

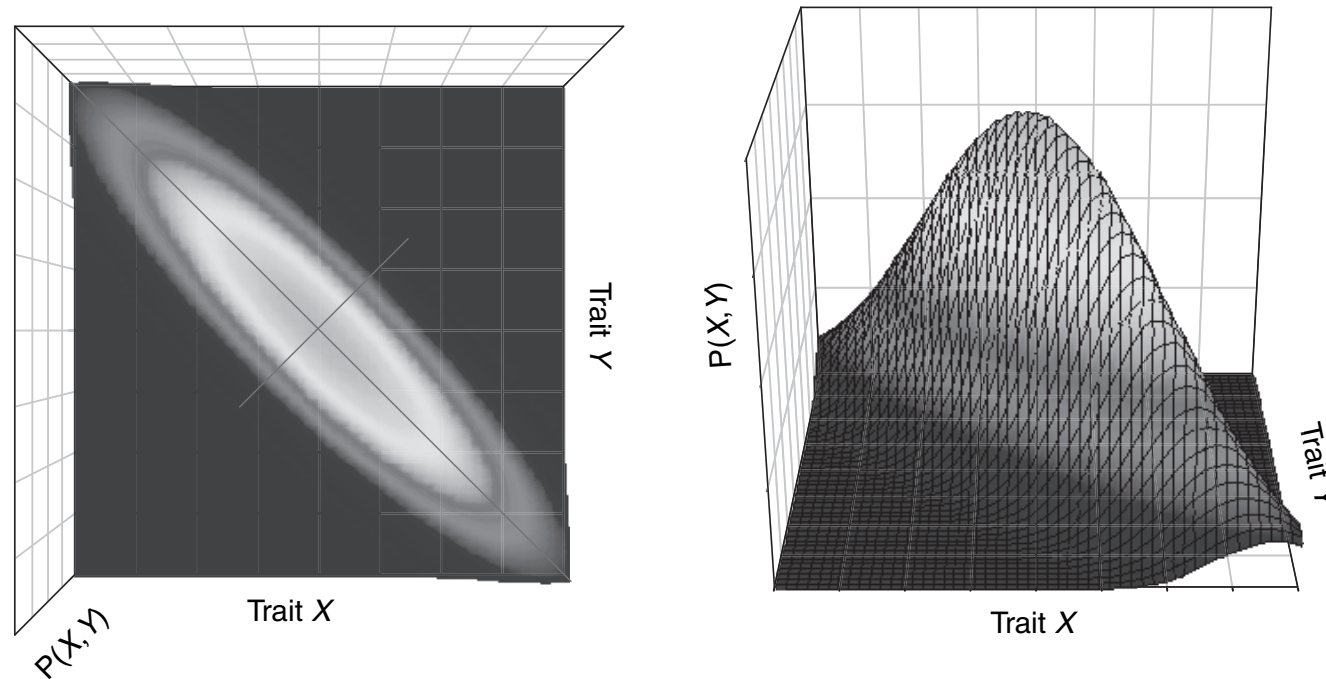
Predicting metabolic adaptation from networks of mutational paths

Peter Swain
Biological Sciences
University of Edinburgh



Trade-offs are common: improving performance in one task undermines performance in another

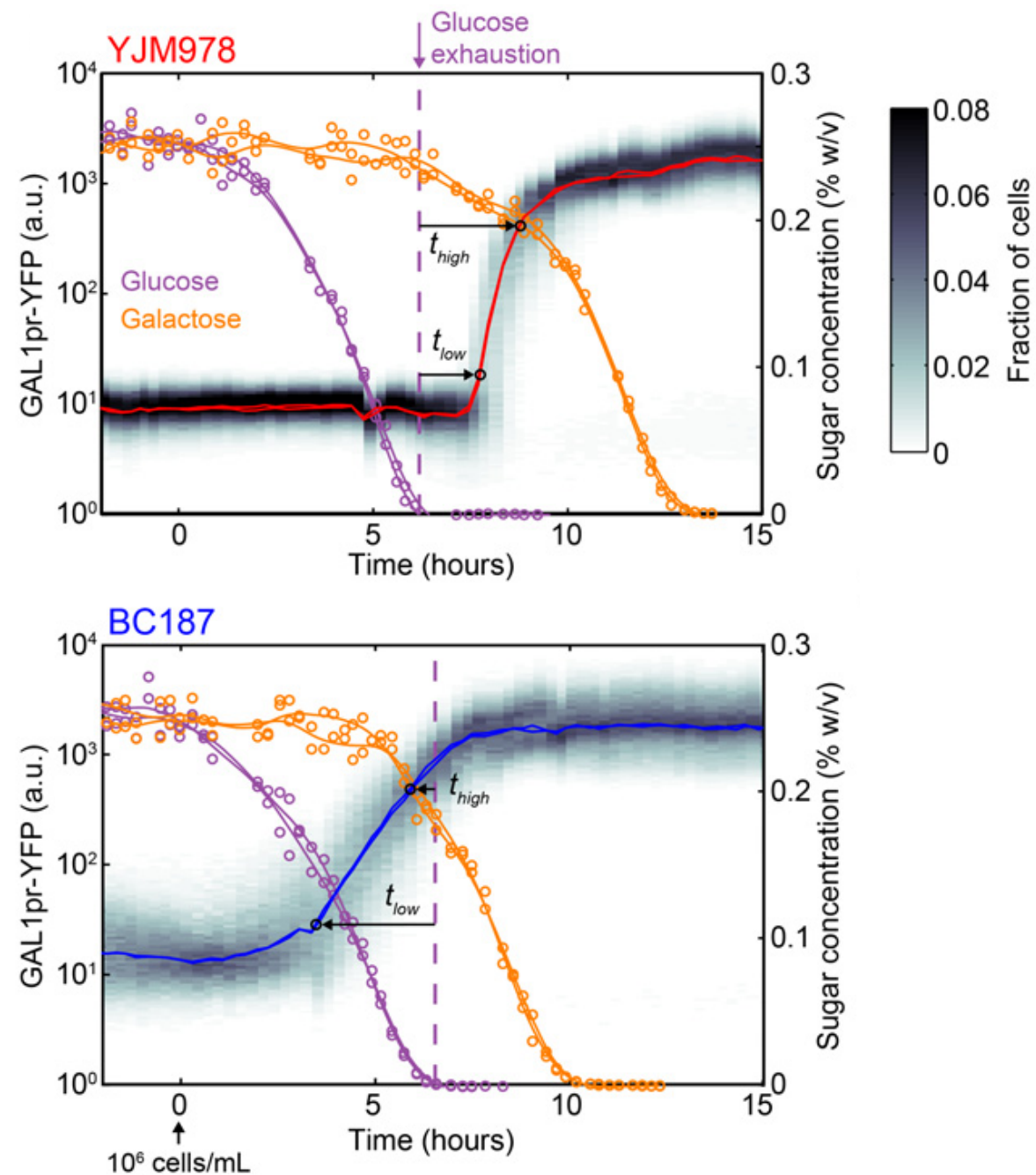
The whole organism is so tied together that when slight variations in one part occur, and are accumulated through natural selection, other parts become modified (Darwin)



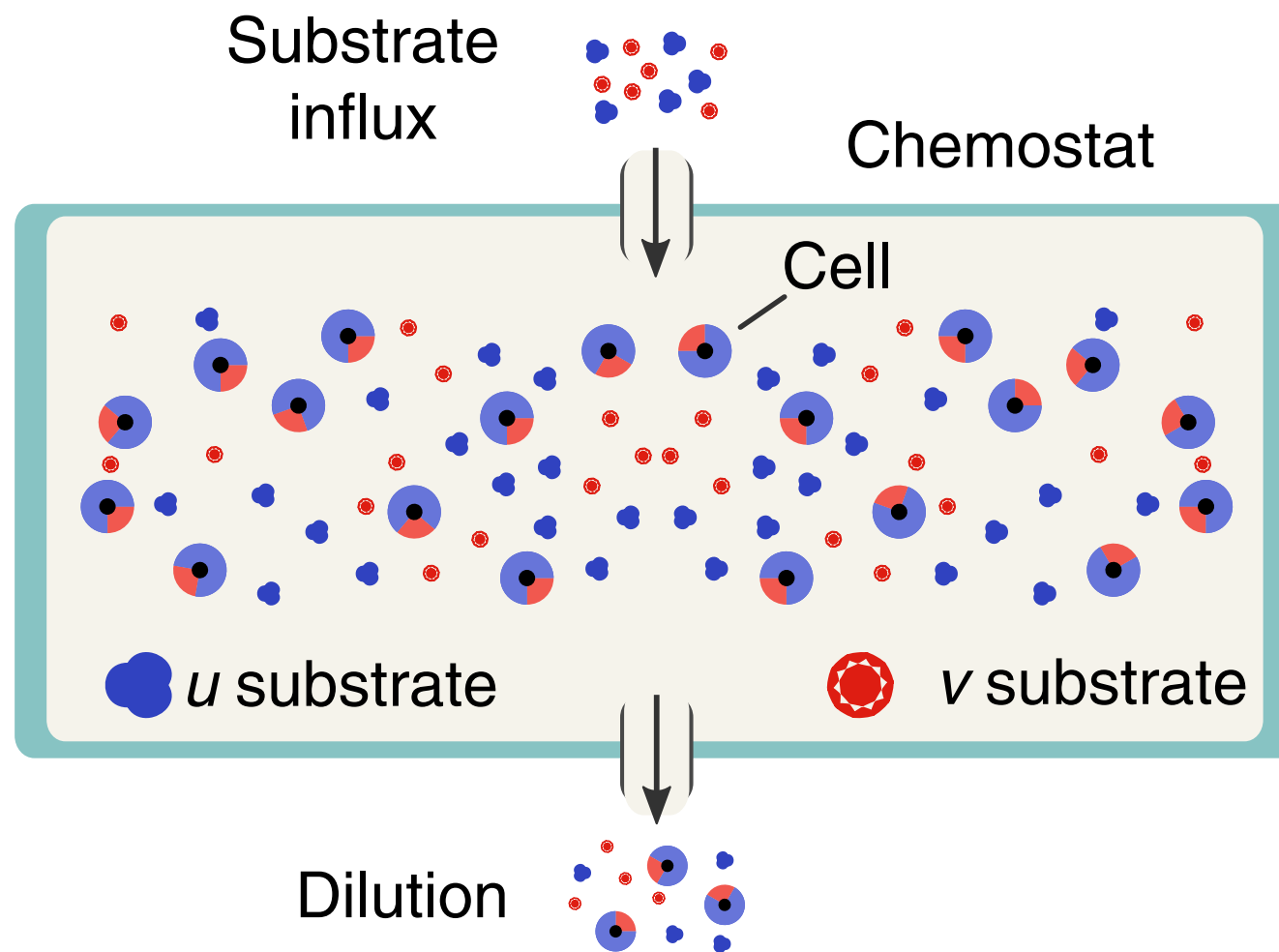
from Roff & Fairbairn, *J Evol Biol* 2006

Trade-offs are often driven by allocation of a finite resource (the Y-model).

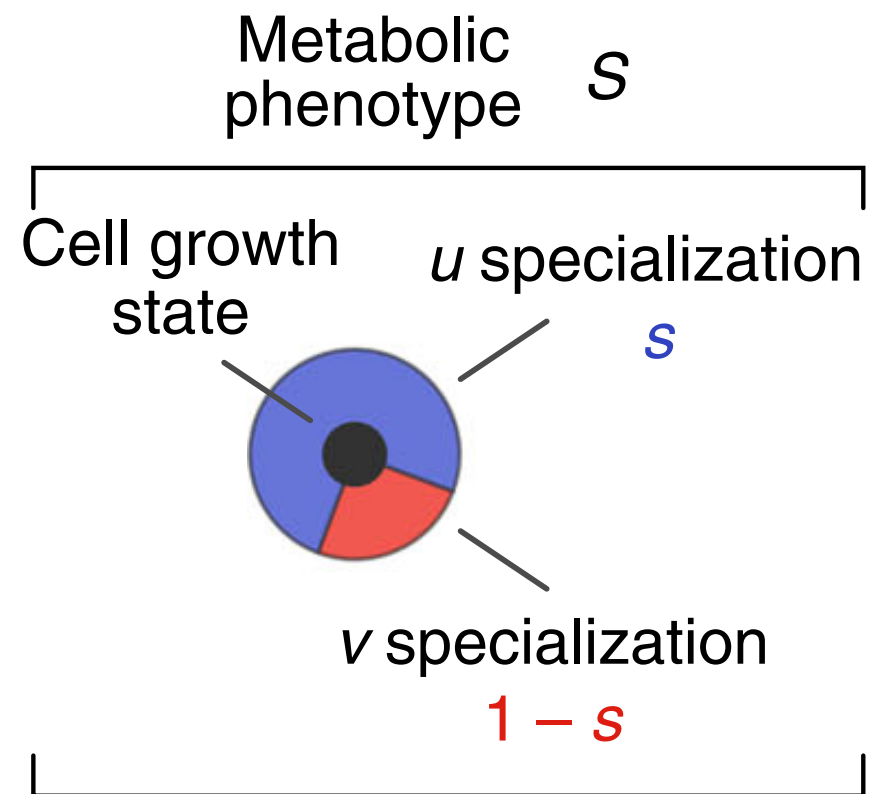
Partly inspired by growth of yeast on mixtures of sugars, we wish to better understand how trade-offs determine evolution



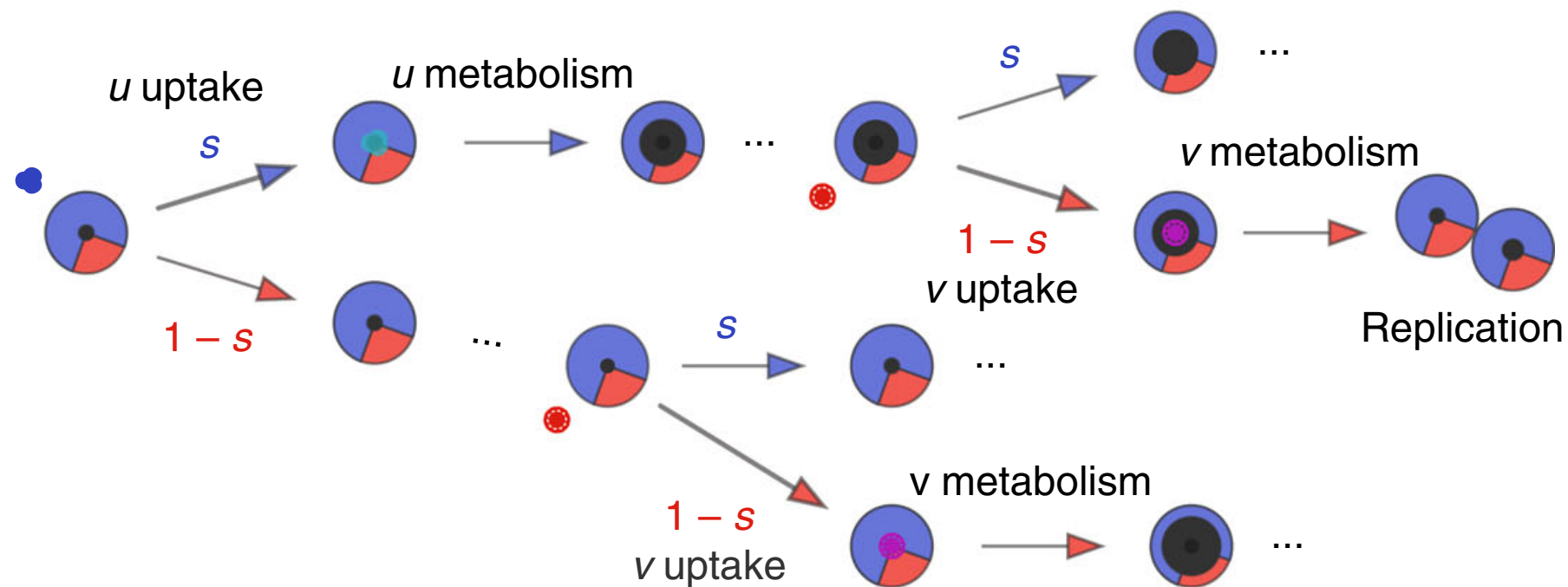
We consider growth in a chemostat with two potentially rate-limiting nutrients



Cells have a discrete phenotype – the degree of specialization for each of the two nutrients – that exhibits a trade-off

