Rapid speciation despite conservation of phenotype

Joshua S. Schiffman[†] Peter L. Ralph[†‡]
[†]University of Southern California, Los Angeles, California [‡]University of Oregon, Eugene, Oregon
jsschiff@usc.edu plr@uoregon.edu

Abstract

We introduce an analytical theory to study the evolution of biological systems, such as gene regulatory networks. The evolutionary conservation of phenotype under selective and environmental stasis does not necessitate conservation of the underlying mechanism, as distinct molecular pathways can realize identical phenotypes. Here we give an exact expression for the set of all linear mechanisms with identical phenotypes, and expect evolution under neutrality to explore this set. We employ a quantitative genetic approach to model evolution under neutrality as a random process over the set of all phenotype-invariant mechanisms: only mutational tweaks to the pathway that leave the phenotype invariant are optimally fit. We show that there is never a unique linear system architecture for any phenotype and that the evolutionary exploration of these distinct and mutationally connected mechanisms can lead to the rapid accumulation of hybrid incompatibilities between allopatric populations and thus lead to the rapid formation of new species – in fewer generations than there are breeding individuals in a population.

Additional ideas to consider adding:

- discuss linearity, linearization, and canalization in introduction

Introduction

Bridging the gulf between an organism's genome and phenotype is a poorly understood and complex molecular machinery. Progress in a suite of biological subdisciplines is stalled by our general lack of understanding of this molecular machinery: with respect to both its function and evolution. There does exist 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, Pavlicev and Wagner, 2012, Weiss and Fullerton, 2000, Edelman and Gally, 2001]. However, verbal theories are often insufficient, if not downright misleading [Servedio et al., 2014]. This is especially pertinent given the staggering complexity and scope of contemporary research programs. This outlook necessitates the advancement of conceptual frameworks of such precision, only mathematics will suffice, as models allow the development of concrete numerical predictions.

probably some more recent Wagner papers in this line as well?

The molecular machinery, interacting with the environment, and bridging genotype to phenotype can be mathematically described as a dynamical system – or a system of differential equations [Jaeger et al., 2015]. Movement in this direction is ongoing, as researchers have begun to study the evolution of both abstract [Wagner, 1994, 1996, Siegal and Bergman, 2002, Bergman and Siegal, 2003, Draghi and Whitlock, 2015] and empirically inspired computational and mathematical models of gene regulatory networks [Mjolsness et al., 1991, Jaeger et al., 2004, Kozlov et al., 2012, 2015, 2014, Crombach et al., 2016, Wotton et al., 2015, Chertkova et al., 2017]. If we allow the reasonable assumption that the genotype-phenotype map can be represented as a system of differential equations, we can immediately discuss its evolution and function in a much more mechanistic, yet general, manner.

In some fields that seek to fit parametric models to experimental data, such as control theory, chemical engineering, and statistics, it is well known that mathematical models can

fundamentally be *unidentifiable* and/or *indistinguishable* – meaning that there can be uncertainty about an inferred model's parameters or even its claims about causal structure, even with access to complete and perfect data [Bellman and Åström, 1970, Grewal and Glover, 1976, Walter et al., 1984]. Models with different parameter schemes, or even different mechanics can be equally accurate, but still not *actually* agree with what is being modelled. 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 on the black box are performed [Kalman, 1963, Anderson et al., 1966, Zadeh and Deoser, 1976]. In chemical engineering, those who study chemical reaction networks sometimes refer to the fundamental unidentifiability of these networks as "the fundamental dogma of chemical kinetics" [Craciun and Pantea, 2008]. In computer science, this is framed as the relationship among processes that simulate one another [Van der Schaft, 2004]. Although this may frustrate the occasional engineer or scientist, viewed from another angle, the concepts of unidentifiability and indistinguishability can provide a starting point for thinking about externally equivalent systems – systems that evolution can explore, so long as the parameters and structures can be realized biologically. In fact, evolutionary biologists who study convergent versus parallel evolution, and developmental homology and analogy, are very familiar with such functional symmetries; macroscopically identical phenotypes in even very closely related species can in fact be divergent at the molecular and sequence level [True and Haag, 2001, Tsong et al., 2006, Hare et al., 2008, **?**, Stergachis et al., 2014, Taylor et al., 2016, Matsui et al., 2015].

In this paper we outline a theoretical framework to study the evolution of biological systems, such as gene regulatory networks. We focus primarily on the neutral scenario: that is where phenotype is conserved over evolutionary time. We present an analytical description of the set of all linear biological systems with identical phenotypes – that is we describe the set of all gene network architectures that yield identical phenotypes in equivalent environments, and show that all linear biological systems can, in principal, can undergo systems drift and rewire. Under neutrality, a population will drift through the set of all possible phenotypically equivalent linear gene networks. Consequentially, two phenotypically equivlanet yet allopatrically isolated populations, will likely produce incompatibile hybrids, despite the absence of adaptation, directional selection, or environmental change. Under these conditions, speciation typically occurs on timescales approximately on the order of N_e generations, where N_e is the effective population size.

Gene networks as linear dynamical systems

Systems theory Organisms' phenotypes are constructred by gene by gene and gene by environment interactions. Here we simply define the *phenotype* to be the temporal molecular dynamics directly under natural selection. The *what*, *when*, and *how much*, of an organism's molecules that are physiologically or otherwise relevant to survival.

Thus an organism's phenotype $\phi(t)$ – a vector of molecular concentrations at time t – is determined both by the structure and organization of a biological system (*e.g.* a gene regualtory network), given by the triple (A, B, C), and by its environment $u(t)$.

Such a biological system $\mathcal{S}$ is given by,

$$\mathcal{S} := \begin{cases} \dot{\kappa}(t) & = A\kappa(t) + Bu(t) \\ \phi(t) & = C\kappa(t) \end{cases} \tag{1}$$

Generally A is any real $n \times n$ matrix, B any $n \times \ell$, and C any $\ell \times n$ dimensional matrix. Although many different biological systems can be modeled with this approach, for clarity, we focus on gene regulatory networks. Each ith row of A describes the *cis*-regulatory module for gene i, and each jth entry, the specific regulatory influence of gene j on gene i. As such,

A_{ij} is the magnitude at which transcription factor $\kappa_j(t)$ regulates transcription factor $\kappa_i(t)$, and if $A_{ij} > 0$, we say that κ_j upregulates κ_i. If $A_{ij} < 0$, we say that κ_j down-regulates κ_i.

The form of B determines precisely how the environment influences the organism – that is B filters and translates the input to the system. C filters and translates the dynamics of the system and precesily determines the output, that is, what is visible to selection. *E.g.* for a metabolic system, C_{ij} is the amount the jth metabolite u_j affects the production of the ith enzyme.

Furthermore, whereas the phenotype is a subset of molecules *visible* to selection, the *kryptotype* includes the molecular dynamics *hidden* from selection. That is $\kappa(t)$ is a vector of the system's molecular concentrations at time t.

This "Thus" needs expanding, e.g., "the solution to this equation is unique and given by" with a reference.

