High-resolution antibody dynamics of vaccine-induced immune responses

Uri Laserson[a,b,c,1], Francois Vigneault[b,d,1], Daniel Gadala-Maria[e], Gur Yaari[f], Mohamed Uduman[e], Jason A. Vander Heiden[e], William Kelton[g], Sang Taek Jung[g], Yi Liu[h], Jonathan Laserson[h], Raj Chari[b], Je-Hyuk Lee[b], Ido Bachelet[b], Brendan Hickey[i], Erez Lieberman-Aiden[j], Bozena Hanczaruk[k], Birgitte B. Simen[k], Michael Egholm[k], Daphne Koller[h], George Georgiou[g], Steven H. Kleinstein[e,f], and George M. Church[a,b,2]

[a]Harvard–MIT Division of Health Sciences and Technology, Cambridge, MA 02139; [b]Department of Genetics, Harvard Medical School, Boston, MA 02115; [c]Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139; [d]AbVitro, Boston, MA 02210; [e]Interdepartmental Program in Computational Biology and Bioinformatics, Yale University, New Haven CT 06520; [f]Department of Pathology, Yale University School of Medicine, New Haven, CT 06520; [g]Department of Chemical Engineering, University of Texas at Austin, Austin, TX 78712; [h]Department of Computer Science, Stanford University, Stanford, CA 94305; [i]Division of Rheumatology, Immunology, and Allergy, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115; [j]Harvard Society of Fellows, Harvard University, Cambridge, MA 02138; and [k]454 Life Sciences, A Roche Company, Branford, CT 06405

The adaptive immune system confers protection by generating a diverse repertoire of antibody receptors that are rapidly expanded and contracted in response to specific targets. Next-generation DNA sequencing now provides the opportunity to survey this complex and vast repertoire. In the present work, we describe a set of tools for the analysis of antibody repertoires and their application to elucidating the dynamics of the response to viral vaccination in human volunteers. By analyzing data from 38 separate blood samples across 2 y, we found that the use of the germ-line library of V and J segments is conserved between individuals over time. Surprisingly, there appeared to be no correlation between the use level of a particular VJ combination and degree of expansion. We found the antibody RNA repertoire in each volunteer to be highly dynamic, with each individual displaying qualitatively different response dynamics. By using combinatorial phage display, we screened selected VH genes paired with their corresponding VL library for affinity against the vaccine antigens. Altogether, this work presents an additional set of tools for profiling the human antibody repertoire and demonstrates characterization of the fast repertoire dynamics through time in multiple individuals responding to an immune challenge.

next-generation sequencing | influenza | immunology

The immune system is able to rapidly sense and respond to a vast array of invading organisms. Its arsenal contains systems that are immediately effective against commonly seen patterns (innate immunity) and systems that are capable of responding to novel invaders (adaptive immunity). Given the acute nature and diversity of infections, the immune system must be capable of rapid recognition of a pathogen, amplification of the response, and subsequent contraction of the response after the resolution of the infection. Adaptive immune responses rely on the continuous selection and amplification of specific clones from an enormous library of immune receptors (antibodies and T cell receptors). Specifically, stimulation of B-cell immunity results in the synthesis of antibodies that are secreted into the blood stream or into the mucosa as well as the programming of B memory cells that play a crucial role in the generation of rapid protective responses upon reinfection.

Currently, many immunology studies depend on characterizing lymphocyte subsets (e.g., assaying cell-surface receptors) and the ability to correlate them to encoded genetic information (1). Recent advances in next-generation sequencing (NGS) (2) have enabled any DNA-encodable assay to produce massive amounts of data. Indeed, NGS has enabled unprecedented views into the immune repertoire, as its immune receptor diversity is genetically encoded within a complex collection of lymphocytes (3–8).

The present study set out to dissect the rapid dynamics of the complete human peripheral antibody response against a controlled immune challenge (vaccination), without the a priori notion of cell state markers or functions. We vaccinated three individuals with approved subunit vaccines designed to confer protection to viral antigens and banked blood samples at multiple time points before and after the vaccinations. The dynamic behavior of the immune repertoire in response to the vaccination was analyzed by RT-PCR of heavy chain V (VH) genes followed by DNA sequencing by using the 454 platform. We found that the VH gene repertoire is highly dynamic, and the response in each vaccination is qualitatively different. In contrast, we found that each individual uses the germ-line–encoded library of antibody components in very similar ways.

