Evolution

Vom Ursprung des Lebens bis zur Gegenwart

Peter Schuster

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

150 Jahre Evolutionstheorie – Darwin und die Folgen

Klagenfurt, 25.– 27.06.2009

Web-Page for further information:

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

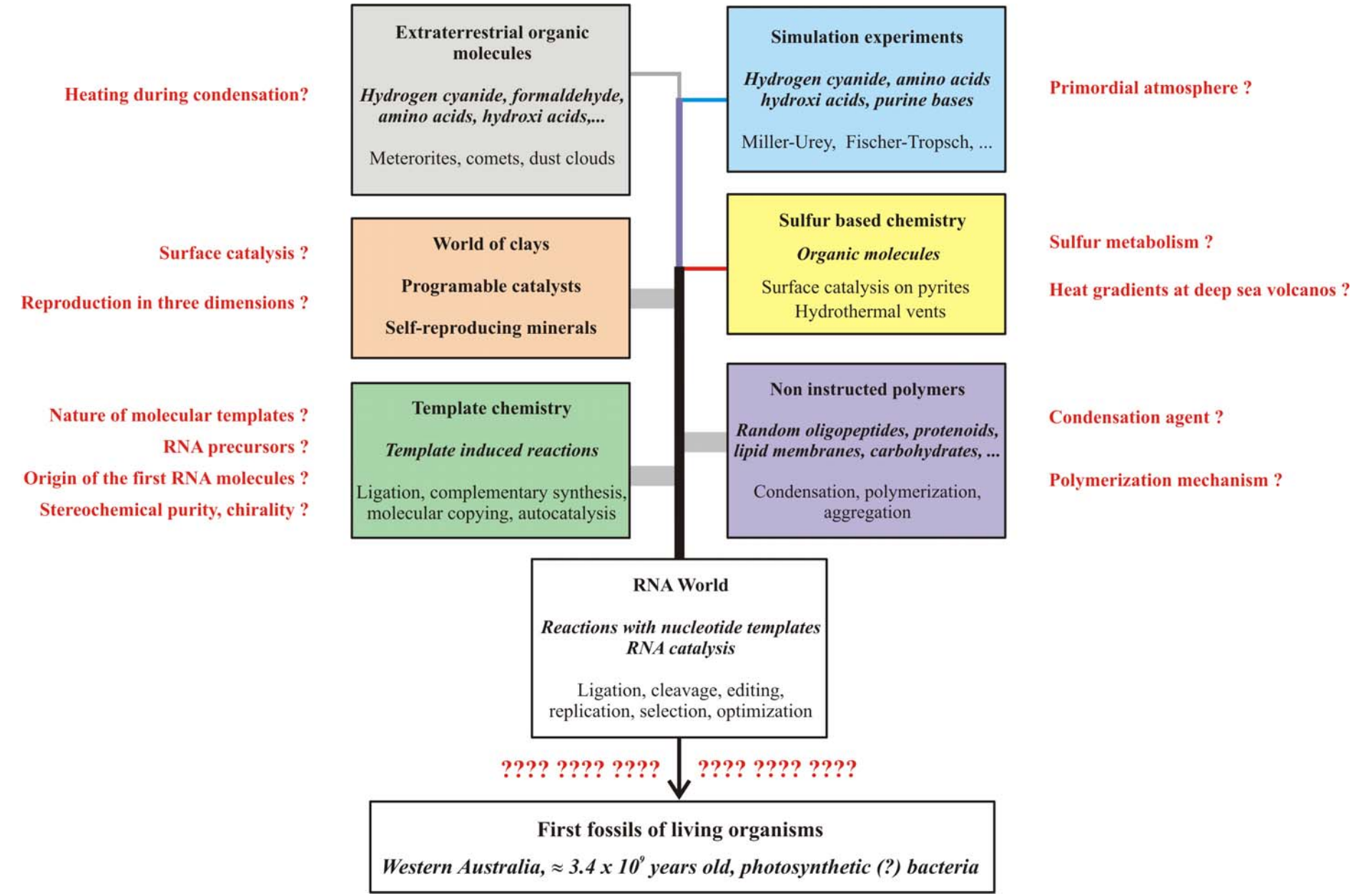
1. Was ist Leben?
2. Chemische Evolution
3. Darwins tiefe Einsichten
4. Der Ursprung biologischer Information
5. Darwinsche Evolution mit Molekülen
6. Evolutionäre Biotechnologie
7. Die DNA + Protein Welt
8. Evolution bis zum Menschen

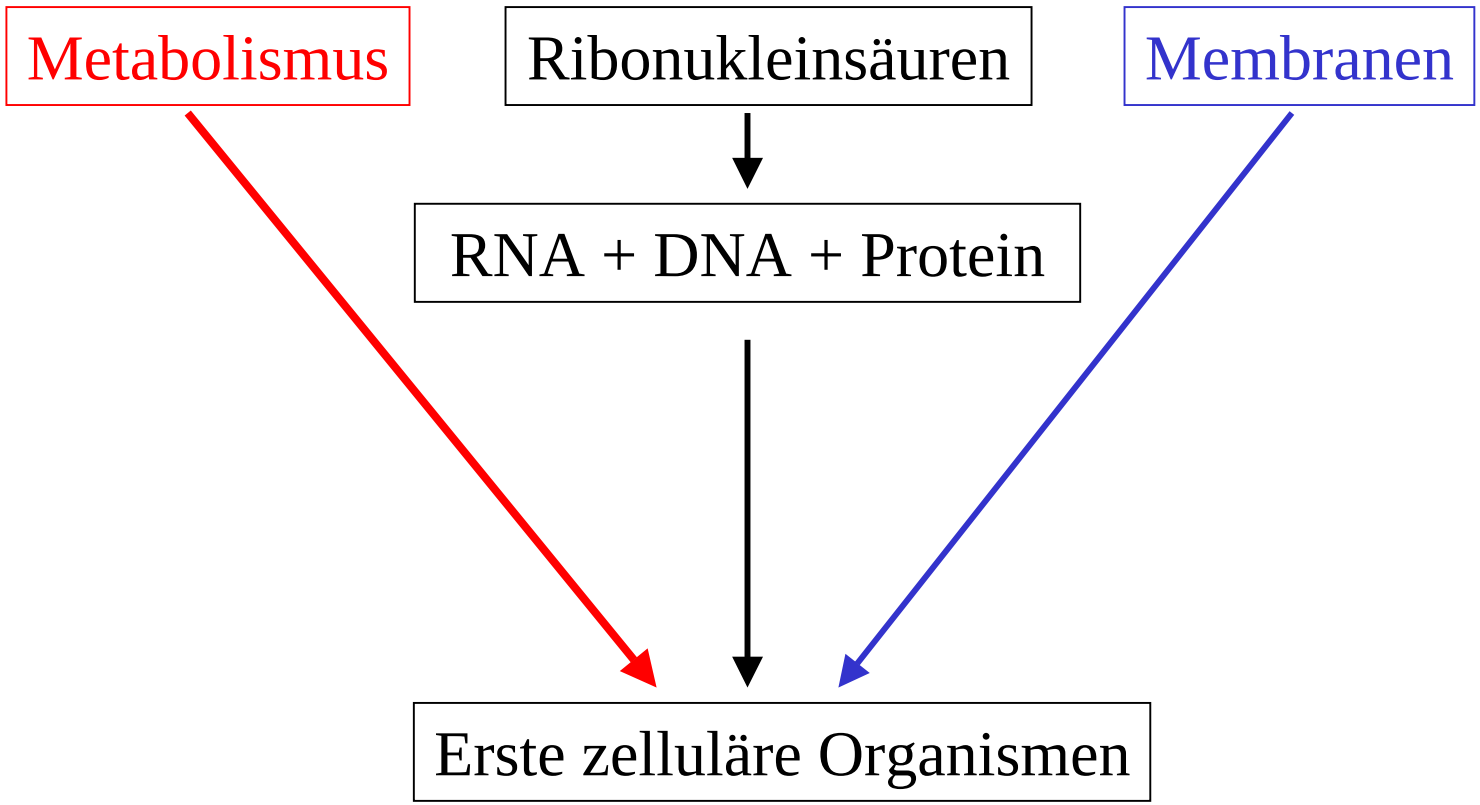
1. **Was ist Leben?**
2. Chemische Evolution
3. Darwins tiefe Einsichten
4. Der Ursprung biologischer Information
5. Darwinsche Evolution mit Molekülen
6. Evolutionäre Biotechnologie
7. Die DNA + Protein Welt
8. Evolution bis zum Menschen

Kriterien zur Unterscheidung zwischen belebter und unbelebter Materie

1. Befähigung zu **Vermehrung** und **Variation**
2. Befähigung zu **Darwinscher Evolution** durch **Variation** und **Selektion**
3. Befähigung zu **universeller Mutation** durch **digitalisierte Information**
4. **Abgrenzung** durch **Membranen**, **Zellwände**, **Häute** oder **Verhalten**
5. **Stoffwechsel** zur Erzeugung der **Bausteine** von **Biomolekülen**
6. **Autopoiese** als Selbsterhalt mit Hilfe des **Stoffwechsels**
7. **Arbeitsteilung** durch **Zelldifferenzierung**
8. **Befähigung** zum **individuellen Lernen**
9. **Übertragung erworbener Eigenschaften** durch **Erziehung**
10. **Sprache**, **Schrift** und **Kultur**

1. Was ist Leben?
2. **Chemische Evolution**
3. Darwins tiefe Einsichten
4. Der Ursprung biologischer Information
5. Darwinsche Evolution mit Molekülen
6. Evolutionäre Biotechnologie
7. Die DNA + Protein Welt
8. Evolution bis zum Menschen


Extraterrestrial organic molecules
Hydrogen cyanide, formaldehyde, amino acids, hydroxi acids,...
Meterorites, comets, dust clouds
Heating during condensation?
Simulation experiments
Hydrogen cyanide, amino acids hydroxi acids, purine bases
Miller-Urey, Fischer-Tropsch, ...
Primordial atmosphere ?
World of clays
Programable catalysts
Self-reproducing minerals
Surface catalysis ?
Reproduction in three dimensions ?
Sulfur based chemistry
Organic molecules
Surface catalysis on pyrites
Hydrothermal vents
Sulfur metabolism ?
Heat gradients at deep sea volcanos ?
Template chemistry
Template induced reactions
Ligation, complementary synthesis, molecular copying, autocatalysis
Nature of molecular templates ?
RNA precursors ?
Origin of the first RNA molecules ?
Stereochemical purity, chirality ?
Non instructed polymers
Random oligopeptides, protenoids, lipid membranes, carbohydrates, ...
Condensation, polymerization, aggregation
Condensation agent ?
Polymerization mechanism ?
RNA World
Reactions with nucleotide templates
RNA catalysis
Ligation, cleavage, editing, replication, selection, optimization
???? ???? ????
???? ???? ????
First fossils of living organisms
Western Australia, ≈ 3.4 x 10⁹ years old, photosynthetic (?) bacteria


Metabolismus
Ribonukleinsäuren
Membranen
RNA + DNA + Protein
Erste zelluläre Organismen

Sag niemals nie !

(i) Synthese von **Pyrimidinnukleotiden** durch einfache Reaktionen

(ii) Synthese von optisch reinen **chiralen Antipoden**

(iii) **Darwinsche Evolution** im Reagensglas **ohne Enzymkatalyse**

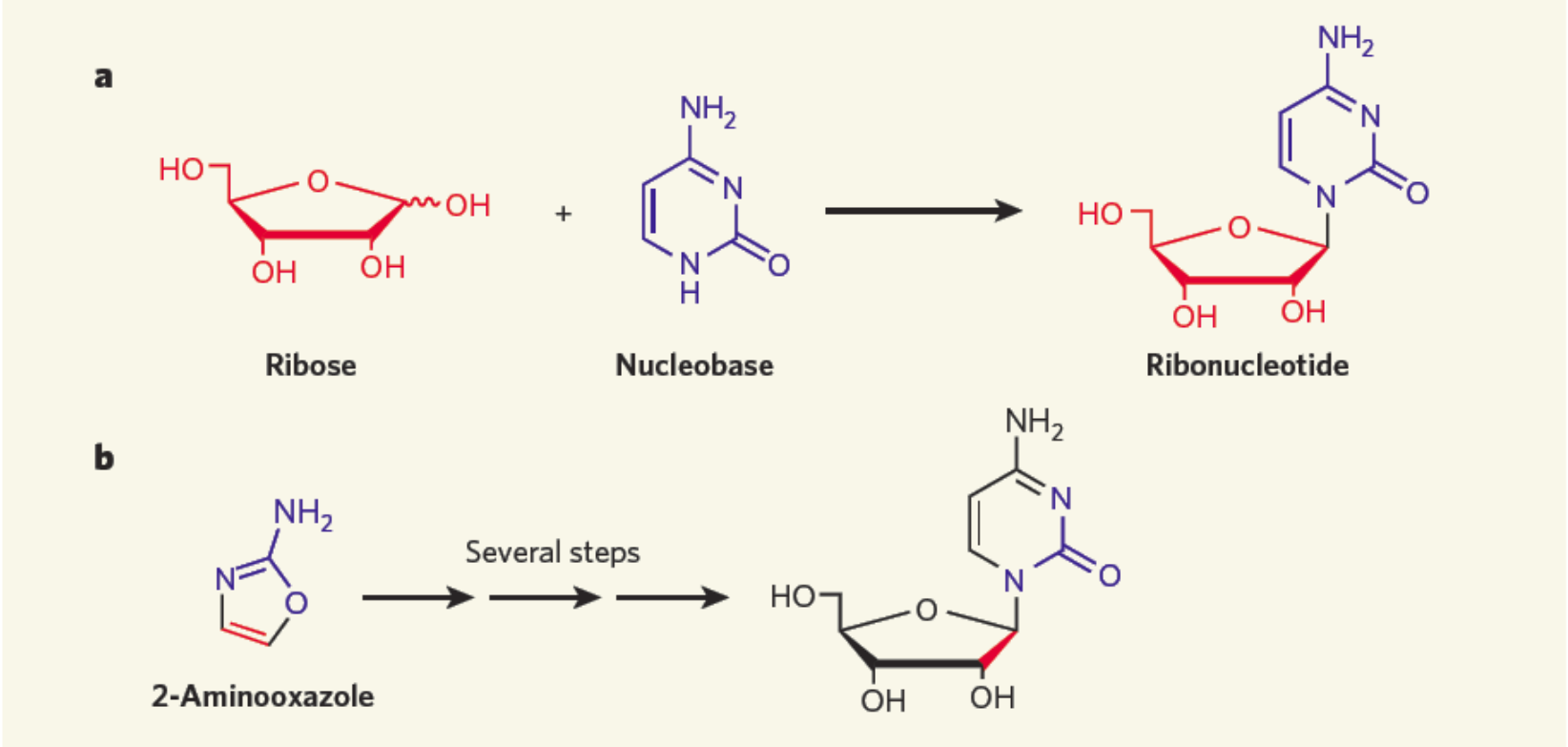
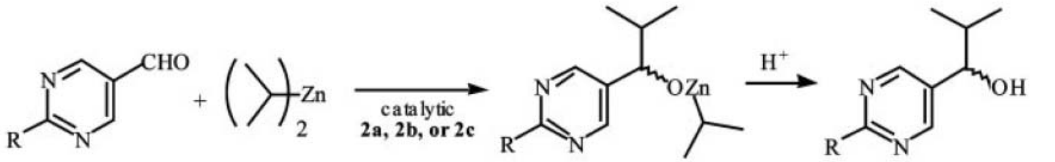


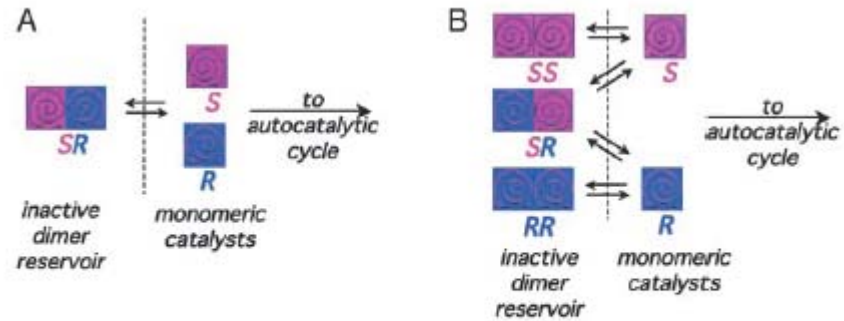
Figure 1 | Theories of prebiotic syntheses of pyrimidine ribonucleotides. The idea that RNA might have formed spontaneously on early Earth has inspired a search for feasible prebiotic syntheses of ribonucleotides, the building blocks of RNA. **a**, The traditional view is that the ribose sugar and nucleobase components of ribonucleotides formed separately, and then combined. But no plausible reactions have been found in which the two components could have joined together. **b**, Powner *et al.*[2] show that a single 2-aminooxazole intermediate could have contributed atoms to both the sugar and nucleobase portions of pyrimidine ribonucleotides, so that components did not have to form separately. For a more detailed overview of the pathways depicted here, see Figure 1 on page 239.



1a R = -H
1b R = $-CH_3$
1c R = $-C{\equiv}CSi(CH_3)_3$

2a R = -H
2b R = $-CH_3$
2c R = $-C{\equiv}CSi(CH_3)_3$

Scheme 1. The Soai autocatalytic reaction.



Scheme 2. Models for including mutual antagonism in autocatalytic systems. (*A*) Specific mutual antagonism: enantiomeric *R* and *S* catalysts form a reservoir of inactive heterochiral dimers. If the initial ratio of *S*:*R* enantiomers is not 1:1, a greater fraction of the minor enantiomer is extracted into the dimer reservoir, which has total *S*:total *R* ratio equal to 1:1. (*B*) Unspecific mutual antagonism: enantiomeric *R* and *S* catalysts form a reservoir of inactive homochiral and heterochiral dimers in statistical proportions.

Asymmetric autocatalysis and its implications for the origin of homochirality

Donna G. Blackmond*
Department of Chemistry, Imperial College, London SW7 2AZ, United Kingdom

Edited by Jack Halpern, University of Chicago, Chicago, IL, and approved February 20, 2004 (received for review December 18, 2003)

An autocatalytic reaction in which the reaction product serves as a catalyst to produce more of itself and to suppress production of its enantiomer serves as a mechanistic model for the evolution of homochirality. The Soai reaction provided experimental confirmation of this concept, nearly 50 years after it was first proposed. This Perspective offers a rationalization of the Soai autocatalytic reaction; accounting for enantiomeric excess and rate observations, that is both simple as well as gratifying in its implications for the chemical origin of life.

The Soai reaction

Proc.Natl.Acad.Sci.USA **101**:5732-5736, 2004

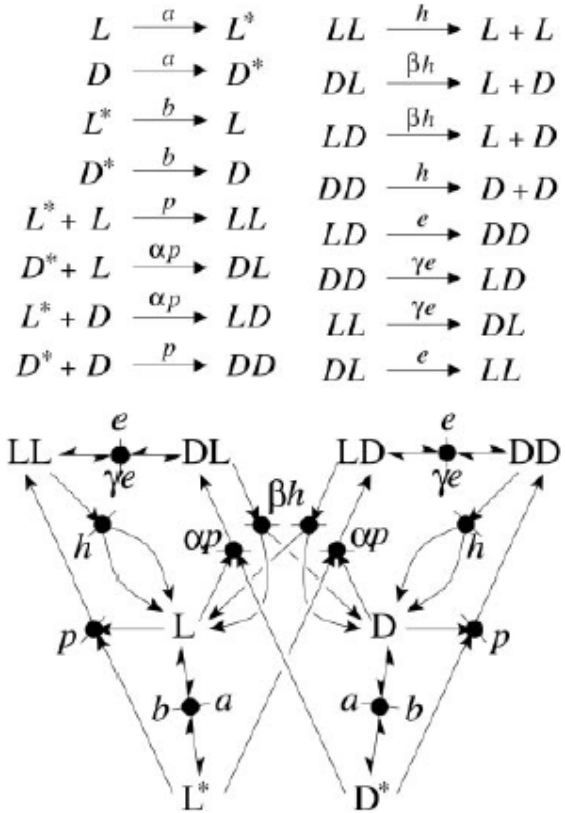


Fig. 1. Minimal APED system limited to dimerizations of L and D residues. (*Upper*) Chemical reactions. (*Lower*) Reaction network. *a*, Activation; *b*, deactivation; *p*, homochiral polymerization; αp, heterochiral polymerization; *h*, homochiral hydrolysis; βh, heterochiral hydrolysis; *e*, homochiral epimerization; γe, heterochiral epimerization.

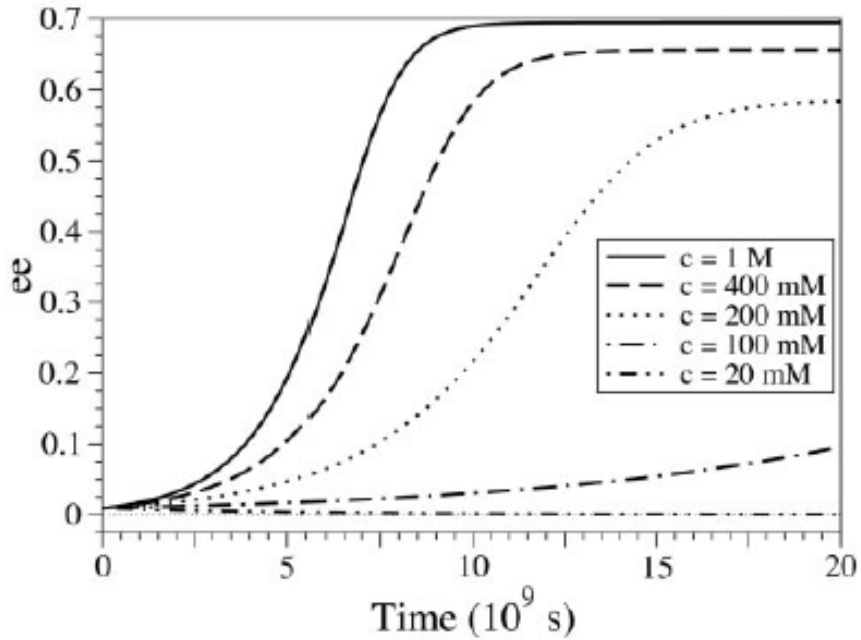
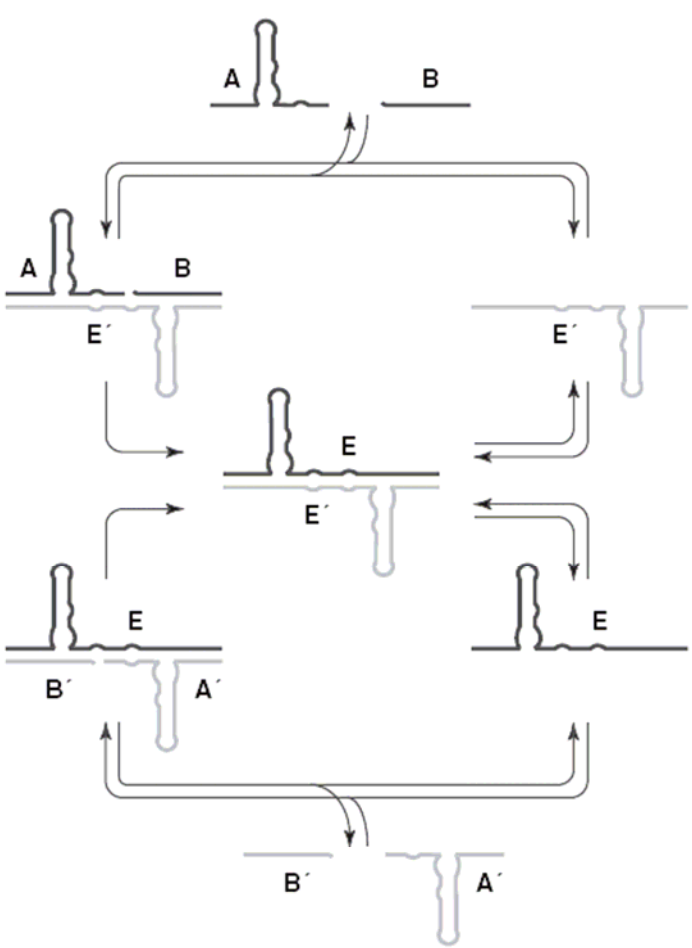


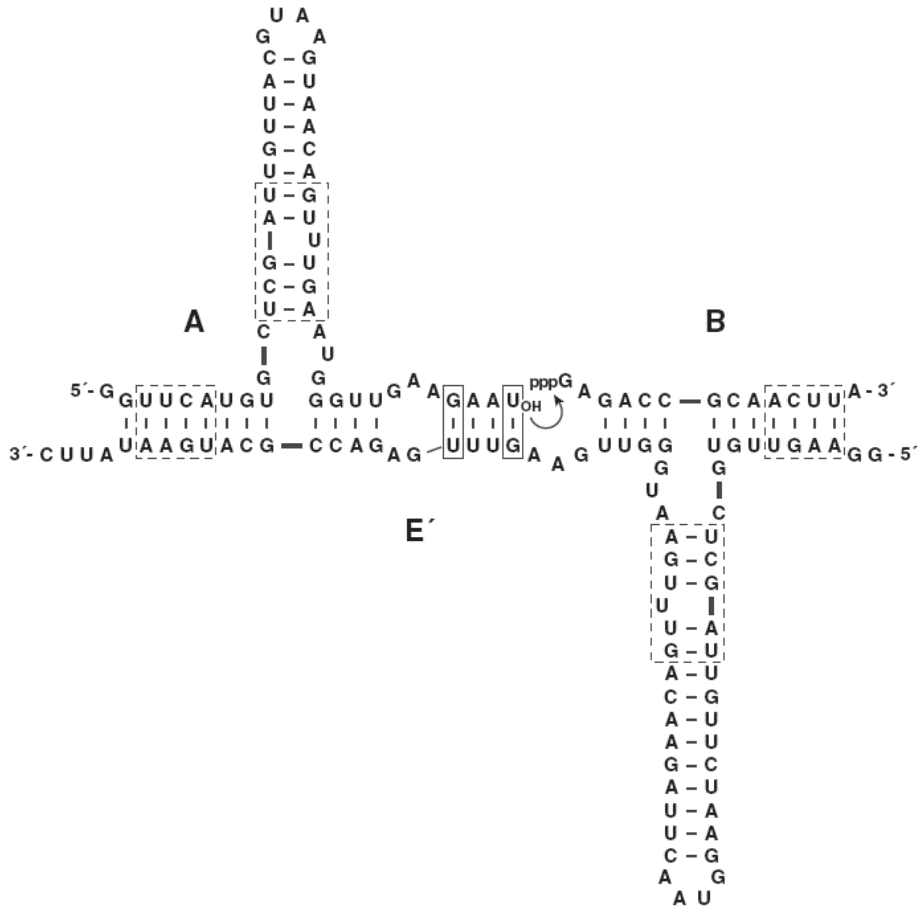
Fig. 5. Time evolution of the enantiomeric excess of APED systems based on experimental data, for different concentrations in residues *c*. Calculated for $ee^{ini} = 0.01$; $a = 10^{-8}\,s^{-1}$, $b = 5{\cdot}10^{-4}\,s^{-1}$, $p = 2{\cdot}10^{2}\,s^{-1}{\cdot}M^{-1}$, $\alpha = 0.35$, $h = 10^{-7}\,s^{-1}$, $\beta = 0.2$, $e = 10^{-7}\,s^{-1}$, $\gamma = 0.3$.

