Determinism and stochasticity during maturation of the zebrafish antibody repertoire

Ning Jiang[a,1], Joshua A. Weinstein[b,1], Lolita Penland[a], Richard A. White III[c,2], Daniel S. Fisher[d,e], and Stephen R. Quake[a,b,c,d,3]

[a]Department of Bioengineering, [b]Biophysics Graduate Program, [c]Howard Hughes Medical Institute, [d]Department of Applied Physics, and [e]Department of Biology, Stanford University, Stanford, CA 94305

It is thought that the adaptive immune system of immature organisms follows a more deterministic program of antibody creation than is found in adults. We used high-throughput sequencing to characterize the diversifying antibody repertoire in zebrafish over five developmental time points. We found that the immune system begins in a highly stereotyped state with preferential use of a small number of V (variable) D (diverse) J (joining) gene segment combinations, but that this stereotypy decreases dramatically as the zebrafish mature, with many of the top VDJ combinations observed in 2-wk-old zebrafish virtually disappearing by 1 mo. However, we discovered that, in the primary repertoire, there are strong correlations in VDJ use that increase with zebrafish maturity, suggesting that VDJ recombination involves a level of deterministic programming that is unexpected. This stereotypy is masked by the complex diversification processes of antibody maturation; the variation and lack of correlation in full repertoires between individuals appears to be derived from randomness in clonal expansion during the affinity maturation process. These data provide a window into the mechanisms of VDJ recombination and diversity creation and allow us to better understand how the adaptive immune system achieves diversity.

immunology | repertoire development | secondary repertoire

It is well known that, in human infants, the adaptive immune system is not fully functional: infants receive maternal antibodies through their mother's milk, and, in some cultures, infants and their mothers are sequestered for the first month of life to prevent inadvertent exposure to pathogens. Young children are at higher risk from infectious disease (1) such as influenza and receive priority in vaccination policies (2). At the molecular level, it has been observed that there are significant differences between infants and adults in the constitution of the immune repertoire (3). Although it is a common view that initial Ig diversity is primarily generated as a random process, there is evidence for deterministic, programmed repertoire development in fetal repertoires: certain favored V segments are vastly overrepresented in both the mouse and in human beings (4–7). Early studies of segment use bias in murine tissues concluded that the fetal repertoire is stereotyped relative to adult repertoire (4, 5). Preferential segment use in adult repertoires were distinct from the fetal stereotypy (8). It has also been shown that the newly generated repertoire is similarly biased in both fetal liver and adult bone marrow B lineage cells (9, 10). Little progress has been made in our understanding of the early repertoire since these studies were published. The most prominent mechanisms that have been implicated in the determinism versus stochasticity debate are as follows: (i) the role of variation in recombination signal sequences, which flank each of the V (variable), D (diverse), and J (joining) gene segments and are recognized by VDJ recombinase, thereby causing certain segments to be favored (11); and (ii) the observation that expression of the terminal deoxyribonucleotidyl transferase enzyme is suppressed in infants, thus reducing the creation of diversity through N addition

(12, 13), i.e., nontemplated random nucleotide insertion between VD and DJ junctions.

Despite the medical and biological importance of these issues, there is still much to learn about how the adaptive immune system diversifies in parallel with the development of the organism. How is the immune repertoire created: Is it random, or is there a common program shared among individuals? Whether the VDJ repertoire of mature B cells in adults is constitutively biased, and, if so, whether similarly or differently from the fetus is a key open question. To address this, two mechanisms affecting immune repertoire measurement must be separated. The first is the creation of the primary repertoire, which we define as the diversity that is the product purely of VDJ recombination and associated junctional diversity. Each such recombination creates a B cell clone that can expand to form a lineage or die out. Each lineage can then be diversified via somatic hypermutation, and some of the resulting clonal populations expanded to varying degrees by antigen stimulation. We call the resulting spectrum of clones the secondary repertoire. The secondary repertoire thus reflects a diversified, biased, and distorted view of the primary repertoire, caused by somatic hypermutation and clonal expansion.

