Is the Concept of Error Catastrophy Relevant for Viruses?

Quasispecies and error thresholds on realistic landscapes

Peter Schuster

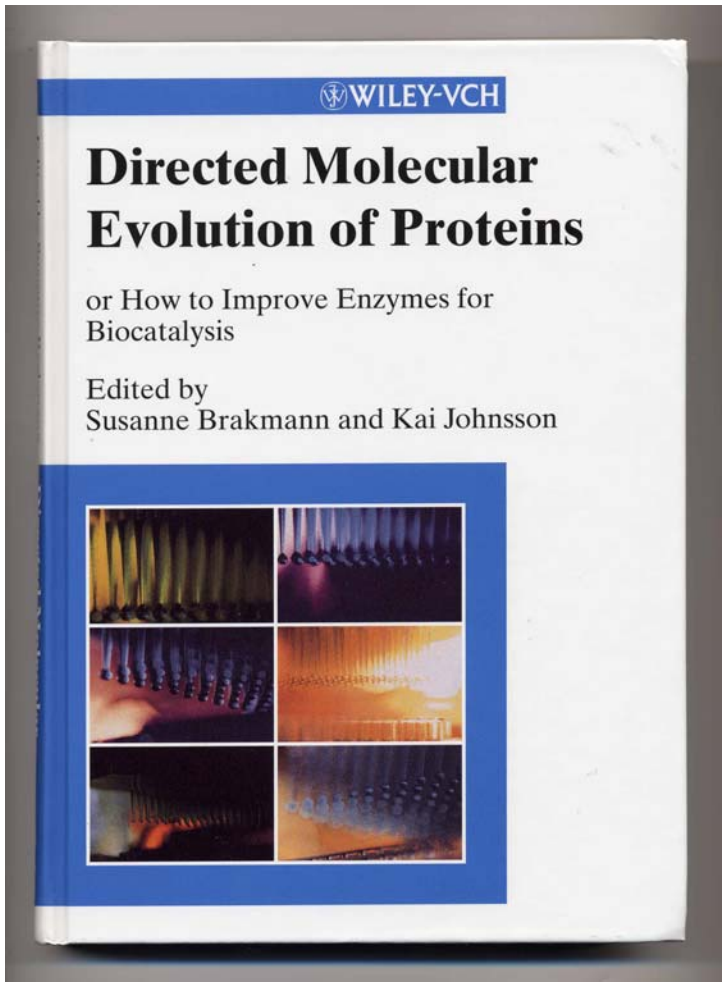
Institut für Theoretische Chemie, Universität Wien, Austria
and
The Santa Fe Institute, Santa Fe, New Mexico, USA

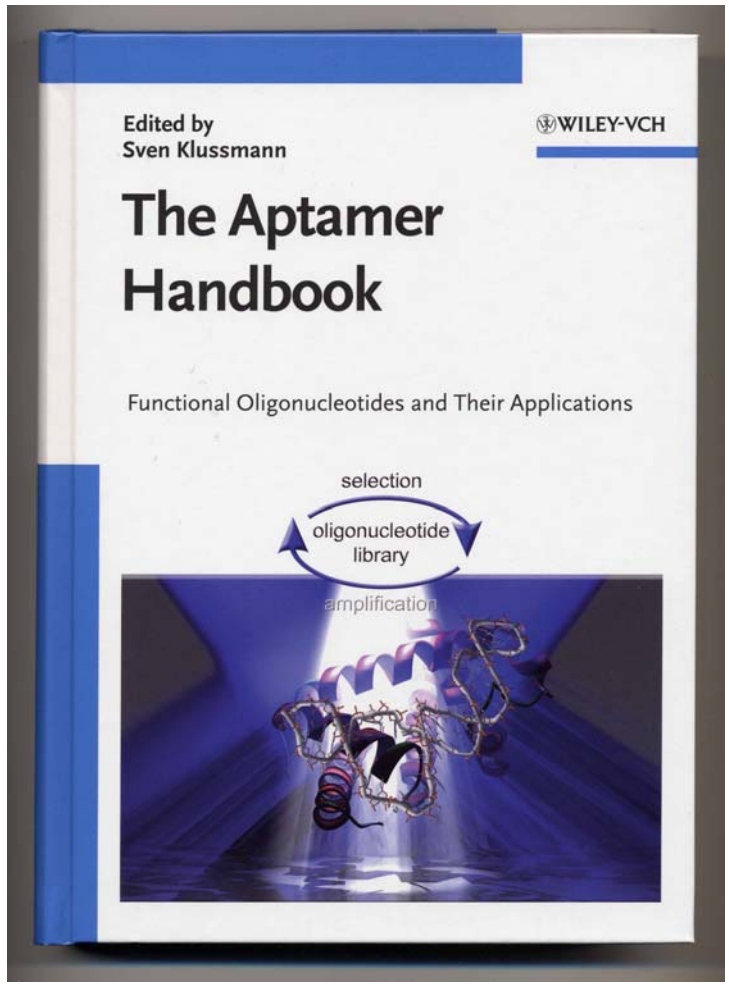
Interdisziplinäres Zentrum für Bioinformatik (IZBI)

Universität Leipzig, 11.04.2008

Web-Page for further information:

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





Application of molecular evolution to problems in biotechnology

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

0168-1702/$ – see front matter
doi:10.1016/j.virusres.2004.11.001

BMC Evolutionary Biology

BioMed Central

Correspondence

Open Access

Quasispecies theory in the context of population genetics

Claus O Wilke*[1,2]

Address: [1]Keck Graduate Institute of Applied Life Sciences, 535 WatsonDrive, Claremont, California 91711, USA and [2]Digital Life Laboratory, California Institute of Technology, Mail Code 136-93, Pasadena, California 91125, USA

Email: Claus O Wilke* - wilke@kgi.edu

* Corresponding author

Published: 17 August 2005

BMC Evolutionary Biology 2005, 5:44 doi:10.1186/1471-2148-5-44

Received: 16 February 2005

Accepted: 17 August 2005

This article is available from: http://www.biomedcentral.com/1471-2148/5/44

Abstract

Background: A number of recent papers have cast doubt on the applicability of the quasispecies concept to virus evolution, and have argued that population genetics is a more appropriate framework to describe virus evolution than quasispecies theory.

Results: I review the pertinent literature, and demonstrate for a number of cases that the quasispecies concept is equivalent to the concept of mutation-selection balance developed in population genetics, and that there is no disagreement between the population genetics of haploid, asexually-replicating organisms and quasispecies theory.

Conclusion: Since quasispecies theory and mutation-selection balance are two sides of the same medal, the discussion about which is more appropriate to describe virus evolution is moot. In future work on virus evolution, we would do good to focus on the important questions, such as whether we can develop accurate, quantitative models of virus evolution, and to leave aside discussions about the relative merits of perfectly equivalent concepts.

OPEN ACCESS Freely available online

PLoS COMPUTATIONAL BIOLOGY

Review

Quasispecies Made Simple

J. J. Bull, Lauren Ancel Meyers, Michael Lachmann*

ABSTRACT

Quasispecies are clouds of genotypes that appear in a population at mutation–selection balance. This concept has recently attracted the attention of virologists, because many RNA viruses appear to generate high levels of genetic variation that may enhance the evolution of drug resistance and immune escape. The literature on these important evolutionary processes is, however, quite challenging. Here we use simple models to link mutation–selection balance theory to the most novel property of quasispecies: the error threshold—a mutation rate below which populations equilibrate in a traditional mutation–selection balance and above which the population experiences an error catastrophe, that is, the loss of the favored genotype through frequent deleterious mutations. These models show that a single fitness landscape may contain multiple, hierarchically organized error thresholds and that an error threshold is affected by the extent of back mutation and redundancy in the genotype-to-phenotype map. Importantly, an error threshold is distinct from an extinction threshold, which is the complete loss of the population through lethal mutations. Based on this framework, we argue that the lethal mutagenesis of a viral infection by mutation-inducing drugs is not a true error catastrophe, but is an extinction catastrophe.

Introduction

The concept of a mutation–selection balance is one of the oldest and most fundamental pillars of population genetics: natural selection increases the frequency of fit variants while mutations introduce unfit variants, giving rise to an equilibrium distribution balanced between these two effects. Mutation–selection balance has been invoked to explain the persistence of undesirable genes, for example, those underlying inbreeding depression, genetic diseases, and even senescence. Despite the long history of the concept, some of its consequences were only realized in 1971, when Manfred Eigen studied mutation–selection dynamics in long genomes [1]. He found that populations do not necessarily attain classic mutation–selection balances in which the wild-type allele is most common, but rather attain an equilibrium with an abundant assemblage of mutant genotypes and a rare wild-type. He and Peter Schuster later called this collection of genotypes at equilibrium a quasispecies [2]. This concept offered not only an intuitive extension of the mutation–selection theory based on simple one- or two-locus systems, but also a novel insight into the impact of mutation rate on evolutionary dynamics. In particular, Eigen found that there are states in which a trivial boost in the mutation rate can lead to a fundamental change in the composition of genotypes in the population. This change, a phase transition in physics terms, is called the error catastrophe.

The error catastrophe has been applied liberally as a metaphor for complications of high mutation rates, as likely plagued primordial life [1] and currently challenges extant viruses with RNA genomes [3]. The error-catastrophe model inspired treatments to extirpate viral populations by mutation enhancement [4,5], and the model has been generalized to explain the attraction of populations to mutationally robust regions of fitness landscapes [6]. The error catastrophe has imparted a mystique to the quasispecies concept, and much of the literature on RNA virus evolution now uses quasispecies as an enriched synonym for a high mutation rate. An excellent and short review of the topic and its relationship to population genetics theory is provided by Wilke [7].

Eigen's insights were developed in the context of genomes with many loci, each of which suffered mutation. Appropriately, the quasispecies has since been considered in this large-genome context. Yet many of its concepts are easily illustrated in the much simpler case of few genotypes, which is our approach here. Our results are not new, per se, but our models should convey quasispecies and error-catastrophe concepts to a broad audience and correct some common misunderstandings.

The Simplest Quasispecies

Our basic model has the fewest number of genotypes needed to demonstrate a quasispecies and an error threshold: two [8]. Genotype A_1 has fitness w_1, and of those w_1 offspring a fraction $1 - \mu_1$ retain the A_1 genotype (Figure 1). Its mutants are converted into the other genotype, A_2, which has the lower fitness w_2. A_2 reproduces its genotype with fidelity $1 - \mu_2$, and all of its mutants die.

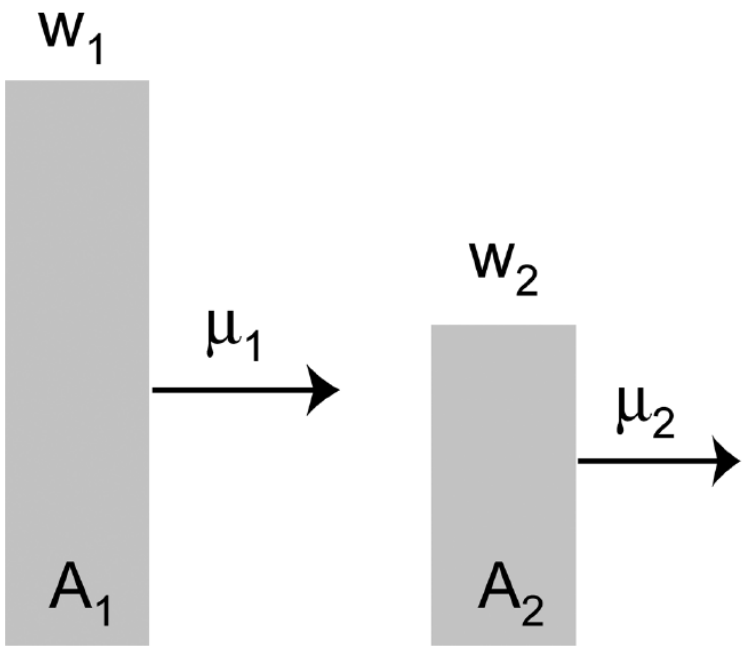
Quasispecies concepts address equilibia, that is, the final distributions of genotypes in populations that have evolved to a stable state. In the quasispecies model, mutation and natural selection steer the population toward the equilibrium distribution, regardless of the initial distribution of genotypes. If the population does not start at the equilibrium, then mutation and natural selection steer it toward equilibrium in the quasispecies model. Mutations introduce new types with various fitnesses while natural selection causes

Citation: Bull JJ, Meyers LA, Lachmann M (2005) Quasispecies made simple. PLoS Comput Biol 1(6): e61.

J. J. Bull is at the Institute for Cellular and Molecular Biology, Section of Integrative Biology, University of Texas, Austin, Texas, United States of America. Lauren Ancel Meyers is at the Institute for Cellular and Molecular Biology, Section of Integrative Biology, University of Texas, Austin, Texas, United States of America, and at the Santa Fe Institute, Santa Fe, New Mexico, United States of America. Michael Lachmann is at the Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany.

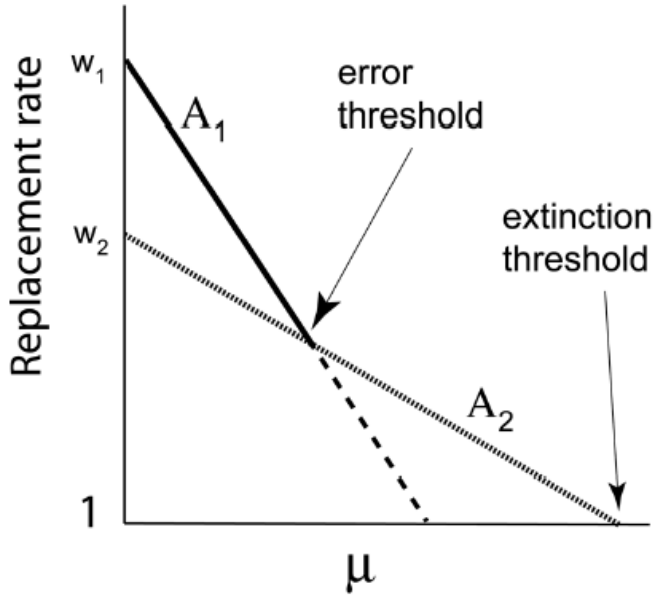
*To whom correspondence should be addressed. E-mail: lachmann@eva.mpg.de

DOI: 10.1371/journal.pcbi.0010061

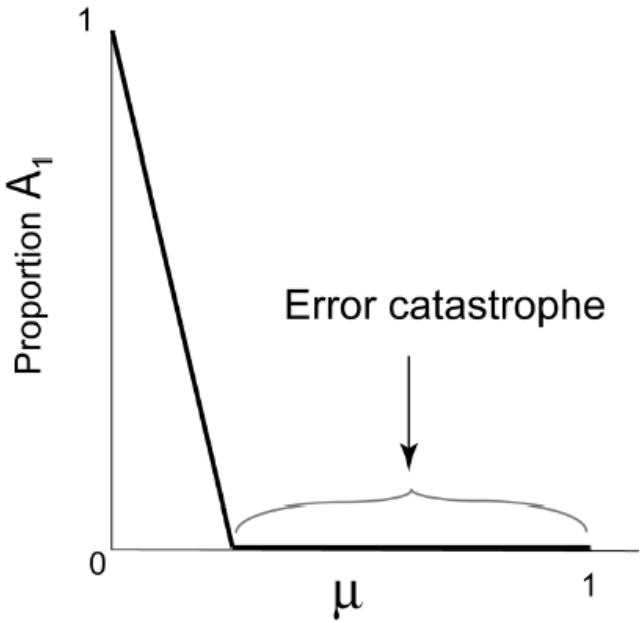


DOI: 10.1371/journal.pcbi.0010061.g001

Figure 1. Model of Two Genotypes with Forward Mutation
Each genotype A_i has its own fitness w_i and mutational loss μ_i. Mutation is asymmetric, so that A_1 gives rise to A_2, but not vice versa.



DOI: 10.1371/journal.pcbi.0010061.g003

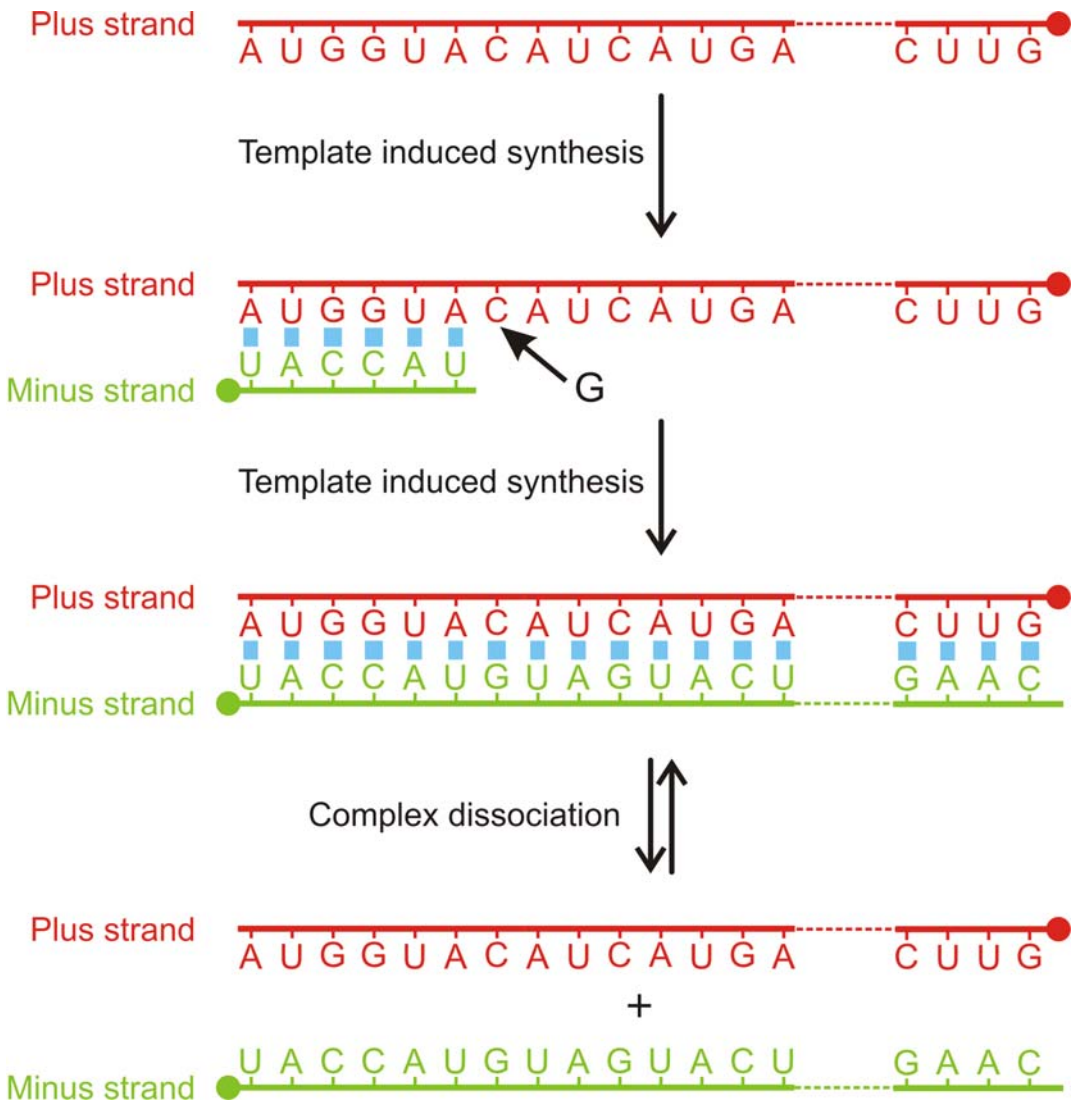


Bull, Ancel Myers and Lachmann

PLoS Computational Biology 1:e61, 2005

1. Replication and mutation
2. Quasispecies and error thresholds
3. Fitness landscapes and randomization
4. Lethal mutations

1. **Replication and mutation**
2. Quasispecies and error thresholds
3. Fitness landscapes and randomization
4. Lethal mutations



Complementary replication is the simplest copying mechanism of RNA.
Complementarity is determined by Watson-Crick base pairs:

G≡C and **A=U**

DIE NATURWISSENSCHAFTEN

58. Jahrgang, 1971 — Heft 10 Oktober

Selforganization of Matter and the Evolution of Biological Macromolecules

MANFRED EIGEN*

Max-Planck-Institut für Biophysikalische Chemie, Karl-Friedrich-Bonhoeffer-Institut, Göttingen-Nikolausberg

I. Introduction

I.1. „Cause and Effect"

The question about the origin of life often appears as a question about "cause and effect". Physical theories of macroscopic processes usually involve answers to such questions, even if a statistical interpretation is given to the relation between "cause" and "effect". It is mainly due to the nature of this question that many scientists believe that our present physics does not offer any obvious explanation for the existence of life, which even in its simplest forms always appears to be associated with complex macroscopic (i.e. multimolecular) systems, such as the living cell.

As a consequence of the exciting discoveries of "molecular biology", a common version of the above question is: *Which came first, the protein or the nucleic acid?* – a modern variant of the old "chicken-and-the-egg" problem. The term "first" is usually meant to define a causal rather than a temporal relationship, and the words "protein" and "nucleic acid" may be substituted by "function" and "information". The question in this form, when applied to the interplay of nucleic acids and proteins as presently encountered in the living cell, leads ad absurdum, because "function"

* Partly presented as the "Robbins Lectures" at Pomona College, California, in spring 1970.

Die Naturwissenschaften 64. Jahrgang Heft 11 November 1977

The Hypercycle

A Principle of Natural Self-Organization

Part A: Emergence of the Hypercycle

Manfred Eigen
Max-Planck-Institut für biophysikalische Chemie, D-3400 Göttingen

Peter Schuster
Institut für theoretische Chemie und Strahlenchemie der Universität, A-1090 Wien

This paper is the first part of a trilogy, which comprises a detailed study of a special type of functional organization and demonstrates its relevance with respect to the origin and evolution of life. Self-replicative macromolecules, such as RNA or DNA in a suitable environment exhibit a behavior, which we may call Darwinian and which can be formally represented by the concept of the quasi-species. A quasi-species is defined as a given distribution of macromolecular species with closely interrelated sequences, dominated by one or several (degenerate) master copies. External constraints enforce the selection of the best adapted distribution, commonly referred to as the wild-type. Most important for Darwinian behavior are the criteria for internal stability of the quasi-species. If these criteria are violated, the information stored in the nucleotide sequence of the master copy will deteriorate irreversibly leading to an error catastrophy. As a consequence, selection and evolution of RNA or DNA molecules is limited with respect to the amount of information that can be stored in a single replicative unit. An analysis of experimental data regarding RNA and DNA replication at various levels of organization reveals, that a sufficient amount of information for the build up of a translation machinery can be gained only via integration of several different replicative units (or reproductive cycles) through functional linkages. A stable functional integration then will raise the system to a new level of organization and thereby enlarge its information capacity considerably. The hypercycle appears to be such a form of organization.

Preview on Part B: The Abstract Hypercycle

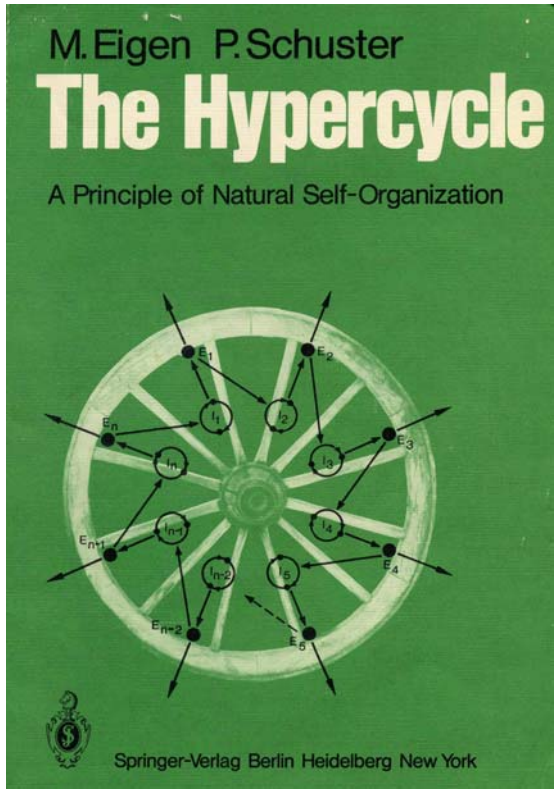
The mathematical analysis of dynamical systems using methods of differential topology, yields the result that there is only one type of mechanisms which fulfills the following requirements: The information stored in each single replicative unit (or reproductive cycle) must be maintained, i.e., the respective master copies must compete favorably with their error distributions. Despite their competitive behavior these units must establish a cooperation which includes all functionally integrated species. On the other hand, the cycle as a whole must continue to compete strongly with any other single entity or linked ensemble which does not contribute to its integrated functions.
These requirements are crucial for a selection of the best adapted functionally linked ensemble and its evolutive optimization. Only hypercyclic organizations are able to fulfil these requirements. Non-cyclic linkages among the autonomous reproduction cycles, such as chains or branched, tree-like networks are devoid of such properties.
The mathematical methods used for proving these assertions are fixed-point, Lyapunov- and trajectorial analysis in higher-dimensional phase spaces, spanned by the concentration coordinates of the cooperating partners. The self-organizing properties of hypercycles are elucidated, using analytical as well as numerical techniques.

