### Selection and neutral mutations drive pervasive mutability losses in long-lived B cell lineages

Marcos Vieira, Daniel Zinder & Sarah Cobey<sup>\*</sup>,

Department of Ecology and Evolution, University of Chicago **\*** cobey@uchicago.edu

## Abstract

High-affinity antibodies arise within weeks of infection from the evolution of B cell receptors under selection to improve antigen recognition. This rapid adaptation is enabled by the frequency and distribution of highly mutable "hotspot" motifs in B cell receptor genes. High mutability in antigen binding regions (CDRs) creates variation in binding affinity, whereas low mutability in structurally important regions (FRs) may reduce the frequency of destabilizing mutations. During the response, the loss of mutational hotspots and changes in their distribution across CDRs and FRs are predicted to compromise the adaptability of B cell receptors, yet the contributions of different mechanisms to gains and losses of hotspots remain unclear. We reconstructed changes in anti-HIV B cell receptor sequences and show that mutability losses were about 60% more frequent than gains in both CDRs and FRs, with the higher relative mutability of CDRs maintained throughout the response. At least 34% of the mutability losses were caused by synonymous mutations. However, non-synonymous substitutions caused most of the mutability loss in CDRs. Because CDRs also show strong positive selection, this result suggests positive selection contributed to as much as 66% of the mutability loss in those regions. Although recurrent adaptation to the evolving virus could indirectly select for high mutation rates, we found no evidence of indirect selection to increase or retain hotspots. Our results suggest mutability losses are intrinsic to the neutral and adaptive evolution of B cell populations and might constrain their adaptation to rapidly evolving pathogens such as HIV and influenza.

## Introduction

High-affinity antibodies arise during the adaptive immune response from the very process that gave vertebrates an adaptive immune system in the first place: adaptation by natural selection. In response to infection or vaccination, mutagenic enzymes and error-prone polymerases cause somatic hypermutation of B cell receptors and thus create variation in their ability to bind antigen [1]. B cells with high-affinity receptors are more likely to receive survival and replication signals from helper T cells, and thus selection for improved antigen binding drives the development of high-affinity B cell receptors that are later secreted as antibodies [2–5]. Understanding how the immune system evolved to facilitate the rapid adaptation of B cell receptors during infection may provide general insights into the evolution of adaptability.

Two features of B cell receptor genes suggest their long-term evolution has been shaped by selection for adaptability during infection. First, the germline genes that recombine to produce B cell receptors are enriched for nucleotide motifs that are

1

1

2

6

8

10

11

12

13

targeted with high frequency by the mutagenic enzymes involved in somatic hypermutation [6–8]. High mutation rates provide the genetic variation required for B cell adaptation, and low B cell mutation rates have been linked to immunodeficiency disorders [9,10] and to the decline in immune function with age [11]. Second, mutational "hotspots" occur where mutations are most likely to be beneficial. Hotspots are concentrated in loops of the B cell receptor protein that are directly involved in antigen recognition (complementarity determining regions, CDRs) [12]. In contrast, structurally important regions of the B cell receptor (framework regions, FRs), which are usually less directly involved in antigen binding, are enriched with motifs that have low mutability [12]. In addition, mutations in CDRs are more likely to be non-synonymous and to produce non-conservative amino acid changes than mutations in FRs [13–16]. The differential mutability of CDRs and FRs appears to focus mutations to regions where they are likely to produce variation in antigen affinity without destabilizing the protein. The frequency and distribution of mutational hotspots in B cell receptor genes therefore seem to contribute to their adaptability during immune responses.

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60

61

62

63

64

65

66

As B cells mutate during the immune response, however, changes in the frequency and distribution of mutational hotspots might affect the subsequent adaptability of B cell receptors. This change in adaptability may be especially important in B cell lineages that coevolve with pathogens like HIV and influenza. Experimental removal of hotspots decreases somatic hypermutation rates in vitro and in laboratory B cell lines [8]. Loss of highly mutable motifs has been hypothesized to occur during the immune response due to motifs' propensity to mutate [17, 18], and decreased mutation rates due to such "hotspot decay" might explain declines in the evolutionary rates of B cell lineages over several years of HIV infection [19, 20]. In addition, changes in the distribution of highly mutable motifs across FRs and CDRs might increase the frequency of deleterious mutations in the former and decrease the frequency of beneficial mutations in the latter. However, evidence for consistent changes in mutability during the evolution of B cell lineages is inconclusive. Several metrics have been developed to estimate somatic hypermutation rates based on differences in the frequency of mutations across nucleotide motifs, and they range from the discrete classification of motifs into hotspots and non-hotspots [6,7] to more recent models that quantify relative mutation rates on a continuous scale [12, 21, 22]. While the average mutability of B cell receptor sequences decreases during HIV infection when measured by the number of classically defined hotspots [17], long-term trends are less consistent for more recent metrics [20].

Factors other than the decay of hotspots through random mutations might affect the mutability of B cell receptors. First, highly mutable motifs might be regained through mutation. Second, selection for affinity and protein stability might favor non-synonymous mutations that incidentally increase or decrease mutability. Finally, selection might act on somatic hypermutation rates themselves. Although selection in theory can favor low mutation rates due to the reduced frequency of deleterious mutation rate through its association with beneficial mutations [23,24], rapidly changing environments may indirectly select for a higher mutation rate through its association with beneficial mutations [25–28]. Thus, although short-term differences in fitness among B cells arise from differences in the affinity of their receptors, B cells with more mutable CDRs might have a higher probability of producing high-affinity descendants able to keep up with evolving antigens in the long term. Indirect selection for mutability might therefore retain or increase the frequency of highly mutable motifs in CDRs, while mutability losses in FRs might be selected due to the reduced frequency of destabilizing mutations.

Understanding changes in mutability may reveal constraints on B cell adaptation, but the contributions of different mechanisms to changes in the frequency and distribution of mutational hotspots during B cell responses are largely unknown. We investigated the evolution of B cell receptor mutability by fitting phylogenetic models to sequences from long-lived anti-HIV B cell lineages. In characterizing mutability, we 67 considered two sequence-based features that appear to have been strongly selected in the 68 evolution of the adaptive immune system: overall mutability (the density of mutagenic 69 nucleotide motifs) and also changes in the mutability of CDRs relative to FRs. First, we 70 examined B cell mutability in the unmutated common ancestors of anti-HIV antibodies. 71 Next, we investigated the effects of random mutations and positive selection for amino 72 acid substitutions that increase affinity for antigen. Finally, we tested for selection to 73 increase, retain or decrease the frequency of highly mutable motifs.

74

75

76

77

79

80

81

82

83

85

86

87

88

89

90

91

92

93

94

95

96

97

98

99

100

101

102

103

104

105

106

107

108

109

110

111

112

113

114

115

# Results

#### Ancestral B cells have higher mutability in CDRs than FRs

To characterize changes in mutability during B cell evolution, we inferred the evolutionary histories of previously reported B cell lineages from three HIV-1 patients. The CH103 and VRC26 lineages comprise heavy and light chain B cell receptor sequences obtained from high-throughput sequencing over 144 and 206 weeks of infection in two patients, respectively [29, 30]. We also analyzed heavy chain sequences of three lineages from the VRC01 dataset, which was sampled from a third patient over a 15-year period [19]. The lineages we analyzed were originally investigated for having evolved the ability to neutralize diverse HIV strains.

To infer changes in mutability over time, we used Bayesian phylogenetic analyses (Methods) to obtain a sample of time-resolved trees from the posterior distribution of each lineage's genealogy for the heavy and light chains separately, and to estimate the nucleotide sequences of all internal nodes. Mutabilities of the observed and the inferred internal sequences were estimated using the S5F model. This model assigns relative mutation rates to all five-nucleotide DNA motifs and is based on a large independent dataset of antigen-experienced B cells [21]. The mutability of each sequence was defined as the average S5F score across all sites in the B cell receptor sequence. We estimated the number, magnitude and distribution of mutability changes on all branches by contrasting the mutability of all pairs of parent-descendant nodes. Figure 1 illustrates mutability evolution in the heavy chain of lineage CH103.

