To do, soon:

1. *(josh)* Add biology to introduction. (keep an eye out for concrete numbers we can use!) *added some relevant biology citations to introduction. Still needs a lot of refinement.* 1

2. Do some order-of-magnitude calculation to figure out how fast systems should drift.

    - *(peter)* Work out the math for how to do this.
    - *(josh)* Find real parameters to put into this. *added some notes for TFBS gain and loss calculations from Lynch* 4

3. *(josh)* Write out simple examples in biological words. 13

4. *(josh)* Rephrase hybrid inviability example in biological terms. Discuss F1's (or heterozygotes) versus F2's.

5. *(josh)* Clarify our assumptions around $x(0)$.

6. *(josh)* Note we don't include D for simplicity but don't.

7. *(peter)* Figure out how to specify P for Kalman decomposition.

8. *(josh ✓,peter)* Prove or find citation for uniqueness of minimal system up to basis change. *See [Kalman et al., 1969, Chapter 10.6, Theorem 6.9: Fundamental Theorem of Linear Realization Theory.]*

9. Work examples (minimal; not minimal; with/without B,C fixed) of Kalman decomposition and description of moving along the walk.

    - *(josh)* Work out parameterization.
    - *(peter)* Describe Brownian motion.

10. *(peter)* Write down and decide between options for parameterizing the random walk on the neutral manifold: use Kalman decomposition, and parameterize Brownian motion so that it has constant speed; OR consider Josh's parameteriziation like this; OR just say we do Brownian motion on the manifold (more or less implicit characterization).

11. *(peter)* Describe model leading to constraints on regulation so we have a stationary distribution. Ask questions of stationary distribution: robustness, connectivity, xxx?

12. Consider options for fitness function: Euclidean? Sign of eigenvalues? Time domain or frequency? First try on simple concrete examples. *see 12*

13. Figure out coherent model for changing dimension: do coefficients stay stuck at zero? In what parameterization? Do we need to allow deleterious variation to get deletion?

14. Write something in discussion about nonlinear systems. Do an example if there is a feasible one.

15. Simulate.

The Evolution of Phenotypically Invariant Genetic Networks

Joshua S. Schiffman[1] and Peter L. Ralph[1,2]

[1]University of Southern California, Los Angeles, California
[2]University of Oregon, Eugene, Oregon
jsschiff@usc.edu plr@uoregon.edu

April 17, 2017

Abstract

I will outline an analytical theory to study the evolution of biological systems such as gene regulatory networks, borrowing insight and tools from control engineering, systems identification, and dynamical systems theory. I will describe a null model of regulatory network evolution by analytically describing the set of all linear gene networks (of any size) that produce identical phenotypes – and the evolutionary paths connecting them. In the idealized case of a perfectly adapted population, constant selection, and a static environment, we observe neutral evolution as a random walk over the phenotypically- invariant network-space. Under neutral conditions, this model can provide descriptions of expected network size and connectivity under mutation-selection equilibrium, estimate the rate of regulatory rewiring, and the rates at which Dobzhansky-Muller incompatibilities arise in reproductively isolated populations. This analysis provides insight into the mechanisms and parameters important for understanding developmental systems drift, network rewiring, evolvability, epistasis, and speciation, as well as the tenuous connection between network architecture and function.

1 Background

Review of relevant other stuff.

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 experiment and 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]. However, as Hardy and Weinberg taught us over a century ago, verbal theories are often insufficient, if not downright misleading [Hardy, 1908, Weinberg, 1908, 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. Previously it has been suggested that any idealized study of evolution is incomplete without a mathematically sufficient description of the genotype, phenotype, and transformation from one to the other [Lewontin et al., 1974].

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] and empirically inspired computational and mathematical models of gene regulatory networks (GRNs) [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 uncertaintity 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 frusturate 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 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, Vierstra et al., 2014, Stergachis et al., 2014, Taylor et al., 2016, Matsui et al., 2015].

In this paper we propose a framework to study the evolution of biological systems. To begin, we focus on the evolution of an idealized population. We consider the evolution of a perfectly adapted, large population, evolving in a static environment for an infinite number of generations. Under these ideal circumstances, we expect to observe a "conservation of phenotype," where the population explores the manifold of phenotypically-invariant (or symmetric) genetic and developmental architecutres. We would like to understand which parameters influence the distribution of a population along the manifold of phenotypically invariant genetic systems. Further, we can show how dispersion along this manifold contributes to speciation and evolvability.

phenomenologically or dynamically equivalent? The set of all systems – that is all model structures and all parameter schemes – that are externally, or phenomenologically, equivalent, biologically realizable, and mutationally accessible,

What are typically referred to as "design principles" or some sort of clear structure-function relationships are commonly saught [Milo et al., 2002, Alon, 2007, Ma et al., 2009].

Seeking to attribute functionality, uninformed by evolution, can lead to spurious claims (i.e.

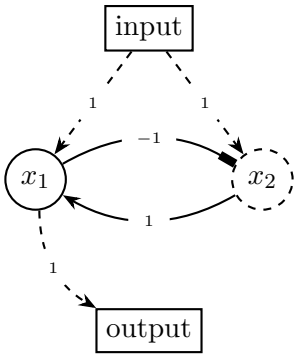


Figure 1: Diagram of Example 1 in the text. *explain what arrows mean if nec*

see Graur [Lynch, 2007a,b, Graur et al., 2013]).

