Optimality

Constraints

Predictability

http://www.imagequest3d.com/pages2/shop/Posters/images/biodiversity.htm

Today …

I. Natural selection

II. Fitness landscapes

III. A network perspective on fitness landscapes

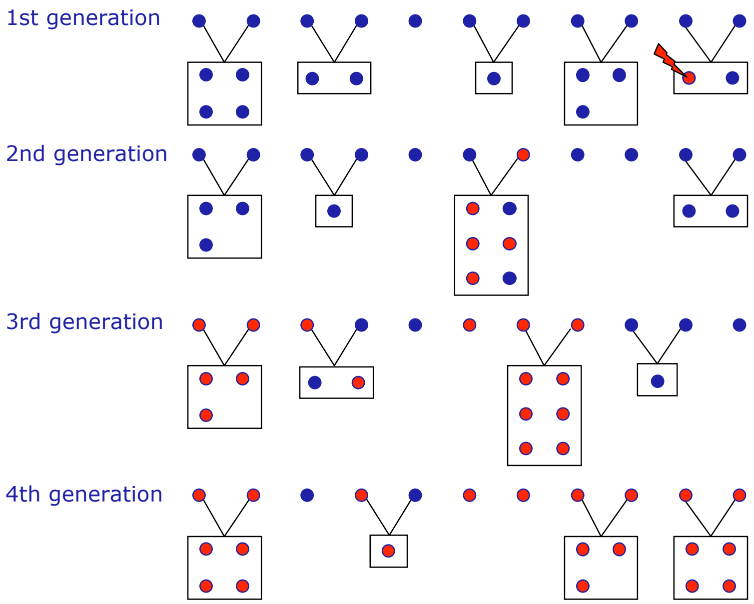
IV. My favorite molecule: RNA

V. Are biological systems optimal?

- Constrained by mutation
- Evolutionary traps

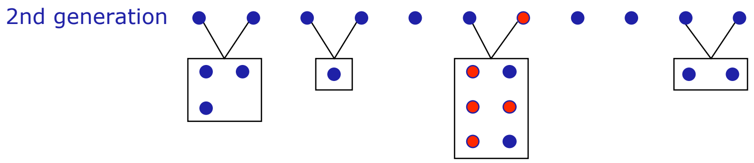
VI. Conclusions

Evolution by Natural Selection



Evolution by Natural Selection

Three requirements?

