Invasion Rate Versus Diversity in Population Dynamics with Catastrophes

A. Melka[1], N. Dori,[2] and Y. Louzoun[1,2,*]

[1]*Department of Mathematics, Bar-Ilan University, Ramat Gan 52900, Israel*
[2]*Gonda Brain Research Center, Bar-Ilan University, Ramat Gan 52900, Israel*

A key question in the current diversity crisis is how diversity has been maintained throughout evolution and how to preserve it. Modern coexistence theories suggest that a high invasion rate of rare new types is directly related to diversity. We show that adding almost any mechanism of catastrophes to a stochastic birth, death, and mutation process with limited carrying capacity induces a novel phase transition characterized by a positive invasion rate but a low diversity. In this phase, new types emerge and grow rapidly, but the resulting growth of very large types decreases diversity. This model also resolves two major drawbacks of neutral evolution models: their failure to explain balancing selection without resorting to fitness differences and the unrealistic time required for the creation of the observed large types. We test this model on a classical case of genetic polymorphism: the HLA locus.

The coexistence of numerous types, their diversity and their growth time have been longstanding subjects of interest and modeling [1,2]. Complex models have been developed to explain coexistence, such as the Allee effect [3,4], the rescue effect [5], depensation [6], or density dependent growth [7,8]. Ecologists such as Chesson [9,10], and more recently Ellner *et al.* [11] have merged many of those into a consistent model named the modern coexistence theory (MCT). They define the invasion rate as the probability for new types to grow in the presence of other competitive types. They suggest that mechanisms increasing the invasion rate contribute to coexistence and diversity. Diversity is commonly defined as the number of types over the total population [12–14].

While it would seem intuitive that diversity is indeed maintained through the emergence of new types, the mechanisms driving this emergence can in parallel promote the growth of very large types. Assuming that at equilibrium, the total population is fairly constant, and assuming a heavy-tailed distribution of type sizes, as is often observed and predicted [15,16], those large types may occupy a macroscopic fraction of the total population and lower the diversity. We here use an extension of a classical birth, death, innovation (or mutation) neutral model (BDIM) including catastrophes (BDICM) [17]. We demonstrate that the addition of catastrophes can simultaneously increase the invasion rate and decrease diversity, breaking the paradigm. A catastrophe represents a major deletion event where a type is fully or partly eliminated. In a constant resources environment, the total population size is balanced by equal average birth and death rates. However, when catastrophes are introduced, this balance is broken, and the average death rate of each type not hit by a catastrophe is less than its average birth rate. This results in

net growth for each individual and exponential growth for all types not yet hit. Part of those births are mutations yielding new types of size 1. Their positive net growth is equivalent to a positive invasion rate. However, exponential growth also leads to very large types. The presence of those large types forces a smaller total number of types. We, therefore, observe a lower diversity despite a positive invasion rate.

We also show that BDICM resolves critical issues raised in classical neutral models (e.g., the Moran model [18,19]). They indeed fail to explain the coexistence of several alleles at frequencies higher than prediction from genetic drift. The genetic drift is the change in the genetic composition of a population through random sampling of the current members dying or duplicating. This coexistence is commonly referred to as balancing selection [20]. It is usually explained by frequency dependent selection (FDS), where the fitness of a genotype decreases as it becomes more common [21,22]. While complex models have been proposed for balancing selection or negative FDS using fitness differences [21,23,24], there is limited evidence for any of those. Classical models also fail to explain the emergence of large types and usually require an unrealistically long time for their growth. This problem, raised by Karev *et al.* [25,26], was partly solved in their nonlinear model, but it assumes that individual net growth rates are affected by their type size, which is not practical for genes or any other nonspatial characteristics. An interesting aspect of BDICM is that, in the presence of catastrophes, life expectancies become much lower and uniform across almost all type sizes since all types can get hit with equal probability. We show that this aspect addresses the issue of long expected growth time. It also replicates features of balancing selection in a purely neutral scenario since it removes the advantage of large types.

