# Evolution and Design

Peter Schuster

Institut für Theoretische Chemie, Universität Wien, Austria
and
The Santa Fe Institute, Santa Fe, New Mexico, USA


Traunkirchner Gedankenexperimente

Traunkirchen, 13.09.2005

Web-Page for further information:

http://www.tbi.univie.ac.at/~pks

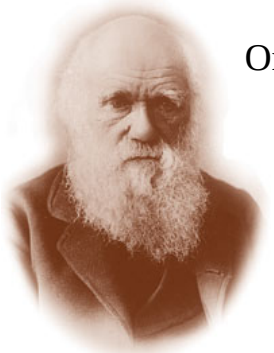
Cardinal Christoph Schönborn: ***The New York Times***, July 07, 2005

… Any system of thought that denies or seeks to explain away the overwhelming evidence for design in biology is ideology, not science.

… Scientific theories that try to explain away the appearance of design as the result of "chance and necessity" are not scientific at all, but, as John Paul put it, an abdication of human intelligence.

1. History of evolutionary thinking
2. Probabilities in biology
3. Complex patterns from simple rules
4. Mechanisms of evolution
5. Origins of complexity – The eye

1. **History of evolutionary thinking**
2. Probabilities in biology
3. Complex patterns from simple rules
4. Mechanisms of evolution
5. Origins of complexity – The eye

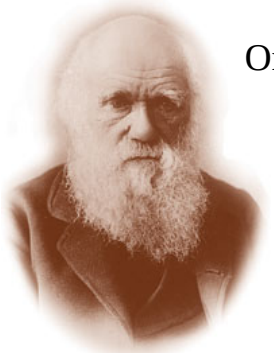


Charles Darwin

Origin of evolutionary biology
1859

Origin of genetics
1865

Gregor Mendel



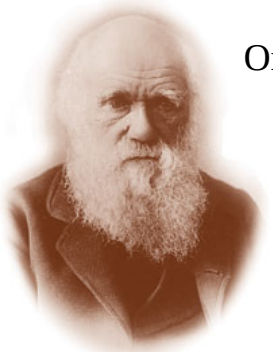
Charles Darwin

Origin of evolutionary biology
1859

Origin of genetics
1865

‘Rediscovery’ 1900

Gregor Mendel



Charles Darwin

Origin of evolutionary biology
1859

Origin of genetics
1865

Gregor Mendel

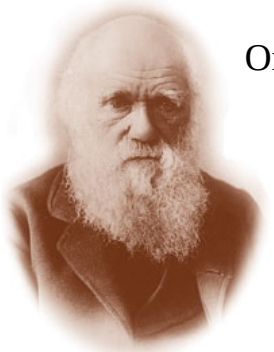
‘Rediscovery’ 1900

First unification: Population genetics 1930

Ronald Fisher

JSB Haldane

Sewall Wright



Charles Darwin

Origin of evolutionary biology
1859

Origin of genetics
1865

Gregor Mendel

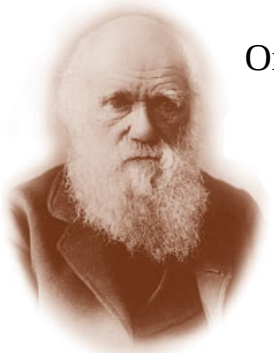
‘Rediscovery’ 1900

First unification: Population genetics 1930

Ernst Mayr

Theodosius
Dobzhansky

Synthetic or
Neo-Darwinian theory
1940 - 1950



Charles Darwin

Origin of evolutionary biology
1859

Origin of genetics
1865

‘Rediscovery’ 1900

Gregor Mendel

Friedrich Woehler

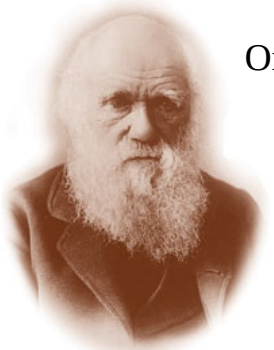
Origin of biochemistry
1828

First unification: Population genetics 1930

Ernst Mayr

Theodosius Dobzhansky

Synthetic or
Neo-Darwinian theory
1940 - 1950



Charles Darwin

Origin of evolutionary biology
1859

Origin of genetics
1865

‘Rediscovery’ 1900

Gregor Mendel

Friedrich Woehler

Origin of biochemistry
1828

Origin of molecular biology 1953

James Watson and Francis Crick

First unification: Population genetics 1930

Ernst Mayr

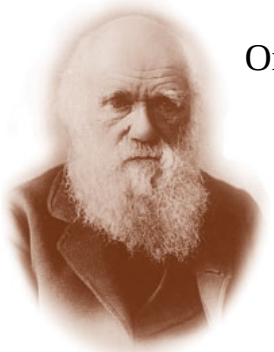
Theodosius Dobzhansky

Synthetic or
Neo-Darwinian theory
1940 - 1950

Max Perutz

John Kendrew

**Biology of the 21st century**



Charles Darwin

Origin of evolutionary biology
1859

Origin of genetics
1865

‘Rediscovery’ 1900

Gregor Mendel

Friedrich Woehler

Origin of biochemistry
1828

First unification: Population genetics 1930

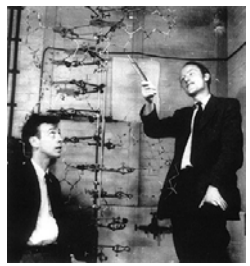
Ernst Mayr

Theodosius Dobzhansky

Synthetic or
Neo-Darwinian theory
1940 - 1950

Origin of molecular biology 1953

Jacques Monod



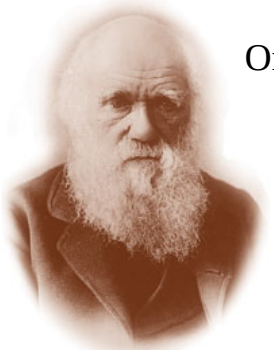
James Watson and Francis Crick

François Jacob

Max Perutz

John Kendrew

**Biology of the 21st century**



Charles Darwin

Origin of evolutionary biology
1859

Origin of genetics
1865

‘Rediscovery’ 1900

Gregor Mendel

Friedrich Woehler

Origin of biochemistry
1828

First unification: Population genetics 1930

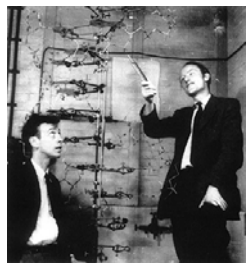
Ernst Mayr

Theodosius Dobzhansky

Synthetic or
Neo-Darwinian theory
1940 - 1950

Origin of molecular biology 1953

Jacques Monod



James Watson and Francis Crick

François Jacob

Max Perutz

Manfred Eigen

Sydney Brenner

John Kendrew

**Biology of the 21st century**

Biomathematics, bioinformatics, … , biophysics, biochemistry, … , molecular genetics, … , systems biology, biomedicine, macroscopic biology, evolutionary biology, sociobiology, anthropology, …

Ernst Mayr and others:

Can we explain the observations in biology without the assumption of a **causa finalis** ?

The answer is **yes**, adaptation through variation and selection leads to the same result as rational design.

"**Teleonomy replaces teleology**"

Evolutionary biotechnology was able to prove this statement.

William of Ockham, ~1285 – 1349

Ockham's razor: *„ ... plurality should not be assumed without necessity,"*

or in modern English: *„ ... keep it simple, unsophisticated, even stupid."*

No concept, construct or variable should be used that is not required for the explanation of phenomena.

If we don't need a causa finalis, we have to dismiss it without replacement.

1. History of evolutionary thinking
2. **Probabilities in biology**
3. Complex patterns from simple rules
4. Mechanisms of evolution
5. Origins of complexity – The eye

Eugene Wigner's argument applied to a bacterium:

**All genomes have equal probability**

5'-end GCGGATTTAGCTCAGTTGGGAGAGCGCCAGACTGAAGATCTGGAGGTCCTGTGTTCGAUCCACAGAATTC.....GCACCA 3'-end

Alphabet size: 4

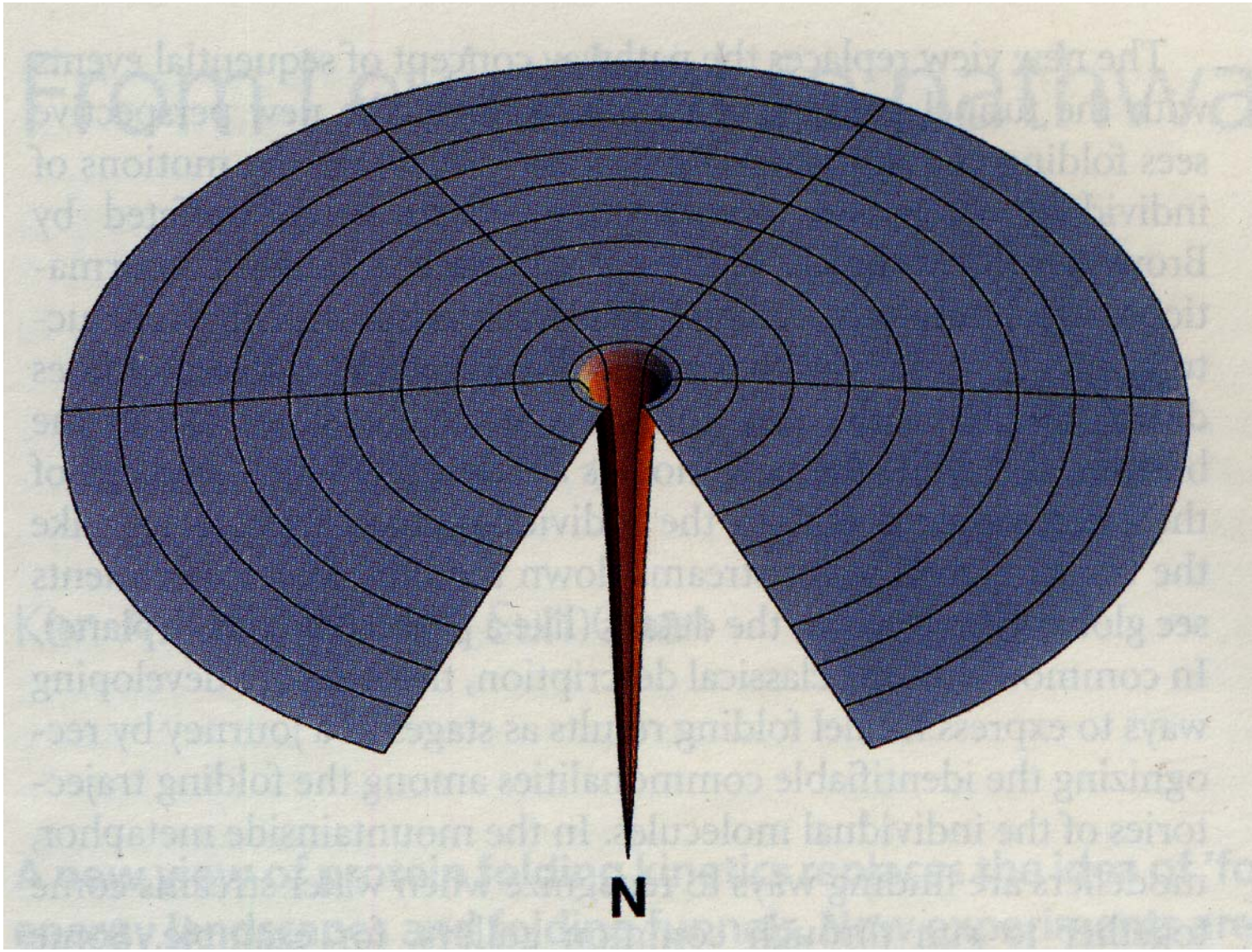
Chain length: ≈ 1 000 000 nucleotides

Number of possible genomes: $4^{1000000}$

Probability to find a given bacterial genome:

$$4^{-1000000} \approx 10^{-600000} = 0.000\ldots\ldots001$$

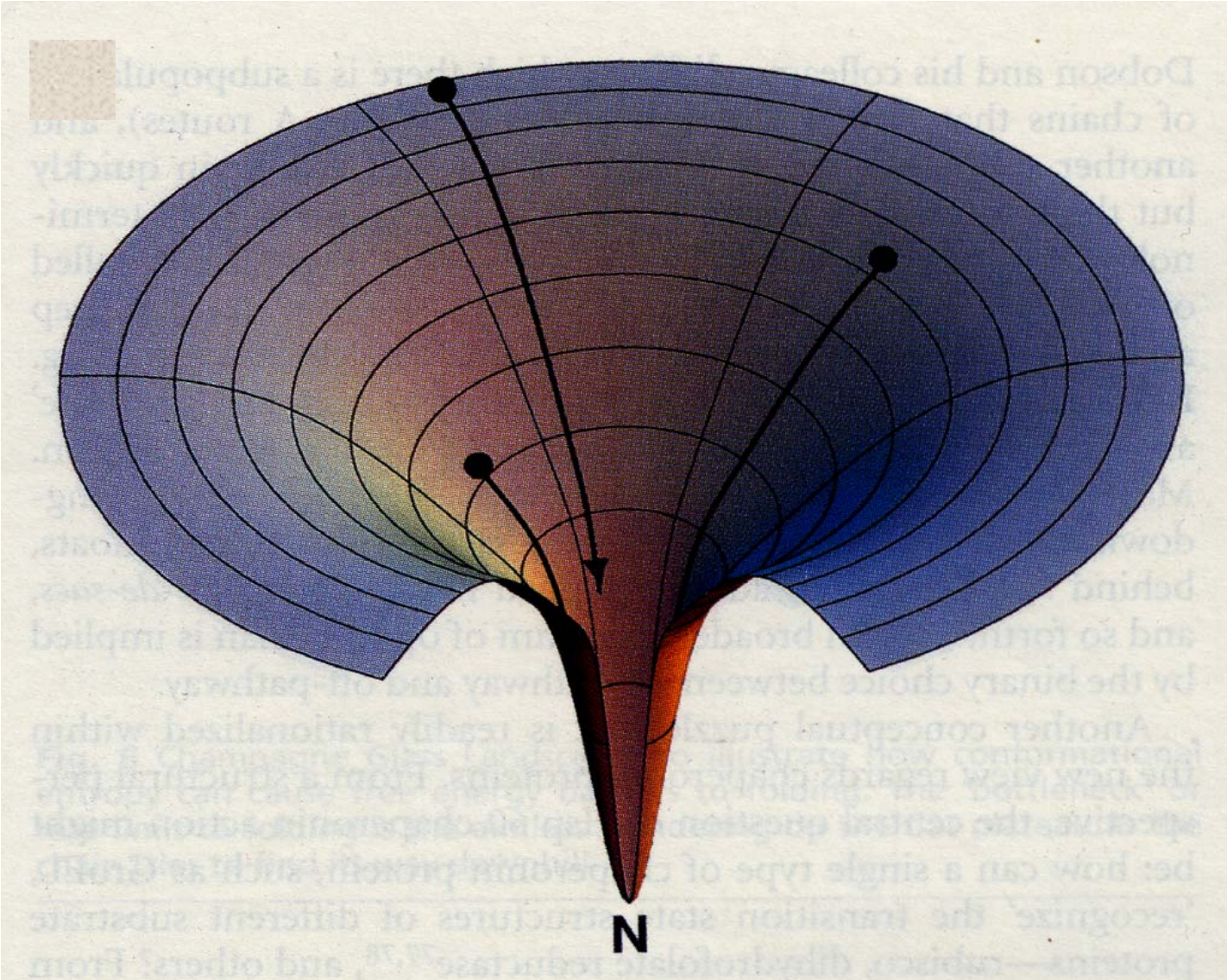
$\Leftarrow$ 600000 $\Rightarrow$




The golf course landscape

Wigner’s paradox

Picture: K.A. Dill, H.S. Chan, Nature Struct. Biol. 4:10-19




The funnel landscape

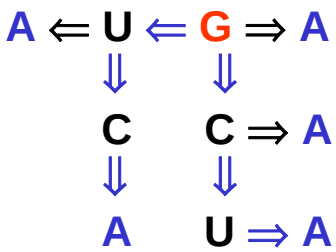
Solution to Wigner’s paradox

Picture: K.A. Dill, H.S. Chan, Nature Struct. Biol. 4:10-19

Eugene Wigner's argument revisited:

**Every single point mutation leads to an improvement and is therefore selected**

5'-end GCGGATTTAGCTCAGTTGGGAGAGCGCCAGACTGAAGATCTGGAGGTCCTGTGTTCGAUCCACAGAATTC.....GCACCA 3'-end



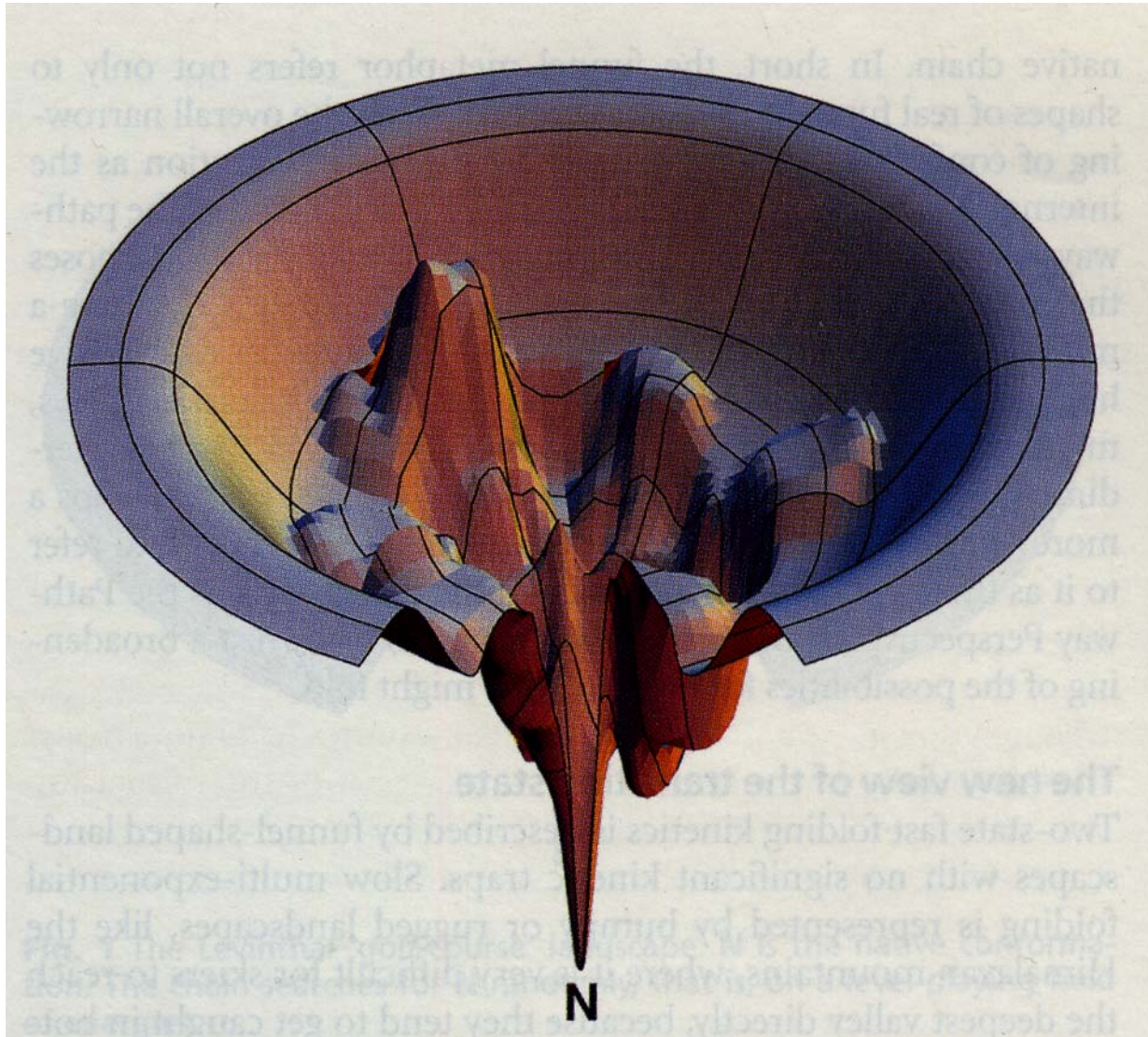

Alphabet size: 4

Chain length: ≈ 1 000 000 nucleotides

Length of longest path to the optimum: $3 \times 1000000$

Probability to find the optimal bacterial genome:

