System drift and speciation

September 1, 2021

Abstract

Even if a species' phenotype does not change over evolutionary time, the underlying mechanism may change, as distinct molecular pathways can realize identical phenotypes. Here we use linear system theory to explore the consequences of this idea, describing how a gene network underlying a conserved phenotype evolves, as the genetic drift of small changes to these molecular pathways cause a population to explore the set of mechanisms with identical phenotypes. To do this, we model an organism's internal state as a linear system of differential equations for which the environment provides input and the phenotype is the output, in which context there exists an exact characterization of the set of all mechanisms that give the same input–output relationship. This characterization implies that selectively neutral directions in genotype space should be common and that the evolutionary exploration of these distinct but equivalent mechanisms can lead to the reproductive incompatibility of independently evolving populations. This evolutionary exploration, or *system drift*, is expected to proceed at a rate proportional to the amount of intrapopulation genetic variation divided by the effective population size (N_e). At biologically reasonable parameter values this could lead to substantial interpopulation incompatibility, and thus speciation, on a time scale of N_e generations. This model also naturally predicts Haldane's rule, thus providing a concrete explanation of why heterogametic hybrids tend to be disrupted more often than homogametes during the early stages of speciation.

Key words: Speciation, Models/Simulations, Genetic Drift

Introduction

It is an overarching goal of many biological subdisciplines to attain a general understanding of the function and evolution of the complex molecular machinery that translates an organism's genome into the characteristics on which natural selection acts, the phenotype. For example, there is a growing body of data on the evolutionary histories and molecular characterizations of particular gene regulatory networks [Jaeger, 2011, Davidson and Erwin, 2006, Israel et al., 2016], as well as thoughtful verbal and conceptual models [True and Haag, 2001, Weiss and Fullerton, 2000, Edelman and Gally, 2001, Pavlicev and Wagner, 2012]. Mathematical models of both particular regulatory networks and the evolution of such systems in general can provide guidance where intuition fails, and thus have the potential to discover general principles in the organization of biological systems as well as provide concrete numerical predictions [Servedio et al., 2014]. There is a substantial amount of work studying the evolution of gene regulatory networks, in frameworks both abstract [Wagner, 1994, 1996, Siegal and Bergman, 2002, Bergman and Siegal, 2003, Draghi and Whitlock, 2015] and empirically inspired [Mjolsness et al., 1991, Jaeger et al., 2004, Kozlov et al., 2015, Crombach et al., 2016, Wotton et al., 2015, Chertkova et al., 2017].

At all levels of biological organization, the problems that biological systems have evolved to solve often do not have single solutions – systems can be structurally different yet remain functionally equivalent [Edelman and Gally, 2001]. Examples can be found across nearly all levels of biological organization from the level of the genetic code itself all the way up to the convergent evolution of adaptive traits. In many cases, these functionally equivalent structures can be explored through small, local changes to the structure that leave the function unchanged. For instance, there are "neutral networks" of nucleic acid sequences that produce the same RNA secondary structure [Grüner et al., 1996] amino acid sequences that fold similarly [Babajide et al., 1997], or proteins with equivalent thermodynamic stability [Hart et al., 2014]. Further examples are found in the vast space of functionally equivalent potential regulatory sequences [Hare et al., 2008], in the logic of transcriptional [Tsong et al., 2006, Matsui et al., 2015, Dalal et al., 2016, Dalal and Johnson, 2017, Jiménez et al., 2017] and neural circuits [Trojanowski et al., 2014], and in developmental systems [von Dassow et al., 2000, True and Haag, 2001].

This capacity for isofunctional yet distinct mechanisms, sometimes called *degeneracy*, is a consequence of a many-to-one mapping between a system's structure and function, a concept that has been explored in many fields beyond biology. For instance, in many contexts mathematical models are fundamentally *nonidentifiable* and/or *indistinguishable* – meaning that there can be uncertainty about an inferred model's parameters or even its claims about causal structure, despite access to complete and perfect data [e.g., Bellman and Åström, 1970, Grewal and Glover, 1976, Walter et al., 1984]. Models with different parameter schemes, or even different mechanics can make equally accurate predictions, but still not actually reflect the internal dynamics of the system being modeled. In control theory, where electrical circuits and mechanical systems are often the focus, it is understood that there can be an infinite number of "realizations", or ways to reverse engineer the dynamics of a "black box", even if all possible input and output experiments are performed [Kalman, 1963, Anderson et al., 1966, Zadeh and Deoser, 1976]. The inherent nonidentifiability of chemical reaction networks is sometimes referred to as "the fundamental dogma of chemical kinetics" [Craciun and Pantea, 2008]. In computer science, this has been framed as the relationship among processes that *simulate* one another [Van der Schaft, 2004]. Finally, the field of *inverse problems* studies those cases in which, despite the existence of a theoretical one-to-one mapping between a model and behavior, tiny amounts of noise make inference problems nonidentifiable in practice [Petrov and Sizikov, 2005].

It has been argued that the ability to modify structure without affecting function is necessary for natural selection [Edelman and Gally, 2001], as it may function as a mechanism for biological robustness and evolvability [reviewed in de Visser et al., 2003], or manifest as *canalization* [Whitacre, 2010]. It may even contribute to the formation of new species [Gavrilets, 2014]. Redundancy of the genetic code, for instance, can make sequences more fault-tolerant to mutations [Sonneborn, 1965], and robustness to modification of genetic networks can allow adaptation without passing through a fitness valley [Wagner, 2008].

In this paper we use results on mathematical nonidentifiability from linear systems theory to study how gene regulatory networks can be modified while retaining the same function, and the possible implications for speciation. If system architectures are not functionally unique, can this open up neutral evolutionary paths, and do species explore these paths through the process termed *developmental system drift* [by True and Haag, 2001]? Is this fast enough to contribute meaningfully to speciation? To do this, we describe results on linear dynamical systems which give an analytical description of the set of all linear gene network architectures that yield identical phenotypes, and use quantitative genetics theory to estimate the speed at which system drift can lead to reproductive incompatibility and hence speciation. In this model, a population diffuses along the neutral ridges of a high-dimensional space of possible system parameters, in a similar vein as *holey landscape* models [Gavrilets, 1997, Yamaguchi and Iwasa, 2013].

The field of population genetics has also explored the consequences of the fact that there is often more than one way to do the same thing, and observed that speciation might be the result of changes that are themselves neutral. Indeed, Bateson [1909] first proposed that what today we call a Bateson-Dobzhansky-Muller incompatibility would likely arise through neutral changes. The potential for speciation has been analyzed in models of traits under stabilizing selection determined additively by alleles at many loci [Wright, 1935, Barton, 1986, 1989, 2001], in related fitness landscape models [Fraïsse et al., 2016], and for pairs of traits that must match but whose value is unconstrained [Sved, 1981]. It has also been shown that population structure can allow long-term stable coexistence of incompatible genotypes encoding identical phenotypes [Phillips, 1996]. However, previous simulations of system drift in regulatory sequences [Tulchinsky et al., 2014] and a regulatory cascade [Porter and Johnson, 2002] found rapid speciation under directional selection but only equivocal support for speciation under models of purely neutral drift. The rate at which hybrid incompatibility accumulates due to genetic drift creating segregation variance between isolated populations is fairly well understood [Slatkin and Lande, 1994, Rosas et al., 2010, Chevin et al., 2014], but model assumptions can strongly affect predictions, including whether variation is due to rare or common alleles [Slatkin and Lande, 1994], and the shape of the fitness landscape [Fraïsse et al., 2016]. Our main aim is to provide a concrete framework that can make natural predictions of these model parameters across a general class of models. Furthermore, tools from system theory allow analytical predictions to be made for large populations with complex phenotypes that would be inaccessible to population simulations.

Results

We use a model of gene regulatory networks that describes the temporal dynamics of a collection of n coregulating molecules – such as transcription factors – as well as external or environmental inputs. We write $\kappa(t)$ for the vector of n molecular concentrations at time t. The vector of m "inputs" determined exogenously to the system is denoted $u(t)$, and the vector of ℓ "outputs" is denoted $\phi(t)$. The output is merely a linear function of the internal state: $\phi_i(t) = \sum_j C_{ij}\kappa_j(t)$ for some matrix C. Since ϕ is what natural selection acts on, we refer to it as the *phenotype* (meaning the "visible" aspects of the organism), and in contrast refer to κ as the *kryptotype*, as it is "hidden" from direct selection. Although ϕ may depend on all entries of κ, it is usually of lower dimension than κ, and we tend to think of it as the subset of molecules relevant for survival. The dynamics are determined by the matrix of regulatory coefficients, A, a time-varying vector of inputs $u(t)$, and a matrix B that encodes the effect of each entry of u on the elements of the kryptotype. The rate at which the ith concentration changes is a weighted sum of the concentrations as well as the input:

$$\begin{aligned} \dot{\kappa}(t) &= A\kappa(t) + Bu(t) \\ \phi(t) &= C\kappa(t). \end{aligned} \tag{1}$$

Furthermore, we always assume that $\kappa(0) = 0$, so that the kryptotype measures deviations from initial concentrations. Here A can be any $n \times n$ matrix, B any $n \times m$, and C any $\ell \times n$ dimensional matrix, with usually ℓ and m less than n. We think of the system as the triple (A, B, C), which translates (time-varying) m-dimensional input $u(t)$ into the ℓ-dimensional output $\phi(t)$. Under quite general assumptions on the input (e.g., $|u(t)|$ is integrable) we can write the phenotype as

$$\phi(t) = \int_0^t Ce^{(t-s)A}Bu(s)ds, \tag{2}$$

which is a convolution of the input $u(t)$ with the system's *impulse response*, which we denote as $h(t) := Ce^{At}B$.

In terms of gene regulatory networks, A_{ij} determines how the jth transcription factor regulates the ith transcription factor. If $A_{ij} > 0$, then κ_j upregulates κ_i, while if $A_{ij} < 0$, then κ_j downregulates κ_i. The ith row of A is therefore determined by genetic features such as the strength of j-binding sites in the promoter of gene i, factors affecting chromatin accessibility near gene i, or basal transcription machinery activity. The form of B determines how the environment influences transcription factor expression levels, and C might determine the rate of production of downstream enzymes.

Wagner [1994] and others have used a similar discrete-time model (that might be written $\phi_{t+1} = f(A\phi_t)$, where f is a sigmoid). Our choice of continuous time does not affect the points we make here, but our restriction to *linear* systems is a stronger assumption (see the Discussion).

To demonstrate the model, we construct a simple gene network in Example 1 below.

Example 1 (An oscillator)**.** *For illustration, we consider an extremely simplified model of oscillating gene transcription, as for instance is found in cell cycle control or the circadian rhythm. There are two genes, whose transcript concentrations are given by $\kappa_1(t)$ and $\kappa_2(t)$, and gene-2 upregulates gene-1, while gene-1 downregulates gene-2 with equal strength. Only the dynamics of gene-1 are consequential to the oscillator (perhaps the amount of gene-1 activates another downstream gene network). Lastly, both genes are equally upregulated by an exogenous signal. The dynamics of the system are described by*

$$\begin{aligned} \dot{\kappa_1}(t) &= \kappa_2(t) + u(t) \\ \dot{\kappa_2}(t) &= -\kappa_1(t) + u(t) \\ \phi(t) &= \kappa_1(t). \end{aligned}$$

In matrix form the system regulatory coefficients are given as, $A = \left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right]$, $B = \left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]$, and $C = \left[\begin{smallmatrix} 1 & 0 \end{smallmatrix}\right]$. If the input is an impulse at time zero (a delta function), then the phenotype is equal to the impulse response:

$$\phi(t) = h(t) = \sin t + \cos t.$$

The system and its dynamics are referred to in Figure 1. We return to the evolution of such a system below.

