bioRxiv preprint doi: https://doi.org/10.1101/558825; this version posted February 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Repertoire-wide phylogenetic models of B cell molecular evolution reveal evolutionary signatures of aging and vaccination

Kenneth B Hoehn[1], Jason A. Vander Heiden[2], Julian Q. Zhou[3], Gerton Lunter[4], Oliver G. Pybus[5*+], Steven H. Kleinstein[1,3*+]

[1]Department of Pathology, Yale School of Medicine, New Haven, CT, USA

[2] Department of Bioinformatics & Computational Biology, Genentech, South San Francisco, CA, USA

[3]Interdepartmental Program in Computational Biology and Bioinformatics, Yale University, New Haven, CT, USA

[4]Wellcome Centre for Human Genetics, Roosevelt Drive, Oxford OX3 7BN, UK

[5]Department of Zoology, University of Oxford, OX1 3PS, United Kingdom

* Authors contributed equally

[+] To whom correspondence should be addressed

bioRxiv preprint doi: https://doi.org/10.1101/558825; this version posted February 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Abstract

In order to produce effective antibodies, B cells undergo rapid somatic hypermutation (SHM) and selection for binding affinity to antigen via a process called affinity maturation. The similarities between this process and evolution by natural selection have led many groups to use evolutionary and phylogenetic methods to characterize the development of immunological memory, vaccination, and other processes that depend on affinity maturation. However, these applications are limited by several features of affinity maturation that violate assumptions in standard phylogenetic models. Further, most phylogenetic models are designed to be applied to individual lineages comprising genetically diverse sequences, while B cell repertoires often consist of hundreds to thousands of separate low-diversity lineages. Here, we introduce a hierarchical phylogenetic framework that incorporates the unique features of SHM, and integrates information from all lineages in a repertoire to more precisely estimate model parameters. We demonstrate the power of this approach by characterizing previously un-described phenomena in affinity maturation. First, we find evidence consistent with age related changes in SHM hot- and cold-spot motifs. Second, we identify a consistent relationship between increased tree length and signs of increased negative selection, apparent in the repertoires of both healthy subjects and those undergoing active immune responses. This suggests that B cell lineages shift towards negative selection over time as a general feature of affinity maturation. Our study provides a framework for undertaking repertoire-wide phylogenetic testing of SHM hypotheses, and provides a new means of charactering the process of mutation and selection during affinity maturation.

bioRxiv preprint doi: https://doi.org/10.1101/558825; this version posted February 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Introduction

B cell receptors (BCRs) are membrane-bound immunoglobulins (Ig) expressed on the surfaces of B cells that bind to antigen and may be released as antibodies to fight infection. BCRs are generated through the shuffling of Ig gene segments by V(D)J recombination and, if the cell expressing them is activated, by a second process of BCR modification called affinity maturation (Murphy *et al.* 2012). Affinity maturation consists of repeated rounds of somatic hypermutation (SHM) of the BCR, cell proliferation, and selection for antigen binding affinity (Murphy *et al.* 2012). These processes give rise to clonal lineages of B cells that each descend from a progenitor cell, from which they differ predominately by point mutations. The BCR sequence, and the nature of the mutations introduced during affinity maturation, can be investigated in detail using high-throughput next generation BCR sequencing (Boyd *et al.* 2009; Greiff *et al.* 2015b; Yaari and Kleinstein 2015).

The fact that affinity maturation is analogous to the process of evolution by natural selection suggests that methods from molecular evolutionary biology, particularly phylogenetics, could have broad utility in studying affinity maturation. This has stimulated the development of methods of evolutionary sequence analysis designed specifically for BCR sequences, particularly phylogenetic approaches (Barak *et al.* 2008; Kepler 2013; Hoehn *et al.* 2017; DeWitt *et al.* 2018). These methods have shown promise in elucidating information about the adaptive immune response in humans, such as the sequence of mutations that occur during antibody co-evolution with HIV (Liao *et al.* 2013) and the migration of B cells in multiple sclerosis (Stern *et al.* 2014).

bioRxiv preprint doi: https://doi.org/10.1101/558825; this version posted February 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

There are a number of challenges in adapting phylogenetic techniques to B cell clonal lineage analysis. For example, the biology of affinity maturation violates fundamental assumptions in most phylogenetic substitution models, such as independent change at each nucleotide site and time-reversibility of substitution rates (Hoehn *et al.* 2016). To address these issues we previously introduced the HLP17 substitution model, which weights codon substitutions by the presence of SHM hot/cold-spot motifs, and does not assume that substitution rates are time reversible (Hoehn *et al.* 2017).

Phylogenetic techniques are typically used to analyze individual, genetically-diverse B cell lineages (Felsenstein 1981; Liao *et al.* 2013). However, B cell repertoires are typically profiled by next-generation sequencing and consist of many – potentially thousands of – expanded clonal lineages, each of which may contain only a few unique sequences (Weinstein *et al.* 2009; Jiang *et al.* 2013; Greiff *et al.* 2015a). In light of this, many analyses have used non-phylogenetic summary statistics to characterize BCR repertoires, such as the distribution of sequences per clone (e.g. Bashford-Rogers *et al.* 2013; Greiff *et al.* 2015a; Hoehn *et al.* 2015). Others have used non-phylogenetic approaches to study somatic hypermutation biases (Yaari *et al.* 2013) and signatures of clonal selection (Yaari *et al.* 2012), which represent B cell clonal lineages by selecting or constructing a single representative sequence. SHM biases have also recently been incorporated into a phylogenetic framework under Feng *et al.* (2017). By explicitly modeling shared ancestry among sequences within the same clone, phylogenetic approaches potentially offer a more powerful means of understanding somatic hypermutation and affinity maturation by using the full set of substitutions expected to have occurred in a repertoire. However, standard

bioRxiv preprint doi: https://doi.org/10.1101/558825; this version posted February 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

phylogenetic approaches are limited to single lineages, and give imprecise parameter estimates, except when applied to unusually long-lived or highly-diverse B cell lineages (e.g. Hoehn *et al.* 2017).

We propose here that it is possible to combine some of the benefits of phylogenetic and summary statistic approaches of B cell repertoire analysis, by using hierarchical phylogenetic models. These approaches contain multiple levels of parameters, some of which are shared among lineages, while others are estimated for each lineage individually (Suchard *et al.* 2003). For example, Rodrigo et al. (2003) applied one such hierarchical phylogenetic approach to a set of HIV sequences from infected patients in order to jointly estimate both the virus substitution rate and the proportion of individuals that did not respond to antiretroviral therapy. However, previous applications of hierarchical phylogenetic models to virus genomes do not address the abovementioned model assumptions that are violated by the biology of B cell affinity maturation.

A hierarchical approach that is specifically tailored to B cell sequence evolution has the potential to dramatically improve accuracy of parameter estimation. By assuming that B cell lineages within a particular repertoire experience broadly similar patterns of substitution (e.g., hot- and cold-spot sequence motifs that experience altered mutation rates under SHM), a hierarchical approach is able to share information across B cell lineages and thereby take advantage of the large genetic diversity of the entire repertoire, despite the fact that each individual lineage within the BCR repertoire data may exhibit low diversity. Here, we develop a hierarchical phylogenetic framework capable of characterizing entire B cell repertoires by jointly estimating parameters and lineage tree topologies for all lineages within a repertoire. We first introduce the HLP19

bioRxiv preprint doi: https://doi.org/10.1101/558825; this version posted February 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

model, an improved form of the HLP17 model that accounts for changing codon frequencies during affinity maturation, and validate our hierarchical approach through simulation. We then apply this new framework to characterize the effects of aging on B cell repertoire development and B cell responses to influenza vaccination. We demonstrate that hierarchical approaches can quantify variation in SHM features both across individuals and within the same individual through time. Our results reveal previously uncharacterized immunological phenomena underlying aging and vaccination. We discover (i) evidence of changes in SHM hot/cold-spot mutation biases associated with age, (ii) evidence of negative selection acting on complementarity-determining regions (CDRs) associated with the human immune response to influenza vaccination, and (iii) a consistent relationship between increased lineage tree length and signatures of negative selection across our datasets.

bioRxiv preprint doi: https://doi.org/10.1101/558825; this version posted February 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Methods

A nonstationary, nonreversible phylogenetic substitution model for B cell evolution

The process of nucleotide change along a given phylogenetic tree is modeled as a Markov process, such that the rate of transitioning into any state at each instant in time is dependent only on the current state of the model (Felsenstein 1981). Here, we characterize codon change in Ig sequences using a 61x61 element matrix (**Q** matrix) that describes the instantaneous rates of change between all non-stop codons. In the HLP17 model (Hoehn *et al.* 2017), these instantaneous rates are parameterized by the nonsynonymous/synonymous mutation rate ratio (ω), transition/transversion mutation rate ratio (κ), codon frequencies ($\boldsymbol{\pi}$), and modified substitution rates h^a where a is an SHM hot- or cold-spot motif, such as WRC (Peled *et al.* 2008); W=A/T, R=A/G; only the underlined base experiences increased substitution).