Recycling Frank: Spontaneous emergence of homochirality in noncatalytic systems

Raphaël Plasson, Hugues Bersini, and Auguste Commeyras

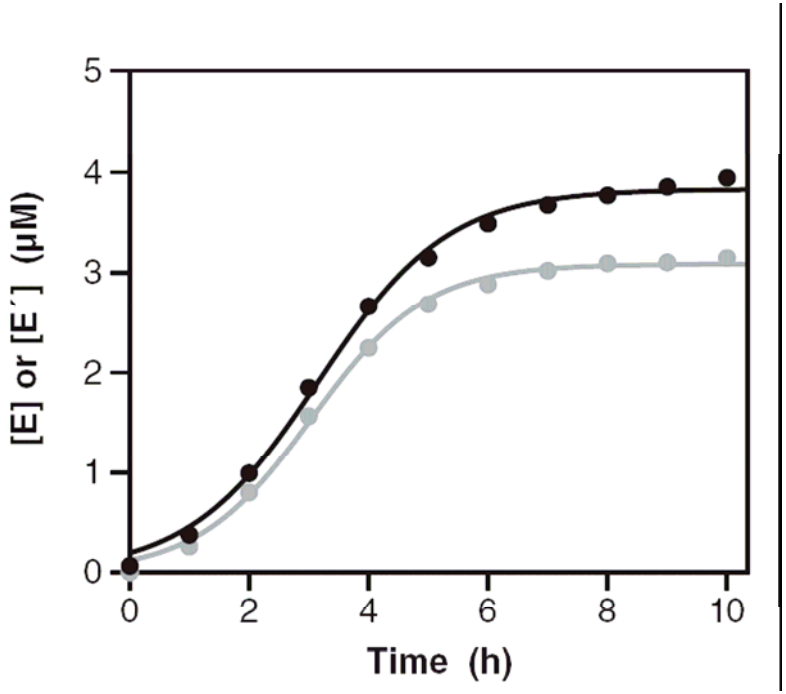
PNAS 2004;101;16733-16738; originally published online Nov 17, 2004;
doi:10.1073/pnas.0405293101

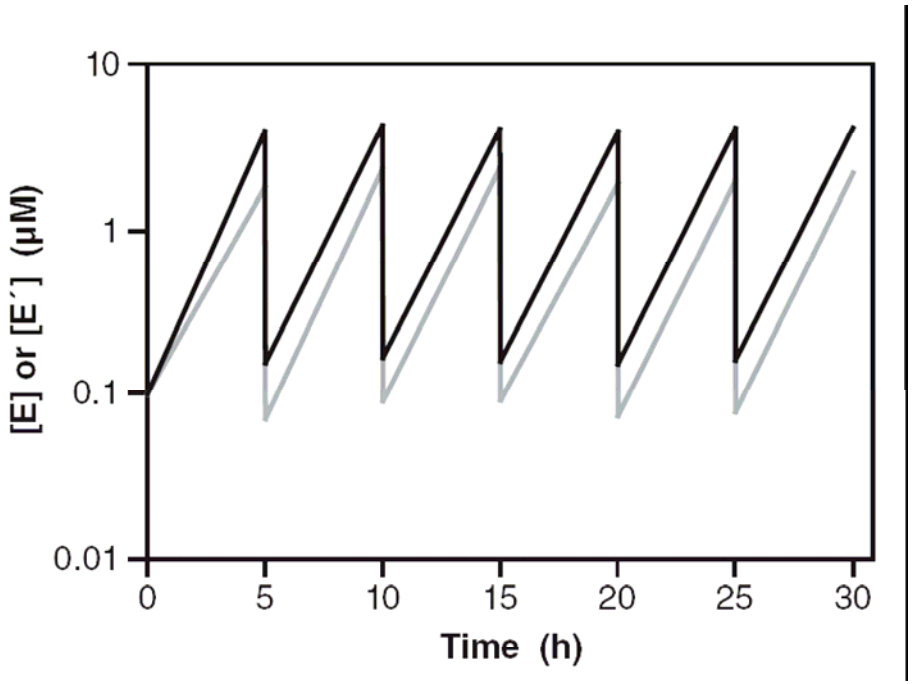




Cross-catalysis of two RNA enzymes leads to self-sustained replication

Tracey A. Lincoln, Gerald F. Joyce, *Science* **323**, 1229-1232, 2009



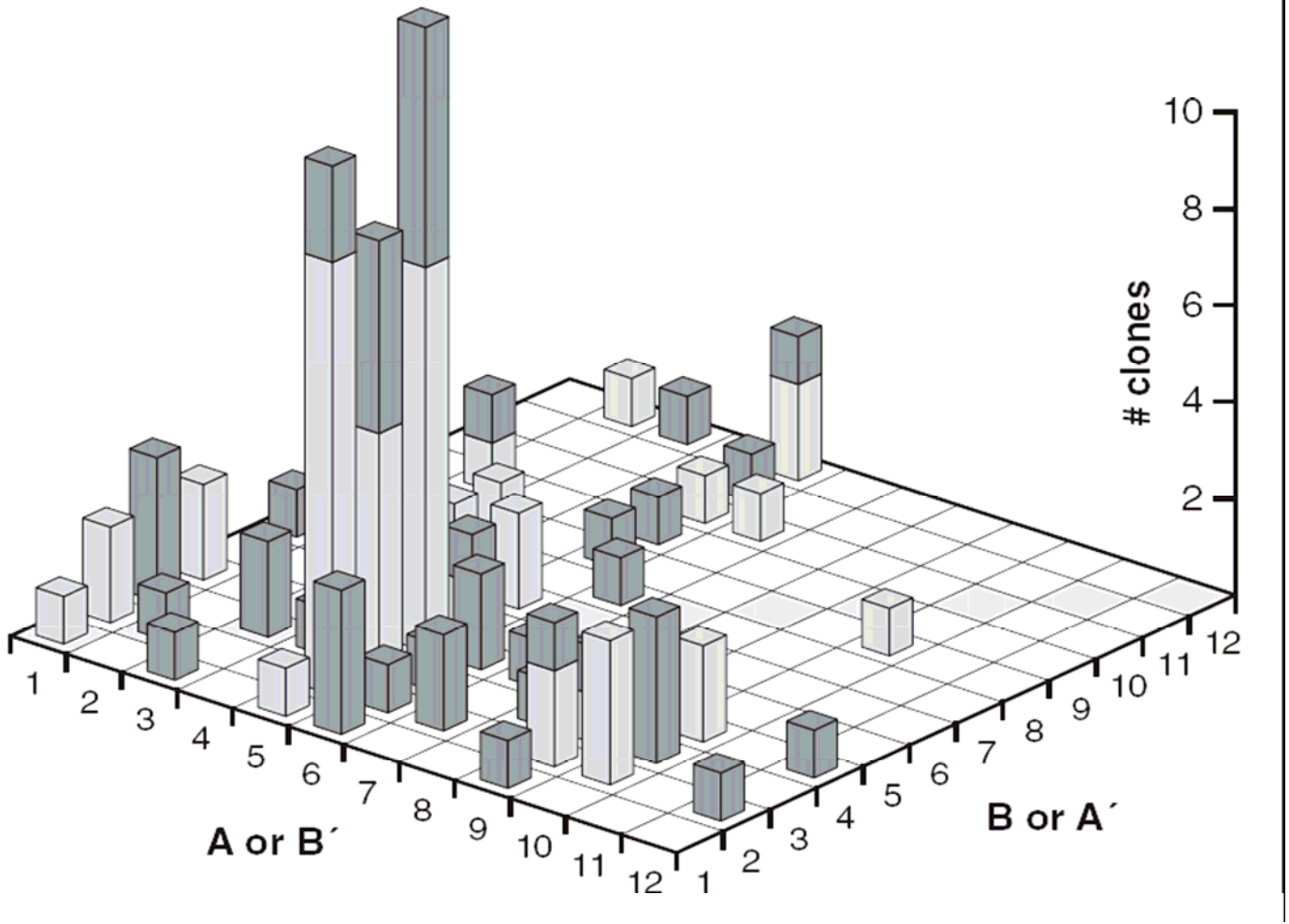


Amplification: $1.5 \mu 10^{10}$

Exponential growth levels off when the reservoir is exhausted (l.h.s.).

RNA production in serial transfer experiments (r.h.s.)

Tracey A. Lincoln, Gerald F. Joyce, *Science* **323**, 1229-1232, 2009



RNA evolution of recombinant replicators

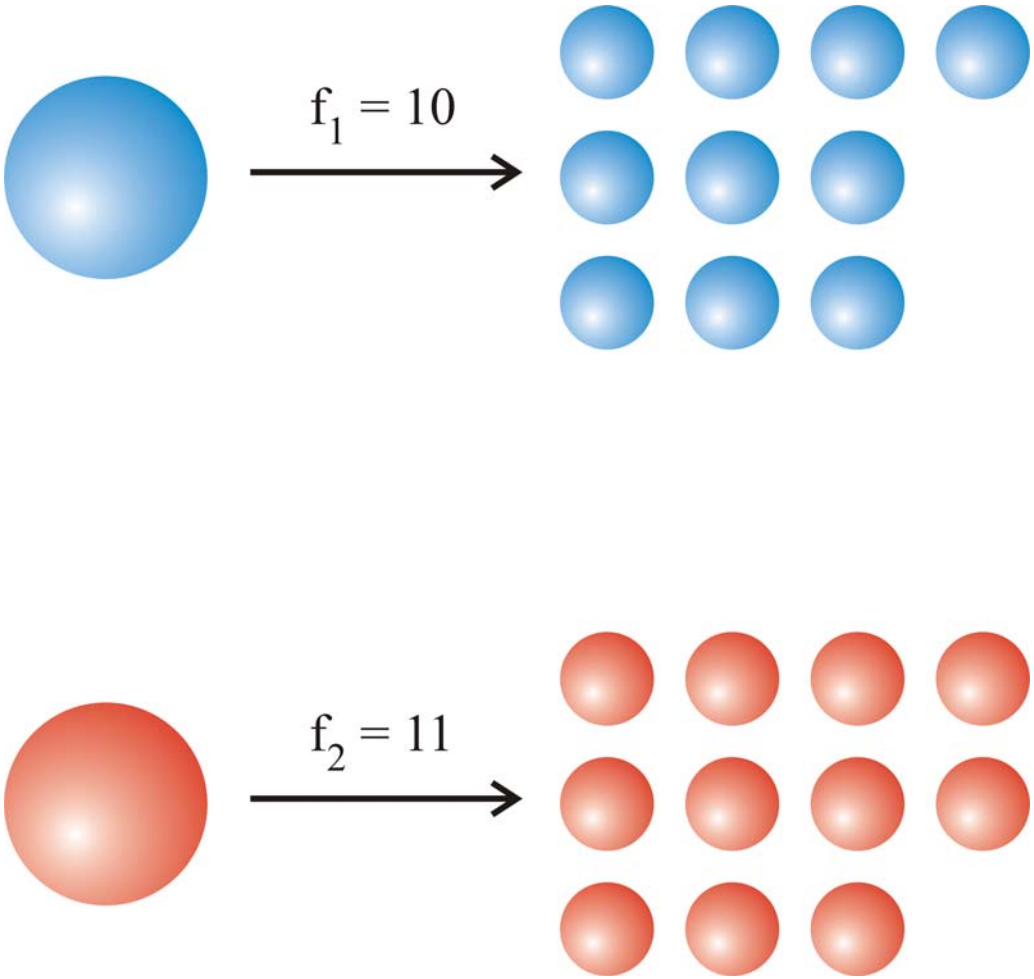
Tracey A. Lincoln, Gerald F. Joyce, *Science* **323**, 1229-1232, 2009

1. Was ist Leben?
2. Chemische Evolution
3. **Darwins tiefe Einsichten**
4. Der Ursprung biologischer Information
5. Darwinsche Evolution mit Molekülen
6. Evolutionäre Biotechnologie
7. Die DNA + Protein Welt
8. Evolution bis zum Menschen

Drei notwendige Bedingungen für Darwinsche Evolution:

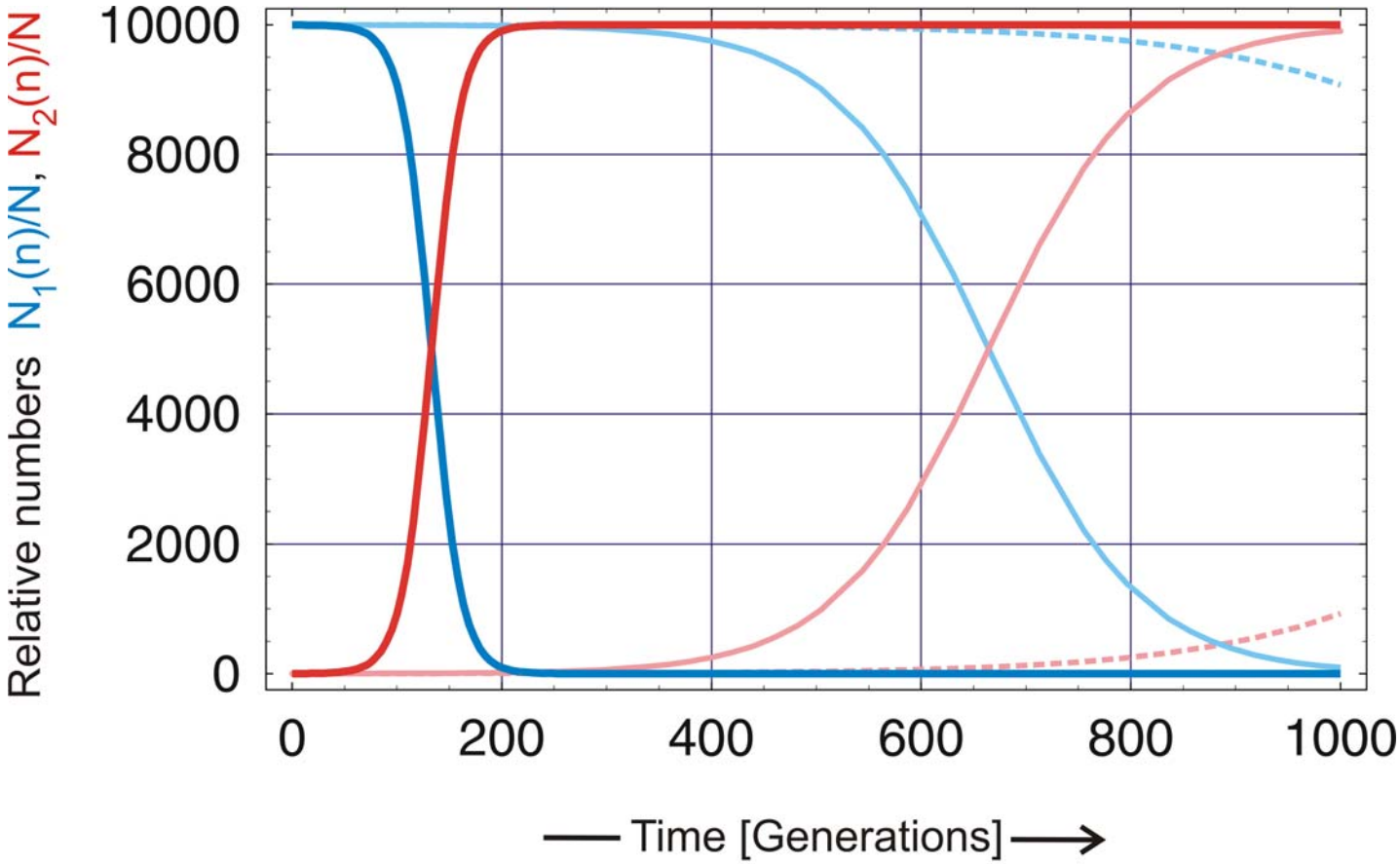
1. **Vermehrung**
2. **Variation**
3. **Selektion**

Empirisch erkanntes Prinzip der natürlichen Auslese



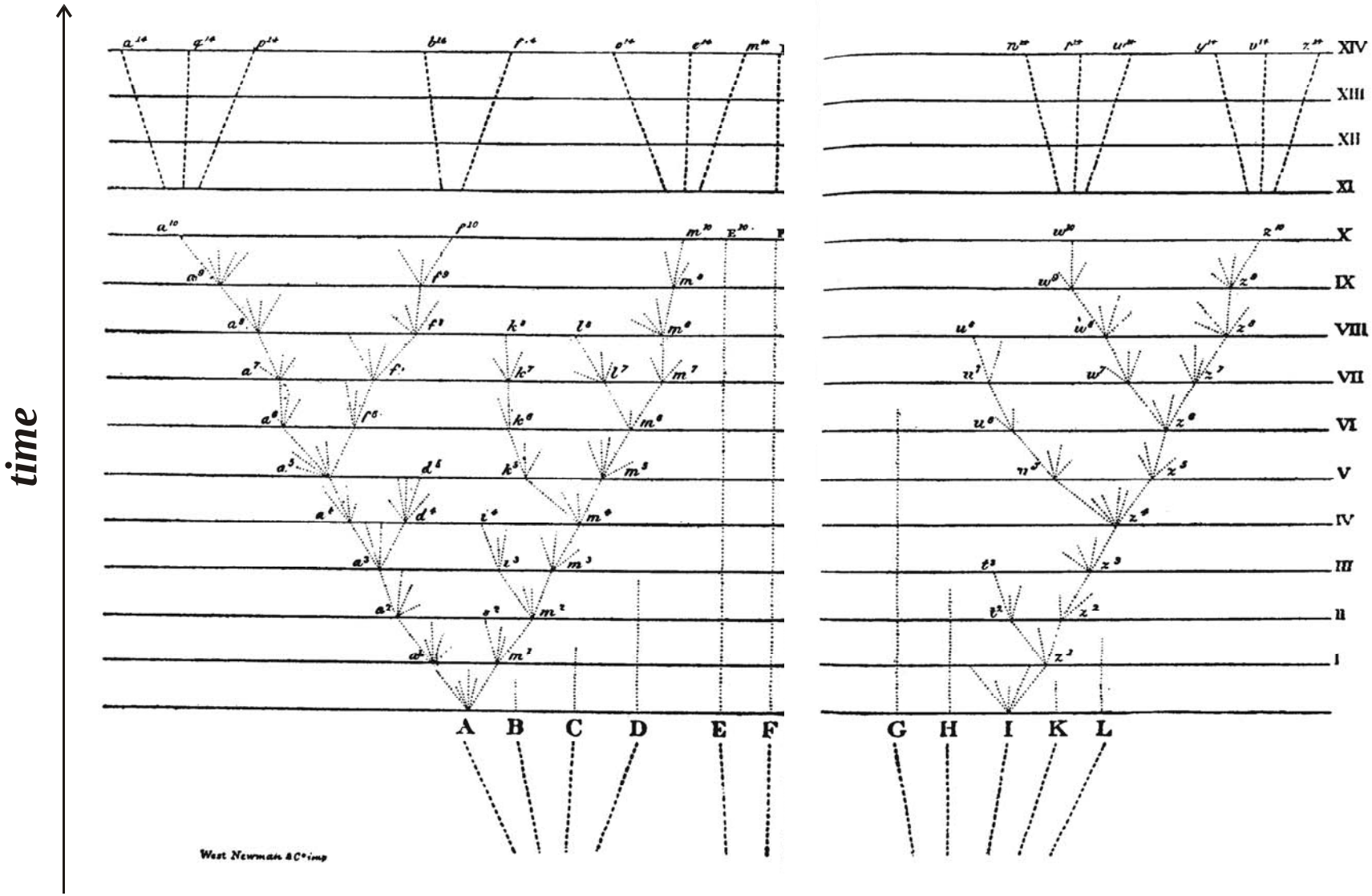
$$s = \frac{f_2 - f_1}{f_1} = 0.1$$

Two variants with a mean progeny of ten or eleven descendants

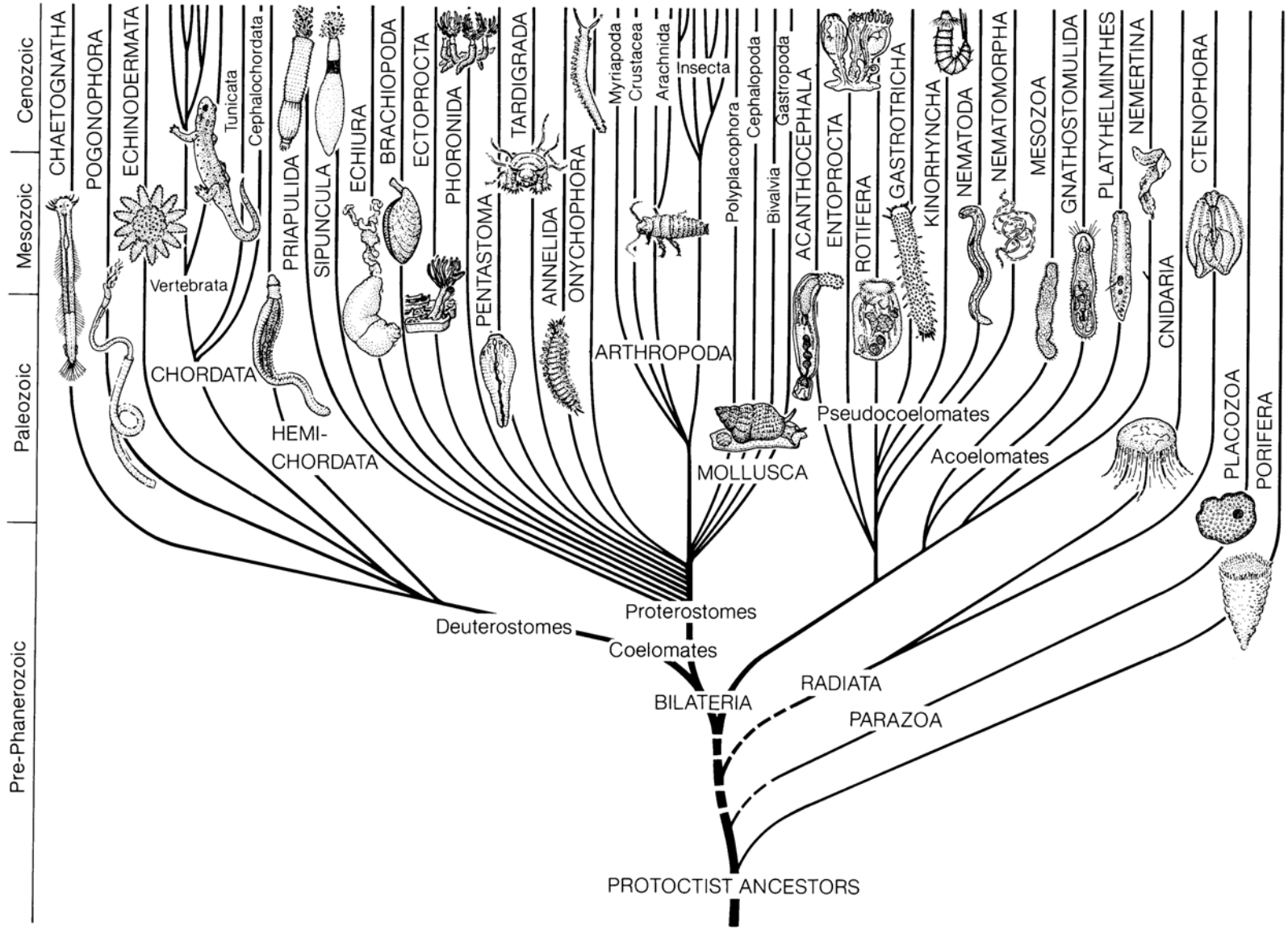


$$N_1(0) = 9999,\ N_2(0) = 1; \quad s = 0.1, 0.02, 0.01$$

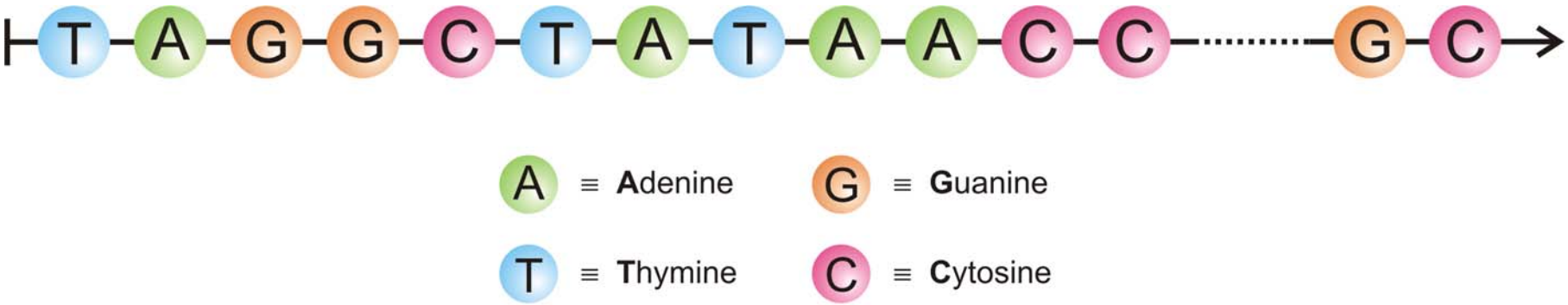
Selection of advantageous mutants in populations of N = 10 000 individuals



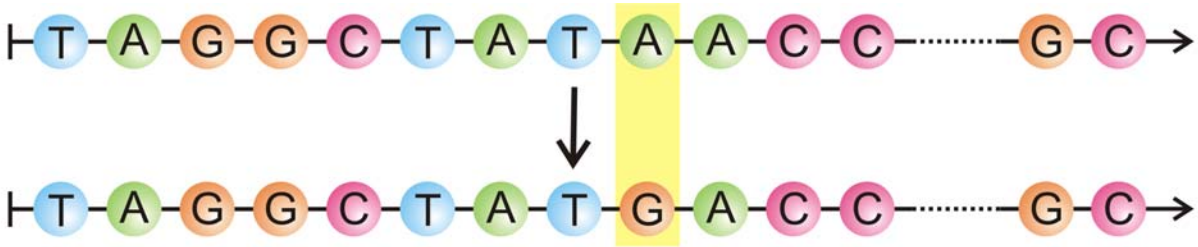
Charles Darwin, *The Origin of Species*, 6th edition.
Everyman‘s Library, Vol.811, Dent London, pp.121-122.



