The Mathematics of Biological Evolution

Peter Schuster

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

ÖMG-DMV Congress 2013

Minisymposium: Mathematics of the Planet Earth

Innsbruck, 23.– 24.09.2013

Web-Page for further information:

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

1. Prologue
2. Mathematics of Darwin's natural selection
3. Mendel, Fisher and population genetics
4. Mutations and selection
5. What means neutrality in evolution?
6. Evolution in simple systems
7. Some origins of complexity in biology

1. **Prologue**
2. Mathematics of Darwin's natural selection
3. Mendel, Fisher and population genetics
4. Mutations and selection
5. What means neutrality in evolution?
6. Evolution in simple systems
7. Some origins of complexity in biology

$A + X \rightarrow 2X$ …… asexual reproduction

viruses, bacteria, some higher organisms (eukaryotes)

$A + X + Y \rightarrow X + Y + Z \in \{\mathbb{X}, \mathbb{Y}\}$ …… sexual reproduction

most higher organisms (eukaryotes) …… obligatory with mammals

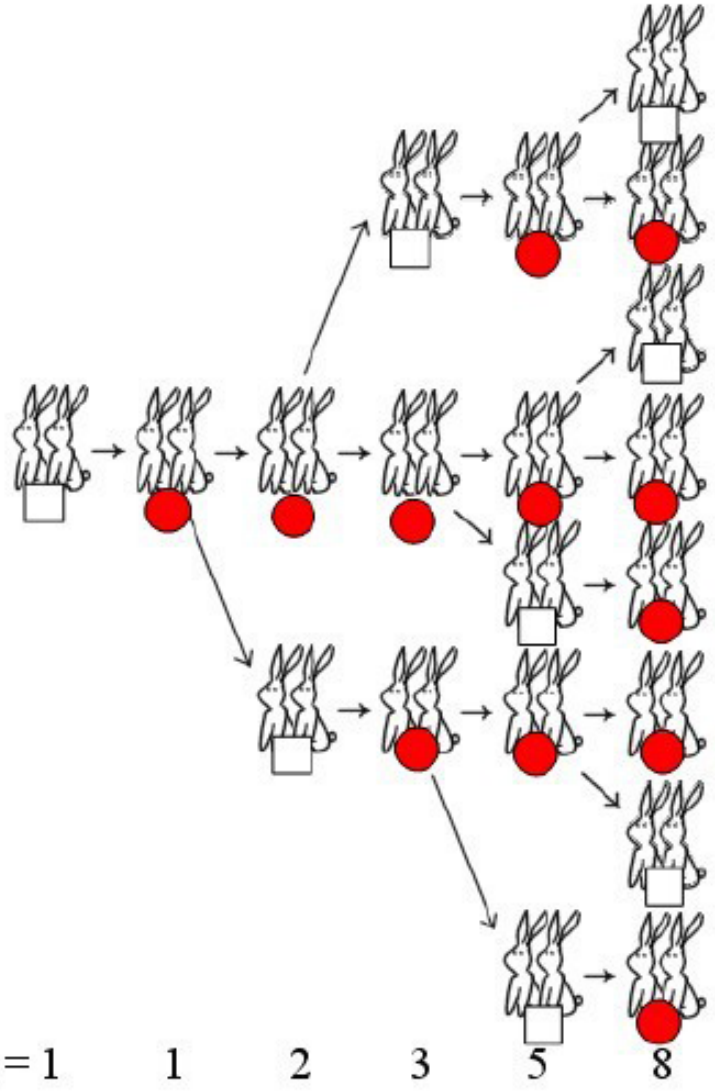
Two modes of reproduction in biology

Leonardo da Pisa
„Fibonacci“

~1180 – ~1240

$$F_{n+1} = F_n + F_{n-1}$$

$$F_0 = 0,\ F_1 = 1$$



The history of exponential growth

Thomas Robert Malthus
1766 – 1834

Leonhard Euler, 1717 - 1783

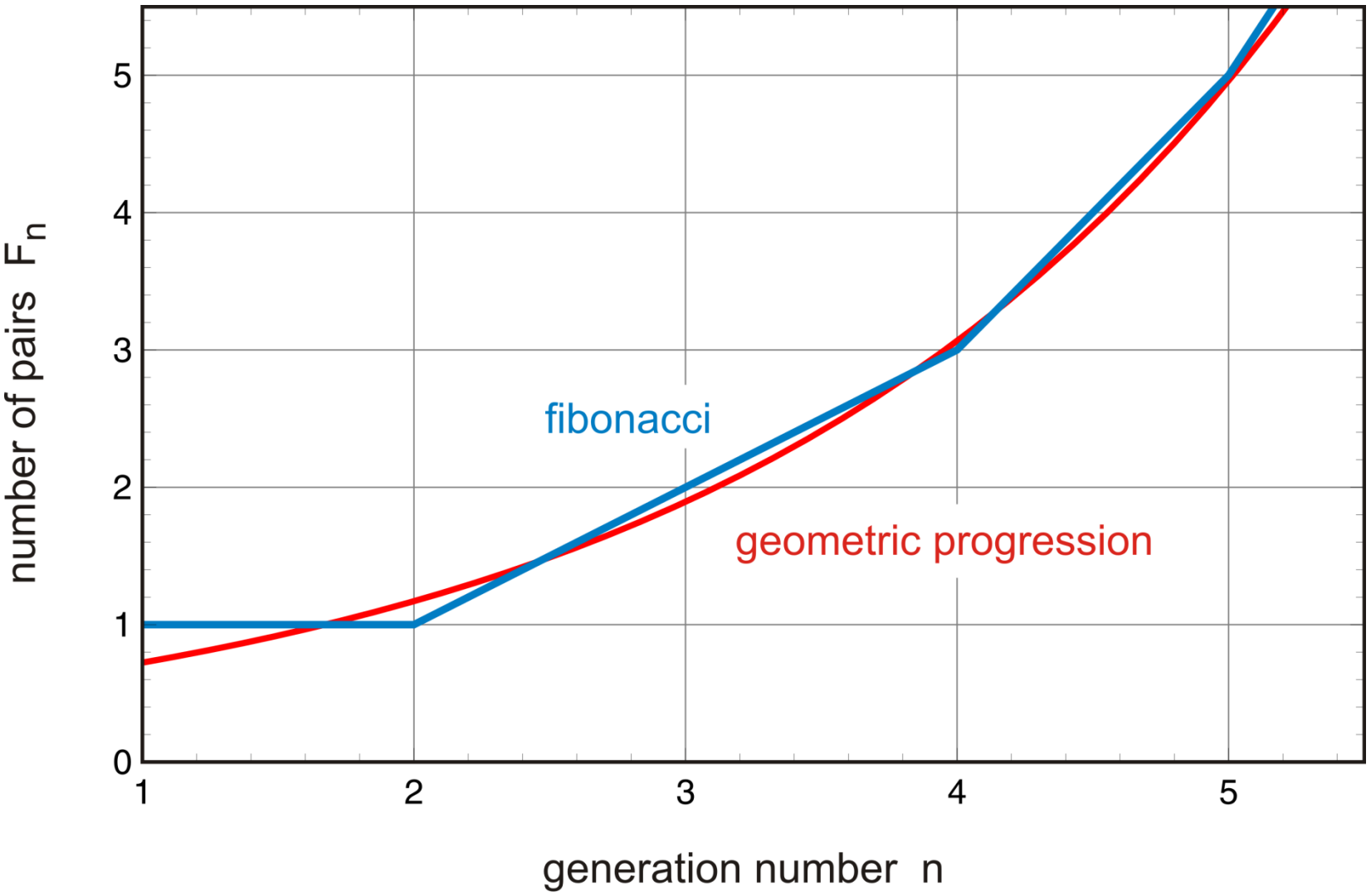
1, 2 , 4 , 8 ,16 , 32 , 64, 128 , ...

geometric progression

$$\exp(x) \equiv \lim_{n\to\infty} (1+\frac{x}{n})^n$$

exponential function

The history of exponential growth



The history of exponential growth

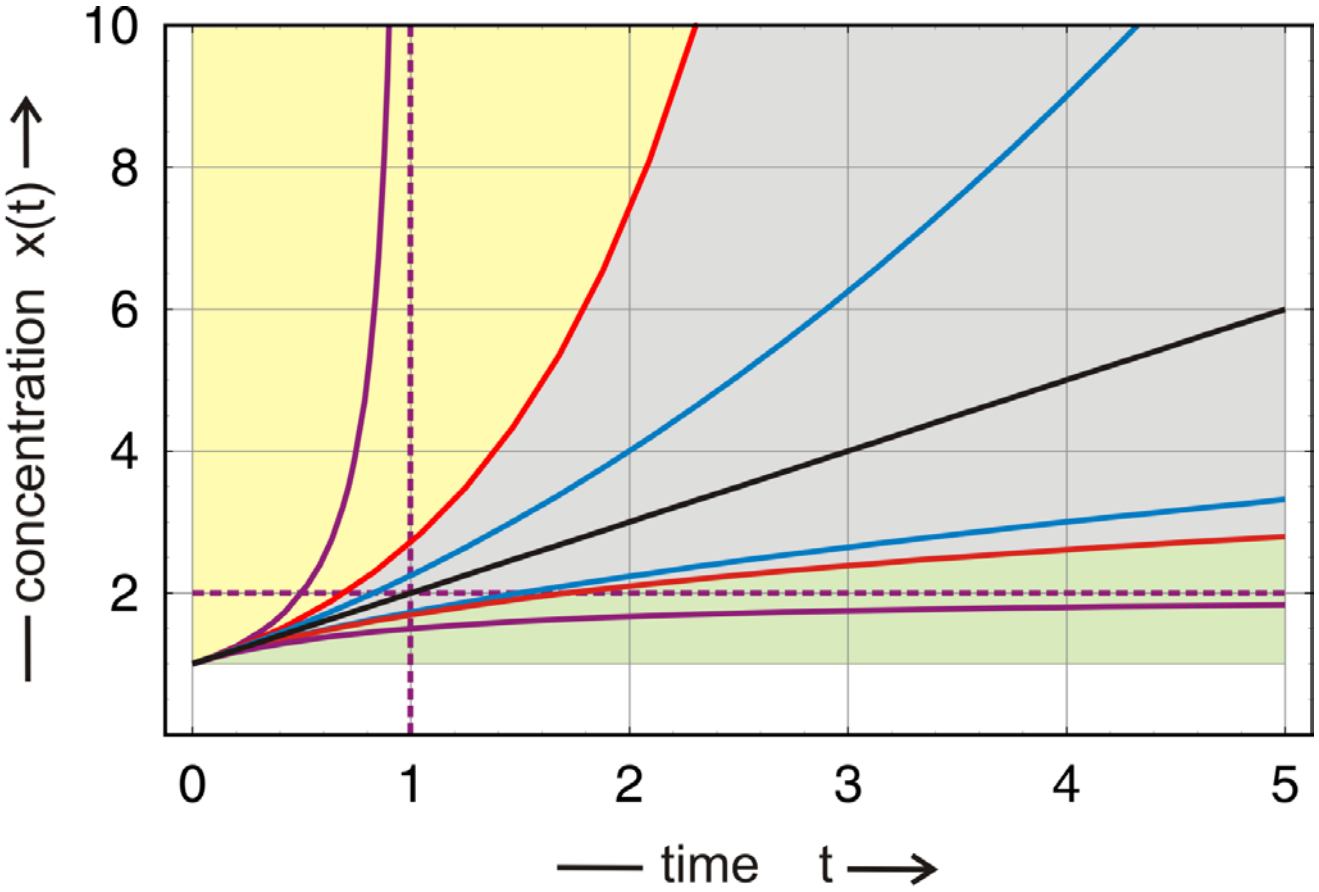
$A + X \rightarrow 2X$ …… asexual reproduction

$[A] = a \; ; \quad [X] = x$

$$\frac{dx}{dt} = k\,a\,x \quad \Rightarrow \quad [A] = a_0 = \text{const} \quad \Rightarrow \quad k\,a_0 = f \; \ldots \text{fitness}$$

$$\frac{dx}{dt} = f\,x \quad \Rightarrow \quad x(t) = x(0)\exp(f\,t) \ldots \text{exponential growth}$$

Reproduction and exponential growth



$$\frac{dx}{dt} = k_\alpha \; x^\alpha \quad \text{normalized to} \quad x(0)=1 \;\text{ and }\; \left.\frac{dx}{dt}\right|_{t=0} = 1$$

Comparison of curves for unlimited growth

1. Prologue
2. **Mathematics of Darwin's natural selection**
3. Mendel, Fisher and population genetics
4. Mutations and selection
5. What means neutrality in evolution?
6. Evolution in simple systems
7. Some origins of complexity in biology

Three necessary conditions for Darwinian evolution are:

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

Darwin discovered the principle of **natural selection** from empirical observations in nature.

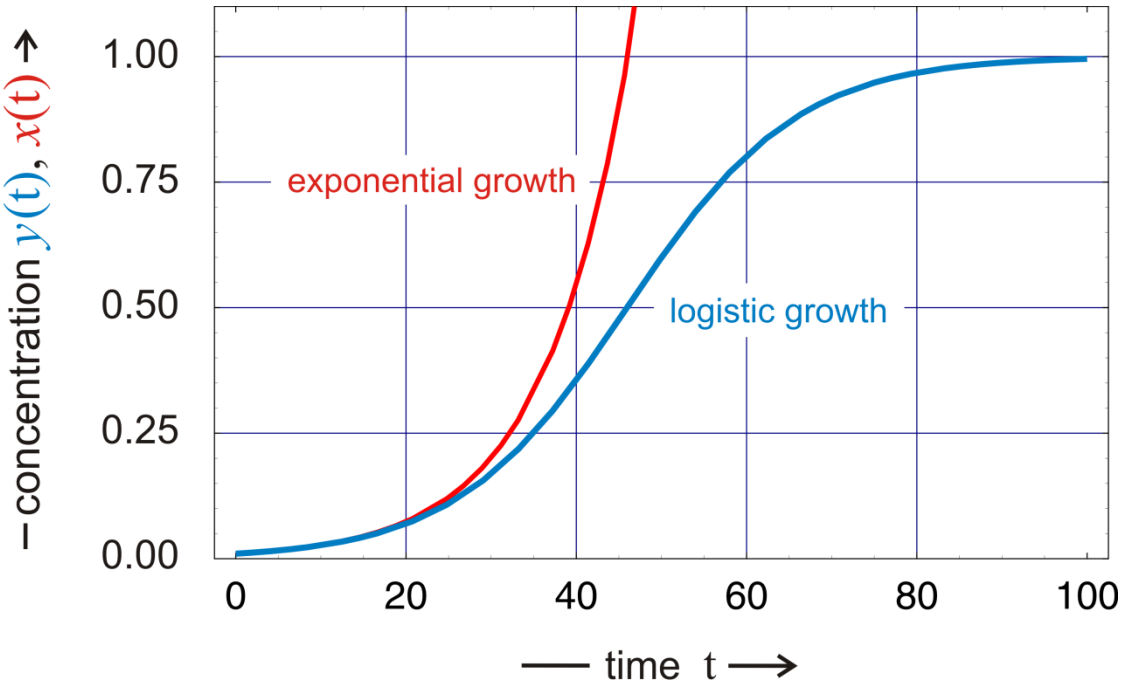
No attempt has been made to cast the principle into theorems.

