

Using the RNA sequence-to-structure map

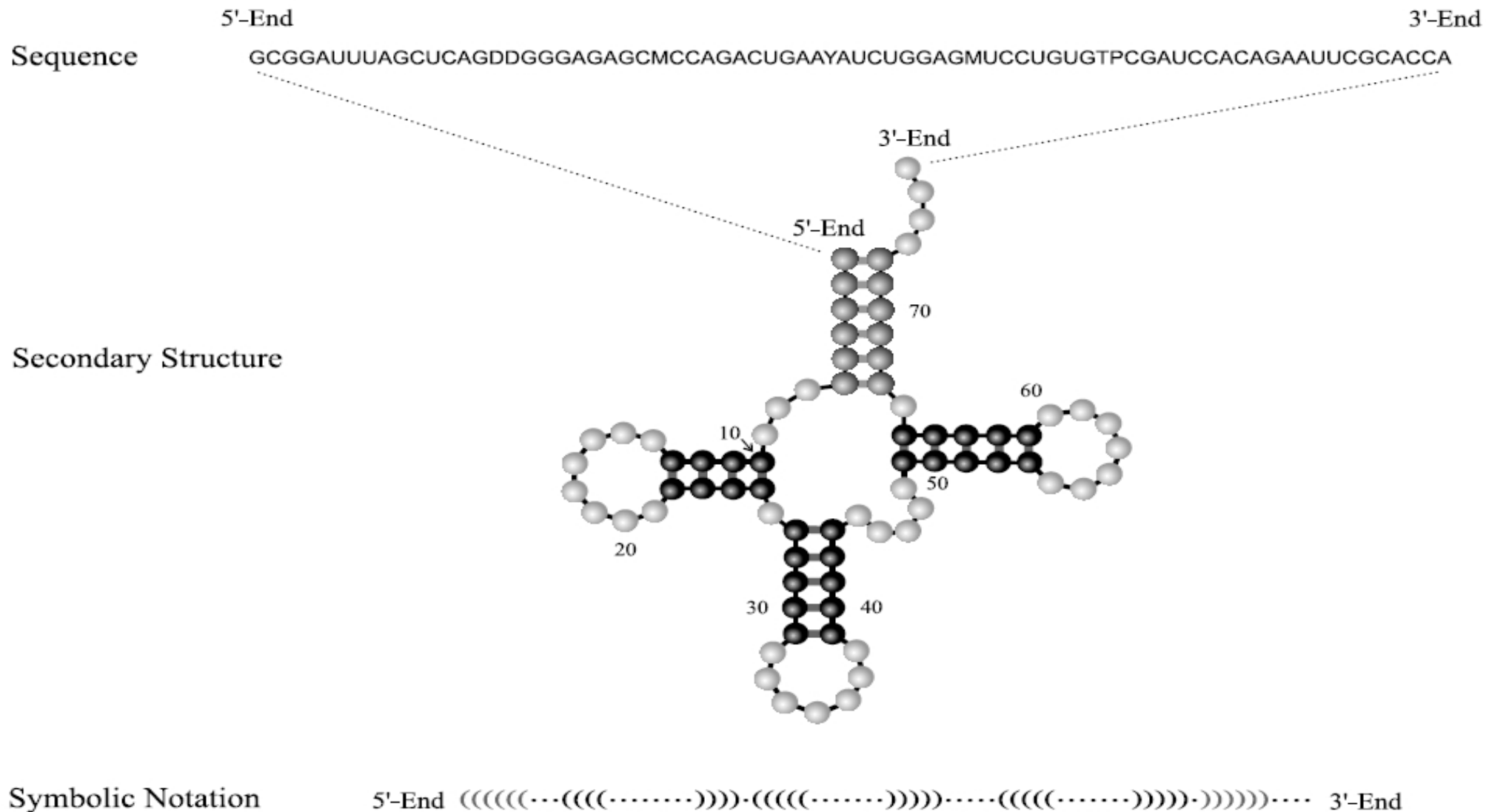
for functional evolution of ribozyme-catalyzed metabolisms

**Alexander Ullrich, Christoph Flamm and Lukas Endler
University of Vienna, Theoretical Chemistry Group
aullrich@tbi.univie.ac.at**

Outline

1. The RNA sequence-to-structure map
2. Using the sequence-to-structure map
3. Evolution of metabolic networks
4. Integrating the sequence-to-structure map
5. Results / Comparison
6. Application / Outlook

The RNA sequence-to-structure map



The RNA sequence-to-structure map

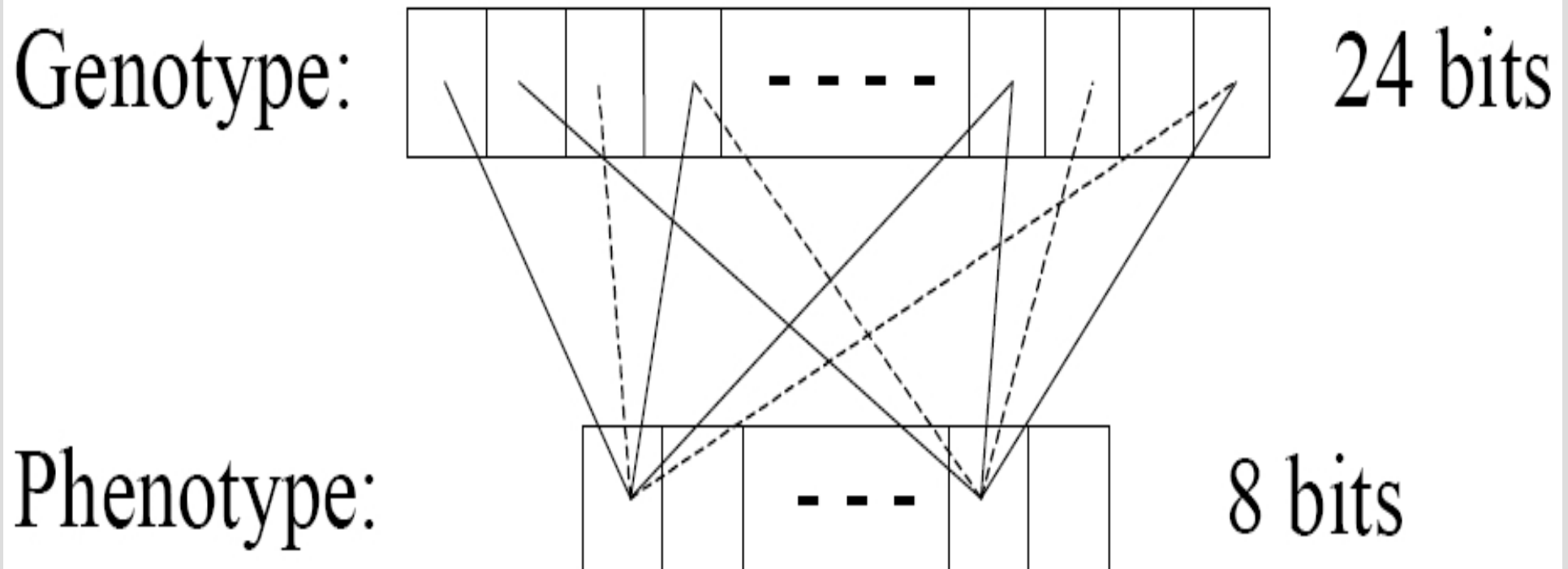
- Many-to-One Property:
 - More sequences than structures
 - Many sequences map to the same structure
- Different Frequencies:
 - Highly frequent and rare structures
- Neutral Networks:
 - Sequences with the same structure build extensive connected networks
 - Allows genetic drift along these networks
- Intertwined Neighborhood
 - All frequent structures lie in close proximity
 - e.g.: length = 100 --> radius = 15 mutations

Using the sequence-to-structure map

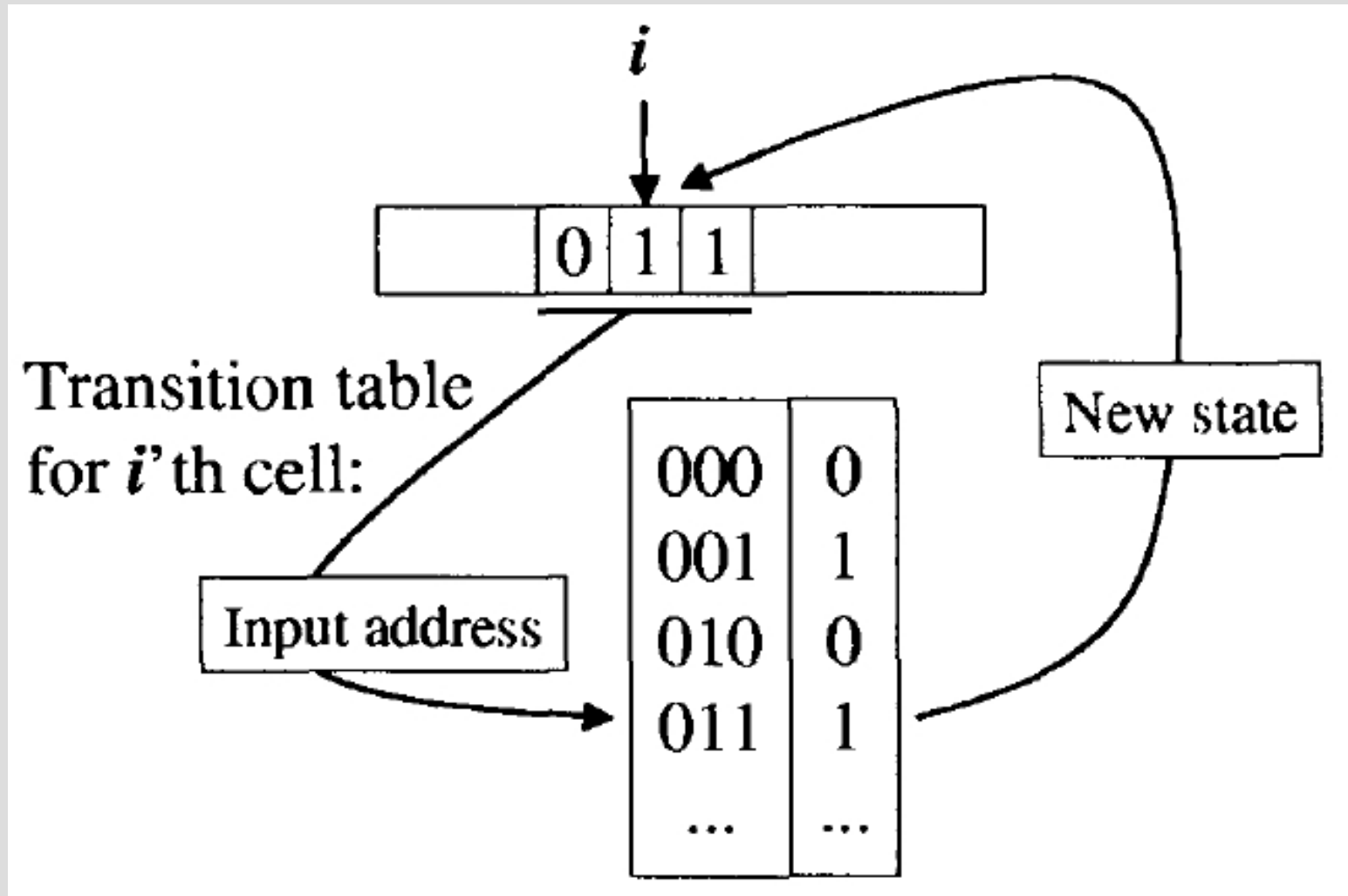
- Evolutionary Algorithms:
 - Reproduction, Variation, Selection (Fitness)
 - Population of Individuals with Genotype and Phenotype
 - Basic EA: Genotype = Phenotype
 - Others: Mapping from Genotype to Phenotype
 - Problem: sub-optimal traps
 - Solution: redundant mapping

