

Genotype-phenotype mapping, continued

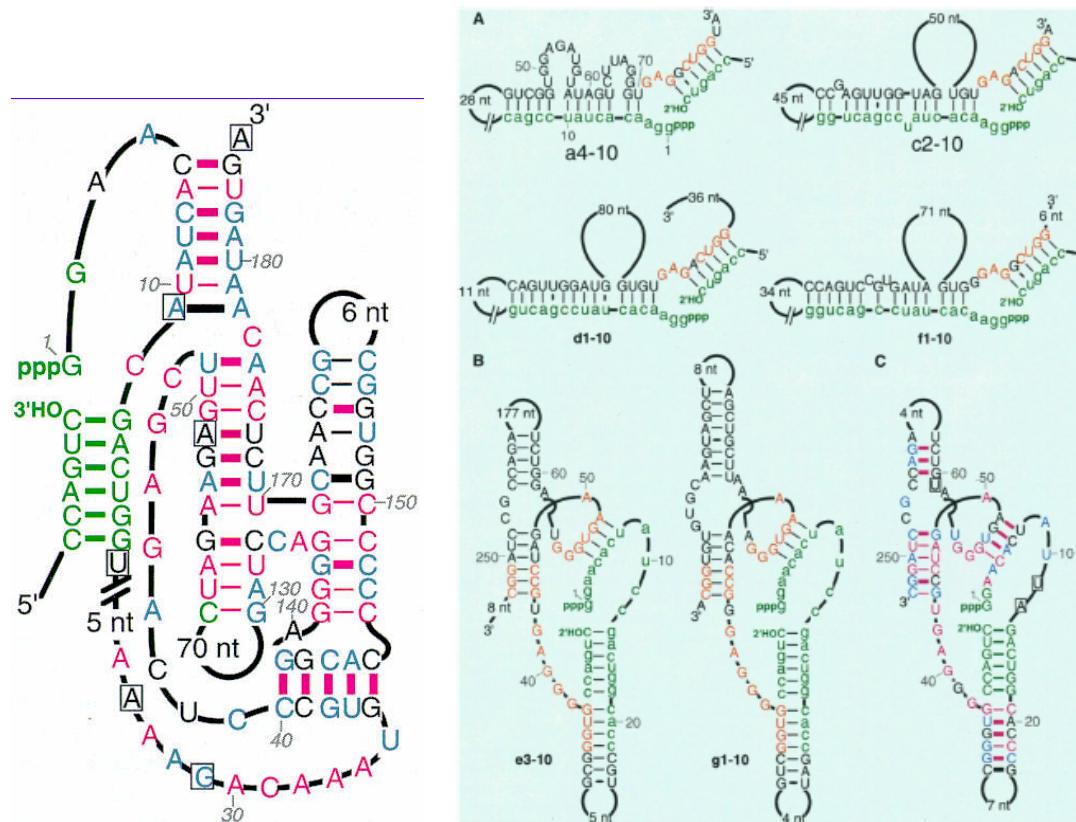
Course computational biology 2018/2019; Paulien Hogeweg;
Theoretical Biology and Bioinformatics Grp Utrecht University
₁

RNA genotype-phenotype mapping so far...

- “smoothness within ruggedness”
single mutation can be neutral and can change ‘everything’
- percolating and intercalating neutral networks
from smooth-rugged towards neutral networks
- no local peaks: detours
- phenotypic vs genotypic information threshold
- diffusion on neutral networks (D prop.to λ)
- adaptive walk with majority of neutral mutations
- reconciliation neutral and adaptive evolution
- RNA landscape “ideal” for evolution
- +++) (today)

(cf VERY efficient in vitro evolution (????))

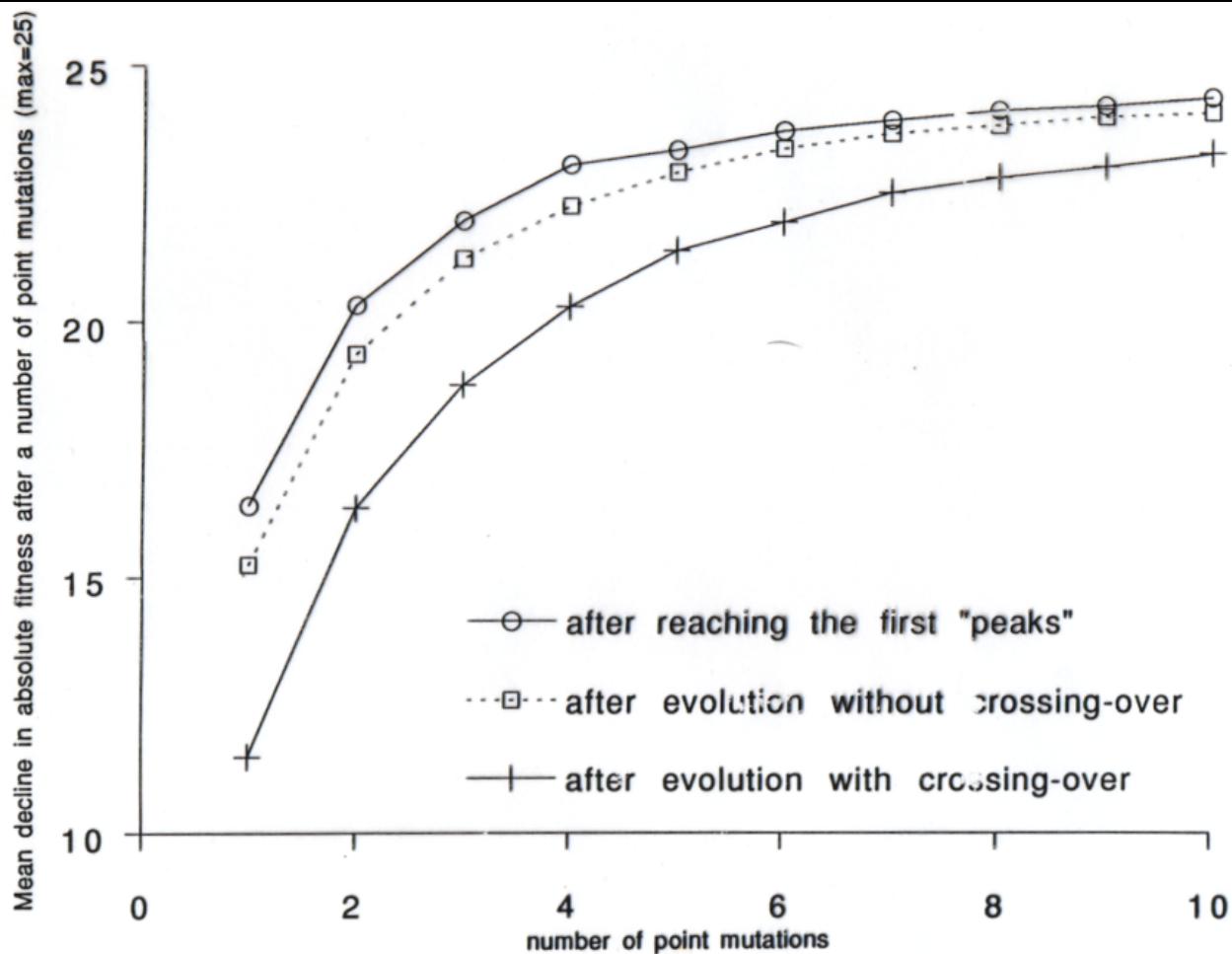
MOREOVER: phenotype → function mapping



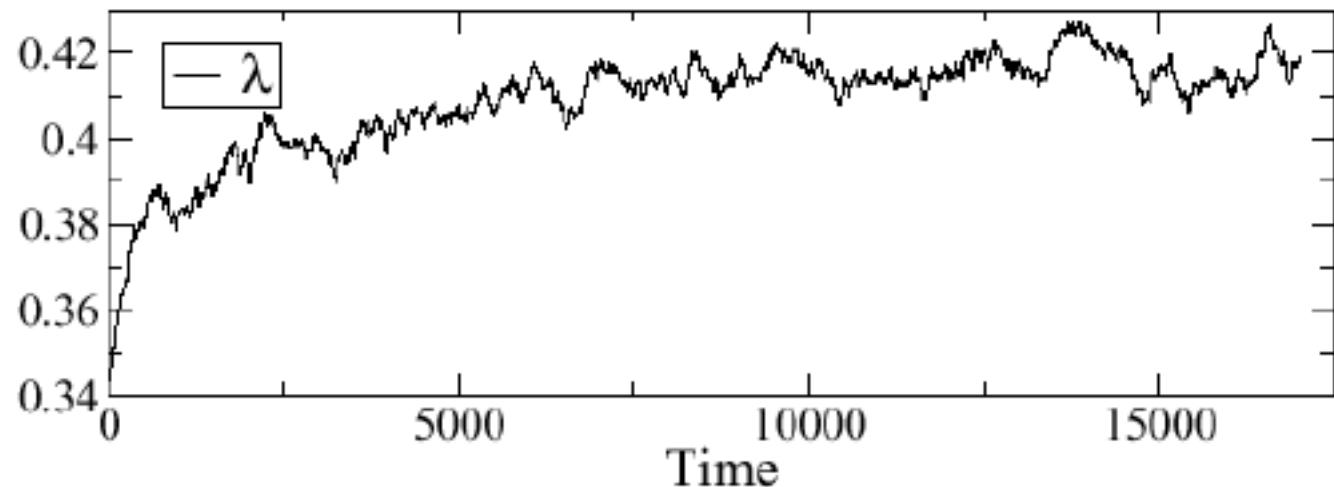
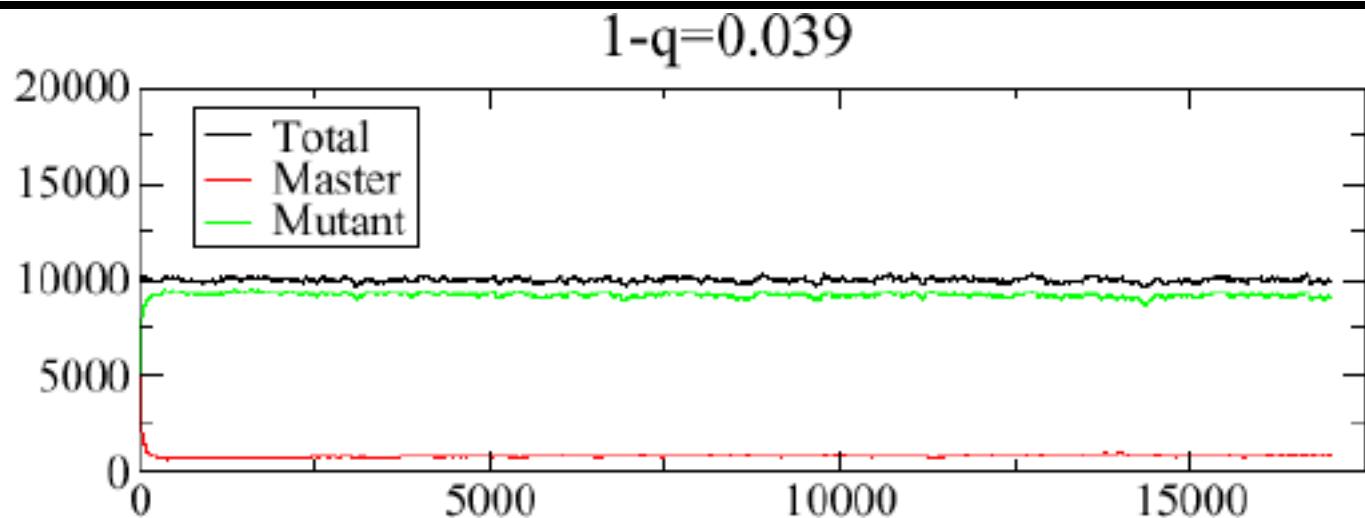
Alternative ligases (Ekland et al 1995)