The model.—Formally, we expanded upon a neutral birth, death and, innovation process, by introducing large-scale events (catastrophes). We assume asexual reproduction. The number of individuals of each type is denoted by k and the number of types of this size is denoted by N_k. The moments of the distribution are $m_j = \sum_k k^j N_k$, $j = 0, 1, 2$, where j is the moment's order. Using this definition, m_0 is the number of types and m_1 is the number of individuals in the population.

Birth events increase by 1 the number of individuals of a type and occur at a rate of α per individual. A constant fraction μ of all births leads through mutations to new types of size 1. Death events decrease by 1 the number of individuals of a type and occur at a rate of $(m_1/\bar{N})$ per individual. $\bar{N}$ is the expected population size in equilibrium in the absence of catastrophes [see Eq. (2) in the Supplemental Material [27]]. The model assumes that the death rate is proportional to the total population size to balance the total population as in the standard nutrient restricted logistic model. A catastrophe deletes all the individuals of a given type with a total rate of γ (we further show that models with only partial deletion yields similar results). The probability that a type would be extinct in a catastrophe is not affected by its size. $\alpha, \gamma, \bar{N}$, and μ are free parameters. Without loss of generality, α can be set equal to 1 through a time rescaling (see [17] and Fig. 1 for a detailed derivation of the model).

Time is discretized with time steps $1/m_1$. Within this short time interval, a type of size k will either not change, grow to a type of size $k + 1$, lose one of its members and become a type of size $k - 1$ or disappear completely through a catastrophe or through a death if its size was 1. In equilibrium, total births and deaths are equal: $(\alpha/m_1) = (1/\bar{N}) + (\gamma/m_0)$. Since only nonmutation birth events increase the current type size, the probability to increase is $k([\alpha(1 - \mu)]/m_1)$. Similarly, the probability for a type to decrease by death is $(k/\bar{N}) = k([\alpha/m_1] - [\gamma/m_0])$ and by catastrophe is (γ/m_0). Denoting by T_k the average time to extinction for a type of initial size k, we obtain Eq. (1). We solved this system with a matrix inversion (see the Supplemental Material [27] for derivations). To avoid
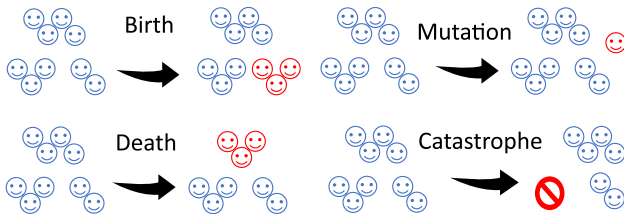

FIG. 1. Description of the model dynamics. Birth: the number of individuals in a given type increases by 1. Death: the number of individuals in a given type decreases by 1. Mutation: a constant fraction of all birth events leads to the creation of new types of size 1. Catastrophe: all individuals in a type are deleted.

restrictive assumptions on the type size distribution (as in [17]), we used numerical estimates for m_0 and m_1 obtained from simulations reaching steady state, for each value of γ.

$$T_1 = 1 + \frac{\alpha(1 - \mu)}{m_1} T_2 + \left[1 - \frac{\alpha(1 - \mu)}{m_1} - \left(\frac{\alpha}{m_1} - \frac{\gamma}{m_0}\right)\right] T_1,$$

$$T_k = 1 + k\frac{\alpha(1 - \mu)}{m_1} T_{k+1} + k\left(\frac{\alpha}{m_1} - \frac{\gamma}{m_0}\right) T_{k-1}$$
$$+ \left[1 - k\frac{\alpha(1 - \mu)}{m_1} - k\left(\frac{\alpha}{m_1} - \frac{\gamma}{m_0}\right) - \frac{\gamma}{m_0}\right] T_k. \quad (1)$$