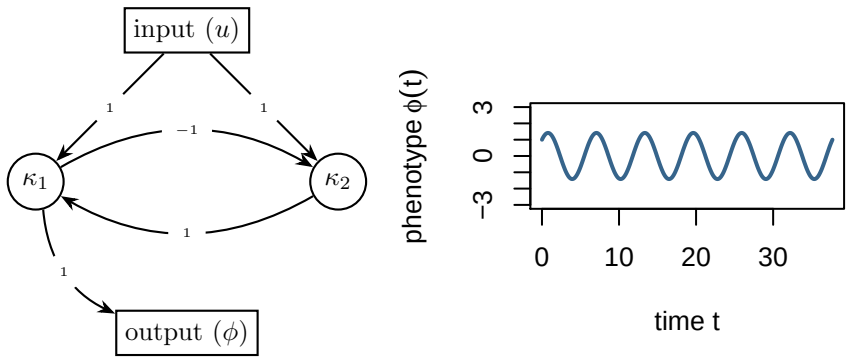

Figure 1: (Left) Diagram of the gene network in Example 1 , and (right) plot of the phenotype $\phi(t)$ against time t .

Equivalent gene networks

As reviewed above, some systems with identical phenotypes are known to differ, sometimes substantially, at the molecular level; systems with identical phenotypes do not necessarily have identical kryptotypes. How many different mechanisms perform the same function?

Two systems are equivalent if they produce the same phenotype given the same input, i.e., have the same input–output relationship. We say that the systems defined by (A, B, C) and $(\bar{A}, \bar{B}, \bar{C})$ are **phenotypically equivalent** if their impulse response functions are the same: $h(t) = \bar{h}(t)$ for all $t \geq 0$. This implies that for any acceptable input $u(t)$, if $(\kappa_u(t), \phi_u(t))$ and $(\bar{\kappa}_u(t), \bar{\phi}_u(t))$ are the solutions to equation (1) of these two systems, respectively, then

$$\phi_u(t) = \bar{\phi}_u(t) \qquad \text{for all } t \geq 0.$$

In other words, phenotypically equivalent systems respond identically for *any* input.

One way to find other systems phenotypically equivalent to a given one is by change of coordinates: if V is an invertible matrix, then the systems (A, B, C) and (VAV^{-1}, VB, CV^{-1}) are phenotypically equivalent because their impulse response functions are equal:

$$\begin{aligned} h(t) &= Ce^{At}B = CV^{-1}Ve^{At}V^{-1}VB \\ &= CV^{-1}e^{VAV^{-1}t}VB = \bar{C}e^{\bar{A}t}\bar{B} = \bar{h}(t). \end{aligned} \tag{3}$$

These "changes of coordinates" are not simply different ways of looking at the same system – if each dimension of the kryptotype corresponds to the concentration of a particular transcription factor, changing A corresponds to changing the strengths of regulatory interactions. We will even see below that interactions may change sign. However, not all phenotypically equivalent systems are of this form: systems can have identical impulse responses without being coordinate changes of each other. In fact, systems with identical impulse responses can involve interactions between different numbers of molecules, and thus have kryptotypes in different dimensions altogether.

This implies that most systems have at least n^2 degrees of freedom, where recall n is the number of components of the kryptotype vector. This is because for an arbitrary $n \times n$ matrix Z, taking V to be the identity matrix plus a small perturbation in the direction of Z above implies that moving A in the direction of $ZA - AZ$ while also moving B in the direction of ZB and C in the direction of $-CZ$ will leave the phenotype unchanged to second order in the size of the perturbation. If the columns of B and the rows of C are not all eigenvectors of A, then any such Z will result in a different system.

It turns out that in general, there are more degrees of freedom, except if the system is *minimal* – meaning, informally, that it uses the smallest possible number of components to achieve the desired dynamics. Results in system theory show that any time-invariant system can be realized in a particular minimal dimension (the dimension of the kryptotype, $n_{\min}$), and that any two phenotypically equivalent systems of dimension $n_{\min}$ are related by a change of coordinates. Since gene networks can grow or shrink following gene duplications and deletions, these additional degrees of freedom can apply, in principle, to any system.

Even if the system is not minimal, results from systems theory explicitly describe the set of all phenotypically equivalent systems. We refer to $\mathcal{N}(A_0, B_0, C_0)$ as the set of all systems phenotypically equivalent to the system defined by (A_0, B_0, C_0):

$$\mathcal{N}(A_0, B_0, C_0) = \left\{(A, B, C) : Ce^{At}B = C_0 e^{A_0 t} B_0 \text{ for } t \geq 0\right\}. \tag{4}$$

These systems need not have the same kryptotypic dimension n, but must have the same input and output dimensions (ℓ and m, respectively).

The Kalman decomposition, which we now describe informally, elegantly characterizes this set [Kalman, 1963, Kalman et al., 1969, Anderson et al., 1966]. To motivate this, first note that the input $u(t)$ only directly pushes the system in certain directions (those lying in the span of the columns of B). As a result, different combinations of input can move the system in any direction that lies in what is known as the *reachable subspace*. Analogously, we can only observe motion of the system in certain directions (those lying in the span of the rows of C), and so can only infer motion in what is known as the *observable subspace*. The Kalman decomposition then classifies each direction in kryptotype space as either reachable or unreachable, and as either observable or unobservable. Only the components that are both reachable and observable determine the system's phenotype – that is, components that both respond to an input and produce an observable output.

Concretely, the **Kalman decomposition** of a system (A, B, C) gives a change of basis P such that the transformed system (PAP^{-1}, PB, CP^{-1}) can be written in block matrix form:

$$PAP^{-1} = \begin{bmatrix} A_{r\bar{o}} & A_{r\bar{o},ro} & A_{r\bar{o},\bar{r}\bar{o}} & A_{r\bar{o},\bar{r}o} \\ 0 & A_{ro} & 0 & A_{ro,\bar{r}o} \\ 0 & 0 & A_{\bar{r}\bar{o}} & A_{\bar{r}\bar{o},\bar{r}o} \\ 0 & 0 & 0 & A_{\bar{r}o} \end{bmatrix},$$

and

$$PB = \begin{bmatrix} B_{r\bar{o}} \\ B_{ro} \\ 0 \\ 0 \end{bmatrix} \qquad\qquad (CP^{-1})^T = \begin{bmatrix} 0 \\ C_{ro}^T \\ 0 \\ C_{\bar{r}o}^T \end{bmatrix}.$$

The n-dimensional system has been divided into subspaces of dimensions $n_{r\bar{o}} + n_{ro} + n_{\bar{r}\bar{o}} + n_{\bar{r}o} = n$, and so, for instance, $A_{r\bar{o}}$ is the $n_{r\bar{o}} \times n_{r\bar{o}}$ square matrix in the top left corner of PAP^{-1}. The impulse response of the system is given by

$$h(t) = C_{ro} e^{A_{ro} t} B_{ro},$$

and therefore, the system is phenotypically equivalent to the *minimal* system (A_{ro}, B_{ro}, C_{ro}).

This decomposition is unique up to a change of basis that preserves the block structure. In particular, the minimal subsystem obtained by the Kalman decomposition is unique up to a change of coordinates. This implies that there is no equivalent system with a smaller number of kryptotypic dimensions than the dimension of the minimal system. It is remarkable that the gene regulatory network architecture to achieve a given input–output map is never unique – both the change of basis used to obtain the decomposition and, once in this form, all submatrices other than A_{ro}, B_{ro}, and C_{ro} can be changed without affecting the phenotype, and so represent degrees of freedom.

Although the other submatrices do not affect the phenotype through the impulse response function $h(t)$, they may affect individual fitness in other ways (which we don't consider here). For instance, the "unreachable" subspaces cannot be affected by any of the possible variations in input produced by varying $u(t)$. However, if the nature of the input changed (for instance, by perturbing a previously constant environmental variable), differences in the unreachable components might lead to differences in phenotypes and thus represent cryptic genetic variation. Similarly, "unobservable" subspaces might represent aspects of phenotype that are not constrained by selection, but might become important if conditions change. Additionally, all subspaces may affect how the system deals with noise.

Note on implementation: The *reachable subspace* is defined to be the closure of span(B) under applying A (or equivalently, the span of $B, AB, A^2B, \ldots A^{n-1}B$), and the *unobservable subspace* is the largest A-invariant subspace contained in the null space of C. The four subspaces, $r\bar{o}$, ro, $\bar{r}\bar{o}$, and $\bar{r}o$ are defined from these by intersections and orthogonal complements – ro refers to the both *reachable and observable* subspace, while $\bar{r}\bar{o}$ refers to the *unreachable and unobservable* subspace, and similarly for $\bar{r}o$ and $r\bar{o}$.

For the remainder of the paper, we interpret $\mathcal{N}$ as the neutral set in the fitness landscape, along which a large population will drift under environmental and selective stasis. This drift need not be purely neutral – for instance, second-order selection on robustness will push the species towards "flatter" areas of genotype space [Rice, 1998, Hermisson et al., 2003]. Even if the phenotype is constrained and remains constant through evolutionary time, the molecular mechanism underpinning it is not constrained and likely will not be conserved.

Finally, note that if B and C are held constant – i.e., if the relationships between environment, kryptotype, and phenotype do not change – there are *still* usually degrees of freedom. Example 2 gives the set of minimal systems equivalent to the oscillator of Example 1, that all share common B and C matrices. The oscillator can also be equivalently realized by a three-gene (or larger) network, and will have even more evolutionary degrees of freedom available, as in Figure 3.

Example 2 (All equivalent rewirings of the oscillator)**.** *The oscillator of Example 1 is minimal, and so any equivalent system is a change of coordinates by an invertible matrix V. If we further require B and C to be invariant then we need $VB = B$ and $CV = C$. Therefore the following one-parameter family $(A(\tau), B, C)$ describes the set of all two-gene systems phenotypically equivalent to the oscillator:*

$$A(\tau) = \frac{-1}{\tau+1}\begin{bmatrix} -\tau & -1 \\ 2\tau(\tau+1)+1 & \tau \end{bmatrix} \text{ for } \tau \neq -1.$$

The resulting set of systems are depicted in Figure 2.

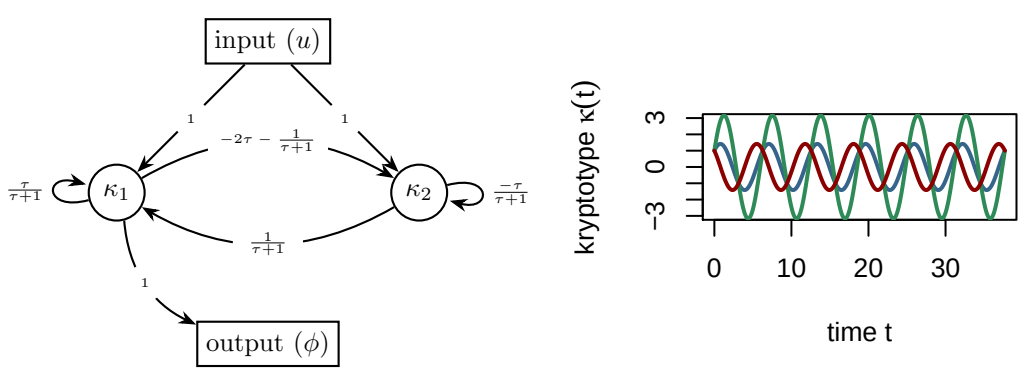

Figure 2: (Left) $A(\tau)$, the set of all phenotype-equivalent cell cycle control networks. (Right) Gene-1 dynamics (blue) for both systems $A(0)$ and $A(-2)$ are identical, however, $A(0)$ gene-2 dynamics (red) differ from $A(-2)$ (green).

