Using cellular automata to uncover evolutionary landscapes

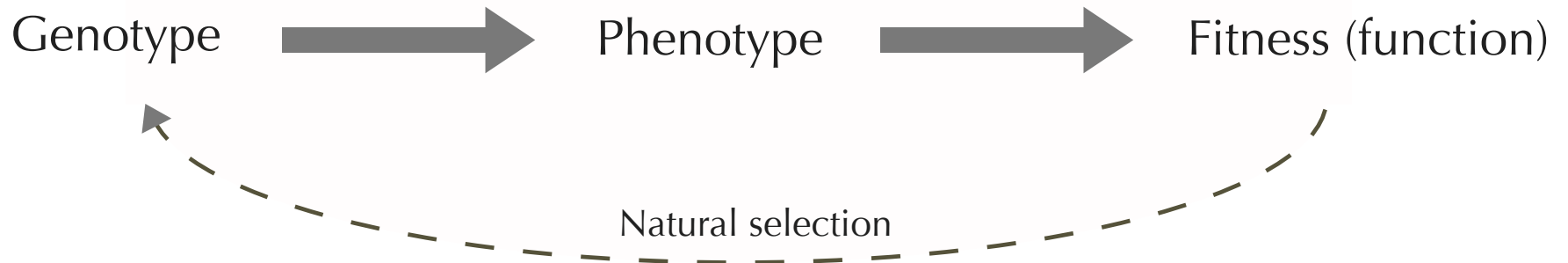Trevor Bedford   Hartl Laboratory   Harvard University

The origins of variation

Dynamics vs. kinetics in evolutionary theory

Evolution in a nutshell

Genotype $\longrightarrow$ Phenotype $\longrightarrow$ Fitness (function)

Natural selection

Most of evolutionary theory has concerned itself with the dynamics of evolution -- what is the probability that a new mutant with fitness X takes over in the population?

However, very little is known the kinetics of evolution -- what are the chances of that mutant of fitness X appearing in the first place?

Genetic neighborhoods

Mutation gives rise to genetic distance

Wild-type sequence   0 1 0 0

One-step neighbors   1 1 0 0   0 0 0 0   0 1 1 0   0 1 0 1
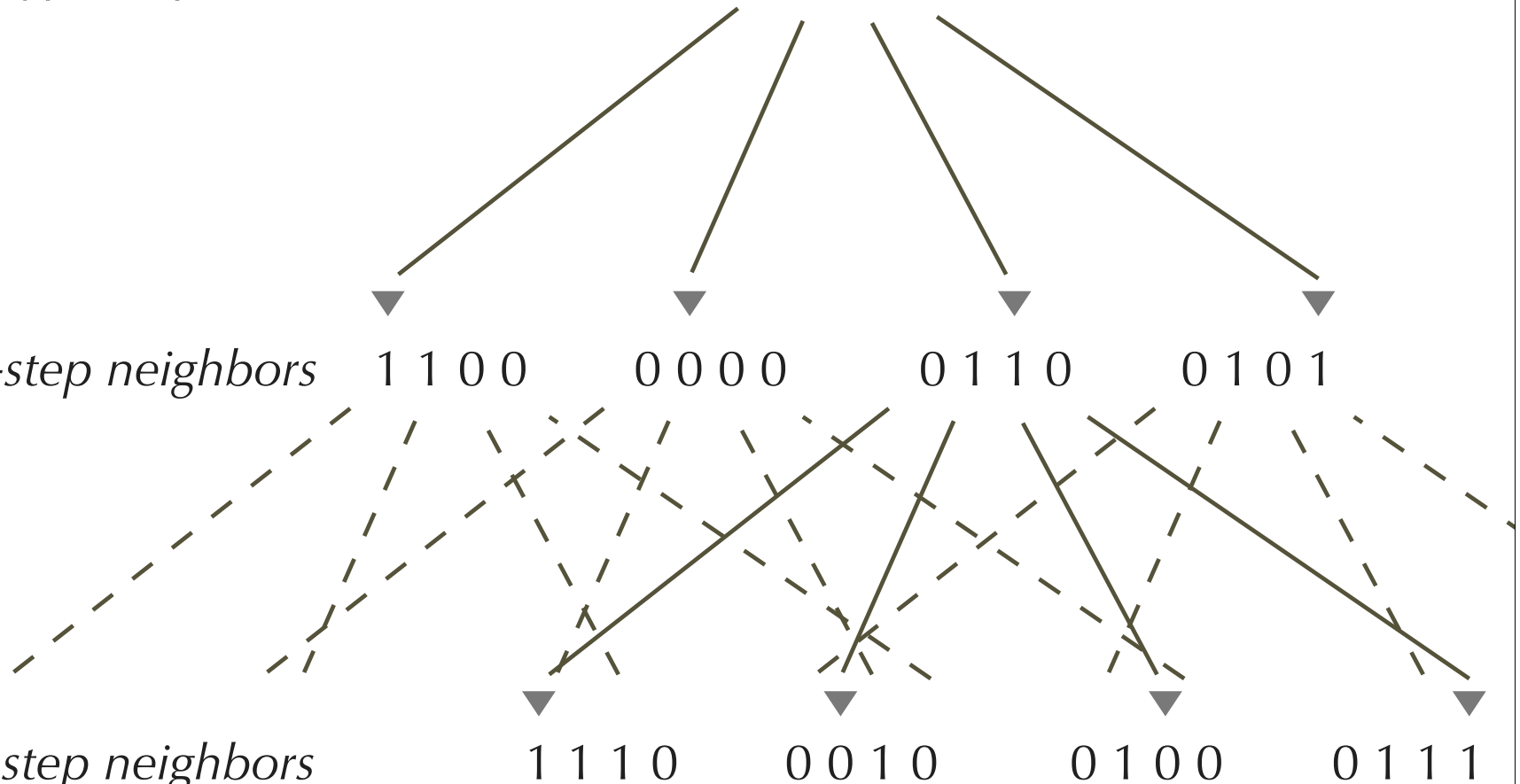
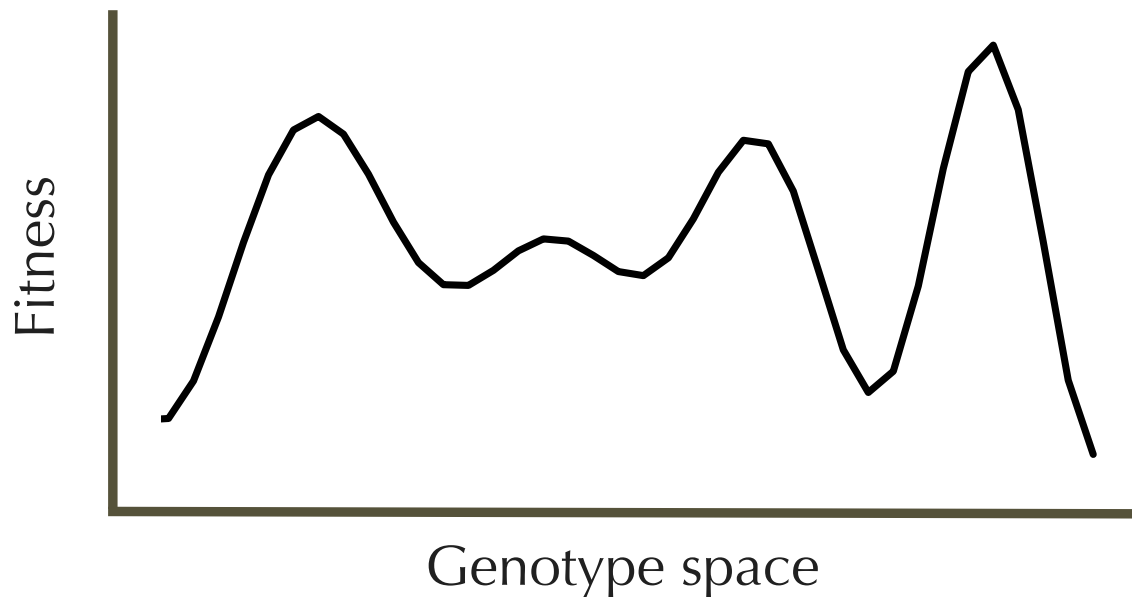Two-step neighbors   1 1 1 0   0 0 1 0   0 1 0 0   0 1 1 1

