

RESEARCH ARTICLE

Impact of HLA type, age and chronic viral infection on peripheral T-cell receptor sharing between unrelated individuals

Sarah A. Johnson , Spencer L. Seale , Rachel M. Gittelman , Julie A. Rytlewski , Harlan S. Robins , Paul A. Fields *

Adaptive Biotechnologies, Seattle, Washington, United States of America

* pfields@adaptivebiotech.com



Abstract

The human adaptive immune system must generate extraordinary diversity to be able to respond to all possible pathogens. The T-cell repertoire derives this high diversity through somatic recombination of the T-cell receptor (TCR) locus, a random process that results in repertoires that are largely private to each individual. However, factors such as thymic selection and T-cell proliferation upon antigen exposure can affect TCR sharing among individuals. By immunosequencing the TCR β variable region of 426 healthy individuals, we find that, on average, fewer than 1% of TCR β clones are shared between individuals, consistent with largely private TCR β repertoires. However, we detect a significant correlation between increased HLA allele sharing and increased number of shared TCR β clones, with each additional shared HLA allele contributing to an increase in ~0.01% of the total shared TCR β clones, supporting a key role for HLA type in shaping the immune repertoire. Surprisingly, we find that shared antigen exposure to CMV leads to fewer shared TCR β clones, even after controlling for HLA, indicative of a largely private response to major viral antigenic exposure. Consistent with this hypothesis, we find that increased age is correlated with decreased overall TCR β clone sharing, indicating that the pattern of private TCR β clonal expansion is a general feature of the T-cell response to other infectious antigens as well. However, increased age also correlates with increased sharing among the lowest frequency clones, consistent with decreased repertoire diversity in older individuals. Together, all of these factors contribute to shaping the TCR β repertoire, and understanding their interplay has important implications for the use of T cells for therapeutics and diagnostics.

OPEN ACCESS

Citation: Johnson SA, Seale SL, Gittelman RM, Rytlewski JA, Robins HS, Fields PA (2021) Impact of HLA type, age and chronic viral infection on peripheral T-cell receptor sharing between unrelated individuals. PLoS ONE 16(8): e0249484. <https://doi.org/10.1371/journal.pone.0249484>

Editor: Jianming Tang, University of Alabama at Birmingham, UNITED STATES

Received: March 23, 2021

Accepted: July 29, 2021

Published: August 30, 2021

Copyright: © 2021 Johnson et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its [Supporting information](#) files. Previously published immunosequencing data is available at <https://clients.adaptivebiotech.com/pub/Emerson-2017-NatGen>.

Funding: Adaptive Biotechnologies provided support in the form of salaries for authors SJ,SS, RG, JR, PF and HR but did not have any additional role in the study design, data collection and

Introduction

T cells make up a key component of the adaptive immune response and allow the body to respond to the diverse range of pathogens it may encounter. The adaptive immune system of a healthy adult includes up to 10^{15} highly diverse T cells [1, 2]. Antigen recognition depends on both T-cell specificity and the major histocompatibility complex (MHC), encoded by the polymorphic human leukocyte antigen (HLA) gene complex, presenting the antigen. The vast

analysis, decision to publish, or preparation of the manuscript.

Competing interests: SJ,SS,RG,PF and HR own stocks or shares and are paid employees of Adaptive Biotechnologies. JR is a former employee of Adaptive Biotechnologies. HR holds patents associated with Adaptive Biotechnologies including 9,150,095; 9,181,591; 9,371,558; 9,809,813; 10,221,461. This does not alter our adherence to PLOS ONE policies on sharing data and materials.

diversity of possible T-cell receptors (TCR) is generated by the recombination of genes in the third complementary determining regions (CDR3) within the α and β chains of a TCR. In the recombination process of the β chain, loci of the variable (V), diversity (D), and joining (J) regions are randomly spliced together with non-templated insertions and deletions occurring between each junction, resulting in up to 10^{11} possible sequences [3, 4]. Recombination of the TCR α chain includes only V and J gene segments, resulting in fewer possible rearrangements and making the TCR β chain a more suitable target for identifying unique T cells. Recombined T cells that are capable of recognizing cognate antigens when presented by the MHC are selected for maturation and released from the thymus [5]. As a result, VDJ recombination, HLA restriction, and antigen exposure collectively contribute to a largely private TCR β repertoire.

Despite the large space of potential TCR β rearrangements, the existence of public clones has been well characterized and occurs more frequently than would be expected by chance [6–8]. Public TCR β clones can arise either via convergent recombination—due to highly-probable rearrangements, or via convergent selection—due to proliferation after common antigen exposure. Biases in VDJ recombination and junctional indel patterns suggest that CDR3 sequences that more closely resemble the germline-encoded nucleotide sequence are more likely to occur and thus be shared between individuals [9, 10]. Additionally, publicly expanded antigen-specific T cells to a variety of pathogens including CMV and SARS-CoV-2 have been identified [11–14], supporting convergent selection. Together, both of these processes contribute to the existence of public clones, but much remains to be discovered about how factors such as additional antigen exposure and HLA type modify them.

Understanding the forces underlying the inherent diversity of the TCR β repertoire of an individual and the public sharing of TCR β clones is of great clinical interest. Decreased diversity of TCR β clones in older individuals has been associated with reduced immune function [15]. Increased evolutionary divergence of HLA class I alleles has been correlated with better responses to immune checkpoint inhibitors in cancer patients [16]. Public TCR β s that respond to specific antigens have the potential to be used diagnostically to identify an individual's antigen exposure [11–13] and have therapeutic potential as next-generation CAR-T cells [17, 18]. Additionally, HLA-restriction of TCR β s is clinically relevant to evaluating histocompatibility for the purposes of bone-marrow and solid organ transplants [19, 20]. Continued development of such immunological and medical advancements depends on fully understanding the determinants that shape public versus private TCR β s.

To explore the influence of biological and environmental forces on the dynamics of TCR β clone sharing, we utilized a published set of TCR β repertoires from healthy human subjects [8, 11]. We assessed the role of HLA zygosity, HLA allele sharing, CMV exposure, and age both in shaping the immune repertoire, and in the sharing of TCR β clones between individuals. By analyzing both the highest- and lowest- (single-copy) frequency TCR β clones, we identified a consistent positive association between numbers of shared HLA alleles and TCR β clones. Additionally, we found that CMV exposure and increased age result in more private TCR β repertoires, in particular among high frequency clones. Our results demonstrate the impact of both age and HLA type on TCR β clone generation and maturation, and indicate that pathogenic antigen exposure leads to the expansion of largely private and HLA-restricted TCR β clones.

Results

Determinants of TCR β repertoire diversity in healthy Individuals

We first investigated how HLA type influences diversity of the TCR β repertoire. The divergent allele hypothesis suggests that greater diversity of HLA alleles leads to a greater diversity of

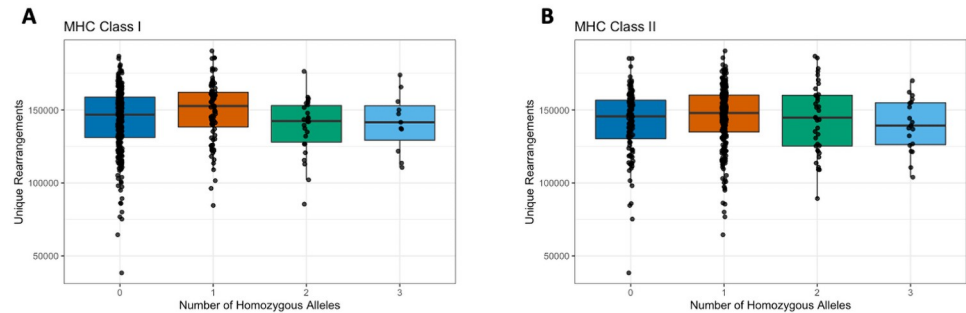


Fig 1. Correlation of TCR β repertoire richness and HLA zygosity. TCR β richness was quantified as the number of unique nucleotide rearrangements in a repertoire computationally down-sampled to 200,000 productive templates. HLA zygosity at neither (A) class I loci nor (B) class II loci correlated with diversity (Spearman rho = 0.035, p = 0.47 and rho = 0.021, p = 0.66 respectively).

