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:

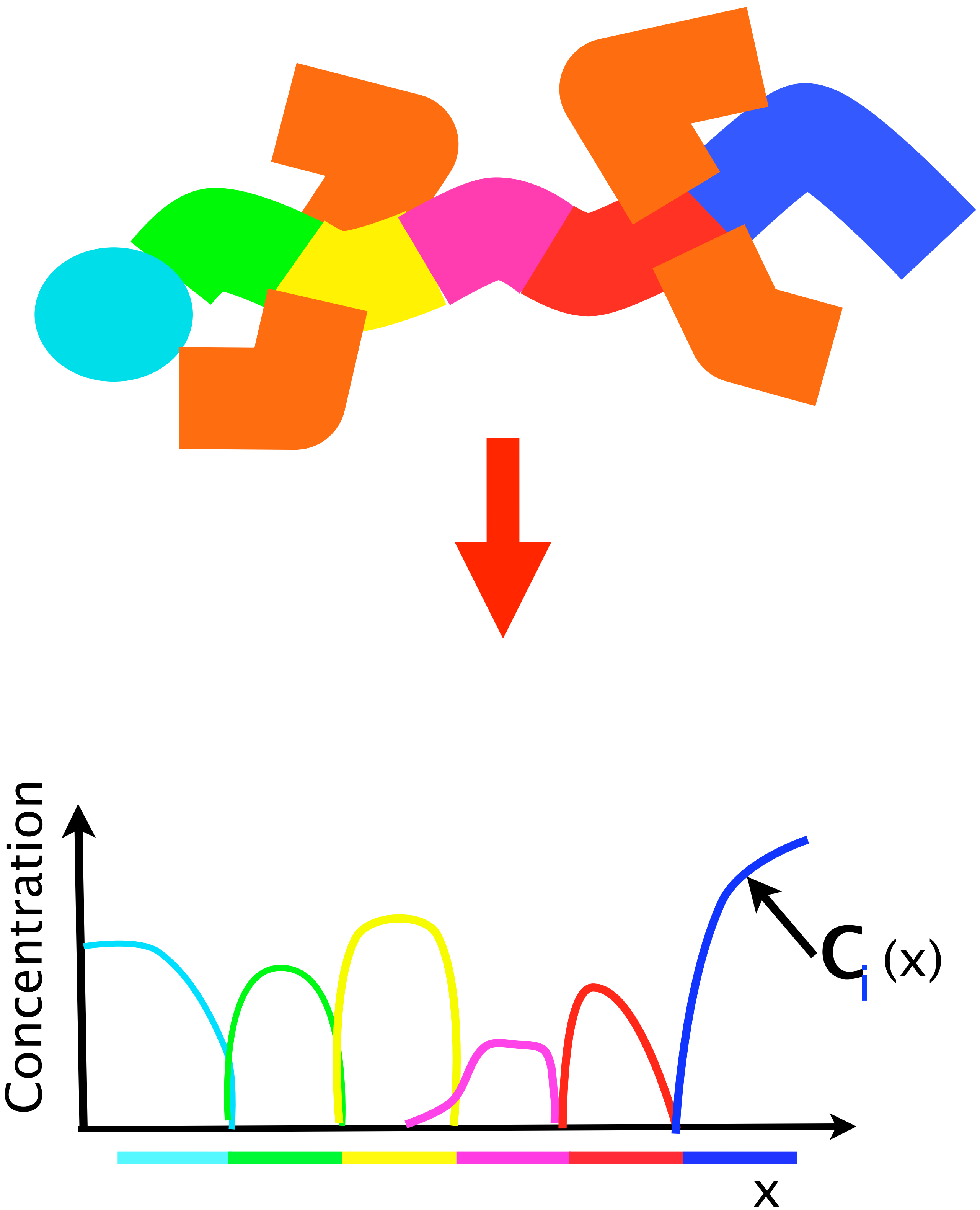
How much does the fitness matter?
Mutation rates matter? (alleviated if dominated by positive selection)

Non problems:

All systems small; no issue of *complexity* with 'N'.
No biological reason to insist learn all members of *concept class*.

Ex: Nilsson&Pelger 1994, "~Evolution eye" *Quant Genetics* (Barton).
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

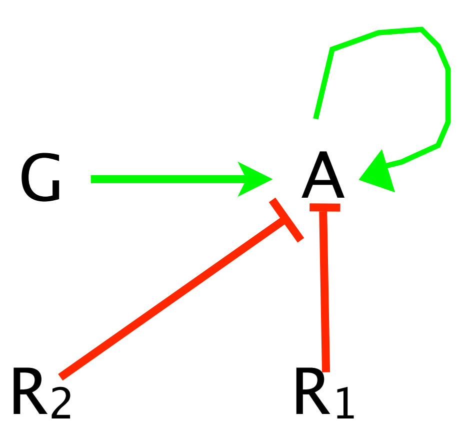
Evolve gene networks via mutation-selection of both network topology, parameters, and 'outputs' (from which fitness calculated)
(~ simulated annealing, no recombination... CP)

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 R_1 R_2 (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(activators) *∏(repressors) – 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:

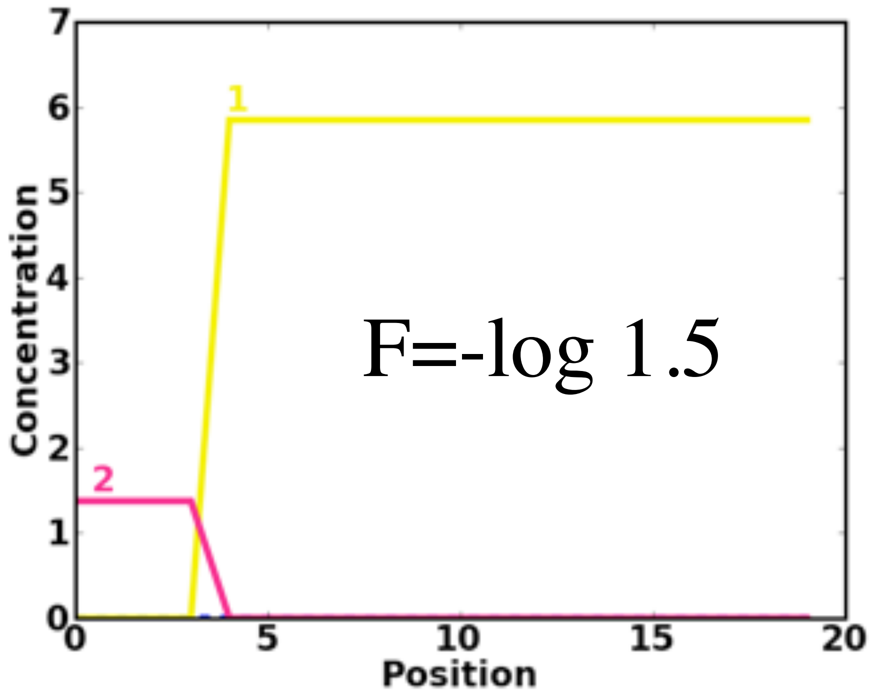
$P(i \mid x) = C_i(x) / \sum_i C_i(x)$ (only relative concentrations matter)
$P(x) = 1/L$ (uniform probability on cells)
$P(i,x) = P(x)*P(i \mid x)$

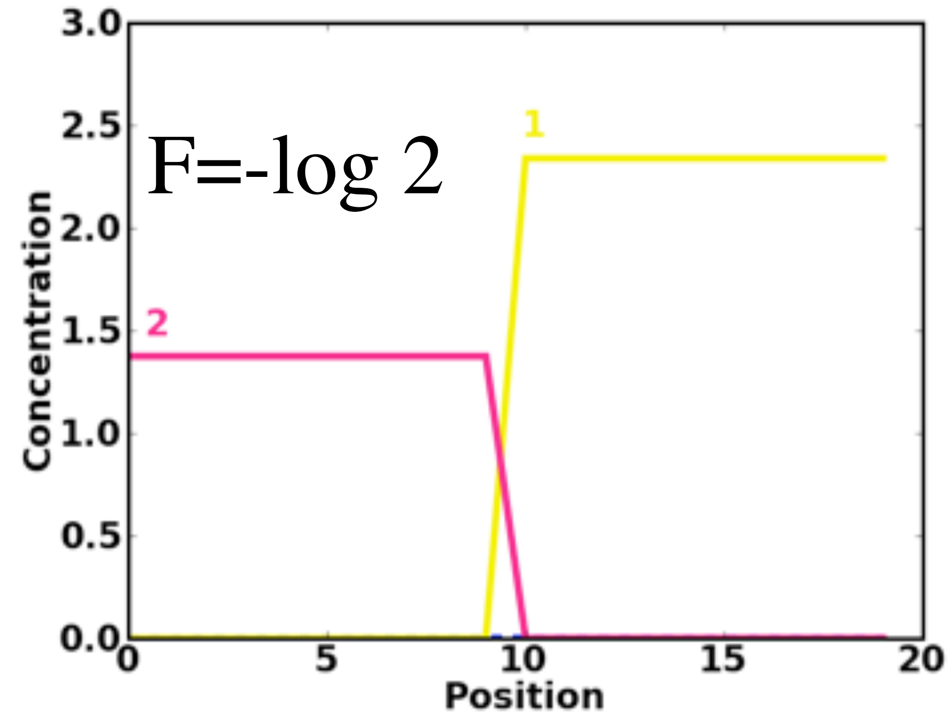
Fitness favors:
1. 'Max. diversity' -> Max entropy, $\mathbf{S_1}$, of P(i): $(-\sum_i P(i) \, Log(P(i))$)
2. 'Min. diversity given x' -> Min entropy, $\mathbf{S_2}$ of $P(i \mid x)$:

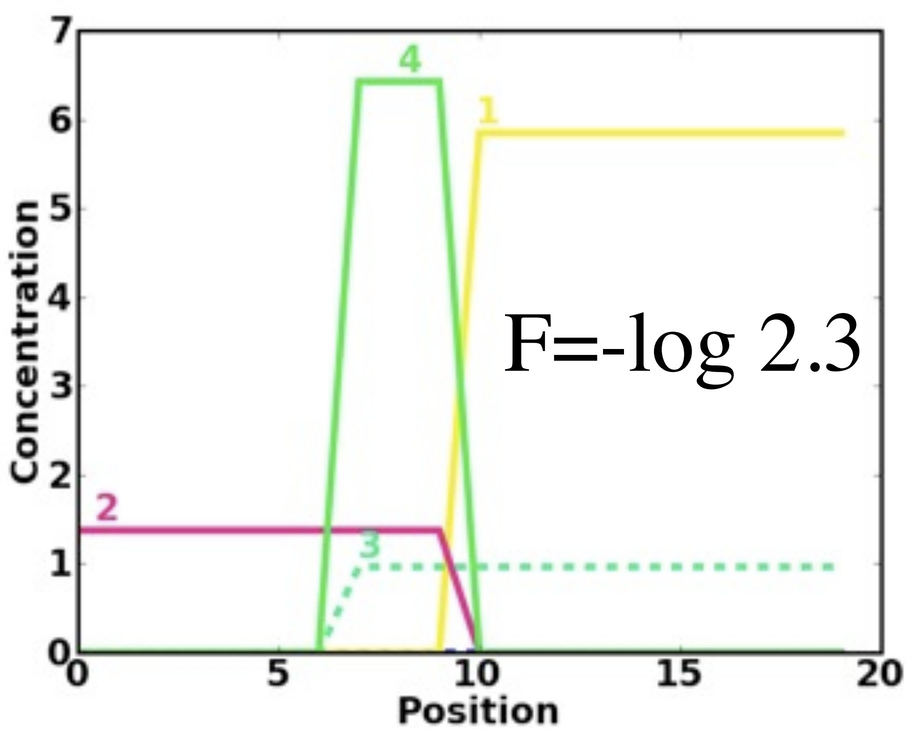
How to combine 2 terms?
3. Assume gene duplication neutral ->
'fitness' = $-\mathbf{S_1} + \mathbf{S_2}$ = – mutual information (i, x). C_i <=> x
(best fitness ~ free energy is most negative)