Voting Mapping

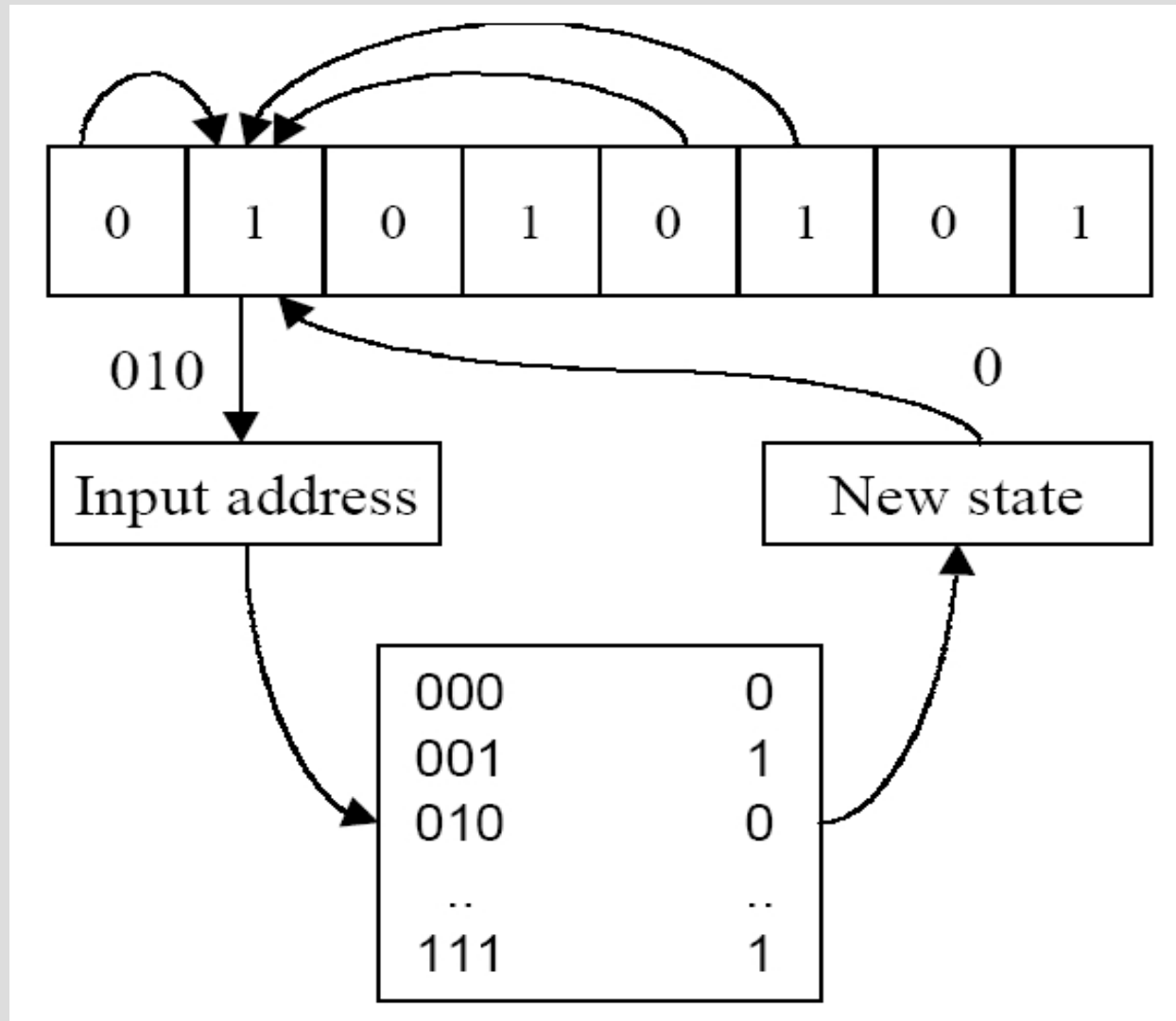
- Trivial voting performs like direct mapping
- Better standard voting mapping



Cellular Automaton (CA)



Random Boolean Network (RBN)



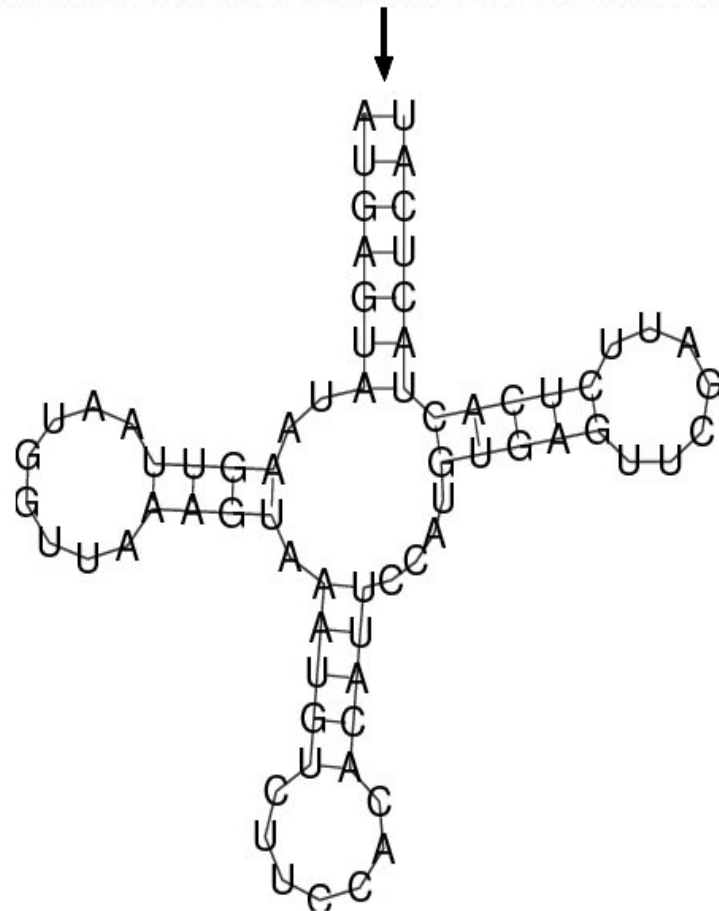
Mappings based on the sequence-to-structure map

- First, the RNA-sequence is folded into its minimum free energy structure
- Option 1:
 - Equidistant target structures or structure templates are generated
 - One target for each phenotype
 - Map to the phenotype with the closest target
- Option 2:
 - Extract features from the structure (and sequence)
 - Trivial voting from features to phenotype