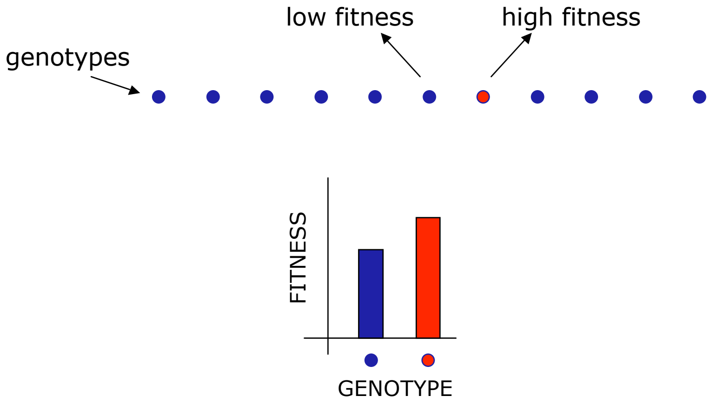


HERITABILITY

VARIATION

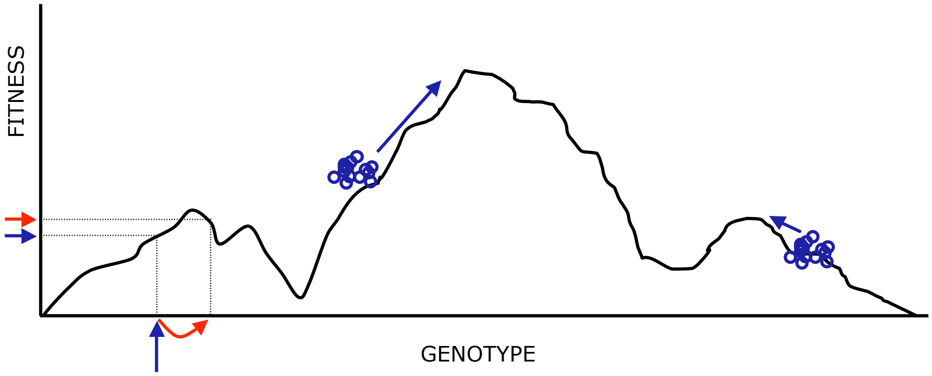
DIFFERENTIAL REPRODUCTIVE SUCCESS

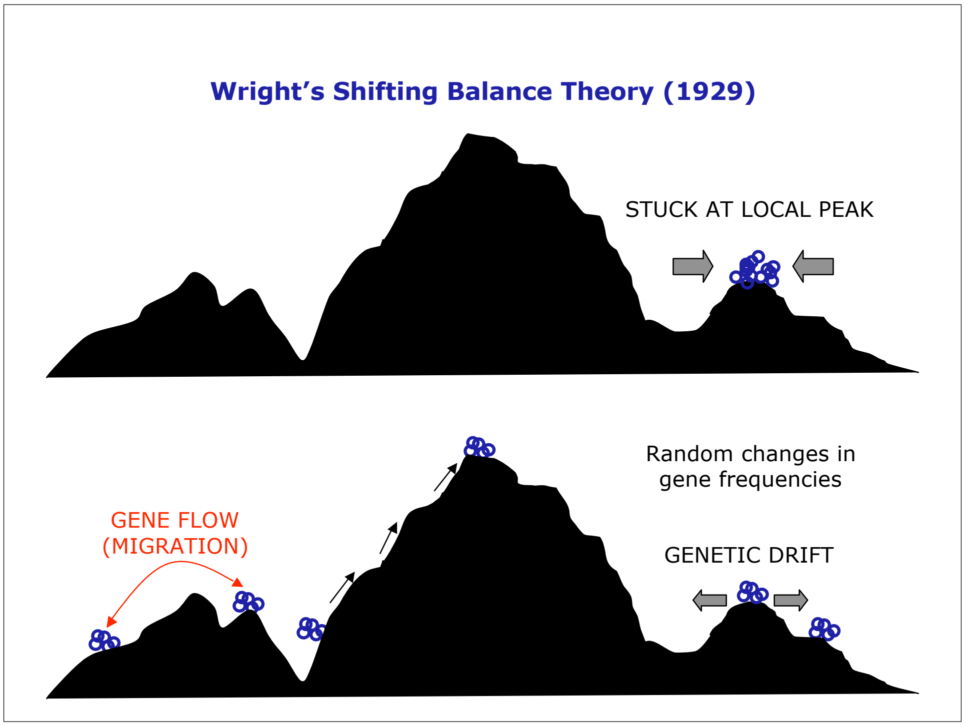
Fitness Landscape



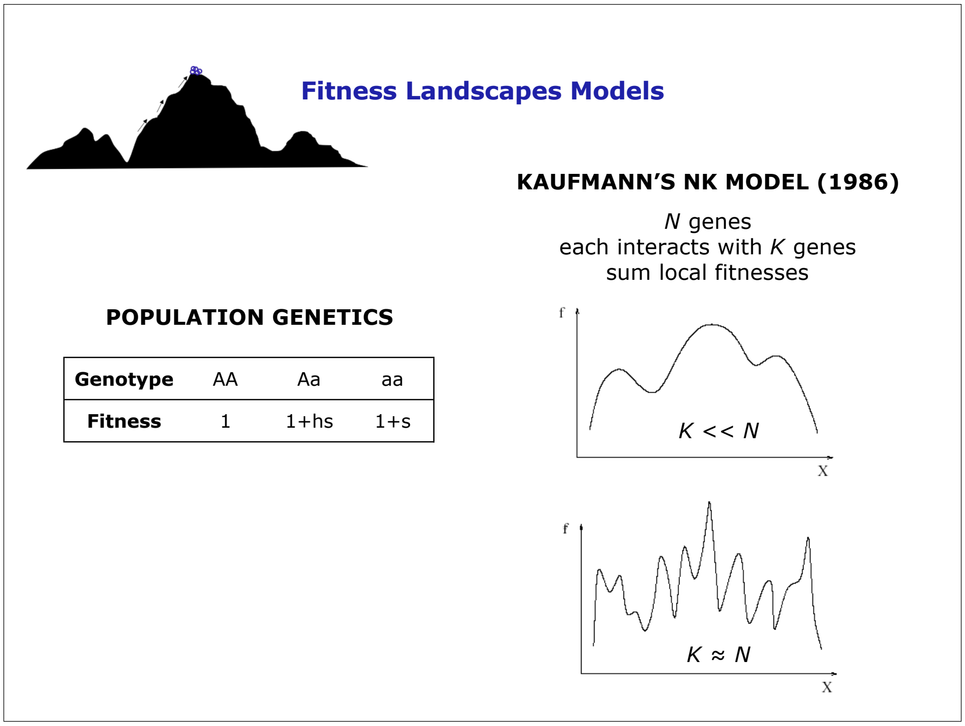
Map from the space of all genotypes to fitness ….

Fitness Landscape





Wright’s Shifting Balance Theory (1929)



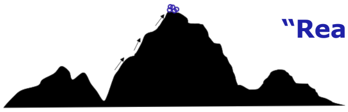
Fitness Landscapes Models

KAUFMANN’S NK MODEL (1986)

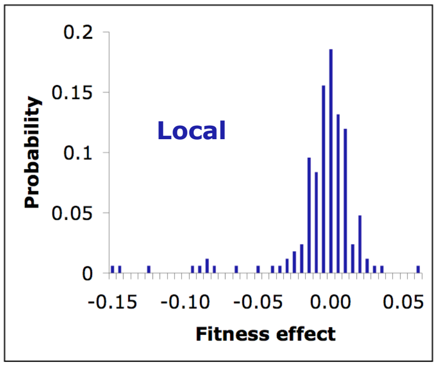
N genes
each interacts with *K* genes
sum local fitnesses

POPULATION GENETICS

Genotype	AA	Aa	aa
Fitness	1	1+hs	1+s

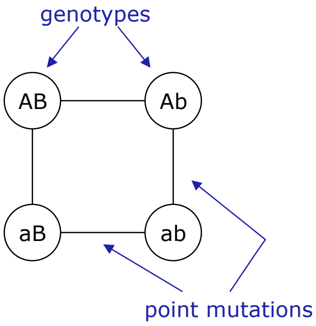


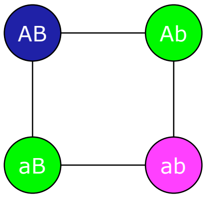
“Real” Fitness Landscapes



Mutational Networks:

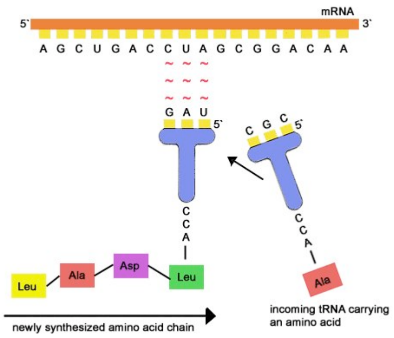
Another way to represent fitness landscapes





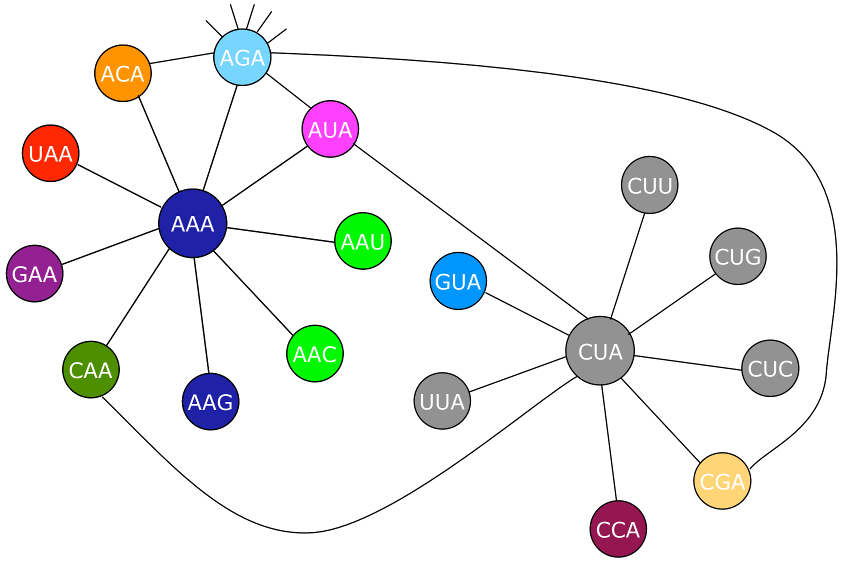
two-locus, two-allele haploid

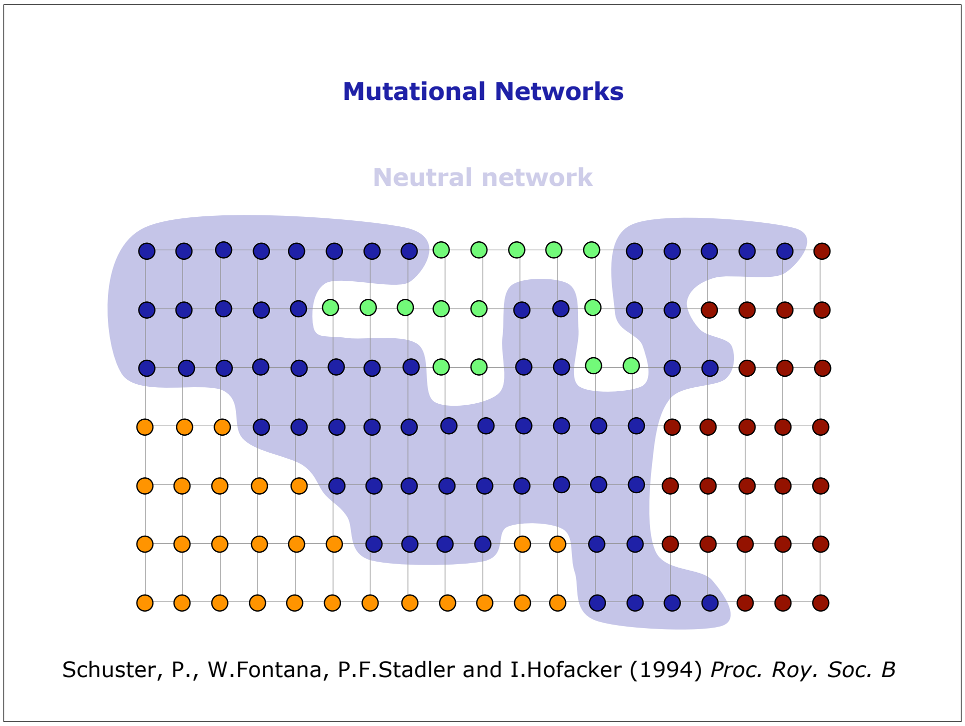
Making proteins from DNA



	T	C	A	G
T	TTT Phe TTC Phe TTA Leu TTG Leu	TCT Ser TCC Ser TCA Ser TCG Ser	TAT Tyr TAC Tyr TAA Och TAG Amb	TGT Cys TGC Cys TGA Opa TGG Trp
C	CTT Leu CTC Leu CTA Leu CTG Leu	CCT Pro CCC Pro CCA Pro CCG Pro	CAT His CAC His CAA Gln CAG Gln	CGT Arg CGC Arg CGA Arg CGG Arg
A	ATT Ile ATC Ile ATA Ile ATG Met	ACT Thr ACC Thr ACA Thr ACG Thr	AAT Asn AAC Asn AAA Lys AAG Lys	AGT Ser AGC Ser AGA Arg AGG Arg
G	GTT Val GTC Val GTA Val GTG Val	GCT Ala GCC Ala GCA Ala GCG Ala	GAT Asp GAC Asp GAA Glu GAG Glu	GGT Gly GGC Gly GGA Gly GGG Gly