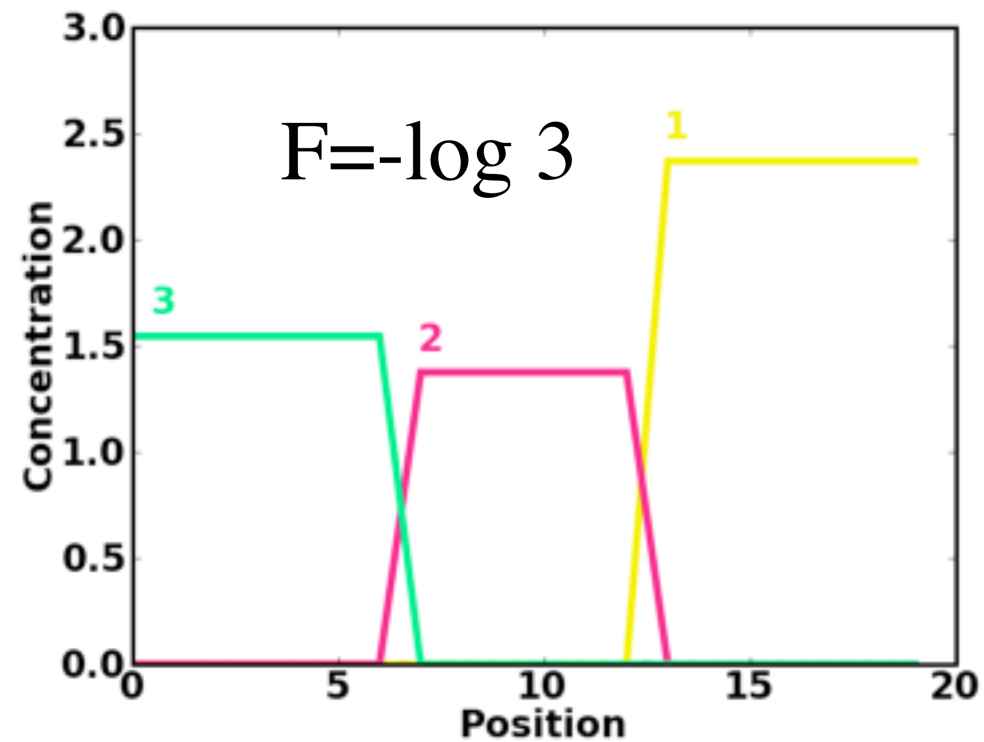
For N selector genes fitness ≥ –log(N).

Mutual entropy fitness

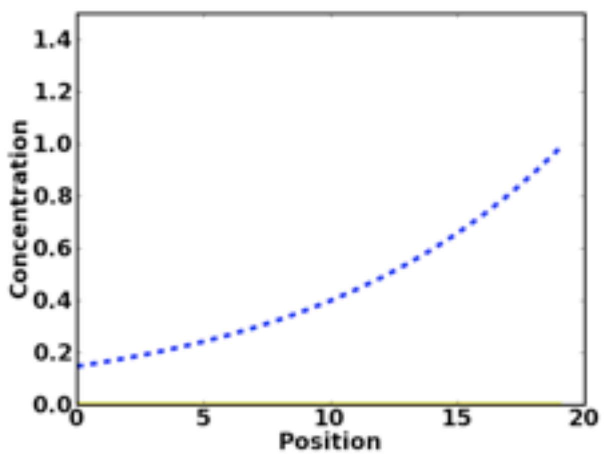




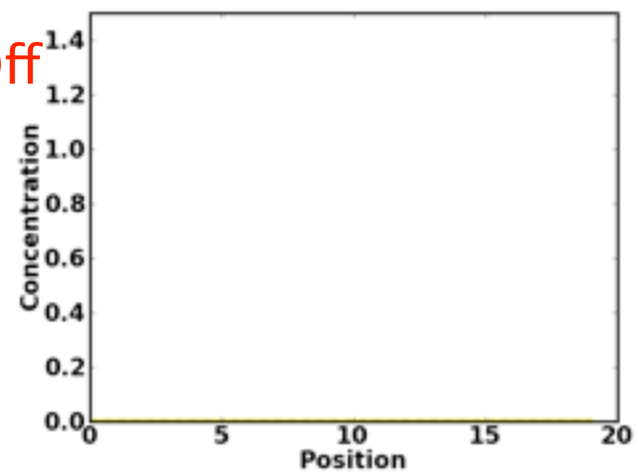


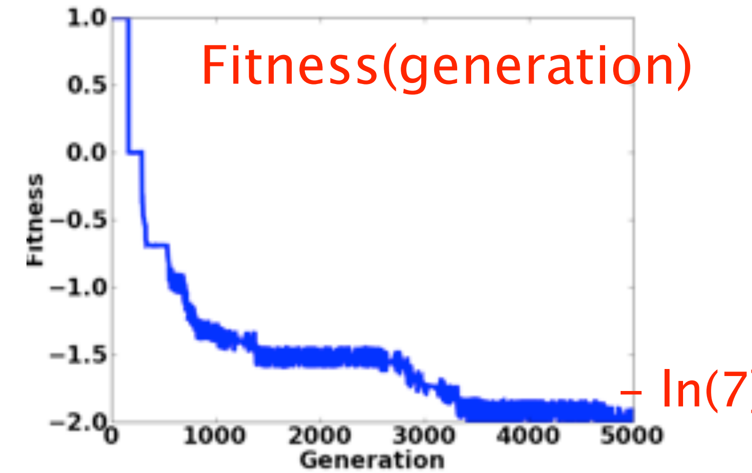


Networks for static ‘morphogen’