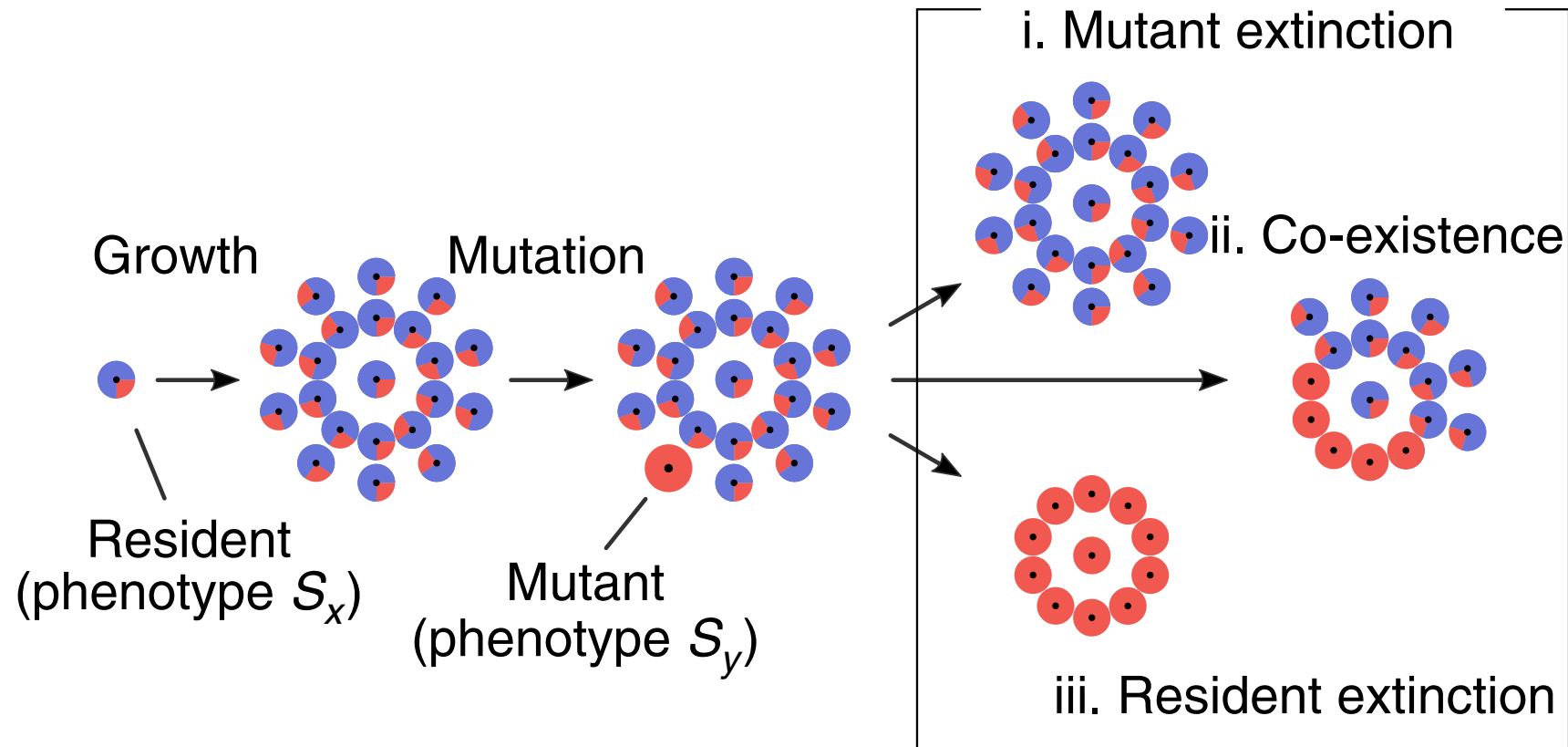


To allow the substrates to have different yields, cells grow by passing through discrete states of growth



The yield of a substrate is the amount by which the cells's state of growth increases on metabolizing the substrate.

Evolution is modelled with a weak rate of mutation



Mutations only arise after the chemostat reaches steady-state and there is then competition between the resident and mutant populations.

There are 10 parameters:

Dilution rate of the chemostat

Influx rate into the chemostat of substrates u and v

Maximum rates of import of u and v

Rates of metabolism of u and v

Yields of u and v

Maximum size of mutation

A mutation from s_x is equally likely to any s_y that obeys

$$|s_y - s_x| \leq \Delta S_{\max}$$

Cell growth and competition between cells of different phenotypes are both deterministic

For N_s states of growth, $\mathbf{n}_x(t) = (n_{1,x}(t), n_{2,x}(t), \dots, n_{N_s,x}(t))$

and

$$\dot{\mathbf{n}}_x(t) = -\mathbf{n}_x [uk_u s_x + vk_v (1 - s_x) + D] \\ + m_u [\mathbf{\Gamma}^{(u)}]^T \mathbf{n}_x^{(u)} + m_v [\mathbf{\Gamma}^{(v)}]^T \mathbf{n}_x^{(v)}$$

$$\dot{\mathbf{n}}_x^{(u)}(t) = \mathbf{n}_x uk_u s_x - \mathbf{n}_x^{(u)} (m_u + D)$$

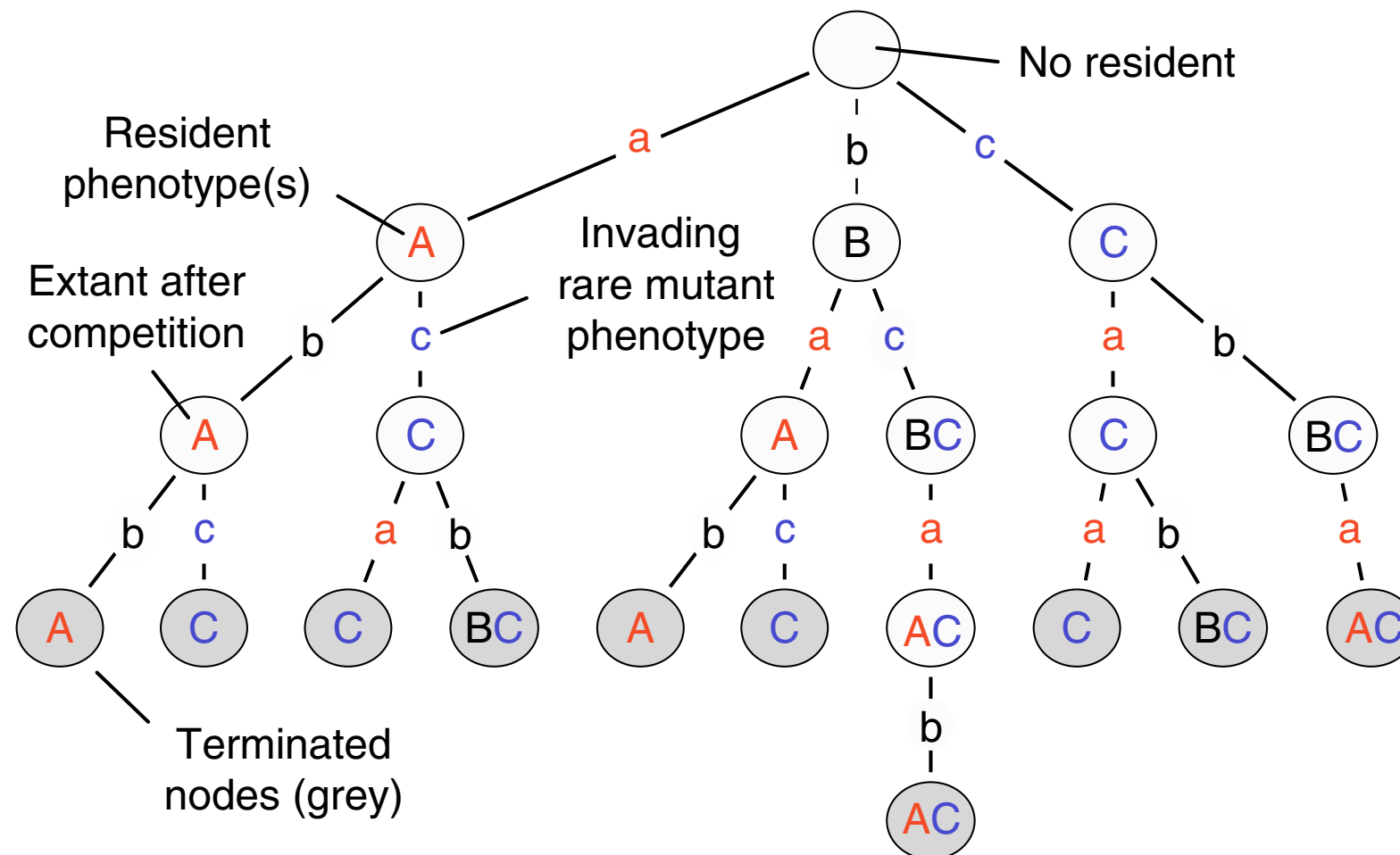
$$\dot{\mathbf{n}}_x^{(v)}(t) = \mathbf{n}_x vk_v (1 - s_x) - \mathbf{n}_x^{(v)} (m_v + D)$$

$$\dot{u}(t) = u^{(0)} - u \left[D + k_u \sum_x \mathbf{1}^T \mathbf{n}_x s_x \right]$$

$$\dot{v}(t) = v^{(0)} - v \left[D + k_v \sum_x \mathbf{1}^T \mathbf{n}_x (1 - s_x) \right]$$

To find the network of mutational paths for a given set of parameters, we first determine the invasion map

An illustrative example with 3 phenotypes:



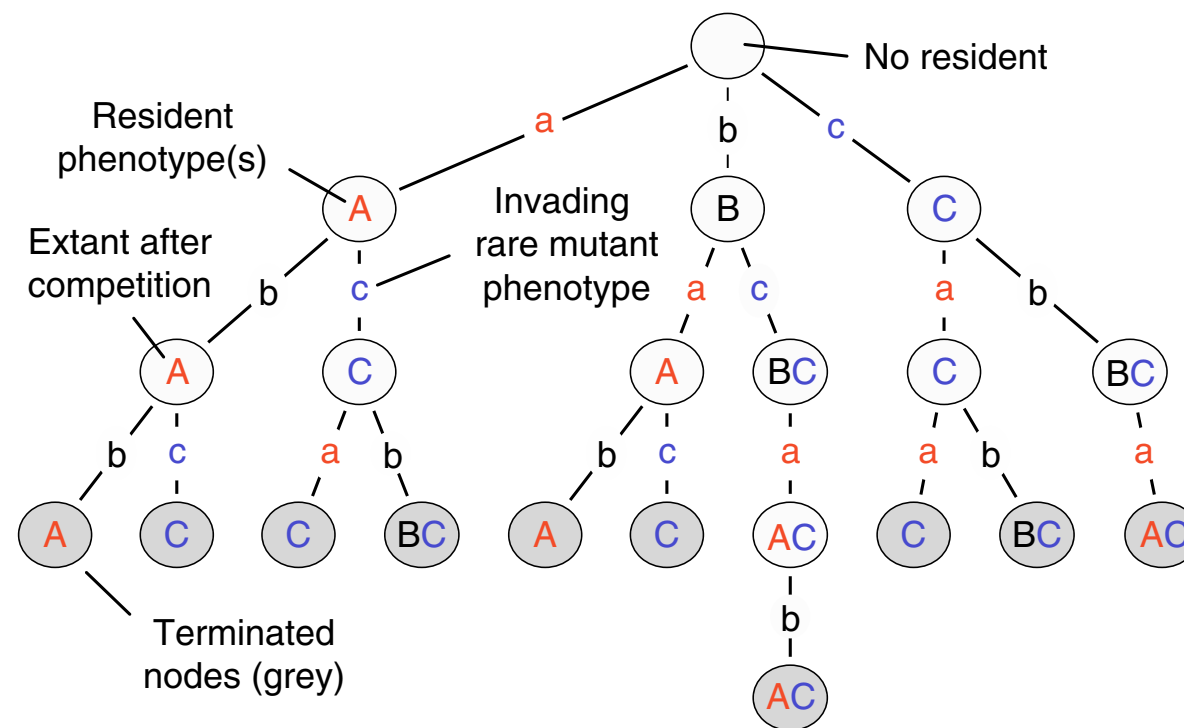
We next determine the conditional probabilities that one phenotype can be reached from another by mutation

An example with a small maximum size of mutation:

$$\begin{array}{c} \text{A} \\ \text{B} \\ \text{C} \end{array} \begin{array}{c} \text{a} \quad \text{b} \quad \text{c} \\ \left[\begin{array}{ccc} - & 1 & 0 \\ 1/2 & - & 1/2 \\ 0 & 1 & - \end{array} \right] \end{array} \begin{array}{c} \text{I} \\ \text{I} \\ \text{I} \end{array} \Delta S_{\max}$$

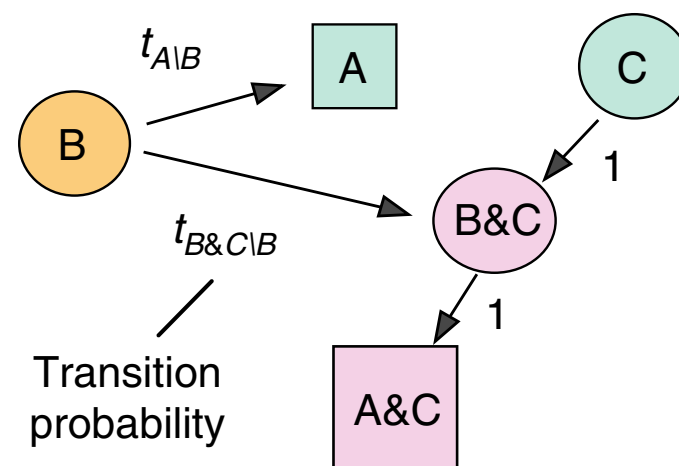
E.g. phenotype A can only mutate into phenotype B.

Combining the invasion map with the matrix of possible mutations gives the network of mutational paths

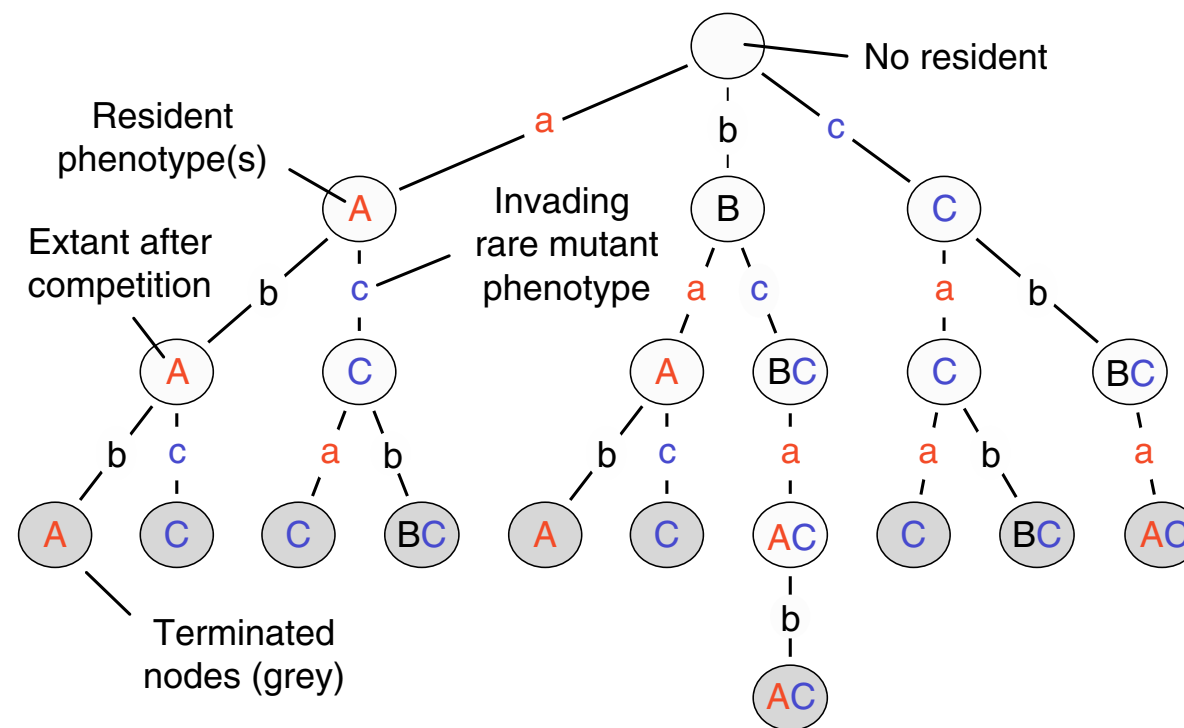


$$\begin{matrix}
 & \begin{matrix} a & b & c \end{matrix} \\
 \begin{matrix} A \\ B \\ C \end{matrix} & \begin{bmatrix} - & 1 & 0 \\ 1/2 & - & 1/2 \\ 0 & 1 & - \end{bmatrix}
 \end{matrix} \Bigg| \Delta S_{\max}$$

This network of mutational paths has two recurrent states (squares):

