

V. Sree Hari Rao · Ravi Durvasula *Editors*

# Dynamic Models of Infectious Diseases

Volume 2: Non Vector-Borne Diseases

# Dynamic Models of Infectious Diseases



V. Sree Hari Rao • Ravi Durvasula  
Editors

# Dynamic Models of Infectious Diseases

Volume 2: Non Vector-Borne Diseases



Springer

*Editors*

V. Sree Hari Rao  
Foundation for Scientific Research  
and Technological Innovation  
Hyderabad, AP, India

Ravi Durvasula  
University of New Mexico  
School of Medicine  
Albuquerque, NM, USA

ISBN 978-1-4614-9223-8

ISBN 978-1-4614-9224-5 (eBook)

DOI 10.1007/978-1-4614-9224-5

Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2012939115

© Springer Science+Business Media New York 2013

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media ([www.springer.com](http://www.springer.com))

# Preface

In Volume 1 of “Dynamic Models of Infectious Diseases,” ISBN 978-1-4614-3960-8 and 978-1-4614-3961-5(eBook), we assembled eight chapters from highly acclaimed international scientists who address several of the major insect vector-borne diseases. A diverse and interdisciplinary group of authors was selected with expertise in clinical infectious diseases, epidemiology, molecular biology, human genetics, and mathematical modeling. Indeed, we believe this collection of chapters is unique and should provide a valuable perspective to a wide audience. Though diverse in approach, all the authors address critical elements of disease control. Myriad tools, whether in the realm of molecular engineering, genomic analysis, predictive modeling, or information technology to improve surveillance, are presented in this collection to provide the reader with a current understanding of research methods directed at control of vector-borne diseases.

The present volume is a culmination of a similar effort in the presentation of certain vital non-vector-borne infectious diseases. A variety of intelligent system technology and machine-learning-based methodologies and models have emerged useful in the development of more accurate predictive tools for the early diagnosis of these diseases. There are nine chapters in this book.

Understanding the dynamics of an infectious disease holds the key for designing effective control strategies from a public health perspective. Modeling strategies and effective decision making is the subject of Chap. 1, by V. Sree Hari Rao and M. Naresh Kumar.

This chapter primarily focuses on evolving tools of mathematical modeling as strategies and is generic in nature. Several aspects in the modeling of infectious diseases have been discussed. The authors present new predictive models aimed at better characterization of human susceptibility, disease severity, and control.

Percolation and most epidemics models are concerned with the spatial features of a random subset of some network, with the fundamental difference that while the spread of infectious diseases consists of a dynamical process, the mathematical theory of percolation is concerned with a static random object. In Chap. 2, Alberto Gandolfi has provided a detailed account of the applicability of percolation methods for SEIR epidemic models. This chapter analyzes the role of percolation in

representing or approximating an epidemic model and also discusses the role of percolation in modeling the random networks on which the spread of the infectious diseases takes place.

*Mycobacterium tuberculosis* (*M. tuberculosis*) remains as one of the most dangerous infectious diseases that causes heavy burden on the economy of developing countries and is responsible for more than three million deaths world-wide annually. This is currently considered reemerging and is the subject of discussion in the next three chapters. In Chap. 3, the authors Daniel Okuonghae and Andrei Korobeinikov discuss issues related to the applicability of the WHO-proposed direct observation therapy strategy (DOTS) for the control of this infectious disease with special reference to the developing country Nigeria. This chapter proposes to study the effectiveness of the DOTS strategy through the development of appropriate dynamic mathematical models and also to make practically relevant recommendations for the health authorities.

Classical methods of studying two component systems (TCS) have relied on utilizing biochemical and genetic methods viz. protein characterization and gene inactivation. These methods have been very successful in elucidating the role of various TCSs in bacterial physiology including adaptation to atypical conditions such as pathogenesis. The authors Agrawal, Narayan, and Saini have presented a detailed account of these methodologies in Chap. 4.

Molecular epidemiology tools play an integrative part in tuberculosis control programs in many developed countries at a local and national level. In Chap. 5 Burgos discusses various aspects of *M. tuberculosis*. This chapter presents an overview of the evolution, host-pathogen interactions, and also implications for the control of tuberculosis.

HIV transmission study is the main subject of Chap. 6 by Zhang, Chow, and Wilson. This chapter focusses on knowledge extraction and data modeling based on the trends in HIV transmission according to differences in numbers of sexual partnerships among men who have sex with men in China.