$$0.333.. \times 10^{-6} = 0.000000333..$$

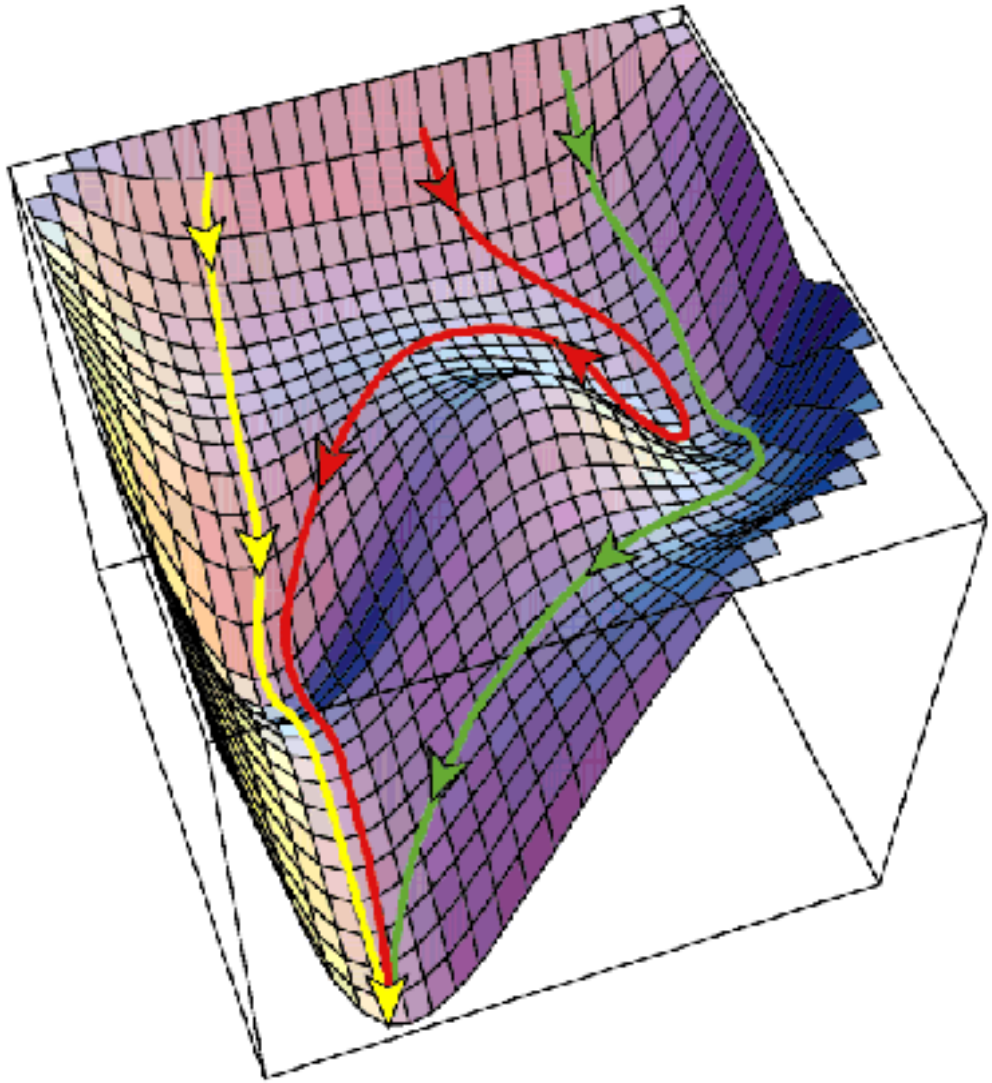



The structured funnel landscape

Solution to Wigner’s paradox

Picture: K.A. Dill, H.S. Chan, Nature Struct. Biol. 4:10-19

# An “all-roads-lead-to-Rome” landscape



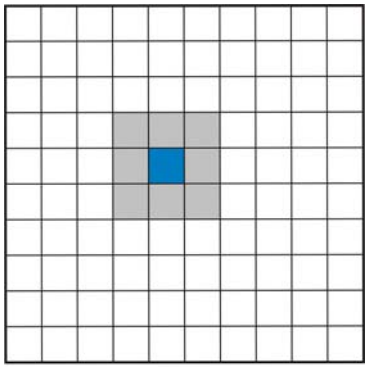
The reconstructed folding landscape
of a real biomolecule: “Lysozyme”

Picture: C.M. Dobson, A. Šali, and M. Karplus, Angew.Chem.Internat.Ed. 37: 868-893, 1988

**But (!)** landscapes of evolution in nature and in the laboratory are unlike all the four examples shown here!

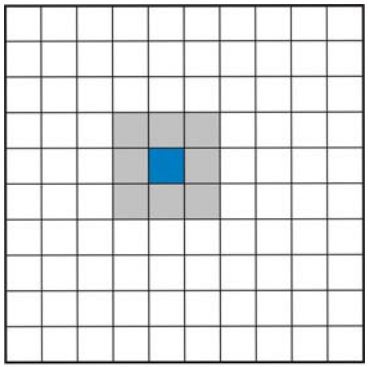
1. History of evolutionary thinking
2. Probabilities in biology
3. **Complex patterns from simple rules**
4. Mechanisms of evolution
5. Origins of complexity – The eye

# John Horton Conway's Game-of-Life

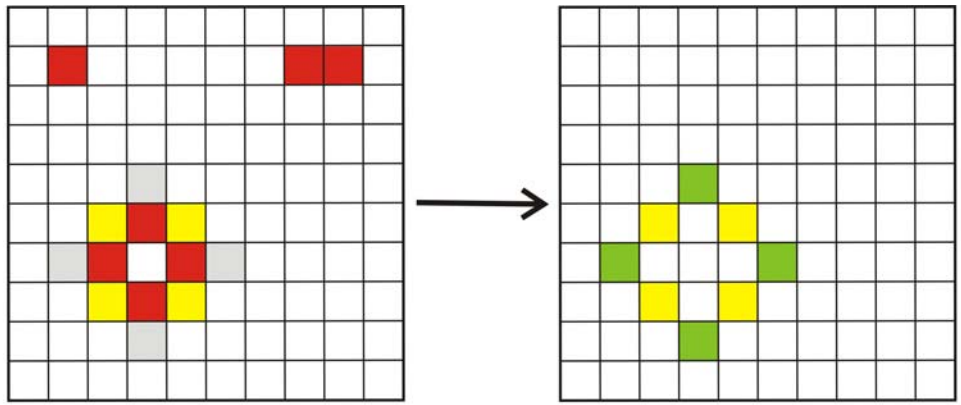


Cell and neigborhood

# John Horton Conway's Game-of-Life



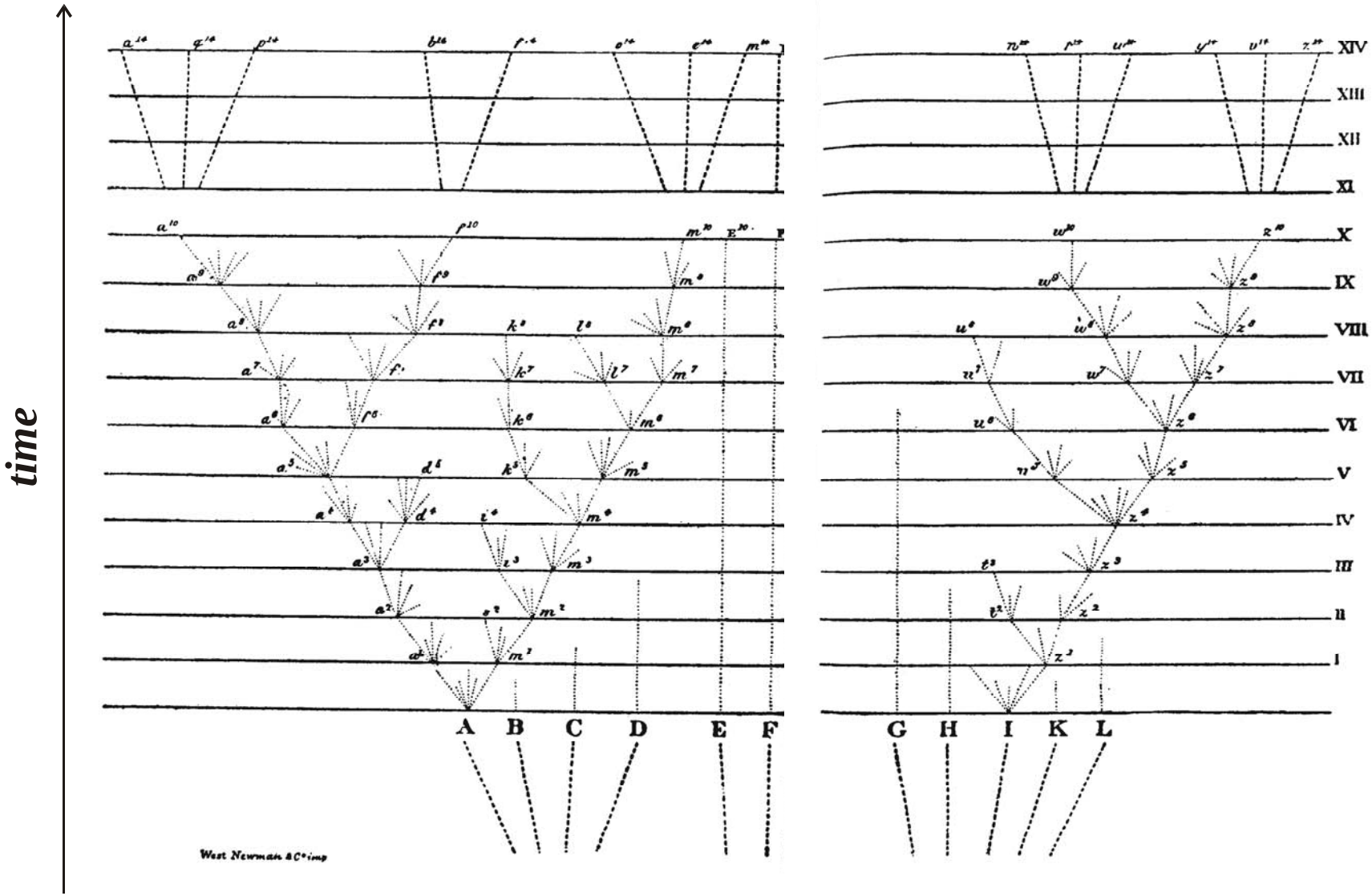
Cell and neigborhood



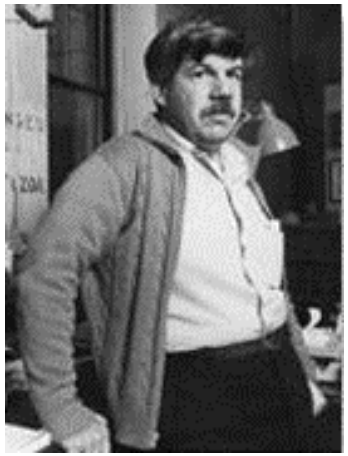
Populated cell: (1) each cell with one or no neighbors dies
(2) each cell with two or three neighbors suvives
(3) each cell with four or more neighbors dies

Empty cell: (4) each empty cell with three neighbors becomes populated.

1. History of evolutionary thinking
2. Probabilities in biology
3. Complex patterns from simple rules
4. **Mechanisms of evolution**
5. Origins of complexity – The eye




Charles Darwin, *The Origin of Species*, 6th edition.
Everyman‘s Library, Vol.811, Dent London, pp.121-122.

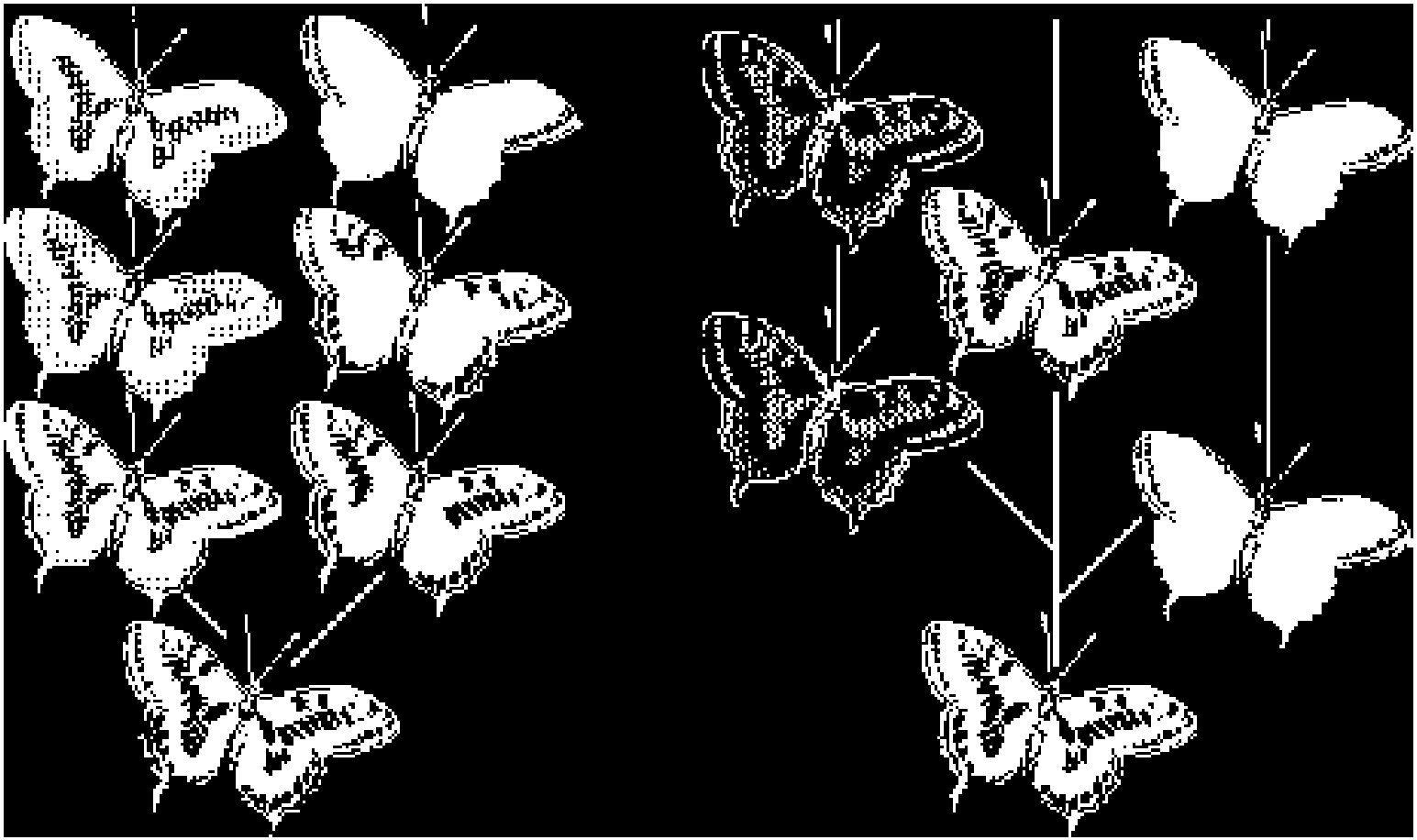


Stephen Jay Gould, 1941 – 2002

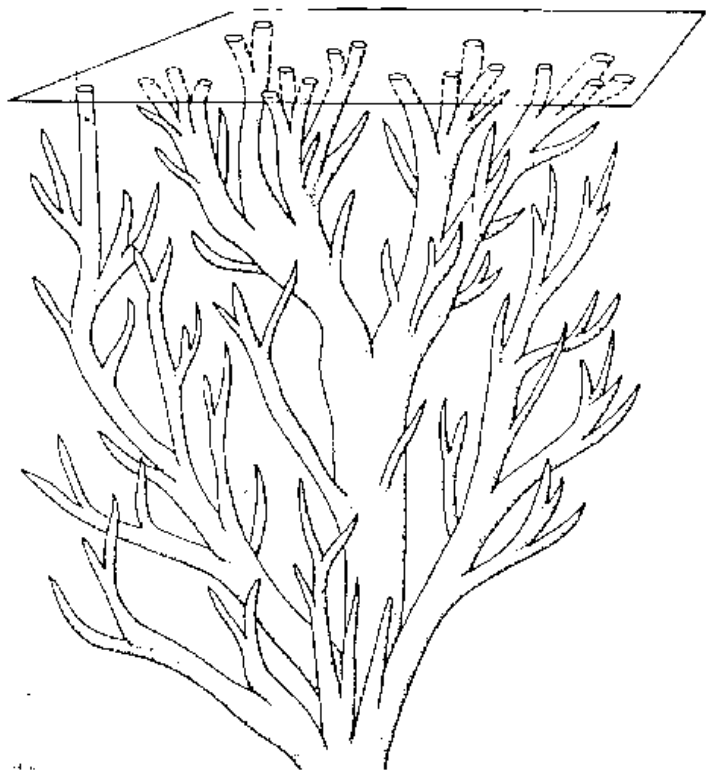
**Punctuated Equilibrium:**

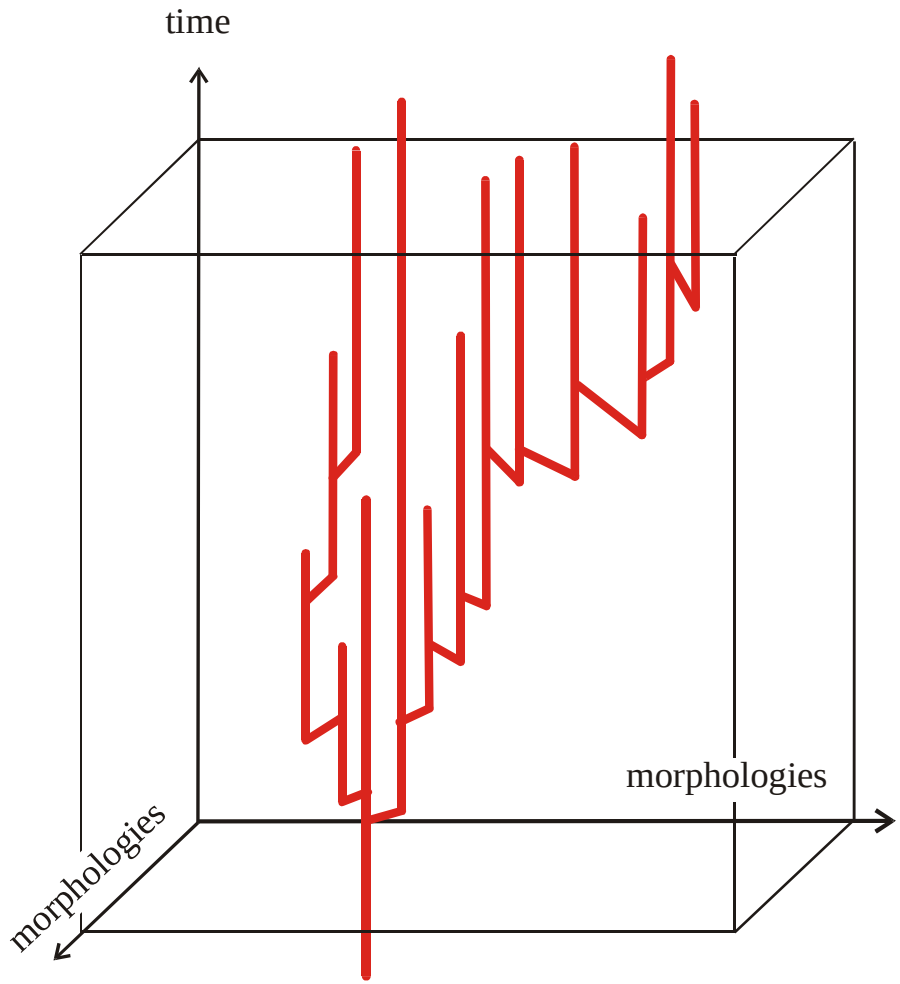
Evolution occurs through abrupt changes and not gradual.

