What did the molecular connection contribute to an understanding of biological evolution?

Insights from Watson-Crick to systems biology

Peter Schuster

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

BioQuant Seminar

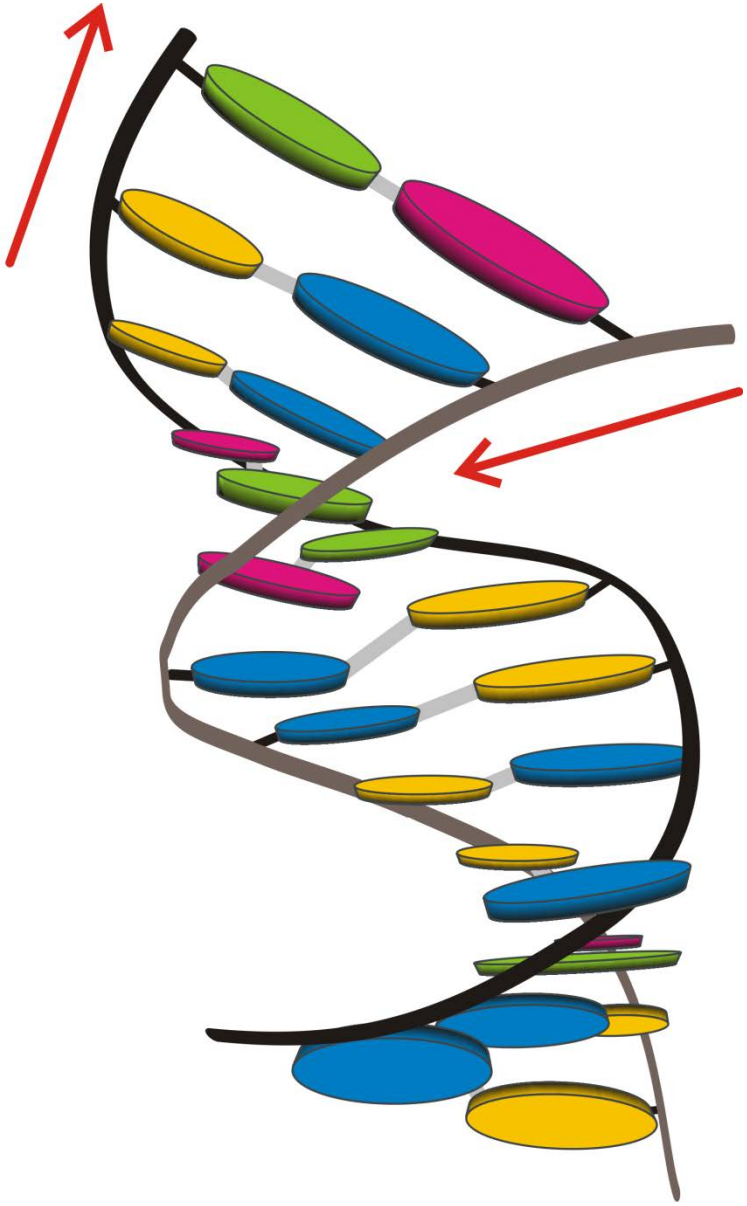
Heidelberg, 12.06.2012

Web-Page for further information:

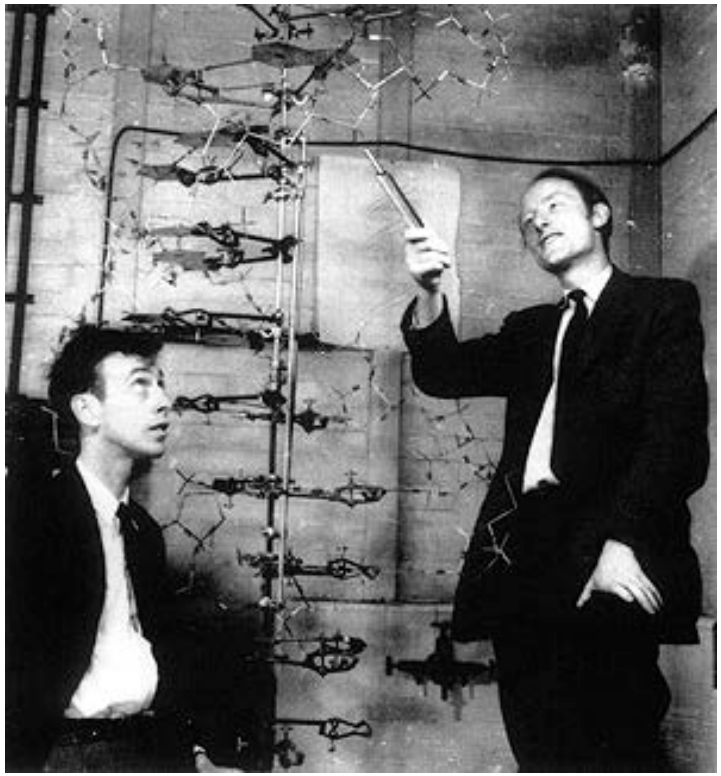
http://www.tbi.univie.ac.at/~pks

1. Prologue
2. Darwin and replicating molecules
3. In vitro evolution
4. „Simple" landscapes
5. „Realistic" landscapes
6. Neutrality in evolution
7. Perspectives of systems biology

1. **Prologue**
2. Darwin and replicating molecules
3. In vitro evolution
4. „Simple“ landscapes
5. „Realistic“ landscapes
6. Neutrality in evolution
7. Perspectives of systems biology



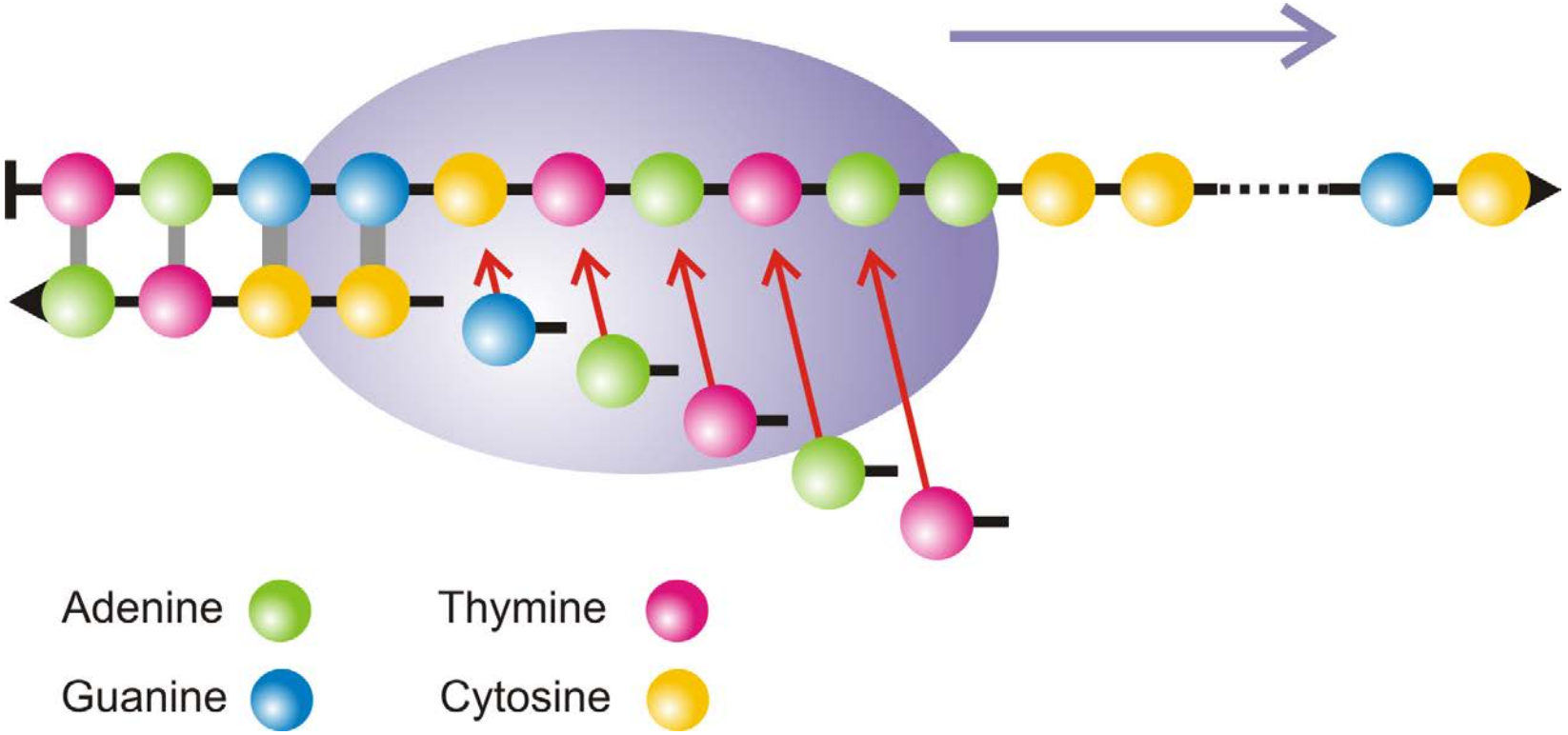
The three-dimensional structure of a short double helical stack of B-DNA



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

It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material.

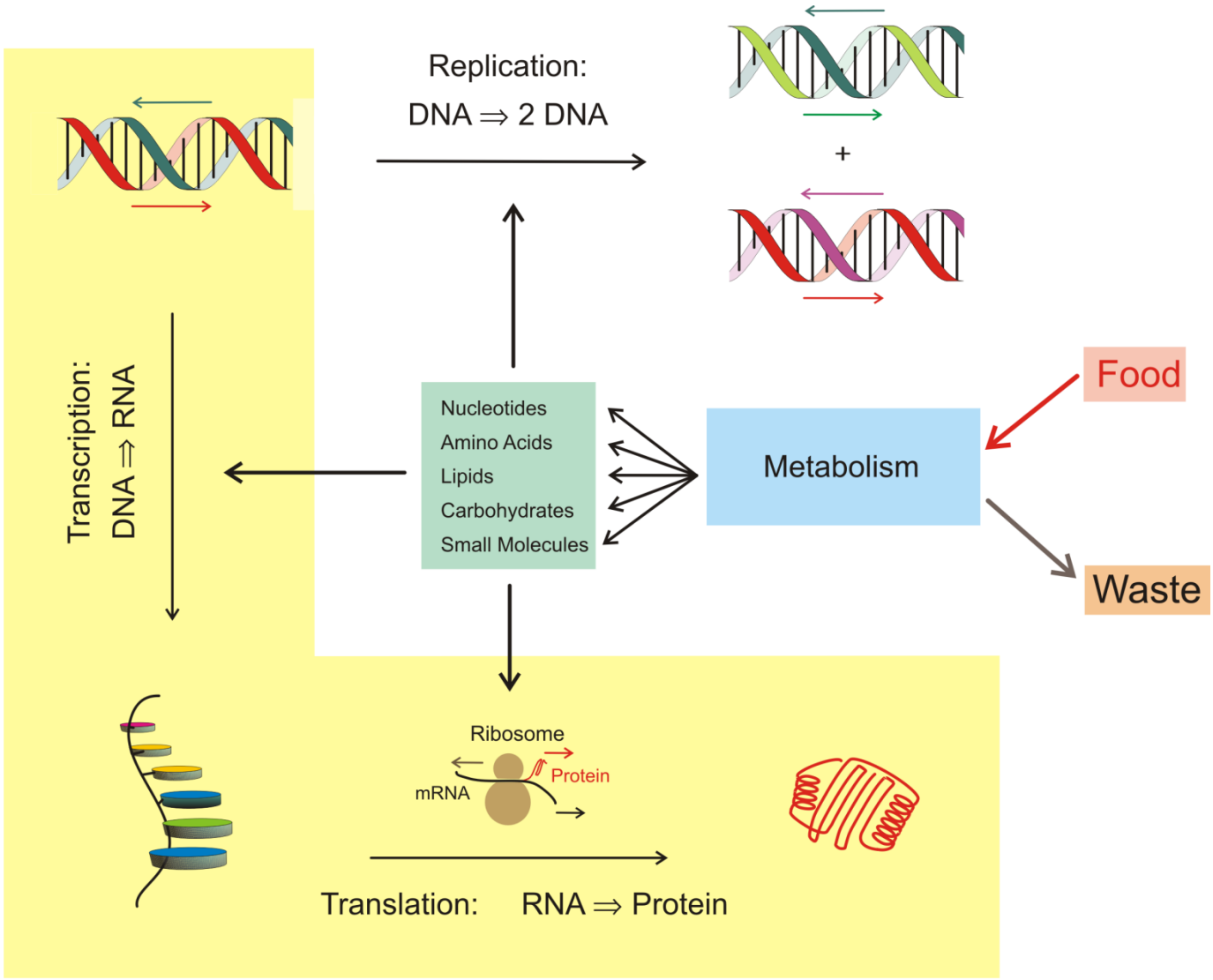
A Structure for Deoxyribose Nucleic Acid
Nature 171:737-738 (1953)



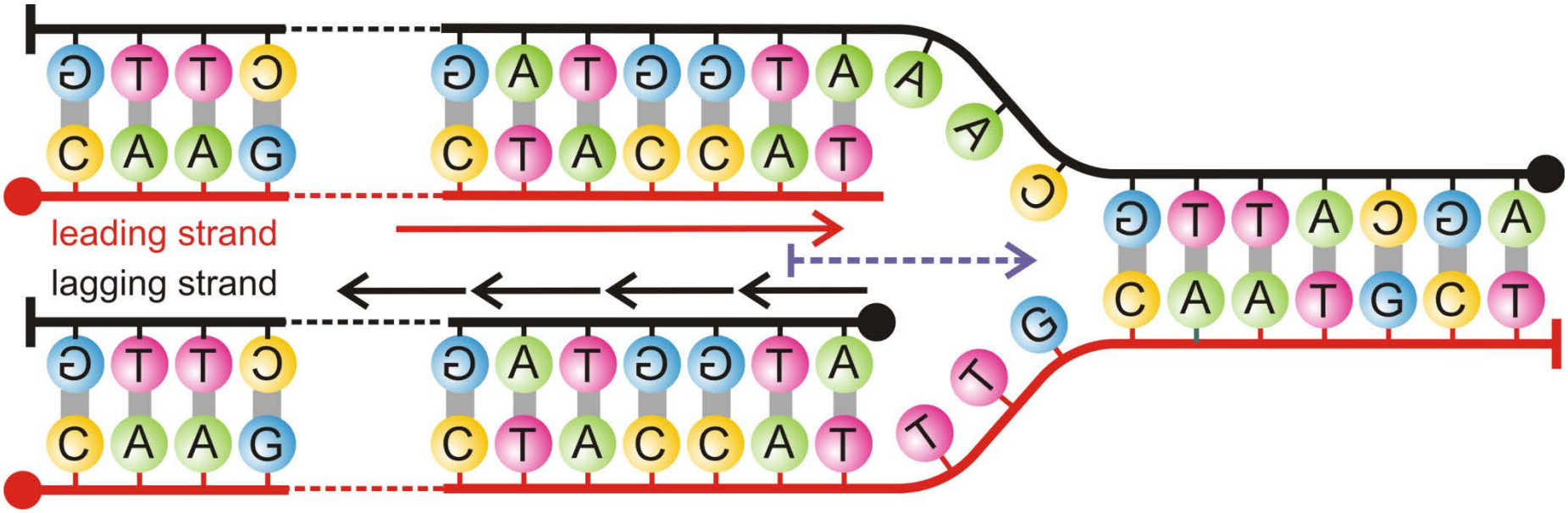
The replication of DNA by Thermophilus aquaticus polymerase (PCR)

Accuracy of replication: $Q = q_1 \cdot q_2 \cdot q_3 \cdot q_4 \cdot \ldots$

The logics of DNA (or RNA) replication



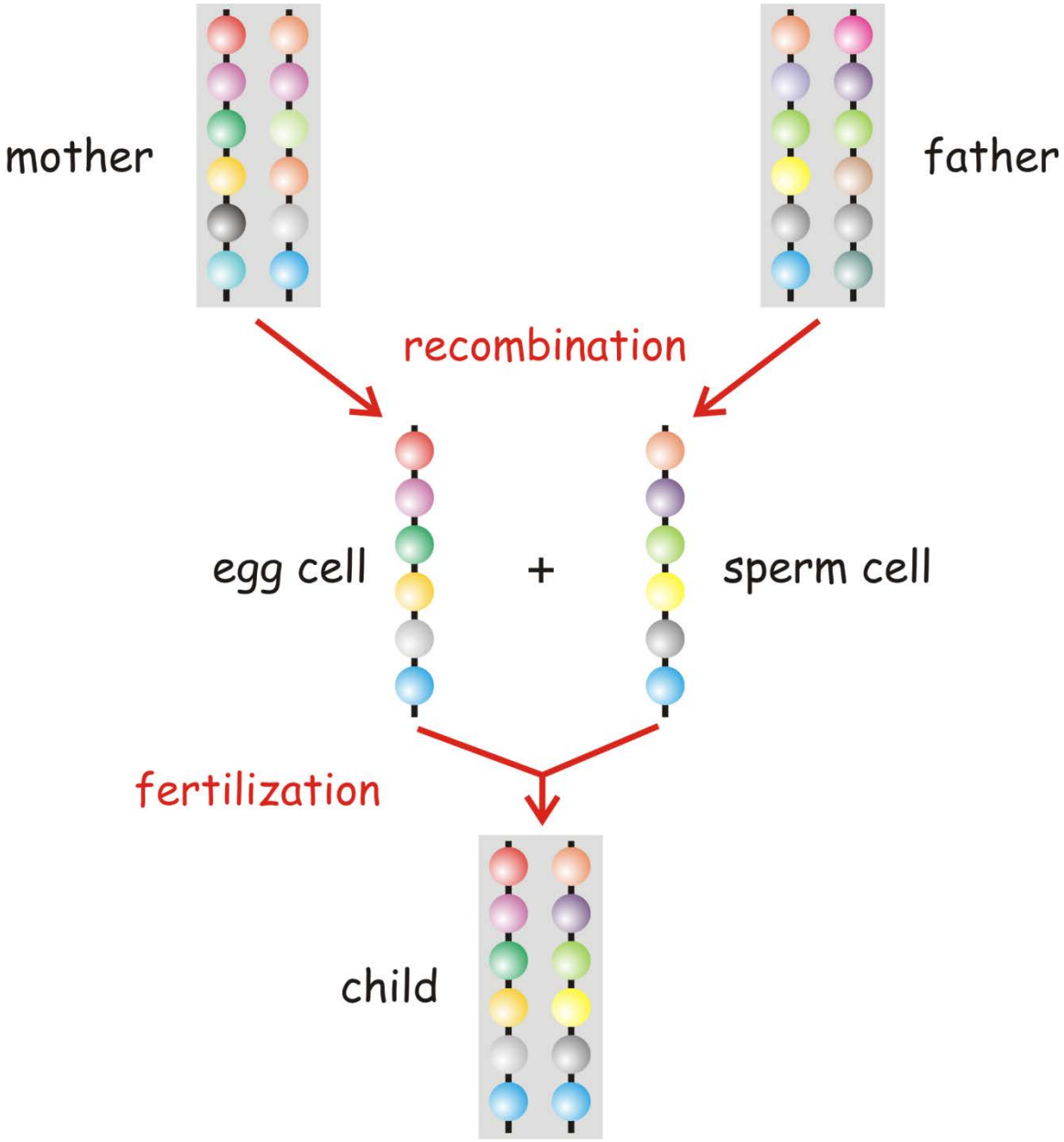
The core metabolism of the cell



Replication fork in DNA double-strand to double-strand replication.

The replication complex involves some twenty enzymes.

Gregor Mendel
1822 - 1884



Recombination in Mendelian genetics

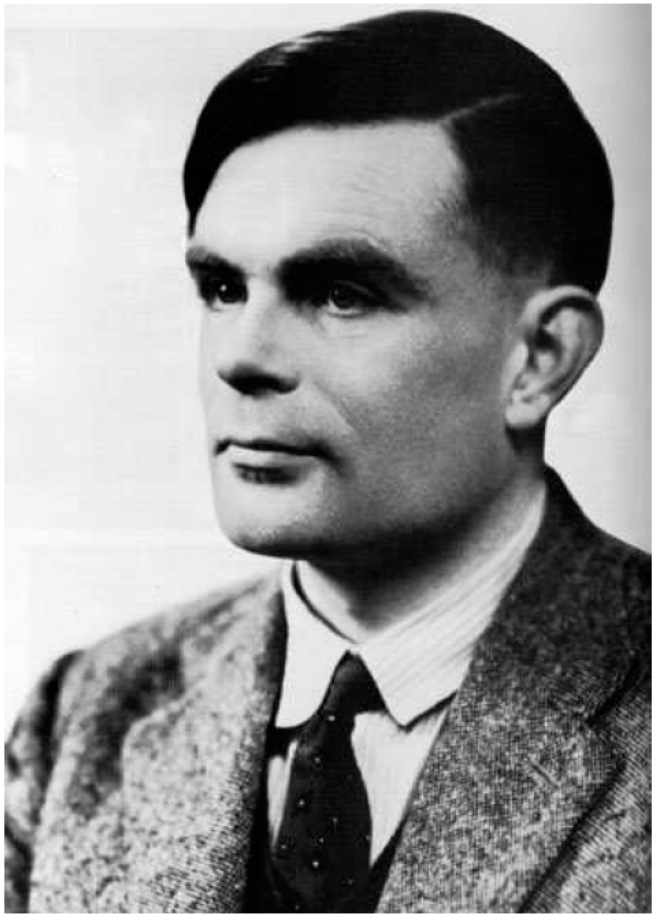
$$\frac{\partial u}{\partial t} = D_u \nabla^2 u + f(u,v)$$

$$\frac{\partial v}{\partial t} = D_v \nabla^2 v + g(u,v)$$

$$u = u(x,y,z,t) \quad \text{and} \quad v = v(x,y,z,t)$$

Change in local concentration =

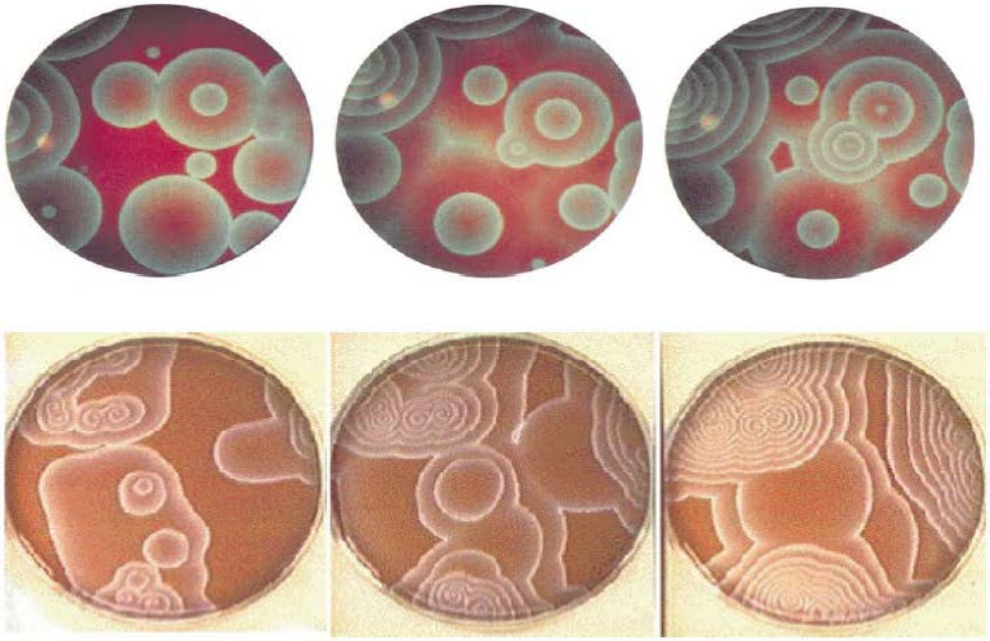
= diffusion + chemical reaction



Alan M. Turing, 1912-1954

A.M. Turing. 1952. The chemical basis of morphogenesis. *Phil.Trans.Roy.Soc.*London B **237**:37-72.

