The Impact of Gene Order on Evolution

Hao Shen
Mentor Anton Goloborodko

MIT PRIMES Conference
Sunday, May 19

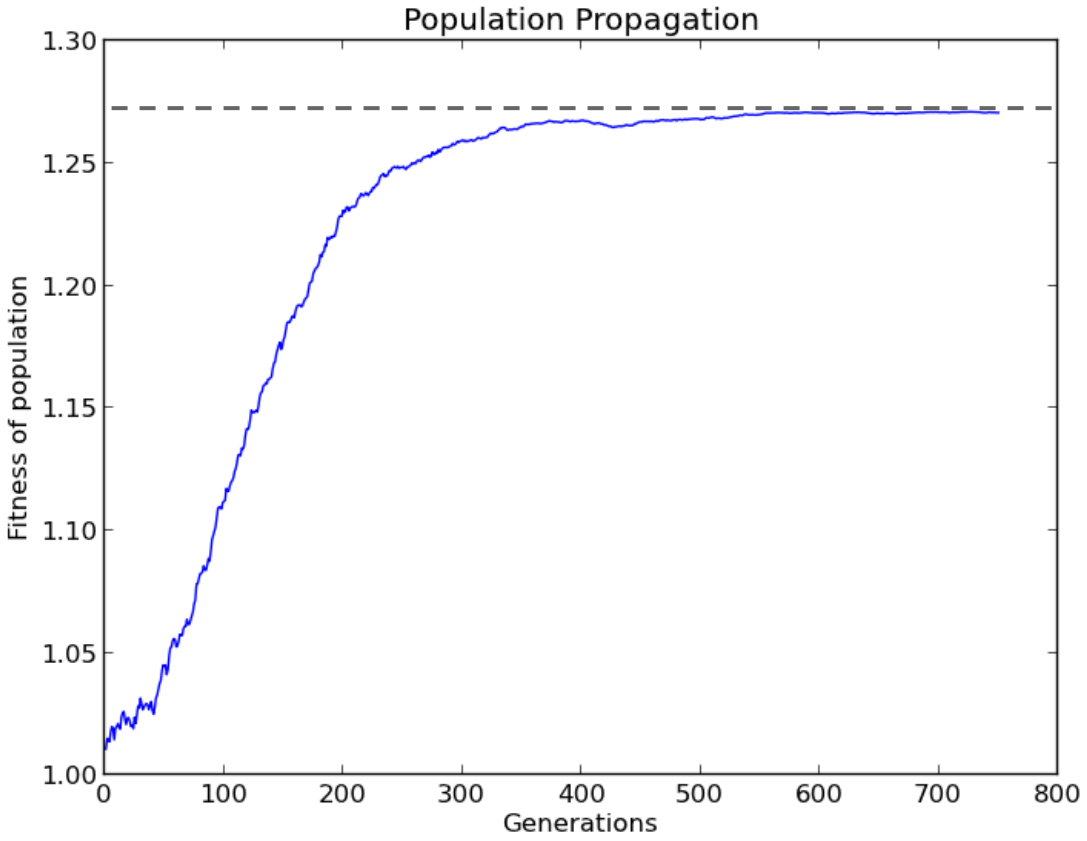
What is Population Genetics?

- Study of changes in distribution of allele frequencies
- Used to explain and confirm the general theory of evolution with genetics
- Part of modern evolutionary synthesis