Previous studies of the development of immune repertoires have been limited by sampling statistics; before the advent of rapid and highly parallel sequencing technologies, it was not possible to extensively characterize the immune repertoire (14). We used high-throughput sequencing to study the evolving immune repertoire of zebrafish as a function of developmental state. The zebrafish is a powerful immunological model system because it shares essential elements of the adaptive immune system with the mouse and human being, including VDJ recombination, junctional diversity, and somatic hypermutation (15). Even the antibody heavy chain locus arrangement is similar to that in the mouse and human being: V gene segments, shared between two heavy chains μ and ζ, are followed by D and J gene segments for both μ and ζ heavy chains (16). However, the number of B cells is many orders of magnitude smaller, enabling far more complete studies than are feasible in mice, let alone human beings. We measured repertoires at a depth that allowed us to analyze Pearson correlations between

IMMUNOLOGY

individuals' VDJ usage and fuller measures of the repertoire, as well as junctional diversity and somatic hypermutation. This allowed us to analyze the repertoire on three levels—VDJ, primary, and secondary—at a level of specificity and completeness far beyond previous measurements of preferential segment usage. We characterized the specific recombinational biases, revealing a much more highly stereotyped developmental repertoire than has previously been suspected. Surprisingly, this conclusion is valid not just for immature animals. Although correlations in VDJ abundances between individual fish disappear in adult repertoires, we discovered previously unknown correlations in their primary repertoires. Our analysis suggests that the main process that causes differentiation between adult repertoires is apparently random clonal expansion, or more generally the increased transcription of antibody mRNA, in the secondary repertoire. This is also reflected in the specific antibody sequences in the repertoire; a key finding is that in adult fish the most highly expressed sequences have multiple mutations, suggesting clonal selection, whereas in young fish the repertoire is dominated by sequences with few mutations.

Results

Highly Stereotyped VDJ Usage in Young Animals. A total of 51 fish were collected over five time points; about half of them were from the same family (*SI Appendix,* Table S2). Each family started with a single clutch of eggs of more than 60. Fish shared development with their family until time of collection.

We have previously described the method used to generate and sequence antibody heavy chain cDNA libraries for each fish (14). A total of 19 sequencing runs of 51 libraries yielded an average of 82,000 useful reads for each fish (*SI Appendix,* Table S2). In zebrafish, differential expression of the two antibody isotypes has been documented, and there is some knowledge about their gross developmental progression; IgZ expresses as early as 4 d postfertilization and IgM starts 2 d later (16). We focused our analysis on IgM because its abundance grows quickly during maturation to a large majority over IgZ (*SI Appendix,* Table S2), and class switching has not been found between these two isotypes.

The number of V (39), D (5), and J (5) gene segments gives rise to 975 possible VDJ combinations. The collection of all VDJ combinations used is referred to as the VDJ repertoire. Individual fish VDJ repertoires can be visualized in graphical form.

A few dominant combinations are apparent (Fig. 1); nevertheless, ~200 (20%) of the possible VDJ combinations are expressed as IgM in an individual 2-wk-old fish at a sequencing depth of 32,000 reads. This number rapidly increases over time to ~700 combinations, which is ~70% of the total possible VDJ repertoire (Fig. 2*A*). Interestingly, we observed decreased VDJ usage in 1-y-old fish, which could be a sign of immune senescence. A similar trend in VDJ repertoire usage was observed for IgZ (*SI Appendix,* Fig. S4).

From the visualizations of the VDJ repertoires, it is apparent that young fish (2 wk) have striking similarities in their VDJ use: the same handful of combinations dominate the repertoire (Fig. 1, *Top Three Panels*). In adults, by contrast, this does not appear to occur (Fig. 1, *Bottom Three Panels*), with broader VDJ occupancy accompanying a fewer number of highly expressed VDJ combinations. To quantify the biases in VDJ combination use, we compared the set of fish in two ways: in terms of the most abundant VDJ combinations, and in terms of overall usage. We examined the most populated VDJ combinations in each age group and measured the overlap between individuals. Consistent with the notion that the antibody repertoire is stereotyped early in age, among young fish there is strong overlap between the top 10 VDJ combinations in individuals; this overlap decreases dramatically in mature fish (Fig. 2*B*). This preferential use can also be quantified with the complete repertoire (beyond the dominant combinations): Pearson correlations between VDJ repertoires show that as early as 2 wk, there is a high degree of correlation between VDJ usage by individuals (red line, Fig. 2*D*) which is largely dictated by the most abundant VDJ combinations (green and blue lines, Fig. 2*D*).