Niles Eldredge, 1943 –

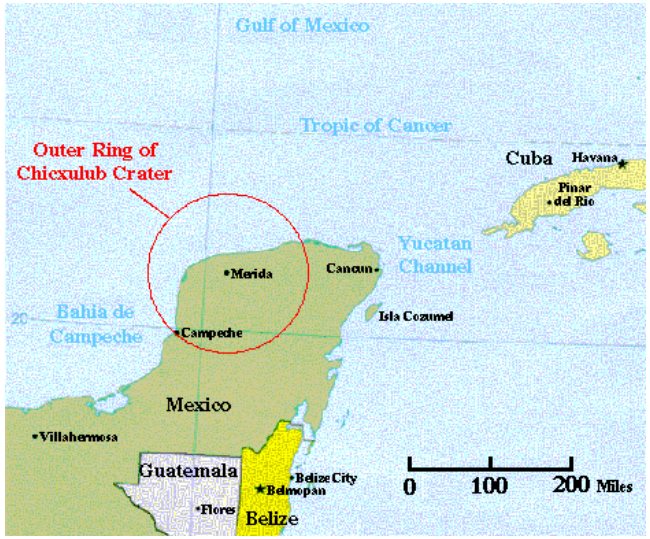


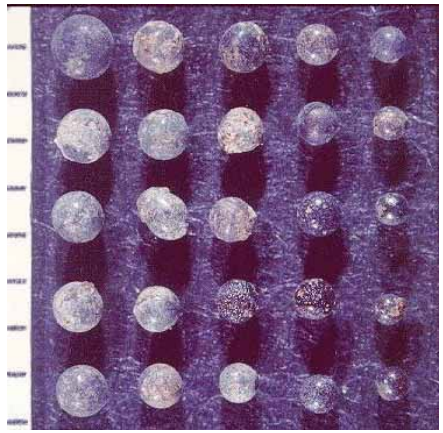
Gradual change versus punctuated equilibrium in butterfly colors

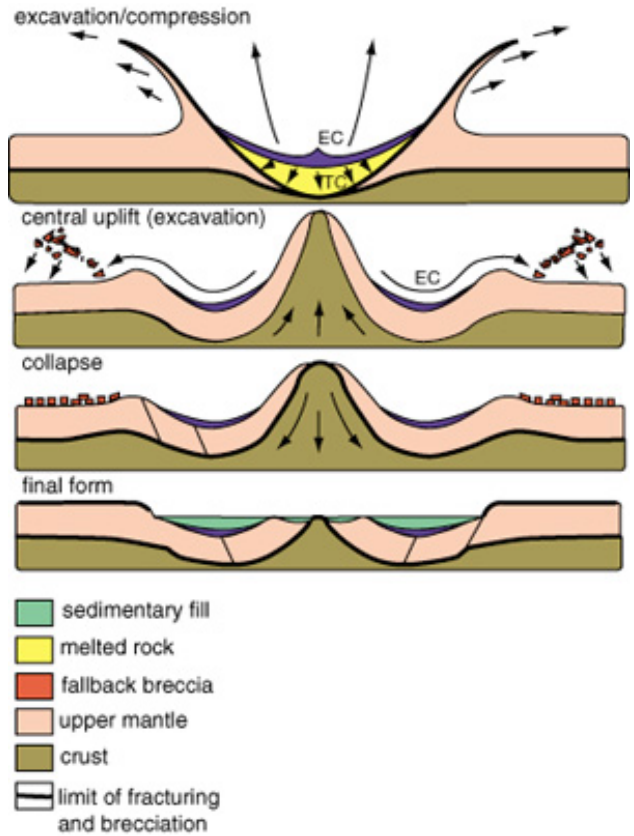





Phyletic tree as pictured by the gradualists' and the punctuated equilibrium approach






Falling meterorites:
An example is the Chicxulub crater in Mexico dated 65 million years ago

L.W.Alvarez, *Mass Extinctions caused by large bolide impacts*. Physics Today **40**: 24-33, 1987

| | Generation time | Selection and adaptation 10 000 generations | Genetic drift in small populations $10^6$ generations | Genetic drift in large populations $10^7$ generations |
|---|---|---|---|---|
| RNA molecules | 10 sec<br>1 min | 27.8 h = 1.16 d<br>6.94 d | 115.7 d<br>1.90 a | 3.17 a<br>19.01 a |
| Bacteria | 20 min<br>10 h | 138.9 d<br>11.40 a | 38.03 a<br>1 140 a | 380 a<br>11 408 a |
| Multicelluar organisms | 10 d<br>20 a | 274 a<br>20 000 a | 27 380 a<br>$2 \times 10^7$ a | 273 800 a<br>$2 \times 10^8$ a |

Time scales of evolutionary change

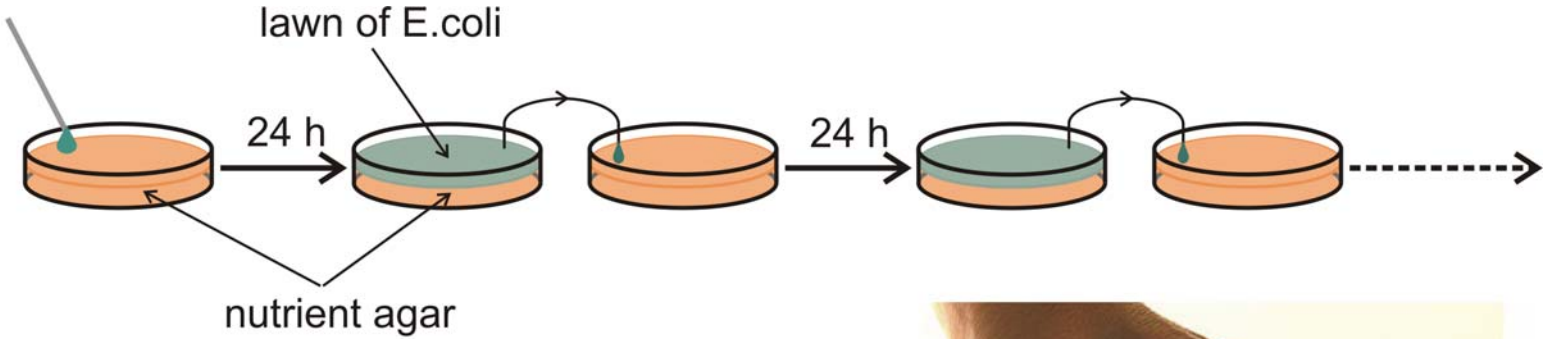
## Bacterial Evolution

S. F. Elena, V. S. Cooper, R. E. Lenski. ***Punctuated evolution caused by selection of rare beneficial mutants***. Science **272** (1996), 1802-1804

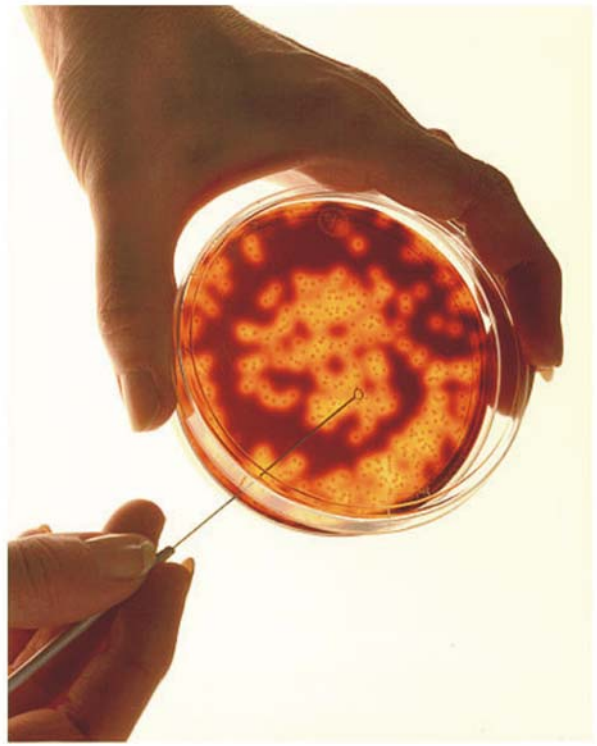
D. Papadopoulos, D. Schneider, J. Meier-Eiss, W. Arber, R. E. Lenski, M. Blot. ***Genomic evolution during a 10,000-generation experiment with bacteria***. Proc.Natl.Acad.Sci.USA **96** (1999), 3807-3812

S. F. Elena, R. E. Lenski. ***Evolution experiments with microorganisms: The dynamics and genetic bases of adaptation***. Nature Review Genetics **4** (2003), 457-469

C. Borland, R. E. Lenski. ***Spontaneous evolution of citrate utilization in* Escherichia coli *after 30000 generations***. Evolution Conference 2004, Fort Collins, Colorado




1 day » 6.67 generations
1 month » 200 generations
1 year » 2400 generations



Serial transfer of Escherichia coli cultures in Petri dishes

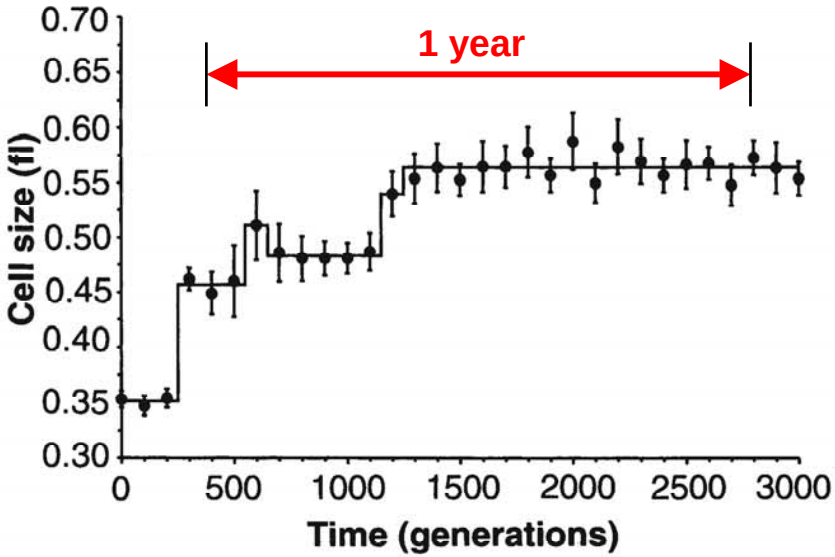



**Fig. 1.** Change in average cell size (1 fl = $10^{-15}$ L) in a population of *E. coli* during 3000 generations of experimental evolution. Each point is the mean of 10 replicate assays (*22*). Error bars indicate 95% confidence intervals. The solid line shows the best fit of a step-function model to these data (Table 1).

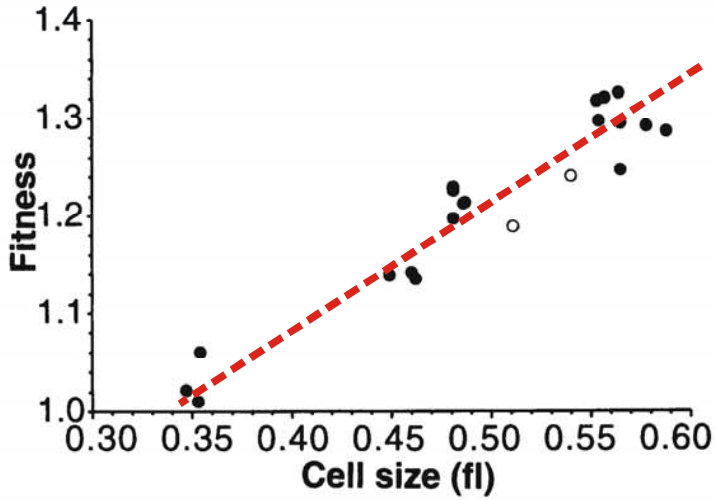
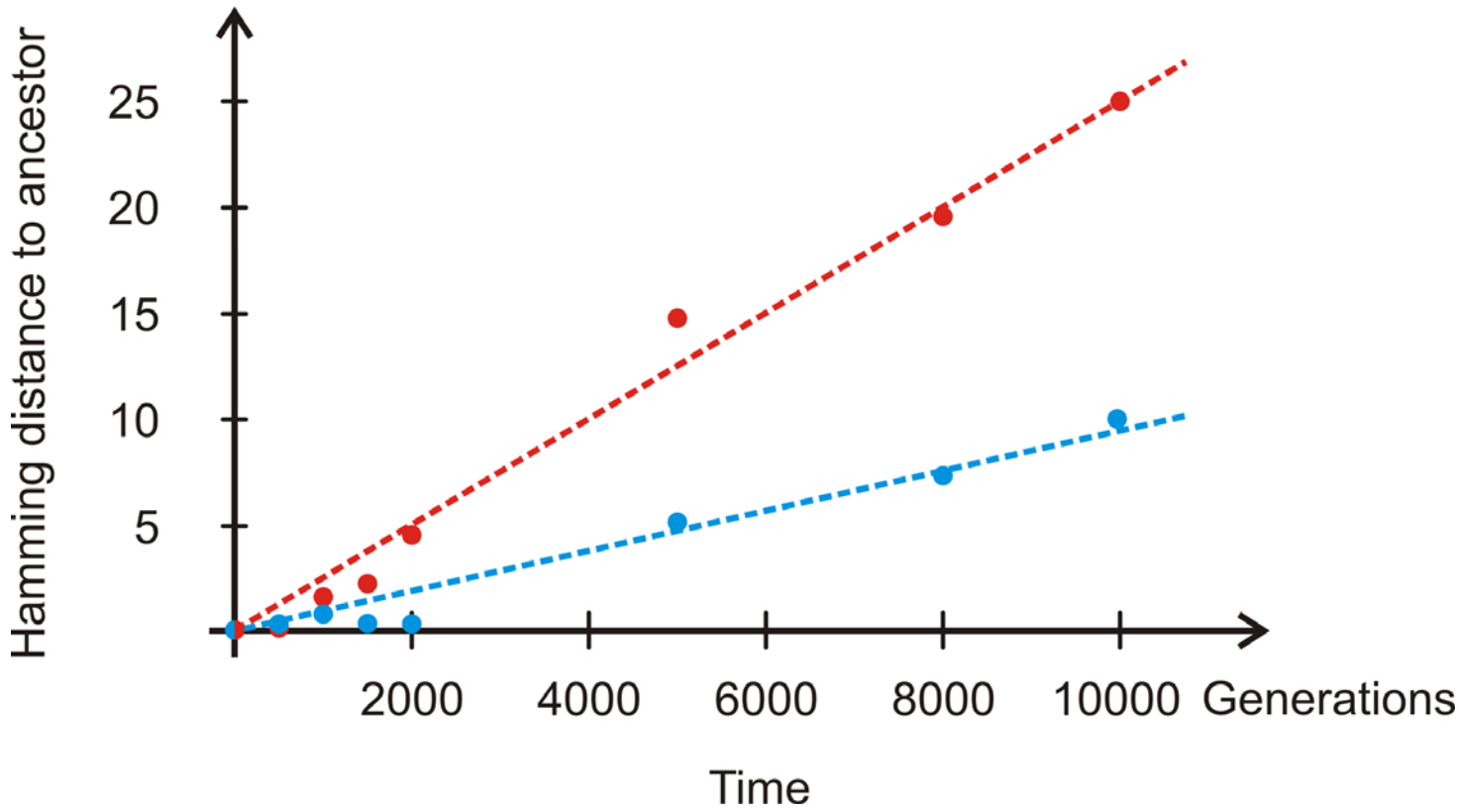



**Fig. 2.** Correlation between average cell size and mean fitness, each measured at 100-generation intervals for 2000 generations. Fitness is expressed relative to the ancestral genotype and was obtained from competition experiments between derived and ancestral cells (*6, 7*). The open symbols indicate the only two samples assigned to different steps by the cell size and fitness data.

Epochal evolution of bacteria in serial transfer experiments under constant conditions

S. F. Elena, V. S. Cooper, R. E. Lenski. *Punctuated evolution caused by selection of rare beneficial mutants*. Science **272** (1996), 1802-1804




Variation of genotypes in a bacterial serial transfer experiment

D. Papadopoulos, D. Schneider, J. Meier-Eiss, W. Arber, R. E. Lenski, M. Blot. *Genomic evolution during a 10,000-generation experiment with bacteria*. Proc.Natl.Acad.Sci.USA **96** (1999), 3807-3812

**Innovation after 33 000 generations:**

One out of 12 *Escherichia coli* colonies adapts to the environment and starts spontaneously to utilize citrate in the medium.