<https://doi.org/10.1371/journal.pone.0249484.g001>

presented peptides [21, 22], and thus potentially greater diversity of the TCR β repertoire. To address this question, we utilized previously published TCR β repertoire data from over 600 individuals. We restricted analysis to individuals in the cohort with full 4-digit resolution at 6 HLA loci (HLA-A, HLA-B, HLA-C, HLA-DRB1, HLA-DQB1, HLA-DPA1) and known CMV serostatus. To control for any technical variation due to differences in T-cell fraction or input material, we further excluded samples with less than 200,000 total T cells. After filtering, our analyses included repertoires from 426 subjects. We determined the zygosity of these individuals at the included HLA loci and quantified repertoire diversity with two metrics: richness and clonality. The richness of each repertoire was calculated as the number of unique TCR β nucleotide rearrangements after computationally down-sampling all repertoires to 200,000 productive templates. Simpson clonality was also calculated for each repertoire to assess the dominance of high frequency clones in the repertoire. Within this cohort, 44% of individuals are homozygous at the HLA-DPA1 loci, and heterozygous at all other loci (S1 Fig).

In contrast to our expectations, there was not a significant relationship between the number of homozygous class I or class II HLA alleles and the richness or clonality of an individual's repertoire in this cohort (Fig 1A and 1B, and S2A and S2B Fig). Similarly, there is not an overall correlation across all 6 included loci (S2C and S2D Fig). This suggests that, in addition to not influencing the overall richness, HLA zygosity does not influence the extent of oligo-clonal dominance in the repertoire. One limitation here is the requirement for an exact match of HLA alleles, which does not account for the potential similarity between HLA alleles that could impact the overall repertoire diversity.

We additionally examined the influence of age and CMV exposure on the diversity of the TCR β repertoire. Consistent with prior studies, we found that CMV+ individuals have increased clonality and decreased repertoire richness, indicating a more focused repertoire after CMV exposure (S3A and S3B Fig). We also observed a negative correlation between age and repertoire diversity, including after controlling for CMV status (S3C and S3D Fig) [23–25]. Within our cohort, we did not see a correlation between overall HLA zygosity and diversity among either CMV+ or CMV- individuals (S2E and S2F Fig). Taken together, these data demonstrate that the diversity of a person's TCR β repertoire is largely independent of HLA zygosity and may be driven more by age and exposure to antigens.

Healthy individuals with more shared HLA alleles share more TCR β clones

While HLA zygosity does not affect the overall diversity of a TCR β repertoire, positive thymic selection selects for TCR β clones with affinity to the specific set of HLA alleles of an individual

[26]. We thus hypothesized that shared HLA alleles between unrelated individuals may contribute to the sharing of unique clones. The number of shared clones between each pairwise combination of repertoires was determined by comparing TCR β clones by their functional identity (the V gene family, CDR3 amino acid sequence, and J gene). This allows clone sharing to be detected when different individuals generate TCR β clones with the same specificity but distinct nucleotide sequences through VDJ recombination. An exact match was required to be considered a shared TCR β clone. The HLA allele sharing between individuals was determined regardless of HLA zygosity, so that two individuals homozygous for a specific HLA allele were considered to share two alleles.

We saw a significant correlation between the number of HLA alleles shared between two individuals and the number of shared unique down-sampled TCR β clones ($p < 0.001$, Mantel test, Fig 2A). Using a linear mixed-effects model, individuals with no HLA alleles in common shared on average 1,884 of their 200,000 down-sampled clones (0.94%), while each additional shared HLA allele resulted in an increase of 14 shared clones. This suggests a baseline population of public clones found across all subjects regardless of HLA type, consistent with the existence of bystander public clones with high generation probability, in addition to a subset of clones that are shared based on HLA type [10, 27]. While there was a similar relationship between shared alleles and shared TCR β clones within both class I and class II alleles, this correlation was slightly stronger for HLAII alleles (Fig 2B and 2C). This is consistent with CD4+ T cells (which bind class II alleles) making up a greater fraction of the repertoire than CD8 + T cells (which bind class I alleles) [28].

Sharing of the low frequency and expanded TCR β clones is impacted by HLA type

We next wanted to characterize the extent to which the number of shared HLA alleles relates to the sharing of both expanded and low frequency clones. While immunosequencing cannot distinguish naïve vs. memory T cells, we hypothesized that many singletons (TCR β clones seen once in a sample) correspond to naïve clones, while the highest frequency rearrangements correspond to clones that have likely expanded in response to prior antigen exposure. To characterize the singleton repertoire, we randomly selected 80,000 singletons from the repertoire of each individual. Singleton clones were selected by nucleotide sequence to best capture rearrangements that were present once in the full repertoire of an individual, but compared between individuals using functional identity. While this process may exclude rearrangements with high generation probability that are present multiple times in an individual [29], it captures the lowest frequency clones in a repertoire. Similar to what we observed when analyzing the down-sampled repertoires, there was a significant correlation between the number of shared HLA alleles and number of shared singletons between individuals ($p < 0.001$, Mantel test, Fig 2D). Notably, individuals with no shared alleles still shared on average 613 of 80,000 (0.77%) singletons. Since T cells that are present once in a repertoire are less likely to have encountered antigens, this overlap of singletons between individuals may be attributed to convergent recombination of frequently-generated TCR β rearrangements. Each shared HLA allele was correlated with an increase of 6 shared singletons, consistent with a role for HLA type in the maturation of the TCR β repertoire. However, even when individuals shared 10 HLA alleles, their repertoires were still largely private, sharing on average 660 of 80,000 singletons, which is still less than 1%. Thus, while HLA type plays an important role in the maturation of T cells, random VDJ somatic recombination results in most individuals having a highly private TCR β repertoire.

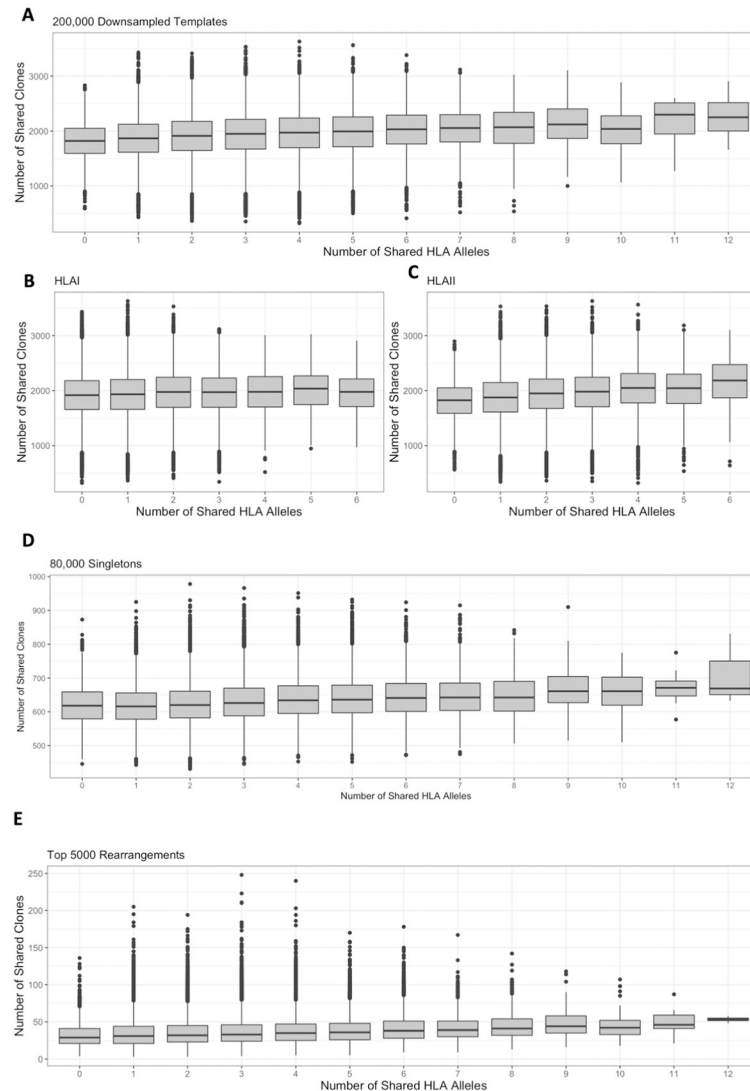


Fig 2. TCR β clone sharing is correlated with HLA allele sharing. Among repertoires computationally down-sampled to 200,000 productive templates, clone sharing positively correlated with increased numbers of (A) shared HLA alleles overall (Mantel rho = 0.17, $p < 1e-3$, linear mixed effects model slope = 14, intercept = 1884), and of both HLA (B) class I (Mantel rho = 0.07, $p < 1e-3$, linear mixed effects model slope = 9, intercept = 1920) and (C) class II alleles (Mantel rho = 0.19, $p < 1e-3$, linear mixed effects model slope = 21, intercept = 1881). (D) The number of singletons shared between two individuals was positively and significantly correlated with the number of HLA alleles shared (Mantel rho = 0.14, $p < 1e-3$, linear mixed effects model slope = 6, intercept = 613). (E) The number of clones shared among the top 5,000 unique rearrangements was positively and significantly correlated with the number of HLA alleles shared between two individuals (Mantel rho = 0.13, $p < 1e-3$, linear mixed effects model slope = 1.3, intercept = 34).