Because the immune system is shaped by selective pressures at the population and somatic levels (9), we observed statistical evidence that the immune system is equally likely to expend energy on expanding nondiversifying VJ combinations (innate action) as it is on expanding diversifying and class-switching VJ combinations (adaptive action). Finally, we synthesized a collection of the strongest-responding VH genes 1 wk after vaccination, paired them to their corresponding VL gene libraries by combinatorial phage display, and tested them for affinity against the vaccine antigens.

Results

Vaccination Time-Course Design. We characterized the antibody repertoire of three Personal Genome Project volunteers (three authors, G.M.C., I.B., and F.V.) in response to vaccination. In 2008, G.M.C. was vaccinated against seasonal influenza,

Table 1. Overview of vaccination experiment

Category	Sequencing data		
	GMC (%)	IB (%)	FV (%)
Raw reads	3,752,117	1,220,302	883,079
Filtered reads*	2,261,155 (60)	1,008,912 (83)	703,192 (80)
With isotype	1,462,059	872,110	590,291
IgM	658,154 (45)	428,070 (49)	174,874 (30)
IgG	273,057 (19)	168,321 (19)	174,692 (30)
IgA	415,102 (28)	156,908 (18)	206,089 (35)
IgD	115,668 (8)	119,587 (14)	34,553 (6)
IgE	78 (0)	224 (0)	83 (0)
No. of clones	725,202	526,838	174,593
With ≥3 reads	91,672	41,459	20,881
With ≥ 10000 reads	13	1	4
Seen in ≥2 time points	58,941	12,569	11,850
Seen in all time points	98	87	72

*Size-selected, VJ-filtered.

hepatitis A, and hepatitis B; in 2009, G.M.C., I.B., and F.V. all received the seasonal flu vaccine. Blood samples were collected at different times before and after the vaccination, as specified in Table 1 and *SI Appendix*, Fig. S1. One of the goals of the study was to track the response of the immune system exclusively through the analysis of the variable gene repertoire. Although we expect the distribution of immune cell types to remain fairly constant [immature, ~5%; naïve, ~60%; memory, ~30%; and plasmablast, 1.8% (10, 11)], we expect that the amount of vaccine-specific B cells will fluctuate in a detectable way (i.e., signal increases while noise remains constant). Total RNA was extracted from unsorted peripheral blood mononucleated cells and VH genes were amplified by RT-PCR by using primers that span the leader exon–exon junction on the V-segment side and sit inside the constant region on the J side. Therefore, our operational definition of "antibody repertoire" is derived purely from the expression levels of the various antibody mRNAs in unsorted peripheral blood mononucleated cells. This primer set allowed us to sequence the entire VH gene sequence and to determine the antibody isotype in the majority of cases. Each sample was uniquely barcoded by ligation following the PCR and subjected to Roche 454 sequencing and analysis.

Reproducibility and Quantitation. We obtained ~4.3 million reads that successfully aligned to the international immunogenetics information system (IMGT) germ-line reference database (Table 1 and *SI Appendix*, Fig. S1). We sequenced one VH gene library twice (generating sequencing replicates SR1 and SR2) and also sequenced an independent VH library from the same RNA sample (technical replicate TR1). Between these three sequencing runs, 477,118 unique VH genes (i.e., unique antibody sequences) were identified, of which only 3% were shared between the three runs and 14% were observed in at least two runs (*SI Appendix*, Fig. S2A). However, those shared clones accounted for 59% and 71% of all reads, suggesting that the highly expressed clones are actually sampled significantly between replicate runs. This was further validated by a strong correlation between technical replicate samples, confirming technical reproducibility of our approach (*SI Appendix*, Fig. S2B). Furthermore, resampling our data showed that 10^4 to 10^5 reads are sufficient to properly characterize a sample and obtain high correlations between replicates (*SI Appendix*, Fig. S2C).