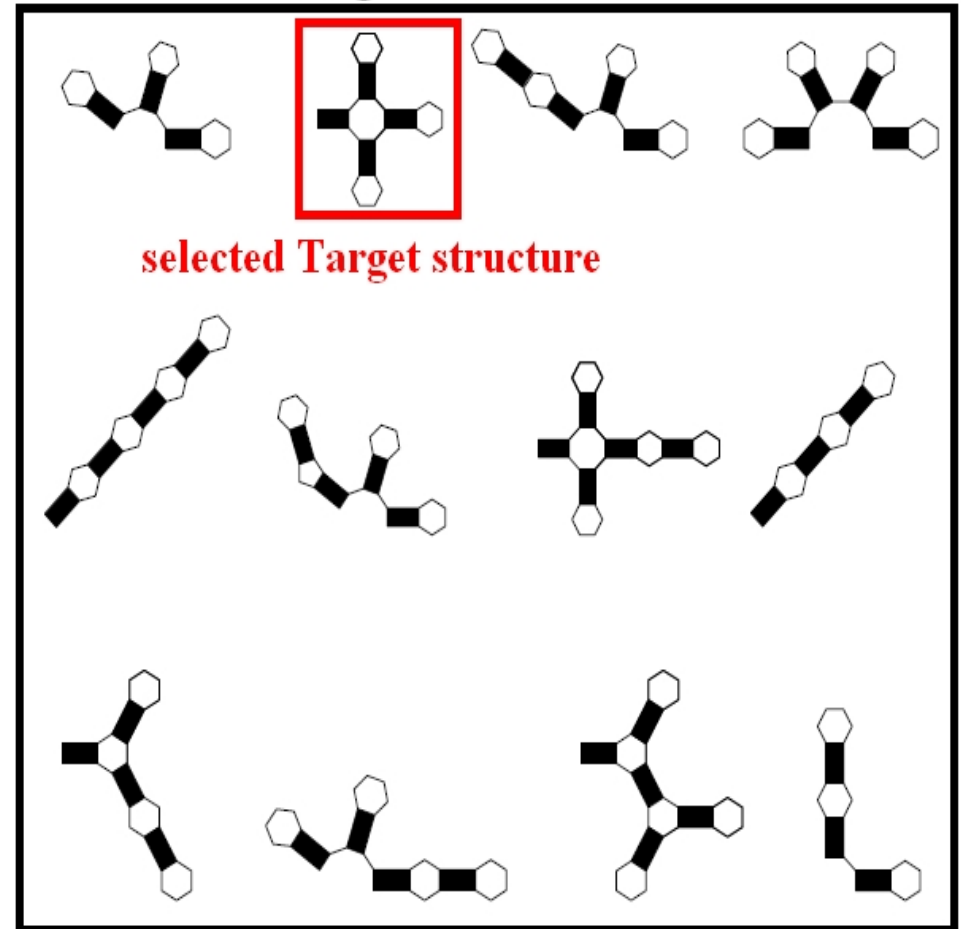
Option 1

- 2 versions: bp distance, ensemble distance

AUGAGUAAAGUUAAUGGUUAAAGUAAAUGUCUUCGACAGAUUCCAUGUGAGUUCGAUUCUGACUACUGAU

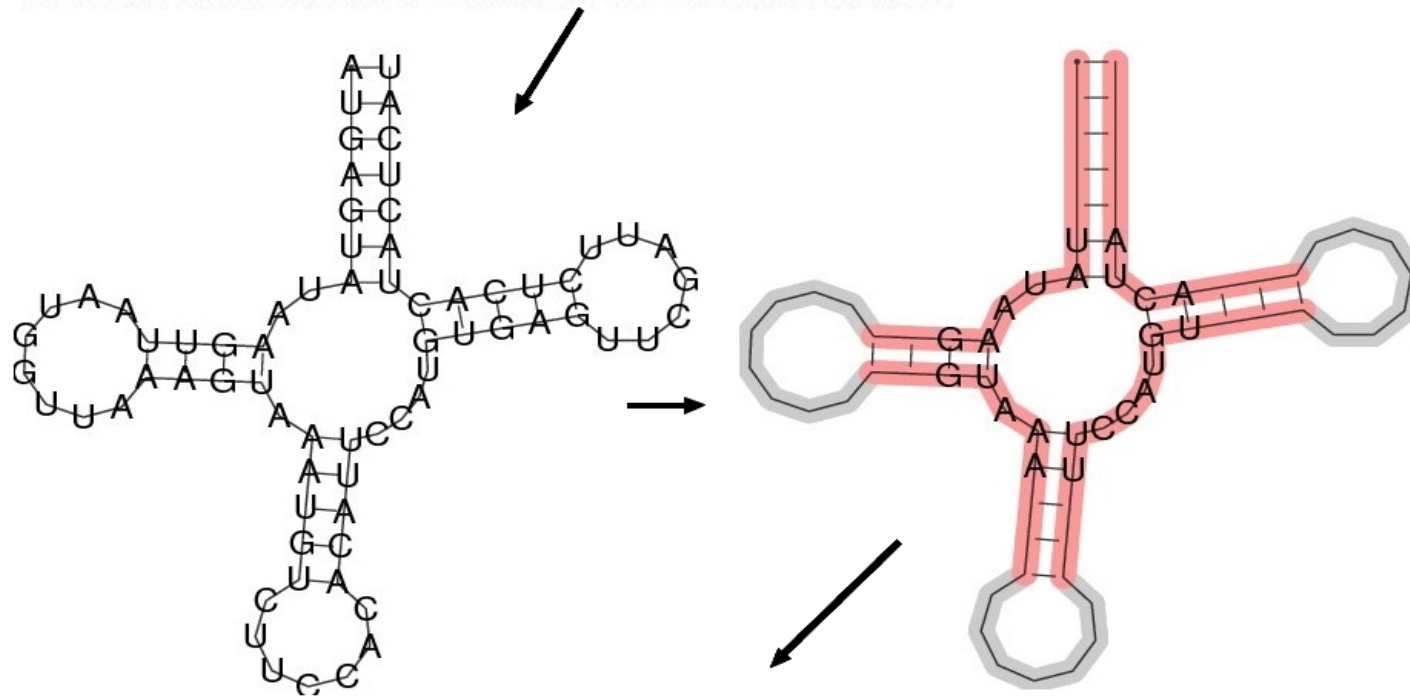


Target structures



Option 2

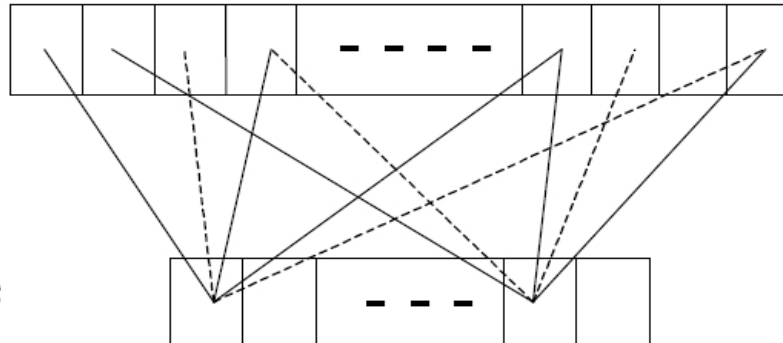
AUGAGUAAAGUUAAUGGUAAAGUAAAUGUCUCCACACAUUCCAUGAGAGUUCGAUUUCUCACUACUCAU



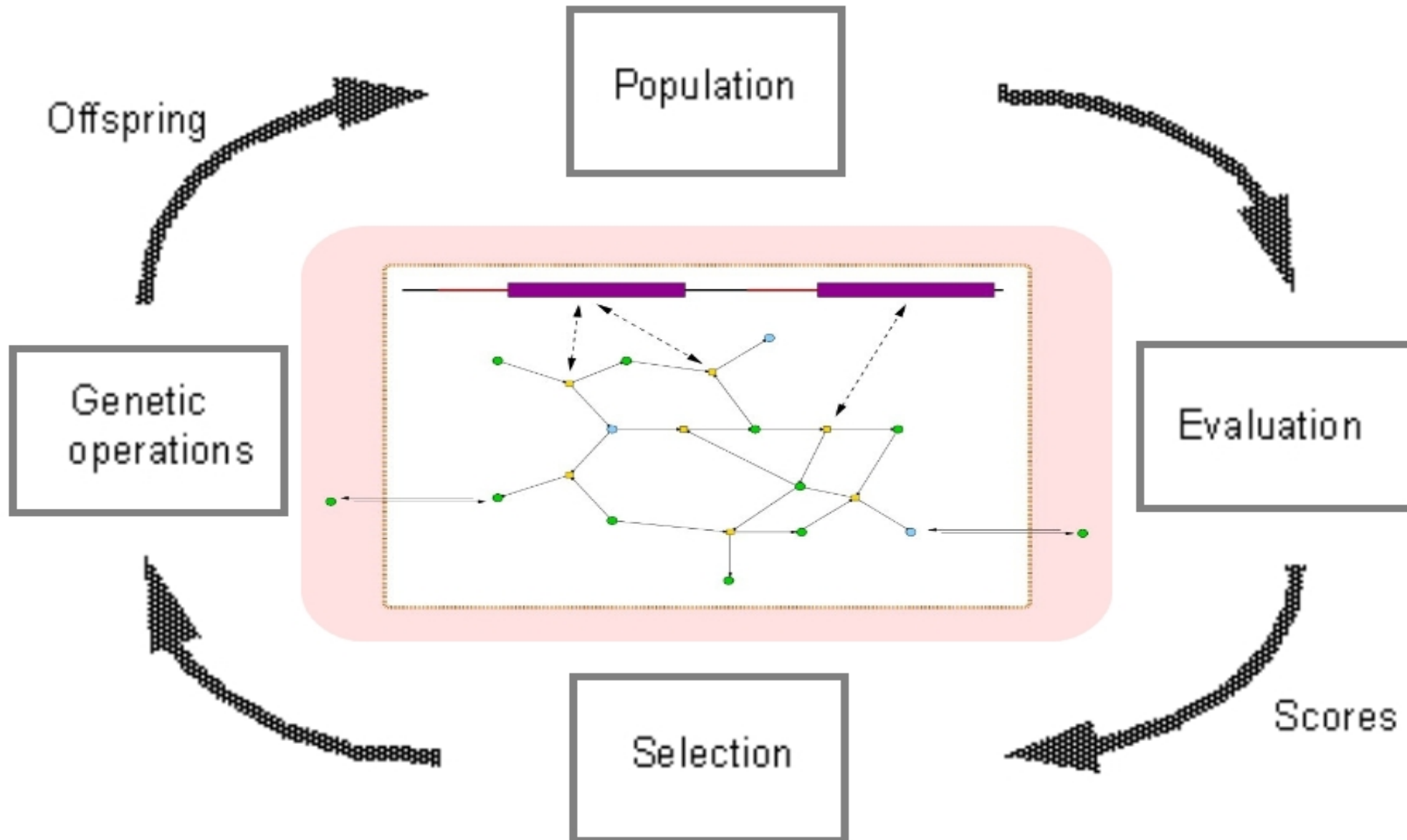
Features



Phenotype



Functional evolution of metabolic networks



Functional evolution of metabolic networks

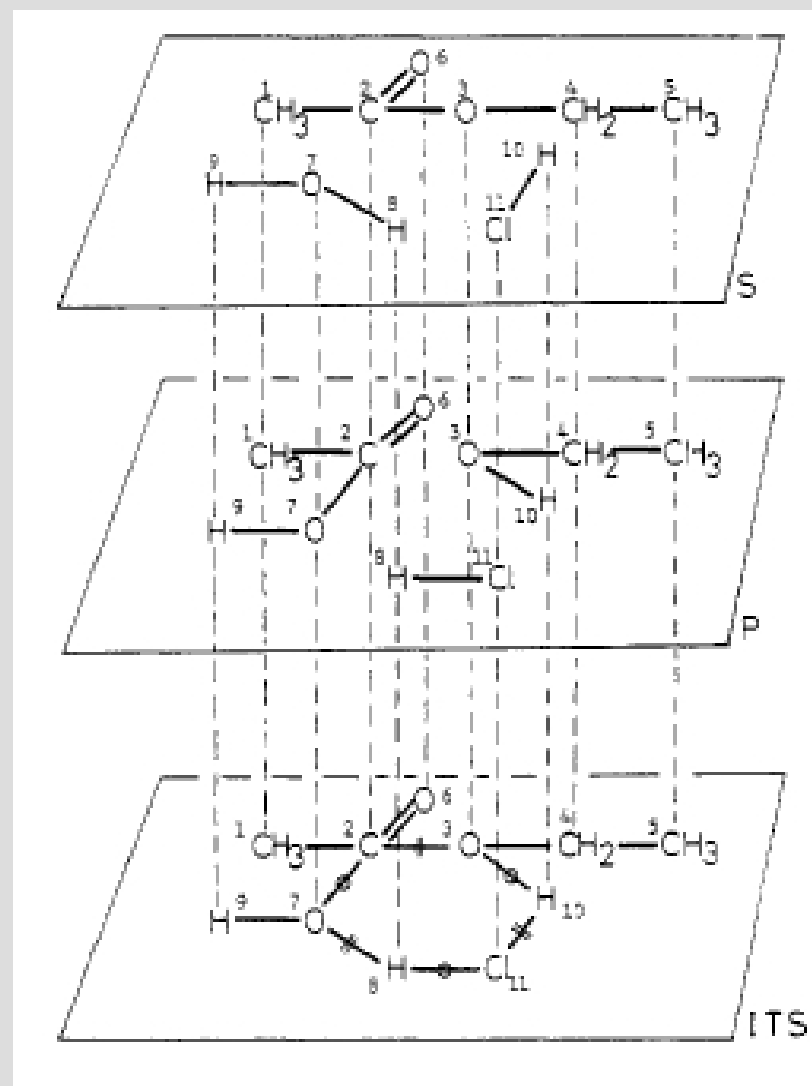
- Genome = RNA-sequence (TATA-box + genes)
- Genes (Genotype) = RNA-sequence
- Metabolism = bipartite graph (metabolites + enzymes)
- Metabolites = chemical graph (atom type + chemical bond)
- Enzymes (Phenotype) = superimposition of two chemical graphs (transition state structure)
 - Chemical reaction = Graph rewrite rule

Functional evolution of metabolic networks

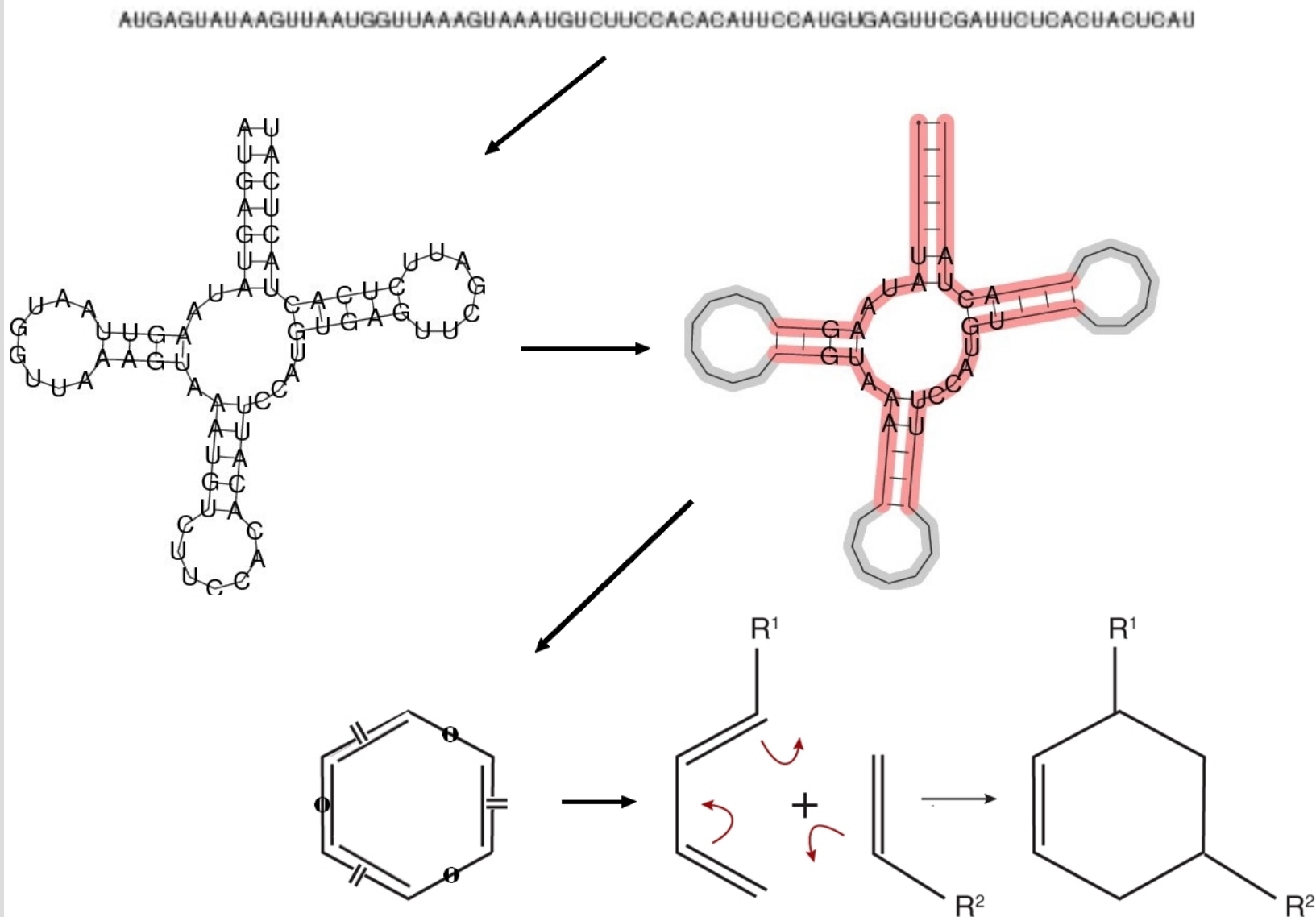
- Evaluation = metabolic pathway analysis
- Scores = metabolic yield
- Selection = 50% with highest yield
- Genetic operations = Mutation, Duplication
- Reproduction = asexual
- Protocol
 - Genetic events
 - Phenotypic events
 - Metabolic network
 - Network properties

Integrating the sequence-to-structure map

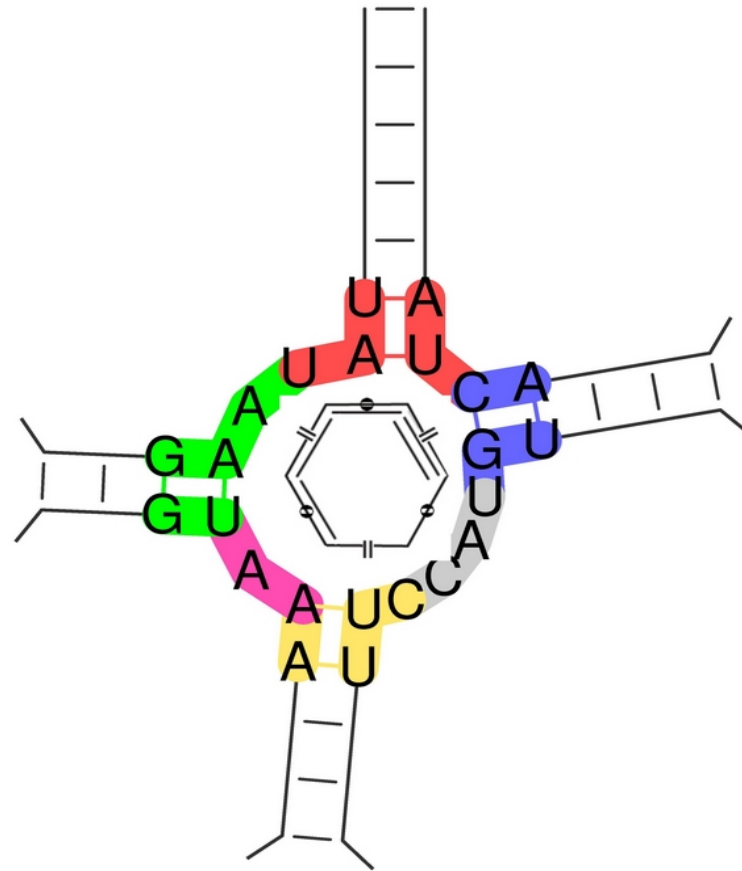
- Genotype = RNA-sequence
- Phenotype = Graph rewrite rule
- Mapping:
 - Sequence
 - > Structure
 - > reduced Structure
 - > Features
 - > ITS ID
 - > ITS
 - > Graph rewrite rule