We first characterized the mutability of each lineage's unmutated ancestor. In the ancestors of all heavy and light chains, sites in CDRs had higher average mutability than sites in FRs. On average across the seven heavy and light chain ancestors, mutability was approximately 27% higher in CDRs than in FRs (range: 14-34%; Figure S1). Previous analyses of germline V genes (which, together with D and J genes, recombine to produce mature B cell receptors) showed that mutability was lower in FRs and higher in CDRs than expected based on their amino acid sequences, suggesting selection to increase the frequency of highly mutable motifs in CDRs and decrease their frequency in FRs [12, 15]. Consistent with those analyses, we found that the FRs of lineage ancestors had lower mutability than expected based on their amino acid sequences. On average across all heavy and light chains, the mean FR mutability of ancestral B cell receptors was lower than 99% of sequences obtained by randomizing their codons (according to usage frequencies in humans [31]) while keeping the amino acid sequences constant (range: 95-100%; Figure S1). However, while V genes often have higher CDR mutability than expected based on their amino acid sequences [12], mean CDR mutability in ancestral B cell receptors was, on average, greater than only 54% of randomized sequences with the same amino acid sequence (range: 18-95%, Figure S1). These patterns suggest that adaptation in these cells may on average be constrained more by destabilizing mutations in the FRs than a shortage of beneficial mutations in the CDRs.



Fig 1. Evolution of S5F-mutability in the heavy-chain CH103 B cell lineage. a) Long-term declines in average mutability across the maximum clade-credibility tree. Nodes are colored according to the average S5F-mutability of the whole sequence. Branches measured in weeks are annotated to indicate gains (+) or losses (-) of mutability. b) Average mutability over time for a combined sample of 100 trees from the posterior distribution. Blue points correspond to terminal nodes (observed sequences), and black points correspond to inferred internal nodes. The red line represents an average of regression lines calculated for each tree in a sample of 1000 trees. c) Distribution of changes in mutability for a sample of 100 trees from the posterior distribution shown as overlaid densities. The inset plot shows the posterior distribution of the fraction of changes on a tree that were losses constructed from a sample of 1000 trees, with the 95% highest-posterior density interval shown in gray. Red lines indicate the means of the distributions.



Fig 2. Frequency of losses relative to the total number of changes in S5F-mutability during the evolution of anti-HIV B cell lineages. Results are shown for the entire analyzed region of the B cell receptor, and separately for framework regions (FRs) and complementarity determining regions (CDRs). Each point denotes the fraction of changes in mutability that were losses, averaged across a sample of 1000 trees from the posterior distribution. Vertical red lines indicate the 95% highest-posterior density interval.

116

117

118

119

120

121

122

123

124

125

126

127

128

129

130

#### Mutability is more often lost than gained

Next we investigated long-term changes in mutability as B cell receptors evolved from their ancestors. Consistent with a previous analysis [17], average mutability decreased with time (Figure 1a,b, Figure S2). To investigate changes in mutability in greater detail, we analyzed the frequency of mutability gains and losses across branches. Mutability losses should arise from hotspot decay, positive selection of the amino acid changes that incidentally decrease mutability, selection for lower mutability due to the reduction in the frequency of deleterious mutations, or a combination of those factors. Mutability gains should reflect gains through mutation, positive selection of amino acid changes that incidentally increase mutability, indirect selection for higher mutation rates by association with beneficial mutations, or a combination of those factors. To summarize the net contributions of mutability-decreasing and mutability-increasing mechanisms, we computed the fraction of branches with mutability losses out of all branches with mutability changes. By computing the frequency of mutability losses for each tree in the posterior sample, we estimated its posterior distribution (Figure 1c).

Across the entire BCR sequence, mutability losses occurred more frequently than 131 gains in four of the five lineages (Figure 1a,c, Figure 2). On average across the seven 132 heavy and light chains, approximately 62% of changes in mutability were losses (range: 133 45% - 71%). Mutability losses were more frequent in the light chains (70% and 63%) 134 than in the heavy chains (60% and 62%) of lineages CH103 and VRC26. The frequency 135 of mutability was not significantly different from the frequency of gains in the heavy 136 chain of VRC26 (95% highest posterior interval 48-74%) and was slightly lower than the 137 frequency of gains in lineage VRC01-01 (42-49%). 138

Long-term declines in mutability occurred both in FRs (five of seven heavy and light chains; Figure S3) and CDRs (six of seven heavy and light chains; Figure S4). <sup>140</sup> Consistent with long-term declines in the average mutability of both regions, we found no long-term changes in the difference between CDR and FR mutabilities (Figure S5). <sup>142</sup> Despite considerable amino acid divergence from the unmutated ancestors in FRs <sup>143</sup>



Fig 3. Selection in framework regions (FRs) and complementarity determining regions (CDRs) of B cell receptors from anti-HIV B cell lineages. Selection strength is measured as the log odds ratio between the observed ratio of non-synonymous to synonymous substitution frequencies,  $\pi/(1-\pi)$ , and the ratio expected under the S5F mutability model in the absence of selection,  $\hat{\pi}/(1-\hat{\pi})$  [21,32]. Selection strength values greater than zero indicate positive selection, and values smaller than zero indicate purifying selection.

(16-67% for different heavy and light chains), sequences from the last sampling time in 144 each dataset had lower FR mutability than 95% of randomizations with constant amino 145 acid sequence (range: 70-99.9%, fig. S6). The lower mutability of ancestral FRs relative 146 to their amino acid sequences was therefore retained throughout the evolution of the B 147 cell lineages. CDRs, however, became less mutable relative to their amino acid 148 sequences than the ancestors. On average across the seven heavy and light chains, 149 CDRs of sequences from the last sampling times were more mutable than 28% of their 150 corresponding randomizations (range: 8-64%, Figure S6), down from 54% in ancestral 151 CDRs. Consistent with the net long-term trends, both CDRs and FRs had similar 152 frequencies of mutability losses (FR average 57%, range: 44% – 69%; CDR average 56%, 153 range: 44–64%; Figure 2). 154

155

156

157

158

159

160

161

162

163

164

165

166

#### Hotspot decay and selection for amino acid substitutions contribute to mutability losses

Highly mutable motifs have been hypothesized to decay due to their propensity to mutate [17,18], but motifs should also be influenced by positive selection. The first case is straightforward. For instance, the hypothetical sequence CAGCTT contains the highly mutable cytosine at the center of the AGCTT motif [7,8,21], and a  $C \rightarrow T$ mutation in the underlined position would decrease the mutability of the site approximately 5-fold [21]. Positive selection on amino acid substitutions should influence motifs in two ways. Mutational hotspots can be disrupted if selection favors amino acid sequences whose codons happen to be, on average, less mutable than codons in the ancestral sequence. Selection can also affect mutability in neighboring codons. For example, if selection led to the replacement of CAG (glutamine) for CGG (arginine) in the sequence CAG<u>C</u>TT, the mutability of the underlined C nucleotide (not involved <sup>167</sup> in the substitution) would decrease approximately 13-fold [21].

To evaluate the strength of positive selection in the B cell lineages, we used 169 BASELINE [32] to quantify deviations in the frequency of non-synonymous substitutions 170 from its expected value in the absence of selection (and under the mutational biases 171 captured by the S5F model). In line with previous analyses of the same datasets [20] 172 and of B cell lineages from healthy donors [33], we detected positive selection in the 173 CDRs of four of the seven heavy and light chains (lineages CH103 and VRC26) and 174 purifying selection in the FRs of all heavy and light chains. Framework regions had 175 fewer non-synonymous substitutions than expected under neutral evolution, while CDRs 176 had more non-synonymous substitutions than the neutral expectation (Figure 3). This 177 result contrasts with repertoire-level analyses showing predominantly purifying selection 178 across both types of regions [34]. 179

180

181

182

183

184

185

186

187

188

189

190

191

To estimate how much of the mutability losses in the CDRs is attributable to selection for amino acid substitutions, we partitioned inferred changes in mutability into changes caused by synonymous substitutions and changes caused by non-synonymous substitutions. Summed across all branches of the maximum-clade-credibility (MCC) trees, non-synonymous substitutions caused a net loss in the average CDR mutability of the same four heavy and light chains where positive selection occurred (Figure 4a). Similarly, synonymous substitutions caused a net decrease in the average CDR mutability of the four lineages whose CDRs were under positive selection. On average across those four lineages, non-synonymous substitutions accounted for approximately 66% of the inferred mutability loss in CDRs (range: 40% - 93%). On average, selection for amino acid substitutions may therefore have contributed to as much as 66% of the loss of mutability in the CDRs.