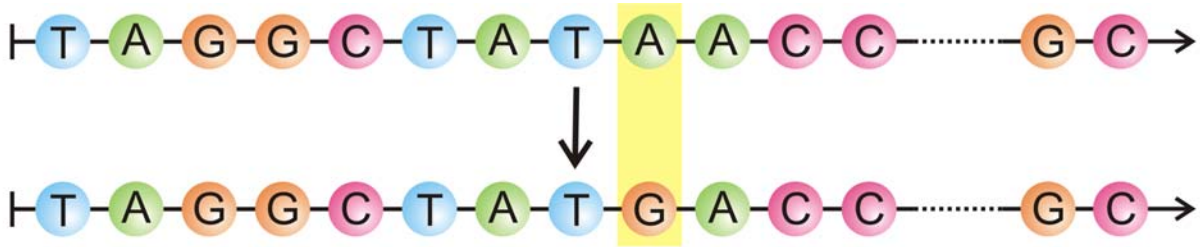
Modern phylogenetic tree: Lynn Margulis, Karlene V. Schwartz. *Five Kingdoms. An Illustrated Guide to the Phyla of Life on Earth*. W.H. Freeman, San Francisco, 1982.



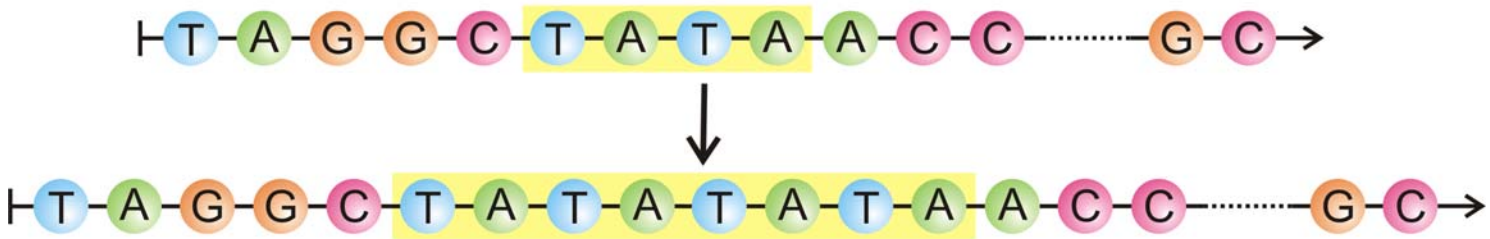
Deoxyribonucleic acid - DNA



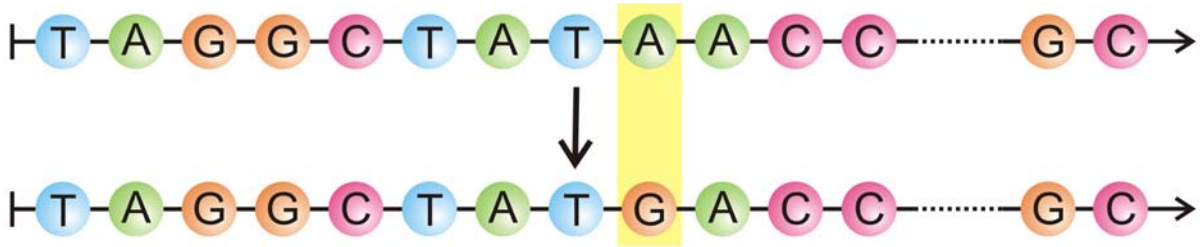
Punktmutation



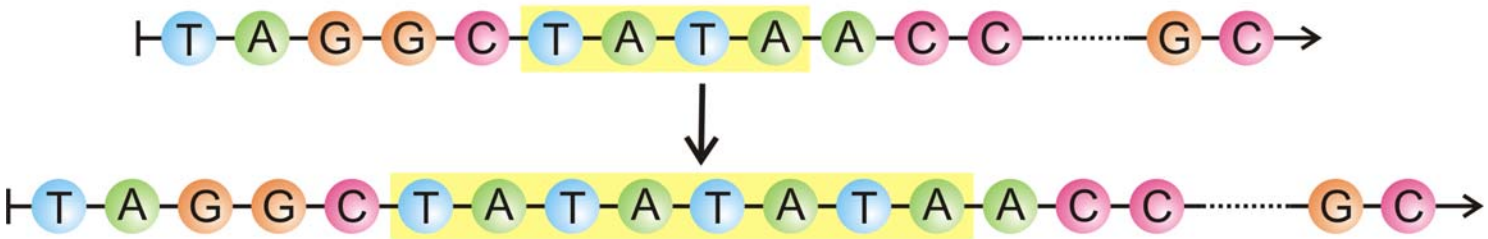
Punktmutation



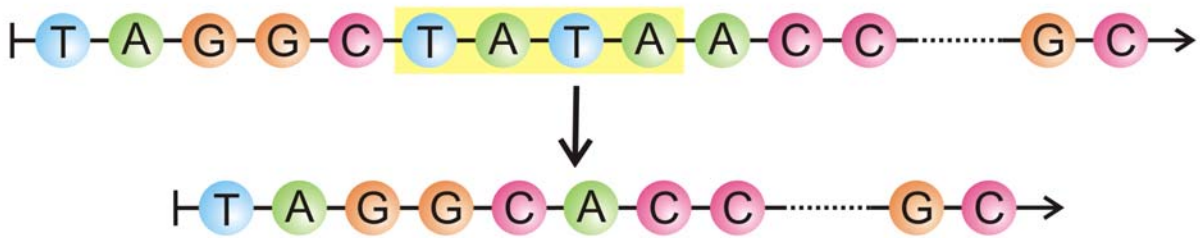
Insertion



Punktmutation



Insertion



Deletion

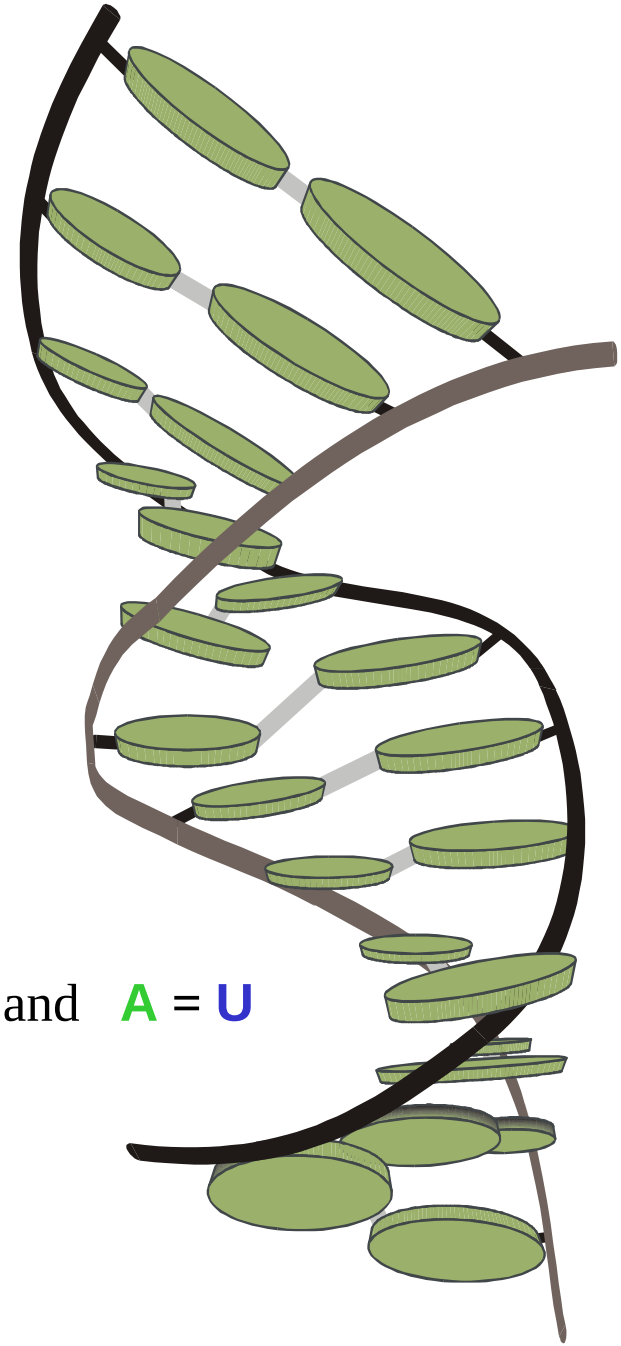
Reconstruction of phylogenies through comparison of molecular sequence data

Molekulare Evolutionsforschung durch DNA-Sequenzierung

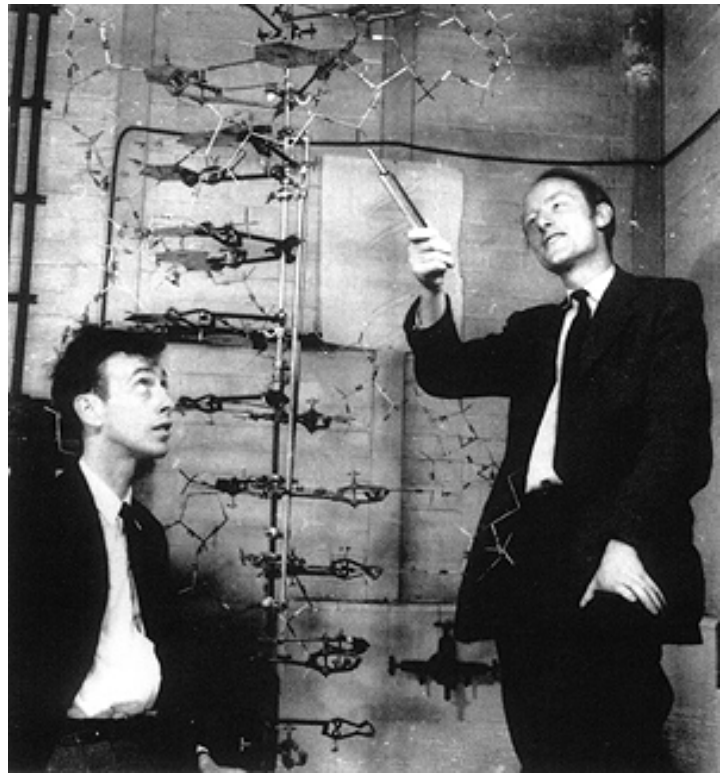
Aus dem Vergleich der heutigen DNA-Sequenzen kann die geschichtliche Abfolge der Mutationen rekonstruiert werden und diese ergibt phylogenetische Bäume, die jenen aus der vergleichenden Morphologie, welche durch Betrachtung von Formen und Gestalten der Organismen gewonnen wurden, weitest gehend entsprechen.

Eine in der Vergangenheit postulierte **molekulare Uhr der Evolution** verlangt, dass die Mutationshäufigkeiten auf den verschiedenen Ästen der phylogenetischen Bäume gleich groß sind. Die **molekulare Uhr** ist bei Wirbeltieren recht gut erfüllt, trifft aber für die wirbellosen Tiere nicht zu.

1. Was ist Leben?
2. Chemische Evolution
3. Darwins tiefe Einsichten
4. **Der Ursprung biologischer Information**
5. Darwinsche Evolution mit Molekülen
6. Evolutionäre Biotechnologie
7. Die DNA + Protein Welt
8. Evolution bis zum Menschen



G ≡ C and **A = U**

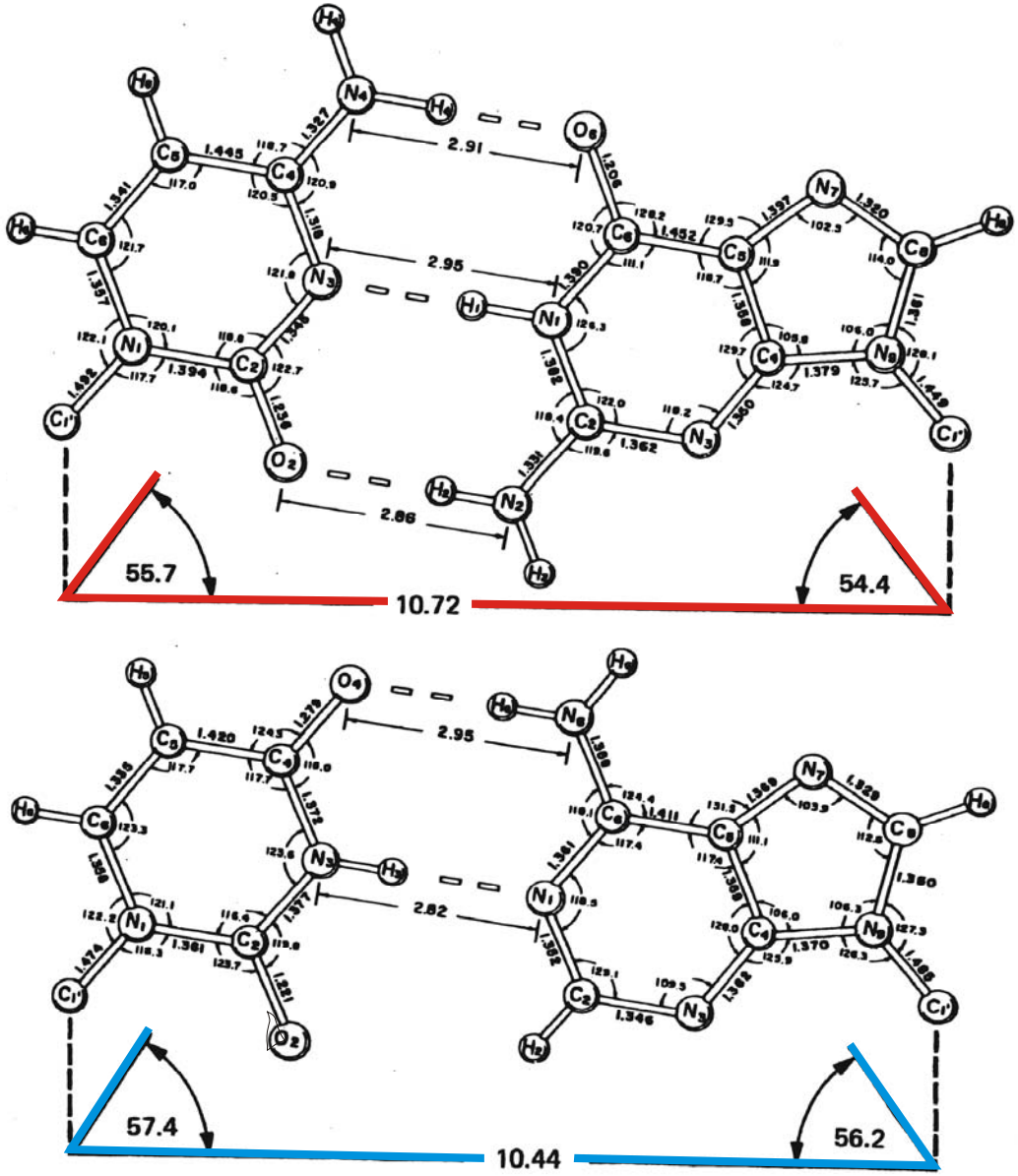


James D. Watson, 1928- , and Francis Crick, 1916-2004,
Nobel Prize 1962

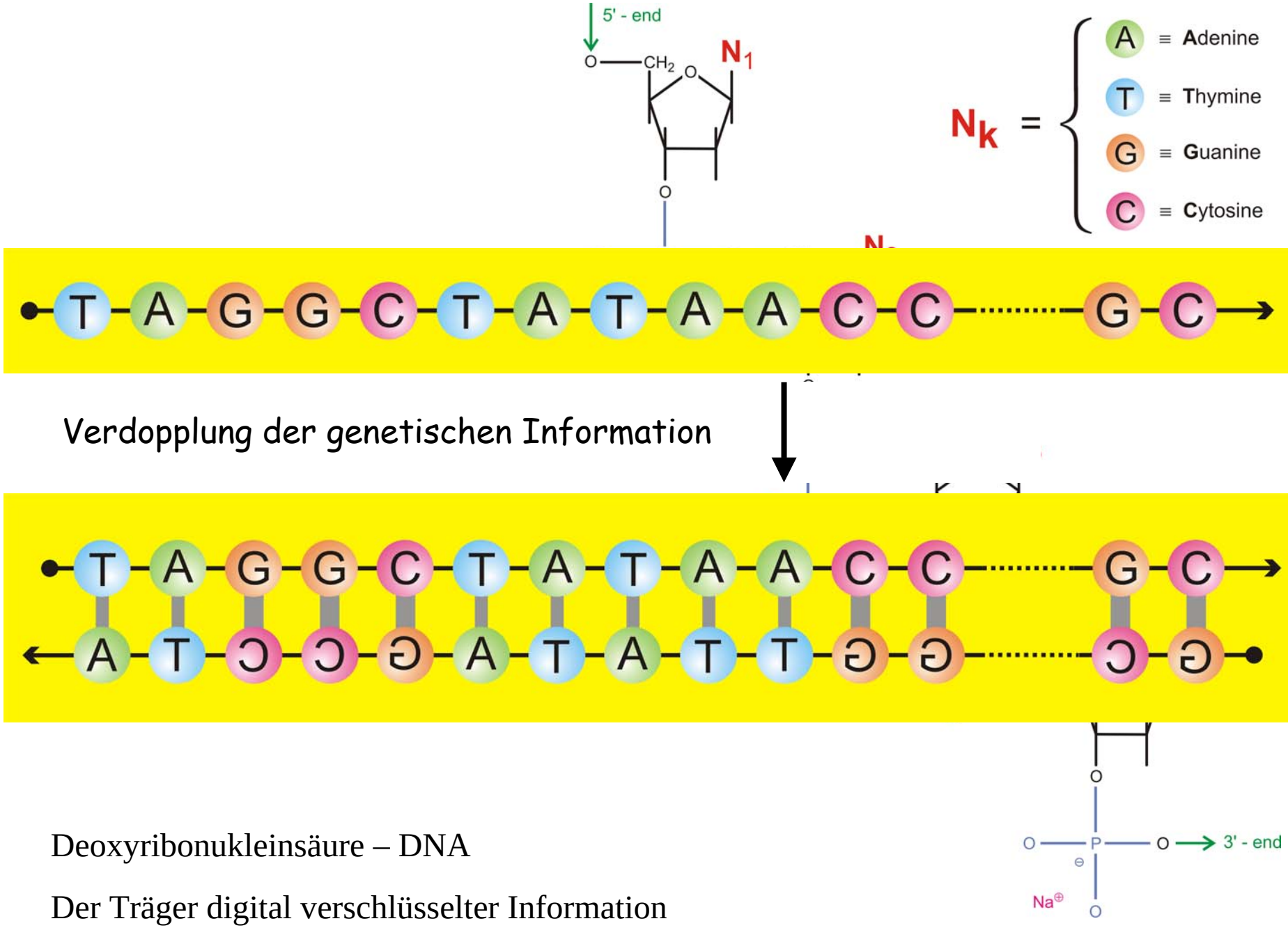
The three-dimensional structure of a
short double helical stack of B-DNA

Canonical Watson-Crick base pairs:

cytosine – guanine
uracil – adenine

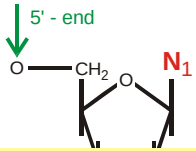


W.Saenger, Principles of Nucleic Acid Structure, Springer, Berlin 1984

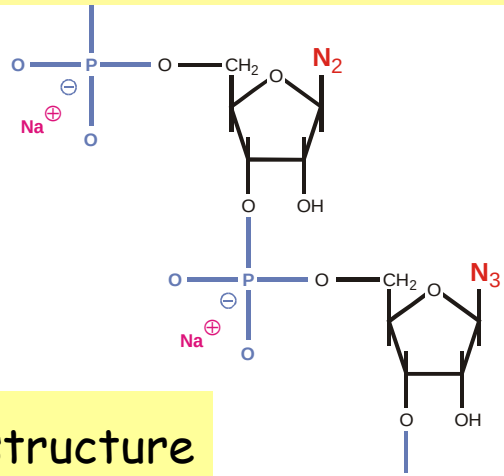


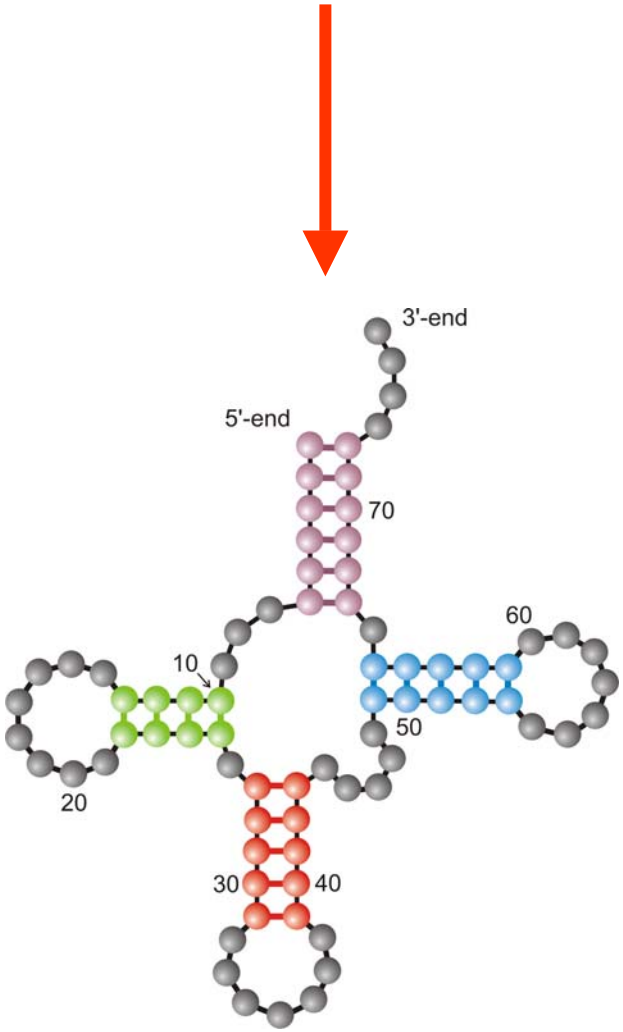
Deoxyribonukleinsäure – DNA

Der Träger digital verschlüsselter Information

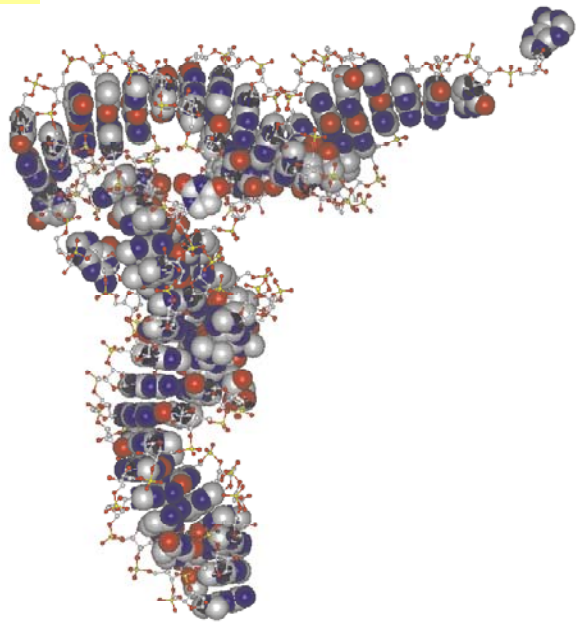


5'-end GCGGAUUUAGCUCAGUUGGGAGAGCGCCAGACUGAAGAUCUGGAGGUCCUGUGUUCGAUCCACAGAAUUCGCACCA 3'-end



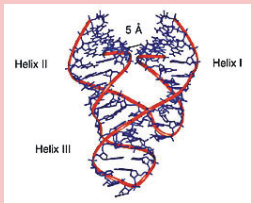


Definition of RNA structure



RNA

RNA *as catalyst*



Ribozyme

RNA *as transmitter of genetic information*

DNA

transcription

...AGAGCGCCAGACUGAAGAUCUGGAGGUCCUGUGUUC...