Integrating the sequence-to-structure map



Integrating the sequence-to-structure map

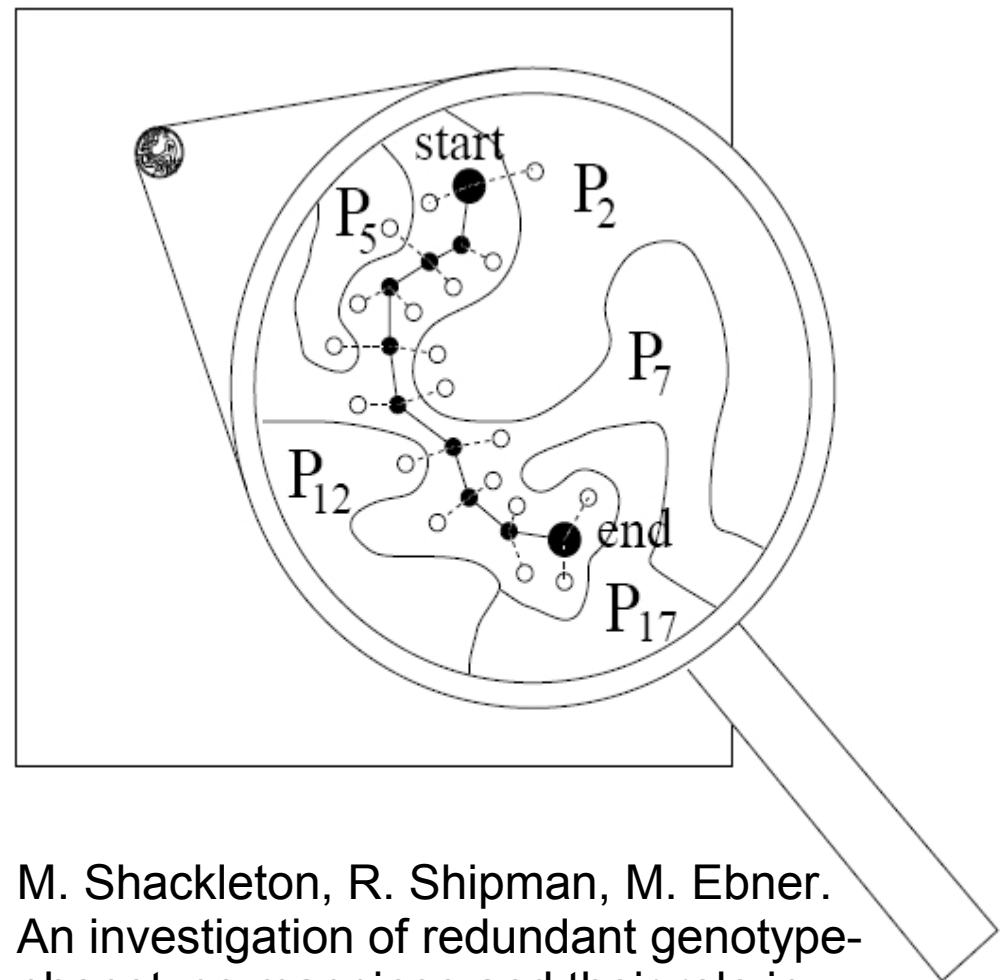


Section	Loop	C-G pair	Neighbor > 5 bp	Bond	Valence	Seq. (loop)	Sequence
1 (red)	yes	0	yes (+1)	1 " - "	3	4	4 = C
2 (blue)	yes	1	yes (+1)	2 " = "	4	1	4 = C
3 (gray)	no	-	no	0 " "	3	4	4 = C
4 (yellow)	yes	0	no	0 " "	1	4	4 = C
5 (pink)	no	-	yes (+1)	1 " - "	2	2	2 = O
6 (green)	yes	1	no	1 " - "	3	3	3 = N

Results / Comparison

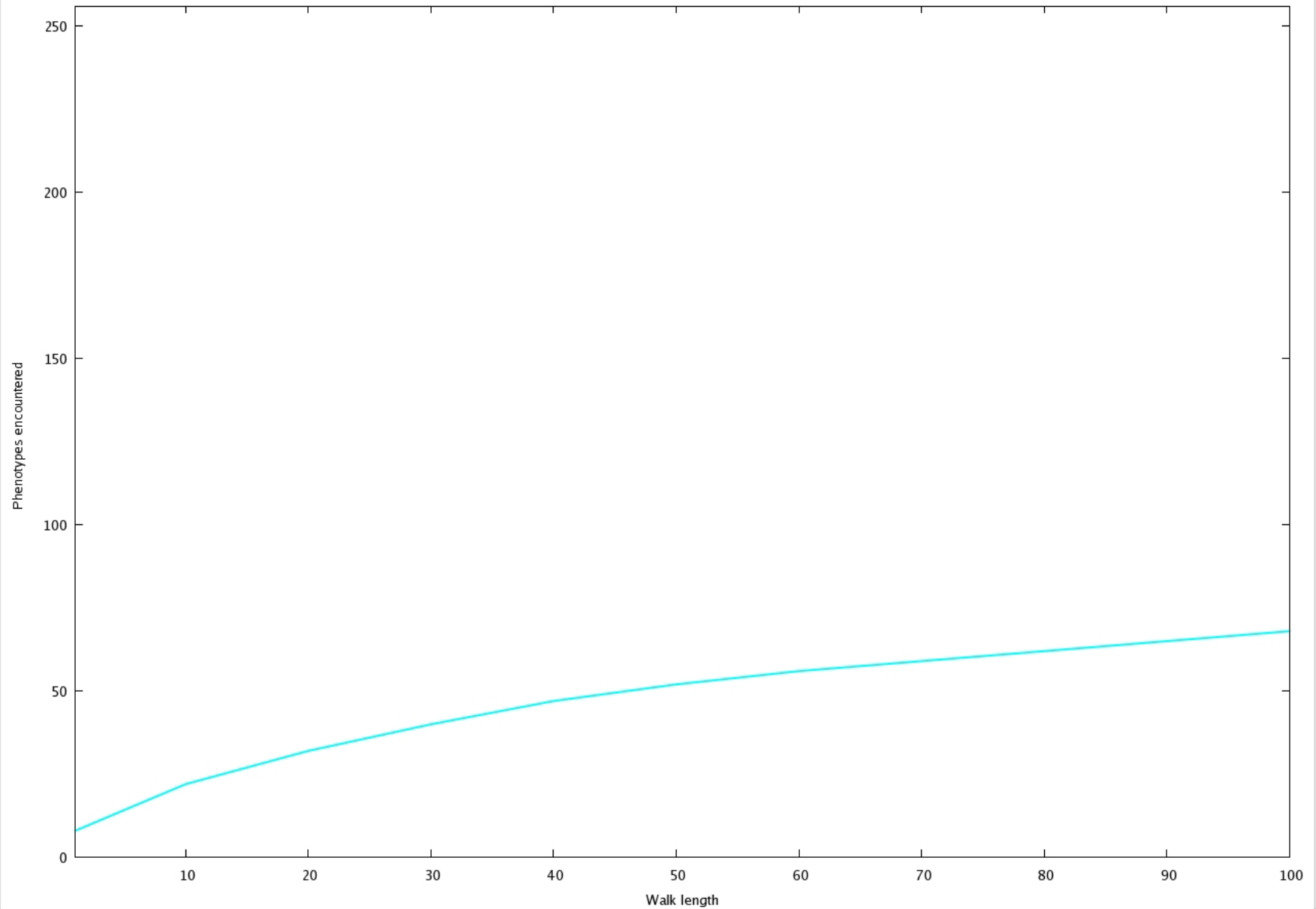
- Comparison of different mappings
 - Neutral walk
 - Neutral networks
 - Redundancy
 - Extension
 - Entanglement
- Results from the simulation
 - Networks
 - Connectivity
 - Enzymes
 - Evolution

Genotype Space

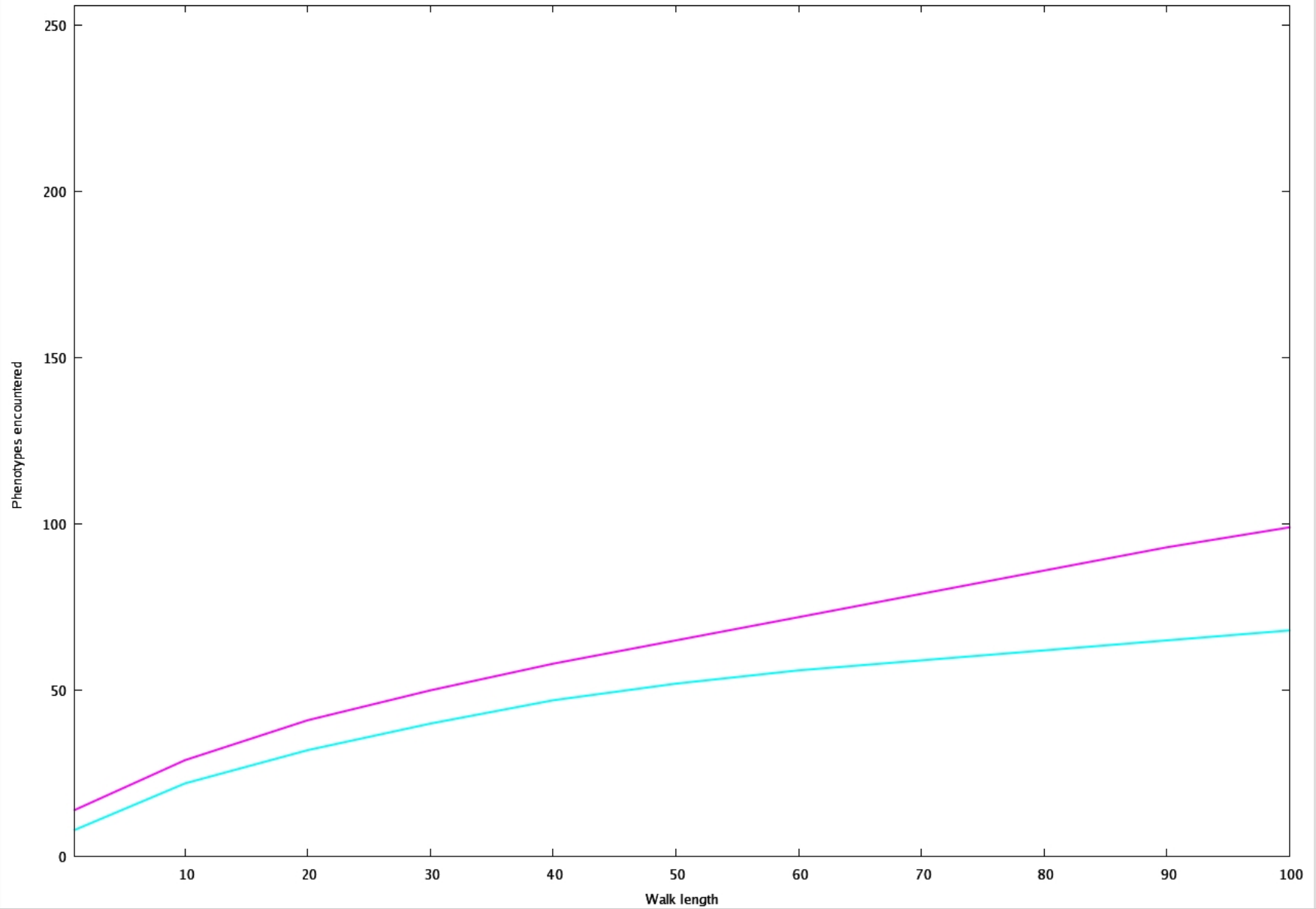


M. Shackleton, R. Shipman, M. Ebner.
An investigation of redundant genotype-phenotype mappings and their role in evolutionary search.