The HLP17 model makes a salient approximation (as do almost all other substitution models) that nucleotide or codon frequencies are constant over time at a stationary distribution. However, the codon composition of B cell sequences begins substantially far away from equilibrium and changes over time (Sheng *et al.* 2016), making this assumption inappropriate. Hoehn *et al.* (2017) attempted to address this problem by using maximum likelihood (ML) to estimate codon frequencies. Whilst this approach may be better than empirical estimates of codon frequencies, at least in some instances, it more than doubles the number of model parameters. Here, we introduce HLP19, a modified HLP17 substitution model which, among other minor adjustments (see **Supplemental File 1**), uses the predicted codon frequencies at the midpoint of phylogeny in question. HLP19 is fully detailed in **Supplemental File 1**. Overall, HLP19 has less than half the

bioRxiv preprint doi: https://doi.org/10.1101/558825; this version posted February 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

number of free parameters as HLP17 and exhibits improved branch length estimates, generally better estimates of certain substitution model parameters such as ω (**Supplemental File 3**), and is structurally more similar to other non-reversible substitution models (Yang 1994; Kaehler *et al.* 2017). Further, certain approximations in HLP19 (detailed in **Supplemental File 1**) give significantly improved runtime compared to HLP17.

Hierarchical phylogenetic models for B cell repertoires

Under standard ML phylogenetics (Felsenstein 1981), a single multiple sequence alignment $\boldsymbol{X}$ is specified, and the goal is to find the tree topology and branch lengths $\boldsymbol{T}$, and the set of substitution parameters, that maximize the likelihood of $\boldsymbol{X}$. For B cell lineage phylogenies, the sequence alignment is supplemented with a predicted germline sequence $\boldsymbol{G}$ that acts as an outgroup and adds direction to the tree. In this study we extend this approach by calculating the likelihood of the entire B cell repertoire, which we define as the product of the tree likelihoods for each of n lineages, using each lineage i's tree topology (T_i), substitution parameters (ω_i, κ_i, $\boldsymbol{h}_i$), sequence data (X_i) and predicted germline sequence (G_i) (equation 1). This approach therefore assumes that mutations in each lineage are independent from each other:

$$L(\boldsymbol{T}, \boldsymbol{\omega}, \boldsymbol{\kappa}, \boldsymbol{h} | \boldsymbol{X}, \boldsymbol{G}) = \prod_{i=1}^{n} L(T_i, \omega_i, \kappa_i, \boldsymbol{h_i} | \boldsymbol{X}_i, G_i) \quad (1)$$

The goal of our phylogenetic repertoire analysis is to find the tree topologies, branch lengths, and substitution parameters that maximize the product on the left-hand side of equation 1, whereas the goal of typical ML phylogenetic analysis is to maximize individually each phylogenetic

bioRxiv preprint doi: https://doi.org/10.1101/558825; this version posted February 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

likelihood on the right-hand side. In hierarchical models, parameters may be constrained to be identical across lineages, allowing them to be estimated at the repertoire level. For instance, we may estimate a repertoire-wide transition/transversion rate ratio by constraining $\kappa_1 = \kappa_2 = \ldots \kappa_n$. Constraining parameters in this way will lower the overall likelihood of the repertoire compared to optimizing parameters for each lineage individually (because there will be fewer degrees of freedom) but will decrease the number of parameters and thereby reduce parameter estimation variance. For the analyses presented here, we constrain all substitution parameters to be identical across lineages within a repertoire.

B cell repertoire datasets

We use hierarchical phylogenetic models to characterize B cell repertoires in two previously published data sets obtained from peripheral blood samples. The first dataset (Age) consists of samples taken from 27 healthy individuals in two consecutive years (Wang *et al.* 2013). Subjects varied in age from 20 to 81 years old and both male and female subjects were included. Our second dataset (Vaccine) consists of samples from three male donors at 10 time points: -8 days, -2 days, -1 hour, +1 hour, +1 day, +3 days, +7 days, +14 days, and +28 days relative to seasonal influenza vaccination (Laserson *et al.* 2014). Quality control and data processing for both of these datasets is detailed in **Supplemental File 1**. Samples in the Vaccine dataset had between 141 and 15,763 (mean 6,272.7) unique sequences in non-singleton clones (i.e. clones containing >1 unique sequence). Because of this large variation, the repertoires in the Vaccine dataset were subsampled to a depth of 3,000 sequences in non-singleton clones (**Supplemental File 1**). Sequence depth in the Age dataset was more even, with between 370 and 2,065 (mean 1,126)

bioRxiv preprint doi: https://doi.org/10.1101/558825; this version posted February 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

unique sequences in non-singleton clones, so the repertoires in the Age dataset were not subsampled.

Phylogenetic model parameter and topology estimation

We used a single-linkage hierarchical clustering approach, detailed in **Supplemental File 1,** to assign sequences into clonal lineages, each of which were assumed to descend from a single naïve B cell ancestor. Because we were not able to reliably predict the junction regions of germline sequences (Gupta *et al.* 2015), we removed the CDR3 from all sequences analyzed. We then used the hierarchical phylogenetic model described above to quantify effects of BCR mutation and selection in the Age and Vaccine datasets.

Phylogenetic model parameters are an important source of information about evolutionary dynamics. For example, the amino acid replacement vs. silent mutation rate ratio (ω) can be used to distinguish positive and negative selection (Nielsen and Yang 1998), while the relative rate of transitions to transversions (κ) can be informative about mutation biases. We first estimated maximum likelihood tree topologies and branch lengths for each B cell lineage using the GY94 (Goldman and Yang 1994; Nielsen and Yang 1998) substitution model, in which single, shared ω and κ parameters were estimated for each repertoire, and codon frequencies were set to their empirical frequencies across all sequences within each repertoire. For computational efficiency, we used these estimated topologies to estimate branch lengths and substitution parameters of the HLP19 model at the repertoire level; namely, we estimated κ, ω_{FWR}, ω_{CDR} (separate ω values for CDR and FWRs), and individual h^a values (altered relative mutation rate) for WRC, GYW, WA,

bioRxiv preprint doi: https://doi.org/10.1101/558825; this version posted February 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

TW, SYC, and GRS hot- and cold-spot motifs (see Hoehn *et al.* (2017) for more details on these parameters).

Hypotheses concerning substitution model parameter estimates can be tested in a phylogenetic framework using a likelihood ratio test (Huelsenbeck and Rannala 1997). For models that differ only by one free parameter, a *p* value of 0.05 corresponds to a log-likelihood difference of 1.92 between the alternative (ML estimated) and null (fixed value) model (Huelsenbeck and Rannala 1997). The log-likelihood ratio test allows estimation of 95% confidence intervals for parameter estimates using profile likelihood curves. Each point on a profile likelihood curve is created by calculating the maximum likelihood obtained when the parameter of interest is fixed to a particular value and all other parameters are optimized. We used a straightforward binary search and linear interpolation approach to estimate the 95% confidence interval either side of the ML estimate.

Dataset simulation

As a means of validation, simulations (detailed in **Supplemental File 2**) were performed to test (i) the performance of the HLP19 model relative to the previous HLP17 and GY94 models (**Supplemental File 3**) and (ii) the effects of estimating parameters using a hierarchical phylogenetic model compared to inference from individual lineage trees (**Supplemental File 4**). To verify that the trends we observe in the Age and Vaccine datasets are not simply the result of biases in our parameter estimation procedure, we performed simulations using pre-specified substitution parameters ($\kappa = 2$, $\omega_{FWR} = 0.4$, $\omega_{CDR} = 0.7$, $h^{WRC} = 4$, $h^{GYW} = 6$, $h^{WA} = 4$, $h^{TW} = 2$, $h^{SYC} = -0.6$, and $h^{GRS} = -0.6$) and the same tree topologies and branch lengths as the empirical

bioRxiv preprint doi: https://doi.org/10.1101/558825; this version posted February 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

trees from the Age and Vaccine datasets. We then repeated the analyses performed in each section on these simulated datasets (**Supplemental File 6**).

bioRxiv preprint doi: https://doi.org/10.1101/558825; this version posted February 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Results

Repertoire-wide phylogenetic models improve parameter estimation

Phylogenetic substitution model parameters can be an important source of information about the evolutionary dynamics of lineages; for instance, the amino acid replacement vs. silent mutation rate ratio (ω) is used to characterize natural selection operating on genetic sequences (Nielsen and Yang 1998). The HLP19 model parameters are informative about the process of B cell affinity maturation. The model includes separate ω parameters for the framework (FWR) and complementarity determining (CDR) regions (ω_{FWR} and ω_{CDR}), the transition/transversion rate ratio (κ), and a set of altered substitution rates at SHM hot/cold-spot motifs (h^{WRC}, h^{GYW}, h^{WA}, h^{TW}, h^{SYC}, h^{GRS}; nucleotides represented using the International Union of Pure and Applied Chemistry (IUPAC) coding scheme, only underlined bases experience altered rates).