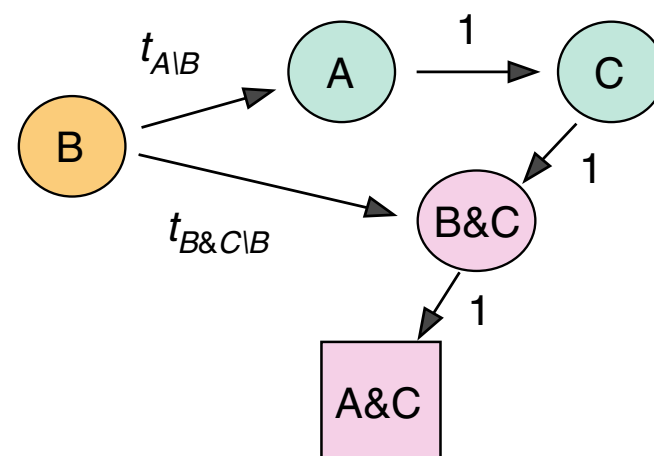


The network of mutational paths changes if the maximum size of mutation changes

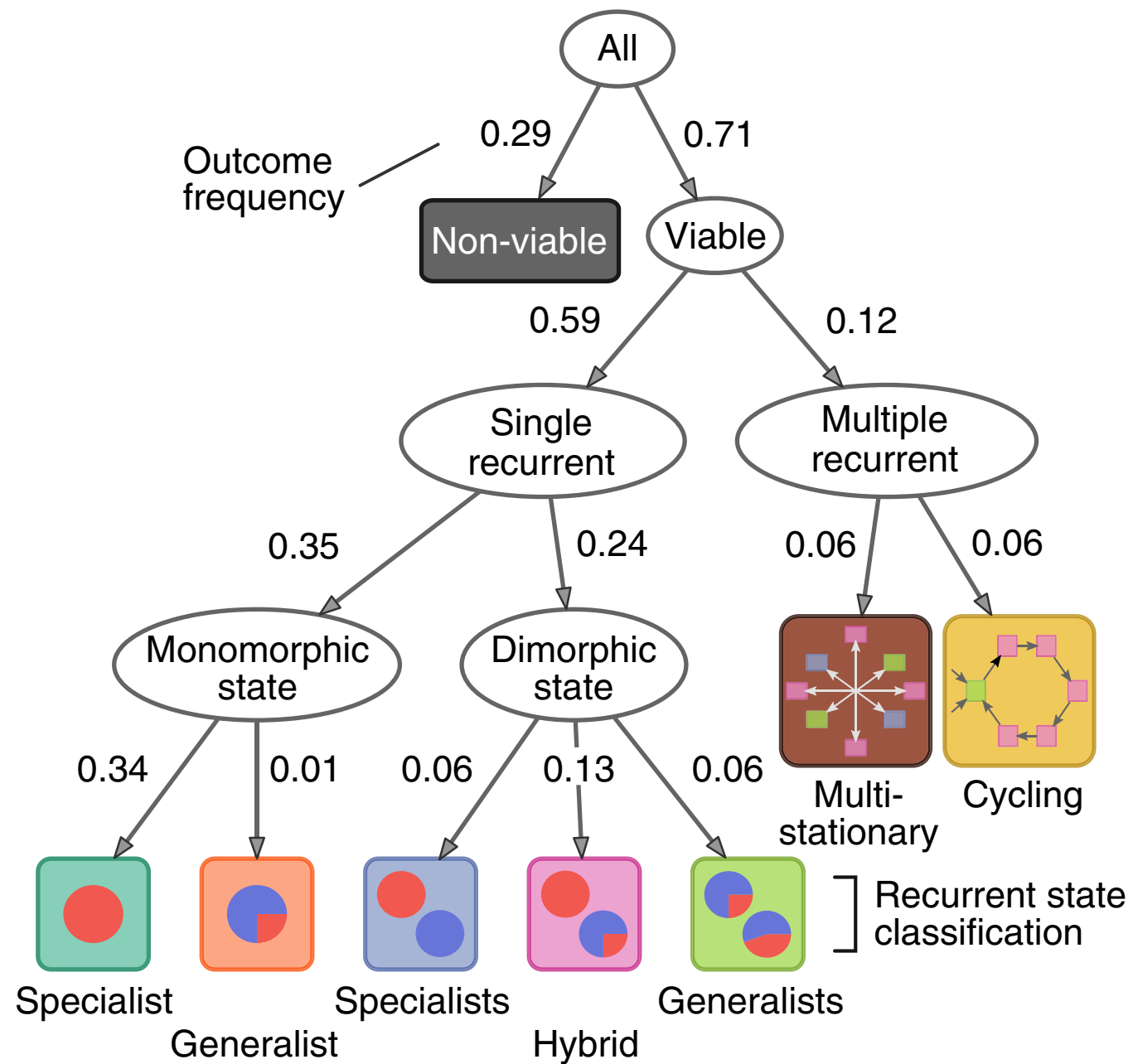


$$\begin{matrix}
 & \begin{matrix} a & b & c \end{matrix} \\
 \begin{matrix} A \\ B \\ C \end{matrix} & \begin{bmatrix} - & 1/2 & 1/2 \\ 1/2 & - & 1/2 \\ 1/2 & 1/2 & - \end{bmatrix} \\
 & \left. \vphantom{\begin{bmatrix} - & 1/2 & 1/2 \\ 1/2 & - & 1/2 \\ 1/2 & 1/2 & - \end{bmatrix}} \right] \Delta S_{\max}
 \end{matrix}$$

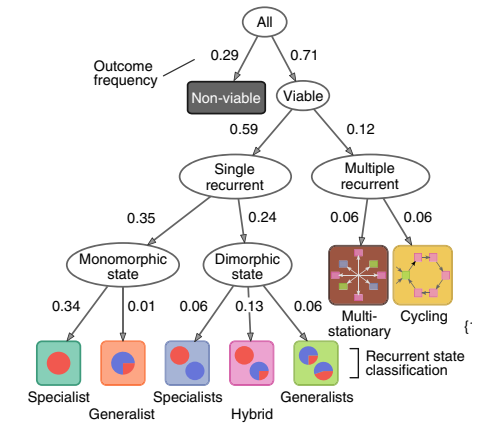
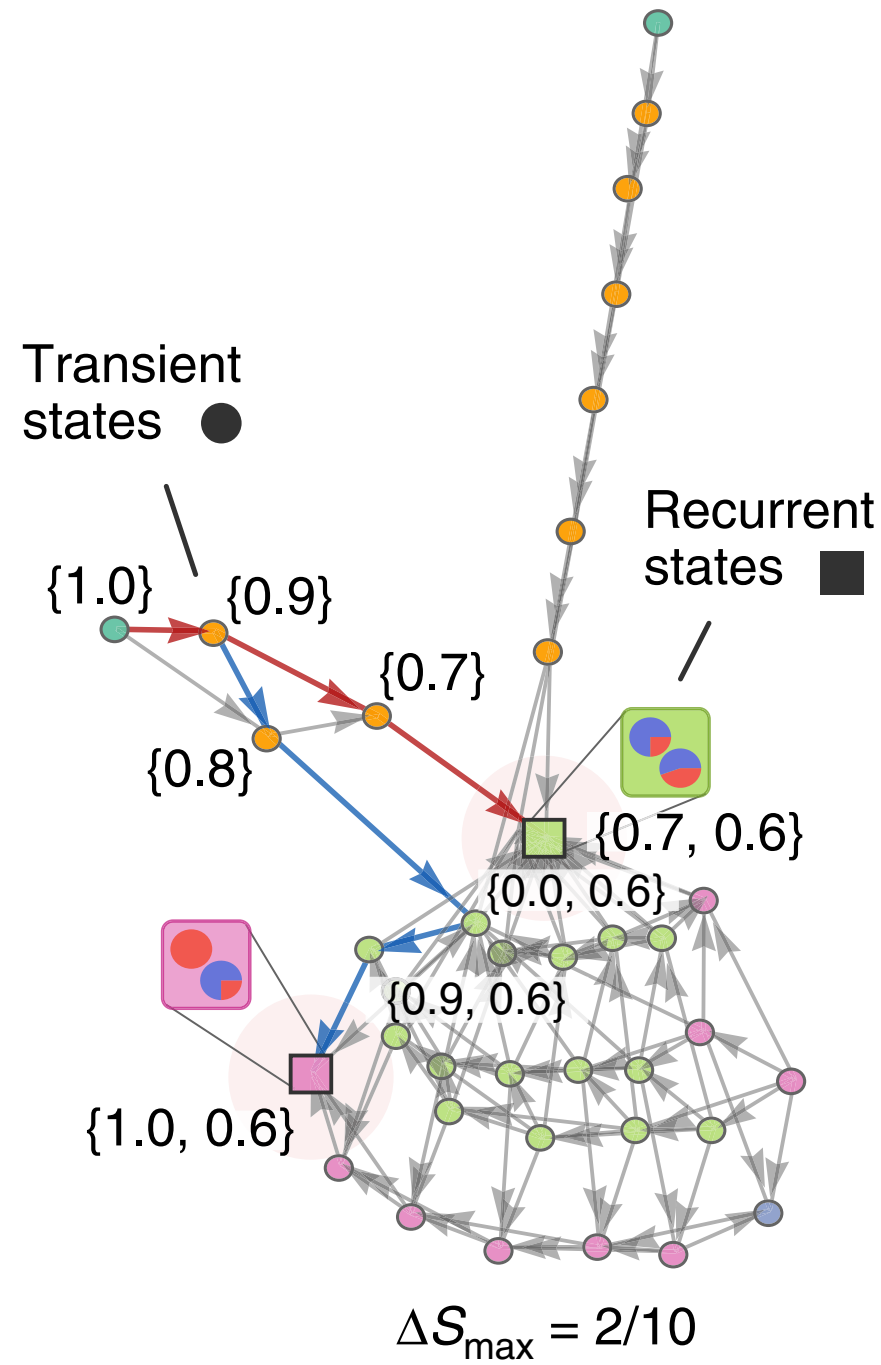
The network of mutational paths has now one recurrent state (square):



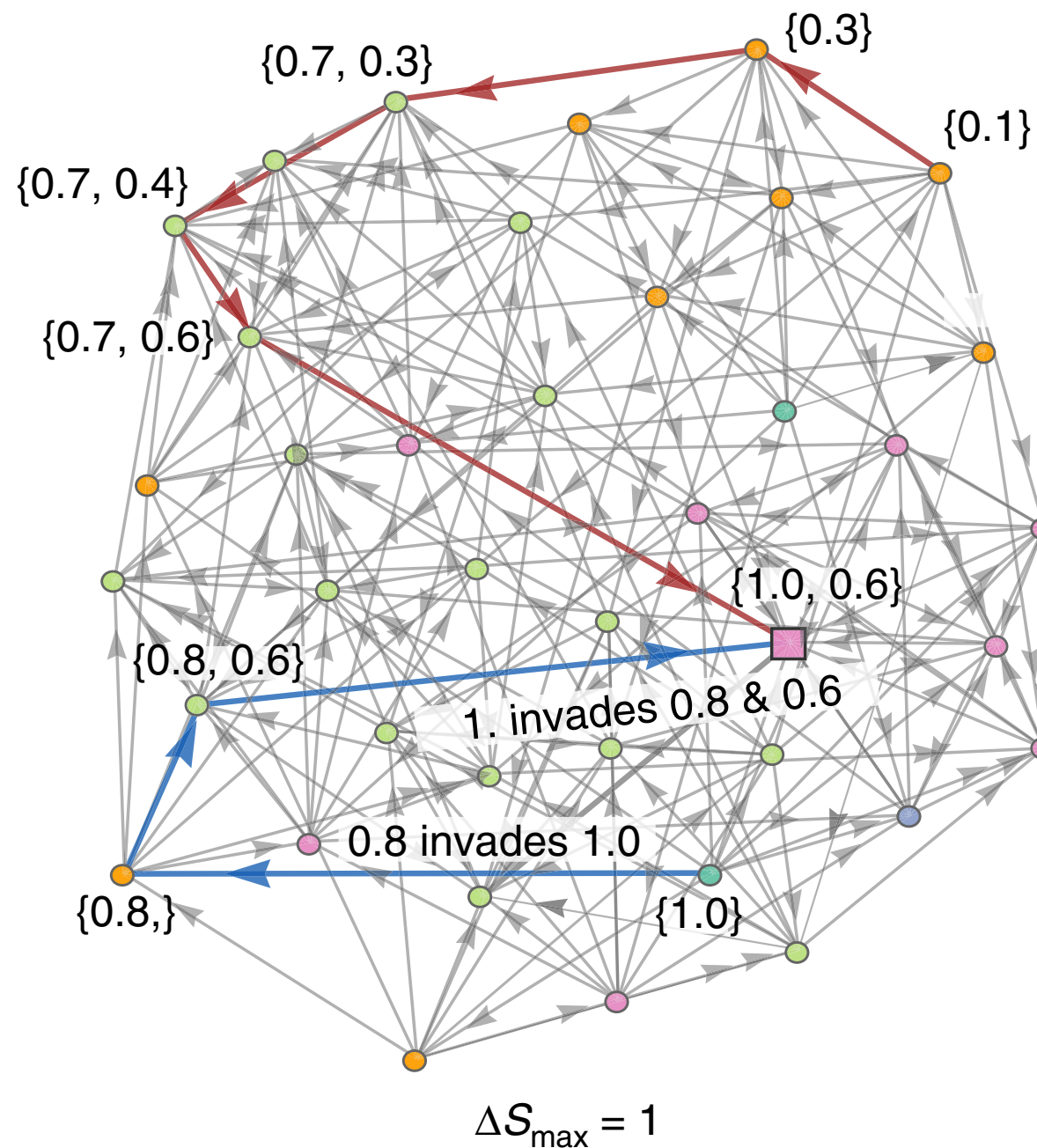
We characterize networks of mutational paths by their long-term, evolutionary outcomes (types of recurrent states)



Example: multiple recurrent states are possible

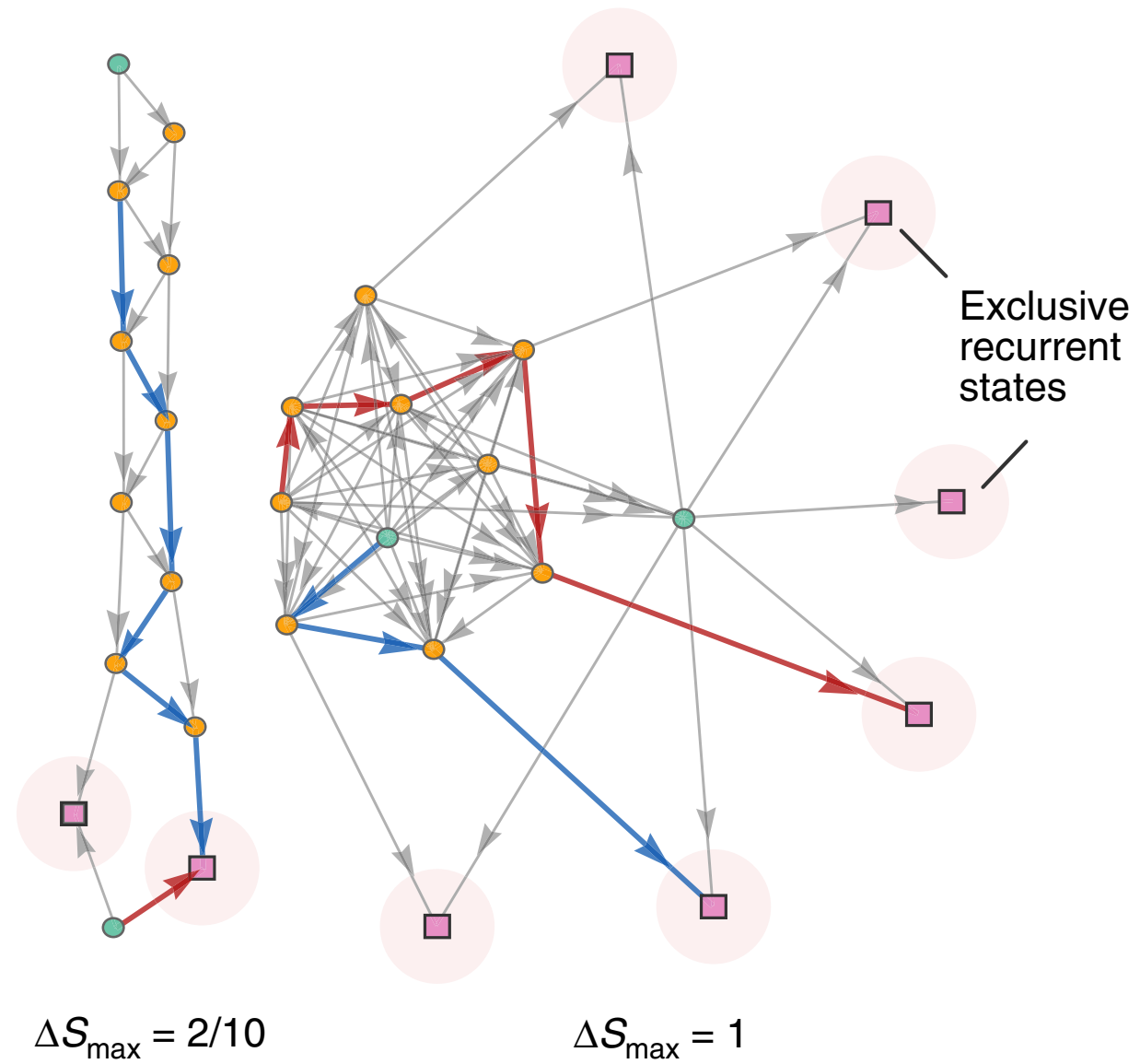


Increasing the maximum size of mutations changes the network and the recurrent states