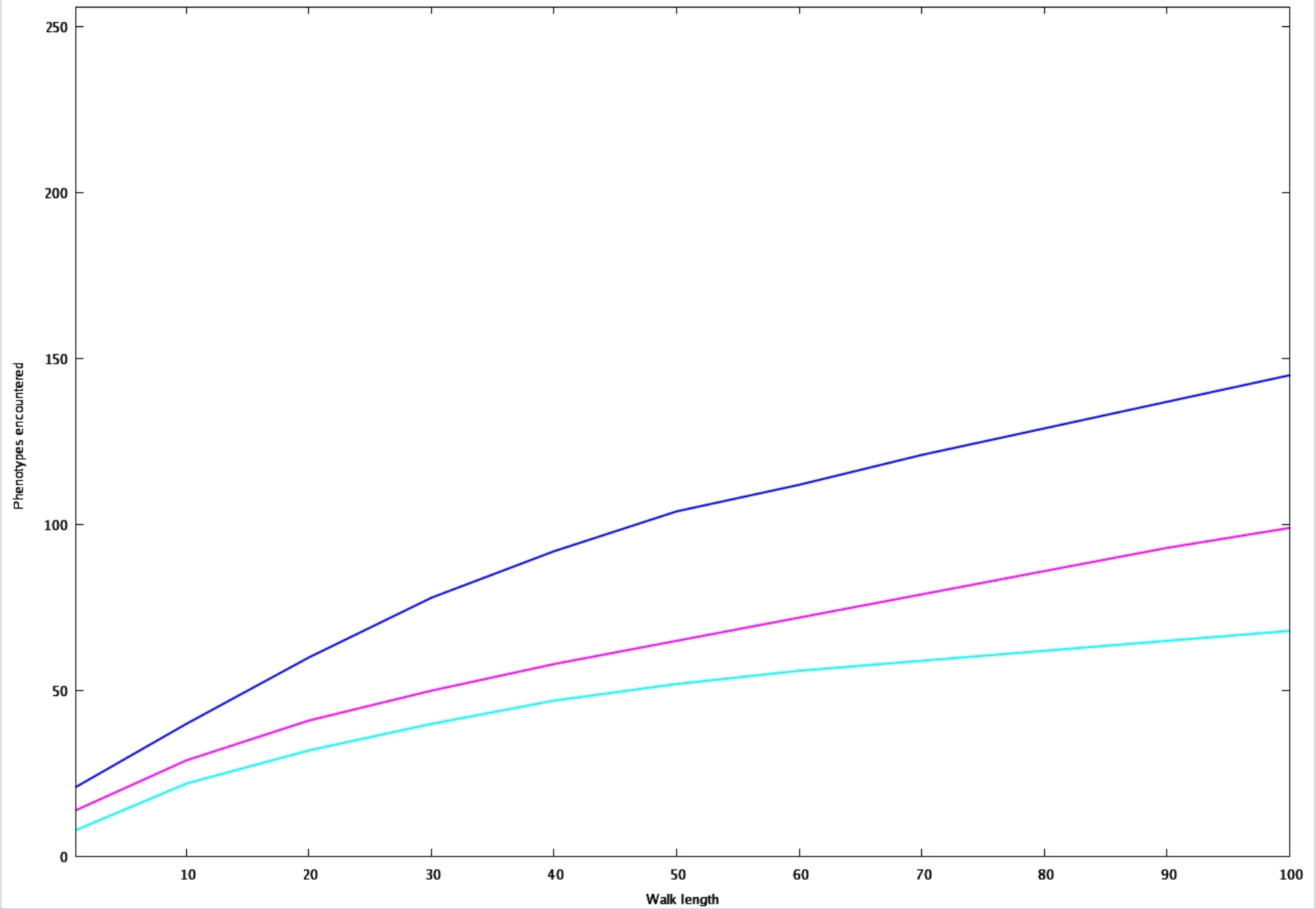
Results / Comparison



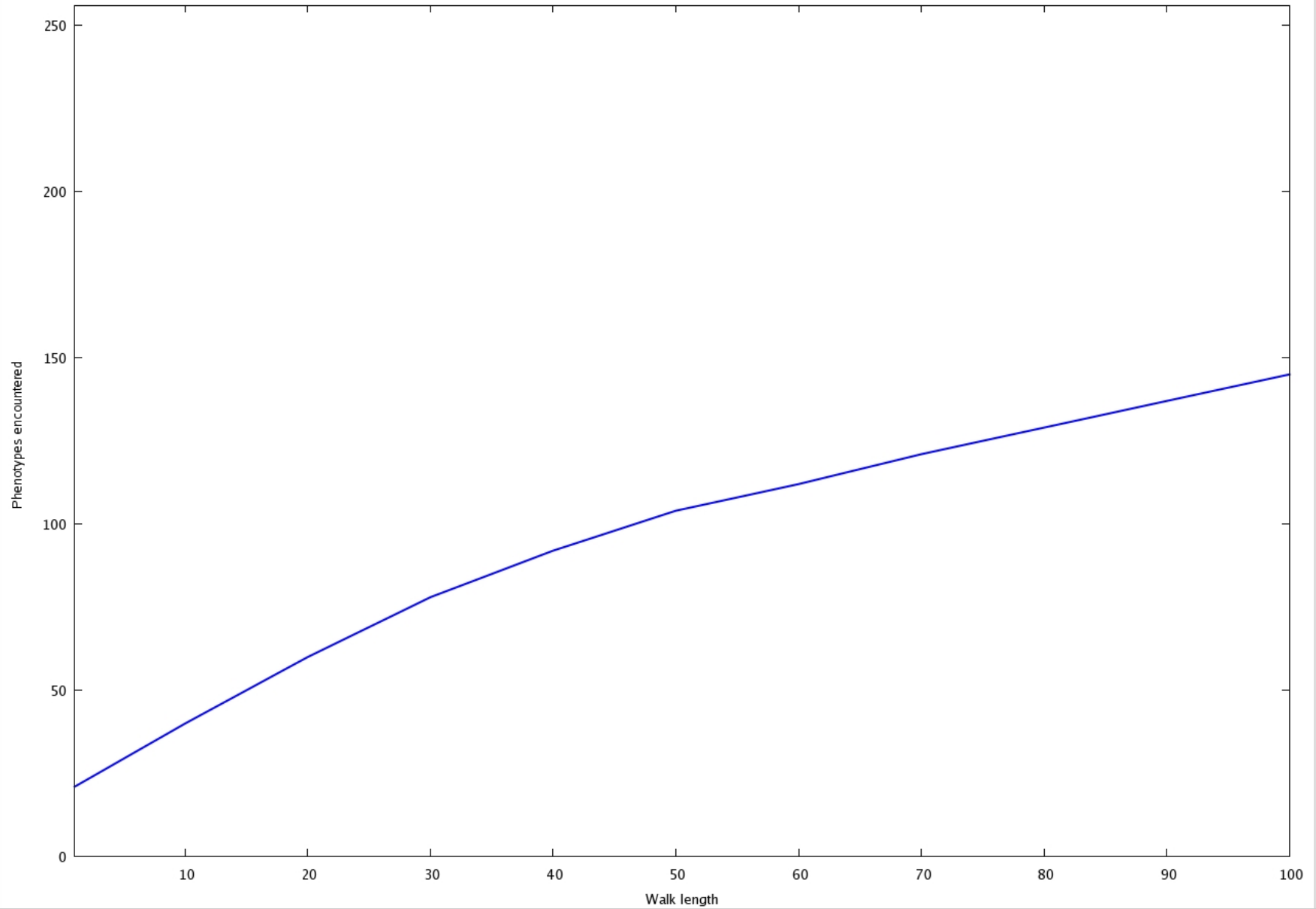
Results / Comparison



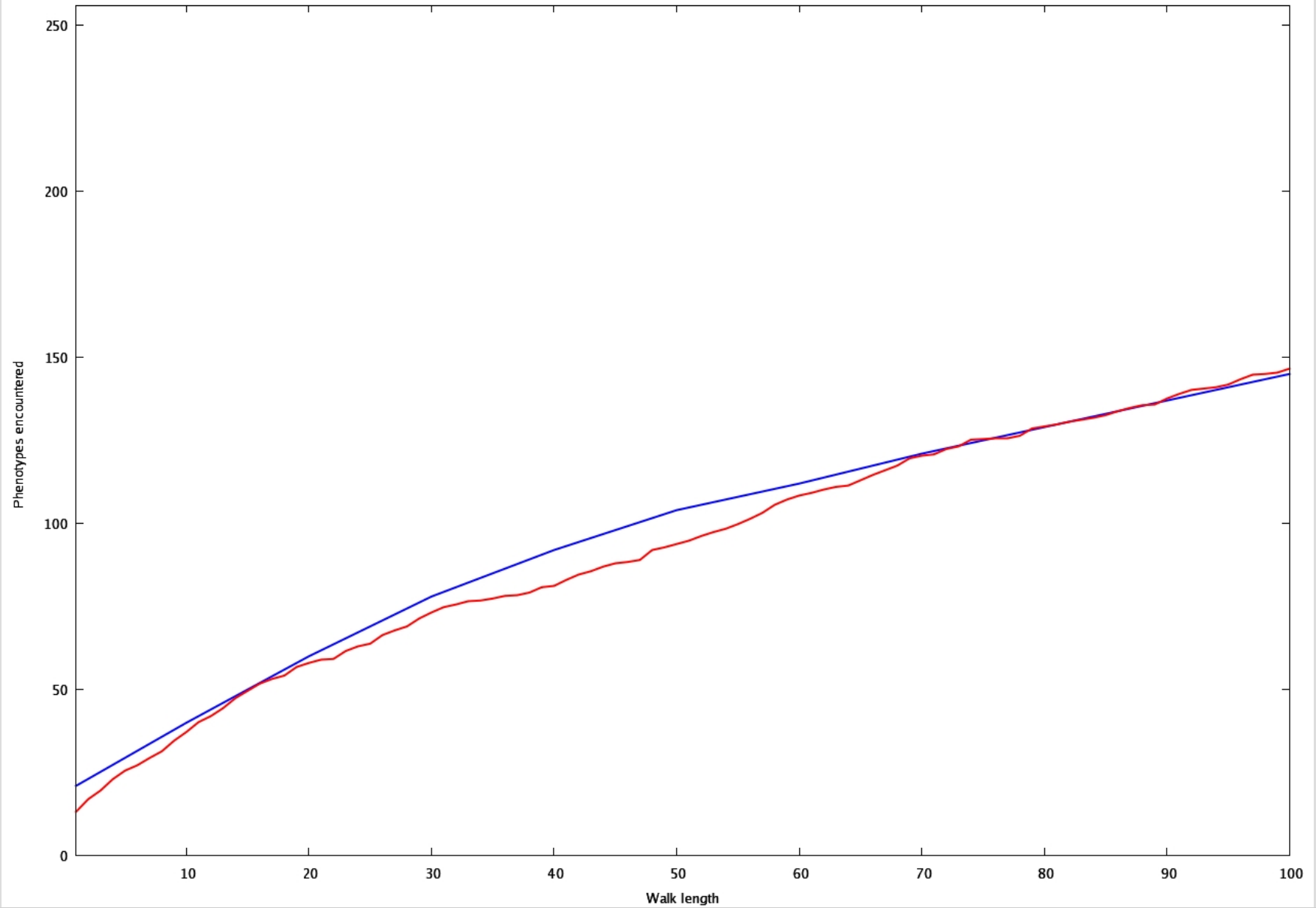
Results / Comparison



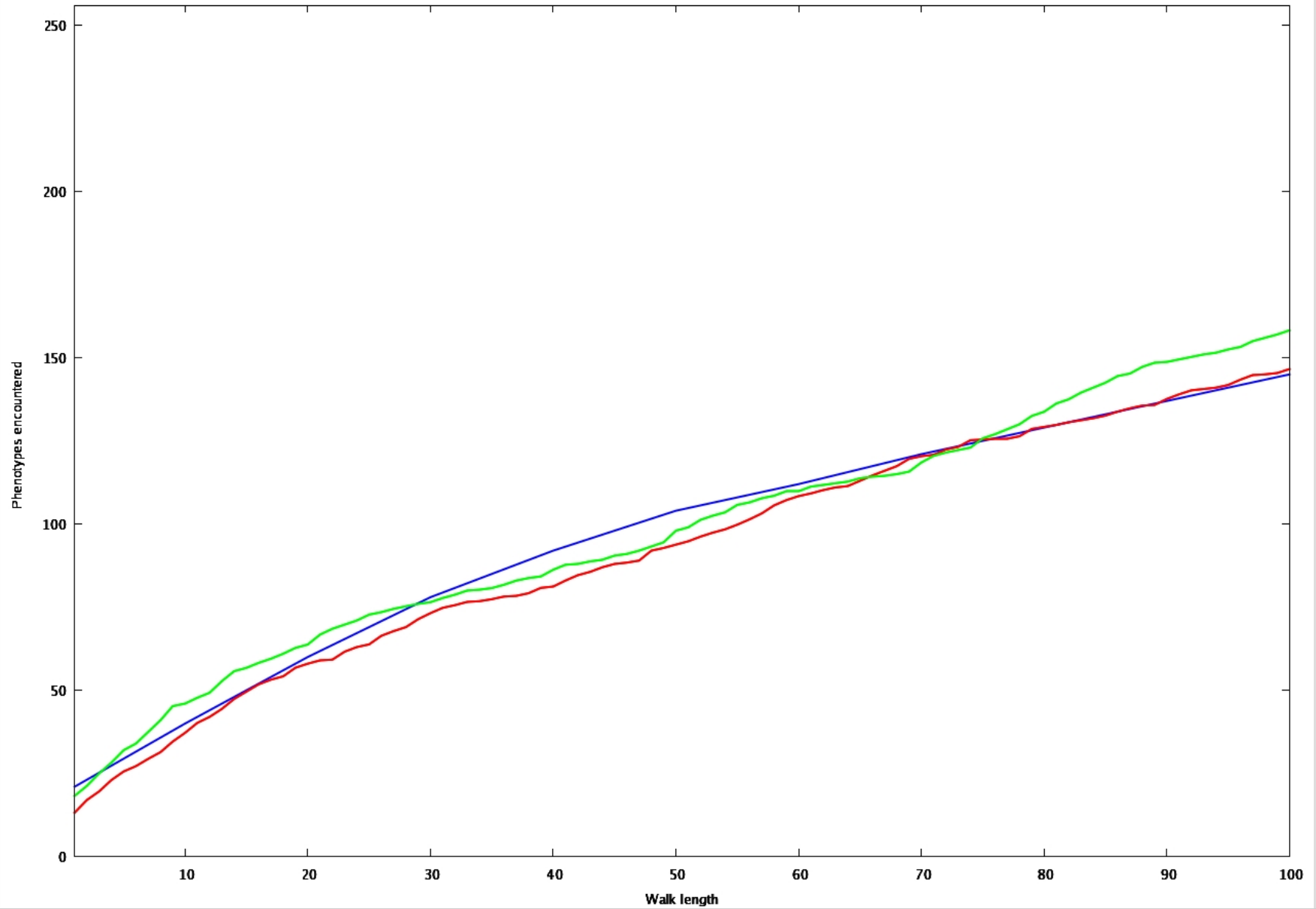
Results / Comparison



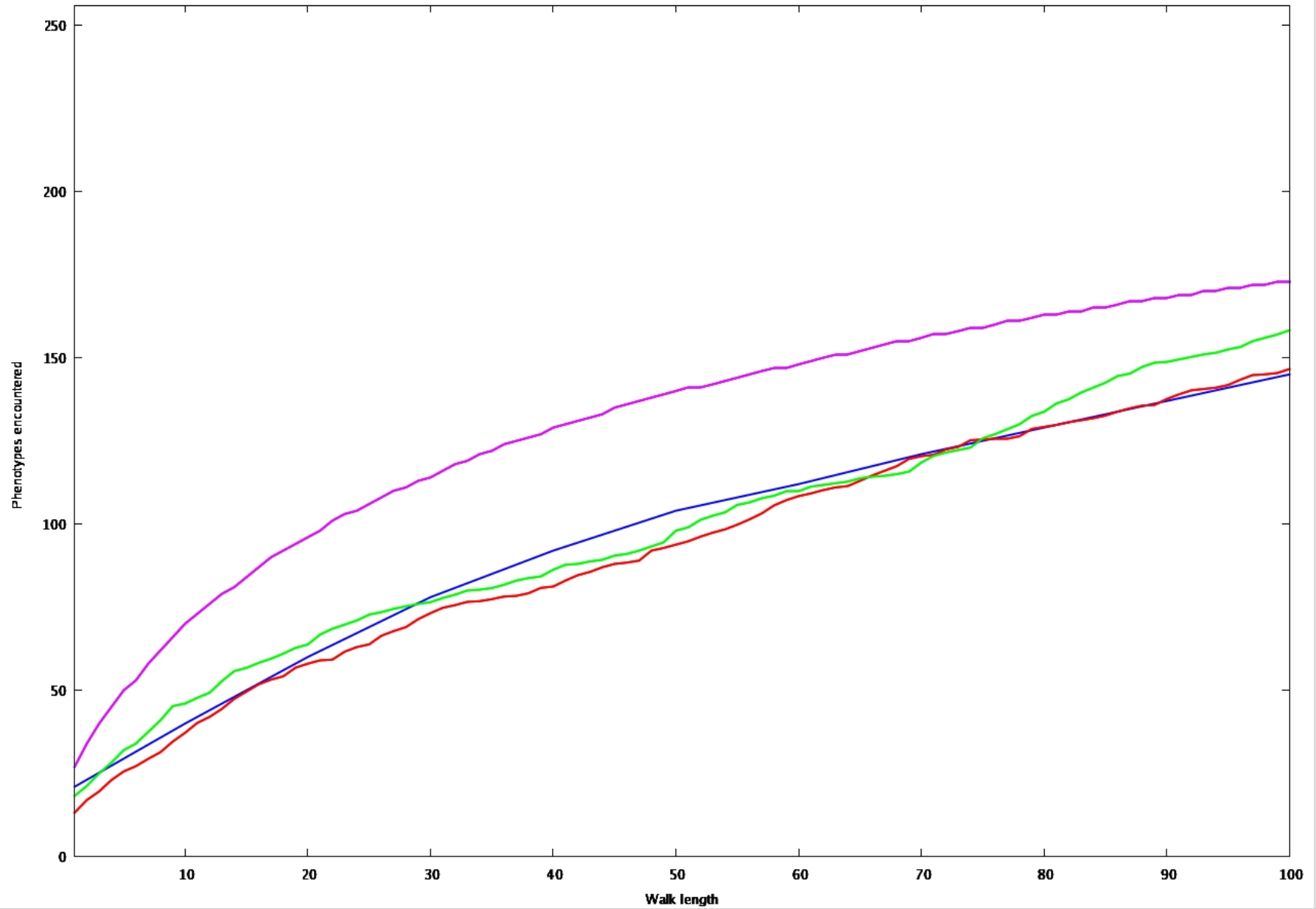
Results / Comparison



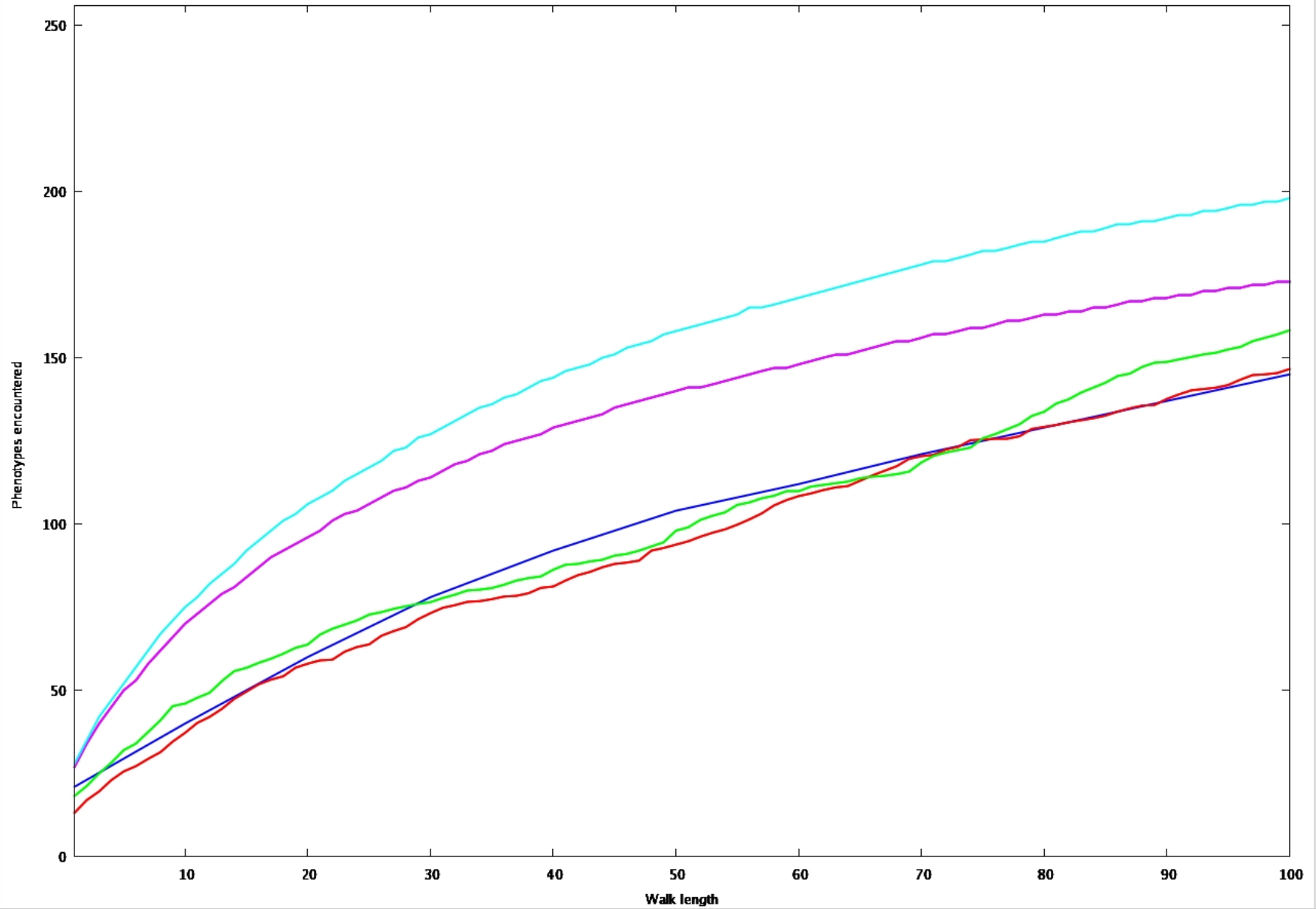
