Human B cell lineages associated with germinal centers following influenza vaccination are measurably evolving

Kenneth B Hoehn[1], Jackson S Turner[2], Frederick I Miller[3], Ruoyi Jiang[4], Oliver G Pybus[5], Ali H Ellebedy[2,6], Steven H Kleinstein[1,4,7]*

[1]Department of Pathology, Yale School of Medicine, New Haven, United States; [2]Department of Pathology and Immunology, Washington University School of Medicine, St Louis, United States; [3]Worcester Polytechnic Institute, Worcester, United States; [4]Department of Immunobiology, Yale School of Medicine, New Haven, United States; [5]Department of Zoology, University of Oxford, Oxford, United Kingdom; [6]The Andrew M. and Jane M. Bursky Center for Human Immunology and Immunotherapy Programs, Washington University School of Medicine, St Louis, United States; [7]Interdepartmental Program in Computational Biology & Bioinformatics, Yale University, New Haven, United States

*For correspondence:
steven.kleinstein@yale.edu

Abstract The poor efficacy of seasonal influenza virus vaccines is often attributed to pre-existing immunity interfering with the persistence and maturation of vaccine-induced B cell responses. We previously showed that a subset of vaccine-induced B cell lineages are recruited into germinal centers (GCs) following vaccination, suggesting that affinity maturation of these lineages against vaccine antigens can occur. However, it remains to be determined whether seasonal influenza vaccination stimulates additional evolution of vaccine-specific lineages, and previous work has found no significant increase in somatic hypermutation among influenza-binding lineages sampled from the blood following seasonal vaccination in humans. Here, we investigate this issue using a phylogenetic test of measurable immunoglobulin sequence evolution. We first validate this test through simulations and survey measurable evolution across multiple conditions. We find significant heterogeneity in measurable B cell evolution across conditions, with enrichment in primary response conditions such as HIV infection and early childhood development. We then show that measurable evolution following influenza vaccination is highly compartmentalized: while lineages in the blood are rarely measurably evolving following influenza vaccination, lineages containing GC B cells are frequently measurably evolving. Many of these lineages appear to derive from memory B cells. We conclude from these findings that seasonal influenza virus vaccination can stimulate additional evolution of responding B cell lineages, and imply that the poor efficacy of seasonal influenza vaccination is not due to a complete inhibition of vaccine-specific B cell evolution.

Editor's evaluation

The manuscript by Hoehn et al., introduces a novel approach to measure evolution in B cell responses, and apply it to a wide variety of data sets. The work provides significant new insight into which stimuli induce effective immune responses, and which has the potential to improve vaccine design. This will be of interest to those interested in B cell responses, especially in the case of vaccinations that induce poor immune responses.

Introduction

Measurably evolving populations are systems that undergo evolution rapidly enough for significant genetic differences to be detected in longitudinally sampled timepoints (*Drummond et al., 2003*). While this concept is frequently applied to viruses such as HIV (*Rambaut et al., 2004*) and SARS-CoV-2 (e.g., *du Plessis et al., 2021*), B cells experience similarly rapid evolution during affinity maturation. B cell affinity maturation is critical for developing high-affinity antibodies in response to infection and vaccination (*Shlomchik and Weisel, 2012*; *Victora and Nussenzweig, 2012*). During affinity maturation, somatic hypermutation (SHM) introduces mutations into the B cell receptor (BCR) loci at a rate orders of magnitude higher than the background rate of somatic mutations (*McKean et al., 1984*; *Murphy et al., 2008*). These modified BCRs are selected based on their binding affinity, and the process repeats cyclically within germinal centers (GCs *Teng and Papavasiliou, 2007*; *Victora and Nussenzweig, 2012*). Infection or vaccination can also stimulate pre-existing memory B cells that rapidly differentiate into antibody secreting plasmablasts or possibly re-enter GCs to undergo additional affinity maturation (*Ellebedy, 2018*; *Mesin et al., 2020*). A lack of vaccine-specific affinity maturation is thought to underlie the poor efficacy of seasonal influenza virus vaccination (*Arevalo et al., 2020*; *Ellebedy, 2018*). While recent work has shown that antigen-specific B cell lineages can be recruited into GCs following influenza vaccination (*Turner et al., 2020*), other work has been unable to detect significant increases in SHM frequency among circulating influenza-binding antibody lineages following vaccination (*Ellebedy et al., 2016*).

Whether seasonal influenza vaccination stimulates an increase in SHM frequency can be answered by determining whether influenza-binding B cell lineages found in GCs are measurably evolving following vaccination. This is distinct from simply quantifying SHM frequency. While influenza vaccination stimulates memory B cell lineages with high SHM frequency (*Laserson et al., 2014*; *Wrammert et al., 2008*), these lineages are only measurably evolving if their level of SHM detectably increases during the sampling interval surrounding vaccination. In this study, we show how a phylogenetic test of measurable evolution can be a powerful tool to detect increasing SHM frequency in longitudinally sampled BCR sequence datasets (*Duchêne et al., 2015*; *Murray et al., 2016*). We validate this approach through simulations and a survey of measurable evolution in B cell repertoires across a wide range of infections and vaccinations. We document significant heterogeneity among conditions, with some like HIV infection and primary hepatitis B vaccination enriched for measurably evolving lineages in the blood. We further show that while most circulating lineages following influenza virus vaccination are not measurably evolving, a subset of memory B cell lineages re-enter GCs and increase in SHM frequency.

Results

Detecting measurable evolution in longitudinally sampled BCR repertoires

We develop a framework to test for measurable evolution in B cells based on longitudinally sampled sequence data from the BCR variable region. After preprocessing the sequencing data, we first identify clonal lineages – B cells that descend from a common V(D)J rearrangement – using clustering based on nucleotide sequence similarity, which we have previously shown detects clonal relationships with high confidence (*Gupta et al., 2017*; *Zhou and Kleinstein, 2019*). The pattern of shared SHM among BCR sequences within a lineage is then used to build a B cell lineage tree, which represents a lineage's history of SHM. Branch lengths within these trees represent SHM per site. The divergence of each tip is the sum of branch lengths leading back to the lineage's most recent common ancestor. In evolving lineages, sequences sampled at later timepoints are expected to have higher divergence than those from earlier timepoints (*Figure 1A*). To estimate the rate of evolution over time, we calculate the slope of the regression line between timepoint (weeks) and divergence (SHM/site) for each tip (*Figure 1B, E*; *Rambaut et al., 2016*). Because tips are not independent, standard linear regression p values are improper. We instead quantify significance using a modified phylogenetic date randomization test (*Duchêne et al., 2015*; *Murray et al., 2016*). This tests whether the Pearson's correlation between divergence and time is significantly greater than that observed in the same tree with timepoints randomized among tips (*Figure 1C, F*). To account for population structure and sequencing error, we permute timepoints among single-timepoint monophyletic clusters of tips rather than

eLife digest When the immune system encounters a disease-causing pathogen, it releases antibodies that can bind to specific regions of the bacterium or virus and help to clear the infection. These proteins are generated by B cells which, upon detecting the pathogen, can begin to mutate and alter the structure of the antibody they produce: the better the antibody is at binding to the pathogen, the more likely the B cell is to survive. This process of evolution produces B cells that make more effective antibodies. After the infection, some of these cells become 'memory B cells' which can be stimulated in to action when the pathogen invades again.

Many vaccines also depend on this process to trigger the production of memory B cells that can fight off a specific disease-causing agent. However, it is unclear to what extent memory B cells that already exist are able to continue to evolve and modify their antibodies. This is particularly important for the flu vaccine, as the virus that causes influenza rapidly mutates. To provide high levels of protection, the memory B cells formed following the vaccine may therefore need to evolve to make different antibodies that recognize mutated forms of the virus.

It is thought that the low effectiveness of the flu vaccine is partially because the response it triggers does not stimulate additional evolution of memory B cells. To test this theory, Hoehn et al. developed a computational method that can detect the evolution of B cells over time. The tool was applied to samples collected from the blood and lymph nodes (organ where immune cells reside) of people who recently received the flu vaccine. The results were then compared to B cells taken from people after different infections, vaccinations, and other conditions.

Hoehn et al. found the degree to which B cells evolve varies significantly between conditions. For example, B cells produced during chronic HIV infections frequently evolved over time, while such evolution was rarely observed during the autoimmune disease myasthenia gravis. The analysis also showed that memory B cells produced by the flu vaccine were able to evolve if recruited to the lymph nodes, but this was rarely detected in B cells in the blood.

These findings suggest the low efficacy of the flu vaccine is not due to a complete lack of B cell evolution, but likely due to other factors. For instance, it is possible the evolutionary process it stimulates is not as robust as in other conditions, or is less likely to produce long-lived B cells that release antibodies. More research is needed to explore these ideas and could lead to the development of more effective flu vaccines.

individual tips (*Figure 1—figure supplements 1 and 2*; *Duchêne et al., 2015*; *Murray et al., 2016*). Further, it is possible that the combined effects of PCR and sequencing error will generate tree structures with multiple spurious tips radiating from a single node. This could increase the error rate of the date randomization test. Because trees are strictly binary, this would produce clusters of zero-length branches (soft polytomies) that could increase the error rate. To limit potential effects of this source of error, we resolve polytomies into the fewest number of single-timepoint monophyletic clades possible (*Figure 1—figure supplements 1 and 2*). We refer to lineages with a date randomization test p < 0.05 as 'measurably evolving'. To limit our analyses to lineages with adequate statistical power, we include only lineages with ≥15 total sequences sampled over at least 3 weeks, and have a minimum possible p value <0.05 based on the number of distinct permutations. Because we use a p value cutoff of 0.05, we expect a false positive rate of approximately 5% if no measurable evolution is occurring. We therefore refer to datasets with >5% measurably evolving lineages as 'enriched' for measurable evolution. This test is implemented within the Immcantation.org framework in the R package *dowser* (*Hoehn et al., 2020*).

To determine the necessary sampling interval to detect B cell evolution, we benchmarked the date randomization test using affinity maturation simulations performed with the package *bcr-phylo* (*Davidsen and Matsen, 2018*; *Ralph and Matsen, 2020*). This simulates alternating GC cycles of B cell proliferation, SHM, and selection based on amino acid similarity to a target sequence. Within these simulations, each lineage was first sampled after 10 simulated GC cycles, and then sampled a second time after a variable number of additional cycles. Using this framework in which all lineages are evolving, the date randomization test detected measurable evolution in 47% of lineages after 10 additional GC cycles, and 77% after 15 additional cycles (*Figure 1G*). Given a GC cycle time of
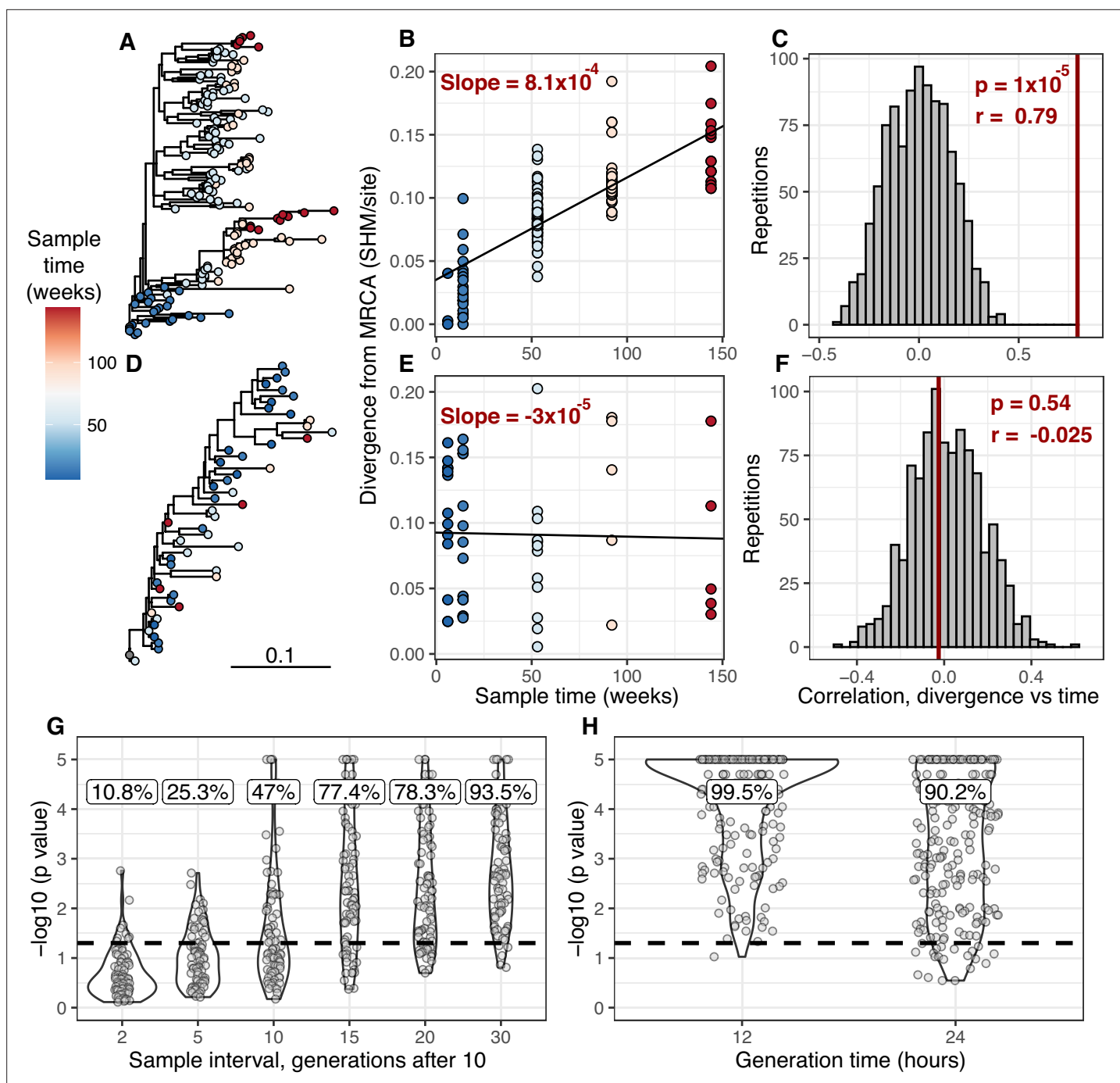
Figure 1. Detecting measurable evolution in B cell lineages. (**A**) Example B cell lineage tree from *Liao et al., 2013* showing increasing divergence with sample time. Branch lengths show somatic hypermutation (SHM)/site according to scale bar in (**D**). (**B**) Rate of SHM accumulation over time estimated using a regression of divergence vs time in tree (**A**). (**C**) Significance of the relationship between divergence and time estimated using a date randomization test comparing the Pearson's correlation (*r*) between divergence and time in tree (**A**). (**D–F**) Same plots as (**A–C**) but on a tree that is not measurably evolving. (**G**) Simulation-based power analysis shows the permutation test has high power over an interval of at least 10–30 GC cycles (generations). Lineages were sampled once at generation 10, and a second time after the specified number of additional generations have elapsed. Percentage of lineages with p < 0.05 are listed above, rounded to three significant digits. The dotted line corresponds to p = 0.05. (**H**) Simulation-based analysis reproducing the sampling of *Laserson et al., 2014* shows the test has high power even at slow (24 hr) GC cycle times.