Finally, we can write the phenotype as a convolution of the system organization and the environment,

$$\phi(t) = Ce^{At}\kappa(0) + \int_0^t Ce^{A(t-s)}Bu(s)ds, \tag{2}$$

where we refer to $h(t) := Ce^{At}B$ as the system's *impulse response*.

Example 1 (Oscillating gene network: cell cycle control)**.** *Cellular division is governed by many different processes, however it is thought that its rhythm is partially controlled by oscillating gene transcription [Orlando et al., 2008]. Here we consider a simplified model of oscillating gene transcription.*

Suppose gene-2 up-regulates the transcription of gene-1 and that gene-1 down-regulates gene-2 with equal magnitudes, whose concentrations are given by $\kappa_1(t)$ and $\kappa_2(t)$. Furthermore, suppose that only the dynamics of gene-1 are consequential to the cell cycle (perhaps the amount of gene-1 activates another downstream gene network). Lastly suppose that the production of both genes is equally stimulated by an impulse of a molecule present immediately after division.

If the rate each of these genes is expressed is a linear function of their concentrations, the dynamics of the system is given by

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

Equivalently, in matrix form the system regulatory coefficients are given as, $A = \left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right]$, $B = \left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]$, $C = \left[\begin{smallmatrix} 1 & 0 \end{smallmatrix}\right]$

Since the input is simply an impulse, its phenotype is equivalent to its impulse response,

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

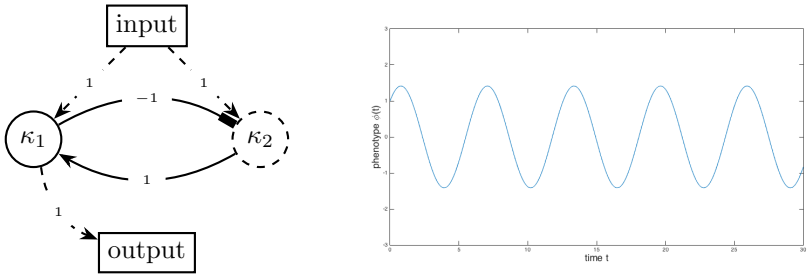

Figure 1: (Left) Graphical representation of the cell cycle control gene network, and (right) plot of the phenotype $\phi(t)$ against time t.

We return to the evolution of such a system below.

Phenotypically equivalent gene networks Despite a symmetry in functionality or phenotype systems can often differ, sometimes substantially, at the molecular level. How many different mechanisms have the same function? Gene regulatory networks with identical phenotypes do not necessarily have identical kryptotypes.

Definition 1 (Phenotypic equivalence of systems)**.** *Let* $(\kappa(t), \phi(t))$ *and* $(\bar{\kappa}(t), \bar{\phi}(t))$ *be the solutions to* (1) *with coefficient matrices* (A, B, C) *and* $(\bar{A}, \bar{B}, \bar{C})$ *respectively, and both* $\kappa(0)$ *and* $\bar{\kappa}(0)$ *are zero. The systems defined by* (A, B, C) *and* $(\bar{A}, \bar{B}, \bar{C})$ *are* ***phenotypically equivalent*** *if*

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

Equivalently, this occurs if and only if

$$h(t) = \bar{h}(t) \qquad \textit{for all } t \geq 0,$$

where h *and* $\bar{h}$ *are the impulse responses of the two systems.*

One way to find other systems 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}) have the same dynamics because their impulse responses are equal:

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

However, the converse is not necessarily true: systems can have identical impulse responses without being changes of coordinates of each other. In fact, systems with identical impulse responses can involve interactions between different numbers of molecular species, and thus be in different dimensions.

We refer to $\mathcal{A}(\mathcal{S}_0)$ as the set of all systems equivalent to $\mathcal{S}_0$, regardless of dimension:

$$\begin{aligned} \mathcal{A}(\mathcal{S}_0) &= \left\{(A, B, C) : Ce^{At}B = C_0e^{A_0t}B_0 \text{ for } t \geq 0\right\} \\ &= \left\{(A, B, C) : CA^rB = C_0A_0^rB_0 \text{ for } 1 \leq r \leq n-1\right\}. \end{aligned} \tag{4}$$

Equivalence of the two characterizations follows from the Cayley-Hamilton theorem.

The set of all systems phenotypically equivalent to a given system (A, B, C) is elegantly described using the Kalman decomposition, which also clarifies the system dynamics. To motivate this, first note that the input $u(t)$ only directly pushes the system in directions 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 the *reachable subspace*, which we denote by $\mathcal{R}$, and is defined to be the closure of span(B) under applying A (or equivalently, the span of $B, AB, A^2B, \ldots A^{n-1}B$). Analogously to this, we define the *observable subspace*, $\mathcal{O}$, to be the closure of span(C^T) under applying A. (Or: $\bar{\mathcal{O}}$ is the largest A-invariant subspace contained in the null space of C; and $\mathcal{R}$ is the largest A-invariant subspace contained in the image of B.)

If we define

1. The columns of $P_{r\bar{o}}$ are an orthonormal basis for $\mathcal{R} \cap \bar{\mathcal{O}}$.
2. The columns of P_{ro} are an orthonormal basis of the complement of $\mathcal{R} \cap \bar{\mathcal{O}}$ in $\mathcal{R}$.
3. The columns of $P_{\bar{r}o}$ are an orthonormal basis of the complement of $\mathcal{R} \cap \bar{\mathcal{O}}$ in $\bar{\mathcal{O}}$.

4. The columns of $P_{\bar{r}\bar{o}}$ are an orthonormal basis of the remainder of $\mathbb{R}^n$.

If we then define

$$P = \left[\begin{array}{c|c|c|c} P_{r\bar{o}} & P_{ro} & P_{\bar{r}o} & P_{\bar{r}\bar{o}} \end{array}\right],$$

then

$$P^T P = \left[\begin{array}{c|c|c|c} I & 0 & 0 & 0 \\ \hline 0 & I & U & 0 \\ \hline 0 & V & I & 0 \\ \hline 0 & 0 & 0 & I \end{array}\right].$$

Check this. Can we get $U = V = 0$?

The following theorem can be found in SOME REFERENCE.

Theorem 1 (Kalman decomposition)**.** *For any system (A, B, C) with corresponding Kalman basis matrix P, the transformed system (PAP^{-1}, PB, CP^{-1}) has the following form:*

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

and

$$\widehat{B} = PB = \begin{bmatrix} B_{r\bar{o}} \\ B_{ro} \\ 0 \\ 0 \end{bmatrix},$$

and

$$\widehat{C} = CP^{-1} = \begin{bmatrix} 0 & C_{ro} & C_{\bar{r}\bar{o}} & 0 \end{bmatrix}.$$

The impulse response of both systems is given by

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

In the latter case, we say that the system is *minimal* – there is no equivalent system with a smaller number of molecular species. Note that this says that any two equivalent minimal systems are changes of coordinates of each other.

Since any system can be put into this form, and once in this form, its impulse response is determined only by C_{ro}, A_{ro}, and B_{ro}, therefore, the set of all equivalent systems are parameterized by the dimension n, the choice of basis (P), the remaining submatrices in $\widehat{A}$, $\widehat{B}$, and $\widehat{C}$ (which are unconstrained), and a invertible transformation of span(P_{ro}), which we call T_{ro}.