As the fish mature, the correlations decrease; distinct and less correlated sets of VDJ combinations begin to dominate each individual (Figs. 1 and 2 *B* and *C*). The average correlation coefficient calculated between fish pairs in each age group was much higher for 2-wk samples (0.8) than for any other age groups (0.06–0.08, Fig. 2*D*) or the average among all pairs of fish (0.09). Such a structured VDJ repertoire in young organisms is surprising and has not been demonstrated before, probably because of technical limitations of cloning or hybridization methods (4–7). Consistently, the VDJ combinations with the first, second, third, and fifth most reads at 2 wk (V4.6-D1-J3, V2.1-D5-J1, V4.3-D5-J5, and
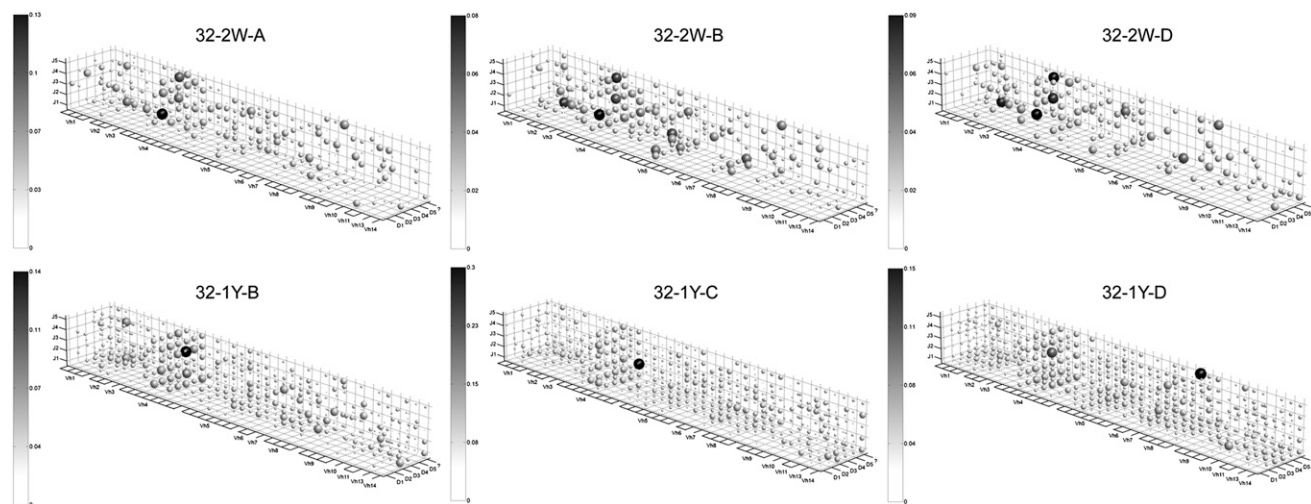
Fig. 1. 3D representations of VDJ repertoires for subsamples of 2-wk-old (*Top* three panels) and 1-y-old (*Bottom* three panels) samples from a single family. [Sample name is ordered: family-age-(letter ID)]. Dot size scales logarithmically with bias-normalized abundance, whereas coloring scales linearly. Color bars indicate fraction of total abundance. Stereotypy in abundance among a large number of VDJ combinations is found to disappear. More generally, although at an early age multiple VDJ combinations are found to share high abundances, later ages display great abundance disparity, with a far smaller number of VDJ combinations holding a greater amount of abundance.