Results / Comparison



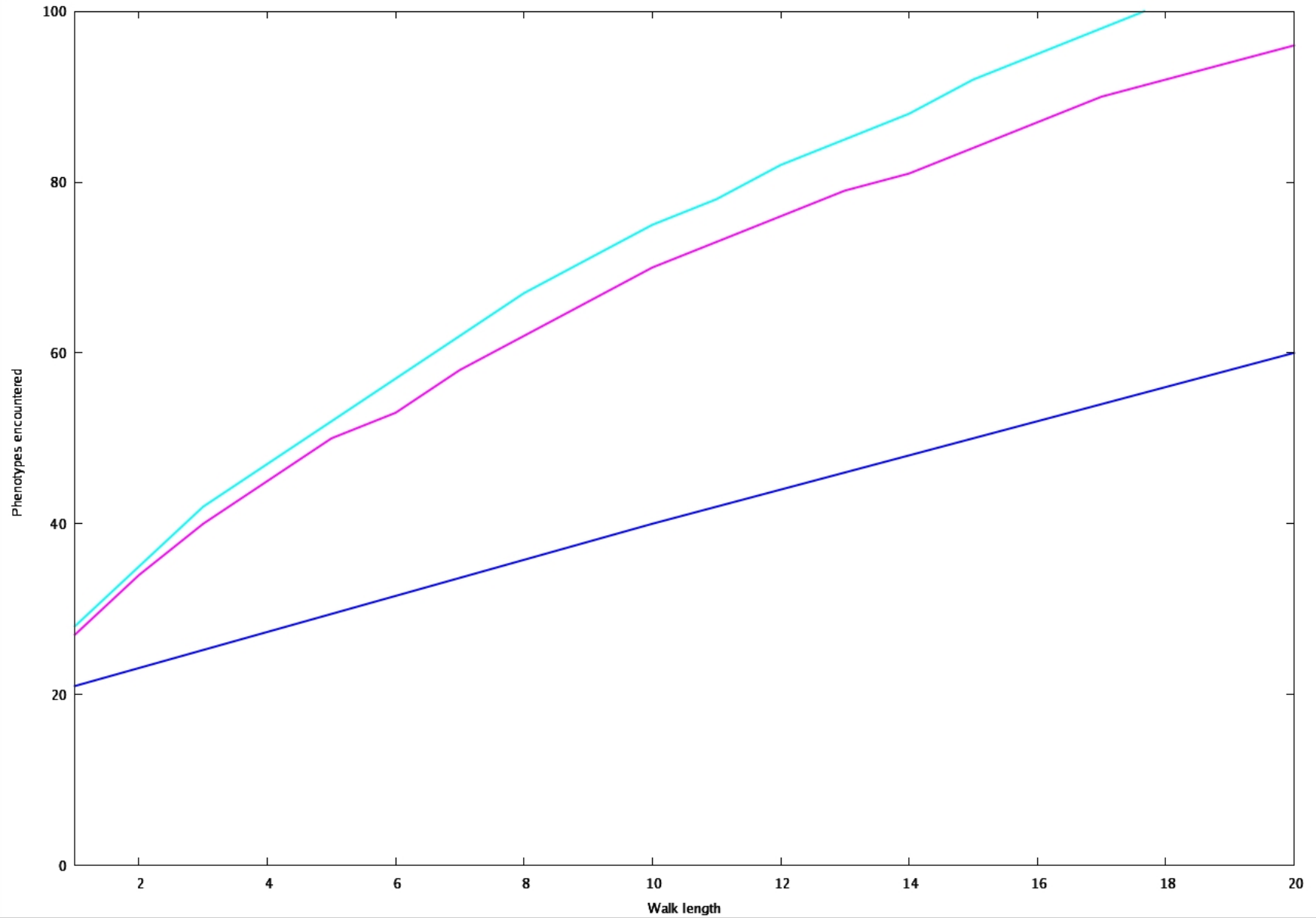
Results / Comparison



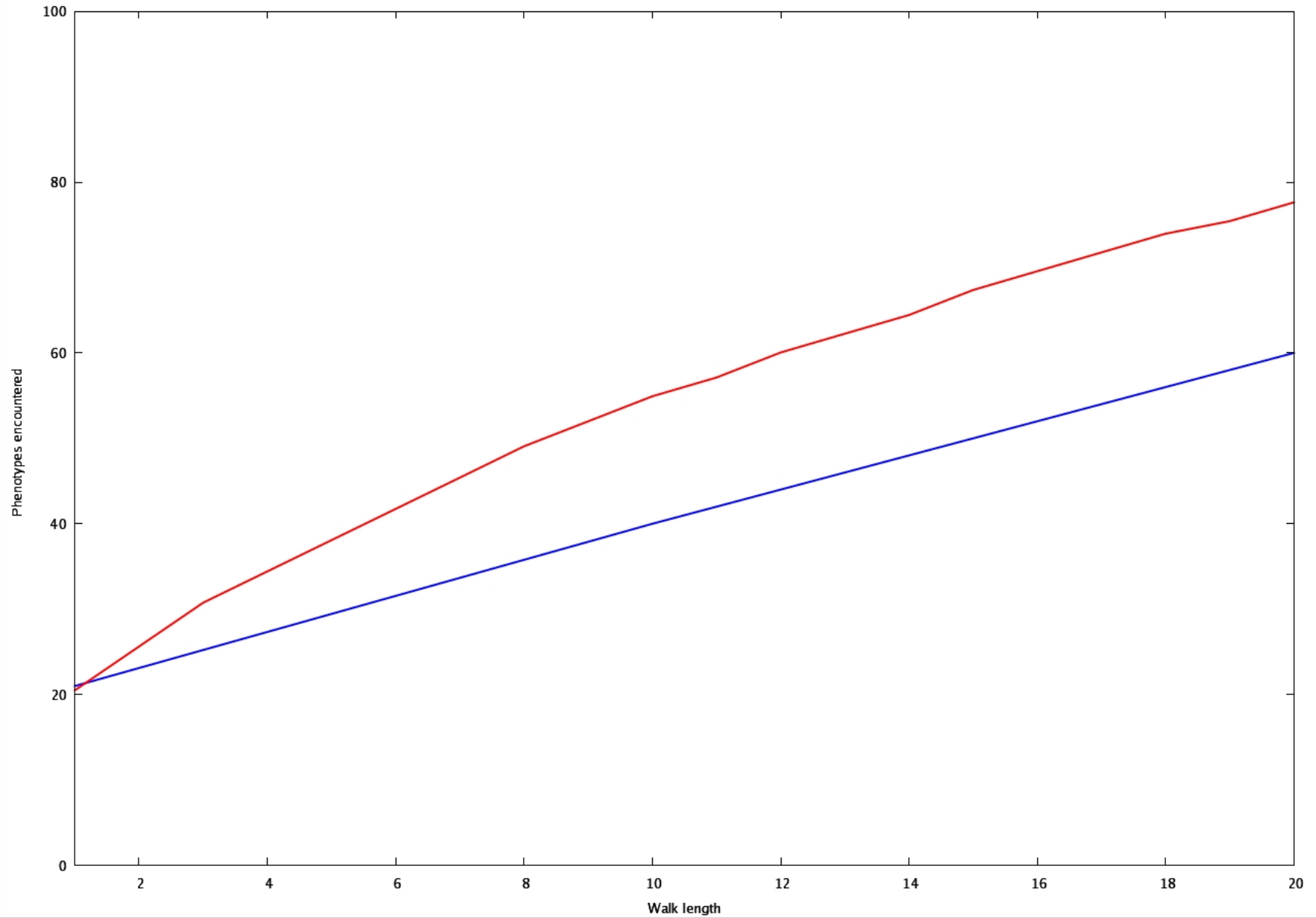
Results / Comparison



Results / Comparison



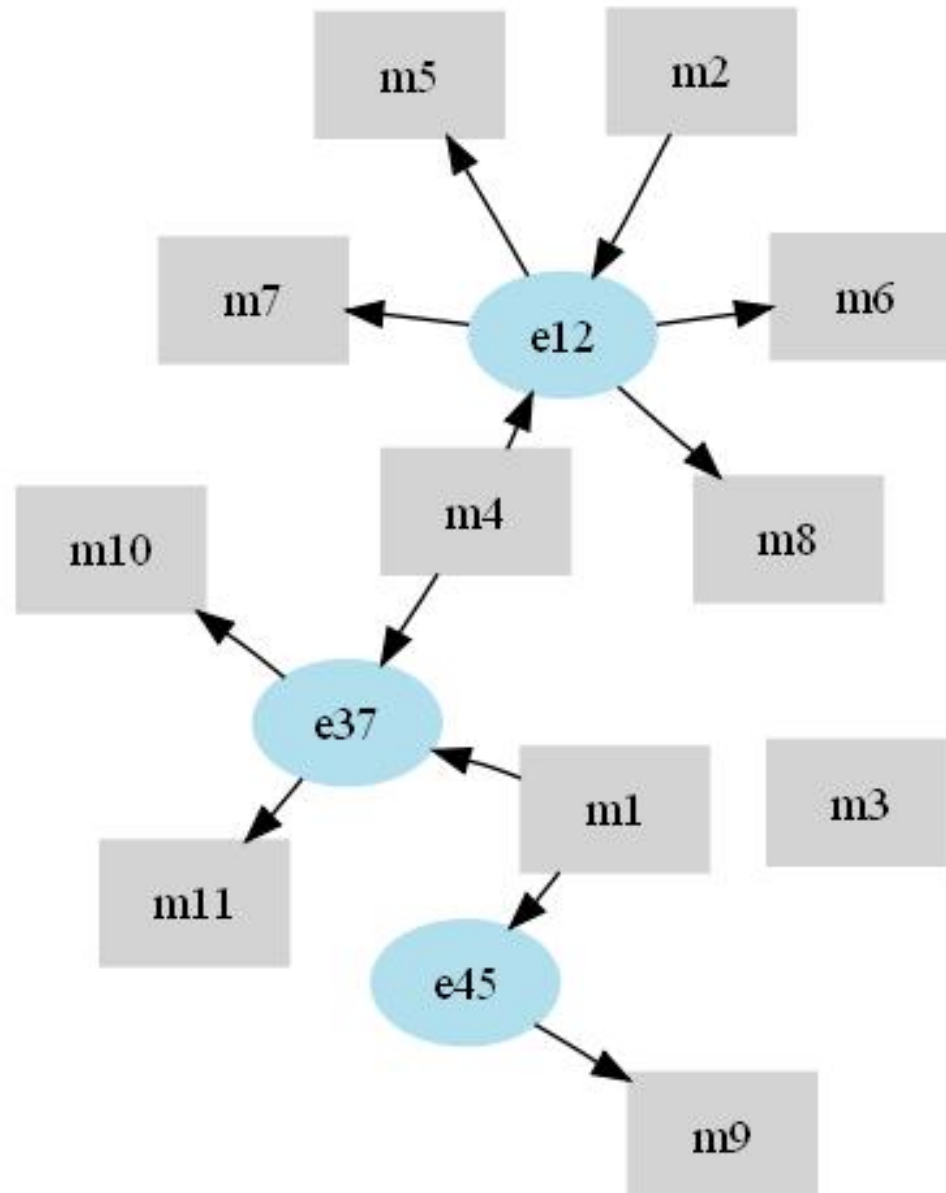
Results / Comparison



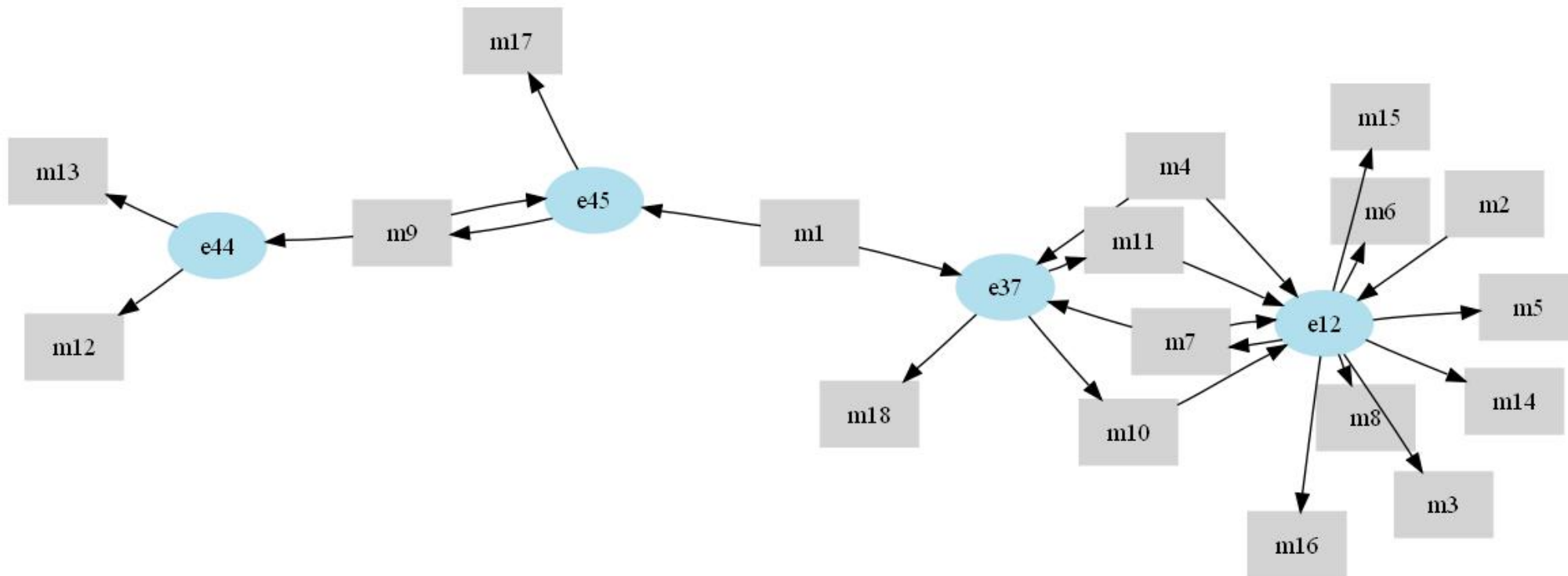
Results / Comparison

- **Neutrality:**
 - RNA : 50%
 - CA : 42%
 - **RBN : 58% --> most robust, but!**
- **Encountered Phenotypes (length=100)**
 - **RNA : 198 --> most evolvable**
 - CA : 99
 - RBN : 145
- **Encountered Phenotypes (length=20)**
 - **RNA : 106 --> most innovative**
 - CA : 41
 - RBN : 60
- **Access to phenotypes via one-point-mutation:**
 - **RNA : 28 --> most intertwined !!!**
 - CA : 14
 - RBN : 21