Pierre-François Verhulst,
1804-1849

$$\frac{dy}{dt} = f\,y\left(1-\frac{y}{C}\right), \quad y(t) = \frac{y(0)\,C}{y(0)+\left(C-y(0)\right)e^{-f\,t}}$$

C …… carrying capacity of the ecosystem

Was known 30 years
before the
'Origin of Species'



The logistic equation, 1828

$$\frac{\mathrm{d}x}{\mathrm{dt}} = f\ x\left(1-\frac{x}{C}\right) \quad \Rightarrow \quad \frac{\mathrm{d}x}{\mathrm{dt}} = f\ x - \frac{x}{C}\ f\ x$$

$$f\ x \equiv \Phi(\mathrm{t}), C = 1: \quad \frac{\mathrm{d}x}{\mathrm{dt}} = x\left(f - \Phi\right)$$

Generalization of the logistic equation to *n* variables yields selection

$$\mathrm{A} + \mathrm{X}_i \rightarrow 2\,\mathrm{X}_i; \quad i = 1, 2, \ldots, \mathrm{n}$$

$$\mathrm{X}_1, \mathrm{X}_2, \ldots, \mathrm{X}_n: \quad [\mathrm{X}_i] = x_i; \quad \sum_{i=1}^{n} x_i = C = 1; \quad f_i = f(\mathrm{X}_i)$$

$$\frac{\mathrm{d}x_j}{\mathrm{dt}} = x_j \left(f_j - \sum_{i=1}^{n} f_i\, x_i \right) = x_j \left(f_j - \Phi \right) \; ; \quad \Phi = \sum_{i=1}^{n} f_i\, x_i$$

Darwin

$$\frac{\mathrm{d}\Phi}{\mathrm{dt}} = < f^2 > - < \bar{f} >^2 = \mathrm{var}\{f\} \geq 0$$

Generalization of the logistic equation to *n* variables yields selection

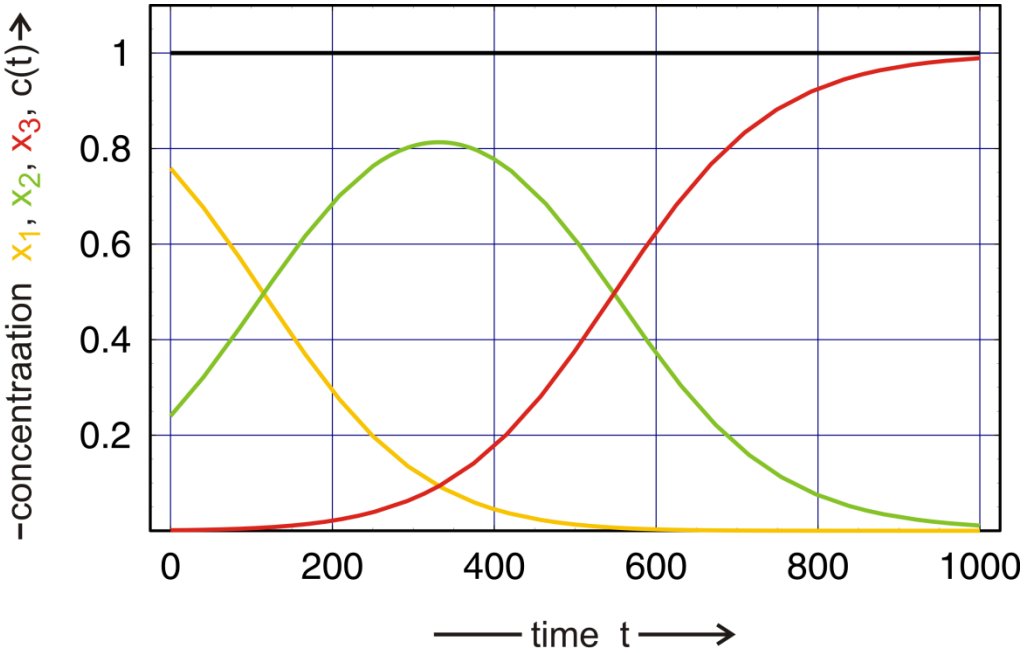
$$\lim_{t\to\infty} \Phi(t) = \Phi_{\max} = f_m$$

$$f_m = \max\{f_1, \ldots f_n\}$$

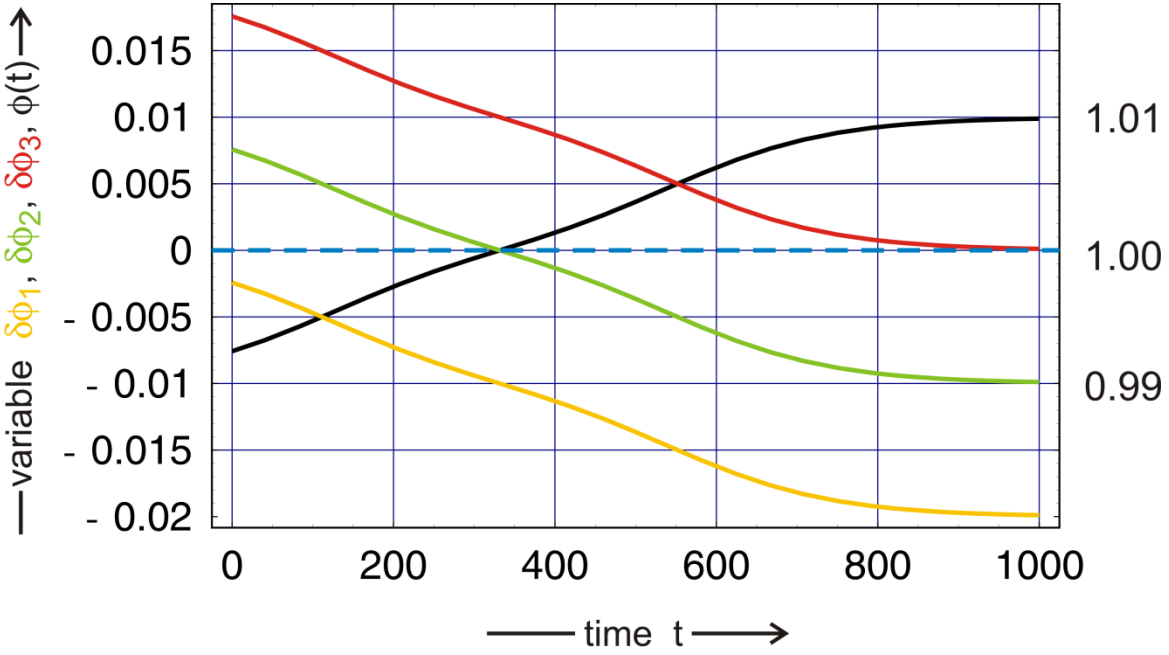
$$\lim_{t\to\infty} x_m = 1$$

$$\lim_{t\to\infty} x_{j,j\neq m} = 0$$

$f_1 = 0.99$, $f_2 = 1.00$, $f_3 = 1.01$



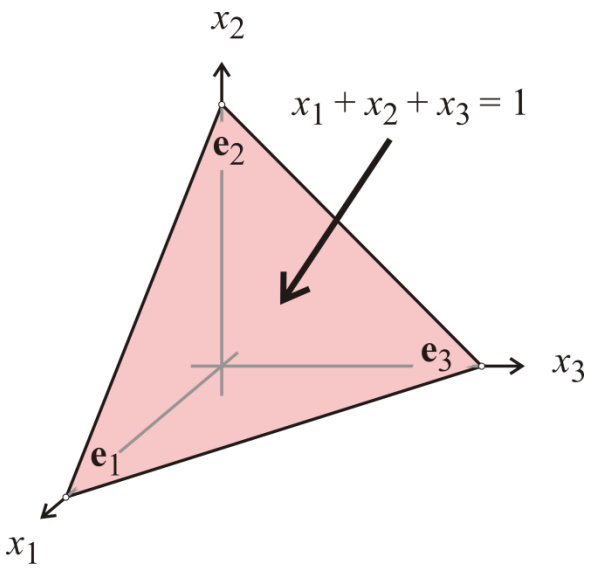
$$\delta\phi_j(t) = f_j - \Phi(t) = f_j - \bar{f}$$

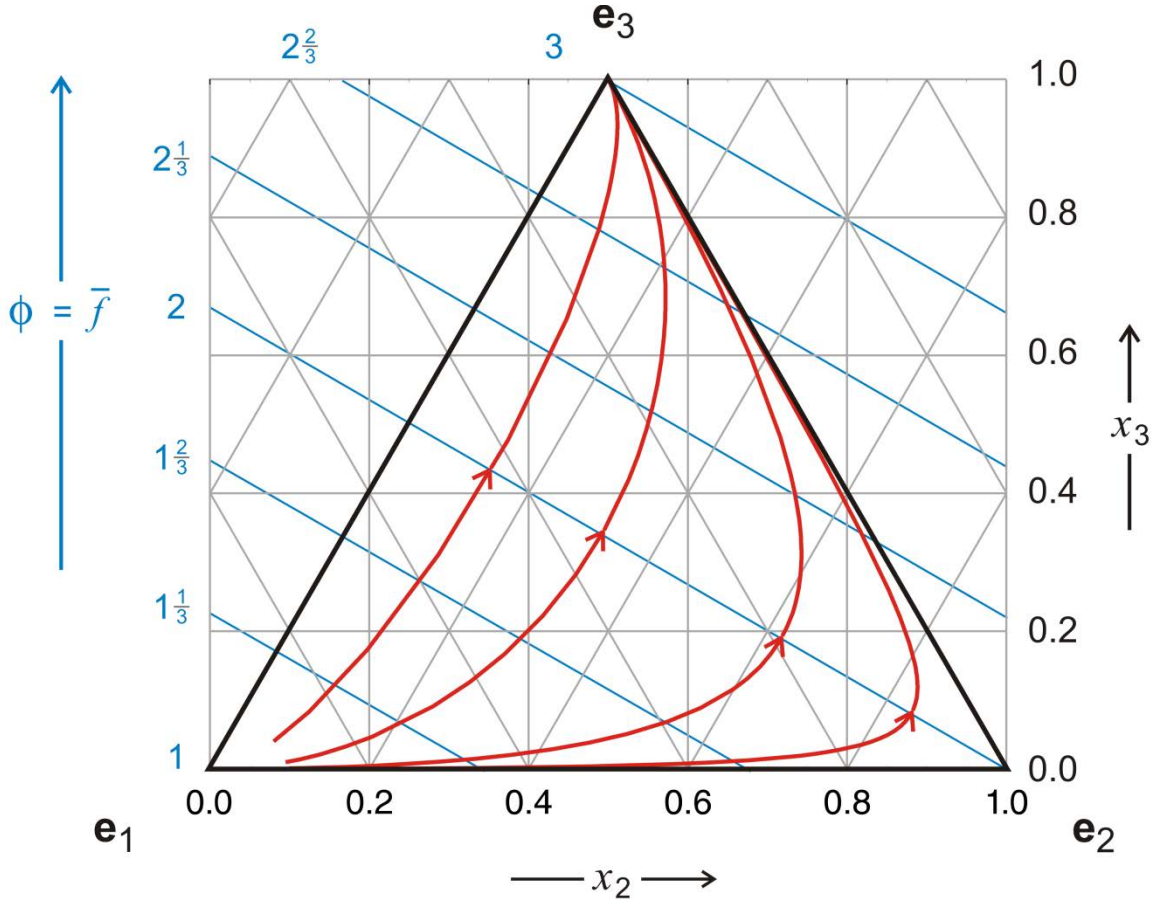


Evolutionary optimization

$$\frac{dx_k}{dt} = (f_k - \Phi)\, x_k \,;\, k = 1,2,3$$

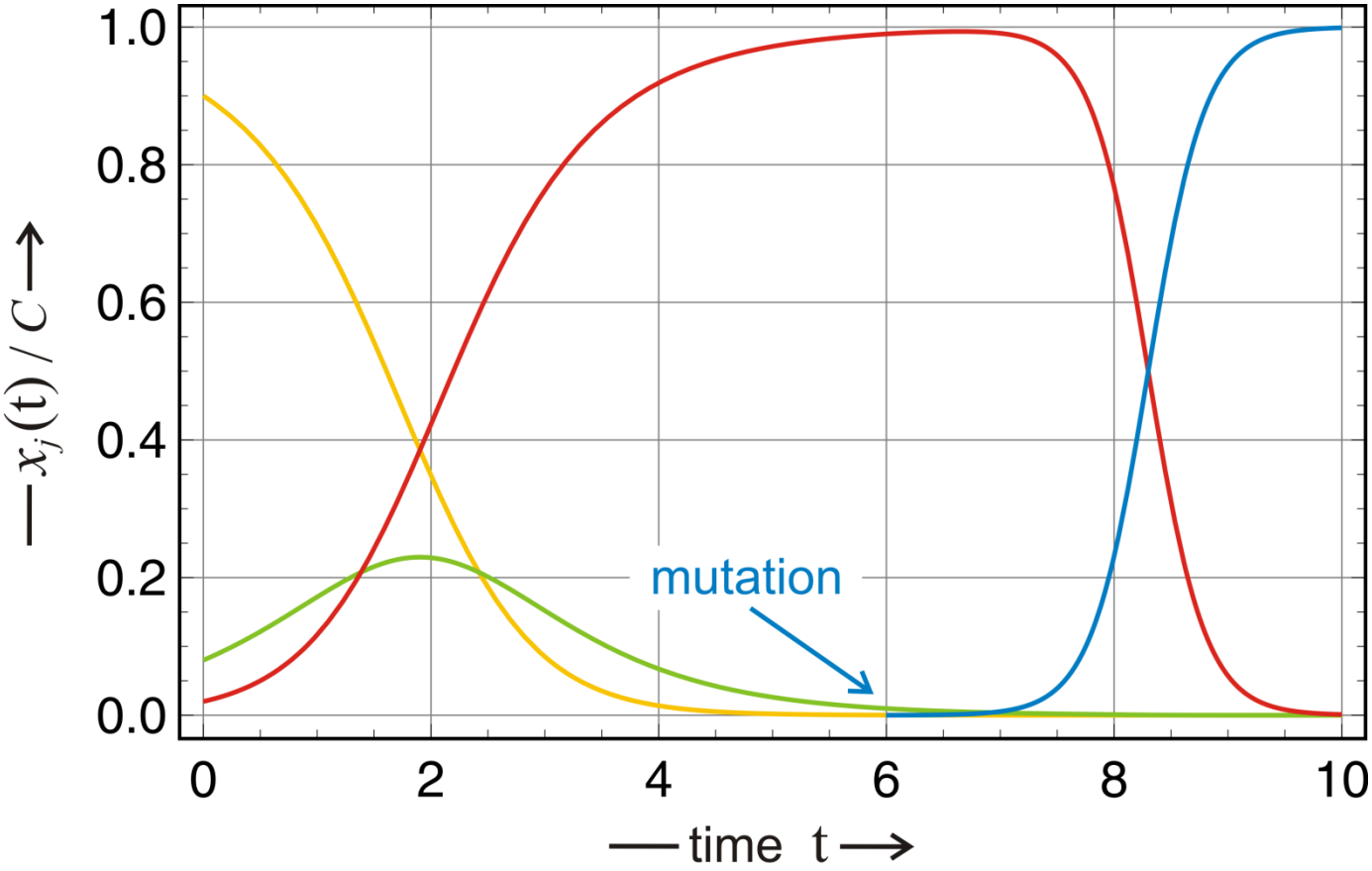
$$\Phi = f_1 x_1 + f_2 x_2 + f_3 x_3 = \bar{f}$$





simplex $\mathbf{S}_3$ is an invariant set $\Rightarrow$ $\mathbf{e}_3$ …… “corner equilibrium”

Phase diagram of Darwinian selection

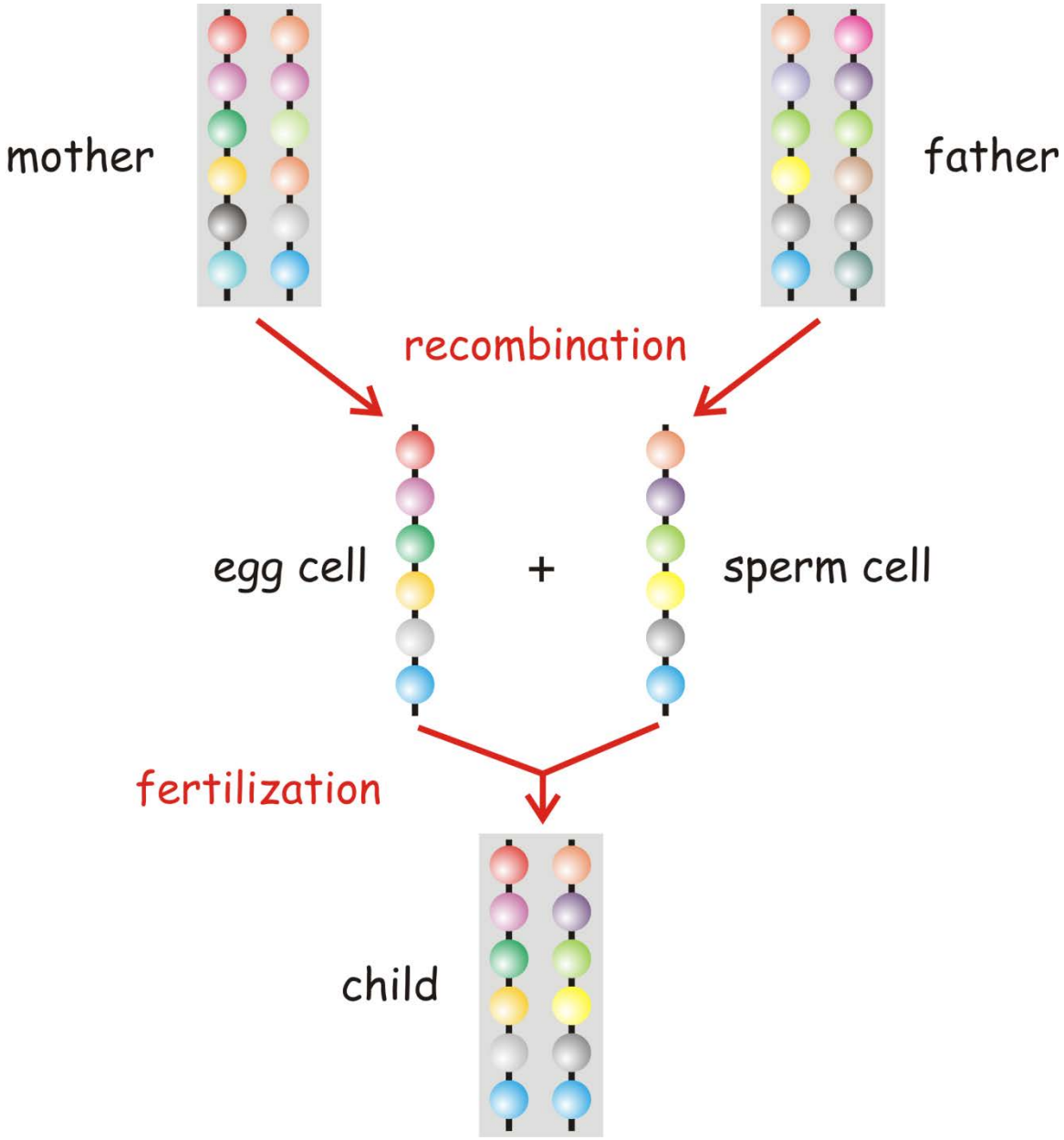


$f_1 = 1$, $f_2 = 2$, $f_3 = 3$, $f_4 = 7$

Before the development of molecular biology mutation was treated as a "deus ex machina"

1. Prologue
2. Mathematics of Darwin's natural selection
3. **Mendel, Fisher and population genetics**
4. Mutations and selection
5. What means neutrality in evolution?
6. Evolution in simple systems
7. Some origins of complexity in biology

