How a well-adapting immune system remembers

Andreas Mayer[a,b], Vijay Balasubramanian[c,d], Aleksandra M. Walczak[b,1,2], and Thierry Mora[b,1,2]

[a]Lewis–Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ 08544 [b]Laboratoire de Physique de l'École Normale Supérieure, PSL University, Centre National de la Recherche Scientifique, Sorbonne University, University Paris-Diderot, 75005 Paris, France; [c]David Rittenhouse Laboratories, University of Pennsylvania, Philadelphia, PA 19104; and [d]Initiative for the Theoretical Sciences, The Graduate Center, The City University of New York, New York, NY 10016

Edited by Herbert Levine, Rice University, Houston, TX, and approved March 21, 2019 (received for review July 25, 2018)

An adaptive agent predicting the future state of an environment must weigh trust in new observations against prior experiences. In this light, we propose a view of the adaptive immune system as a dynamic Bayesian machinery that updates its memory repertoire by balancing evidence from new pathogen encounters against past experience of infection to predict and prepare for future threats. This framework links the observed initial rapid increase of the memory pool early in life followed by a midlife plateau to the ease of learning salient features of sparse environments. We also derive a modulated memory pool update rule in agreement with current vaccine-response experiments. Our results suggest that pathogenic environments are sparse and that memory repertoires significantly decrease infection costs, even with moderate sampling. The predicted optimal update scheme maps onto commonly considered competitive dynamics for antigen receptors.

immune repertoire | Bayesian prediction | biophysics | immune memory | stochastic dynamics

All living systems sense the environment, learn from the past, and adapt predictively to prepare for the future. Their task is challenging because environments change constantly, and it is impossible to sample them completely. Thus, a key question is how much weight should be given to new observations vs. accumulated past experience. Because evidence from the world is generally uncertain, it is convenient to cast this problem in the language of probabilistic inference where past experience is encapsulated in a prior probability distribution which is updated according to sampled evidence. This framework has been successfully used to understand aspects of cellular (1–4) and neural (5–8) sensing. Here, we propose that the dynamics of the adaptive immune repertoires of vertebrates can be similarly understood as a system for probabilistic inference of pathogen statistics.

The adaptive immune system relies on a diverse repertoire of B- and T-cell receptors to protect the host organism from a wide range of pathogens. These receptors are expressed on clones of receptor-carrying cells present in varying copy numbers. A defining feature of the adaptive immune system is its ability to change its clone composition throughout the lifetime of an individual, in particular via the formation of memory repertoires of B and T cells following pathogen encounters (9–14). In detail, after a proliferation event that follows successful recognition of a foreign antigen, some cells of the newly expanded clone acquire a memory phenotype. These cells make up the memory repertoire compartment that is governed by its own homeostasis, separate from the inexperienced naive cells from which they came. Upon reinfection by a similar antigen, memory guarantees a fast immune response. With time, our immune repertoire thus becomes specific to the history of infections and adapted to the environments we live in. However, the commitment of part of the repertoire to maintaining memory must be balanced against the need to also provide broad protection from as-yet-unseen threats. What is more, memory will lose its usefulness over time as pathogens evolve to evade recognition.

How much benefit can immunological memory provide to an organism? How much memory should be kept to minimize harm from infections? How much should each pathogen encounter affect the distribution of receptor clones? To answer these questions, we extend a framework for predicting optimal repertoires given pathogen statistics (15) by explicitly considering the inference of pathogen frequencies as a Bayesian forecasting problem (16). We derive the optimal repertoire dynamics in a temporally varying environment. This approach can complement more mechanistic studies of the dynamics and regulation of immune responses (12, 17–22) by revealing adaptive rationales underlying particular features of the dynamics. In particular, we link the amount of memory production to the variability of the environment and show that there exists an optimal time scale for memory attrition. Additionally, we demonstrate how biologically realistic population dynamics can approximate the optimal inference process and analyze conditions under which memory provides a benefit. Comparing predictions of our theory to experiment, we argue for a view in which the adaptive immune system can be interpreted as a machinery for learning a highly sparse distribution of antigens.

Theory of Optimal Immune Prediction

The pathogenic environment is enormous, and the immune system can only sample it sparsely, as pathogens enter into contact with it at some rate λ. We consider an antigenic space of K different pathogens with time-varying frequencies

Significance

The adaptive immune system is able to protect us from a large variety of pathogens, even ones it has not seen yet. Can predicting the future pathogen distribution help in protection? We find that a combination of probabilistic forecasting and occasional sampling of the current environment reduces infection costs—a scheme easily implemented by the memory repertoire. The proposed theoretical framework offers a modular recipe for updating the memory repertoire, which quantitatively predicts the strength of the immune response in flu-vaccination experiments, unlike other update schemes. It also links the observed early life dynamics of the memory pool to the sparseness properties of the pathogen distribution and competitive receptor dynamics for pathogens.

Author contributions: A.M., V.B., A.M.W., and T.M. designed research; A.M., V.B., A.M.W., and T.M. performed research; A.M., V.B., A.M.W., and T.M. contributed new reagents/analytic tools; A.M., V.B., A.M.W., and T.M. analyzed data; and A.M., V.B., A.M.W., and T.M. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]A.M.W. and T.M. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: awalczak@lpt.ens.fr or tmora@lps.ens.fr.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1812810116/-/DCSupplemental.

Published online April 15, 2019.

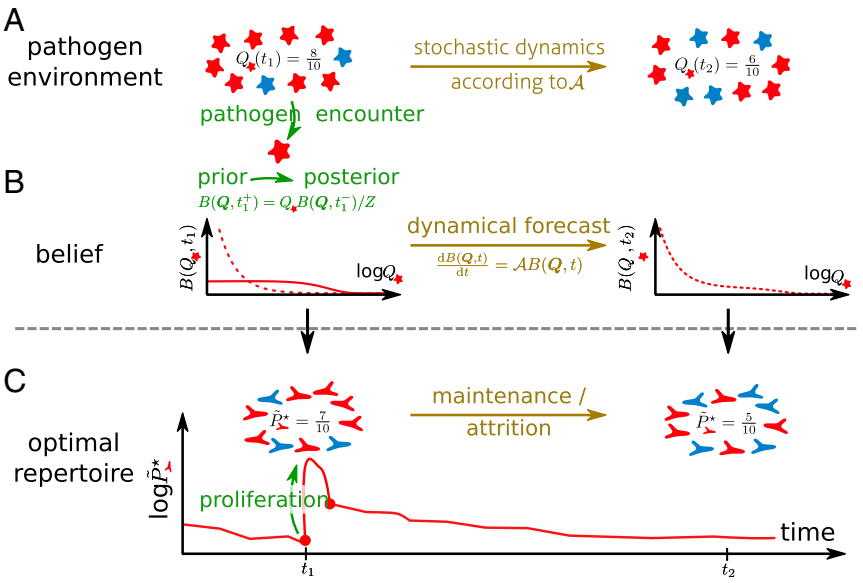

Fig. 1. Sketch of a model of immune repertoire dynamics as a sequential inference process about a time-varying pathogen distribution. (*A*) The organism lives in a pathogenic environment with frequencies of different pathogen strains that change over time. (*B*) Past pathogen encounters provide an avenue for the immune system to learn the pathogen distribution. Using sequential Bayesian inference provides an optimal way to update the beliefs about the frequencies of different pathogens over time. (*C*) Based on its beliefs about the prevalence of pathogens, the optimal immune dynamics allocates lymphocytes across different pathogens to minimize the expected harm from infections. Broadly, the more frequent a pathogen is, the more the organism should be covered. This resource allocation maps the changes in beliefs to the changes in the repertoire composition.

$\boldsymbol{Q}(t) = (Q_1(t), \ldots, Q_K(t))$. These frequencies are unknown to the organism and evolve stochastically. Their dynamics is formally described by a Fokker–Planck operator $\mathcal{A}$ encoding how pathogenic frequencies change (Fig. 1*A* and Materials and Methods). We reason that the immune system should efficiently use the information available through these encounters, along with prior knowledge of how pathogens evolve encoded in the system dynamics, to build an internal representation of the environment (Fig. 1*B*). Biologically, we can think about this representation as being encoded in the composition of the adaptive immune repertoire (the size and specificity of naive and memory lymphocyte clones), but generally further cellular memory mechanisms might also contribute. Based on this representation of the world, the immune system should organize its defenses to minimize harm from future infections (Fig. 1*C*).