Liesegang Ringe 1895



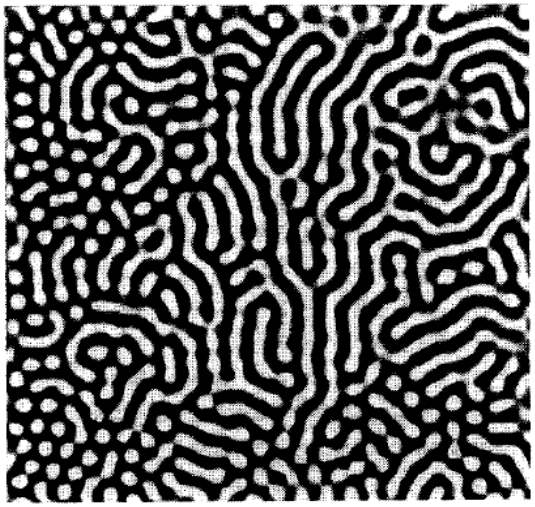
Belousov-Zhabotinskii Reaktion 1959

Musterbildung durch chemische Selbstorganisation:

Liesegang Ringe durch Fällung aus übersättigten Lösungen,

Raum-Zeit-Muster in der Belousov-Zhabotinskii Reaktion,

und stationäre Turing Muster.



Turing Muster:
Boissonade, De Kepper 1990

Philip K. Maini, 1959 -

More recently, detailed experimental work on *Drosophila* has shown that the pattern forming process is not, in fact, *via* reaction diffusion, but due to a cascade of gene switching, where certain gene proteins are expressed and, in turn, influence subsequent gene expression patterns. **Therefore, although reaction diffusion theory provides a very elegant mechanism for segmentation nature has chosen a much less elegant way of doing it!**

Philip K. Maini, Kevin J. Painter, and Helene Nguyen Phong Chau. 1997.
Spatial Pattern Formation in Chemical and Biological Systems
J.Chem.Soc., Faraday Transactions **93**:3601-3610.

Evelyn Fox Keller, 1936 -

„Untimely Birth of a Mathematical Biology"

Evelyn Fox Keller. 2002. Making Sense of Life.
Explaining Biological Development with Models,
Metaphors and Machines.
Harvard University Press. Cambridge, MA.

1. Prologue
2. **Darwin and replicating molecules**
3. In vitro evolution
4. „Simple“ landscapes
5. „Realistic“ landscapes
6. Neutrality in evolution
7. Perspectives of systems biology

Three necessary conditions for Darwinian evolution are:

1. **Multiplication,**
1. **Variation**, and
1. **Selection**.

Biologists distinguish the **genotype** – the genetic information – and the **phenotype** – the organisms and all its properties. The **genotype** is unfolded in development and yields the **phenotype**.

Variation operates on the **genotype** – through mutation and recombination – whereas the **phenotype** is the target of **selection**.

The Darwinian mechanism requires **no process** that could not be implemented in **cell-free molecular systems**.

DOI: 10.1002/anie.200701369

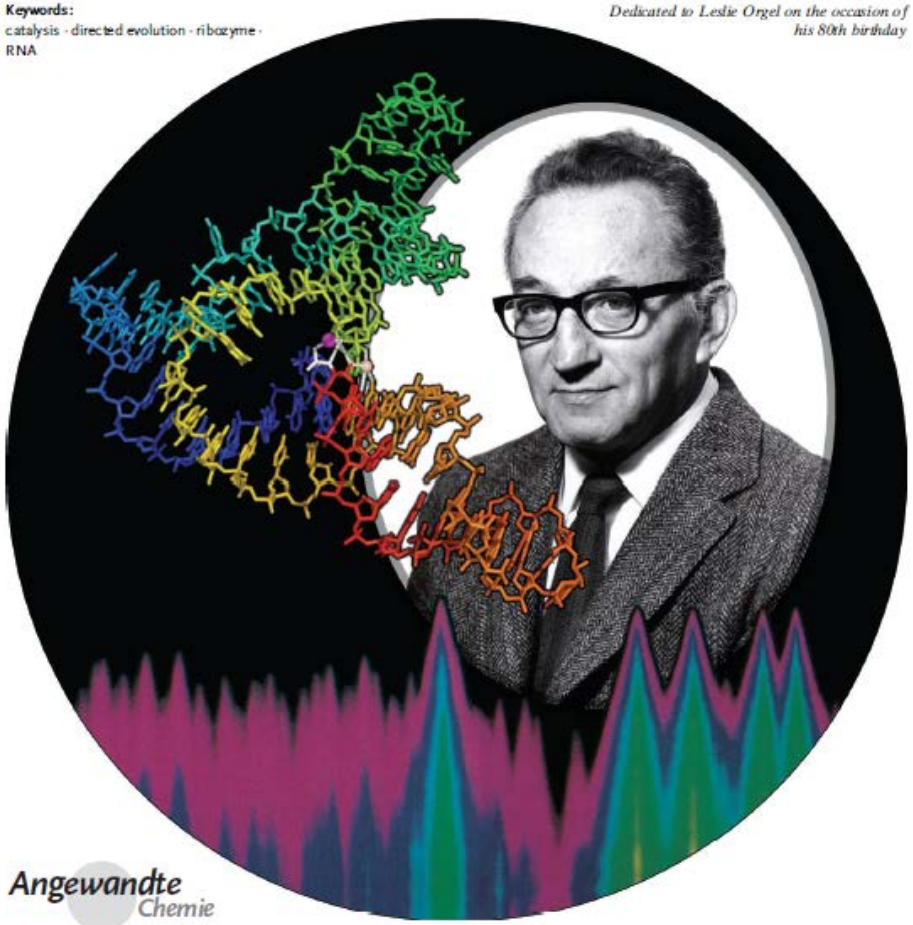
Molecular Evolution

Forty Years of In Vitro Evolution**

*Gerald F. Joyce**

Keywords:
catalysis · directed evolution · ribozyme · RNA

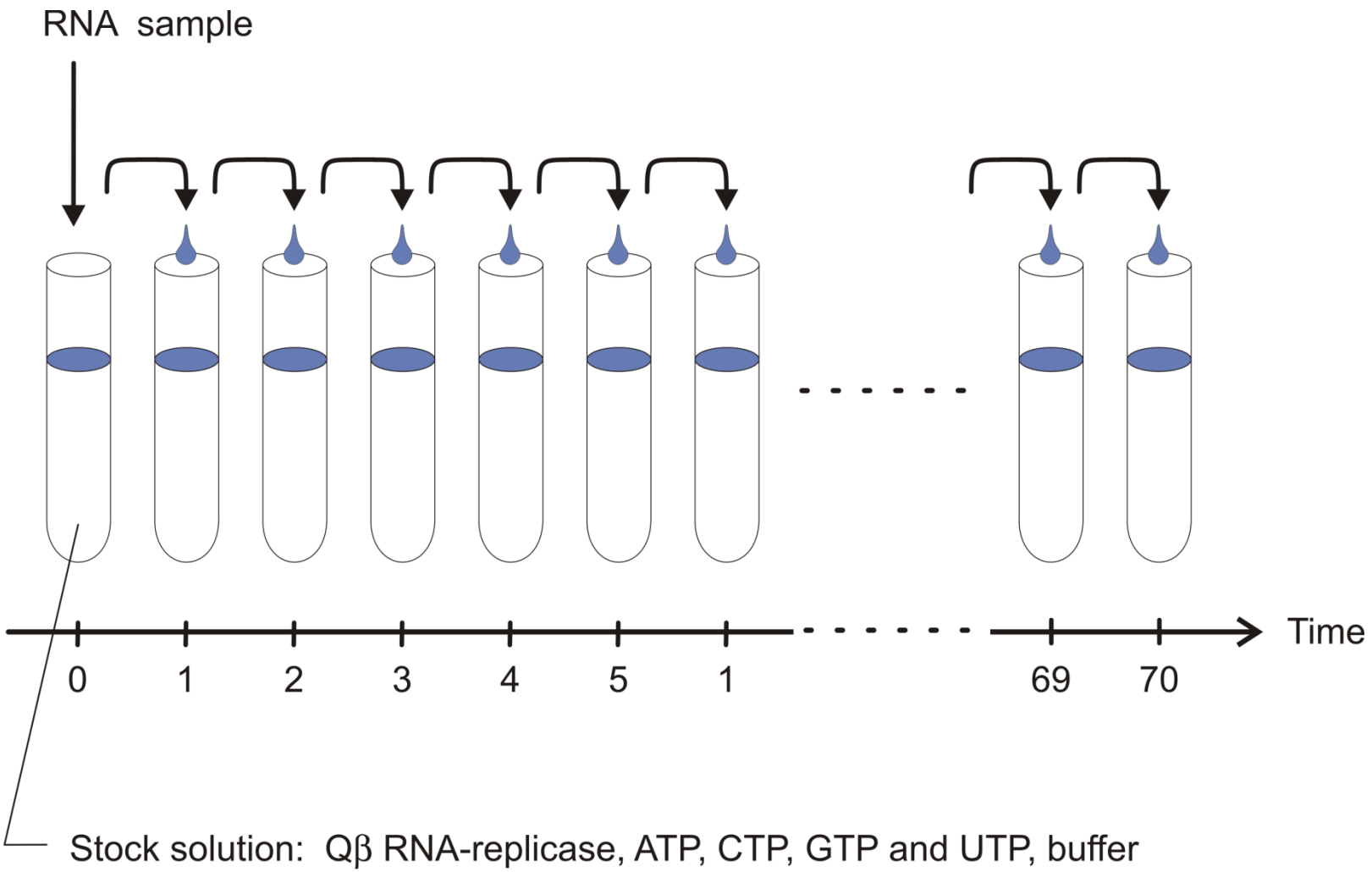
Dedicated to Leslie Orgel on the occasion of his 80th birthday



Sol Spiegelman,
1914 - 1983

Evolution in the test tube:

G.F. Joyce, *Angew.Chem.Int.Ed.* **46** (2007), 6420-6436


RNA sample
Time
0
1
2
3
4
5
1
69
70
Stock solution: Qβ RNA-replicase, ATP, CTP, GTP and UTP, buffer

Reproduction of the original figure of the serial transfer experiment with Qβ RNA

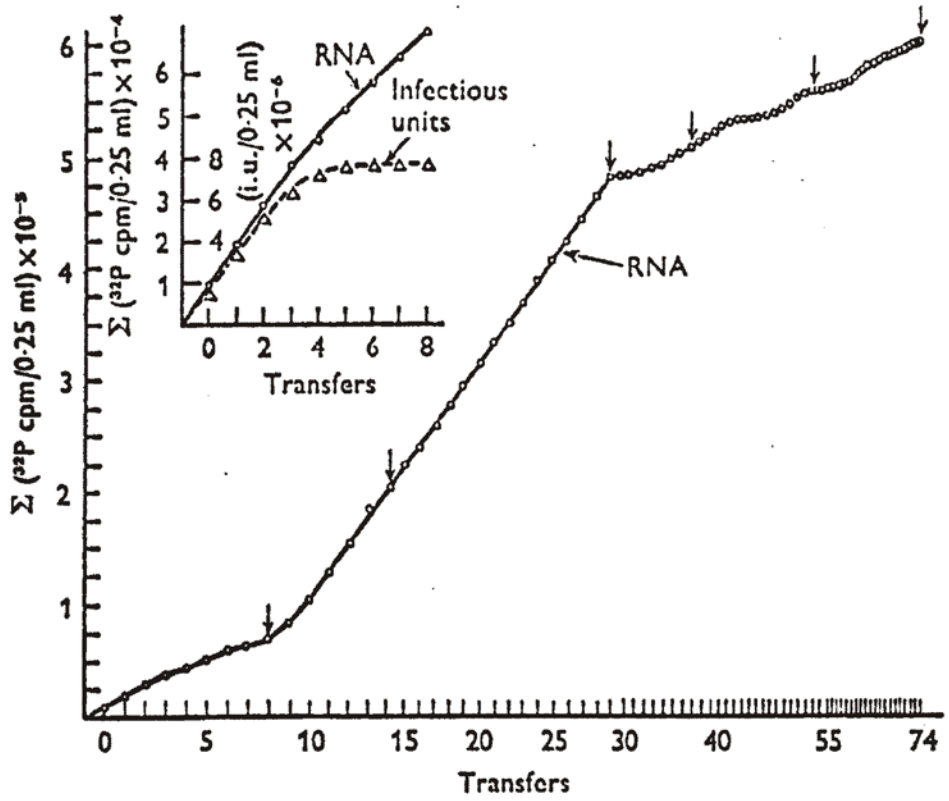


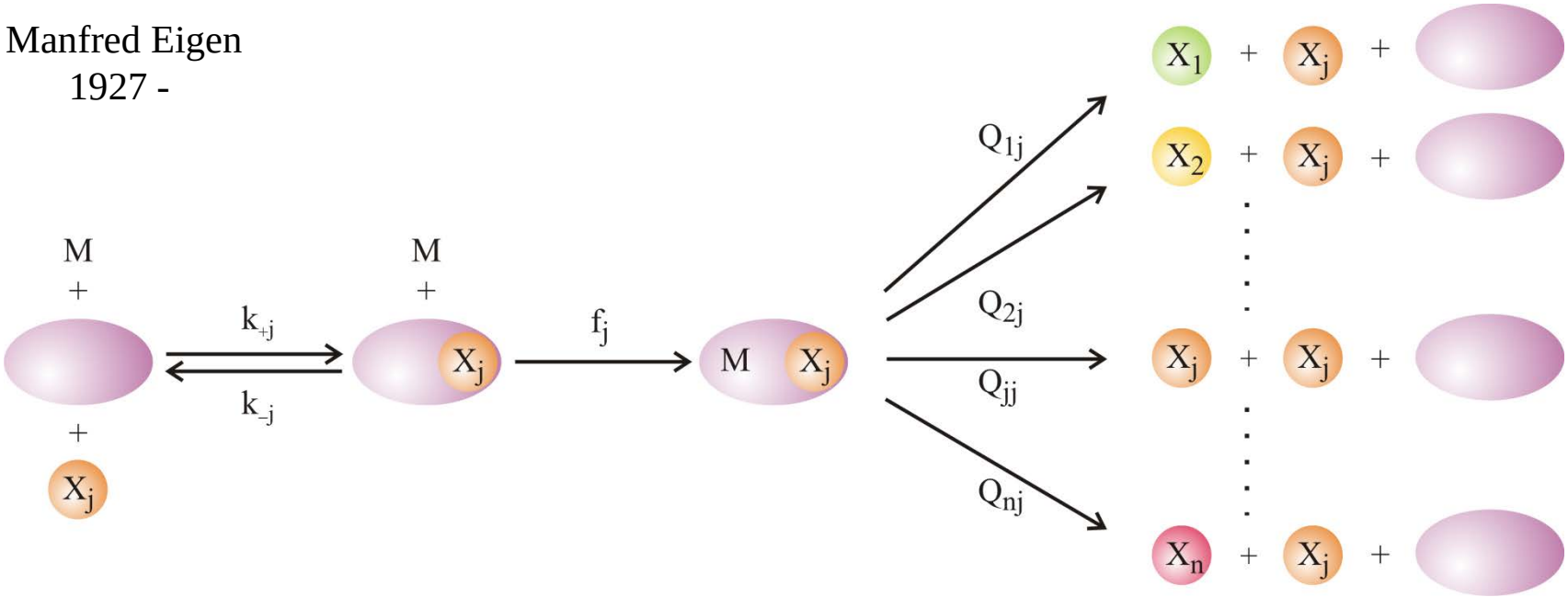
Fig. 9. Serial transfer experiment. Each 0·25 ml standard reaction mixture contained 40 μg of Qβ replicase and ^{32}P-UTP. The first reaction (0 transfer) was initiated by the addition of 0·2 μg ts-1 (temperature-sensitive RNA) and incubated at 35 °C for 20 min, whereupon 0·02 ml was drawn for counting and 0·02 ml was used to prime the second reaction (first transfer), and so on. After the first 13 reactions, the incubation periods were reduced to 15 min (transfers 14–29). Transfers 30–38 were incubated for 10 min. Transfers 39–52 were incubated for 7 min, and transfers 53–74 were incubated for 5 min. The arrows above certain transfers (0, 8, 14, 29, 37, 53, and 73) indicate where 0·001–0·1 ml of product was removed and used to prime reactions for sedimentation analysis on sucrose. The inset examines both infectious and total RNA. The results show that biologically competent RNA ceases to appear after the 4th transfer (Mills *et al.* 1967).

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

Manfred Eigen
1927 -

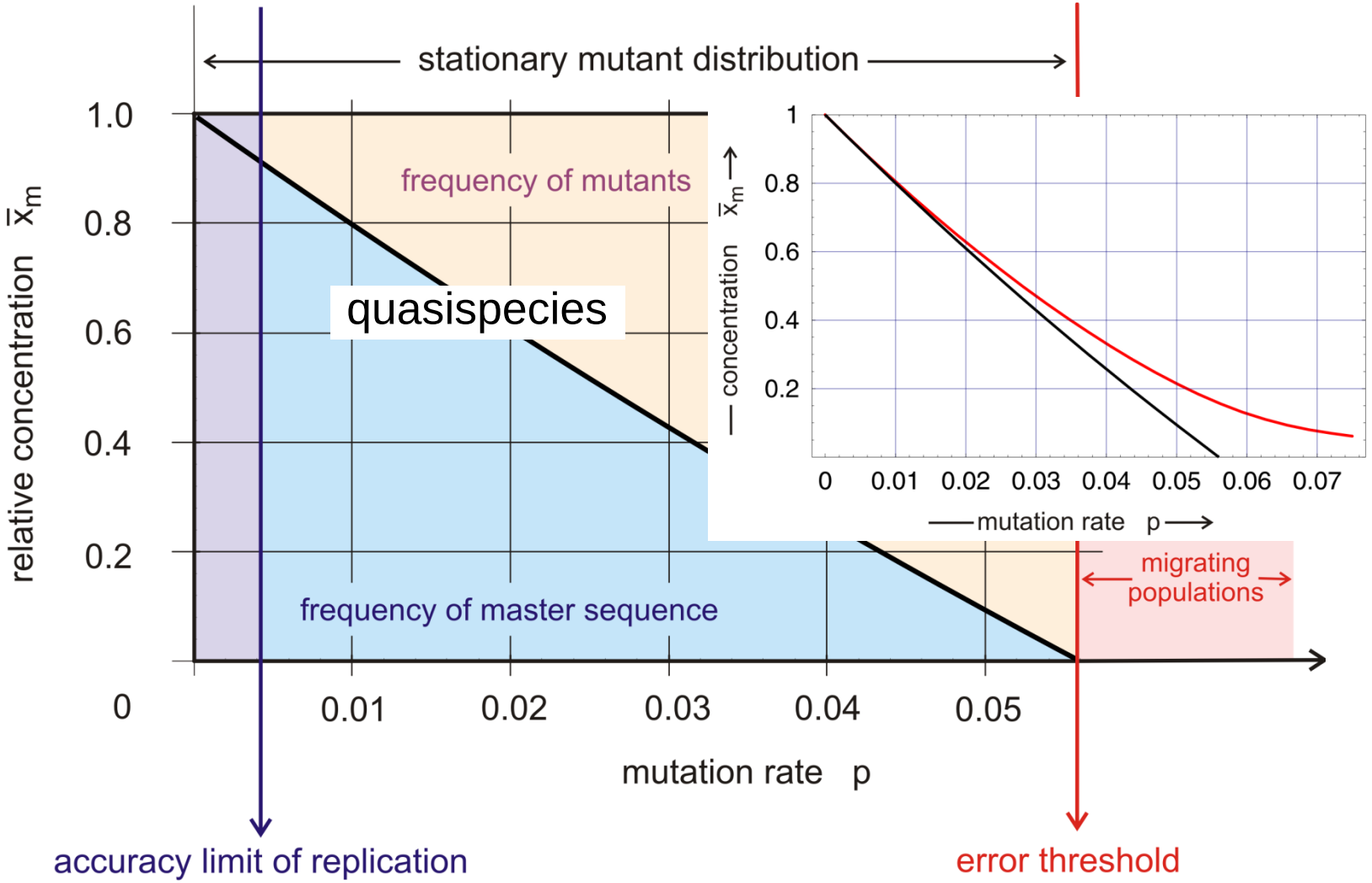
$$\frac{dx_j}{dt} = \sum_{i=1}^{n} W_{ji}\, x_i - x_j\, \Phi\,; \quad j = 1, 2, \ldots, n$$

$$\Phi = \sum_{i=1}^{n} f_i\, x_i \Big/ \sum_{i=1}^{n} x_i$$



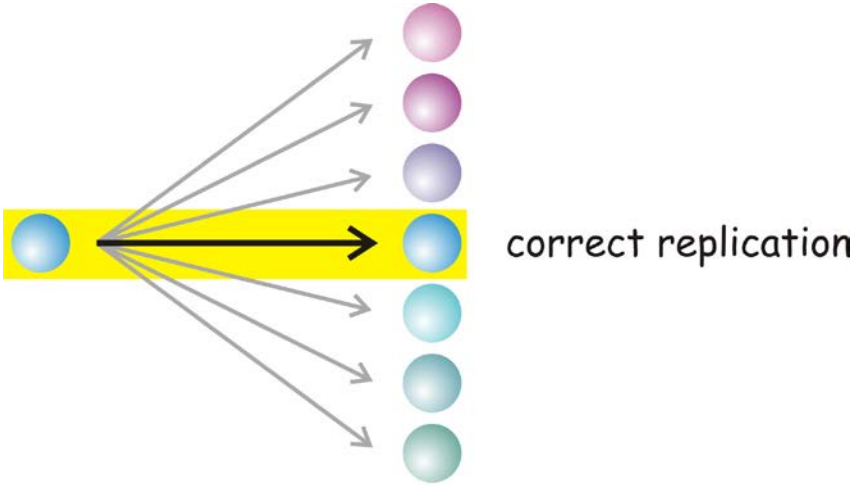
Mutation and (correct) replication as parallel chemical reactions

M. Eigen. 1971. *Naturwissenschaften* 58:465,
M. Eigen & P. Schuster.1977. *Naturwissenschaften* 64:541, 65:7 und 65:341



The error threshold in replication and mutation

Hermann J. Muller
1890 - 1967



Thomas H. Morgan
1866 - 1945

organism	mutation rate per genome	reproduction event
RNA virus	1	replication
retroviruses	0.1	replication
bacteria	0.003	replication
eukaryotes	0.003	cell division
eukaryotes	0.01 – 0.1	sexual reproduction

John W. Drake, Brian Charlesworth, Deborah Charlesworth and James F. Crow. 1998. Rates of spontaneous mutation. *Genetics* 148:1667-1686.

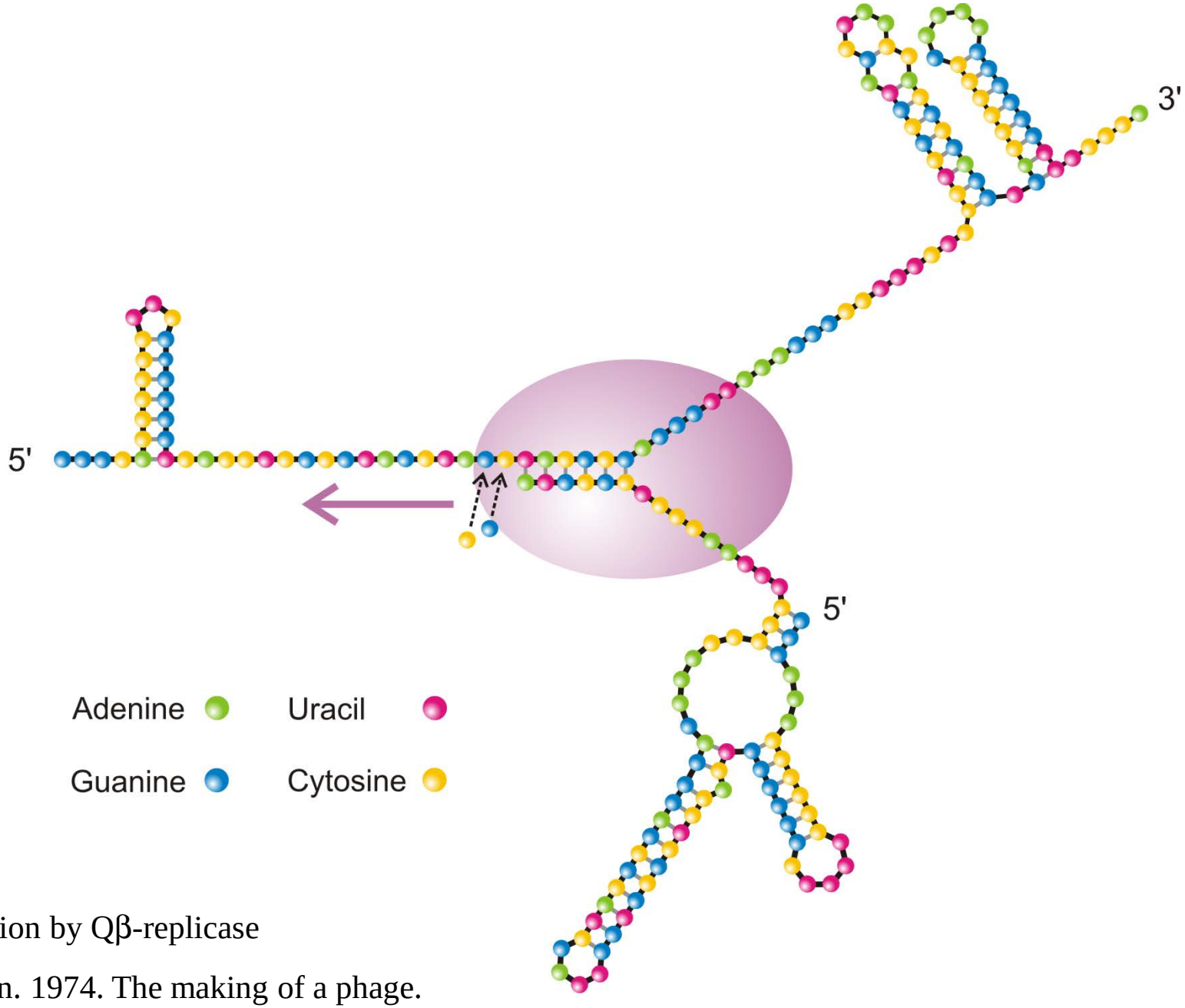
Results of the kinetic theory of evolution

1. Not a single "wild type" is selected but a fittest genotype together with its mutant cloud forming a **quasispecies**.