The small size of most B cell lineages poses a problem for accurate estimation of phylogenetic model parameters for individual B cell lineages. Namely, the size distributions of clonal lineages within B cell repertoires, particularly those derived from blood samples, typically follow a power-law distribution and are dominated by many lineages that each carry only a few unique sequences (Weinstein *et al.* 2009; Jiang *et al.* 2013). We confirmed this pattern using blood sample-derived BCR repertoires from 27 healthy subjects (Age dataset; Wang *et al.* 2013). Across these subjects 88 to 96% (mean: 92.3%) of lineages comprised a single unique sequence, and between 98 to 99.8% (mean: 99.3%) of lineages contained < 5 unique sequences. To test our ability to estimate phylogenetic substitution model parameters from small lineages such as these,

bioRxiv preprint doi: https://doi.org/10.1101/558825; this version posted February 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

we simulated B cell repertoire datasets using a model of SHM sequence evolution, and re-estimated HLP19 model parameters for each lineage individually (**Supplemental File 4**). These results confirm that model parameter estimation is highly inaccurate for most individual B cell lineages within this dataset, with mean proportional error rates ranging from 3.37 for $h^{\underline{GYW}}$ to 211.35 for ω_{CDR} (**Figure 1**; **Supplemental File 4**).

Given the overrepresentation of small lineages in B cell repertoires, we sought to increase the precision of parameter estimation using a hierarchical phylogenetic framework. As outlined in **Methods**, this is done by constraining model parameters to be identical among lineages, and then estimating the set of parameter values that maximizes the likelihood of the entire repertoire. Applying this hierarchical framework to the same simulation data as above (**Supplemental File 4**), we found that sharing parameters across lineages substantially reduced error in estimates for all substitution model parameters compared to individual estimates, often by multiple orders of magnitude (**Figure 1; Supplemental File 4**). Thus, a repertoire-wide phylogenetic approach leverages the information available across all lineages in the repertoire and more precisely estimates biological parameters relating to affinity maturation.

To test the strengths and weaknesses of different phylogenetic models in our hierarchical framework, we compared the performance of three codon substitution models: GY94 (Goldman and Yang 1994), which does not include SHM hot- or cold-spot motifs, HLP17 (Hoehn et al 2017), which is a modification of the GY94 model that incorporates hot- and cold-spot biases, and HLP19, which is a modification of the HLP17 introduced herein (**Supplemental File 1**) that incorporates codon frequencies in a manner that is more interpretable and formally similar to

bioRxiv preprint doi: https://doi.org/10.1101/558825; this version posted February 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

previous nonreversible models (Kaehler *et al.* 2017). Simulation analyses performed using multiple tree topologies and parameter values (**Supplemental File 3)** revealed that parameter estimates under HLP19 had a lower mean absolute bias across all parameters (0.03) than HLP17 (0.08) and GY94 (0.16; **Supplemental File 3**). For parameters relating to mutation (κ, *h* values), HLP19 showed a slightly lower mean absolute bias (0.04) compared to HLP17 (0.06; **Supplemental File 3**). HLP19 and GY94 models had similar absolute bias in branch lengths (< 0.002), which was lower than that of HLP17 (0.11; **Supplemental File 3**). HLP17 performed worse than GY94 in branch length estimation, which is surprising given Hoehn *et al.* (2017) showed that the HLP17 model improved branch length estimates compared to GY94. However, we have since determined that the simulations performed in Hoehn *et al.* (2017) were unintentionally but unfairly biased towards the HLP17 model (see detailed explanation in **Supplemental File 2)**. The simulations performed here do not have this issue, and show that HLP19 largely addresses the weaknesses of on HLP17 in branch length estimation (**Supplemental File 3**). Perhaps most importantly, for parameters relating to selection (ω_{FWR} and ω_{CDR}) HLP19 showed significantly lower mean absolute bias (0.02) compared to HLP17 (0.1) and GY94 (0.29; **Supplemental File 3**). Mean bias of ω_{CDR} estimates were especially high under the GY94 model (range: 0.38 to 0.59) and increased in simulations with higher hot-spot mutation rates and longer branch lengths (**Supplemental File 3**). This echoes previous findings; models that fail to account for altered mutation rates of SHM motifs (e.g. GY94) can significantly bias estimates of ω (dN/dS) in BCR lineages towards detecting positive selection in the CDRs (Dunn-Walters and Spencer 1998; Hershberg *et al.* 2008). Overall, we found the HLP19 model improves upon the GY94 and HLP17 models, particularly when estimating ω_{CDR} and branch lengths, respectively.

bioRxiv preprint doi: https://doi.org/10.1101/558825; this version posted February 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

To further compare the appropriateness of the GY94, HLP17, and HLP19 models when applied to BCR repertoire data, we estimated how well each model fit our empirical datasets using the Akaike information criterion (AIC; Akaike 1974). The AIC uses the maximum log-likelihood estimated using a model, penalized by the number of freely estimated parameters. Smaller AIC values are generally interpreted as better model fit. In all 27 subjects AIC was highest under GY94 and lowest under HLP19, indicating that the HLP19 model had a significantly better fit to all subjects compared to the GY94 and HLP17 models (**Supplemental File 5**).

Variation of model parameters within and among subjects

We tested whether repertoire-wide parameter estimates can reproduce known features of somatic hypermutation targeting, such as hot/cold-spot targeting (Yaari *et al.* 2013) by estimating HLP19 model parameters from BCR repertoire data that was obtained from 27 healthy individuals of varying age and sex (Age dataset; Wang *et al.* 2013). While the values of parameter estimates varied, all subjects exhibited the same overall pattern in model parameters that relate to SHM targeting (**Figure 2**). In all subjects, G<u>Y</u>W motifs exhibited the largest substitution rate increases of the all motifs considered ($h^{\underline{GYW}}$ values were 4 to 6), followed by the WR<u>C</u> ($h^{\underline{WRC}}$ ~3), W<u>A</u> ($h^{\underline{WA}}$ ~3) and <u>T</u>W ($h^{\underline{TW}}$~1) motifs. Symmetrical SY<u>C</u> and <u>G</u>RS motifs were estimated to be mutational cold-spots ($h^{\underline{SYC}}$ and $h^{\underline{GRS}}$ ~ -0.6). We compared these parameter estimates to mutability estimates under the S5F model (Yaari *et al.* 2013), which describes the relative mutation rate of sequence pentamers during SHM in an independent and separate cohort of healthy subjects. When averaging over pentamers within particular SHM motifs under uniform pentamer frequencies, the S5F model predicts the same ranking as we obtained using the HLP19 model: G<u>Y</u>W (mean mutability= 2.46) > <u>W</u>RC (1.87) ≈ W<u>A</u> (1.71) > <u>T</u>W (1.19) > SY<u>C</u> (0.23) ≈

bioRxiv preprint doi: https://doi.org/10.1101/558825; this version posted February 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

GRS (0.22).The transition/transversion rate ratio (κ) estimated by our repertoire-wide model was ~2, which is also consistent with previous findings (Betz *et al.* 1993; Cowell and Kepler 2000). Overall, these results show that repertoire-wide parameter estimates obtained using a hierarchical phylogenetic approach are broadly consistent with previous expectations in healthy individuals.

Age is associated with changes in SHM mutation biases

Age and sex are associated with substantial differences in the immune system; for example, older individuals are more vulnerable to infection (Castle 2000; Fink *et al.* 2018), while females are at a higher risk of developing auto-immune diseases (Ansar Ahmed *et al.* 1985). We sought to investigate whether the mutation and selection processes underlying SHM might contribute to these differences.

To investigate potential age- and sex-related differences in SHM targeting, we analyzed the 27 subjects surveyed by Wang et al (2013; Age dataset), which included both male and female subjects with an age range of 21 to 88 years at the time of sampling. We used multiple linear regression to investigate the effects of age and sex on estimated model parameters. Age and sex were modeled as interaction variables against the estimated substitution rate biases of SHM motifs (i.e. the HLP19 model *h* values; **Supplemental File 1**). Because we conducted 18 tests in all (two dependent and nine independent variables), we used Benjamini-Hochberg (Benjamini and Hochberg 1995) multiple hypothesis test correction to adjust *p*-values. Substitution rates in WA (h^{WA}) were significantly negatively associated with age in both male (coefficient=-0.011; adjusted p=0.001), and female subjects (coefficient=-0.006; adjusted p=0.03). No other parameter showed a significant interaction with age in either sex after *p*-value adjustment. These

bioRxiv preprint doi: https://doi.org/10.1101/558825; this version posted February 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

results are consistent with a model in which older individuals have reduced mutation bias towards WA hot-spots, possibly reflecting a difference in SHM mechanism in these individuals.

We performed simulation analyses to test whether the observed trends between h^{WA} and age could be due to biases in our parameter estimation procedure (*Methods*; **Supplemental File 6**). For all 20 simulated repetitions of the Age dataset, the h^{WA} slope coefficients for males and females were closer to zero than their respective empirical estimates (**Supplemental File 6a**). These results demonstrate that these trends are due to factors other than biases in parameter estimation, given the underlying structure of our datasets and predicted germline sequences.