A linear time invariant (LTI) dynamical system is a system of linear differential equations that describes a physical process. LTI systems are used in electrical and control engineering to model a myriad of phenomena including circuits. One can use this methodology, under a set of assumptions, to reverse engineer a mechanism from impulse experiments (input/output data). The idea is that given a black box, and it's experimental manipulation, can we describe the internal mechanisms?

2 The problem

Organisms often have to respond to their environments, and the way this happens is that external input triggers a cascade of molecular signals whose eventual result is a reaction, whose appropriateness for the situation determines fitness. *A similar situation occurs in development (...?).* However, there are many ways to skin a cat *(there are also many ways to develop a cat)*: especially given a number of possible molecular intermediaries, there may be a large number of ways to construct such a molecular signalling system that has precisely the *same* input-output relationships, and hence the same function. Therefore, any mutation that changes one such system to another, equivalent version will be neutral, and may well become polymorphic in a population if there are no ill effects in heterozygotes. Certainly, mutations that move a system away from optimal functioning will also appear, but in very large populations will not rise to high frequency. *Deleterious things are quite important but somehow justify starting with the neutral ones.*

More stuff on why we care about this: Provide neutral model; can address questions. Main point of paper.

Example 1 (Example: Oscillator.) *Motivate the system:*

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

as a special case of the below.

Example 2 (Example: two-component systems.) *As a more general example, suppose that a cell produces two molecular species, S_1 and S_2 whose production are both stimulated by the presence*

of an external "input" substance. call them TF? Suppose that the two species self-regulate with strengths λ_1 and λ_2, respectively, and the second species also regulates the first species with strength γ. Rephrase in terms of binding to promoter sequences? However, only the time course of the concentration of the first one of these species determines the fitness of the organism. (As happens for instance in XXX.) Write $x(t) = (x_1(t), x_2(t))$ for the concentrations of the two species at time t, and write $u(t)$ for the concentration at time t of the input substance. Then, if the rates at which each species are produced are linear functions of the concentrations (look up way this is usually said in the literature.), then the dynamics of the system are given by

$$\dot{x}_1(t) = \lambda_1 x_1(t) + \gamma x_2(t) + u(t)$$
$$\dot{x}_2(t) = \lambda_2 x_1(t) + u(t),$$

where $\dot{x}$ denotes the time derivative. The initial conditions, $x_1(0)$ and $x_2(0)$, and the input $u(t)$ then determine the concentrations through time. If we record the regulatory coefficients in the matrix:

$$A = \begin{bmatrix} \lambda_1 & \gamma \\ 0 & \lambda_2 \end{bmatrix},$$

and define the column vector $B = [1, 1]^T$, then in matrix notation the dynamics are

$$\dot{x}(t) = Ax(t) + Bu(t).$$

If these are all transcription factors, and we suppose their binding motifs are fixed, then the i^{th} row of A is determined by the i^{th} promoter sequence.

Since the system is linear, the state of the system at any time is the superposition of its responses to all previous inputs. This implies that if $G(t)$ is the transient response of the system to a unit impulse, t units of time later, and $x(0) = 0$, then the time course of the concentration of the thing we care about, $x_1(t)$, can be written as

$$x_1(t) = \int_0^t G(t-s)u(s)ds. \tag{1}$$

However, it turns out that *look for simple demontration this is true*

$$P_p = \begin{bmatrix} 1 & 0 \\ p & 1-p \end{bmatrix},$$

then for any $p \neq 1$, if we replace the regulatory matrix A with

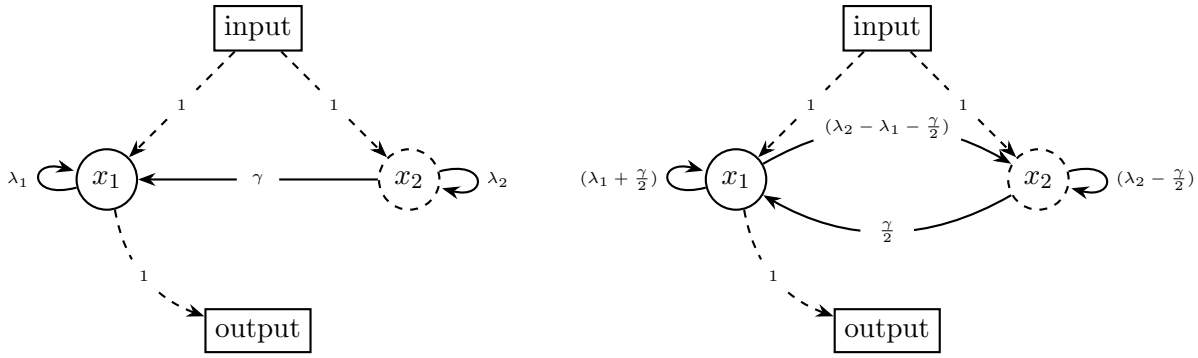
$$A(p) = PAP^{-1},$$

then the response of $x_1(t)$ to any particular input $u(t)$ is identical to the original system. (However, x_2 may well be different!)