Mutational Network

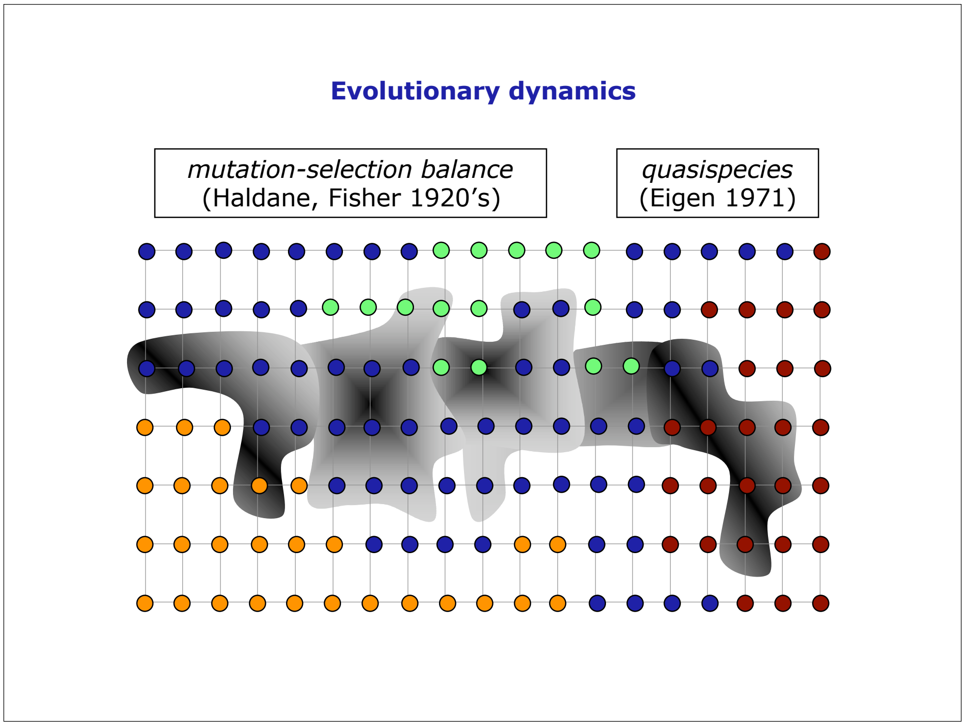




Mutational Networks

Neutral network

Schuster, P., W.Fontana, P.F.Stadler and I.Hofacker (1994) *Proc. Roy. Soc. B*



Evolutionary dynamics

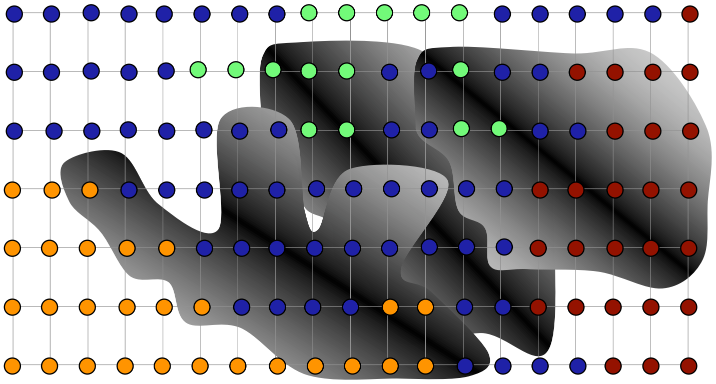
mutation-selection balance
(Haldane, Fisher 1920's)

quasispecies
(Eigen 1971)

Evolutionary dynamics

mutation-selection balance
(Haldane, Fisher 1920's)

quasispecies
(Eigen 1971)

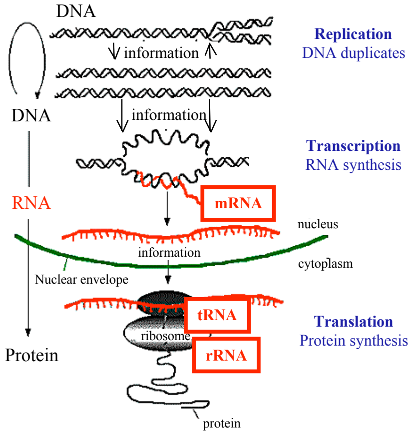


Mutation >> selection

error catastrophe

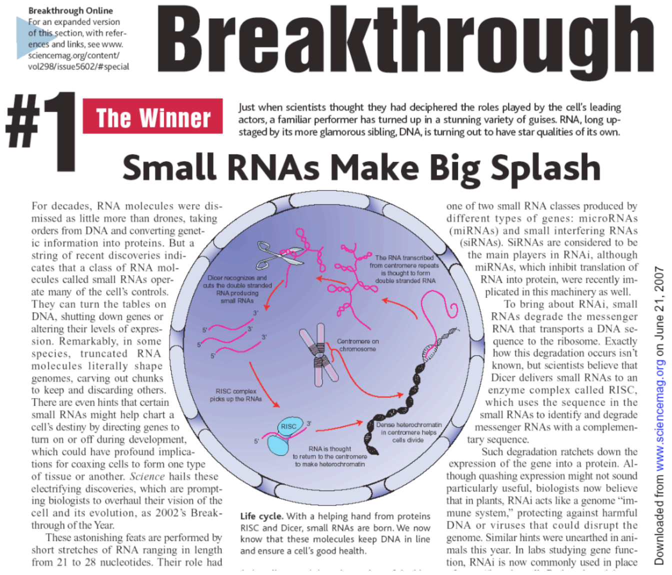
RNA: An underappreciated molecule

Essential machinery of life



The Central Dogma of Molecular Biology

RNA



Breakthrough Online
For an expanded version of this section, with references and links, see www.sciencemag.org/content/vol298/issue5602/#special

Breakthrough

#1 The Winner

Just when scientists thought they had deciphered the roles played by the cell's leading actors, a familiar performer has turned up in a stunning variety of guises. RNA, long upstaged by its more glamorous sibling, DNA, is turning out to have star qualities of its own.

Small RNAs Make Big Splash

For decades, RNA molecules were dismissed as little more than drones, taking orders from DNA and converting genetic information into proteins. But a string of recent discoveries indicates that a class of RNA molecules called small RNAs operate many of the cell's controls. They can turn the tables on DNA, shutting down genes or altering their levels of expression. Remarkably, in some species, truncated RNA molecules literally shape genomes, carving out chunks to keep and discarding others. There are even hints that certain small RNAs might help chart a cell's destiny by directing genes to turn on or off during development, which could have profound implications for coaxing cells to form one type of tissue or another. *Science* hails these electrifying discoveries, which are prompting biologists to overhaul their vision of the cell and its evolution, as 2002's Breakthrough of the Year.