How could the immune system leverage a representation of beliefs about pathogen frequencies to provide effective immunity? Each lymphocyte (B or T cell) of the adaptive immune system expresses on its surface a single receptor r out of L possible receptors. This receptor endows the lymphocyte with the ability to specifically recognize pathogens (labeled a) with probability $f_{a,r}$. The immune repertoire is defined by frequencies of these receptors across the lymphocyte population, denoted by $\boldsymbol{P} = (P_1, \ldots P_L)$. These frequencies sum up to one, which implies a resource allocation trade-off between the different receptor types—having more of one in the repertoire implies having less of others. How much harm an infection inflicts depends on how much resources the immune system has devoted to fighting the infection, i.e., the fraction $\tilde{P}_a(t) = \sum_r f_{a,r} P_r$ of the repertoire specific to antigen a, which we will refer to as the coverage of the antigen. Given the pathogen frequencies $\boldsymbol{Q}(t)$ and repertoire distribution $\boldsymbol{P}(t)$, the mean harm cause by the next infection is given by $\sum_a Q_a \cdot c(\tilde{P}_a)$, where c is decreasing function of the fraction of the repertoire specific to the infection (15). The host organism does not know $\boldsymbol{Q}$ with certainty, but has an internal belief $B(\boldsymbol{Q}, t)$ about the frequencies learned through sampling during previous infections. An optimal immune system can then distribute its resources to minimize the expected harm of the next infection:

$$\boldsymbol{P}^\star(t) = \underset{\boldsymbol{P}}{\operatorname{argmin}} \sum_a \hat{Q}_a(t) \cdot c(\tilde{P}_a) \equiv G(\hat{\boldsymbol{Q}}(t)), \qquad [1]$$

where $\hat{\boldsymbol{Q}}(t) \equiv \langle \boldsymbol{Q} \rangle_{B(\boldsymbol{Q},t)}$ are the expected frequencies of pathogens. Note that, because in our framework the state of the immune system does not affect the stochastic dynamics of the pathogenic environment, this strategy is the best possible one, even for minimizing harm against all future infections, provided that the repertoire is allowed to revise its composition instantaneously. Although the function G may be complicated, it generally implies that receptors that are specific to frequent infections (high $\hat{Q}_a$) should be well represented in the optimal repertoire (high $\tilde{P}^\star_a$) (ref. 15 and Materials and Methods). In this framework, we have assumed that infections, their clearing by the immune system, and the subsequent update of the repertoire are all fast compared with changes in the environment, which occur over a time τ, and to the mean time between pathogenic encounters (λ^{-1}). In our optimization, we keep the specificity of the receptors ($f_{a,r}$) fixed, assuming that this specificity is set independently by biophysical constraints and by the need to allow discrimination between antigens, especially between self and foreign ones (12, 23).

The internal representation of the environment can be regarded as a system of beliefs, or guesses, about pathogen frequencies. Formally, these beliefs can be represented in the form of a probability distribution function $B(\boldsymbol{Q}, t)$ over pathogen frequencies, which the host implicitly computes using all of the information it has garnered over time. Optimally, these beliefs are computed by the rules of Bayesian sequential forecasting, by combining the memory of past encounters with knowledge of the stochastic rules under which the pathogenic environment evolves (Materials and Methods). Optimally, the belief distribution should be initialized at birth to reflect the steady-state distribution of the dynamics,

$$B(\boldsymbol{Q},0)=\rho_s(\boldsymbol{Q}), \quad [2]$$

where $\mathcal{A}\rho_s=0$. Upon encountering a pathogen a at time t, the prior belief distribution $B(\boldsymbol{Q},t^-)$ is combined with the likelihood of the observed pathogen Q_a to compute the postencounter belief $B(\boldsymbol{Q},t^+)$ according to Bayes rule (16):

$$B(\boldsymbol{Q},t^+)=\frac{Q_a\,B(\boldsymbol{Q},t^-)}{\int \mathrm{d}\boldsymbol{Q}'\,Q'_a\,B(\boldsymbol{Q}',t^-)}. \quad [3]$$

Between encounters, the immune system should continue to update its beliefs by forecasting how pathogen frequencies change with time. The optimal way to do so is to project the old belief distribution forward in time using (16)

$$\frac{\mathrm{d}B(\boldsymbol{Q},t)}{\mathrm{d}t}=\mathcal{A}B(\boldsymbol{Q},t). \quad [4]$$

This prediction step, which is performed in the absence of any new information, relies on the immune system "knowing," i.e., having learned over long evolutionary time scales, the probability laws governing the stochastic evolution of the environment—but not, of course, the actual path that it takes. In Results, we show how Eqs. **2**–**4** can be turned from abstract belief updates into dynamical equations for a well-adapting immune repertoire.

The Bayesian forecasting framework provides a broad account of the possible adaptive value of many features of the adaptive immune system without the need for additional assumptions. Immune memory formed after a pathogenic challenge is explained as an increase in optimal protection level resulting from an increase in estimated pathogen frequency, following Eq. **3**. Attrition of immune memory is also adaptive, because it allows the immune repertoire to forget about previously seen pathogens which it should do in a dynamically changing environment (Eq. **4**). Lastly, some of the biases in the recombination machinery and initial selection mechanisms (24) represent an evolutionary prior (Eq. **2**) which tilts the naive repertoire toward important regions of antigenic space.

We are proposing an interpretive framework for understanding adaptive immunity as a scheme of sequential inference. This view provides two key insights. First, it confirms the intuition that new experience should be balanced against previous memory and against unknown threats in order for adaptive immunity to work well. Second, it suggests a particular dynamics of implicit belief updates that can globally reorganize the immune repertoire to minimize harm from the pathogenic environment. Going beyond these broad ideas, in Results, we analyze in detail a model for optimal immune prediction in which all these statements can be made mathematically precise. We also show a plausible implementation that the immune system could follow to approximate optimal Bayesian inference, and we compare the resulting dynamics with specific features of the adaptive immune system.

Results

A Lymphocyte Dynamics for Approximating Optimal Sequential Inference. For concreteness, we consider a drift-diffusion model of environmental change (Eq. **11**). The drift-diffusion model, while clearly a much-simplified model of real evolution, captures two key features of changing pathogenic environments: the coexistence of diverse pathogens and the temporal turnover of dominant pathogen strains. The aim of this model of pathogen evolution is not to provide a realistic description of short-term pathogen dynamics within a population such as during an epidemic, but rather to capture overall features of the long-term dynamics of many pathogens over a host's lifetime. The drift-diffusion model is mathematically equivalent to a classical neutral stochastic evolution of pathogens (25) driven by genetic drift happening on a characteristic time scale τ and immigration from an external pool with immigration parameters $\boldsymbol{\theta}=(\theta_1,\ldots,\theta_K)$ (Eq. **11**). Generally, pathogens are under selective pressure to evade host immunity, and strains are replaced faster than under the sole action of genetic drift. Matching the time scale of pathogen change to those observed experimentally, the model then provides a simple, effective description of the pathogen dynamics.

Can a plausible dynamics of lymphocyte receptor clones approximate the optimal repertoire dynamics? In particular, is there an approximate autonomous dynamics for the repertoire composition, which does not require access to the full latent high-dimensional belief distribution $B(\boldsymbol{Q},t)$? Consider the simple case of a logarithmic cost function, $c(\tilde{P}_a)=-\ln\tilde{P}_a$, and uniquely specific receptors, $f_{a,r}=\delta_{a,r}$, in which the optimal match between the receptor and pathogen distributions (Eq. **1**) is the identity, $P^*_a=\hat{Q}_a$ (ref. 15 and Materials and Methods).