# Evolution of RNA molecules based on Qβ phage


D.R.Mills, R.L.Peterson, S.Spiegelman, ***An extracellular Darwinian experiment with a self-duplicating nucleic acid molecule***. Proc.Natl.Acad.Sci.USA **58** (1967), 217-224

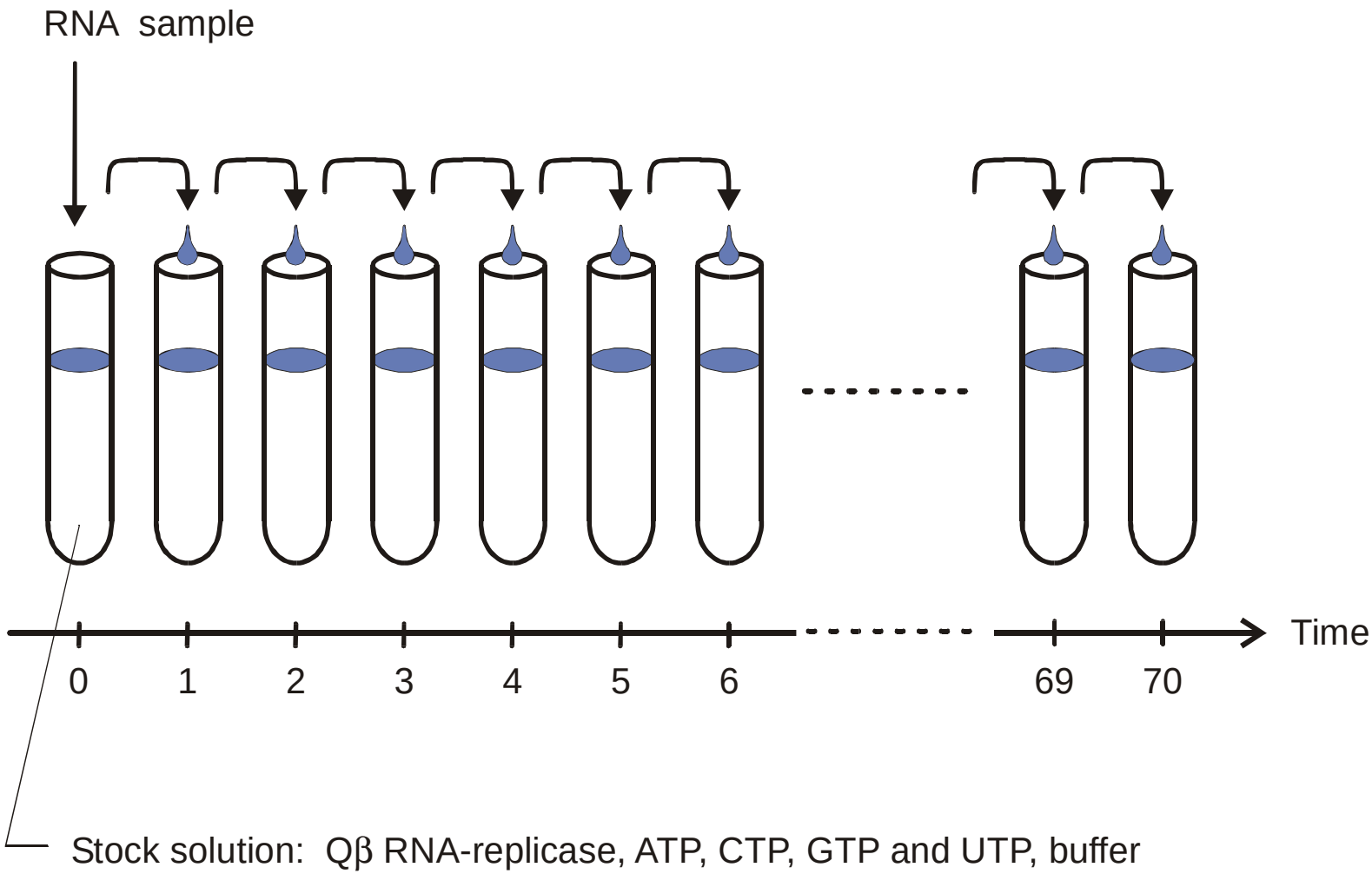
S.Spiegelman, ***An approach to the experimental analysis of precellular evolution***. Quart.Rev.Biophys. **4** (1971), 213-253

C.K.Biebricher, ***Darwinian selection of self-replicating RNA molecules***. Evolutionary Biology **16** (1983), 1-52

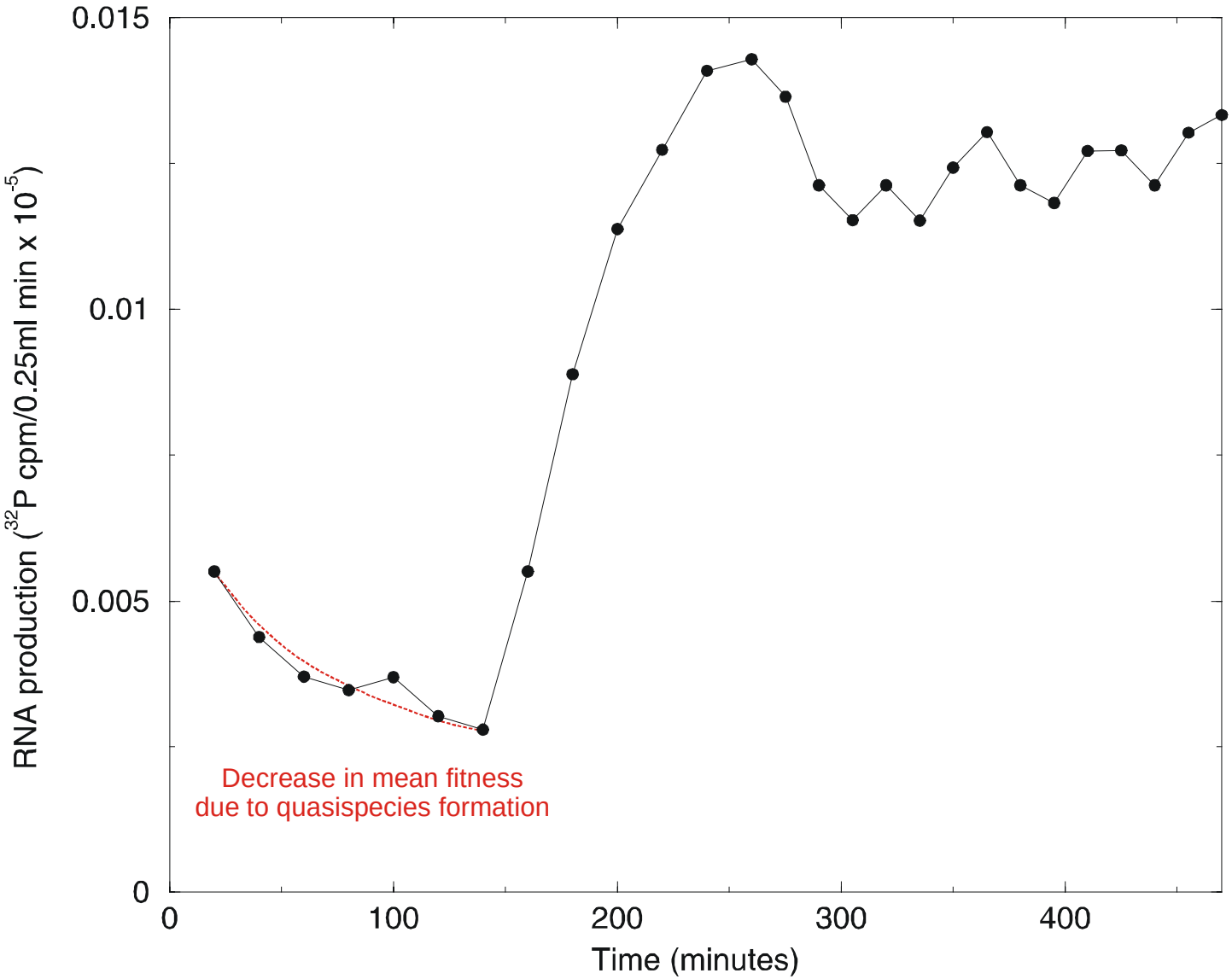
G.Bauer, H.Otten, J.S.McCaskill, ***Travelling waves of* in vitro *evolving RNA.*** *Proc.Natl.Acad.Sci.USA* **86** (1989), 7937-7941

C.K.Biebricher, W.C.Gardiner, ***Molecular evolution of RNA* in vitro**. Biophysical Chemistry **66** (1997), 179-192

G.Strunk, T.Ederhof, ***Machines for automated evolution experiments* in vitro *based on the serial transfer concept***. Biophysical Chemistry **66** (1997), 193-202




The serial transfer technique applied to RNA evolution ***in vitro***




The increase in RNA production rate during a serial transfer experiment

# Evolutionary design of RNA molecules


D.B.Bartel, J.W.Szostak, **In vitro *selection of RNA molecules that bind specific ligands***. Nature **346** (1990), 818-822

C.Tuerk, L.Gold, **SELEX - *Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage* T4 *DNA polymerase***. Science **249** (1990), 505-510

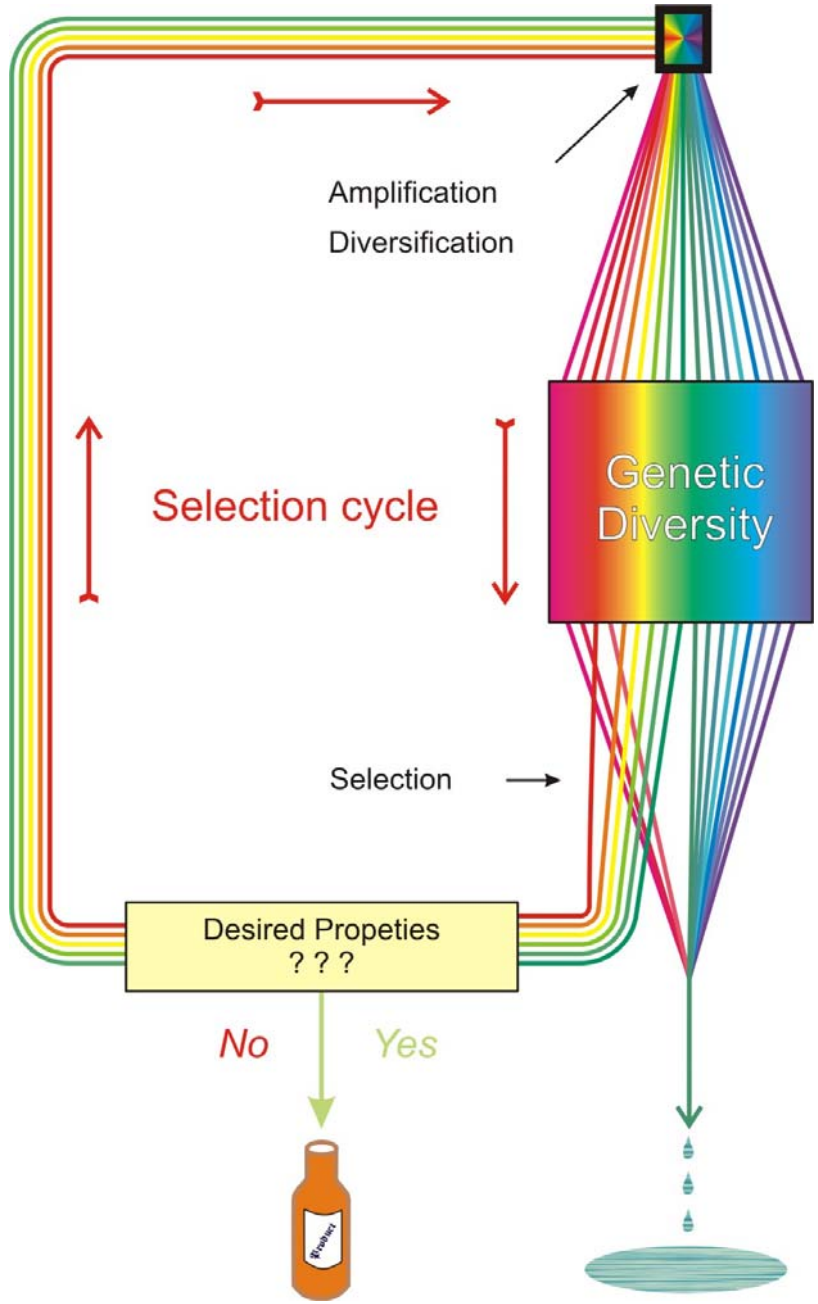
D.P.Bartel, J.W.Szostak, ***Isolation of new ribozymes from a large pool of random sequences***. Science **261** (1993), 1411-1418

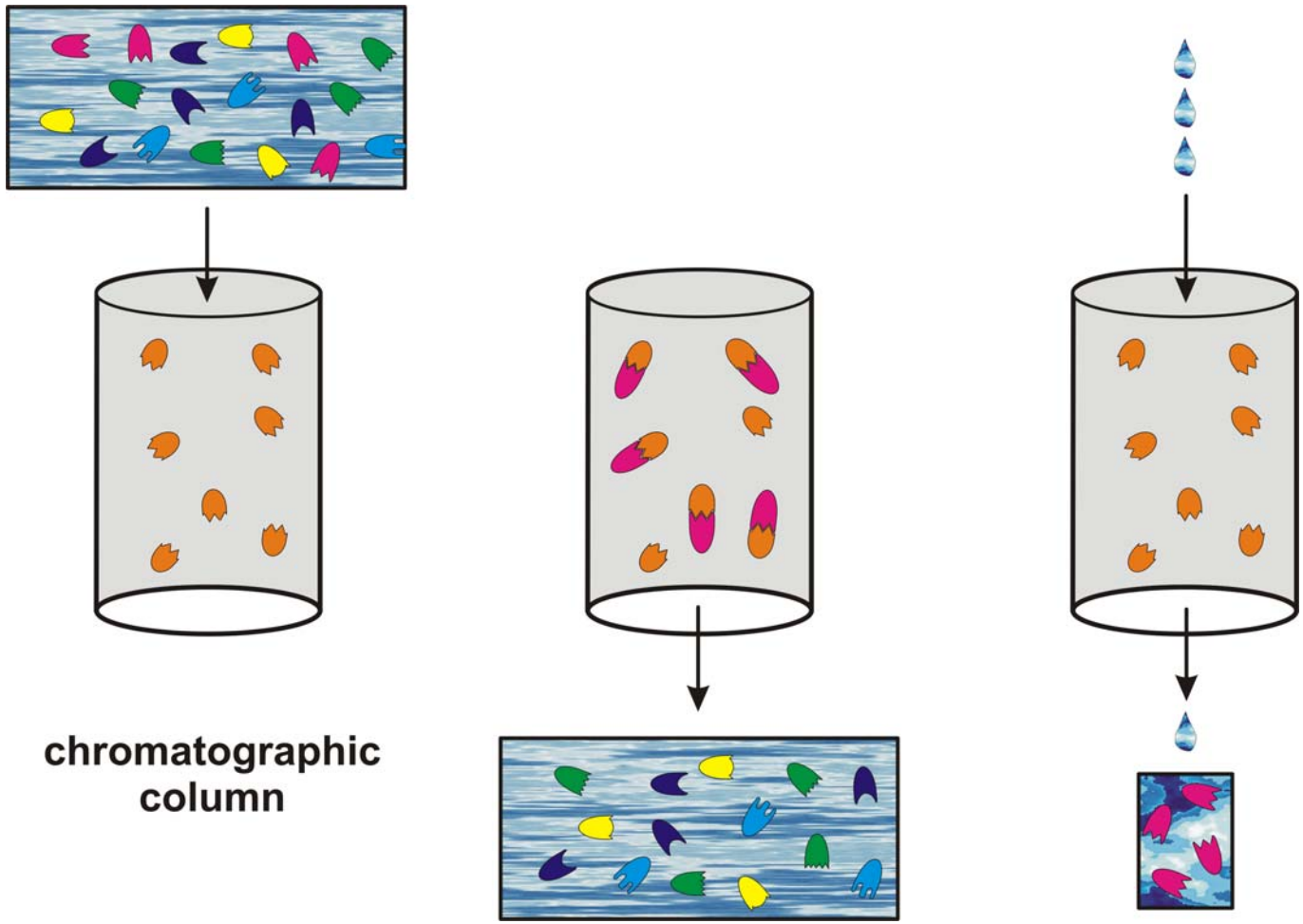
R.D.Jenison, S.C.Gill, A.Pardi, B.Poliski, ***High-resolution molecular discrimination by RNA***. Science **263** (1994), 1425-1429

Y. Wang, R.R.Rando, ***Specific binding of aminoglycoside antibiotics to RNA***. Chemistry & Biology **2** (1995), 281-290

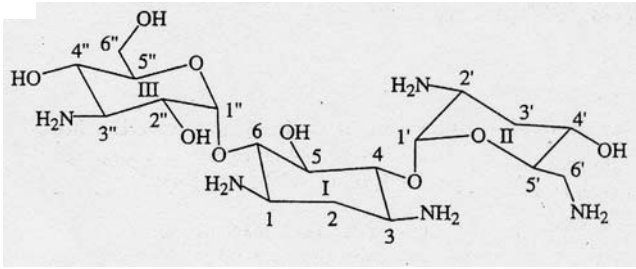
Jiang, A. K. Suri, R. Fiala, D. J. Patel, ***Saccharide-RNA recognition in an aminoglycoside antibiotic-RNA aptamer complex***. Chemistry & Biology **4** (1997), 35-50

An example of ‘artificial selection’ with RNA molecules or ‘breeding’ of biomolecules



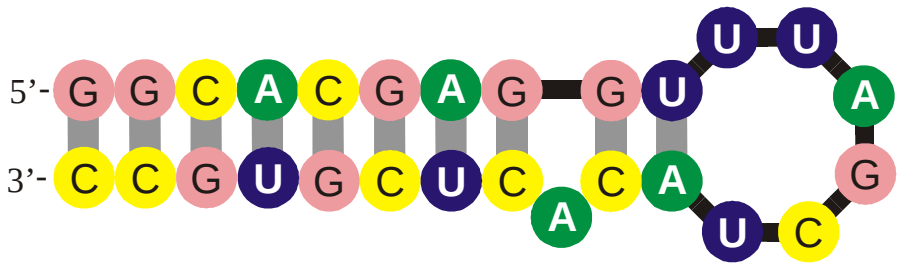



The SELEX technique for the evolutionary preparation of aptamers




tobramycin

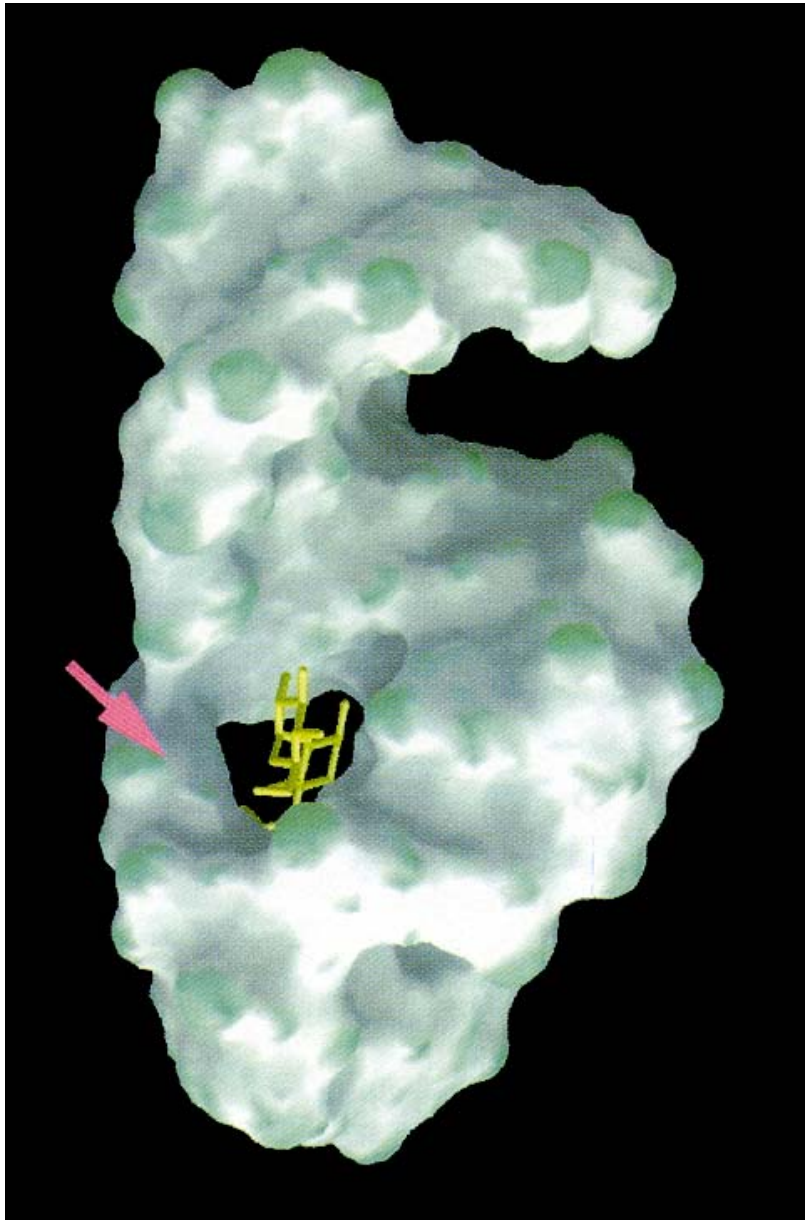
5’-GGCACGAGGUUUAGCUACACUCGUGCC-3’