Theorem 2 (Parameterization of equivalent systems)**.** *Let (A, B, C) be a minimal system.*

(a) Every equivalent system is of the form given in Theorem 3, i.e., can be specified by choosing a dimension, n; submatrices in $\widehat{A}$, $\widehat{B}$, and $\widehat{C}$ except for $A_{ro} = A$, $B_{ro} = B$, and $C_{ro} = C$; and choosing an invertible matrix P.

(b) The parameterization is unique if P is furthermore chosen so that each P_x other than P_{ro} is a projection matrix, and that

$$0 = P_x^T P_y$$

for all (x, y) except $(ro, \bar{r}\bar{o})$.

conjecture:

It is remarkable to note that even with the relationships between environment, kryptotype, and phenotype constant, and in the minimal dimension, there are still almost always degrees of freedom. These correspond to distinct genetic networks that perform indistinguishable functions. For example, the equivalent systems in Example 2 are minimal, and share common B and C matrices.

It is also remarkable to note that there is no unique triple (A, B, C) for the mapping $h : u \mapsto \phi$. This implies that the gene regulatory network architecture per phenotype/environment pair is never unique – that is under these assumptions, there is always more than one possible gene regulatory network architecture per phenotype.

For the remainder of the paper, we interpret $\mathcal{A}$ as the phenotypially neutral landscape, wherein a large population will drift under environmental and selective stasis. 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.

Example 2 (All Phenotypically Equivalent Cell Cycle Control Networks)**.** *The set of all two-gene regulatory networks phenotypically equivalent to the cell cycle control network in 1, where only A can vary, is given by* $\mathcal{S}(\tau) = (A(\tau), B, C)$*, such that,*

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

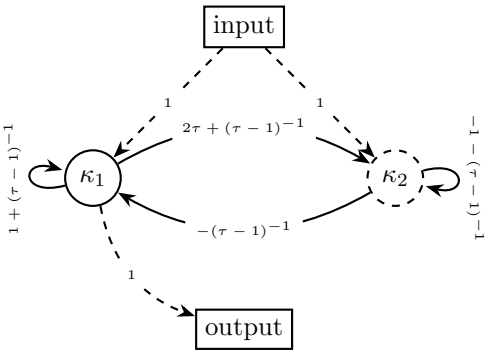

Figure 2: $\mathcal{S}(\tau)$, the set of all phenotype-equivalent cell cycle control networks $\mathcal{S}(\tau) = \{\mathcal{A}_2 : B = [1\ 1]^T,\ C = [1\ 0]\}$.

Despite the phenotypic equivalence of all instantiations of $\mathcal{S}(\tau)$*, the kryptotypes, vary as a function of* τ*.*

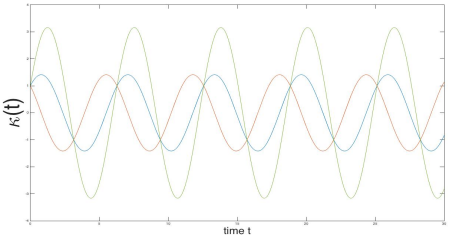

Figure 3: Gene-1 dynamics (blue) for both systems $\mathcal{S}(0)$ and $\mathcal{S}(2)$ are identical, however, $\mathcal{S}(0)$ gene-2 dynamics (orange) differ from $\mathcal{S}(2)$ (green). Both gene-1 dynamics are given by $\kappa_1 = \sin t + \cos t$, and gene-2 by $\kappa_2 = \cos t - \sin t$ $(\mathcal{S}(0))$ and $\kappa_2' = \cos t + 3\sin t$ $(\mathcal{S}(2))$.

Example 3 (Metabolic network)**.** *Consider an organism that can metabolize two different sugars (present at logarithmic concentrations u_1 and u_2), with two enzymes (at log concentrations ϕ_1 and ϕ_2). Further suppose that the second sugar is prefered, and that depending on u_1 and u_2 the organism will synthesize an appropriate ϕ_1 and ϕ_2. Furthermore, assume this system contains at least two transcription factors, whose log concentrations are $\kappa_1, \kappa_2, \ldots \kappa_n$. Minimally such a system may have the architecture, $\mathcal{S}_{\min}(U) = (U \left[\begin{smallmatrix} 0 & -1 \\ 0 & 0 \end{smallmatrix}\right] U^{-1}, U, U^{-1})$. We can find alternative equivalent systems by changning coordinates ($U \to U'$) or, more generally, by applying the Kalman decomposition A. To illustrate that phenotypic invariance does not require dimensional invariance, we apply the Kalman decomposition to $\mathcal{S}_{\min}$ to find $\mathcal{S}(D_{1-3}, V) = (\ V \left[\begin{smallmatrix} D_1 & D_2 \\ 0 & A \end{smallmatrix}\right] V^{-1}, V \left[\begin{smallmatrix} D_3 \\ B \end{smallmatrix}\right], [\,0\ C\,] V^{-1}\ )$, both of which are in $\mathcal{A}$ (V can be any 4-dimensional and invertible matrix, $D_{1-3} \in \mathbb{R}^{2\times 2}$, and $(A, B, C) \in \mathcal{S}_{\min}$).*

Sexual recombination and hybrid incompatibility Phenotypically identical yet kryptotypically variant gene networks, when brought together and/or recombined during sexual reproduction, can lead to phenotypic inviability.

For instance, the diploid $\{\mathcal{S}, \mathcal{S}'\}$ has system architecture $\mathcal{S}_{\text{diploid}} = (\mathcal{S} + \mathcal{S}')/2$, the average of its two haploid systems.

During meiosis, a diploid recombines its two genomes to form a gamete by randomly swapping rows or columns between its two systems. *E.g.* gamete systems (A'', B'', C''), produced by the diploid $\{\mathcal{S}, \mathcal{S}'\}$, are formed by recombining two systems $\mathcal{S} = (A, B, C)$ and $\mathcal{S}' = (A', B', C')$ to form $\mathcal{S}''$ such that, $A'' = MA + (I - M)A'$, $B'' = MB + (I - M)B'$, and $C'' = CM + C'(I - M)$. M is a diagonal matrix where each diagonal element is a Bernoulli random variable ($M_{ii} = 0$ or 1 with equal probability, and $M_{ij} = 0$ if $i \neq j$). If systems are different dimensions the smaller system elements can be augmented with 0s (*e.g.* $\left[\begin{smallmatrix} A & 0 \\ 0 & 0 \end{smallmatrix}\right], \left[\begin{smallmatrix} B \\ 0 \end{smallmatrix}\right], [\,C\ 0\,]$).

Even if two systems yield the same phenotype, their diploid and recombinant systems may not. This is because $\mathcal{A}$ is not required to be closed under averaging or recombination.

If sexual recombination among systems drawn from $\mathcal{A}$ yields systems with divergent phenotypes, populations containing significant diversity in $\mathcal{A}$ can carry genetic load, and isolated populations may fail to produce hybrids with viable phenotypes (Figure 4).