We derived similar results using a Galton-Watson (GW) process. In a typical GW process, a population is assumed to have an initial size of 1. At each time step, each member can be replaced by i identical descendants. Here, we set $i \in 0, 1, 2$. Death corresponds to $i = 0$, no event to $i = 1$, and birth to $i = 2$. We denote by p_i the probability of getting i descendants. The time to extinction in a regular GW process, namely when the population reaches a size of 0, is obtained through the probability generating function $G(x) = p_0 + p_1 x + p_2 x^2$. Denoting by d_t the probability for this type to be extinct at time t, the GW model states that $d_t = G(d_{t-1})$. If the initial size is k instead of 1, then the probability to be extinct at time t is d_t^k since all k members behave like independent types of size 1, ignoring the possible interactions through the logistic term and assuming a large total population. Since the GW process only describes the events of regular birth or death and does not account for the catastrophes, the survival probability must be multiplied by the probability that no catastrophe occurred. Denoting by $\tilde{d}_{t,k} = 1 - (1 - d_t^k)e^{(-\gamma t/m_0)}$, the catastrophe corrected probability of extinction, one gets:

$$p_0 = \frac{\alpha}{m_1} - \frac{\gamma}{m_0}, \qquad p_1 = 1 - p_0 - p_2, \qquad p_2 = \frac{\alpha(1 - \mu)}{m_1},$$
$$T_k = \sum_t t(\tilde{d}_{t,k} - \tilde{d}_{t-1,k}). \quad (2)$$

As expected, the GW process leads to the same times to extinction as in the analytical method. Finally, we show the accuracy of our two approaches by testing them against simulations (lower plots in Fig. 2), with similar results (see the Supplemental Material [27] for methodologies).

In classical BDIM (i.e., no catastrophe) [18,25], there is a natural extinction survival transition at a zero net growth rate, and the times to extinction increase with respect to the initial type size (green lines in upper right plot of Fig. 2). However, in the BDICM, two other transitions appear. If γ is higher than μ, a transition to extinction occurs since fewer new types are created by mutations than the ones destroyed. As γ approaches μ, another intermediate transition is observed, where the times to extinction not only decrease but also become uniform across all initial type
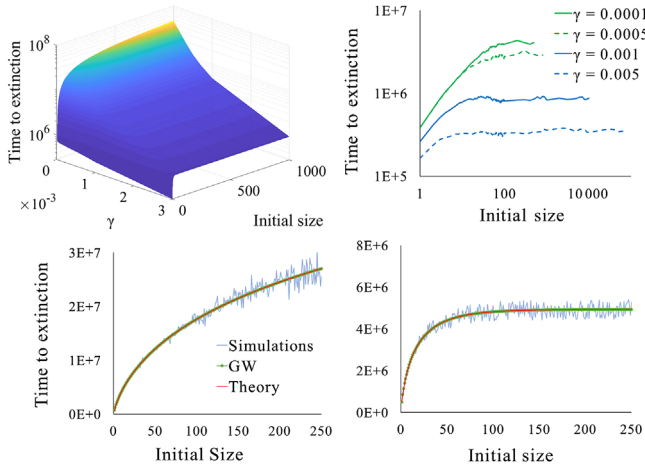
FIG. 2. Upper left plot: times to extinction with respect to the catastrophe rate γ and the initial type size. Upper right plot: times to extinction for different values of γ. Green curves correspond to the "negative invasion rate" phase where large types live longer but never reach a high size. Blue curves correspond to the "positive invasion rate" phase where catastrophes level the lifetime expectancy but types can reach higher sizes. Lower plots: comparison between simulations, GW process, and analytical expected times to extinction. The left plot is for $\gamma = 0$ and the right plot for $\gamma = 0.001$. Without catastrophe, times to extinction grow with type size and can reach high values. As γ increases, times to extinction are reduced and become uniform across type sizes for sufficiently large types.