Preview on Part C: The Realistic Hypercycle

A realistic model of a hypercycle relevant with respect to the origin of the genetic code and the translation machinery is presented. It includes the following features referring to natural systems:
1) The hypercycle has a sufficiently simple structure to admit an origination with finite probability under prebiotic conditions.
2) It permits a continuous emergence from closely interrelated (t-RNA-like) precursors, originally being members of a stable RNA quasi-species and having been amplified to a level of higher abundance.
3) The organizational structure and the properties of single functional units of this hypercycle are still reflected in the present genetic code in the translation apparatus of the prokaryotic cell, as well as in certain bacterial viruses.

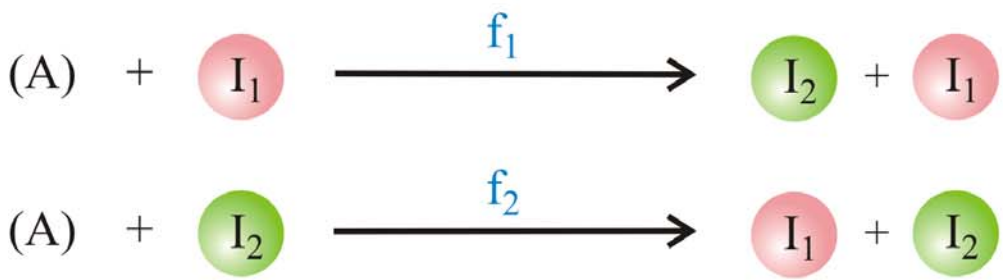
I. The Paradigm of Unity and Diversity in Evolution

Why do millions of species, plants and animals, exist, while there is only one basic molecular machinery of the cell: one universal genetic code and unique chiralities of the macromolecules?
The geneticists of our day would not hesitate to give an immediate answere to the first part of this question. Diversity of species is the outcome of the tremendous branching process of evolution with its myriads of single steps of reproduction and mutation. It in-



Chemical kinetics of molecular evolution

M. Eigen, P. Schuster, `The Hypercycle´, Springer-Verlag, Berlin 1979



$$dx_1 / dt = f_2 x_2 - x_1 \Phi$$
$$dx_2 / dt = f_1 x_1 - x_2 \Phi$$

$$\Phi = \Sigma_i f_i x_i ; \quad \Sigma_i x_i = 1 ; \quad i = 1,2$$

Complementary replication as the simplest molecular mechanism of reproduction

Equation for complementary replication: $[I_i] = x_i \geq 0$, $f_i > 0$; i=1,2

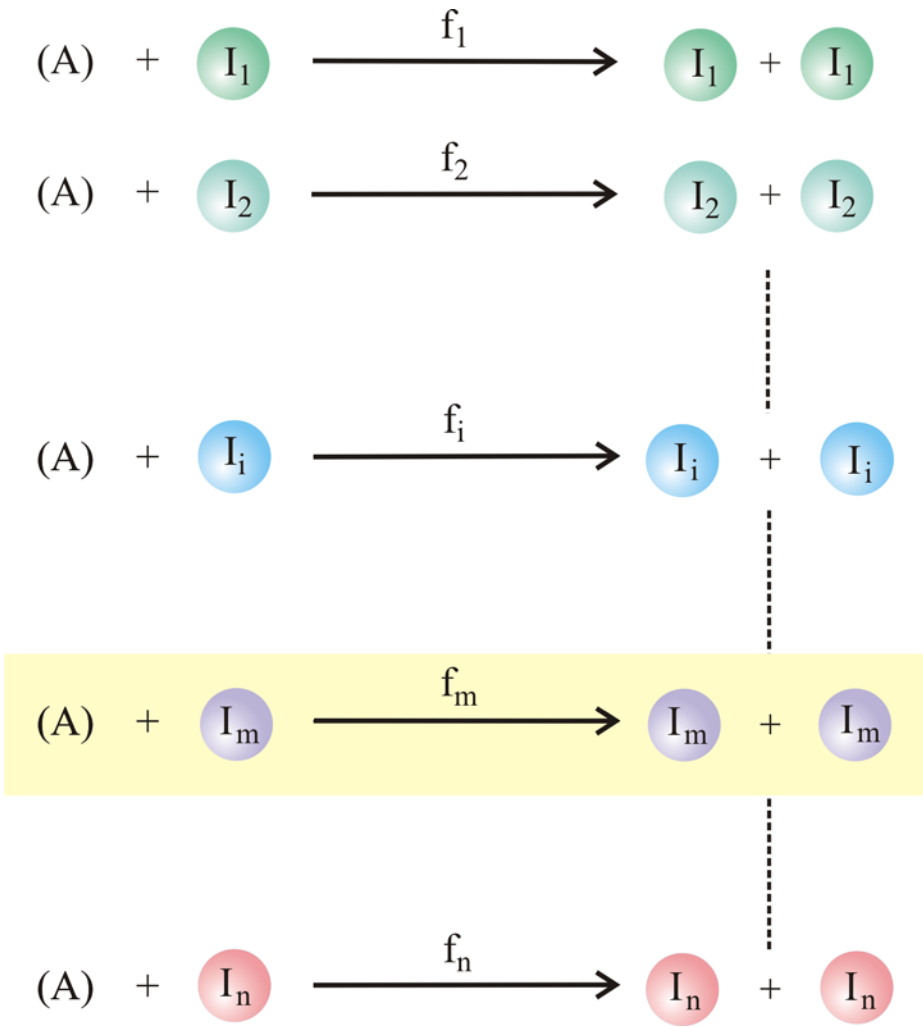
$$\frac{dx_1}{dt} = f_2 x_2 - x_1 \phi , \quad \frac{dx_2}{dt} = f_1 x_1 - x_2 \phi , \quad \phi = f_1 x_1 + f_2 x_2 = \overline{f}$$

Solutions are obtained by integrating factor transformation

$$x_{1,2}(t) = \frac{\sqrt{f_{2,1}}\left(\gamma_1(0)\cdot\exp(ft) + \gamma_2(0)\cdot\exp(-ft)\right)}{(\sqrt{f_1}+\sqrt{f_2})\,\gamma_1(0)\cdot\exp(ft) - (\sqrt{f_1}-\sqrt{f_2})\,\gamma_1(0)\cdot\exp(-ft)}$$

$$\gamma_1(0) = \sqrt{f_1}\,x_1(0) + \sqrt{f_2}\,x_2(0), \gamma_2(0) = \sqrt{f_1}\,x_1(0) - \sqrt{f_2}\,x_2(0), f = \sqrt{f_1 f_2}$$

$$x_1(t) \to \frac{\sqrt{f_2}}{\sqrt{f_1}+\sqrt{f_2}} \text{ and } x_2(t) \to \frac{\sqrt{f_1}}{\sqrt{f_1}+\sqrt{f_2}} \text{ as } \exp(-ft) \to 0$$



$$dx_i / dt = f_i x_i - x_i \Phi = x_i (f_i - \Phi)$$

$$\Phi = \Sigma_j f_j x_j ; \quad \Sigma_j x_j = 1 ; \quad i,j = 1,2,...,n$$

$$[I_i] = x_i \geq 0 ; \quad i = 1,2,...,n ;$$

$$[A] = a = \text{constant}$$

$$f_m = \max \{f_j; j=1,2,...,n\}$$

$$x_m(t) \rightarrow 1 \quad \text{for} \quad t \rightarrow \infty$$

Reproduction of organisms or replication of molecules as the basis of selection

Selection equation: $[\mathrm{I}_i] = x_i \geq 0\ ,\ f_i > 0$

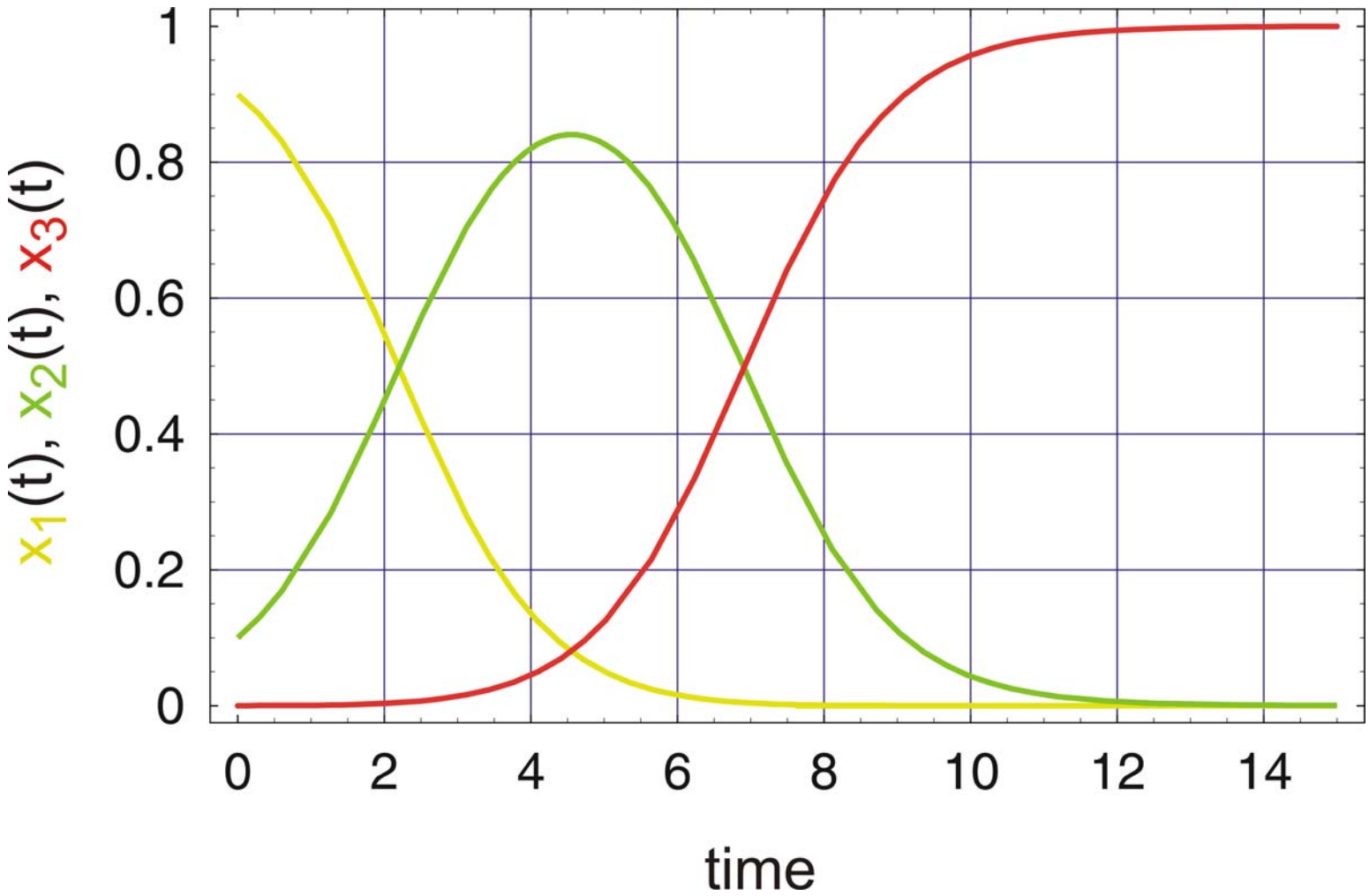
$$\frac{dx_i}{dt} = x_i\left(f_i - \phi\right), \quad i=1,2,\cdots,n; \quad \sum_{i=1}^{n} x_i = 1; \quad \phi = \sum_{j=1}^{n} f_j x_j = \overline{f}$$

Mean fitness or dilution flux, $\phi(\mathrm{t})$, is a non-decreasing function of time,

$$\frac{d\phi}{dt} = \sum_{i=1}^{n} f_i \frac{dx_i}{dt} = \overline{f^2} - \left(\overline{f}\right)^2 = \mathrm{var}\{f\} \geq 0$$

Solutions are obtained by integrating factor transformation

$$x_i(t) = \frac{x_i(0)\cdot \exp\left(f_i t\right)}{\sum_{j=1}^{n} x_j(0)\cdot \exp\left(f_j t\right)}; \quad i = 1,2,\cdots,n$$



Selection between three species with $f_1 = 1$, $f_2 = 2$, and $f_3 = 3$

Stock solution:

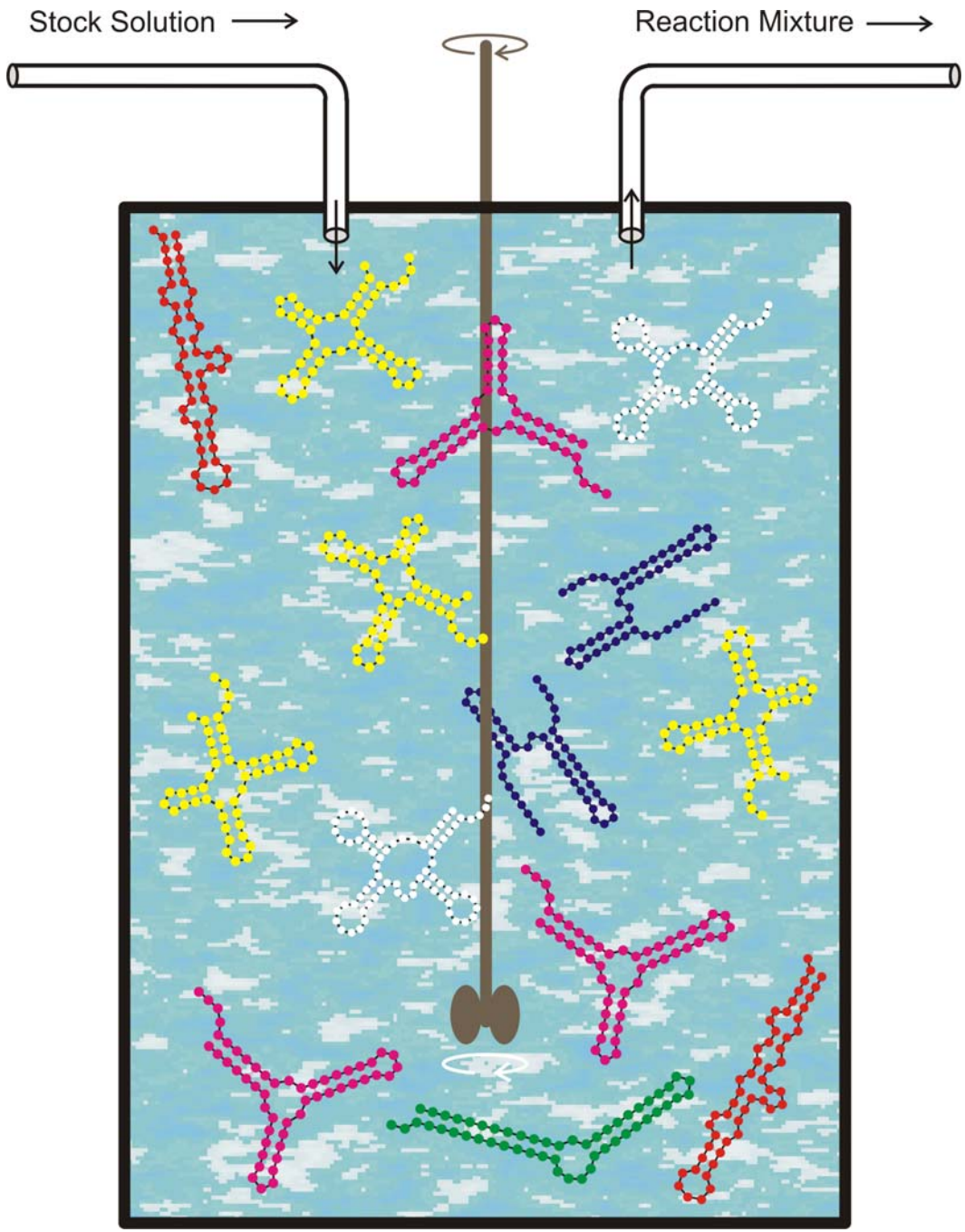
activated monomers, **ATP, CTP, GTP, UTP (TTP);**
a replicase, an enzyme that performs complemantary replication;
buffer solution

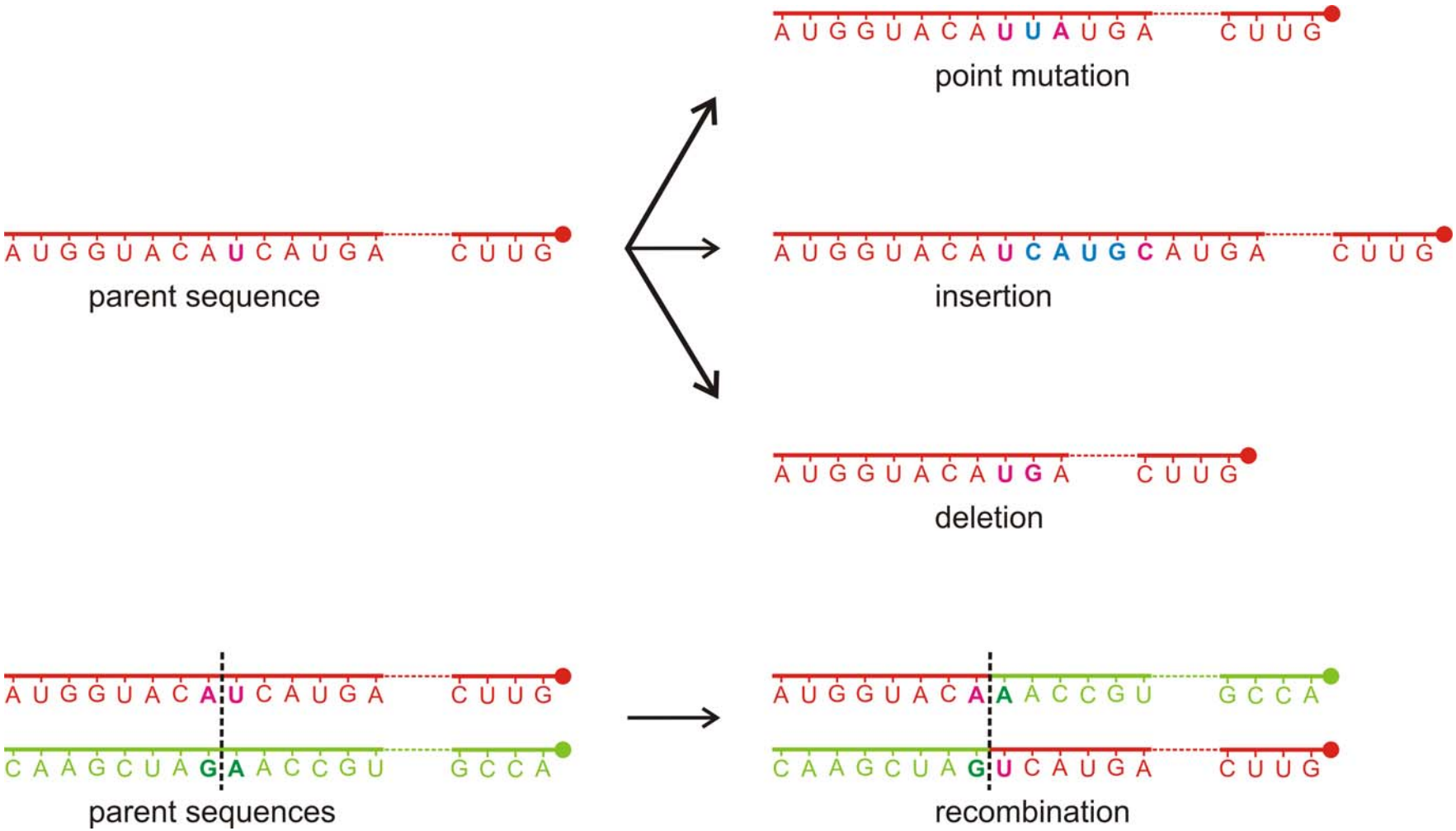
Flow rate: $r = \tau_R^{-1}$

The population size N, the number of polynucleotide molecules, is controlled by the flow r

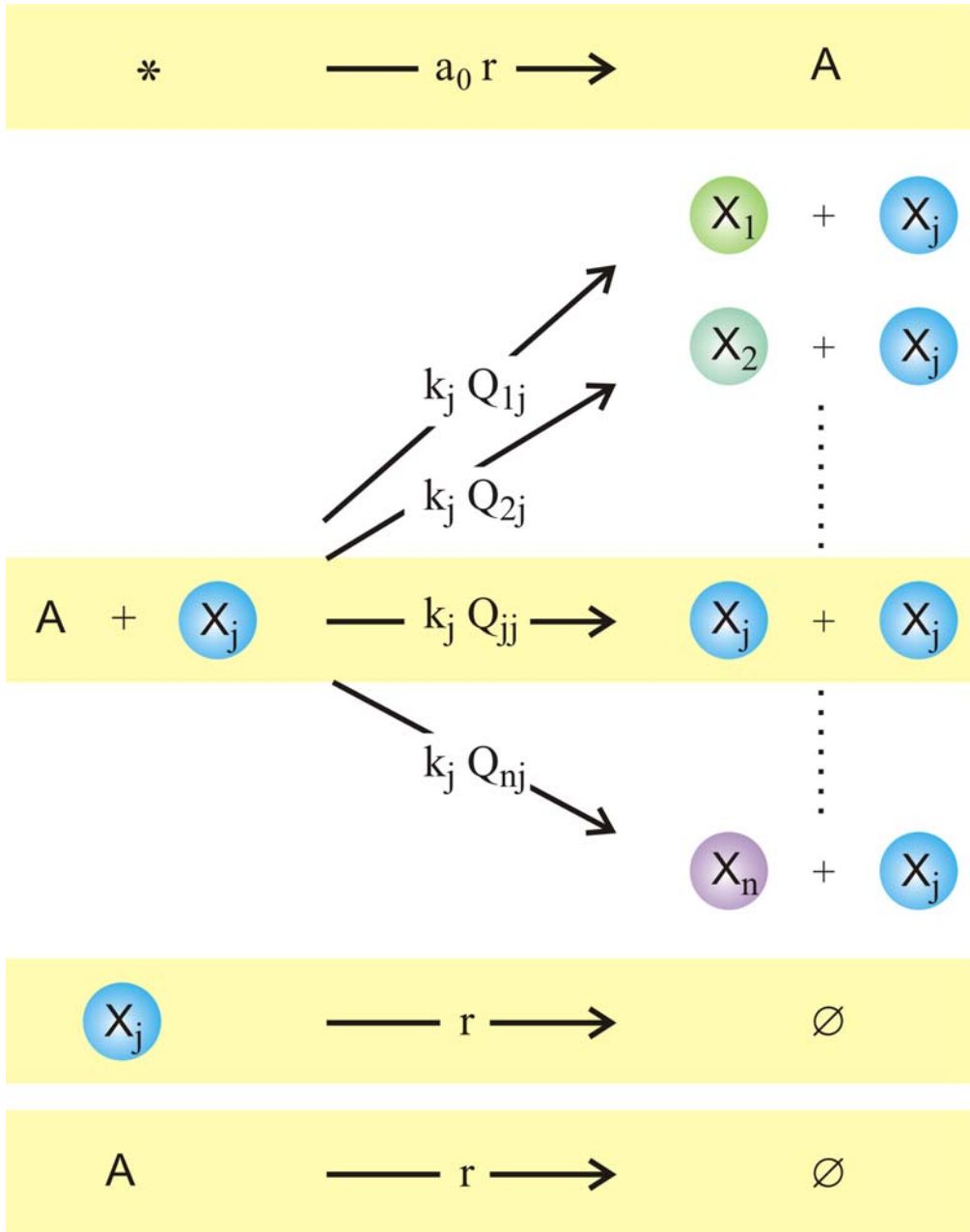
$$N(t) \approx \overline{N} \pm \sqrt{\overline{N}}$$