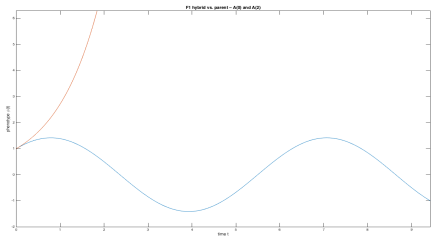

Figure 4: **Hybrid phenotypic breakdown** F_1 hybrid (orange) and parental (blue) phenotypic cell-cycle control (oscillator) dynamics for an $\mathcal{S}(0)$ by $\mathcal{S}(2)$ cross from Example 2. The F_1 hybrid gene expression fails to oscillate and instead grows exponentialy ($\phi_{F_0} = \sin t + \cos t$, however $\phi_{F_1} = e^t$).

Example 4 (Hybrid Incompatibility in an Oscillating Gene Network)**.** *Here we compare the phenotypes for F_2 hybrids formed by crossing oscillators with different values for τ from 2, $\mathcal{S}(2)$ with $\mathcal{S}(2.01)$, $\mathcal{S}(2.1)$, and $\mathcal{S}(2.5)$. Systems have identical phenotypes, however some of the hybrids exhibit markedly different dynamics.*

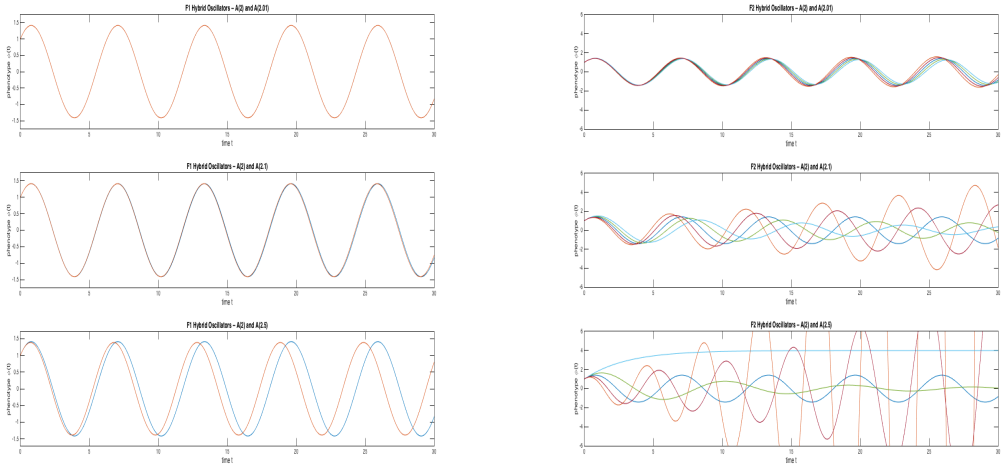

Figure 5: F_1 (left) and F_2 (right) hybrids crossing $\mathcal{S}(2)$ with $\mathcal{S}(2.01)$ (top), $\mathcal{S}(2.1)$ (middle), and $\mathcal{S}(2.5)$ (bottom). F_2 hybrids display more phenotypic divergence than F_1s, on average. Further, some F_2s completely fail to oscillate, as seen in an $\mathcal{S}(2.5)$ F_2 (light blue). *make axis labels bigger*

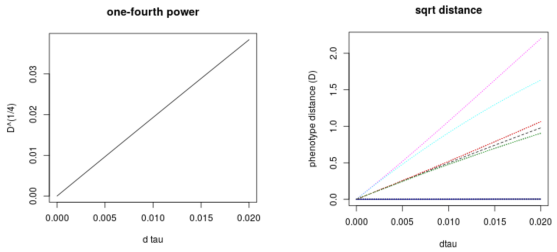

Figure 6: (Left) Squared F_1 hybrid phenotypes from Figure **??** diverge quartically, and (right) F_2 hybrid phenotypes diverge quadratically as a function of time in allopatry.

Systems drift and the accumulation of incompatibilities

Thus far we have shown that many distinct molecular mechanisms can realize identical phenotypes and that these mechanisms may fail to produce viable hybrids. Here we discuss the rate at which evolution shifts molecular mechanisms and ask whether this process is fast enough to be a significant driver of speciation. To do so, we explore a general quantitative genetic model in which a population drifts stochastically near a set of equivalent and optimal systems. We work with a population of effective size N_e.

At any given time, there will be a range of regulatory coefficients present in the population due to segregating genetic variants. Over many generations, even if selective pressures do not change, this range of networks will shift as recombination, mutation, and demographic noise create new alleles and shift allele frequencies.

Suppose that a set x of coefficients that determine a system (this is $\mathcal{S}$ above), produce a phenotype $\Phi(x)$ (the time course of $\phi(t)$). There is an optimal phenotype Φ_0, and a set $\mathcal{X}$ of "optimal" coefficients that produce this phenotype. Fitness depends on distance to the

optimal phenotype – we will write the "distance" between phenotypes ϕ and ψ as $d(\phi, \psi)$, measured so that the fitness of an organism with coefficents x is $\mathcal{F}(x) = \exp(-d(\Phi(x), \Phi_0)^2)$. We will assume that the map Φ is smooth and that the optimal set $\mathcal{X}$ is locally isomorphic to $\mathbb{R}^m$.

say that better

System drift If the regulatory coefficient population variation has standard deviation σ, since subsequent generations resample from this diversity, the population mean coefficient will move a random distance of size $\sigma/\sqrt{N_e}$ per generation, simply because this is the the standard deviation of the mean of a random sample [**?**]. Selection will tend to restrain this motion, but mean movement along the optimal set $\mathcal{X}$ is unconstrained. The amount of variance in particular directions in coefficient space depend on constraints imposed by selection and the covariance between genetic variation between different coefficients (the G matrix [**?**]).

We denote σ_N and σ_S, the standard deviations of coefficient variation along and perpendicular to $\mathcal{X}$ respectively, and γ, the scale on which phenotype changes moving away from $\mathcal{X}$. Concretely, γ is the inverse of the derivative of $d(\Phi(x + uz), \Phi(x))$ with respect to u for $x \in \mathcal{X}$ and z perpendicular to the tangent space at x. With these parameters, a typical individual will have a fitness of around $\exp(-(\sigma_S/\gamma)^2)$.

Hybridization The means of two allopatric populations separated for T generations will be a distance of order $2\sigma_N\sqrt{T/N_e}$ apart along $\mathcal{X}$. A population of F_1 hybrids has one haploid genome from each, whose coefficients are averaged, and so will have mean system coefficients at the midpoint between their means, and variance equal to σ^2. Each F_2 hybrid will be homozygous for one parental allele on averge at half of the loci in the genome, so the distribution of F_2s will have mean at the average of the two populations, as before, but variance equal to $\sigma^2 + z^2/2$, where z is the distance between the parental populations. These are depicted in figure 7.