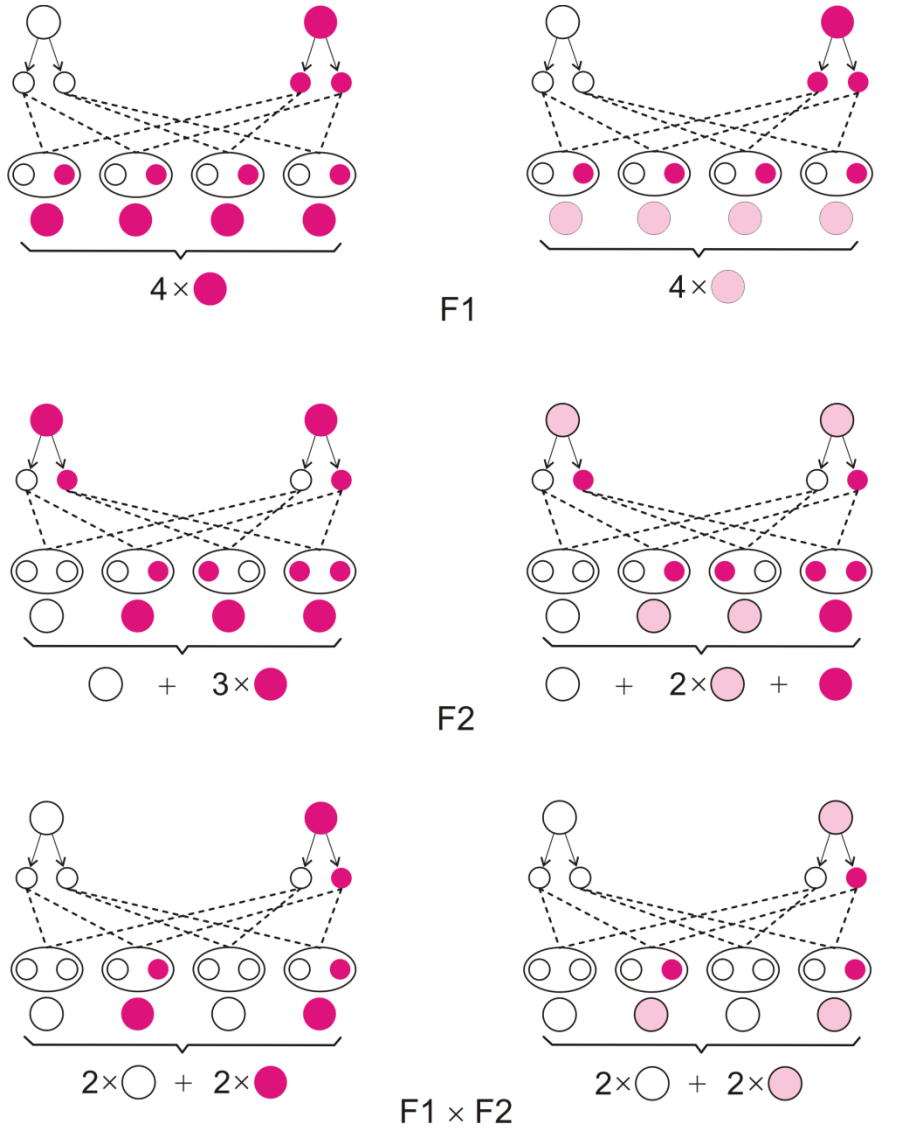
Gregor Mendel
1822 - 1884



Recombination in Mendelian genetics

Mendelian genetics

The 1:3 rule



dominant/recessive pair of alleles

dominance

intermediate pair of alleles

semi-dominance

Char.	Parental phenotype	F1	F2	F2 ratio
1	round × wrinkled seeds	all round	5174 / 1859	2.96
2	yellow × green seeds	all yellow	6022 / 2001	3.01
3	purple × white petals	all purple	705 / 244	3.15
4	inflated × pinched pods	all inflated	882 / 299	2.95
5	green × yellow pods	all green	428 / 152	2.82
6	axial × terminal flowers	all axial	651 / 207	3.14
7	long × short stems	all axial	787 / 277	2.84

The results of the individual experiments Gregor Mendel did with the garden pea *pisum sativum*.

Ronald Fisher (1890-1962)

alleles: $A_1, A_2, \ldots, A_n$

frequencies: $x_i = [A_i]$; genotypes: $A_i \cdot A_j$

fitness values: $a_{ij} = f(A_i \cdot A_j)$, $a_{ij} = a_{ji}$

Mendel

Darwin

$$\frac{dx_j}{dt} = \sum_{i=1}^{n} a_{ji}\, x_i\, x_j - \Phi\, x_j = x_j \left(\sum_{i=1}^{n} a_{ji}\, x_i - \Phi \right), \quad j = 1, 2, \ldots, n$$

$$\text{mit} \quad \Phi(t) = \sum_{j=1}^{n} \sum_{i=1}^{n} a_{ji}\, x_i\, x_j \quad \text{und} \quad \sum_{j=1}^{n} x_j = 1$$

$$\frac{d\Phi}{dt} = 2\left(< \bar{a}^2 > - < \bar{a} >^2 \right) = 2\, \mathrm{var}\{\bar{a}\} \geq 0$$

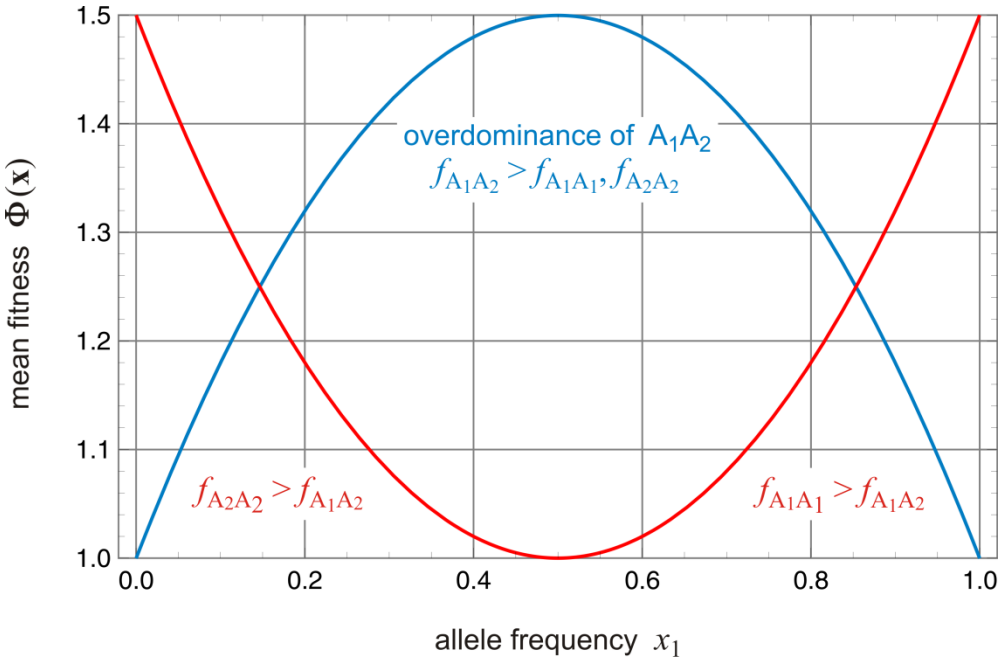
Ronald Fisher‘s selection equation: The genetical theory of natural selection. Oxford, UK, Clarendon Press, 1930.

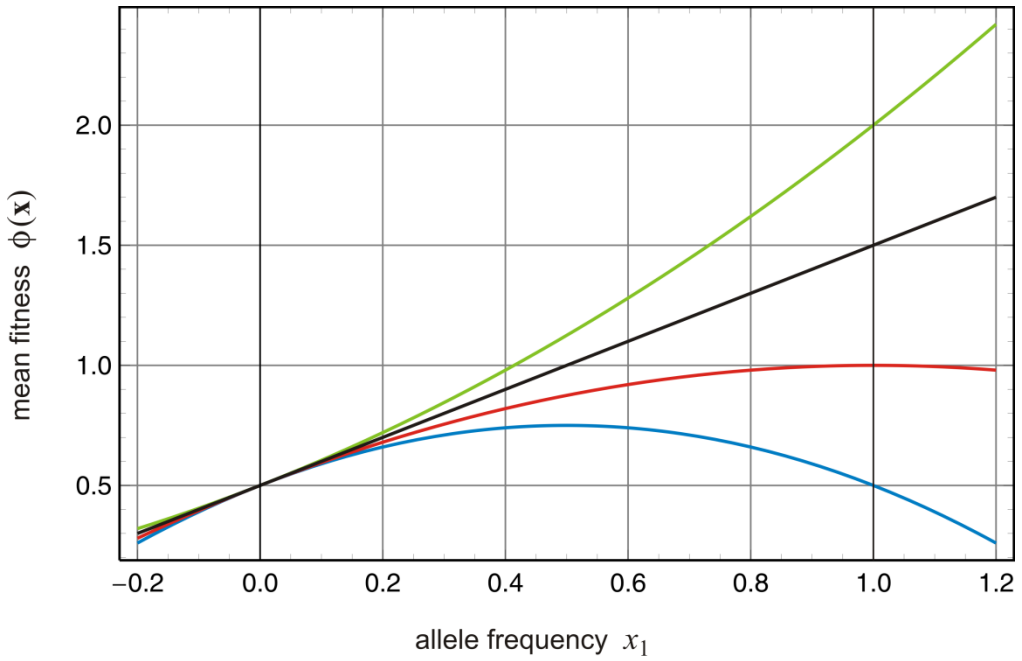
mean fitness

$$\Phi(\mathbf{x}) = a_{11}\, x_1^2 + 2\, a_{12}\, x_1 x_2 + a_{22}\, x_2^2$$

$$a_{ij} = f_{\mathrm{X}_i \mathrm{X}_j}$$

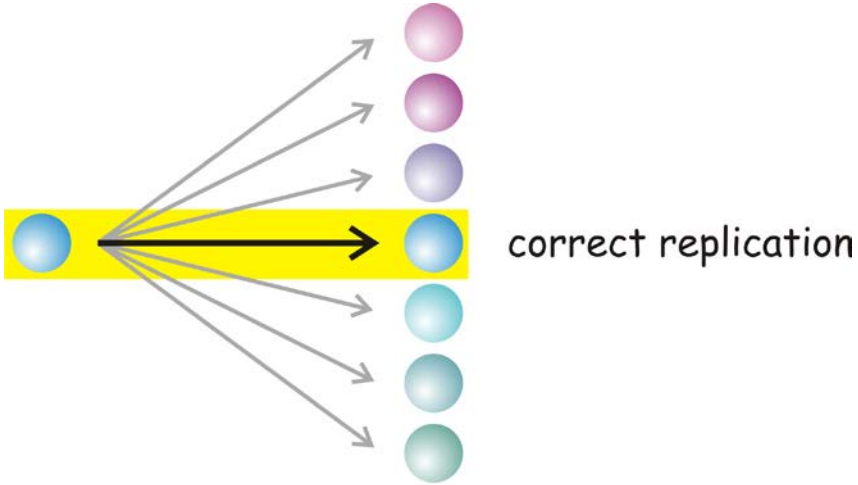
Fitness in Fisher's selection equation





1. Prologue
2. Mathematics of Darwin's natural selection
3. Mendel, Fisher and population genetics
4. **Mutations and selection**
5. What means neutrality in evolution?
6. Evolution in simple systems
7. Some origins of complexity in biology