*Cryptococcus gattii* is an emerging infectious disease with an expanding geographic range that gained increased attention in recent years due its frequent outbreaks in North America. This disease is discussed in Chap. 7 by Walraven, Jahng, Davenport, Rane, and Lee. This chapter presents a review on the global and molecular epidemiology, taxonomy, microbial pathogenesis and immunology, and clinical considerations of this emerging fungal pathogen. Further, environmental modeling of the potential ecological niches of *C. gattii* and speculative measures for avoidance and control are discussed.

In Chap. 8, Steinbrück and McHardy discuss the evolutionary dynamics of human viral populations by presenting a technique that leads to the construction of allele dynamics plots (AD-plots) as a method for visualizing the evolutionary dynamics of a gene in a population. The main focus of this chapter rests in the construction of AD-plots for human influenza A viruses such as swine-origin influenza A (H1N1) virus (“swine flu”) and seasonal influenza A (H3N2) virus.

Finally, avian influenza, commonly known as bird flu, is an epidemic caused by H5N1 virus that primarily affects birds such as chickens, wild water birds, ducks,

and swans, and on rare occasions pigs and humans. In recent years this epidemic has emerged as a major global health concern due to its propensity for explosive outbreaks and is the main focus of Chap. 9 by V. Sree Hari Rao and Ranjit Kumar Upadhyay. This chapter aims at developing mathematical models that predict the spread and outbreak diversity of low pathogenic avian influenza virus by presenting a deterministic mathematical model which deals with the dynamics of human infection by avian influenza both in birds and in human, a discrete dynamical model for the spread of H5N1 and the statistical-transmission model of bird flu taking into account the factors that affect the epidemic transmission such as source of infection and social and natural factors.

We have immense pleasure in expressing our appreciation to all those who have directly or indirectly influenced this work. Specifically, we thank all the chapter contributors and the reviewers who untiringly responded to our request by providing useful and thought-provoking reviews. We are grateful to the editorial staff at Springer, New York, for their interest, initiative, and enthusiasm in bringing out this publication. In particular our special thanks go to Ms. Melanie Tucker, Editor, and Ms. Meredith Clinton, Assistant Editor, Springer Science + Business Media, New York, for their very efficient handling of this manuscript.

The first author (VSHR) gratefully acknowledges the research support received from the Foundation for Scientific Research and Technological Innovation (FSRTI)—a constituent division of Sri Vadrevu Seshagiri Rao Memorial Charitable Trust, Hyderabad, India.

The second author (RVD) acknowledges the continued research support provided by the National Institutes of Health (USA), United States Department of Agriculture, and The Bill and Melinda Gates Foundation. Additionally, the support provided by The University of New Mexico School of Medicine and The Raymond G. Murphy Veterans Administration Hospital, both located in Albuquerque, New Mexico, USA, is gratefully acknowledged.

Hyderabad, AP, India  
Albuquerque, NM, USA

V. Sree Hari Rao  
Ravi Durvasula



# Contents

<b>1</b>	<b>Control of Infectious Diseases: Dynamics and Informatics . . . . .</b>	1
	V. Sree Hari Rao and M. Naresh Kumar	
<b>2</b>	<b>Percolation Methods for SEIR Epidemics on Graphs . . . . .</b>	31
	Alberto Gandolfi	
<b>3</b>	<b>Dynamics of Tuberculosis in a Developing Country: Nigeria as a Case Study . . . . .</b>	59
	Daniel Okuonghae and Andrei Korobeinikov	
<b>4</b>	<b>Two-Component Signalling Systems of <i>M. tuberculosis</i>: Regulators of Pathogenicity and More . . . . .</b>	79
	Ruchi Agrawal, Vignesh H. Narayan, and Deepak Kumar Saini	
<b>5</b>	<b><i>Mycobacterium tuberculosis</i>: Evolution, Host–Pathogen Interactions, and Implications for Tuberculosis Control . . . . .</b>	111
	Marcos Burgos	
<b>6</b>	<b>Trends in HIV Transmission According to Differences in Numbers of Sexual Partnerships Among Men Who Have Sex with Men in China . . . . .</b>	147
	Lei Zhang, Eric P.F. Chow, and David P. Wilson	
<b>7</b>	<b>The Impact of <i>Cryptococcus gattii</i> with a Focus on the Outbreak in North America . . . . .</b>	177
	Carla J. Walraven, Maximillian Jahng, Gregory C. Davenport, Hallie Rane, and Samuel A. Lee	
<b>8</b>	<b>Evaluating the Evolutionary Dynamics of Viral Populations . . . . .</b>	205
	Lars Steinbrück and Alice Carolyn McHardy	