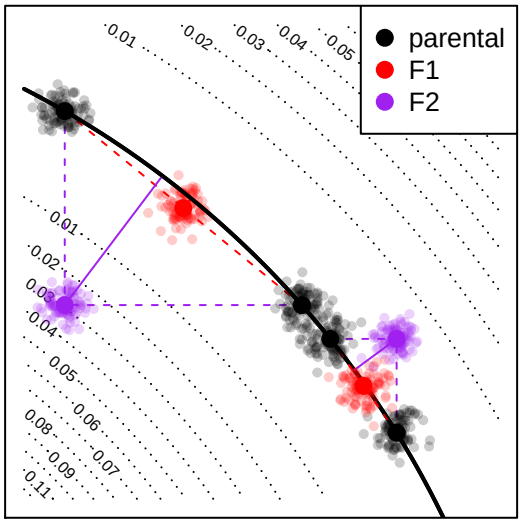

Figure 7: A conceptual figure of the fitness consequences of hybridization: axes represent system coefficients (i.e., entries of A); the line of optimal system coefficients is down in black; dotted lines give phenotypic distances to the optimum. Two pairs of parental populations are shown in black, along the optimum; a hypothetical population of F_1s are shown for each in red, and the distribution of one type of F_2 is shown in purple (other types of F_2 are not shown). Solid lines depict the distance of the F_2 to optimum. *Should show all types of F2? would be messy.*

improve figure by putting labels on from the following Suppose that two populations have drifted independently to differ by z, and that z is of the same order as σ but is smaller than γ. The mean F_1 is the average of the parental means, and since the first-order terms in the Taylor series vanish, has phenotype differing from the optimum by a distance of order $\|z\|^2$ (see appendix D). The mean F_2 is the same, but the standard deviation is of order z, so that up to lower order terms, while the typical fitness of an individual in the original population is $\mathcal{F}_0 = \exp(-(\sigma_s/\gamma)^2)$; of an F_1 is $\mathcal{F}_1/\mathcal{F}_0 = \exp(-(c_1\sigma_N^2 T/N_e)^2)$; and of an F_2 is $\mathcal{F}_2/\mathcal{F}_0 = \exp(-T/(N_e\gamma^2))$.

Note that this follows directly from the fact that F_1 and F_2 hybrid phenotypes diverge from F_0 phenotypes, linearly and inverse-quadratically, with respect to time (more precisely T/N_e). As such, removing assumptions about the form of the fitness function, this model still predicts hybrid F_2 phenotypes to diverge much faster than F_1s, initially.

Speciation rates under neutrality If we assume that perpendicular variation segregating within a population is constrained by selection such that $\sigma_S/\gamma = \xi$, and that neutral variation is on the order of perpendicular variation, such that $\sigma_N = \sigma_S$, then the typical distance between two individuals drawn from separate allopatric populations will be $z = 2\sigma_N\sqrt{T/N_e}$. We will expect a drop in hybrid fitness to be,

$$\mathcal{F}_2/\mathcal{F}_0 = \exp\left(-4\xi^2\frac{T}{N_e}\right). \tag{5}$$

Hybrid fitness depression will be significant when $z > \gamma$. If ξ is in $[1/5, 1]$, speciation will occur on the order of N_e generations, and possibly much faster. Note, however, that if intrapopulation variation is very small such that $\sigma_S \ll \gamma$, speciation rates will be slower.

Discussion

The complexity of biological systems has limited our understanding of their function and evolution. Above we outline an approach, a first step, towards untangling this complexity in reference to function and evolution. This methodology borrows successfully applied tools from engineering and aims to synthesize these with the concepts and tools of molecular and evolutionary biology.

Theoretical models in evolution and population genetics often lack the molecular details of physiology or of the genotype-phenotype map. Here, we offer a tractable and simple model which includes these missing features. Further, we provide, in clear mathematical language, an analytical description of phenomena hitherto discussed verbally and conceptually (phenogenetic drift [Weiss and Fullerton, 2000], developmental systems drift [True and Haag, 2001], biological degeneracy [Edelman and Gally, 2001], *etc.*). The tractability and relative simplicity of this exposition enables the interested biologist to work out by hand, if desired, the dynamics of a genetic system, as well as perturbations to the system – an attribute not likely to be found in less tractable models and simulations.

We have suggested an interpretation of system identification: to see it as an evolutionarily neutral manifold, and not simply a computational nuisance. We have demonstrated a method to analytically determine the set of all phenotypically invariant gene networks; by a simple change of coordinates in the minimal configuration, or more generally by applying the Kalman decomposition in higher dimensions. This set is explored over evolutionary time when phenotype is conserved, and can lead to a diverse set of consequences, including the accumulation of Dobzhansky-Muller incompatibilities. We emphasize that these incompatibilities are a consequence of recombining different, yet functionally equivalent, mechanisms.

check original Bateson/DM papers to see if this accords with those defns

I read through Orr's review of those papers, which included excerpts. It seems like the DMI definition is very general and this accords.

Furthermore, using a quantitative genetic approach, we estimated that a genetically variable population will drift in neutral system space at a rate determined by its intra-population variation and its effective population size. Because mechanistically distinct yet phenotypically equivalent biological systems can fail to produce viable hybrids, we predict allopatric populations to accumulate genetic incompatibilities at a rate on the order of N_e if intrapopulation regulatory variation is constrained by selection. Additionally we see second-generation hybrid fitness plummet much faster than that of first-generation hybrids. This result is also consistent with Haldane's rule; that if only one hybrid sex is inviable or sterile it is likely the heterogametic sex. The consistency comes from gene networks localized to the sex chromosomes functioning as an F_2 hybrid cross within a diploid F_1 heterogamete as there is only one sex chromosome.

refer to Turelli here

Lastly, we show that hybrid gene networks, under neutral processes, break down as function of genetic distance, and may, in part, explain broad patterns of reproductive isolation among diverse phyla [Roux et al., 2016].

Acknowledgements

We would like to thank Sergey Nuzhdin, Stevan Arnold, Erik Lundgren, and Hossein Asgharian for valuable discussion.

References

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

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

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

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

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

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, 10

J Felsenstein. Phylogenies and quantitative characters. *Annual Review of Ecology and Systematics*, 19(1):445–471, 1988. doi: 10.1146/annurev.es.19.110188.002305. URL https://doi.org/10.1146/annurev.es.19.110188.002305.

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

Thomas F. Hansen and Emilia P. Martins. Translating between microevolutionary process and macroevolutionary patterns: The correlation structure of interspecific data. *Evolution*, 50(4):1404–1417, 1996. ISSN 00143820, 15585646. URL http://www.jstor.org/stable/2410878. 17, 20

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 Genet*, 4(6):e1000106, 2008. 2

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 Biol*, 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

Johannes Jaeger, Manfred Laubichler, and Werner Callebaut. The comet cometh: evolving developmental systems. *Biological theory*, 10(1):36–49, 2015. 1

Rudolf Emil Kalman. Mathematical description of linear dynamical systems. *J.S.I.A.M Control*, 1963. 2

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.

