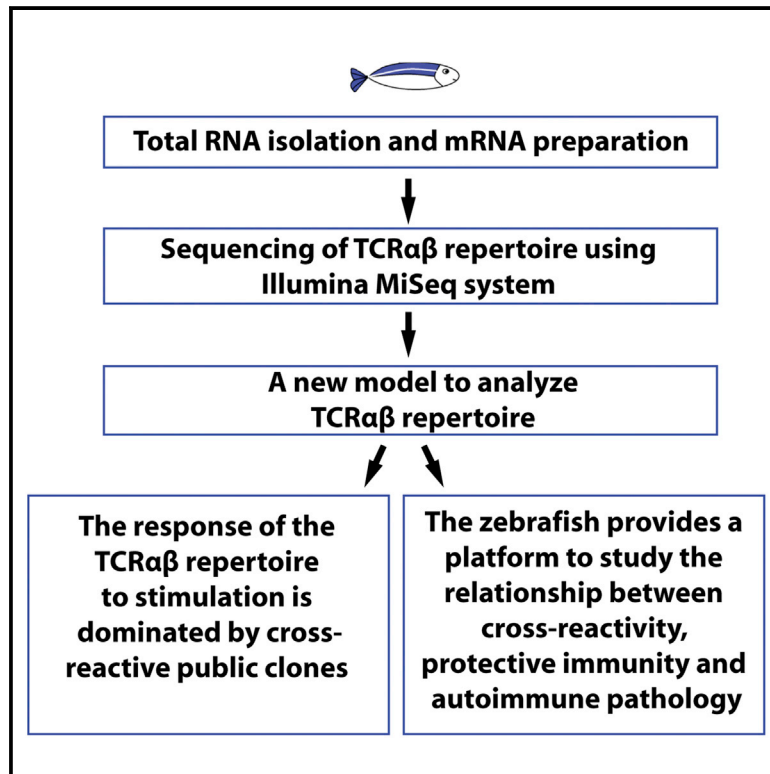


System-wide Analysis of the T Cell Response

Graphical Abstract



Authors

Ruxandra Covacu, Hagit Philip, Merja Jaronen, ..., Daniel C. Douek, Sol Efroni, Francisco J. Quintana

Correspondence

ddouek@mail.nih.gov (D.C.D.),
sol.efroni@biu.ac.il (S.E.),
fquintana@rics.bwh.harvard.edu (F.J.Q.)

In Brief

The T cell receptor (TCR) controls the adaptive immune response, but our understanding of the TCR repertoire is incomplete. Covacu et al. study the zebrafish TCR repertoire and its response to self and foreign antigens and find that public cross-reactive clones dominate the TCR repertoire, facilitating the rapid generation of T cell immunity.

Highlights

- We study the response of the zebrafish TCR $\alpha\beta$ repertoire to antigenic stimulation
- The zebrafish TCR $\alpha\beta$ repertoire is dominated by cross-reactive public clones
- Public T cells facilitate the rapid generation of protective T cell immunity
- The zebrafish provides a model to study the T cell response at a systems levels

Accession Numbers

SRP070056



System-wide Analysis of the T Cell Response

Ruxandra Covacu,^{1,9} Hagit Philip,^{2,9} Merja Jaronen,^{1,9} Jorge Almeida,³ Jessica E. Kenison,¹ Samuel Darko,³ Chun-Cheih Chao,¹ Gur Yaari,⁴ Yoram Louzoun,^{5,6} Liran Carmel,⁷ Daniel C. Douek,^{3,*} Sol Efroni,^{2,*} and Francisco J. Quintana^{1,8,*}

¹Ann Romney Center for Neurologic Diseases, Brigham and Women's Hospital, Harvard Medical School, 77 Avenue Louis Pasteur, Boston, MA 02115, USA

²The Mina & Everard Goodman Faculty of Life Sciences, Bar-Ilan University, Ramat Gan 52900, Israel

³Human Immunology Section, Vaccine Research Center, National Institute of Allergy and Infectious Diseases, US National Institutes of Health, Bethesda, MD 20892, USA

⁴Bioengineering Program, Faculty of Engineering

⁵Department of Mathematics

⁶Gonda Brain Research Center

Bar-Ilan University, Ramat Gan 52900, Israel

⁷Department of Genetics, The Alexander Silberman Institute of Life Sciences, Faculty of Science, The Hebrew University of Jerusalem, Edmond J. Safra Campus, Givat Ram, Jerusalem 91904, Israel

⁸Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA

⁹Co-first author

*Correspondence: ddouek@mail.nih.gov (D.C.D.), sol.efroni@biu.ac.il (S.E.), fquintana@rics.bwh.harvard.edu (F.J.Q.)

<http://dx.doi.org/10.1016/j.celrep.2016.02.056>

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

SUMMARY

The T cell receptor (TCR) controls the cellular adaptive immune response to antigens, but our understanding of TCR repertoire diversity and response to challenge is still incomplete. For example, TCR clones shared by different individuals with minimal alteration to germline gene sequences (public clones) are detectable in all vertebrates, but their significance is unknown. Although small in size, the zebrafish TCR repertoire is controlled by processes similar to those operating in mammals. Thus, we studied the zebrafish TCR repertoire and its response to stimulation with self and foreign antigens. We found that cross-reactive public TCRs dominate the T cell response, endowing a limited TCR repertoire with the ability to cope with diverse antigenic challenges. These features of vertebrate public TCRs might provide a mechanism for the rapid generation of protective T cell immunity, allowing a short temporal window for the development of more specific private T cell responses.

INTRODUCTION

The T cell repertoire, constituted by the pool of T cell receptor (TCR) specificities, governs the ability of the immune system to respond to both foreign and self-derived immune challenges (Linnemann et al., 2013; Newell and Davis, 2014; Nikolich-Zugich et al., 2004; Turner et al., 2009). Ninety five percent of the TCRs are composed of an α and a β protein chain. The antigen specificity of the TCR is primarily determined by the CDR3 of α and β chains (Rudolph et al., 2006), which interacts with the peptide

MHC complex (Davis and Bjorkman, 1988). Indeed, the majority of TCR variation is localized in the third complementarity-determining region (CDR3) as a result of the recombination of variable (V), diversity (D), and joining (J) segments and the incorporation of multiple nucleotide insertions and deletions. Thus, the study of CDR3 sequences provides information about the fraction of the TCR repertoire relevant for antigen recognition. However, the complexity and dynamics of the TCR repertoire remain unknown because of the limited power of the tools used for its investigation.

Previous studies estimated CDR3 D based on the analysis of a relatively small number of T cells. These studies are based on a solution for the “unseen species problem” developed to estimate the total number of species in a given population based on random samples of species (Efron and Thisted, 1976; Fisher et al., 1943). This method assumes that the number of TCR clones follows a Poisson distribution, however, recent studies found a power law distribution instead (Weinstein et al., 2009). Indeed, studies based on the sequencing of small T cell samples produced estimates of TCR D that were directly proportional to the number of sequences analyzed, suggesting that these methods do not capture the complete TCR repertoire D (Freeman et al., 2009). Even when advanced methods are used to study the TCR repertoire, these methods are still limited by their lack of consideration of tissue resident T cells (Burzyn et al., 2013; Park and Kupper, 2015). Because of these limitations, it is still unclear what fraction of the potential T cell repertoire is expressed, and how similar are the repertoires of different individuals in the quiescent state and during the course of an immune response. In addition, TCR sequences shared by different individuals (termed public TCR sequences) are detected in all vertebrates in multiple biological contexts, a surprising finding when the number of potential unique CDR3 sequences generated by VDJ recombination is considered (McBerry et al., 2012; Venturi et al., 2008). However, the significance of public