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





Variation of genotypes through mutation and recombination



$j = 1,2, \ldots ,n$

$$
\begin{aligned}
* &\xrightarrow{a_0 \cdot r} \mathbf{A} && : \text{influx} \\
\mathbf{A} + \mathbf{X}_i &\xrightarrow{k_i\, Q_{ii}} 2\,\mathbf{X}_i \;; \quad i = 1, \ldots, n && : \text{replication} \\
\mathbf{A} + \mathbf{X}_i &\xrightarrow{k_i\, Q_{ji}} \mathbf{X}_i + \mathbf{X}_j \quad i, j = 1, \ldots, n;\; i \neq j && : \text{mutation} \\
\mathbf{A} &\xrightarrow{r} \emptyset && : \text{outflux} \\
\mathbf{X}_j &\xrightarrow{r} \emptyset \;; \quad j = 1, \ldots, n && : \text{outflux}
\end{aligned}
$$

$$
\frac{da}{dt} \;=\; -\,a \sum_{i=1}^{n} \sum_{j=1}^{n} k_i Q_{ji}\, x_i \;+\; r\,(a_0 - a) \;=\; -\,a \sum_{i=1}^{n} k_i\, x_i \;+\; r\,(a_0 - a)
$$

$$
\frac{dx_j}{dt} \;=\; a \sum_{i=1}^{n} k_i Q_{ji}\, x_i \;-\; r\, x_j
$$

Origin of the replication-mutation equation from the flowreactor

Stationary solutions of the flow reactor:

$$\frac{da}{dt} = 0 = -\tilde{a}\left(\sum_{i=1}^{n} k_i\,\tilde{x}_i + r\right) + r\,\tilde{a}$$

$$\frac{dx_j}{dt} = 0 = \tilde{a}\sum_{i=1}^{n} k_i Q_{ji}\tilde{x}_i - r\,\tilde{x}_j\,;\; c = \sum_{i=1}^{n} x_i\,;\; \bar{k} = \frac{\sum_{i=1}^{n} k_i\,x_i}{c}$$

$$\frac{dc}{dt} = 0 = \tilde{c}\left(\bar{k}\,\tilde{a} - r\right)$$

Stationary solutions: 1. active state

$$r < \bar{k}\,a_0$$

$$\tilde{a} = \frac{r}{\bar{k}}$$

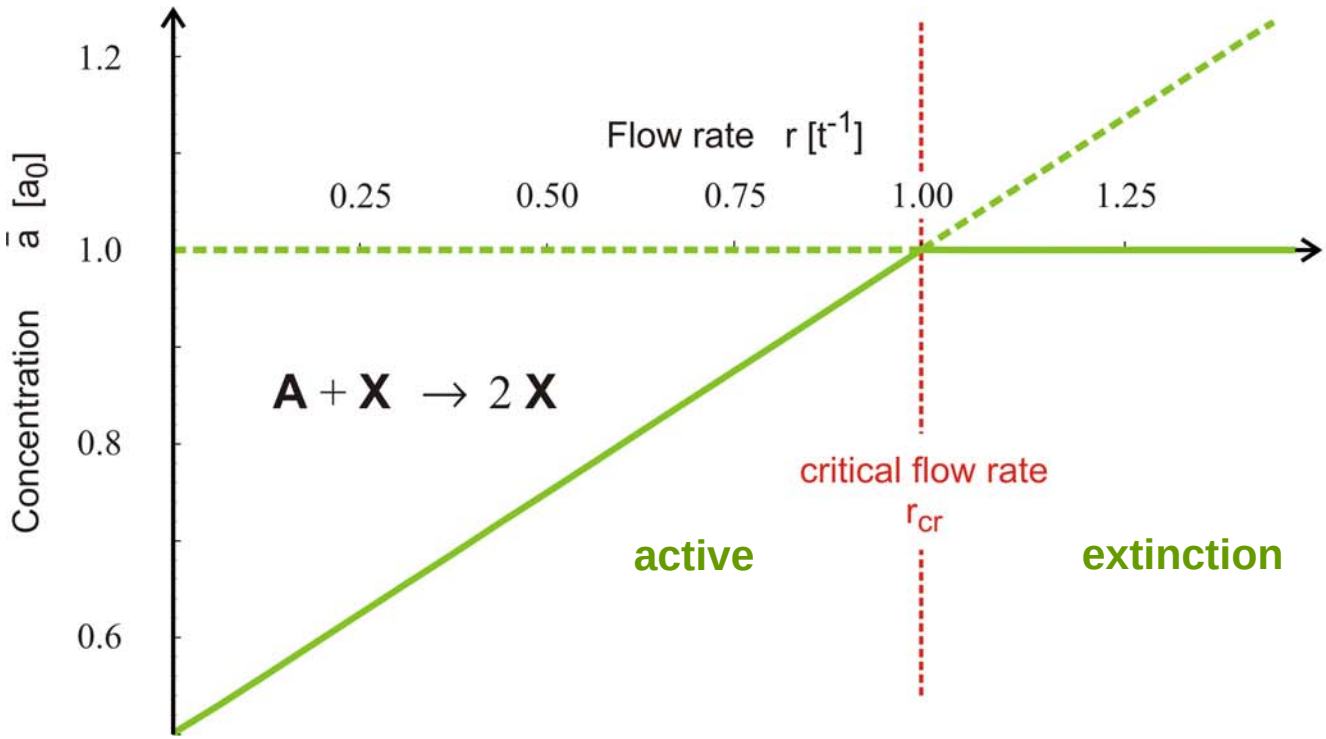
$$\tilde{c} = \frac{\bar{k}\,a_0 - r}{\bar{k}}$$

Stationary solutions: 2. extinction

$$r > \bar{k}\,a_0$$

$$\tilde{a} = a_0$$

$$\tilde{x}_j = 0\,;\; j = 1, 2, \ldots, n$$


1.2
1.0
0.8
0.6
Concentration $\bar{a}$ [a_0]
Flow rate r [t^{-1}]
0.25
0.50
0.75
1.00
1.25
A + X → 2 X
critical flow rate
r_{cr}
active
extinction

Find $r(t)$ such that $a(t) = \bar{a} = const.$

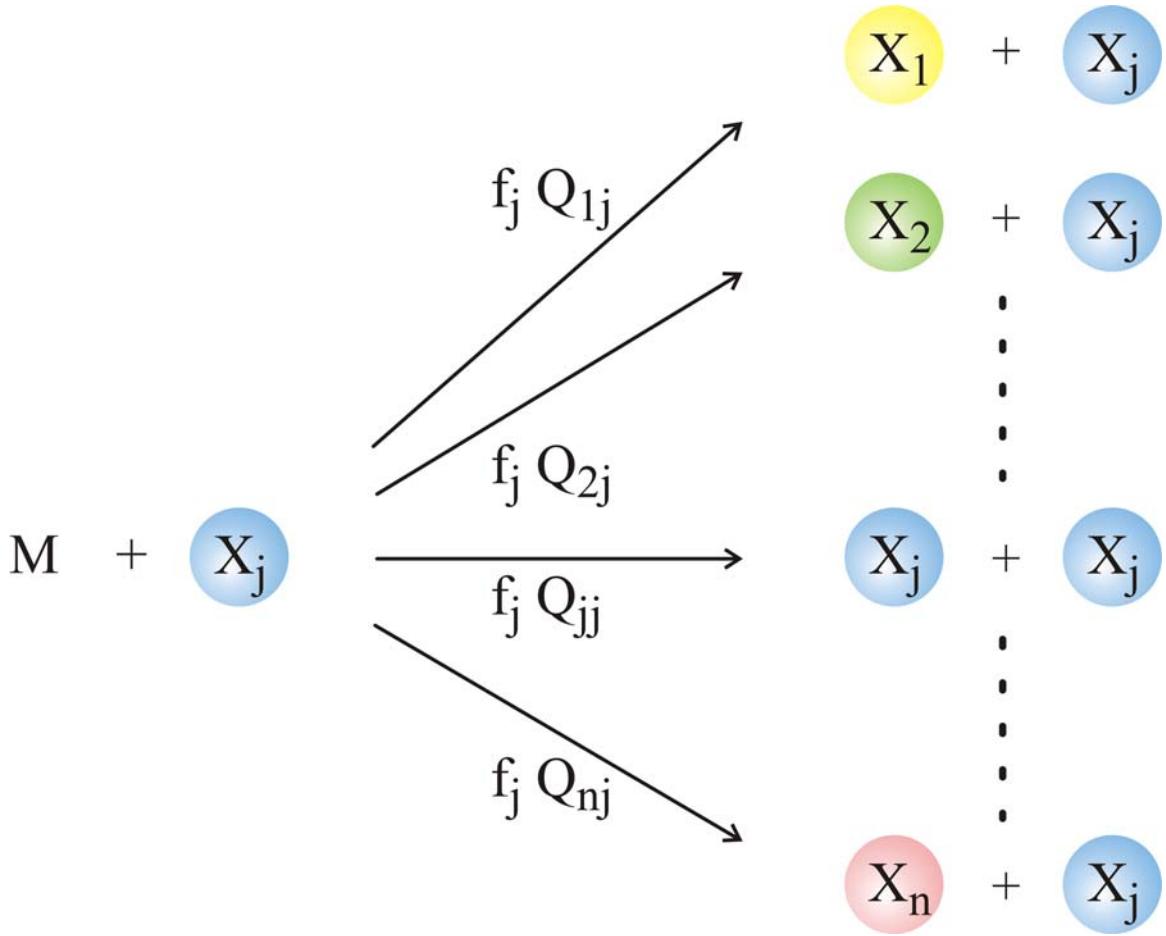
$$\frac{da}{dt} \;=\; 0 \;=\; -\,\bar{a}\sum_{i=1}^{n}\sum_{j=1}^{n} k_i Q_{ji}\, x_i \;+\; r(t)\,(a_0 - \bar{a})$$

$$r(t) \;=\; \frac{\bar{a}}{a_0 - \bar{a}}\sum_{i=1}^{n} k_i\, x_i\,;\;\; f_i \;=\; k_i\,\bar{a}$$

$$\frac{dx_j}{dt} \;=\; \sum_{i=1}^{n} f_i Q_{ji}\, x_i \;-\; x_j\,\frac{\sum_{i=1}^{n} f_i\, x_i}{\sum_{i=1}^{n} x_i} \;=\; \sum_{i=1}^{n} f_i Q_{ji}\, x_i \;-\; x_j\,\bar{f}$$

Origin of the replication-mutation equation from the flowreactor

1. Replication and mutation
2. **Quasispecies and error thresholds**
3. Fitness landscapes and randomization
4. Lethal mutations



Chemical kinetics of replication and mutation as parallel reactions

$$\frac{dx_i}{dt} = \sum_{j=1}^{n} Q_{ij} f_j x_j - x_i \Phi$$

with $\Phi = \sum_{i=1}^{n} f_i x_i$ and $\sum_{i=1}^{n} x_i = 1$

$$\sum_{i=1}^{n} Q_{ij} = 1$$

The replication-mutation equation

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

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

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

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

$$W \div \{f_i Q_{ij};\ i, j = 1,2,\cdots,n\};\ L = \{\ell_{ij};\ i, j = 1,2,\cdots,n\};\ L^{-1} = H = \{h_{ij};\ i, j = 1,2,\cdots,n\}$$

$$L^{-1} \cdot W \cdot L \quad = \quad \Lambda \quad = \quad \{\lambda_k; k = 0,1,\cdots,n-1\}$$

Matrix W and Frobenius theorem:

$$\mathrm{W} = \begin{pmatrix} w_{11} & w_{12} & \dots & w_{1n} \\ w_{21} & w_{22} & \dots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \dots & w_{nn} \end{pmatrix}$$

Primitive matrix W:

A nonnegative square matrix $\mathrm{W} = \{w_{ij}\}$ is said to be a primitive matrix if there exists k such that $\mathrm{W}^k \gg 0$, i.e., if there exists k such that for all i, j, the (i, j) entry of W^k is positive.

Perron-Frobenius theorem applied to the value matrix W

W is primitive: (i) λ_0 is real and strictly positive

(ii) $\lambda_0 > |\lambda_k|$ for all $k \neq 0$

(iii) λ_0 is associated with strictly positive eigenvectors

(iv) λ_0 is a simple root of the characteristic equation of W

(v-vi) etc.

W is irreducible: (i), (iii), (iv), etc. as above

(ii) $\lambda_0 \geq |\lambda_k|$ for all $k \neq 0$

Decomposition of matrix W

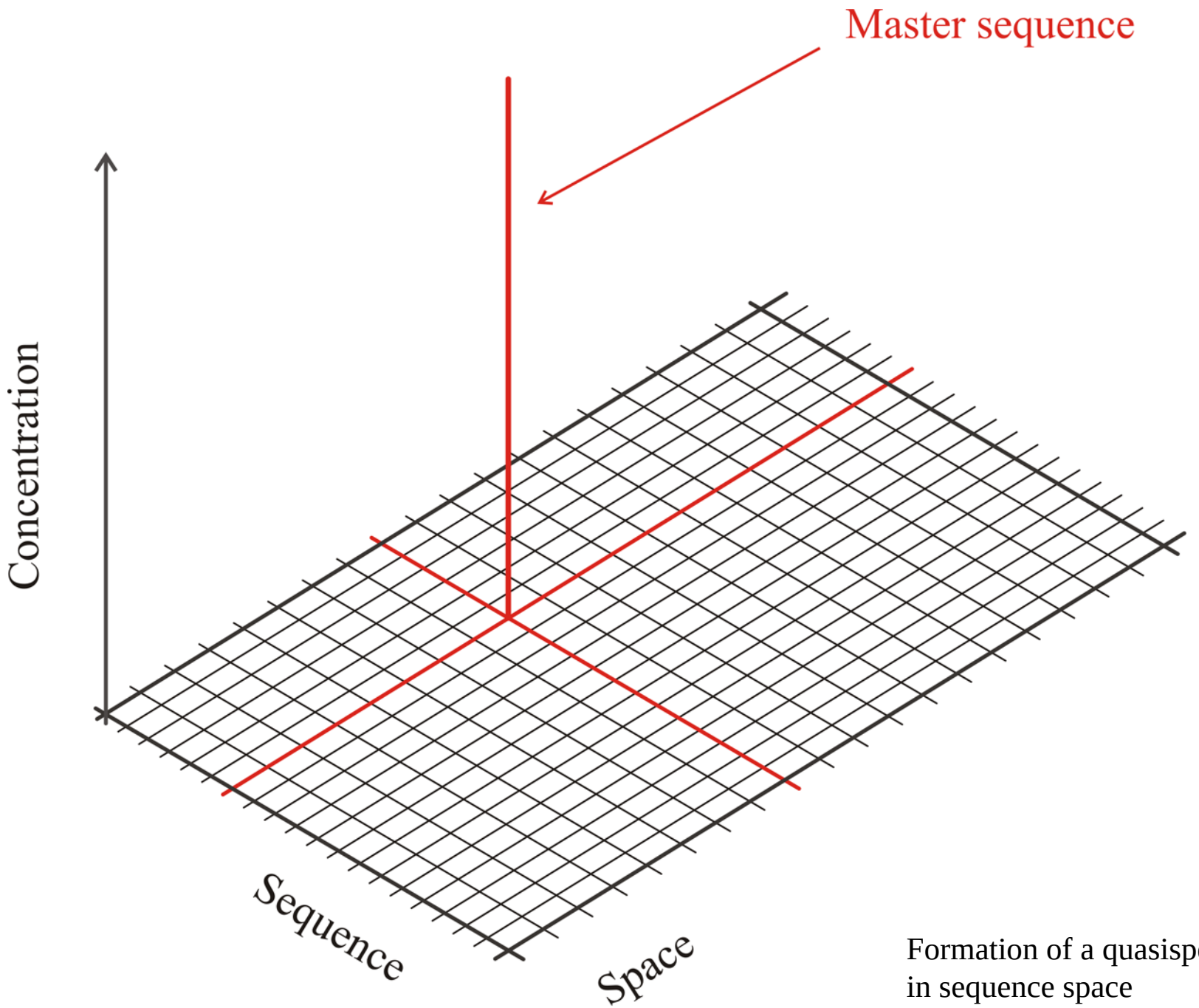
$$\mathrm{W} = \begin{pmatrix} w_{11} & w_{12} & \dots & w_{1n} \\ w_{21} & w_{22} & \dots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \dots & w_{nn} \end{pmatrix} = \mathrm{Q} \cdot \mathrm{F} \quad \text{with}$$

$$\mathrm{Q} = \begin{pmatrix} Q_{11} & Q_{12} & \dots & Q_{1n} \\ Q_{21} & Q_{22} & \dots & Q_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ Q_{n1} & Q_{n2} & \dots & Q_{nn} \end{pmatrix} \quad \text{and} \quad \mathrm{F} = \begin{pmatrix} f_1 & 0 & \dots & 0 \\ 0 & f_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & f_n \end{pmatrix}$$

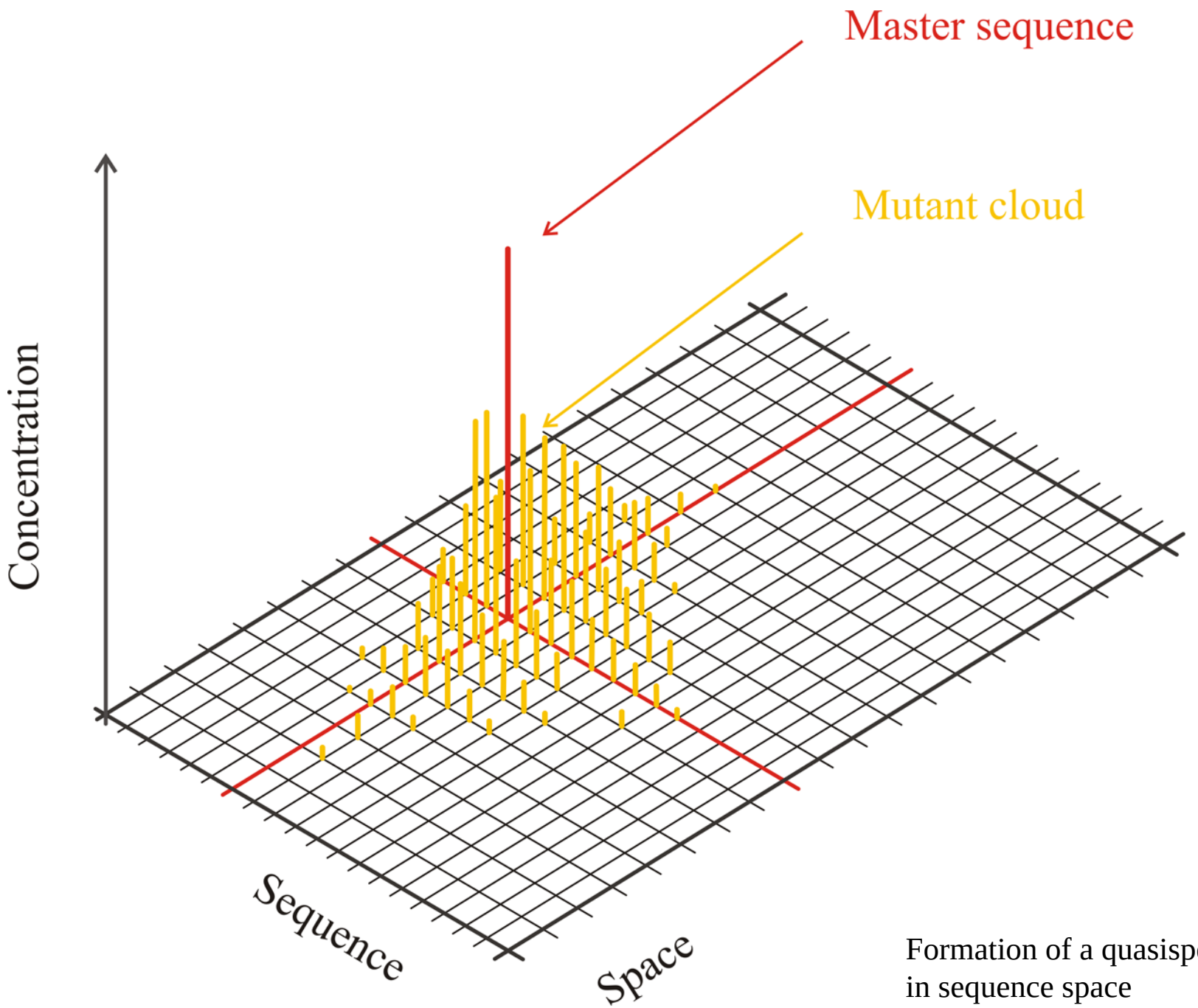
Uniform error rate model:

$$Q_{ij} \;=\; p^{d_H(\mathbf{X}_i,\mathbf{X}_j)}\,(1-p)^{\left(n-d_H(\mathbf{X}_i,\mathbf{X}_j)\right)}$$

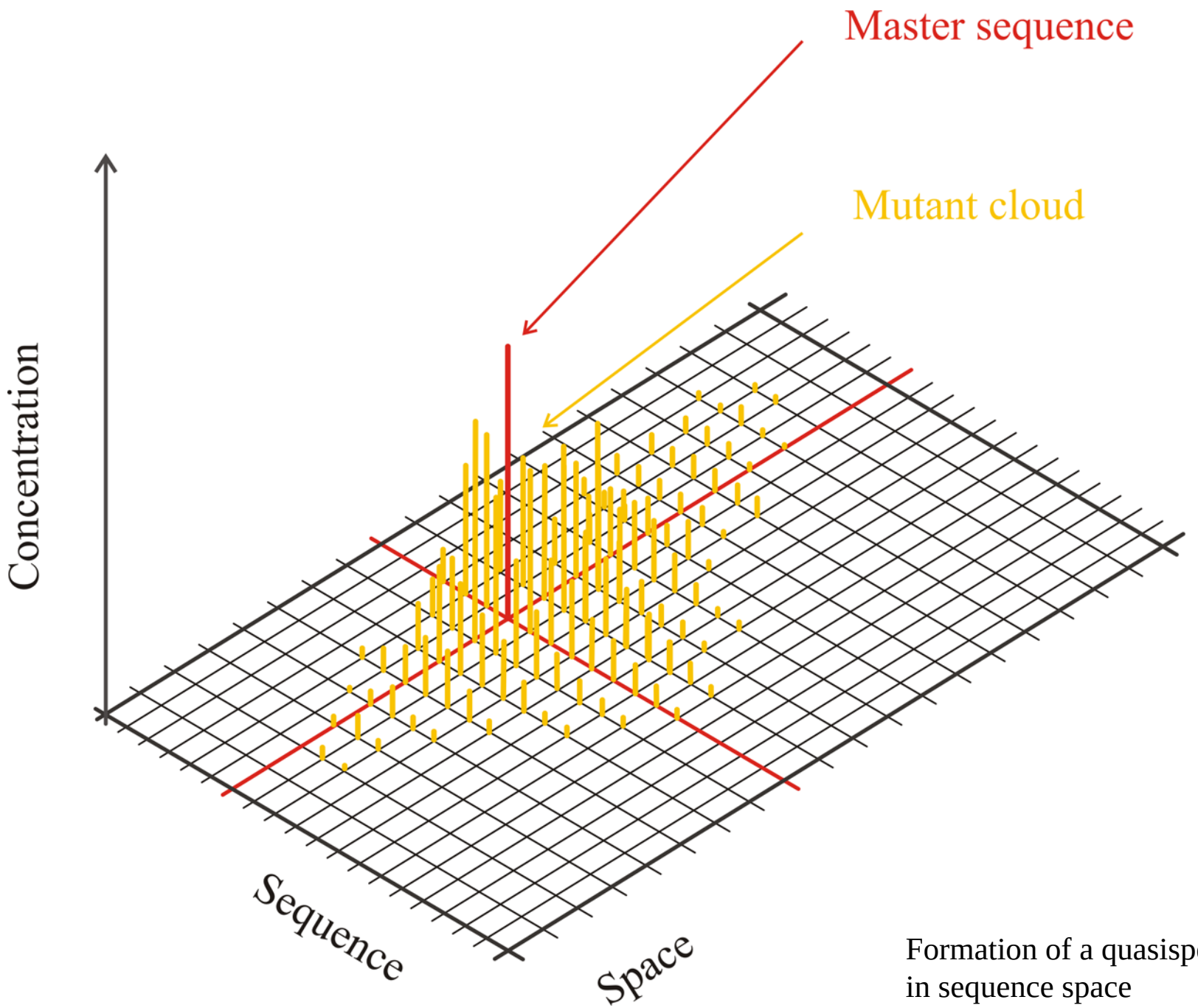
$d_H(\mathbf{X}_i, \mathbf{X}_j)$ ... Hamming distance



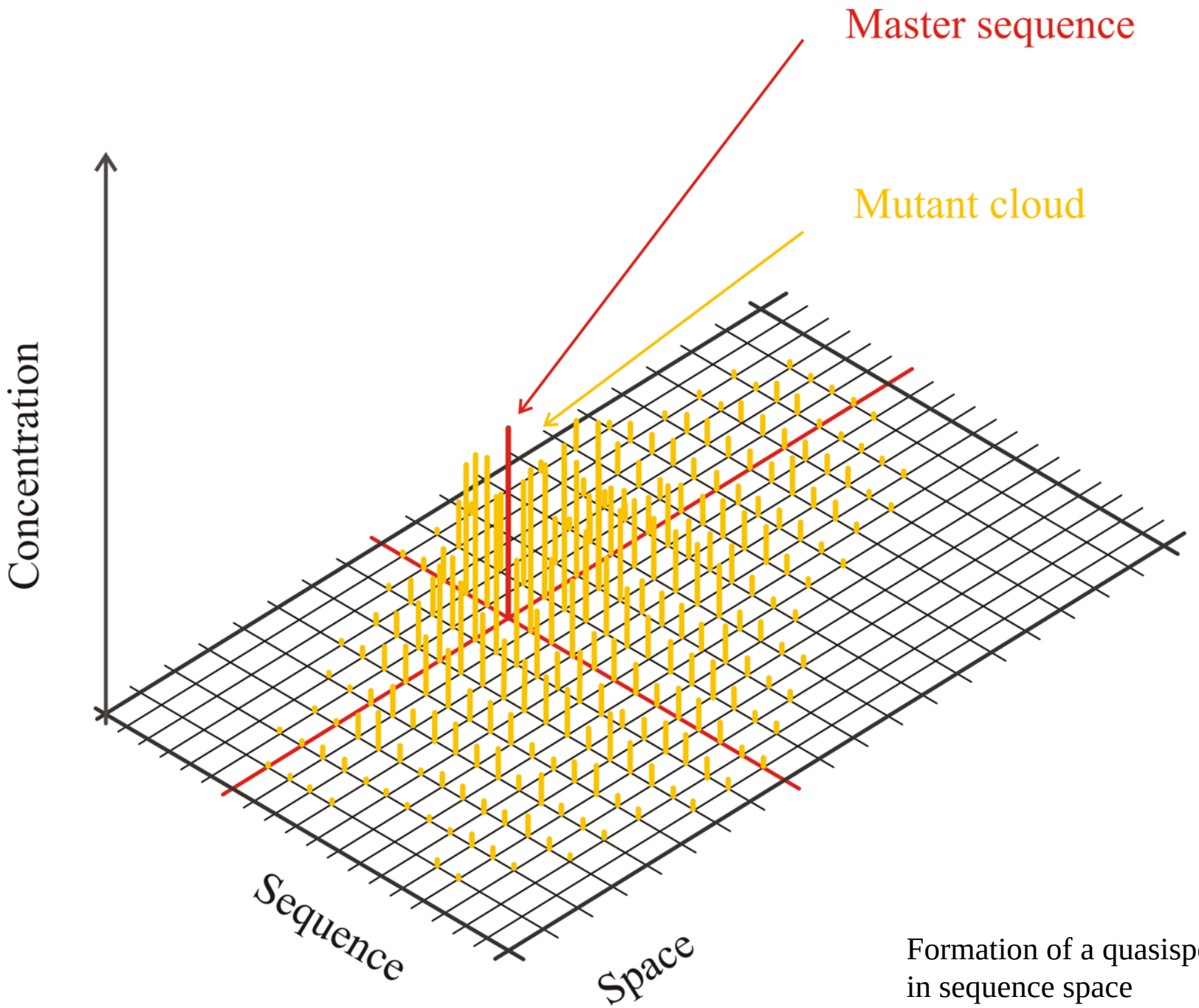
Formation of a quasispecies in sequence space