2. Mutation rates are limited by an **error threshold** above which genetic information is unstable.

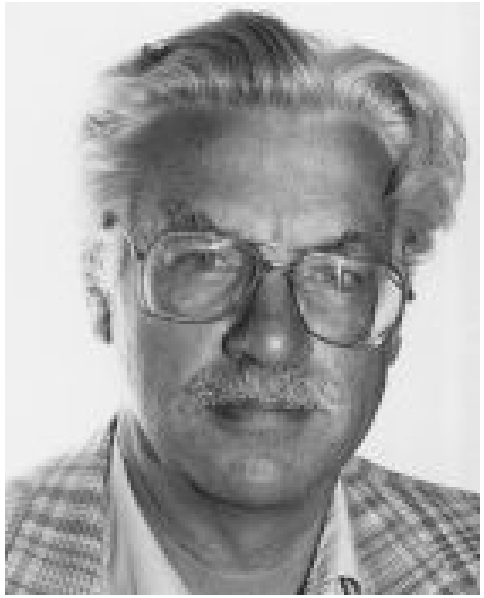
3. For a given replication machinery the error threshold sets a limit to the length of genomes.

1. Prologue
2. Darwin and replicating molecules
3. **In vitro evolution**
4. „Simple" landscapes
5. „Realistic" landscapes
6. Neutrality in evolution
7. Perspectives of systems biology

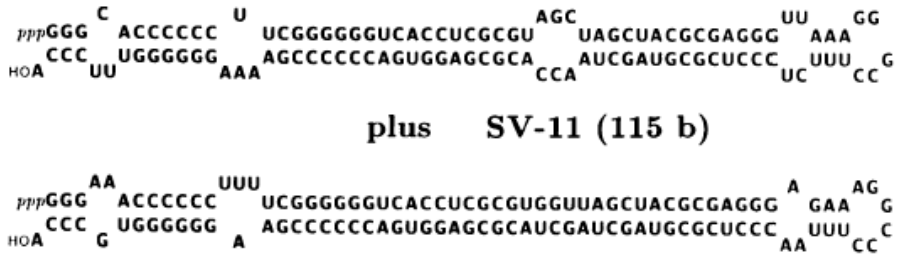


RNA replication by Qβ-replicase

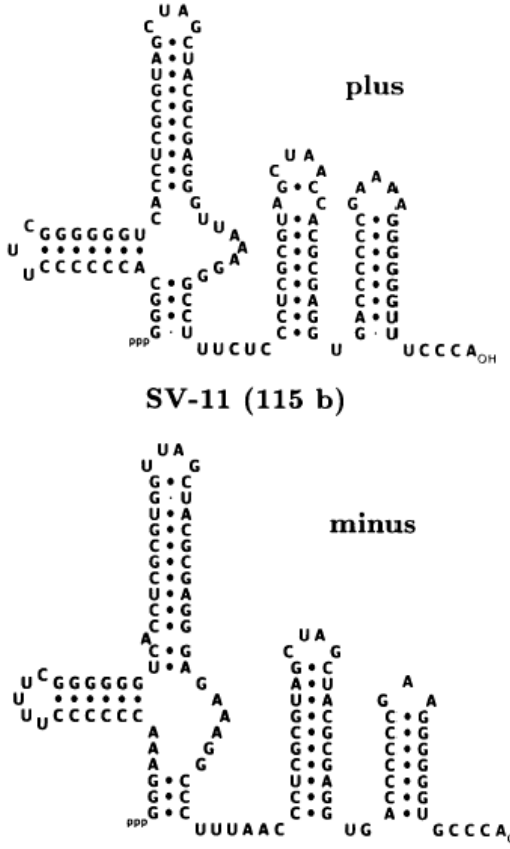
C. Weissmann. 1974. The making of a phage.
FEBS Letters **40**:S10-S18



Christof K. Biebricher 1941-2009

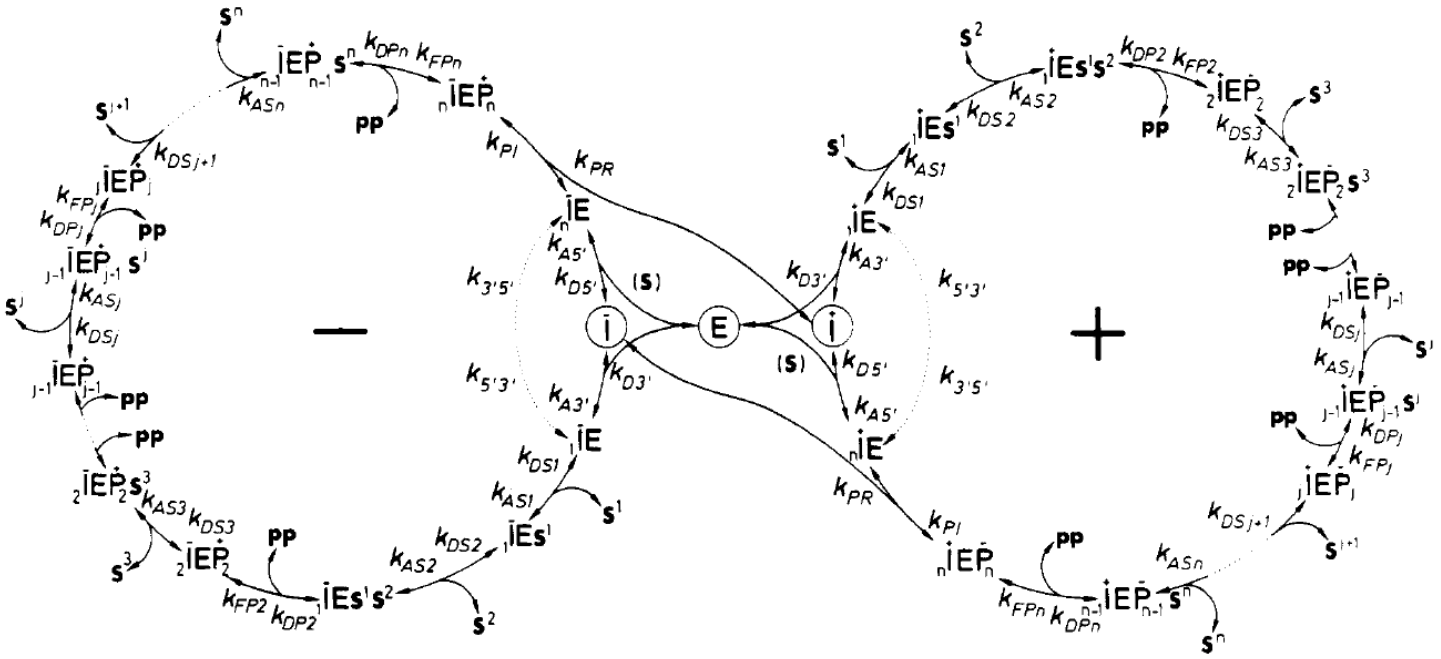


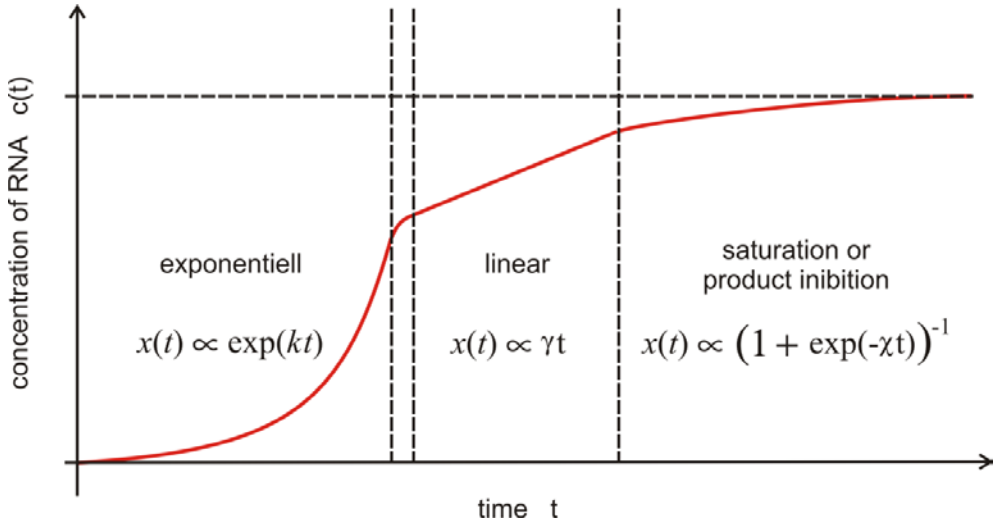
stable



metastable

C.K. Biebricher, R. Luce. 1992. *In vitro* recombination and terminal recombination of RNA by Qβ replicase. *The EMBO Journal* 11:5129-5135.





Kinetics of RNA replication

C.K. Biebricher, M. Eigen, W.C. Gardiner, Jr.
Biochemistry **22**:2544-2559, 1983

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Virus Research 107 (2005) 115–116
Virus Research
www.elsevier.com/locate/virusres

Preface

Antiviral strategy on the horizon

Error catastrophe had its conceptual origins in the middle of the XXth century, when the consequences of mutations on enzymes involved in protein synthesis, as a theory of aging. In those times biological processes were generally perceived differently from today. Infectious diseases were regarded as a fleeting nuisance which would be eliminated through the use of antibiotics and antiviral agents. Microbial variation, although known in some cases, was not thought to be a significant problem for disease control. Variation in differentiated organisms was seen as resulting essentially from exchanges of genetic material associated with sexual reproduction. The problem was to unveil the mechanisms of inheritance, expression of genetic information and metabolism. Few saw that genetic change is occurring at present in all organisms, and still fewer recognized Darwinian principles as essential to the biology of pathogenic viruses and cells. Population geneticists rarely used bacteria or viruses as experimental systems to define concepts in biological evolution. The extent of genetic polymorphism among individuals of the same biological species came as a surprise when the first results on comparison of electrophoretic mobility of enzymes were obtained. With the advent of in vitro DNA recombination, and rapid nucleic acid sequencing techniques, molecular analyses of genomes reinforced the conclusion of extreme inter-individual genetic variation within the same species. Now, due largely to spectacular progress in comparative genomics, we see cellular DNAs, both prokaryotic and eukaryotic, as highly dynamic. Most cellular processes, including such essential information-bearing and transferring events as genome replication, transcription and translation, are increasingly perceived as inherently inaccurate. Viruses, and in particular RNA viruses, are among the most extreme examples of exploitation of replication inaccuracy for survival.

Error catastrophe, or the loss of meaningful genetic information through excess genetic variation, was formulated in quantitative terms as a consequence of quasispecies theory, which was first developed to explain self-organization and adaptability of primitive replicons in early stages of life. Recently, a conceptual extension of error catastrophe that could be defined as "induced genetic deterioration" has emerged as a possible antiviral strategy. This is the topic of the current special issue of ***Virus Research***.

Few would nowadays doubt that one of the major obstacles for the control of viral disease is short-term adaptability of viral pathogens. Adaptability of viruses follows the same Darwinian principles that have shaped biological evolution over eons, that is, repeated rounds of reproduction with genetic variation, competition and selection, often perturbed by random events such as statistical fluctuations in population size. However, with viruses the consequences of the operation of these very same Darwinian principles are felt within very short times. Short-term evolution (within hours and days) can be also observed with some cellular pathogens, with subsets of normal cells, and cancer cells. The nature of RNA viral pathogens begs for alternative antiviral strategies, and forcing the virus to cross the critical error threshold for maintenance of genetic information is one of them.

The contributions to this volume have been chosen to reflect different lines of evidence (both theoretical and experimental) on which antiviral designs based on genetic deterioration inflicted upon viruses are being constructed. Theoretical studies have explored the copying fidelity conditions that must be fulfilled by any information-bearing replication system for the essential genetic information to be transmitted to progeny. Closely related to the theoretical developments have been numerous experimental studies on quasispecies dynamics and their multiple biological manifestations. The latter can be summarized by saying that RNA viruses, by virtue of existing as mutant spectra rather than defined genetic entities, remarkably expand their potential to overcome selective pressures intended to limit their replication. Indeed, the use of antiviral inhibitors in clinical practice and the design of vaccines for a number of major RNA virus-associated diseases, are currently presided by a sense of uncertainty. Another line of growing research is the enzymology of copying fidelity by viral replicases, aimed at understanding the molecular basis of mutagenic activities. Error catastrophe as a potential new antiviral strategy received an important impulse by the observation that ribavirin (a licensed antiviral nucleoside analogue) may be exerting, in some systems, its antiviral activity through enhanced mutagenesis. This has encouraged investigations on new mutagenic base analogues, some of them used in anticancer chemotherapy. Some chapters summarize these important biochemical studies on cell entry pathways and metabolism of mutagenic agents, that may find new applications as antiviral agents.

This volume intends to be basically a progress report, an introduction to a new avenue of research, and a realistic appraisal of the many issues that remain to be investigated. In this respect, I can envisage (not without many uncertainties) at least three lines of needed research: (i) One on further understanding of quasispecies dynamics in infected individuals to learn more on how to apply combinations of virus-specific mutagens and inhibitors in an effective way, finding synergistic combinations and avoiding antagonistic ones as well as severe clinical side effects. (ii) Another on a deeper understanding of the metabolism of mutagenic agents, in particular base and nucleoside analogues. This includes identification of the transporters that carry them into cells, an understanding of their metabolic processing, intracellular stability and alterations of nucleotide pools, among other issues. (iii) Still another line of needed research is the development of new mutagenic agents specific for viruses, showing no (or limited) toxicity for cells. Some advances may come from links with anticancer research, but others should result from the designs of new molecules, based on the structures of viral polymerases. I really hope that the reader finds this issue not only to be an interesting and useful review of the current situation in the field, but also a stimulating exposure to the major problems to be faced.

The idea to prepare this special issue came as a kind invitation of Ulrich Desselberger, former Editor of ***Virus Research***, and then taken enthusiastically by Luis Enjuanes, recently appointed as Editor of ***Virus Research***. I take this opportunity to thank Ulrich, Luis and the Editor-in-Chief of ***Virus Research***, Brian Mahy, for their continued interest and support to the research on virus evolution over the years.

My thanks go also to the 19 authors who despite their busy schedules have taken time to prepare excellent manuscripts, to Elsevier staff for their prompt responses to my requests, and, last but not least, to Ms. Lucía Horrillo from Centro de Biología Molecular "Severo Ochoa" for her patient dealing with the correspondence with authors and the final organization of the issue.

Esteban Domingo
Universidad Autónoma de Madrid
Centro de Biología Molecular "Severo Ochoa"
Consejo Superior de Investigaciones Científicas
Cantoblanco and Valdeolmos
Madrid, Spain
Tel.: +34 91 497 84858/9; fax: +34 91 497 4799
E-mail address: edomingo@cbm.uam.es

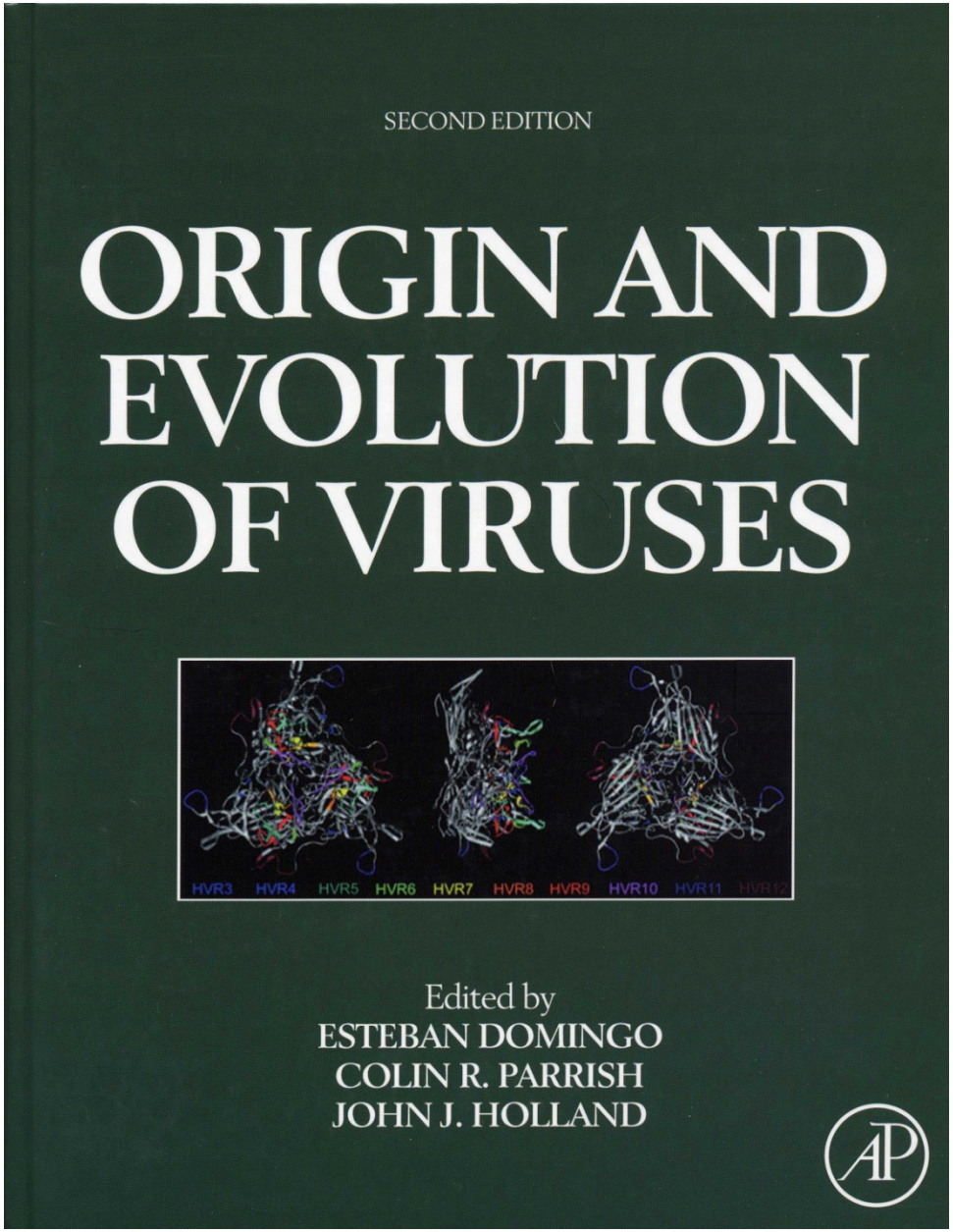
Available online 8 December 2004

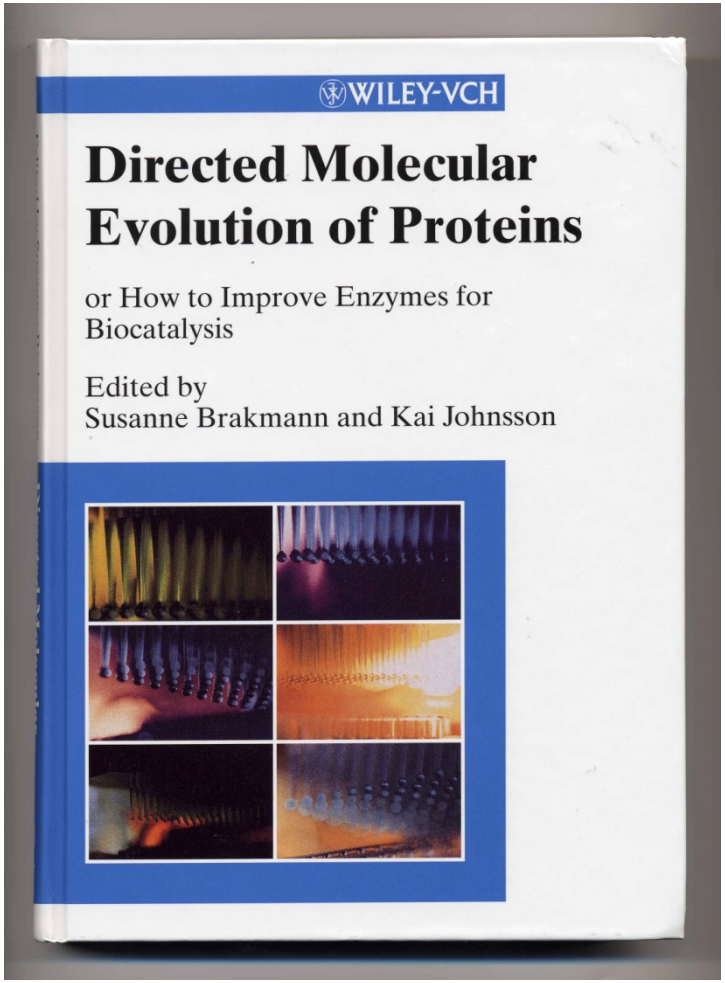
doi:10.1016/j.virusres.2004.11.001

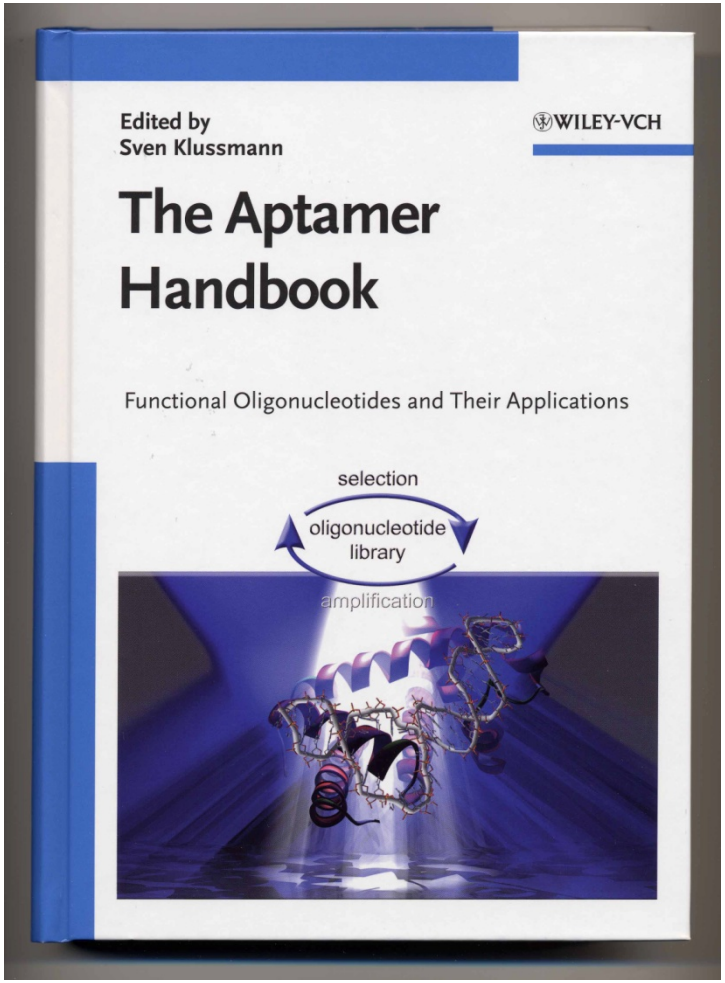
Esteban Domingo
1943 -

Application of quasispecies theory to the fight against viruses

Molecular evolution of viruses

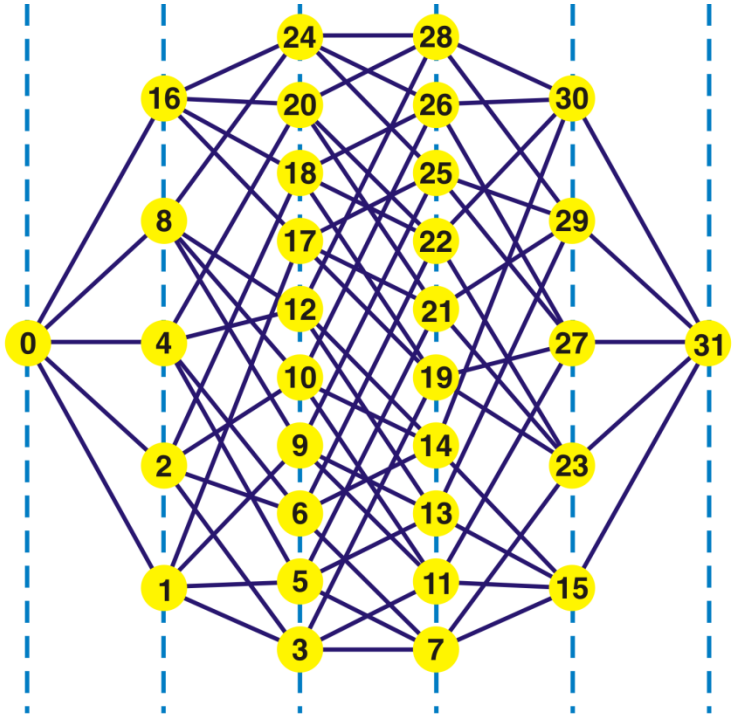






Application of molecular evolution to problems in biotechnology

1. Prologue
2. Darwin and replicating molecules
3. In vitro evolution
4. **„Simple" landscapes**
5. „Realistic" landscapes
6. Neutrality in evolution
7. Perspectives of systems biology