Definitions

Fitness: describes the ability for an organism to survive and reproduce

Terminal Fitness: the eventual average fitness of a population



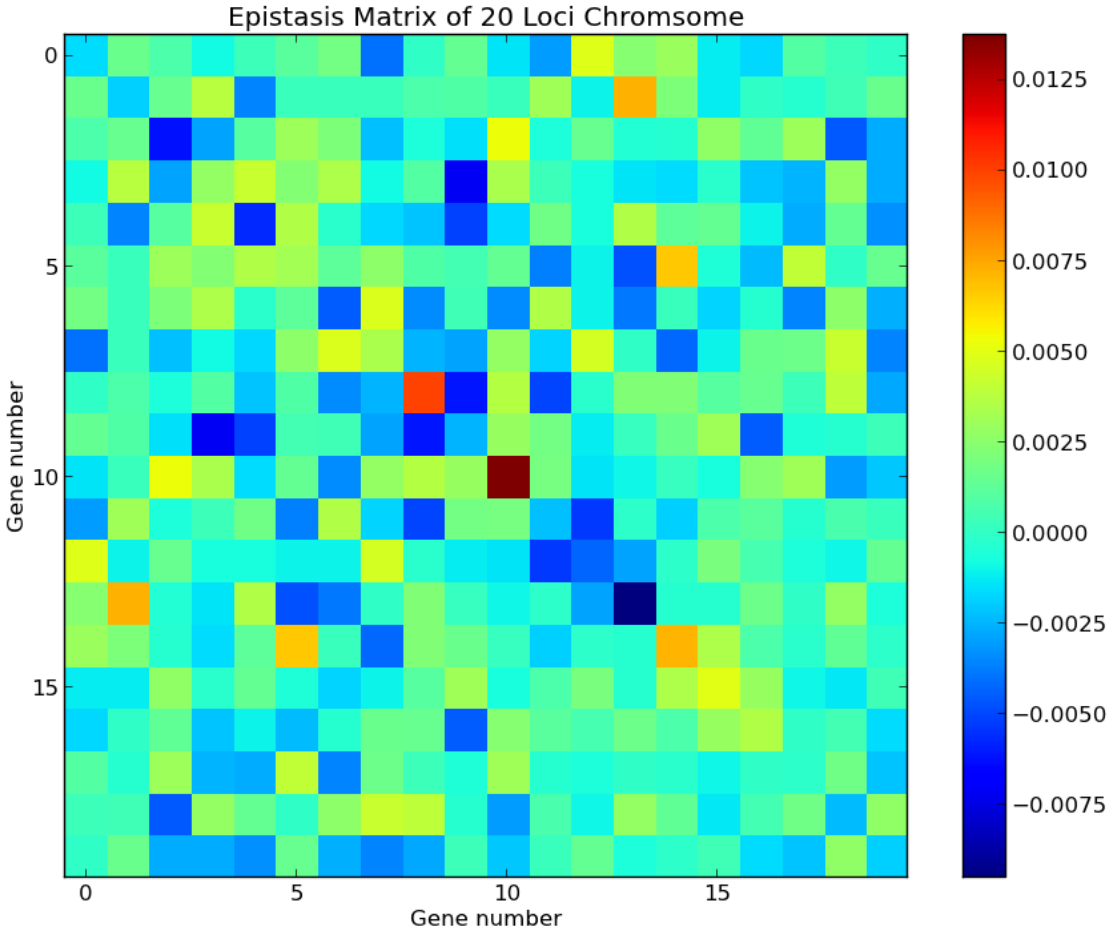
Definitions

Epistasis: the cumulative effects of combinations of alleles on fitness

Epistasis matrix: a matrix of numerical values of epistasis

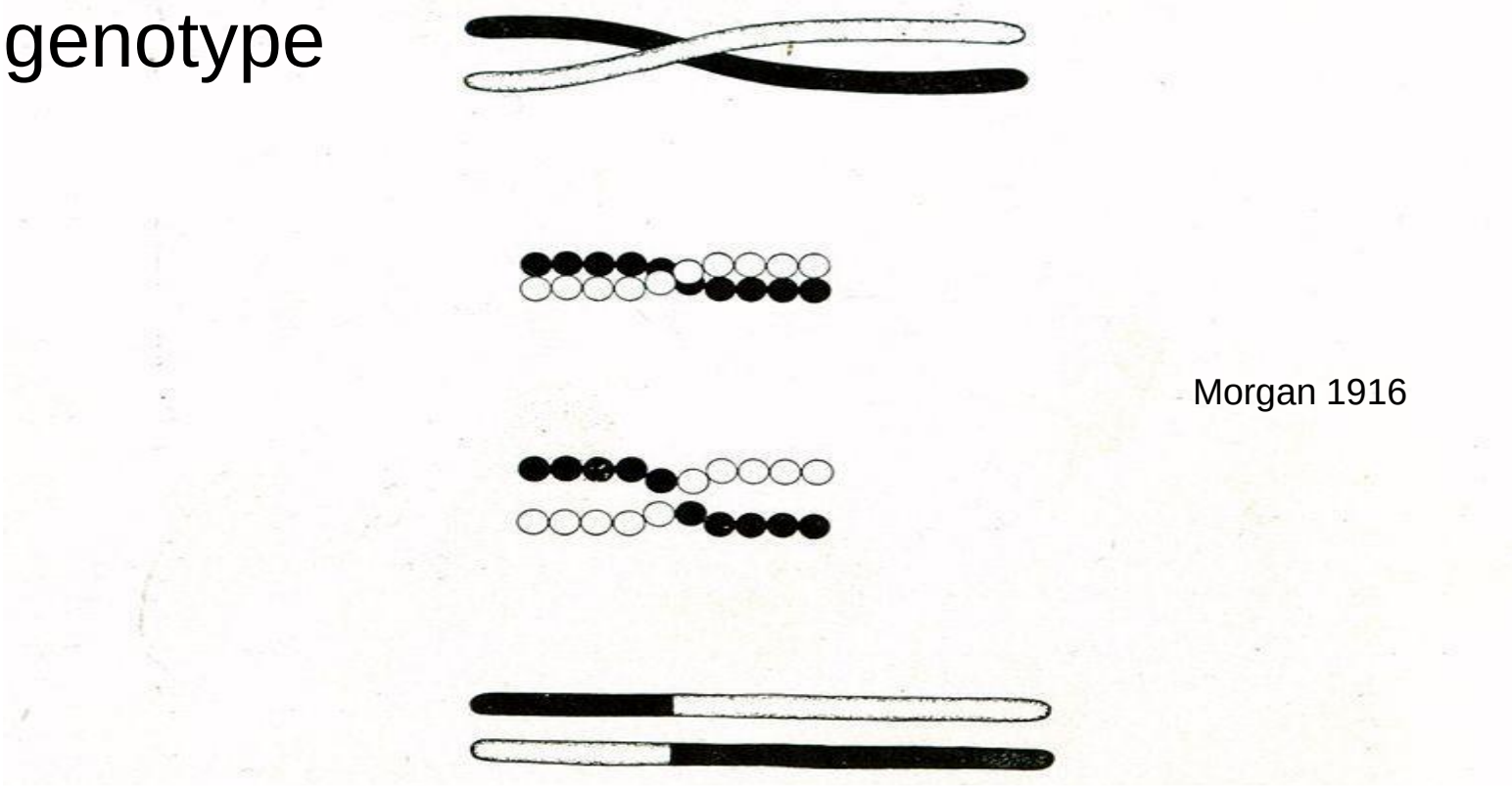
Gene Number	1	2	3	4
1	0.2	0.1	0.9	-0.5
2	0.1	0.4	1.1	0.6
3	0.9	1.1	-0.2	0.7
4	-0.5	0.6	0.7	0.3

Example of an Epistasis Matrix



Definitions

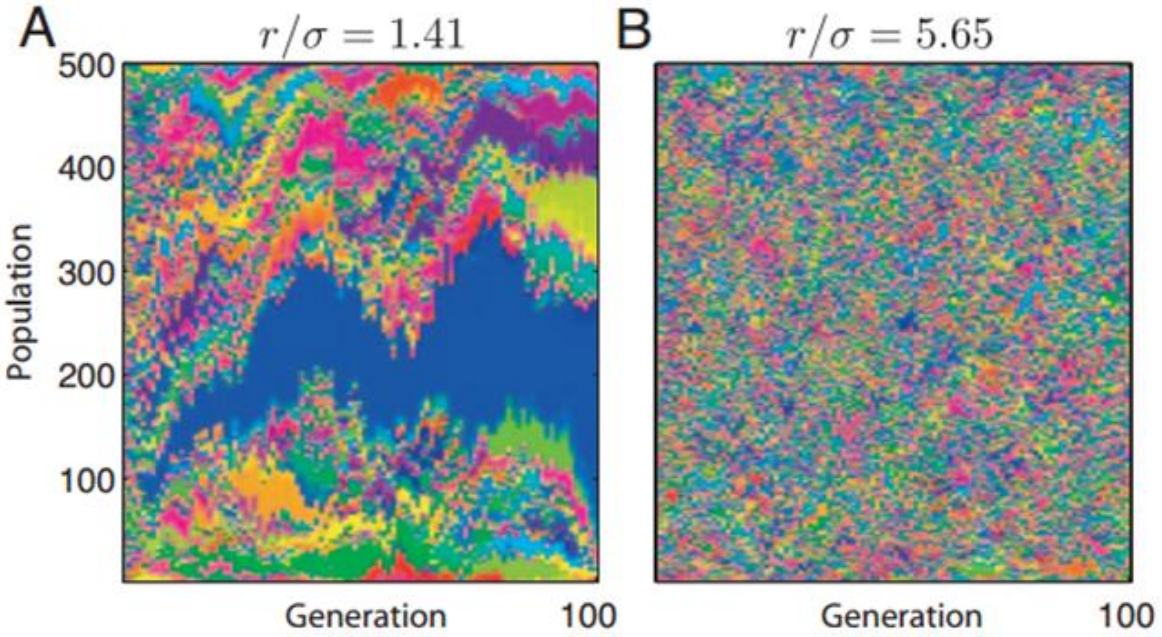
Recombination: the process by which two parents swap DNA to produce a new genotype