For instance, setting $p = -1$, the system with

$$A(-1) = \begin{bmatrix} \lambda_1 + \gamma/2 & \gamma/2 \\ \lambda_2 - \lambda_1 - \gamma/2 & \lambda_2 - \gamma/2 \end{bmatrix},$$

looks very different, but gives the same input-output relationship. Although these systems are equivalent, hybrids between them may not be: *do example*.



(a) Gene Regulatory Network $A(0)$.

(b) Gene Regulatory Network $A(-1)$.

Figure 2: Neutral rewiring of gene network $A(p)$ from Example 2. Not only are the regulatory coefficients different between $A(0)$ and $A(-1)$, there is also a new regulatory connection.

Example 3 (Hyrbid Inviability) *Let,*

$$A(0) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \text{ and } A(2) = \begin{bmatrix} 2 & -1 \\ 5 & -2 \end{bmatrix}$$

from Example 1 ($B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $C = \begin{bmatrix} 1 & 0 \end{bmatrix}$). If hybrids are formed by exchanging genes (swapping rows), the hybrid's gene network is,

$$\textit{hybrid network } := \mathcal{H}(A(0), A(2)) = \begin{bmatrix} 2 & -1 \\ -1 & 0 \end{bmatrix},$$

Or should it be:

$$\textit{hybrid network } := \textit{mean}(A(0), A(2)) = \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix},$$

It's transfer function will be $H(z, \mathcal{H}) = \frac{z-2}{z^2-3z+1}$, which produces very different dynamics than its parents, $H(z, A(0)) = H(z, A(2)) = \frac{z+1}{z^2+1}$. The time dynamcs of molecular species 1 in the hybrid is $y^{(\mathcal{H})} = e^t \cosh(\sqrt{2}t)$ which is approx. $\approx e^t$, with the concentration of species 1 rapidly heading towards infinity. The dynamics of species 1 in the parental networks is $y = \sin(t) + \cos(t)$, where it continuously oscillates.

The transfer function $H(z, \mathcal{H})$, by averaging networks, is $\frac{1}{z-1}$ and the time dynamics are: $y^{(\mathcal{H})} = e^t$.

Definition 1 (Biological Species Concept) *Let $\mathcal{A}$ be the set of all phenotypically symmetric gene networks with transfer funtion $H(z)$. Choose A and $\widetilde{A}$, both in $\mathcal{A}$. If $\mathcal{H}(A, \widetilde{A}) \notin \mathcal{A}$ then A and $\widetilde{A}$ are distinct species.*

3 Regulatory networks as linear, time-invariant systems

Make sure to have strongly worded disclaimer somewhere that we know real networks aren't linear but they are locally. Plus, it works in engineering.

Maybe we should discuss the Hartman-Grobman Theorem?

We will now lay out the model in more general terms. Suppose that the *internal state* of the system is parameterized by the concentrations of a collection of n molecular species, $S_1, \ldots, S_n$, and the vector of concentrations at time t we denote $x(t) = (x_1(t), \ldots, x_n(t))$. There are also m "input" species, whose concentrations are determined exogenously to the system, and are denoted $u(t) = (u_1(t), \ldots, u_m(t))$, and ℓ "output" species, whose concentrations are denoted $y(t) = (y_1(t), \ldots, y_\ell(t))$. The output is merely a linear function of the internal state:

$$y_i(t) = \sum_j C_{ij} x_i(t).$$

Since y is what natural selection acts on, we refer to it as the *phenotype*, and sometimes in contrast refer to x as the *kryptotype*, as it is "hidden" from direct selection. *Maybe.* The rate at which the i^{th} species is produced is a weighted sum of the concentrations of the other species as well as the input:

$$\dot{x}_i(t) = \sum_j A_{ij} x_j(t) + \sum_k B_{ik} u_k(t).$$

In matrix notation, this is written more concisely as

$$\dot{x}(t) = Ax(t) + Bu(t) \tag{2}$$
$$y(t) = Cx(t). \tag{3}$$

Everything we say I think is assuming $x(0) = 0$). Figure this out. "Zero-state equivalence."

Given an initial condition and an input, it is possible to write the solution $x(t)$ as a convolution with an input-response kernel as in the example above (equation (1)). An alternative way of describing the input-output relationship, more common in engineering, is instead to find the *transfer function*

$$H(z) = C(zI - A)^{-1}B, \tag{4}$$

which is more mathematically readable *if you know what it means*. This can be thought of as the system's response to input at "frequency" z. The fact that the transfer function uniquely determines the system's input-output relationship (i.e., the mapping $u \mapsto y$) follows from the fact that if $\tilde{y}(z) = \int_0^\infty e^{-zt} y(t) dt$ is the Laplace transform of y, and $\tilde{u}(z)$ is the Laplace transform of u, then

$$\tilde{y}(z) = H(z)\tilde{u}(z).$$

(This happens because the Laplace transform takes convolutions to products, and H is the Laplace transform of the kernel G.) Concretely,

Should we remind what dimensions everything is?

The state space representation is considered an internal description of the system, whereas the transfer function is an external description. "Realizations" are something.

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

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

Equivalently, this occurs if and only if

$$H(z) = \bar{H}(z) \qquad \text{for all } z \geq 0,$$

where H and $\bar{H}$ are the transfer functions of the two systems.