'tyranny' of small motifs... or complex structures?

'drift' on neutral network not 'neutral':
(1) Longterm RNA evolution: fitness of mutants



(2) Evolution towards high lambda



redundant genotype-phenotype mapping: choice of coding

Evolution towards 'flatter parts'

== Mutational robustness

== high connectivity of neutral network

== MAX EIGENVECTOR OF CONNECTION MATRIX
(van Nimwegen 2000)

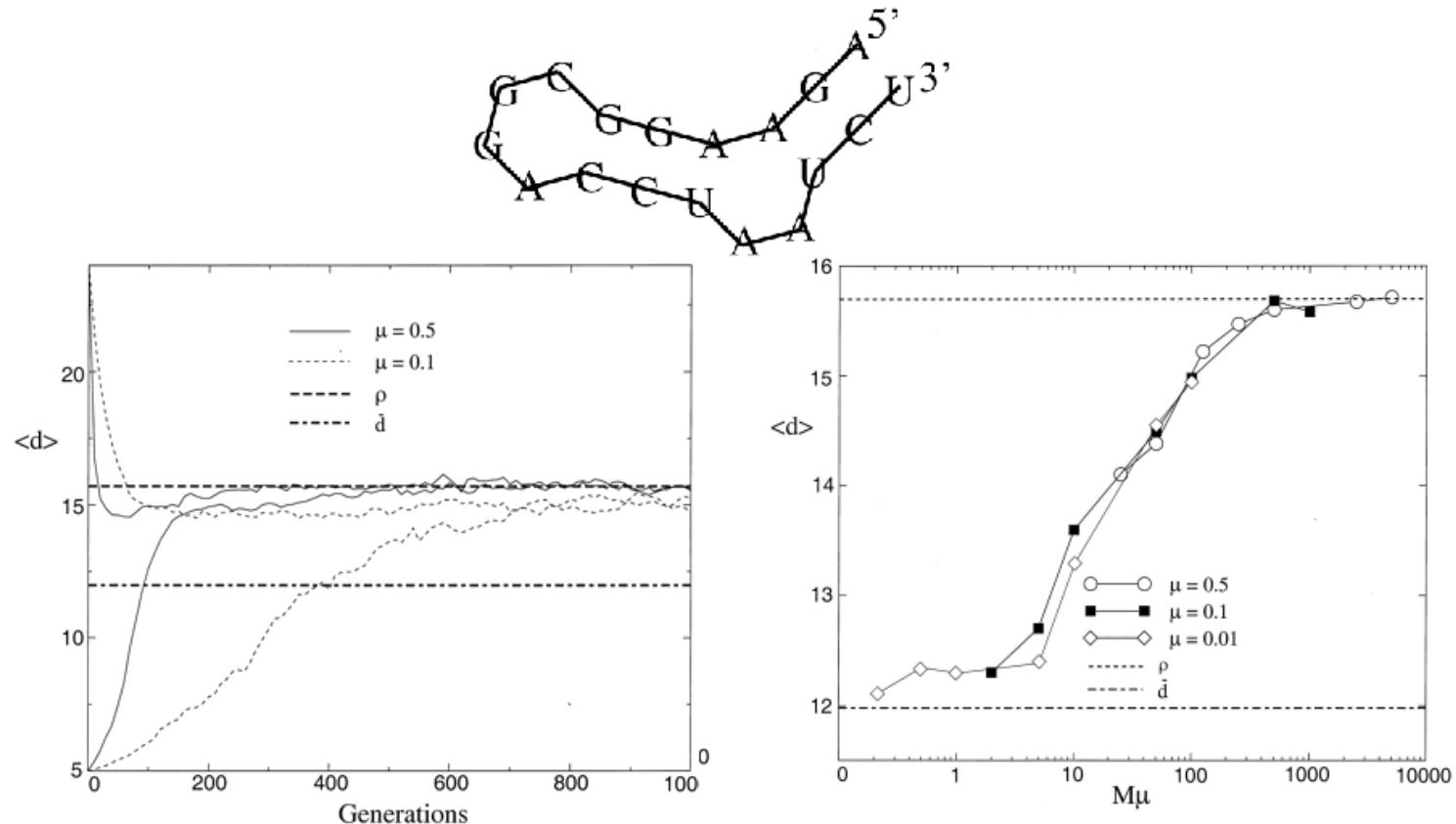
compare blind ant (moves with prob. rel neutral NB)

-- > same freq in each node)

myopic ant (moves with fixed probability)

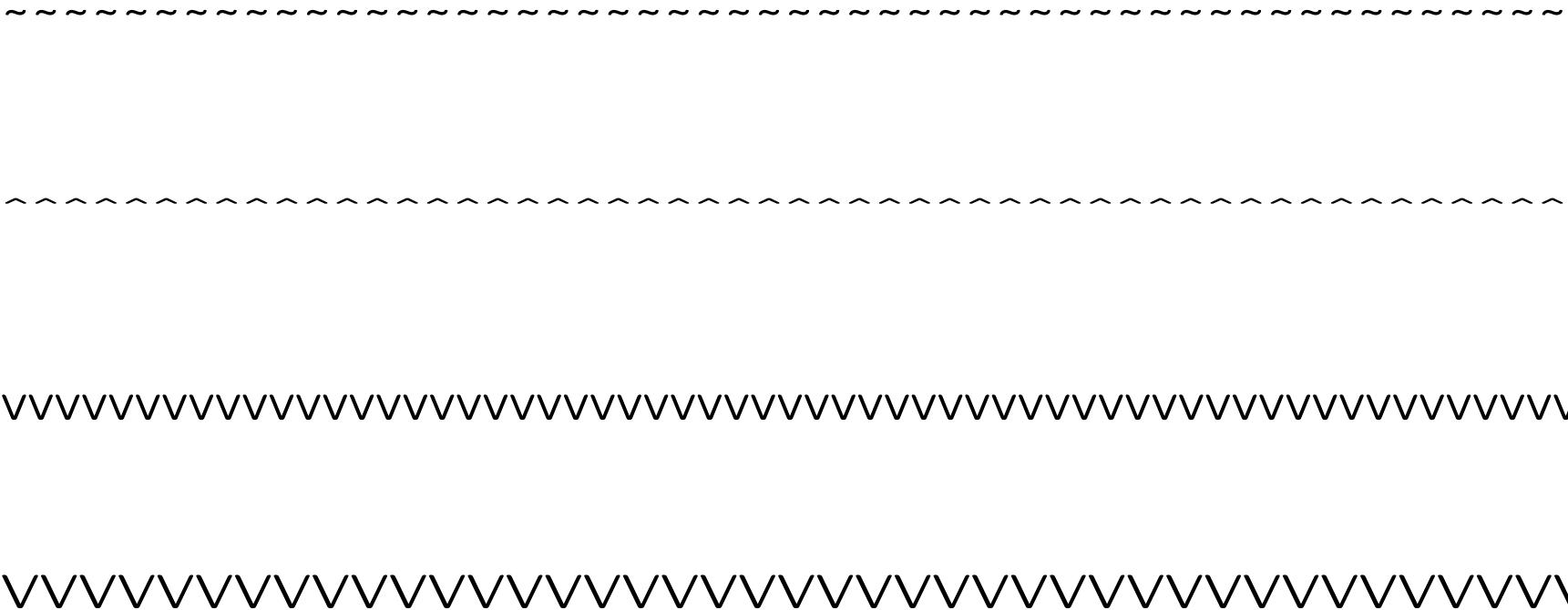
-- > $D = \hat{d} + Var(d)/\hat{d}$

Evolution towards mutational robustness == largest eigenvalue of connection matrix van Nimwegen et al PNAS 1999