The online version of this article includes the following figure supplement(s) for figure 1:

Figure supplement 1. Clustered date randomization and resolution of polytomies.

Figure supplement 2. Comparison of date randomization strategies.

Figure 1 continued on next page

Figure 1 continued

Figure supplement 3. Simulation-based power analysis.

Figure supplement 4. Simulation-based power analysis recreating experimental sampling design.

Figure supplement 5. Mean divergence of lineages from simulations in *Figure 1—figure supplement 3* under neutral evolution (selection = 0) or strong selection (selection = 1).

6–24 hr, 15 cycles corresponds to 4–15 days, within the timeframe of many longitudinal B cell repertoire studies (*Ellebedy et al., 2016*; *Laserson et al., 2014*). Interestingly, the date randomization test had higher power to detect measurable evolution in simulations of neutral evolution than those that included selection (*Figure 1—figure supplements 3–4*). This is likely because selection can reduce the rate of divergence within lineages compared to neutral evolution (*Figure 1—figure supplement 5*). To quantify the false positive rate, we repeated these calculations on the same simulations but with randomized sample time associations. Here, the date randomization test found measurable evolution in <4% in each case, indicating a low false positive rate (*Figure 1—figure supplements 3–4*). These analyses demonstrate that the date randomization test has sufficient sensitivity and specificity to detect ongoing B cell evolution from longitudinally sampled BCR data.

Primary immune responses are enriched for measurably evolving lineages

To further validate our approach, we tested for measurable evolution in cases of known or suspected affinity maturation in humans. We hypothesized that primary immune responses would be enriched for measurably evolving lineages. To test this, we used publicly available data primarily from the Observed Antibody Space (OAS) database (*Kovaltsuk et al., 2018*) to survey measurable evolution in BCR datasets from 99 human subjects in 21 studies spanning 10 conditions including HIV infection, Ebola virus infection, and healthy controls (*Table 1*). We observed considerable heterogeneity in measurable evolution among conditions. Confirming our hypothesis, we observed an enrichment of measurably evolving lineages (>5% of tested lineages) in primary immune responses including HIV infection, meningococcus vaccination, primary but not secondary hepatitis B vaccination, and early childhood development (*Table 1*, *Figure 2A*, and *Figure 2—figure supplement 1*).

Chronic HIV infection stimulates ongoing affinity maturation as B cells evolve to contain viral escape mutants (*Liao et al., 2013*; *Wendel et al., 2020*). Consistent with this arms race, HIV infection was more enriched for measurably evolving lineages than other conditions surveyed, with each study having between 5.9% and 53% of lineages measurably evolving (*Figure 2A*). Lineages from subjects with broadly neutralizing anti-HIV lineages sampled over multiple years (*Doria-Rose et al., 2014*; *Landais et al., 2017*; *Liao et al., 2013*; *Wu et al., 2015*) were particularly enriched (26–53% measurably evolving). Importantly, the HIV studies included were sampled over longer time periods than studies of other conditions (mean = 225 vs 45 weeks, *Table 1*). To determine whether these results were simply due to longer sampling intervals, we repeated our analysis of subjects with HIV using only samples within the first 60 weeks of the study. These truncated datasets were still highly enriched for measurably evolving lineages (6.9–64%) compared to other non-HIV datasets with similar sampling intervals (0–7.2%, *Figure 2A*). This indicates that the observed high frequency of measurably evolving lineages is not simply due to long sampling intervals.

Other primary immune responses were also enriched for measurably evolving lineages (*Table 1*, *Figure 2A*). B cell lineages from healthy children sampled during the first 3 years of life were enriched for measurable evolution (14%), possibly reflecting continual exposure to novel antigens (*Nielsen et al., 2019*). We also observed an enrichment of measurably evolving lineages following primary meningococcus vaccination (10%; *Galson et al., 2015a*) and primary but not secondary hepatitis B vaccination (7.2% vs 2.9%, respectively; *Galson et al., 2016*; *Galson et al., 2015b*). Primary hepatitis B vaccinees were sampled over a longer time period than secondary vaccines, so this difference may also be due to different sampling intervals (*Figure 2—figure supplement 2*). Further, allergen-specific immunotherapy, which stimulates tolerance of allergy-causing antigens through exposure, was also enriched for measurable evolution (6.5%; *Levin et al., 2016*). Interestingly, Ebola virus infection showed a borderline (5%) percentage of measurably evolving lineages (*Table 1*) despite likely being a

Table 1. Summary of datasets.

N shows number of subjects with at least one powered lineage. *Mean range* shows mean total sampling interval across subjects. *Powered lineages* shows the number of lineages that: (1) contained at least 15 sequences, (2) were sampled over at least 3 weeks, and (3) had a minimum possible p value <0.05. The rightmost column shows the percentage of these lineages with p < 0.05, rounded to two significant digits. Studies with at least 5% positive lineages are shown in bold. *Turner et al., 2020* in this table and *Figure 2* included only blood samples. Data from studies marked with an asterisk (*) were obtained from Observed Antibody Space (*Kovaltsuk et al., 2018*).

Study	Condition	N	Mean range (weeks)	Mean sample count	Multi-timepoint lineages	Powered lineages	% lineages p < 0.05
Levin et al., 2016*	Allergy + SIT	9	52	2.7	42	31	6.5
Davis et al., 2019*	Ebola virus	4	36	3.6	1,549	877	5
Wang et al., 2014	Healthy adults	7	52	2	18	10	0
Nielsen et al., 2019	Healthy children	20	69	2.7	262	71	14
Galson et al., 2015b*	Hep. B vaccine (boost)	9	4	4.8	4,923	3,422	2.9
Galson et al., 2016*	Hep. B vaccine (primary)	9	23	6.9	4,426	2,529	7.2
Doria-Rose et al., 2014*		1	190	8	65	48	44
Huang et al., 2016*		1	120	12	388	221	5.9
Johnson et al., 2018*		1	170	5	561	330	23
Landais et al., 2017*		1	160	7	1,084	743	48
Liao et al., 2013*		1	140	5	205	151	53
Schanz et al., 2014*		1	120	3	147	54	11
Setliff et al., 2018*		6	170	3	787	173	9.8
Wu et al., 2011, Wu et al., 2015*	HIV	1	730	7	393	305	26
Ellebedy et al., 2016*		8	13	5	1966	1,479	5.2
Laserson et al., 2014*		3	4	9	1,182	639	4.9
Turner et al., 2020	Influenza vaccine	1	8.6	5	168	104	2.9
Galson et al., 2015a*	Meningococcus vaccine	7	4	3	483	80	10

Table 1 continued on next page

Table 1 continued

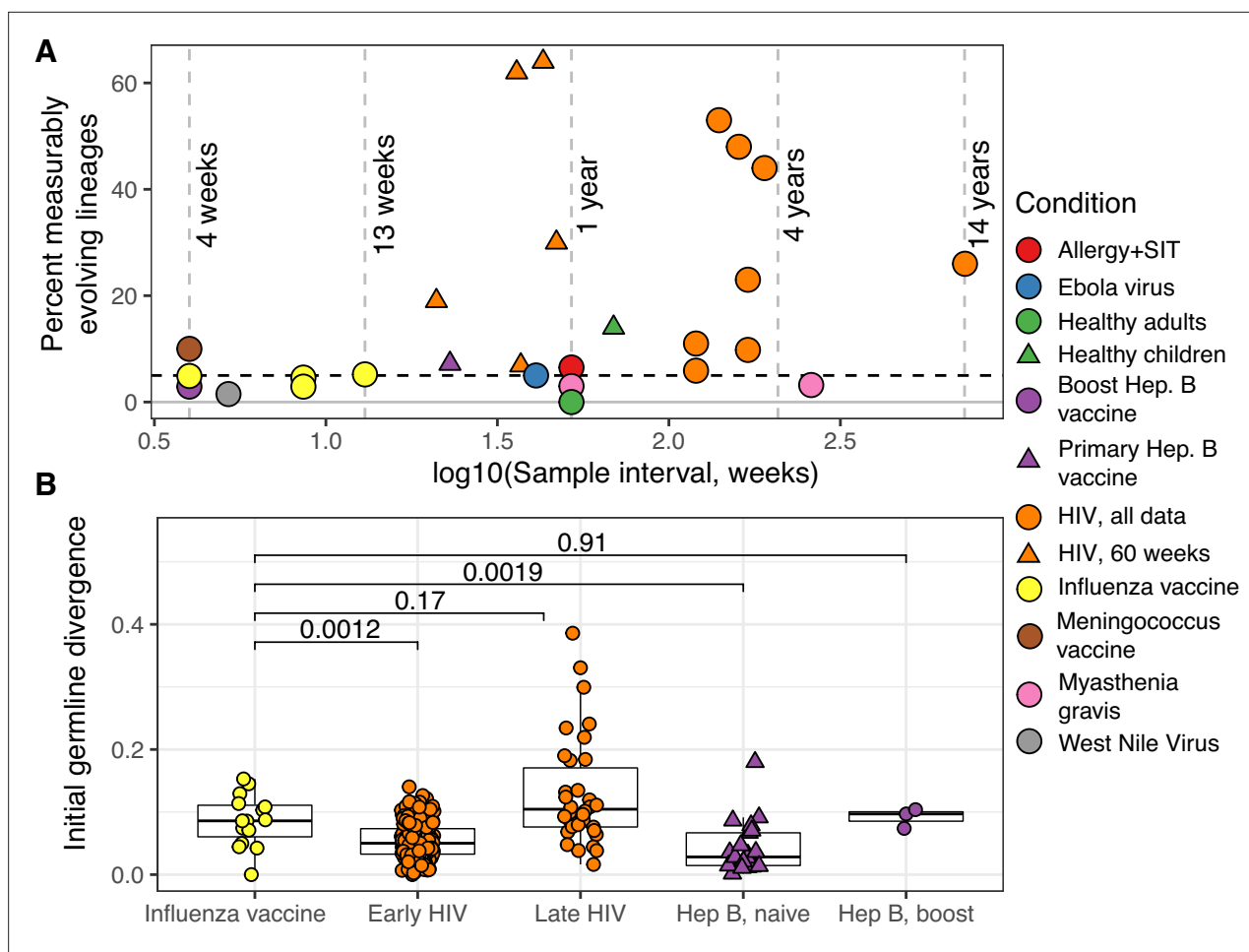| Study	Condition	N	Mean range (weeks)	Mean sample count	Multi-timepoint lineages	Powered lineages	% lineages p < 0.05
Jiang et al., 2020a		3	260	3.6	110	62	3.2
Jiang et al., 2020b	Myasthenia gravis	1	52	2	46	33	3
Tsioris et al., 2015	West Nile virus	6	5.2	2	151	65	1.5

Figure 2. Measurable evolution in B cell lineages across time and conditions. (**A**) Percentage of lineages that are measurably evolving within each study (Table 1, **Figure 1C**). The dotted line indicates 5%, the percent expected under the null hypothesis that there is no measurable evolution occurring in a given dataset. Orange triangles indicate HIV datasets truncated to only include data within the first 60-week sampling interval. Note that three HIV studies were not truncated because they contained <2 sample timepoints within the first 60 weeks of sampling (**Huang et al., 2016**; **Schanz et al., 2014**; **Wu et al., 2015**). (**B**) Mean initial germline divergence (sum of branch lengths) from germline to sequences from each adjusted measurably evolving lineage's first timepoint. Note that 'Early/Late' HIV in (**B**) separates studies by time since initial infection, while 'HIV, first 60 weeks' in (**A**) includes only samples taken over the first 60 weeks of sampling. Each point is a measurably evolving lineage with a Benjamini–Hochberg adjusted p value <0.1. Wilcoxon tests were used to compare divergence levels among datasets.