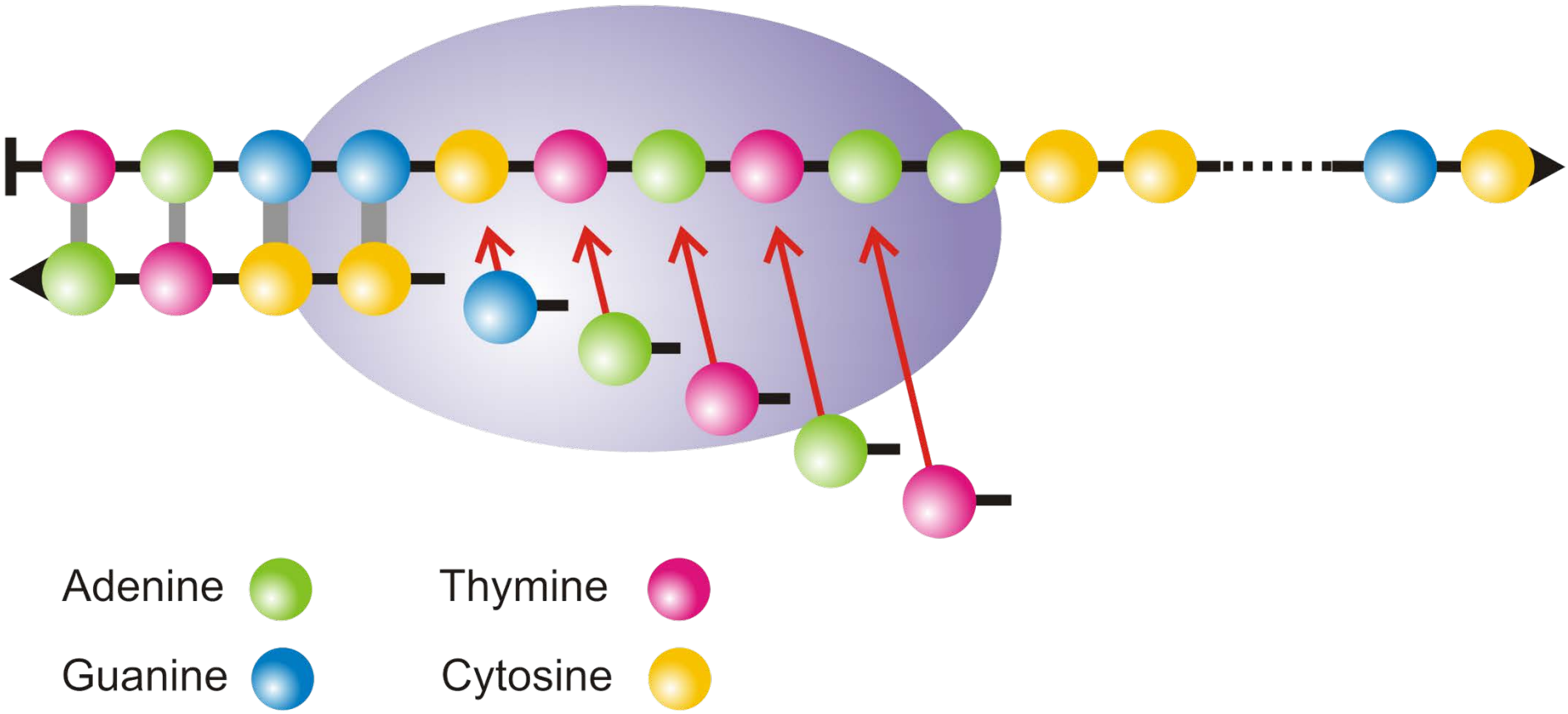
Hermann J. Muller
1890 - 1967



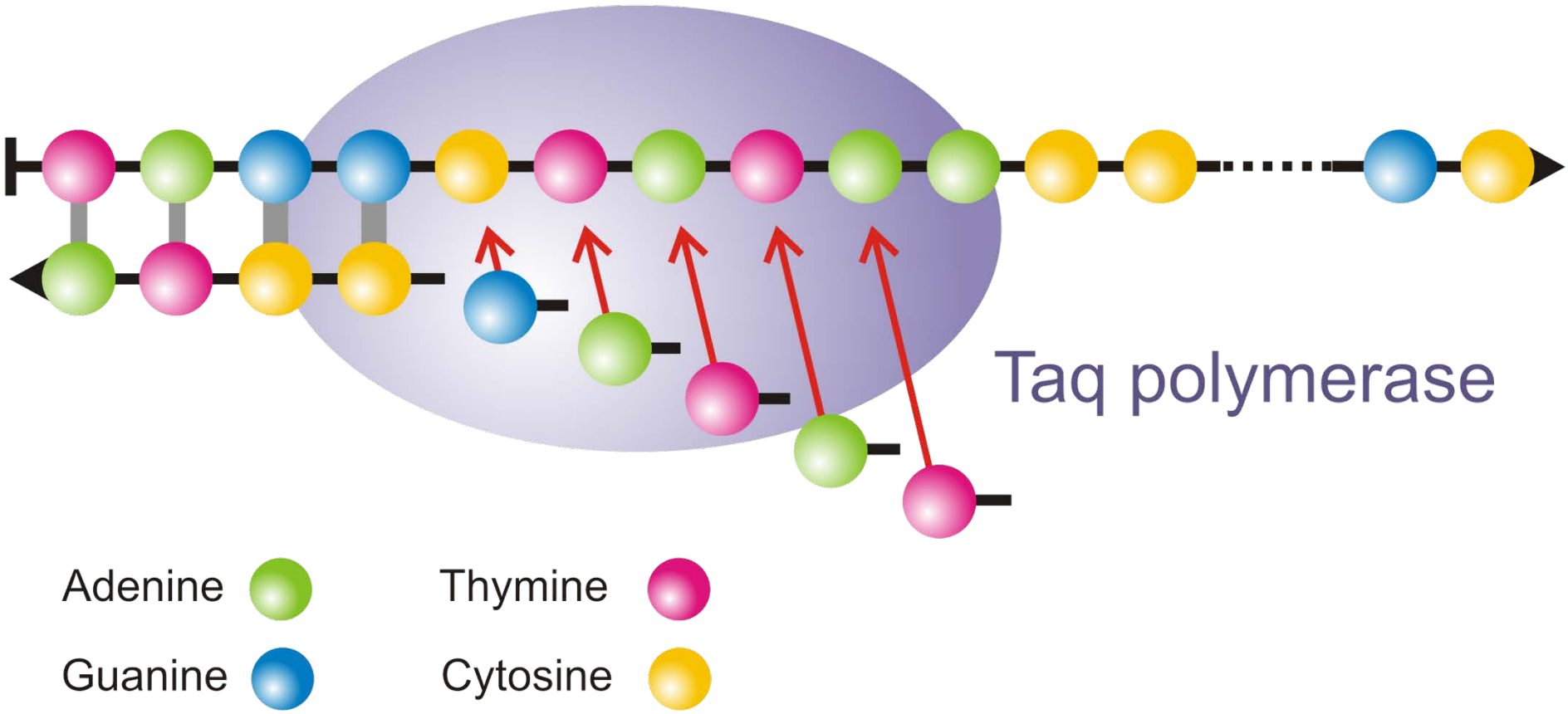
Thomas H. Morgan
1866 - 1945

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

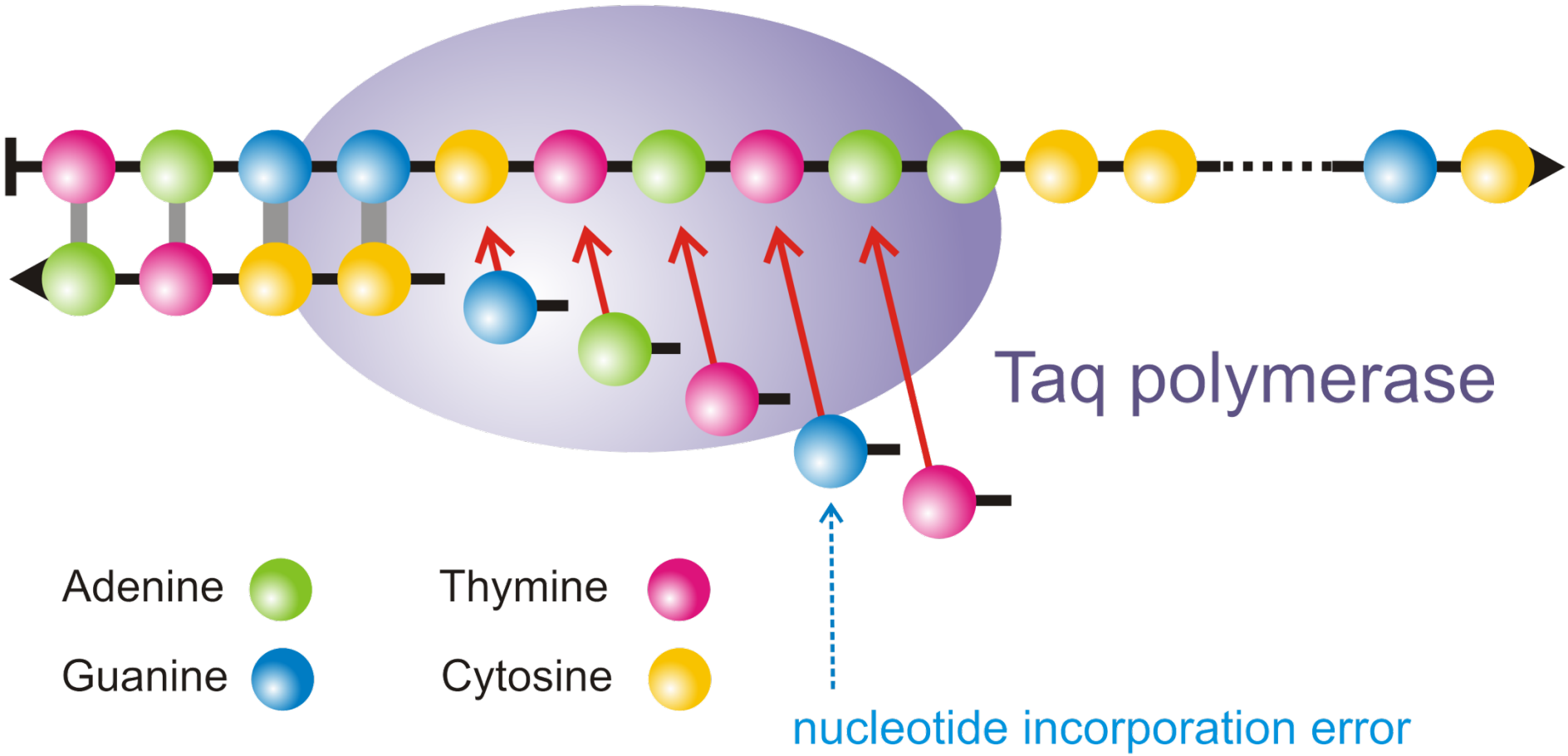
John W. Drake, Brian Charlesworth, Deborah Charlesworth and James F. Crow. 1998. Rates of spontaneous mutation. *Genetics* 148:1667-1686.



The logic of DNA (or RNA) replication and mutation



The logic of DNA (or RNA) replication and mutation

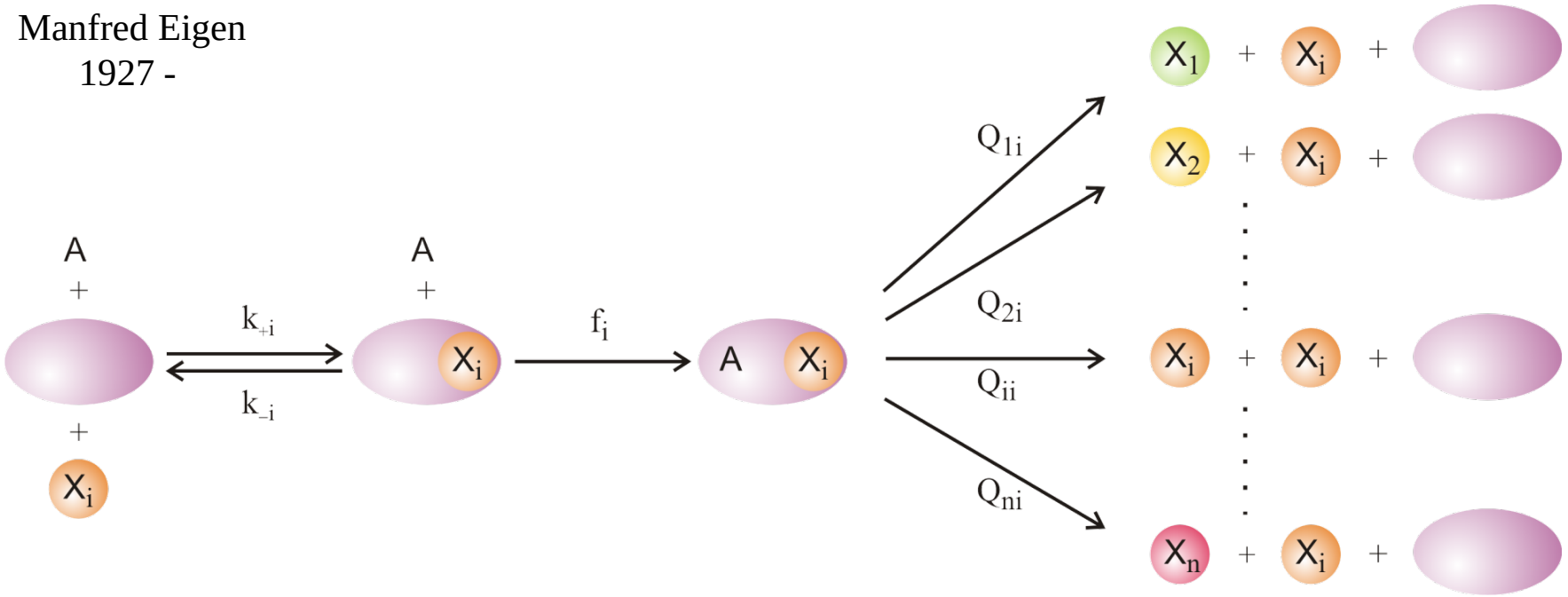


The logic of DNA (or RNA) replication and mutation

Manfred Eigen
1927 -

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

$$W_{ji} = Q_{ji} \cdot f_i\,, \quad \sum_{i=1}^{n} x_i = 1, \quad \Phi = \sum_{i=1}^{n} f_i\, x_i$$



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

Mutation-selection equation: $[\mathrm{I}_i] = x_i \geq 0,\ f_i > 0,\ Q_{ij} \geq 0$

$$\frac{dx_i}{dt} = \sum_{j=1}^{n} f_j Q_{ji}\, x_j - x_i \Phi, \quad i = 1,2,\cdots,n; \quad \sum_{i=1}^{n} x_i = 1; \quad \Phi = \sum_{j=1}^{n} f_j x_j = \overline{f}$$

Solutions are obtained after integrating factor transformation by means of an eigenvalue problem

$$x_i(t) = \frac{\sum_{k=0}^{n-1} b_{ik} \cdot c_k(0) \cdot \exp(\lambda_k t)}{\sum_{j=1}^{n} \sum_{k=0}^{n-1} b_{jk} \cdot c_k(0) \cdot \exp(\lambda_k t)}; \quad i = 1,2,\cdots,n; \quad c_k(0) = \sum_{i=1}^{n} h_{ki}\, x_i(0)$$

$$W \div \left\{ f_i Q_{ij};\ i, j = 1,\cdots,n \right\};\ B = \left\{ b_{ij};\ i, j = 1,\cdots,n \right\};\ B^{-1} = H = \left\{ h_{ij};\ i, j = 1,\cdots,n \right\}$$

$$B^{-1} \cdot W \cdot B = \Lambda = \left\{ \lambda_k; k = 0,1,\cdots,n-1 \right\}$$

$$\frac{d\mathbf{b}_k}{dt} = \mathbf{b}_k (\lambda_k - \Phi);\ k = 0, 1, \ldots, n-1$$

Perron – Frobenius theorem: $\lambda_0 > \lambda_1 \geq \lambda_2 \geq \lambda_3 \cdots$

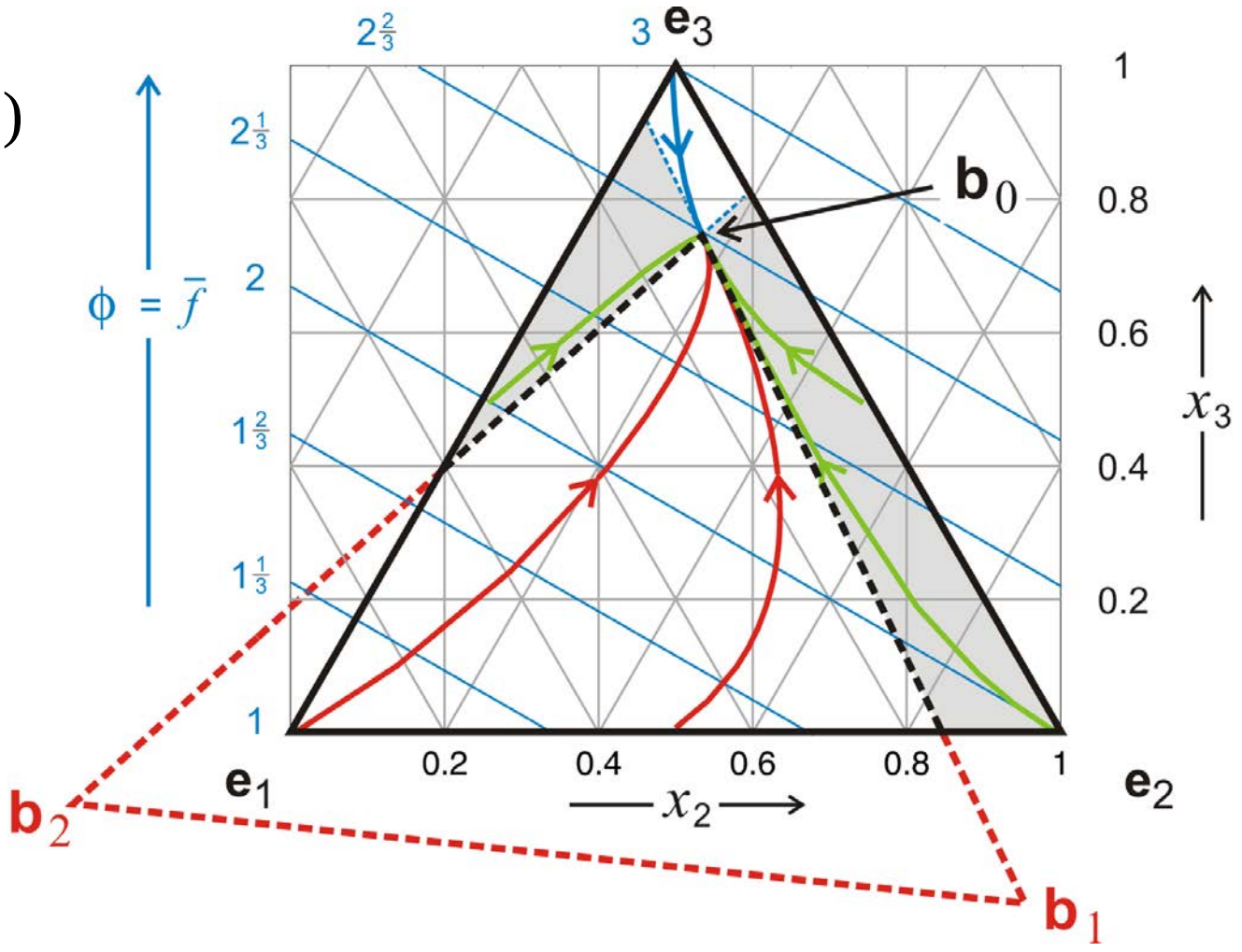
$$\lambda_0 \quad \Leftrightarrow \quad \mathbf{b}_0 = \begin{pmatrix} b_{10} \\ b_{20} \\ \vdots \\ b_{n0} \end{pmatrix}$$

$$\lim_{t\to\infty} x_i(t) = \overline{x}_i = \frac{b_{i0}\, c_0(0)}{\sum_{j=1}^{n} b_{j0}\, c_0(0)}; \quad i = 1,2,\cdots,n; \quad c_0(0) = \sum_{i=1}^{n} h_{0i}\, x_i(0)$$

The *quasispecies* is the long-time solution of the mutation-selection equation.

Definition of quasispecies

$$\frac{d\mathbf{b}_0}{dt} = \mathbf{b}_0(\lambda_0 - \Phi)$$



Phase diagram of the mutation-selection system

	Selection	Selection-recombination	Selection-mutation
Method of solution	integrating factor	qualitative analysis	integrating factor, eigenvalue problem
Linearity	yes	no	yes
Optimization of Φ	yes	yes	no
Unique optimum	yes	no	no optimum
Invariance of S_n	yes	yes	no
Uniqueness of solution	yes	no	yes
Selection of	fittest	fittest/coexistence	quasispecies

Comparison of mathematical models of evolution

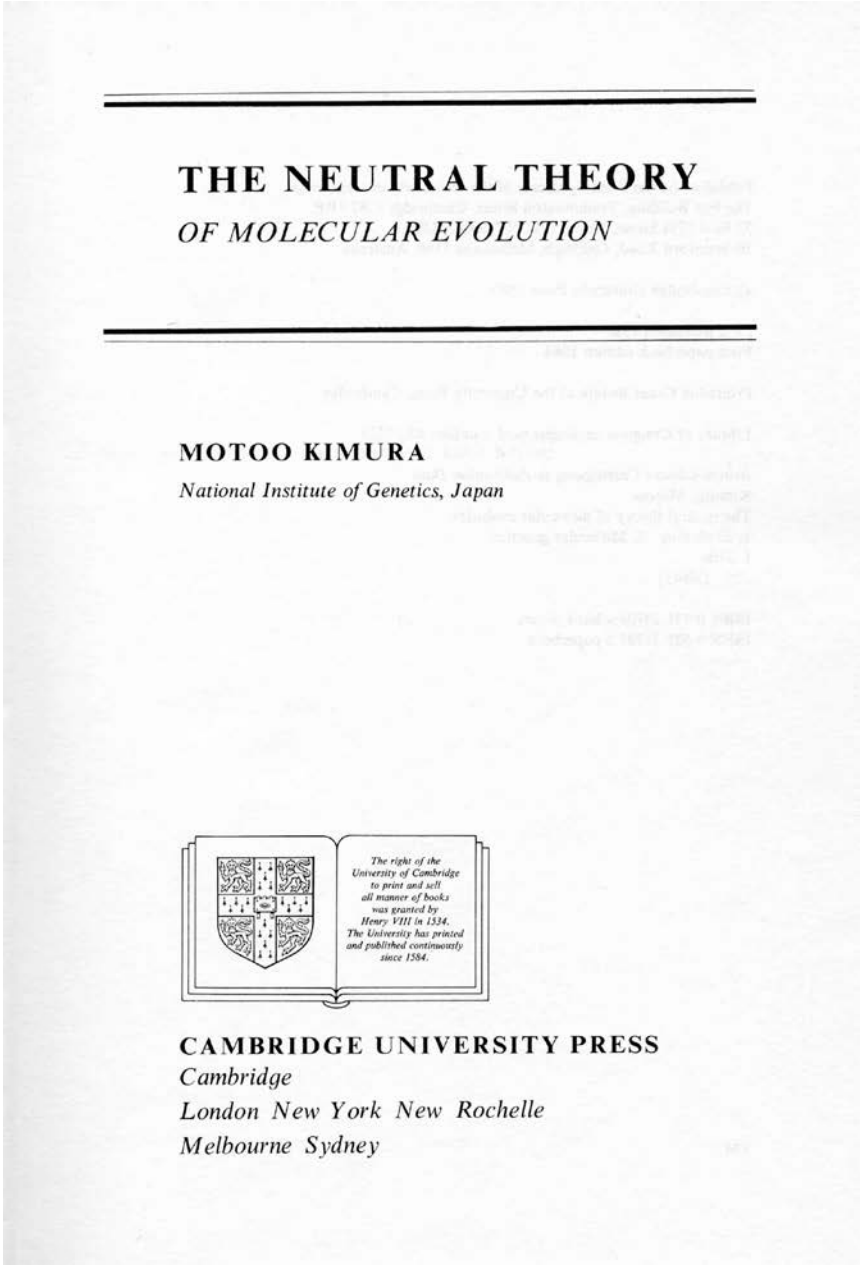
1. Prologue
2. Mathematics of Darwin's natural selection
3. Mendel, Fisher and population genetics
4. Mutations and selection
5. **What means neutrality in evolution?**
6. Evolution in simple systems
7. Some origins of complexity in biology