For the pathogen dynamics of Eq. **10**, we show, using a decomposition of $B(\boldsymbol{Q},t)$ into a mixture of Dirichlet distributions (*SI Appendix*, section 1B), that there exists an approximate but autonomous dynamics for the number of lymphocyte receptors $\boldsymbol{N}$. It takes the following simple form (parallel to Eqs. **2**–**4**):

$$N_a(0)=\chi\theta_a \quad \text{for all } a, \quad [5]$$

$$N_a(t^+)=N_a(t^-)+\chi \quad \text{when } a \text{ is encountered}, \quad [6]$$

$$\tau\frac{\mathrm{d}N_a}{\mathrm{d}t}=-\frac{1}{2}(\chi^{-1}|\boldsymbol{N}|-1)(N_a-\chi\theta_a) \quad \text{for all } a, \quad [7]$$

where χ is a scale factor controlling the total population size and where $\boldsymbol{P}=\boldsymbol{N}/|\boldsymbol{N}|$. Here and in the following, we use $|\boldsymbol{x}|=\sum_i x_i$ to denote the l_1-norm of a vector $\boldsymbol{x}$. We compared these approximate dynamics to an exact solution computed by spectrally expanding the generator of the stochastic dynamics (*SI Appendix*, section 1C) and found that they approximate the optimal dynamics closely (Fig. 2).

These dynamics have a plausible biological implementation. Each pathogen encounter leads to a fixed increment χ of the number of specific lymphocytes (Eq. **6**), implying a regulation mechanism that controls the number of cell divisions upon clonal expansion as a function of precursor frequency. Through thymic

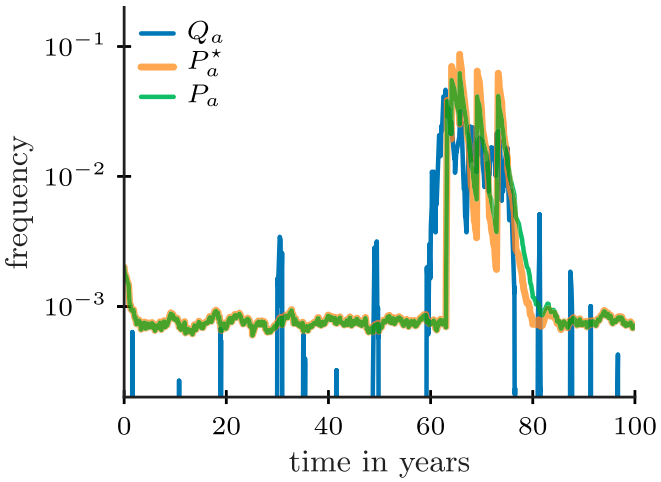

Fig. 2. A plausible repertoire dynamics can implement approximate Bayesian inference. Frequency Q_a of a pathogen over time (blue line) in a dynamically changing environment along with the exact optimal repertoire dynamics (orange line) and the approximate dynamics based on Eqs. **5**–**7** (P_a, green line). The approximate dynamics closely follows the exact Bayesian dynamics implemented as described in *SI Appendix*, section 1C ($P_a^\star$, orange line). Parameters are as follows: $K=500$, $\theta=0.02$, $\tau=20$ year, $\lambda=10$/year, $c(\tilde{P}_a)=-\ln\tilde{P}_a$, and $f_{a,r}=\delta_{a,r}$.

output, the repertoire starts at and is renewed by the naive repertoire—encoded in $\chi\theta_a$ (Eqs. **5** and **7**). By definition, this naive repertoire is optimally tuned to the baseline frequencies of pathogens $\theta_a/|\boldsymbol{\theta}|$. The total number of lymphocytes $|\boldsymbol{N}|$ regulates the time scale of cell decay and renewal (Eq. **7**), which can be done through some global homeostatic mechanisms.

To gain intuition about what the dynamics of Eqs. **5**–**7** mean, let us define the rescaled variables $n_a = \chi^{-1} N_a$. Eq. **6** increases n_a by 1 whenever pathogen a is sampled. $\boldsymbol{n}$ can thus be interpreted as counting the number of times the different pathogens were sampled and their initial value $n_a(0) = \theta_a$ as pseudocounts corresponding to the prior belief in the pathogen frequencies. In a changing environment, experience gained from previous encounters eventually loses its predictive power, and thus Eq. **7** discounts these old encounters, giving more relative weight to recent ones. If there are no new pathogen encounters, n_a eventually relaxes back to the prior θ_a.

The assumption that $P_a^* = \hat{Q}_a$ can be relaxed, provided that the nonlinear relation G between $\boldsymbol{P}$ and $\boldsymbol{Q}$ (Eq. **1**) remains approximately invertible, leading to autonomous dynamics of the lymphocyte clone sizes $N_a(t)$ similar to Eqs. **5**–**7** (*SI Appendix*, section 2C).

Learnability of Pathogen Distribution Implies a Sparse Pathogenic Landscape. The immune system must be prepared to protect us not just from one pathogen but a whole distribution of them. Even restricting recognition to short peptides and accounting for cross-reactivity (26), estimates based on precursor frequencies for common viruses give an effective antigen environment of size $K \sim 10^5$–10^7 (27). How can the immune system learn anything useful about such a high-dimensional distribution from a limited number of pathogenic encounters? Naively, one might expect that the number of samples needed to learn the distribution of pathogens must be larger than of the number of pathogens, i.e., $\lambda t \sim K$, where t is the time over which learning takes place. Although little is known about the receptor–antigen encounter rate λ, this estimate suggests that the pathogenic environment is not easily learnable, and therefore memory has limited utility.

This apparent paradox can be resolved by the fact that the pathogenic environment may be sparse, meaning that only a small fraction of the possible pathogens are present at any given time. In our model of the pathogen dynamics, this sparsity is controlled by the parameter $\boldsymbol{\theta}$. In the scenario that we are considering, typical pathogen landscapes $\boldsymbol{Q}$ are drawn from the steady-state distribution $\rho_s(\boldsymbol{Q})$ of the immigration-drift dynamics, which is a Dirichlet distribution parametrized by $\boldsymbol{\theta}$ (*SI Appendix*, section 1, Eq. **S1**). When θ_a is small, the distribution is peaked at $Q_a = 0$, meaning that pathogen a is absent a majority of the time. For instance, for uniform $\theta_a \equiv \theta \ll 1$, the effective number of pathogens present at any given time is $K\theta$ (*SI Appendix*, section 3C). Since the system only needs to learn about the pathogens that are present, the condition for efficient learning should naively be $\lambda t \sim K\theta$, which is much easier to achieve realistically for small θ.

Our theory can be used to quantify the benefit of memory as a function of the different immunological parameters. We compute the optimized cost function $c(t) = \sum_a Q_a(t) c(\tilde{P}_a^\star(t))$ and study how it decreases as a function of age, t, relative to the cost at birth $c_0 = c(t = 0)$, as the organism learns from pathogen encounters. This relative cost also depends on the encounter rate λ, the size of the pathogenic space K, the sparsity of pathogenic space θ, and time scale of change in the environment τ. Fig. 3*A* shows that the benefit of memory increases with pathogen sparsity—when θ is small, even a few encounters suffice to seed enough memory to reduce the cost of future infections. The cost saturates with age to a value c_∞, either because memory approaches optimality or because memory eventually gets discarded and renewed as the environment changes. Fast-changing environments lead to an earlier and higher saturation of the cost with age (Fig. 3*B*) since learning and prediction are limited by decorrelation of the environment. The pathogen dynamics is sped up when there are strong selection pressures to evade immunity. Faster dynamics decrease learning efficiency and in turn reduce selective pressures. The effective time scale should in practice be set by a coevolutionary balance between both effects.