Fitness landscapes

## Genotype to fitness mapping results in 'fitness landscapes'

Visual metaphor introduced by Sewell Wright -- natural selection pushes a population toward higher regions of the landscape

Fitness

Genotype space

This landscape has only one dimension, actual genic landscapes have thousands of dimensions

Genotype space in CA

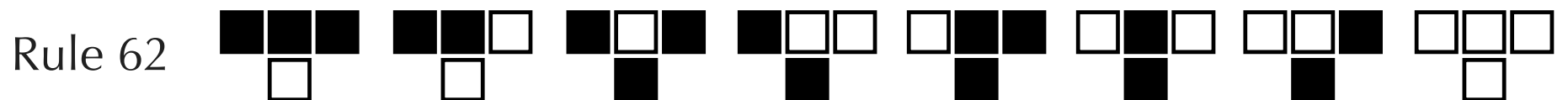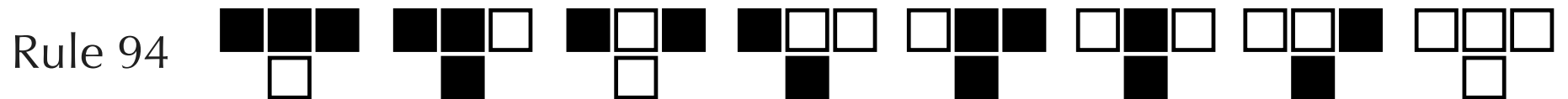## One-step neighborhood of an elementary CA
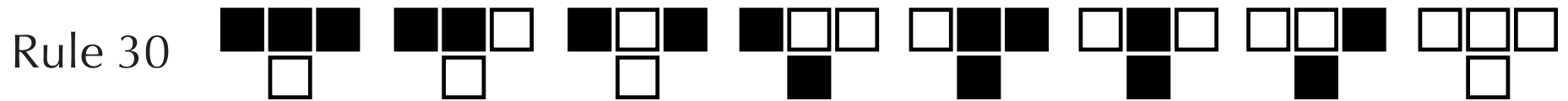
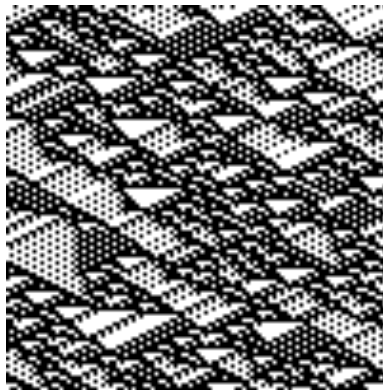# Genotype ⇒ phenotype mapping