walk along neutral path not neutral

**walk along neutral path not neutral....
how neutral is neutral**



**walk along neutral path not neutral....
how neutral is neutral**



~~~~~

neutral if above the information threshold!

## Implications evolution towards higher robustness

---

- more robustness —> more exploration ( $D \lambda$ )
- evolution of evolvability

cf Andreas Wagner e.g.

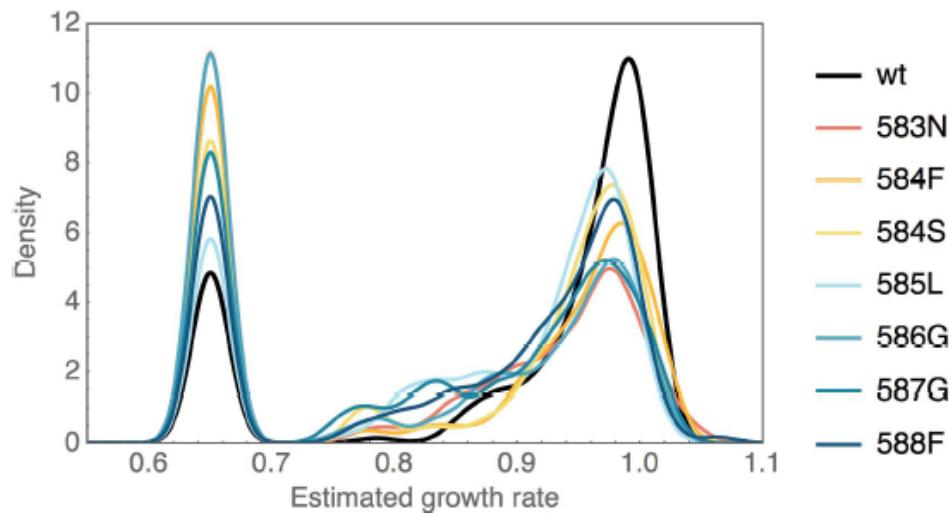
2008 Robustness and evolvability, a paradox resolved

2013 Robustness and evolvability in living systems

*high mutational load of recently evolved strains well known from traditional evolutionary experiments (Scharloo 1999: canalization)*

## example of intra-molecular evolved landscape negative epistasis

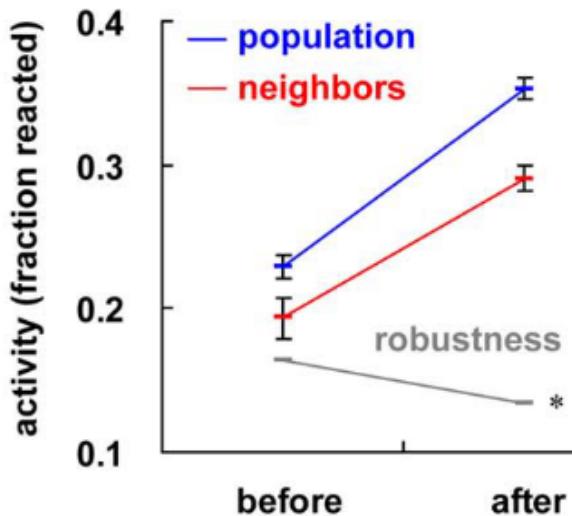
---



Hsp90,582-590 Effect on growth-rate of single point mutations from wild-type and from 7 (almost) neutral mutations

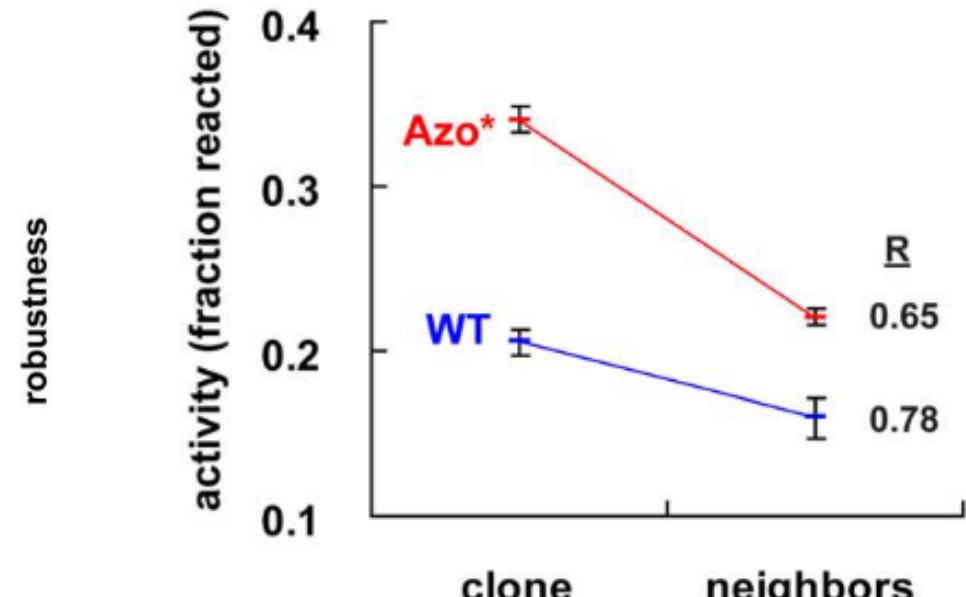
# experimental verification of evolution toward robustness ???

A



population level

B



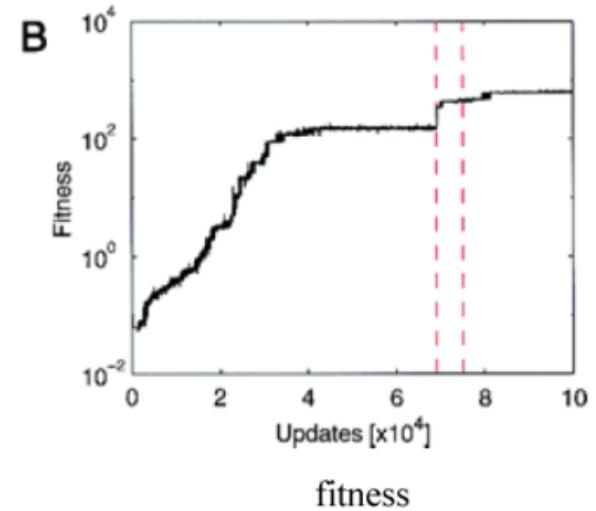
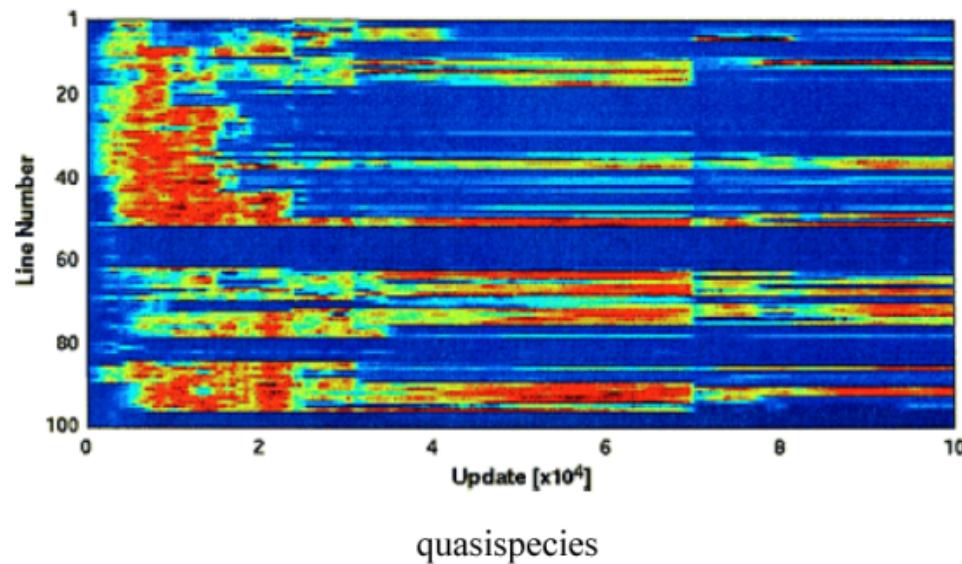
WT and evolved mutant (3 muts)

Directional selection causes decanalization in a group I ribozyme. Hayden EJ, Weikert C, Wagner A. PLoS One. 2012;

*fair comparison?*

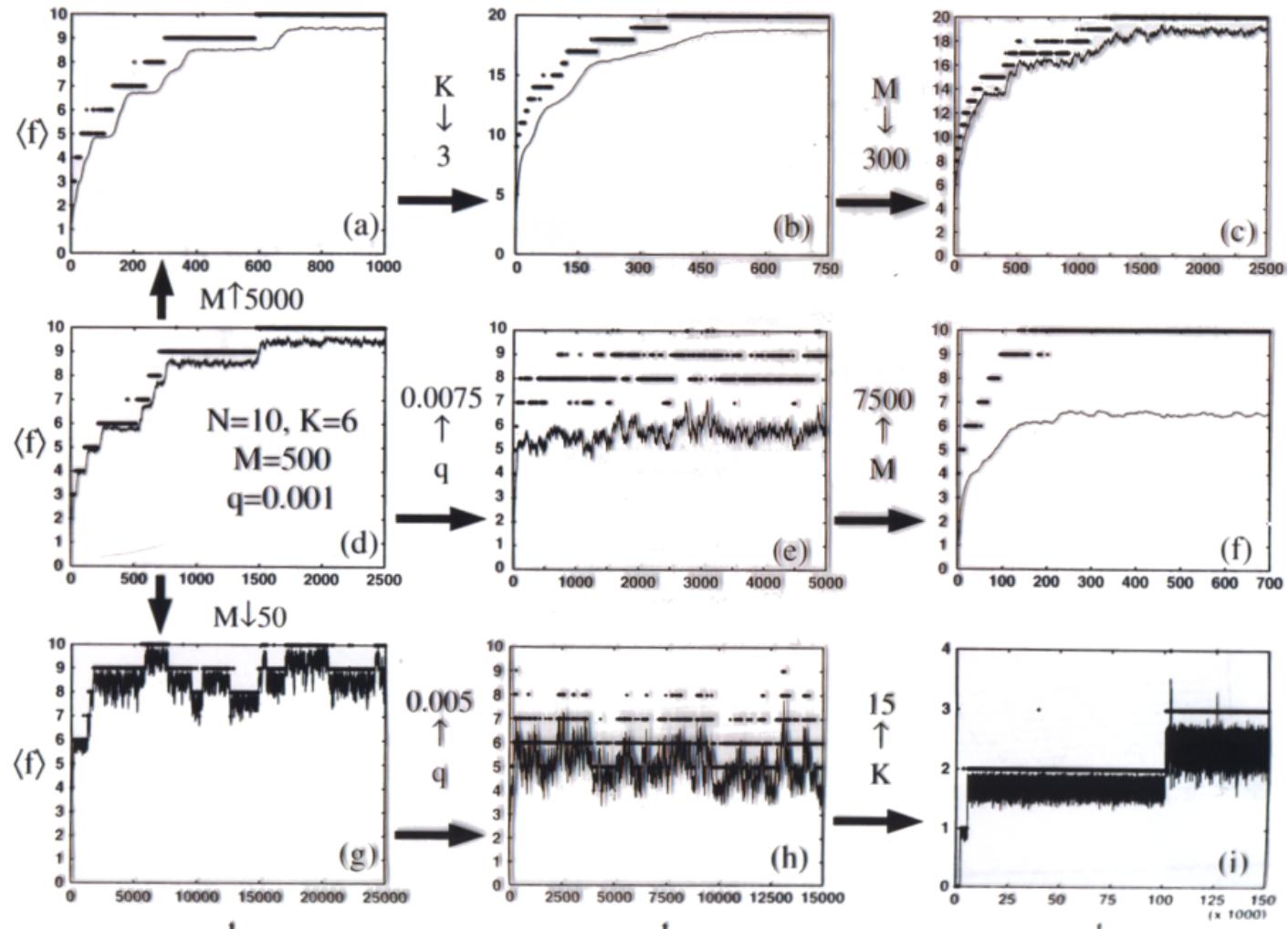
## Robustness, population diversity and evolutionary optimization

AVIDA: Self-replicating computer program (Adami et al)



