals (18).

The T cell repertoire can now be probed with sufficient depth using new sequencing technology (22–24) and, when combined with high-definition polychromatic flow cytometric sorting (25), one may interrogate the clonotypic composition of stringently defined memory and naive T cell populations. In this study, we addressed the following hypotheses: that TCRβ clonotype frequency in the memory and naive CD8⁺ T cell pools is highly influenced by production frequency as predicted by convergent recombination; that the hierarchy of TCRβ clonotype frequencies established by *V(D)J* gene recombination is maintained through pairing with TCRα-chains, thymic selection, and peripheral expansion into the naive pool and subsequently into the memory T cell repertoire; and that convergent recombination is a fundamental determinant of TCRβ clonotype sharing between individuals and between the memory and naive T cell pools within individuals.

Materials and Methods

Subjects

Donors were healthy HLA-A*0201⁺ blood bank volunteers, who provided consent in accordance with the Institution Review Board of the Vaccine Research Center. Age, sex, and CMV serostatus are described in Table I. A total of four donors was studied.

Isolation of cell subsets

PBMCs were obtained by apheresis and enriched for CD8⁺ T cells using an Ab-depletion mixture containing reagents specific for CD4, CD14, CD19, CD36, CD66b, and CD235a (Stemcell Technologies, Vancouver, British Columbia, Canada). The CD8⁺ T cell-enriched sample was stained with a panel of directly conjugated mAbs specific for the following surface markers: CD3–allophycocyanin–H7 (BD Pharmingen), CD27–PECy5, CD127–PECy5.5 (Beckman Coulter), CD4–QD705, CD8–Ax594–PE, CD14–Pacific Blue, CD19–Pacific Blue, CD45RO–QD585, CD57–QD545, and CCR7–Ax680 [all provided by the Vaccine Research Center flow cytometry core and TCR Vβ8–FITC, which corresponds to TCR Vβ12 in the ImMunoGeneTics (IMGT) nomenclature (26) (Beckman Coulter clone 56C5.2)]. Dead cells were excluded using the amine viability dye ViViD (Invitrogen). Flow cytometric cell sorting of distinct TCR Vβ12⁺ phenotypic subsets was conducted using a modified FACSAria (BD Biosciences), as shown in Fig. 1; postsort purity was >99% in all cases. Electronic compensation was performed with Ab-capture beads (BD Biosciences) stained with the individual reagents used for the experimental samples. Data were analyzed using Flow Jo v9.0.1 (Tree Star). Each isolated sample contained from 1.95×10^5 to 2×10^6 TCR Vβ12⁺CD8⁺ T cells.

Pyrosequencing of TRB PCR products

mRNA was extracted from cell samples using the Qiagen Oligotex Direct mRNA extraction kit, and the entire sample was subjected to cDNA synthesis. All *TRBV12-4/TRBJ1-2* gene rearrangements were amplified by PCR using primers specific for *TRBV12-4* (5′-CTGAAGATCCAG-CCCTCAGA-3′) and *TRBJ1-2* (5′-GTTAACCTGGTCCCCGAAC-3′). Each PCR contained $1 \times$ HiFi Buffer, 3 mM MgSO₄, 200 μM 2′-deoxy-

nucleoside 5′-triphosphates, 2.8 U platinum Taq Hi-Fi DNA polymerase (Invitrogen), and 10 pmol of each primer in a final volume of 50 μl. Cycling parameters were as follows: 95°C for 5 min; 95°C for 30 s, 68°C for 1 min (2 cycles); 95°C for 30 s, 65°C for 30 s, 68°C for 30 s (3 cycles); and 95°C for 30 s, 60°C for 30 s, 68°C for 30 s (30 cycles). Purified amplicons were sequenced using GS Titanium technology (Roche) (27).

Analysis of Ag-specific CD8 T cell populations

CMV pp65 NV9-specific CD8⁺ T cells were stained with an allophycocyanin-labeled MHC class I tetramer and sorted by flow cytometry. Molecular analysis of all expressed TCR locus (TRB) gene products was conducted using a template-switch anchored RT-PCR, as described previously (28).

Sequence alignment and CDR3 identification

Analysis of TCRβ sequences was conducted with reference to the *01 allele sequences for the human *TRBV12-4*, *TRBJ1-2*, and *TRBD1* genes, which were obtained from the international IMGT information system (<http://imgt.cines.fr/>); the IMGT nomenclature is used throughout for all TCR genes (26). Raw sequences were aligned against the *TRBV12-4* and *TRBJ1-2* sequences using blastn (29). The CDR3 sequence extending between the YFC and FGSG amino acid motifs in the *TRBV12-4* and *TRBJ1-2* gene-encoded region of the TCRβ sequences was then extracted in each case, including the flanking motifs. Out-of-frame TCRβ sequences were removed. Homopolymers containing >10 A, C, or T nucleotides or >15 G nucleotides were attributed to sequencing error, and TCRβ sequences containing these homopolymers were removed. Sequences that contained multiple *TRBV* and *TRBJ* regions, as well as sequences with lengths >80 bp that contained multiple *TRBV* or *TRBJ* regions, most likely represented ligated products formed during PCR amplification and were also removed. Successful alignments were made in 85% of cases. The *TRBV12-4* and *TRBJ1-2* genes were sequentially aligned to the 5′ and 3′ ends of the TRB sequence, respectively. Unaligned sequence in the *TRBV12-4/TRBJ1-2* junctional region was attributed to the *TRBD1* gene for the longest consecutive match of at least two nucleotides. Any further unaligned nucleotides at the *TRBV12-4/TRBD1* and *TRBD1/TRBJ1-2* junctions were considered nucleotide additions.

Statistical analysis

Statistical analyses were conducted using GraphPad Prism software (GraphPad) and R v2.8.1 (R Development Core 2008). Correlations were performed using the Spearman Rank test, and comparisons between two groups of data were performed using the Mann–Whitney *U* test.

Results

Deep sequencing enables detailed repertoire analysis of stringently defined T cell subsets

We investigated the TCRβ repertoires of precisely defined memory and naive CD8⁺ T cell populations from four healthy HLA-A*0201⁺ donors, focusing on the specific portions of the TCR repertoire defined by the single gene combination *TRBV12-4/TRBJ1-2*. This approach enabled sufficient sequencing depth to facilitate a detailed comparative analysis of TCRβ repertoire composition. Importantly, potential biases arising from disproportionate usage of the many possible *TRBV/TRBJ* gene combinations within the overall TCR repertoire were eliminated with this strategy. The *TRBV12-4/TRBJ1-2* gene combination was identified in previous studies as being used by the most shared human TCRβ clonotype in CD8⁺ T cell responses to the CMV-NV9 epitope (10, 21). However, this choice of *TRBV/TRBJ* gene combination should not influence the hypotheses being tested in this study, because the most shared *TRBV12-4/TRBJ1-2* clonotype was only observed in 8 of 23 donors (21), and the *TRBV12-4/TRBJ1-2* gene combination did not dominate many of the CMV-NV9-specific TCRβ repertoires previously studied (10). Moreover, to determine whether CMV serostatus was an influencing factor, a CMV-seronegative donor (Donor 2) was included in this study (Table I). The memory and naive compartments of CD8⁺ T cell populations were isolated using polychromatic flow cyto-

Table I. Summary of TRBV12-4/TRBJ1-2 repertoire data obtained from four donors

	Donor 1		Donor 2		Donor 3		Donor 4	
Age (y)	25		24		42		51	
Sex	Male		Male		Male		Male	
CMV serostatus	Positive		Negative		Positive		Positive	
	Memory	Naive	Memory	Naive	Memory	Naive	Memory	Naive
TCR sequences (<i>n</i>)	160,014	266,362	132,330	277,114	189,056	254,899	386,934	206,424
Unique a.a. clonotypes (<i>n</i>)	1,431	34,377	2,961	20,422	2,576	23,129	3,006	12,788
Unique n.t. clonotypes (<i>n</i>)	1,770	39,926	3,348	23,364	3,001	26,654	3,819	14,934

a.a., amino acid; n.t., nucleotide.