<b>9 Modeling the Spread and Outbreak Dynamics of Avian Influenza (H5N1) Virus and Its Possible Control . . . . .</b>	<b>227</b>
V. Sree Hari Rao and Ranjit Kumar Upadhyay	
<b>Index . . . . .</b>	<b>251</b>

# Contributors

**Ruchi Agrawal** Department of Molecular Reproduction, Development and Genetics, Indian Institute of Science, Bangalore, India

**Marcos Burgos** Division of Infectious Diseases, Department of Medicine, New Mexico Veterans Healthcare System, Albuquerque, NM, USA

Division of Infectious Diseases, University of New Mexico Health Science Center, Albuquerque, NM, USA

**Eric P.F. Chow** The Kirby Institute, The University of New South Wales, Sydney, NSW, Australia

**Gregory C. Davenport** University of New Mexico Health Science Center, Albuquerque, NM, USA

**Alberto Gandolfi** Dipartimento di Matematicae Informatica U. Dini, Università di Firenze, Firenze, Italy

**Maximillian Jahng** Division of Infectious Diseases, New Mexico Veterans Healthcare System, Albuquerque, NM, USA

**Andrei Korobeinikov** Centre de Recerca Matematica, Campus de Bellaterra, Bellaterra, Barcelona, Spain

**Samuel A. Lee** Division of Infectious Diseases, New Mexico Veterans Healthcare System, Albuquerque, NM, USA

University of New Mexico Health Science Center, Albuquerque, NM, USA

**Alice Carolyn McHardy** Max-Planck Research Group for Computational Genomics and Epidemiology, Max-Planck Institute for Informatics, Saarbrücken, Germany

Department for Algorithmic Bioinformatics, Heinrich-Heine-University, Düsseldorf, Germany

**Vignesh H. Narayan** Department of Molecular Reproduction, Development and Genetics, Indian Institute of Science, Bangalore, India

**M. Naresh Kumar** National Remote Sensing Center (ISRO), Hyderabad, Andra Pradesh, India

**Daniel Okuonghae** Department of Mathematics, University of Benin, Benin City, Edo State, Nigeria

**Hallie Rane** Division of Infectious Diseases, New Mexico Veterans Healthcare System, Albuquerque, NM, USA

**Deepak Kumar Saini** Department of Molecular Reproduction, Development and Genetics, Indian Institute of Science, Bangalore, India

**V. Sree Hari Rao** Foundation for Scientific Research and Technological Innovation, Hyderabad, Andra Pradesh, India

**Lars Steinbrück** Max Planck Institut Informatik, Saarbrücken, Germany

**Ranjit Kumar Upadhyay** Department of Applied Mathematics, Indian School of Mines, Dhanbad, India

**Carla J. Walraven** University of New Mexico Health Science Center, Albuquerque, NM, USA

**David P. Wilson** The Kirby Institute, The University of New South Wales, Sydney, NSW, Australia

**Lei Zhang** The Kirby Institute, The University of New South Wales, Sydney, NSW, Australia

# Chapter 1

## Control of Infectious Diseases: Dynamics and Informatics

V. Sree Hari Rao and M. Naresh Kumar

### 1.1 Introduction

Information science or informatics plays a crucial role in the design and development of effective strategies to respond, predict, manage, and control the infectious disease outbreaks [5, 8, 6, 9, 47]. Current research in this area includes development of analytical and statistical models to analyze the data for disease outbreak prediction and build disease surveillance systems [17, 23, 24, 25, 26, 28, 37, 50]. The statistical models are data driven and hence their predictive capabilities are restricted to the quality and availability of data from the real-world processes. In contrast the dynamical models are capable of simulating complex phenomena and can generate the required data for statistical analysis. In the present chapter we develop mathematical models for simple infectious diseases and build hybrid procedures for assimilating the data generated from simulations of these models into statistical models that depict improved predictive capabilities.