<https://doi.org/10.1371/journal.pone.0249484.g002>

To characterize the expanded repertoire of each subject, we selected the top 5,000 unique clones by frequency from each full repertoire prior to down-sampling. As chronic infections such as CMV increase the clonality of TCR β repertoires [11, 30] (S3A Fig), we selected a common number of rearrangements to mitigate the effect of a few largely expanded clones. Similar to the results among the down-sampled and singleton repertoires, there was a significant relationship between the number of HLA alleles shared and number of high-frequency clones shared ($p < 0.001$, Mantel test, Fig 2E). On average, individuals that shared no HLA alleles

shared ~34 of their top 5,000 clones (0.68%), with each additional shared allele leading to an additional 1.3 shared clones. Together, these findings suggest that while the HLA type of an individual does not impact the overall diversity of their TCR β repertoire, it does have a small but significant impact on the specific TCR β clones selected for maturation, as well as on those that expand in response to antigens.

Shared HLA alleles increases sharing of rare TCR β clones

Given the role of convergent recombination in influencing TCR β clone sharing, we next sought to assess whether HLA type influences sharing of clones with high generation probability differently than that of those with lower generation probability. To do this, we determined the generation probabilities of all TCR β rearrangements using OLGA [9], which correlated strongly with CDR3 length and the number of indels (S5A–S5C Fig). To distinguish rare from common TCR β clones, we established a cutoff point ($P_{\text{gen}} = 1.66\text{e-}09$) by finding the intersection point of generation probabilities between public TCR β clones that occurred in more than one subject and private TCR β clones found in only one subject (Fig 3A, S4D Fig). Each TCR β clone was thus classified as “rare,” with generation probabilities lower than this cutoff, or “common” otherwise. Sharing of each subset of clones between individuals was examined in both the down-sampled, singleton, and top 5,000 rearrangements repertoires. While sharing of both common and rare clones were significantly correlated with increased sharing of HLA alleles, the relationship was stronger among rare clones (Mantel rho 0.28) compared to common clones (Mantel rho 0.12) ($p < 0.001$, Mantel test, Fig 3B and 3C). This pattern persisted within the top 5,000 rearrangements (S4E and S4F Fig), though this analysis is limited by the small number of rare high frequency clones. As rare TCR β clones are those with rearrangements not commonly generated during VDJ recombination, these results suggest that their selection for maturation may be modulated by HLA genotype.

Interestingly, and in contrast to the down-sampled repertoires, the singleton TCR β repertoires exhibited similar positive associations between HLA allele sharing and both shared common (Mantel rho = 0.14) and rare (Mantel rho = 0.18) clones ($p < 0.001$, Mantel test, Fig 3D and 3E). These results suggest that HLA allele sharing affects both rare and common TCR β clone sharing, but the degree of impact depends on the frequencies of those clones within an individual repertoire. TCR β repertoire frequency is largely determined by antigen exposure, suggesting that expansion of rare clones may increase clone sharing among individuals with matched HLA types, although these clones only make up a small proportion of the overall repertoire.

CMV negative individuals share more TCR β clones than CMV positive individuals

Our prior analyses focused on clone sharing independent of antigen exposure. We next sought to determine whether common viral antigenic exposure would lead to differences in clone sharing. We focused our analysis on CMV, as our analysis cohort was split roughly equally between CMV- and CMV+ individuals (i.e. subjects with past CMV infection) [11]. We hypothesized that CMV+ individuals may share more clones than CMV- individuals. Surprisingly, we observed the opposite, with CMV- individuals sharing more down-sampled TCR β clones than CMV+ subjects (Fig 4A), and significantly more clones than a permuted null distribution (S5 Fig; permuted $P = 0.002$). The pattern of increased clone sharing between CMV- individuals persisted among both rare and common TCR β clones, regardless of the number of HLA clones shared (S6A and S6B Fig). This suggests that CMV infection leads to significantly fewer shared clones between individuals.

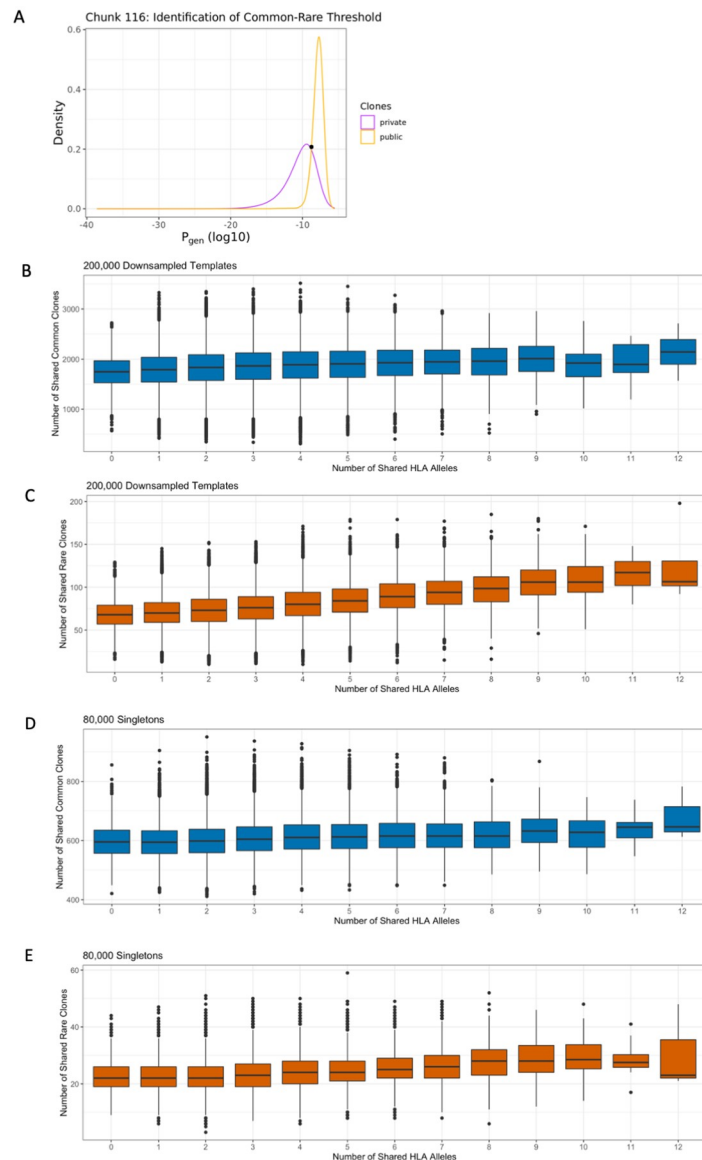


Fig 3. TCR β clone sharing among common and rare clones. Clones were classified as common versus rare by identifying a generation probability cutoff point between clones occurring once and those occurring more than once in 31 different “chunks” representing the entirety of TCR β rearrangements observed in the down-sampled data. (A) Intersection point from a single representative “chunk.” The number of shared HLA alleles was significantly correlated with clone sharing among both (B) common (Mantel rho = 0.12, $p < 1e-3$) and (C) rare clones (Mantel rho = 0.28, $p < 1e-3$). Similarly, the number of shared HLA alleles was significantly correlated with sharing of singletons among (D) common (Mantel rho = 0.14, $p < 1e-3$) and (E) rare (Mantel rho = 0.18, $p < 1e-3$) clones.

