Gene regulatory network drift and speciation

Joshua S. Schiffman[†] and Peter L. Ralph[†‡]

[†]University of Southern California, Los Angeles, California

[‡]University of Oregon, Eugene, Oregon

Background

- Developmental systems drift: many different molecular pathways can yield identical phenotypes. For example, in *Drosophila melanogaster* and *D. simulans*, bristle patterning is identical, yet in hybrid crosses, bristle patterns are different, suggesting divergent molecular mechanisms[1].
- Here we model gene regulatory networks as linear dynamical systems.
- We analytically describe the set of all phenotypically equivalent gene regulatory network architectures.
- Evolution can explore this set, despite selective and environmental stasis.
- Using quantitative genetics, we show that over time, this neutral process can lead to rapid hybrid phenotypic divergence and incompatibility (~ on the order of N_e).

The set of all phenotypically identical gene networks of any size

$$\begin{aligned}\mathcal{A}_n(A_0) &= \{A : Ce^{At}B = Ce^{A_0t}B \text{ for } t \geq 0\} \\ &= \{A : CA^kB = CA_0^kB \text{ for } 1 \leq k \leq n-1\}\end{aligned}$$

We denote $\mathcal{A}_n(A_0)$ as the set of all n-dimensional gene network architectures equivalent to A_0, where A_0 is any linear gene network.

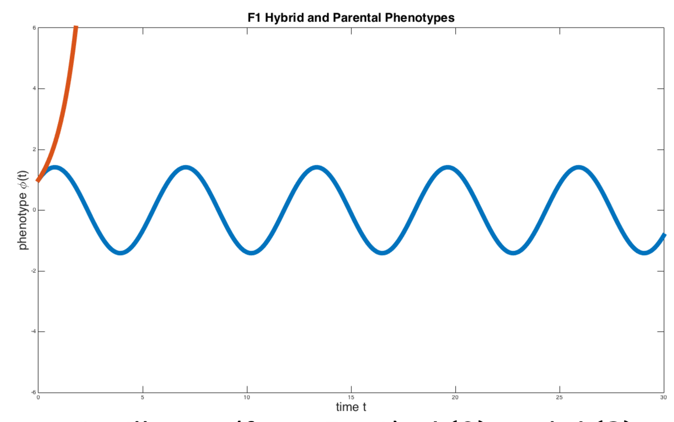
Gene regulatory networks as linear dynamical systems

$$\begin{cases} \dot{\kappa}(t) &= A\kappa(t) + Bu(t) \\ \phi(t) &= C\kappa(t) \end{cases}$$

- A: gene regulatory network ($n \times n$ matrix); rows are *cis*-regulatory modules.
- $u(t)$: environmental input at time t.
- B: how the organism processes its input ($n \times l$ matrix).
- $\kappa(t)$: all molecular concentrations at time t – the **kryptotype**.
- C: filtered molecular dynamics relevant to survival ($l \times n$ matrix).
- $\phi(t)$: molecular concentrations, visible to selection, at time t – the **phenotype**.

Neutral evolution leads to hybrid incompatibility

- Diploid F_1 hybrids are formed by averaging the two parental gene networks A and A'.
- F_2 hybrids are formed by first recombining genes (swapping rows between A and A'), then next, two gametes are chosen and averaged.
- Fitness is scored as a Gaussian function of phenotypic distance.

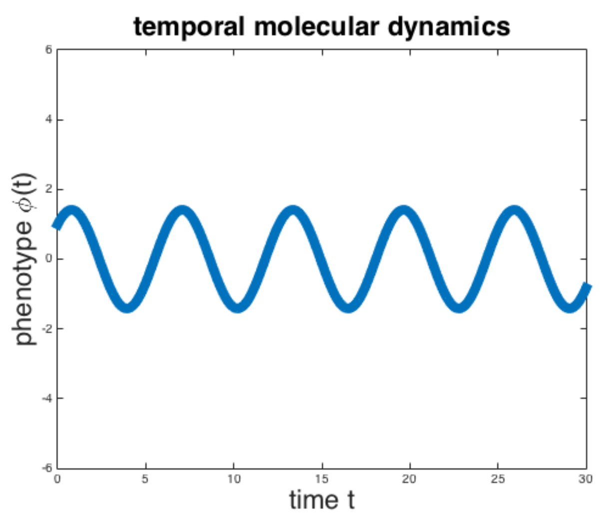


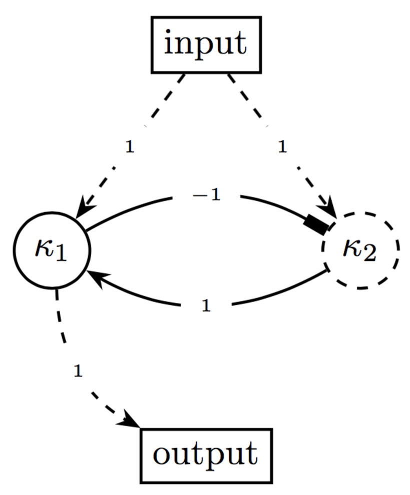
- Oscillators (from Ex. 1) $A(0)$ and $A(2)$ F_1s
- Parents (blue), hybrids (orange).
- Hybrid $\phi_{F_1}(t) = e^t$