Analytical arguments show that in the limit of few samples the relative cost c/c_0 achievable in a static environment scales as $\lambda t/K\theta$ (*SI Appendix*, section 3). In general, we find that the cost is a function of $\lambda t_e/K\theta$, where the effective time t_e is defined via $\lambda t_e = |\boldsymbol{n}(t)| - |\boldsymbol{n}(0)| \approx \sqrt{2\lambda\tau}\tanh(\lambda t/\sqrt{2\lambda\tau})$ with $\boldsymbol{n}(t)$ being the vector of the encounter counts discussed above (*SI Appendix*, section 1D for derivation from Eqs. **6** and **7**). Plotted in terms of this variable, the relative cost gap as function of time, $(c - c_\infty)/(c_0 - c_\infty)$, collapses onto a single curve for all parameter choices (Fig. 3*C*). Fig. 3*C* shows that the cost drops by a factor of ~ 2, when $\lambda t_e/K\theta \sim 1$. Thus, there is a substantial benefit to memory already when the effective number of encounters is comparable to the effective number of pathogens. At young ages (small t) or with slowly changing environments (large τ),

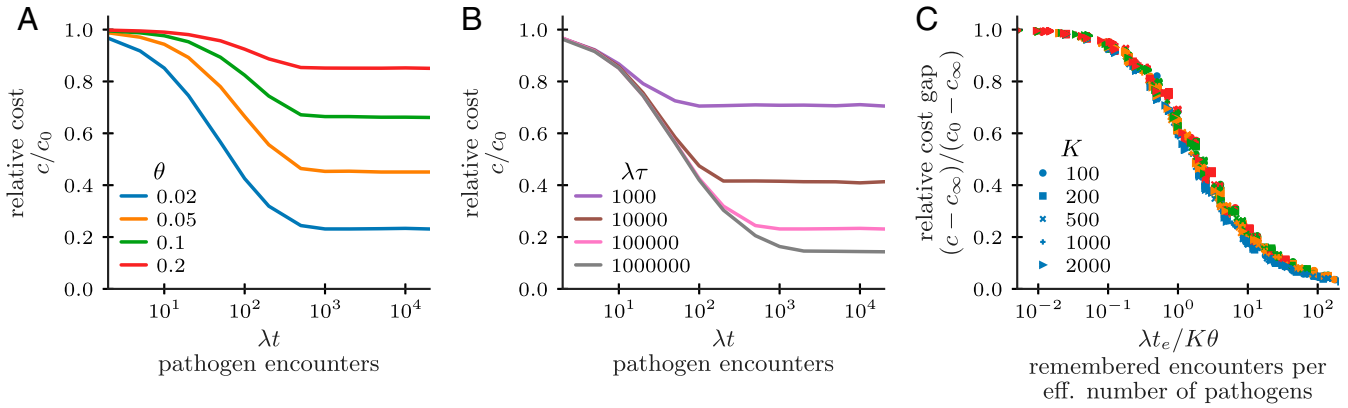

Fig. 3. Advantage of immunological memory depends on sufficient sampling. The mean expected cost of an infection in a changing environment is a function of the age of the organism t, the time scale τ on which the environment changes, and the sparsity $1/\theta$ of the environment. (*A*) Relative cost as a function of age for environments with different sparsity (for fixed $\lambda\tau = 10^5$, K = 2,000). (*B*) Relative cost as a function of age for environments changing more or less rapidly (for fixed $\theta = 0.02$, K = 2,000). (*C*) Collapse of the data by plotting the relative cost gap against the number of samples per effective (eff.) dimension. Simulations with logarithmically spaced parameters in the ranges λt from 1 to 5,000, $\lambda\tau$ from 1 to 1,000,000, K from 100 to 2,000, and θ from 0.01 to 0.2 (color-coded as in *A*). Cost is $c(\tilde{P}_a) = 1/\tilde{P}_a$, and receptors are assumed to be uniquely specific, $f_{a,r} = \delta_{a,r}$.

$t_e \approx t$, and so this condition is simply $\lambda t \sim K\theta$, i.e., the total number of encounters should be comparable to effective number of pathogens that are present.

Optimal Attrition Time Scale. Our theory suggests that there is an optimal time scale for forgetting about old infections which is related to the time scale over which the environment varies. Eq. **7** shows that memory should optimally be discounted on an effective time scale $\tau_{\rm mem} = 2\tau/(|\boldsymbol{n}| - 1)$. Comparing this to the slowest time scale of environmental variation, $\tau_c = 2\tau/|\boldsymbol{\theta}|$ (*SI Appendix*, Eq. **37**), where $|\boldsymbol{\theta}| = \sum_a \theta_a$, we have

$$\tau_{\rm mem} = \tau_c \frac{|\boldsymbol{\theta}|}{|\boldsymbol{n}| - 1}. \qquad [8]$$

The time scale on which old memories should be forgotten scales with the environmental correlation time scale. The two time scales are equivalent when the immune system has little information about the pathogenic environment ($|\boldsymbol{n}| \sim |\boldsymbol{\theta}|$). Given the long time scales over which many relevant pathogens change, immune memory should generally be long-lived (with the time scale of decay being of the order of years or decades). Indeed, despite the relatively short life span of memory cells (28), constant balanced turnover keeps elevated levels of protection for decades after an infection, even in the absence of persistent antigens (29–31).

Interestingly, our theory predicts that memory should be discounted more quickly when the immune system has gathered more information (larger $|\boldsymbol{n}|$). Using the mean-field equations for $|\boldsymbol{n}(t)| - |\boldsymbol{n}(0)|$ from the previous section, we can derive how the memory time scales at steady state at high sampling rate. Using that for large times $|\boldsymbol{n}(t)| \gg |\boldsymbol{n}(0)|$ holds in the high sampling rate limit, one can simplify the mean-field result to $|\boldsymbol{n}| \sim \sqrt{2\lambda\tau}$ in steady state ($t \to \infty$). Combined with Eq. **8**, $\tau_{\rm mem} = \sqrt{\tau_c |\boldsymbol{\theta}|/\lambda}$ follows, which shows that a larger sampling rate leads to a faster discounting of past evidence. This is reminiscent of results in optimal cellular signaling where there are similar trade-offs between noise averaging and responsiveness to changes in the input signal (32).

Memory Production in Sparse Environments Should Be Large and Decrease with Prior Exposure. The theory can be used to make quantitative and testable predictions about the change in the level of protection that should follow a pathogen encounter. Consider an infection cost function that depends as a power law on the coverage, $c(\tilde{P}_a) = 1/\tilde{P}_a^{\alpha}$, with a cost exponent α that sets how much attention the immune system should pay to recognizing rare threats. (Below, we will use the shorthand $\alpha = 0$ to indicate logarithmic cost.) Cost functions of this form can be motivated by considering the time to recognition of an exponentially growing antigen population by the immune system (15) or, alternatively, by considering the time delay of the expansion of the precursor cells to some fixed number of effector cells (*SI Appendix*, section 4).

In the simplest model for repertoire updates, recognition of pathogens leads to proliferation proportionally to the number of specific precursor cells, followed by a homeostatic decrease of the memory pool (18, 33). Thus, the fold change $\tilde{P}_a(t^+)/\tilde{P}_a(t^-) = \text{const}$ where t^-, t^+ are times just before and after the encounter. By contrast, our Bayesian theory predicts that the fold change upon encountering pathogen a should be

$$\frac{\tilde{P}_a^\star(t^+)}{\tilde{P}_a^\star(t^-)} = (1 + \kappa/\tilde{P}_a^\star(t^-)^{(1+\alpha)})^{1/(1+\alpha)}, \qquad [9]$$

where κ depends on prior expectations about the antigenic environment and previous pathogen encounters (*Materials and Methods*). Setting $\alpha = 0$ gives the result for a logarithmic cost function.

To understand this prediction, first consider the effect of a primary infection on a naive repertoire, $\theta_a \equiv \theta$, $\tilde{P}^\star(0) = 1/K$, and $|\boldsymbol{n}(0)| = K\theta$, where the receptors are uniquely specific ($f_{a,r} = \delta_{a,r}$). In this case, $\kappa = 1/K^{1+\alpha}\theta$ (*SI Appendix*, section 2A) and Eq. **9** predicts a fold change of $(1 + 1/\theta)^{1/(1+\alpha)}$. We have argued previously that their learnability implies that pathogenic environments are sparse, i.e., $\theta \ll 1$. Therefore, we predict that primary antigenic encounters should lead to a large memory production. Experimentally, memory production typically leads to the proliferation of antigen-specific cells by a factor of 100- to 1,000-fold (14), in qualitative agreement with this prediction. Turning the argument around, such a large increase in protection upon an encounter is only adaptive in highly sparse environments. Quantitatively, it implies a sparsity parameter $\theta \sim 10^{-6}$–10^{-4} (here, taking $\alpha = 1$ for definiteness) (Fig. 4*B*). Combined with the estimate $K \sim 10^5$–10^7 (27), this suggests that the effective number of pathogens at any given time ranges from $K\theta = 0.1$ to $1,000$.