During an epidemic outbreak due to contact with the infected individuals after a period of time, two groups are formed in the population: those who have not acquired the disease but are likely to contract (susceptible population) and those who are infected (infective population), capable of spreading the disease. The fundamental characteristic of the simple epidemic is that susceptible individuals contract the disease only by getting in contact with the infective individuals. Also, the average time constant or the latency time of the disease depends on the nature of the epidemic. Further, cured individuals do not contract the disease again

---

V. Sree Hari Rao

Foundation for Scientific Research and Technological Innovation, Vadrevu Nilayam,  
13-405, Road No.14, Alakapuri, Hyderabad, Andhra Pradesh 500 035, India  
e-mail: [vshrao@gmail.com](mailto:vshrao@gmail.com)

M. Naresh Kumar (✉)

National Remote Sensing Center (ISRO), Hyderabad, Andhra Pradesh 500 625, India  
e-mail: [mnareshk1@gmail.com](mailto:mnareshk1@gmail.com)

during the same period. Generally, such epidemics will be treated by appropriate vaccination and/or other efforts, which may be viewed as control efforts to contain the spread of disease. The vaccination effort is regarded as a vital parameter in the mathematical models. So it is desirable to have a realistic mathematical model that describes the dynamical interactions between these two groups of populations, namely the susceptible and the infective. For a few earlier studies on epidemiological problems, we refer the readers to [2, 3, 4, 12, 31, 32, 38, 45, 48]. This is the starting point for our investigations in this chapter and accordingly a simplified epidemic model without vital dynamics and relapse is considered. We introduce nonlinearities in the interacting populations through mutual interference parameters. Our main result provides suitable ranges for these parameters.

The present chapter is organized as follows: In Sect. 1.2 we present a mathematical model for a simple infection and study its qualitative properties. We provide a brief account of some useful statistical and data mining methods which will be used in building models from the data, in Sect. 1.3. In Sect. 1.4 the experimental study and the simulation results are discussed. Finally, a discussion is presented in Sect. 1.5.

## 1.2 Models and System Dynamics

In this section we present a model proposed in [39] based on the consideration that the process of the spread of the susceptible population turning into infective population is a nonlinear phenomenon. The following are the underlying biological principles

- (1) The total population is fixed and initially every individual is susceptible to the disease
- (2) The disease is spread through the direct contact of susceptible individuals with the infective individuals
- (3) Every individual who has contracted the disease and has recovered is regarded as immune

These principles when translated into the mathematical framework yield the following system of ordinary differential equations and these equations describe the dynamics of the interacting populations.

$$\begin{aligned} x'_1 &= -\beta x_1^{m_1}(t) x_2^{m_2}(t) - \alpha S(x_1(t)) \\ x'_2 &= \beta x_1^{m_1}(t) x_2^{m_2}(t) - \gamma P(x_2(t)) \end{aligned} \tag{1.1}$$

In Model (1.1)  $x_1$  represents the number of susceptible individuals,  $x_2$  represents the number of infective individuals,  $\beta$  is the infection parameter,  $\gamma$  denotes a parameter related to the average time constant of the disease,  $S$  describes the control input and it is assumed to be proportional to the vaccination effort, and  $P$  corresponds to those individuals who have contracted the disease and are recovered (regarded as individuals with acquired immunity).

The functions  $S$  and  $P$  are assumed to satisfy the following mathematical conditions that arise from epidemiological setting:

$$S(0) \geq 0; \frac{dS}{dx_1} \geq 0; P(0) = 0; \frac{dP}{dx_2} \geq 0 \quad (1.2)$$

The parameters  $m_1$  and  $m_2$  appearing in (1.1) represent the indexes of the interacting populations arising out of the nonlinear considerations of the epidemic phenomenon. Prototypes of the Model (1.1), in which the nonlinear interactions  $m_1 = m_2 = 1$  and when a fixed (constant) control effort is invoked, reduce the above model equations to

$$\begin{aligned} x'_1 &= -\beta x_1 x_2 - u \\ x'_2 &= \beta x_1 x_2 - \gamma x_2 \end{aligned} \quad (1.3)$$

and this model has been studied in [1]. More recent studies on the better vaccination efforts may be found in the papers [31, 38]. Clearly, the present Model (1.1) is more realistic, as the control effort essentially varies with regard to the size of susceptible population rather than being fixed. The Model (1.3) rests on the simple considerations that if each infected individual  $x_2$  converts one susceptible individual into an infective individual, then  $x_2$  susceptible get converted into  $x_1 x_2$  infectives, and this describes the interactions between these populations (simple nonlinear interactions). Our model instead of considering only simple nonlinear interactions also addresses the sublinear interactions between the two populations.