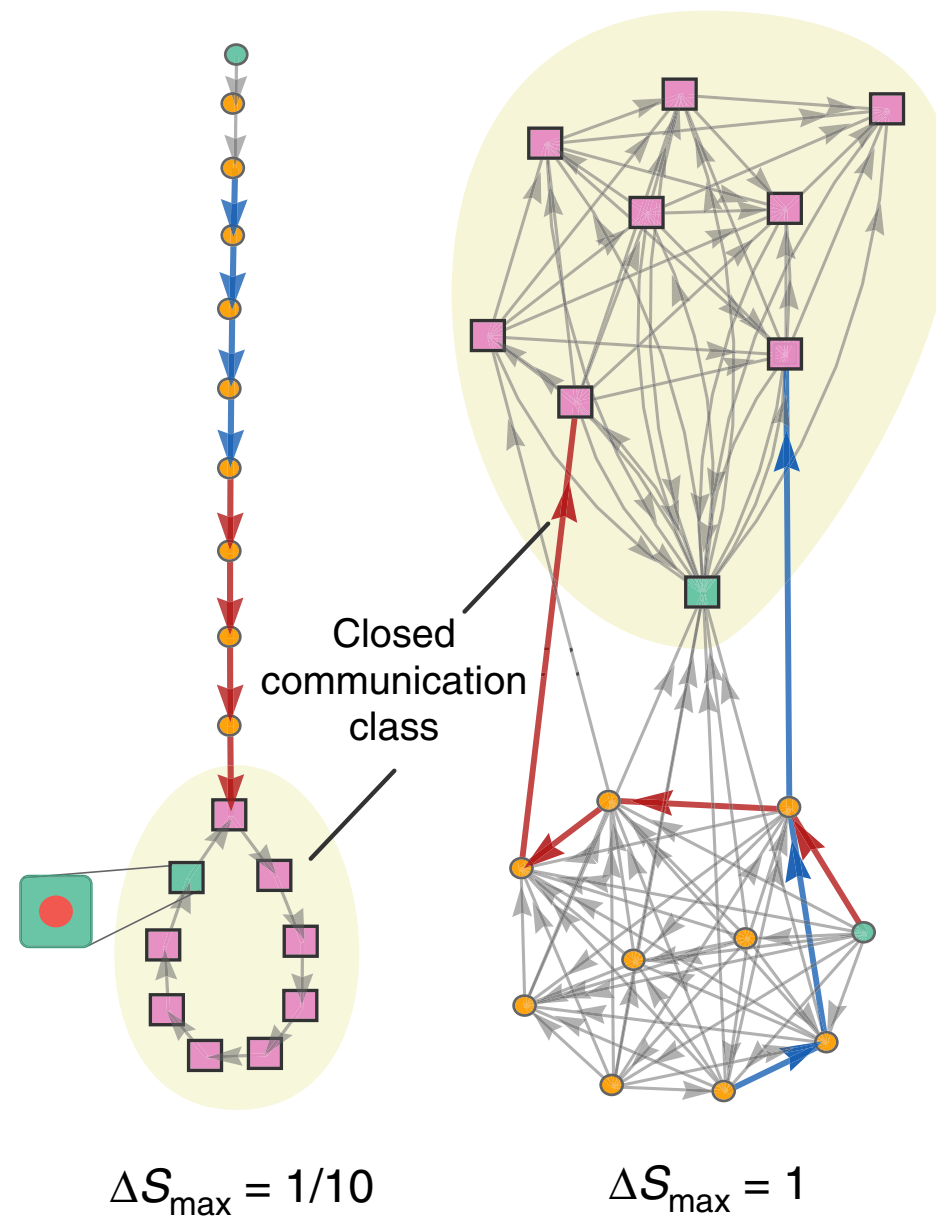


Context (the state of the environment) partly determines the successful phenotypes.

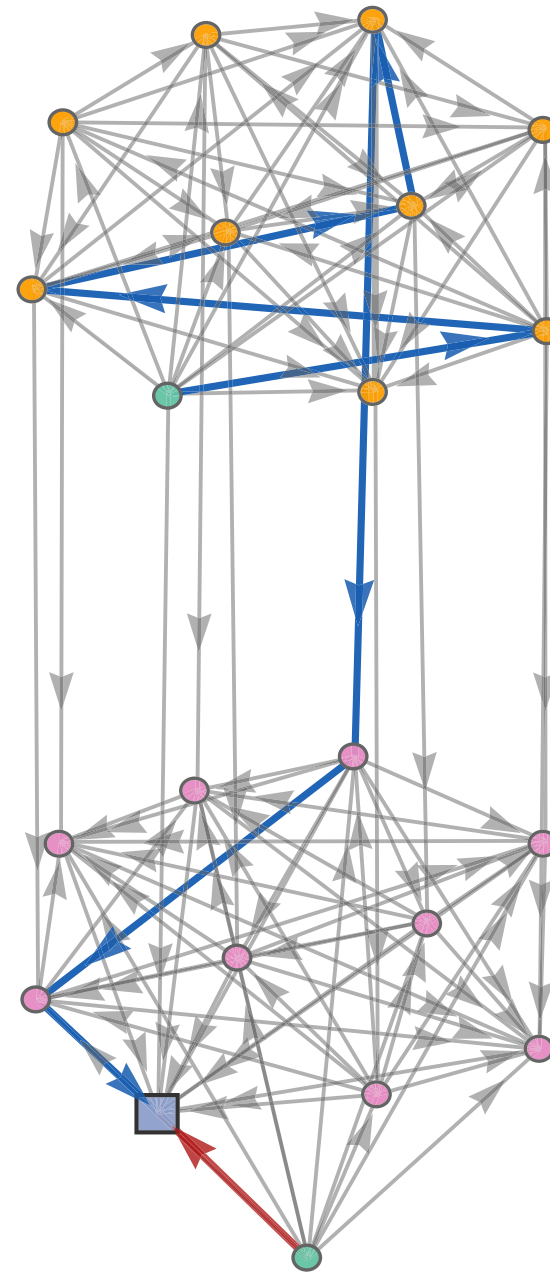
Multiple recurrent states are possible



Multiple recurrent states that are connected undergo evolutionary cycling



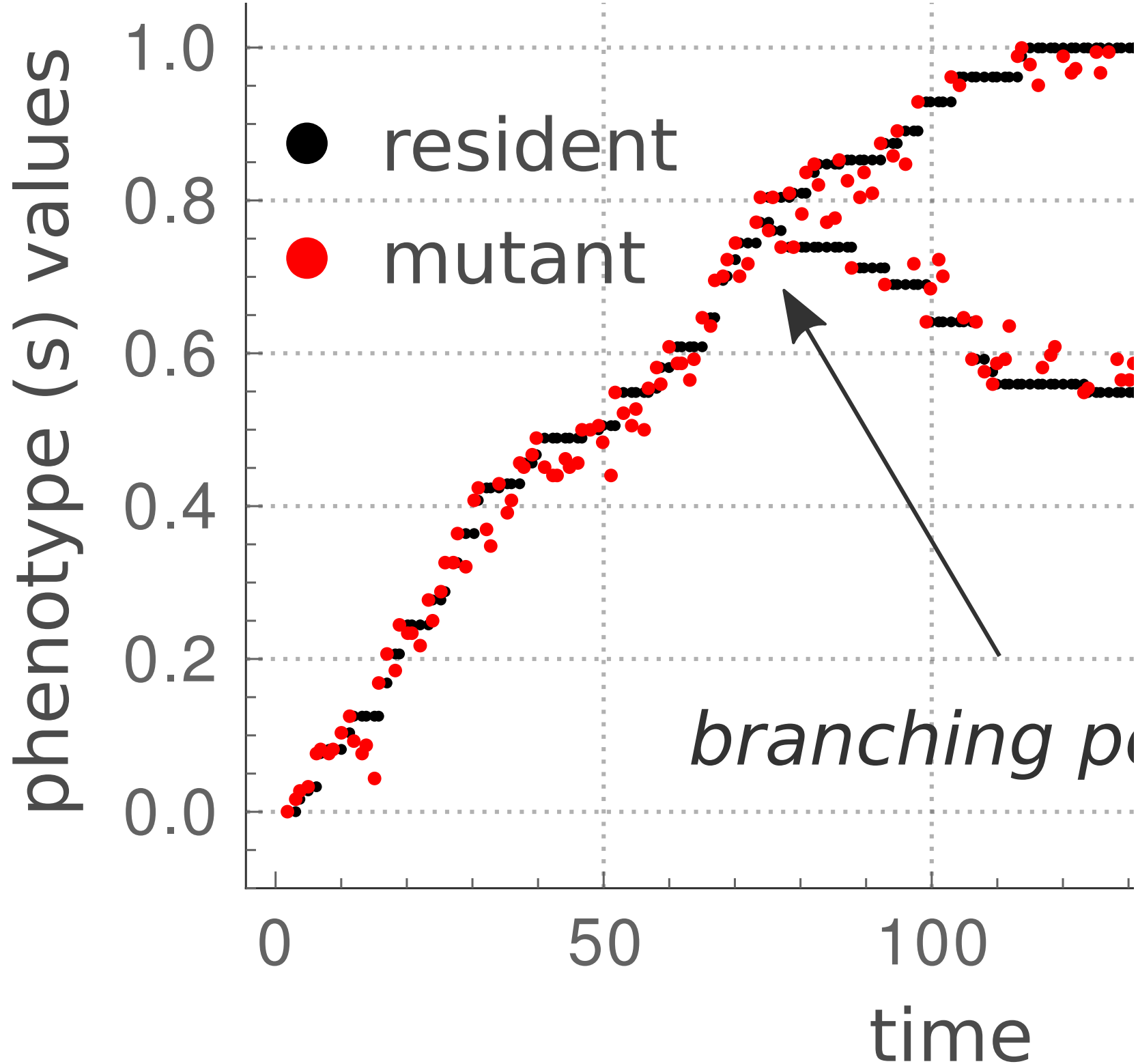
Evolution can be bottle-necked



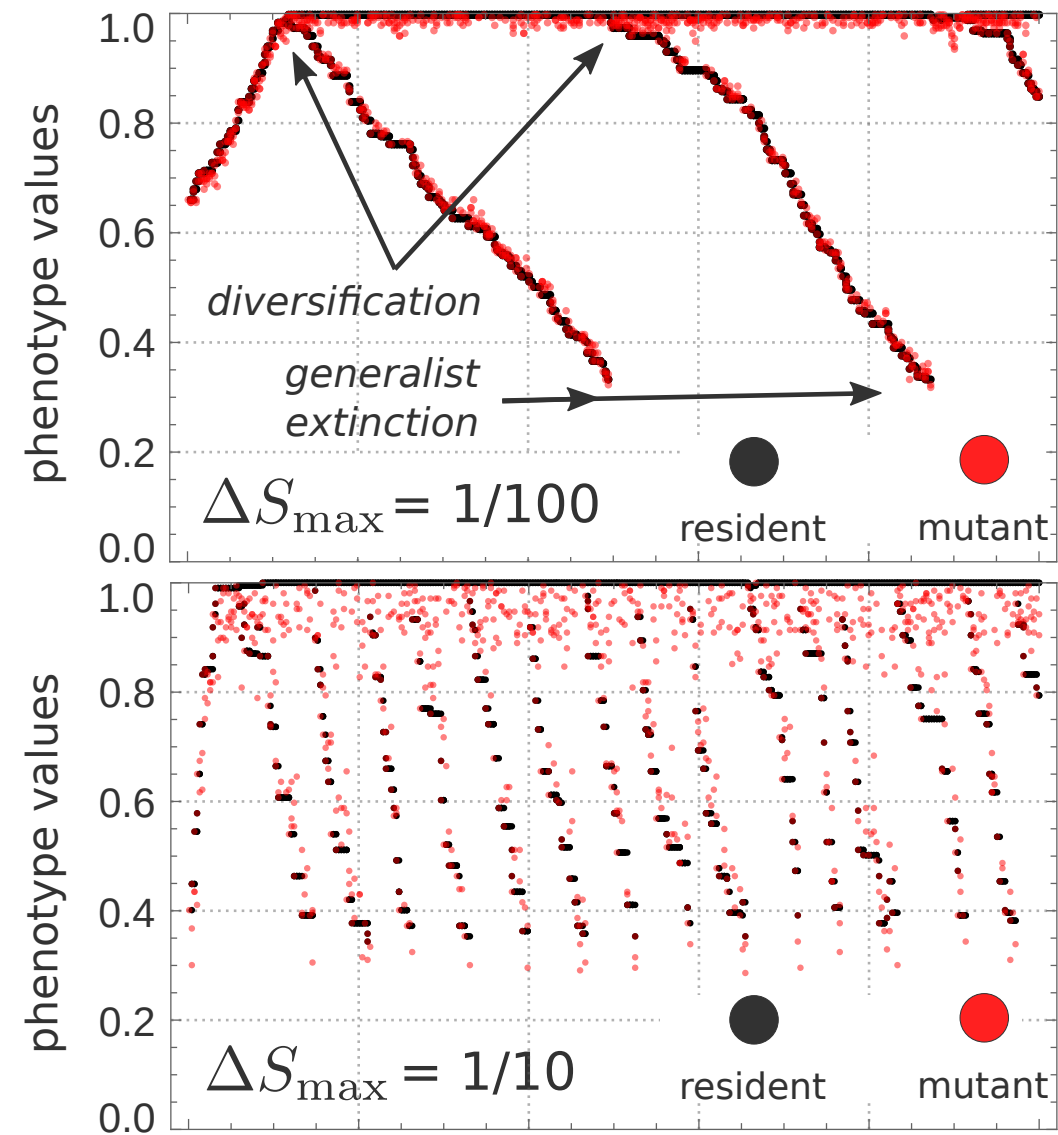
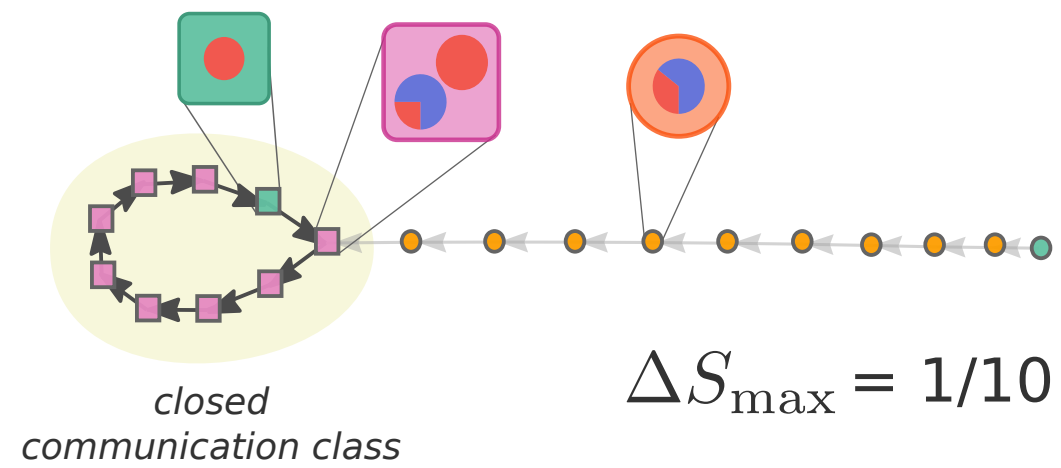
$$\Delta S_{\max} = 1$$

Two highly connected subnetworks are themselves connected by only a few mutation-invasion events.