One way to find other systems equvalent to a given one is by change of coordinates ("algebraic equivalance"): if T is an invertible matrix, then the systems (A, B, C) and (TAT^{-1}, TB, CT^{-1}) have the same dyanamics because their transfer functions are equal:

$$CT^{-1}(zI - TAT^{-1})^{-1}TB = CT^{-1}T(zI - A)^{-1}T^{-1}TB = C(zI - A)^{-1}B.$$

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 phenomenologically 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? *or something* To motivate this, first note that the 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 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 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 2 (Parameterization of equivalent systems) *Let (A, B, C) be a minimal system.*

(a) Every equivalent system is of the form given in Theorem 1, 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) conjecture: The parameterization is unique if P is futhermore 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 $(ro, \bar{r}\bar{o})$.

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?

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.

Lemma 1 *Let (A, B, C) be a system. Every matrix $\widehat{A}$ such that $(\widehat{A}, B, C)$ is equivalent to (A, B, C) can be specified in the form given in Theorem 1, with them having the same ro bit and satisfying*

$$\widehat{B} = PB$$
$$\widehat{C}P = C$$

It is remarkable to note that even with the relationships between environment, cryptotype, 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 of Example 2 above are minimal, and share common B and C matrices.

Maybe say that stability to noise can be determined by $A_{\bar{r}o}$ (and others?), while complexity of the internal, unobserved dynamics are somehow given by the difference between the dimension of y and the dimension of A_{ro}.

General argument why Brownian motion on set of equivalent things is maybe ok.

Example 4 *The system in Example 2 is minimal, so $A = A_{ro}$, and the only free parameter is P. Suppose also that the effects of the input and output are strongly constrained, so that as in that example, B and C cannot change, and neutral evolution moves through the set of of P such that $PB = B$ and $CP = C$. Clarify we mean with no gene duplication and only neutral drift. How will this system change over evolutionary time in the manner described above? An unbiased Brownian motion on this set can be written in terms of a one-dimensional Brownian motion β_τ with $\beta_0 = 0$ as*

$$P(\tau) = \begin{bmatrix} 1 & 0 \\ e^{\beta_\tau} - 1 & e^{\beta_\tau} \end{bmatrix}.$$

Therefore, the state of the system after (evolutionary) time τ is $(P(\tau)AP(\tau)^{-1}, B, C)$.

4 Estimates of network rewiring and developmenal systems drift.

Parameter estimates for DSD ↓

What parameters (and other relevant factors) do we need from the literature and experiment?

- mutation rate
- Transcription factor binding site appearence and loss
  - In [Lynch, 2007a] TF binding site gain (u_g) and loss (u_l) is estimated to be,

$$u_g = \frac{(L - n + 1)nu}{4^n}$$
$$u_l = nu$$

Where L is the length of the promoter, n is the number of nucleotides in a TFBS, and u is the mutation rate per base pair per generation.

- Lynch [2007a] further states that in prokaryotes L will usually be ≤ 100 and in eukaryotes L can range from 10^3 to 10^6. He also cites [Harbison et al., 2004] that TFBSs range from $5-10$bps in length.
- in [Lynch and Hagner, 2014], they estimate

$$\tilde{P}(m) = C\left[3^{\ell-m}\begin{pmatrix}\ell\\ m\end{pmatrix}\right]e^{2N_e s(m)},$$

where $\tilde{P}(m)$ is the stationary distribution of a TFBS with match (and strength) m. C is the normalizer constant, m is the number of matches, ℓ is the length of the TFBS motif. The term on the right is the ratio of fixation probabilities. Or in their own words, there are two interpretations:

> "(i) For a single TFBS, it represents the long-term proportion of evolutionary time spent in the various matching states and (ii) for a set of different TFBSs under the same selective constraints, it represents the expected distribution of states at any point in time."

- Tuğrul et al. [2015], Anderson et al. [2015] also may be helpful.

- rate of both *trans-* and *cis-* regulatory evolution. (see Sergey's papers). i.e. [Nuzhdin et al., 2004, Fear et al., 2016, Graze et al., 2009, 2012, Genissel et al., 2008]
- The PWN-score landscape (does binding strength increase linearly, exponentially, have a threshold, etc.) with respect to SNPs.
- Others?

5 Algorithm for Gene Network Growth

Let $A^(d)$ be a gene network with index (d); (d) is the difference between the specific realization's dimension and the dimension of a minimal realization of the requisite transfer function, such that $A^{(0)}$ is a minimal realization, $A^{(1)}$ has one superfluous dimension and $A^{(-1)}$ is one short.

Algorithm for constructing $A^{(d+1)}$ from $A^{(d)}$,

$$A^{(d+1)} = \left[\begin{array}{c|c} A^{(d)} & 0 \\ r_1 \dots r_{n+d} & r_{n+d+1}\end{array}\right].$$

Two systems have equivalent external dynamics iff their transfer functions are identical. Two transfer functions are identical iff $CA^kB = \bar{C}\bar{A}^k\bar{B} \qquad k = 0, 1, \dots$.