The online version of this article includes the following figure supplement(s) for figure 2:

Figure supplement 1. Date randomization p value histograms from blood-derived lineages across all studies.

Figure supplement 2. Comparison of measurably evolving lineages among hepatitis B vaccine schedules.

Figure supplement 3. Enrichment of antigen-binding monoclonal antibody (mAb) sequences across studies.

Figure supplement 4. Initial germline divergence with alternate p value thresholds.

primary infection. Overall, however, these results confirm that the date randomization test can detect ongoing SHM in empirical datasets where it is expected to be occurring.

We next investigated whether measurable evolution was associated with antigen-binding lineages. While antigen-binding information was not available for most B cell lineages surveyed, some studies included experimentally validated monoclonal antibody sequences (mAbs). Lineages containing these sequences thus contain B cells that bind to the antigen under study. Experimentally validated mAbs were included from six studies: four in HIV (**Doria-Rose et al., 2014**; **Landais et al., 2017**; **Liao et al., 2013**; **Wu et al., 2015**), one in Ebola virus infection (**Davis et al., 2019**), and one in influenza vaccine response (**Turner et al., 2020**). We found that across these studies measurably evolving lineages were more likely to contain mAbs than nonmeasurably evolving lineages (p = 0.031, Wilcoxon test,

Figure 2—figure supplement 3). This is consistent with the hypothesis that measurably evolving lineages are actively responding to antigens relevant to the condition being studied.

Measurably evolving lineages are rare in peripheral blood following influenza vaccination

Seasonal influenza vaccination is believed to trigger a memory B cell response in adults. If memory B cells rarely re-enter GCs to undergo additional affinity maturation (*Mesin et al., 2020*), and there is little evolution of naive B cell lineages, we expect little measurable evolution in the blood following vaccination. To test this, we applied the date randomization test to three longitudinally sampled adult influenza vaccine datasets. The first comprised three adults sampled seven times between 1 hr and 28 days postvaccination (*Gupta et al., 2017*; *Laserson et al., 2014*); the second contained eight adults sampled five times between 0 and 90 days postvaccination (*Ellebedy et al., 2016*) the third used blood samples from a single individual sampled five times between 0 and 60 days postvaccination (*Turner et al., 2020*). Across subjects in each study, between only 2.9% and 5.2% of lineages were measurably evolving (*Table 1*). These values are approximately as expected under the null hypothesis of no measurable evolution, and histograms of p values from these datasets are roughly uniform, suggesting the measurably evolving lineages identified are mostly false positives from multiple testing (*Figure 2—figure supplement 1*). Distributions of p values for all datasets are also available in *Figure 2—figure supplement 1*. To verify the 4- to 13-week sampling range of these studies was sufficient to detect measurable evolution, we performed simulation analyses replicating the sampling strategy of the influenza dataset with the shortest sampling range (*Figure 1—figure supplement 4*; *Laserson et al., 2014*). These simulations show this timescale was sufficiently long to detect ongoing affinity maturation with high sensitivity (>90%, *Figure 1H*). Overall, these results indicate B cell lineages present in blood infrequently undergo additional evolution within 13 weeks following influenza vaccination, consistent with a primarily GC-independent memory B cell response and/or rarity of antigen-specific lineages in the peripheral blood (*Wrammert et al., 2008*).

Measurably evolving lineages following influenza vaccination include memory B cell origin

While measurably evolving lineages do not occur at high frequency in the blood following influenza vaccination, we checked if any could be identified after adjustment for multiple testing. To adjust for multiple hypothesis tests, we pooled lineages across all studies and adjusted their p values using the Benjamini–Hochberg procedure (*Benjamini and Hochberg, 1995*). We identified 15 lineages in influenza datasets, and 354 lineages in other conditions, with adjusted date randomization p values < 0.1. We investigated if these 'adjusted' measurably evolving lineages were derived from naive or pre-existing memory B cells. Because memory B cell lineages have already undergone affinity maturation, they are expected to have higher initial SHM levels compared to naive B cell lineages. To test this, we compared germline sequence divergence in adjusted measurably evolving lineages from influenza vaccination to other conditions. Consistent with memory B cell reactivation, lineages from influenza vaccination had significantly higher initial divergence (median = 8.6%) than those from primary responses such as early HIV infection (median = 5%, p = 0.0012) and primary hepatitis B vaccination (median = 2.8%, p = 0.0019) (*Figure 2B*). Further, these influenza lineages had initial divergence levels similar to lineages from subjects with HIV first sampled >5 years after infection (*Huang et al., 2016*; *Wu et al., 2015*), and hepatitis B booster vaccination subjects (*Figure 2B*; *Galson et al., 2015b*). Ebola virus infection, meningococcus vaccination, and early childhood development had median initial divergence levels of 0.4%, 6.6%, and 2.0%, respectively, but contained less than three adjusted measurably evolving lineages each. To understand the effect of multiple hypothesis correction on these results, we repeated the comparisons in *Figure 2B* using all measurably evolving lineages (unadjusted p < 0.05) from the same datasets. Considering this larger set of lineages, initial divergence of lineages from influenza vaccination studies was significantly higher than those in all other conditions except late HIV infection (*Figure 2—figure supplement 4*). The same pattern from *Figure 2B* was also found when repeating these comparisons with a more strict cutoff (adjusted p < 0.05, *Figure 2—figure supplement 4*). Overall, these results are consistent with measurably evolving lineages from influenza vaccination arising mainly from pre-existing memory B cells.

Measurably evolving lineages show signs of purifying selection

We next investigated the type and degree of selection operating on measurably evolving B cell lineages. One way to detect natural selection in DNA sequences is to estimate the ratio of nonsynonymous (amino acid replacement) to synonymous (silent) mutation rates. This ratio is often called ω (**Nielsen and Yang, 1998**). Neutral evolution, where amino acid replacements are not selected for or against, should result in $\omega = 1$. Purifying selection, where amino acid replacements are disfavored, should result in $\omega < 1$. Diversifying selection, where amino acid replacements are favored, should result in $\omega > 1$. In B cell lineages, ω is often estimated separately for complementarity-determining regions (CDRs) involved in antigen binding, and framework regions (FWRs), which are more structural. Further, it is important to estimate ω or similar metrics using models that account for intrinsic hot- and cold-spot biases of SHM (**Hoehn et al., 2017**; **Uduman et al., 2011**; **Yaari et al., 2012**). To understand what kind of selection operated on measurably evolving lineages, we estimated separate ω values for CDR and FWR regions (ω_{CDR} and ω_{FWR}) of the adjusted measurably evolving lineages (**Figure 2B**) using the HLP19 model in IgPhyML (**Hoehn et al., 2019**). Model parameters were shared among lineages within the same subject, and only subjects with at least two adjusted measurably evolving lineage were included to reduce noise. Across all conditions surveyed, we found evidence of purifying selection operating on adjusted measurably evolving lineages (mean $\omega_{CDR} = 0.58$, mean $\omega_{FWR} = 0.48$, **Table 2**). We estimated the significance of these results using a phylogenetic likelihood ratio test (**Huelsenbeck and Rannala, 1997**). We found that ω_{CDR} was significantly <1 in 10/13 subjects (significantly >1 in none) and ω_{FWR} was significantly <1 in 13/13 subjects (**Table 2**). This signal of purifying selection was particularly strong in both early and late HIV. Influenza vaccination showed higher ω values, comparable to primary hepatitis B vaccination.

Influenza-binding lineages associated with GCs are measurably evolving

While we found little measurable evolution in the blood following seasonal influenza vaccination, influenza vaccination has been shown to stimulate both naive and memory B cells to enter GCs (**Turner et al., 2020**). This raises the possibility that additional affinity maturation could be occurring in GCs, but its products are not enriched in the blood. Data from **Turner et al., 2020** provided both blood samples and fine-needle aspirations of lymph nodes (including GCs) from the same subject. By combining these samples, we identified 53 powered B cell lineages containing at least one GC B cell following influenza vaccination, and 100 powered lineages that contained none. We refer to lineages containing one or more GC B cells as 'GC-associated'. To determine whether GC-associated lineages were undergoing additional SHM, we tested whether they were enriched for measurable evolution. We found that 7.5% of lineages containing sequences from GC B cells were measurably evolving, compared to only 3.0% of lineages with no identified GC sequences. This signal of measurable evolution increased with the fraction of GC sequences. For instance, while 10% of lineages containing ≥10% GC sequences were measurably evolving, 38% (3/8) of those with ≥25% GC sequences were measurably evolving (**Figure 3A**). Lineages with higher proportions of GC sequences also had a higher correlation between divergence and time (linear regression slope = 1.1, p = 8.9 × 10^{-13}, **Figure 3—figure supplement 1**). We further estimated the significance of this positive relationship by bootstrapping our data using 10,000 resampling repetitions with replacement. We found that in all 10,000 resampling repetitions, the slope of the linear regression between GC sequence proportion and the correlation between divergence and time was positive, with 95% of repetitions having a slope between 0.81 and 1.3 (**Figure 3—figure supplement 1**). Measurably evolving lineages in this dataset did not contain significantly more sequences than other lineages, indicating these results were not significantly confounded by lineage size (**Figure 3—figure supplement 2**). Finally, the measurably evolving lineages with the highest proportion of GC sequences contained mAbs that bound to vaccine antigens (**Figure 3B, C**). These lineages show signs of origin from memory B cells, such as clonal relatedness to blood plasmablasts sampled 5 days postvaccination, and high mean germline divergence at their first sampled timepoint (6.3%, 7.2%, **Figure 3B, C**, respectively). To test whether GC-associated lineages accumulated new amino acid replacement mutations rather than just silent mutations, we repeated the date randomization test but calculated the divergence of each tip as the number of amino acid differences between that tip's sequence and the unmutated germline ancestor. This amino acid-based correlation analysis also showed a strong positive relationship between the

Table 2. Analysis of selection on adjusted measurably evolving lineages.

Repertoire-wide estimates of ω for CDRs (ω_{CDR}) and FWRs (ω_{FWR}) for adjusted measurably evolving lineages within different subjects are shown. L indicates the maximum log-likelihood obtained when both ω_{CDR} and ω_{FWR} were estimated by maximum likelihood. $L_{CDR=1}$ indicates the maximum log-likelihood obtained when ω_{FWR} was estimated by maximum likelihood but ω_{CDR} was fixed at 1. $L_{FWR=1}$ indicates the maximum log-likelihood obtained when ω_{CDR} was estimated by maximum likelihood but ω_{FWR} was fixed at 1. The likelihood ratio statistic (LRS) was calculated as either $2\times(L - L_{CDR=1})$ for CDRs or $2\times(L - L_{FWR=1})$ for FWRs, and p values were calculated using a likelihood ratio test with one degree of freedom (see Methods). L values were rounded to two decimal places, LRS values are reported to three significant digits. Significant p values are in bold. p values below the numerical limit for double values are reported as <2E-16.