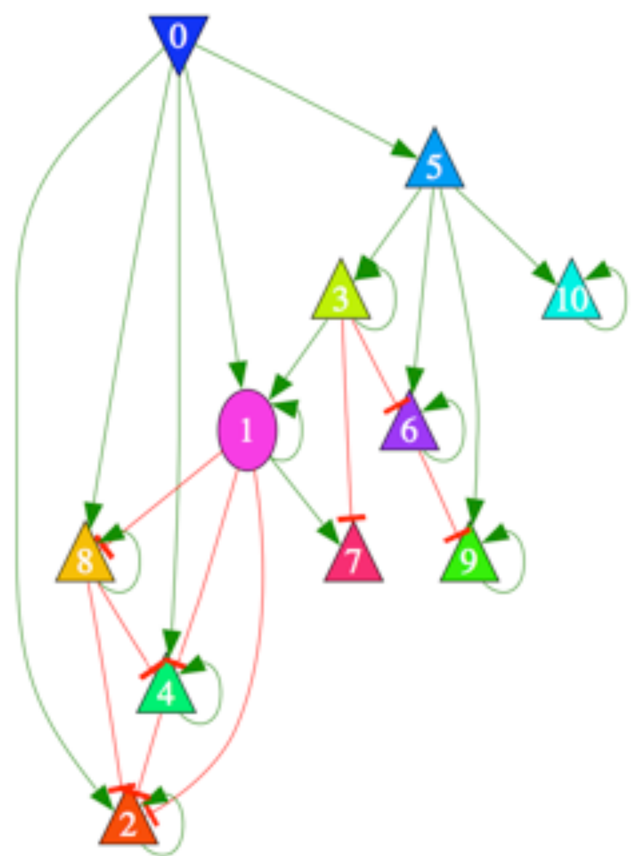


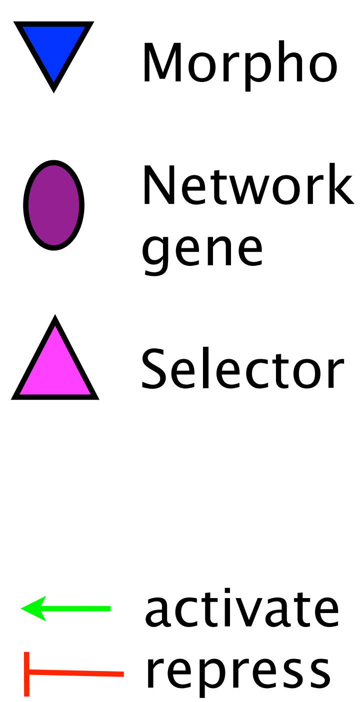
Morp On->Off



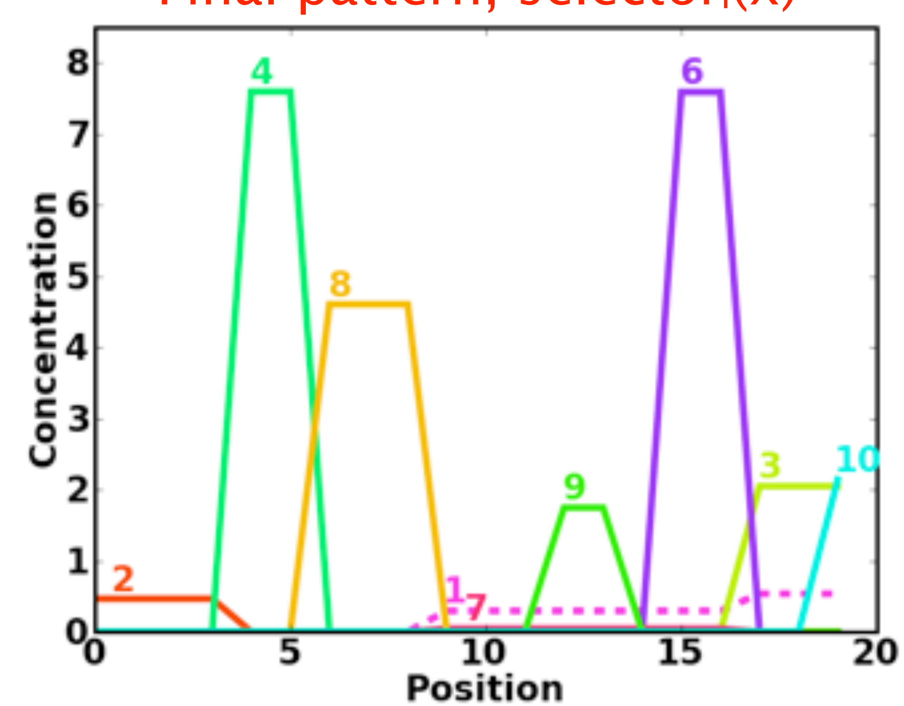


Final network

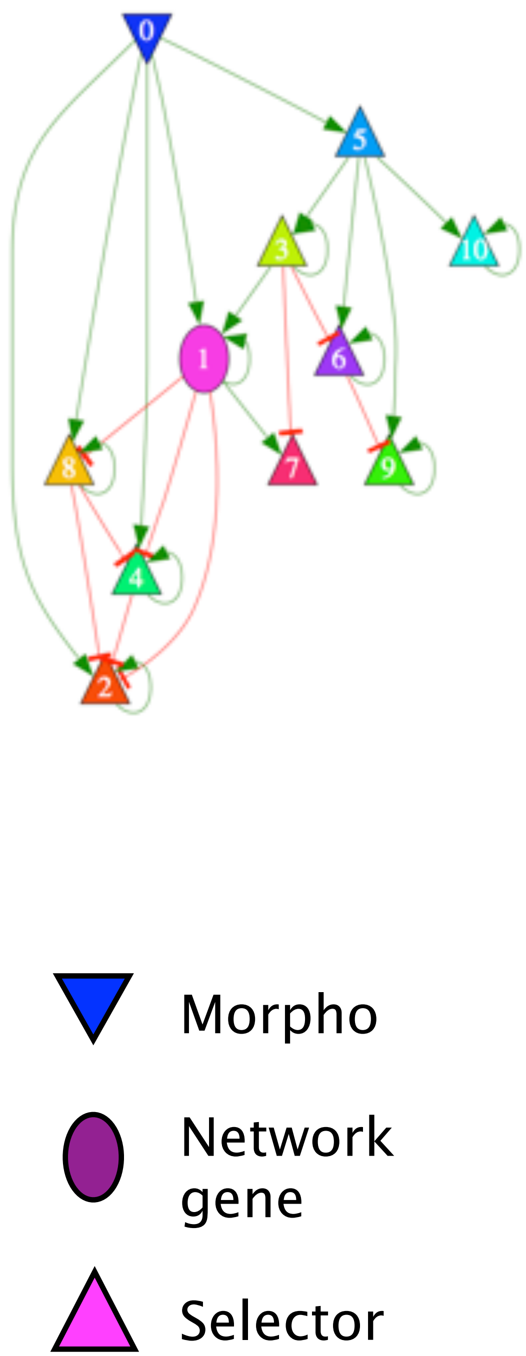


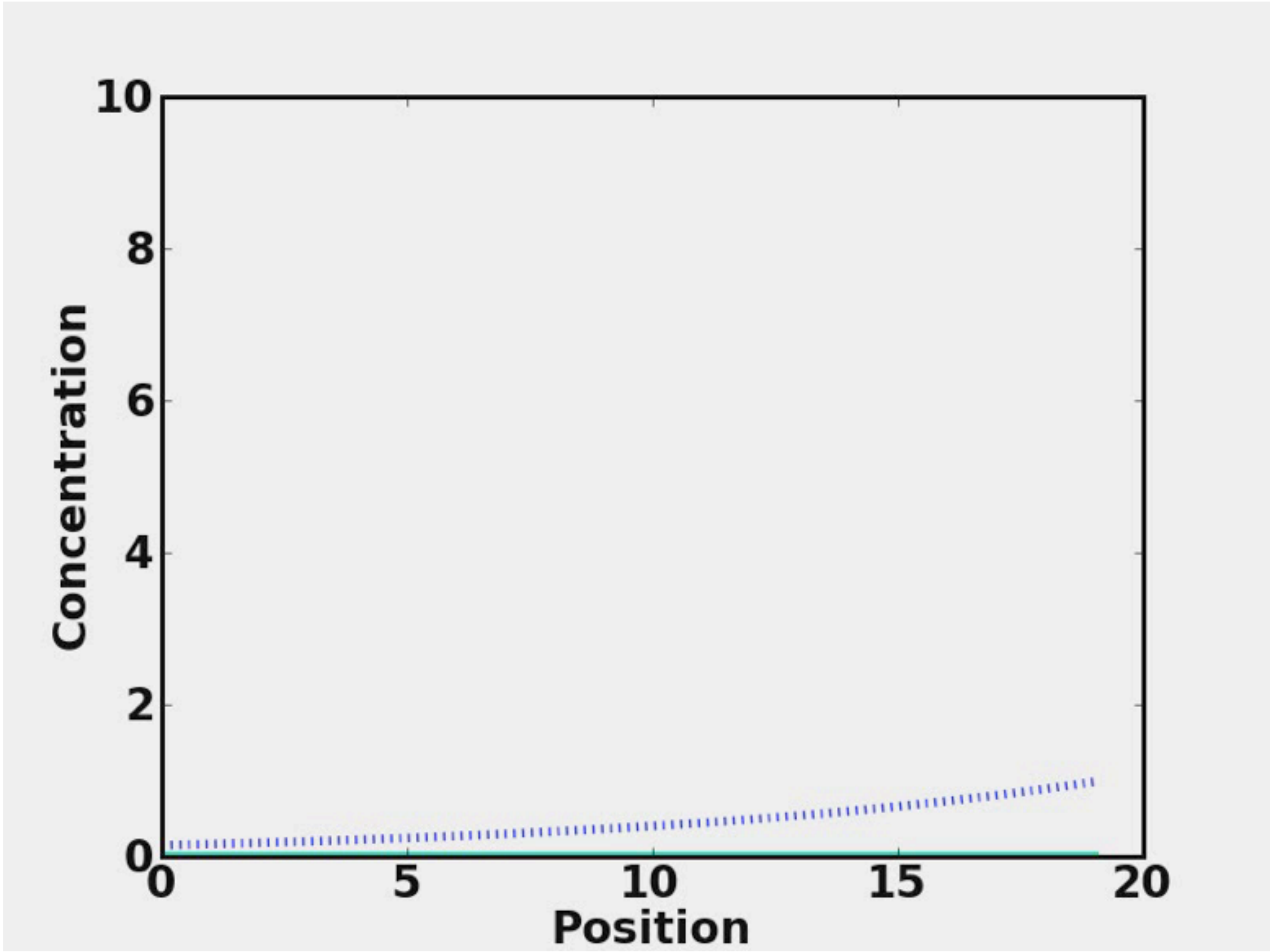


Final pattern, $selector_i(x)$

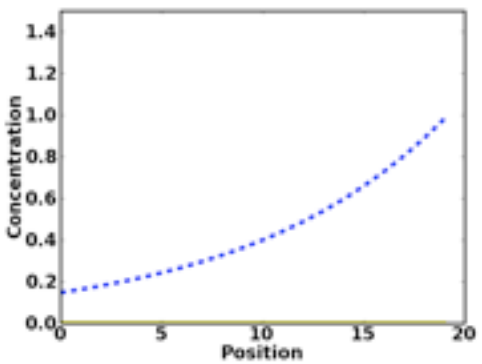


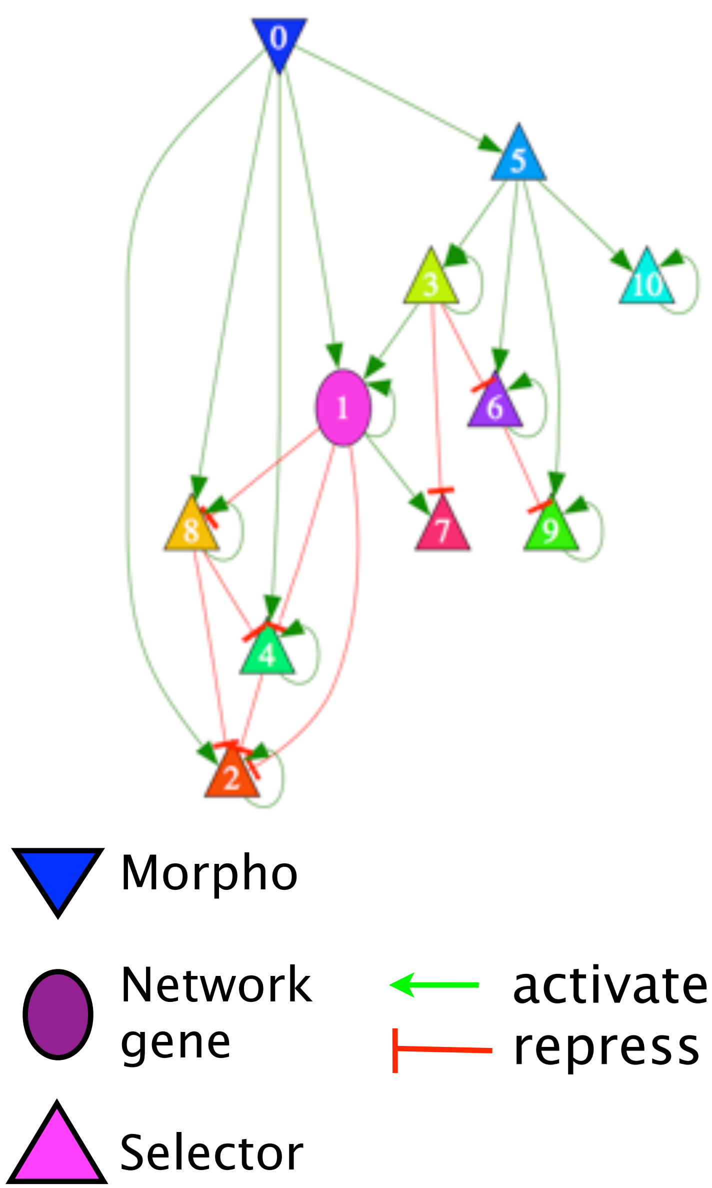
Time (development) dynamics in evolved static morphogen network