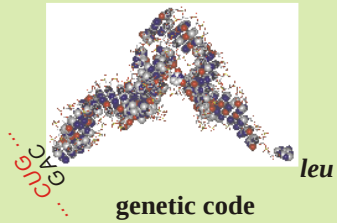
messenger-**RNA**

translation

protein

RNA as **working copy** of genetic information

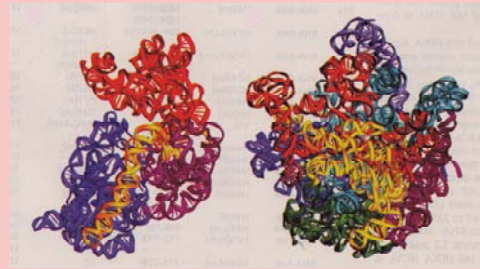
RNA *as adapter molecule*



leu

genetic code

RNA *is the catalytic subunit in supramolecular complexes*



The **ribosome** is a **ribozyme !**

The RNA *world as a precursor of the current* DNA + protein *biology*

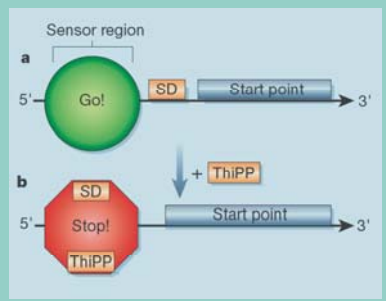
RNA *as carrier of genetic information*

RNA viruses and retroviruses

RNA evolution *in vitro*

Evolutionary biotechnology
RNA aptamers, artificial ribozymes, allosteric ribozymes

Allosteric control of transcribed RNA



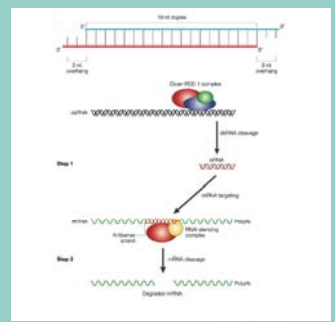
Riboswitches controlling transcription and translation through **metabolites**

RNA *is modified by epigenetic control*

RNA editing

Alternative splicing of messenger **RNA**

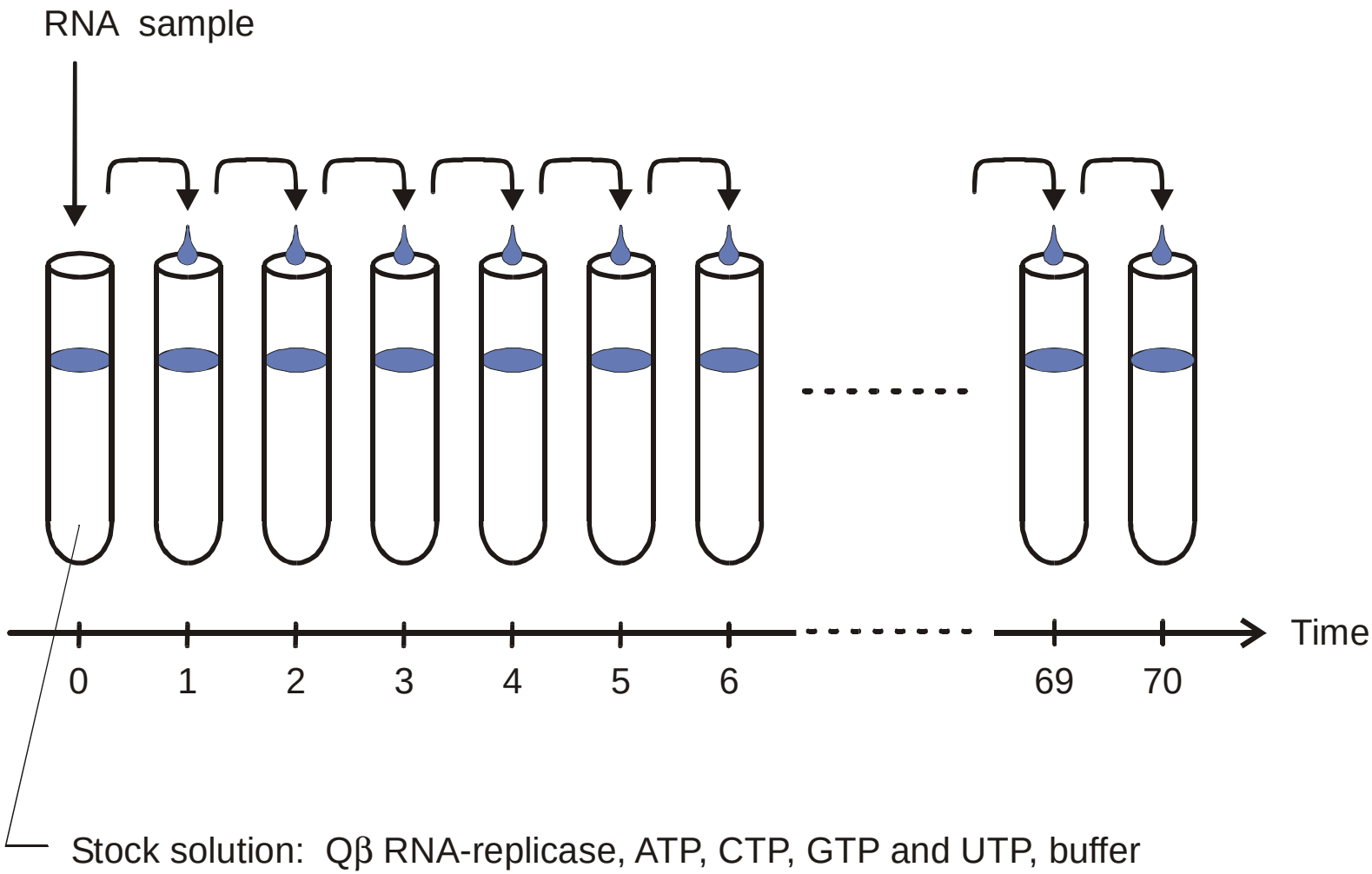
RNA *as regulator of gene expression*



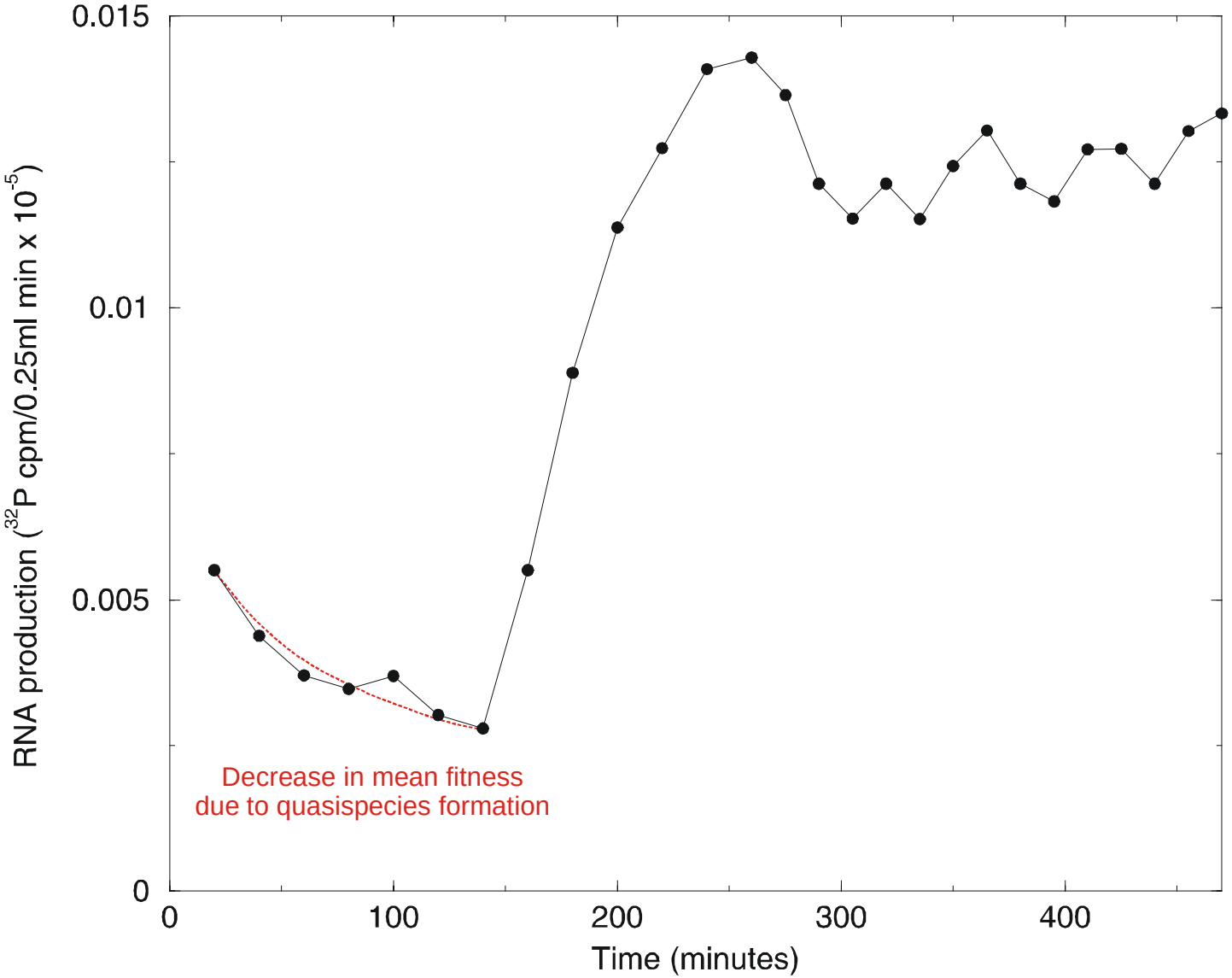
Gene silencing by small interfering RNAs

Functions of RNA molecules

1. Was ist Leben?
2. Chemische Evolution
3. Darwins tiefe Einsichten
4. Der Ursprung biologischer Information
5. **Darwinsche Evolution mit Molekülen**
6. Evolutionäre Biotechnologie
7. Die DNA + Protein Welt
8. Evolution bis zum Menschen



Anwendung der seriellen Überimpfungstechnik auf RNA-Evolution in Reagenzglas



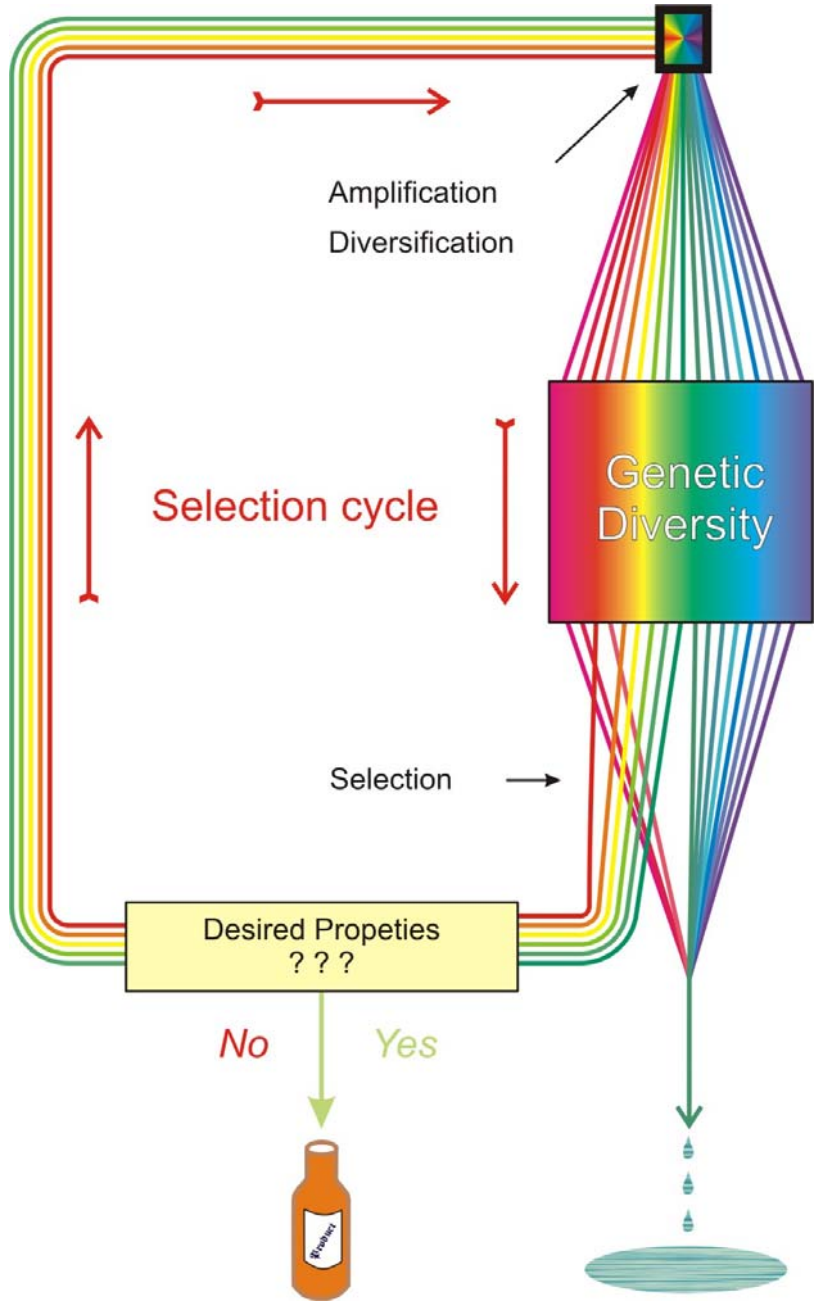
The increase in RNA production rate during a serial transfer experiment

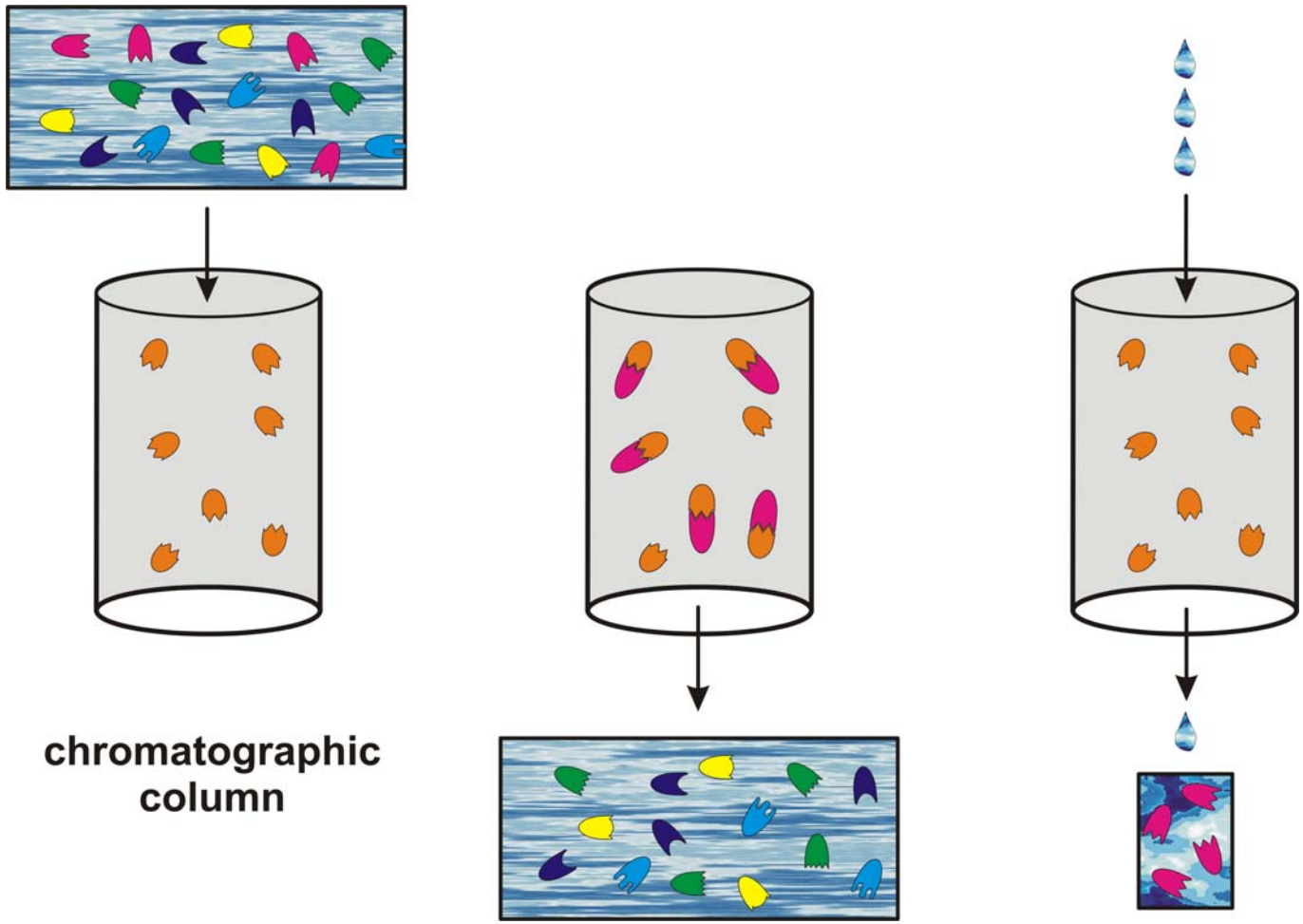
	Generation time	Selection and adaptation 10 000 generations	Genetic drift in small populations 10^6 generations	Genetic drift in large populations 10^7 generations
RNA molecules	10 sec 1 min	27.8 h = 1.16 d 6.94 d	115.7 d 1.90 a	3.17 a 19.01 a
Bacteria	20 min 10 h	138.9 d 11.40 a	38.03 a 1 140 a	380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

Time scales of evolutionary change

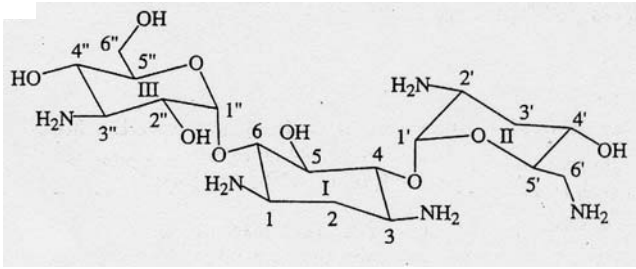
1. Was ist Leben?
2. Chemische Evolution
3. Darwins tiefe Einsichten
4. Der Ursprung biologischer Information
5. Darwinsche Evolution mit Molekülen
6. **Evolutionäre Biotechnologie**
7. Die DNA + Protein Welt
8. Evolution bis zum Menschen

Ein Beispiel für Selektion von Molekülen mit vorbestimmbaren Eigenschaften im Laborexpriment

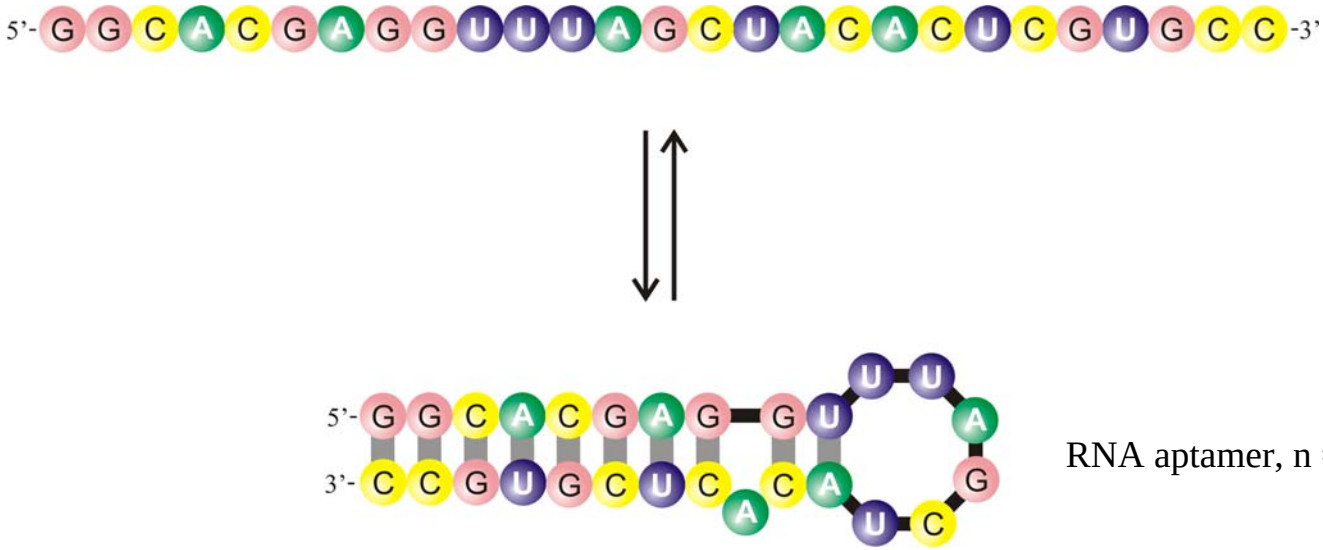




Die SELEX-Technik zur evolutionären Erzeugung von stark bindenden Molekülen



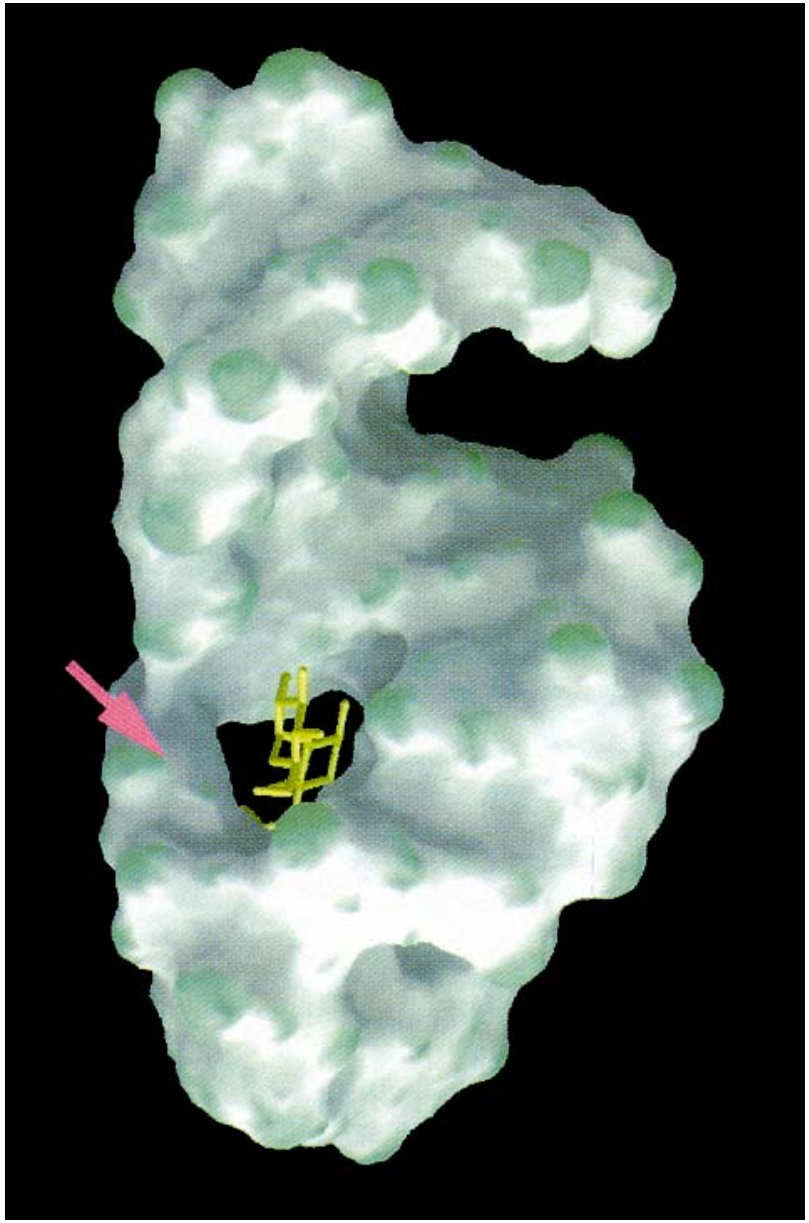
tobramycin



RNA aptamer, n = 27

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)