Morgan 1916

Recombination vs Epistasis

- Epistasis favors particular combination of genes
- Recombination breaks apart genomes



Neher & Shraiman, 2009

Base model: Wright-Fisher

Simulation Process

1. Create a population of genomes
2. At each generation, randomly select N genomes, with replacement, where N is the population size

Our Model [Neher & Shraiman, 2009]

- extension of Wright Fisher model

Addition of:

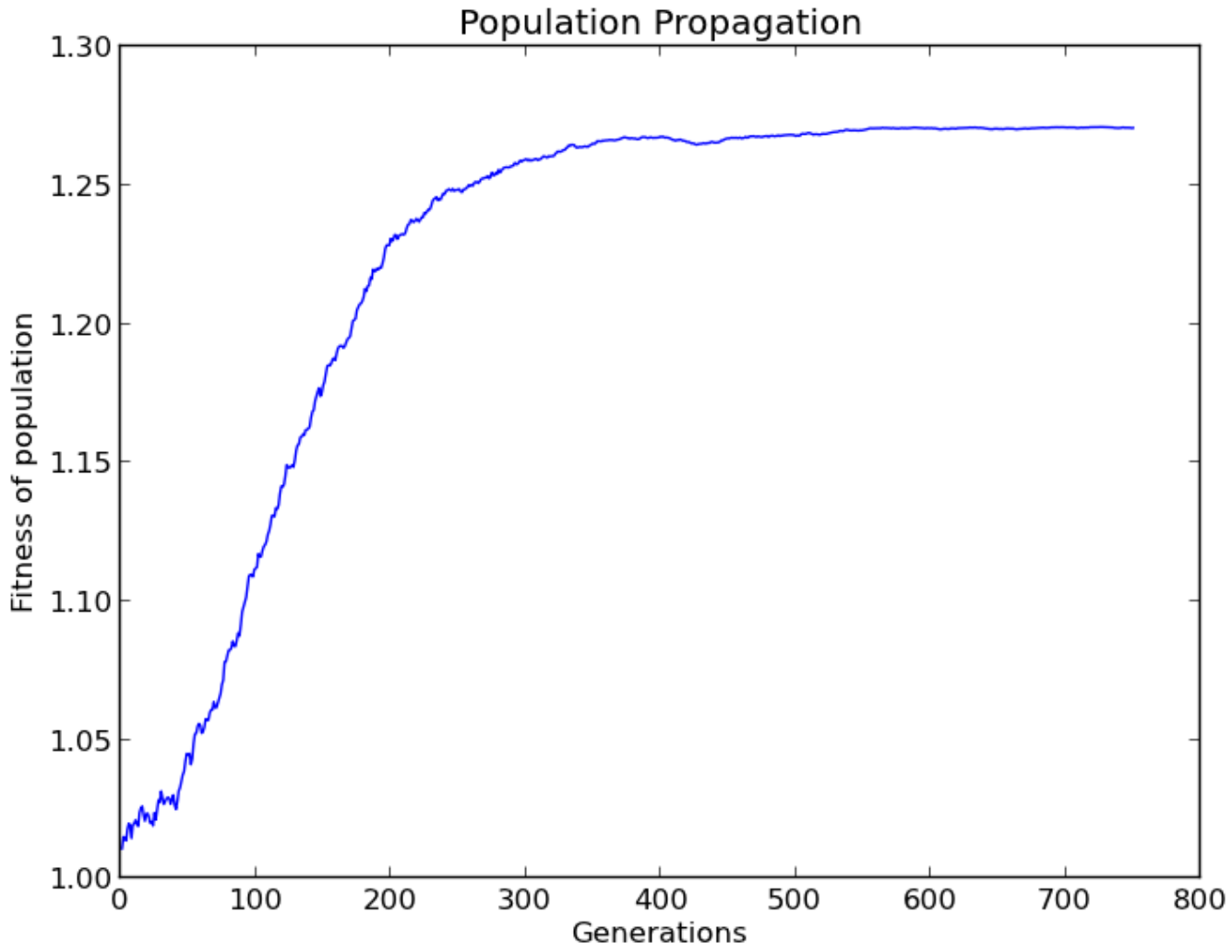
- gene interactions (epistasis)

- recombination

Simulation process

1. Generate population
2. Randomly pick 2 parents to mate based on fitness
3. Randomly recombine their genomes
4. Randomly pick a created genome to give to offspring
5. Repeat steps 2-4, N times, where N is the size of the population