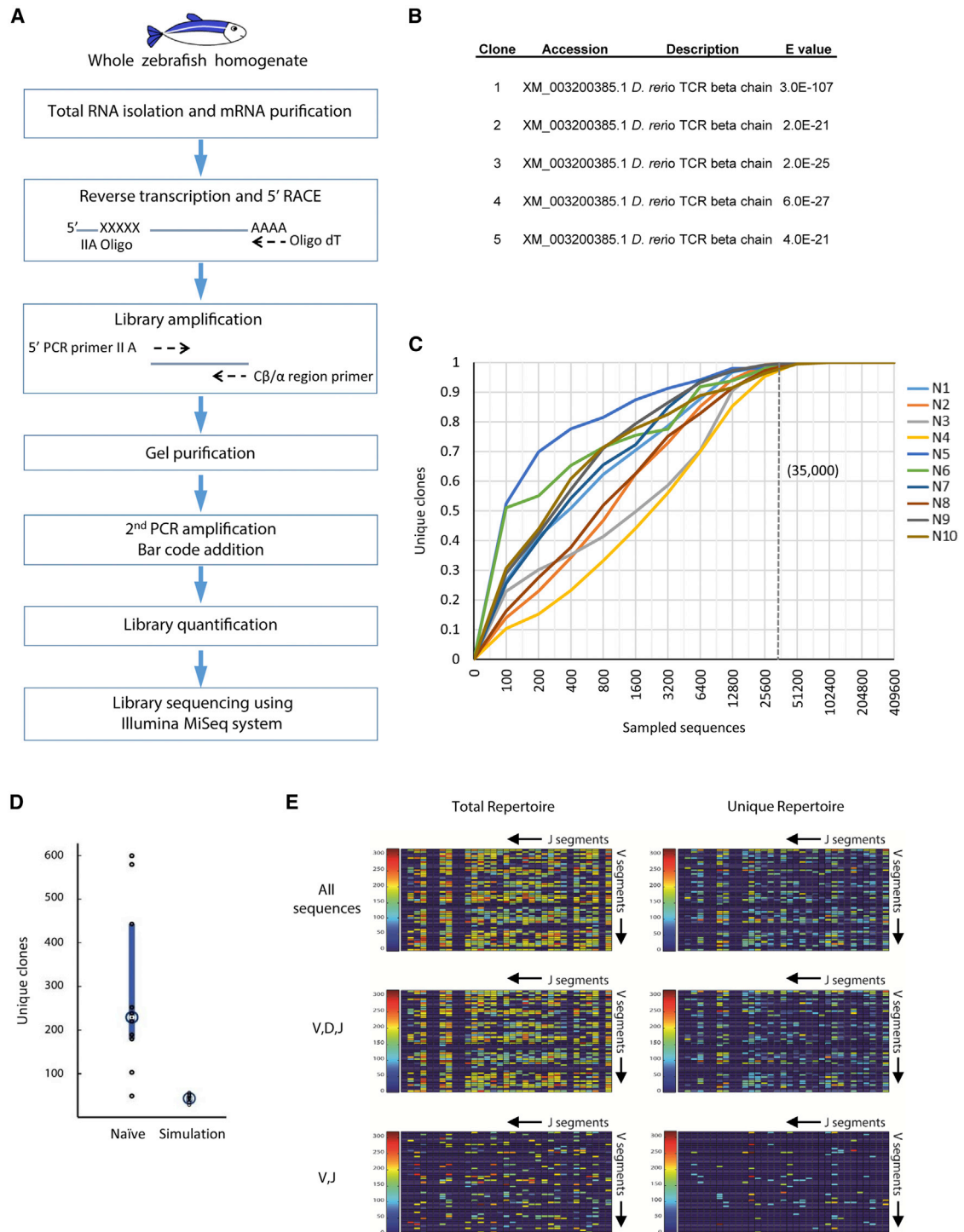


Figure 1. Sequencing of the Zebrafish TCR Repertoire

(A) Flow diagram of the experimental protocols used to sequence the zebrafish TCRβ1 (and TCRα) repertoire. mRNA was purified from whole zebrafish, each fish separately. The RT was performed using an oligo dT primer, and the 5' RACE was obtained using oligonucleotide IIA (IIA Oligo). The library amplification was done using a primer specific for the Cβ1 or Cα region (Cβ1/α region primer) and a primer complementary to IIA Oligo (5' PCR primer IIA). The library was gel purified and barcodes were added. The library was quantified and sequenced.

(B) Verification of TCRβ1 sequences obtained after the library amplification step. There were five clones that were sequenced and the top hits of the blasted sequences are presented.

(legend continued on next page)

TCRs on the repertoire, as well as their response to stimulation is unknown.

Zebrafish (*Danio rerio*) is an ideal immunological model system to study the TCR repertoire because its adaptive immune system shares important features with its mammalian counterpart. Examples of these shared elements are the presence of a recombination activating gene (RAG), a combinatorial rearrangement of V, D, and J gene segments, junctional D during recombination and somatic hypermutation (Lieschke and Trede, 2009; Trede et al., 2004). In addition, the number of T cells in the zebrafish has been approximated to about 2×10^5 cells, a 10^6 -fold lower number compared to the T cell numbers found in mice. Therefore, in contrast to TCR sequencing studies performed in mammals using isolated T cell populations, the zebrafish offers the possibility to perform far more complete immune repertoire studies. In this work, we combine the experimental advantages offered by the zebrafish with high-coverage sequencing and computational approaches to investigate the full diversity of the TCR repertoire under homeostatic conditions and its response to challenge with self and non-self antigens.

RESULTS

The TCR β -chain repertoire provides an accurate representation of TCR diversity (Miles et al., 2011). Moreover, although two C-region TCR β -chain genes have been identified in the zebrafish, C β 1 and C β 2, transcripts of the C β 2 segment are very rare (Meeker et al., 2010). Thus, we focused our efforts on the analysis of the zebrafish TCR β 1 repertoire. To analyze the TCR repertoire in zebrafish, we developed a method for TCR library generation from whole zebrafish mRNA based on 5' rapid amplification of cDNA ends (RACE) amplification from a single primer annealing to the constant TCR region (Douek et al., 2002) (Figure 1A). This method uses a single constant region (C-region) and a 5'-anchor primer rather than multiple J or V region primers to avoid differential PCR amplification efficiencies and subsequent library bias (Boyd et al., 2009; Robins et al., 2009; Wang et al., 2010). To confirm the specificity of the method, we cloned and sequenced the amplification products. We found that 100% of the amplification products correspond to TCR β 1 sequences, demonstrating that this amplification method utilizing a single C-region primer is specific (Figure 1B).

To analyze the number of sequences required to provide a comprehensive coverage of the TCR β 1 repertoire with our sequencing strategy, we performed rarefaction studies using partial samples of the full TCR β 1 sequencing data. We found that the total number of V(D)J combinations detected was asymptotic toward saturation, with all of the expressed V(D)J species predicted to be detected by sampling 35,000 sequences or more (Figure 1C).

Small TCR β 1 Repertoire in Adult Zebrafish

The number of potential TCR combinatorial possibilities exceeds the total number of peripheral T cells in an individual (Davis and Bjorkman, 1988), suggesting that only a fraction of the potential TCR repertoire is actually expressed. Indeed, the repertoire of unique $\alpha\beta$ TCRs has been estimated at $\sim 10^7$ clones in the human (Arstila et al., 1999) and $\sim 10^6$ in the mouse (Casrouge et al., 2000), a small fraction of the 10^{15} – 10^{20} unique $\alpha\beta$ TCRs repertoire that could be potentially generated by these mammalian immune systems. To determine the fraction of the potential TCR β 1 repertoire expressed by zebrafish, we first constructed a computational model of TCR recombination in the zebrafish based on available sequences and our own data on V(D)J recombination, deletions, insertions, and substitutions in TCR β 1 sequences (Figure S1). This simulation, which considers the biophysical properties of recombination, asymptotically estimates an upper limit of 400,000 unique sequences. Because of the limitations imposed by the biophysical features of the recombination process, this estimate is smaller than the 10^8 – 10^{20} sequences that could result from all the potential V(D)J combinations (Benichou et al., 2012).

We then used our computational model to simulate the TCR β 1 repertoire for ten individual fish, considering not only the number of unique sequences detected, but also, the frequency of these sequences in the total TCR β 1 repertoire. Surprisingly, the model predicted a TCR β 1 repertoire consisting of only 40 unique TCR β 1 clones per fish. This low predicted number is comparable to the zebrafish TCR repertoire detected in our sequencing efforts, in which we detected 49–599 unique TCR β 1 CDR3 sequences per individual fish (Figure 1D). This small number of unique T cell clones is in agreement with previous studies of the zebrafish B cell repertoire, which has been estimated to harbor 9–200 unique V(D)J sequences generated by recombination, expanded to 1,200–6,000 antibody clones by somatic hypermutation (Weinstein et al., 2009). Taken together, these data suggest that only a small fraction of the potential TCR β 1 repertoire is actually expressed in adult zebrafish, a fraction significantly lower than the one estimated for the murine and human immune system.

Biased TCR β 1 Repertoire in Naive Zebrafish

The V(D)J recombination system generates a diverse TCR β 1 repertoire based on the stochastic use of V, D, and J gene segments and the deletion and insertion of nucleotides (Davis and Bjorkman, 1988; Fujimoto and Yamagishi, 1987; Malissen et al., 1986; Okazaki et al., 1987; Tonegawa, 1983). Although each V, D, and J segment has a theoretically equal chance of being incorporated into a mature TCR, the murine and human TCR repertoires are not evenly distributed and specific V, D, and

(C) Rarefaction analysis of TCR β 1 D. Each curve gives the fraction of the observed repertoire as a function of the number of obtained sequences in each of ten naive (non-immunized) fish. The dotted line indicates the point at which all fish reach 99% coverage of the total unique sequences, corresponding to 35,000 sequences.

(D) Number of unique clones in ten naive fish and ten simulated repertoires. The circle indicates the median. The small number of sequences in the simulated repertoire results from adherence to the distribution of clones within each fish. The abundance of specific clones (learned from the natural occurring repertoire in the naive fish) dictates a distribution with a few highly represented clones.

(E) Heatmaps of V and J combination abundance in the total and the unique repertoire of naive zebrafish. The data are segregated based on the usage of the D segment.