Study	Subject	N	Region	ω	L	$L_{CDR=1}$	$L_{FWR=1}$	LRS	p
			CDR	0.624	−5753.04	−5759.65		13.2	**2.80E−04**
	Donor-4	4	FWR	0.503	−5753.04		−5783.9	61.7	**4.00E−15**
			CDR	0.979	−5074.01	−5074.02		0.0212	8.84E−01
Ellebedy et al., 2016	Donor-5	5	FWR	0.584	−5074.01		−5090.55	33.1	**8.75E−09**
			CDR	0.583	−9162.11	−9173.39		22.6	**1.99E−06**
Laserson et al., 2014	FV	4	FWR	0.508	−9162.11		−9217.38	111	**<2E−16**
			CDR	0.424	−18,976.96	−19,045.96		138	**<2E−16**
Doria-Rose et al., 2014	CAP256	9	FWR	0.451	−18,976.96		−19,151.29	349	**<2E−16**
			CDR	0.423	−46,720.28	−46,858.67		277	**<2E−16**
Johnson et al., 2018	CAP256	26	FWR	0.408	−46,720.28		−47,243.96	1,050	**<2E−16**
			CDR	0.39	−416,489.06	−418,035.9		3,090	**<2E−16**
Landais et al., 2017	PC064	188	FWR	0.386	−416,489.06		−422,043.68	11,100	**<2E−16**
			CDR	0.527	−102,131.66	−102,307.3		351	**<2E−16**
Liao et al., 2013	CH505	51	FWR	0.417	−102,131.66		−103,240.86	2,220	**<2E−16**
			CDR	0.657	−5573.47	−5577.12		7.29	**6.93E−03**
Schanz et al., 2014	ZA159	2	FWR	0.466	−5573.47		−5617.91	88.9	**<2E−16**

Influenza vaccine

Early HIV

Table 2 continued on next page

Table 2 continued

Study	Subject	N	Region	ω	L	$L_{CDR=1}$	$L_{FWR=1}$	LRS	p
Huang et al., 2016	Donor-Z258	2	CDR	0.345	−3162.12	−3172.87		21.5	3.54E−06
			FWR	0.401	−3162.12		−3196.5	68.7	1.11E−16
Wu et al., 2015 (Late HIV)	Donor-45	32	CDR	0.451	−63,527.81	−63,687		318	<2E−16
			FWR	0.378	−63,527.81		−64,444.73	1,830	<2E−16
	Subject-2277	6	CDR	0.885	−5498.77	−5499.15		0.766	3.81E−01
			FWR	0.668	−5498.77		−5508.36	19.2	1.18E−05
	Subject-2752	2	CDR	0.533	−1112.1	−1113.82		3.44	6.36E−02
			FWR	0.461	−1112.1		−1120.06	15.9	6.68E−05
Galson et al., 2016 (Hep B. vaccine, primary)	Subject-2954	9	CDR	0.711	−6023.13	−6026.4		6.56	1.04E−02
			FWR	0.545	−6023.13		−6048.08	49.9	1.62E−12
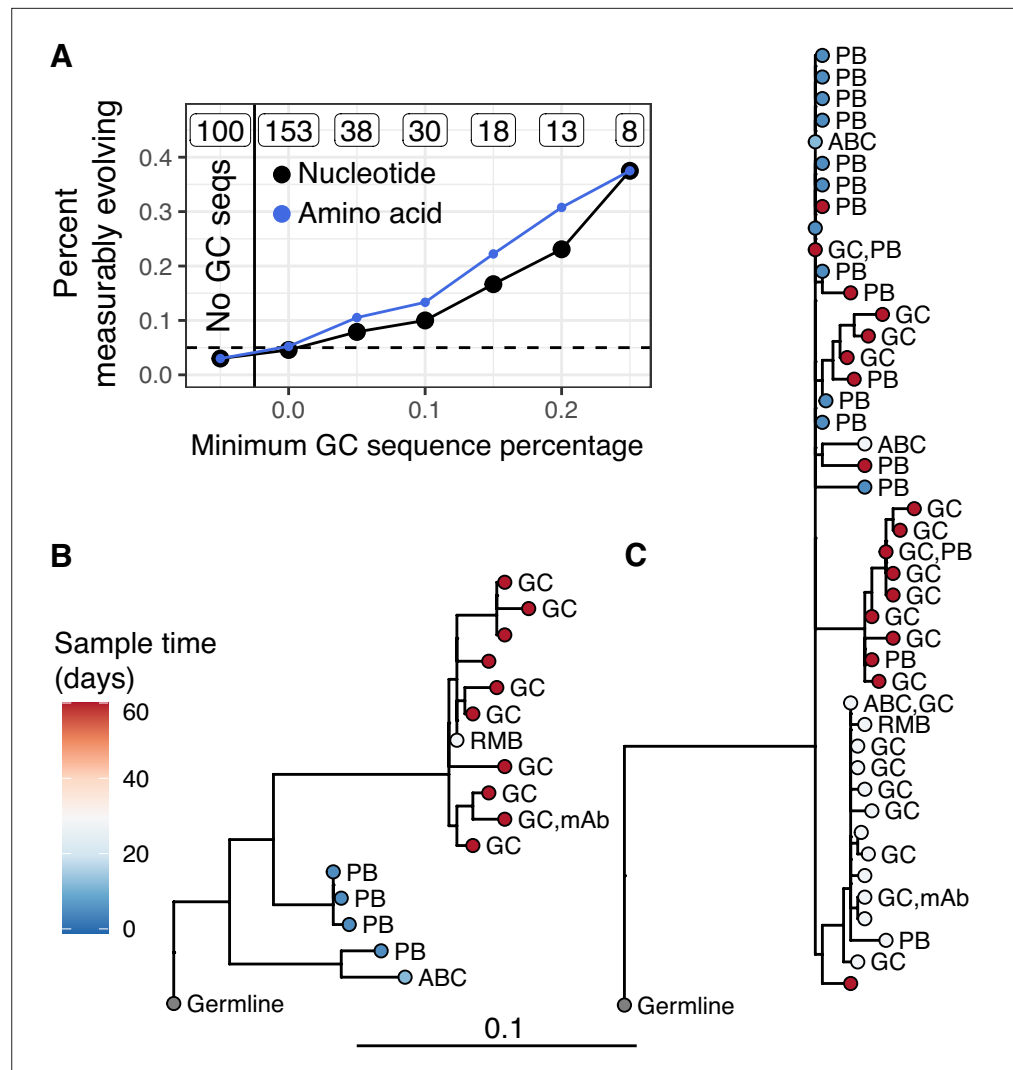
Figure 3. Germinal center (GC) association is positively related to measurable evolution following influenza vaccination. (**A**) Percent of lineages that are measurably evolving given a minimum percentage of GC sequences. The minimum (inclusive) percent of GC sequences within a clone is shown on the *x* axis. The origin shows the percentage of measurably evolving lineages across all lineages. The left-most point shows lineages without any GC sequences. The total number of lineages in each category are listed above each point. The dashed line shows 5%, the expected false positive rate under the null hypothesis. Results are shown for clustered date randomization tests using divergence values calculated either as the sum of nucleotide-based phylogenetic branch lengths (nucleotide), and the amino acid Hamming distance of each sequence to the germline (amino acid). (**B, C**) Lineage trees showing measurably evolving lineages with the highest proportion of GC sequences. Tips are labeled by cell type if available. ABC, activated B cell; GC, germinal center; PB, plasmablast; RMB, resting memory B; and unlabeled tips are from bulk PBMC sequencing. mAb = influenza-binding monoclonal antibody sequence (2018/2019 quadrivalent inactivated influenza virus vaccine). Branch lengths represent somatic hypermutation (SHM)/site, as shown by the shared scale bar.

The online version of this article includes the following figure supplement(s) for figure 3:

Figure supplement 1. Germinal center (GC) engagement is positively related to measurable evolution following influenza vaccination.

Figure supplement 2. Number of sequences per lineage in measurably evolving vs nonmeasurably evolving lineages.

Figure supplement 3. Percent of lineages that are measurably evolving given a minimum percentage of germinal center (GC) sequences.

proportion of lineages that were measurably evolving and the percentage of sequences derived from GC B cells (*Figure 3A*). This indicates that these GC-associated lineages accumulated new amino acid mutations as well as nucleotide mutations over the study interval. Overall, these analyses demonstrate that influenza-binding, GC-associated B cell lineages undergo additional, measurable evolution following vaccination.

A possible alternative explanation for measurable evolution following influenza vaccination is that SHM is not occurring over the sampled time interval, but that highly mutated B cells were preferentially recalled due to their higher binding affinity. Preferential recall of highly mutated B cells would likely result in a positive correlation between divergence and sample time. While difficult to directly test, we believe this explanation is unlikely to be the sole source of measurable evolution in our data. First, blood samples taken 5 days postvaccination represent the breadth of the pool of memory B cells. If measurable evolution was simply due to expansion of mutated memory B cells, we would expect divergence of later-sampled B cells to be within the range of day 5 plasmablasts. Instead, many later-sampled GC sequences are clearly more diverged than earlier-sampled sequences within the measurably evolving, influenza-binding lineages we observed (*Figure 3B, C*). Second, if measurable evolution were due simply to preferential expansion of more mutated B cells, we would expect to observe measurable evolution within influenza-binding lineages in both the blood and GC. This is not tested in *Figure 3A* because that analysis includes all lineages, not just those that bind to influenza. To adjust for this, we repeated the analysis in *Figure 3A* while only including lineages that contained influenza-binding mAbs. Even among influenza-binding lineages, we still observed an association between GC cells and measurable evolution. While 1/10 of influenza-binding lineages found only in the blood were measurably evolving, 2/5 lineages with >25% GC sequences were measurably evolving (*Figure 3—figure supplement 3*). Overall, while we cannot rule out preferential expansion of highly mutated memory B cells, these results are more easily interpretable as the result of ongoing SHM in GCs.

Discussion

The extent to which seasonal influenza vaccination stimulates affinity maturation against vaccine antigens is unclear, and poor efficacy of seasonal influenza virus vaccination is often attributed to stimulation of pre-existing memory B cells interfering with novel responses to vaccine antigens (*Ellebedy, 2018*). While a prior study has shown that influenza-binding B cell lineages are found in GCs following seasonal influenza vaccination (*Turner et al., 2020*), other work has suggested that circulating influenza-binding B cell lineages do not accumulate additional SHM following vaccination (*Ellebedy et al., 2016*). To determine whether seasonal influenza vaccination stimulates additional evolution in B cell lineages, we developed and validated a framework to detect measurable evolution using longitudinally sampled BCR sequencing data. This phylogenetic test can be a powerful tool to detect ongoing B cell evolution using longitudinally sampled BCR datasets across a wide array of immunological conditions, including influenza virus vaccine responses. Our results confirm prior findings that there is little evidence of B cell evolution among lineages sampled in the peripheral blood following seasonal influenza vaccination (*Ellebedy et al., 2016*). However, we also show that seasonal influenza vaccination is capable of stimulating measurable evolution in influenza-binding, GC-associated B cell lineages.

To place our analyses of the influenza vaccination response in a broader context, we surveyed measurable evolution across a broad range of infections and vaccinations. Prior work has shown that chronic HIV infection induces long-term affinity maturation of broadly neutralizing antibody lineages in response to viral escape mutants (*Liao et al., 2013*; *Vieira et al., 2018*; *Wu et al., 2015*). Our results show that HIV infection is associated with an exceptionally strong signature of B cell evolution over time. This signature is not limited to single lineages. Rather, a substantial fraction of longitudinally sampled B cell lineages within the repertoires of subjects with HIV are measurably evolving, consistent with clonal competition among B cell lineages during HIV infection (*Nourmohammad et al., 2019*). Early childhood development during the first 3 years of life (*Nielsen et al., 2019*) showed the second-highest enrichment of measurably evolving lineages among surveyed conditions. This possibly reflects continual exposure to novel antigens during childhood. Further, primary vaccinations (meningococcus, primary hepatitis B)(*Galson et al., 2016*; *Galson et al., 2015a*) showed stronger signatures of measurable evolution than secondary vaccinations (adult seasonal influenza, hepatitis B booster) (*Ellebedy*

et al., 2016; *Galson et al., 2015b*; *Laserson et al., 2014*; *Turner et al., 2020*). Overall, our results are consistent with the hypothesis that GC responses are stronger in response to novel antigens.

In addition to detecting measurable evolution, we also characterized selection operating on measurably evolving B cell lineages. We found that measurably evolving lineages showed evidence of purifying selection ($\omega_{CDR} < 1$) (*Table 2*). Though perhaps counterintuitive, a strong signal of purifying selection is a straightforward prediction of evolution toward an adaptive peak (*Hoehn et al., 2019*). Similar evidence of purifying selection during affinity maturation has been observed in other studies including influenza vaccination, HIV infection, and healthy controls (*Cizmeci et al., 2021*; *Hoehn et al., 2019*; *Sheng et al., 2016*; *Yaari et al., 2015*). Importantly, ω estimates are an average across all codons within CDRs or FWRs. It is possible that positive selection operated on a small number of codon sites, but that this signal was outweighed by the larger number of sites under purifying selection. Codon-specific models may be useful in future analyses to identify these sites under positive selection (e.g., *Yang et al., 2000*). When interpreting these results, it is also important to note that parameters were estimated using all mutations represented by each lineage tree, including those that potentially occurred before the first sampled timepoint. In all, these results are consistent with typical forces of selection having operated on measurably evolving lineages.

There are several limitations to this study. Data from different studies were sampled according to different schedules and time intervals. Because the power to detect measurable evolution should increase over time (*Figure 1G*), this could confound comparisons among datasets. We note however that multiple influenza and HIV datasets were surveyed, and enrichment of measurable evolution within these conditions was not strongly related to the sampling range (*Figure 2A*). This suggests immunological condition, rather than sample range, was the primary determinant of observed differences. By including monoclonal antibody (mAb) sequences with experimentally validated binding in several datasets, we were able to show that measurably evolving lineages are more likely than nonmeasurably evolving lineages to contain mAb sequences. This is consistent with the idea that measurably evolving lineages are actively responding to antigen. These results should be interpreted cautiously, however. With the exception of *Turner et al., 2020* only a small number mAb sequences were found in the lineages analyzed (mean = 3.6 per study). Further, lineages containing GC sequences were preferentially selected for mAb generation in *Turner et al., 2020*, which may artificially increase the likelihood that measurably evolving lineages contain mAbs. While intriguing and biologically plausible, conclusively determining whether measurable evolution predicts antigen binding is beyond the scope of this study. Another limitation is that maximum parsimony was used to estimate lineage tree topologies and branch lengths. While more sophisticated methods are available for inferring B cell lineage trees (e.g., *Hoehn et al., 2019*), maximum parsimony often has competitive performance for topology estimation (*Davidsen and Matsen, 2018*) and is faster than more complex maximum likelihood models designed for B cell lineages. Computational efficiency was particularly important as our analyses required constructing more than 20,000 lineages trees spanning approximately 1,100,000 BCR sequences. Finally, our analysis of GC-associated B cell lineages was limited to data sampled from a single subject. Thus, while the results here demonstrate that influenza vaccination is capable of inducing measurable evolution, it remains unclear whether this is a general feature of influenza vaccination.