Example of Time Evolution

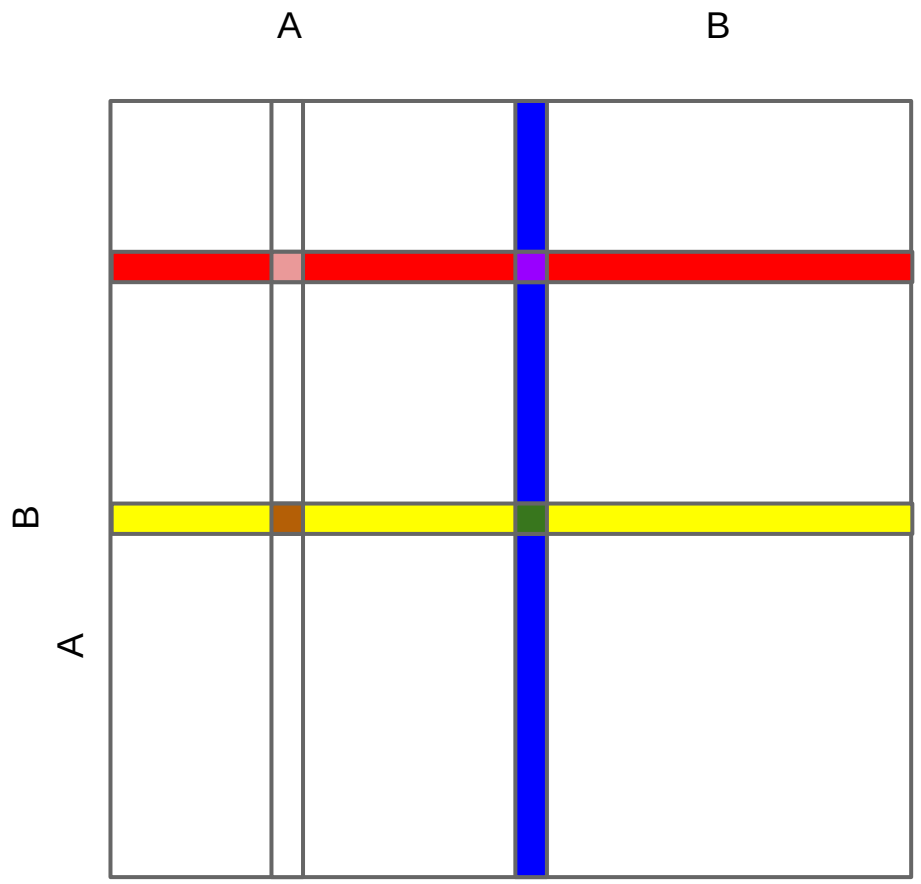


Hypothesis

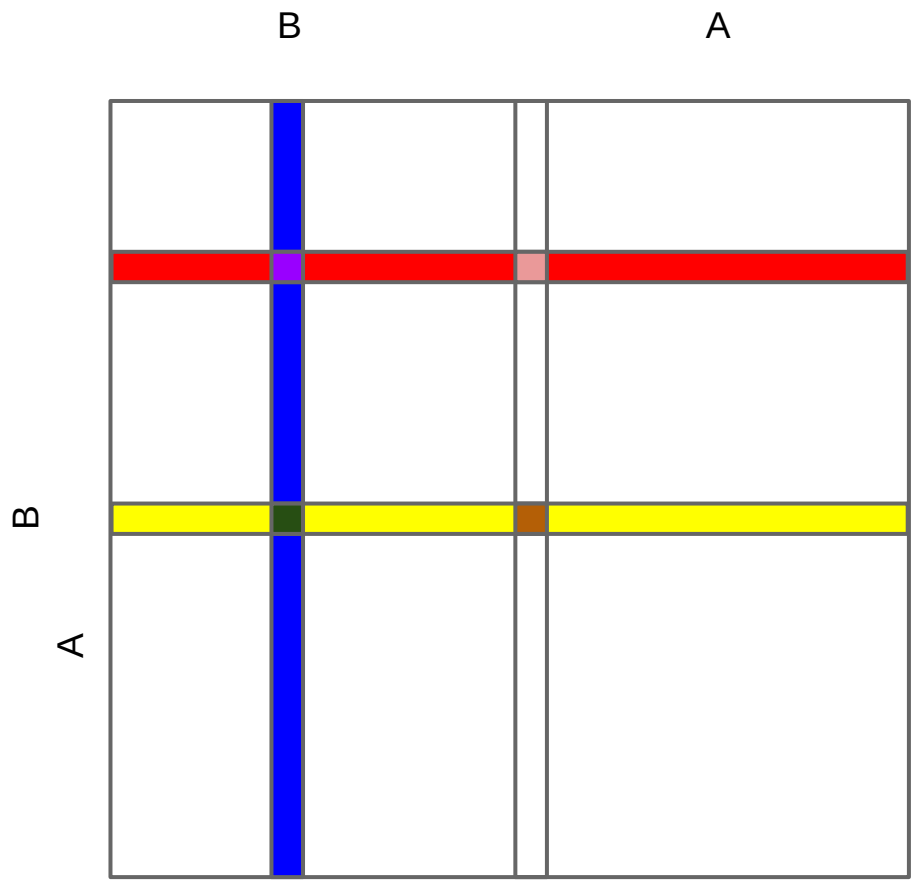
1. Epistasis and recombination will make populations sensitive to gene order

2. A gene ordering that produces an epistasis matrix with greater values closer to the diagonal will allow populations to become more fit.