Characteristics of the VH Repertoire. Overall use of the individual V and J components was highly nonuniform within a given individual (*SI Appendix*, Fig. S3). The most frequently observed V segments were IGHV3-23 (11% of all reads), IGHV3-30 (8%), IGHV4-59 (7%), and IGHV1-69 (6%), and the most frequent J segments were IGHJ4 (41%) and IGHJ6 (31%), consistent with previous studies (7, 12). Use of the germ-line–encoded Variable, Diverse, and Joining (VDJ) gene library was quite similar across
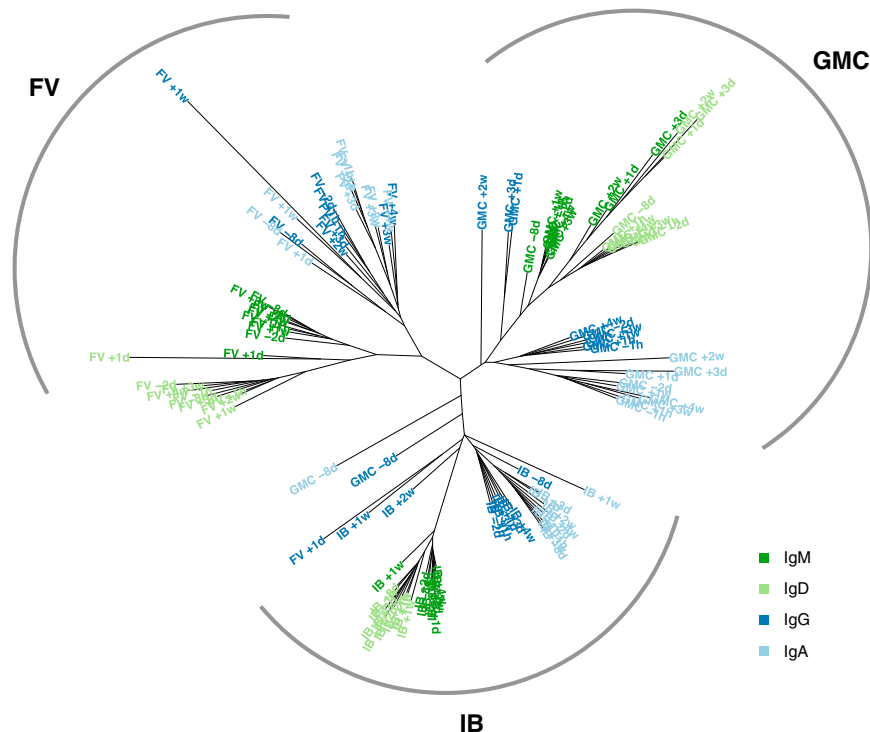

Fig. 1. Neighbor-joining tree of V-use vectors. V-use vectors are calculated for each individual–isotype combination, and clustered by using the neighbor-joining algorithm. Each isotype is colored according to the legend. The tree naturally clusters by individual and then by isotype.
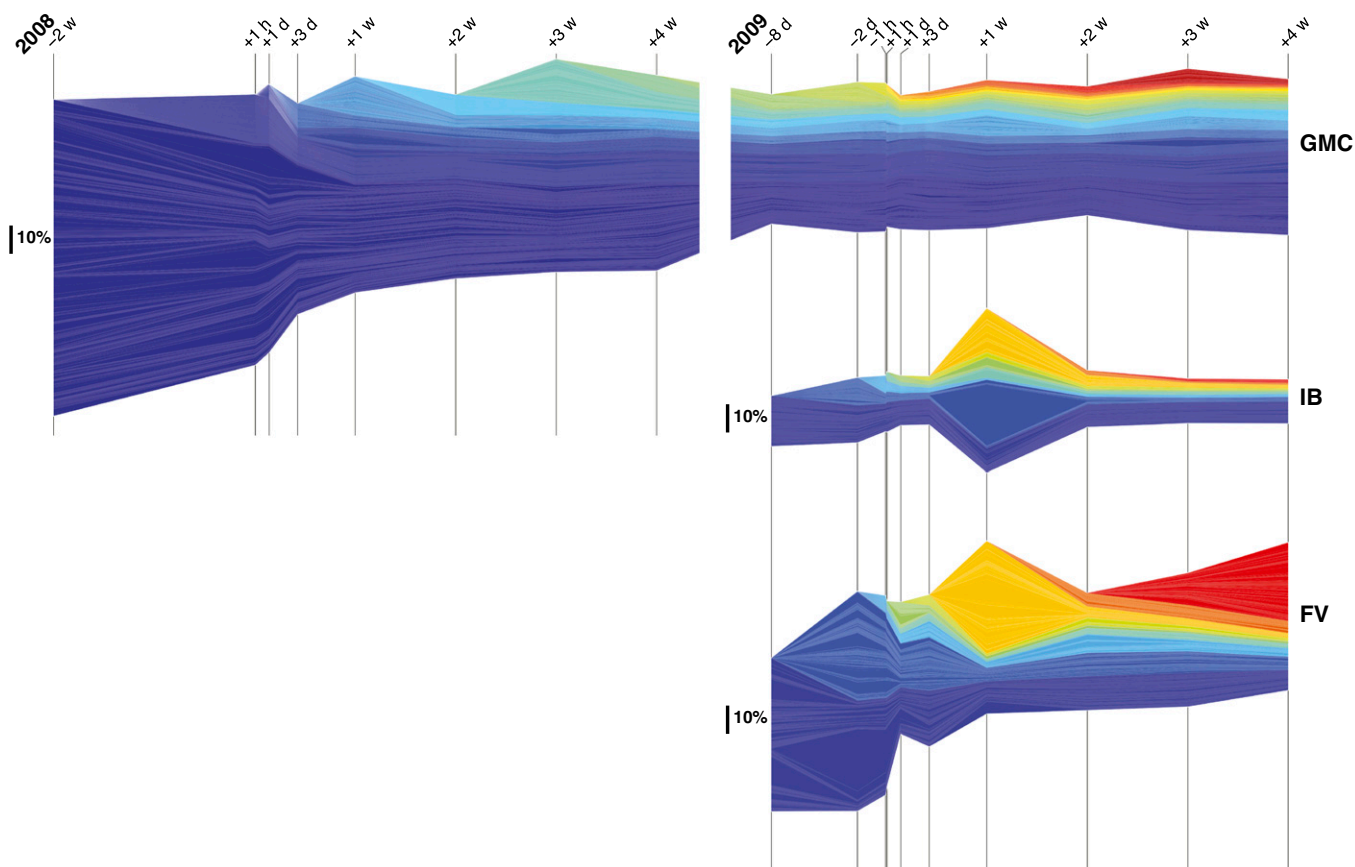
Fig. 2. Probability of clone activation by VJ use. (*A*) Each clone is labeled according to its VJ use and whether it is naïve (IgM-only with low mutation) or activated (IgG or IgA with high mutation). The probability of observing naïve or activated clones is estimated assuming a binomial distribution. The line plots the estimated probability of activation, and bars represent ±1 SD. VJ combinations are ordered according to G.M.C. (*B*) For each VJ combination, we plot its VJ-use frequency rank against its rank of probability of activation. This is filtered on VJ combinations for which we obtain at least 100 clones that are classifiable as naïve or activated.