Rules that are genetically similar tend to have similar phenotypes

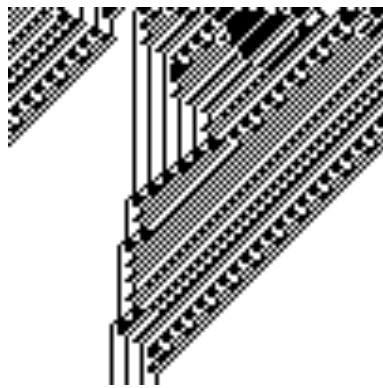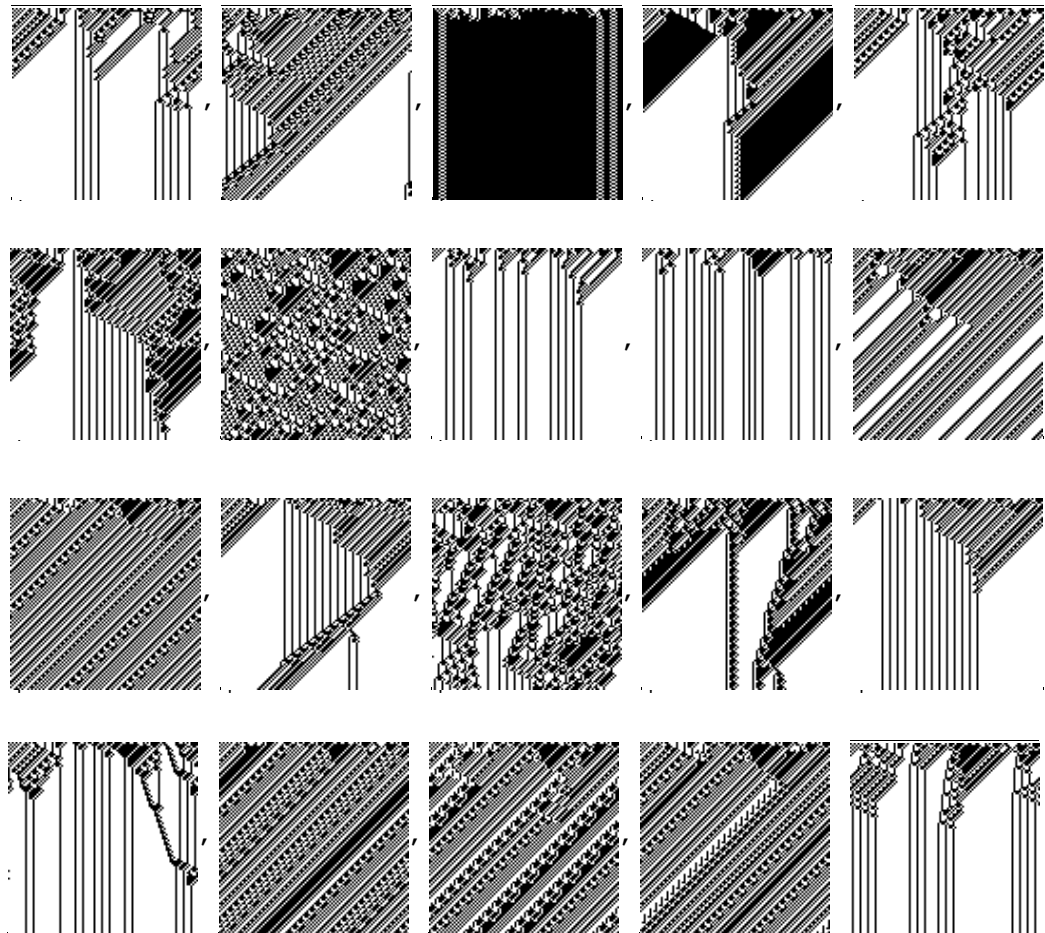One-step neighbors