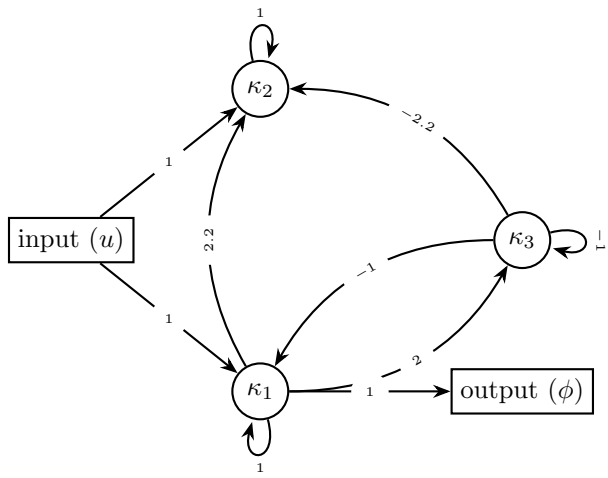

Figure 3: A possible non-minimal three-gene oscillator, phenotypically equivalent to $A(\tau)$, the systems in Examples 1 and 2.

Sexual reproduction and recombination Parents with phenotypically equivalent yet differently wired gene networks may produce offspring with dramatically different phenotypes. If the phenotypes are significantly divergent then the offspring may be inviable or otherwise dysfunctional, despite both parents being well adapted. If this is consistent for the entire population, we would consider them to be separate species, in accord with the biological species concept [Mayr, 2000].

First, we must specify how sexual reproduction acts on these systems. Suppose that each of a diploid organisms' two genomes encodes a set of system coefficients with the same kryptotype dimension. We assume that a diploid which has inherited systems (A', B', C') and (A'', B'', C'') from its two parents has phenotype determined by the system that averages these two, $((A' + A'')/2, (B' + B'')/2, (C' + C'')/2)$. In general, how parental systems are recombined and inherited depends on the genetic architecture of the segregating variation. For the purposes of examples in this paper (Example 3 and Figure 5), we take the simplest assumption that each system coefficient is encoded by a nonrecombining locus, and that these loci are unlinked. For instance, if one genome's coefficients are written with Greek and the other with Latin, a possible oscillator gamete might be produced as follows: $\begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \times \begin{bmatrix} a & b \\ c & d \end{bmatrix} \to \begin{bmatrix} \alpha & b \\ c & \delta \end{bmatrix}$.

Each genome an organism inherits is generated by meiosis, in which both of its diploid parents recombine their two genomes, and so an F_1 offspring carries one system copy from each parent, and an F_2 is an offspring of two independently formed F_1s. If the parents are from distinct populations, these are simply first– and second–generation hybrids, respectively.

Exactly how the coefficients (i.e., entries of A, B and C) of a haploid system inherited by an offspring from a diploid parent are determined by the parent's two systems depends on the genetic basis of any variation in the coefficients. Thanks to the randomness of meiotic segregation, the result is random to the extent that each parent is heterozygous for alleles that affect the coefficients. Since the ith row of A summarizes how each gene regulates gene i, and hence is determined by the promoter region of gene i, the elements of a row of A tend to be inherited together, which will create covariance between entries of the same row. It is, however, a quite general observation that the variation seen among recombinant systems is proportional to the difference between the two parental systems.

Offspring formed from two phenotypically identical systems do not necessarily exhibit the same phenotype as both of its parents – in other words $\mathcal{N}$, the set of all systems phenotypically equivalent to a given one, is not, in general, closed under averaging or recombination. If sexual recombination among systems drawn from $\mathcal{N}$ yields systems with divergent phenotypes, populations containing significant diversity in $\mathcal{N}$ can carry genetic load, and isolated populations may fail to produce hybrids with viable phenotypes.

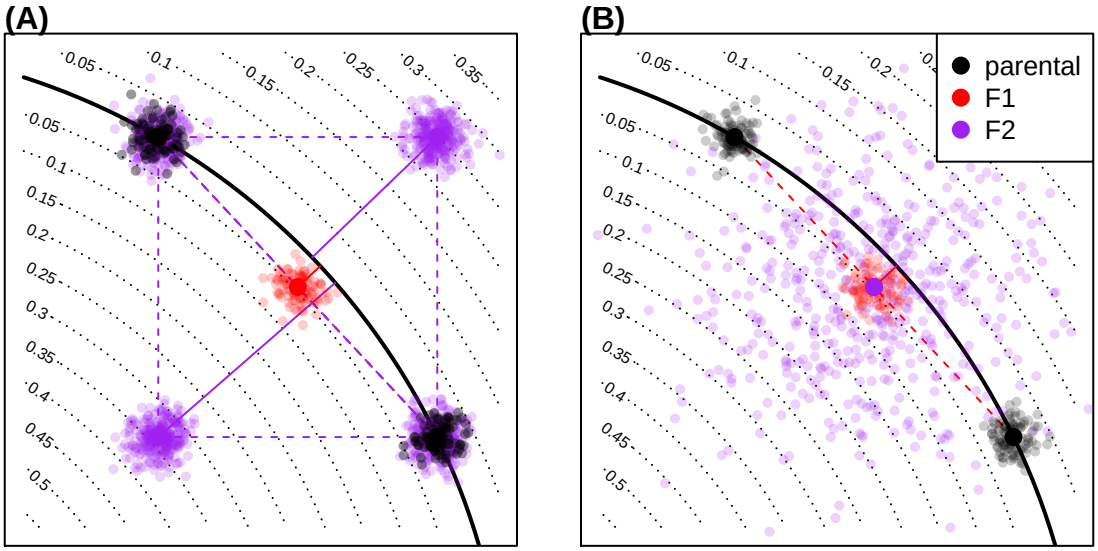

Figure 4: A conceptual figure of the fitness consequences of hybridization: axes represent system coefficients (i.e., entries of A); the line of optimal system coefficients is down in black; solid lines give phenotypic distances to the optimum. A pair of parental populations are shown in black, along the optimum; a hypothetical population of F_1s are shown in red, and the distribution of F_2s is shown in purple. The two figures differ in the genetic basis, and hence, the distribution of F_2 phenotypes: **(A)** F_2s compose all four mixed homozygotes if variation at both traits has a simple, one-locus genetic basis in both populations; and **(B)** F_2 show a much wider distribution of phenotypes if the genetic basis of variation in each population is polygenic.

Hybrid incompatibility

Two parents with the optimal phenotype can produce offspring whose phenotype is suboptimal if the parents have different underlying systems. Hybrid phenotypic break down, as a function of genetic distance between phenotypically equivalent parental oscillators (described in Example 2) is illustrated in Example 3. How quickly do hybrid phenotypes break down as genetic distance between parents increases? We will quantify how far a system's phenotype is from optimal using a weighted difference between impulse response functions. Suppose that $\rho(t)$ is a nonnegative weighting function, $h_0(t)$ is the *optimal* impulse response function and define the "distance to optimum" of another impulse response function to be

$$D(h) = \left(\int_0^\infty \rho(t) \|h(t) - h_0(t)\|^2 dt \right)^{1/2}. \tag{5}$$

In practice, we take $\rho(t) = \exp(-t/4\pi)$, so that fitness is determined by the dynamics of the system over a few multiples of 2π, but not longer. Consider reproduction between a parent with system (A, B, C) and another displaced by distance ϵ in the direction (X, Y, Z), i.e., having system $(A + \epsilon X, B + \epsilon Y, C + \epsilon Z)$. We assume both are "perfectly adapted" systems, i.e., having impulse response function $h_0(t)$, and their offspring has impulse response function $h_\epsilon(t)$. A Taylor expansion of $D(h_\epsilon)$ in ϵ is explicitly worked out in Appendix A, and shows that the phenotype of an F_1 hybrid between these two is at distance proportional to ϵ^2 from optimal, while F_2 hybrids are at distance proportional to ϵ. This is because an F_1 hybrid has one copy of each parental system, and therefore lies directly between the parental systems (see Figure 4) – the parents both lie in $\mathcal{N}$, which is the valley defined by D, and so their midpoint only differs from optimal due to curvature of $\mathcal{N}$. In contrast, an F_2 hybrid may be homozygous for one parental type in some coefficients and homozygous for the other parental type in others; this means that each coefficient of an F_2 may be equal to either one of the parents, or intermediate between the two; this means that possible F_2 systems may be as

far from the optimal set, $\mathcal{N}$, as the distance between the parents. The precise rate at which the phenotype of a hybrid diverges depends on the geometry of the optimal set $\mathcal{N}$ relative to segregating genetic variation.

Example 3 (Hybrid incompatibility: misregulation due to system drift)**.** *Offspring of two equivalent systems from Example 2 can easily fail to oscillate. For instance, the F_1 offspring between homozygous parents at $\tau = 0$ and $\tau = -2$ has phenotype $\phi_{F_1}(t) = e^t$, rather than $\phi(t) = \sin t + \cos t$. However, the coefficients of these two parental systems differ substantially, probably more than would be observed between diverging populations. In Figure 5 we compare the phenotypes for F_1 and F_2 hybrids between more similar parents, and see increasingly divergent phenotypes as the difference between the parental systems increases. (In this example, the coefficients of $A(\epsilon)$ differ from those of $A(0)$ by an average factor of $1 + \epsilon/2$; such small differences could plausibly be caused by changes to promoter sequences.) This divergence is quantified in Figure 6, which shows that mean distance to optimum phenotype of the F_1 and F_2 hybrid offspring between $A(0)$ and $A(\epsilon)$ increases with ϵ^2 and ϵ, respectively.*

