Current Topics in Microbiology and Immunology

Esteban Domingo
Peter Schuster *Editors*

Quasispecies: From Theory to Experimental Systems

Springer

Current Topics in Microbiology and Immunology

Volume 392

More information about this series at http://www.springer.com/series/82

Esteban Domingo · Peter Schuster
Editors

Quasispecies: From Theory to Experimental Systems

Responsible Series Editor: Michael B.A. Oldstone

Springer

Editors
Esteban Domingo
Centro de Biología Molecular Severo Ochoa (CSIC-UAM)
Campus de Cantoblanco
Madrid
Spain

Peter Schuster
Institut für Theoretische Chemie
Universität Wien
Vienna
Austria

and

The Santa Fe Institute
Santa Fe, NM
USA

ISSN 0070-217X ISSN 2196-9965 (electronic)
Current Topics in Microbiology and Immunology
ISBN 978-3-319-23897-5 ISBN 978-3-319-23898-2 (eBook)
DOI 10.1007/978-3-319-23898-2

Library of Congress Control Number: 2016930545

Printed on acid-free paper

This Springer imprint is published by SpringerNature
The registered company is Springer International Publishing AG Switzerland

In memory of
Christof Biebricher and
Emmanuel Tannenbaum

Foreword

The concept of the quasispecies will soon be 50 years old. This term I introduced in the late 1960s in my considerations on self-organization of matter and the evolution of biological macromolecules.

The idea of heterogeneous populations has been quite uncommon in biology. Population biology considered mutation as a rare event and even in the absence of selective differences Kimura's theory of neutral evolution predicted a very low fraction of mutants. Molecular biology, on the other hand, was showing that correct reproduction and mutation are parallel reaction channels, which inevitably result in a distribution of genotypes. Heterogeneity in populations of bacteriophages has been verified experimentally in the laboratory of Charles Weissmann at about the same time. Chemical kinetics of replication and mutation built a firm bridge from molecular biology to evolutionary theory and provided the basis for the design of evolution experiments in the test tube. These included the work of Sol Spiegelman and his group as well as that of Christof Biebricher, who conducted research on replication of RNA from the bacteriophage Qβ in my laboratory. Biebricher's fundamental experiments on exploring the chemical kinetics of in vitro evolution laid the foundation for the forthcoming applications of the theory in evolutionary biotechnology and antiviral pharmacology. Without his careful and detailed work on Qβ-phage RNA, we would not now be able to understand Darwinian evolution in the test tube.

The dedication of the volume to the memory of Christof Biebricher recognizes his pioneering research on quasispecies. Esteban Domingo and Peter Schuster present fourteen selected chapters written by experts who revisit the concept of quasispecies, review the current status in theory and experiment, and provide an overview of its application to virus evolution. This book should find a place in every Life Science collection.

Göttingen
July 2015

Manfred Eigen

Preface

Quasispecies theory has come of age, and regular updates of the concept of mutation-caused diversity of populations are appropriate in order to provide straightforward access to information on recent progress in theory and applications to the real-world problems. Among a great variety of other applications, the concept of viral quasispecies, the limitation of sustainable mutation rates through error thresholds, and its usage in the development of antiviral therapies are most prominent. Indeed viral infection of hosts, epidemic spread of viral diseases as well as evolution of virus species can hardly be understood in depth without the notion of quasispecies and sufficient knowledge on their evolutionary dynamics.

Within the last decade, progress in the theory of quasispecies came mainly from two developments: (i) the accessibility of real fitness landscapes due to the enormous technical improvements in sequencing and high-throughput techniques and (ii) the exploitation of the formal mathematical analogy of quasispecies dynamics and statistical mechanics of classical and quantum spin systems. In the latter case, Eigen's quasispecies concept and Crow and Kimura's mutation-selection model give rise to identical mathematical problems, and therefore, the distinction between the two approaches is often not made with sufficient clarity: Mutation in the quasispecies theory is a parallel process to correct replication and happens during reproduction, whereas mutation and replication are entirely two separate processes in the Crow–Kimura model and occur at different instants. In virus reproduction, the former case applies and, accordingly, only the quasispecies concept is the appropriate model. Another important issue concerns the mechanism of mutagen-induced extinction of viral populations. A change in the mutation rate through certain pharmaceutical compounds, notably nucleotide analogues, is considered as the causing principle. This antiviral mechanism has been popularized under the name "lethal mutagenesis." Quasispecies theory predicts a maximal error rate that is compatible with a stable virus population. The antiviral strategy is to destabilize and extinguish the virus population by a drug-induced increase of the mutation rate that generates defective genomes that drive the virus population through the error threshold. There is also a second mechanism of extinction which

consists in increasing the fraction of lethal variants in the population above the maximum required for survival. As shown in several chapters of this book, both mechanisms are in operation and by now the interplay of error threshold phenomena and lethal mutagenesis is well established.