Population variability per position (gene)  $p_i \log(p_i)$

## Neutrality and information accumulation (royal road)



information accumulation upto information threshold..

# Genotype-phenotype mapping: Coding structure

---

3 questions/answers:

**Given code –> which evolutionary dynamics?**

eg RNA folding: punctuated evolution etc.

**Given problem –> how to code?**

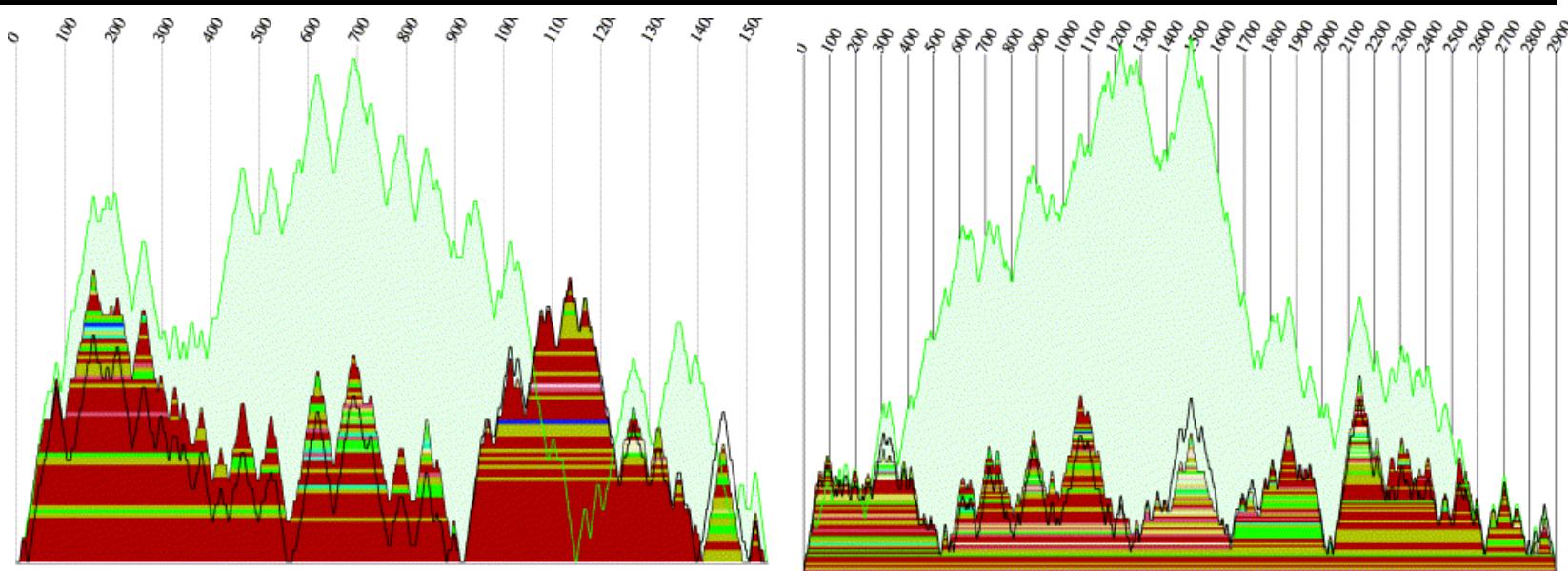
expectation: smooth, non-redundant;  
found intertwining neutral paths

**Given evolutionary dynamics –> which code?**

towards robustness, hence evolvability

*reconciliation adaptive and neutral evolution  
cf information threshold/ survival of the flattest*

# RNA secondary structure as paradigm for genotype-phenotype mapping computable??



16S RNA                    23S RNA  
Min. Energy folding vs Conserved folding

*A. globiformis, Anabaena sp., A. tumefaciens, B. japonicum, E. coli, B. subtilis, T. thermoph, Pir. marina, Rb. sphaero*

## alternative basins of attraction: replicator in nonMFE state

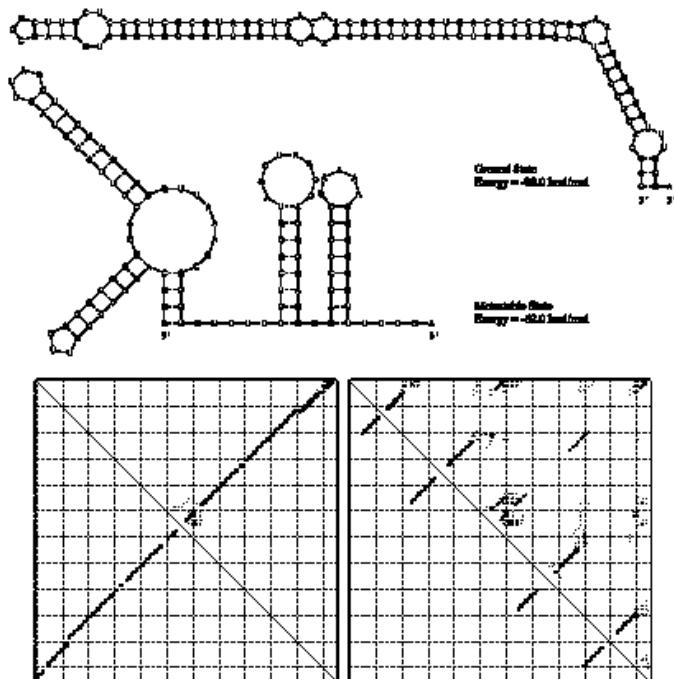


FIGURE 12. Structures and base pairing density plots for the mfe structure and the metastable conformation of the Q $\beta$  variant SV11. The secondary structures and their free energies are shown in the upper part. In the lower half we show the matrix of base pair probabilities as obtained from the thermodynamic partition function (McCaskill, 1990; Hofacker et al., 1994) (left) and from kinetic trajectories (right).