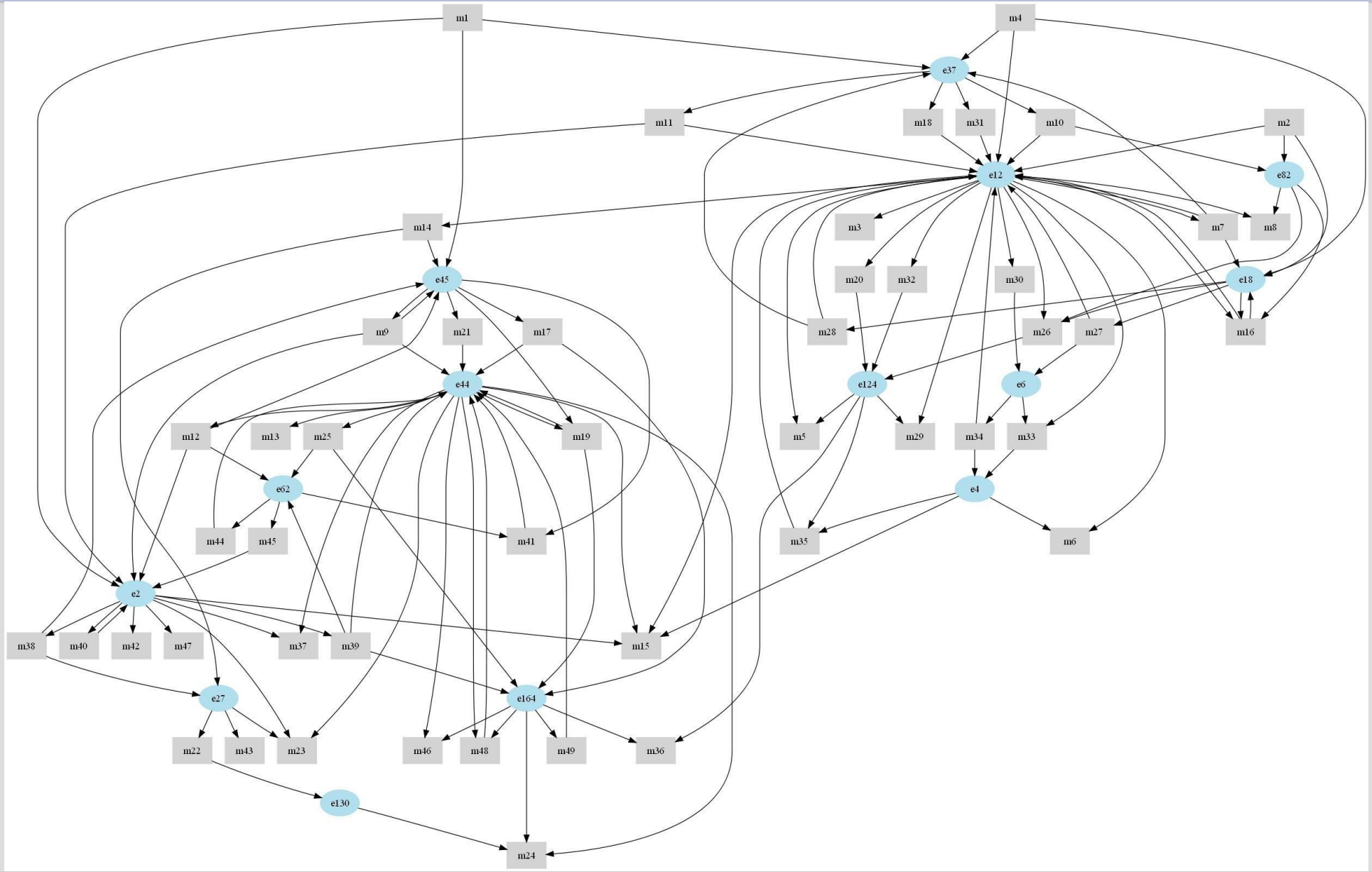
Results / Comparison



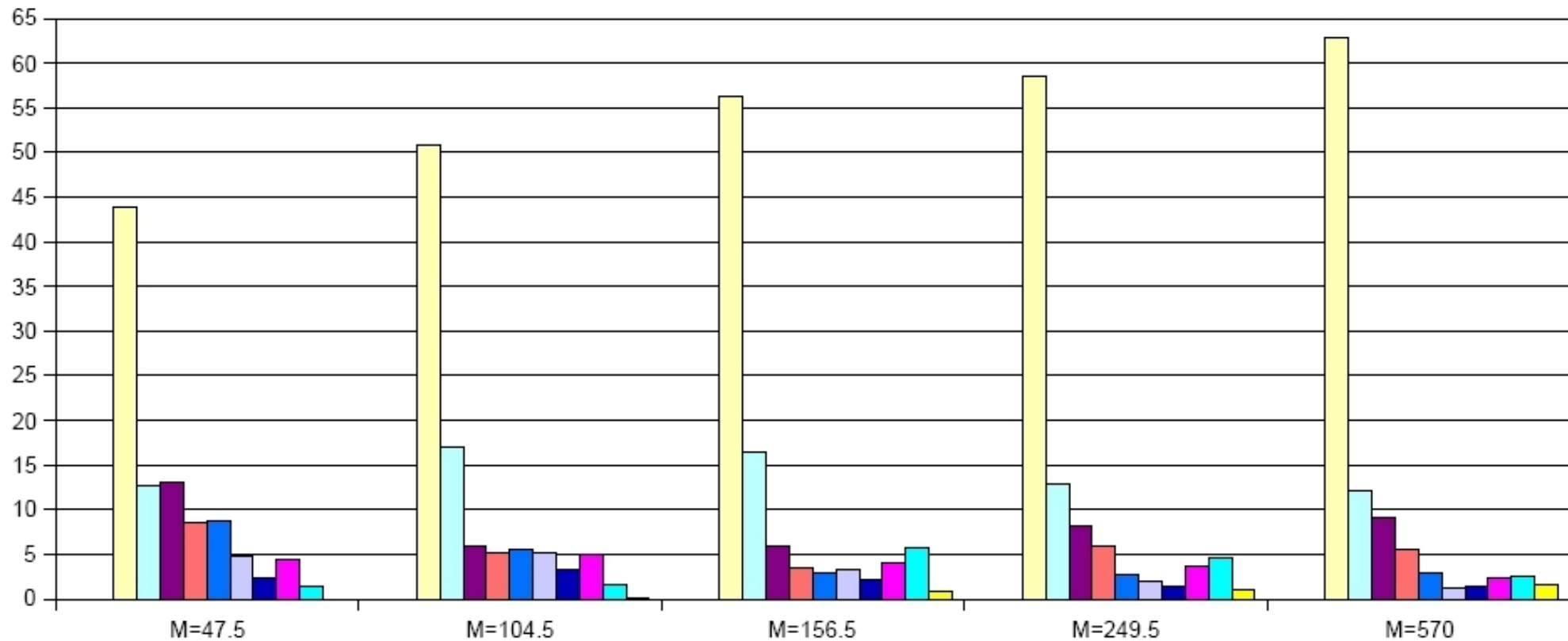
Results / Comparison



Results / Comparison



Results / Comparison

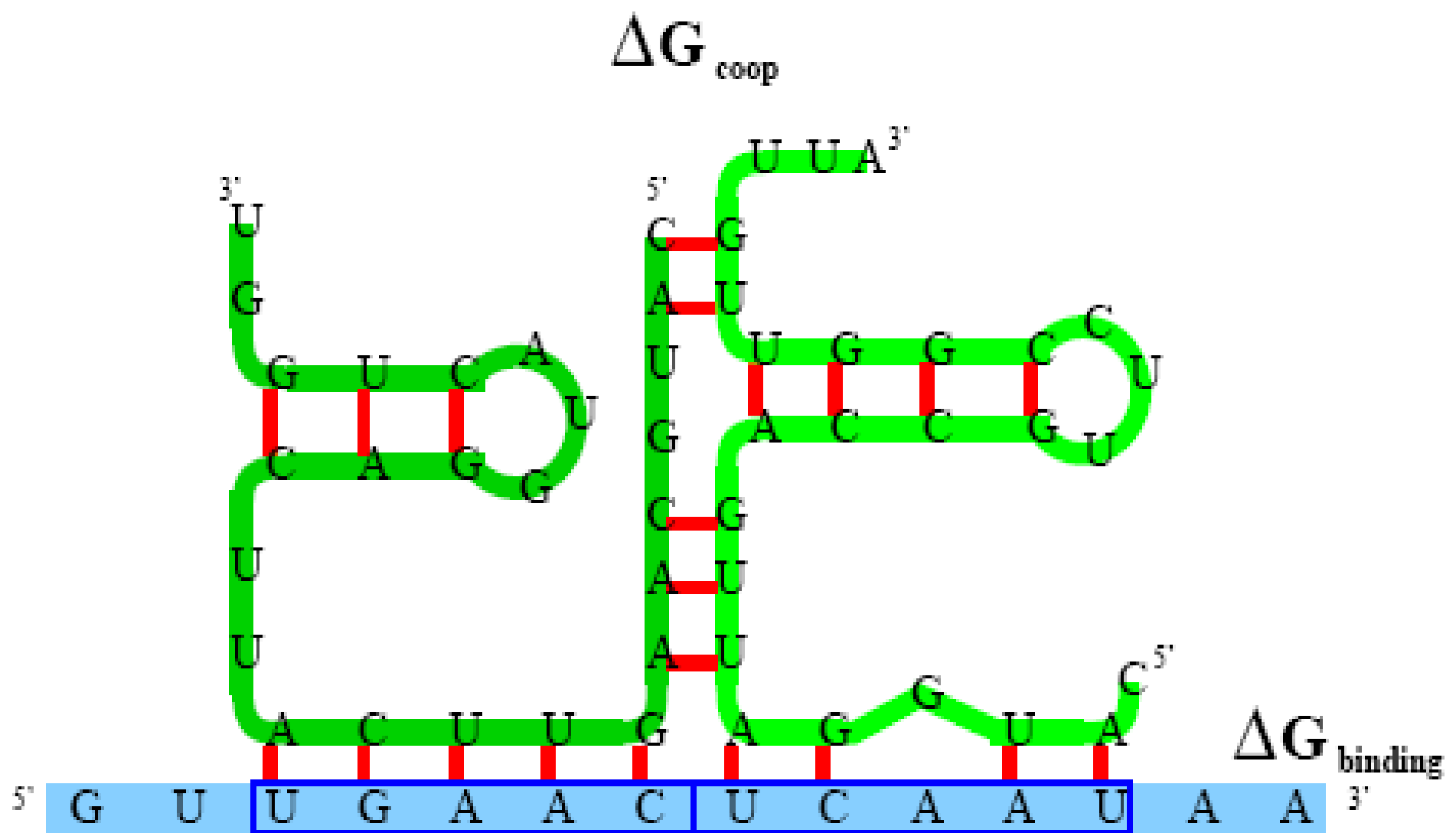


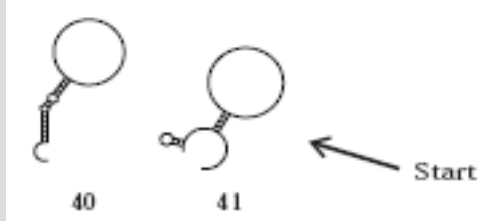
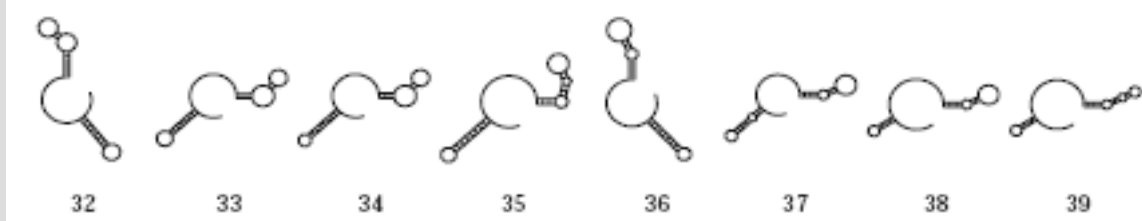
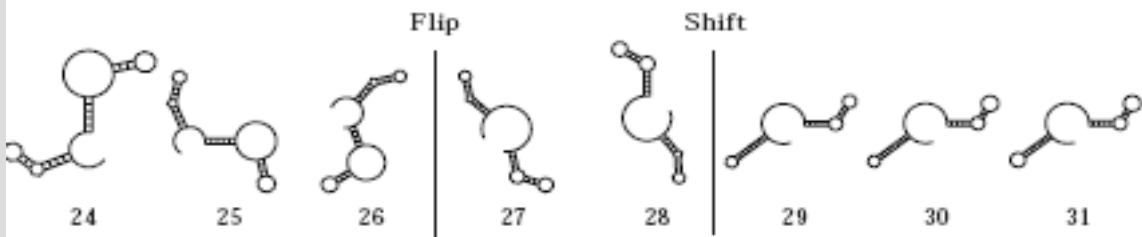
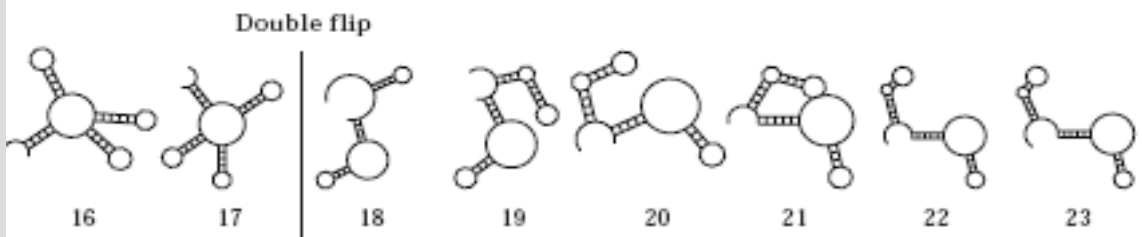
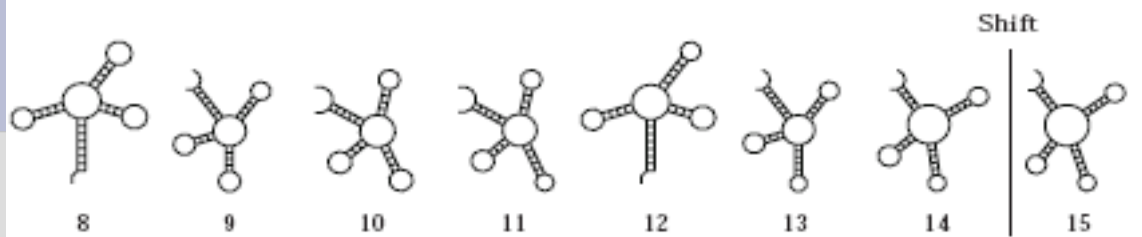
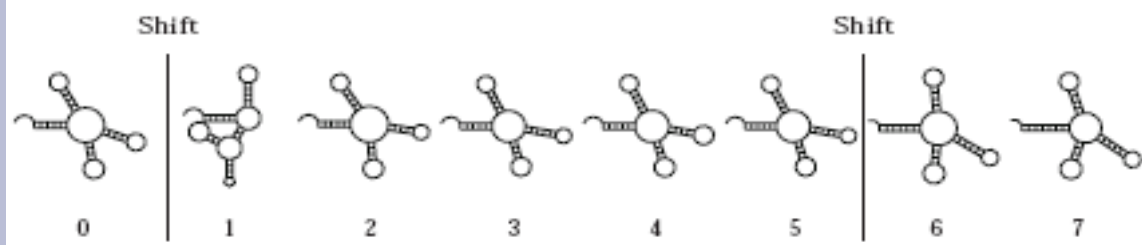
Enzyme	e37	e45	e12	e44	e18	e27	e6	e82	e4	e62	e124	e130	e2
Generation	1	1	1	2	3	3	7	10	14	42	44	51	53
Connectivity	4	5	12	9	4	2	2	2	2	2	3	1	6

Specificity of enzymes in the example network

Application / Outlook

- Evolutionary Algorithms
- Evolutionary studies
 - Metabolic networks, catalytic elements
 - Gene regulatory networks
- Simulation
 - Mapping
 - RNA
 - Evolution
 - Selection vs neutral theory
 - Smooth vs rugged parts in phenotype / fitness landscape
 - Network / System properties





Acknowledgement

- University of Vienna
 - Christoph Flamm
 - TBI Group
- University of Leipzig
 - Peter Stadler
- EBI
 - Lukas Endler
- Funding
 - Vienna Science and Technology Foundation
WWTF

Thanks for your attention.

Questions? Comments? Suggestions?