FIG. 3. On each plot the right scale represents the diversity as defined as m_0/m_1 and the entry rate as defined as the number of mutations minus the deaths by catastrophe. The left scale represents the invasion rate as the log of the average growth rate of types after one step. (a) Complete deletion for each catastrophe as baseline model. (b) Deletion of a fixed fraction of a type in each catastrophe. (c) Catastrophe rate proportional to the population size. (d) Deletion of a random fraction of a type. We observe in all models that above a given catastrophe rate, the invasion rate becomes positive while diversity decreases.

sizes (upper left plot in Fig. 2) resulting in a change in the balance between small and large types. As will be explained in the next section, this transition is characterized by a positive invasion rate since large types can now disappear just as fast as small ones and faster than for $\gamma = 0$, creating room for new types to appear and grow. It is also characterized by a lower diversity since types can grow rapidly, before getting hit, to attain large sizes (longer range in the blue lines in upper right plot of Fig. 2) and hold a large fraction of the total population.

Invasion rate vs diversity.—Using the GW model, one can compute the average growth after one step as:

$$\mathbb{E}[\text{growth after one step}] = \ln\left[\frac{1}{k}\frac{d[G(x)]^k}{dx}\bigg|_{x=1}\right],$$
$$= \ln\left(1 - p_0 + p_2\right). \quad (3)$$

Note that this growth rate is independent of the initial type size k and its sign is entirely determined by the sign of $p_2 - p_0$. Without catastrophes, on average, existing type sizes would consistently decrease (since new types are created, and the total birth and death rate must equilibrate). This decrease leads them to the 0 absorbing state. The presence of catastrophes ensures that, in equilibrium, a transition occurs for $\gamma = [(\alpha\mu m_0)/m_1]$, where the average growth rate of all types becomes positive until their
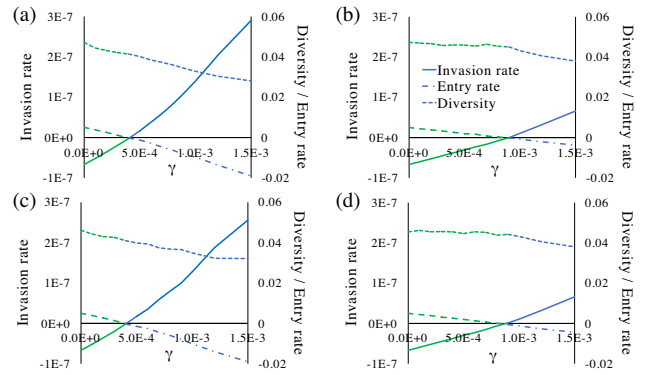
annihilation by a catastrophe as shown in Fig. 3. For this same γ value, we also observe from the simulations that the slope of the diversity (m_0/m_1) becomes more negative and the entry rate becomes negative, which confirms the results obtained with the GW model. The entry rate is defined as the difference between the number of births by mutation and the number of deaths by catastrophe. Since the equilibrium ensures that regular births plus mutations equal natural deaths plus catastrophes, the entry rate is opposite to the growth rate.

This transition, although closely related, is different from the one observed in [17]. The transition described there is between "low variance" and "high variance" phases for $\gamma = (2\alpha\mu m_0)/m_1$. Our transition appears at a lower γ, and the diversity starts to decrease before the second moment of the type size distribution diverges.

So far, the catastrophe rate was neither a function of the population size nor the number of types. When a catastrophe event occurs, a type is randomly chosen to be eliminated with equal probability over all types. To assess the robustness of our results, we compared other catastrophe models: two in which, for each catastrophe, only a fraction (fixed or random) of the population of one type is destroyed, and one where the catastrophe rate is proportional to m_1 like the regular death rate (see the Supplemental Material [27]). As shown in Fig. 3, as γ increases, the fraction of deaths induced by catastrophes becomes large enough and the transition emerges.

To summarize, in BDICM, in the "positive invasion rate" phase, although the lifespan of each type is on average low since they can be deleted via catastrophes, the growth rate is positive across all types. Therefore, several new types emerge and rapidly invade the system. At the same time,

surviving types grow exponentially and become so large that the diversity decreases, breaking the paradigm that a positive invasion rate implies higher diversity.