individuals (e.g., G.M.C. vs. F.V.) and also similar over time within an individual (e.g., G.M.C. in 2008 vs. 2009). For each time point, we compute a vector of use for the possible combinations of the VJ gene segments. Indeed, the Spearman correlation between VJ-use vectors was consistently high across time points and individuals (*SI Appendix*, Fig. S4) and VJ-use

Fig. 3. Vaccination clone dynamics colored by mutation. (*A*) Each layer represents a clone. Time is shown by the grid lines on the *x* axis, and labeled relative to the two vaccination events. The thickness of each layer is proportional to the frequency of that clone at that time point. Each clone is colored based on the average mutation level of the corresponding reads (see color bar for *B*). Only clones seen in at least two time points are shown here. (*B*) Histogram of the average mutation level of all of the clones. Each clone is counted once (i.e., clones are not weighted by the number of corresponding reads). Stream graphs are stacked bar charts with moving baselines; if we did not filter out clones in only a single time point, the total thickness of the graph would add up to unity.

Fig. 4. Vaccination clone dynamics colored by onset time. Same as Fig. 3, except clones are colored by onset time. Onset times are ordered spectrally, so that all clones seen in the first time points are blue, followed by cyan, etc.

time series are remarkably stable as well (*SI Appendix*, Fig. S5). Finally, we built a neighbor-joining tree by using V-use vectors (as opposed to VJ-use) of each of the samples to see how V use is structured. For the most part, V use clustered first by individual, and then by isotype, implying that, even though V use is grossly similar across individuals, each individual still has a unique signature (Fig. 1).

For a majority of the reads, we were able to genetically discern the antibody isotype. We found that IgM antibody transcripts were the most abundant (43% of all reads), followed by IgA (27%), IgG (21%), IgD (9%), and IgE (0.01%; Table 1 and *SI Appendix*, Fig. S1). The level of IgA RNA is higher than the amount of IgA protein typically observed in serum. We believe this is observed because serum IgA are cleared more rapidly than IgG and also because the blood carries IgA cells in transit to mucosal sites. However, the isotype use varied significantly between time points (*SI Appendix*, Fig. S6).

Mutation levels were also measured across each of the reads. As expected, mutation rates were higher in the complementarity determining regions (CDRs) of the antibodies, and were much higher in IgG and IgA antibodies (*SI Appendix*, Fig. S7). We further processed our reads through the Bayesian estimation of antigen-driven selection (BASELINe) pipeline that estimates selection pressure on the antibodies (13). Framework regions (FWR) were universally negatively selected, whereas CDR regions showed neutral to slightly negative selection on average; however, CDR selection values were always more positive than FWR selection values (*SI Appendix*, Fig. S8).