Our analyses of measurable evolution involved a series of hypothesis tests, and the definition for 'enrichment' of measurable evolution (>5% of lineages) was chosen based on the expected false positive rate under the null hypothesis. This enrichment measure was chosen to compare the relative frequency of measurably evolving lineages among datasets. A lack of enrichment does not indicate a complete lack of measurably evolving lineages. Conversely, slight enrichment of measurably evolving lineages (~5%) should not be interpreted as proof of ongoing affinity maturation in a set of lineages. The analysis in *Figure 2A* was not intended to test the null hypothesis that no lineages are measurably evolving in a particular dataset. Because multiple studies were surveyed (*Table 1*), it is possible our results contain false positives. Indeed, while some conditions such as HIV have a strong signal of measurable evolution across multiple studies, others such as naive hepatitis B vaccination and allergen-specific immunotherapy are just above the significance threshold (≤7.2% lineages with p < 0.05) and are each represented by a single study (*Figure 2A*, *Table 1*). These latter datasets should be interpreted cautiously and with the understanding that the vast majority of their lineages were not measurably evolving. To limit the influence of false positives, a multiple testing correction (false

discovery rate < 0.1) was performed in analyses investigating the properties of measurably evolving lineages (*Figure 2B*, *Table 2*). We note, however, that repeating our analyses of initial germline divergence using all measurably evolving lineages (unadjusted p < 0.05) or a more strict p value cutoff (adjusted p < 0.05) yielded similar results (*Figure 2—figure supplement 4*), indicating our results are robust to the thresholds used. Finally, we validated the specificity of the date randomization test in empirical data (*Figure 1—figure supplement 2*). While a significant decrease in divergence over time is biologically unlikely, false positives due to multiple testing or sequencing error should produce a similar number of lineages with a significant correlation in either direction. We quantified correlation in either direction using a two-tailed version of the clustered, resolved date randomization test with a critical value of 0.025 (see Methods). Under a null hypothesis of no ongoing evolution, 2.5% of lineages were expected to have a significant negative correlation between divergence and time (all false positives). However, we found such 'negatively evolving' lineages at a mean frequency of only 1.2% (median = 1.3%, range = 0–2.8%) across all datasets when using clustered permutations and resolved polytomies (*Figure 1—figure supplement 2*). This indicates the date randomization test we used is conservative, and that our chosen thresholds are likely more strict than necessary.

Beyond multiple testing, there are some biological scenarios that could plausibly give rise to a signal of measurable evolution without additional SHM occurring during the sampling interval. For instance, it is possible that all SHM within a lineage occurred before the first sampled timepoint in a study, but that more mutated, higher-affinity BCRs were preferentially stimulated and sampled in later timepoints. Such a scenario would likely result in a positive correlation between divergence and sample time. However, our results are easier to explain if measurable B cell evolution results at least in part from ongoing SHM. For instance, our analysis of measurably evolving lineages following influenza vaccination showed (1) many later-sampled GC B cells had higher divergence than any sampled day 5 plasmablast sequence in the same lineage and (2) continued association between GC cells (rather than the blood) and measurable evolution even among influenza-binding lineages. Nonetheless, it is still theoretically possible that more mutated B cells were not sampled at early timepoints, and were preferentially expanded in GCs compared to the peripheral blood. Thus, while our interpretation is that measurable B cell evolution most likely represents an ongoing SHM process, we cannot conclusively rule out biased selection of more mutated sequences that were generated before the sampling interval.

Affinity maturation is a rapid evolutionary process. It is perhaps surprising that, while we identified conditions enriched for measurably evolving lineages, most lineages in circulation were not measurably evolving (*Table 1*). One explanation is that our analyses did not use sufficiently long sampling intervals to detect affinity maturation, though we believe this is unlikely. Studies in mice have estimated that SHM occurs at ~10^{-3} SHM/bp/division (*Kleinstein et al., 2003*; *McKean et al., 1984*), and that GC B cells cycle every 6–12 hr (*Allen et al., 2007*; *Hauser et al., 2007*; *Victora and Nussenzweig, 2012*). Simulations using conservative assumptions (strong selection, 24-hr cell cycle) and replicating the sample structure of our shortest-term influenza dataset (4 weeks), showed high power with >90% true positive rate (*Figure 1H*). Further, we found an enrichment of blood-derived measurably evolving lineages after only 4 weeks in one study (*Galson et al., 2015a*), and after 8 weeks corresponding to a known context of affinity maturation (GC entry, *Figure 3*). Overall, these results show that the sample times of our surveyed datasets should be sufficient to detect ongoing B cell evolution if it were occurring. However, it is possible that lineages may not remain in GCs continuously, which would slow the rate of evolution compared to our simulations. A more plausible explanation for the lack of measurable evolution is that most lineages in the blood are either nonspecific to the condition being studied, or derive from a GC-independent response (*Mesin et al., 2020*; *Takemori et al., 2014*; *Taylor et al., 2012*). It is also possible that lineages relevant to the condition being studied are inefficiently stimulated.

We find that seasonal influenza virus vaccination in young adults induces a GC reaction where maturation of vaccine-specific B cell lineages occurs, including those likely recruited from the pre-existing memory B cell compartment. These results imply that poor efficacy of seasonal influenza vaccination does not result from a complete lack of vaccine-induced B cell evolution. While we showed that B cells in these evolving lineages increased in amino acid replacement mutation frequency, it remains possible that this evolution is less able to select affinity-increasing mutations (*Hoehn et al., 2019*), that the overall number of evolving lineages is reduced, or that the products of this vaccine-induced

evolution are not efficiently translated into memory and long-lived plasma cells. These latter two explanations are consistent with the results of our survey of longitudinally sampled peripheral blood datasets, which found an enrichment of measurably evolving lineages in some primary immune response conditions, but not influenza vaccination. Future studies will be needed to fully test these hypotheses about the causes of poor efficacy of seasonal influenza vaccination.

Materials and methods

Study design

The goal of this study was to determine whether B cell lineages found in GCs following influenza vaccination evolved over a given sample interval. This necessitated describing and validating a test for measurable evolution from longitudinally sampled BCR sequencing data. Simulation-based power analyses determined that this date randomization test has sufficient sensitivity to detect evolving B cell populations over a sampling interval of approximately 2 weeks. To determine whether the date randomization test also worked on known examples of affinity maturation, all longitudinally sampled datasets hosted on OAS (as of 6/2020) were downloaded and tested. To cover as wide a variety of conditions as possible, these datasets were supplemented with processed, publicly available datasets from other prior studies. To ensure datasets were appropriately powered, datasets were only included if they contained at least 10 B cell lineages with at least 15 sequences sampled over 3 weeks and a minimum possible date randomization test p value <0.05. BCR data from blood and fine-needle aspirations following influenza vaccination were obtained from *Turner et al., 2020*.

BCR sequence datasets and preprocessing

All longitudinally sampled BCR repertoire datasets were publicly available and obtained both from primary publications and through the OAS database (antibodymap.org, accessed 6/2020; *Kovaltsuk et al., 2018*). Both assembled nucleotide sequences and deduplicated amino acid sequences were obtained from OAS. To reduce the effect of sequencing error in OAS datasets, only nucleotide sequences corresponding to an amino acid sequence with a multiplicity of at least two were included. Datasets obtained from OAS are labeled in *Table 1*. Raw sequence data obtained from *Nielsen et al., 2019* were preprocessed with pRESTO v0.5.13 (*Vander Heiden et al., 2014*). Quality control was performed by first removing all sequences with a Phred quality score <20, length <300 bp, or any missing ('N') nucleotides. The 3′ and 5′ ends of each read were matched to forward and constant region primers with a maximum error rate of 0.1. The region adjacent to the constant region primer was exactly matched to subisotype-specific internal constant region sequences. Only sequences with the same isotype predicted by their constant region primer and internal constant region sequence were retained. Identical reads within the same isotype were collapsed and sequences observed only once were discarded. All other datasets used processed BCR sequence data provided by the authors of their respective publications. Data from *Wang et al., 2014* were processed in *Hoehn et al., 2019*. Data from *Jiang et al., 2020b* used only blood samples.

BCR sequence processing, genotyping, and clonal clustering

Datasets were processed using the Immcantation framework (immcantation.org). V(D)J gene assignment on data obtained from *Nielsen et al., 2019* was performed using IgBLAST v1.13 (*Ye et al., 2013*) against the IMGT human germline reference database (*Giudicelli et al., 2005*) (IMGT/GENE-DB v3.1.24; retrieved August 3, 2019). V(D)J gene assignments and clonal cluster assignments were already available in all other non-OAS datasets and were retained. Nonproductively rearranged sequences were excluded. Using Change-O v1.0.0 (*Gupta et al., 2015*), the V and J genes of unmutated germline ancestors for each sequence were constructed with D segment and N/P regions masked by 'N' nucleotides. Sequence chimeras were filtered by removing any sequence with more than six mutations in any 10 nucleotide window. Individual immunoglobulin genotypes were computationally inferred using TIgGER v1.0.0 and used to finalize V(D)J annotations (*Gadala-Maria et al., 2015*). To infer clonal clusters, sequences were first partitioned based on common V and J gene annotations, and junction region length. Within these groups, sequences differing from one another by a specified Hamming distance threshold within the junction region were clustered into clones using single linkage hierarchical clustering (*Gupta et al., 2017*). The Hamming distance threshold was determined by finding

the local minimum of a bimodal distance to nearest sequence neighbor plot using SHazaM v1.0.2.999 (*Yaari et al., 2013*). In cases where automated threshold detection failed, usually because the distance to nearest neighbor distribution was not bimodal, the threshold was set to 0.1 and verified by manual inspection to ensure that a threshold of 0.1 was near a local minimum. Finally, the V and J genes of unmutated germline ancestors for each clone were constructed. Within these unmutated ancestral sequences, D segments and N/P regions were masked using ambiguous 'N' nucleotides.

Testing for measurable evolution

Testing for measurable evolution begins with building B cell lineage trees. Within each B cell clone, identical sequences or those differing only by ambiguous nucleotides were collapsed unless they were sampled at different timepoints. To reduce computational complexity, lineages were randomly down-sampled to at most 500 sequences each. B cell lineage tree topologies and branch lengths were estimated using maximum parsimony using the pratchet function of the R package *phangorn* v2.5.5 (*Schliep, 2011*). R packages *dowser* v0.0.3 (*Hoehn et al., 2020*), *alakazam* v1.0.2.999 (*Gupta et al., 2015*), and *ape* v5.4-1 (*Paradis et al., 2004*) were used for phylogenetic analysis. Trees were visualized using *ggtree* v2.4.2 (*Yu et al., 2016*), and other figures were generated using *ggplot2* v3.3.5 (*Wickham, 2016*) and *ggpubr* v0.4.0 (*Kassambara, 2020*). R v3.6.1 (*R Development Core Team, 2017*) was used for analysis of measurable evolution except for data from *Davis et al., 2019*. Due to technical upgrades, figure generation and selection analysis were performed using R v4.0.3, as well as *ape* v5.5, *phangorn* v2.7.1, *shazam* v1.1.0, *alakazam* v1.1.0, and *dowser* v0.1.0. Data from *Davis et al., 2019* were also analyzed using these updated packages.

To test for measurable evolution over time, we use a modified version of the previously described phylogenetic date randomization test (*Duchêne et al., 2015*; *Murray et al., 2016*) implemented in *dowser* v0.0.3 (*Hoehn et al., 2020*). Briefly, for a given tree the divergence of each tip was calculated as the sum of branch lengths leading to the tree's most recent common ancestor (MRCA). Only branches directly between a tip and the tree's most recent common ancestor were used to calculate divergence. We next calculated the Pearson's correlation between the divergence and sampling time of each tip. Measurably evolving lineages should show a positive correlation between divergence and time (*Figure 1A*). Divergence from the lineage's predicted unmutated ancestral sequence rather than the MRCA could also be used. Because all sequences relate to the unmutated ancestral sequence through the MRCA node, this would add a constant additional divergence to all sequences, resulting in the same correlation as when the MRCA is used. We next identified monophyletic clades containing only sequences from a single timepoint (here referred to as 'clusters'). We then randomly permuted sampling times among clusters, such that all sequences within each cluster had the same, randomly chosen timepoint. We next measured the correlation between divergence and time in this randomized tree, and repeated the process 100,000 times. We then estimated the p value that the observed correlation between divergence and time was no greater than expected from random distribution of times among clusters. This p value was calculated as the proportion of permutation replicates that had an equal or higher correlation than in the observed tree. We used a pseudocount of one for this calculation. The minimum possible p value for a lineage was calculated as one divided by possible number of distinct cluster permutations.