RNA aptamer

Formation of secondary structure of the tobramycin binding RNA aptamer with $K_D$ = 9 nM

L. Jiang, A. K. Suri, R. Fiala, D. J. Patel, ***Saccharide-RNA recognition in an aminoglycoside antibiotic-RNA aptamer complex.*** Chemistry & Biology **4**:35-50 (1997)



The three-dimensional structure of the tobramycin aptamer complex

L. Jiang, A. K. Suri, R. Fiala, D. J. Patel, Chemistry & Biology **4**:35-50 (1997)

# Computer simulation of RNA optimization

Walter Fontana and Peter Schuster,
Biophysical Chemistry 26:123-147, 1987

Walter Fontana, Wolfgang Schnabl, and
Peter Schuster, Phys.Rev.A 40:3301-3321, 1989

PHYSICAL REVIEW A VOLUME 40, NUMBER 6 SEPTEMBER 15, 1989

## Physical aspects of evolutionary optimization and adaptation

Walter Fontana, Wolfgang Schnabl, and Peter Schuster*
*Institut für Theoretische Chemie der Universität Wien, Währingerstrasse 17, A 1090 Wien, Austria*
(Received 2 February 1989; revised manuscript received 5 May 1989)

A model of an objective function based on polynucleotide folding is used to investigate the dynamics of evolutionary adaptation in finite populations. Binary sequences are optimized with respect to their kinetic properties through a stochastic process involving mutation and selection. The objective function consists in a mapping from the set of all binary strings with given length into a set of two-dimensional structures. These structures then encode the kinetic properties, expressed in terms of parameters of reaction probability distributions. The objective function obtained thereby represents a realistic example of a highly "rugged landscape." Ensembles of molecular strings adapting to this landscape are studied by tracing their escape path from local optima and by applying multivariate analysis. Effects of small population numbers in the tail of the sequence distribution are discussed quantitatively. Close upper bounds to the number of distinct values produced by our objective function are given. The distribution of values is explored by means of simulated annealing and reveals a random scatter in the locations of optima in the space of all sequences. The genetic optimization protocol is applied to the "traveling salesman" problem.

Biophysical Chemistry 26 (1987) 123–147
Elsevier

BPC 01133



## A computer model of evolutionary optimization

Walter Fontana and Peter Schuster
*Institut für theoretische Chemie und Strahlenchemie der Universität Wien, Währingerstraße 17, A-1090 Wien, Austria*

Accepted 27 February 1987

Molecular evolution; Optimization; Polyribonucleotide folding; Quasi-species; Selective value; Stochastic reaction kinetics

Molecular evolution is viewed as a typical combinatorial optimization problem. We analyse a chemical reaction model which considers RNA replication including correct copying and point mutations together with hydrolytic degradation and the dilution flux of a flow reactor. The corresponding stochastic reaction network is implemented on a computer in order to investigate some basic features of evolutionary optimization dynamics. Characteristic features of real molecular systems are mimicked by folding binary sequences into unknotted two-dimensional structures. Selective values are derived from these molecular 'phenotypes' by an evaluation procedure which assigns numerical values to different elements of the secondary structure. The fitness function obtained thereby contains nontrivial long-range interactions which are typical for real systems. The fitness landscape also reveals quite involved and bizarre local topologies which we consider also representative of polynucleotide replication in actually occurring systems. Optimization operates on an ensemble of sequences via mutation and natural selection. The strategy observed in the simulation experiments is fairly general and resembles closely a heuristic widely applied in operations research areas. Despite the relative smallness of the system – we study 2000 molecules of chain length $\nu = 70$ in a typical simulation experiment – features typical for the evolution of real populations are observed as there are error thresholds for replication, evolutionary steps and quasistationary sequence distributions. The relative importance of selectively neutral or almost neutral variants is discussed quantitatively. Four characteristic ensemble properties, entropy of the distribution, ensemble correlation, mean Hamming distance and diversity of the population, are computed and checked for their sensitivity in recording major optimization events during the simulation.

### 1. Molecular evolution and optimization

Conventional population genetics treats mutation as an external stochastic source. Moreover, mutations are considered as very rare events. In the absence of genetic recombination populations of haploid organisms are expected to be usually homogeneous. Experimental evidence on viral and bacterial populations is available now and it contradicts these expectations. Mutations appear much more frequently than was originally assumed.

Dedicated to Professor Manfred Eigen on the occasion of his 60th birthday.
Correspondence address: P. Schuster, Institut für theoretische Chemie und Strahlenchemie der Universität Wien, Währingerstraße 17, A-1090 Wien, Austria.

The molecular approach considers error-free replication and mutation as parallel reactions within the same mechanism. Detailed information on the molecular mechanisms of polynucleotide replication provides direct insight into the nature of mutations and their role in evolution. Several classes of mutations are properly distinguished: point mutations, deletions and insertions. Point mutations are of special importance: they represent the most frequent mutations and are easily incorporated into theoretical models of molecular evolution. This does not mean, however, that the other classes of mutations are not important in evolution. To give an example: there is a general belief that insertions leading to gene duplication played a major role in the development of present day enzyme families.

The first theoretical model of molecular evolu-

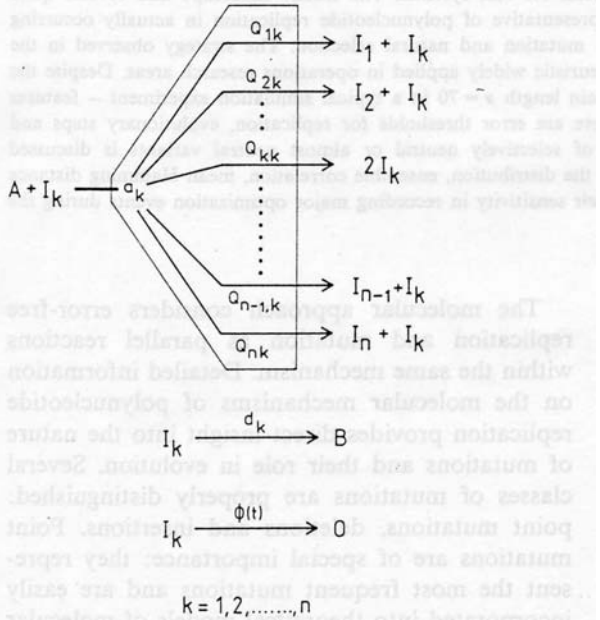



Fig. 1. The reaction network. Synthesis on template $I_k$ proceeds with the rate constant $a_k$ and leads with frequency $Q_{ik}$ to a new template $I_i$ preserving the old copy. Materials A needed for polymerization are assumed to be buffered. Degradation to waste products B occurs with rate $d_k$ and a controlled unspecific flux $\Phi(t)$ removes templates from the system.

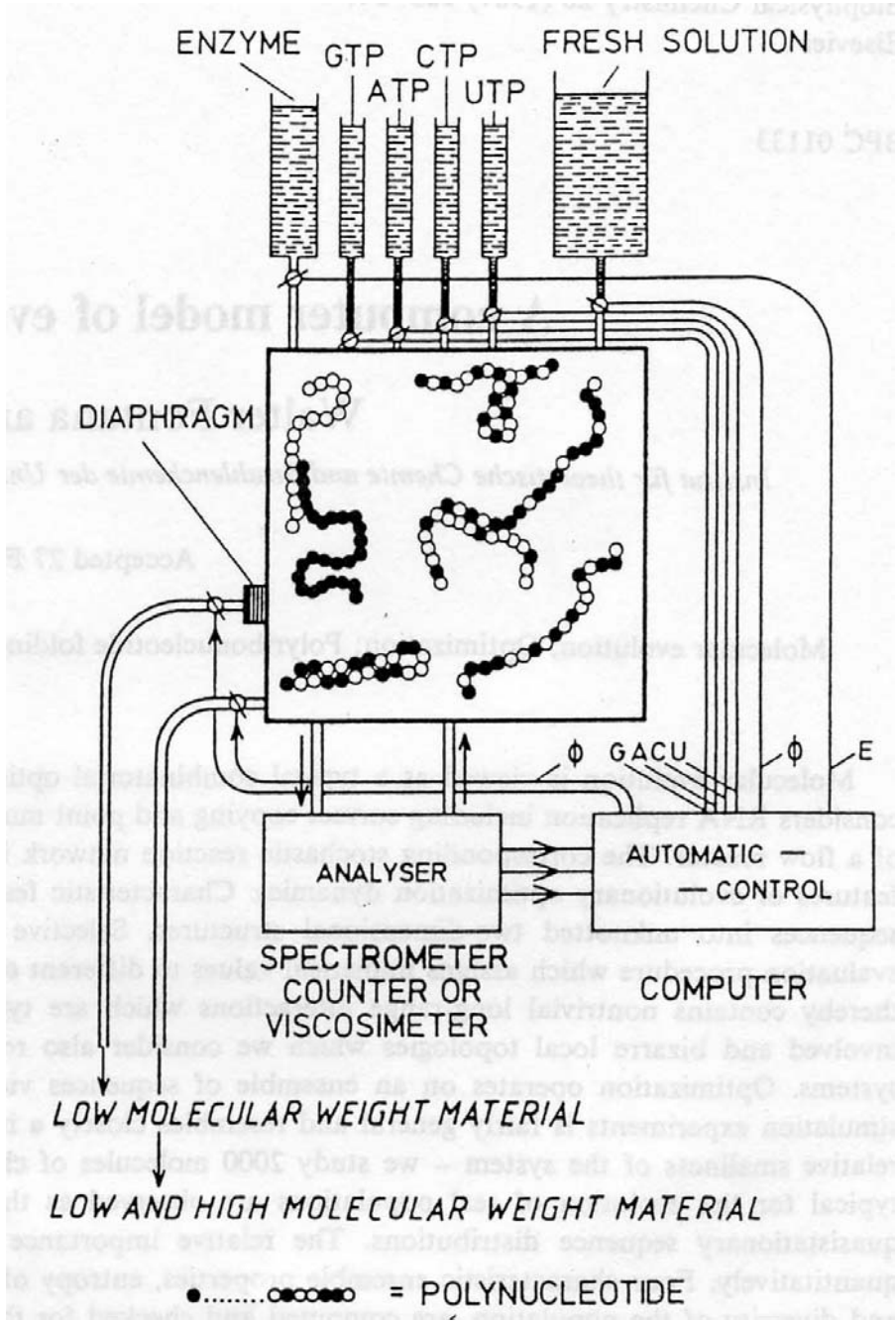
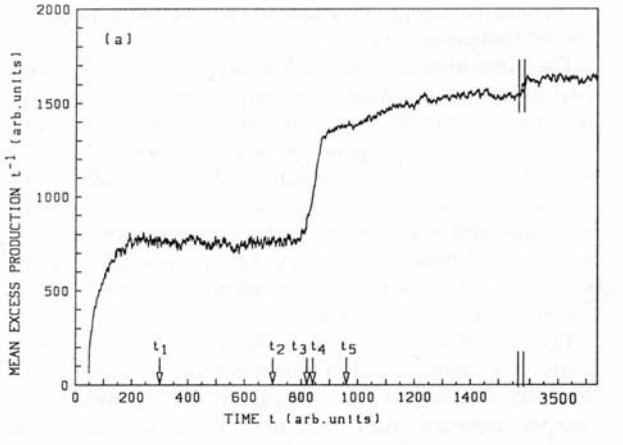
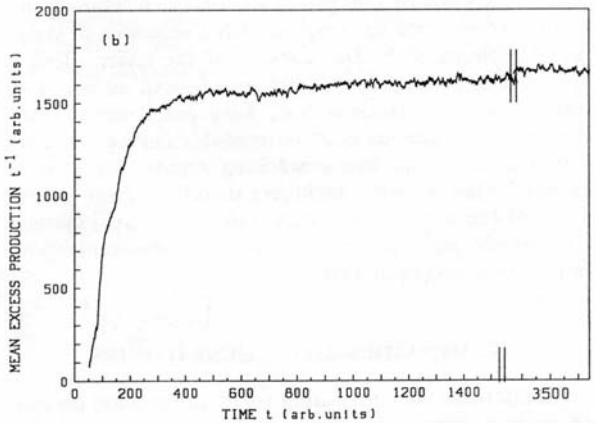
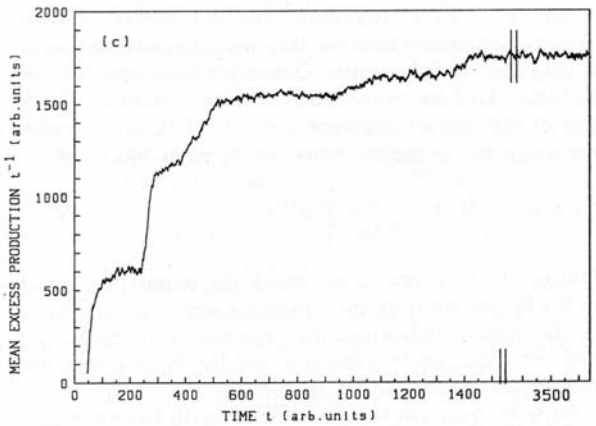
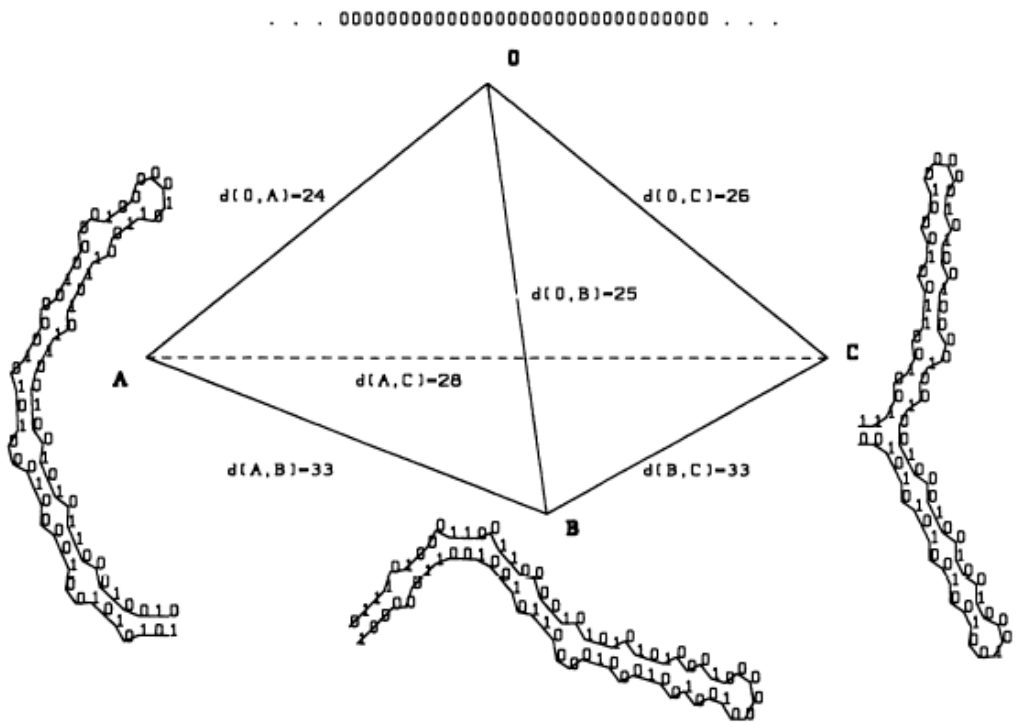



Fig. 2. The evolution reactor. This kind of flow reactor consists of a reaction vessel which allows for temperature and pressure control. Its walls are impermeable to polynucleotides. Energy-rich material is poured from the environment into the reactor. The degradation products are removed steadily. Material transport is adjusted in such a way that the concentration of monomers is constant in the reactor. A dilution flux $\Phi$ is installed in order to remove excess of polynucleotides produced by replication. Thus, the sum of the numbers of individual particles $\sum_i X_i(t) = N(t)$ may be controlled by the flux $\Phi$. Under 'constant organization' $\Phi$ is adjusted such that $N(t) = \Theta$ is essentially constant. By this we indicate that fluctuations with standard deviation $\sigma = \sqrt{N}$ occur regularly. The regulation of $\Phi$ requires internal control, which can be achieved by logistic coupling.




Walter Fontana, Wolfgang Schnabl, and
Peter Schuster, Phys.Rev.A 40:3301-3321, 1989

# Evolution *in silico*

W. Fontana, P. Schuster,
*Science* **280** (1998), 1451-1455



random individuals. The primer pair used for genomic DNA amplification is 5′-TCTCCCTGGATTCT-CATTTA-3′ (forward) and 5′-TCTTTGTCTTCTGT-TCCACC-3′ (reverse). Reactions were performed in 25 μl using 1 unit of Taq DNA polymerase with each primer at 0.4 μM; 200 μM each dATP, dTTP, dGTP, and dCTP; and PCR buffer [10 mM tris-HCl (pH 8.3), 50 mM $KCl_2$,1.5 mM $MgCl_2$] in a cycle condition of 94°C for 1 min and then 35 cycles of 94°C for 30 s, 55°C for 30 s, and 72°C for 30 s followed by 72°C for 6 min. PCR products were purified (Qiagen), digested with Xmn I, and separated in a 2% agarose gel.

32. A nonsense mutation may affect mRNA stability and result in degradation of the transcript [L. Maquat, *Am. J. Hum. Genet.* **59**, 279 (1996)].
33. Data not shown; a dot blot with poly (A)$^+$ RNA from 50 human tissues (The Human RNA Master Blot, 7770-1, Clontech Laboratories) was hybridized with a probe from exons 29 to 47 of *MYO15* using the same condition as Northern blot analysis (*13*).
34. Smith–Magenis syndrome (SMS) is due to deletions of 17p11.2 of various sizes, the smallest of which includes *MYO15* and perhaps 20 other genes [(*6*); K-S Chen, L. Potocki, J. R. Lupski, *MRDD Res. Rev.* **2**, 122 (1996)]. *MYO15* expression is easily detected in the pituitary gland (data not shown). Haploinsufficiency for *MYO15* may explain a portion of the SMS phenotype such as short stature. Moreover, a few SMS patients have sensorineural hearing loss, possibly because of a point mutation in *MYO15* in trans to the SMS 17p11.2 deletion.
35. R. A. Fridell, data not shown.
36. K. B. Avraham *et al.*, *Nature Genet.* **11**, 369 (1995); X-Z. Liu *et al.*, *ibid.* **17**, 268 (1997); F. Gibson *et al.*, *Nature* **374**, 62 (1995); D. Weil *et al.*, *ibid.*, p. 60.
37. RNA was extracted from cochlea (membranous labyrinths) obtained from human fetuses at 18 to 22 weeks of development in accordance with guidelines established by the Human Research Committee at the Brigham and Women's Hospital. Only samples without evidence of degradation were pooled for poly (A)$^+$ selection over oligo(dT) columns. First-strand cDNA was prepared using an Advantage RT-for-PCR kit (Clontech Laboratories). A portion of the first-strand cDNA (4%) was amplified by PCR with Advantage cDNA polymerase mix (Clontech Laboratories) using human *MYO15*-specific oligonucleotide primers (forward, 5′-GCATGACCTGCCGGCTAAT-GGG-3′; reverse, 5′-CTCACGGCTTCTGCATGGT-GCTCGGCTGGC-3′). Cycling conditions were 40 s at 94°C; 40 s at 66°C (3 cycles), 60°C (5 cycles), and 55°C (29 cycles); and 45 s at 68°C. PCR products were visualized by ethidium bromide staining after fractionation in a 1% agarose gel. A 688-bp PCR product is expected from amplification of the human *MYO15* cDNA. Amplification of human genomic DNA with this primer pair would result in a 2903-bp fragment.