Wild-type

r=2 2-color CA
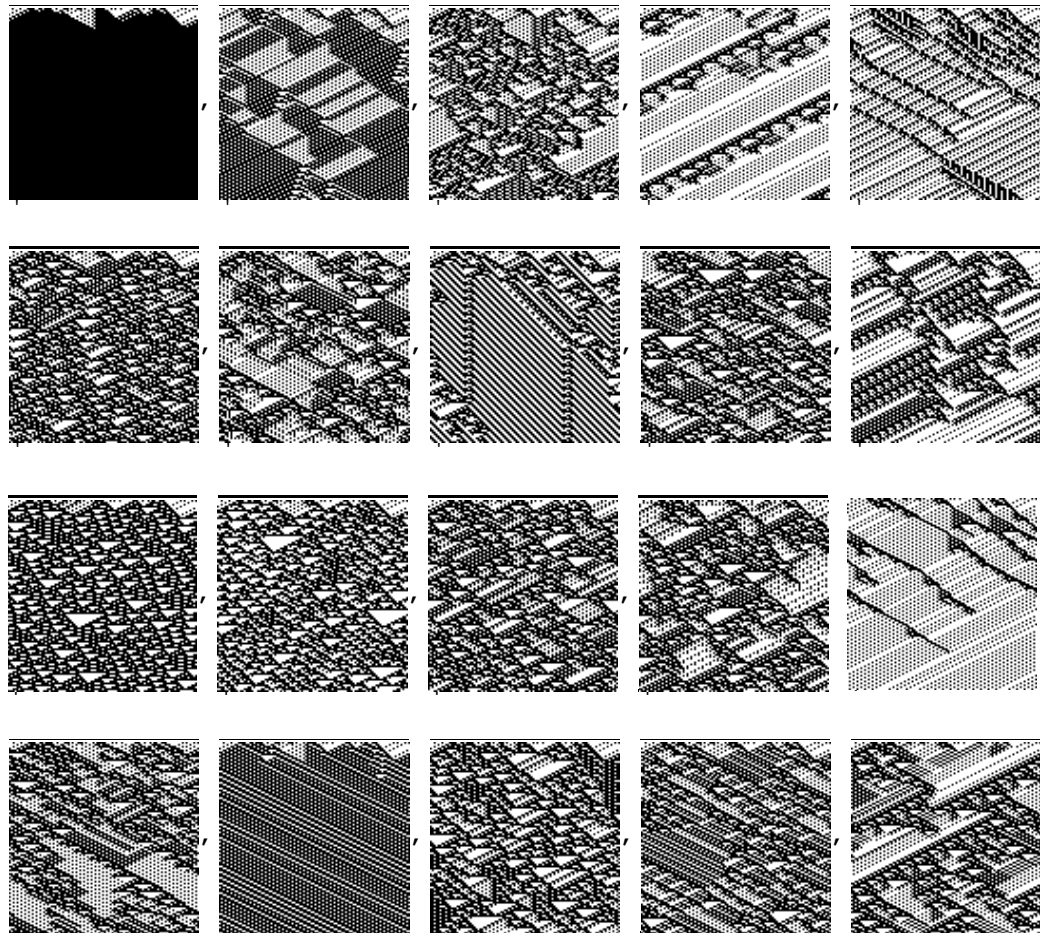
2^5 = 32 neighbors

Genotype ⇒ phenotype mapping

Rules that are genetically similar tend to have similar phenotypes

One-step neighbors

Wild-type

r=2 2-color CA

2^5 = 32 neighbors

Genotype ⇒ phenotype mapping

Rules that are genetically similar tend to have similar phenotypes

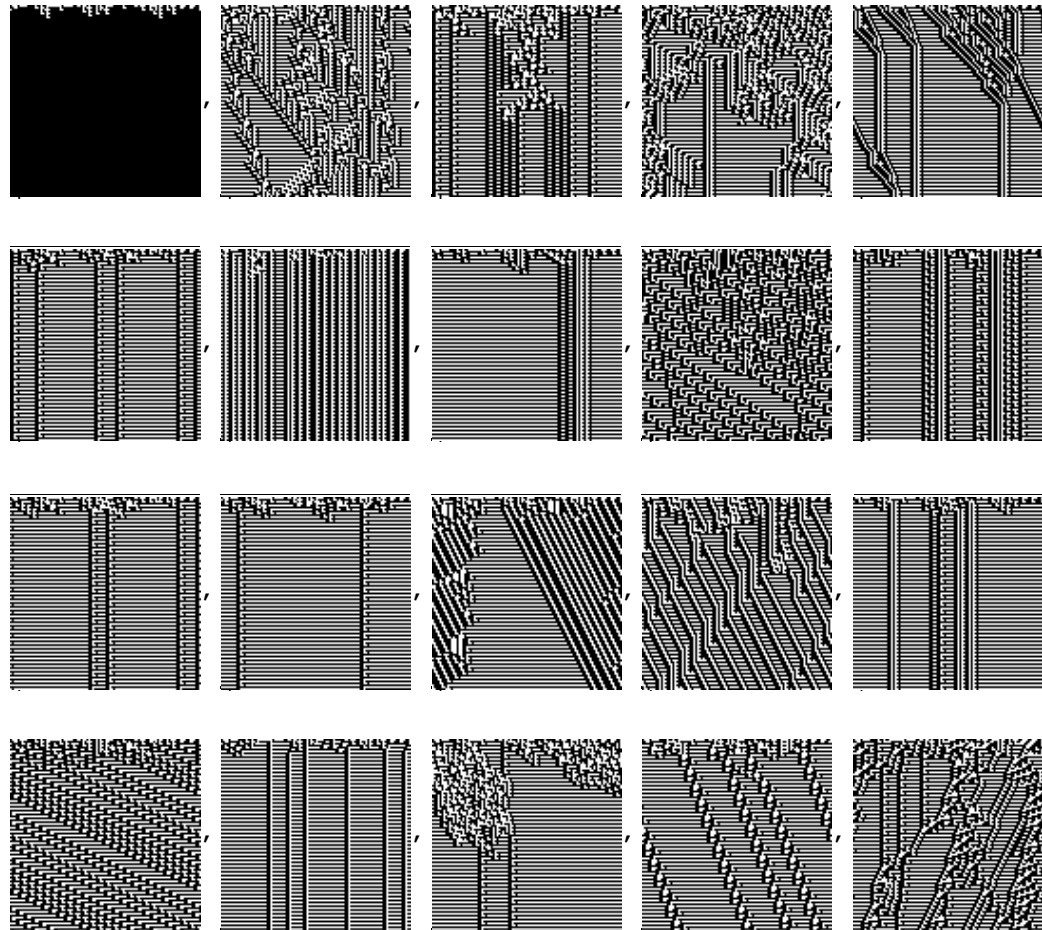One-step neighbors

Wild-type

r=2 2-color CA

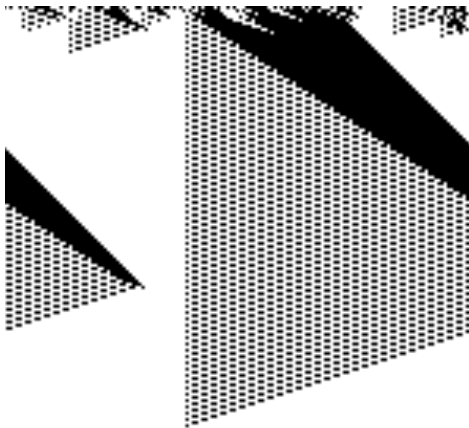2^5 = 32 neighbors

Density classification CA

Assays function rather than form
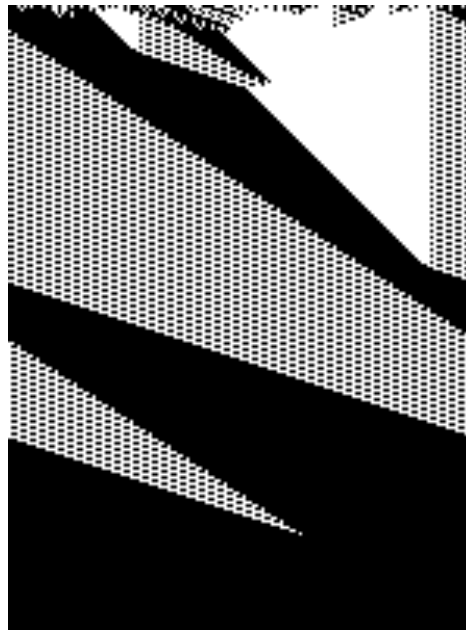
Fitness as proportion of initial conditions correctly assigned

$r = 3$ 2-color cellular automata -- 128 one-step neighbors

Correctly identifies initial density 74.4% of the time -- fitness = 0.744
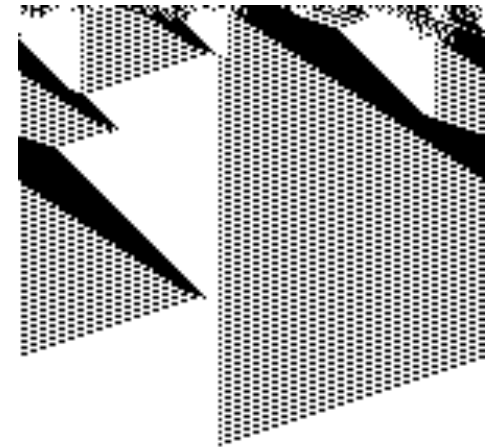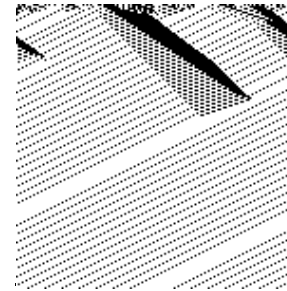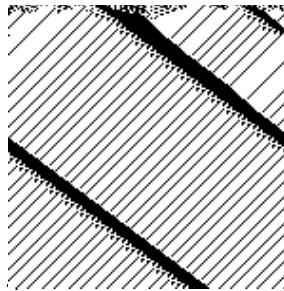