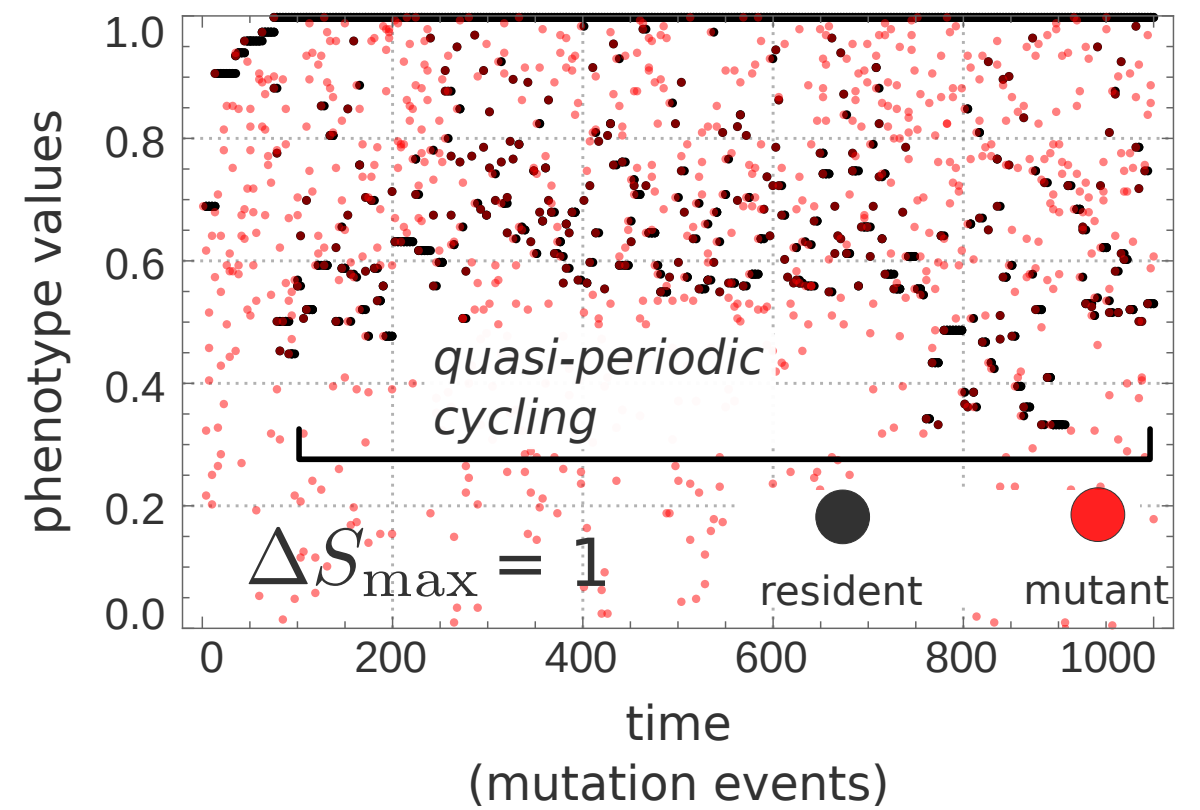
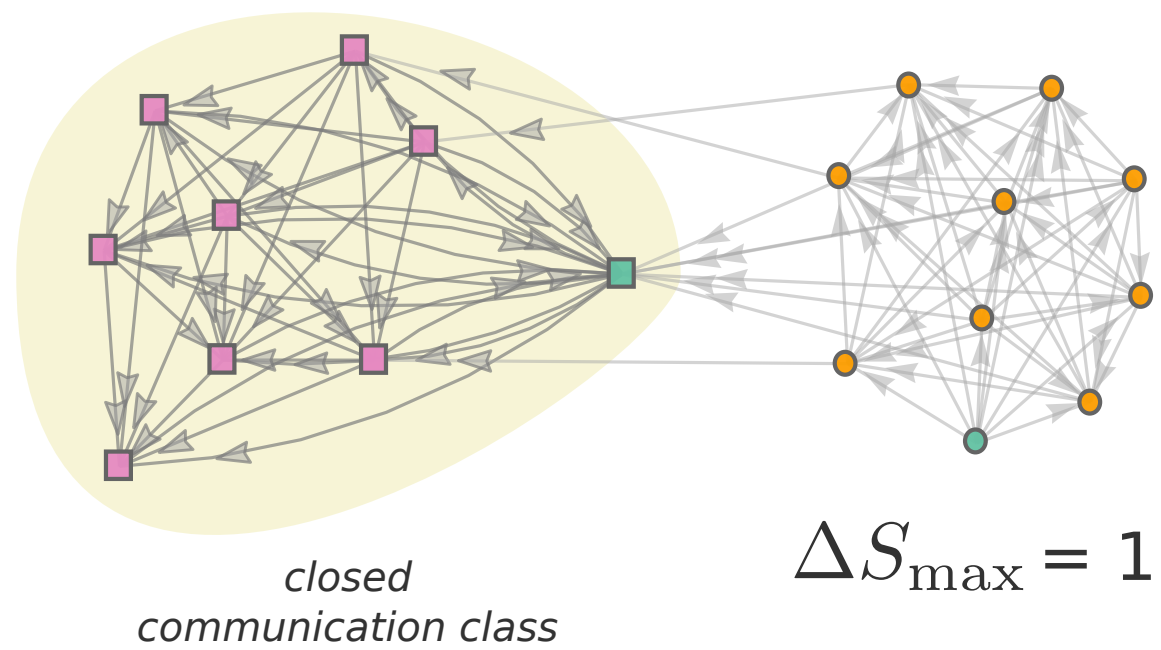
As a control, we run stochastic simulations with a continuous phenotype



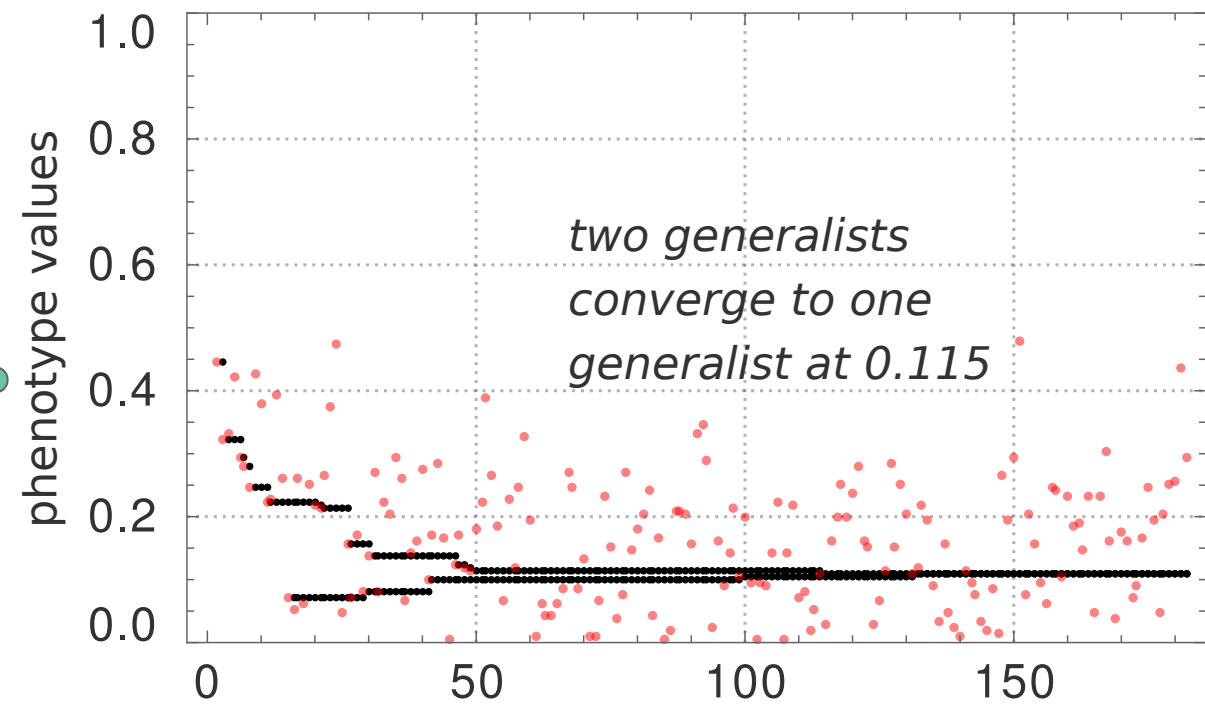
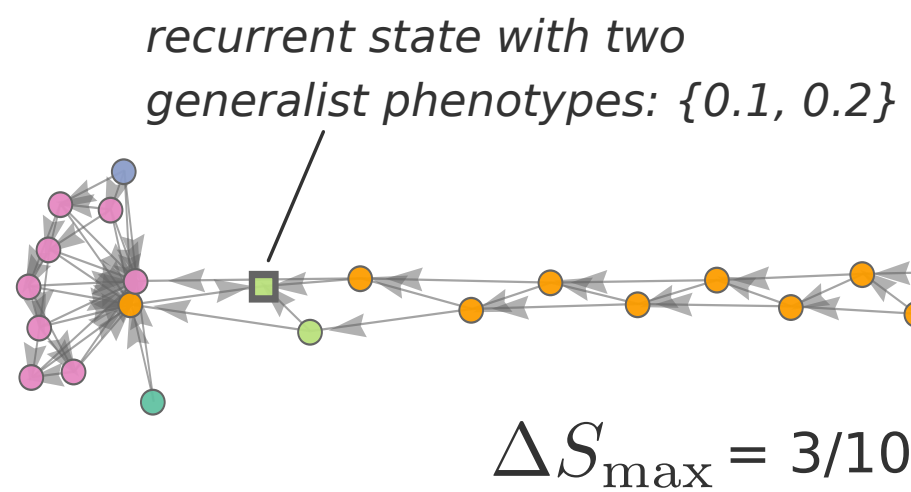
The recurrent states persist



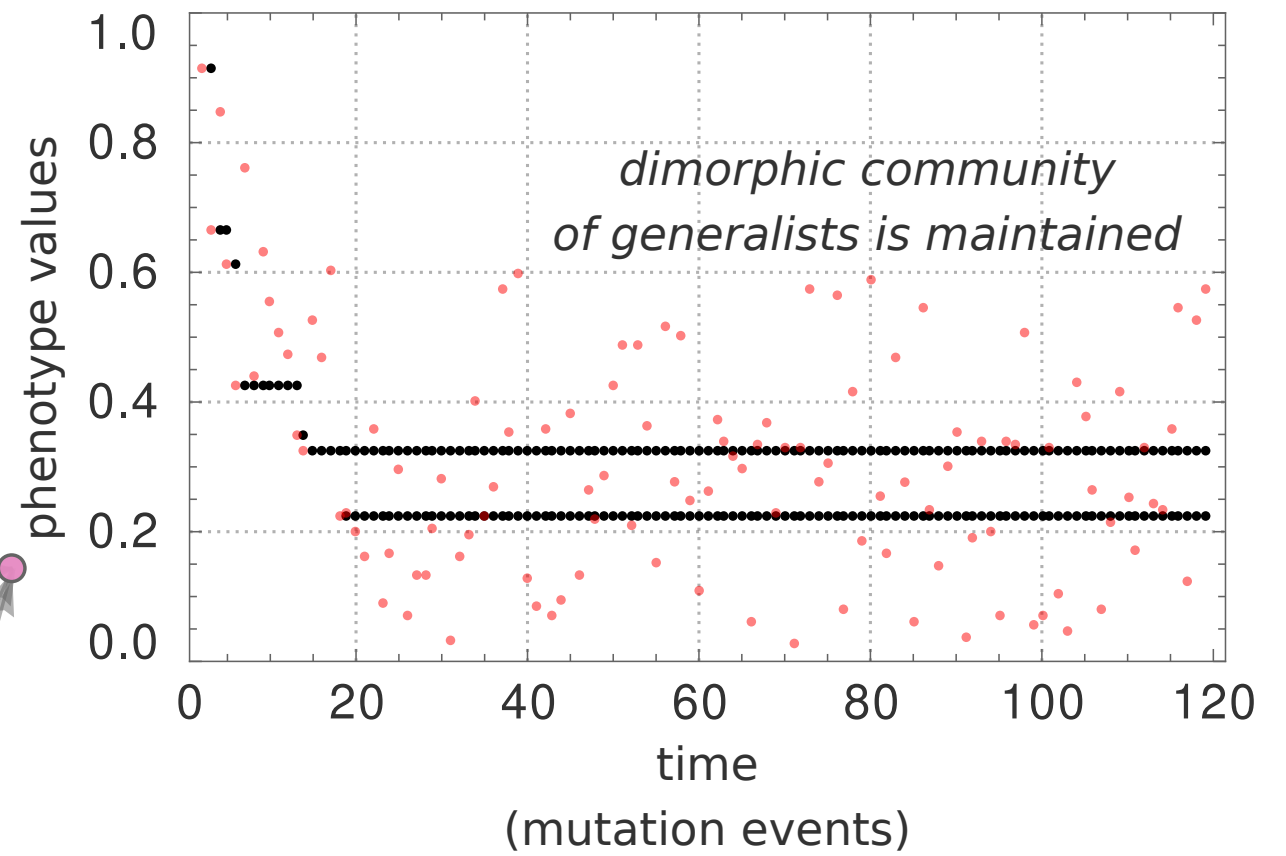
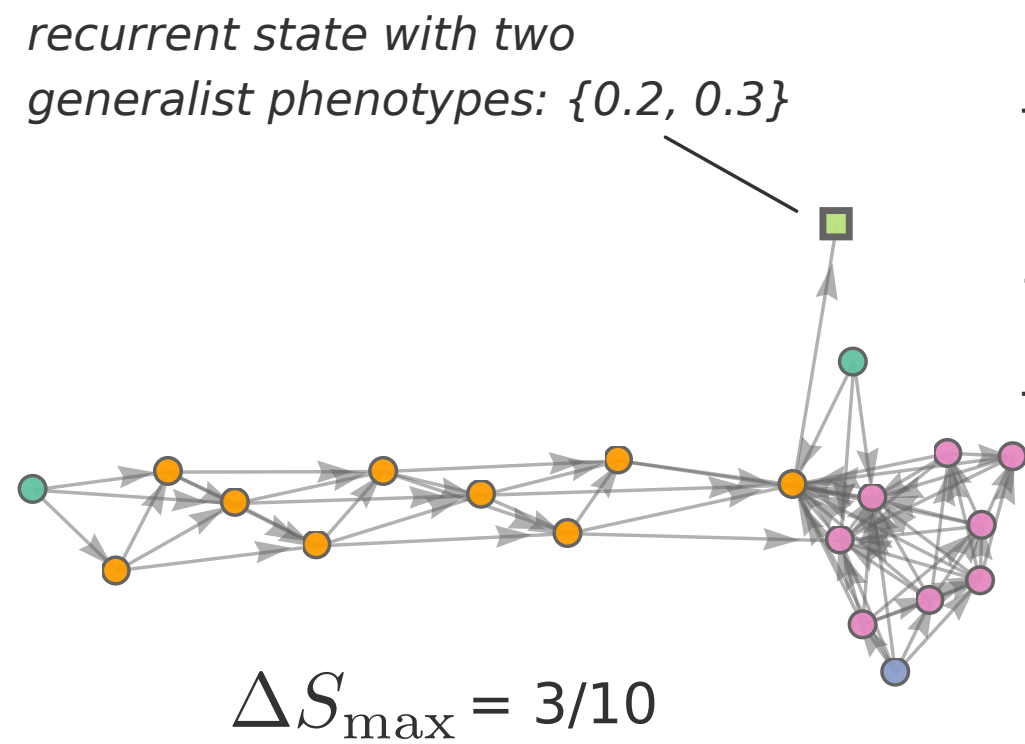
The recurrent states persist