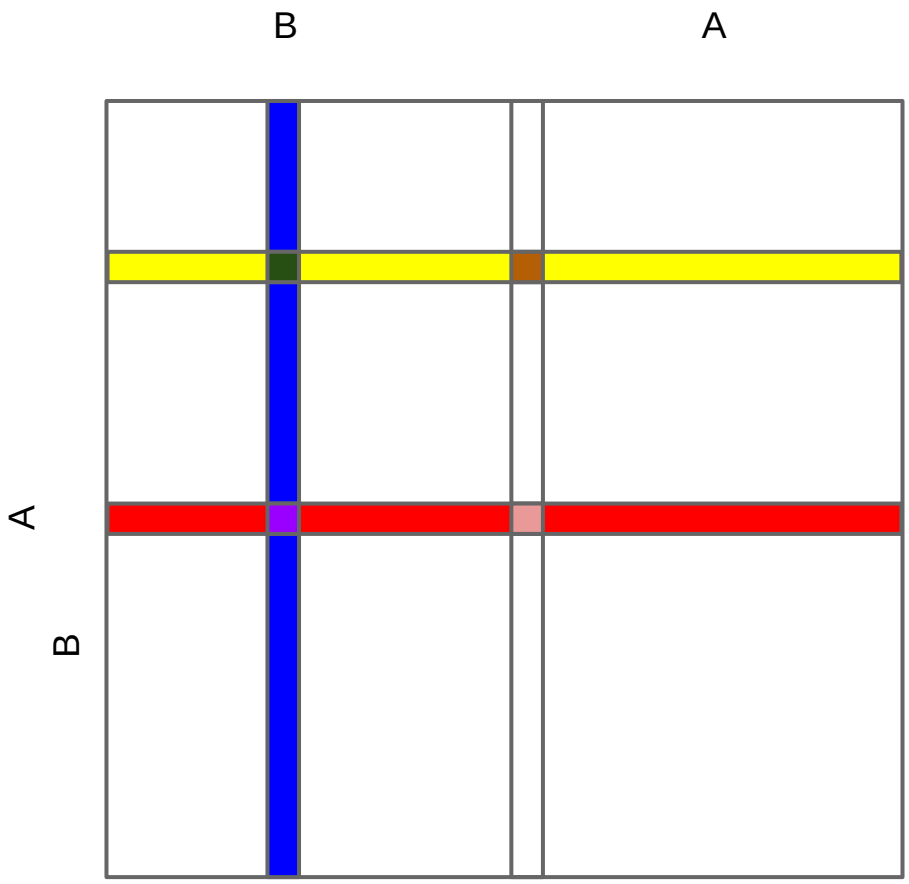
Swapping Gene Ordering



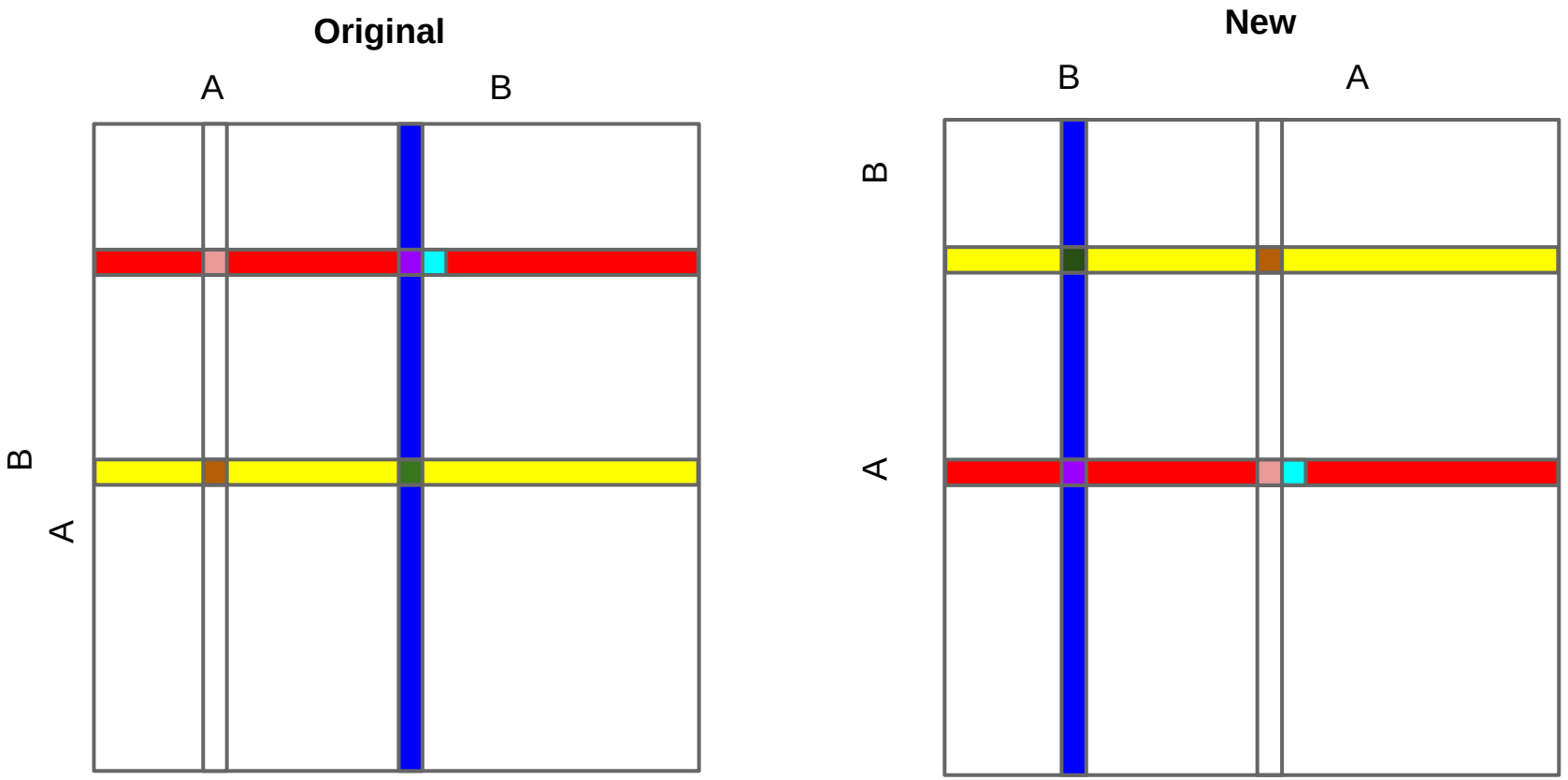
Swapping Gene Ordering



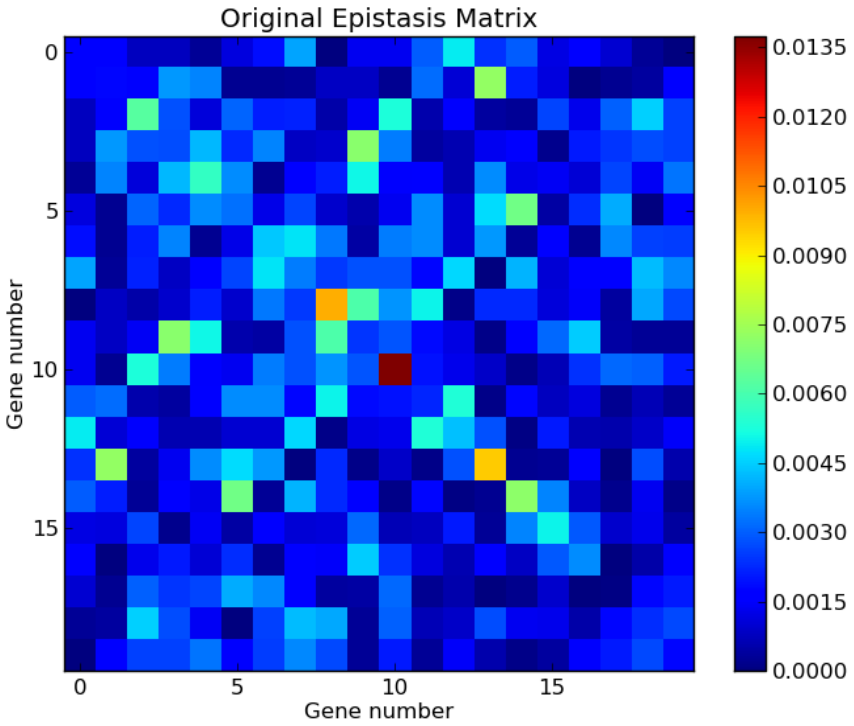