To test if the observed synonymous mutability losses were consistent with the decay 192 of mutational hotspots under their predicted S5F mutability, we simulated B cell 193 receptor evolution under a model that allows for variation in mutation rates across 194 motifs based on their S5F mutability scores [21] and thus captures the propensity for 195 certain motifs to mutate (Methods). We compared changes in mutability caused by 196 synonymous substitutions simulated under the S5F-based model to changes under 197 models that do not allow for motif-driven mutation rate variation. Instead, these models 198 assume that the mutation rate is identical across all sites ("uniform" model) or depends 199 on the position within a codon ("codon-position" model), with the relative rate for each 200 position estimated from the data (Methods). For each branch on the MCC tree of each 201 lineage, we started from the inferred nucleotide sequence of the branch's parent node 202 and simulated 100 replicates of a descendant sequence while constraining the simulations 203 to produce the exact same amino acid substitutions inferred from the data. We also 204 constrained simulated descendant sequences to have the same number of synonymous 205 substitutions as inferred from the data but allowed them to occur at different positions. 206



Fig 4. Changes in S5F mutability due to synonymous and non-synonymous substitutions in anti-HIV B cell lineages. A) Changes in mutability caused by synonymous and non-synonymous substitutions, summed across all branches of B cell genealogies. B) Changes in mutability due to synonymous substitutions, averaged across all branches. Gray indicates values inferred from the data, red indicates an S5F-based model where different nucleotide motifs mutate with different rates, dark blue indicates a model with no mutation rate variation across sites, and light blue indicates a model with different mutation rates for each position of a codon. Simulations were performed independently for each branch on the MCC tree of different anti-HIV B cell lineages, starting from the inferred sequence of the parent node. Each simulated sequence was constrained to have the same amino acid sequence as the inferred sequence for the descendant node. Vertical bars indicate the 95% range obtained from 100 simulations per model.



Fig 5. Frequency of losses relative to the total number of changes in whole-sequence S5F-mutability caused by synonymous substitutions during the evolution of anti-HIV B cell lineages. Blue indicates changes on terminal branches, and black indicates changes on internal branches. Results are shown separately for framework regions (FRs) and complementarity determining regions (CDRs). Each point denotes the frequency of changes in mutability that were losses, averaged across a sample of 1000 trees from the posterior distribution. Vertical lines indicate the 95% highest-posterior density interval.

Inferred mutability losses due to synonymous substitutions were consistent with the decay of mutational hotspots. On average across all branches, synonymous substitutions simulated under hotspot decay resulted in mutability losses in all seven heavy and light chains (Figure 4b). In contrast, synonymous substitutions simulated with constant mutation rates across motifs (uniform and codon-position models) increased mutability in all lineages.

207

208

209

210

211

212

213

220

221

222

223

224

225

226

227

228

No evidence of indirect selection to retain mutability

Under persistent or recurrent selective pressures, alleles that increase the mutation rate may be indirectly selected due to their association with beneficial mutations [25–28]. Indirect selection for increased mutation rates might therefore lead to the conservation of highly mutable motifs in CDRs, where mutations that improve affinity are selected for during B cell evolution. In contrast, selection to reduce the frequency of deleterious mutations [23, 24] might directly favor mutability losses in FRs. 219

To test if mutability is subject to direct or indirect selection over the long term, we compared the frequency of synonymous changes in mutability between terminal and internal branches of the B cell trees. Terminal branches are expected to be enriched for deleterious mutations [35–38]. If mutability losses are deleterious in the long term, branches where mutability losses occurred should be less likely to contribute descendants to the B cell population, and would therefore be more likely terminal, rather than internal branches. In contrast, mutability losses should be more frequent on internal branches if mutability losses are beneficial. Because changes in mutability due to non-synonymous substitutions may be driven by selection for B cell receptor affinity

and stability, we first restricted the analysis to mutability changes from synonymous substitutions.

229

230

231

232

233

234

235

236

237

238

239

240

241

242

243

244

245

246

247

248

249

250

251

252

253

254

260

We found no consistent evidence of selection to increase, retain, or decrease mutability (Figure 5). In one instance (the CDRs of light chain VRC26), mutability losses were more frequent on terminal branches than on internal branches, possibly indicating selection against mutability losses. In another instance (FRs of heavy chains VRC26), mutability losses were more frequent on internal branches than on internal branches, suggesting selection for mutability losses. In the remaining cases, the frequencies of synonymous mutability losses were similar in terminal and internal branches for both FRs and CDRs. The same lack of general trends was seen when we compared terminal branches, internal branches that belong to the main "backbone" of the tree, and other internal branches [38] (Figure S7), and also when we included mutability changes due to non-synonymous substitutions (results not shown).

#### Results are consistent for three of four mutability metrics

In addition to the S5F model [21], we repeated the analyses using three other mutability metrics. The different mutability metrics were estimated from databases of somatic mutations not subject to the effect of selection, such as synonymous mutations, mutations in non-coding flanking regions of V genes, and mutations in unproductive B cell receptor genes that are not expressed by B cells but still undergo somatic hypermutation. Two of the alternative metrics are based on a discrete classification of DNA motifs into either hotspots or regular motifs: the number of WRCH and DGYW hotspots [7] or the number of "overlapping" (WGCW) hotspots [8], where  $W = \{A/T\}$ ,  $R = \{A/G\}, H = \{A/T/C\}, D = \{A/T/G\}$  and  $Y = \{C/T\}$ . We also quantified mutability using the 7-mer model, which, similar to the S5F model, assigns relative mutation rates to different motifs on a continuous scale, but does so for seven- instead of five-nucleotide motifs [22].

Consistent with the results for S5F mutability, hotspots and overlapping hotspots decreased in number over time (Figure S8 and Figure S9) and were lost more frequently than gained (Figure S11). In contrast, average 7-mer mutability increased over time in four of five lineages (Figure S10), and losses of 7-mer mutability were approximately as frequent as gains across the entire B cell receptor sequence (fig. S11).

## Discussion

We used phylogenetic methods to reconstruct changes in B cell receptor sequences and 261 found consistent loss of mutational hotspots over the course of the adaptive immune 262 response. Our analyses shed light on previous studies of long-term trends in 263 mutability [17, 20] by quantifying the contributions of different mechanisms to these 264 losses. Selection for amino acid substitutions appears to have driven most of the 265 mutability losses in the CDRs—precisely the regions where high mutation rates 266 contribute the most to adaptation. However, mutability changes caused by synonymous 267 substitutions throughout the sequence demonstrate mutability losses also arose from the 268 spontaneous decay of highly mutable motifs through neutral mutations. We also show 269 that both FRs and CDRs lost mutability frequently, and the higher mutability of CDRs 270 relative to FRs observed in ancestral receptors persisted over at least several years of B 271 cell evolution. Disruption of highly mutable motifs through selection and hotspot decay 272 therefore seems to be intrinsic to the evolution of B cells during affinity maturation. 273 These factors counteract the high mutation rate selected in the evolution of 274 immunoglobulin genes, and they suggest adaptability may inevitably become 275 compromised in evolving B cell lineages. However, the relatively high mutation rate in 276 the CDRs compared to the FRs was preserved in these lineages, suggesting that some degree of adaptability may still be maintained. (These observations may reflect survival bias: lineages in which mutability differences between CDRs and FRs decrease over time may be less likely to persist in the long term.)