metric sorting (Fig. 1). Naive CD8⁺ T cells were defined as CD27⁺CD45RO⁻CD57⁻CD127⁺CCR7⁺ and memory CD8⁺ T cells included the CD27⁺CD45RO⁺, CD27⁻CD45RO⁺, and CD27⁻CD45RO⁻ subsets. The use of multiple phenotypic markers was critical to ensure greater homogeneity within, and reduced artifactual overlap between, these populations. TCRβ sequences were obtained from the entire population of sorted cells in each case by PCR amplification using primers specific for TRBV12-4 and TRBJ1-2, followed by pyrosequencing (Roche 454). A total of 1,873,133 TRBV12-4/TRBJ1-2 sequences was obtained across the memory and naive CD8⁺ T cell pools of all four donors. At least 1.32×10^5 and 2.06×10^5 TCRβ sequences were obtained for the memory and naive CD8⁺ T cell populations, respectively, in each donor (Table I). Each memory repertoire contained $\geq 1,431$ and 1,770 unique TCRβ clonotypes at the amino acid and nucleotide level, respectively. More than 12,788 and 14,934 unique TCRβ clonotypes at the amino acid and nucleotide level, respectively, contributed to each naive repertoire (Table I). In the analysis that follows, we investigated the features of the unique TCRβ clono-

types contributing to each donor's memory and naive pools and the total memory and naive TCRβ repertoires, where the latter accounts for clonotype size.

Clonotype size distributions differ between the memory and naive TCRβ repertoires

The memory T cell pool consists of clonotypes that have previously expanded in response to many different Ags encountered during an individual's lifetime. Thus, we would expect clonotype representation in the memory pool to differ from the corresponding size distribution in the naive pool. Indeed, the distributions of amino acid clonotype sizes differed substantially between the memory and naive TCRβ repertoires in all four donors, with greater skewing in the memory pool (Fig. 2A, Supplemental Fig. 1A–F). For example, the 10% of memory TCRβ amino acid clonotypes with the largest sizes in Donor 1 contributed 96.5% of the memory repertoire; in contrast, the 10% of naive TCRβ amino acid clonotypes with the largest sizes contributed only 46.7% of the naive repertoire. Thus, the unevenness of memory TCRβ amino acid clonotype size distributions was largely due to several highly dominant clonotypes (Fig. 2B). Interindividual differences in TCRβ amino acid clonotype size distributions were also observed, albeit on a smaller scale compared with the differences between the memory and naive pools. In particular, the sizes of naive TCRβ amino acid clonotypes were less evenly distributed in Donor 4, and the sizes of memory TCRβ amino acid clonotypes were more evenly distributed in Donor 2, compared with the other three donors (Supplemental Fig. 1A, 1B).

The most dominant memory TCRβ clonotypes made up 61, 35, 20, and 24% of the memory TCRβ repertoires of Donors 1, 2, 3, and 4, respectively. We investigated whether persistent CMV infection could potentially play a role in the surprising dominance of the memory TCRβ clonotypes in the three CMV-seropositive donors. A search of the memory TCRβ repertoires of Donors 1, 3, and 4 for TCRβ clonotypes specific for the immunodominant CMV-NV9 epitope from previous studies (10, 21, 30–34) revealed only three previously identified CMV-NV9-specific TCRβ clonotypes. All three of these previously identified CMV-NV9-specific TCRβ clonotypes were found in Donor 1 and included the second most dominant memory TCRβ clonotype, CASSLVG-GRYGTF. We determined that the most dominant memory TCRβ clonotype in Donor 1, CASAYGAYNGYTTF, was also specific for the CMV-NV9 epitope by sequencing the CMV-NV9-specific TCRβ repertoire of Donor 1 using conventional Sanger sequencing. Although CMV-specific T cell responses seem to play a role in the dominance of memory TRBV12-4/TRBJ1-2 clonotypes in Donor 1, the prevalence of the most dominant memory TCRβ clonotype in CMV-seronegative Donor 2, which was comparable to that of the most dominant memory TCRβ clonotypes in CMV-seropositive Donors 3 and 4, suggests that persistent CMV infection is not the

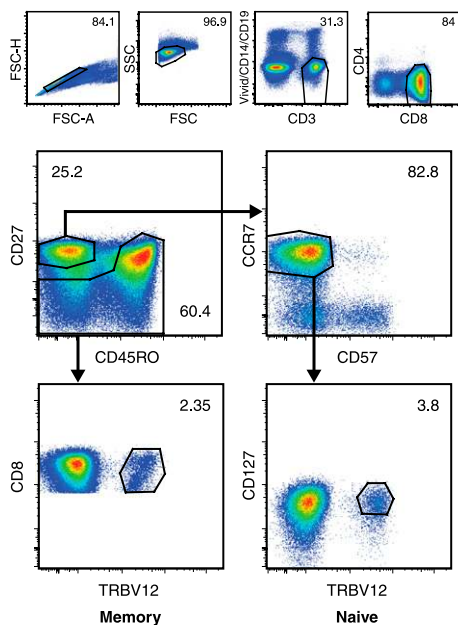


FIGURE 1. Isolation of memory and naive CD8⁺ T cells by high-definition flow cytometric sorting. Sorting scheme for the isolation of memory and naive TCR Vβ12⁺CD8⁺ T cells from Donor 4. *Top row*, The four plots delineate singlet, small lymphocyte, live CD3⁺, and CD8⁺CD4⁻ cells (from left to right). These cells were subsequently analyzed on the basis of multiple cell surface marker expression (*middle and bottom rows*). Naive cells were defined as CD27⁺CD45RO⁻CD57⁻CD127⁺CCR7⁺, and memory cells were defined as CD27⁺CD45RO⁺, CD27⁻CD45RO⁺, and CD27⁻CD45RO⁻. Both subsets expressed TCR Vβ12.