These astonishing feats are performed by short stretches of RNA ranging in length from 21 to 28 nucleotides. Their role had

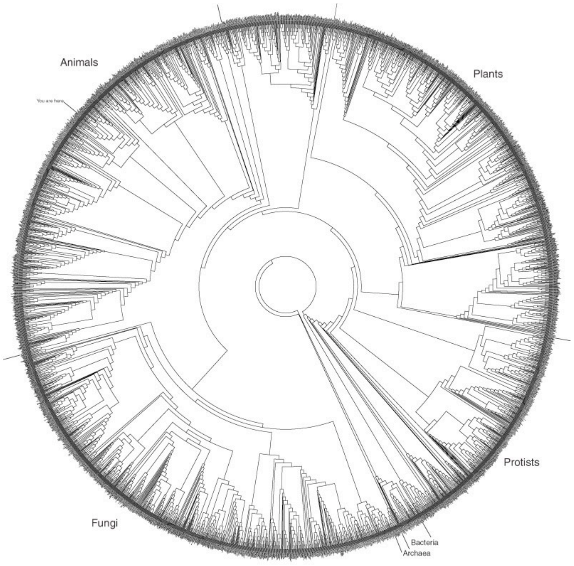
Life cycle. With a helping hand from proteins RISC and Dicer, small RNAs are born. We now know that these molecules keep DNA in line and ensure a cell's good health.

one of two small RNA classes produced by different types of genes: microRNAs (miRNAs) and small interfering RNAs (siRNAs). SiRNAs are considered to be the main players in RNAi, although miRNAs, which inhibit translation of RNA into protein, were recently implicated in this machinery as well.

To bring about RNAi, small RNAs degrade the messenger RNA that transports a DNA sequence to the ribosome. Exactly how this degradation occurs isn't known, but scientists believe that Dicer delivers small RNAs to an enzyme complex called RISC, which uses the sequence in the small RNAs to identify and degrade messenger RNAs with a complementary sequence.

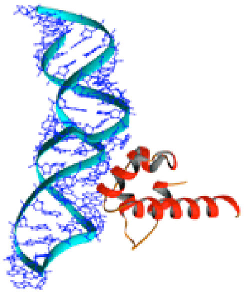
Such degradation ratchets down the expression of the gene into a protein. Although quashing expression might not sound particularly useful, biologists now believe that in plants, RNAi acts like a genome "immune system," protecting against harmful DNA or viruses that could disrupt the genome. Similar hints were unearthed in animals this year. In labs studying gene function, RNAi is now commonly used in place

RNA



Science **2003, 300: 1692-1697**

RNA



www.monmouth.com/~spidersigns/RNA/

Essential machinery of life

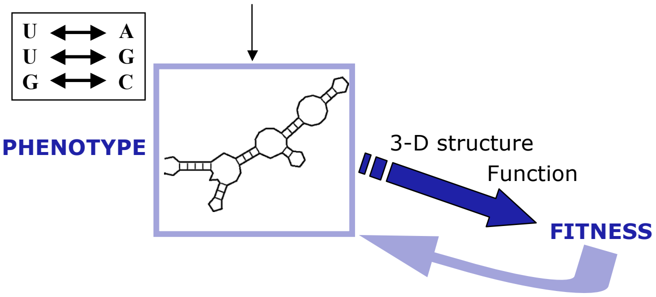
“Molecule of the year” (2002)

Phylogenetics

Bio-engineering

Tractable model of a complex biological system

Modeling the evolution of RNA structure



GENOTYPE AUCGGCGCUCCGUACUACGCUUAAAAAACAGGA … UACUUGCAAAA

U ↔ A
U ↔ G
G ↔ C

PHENOTYPE

3-D structure

Function

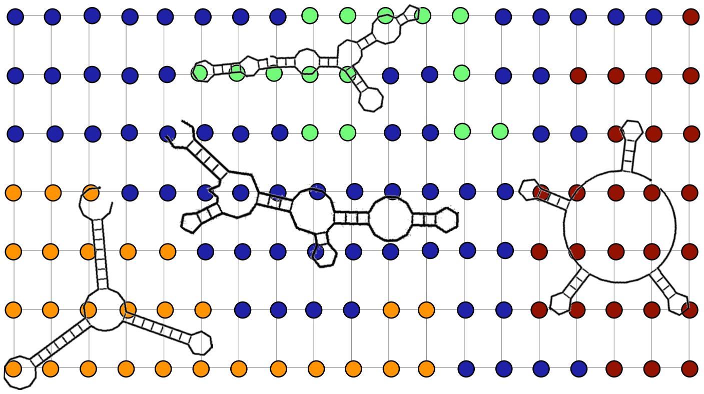
FITNESS

Mutational surveys: Fold millions of molecules and all of their mutants to systematically explore the mutational network.

In silico evolution: In the computer, select populations of RNA molecules according to structural criteria, and allow them to mutate.

RNA Mutational Networks

AUCGGCGCUCCGUACUACGCUUAAAAAACAGGA … UACUUGCAAAA



Are RNA molecules optimal?

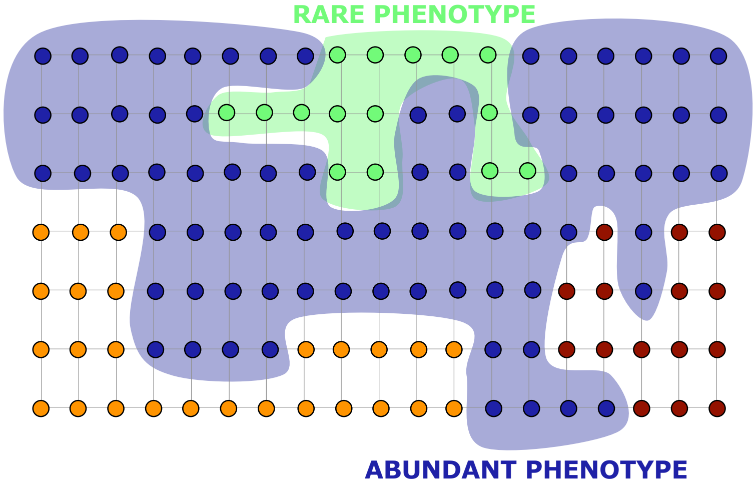
Motivation

“tyranny of the short”
super-optimal laboratory molecules
Is anything optimal?
If we re-ran the tape, where would we end up?

Characterizing entire mutational networks

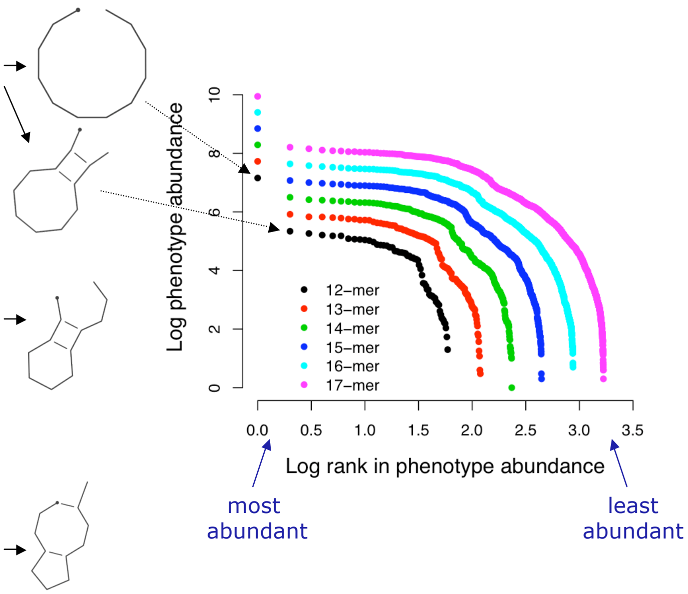
Molecule Length	Number of genotypes	Number of phenotypes
12	4^{12} = 16,777,216	59
13	4^{13} = 67,108,864	119
14	4^{14} = 268,435,456	234
15	4^{15} = 1,073,741,824	443
16	4^{16} = 4,294,967,296	872
17	4^{17} = 17,179,869,184	1673
18	4^{18} = 68,719,476,736	