277

278

279

280

281

282

283

284

285

286

287

288

289

290

291

292

293

294

295

296

297

298

299

300

301

302

303

304

305

306

307

308

309

310

311

312

313

314

315

316

317

318

319

320

321

322

Theoretical simulations [26] and experimental evolution of bacteria [27, 28] suggest recurrent adaptation in asexual populations can select for increased mutation rates over the long term, but we found no consistent evidence of selection to increase or retain mutability during B cell evolution. Any long-term benefits of high mutation rates (in terms of increased chance of producing beneficial mutations) appear insufficient to overcome the rapid decay of highly mutable motifs through mutation and/or positive selection on amino acid substitutions. We also did not detect selection for reduced mutability in FRs, despite the potential for mutability losses to decrease the rate of destabilizing mutations. Because mutability is already low in the FRs of V genes [12,15] and ancestral B cell receptors (this study), it is possible that the fitness effect of mutability losses in FRs is so small that they are nearly neutral given the effective size of the B cell population. Under this "drift-barrier" hypothesis [24], selection would therefore not be able to fix mutability losses.

Losses of mutational hotspots in CDRs may reduce the adaptability of antibody responses since lower mutation rates reduce the supply of genetic variation available for selection. Declines in mutation rates caused by hotspot losses may have contributed to reported declines in the evolutionary rates of broadly neutralizing B cell lineages during chronic HIV infection [20]. However, these losses may not be important for lineages that bind to conserved sites. As B cell receptors evolve high affinity for conserved sites, beneficial substitutions become rare, and the corresponding transition from positive to purifying selection should cause substitution rates to fall [20]. In line with the previous analysis [20], we found the longest-lived lineages to be under the strongest purifying selection in the CDRs. Those lineages also had minimal cumulative changes in mutability, suggesting most of the mutability loss occurred during early adaptation. However, as a result of hotspot loss, these lineages might adapt poorly if formerly conserved antigenic sites suddenly acquire mutations.

A direct link between motif-based mutability scores and absolute rates of B cell mutation in vivo is still lacking. Hotspot loss is associated with reduced substitution rates [19,20], which are influenced not only by the underlying mutation rate but also by changes in selective pressures and generation times [20]. Using robust counting [39] and a random local clock model [40], we found no evidence for consistent declines in the synonymous or non-synonymous substitution rates of these lineages (Figures S12 to S15). Simulation-based power analysis (Supplementary Methods) revealed that synonymous and non-synonymous substitution rates had to decline at least 50% to be detected with this method (Figures S16 and S17). Observed long-term declines in average S5F mutability were less than 50% (Figure S2) and suggest their effects on substitution rates would not have been detected, assuming a one-to-one relationship between S5F mutability and the actual mutation rate. Additionally, inference of B cell genealogies may be affected by non-equilibrium nucleotide frequencies, distorting inferences of absolute rates [17]. Analyzing B cell sequences from controlled experimental infections with customized substitution models (e.g., [17, 34, 41]) could help quantify the relationship between mutability metrics and absolute mutation rates.

Our results suggest a trade-off between the short-term and long-term adaptability of antibody responses. Repeated infections by antigenically related pathogens, such as influenza viruses, often recall memory B cell populations [42–45]. Preferential recruitment of less mutable B cells at the expense of naive (or perhaps younger memory) B cells may compromise the long-term adaptability of the immune repertoire to these pathogens. Protection against such pathogens may rely on a robust naive response that

> can complement preexisting B cell lineages once they fail to adapt to new antigens. Finally, some strategies for eliciting universal responses against HIV and influenza attempt to recapitulate the evolution of long-lived, highly mutated B cell lineages that were found to produce broadly neutralizing antibodies in infected patients [46]. Our results suggest those approaches may be hindered by a decrease in the adaptability of highly mutated antibodies. 329

# Materials and Methods

#### Sequence data

We analyzed BCR sequences from three published studies of B cell lineages sampled longitudinally in individual HIV-1 patients not subject to antiretroviral therapy. The CH103 lineage comprises broadly neutralizing antibodies (bnAbs) isolated from individual B cells and clonally related sequences bioinformatically isolated from high-throughput sequencing over 144 weeks of HIV-1 infection in a single donor [29]. Likewise, the VRC26 lineage comprises both bnAbs isolated from individually sorted B cells and clonally related heavy-chain sequences obtained from high-throughput sequencing over 206 weeks of HIV-1 infection in a different patient [30]. In both cases, BCR sequences obtained from high-throughput sequencing were classified as clonally related to the isolated antibodies based on V and J gene usage. A third lineage, VRC01, was sampled longitudinally from a third HIV patient for 15 years starting from approximately five years after the date of infection. VRC01 is a large lineage that possibly consists of multiple independent B cell lineages [19,20]. We therefore used PARTIS, a recently developed hidden-markov-model-based method [47] to partition the heavy-chain VRC01 dataset into sets of sequences likely to have descended from the same naive B cell, and to identify their most likely germline V and J genes. We analyzed the three largest lineages identified in this way, hereafter VRC01-13, VRC01-01 and VRC01-19. We aligned each set of sequences in MACSE v1.01b [48] along with their concatenated V and J genes (for the heavy chains) or along with their V genes only (light chains) (Table 1). Including the J genes in the light chain datasets produced bad alignments across the J region.

**Table 1.** BCR alignments analyzed. The number of sequences includes the V+J or V germline sequences.

Lineage (chain)	V gene	J gene	N sequences	N sites
CH103 (heavy)	IGHV4-59*01	IGHJ4*01	460	321
CH103 (light)	IGLV3-1*01	_	175	285
VRC26 (heavy)	IGHV3-30*18	IGHJ3*01	681	489
VRC26 (light)	IGLV1-51*02	_	464	294
VRC01-13 (heavy)	IGHV1-2*04	IGHJ1*01	157	351
VRC01-01 (heavy)	IGHV-ORF15-1*04	IGHJ4*03	124	372
VRC01-19 (heavy)	IGHV1-2*04	IGHJ1*01	110	438

### Phylogenetic inference

Using BEAST v.1.8.2 [49], we fit Bayesian phylogenetic models to the BCR sequence data in order to estimate time-resolved genealogies and internal node sequences for heavy and light chains separately. We used a GTR nucleotide substitution model, assumed a random local clock model to account for potential variation in substitution rates [40], and enabled robust counting [39] to estimate the numbers of synonymous and non-synonymous changes on each branch. To reduce the number of parameters, we used

335

336

337

338

339

340

341

342

343

344

345

346

347

348

349

350

351

352

353

354

355

356

357

empirical base frequencies and assumed a shared nucleotide transition matrix across 365 codon positions for the robust counting inference. The inferred dynamics of mutability 366 losses and gains were qualitatively robust to the choice of demographic model (constant 367 population size, logistic or exponential growth) used to calculate coalescent prior 368 probabilities (Supplementary Methods). We set the V+J or V germline sequence of each 369 alignment as the outgroup and assigned them a sampling time of zero to represent the 370 assumption that, at the start of the infection, the ancestral BCR sequence was close to 371 its corresponding germline genes (except for insertions and deletions at the junctions). 372 For each dataset, four independent MCMC chains were set to run for 500 million steps 373 and sampled every 1,000 steps (for parameter values) and every 10,000 steps (for trees). 374 We downsampled the set of trees recovered by the MCMC chains to obtain 1.000 trees 375 per chain, sampled at regular intervals between the end of the the burn-in and the end 376 of the entire chain. 377

Because of the long computation times for the larger datasets (lineages VRC26 and CH103), their MCMC chains were interrupted before 500 million steps had been reached. With the exception of heavy-chain VRC26, interrupted chains had ESSs close to or greater than 200 for most parameters and for the likelihood, prior and posterior probabilities. Although MCMC chains failed to converge for the heavy chain data of lineage VRC26, estimates of the parameters of interest (mean mutability changes and fraction of mutability losses across the tree) were numerically close across replicate MCMC chains, and we therefore present the results in the main text.