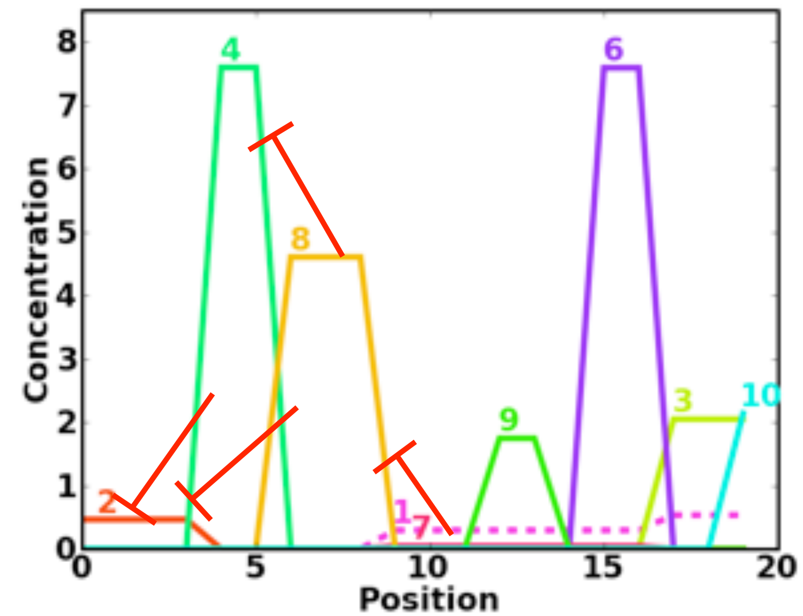




Properties of Networks static gradient







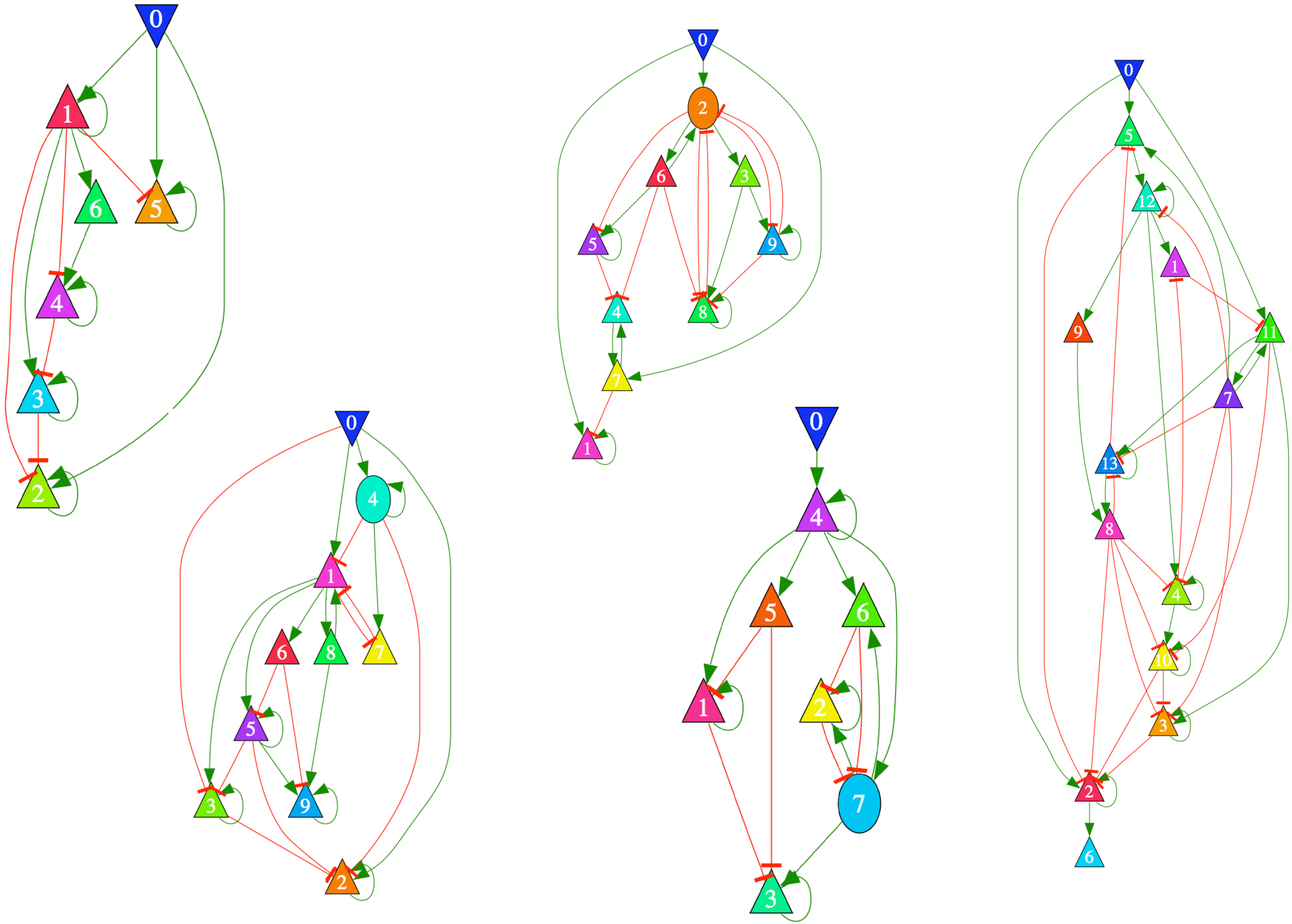
Networks 'cell autonomous' (no communication between cells)-> morphogen defines cell position.

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

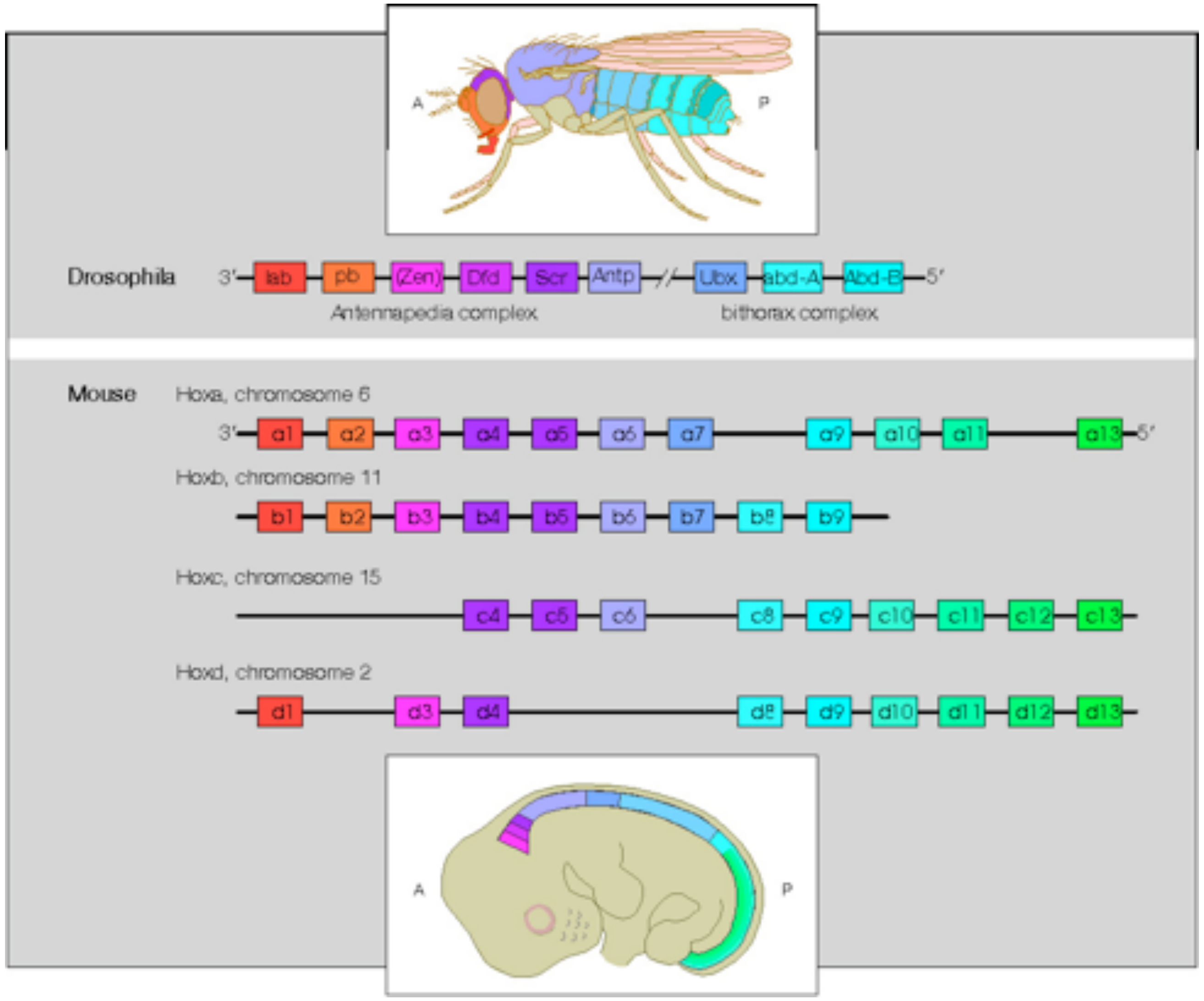
Multi-stability -> order of gene expression matters & numbers determine final state.

Morphogen sets anterior boundaries, repression sets posterior boundaries -> statistical char. of evolved networks.

Topology ≠ function



Anterior–Posterior patterning



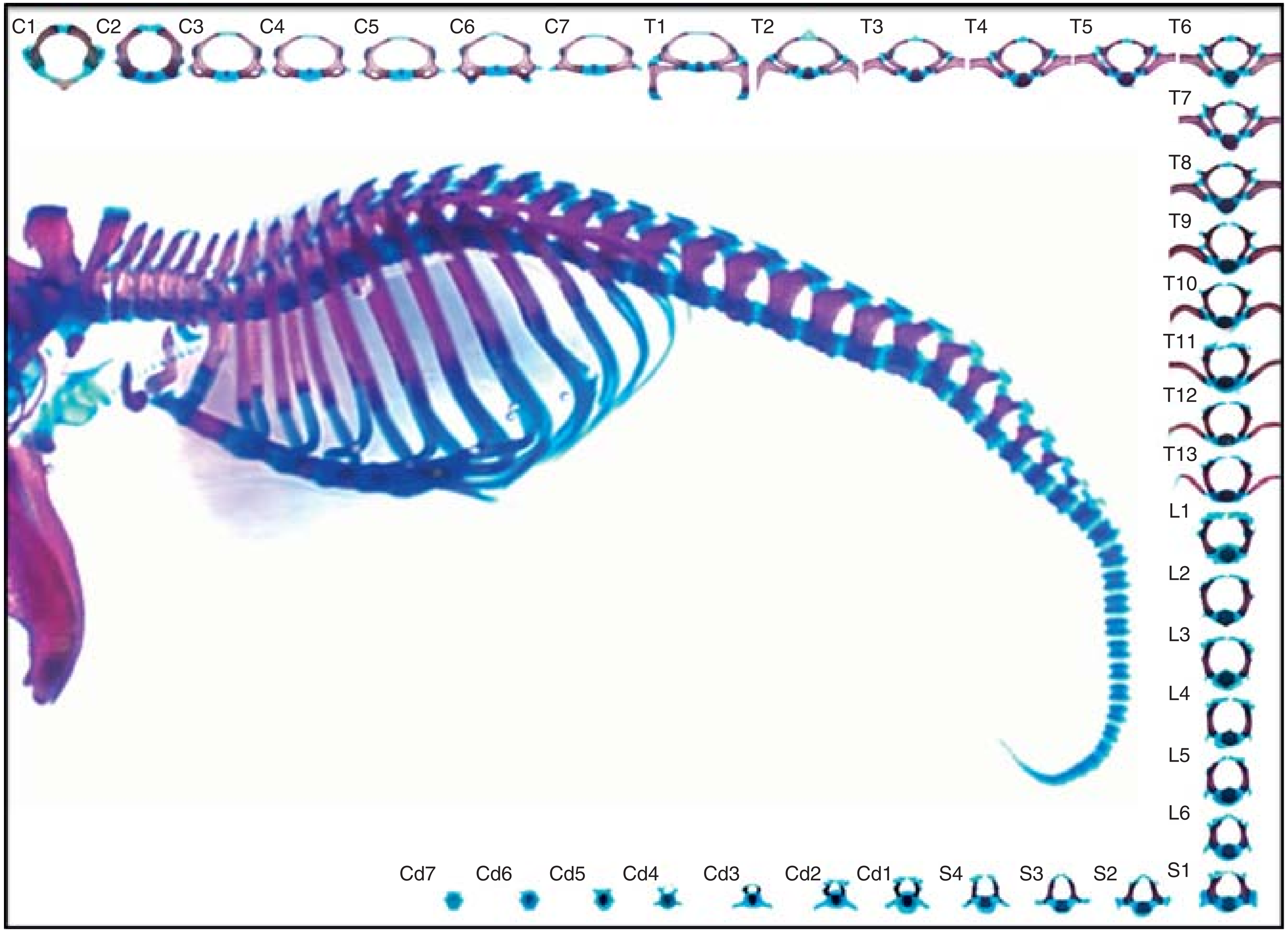
Hox genes conserved in bilaterians

Define coarse AP coordinates

Cellular “Zip code” controls master regulatory genes

Biochem of regulation very complex, but simple phenomenology

Mouse vertebrae reflect Hox territories



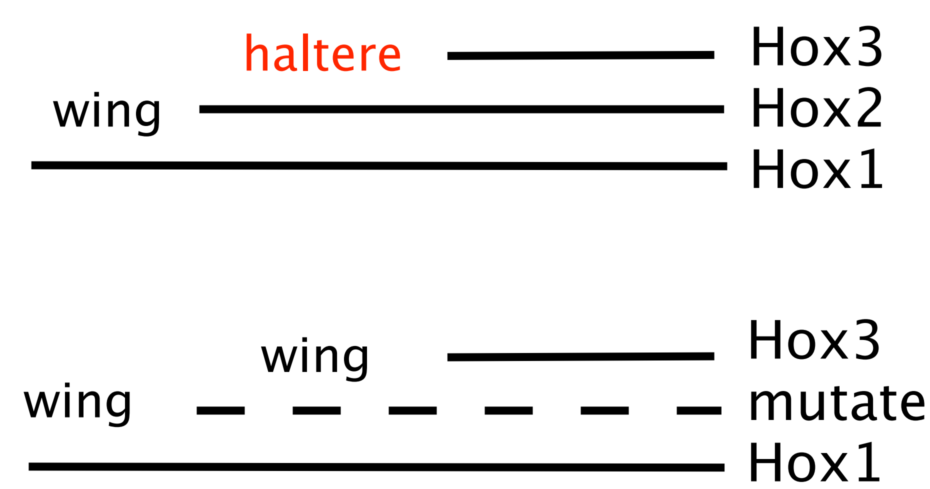
DM Wellik 2009

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



wing = (1 AND NOT(2,3..))
haltere = (2 AND NOT(3,..))