Definition: Phenotype Abundance

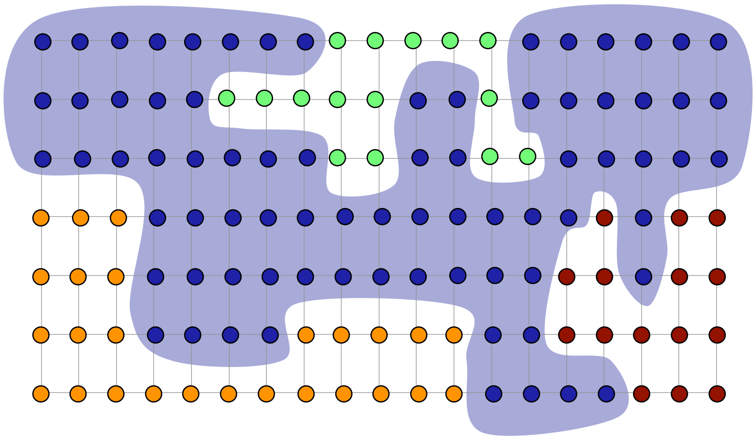


Mutational Landscape Structure

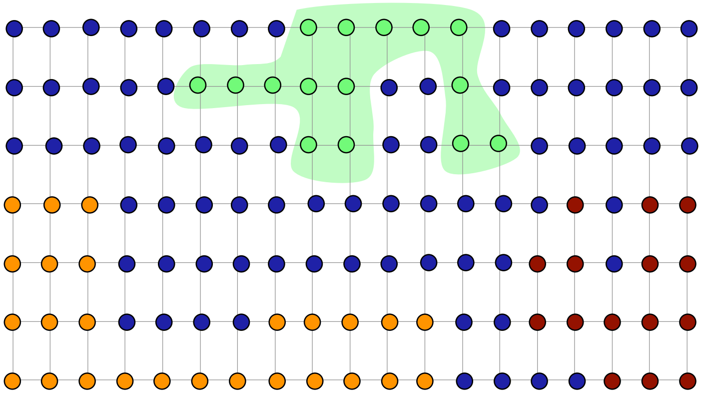
Rank	Abundance
1	14325304
2	218567
3	183791
4	162009
5	152393
6	151705
7	122332
8	117213
9	113896
10	110509
11	105538
12	93866
.	.
.	.
.	.
53	220
54	197
55	165
56	153
57	109
58	54
59	20



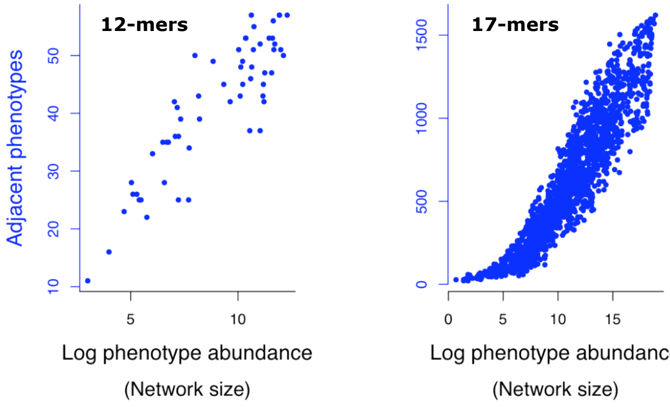
Adjacent Phenotypes



Adjacent Phenotypes



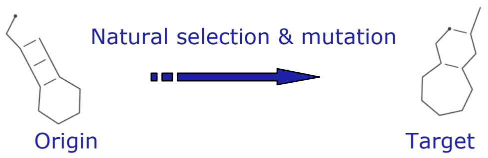
Abundant phenotypes can access more novelty



Does network structure actually impact evolution?

Evolutionary simulation:

- Populations of 1000 molecules
- Choose a target shape and select for similarity to the target
- Point mutation



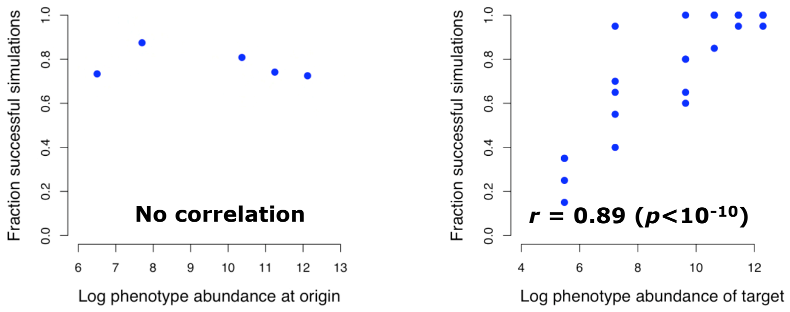
Question:

Is the evolutionary fate of the population influenced by

(a)The abundance of the **origin** phenotype?

(b)The abundance of the **target** phenotype?

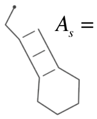
Yes, network structure constrains evolution



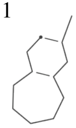
Hypothesis: Real RNA are not necessarily optimal. Naturally occurring RNA molecules may be biased towards phenotypes with high abundance, that is, towards shapes that are produced by many different sequences.

A shape statistic that correlates with phenotype abundance

$$A_s = \frac{\text{total length stem-loop regions + number base pairs}}{\text{number of contiguous stacks}}$$

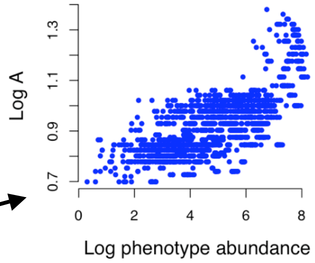


$$A_s = \frac{10+3}{1} = 13$$



$$A_s = \frac{11+2}{2} = 6.5$$

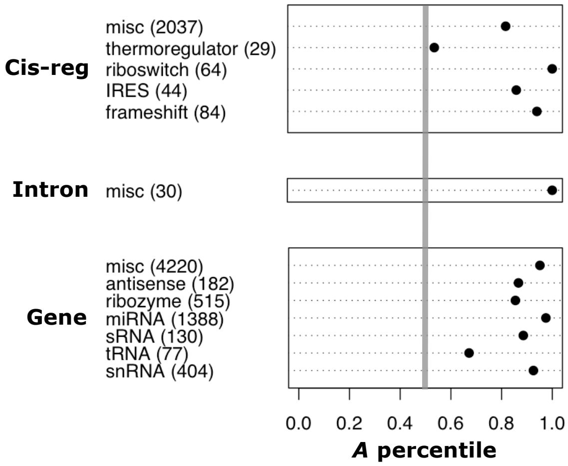
Size	*r*	*p*-value
12-mer	0.712	$<10^{-10}$
13-mer	0.721	$<10^{-16}$
14-mer	0.724	$<10^{-16}$
15-mer	0.703	$<10^{-16}$
16-mer	0.692	$<10^{-16}$
17-mer	0.692	$<10^{-16}$



Question: Do natural RNA have statistically high values of *A*?

If so, then maybe there is a bias toward abundant (rather than best) structures.