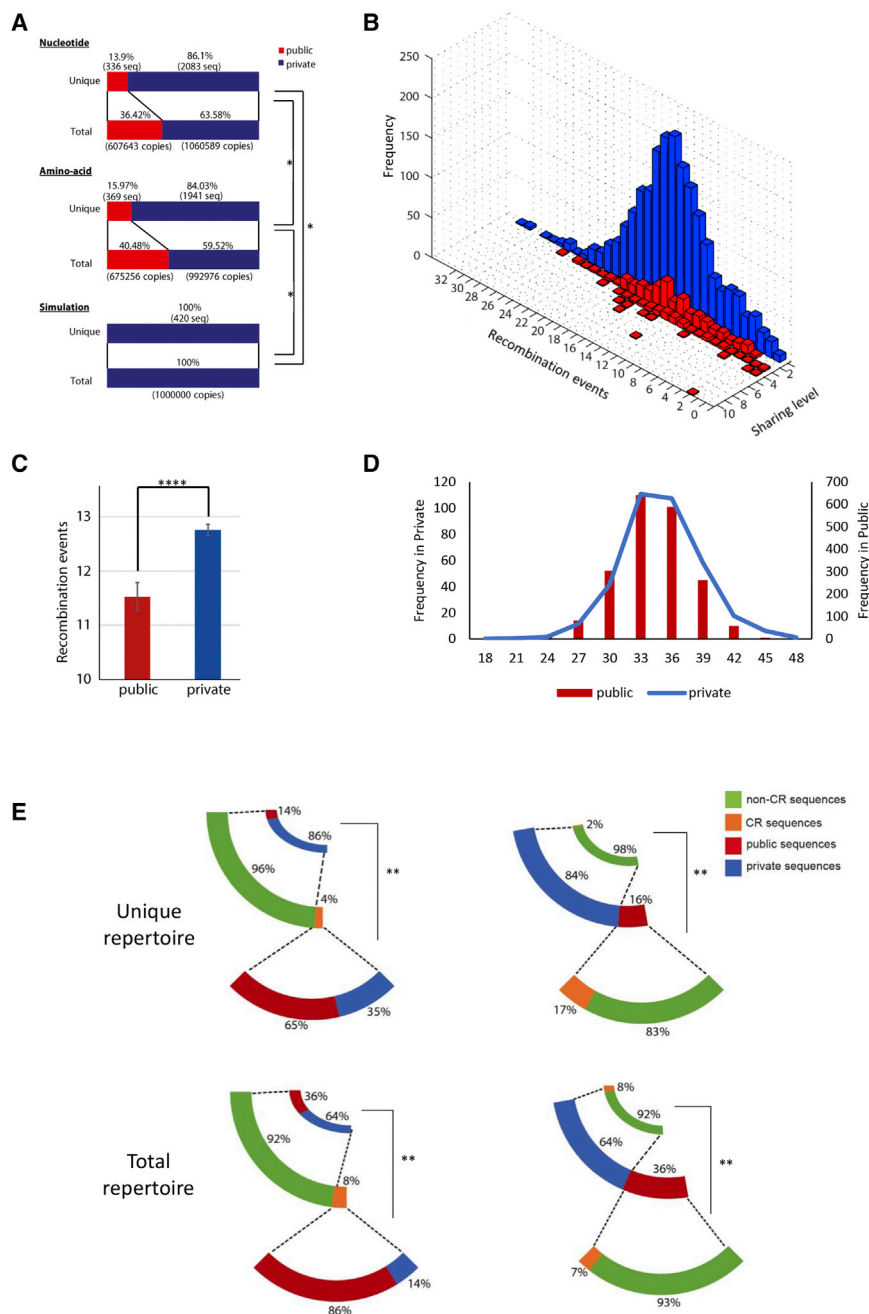


Figure 2. Convergent Recombination Characterizes Public TCR Clones that Dominate the Zebrafish TCRβ1 Repertoire

(A) Observed and predicted public and private TCRβ1 sequences in naive fish.

(B) Recombination events in the private and in the public TCRβ1 repertoires and their association with their sharing. The height of the histogram bars represents the frequency of the clone in the pooled repertoire (red, public and blue, private).

(C) Statistical analysis of recombination events in public and private TCRβ1 clones (****p < 1 × 10⁻¹²).

(D) Distribution of CDR3 lengths in the private and public TCRβ1 clones.

(E) Fraction of public and private TCRβ1 clones within the unique TCRβ1 repertoire. The fraction of sequences generated by convergent recombination is shown (**p < 0.01).

TCRβ1 repertoire when sequences containing the D segment were analyzed (Figure 1E). However, only a subset of pairs was utilized when the D segment was not included, suggesting that only a limited set of available VJ combinations overcomes the limitations imposed by the 12/23 rule (Akira et al., 1987; Yancopoulos et al., 1986). These data suggest that biases in gene segment usage characterize the zebrafish TCRβ1 repertoire. In addition, the analysis of the total repertoire, that is the collection of TCRβ1 unique sequences adjusted for their frequency, revealed the over-representation of specific V and J pairs (Figure 1E), suggesting that their expansion results from antigenic stimulation.

Convergent Recombination Drives the Generation of Zebrafish Public TCR Clones

Public T cell clones express TCR sequence motifs shared by different individuals and are often expanded by immunization, infection, or autoimmunity (McBerry et al., 2012; Venturi et al., 2008). To study the role of public reper-

J genes are used more often (Argaet et al., 1994; Cibotti et al., 1994; Moss et al., 1991). Thus, we studied the TCRβ1 repertoire in ten naive zebrafish for V, D, and J usage bias. We analyzed between 97,503 and 232,193 sequences per fish, a number significantly higher than the 35,000 sequences required to cover the whole TCRβ1 repertoire. Each sequence was aligned with reference sequences (Howe et al., 2013; Kettleborough et al., 2013) to identify specific V, D, and J genes, and then the frequency of each V and J combination, including or not the D segment, was estimated (Turner et al., 2006). We found that almost all possible V and J pairs were used in the zebrafish

toires in zebrafish, we defined a public sequence as one appearing in at least two different individuals, as previously defined in other studies (Li et al., 2012). We found that public clones represent 36% of the total TCRβ1 CDR3 nucleotide sequences and 40% of the amino acid sequences (Figure 2A). Conversely, our computational model predicted no sharing of TCR sequences between individual fish (p < 10⁻⁹). Thus, the zebrafish TCRβ1 repertoire contains a relatively low number of unique sequences, many of which are shared between different individuals.

To study the genetic mechanisms involved in the generation of public T cell clones, we analyzed the frequency of recombination

events in public and private TCR β 1 sequences. We found significantly fewer recombination events in public TCR β 1 sequences (Figures 2B and 2C), in agreement with previous reports of public TCRs being closer to germline configurations (Ishizuka et al., 2008; Miles et al., 2010; Vermijlen et al., 2010). Of note, public and private clones do not differ in their CDR3 length (Figure 2D), suggesting that the reduction in recombination events in public clones is not a byproduct of shorter CDR3 sequences.

Convergent recombination, the process by which multiple recombination events produce the same nucleotide sequence and multiple nucleotide sequences encode the same amino acid sequence, is considered an important driving force in the generation of public T cell responses (Quigley et al., 2010; Venturi et al., 2006). To study the contribution of convergent recombination in zebrafish public T cell responses, we searched for identical TCR β 1 amino acid sequences originating from different nucleotide sequences in naive zebrafish. We found a significant contribution of convergent recombination to the public TCR repertoire of naive fish. Four percent of the amino acid sequences in the unique TCR β 1 repertoire are produced by convergent recombination (Figure 2E, left). Strikingly, 65% of TCR β 1 sequences generated by convergent recombination are public (Figure 2E, left), suggesting that convergent recombination plays a significant role during the generation of public TCR β 1 sequences. Indeed, 17% of the public TCR β 1 sequences in the unique repertoire were generated by convergent recombination, as opposed to 2% in private sequences (Figure 2E, right). Taken together, these data show that convergent recombination drives the generation of the public repertoire in naive zebrafish.

Antigenic Stimulation Expands Public TCR β 1 Clones

The frequency of a specific TCR in the total repertoire reflects the number of T cells bearing that specific TCR and the amount of mRNA produced by each T cell, both of which are controlled by the stimulation of T cells by their cognate antigen. The size of public clones in the naive total TCR β 1 repertoire was directly correlated with their usage by different individuals (Figure 3A), suggesting that the same clones are expanded in different individuals in response to antigenic stimulation. Thus, to study the effect of antigenic stimulation on private and public T cell responses, we analyzed the TCR β 1 repertoire 21 days after immunization of naive zebrafish with the self-antigen calmodulin (CALM; Figure S2A), the non-self antigen keyhole limpet hemocyanin (KLH), or administration of the common polyclonal stimulus lectin from *Phaseolus vulgaris* (PHA). Only 16% of the TCR β 1 clones expanded by PHA administration were expanded by immunization with KLH or CALM, suggesting that PHA activates a larger number of TCR β 1 bearing T cells than protein antigens (data not shown).