Rule 340,281,450,309,255,942,604,150,056,210,657,181,704 discovered by Crutchfield and Mitchell (1996)

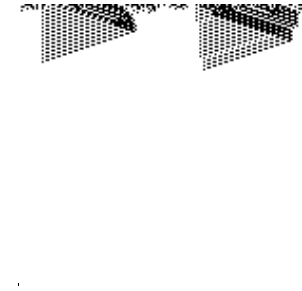# Alternate phenotypes
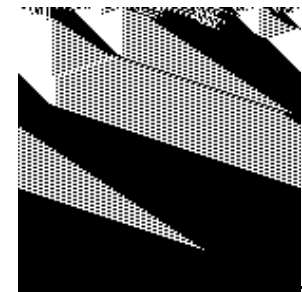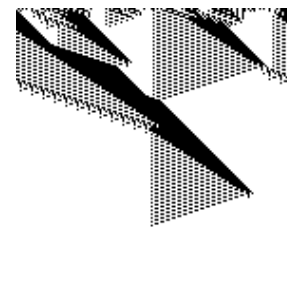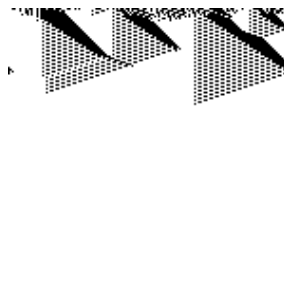
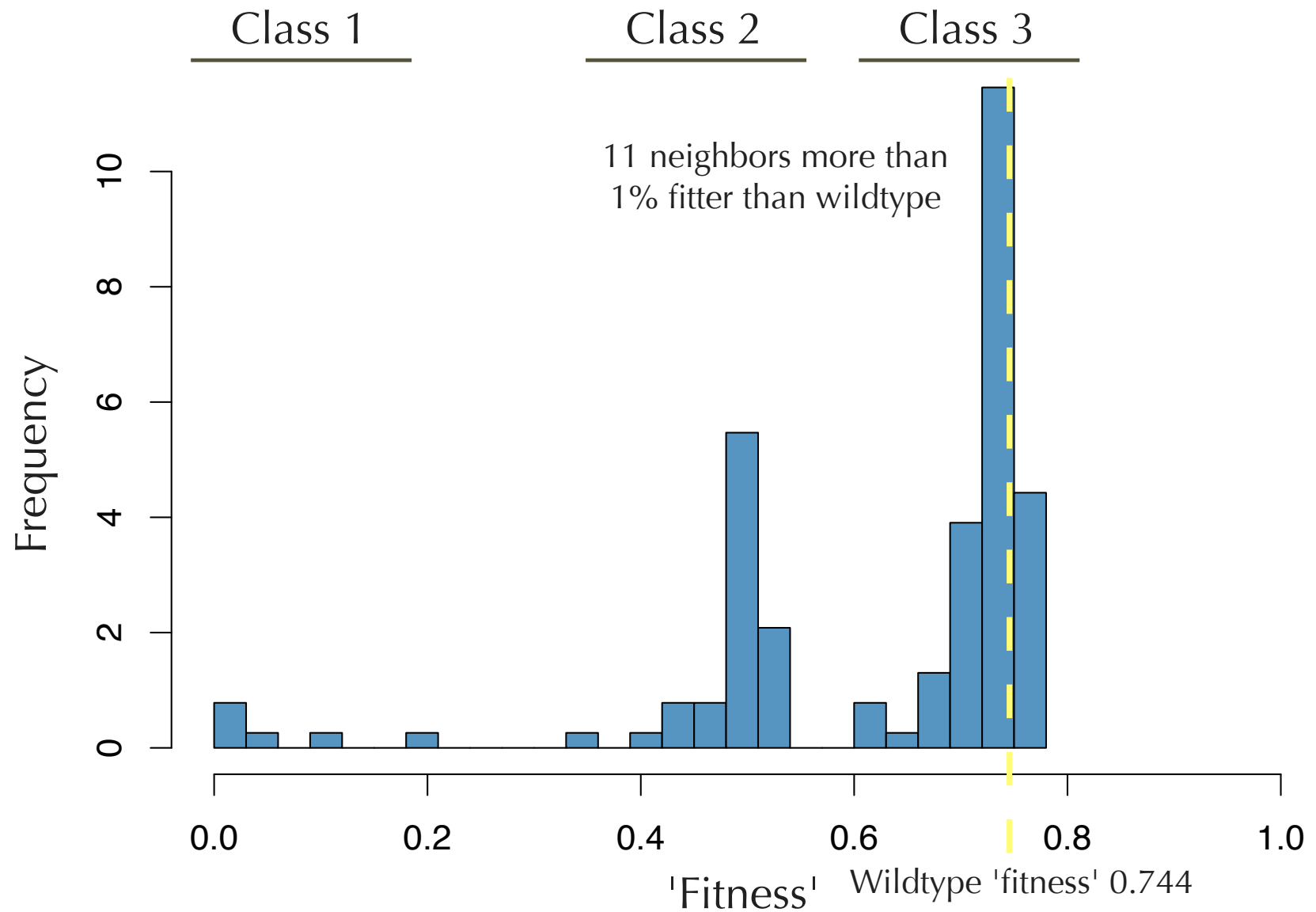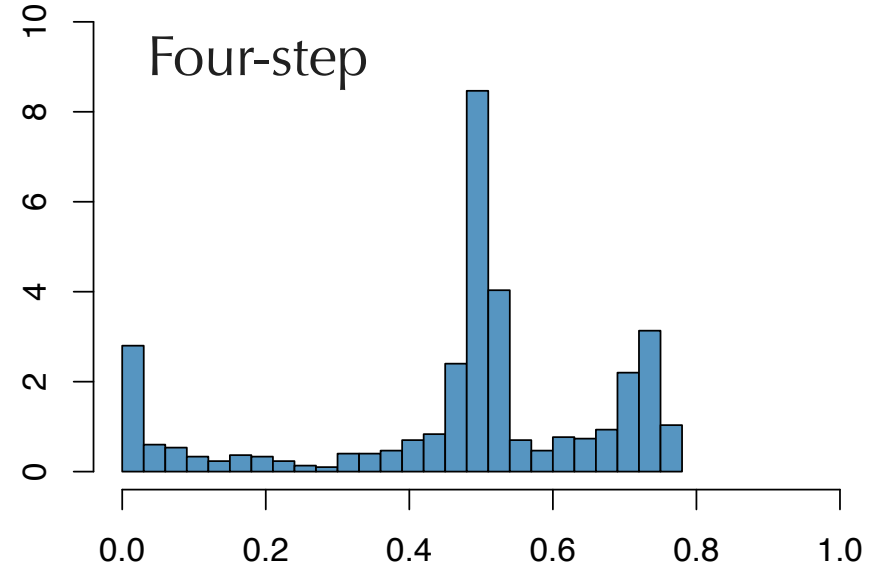Phenotypic effects fall into three broad categories