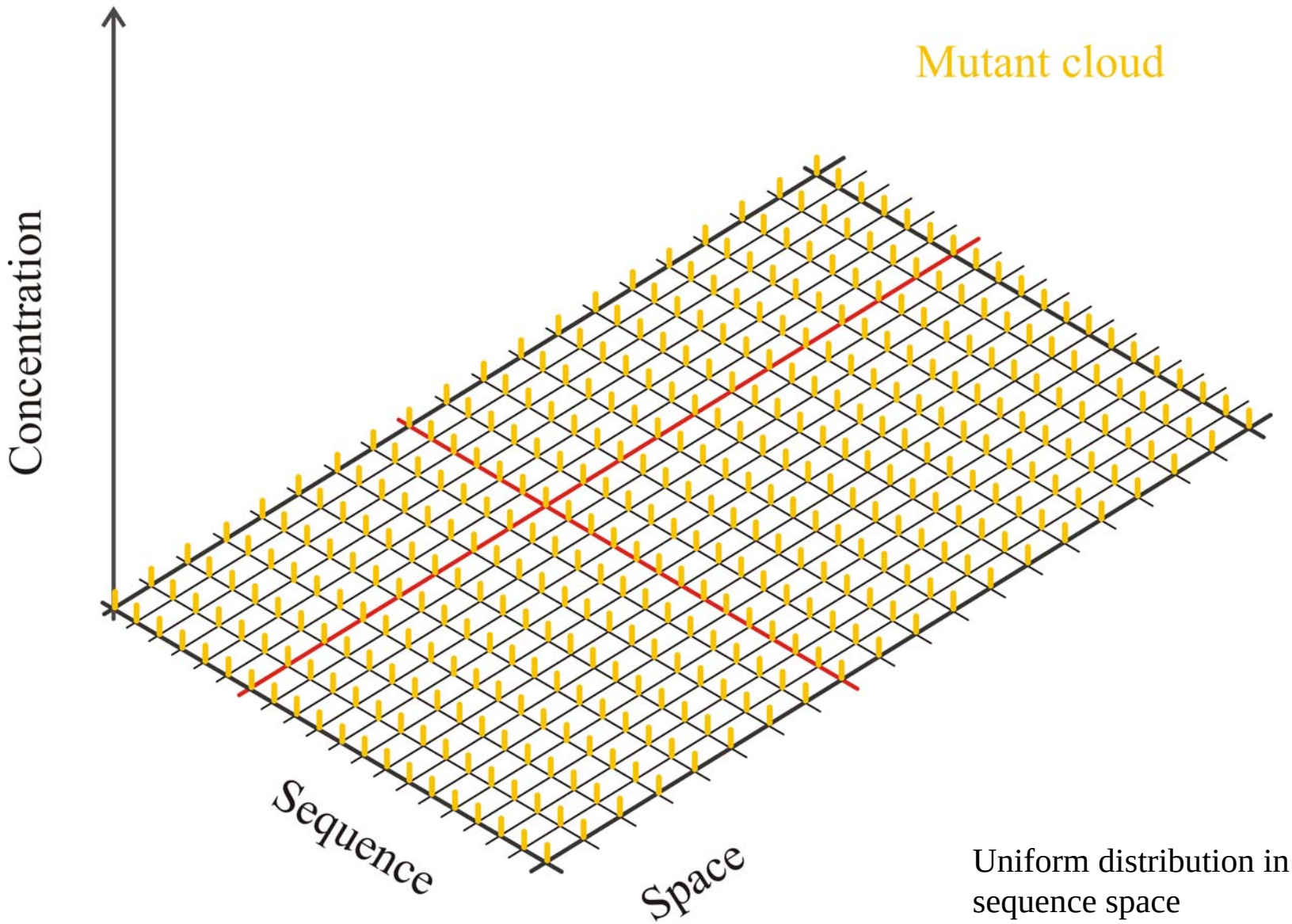
Formation of a quasispecies in sequence space



Formation of a quasispecies in sequence space



Formation of a quasispecies in sequence space



Uniform distribution in sequence space

Biophysical Chemistry 16 (1982) 329–345
Elsevier Biomedical Press

SELF-REPLICATION WITH ERRORS

A MODEL FOR POLYNUCLEOTIDE REPLICATION **

Jörg SWETINA and Peter SCHUSTER *
Institut für Theoretische Chemie und Strahlenchemie der Universität, Währingerstraße 17, A-1090 Wien, Austria

Received 4th June 1982
Revised manuscript received 23rd August 1982
Accepted 30th August 1982

Key words: Polynucleotide replication; Quasi-species; Point mutation; Mutant class; Stochastic replication

A model for polynucleotide replication is presented and analyzed by means of perturbation theory. Two basic assumptions allow handling of sequences up to a chain length of $\nu \approx 80$ explicitly: point mutations are restricted to a two-digit model and individual sequences are subsumed into mutant classes. Perturbation theory is in excellent agreement with the exact results for long enough sequences ($\nu > 20$).

1. Introduction

Eigen [8] proposed a formal kinetic equation (eq. 1) which describes self-replication under the constraint of constant total population size:

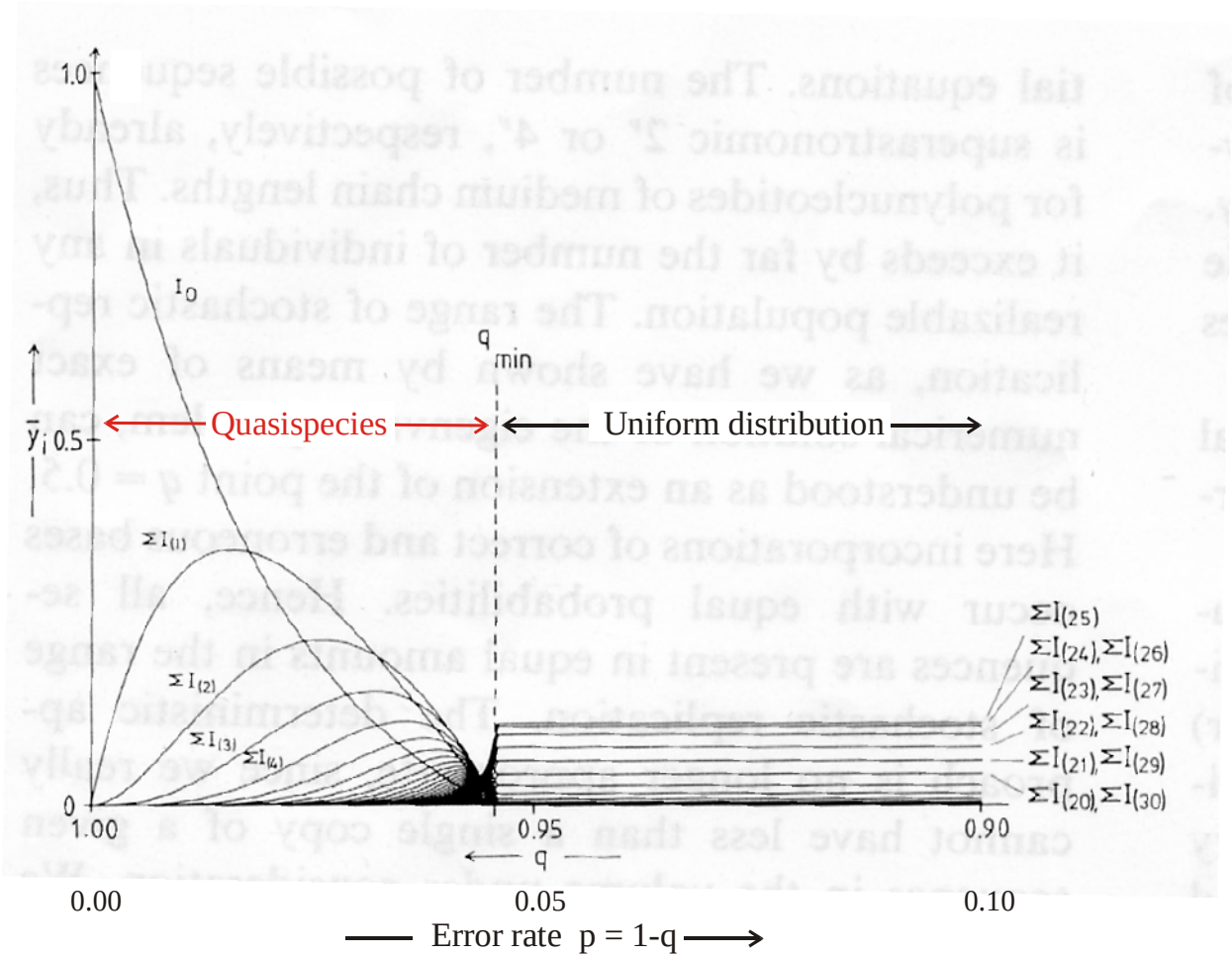
$$\frac{dx_i}{dt} = \dot{x}_i = \sum_j w_{ij} x_j - \frac{x_i}{c}\phi; \; i = 1, \ldots, n \; ^{\dagger} \qquad (1)$$

By x_i we denote the population number or concentration of the self-replicating element I_i, i.e., $x_i = [I_i]$. The total population size or total concentration $c = \sum_i x_i$ is kept constant by proper adjustment of the constraint ϕ: $\phi = \sum_i \sum_j w_{ij} x_j$. Characteristically, this constraint has been called 'constant organization'. The relative values of diagonal (w_{ii}) and off-diagonal (w_{ij}, $i \neq j$) rates, as we shall see in detail in section 2, are related to the accuracy of the replication process. The specific properties of eq. 1 are essentially based on the fact that it leads to exponential growth in the absence of constraints ($\phi = 0$) and competitors ($n = 1$).

The non-linear differential equation, eq. 1 – the non-linearity is introduced by the definition of ϕ at constant organization – shows a remarkable feature: it leads to selection of a defined ensemble of self-replicating elements above a certain accuracy threshold. This ensemble of a master and its most frequent mutants is a so-called 'quasi-species' [9]. Below this threshold, however, no selection takes place and the frequencies of the individual elements are determined exclusively by their statistical weights.

Rigorous mathematical analysis has been performed on eq. 1 [7,15,24,26]. In particular, it was shown that the non-linearity of eq. 1 can be removed by an appropriate transformation. The eigenvalue problem of the linear differential equation obtained thereby may be solved approximately by the conventional perturbation technique

* Dedicated to the late Professor B.L. Jones who was among the first to do rigorous mathematical analysis on the problems described here.
** This paper is considered as part II of Model Studies on RNA replication. Part I is by Gassner and Schuster [14].
† All summations throughout this paper run from 1 to n unless specified differently: $\sum_i = \sum_{i=1}^{n}$ and $\sum_{i,i\neq j} = \sum_{i=1}^{j-1} + \sum_{i=j+1}^{n}$, respectively.



Quasispecies as a function of the replication accuracy q

Chain length and error threshold

$$Q \cdot \sigma \quad = \quad (1-p)^n \cdot \sigma \quad \geq \quad 1 \quad \Rightarrow \quad n \cdot \ln(1-p) \geq -\ln\sigma$$

$$n \ \ldots \ \text{constant} : \quad p_{\max} \approx \frac{\ln\sigma}{n}$$

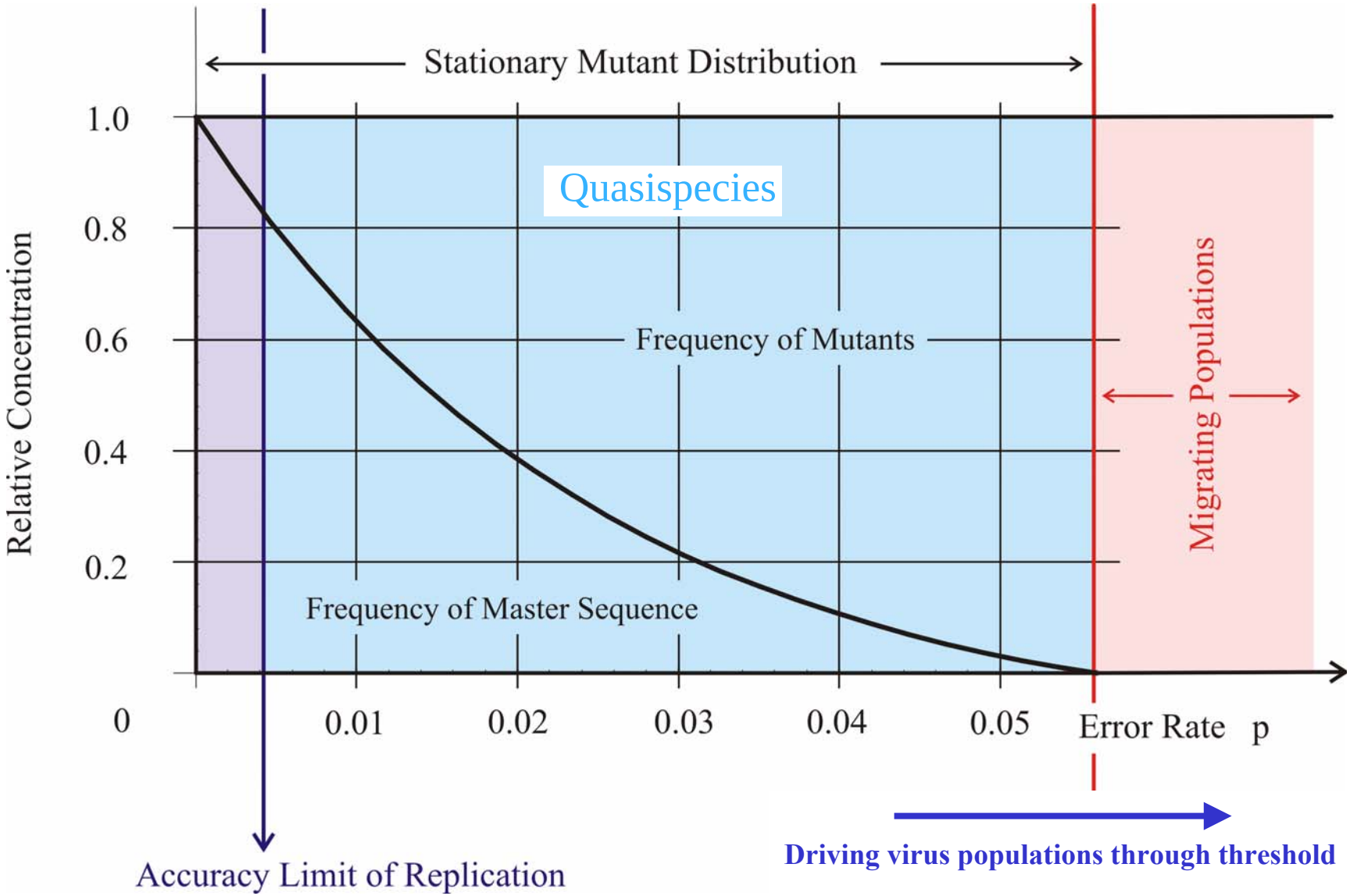
$$p \ \ldots \ \text{constant} : \quad n_{\max} \approx \frac{\ln\sigma}{p}$$

$Q = (1-p)^n$ … replication accuracy

p … error rate

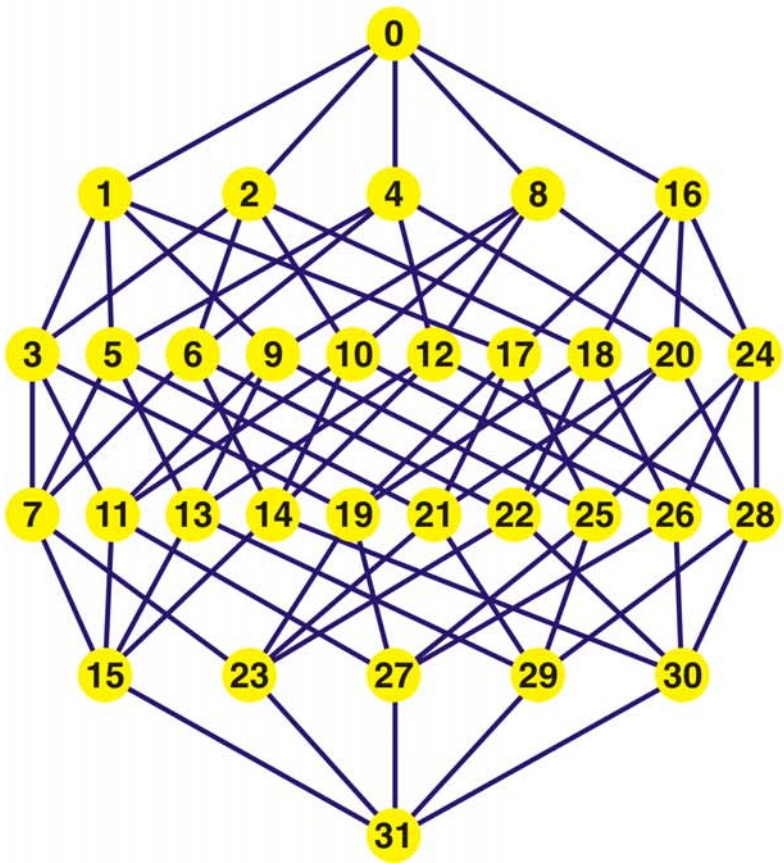
n … chain length

$\sigma = \dfrac{f_m}{(1-x_m)\sum_{j \neq m} f_j}$ … superiority of master sequence



The error threshold in replication

1. Replication and mutation
2. Quasispecies and error thresholds
3. **Fitness landscapes and randomization**
4. Lethal mutations



Binary sequences can be encoded by their decimal equivalents:

C = 0 and **G = 1**, for example,

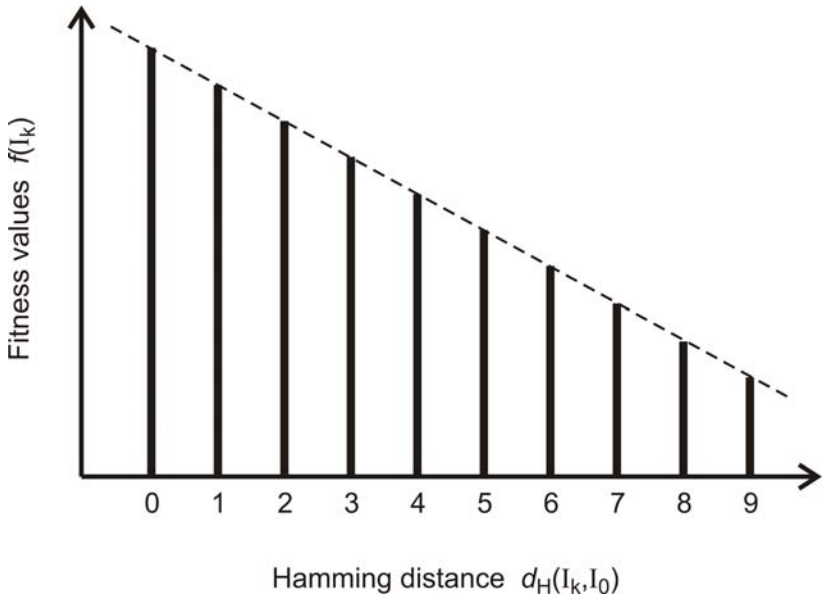
"0" ≡ 00000 = CCCCC,

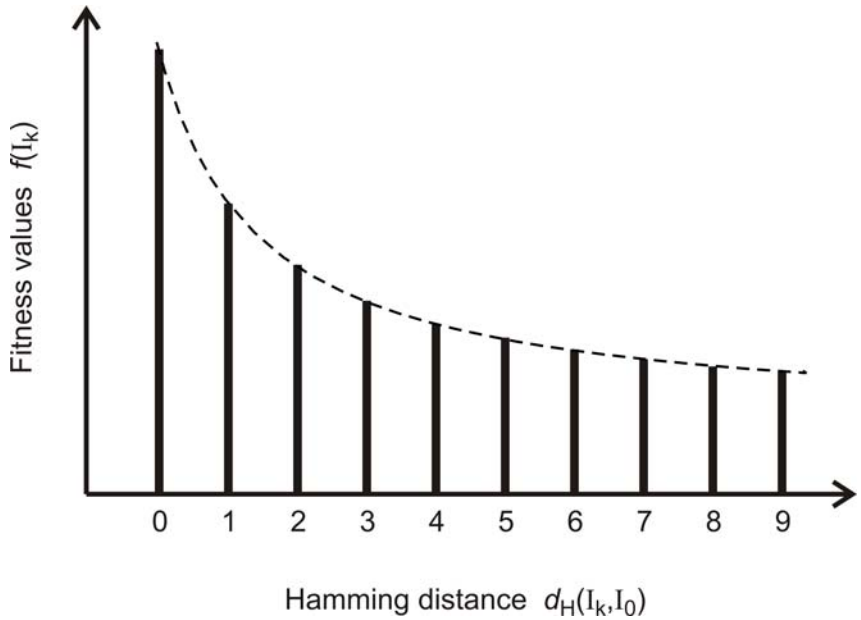
"14" ≡ 01110 = CGGGC,

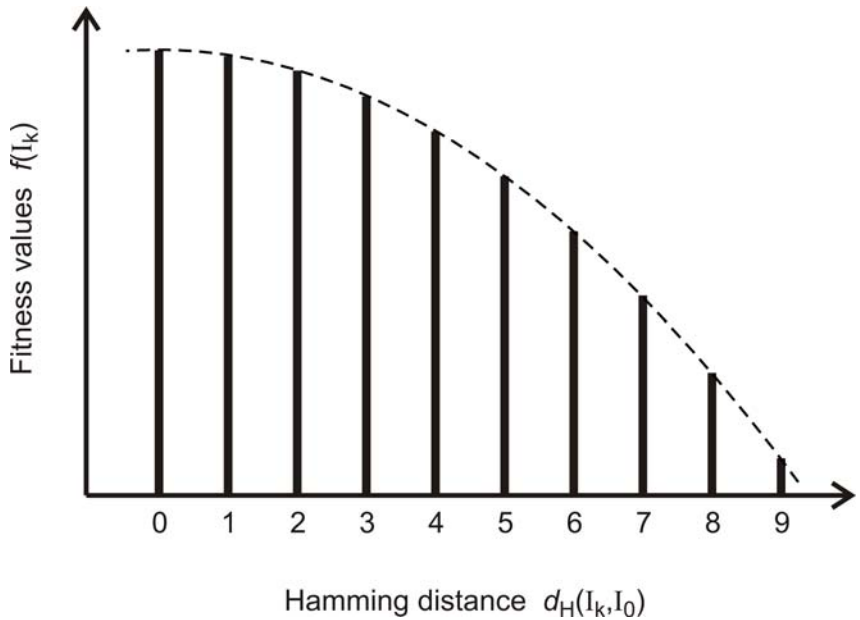
"29" ≡ 11101 = GGGCG, etc.