<https://doi.org/10.1371/journal.pone.0249484.g003>

CMV+ individuals have significantly higher clonality and lower richness than CMV- individuals (S3A and S3B Fig), effectively decreasing the number of unique clones that could be shared in CMV+ individuals [25]. To account for this disparity in diversity, we again looked at clone sharing among the top 5,000 unique TCR β rearrangements. There was also greater sharing of top TCR β clones between CMV- individuals compared to CMV+ individuals (Fig 4B) and compared to the null (permuted $P = 0.002$). This suggests that TCR β repertoires become more private after CMV antigen exposure, likely through the expansion of mostly private

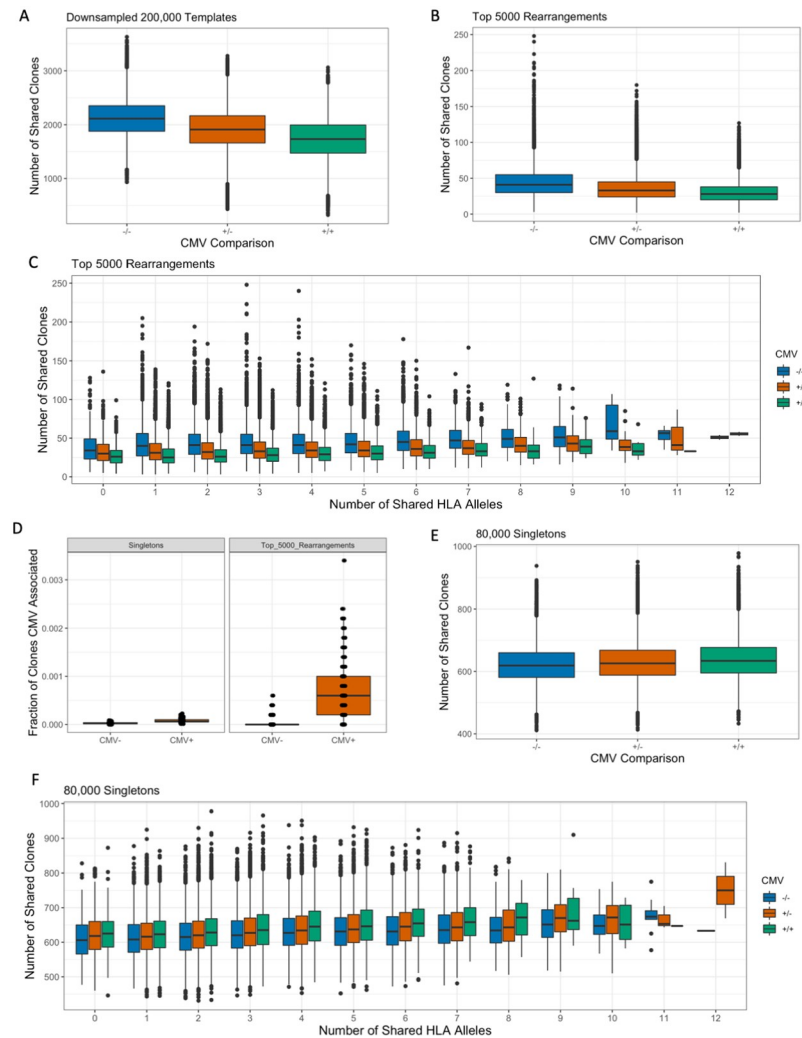


Fig 4. CMV status influences TCR β clone sharing. (A) CMV⁻ individuals shared more TCR β clones within their down-sampled repertoires compared to CMV⁺ individuals. (B) CMV⁻ individuals shared more of their top 5,000 rearrangements compared to CMV⁺ individuals. (C) CMV⁻ individuals shared more top rearrangements than CMV⁺ individuals do, regardless of the number of HLA alleles shared. (D) There was a greater fraction of clones in CMV⁺ individuals that were among the 164 CMV associated clones identified in Emerson et al. [11] compared to CMV⁻ individuals, among both singletons and top 5,000 rearrangements (Wilcoxon rank sum tests, $p < 2.2e-16$ for both). (E) There was a similar number of singletons shared among CMV⁻ and CMV⁺ individuals. (F) CMV⁺ individuals with a common number of shared HLA alleles shared slightly more singletons than CMV⁻ individuals do.

<https://doi.org/10.1371/journal.pone.0249484.g004>

clones. Notably, this trend is consistent independent of shared HLA alleles (Fig 4C). Despite the overall decrease in shared TCR β clones compared to CMV⁻ individuals, CMV⁺ individuals showed significantly greater enrichment in their top 5,000 clones of CMV-associated TCR β clones, as identified in Emerson et al. (Fig 4D, $P < 2.2e-16$, Unpaired Wilcoxon test). This suggests that while some public clones expand based on common viral antigen exposure, most of the expanded antigen-specific clones are private to a given repertoire.

In contrast to the most expanded clones, individuals shared a comparable number of singletons regardless of CMV status (Fig 4E). We did observe a significant enrichment of CMV-associated clones among the singletons of CMV⁺ individuals (Fig 4D, $P < 2.2e-16$, Unpaired Wilcoxon test), which may reflect some antigen exposed TCR β clones among the singletons.

When stratified by shared HLA alleles, CMV+ individuals share slightly more singletons than CMV- individuals (Fig 4F), which can likely be explained by differences in age between CMV+ and CMV- individuals. We observed that CMV+ individuals in this cohort tend to be older than CMV- individuals (S3E Fig), and that older individuals had a greater proportion of common singletons than younger individuals (S7C and S9C Figs), making increased singleton sharing between CMV+ individuals consistent with the less diverse and more public singleton repertoires of older individuals [15].

TCR β repertoires become increasingly private with age

Given the association between CMV status and TCR β clone sharing, we hypothesized that antigen exposure in general may lead to more private TCR β repertoires. Older individuals are likely to have encountered more antigens over the course of their life, consistent with more clonal repertoires in older subjects (S3F Fig) [31]. To examine whether age, as a proxy for antigen exposure, is a contributing factor in clone sharing, we divided individuals into above/equal to or below the median age (42 years) and compared TCR β clone sharing within the down-sampled, singleton, and top 5,000 rearrangements repertoires. We found that older individuals shared fewer down-sampled clones than younger individuals (median 1842 vs. 2136) (Fig 5A), and significantly fewer clones than the null (permuted $P = 0.002$). Since CMV+ individuals tend to be older (S3E Fig), we also looked at the impact of age within CMV+ and CMV- individuals. Older individuals continued to share fewer clones independent of CMV status (Fig 5B). Additionally, clone sharing between younger individuals continued to be

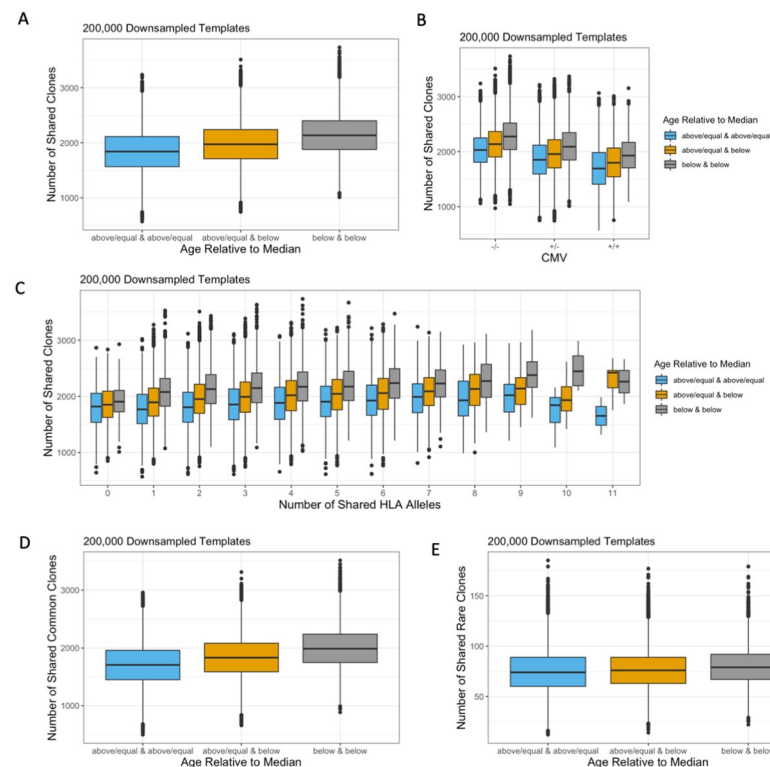


Fig 5. Impact of age on TCR β clone sharing. Individuals were stratified by the median age (42). (A) Younger individuals shared more clones in their down-sampled repertoires compared to older individuals, independent of (B) CMV exposure and (C) shared HLA alleles. Among (D) common and (E) rare clones, younger individuals shared more clones in their down-sampled repertoires than older individuals.

<https://doi.org/10.1371/journal.pone.0249484.g005>

greater than clone sharing between older individuals regardless of the number of HLA alleles shared (Fig 5C). We observe the same trend of younger individuals sharing more clones within the top 5,000 rearrangements (S8A and S8B Fig). This suggests that exposure to more pathogenic antigens, such as CMV, results in expansion of a set of largely private TCR β clones.

To evaluate whether age impacts shared clones differently based on their generative probability, we again split up clonotypes into rare and common subsets using the previously defined cutoff (Fig 3A, S4D Fig). Older individuals shared fewer common clones than younger individuals (median 1707 vs. 1989) (Fig 5D), and significantly fewer common clones than the null (permuted $P = 0.002$). Older individuals additionally shared fewer rare clones than younger individuals (median 74 vs 79) (Fig 5E), although not significantly fewer rare clones than the null (permuted $P > 0.05$). Within the top 5,000 rearrangements, older subjects continue to share fewer common clones than younger subjects (median 9 vs. 18) (S8C Fig), and significantly fewer than the null (permuted $P = 0.002$). Rare clone sharing within the top 5,000 rearrangements (S8D Fig) is unevaluable due to the limited sample size of these subsets. This demonstrates that older individuals share fewer expanded TCR β clones independent of generation probability.