Class 1 ⇒ Indeterminate

Class 2 ⇒ Insensitive

Class 3 ⇒ Successful

Fitness distributions

Class 1   Class 2   Class 3

11 neighbors more than
1% fitter than wildtype

Frequency

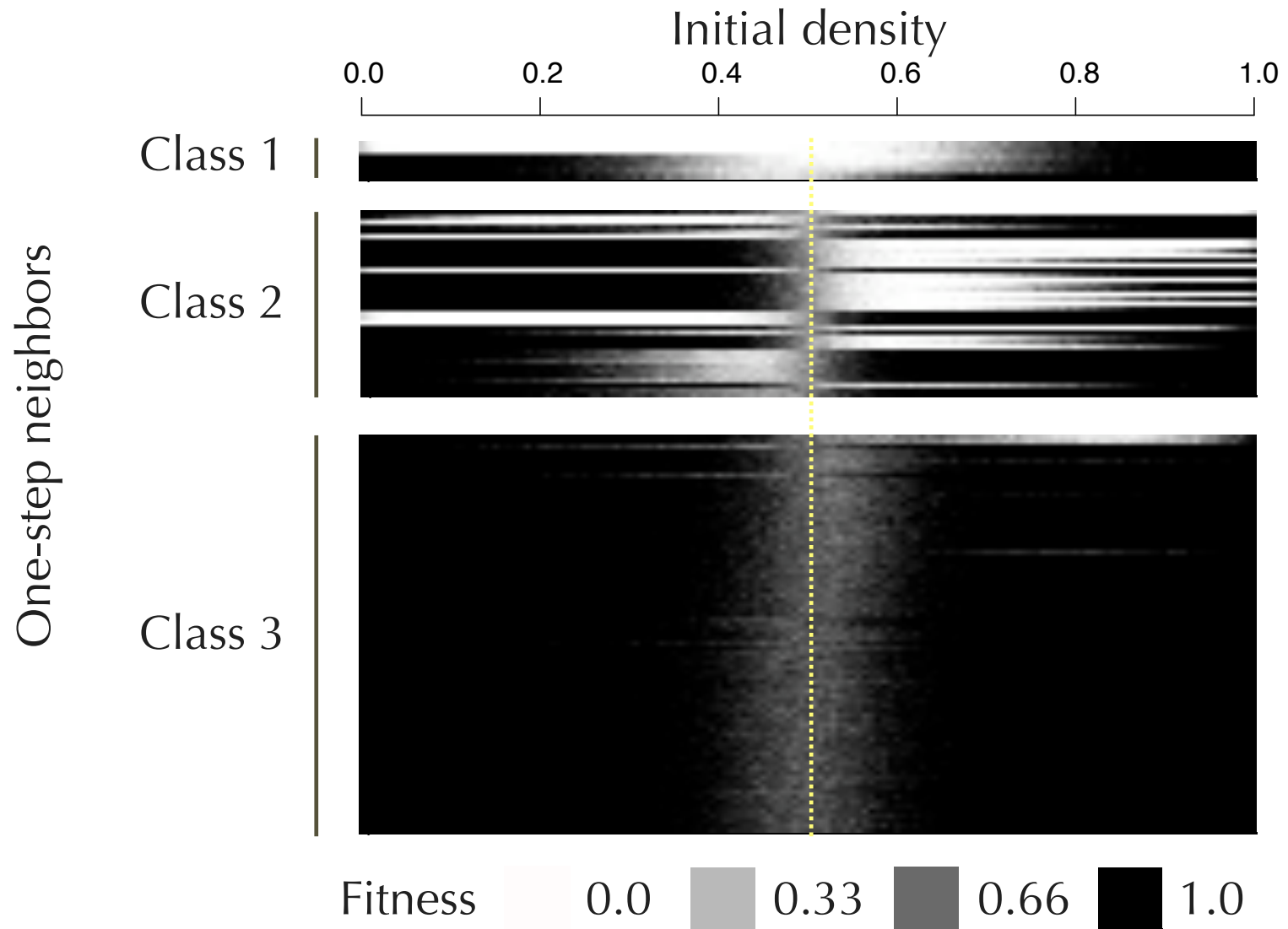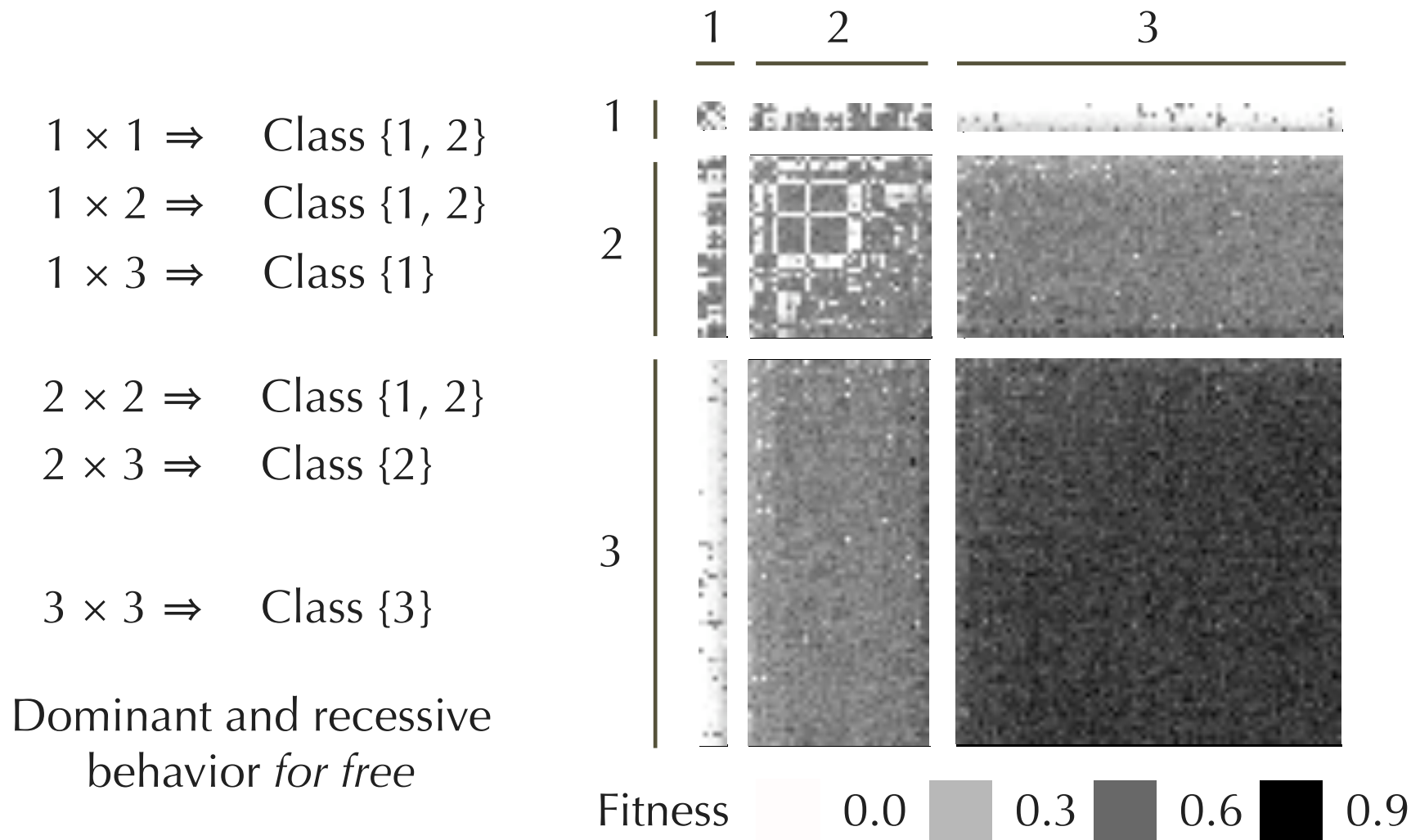'Fitness'   Wildtype 'fitness' 0.744