Discrete phenotypes can however preserve some recurrent states

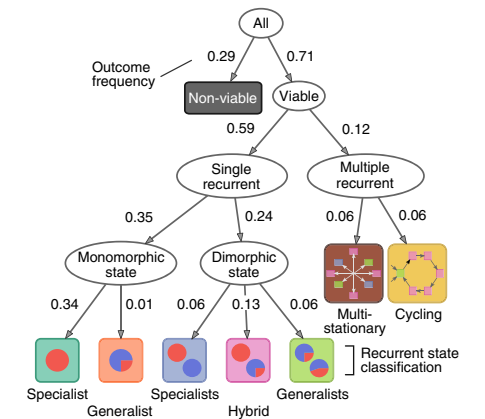
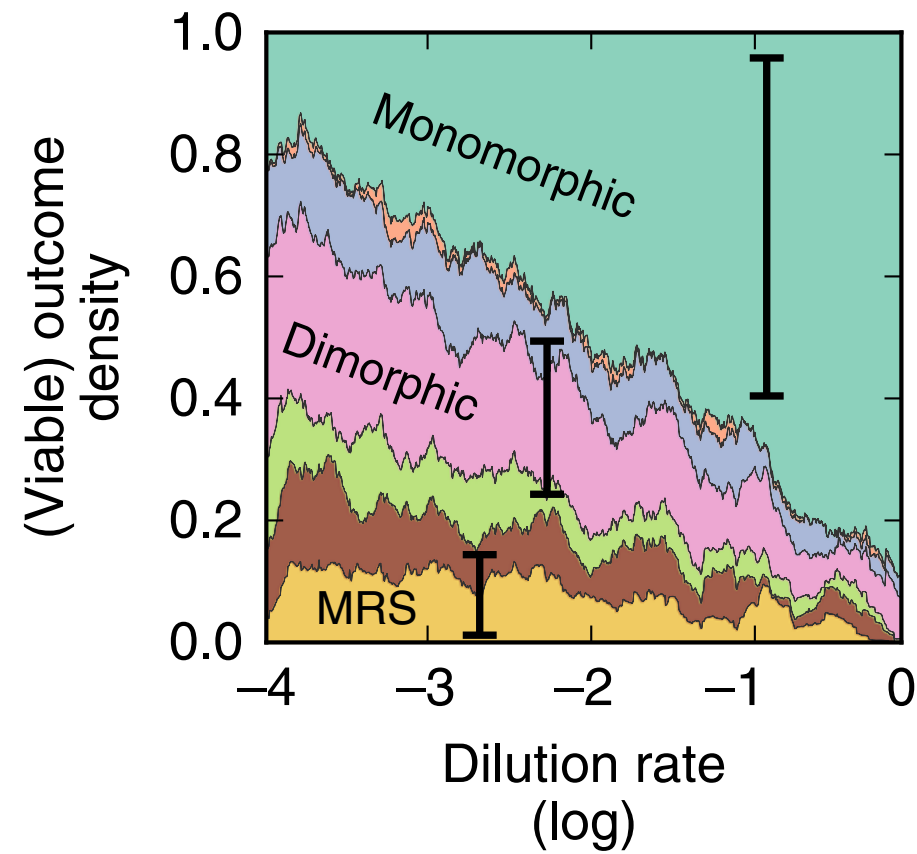


but need not

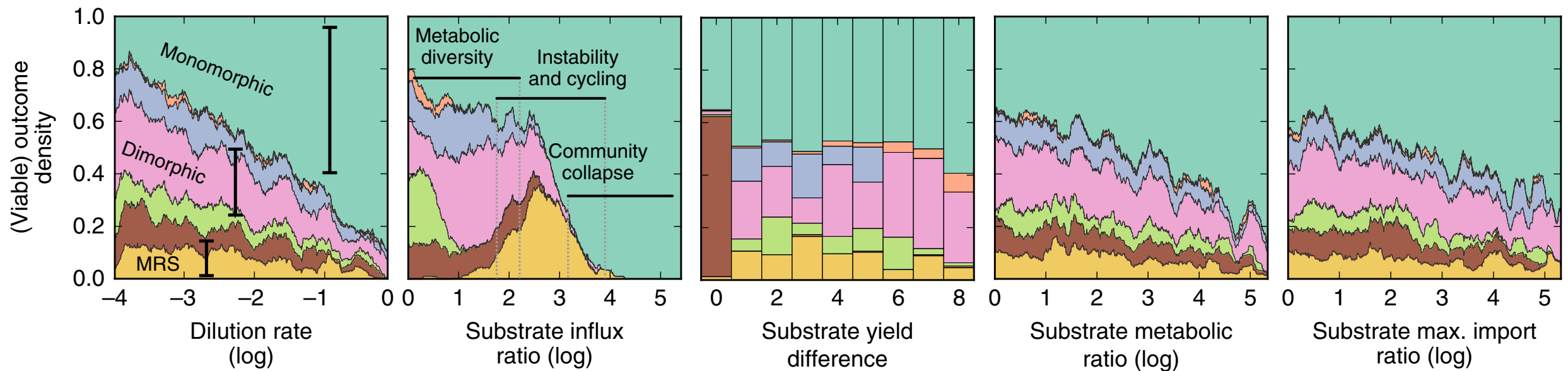
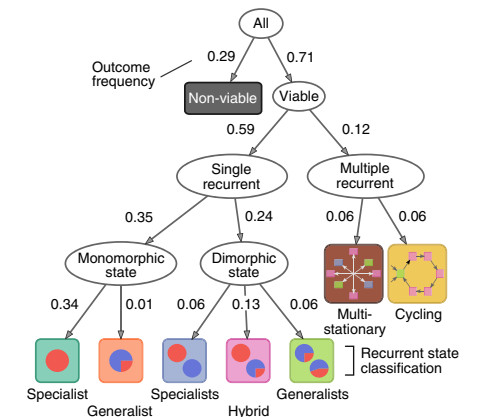


Is the type of recurrent state determined by the environment and nutrients?

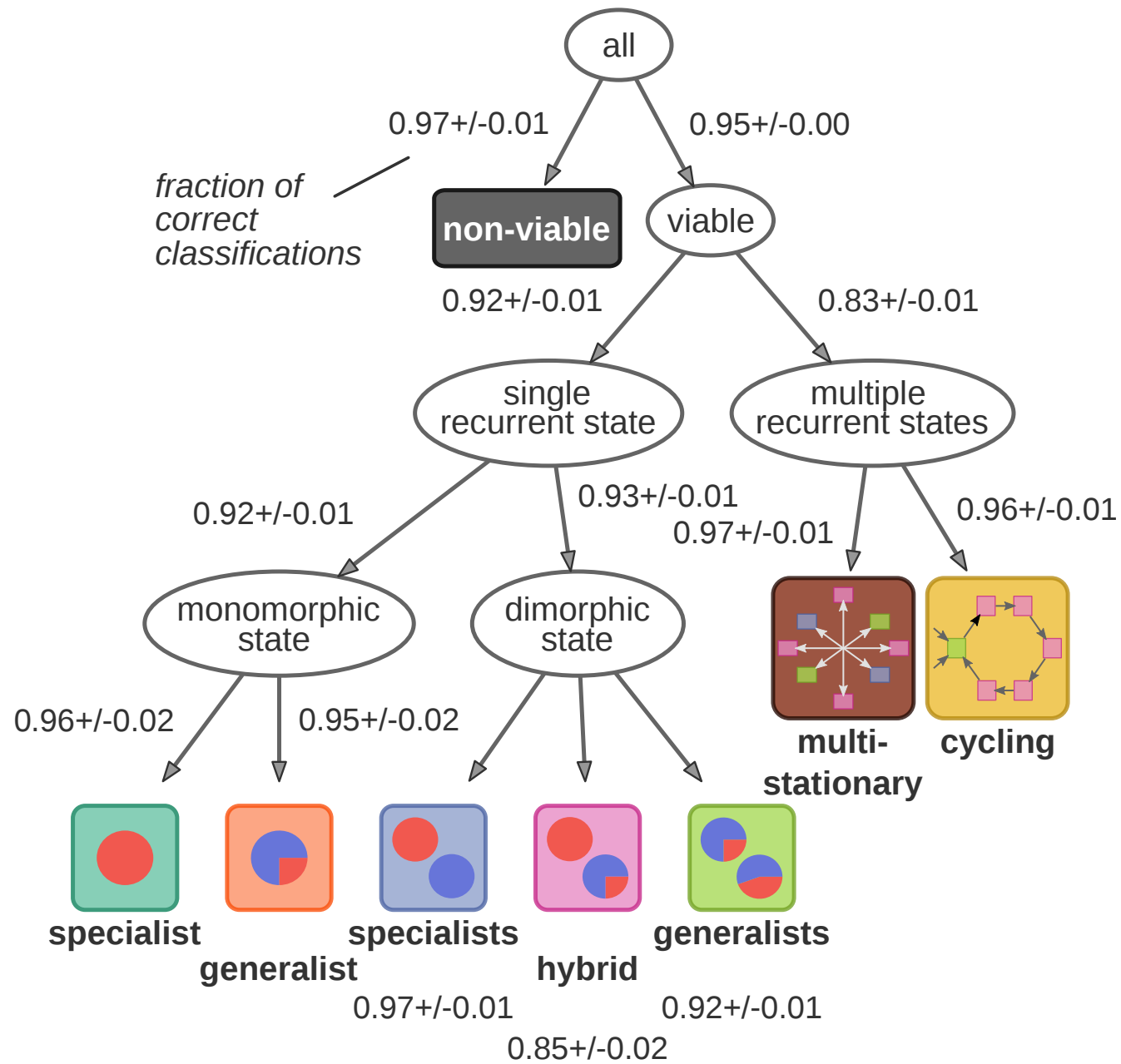
Fixing one parameter and varying the others, there is no clear pattern



Fixing one parameter and varying the others, there is no clear pattern

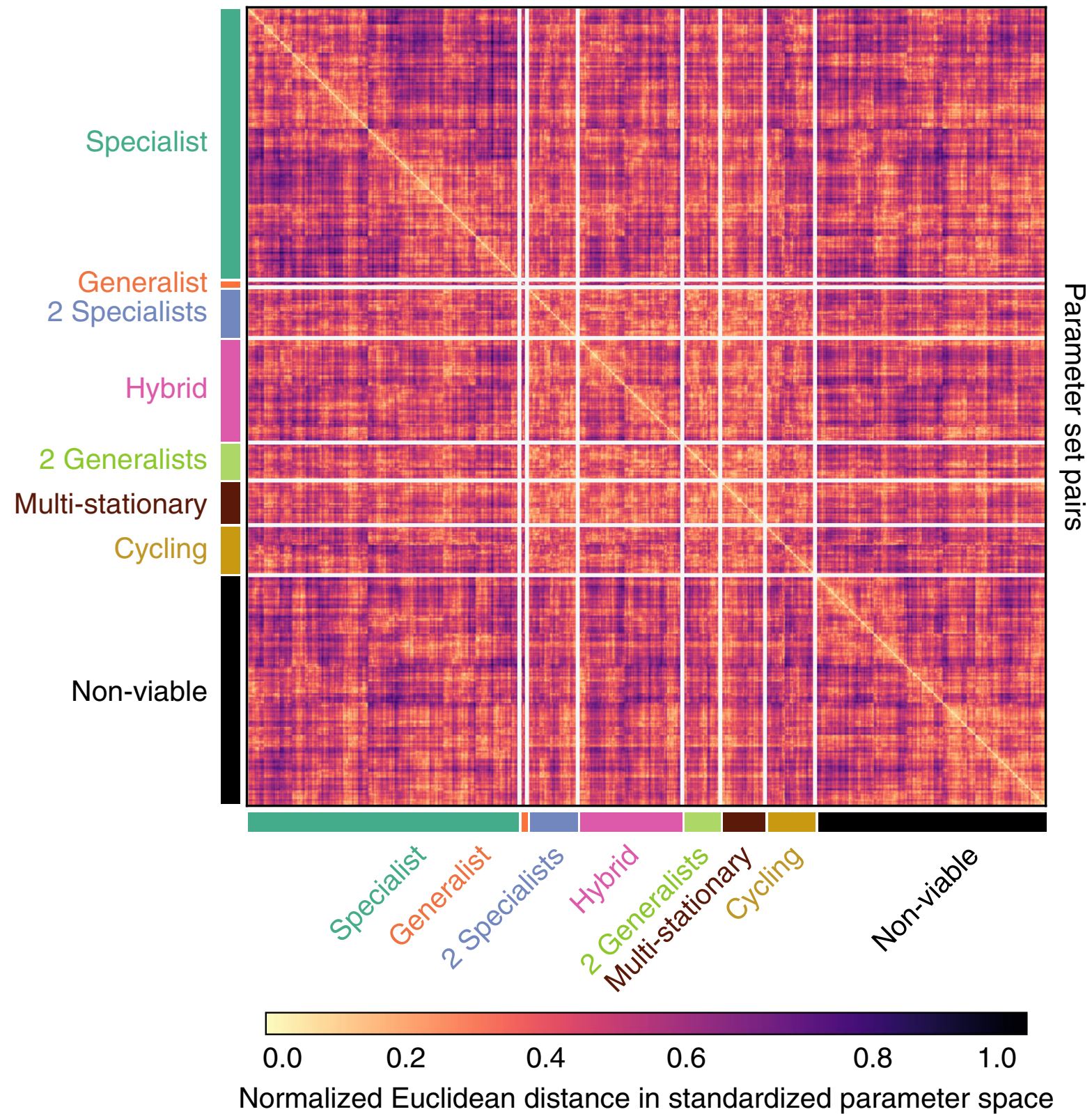


Even a hierarchical classification scheme performs poorly

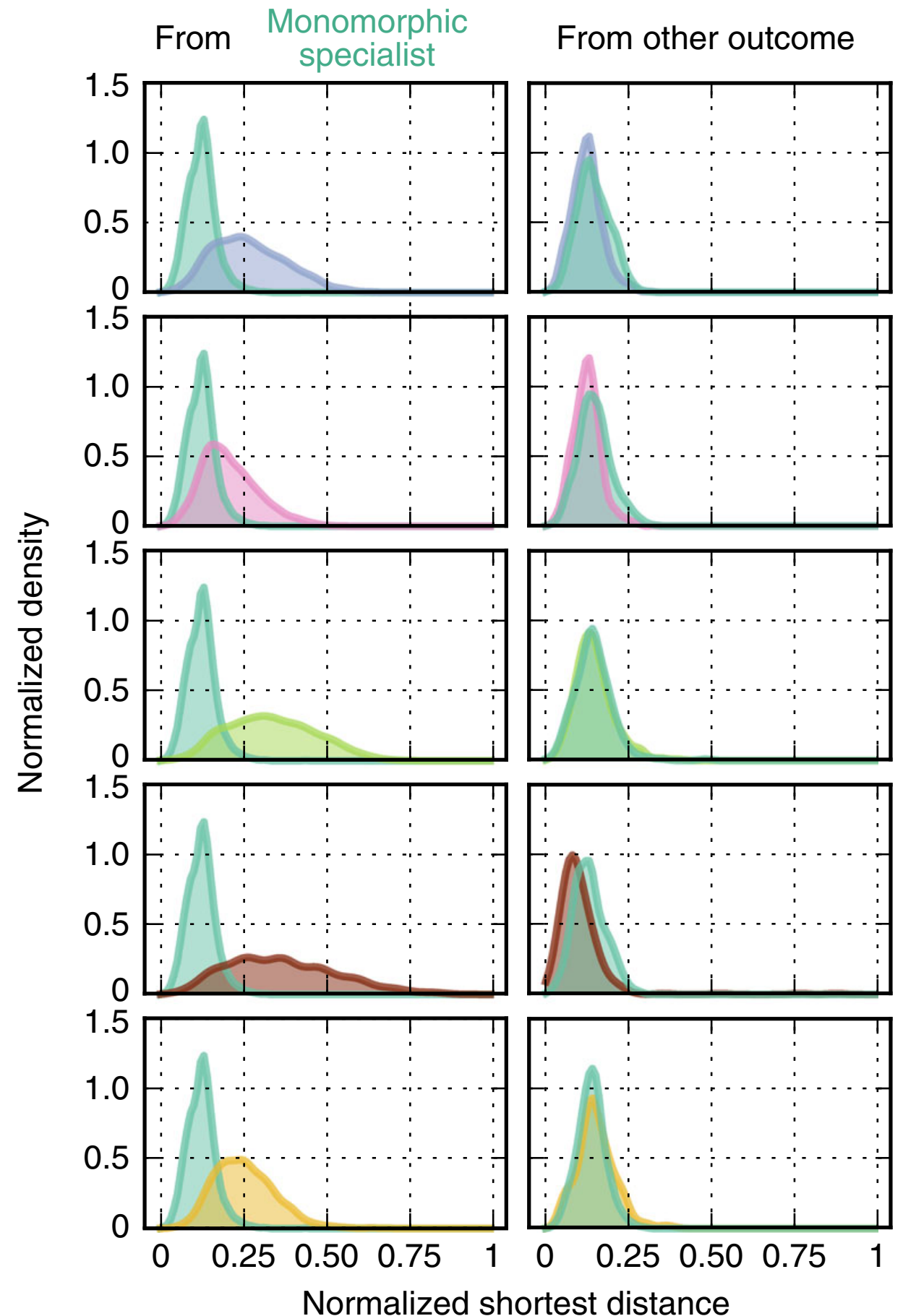


The average recall is 0.78.