To test Eq. **9** on immunological data, we fit the Bayesian update model to experiments reporting fold changes in antigen titers upon booster vaccinations against influenza from ref. 34

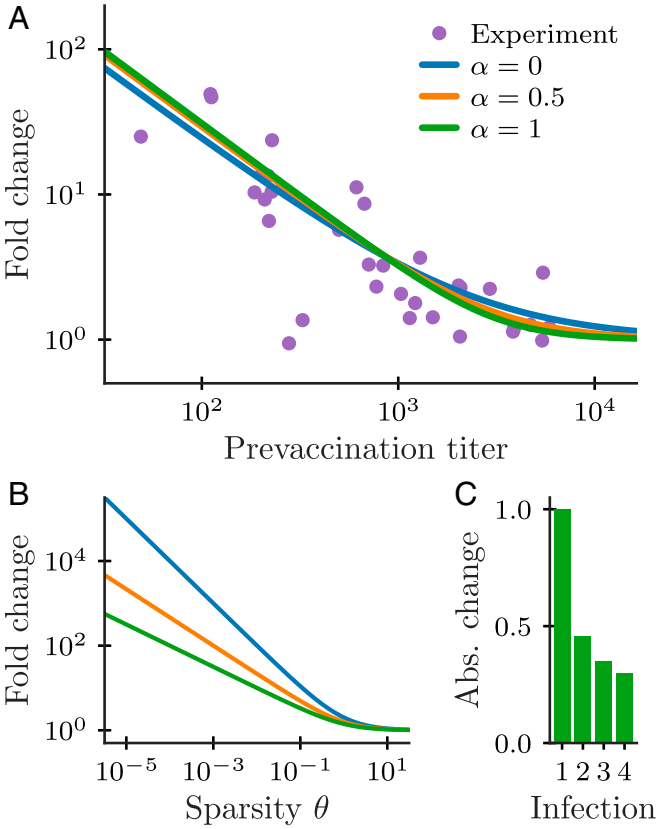

Fig. 4. Changes in protection levels upon infections for cost functions $c(\tilde{P}_a) = 1/\tilde{P}_a^{\alpha}$ ($\alpha = 0$ indicates logarithmic cost). (*A*) The prediction of the Bayesian model (Eq. **9** for different α) closely fits experimental data on antibody titers, while a constant fold-change model (which would correspond to a straight horizontal line) does not. Data on prevaccination and postvaccination antibody titers against stem and head hemagglutinin epitopes following a booster vaccination with inactivated H5N1 in humans from ref. 34. (*B*) Optimal fold change of coverage for $n_i(t^-) = \theta$ for all i. The fold change increases with the sparsity of the environment controlled by parameter θ. (*C*) Absolute (Abs.) change in coverage for primary infection and reinfections (for $\alpha = 1$), normalized such that the change for the primary infection is 1 and neglecting attrition.

(Fig. 4*A*) using least squares. Titers correspond to the concentration of antibodies that are specific to the antigen a and can thus be viewed as an experimental estimate of $\tilde{P}_a$. The optimal Bayesian strategy explains the data, accounting for the larger boosting at small prevaccination titers and showing no increase for large titers, while the proportional model predicts constant boosting for all titers. Similar experimental results have been reported for antibody titers before and after a shingles vaccination (35). Mechanistic models have been proposed to explain how the population dynamics of expanding lymphocytes might give rise to nonproportional boosting for both B and T cells (22, 33, 36, 37).

Interestingly, for T cells, Quiel et al. (38) have shown that fold expansion to peak cell numbers in an adoptive transfer experiment depends on the initial number of T cells as a power law with exponent $\sim -1/2$. That scaling, which is for the peak expansion, predicts more expansion at high precursor number than Eq. **9**, which is for memory production. This implies a nonlinear relationship between peak T-cell level and memory production, which further suppresses memory production at high precursor numbers. This prediction might be checked in experiments measuring memory production after infection clearance, as well as the expansion peak.

Long-Term Dynamics of a Well-Adapting Repertoire. Our model makes predictions for the dynamics of growth and attrition of memory over time, with consequences for immunity and for the diversity of the immune repertoire. We quantify the dynamics in terms of a memory fraction defined as a sum of the coverage fractions $\tilde{P}_{a_i}$ over all previously encountered pathogens $\{a_i\}$. The memory fraction measures the size of memory relative to the size of the whole immune repertoire. Early in life, every infection is new, and even modest increases in the memory fraction lead to large drops in infection susceptibility (measured by the expected cost of new infections in Fig. 5*A*). At the same time, the memory fraction increases rapidly (Fig. 5*B*), but the growth of memory slows as subsequent infections lead to less memory production following the optimal fold-change rule in Eq. **9** and as attrition begins to play a role. The fraction of the repertoire devoted to memory in midlife is largely determined by how the cost of infections scales with coverage. The observed memory fraction of $\sim 50\%$ at midlife suggests a cost exponent of $\alpha \approx 0.5$ (Fig. 5*B*). The diversity of the memory repertoire increases with time at a rate that slows with age (quantified in Fig. 5*C* by richness, which measures the number of unique specificities and the Shannon entropy of the repertoire frequency distribution).

To gain insight into these dynamics of our model, we average the stochastic equations over the statistics of pathogen encounters. We show in *SI Appendix*, section 2B that this mean-field approximation yields a differential equation for the population fraction of different clones with two opposing contributions which balance alignment of the immune repertoire with the current pathogenic environment (i.e., memory production) against alignment with the long-term mean environment (i.e., attrition). Interestingly, the mean-field equation broadly coincides with dynamics that were proposed in ref. 15 to self-organize an optimal immune repertoire. The essential difference here is that the time scale of learning slows down with increasing experience following the rules of optimal sequential update in Eq. **9**.

We then asked which features of the proposed repertoire dynamics are most relevant to ensure its effectiveness. How important is the negative correlation between fold expansion and prior immune levels, and how important is attrition? Furthermore, if the immune system follows Bayesian dynamics, it must have integrated on an evolutionary time scale a prior about composition and evolution of the pathogen environment through the parameters $\boldsymbol{\theta}$ and τ—however, the prior may be inaccurate. How robust is the benefit of memory to imperfections of the host's prior assumptions about pathogen evolution? To answer these questions, we compare the long-term immune-repertoire dynamics using the optimal Bayesian scheme to other