Older individuals share more singletons than younger individuals

Next, we examined how age shapes the sharing of singletons. In contrast to the down-sampled repertoire, older individuals shared more singletons than younger individuals (median 634 vs 617) (Fig 6A), and significantly more clones than the null (permuted $P < 0.05$). This trend

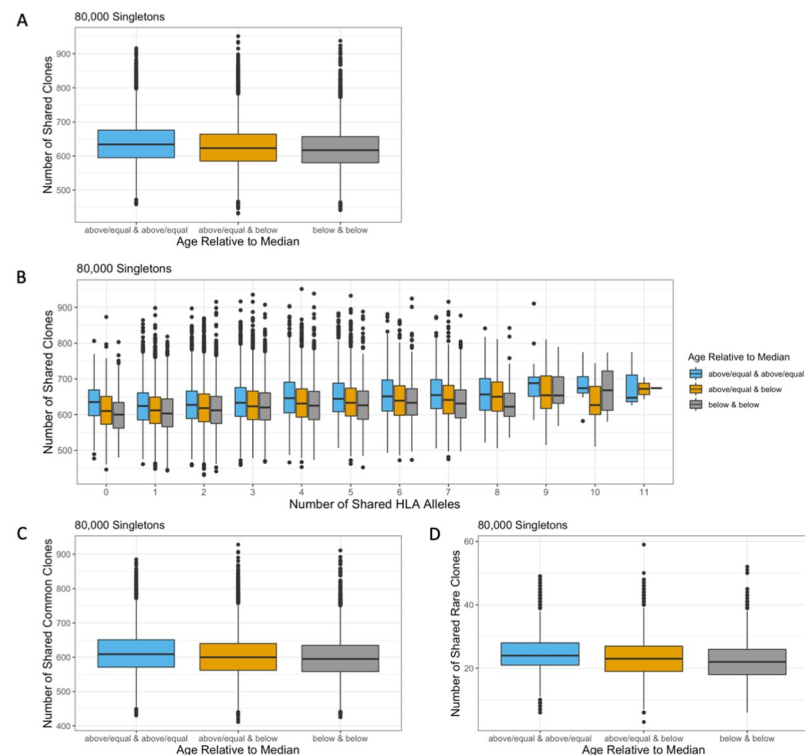


Fig 6. Impact of age on sharing of singleton TCR β clones. Individuals were stratified by the median age (42). (A) Older individuals shared more singletons than younger individuals. (B) Age impacted the sharing of singletons regardless of the number of HLA alleles. Older individuals shared more (C) common and (D) rare singletons than younger individuals.

<https://doi.org/10.1371/journal.pone.0249484.g006>

held true regardless of the number of HLA alleles shared (Fig 6B). Additionally, older individuals shared more singletons in both the rare and the common subsets (Fig 6C and 6D), and significantly more rare and common singletons than the null (permuted (rare) $P < 0.05$ and (common) $P = 0.002$). While the expansion of unique clones in response to antigen exposure can explain the decreased sharing in the down-sampled repertoire of older individuals, greater TCR β overlap within the singleton repertoire of older individuals is consistent with the diminished diversity of newly generated TCR β clones in aging immune systems caused by decreased thymopoiesis [32, 33]. This is furthermore supported by the proportion of rare and common clones within older and younger individuals. Younger and older individuals have a comparable proportion of rare clones among their down-sampled repertoire ($p = 0.5$, Wilcox test) (S9A Fig). However, among the 5000 most abundant TCR β rearrangements, age was significantly correlated with lower generation probabilities (S7B Fig, Spearman rho = -0.37, $p = 8.6e-15$) and a higher proportion of rare TCR β clones (S9B Fig, $p = 3.6e-10$, Wilcox test). This suggests that antigen-specific TCR β clones, including those with low generation probability, proliferate due to antigen exposure and thus are enriched in older individuals. In contrast, median generation probabilities of singletons increased significantly with age (S7C Fig, Spearman rho = 0.17, $p = 0.0015$), and older individuals had a significantly lower proportion of rare singletons compared to younger individuals (S9C Fig, $p = 0.03$, Wilcox test). Together, these findings support a model whereby antigen exposure to pathogens such as CMV during the life of an individual leads to expansion of mostly rare, private clones, while age-related changes in the thymus reduce the complexity and diversity of singletons among older individuals.

Discussion

The composition of an individual's TCR β repertoire is the result of many forces, both biological and environmental [34]. While several of these factors have been well characterized, their interaction within the context of HLA type has not been thoroughly explored. Here we looked at the extent to which HLA type, antigen exposure, and age impact the sharing of TCR β clones between unrelated individuals. We assessed these factors within the context of TCR β generation probabilities to determine how clones of varying prevalence are affected by them. By separately surveying clone sharing among singletons and high-frequency clones, we distinguished between HLA-restricted selection and HLA-restricted expansion. Our data support a broad model of TCR β proliferation wherein HLA allele sharing increases sharing of TCR β clones across all frequencies, while increased age correlates with the generation of more public low frequency TCR β clones and antigenic exposure results in a more private expanded TCR β repertoire.

Consistent with previous work characterizing the diversity of the immune repertoire [10, 11, 14], our data reinforces that the TCR β repertoire of an individual is predominantly private. Even when several HLA alleles are shared between two unrelated individuals, we see that over 98% of each repertoire is comprised of unique clones. Within the singleton repertoire, which best reflects the inherent diversity of the naïve TCR β repertoire, less than 1% of rearrangements are shared between individuals, demonstrating the high variability of VDJ somatic recombination. We show that HLA zygosity of an individual does not strongly affect the overall diversity of their TCR β repertoire. While previous work with a similar dataset found a correlation between HLA class I zygosity and decreased repertoire richness among CMV-individuals, that analysis quantified diversity with different metrics and did not control for sampling depth, which could impact the finding [30]. One limitation of this cohort is that over half of the individuals are of European descent, and a more diverse cohort may have a different distribution of HLA zygosity. Nonetheless, these findings suggest that the space of potential

TCR β clones is diverse enough that the HLA genotype homozygosity of an individual does not affect the overall diversity of their repertoire.

The divergent allele advantage hypothesis suggests that more heterozygous HLA loci and greater evolutionary divergence among HLA alleles lead to an increased number of antigens that can be presented to T cells by the MHC [21]. Even within fully heterozygous individuals, evolutionary divergence of HLA alleles is positively correlated with the number of peptides predicted to be recognized by the TCR β repertoire of an individual [16, 22]. As such, it is perhaps not surprising that HLA zygosity alone is not enough to alter the shape of the TCR β repertoire. Future work incorporating phylogeny-aware or distance-based clustering of HLA alleles [16, 35, 36] could better assess the functional diversity of the HLA alleles of an individual and thus the relationship between potential bound peptides and diversity. However, it is also important to note that decreased numbers of bound peptides may not translate to a decrease in the number of unique T cells generated, as HLA restriction during T-cell maturation functions independently of antigen exposure.

Despite the large inherent diversity of the TCR β repertoire, comparisons between unrelated individuals always contain a portion of public clones. We show that individuals sharing more HLA alleles share more TCR β clones, suggesting that HLA type impacts the specific T cells selected for maturation during thymic selection. Previous work has also shown that public clones occur more frequently than would be expected by chance due to biases in VDJ recombination and indels [1, 6, 10, 37]. Our analyses support these findings, as individuals sharing no HLA alleles still share TCR β clones. Moreover, we see a stronger correlation between HLA sharing and the sharing of rare clones compared to the sharing of common clones, suggesting a background of easily generated clones while rare clones are more HLA restricted. Notably, even among commonly recombined clones, HLA restriction is significantly correlated with clone sharing, suggesting that convergent recombination alone cannot explain the prevalence of all public clones. Further work could also assess the similarity of HLA alleles between individuals to better quantify their overall ability to present common antigen.

While high frequency clones may be shared between individuals after exposure to a common antigen, the lowest frequency clones within an individual should be largely antigen naïve. We see that the number of shared HLA alleles is positively correlated with the number of shared singletons between individuals, further supporting our assertion that sharing of TCR β clones of all frequencies is influenced by HLA allele overlap. Additionally, our data showing that older individuals share more singleton clones than younger individuals is consistent with thymic involution resulting in decreased T-cell diversity [32]. As such, our work highlights the importance of incorporating HLA genotype and age in models examining public clone sharing, as well as the important distinctions between the naïve and memory compartments. This work further supports the hypothesis that TCR β repertoires can be used to infer the HLA type of an individual [11], however age may be a confounding factor and including younger individuals in the training data may increase model specificity and sensitivity.