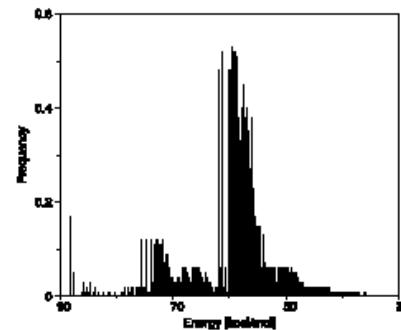


FIGURE 13. Fraction of folding paths visiting local minima in the Q $\beta$  variant SV11. The majority of paths visits the local minima in the basin of the metastable structure where the paths get trapped. Only about 16% reach the ground state.

**Derived properties  
JUST RNA?  
or even just by wrongly computed (and 2 D) folding?**

---

*percolating neutral path; innovations  
evolution toward robustness*

NO.....

similar (mutatis mutandis) properties in

Gene regulatory networks ( A Wagner 2007a,b)

Protein folding A Wagner 2010

Metabolic networks (A.Wagner 2012

see also books by A. Wagner

# From paradigm systems to general conclusions vs Studying “all” cases

---

NK landscapes (Kauffman):

Class of models to study impact of GP mapping on evolutionary dynamics.

N: number of properties (e.g. sequence length)

K: number of ‘epistatic’ interactions

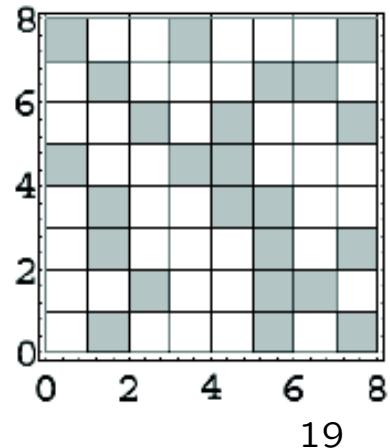
most often 2 states per position

Fitness contribution of each  $N \cdot 2^K$  states  
chosen randomly. Fitness is sum of those

Calculate e.g. pathlength to local peak  
height of optima reached (etc.)

NO percolating, intercalating neutral paths  
**and its evolutionary consequences**

versions include neutrality BUT



## **Neutral Paths, Causal Drift, Robust Signaling, and Complex Disease Andreas Wagner 2015 PLONE**

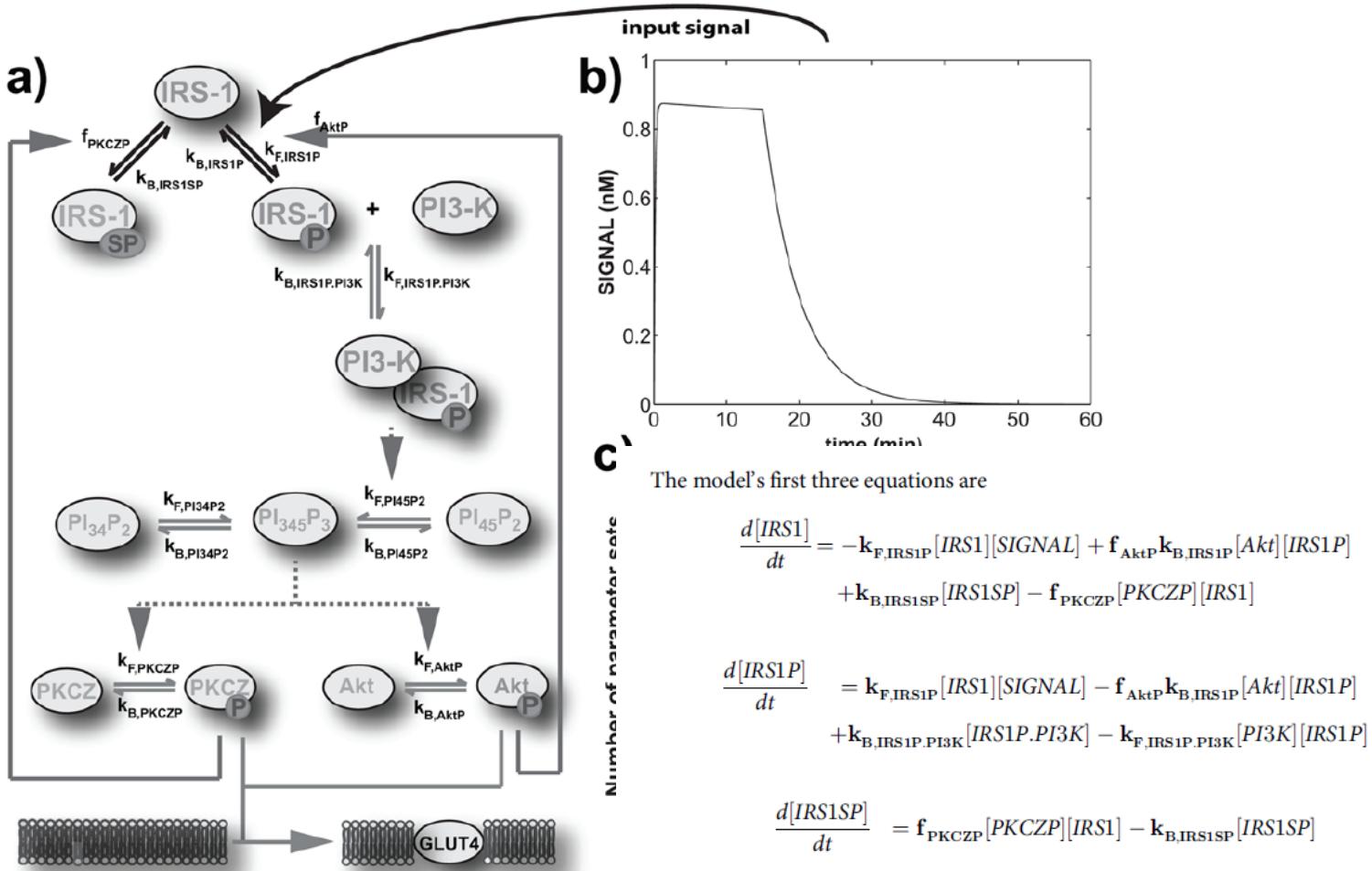
---

Explicit model of Insulin signaling pathway

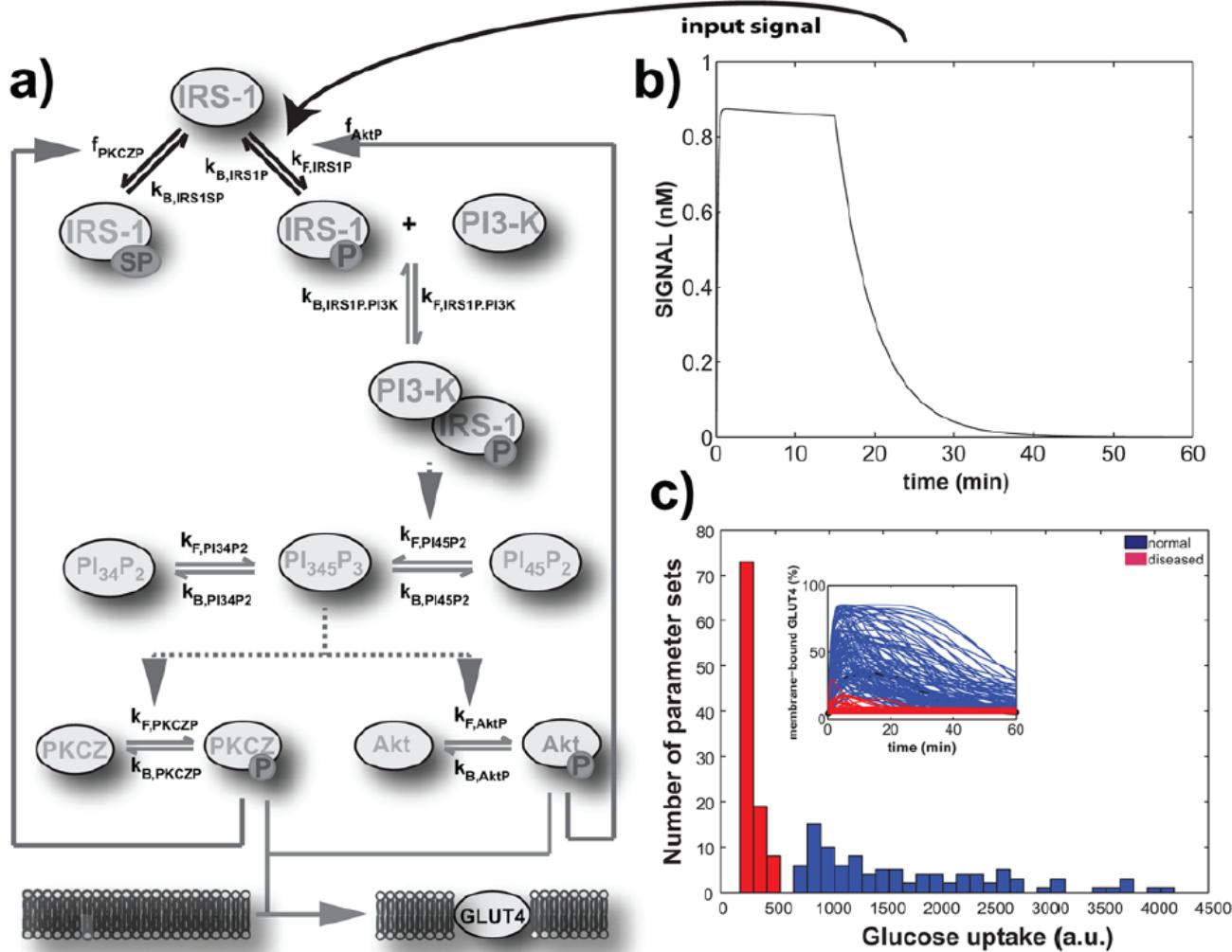