Variation in signatures of selection is uncorrelated with age, sex, EBV, and CMV status

Antigen-driven selection plays a major role in shaping BCR repertoire diversity. In molecular evolutionary biology, selective dynamics are often characterized by estimating the relative rate of substitutions that change amino acids versus those that do not, often called dN/dS or ω (Nielsen and Yang 1998). Low ω values are indicative of fewer amino acid changes than expected, which is generally interpreted as resulting from negative selection. We estimate ω separately for the complementarity determining regions (CDRs) and framework regions (FWR). Estimates of ω_{FWR} are expected to be lower than those of ω_{CDR} because FWRs are more structurally constrained than CDRs (Shlomchik *et al.* 1989), which are primarily used in antigen binding (Murphy *et al.* 2012; Yaari *et al.* 2012). Consistent with this expectation, we found that in the Age dataset, estimated ω_{CDR} values (range: 0.52 – 0.87, mean: 0.68) were higher than estimated ω_{FWR} values (range: 0.44 – 0.56, mean: 0.51) in all 27 subjects ($p < 0.001$; paired Wilcoxon test; **Figure 2**). ω_{CDR} estimates were also more varied among subjects than ω_{FWR}

bioRxiv preprint doi: https://doi.org/10.1101/558825; this version posted February 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

values, perhaps representing different individual histories of antigenic stimulation. However, we were unable to find a clear biological correlate of ω_{CDR} in the Age dataset among the variables provided with the data (Wang *et al.* 2013). Specifically, values of ω_{CDR} did not show a significant relationship with age (slope *p*-value = 0.66; least squares regression; **Figure 3**), sex (*p* = 1.0; Wilcoxon rank sum test), Epstein-Barr virus seropositivity (*p* = 0.19; Wilcoxon rank sum test), or cytomegalovirus seropositivity (*p* = 0.19; Wilcoxon rank sum test).

Post-influenza vaccination repertoires show signs of negative selection and longer tree length

Influenza vaccination substantially perturbs the B cell repertoire. A large, antigen-specific plasmablast response is observed in the blood ~7 days post-vaccination which subsides approximately one week later (Wrammert *et al.* 2008; Jackson *et al.* 2014). To investigate the selective dynamics of this process, we estimated HLP19 substitution model parameters using the repertoires of three otherwise healthy subjects who were sampled 10 times over the course of influenza vaccination, beginning 8 days prior to vaccination and ending 28 days afterwards (Laserson *et al.* 2014). Because we were primarily interested in selection and genetic diversity of these samples, we focused on changes in ω_{CDR}, the relative rate of nonsynonymous/synonymous substitutions, and tree length (the total expected substitutions per site within an individual lineage phylogeny).

We found a variety of responses among the subjects. PGP1, the oldest subject of the three, did not show any clear patterns of change over time, in either mean tree length or ω_{CDR}. Notably, this subject at day +14 had only 141 sequences and consequently very wide 95% confidence

bioRxiv preprint doi: https://doi.org/10.1101/558825; this version posted February 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

intervals, illustrating the importance of correctly estimating model uncertainty in analysis of BCR sequence data.

In contrast to PGP1, subjects 420IV and hu420143 both showed increased tree length at day+7 compared to one hour prior to vaccination (-1h), consistent with the expected burst of BCR genetic diversity 7 days post-vaccination (Wrammert *et al.* 2008). The estimated mean tree length within a sample was highest at day+7 for subjects 420IV and hu420143, with a fold increase of 2.38 and 1.18 compared to one hour prior to vaccination (-1h) (**Figure 4**). Consistent with this, multiple large clones in subjects 420IV and hu420143 arose at day+7 (**Supplemental File 8**). In addition to increased tree length, day+7 was associated with a significant decrease in ω_{CDR} in both these subjects (**Figure 4**). For 420IV at -1h, ω_{CDR} = 0.64 (95% CI: 0.6, 0.66) and at day+7 ω_{CDR} = 0.47 (95% CI: 0.45, 0.50). For hu420143 at -1h, ω_{CDR} = 0.57 (95% CI: 0.54, 0.59) and at day +7 ω_{CDR} = 0.49 (95% CI: 0.46, 0.51). Interestingly, though 420IV and hu420139 had different pre-vaccination estimates of ω_{CDR} (0.64 and 0.57, respectively) their estimates were similar at day+7 (0.47 and 0.49), day+14 (0.62 and 0.61), and day+21 (0.60 and 0.60; **Figure 4**). Overall, this indicates that, at the expected date of peak vaccine response, the repertoires of these two subjects were characterized by an increase in genetically-diverse BCR lineages and signatures of increased negative selection.

We performed simulation analyses to test whether decreased ω_{CDR} at day+7 in subjects hu420139 and 420IV were due to biases in our parameter estimation procedure (*Methods*, **Supplemental File 6**). None of the 20 simulation repetitions performed using the Vaccine dataset were able to reproduce the observed change in ω_{CDR} at day+7 compared to the pre-

bioRxiv preprint doi: https://doi.org/10.1101/558825; this version posted February 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

vaccination time point (-1h; **Supplemental File 6b** and **6c**), demonstrating that these trends are due to factors besides biases in parameter estimation, given the underlying structure of our datasets and their predicted germline sequences.

Increased tree length is associated with signatures of negative selection

Our analysis of the Vaccine dataset indicated that, in two subjects, there was a concurrent increase in mean tree length and decrease in ω_{CDR}, at day +7 following influenza vaccination. We hypothesized that this relationship between ω_{CDR} and tree length might be more general, and tested this hypothesis using log-linear regression across all 27 subjects of the Age dataset and all 30 samples (10 time points from three subjects) of the Vaccine dataset. Across both datasets we observed a consistent and significant negative relationship between both ω_{CDR} and ω_{FWR}, and mean repertoire tree length (i.e. the average expected substitutions per codon site across all lineages within the repertoire; **Figure 5**). This trend was surprisingly similar between datasets, with slopes of linear regressions having overlapping 95% confidence intervals, and was particularly strong in the CDRs. For the Age dataset, the slope of a linear regression of ω_{CDR} against the ln(mean tree length) was -0.24 (95% CI = -0.35, -0.14; $p < 6x10^{-5}$), while for the Vaccine dataset the corresponding slope was -0.26 (95% CI = -0.29, -0.23; $p < 4x10^{-16}$). Overall, these regressions show a 32.1% and 41.4% decrease in ω_{CDR} over the range of mean tree length observed in the Age and Vaccine datasets, respectively. A similar, if weaker, relationship was found between ω_{FWR} and ln(mean tree length) (**Figure 5**; details in caption). This indicates that repertoires with longer (i.e. more diverged) lineages are associated with signatures of increased negative selection, particularly in the CDRs.

bioRxiv preprint doi: https://doi.org/10.1101/558825; this version posted February 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

We performed simulation analyses to test whether the observed trends between ω and mean tree length were due to biases in our parameter estimation procedure (*Methods,* **Supplemental File 6**). In none of 20 simulations, using both datasets, did we observe a significant relationship between of ω_{CDR} and mean tree length or ω_{FWR} and mean tree length (**Supplemental File 6d-e)**. However, the simulations in **Supplemental File 6** were performed under a fully-context dependent version of the HLP19 model, which does not completely represent the biased nature of SHM. To test whether a richer model of SHM could potentially reproduce our results, we performed simulations using the S5F model (Yaari *et al.* 2013), using the tree topologies and branch lengths estimated using maximum parsimony (dnapars v3.679; Felsenstein 2002) and the predicted germline sequences of the Age dataset (detailed in **Supplemental File 7**). None of 50 such simulation repetitions showed a negative slope between ω_{CDR} and mean tree length as large as that observed for the empirical data; only one simulation repetition showed a more negative slope than observed between ω_{FWR} and mean tree length (**Supplemental File 7**). We therefore conclude that the negative relationship between mean tree length and ω_{CDR} observed in **Figure 5** is not due simply to inherent biases in our parameter estimation procedure.

bioRxiv preprint doi: https://doi.org/10.1101/558825; this version posted February 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Discussion

Phylogenetic techniques have been used to study B cell lineages for many years (Shlomchik *et al.* 1987) and continue to be a powerful tool in understanding affinity maturation (Vieira *et al.* 2018). Two fundamental issues that arise from the application of phylogenetic techniques to B cell repertoires are (i) the biology underlying B cell affinity maturation violates key assumptions of most phylogenetic models, and (ii) phylogenetic models are designed typically to work on lineages originating from a common ancestor, but B cell repertoires are composed of multiple lineages with separate ancestries, many of which are composed of only a few unique sequences. If such lineages are each analyzed independently then parameter estimates will be noisy and highly uncertain. Here we introduce a hierarchical approach to B cell phylogenetics that addresses these issues. We extend phylogenetic models of SHM evolution so that they are consistent with the known biology of B cell affinity maturation, and can share parameters of the sequence evolution process across all lineages within a repertoire. This approach outperforms the alternative of estimating parameters for each lineage individually. By applying our new approach to empirical data we find evidence consistent with dysregulation of SHM in older subjects, increased negative selection during influenza vaccine response, and a relationship between mean tree lengths and signatures of negative selection.