If system $\bar{A} := A^{(1)}$ is an extension of the minimal system $A^{(0)}$ then it must be of the form,

$$\left[\begin{array}{c|c} C & 0\end{array}\right]\begin{bmatrix} V & \Psi \\ \Phi & \Gamma\end{bmatrix}\begin{bmatrix}\Lambda & 0\\ 0 & I\end{bmatrix}\begin{bmatrix} V^{-1} & \Omega \\ \Theta & I\end{bmatrix}\left[\begin{array}{c} B \\ \hline 0\end{array}\right]$$

$$\Phi\Theta = 0$$

$$\bar{C}\bar{A}^k\bar{B} = (CV\Lambda^kV^{-1}B + C\Psi\Theta B).$$

6 Adaptation

Suppose a population is perfectly adapted to its environment – it expresses a metabolic enzyme in response to environmental sugar availability at a rate that balances the energetic costs and benefits of the enzyme's synthesis and utility. Then, the population suddenly finds itself in a novel environment and exhibits suboptimal expression rates. How does the population adapt? Are some network mechanisms mutationally closer or farther from the desired new network connectivity?

Let $A_\alpha(p)$ be the pth system realization perfectly adapted to the first environment (α), and let $A_\beta(q)$ be the qth system realization perfectly adapted to the second environment (β).

$$\{p, q : \min \sum_{i,j}^{n} |A_\alpha(p)_{i,j} - A_\beta(q)_{i,j}|\}$$

In an asexual population, if there is variation in network topology, the A_α network closest to the organization of a A_β network might be the fastest to adapt and outcompete all of its neighbors.

7 Mutational Robustness

Are any network realizations $A(p)$ more mutationally robust than others? If $s_{i,j}$ is the fitness cost of mutating element $A_{i,j}$ and $\mathbb{E}_p(s) = \frac{1}{n} \sum_{i,j}^{n} s_{i,j}$ is the mutationaly susceptibility of realization p, is there a p such that $\mathbb{E}(s)$ is minimal? Will the population find this network orgnization?

8 Network Ratchet

If a network evolves to include more molecules (such as transcription factors) and thus becomes higher dimensional, can the network also remove these originally superfluous interactions or will removing these typically lead to catastrophic network dysfunction after some amount of evolutionary time and system drift has occured?

9 Network Minimality versus Rube Goldbergness

Are there evolutionary differences between realizing a transfer function minimally (with a reachable and observable system) or non-minimally (a system with unreachable and/or unobservable subspaces)? Can reducible systems be more or less mutationally robust? Are minimal systems energetically more efficient? Are reducible systems more or less evolvable?

10 (Old Notes) What biological question are we asking?

How do gene regulatory networks (GRNs) change through evolutionary time? Specifically, under constant selection and environmental pressures, how does a maximally fit population's GRN change and what molecular and population parameters are significant to the process? How frequent does gene network rewiring occur in the absense of genetic drift or adaptation? Is rewiring typically compensatory and ususally follow the fixation of a deleterious allele, or can networks neutrally

reorganize? How does the size, average degree, function, peiotropy, etc of a network influence its likelihood to drift? If networks can drift, how much topological heterogeneity can a well-mixed population tolerate, if any? Does topoligical heterogeneity and/or the size of the neutral genotype-set, if consequences of network organization, constrain or entail exaptation and evolvability? Are some network organizations (maybe higher dimensional realizations?) able to more rapidly evolve to construct novel phenotypes and meet the demands of changes in selection pressures?

11 (Old Notes) How can we mathematically model this question?

In some respect an organism can be thought of as a black box that responds to some set of inputs (it's environment and other initial conditions) and outputs a phenotype. The phenotype is then evaluated by natural selection. The black box metaphor holds, because much of the details of what happens to construct a phenotype are unimportant as far as selection is concered, so long as the phenotype reliably performs its function. If multiple internal mechanisms can map the same set of inputs to the same set of outputs, and these mechanisms are mutationally nearby (in genotype space), then evolution may drift from mechanism to mechanism.

In this simple model we can say that $B\vec{u}$ is a list of initial GRN protein concentrations determined by the more general initial environmental conditions $\vec{u}$. A represents all the genetic interactions of a particular GRN, and $C\vec{x}$ is an organism's phenotype. Fitness could be calculated by comparing a subset of the phenotype to an optimum. Maybe, $f(\cdot) = e^{-\int_a^b \|G(s)-G^*(s)\|ds}$ Alternatively, fitness scores could simply assess whether or not a phenotype is within some range or breaks some threshold by some time point. Or possbily the simplest: any phenotype not exactly equal to $G(s)$ is lethal, and $G(s)$ is perfectly fit.

Next, we will have to define mutation and recombination. Mutation needs to add and remove genetic interactions and recombination needs to shuffle genes during sexual reproduction. We could adapt methods from Lynch 2007 and Lynch and Hagner 2015 to model the probability of a TFBS appearing or disappearing due to mutational pressure. Every time a new binding site is gained or lost, the values within the A matrix are either modified or rows and columns are added removed following a specific set of rules. I am not sure if it would be easier to maintain a matrix size (say $m \times m$) where m is arbitrarily large, with most entries being zero, or if a matrix should only have as many dimensions as the number of active TFs. Recombination should then shuffle rows of the A matrix randomly, as each row represents the regulatory region of a given gene (assuming only cis-regulation, and that the regulatory element is small enough to not break up during recombination).

To study some of the above posed questions, we can simulate (or analytically determine) the rates of GRN change, and how frequently these changes lead to significant rewiring or speciation. We would want to know the dynamics of a brownian motion over the set of all realizations. Another interesting question would be to determine the size of the neutral genotype space as well as the number of connections from the neutral genotype space to the non-neutral genotype space. Maybe higher dimensional realizations of a GRN will have significantly more connections to non-neutral genotypes and therefore be more "evolvable" or exapted to novel environments?