Every point in sequence space is equivalent

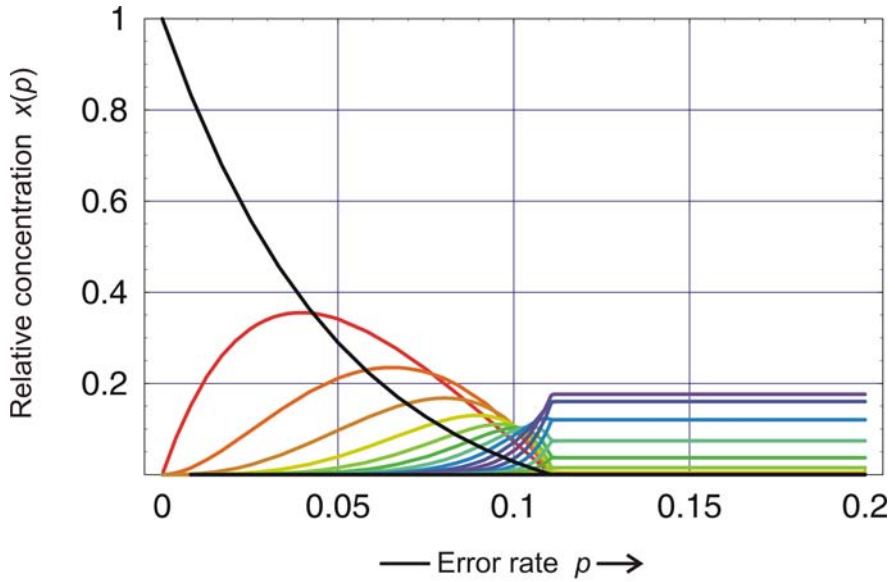
Sequence space of binary sequences with chain length $n = 5$

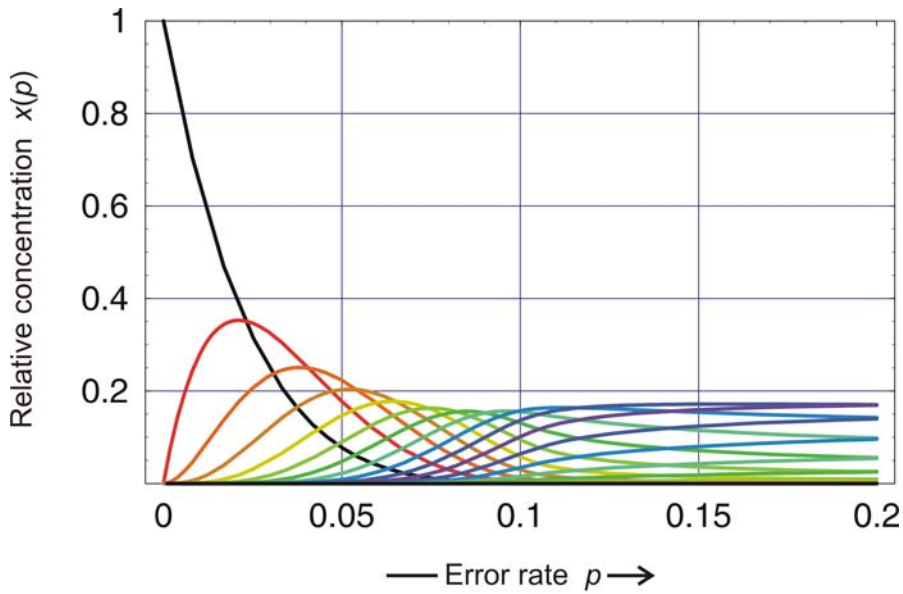






Fitness landscapes **not** showing error thresholds





Error thresholds and gradual transitions

n = 20 and $\sigma = 10$

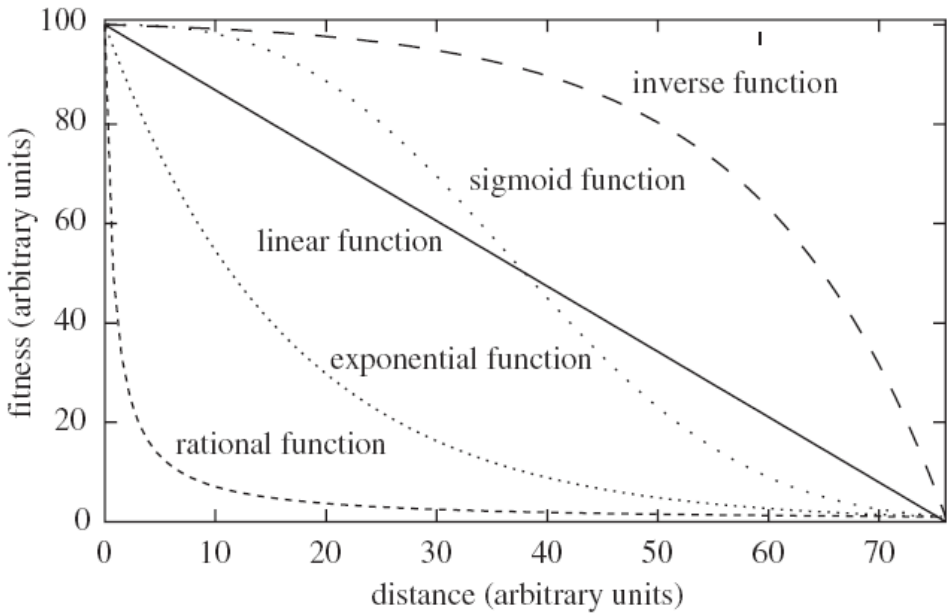
(1) linear $f^1_{scale}(d) = 100(1 - d/l)$,

(2) exponential $f^2_{scale}(d) = 100^{1-d/l}$,

(3) rational $f^3_{scale}(d) = \dfrac{1}{0.01 + d/l}$,

(4) sigmoid $f^4_{scale}(d) = 100^{1-(d/l)^\sigma}$,

(5) inverse $f^5_{scale}(d) = 100 - 100^{d/l} + 1$.



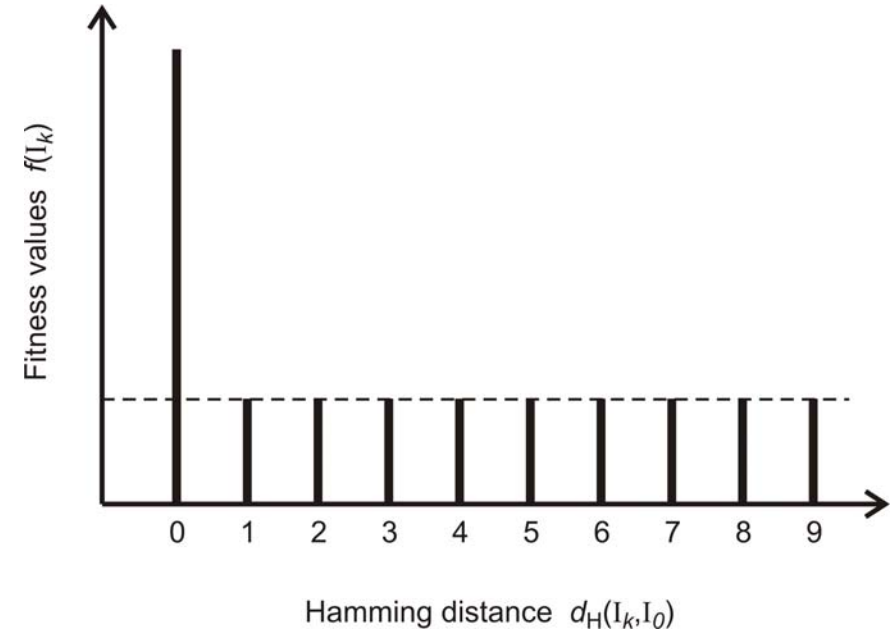
Anne Kupczok, Peter Dittrich, Determinats of simulated RNA evolution.
J.Theor.Biol. **238**:726-735, 2006

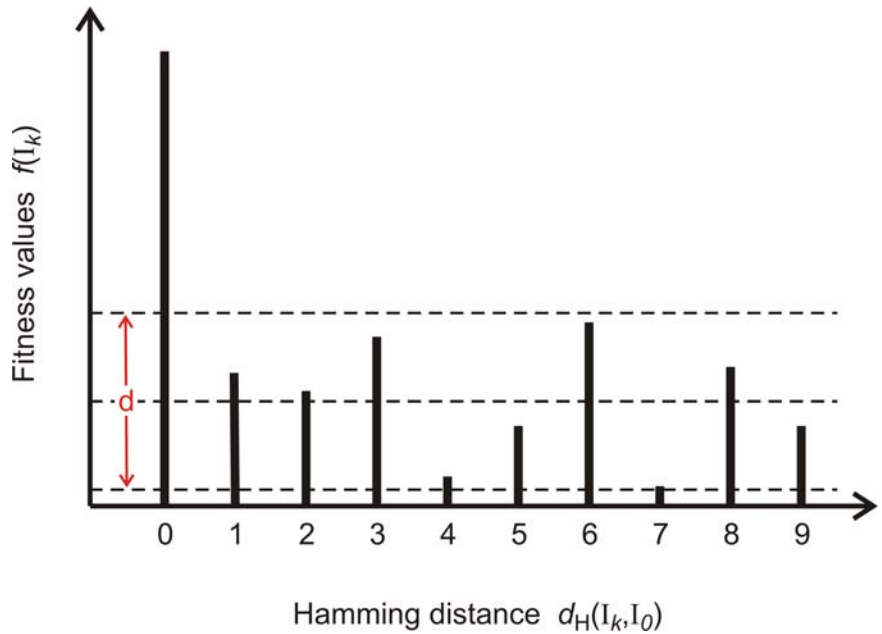
Three sources of ruggedness:

1. Variation in fitness values
2. Deviations from uniform error rates
3. Neutrality

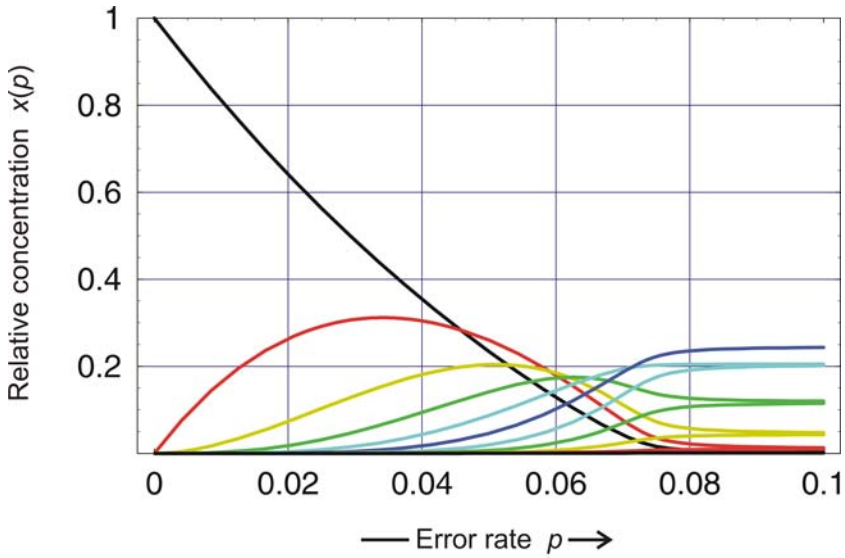
Three sources of ruggedness:

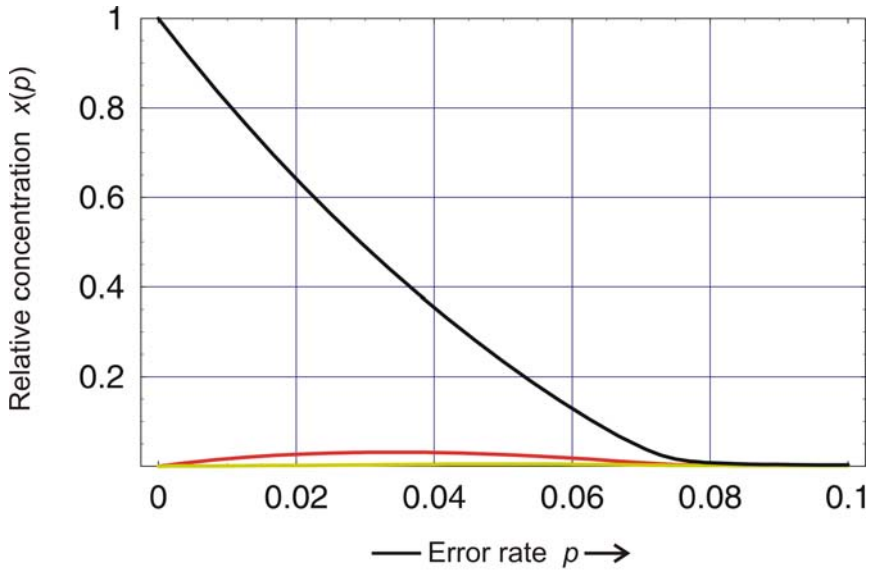
1. **Variation in fitness values**
2. Deviations from uniform error rates
3. Neutrality