Yes, natural RNA are shaped by mutational networks



Rfam: Griffiths-Jones, S., A. Bateman, M. Marshall, A. Khanna, S.R. Eddy (2003) *Nucleic Acids Research*

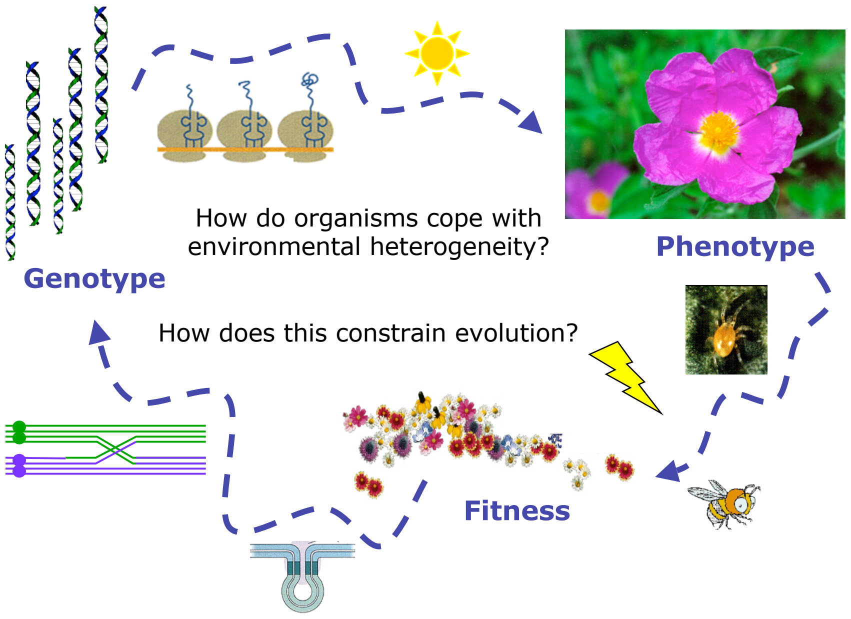
Evolution of the abundant …

Are RNA molecules optimal?

Not necessarily

For a function to evolve, it has to be *accessible*

But life, even for an RNA molecule, is more complex …



Genotype

Phenotype

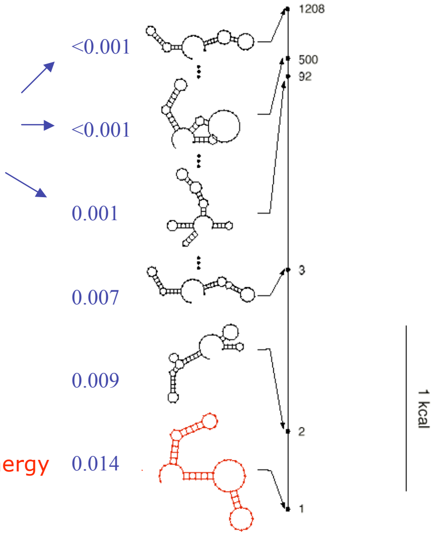
Fitness

How do organisms cope with environmental heterogeneity?

How does this constrain evolution?

Plasticity in RNA Secondary Structure

Proportion of
time spent in each
structure

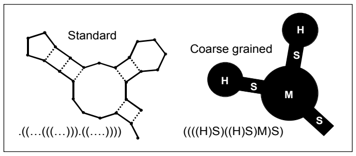


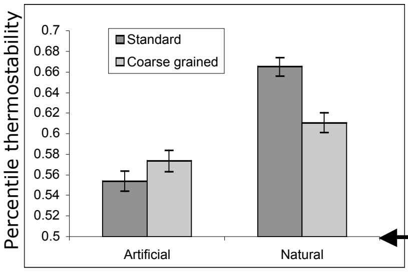
Minimum free energy
structure

GCUGUUAUCGGCGCUCCGUACUACGCUUAAAAAACAGGACAGUUGGGAUACUUGCAAAACCAGGUUCAUCUUGUGA

RNA have evolved thermostability

Levels of structural resolution

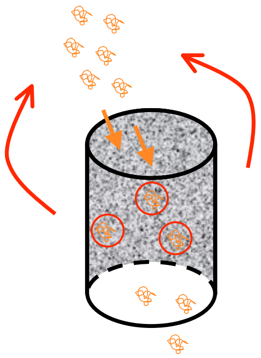




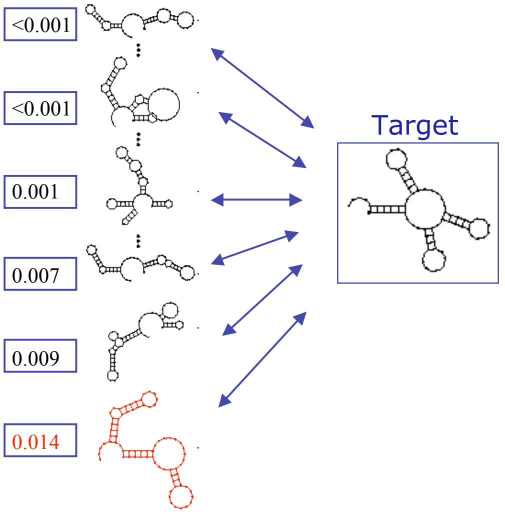
Average thermostability for
length and base composition

Lets model this process …

Evolutionary model (loosely) based on artificial selection of RNA



Ensemble fitness



Selecting for
function (shape)
and
thermostability

GCUGUUAUCGGCGCUCCGUACUACGCUUAAAAAACAGGACAGUUGGGAUACUUGCAAAACCAGGUUCAUCUUGUGA

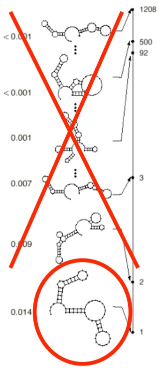
Evolutionary Simulations

Population of ~1000 sequences in a chemical flow reactor

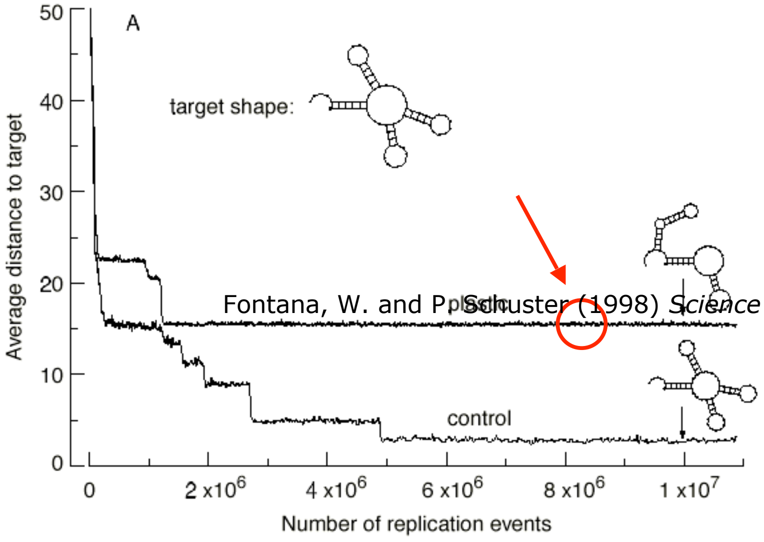
Replicate sequences in proportion to their fitnesses*

Constant mutation rate (0.001 per position)