The CDR3 length distribution we observed was consistent with IMGT/LIGM data (14) (*SI Appendix*, Fig. S9; CDR3 extraction is described in *Methods*). The 5th and 95th percentiles of the observed CDR3 lengths are 36 nt and 75 nt, with median length 54 nt (with longest observed CDR3 at 140 nt).

B cells and their associated antibody transcripts can be present at vastly different quantities, depending on which lymphocyte subset they are derived from and the degree of expansion of particular clones. In particular, our method does not allow for functional differentiation between naïve, memory, or antibody-secreting cells (ASCs). Therefore, the variation in antibody expression may influence our view of the immune repertoire (e.g., ASCs, although rare in the periphery, exhibit very high expression of antibodies). Because the VDJ recombination process introduces so much diversity, the CDR3 sequence effectively functions as a natural barcode for a particular clone (15). To functionally define antibody clones, we perform clustering of the CDR3 sequences and define two reads as derived from the same clone if their CDR3 sequences are highly similar (same VJ combination and CDR3 within four nucleotide changes), as it is unlikely that two independent B cells will generate the same nucleotide sequence. In total, we observe >1.4 million clones across all of our data; however, only ~150,000 clones had more than three reads each and only 24 clones with >10,000 reads each. Summed over all three subjects, ~84,000 clones were seen in at least two separate time points, whereas only 257 heavy-chain clones were seen in every time point (*SI Appendix*, Fig. S1).

Each clone is labeled as "activated" or "naïve" based on its isotype and mutation level. For each VJ combination, we estimated the probability that a clone is activated (assuming a binomial distribution) and found that it is reproducibly biased by VJ use (Fig. 2*A*). The V regions most likely to become activated are dominated by IGHV4- and IGHV5-family genes, and the three individuals have highly correlated biases in the VJ-activation probabilities (Spearman correlation of ~0.7). Nevertheless,
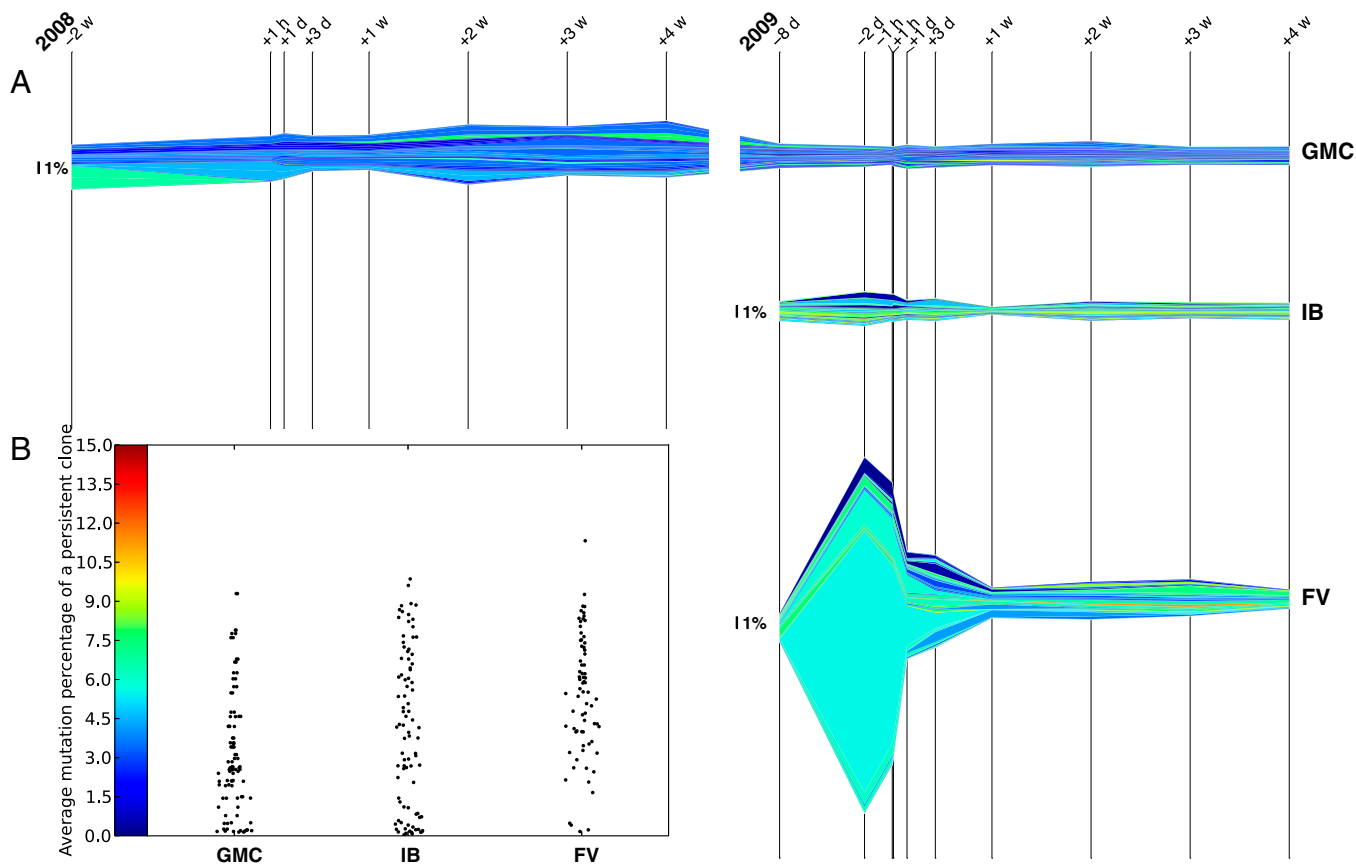
IMMUNOLOGY