38. We are grateful to the people of Bengkala, Bali, and the two families from India. We thank J. R. Lupski and K.-S. Chen for providing the human chromosome 17 cosmid library. For technical and computational assistance, we thank N. Dietrich, M. Fergusson, A. Gupta, E. Sorbello, R. Torkzadeh, C. Varner, M. Walker, G. Bouffard, and S. Beckstrom-Sternberg (National Institutes of Health Intramural Sequencing Center). We thank J. T. Hinnant, I. N. Arhya, and S. Winata for assistance in Bali, and T. Barber, S. Sullivan, E. Green, D. Drayna, and J. Battey for helpful comments on this manuscript. Supported by the National Institute on Deafness and Other Communication Disorders (NIDCD) (Z01 DC 00035-01 and Z01 DC 00038-01 to T.B.F. and E.R.W. and R01 DC 03402 to C.C.M.), the National Institute of Child Health and Human Development (R01 HD30428 to S.A.C.) and a National Science Foundation Graduate Research Fellowship to F.J.P. This paper is dedicated to J. B. Snow Jr. on his retirement as the Director of the NIDCD.

9 March 1998; accepted 17 April 1998

## Continuity in Evolution: On the Nature of Transitions

Walter Fontana and Peter Schuster

To distinguish continuous from discontinuous evolutionary change, a relation of nearness between phenotypes is needed. Such a relation is based on the probability of one phenotype being accessible from another through changes in the genotype. This nearness relation is exemplified by calculating the shape neighborhood of a transfer RNA secondary structure and provides a characterization of discontinuous shape transformations in RNA. The simulation of replicating and mutating RNA populations under selection shows that sudden adaptive progress coincides mostly, but not always, with discontinuous shape transformations. The nature of these transformations illuminates the key role of neutral genetic drift in their realization.

A much-debated issue in evolutionary biology concerns the extent to which the history of life has proceeded gradually or has been punctuated by discontinuous transitions at the level of phenotypes (*1*). Our goal is to make the notion of a discontinuous transition more precise and to understand how it arises in a model of evolutionary adaptation.

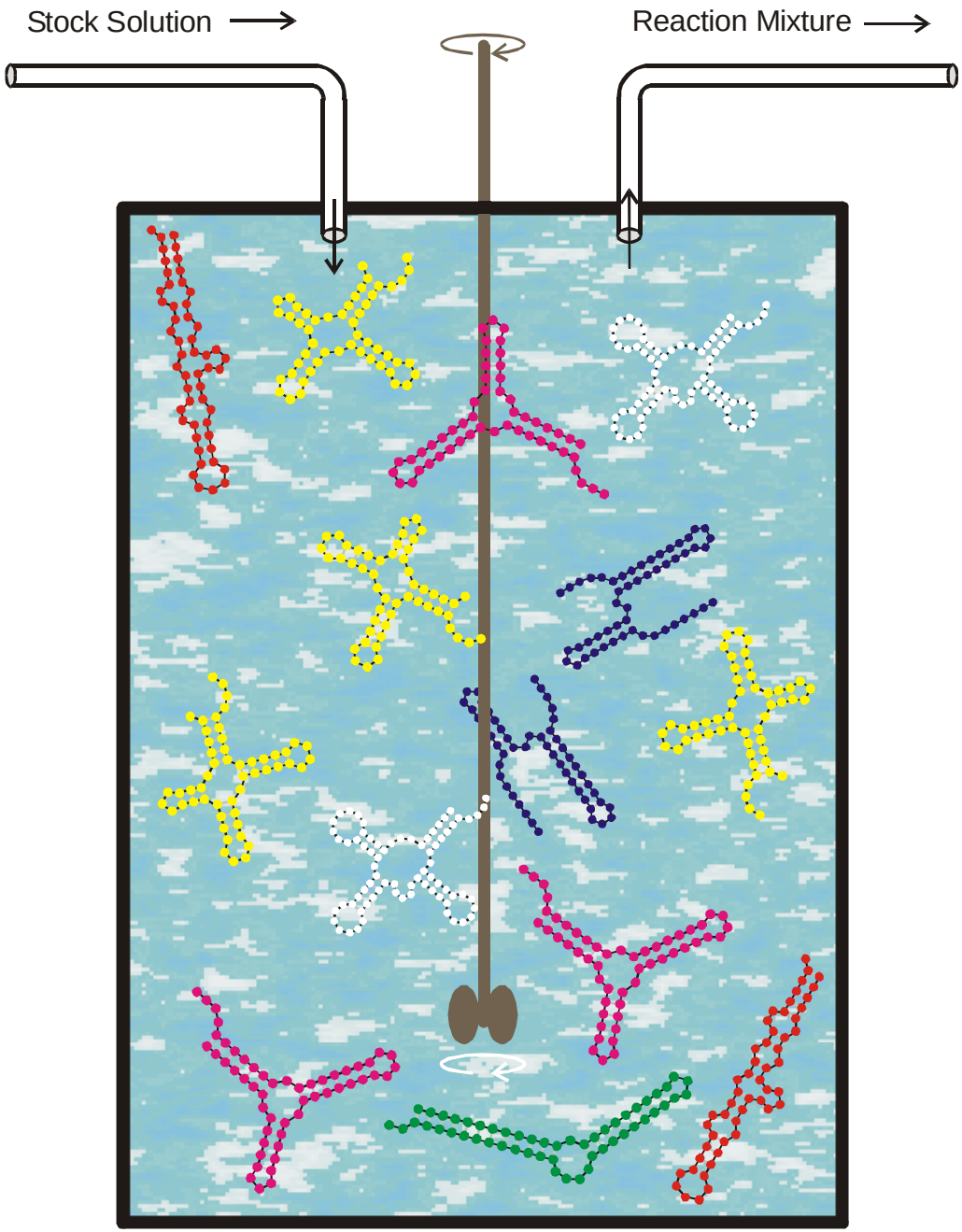
We focus on the narrow domain of RNA secondary structure, which is currently the simplest computationally tractable, yet realistic phenotype (*2*). This choice enables the definition and exploration of concepts that may prove useful in a wider context. RNA secondary structures represent a coarse level of analysis compared with the three-dimensional structure at atomic resolution. Yet, secondary structures are empirically well defined and obtain their biophysical and biochemical importance from being a scaffold for the tertiary structure. For the sake of brevity, we shall refer to secondary structures as "shapes." RNA combines in a single molecule both genotype (replicatable sequence) and phenotype (selectable shape), making it ideally suited for in vitro evolution experiments (*3*, *4*).

To generate evolutionary histories, we used a stochastic continuous time model of an RNA population replicating and mutating in a capacity-constrained flow reactor under selection (*5*, *6*). In the laboratory, a goal might be to find an RNA aptamer binding specifically to a molecule (*4*). Although in the experiment the evolutionary end product was unknown, we thought of its shape as being specified implicitly by the imposed selection criterion. Because our intent is to study evolutionary histories rather than end products, we defined a target shape in advance and assumed the replication rate of a sequence to be a function of the similarity between its shape and the target. An actual situation may involve more than one best shape, but this does not affect our conclusions.

An instance representing in its qualitative features all the simulations we performed is shown in Fig. 1A. Starting with identical sequences folding into a random shape, the simulation was stopped when the population became dominated by the target, here a canonical tRNA shape. The black curve traces the average distance to the target (inversely related to fitness) in the population against time. Aside from a short initial phase, the entire history is dominated by steps, that is, flat periods of no apparent adaptive progress, interrupted by sudden approaches toward the target structure (*7*). However, the dominant shapes in the population not only change at these marked events but undergo several fitness-neutral transformations during the periods of no apparent progress. Although discontinuities in the fitness trace are evident, it is entirely unclear when and on the basis of what the series of successive phenotypes itself can be called continuous or discontinuous.

A set of entities is organized into a (topological) space by assigning to each entity a system of neighborhoods. In the present case, there are two kinds of entities: sequences and shapes, which are related by a thermodynamic folding procedure. The set of possible sequences (of fixed length) is naturally organized into a space because point mutations induce a canonical neighborhood. The neighborhood of a sequence consists of all its one-error mutants. The problem is how to organize the set of possible shapes into a space. The issue arises because, in contrast to sequences, there are

Institut für Theoretische Chemie, Universität Wien, Währingerstrasse 17, A-1090 Wien, Austria, Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA, and International Institute for Applied Systems Analysis (IIASA), A-2361 Laxenburg, Austria.




Replication rate constant:

$$f_k = \gamma / [\alpha + \Delta d_S{}^{(k)}]$$

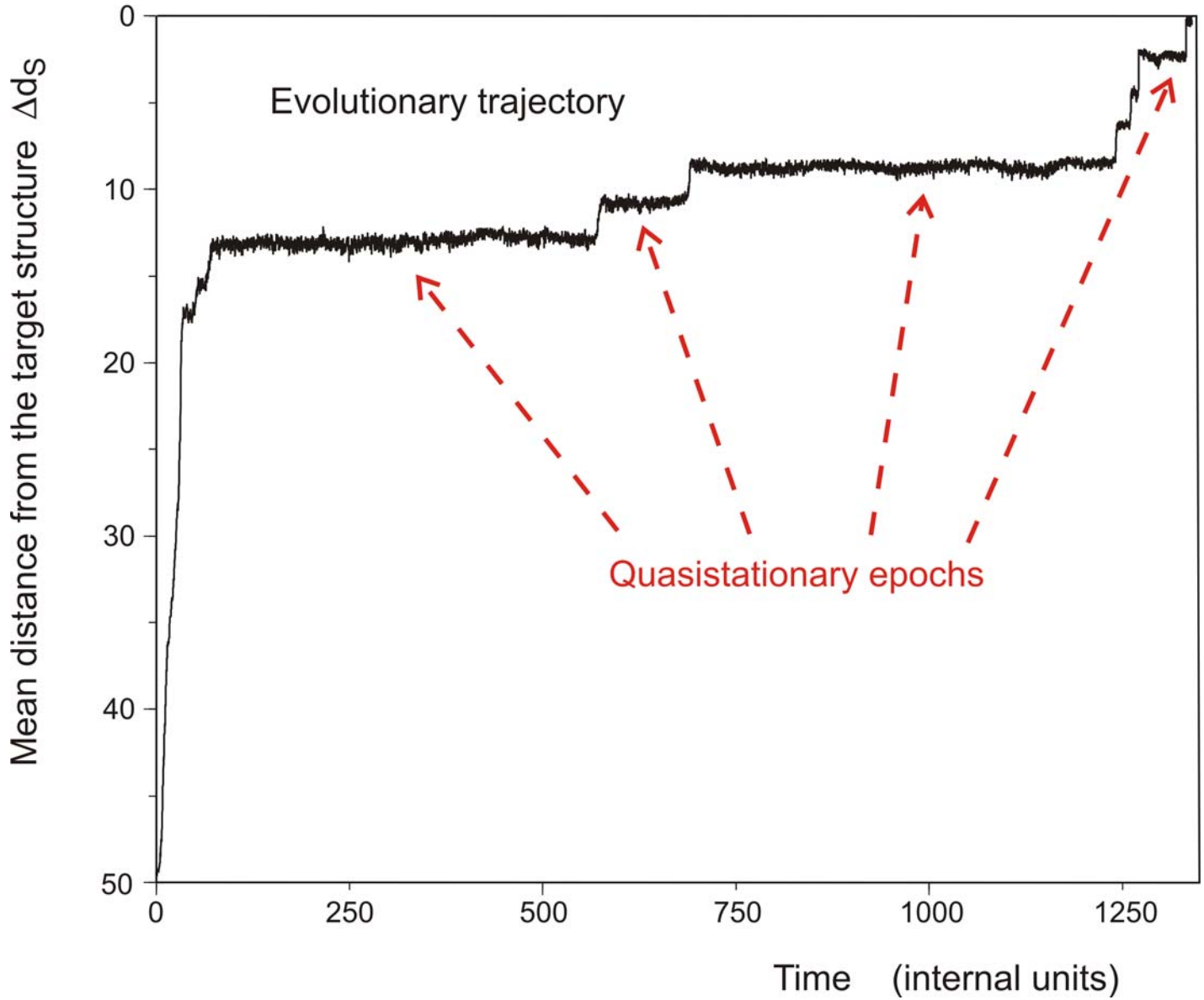
$$\Delta d_S{}^{(k)} = d_H(S_k, S_\tau)$$

Selection constraint:

# RNA molecules is controlled by the flow

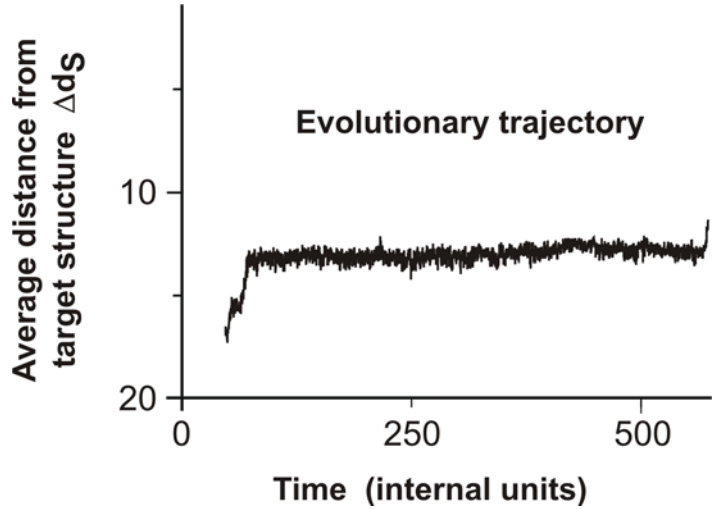
$$N(t) \approx \overline{N} \pm \sqrt{\overline{N}}$$

The flowreactor as a device for studies of evolution ***in vitro*** and ***in silico***




***In silico*** optimization in the flow reactor: Evolutionary Trajectory

**28 neutral point mutations during a long quasi-stationary epoch**




```
entry   GGUAUGGGCGUUGAAUAGUAGGGUUUAAACCAAUCGGCCAACGAUCUCGUGUGCGCAUUUCAUAUCCCGUACAGAA
8       .((((((((((((.........(((....)))........)))))).....((((.........))))))))))....
exit    GGUAUGGGCGUUGAAUAAUAGGGUUUAAACCAAUCGGCCAACGAUCUCGUGUGCGCAUUUCAUAUCCCAUACAGAA
entry   GGUAUGGGCGUUGAAUAAUAGGGUUUAAACCAAUCGGCCAACGAUCUCGUGUGCGCAUUUCAUAUACCAUACAGAA
9       .((((((.((((.........(((....)))........)))))).....((((.........))))).))))))....
exit    UGGAUGGACGUUGAAUAACAAGGUAUCGACCAAACAACCAACGAGUAAGUGUGUACGCCCCACACACCGUCCCAAG
entry   UGGAUGGACGUUGAAUAACAAGGUAUCGACCAAACAACCAACGAGUAAGUGUGUACGCCCCACACAGCGUCCCAAG
10      .(((((..((((.........(((....)))........)))))).....((((.........))))).)))))....
exit    UGGAUGGACGUUGAAUAACAAGGUAUCGACCAAACAACCAACGAGUAAGUGUGUACGCCCCACACAGCGUCCCAAG
```

**Transition inducing point mutations change the molecular structure**

**Neutral point mutations leave the molecular structure unchanged**

**Neutral genotype evolution** during phenotypic stasis

*„...Variations neither useful not injurious would not be affected by natural selection, and would be left either a fluctuating element, as perhaps we see in certain polymorphic species, or would ultimately become fixed, owing to the nature of the organism and the nature of the conditions. ...“*

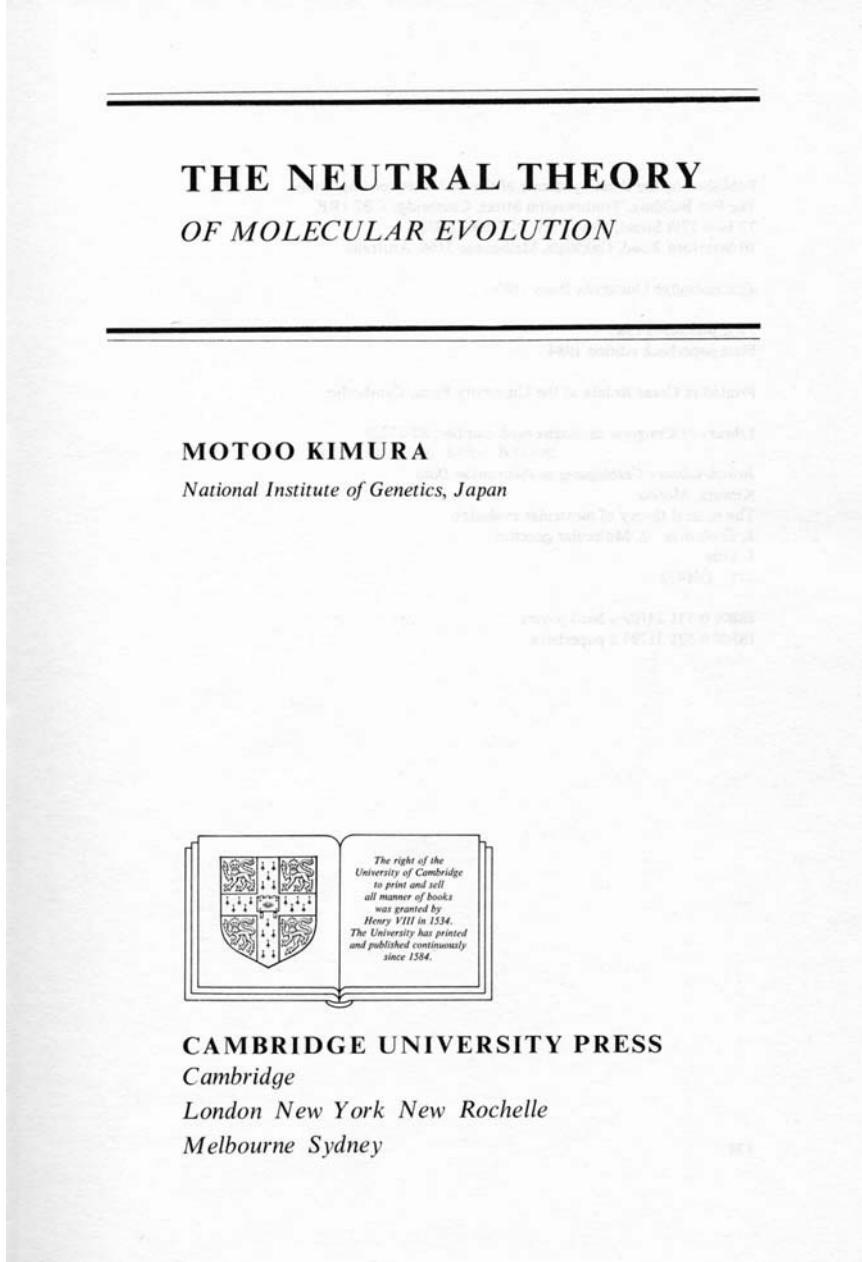
Charles Darwin, Origin of species (1859)

MOTOO KIMURA

Motoo Kimura's Population genetics of neutral evolution.