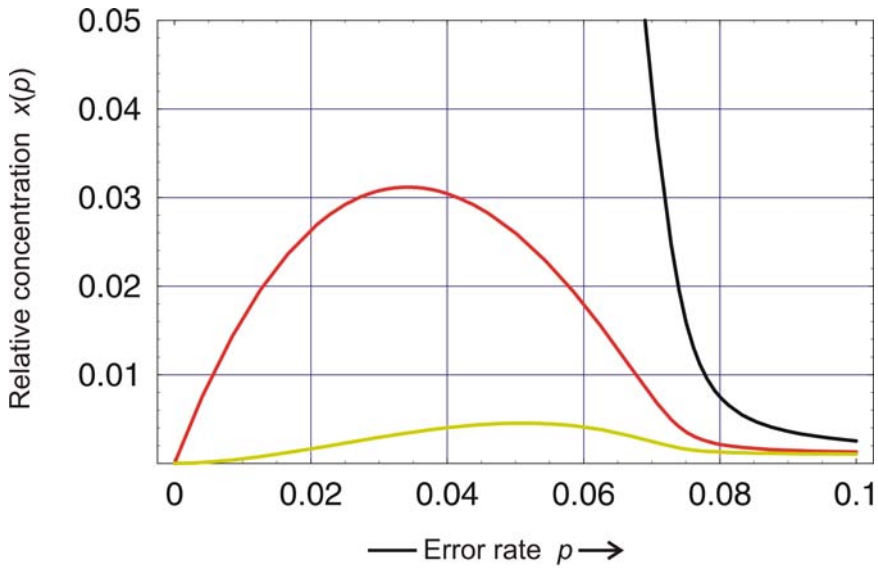




Fitness landscapes showing error thresholds

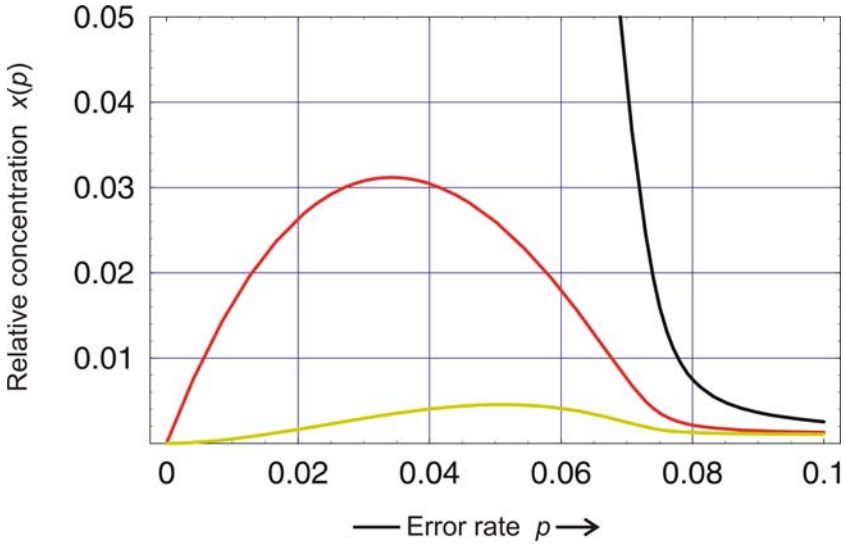


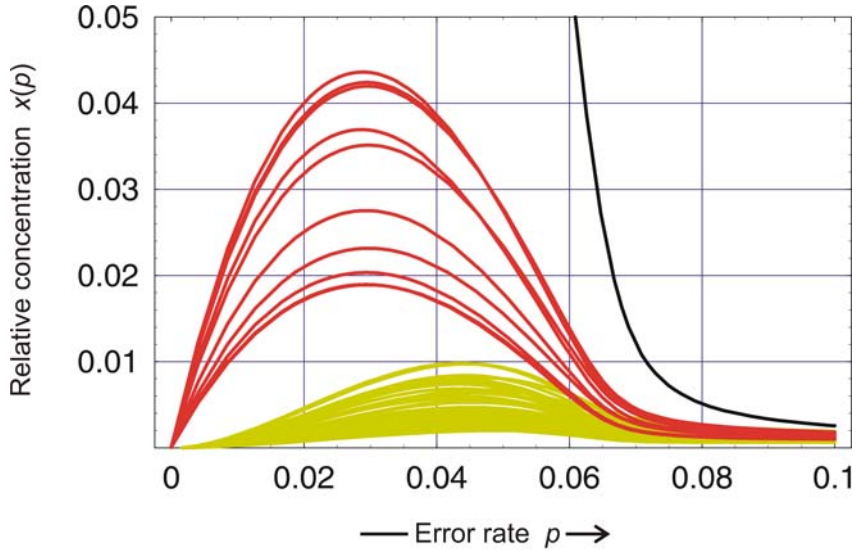


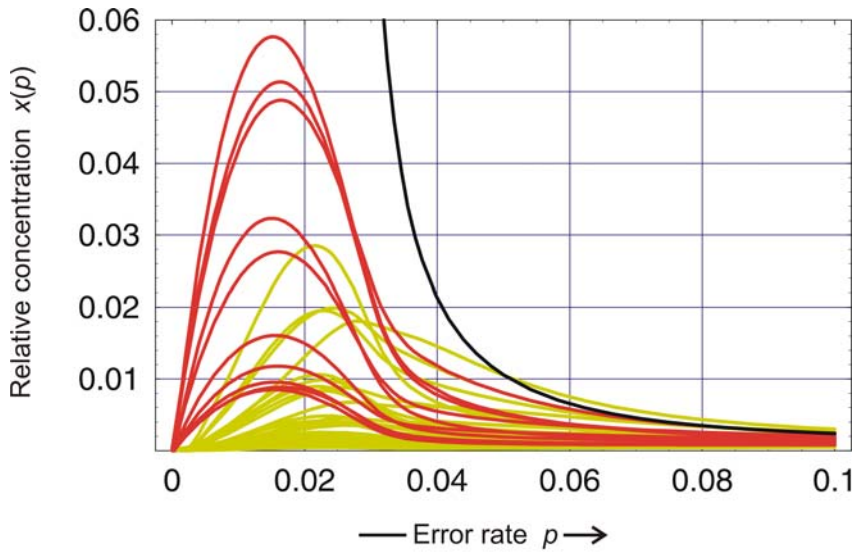


Error threshold: Error classes and individual sequences

n = 10 and $\sigma = 2$

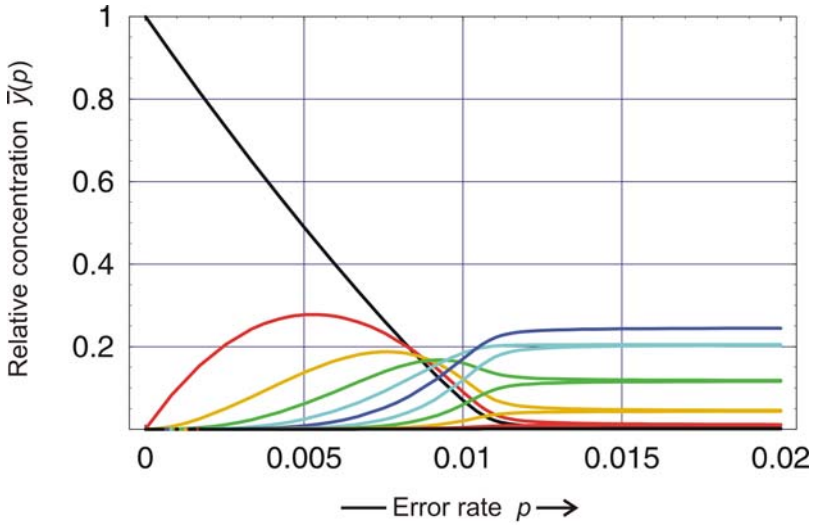


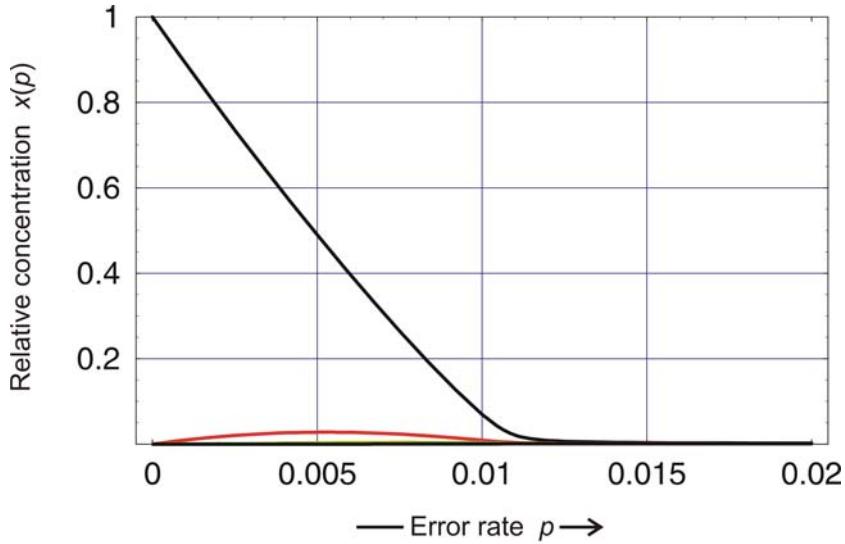


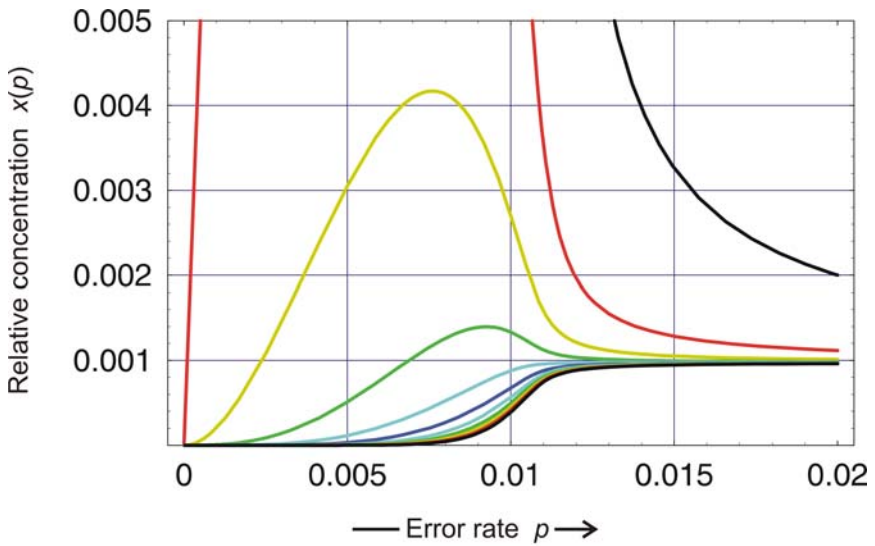


Error threshold: Individual sequences

n = 10, σ = 2 and d = 0, 1.0, 1.85

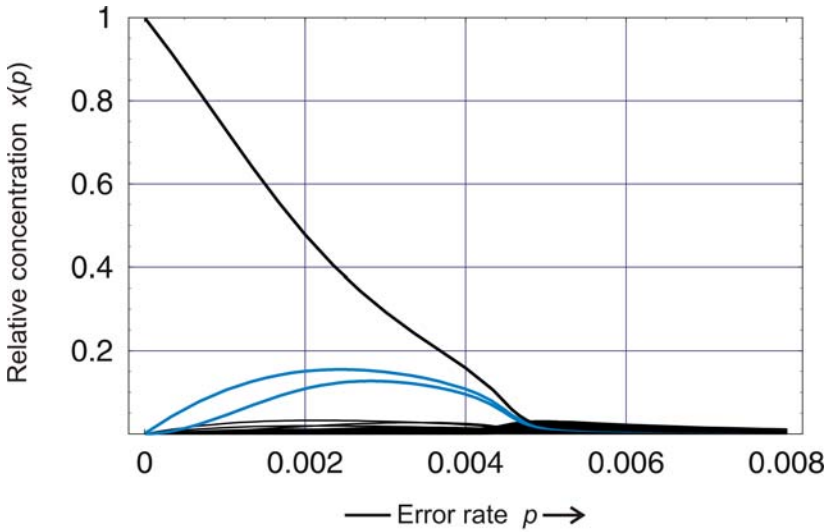


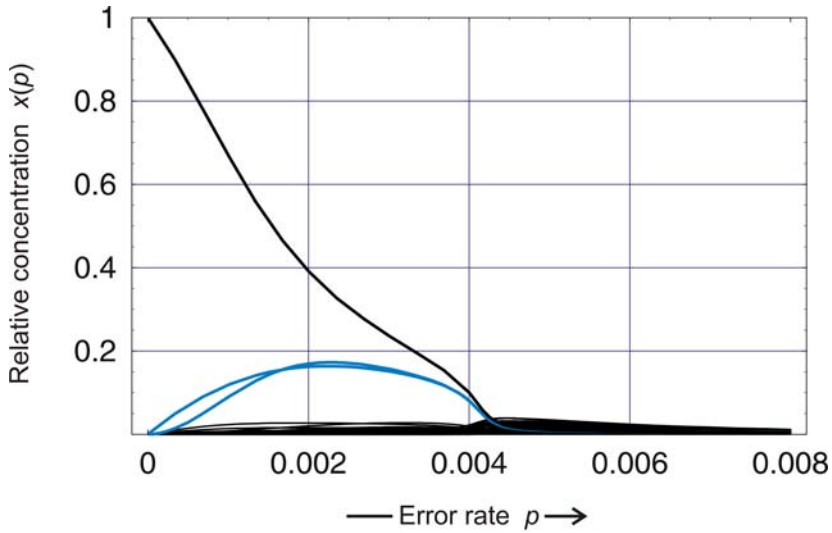


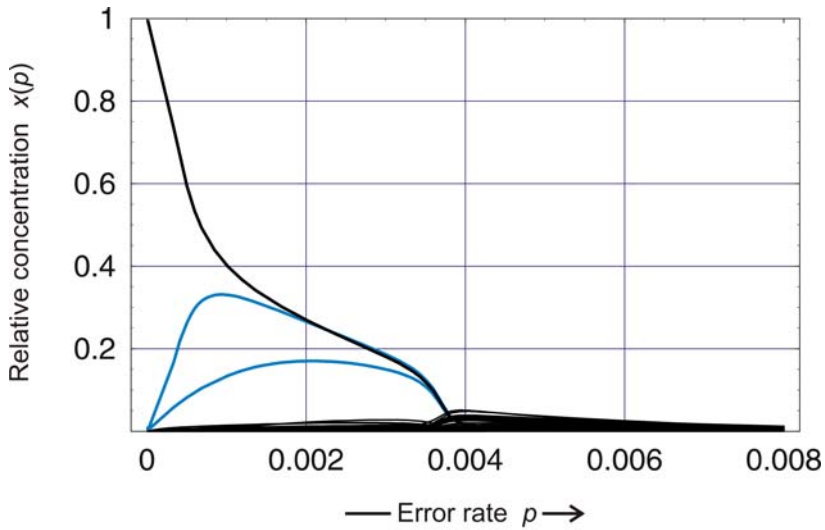


Error threshold: Error classes and individual sequences

n = 10 and σ = 1.1

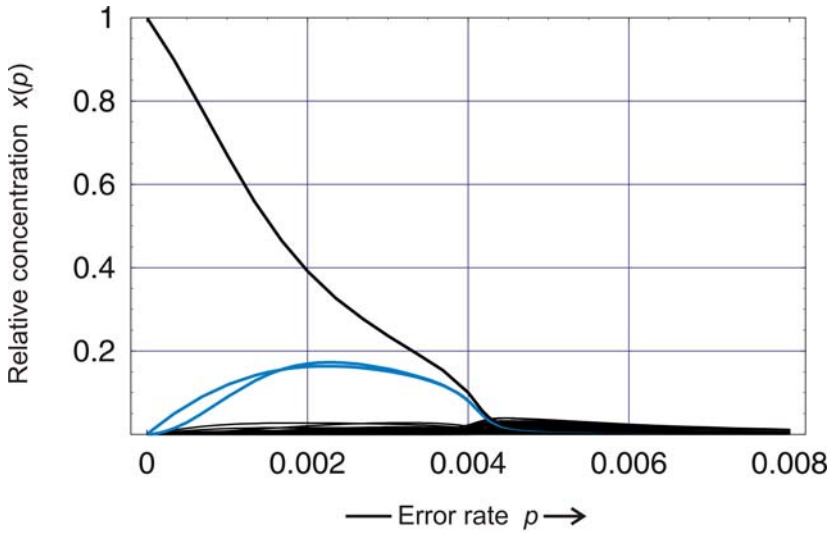


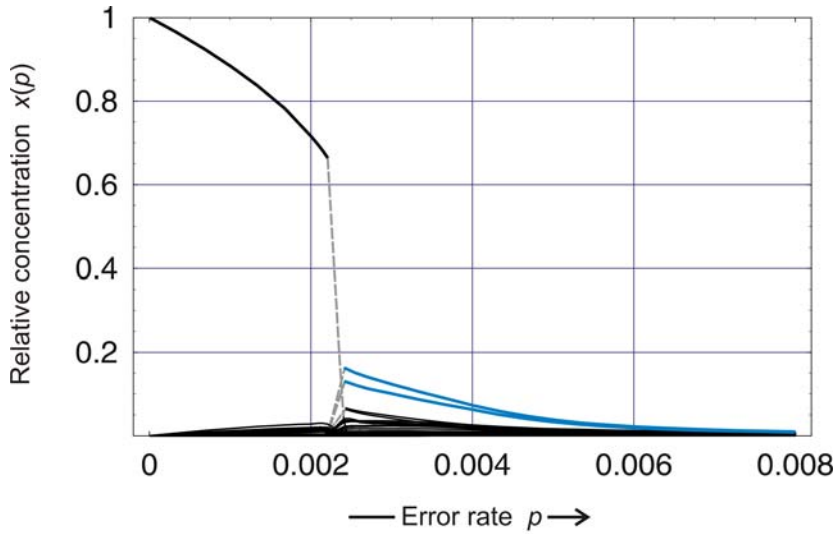


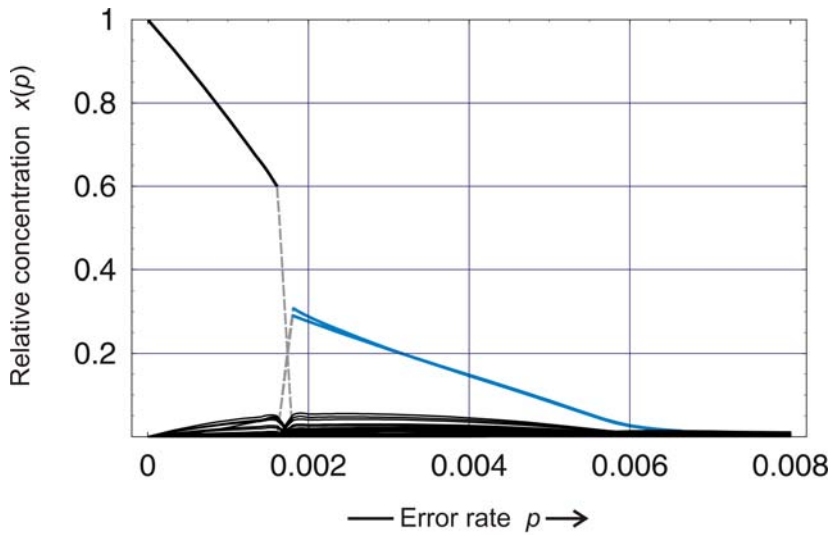


Error threshold: Individual sequences

n = 10, σ = 1.1, d = 1.95, 1.975, 2.00 and seed = 877





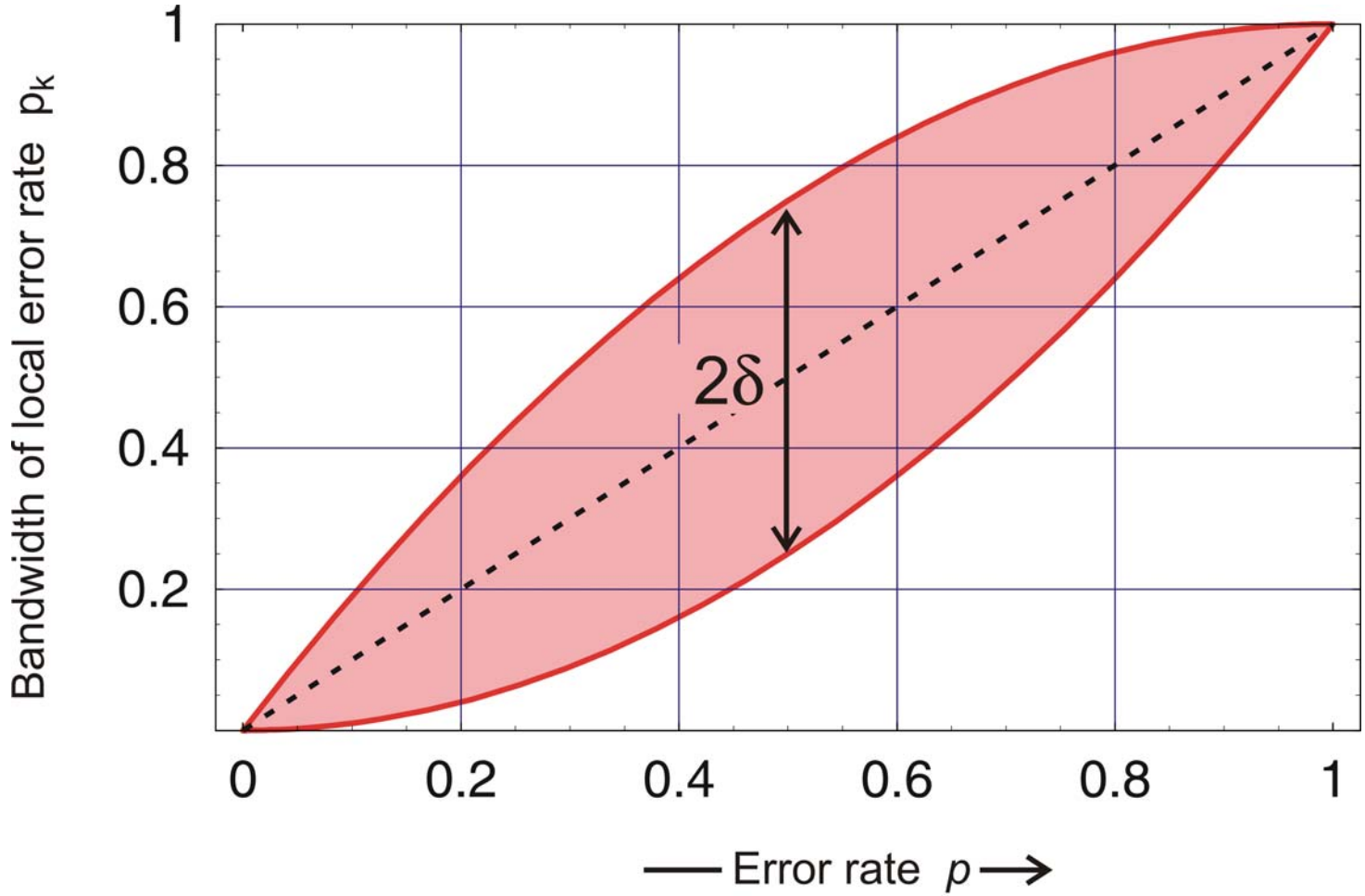


Error threshold: Individual sequences

n = 10, σ = 1.1, d = 1.975, and seed = 877, 637, 491

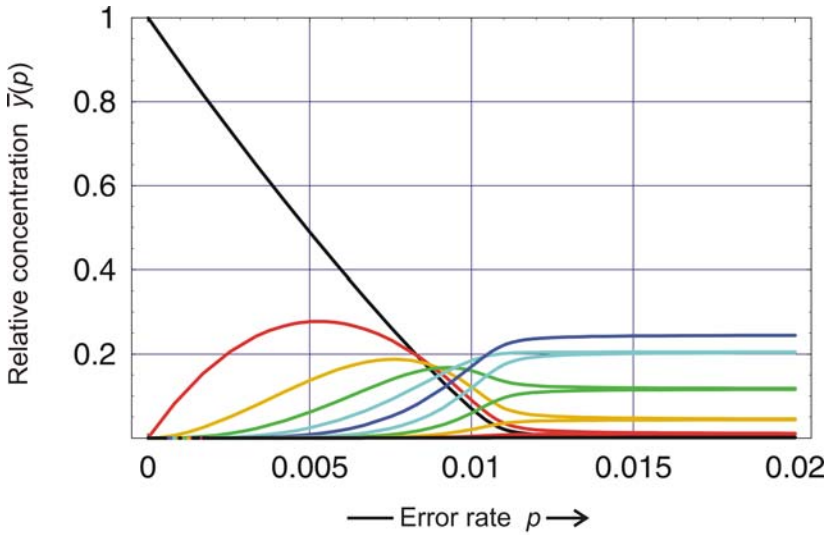
Three sources of ruggedness:

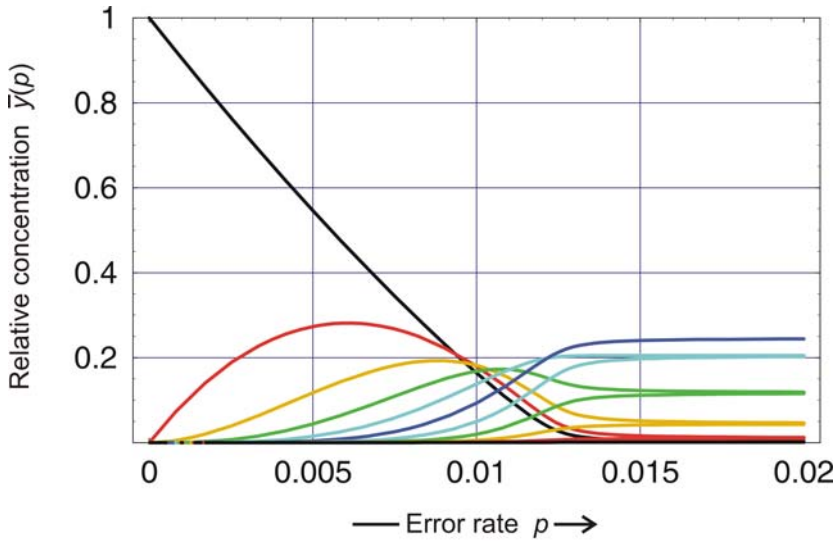
1. Variation in fitness values
2. **Deviations from uniform error rates**
3. Neutrality

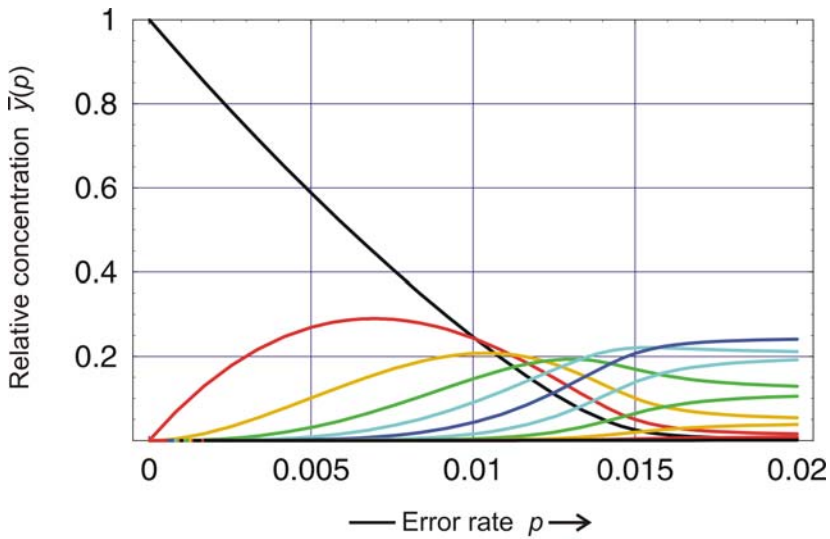


Local replication accuracy p_k:

$$p_k = p + 4\,\delta\,p(1-p)\,(X_{rnd}-0.5)\,,\quad k = 1,2,\ldots,2^{\nu}$$

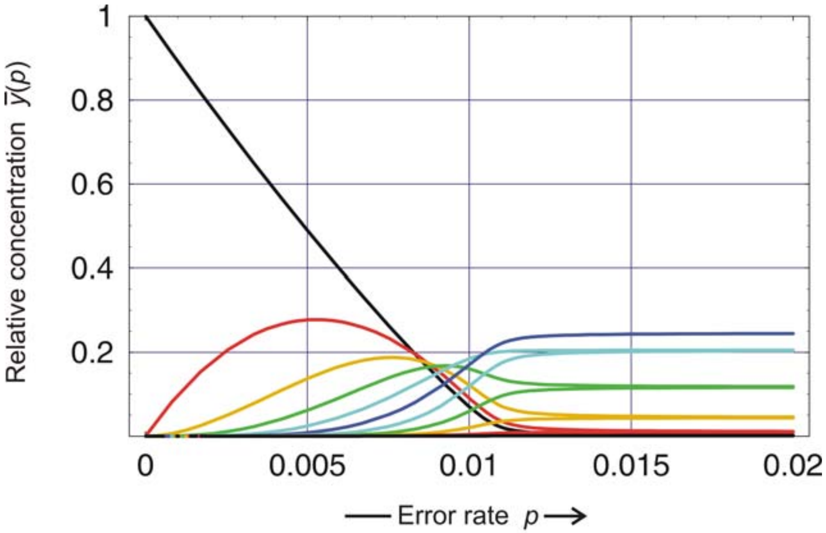


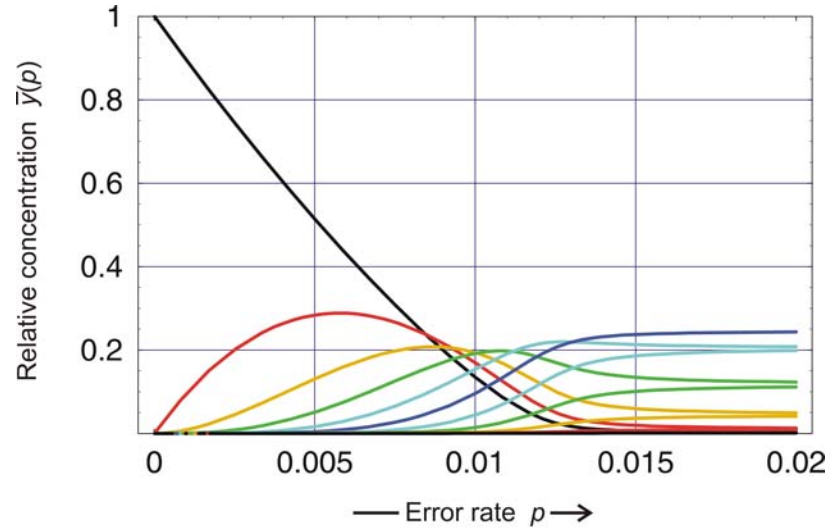


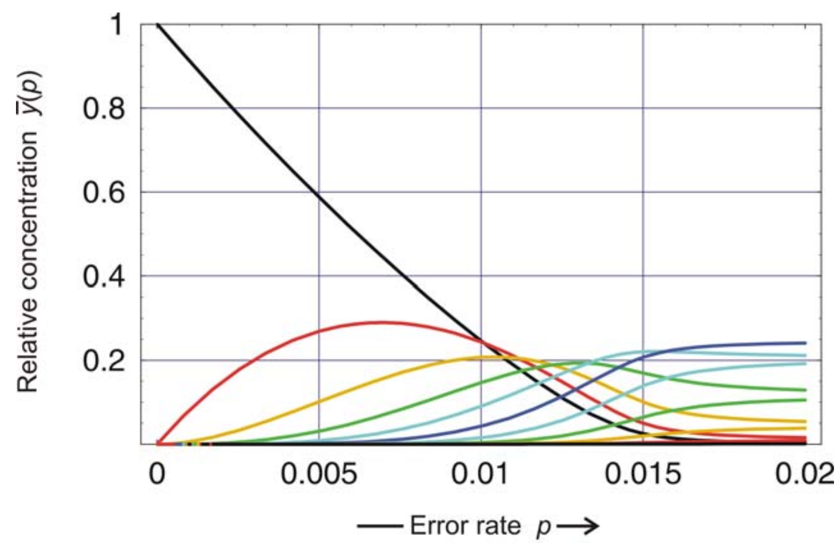


Error threshold: Classes

n = 10, σ = 1.1, δ = 0, 0.3, 0.5, and seed = 877





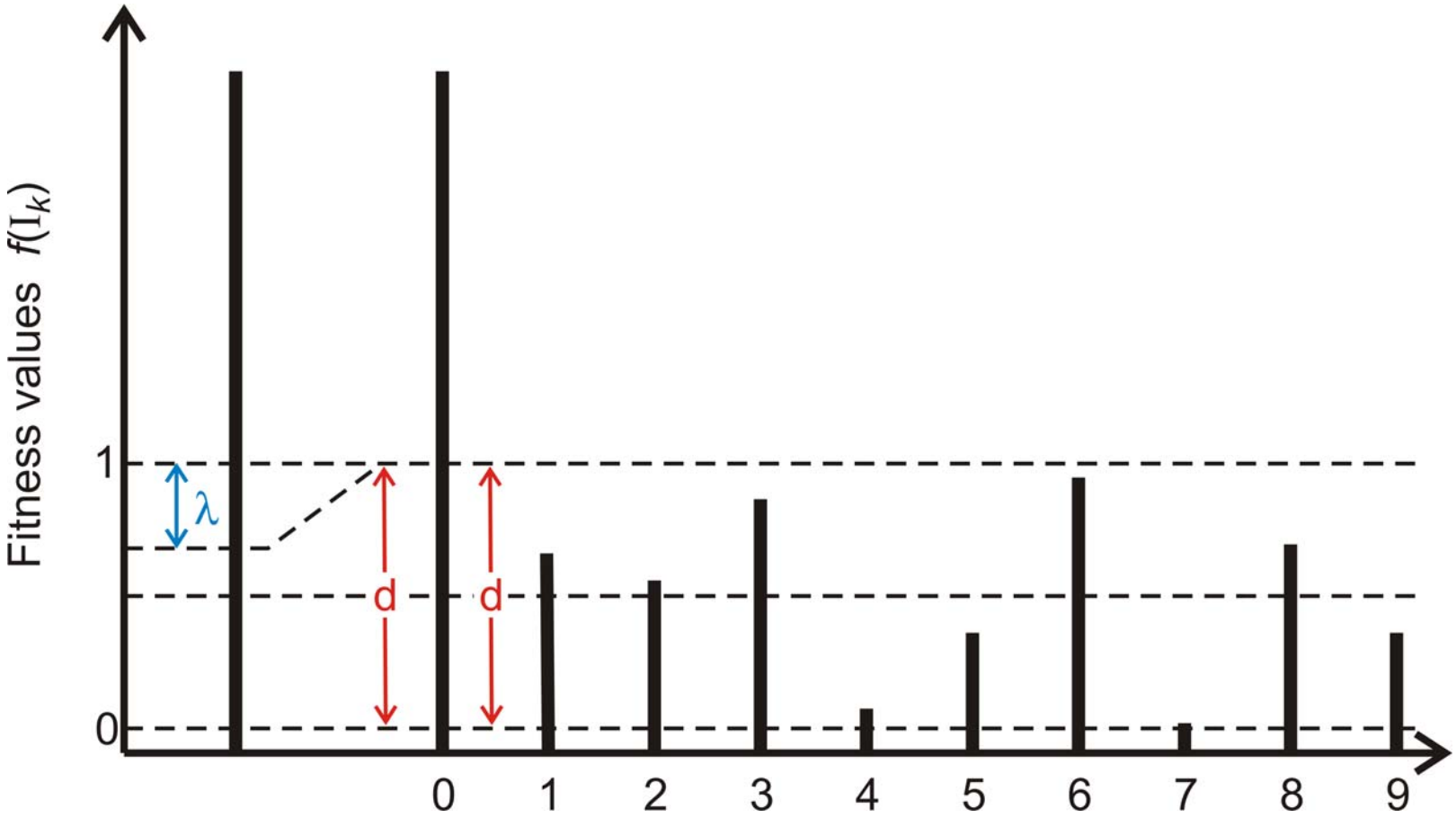


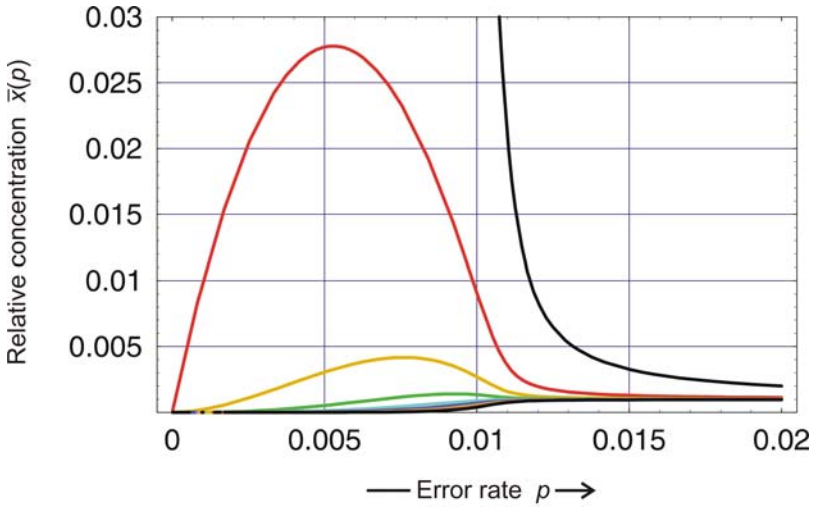
Error threshold: Classes

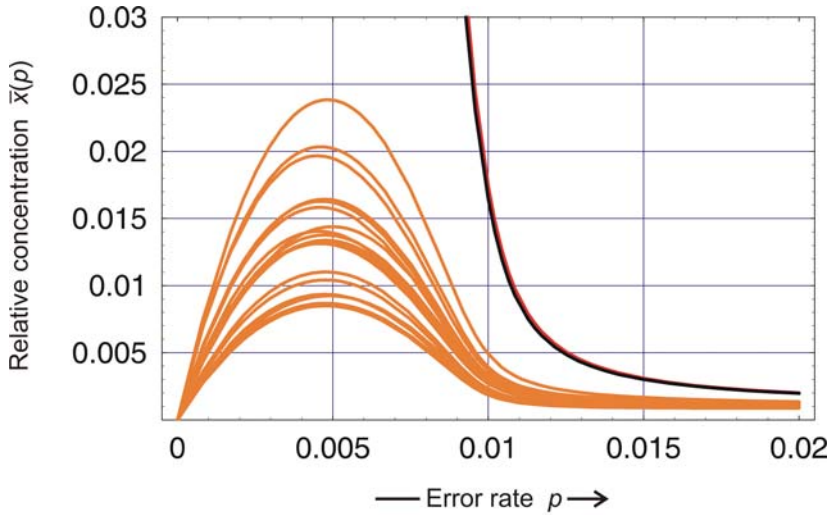
$n = 10$, $\sigma = 1.1$, $\delta = 0, 0.5$, and seed = 299, 877