Finally, while there has been demonstrated success in the diagnostic potential of public antigen-specific clones, our work suggests a largely private expansion of TCR β clones in response to CMV antigen exposure. While public TCR β clones exist, they may represent the minority of the overall antigenic response. We show that CMV exposure actually decreases overall clone sharing, regardless of the number of HLA alleles shared. While CMV is a unique virus in that it often remains dormant without being fully cleared, by extending this analysis to include the age of each individual as a proxy for continued pathogenic exposure, our data indicate that the expansion of private clones in response to antigen exposure is not unique to CMV. Older individuals have likely been exposed to a broader range of pathogens than younger individuals, shaping their TCR β repertoires in an increasingly private manner as T cells specific to

encountered antigens expand and remain in the memory compartment. Interestingly, work in identical twins has similarly shown predominantly private responses to common antigens across individuals with the same genetic background [38–40]. However additional studies confirming this pattern in datasets with other known viruses, such as EBV or HBV or pre-post vaccination would provide further support for the expansion of private clones to specific pathogens. These data underscore how finding public, antigen-specific clones is a difficult problem requiring large datasets. However, the individual nature of TCR β repertoires provides utility in tracking clones within individuals over time or between individuals with allogeneic T-cell transplants. While posing a considerable computational challenge, important future work will include defining motifs or implementing clustering algorithms to identify TCR β clones that bind the same antigen [36, 41]. Together, this work emphasizes the inherent diversity and private nature of human TCR β repertoires, as well as the importance of incorporating HLA genotypes into models predicting both public TCR β sharing and antigen-specific expansion.

Methods

Sample details

All samples analyzed in this study were previously published in Emerson et al. [11]. Of the 666 immunosequenced healthy adult PBMC repertoires, we included 426 individuals in this analysis. These individuals all have complete 4-digit HLA resolution at the 6 included HLA loci (HLA-A, HLA-B, HLA-C, HLA-DPA, HLA-DQB, HLA-DRB), CMV serotyping, and 200,000 or more productive templates sequenced. HLA typing and CMV serostatus testing was conducted at the time of sample collection. Additionally, 15 samples were removed due to unexpected sharing of high-frequency TCR β clones. Of these 426 samples, we know the age at collection of 366 of the subjects who were thus included in the age-related analyses. Samples were computationally down-sampled to 200,000 productive nucleotide templates with replacement to minimize skewing in the number of singletons. TCR β functional identity was determined using V-family and J-gene, as those gave the most robust resolutions of TCR β clones.

Calculating simpson clonality

Simpson Clonality was defined as: $Simpson\ clonality = \sqrt{\sum p_i^2}$ and was calculated on productive nucleotide rearrangements, where p_i is the proportional abundance of rearrangement i and N is the total number of rearrangements. Clonality values range from 0 to 1 and describe the shape of the frequency distribution. Clonality values approaching 0 indicate a very even distribution of frequencies, whereas values approaching 1 indicate an increasingly asymmetric distribution in which a few clones are present at high frequencies.

Calculating a generation probability cutoff for common and rare clone distinction

To calculate the generation probability, the TCR β CDR3_{aa} sequence, and if available V family and/or J gene, of each clone were input into OLGA, an algorithm that calculates generation probabilities using a generative model for VDJ recombination [9]. The unique rearrangements (a unique combination of TCR β CDR3_{aa}, V family, and J gene) from each sample were aggregated into a single data table and randomly shuffled. To optimize downstream computation efficiency, 31 sub tables (“chunks”) containing ~2 million rearrangements were created. Within each sub table, rearrangements occurring greater than once were labeled as “common” and those occurring once labeled as “rare.” The generation probabilities of each sub table’s common and rare clones were plotted as density curves and their points of intersection

identified. The median value of the 31 intersection points was used downstream as the universal cutoff point to identify individual rearrangements as either common or rare.

Statistical analysis for association between HLA allele similarity and TCR β clone sharing

Mantel tests were conducted to evaluate the correlation between the pairwise number of shared HLA alleles and number of shared TCR β clones. This correlation statistic is calculated by first creating two dissimilarity matrices, one for each variable being compared. The correlation between these two matrices is subsequently measured and then one matrix is repetitively shuffled to determine how often randomization of one matrix leads to increased correlation between the two matrices and thus variables. This permutation test evaluates the significance of the correlation between the observed dissimilarity matrices.

Linear mixed effects models were additionally created to measure the dependency of TCR β clone sharing on HLA allele sharing by generating an intercept and slope. The mixed effects model format was selected to incorporate the impact of the numerous different sample comparisons between separate individuals resulting in random effects altering the regression.

All statistical analyses and visualizations were conducted using R version 3.6 (<https://www.r-project.org/>).

Creating null distribution of TCR β clone sharing across groups

To test for the significance of the association between CMV or age with TCR β clone sharing, we employed a permutation test in which the labels of interest were shuffled 1,000 times and the shuffled medians for each group were compared to the corresponding observed median. We then report the rank of each empirical median relative to the 1,000 shuffled medians and inclusive of the observed median, as well as the p-value for each group (S5A–S5C Fig). Rank values closer to 1000 signify that the observed median is greater than the permuted median in most or all of the permutations, while values closer to 0 indicate that the measured median was below most of the permuted medians. The p-value was determined as the number of permutations more extreme than the observed value plus 1, to be inclusive of the observed value, out of 1001 and multiplied by 2 to account for a two-sided test.

Supporting information

S1 Fig. Distribution of HLA alleles in study cohort. (A) Count of individuals, among the cohort of 426, based on number of homozygous HLA loci. (B) Count of individuals based on number of homozygous HLA class I loci: HLA-A, HLA-B, HLA-C. Most individuals are heterozygous at all HLA Class I loci. (C) Count of individuals based on number of homozygous HLA class II loci: HLA-DPA1, HLA-DQB1, HLA-DRB1. (D) The distribution of individuals with each combination of homozygous loci. 44% of individuals in this cohort are homozygous at HLA-DPA1.
(TIF)

S2 Fig. Impact of HLA zygosity on repertoire richness and clonality. HLA zygosity at neither (A) class I loci (Spearman rho = 0.011, p = 0.08) nor (B) class II loci were correlated with Simpson clonality (-0.023, p = 0.64). Overall, HLA zygosity in this cohort was not correlated with (C) richness (Spearman rho = 0.026, p = 0.60) or (D) Simpson clonality (Spearman rho = 0.0005, p = 0.99). (E) HLA zygosity was not correlated with richness among CMV- individuals (Spearman rho = 0.018, p = 0.79) or CMV+ individuals (Spearman rho = 0.086, p = 0.23). (F) HLA zygosity was not correlated with Simpson clonality among CMV- individuals (Spearman

$\rho = -0.064$, $p = 0.34$) or CMV+ individuals (Spearman $\rho = 0.023$, $p = 0.75$).
(TIF)

S3 Fig. Impact of age and CMV status on repertoire clonality and richness. CMV+ Individuals had significantly (A) higher Simpson clonality (Wilcoxon rank sum test, $p < 2.2e-16$) and (B) lower richness (Wilcoxon rank sum test, $p = 4.2e-14$) than CMV- individuals. (C) Simpson clonality was positively correlated with age overall (Spearman $\rho = 0.34$, $p = 3.4e-11$) and within both CMV+ (Spearman $\rho = 0.23$, $p = 0.003$) and CMV- individuals (Spearman $\rho = 0.31$, $p = 1.4e-3$). (D) Repertoire richness was inversely correlated with age overall (Spearman $\rho = -0.35$, $p = 8.6e-12$) and within both CMV+ (Spearman $\rho = -0.23$, $p = 0.002$) and CMV- individuals (Spearman $\rho = -0.39$, $p = 3.4e-8$). (E) CMV+ individuals were significantly older than CMV- individuals (Wilcoxon rank sum test, $p = 2.9e-5$). (F) Subjects older than the overall median age had significantly greater Simpson clonality values than individuals that are younger than or equal to the median age (Wilcoxon rank sum test, $p = 2.0e-7$).
(TIF)

S4 Fig. Distributions of TCR β generation probabilities. (A) Distribution of generative probabilities of TCR β clones from 10 representative individuals. Generation probability was significantly correlated with (B) CDR3 nucleotide sequence length (Spearman $\rho = -0.48$, $p < 2.2e-16$) and (C) number of insertions and deletions (indels) (Spearman $\rho = -0.63$, $p < 2.2e-16$), figure from single representative repertoire. (D) Identification of a generation probability cutoff point to distinguish between common and rare clones. Generation probability intersection points from the 31 chunks ranged from -8.81 to -8.76 (log₁₀ transformed values) and the median -8.78 (1.66e-9) was selected as the cutoff point. Within the 5000 most abundant rearrangement subset, the number of shared HLA alleles was significantly correlated with clone sharing among both (E) common (Mantel $\rho = 0.19$, $p < 1e-3$) and (F) rare clones (Mantel $\rho = 0.28$, $p < 1e-3$).
(TIF)