Fig. 5. Dynamics of persistent clones. (*A*) Stream graphs of only clones that are observed in every single time point for a given individual. They are colored as in Fig. 3. (*B*) Distribution of average mutation level of the persistent clones.

we find that there is virtually no correlation between whether a gene derived from a particular VJ germ-line combination is likely to become activated (i.e., subjected to somatic hypermutation and clonally expanded) and whether it is highly used (Fig. 2*B*). One hypothesis consistent with these results is that the immune system is being selected for naïve antibodies that confer protection. If, instead, the sole function of antibodies is in adaptive immunity, it seems wasteful to highly express VJ combinations that are less likely to become activated. However, we observed no correlation between the level of VJ combination use and activation. We interpret this as consistent with the hypothesis that the antibody repertoire is shaped by selective forces at population and somatic time scales. More precisely, use of the VDJ germ-line library may be optimized for naïve interactions with common pathogens at the population scale, and the propensity of any given germ-line gene to somatically mutate may be optimized for the evolvability of the target organisms. In the context of our hypothesis, the population of antibodies can be seen to occupy an innate-adaptive spectrum (9, 16).

Antibody Repertoire Dynamics. Each of the three volunteers was given the clinically-indicated seasonal flu vaccine, and G.M.C. was also given boosters to hepatitis A/B in 2008. None of the subjects were naïve to the antigens at the time of vaccination (through prior vaccination or infection). Each read was assigned to a clone and a time point, allowing us to compute time series. The clone frequencies were tracked across all 38 time points to produce >20 million clone-frequency measurements (although the matrix is sparse). In contrast to the relative stability of the VJ use, antibody clones were highly dynamic and variable across individuals (Figs. 3 and 4).

We used expression-level/burstiness as a proxy for immune response strength. Responses to each of the four vaccination events were qualitatively different: IB produced a "textbook" response with large proliferating clones 7 d after vaccination; FV displayed high-frequency clones just before vaccination as well as large clonal expansions 7 d postvaccination; the repertoire analysis suggested that G.M.C. showed weak response from his first vaccination, and his second vaccination appeared to have produced no high-frequency responses. Taken together, these data show that individuals can demonstrate a highly varying array of responses against an identical immune challenge, likely influenced by prior exposure, age, and other concurrent immune responses during the course of this experiment, among other factors.

We verified that samples that are closer in time share more unique clones. We computed the number of shared CDR3 sequences between all 703 possible pairs of samples across all 38 time points, and observed that closer time points within an individual indeed share a larger number of unique CDR3s (*SI Appendix*, Fig. S10). Consistent with this, interindividual comparisons between time points showed very little CDR3 overlap.

We also quantified the range of dynamic behavior of the clones, finding that clones generally fluctuate wildly (*SI Appendix*, Fig. S11). Interestingly, each individual had a number of clones that were present at every time point sampled, including the samples separated by more than 1 y (257 clones total with median clone frequency 1.1×10^{-4}; Fig. 5*A*). It is possible that these clones correspond to a long-lived, expanded B-cell memory population or are chronically responding to antigens (foreign or auto-) that are always present; indeed, these clones include sequences that are highly mutated (Fig. 5*B*).