Fig. 2. Characteristics of the VDJ use. (*A*) Time course of VDJ use for IgM heavy chains at a fixed sequencing depth of 32,000 reads among the 975 possible IgM VDJ combinations. Black line represents bounds by the most- and least-represented fish in that age group; solid blue line represents median value within that age group; and dotted blue line represents sample SD away from that value. (*B*) Top 10 VDJ (ranked by bias-normalized abundance) from fish subsampled to 32,000 reads at 2 wk and 1 mo of age. Individual fish are represented by columns. Differently colored shapes represent VDJ combinations that occur more than once in the diagram (the same symbol occurring anywhere in the two subplots represents the same VDJ combination); black dots represent VDJ combinations that occur only in one individual's top 10 VDJs. (*C*) Top 5 VDJ combinations (ranked by read number) at 2 wk tracked over subsequent ages. Data from fish that are sampled to a depth of 32,000 reads are listed. (*D*) Average and sample SD of time course correlating the VDJ repertoires of pairs of individuals for different age groups. VDJ abundances (red) are bias normalized. Correlations of th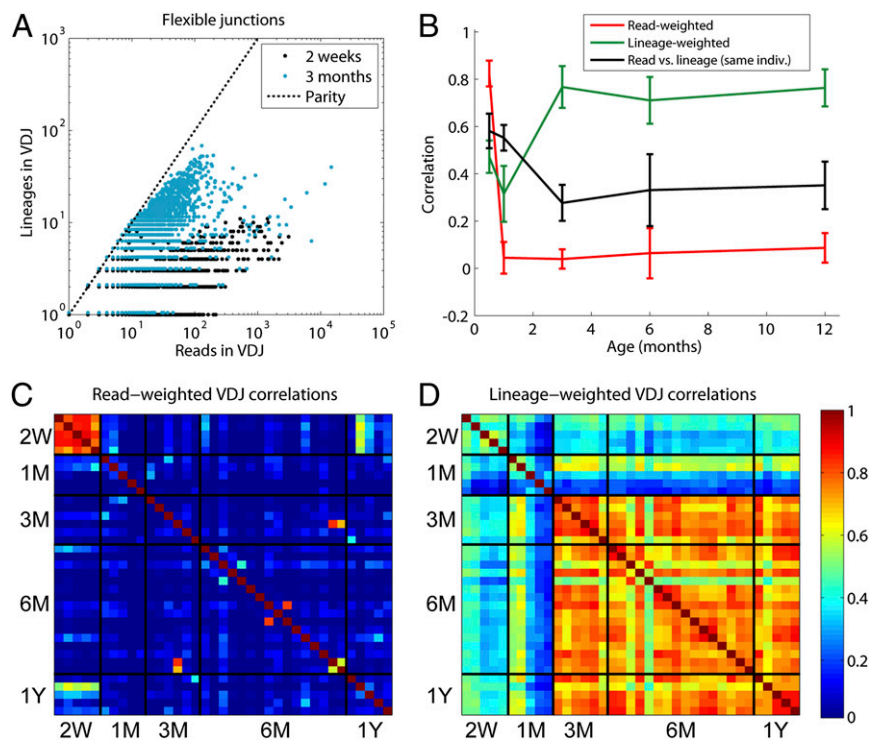e square roots of VDJ bias-normalized abundances (green) reduce the impact of the most abundant VDJ combination. To the same end, correlations of VDJ presence (cyan) are performed by assigning each VDJ combination a 1 or 0 depending on its being observed.

V5.4-D5-J4) are virtually nonexistent at 1 mo (Fig. 2*C*). The 3D representation of the single highly correlated VDJ repertoire-pair at 1 y (*SI Appendix*, Fig. S3) reveals a much lower degree of stereotypy for all but the most abundant VDJ combination, suggesting that the similarity itself is of a qualitatively different nature than its early-development counterparts.

VDJ Usage of Antibody Primary Repertoire Is Also Stereotyped in Mature Animals. Although a main focus in the literature has been VDJ-segmental use, this neglects a crucial part of the primary repertoire: The junctional diversity that differs among the set of lineages created within the same VDJ combinations. To explore this, we developed an informatic approach to characterize the primary repertoire from our sequence data. Because of sequencing and PCR errors and hypermutations, clustering sequences into lineages is not trivial. We first took a liberal approach and used single linkage clustering to allow sequences of the same VJ combination that differed slightly at the VDJ-junction, by at most 20%, due to mutation and error to be grouped together (*SI Appendix*). We mean-

while grouped sequences from the same VJ combination that varied outside the VDJ junction alone. This aimed to eliminate the effect of differences in VDJ junction abundance. To place bounds on the uncertainties in this clustering, we compared the results with a conservative approach of assigning sequences to the same lineage only if they were identical in their junction regions (as well as their VJ combinations). The qualitative and semiquantitative results were not strongly affected by the differences among these approaches (Fig. 3 *C* and *D* and *SI Appendix*, Fig. S8), ensuring that the conclusions are robust.