Binary sequences are encoded by their decimal equivalents:

C = 0 and G = 1, for example,

"0" ≡ 00000 = CCCCC,

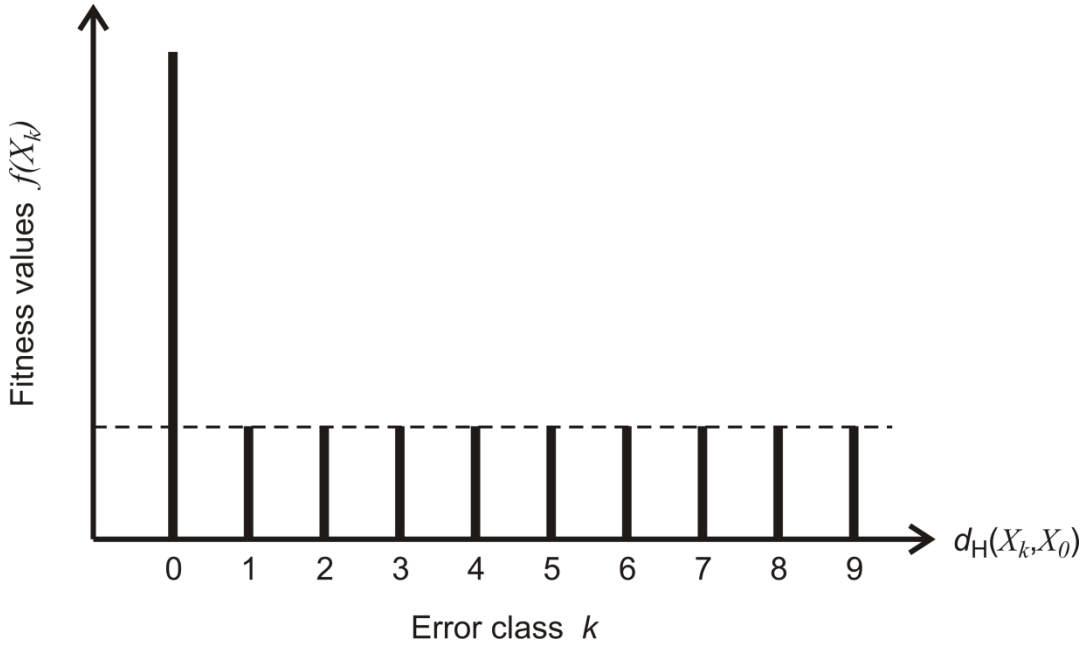
"14" ≡ 01110 = CGGGC,

"29" ≡ 11101 = GGGCG, etc.

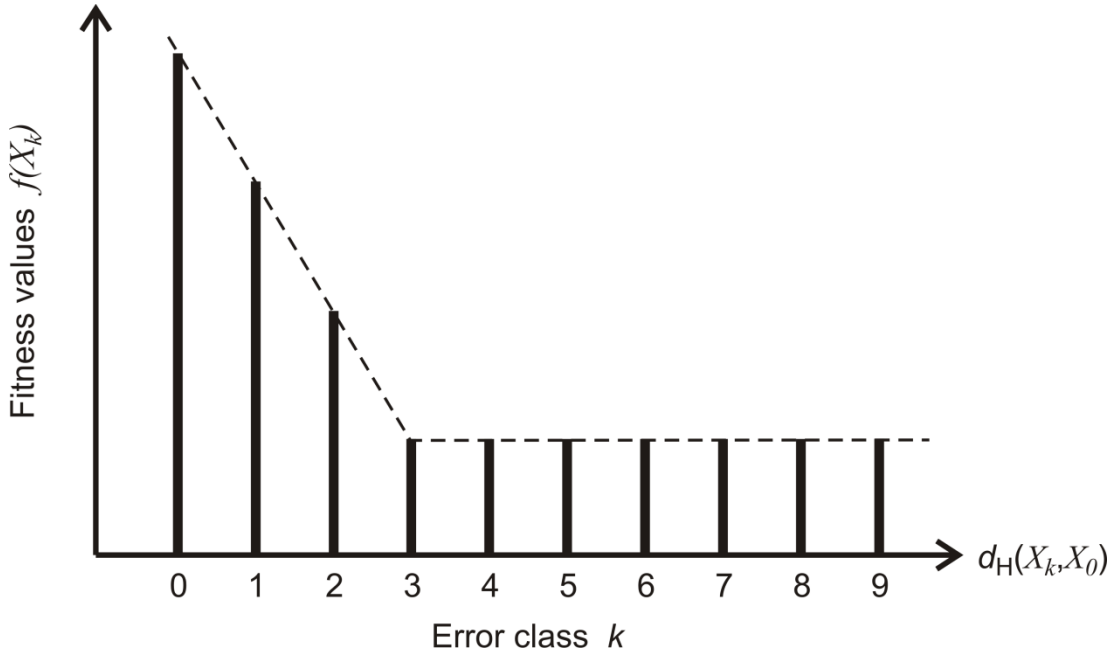
Concentrations of entire error classes: $[\Gamma_k] = y_k(p)\,,\ k = 0, 1, \ldots, n$

$$y_k(p) = \sum_{i=1,\ d_{\mathrm{H}}(\mathsf{X}_i,\mathsf{X}_k)=k}^{N} x_i(p)\,, \quad |\Gamma_k| = \binom{n}{k}$$

single peak landscape

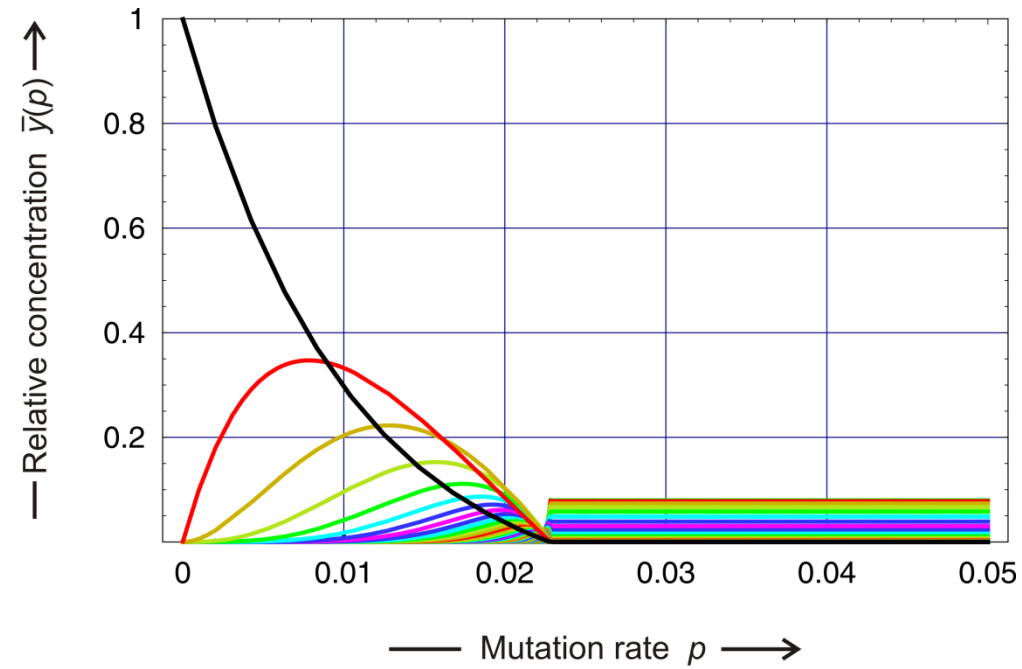


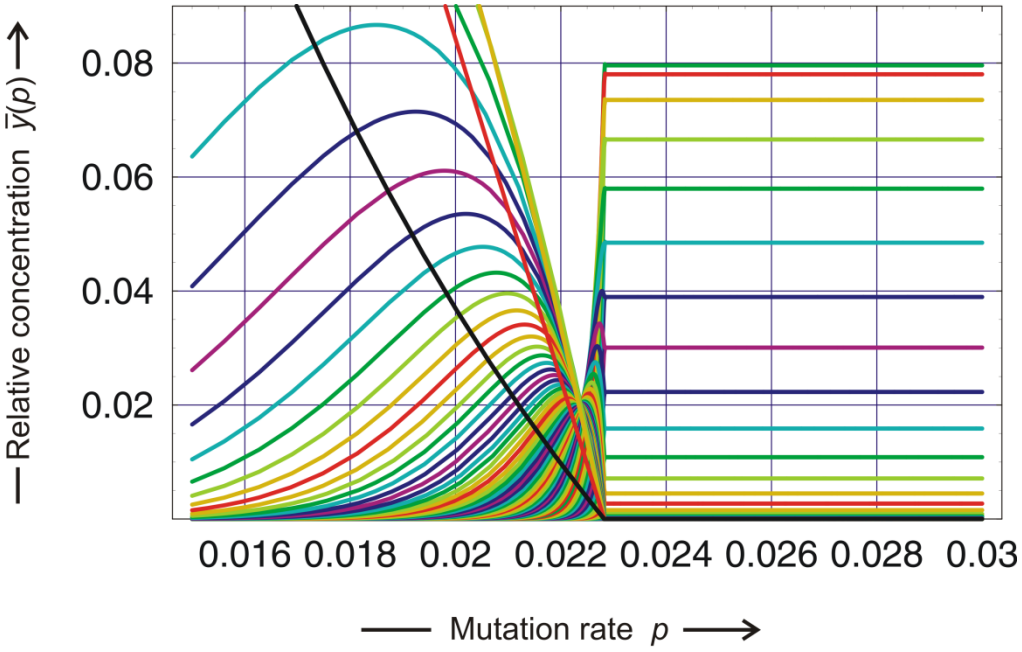
step linear landscape



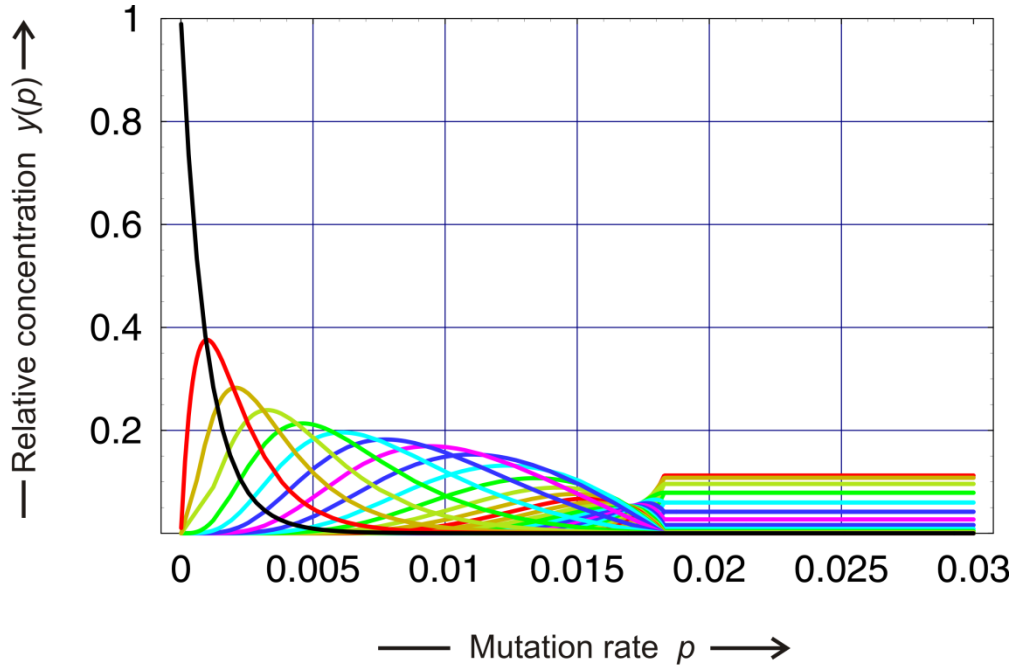
Model fitness landscapes I

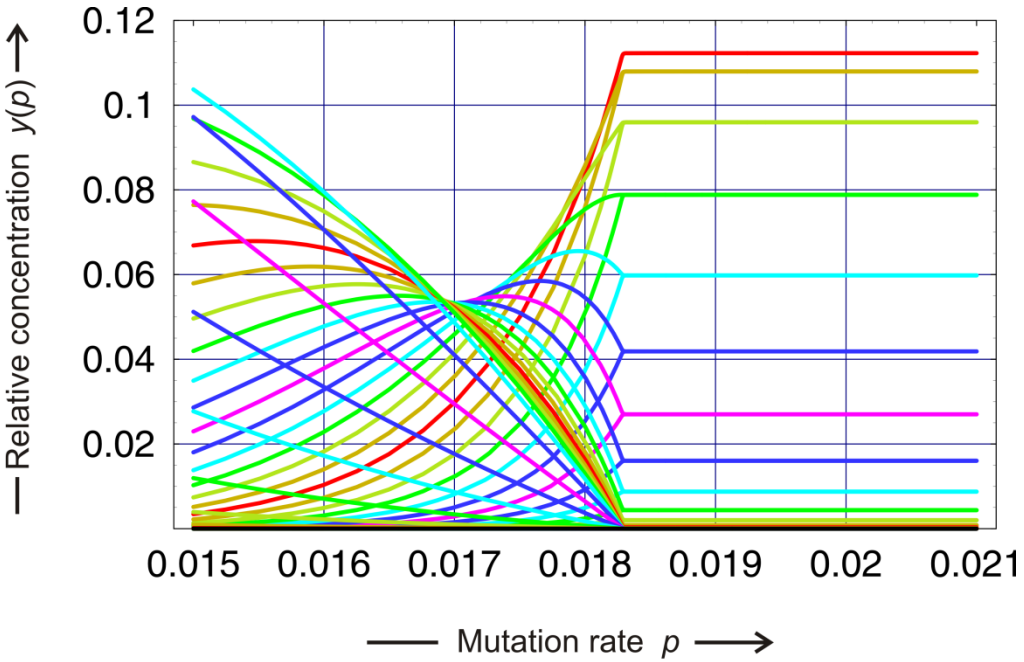
Error threshold on the single peak landscape





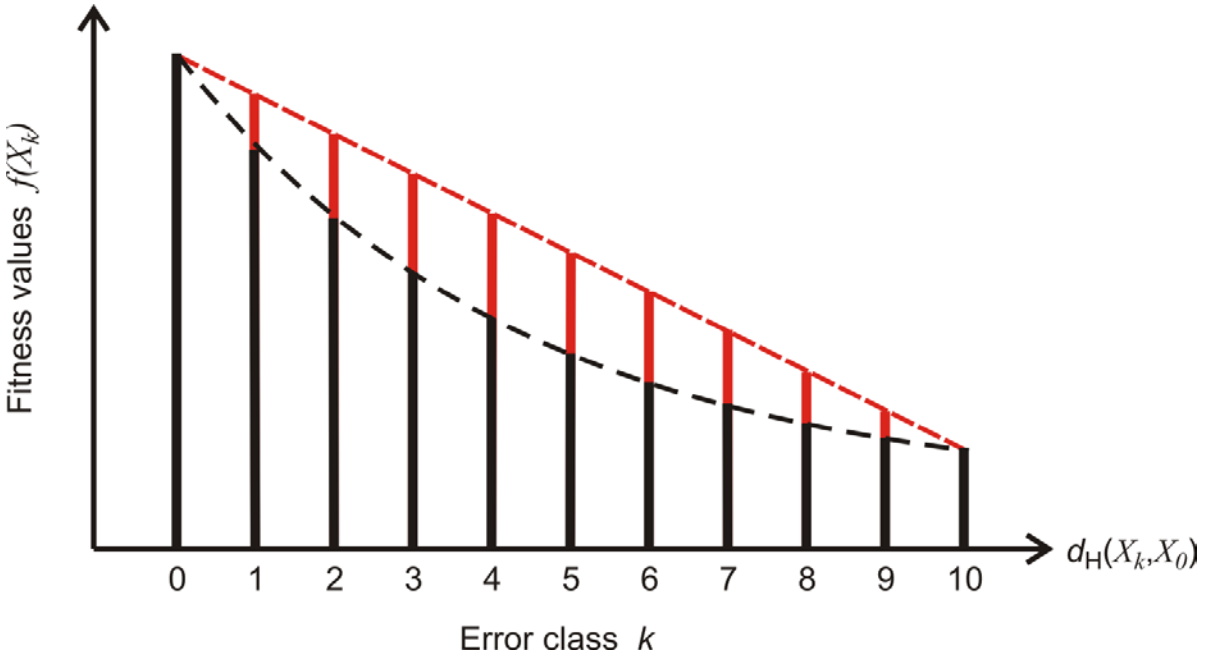
Error threshold on the step linear landscape



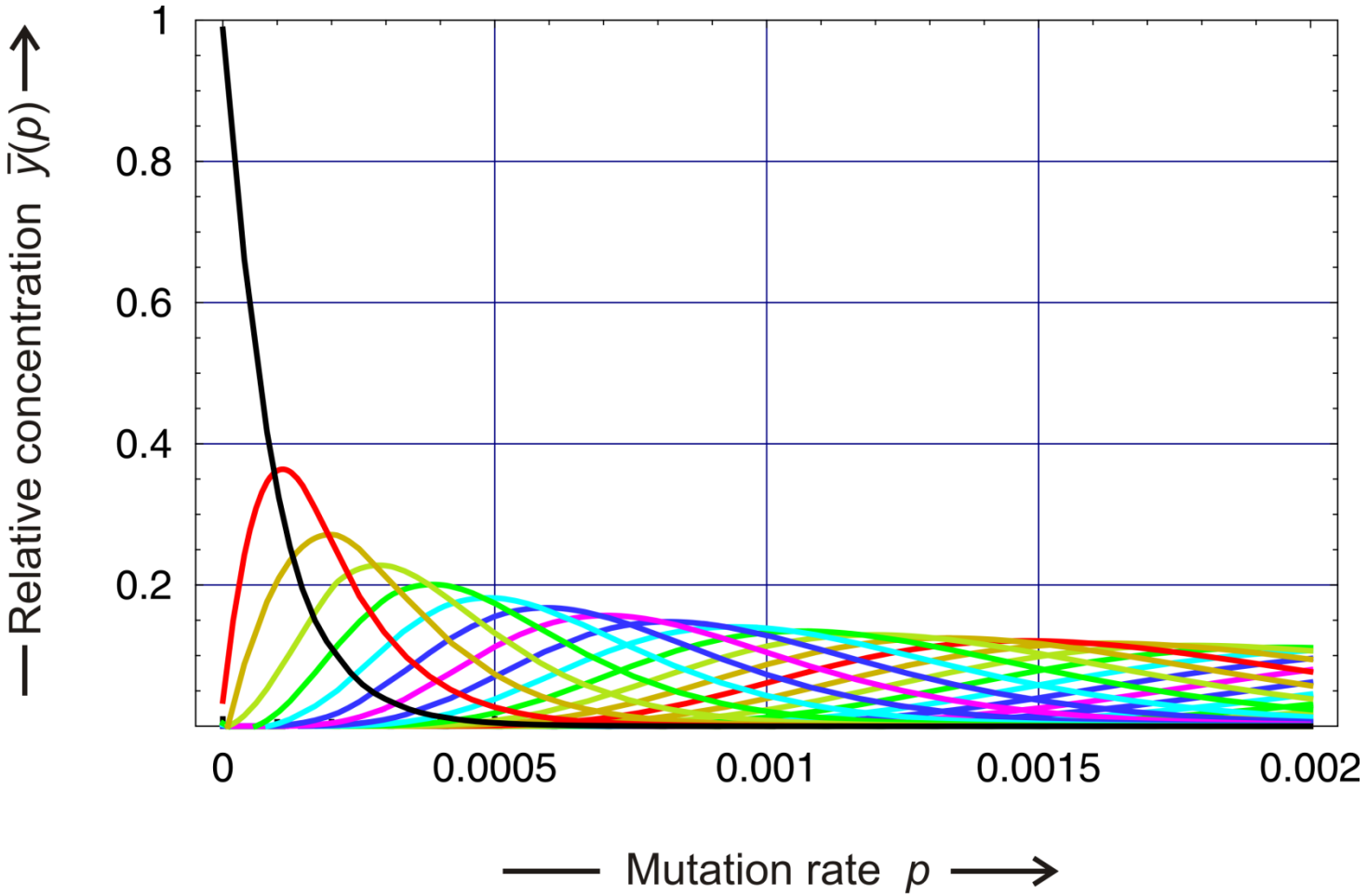


Thomas Wiehe. 1997. Model dependency of error thresholds: The role of fitness functions and contrasts between the finite and infinite sites models. *Genet. Res. Camb.* 69:127-136

linear and multiplicative

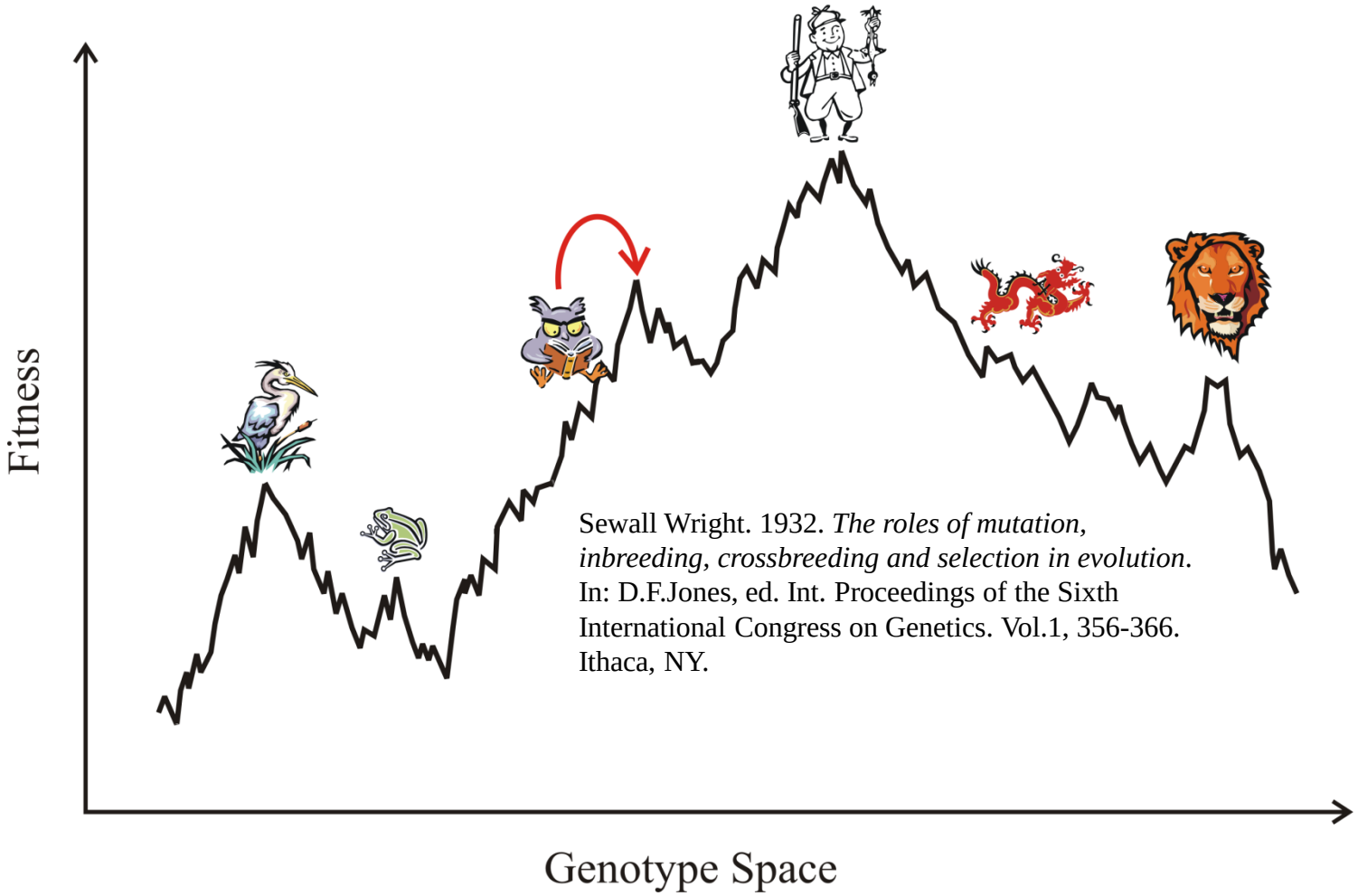


Model fitness landscapes II



The linear fitness landscape shows no error threshold

1. Prologue
2. Darwin and replicating molecules
3. In vitro evolution
4. „Simple" landscapes
5. **„Realistic" landscapes**
6. Neutrality in evolution
7. Perspectives of systems biology