Example 1: oscillating two-gene network

- A two-gene regulatory network with oscillating gene-1 expression.
- Environmental input is simply an impulse (a dirac delta).

$$\begin{aligned}\dot{\kappa}(t) &= \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\kappa(t) + \begin{bmatrix} 1 \\ 1 \end{bmatrix}u(t) \\ \phi(t) &= \begin{bmatrix} 1 & 0 \end{bmatrix}\kappa(t)\end{aligned}$$





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

Different network architectures produce identical phenotypes

$$\begin{aligned}\dot{\kappa}(t) &= \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\kappa(t) + \begin{bmatrix} 1 \\ 1 \end{bmatrix}u(t) \\ \phi(t) &= \begin{bmatrix} 1 & 0 \end{bmatrix}\kappa(t)\end{aligned}$$

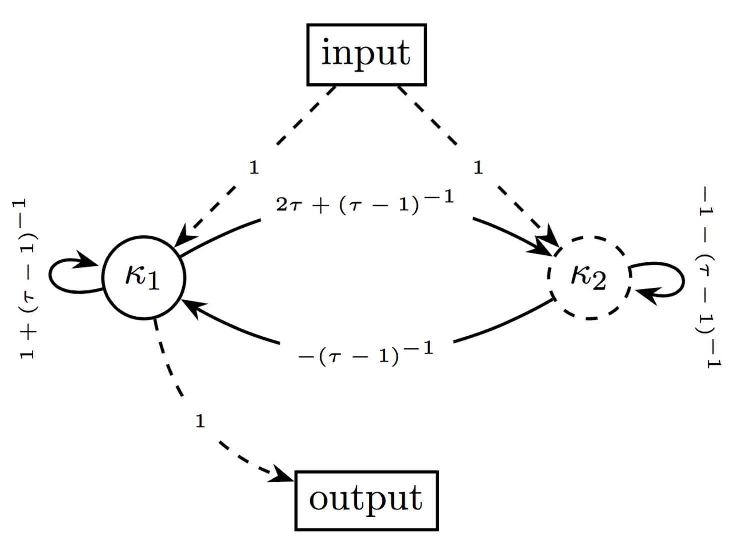
$$\begin{aligned}\dot{\widehat{\kappa}}(t) &= \begin{bmatrix} 2 & -1 \\ 5 & -2 \end{bmatrix}\widehat{\kappa}(t) + \begin{bmatrix} 1 \\ 1 \end{bmatrix}u(t) \\ \widehat{\phi}(t) &= \begin{bmatrix} 1 & 0 \end{bmatrix}\widehat{\kappa}(t)\end{aligned}$$

$$\phi(t) = \widehat{\phi}(t) = \sin(t) + \cos(t)$$

The set of all two-gene oscillators with identical phenotypes

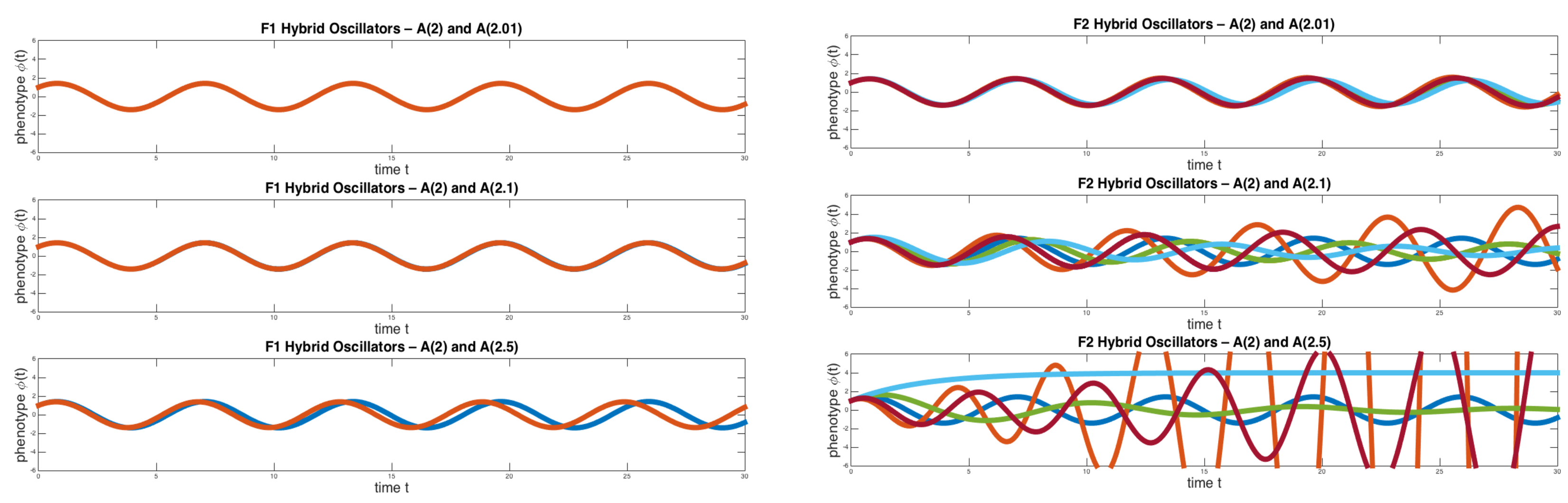
$$P(\tau) := \begin{bmatrix} 1 & 0 \\ \tau & 1-\tau \end{bmatrix} \qquad \tau \neq 1$$