Three sources of ruggedness:

1. Variation in fitness values
2. Deviations from uniform error rates
3. **Neutrality**

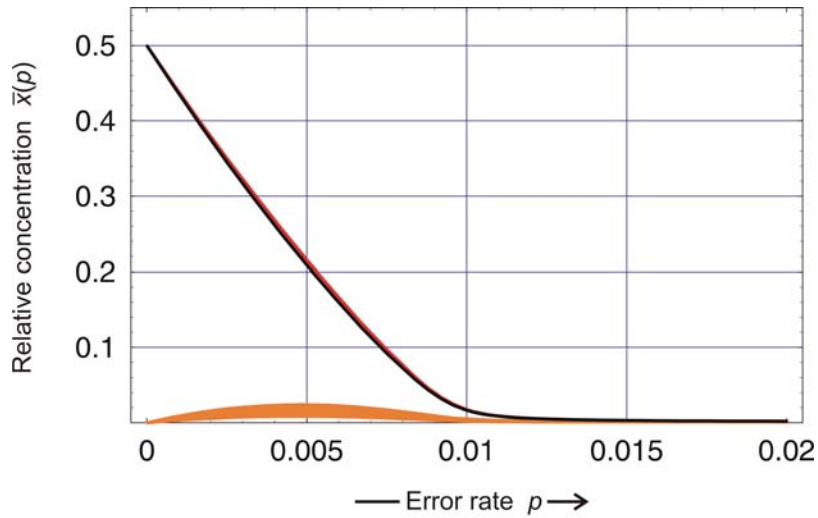

Fitness values $f(I_k)$
1
0
λ
d
d
0
1
2
3
4
5
6
7
8
9





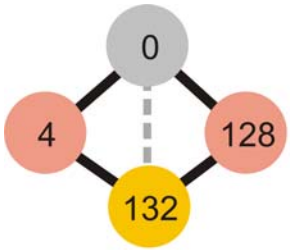
Neutral network

$\lambda = 0.01,\ s = 367$



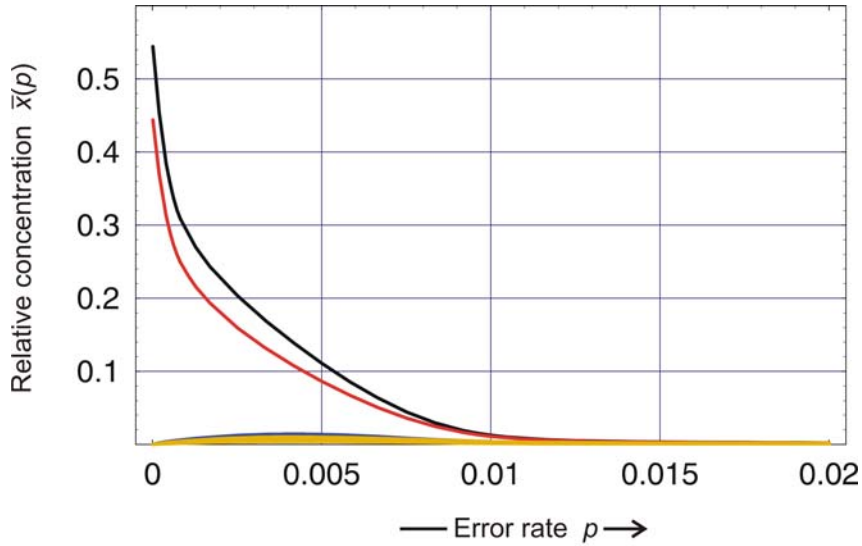
Error threshold: Individual sequences

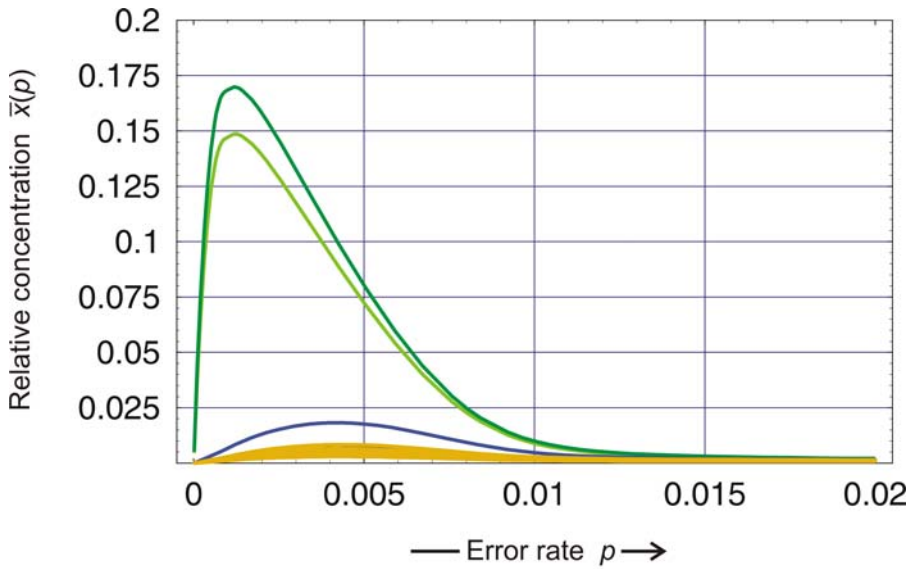
$n = 10,\ \sigma = 1.1,\ d = 1.0$



Neutral networks

λ = 0.01, s = 877





Error threshold: Individual sequences

n = 10, σ = 1.1, d = 1.0

Bulletin of Mathematical Biology Vol. 50, No. 6, pp. 635–660, 1988.
Printed in Great Britain.

0092-8240/88$3.00 + 0.00
Pergamon Press plc
Society for Mathematical Biology

STATIONARY MUTANT DISTRIBUTIONS AND EVOLUTIONARY OPTIMIZATION

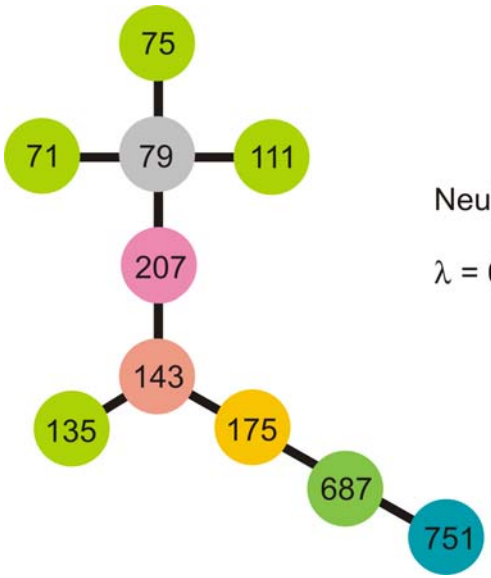
■ PETER SCHUSTER and JÖRG SWETINA
Institut für theoretische Chemie
und Strahlenchemie der Universität Wien,
Währingerstraße 17,
A 1090 Wien,
Austria

Molecular evolution is modelled by erroneous replication of binary sequences. We show how the selection of two species of equal or almost equal selective value is influenced by its nearest neighbours in sequence space. In the case of perfect neutrality and sufficiently small error rates we find that the Hamming distance between the species determines selection. As the error rate increases the fitness parameters of neighbouring species become more and more important. In the case of almost neutral sequences we observe a critical replication accuracy at which a drastic change in the "quasispecies", in the stationary mutant distribution occurs. Thus, in frequently mutating populations fitness turns out to be an ensemble property rather than an attribute of the individual.

In addition we investigate the time dependence of the mean excess production as a function of initial conditions. Although it is optimized under most conditions, cases can be found which are characterized by decrease or non-monotonous change in mean excess productions.

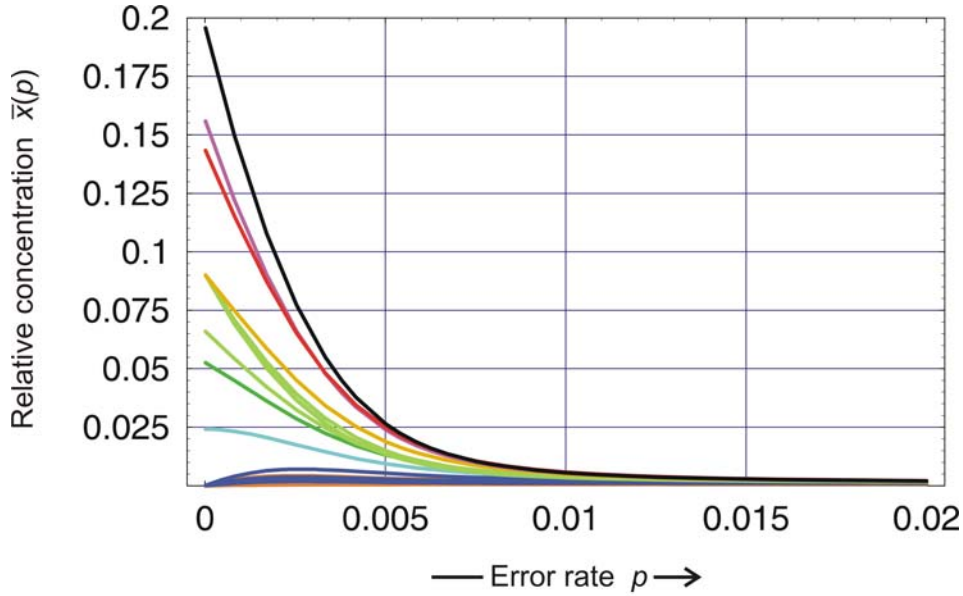
1. Introduction. Recent data from populations of RNA viruses provided direct evidence for vast sequence heterogeneity (Domingo *et al.*, 1987). The origin of this diversity is not yet completely known. It may be caused by the low replication accuracy of the polymerizing enzyme, commonly a virus specific, RNA dependent RNA synthetase, or it may be the result of a high degree of selective neutrality of polynucleotide sequences. Eventually, both factors contribute to the heterogeneity observed. Indeed, mutations occur much more frequently than previously assumed in microbiology. They are by no means rare events and hence, neither the methods of conventional population genetics (Ewens, 1979) nor the neutral theory (Kimura, 1983) can be applied to these virus populations. Selectively neutral variants may be close with respect to Hamming distance and then the commonly made assumption that the mutation backflow from the mutants to the wilde type is negligible does not apply.

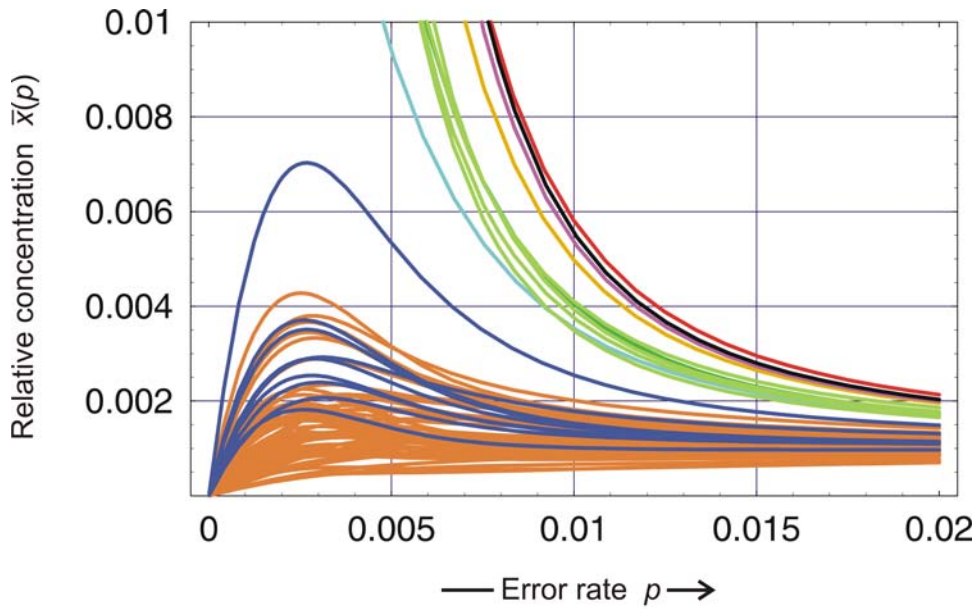
A kinetic theory of polynucleotide evolution which was developed during the past 15 years (Eigen, 1971; 1985; Eigen and Schuster, 1979; Eigen *et al.*, 1987; Schuster, 1986); Schuster and Sigmund, 1985) treats correct replication and mutation as parallel reactions within one and the same reaction network