To separate the effects of varying probabilities of creation of different VDJ combinations from differences in the subsequent successes of the lineages produced, we need to ask whether large numbers of reads of a given VDJ combination are associated with a proportionally large number of distinct lineages. If there were a very close correspondence, it might, on one hand, implicate the rates of formation of the VDJ combinations in their subsequent abundances. On the other hand, such an observation

Fig. 3. Characteristics of primary repertoire. (*A*) Scatter plot of number of lineages and number of reads observed across the 2-wk and 1-mo age groups (a small offset along the *y* axis allows both sets of data points to be seen), revealing more lineages in mature fish. (*B*) Average and sample SDs of correlations between different members of the same age-group with VDJ repertoires weighted by raw reads (red), without bias-normalization, and lineages (green). For each individual, read-weighted and lineage-weighted VDJ repertoires are further correlated (black). Color maps of read-weighted (*C*) and lineage-weighted (*D*) VDJ correlations between fish sampled to 40,000 read depths. Age groups are delineated as 2 wk (2W), 1 mo (1M), 3 mo (3M), 6 mo (6M), and 1 y (1Y).

might also implicate sequencing artifact: The discovery of new lineages within a VDJ combination could strongly depend on sequencing depth.

To address these questions, we compared the number of raw reads within each VDJ combination to the number of distinct lineages discovered in them. Scatter plots of early and late development VDJ read abundance and VDJ lineage diversity (Fig. 3*A*, subsampling done to 40,000 identifiable IgM reads for all individuals with at least that many) point to both an increase with age of the total lineage diversity and to a trend, with increasing age, where the VDJ combinations with the most abundance are not in general the VDJ combinations with the most lineage diversity. We sought to further visualize this by plotting the correlations between these two quantities within each individual (black line, Fig. 3*B*). The decaying correlation demonstrates that in general, with increasing age, a VDJ's diversity becomes less predictive of its abundance. This result recapitulates the loss of VDJ abundance stereotypy observed earlier. However, it leaves open the question of whether, although stereotypy in abundance (associated with the secondary repertoire) decreases, the primary repertoire itself undergoes a change in stereotypy.

To answer this, we compared the VDJ lineage diversities between individuals of the same age group. Surprisingly, stereotypy in VDJ lineage diversities, measured by correlating pairs of individuals within the same age group, undergoes an abrupt increase after 3 mo (green line, Fig. 3*B*). To ascertain whether the same VDJ lineage diversity stereotypy is conserved across older age groups, we plotted correlations of VDJ read abundances and VDJ lineage-diversities as color-maps (Figs. 3*C* and 3*D*) across all fish subsampled to 40,000 reads. The color maps establish that whereas the characteristic 2-wk abundance stereotypy disappears immediately, the primary repertoire exhibits the very opposite behavior, with individuals aged 3 mo and older retaining strong

stereotypy. This observation eliminates the possibility that differences in observed lineage diversities are dominated by different VDJ sequencing depths.

Dynamics of Antibody Secondary Repertoire Development. Having characterized aspects of the primary repertoire, we now turn to the secondary repertoire: This is formed from the primary repertoire by the interplay between hypermutation and clonal expansion caused by selection for particular antibodies. We characterized the secondary repertoire by computing the diversity in terms of amino acid use for the most highly represented VJ combination in each fish (Fig. 4 *A* and *B*), the combinations themselves subsampled to 5,000 reads each. The number of amino acids used fluctuates at a constant level throughout the entire sequence in the 2-wk samples. In other age groups, this number has a spike close to the junction which gradually increases in time. The fluctuation along the rest of the sequences increases as well.

Early in development, we expect little correspondence between the number of amino acid changes from hypermutation and the abundance of clones. However, exposure to antigens is likely to result in proliferation and more amino acid changes. This trend emerges in plots of numbers of amino acid mutations as a function of antibody abundance (Fig. 4 *C* and *D*). In 2-wk and 1-mo samples, the average number of amino acid mutations per sequence decreases monotonically as the sequence abundance over which the mutation average is taken increases. However, with increasing age (up through 1 y), highly abundant sequences become increasingly likely to harbor higher numbers of mutations.