Xenopus development (2)

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

Gastrulation of Xenopus

1.2mm egg

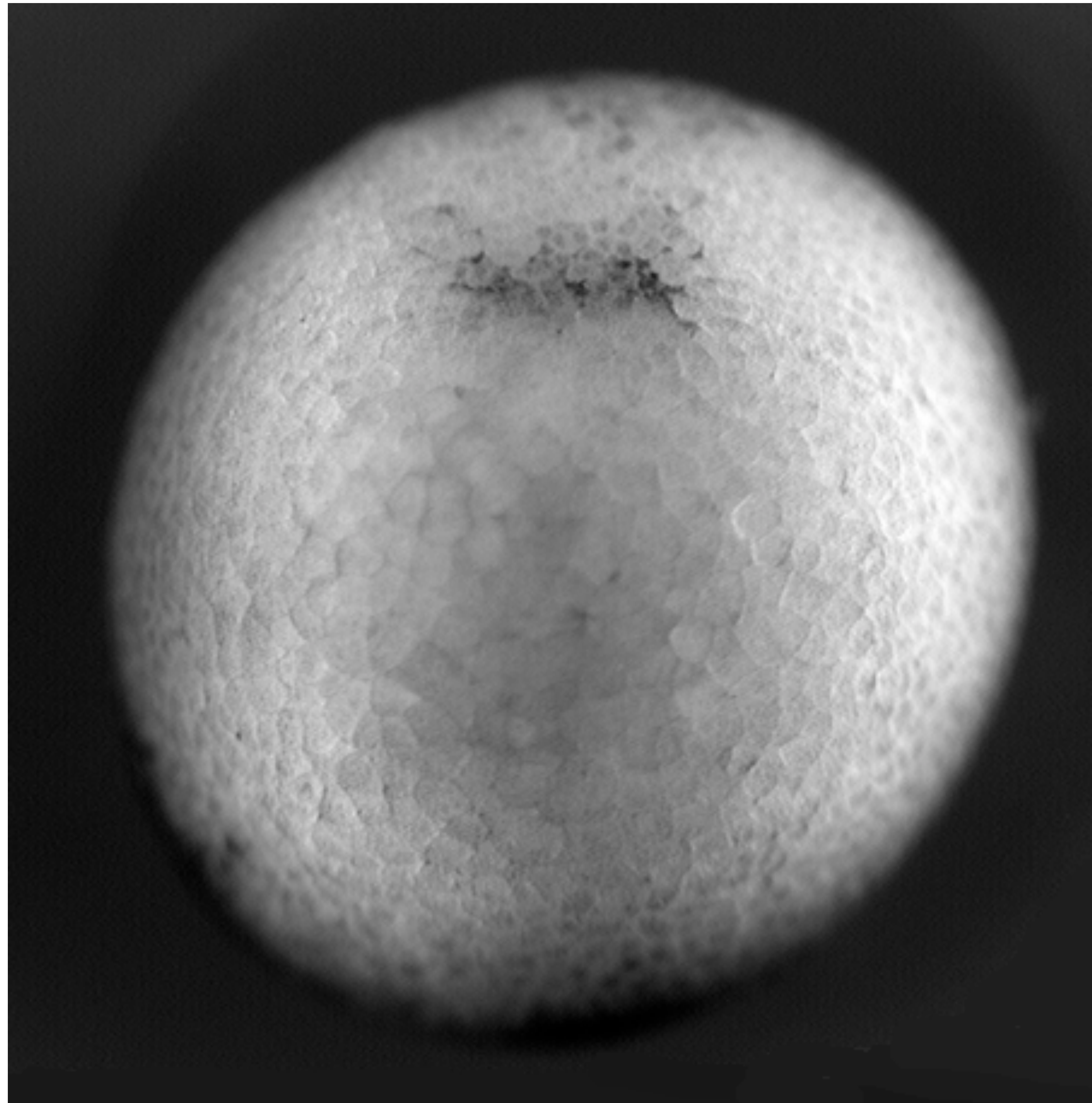
5 hrs fertilization to Movie0
4000+ cells

17hrs @23C Movie

Anterior

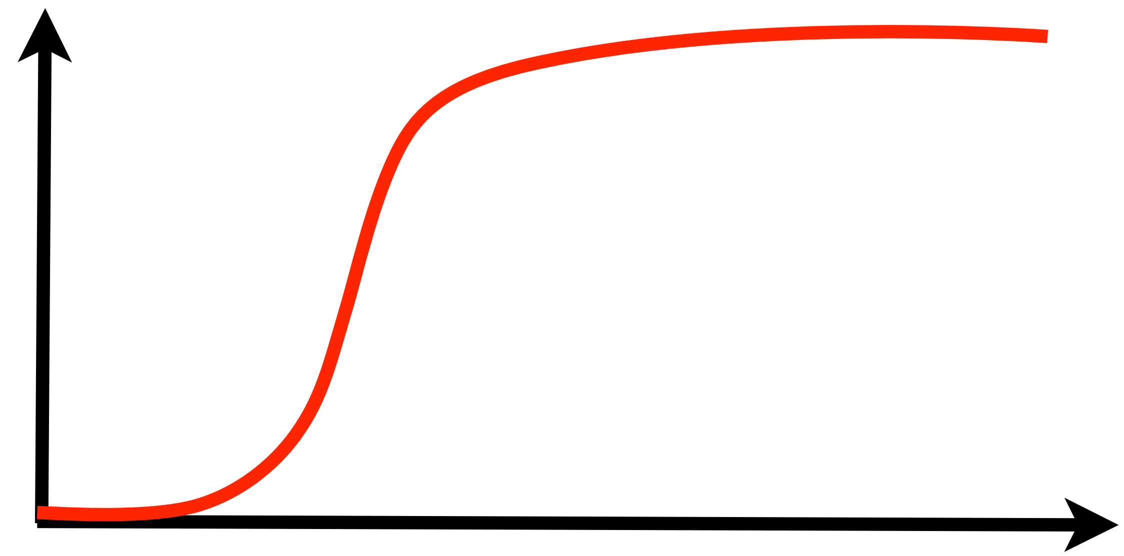
Dorsal view

Posterior

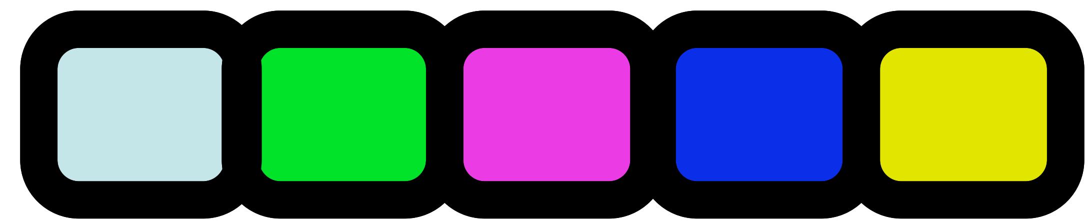


Patterning a field of cells : AP growth

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

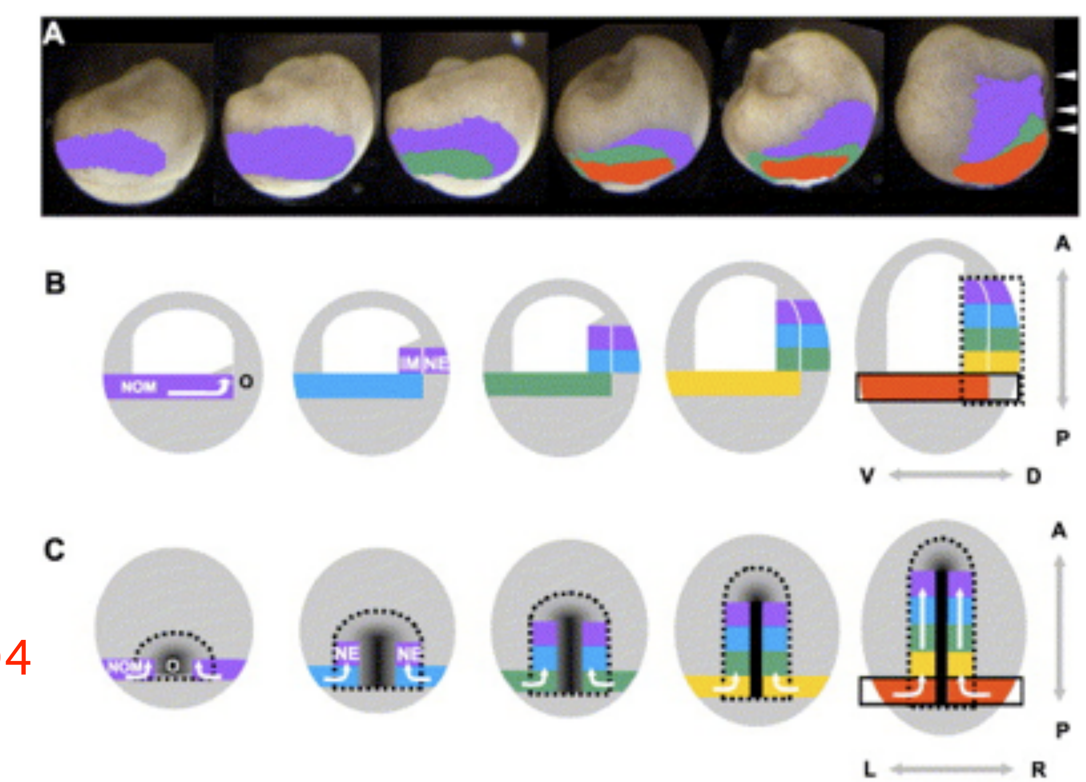


“Anterior” “Posterior”