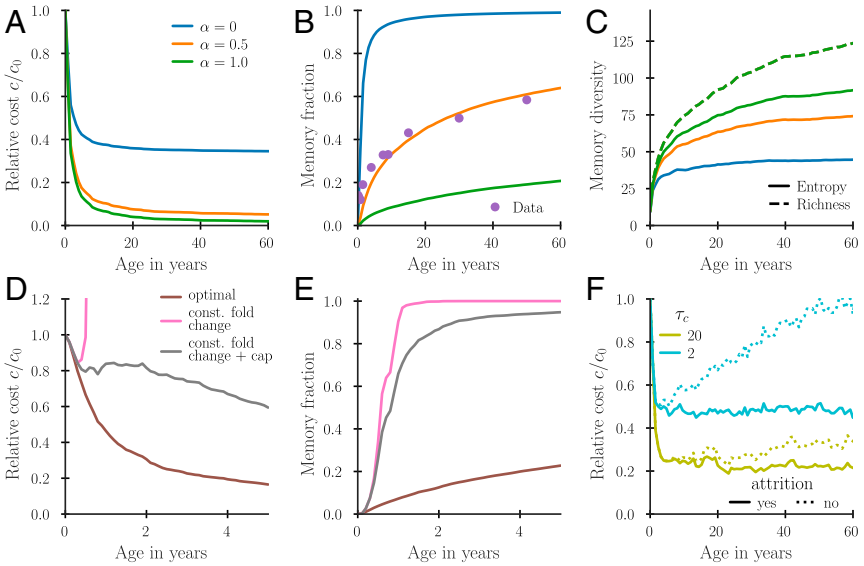

Fig. 5. Relative cost (*A* and *D*), memory-cell fraction (*B* and *E*), and memory diversity (*C*) as a function of age. *A*–*C* show long-term dynamics for a repertoire following optimal Bayesian update dynamics for three different α. Memory diversity is plotted as richness, i.e., the number of unique memory specificities, as well as the exponential of Shannon entropy *S* of the memory compartment defined for a probability distribution p_i with $\sum_i p_i = 1$ as $S = -\sum p_i \ln p_i$. *D* and *E* compare the optimal Bayesian dynamics with a constant fold change update by a factor of 30 and the same proportional update but with a cap to 10^5 of the total fold expansion for any clone. (*F*) Comparison of the cost as a function of age for the optimal dynamics (solid lines) and a dynamics without attrition (dotted lines) for two environmental correlation time scales $\tau_c = 2\tau/K\theta$ (*SI Appendix*, section 1C). Parameters are as follows: encounter rate $\lambda = 40$/year, antigen space dimensionality $K = 10^5$, and antigen sparsity $\theta = 2.5 \cdot 10^{-4}$. In *A*–*E*, we used $\tau \to \infty$, and in *D*–*F*, $\alpha = 0.5$. To reduce fluctuations, the statistics were averaged over multiple runs of the dynamics. Data in *B* are from refs. 39 and 40.

simplified schemes. We find that a constant fold expansion dynamics quickly leads to very suboptimal repertoire compositions (Fig. 5*D*, pink line), since the exponential amplification of cells specific to recurrent threats quickly leads to a very large fraction of the repertoire consisting of memory of those pathogens (Fig. 5*E*, pink line). This suboptimality persists, even if we assume that some global regulation caps the constant fold expansion such that no individual receptor clone can take over all of the repertoire (Fig. 5 *D* and *E*, gray line). Thus, negative feedback in T-cell expansion to individual antigens is very important to maintain a properly balanced diverse repertoire. In contrast, within dynamics with a negative correlation, the precise levels of updating do not need to be finely tuned to the environmental statistics: Varying the assumed sparsity of the pathogen distribution, which controls fold expansion upon primary infection in the optimal dynamics, leads to a relatively modest deterioration of the convergence speed of the learning process (*SI Appendix*, Fig. S2*A*) and does not matter asymptotically (*SI Appendix*, Fig. S2*B*). Attrition does not matter at young age, but can play an important role for long-term adaptation to relatively rapidly changing pathogen distributions (Fig. 5*F*). However, the attrition time scale need not be finely tuned to get close to optimal dynamics (*SI Appendix*, Fig. S3).

Adapting a Cross-Reactive Repertoire. Above, we described adaptation of immune repertoires in terms of changes in the effective coverage $\tilde{P}_a = \sum_r f_{a,r} P_r$, where the cross-reactivity matrix $F = (f_{a,r})$ reflects the ability of each receptor to recognize many antigens and also the propensity of each antigen to bind to many receptors (26). Because of cross-reactivity, each pathogen encounter should result in the expansion of not just one but potentially many receptor clones. Here, we ask how the optimal immune response is distributed among clones with different affinities.

Following Perelson and Oster (41), we will represent the interaction of receptors and antigens by embedding both in a multidimensional metric recognition "shape space," where receptors are points surrounded by recognition balls. Antigens that fall within a ball's radius will be recognized by the corresponding receptor. In this presentation, a and r are the coordinates of antigens and receptors, respectively, and their recognition propensity depends on their distance, $f_{a,r} = f(|a - r|)$.

Earlier sections have already discussed the optimal dynamics of the coverage $\tilde{P}_a$, which is a convolution of the cross-reactivity matrix with the receptor clone distribution P_r. Thus, the optimal dynamics of the clone distribution can be derived by deconvolving the cross-reactivity subject to the constraint that P_r cannot be negative. Carrying out this analysis in *SI Appendix*, section 2C reveals a general qualitative phenomenon—competitive exclusion between clones expressed in the repertoire and their close neighbors within the cross-reactivity radius (*SI Appendix*, Fig. S4, blue line). This exclusion is not an assumption of the model, but rather stems from the optimal Bayesian theory. Given a receptor clone that covers one region of antigenic shape space, the global likelihood of detecting infections increases by placing other clones to cover other regions. This can be shown analytically when cross-reactivity is limited, memory updates are small in magnitude, and the pathogen distribution is assumed to be uncorrelated (*SI Appendix*, section 2C).

In general, the frequencies of pathogens might be correlated in antigenic space, for example, because mutations from a dominant strain give rise to new neighboring strains. An optimally adapting immune system should incorporate such correlations as a prior probability favoring smoothness of the pathogen distribution. Such priors work their way through the optimal belief update scheme that we have described and weaken the competitive exclusion between clones with overlapping cross-reactivity (*SI Appendix*, Fig. S4, orange line).

In general, when cross-reactivity is wide or the required clone fraction update is large, numerical analysis shows that achieving optimally predictive immunity after a pathogen encounter requires a global reorganization of the entire repertoire (*SI Appendix*, Fig. S5, blue line). There is no plausible mechanism for such a large-scale reorganization since it would involve up- and down-regulation, even of unspecific clones. However, in *SI Appendix*, section 2C we show that the optimal update can be well-approximated by changes just to the populations of specific clones with pathogen-binding propensities $f_{a,r}$ that exceed a threshold. The optimizing dynamics with this constraint exhibit strong competitive exclusion, where only the highest-affinity clones proliferate, while nearby clones with lower affinity are depleted from the repertoire (*SI Appendix*, Fig. S5, orange line). The local update rule provides protection that comes within 1% of the cost achievable by the best global update. Thus, reorganization of pathogen-specific receptor clone populations following an infection, as seen in vertebrates, can suffice to achieve near-optimal predictive adaptation of the immune repertoire.

Discussion

The adaptive immune system has long been viewed as a system for learning the pathogenic environment (10). We developed a mathematical framework in which this notion can be made precise. In particular, we derived a procedure for inferring the frequencies of pathogens undergoing an immigration-drift dynamics and showed how such inference might approximately be performed by a plausible population dynamics of lymphocyte clones. Additionally, we analyzed how quickly the immune system can learn about its environment and find that the antigenic environment must be effectively sparse to be learnable with a realistic rate of pathogen encounters.

The optimal repertoire dynamics recapitulate a number of properties of real adaptive immune systems. Two repertoire compartments emerge from the theory: memory, which encodes lived experience, and naive, which encodes prior expectations. Memory is effective in reducing harm from infection despite the high dimensionality of pathogenic space; having encountered circulating pathogens only once on average can reduce the overall cost of infections by half. The first encounter of a naive individual with a pathogen leads to a large response, which increases protection levels by several orders of magnitude. Memory production depends on prior protection levels and eventually is predicted to saturate, as seen in vaccination experiments (34). This saturation is predicted to lead to a sublinear increase with age of the fraction of the total repertoire taken up by memory, consistent with observations in human cohorts (39, 40).

Our work makes a number of concrete predictions amenable to further testing. We make quantitative predictions about how much memory should be formed following an infection, as a function of number of pathogens effectively present in the environment and the number of previous infections. These dependencies might be tested by using a comparative approach, which relates the amount of formed memory in different species to how many pathogens to which they are susceptible. Our model also makes predictions about changes in the size and diversity of the memory compartment with aging, which might be tested in future repertoire-sequencing studies similar to those done by Britanova et al. (42), by sorting cells into naive and memory types before sequencing. Ultimately, we hope that our work might help motivate studies with longitudinal tracking of the long-term repertoire dynamics in model organisms living in controlled pathogenic environments.