We first used our hierarchical framework to demonstrate a negative association between age and the estimated mutability of WA hot-spot motifs in males and females. Previous studies have shown that aging is associated with a decrease in the affinity, specificity and diversity of antibodies produced (LeMaoult *et al.* 1997; Dunn-Walters *et al.* 2003; Weiskopf *et al.* 2009;

bioRxiv preprint doi: https://doi.org/10.1101/558825; this version posted February 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Dunn-Walters 2016), as well as with a number of changes at the repertoire level, including longer CDR3s, higher levels of SHM, and persistent clonal lineages in the blood (Wang et al. 2013). It is possible that age-related dysregulation of SHM machinery plays a role in phenomena associated with immunosenescence. Our finding that older individuals tend to have altered mutability of WA motifs is consistent with this hypothesis. However, we note that this effect is observed only in a small cohort, and while the linear regression slopes we observed were significant, the overall change in WA mutability was modest. While we were unable to reproduce this relationship in simulations under a null model (**Supplemental File 6**), it is possible this trend is driven by other confounding factors we have not considered. Future analyses with more subjects are needed to validate this trend generally.

We further used our hierarchical phylogenetic approach to characterize BCR molecular evolution during vaccination. B cell receptors during affinity maturation are subject to multiple selective pressures; positive selection to introduce new affinity-increasing amino acid variants (higher ω) and negative selection to remove affinity-decreasing variants (lower ω). *A priori*, we might expect positive selection to predominate during vaccine response. However, in our analysis we found that lineages present at the time of peak influenza vaccine response show signs of increased negative selection (lower ω) on CDRs. We suggest this is because B-cell lineages with a history of affinity maturation during influenza infection/vaccination will likely have already evolved effective or nearly effective neutralization at the time of vaccination, resulting in a greater proportion of amino acid changes being deleterious (Clarke *et al.* 1985). This would result in lower ω_{CDR} values, which may be particularly marked during influenza vaccine

bioRxiv preprint doi: https://doi.org/10.1101/558825; this version posted February 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

response, since B cells activated by influenza vaccination in adults are expected to derive from re-activated memory B cell lineages (Vollmers *et al.* 2013).

We also observed a negative relationship between tree length and ω_{CDR} across both of our datasets (**Figure 5**). This relationship is remarkably consistent given that our combined datasets contain a total of 30 subjects of different age, sex, and treatment status. None of the simulation analyses performed under a null model were able to reproduce this result (**Supplemental File 6** and **7**), so this relationship is unlikely to be due to a bias or intrinsic correlation between these variables in our estimation procedure. In the absence of other obvious confounding factors, we posit a simple biological explanation: as B cell clones accumulate mutations through repeated rounds of affinity maturation, their binding affinity to target antigen increases and consequently the benefit of random amino acid changes (i.e. new mutations) decreases (Clarke *et al.* 1985). This idea that the rate of fitness-increasing mutations decreases is a straightforward implication of a population nearing a "peak" within a fitness landscape (Wright 1932). Sheng *et al.* (2016) demonstrated evidence of this process (which they termed the "affinity maturation selection" model) in anti-HIV broadly neutralizing antibody (bnAb) lineages. This explanation is also consistent with the findings of Yaari *et al.* (2015) showing that mutations earlier in B cell lineage trees from healthy subjects show clearer signs of positive selection than more recent mutations. Our results suggest that decreased rates of non-synonymous mutations relative to synonymous mutations, as observed in HIV bnAb lineages (Sheng *et al.* 2016), BCR repertoires during vaccine response (**Figures 4** and **5**), and even in healthy subjects with no obvious signs of infection (**Figure 5**; Yaari et al. 2015), are all special cases of a general feature of affinity maturation.

bioRxiv preprint doi: https://doi.org/10.1101/558825; this version posted February 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

There are several limitations to the hierarchical phylogenetic approach introduced here. The most obvious is that by constraining parameter values so that they are identical for all lineages within a repertoire, we mask any potential parameter variation among lineages. Estimating separate parameter values for one or a set of lineages within a repertoire may be easily accommodated within the hierarchical framework (Suchard *et al.* 2003). This may be useful for parameters such as ω_{CDR}, which might reflect lineage-specific histories of antigen-driven selection (Horns *et al.* 2019). As an example of this approach, we explore heterogeneity in estimates of ω_{CDR} among lineages of different sizes for one repertoire (**Supplemental File 9**). It is unclear whether estimation of individual κ and h^a values would yield useful insights, since these parameters relate primarily to biases resulting from SHM, and there is little *a priori* reason to believe they might vary among B cell lineages within an individual. However, it is clear that estimating parameters (e.g. ω_{CDR}) for each lineage individually will lead to issues with over-fitting (e.g., when all CDR mutations within a lineage are non-synonymous). Further work will be needed to resolve lineage heterogeneity within individual repertoires.

A hierarchical phylogenetic approach to BCR phylogenetics is justified theoretically and represents a step forward in the statistical analysis of B cell repertoires. Our new methods are implemented in the program IgPhyML (v1.0.7; https://bitbucket.org/kbhoehn/igphyml), which is freely available and integrated into the Immcantation suite (http://immcantation.org).

bioRxiv preprint doi: https://doi.org/10.1101/558825; this version posted February 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figures

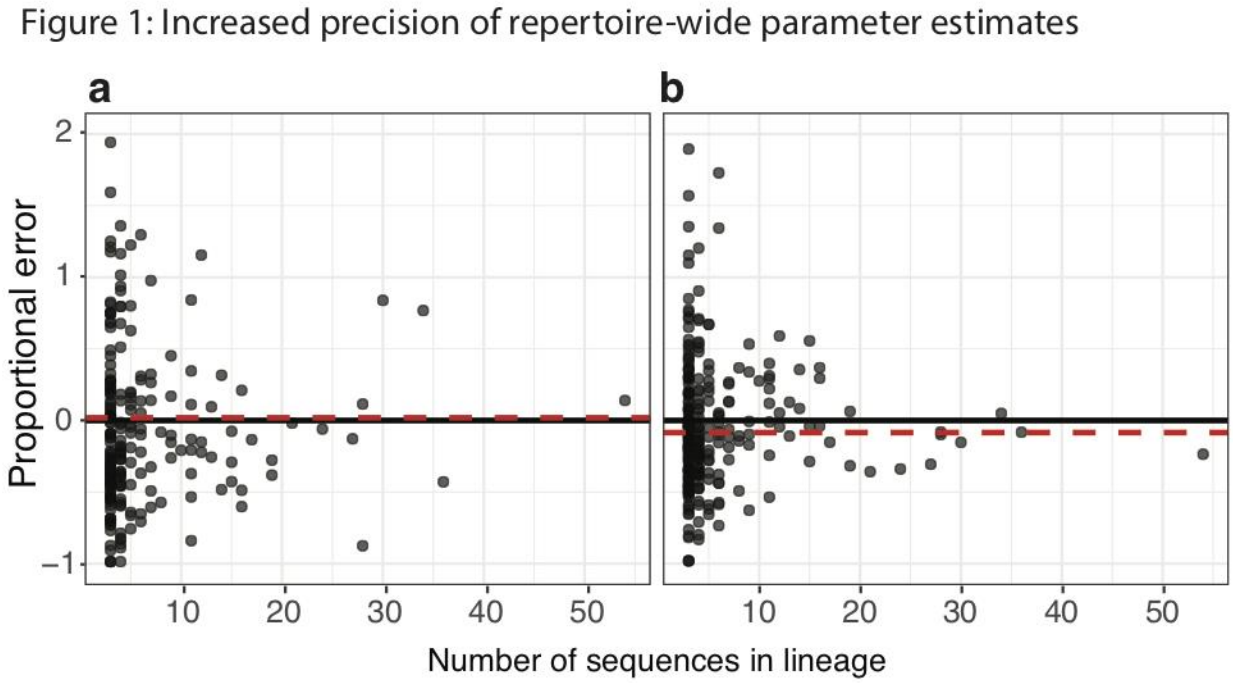

Figure 1: Increased precision of repertoire-wide parameter estimates (**a**) Proportional error in estimates of the ω_{CDR} parameter under the HLP19 model. (**b**) Proportional error in estimates of the ω_{FWR} parameter under the HLP19 model. In both panels, the black dots show the values estimated from each individual lineage B cell lineage and the red lines shows the estimate obtained from all lineages combined using a hierarchical model. Data were generated from a simulated repertoire using tree topologies from subject 97 in the Age dataset (see **Supplemental File 4** for full details and results).

bioRxiv preprint doi: https://doi.org/10.1101/558825; this version posted February 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