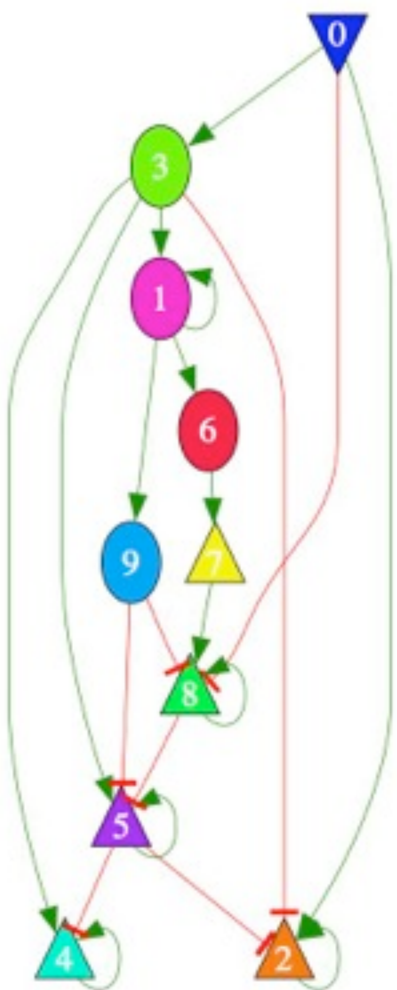
‘organizer’ is point where converging equator -> extending AP

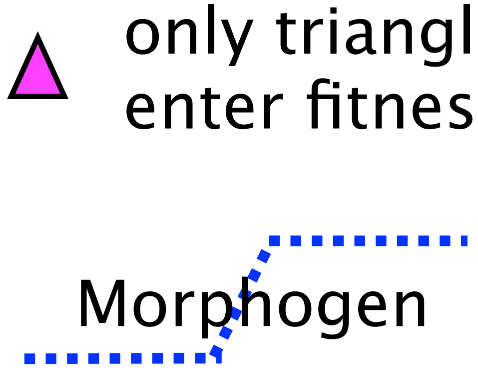
morphogen step ~ organizer

Wacher 2004

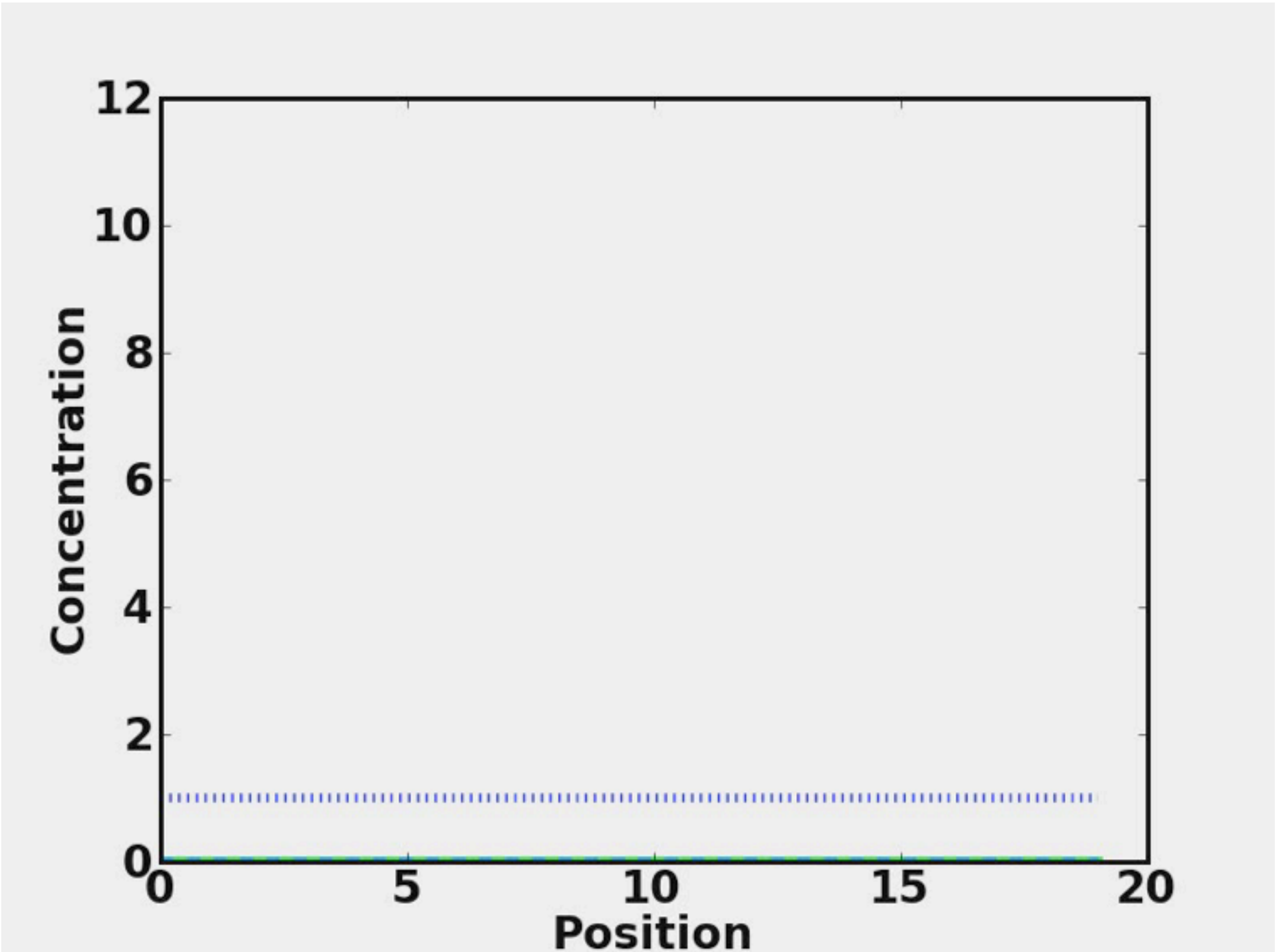


Sliding morphogen (2)

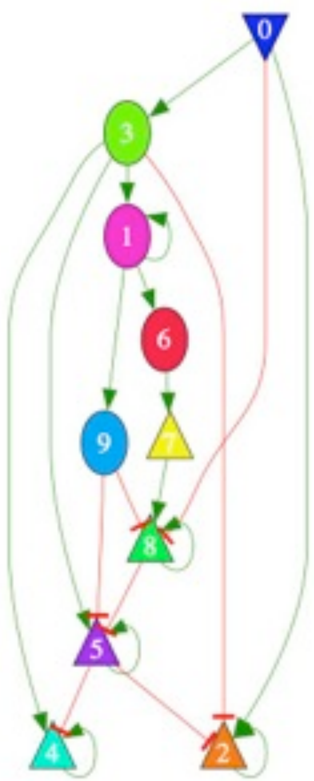


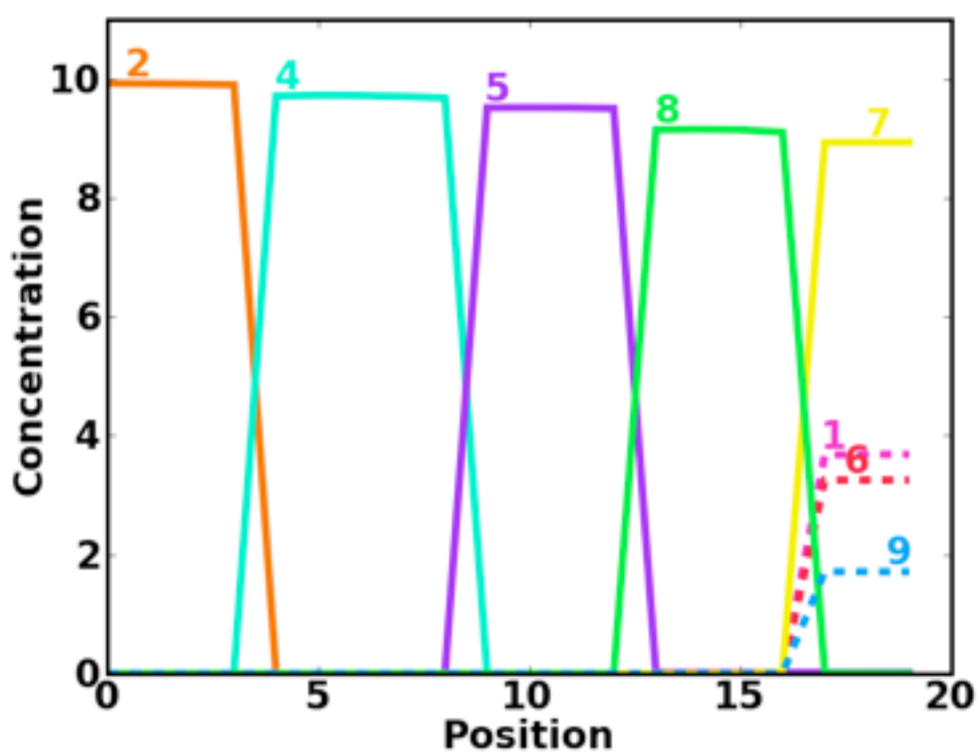


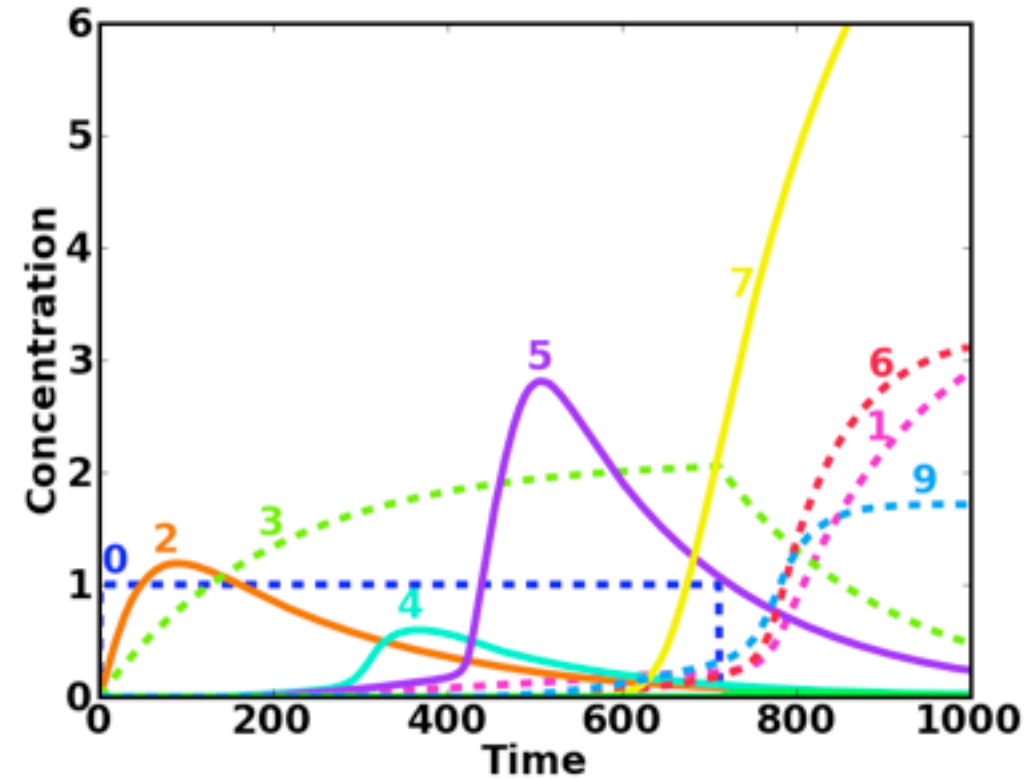
Time development of final evolved network