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

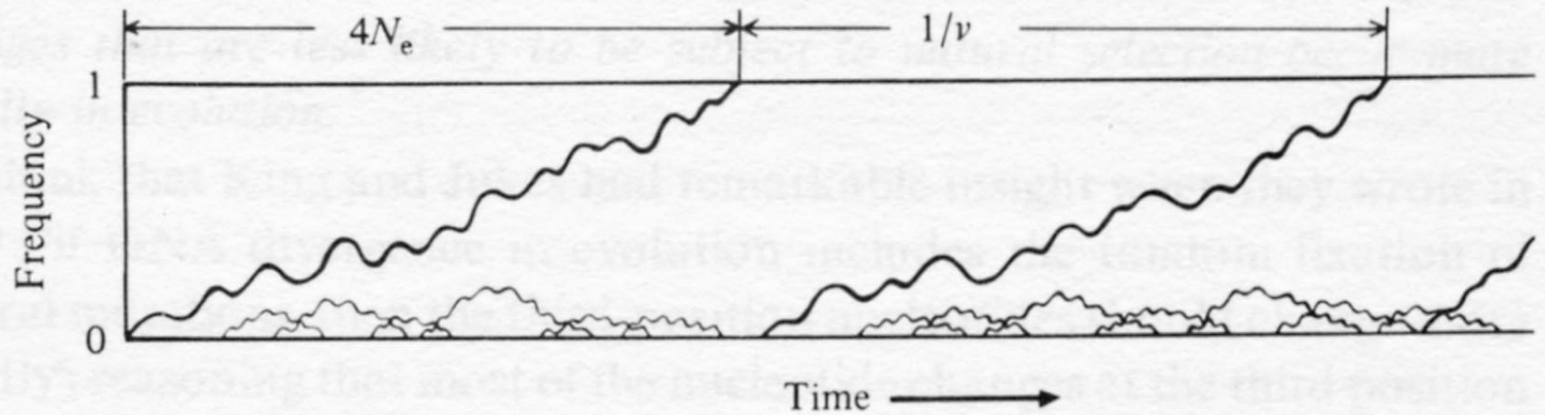
OF MOLECULAR EVOLUTION

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

mean time of fixation: $4\,N_e$ … effective population size

mean time of replacement: ν^{-1} … reciprocal mutation rate

$$\mathrm{A} + \mathrm{X}_k \xrightarrow{\lambda} 2\,\mathrm{X}_k\,; \quad k = 1, \ldots, m$$

$$\mathrm{X}_k \xrightarrow{\lambda} \mathrm{B}\,; \quad k = 1, \ldots, m$$

$$P_{n_1, n_2, \cdots, n_m}(t) = P\left(\mathcal{X}_1(t) = n_1, \mathcal{X}_2(t) = n_2, \cdots, \mathcal{X}_m(t) = n_m\right) =$$

$$= P_{n_1}(t) \cdot P_{n_2}(t) \cdot \ldots \cdot P_{n_m}(t)$$

neutrality: all probabilities are equivalent and the densities identical

$$\frac{dP_n}{dt} = \lambda\left((n-1)\,P_{n-1}(t) + (n+1)P_{n+1}(t) - 2\,n\,P_n(t)\right);\; P_n(t) = P\left(\mathcal{X}(t) = n\right)$$

$$P_n(t) = \left(\frac{\lambda t}{1+\lambda t}\right)^{n_0+n} \sum_{k=0}^{\min(n_0,n)} \binom{n_0 + n - k - 1}{n-k} \binom{n_0}{k} \left(\frac{1-\lambda^2 t^2}{\lambda^2 t^2}\right)^k$$

$$A + X \xrightarrow{\lambda} 2X$$

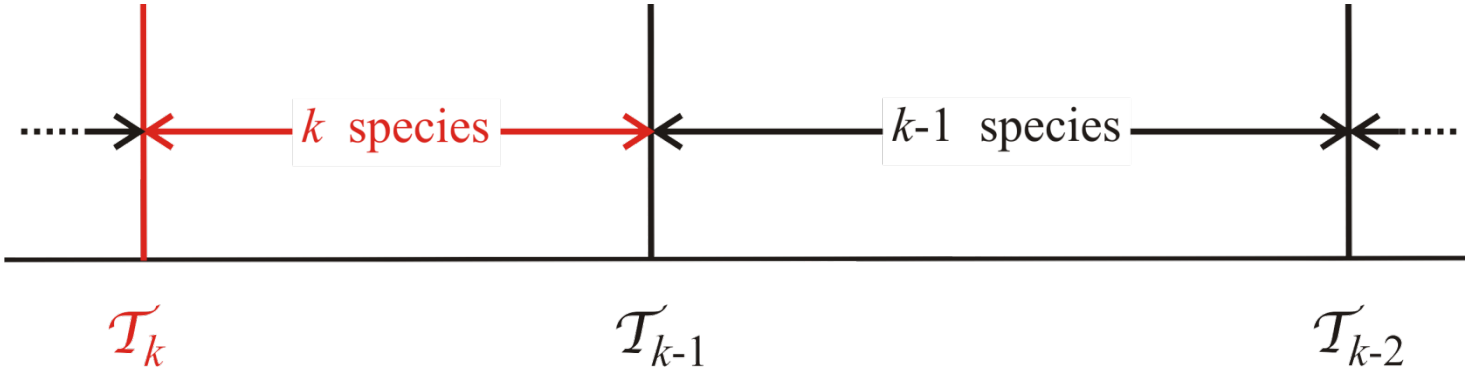
$$X \xrightarrow{\lambda} B$$

$$\frac{dP_n}{dt} = \lambda\big((n-1)P_{n-1}(t) + (n+1)P_{n+1}(t) - 2nP_n(t)\big);\ P_n(t) = P\big(\mathcal{X}(t) = n\big)$$

$$P_n(t) = \left(\frac{\lambda t}{1+\lambda t}\right)^{n_0+n} \sum_{k=0}^{\min(n_0,n)} \binom{n_0+n-k-1}{n-k}\binom{n_0}{k}\left(\frac{1-\lambda^2 t^2}{\lambda^2 t^2}\right)^k$$

$$E\big(\mathcal{X}(t)\big) = n_0, \quad \sigma^2\big(\mathcal{X}(t)\big) = 2n_0\lambda t, \quad P_0(t) = \left(\frac{\lambda t}{1+\lambda t}\right)^{n_0}$$

$$\lim_{t\to\infty} P_0(t) = 1$$



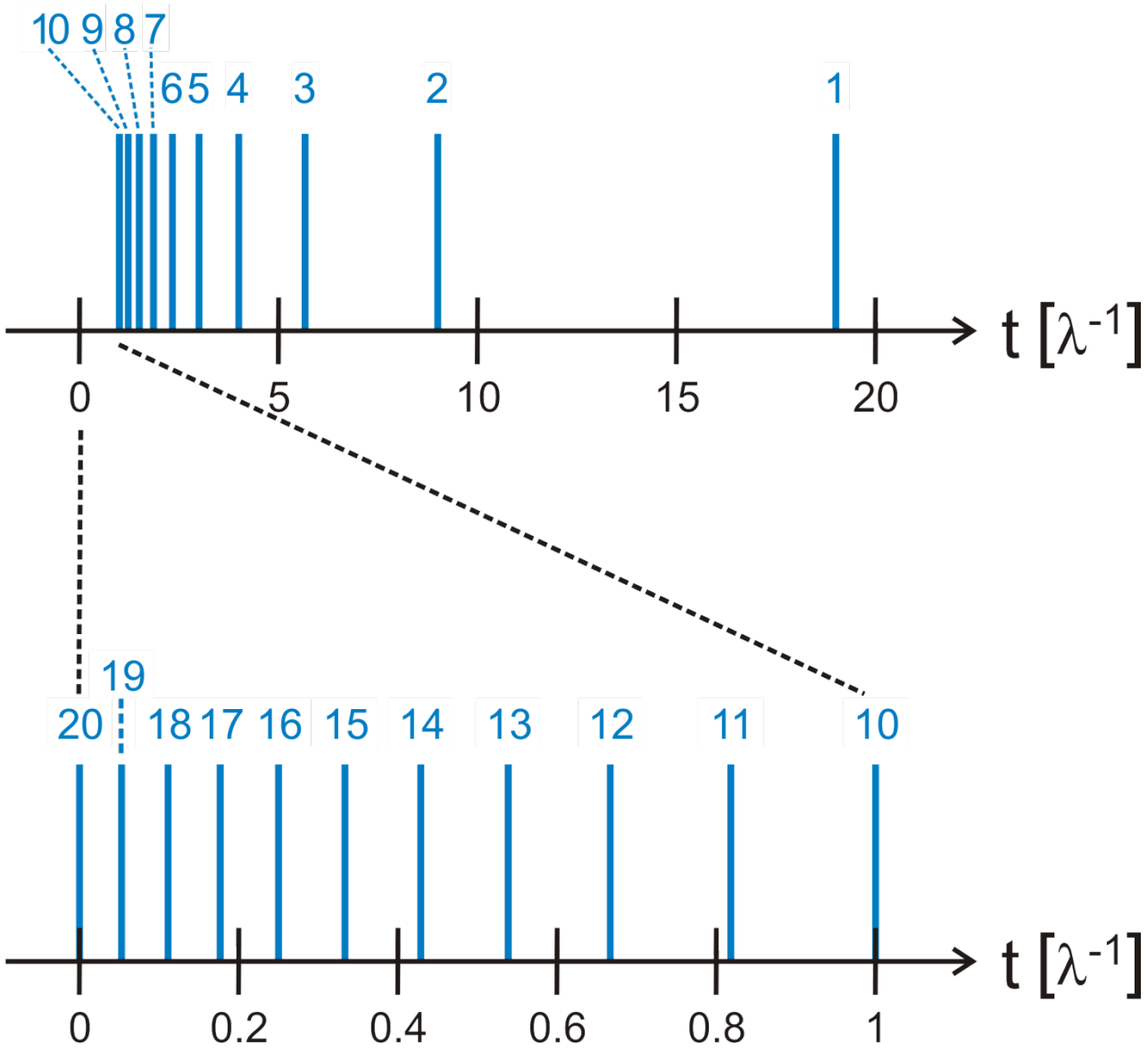
$$H_k(t) = P\left(\mathcal{T}_k < t\right), \quad H_0(t) = P_{0,0,\cdots,0}(t) = \left(\frac{\lambda t}{1+\lambda t}\right)^m$$

$$H_k(t) = \sum_{j=0}^{k} \binom{m}{j} \frac{(\lambda t)^{m-j}}{(1+\lambda t)^m}$$

Sequential extinction times $\mathcal{T}_k$

$$E(\mathcal{T}_k) = \frac{m-k}{k} \cdot \frac{1}{\lambda}$$

$$E(\mathcal{T}_1) = \frac{m-1}{\lambda} \approx \frac{m}{\lambda}$$

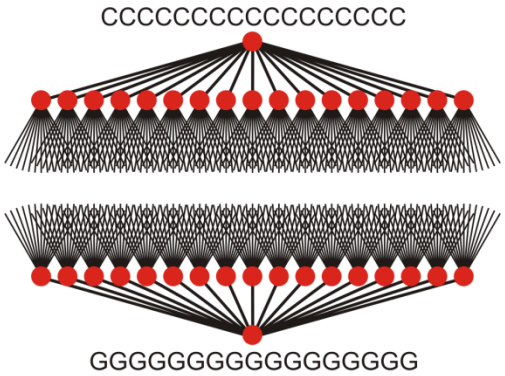


Sequential extinction times $\mathcal{T}_k$

Sewall Wright, 1889 - 1988

Sewall Wright. 1932. *The roles of mutation, inbreeding, crossbreeding and selection in evolution*. In: D.F.Jones, ed. Int. Proceedings of the Sixth International Congress on Genetics. Vol.1, 356-366. Ithaca, NY.

Sewall Wrights fitness landscape as metaphor for Darwinian evolution

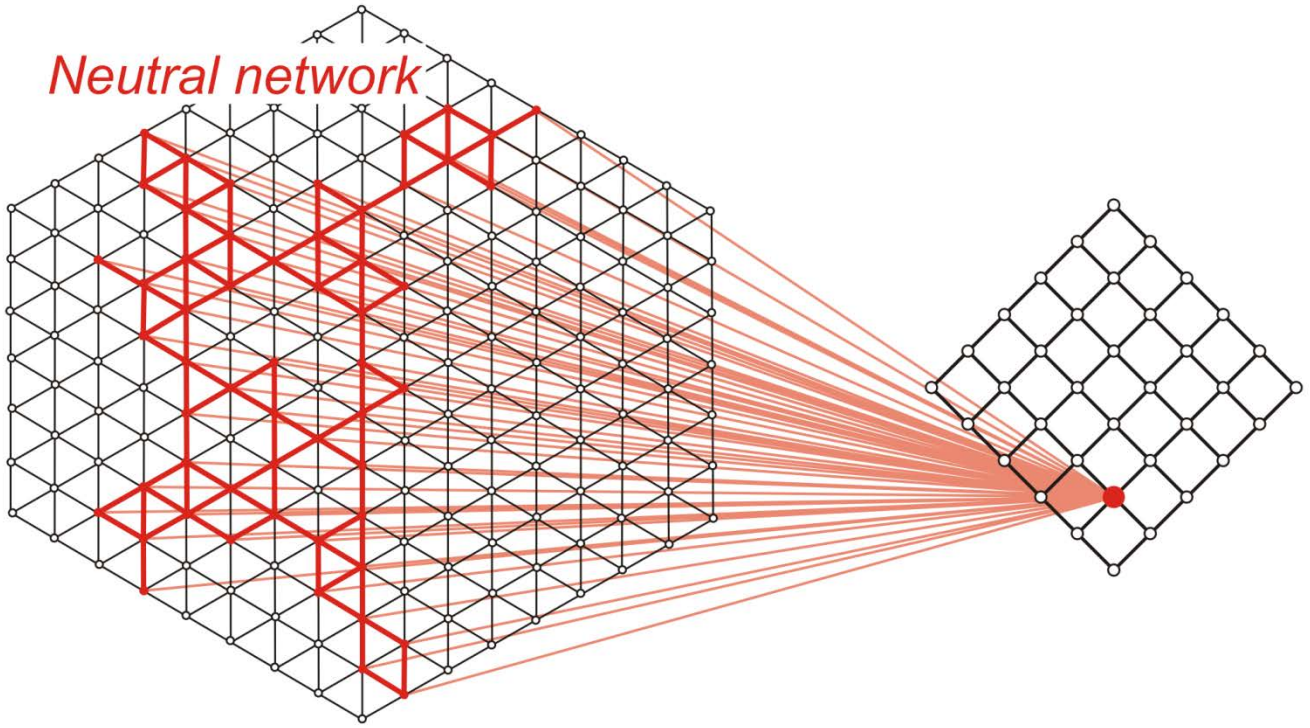


sequence space

Q

genotype → phenotype → selection

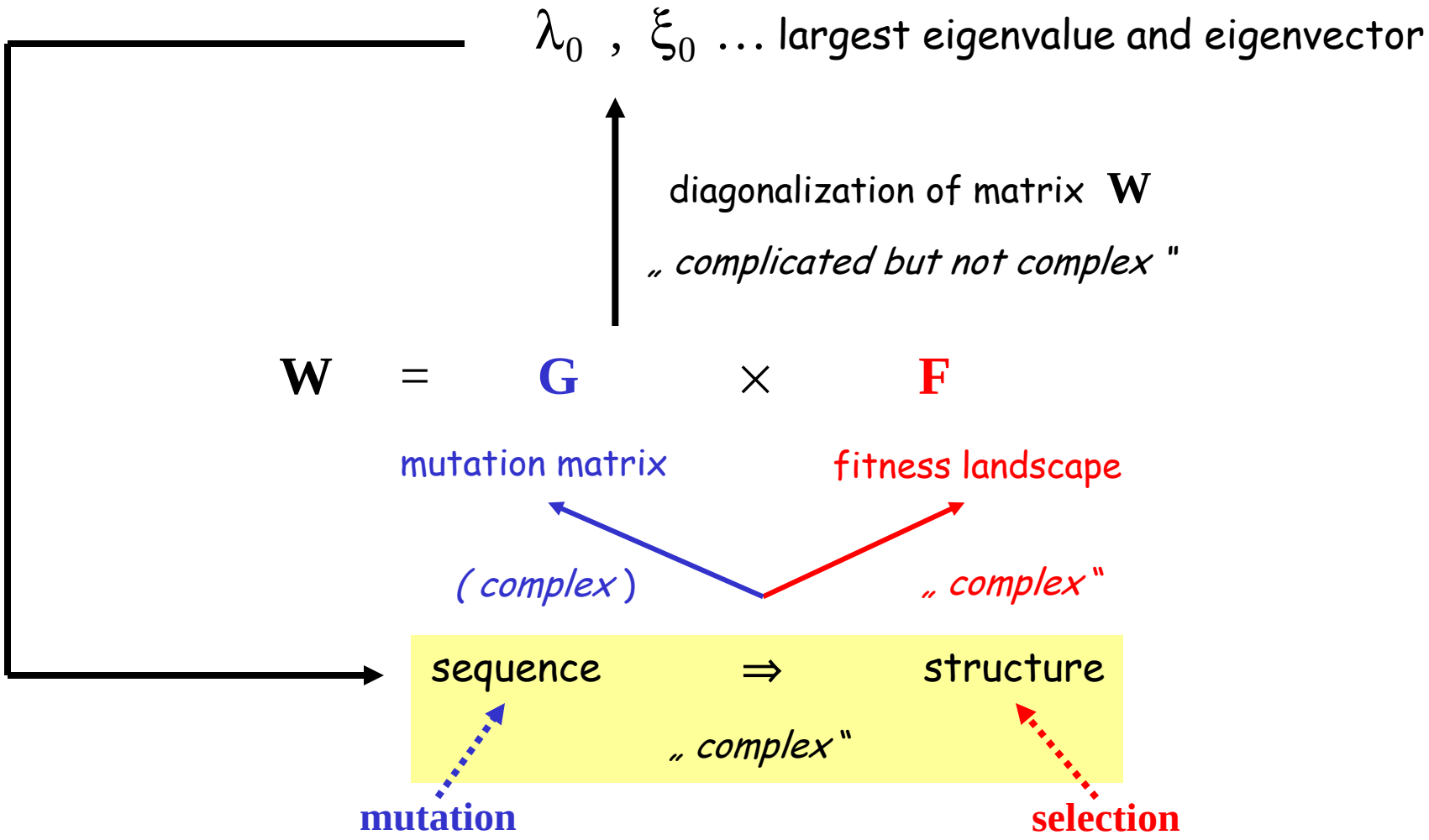
Evolution as a global phenomenon in genotype space