Evolutionary rate at the molecular level. *Nature* **217**: 624-626, 1955.

*The Neutral Theory of Molecular Evolution*. Cambridge University Press. Cambridge, UK, 1983.



THE NEUTRAL THEORY

*OF MOLECULAR EVOLUTION*

MOTOO KIMURA

*National Institute of Genetics, Japan*

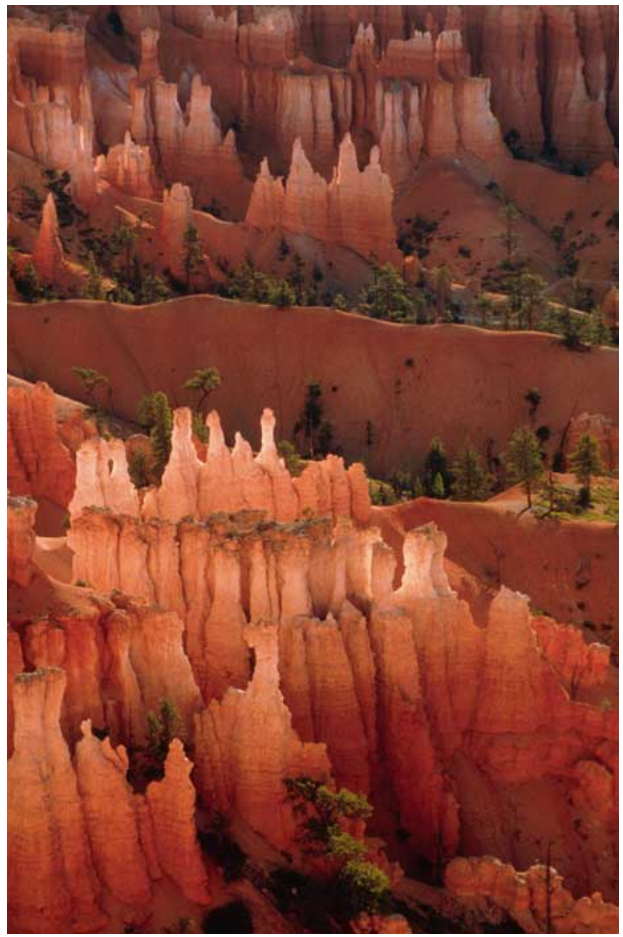


CAMBRIDGE UNIVERSITY PRESS
*Cambridge*
*London New York New Rochelle*
*Melbourne Sydney*

Mount Fuji

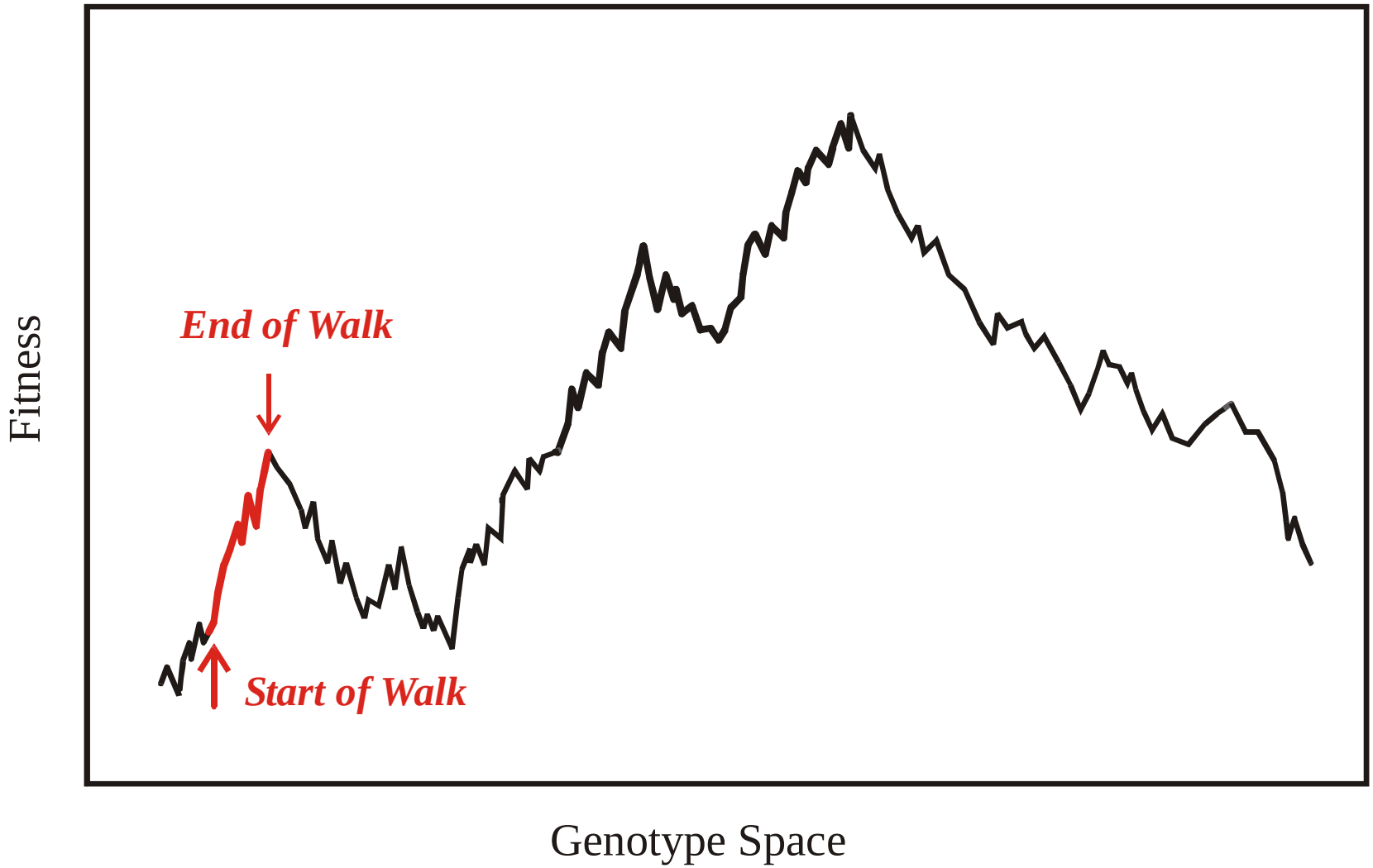
Example of  a smooth landscape on Earth

**Dolomites**

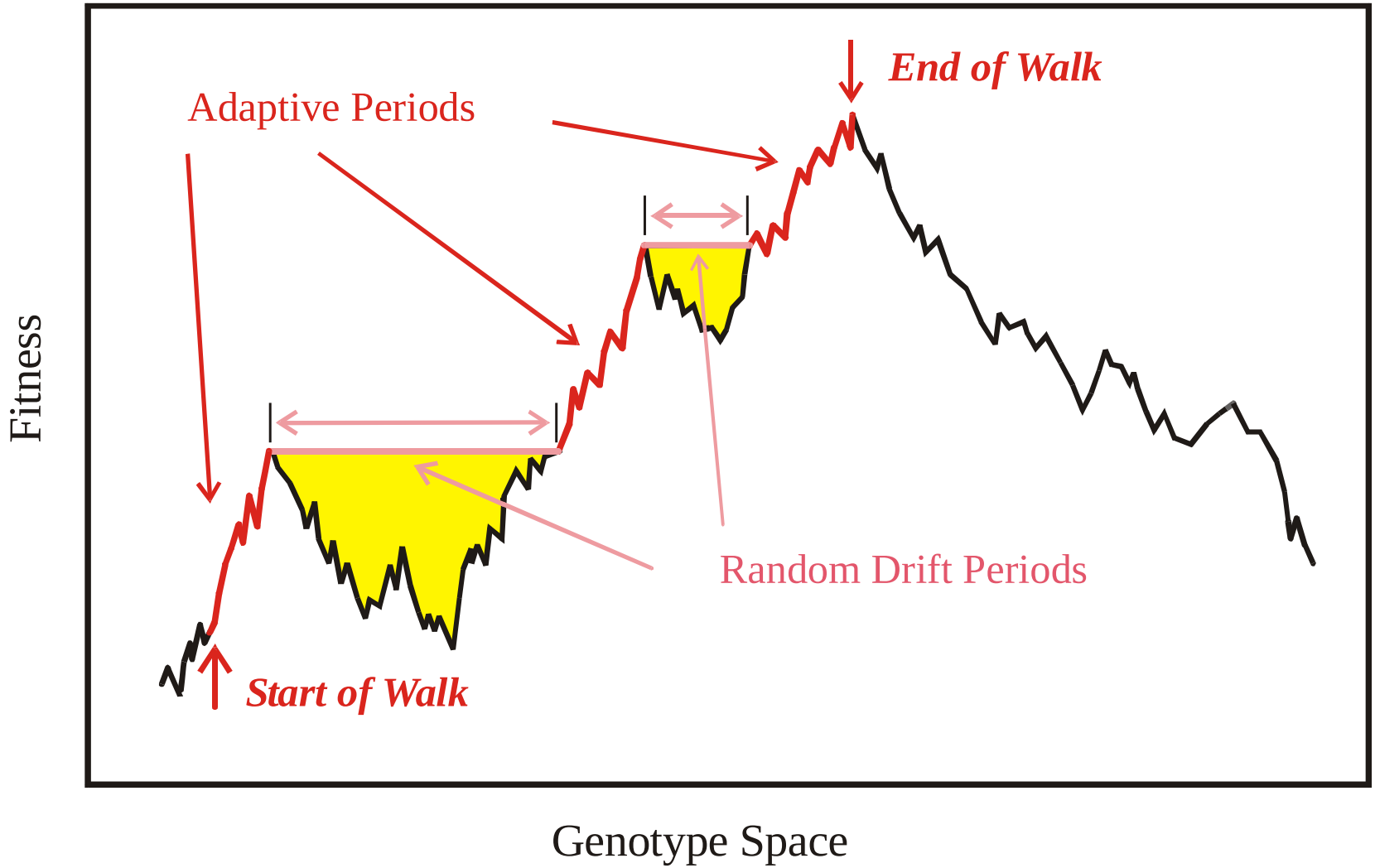


**Bryce Canyon**

Examples of rugged landscapes on Earth




Evolutionary optimization in absence of neutral paths in sequence space

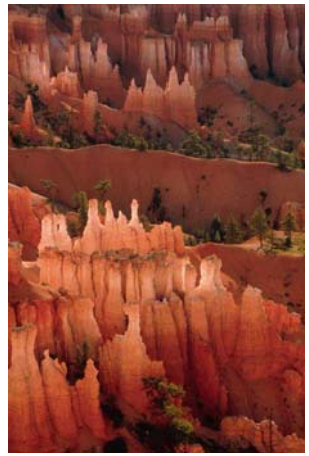



Evolutionary optimization including neutral paths in sequence space

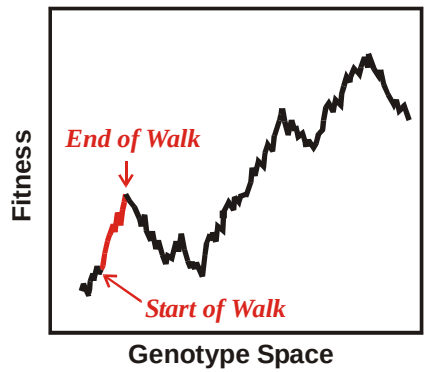
Grand Canyon

Example of a landscape on Earth with ‘neutral’ ridges and plateaus

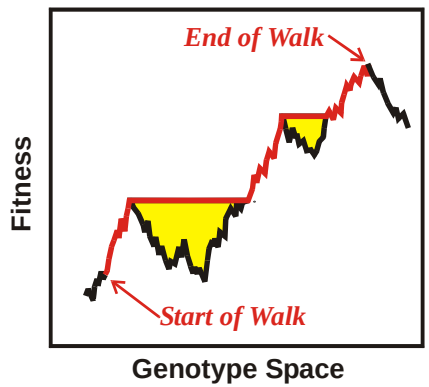
- Conformational and mutational landscapes of biomolecules as well as fitness landscapes of evolutionary biology are **rugged**.



- **Adaptive** or **non-descending walks** on rugged landscapes end commonly at one of the low lying local maxima.




- Selective neutrality in the form of **neutral networks** plays an active role in evolutionary optimization and enables populations to reach high local maxima or even the global optimum.



1. History of evolutionary thinking
2. Probabilities in biology
3. Complex patterns from simple rules
4. Mechanisms of evolution
5. **Origins of complexity – The eye**

Walter Gehring, Biozentrum, Universität Basel

Molecular genetics shows that the development of all different forms of eyes have the same evolutionary origin, which can be traced back to a simple form of light-sensitivity found already in primitive bacteria.

W. J. Gehring. The genetic control of eye development and its implications for the evolution of the various eye-types. *Zoology* **104** (2001):171-183

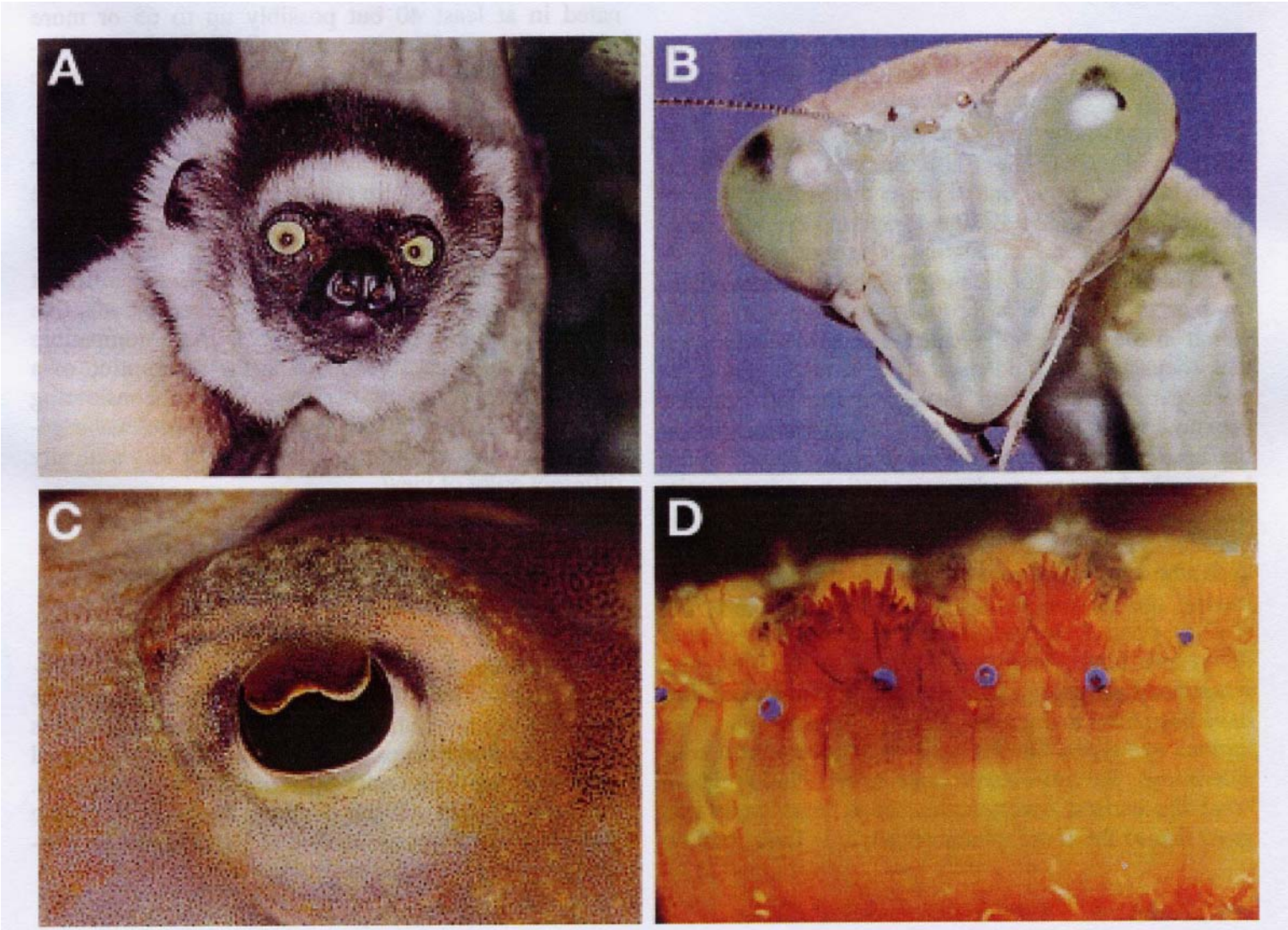



**Fig. 1.** Different types of eyes. **(A)** Camera-type eye from the Lemur *Propithecus verrauxi.* **(B)** Compound eye of the praying *Mantis.* **(C)** Camera-type eye from the Cephalopod *Sepia erostrata.* **(D)** Mirror eye from the clam *Chlamys nobilis.* (Courtesy of Dr. Kazuto Kato; photographs kindly provided by Masahiro Iijima, Susumu Yamaguchi and Isamu Soyama).