We modified the date randomization test to account for the high degree of topological uncertainty of many B cell lineage trees. More specifically, B cell lineage trees often contain large clusters of zero-length branches (soft polytomies) that represent high uncertainty in branching order (e.g., *Figure 1— figure supplement 1*). In bulk BCR data, these polytomies may be due to PCR error or sequencing error. If polytomies are resolved randomly into bifurcations, this can produce more single-timepoint monophyletic clades than necessary and lead to a high false positive rate of the date randomization test (*Figure 1—figure supplements 1 and 2*). To ensure this source of uncertainty did not increase the false positive rate of our analyses, we resolved bifurcations within each polytomy such that sequences from the same timepoint were grouped into the fewest possible number of single-timepoint monophyletic clades before performing permutations. While we do not have direct evidence that polytomies in B cell lineages trees are produced from PCR error, the fact that resolving them reduces the rate of lineages with a significant negative correlation between divergence and time (a biologically implausible result, *Figure 1—figure supplement 2*) suggests they are at least in part due to technical artifacts.

The clustered date randomization approach is more conservative than tests that permute tips uniformly (e.g., *Unterman et al., 2020*), but has been shown to be less biased if different subpopulations are sampled at each timepoint (*Murray et al., 2016*). To explore the effect of this modeling choice, we repeated the analyses in *Table 1* using two-tailed clustered and uniform date randomization tests (*Figure 1—figure supplement 2*). Two-tailed tests can identify lineages with a significant positive or negative correlation between divergence and time. This is useful because a significant negative correlation between divergence and time is biologically implausible and represents a likely false positive result. Due to multiple testing under an alpha value of 0.025, we expect no more than 2.5% of lineages to have a significant negative correlation from these two-tailed tests. We found the uniform permutation test had a high rate of negatively evolving lineages (mean = 8.3%), indicating a high false positive rate. By contrast, the clustered permutation test without resolved polytomies had a mean rate of only 2.2% negatively evolving lineages, approximately as expected given an alpha value of 0.025. Resolving polytomies and then performing the clustered permutation test improved performance even more, with a mean rate of 1.2% negatively evolving lineages and no dataset having more than 2.8% of lineages negatively evolving. This analysis shows the uniform date randomization test is prone to false positives in empirical B cell data, while the clustered date randomization test with resolved polytomies corrects this issue. All other tests performed in this study used a one-tailed, clustered date randomization test with resolved polytomies and an alpha value of 0.05.

To identify and characterize measurably evolving lineages while adjusting for multiple testing, all lineages tested were pooled together and p values were adjusted using the Benjamini–Hochberg procedure (*Benjamini and Hochberg, 1995*) implemented in the function p.adjust (*R Development Core Team, 2017*). Lineages with adjusted p values <0.1 were referred to as adjusted measurably evolving lineages (*Figure 2B*). To determine whether lineages were measurably increasing in amino acid divergence (*Figure 3A*), we repeated the clustered date randomization test for each tree. However, instead of calculating divergence as the sum of phylogenetic branch lengths leading from each tip to the most recent common ancestor of the lineage, we calculated divergence as the number of nonambiguous amino acid differences between each tip and the lineage's clonal germline. The clustered permutation test then proceeded as before, using the same cluster assignments as in the nucleotide-based test. This tested whether sequences at later timepoints had more amino acid substitutions compared to the germline than sequences at earlier timepoints.

It is possible that the results reported are affected by the size (number of sequences) of lineages in each dataset. A large number of lineages without adequate power could result in a spurious lack of measurable evolution. To ensure the lineages included in each study were adequately powered, we included only lineages with at least 15 sequences, that were sampled over at least 3 weeks, and had a minimum possible p value <0.05 based on the number of distinct permutations of timepoints among clusters. If measurable evolution were still strongly confounded by lineage size even after these filtering steps, we would expect measurably evolving lineages to be larger on average than nonmeasurably evolving lineages. By contrast, measurably evolving lineages were significantly larger than nonmeasurably evolving lineages in only 5/21 datasets surveyed (*Figure 3—figure supplement 2*), indicating our results are not strongly confounded by lineage size.

Inclusion of experimentally validated mAbs

To identify B cell lineages that likely bind to the antigen under study, we included experimentally validated monoclonal antibody (mAb) heavy chain sequences provided from multiple studies. This included multiple anti-HIV mAbs: 11 from *Liao et al., 2013*, 12 from *Doria-Rose et al., 2014*, 7 from *Johnson et al., 2018*, 42 from *Landais et al., 2017*, 31 from *Wu et al., 2015*, and 4 from *Huang et al., 2016*. *Doria-Rose et al., 2014* and *Wu et al., 2015* also provided 680 and 1033 bulk BCR sequences, respectively, identified as clonally related to the provided anti-HIV broadly neutralizing mAbs. These sequences were also included in processing and clonal clustering but were not labeled as experimentally validated mAbs. *Davis et al., 2019* provided 885 mAb heavy chain sequences, some of which were tested for binding against Ebola virus proteins. All of these sequences were included in processing and clonal clustering, but only 368 validated by ELISA to bind to Ebola virus were labeled as EBV-binding mAbs. All of the above sequences were processed in the same manner as bulk sequences from OAS, except they were not filtered as potential PCR chimeras. Clonal lineages containing experimentally validated mAbs were labeled as antigen-binding; however, because sample

timepoints were not always apparent, mAb sequences themselves were removed before lineage tree inference for the abovementioned studies. Processed data from *Turner et al., 2020* also included 196 anti-influenza mAbs. These sequences were retained during tree inference because they were explicitly labeled by timepoint and usually cloned from previously identified sequences within the data. Of all mAbs included, only the 58 clonally clustered within powered lineages (at least 15 sequences sampled over 3 weeks, and minimum p value <0.05) were included in tests of mAb enrichment (*Figure 2—figure supplement 3*).

Simulation-based power analysis

We used simulations to determine whether the clustered date randomization test was sufficiently powered to detect ongoing B cell evolution. These analyses used the *bcr-phylo* package accessed 9/21/2020 (*Davidsen and Matsen, 2018*; *Ralph and Matsen, 2020*), which simulates clonal lineages of B cells undergoing affinity maturation against a target sequence. For all simulations, a random naive heavy chain sequence was chosen from those provided in *bcr-phylo* and the rate of SHM was set to the default of $\lambda = 0.356$, which corresponds to an SHM rate of ~0.001 SHM/site/division (*Teng and Papavasiliou, 2007*). Mutations were introduced according to the S5F model (*Yaari et al., 2013*). Selection strength was chosen to be either 0 (neutral) or 1 (entirely affinity driven). A single target sequence was chosen for affinity maturation. All other parameters were set to their default.

We performed two sets of simulations. In the first, we simulated single B cell lineages from which 50 cells were sampled at generation 10, and 50 more cells were sampled after a specified number of additional generations (*Figure 1G*, *Figure 1—figure supplement 3*). In the second type of simulation, we replicated the sampling strategy of *Laserson et al., 2014*. Briefly, for each clone in subject *hu420143* from *Laserson et al., 2014*, we simulated one lineage with the same number of cells (if enough cells had been generated) sampled after the number of generations corresponding to 1, 3, 7, 14, 21, and 28 days (*Figure 1—figure supplement 4*). The number of generations corresponding to each sample day was calculated using a strict generation time of either 12 or 24 hr, which are conservative given previous GC cycle estimates of 6–12 hr (*Allen et al., 2007*; *Hauser et al., 2007*; *Victora and Nussenzweig, 2012*). These simulations used a selection strength of 1, which gave more conservative results in previous simulations (*Figure 1—figure supplement 3*).

To account for possible issues with clonal clustering, we did not preserve clonal identities among simulated sequences in either simulation type. Instead, we pooled sequences from all simulation repetitions under a particular parameter set and used the same clonal clustering method used for empirical data analyses to group them into clonal clusters. We did not repeat the genotyping or chimera filtering steps done on empirical data analyses as genotyped individuals and sequence chimeras were not part of the simulations. We performed the clustered date randomization test with resolved polytomies on each lineage with a minimum possible p value <0.05. Because all sequences were simulated under affinity maturation, the proportion of lineages with p < 0.05 indicated the true positive rate of the test. To determine the false positive rate, we randomized sample times among tips within each tree and repeated the date randomization test (*Figure 1—figure supplements 3 and 4*). Here, the proportion with p < 0.05 indicated the false positive rate.

Analysis of selection

To understand the force of selection operating on B cell lineages, we first separated all adjusted measurably evolving lineages into their respective subjects within each study. We then excluded all subjects with only one measurably evolving lineage. While all sequences included were labeled as productive by IgBlast, three contained premature stop codons in their IMGT-aligned sequences, likely due to insertions that were removed during alignment. These sequences were removed. For computational efficiency, all lineages were down-sampled to a maximum size of 100 sequences. Due to uncertainty in germline D-region assignment, only V-gene (IMGT positions 1–312) nucleotides were included for analyses of selection, similar to *Hoehn et al., 2017*. We then estimated lineage tree topologies, branch lengths, and subject-wide substitution model parameters under the GY94 model (*Hoehn et al., 2019*; *Nielsen and Yang, 1998*). Using fixed tree topologies estimated from the GY94 model, we then estimated branch lengths, subject-wide ω values for CDR and FWR partitions (ω_{CDR} and ω_{FWR}), and all six canonical SHM hot- and cold-spot motif parameters under the HLP19 model in IgPhyML v1.1.3 (*Hoehn et al., 2019*) for all adjusted measurably evolving lineages. Significance of ω

estimates was determined using two phylogenetic likelihood ratio tests, similar to *Hoehn et al., 2017*. To determine the significance of ω_{CDR} estimates, we compared the maximum log-likelihood obtained when both ω_{CDR} and ω_{FWR} were estimated by maximum likelihood (L) to that obtained when ω_{FWR} was estimated by maximum likelihood but ω_{CDR} was fixed at 1 ($L_{CDR=1}$). The likelihood ratio statistic (LRS) for this test was calculated as $2\times(L - L_{CDR=1})$. Because these models differ by one freely estimated parameter, the LRS will be approximately chi-squared distributed with one degree of freedom under the null hypothesis that $\omega_{CDR} = 1$, which allows for p value calculation (*Huelsenbeck and Rannala, 1997*). To determine significance of ω_{FWR} estimates, the process is the same except LRS = $2\times(L - L_{FWR=1})$, where $L_{FWR=1}$ is the maximum log-likelihood obtained when ω_{CDR} was estimated by maximum likelihood but ω_{FWR} was fixed at 1 ($L_{FWR=1}$). All of the above statistics are reported in *Table 2*.

Data and material availability

All data are publically available from prior publications. Script to reproduce all analyses performed are available at https://bitbucket.org/kleinstein/projects.git (*Kleinstein Lab, 2021*; copy archived at swh:1:rev:1ca83cda5d1baac880c71c314b0adc359314f6fa).

Acknowledgements

Additional information

Competing interests

Kenneth B Hoehn: receives consulting fees from Prellis Biologics. Jackson S Turner: is the recipient of a licensing agreement with Abbvie and has received consulting fees from Gerson Lehman Group. Ali H Ellebedy: The Ellebedy laboratory received funding under sponsored research agreements from Emergent BioSolutions and AbbVie. Steven H Kleinstein: receives consulting fees from Northrop Grumman and Peraton. The other authors declare that no competing interests exist.

Funding

Author contributions
Kenneth B Hoehn, Conceptualization, Data curation, Formal analysis, Methodology, Software, Writing – original draft, Writing – review and editing; Jackson S Turner, Ruoyi Jiang, Data curation, Resources, Writing – review and editing; Frederick I Miller, Data curation, Software, Writing – review and editing; Oliver G Pybus, Methodology, Writing – review and editing; Ali H Ellebedy, Resources, Supervision, Writing – review and editing; Steven H Kleinstein, Conceptualization, Funding acquisition, Supervision, Writing – original draft, Writing – review and editing

Author ORCIDs
Kenneth B Hoehn http://orcid.org/0000-0003-0411-4307
Steven H Kleinstein http://orcid.org/0000-0003-4957-1544

Decision letter and Author response
Decision letter https://doi.org/10.7554/eLife.70873.sa1
Author response https://doi.org/10.7554/eLife.70873.sa2

Additional files

Supplementary files
• Transparent reporting form

Data availability
The manuscript is a computational study. All data used are publicaly available. Source code are available at https://bitbucket.org/kleinstein/projects (copy archived at https://archive.softwareheritage.org/swh:1:rev:1ca83cda5d1baac880c71c314b0adc359314f6fa). All of the OAS datasets are available at this URL: http://opig.stats.ox.ac.uk/webapps/oas.