Sequence space

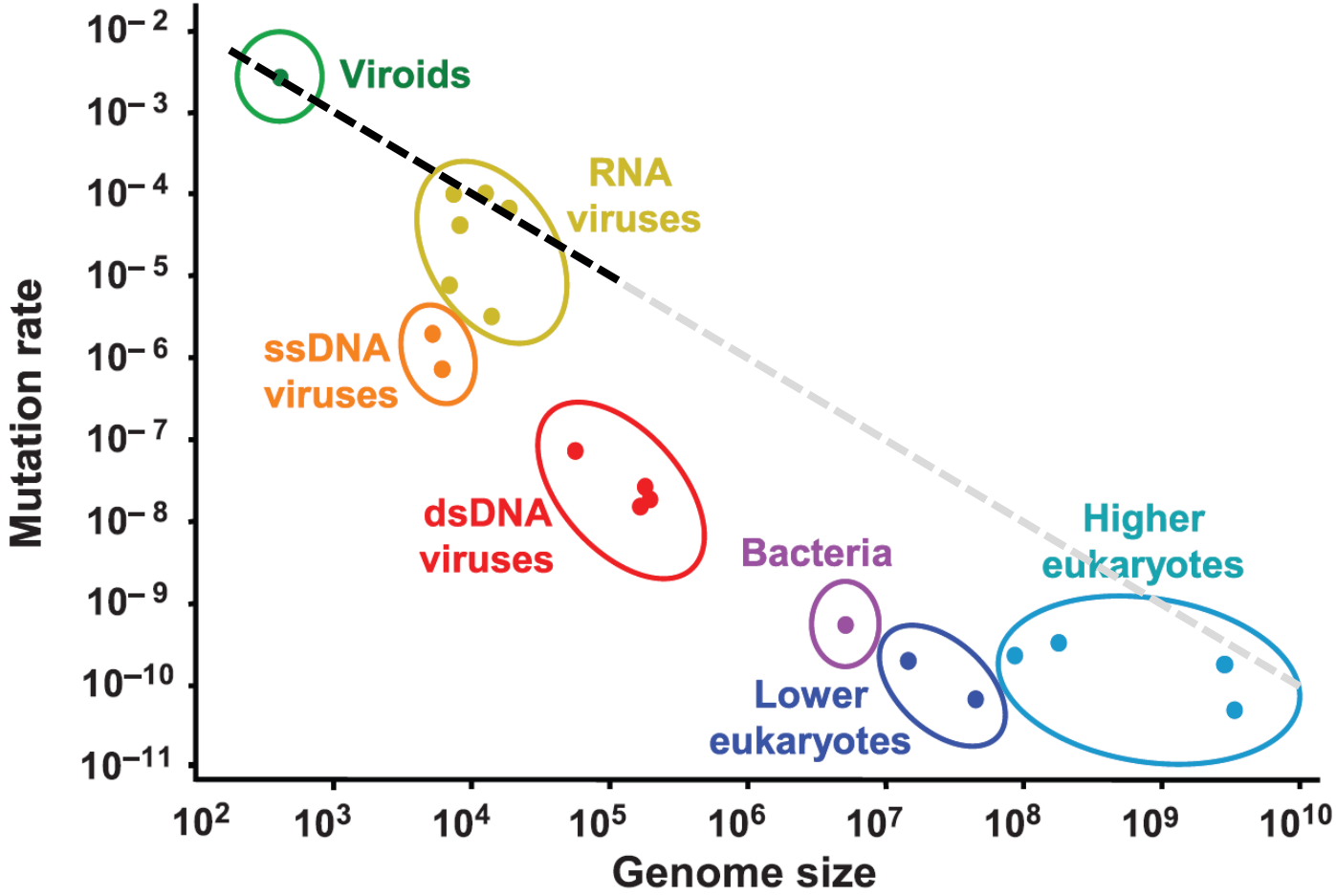
Structure space

many genotypes ⇒ one phenotype



Complexity in molecular evolution

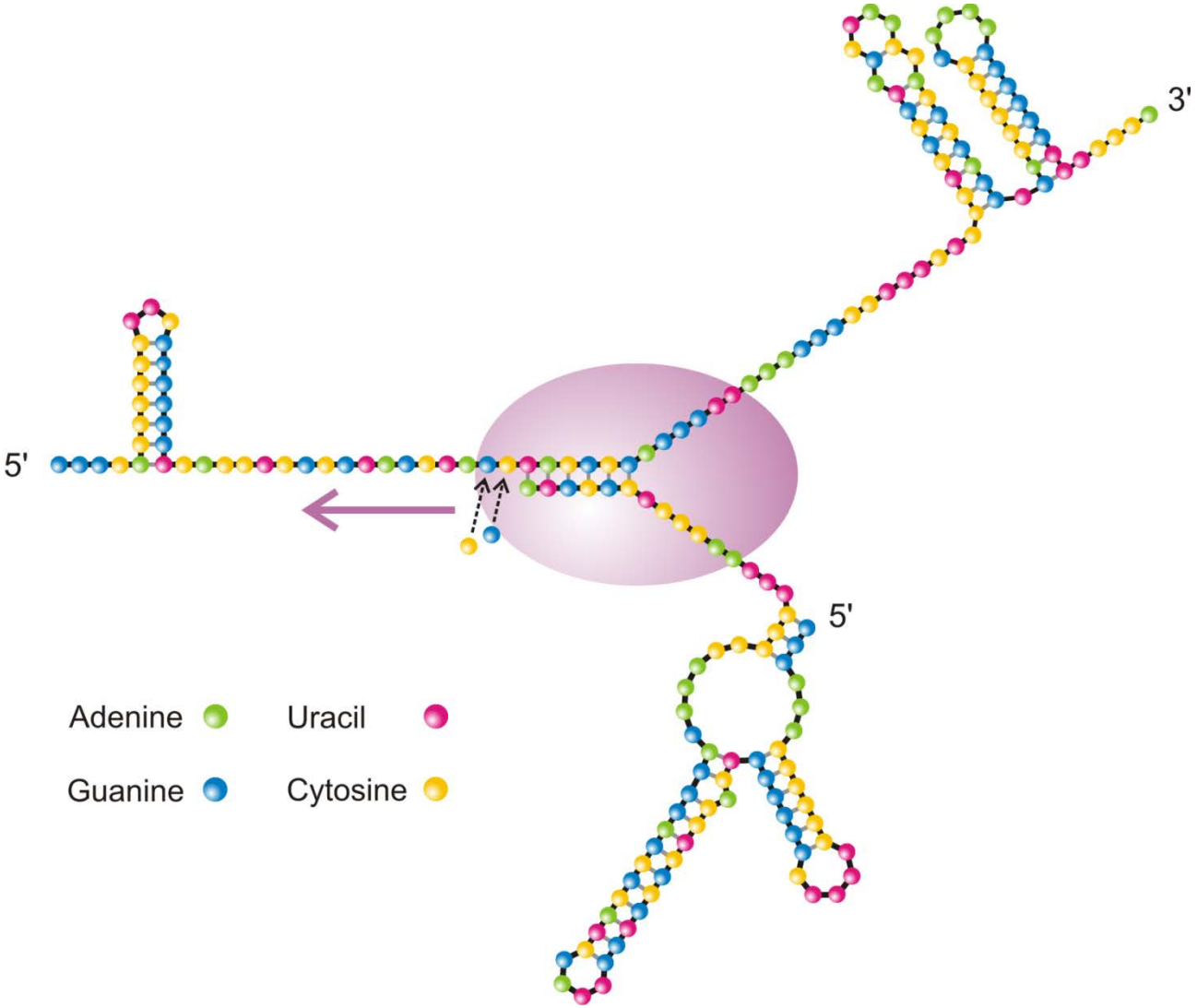
1. Prologue
2. Mathematics of Darwin's natural selection
3. Mendel, Fisher and population genetics
4. Mutations and selection
5. What means neutrality in evolution?
6. **Evolution in simple systems**
7. Some origins of complexity in biology

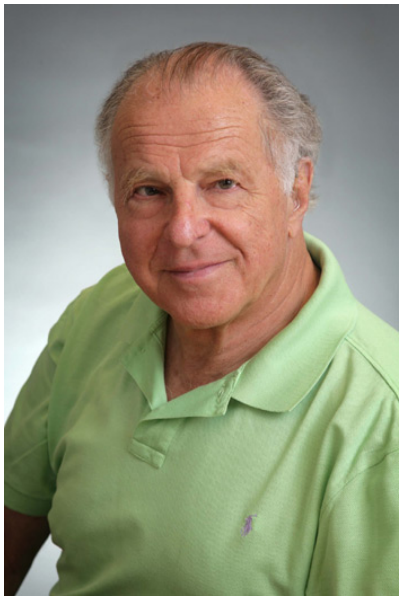


Selma Gago, Santiago F. Elena, Ricardo Flores, Rafael Sanjuán. 2009. Extremely high mutation rate of a hammerhead viroid. Science 323:1308.

Mutation rate and genome size

Replicating molecules



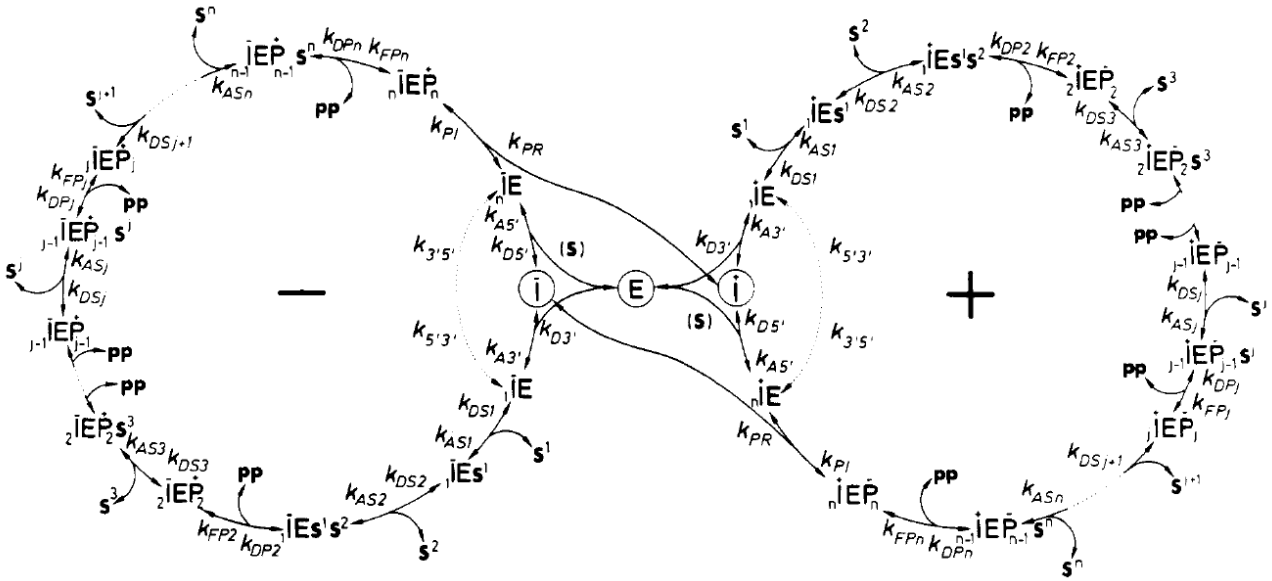


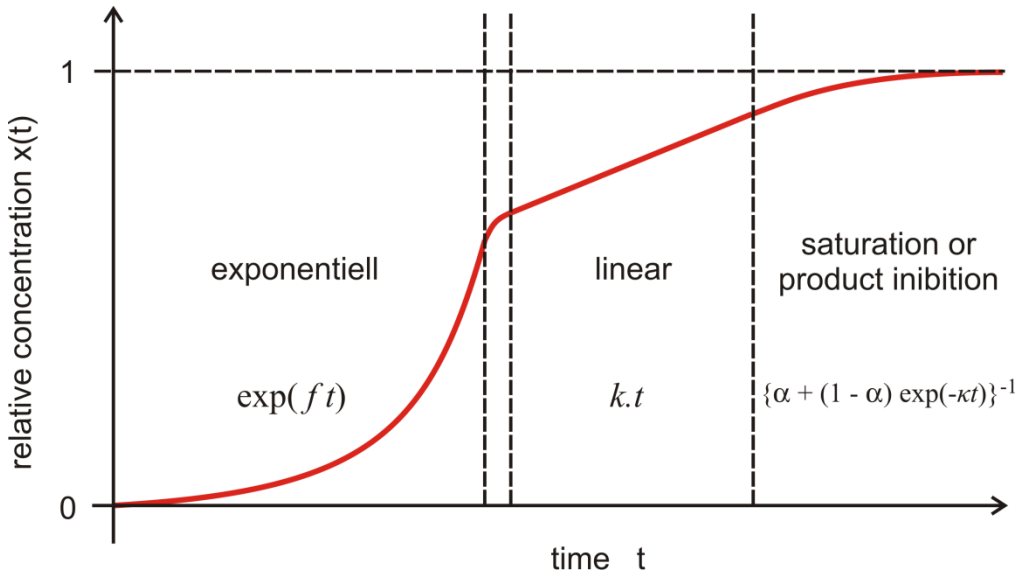
Charles Weissmann
1931-

RNA replication by Qβ-replicase

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

Christof K. Biebricher,
1941-2009





Kinetics of RNA replication

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

$$
\begin{aligned}
X_+ + E &\underset{h_2^+}{\overset{h_1^+}{\rightleftharpoons}} EX_+ \,, \\
EX_+ + 2\,A &\xrightarrow{g_+} I_-EX_+ \,, \\
I_-EX_+ + (n-2)\,A &\xrightarrow{k_+} X_-EX_+ \,, \\
X_-EX_+ &\underset{d_2^+}{\overset{d_1^+}{\rightleftharpoons}} X_- + EX'_+ \,, \\
EX'_+ &\underset{b_2^+}{\overset{b_1^+}{\rightleftharpoons}} X_+ + E \,, \\
X_- + E &\underset{h_2^-}{\overset{h_1^-}{\rightleftharpoons}} EX_- \,, \\
EX_- + 2\,A &\xrightarrow{g_-} I_+EX_- \,, \\
I_+EX_- + (n-2)\,A &\xrightarrow{k_-} X_-EX_+ \,, \\
X_+EX_- &\underset{d_2^-}{\overset{d_1^-}{\rightleftharpoons}} X_+ + EX'_- \,, \\
EX'_- &\underset{b_2^-}{\overset{b_1^-}{\rightleftharpoons}} X_- + E \,.
\end{aligned}
$$

$$
\begin{aligned}
\frac{da}{dt} &= -\,2\,(g_+y_+ + g_-y_-)\,a^2 \,-\, (n-2)\,(k_+m_+ + k_-m_-)\,a^{n-2} \\
\frac{de}{dt} &= -\,(h_1^+x_+ + h_1^-x_- + b_2^+x_+ + b_2^-x_-)\,e + \\
&\quad + h_2^+y_+ + h_2^-y_- + b_1^+z_+ + b_1^-z_- \\
\frac{dx_+}{dt} &= -\,(h_1^+e + b_2^+e + d_2^-z_-)\,x_+ \,+\, h_2^+y_+ + b_1^+z_+ + d_1^-w_- \\
\frac{dy_+}{dt} &= -\,(h_2^+ + g_+a^2)\,y_+ \,+\, h_1^+x_+e \\
\frac{dm_+}{dt} &= -\,k_+a^{n-2}m_+ \,+\, g_+a^2y_+ \\
\frac{dw_+}{dt} &= -\,d_1^+w_+ \,+\, d_2^+x_-z_+ \,+\, k_+a^{n-2}m_+ \\
\frac{dz_+}{dt} &= -\,(b_1^+ + d_2^+x_-)\,z_+ \,+\, d_1^+w_+ \,+\, b_2^+x_+e \\
\frac{dx_-}{dt} &= -\,(h_1^-e + b_2^-e + d_2^+z_+)\,x_- \,+\, h_2^-y_- + b_1^-z_- + d_1^+w_+ \\
\frac{dy_-}{dt} &= -\,(h_2^- + g_-a^2)\,y_- \,+\, h_1^-x_-e \\
\frac{dm_-}{dt} &= -\,k_-a^{n-2}m_- \,+\, g_-a^2y_- \\
\frac{dw_-}{dt} &= -\,d_1^-w_- \,+\, d_2^-x_+z_- \,+\, k_-a^{n-2}m_- \\
\frac{dz_-}{dt} &= -\,(b_1^- + d_2^-x_+)\,z_- \,+\, d_1^-w_- \,+\, b_2^-x_-e \,.
\end{aligned}
$$