To investigate their ability to detect consistent changes in mutability, we repeated the phylogenetic analyses on alignments simulated using different population genetic models (Supplementary Methods). For alignments simulated under a model where all sites are equally likely to mutate, mutability losses were estimated to be approximately as frequent as gains. As expected, mutability losses were estimated to be more frequent than gains in datasets simulated under an S5F-based model, where motifs with high S5F score have a high probability of mutating.

#### Simulations of sequence evolution on MCC trees

For each branch on an MCC tree, we compared the nucleotide sequences of the parent and descendant nodes to identify the number of codon sites with non-synonymous and synonymous differences. Starting from the parent sequence, we simulated an alternative descendant sequence constrained to have 1) the same number of non-synonymous and synonymous differences from the ancestor and 2) the same amino acid sequence as the true descendant.

Each simulation was done in two steps. First, for each codon site with a non-synonymous difference, we identified all codons coding for same amino acid and with the same number of nucleotide differences from the ancestral codon as the true descendant codon. We used different models to compute the probability of each possible descendant codon. Let A be the ancestral codon and D one of the possible descendant codons. For the case where A and D differ by a single nucleotide, we define the transition rate from A to D as:

$$R_{A \to D} = \mu_i \times P(a_i \to d_i | \text{ mutation at } i) \tag{1}$$

where  $i \in \{1, 2, ..., L\}$  is the position where codon A differs from codon D, L is the length of the nucleotide sequence,  $\mu_i$  is the mutation rate at position i, and  $a_i$  and  $d_i$ are the nucleotides at position i in A and D, respectively. For the case where the transition from A to D requires more than one nucleotide substitution, we define  $R_{A \to D}$ as the product of the right-hand side of Equation (1) applied to all intermediary one-nucleotide steps between A and D, summed across all the possible orders in which

393

378

379

380

381

382

383

384

385

386

387

388

389

390

391

the required mutations can be introduced. For example:

$$R_{AAC \to AGT} = R_{AAC \to AGC} \times R_{AGC \to AGT} + R_{AAC \to AAT} \times R_{AAT \to AGT}$$
(2)

We simulate the evolution of a descendant sequence by sampling a codon from the set of possible descendant codons according to their transition rates from the ancestral codon. The probability of sampling codon  $D_k$  is given by: 416

$$P(D_k) = \frac{R_{A \to D_k}}{\sum_j R_{A \to D_j}} \tag{3}$$

where k and j index the possible codons.

After simulating evolution at all codon sites where non-synonymous substitutions ocurred, we introduced the same number of synonymous mutations as the number of synonymous codon differences between the ancestral sequence and the true descendant sequence. We allowed synonymous mutations to occur at any codon site except for sites where non-synonymous mutations occurred. To introduce each synonymous mutation, we randomly sampled a nucleotide site, replaced its nucleotide, and either kept the mutation (if it was synonymous) or rejected it (if it was non-synonymous). To simplify implementation, we did not allow for synonymous differences involving more than one nucleotide difference to be simulated – once a codon site was hit by a synonymous mutation, nucleotide sites in that codon could no longer be sampled for subsequent mutations. On average across all datasets, 99.1% of inferred synonymous changes involved a single nucleotide difference (min. 96.1%, max. 99.8%).

Using different models to parameterize Equation (1) and to sample sites for synonymous mutations, we created scenarios where the mutation rate at a site either depends on the nucleotide motif centered at that site, or is independent of nucleotide motifs (see below). To model nucleotide transitions, we made use of the fact that, in addition to estimating the relative mutation rate of different motifs, the S5F model includes the probability of transitions between nucleotides, given the motif where a mutation occurs. All simulations used the S5F-based nucleotide transition probabilities from [21], regardless of whether they assumed variable or constant mutation rates across motifs.

To simulate sequence evolution under an S5F-based model, we let  $\mu_i$  in Equation (1) be given by:

$$_{i} = S5F(S_{i}) \tag{4}$$

where  $S5F(S_i)$  is the S5F-score of the five nucleotide motif  $S_i$  centered at *i*, from [21].

μ

In addition to the S5F-based model, we performed simulations under two models 442 that assume mutation rates do not depend on the local nucleotide sequence around each 443 site. First, we simulated sequence evolution under a uniform model where all sites are 444 equally likely to mutate ( $\mu_i = 1$ , regardless of the nucleotide occupying *i* and the 445 neighboring positions). Second, we used a codon-position-based model, where codon 446 positions 1, 2 and 3 have different mutation rates. Estimates of the relative mutation 447 rates of each codon position were obtained from the robust counting inference 448 performed in BEAST and used to parameterize the simulations. 449

### Acknowledgements

This work was supported in part by the National Institute of Allergy and Infectious 451 Diseases (NIAID) of the National Institutes of Health National (DP2AI117921) and by 452 a Complex Systems Scholar Award from the James S. McDonnell Foundation. The 453 content is solely the responsibility of the authors and does not necessarily represent the 454

413

417 418

419

420

421

422

423

424

425

426

427

428

429

430

431

432

433

434

435

436

437

439

440

441

official views of the NIAID or the National Institutes of Health. This work was completed in part with resources provided by the University of Chicago Research Computing Center. Code implementing the analyses and figures is available at https://github.com/cobeylab/evolution\_of\_mutability.

455

456

457

458

### References

- Teng G, Papavasiliou FN. Immunoglobulin somatic hypermutation. Annual Review of Genetics. 2007;41:107–120. doi:10.1146/annurev.genet.41.110306.130340.
- 2. Griffiths GM, Berek C, Kaartinen M, Milstein C. Somatic mutation and the maturation of immune response to 2-phenyl oxazolone. Nature. 1984;312:271–275.
- MaClennan ICM, Gray D. Antigen-Driven Selection of Virgin and Memory B Cells. Immunological Reviews. 1986;91(1):61–86. doi:10.1111/j.1600-065X.1986.tb01484.x.
- Liu YJ, Joshua DE, Williams GT, Smith CA, Gordon J, MacLennan ICM. Mechanism of antigen-driven selection in germinal centres. Nature. 1989;342(6252):929–931. doi:10.1038/342929a0.
- Victora GD, Nussenzweig MC. Germinal Centers. Annual Review of Immunology. 2012;30(1):429–457. doi:10.1146/annurev-immunol-020711-075032.
- Rogozin IB, Kolchanov NA. Somatic hypermutagenesis in immunoglobulin genes. Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression. 1992;1171(1):11–18. doi:10.1016/0167-4781(92)90134-L.
- Rogozin IB, Diaz M. Cutting Edge: DGYW/WRCH Is a Better Predictor of Mutability at G:C Bases in Ig Hypermutation Than the Widely Accepted RGYW/WRCY Motif and Probably Reflects a Two-Step Activation-Induced Cytidine Deaminase-Triggered Process. The Journal of Immunology. 2004;172(6):3382–3384. doi:10.4049/jimmunol.172.6.3382.
- Wei L, Chahwan R, Wang S, Wang X, Pham PT, Goodman MF, et al. Overlapping hotspots in CDRs are critical sites for V region diversification. Proceedings of the National Academy of Sciences. 2015; p. E728–E737. doi:10.1073/pnas.1500788112.
- Levy Y, Gupta N, Le Deist F, Garcia C, Fischer A, Weill JC, et al. Defect in IgV gene somatic hypermutation in Common Variable Immuno-Deficiency syndrome. Proceedings of the National Academy of Sciences. 1998;95(13135-13140).
- Bonhomme D, Hammarström L, Webster D, Chapel H, Hermine O, Le Deist F, et al. Impaired antibody affinity maturation process characterizes a subset of patients with common variable immunodeficiency. Journal of Immunology. 2000;165(8):4725–30.
- 11. Frasca D, Blomberg BB. Effects of aging on B cell function. Current Opinion in Immunology. 2009;21(4):425–430. doi:10.1016/j.coi.2009.06.001.
- Oprea M, Kepler TB. Genetic Plasticity of V Genes Under Somatic Hypermutation : Statistical Analyses Using a New Resampling-Based Methodology. Genome Research. 1999;919:1294–1304.