Let $M \in \mathbb{R}^{n\times n}$ be a least dimensional realization of A, and let M' bean $(n+m \times n+m)$ matrix where $M'_{(i,j)} = M_{(i,j)}$ for $i, j \leq n,$ and 0 otherwise. Let P be any $(n+m \times n+m)$ matrix with rank n. Then for any matrix B such that $BP = PM'$, B is a realization of A.

12 Including Deleterious Mutations and a Fitness Function

At the moment we focus strictly on the case where the minimal realization is in 2 dimensions and both of its eigenvalues are real and negative (so a special case of the system from Example 2). *(Note: this is not to say the ideas below will not be generalizable).*
The system of interest has the form,

$$\Sigma := \left\{ A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, C = \begin{bmatrix} 1 & 0 \end{bmatrix} \right\}$$

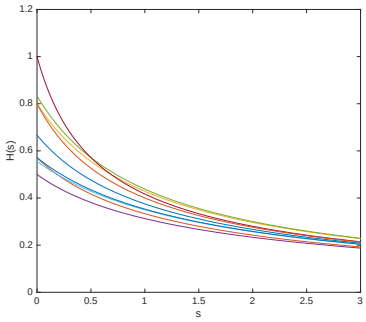


Figure 3: The transfer function for $A = \begin{bmatrix} -3 & 1 \\ 0 & -1 \end{bmatrix}$ and all eight $\mu(\Sigma, A_{ij})$ mutations, where $\upsilon = \pm 0.5$

Define the fitness function to be,

$$\Phi(\bar{\Sigma}) := \begin{cases} e^{-\left|H(0)-\tilde{H}(0)\right| = -\left|\frac{k}{\lambda_1\lambda_2} - \frac{\tilde{k}}{\tilde{\lambda}_1\tilde{\lambda}_2}\right|} & \text{if } \lambda_{1,2} < 0 \text{ and } \in \mathbb{R},\ \det(A) \neq 0 \\ 0 & \text{if } \lambda_{1,2} \geq 0 \text{ or } \notin \mathbb{R},\ \det(A) = 0 \end{cases} \tag{5}$$

and let υ be the magnitude and direction (sign) of a mutation on regulatory interaction strength. The fitness function (5) may me appropriate if mutations only slightly alter the dynamics of the system, as in Figure 3.

The transfer function at $z = 0$ for any GRN with mutations is,

$$H(0, \begin{bmatrix} A_{11} + \upsilon & A_{12} + \varepsilon \\ A_{21} + \delta & A_{22} + \omega \end{bmatrix}) = \frac{A_{22} - A_{12} + \omega - \varepsilon}{\det(A) + \omega A_{11} - \delta A_{12} - \varepsilon A_{21} + \upsilon A_{22} + \omega\upsilon - \varepsilon\delta}$$

If we want to know how new mutations fix in the population, we can explicitly show all mutational paths that could compensate for fixed mutations at sites A_{ij}. In a population, not at a phenotypic optimum, we are interested in the combination of mutations with varying magnitudes and locations that are most likely to fix.

$$\frac{(A_{22} - A_{12}) + \left(\sum_{\jmath=1}^{\mathbb{1}_\omega} \omega_\jmath - \sum_{\jmath=1}^{\mathbb{1}_\varepsilon} \varepsilon_\jmath \right)}{\det(A) + \sum_{\mu=\{\omega,\delta,\varepsilon,\upsilon\}}^{\mathbb{1}_\mu} \mu_\jmath (-1)^{i+j} A_{ij} + \sum_{\jmath=1}^{\mathbb{1}_\omega} \omega_\jmath \sum_{\jmath=1}^{\mathbb{1}_\upsilon} \upsilon_\jmath - \sum_{\jmath=1}^{\mathbb{1}_\varepsilon} \varepsilon_\jmath \sum_{\jmath=1}^{\mathbb{1}_\delta} \delta_\jmath}$$

13 Biological Examples

Example 5 *Negative feedback and Positive feedback systems:*

$$T(z) = \frac{H(z)}{1 + G(z)H(z)}$$
$$T(z) = \frac{H(z)}{1 - G(z)H(z)}$$

Example 6 *Systems in serial* $(u \to \Sigma_1\{A_1, B_1, C_1\} \to \Sigma_2\{A_2, B_2, C_2\} \to y)$*:*

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \left[\begin{array}{c|c} A_1 & 0 \\ \hline B_2 C_2 & A_2 \end{array}\right] \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} + \begin{bmatrix} B_1 \\ 0 \end{bmatrix} u$$
$$y = \begin{bmatrix} 0 & C_2 \end{bmatrix} \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix}.$$

Systems in serial could be applied to model a gene network A_1 *and the subsequent degradation rates of its outputs at rate* A_2*.*

Example 7

$$\dot{x} = \begin{bmatrix} + & - \\ 0 & + \end{bmatrix} \vec{x} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \vec{u}$$
$$y(t) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \vec{x}$$
$$E(t) = \begin{bmatrix} -y_1(t) & 0 & 0 \\ y_1(t) & -y_2(t) & 0 \\ -\sigma_1 y_1(t) & (1-\sigma_2) y_2(t) & 1 \end{bmatrix} \vec{\zeta} + \begin{bmatrix} u_1 \\ u_2 \\ \epsilon \end{bmatrix}.$$