Comparison to FDS models.—This "positive invasion rate" phase is similar to balancing selection in that it promotes the growth of rare types. Moreover, in BDIM, small types have a disadvantage since they have a higher probability to reach the 0 absorbing state. However, when death is dominated by catastrophes, large types also have a high probability to reach the absorbing state. This removes their advantage over small types and gives the later a higher relative chance to grow. Other models were developed to explain balancing selection. For instance, Karev *et al.* [26] proposed a nonlinear BDIM, which corresponds to FDS. In their linear BDIM (without FDS), birth and death rates are equal and proportional to the type size. This model does not explain balancing selection and it requires extremely long times to reach equilibrium. Those issues are resolved in their nonlinear FDS model. However, this model implies the unrealistic assumption that individuals die or give birth knowing the size of their type. Such an assumption might occur if each type had a distinct resource, but it is highly improbable in a model with hundreds of types.

The BDICM with $\gamma = 0$ reproduces the same times to extinction as the linear model [26]. In the nonlinear model, a positive fixed birth rate a (type size independent) is assumed for each type and the individual birth rate α is set to be lower to achieve equilibrium $[\alpha(k + a)]/m_1 = (k/\bar{N})$ (see the Supplemental Material [27] for details). The times to extinction with FDS are equivalent to the ones obtained for a weak catastrophe rate, suggesting a similarity between classical negative FDS and the BDICM (Fig. 4).

Populations in HLA.—One of the systems where the contrast between diversity and a positive invasion rate is the

clearest is the MHC locus [28]. MHC proteins present self and foreign peptides to immune system cells. The genes coding for these proteins (denoted HLA in humans) are by far the most polymorphic in the human genome with thousands of alleles of each of the main classical HLA genes [23,29]. The main arguments currently used to explain the origins of this extreme genetic diversity in a specific locus are different aspects of balancing selection where new types have an advantage over existing types [21]. However, recent evidence suggests that the HLA haplotype distribution is biased toward excessive large types [24,30–32].

To test whether BDICM is consistent with the HLA frequencies, we computed the HLA-A frequency distribution in different populations in the United States, and the catastrophe rate producing the best fit for this distribution. A catastrophe rate of 0 cannot explain the observed distribution. However, a low catastrophe rate can produce an excellent fit (see Fig. 4 for the Caucasian population). This is not evidence that catastrophes are the mechanism explaining the distribution, but it shows their plausibility. In this context, catastrophes would represent a destructive allele, that, in some conditions, can kill all individuals that carry it. The distributions for other alleles and populations are similar [23,24,30].

Conclusions.—Catastrophes partly or fully erasing populations are frequent in many domains ranging from ecology to market dynamics [33–35]. In the population dynamics context, such events can be the eradication of local populations through weather events or diseases. In the genetic context, this can happen if a gene induces susceptibility to a certain pathogen. Introducing even a small amount of catastrophes to a BDIM inherently changes the dynamics. Specifically, there is a transition to an "invasion" regime, where the lifespan of types is mainly determined by catastrophes and not by the diffusion to the 0 absorbing state.

The intuition is simple. In the absence of catastrophes, births and deaths must be balanced leading to an Ewens-like type-size distribution [15]. However, in the presence of catastrophes, the average birth rate of each type can be higher than its death rate, with the total population balanced by catastrophes. This results in a positive growth rate for all type sizes and a positive invasion rate. In such a case, the sizes of types are determined by the combination of exponential growth and geometric distribution of time to catastrophe, leading to a skewing of the distribution toward large types. These large types can exist in parallel with a high invasion rate since the lifespan of each type is very low.

