Evolution and optimization

Paul Francois (McGill)

and V. Hakim (ENS Paris), F. Corson (ENS), M. Vergassola (Pasteur, UCSD)

Papers in Pubmed
Phenotypic models of development:

Development happens in time, progressive refinement of pattern
(genomes + ChIP-seq, RNA-seq tells us nothing about morphogenesis)

‘Geometry’ (& bifurcations) of dynamical systems define the phenotypic models

Can evolution alone predict the phenotypic properties of genetic networks (common to phyla)?

Phenotypic evolution (19th C Darwinism) by positive selection
Implications:
  What we see is what evolved quickly, survival of the fleetist.
  Phenotypic evolution convergent (molecular implementation contingent)
Setup

Fitness:
Formulated as quality of pattern, case by case, not reproductive fitness
Evolve features common to phyla, not particular species.

Representation:
Dynamical systems made from interaction of pseudo–biochem parts.
Impose simple dynamics i.e., Morse–Smale
Change network & parameters, only neutral and fitness incr. changes.
Time evolve network + boundary conditions, at \( T=\infty \), evaluate fitness.

Problems:
How much does the fitness matter?
Mutation rates matter? (alleviated if dominated by positive selection)

Non problems:
All systems small; no issue of *complexity* with ‘N’.
No biological reason to insist learn all members of *concept class*.

Fitness= acuity via physical optics, shape and refractive index change,
Fitness for embryonic patterning (1)

Body axes Cartesian: AP x DV

Development via pattern refinement

Selector gene hypothesis:
Define compartments/segments, tracks cell lineage, cell autonomous

‘Morphology’ -> network that positions the selector genes
Methods

Evolve gene networks via mutation–selection of both network topology, parameters, and ‘outputs’ (from which fitness calculated) (~ simulated annealing, no recombination... CP)

Embryo a line of ‘cells’

Network functions identically in all cells, which differ only in exposure to morphogen G (external protein whose spacial profile determines fate), no direct cell–cell communication in example here.

Interactions either activate & **add**, or repress & **multiply**

eg A auto–activates, repressed by R₁ R₂  (dropping *csts*), G(time)

\[
\dot{A} = \max(G(t), \frac{A^{n_1}}{1 + A^{n_1}}) \frac{1}{1 + R_1^{n_2}} \frac{1}{1 + R_2^{n_3}} - A
\]

\[
= \max(\text{activators}) \times \prod(\text{repressors}) - \text{degradation}
\]
Fitness for embryonic patterning (1Dim)

Require for fitness:

1. Assign a number to any collection of selector genes $C_i(x)$
2. Max. diversity... many selector genes expressed in embryo
3. Min. diversity for given $x$... (unique fate)
4. Smooth function that rewards a little bit of pattern
Fitness as mutual entropy:

\[ P(i \mid x) = \frac{C_i(x)}{\sum_i C_i(x)} \] (only relative concentrations matter)
\[ P(x) = \frac{1}{L} \] (uniform probability on cells)
\[ P(i, x) = P(x) \cdot P(i \mid x) \]

Fitness favors:
1. ‘Max. diversity’ \(\rightarrow\) Max entropy, \(S_1\), of \(P(i): \left(-\sum_i P(i) \log(P(i)\right)\)
2. ‘Min. diversity given \(x\)’ \(\rightarrow\) Min entropy, \(S_2\) of \(P(i \mid x)\): 

How to combine 2 terms?
3. Assume gene duplication neutral \(\rightarrow\)
   ‘fitness’ = \(-S_1 + S_2 = -\) mutual information \((i, x)\). \(C_i \iff x\)
   (best fitness ~ free energy is most negative)

For \(N\) selector genes fitness \(\geq -\log(N)\).
Mutual entropy fitness

F = -log 1.5

F = -log 2

F = -log 3

F = -log 2.3
Networks for static ‘morphogen’

Morp On→Off

Fitness(generation) = \ln(7)

Final network

- Morpho
- Network gene
- Selector

activate
repress

Final pattern, selector_i(x)
Time (development) dynamics in evolved static morphogen network

![Diagram of morphogen network with nodes labeled as Morpho, Network gene, and Selector. The graph shows a concentration vs. position plot.](image)
Properties of Networks static gradient

Networks ‘cell autonomous’ (no communication between cells) → morphogen defines cell position.

Morphogen disappears → multi-stability → sharp boundaries & only need repression between ~adjacent domains

Multi-stability → order of gene expression matters & numbers determine final state.