We then investigated the effect of immunization on the public TCR β 1 repertoire. Immunization with KLH or CALM or polyclonal activation with PHA expanded public clones in the unique and the total TCR β 1 repertoire (Figures 3B and 3C). Indeed, our analyses identified two groups within the public clones: *general public clones*, consisting of TCR β 1 sequences shared by all immunization groups, and *special public clones*, consisting of TCR β 1 sequences shared only by fish that received the same antigenic stimulation. Special public clones were detected in

the unique and total TCR β 1 repertoire following KLH and CALM immunization and also PHA treatment (Figures 3B and 3C). However, immunization with the self-antigen CALM resulted in higher special public responses, suggesting that public TCR β 1 clones are enriched for self reactive elements. Immunization did not affect TCR D, as indicated by the analysis of the Gini coefficient (Figure S3A). Thus, immunization with self and non-self antigens stimulates public T cell responses, which are partially cross-reactive because 41% of the public TCR β 1 expanded by CALM were also expanded by KLH immunization.

The expansion of the public repertoire in response to immunization resulted in part from the expansion of TCR β 1 T cell clones identified as public in naive zebrafish and also from the sharing of TCR β 1 sequences previously identified as private in naive zebrafish (Figures 3D and 3E). Interestingly, most of the public T cell clones in the unique and the total TCR β 1 repertoire, both general and specific, were generated by convergent recombination (Figures 3F, 3G, and S3B). Taken together, these data identify convergent recombination as an important mechanism for the generation of public clones responsive to self and foreign antigens.

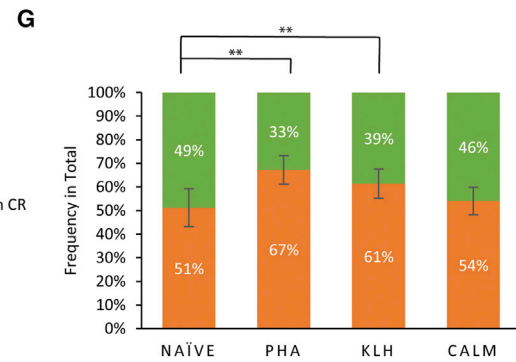
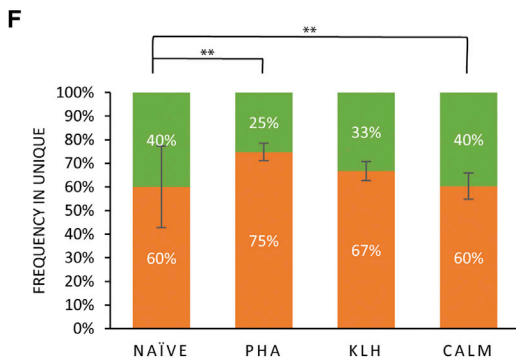
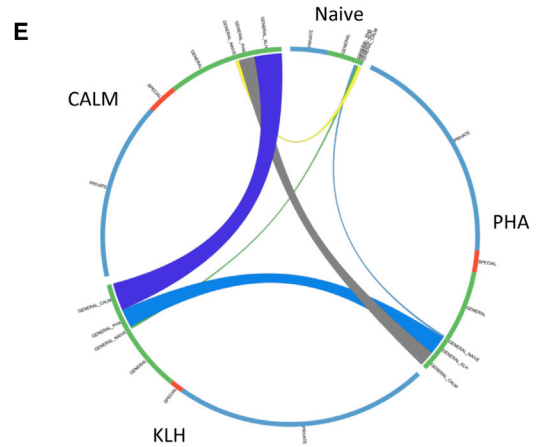
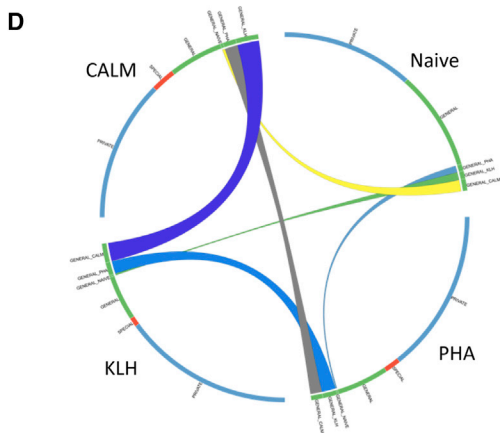
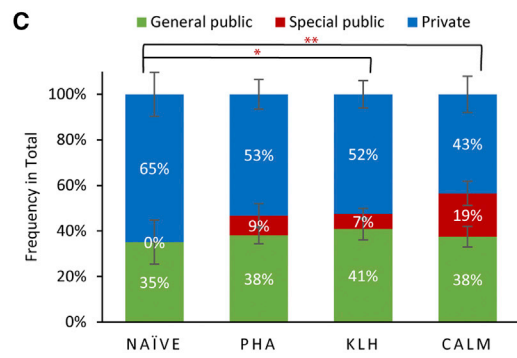
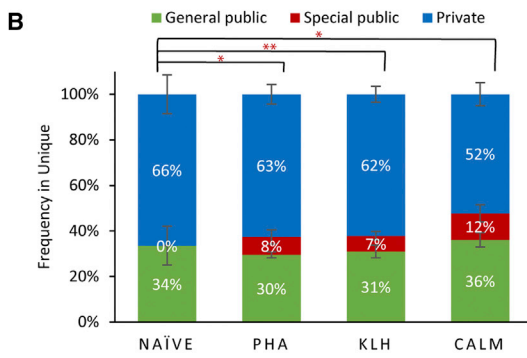
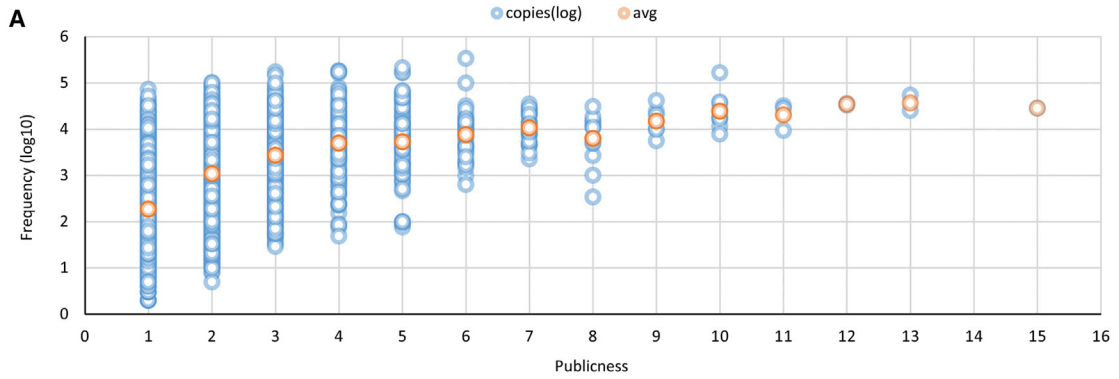
We then investigated the origin of public clones. We found that public clones in KLH-immunized fish or those treated with PHA originated mostly from low-frequency clones in naive zebrafish (Figures 4A and 4B). However, public clones in CALM-immunized fish originated from both high- and low-frequency clones in naive zebrafish, suggesting that self-reactive public T clones are major components of the adult T cell repertoire in naive fish.

Time Course Analysis of the TCR β 1 Repertoire in Response to Immunization

To further elucidate the effect of antigenic stimulation on the T cell response, we analyzed the TCR β 1 repertoire 14, 21, and 28 days after immunization of naive zebrafish with PHA, KLH, and CALM. In these experiments, the zebrafish were boosted by immunization at day 14. Special public TCR β 1 clones were identified at all time points after immunization or PHA administration (Figure 5). In agreement with our previous findings, the T cell response to PHA stimulation or immunization with KLH or CALM was dominated by public TCR β 1 clones generated by convergent recombination (Figures S3C–S3F). However, the clonal responses induced by the different stimuli showed differences in their kinetics. In KLH-immunized and PHA-treated fish, the frequency of special public TCR β 1 clones peaked 14 days after treatment. Immunization with CALM resulted in higher frequencies of both general and special public clones (Figure 5). However, the peak in the number of special public clones expanded by CALM immunization was delayed and was only observed 1 week after boosting, probably reflecting the need for additional antigenic stimulation needed to break self-tolerance (Figure 5). Of note, immunizations over time had no effect on TCR β 1 D, as measured by the Gini coefficient (Figure S3G).