Walter J. Gehring, The genetic control of eye development and its implications for the evolution of the various eye-types. Zoology 104 (2001), 171-183

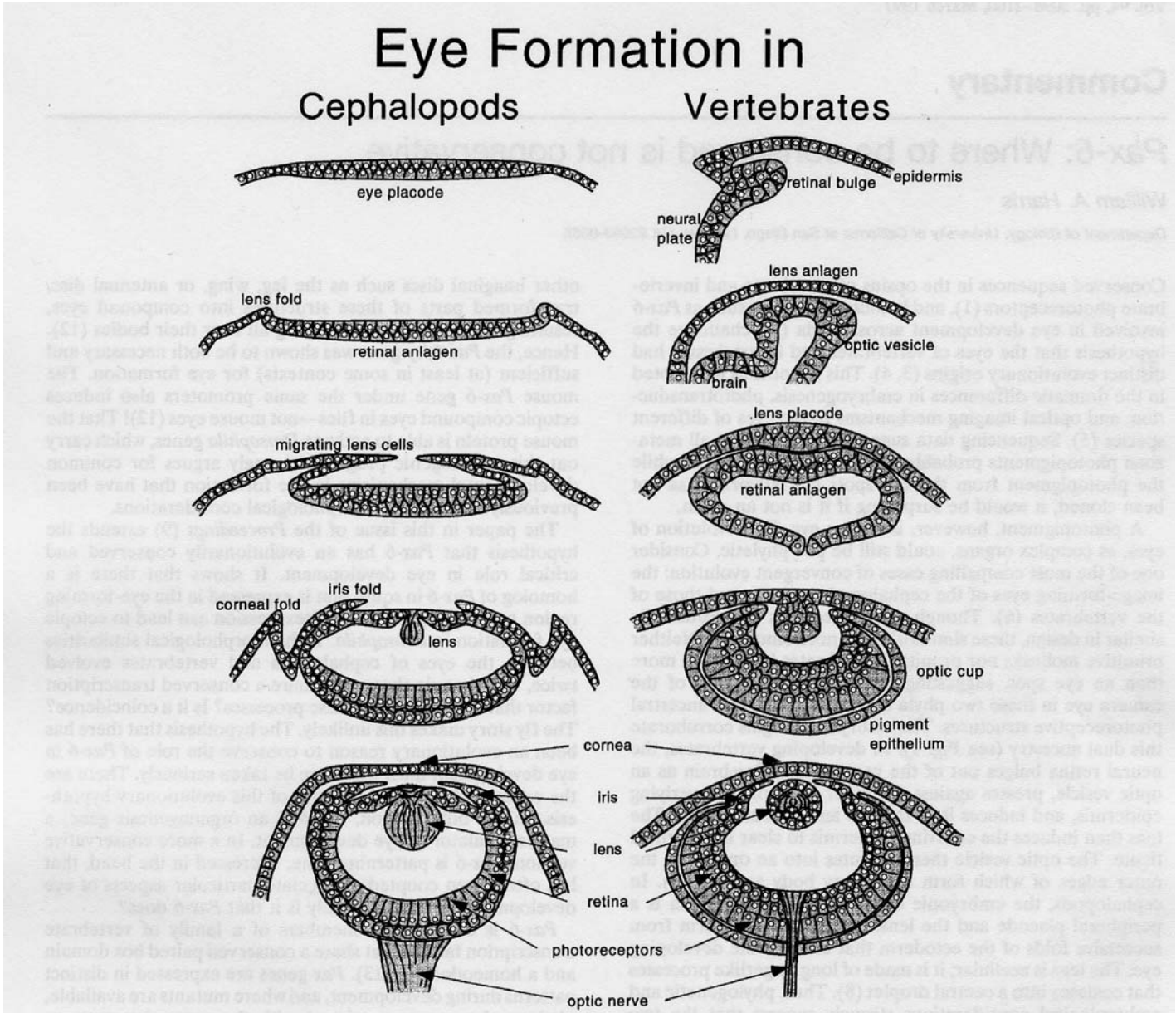
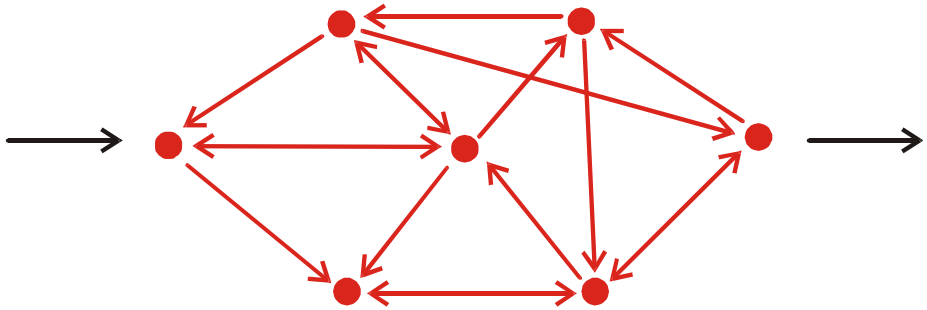



FIG. 1. Schematic diagram of cephalopod eye development (*Left*) and vertebrate eye development (*Right*) as explained in more detail in refs. 7 and 8. Development proceeds from top to bottom. Even though the adult structures are fairly similar, excepting certain obvious features such as the placement of the photoreceptors and lentigenic cells, the development is very different. The cephalopod eye forms from an epidermal placode through a series of successive infoldings, while the vertebrate eye emerges from the neural plate and induces the overlying epidermis to form the lens.

**William A. Harris,**
**Proc.Natl.Acad.Sci.USA**
**94 (1997), 2098-2100**

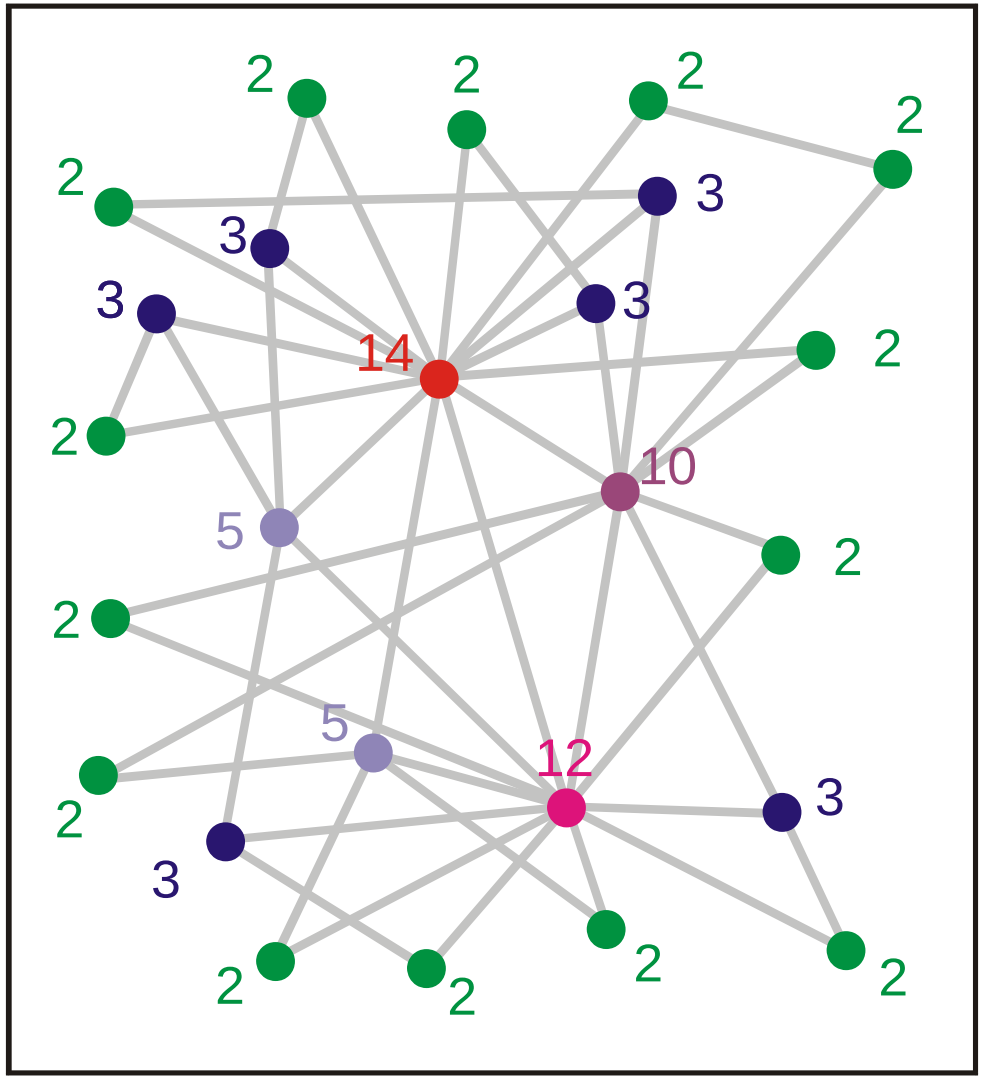
Linear chain



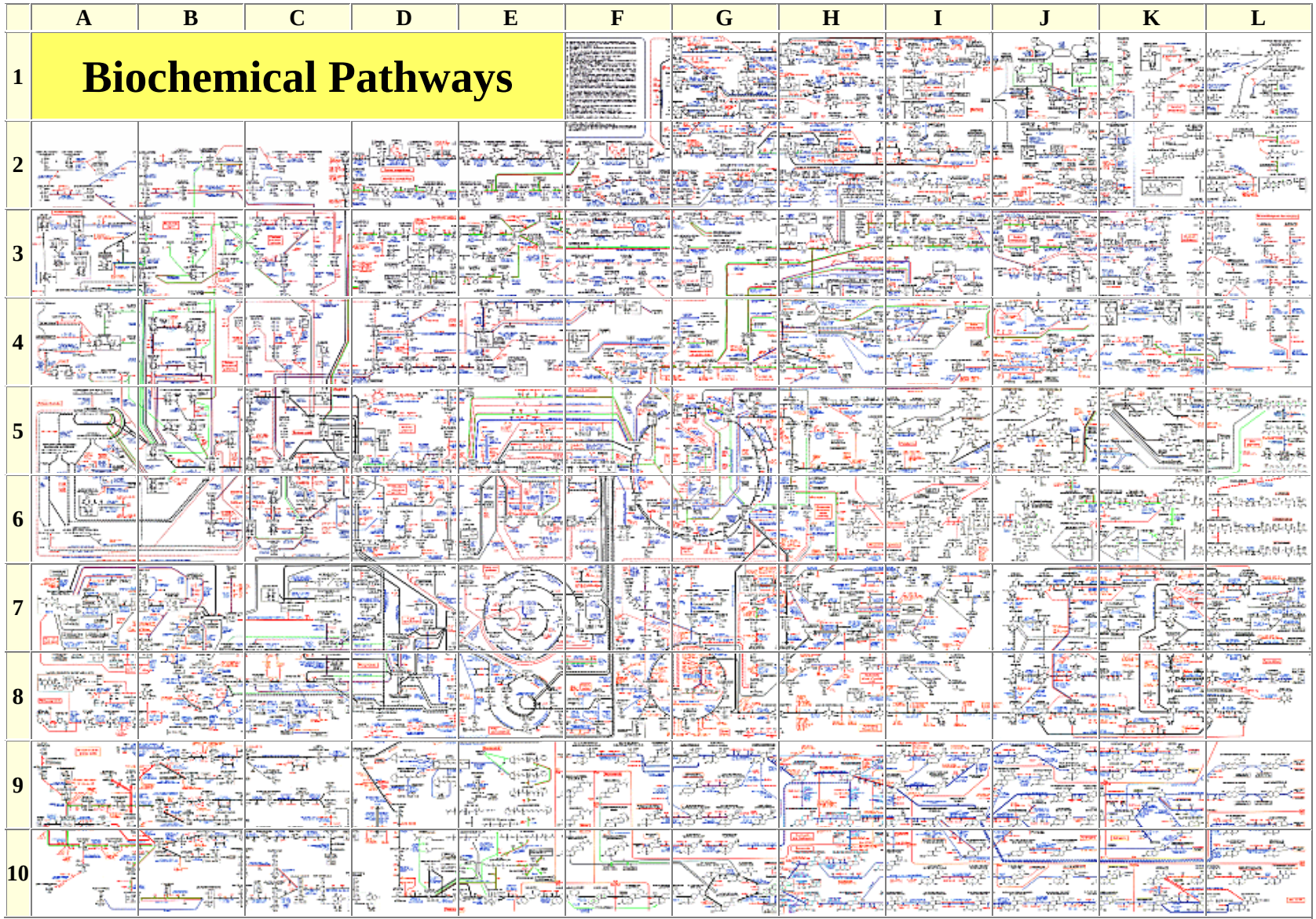
Network

Processing of information in cascades and networks

| links | # nodes |
|---|---|
| 2 | 14 |
| 3 | 6 |
| 5 | 2 |
| 10 | 1 |
| 12 | 1 |
| 14 | 1 |

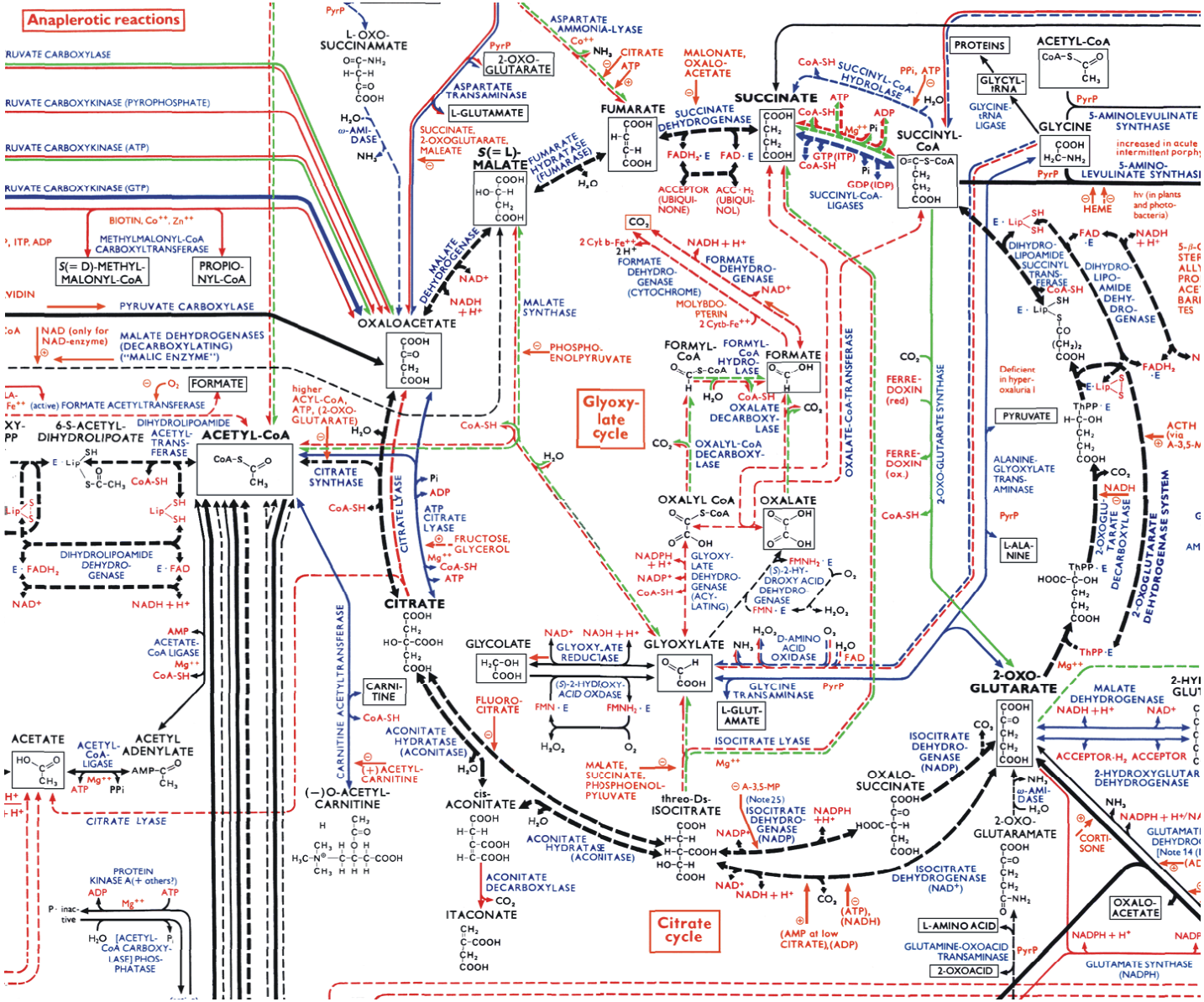


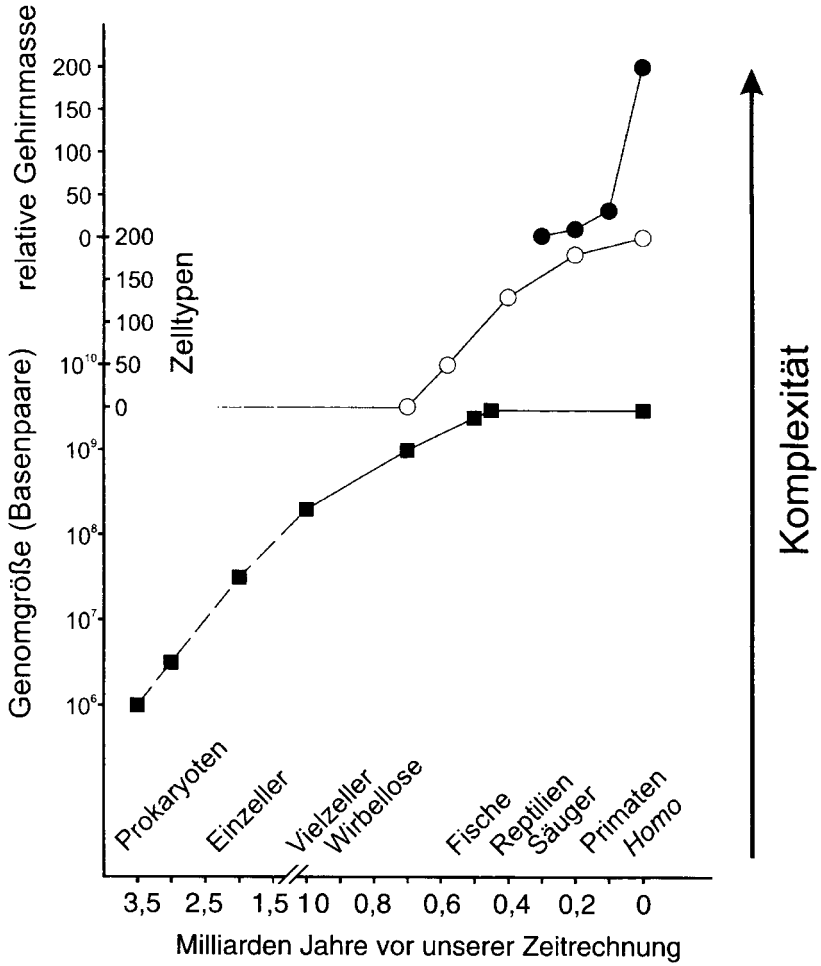

Analysis of nodes and links in a step by step evolved network




The reaction network of cellular metabolism published by Boehringer-Ingelheim.

The citric acid or Krebs cycle (enlarged from previous slide).






**4.10** Die Zunahme der Komplexität ist ein wesentlicher Aspekt der biologischen Evolution, wobei höhere Komplexität sowohl durch Vergrößerung der Zahl von miteinander in Wechselwirkung stehenden Elementen als auch durch Differenzierung der Funktionen dieser Elemente entstehen kann. In dieser Abbildung wird zwischen drei Phasen oder Strategien der Evolution von Komplexität unterschieden. *Untere Kurve*: Zunahme der Genomgröße; logarithmische Auftragung der Zahl der Basenpaare im Genom von Zellen seit Beginn der biologischen Evolution (Daten aus Abbildung 2.3). *Mittlere Kurve*: Zunahme der Zahl der Zelltypen in der Evolution der Metazoa (Daten aus Abbildung 4.8). *Obere Kurve*: Zunahme des relativen Gehirngewichts (bezogen auf die Körperoberfläche) bei Säugetieren (Daten aus Wilson 1985). Für die Abszisse wurden zwei Skaleneinteilungen verwendet, eine für den Zeitraum $>10^9$ Jahre, eine andere für den Zeitraum $<10^9$ Jahre vor der Gegenwart. Oberhalb der Abszisse sind die Namen einiger wichtiger taxonomischer Einheiten angeführt, deren Evolution in etwa beim jeweiligen Wortbeginn einsetzt.

Wolfgang Wieser. Die Erfindung der Individualität oder die zwei Gesichter der Evolution. Spektrum Akademischer Verlag, Heidelberg 1998.

A.C.Wilson. The Molecular Basis of Evolution. Scientific American, Oct.1985, 164-173.

Web-Page for further information:

http://www.tbi.univie.ac.at/~pks