Morphogen sets anterior boundaries, repression sets posterior boundaries → statistical char. of evolved networks.
Topology $\neq$ function

Figure S1. Network evolved under control of static gradient, Posterior Index: 5, Anterior Index: 0, Fitness: $1.54 \times \log(4.7)$

Figure S3. Network evolved under control of static gradient, Posterior Index: 8, Anterior Index: 2, Fitness: $1.62 \times \log(5.1)$

Figure S4. Network evolved under control of static gradient, Posterior Index: 8, Anterior Index: 2, Fitness: $1.7 \times \log(5.47)$

Figure S5. Network evolved under control of static gradient, Posterior Index: 13, Anterior Index: 1, Fitness: $1.79 \times \log(6)$
Anterior–Posterior patterning

Hox genes conserved in bilaterians

Define coarse AP coordinates

Cellular “Zip code” controls master regulatory genes

Biochem of regulation very complex, but simple phenomenology
Mouse vertebrae reflect Hox territories

2.2. Hox expression

The earliest indication that vertebrate Hox genes might play a role in vertebrate axial patterning was obtained by in situ hybridization analyses. Hox genes are expressed from 30 to 50 in the clusters, with the earliest genes expressed in the posterior primitive streak at late streak stages, and more 50 genes expressed at progressively later stages (Dressler and Gruss, 1989; Duboule and Dolle, 1989; Gaunt, 1991; Gaunt and Strachan, 1996; Gaunt et al., 1986, 1990; Graham et al., 1989; Izpisua-Belmonte et al., 1991). This temporal control of Hox expression onset, coupled with growth and elongation of the embryo, results in spatially graded anterior boundaries of expression where 30 genes (Hox1 and Hox2) display anterior expression limits in the hindbrain region of the embryo and increasingly 50 genes demonstrate increasingly posterior limits.

Figure 9.2 A lateral view of an E18.5 mouse skeleton stained with Alcian blue and alizarin red; anterior to the left, posterior to the right. Circumferentially around the edges of the panel, the individual vertebral elements are pictures, beginning with the first cervical vertebra, the atlas, in the top, left side with the elements in order, clockwise. C, cervical, T, thoracic, L, lumbar, S, sacral and Cd, caudal; numbers reflect their position in the skeleton. (Only 7 of the approximately 30 caudal vertebrae are shown).

DM Wellik 2009
Phenomenology of Hox expression

1. Spacial colinearity: 3’ to 5’ genome order follows A to P expr.
2. Temporal colinearity: (vertebrates) temporal order follows A to P
3. Posterior prevalence rule: most posterior Hox gene imposes fate on all anterior genes

Hox mutation haltere→wing

Hox expression A to P

\[
\begin{align*}
\text{haltere} & \quad \text{Hox3} \\
\text{wing} & \quad \text{Hox2} \\
\text{Hox1} &
\end{align*}
\]

\[
\begin{align*}
\text{wing} & \quad \text{Hox3} \\
\text{mutate} & \quad \text{Hox1} \\
\text{wing} = (1 \text{ AND NOT}(2,3..)) \\
\text{haltere} = (2 \text{ AND NOT}(3,..))
\end{align*}
\]
Xenopus development (2)

1.2mm egg, 7hrs stage 9 4000+ cells; 17hrs stage 15; 40hrs stage 32 @23C
Gastrulation of Xenopus

1.2mm egg

5 hrs fertilization to Movie0
4000+ cells

17hrs @23C Movie

[Image of a fertilized egg with labels for Anterior, Dorsal view, and Posterior]
Patterning a field of cells: AP growth

Model ‘patterning during growth’ as sliding morphogen that marks boundary between growth zone and patterned tissue.

Hox expression marked as colors. Temporal sequence of expression on equator → spacial domains AP

‘organizer’ is point where converging equator → extending AP morphogen step ~ organizer

Wacher 2004
Sliding morphogen (2)

Time development of final evolved network

only triangles enter fitness

Morphogen
Temporal colinearity

Temporal colinearity: Hox(time) fixed posterior cell --> Anterior–Posterior progression
Anterior Homeotic Mutation (2)
Properties of Networks with Sliding Gradient

Recall static morphogen: anterior boundaries positioned from morphogen. Analogue for sliding gradient?

Position == time exposed to morphogen: ‘Timer’ gene 3 converts time in morphogen to morphogen level + cell autonomy. Static Morph <-> Sliding Morph.