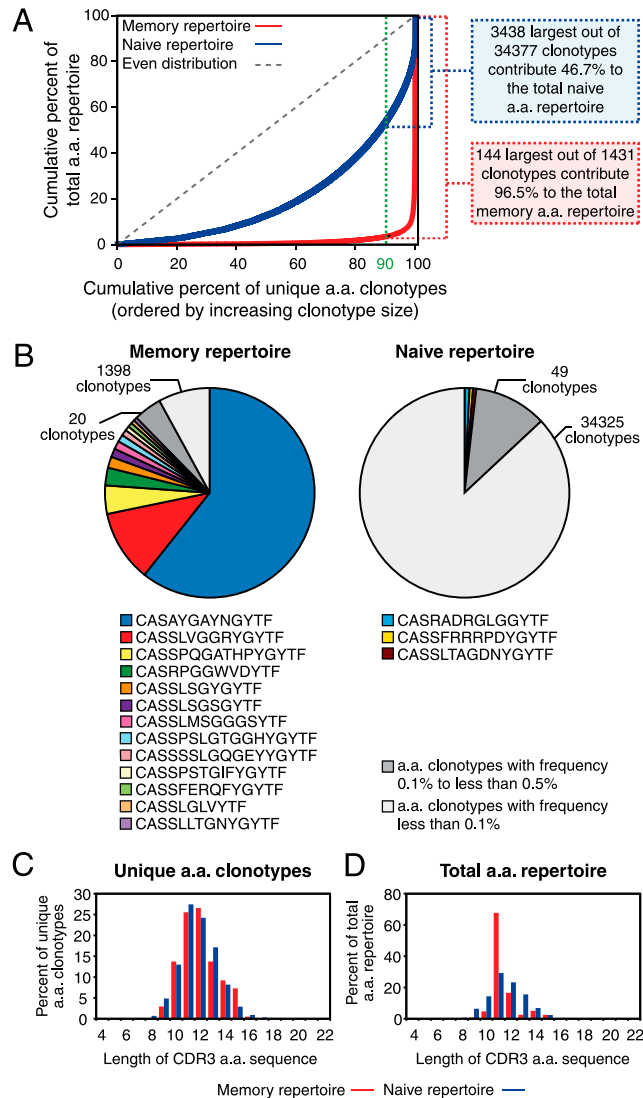


FIGURE 2. Composition of the memory and naive TCR β amino acid repertoires. **A**, Distributions of clonotype sizes in the memory and naive repertoires of Donor 1 represented as the cumulative percentage of the total TCR β amino acid (a.a.) repertoire versus the cumulative percent of unique TCR β amino acid clonotypes ordered by increasing clonotype size. In this representation, an even clonotype size distribution (i.e., all clonotypes being the same size) would be indicated by the gray dashed line, and the extent of deviation from this line indicates the unevenness of the distribution. This plot shows, for example, that the 10% of naive TCR β amino acid clonotypes with the largest sizes contribute 46.7% of the total naive TCR β repertoire; in contrast, the 10% of memory TCR β clonotypes with the largest sizes contribute 96.5% of the total memory TCR β repertoire. **B**, Individual dominant TCR β amino acid clonotypes in the memory (*left panel*) and naive (*right panel*) repertoires of Donor 1. Clonotypes that make up $\geq 0.5\%$ of the total TCR β amino acid repertoire are shown as color-shaded segments. The TCR β amino acid clonotypes are represented by the CDR3 amino acid sequence between the conserved cysteine and phenylalanine. The light and dark gray segments are composed of a number of TCR β amino acid clonotypes, as indicated on the plots. Segments shaded in dark gray correspond to TCR β sequences making up $< 0.5\%$ and $\geq 0.1\%$ of the total amino acid repertoires. Segments shaded in light gray correspond to TCR β sequences making up $< 0.1\%$ of the total amino acid repertoires. **C**, Distribution of CDR3 lengths among the unique TCR β amino acid clonotypes in the memory and naive repertoires of Donor 1. **D**, Distribution of CDR3 lengths for the total TCR β amino acid memory and naive repertoires (i.e., accounting for clonotype size) of Donor 1. CDR3 length is the number of residues between the conserved cysteine and phenylalanine residues, as defined by IMGT (26).

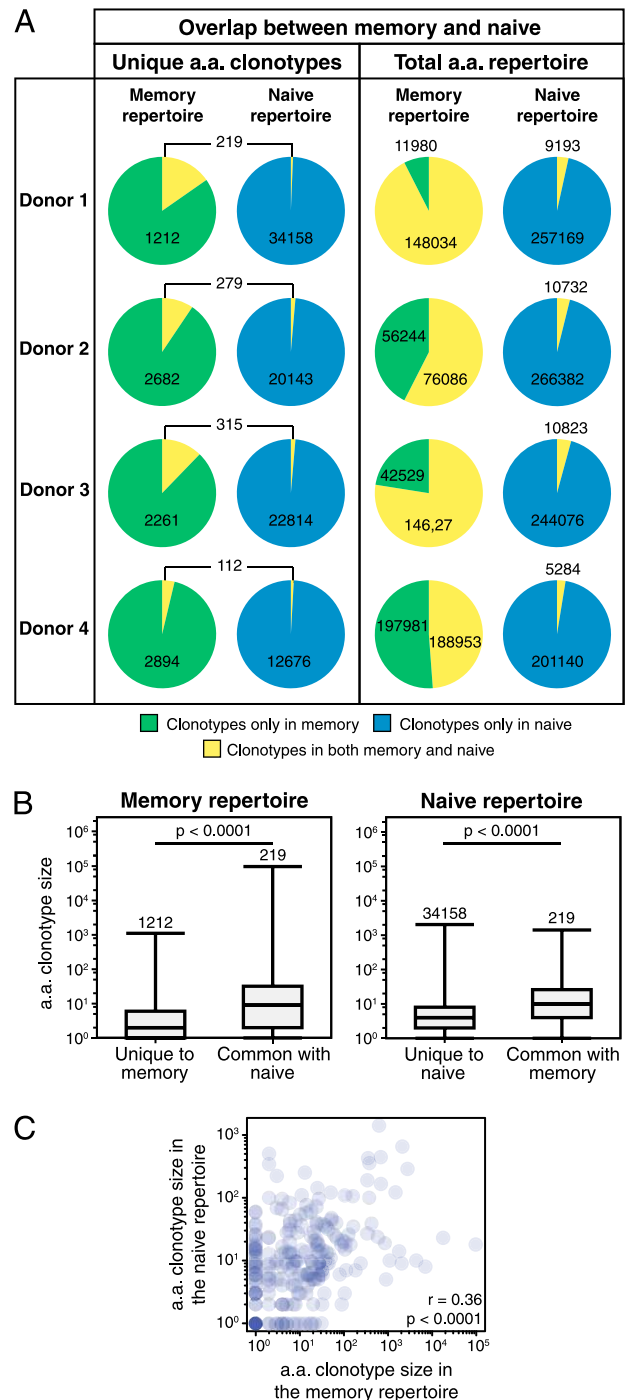


FIGURE 3. TCR β amino acid clonotypes common to the memory and naive repertoires. **A**, *Left panel*, For each of the four donors, the proportion of unique memory and naive TCR β amino acid (a.a.) clonotypes that are common to the memory and naive pools. *Right panel*, For each of the four donors, the proportion of the total memory and naive TCR β repertoires attributable to TCR β amino acid clonotypes that are common to the memory and naive pools. **B**, *Left panel*, Comparison of TCR β amino acid clonotype sizes in the memory repertoire of Donor 1 between clonotypes unique to the memory pool and those also present in the naive pool. *Right panel*, Comparison of the TCR β amino acid clonotype sizes in the naive repertoire of Donor 1 between clonotypes unique to the naive pool and those also present in the memory pool. Statistics are based on the Mann-Whitney U test. **C**, Relationship between TCR β amino acid clonotype sizes in the naive and memory repertoires for clonotypes common to both pools in Donor 1. The opacity of the circles indicates the density of points plotted. Correlations are based on the Spearman rank test.

sole determinant of the surprising dominance of these memory clonotypes.

CDR3 length distributions differ between the memory and naive TCRβ repertoires

The distribution of CDR3 sequence lengths is another feature that provides an overall view of repertoire composition. Biases in CDR3 length are often observed in epitope-specific T cell repertoires, suggesting that CDR3 length distributions might differ substantially between the memory and naive T cell pools. Surprisingly, we found that the distributions of CDR3 lengths among unique TCRβ amino acid clonotypes were similar in the memory and naive pools within individuals and between individuals. The median CDR3 length was 12 aa in the memory and naive pools of all donors (Fig. 2C, Supplemental Fig. 1G–J). However, this parameter provides little indication of the CDR3 lengths that predominate within the total repertoires. Therefore, we also assessed the distributions of CDR3 lengths across the total memory and naive TCRβ amino acid repertoires (i.e., including the size of each clonotype). We found substantial CDR3 length distribution differences between the memory and naive pools, largely due to the prevalence of one or two CDR3 lengths in the memory TCRβ repertoires of each donor (Fig. 2D, Supplemental Fig. 1K–N). The preponderance of these few CDR3 lengths in the memory repertoires was largely associated with the most dominant memory TCRβ clonotypes. For example, the dominance of the CASAY-GAYNGYTF clonotype in the memory repertoire of Donor 1 (Fig. 2B) was largely responsible for the peak at 11 aa (Fig. 2D). There were also notable interindividual differences in CDR3 length distributions among all memory TCRβ amino acid sequences between the four donors (Supplemental Fig. 1K–N). Thus, al-

though the memory and naive repertoires consisted of TCRβ amino acid clonotypes with similar CDR3 length features, most likely owing to the large variety of Ag specificities of T cells in the memory pool, differences in the most prevalent CDR3 lengths were observed between the memory and naive pools and between individuals.

Dominant memory TCRβ clonotypes are highly represented in the naive pool

Next, we examined the extent of overlap of TCRβ amino acid clonotypes between the memory and naive CD8⁺ T cell compartments. In each donor, a subset of TCRβ amino acid clonotypes was common to the memory and naive pools. Of the unique TCRβ amino acid clonotypes in the memory pool of a donor, a mean of 10.2% (range: 3.7–15.3%) was also present in the naive pool (Fig. 3A). A much smaller percentage of the TCRβ amino acid clonotypes in the naive repertoires was observed in the memory pools (mean, 1.1%; range: 0.6–1.4%). We then determined the contribution that these clonotypes present in memory and naive pools made to the total memory and total naive TCRβ repertoires (i.e., the extent of overlap of the total memory and total naive repertoires). Despite large interindividual variations, a substantial proportion of each donor's total memory TCRβ repertoire consisted of amino acid clonotypes that were also present in the naive pool (mean, 69.1%; range: 48.8–92.5%; Fig. 3A). In contrast, a much smaller proportion of an individual's total naive TCRβ repertoire overlapped with the memory pool (mean, 3.5%; range: 2.6–4.2%).