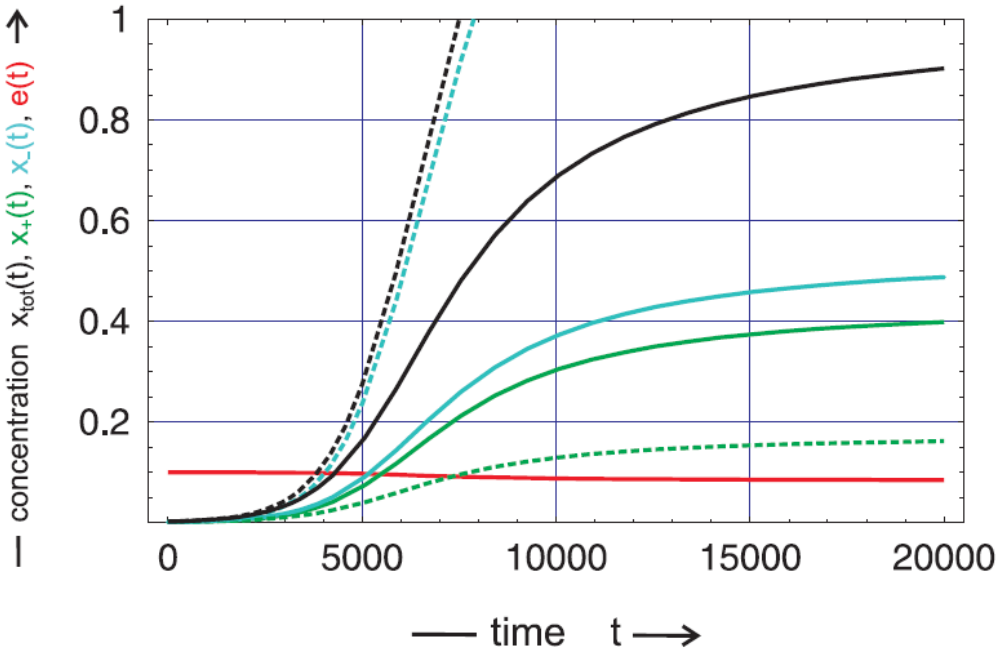
Paul E. Phillipson, Peter Schuster. 2009.
Modeling by nonlinear differential equations.
Dissipative and conservative processes.
World Scientific Publishing, Hackensack, NJ.

replicase e(t)

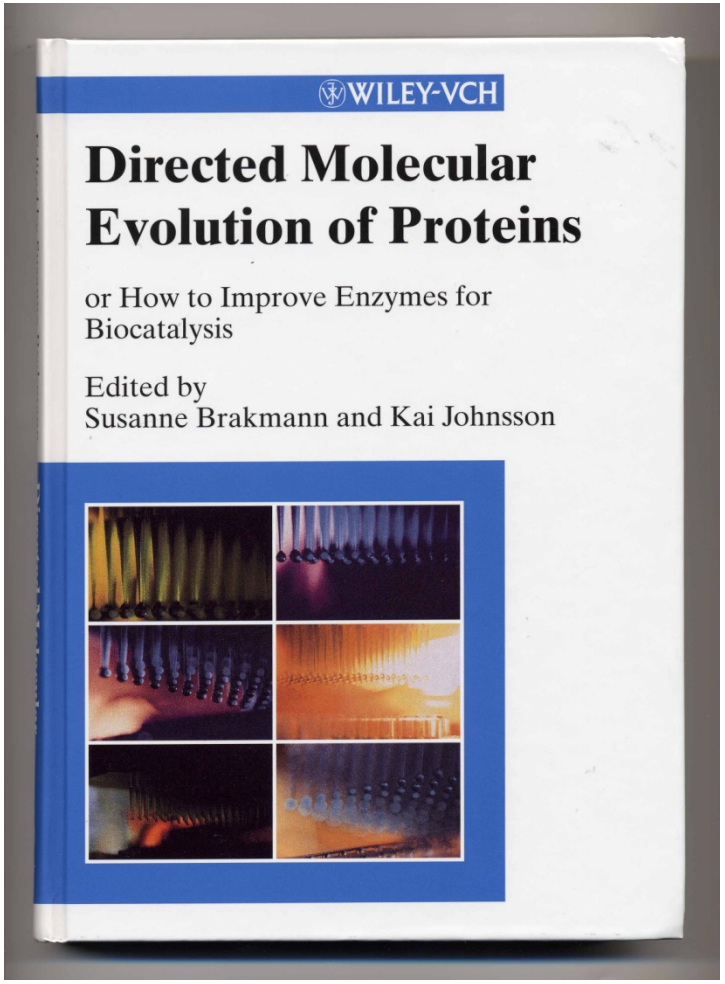
plus strand $x_+(t)$
minus strand $x_-(t)$

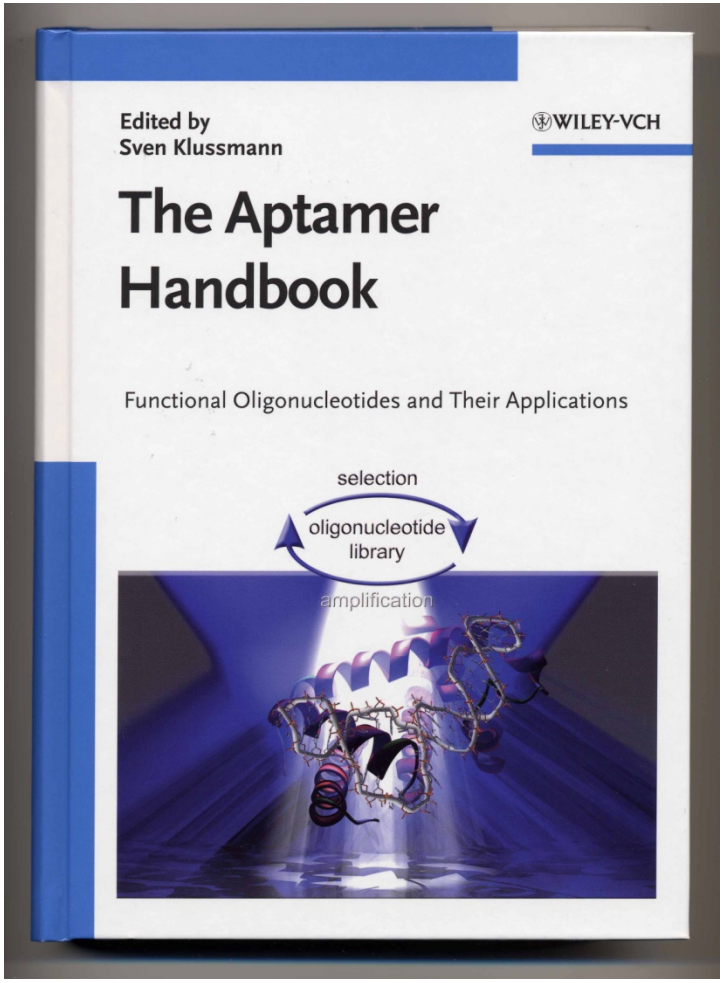
total RNA concentration
$x_{tot}(t) = x_+(t) + x_-(t)$

complemetary replication



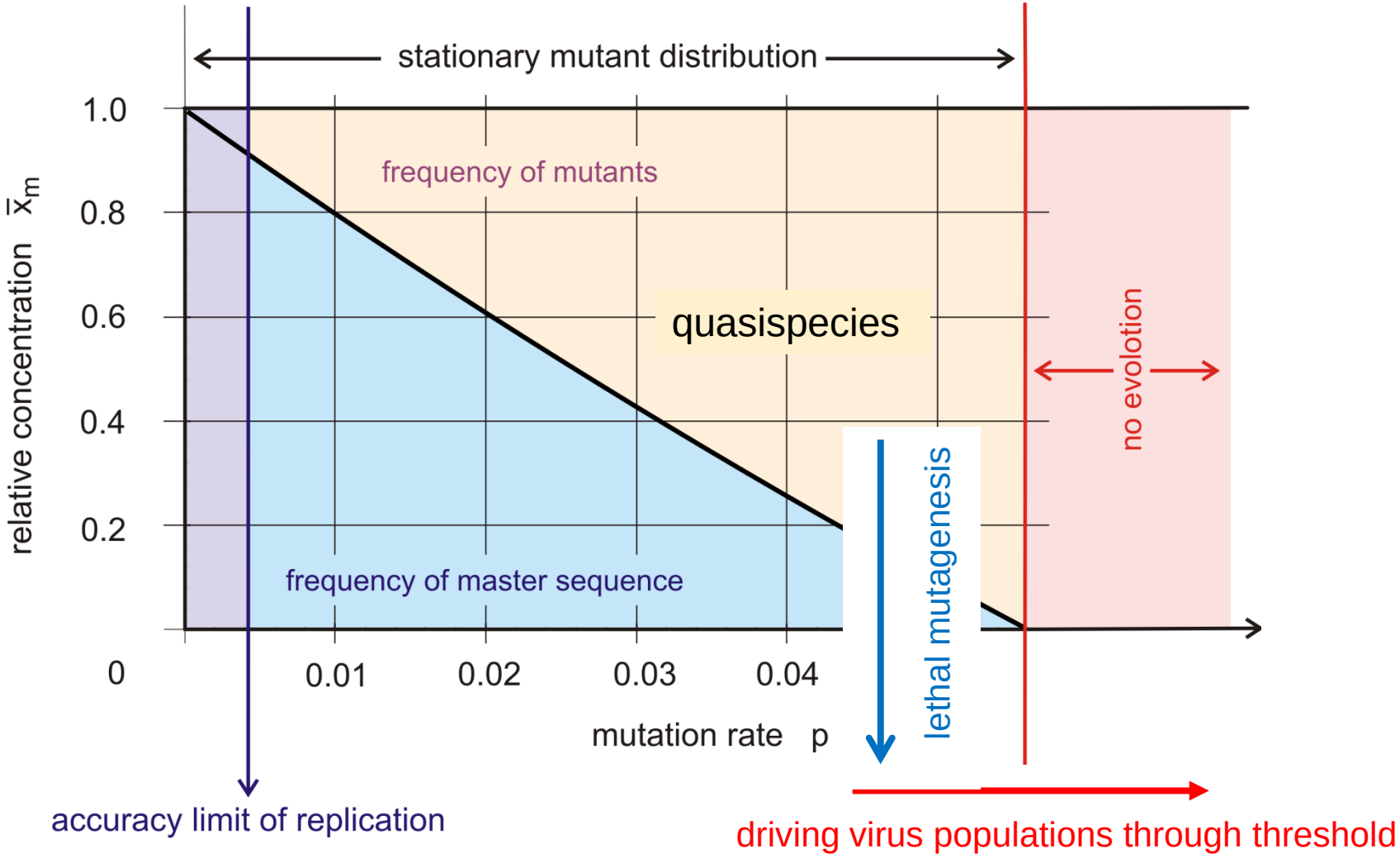
Paul E. Phillipson, Peter Schuster. 2009.
Modeling by nonlinear differential equations.
Dissipative and conservative processes.
World Scientific Publishing, Hackensack, NJ.





Application of molecular evolution to problems in biotechnology

Viruses



The error threshold in replication

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

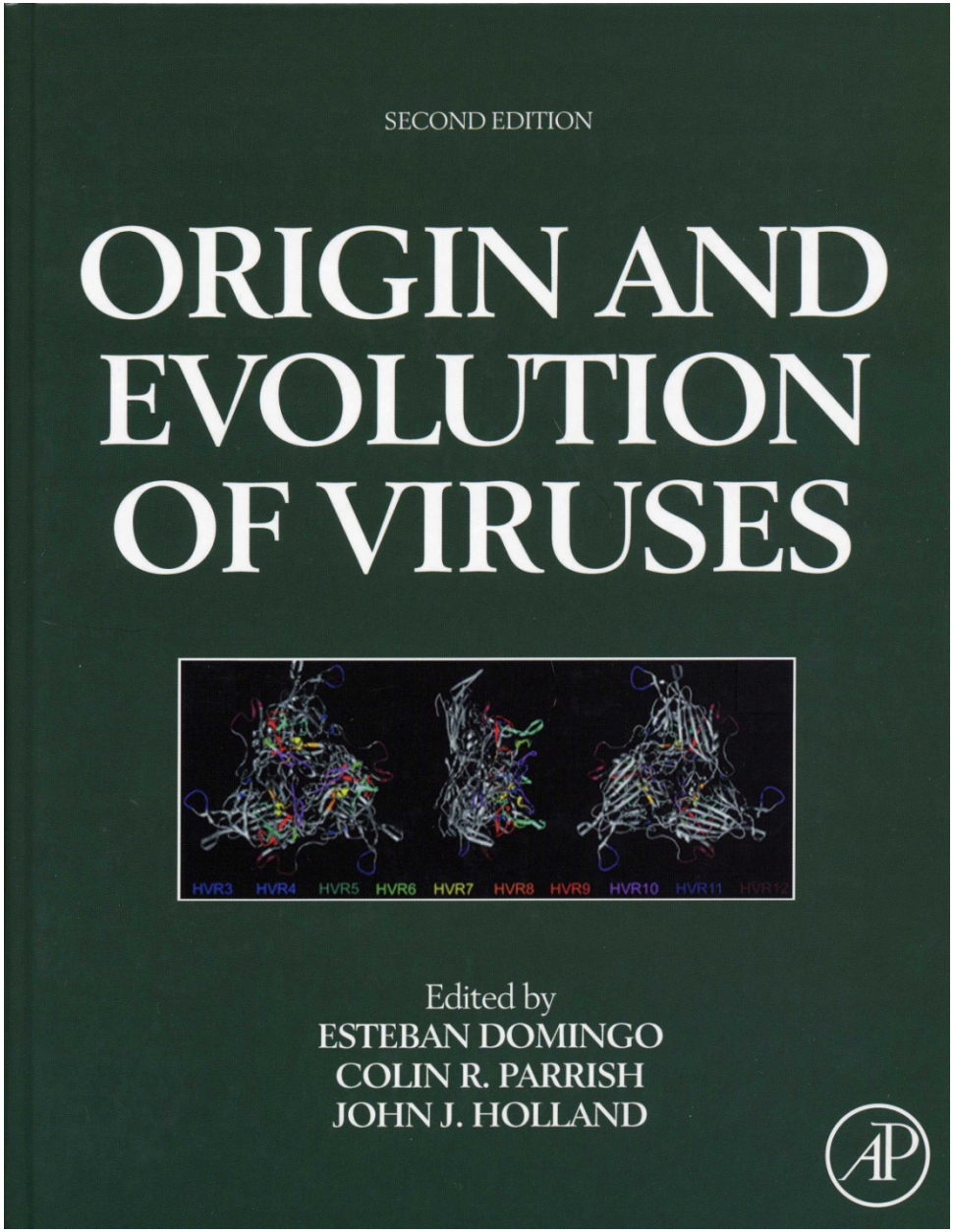
Available online 8 December 2004

doi:10.1016/j.virusres.2004.11.001

Esteban Domingo
1943 -

Application of quasispecies theory to the fight against viruses

Molecular evolution of viruses



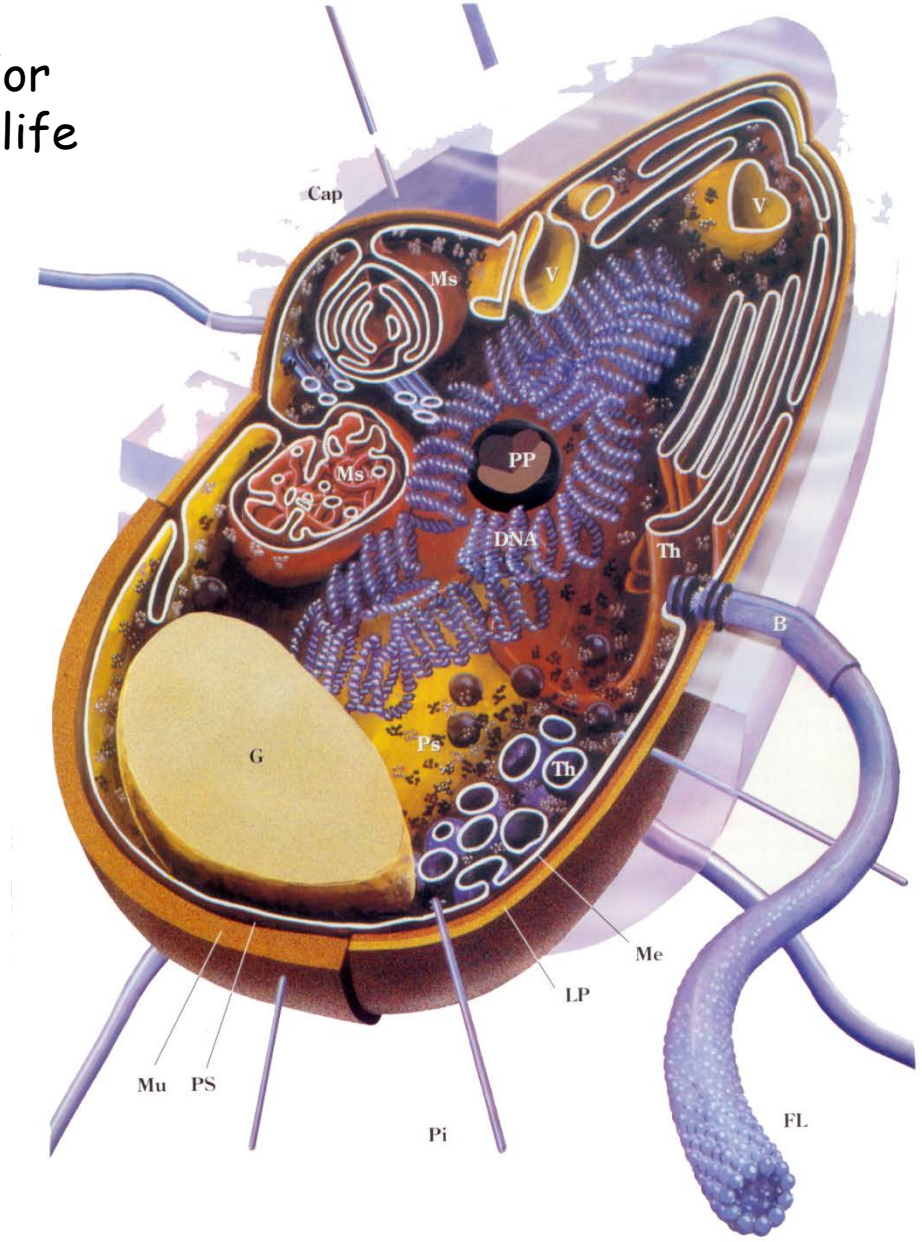
1. Prologue
2. Mathematics of Darwin's natural selection
3. Mendel, Fisher and population genetics
4. Mutations and selection
5. What means neutrality in evolution?
6. Evolution in simple systems
7. **Some origins of complexity in biology**