The evolutionary outcomes form clusters, but these clusters are not distinct



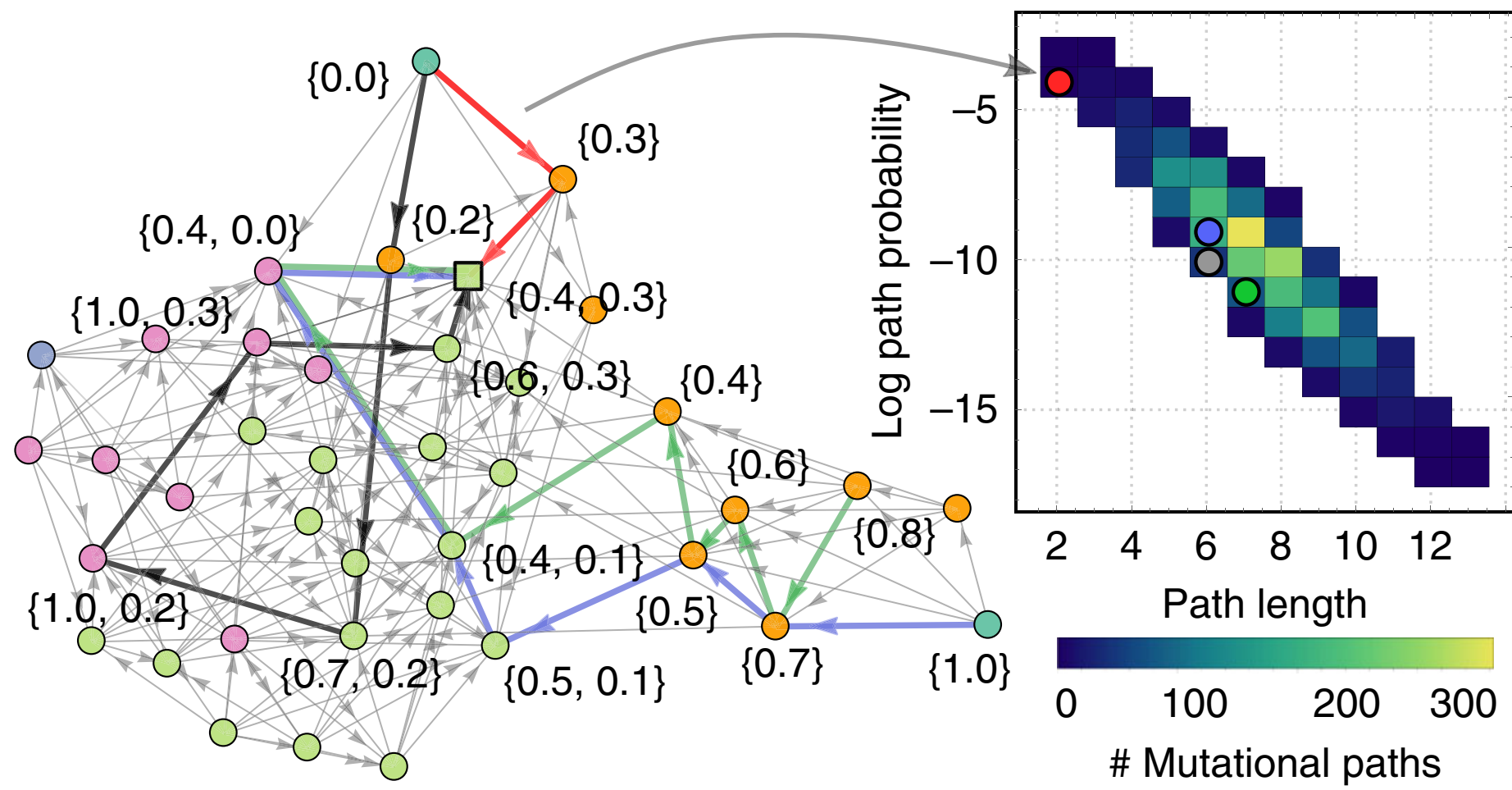
Monomorphic specialists permeate parameter space and are typically close to all other outcomes



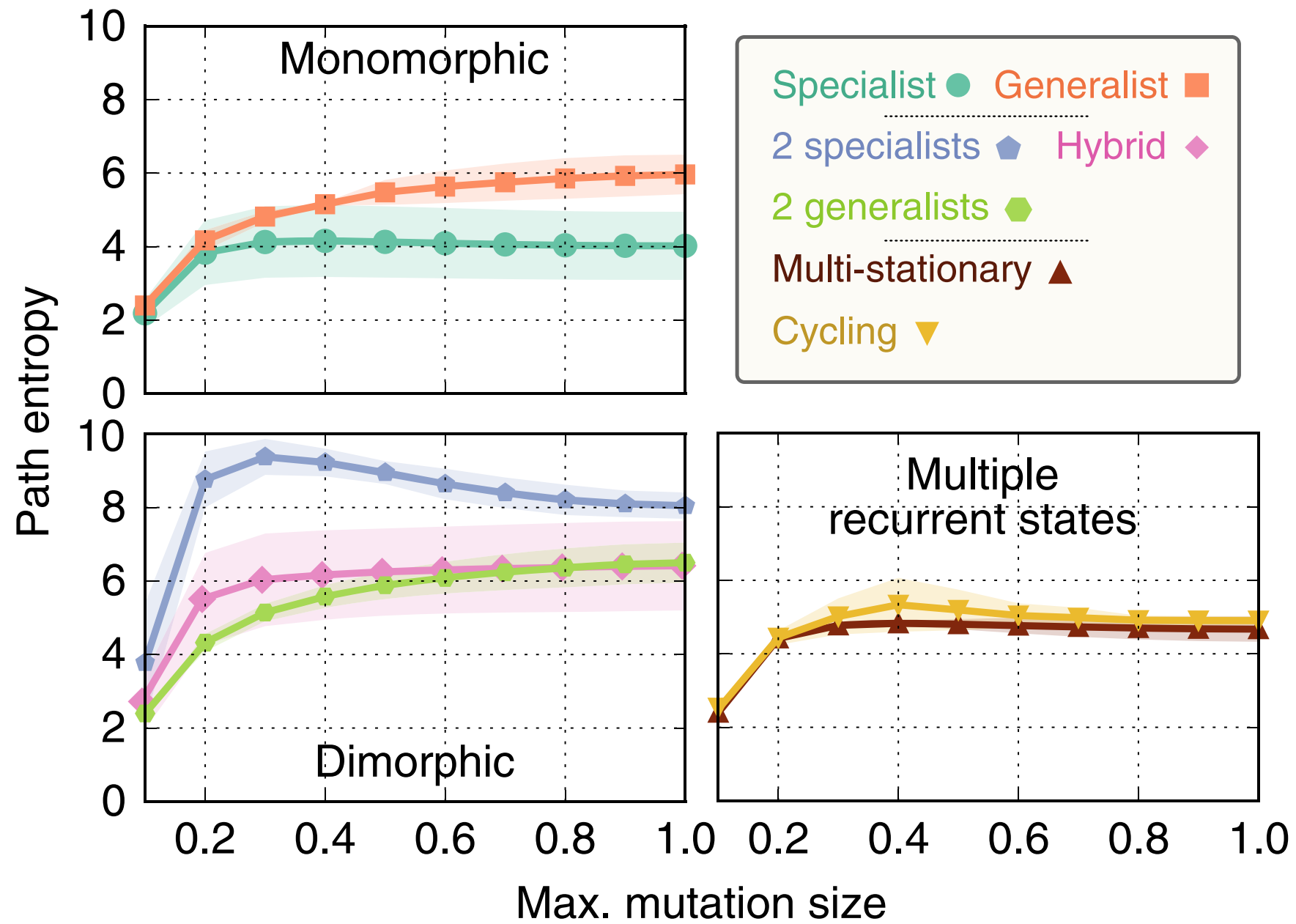
Environmental perturbations therefore are more likely to reduce rather than increase diversity.

Is the evolutionary outcome determined by the network's mutational paths?

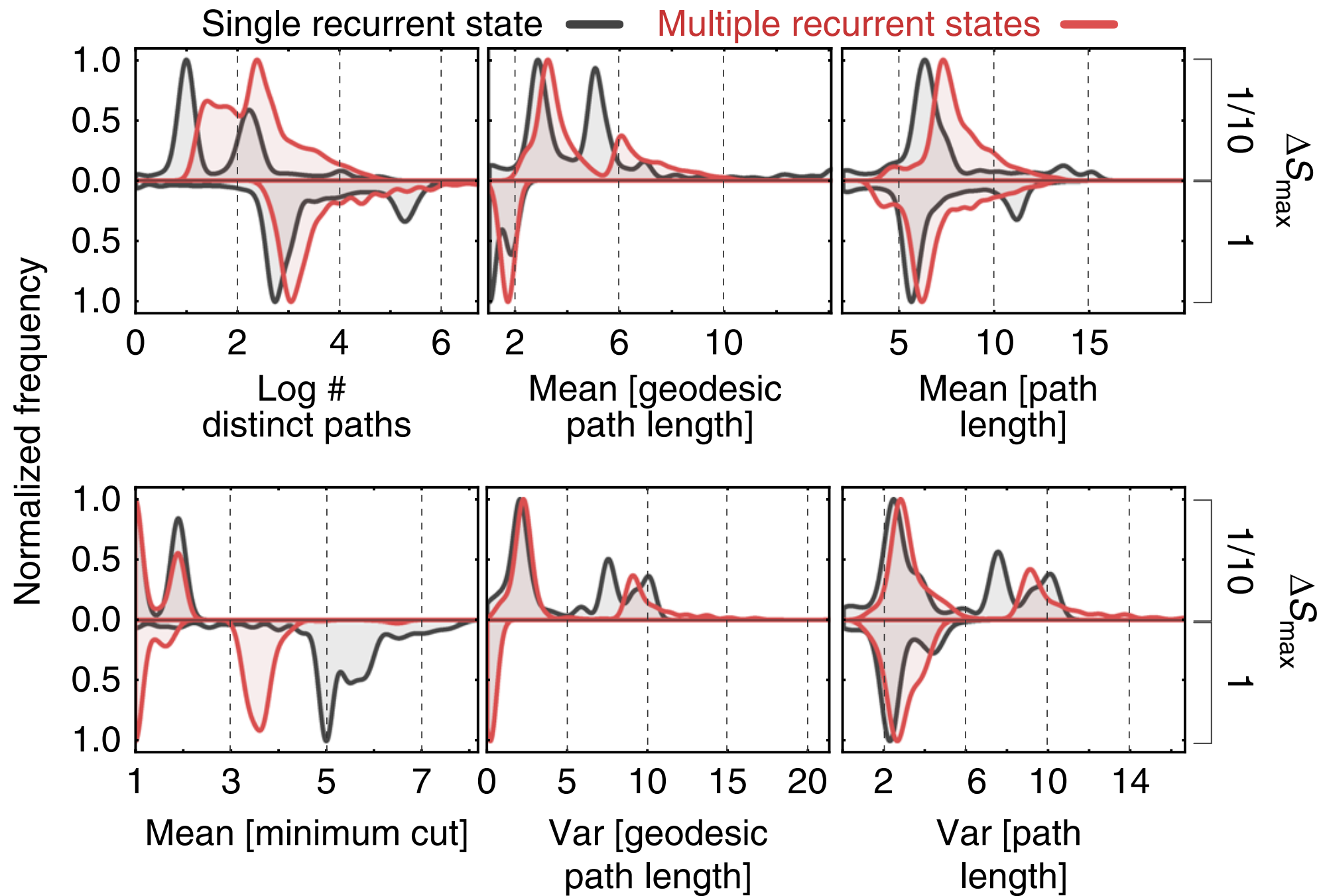
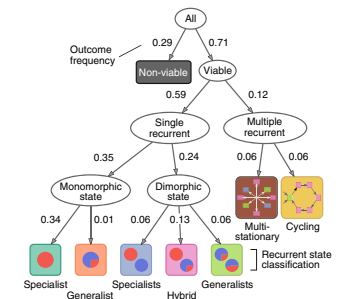
We characterize a network by its distribution of mutational paths



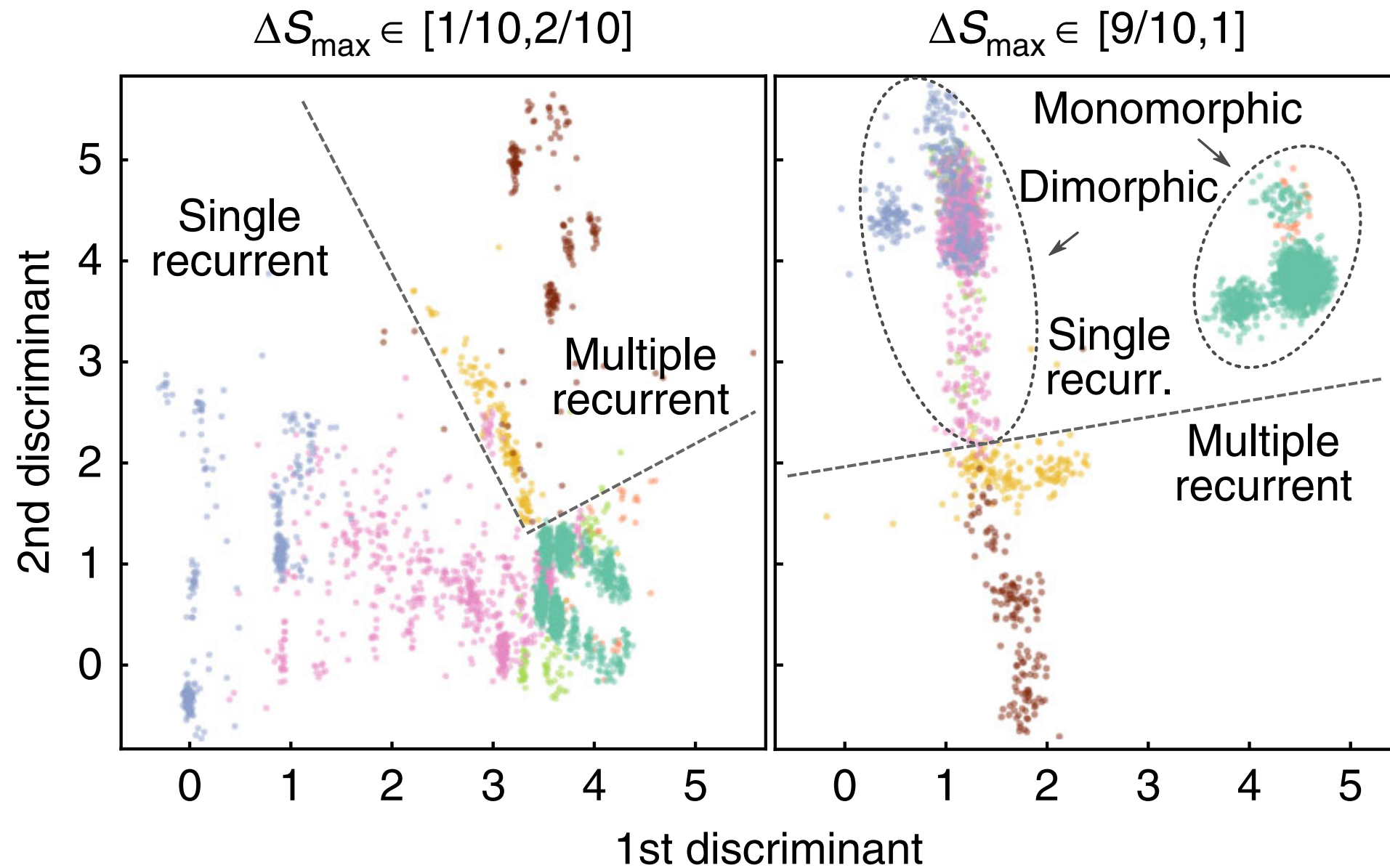
The repeatability of adaptation (the path entropy) depends on the evolutionary outcome



We use six properties to characterize the distribution of mutational paths

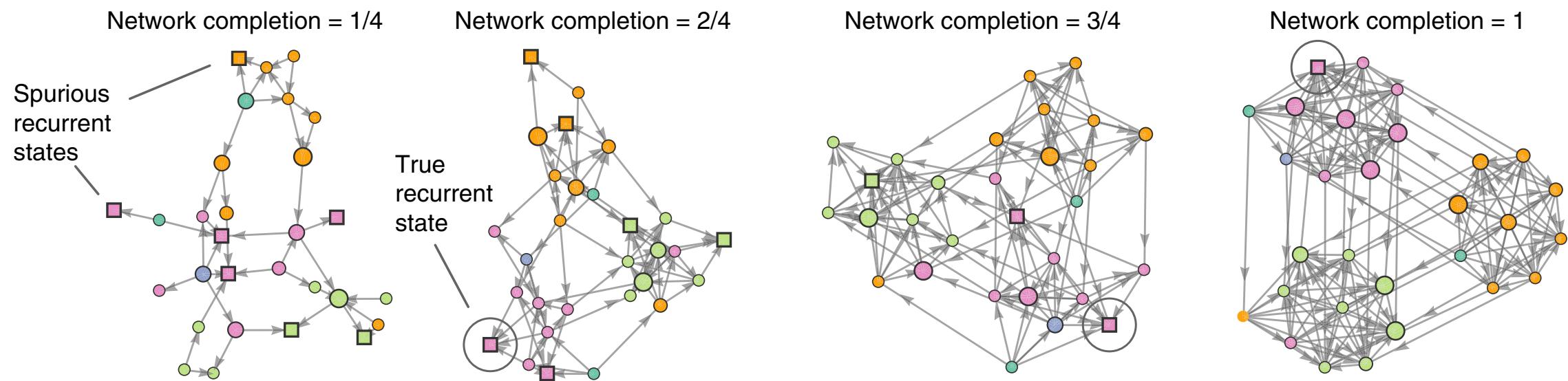


Combining the six properties of the mutational pathways, we can distinguish different evolutionary outcomes



We use a linear discriminant analysis.

Can we predict the evolutionary outcome of an incomplete network?