Konstantin Kozlov, Svetlana Surkova, Ekaterina Myasnikova, John Reinitz, and Maria Samsonova. Modeling of gap gene expression in *Drosophila Kruppel* mutants. *PLoS Computational Biology*, 2012. 1

Konstantin Kozlov, Vitaly Gursky, Ivan Kulakovskiy, and Maria Samsonova. Sequnce-based model of gap gene regulatory network. *BMC Genomics*, 2014. 1

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. Models of speciation by sexual selection on polygenic traits. *Proceedings of the National Academy of Sciences*, 78(6):3721–3725, 1981. URL http://www.pnas.org/content/78/6/3721.abstract. 17

Tuuli Lappalainen, Michael Sammeth, Marc R Friedländer, Peter ACt 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.

Richard Levins. The strategy of model building in population biology. *American scientist*, 54(4):421–431, 1966.

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. 2

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

David A Orlando, Charles Y Lin, Allister Bernard, Jean Y Wang, Joshua ES Socolar, Edwin S Iversen, Alexander J Hartemink, and Steven B Haase. Global control of cell-cycle transcription by coupled cdk and network oscillators. *Nature*, 453(7197):944–947, 2008. 3

Mihaela Pavlicev and Gunter P Wagner. A model of developmental evolution: selection, peliotropy and compensation. *Trends in Ecology & Evolution*, 2012. 1

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. 11

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 theorythe utility of mathematical models in evolutionary biology. *PLoS Biol*, 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

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.

Andrew B. Stergachis, Shane Neph, Richard Sandstrom, Eric Haugen, Alex P. Reynolds, Miaohua Zhang, Rachel Byron, Theresa Canfield, Sandra Stelhing-Sun, Kristen Lee, Robert E. Thurman, Shinny Vong, Daniel Bates, Fidencio Neri, Morgan Diegel, Erika Giste, Douglas Dunn, Jeff Vierstra, R. Scott Hansen, Audra K. Johnson, Peter J. Sabo, Matthew S. Wilken, Thomas A. Reh, Piper M. Treuting, Rajinder Kaul, Mark Groudine, M. A. Bender, Elhanan Borenstein, and John A. Stamatoyannopoulos. Conservation of trans-acting circuitry during mammalian regulatory evolution. *Nature*, 515(7527):365–370, November 2014. ISSN 00280836. URL http://dx.doi.org/10.1038/nature13972. 2

Matthew B Taylor, Joann Phan, Jonathan T Lee, Madelyn McCadden, and Ian M Ehrenreich. Diverse genetic architectures lead to the same cryptic phenotype in a yeast cross. *Nature communications*, 7, 2016. 2

John R True and Eric S Haag. Developmental system drift and flexibility in evolutionary trajectories. *Evolution & development*, 3(2):109–119, 2001. 1, 2, 10

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. 2

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.

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

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

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

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, 10

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

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

Examples

A Kalman Decomposition

Definition 2 (Phenotypic equivalence of systems)**.** *Let* $(\kappa(t), \phi(t))$ *and* $(\bar{\kappa}(t), \bar{\phi}(t))$ *be the solutions to* (**??**) *with coefficient matrices* (A, B, C) *and* $(\bar{A}, \bar{B}, \bar{C})$ *respectively, and both* $\kappa(0)$ *and* $\bar{\kappa}(0)$ *are zero. The systems defined by* (A, B, C) *and* $(\bar{A}, \bar{B}, \bar{C})$ *are* ***phenotypically equivalent*** *if*

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

Equivalently, this occurs if and only if

$$h(t) = \bar{h}(t) \qquad \text{for all } t \geq 0,$$

where h *and* $\bar{h}$ *are the impulse responses of the two systems.*

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

$$CV^{-1}(zI - VAV^{-1})^{-1}VB = CV^{-1}V(zI - A)^{-1}V^{-1}VB = C(zI - A)^{-1}B.$$

However, the converse is not necessarily true: systems can have identical transfer functions without being changes of coordinates of each other. In fact, systems with identical transfer functions can involve interactions between different numbers of molecular species.

The set of all systems phenotypically equivalent to a given system (A, B, C) is elegantly described using the Kalman decomposition, which also clarifies the system dynamics? tells us a lot about how it works? To motivate this, first note that the input $u(t)$ only directly pushes the system in directions 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 the *reachable subspace*, which we denote by $\mathcal{R}$, and is defined to be the closure of span(B) under applying A (or equivalently, the span of $B, AB, A^2B, \ldots A^{n-1}B$). Analogously to this, we define the *observable subspace*, $\mathcal{O}$, to be the closure of span(C^T) under applying A. (Or: $\bar{\mathcal{O}}$ is the largest A-invariant subspace contained in the null space of C; and $\mathcal{R}$ is the largest A-invariant subspace contained in the image of B.)

or something

If we define

1. The columns of $P_{r\bar{o}}$ are an orthonormal basis for $\mathcal{R} \cap \bar{\mathcal{O}}$.
2. The columns of P_{ro} are an orthonormal basis of the complement of $\mathcal{R} \cap \bar{\mathcal{O}}$ in $\mathcal{R}$.
3. The columns of $P_{\bar{r}o}$ are an orthonormal basis of the complement of $\mathcal{R} \cap \bar{\mathcal{O}}$ in $\bar{\mathcal{O}}$.
4. The columns of $P_{\bar{r}\bar{o}}$ are an orthonormal basis of the remainder of $\mathbb{R}^n$.

If we then define

$$P = \left[\begin{array}{c|c|c|c} P_{r\bar{o}} & P_{ro} & P_{\bar{r}o} & P_{\bar{r}\bar{o}} \end{array} \right],$$

then

$$P^T P = \left[\begin{array}{c|c|c|c} I & 0 & 0 & 0 \\ \hline 0 & I & U & 0 \\ \hline 0 & V & I & 0 \\ \hline 0 & 0 & 0 & I \end{array} \right].$$

Check this. Can we get $U = V = 0$?

The following theorem can be found in SOME REFERENCE.

Theorem 3 (Kalman decomposition)**.** *For any system (A, B, C) with corresponding Kalman basis matrix P, the transformed system (PAP^{-1}, PB, CP^{-1}) has the following form:*

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

and

$$\widehat{B} = PB = \begin{bmatrix} B_{r\bar{o}} \\ B_{ro} \\ 0 \\ 0 \end{bmatrix},$$

and

$$\widehat{C} = CP^{-1} = \begin{bmatrix} 0 & C_{ro} & C_{\bar{r}\bar{o}} & 0 \end{bmatrix}.$$

The transfer function of both systems is given by

$$H(z) = C_{ro}(zI - A_{ro})^{-1}B_{ro}.$$

In the latter case, we say that the system is *minimal* – there is no equivalent system with a smaller number of species. Note that this says that any two equivalent minimal systems are changes of basis of each other.

Since any system can be put into this form, and once in this form, its transfer function is determined only by C_{ro}, A_{ro}, and B_{ro}, therefore, the set of all equivalent systems are parameterized by the dimension n, the choice of basis (P), the remaining submatrices in $\widehat{A}$, $\widehat{B}$, and $\widehat{C}$ (which are unconstrained), and a invertible transformation of span(P_{ro}), which we call T_{ro}.