Discussion

VDJ recombination creates the first level of antibody repertoire diversity. Developmentally it has been shown that V gene segments
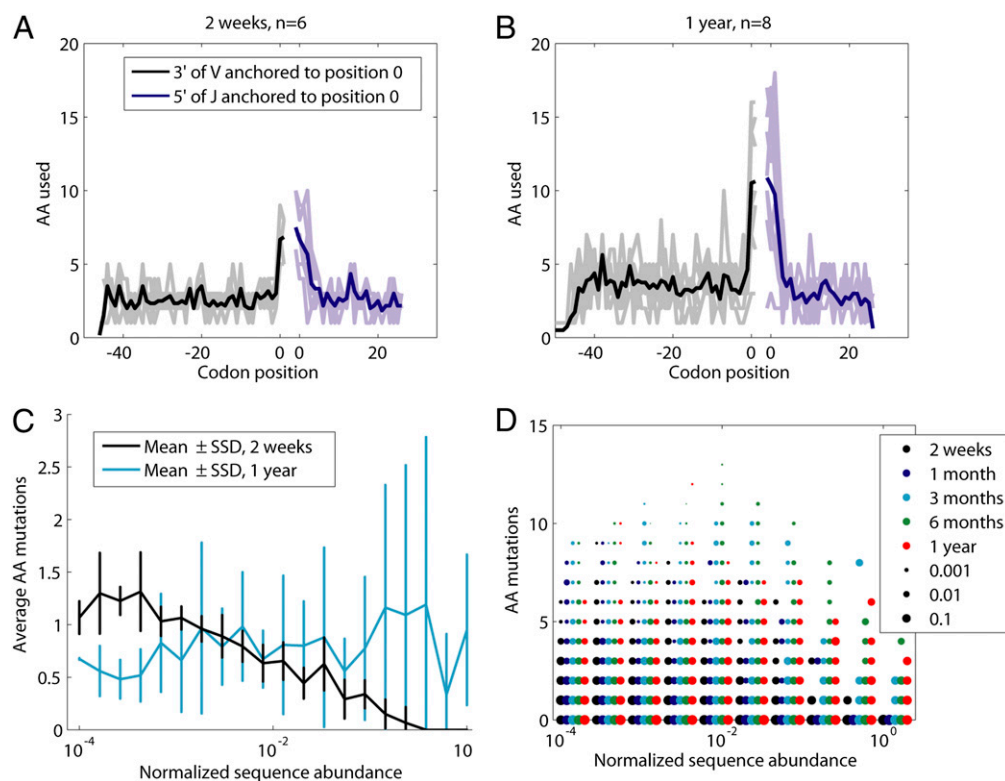
Fig. 4. Characteristics of the secondary repertoire. Amino acid diversification at 2 wk (*A*) and 1 y (*B*) for the most abundant VJ combination in each fish, the combination itself subsampled, to 5,000 reads. Amino acids are counted at a particular position if they occur at least once in the data set. Observed D-gene segment diversity across these two datasets averages 3.8 and 4.0, respectively. The 3-prime end of V gene segments (black) and the 5-prime end of J gene segments (blue) are anchored to position zero. Dark colors denote the average over the amino acid use of individual fish (drawn in light colors). (*C*) Average amino acid mutations as a function of sequence abundance at 2-wk and 1-y time points, using lineage analysis. Mutation is defined as the amino acid differences between one sequence and the sequence within the same lineage that most closely resembles the reference genomic sequences. Sequencing depth is fixed at 40,000 reads. Abundances are both bias normalized and divided by the most abundant sequence in the repertoire. Twenty log-spaced bins are spaced evenly between 0.0001 and 1. (*D*) Using the 10 bins in the same interval, the distribution of amino acid mutations is plotted with dot sizes that scale as the fraction of the total unique sequences at a particular abundance across all fish in the color-coded age group.

that are close to the D and J gene segments are preferentially used by young organisms (4–7). Using high-throughput sequencing, we analyzed the antibody heavy chain repertoire for 51 zebrafish spanning five developmental time points. We noticed that the same handful of VDJ combinations dominate the young repertoire and that this trend disappears as the fish mature. However, there is no obvious connection between V gene segments in these early VDJ combinations and their relative positions to D and J gene segments on the heavy chain locus (16); perhaps this restriction is related to chromatin structure that makes certain gene segments more accessible than others (17). It also suggests that either there is a common mechanism to constrain the ontogeny of B cell development, or else possibly that the adaptive immune system first trains on a common set of autoantigens (18), and the corresponding autoreactive B cells are eliminated later by selection processes (19). Consistent with this early stereotypy is the narrower distribution of N-nucleotide insertions observed in 2-wk samples compared to a much wider distribution in later age groups (*SI Appendix*, Fig. S6).