*Plastic fitness
vs.
Control fitness



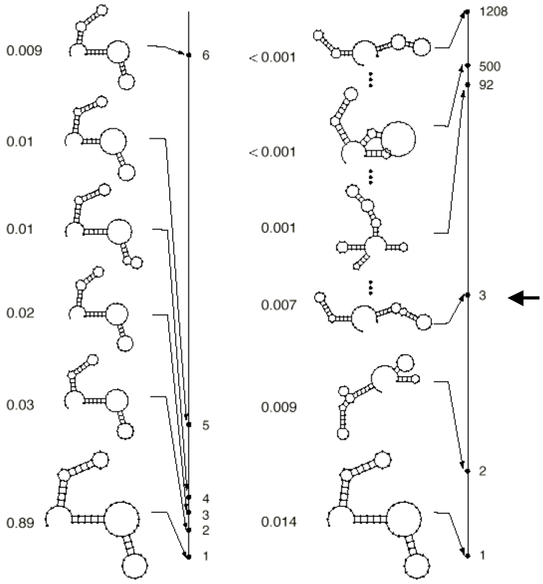
Evolutionary Trajectories



Fontana, W. and P. Schuster (1998) *Science*

Evolution of Thermodynamic Stability

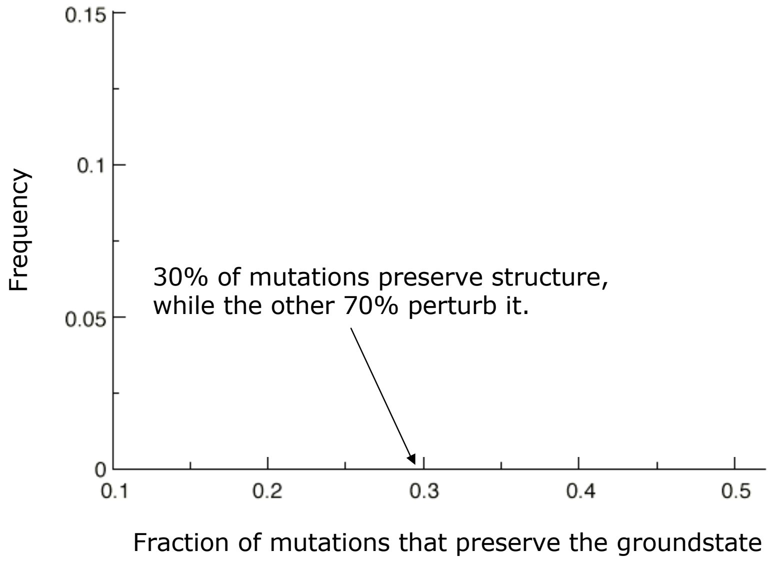
Evolved →



← Random

Evolution of mutational stability

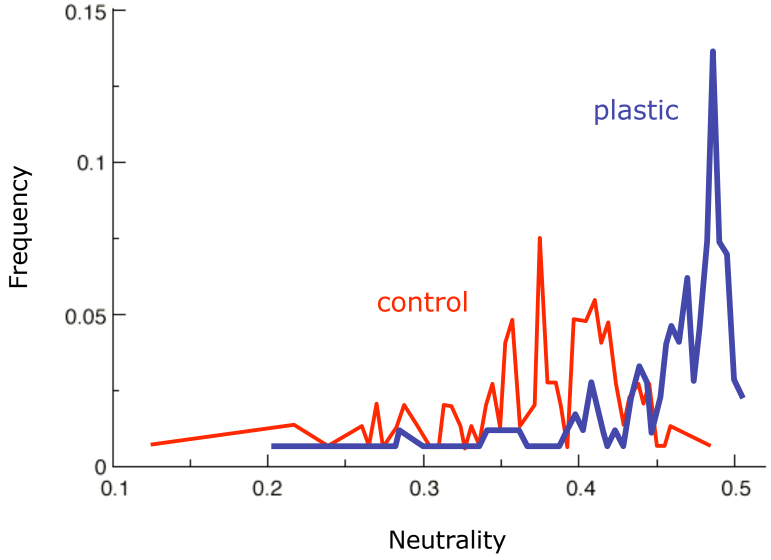
Can mutations create novelty?



30% of mutations preserve structure,
while the other 70% perturb it.

Evolution of mutational stability

Can mutations create novelty?



How do these populations get stuck?

Thermostability is favored

↓

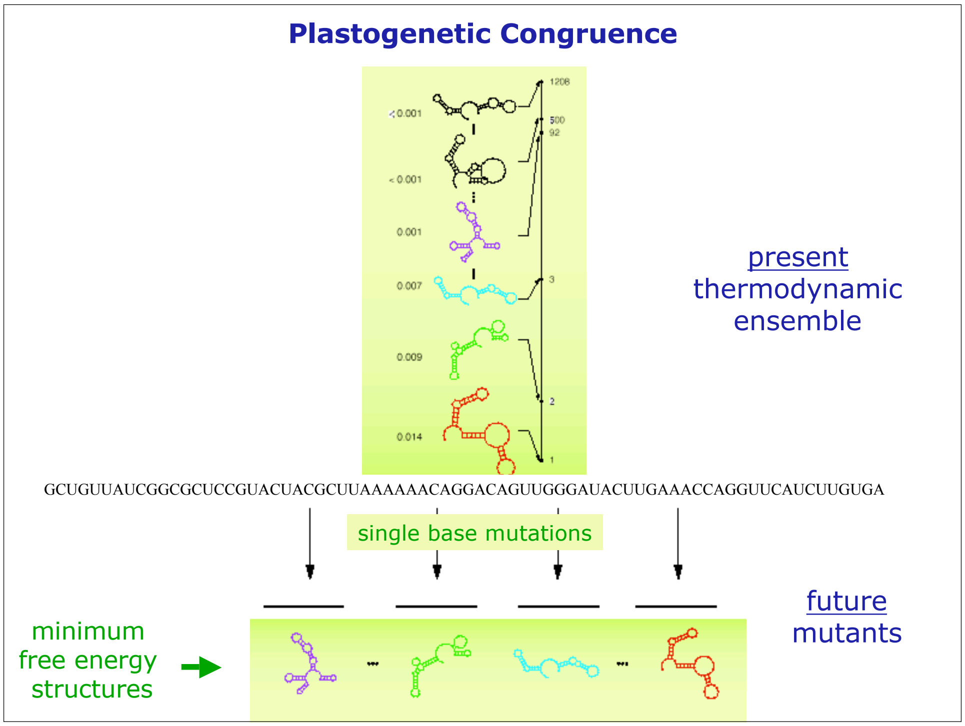
Natural selection increases thermostability

?

Sequences become insensitive to mutation

↓

Populations cannot access novelty through mutation,
and therefore cannot evolve



Plastogenetic Congruence

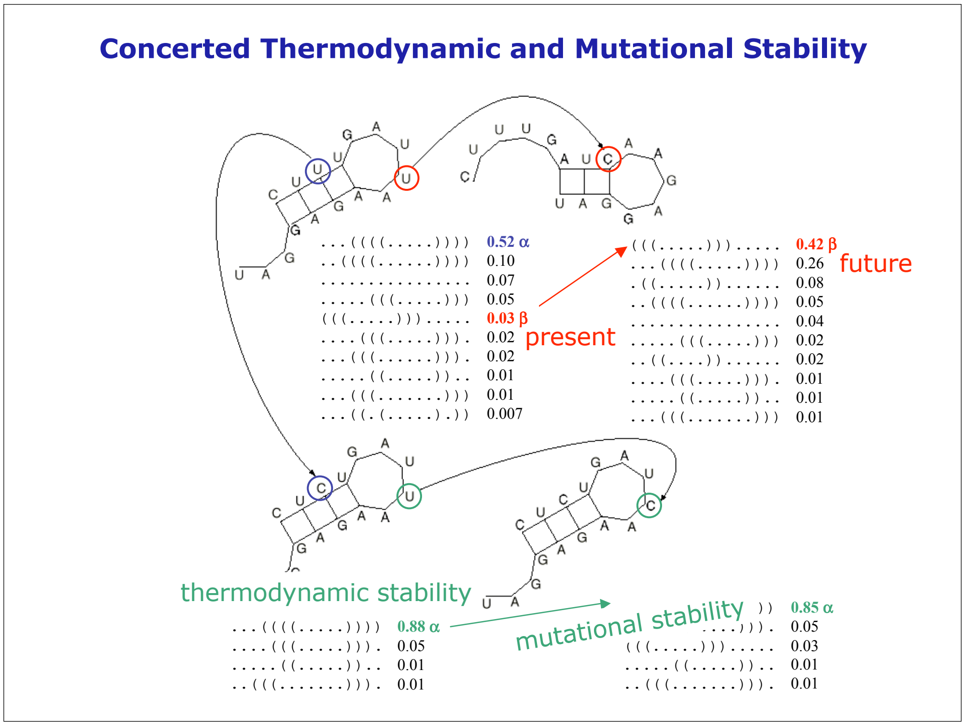
present thermodynamic ensemble

GCUGUUAUCGGCGCUCCGUACUACGCUUAAAAAACAGGACAGUUGGGAUACUUGAAACCAGGUUCAUCUUGUGA

single base mutations

minimum free energy structures

future mutants



Concerted Thermodynamic and Mutational Stability

...((((.....))))	**0.52 α**
..((((......))))	0.10
....	0.07
.....(((.....)))	0.05
(((.....))).....	**0.03 β**
....(((.....))).	0.02
...(((......))).	0.02
.....((.....))..	0.01
...(((.......)))	0.01
...((.(.....).))	0.007

present

(((.....))).....	**0.42 β**
...((((.....))))	0.26
.((.....))......	0.08
..((((......))))	0.05
....	0.04
.....(((.....)))	0.02
..((....))......	0.02
....(((.....))).	0.01
.....((.....))..	0.01
...(((.......)))	0.01

future

thermodynamic stability

...((((.....))))	**0.88 α**
....(((.....))).	0.05
.....((.....))..	0.01
..(((.......))).	0.01

mutational stability

))	**0.85 α**
....))).	0.05
(((.....))).....	0.03
.....((.....))..	0.01
..(((.......))).	0.01