Theorem 4 (Parameterization of equivalent systems)**.** *Let (A, B, C) be a minimal system.*

(a) Every equivalent system is of the form given in Theorem 3, i.e., can be specified by choosing a dimension, n; submatrices in $\widehat{A}$, $\widehat{B}$, and $\widehat{C}$ except for $A_{ro} = A$, $B_{ro} = B$, and $C_{ro} = C$; and choosing an invertible matrix P.

(b) The parameterization is unique if P is furthermore chosen so that each P_x other than P_{ro} is a projection matrix, and that

$$0 = P_x^T P_y$$

for all (x, y) except $(ro, \bar{r}\bar{o})$.

conjecture:

In some situations we may be interested in only "network rewiring", where A changes while B and C do not. For instance, if all non-regulatory functions of each molecule are strongly constrained, then C cannot change. Likewise, if responses of each molecule to the external inputs are not changed by evolution, then B does not change.

Another way of saying it: pick the $\mathcal{R}$ and $\bar{\mathcal{O}}$ subspaces, that must intersect in something of the minimal dimension; then let P be the appropriate basis?

A.1 Neutral directions from the Kalman decomposition

The Kalman decomposition above says that any system (A, B, C) can be decomposed into $(P, \hat{A}, \hat{B}, \hat{C})$ so that

$$A = P^{-1}\hat{A}P$$
$$B = P^{-1}\hat{B}$$
$$C = \hat{C}P,$$

and we know precisely how we can change these to preserve the transfer function:

1. $P \to P + \epsilon Q$ as long as the result is still invertible,
2. $\hat{A} \to A + \epsilon X$ as long as X is zero in the correct places,
3. $\hat{B} \to B + \epsilon Y$ as long as Y is zero in the correct places,
4. $\hat{C} \to C + \epsilon Y$ as long as Z is zero in the correct places.

By taking $\epsilon \to 0$, these tell us the local directions we can move a system (A, B, C) in. All statements below are up to first order in ϵ, omitting terms of order ϵ^2.

First, since $(P + \epsilon Q)^{-1} = P^{-1} + \epsilon P^{-1} Q P^{-1}$, modifying $P \to P + \epsilon Q$ changes

$$\begin{aligned} A &\to A + \epsilon P^{-1} \hat{A} Q - \epsilon P^{-1} Q P^{-1} \hat{A} P \\ &= A + \epsilon \left(A P^{-1} Q - P^{-1} Q A \right), \\ B &\to B - \epsilon P^{-1} Q B \\ C &\to C + \epsilon C P^{-1} Q. \end{aligned}$$

Since P is invertible and Q can be anything (if ϵ is small enough), this allows changes in the direction of an arbitrary W:

$$\begin{aligned} A &= A + \epsilon \left(A W - W A \right), \\ B &\to B - \epsilon W B \\ C &\to C + \epsilon C W. \end{aligned}$$

Then, $\hat{A} \to A + \epsilon X$ does

$$A \to A + \epsilon P^{-1} X P$$

and $\hat{B} \to B + \epsilon Y$ does

$$B \to B + \epsilon P^{-1} Y$$

and $\hat{C} \to C + \epsilon Z$ does

$$C \to C + \epsilon Z P.$$

These degrees of freedom look like they depend on P, which is not unique, but for any two choices of P there are corresponding choices of X that give the same actual change in A (and likewise for Y and Z).

Therefore, this gives us an upper bound on the number of degrees of freedom, in terms of the dimensions of the blocks in the Kalman decomposition (n_{ro} etc) and the dimensions of B and C (n_B and n_C respectively): namely, for W, X, Y, and Z respectively:

$$n^2 + (n_{r\bar{o}} + n_{ro} n_{\bar{r}o} + n_{\bar{r}\bar{o}}(n_{\bar{r}\bar{o}} + n_{\bar{r}o}) + n_{\bar{r}o}^2) + n_B n_{r\bar{o}} + n_C n_{\bar{r}\bar{o}}.$$

However, some of these may be redundant. For instance, changing P in the direction of a Q that satisfies both $AP^{-1}Q = P^{-1}QA$ and $CP^{-1}Q = 0$ is equivalent to changing B by $Y = QB$.

B Meiotic recombination in linear systems

Recombination is performed by taking two analogous system components from $\mathcal{S}$ and $\mathcal{S}'$ and randomly swapping rows or columns. *E.g.* gamete systems (A'', B'', C''), produced by the diploid $\{\mathcal{S}, \mathcal{S}'\}$, are formed by recombining (randomly swapping rows or columns) between two, possibly distinct, systems $\mathcal{S} = (A, B, C)$ and $\mathcal{S}' = (A', B', C')$ such that,

$$\mathcal{S}'' = \begin{pmatrix} A'' & = MA + (I - M)A', \\ B'' & = MB + (I - M)B', \\ C'' & = CM + C'(I - M) \end{pmatrix}$$

where M is a diagonal matrix where each diagonal element is a Bernoulli random variable ($M_{ii} = 0$ or 1 with equal probability, and $M_{ij} = 0$ if $i \neq j$). If systems are different dimensions the smaller system elements can be augmented with 0s (*e.g.* $\left[\begin{smallmatrix} A & 0 \\ 0 & 0 \end{smallmatrix}\right], \left[\begin{smallmatrix} B \\ 0 \end{smallmatrix}\right], \left[\begin{smallmatrix} C & 0 \end{smallmatrix}\right]$).

C Genetic drift with a multivariate trait

For completeness, we provide a brief argument of how the population mean moves under genetic drift with a quantitative genetics model, as in Lande [1981] or Hansen and Martins [1996]. These ignore details of the underlying genetic basis, but developing a more accurate model is beyond the scope of this paper.

Completing the square First note that

$$(x-y)^T A(x-y) = x^T A\,(x-2y) + y^T A y,$$

and so

$$\begin{aligned}(x-y)^T A(x-y) + x^T B x &= x^T (A+B)\left(x - 2(A+B)^{-1}Ay\right) + y^T A y \\ &= \left(x-(A+B)^{-1}Ay\right)^T (A+B)\left(x-(A+B)^{-1}Ay\right) + (\text{terms that don't depend on } x).\end{aligned}$$

Therefore, if $f(x;\Sigma,y)$ is the density of a Gaussian with mean y and covariance matrix Σ then substituting $A=\Sigma^{-1}$ and $B=U^{-1}$ above,

$$\frac{f(x;\Sigma,y)f(x;U,0)}{\int_x f(z;\Sigma,y)f(z;U,0)dz} = f(x;(\Sigma^{-1}+U^{-1})^{-1}, (\Sigma^{-1}+U^{-1})^{-1}\Sigma^{-1}y).$$

Now suppose that the population is distributed in genotype space as a Gaussian with covariance matrix Σ and mean y. Selection has the effect of multiplying this density by the fitness function and renormalizing, so that if expected fitness of x is proportional to $f(x;U,z)$ then the above argument shows that the next generation will be sampled from a Gaussian distribution with covariance matrix $(\Sigma^{-1}+U^{-1})^{-1}$ and mean $z+(\Sigma^{-1}+U^{-1})^{-1}\Sigma^{-1}(y-z)$ Taking a sample of size N to construct the next generation will produce something close to this but with a slightly (stochastically) deviating mean. The next generation's mean is drawn from a Gaussian distribution with mean with covariance matrix $(\Sigma^{-1}+U^{-1})^{-1}/N$ and mean $z+(\Sigma^{-1}+U^{-1})^{-1}\Sigma^{-1}(y-z)$.