This model demonstrates that the paradigm in which the invasion rate induces diversity does not hold. Moreover, it shows that, in order to maintain diversity, the focus should be on limiting the size of large types instead of boosting the invasion rate. The presence of catastrophes can explain
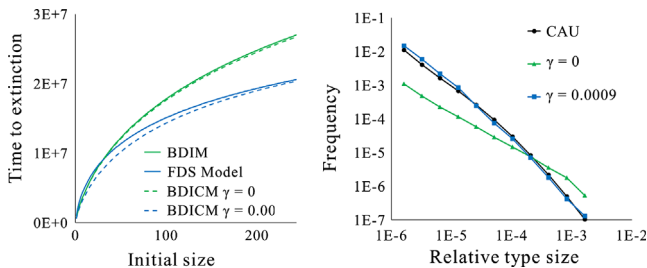

FIG. 4. Left plot: simulated and analytical times to extinction computed with the modified model with FDS introduced in the form of a constant birth rate a and a reduced relative birth rate α. For $a = 0$, the curves fit with our model without catastrophe. When adding balancing selection, we observe that the behavior is similar to adding catastrophes. Hence, the catastrophe model replicates balancing selection. Right plot: allele frequencies. The CAU curves represent the frequencies of HLA alleles for the Caucasian population. We compared it with simulations with and without catastrophes. We observe that the presence of catastrophes reproduces the frequencies of the alleles and could explain their distribution.

balancing selection without resorting to fitness differences or nonlinear models. It also explains the emergence and growth of large types within short times. Since this transition is only the result of a change in the balance between the total birth and death events, it is not affected by the details of the catastrophes. Similarly, Wilcox *et al.* [36] showed that density dependent catastrophes do not change the conclusions of their model.

––––––––––––

*Corresponding author.
louzouy@math.biu.ac.il

[1] I. Hanski, A practical model of metapopulation dynamics, J. Anim. Ecol. **63**, 151 (1994).

[2] F. Smith, Population dynamics in Daphnia magna and a new model for population growth, Ecology **44**, 651 (1963).

[3] P. Stephens, W. Sutherland, and R. Freckleton, What is the Allee effect?, **87**, Oikos, 185 (1999).

[4] S. J. Schreiber, Allee effects, extinctions, and chaotic transients in simple population models, Theor. Popul. Biol. **64**, 201 (2003).

[5] N. J. Gotelli, Metapopulation models: the rescue effect, the propagule rain, and the core-satellite hypothesis, Am. Nat. **138**, 768 (1991).

[6] M. Liermann and R. Hilborn, Depensation: evidence, models and implications, Fish Fisher. **2**, 33 (2001), http://www.seaturtle.org/PDF/LiermannM_2001_FishFish.pdf.

[7] J. C. Fay, G. Wyckoff, and C. Wu, Positive and negative selection on the human genome, Genetics **158**, 1227 (2001), https://www.genetics.org/content/158/3/1227.short.

[8] M. P. Hassell, Density dependence in single species populations, J. Anim. Ecol. **44**, 283 (1975).

[9] P. Chesson, General theory of competitive coexistence in spatially varying environments, Theor. Popul. Biol. **58**, 211 (2000).

[10] P. Chesson, Mechanisms of maintenance of species diversity, Annu. Rev. Ecol. Syst. **31**, 343 (2000).

[11] S. P. Ellner, R. E. Snyder, P. B. Adle, and G. Hooker, An expanded modern coexistence theory for empirical applications, Ecol. Lett. **22**, 3 (2019).

[12] E. Simpson, Measurement of diversity, Nature (London) **163**, 688 (1949).

[13] M. Nei, Analysis of gene diversity in subdivided populations, Proc. Natl. Acad. Sci. U. S. A. **70**, 3321 (1973).

[14] J. Hughes, G. Daily, and P. Ehrlich, Population diversity: its extent and extinction, Science **278**, 689 (1997).

[15] W. Ewens, The sampling theory of selectively neutral alleles, Theor. Popul. Biol. **3**, 87 (1972).

[16] K. Yuan, P. Bentler, and W. Chan, Structural equation modeling with heavy tailed distributions, Psychometrika **69**, 421 (2004).