Since the CTMI volume 299 on Quasispecies published in 2006, major progress has been based in developments on the scope of applicability of quasispecies theory, the implementation of ultra-deep sequencing to analyze mutant spectra of viral populations, and the confirmation of the profound biological effects that changes in replication fidelity have on virus behavior. John Holland wrote in the 2006 volume the closing chapter on a historical perspective of major transitions in the understanding of RNA viruses. He emphasized the recognition of viral quasispecies dynamics as a major development in RNA virology. Sadly, John passed away on October 11, 2013. As people convinced of the role of viral dynamics to understand viral disease, we were very fortunate that John's laboratory became involved in this area of research, after a long and productive career that resulted in fundamental contributions in virology. Indeed, John Holland established the concept of cellular receptors as determinants of tissue tropism and did pioneering work on viral polyprotein synthesis and processing, and in the characterization and biological activities of defective interfering (DI) RNAs and particles (see the "In Memoriam" note by K.R. Spindler and B.L. Semler in Journal of Virology 88: 5903–5905, 2014). He recognized high mutation rates as key to the understanding of competition dynamics between vesicular stomatitis virus and its DIs. Expanding on this, his laboratory made key contributions to measurements of viral mutation rates, to the understanding of viral quasispecies, and to establish connections between quasispecies and several concepts from classic population genetics. His work permeates most of the chapters on experimental quasispecies in the present volume, with a number of topics which are now pursued with the new experimental, theoretical, and bioinformatic tools that have become available over the last years.

The present volume consists of fourteen chapters. Chapter "What Is a Quasispecies? Historical Origins and Current Scope" provides an introduction into the concept of quasispecies and its applications. The next chapter "The Nucleation of Semantic Information in Prebiotic Matter" deals with the concept of semantic information and its origin in biology. The simplest systems that can be studied by both extensive computation and in vitro experiments showing adaptation and evolution are small RNA molecules that are described in chapter "Evolution of RNA-Based Networks." The interplay of fitness landscapes and mutation-selection dynamics in particular with respect to quasispecies formation and the existence of error thresholds is treated in chapter "Quasispecies on Fitness Landscapes." Chapter "Mathematical Models of Quasispecies Theory and Exact Results for the Dynamics" deals with the application of methods from quantum statistical mechanics to derive solutions for the quasispecies concept and the Crow–Kimura mutation-selection model. The quasispecies concept can be extended to much more complex reproduction mechanisms than asexual reproduction and mutation as is shown in chapter "Theoretical Models of Generalized Quasispecies." Chapter "Theories of Lethal Mutagenesis: From Error Catastrophe to Lethal Defection" reviews models of lethal

mutagenesis, with a critical assessment of their relevance to experimental observations. Chapter "Estimating Fitness of Viral Quasispecies from Next-Generation Sequencing Data" reveals how new developments in deep sequencing can provide an increasingly accurate picture of viral population structure, and fitness landscapes based on the quasispecies model. Chapter "Getting to Know Viral Evolutionary Strategies: Towards the Next Generation of Quasispecies Models" discusses the difficulties of capturing the implications of complex mutant spectra, and the need to integrate theoretical and experimental approaches. In chapter "Cooperative Interaction Within RNA Virus Mutant Spectra," the fundamental issue of biological implications of interactions among components of a mutant spectrum is addressed. In chapter "Arenavirus Quasispecies and Their Biological Implications," replication of the important group of arenavirus pathogens is dissected at the molecular level to reveal the impact of genome variation; this chapter is a paradigm of the multiple facets of genome variation in viral pathogenesis and how to approach the problem experimentally. Chapter "Models of Viral Population Dynamics" connects theory and observation in the important area of dynamics of drug resistance in viral populations, centered on research on HIV-1. Chapter "Fidelity Variants and RNA Quasispecies" reviews the increasingly important field of copying fidelity mutants in viruses, and the book closes with chapter "Antiviral Strategies Based on Lethal Mutagenesis and Error Threshold" which is a review of recent developments on antiviral treatment designs based on lethal mutagenesis. We have tried to bring to the reader an updated account of quasispecies theory and experiment, and to reduce the gap between these two branches of research.