Swapping Gene Ordering



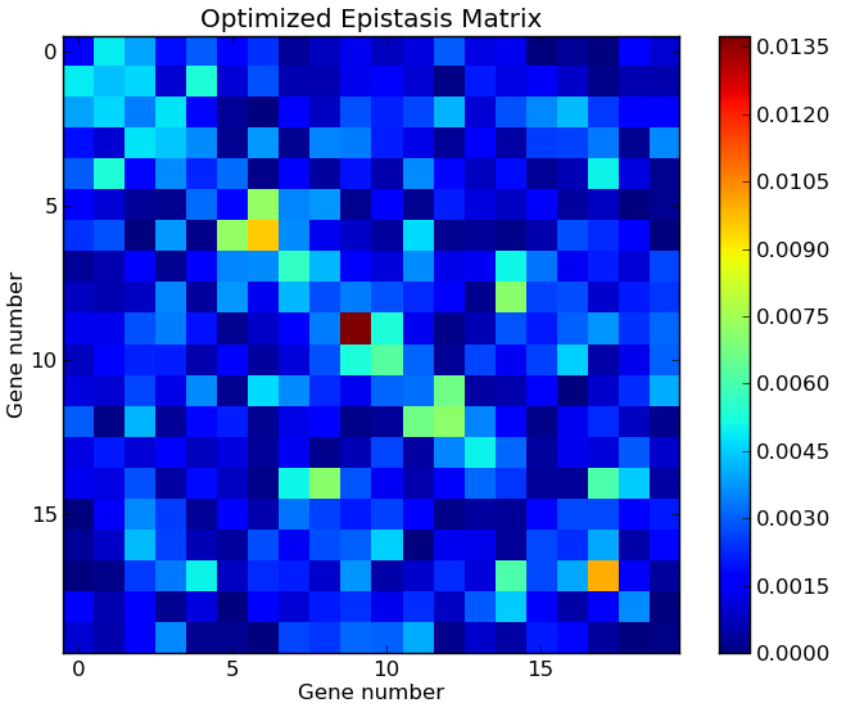
Swapping Gene Ordering



Results

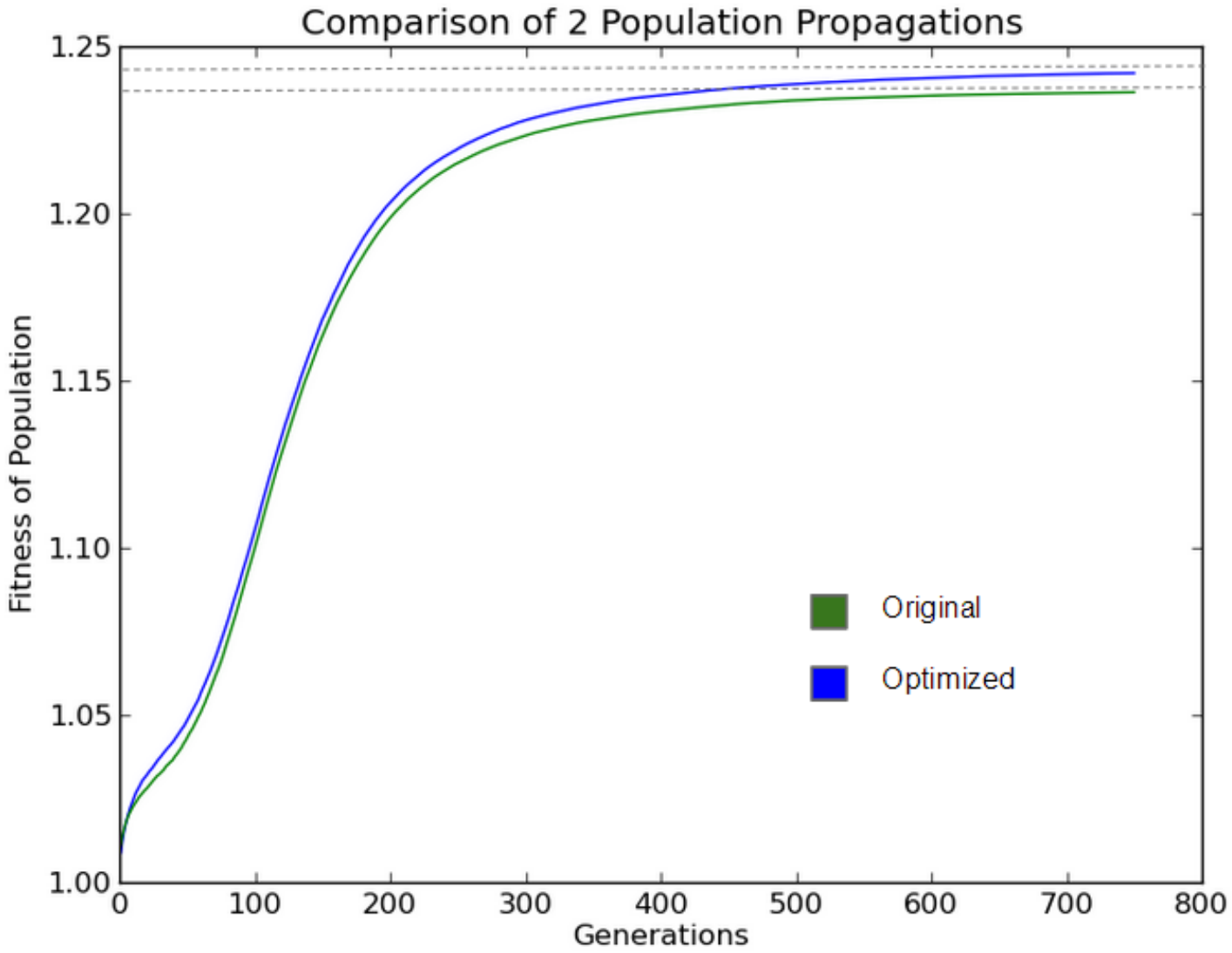


50.28

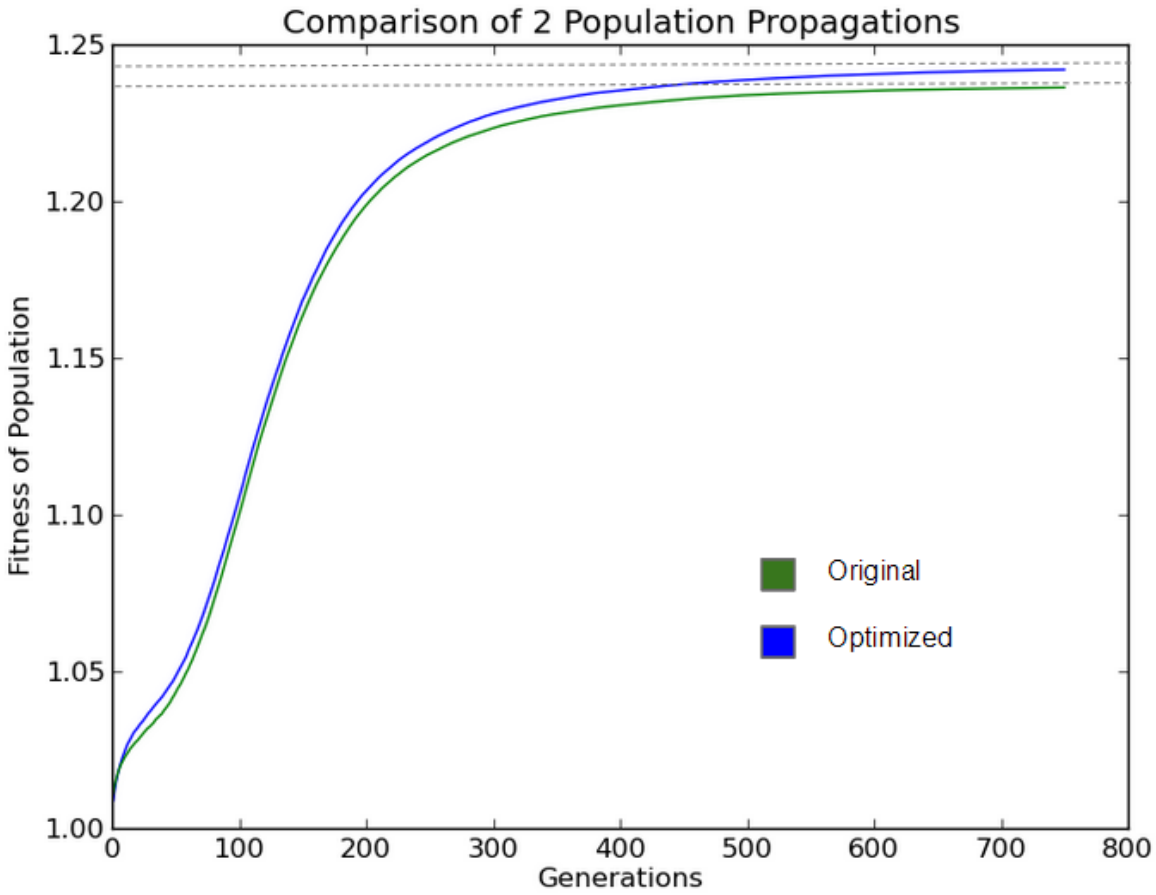


42.05

Results


Comparison of 2 Population Propagations
Fitness of Population
Generations
1.25
1.20
1.15
1.10
1.05
1.00
0
100
200
300
400
500
600
700
800
Original
Optimized