Where u_1 *is input of* lactose*,* u_2 *is the input of* glucose*,* y_1 *is the enzyme* lactase *and* y_2 *is enzyme* "glucase"*.* $E(t)$ *is metabolism of the organism, converting sugars to energy as a function of enzyme concentrations. Lactase converts lactose to glucose, and the enzyme is produced at energy cost* σ_1*; glucase converts glucose to energy and is produced at cost* σ_2*. Fitness is a function of the total energy of the system, given inputs, possibly:* $\int_0^T E(t)_3 dt$ *after some time* T *has passed. Or perhaps, once the total metabolic energy produced passes some threshold, the organism undergoes fission. Thus, fitness selects for maximizing energy as quickly as possible.*

Example 8 *Gene Regulatory and Metabolic Network Cascade:*
Gene Regulatory Network

$$\dot{x}(t) = Ax(t) + B_1 u(t)$$
$$y(t) = C_1 x(t)$$

Metabolism

$$\dot{\zeta}(t) = M\big(y(t), \sigma\big)\zeta(t) + B_2 \begin{bmatrix} u(t) \\ v(t) \end{bmatrix}$$
$$m(t) = C_2 \zeta(t)$$
$$b(t) = C_3 \zeta(t)$$

Fitness

$$\Phi(m(t), p(t))$$

The GRN produces time varying proteins involved in the metabolism. The metabilic network (MN), is a function of protein concentrations $y(t)$ and production costs σ. The metabolic network produces energy $m(t)$ and toxic byproducts $b(t)$. Fitness Φ, or maybe something more concrete, like fission time and/or cell death, is a function of both energy $m(t)$ and toxins $b(t)$.

Example 9 *Rough yeast colony morphology phenotype / Simplified Ras pathway. Rough colony morphology is a complex phenotypic trait influenced by many different genetic variants. Ultimately it appears that colony roughness is a function of RNA Polymerase II activity, downstream of Ras and Gpr1, and acting on FLO11. If RNA pol II activity breaks a threshold level, FLO11 is expressed and colony roughness is observed.*

$$\dot{x} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ + & 0 & \zeta & 0 & 0 & 0 \\ + & \eta & 0 & 0 & 0 & 0 \\ 0 & + & 0 & 0 & 0 & 0 \\ 0 & + & 0 & 0 & 0 & 0 \\ 0 & 0 & - & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} Cyr1 \\ Tpk1 \\ Tpk3 \\ Flo8 \\ Mga1 \\ Sfl1 \end{bmatrix} + \begin{bmatrix} + & + & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & + \end{bmatrix} \begin{bmatrix} Gpa2 & Ras & Sfl1 \end{bmatrix}$$
$$y = \textit{cumulative impact on RNA Pol II activity} = \begin{bmatrix} 0 & 0 & 0 & + & + & - \end{bmatrix} \vec{x}$$

14 Notes on what to do when allowing deleterious mutations

1. Epistasis, coming from mutations deleterious in some networks but not in others (come from incompatibilities).
2. Describe heterosis.
3. Distribution of dominance coefficients.
4. Evolvability: easier for bigger networks to adapt to a new environment?
5. Do short-term responses to strong directional selection have deleterious epistatic consequences in other traits?

References

Uri Alon. Network motifs: theory and experimental approaches. *Nature Reviews Genetics*, 8(6): 450–461, 2007. 3

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. 3

Dave W Anderson, Alesia N McKeown, and Joseph W Thornton. Intermolecular epistasis shaped the function and evolution of an ancient transcription factor and its dna binding sites. *Elife*, 4: e07864, 2015. 8

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

Aviv Bergman and Mark L Siegal. Evolutionary capacitance as a general feature of complex gene networks. *Nature*, 424(6948):549–552, 2003. 2

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. 2

Gheorghe Craciun and Casian Pantea. Identifiability of chemical reaction networks. *Journal of Mathematical Chemistry*, 44(1):244–259, 2008. 3

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. 2

Eric H Davidson and Douglas H Erwin. Gene regulatory networks and the evolution of animal body plans. *Science*, 311(5762):796–800, 2006. 2

Justin M Fear, Luis G León-Novelo, Alison M Morse, Alison R Gerken, Kjong Van Lehmann, John Tower, Sergey V Nuzhdin, and Lauren M McIntyre. Buffering of genetic regulatory networks in drosophila melanogaster. *Genetics*, 203(3):1177–1190, 2016. 8

Anne Genissel, Lauren M McIntyre, Marta L Wayne, and Sergey V Nuzhdin. Cis and trans regulatory effects contribute to natural variation in transcriptome of drosophila melanogaster. *Molecular biology and evolution*, 25(1):101–110, 2008. 8

Dan Graur, Yichen Zheng, Nicholas Price, Ricardo BR Azevedo, Rebecca A Zufall, and Eran Elhaik. On the immortality of television sets:function in the human genome according to the evolution-free gospel of encode. *Genome biology and evolution*, 5(3):578–590, 2013. 3

Rita M Graze, Lauren M McIntyre, Bradley J Main, Marta L Wayne, and Sergey V Nuzhdin. Regulatory divergence in drosophila melanogaster and d. simulans, a genomewide analysis of allele-specific expression. *Genetics*, 183(2):547–561, 2009. 8