[17] N. Dori, H. Behar, H. Brot, and Y. Louzoun, Family size variability grows with collapse rate in a birth death catastrophe model, Phys. Rev. E **98**, 012416 (2018).

[18] S. Karlin and J. McGregor, On a genetics model of Moran, Math. Proc. Cambridge Philos. Soc. **58**, 299 (1962).

[19] F. Kelly, Exact results for the Moran neutral allele model, Adv. Appl. Probab. **9**, 197 (1977).

[20] F. Key, J. Teixeira, C. de Filippo, and A. Andrés, Advantageous diversity maintained by balancing selection in humans, Curr. Opin. Genet. Dev. **29**, 45 (2014).

[21] P. Hedrick and G. Thomson, Evidence for balancing selection at HLA, Genetics **104**, 449 (1983), https://www.genetics.org/content/104/3/449.short.

[22] M. Fitzpatrick, E. Feder, L. Rowe, and M. Sokolowski, Maintaining a behaviour polymorphism by frequency dependent selection on a single gene, Nature (London) **447**, 210 (2007).

[23] N. Slater, Y. Louzoun, L. Gragert, M. Maiers, A. Chatterjee, and M. Albrecht, Power laws for heavy tailed distributions: modeling allele and haplotype diversity for the national marrow donor program, PLoS Comput. Biol. **11**, e1004204 (2015).

[24] I. Alter, L. Gragert, S. Fingerson, M. Maiers, and Y. Louzoun, HLA class I haplotype diversity is consistent with selection for frequent existing haplotypes, PLoS Comput. Biol. **13**, e1005693 (2017).

[25] G. Karev, Y. Wolf, A. Rzhetsky, F. Berezovskaya, and E. Koonin, Birth and death of protein domains: a simple model of evolution explains power law behavior, BMC Evol. Biol. **2**, 18 (2002).

[26] G. Karev, Y. Wolf, and E. Koonin, Simple stochastic birth and death models of genome evolution: was there enough time for us to evolve?, Bioinformatics **19**, 1889 (2003).

[27] See the Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.124.158301 for model derivations and methodologies.

[28] J. de Castro, J. Strominger, D. Strong, and H. Orr, Structure of crossreactive human histocompatibility antigens HLA A28 and HLA A2: possible implications for the generation of HLA polymorphism, Proc. Natl. Acad. Sci. U.S.A. **79**, 3813 (1982).

[29] A. Zachary, A. Steinberg, W. Bias, and M. Leffell, The frequencies of HLA alleles and haplotypes and their distribution among donors and renal patients in the UNOS registry 1, Transplantation **62**, 272 (1996).

[30] A. Lobkovsky, L. Levi, Y. Wolf, M. Maiers, L. Gragert, I. Alter, Y. Louzoun, and E. Koonin, Multiplicative fitness, rapid haplotype discovery, and fitness decay explain evolution of human MHC, Proc. Natl. Acad. Sci. U.S.A. **116**, 201714436 (2019), https://www.pnas.org/content/116/28/14098.short.

[31] C. Van Oosterhout, A new theory of MHC evolution: beyond selection on the immune genes, Proc. R. Soc. B **276**, 657 (2009).

[32] S. Sommer, The importance of immune gene variability (MHC) in evolutionary ecology and conservation, Front. Zool. **2**, 16 (2005).

[33] C. Hsieh, S. Glaser, A. Lucas, and G. Sugihara, Distinguishing random environmental fluctuations from ecological catastrophes for the North Pacific Ocean, Nature (London) **435**, 336 (2005).

[34] R. Knight and D. Pretty, Corporate catastrophes, stock returns, and trading volume, Corp. Rep. Rev. **2**, 363 (1999).

[35] C. Thomann, , The impact of catastrophes on insurer stock volatility, J. Risk Insur. **80**, 65 (2013).

[36] C. Wilcox and B. Elderd, The effect of density dependent catastrophes on population persistence time, J. Appl. Ecol. **40**, 859 (2003), https://www.jstor.org/stable/3505769?seq=1.