Good for growth control, change all rates get same pattern (Deschamps et al timer ~ CAUDAL, CDX2)

Hox phenomenology: temporal colinearity, anterior homeotic mutation

Evolution of long from short germ band insects.
Phylogeny of short (seq) and long germ insects (seq ~ vertebrates, pattern during growth)

<table>
<thead>
<tr>
<th>order</th>
<th>species</th>
<th>hemi or holo metamorphosis</th>
<th>extended syncytial blastoderm</th>
<th>sequential or long germ segmentation</th>
<th>anterior localization of maternal mRNAs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Orthoptera</td>
<td><em>Schistocerca</em> sp. <em>Gryllus bimaculatus</em></td>
<td>hemi</td>
<td>no</td>
<td>seq</td>
<td>?</td>
</tr>
<tr>
<td>Hemiptera</td>
<td><em>Oncopeltus fasciatus</em></td>
<td>hemi</td>
<td>?</td>
<td>seq</td>
<td>?</td>
</tr>
<tr>
<td>Hymenoptera</td>
<td><em>Nasonia vitripennis</em></td>
<td>holo</td>
<td>yes</td>
<td>long</td>
<td>yes</td>
</tr>
<tr>
<td></td>
<td><em>Bracon hebetor</em></td>
<td></td>
<td>yes</td>
<td>long</td>
<td>?</td>
</tr>
<tr>
<td></td>
<td><em>Aphidius ervi</em></td>
<td></td>
<td>no</td>
<td>seq</td>
<td>?</td>
</tr>
<tr>
<td></td>
<td><em>Apis mellifera</em></td>
<td></td>
<td>yes</td>
<td>long*</td>
<td>?</td>
</tr>
<tr>
<td>Coleoptera</td>
<td><em>Tribolium castaneum</em></td>
<td>holo</td>
<td>yes</td>
<td>seq*</td>
<td>yes</td>
</tr>
<tr>
<td></td>
<td><em>Callosobruchus maculatus</em></td>
<td></td>
<td></td>
<td></td>
<td>?</td>
</tr>
<tr>
<td>Lepidoptera</td>
<td><em>Bombyx mori</em></td>
<td>holo</td>
<td>no</td>
<td>seq*</td>
<td>?</td>
</tr>
<tr>
<td></td>
<td><em>Manduca sexta</em></td>
<td></td>
<td>?</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Drosophila melanogaster</em></td>
<td>holo</td>
<td>yes</td>
<td>long</td>
<td>yes</td>
</tr>
</tbody>
</table>

AD Peel, Phil. Trans. R. Soc 2008

Short to long germ: *timer* gene \(\rightarrow\) static morphogen. Down stream network invariant

Insect evolution focuses on segmentation, but Hox supplies identity.
Other systems evolved:

- Clocks and bistable systems: (Francois & Hakim PNAS 2004)
- AP–Hox patterning (Francois & ES, Development 2010)
- Temperature compensated clocks that entrain (Francois & ES, PLoS Comp Bio)
- Networks that take a spatial derivative of transient morphogen
- Fit genes to topology (Corson & ES PNAS 2012)

A few other applications in brief
T Cell Activation

T cell receptors will respond to a few molecules of agonist and ignore a $> 10^4$ higher concentration of ‘self’ proteins, based on a 3-4x slower off rate from the receptor. (Kinetic proof reading will not explain this).


System also evolved by Lalanne & Francois PRL 2013 (see also Francois etal PNAS 2013)
Optimal decision theory (Explore–Exploit)

1. Given a stream of data from distribution A or B, what is minimum average decision time to identify the source for a given error rate, and what is the algorithm that realizes it?? (Wald 1945)

2. A stream of data changes from type A to type B at an unknown time T. What is the minimum average time lag in detecting the change point, for a prescribed false positive rate??

Plausible constraint on sensory systems, from cells in an embryo to higher cognition, decision speed matters.

Refs:
Neural MN Shadlen ~2006
Cellular, Kobayashi 2010;
Vergassola & EDS 2013, Simple biochemical networks can optimally solve 1 & 2 and the parameters fit via local search.
‘Saddle points’ or last common ancestor (or how to turn a fly into a mosquito)

gap genes regulate *eve*. gap genes move, *eve* fixed and essential
Goltsev *Dev. Bio*. 2004

Engrailed (and wg) mark segment boundaries
N H Patel *Development Suppl* 201-207 1994
Are there any biophysical principles such as dynamical behavior that control where/when certain pathways used? Could evolution simulations define discrete dynamical types?

(Brivanlou and Darnell 2002, Science)
Characteristics of evolved models

• Close to dynamical system picture, evolve topology of flow, not genes → visualize minimal parameter description (→ genes to be fit). Evolution as cascade of bifurcations.

• Network and parameters evolve together, de novo fitting of all parameters in final network could be hard.

• Networks work by sloppy confluence of opposing activities; with tuned rates; no time scale separation ≠ 19th C applied math. **BUT** simple in that parameters follow by gradient search.

• Evolved models not obvious, like genetic screen

• Relevance to experiment, hi level (static ↔ dynam morpho), lo level (fit parameters)

• 19th C Darwinism → grad search, Useful engineering principle for specific systems.
The End