To gain a better understanding of the overlap between the memory and naive repertoires, we also investigated the commonality of TCRβ clonotypes at the nucleotide level between the

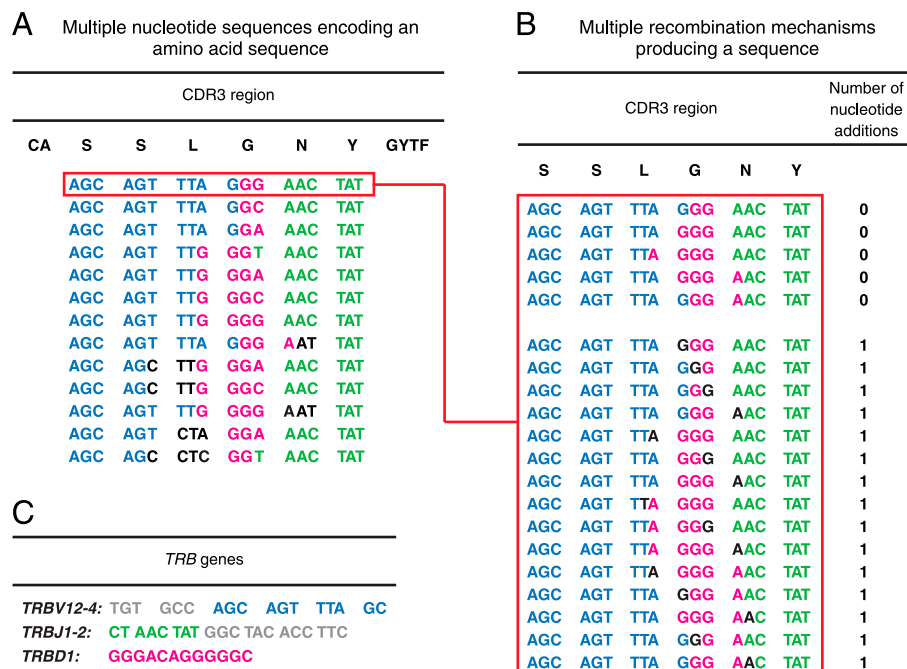


FIGURE 4. Convergent recombination. Using examples of TCRβ clonotypes prevalent in the TRBV12-4/TRBJ1-2 repertoires, the various levels of convergent recombination are illustrated. *A*, At the level of the amino acid sequence, multiple nucleotide sequences converge to code for the same TCRβ amino acid clonotype. The TCRβ clonotype, CASSLGN YGYTF, was observed in all four donors, and it was common between the naive and memory compartments in two of these donors. This clonotype was found to be encoded by a variety of nucleotide sequences across all donors' repertoires. One possible recombination mechanism is shown for each nucleotide sequence involving the *TRBV12-4* (blue), *TRBD1* (pink), and *TRBJ1-2* (green) genes and a minimal number of nucleotide additions (black). *B*, At the level of the nucleotide sequence, multiple recombination mechanisms, involving different gene contributions and nucleotide additions, converge to produce the same nucleotide sequence. This is shown for one of the nucleotide sequences (red box) encoding CASSLGN YGYTF using a variety of recombination mechanisms requiring similarly low numbers of nucleotide additions. This particular nucleotide sequence was shared among three of the four donors. *C*, The relevant human TRB gene sequences.

memory and naive compartments of a donor's repertoire. A large proportion of the total memory repertoire (mean, 62.9%; range: 33.8–90.4%) consisted of TCR β nucleotide clonotypes that were also present in the naive pool. One explanation of this overlap is simply contamination of naive cells in the memory repertoire or vice versa. However, more than half (mean, 56.8%) of the TCR β amino acid clonotypes common to memory and naive pools were encoded by at least one nucleotide sequence that was present in only one of the compartments. This suggests that, although there is substantial overlap between the memory and naive pools at the nucleotide sequence level, this only partially contributes to the substantial overlap observed at the TCR β amino acid sequence level.

To assess quantitatively the dominance of the TCR β amino acid clonotypes common to the memory and naive pools within an individual, we compared the size of these clonotypes with those that were unique to one of the compartments. In the memory repertoires, TCR β amino acid clonotypes unique to the memory pool had a significantly smaller size compared with those that were also found in the naive pool (Fig. 3B, Supplemental Fig. 2A–D). Similarly, in the naive repertoires, the size of TCR β amino acid clonotypes unique to the naive pool was significantly smaller compared with those also present in the memory pool (Fig. 3B, Supplemental Fig. 2E–H). Furthermore, among the TCR β amino acid clonotypes common to the memory and naive pools within a donor, we observed a positive correlation between clonotype size in the naive pool and clonotype size in the memory pool (Fig. 3C, Supplemental Fig. 2I–L). Thus, a substantial portion of the memory TCR β repertoire at the amino acid level overlapped with the naive TCR β repertoire within each donor. The extent of the overlap was largely due to the dominance in the memory pool of clonotypes also present in the naive pool.

Convergent recombination is a determinant of relative TCR β clonotype size

To establish a mechanistic basis for the observed TCR β clonotype dominance hierarchies, we investigated whether the process of convergent recombination could predict relative clonotype sizes in the memory and naive CD8⁺ T cell pools. Our previous studies of epitope-specific memory repertoires showed that convergent recombination enables particular TCR β clonotypes to be produced more easily than others during gene rearrangement (18–21). This process operates at two levels: particular TCR nucleotide sequences can be produced more efficiently by one or more frequently occurring V(D)J recombination events (35) and/or by many closely related recombination events (i.e., different contributions of the gene segments and nucleotide additions) (18–21) and particular TCR amino acid sequences can be easily made if they can be encoded by nucleotide sequences that are efficiently produced and/or encoded by many different nucleotide sequences. The latter is largely determined by codon degeneracy of specific amino acids in the CDR3 sequence. The convergent recombination process is demonstrated in Fig. 4 using, as an example, a *TRBV12-4/TRBJ1-2* amino acid clonotype that was observed in all four individuals. A resulting prediction is that the frequencies at which TCR sequences are produced by V(D)J recombination should shape clonotype dominance hierarchies in the naive T cell repertoire.

A key indicator that a TCR nucleotide sequence has the potential to be produced efficiently by V(D)J recombination is that it requires fewer nucleotide additions and, thus, comprises less of a random element. To determine whether the relative TCR β clonotype sizes in the peripheral repertoires were associated with the number of nucleotide additions, we calculated the minimal number of nu-

cleotide additions required to produce each of the TCR β sequences by sequentially aligning the *TRBV12-4*, *TRBJ1-2*, and *TRBD1* genes (Fig. 5A, Supplemental Fig. 3A–D). For the memory and naive CD8⁺ T cell populations in each donor, the estimated minimal number of nucleotide additions required to produce a TCR β nucleotide clonotype was negatively correlated with the size of the TCR β nucleotide clonotype in the repertoire (Fig. 5C, Supplemental Fig. 3E–L). Comparisons between the TCR β amino acid clonotypes unique to memory pool and those also present in the naive pool and between the TCR β amino acid clonotypes unique to naive pool and those also present in the memory pool revealed that nucleotide sequences coding for TCR β amino acid