S5 Fig. Determination of p-values by permutation tests. Within each set of pairwise comparisons, we compared the observed median value to the medians of 1,000 permuted comparisons and reported both rank and p-value for distribution indicated. (A) The observed median number of clones shared between CMV+ individuals was lower than the median of all shuffled comparisons, yielding a significant empirical p-value of 0.002. (B) The observed median number of clones shared between CMV+ and CMV- individuals was lower than the median of all shuffled comparisons, yielding a significant empirical p-value of 0.002. (C) The observed median number of clones shared between CMV- individuals was greater than the median of all shuffled comparisons, yielding a significant empirical p-value of 0.002. (D) Table containing the rank and empirical p-values of all clone sharing comparisons stratified by CMV serostatus. (E) Table containing the rank and empirical p-values of all clone sharing comparisons stratified by age relative to median age (42).
(TIF)

S6 Fig. Impact of shared HLA alleles and CMV status on the sharing of common and rare TCR β clones. (A) Individuals that were both CMV- shared more common clones than individuals that were both CMV+, regardless of the number of HLA alleles shared. (B) Individuals that were both CMV- shared more rare clones than individuals that were both CMV+, regardless of the number of HLA alleles shared. Individuals that were both CMV+ shared more (C) common and (D) rare singletons than individuals that were both CMV-, regardless of the number of HLA alleles shared.
(TIF)

S7 Fig. Association between age and median TCR β repertoire generation probability. Age was significantly correlated with (A) higher median generation probability within the down-sampled repertoires (Spearman rho = 0.12, p = 0.022), (B) lower median generation probability within the top 5,000 clones (Spearman rho = -0.37, p = 8.6e-15), and (C) higher median generation probability within the singleton repertoires (Spearman rho = 0.17, p = 0.0015). Only subjects with available age data were included in this analysis.
(TIF)

S8 Fig. Impact of age on sharing of top 5000 TCR β clones. Individuals were stratified by the median age (42). (A) Younger individuals shared more clones than older individuals. (B) Age impacted the sharing of the top 5000 clones regardless of the number of HLA alleles shared. Younger individuals shared more (C) common clones, but no difference between age groups in (D) rare clones sharing was observed.
(TIF)

S9 Fig. Proportion of rare TCR β clones in repertoires. (A) There was not a significant difference in proportion of rare clones between older and younger individuals within the down-sampled repertoires (Wilcoxon rank sum test, p = 0.5). (B) Older individuals had a significantly greater proportion of rare clones among their top 5,000 most abundant rearrangements than younger individuals (Wilcoxon rank sum test, p = 3.6e-10). (C) Older individuals had a significantly lower proportion of rare singletons compared to younger individuals (Wilcoxon rank sum test, p = 3.0e-2).
(TIF)

Acknowledgments

We would like to thank all members of the Adaptive Biotechnologies Computational Biology Group who provided valuable insights during the project development and data analysis, in particular Erik Yusko, Bryan Howie, and Marissa Vignali.

Author Contributions

Conceptualization: Sarah A. Johnson, Rachel M. Gittelman, Julie A. Rytlewski, Paul A. Fields.

Formal analysis: Sarah A. Johnson, Spencer L. Seale, Rachel M. Gittelman, Julie A. Rytlewski.

Investigation: Sarah A. Johnson, Spencer L. Seale, Rachel M. Gittelman, Julie A. Rytlewski.

Methodology: Sarah A. Johnson, Spencer L. Seale, Rachel M. Gittelman, Julie A. Rytlewski.

Supervision: Harlan S. Robins, Paul A. Fields.

Visualization: Sarah A. Johnson, Spencer L. Seale, Rachel M. Gittelman, Julie A. Rytlewski.

Writing – original draft: Sarah A. Johnson, Spencer L. Seale, Paul A. Fields.

Writing – review & editing: Sarah A. Johnson, Spencer L. Seale, Rachel M. Gittelman, Harlan S. Robins, Paul A. Fields.

References

1. Robins HS, Campreggher PV, Srivastava SK, Wachter A, Turtle CJ, Kahsai O, et al. Comprehensive assessment of T-cell receptor beta-chain diversity in alphabeta T cells. *Blood*. 2009 Nov 5; 114(19):4099–107. <https://doi.org/10.1182/blood-2009-04-217604> PMID: 19706884

2. Huppa JB, Davis MM. T-cell-antigen recognition and the immunological synapse. *Nat Rev Immunol*. 2003; 3(12):973–83. <https://doi.org/10.1038/nri1245> PMID: 14647479
3. Kirsch I, Vignali M, Robins H. T-cell receptor profiling in cancer. *Mol Oncol*. 2015 Dec; 9(10):2063–70. <https://doi.org/10.1016/j.molonc.2015.09.003> PMID: 26404496
4. Robins HS, Srivastava SK, Campregher PV, Turtle CJ, Andriesen J, Riddell SR, et al. Overlap and effective size of the human CD8+ T-cell receptor repertoire. *Sci Transl Med*. 2010 Sep 1; 2(47):47ra64. <https://doi.org/10.1126/scitranslmed.3001442> PMID: 20811043
5. Yates A. Theories and Quantification of Thymic Selection. *Front Immunol* [Internet]. 2014 [cited 2020 Oct 23]; 5. <https://www.frontiersin.org/articles/10.3389/fimmu.2014.00013/full> PMID: 24550908
6. Pogorely MV, Minervina AA, Chudakov DM, Mamedov IZ, Lebedev YB, Mora T, et al. Method for identification of condition-associated public antigen receptor sequences. Chakraborty AK, editor. *eLife*. 2018 Mar 13; 7:e33050. <https://doi.org/10.7554/eLife.33050> PMID: 29533178
7. Madi A, Shifrut E, Reich-Zeliger S, Gal H, Best K, Ndifon W, 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 Oct; 24(10):1603–12. <https://doi.org/10.1101/gr.170753.113> PMID: 25024161
8. DeWitt WS III, Smith A, Schoch G, Hansen JA, Matsen FA IV, Bradley P. Human T cell receptor occurrence patterns encode immune history, genetic background, and receptor specificity. Walczak AM, Chakraborty AK, Elhanati Y, Gerritsen B, editors. *eLife*. 2018 Aug 28; 7:e38358. <https://doi.org/10.7554/eLife.38358> PMID: 30152754
9. Sethna Z, Elhanati Y, Callan CG, Walczak AM, Mora T. OLGA: fast computation of generation probabilities of B- and T-cell receptor amino acid sequences and motifs. *Bioinformatics*. 2019 Sep 1; 35(17):2974–81. <https://doi.org/10.1093/bioinformatics/btz035> PMID: 30657870
10. Elhanati Y, Sethna Z, Callan CG, Mora T, Walczak AM. Predicting the spectrum of TCR repertoire sharing with a data-driven model of recombination. *Immunol Rev*. 2018 Jul; 284(1):167–79. <https://doi.org/10.1111/imr.12665> PMID: 29944757
11. Emerson RO, DeWitt WS, Vignali M, Gravley J, Hu JK, Osborne EJ, et al. Immunosequencing identifies signatures of cytomegalovirus exposure history and HLA-mediated effects on the T cell repertoire. *Nat Genet*. 2017 May; 49(5):659–65. <https://doi.org/10.1038/ng.3822> PMID: 28369038
12. DeWitt WS, Emerson RO, Lindau P, Vignali M, Snyder TM, Desmarais C, et al. Dynamics of the cytotoxic T cell response to a model of acute viral infection. *J Virol*. 2015 Apr; 89(8):4517–26. <https://doi.org/10.1128/JVI.03474-14> PMID: 25653453
13. Snyder TM, Gittelman RM, Klinger M, May DH, Osborne EJ, Taniguchi R, et al. Magnitude and Dynamics of the T-Cell Response to SARS-CoV-2 Infection at Both Individual and Population Levels. *medRxiv*. 2020 Sep 17;2020.07.31.20165647. <https://doi.org/10.1101/2020.07.31.20165647> PMID: 32793919
14. Venturi V, Chin HY, Asher TE, Ladell K, Scheinberg P, Bornstein E, et al. TCR β -Chain Sharing in Human CD8+ T Cell Responses to Cytomegalovirus and EBV. *J Immunol*. 2008 Dec 1; 181(11):7853–62. <https://doi.org/10.4049/jimmunol.181.11.7853> PMID: 19017975
15. Britanova OV, Shugay M, Merzlyak EM, Staroverov DB, Putintseva EV, Turchaninova MA, et al. Dynamics of Individual T Cell Repertoires: From Cord Blood to Centenarians. *J Immunol*. 2016 Jun 15; 196(12):5005–13. <https://doi.org/10.4049/jimmunol.1600005> PMID: 27183615
16. Chowell D, Krishna C, Pierini F, Makarov V, Rizvi NA, Kuo F, et al. Evolutionary divergence of HLA class I genotype impacts efficacy of cancer immunotherapy. *Nat Med*. 2019 Nov; 25(11):1715–20. <https://doi.org/10.1038/s41591-019-0639-4> PMID: 31700181
17. Wall DA, Krueger J. Chimeric antigen receptor T cell therapy comes to clinical practice. *Curr Oncol*. 2020 Apr; 27(Suppl 2):S115–23. <https://doi.org/10.3747/co.27.5283> PMID: 32368181
18. Porter DL, Levine BL, Kalos M, Bagg A, June CH. Chimeric Antigen Receptor–Modified T Cells in Chronic Lymphoid Leukemia. *N Engl J Med*. 2011 Aug 25; 365(8):725–33. <https://doi.org/10.1056/NEJMoa1103849> PMID: 21830940
19. Nakamura T, Shirouzu T, Nakata K, Yoshimura N, Ushigome H. The Role of Major Histocompatibility Complex in Organ Transplantation- Donor Specific Anti-Major Histocompatibility Complex Antibodies Analysis Goes to the Next Stage-. *Int J Mol Sci* [Internet]. 2019 Sep 13 [cited 2020 Nov 6]; 20(18). <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6769817/> PMID: 31540289
20. Bertaina A, Andreani M. Major Histocompatibility Complex and Hematopoietic Stem Cell Transplantation: Beyond the Classical HLA Polymorphism. *Int J Mol Sci* [Internet]. 2018 Feb 22 [cited 2020 Nov 6]; 19(2). Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5855843/> PMID: 29470425
21. Wakeland EK, Boehme S, She JX, Lu C-C, McIndoe RA, Cheng I, et al. Ancestral polymorphisms of MHC class II genes: Divergent allele advantage. *Immunol Res*. 1990 Jun 1; 9(2):115–22. <https://doi.org/10.1007/BF02918202> PMID: 2189934