Temporal colinearity

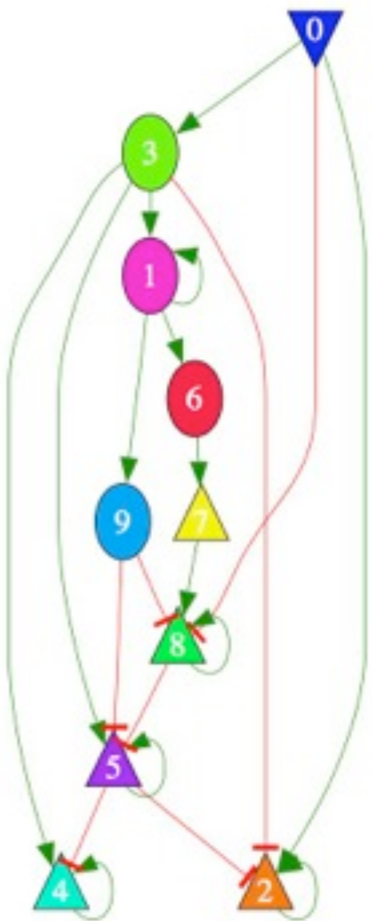


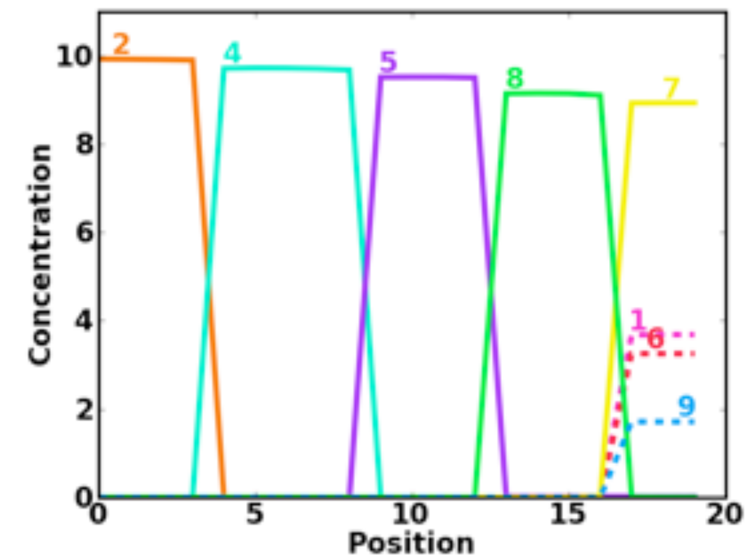


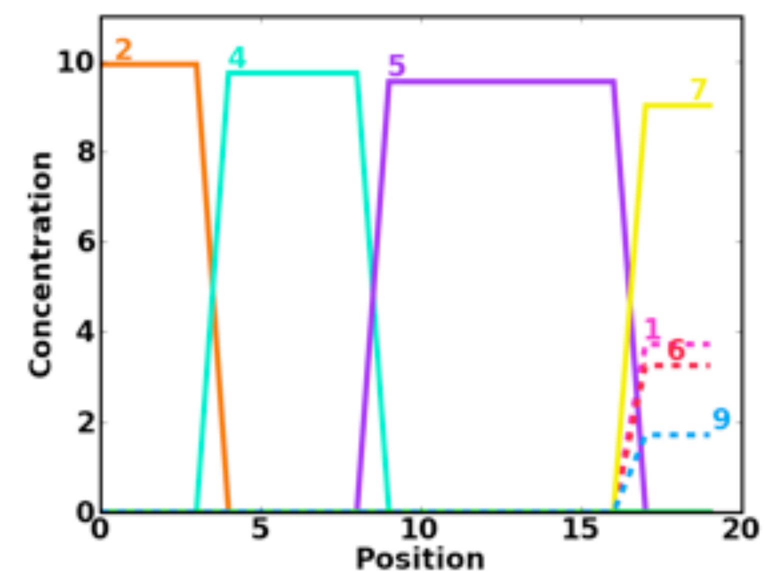


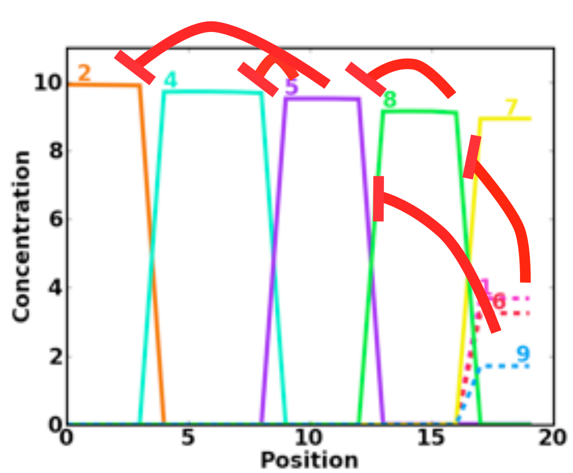
Temporal colinearity: Hox(time) fixed posterior cell -->
Anterior-Posterior progression

Anterior Homeotic Mutation (2)









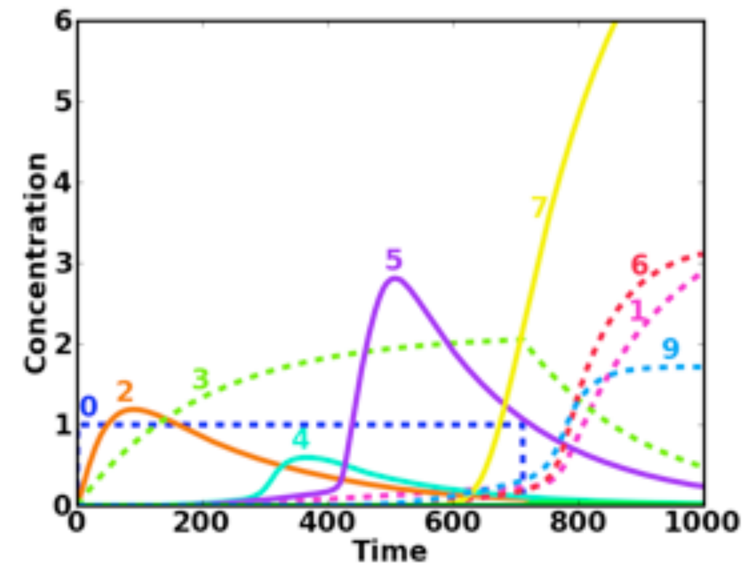
Properties of Networks with Sliding Gradient

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 etal timer ~ CAUDAL, CDX2)

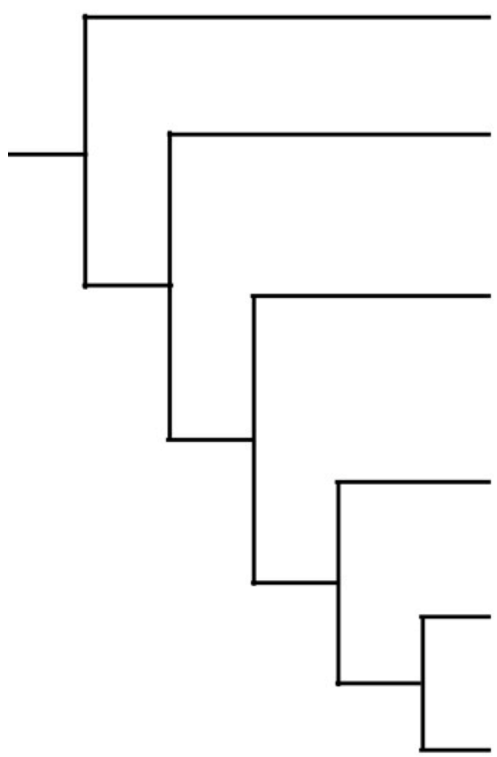
Genes(time) 1 cell



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)



order	species	hemi- or holo-metabolous	extended syncytial blastoderm	sequential or long germ segmentation	anterior localization of maternal mRNAs
Orthoptera	*Schistocerca* sp. *Gryllus bimaculatus*	hemi	no	seq	?
Hemiptera	*Oncopeltus fasciatus*	hemi	?	seq	?
Hymenoptera	*Nasonia vitripennis* *Bracon hebetor* *Aphidius ervi* *Apis mellifera*	holo	yes yes no yes	long long seq long*	yes ? ?
Coleoptera	*Tribolium castaneum* *Callosobruchus maculates*	holo	yes	seq long*	yes ?
Lepidoptera	*Bombyx mori* *Manduca sexta*	holo	no ?	seq* long*	?
Diptera	*Drosophila melanogaster* *Anopheles gambiae*	holo	yes	long	yes

AD Peel, Phil. Trans. R. Soc 2008

Short to long germ: *timer* gene -> 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)
- Somitogenesis (eg vertebrae): (Francois, Hakim, ES Mol.Sys.Bio. 2007)
- Adaptation to temporal signal (Francois & ES, Phys.Bio. 2008)
- 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)

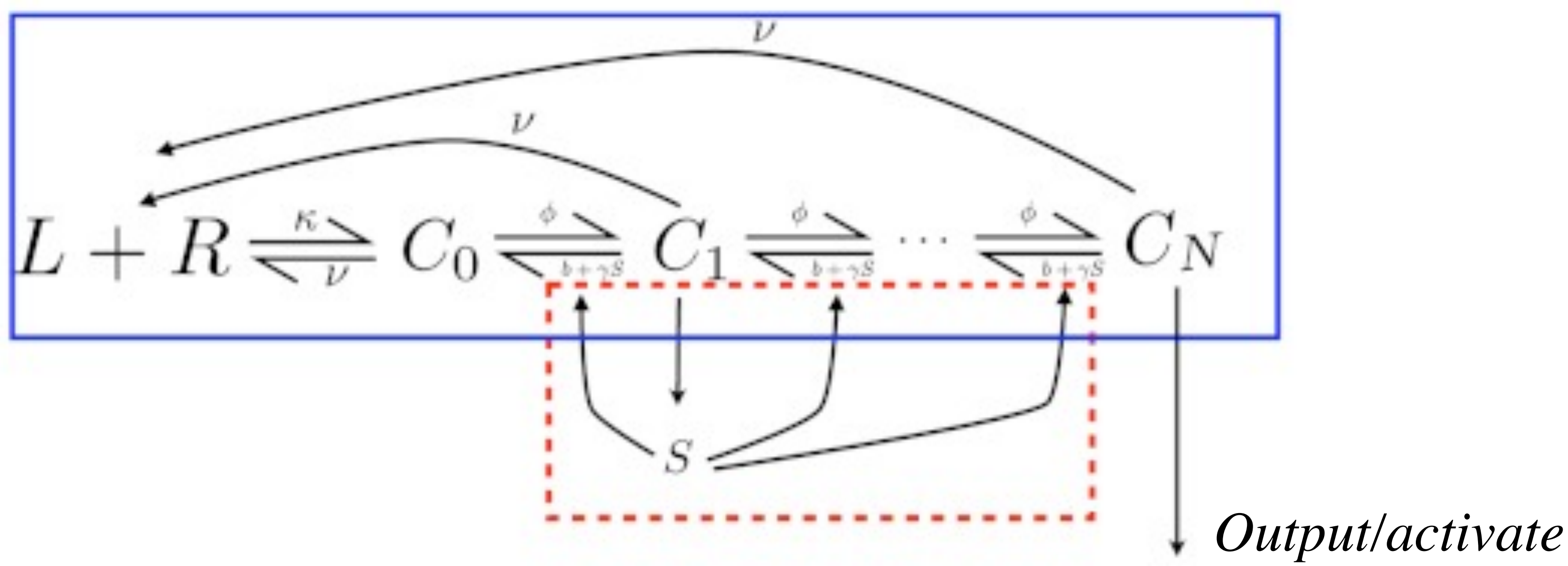
A few other applications in brief

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).