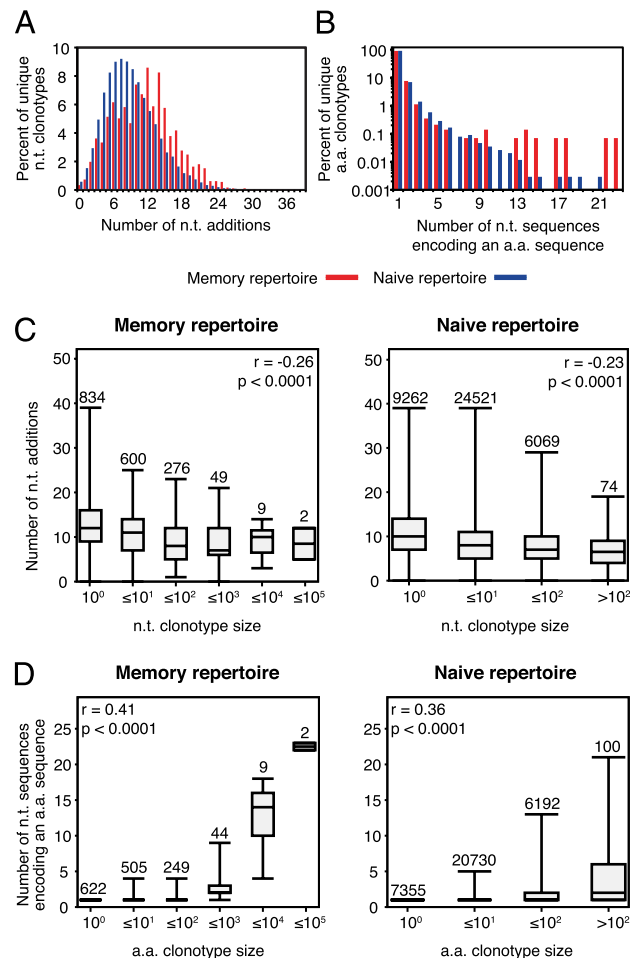


FIGURE 5. Relationship between convergent recombination and TCR β clonotype size in the memory and naive repertoires. *A*, Distribution of the minimal number of nucleotide (n.t.) additions required for the production of memory and naive TCR β clonotypes from Donor 1. *B*, Distribution of the number of different nucleotide sequences encoding memory and naive TCR β amino acid (a.a.) clonotypes from Donor 1. The percentages of unique TCR β amino acid clonotypes at which the distributions plateau at higher numbers of encoding nucleotide sequences correspond to one TCR β amino acid clonotype. *C*, Relationship between the minimal number of nucleotide additions and TCR β nucleotide clonotype size in the memory (*left*) and naive (*right*) repertoires of Donor 1. *D*, Relationship between the number of different nucleotide sequences observed to encode a TCR β amino acid clonotype and its size in the memory (*left*) and naive (*right*) repertoires of Donor 1. TCR β clonotype sizes are grouped according to the nonoverlapping categories on the horizontal axis. For example, $\leq 10^2$ refers to a number of copies >10 and $\leq 10^2$. Correlations are based on the Spearman rank test applied to the noncategorized data. In *C* and *D*, the total numbers of unique TCR β clonotypes are indicated for each group of data above the box representing the data group.

clonotypes common to the memory and naive pools required significantly fewer nucleotide additions ($p < 0.01$ for each comparison for Donors 1–4; Mann–Whitney U test).

We also investigated whether the relative sizes of memory and naive TCR β amino acid clonotypes were associated with the variety of encoding nucleotide sequences, which is another key indicator of TCR β amino acid sequence production efficiency. The distributions of the number of nucleotide sequences observed to encode each memory and naive TCR β amino acid clonotype are shown in Fig. 5B and Supplemental Fig. 3M–P. In all four donors, the number of different encoding nucleotide sequences was positively correlated with TCR β amino acid clonotype size in the memory and naive repertoires (Fig. 5D, Supplemental Fig. 3Q–X). Comparisons between the TCR β amino acid clonotypes unique to memory pool and those also present in the naive pool and between the TCR β amino acid clonotypes unique to naive pool and those also present in the memory pool revealed that TCR β amino acid clonotypes common to the memory and naive pools were encoded by a greater variety of nucleotide sequences ($p < 0.0001$ for each comparison for Donors 1–4; Mann–Whitney U test).

Thus, TCR β clonotypes with the potential to be produced more efficiently during gene recombination are more likely to be present at a higher frequency in the naive repertoire. Remarkably, this also holds for the memory repertoire, suggesting that the dominance hierarchy established by convergent recombination during TCR production is, to some extent, preserved through the generation of memory populations.

Clonotype size is a determinant of interindividual TCR β clonotype sharing

The many previous descriptions of interindividual sharing of identical epitope-specific clonotypes (17) imply that there must be a degree of TCR repertoire overlap between different individuals. However, the extent of overlap between the total peripheral repertoires is not known. Therefore, we initially determined the TCR β amino acid clonotypes that were shared between each of the six possible pairings among the four donors (i.e., Donors 1 and 2, Donors 1 and 3, Donors 1 and 4, Donors 2 and 3, Donors 2 and 4, and Donors 3 and 4). We considered two measures of the extent of TCR β sharing between pairs of donor. First, we evaluated the proportion of unique TCR β clonotypes across the memory and

naive pools of a donor that were shared with each of the other three donors. Between 3.8 and 9.8% of the unique TCR β amino acid clonotypes across the memory or naive CD8⁺ T cell populations of a donor were shared with another donor (Fig. 6A). Second, as a measure of the interindividual overlap between total repertoires (i.e., allowing for clonotype sizes), we considered the proportions of the total memory and total naive repertoires in a donor that were attributable to TCR β amino acid clonotypes shared with each of the other three donors (Fig. 6B, 6C). The proportion of a donor's total naive repertoire that was contributed by TCR β amino acid clonotypes shared with each of the other three donors ranged between 10.6 and 20.3% (mean: 15.8%; Fig. 6C). In contrast, substantially greater interindividual variability was observed in the memory repertoires; the percentages of total memory TCR β repertoires that were contributed by TCR β amino acid clonotypes present in each of the other three donors varied between 2.1 and 83.9%, with a mean across all six pairwise comparisons between donors of 37.5% (Fig. 6B).

To examine the extent of sharing of TCR β amino acid clonotypes across all four donors, we determined the number of individuals in which each TCR β amino acid clonotype was present. Identical TCR β amino acid clonotypes could be found in two, three, or all four individuals (Fig. 7). As with the comparisons between pairs of donors above, we considered two measures of the extent of TCR β sharing across the four donors: the proportion of unique TCR β clonotypes in the memory and naive pools that were shared with other individuals (Fig. 7A) and the proportions of the total memory and total naive repertoires that were attributable to shared TCR β amino acid clonotypes (Fig. 7B). Notably, TCR β amino acid clonotype size, in the memory and naive pools of an individual, was positively correlated with the number of individuals in which that TCR β clonotype was observed (Fig. 7C, Supplemental Fig. 4). Thus, TCR β amino acid clonotypes that are prevalent in individual memory and naive repertoires are more likely to be observed in many individuals.

Many TCR β clonotypes common to the memory and naive pools within individuals are also shared between individuals

Next, we investigated the extent of interindividual sharing of TCR β clonotypes common to the memory and naive repertoires within individuals. The overlapping portions of a donor's memory

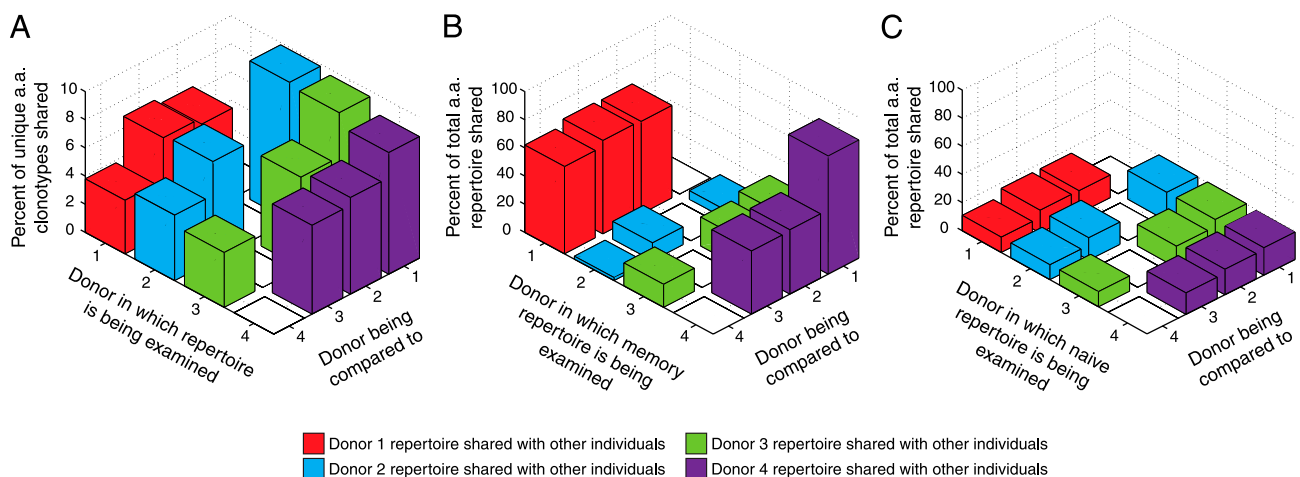


FIGURE 6. Sharing between pairs of donors of TCR β amino acid clonotypes in the memory and naive repertoires. The TCR β amino acid (a.a.) clonotypes shared between each pair of donors was evaluated, regardless of whether they were found in the memory and/or naive pool. **A**, The proportion of unique TCR β amino acid clonotypes in Donors 1–4 (i.e., pooled across the memory and naive repertoires) that were observed in each of the other three donors. The proportions of the total memory (**B**) and naive (**C**) TCR β repertoires of Donors 1–4 that were attributable to TCR β amino acid clonotypes observed in each of the other three donors. For example, **B** shows that 83.9% of the memory TCR β repertoire in Donor 4 was made up of TCR β amino acid clonotypes that were also observed in Donor 1.