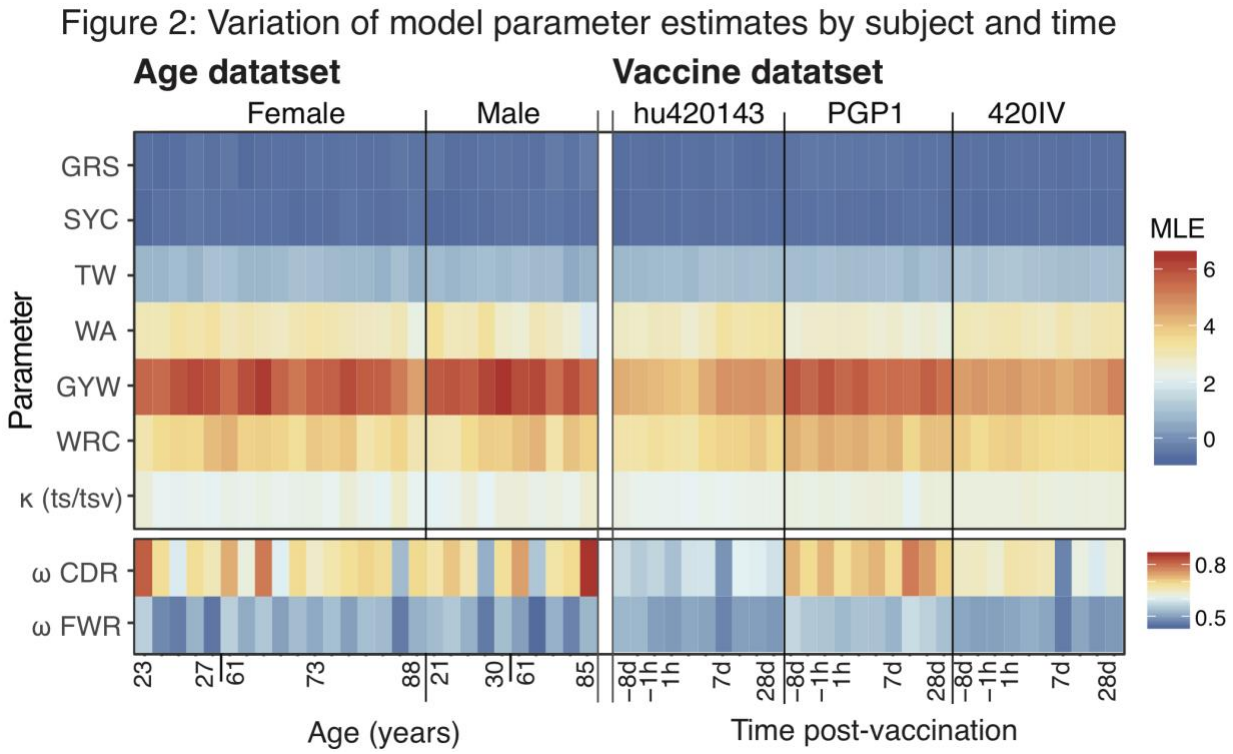

Figure 2: Variation of parameter estimates by subject and time in the Age and Vaccine datasets (**a)** HLP19 parameter estimates from each subject in the Age dataset, ordered by sex and age. (**b**) HLP19 parameter estimates for the Vaccine dataset, ordered by subject and sample time relative to influenza vaccination. The upper box in both (a) and (b) shows the model parameters that relate to somatic hypermutation (motif targeting, transition/transversion ratio). The lower box shows estimates of ω_{CDR} and ω_{FWR}, which relate to selection. Note that values in the lower box are scaled differently from those in the upper panel (see legends on right hand side).

bioRxiv preprint doi: https://doi.org/10.1101/558825; this version posted February 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

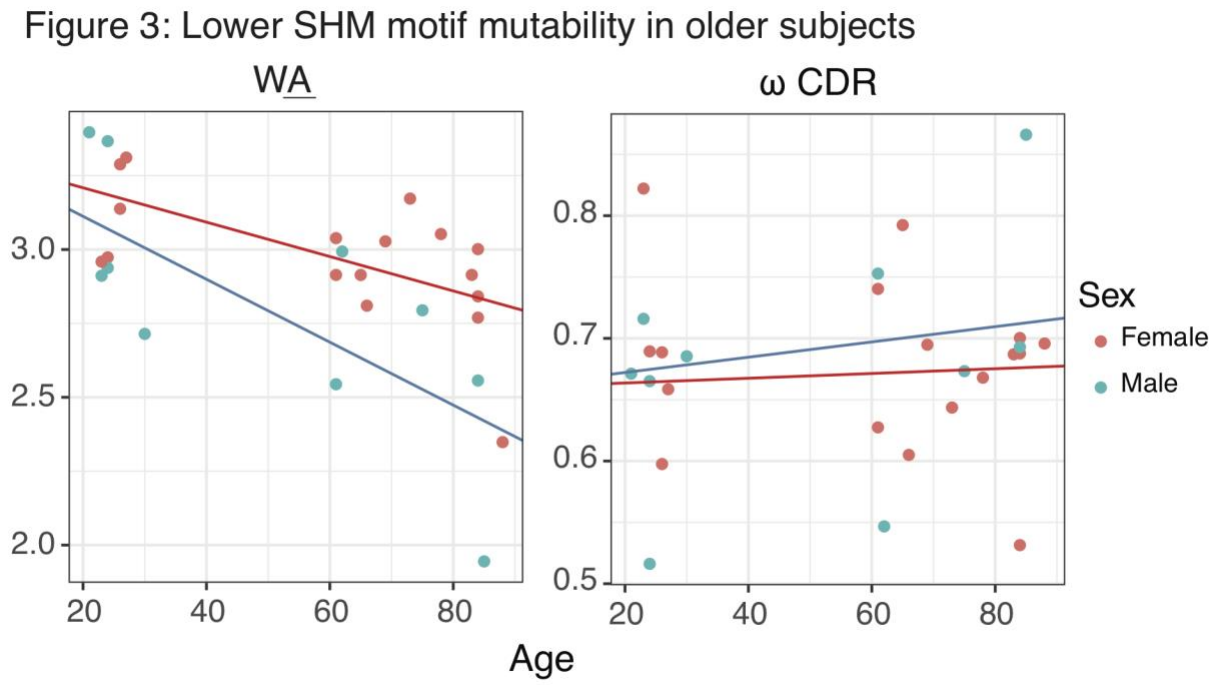

Figure 3: Change in SHM motif mutation rates with age. Linear regressions of age against model parameters under the HLP19 model, estimated using maximum likelihood. Separate regressions are shown for male (blue) and female (red) subjects. The h^{WA} parameter shows a significant decrease with age in males and females.

bioRxiv preprint doi: https://doi.org/10.1101/558825; this version posted February 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

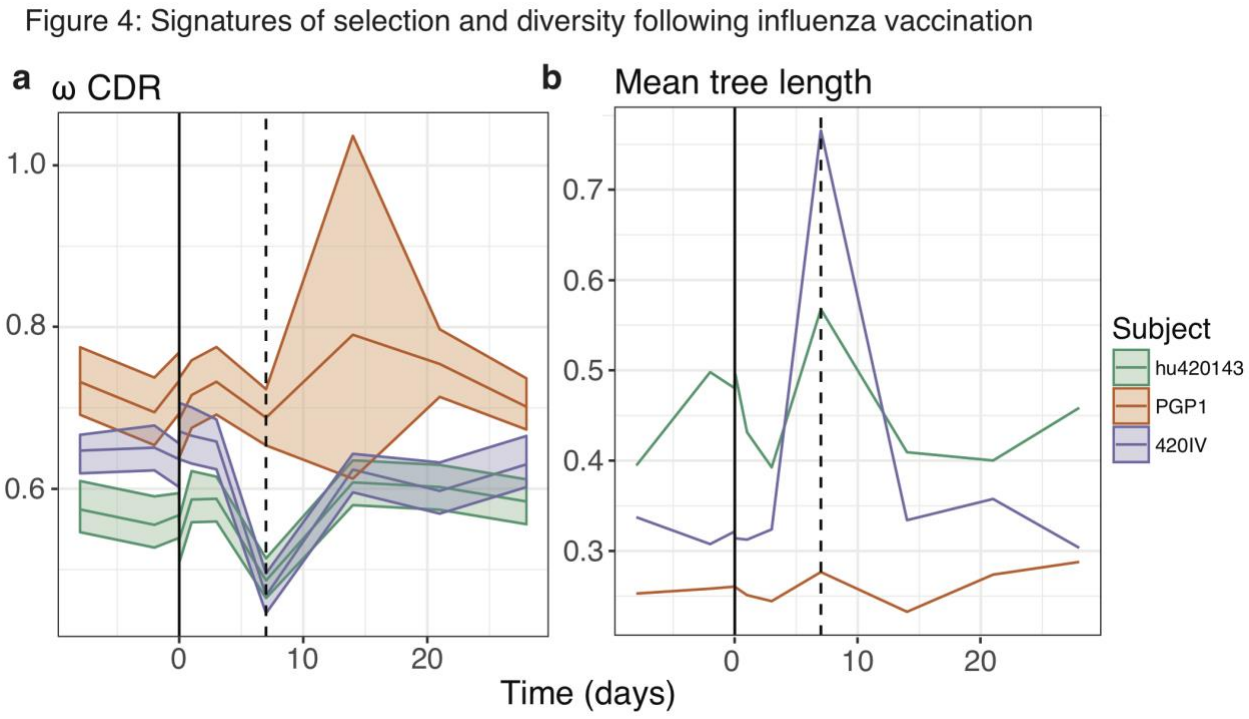

Figure 4: Signatures of selection and diversity following influenza vaccination (**a**) Estimates of the ω_{CDR} parameter of the HLP19 model over the course of influenza vaccination. (**b**) Estimates of mean tree length (total substitutions per codon within a lineage, averaged across all lineages within a repertoire) over the course of influenza vaccination. The x-axis shows the number of days since vaccination; the vertical dashed line represents day 7 post-vaccination. The shaded areas in panel (a) represent the 95% confidence intervals for each parameter estimate, calculated using profile likelihood curves.