References

Allen CDC, Okada T, Tang HL, Cyster JG. 2007. Imaging of germinal center selection events during affinity maturation. *Science* **315**:528–531. DOI: https://doi.org/10.1126/science.1136736, PMID: 17185562

Arevalo P, McLean HQ, Belongia EA, Cobey S. 2020. Earliest infections predict the age distribution of seasonal influenza A cases. *eLife* Vol. 9:e50060. DOI: https://doi.org/10.7554/eLife.50060, PMID: 32633233

Benjamini Y, Hochberg Y. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society* **57**:289–300. DOI: https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Cizmeci D, Lofano G, Rossignol E, Dugast AS, Kim D, Cavet G, Nguyen N, Tan YC, Seaman MS, Alter G, Julg B. 2021. Distinct clonal evolution of B-cells in HIV controllers with neutralizing antibody breadth. *eLife* **10**:e62648. DOI: https://doi.org/10.7554/eLife.62648, PMID: 33843586

Davidsen K, Matsen FA. 2018. Benchmarking Tree and Ancestral Sequence Inference for B Cell Receptor Sequences. *Frontiers in Immunology* **9**:2451. DOI: https://doi.org/10.3389/fimmu.2018.02451, PMID: 30429847

Davis CW, Jackson KJL, McElroy AK, Halfmann P, Huang J, Chennareddy C, Piper AE, Leung Y, Albariño CG, Crozier I, Ellebedy AH, Sidney J, Sette A, Yu T, Nielsen SCA, Goff AJ, Spiropoulou CF, Saphire EO, Cavet G, Kawaoka Y, et al. 2019. Longitudinal Analysis of the Human B Cell Response to Ebola Virus Infection. *Cell* **177**:1566-1582.. DOI: https://doi.org/10.1016/j.cell.2019.04.036, PMID: 31104840

Doria-Rose NA, Schramm CA, Gorman J, Moore PL, Bhiman JN, DeKosky BJ, Ernandes MJ, Georgiev IS, Kim HJ, Pancera M, Staupe RP, Altae-Tran HR, Bailer RT, Crooks ET, Cupo A, Druz A, Garrett NJ, Hoi KH, Kong R, Louder MK, et al. 2014. Developmental pathway for potent V1V2-directed HIV-neutralizing antibodies. *Nature* **509**:55–62. DOI: https://doi.org/10.1038/nature13036, PMID: 24590074

Drummond AJ, Pybus OG, Rambaut A, Forsberg R, Rodrigo AG. 2003. Measurably evolving populations. *Trends in Ecology & Evolution* **18**:481–488. DOI: https://doi.org/10.1016/S0169-5347(03)00216-7

du Plessis L, McCrone JT, Zarebski AE, Hill V, Ruis C, Gutierrez B, Raghwani J, Ashworth J, Colquhoun R, Connor TR, Faria NR, Jackson B, Loman NJ, O'Toole Á, Nicholls SM, Parag KV, Scher E, Vasylyeva TI, Volz EM, Watts A, et al. 2021. Establishment and lineage dynamics of the SARS-CoV-2 epidemic in the UK. *Science* **371**:708–712. DOI: https://doi.org/10.1126/science.abf2946, PMID: 33419936

Duchêne S, Duchêne D, Holmes EC, Ho SYW. 2015. The Performance of the Date-Randomization Test in Phylogenetic Analyses of Time-Structured Virus Data. *Molecular Biology and Evolution* **32**:1895–1906. DOI: https://doi.org/10.1093/molbev/msv056, PMID: 25771196

Ellebedy AH, Jackson KJL, Kissick HT, Nakaya HI, Davis CW, Roskin KM, McElroy AK, Oshansky CM, Elbein R, Thomas S, Lyon GM, Spiropoulou CF, Mehta AK, Thomas PG, Boyd SD, Ahmed R. 2016. Defining antigen-specific plasmablast and memory B cell subsets in human blood after viral infection or vaccination. *Nature Immunology* **17**:1226–1234. DOI: https://doi.org/10.1038/ni.3533, PMID: 27525369

Ellebedy AH. 2018. Immunizing the Immune: Can We Overcome Influenza's Most Formidable Challenge? *Vaccines* **6**:68. DOI: https://doi.org/10.3390/vaccines6040068, PMID: 30248996

Gadala-Maria D, Yaari G, Uduman M, Kleinstein SH. 2015. Automated analysis of high-throughput B-cell sequencing data reveals a high frequency of novel immunoglobulin V gene segment alleles. *PNAS* **112**:E862-E870. DOI: https://doi.org/10.1073/pnas.1417683112, PMID: 25675496

Galson JD, Clutterbuck EA, Trück J, Ramasamy MN, Münz M, Fowler A, Cerundolo V, Pollard AJ, Lunter G, Kelly DF. 2015a. BCR repertoire sequencing: different patterns of B-cell activation after two Meningococcal vaccines. *Immunology and Cell Biology* **93**:885–895. DOI: https://doi.org/10.1038/icb.2015.57, PMID: 25976772

Galson JD, Trück J, Fowler A, Clutterbuck EA, Münz M, Cerundolo V, Reinhard C, van der Most R, Pollard AJ, Lunter G, Kelly DF. 2015b. Analysis of B Cell Repertoire Dynamics Following Hepatitis B Vaccination in Humans, and Enrichment of Vaccine-Specific Antibody Sequences. *EBioMedicine* **2**:2070–2079. DOI: https://doi.org/10.1016/j.ebiom.2015.11.034

Galson JD, Trück J, Clutterbuck EA, Fowler A, Cerundolo V, Pollard AJ, Lunter G, Kelly DF. 2016. B-cell repertoire dynamics after sequential hepatitis B vaccination and evidence for cross-reactive B-cell activation. *Genome Medicine* **8**:68. DOI: https://doi.org/10.1186/s13073-016-0322-z, PMID: 27312086

Giudicelli V, Chaume D, Lefranc MP. 2005. IMGT/GENE-DB: a comprehensive database for human and mouse immunoglobulin and T cell receptor genes. *Nucleic Acids Research* **33**:D256-D261. DOI: https://doi.org/10.1093/nar/gki010, PMID: 15608191

Gupta NT, Vander Heiden JA, Uduman M, Gadala-Maria D, Yaari G, Kleinstein SH. 2015. Change-O: a toolkit for analyzing large-scale B cell immunoglobulin repertoire sequencing data. *Bioinformatics* **31**:3356–3358. DOI: https://doi.org/10.1093/bioinformatics/btv359, PMID: 26069265

Gupta NT, Adams KD, Briggs AW, Timberlake SC, Vigneault F, Kleinstein SH. 2017. Hierarchical Clustering Can Identify B Cell Clones with High Confidence in Ig Repertoire Sequencing Data *Journal of Immunology* **198**:2489–2499. DOI: https://doi.org/10.4049/jimmunol.1601850, PMID: 28179494

Hauser AE, Junt T, Mempel TR, Sneddon MW, Kleinstein SH, Henrickson SE, von Andrian UH, Shlomchik MJ, Haberman AM. 2007. Definition of Germinal-Center B Cell Migration In Vivo Reveals Predominant Intrazonal Circulation Patterns. *Immunity* **26**:655–667. DOI: https://doi.org/10.1016/j.immuni.2007.04.008

Hoehn KB, Lunter G, Pybus OG. 2017. A Phylogenetic Codon Substitution Model for Antibody Lineages. *Genetics* **206**:417–427. DOI: https://doi.org/10.1534/genetics.116.196303, PMID: 28315836

Hoehn KB, Vander Heiden JA, Zhou JQ, Lunter G, Pybus OG, Kleinstein SH. 2019. Repertoire-wide phylogenetic models of B cell molecular evolution reveal evolutionary signatures of aging and vaccination. *PNAS* Vol. 116:22664–22672. DOI: https://doi.org/10.1073/pnas.1906020116, PMID: 31636219

Hoehn KB, Pybus OG, Kleinstein SH. 2020. Phylogenetic Analysis of Migration, Differentiation, and Class Switching in B Cells. [bioRxiv] DOI: https://doi.org/10.1101/2020.05.30.124446

Huang J, Kang BH, Ishida E, Zhou T, Griesman T, Sheng Z, Wu F, Doria-Rose NA, Zhang B, McKee K, O'Dell S, Chuang G-Y, Druz A, Georgiev IS, Schramm CA, Zheng A, Joyce MG, Asokan M, Ransier A, Darko S, et al. 2016. Identification of a CD4-Binding-Site Antibody to HIV that Evolved Near-Pan Neutralization Breadth. *Immunity* **45**:1108–1121. DOI: https://doi.org/10.1016/j.immuni.2016.10.027, PMID: 27851912

Huelsenbeck JP, Rannala B. 1997. Phylogenetic methods come of age: testing hypotheses in an evolutionary context. *Science* **276**:227–232. DOI: https://doi.org/10.1126/science.276.5310.227, PMID: 9092465

Jiang R, Fichtner ML, Hoehn KB, Pham MC, Stathopoulos P, Nowak RJ, Kleinstein SH, O'Connor KC. 2020a. Single-cell repertoire tracing identifies rituximab-resistant B cells during myasthenia gravis relapses. *JCI Insight* Vol. 5:136471. DOI: https://doi.org/10.1172/jci.insight.136471, PMID: 32573488

Jiang R, Hoehn KB, Lee CS, Pham MC, Homer RJ, Detterbeck FC, Aban I, Jacobson L, Vincent A, Nowak RJ, Kaminski HJ, Kleinstein SH, O'Connor KC. 2020b. Thymus-derived B cell clones persist in the circulation after thymectomy in myasthenia gravis. *PNAS* **117**:30649–30660. DOI: https://doi.org/10.1073/pnas.2007206117, PMID: 33199596

Johnson EL, Doria-Rose NA, Gorman J, Bhiman JN, Schramm CA, Vu AQ, Law WH, Zhang B, Bekker V, Abdool Karim SS, Ippolito GC, Morris L, Moore PL, Kwong PD, Mascola JR, Georgiou G. 2018. Sequencing HIV-neutralizing antibody exons and introns reveals detailed aspects of lineage maturation. *Nature Communications* **9**:4136. DOI: https://doi.org/10.1038/s41467-018-06424-6, PMID: 30297708

Kassambara A. 2020. ggpubr: ggplot2 Based Publication Ready Plots. 0.4.0. Ggpubr. https://CRAN.R-project.org/package=ggpubr

Kleinstein SH, Louzoun Y, Shlomchik MJ. 2003. Estimating hypermutation rates from clonal tree data *Journal of Immunology* **171**:4639–4649. DOI: https://doi.org/10.4049/jimmunol.171.9.4639, PMID: 14568938

Kleinstein Lab. 2021. Repository of publication scripts. swh:1:rev:1ca83cda5d1baac880c71c314b0adc359314f6fa. Software Heritage. https://archive.softwareheritage.org/swh:1:dir:87cffd9bb02b26906cfb0cdc056dbdf5bc61981f;origin=https://bitbucket.org/kleinstein/projects.git;visit=swh:1:snp:ae1a4d6e82099e0e8c05ec965dd7c2d82206e3a3;anchor=swh:1:rev:1ca83cda5d1baac880c71c314b0adc359314f6fa

Kovaltsuk A, Leem J, Kelm S, Snowden J, Deane CM, Krawczyk K. 2018. Observed Antibody Space: A Resource for Data Mining Next-Generation Sequencing of Antibody Repertoires. *Journal of Immunology* **201**:2502–2509. DOI: https://doi.org/10.4049/jimmunol.1800708, PMID: 30217829

Landais E, Murrell B, Briney B, Murrell S, Rantalainen K, Berndsen ZT, Ramos A, Wickramasinghe L, Smith ML, Eren K, de Val N, Wu M, Cappelletti A, Umotoy J, Lie Y, Wrin T, Algate P, Chan-Hui P-Y, Karita E, IAVI Protocol C Investigators, et al. 2017. HIV Envelope Glycoform Heterogeneity and Localized Diversity Govern the

Initiation and Maturation of a V2 Apex Broadly Neutralizing Antibody Lineage. *Immunity* **47**:990-1003.. DOI: https://doi.org/10.1016/j.immuni.2017.11.002, PMID: 29166592

Laserson U, Vigneault F, Gadala-Maria D, Yaari G, Uduman M, Vander Heiden JA, Kelton W, Taek Jung S, Liu Y, Laserson J, Chari R, Lee J-H, Bachelet I, Hickey B, Lieberman-Aiden E, Hanczaruk B, Simen BB, Egholm M, Koller D, Georgiou G, et al. 2014. High-resolution antibody dynamics of vaccine-induced immune responses. *PNAS* **111**:4928–4933. DOI: https://doi.org/10.1073/pnas.1323862111, PMID: 24639495

Levin M, King JJ, Glanville J, Jackson KJL, Looney TJ, Hoh RA, Mari A, Andersson M, Greiff L, Fire AZ, Boyd SD, Ohlin M. 2016. Persistence and evolution of allergen-specific IgE repertoires during subcutaneous specific immunotherapy. *The Journal of Allergy and Clinical Immunology* **137**:1535–1544. DOI: https://doi.org/10.1016/j.jaci.2015.09.027, PMID: 26559321

Liao H-X, Lynch R, Zhou T, Gao F, Alam SM, Boyd SD, Fire AZ, Roskin KM, Schramm CA, Zhang Z, Zhu J, Shapiro L, NISC Comparative Sequencing Program, Mullikin JC, Gnanakaran S, Hraber P, Wiehe K, Kelsoe G, Yang G, Xia S-M, et al. 2013. Co-evolution of a broadly neutralizing HIV-1 antibody and founder virus. *Nature* **496**:469–476. DOI: https://doi.org/10.1038/nature12053, PMID: 23552890