Sewall Wrights fitness landscape as metaphor for Darwinian evolution

Sewall Wright, 1889 - 1988

+ …….. wild type
a ..... alternative allele on locus A
:
:
:
abcde … alternative alleles on all five loci

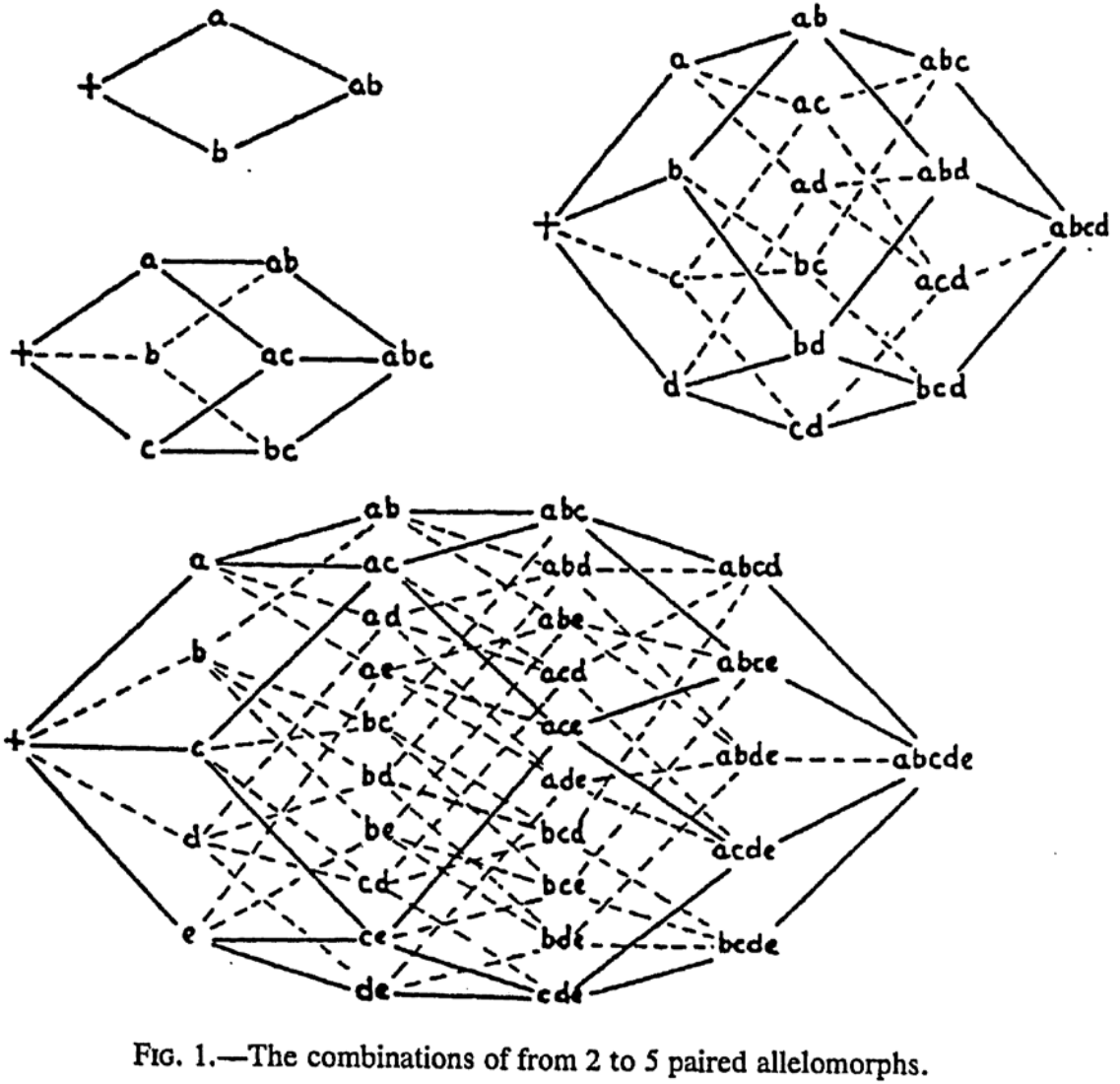


FIG. 1.—The combinations of from 2 to 5 paired allelomorphs.

The multiplicity of gene replacements with two alleles on each locus

Sewall Wright. 1988. Surfaces of selective value revisited. American Naturalist 131:115-123

Fitness landscapes became experimentally accessible!

Protein landscapes: Yuuki Hayashi, Takuyo Aita, Hitoshi Toyota, Yuzuru Husimi, Itaru Urabe, Tetsuya Yomo. 2006. Experimental rugged fitness landscape in protein seqeunce space. *PLoS One* 1:e96.

RNA landscapes: Sven Klussman, Ed. 2005. The aptamer handbook. Wiley-VCh, Weinheim (Bergstraße), DE.
Jason N. Pitt, Adrian Ferré-D'Amaré. 2010. Rapid construction of empirical RNA fitness landscapes. *Science* 330:376-379.

RNA viruses: Esteban Domingo, Colin R. Parrish, John J. Holland, Eds. 2007. Origin and evolution of viruses. Second edition. Elesvier, San Diego, CA.

Retroviruses: Roger D. Kouyos, Gabriel E. Leventhal, Trevor Hinkley, Mojgan Haddad, Jeannette M. Whitcomb, Christos J. Petropoulos, Sebastian Bonhoeffer. 2012. Exploring the complexity of the HIV-I fitness landscape. *PLoS Genetics* 8:e1002551

Realistic fitness landscapes

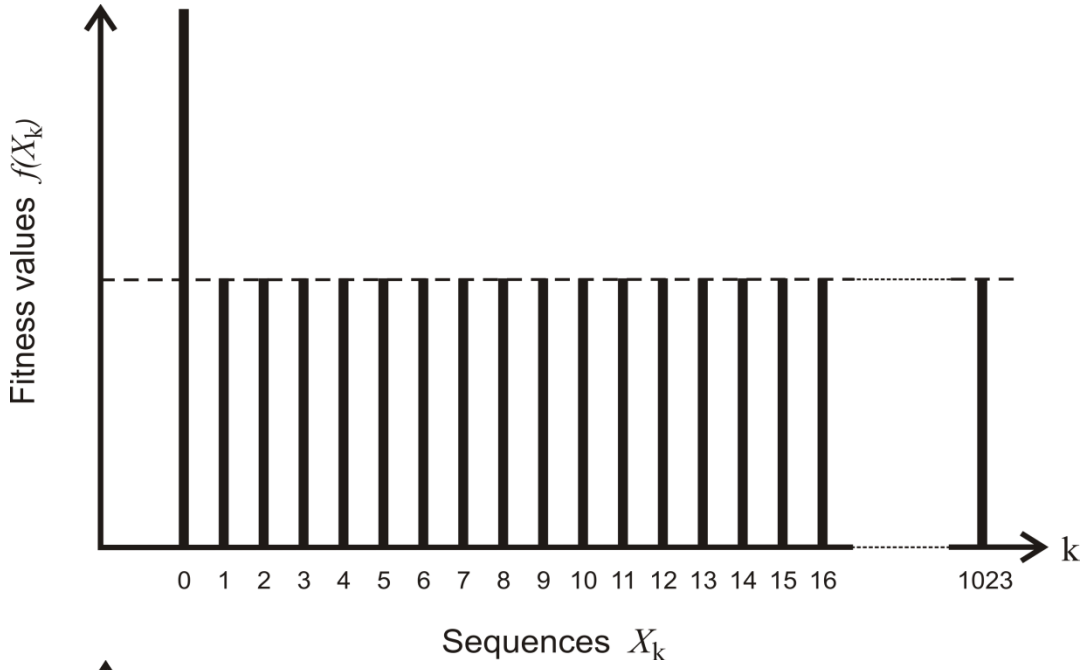
1. Ruggedness: nearby lying genotypes may develop into very different phenotypes

2. Neutrality: many different genotypes give rise to phenotypes with identical selection behavior

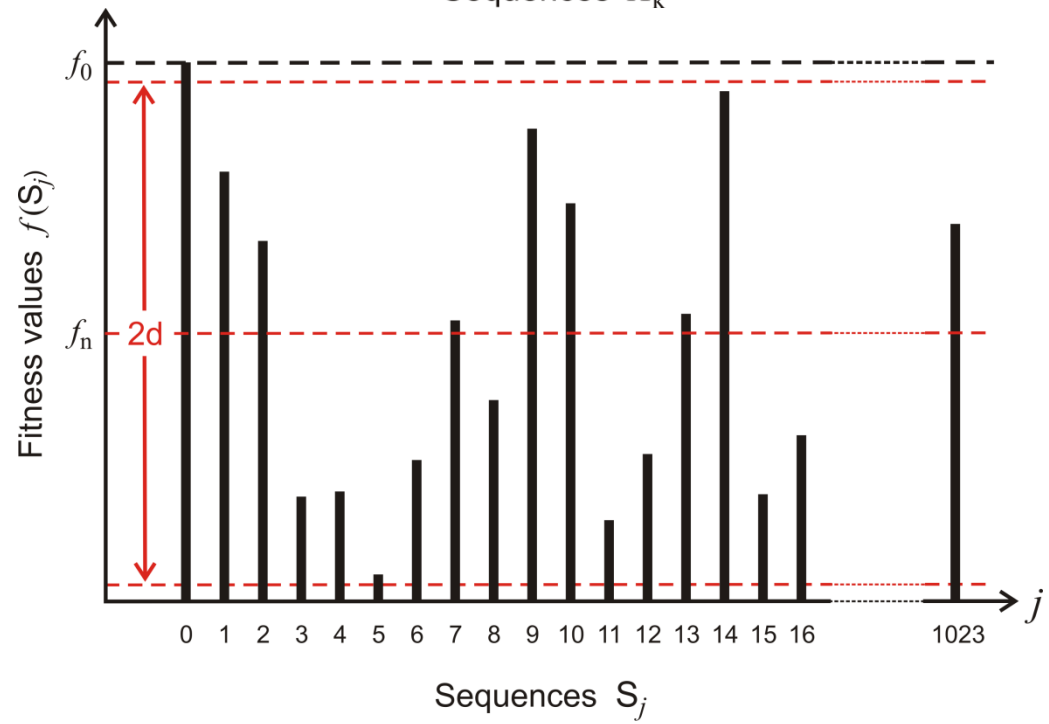
3. Combinatorial explosion: the number of possible genomes is prohibitive for systematic searches

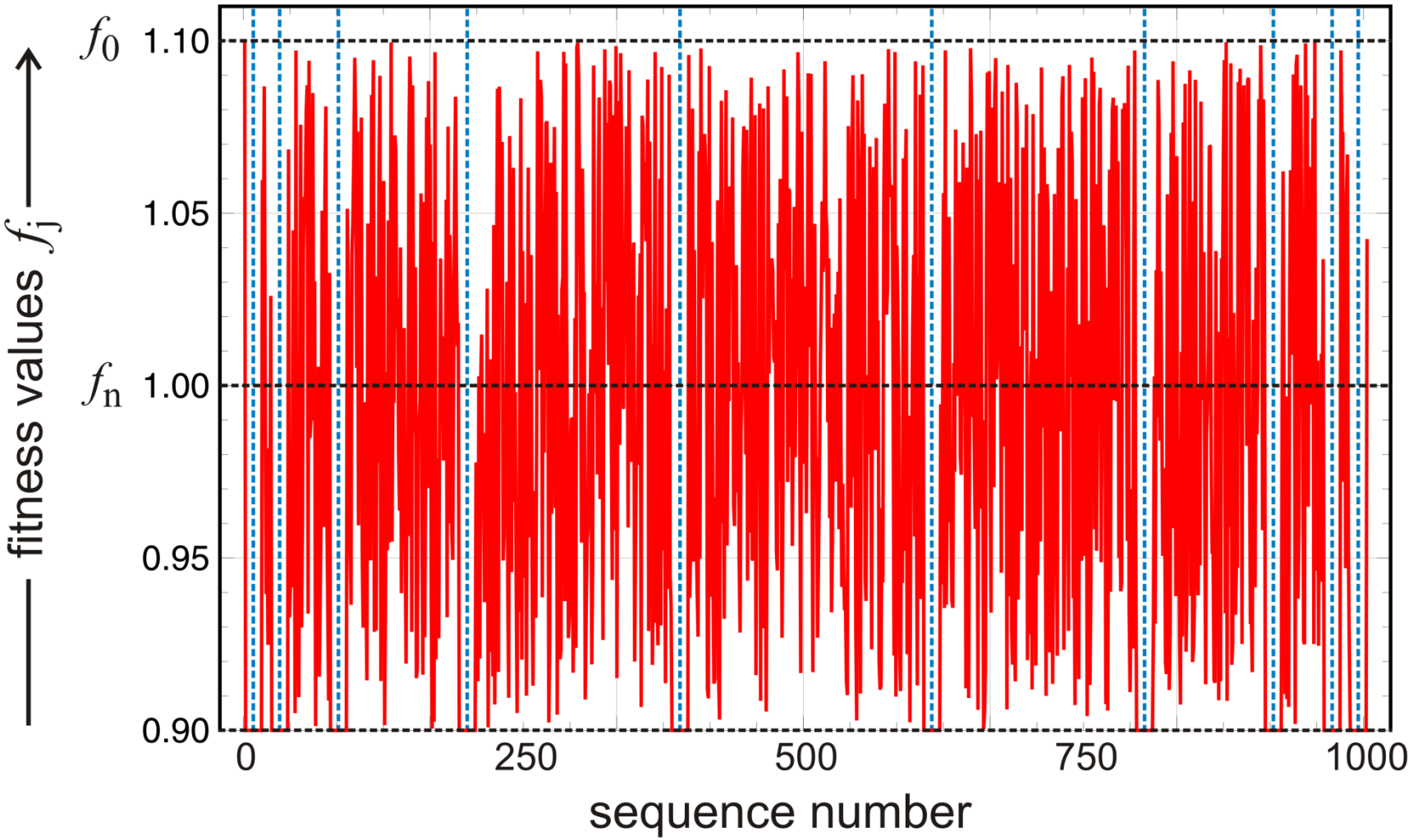
Facit: Any successful and applicable theory of molecular evolution must be able to predict evolutionary dynamics from a small or at least in practice measurable number of fitness values.