Model of phosphorylation cascade + self activated kinase/phosphatase can: (Altan-Bonnet & Germain PLoS Bio 2005)

System also evolved by Lalanne & Francois PRL 2013 (see also Francois etal PNAS 2013)



Optimal decision theory (Explore-Exploit)

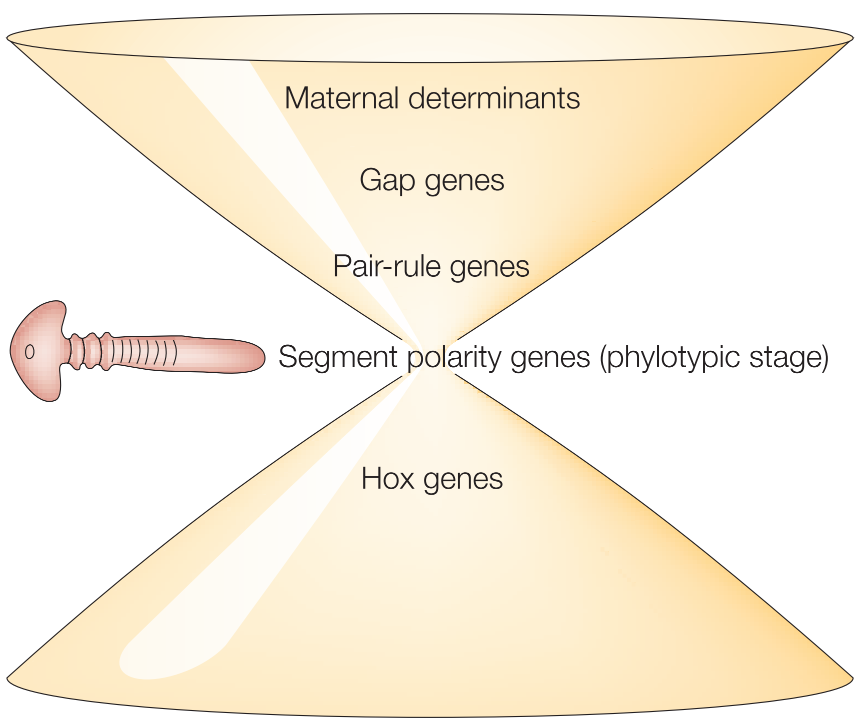
1. Given a stream of data from distribution A or B, what is minimum average decision time to identify the source for a given error rate, and what is the algorithm that realizes it?? (Wald 1945)

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.

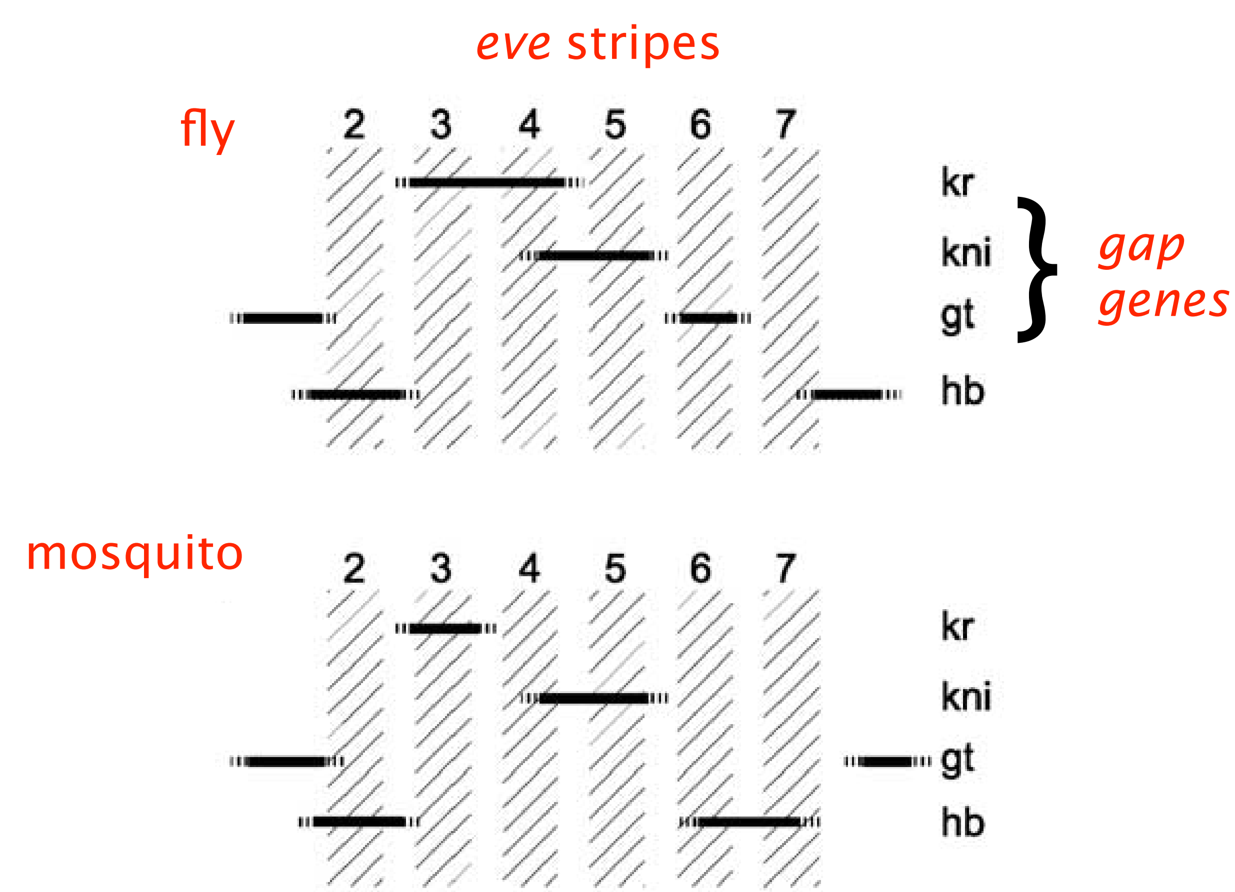
Refs:
Neural MN Shadlen ~2006
Cellular, Kobayashi 2010;
Vergassola & EDS 2013, Simple biochemical networks can optimally solve 1 & 2 and the parameters fit via local search.

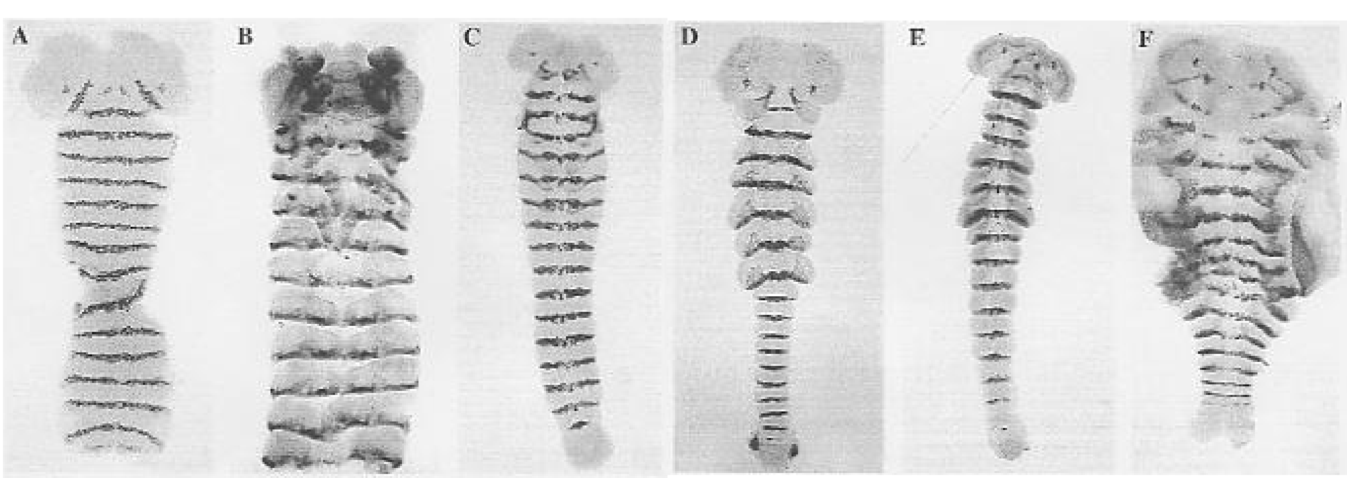
‘Saddle points’ or last common ancestor
(or how to turn a fly into a mosquito)



Peel, *Nat. Rev. Genet* 2005

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

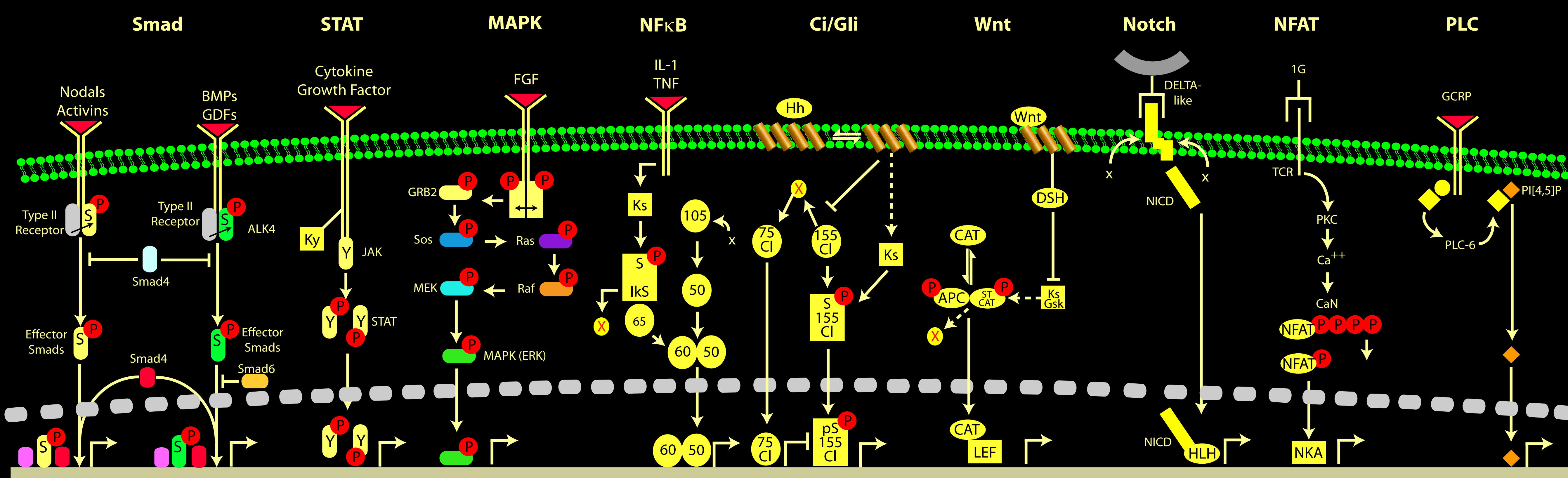




fly (dissected) moth beetle cricket grasshopper crayfish

Engrailed (and wg) mark segment boundaries
N H Patel *Development Suppl* 201-207 1994

Signaling pathways involved in cell fate determination



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