A ribozyme switch

E.A.Schultes, D.B.Bartel, Science **289** (2000), 448-452

minus the background levels observed in the HSP in the control (Sar1-GDP–containing) incubation that prevents COPII vesicle formation. In the microsome control, the level of p115-SNARE associations was less than 0.1%.
46. C. M. Carr, E. Grote, M. Munson, F. M. Hughson, P. J. Novick, *J. Cell Biol.* **146**, 333 (1999).
47. C. Ungermann, B. J. Nichols, H. R. Pelham, W. Wickner, *J. Cell Biol.* **140**, 61 (1998).
48. E. Grote and P. J. Novick, *Mol. Biol. Cell* **10**, 4149 (1999).
49. P. Uetz *et al.*, *Nature* **403**, 623 (2000).
50. GST-SNARE proteins were expressed in bacteria and purified on glutathione-Sepharose beads using standard methods. Immobilized GST-SNARE protein (0.5 μM) was incubated with rat liver cytosol (20 mg) or purified recombinant p115 (0.5 μM) in 1 ml of NS buffer containing 1% BSA for 2 hours at 4°C with rotation. Beads were briefly spun (3000 rpm for 10 s) and sequentially washed three times with NS buffer and three times with NS buffer supplemented with 150 mM NaCl. Bound proteins were eluted three times in 50 μl of 50 mM tris-HCl (pH 8.5), 50 mM reduced glutathione, 150 mM NaCl, and 0.1% Triton X-100 for 15 min at 4°C with intermittent mixing, and elutes were pooled. Proteins were precipitated by MeOH/CH_3Cl and separated by SDS–polyacrylamide gel electrophoresis (PAGE) followed by immunoblotting using p115 mAb 13F12.
51. V. Rybin *et al.*, *Nature* **383**, 266 (1996).
52. K. G. Hardwick and H. R. Pelham, *J. Cell Biol.* **119**, 513 (1992).
53. A. P. Newman, M. E. Groesch, S. Ferro-Novick, *EMBO J.* **11**, 3609 (1992).
54. A. Spang and R. Schekman, *J. Cell Biol.* **143**, 589 (1998).
55. M. F. Rexach, M. Latterich, R. W. Schekman, *J. Cell Biol.* **126**, 1133 (1994).
56. A. Mayer and W. Wickner, *J. Cell Biol.* **136**, 307 (1997).
57. M. D. Turner, H. Plutner, W. E. Balch, *J. Biol. Chem.* **272**, 13479 (1997).
58. A. Price, D. Seals, W. Wickner, C. Ungermann, *J. Cell Biol.* **148**, 1231 (2000).
59. X. Cao and C. Barlowe, *J. Cell Biol.* **149**, 55 (2000).
60. G. G. Tall, H. Hama, D. B. DeWald, B. F. Horazdovsky, *Mol. Biol. Cell* **10**, 1873 (1999).
61. C. G. Burd, M. Peterson, C. R. Cowles, S. D. Emr, *Mol. Biol. Cell* **8**, 1089 (1997).
62. M. R. Peterson, C. G. Burd, S. D. Emr, *Curr. Biol.* **9**, 159 (1999).
63. M. G. Waters, D. O. Clary, J. E. Rothman, *J. Cell Biol.* **118**, 1015 (1992).
64. D. M. Walter, K. S. Paul, M. G. Waters, *J. Biol. Chem.* **273**, 29565 (1998).
65. N. Hui *et al.*, *Mol. Biol. Cell* **8**, 1777 (1997).
66. T. E. Kreis, *EMBO J.* **5**, 931 (1986).
67. H. Plutner, H. W. Davidson, J. Saraste, W. E. Balch, *J. Cell Biol.* **119**, 1097 (1992).
68. D. S. Nelson *et al.*, *J. Cell Biol.* **143**, 319 (1998).
69. We thank G. Waters for p115 cDNA and p115 mAbs; G. Warren for p97 and p47 antibodies; R. Scheller for rbet1, membrin, and sec22 cDNAs; H. Plutner for excellent technical assistance; and P. Tan for help during the initial phase of this work. Supported by NIH grants GM 33301 and GM42336 and National Cancer Institute grant CA58689 (W.E.B.), a NIH National Research Service Award (B.D.M.), and a Wellcome Trust International Traveling Fellowship (B.B.A.).

20 March 2000; accepted 22 May 2000

One Sequence, Two Ribozymes: Implications for the Emergence of New Ribozyme Folds

Erik A. Schultes and David P. Bartel*

We describe a single RNA sequence that can assume either of two ribozyme folds and catalyze the two respective reactions. The two ribozyme folds share no evolutionary history and are completely different, with no base pairs (and probably no hydrogen bonds) in common. Minor variants of this sequence are highly active for one or the other reaction, and can be accessed from prototype ribozymes through a series of neutral mutations. Thus, in the course of evolution, new RNA folds could arise from preexisting folds, without the need to carry inactive intermediate sequences. This raises the possibility that biological RNAs having no structural or functional similarity might share a common ancestry. Furthermore, functional and structural divergence might, in some cases, precede rather than follow gene duplication.

Related protein or RNA sequences with the same folded conformation can often perform very different biochemical functions, indicating that new biochemical functions can arise from preexisting folds. But what evolutionary mechanisms give rise to sequences with new macromolecular folds? When considering the origin of new folds, it is useful to picture, among all sequence possibilities, the distribution of sequences with a particular fold and function. This distribution can range very far in sequence space (*1*). For example, only seven nucleotides are strictly conserved among the group I self-splicing introns, yet secondary (and presumably tertiary) structure within the core of the ribozyme is preserved (*2*). Because these disparate isolates have the same fold and function, it is thought that they descended from a common ancestor through a series of mutational variants that were each functional. Hence, sequence heterogeneity among divergent isolates implies the existence of paths through sequence space that have allowed neutral drift from the ancestral sequence to each isolate. The set of all possible neutral paths composes a "neutral network," connecting in sequence space those widely dispersed sequences sharing a particular fold and activity, such that any sequence on the network can potentially access very distant sequences by neutral mutations (*3–5*).

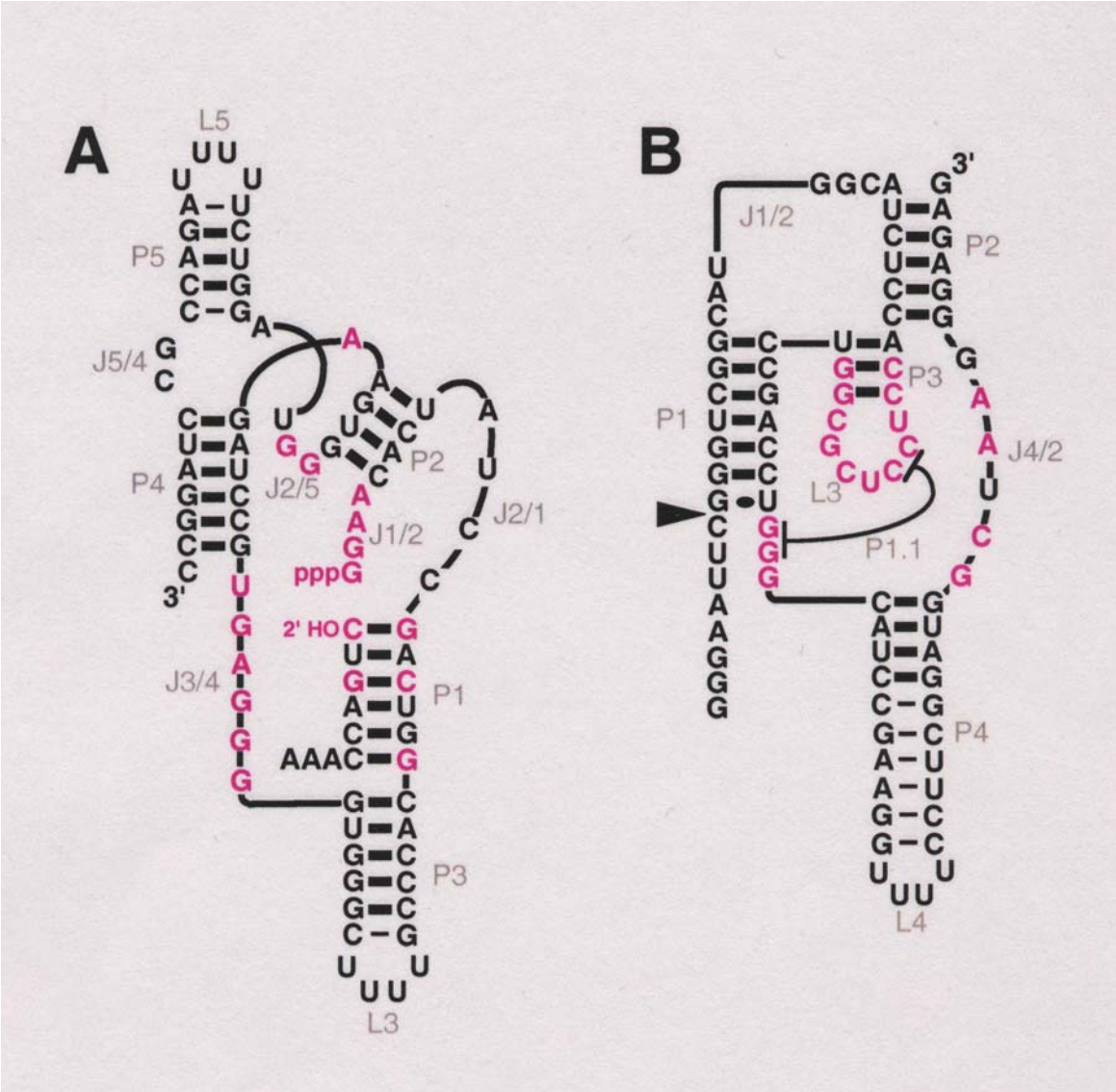
Theoretical analyses using algorithms for predicting RNA secondary structure have suggested that different neutral networks are interwoven and can approach each other very closely (*3*, *5–8*). Of particular interest is whether ribozyme neutral networks approach each other so closely that they intersect. If so, a single sequence would be capable of folding into two different conformations, would have two different catalytic activities, and could access by neutral drift every sequence on both networks. With intersecting networks, RNAs with novel structures and activities could arise from previously existing ribozymes, without the need to carry nonfunctional sequences as evolutionary intermediates. Here, we explore the proximity of neutral networks experimentally, at the level of RNA function. We describe a close apposition of the neutral networks for the hepatitis delta virus (HDV) self-cleaving ribozyme and the class III self-ligating ribozyme.

In choosing the two ribozymes for this investigation, an important criterion was that they share no evolutionary history that might confound the evolutionary interpretations of our results. Choosing at least one artificial ribozyme ensured independent evolutionary histories. The class III ligase is a synthetic ribozyme isolated previously from a pool of random RNA sequences (*9*). It joins an oligonucleotide substrate to its 5′ terminus. The prototype ligase sequence (Fig. 1A) is a shortened version of the most active class III variant isolated after 10 cycles of in vitro selection and evolution. This minimal construct retains the activity of the full-length isolate (*10*). The HDV ribozyme carries out the site-specific self-cleavage reactions needed during the life cycle of HDV, a satellite virus of hepatitis B with a circular, single-stranded RNA genome (*11*). The prototype HDV construct for our study (Fig. 1B) is a shortened version of the antigenomic HDV ribozyme (*12*), which undergoes self-cleavage at a rate similar to that reported for other antigenomic constructs (*13*, *14*).

The prototype class III and HDV ribozymes have no more than the 25% sequence identity expected by chance and no fortuitous structural similarities that might favor an intersection of their two neutral networks. Nevertheless, sequences can be designed that simultaneously satisfy the base-pairing requirements

Whitehead Institute for Biomedical Research and Department of Biology, Massachusetts Institute of Technology, 9 Cambridge Center, Cambridge, MA 02142, USA.

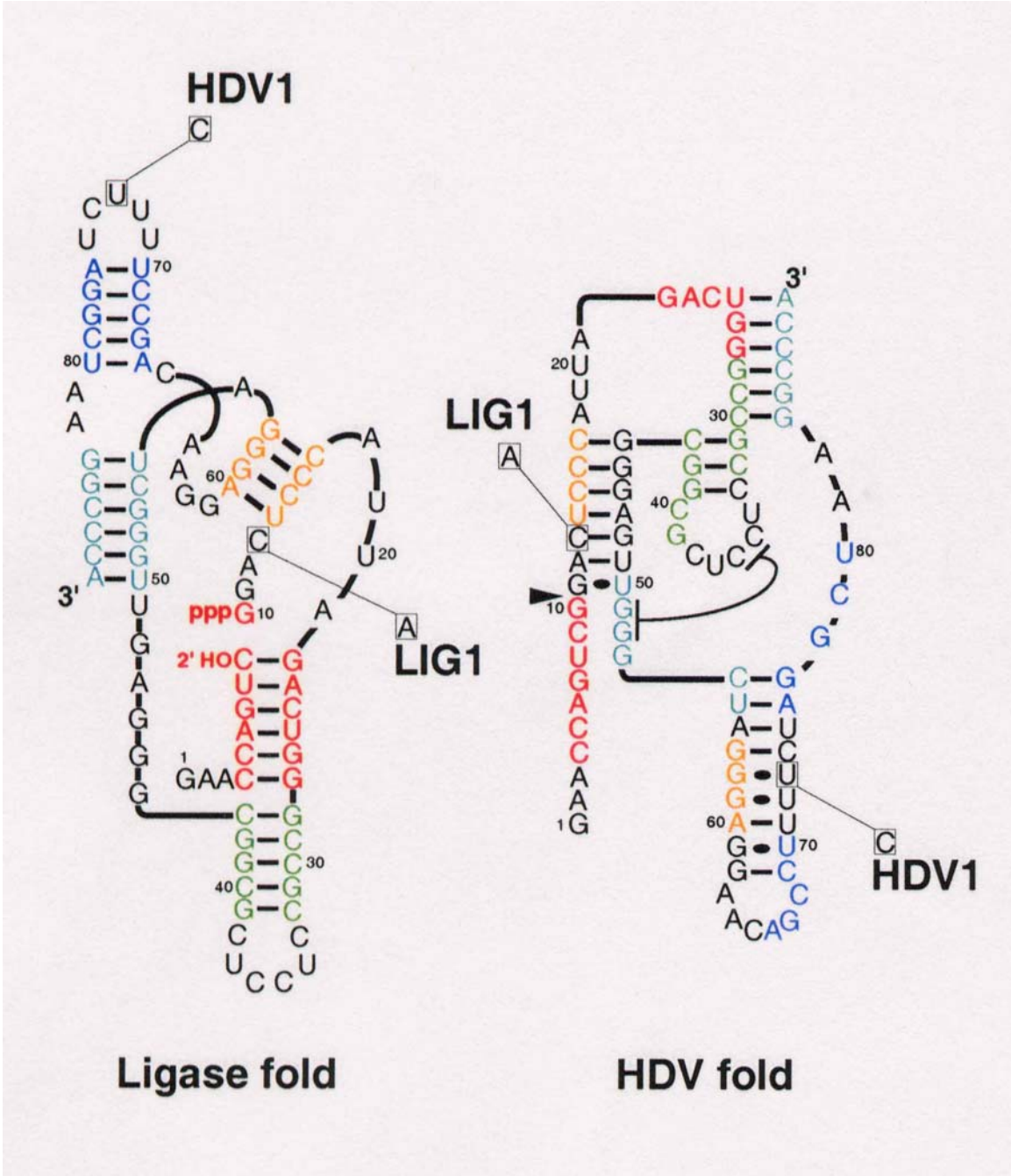
*To whom correspondence should be addressed. E-mail: dbartel@wi.mit.edu

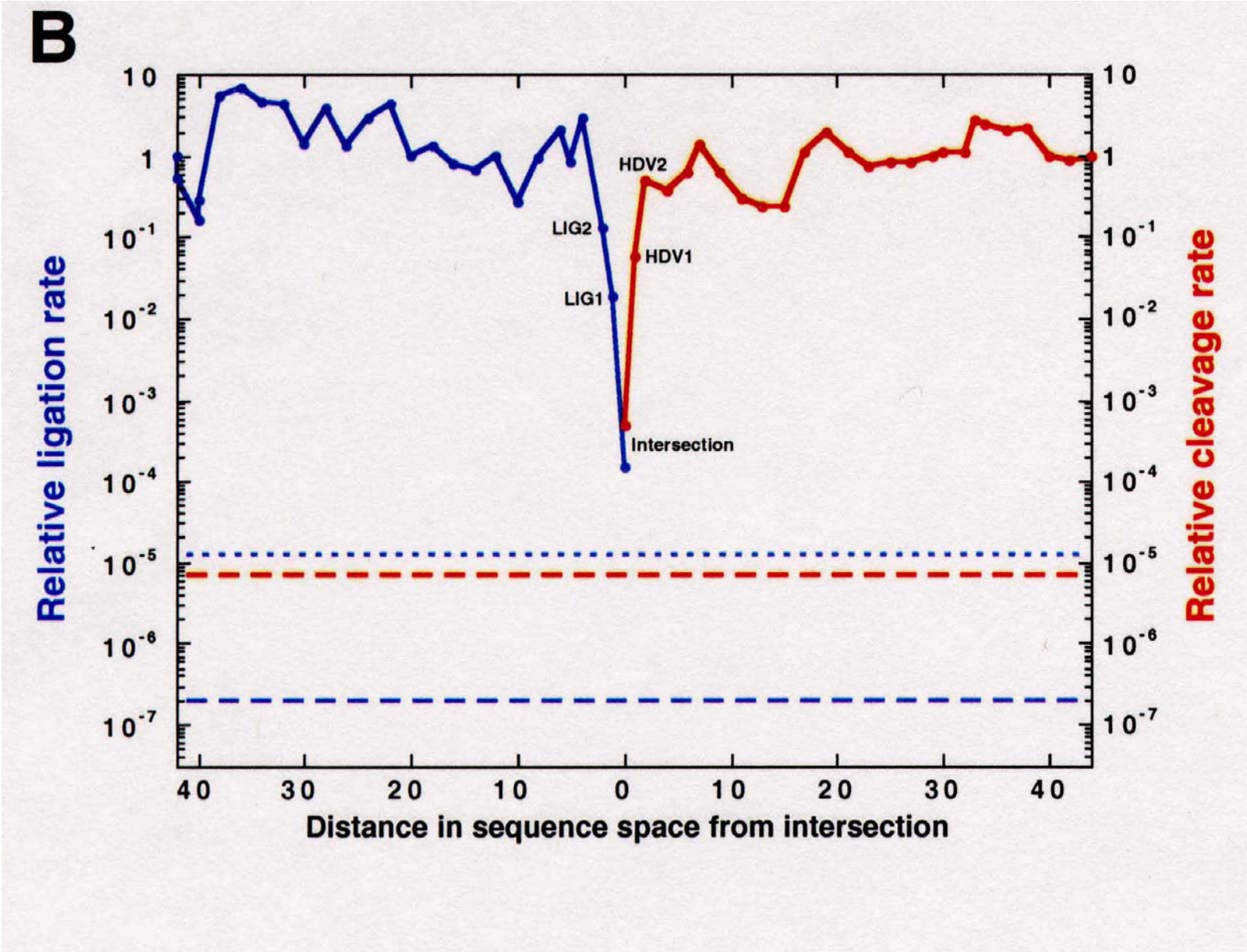


Two ribozymes of chain lengths n = 88 nucleotides: An artificial ligase (**A**) and a natural cleavage ribozyme of hepatitis-δ-virus (**B**)

The sequence at the ***intersection***:

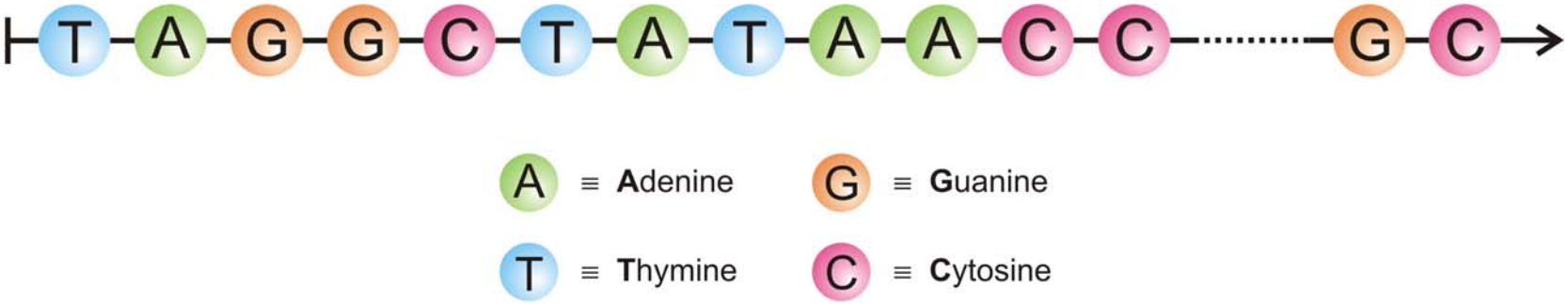
An RNA molecules which is 88 nucleotides long and can form both structures

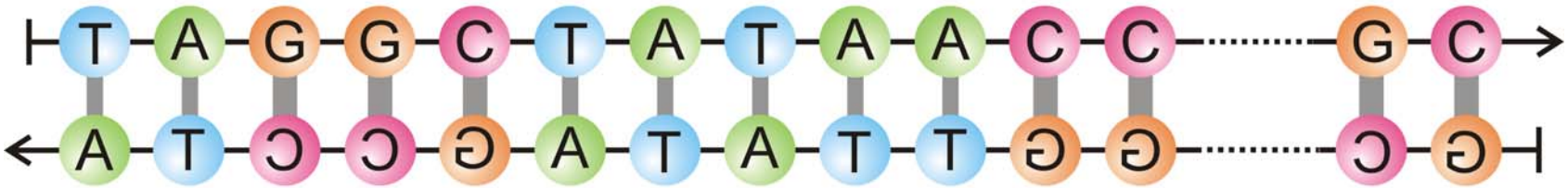




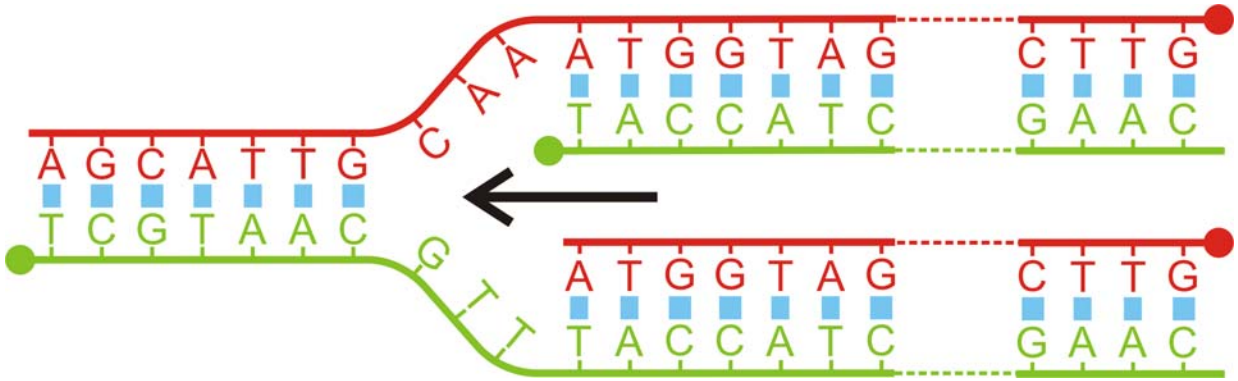
Two neutral walks through sequence space with conservation of structure and catalytic activity