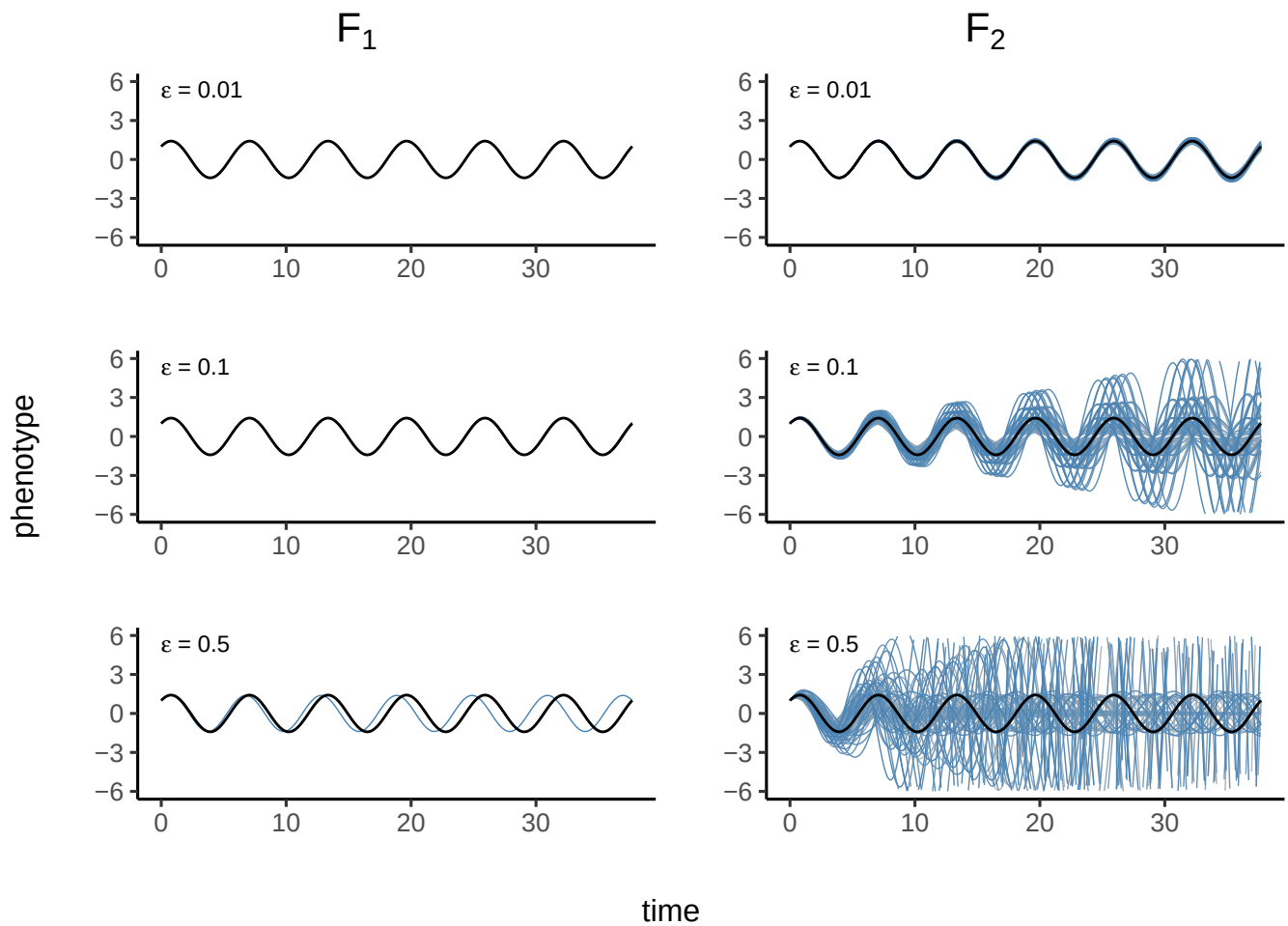

Figure 5: **(left)** Phenotypes of F_1 hybrids between a homozygous $A(0)$ parent and, top-to-bottom, homozygous $A(1/100)$, $A(1/10)$, and $A(1/2)$ parents, where $A(\epsilon)$ is defined in Figure 2; parental coefficients differ by around 0.5%, 5%, and 25% respectively. Parental phenotypes ($\sin t + \cos t$) are shown in solid black, and hybrid phenotypes in blue. **(right)** Phenotypes of all $3^4 = 81$ possible F_2 hybrids between the same set of parents, with parental phenotype again in black. F_2 hybrids, many of which show complete breakdown, are shown in shades of blue and grey.

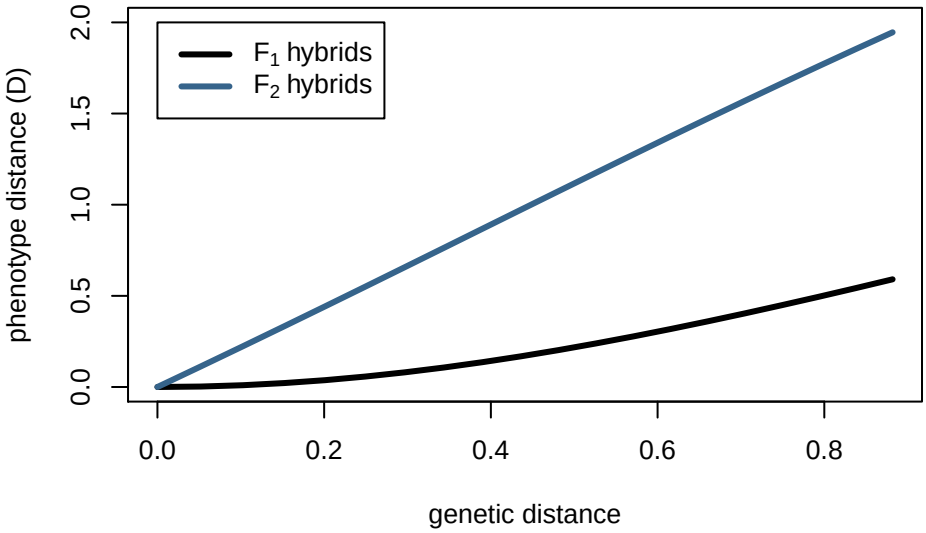

Figure 6: Mean hybrid phenotypic distance from optimum computed with equation (5), using $\rho(t) = \exp(-t/4\pi)$ for F_1 (black) and F_2 (blue) hybrids between $A(0)$ and $A(\epsilon)$ parent oscillators. Genetic distance is computed as $\left(\sum_{ij}(A_{ij}(0) - A_{ij}(\epsilon))^2\right)^{1/2}$.

Haldane's rule This model naturally predicts Haldane's rule, the observation that if only one hybrid sex is sterile or inviable it is likely the heterogametic sex (e.g., the male in XY sex determination systems) [Haldane, 1922, Orr, 1997]. For example, consider an XY species with a two-gene network where the first gene resides on an autosome and the second gene on the X chromosome. A male whose pair of haplotypes is $\left(\left[\begin{smallmatrix} A_1 & A_2 \\ \cdot & \cdot \end{smallmatrix}\right], \left[\begin{smallmatrix} A_1 & A_2 \\ X_1 & X_2 \end{smallmatrix}\right]\right)$ has phenotype determined by $A = \left[\begin{smallmatrix} A_1 & A_2 \\ X_1 & X_2 \end{smallmatrix}\right]$, if dosage compensation upregulates heterogametes by a factor of two relative to homogametes (as with *Drosophila*), while a female homozygous for the haplotype $\left[\begin{smallmatrix} \bar{A}_1 & \bar{A}_2 \\ \bar{X}_1 & \bar{X}_2 \end{smallmatrix}\right]$, has phenotype determined by $A = \left[\begin{smallmatrix} \bar{A}_1 & \bar{A}_2 \\ \bar{X}_1 & \bar{X}_2 \end{smallmatrix}\right]$. An F_1 male offspring of these two will have its phenotype determined by $\left[\begin{smallmatrix} (A_1+\bar{A}_1)/2 & (A_2+\bar{A}_2)/2 \\ \bar{X}_1 & \bar{X}_2 \end{smallmatrix}\right]$. If both genes resided on the autosomes, this system would only be possible in an F_2 cross. More generally, the contribution of sex chromosomes to the system coefficients of F_1 males are effectively equivalent to those of purely autosomal F_2 hybrids, which are less fit on average than F_1s (see Figure 6). Note, however, that presence of genetic variation on sex chromosomes does not alone guarantee Haldane's rule: under this model, Haldane's rule will only occur when there is sufficient genetic distance separating parentals and when the nature of genetic variation produces F_2 hybrid breakdown.

The speed of speciation

We have shown that system drift can lead to speciation in principle, but is it rapid enough to be an important factor in practice? In other words, after what period of time would we expect the fitness of hybrids between two allopatric populations to be substantially lower than the parentals? Selection – on pleiotropic traits or on robustness – may actively push even a strongly constrained system along neutral directions, but even the calculations under purely neutral drift are informative. The population mean of an unconstrained quantitative trait with additive genetic variance V_G in a population with effective size N_e will move in t generations a random amount whose variance is tV_G/N_e [Lande, 1976]. The mean difference between two such populations has twice the variance. Although this mean difference is along neutral directions, we would in many cases expect the range of variation among F_2s in *all* directions to be of the same order as the differences between the populations, as depicted in Figure 4. This suggests that, naively, two such

populations that have been separated for t generations will produce F_2 offspring that differ from optimal by an amount proportional to $\sqrt{tV_G/N_e}$. Since we assume they are at a local fitness optimum, without much loss of generality we can assume that fitness is locally quadratic, and so F_2 fitness decays linearly in time: proportionally to tV_G/N_e – fastest in small, diverse populations, This predicts that we need only wait some multiple of N_e generations until substantial incompatibility has been accumulated.

It is useful to think in more detail about the assumptions in the rough argument above. The key aspect is how population differences in neutral directions (along the fitness ridge) translate to segregation variance in F_2s in selectively constrained directions. To move the system (the A matrix) a given distance generally involves moving many individual interaction coefficients (the entries A_{ij}). The movements must be coordinated, for the population to stay near the fitness ridge. However, mixing elements between systems that have made independent sets of coordinated changes to remain on the fitness ridge is unlikely to produce a set of coordinated changes; and the resulting system could move away from the ridge in almost any direction.

Genetic variation in empirical regulatory systems

What is known about the key quantity above, the amount of heritable variation in real regulatory networks? The coefficient A_{ij} from the system (1) measures how much the rate of net production of i changes per change in concentration of j. It is generally thought that regulatory sequence change contributes much more to inter- and intraspecific variation than does coding sequence change affecting molecular structure [Schmidt et al., 2010]. In the context of transcription factor networks this may be affected not only by the binding strength of molecule j to the promoter region of gene i but also the effects of other transcription factors (e.g., cooperativity) and local chromatin accessibility [Stefflova et al., 2013]. For this reason, the mutational target size for variation in A_{ij} may be much larger than the dozens of base pairs typically implicated in the handful of binding sites for transcription factor j of a typical promoter region, and single variants may affect many entries of $\mathcal{N}$ simultaneously.

Variation in binding site occupancy may overestimate variation in A, since it does not capture buffering effects (if for instance only one site of many needs to be occupied for transcription to begin), and variation in expression level measures changes in steady-state concentration (our κ_i) rather than the *rate* of change. Nonetheless, these measures likely give us an idea of the scale of variability. It has been shown that between human individuals, there is differential occupancy in 7.5% of binding sites of a transcription factor (p65) [Kasowski et al., 2010]. It has also been inferred that cis-regulatory variation accounts for around 2–6% of expression variation in human blood-derived primary cells [Verlaan et al., 2009], and that human population variation explained about 3% of expression variation [Lappalainen et al., 2013]. Allele-specific expression is indicative of standing genetic *cis*-regulatory variation; allele-specific expression in 7.2–8.5% of transcripts of a flycatcher species has been observed [Wang et al., 2017], as well as allele-specific expression in 23.4% of genes studied in a baboon species [Tung et al., 2015]. Taken together, this suggests that variation in the entries of A may be on the order of at least a few percent between individuals of a population – doubtless varying substantially between species and between genes.

Discussion

In this paper, we use tools from linear system theory and quantitative genetics to study the evolution of a mechanistic model of the genotype-phenotype map, in which the phenotype is subject to stabilizing selection. In so doing, we provide an explicit model of phenogenetic drift [Weiss and Fullerton, 2000] and developmental system drift [True and Haag, 2001]. In this context, the Kalman decomposition [Kalman, 1963] gives an analytical description of all phenotypically equivalent gene networks. This description shows that the space of functionally equivalent network architectures increases with the square of a network's size, and that this space increases further if networks grow larger than absolutely necessary – that is use more interacting components than the most efficient potential architectures. In this framework, even minimal gene network architectures – efficient architectures that contain only the requisite number of interacting parts, are not structurally unique with respect to function. Functionally equivalent architectures are often related by

continuous parameter changes, suggesting that equivalent networks might be mutationally connected, and that there exist axes of genetic variation unconstrained by natural selection. The independent movement of separated populations along these axes by genetic drift can lead to a significant reduction in hybrid viability, and thus precipitate speciation, at a speed dependent on the effective population size and the amount of genetic variation. In this model, at biologically reasonable parameter values, system drift is a significant – and possibly rapid – driver of speciation. This may be surprising because hybrid inviability appears as a consequence of recombining different, yet functionally equivalent, mechanisms, and since species are often defined by their unique adaptations or morphologies.

Consistent with empirical observation of hybrid breakdown we see that the fitnesses of F_2 hybrids drop at a much faster rate than those of F_1s. Another natural consequence of the model is Haldane's rule, that if only one F_1 hybrid sex is inviable or sterile it is likely to be the heterogametic sex. This occurs because if the genes underlying a regulatory network are distributed among both autosomes and the sex chromosome, then heterogametic F_1s show variation (and fitnesses) similar to that seen in F_2 hybrids. This observation appears to be similar to the extreme hybrid phenotypes produced by transgressive segregation [Rieseberg et al., 1999], which can manifest in F_1s when only one (dominant) parental allele is expressed at heterozygous loci; this was observed in hybrid gene expression patterns, and increased as a function of parental genetic distance [Stelkens and Seehausen, 2009].