Using single linkage clustering, we were able to informatically reconstruct the primary repertoire from the secondary repertoire, thereby removing the effect of clonal expansion. Surprisingly, the VDJ usage of the primary repertoire exhibits a unique type of stereotypy with a smaller correlation at 2 wk followed by a decrease at 1 mo then a high correlation from 3 mo onward. The emerging new stereotypy for older animals is unexpected. This suggests that VDJ recombination is an ordered process in generating antibody repertoire throughout life, even though clonal expansion may distort VDJ abundances. Simple differences in the recombination

signal sequences may not explain VDJ correlation dynamics we have observed, as RSS sequences in fish do not change over development. VDJ recombination must therefore be examined in the context of other factors that do change during development.

The smaller correlation between the primary repertoires of individuals at a young age does not necessarily mean that their VDJ formation programs are less stereotyped than at older ages. In fact, the significant early-age correlations between the read-abundance and lineage-diversity in each individual's VDJ combinations suggest a close correspondence between the primary repertoire and the stereotyped secondary repertoire in young individuals.

In this study we also implemented VDJ junction and mutation analysis and compared junctional diversity and sequence mutations as a function of sequence abundance among different age groups. Our results suggest that most of the highly represented repertoire in young fish comprises naive B cells with few mutations, and that only a small fraction of the B cells have undergone somatic hypermutation and antigen selection. On longer time scales, interaction with a fish's environment changes the composition of its repertoire. As a result, mutated sequences dominate the part of the repertoire that is highly expressed. Our analysis of the spread of lineage diversities and gross amino acid usage suggests that diversification of the repertoire in older fish may also involve a higher rate of nucleotide deletions and insertions at the junctions during VDJ recombination. The higher rates of nucleotide insertion may correspond to an increase in terminal deoxyribonucleotidyl transferase expression, as previously noted (12, 13).

By combining high-throughput sequencing with informatic analysis, we were able to deconvolve the primary repertoire from the secondary repertoire without physically separating naïve B cells from activated or memory B cells. We discovered that the first step in somatic diversification of the antibody repertoire, VDJ recombination, has a deterministic component whose effects are evident not only in the early development of the immune system but also in the primary repertoire of mature adults. However, the stereotypy is different between young and old fish, suggesting that developmentally the B cell repertoire goes through a dramatic restructuring before the organism reaches sexual maturity. We observed that VDJ usage gradually saturates over time and the drop of VDJ-abundance correlation coefficients at 1 mo coincides with the disappearance of the most abundant VDJ families used at 2 wks. Finally, as the fish enter adulthood the mutation content of the repertoire qualitatively shifts from one in which mutated sequences are marginally expressed to one in which they are highly expressed. Collectively, these results demonstrate how many aspects of B cell repertoire development are balanced between deterministic and stochastic processes.

Materials and Methods

Zebrafish. Five wild-type WIK zebrafish families were raised at the same time in separate tanks, and 51 zebrafish were collected from these families at five developmental time points. Fish were sacrificed according to an animal protocol approved by the Stanford University administrative panel on laboratory animal care, and were snap-frozen in liquid nitrogen and stored at −80 °C.

mRNA Preparation, cDNA Synthesis, and PCR. mRNA preparation, cDNA synthesis, and PCR were performed as described previously (14). Briefly, total RNA from each fish was purified using TRIzol Plus RNA Purification System (Invitrogen). The mRNA was further purified using Oligotex mRNA Kit (Qiagen). Manufacturers' protocols were followed during these processes. cDNA was synthesized using SuperScript III reverse transcriptase (Invitrogen). PCRs were set up for each fish using 27 forward and two reverse primers. The PCR program began with an initial denaturation at 94 °C for 2 min, followed by 35 cycles of denaturation at 94 °C for 30 s, annealing of primer to DNA at 60 °C for 30 s, and extension by Platinum Taq DNA Polymerase High Fidelity (Invitrogen) at 68 °C for 2 min. PCR products were cleaned using QIAquick PCR Purification Kit (Qiagen) and followed by electrophoresis and gel extraction (Qiagen).