How do these populations get stuck?

Thermostability is favored

↓

Natural selection increases thermostability

Plastogenetic congruence

Sequences become insensitive to mutation

↓

Populations cannot access novelty through mutation,
and therefore cannot evolve

General hypothesis:

Evolution of environmental stability hinders evolution?!

Selection to withstand environmental variability

Plastogenetic congruence (proteins, phenocopies, …)

Another byproduct of natural selection
for thermodynamic robustness ...

Modularity

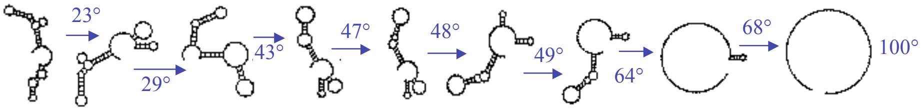
(1) Thermodynamic Modularity

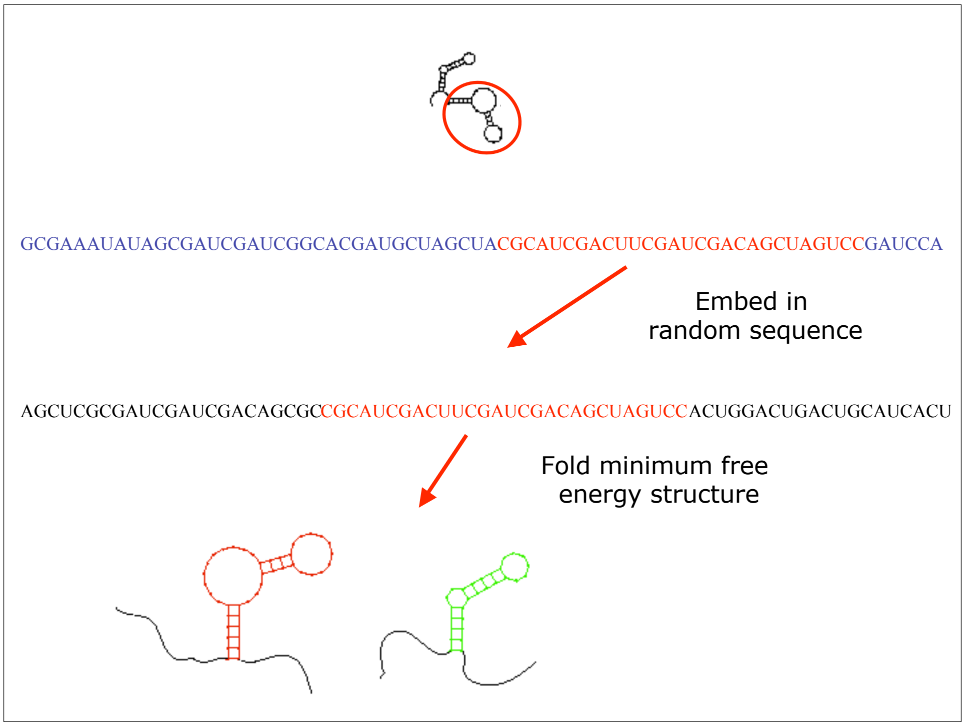
(2) Genetic Modularity

Melting Profile

Evolved Sequence

Random Sequence



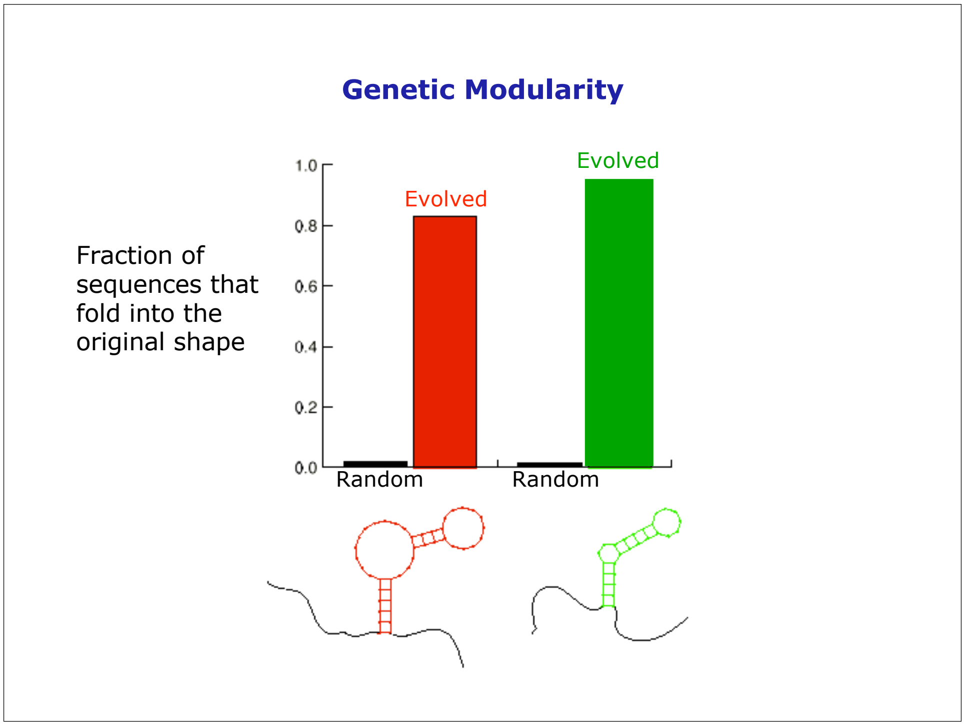


GCGAAAUAUAGCGAUCGAUCGGCACGAUGCUAGCUACGCAUCGACUUCGAUCGACAGCUAGUCCGAUCCA

Embed in
random sequence

AGCUCGCGAUCGAUCGACAGCGCCGCAUCGACUUCGAUCGACAGCUAGUCCACUGGACUGACUGCAUCACU

Fold minimum free
energy structure



Genetic Modularity

Fraction of
sequences that
fold into the
original shape

Conclusions

New evolutionary hypotheses

1. Plastogenetic congruence: A *biophysical* relationship between environmental and mutational stability
2. Selection to withstand environmental variability may impede adaptation
3. Selection to withstand environmental variability may explain the origins of modularity

Morals and Challenges

Simulations of complex biological systems can give new insight:

Phenomena not even accessible from traditional population genetic models

Inspire new generation of models, better engineering, and experimental tests

Building even more realistic models

Extracting the essential features/dynamics of a system

Testing these ideas

Extrapolating to other biological systems

Acknowledgements

Matt Cowperthwaite, University of Texas

Andy Ellington, University of Texas

Walter Fontana, Harvard University

James S. McDonnell Foundation

Santa Fe Institute

NSF