- Wagner SD, Milstein C, Neuberger MS. Codon bias targets mutation. Nature. 1995;376(6543):732–732. doi:10.1038/376732a0.
- Kepler TB. Codon bias and plasticity in immunoglobulins. Molecular Biology and Evolution. 1997;14(6):637–643. doi:10.1093/oxfordjournals.molbev.a025803.
- Saini J, Hershberg U. B cell Variable genes have evolved their codon usage to focus the targeted patterns of somatic mutation on the complementarity determining regions. Molecular Immunology. 2015;65(1):157–167. doi:10.1016/j.molimm.2015.01.001.
- 16. Hershberg U, Shlomchik MJ. Differences in potential for amino acid change after mutation reveals distinct strategies for kappa and lambda light-chain variation. Proceedings of the National Academy of Sciences of the United States of America. 2006;103(43):15963–8. doi:10.1073/pnas.0607581103.
- Hoehn KB, Lunter G, Pybus O. A Phylogenetic Codon Substitution Model for Antibody Lineages. Genetics. 2017;206:417–427.
- Stamper CT, Wilson PC. What Are the Primary Limitations in B-Cell Affinity Maturation, and How Much Affinity Maturation Can We Drive with Vaccination? Is Affinity Maturation a Self-Defeating Process for Eliciting Broad Protection? Cold Spring Harbor Perspectives in Biology. 2017; p. a028803. doi:10.1101/cshperspect.a028803.
- Wu X, Zhang Z, Schramm CA, Joyce MG, Do Kwon Y, Zhou T, et al. Maturation and Diversity of the VRC01-Antibody Lineage over 15 Years of Chronic HIV-1 Infection. Cell. 2015;161(3):470–485. doi:10.1016/j.cell.2015.03.004.
- Sheng Z, Schramm CA, Connors M, Morris L, Mascola JR, Kwong PD, et al. Effects of Darwinian Selection and Mutability on Rate of Broadly Neutralizing Antibody Evolution during HIV-1 Infection. PLOS Computational Biology. 2016;12(5):e1004940. doi:10.1371/journal.pcbi.1004940.
- 21. Yaari G, Vander Heiden JA, Uduman M, Gadala-Maria D, Gupta N, Stern JNH, et al. Models of somatic hypermutation targeting and substitution based on synonymous mutations from high-throughput immunoglobulin sequencing data. Frontiers in Immunology. 2013;4:358. doi:10.3389/fimmu.2013.00358.
- Elhanati Y, Sethna Z, Marcou Q, Callan CG, Mora T, Walczak AM. Inferring processes underlying B-cell repertoire diversity. Philosophical transactions of the Royal Society of London Series B, Biological sciences. 2015;370(1676):20140243. doi:10.1098/rstb.2014.0243.
- 23. Kimura M. On the evolutionary adjustment of spontaneous mutation rates. Genetics Research. 1967;9:23–34.
- Lynch M, Ackerman MS, Gout JF, Long H, Sung W, Thomas WK, et al. Genetic drift, selection and the evolution of the mutation rate. Nature Reviews Genetics. 2016;17(11):704–714. doi:10.1038/nrg.2016.104.
- Leigh EG. Natural Selection and Mutability. The American Naturalist. 1970;104(937):301–305. doi:10.1086/282663.
- Taddei F, Radman M, Maynard-Smith J, Toupance B, Gouyon PH, Godelle B. Role of mutator alleles in adaptive evolution. Nature. 1997;387(6634):700–702. doi:10.1038/42696.

- 27. Giraud A, Matic I, Tenaillon O, Clara A, Radman M, Fons M, et al. Costs and Benefits of High Mutation Rates: Adaptive Evolution of Bacteria in the Mouse Gut. Science. 2001;291(5513).
- Pal C, Maciá MD, Oliver A, Schachar I, Buckling A. Coevolution with viruses drives the evolution of bacterial mutation rates. Nature. 2007;450(7172):1079–1081. doi:10.1038/nature06350.
- Liao HX, Lynch R, Zhou T, Gao F, Alam SM, Boyd SD, et al. Co-evolution of a broadly neutralizing HIV-1 antibody and founder virus. Nature. 2013;496(7446):469–476. doi:10.1038/nature12053.
- Doria-Rose NA, Schramm CA, Gorman J, Moore PL, Bhiman JN, DeKosky BJ, et al. Developmental pathway for potent V1V2-directed HIV-neutralizing antibodies. Nature. 2014;509(7498):55–62. doi:10.1038/nature13036.
- Nakamura Y, Gojobori T, Ikemura T. Codon usage tabulated from international DNA sequence databases: status for the year 2000. Nucleic acids research. 2000;28(1):292.
- Yaari G, Uduman M, Kleinstein SH. Quantifying selection in high-throughput Immunoglobulin sequencing data sets. Nucleic Acids Research. 2012;40(17):e134. doi:10.1093/nar/gks457.
- 33. Yaari G, Benichou JIC, Vander Heiden JA, Kleinstein SH, Louzoun Y. The mutation patterns in B-cell immunoglobulin receptors reflect the influence of selection acting at multiple time-scales. Philosophical Transactions of the Royal Society B: Biological Sciences. 2015;370(1676).
- McCoy CO, Bedford T, Minin VN, Bradley P, Robins H, Matsen FA. Quantifying evolutionary constraints on B-cell affinity maturation. Philosophical transactions of the Royal Society of London Series B, Biological sciences. 2015;370(1676):20140244–. doi:10.1098/rstb.2014.0244.
- Williamson S, Orive ME. The Genealogy of a Sequence Subject to Purifying Selection at Multiple Sites. Molecular Biology and Evolution. 2002;19(8):1376–1384. doi:10.1093/oxfordjournals.molbev.a004199.
- Kosakovsky Pond SL, Frost SDW, Grossman Z, Gravenor MB, Richman DD, Brown AJL. Adaptation to Different Human Populations by HIV-1 Revealed by Codon-Based Analyses. PLoS Computational Biology. 2006;2(6):e62. doi:10.1371/journal.pcbi.0020062.
- 37. Edwards CTT, Holmes EC, Pybus OG, Wilson DJ, Viscidi RP, Abrams EJ, et al. Evolution of the Human Immunodeficiency Virus Envelope Gene Is Dominated by Purifying Selection. Genetics. 2006;174(3).
- Lemey P, Kosakovsky Pond SL, Drummond AJ, Pybus OG, Shapiro B, Barroso H, et al. Synonymous substitution rates predict HIV disease progression as a result of underlying replication dynamics. PLoS computational biology. 2007;3(2):e29. doi:10.1371/journal.pcbi.0030029.
- O'Brien JD, Minin VN, Suchard MA. Learning to count: robust estimates for labeled distances between molecular sequences. Molecular biology and evolution. 2009;26(4):801–14. doi:10.1093/molbev/msp003.
- Drummond AJ, Suchard MA. Bayesian random local clocks, or one rate to rule them al. BMC Biology. 2010;8(114). doi:10.1080/10635150590947041.