Antigen Stimulation Expands Public Clones in the TCR α Repertoire

A diverse repertoire has also been described for the zebrafish TCR α (Haire et al., 2000). Thus, we analyzed the TCR α repertoire using a primer specific for the C-region of the TCR α chain as



(legend on next page)

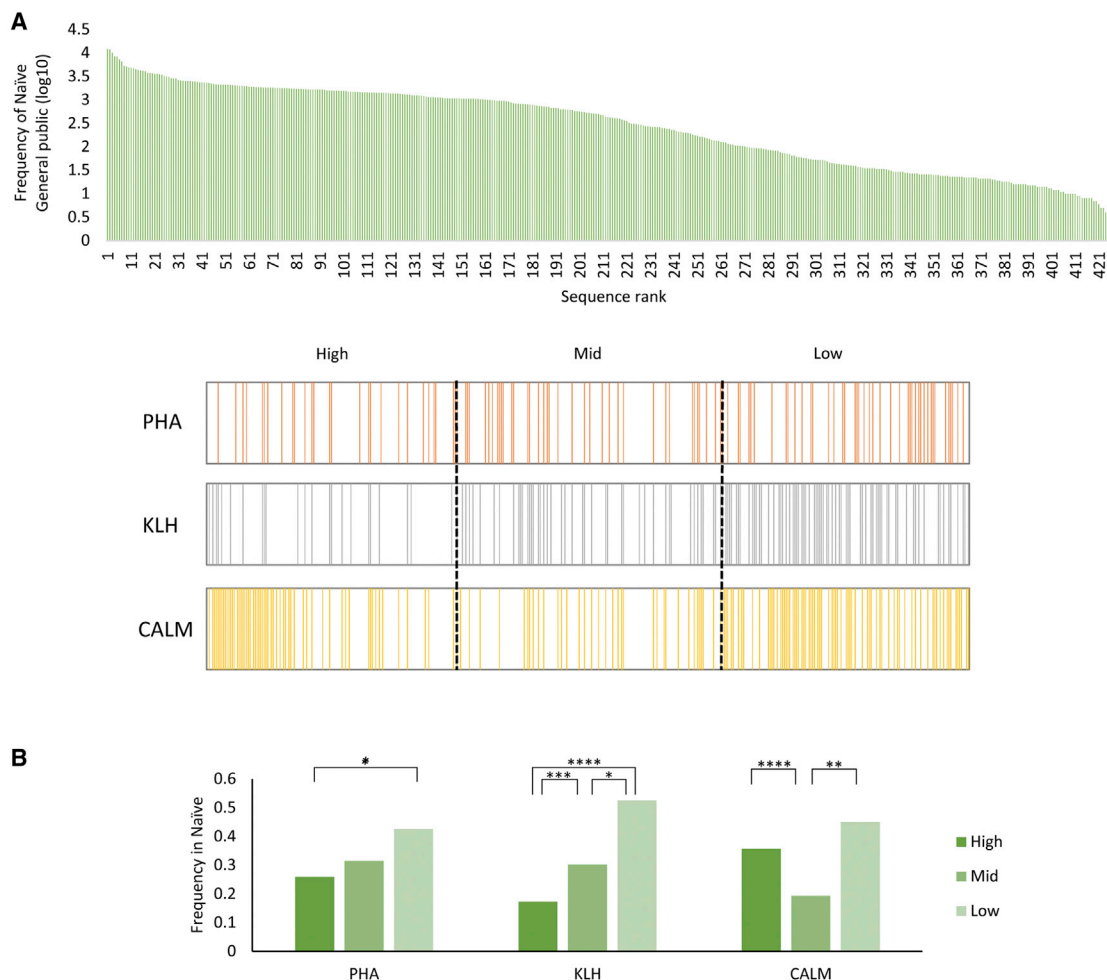


Figure 4. Origin of General Public TCR β 1 Clones

(A) TCR β 1 general public sequences in the naive repertoire, ranked according to their frequency. The lower images show general public sequences in the TCR β 1 repertoire of fish in the PHA, KLH, and CALM groups.

(B) To analyze the origin of general public sequences detected in the immunized repertoire, sequences were classified in the naive TCR β 1 repertoire into low-, mid-, and high-abundant clones and then quantified for the fraction of these groups in the immunized repertoire.

described in Figure 1A. This method was specific because 100% of the amplification products corresponded to TCR α sequences (Figure 6A). Similarly to our observations on the TCR β 1 repertoire, we found that the size of public clones in the naive total TCR α repertoire was directly correlated with their usage by different individuals (Figure 6B).

We then investigated the response of the TCR α repertoire to immunization with KLH or CALM or to PHA administration. Immunization with CALM or KLH or PHA administration did

not affect the number of TCR α unique sequences (Figure S4). Sequence sharing in the TCR α repertoire, however, is at least as strong as that detected in TCR β 1 repertoire, with most TCR α clones being general public sequences (Figures 6C and 6D). We also identified special public TCR α clones following immunization, which were more prevalent in the KLH-immunized fish (Figure 6D). Taken together, these observations suggest that public TCR α sequences in the naive repertoire are expanded in response to antigenic stimulation.

Figure 3. Public Clones Dominate the TCR β 1 Repertoire

(A) Relationship between sequence sharing between different individuals and the number of copies of each TCR β 1 clone. The orange dot for each sharing level gives the average frequency of clones at that specific sharing level.

(B and C) Contribution of private, general public, and special public sequences to the unique (B) and total (C) TCR β 1 repertoire following immunization with KLH, CALM, or polyclonal stimulation with PHA.

(D and E) Sharing of TCR β 1 sequences between the different groups. Each group occupies the same fraction of the circle, regardless of repertoire size, while in (E), each sequence occupies the same portion of the circle. The circle is colored based on whether the TCR β 1 clone is private (blue), special public (red), or general public (green). The edges represent sequences shared between each two groups.

(F and G) Fraction of TCR β 1 clones generated by convergent recombination in the unique repertoire (F) and the total (G) TCR β 1 public repertoire.

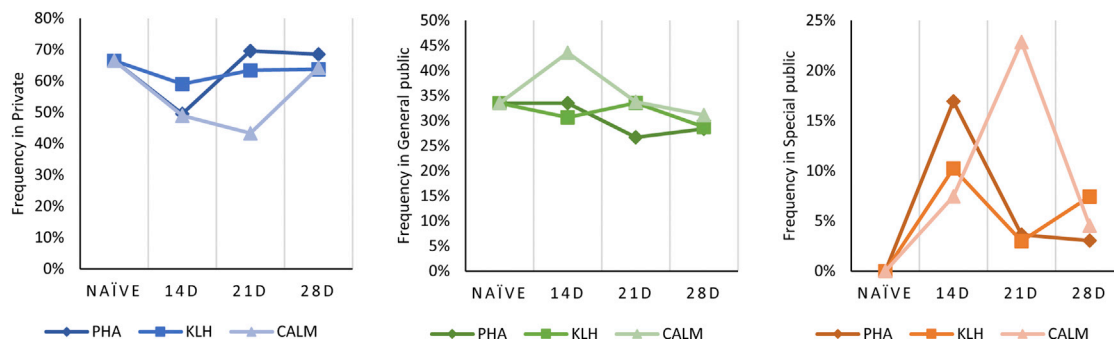


Figure 5. Time Course Analysis of the TCRβ1 Repertoire

Changes in the private, general public, and special public fractions of the TCRβ1 repertoire following treatment with PHA or immunization with KLH or CALM.

We detected a large percentage of clones generated by convergent recombination in the unique TCRα repertoire across all the immunization and treatment groups (Figure 6E), while in the total repertoire, the percentages were even higher, suggesting that T cell clones bearing TCRα public sequences are larger (Figure 6F). Similar to our previous observations with TCRβ public clones, TCRα public clones were mostly generated by CR (Figures 6G–6J). Collectively, these observations suggest that the TCRα repertoire follows the same rules that we described for the TCRβ, being dominated by public clones, some of which are expanded by antigenic stimulation.

Finally, based on the frequency of each TCRα and TCRβ zebrafish clone detected in our sequencing efforts, we constructed a probabilistic model of all potential TCRαβ combinations (Figure 7); a list of the most frequent TCRα and TCRβ sequences used in the construction of the model is provided in Table S2. Based on this model, we estimate that the 2×10^5 T cells present in an adult zebrafish contain at least unique 1.5×10^4 TCRαβ pairs, present in low frequency in the zebrafish TCRαβ repertoire. It should be noted that this is a lower estimate and TCRαβ D may be higher, for example as a result of the expression of more than one α chain by T cells described in humans and other vertebrates (Padovan et al., 1993).

DISCUSSION

In this work, we analyzed the zebrafish TCRβ1 and TCRα repertoire and its response to immunization with self and non-self antigens. We found that the zebrafish TCR repertoire is small and biased toward the use of certain V-J combinations, in a similar manner to what is known from partial repertoire analyses in other vertebrates (Miles et al., 2011). Moreover, the analysis of the TCR repertoire revealed the over-representation of specific V-J pairs, suggestive of clonal expansion in response to antigenic stimulation. These observations suggest that the zebrafish TCR repertoire is shaped by the balance between T cell expansion in response to self and non-self antigens and T cell competition for limited growth and survival factors. In addition, since decreased repertoire D is linked to impaired T cell immunity (Yager et al., 2008), these data suggest that compensatory mechanisms operate in zebrafish to provide protective immunity against pathogens.

Public T cell clones encoded with minimal alteration to germline gene sequences characterize the TCR repertoire of vertebrates (McBerry et al., 2012; Venturi et al., 2008). We detected high frequencies of public TCRαβ sequences in the zebrafish. High-frequency sharing of antibody sequences has also been reported in the zebrafish antibody repertoire, which is also characterized by its small size in agreement with our observations on the TCRαβ repertoire (Jiang et al., 2011; Weinstein et al., 2009). Public T cell clones have been shown to contribute to anti-viral immune responses (Miles et al., 2011). We identified public T cell clones as major components of the zebrafish response to immunization. Moreover, we detected a significant overlap in the public TCRαβ sequences expanded in response to self and foreign antigens that share no sequence homology, suggesting that public T cell clones are highly cross-reactive. Cross-reactivity with self antigens plays an important role in the development of the TCR repertoire and T cell responses to foreign antigens (Birnbbaum et al., 2014; Fulton et al., 2015; Krogsgaard et al., 2005; Mandl et al., 2013; Stefanová et al., 2002). Collectively, these observations suggest that public T cell responses allow a relatively small TCR repertoire to cope with the diverse range of antigens presented by pathogens.