1. Was ist Leben?
2. Chemische Evolution
3. Darwins tiefe Einsichten
4. Der Ursprung biologischer Information
5. Darwinsche Evolution mit Molekülen
6. Evolutionäre Biotechnologie
7. **Die DNA + Protein Welt**
8. Evolution bis zum Menschen



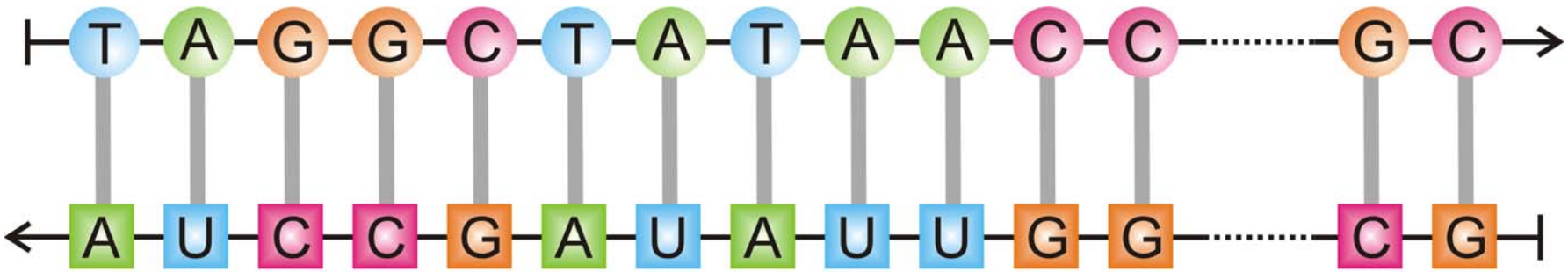


Deoxyribonucleic acid - DNA

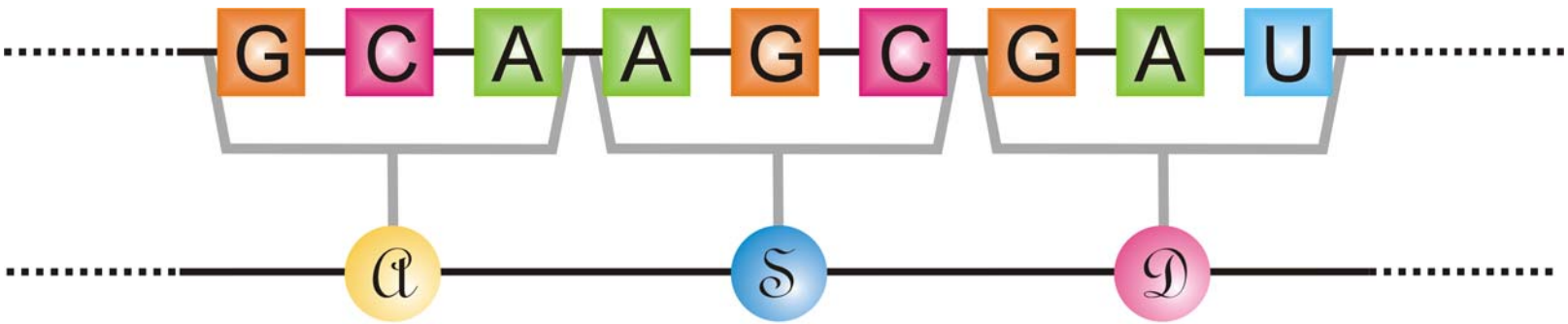


Die “Replikationsgabel”

Mechanismus der Replikation von doppelsträngigen DNA-Molekülen

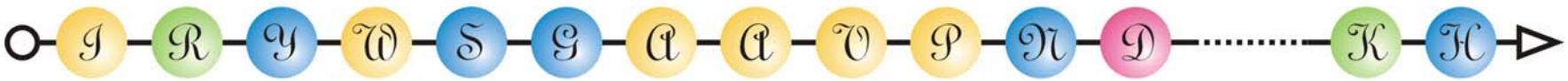


Transcription - DNA → RNA



Translation - RNA → Protein

Redundancy of the code: 4^3 = **64** codons versus **20** amino acids



𝔄 ≡ **ala**nine

𝔅 ≡ **cys**teine

𝔇 ≡ **asp**artic acid

𝔈 ≡ **glu**tamic acid

𝔉 ≡ **phe**nyl alanine

𝔊 ≡ **gly**cine

ℌ ≡ **his**tidine

ℑ ≡ **iso**l**eu**cine

𝔎 ≡ **lys**ine

𝔏 ≡ **leu**cine

𝔐 ≡ **met**hionine

𝔑 ≡ **as**paragi**n**e

𝔓 ≡ **pro**line

𝔔 ≡ **glu**tami**n**e

ℜ ≡ **arg**inine

𝔖 ≡ **ser**ine

𝔗 ≡ **thr**eonine

𝔙 ≡ **val**ine

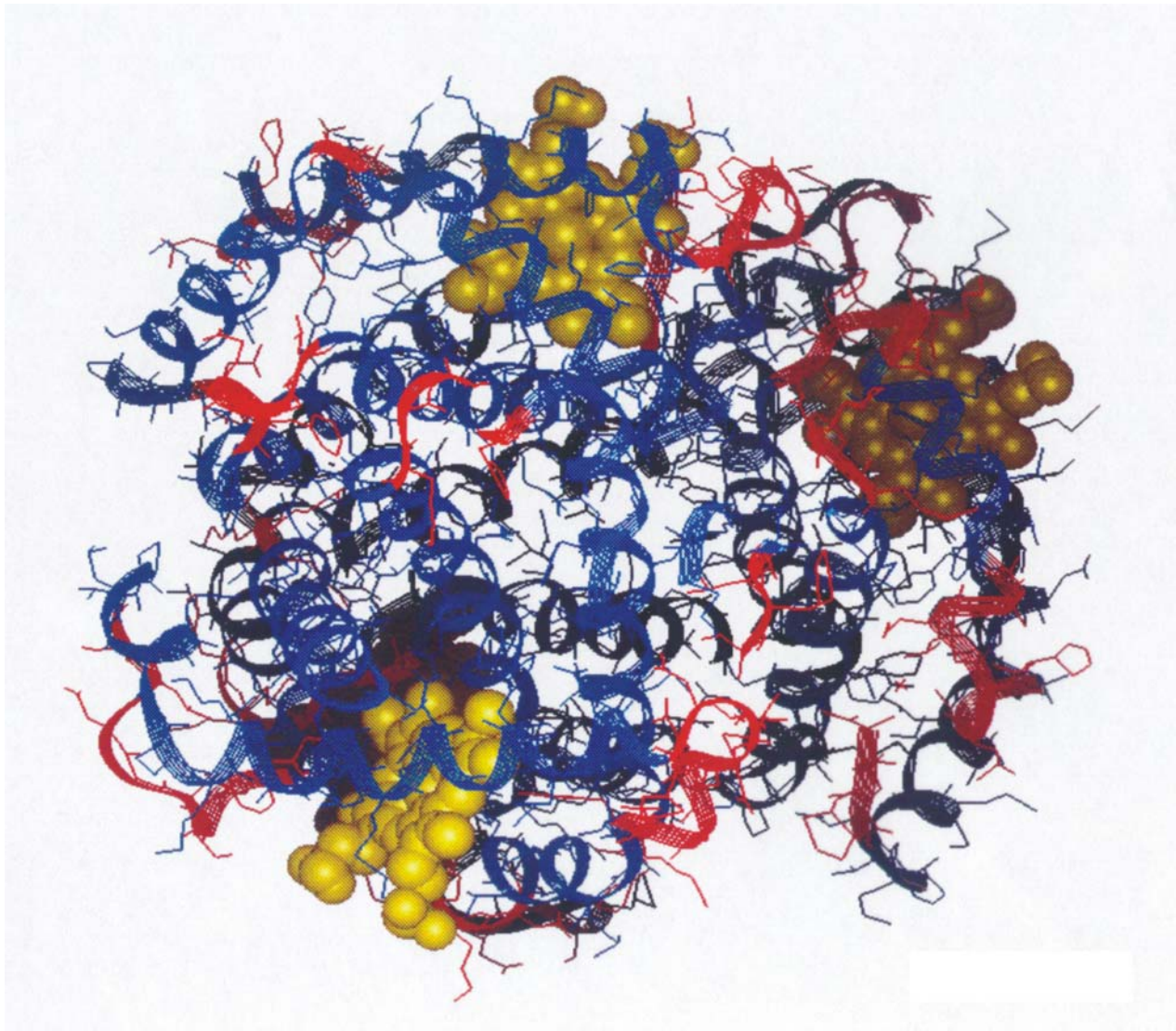
𝔚 ≡ **tryp**tophane

𝔜 ≡ **tyr**osine

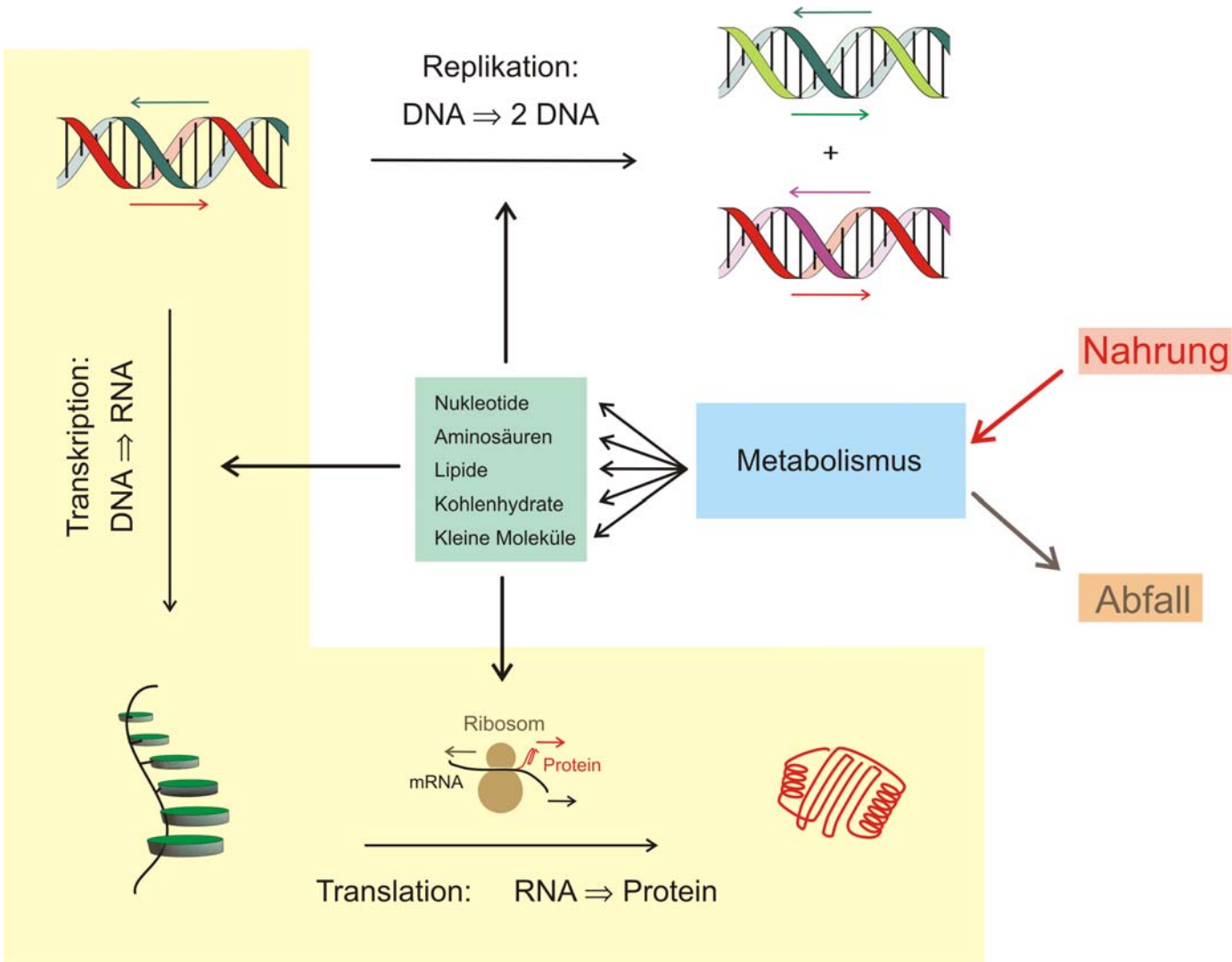
Protein

Max F. Perutz 1914-2002

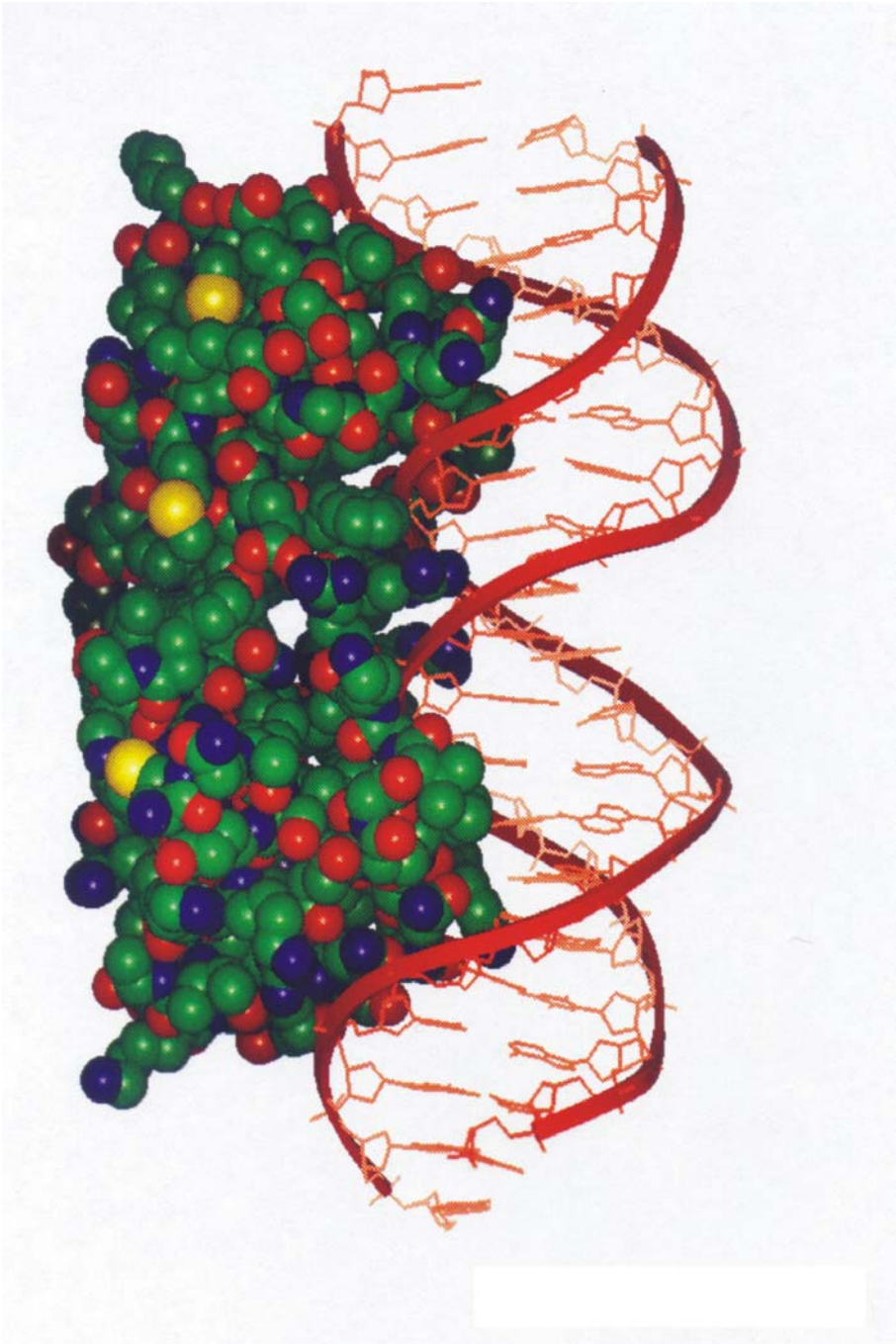
Nobel prize 1962



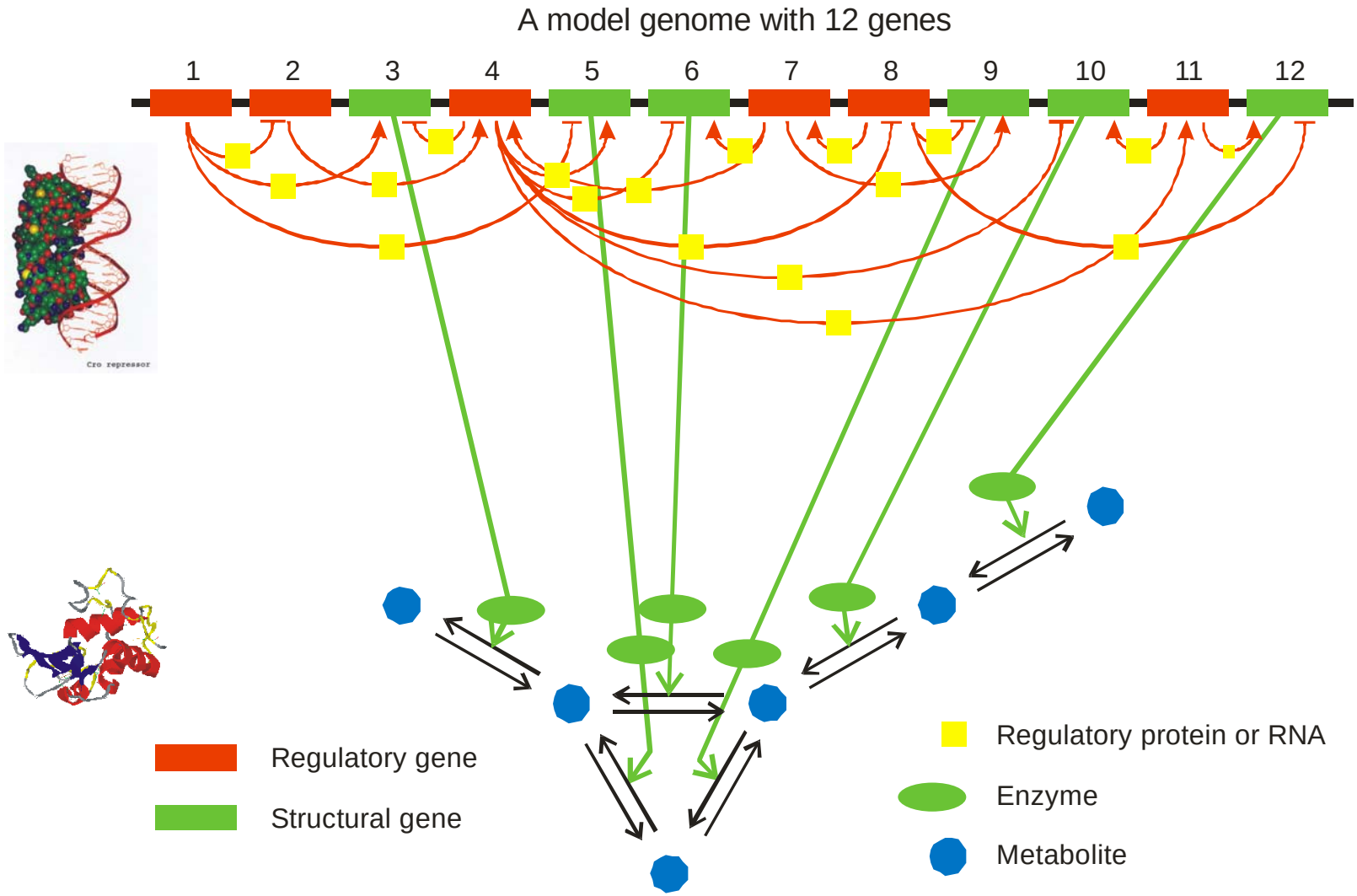
Three-dimensional
structure of hemoglobin



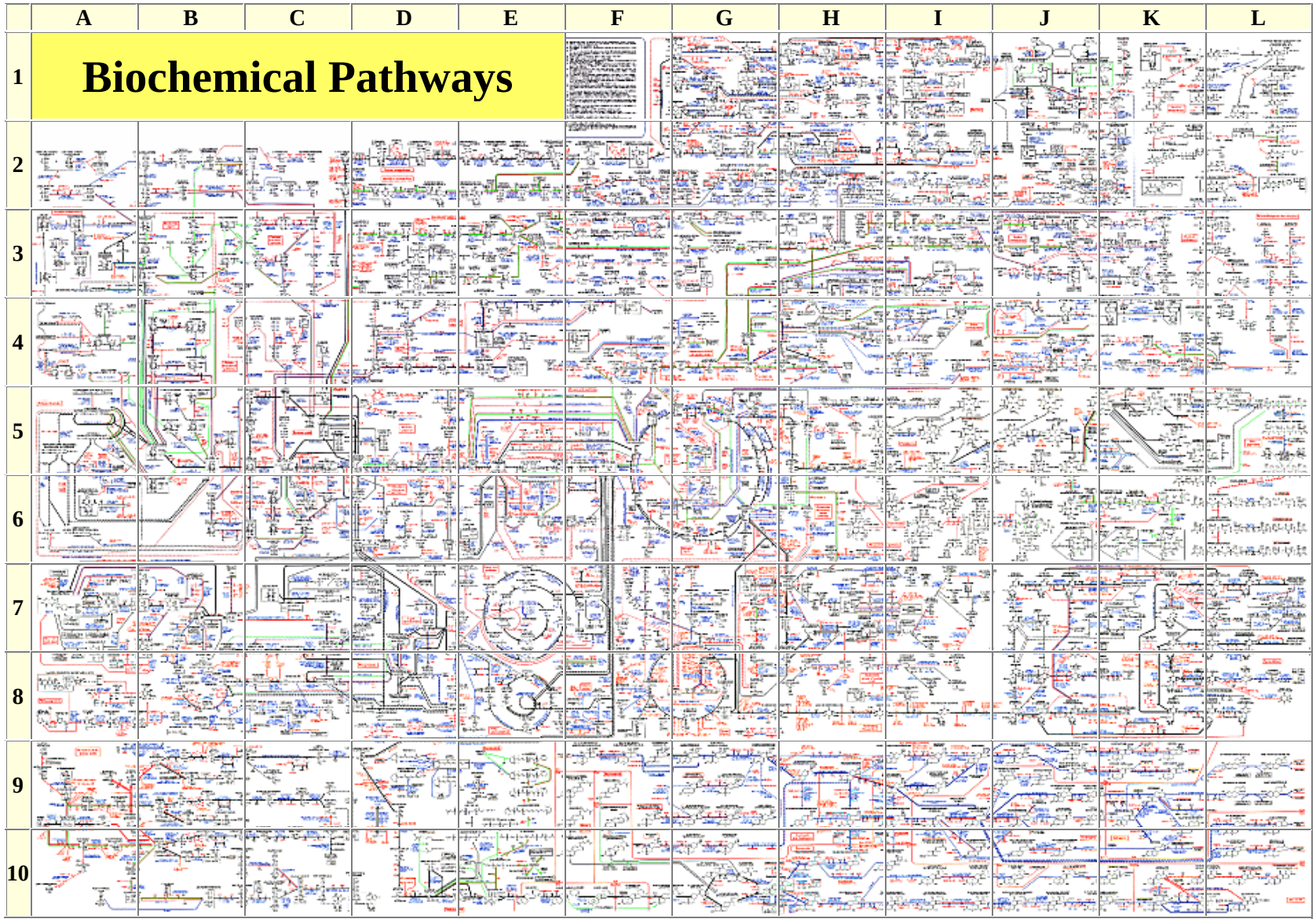
Skizze des zellulären Stoffwechsels



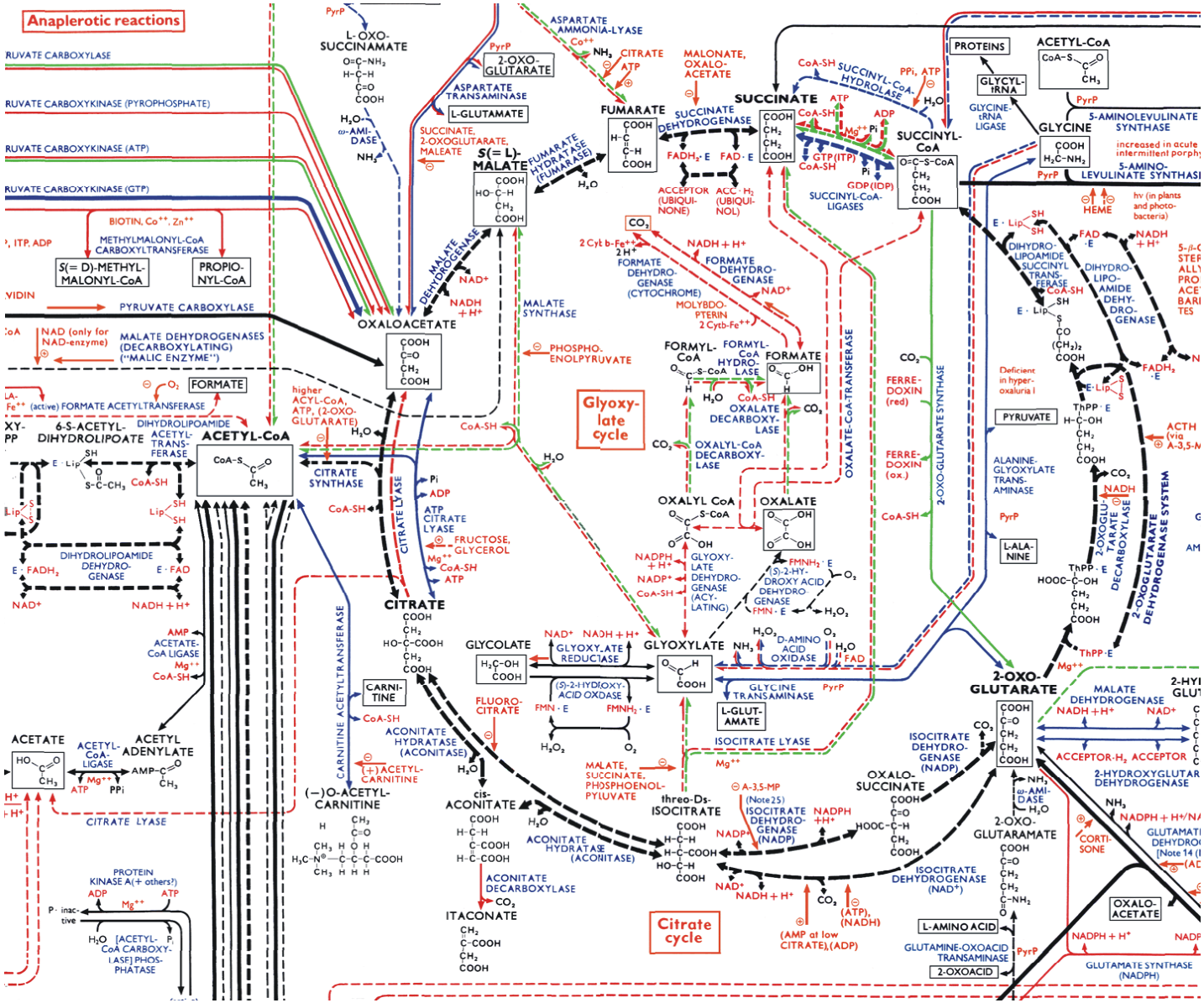
Three-dimensional structure of the complex between a specific DNA binding site and the regulatory protein cro-repressor



Skizze eines einfachen genetisch-metabolischen Regulationsnetzwerkes



Das Reaktionsnetzwerk des zellulären Stoffwechsels publiziert von Boehringer-Ingelheim.