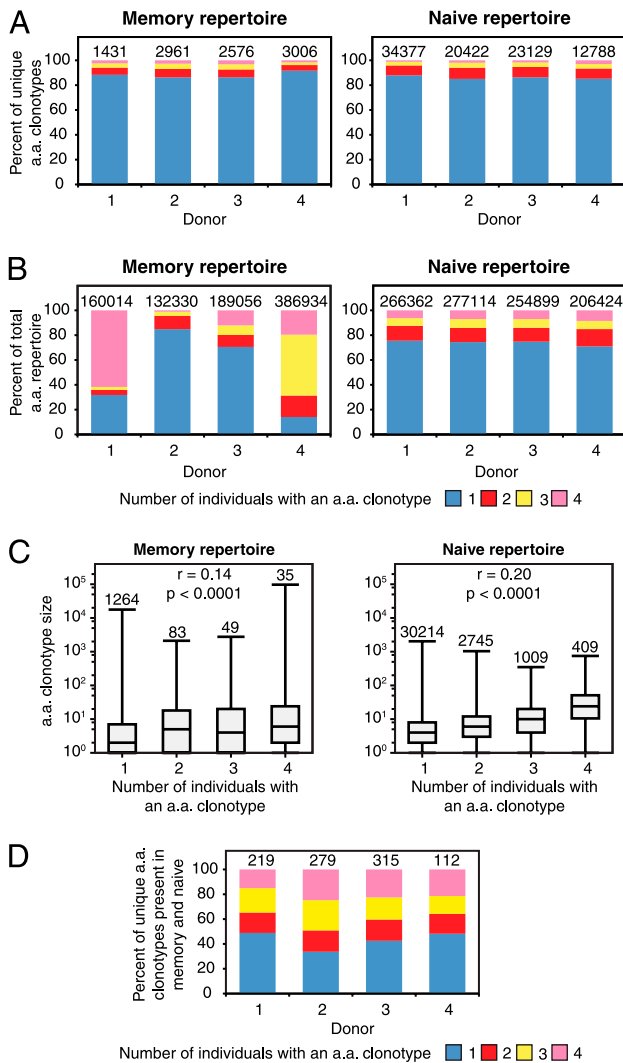


FIGURE 7. Interindividual sharing across all four donors of TCR β amino acid clonotypes in the memory and naive repertoires. The number of individuals in which a TCR β amino acid (a.a.) clonotype was observed was evaluated, regardless of whether it was found in the memory and/or naive pool. *A*, The proportions of unique TCR β amino acid clonotypes in the memory (*left*) and naive (*right*) repertoires of Donors 1–4 that were observed in one, two, three, or all four individuals. *B*, The proportions of the total memory and naive TCR β repertoires of Donors 1–4 that were attributable to TCR β amino acid clonotypes observed in one, two, three, or all four individuals. *C*, Relationship between TCR β amino acid clonotype size in the memory (*left*) and naive (*right*) repertoires of Donor 1 and the number of individuals in which the TCR β amino acid clonotype was present. Correlations are based on the Spearman rank test. *D*, The proportions of unique TCR β amino acid clonotypes in the memory and in the naive repertoires of Donors 1–4 that were observed in one, two, three, or all four individuals. In all panels, the total numbers of unique TCR β clonotypes are indicated for each group of data above the bar or box representing the data group.

and naive pools consisted of a mean of 56.7% (range: 51.1–66.3%) unique TCR β amino acid clonotypes that were also shared between individuals (Fig. 7*D*). Indeed, a substantial proportion of TCR β amino acid clonotypes present in the memory and naive pools within a donor were also found in more than two individuals. Thus, a higher degree of TCR β amino acid clonotype sharing was observed among clonotypes common to the memory and naive pools of an individual than among those unique to either pool.

Discussion

Although recent studies highlighted the importance of T cell repertoire composition for immune recognition of specific pathogens (3, 36–38) and the maintenance of immune efficacy with age (39), our current understanding of these processes is derived predominantly from analyses of small samples of epitope-specific repertoires. Next-generation sequencing technologies provide an opportunity to study complete memory and naive T cell repertoires in depth. In this study, we rigorously sorted the memory and naive CD8⁺ T cell populations from four donors and pyrosequenced the portions of the corresponding TCR β repertoires with *TRBV12-4/TRBJ1-2* gene rearrangements. We found that there is a high degree of overlap between the memory and naive repertoires within individuals; TCR β clonotypes common to the memory and naive pools tend to be more frequent in both pools; a substantial proportion of the memory and naive repertoires consist of TCR β clonotypes that are shared between individuals; shared TCR β clonotypes tend to have larger sizes in the memory and naive repertoires; and the process of convergent recombination is an important determinant of the relative TCR β clonotype size in the memory and naive repertoires.

Overall, these data provide insights into the complex and dynamic relationship between the memory and naive T cell populations, which involves thymic output into the naive pool, homeostatic maintenance of the peripheral repertoire, and the ongoing recruitment of T cells from the naive pool during Ag encounter (7, 8, 40). Although there were substantial differences in clonotype size and CDR3 length distribution between the memory and naive pools, a subset of TCR β clonotypes was common to both pools. Many of these TCR β clonotypes were relatively dominant in the memory and naive pools, with clonotype size in the memory pool being significantly associated with clonotype size in the naive pool. However, the more pronounced dominance hierarchy among memory TCR β clonotypes compared with naive TCR β clonotypes resulted in a surprisingly high degree of overlap of the total memory repertoire with the total naive repertoire within individuals.

The presence of particular TCR β clonotypes in the memory and naive pools evokes several interesting interpretations. One scenario is that naive T cells bearing the same TCR are entirely recruited into the memory pool during an Ag-specific response. The presence of those same TCR β clonotypes in the naive pool would then derive from thymic replenishment of the naive repertoire with identical TCR sequences. Indeed, recent studies in mice showed that, although governed by affinity for peptide–MHC (11, 41), recruitment of Ag-specific CD8⁺ T cells from the naive pool is markedly efficient (9). This scenario is supported by our observations that TCR β clonotypes common to the memory and naive pools tended to have larger sizes in both pools and are more efficiently generated by convergent recombination. Moreover, the high degree of interindividual TCR β clonotype sharing between the memory and naive pools strongly suggests that these TCR β clonotypes were initially present at a relatively high frequency in the naive TCR repertoires of each of the donors. Alternatively, it might be that not all cells with a particular TCR β clonotype are recruited from the naive pool by a particular Ag. This could occur as a result of suboptimal clonotypic TCR affinities for peptide–MHC, differential TCR α -chain pairing with altered fine specificity, or simply a lack of Ag encounter. Another interpretation is that asymmetric division and differentiation after Ag encounter result in T cells bearing the same TCR β clonotypes having memory and naive phenotypes. Such a model has been proposed in mice (42) and is supported by the overlap between memory and naive

repertoires at the level of nucleotide sequence. However, the observation that different nucleotide sequences in the memory and naive repertoires also encoded many common TCR β amino acid clonotypes makes it quite clear that these common TCR β clonotypes originated from different cells, which runs somewhat contrary to this explanation. Importantly, these various interpretations are not mutually exclusive and may all contribute to TCR β clonotype sharing between the memory and naive pools.