We consider six centralities for each vertex (a measure of its importance)

In-degree: number of edges arriving

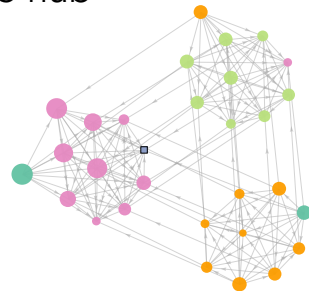
Out-degree: number of edges leaving

Closeness: reciprocal of the average shortest distance to all other vertices (potential of diversification)

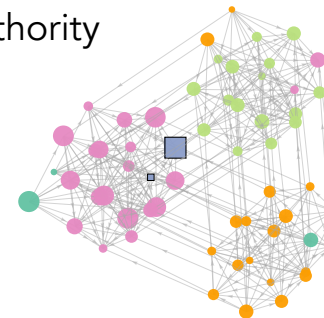
Betweenness: a measure of the number of shortest paths that pass through a vertex (potential of bottlenecking)

HITS hub & authority: hubs can be invaded by multiple mutants; authorities are the outcomes of multiple different invasion events

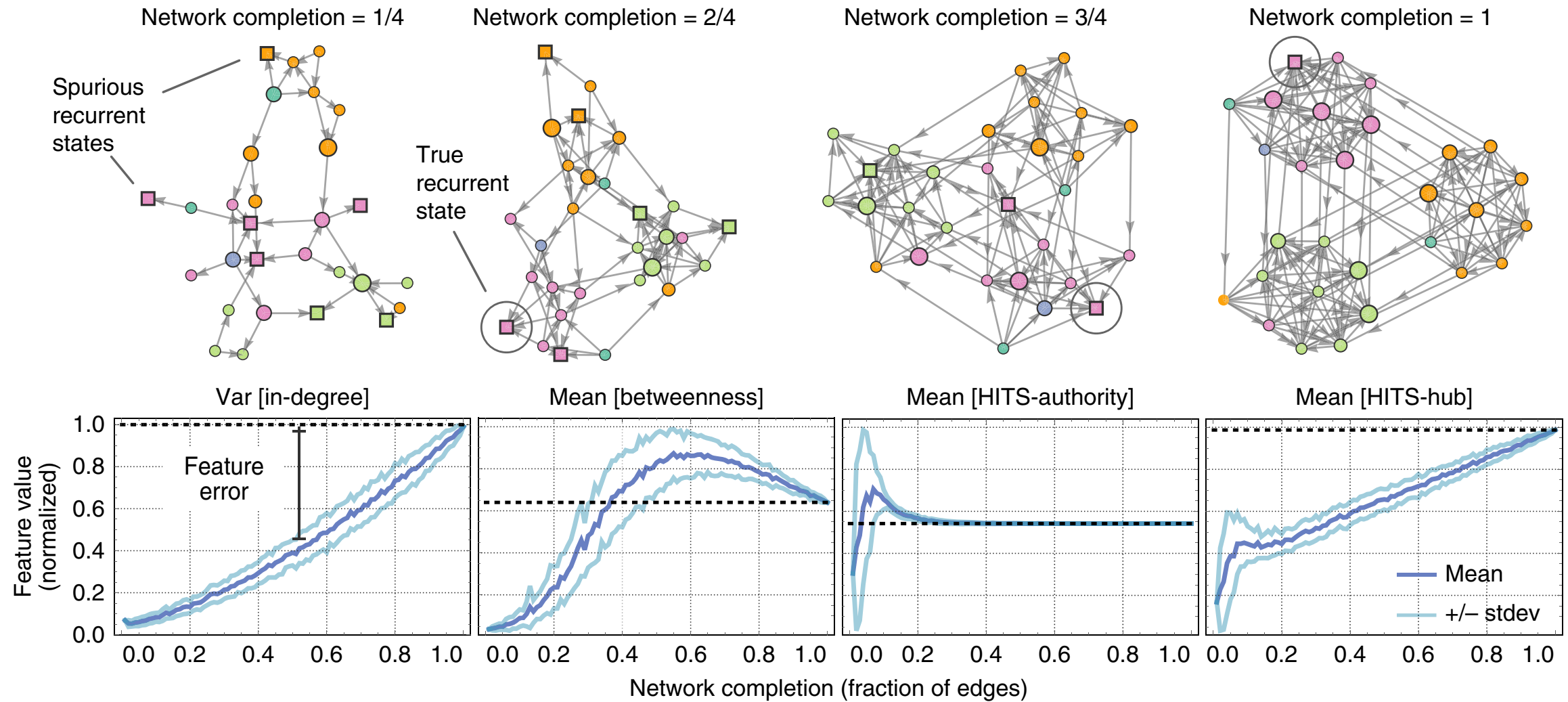
HITS hub



HITS authority



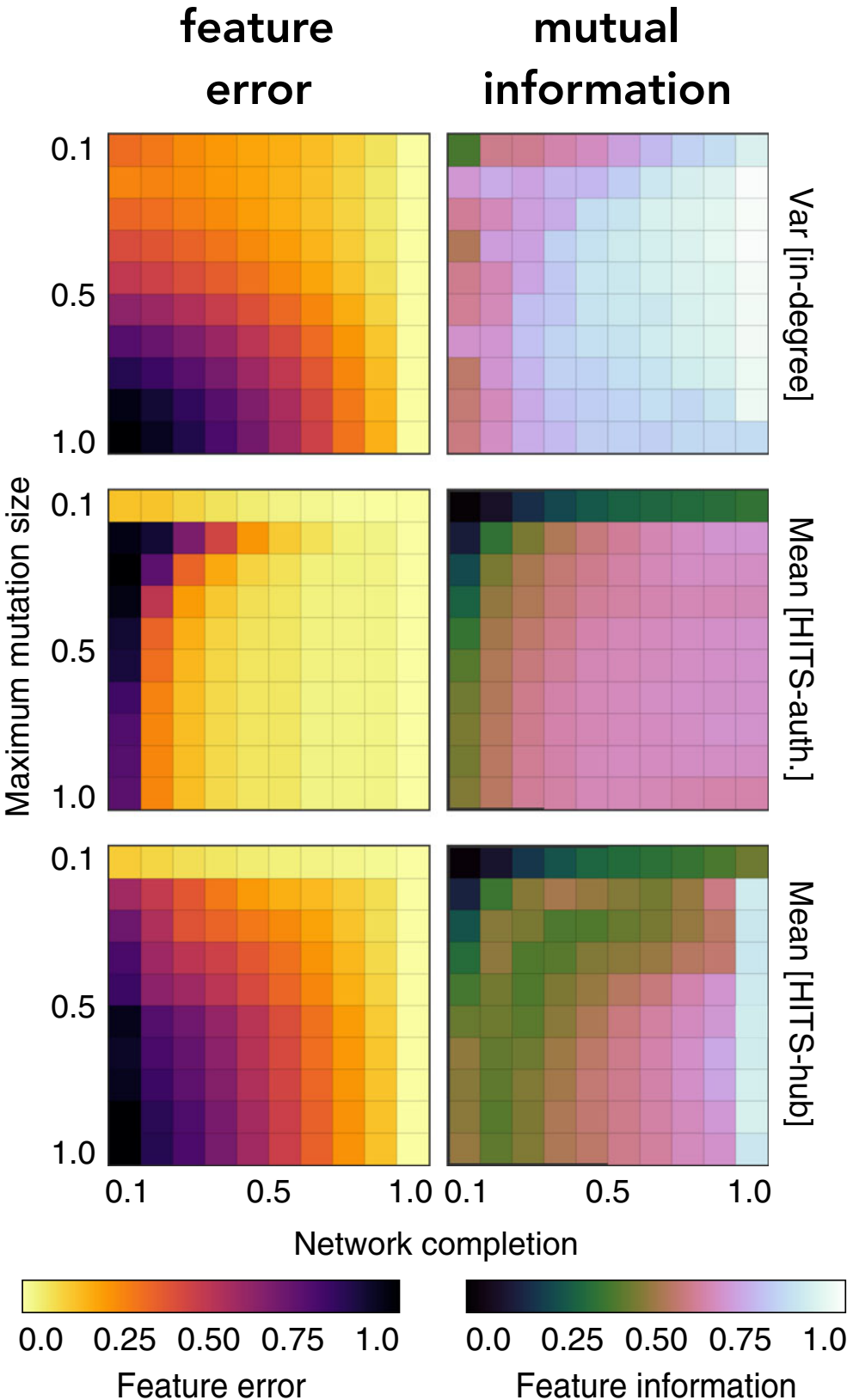
Changes in the network's topology are reflected in the statistics of the centralities



A good statistic converges quickly and is discriminatory

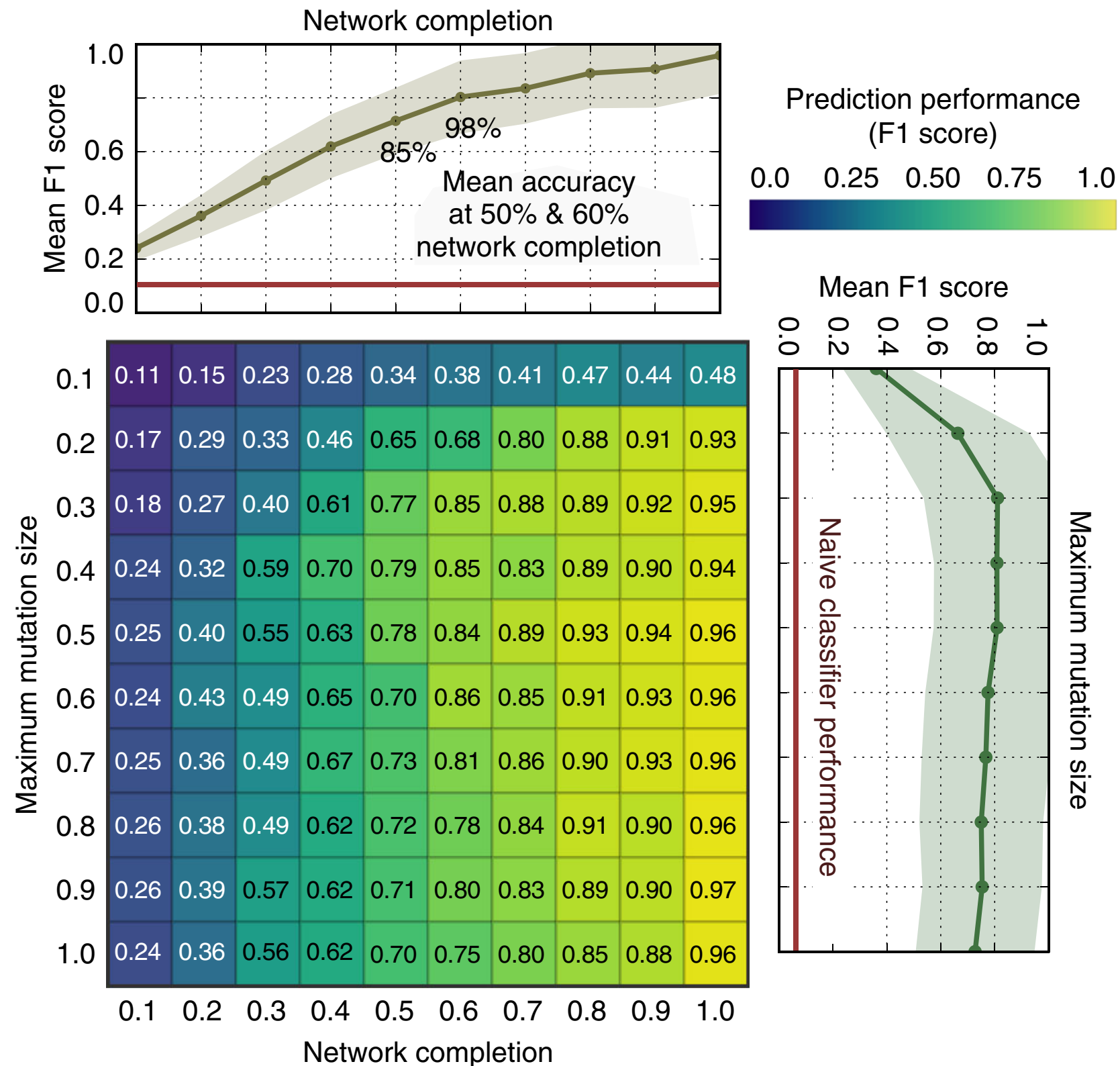
Three examples:

Mean of the feature error
(the distance between the statistic's current value and its value on completion of the network)



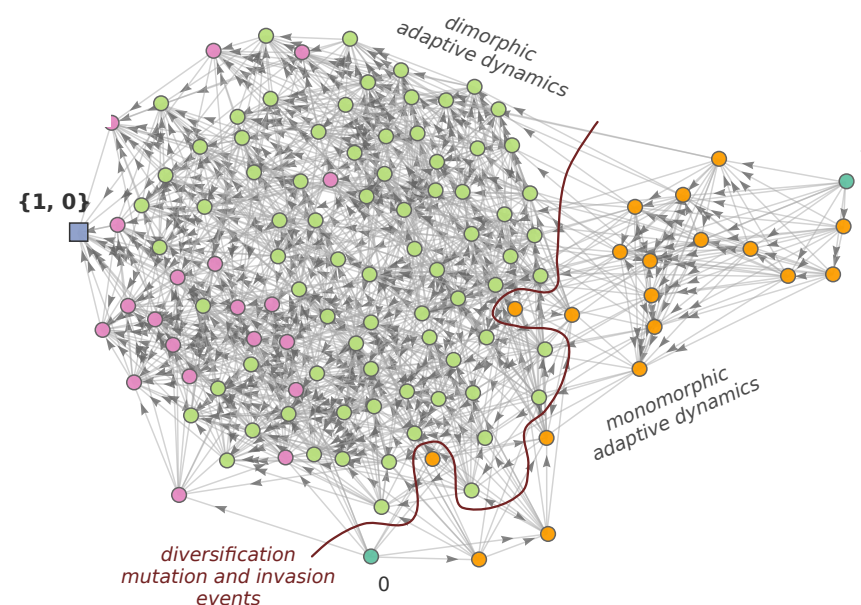
Normalized mutual information between a statistic and the evolutionary outcomes

Evolutionary outcomes can be predicted from incomplete networks



Summary

- We develop an eco-evolutionary model with a trade-off that has surprisingly diverse evolutionary outcomes
- Working with the networks of mutational paths is insightful
- Evolutionary outcomes cannot be determined from environmental conditions, but can be determined from their networks of mutational paths even for incomplete networks
- The dynamics of adaptation is a key variable for predicting long-term evolutionary behaviours



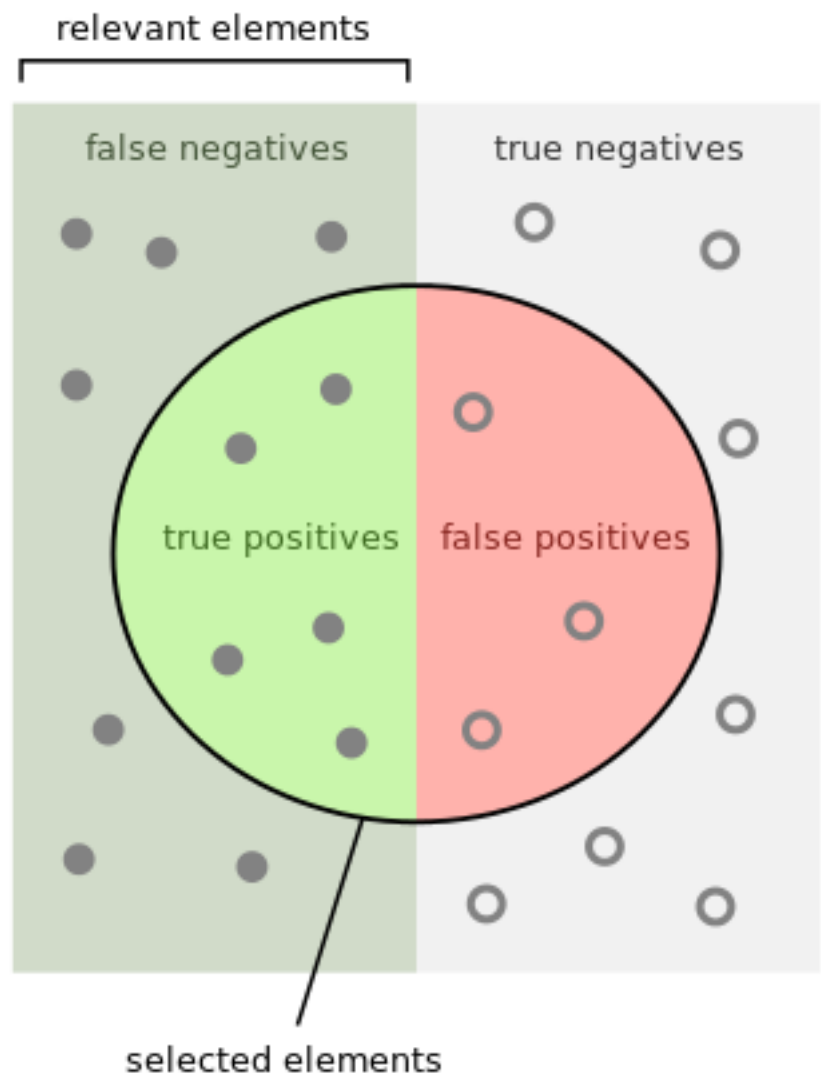
swainlab.bio.ed.ac.uk

Josephides & Swain,
Nat Commun 2017

Chris Josephides

Luke McNally
Nikola Popovic





How many selected items are relevant?

$$\text{Precision} = \frac{\text{true positives}}{\text{true positives} + \text{false positives}}$$

How many relevant items are selected?

$$\text{Recall} = \frac{\text{true positives}}{\text{true positives} + \text{false negatives}}$$