bioRxiv preprint doi: https://doi.org/10.1101/558825; this version posted February 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 5: Negative relationship between ω and mean tree length

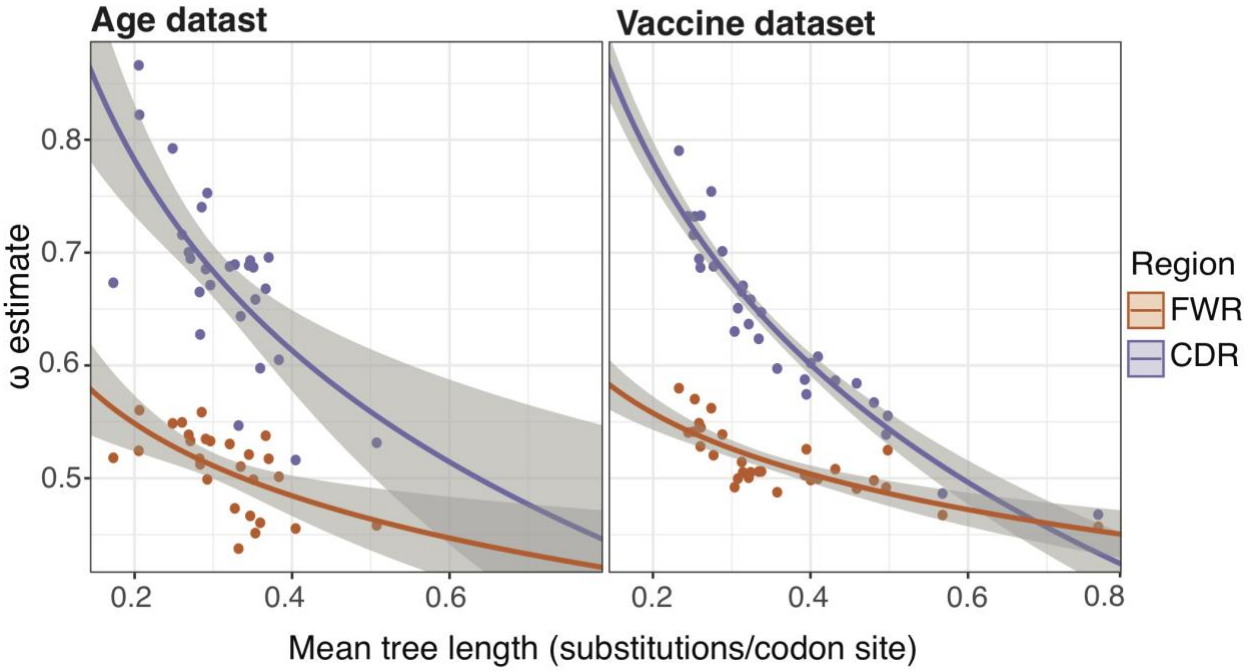

Figure 5: Negative relationship between ω and mean tree length (**a**) Linear regression between estimates of ω_{CDR} (purple) and ω_{FWR} (orange) and the natural log of mean tree length for each subject in the Age dataset. The slope and intercept of ω_{CDR} against ln(mean tree length) were -0.24 (95% CI = -0.35, -0.14) and 0.39, respectively ($p < 6x10^{-5}$ for both). The corresponding slope and intercept of ω_{FWR} were -0.09 (95% CI = -0.14, -0.04) and 0.4 ($p < 0.002$ for both). (**b**) Linear regression between estimates of ω_{CDR} (purple) and ω_{FWR} (orange) and the natural log of mean tree length for each sample in the Vaccine dataset (three subjects, 10 samples each). The slope and intercept of ω_{CDR} against ln(mean tree length) were -0.26 (95% CI = -0.29, -0.23) and 0.36, respectively ($p < 4x10^{-16}$ for both). The corresponding slope and intercept of ω_{FWR} were -0.08 (95% CI = -0.1, -0.05) and 0.43 ($p < 4x10^{-7}$ for both). Grey shaded areas in both panels show standard error estimates of the log-linear regression.

bioRxiv preprint doi: https://doi.org/10.1101/558825; this version posted February 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

References

Akaike H., 1974 A new look at the statistical model identification. IEEE Trans. Autom. Control 19: 716–723.

Ansar Ahmed S., W. J. Penhale, and N. Talal, 1985 Sex hormones, immune responses, and autoimmune diseases. Mechanisms of sex hormone action. Am. J. Pathol. 121: 531–551.

Barak M., N. S. Zuckerman, H. Edelman, R. Unger, and R. Mehr, 2008 IgTree©: Creating Immunoglobulin variable region gene lineage trees. J. Immunol. Methods 338: 67–74. https://doi.org/10.1016/j.jim.2008.06.006

Bashford-Rogers R. J. M., A. L. Palser, B. J. Huntly, R. Rance, G. S. Vassiliou, *et al.*, 2013 Network properties derived from deep sequencing of human B-cell receptor repertoires delineate B-cell populations. Genome Res. 23: 1874–1884. https://doi.org/10.1101/gr.154815.113

Benjamini Y., and Y. Hochberg, 1995 Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J. R. Stat. Soc. Ser. B Methodol. 57: 289–300.

Betz A. G., C. Rada, R. Pannell, C. Milstein, and M. S. Neuberger, 1993 Passenger transgenes reveal intrinsic specificity of the antibody hypermutation mechanism: clustering, polarity, and specific hot spots. Proc. Natl. Acad. Sci. 90: 2385–2388. https://doi.org/10.1073/pnas.90.6.2385

Boyd S. D., E. L. Marshall, J. D. Merker, J. M. Maniar, L. N. Zhang, *et al.*, 2009 Measurement and Clinical Monitoring of Human Lymphocyte Clonality by Massively Parallel V-D-J Pyrosequencing. Sci. Transl. Med. 1: 12ra23-12ra23. https://doi.org/10.1126/scitranslmed.3000540

Castle S. C., 2000 Clinical Relevance of Age-Related Immune Dysfunction. Clin. Infect. Dis. 31: 578–585. https://doi.org/10.1086/313947

Clarke S. H., K. Huppi, D. Ruezinsky, L. Staudt, W. Gerhard, *et al.*, 1985 Inter- and intraclonal diversity in the antibody response to influenza hemagglutinin. J. Exp. Med. 161: 687–704. https://doi.org/10.1084/jem.161.4.687

Cowell L. G., and T. B. Kepler, 2000 The Nucleotide-Replacement Spectrum Under Somatic Hypermutation Exhibits Microsequence Dependence That Is Strand-Symmetric and Distinct from That Under Germline Mutation. J. Immunol. 164: 1971–1976. https://doi.org/10.4049/jimmunol.164.4.1971

DeWitt W. S., L. Mesin, G. D. Victora, V. N. Minin, F. A. Matsen, *et al.*, 2018 Using Genotype Abundance to Improve Phylogenetic Inference. Mol. Biol. Evol. 35: 1253–1265. https://doi.org/10.1093/molbev/msy020

bioRxiv preprint doi: https://doi.org/10.1101/558825; this version posted February 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Dunn-Walters D. K., and J. Spencer, 1998 Strong intrinsic biases towards mutation and conservation of bases in human IgVH genes during somatic hypermutation prevent statistical analysis of antigen selection. Immunology 95: 339–345.

Dunn-Walters D. K., M. Banerjee, and R. Mehr, 2003 Effects of age on antibody affinity maturation. Biochem. Soc. Trans. 31: 447–448. https://doi.org/10.1042/bst0310447

Dunn-Walters D. K., 2016 The ageing human B cell repertoire: a failure of selection?: Ageing B cell repertoire. Clin. Exp. Immunol. 183: 50–56. https://doi.org/10.1111/cei.12700

Felsenstein J., 1981 Evolutionary trees from DNA sequences: A maximum likelihood approach. J. Mol. Evol. 17: 368–376. https://doi.org/10.1007/BF01734359

Felsenstein J., 2002 *{PHYLIP} (Phylogeny Inference Package) version 3.6a3*.

Feng J., D. A. Shaw, V. N. Minin, N. Simon, and F. A. Matsen IV, 2017 Survival analysis of DNA mutation motifs with penalized proportional hazards. ArXiv171104057 Q-Bio Stat.

Fink A. L., K. Engle, R. L. Ursin, W.-Y. Tang, and S. L. Klein, 2018 Biological sex affects vaccine efficacy and protection against influenza in mice. Proc. Natl. Acad. Sci. 115: 12477–12482. https://doi.org/10.1073/pnas.1805268115

Goldman N., and Z. Yang, 1994 A codon-based model of nucleotide substitution for protein-coding DNA sequences. Mol. Biol. Evol. 11: 725–736.