single peak landscape

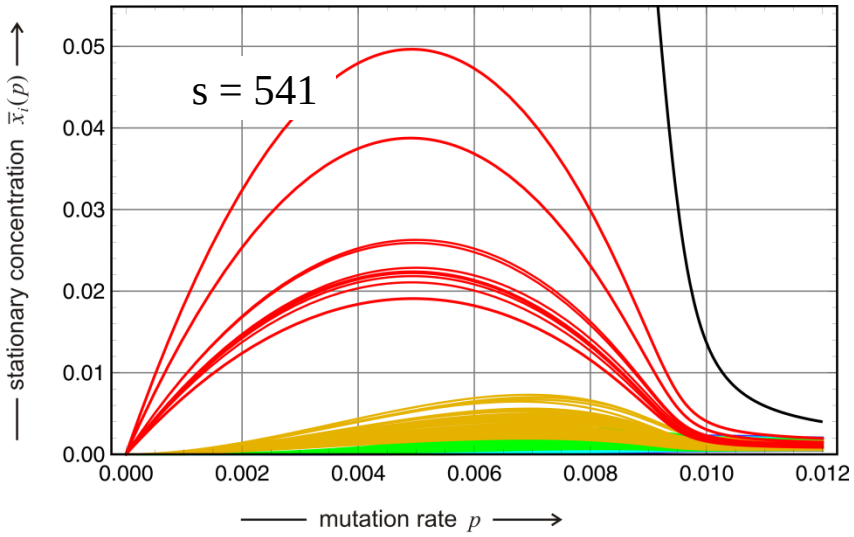


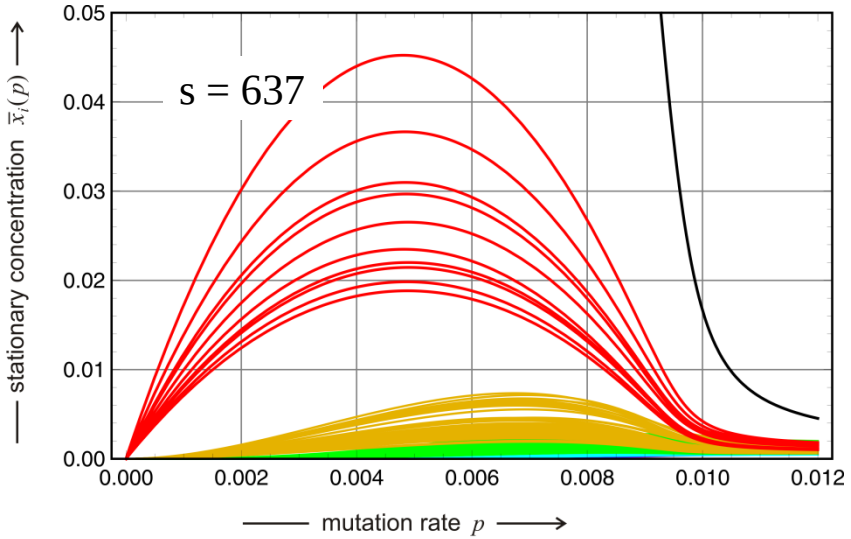
„realistic“ landscape

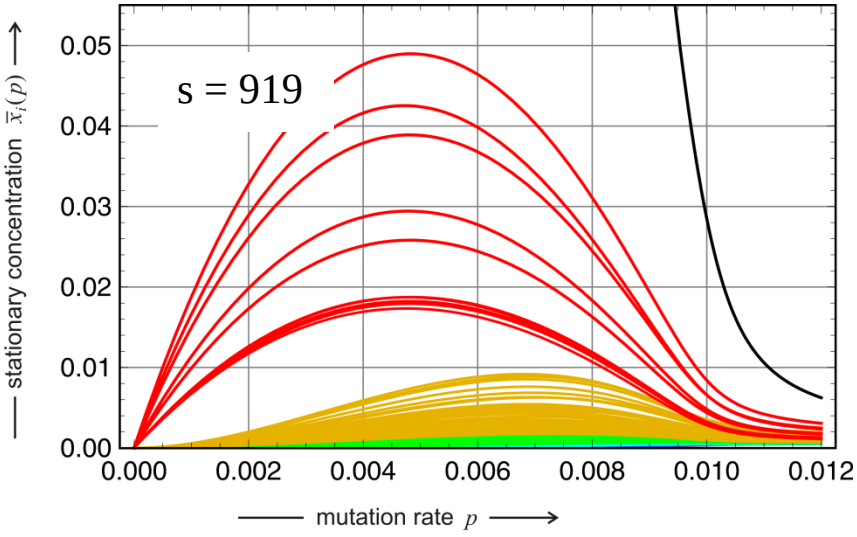




Random distribution of fitness values: d = 1.0 and s = 637

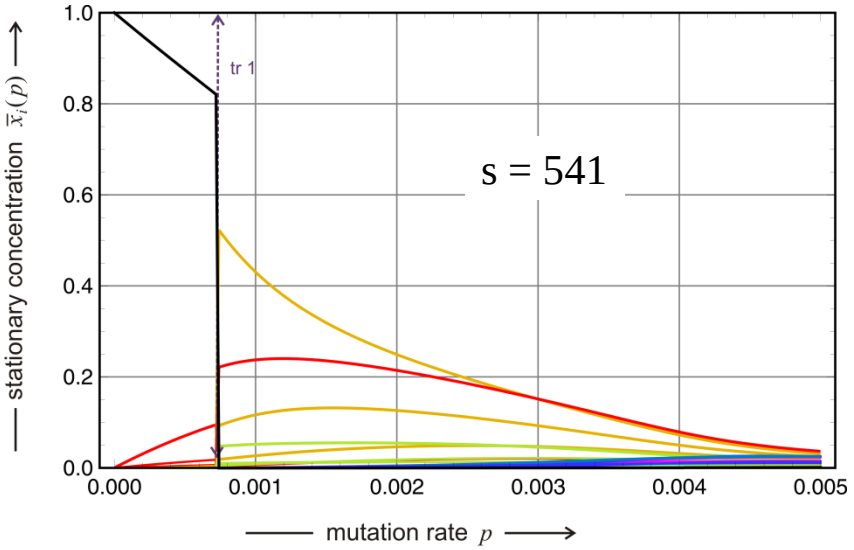


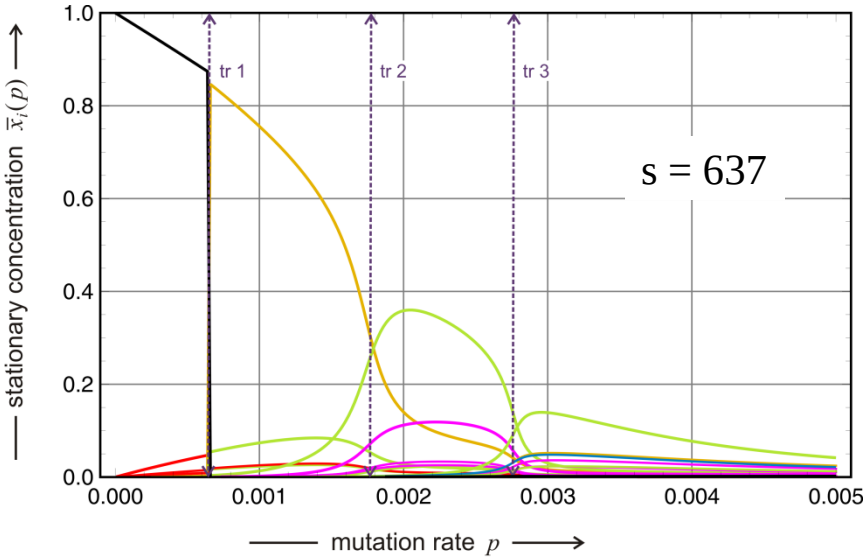


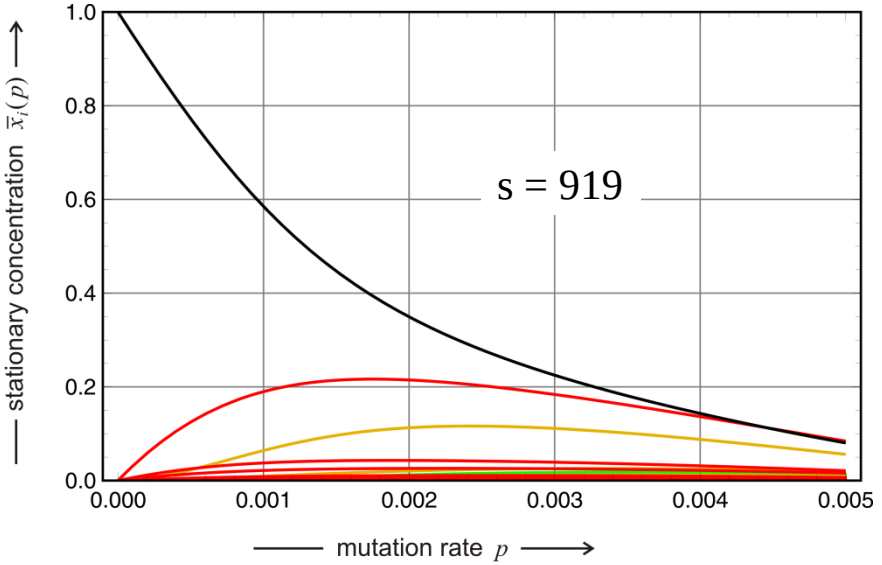


Error threshold on ‚realistic' landscapes

n = 10, $f_0 = 1.1$, $f_n = 1.0$, d = 0.5

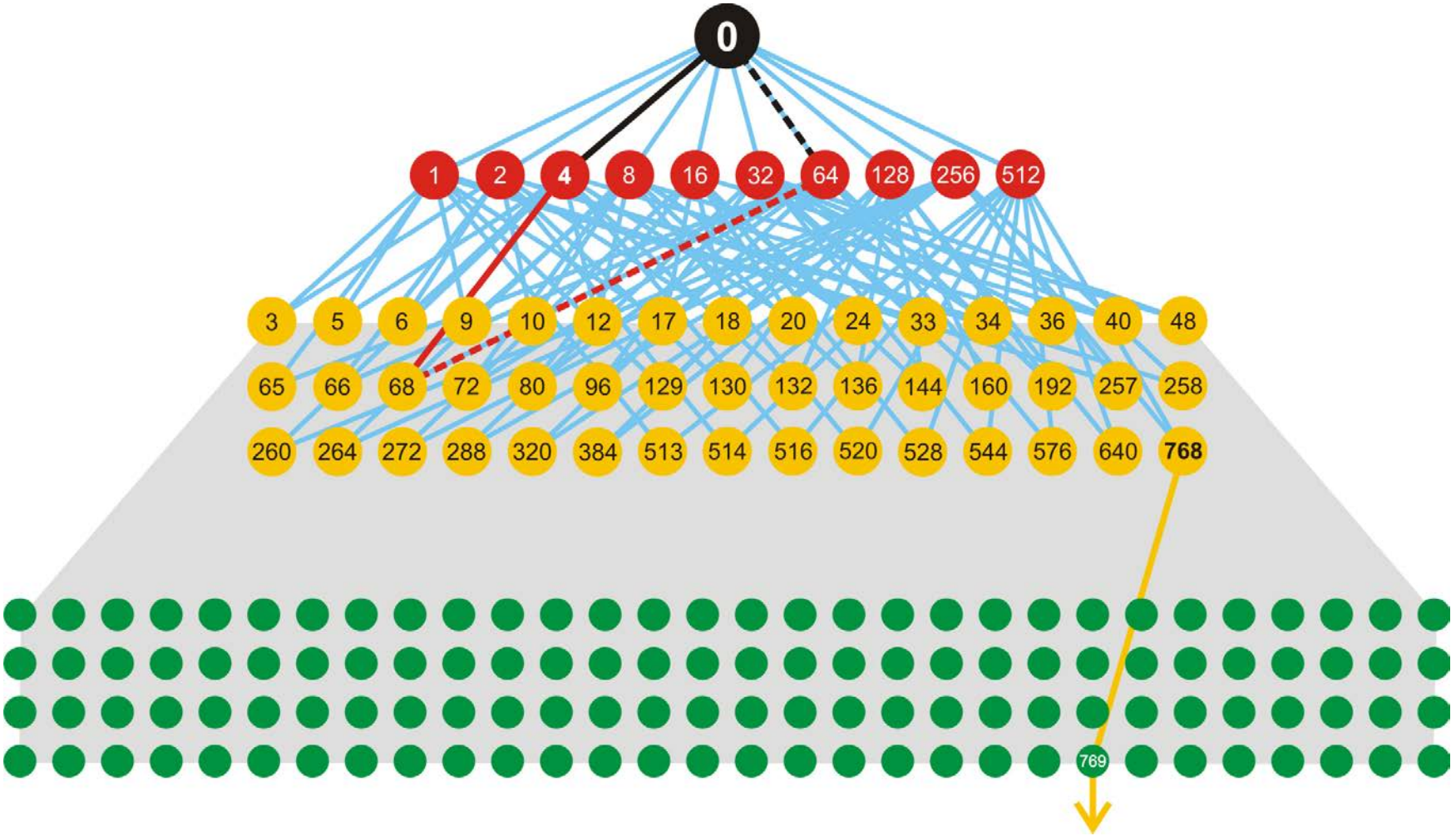




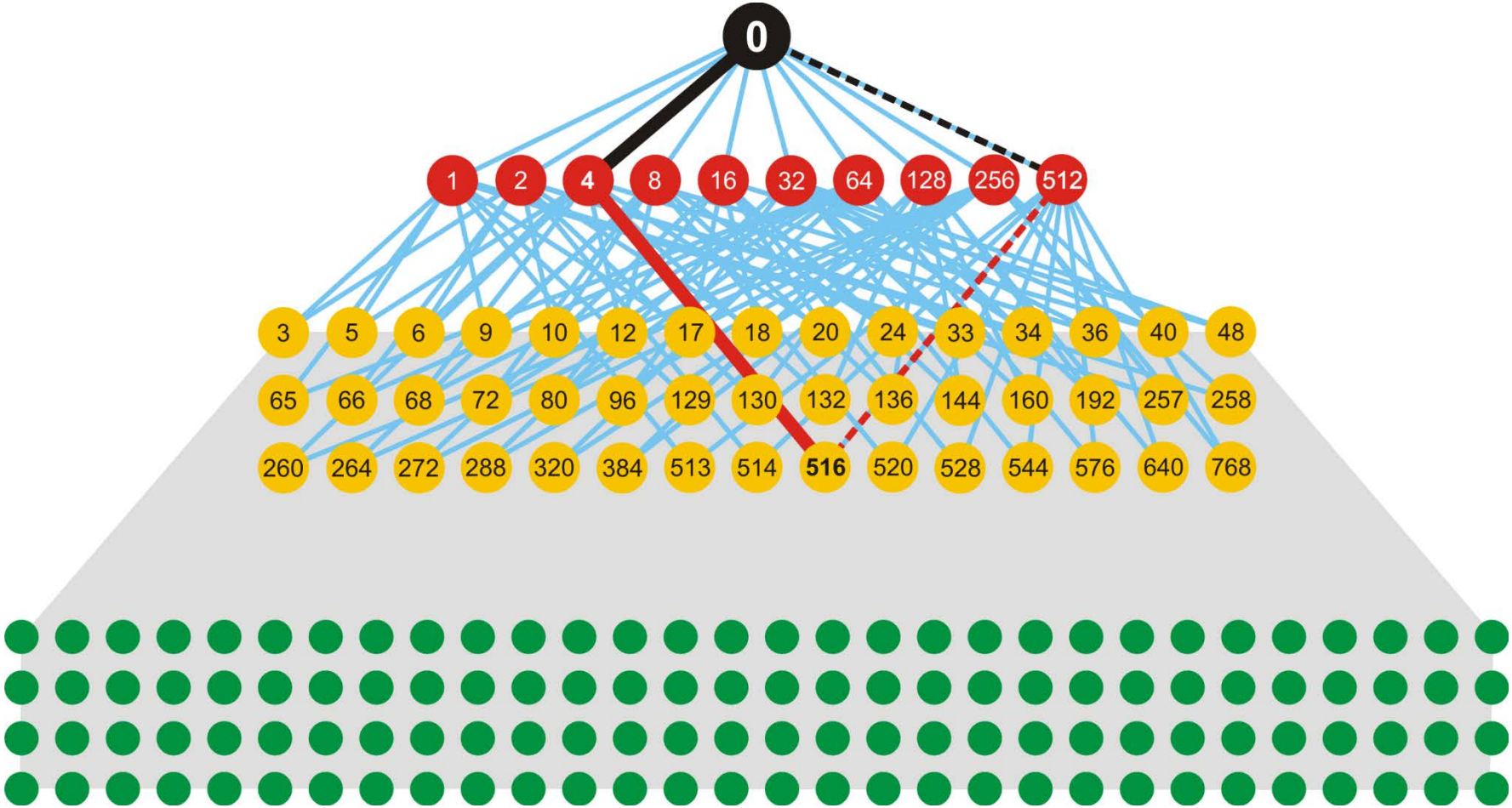


Error threshold on ‚realistic' landscapes

n = 10, $f_0 = 1.1$, $f_n = 1.0$, d = 1.0



Determination of the dominant mutation flow: $d = 1$, $s = 613$



Determination of the dominant mutation flow: $d = 1$, $s = 919$

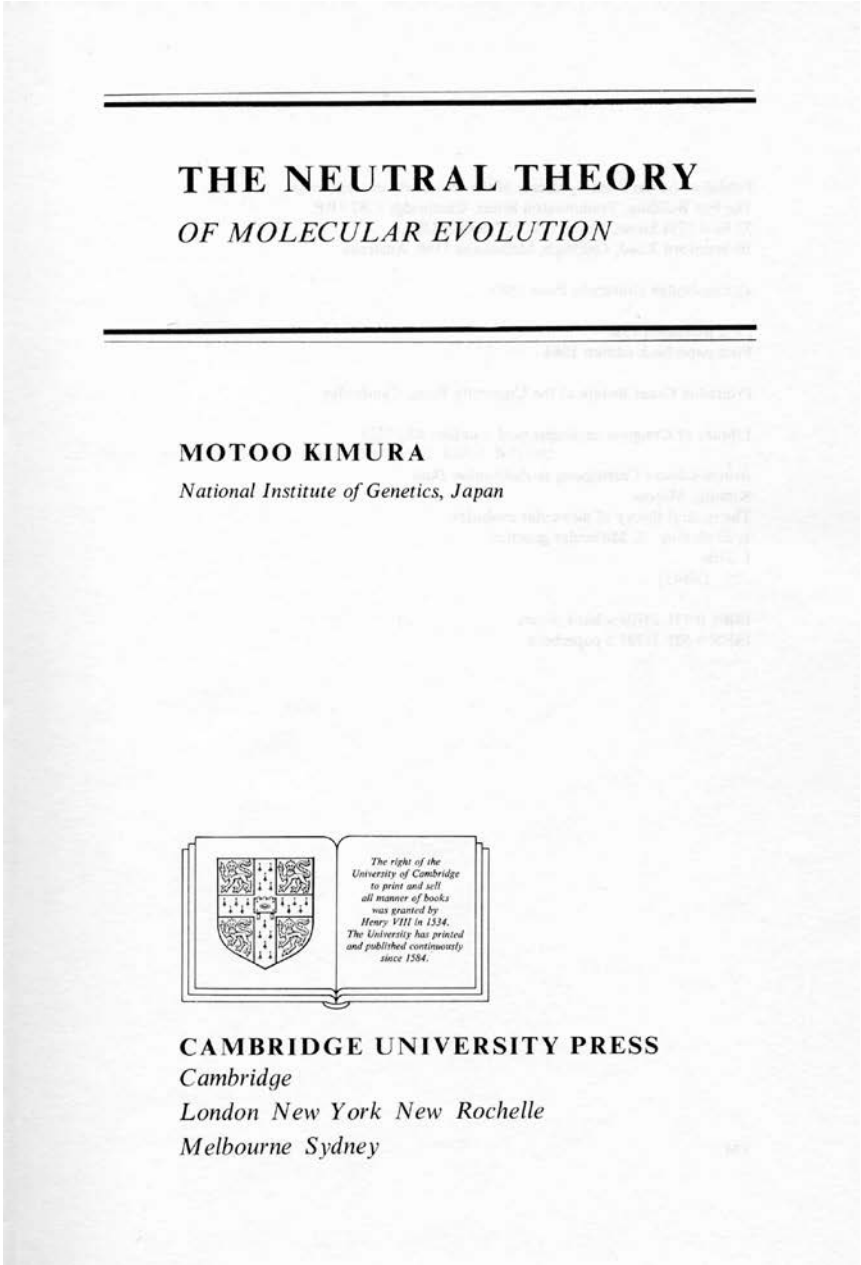
1. Prologue
2. Darwin and replicating molecules
3. In vitro evolution
4. „Simple“ landscapes
5. „Realistic“ landscapes
6. **Neutrality in evolution**
7. Perspectives of systems biology

Motoo Kimura, 1924 - 1994

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

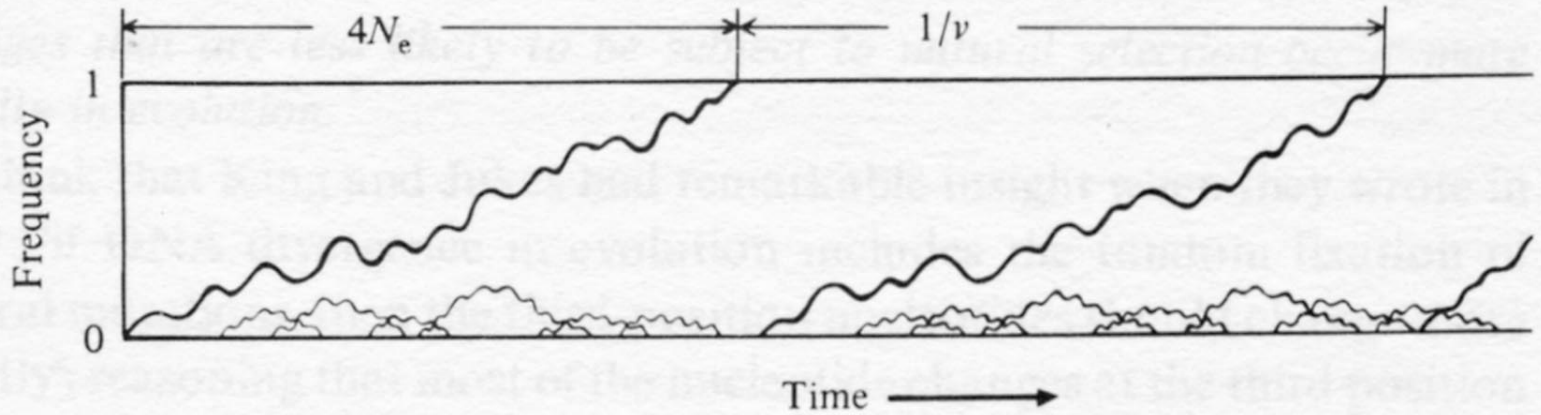
OF MOLECULAR EVOLUTION

MOTOO KIMURA

National Institute of Genetics, Japan

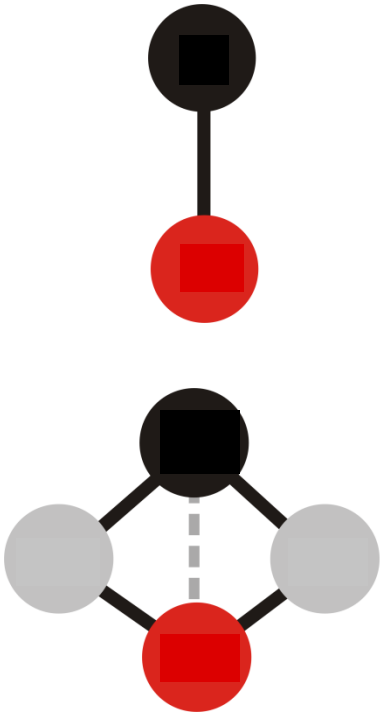
CAMBRIDGE UNIVERSITY PRESS
Cambridge
London New York New Rochelle
Melbourne Sydney

Fig. 3.1. Behavior of mutant genes following their appearance in a finite population. Courses of change in the frequencies of mutants destined to fixation are depicted by thick paths. N_e stands for the effective population size and v is the mutation rate.



Motoo Kimura

Is the Kimura scenario correct for frequent mutations?



$\mathbf{d_H = 1}$

$$\lim_{p\to 0} x_1(p) = x_2(p) = 0.5$$

$\mathbf{d_H = 2}$

$$\lim_{p\to 0} x_1(p) = \alpha/(1+\alpha)$$

$$\lim_{p\to 0} x_2(p) = 1/(1+\alpha)$$

$\mathbf{d_H \geq 3}$

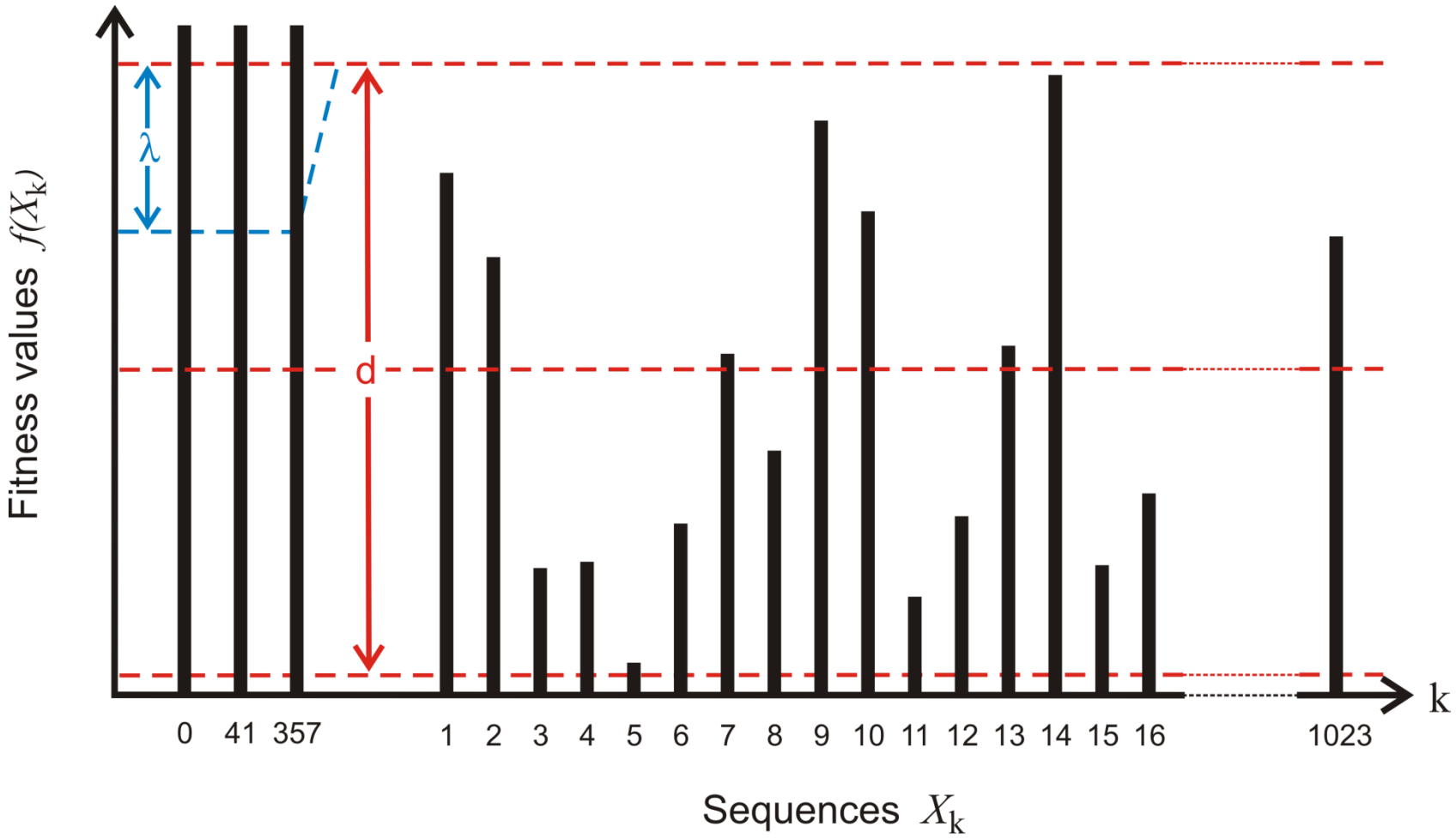
$$\lim_{p\to 0} x_1(p) = 1, \lim_{p\to 0} x_2(p) = 0 \quad \text{or}$$

$$\lim_{p\to 0} x_1(p) = 0, \lim_{p\to 0} x_2(p) = 1$$

Pairs of neutral sequences in replication networks

Random fixation in the sense of Motoo Kimura

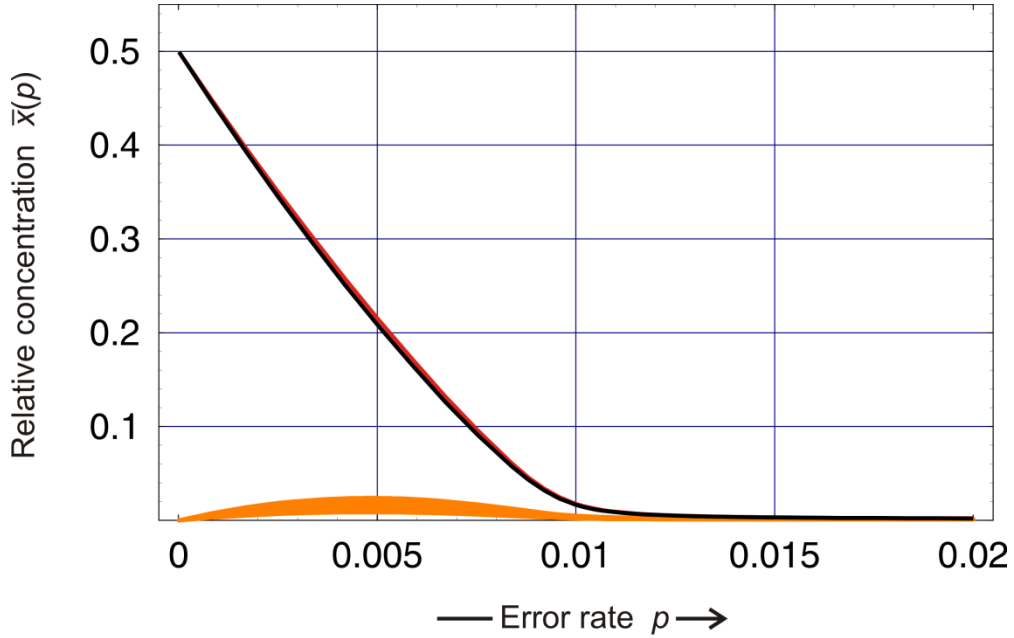
P. Schuster, J. Swetina. 1988. Bull. Math. Biol. 50:635-650