The dominant role of public T cell responses in zebrafish might represent an early step during the evolution of adaptive immunity (Boehm et al., 2012; Cooper and Herrin, 2010; Flajnik and Kasahara, 2010; Guo et al., 2009). However, while the cross-reactivity of public T cell clones could potentially compensate for the small size of the TCR repertoire, it might also increase the risk for the development of pathogenic autoimmunity. Interestingly, a FoxP3 homolog is detectable in zebrafish (Quintana et al., 2010), suggesting that the potential for the development of autoimmunity was co-selected with mechanisms of immune regulation.

Public T cell clones do not constitute a dominant fraction of the TCR repertoire of mice, humans, and other mammals (Miles et al., 2011). However, polyfunctional and cross-reactive public T cell clones are detected in HIV-1 controllers (Chen et al., 2012; Kosmrlj et al., 2010), and similar observations have been made in the context of infection with herpes virus (Zhu et al., 2013), as well as shared self-peptide MHC-specific clones in healthy individuals (Yu et al., 2015). Although it is still unknown whether public TCRs directly control the polyfunctionality and

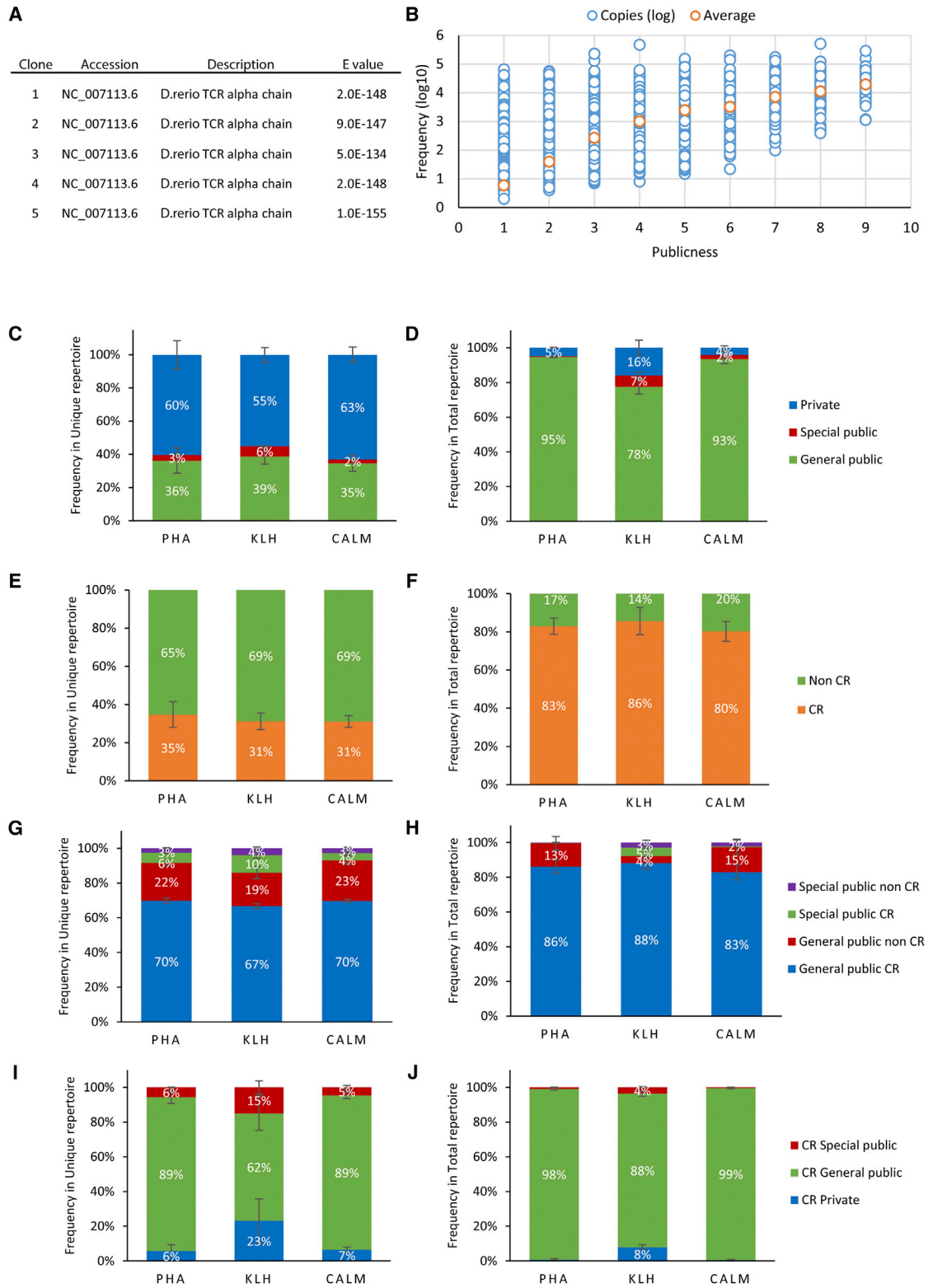


Figure 6. The Zebrafish TCR α Repertoire

(A) Verification of TCR α sequences obtained after library amplification, five clones were sequenced, and the top hits of the blasted sequences are presented in the table.

(legend continued on next page)

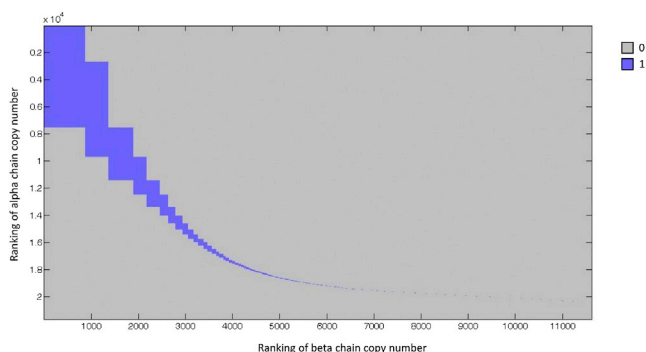


Figure 7. A Computational Model of the TCR $\alpha\beta$ Repertoire

The horizontal axis represents TCR β chains, while the vertical axis represents TCR α chains. Each blue square in the figure represents a potential TCR $\alpha\beta$ pair.

polyreactivity of public T cells (Tubo et al., 2013), these observations suggest that cross-reactive public T cell clones contribute to pathogen control in organisms with larger TCR repertoires. Their restricted D and sharing by different individuals, together with their ability to respond to diverse self and non-self molecules, are features of vertebrate public TCRs that might provide a mechanism for the rapid generation of protective T cell immunity, allowing a short temporal window for the development of more specific private T cell responses.

EXPERIMENTAL PROCEDURES

Fish Maintenance

1-year-old male zebrafish (AB strain) were maintained in a 28–30°C system with a 14/10 hr light/dark cycle in accordance with guidelines by the Institutional Animal Care and Use Committee of Harvard Medical School.

Immunization

Fish were anaesthetized using 0.02% Tricaine methanesulfonate (Sigma-Aldrich) and immunized intraperitoneally (i.p.) with a 10 μ l emulsion containing 1:1 Incomplete Freund's Adjuvant (IFA, Difco Laboratories) and 90% PBS (Invitrogen), 0.25 μ g lipopolysaccharide (ultrapure LPS, InvivoGen), 0.7 μ g CpG oligonucleotide ODN 1826 (InvivoGen), and 2 μ g of either PHA (Sigma-Aldrich), KLH (Sigma-Aldrich), or CALM (Creative BioMart). At 2 weeks later, the fish were boosted with PHA, KLH, or CALM in 1:1 IFA: 90% PBS.

TCR $\alpha\beta$ Sequencing and Annotation

Total RNA was extracted from whole fish homogenate and cDNA was generated. cDNA from each of fish was used for TCR β/α chain library amplification using the 5'PCR primer IIA from the SMARTer Pico PCR cDNA Synthesis Kit (Clontech) and the C-region primer (Table S1). The library was gel-purified and barcodes were added using the same reaction as for the library amplification and the primers listed in Table S1.

TCR β and TCR α annotation was performed by using NCBI BLAST+ to identify the V and J germline genes of a TCR read, and then the CDR3 was

determined by finding the conserved cysteine at the 5' end of the CDR3 and the conserved phenylalanine at the 3' end of the CDR3.

ACCESSION NUMBERS

Sequences have been uploaded to Sequence Read Archive (<http://www.ncbi.nlm.nih.gov/sra>) and can be found under accession number SRA: SRP070056.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, four figures, and two tables and can be found with this article online at <http://dx.doi.org/10.1016/j.celrep.2016.02.056>.