Der Zitronensäure- oder Krebszyklus (vergrößert aus der vorigen Abbildung).

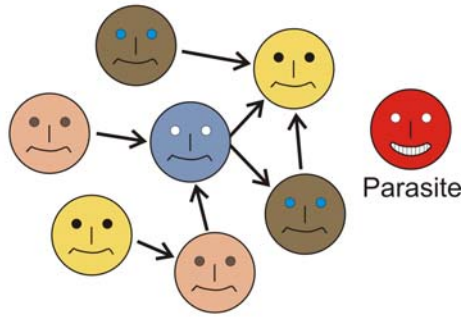
1. Was ist Leben?
2. Chemische Evolution
3. Darwins tiefe Einsichten
4. Der Ursprung biologischer Information
5. Darwinsche Evolution mit Molekülen
6. Evolutionäre Biotechnologie
7. Die DNA + Protein Welt
8. **Evolution bis zum Menschen**

Die großen Evolutionsschritte (nach John Maynard Smith und Eörs Szathmáry)

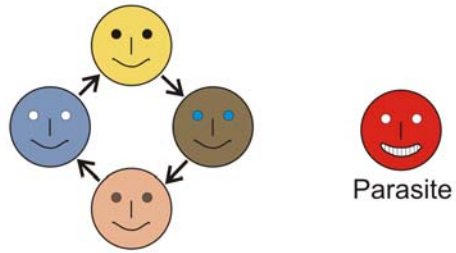
	Membranen, organisierte Teilung		
Replizierende Moleküle	⇒	Moleküle in Kompartments	
	Molekülverkettung, gemeinsame Replikation		
Unabhängige Replikatoren	⇒	Chromosomen	
	genetischer Code, Ribosom		
RNA als Gen und Enzyme	⇒	DNA und Protein	
	Zusammenschluß durch Endosymbiose		
Prokaryoten	⇒	Eukaryoten	
	Ursprung der sexuellen Vermehrung		
Asexuell vermehrende Klone	⇒	Sexuell vermehrende Populationen	
	Zelldifferenzierung und Entwicklung		
Protisten	⇒	Pflanzen, Pilze und Tiere	
	Entstehung nicht-reproduktiver Kasten		
Einzeln lebende Individuen	⇒	Tierkolonien	
	Sprache, Schrift, Kultur, …		
Primatengesellschaften	⇒	menschliche Gesellschaften	

Stufe I:
Unabhängige Replikatoren
in Konkurrenz

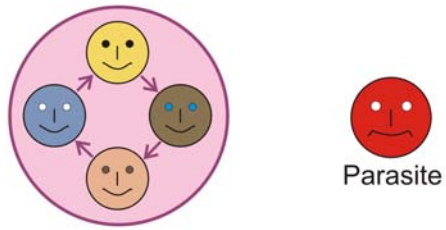
Stufe II:
Katalyse und Konkurrenz
bei der Replikation



Stufe III:
Funktionell verknüpfte
Replikatoren



Stufe IV:
Neue Einheit der
Selektion



Stufe V:
Unabhängige Einheiten
in Konkurrenz

Ein Mechanismus zur Überwindung
hierarchischer Stufen in der Evolution
(nach Manfred Eigen und Peter Schuster)

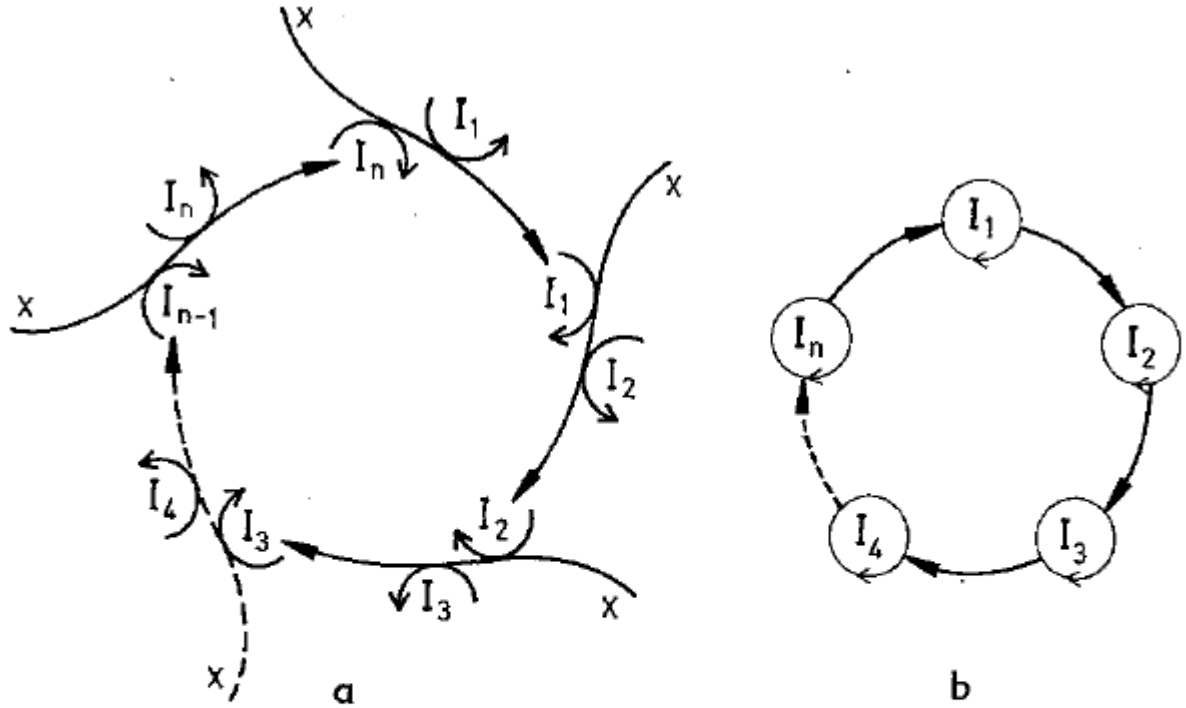
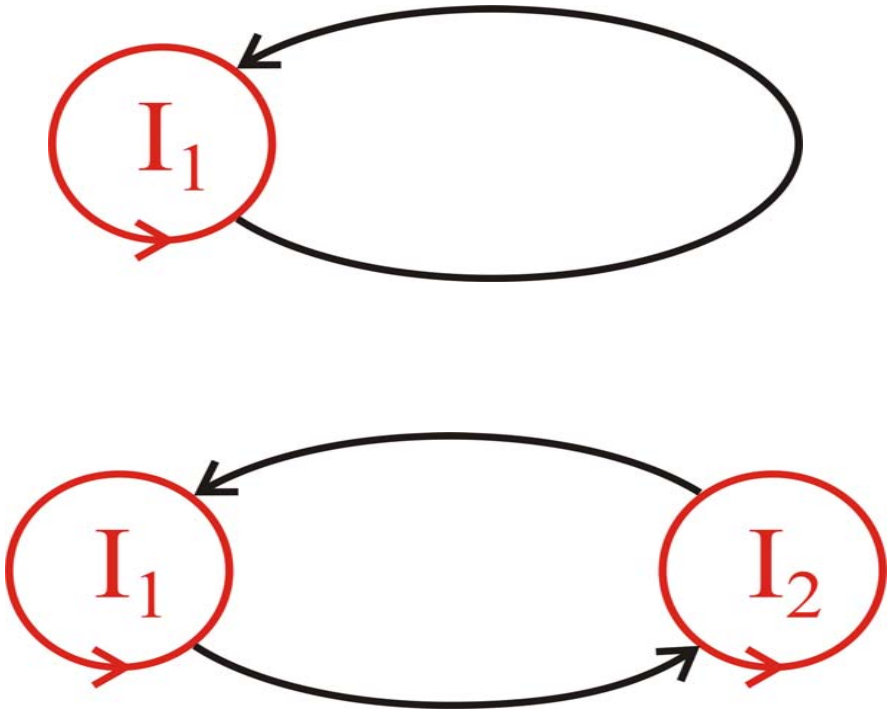
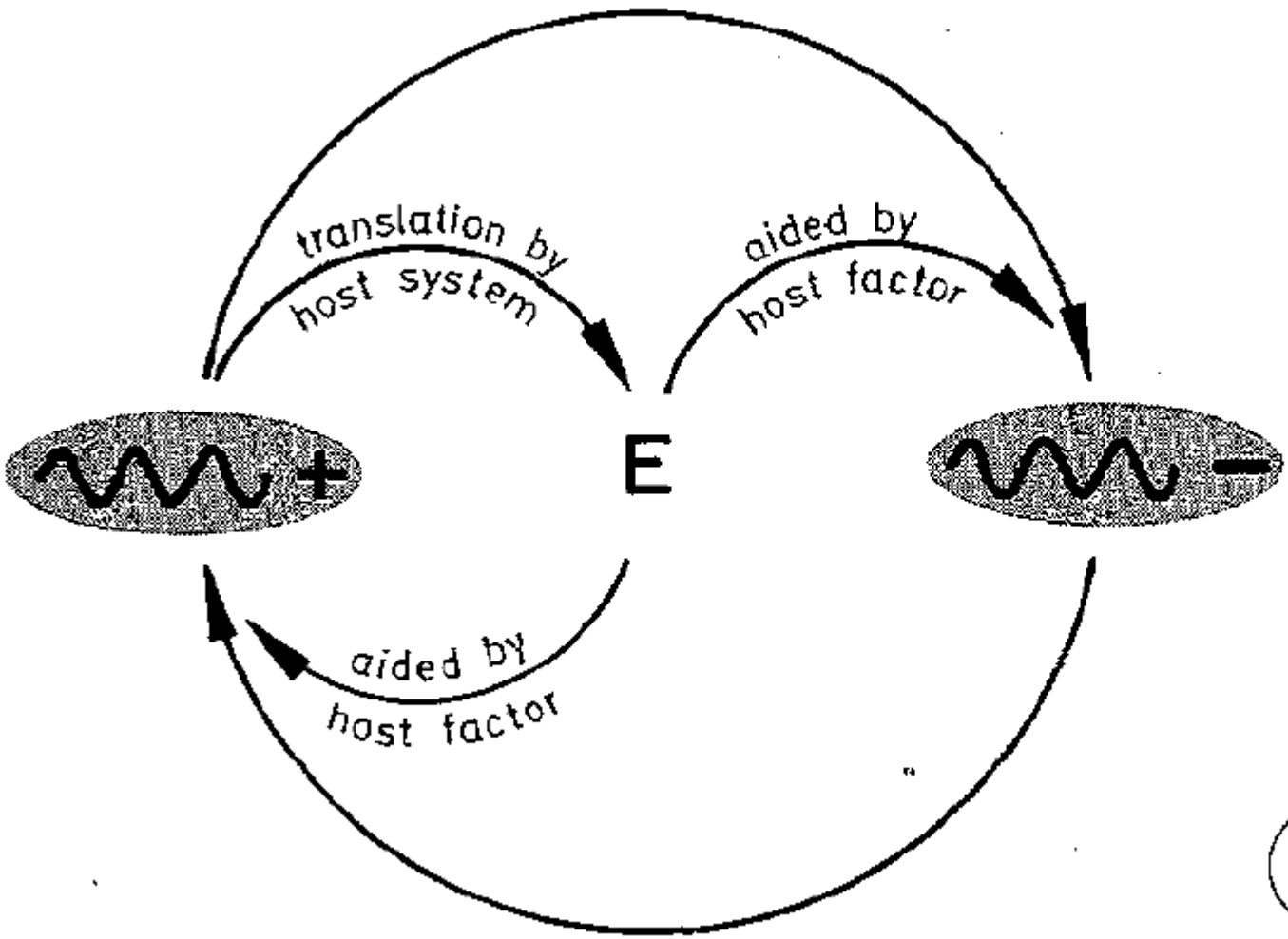


Fig. 7. *A catalytic hypercycle* consists of self-instructive units I_i with two-fold catalytic functions. As autocatalysts or—more generally—as catalytic cycles the intermediates I_i are able to instruct their own reproduction and, in addition, provide catalytic support for the reproduction of the subsequent intermediate (using the energy-rich building material X). The simplified graph (b) indicates the cyclic hierarchy



Hypercycles with one and two members are common in nature.


translation by
host system
aided by
host factor
+
E
−
aided by
host factor

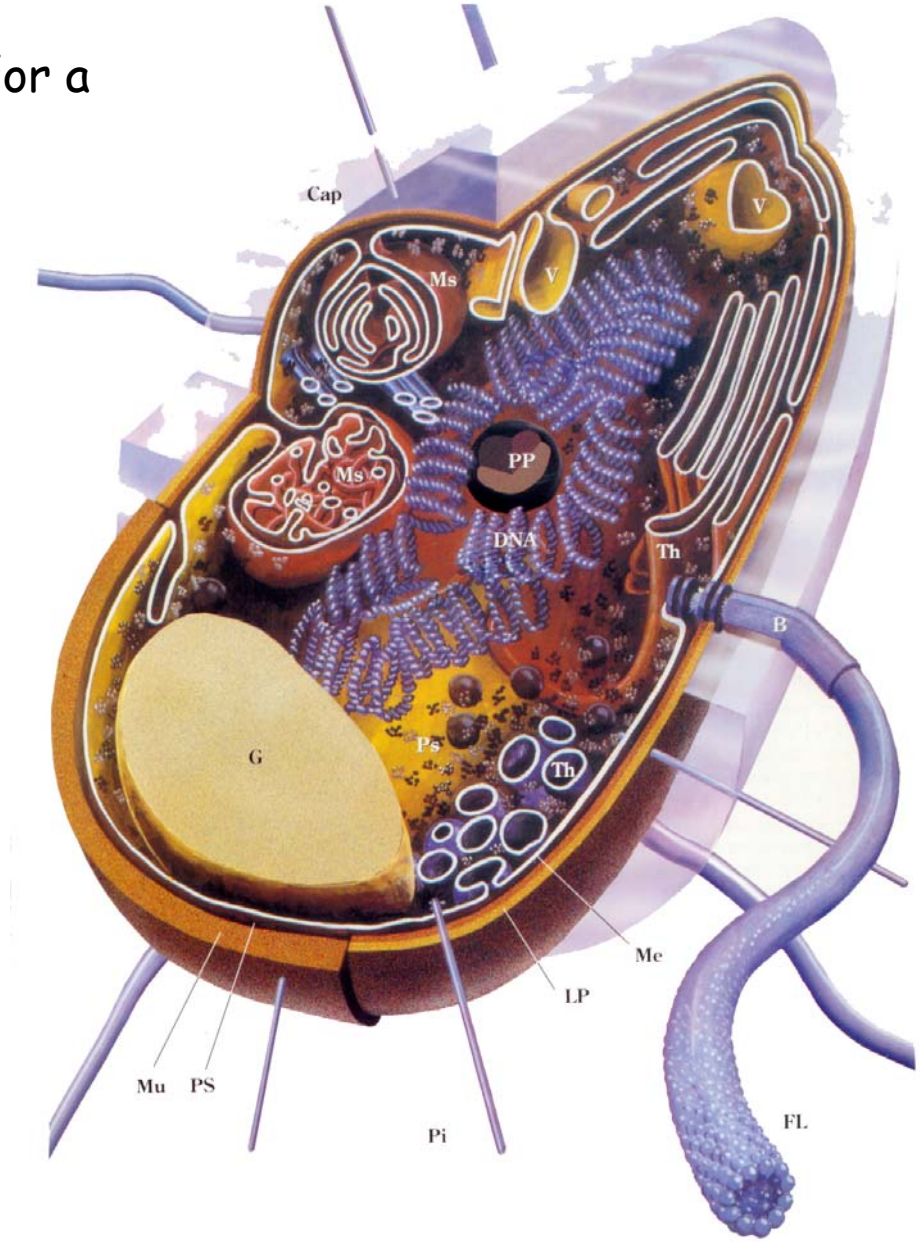

I
E

The bacterial cell as an example for a simple form of autonomous life

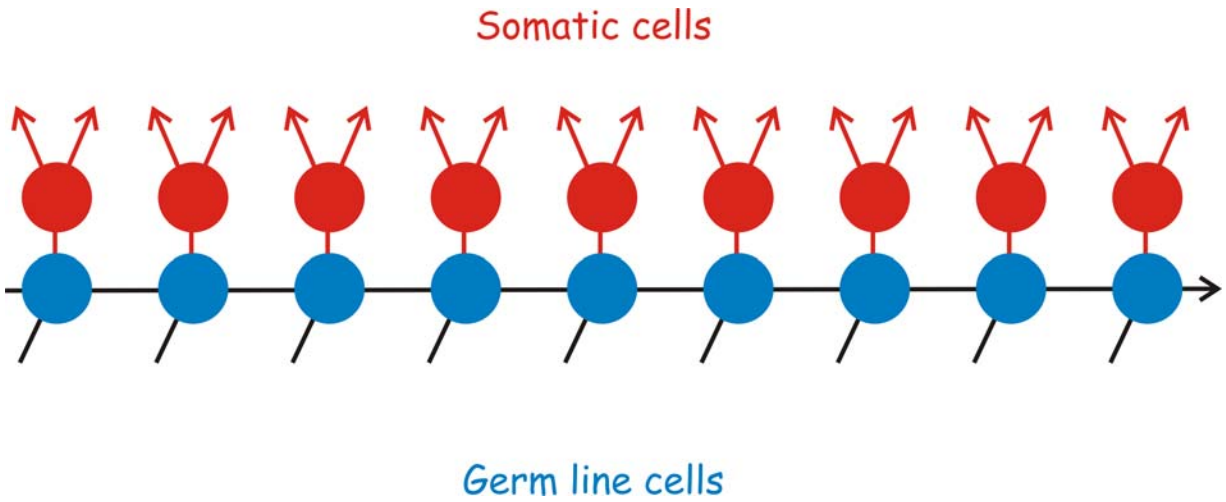
Escherichia coli genome:

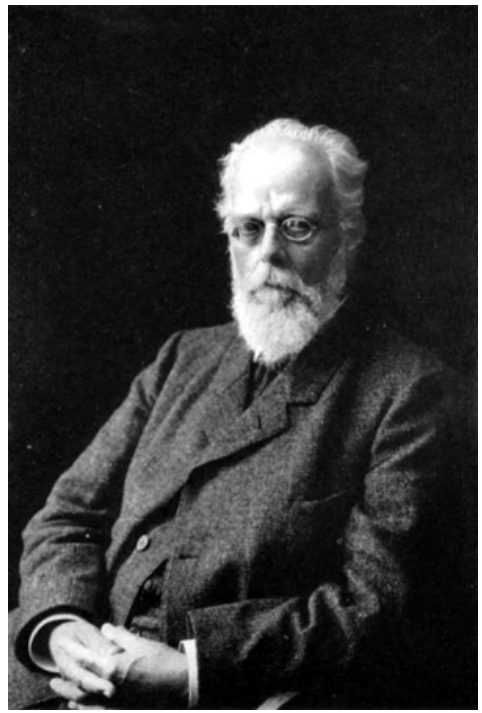
4 million nucleotides

4460 genes



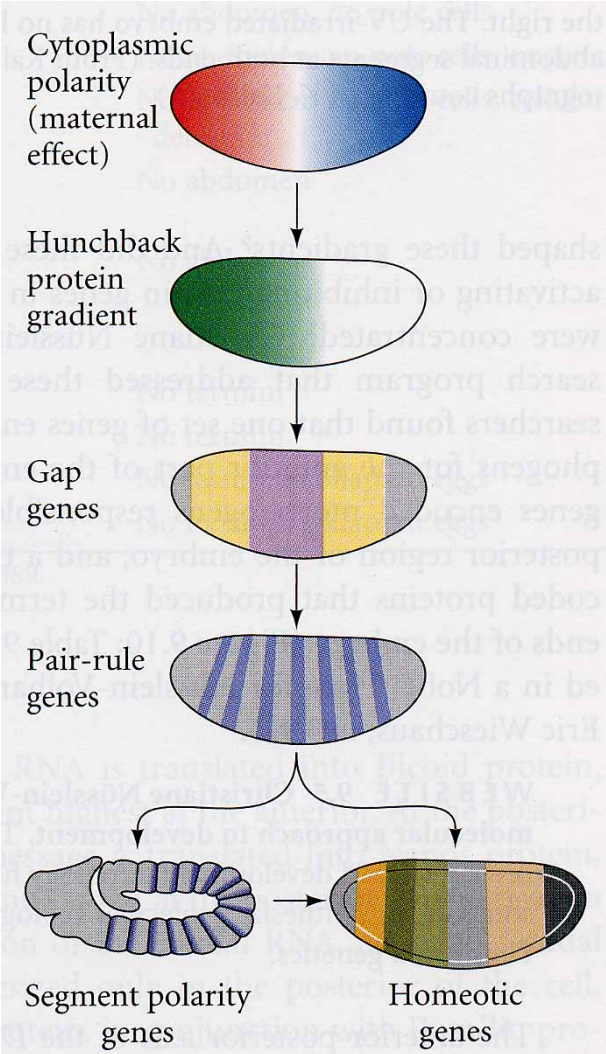
The structure of the bacterium *Escherichia coli*

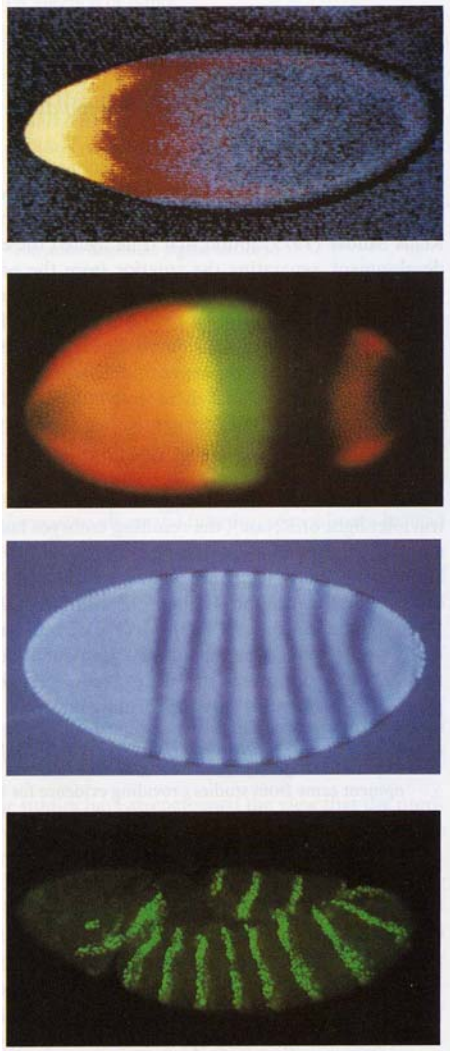


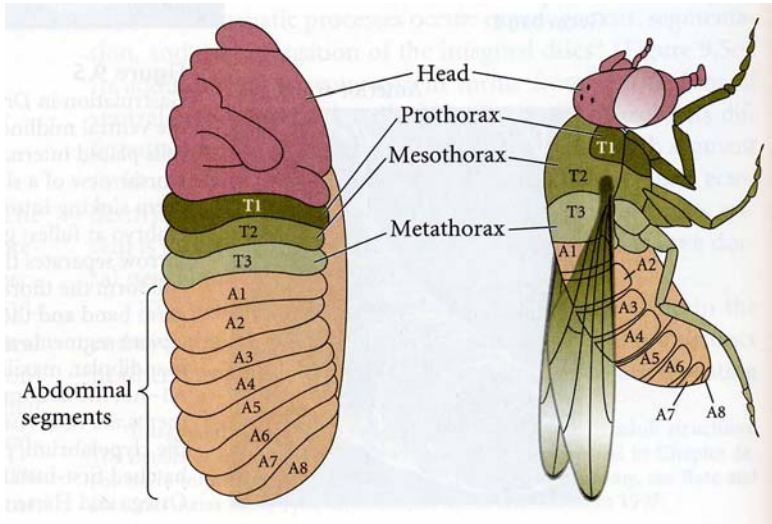


August Weismann, 1834-1914

Separation of germ line and soma







Cascades, A ⇒ B ⇒ C ⇒ ... , and networks of genetic control

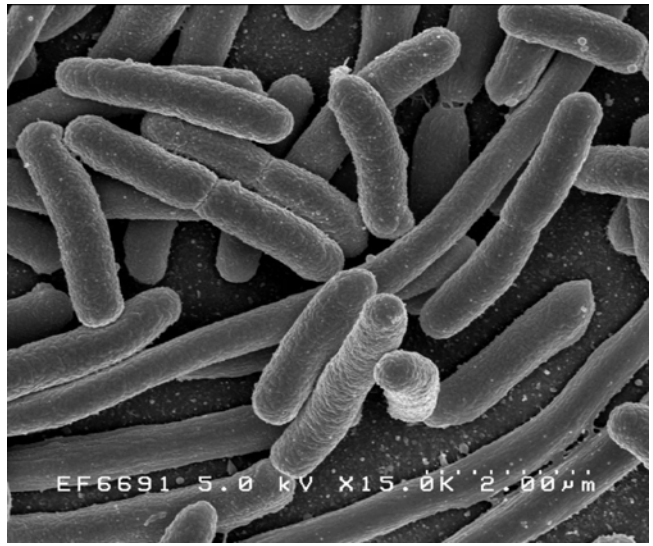
Turing pattern resulting from reaction-diffusion equation ?