A fitness landscape including neutrality

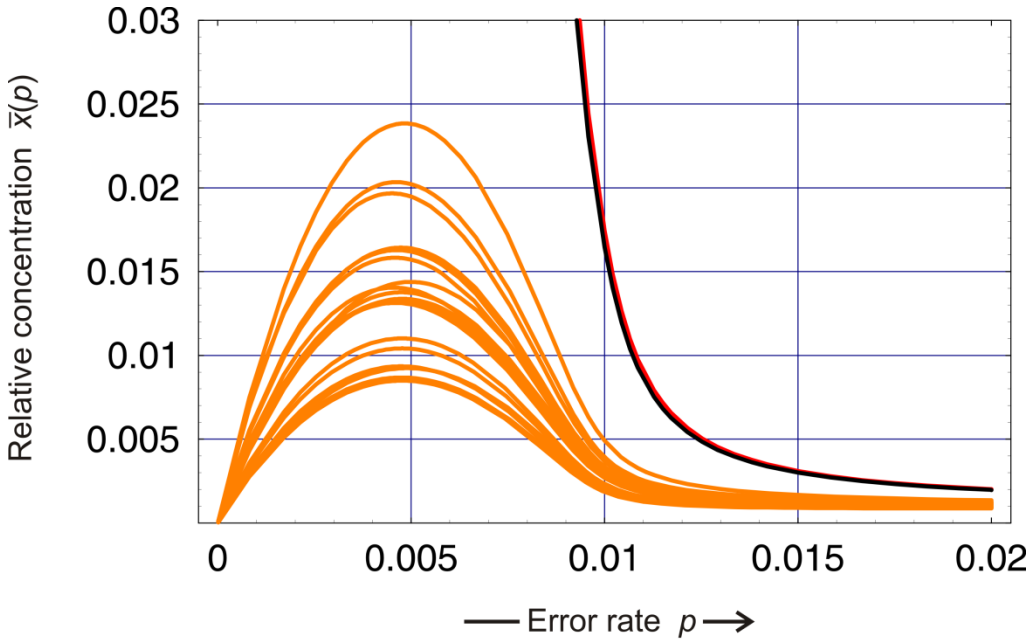
Neutral network

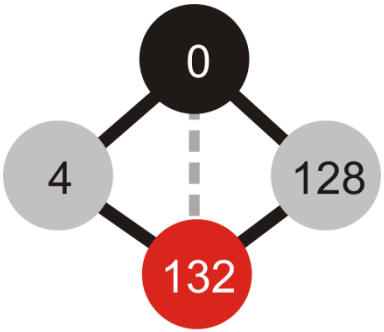
$\lambda = 0.01,\ s = 367$



Neutral network: Individual sequences

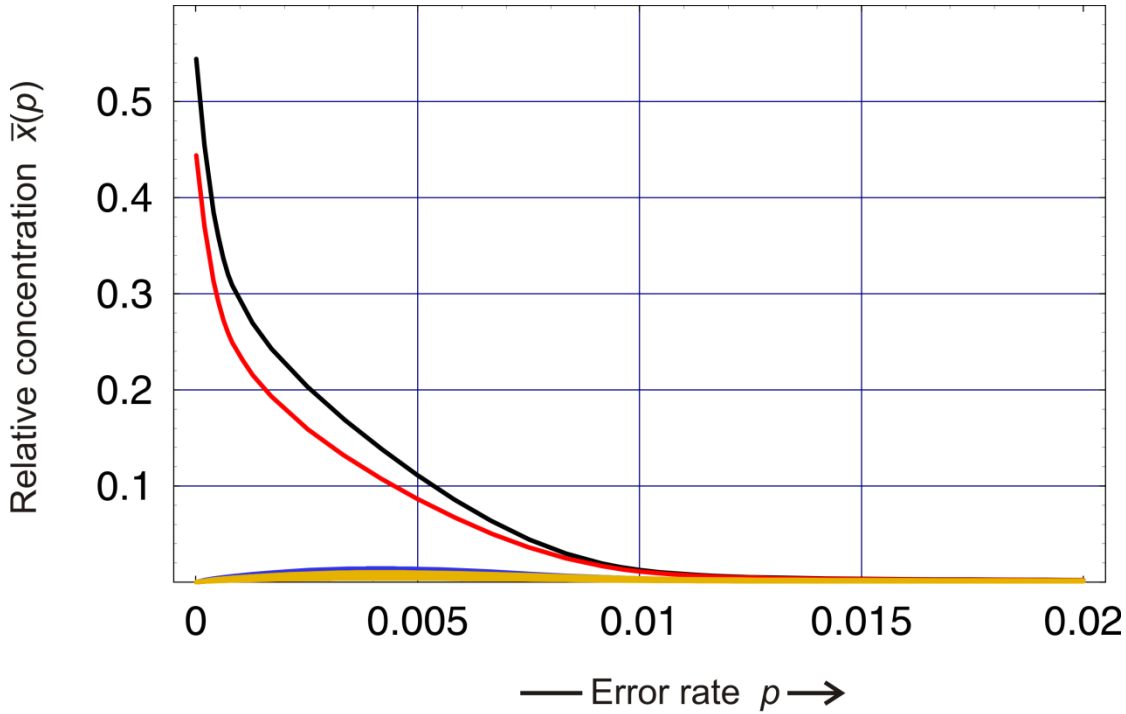
$n = 10,\ \sigma = 1.1,\ d = 1.0$





Neutral network

$\lambda = 0.01,\ s = 877$

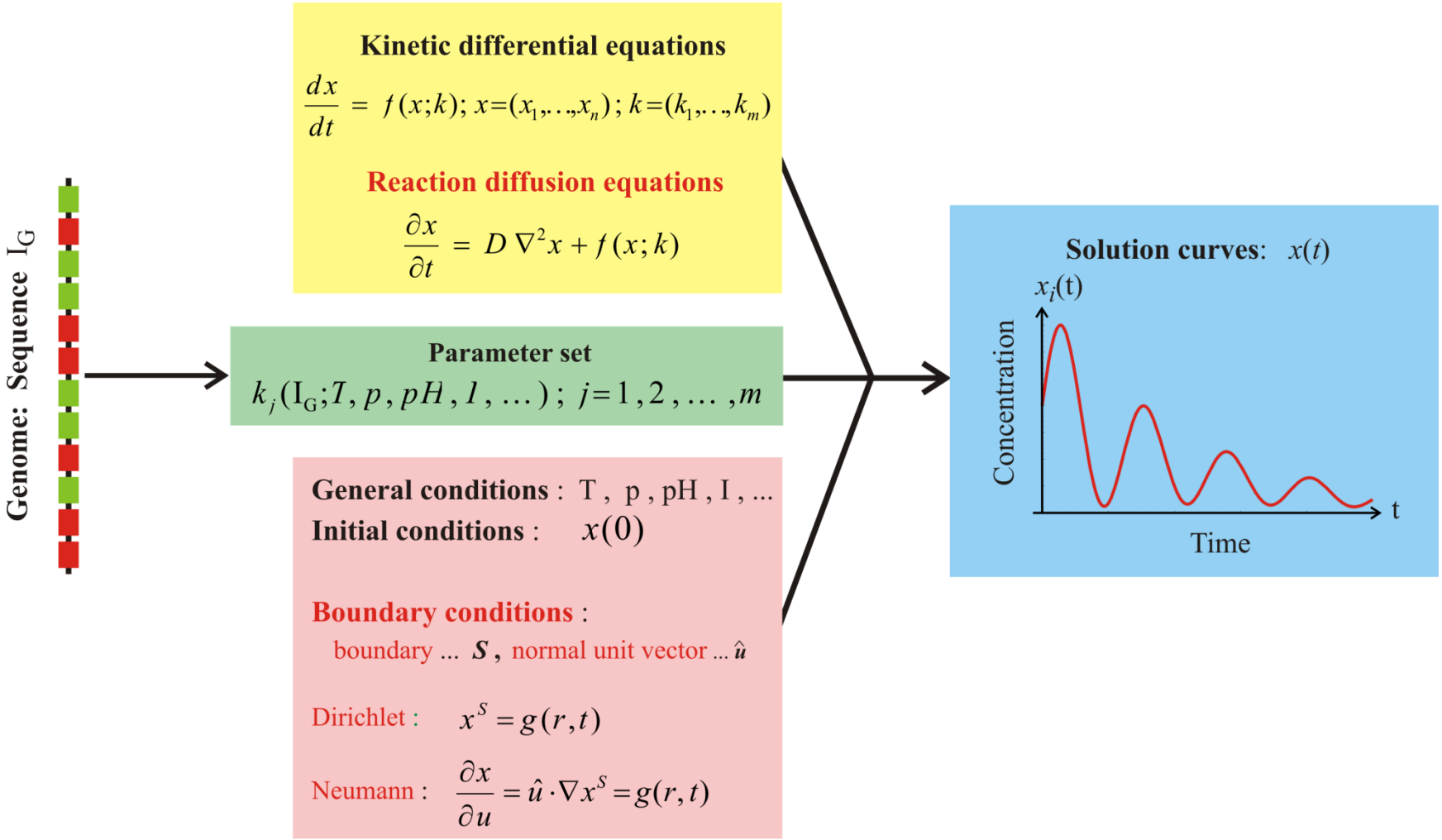


Neutral network: Individual sequences

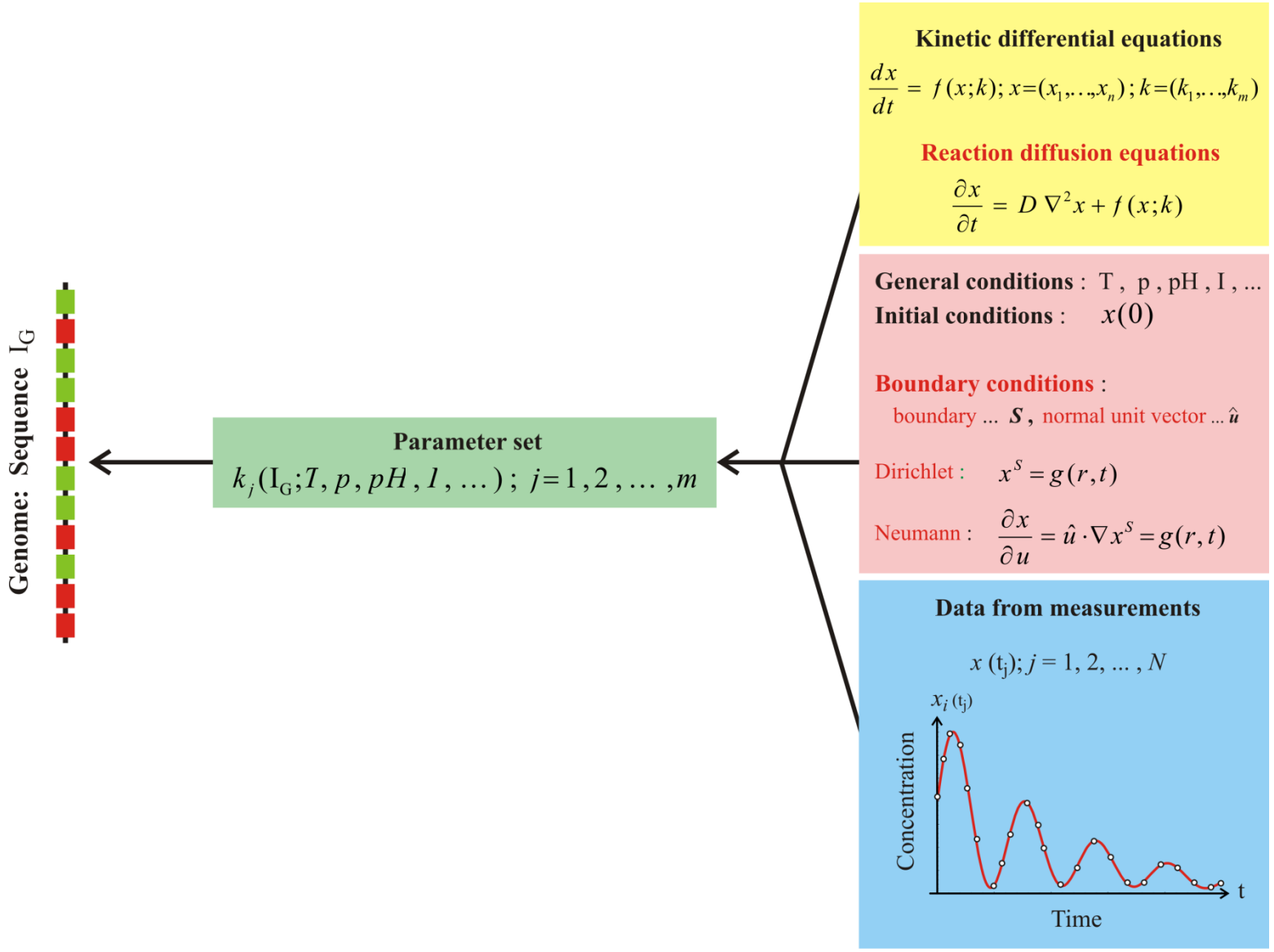
$n = 10,\ \sigma = 1.1,\ d = 1.0$

Consensus sequences of a quasispecies of two strongly coupled sequences of Hamming distance $d_H(X_i,X_j) = 1$ and 2.

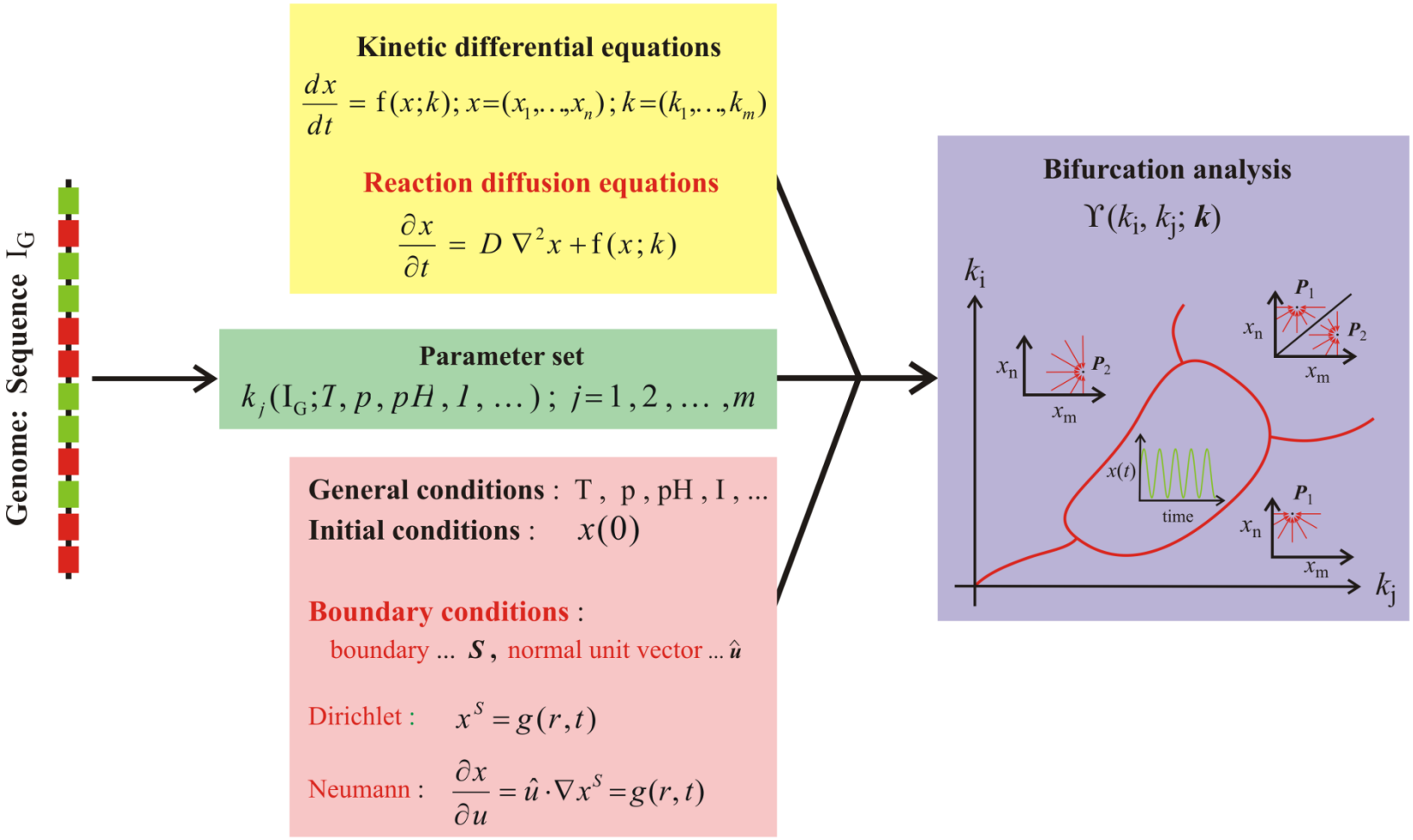
1. Prologue
2. Darwin and replicating molecules
3. In vitro evolution
4. „Simple“ landscapes
5. „Realistic“ landscapes
6. Neutrality in evolution
7. **Perspectives of systems biology**



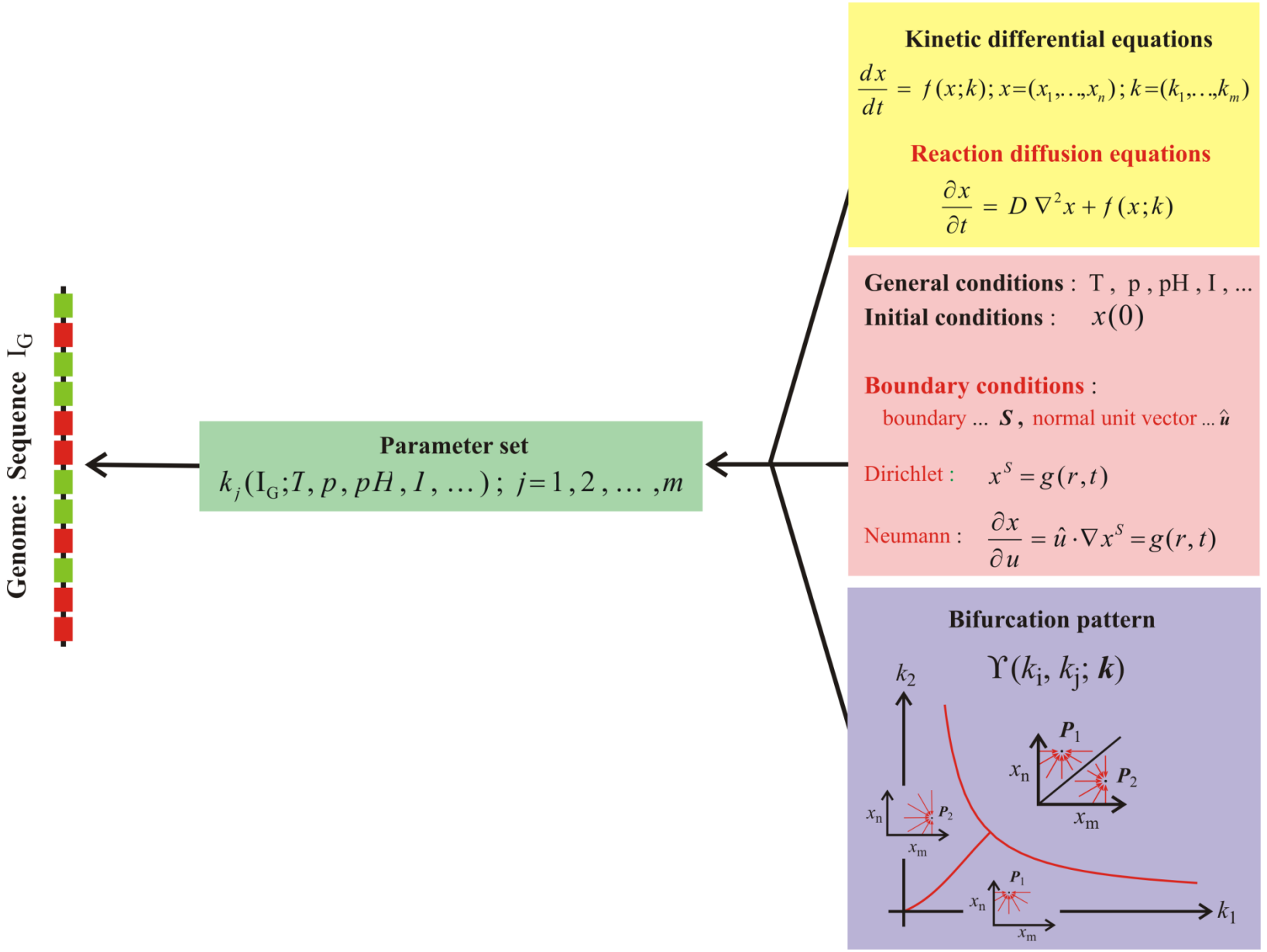
H. Engl, C. Flamm, P. Kügler, J. Lu, S. Müller, P. Schuster. 2009.
Inverse problems in systems biology. Inverse Problems **25**:e123014.



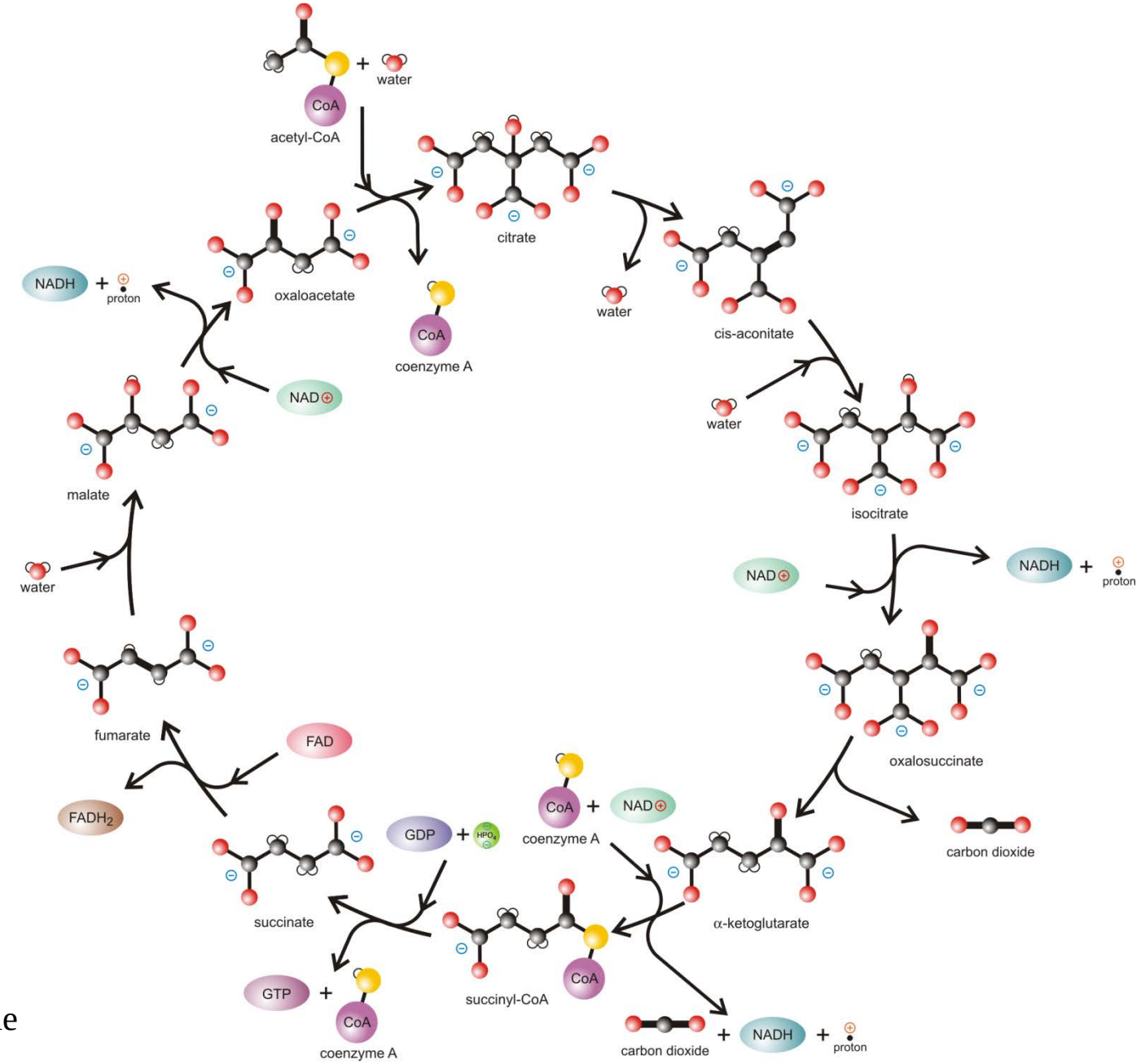
H. Engl, C. Flamm, P. Kügler, J. Lu, S. Müller, P. Schuster. 2009.
Inverse problems in systems biology. Inverse Problems **25**:e123014.