$$A(\tau) := P(\tau)AP^{-1}(\tau)$$



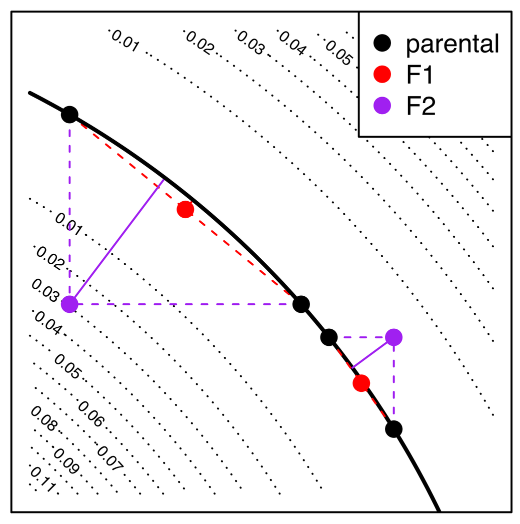
Here the set of all phenotypically identical two-gene oscillator networks is given by any coordinate change that preserves B and C. More generally, without preserving input or output matrices, or network size, this set is given by the Kalman decomposition[2].

Example 2: hybrid phenotypic divergence in a two-gene oscillator



Phenotypes diverge quadratically in F_2 and quartically in F_1 hybrids, with respect to change in network architecture.

Example of a phenotypic landscape. The dark **black** line is the phenotypically identical gene network space, and numbers are phenotypic divergence.



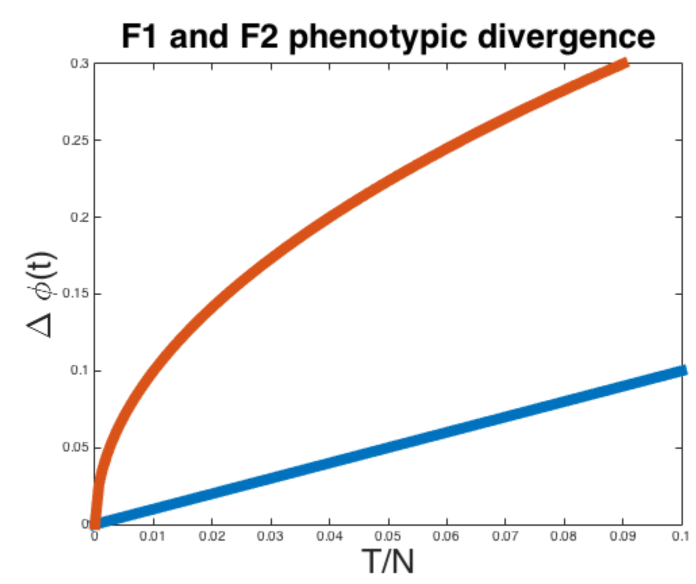
Expected phenotypic divergence rate in allopatric populations under static and identical selective and environmental pressures

Using a quantitative genetics model and a Gaussian fitness function, the phenotypic divergence (locally) is,

$\Delta\phi_{F_1}(t) \approx c_1 \frac{T}{N}$ in F_1 (blue),

$\Delta\phi_{F_2}(t) \approx c_2 \sqrt{\frac{T}{N}}$ in F_2 (orange)

formed by mating allopatric populations of size N isolated for T generations, where c_1 and c_2 are constants.



Future work

Apply model to study other evolutionary phenomena such as the necessity of complexity. Are gene regulatory networks Rube Goldberg machines? Is there a gene network ratchet?

References

1. JR True & ES Haag. Developmental system drift and flexibility in evolutionary trajectories. *Evolution & development*, 3(2):109–119, 2001.
2. RE Kalman. Mathematical description of linear dynamical systems. *J.S.I.A.M Control*, 1963.