AUTHOR CONTRIBUTIONS

R.C., M.J., J.K., and C.-C.C. performed in vitro and in vivo experiments in zebrafish; J.A. sequenced the zebrafish TCR libraries; H.P., S.D., G.Y., Y.L., L.C., and S.E. performed bioinformatics analysis; R.C., M.J., D.C.D., and F.J.Q. conceived the experimental design; R.C., M.J., D.C.D., S.E., and F.J.Q. wrote the manuscript; R.C., M.J., D.C.D., S.E., and F.J.Q. conceived the study; D.C.D. and F.J.Q. supervised the TCR sequencing; S.E. and F.J.Q. supervised the bioinformatic analysis; and F.J.Q. supervised the overall project.

ACKNOWLEDGMENTS

This work was supported by grants AI075285 and AI093903 from the NIH and a Harry Weaver Scholar Award and grant RG4111A1 from the National Multiple Sclerosis Society to F.J.Q. and grant 2011154 from the BSF to F.J.Q. and S.E. F.J.Q. thanks Maria Ethel del Aguila for useful discussions and support. R.C. is supported by a postdoctoral fellowship from the Swedish Research Council. M.J. is supported by a Sigrid Juselius fellowship, The Paulo Foundation, The Finnish Multiple Sclerosis Foundation, Orion-Farmos Research Foundation, and the Saastamoinen Foundation. This work was funded in part through the intramural program of the National Institute of Allergy and Infectious Diseases, NIH.

Received: April 18, 2015

Revised: January 12, 2016

Accepted: February 2, 2016

Published: March 10, 2016

REFERENCES

- Akira, S., Okazaki, K., and Sakano, H. (1987). Two pairs of recombination signals are sufficient to cause immunoglobulin V-(D)-J joining. *Science* 238, 1134–1138.
- Argaet, V.P., Schmidt, C.W., Burrows, S.R., Silins, S.L., Kurilla, M.G., Doolan, D.L., Suhrbier, A., Moss, D.J., Kieff, E., Sculley, T.B., and Misko, I.S. (1994). Dominant selection of an invariant T cell antigen receptor in response to persistent infection by Epstein-Barr virus. *J. Exp. Med.* 180, 2335–2340.
- Arstila, T.P., Casrouge, A., Baron, V., Even, J., Kanellopoulos, J., and Kourilsky, P. (1999). A direct estimate of the human alphabeta T cell receptor diversity. *Science* 286, 958–961.
- Benichou, J., Ben-Hamo, R., Louzoun, Y., and Efroni, S. (2012). Rep-Seq: uncovering the immunological repertoire through next-generation sequencing. *Immunology* 135, 183–191.

(B) Relationship between sequence sharing and the number of copies of each TCR α clone. The orange dot for each sharing level gives the average frequency of clones at that specific sharing level.

(C and D) Effect of PHA administration or immunization with KLH or CALM in the private, general public, and special public fractions of the unique (C) and total (D) TCR α repertoire.

(E and F) Frequency of clones generated by CR in the unique (E) and total (F) TCR α repertoire following PHA administration or immunization with KLH or CALM.

(G and H) Fraction of CR and non-CR clones in the general public and special public unique (G) and total TCR α repertoire (H).

(I and J) Fractions of private, general public, and special public clones within CR and non-CR clones in the unique (I) and total (J) TCR α repertoire.

- Birnbaum, M.E., Mendoza, J.L., Sethi, D.K., Dong, S., Glanville, J., Dobbins, J., Ozkan, E., Davis, M.M., Wucherpfennig, K.W., and Garcia, K.C. (2014). Deconstructing the peptide-MHC specificity of T cell recognition. *Cell* 157, 1073–1087.
- Boehm, T., Iwanami, N., and Hess, I. (2012). Evolution of the immune system in the lower vertebrates. *Annu. Rev. Genomics Hum. Genet.* 13, 127–149.
- Boyd, S.D., Marshall, E.L., Merker, J.D., Maniar, J.M., Zhang, L.N., Sahaf, B., Jones, C.D., Simen, B.B., Hanczaruk, B., Nguyen, K.D., et al. (2009). Measurement and clinical monitoring of human lymphocyte clonality by massively parallel VDJ pyrosequencing. *Sci. Transl. Med.* 1, 12ra23.
- Burzyn, D., Benoist, C., and Mathis, D. (2013). Regulatory T cells in nonlymphoid tissues. *Nat. Immunol.* 14, 1007–1013.
- Casrouge, A., Beaudoin, E., Dalle, S., Pannetier, C., Kanellopoulos, J., and Kourilsky, P. (2000). Size estimate of the alpha beta TCR repertoire of naive mouse splenocytes. *J. Immunol.* 164, 5782–5787.
- Chen, H., Ndhlovu, Z.M., Liu, D., Porter, L.C., Fang, J.W., Darko, S., Brockman, M.A., Miura, T., Brumme, Z.L., Schneidewind, A., et al. (2012). TCR clonotypes modulate the protective effect of HLA class I molecules in HIV-1 infection. *Nat. Immunol.* 13, 691–700.
- Cibotti, R., Cabaniols, J.P., Pannetier, C., Delarbre, C., Vergnon, I., Kanellopoulos, J.M., and Kourilsky, P. (1994). Public and private V beta T cell receptor repertoires against hen egg white lysozyme (HEL) in nontransgenic versus HEL transgenic mice. *J. Exp. Med.* 180, 861–872.
- Cooper, M.D., and Herrin, B.R. (2010). How did our complex immune system evolve? *Nat. Rev. Immunol.* 10, 2–3.
- Davis, M.M., and Bjorkman, P.J. (1988). T-cell antigen receptor genes and T-cell recognition. *Nature* 334, 395–402.
- Douek, D.C., Betts, M.R., Brenchley, J.M., Hill, B.J., Ambrozak, D.R., Ngai, K.L., Karandikar, N.J., Casazza, J.P., and Koup, R.A. (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.
- Efron, B., and Thisted, R. (1976). Estimating the number of unseen species: How many words did Shakespeare know? *Biometrika* 63, 435–447.
- Fisher, R.A., Corbet, A.S., and Williams, C.B. (1943). The relation between the number of species and the number of individuals in a random sample of an animal population. *J. Anim. Ecol.* 12, 42–58.
- Flajnik, M.F., and Kasahara, M. (2010). Origin and evolution of the adaptive immune system: genetic events and selective pressures. *Nat. Rev. Genet.* 11, 47–59.
- Freeman, J.D., Warren, R.L., Webb, J.R., Nelson, B.H., and Holt, R.A. (2009). Profiling the T-cell receptor beta-chain repertoire by massively parallel sequencing. *Genome Res.* 19, 1817–1824.
- Fujimoto, S., and Yamagishi, H. (1987). Isolation of an excision product of T-cell receptor alpha-chain gene rearrangements. *Nature* 327, 242–243.
- Fulton, R.B., Hamilton, S.E., Xing, Y., Best, J.A., Goldrath, A.W., Hogquist, K.A., and Jameson, S.C. (2015). The TCR's sensitivity to self peptide-MHC dictates the ability of naive CD8(+) T cells to respond to foreign antigens. *Nat. Immunol.* 16, 107–117.
- Guo, P., Hirano, M., Herrin, B.R., Li, J., Yu, C., Sadlonova, A., and Cooper, M.D. (2009). Dual nature of the adaptive immune system in lampreys. *Nature* 459, 796–801.
- Haire, R.N., Rast, J.P., Litman, R.T., and Litman, G.W. (2000). Characterization of three isotypes of immunoglobulin light chains and T-cell antigen receptor alpha in zebrafish. *Immunogenetics* 51, 915–923.
- Howe, K., Clark, M.D., Torroja, C.F., Tarrance, J., Berthelot, C., Muffato, M., Collins, J.E., Humphray, S., McLaren, K., Matthews, L., et al. (2013). The zebrafish reference genome sequence and its relationship to the human genome. *Nature* 496, 498–503.
- Ishizuka, J., Stewart-Jones, G.B., van der Merwe, A., Bell, J.I., McMichael, A.J., and Jones, E.Y. (2008). The structural dynamics and energetics of an immunodominant T cell receptor are programmed by its Vbeta domain. *Immunity* 28, 171–182.
- Jiang, N., Weinstein, J.A., Penland, L., White, R.A., 3rd, Fisher, D.S., and Quake, S.R. (2011). Determinism and stochasticity during maturation of the zebrafish antibody repertoire. *Proc. Natl. Acad. Sci. USA* 108, 5348–5353.
- Kettleborough, R.N., Busch-Nentwich, E.M., Harvey, S.A., Dooley, C.M., de Bruijn, E., van Eeden, F., Sealy, I., White, R.J., Herd, C., Nijman, I.J., et al. (2013). A systematic genome-wide analysis of zebrafish protein-coding gene function. *Nature* 496, 494–497.
- Kosmrlj, A., Read, E.L., Qi, Y., Allen, T.M., Altfeld, M., Deeks, S.G., Pereyra, F., Carrington, M., Walker, B.D., and Chakraborty, A.K. (2010). Effects of thymic selection of the T-cell repertoire on HLA class I-associated control of HIV infection. *Nature* 465, 350–354.
- Krogsgaard, M., Li, Q.J., Sumen, C., Huppa, J.B., Huse, M., and Davis, M.M. (2005). Agonist/endogenous peptide-MHC heterodimers drive T cell activation and sensitivity. *Nature* 434, 238–243.
- Li, H., Ye, C., Ji, G., Wu, X., Xiang, Z., Li, Y., Cao, Y., Liu, X., Douek, D.C., Price, D.A., and Han, J. (2012). Recombinatorial biases and convergent recombination determine interindividual TCRbeta sharing in murine thymocytes. *J. Immunol.* 189, 2404–2413.
- Lieschke, G.J., and Trede, N.S. (2009). Fish immunology. *Curr. Biol.* 19, R678–R682.
- Linnemann, C., Heemskerck, B., Kvistborg, P., Kluijn, R.J., Bolotin, D.A., Chen, X., Bresser, K., Nieuwland, M., Schotte, R., Michels, S., et al. (2013). High-throughput identification of antigen-specific TCRs by TCR gene capture. *Nat. Med.* 19, 1534–1541.
- Malissen, M., McCoy, C., Blanc, D., Trucy, J., Devaux, C., Schmitt-Verhulst, A.M., Fitch, F., Hood, L., and Malissen, B. (1986). Direct evidence for chromosomal inversion during T-cell receptor beta-gene rearrangements. *Nature* 319, 28–33.
- Mandl, J.N., Monteiro, J.P., Vrisekoop, N., and Germain, R.N. (2013). T cell-positive selection uses self-ligand binding strength to optimize repertoire recognition of foreign antigens. *Immunity* 38, 263–274.
- McBerry, C., Gonzalez, R.M., Shryock, N., Dias, A., and Aliberti, J. (2012). SOCS2-induced proteasome-dependent TRAF6 degradation: a common anti-inflammatory pathway for control of innate immune responses. *PLoS ONE* 7, e38384.
- Meeker, N.D., Smith, A.C., Frazer, J.K., Bradley, D.F., Rudner, L.A., Love, C., and Trede, N.S. (2010). Characterization of the zebrafish T cell receptor beta locus. *Immunogenetics* 62, 23–29.
- Miles, J.J., Bulek, A.M., Cole, D.K., Gostick, E., Schauenburg, A.J., Dolton, G., Venturi, V., Davenport, M.P., Tan, M.P., Burrows, S.R., et al. (2010). Genetic and structural basis for selection of a ubiquitous T cell receptor deployed in Epstein-Barr virus infection. *PLoS Pathog.* 6, e1001198.
- Miles, J.J., Douek, D.C., and Price, D.A. (2011). Bias in the $\alpha\beta$ T-cell repertoire: implications for disease pathogenesis and vaccination. *Immunol. Cell Biol.* 89, 375–387.
- Moss, P.A., Moots, R.J., Rosenberg, W.M., Rowland-Jones, S.J., Bodmer, H.C., McMichael, A.J., and Bell, J.I. (1991). Extensive conservation of alpha and beta chains of the human T-cell antigen receptor recognizing HLA-A2 and influenza A matrix peptide. *Proc. Natl. Acad. Sci. USA* 88, 8987–8990.
- Newell, E.W., and Davis, M.M. (2014). Beyond model antigens: high-dimensional methods for the analysis of antigen-specific T cells. *Nat. Biotechnol.* 32, 149–157.
- Nikolich-Zugich, J., Slifka, M.K., and Messaoudi, I. (2004). The many important facets of T-cell repertoire diversity. *Nat. Rev. Immunol.* 4, 123–132.
- Okazaki, K., Davis, D.D., and Sakano, H. (1987). T cell receptor beta gene sequences in the circular DNA of thymocyte nuclei: direct evidence for intramolecular DNA deletion in V-D-J joining. *Cell* 49, 477–485.
- Padovan, E., Casorati, G., Dellabona, P., Meyer, S., Brockhaus, M., and Lanzavecchia, A. (1993). Expression of two T cell receptor alpha chains: dual receptor T cells. *Science* 262, 422–424.
- Park, C.O., and Kupper, T.S. (2015). The emerging role of resident memory T cells in protective immunity and inflammatory disease. *Nat. Med.* 21, 688–697.