RM Graze, LL Novelo, V Amin, JM Fear, G Casella, SV Nuzhdin, and LM McIntyre. Allelic imbalance in drosophila hybrid heads: exons, isoforms, and evolution. *Molecular biology and evolution*, 29(6):1521–1532, 2012. 8

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. 3

Christopher T Harbison, D Benjamin Gordon, Tong Ihn Lee, Nicola J Rinaldi, Kenzie D Macisaac, Timothy W Danford, Nancy M Hannett, Jean-Bosco Tagne, David B Reynolds, Jane Yoo, et al. Transcriptional regulatory code of a eukaryotic genome. *Nature*, 431(7004):99–104, 2004. 8

GH Hardy. Mendelian proportions in a mixed population. *Science*, 28(706):49–50, 1908. 2

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. 3

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. 2

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

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. 2

Johannes Jaeger, Manfred Laubichler, and Werner Callebaut. The comet cometh: evolving developmental systems. *Biological theory*, 10(1):36–49, 2015. 2

R. E. 1930-(Rudolf Emil) Kalman, Peter L. Falb, and Michael A. Arbib. *Topics in mathematical system theory*. McGraw-Hill, New York, 1969. ISBN 0754321069. 1

Rudolf Emil Kalman. Mathematical description of linear dynamical systems. *J.S.I.A.M Control*, 1963. 3

Konstantin Kozlov, Svetlana Surkova, Ekaterina Myasnikova, John Reinitz, and Maria Samsonova. Modeling of gap gene expression in *Drosophila Kruppel* mutants. *PLoS Computational Biology*, 2012. 2

Konstantin Kozlov, Vitaly Gursky, Ivan Kulakovskiy, and Maria Samsonova. Sequnce-based model of gap gene regulatory network. *BMC Genomics*, 2014. 2

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. 2

Richard C Lewontin et al. *The genetic basis of evolutionary change*, volume 560, chapter The Structure of Evolutionary Genetics, pages 12–16. Columbia University Press New York, 1974. 2

Michael Lynch. The evolution of genetic networks by non-adaptive processes. *Nature Reviews Genetics*, 8(10):803–813, 2007a. 3, 7, 8

Michael Lynch. The frailty of adaptive hypotheses for the origins of organismal complexity. *Proceedings of the National Academy of Sciences*, 104(suppl 1):8597–8604, 2007b. 3

Michael Lynch and Kyle Hagner. Evolutionary meandering of intermolecular interactions along the drift barrier. *PNAS*, 2014. 8

Wenzhe Ma, Ala Trusina, Hana El-Samad, Wendell A Lim, and Chao Tang. Defining network topologies that can achieve biochemical adaptation. *Cell*, 138(4):760–773, 2009. 3

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. 3

Ron Milo, Shai Shen-Orr, Shalev Itzkovitz, Nadav Kashtan, Dmitri Chklovskii, and Uri Alon. Network motifs: simple building blocks of complex networks. *Science*, 298(5594):824–827, 2002. 3

Eric Mjolsness, David H Sharp, and John Reinitz. A connectionist model of development. *Journal of theoretical Biology*, 152(4):429–453, 1991. 2

Sergey V Nuzhdin, Marta L Wayne, Kristy L Harmon, and Lauren M McIntyre. Common pattern of evolution of gene expression level and protein sequence in drosophila. *Molecular biology and evolution*, 21(7):1308–1317, 2004. 8

Mihaela Pavlicev and Gunter P Wagner. A model of developmental evolution: selection, peliotropy and compensation. *Trends in Ecology & Evolution*, 2012. 2

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. 2

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. 2

Andrew B Stergachis, Shane Neph, Richard Sandstrom, Eric Haugen, Alex P Reynolds, Miaohua Zhang, Rachel Byron, Theresa Canfield, Sandra Stelhing-Sun, Kristen Lee, et al. Conservation of trans-acting circuitry during mammalian regulatory evolution. *Nature*, 515(7527):365–370, 2014. 3

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. 3

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

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. 3

Murat Tuğrul, Tiago Paixao, Nicholas H Barton, and Gašper Tkačik. Dynamics of transcription factor binding site evolution. *PLoS Genet*, 11(11):e1005639, 2015. 8

AJ Van der Schaft. Equivalence of dynamical systems by bisimulation. *IEEE transactions on automatic control*, 49(12):2160–2172, 2004. 3

Jeff Vierstra, Eric Rynes, Richard Sandstrom, Miaohua Zhang, Theresa Canfield, R Scott Hansen, Sandra Stehling-Sun, Peter J Sabo, Rachel Byron, Richard Humbert, et al. Mouse regulatory dna landscapes reveal global principles of cis-regulatory evolution. *Science*, 346(6212):1007–1012, 2014. 3

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. 2

Andreas Wagner. Does evolutionary plasticity evolve? *Evolution*, pages 1008–1023, 1996. 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. 3

Wilhelm Weinberg. Über vererbungsgesetze beim menschen. *Zeitschrift für induktive Abstammungs- und Vererbungslehre*, 1(1):440–460, 1908. 2

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. 2

Lotfi A Zadeh and Charles A Deoser. *Linear system theory*. Robert E. Krieger Publishing Company Huntington, 1976. 3