Intercelluar communication creating positional information

Development of the fruit fly *drosophila melanogaster*: Genetics, experiment, and imago

E. coli:	Genome length	4×10^6 nucleotides
	Number of cell types	1
	Number of genes	4 460

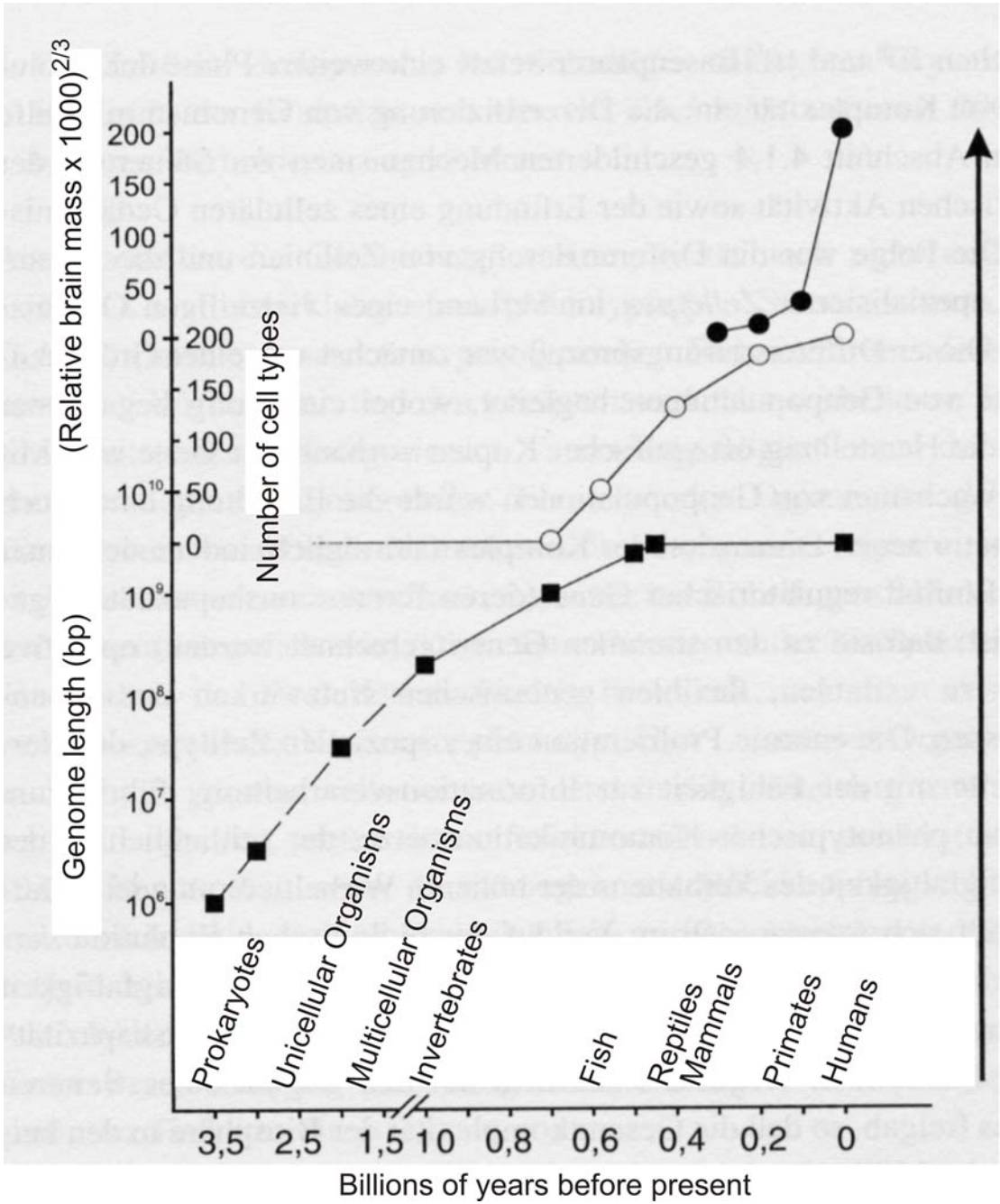
Four books, 300 pages each



Man:	Genome length	3×10^9 nucleotides
	Number of cell types	200
	Number of genes	≈ 30 000

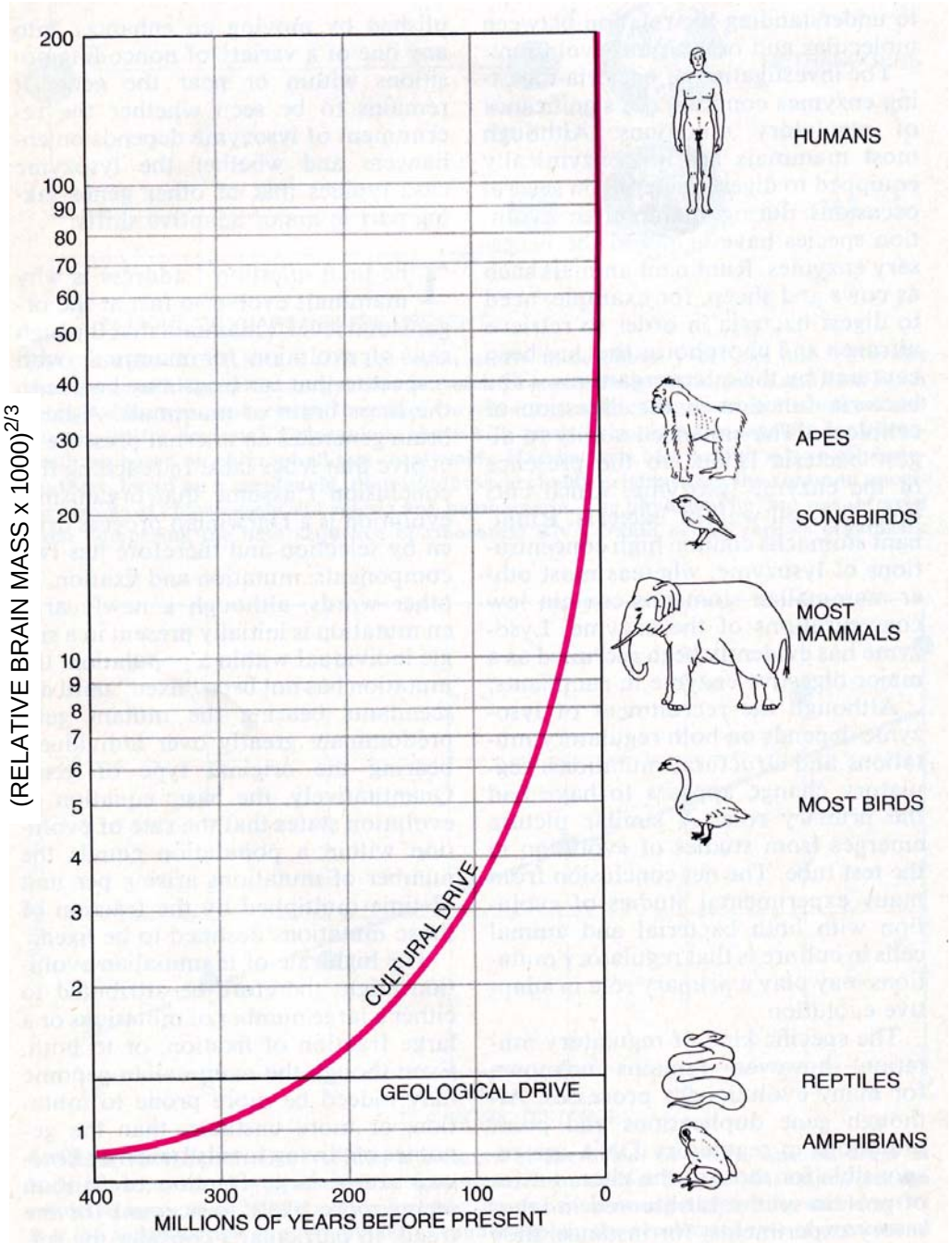
A library of 3000 volumes,
300 pages each

Complexity in biology

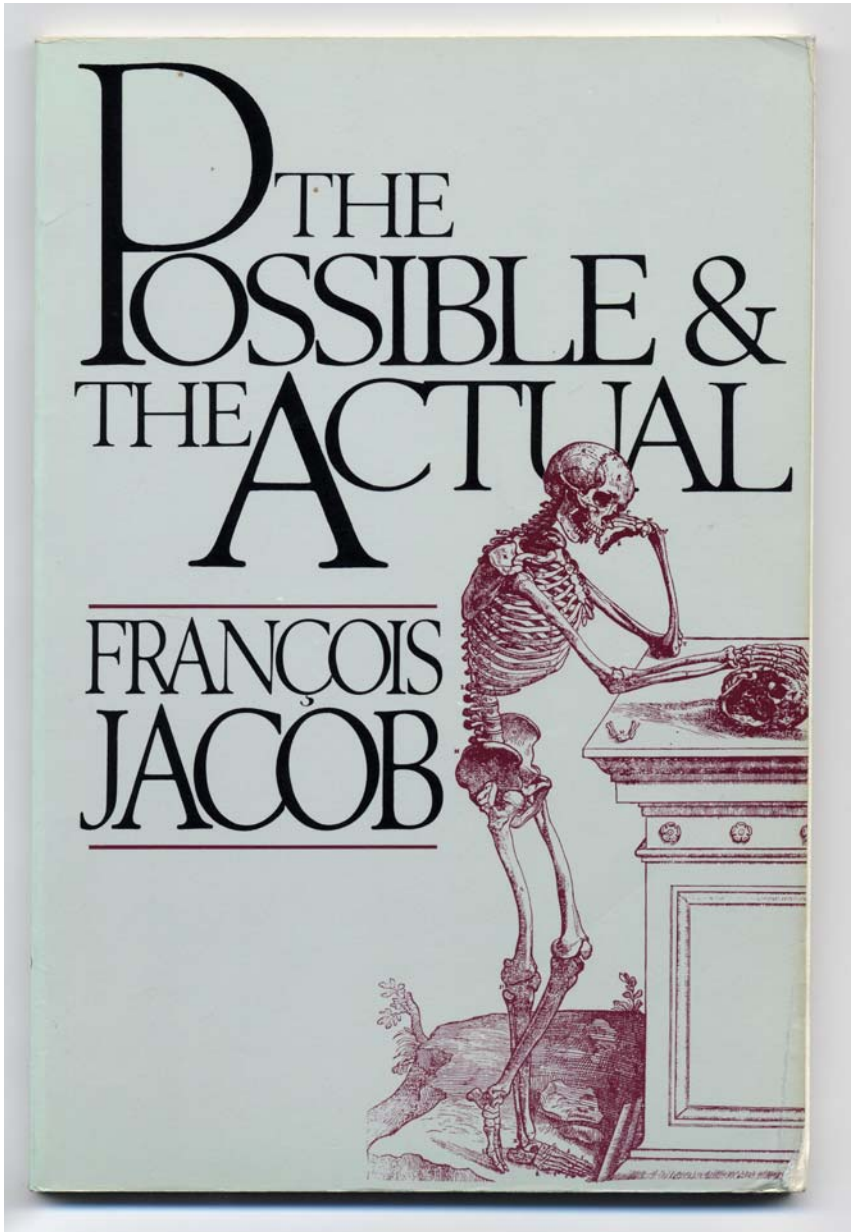


Wolfgang Wieser. 1998. ‚*Die Erfindung der Individualität*‘ oder ‚*Die zwei Gesichter der Evolution*‘. Spektrum Akademischer Verlag, Heidelberg 1998

BRITISH TIT



Alan C. Wilson.1985. The molecular basis of evolution.
Scientific American **253**(4):148-157.

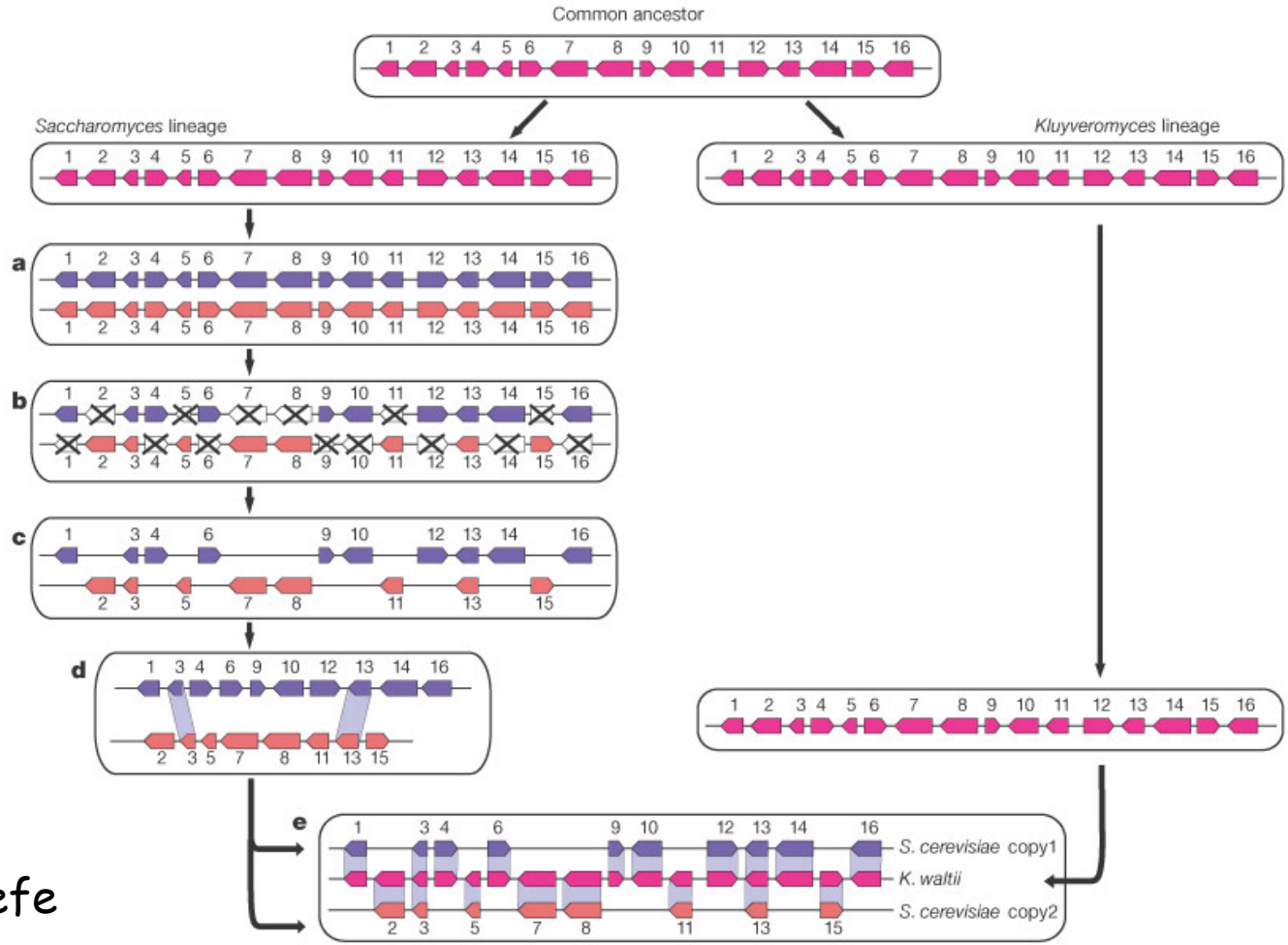


Evolution does not design with the eyes of an engineer, evolution works like a tinkerer.

François Jacob. *The Possible and the Actual*. Pantheon Books, New York, 1982, and

Evolutionary tinkering. *Science* **196** (1977), 1161-1166.

Ein Modell für die Genverdopplung in Hefe vor ≈ 1×10^8 Jahren



Manolis Kellis, Bruce W. Birren, and Eric S. Lander. Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*. *Nature* **428**: 617-624, 2004

The difficulty to define the notion of „gene".

Helen Pearson,
Nature **441**: 399-401, 2006

WHAT IS A GENE?

The idea of genes as beads on a DNA string is fast fading. Protein-coding sequences have no clear beginning or end and RNA is a key part of the information package, reports **Helen Pearson**.

'Gene' is not a typical four-letter word. It is not offensive. It is never bleeped out of TV shows. And where the meaning of most four-letter words is all too clear, that of gene is not. The more expert scientists become in molecular genetics, the less easy it is to be sure about what, if anything, a gene actually is.

Rick Young, a geneticist at the Whitehead Institute in Cambridge, Massachusetts, says that when he first started teaching as a young professor two decades ago, it took him about two hours to teach fresh-faced undergraduates what a gene was and the nuts and bolts of how it worked. Today, he and his colleagues need three months of lectures to convey the concept of the gene, and that's not because the students are any less bright. "It takes a whole semester to teach this stuff to talented graduates," Young says. "It used to be we could give a one-off definition and now it's much more complicated."

In classical genetics, a gene was an abstract concept — a unit of inheritance that ferried a characteristic from parent to child. As biochemistry came into its own, those characteristics were associated with enzymes or proteins, one for each gene. And with the advent of molecular biology, genes became real, physical things — sequences of DNA which when converted into strands of so-called messenger RNA could be used as the basis for building their associated protein piece by piece. The great coiled DNA molecules of the chromosomes were seen as long strings on which gene sequences sat like discrete beads.

This picture is still the working model for many scientists. But those at the forefront of genetic research see it as increasingly old-fashioned — a crude approximation that, at best, hides fascinating new complexities and, at worst, blinds its users to useful new paths of enquiry.

Information, it seems, is parceled out along chromosomes in a much more complex way than was originally supposed. RNA molecules are not just passive conduits through which the gene's message flows into the world but active regulators of cellular processes. In some cases, RNA may even pass information across generations — normally the sole preserve of DNA.

An eye-opening study last year raised the possibility that plants sometimes rewrite their DNA on the basis of RNA messages inherited from generations past[1]. A study on page 469 of this issue suggests that a comparable phenomenon might occur in mice, and by implication in other mammals[2]. If this type of phenomenon is indeed widespread, it "would have huge implications," says evolutionary geneticist Laurence Hurst at the University of Bath, UK.

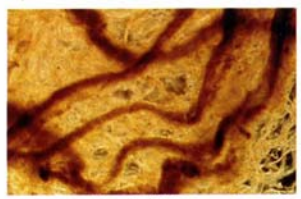
"All of that information seriously challenges our conventional definition of a gene," says molecular biologist Bing Ren at the University of California, San Diego. And the information challenge is about to get even tougher. Later this year, a glut of data will be released from the international Encyclopedia of DNA Elements (ENCODE) project. The pilot phase of ENCODE involves scrutinizing roughly 1% of the human genome in unprecedented detail; the aim is to find all the sequences that serve a useful purpose and explain what that purpose is. "When we started the ENCODE project I had a different view of what a gene was," says contributing researcher Roderic Guigo at the Center for Genomic Regulation in Barcelona. "The degree of complexity we've seen was not anticipated."

"We've come to the realization that the genome is full of overlapping transcripts."
— Phillip Kapranov

Under fire

The first of the complexities to challenge molecular biology's paradigm of a single DNA sequence encoding a single protein was alternative splicing, discovered in viruses in 1977 (see 'Hard to track', overleaf). Most of the DNA sequences describing proteins in humans have a modular arrangement in which exons, which carry the instructions for making proteins, are interspersed with non-coding introns. In alternative splicing, the cell snips out introns and sews together the exons in various different orders, creating messages that can code for different proteins. Over the years geneticists have also documented overlapping genes, genes within genes and countless other weird arrangements (see 'Muddling over genes', overleaf).

Alternative splicing, however, did not in itself require a drastic reappraisal of the notion of a gene; it just showed that some DNA sequences could describe more than one protein. Today's assault on the gene concept is more far reaching, fuelled largely by studies that show the previously unimagined scope of RNA.



Spools of DNA (above) still harbour surprises, with one protein-coding gene often overlapping the next.

P. PLAILLY/SPL

The one gene, one protein idea is coming under particular assault from researchers who are comprehensively extracting and analysing the RNA messages, or transcripts, manufactured by genomes, including the human and mouse genome. Researchers led by Thomas Gingeras at the company Affymetrix in Santa Clara, California, for example, recently studied all the transcripts from ten chromosomes across eight human cell lines and worked out precisely where on the chromosomes each of the transcripts came from[3].

The picture these studies paint is one of mind-boggling complexity. Instead of discrete genes dutifully mass-producing identical RNA transcripts, a teeming mass of transcription converts many segments of the genome into multiple RNA ribbons of differing lengths. These ribbons can be generated from both strands of DNA, rather than from just one as was conventionally thought. Some of these transcripts come from regions of DNA previously identified as holding protein-coding genes. But many do not. "It's somewhat revolutionary," says Gingeras's colleague Phillip Kapranov. "We've come to the realization that the genome is full of overlapping transcripts."

Other studies, one by Guigo's team[4], and one by geneticist Rotem Sorek[5], now at Tel Aviv University, Israel, and his colleagues, have hinted at the reasons behind the mass of transcription. The two teams investigated occasional reports that transcription can start at a DNA sequence associated with one protein and run straight through into the gene for a completely different protein, producing a fused transcript. By delving into databases of human RNA transcripts, Guigo's team estimate that 4–5% of the DNA in regions conventionally recognized as genes is transcribed in this way. Producing fused transcripts could be one way for a cell to generate a greater variety of proteins from a limited number of exons, the researchers say.

Many scientists are now starting to think that the descriptions of proteins encoded in DNA know no borders — that each sequence reaches into the next and beyond. This idea will be one of the central points to emerge from the ENCODE project when its results are published later this year.

Kapranov and others say that they have documented many examples of transcripts in which protein-coding exons from one part of the genome combine with exons from another

ENCODE stands for
ENCyclopedia Of DNA Elements.

ENCODE Project Consortium.
Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project.
Nature **447**:799-816, 2007

Web-Page for further information:

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