Madrid
Vienna
March 2016

Esteban Domingo
Peter Schuster

Contents

What Is a Quasispecies? Historical Origins and Current Scope 1
Esteban Domingo and Peter Schuster

The Nucleation of Semantic Information in Prebiotic Matter 23
Bernd-Olaf Küppers

Evolution of RNA-Based Networks 43
Peter F. Stadler

Quasispecies on Fitness Landscapes 61
Peter Schuster

Mathematical Models of Quasi-Species Theory and Exact Results for the Dynamics 121
David B. Saakian and Chin-Kun Hu

Theoretical Models of Generalized Quasispecies 141
Nathaniel Wagner, Yoav Atsmon-Raz and Gonen Ashkenasy

Theories of Lethal Mutagenesis: From Error Catastrophe to Lethal Defection 161
Héctor Tejero, Francisco Montero and Juan Carlos Nuño

Estimating Fitness of Viral Quasispecies from Next-Generation Sequencing Data 181
David Seifert and Niko Beerenwinkel

Getting to Know Viral Evolutionary Strategies: Towards the Next Generation of Quasispecies Models 201
Susanna Manrubia and Ester Lázaro

Cooperative Interaction Within RNA Virus Mutant Spectra 219
Yuta Shirogane, Shumpei Watanabe and Yusuke Yanagi

Arenavirus Quasispecies and Their Biological Implications 231
Ana Grande-Pérez, Veronica Martin, Hector Moreno
and Juan C. de la Torre

Models of Viral Population Dynamics 277
Pranesh Padmanabhan and Narendra M. Dixit

Fidelity Variants and RNA Quasispecies 303
Antonio V. Bordería, Kathryn Rozen-Gagnon and Marco Vignuzzi

Antiviral Strategies Based on Lethal Mutagenesis and Error Threshold 323
Celia Perales and Esteban Domingo

Author Index 341

Subject Index 351

What Is a Quasispecies? Historical Origins and Current Scope

Esteban Domingo and Peter Schuster

Abstract The quasispecies concept is introduced by means of a simple theoretical model that uses as little chemical kinetics and mathematics as possible but fully in the spirit of Albert Einstein who said: "Things should be made as simple as possible but not simpler." More elaborate treatments follow in the forthcoming chapters. It is shown that the most important results of the theory, in particular the existence of error thresholds, are not dependent on simplifying assumptions concerning the distribution of fitness values. Error thresholds are regularly found on landscapes with large and irregular scatter of fitness. After the introduction to theory, it will be shown how experimental data on the evolution of molecules or viruses may be fit to the theoretical model.

Contents

1 Evolution on the Cross-Road of Chemistry and Biology 2
2 A Few Quantitative Relations 6
3 What Happens Beyond Error Threshold? 8
4 Origin of the Experimental Quasispecies Concept 10
5 Quasispecies Theory and Experimental RNA Virus Quasispecies 12

E. Domingo (✉)
Centro de Biología Molecular "Severo Ochoa" (CSIC-UAM), Consejo Superior de Investigaciones Científicas (CSIC), Campus de Cantoblanco, 28049 Madrid, Spain
e-mail: edomingo@cbm.csic.es

E. Domingo
Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas (CIBERehd), Barcelona, Spain

P. Schuster (✉)
Institut für Theoretische Chemie der Universität Wien, Währingerstraβe 17, 1090 Vienna, Austria
e-mail: pks@tbi.univie.ac.at

P. Schuster
The Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA

Current Topics in Microbiology and Immunology (2016) 392: 1–22
DOI 10.1007/82_2015_453

Published Online: 30 August 2015

6 Extensions of Quasispecies to Non-viral Systems 15
7 Limitations and Strengths of the Quasispecies Concept 18
References 19

1 Evolution on the Cross-Road of Chemistry and Biology

A theory of evolution at the molecular level was conceived by Manfred Eigen (Eigen 1971; Eigen and Schuster 1977, 1978a, 1978b) through merging population dynamics with the knowledge of molecular biology. In this way, evolution could be integrated into chemical kinetics without losing the characteristic features of biology, in particular the role of genetic information stored in nucleic acid molecules and the nature of mutations were fully preserved. The key to evolution is reproduction, and at the level of DNA or RNA, reproduction is replication, which can be simply understood as copying of genetic information, which is error prone in general but can be error free or correct in a particular replication event. Modeling the basic property of molecular copying mechanisms, correct replication and mutation are parallel chemical reaction channels (Fig. 1) and accordingly, the same model assumptions hold for low and high mutation rates. The assumption that mutation is a byproduct of replication is straightforward for virus populations. One important consequence of this assumption is the factorization of fitness and mutation effects that is indispensable for the fitness landscape concept, which turned out to be very useful in understanding viral infection (see, e.g., Kouyos et al. 2012). In population genetics, for particular in the Crow–Kimura model of asexual evolution (Crow and Kimura 1970, p. 265), replication and mutation are considered as independent events, but there an entirely different mechanism is assumed: Mutation is not related to reproduction and occurs by external action during the whole lifetime of the organism. In order to be able to study evolution of molecules, environmental conditions may be kept constant in the model, but the extension to changing condition is straightforward.