TCR β clonotypes that are shared between individuals are thought to play an important role in the efficacy of pathogen-specific responses and the control of infection (3, 36). Thus, an understanding of the mechanisms that determine interindividual TCR sharing in humans is likely salient to vaccine development. We found that a substantial proportion (>24.5%) of the naive TCR β amino acid repertoire within an individual was shared with at least one of the other donors in this study. Thus, the extent of TCR β sharing between larger groups of individuals should be potentially much greater. Indeed, studies in syngeneic mice showed that up to 27% of the peripheral repertoire of one naive mouse overlap with that of another (16). Our previous studies of epitope-specific TCR β repertoires suggested that a process of convergent recombination enables some TCR β clonotypes to be produced by V(D)J recombination in the thymus more frequently than others. Such TCR β clonotypes are predicted to be present at a greater frequency in the naive repertoire and have a greater likelihood of being shared between individuals (18–21). Our investigations of the relationship between the sharing of TCR β clonotypes between the donors in this study and clonotype sizes in the memory and naive pools provide strong evidence for convergent recombination as a molecular basis for interindividual TCR sharing.

We also determined to what extent the hierarchy of TCR β production frequency predicted by convergent recombination is modified by other processes, such as TCR α -chain pairing, thymic selection, and peripheral expansion, which generate the peripheral T cell pools. Using the number of nucleotide additions and the variety of nucleotide sequences encoding each CDR3 β amino acid sequence as key indicators of convergent recombination, we found that convergent recombination is a significant predictor of relative TCR β clonotype sizes in the memory and naive repertoires. This suggests that, at the level of the T cell population, the relative clonotype frequencies are substantially preserved through Ag selection and the formation of the memory pool. This finding is surprising given that the memory pool contains T cells that have responded to a variety of different Ags, both transient and persistent, with varying levels of immunodominance over the lifetime of each individual.

It is worth noting that there has been some confusion in the literature over the use of the term “convergent” in relation to the development of the T cell repertoire. Robins et al. (24) recently introduced the concept of “convergent evolution,” defined as “the possibility that a diverse set of TCRs rearranges in the thymus, and the positive and negative selection process favors the same lower diversity subset of TCRs in each individual.” Although Robins et al. (24) find no evidence of convergent evolution, their results clearly support our findings on the role of convergent recombination (illustrated in Fig. 4) in shaping the TCR repertoire, albeit using data with a more limited depth of sequencing, in terms of the number of sequences using a specific V and J gene combination.

Taken together, our data showed that the production frequency of TCR β clonotypes in the thymus, as predicted by convergent recombination, is a fundamental determinant of clonotype size in the memory and naive T cell pools. Furthermore, TCR β clonotype

size influences the overlap between the memory and naive TCR repertoires, as well as interindividual clonotype sharing. Interestingly, outliers in the data suggest the involvement of other factors in shaping the repertoire. That is, not all TCR β clonotypes predicted to be easily generated by convergent recombination were present in large numbers in the memory and/or naive repertoires of multiple individuals. The variety of TCR α -chains with which the TCR β -chains pair, the efficiency with which these TCR α -chains are made, the proportion of these TCR $\alpha\beta$ heterodimers that survive thymic selection, and the efficiency of their recruitment into the memory pool are all important factors that determine the presence and size of TCR β clonotypes in the memory and naive pools. Nonetheless, the associations identified in this study provide new insights into the mechanisms that shape the peripheral T cell repertoires and reveal the profound influence of differential clonotype production frequencies in the thymus as a consequence of the process of convergent recombination. This mechanistic understanding may prove invaluable for the design of effective T cell-based vaccines that aim to exploit established correlates of immune control (3, 43).

Acknowledgments

We thank D. Ambrozak, R. Nguyen, and S. Perfetto for expert assistance with flow cytometry, and M. Roederer and J. Yu of the Vaccine Research Center Flow Cytometry Core for production and conjugation of Abs.

Disclosures

The authors have no financial conflicts of interest.

References

- Nikolich-Zugich, J., M. K. Slika, and I. Messaoudi. 2004. The many important facets of T-cell repertoire diversity. *Nat. Rev. Immunol.* 4: 123–132.
- Davenport, M. P., D. A. Price, and A. J. McMichael. 2007. The T cell repertoire in infection and vaccination: implications for control of persistent viruses. *Curr. Opin. Immunol.* 19: 294–300.
- Price, D. A., T. E. Asher, N. A. Wilson, M. C. Nason, J. M. Brenchley, I. S. Metzler, V. Venturi, E. Gostick, P. K. Chattopadhyay, M. Roederer, et al. 2009. Public clonotype usage identifies protective Gag-specific CD8+ T cell responses in SIV infection. *J. Exp. Med.* 206: 923–936.
- Arstila, T. P., A. Casrouge, V. Baron, J. Even, J. Kanellopoulos, and P. Kourilsky. 1999. A direct estimate of the human alpha beta T cell receptor diversity. *Science* 286: 958–961.
- Arstila, T. P., A. Casrouge, V. Baron, J. Even, J. Kanellopoulos, and P. Kourilsky. 2000. Diversity of human alpha beta T cell receptors. *Science* 288: 1135.
- Mason, D. 1998. A very high level of crossreactivity is an essential feature of the T-cell receptor. *Immunol. Today* 19: 395–404.
- Snyder, C. M., K. S. Cho, E. L. Bonnett, S. van Dommelen, G. R. Shellam, and A. B. Hill. 2008. Memory inflation during chronic viral infection is maintained by continuous production of short-lived, functional T cells. *Immunity* 29: 650–659.
- Vezys, V., D. Masopust, C. C. Kambal, D. L. Barber, L. A. O'Mara, C. P. Larsen, T. C. Pearson, R. Ahmed, and A. E. Lukacher. 2006. Continuous recruitment of naive T cells contributes to heterogeneity of antiviral CD8 T cells during persistent infection. *J. Exp. Med.* 203: 2263–2269.
- van Heijst, J. W., C. Gerlach, E. Swart, D. Sie, C. Nunes-Alves, R. M. Kerkhoven, R. Arens, M. Correia-Neves, K. Schepers, and T. N. Schumacher. 2009. Recruitment of antigen-specific CD8+ T cells in response to infection is markedly efficient. *Science* 325: 1265–1269.
- Price, D. A., J. M. Brenchley, L. E. Ruff, M. R. Betts, B. J. Hill, M. Roederer, R. A. Koup, S. A. Migueles, E. Gostick, L. Wooldridge, et al. 2005. Avidity for antigen shapes clonal dominance in CD8+ T cell populations specific for persistent DNA viruses. *J. Exp. Med.* 202: 1349–1361.
- Zehn, D., S. Y. Lee, and M. J. Bevan. 2009. Complete but curtailed T-cell response to very low-affinity antigen. *Nature* 458: 211–214.
- Correia-Neves, M., C. Waltzinger, D. Mathis, and C. Benoist. 2001. The shaping of the T cell repertoire. *Immunity* 14: 21–32.
- Furmanski, A. L., C. Ferreira, I. Bartok, S. Dimakou, J. Rice, F. K. Stevenson, M. M. Millrain, E. Simpson, and J. Dyson. 2008. Public T cell receptor beta-chains are not advantaged during positive selection. *J. Immunol.* 180: 1029–1039.
- Bouneaud, C., Z. Garcia, P. Kourilsky, and C. Pannetier. 2005. Lineage relationships, homeostasis, and recall capacities of central- and effector-memory CD8 T cells in vivo. *J. Exp. Med.* 201: 579–590.
- Baron, V., C. Bouneaud, A. Cumano, A. Lim, T. P. Arstila, P. Kourilsky, L. Ferradini, and C. Pannetier. 2003. The repertoires of circulating human CD8