Random sampling of 15 kinetic parameters  $10^{-3} – 10^3$   
and evolving populations by mutating these parameters

Classifying behavior as “normal” or ”diseased”  
(based on glucose uptake-curve in time)

Determine sensitivity of parameters in different populations  
and during evolution. (log sampling of parameters)



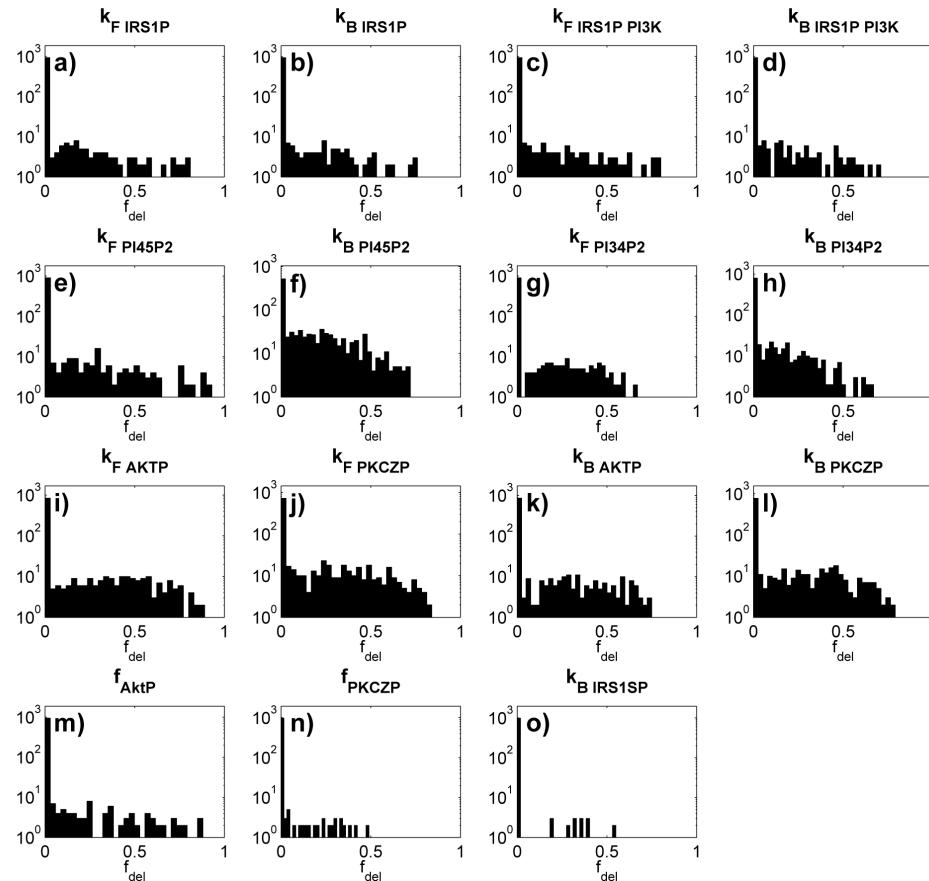
**Fig 1. Insulin signaling model, input and output. a)** Molecular interactions in the signaling pathway modeled here. Briefly, extracellular insulin leads to phosphorylation of the insulin receptor, which promotes the phosphorylation of IRS1 to yield IRS1P. The latter molecule associates with PI3K in a complex that triggers production of the second messenger PI<sub>345</sub>P<sub>3</sub>, which activates the protein kinases Akt and PKCZ. These kinases then promote the translocator of the glucose transporter GLUT4 to the membrane, where it helps import glucose into the cell. Mass-action parameters that determine the rates of the respective reactions are indicated by a 'k' followed by a subscript. Activated PKCZ and Akt exert feedback on the production of two different phosphorylated forms of IRS1 (IRS1SP and IRS1P). The strength of this feedback is encapsulated by parameters  $f_{PKCZP}$  and  $f_{AktP}$ , respectively. See [Methods](#) for details. **b)**



**Fig 1. Insulin signaling model, input and output. a)** Molecular interactions in the signaling pathway modeled here. Briefly, extracellular