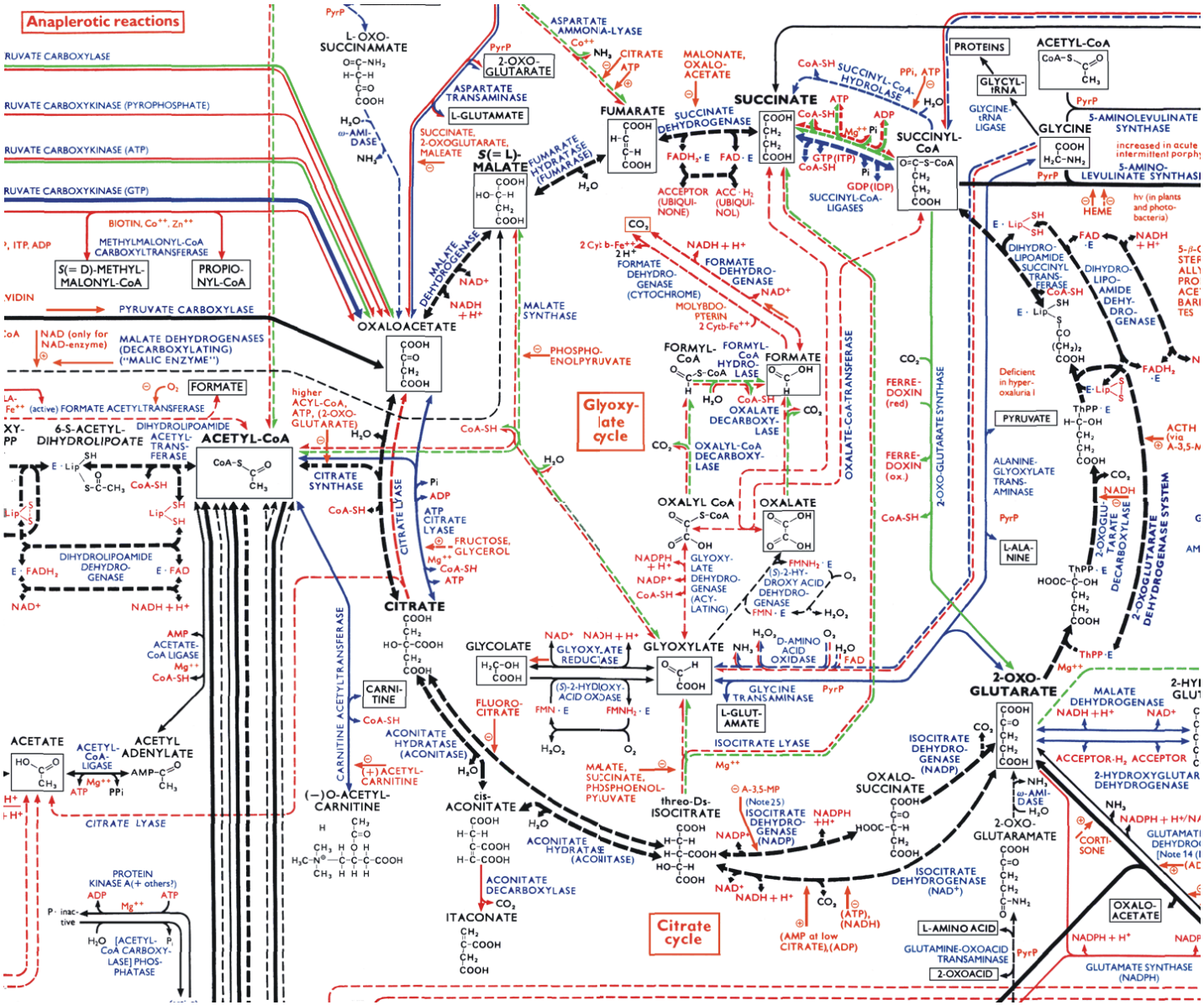
H. Engl, C. Flamm, P. Kügler, J. Lu, S. Müller, P. Schuster. 2009.
Inverse problems in systems biology. Inverse Problems **25**:e123014.



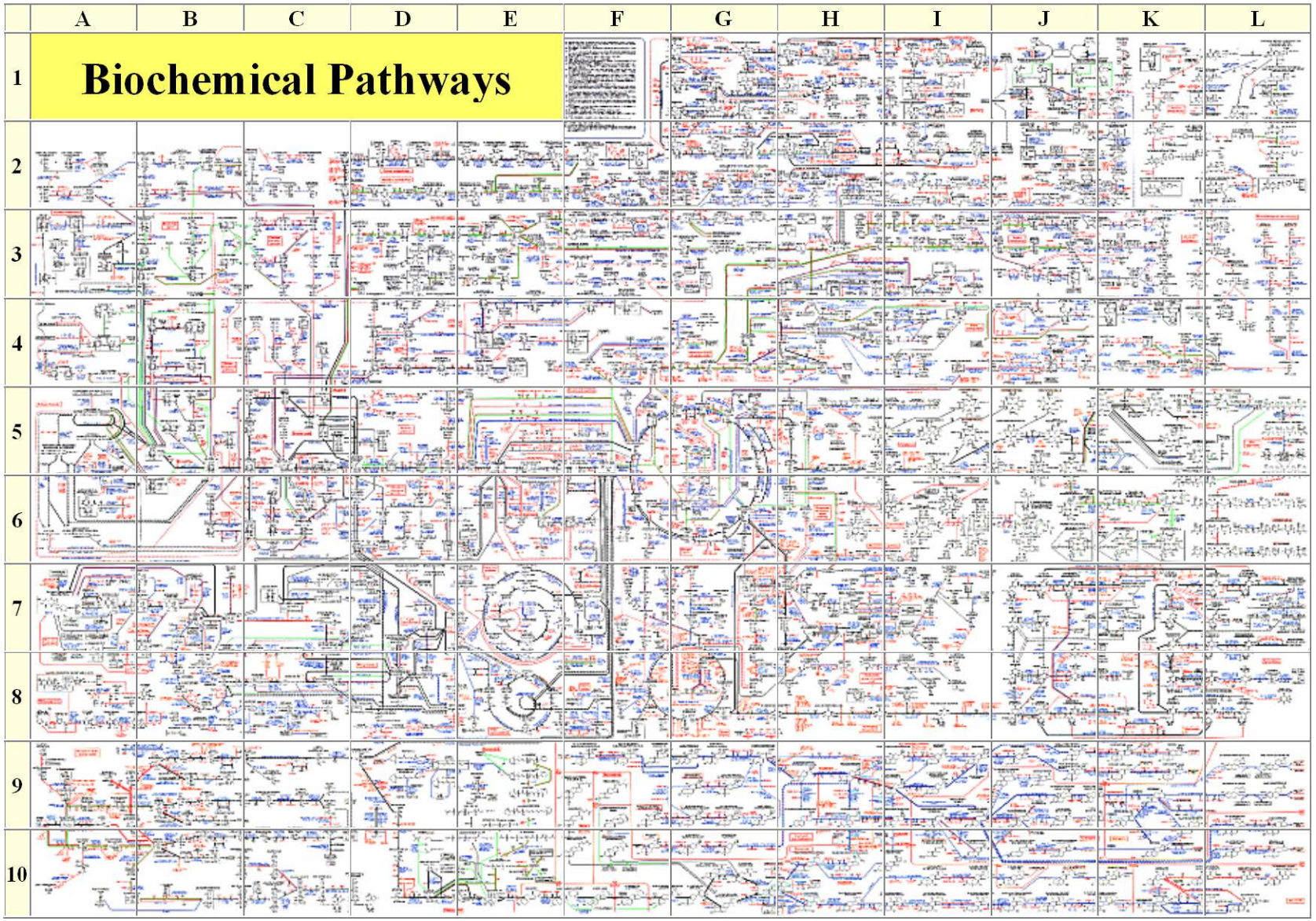
H. Engl, C. Flamm, P. Kügler, J. Lu, S. Müller, P. Schuster. 2009. Inverse problems in systems biology. Inverse Problems **25**:e123014.



Cellular citrate cycle



The citrate or Krebs cycle



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

ENCODE stands for
ENCyclopedia Of DNA Elements.

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

Mycoplasma pneumoniae:	genome length	820 000 bp
	# genes:	733
	# proteins (ORF):	689
	# tRNAs	37
	# rRNAs	3
	# other RNAs	4

S. Kühner, V. van Noort, M. J. Betts, A. Leo-Macias, C. Batisse, M. Rode, T. Yamada, T. Maier, S. Bader, P. Beltran-Alvarez, D. Castaño-Diez, W.-H. Chen, D. Devos, M. Güell, T. Norambuena, I. Racke, V. Rybin, A. Schmidt, E. Yus, R. Aebersold, R. Herrmann, B. Böttcher, A. S. Frangakis, R. B. Russell, L. Serrano, P. Bork, and A.-C. Gavin. 2009.
Proteome organization in a genome-reduced bacterium. Science **326**:1235–1240.

E. Yus, T. Maier, K. Michalodimitrakis, V. van Noort, T. Yamada, W.-H. Chen, J. A. Wodke, M. Güell, S. Martínez, R. Bourgeois, S. Kühner, E. Raineri, I. Letunic, O. V. Kalinina, M. Rode, R. Herrmann, R. Gutiérez-Gallego, R. B. Russell, A.-C. Gavin, P. Bork, and L. Serrano. 2009.
Impact of genome reduction on bacterial metabolism and its regulation. Science **326**:1263–1268.

M. Güell, V. van Noort, E. Yus, W.-H. Chen, J. Leigh-Bell, K. Michalodimitrakis, T. Yamada, M. Arumugam, T. Doerks, S. Kühner, M. Rode, M. Suyama, S. Schmidt, A.-C. Gavin, P. Bork, and L. Serrano. 2009.
Transcriptome complexity in a genome-reduced bacterium. Science **326**:1268–1271.

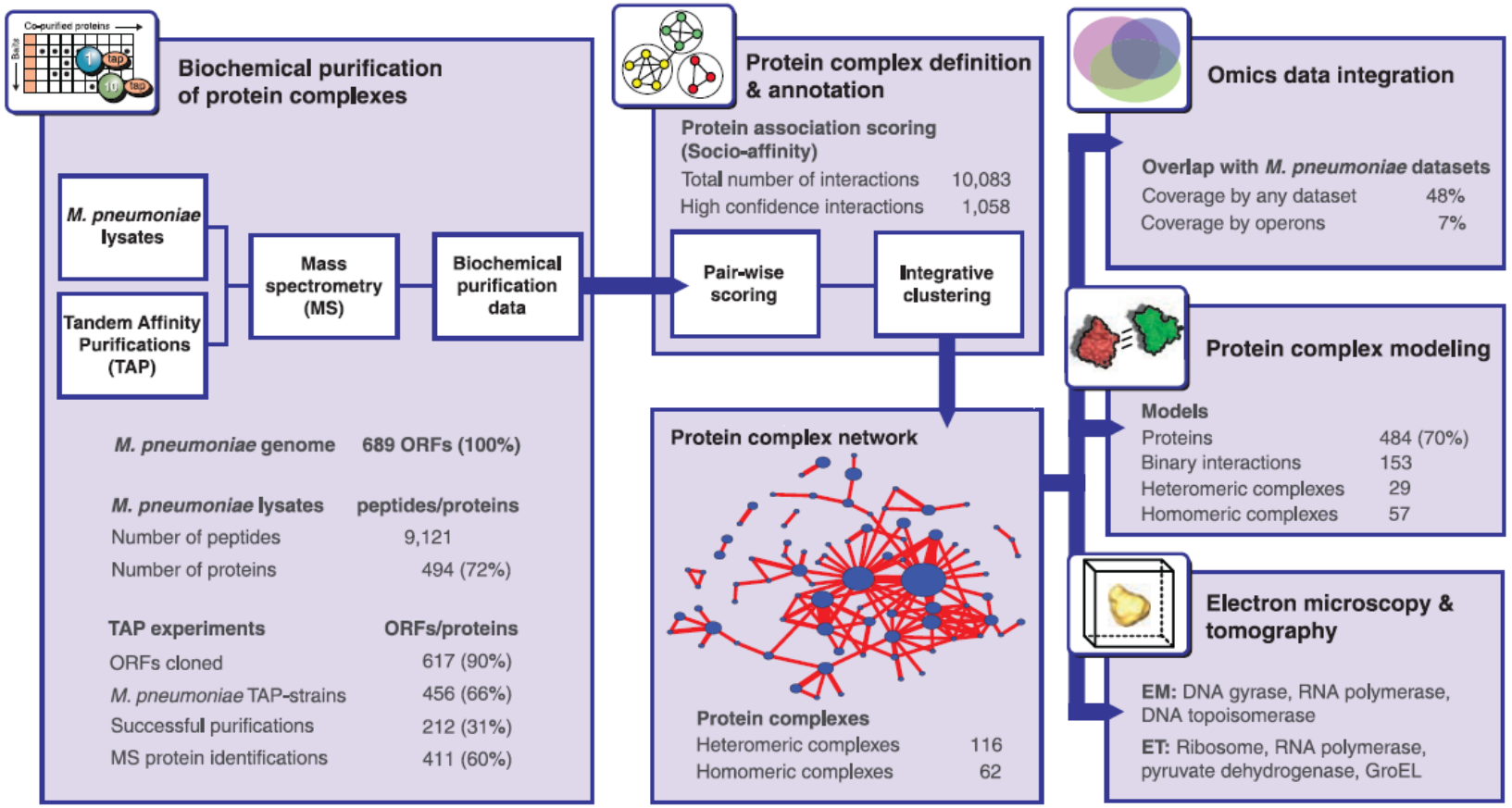


Fig. 1. Synopsis of the genome-wide screen of complexes in *M. pneumoniae*.

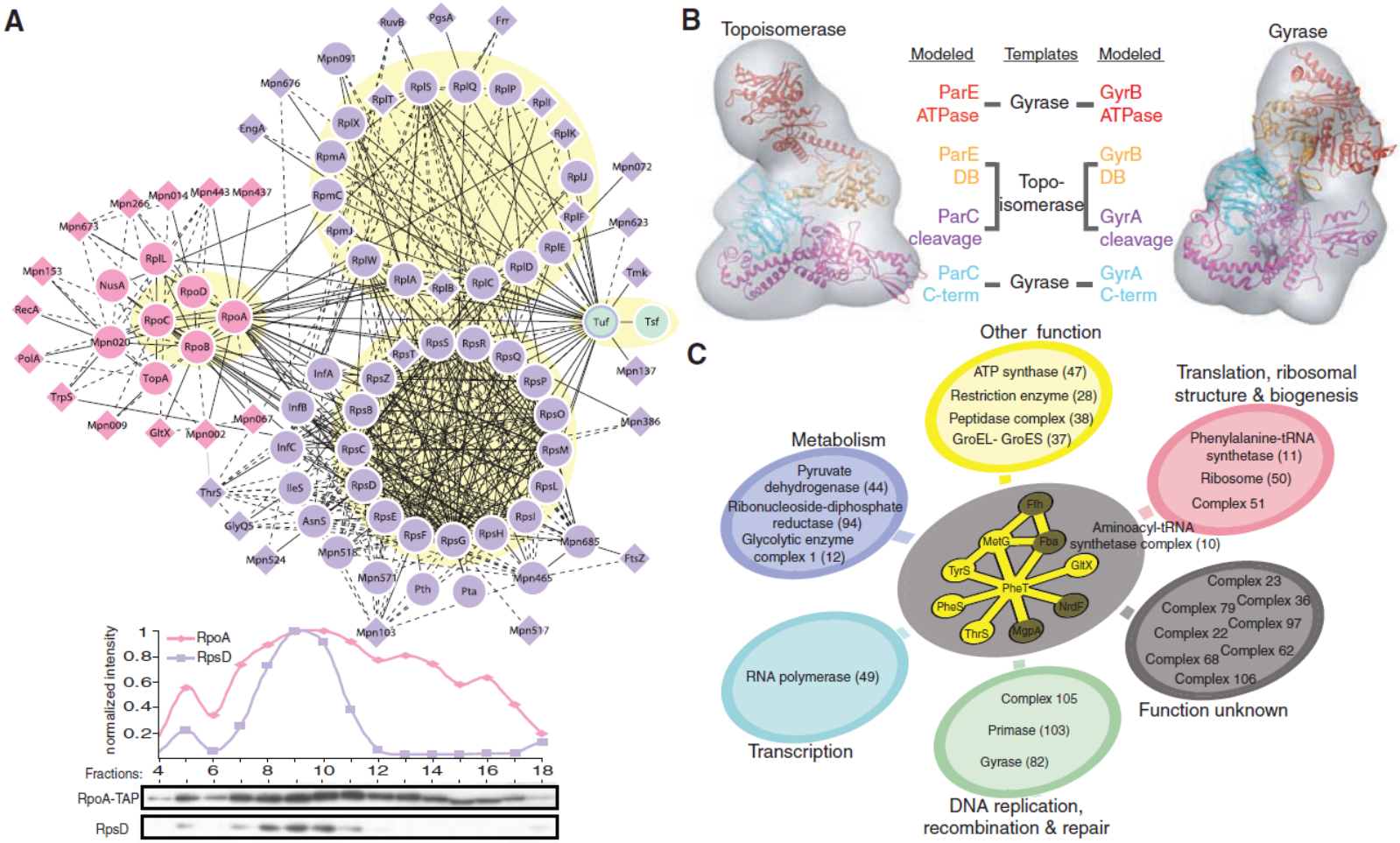


Fig. 3. Higher level of proteome organization. (**A**) The RNA polymerase–ribosome assembly. Core components are represented by circles, attachments by diamonds. The line attribute corresponds to socio-affinity indices: dashed lines, 0.5 to 0.86; plain lines, >0.86. Color code and shaded yellow circles around groups of proteins refer to individual complexes: RNA polymerase (pink), ribosome (purple), and translation elongation factor (green). The bottom graph shows that the ribosomal protein RpsD (23 kD) and the α subunit of the RNA polymerase, RpoA-TAP (57 kD), co-elute in high molecular weight fractions (MD range) during gel filtration chromatography. (**B**) DNA topoisomerase (diameter ~ 12 nm) is a heterodimer in bacteria: ParE (ATPase and DNA binding domains) and ParC (cleavage and C-terminal domains). The interaction between ParE-DNA–binding and ParC–cleavage domains was modeled by using yeast topoisomerase II as a template [Protein Data Bank (PDB) code 2rgr], and ParE-ATPase and ParC–C-terminal domains were modeled separately on structures of gyrase homologs (PDB 1kij and 1suu). All four domains were fitted into the electron microscopy density. Gyrase (~12 nm) is similarly split in bacteria into GyrA/GyrB, which are paralogs of ParE/ParC, and was modeled and fitted by using PDB 1bjt as a template for the GyrB-DNA–binding and GyrA-cleavage domains interaction. (**C**) Protein multifunctionality in *M. pneumoniae* illustrated with the AARS complexes.

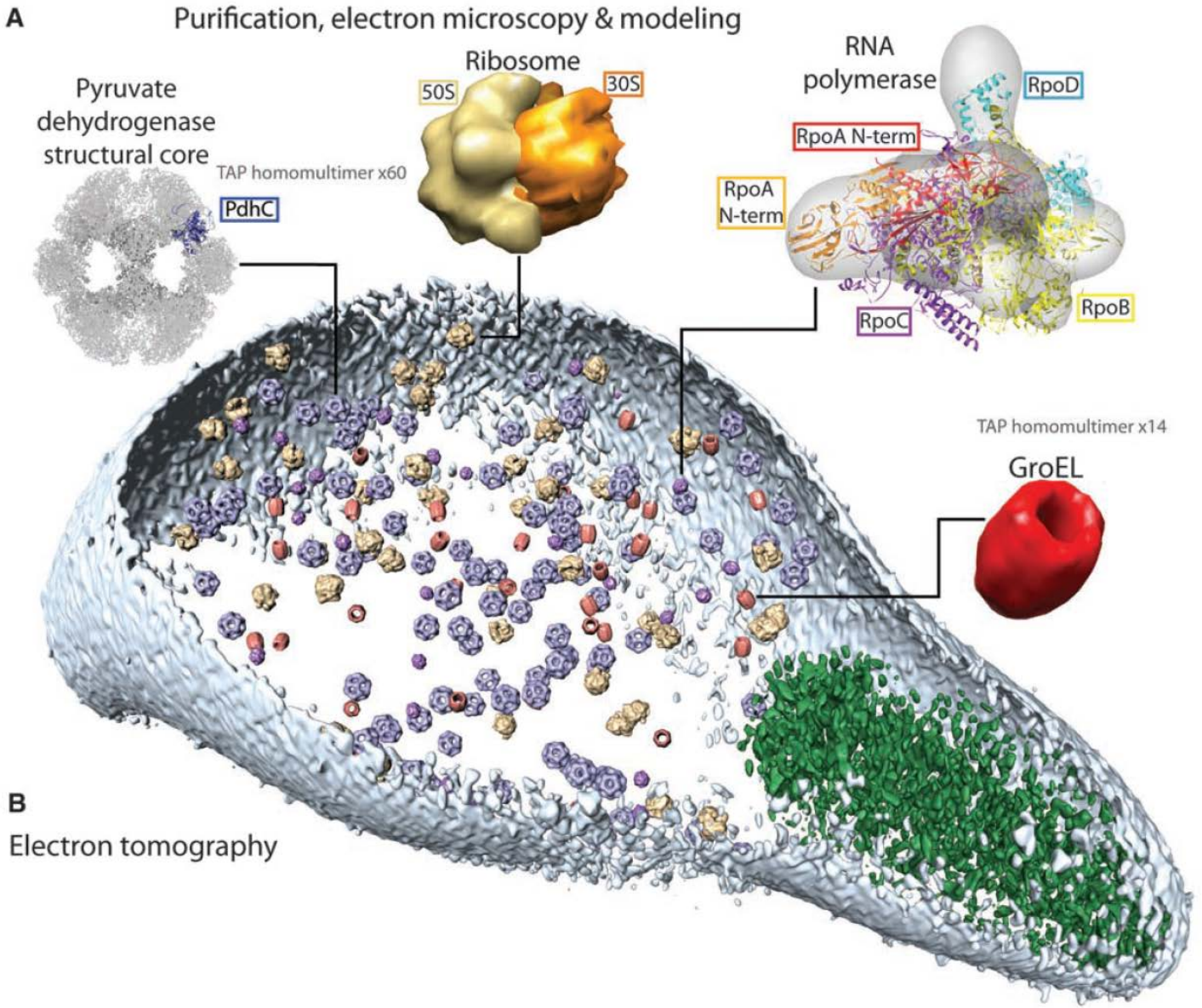


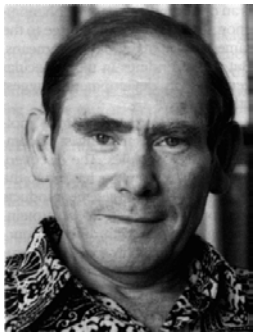
Fig. 4. From proteomics to the cell. By a combination of pattern recognition and classification algorithms, the following TAP-identified complexes from *M. pneumoniae*, matching to existing electron microscopy and x-ray and tomogram structures (**A**), were placed in a whole-cell tomogram (**B**): the structural core of pyruvate dehydrogenase in blue (~23 nm), the ribosome in yellow (~26 nm), RNA polymerase in purple (~17 nm), and GroEL homomultimer in red (~20 nm). Cell dimensions are ~300 nm by 700 nm. The cell membrane is shown in light blue. The rod, a prominent structure filling the space of the tip region, is depicted in green. Its major structural elements are HMW2 (Mpn310) in the core and HMW3 (Mpn452) in the periphery, stabilizing the rod (*42*).The individual complexes (A) are not to scale, but they are shown to scale within the bacterial cell (B).

Advantages of the molecular approach

1. Complex reproduction mechanisms are readily included.

2. Gene regulation – DNA or RNA based – is chemical kinetics!

3. Accounting for epigenetic effects requires just the simultaneous consideration of several generations.



Sydney Brenner, 1927 -

What else is epigenetics than a funny form of enzymology ? Each protein, after all, comes from some piece of DNA.

Coworkers

Universität Wien

Peter Stadler, **Bärbel M. Stadler**, Bioinformatik**,** Universität Leipzig, GE

Walter Fontana, Harvard Medical School, MA

Martin Nowak, Harvard University, MA

Sebastian Bonhoeffer, Theoretical Biology, ETH Zürich, CH

Christian Reidys, Mathematics, University of Southern Denmark, Odense, DK

Christian Forst, Southwestern Medical Center, University of Texas, Dallas, TX

Thomas Wiehe, Institut für Genetik, Universität Köln, GE

Ivo L.Hofacker, Theoretische Chemie**,** Universität Wien, AT

Kurt Grünberger, Michael Kospach, **Andreas Wernitznig**, **Ulrike Langhammer**, **Ulrike Mückstein**, Theoretische Chemie, Universität Wien, AT

Universität Wien

Acknowledgement of support

Fonds zur Förderung der wissenschaftlichen Forschung (FWF)

Jubiläumsfonds der Österreichischen Nationalbank

European Commission

Austrian Genome Research Program – GEN-AU

Österreichische Akademie der Wissenschaften

Siemens AG, Austria

Universität Wien and The Santa Fe Institute

Thank you for your attention!

Web-Page for further information:

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