- (+) central and effector memory T cell subsets are largely distinct. *Immunity* 18: 193–204.
16. Bousso, P., A. Casrouge, J. D. Altman, M. Haury, J. Kanellopoulos, J. P. Abastado, and P. Kourilsky. 1998. Individual variations in the murine T cell response to a specific peptide reflect variability in naive repertoires. *Immunity* 9: 169–178.
 17. Turner, S. J., P. C. Doherty, J. McCluskey, and J. Rossjohn. 2006. Structural determinants of T-cell receptor bias in immunity. *Nat. Rev. Immunol.* 6: 883–894.
 18. Venturi, V., D. A. Price, D. C. Douek, and M. P. Davenport. 2008. The molecular basis for public T-cell responses? *Nat. Rev. Immunol.* 8: 231–238.
 19. Venturi, V., K. Kedzierska, D. A. Price, P. C. Doherty, D. C. Douek, S. J. Turner, and M. P. Davenport. 2006. Sharing of T cell receptors in antigen-specific responses is driven by convergent recombination. *Proc. Natl. Acad. Sci. USA* 103: 18691–18696.
 20. Venturi, V., H. Y. Chin, D. A. Price, D. C. Douek, and M. P. Davenport. 2008. The role of production frequency in the sharing of simian immunodeficiency virus-specific CD8+ TCRs between macaques. *J. Immunol.* 181: 2597–2609.
 21. Venturi, V., H. Y. Chin, T. E. Asher, K. Ladell, P. Scheinberg, E. Bornstein, D. van Bockel, A. D. Kelleher, D. C. Douek, D. A. Price, and M. P. Davenport. 2008. TCR β -chain sharing in human CD8+ T cell responses to cytomegalovirus and EBV. *J. Immunol.* 181: 7853–7862.
 22. Freeman, J. D., R. L. Warren, J. R. Webb, B. H. Nelson, and R. A. Holt. 2009. Profiling the T-cell receptor beta-chain repertoire by massively parallel sequencing. *Genome Res.* 19: 1817–1824.
 23. Robins, H. S., P. V. Campregher, S. K. Srivastava, A. Wachter, C. J. Turtle, O. Kahsai, S. R. Riddell, E. H. Warren, and C. S. Carlson. 2009. Comprehensive assessment of T-cell receptor beta-chain diversity in alphabeta T cells. *Blood* 114: 4099–4107.
 24. Robins, H. S., S. K. Srivastava, P. V. Campregher, C. J. Turtle, J. Andriesen, S. R. Riddell, C. S. Carlson, and E. H. Warren. 2010. Overlap and effective size of the human CD8+ T cell receptor repertoire. *Sci. Transl. Med.* 2: 47ra64.
 25. Chattopadhyay, P. K., D. A. Price, T. F. Harper, M. R. Betts, J. Yu, E. Gostick, S. P. Peretto, P. Goepfert, R. A. Koup, S. C. De Rosa, et al. 2006. Quantum dot semiconductor nanocrystals for immunophenotyping by polychromatic flow cytometry. *Nat. Med.* 12: 972–977.
 26. Lefranc, M. P., V. Giudicelli, C. Ginestoux, J. Bodmer, W. Müller, R. Bontrop, M. Lemaître, A. Malik, V. Barbić, and D. Chaume. 1999. IMGT, the international ImMunoGeneTics database. *Nucleic Acids Res.* 27: 209–212.
 27. Margulies, M., M. Egholm, W. E. Altman, S. Attiya, J. S. Bader, L. A. Bemben, J. Berka, M. S. Braverman, Y. J. Chen, Z. Chen, et al. 2005. Genome sequencing in microfabricated high-density picolitre reactors. [Published erratum appears in 2006 *Nature*: 441: 326–327.] *Nature* 437: 376–380.
 28. Douek, D. C., M. R. Betts, J. M. Brenchley, B. J. Hill, D. R. Ambrozak, K. L. Ngai, N. J. Karandikar, J. P. Casazza, and R. A. Koup. 2002. A novel approach to the analysis of specificity, clonality, and frequency of HIV-specific T cell responses reveals a potential mechanism for control of viral escape. *J. Immunol.* 168: 3099–3104.
 29. Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215: 403–410.
 30. Weekes, M. P., M. R. Wills, K. Mynard, A. J. Carmichael, and J. G. Sissons. 1999. The memory cytotoxic T-lymphocyte (CTL) response to human cytomegalovirus infection contains individual peptide-specific CTL clones that have undergone extensive expansion in vivo. *J. Virol.* 73: 2099–2108.
 31. Khan, N., M. Cobbold, R. Keenan, and P. A. Moss. 2002. Comparative analysis of CD8+ T cell responses against human cytomegalovirus proteins pp65 and immediate early 1 shows similarities in precursor frequency, oligoclonality, and phenotype. *J. Infect. Dis.* 185: 1025–1034.
 32. Khan, N., N. Shariff, M. Cobbold, R. Bruton, J. A. Ainsworth, A. J. Sinclair, L. Nayak, and P. A. Moss. 2002. Cytomegalovirus seropositivity drives the CD8 T cell repertoire toward greater clonality in healthy elderly individuals. *J. Immunol.* 169: 1984–1992.
 33. Trautmann, L., M. Rimbart, K. Echasserieau, X. Saulquin, B. Neveu, J. Dechanet, V. Cerundolo, and M. Bonneville. 2005. Selection of T cell clones expressing high-affinity public TCRs within Human cytomegalovirus-specific CD8 T cell responses. *J. Immunol.* 175: 6123–6132.
 34. Day, E. K., A. J. Carmichael, I. J. ten Berge, E. C. Waller, J. G. Sissons, and M. R. Wills. 2007. Rapid CD8+ T cell repertoire focusing and selection of high-affinity clones into memory following primary infection with a persistent human virus: human cytomegalovirus. *J. Immunol.* 179: 3203–3213.
 35. Fazilleau, N., J. P. Cabaniols, F. Lemaître, I. Motta, P. Kourilsky, and J. M. Kanellopoulos. 2005. Valpha and Vbeta public repertoires are highly conserved in terminal deoxynucleotidyl transferase-deficient mice. *J. Immunol.* 174: 345–355.
 36. Price, D. A., S. M. West, M. R. Betts, L. E. Ruff, J. M. Brenchley, D. R. Ambrozak, Y. Edghill-Smith, M. J. Kuroda, D. Bogdan, K. Kunstman, et al. 2004. T cell receptor recognition motifs govern immune escape patterns in acute SIV infection. *Immunity* 21: 793–803.
 37. Cornberg, M., A. T. Chen, L. A. Wilkinson, M. A. Brehm, S. K. Kim, C. Calcagno, D. Gherzi, R. Puzone, F. Celada, R. M. Welsh, and L. K. Selin. 2006. Narrowed TCR repertoire and viral escape as a consequence of heterologous immunity. *J. Clin. Invest.* 116: 1443–1456.
 38. Meyer-Olson, D., N. H. Shoukry, K. W. Brady, H. Kim, D. P. Olson, K. Hartman, A. K. Shintani, C. M. Walker, and S. A. Kalams. 2004. Limited T cell receptor diversity of HCV-specific T cell responses is associated with CTL escape. *J. Exp. Med.* 200: 307–319.
 39. Cicin-Sain, L., S. Smyk-Pearson, N. Currier, L. Byrd, C. Koudelka, T. Robinson, G. Swarbrick, S. Tackitt, A. Legasse, M. Fischer, et al. 2010. Loss of naive T cells and repertoire constriction predict poor response to vaccination in old primates. [Published erratum in 2010 *J. Immunol.* 185: 4509.] *J. Immunol.* 184: 6739–6745.
 40. Nikolich-Zugich, J. 2008. Ageing and life-long maintenance of T-cell subsets in the face of latent persistent infections. *Nat. Rev. Immunol.* 8: 512–522.
 41. Malherbe, L., C. Hausl, L. Teyton, and M. G. McHeyzer-Williams. 2004. Clonal selection of helper T cells is determined by an affinity threshold with no further skewing of TCR binding properties. *Immunity* 21: 669–679.
 42. Chang, J. T., V. R. Palanivel, I. Kinjyo, F. Schambach, A. M. Intlekofer, A. Banerjee, S. A. Longworth, K. E. Vinup, P. Mrass, J. Oliaro, et al. 2007. Asymmetric T lymphocyte division in the initiation of adaptive immune responses. *Science* 315: 1687–1691.
 43. Douek, D. C., P. D. Kwong, and G. J. Nabel. 2006. The rational design of an AIDS vaccine. *Cell* 124: 677–681.