- 41. Cui A, Di Niro R, Vander Heiden JA, Briggs AW, Adams K, Gilbert T, et al. A Model of Somatic Hypermutation Targeting in Mice Based on High-Throughput Ig Sequencing Data. The Journal of Immunology. 2016;197(9):3566–3574.
- 42. Wrammert J, Koutsonanos D, Li GM, Edupuganti S, Sui J, Morrissey M, et al. Broadly cross-reactive antibodies dominate the human B cell response against 2009 pandemic H1N1 influenza virus infection. The Journal of Experimental Medicine. 2011;208(1):181–193. doi:10.1084/jem.20101352.
- 43. Li GM, Chiu C, Wrammert J, McCausland M, Andrews SF, Zheng NY, et al. Pandemic H1N1 influenza vaccine induces a recall response in humans that favors broadly cross-reactive memory B cells. Proceedings of the National Academy of Sciences. 2012;109(23):9047–9052. doi:10.1073/pnas.1118979109.
- 44. Jiang N, He J, Weinstein JA, Penland L, Sasaki S, He XS, et al. Lineage Structure of the Human Antibody Repertoire in Response to Influenza Vaccination. Science Translational Medicine. 2013;5(171):171ra19. doi:10.1126/scitranslmed.3004794.
- 45. Andrews SF, Huang Y, Kaur K, Popova LI, Ho IY, Pauli NT, et al. Immune history profoundly affects broadly protective B cell responses to influenza. Science translational medicine. 2015;7(316):316ra192. doi:10.1126/scitranslmed.aad0522.
- Haynes BF, Kelsoe G, Harrison SC, Kepler TB. B-cell-lineage immunogen design in vaccine development with HIV-1 as a case study. Nature Biotechnology. 2012;30(5):423–433. doi:10.1038/nbt.2197.
- Ralph DK, Matsen FA. Consistency of VDJ Rearrangement and Substitution Parameters Enables Accurate B Cell Receptor Sequence Annotation. PLoS computational biology. 2016;12(1):e1004409. doi:10.1371/journal.pcbi.1004409.
- Ranwez V, Harispe S, Delsuc F, Douzery EJP, Beal K. MACSE: Multiple Alignment of Coding SEquences Accounting for Frameshifts and Stop Codons. PLoS ONE. 2011;6(9):e22594. doi:10.1371/journal.pone.0022594.
- Drummond AJ, Suchard MA, Xie D, Rambaut A. Bayesian Phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology and Evolution. 2012;29(8):1969–1973. doi:10.1093/molbev/mss075.

# Supplementary information

### Simulation of B cell receptor alignments

To assess the power of a random local clock model [40] combined with robust counting [39] to detect changes in synonymous and non-synonymous substitution rates during B cell evolution, we used those methods to analyze B cell alignments simulated under different scenarios of change in the underlying mutation rates:

Scenario 1: uniform mutability, constant mutation rate.

Scenario 2: uniform mutability, decreasing mutation rate.

- 2a: 10% decrease in rate over simulation period
- 2b: 20% decrease in rate
- 2c: 30% decrease in rate
- 2d: 40% decrease in rate
- 2e: 50% decrease in rate

**Scenario 3**: context-dependency in mutation rates and / or transition probabilities. 3a: S5F context dependency in mutation rates and transition probabilities

3b: hotspot context dependency in mutation rates, uniform transition probabilities, hotspots 3x more likely to mutate

3c: hotspot context dependency, hotspots 30x more likely to mutate

For the simulations, we modified a simple forward-time Wright-Fisher model to impose a fitness cost on non-synonymous mutations. The status of a mutation (synonymous or non-synonymous) is defined relative to a fixed reference sequence. As a reference sequence, we used the heavy chain sequence at the root node of lineage CH103, inferred under a logistic growth prior (98-100% identical to sequences inferred in four replicate chains under a constant population size prior). An initial population is generated consisting of a single copy of the reference sequence. At each subsequent generation t,  $N_t$  sequences are produced by replicating sequences from the previous generation with the possibility of mutation. We assumed N grows logistically:

$$N_t = N_{t-1} + rN_{t-1} \left( 1 - \frac{N_{t-1}}{K} \right)$$
(5)

with  $N_0 = 1$  and  $N_t$  rounded to the nearest integer. Logistic growth assumes the B cell population initially expands exponentially and then saturates at the carrying capacity K.

The probability that a newly generated sequence at generation t descends from sequence i in generation t-1 is equal to the fitness of sequence i,  $w_i$ , normalized by the sum of fitness values across the entire population at t-1. Any sequence i whose amino acid translation is the same as that of the reference sequence has fitness  $w_i = 1$ . Each non-synonymous mutation relative to the reference sequence adds s to the value of  $w_i$ , with negative values of s representing a fitness cost ( $w_i$  cannot go below 0). Newly generated sequences undergo mutations at fixed or variable rates, depending on the scenario.

#### Modeling relative and absolute mutation rates

Different models have been proposed to describe variation in somatic hypermutation rates across different nucleotide motifs [7,21,22]. For site *i* in sequence *j*, those models can be used to assign a *relative* mutation rate  $m_{i,j}$  to site *i*, based on its local sequence context in sequence *j*. However, the precise relationship between the *relative* mutability of a site and the site's *absolute* mutation rate is unclear. To model absolute mutation rates as a function of the relative mutability of a sequence's motifs, we assume that the average mutation rate per site per generation for sequence  $j, \overline{\mu}_j$ , is proportional to the sequence's average relative mutability,  $\overline{m}_j$ :

$$\overline{\mu}_j = k \times \overline{m}_j \tag{6}$$

Let  $\overline{m}_0$  be the average relative mutability of the reference sequence. We choose k so that the reference sequence has an average mutation rate per site per generation,  $\overline{\mu}_0$ , equal to 1/4L, where L is the sequence length (in number of nucleotides):

$$k = \frac{1}{\overline{m}_0 4L} \tag{7}$$

In preliminary simulations under default values for other parameters (see below), this choice of  $\overline{\mu}_0 = 1/4L$  produced alignments visually similar to those observed for real B cell lineages in terms of overall nucleotide diversity. For any sequence j, the average mutation rate per site per generation is then given by:

$$\overline{\mu}_j = \frac{\overline{m}_j}{\overline{m}_0 4L} \tag{8}$$

Thus, a sequence with half the average relative mutability of the reference sequence has half the average mutation rate per site per generation. Finally, site-specific mutation rates per generation for sequence j are given by:

$$\mu_{i,j} = m_{i,j} \times k = \frac{m_{i,j}}{\overline{m}_0 4L} \tag{9}$$

Note that scaling site-specific mutation rates to 1/4L does not change the relative mutability of the sites. For example, for sites a and b in the same sequence j:

$$\frac{\mu_{a,j}}{\mu_{b,j}} = \frac{m_{a,j}}{m_{b,j}}$$

To models scenarios 1 and 2 (where mutation rates are independent of sequence context) we let  $\mu_{i,j}(t) = c(t) \times 1/4L$  be the mutation rate per time per generation for all sites in all sequences at generation t, where  $0 \le c(t) \le 1$ . In scenario 1 we let c(t) = 1 for all t, whereas in scenario 2 we choose decreasing values of c(t) for different time intervals

In the S5F-based parameterization (scenario 3a), we set  $m_{i,j}$  to the relative mutability scores from [21], based on a five-nucleotide window centered on site *i*. The first two sites and the last two sites, for which S5F is indeterminate since the neighbors are unknown, are assigned mutability zero. We used motif-specific transition probabilities between nucleotides inferred by the S5F model along with motif-specific mutation rates.

In the "hotspot" parameterization (scenarios 3b-c), we let  $m_{i,j}$  be either 1, if site *i* is not at the central position of a hotspot, or *h*, if it is. Site *i* is at the central position of a hotspot if it is occupied by the underlined nucleotide in a  $WR\underline{C}H$  or a  $D\underline{G}YW$  motif, where  $W = \{A/T\}$ ,  $R = \{A/G\}$ ,  $H = \{A/T/C\}$ ,  $D = \{A/T/G\}$  and  $Y = \{C/T\}$  [7]. A site is assigned mutability 1 f it cannot be determined whether or not that site is at the center of a WR<u>C</u>H or a D<u>G</u>YW motif (for example, the left neighbors of the first site in a sequence and the right neighbors of the last site in a sequence are unknown). We assumed uniform transition probabilities between nucleotides.

We ran all simulations with s = -0.01, r = 0.7 and K = 1000 for 2000 generations, sampling 25 sequences every 250 generations starting at generation 500. We decreased c(t) by 0.2 every 250 generations starting at generation t = 1000 for scenario 2a, by 0.4 every 250 generations starting at generation t = 1000 for scenario 2b, by 0.5 every 250

generations starting at generation t = 750 for scenario 2c, by 0.8 every 250 generations starting at generation t = 1000 for scenario 2d, and by 0.1 every 250 generations starting at generation t = 1000 for scenario 2d.