Is there evidence that this is actually occurring? System drift and network rewiring has been inferred across the tree of life [Wotton et al., 2015, Crombach et al., 2016, Dalal and Johnson, 2017, Johnson, 2017, Ali et al., 2019], and there is often significant regulatory variation segregating within populations. Transcription in hybrids between closely related species with conserved transcriptional patterns can also be divergent [Haerty and Singh, 2006, Maheshwari and Barbash, 2012, Coolon et al., 2014, Michalak and Noor, 2004, Mack and Nachman, 2016], and hybrid incompatibilities have been attributed to cryptic molecular divergence underlying conserved body plans [Gavin-Smyth and Matute, 2013]. Furthermore, in cryptic species complexes (e.g., sun skinks [Barley et al., 2013]), genetically distinct species may be nearly morphologically indistinguishable.

The origin of species not by means of natural selection? As classically formulated, the Dobzhansky-Muller model of hybrid incompatibility is agnostic to the relative importance of neutral versus selected genetic substitutions [Coyne and Orr, 1998], and plausible mechanisms have been proposed whereby Dobzhansky–Muller incompatibilities could originate under neutral genetic drift [Lynch and Force, 2000] or stabilizing selection [Fierst and Hansen, 2009]. The same holds for the "pathway model" [Lindtke and Buerkle, 2015], which is closer to the situation here. However, previous authors have argued that neutral processes are likely too slow to be a significant driver of speciation [Nei et al., 1983, Seehausen et al., 2014]. This has led some to conclude that hybrid incompatibility is typically a byproduct of positive selection [Orr et al., 2004, Schluter, 2009] or a consequence of genetic conflict [Presgraves, 2010, Crespi and Nosil, 2013], two processes that typically act much more rapidly than genetic drift. Supporting this view, experimental approaches designed to study the effect of drift on speciation through induced population bottlenecks found no reproductive isolation in the majority of tested pairs [Fry, White et al., 2020]. However, our calculations suggest that even under strictly neutral processes, hybrid fitness breaks down as a function of genetic distance rapidly enough to play a substantial role in species formation across the tree of life. This is consistent with broad patterns such as the relationship between molecular divergence and genetic isolation seen by Roux et al. [2016], and the clocklike speciation rates observed by Hedges et al. [2015]. More recently, Sánchez-Ramírez et al. [2021] observed widespread transcriptional misregulation in *Drosophila* hybrids, which might be due to the action of system drift.

Although the phrase "system drift" as we use it implies neutral changes, it does not mean that these changes are induced purely by neutral genetic drift. Selection could dramatically accelerate speciation by the same process – e.g., by genetic draft [Gillespie, 2000] or pleiotropic selection. For instance, Mani and Clarke [1990] showed that populations developed incompatibilities as they independently evolved towards the same quantitative trait optimum, as different alleles fix in different populations. Fluctuating selection has similarly been shown to accelerate speciation in some models [Bell, 2010, Barton, 2001]. Under our model,

populations evolving towards different optimal phenotypes will always produce less fit hybrids, as the hybrid phenotypes will be displaced from both parental optima, and so maladapted to both parental environments. This suggests that fluctuating selection – if not synchronized between populations – should diminish hybrid fitnesses. In the case where fluctuations are synchronized, due to the size of the neutral system space, populations may arrive at different system organizations when returning to previous phenotypes at a rate faster than by drift alone.

These explanations are not mutually exclusive. All of these forces – adaptive shifts, conflict and network drift – are plausible drivers of speciation, and may even interact. Many of our observations carry over to models of directional selection – for instance, rapid drift along the set of equivalent systems could be driven by adaptation in a different, pleiotropically coupled system. Or, reinforcement due to local adaptation might provide a selective pressure that speeds up system drift. Furthermore, while the fitness consequences of incompatibility in any one given network may be small, the cumulative impact of system drift across the many different networks an organism relies on may be substantial. It remains to be seen how the relative strengths of these forces compare.

The dimensionality of trait space We have focused on examples of single traits (where the phenotype is one-dimensional), but phenotypes under selection are often high-dimensional, and variation in different traits often share a genetic basis. However, we still expect many degrees of freedom as long as there are components of the system not directly and individually constrained by selection (i.e., a kryptotype). Even in networks where the phenotype and kryptotype are of the same dimension, system theory shows us that there will always be available degrees of freedom as specific system realizations are only unique up to a change of coordinates. Some phenotypes, however, require kryptotypic dimensions to be larger than that of the phenotype. For instance, many systems have minimal realizations (*e.g.*, the oscillator in Example 2) where the dimension of the kryptotype is larger than that of the phenotype, implying that for these phenotypic dynamics to be realized, the kryptotype dimension *has* to be larger than the dimension of the phenotype. Even if components of the system's internal state are directly subject to selection and the mode of action of the environment on the internal state is constrained (so, the input and output matrices B and C are fixed) then one could still perturb A as described above by $ZA - AZ$ if ZB and CZ are both zero, implying a number of degrees of freedom that still grows with n^2 for fixed ℓ and m. Generically, the number of degrees of freedom is $n(n - \ell - m)$, so that in a system of n components, if even one component is not directly constrained, this leads to n degrees of freedom. Whatever the true "dimensionality" of phenotype space of a typical organism, there are undoubtedly aspects of its underlying molecular machinery that are not directly constrained, suggesting large numbers of degrees of freedom. Note that pleiotropy does not directly affect this argument at all – indeed, many phenotypically equivalent changes will lead to denser A matrices and hence more pleiotropy. However, more pleiotropic genes may be more strongly constrained, making it more difficult for systems to make the required compensatory changes for system drift.

Phenotypically equivalent system evolution is probably not only driven by neutral genetic drift. For one thing, movement along the set of equivalent networks is not expected to be completely neutral, since second-order selection pushes populations towards "flatter" regions of the fitness landscape in which a population centered on the optimal set has lower genetic load [as described in different contexts by Rice, 1998, Nimwegen et al., 1999]. If this bias towards more robust networks is strong enough, it may even prevent drift, but it is unclear how strong this effect would be in practice. Our results, on the other hand, do not rely on the flatness of the fitness surface around the phenotypically equivalent set, but rather on the curvature of the equivalent set itself. So long as the phenotypically equivalent set is not closed under sexual recombination, opportunities for incompatibility remain. However the speed at which system drift can generate incompatibilities might diminish if selection for robustness is strong enough to constrain a population to a small section of system space, although the strength of such effects in practice are not known. Likewise, as the speed of system drift relies on segregating genetic variation, any constraints on such variation, possibly due to epistasis, genetic architecture [Hermisson et al., 2003], adaptive inertia [Baatz and Wagner, 1997, Álvarez-Castro et al., 2009] or weak gene flow could plausibly slow it down. More work on specific systems, likely coupled with simulations, will be necessary to identify the biologically relevant parameter regimes.

Geometric models The model we work with describes a continuum of possible systems, parameterized by continuous coefficient matrices. The arguments we make (e.g., Figure 4) are often geometric in nature, and so it would be interesting to ask whether Fisher's geometric model might provide more quantitative, generalizable predictions, as in for instance Fraïsse et al. [2016], Simon et al. [2017]. Martin [2014] argued that such an approximation would be often fruitful. In the other direction, it is also natural to ask about the distribution of effect sizes of segregating mutations in a population evolving under this model with a realistic *de novo* effect size distribution, as well as what the distribution of fitness effects would be (both within and between populations). These are questions that might be best answered by simulation, and could provide much additional information – for instance, are alleles typically dominant in their effects on either system coefficients or fitness?

Nonlinearity and model assumptions Of course, real regulatory networks are not linear dynamical systems. Most notably, physiological limits put upper bounds on expression levels, implying saturating response curves. It remains to be seen how well our results carry over into real systems, but the fact that most nonlinear systems can be locally approximated by a linear one suggests our qualitative results may hold more generally. Furthermore, nonidentifiability (which implies the existence of neutral directions) is often found in practice in moderately complex models of biological systems [e.g., Gutenkunst et al., 2007, Piazza et al., 2008, Jiménez et al., 2017].

Finally, despite our model's precise separation of phenotype and kryptotype, this relationship in nature may be far more complicated as aspects of the kryptotype may be less "hidden" than we currently assume, and the neutral network changes we describe here may only be nearly neutral. For instance, attributes excluded from the phenotype as modeled here ignore the potential energy costs associated with excessively large (non-minimal) kryptotypes, as well as the relationship between a specific network architecture and robustness to mutational, transcriptional, or environmental noise. More precise modeling will require better mechanistic understanding not only of biological systems, but also the nature of selective pressures and genetic variation in these systems.

Acknowledgements

We would like to thank Sergey Nuzhdin, Stevan Arnold, Michael Turelli, Patrick Phillips, Erik Lundgren and Hossein Asgharian for valuable discussion. We would also like to thank Nick Barton, Sarah Signor, Todd Parsons, and Joachim Hermisson for very helpful comments on the manuscript. Work on this project was supported by funds from the Sloan Foundation and the NSF (under DBI-1262645) to PR.

References

Sammi Ali, Sarah A Signor, Konstantin Kozlov, and Sergey V Nuzhdin. Novel approach to quantitative spatial gene expression uncovers genetic stochasticity in the developing drosophila eye. *Evolution & Development*, 21(3):157–171, 2019. 12

José M Álvarez-Castro, Michael Kopp, and Joachim Hermisson. Effects of epistasis and the evolution of genetic architecture: exact results for a 2-locus model. *Theoretical population biology*, 75(2-3):109–122, 2009. 13

BDO Anderson, RW Newcomb, RE Kalman, and DC Youla. Equivalence of linear time-invariant dynamical systems. *Journal of the Franklin Institute*, 281(5):371–378, 1966. 2, 5

M Baatz and G.P Wagner. Adaptive inertia caused by hidden pleiotropic effects. *Theoretical Population Biology*, 51(1):49–66, February 1997. doi: 10.1006/tpbi.1997.1294. 13

Aderonke Babajide, Ivo L Hofacker, Manfred J Sippl, and Peter F Stadler. Neutral networks in protein space: a computational study based on knowledge-based potentials of mean force. *Folding and Design*, 2 (5):261–269, 1997. 1

Anthony J Barley, Jordan White, Arvin C Diesmos, and Rafe M Brown. The challenge of species delimitation at the extremes: diversification without morphological change in Philippine sun skinks. *Evolution*, 67(12): 3556–3572, 2013. 12

Nicholas H. Barton. The maintenance of polygenic variation through a balance between mutation and stabilizing selection. *Genet Res*, 47(3):209–216, June 1986. doi: 10.1017/S0016672300023156. URL https://www.ncbi.nlm.nih.gov/pubmed/3744046. 2

Nicholas H. Barton. The divergence of a polygenic system subject to stabilizing selection, mutation and drift. *Genetics Research*, 54(1):59–78, 1989. 2

Nicholas H. Barton. The role of hybridization in evolution. *Molecular Ecology*, 10(3):551–568, 2001. 2, 12

W. Bateson. *Heredity and Variation in Modern Lights*, pages 85–101. Cambridge University Press, Cambridge, 1909. ISBN 9781108004350. doi: DOI:10.1017/CBO9780511693953.007. 2

Graham Bell. Fluctuating selection: the perpetual renewal of adaptation in variable environments. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1537):87–97, 2010. 12

Richard Ernest Bellman and Karl Johan Åström. On structural identifiability. *Mathematical biosciences*, 7 (3-4):329–339, 1970. 2