Fitness distributions
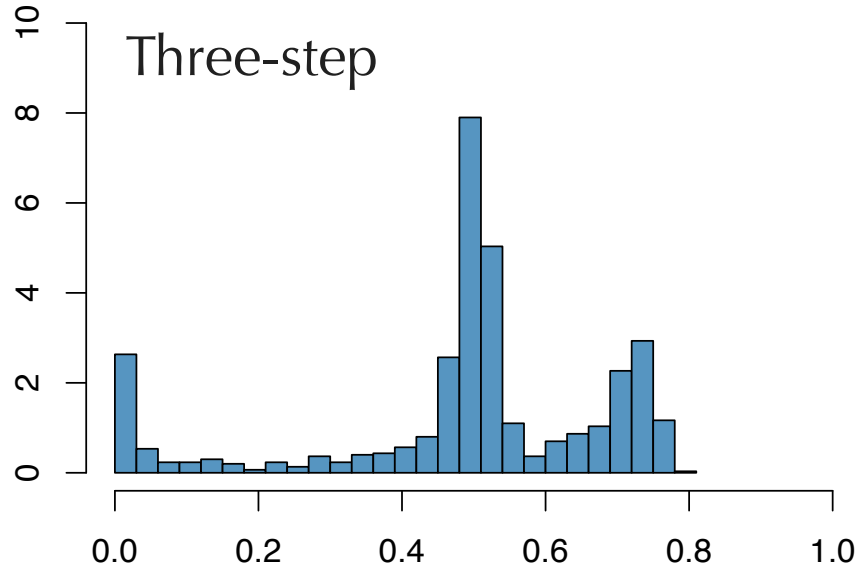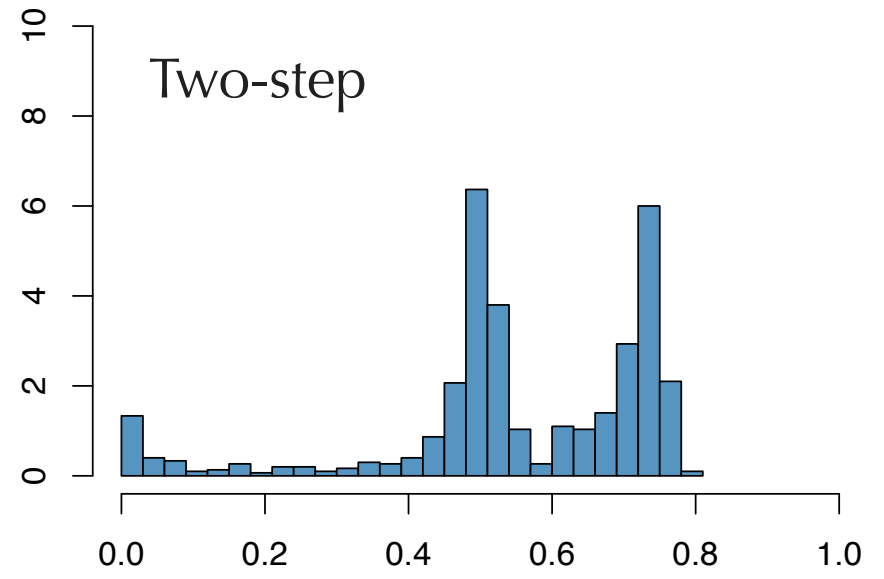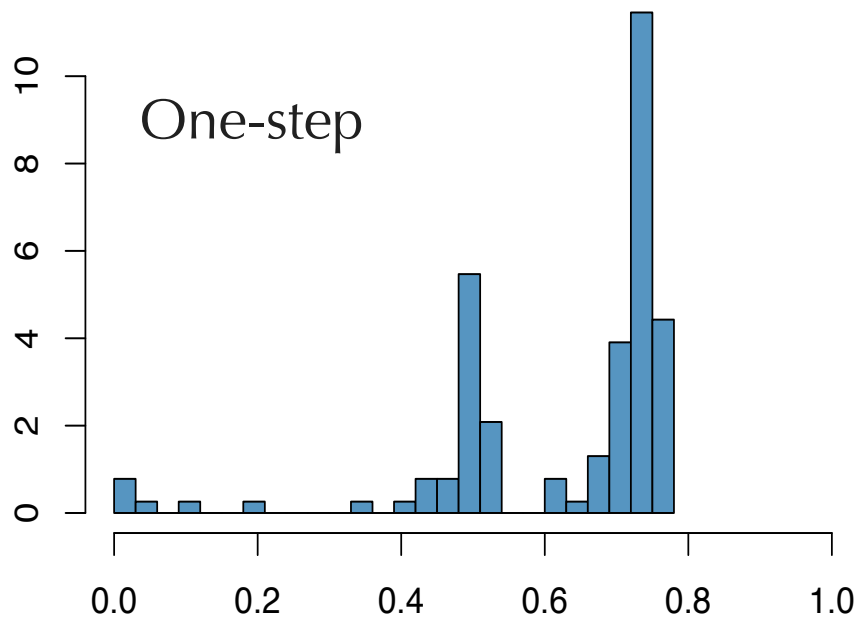
Initial density