General results derived from the theory of molecular evolution in constant environment are as follows:

(i) In error-free replication,

$$\mathbf{A} + \mathbf{X}_k \rightarrow 2\mathbf{X}_k; \quad k = 1, 2, \ldots \mathrm{n}, \tag{1}$$

selection in the sense of Charles Darwin's survival of the fittest results from chemical reactions approaching a stable stationary state, and is a straightforward consequence of the reaction mechanism. The approach toward stationarity is accompanied by optimization of the mean fitness of the population. Accordingly, the mean fitness of the population $\bar{f}$ is steadily increasing during the selection process, the selected molecular species $\mathbf{X}_m$ is the one with the highest fitness value: $f_m = \max(f_1, f_2, \ldots, f_n)$, and survival of the fittest is

tantamount to optimization of the fitness of the entire population. The final result of selection is unique, a stationary homogeneous population containing only the fittest molecular species $\mathbf{X}_m$, no matter what the initial sequence distribution in the population was (provided it contained $\mathbf{X}_m$ at some, maybe very small amount).

(ii) Errors occurring during the replication process,

$$\mathbf{A} + \mathbf{X}_k \rightarrow \mathbf{X}_j + \mathbf{X}_k; \quad j,k = 1,2,\ldots n;\ j \neq k, \tag{2}$$

produce mutations (Fig. 1) and change the features of correct replication kinetics discussed in (i). After sufficiently long time, the replication–mutation process approaches a stationary state, which does not consist of not a single fittest species $\mathbf{X}_m$ only but is a collective of replicating variants, symbolized by γ. The name "quasispecies" has been coined for this longtime sequence distribution in order to point at the fact that asexual reproduction like sexual reproduction forms genetic reservoirs, which provide pools of variants for future evolution. For a given parameter set, the quasispecies is unique: No matter what the population looked like initially the same longtime sequence distribution will result. The question of fitness optimization is more subtle than in the previous case (i): For most initial conditions, fitness will increase during the replication–mutation process and selection of the quasispecies γ is

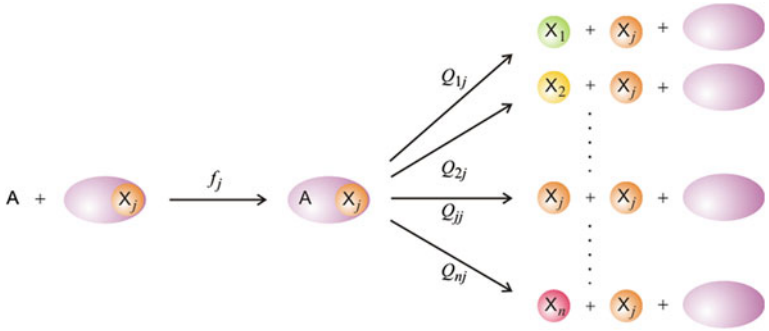

Fig. 1 A molecular view of replication and mutation. The replication device **E** (*violet*), commonly a single replicase molecule—as in polymerase chain reaction (PCR) or in many examples of simple viruses—or a multienzyme complex binds the template DNA or RNA molecule ($\mathbf{X}_j$, *orange*) in order to form a replication complex $\mathbf{E} \cdot \mathbf{X}_j$ and replicates with a rate parameter f_j. During the template-copying process, reaction channels leading to mutations are opened through replication errors. The reaction leads to a correct copy with frequency Q_{jj} and to a mutant $\mathbf{X}_k$ with frequency Q_{kj}. Commonly, we have $Q_{jj} \gg Q_{kj}$ for all $k \neq j$. In other words, correct replication dominates mutant formation. Stoichiometry of replication requires $\sum_{i=1}^{n} Q_{ij} = 1$, since the product has to be either correct or incorrect. The reaction is terminated by full dissociation of the replication complex. The sum of all activated monomers is denoted by **A**. A consequence of the model is factorization of the contributions from fitness and mutation: $w_{kj} = Q_{kj} \cdot f_j$