insulin leads to phosphorylation of the insulin receptor, which promotes the phosphorylation of IRS1 to yield IRS1P. The latter molecule associates with PI3K in a complex that triggers production of the second messenger  $\text{PI}_{345}P_3$ , which activates the protein kinases Akt and PKCZ. These kinases then promote the translocator of the glucose transporter GLUT4 to the membrane, where it helps import glucose into the cell. Mass-action parameters that determine the rates of the respective reactions are indicated by a 'k' followed by a subscript. Activated PKCZ and Akt exert feedback on the production of two different phosphorylated forms of IRS1 (IRS1SP and IRS1P). The strength of this feedback is encapsulated by parameters  $f_{PKCZP}$  and  $f_{AktP}$ , respectively. See [Methods](#) for details. **b)**

very high neutrality of individual parameter changes  
but very different in different parameter sets.

---



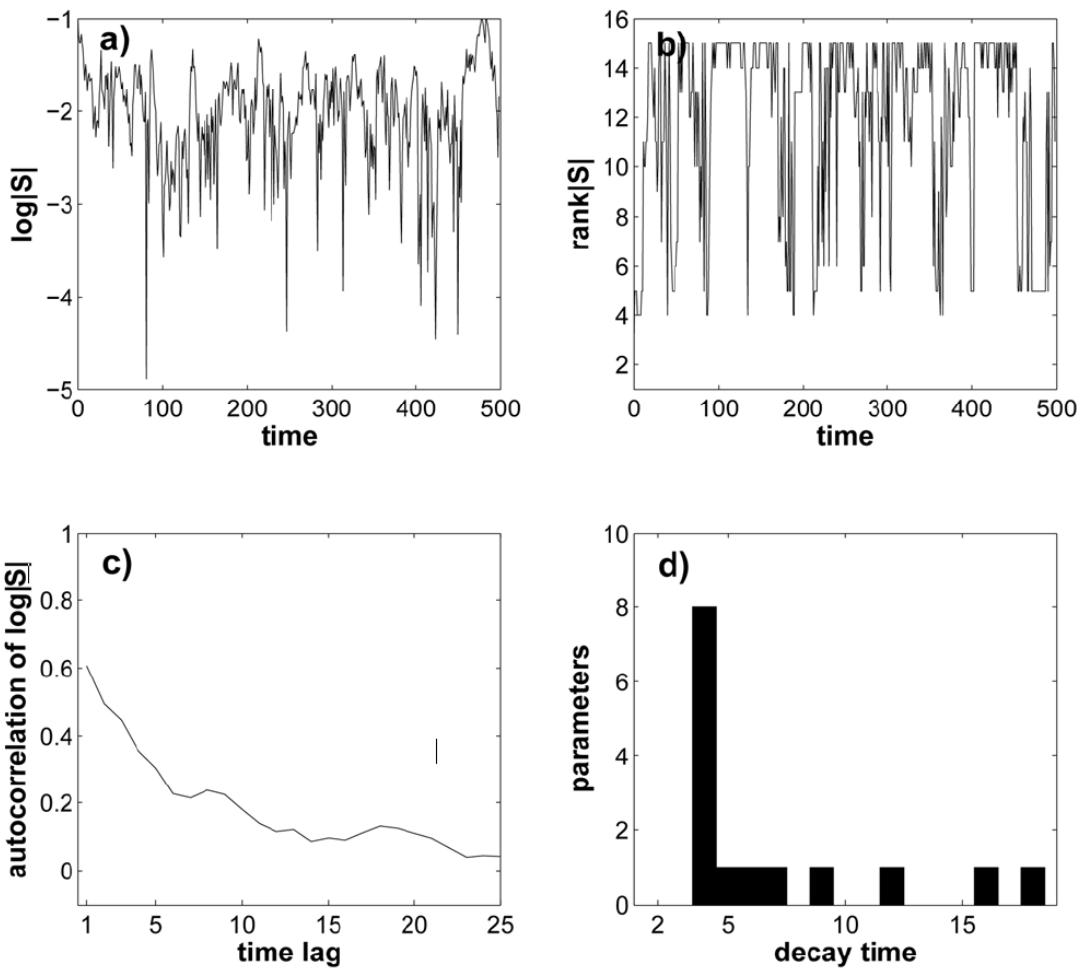
## Rapid “Causal drift”

rapid change of  
sensitivity to  
parameter changes  
(mutations)  
due to neutral drift

*“genetic background  
“cause of disease”*

cf GWAS studies  
50% “explained”

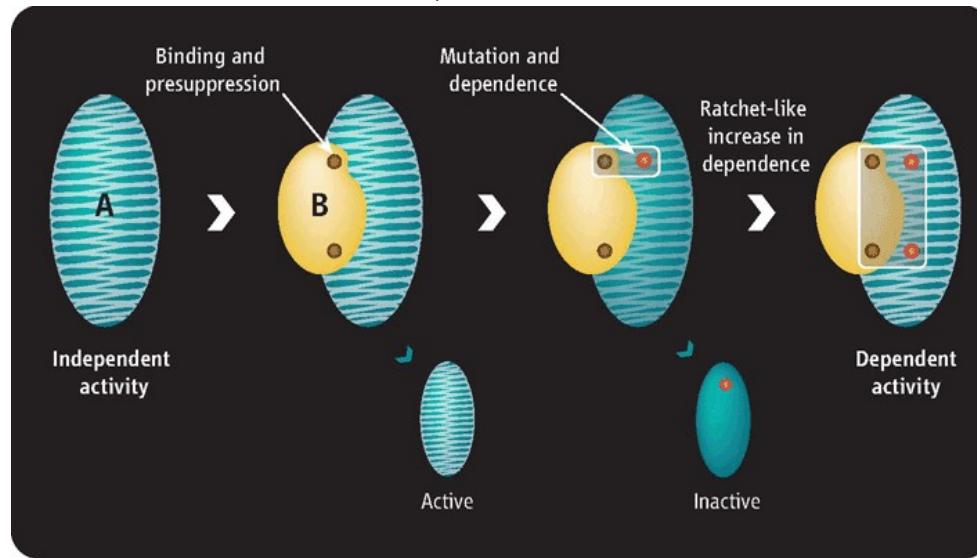
Mouse models



# Neutrality and evolution of complexity neutral ratchet/constructive neutral evolution/irremediable complexity

---

e.g. neutral binding / increase neutrality /  
accumulation of mutations / indispensability of binding