- Quigley, M.F., Greenaway, H.Y., Venturi, V., Lindsay, R., Quinn, K.M., Seder, R.A., Douek, D.C., Davenport, M.P., and Price, D.A. (2010). Convergent recombination shapes the clonotypic landscape of the naive T-cell repertoire. *Proc. Natl. Acad. Sci. USA* *107*, 19414–19419.
- Quintana, F.J., Iglesias, A.H., Farez, M.F., Caccamo, M., Burns, E.J., Kassam, N., Oukka, M., and Weiner, H.L. (2010). Adaptive autoimmunity and Foxp3-based immunoregulation in zebrafish. *PLoS ONE* *5*, e9478.
- Robins, H.S., Campregher, P.V., Srivastava, S.K., Wachter, A., Turtle, C.J., Kahsai, O., Riddell, S.R., Warren, E.H., and Carlson, C.S. (2009). Comprehensive assessment of T-cell receptor beta-chain diversity in alphabeta T cells. *Blood* *114*, 4099–4107.
- Rudolph, M.G., Stanfield, R.L., and Wilson, I.A. (2006). How TCRs bind MHCs, peptides, and coreceptors. *Annu. Rev. Immunol.* *24*, 419–466.
- Stefanová, I., Dorfman, J.R., and Germain, R.N. (2002). Self-recognition promotes the foreign antigen sensitivity of naive T lymphocytes. *Nature* *420*, 429–434.
- Tonegawa, S. (1983). Somatic generation of antibody diversity. *Nature* *302*, 575–581.
- Trede, N.S., Langenau, D.M., Traver, D., Look, A.T., and Zon, L.I. (2004). The use of zebrafish to understand immunity. *Immunity* *20*, 367–379.
- Tubo, N.J., Pagán, A.J., Taylor, J.J., Nelson, R.W., Linehan, J.L., Ertelt, J.M., Huseby, E.S., Way, S.S., and Jenkins, M.K. (2013). Single naive CD4+ T cells from a diverse repertoire produce different effector cell types during infection. *Cell* *153*, 785–796.
- Turner, S.J., Doherty, P.C., McCluskey, J., and Rossjohn, J. (2006). Structural determinants of T-cell receptor bias in immunity. *Nat. Rev. Immunol.* *6*, 883–894.
- Turner, S.J., La Gruta, N.L., Kedzierska, K., Thomas, P.G., and Doherty, P.C. (2009). Functional implications of T cell receptor diversity. *Curr. Opin. Immunol.* *21*, 286–290.
- Venturi, V., Kedzierska, K., Price, D.A., Doherty, P.C., Douek, D.C., Turner, S.J., and Davenport, M.P. (2006). Sharing of T cell receptors in antigen-specific responses is driven by convergent recombination. *Proc. Natl. Acad. Sci. USA* *103*, 18691–18696.
- Venturi, V., Price, D.A., Douek, D.C., and Davenport, M.P. (2008). The molecular basis for public T-cell responses? *Nat. Rev. Immunol.* *8*, 231–238.
- Vermijlen, D., Brouwer, M., Donner, C., Liesnard, C., Tackoen, M., Van Rysselberge, M., Twité, N., Goldman, M., Marchant, A., and Willems, F. (2010). Human cytomegalovirus elicits fetal gammadelta T cell responses in utero. *J. Exp. Med.* *207*, 807–821.
- Wang, C., Sanders, C.M., Yang, Q., Schroeder, H.W., Jr., Wang, E., Babrzadeh, F., Gharizadeh, B., Myers, R.M., Hudson, J.R., Jr., Davis, R.W., and Han, J. (2010). High throughput sequencing reveals a complex pattern of dynamic interrelationships among human T cell subsets. *Proc. Natl. Acad. Sci. USA* *107*, 1518–1523.
- Weinstein, J.A., Jiang, N., White, R.A., 3rd, Fisher, D.S., and Quake, S.R. (2009). High-throughput sequencing of the zebrafish antibody repertoire. *Science* *324*, 807–810.
- Yager, E.J., Ahmed, M., Lanzer, K., Randall, T.D., Woodland, D.L., and Blackman, M.A. (2008). Age-associated decline in T cell repertoire diversity leads to holes in the repertoire and impaired immunity to influenza virus. *J. Exp. Med.* *205*, 711–723.
- Yancopoulos, G.D., Blackwell, T.K., Suh, H., Hood, L., and Alt, F.W. (1986). Introduced T cell receptor variable region gene segments recombine in pre-B cells: evidence that B and T cells use a common recombinase. *Cell* *44*, 251–259.
- Yu, W., Jiang, N., Ebert, P.J., Kidd, B.A., Müller, S., Lund, P.J., Juang, J., Adachi, K., Tse, T., Birnbaum, M.E., et al. (2015). Clonal deletion prunes but does not eliminate self-specific $\alpha\beta$ CD8(+) T lymphocytes. *Immunity* *42*, 929–941.
- Zhu, J., Peng, T., Johnston, C., Phasouk, K., Kask, A.S., Klock, A., Jin, L., Diem, K., Koelle, D.M., Wald, A., et al. (2013). Immune surveillance by CD8 $\alpha\alpha$ + skin-resident T cells in human herpes virus infection. *Nature* *497*, 494–497.