Aviv Bergman and Mark L Siegal. Evolutionary capacitance as a general feature of complex gene networks. *Nature*, 424(6948):549–552, 2003. 1

Aleksandra A. Chertkova, Joshua S. Schiffman, Sergey V. Nuzhdin, Konstantin N. Kozlov, Maria G. Samsonova, and Vitaly V. Gursky. In silico evolution of the Drosophila gap gene regulatory sequence under elevated mutational pressure. *BMC Evolutionary Biology*, 17(1):4, 2017. ISSN 1471-2148. doi: 10.1186/s12862-016-0866-y. URL http://dx.doi.org/10.1186/s12862-016-0866-y. 1

Luis-Miguel Chevin, Guillaume Decorzent, and Thomas Lenormand. Niche dimensionality and the genetics of ecological speciation. *Evolution*, 68(5):1244–1256, 2014. 2

Joseph D Coolon, C Joel McManus, Kraig R Stevenson, Brenton R Graveley, and Patricia J Wittkopp. Tempo and mode of regulatory evolution in Drosophila. *Genome research*, 24(5):797–808, 2014. 12

Jerry A Coyne and H Allen Orr. The evolutionary genetics of speciation. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 353(1366):287–305, 1998. 12

Gheorghe Craciun and Casian Pantea. Identifiability of chemical reaction networks. *Journal of Mathematical Chemistry*, 44(1):244–259, 2008. 2

Bernard Crespi and Patrik Nosil. Conflictual speciation: species formation via genomic conflict. *Trends in Ecology & Evolution*, 28(1):48–57, 2013. 12

Anton Crombach, Karl R Wotton, Eva Jiménez-Guri, and Johannes Jaeger. Gap gene regulatory dynamics evolve along a genotype network. *Molecular biology and evolution*, 33(5):1293–1307, 2016. 1, 12

Chiraj K Dalal and Alexander D Johnson. How transcription circuits explore alternative architectures while maintaining overall circuit output. *Genes & Development*, 31(14):1397–1405, 2017. 1, 12

Chiraj K Dalal, Ignacio A Zuleta, Kaitlin F Mitchell, David R Andes, Hana El-Samad, and Alexander D Johnson. Transcriptional rewiring over evolutionary timescales changes quantitative and qualitative properties of gene expression. *eLife*, 5:e18981, 2016. 1

Eric H Davidson and Douglas H Erwin. Gene regulatory networks and the evolution of animal body plans. *Science*, 311(5762):796–800, 2006. 1

J. Arjan G. M. de Visser, Joachim Hermisson, Günter P. Wagner, Lauren Ancel Meyers, Homayoun Bagheri-Chaichian, Jeffrey L. Blanchard, Lin Chao, James M. Cheverud, Santiago F. Elena, Walter Fontana, Greg Gibson, Thomas F. Hansen, David Krakauer, Richard C. Lewontin, Charles Ofria, Sean H. Rice, George von Dassow, Andreas Wagner, and Michael C. Whitlock. Perspective: Evolution and detection of genetic robustness. *Evolution*, 57(9):1959–1972, 2003. doi: https://doi.org/10.1111/j.0014-3820.2003.tb00377.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.0014-3820.2003.tb00377.x. 2

Jeremy Draghi and Michael Whitlock. Robustness to noise in gene expression evolves despite epistatic constraints in a model of gene networks. *Evolution*, 69(9):2345–2358, 2015. 1

Gerald M Edelman and Joseph A Gally. Degeneracy and complexity in biological systems. *Proceedings of the National Academy of Sciences*, 98(24):13763–13768, 2001. 1, 2

Janna L Fierst and Thomas F Hansen. Genetic architecture and postzygotic reproductive isolation: evolution of Bateson-Dobzhansky-Muller incompatibilities in a polygenic model. *Evolution*, 2009. 12

C Fraïsse, P A Gunnarsson, D Roze, N Bierne, and J J Welch. The genetics of speciation: Insights from Fisher's geometric model. *Evolution*, 70(7):1450–1464, 07 2016. doi: 10.1111/evo.12968. URL https://www.ncbi.nlm.nih.gov/pubmed/27252049. 2, 14

James D Fry. Laboratory experiments on speciation. *Concepts, Methods, and Applications of Selection Experiments*, page 631. 12

Jackie Gavin-Smyth and Daniel R Matute. Embryonic lethality leads to hybrid male inviability in hybrids between *Drosophila melanogaster* and *D. santomea*. *Ecology and Evolution*, 3(6):1580–1589, 2013. 12

Sergey Gavrilets. Evolution and speciation on holey adaptive landscapes. *Trends in ecology & evolution*, 12 (8):307–312, 1997. 2

Sergey Gavrilets. Models of speciation: Where are we now? *Journal of heredity*, 105(S1):743–755, 2014. 2

John H. Gillespie. Genetic drift in an infinite population: The pseudohitchhiking model. *Genetics*, 155(2): 909–919, 2000. ISSN 0016-6731. URL http://www.genetics.org/content/155/2/909. 12

M Grewal and K Glover. Identifiability of linear and nonlinear dynamical systems. *IEEE Transactions on Automatic Control*, 21(6):833–837, Dec 1976. doi: 10.1109/TAC.1976.1101375. 2

Walter Grüner, Robert Giegerich, Dirk Strothmann, Christian Reidys, Jacqueline Weber, Ivo L Hofacker, Peter F Stadler, and Peter Schuster. Analysis of rna sequence structure maps by exhaustive enumeration i. neutral networks. *Monatshefte für Chemie/Chemical Monthly*, 127(4):355–374, 1996. 1

Ryan N Gutenkunst, Joshua J Waterfall, Fergal P Casey, Kevin S Brown, Christopher R Myers, and James P Sethna. Universally sloppy parameter sensitivities in systems biology models. *PLoS Computational Biology*, 3(10):e189, 2007. 14

Wilfried Haerty and Rama S Singh. Gene regulation divergence is a major contributor to the evolution of Dobzhansky–Muller incompatibilities between species of Drosophila. *Molecular Biology and Evolution*, 23 (9):1707–1714, 2006. 12

JBS Haldane. Sex ratio and unisexual sterility in hybrid animals. *Journal of Genetics*, 12(2):101–109, 1922. 10

Emily E Hare, Brant K Peterson, Venky N Iyer, Rudolf Meier, and Michael B Eisen. Sepsid even-skipped enhancers are functionally conserved in Drosophila despite lack of sequence conservation. *PLoS Genetics*, 4(6):e1000106, 2008. 1

Kathryn M Hart, Michael J Harms, Bryan H Schmidt, Carolyn Elya, Joseph W Thornton, and Susan Marqusee. Thermodynamic system drift in protein evolution. *PLoS biology*, 12(11), 2014. 1

S Blair Hedges, Julie Marin, Michael Suleski, Madeline Paymer, and Sudhir Kumar. Tree of life reveals clock-like speciation and diversification. *Molecular Biology and Evolution*, 32(4):835–845, 2015. 12

Joachim Hermisson, Thomas F Hansen, and Günter P Wagner. Epistasis in polygenic traits and the evolution of genetic architecture under stabilizing selection. *The American Naturalist*, 161(5):708–734, 2003. 6, 13

Jennifer W Israel, Megan L Martik, Maria Byrne, Elizabeth C Raff, Rudolf A Raff, David R McClay, and Gregory A Wray. Comparative developmental transcriptomics reveals rewiring of a highly conserved gene regulatory network during a major life history switch in the sea urchin genus Heliocidaris. *PLoS Biology*, 14(3):e1002391, 2016. 1

Johannes Jaeger. The gap gene network. *Cellular and Molecular Life Sciences*, 68(2):243–274, 2011. 1

Johannes Jaeger, Svetlana Surkova, Maxim Blagov, Hilde Janssens, David Kosman, Konstantin N Kozlov, Ekaterina Myasnikova, Carlos E Vanario-Alonso, Maria Samsonova, David H Sharp, et al. Dynamic control of positional information in the early Drosophila embryo. *Nature*, 430(6997):368–371, 2004. 1

Alba Jiménez, James Cotterell, Andreea Munteanu, and James Sharpe. A spectrum of modularity in multi-functional gene circuits. *Molecular Systems Biology*, 13(4):925, April 2017. doi: 10.15252/msb.20167347. 1, 14

Alexander D Johnson. The rewiring of transcription circuits in evolution. *Current Opinion in Genetics & Development*, 47:121–127, 2017. 12

Rudolf E. Kalman. Mathematical description of linear dynamical systems. *J. SIAM Control*, 1963. 2, 5, 11

Rudolf E. Kalman, Peter L. Falb, and Michael A. Arbib. *Topics in mathematical system theory*. McGraw-Hill, New York, 1969. ISBN 0754321069. 5

M. Kasowski, F. Grubert, C. Heffelfinger, M. Hariharan, A. Asabere, S. M. Waszak, L. Habegger, J. Rozowsky, M. Shi, A. E. Urban, M. Y. Hong, K. J. Karczewski, W. Huber, S. M. Weissman, M. B. Gerstein, J. O. Korbel, and M. Snyder. Variation in transcription factor binding among humans. *Science*, 328(5975): 232–235, April 2010. 11

Konstantin Kozlov, Vitaly V Gursky, Ivan V Kulakovskiy, Arina Dymova, and Maria Samsonova. Analysis of functional importance of binding sites in the Drosophila gap gene network model. *BMC Genomics*, 2015. 1

Russell Lande. Natural selection and random genetic drift in phenotypic evolution. *Evolution*, 30(2):pp. 314–334, 1976. ISSN 00143820. URL http://www.jstor.org/stable/2407703. 10

Tuuli Lappalainen, Michael Sammeth, Marc R Friedländer, Peter AC 't Hoen, Jean Monlong, Manuel A Rivas, Mar Gonzalez-Porta, Natalja Kurbatova, Thasso Griebel, Pedro G Ferreira, et al. Transcriptome and genome sequencing uncovers functional variation in humans. *Nature*, 501(7468):506–511, 2013. 11

Dorothea Lindtke and C Alex Buerkle. The genetic architecture of hybrid incompatibilities and their effect on barriers to introgression in secondary contact. *Evolution*, 69(8):1987–2004, 2015. 12

Michael Lynch and Allan G Force. The origin of interspecific genomic incompatibility via gene duplication. *The American Naturalist*, 156(6):590–605, 2000. 12

Katya L Mack and Michael W Nachman. Gene regulation and speciation. *Trends in Genetics*, 2016. 12

Shamoni Maheshwari and Daniel A Barbash. Cis-by-trans regulatory divergence causes the asymmetric lethal effects of an ancestral hybrid incompatibility gene. *PLoS Genetics*, 8(3):e1002597, 2012. 12

GS Mani and Bryan Campbell Clarke. Mutational order: a major stochastic process in evolution. *Proceedings of the Royal Society of London. B. Biological Sciences*, 240(1297):29–37, 1990. 12, 2

Guillaume Martin. Fisher's geometrical model emerges as a property of complex integrated phenotypic networks. *Genetics*, 197(1):237–255, February 2014. doi: 10.1534/genetics.113.160325. URL https://www.genetics.org/content/197/1/237. 14, 2

Takeshi Matsui, Robert Linder, Joann Phan, Fabian Seidl, and Ian M Ehrenreich. Regulatory rewiring in a cross causes extensive genetic heterogeneity. *Genetics*, 201(2):769–777, 2015. 1

Ernst Mayr. The biological species concept. *Species concepts and phylogenetic theory: a debate. Columbia University Press, New York*, pages 17–29, 2000. 7

Pawel Michalak and Mohamed AF Noor. Association of misexpression with sterility in hybrids of *Drosophila simulans* and *D. mauritiana*. *Journal of Molecular Evolution*, 59(2):277–282, 2004. 12