Results

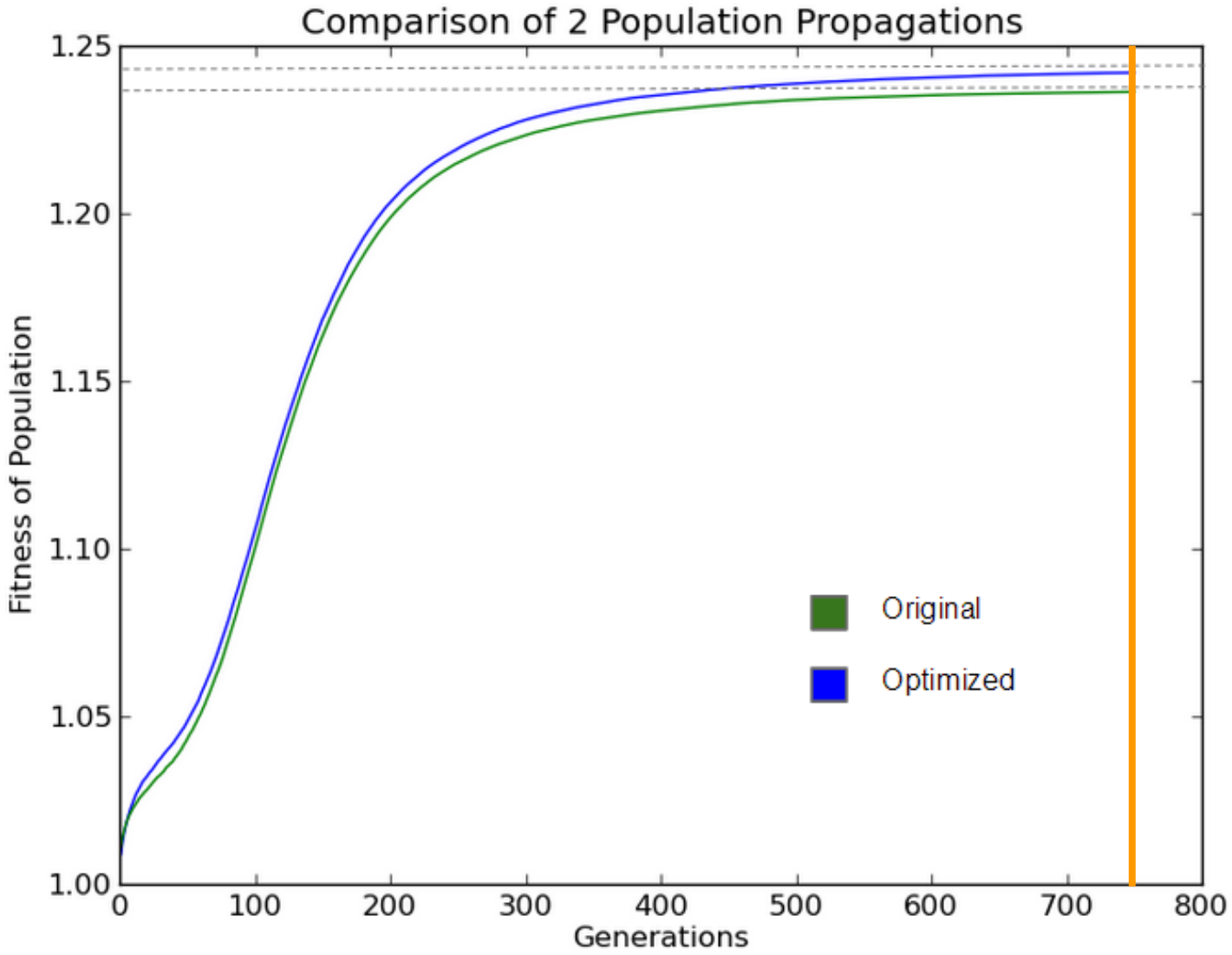


Average Terminal Fitness:

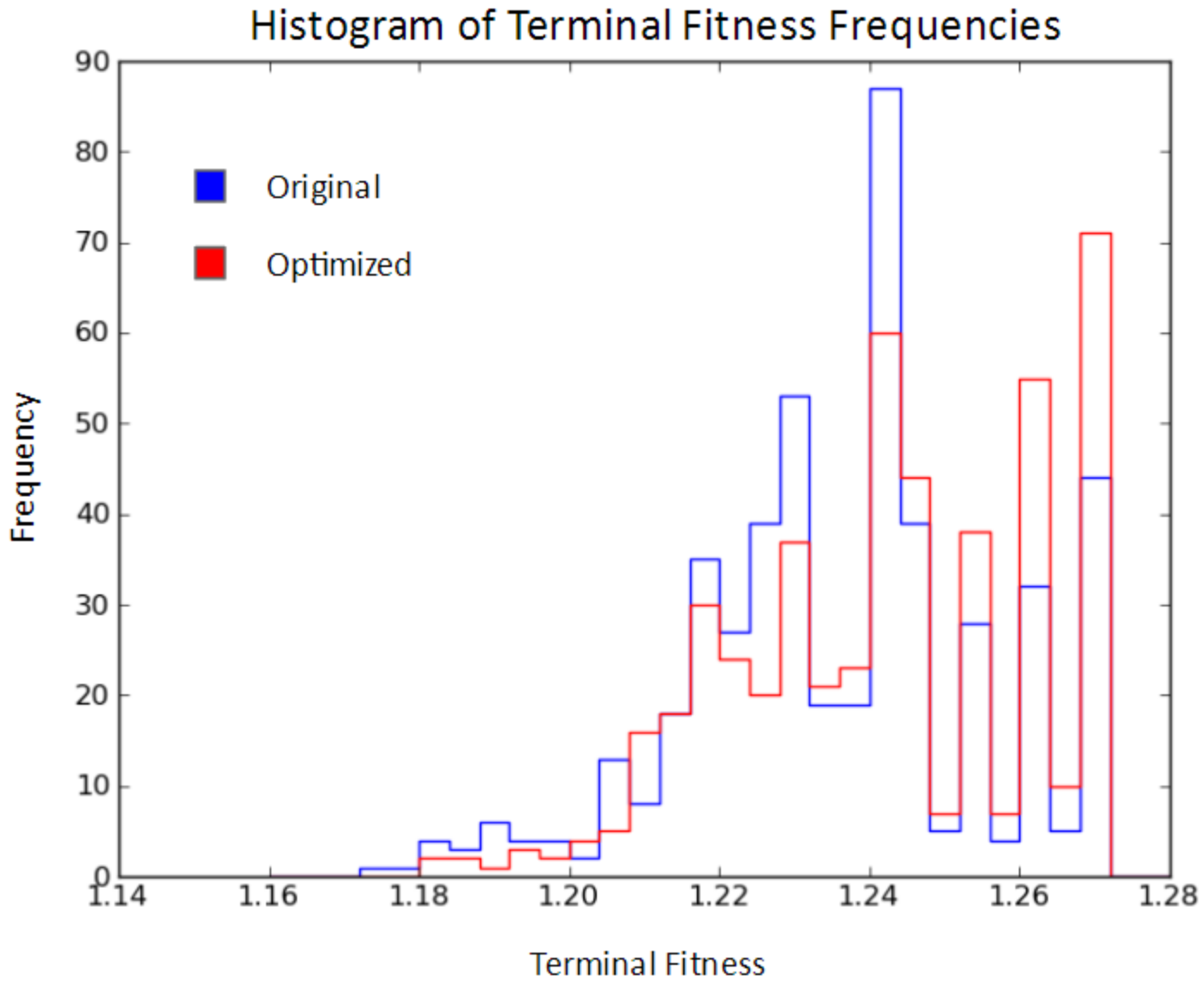
Original:	1.23625
Optimized:	1.24196

P-value: 0.000742

Results


Comparison of 2 Population Propagations
Fitness of Population
Generations
1.25
1.20
1.15
1.10
1.05
1.00
0
100
200
300
400
500
600
700
800
Original
Optimized

Results


Histogram of Terminal Fitness Frequencies
Original
Optimized
Frequency
Terminal Fitness
0
10
20
30
40
50
60
70
80
90
1.14
1.16
1.18
1.20
1.22
1.24
1.26
1.28

Conclusions

- Recombination and epistasis make populations sensitive to gene orders
- Changes in gene ordering can produce significantly different terminal fitnesses
- The overall rate at which populations evolve is similar for different gene orderings

Implications

- Results of this project can be used to analyze the observed gene arrangement

- Understand the history of observed evolving populations

Implications

- Example: People have observed that chromosome 21 genes are different from chromosome 1 genes. Also, chromosome 21 is a lot shorter than chromosome 1. However, both chromosomes recombine on average once per generation.
- Different recombination rates on different chromosomes could predict the relative epistasis strengths of the alleles on those chromosomes

Thanks to...

- Anton Goloborodko
- Chris McFarland
- Leonid Mirny
- Mirny Lab
- MIT PRIMES Program