Roughly, what is this doing? Suppose that the population mean differs from the optimum by ϵ, that $\Sigma=\sigma^2 I$ and $U=I/\beta^2$ (so, stablizing selection happening on a distance scale of β). Then the population mean gets closer to the optimum on average, moving to $\epsilon/(1+\sigma^2\beta^2)$ and adds noise of size $(1/\beta)\sigma/\sqrt{N\sigma^2+N1/\beta^2}$. At equilibrium, these two movements will be of the same order, so that ϵ is of order $(\sigma/\sqrt{N})\sqrt{1+\sigma^2\beta^2}$.

D Away from the optimum

Let two points on $\mathcal{X}$ be x_1 and x_2, let $\bar{x}=(x_1+x_2)/2$, and let $z=(x_2-x_1)/2$. Then with $D\Phi$ and $D^2\Phi$ the first and second derivatives of Φ, respectively, then Taylor expanding about x_1 and x_2 finds that

$$\begin{aligned}\Phi(\bar{x}) &= \Phi(x_1) + D\Phi(x_1)\cdot z + \frac{1}{2}z^T D^2\Phi(x_1) z + O(\|z\|^3) \\ &= \Phi(x_2) - D\Phi(x_2)\cdot z + \frac{1}{2}z^T D^2\Phi(x_2) z + O(\|z\|^3).\end{aligned}$$

Now, since $\Phi(x_1) = \Phi(x_2) = \Phi_0$ and

$$D\Phi(x_2) = D\Phi(x_1) + 2z^T D^2\Phi(x_1) + O(\|z\|^2), \quad \text{and}$$
$$D^2\Phi(x_2) = D^2\Phi(x_1) + O(\|z\|), \quad \text{and}$$

adding together the two equations above and dividing by two gets that

$$\Phi(\bar{x}) = \Phi_0 - \frac{3}{2} z^T D^2\Phi(x_1) z + O(\|z\|^3).$$

E Differentiating the fitness function

Suppose that $\rho(t) \geq 0$ is a weighting function on $[0, \infty)$ so that fitness is a function of $L^2(\rho)$ distance of the impulse response from optimal. With A_0 a representative of the optimal set:

$$\begin{aligned} D(A) &:= \int_0^\infty \rho(t) \left| h_A(t) - h_{A_0}(t) \right|^2 dt \\ &:= \int_0^\infty \rho(t) \left| Ce^{At}B - Ce^{A_0 t}B \right|^2 dt \\ &= \int_0^\infty \rho(t) \left| C\left(e^{At} - e^{A_0 t}\right)B \right|^2 dt \\ &= \int_0^\infty \rho(t) C\left(e^{At} - e^{A_0 t}\right) BB^T \left(e^{At} - e^{A_0 t}\right)^T C^T dt \end{aligned} \tag{6}$$

How does this change with A? Since

$$\frac{d}{du} e^{(A+uZ)t}|_{u=0} = \int_0^t e^{As} Z e^{A(t-s)} ds, \tag{7}$$

we have that

$$\begin{aligned} \frac{d}{du} D(A+uZ)|_{u=0} &= 2\int_0^\infty \rho(t) C \left(\int_0^t e^{As} Z e^{A(t-s)} ds \right) BB^T \left(e^{At} - e^{A_0 t}\right)^T C^T dt \\ &= 2\int_0^\infty \rho(t) C \left(\int_0^t e^{As} Z e^{A(t-s)} ds \right) B \left(h_A(t) - h_{A_0}(t)\right)^T dt \end{aligned} \tag{8}$$

and, by differentiating this and supposing that A is on the optimal set, i.e., $h_A(t) = h_{A_0}(t)$, (so wolog $A = A_0$):

$$\begin{aligned} \mathcal{H}(Y,Z) &:= \frac{1}{2} \frac{d}{du} \frac{d}{dv} D(A_0 + uY + vZ)|_{u=v=0} \\ &= \int_0^\infty \rho(t) C \left(\int_0^t e^{A_0 s} Y e^{A_0(t-s)} ds \right) BB^T \left(\int_0^t e^{A_0 s} Z e^{A_0(t-s)} ds \right)^T C^T dt. \end{aligned} \tag{9}$$

Here $\mathcal{H}$ is the quadratic form underlying the Hamiltonian. 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 is

$$H_{ij,k\ell}(A) := \mathcal{H}(\Delta_{ij}, \Delta_{k\ell}). \tag{10}$$

We could use this to compute the gradient of D, or 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 can diagonalize $A = U\Lambda U^{-1}$. Then

$$\int_0^t e^{As} Z e^{A(t-s)} ds = \int_0^t U e^{\Lambda s} U^{-1} Z U e^{\Lambda(t-s)} U^{-1} ds \tag{11}$$

Now, notice that

$$\int_0^t e^{s\lambda_i} e^{(t-s)\lambda_j} ds = \frac{e^{t\lambda_i} - e^{t\lambda_j}}{\lambda_i - \lambda_j}. \tag{12}$$

Therefore, defining

$$X_{ij}(t, Z) = \left(U^{-1} Z U\right)_{ij} \frac{e^{t\lambda_i} - e^{t\lambda_j}}{\lambda_i - \lambda_j} \tag{13}$$

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

$$\int_0^t e^{As} Z e^{A(t-s)} ds = U X(t, Z) U^{-1}. \tag{14}$$

Following on from above, we see that if $Z = \Delta_{k\ell}$, then

$$X_{ij}^{k\ell}(t) = \frac{e^{t\lambda_i} - e^{t\lambda_j}}{\lambda_i - \lambda_j} (U^{-1})_{\cdot k} U_{\ell \cdot}, \tag{15}$$

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

$$H_{ij,k\ell}(A) = \int_0^\infty \rho(t) C U X^{ij}(t) U^{-1} B B^T (U^{-1})^T X^{k\ell}(t)^T U^T C^T dt. \tag{16}$$

This implies that

$$D(A_0 + \epsilon Z) \approx \epsilon^2 \sum_{ijk\ell} H^{ij,k\ell} Z_{ij} Z_{k\ell} \tag{17}$$

and so

$$D(A_0 + \epsilon Z) \approx \epsilon^2 \sum_{ijk\ell} H^{ij,k\ell} Z_{ij} Z_{k\ell} \tag{18}$$

By section C, if we set $\Sigma = \sigma^2 I$ and $U = H$, then a population at $A_0 + Z$ experiences a restoring force of strength $(I + \sigma^2 H^{-1})^{-1} Z$ (treating Z as a vector and H as an operator on these). If σ^2 is small compared to H^{-1} then this is approximately $-\sigma^2 H^{-1} Z$. This suggests that the population mean follows an Ornstein-Uhlenbeck process, as described (in different terms) in Hansen and Martins [1996].