Eric Mjolsness, David H Sharp, and John Reinitz. A connectionist model of development. *Journal of Theoretical Biology*, 152(4):429–453, 1991. 1

Masatoshi Nei, Takeo Maruyama, and Chung-I Wu. Models of evolution of reproductive isolation. *Genetics*, 103(3):557–579, 1983. 12

Erik Van Nimwegen, James P Crutchfield, and Martijn Huynen. Neutral evolution of mutational robustness. *PNAS*, 1999. 13

H Allen Orr. Haldane's rule. *Annual Review of Ecology and Systematics*, 28(1):195–218, 1997. 10

H Allen Orr, John P Masly, and Daven C Presgraves. Speciation genes. *Current Opinion in Genetics & Development*, 14(6):675–679, 2004. 12

Mihaela Pavlicev and Günter P Wagner. A model of developmental evolution: selection, pleiotropy and compensation. *Trends in Ecology & Evolution*, 27(6):316–322, 2012. 1

Y.P. Petrov and V.S. Sizikov. *Well-posed, ill-posed, and intermediate problems with applications*, volume 49. Walter de Gruyter, 2005. 2

Patrick C Phillips. Maintenance of polygenic variation via a migration–selection balance under uniform selection. *Evolution*, 50(3):1334–1339, 1996. 2

Matthew Piazza, Xiao-Jiang Feng, Joshua D Rabinowitz, and Herschel Rabitz. Diverse metabolic model parameters generate similar methionine cycle dynamics. *Journal of Theoretical Biology*, 251(4):628–639, 2008. 14

Adam H Porter and Norman A Johnson. Speciation despite gene flow when developmental pathways evolve. *Evolution*, 56(11):2103–2111, 2002. 2

Daven C Presgraves. The molecular evolutionary basis of species formation. *Nature Reviews Genetics*, 11 (3):175–180, 2010. 12

Sean H. Rice. The evolution of canalization and the breaking of von Baer's laws: Modeling the evolution of development with epistasis. *Evolution*, 52(3):647–656, 1998. doi: https://doi.org/10.1111/j.1558-5646.1998.tb03690.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1558-5646.1998.tb03690.x. 6, 13

Loren H Rieseberg, Margaret A Archer, and Robert K Wayne. Transgressive segregation, adaptation and speciation. *Heredity*, 83(4):363–372, 1999. 12

Ulises Rosas, Nick H. Barton, Lucy Copsey, Pierre Barbier de Reuille, and Enrico Coen. Cryptic variation between species and the basis of hybrid performance. *PLOS Biology*, 8(7):1–12, 07 2010. doi: 10.1371/journal.pbio.1000429. URL https://doi.org/10.1371/journal.pbio.1000429. 2

Camille Roux, Christelle Fraisse, Jonathan Romiguier, Yoann Anciaux, Nicolas Galtier, and Nicolas Bierne. Shedding light on the grey zone of speciation along a continuum of genomic divergence. *PLoS Biology*, 14 (12):e2000234, 2016. 12

Santiago Sánchez-Ramírez, Jörg G Weiss, Cristel G Thomas, and Asher D Cutter. Widespread misregulation of inter-species hybrid transcriptomes due to sex-specific and sex-chromosome regulatory evolution. *PLoS genetics*, 17(3):e1009409, 2021. 12

Dolph Schluter. Evidence for ecological speciation and its alternative. *Science*, 323(5915):737–741, 2009. 12

D Schmidt, M D Wilson, B Ballester, P C Schwalie, G D Brown, A Marshall, C Kutter, S Watt, C P Martinez-Jimenez, S Mackay, I Talianidis, P Flicek, and D T Odom. Five-vertebrate ChIP-seq reveals the evolutionary dynamics of transcription factor binding. *Science*, 328(5981):1036–1040, May 2010. doi: 10.1126/science.1186176. URL https://www.ncbi.nlm.nih.gov/pubmed/20378774. 11

Ole Seehausen, Roger K Butlin, Irene Keller, Catherine E Wagner, Janette W Boughman, Paul A Hohenlohe, Catherine L Peichel, Glenn-Peter Saetre, Claudia Bank, Åke Brännström, et al. Genomics and the origin of species. *Nature Reviews Genetics*, 15(3):176–192, 2014. 12

Maria R Servedio, Yaniv Brandvain, Sumit Dhole, Courtney L Fitzpatrick, Emma E Goldberg, Caitlin A Stern, Jeremy Van Cleve, and D Justin Yeh. Not just a theory – the utility of mathematical models in evolutionary biology. *PLoS Biology*, 12(12):e1002017, 2014. 1

Mark L Siegal and Aviv Bergman. Waddington's canalization revisited: developmental stability and evolution. *Proceedings of the National Academy of Sciences*, 99(16):10528–10532, 2002. 1

Alexis Simon, Nicolas Bierne, and John J. Welch. Coadapted genomes and selection on hybrids: Fisher's geometric model explains a variety of empirical patterns. *bioRxiv*, 2017. doi: 10.1101/237925. URL https://www.biorxiv.org/content/early/2017/12/27/237925. 14

Montgomery Slatkin and Russell Lande. Segregation variance after hybridization of isolated populations. *Genetics Research*, 64(1):51–56, 1994. 2

TM Sonneborn. Degeneracy of the genetic code: extent, nature, and genetic implications. In *Evolving genes and proteins*, pages 377–397. Elsevier, 1965. 2

K. Stefflova, D. Thybert, M. D. Wilson, I. Streeter, J. Aleksic, P. Karagianni, A. Brazma, D. J. Adams, I. Talianidis, J. C. Marioni, P. Flicek, and D. T. Odom. Cooperativity and rapid evolution of cobound transcription factors in closely related mammals. *Cell*, 154(3):530–540, August 2013. 11

Rike Stelkens and Ole Seehausen. Genetic distance between species predicts novel trait expression in their hybrids. *Evolution: International Journal of Organic Evolution*, 63(4):884–897, 2009. 12

J A Sved. A two-sex polygenic model for the evolution of premating isolation. II. Computer simulation of experimental selection procedures. *Genetics*, 97(1):217–235, January 1981. URL https://www.ncbi.nlm.nih.gov/pubmed/17249074. 2

Nicholas F Trojanowski, Olivia Padovan-Merhar, David M Raizen, and Christopher Fang-Yen. Neural and genetic degeneracy underlies *Caenorhabditis elegans* feeding behavior. *Journal of neurophysiology*, 112(4): 951–961, 2014. 1

John R True and Eric S Haag. Developmental system drift and flexibility in evolutionary trajectories. *Evolution & Development*, 3(2):109–119, 2001. 1, 2, 11

Annie E Tsong, Brian B Tuch, Hao Li, and Alexander D Johnson. Evolution of alternative transcriptional circuits with identical logic. *Nature*, 443(7110):415–420, 2006. 1

Alexander Y Tulchinsky, Norman A Johnson, Ward B Watt, and Adam H Porter. Hybrid incompatibility arises in a sequence-based bioenergetic model of transcription factor binding. *Genetics*, 198(3):1155–1166, 2014. 2

Jenny Tung, Xiang Zhou, Susan C Alberts, Matthew Stephens, and Yoav Gilad. The genetic architecture of gene expression levels in wild baboons. *eLife*, 4:e04729, 2015. 11

AJ Van der Schaft. Equivalence of dynamical systems by bisimulation. *IEEE transactions on automatic control*, 49(12):2160–2172, 2004. 2

Dominique J Verlaan, Bing Ge, Elin Grundberg, Rose Hoberman, Kevin CL Lam, Vonda Koka, Joana Dias, Scott Gurd, Nicolas W Martin, Hans Mallmin, et al. Targeted screening of cis-regulatory variation in human haplotypes. *Genome research*, 19(1):118–127, 2009. 11

George von Dassow, Eli Meir, Edwin M. Munro, and Garrett M. Odell. The segment polarity network is a robust developmental module. *Nature*, 406(6792):188–192, 2000. ISSN 14764687. doi: 10.1038/35018085. URL https://doi.org/10.1038/35018085. 1

Andreas Wagner. Evolution of gene networks by gene duplications: a mathematical model and its implications on genome organization. *Proceedings of the National Academy of Sciences*, 91(10):4387–4391, 1994. 1, 3

Andreas Wagner. Does evolutionary plasticity evolve? *Evolution*, pages 1008–1023, 1996. 1

Andreas Wagner. Robustness and evolvability: a paradox resolved. *Proceedings of the Royal Society B: Biological Sciences*, 275(1630):91–100, 2008. 2

Eric Walter, Yves Lecourtier, and John Happel. On the structural output distinguishability of parametric models, and its relations with structural identifiability. *IEEE Transactions on Automatic Control*, 29(1): 56–57, 1984. 2

Mi Wang, Severin Uebbing, and Hans Ellegren. Bayesian inference of allele-specific gene expression indicates abundant cis-regulatory variation in natural flycatcher populations. *Genome Biology and Evolution*, 9(5): 1266–1279, 2017. 11

Kenneth M Weiss and Stephanie M Fullerton. Phenogenetic drift and the evolution of genotype–phenotype relationships. *Theoretical Population Biology*, 57(3):187–195, 2000. 1, 11

James M Whitacre. Degeneracy: a link between evolvability, robustness and complexity in biological systems. *Theoretical Biology and Medical Modelling*, 7(1):6, 2010. 2

Nathan J White, Rhonda R Snook, and Isobel Eyres. The past and future of experimental speciation. *Trends in Ecology & Evolution*, 35(1):10–21, 2020. 12

Karl R Wotton, Eva Jiménez-Guri, Anton Crombach, Hilde Janssens, Anna Alcaine-Colet, Steffen Lemke, Urs Schmidt-Ott, and Johannes Jaeger. Quantitative system drift compensates for altered maternal inputs to the gap gene network of the scuttle fly *Megaselia abdita*. *eLife*, 4:e04785, 2015. 1, 12

Sewall Wright. Evolution in populations in approximate equilibrium. *Journal of Genetics*, 30(2):257, 1935. URL http://link.springer.com/content/pdf/10.1007/BF02982240.pdf. 2

Ryo Yamaguchi and Yoh Iwasa. First passage time to allopatric speciation. *Interface focus*, 3(6):20130026, 2013. 2

Lotfi A Zadeh and Charles A Deoser. *Linear system theory*. Robert E. Krieger Publishing Company Huntington, 1976. 2

A Local expansion of the fitness surface

The fitness of a system depends on the difference between the system's impulse response and the optimal impulse response, measured as a weighted sum of the distance between the impulse response from optimal. With $\rho(t) \geq 0$ a weighting function on $[0, \infty)$, and $h_0(t) = C_0 e^{tA_0} B_0$ a representative of the optimal set, the distance is equation (5):

$$D(h) = \left(\int_0^\infty \rho(t) \|h(t) - h_0(t)\|^2 dt \right)^{1/2}. \tag{6}$$