McKean D, Huppi K, Bell M, Staudt L, Gerhard W, Weigert M. 1984. Generation of antibody diversity in the immune response of BALB/c mice to influenza virus hemagglutinin. *PNAS* **81**:3180–3184. DOI: https://doi.org/10.1073/pnas.81.10.3180, PMID: 6203114

Mesin L, Schiepers A, Ersching J, Barbulescu A, Cavazzoni CB, Angelini A, Okada T, Kurosaki T, Victora GD. 2020. Restricted Clonality and Limited Germinal Center Reentry Characterize Memory B Cell Reactivation by Boosting. *Cell* **180**:92-106.. DOI: https://doi.org/10.1016/j.cell.2019.11.032, PMID: 31866068

Murphy K, Trivers P, Walport M. 2008. Janeway's Immunobiology. 7th ed. New York, NY: Garland Science. DOI: https://doi.org/10.1097/01.SHK.0000286285.87596.06

Murray GGR, Wang F, Harrison EM, Paterson GK, Mather AE, Harris SR, Holmes MA, Rambaut A, Welch JJ. 2016. The effect of genetic structure on molecular dating and tests for temporal signal. *Methods in Ecology and Evolution* **7**:80–89. DOI: https://doi.org/10.1111/2041-210X.12466, PMID: 27110344

Nielsen R, Yang Z. 1998. Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. *Genetics* **148**:929–936. DOI: https://doi.org/10.1093/genetics/148.3.929, PMID: 9539414

Nielsen SCA, Roskin KM, Jackson KJL, Joshi SA, Nejad P, Lee JY, Wagar LE, Pham TD, Hoh RA, Nguyen KD, Tsunemoto HY, Patel SB, Tibshirani R, Ley C, Davis MM, Parsonnet J, Boyd SD. 2019. Shaping of infant B cell receptor repertoires by environmental factors and infectious disease. *Science Translational Medicine* **11**:eaat2004. DOI: https://doi.org/10.1126/scitranslmed.aat2004, PMID: 30814336

Nourmohammad A, Otwinowski J, Łuksza M, Mora T, Walczak AM. 2019. Fierce Selection and Interference in B-Cell Repertoire Response to Chronic HIV-1. *Molecular Biology and Evolution* **36**:2184–2194. DOI: https://doi.org/10.1093/molbev/msz143, PMID: 31209469

Paradis E, Claude J, Strimmer K. 2004. APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics* **20**:289–290. DOI: https://doi.org/10.1093/bioinformatics/btg412, PMID: 14734327

R Development Core Team. 2017. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/

Ralph DK, Matsen FA. 2020. Using B cell receptor lineage structures to predict affinity. *PLOS Computational Biology* **16**:e1008391. DOI: https://doi.org/10.1371/journal.pcbi.1008391, PMID: 33175831

Rambaut A, Posada D, Crandall KA, Holmes EC. 2004. The causes and consequences of HIV evolution. *Nature Reviews. Genetics* **5**:52–61. DOI: https://doi.org/10.1038/nrg1246, PMID: 14708016

Rambaut A, Lam TT, Max Carvalho L, Pybus OG. 2016. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evolution* **2**:vew007. DOI: https://doi.org/10.1093/ve/vew007, PMID: 27774300

Schanz M, Liechti T, Zagordi O, Miho E, Reddy ST, Günthard HF, Trkola A, Huber M. 2014. High-throughput sequencing of human immunoglobulin variable regions with subtype identification. *PLOS ONE* Vol. 9:e111726. DOI: https://doi.org/10.1371/journal.pone.0111726, PMID: 25364977

Schliep KP. 2011. phangorn: phylogenetic analysis in R. *Bioinformatics* **27**:592–593. DOI: https://doi.org/10.1093/bioinformatics/btq706, PMID: 21169378

Setliff I, McDonnell WJ, Raju N, Bombardi RG, Murji AA, Scheepers C, Ziki R, Mynhardt C, Shepherd BE, Mamchak AA, Garrett N, Karim SA, Mallal SA, Crowe JE, Morris L, Georgiev IS. 2018. Multi-Donor Longitudinal Antibody Repertoire Sequencing Reveals the Existence of Public Antibody Clonotypes in HIV-1 Infection. *Cell Host & Microbe* **23**:845-854.. DOI: https://doi.org/10.1016/j.chom.2018.05.001, PMID: 29861170

Sheng Z, Schramm CA, Connors M, Morris L, Mascola JR, Kwong PD, Shapiro L. 2016. Effects of Darwinian Selection and Mutability on Rate of Broadly Neutralizing Antibody Evolution during HIV-1 Infection. *PLOS Computational Biology* **12**:e1004940. DOI: https://doi.org/10.1371/journal.pcbi.1004940, PMID: 27191167

Shlomchik MJ, Weisel F. 2012. Germinal center selection and the development of memory B and plasma cells. *Immunological Reviews* **247**:52–63. DOI: https://doi.org/10.1111/j.1600-065X.2012.01124.x, PMID: 22500831

Takemori T, Kaji T, Takahashi Y, Shimoda M, Rajewsky K. 2014. Generation of memory B cells inside and outside germinal centers. *European Journal of Immunology* **44**:1258–1264. DOI: https://doi.org/10.1002/eji.201343716, PMID: 24610726

Taylor JJ, Pape KA, Jenkins MK. 2012. A germinal center-independent pathway generates unswitched memory B cells early in the primary response. *The Journal of Experimental Medicine* **209**:597–606. DOI: https://doi.org/10.1084/jem.20111696, PMID: 22370719

Teng G, Papavasiliou FN. 2007. Immunoglobulin somatic hypermutation. *Annual Review of Genetics* **41**:107–120. DOI: https://doi.org/10.1146/annurev.genet.41.110306.130340, PMID: 17576170

Tsioris K, Gupta NT, Ogunniyi AO, Zimnisky RM, Qian F, Yao Y, Wang X, Stern JNH, Chari R, Briggs AW, Clouser CR, Vigneault F, Church GM, Garcia MN, Murray KO, Montgomery RR, Kleinstein SH, Love JC. 2015. Neutralizing antibodies against West Nile virus identified directly from human B cells by single-cell analysis and next generation sequencing. *Integrative Biology* **7**:1587–1597. DOI: https://doi.org/10.1039/c5ib00169b, PMID: 26481611

Turner JS, Zhou JQ, Han J, Schmitz AJ, Rizk AA, Alsoussi WB, Lei T, Amor M, McIntire KM, Meade P, Strohmeier S, Brent RI, Richey ST, Haile A, Yang YR, Klebert MK, Suessen T, Teefey S, Presti RM, Krammer F, et al. 2020. Human germinal centres engage memory and naive B cells after influenza vaccination. *Nature* **586**:127–132. DOI: https://doi.org/10.1038/s41586-020-2711-0, PMID: 32866963

Uduman M, Yaari G, Hershberg U, Stern JA, Shlomchik MJ, Kleinstein SH. 2011. Detecting selection in immunoglobulin sequences. *Nucleic Acids Research* **39**:W499-W504. DOI: https://doi.org/10.1093/nar/gkr413, PMID: 21665923

Unterman A, Sumida TS, Nouri N, Yan X, Zhao AY, Gasque V, Schupp JC, Asashima H, Liu Y, Cosme C, Deng W, Chen M, Raredon MSB, Hoehn K, Wang G, Wang Z, Deiuliis G, Ravindra NG, Li N, Castaldi C, et al. 2020. Single-Cell Omics Reveals Dyssynchrony of the Innate and Adaptive Immune System in Progressive COVID-19. [*medRxiv*]. DOI: https://doi.org/10.1101/2020.07.16.20153437

Vander Heiden JA, Yaari G, Uduman M, Stern JNH, O'Connor KC, Hafler DA, Vigneault F, Kleinstein SH. 2014. pRESTO: a toolkit for processing high-throughput sequencing raw reads of lymphocyte receptor repertoires. *Bioinformatics* **30**:1930–1932. DOI: https://doi.org/10.1093/bioinformatics/btu138, PMID: 24618469

Victora GD, Nussenzweig MC. 2012. Germinal centers. *Annual Review of Immunology* **30**:429–457. DOI: https://doi.org/10.1146/annurev-immunol-020711-075032, PMID: 22224772

Vieira MC, Zinder D, Cobey S. 2018. Selection and Neutral Mutations Drive Pervasive Mutability Losses in Long-Lived Anti-HIV B-Cell Lineages. *Molecular Biology and Evolution* **35**:1135–1146. DOI: https://doi.org/10.1093/molbev/msy024, PMID: 29688540

Wang C, Liu Y, Xu LT, Jackson KJL, Roskin KM, Pham TD, Laserson J, Marshall EL, Seo K, Lee J-Y, Furman D, Koller D, Dekker CL, Davis MM, Fire AZ, Boyd SD. 2014. Effects of aging, cytomegalovirus infection, and EBV infection on human B cell repertoires. *Journal of Immunology* **192**:603–611. DOI: https://doi.org/10.4049/jimmunol.1301384, PMID: 24337376

Wendel BS, Fu Y, He C, Hernandez SM, Qu M, Zhang Z, Jiang Y, Han X, Xu J, Ding H, Jiang N, Shang H. 2020. Rapid HIV Progression Is Associated with Extensive Ongoing Somatic Hypermutation. *Journal of Immunology* **205**:587–594. DOI: https://doi.org/10.4049/jimmunol.1901161, PMID: 32591400

Wickham H. 2016. ggplot2. *Ggplot2: Elegant Graphics for Data Analysis*. Cham: Springer-Verlag. DOI: https://doi.org/10.1007/978-3-319-24277-4

Wrammert J, Smith K, Miller J, Langley WA, Kokko K, Larsen C, Zheng NY, Mays I, Garman L, Helms C, James J, Air GM, Capra JD, Ahmed R, Wilson PC. 2008. Rapid cloning of high-affinity human monoclonal antibodies against influenza virus. *Nature* **453**:667–671. DOI: https://doi.org/10.1038/nature06890, PMID: 18449194

Wu X, Zhou T, Zhu J, Zhang B, Georgiev I, Wang C, Chen X, Longo NS, Louder M, McKee K, O'Dell S, Perfetto S, Schmidt SD, Shi W, Wu L, Yang Y, Yang Z-Y, Yang Z, Zhang Z, Bonsignori M, et al. 2011. Focused evolution of HIV-1 neutralizing antibodies revealed by structures and deep sequencing. *Science* **333**:1593–1602. DOI: https://doi.org/10.1126/science.1207532, PMID: 21835983

Wu X, Zhang Z, Schramm CA, Joyce MG, Kwon YD, Zhou T, Sheng Z, Zhang B, O'Dell S, McKee K, Georgiev IS, Chuang G-Y, Longo NS, Lynch RM, Saunders KO, Soto C, Srivatsan S, Yang Y, Bailer RT, Louder MK, et al. 2015. Maturation and Diversity of the VRC01-Antibody Lineage over 15 Years of Chronic HIV-1 Infection. *Cell* **161**:470–485. DOI: https://doi.org/10.1016/j.cell.2015.03.004, PMID: 25865483

Yaari G, Uduman M, Kleinstein SH. 2012. Quantifying selection in high-throughput Immunoglobulin sequencing data sets. *Nucleic Acids Research* **40**:e134. DOI: https://doi.org/10.1093/nar/gks457, PMID: 22641856

Yaari G, Vander Heiden JA, Uduman M, Gadala-Maria D, Gupta N, Stern JNH, O'Connor KC, Hafler DA, Laserson U, Vigneault F, Kleinstein SH. 2013. Models of somatic hypermutation targeting and substitution based on synonymous mutations from high-throughput immunoglobulin sequencing data. *Frontiers in Immunology* **4**:358. DOI: https://doi.org/10.3389/fimmu.2013.00358, PMID: 24298272

Yaari G, Benichou JIC, Vander Heiden JA, Kleinstein SH, Louzoun Y. 2015. The mutation patterns in B-cell immunoglobulin receptors reflect the influence of selection acting at multiple time-scales. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **370**:20140242. DOI: https://doi.org/10.1098/rstb.2014.0242, PMID: 26194756

Yang Z, Nielsen R, Goldman N, Pedersen AMK. 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics* **155**:431–449. DOI: https://doi.org/10.1093/genetics/155.1.431, PMID: 10790415

Ye J, Ma N, Madden TL, Ostell JM. 2013. IgBLAST: an immunoglobulin variable domain sequence analysis tool. *Nucleic Acids Research* **41**:W34-W40. DOI: https://doi.org/10.1093/nar/gkt382, PMID: 23671333

Yu G, Smith DK, Zhu H, Guan Y, Lam TT, McInerny G. 2016. ggtree : an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution* **8**:28–36. DOI: https://doi.org/10.1111/2041-210X.12628

Zhou JQ, Kleinstein SH. 2019. Cutting Edge: Ig H Chains Are Sufficient to Determine Most B Cell Clonal Relationships. *Journal of Immunology* **203**:1687–1692. DOI: https://doi.org/10.4049/jimmunol.1900666, PMID: 31484734