Neutral network

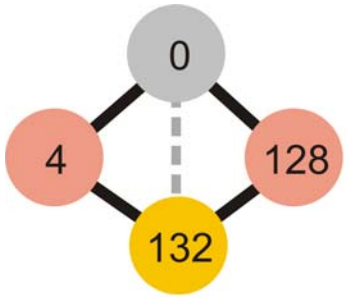
$\lambda = 0.10,\ s = 367$





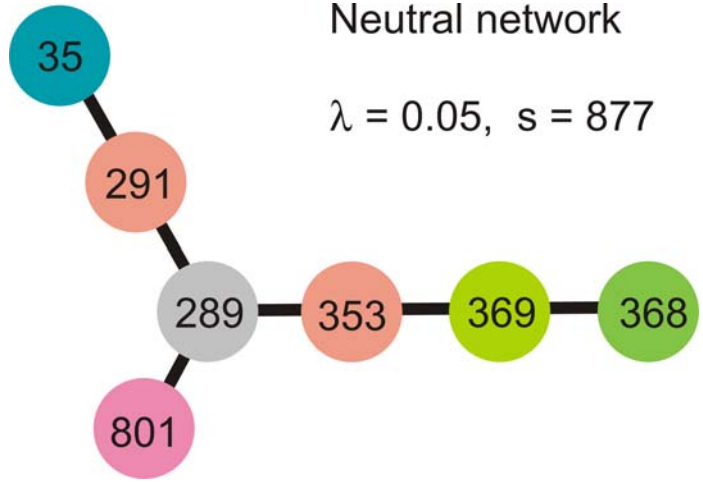
Error threshold: Individual sequences

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



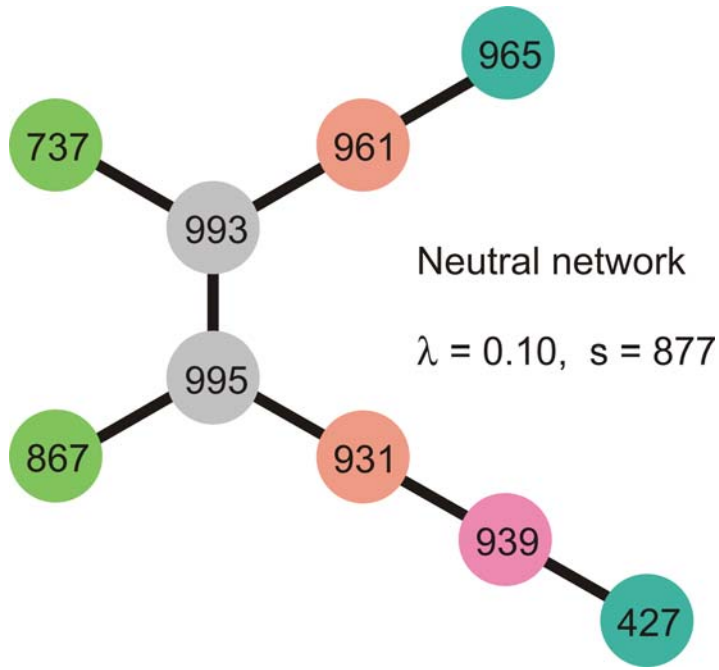
Neutral networks

$\lambda = 0.01,\ s = 877$



Neutral network

$\lambda = 0.05,\ s = 877$

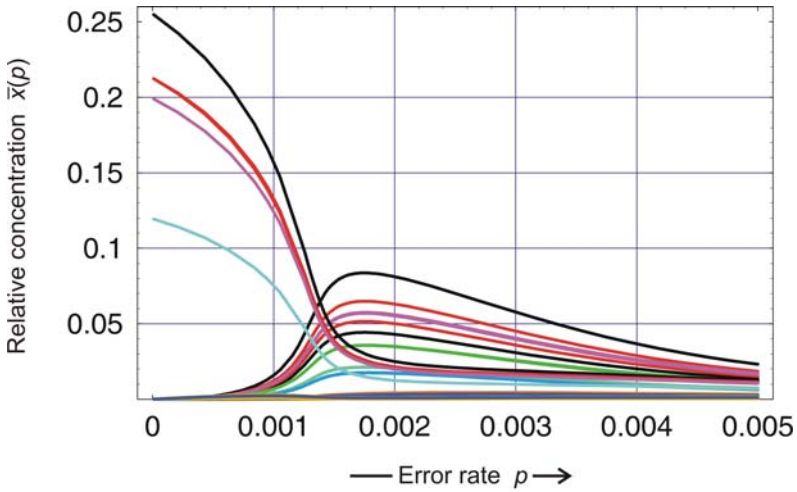


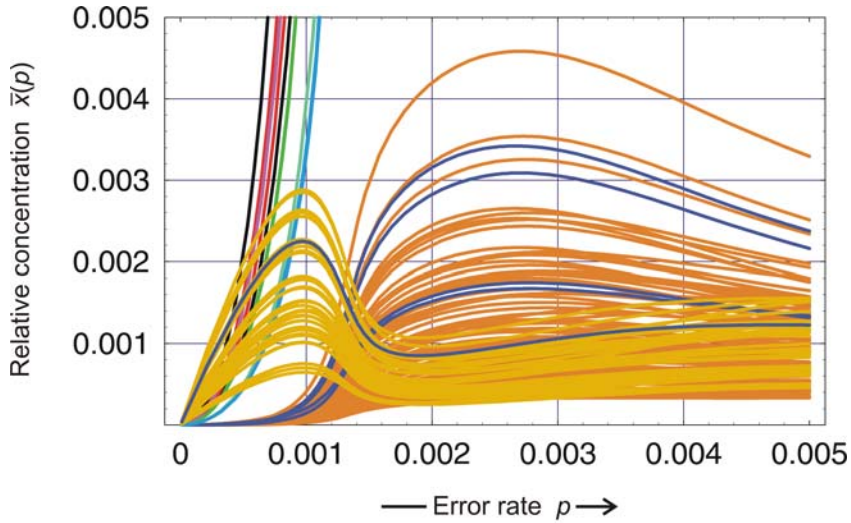
Neutral network

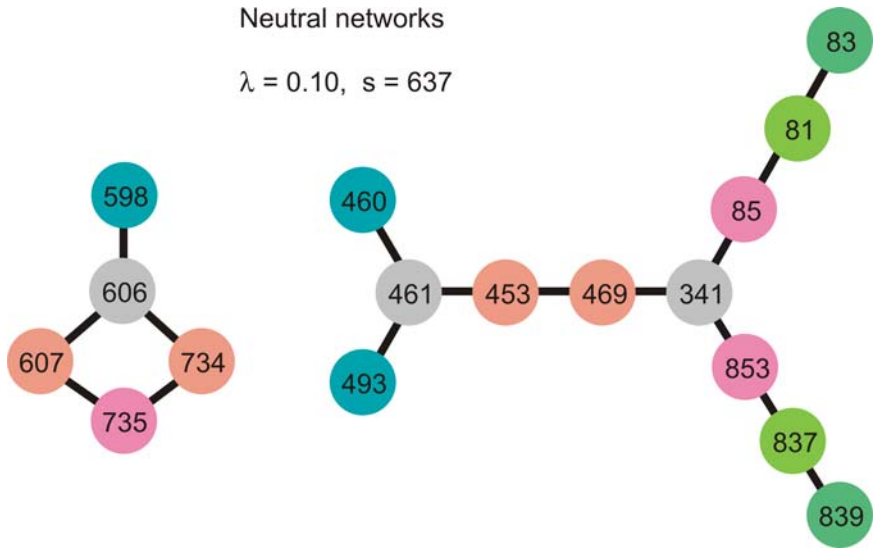
$\lambda = 0.10,\ s = 877$

Error threshold: Individual sequences

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

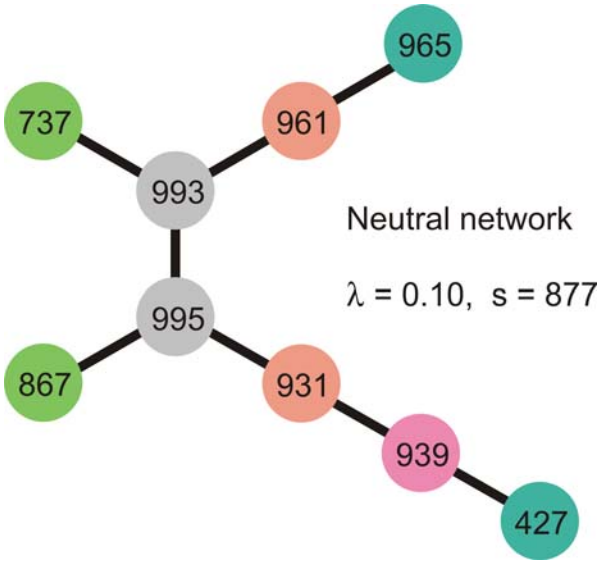






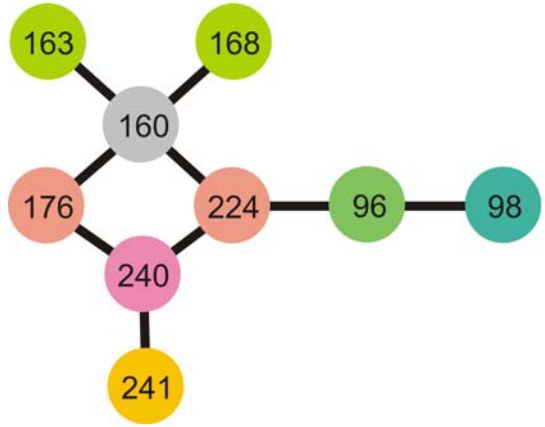
Error threshold: Individual sequences

n = 10, σ = 1.1, d = 1.0



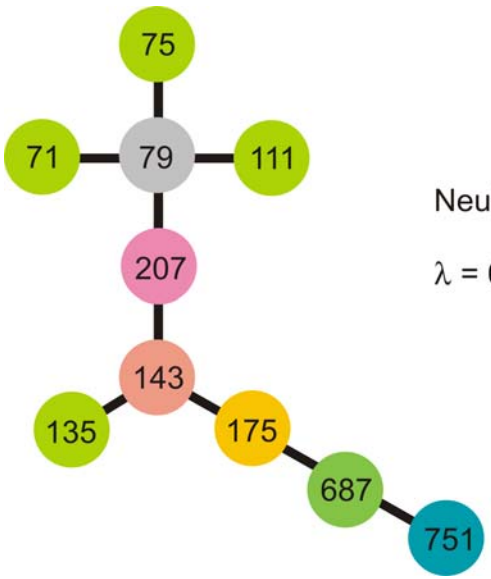
Neutral network

$\lambda = 0.10,\ s = 877$



Neutral network

$\lambda = 0.10,\ s = 491$

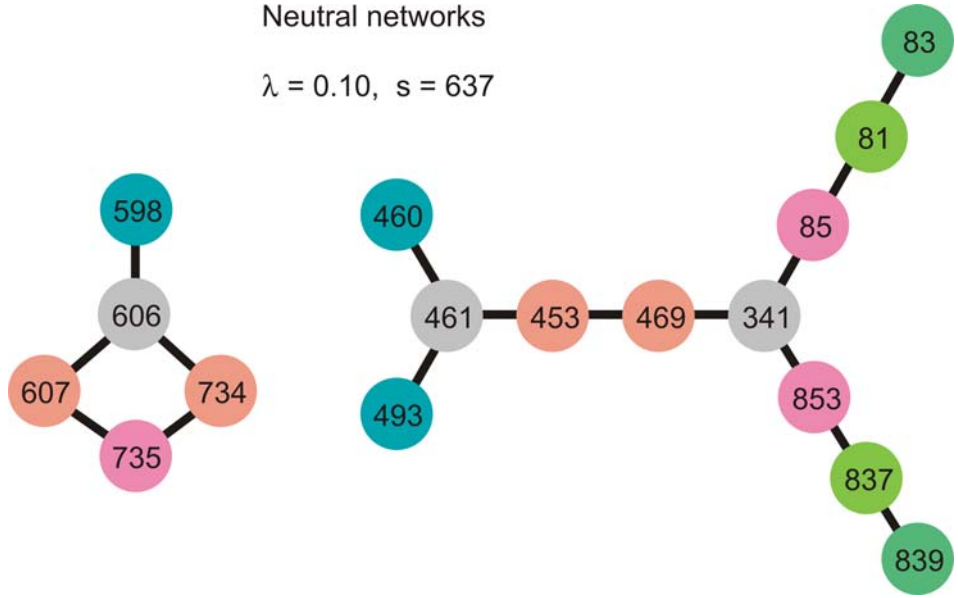


Neutral network

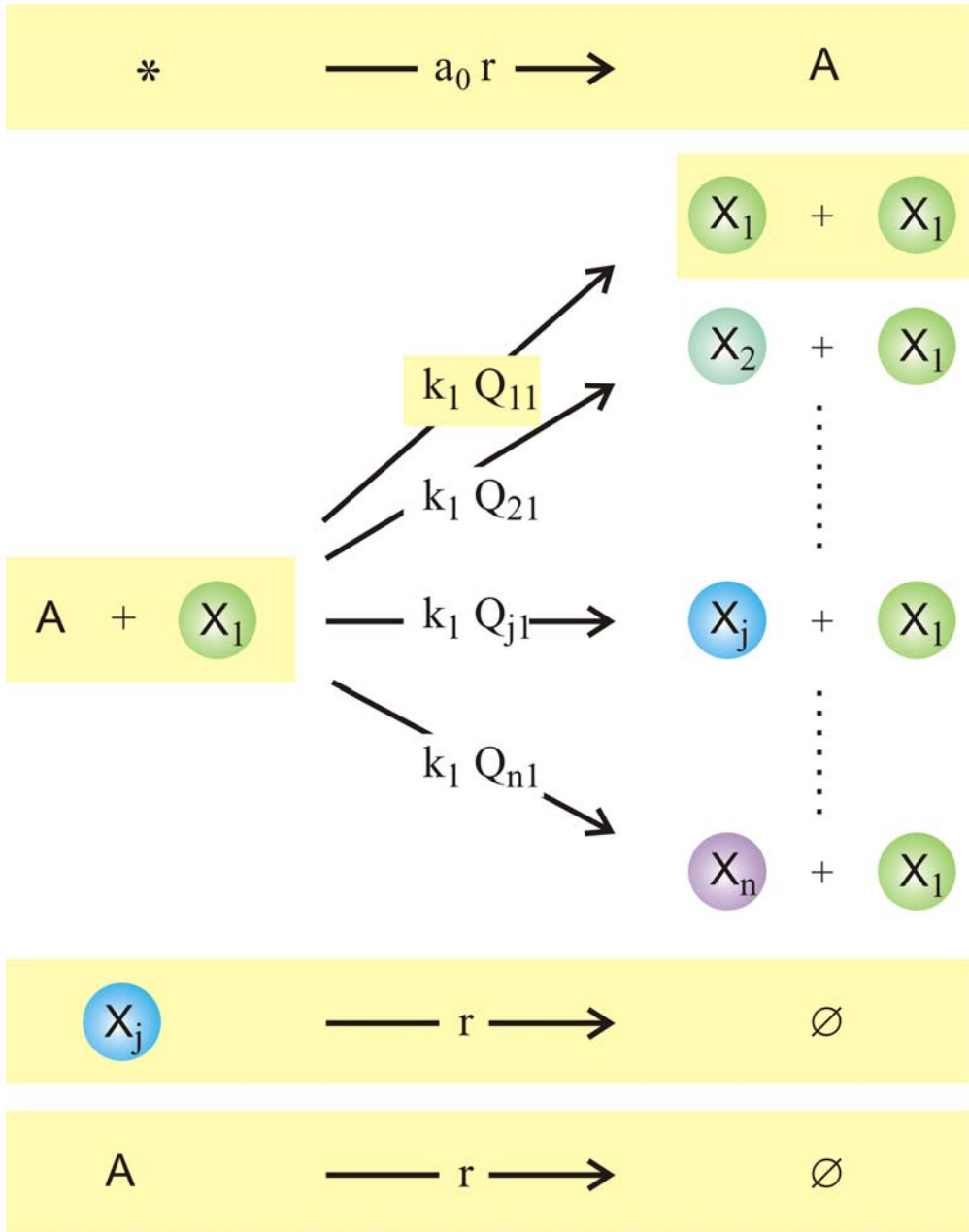
$\lambda = 0.10,\ s = 367$

Neutral networks

$\lambda = 0.10,\ s = 637$



1. Replication and mutation
2. Quasispecies and error thresholds
3. Fitness landscapes and randomization
4. **Lethal mutations**



$j = 1,2, \ldots ,n$

Lethal mutants and Frobenius theorem:

$$\mathrm{W} = \begin{pmatrix} w_{11} & 0 & \dots & 0 \\ w_{21} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & 0 & \dots & 0 \end{pmatrix} = w_{11} \begin{pmatrix} 1 & 0 & \dots & 0 \\ \frac{w_{21}}{w_{11}} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \frac{w_{n1}}{w_{11}} & 0 & \dots & 0 \end{pmatrix}$$

$$\mathrm{W}^k = w_{11}^k \begin{pmatrix} 1 & 0 & \dots & 0 \\ \frac{w_{21}}{w_{11}} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \frac{w_{n1}}{w_{11}} & 0 & \dots & 0 \end{pmatrix}$$

$$
\begin{aligned}
* &\xrightarrow{a_0 \cdot r} \mathbf{A} && : \text{influx} \\
\mathbf{A} + \mathbf{X}_1 &\xrightarrow{k_1 Q_{11}} 2\,\mathbf{X}_1\ ; && : \text{replication} \\
\mathbf{A} + \mathbf{X}_1 &\xrightarrow{k_1 Q_{j1}} \mathbf{X}_1 + \mathbf{X}_j \quad j = 2, \ldots, n && : \text{mutation} \\
\mathbf{A} &\xrightarrow{r} \emptyset && : \text{outflux} \\
\mathbf{X}_j &\xrightarrow{r} \emptyset\ ; \quad j = 1, \ldots, n && : \text{outflux}
\end{aligned}
$$

$$
\begin{aligned}
\frac{da}{dt} &= -\,a \sum_{j=1}^{n} k_1 Q_{j1}\, x_1 \;+\; r\,(a_0 - a) \;=\; -\,a\,k_1\,x_1 \;+\; r\,(a_0 - a) \\
\frac{dx_j}{dt} &= a\,Q_{j1}\,x_1 \;-\; r\,x_j
\end{aligned}
$$

Stationary solutions: 1. active state

$$r < k_1 Q_{11} a_0$$

$$\tilde{a} = \frac{r}{k_1 Q_{11}}$$

$$\tilde{x}_1 = Q_{11}(a_0 - \tilde{a}) = Q_{11} a_0 - \frac{r}{k_1}$$

$$\tilde{x}_j = Q_{j1}(a_0 - \tilde{a}) = Q_{j1}\left(a_0 - \frac{r}{k_1 Q_{11}}\right); \quad j = 2, 3, \ldots, n$$

Stationary solutions: 2. extinction

$$r > k_1 Q_{11} a_0$$

$$\tilde{a} = a_0$$

$$\tilde{x}_j = 0; \; j = 1, 2, \ldots, n$$

Find $r(t)$ such that $a(t) = \bar{a} = const.$

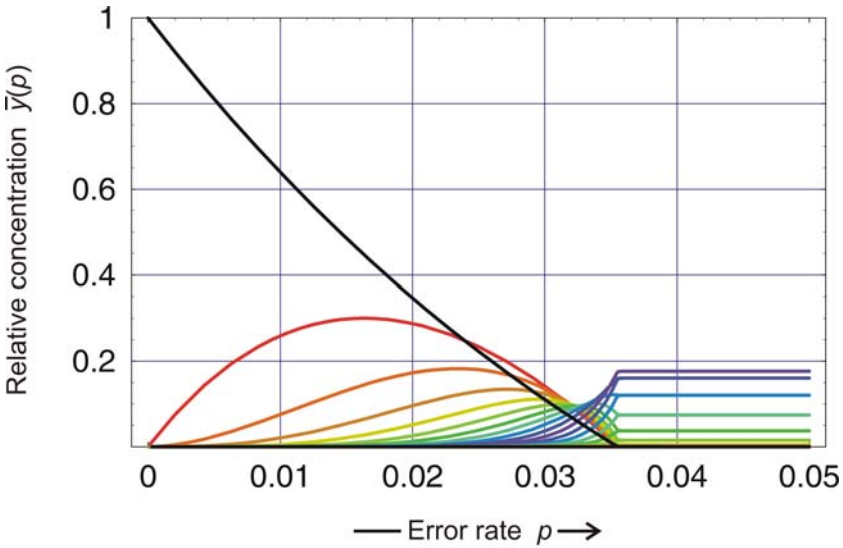
$$\frac{da}{dt} \;=\; 0 \;=\; -\,\bar{a} \sum_{j=1}^{n} k_1 Q_{j1}\, x_1 \;+\; r(t)\,(a_0 - \bar{a})$$

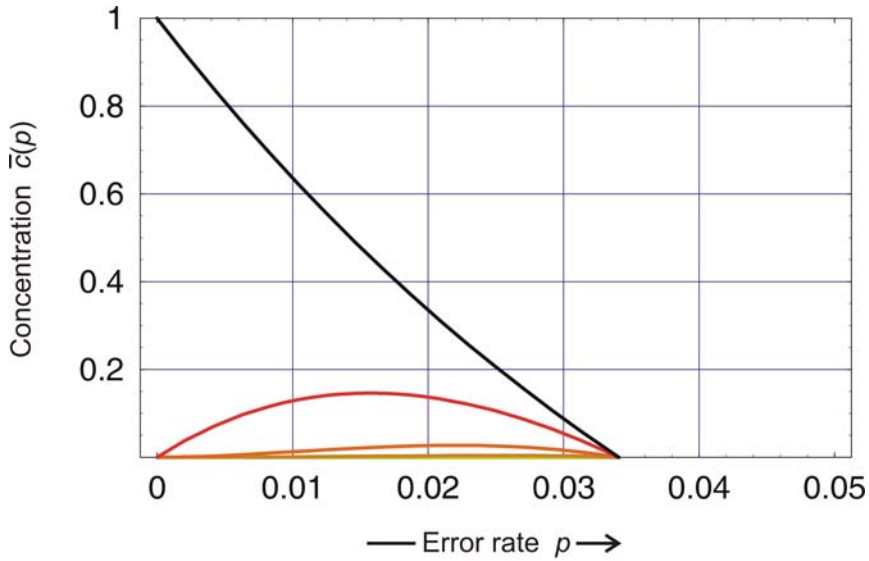
$$r(t) \;=\; \frac{\bar{a}}{a_0 - \bar{a}}\, k_1\, x_1\,;\;\; f_1 \;=\; k_1\,\bar{a}\,;\;\; \sum_{i=1}^{n} x_i \;=\; c \;=\; a_0 - \bar{a}$$

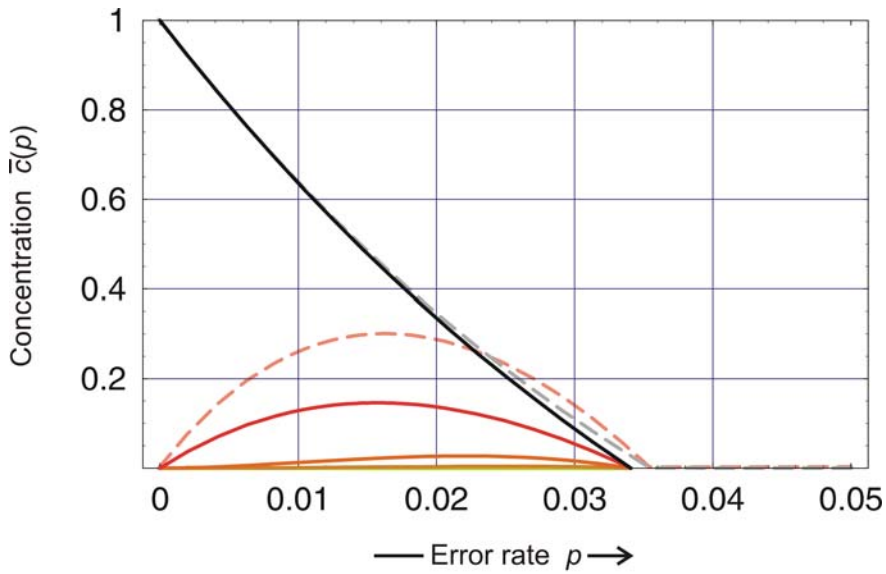
$$\frac{dx_j}{dt} \;=\; f_1 Q_{j1}\, x_1 \;-\; x_j\, \frac{f_1\, x_1}{\sum_{i=1}^{n} x_i} \;=\; f_1\, x_1 \left(Q_{j1} \;-\; \frac{x_j}{c} \right)$$

Stationary solutions:

$$\bar{x}_j \;=\; Q_{j1} \sum_{i=1}^{n} \bar{x}_i \;=\; Q_{ji}\, c$$







Replication-mutation in the flow reactor

One viable species: I_1

$n = 20$, $\sigma = 2$

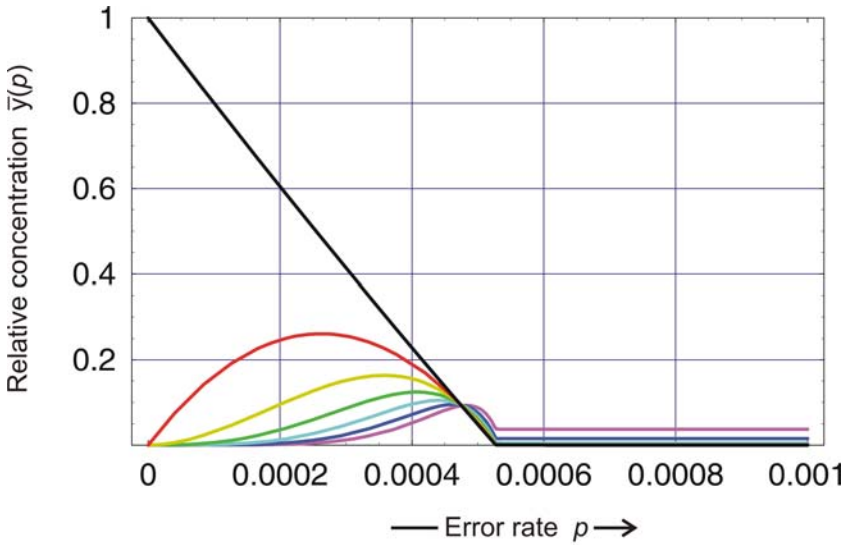
$$
\begin{aligned}
* &\xrightarrow{a_0 \cdot r} \mathbf{A} && : \text{influx} \\
\mathbf{A} + \mathbf{I}_1 &\xrightarrow{Q_{11} k_1} 2\,\mathbf{I}_1 && : \text{replication} \\
\mathbf{A} + \mathbf{I}_1 &\xrightarrow{Q_{21} k_1} \mathbf{I}_2 + \mathbf{I}_1 && : \text{mutation} \\
\mathbf{A} + \mathbf{I}_1 &\xrightarrow{Q_{31} k_1} \mathbf{I}_3 + \mathbf{I}_1 && : \text{mutation} \\
\mathbf{A} + \mathbf{I}_2 &\xrightarrow{Q_{12} k_1} \mathbf{I}_1 + \mathbf{I}_2 && : \text{mutation} \\
\mathbf{A} + \mathbf{I}_2 &\xrightarrow{Q_{22} k_1} 2\,\mathbf{I}_2 && : \text{replication} \\
\mathbf{A} + \mathbf{I}_2 &\xrightarrow{Q_{32} k_1} \mathbf{I}_3 + \mathbf{I}_2 && : \text{mutation} \\
\mathbf{A}, \mathbf{I}_1, \mathbf{I}_2, \mathbf{I}_3 &\xrightarrow{r} \varnothing && : \text{outflux}
\end{aligned}
$$

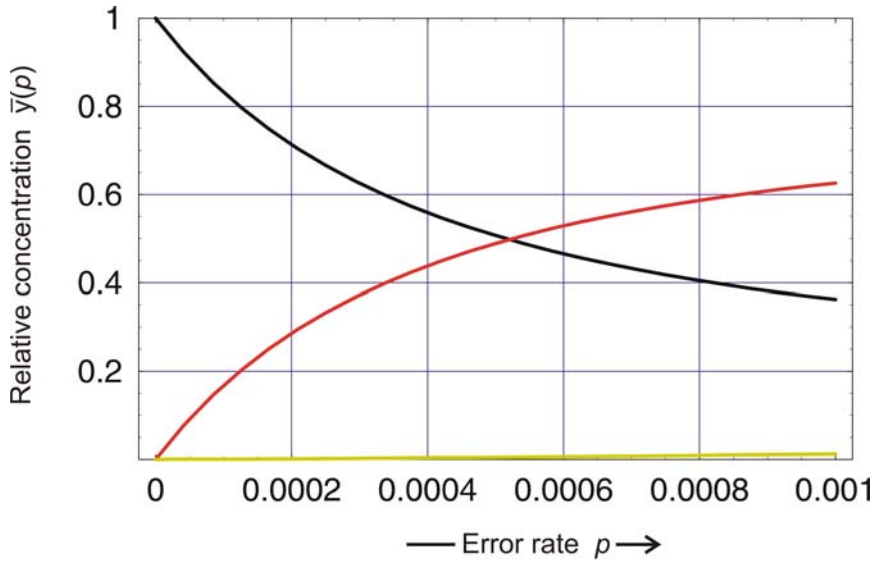
$$
\begin{aligned}
\frac{da}{dt} &= -\left(k_1 c_1 + k_2 c_2\right) a + r\,(a_0 - a) \\
\frac{dc_1}{dt} &= a\left(Q_{11} k_1 c_1 + Q_{12} k_2 c_2\right) - r\,c_1 \\
\frac{dc_2}{dt} &= a\left(Q_{21} k_1 c_1 + Q_{22} k_2 c_2\right) - r\,c_2 \\
\frac{dc_3}{dt} &= a\left(Q_{31} k_1 c_1 + Q_{32} k_2 c_2\right) - r\,c_3
\end{aligned}
$$

$$
\begin{aligned}
\bar{a} &= \frac{a_0\, r}{k_1 \bar{c}_1 + k_2\, \bar{c}_2 + r} \\
\bar{c}_1 &= \frac{a_0\ (Q_{11} k_1 \bar{c}_1 + Q_{12} k_2 \bar{c}_2)}{k_1 \bar{c}_1 + k_2\, \bar{c}_2 + r} \\
\bar{c}_2 &= \frac{a_0\ (Q_{21} k_1 \bar{c}_1 + Q_{22} k_2 \bar{c}_2)}{k_1 \bar{c}_1 + k_2\, \bar{c}_2 + r} \\
\bar{c}_3 &= \frac{a_0\ (Q_{31} k_1 \bar{c}_1 + Q_{32} k_2 \bar{c}_2)}{k_1 \bar{c}_1 + k_2\, \bar{c}_2 + r}
\end{aligned}
$$

Replication-mutation in the flow reactor

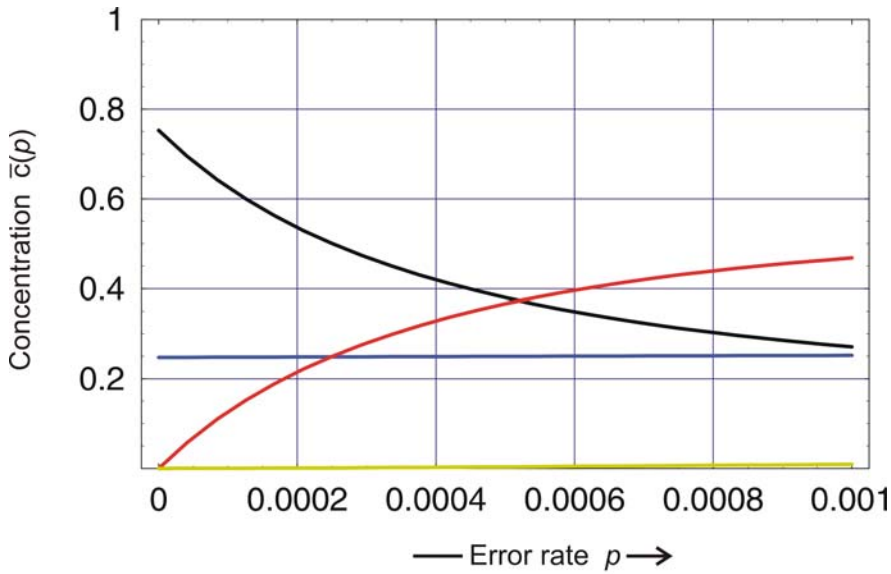
Two viable species: I_1 and I_2





$$p_{\max} = 0.0005$$

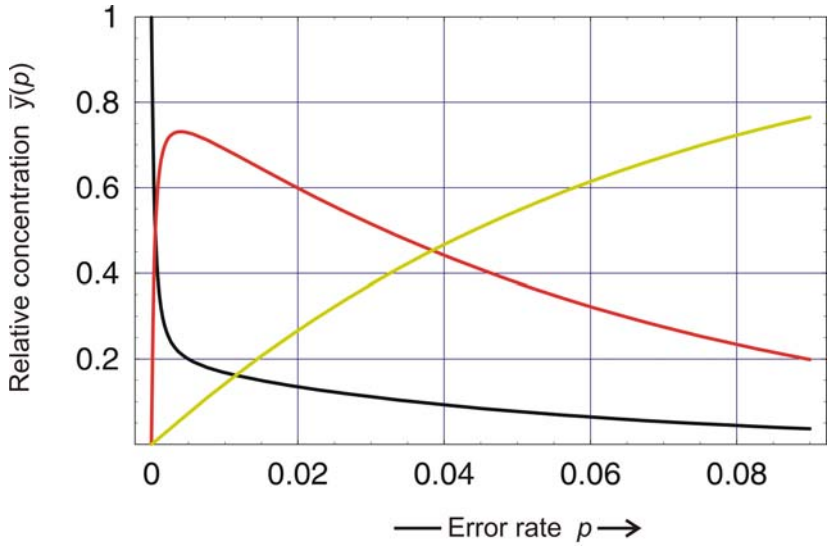
Error threshold: $$p_{\max} = \frac{\ln \sigma}{n}$$

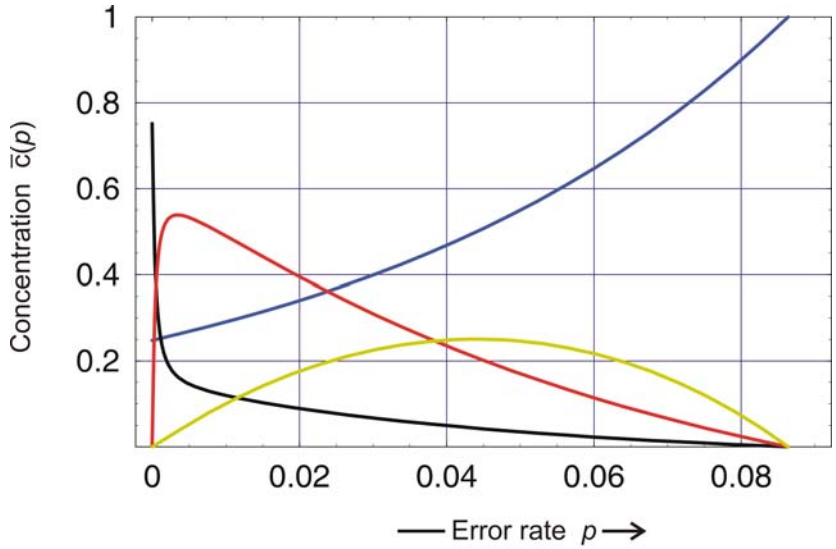


Replication-mutation in the flow reactor

Two viable species: I_1 and I_2

n = 20 , σ = 1.01 , k = 1, a_0 = 1 , r = 0.25





$$p_{\text{ext}} = 0.083$$

Extinction threshold: $(1-p_{\text{ext}})^{n-1}\left(\sigma(1+(n-1)\,p_{\text{ext}})+\sqrt{n}\right)\left(1+\sqrt{n}\right)^{-1} = \dfrac{r}{k\,a_0}$

Replication-mutation in the flow reactor. Two viable species: I_1 and I_2

n = 20 , σ = 1.01 , k = 1 , a_0 = 1 , r = 0.25

Acknowledgement of support

Universität Wien

Fonds zur Förderung der wissenschaftlichen Forschung (FWF)
Projects No. 09942, 10578, 11065, 13093
13887, and 14898

Wiener Wissenschafts-, Forschungs- und Technologiefonds (WWTF)
Project No. Mat05

Jubiläumsfonds der Österreichischen Nationalbank
Project No. Nat-7813

European Commission: Contracts No. 98-0189, 12835 (NEST)

Austrian Genome Research Program – GEN-AU

Siemens AG, Austria

Universität Wien and the Santa Fe Institute

Coworkers

Universität Wien

Walter Fontana, Harvard Medical School, MA

Christian Forst, **Christian Reidys**, Los Alamos National Laboratory, NM

Peter Stadler, **Bärbel Stadler**, Universität Leipzig, GE

Jord Nagel, **Kees Pleij**, Universiteit Leiden, NL

Christoph Flamm, **Ivo L.Hofacker**, **Andreas Svrček-Seiler**,
Universität Wien, AT

Kurt Grünberger, **Michael Kospach**, **Andreas Wernitznig**,
Stefanie Widder, Michael Wolfinger, Stefan Wuchty,Universität Wien, AT

Stefan Bernhart, **Jan Cupal**, **Lukas Endler, Ulrike Langhammer**,
Rainer Machne, Ulrike Mückstein, **Hakim Tafer,** Universität Wien, AT

Ulrike Göbel, **Walter Grüner**, **Stefan Kopp**, **Jaqueline Weber,**
Institut für Molekulare Biotechnologie, Jena, GE

Web-Page for further information:

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