We analyzed the simulated alignments using BEAST v.1.8.2 to test if declines in synonymous and non-synonymous substitution rates are correctly detected in scenarios 2 and 3.

#### Changes in substitution rates over time

For each tree in the posterior distributions inferred for observed and simulated alignments, we computed the estimated synonymous and non-synonymous substitution rates for each branch by dividing estimated counts of synonymous and non-synonymous substitutions (obtained by robust counting) by the branch's length measured in time. For each tree in the posterior distributions of simulated alignments, we estimated pairwise linear regression coefficients between branch times (predictor variable, measured at each branch's parent node) and total, synonymous and non-synonymous substitution rates (response variables).

## Supplementary figures



**Fig S1.** Mutability of the inferred ancestral sequences of long-lived B cell lineages (red squares) compared with the distribution of mutability values obtained by randomizing the ancestral codon sequence while keeping the amino acid sequence constant. Results are shown for the whole sequences (WS) and separately for framework regions (FRs) and complementarity determining regions (CDRs).





g) VRC01-19 heavy chain

Fig S2. Evolution of S5F-mutability in long-lived B cell lineages. Scatterplots show S5F-mutability over time for nodes from a sample of 100 trees from the posterior distribution. Blue points correspond to terminal nodes (observed sequences), and black points correspond to internal nodes whose sequences were inferred statistically. The red line represents an average of regression lines calculated for each tree in a sample of 1000 trees. Solid lines indicate instances indicate significant relationships (where the highest-posterior density of the slope did not overlap zero), and dashed lines indicate non-significant relationships.







Fig S3. Evolution of S5F-mutability in the framework regions (FRs) of long-lived B cell lineages. Scatterplots show S5F-mutability over time for nodes from a sample of 100 trees from the posterior distribution. Blue points correspond to terminal nodes (observed sequences), and black points correspond to internal nodes whose sequences were inferred statistically. The red line represents an average of regression lines calculated for each tree in a sample of 1000 trees. Solid lines indicate instances indicate significant relationships (where the highest-posterior density of the slope did not overlap zero), and dashed lines indicate non-significant relationships.







**Fig S4.** Evolution of S5F-mutability in the complementarity determining regions (CDRs) of long-lived B cell lineages. Scatterplots show S5F-mutability over time for nodes from a sample of 100 trees from the posterior distribution. Blue points correspond to terminal nodes (observed sequences), and black points correspond to internal nodes whose sequences were inferred statistically. The red line represents an average of regression lines calculated for each tree in a sample of 1000 trees. Solid lines indicate instances indicate significant relationships (where the highest-posterior density of the slope did not overlap zero), and dashed lines indicate non-significant relationships.



**Fig S5.** Evolution of the difference in mutability between complementarity determining regions (CDRs) and framework regions (FRs) in long-lived B cell lineages. The relative difference is calculated as the average mutability of CDRs minus the average mutability of FRs, divided by the average mutability of FRs. Each point corresponds to a node in the maximum-clade-credibility tree of each lineage. Linear regression lines are shown in red.



Fig S6. Mutability of B cell receptor sequences from different B cell lineages relative to the expected distribution of mutability values obtained by randomizing codon sequences while keeping the amino sequences constant. The distribution of mutability percentiles obtained for sequences sampled at the last sampling time point in each dataset is shown in gray. The mutability percentile of each lineage's ancestor is shown in red. Results are shown separately for framework regions (FR) and complementarity determining regions (CDRs).



Fig S7. Frequency of mutability losses relative to the total number of changes in S5F mutability caused by synonymous substitutions during the evolution of anti-HIV B cell lineages. Blue indicates changes that on terminal branches, orange indicates changes along the trunk of the tree, and black indicates changes on the remaining internal branches. Results are shown separately for framework regions (FRs) and complementarity determining regions (CDRs). Each point denotes the frequency of changes in mutability that were losses, averaged across a sample of 1000 trees from the posterior distribution. Vertical lines indicate the 95% highest-posterior density interval.







Fig S8. Evolution of the number of WRCH/DGYW hotspots in long-lived B cell lineages. Scatterplots show the number of hotspots over time for nodes from a sample of 100 trees from the posterior distribution. Blue points correspond to terminal nodes (observed sequences), and black points correspond to internal nodes whose sequences were inferred statistically. The red line represents an average of regression lines calculated for each tree in a sample of 1000 trees. Solid lines indicate instances indicate significant relationships (where the highest-posterior density of the slope did not overlap zero), and dashed lines indicate non-significant relationships.



g) VRC01-19 heavy chain

**Fig S9.** Evolution of the number of overlapping hotspots in long-lived B cell lineages. Scatterplots show the number of overlapping hotspots over time for nodes from a sample of 100 trees from the posterior distribution. Blue points correspond to terminal nodes (observed sequences), and black points correspond to internal nodes whose sequences were inferred statistically. The red line represents an average of regression lines calculated for each tree in a sample of 1000 trees. Solid lines indicate instances indicate significant relationships (where the highest-posterior density of the slope did not overlap zero), and dashed lines indicate non-significant relationships.





g) VRC01-19 heavy chain

Fig S10. Evolution of 7-mer mutability in long-lived B cell lineages. Scatterplots show 7-mer mutability over time for nodes from a sample of 100 trees from the posterior distribution. Blue points correspond to terminal nodes (observed sequences), and black points correspond to internal nodes whose sequences were inferred statistically. The red line represents an average of regression lines calculated for each tree in a sample of 1000 trees. Solid lines indicate instances indicate significant relationships (where the highest-posterior density of the slope did not overlap zero), and dashed lines indicate non-significant relationships.



Fig S11. Frequency of mutability losses relative to the total number of changes in mutability during the evolution of anti-HIV B cell lineages. Rows correspond to different mutability metrics, and column contain results obtained for the whole analyzed region of the BCR sequence, and separately for framework regions (FRs) and complementarity determining regions (CDRs). Each point denotes the frequency of changes in mutability that were losses, averaged across a sample of 1000 trees from the posterior distribution. Vertical red lines indicate the 95% highest-posterior density interval.



Fig S12. Total substitution rate inferred from the random local clock model, as a function of time for the observed lineages. Each plot shows the points corresponding to a sample of 100 trees from the posterior distribution inferred by BEAST. Terminal branches are shown in blue, and internal branches are shown in black.



Fig S13. Total substitution rate inferred from robust counting (summing synonymous and non-synonymous rate estimates), as a function of time for the observed lineages. Each plot shows the points corresponding to a sample of 100 trees from the posterior distribution inferred by BEAST. Terminal branches are shown in blue, and internal branches are shown in black.



Fig S14. Robust counting synonymous substitution rate as a function of time for the observed lineages. Each plot shows the points corresponding to a sample of 100 trees from the posterior distribution inferred by BEAST. Terminal branches are shown in blue, and internal branches are shown in black.



Fig S15. Robust counting non-synonymous substitution rate as a function of time for the observed lineages. Each plot shows the points corresponding to a sample of 100 trees from the posterior distribution inferred by BEAST. Terminal branches are shown in blue, and internal branches are shown in black.



Fig S16. Relationship between robust counting substitution rates and time for simulations performed under different levels of decline in the overall mutation rate. Each black line is the linear regression line between branch-specific rates and times for a single tree from the posterior distribution inferred for a simulated alignment using BEAST. Each plot shows a sample of 500 lines. The red lines are the "average" regression lines, with the average intercept and the average slope calculated from a larger sample of 1000 trees from each distribution.



Fig S17. Relationship between robust counting substitution rates and time for simulations performed under models where the mutation rate at each site depends on its S5F mutability or on whether that site is at the center of a WRCH/DGYW hotspot (in which case it mutates either 3 or 30 times more frequently than non-hotspots sites). Each black line is the linear regression line between branch-specific rates and times for a single tree from the posterior distribution inferred for a simulated alignment using BEAST. Each plot shows a sample of 500 lines. The red lines are the "average" regression lines, with the average intercept and the average linear coefficient calculated from a larger sample of 1000 trees from each distribution.