Our framework can be extended to incorporate additional constraints on immune-system function or further aspects of pathogen evolution. An example of a constraint would be to

introduce a maximal rate of change in the optimization to model maximal cellular division and death rates. A more complex pathogen dynamics might model explicitly their mutational dynamics in antigenic space. Such dynamics will lead to correlations in the pathogen distribution, which we showed will influence the structure of the optimal conjugate repertoire. In particular, the optimal response should spread around the currently dominant antigens to also provide protection against potential future mutations. Hypermutations in B cells may play a role in this diversification, in addition to their known function of generating receptors with increased affinity for antigens of current interest. It would also be interesting to extend our framework to other immune-defense mechanisms, including innate immunity, where the role of memory has received recent attention (43).

Although our study was motivated by the adaptive immune system, some of our main results extend to other statistical inference problems. We have extended earlier results on exactly computable solutions to the stochastic filtering problem for Wright–Fisher diffusion processes (44) to derive an efficient approximate inference procedure. This procedure might be of use in other contexts where changing distributions must be inferred from samples at different time points, e.g., in population genetics. Additionally, we have derived the convergence rate for Bayesian inference of categorical distributions in high dimensions in the undersampled regime, showing that effectively sparse distributions can be inferred much more quickly. These results add to the growing literature on high-dimensional inference from few samples (45, 46), which has arisen in the context of the big-data revolution.

We propose that the adaptive immune system balances integration of new evidence against prior knowledge, while discounting previous observations to account for environmental change. Similar frameworks have been developed for other biological systems. In neuroscience, leaky integration of cues has been proposed as an adaptive mechanism to discount old observations in change-point detection tasks (47, 48), and close-to-optimal accumulation and discounting of evidence has been reported in a behavioral study of rat decision-making in dynamic environments (49). Inference from temporally sparse sampling has been considered in the framework of infotaxis, which is relevant for olfactory navigation (50). In the context of immunity, related ideas about inference and prediction of pathogen dynamics have been used to predict flu-strain and cancer-neoantigen evolution in silico (51, 52). Finally, ideas similar to those developed here could be used in ecology or microbiome studies to reconstruct evolutionary or ecological trajectories of population dynamics from incomplete sampling of data at a finite number of time points, e.g., from animal sightings or metagenomics.

Materials and Methods

Modeling Pathogen Dynamics by an Immigration-Drift Process. In our model, we describe the stochastic dynamics of the pathogenic environment (Fig. 1*A*) by a Fokker–Planck equation for the conditional probability distribution $\rho(\boldsymbol{Q}, t)$

$$\frac{\partial \rho(\boldsymbol{Q}, t)}{\partial t} = \mathcal{A}\rho(\boldsymbol{Q}, t), \qquad [10]$$

where $\mathcal{A}$ is a differential operator acting on ρ that controls the dynamics. For concreteness, we consider a population that changes due to genetic drift and immigration from an external reservoir, which we describe by a Wright–Fisher diffusion equation (53, 54)

$$\tau \frac{\partial \rho(\boldsymbol{Q}, t)}{\partial t} = -\frac{1}{2}\sum_a \frac{\partial}{\partial Q_a}[(\theta_a - |\boldsymbol{\theta}|Q_a)\rho(\boldsymbol{Q}, t)] + \frac{1}{2}\sum_{a,b}\frac{\partial^2}{\partial Q_a \partial Q_b}[Q_a(\delta_{a,b} - Q_b)\rho(\boldsymbol{Q}, t)], \qquad [11]$$

where τ sets the time scale of dynamics, θ is a K-dimensional vector of immigration rates, and $\delta_{a,b}$ is the Kronecker delta, which is 1 if $a = b$ and 0 otherwise. To efficiently simulate trajectories according to this dynamics, we sample the new distribution of frequencies directly from the transition density of the stochastic process as described in *SI Appendix*, section 1E. This dynamics retains key features of real pathogen environments. First, at a given point in time, the environment contains many different pathogens with different frequencies determined by genetic drift and immigration. Second, the dominant pathogens change over time, such as is the case for many viruses, e.g., the flu or HIV.

Minimizing the Cost of Infection. To solve the optimization problem Eq. **1** analytically, a set of necessary conditions for optimality, the so-called Karush–Kuhn–Tucker conditions, can be derived. When all receptors are present at a nonzero frequency in the optimal repertoire $P_r^\star > 0$, these conditions imply (15)

$$\left.\frac{\partial \sum_a Q_a c(\tilde{P}_a)}{\partial P_r}\right|_{\boldsymbol{P}^\star} = -\lambda^\star, \qquad [12]$$

where $\lambda^\star$ is set by the condition $\sum_r P_r^\star = 1$. If we further simplify the problem by assuming that there is no cross-reactivity between different pathogens and by considering power-law cost functions, then this simplifies to the explicit solution

$$\tilde{P}_a^\star = \frac{1}{Z} Q_a^{1/(1+\alpha)}, \qquad [13]$$

where Z is a normalization constant. Other cases are discussed in detail in ref. 15, including how to solve the optimization problem numerically by using a projected gradient algorithm in the general case.

Change in Protection upon a Pathogen Encounter. The inference dynamics induces via the mapping from $\hat{\boldsymbol{Q}}$ to $\boldsymbol{P}^\star$ (Eq. **1**) a dynamics of an optimally adapting immune repertoire. To get intuition we derive how the coverage changes in a simple setting in which Eq. **13** holds (further cases are considered in *SI Appendix*, section 2).

Rewriting Eq. **6** in terms of the expected pathogen frequencies $\hat{\boldsymbol{Q}} = \boldsymbol{n}/|\boldsymbol{n}|$, we obtain an update equation for the expected frequencies upon encounter of antigen a as

$$\hat{\boldsymbol{Q}}^+ = \frac{\boldsymbol{n}^- + \boldsymbol{e}_a}{|\boldsymbol{n}^+|} = \frac{|\boldsymbol{n}^-|\hat{\boldsymbol{Q}}^- + \boldsymbol{e}_a}{|\boldsymbol{n}^+|}, \qquad [14]$$

where to simplify notations, we use $\hat{\boldsymbol{Q}}(t^+) = \hat{\boldsymbol{Q}}^+$, and where $|\boldsymbol{n}^+| = |\boldsymbol{n}^-| + 1$. By using Eq. **13**, it follows that coverages are updated as

$$\tilde{P}_b^+ = \frac{[((\tilde{P}_b^-)^{1+\alpha}|\boldsymbol{n}^-| + \delta_{a,b}/(Z^-)^{1+\alpha})/|\boldsymbol{n}^+|]^{1/(1+\alpha)}}{Z^+/Z^-} \qquad [15]$$

Defining $\kappa = 1/(|\boldsymbol{n}^+|(Z^-)^{1+\alpha})$ and neglecting the change in normalization which is of order $1/K$ relative to the update size, we obtain Eq. **9**. To fit the dataset, we note that a proportional rescaling of $\tilde{P}_a$ by a factor k can be subsumed within the model by redefining $\kappa \to \kappa k^{1+\alpha}$. Therefore, the scaling of $\tilde{P}_a$ to an antibody titer can be subsumed within κ.

ACKNOWLEDGMENTS. The work was supported by European Research Council Starting Grant 306312, a Lewis–Sigler fellowship (to A.M.), Simons Foundation Mathematical Modeling of Living Systems Grant 400425, and NSF Grant PHY-1734030. Work on this project at the Aspen Center for Physics was supported by NSF Grant PHY-1607611.