Clone Analysis. It is commonly accepted that expanding clone populations should arise from an immune challenge approximately 7 d after flu vaccination (17). More precisely, one would expect that prevaccination samples are dominated by naïve and memory cells with a large diversity of specificities, whereas flu-specific ASCs would be observed 7 d after vaccination. As our assay does not distinguish between these various cell types, we anticipated that the day 7 emergence of flu-specific ASCs would provide a strong enough signal above the noise of irrelevant memory and naïve cells. Therefore, we picked a subset of the largest clones from multiple time points before and after vaccination (−2 d, +7 d, +21 d), and synthesized, expressed, and panned them by phage display. Although it is expected that Ig expression in *Escherichia coli* using phage display is inefficient, we were surprised to find very few strong binders against the vaccine hemagglutinin antigens. This might also reflect our combinatorial pairing strategy, which may not yield a natural pairing of heavy and light chains. Alternatively, a yeast display approach may have had a better chance for success for expressing human-derived antibody chains, as previously demonstrated (18).

Interestingly, even though G.M.C. showed no significant response in 2009, the strongest binder (GMC J-065) was found in his day +7 response of that year. We then applied the Immuni-Tree algorithm (19) on clone GMC J-065 to infer the most likely evolutionary pathway (19). The tree was also overlaid with selection values estimated by using the BASELINe algorithm (13) as well as mutation levels ([SI Appendix](), Fig. S12). As expected, most nodes in the tree displayed significant negative selection in the FWRs, whereas some of the nodes show significant positive selection in the CDRs. We are currently in the process of analyzing clones of these trees that are more evolved and show signs of greater selection pressure.

Discussion

In this study, we generated a high-throughput profile of the short–time-scale dynamics of the antibody heavy chain repertoire. For proper function, the antibody repertoire has the ability to rapidly expand and contract in a highly dynamic manner, consistent with our observations. We also found evidence that the antibody repertoire functions on an innate-adaptive spectrum, on which use of the germ-line antibody VDJ library is simultaneously shaped by population selection and somatic selection pressures. Indeed, it is apparent that use of the germ-line library is strongly stereotyped between individuals, but particular clones are highly dynamic.

Although we were able to glean significant insights into the immune system from variable gene sequencing alone, it appears that using the information for predictive purposes still requires a significantly greater amount of data (20). This is analogous to the dichotomy between supervised and unsupervised learning in statistics: our (high-throughput) genetics-only data acquisition in contrast with (low-throughput) functional labeling. We hope that such an approach will eventually enable the analysis of immune function and also mining the "fossil record" (21) of individual antigen exposures.

Although we have thus far not been able to realize this vision, we believe this study represents a necessary milestone in a collective effort for the development of new tools to harness the full potential of the immune system. To that extent, we are focusing on developing methodologies for high-throughput capture of paired heavy and light chain sequences from single cells (22). Coupled with significant advances in DNA synthesis technology (23, 24), we should soon be able to assay a large immune repertoire against a large, synthetic library of antigens (e.g., auto-antigens, allergens, infectious agents) (25–28). Doing so will further the development of immune repertoire profiling and facilitate our progress toward the next generation of diagnostics, vaccines, and personalized therapeutic discovery.

Materials and Methods

Experimental methods are detailed in *SI Appendix, SI Materials and Methods*. It includes detailed description of the methods such as: sample collection, primer design, and sequencing library preparation. It also includes detail of data processing such as: data processing overview, VDJ alignment process, sequence clustering, mutation analysis pipeline, analysis of selection pressures, clone phylogeny inference, V-usage clustering, clone synthesis/affinity, and software tools. Supplementary figures and legends are also detailed.