To see how this is affected by small changes to the system, first note that if $\|\cdot\|_\rho$ is the $L^2(\rho)$ norm,

$$\|g\|_\rho^2 := \int_0^\infty \rho(t) \|g(t)\|^2 dt \tag{7}$$

then the distance to optimum of a perturbed system, $h(t) = h_0(t) + \epsilon g(t)$, is

$$D(h_0 + \epsilon g) = \epsilon \|g\|_\rho. \tag{8}$$

Now consider the perturbed systems $(A(\epsilon), B(\epsilon), C(\epsilon)) = (A(0), B(0), C(0)) + \epsilon(U, V, W)$, and let $h_\epsilon(t) = C(\epsilon) e^{A(\epsilon)t} B(\epsilon)$. If two populations with systems $(A(0), B(0), C(0))$ and $(A(\epsilon), B(\epsilon), C(\epsilon))$ interbreed, then the mean resulting system is $(A(\epsilon/2), B(\epsilon/2), C(\epsilon/2))$, but with greater variance in the F_2 than the F_1 offspring (see Figure 4). Indeed, if the difference between (say) $A(0)_{ij}$ and $A(\epsilon)_{ij}$ is due to n fixed differences with absolute additive effect size $\epsilon/\sqrt{n}$ each, then the standard deviation of A_{ij} among the F_1 offspring is still equal to that within each parent population, as each locus is heterozygous. However, if loci are unlinked then an F_2 offspring is homozygous with probability 1/2, and takes either homozygote with equal probability, so the variance of the contribution of each locus is $\epsilon/4\sqrt{n}$, and so the standard deviation across F_2 offspring is $\epsilon/4$ plus a contribution for within-population genetic variance. (This many differences of this size would be expected if the difference was due to drift, for instance.) This implies that the impulse response, $h(t)$, of an F_2 will also differ from optimum by something of order ϵ. However, since the displacement between the populations lies along a ridge, the F_1 offspring are closer: the Taylor expansions for h about both 0 and ϵ are

$$h_{\epsilon/2}(t) = h_0(t) + \frac{\epsilon}{2} \partial_\epsilon h_0(t) + \frac{\epsilon^2}{8} \partial_\epsilon^2 h_0(t) + O(\epsilon^3) \tag{9}$$

$$\text{and} \quad = h_\epsilon(t) - \frac{\epsilon}{2} \partial_\epsilon h_\epsilon(t) + \frac{\epsilon^2}{8} \partial_\epsilon^2 h_\epsilon(t) + O(\epsilon^3) \tag{10}$$

Since $h_\epsilon(t) = h_0(t)$ by assumption, and $\partial_\epsilon h_\epsilon(t) = \partial_\epsilon h_0(t) + \epsilon \partial_\epsilon^2 h_0(t) + O(\epsilon^2)$, combining these expressions we get that

$$h_{\epsilon/2}(t) = h_0(t) - \frac{\epsilon^2}{8} \partial_\epsilon^2 h_0(t) + O(\epsilon^3). \tag{11}$$

In other words, if the parental populations are both optimal but differ by ϵ, then the F_1 offspring differ from optimal by something of order ϵ^2, but the F_2 by something of order ϵ.

Of course, when moving away from the ridge, the fitness landscape is flatter in some directions than others, and how much so depends on the details. For completeness, we describe here how to compute the sensitivity of distance to optimum with respect to each of A, B, and C. Changing notation slightly, let's rewrite the squared distance to optimum of a system (A, B, C):

$$\begin{aligned} D(A,B,C)^2 &:= \int_0^\infty \rho(t)\left|h_A(t) - h_0(t)\right|^2 dt \\ &= \int_0^\infty \rho(t)\left|Ce^{At}B - C_0 e^{A_0 t}B_0\right|^2 dt \\ &= \int_0^\infty \rho(t)\,\mathrm{tr}\left\{\left(Ce^{At}B - C_0 e^{A_0 t}B_0\right)^T\left(Ce^{At}B - C_0 e^{A_0 t}B_0\right)\right\} dt \\ &= \int_0^\infty \rho(t)\,\mathrm{tr}\left\{\left(Ce^{At}B - C_0 e^{A_0 t}B_0\right)\left(Ce^{At}B - C_0 e^{A_0 t}B_0\right)^T\right\} dt, \end{aligned} \tag{12}$$

where $\mathrm{tr}\, X$ denotes the trace of a square matrix X. How does this change as we perturb about (A_0, B_0, C_0)? First we differentiate with respect to A, keeping $B = B_0$ and $C = C_0$ fixed. Since

$$\frac{d}{du} e^{(A+uZ)t}|_{u=0} = \int_0^t e^{As} Z e^{A(t-s)} ds, \tag{13}$$

we have that

$$\begin{aligned} \frac{d}{du} D(A+uZ, B_0, C_0)^2|_{u=0} &= 2\int_0^\infty \rho(t)\,\mathrm{tr}\left\{C_0\left(\int_0^t e^{As} Z e^{A(t-s)} ds\right) B_0 B_0^T \left(e^{At} - e^{A_0 t}\right)^T C_0^T\right\} dt \\ &= 2\int_0^\infty \rho(t)\,\mathrm{tr}\left\{C_0\left(\int_0^t e^{As} Z e^{A(t-s)} ds\right) B_0 \left(h_A(t) - h_0(t)\right)^T\right\} dt \end{aligned} \tag{14}$$

and, by differentiating this and supposing that A is on the optimal set, i.e., $h_A(t) = h_0(t)$, (so without loss of generality, $A = A_0$):

$$\begin{aligned} \mathcal{H}^{A,A}(Y,Z) &:= \frac{1}{2}\frac{d}{du}\frac{d}{dv} D(A_0 + uY + vZ, B_0, C_0)^2|_{u=v=0} \\ &= \int_0^\infty \rho(t)\,\mathrm{tr}\left\{C_0\left(\int_0^t e^{A_0 s} Y e^{A_0(t-s)} ds\right) B_0 B_0^T \left(\int_0^t e^{A_0 s} Z e^{A_0(t-s)} ds\right)^T C_0^T\right\} dt. \end{aligned} \tag{15}$$

The function $\mathcal{H}$ will define a quadratic form. To illustrate the use of this, suppose that B and C are fixed. By defining Δ_{ij} to be the matrix with a 1 in the (i,j)th slot and 0 elsewhere, the coefficients of the quadratic form are

$$H_{ij,k\ell}(A) := \mathcal{H}(\Delta_{ij}, \Delta_{k\ell}). \tag{16}$$

We could use this to get the quadratic approximation to D near the optimal set. To do so, it'd be nice to have a way to compute the inner integral above. Suppose that we diagonalize $A = U\Lambda U^{-1}$. Then

$$\int_0^t e^{As} Z e^{A(t-s)} ds = \int_0^t U e^{\Lambda s} U^{-1} Z U e^{\Lambda(t-s)} U^{-1} ds \tag{17}$$

Now, notice that

$$\int_0^t e^{s\lambda_i} e^{(t-s)\lambda_j} ds = \begin{cases} \frac{e^{t\lambda_i} - e^{t\lambda_j}}{\lambda_i - \lambda_j} & \text{if} \quad i \neq j \\ te^{t\lambda_i} & \text{if} \quad i = j. \end{cases} \tag{18}$$

Therefore, defining

$$X_{ij}(t, Z) = \begin{cases} \left(U^{-1}ZU\right)_{ij} \frac{e^{t\lambda_i} - e^{t\lambda_j}}{\lambda_i - \lambda_j} & \text{if} \quad i \neq j \\ \left(U^{-1}ZU\right)_{ii} te^{t\lambda_i} & \text{if} \quad i = j, \end{cases} \tag{19}$$

moving the U and U^{-1} outside the integral and integrating we get that

$$\int_0^t e^{As} Z e^{A(t-s)} ds = UX(t,Z)U^{-1}. \tag{20}$$

This implies that

$$D(A_0+\epsilon Z)^2 \approx \frac{1}{2}\epsilon^2 \int_0^\infty \rho(t)\,\mathrm{tr}\left\{CUX(t,Z)U^{-1}BB^T(U^{-1})^T X(t,Z)^T U^T C^T\right\} dt. \tag{21}$$

To compute the $n^2 \times n^2$ matrix H, we see that if $Z = \Delta_{k\ell}$, then

$$X_{ij}^{k\ell}(t) = \begin{cases} (U^{-1})_{\cdot k} U_{\ell\cdot} \frac{e^{t\lambda_i} - e^{t\lambda_j}}{\lambda_i - \lambda_j} & \text{if} \quad i \neq j \\ (U^{-1})_{\cdot k} U_{\ell\cdot} t e^{t\lambda_i} & \text{if} \quad i = j, \end{cases} \tag{22}$$

where $U_{k\cdot}$ is the kth row of U, and so

$$H_{ij,k\ell}(A) = \int_0^\infty \rho(t)\,\mathrm{tr}\left\{CUX^{ij}(t)U^{-1}BB^T(U^{-1})^T X^{k\ell}(t)^T U^T C^T\right\} dt. \tag{23}$$

This implies that

$$D(A_0+\epsilon Z)^2 \approx \frac{1}{2}\epsilon^2 \sum_{ijk\ell} H_{ij,k\ell}(A_0) Z_{ij} Z_{k\ell}. \tag{24}$$

More generally, B and C may also change. To extend this we need the remaining second derivatives of D^2. First, in B:

$$\begin{aligned} \mathcal{H}^{B,B}(Y,Z) &:= \frac{1}{2}\frac{d}{du}\frac{d}{dv} D(A_0, B_0 + uY + vZ, C_0)|_{u=v=0} \\ &= \frac{1}{2}\int_0^\infty \rho(t)\,\mathrm{tr}\left\{C_0 e^{tA_0} \frac{d}{du}\frac{d}{dv}(uY+vZ)(uY+vZ)^T|_{u=v=0} e^{tA_0^T} C_0^T\right\} dt \\ &= \frac{1}{2}\int_0^\infty \rho(t)\,\mathrm{tr}\left\{C_0 e^{tA_0}\left(YZ^T + ZY^T\right) e^{tA_0^T} C_0^T\right\} dt. \end{aligned} \tag{25}$$

Next, in C:

$$\begin{aligned} \mathcal{H}^{B,B}(Y,Z) &:= \frac{1}{2}\frac{d}{du}\frac{d}{dv} D(A_0, B_0, C_0 + uY + vZ)|_{u=v=0} \\ &= \frac{1}{2}\int_0^\infty \rho(t)\,\mathrm{tr}\left\{B_0 e^{tA_0^T} \frac{d}{du}\frac{d}{dv}(uY+vZ)^T(uY+vZ)|_{u=v=0} e^{tA_0} B_0\right\} dt \\ &= \frac{1}{2}\int_0^\infty \rho(t)\,\mathrm{tr}\left\{B_0 e^{tA_0^T}\left(YZ^T + ZY^T\right) e^{tA_0} B_0\right\} dt. \end{aligned} \tag{26}$$

Now, the mixed derivatives in B and C:

$$\begin{aligned} \mathcal{H}^{B,C}(Y,Z) &:= \frac{1}{2}\frac{d}{du}\frac{d}{dv} D(A_0, B_0 + uY, C_0 + vZ)|_{u=v=0} \\ &= \int_0^\infty \rho(t)\,\mathrm{tr}\left\{Y e^{tA_0^T} C_0^T Z e^{tA_0} B_0\right\} dt. \end{aligned} \tag{27}$$

In A and B

$$\begin{aligned} \mathcal{H}^{A,B}(Y,Z) &:= \frac{1}{2}\frac{d}{du}\frac{d}{dv} D(A_0 + uY, B_0 + vZ, C_0)|_{u=v=0} \\ &= \int_0^\infty \rho(t)\,\mathrm{tr}\left\{C_0\left(\int_0^t e^{sA_0} Y e^{(t-s)A_0} ds\right) B_0 Z^T e^{tA_0} C_0\right\} dt, \end{aligned} \tag{28}$$

and finally in A and C:

$$\begin{aligned}\mathcal{H}^{A,C}(Y,Z) &:= \frac{1}{2}\frac{d}{du}\frac{d}{dv}D(A_0+uY,B_0,C_0+vZ)|_{u=v=0} \\ &= \int_0^\infty \rho(t)\,\mathrm{tr}\left\{C_0\left(\int_0^t e^{sA_0}Ye^{(t-s)A_0}ds\right)B_0B_0e^{tA_0}Z\right\}dt.\end{aligned} \tag{29}$$

Together, numerical computation of these expressions, along with estimates of genetic covariance within a population, allow precise predictions of evolutionary dynamics of a particular system. The approximation should be good as long as the second-order Taylor approximation holds.