454 Library Preparation and Sequencing. The standard Roche 454 GS Titanium shotgun library protocol was followed. Multiplex Identifier (MID)–containing oligonucleotides were synthesized by Integrated DNA Technologies and were annealed to form 454 adaptor according to the Roche protocol. Libraries were quantified using a digital PCR method (20), which gave the absolute count of DNA molecules in the library. This allowed us to eliminate the manufacturer's suggested titration run.

Computational Methods. Methods of VDJ alignment, classification, and correlation analysis are detailed in *SI Appendix*.

1. Stein KE (1992) Thymus-independent and thymus-dependent responses to polysaccharide antigens. *J Infect Dis* 165(Suppl 1):S49–S52.
2. Bekker A, Chou C, Bernstein HH (2009) Update on universal annual influenza immunization recommendations for children. *Curr Opin Pediatr* 21:122–126.
3. Schroeder HW, Jr. (2006) Similarity and divergence in the development and expression of the mouse and human antibody repertoires. *Dev Comp Immunol* 30:119–135.
4. Yancopoulos GD, et al. (1984) Preferential utilization of the most JH-proximal VH gene segments in pre-B-cell lines. *Nature* 311:727–733.
5. Perlmutter RM, Kearney JF, Chang SP, Hood LE (1985) Developmentally controlled expression of immunoglobulin VH genes. *Science* 227:1597–1601.
6. Schroeder HW, Jr., et al. (1988) Physical linkage of a human immunoglobulin heavy chain variable region gene segment to diversity and joining region elements. *Proc Natl Acad Sci USA* 85:8196–8200.
7. Berman JE, et al. (1991) VH gene usage in humans: Biased usage of the VH6 gene in immature B lymphoid cells. *Eur J Immunol* 21:1311–1314.
8. Yancopoulos GD, Malynn BA, Alt FW (1988) Developmentally regulated and strain-specific expression of murine VH families. *J Exp Med* 168:417–435.
9. Malynn BA, Yancopoulos GD, Barth JE, Bona CA, Alt FW (1990) Biased expression of JH-proximal VH genes occurs in the newly generated repertoire of neonatal and adult mice. *J Exp Med* 171:843–859.
10. Jung D, Giallourakis C, Mostoslavsky R, Alt FW (2006) Mechanism and control of V(D)J recombination at the immunoglobulin heavy chain locus. *Annu Rev Immunol* 24:541–570.
11. Feeney AJ, Tang A, Ogwaro KM (2000) B-cell repertoire formation: Role of the recombination signal sequence in non-random V segment utilization. *Immunol Rev* 175:59–69.
12. Feeney AJ (1990) Lack of N regions in fetal and neonatal mouse immunoglobulin V-D-J junctional sequences. *J Exp Med* 172:1377–1390.
13. Schroeder HW, Jr., Zhang L, Philips JB, 3rd (2001) Slow, programmed maturation of the immunoglobulin HCDR3 repertoire during the third trimester of fetal life. *Blood* 98:2745–2751.
14. Weinstein JA, Jiang N, White RA, 3rd, Fisher DS, Quake SR (2009) High-throughput sequencing of the zebrafish antibody repertoire. *Science* 324:807–810.
15. Litman GW, Cannon JP, Dishaw LJ (2005) Reconstructing immune phylogeny: New perspectives. *Nat Rev Immunol* 5:866–879.
16. Danilova N, Bussmann J, Jekosch K, Steiner LA (2005) The immunoglobulin heavy-chain locus in zebrafish: Identification and expression of a previously unknown isotype, immunoglobulin Z. *Nat Immunol* 6:295–302.
17. Jhunjhunwala S, et al. (2008) The 3D structure of the immunoglobulin heavy-chain locus: Implications for long-range genomic interactions. *Cell* 133:265–279.
18. Merbl Y, Zucker-Toledano M, Quintana FJ, Cohen IR (2007) Newborn humans manifest autoantibodies to defined self molecules detected by antigen microarray informatics. *J Clin Invest* 117:712–718.
19. Wardemann H, et al. (2003) Predominant autoantibody production by early human B cell precursors. *Science* 301:1374–1377.
20. White RA, 3rd, Blainey PC, Fan HC, Quake SR (2009) Digital PCR provides sensitive and absolute calibration for high throughput sequencing. *BMC Genomics* 10:116.