Greiff V., P. Bhat, S. C. Cook, U. Menzel, W. Kang, *et al.*, 2015a A bioinformatic framework for immune repertoire diversity profiling enables detection of immunological status. Genome Med. 7: 49. https://doi.org/10.1186/s13073-015-0169-8

Greiff V., E. Miho, U. Menzel, and S. T. Reddy, 2015b Bioinformatic and Statistical Analysis of Adaptive Immune Repertoires. Trends Immunol. 36: 738–749. https://doi.org/10.1016/j.it.2015.09.006

Gupta N. T., J. A. Vander Heiden, M. Uduman, D. Gadala-Maria, G. Yaari, *et al.*, 2015 Change-O: a toolkit for analyzing large-scale B cell immunoglobulin repertoire sequencing data: Table 1. Bioinformatics btv359. https://doi.org/10.1093/bioinformatics/btv359

Hershberg U., M. Uduman, M. J. Shlomchik, and S. H. Kleinstein, 2008 Improved methods for detecting selection by mutation analysis of Ig V region sequences. Int. Immunol. 20: 683–694. https://doi.org/10.1093/intimm/dxn026

Hoehn K. B., G. Lunter, and O. G. Pybus, 2017 A Phylogenetic Codon Substitution Model for Antibody Lineages. Genetics 206: 417–427. https://doi.org/10.1534/genetics.116.196303

Horns F., C. Vollmers, C. L. Dekker, and S. R. Quake, 2019 Signatures of selection in the human antibody repertoire: Selective sweeps, competing subclones, and neutral drift. Proc. Natl. Acad. Sci. 201814213. https://doi.org/10.1073/pnas.1814213116

bioRxiv preprint doi: https://doi.org/10.1101/558825; this version posted February 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Huelsenbeck J. P., and B. Rannala, 1997 Phylogenetic Methods Come of Age: Testing Hypotheses in an Evolutionary Context. Science 276: 227–232. https://doi.org/10.1126/science.276.5310.227

Jackson K. J. L., Y. Liu, K. M. Roskin, J. Glanville, R. A. Hoh, *et al.*, 2014 Human Responses to Influenza Vaccination Show Seroconversion Signatures and Convergent Antibody Rearrangements. Cell Host Microbe 16: 105–114. https://doi.org/10.1016/j.chom.2014.05.013

Jiang N., J. He, J. A. Weinstein, L. Penland, S. Sasaki, *et al.*, 2013 Lineage Structure of the Human Antibody Repertoire in Response to Influenza Vaccination. Sci. Transl. Med. 5: 171ra19-171ra19. https://doi.org/10.1126/scitranslmed.3004794

Kaehler B. D., V. B. Yap, and G. A. Huttley, 2017 Standard Codon Substitution Models Overestimate Purifying Selection for Nonstationary Data. Genome Biol. Evol. 9: 134–149. https://doi.org/10.1093/gbe/evw308

Kepler T. B., 2013 Reconstructing a B-cell clonal lineage. I. Statistical inference of unobserved ancestors. F1000Research 2. https://doi.org/10.12688/f1000research.2-103.v1

Laserson U., F. Vigneault, D. Gadala-Maria, G. Yaari, M. Uduman, *et al.*, 2014 High-resolution antibody dynamics of vaccine-induced immune responses. Proc. Natl. Acad. Sci. 111: 4928–4933. https://doi.org/10.1073/pnas.1323862111

LeMaoult J., P. Szabo, and M. E. Weksler, 1997 Effect of age on humoral immunity, selection of the B-cell repertoire and B-cell development. Immunol. Rev. 160: 115–126. https://doi.org/10.1111/j.1600-065X.1997.tb01032.x

Liao H.-X., R. Lynch, T. Zhou, F. Gao, S. M. Alam, *et al.*, 2013 Co-evolution of a broadly neutralizing HIV-1 antibody and founder virus. Nature 496: 469–476. https://doi.org/10.1038/nature12053

Murphy K., P. Travers, M. Walport, and C. Janeway, 2012 *Janeway's immunobiology*. Garland Science, New York.

Nielsen R., and Z. Yang, 1998 Likelihood Models for Detecting Positively Selected Amino Acid Sites and Applications to the HIV-1 Envelope Gene. Genetics 148: 929–936.

Peled J. U., F. L. Kuang, M. D. Iglesias-Ussel, S. Roa, S. L. Kalis, *et al.*, 2008 The Biochemistry of Somatic Hypermutation. Annu. Rev. Immunol. 26: 481–511. https://doi.org/10.1146/annurev.immunol.26.021607.090236

Rodrigo A. G., M. Goode, R. Forsberg, H. A. Ross, and A. Drummond, 2003 Inferring Evolutionary Rates Using Serially Sampled Sequences from Several Populations. Mol. Biol. Evol. 20: 2010–2018. https://doi.org/10.1093/molbev/msg215

Sheng Z., C. A. Schramm, M. Connors, L. Morris, J. R. Mascola, *et al.*, 2016 Effects of Darwinian Selection and Mutability on Rate of Broadly Neutralizing Antibody Evolution

bioRxiv preprint doi: https://doi.org/10.1101/558825; this version posted February 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

during HIV-1 Infection. PLOS Comput Biol 12: e1004940. https://doi.org/10.1371/journal.pcbi.1004940

Shlomchik M. J., A. Marshak-Rothstein, C. B. Wolfowicz, T. L. Rothstein, and M. G. Weigert, 1987 The role of clonal selection and somatic mutation in autoimmunity. Nature 328: 805–811. https://doi.org/10.1038/328805a0

Shlomchik M. J., S. Litwin, and M. Weigert, 1989 The Influence of Somatic Mutation on Clonal Expansion, pp. 415–423 in *Progress in Immunology*, edited by Melchers F., Albert E. D., Boehmer H. von, Dierich M. P., Du Pasquier L., *et al.* Springer Berlin Heidelberg.

Stern J. N. H., G. Yaari, J. A. V. Heiden, G. Church, W. F. Donahue, *et al.*, 2014 B cells populating the multiple sclerosis brain mature in the draining cervical lymph nodes. Sci. Transl. Med. 6: 248ra107-248ra107. https://doi.org/10.1126/scitranslmed.3008879

Suchard M. A., C. M. R. Kitchen, J. S. Sinsheimer, R. E. Weiss, and Z. Yang, 2003 Hierarchical Phylogenetic Models for Analyzing Multipartite Sequence Data. Syst. Biol. 52: 649–664. https://doi.org/10.1080/10635150390238879

Vieira M. C., D. Zinder, and S. Cobey, 2018 Selection and Neutral Mutations Drive Pervasive Mutability Losses in Long-Lived Anti-HIV B-Cell Lineages. Mol. Biol. Evol. 35: 1135–1146. https://doi.org/10.1093/molbev/msy024

Vollmers C., R. V. Sit, J. A. Weinstein, C. L. Dekker, and S. R. Quake, 2013 Genetic measurement of memory B-cell recall using antibody repertoire sequencing. Proc. Natl. Acad. Sci. 110: 13463–13468.

Wang C., Y. Liu, L. T. Xu, K. J. L. Jackson, K. M. Roskin, *et al.*, 2013 Effects of Aging, Cytomegalovirus Infection, and EBV Infection on Human B Cell Repertoires. J. Immunol. 1301384. https://doi.org/10.4049/jimmunol.1301384

Weinstein J. A., N. Jiang, R. A. White, D. S. Fisher, and S. R. Quake, 2009 High-Throughput Sequencing of the Zebrafish Antibody Repertoire. Science 324: 807–810. https://doi.org/10.1126/science.1170020

Weiskopf D., B. Weinberger, and B. Grubeck-Loebenstein, 2009 The aging of the immune system. Transpl. Int. 22: 1041–1050. https://doi.org/10.1111/j.1432-2277.2009.00927.x

Wrammert J., K. Smith, J. Miller, W. A. Langley, K. Kokko, *et al.*, 2008 Rapid cloning of high-affinity human monoclonal antibodies against influenza virus. Nature 453: 667–671. https://doi.org/10.1038/nature06890

Wright S., 1932 The Roles of Mutation, Inbreeding, Crossbreeding and Selection in Evolution. Proc. Sixth Int. Congr. Genet. 1: 356–366.

Yaari G., M. Uduman, and S. H. Kleinstein, 2012 Quantifying selection in high-throughput Immunoglobulin sequencing data sets. Nucleic Acids Res. gks457. https://doi.org/10.1093/nar/gks457

bioRxiv preprint doi: https://doi.org/10.1101/558825; this version posted February 25, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Yaari G., J. A. Vander Heiden, M. Uduman, D. Gadala-Maria, N. Gupta, *et al.*, 2013 Models of Somatic Hypermutation Targeting and Substitution Based on Synonymous Mutations from High-Throughput Immunoglobulin Sequencing Data. Front. Immunol. 4. https://doi.org/10.3389/fimmu.2013.00358

Yaari G., J. I. C. Benichou, J. A. Vander Heiden, S. H. Kleinstein, and Y. Louzoun, 2015 The mutation patterns in B-cell immunoglobulin receptors reflect the influence of selection acting at multiple time-scales. Philos. Trans. R. Soc. B Biol. Sci. 370: 20140242. https://doi.org/10.1098/rstb.2014.0242

Yaari G., and S. H. Kleinstein, 2015 Practical guidelines for B-cell receptor repertoire sequencing analysis. Genome Med. 7. https://doi.org/10.1186/s13073-015-0243-2

Yang Z., 1994 Estimating the pattern of nucleotide substitution. J. Mol. Evol. 39: 105–111.