22. Divergent Allele Advantage at Human MHC Genes: Signatures of Past and Ongoing Selection | Molecular Biology and Evolution | Oxford Academic [Internet]. [cited 2020 Oct 26]. <https://academic.oup.com/mbe/article/35/9/2145/5034935>
23. Britanova OV, Putintseva EV, Shugay M, Merzlyak EM, Turchaninova MA, Staroverov DB, et al. Age-Related Decrease in TCR Repertoire Diversity Measured with Deep and Normalized Sequence Profiling. *J Immunol*. 2014 Mar 15; 192(6):2689–98. <https://doi.org/10.4049/jimmunol.1302064> PMID: 24510963
24. Qi Q, Liu Y, Cheng Y, Glanville J, Zhang D, Lee J-Y, et al. Diversity and clonal selection in the human T-cell repertoire. *Proc Natl Acad Sci U S A*. 2014 Sep 9; 111(36):13139–44. <https://doi.org/10.1073/pnas.1409155111> PMID: 25157137
25. Lindau P, Mukherjee R, Gutschow MV, Vignali M, Warren EH, Riddell SR, et al. Cytomegalovirus Exposure in the Elderly Does Not Reduce CD8 T Cell Repertoire Diversity. *J Immunol*. 2019 Jan 15; 202(2):476–83. <https://doi.org/10.4049/jimmunol.1800217> PMID: 30541882
26. Takada K, Takahama Y. Chapter Three—Positive-Selection-Inducing Self-Peptides Displayed by Cortical Thymic Epithelial Cells. In: Alt FW, editor. *Advances in Immunology* [Internet]. Academic Press; 2015 [cited 2020 Nov 2]. p. 87–110. <http://www.sciencedirect.com/science/article/pii/S0065277614000042> PMID: 25591465
27. Venturi V, Kedzierska K, Price DA, Doherty PC, Douek DC, Turner SJ, et al. Sharing of T cell receptors in antigen-specific responses is driven by convergent recombination. *Proc Natl Acad Sci*. 2006 Dec 5; 103(49):18691–6. <https://doi.org/10.1073/pnas.0608907103> PMID: 17130450
28. Jiang W, Kang L, Lu H-Z, Pan X, Lin Q, Pan Q, et al. Normal Values for CD4 and CD8 Lymphocyte Subsets in Healthy Chinese Adults from Shanghai. *Clin Diagn Lab Immunol*. 2004 Jul; 11(4):811–3. <https://doi.org/10.1128/CDLI.11.4.811-813.2004> PMID: 15242966
29. The naive T-cell receptor repertoire has an extremely broad distribution of clone sizes | eLife [Internet]. [cited 2020 Oct 26]. <https://elifesciences.org/articles/49900>
30. Krishna C, Chowell D, Gönen M, Elhanati Y, Chan TA. Genetic and environmental determinants of human TCR repertoire diversity. *Immun Ageing*. 2020 Sep 4; 17(1):26. <https://doi.org/10.1186/s12979-020-00195-9> PMID: 32944053
31. Yoshida K, Cologne JB, Cordova K, Misumi M, Yamaoka M, Kyoizumi S, et al. Aging-related changes in human T-cell repertoire over 20years delineated by deep sequencing of peripheral T-cell receptors. *Exp Gerontol*. 2017 Oct 1; 96:29–37. <https://doi.org/10.1016/j.exger.2017.05.015> PMID: 28535950
32. Fulton RB, Varga SM. Effects of aging on the adaptive immune response to respiratory virus infections. *Aging Health*. 2009 Dec 1; 5(6):775. <https://doi.org/10.2217/ahe.09.69> PMID: 20174457
33. Palmer DB. The effect of age on thymic function. *Front Immunol*. 2013 Oct 7; 4:316. <https://doi.org/10.3389/fimmu.2013.00316> PMID: 24109481
34. Brodin P, Davis MM. Human immune system variation. *Nat Rev Immunol*. 2017 Jan; 17(1):21–9. <https://doi.org/10.1038/nri.2016.125> PMID: 27916977
35. Grantham R. Amino acid difference formula to help explain protein evolution. *Science*. 1974 Sep 6; 185(4154):862–4. <https://doi.org/10.1126/science.185.4154.862> PMID: 4843792
36. Huang H, Wang C, Rubelt F, Scriba TJ, Davis MM. Analyzing the Mycobacterium tuberculosis immune response by T-cell receptor clustering with GLIPH2 and genome-wide antigen screening. *Nat Biotechnol*. 2020 Oct; 38(10):1194–202. <https://doi.org/10.1038/s41587-020-0505-4> PMID: 32341563
37. Venturi V, Quigley MF, Greenaway HY, Ng PC, Ende ZS, McIntosh T, et al. A mechanism for TCR sharing between T cell subsets and individuals revealed by pyrosequencing. *J Immunol Baltim Md 1950*. 2011 Apr 1; 186(7):4285–94.
38. Pogorely MV, Minervina AA, Touzel MP, Sycheva AL, Komech EA, Kovalenko EI, et al. Precise tracking of vaccine-responding T cell clones reveals convergent and personalized response in identical twins. *Proc Natl Acad Sci U S A*. 2018 Dec 11; 115(50):12704–9. <https://doi.org/10.1073/pnas.1809642115> PMID: 30459272
39. Brodin P, Jovic V, Gao T, Bhattacharya S, Angel CJL, Furman D, et al. Variation in the human immune system is largely driven by non-heritable influences. *Cell*. 2015 Jan 15; 160(1–2):37–47. <https://doi.org/10.1016/j.cell.2014.12.020> PMID: 25594173
40. Yu XG, Lichtenfeld M, Williams KL, Martinez-Picado J, Walker BD. Random T-Cell Receptor Recruitment in Human Immunodeficiency Virus Type 1 (HIV-1)-Specific CD8+ T Cells from Genetically Identical Twins Infected with the Same HIV-1 Strain. *J Virol*. 2007 Nov 15; 81(22):12666–9. <https://doi.org/10.1128/JVI.01450-07> PMID: 17804503
41. Dash P, Fiore-Gartland AJ, Hertz T, Wang GC, Sharma S, Souquette A, et al. Quantifiable predictive features define epitope specific T cell receptor repertoires. *Nature*. 2017 Jul 6; 547(7661):89–93. <https://doi.org/10.1038/nature22383> PMID: 28636592