Assaying fitness at different initial densities

Mutation interaction effects

Fitness landscape of the two-step neighbors

$1 \times 1 \Rightarrow$ Class $\{1, 2\}$

$1 \times 2 \Rightarrow$ Class $\{1, 2\}$

$1 \times 3 \Rightarrow$ Class $\{1\}$

$2 \times 2 \Rightarrow$ Class $\{1, 2\}$

$2 \times 3 \Rightarrow$ Class $\{2\}$

$3 \times 3 \Rightarrow$ Class $\{3\}$

Dominant and recessive
behavior *for free*

Fitness   0.0   0.3   0.6   0.9

Further evolution
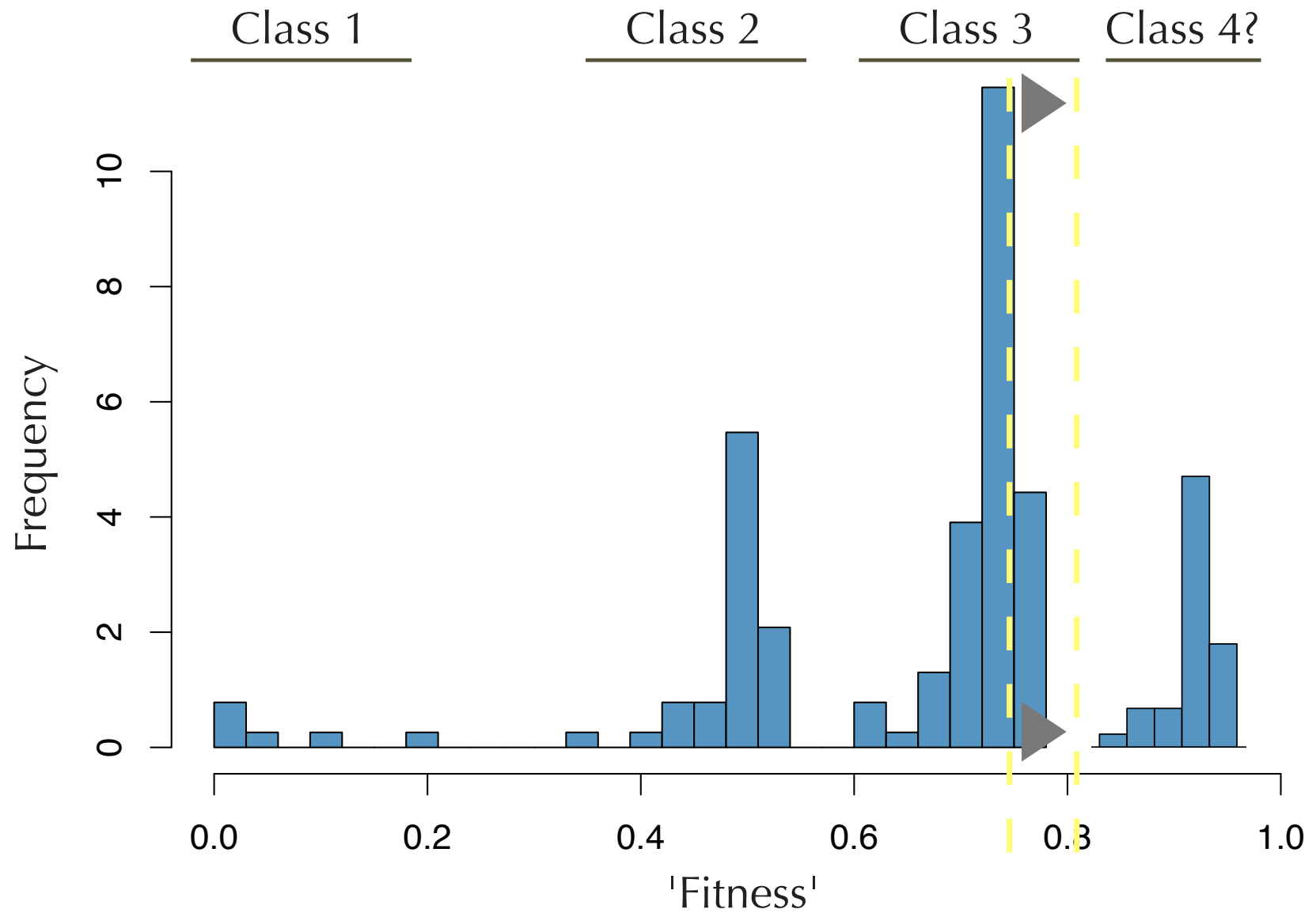
Conclusions

Cellular automata as model systems

Displays 'function' rather than 'form'

Mutations of large phenotypic effect

Dominant and recessive mutations

Epistasis

Phenotypic 'classes'

Acknowledgments

Kovas Boguta

Dan Hartl