cf Covello & Gray 1993, Stoltzfus 1999, Lynch 2007, Gray et al 2010  
23

# Evolution of coding structure cont.

## Evolution of multiple coding in RNA's

*doing more with less*

---

Evolve towards target == set of (25) RNA structures .

ALL other structures (Shapiro) TOXIC

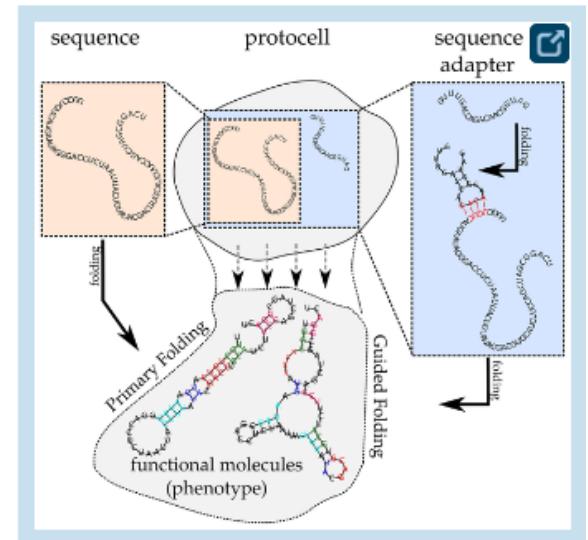
define possible interaction of RNA's:

adaptors (=single hairpin)

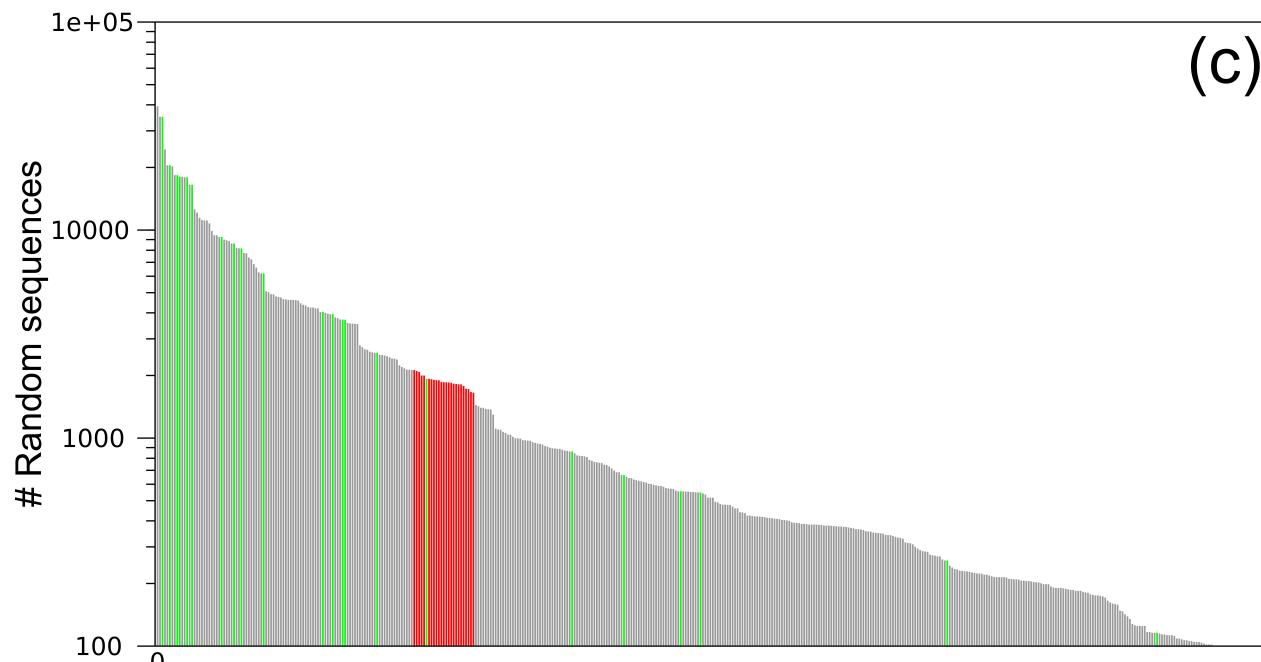
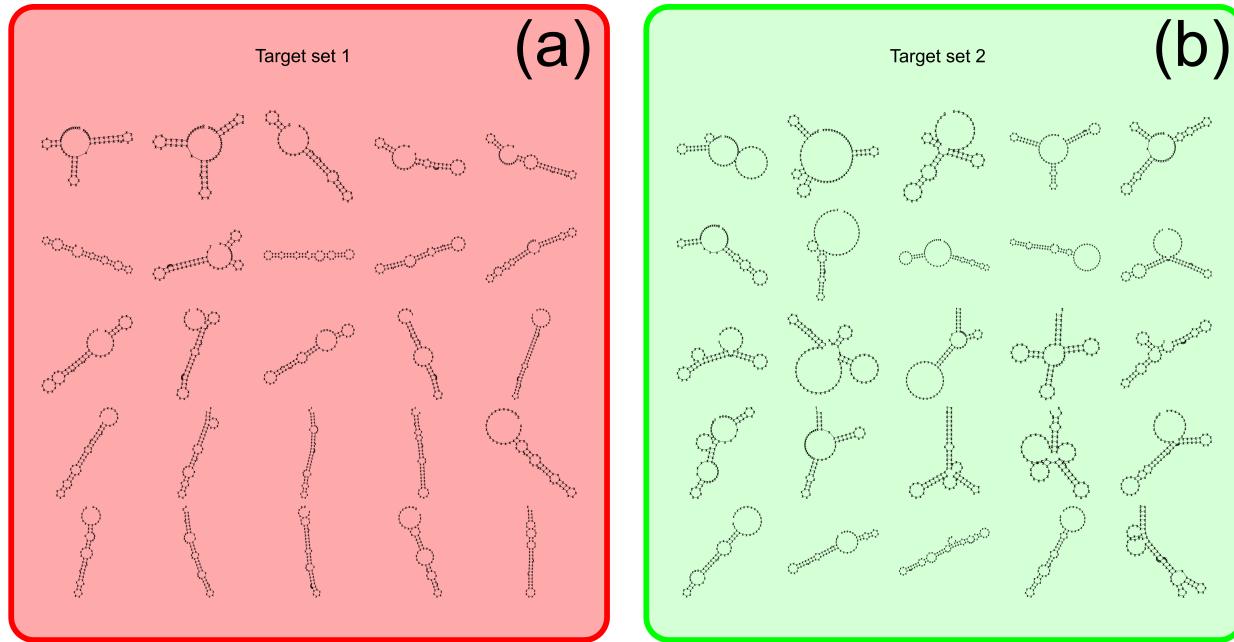
can bind to other RNA

bound (modified) nucl not 'available'  
for folding

fitness of cell: set of struct.  
cells compete in space

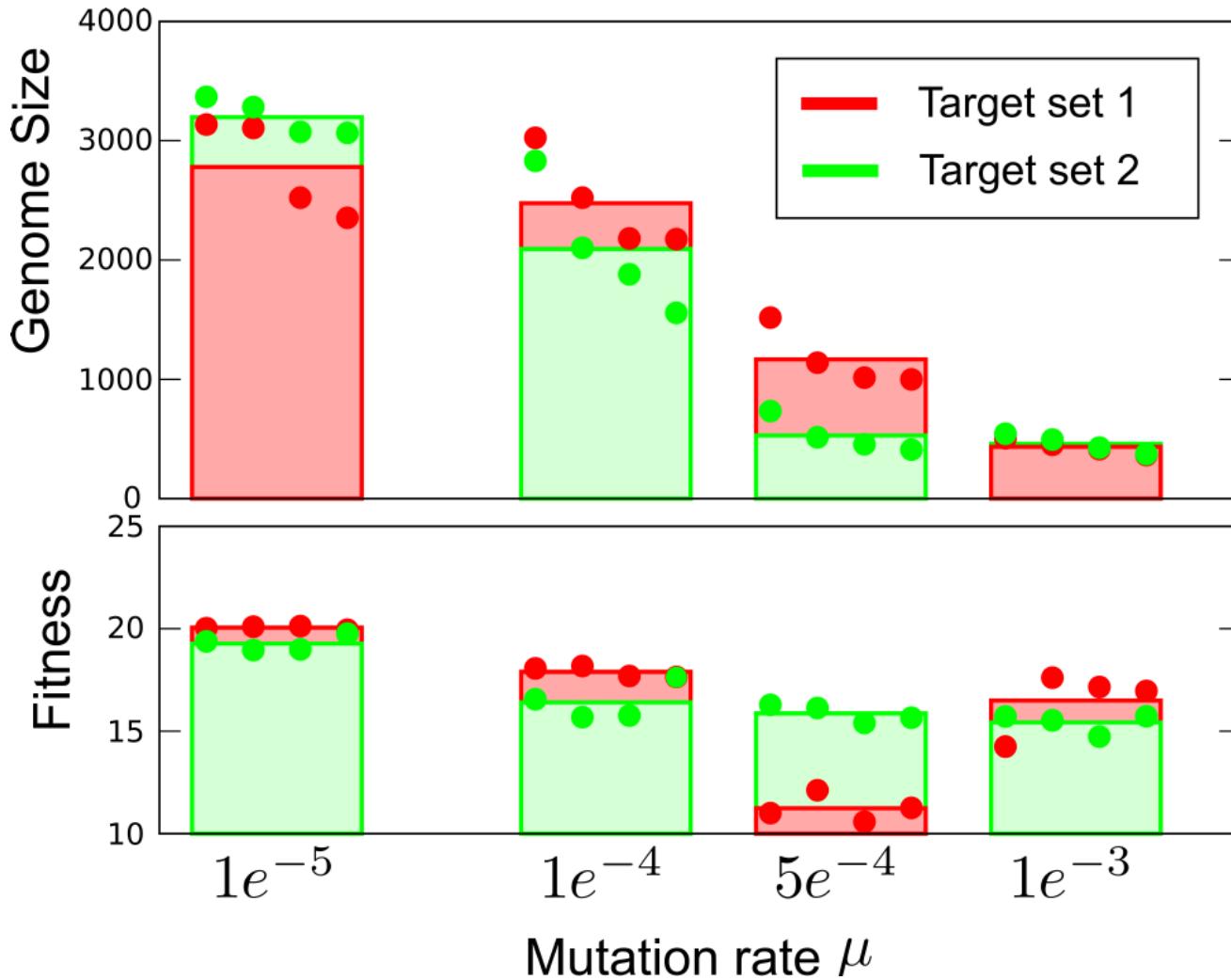


*How to cope with high mutation rates?*

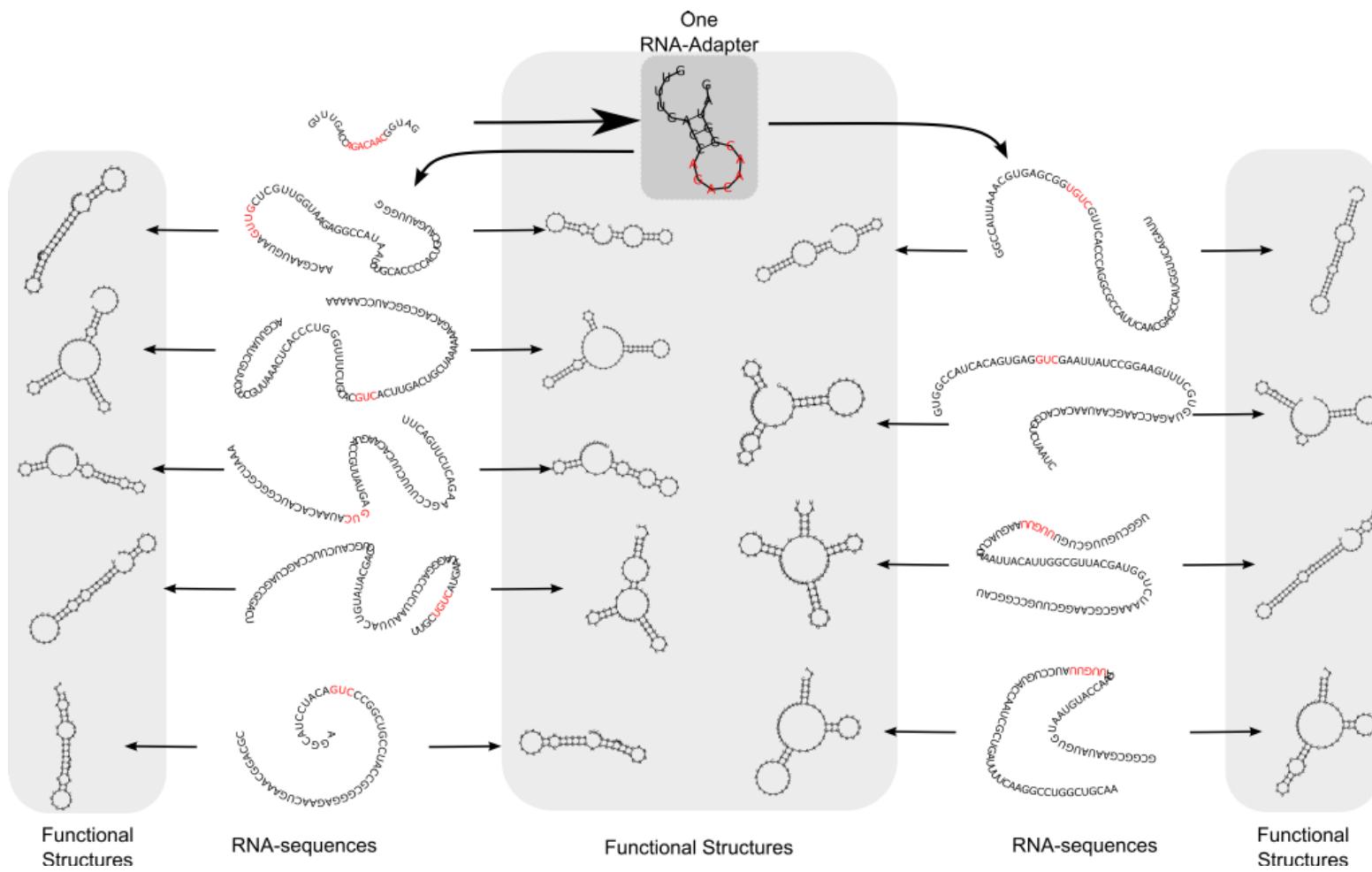


# high mutation rate - short genome - same functionality

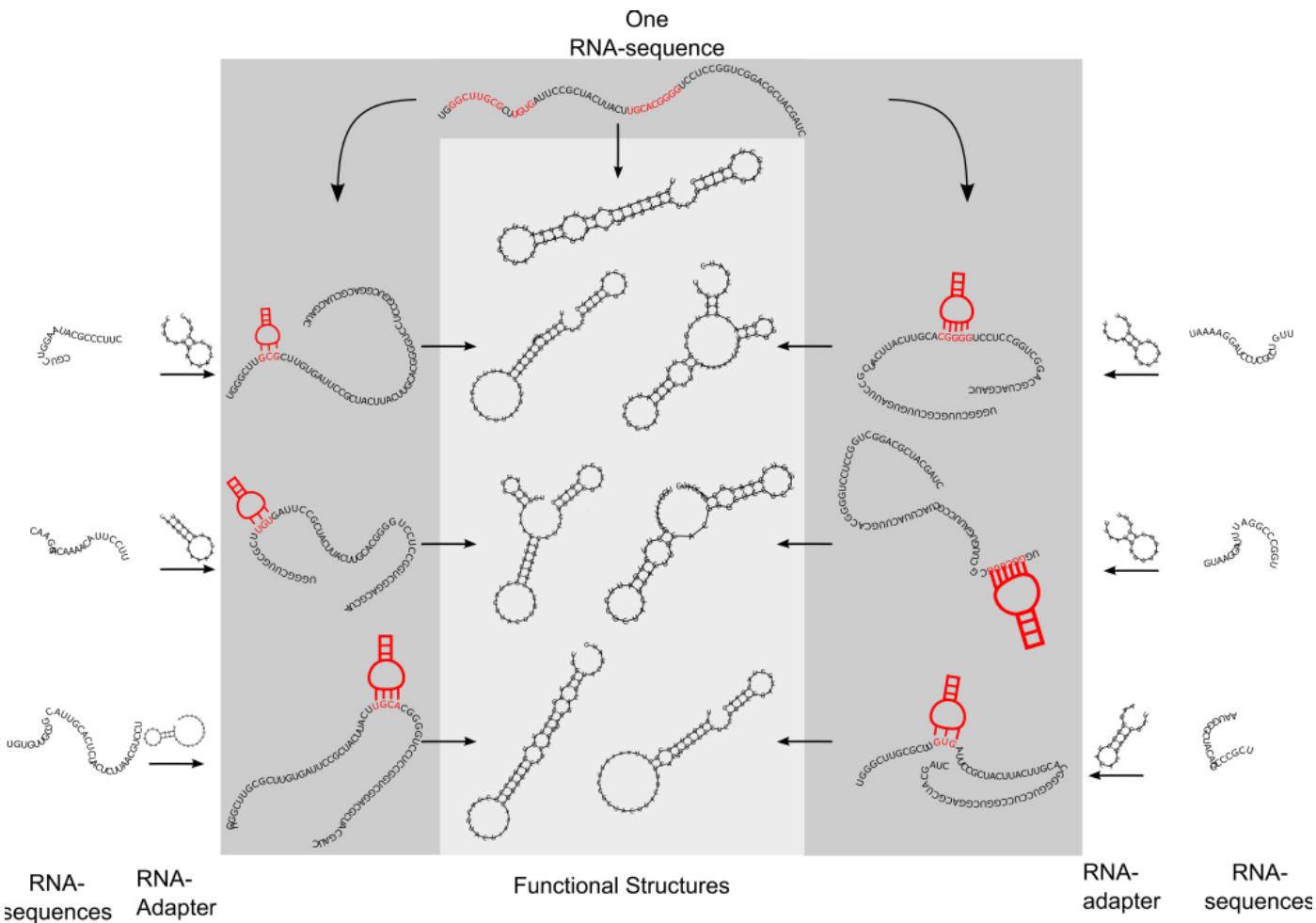
---



**one adaptor used by all sequences**



# many adaptors used by 1 sequence



# Conclusion: multiple coding

---

RNA even more an “ideal evolvable molecule”

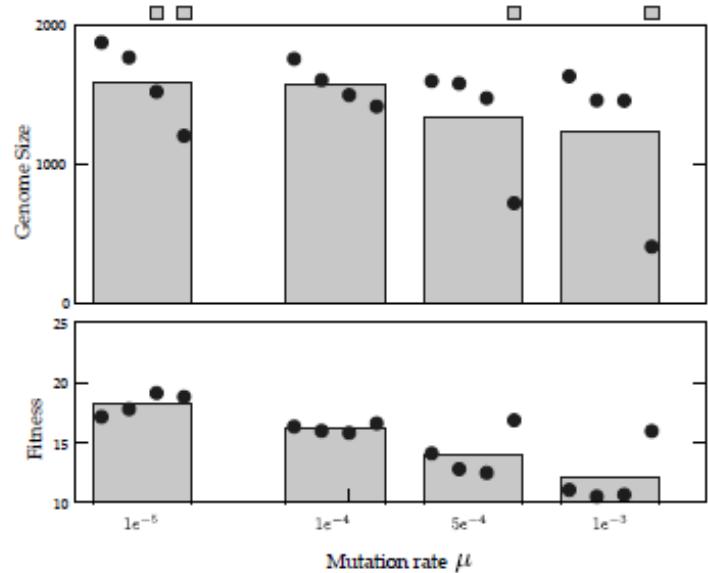
information threshold shapes coding structure  
multiple coding arises and alleviates information threshold

information threshold does not (necessarily) limit functionality

(Similar effects seen with alternative (non-minimal energy) foldings)

Also in this case:

local competition in space helps!



well mixed:

## Conclusions

---

Coding structure adapts to mutation rate  
Coding length, selection strength

Result:

Evolution converges to being  
Close to Information Threshold