1. Perkins TJ, Swain PS (2009) Strategies for cellular decision-making. *Mol Syst Biol* 5:326.
2. Kobayashi TJ (2010) Implementation of dynamic Bayesian decision making by intracellular kinetics. *Phys Rev Lett* 104:228104.
3. Siggia ED, Vergassola M (2013) Decisions on the fly in cellular sensory systems. *Proc Natl Acad Sci USA* 110:E3704–E3712.
4. Sivak DA, Thomson M (2014) Environmental statistics and optimal regulation. *PLoS Comput Biol* 10:e1003826.
5. DeWeese MR, Zador AM (1998) Asymmetric dynamics in optimal variance adaptation. *Neural Comput* 10:1179–1202.
6. Beck JM, et al. (2008) Probabilistic population codes for Bayesian decision making. *Neuron* 60:1142–1152.
7. Deneve S (2008) Bayesian spiking neurons I: Inference. *Neural Comput* 20:91–117.
8. Wark B, Fairhall A, Rieke F (2009) Timescales of inference in visual adaptation. *Neuron* 61:750–761.
9. Murphy K, Travers P, Walport M (2001) *Janeway's Immunobiology* (Garland Science, Abingdon, UK), 7th Ed, Vol 2.
10. Farmer JD, Packard NH, Perelson AS (1986) The immune system, adaptation, and machine learning. *Phys D* 22:187–204.

11. Ahmed R, Gray D (1996) Immunological memory and protective immunity: Understanding their relation. *Science* 272:54–60.
12. Perelson AS, Weisbuch G (1997) Immunology for physicists. *Rev Mod Phys* 69:1219–1268.
13. Farber DL, Yudanin NA, Restifo NP (2014) Human memory T cells: Generation, compartmentalization and homeostasis. *Nat Rev Immunol* 14:24–35.
14. Buchholz VR, Schumacher TNM, Busch DH (2016) T cell fate at the single-cell level. *Annu Rev Immunol* 34:65–92.
15. Mayer A, Balasubramanian V, Mora T, Walczak AMA (2015) How a well-adapted immune system is organized. *Proc Natl Acad Sci USA* 112:5950–5955.
16. Chen ZHE (2003) Bayesian filtering: From Kalman filters to particle filters, and beyond. *Statistics* 182:1–69.
17. Deem MW, Lee HY (2003) Sequence space localization in the immune system response to vaccination and disease. *Phys Rev Lett* 91:068101, and erratum (2003) 91:229902.
18. Antia R, Ganusov VV, Ahmed R (2005) The role of models in understanding CD8+ T-cell memory. *Nat Rev Immunol* 5:101–111.
19. François P, Altan-Bonnet G (2016) The case for absolute ligand discrimination: Modeling information processing and decision by immune T cells. *J Stat Phys* 162:1130–1152.
20. George JT, Kessler DA, Levine H (2017) Effects of thymic selection on T cell recognition of foreign and tumor antigenic peptides. *Proc Natl Acad Sci USA* 114:E7875–E7881.
21. Chakraborty AK (2017) A perspective on the role of computational models in immunology. *Annu Rev Immunol* 35:403–439.
22. Mayer A, Zhang Y, Perelson AS, Wingreen NS (2019) Regulation of T cell expansion by antigen presentation dynamics. *Proc Natl Acad Sci USA* 116:5914–5919.
23. Percus JK, Percus OE, Perelson AS (1993) Predicting the size of the T-cell receptor and antibody combining region from consideration of efficient self-nonself discrimination. *Immunology* 90:1691–1695.
24. Mora T, Walcak A (2019) Quantifying lymphocyte receptor diversity. *Systems Immunology: An Introduction to Modeling Methods for Scientists* (Taylor & Francis, Boca Raton, FL).
25. Kimura M (1964) Diffusion models in population genetics. *J Appl Probab* 1:177–232.
26. Mason D (1998) A very high level of crossreactivity is an essential feature of the T-cell receptor. *Immunol Today* 19:395–404.
27. Zarnitsyna VI, Evavold BD, Schoettle LN, Blattman JN, Antia R (2013) Estimating the diversity, completeness, and cross-reactivity of the T cell repertoire. *Front Immunol* 4:485.
28. Antia R, Pilyugin SS, Ahmed R (1998) Models of immune memory: On the role of cross-reactive stimulation, competition, and homeostasis in maintaining immune memory. *Proc Natl Acad Sci USA* 95:14926–14931.
29. Amanna IJ, Carlson NE, Slifka MK (2007) Duration of humoral immunity to common viral and vaccine antigens. *N Engl J Med* 357:1903–1915.
30. Sallusto F, Lanzavecchia A, Araki K, Ahmed R (2010) From vaccines to memory and back. *Immunity* 33:451–463.
31. Macallan DC, Borghans JA, Asquith B (2017) Human T cell memory: A dynamic view. *Vaccines* 5:E5.
32. Becker NB, Mugler A, Ten Wolde PR (2015) Optimal prediction by cellular signaling networks. *Phys Rev Lett* 115:258103.
33. De Boer RJ, Perelson AS (2013) Antigen-stimulated CD4 T cell expansion can be limited by their grazing of peptide-MHC complexes. *J Immunol* 190:5454–5458.
34. Ellebedy A, et al. (2014) Induction of broadly cross-reactive antibody responses to the influenza HA stem region following H5N1 vaccination in humans. *Proc Natl Acad Sci USA* 111:13133–13138.
35. Li S, et al. (2017) Metabolic phenotypes of response to vaccination in humans. *Cell* 169:862–877.e17.
36. Bocharov G, et al. (2011) Feedback regulation of proliferation vs. differentiation rates explains the dependence of CD4 T-cell expansion on precursor number. *Proc Natl Acad Sci USA* 108:3318–3323.
37. Zarnitsyna VI, Lavine J, Ellebedy A, Ahmed R, Antia R (2016) Multi-epitope models explain how pre-existing antibodies affect the generation of broadly protective responses to influenza. *PLoS Pathog* 12:e1005692.
38. Quiel J, et al. (2011) Antigen-stimulated CD4 T-cell expansion is inversely and log-linearly related to precursor number. *Proc Natl Acad Sci USA* 108:3312–3317.
39. Saule P, et al. (2006) Accumulation of memory T cells from childhood to old age: Central and effector memory cells in CD4+ versus effector memory and terminally differentiated memory cells in CD8+ compartment. *Mech Ageing Dev* 127:274–281.
40. Shearer WT, et al. (2003) Lymphocyte subsets in healthy children from birth through 18 years of age: The Pediatric AIDS Clinical Trials Group P1009 study. *J Allergy Clin Immunol* 112:973–980.
41. Perelson AS, Oster GF (1979) Theoretical studies of clonal selection: Minimal antibody repertoire size and reliability of self-non-self discrimination. *J Theor Biol* 81:645–670.
42. Britanova OV, et al. (2016) Dynamics of individual T cell repertoires: From cord blood to centenarians. *J Immunol* 196:5005–5013.
43. Netea MG, et al. (2016) Trained immunity: A program of innate immune memory in health and disease. *Science* 352:aaf1098.
44. Chaleyat-Maurel M, Genon-Catalot V (2007, July 4) Filtering the Wright-Fisher diffusion. arXiv:0707.0537.
45. Johnstone IM, Titterington DM (2009) Statistical challenges of high-dimensional data. *Philos Trans A Math Phys Eng Sci* 367:4237–4253.
46. Advani M, Ganguli S (2016) Statistical mechanics of optimal convex inference in high dimensions. *Phys Rev X* 6:031034.
47. Glaze CM, Kable JW, Gold JI (2015) Normative evidence accumulation in unpredictable environments. *eLIFE* 4:e08825.
48. Glaze CM, Filipowicz ALS, Kable JW, Balasubramanian V, Gold JI (2018) A bias-variance trade-off governs individual differences in on-line learning in an unpredictable environment. *Nat Hum Behav* 2:213.
49. Piet AT, El Hady A, Brody CD (2018) Rats adopt the optimal timescale for evidence integration in a dynamic environment. *Nat Commun* 9:4265.
50. Vergassola M, Villermaux E, Shraiman BI (2007) 'Infotaxis' as a strategy for searching without gradients. *Nature* 445:406–409.
51. Łuksza M, Lässig M (2014) A predictive fitness model for influenza. *Nature* 507:57–61.
52. Luksza M, et al. (2017) A neoantigen fitness model predicts tumour response to checkpoint blockade immunotherapy. *Nature* 551:517–520.
53. Etheridge A (2012) *Some Mathematical Models from Population Genetics*, Lecture Notes in Mathematics (Springer, Berlin).
54. Griffiths RC, Spano D (2010, March 24) Diffusion processes and coalescent trees. arXiv:1003.4650.