The bacterial cell as an example for the simplest form of autonomous life

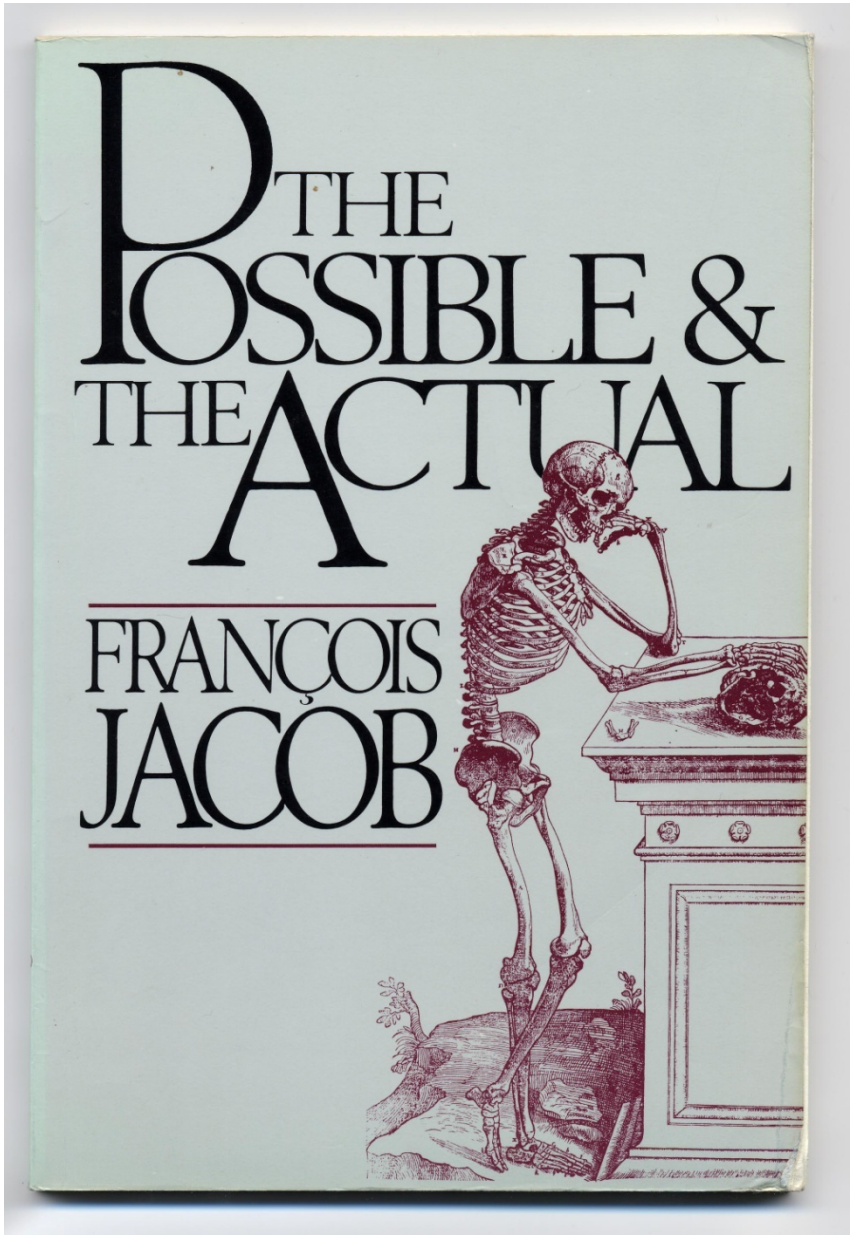
Escherichia coli genome:

4 million nucleotides

4460 genes



The spatial structure of the bacterium *Escherichia coli*

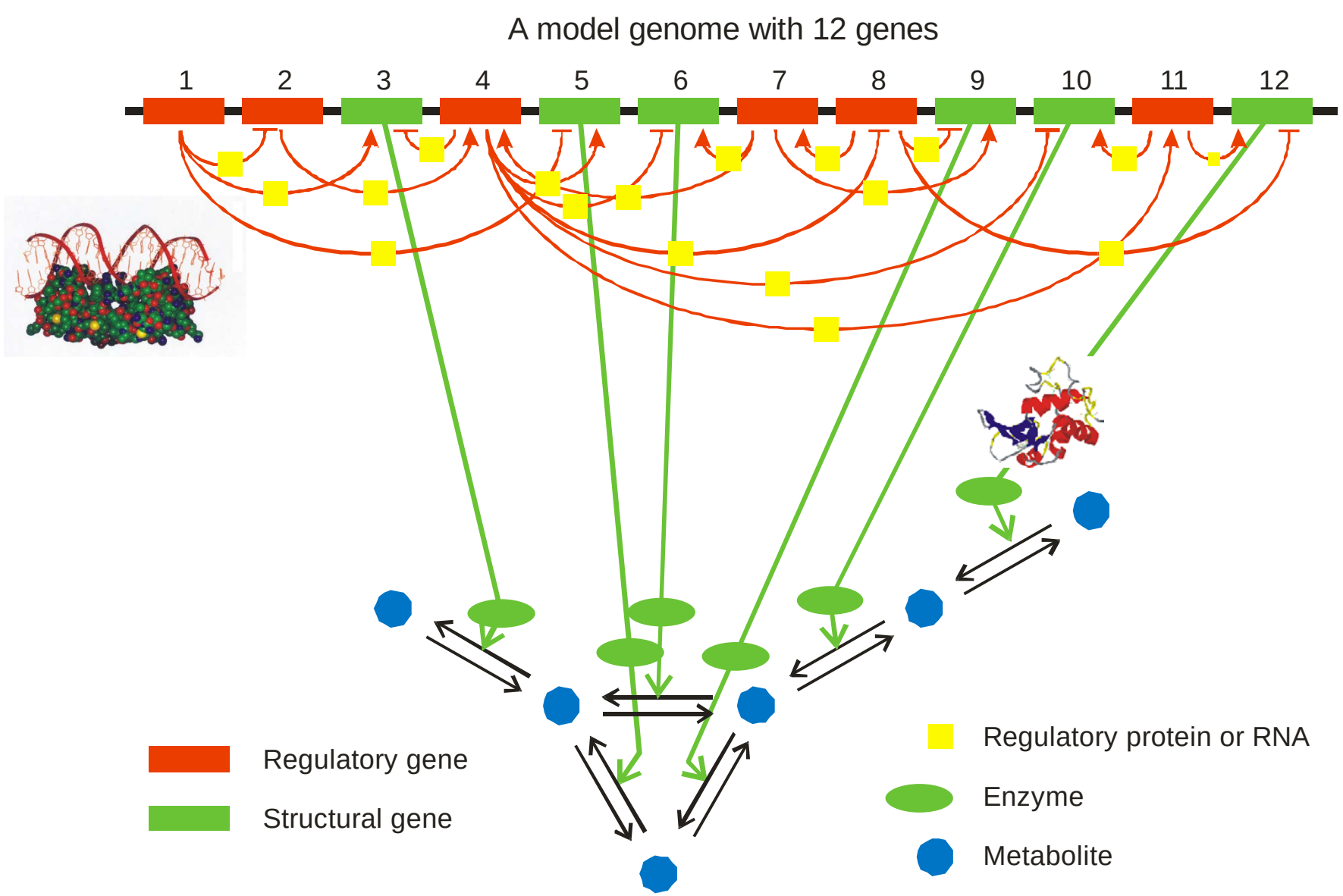


François Jacob, 1920-2013

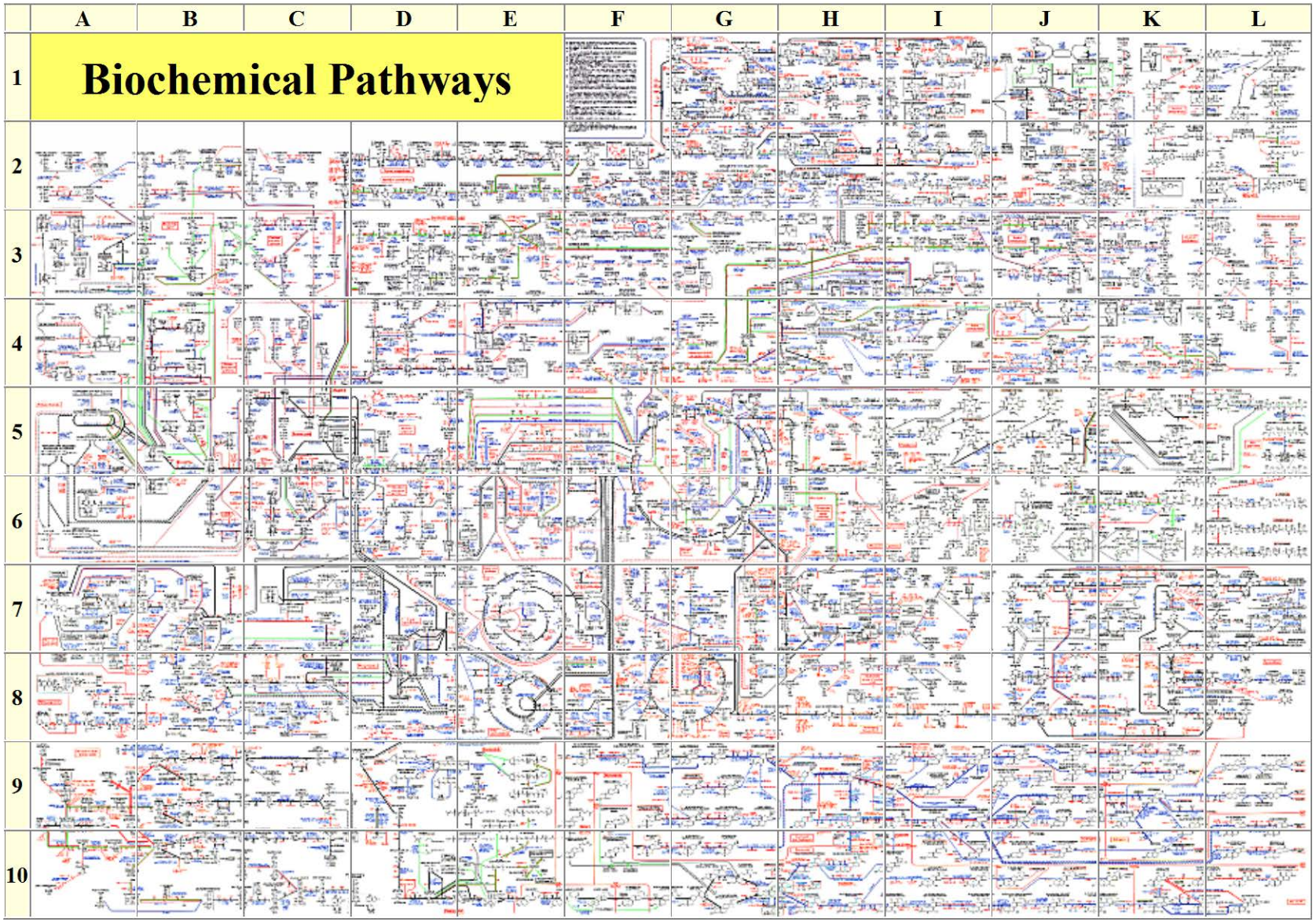
Evolution does not design with the eyes of an engineer, evolution works like a tinkerer.

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

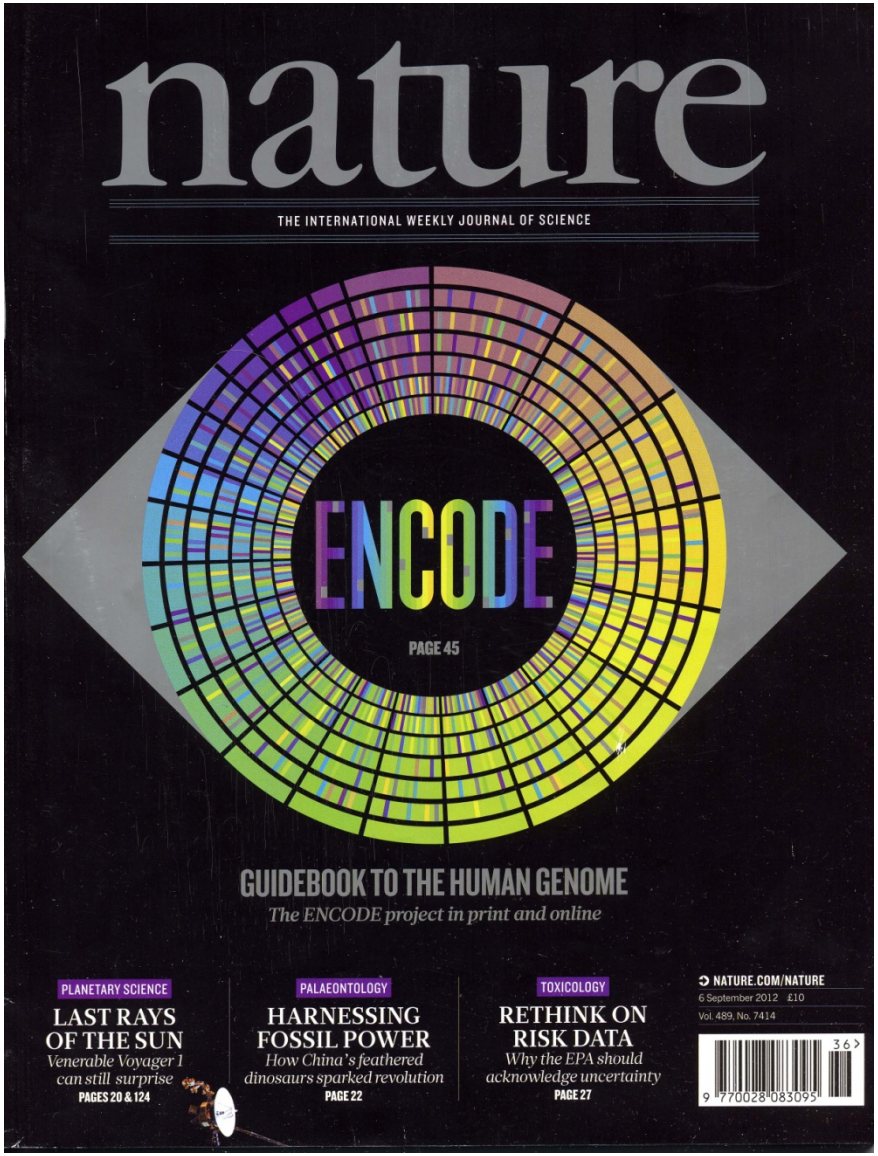
Evolutionary tinkering. *Science* **196** (1977), 1161-1166.



Sketch of a genetic and metabolic network



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



ENCyclopedia **O**f **D**NA **E**lements

Coworkers

Universität Wien

Walter Fontana, Harvard Medical School, MA

Josef Hofbauer, Universität Wien, AT

Martin Nowak, Harvard University, MA

Christian Reidys, University of Southern Denmark, Odense, DK

Karl Sigmund, Universität Wien, AT

Peter Stadler, **Bärbel M. Stadler**, Universität Leipzig, DE

Thomas Wiehe, Universität Köln, DE

Ivo L.Hofacker, **Christoph Flamm**, Universität Wien, AT

Paul Phillipson, University of Boulder at Colorado, CO

Christian Forst, University of Texas, Southwestern Medical School, TX

Erich Bornberg-Bauer, Universität Münster, DE

Thank you for your attention!

Web-Page for further information:

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