1. Maecker HTH, et al. (2012) New tools for classification and monitoring of autoimmune diseases. *Nat Rev Rheumatol* 8(6):317–328.
2. Shendure J, Lieberman Aiden E (2012) The expanding scope of DNA sequencing. *Nat Biotechnol* 30(11):1084–1094.
3. Weinstein JAJ, Jiang NN, White RAR, 3rd, Fisher DSD, Quake SRS (2009) High-throughput sequencing of the zebrafish antibody repertoire. *Science* 324(5928): 807–810.
4. Boyd SD, et al. (2009) Measurement and clinical monitoring of human lymphocyte clonality by massively parallel VDJ pyrosequencing. *Sci Transl Med* 1(12):12ra23.
5. Warren RL, et al. (2011) Exhaustive T-cell repertoire sequencing of human peripheral blood samples reveals signatures of antigen selection and a directly measured repertoire size of at least 1 million clonotypes. *Genome Res* 21(5):790–797.
6. Logan ACA, et al. (2011) High-throughput VDJ sequencing for quantification of minimal residual disease in chronic lymphocytic leukemia and immune reconstitution assessment. *Proc Natl Acad Sci USA* 108(52):21194–21199.
7. Wu Y-C, et al. (2010) High-throughput immunoglobulin repertoire analysis distinguishes between human IgM memory and switched memory B-cell populations. *Blood* 116(7):1070–1078.
8. Vollmers C, Sit RV, Weinstein JA, Dekker CL, Quake SR (2013) Genetic measurement of memory B-cell recall using antibody repertoire sequencing. *Proc Natl Acad Sci USA* 110(33):13463–13468.
9. Cohen IR (2000) *Tending Adam's Garden: Evolving the Cognitive Immune Self* (Academic, London).
10. Perez-Andres M, et al.; Primary Health Care Group of Salamanca for the Study of MBL (2010) Human peripheral blood B-cell compartments: A crossroad in B-cell traffic. *Cytometry B Clin Cytom* 78(Suppl 1):S47–S60.
11. Caraux A, et al.; Myeloma Stem Cell Network (2010) Circulating human B and plasma cells. Age-associated changes in counts and detailed characterization of circulating normal CD138- and CD138+ plasma cells. *Haematologica* 95(6):1016–1020.
12. Brezinschek HP, Brezinschek RI, Lipsky PE (1995) Analysis of the heavy chain repertoire of human peripheral B cells using single-cell polymerase chain reaction. *J Immunol* 155(1):190–202.
13. Yaari G, Uduman M, Kleinstein SH (2012) Quantifying selection in high-throughput immunoglobulin sequencing data sets. *Nucleic Acids Res* 40(17):e134.
14. Giudicelli V, et al. (2006) IMGT/LIGM-DB, the IMGT comprehensive database of immunoglobulin and T cell receptor nucleotide sequences. *Nucleic Acids Res* 34(database issue):D781–D784.
15. Davis MMM, Bjorkman PJP (1988) T-cell antigen receptor genes and T-cell recognition. *Nature* 334(6181):395–402.
16. Ochsenbein AF, et al. (1999) Control of early viral and bacterial distribution and disease by natural antibodies. *Science* 286(5447):2156–2159.
17. Wrammert J, et al. (2008) Rapid cloning of high-affinity human monoclonal antibodies against influenza virus. *Nature* 453(7195):667–671.
18. Bowley DR, Labrijn AF, Zwick MB, Burton DR (2007) Antigen selection from an HIV-1 immune antibody library displayed on yeast yields many novel antibodies compared to selection from the same library displayed on phage. *Protein Eng Des Sel* 20(2): 81–90.
19. Laserson J (2012) Bayesian assembly of reads from high throughput sequencing. PhD Thesis (Stanford Univ, Stanford, CA).
20. Le QV, et al. (2011) Building high-level features using large scale unsupervised learning. arXiv:1112.6209v5.
21. Lerner RA, Barbas CF, 3rd, Kang AS, Burton DR (1991) On the use of combinatorial antibody libraries to clone the "fossil record" of an individual's immune response. *Proc Natl Acad Sci USA* 88(21):9705–9706.
22. DeKosky BJ, et al. (2013) High-throughput sequencing of the paired human immunoglobulin heavy and light chain repertoire. *Nat Biotechnol* 31(2):166–169.
23. Carr PA, Church GM (2009) Genome engineering. *Nat Biotechnol* 27(12):1151–1162.
24. Kosuri S, et al. (2010) Scalable gene synthesis by selective amplification of DNA pools from high-fidelity microchips. *Nat Biotechnol* 28(12):1295–1299.
25. Larman HB, et al. (2011) Autoantigen discovery with a synthetic human peptidome. *Nat Biotechnol* 29(6):535–541.
26. Bowley DR, Jones TM, Burton DR, Lerner RA (2009) Libraries against libraries for combinatorial selection of replicating antigen-antibody pairs. *Proc Natl Acad Sci USA* 106(5):1380–1385.
27. Nirantar SR, Ghadessy FJ (2011) Compartmentalized linkage of genes encoding interacting protein pairs. *Proteomics* 11(7):1335–1339.
28. Zhu J, et al. (2013) Protein interaction discovery using parallel analysis of translated ORFs (PLATO). *Nat Biotechnol* 31(4):331–334.

IMMUNOLOGY