We now consider the model Eq. (1.1) and examine the qualitative properties of its solutions. From the biological point of view, the system in (1.1) describes the dynamical interactions among the two classes of populations such as the susceptible and the infective populations. Clearly, the qualitative study of solutions of this system depends on ensuring conditions that are sufficient to guarantee the existence of unique solutions for initial value problems associated with (1.1). Usually this existence of solutions is determined in a finite interval and the solutions are continued on their maximal intervals. This approach yields continuable solutions. Often the inherent dynamics of the system requires one to pick up a specific point on the trajectory and to move in both forward and backward directions. From this discussion, a mathematical treatment of the model equations requires one to obtain conditions for the existence and uniqueness of continuable solutions for the system given in (1.1). An analogy with ecological problems would render one to regard the community of infective individuals as a sub-population affecting the survival of the susceptible individuals. One aspect of the dynamics of community interactions would be mutual interference among the interacting sub-populations, which in our Model (1.1) is represented by the parameters  $m_1$  and  $m_2$ . It is known that mutual interference is a “stabilizing” process [14, 15, 16]. A question of interest is to describe conditions leading to the persistence/survivability of interacting populations. We observe that the mutual interference introduces sub-linearities

into the system given in (1.1) and it is known that initial value problems for systems with sublinearities have continuable solutions but these solutions are not always unique [10, 21], and hence the underlying system may not be regarded as a dynamical system [16, 36]. Our first result in this direction is to find suitable ranges of values for the parameters  $m_1$  and  $m_2$  so that the solutions of (1.1) form a dynamical system in the sense described above.

Consider the system of equations given by

$$x'_i = g_i(x_1, x_2) \quad (1.4)$$

where  $x_i(0) \geq 0$ , the functions  $g_i : R^+ \rightarrow R^+, R^+ = [0, \infty)$  are continuous for  $i = 1, 2$  that is  $g_i \in C(R^+)$ . Assume that the following conditions are satisfied:

- (H1) There exist constants  $m_j > 0, j = 1, 2$  such that  $h_j \in C(R^+)$  where  $h_j(x_1, x_2) = x_j^{-m_j} g_j(x_1, x_2)$ .  
(H2)  $x_k^{m_k} \frac{\partial}{\partial x_k} h_j(x_1, x_2) \in C(R^+)$  for  $j \neq k = 1, 2$ .

Then the solutions of the system (1.4) form a dynamical system in the sense of [36].

The following change of variables [14] for the system

$$u_1 = x_1^{1-m_1} \quad u_2 = x_2^{1-m_2} \quad (1.5)$$

transforms the system (1) into

$$\begin{aligned} u'_1 &= -(1-m_1) \left[ \beta u_2^{\frac{m_2}{1-m_2}} + S u_1^{\frac{1-m_1}{m_1}} \right] \\ u'_2 &= (1-m_2) \left[ \beta u_1^{\frac{m_1}{1-m_1}} - \gamma P u_2^{\frac{1-m_2}{m_2}} \right], \end{aligned} \quad (1.6)$$

which in turn may be written as

$$\begin{aligned} u'_1 &= -(1-m_1) [\beta x_2^{m_2} + S x_1^{-m_1}] \\ u'_2 &= (1-m_2) [\beta x_1^{m_1} - \gamma P x_2^{-m_2}] \end{aligned} \quad (1.7)$$

To establish that the system (1.7) is a dynamical system we should verify that the hypotheses (H1) and (H2) must hold. This clearly implies that  $\lim_{x_j \rightarrow 0^+} h_j(x_1, x_2)$ , for  $j = 1, 2$  must exist. This implies that  $\lim_{x_1 \rightarrow 0^+} x_1^{-m_1} g_1(x_1, x_2)$ ,  $\lim_{x_2 \rightarrow 0^+} x_2^{-m_2} g_2(x_1, x_2)$ , and  $\lim_{x_1 \rightarrow 0^+} x_1^{-m_1} S(x_1)$  must exist. Also, the initial conditions  $S(0) = 0, S(x_1) > 0, \frac{S(x_1)}{x_1^{m_1}} \rightarrow 0$  as  $x_1 \rightarrow 0$  must be verified. For example, if  $S(x_1) = x_1^p$  with  $p > m_1$  then (H1) is clearly satisfied. Further the initial conditions  $P(0) = 0, P(x_2) > 0, \frac{P(x_2)}{x_2^{m_2}} \rightarrow 0$  as  $x_2 \rightarrow 0$  must also be verified. If  $P(x_2) = x_2^q$  with  $q > m_2$ , then (H1) is satisfied. For (H2) we require that  $\lim_{x_k \rightarrow 0^+} x_k^{m_k} \frac{\partial h_j}{\partial x_k}$  exists. Simple calculations ensure that these limits exist only if  $2m_1 - 1 \geq 0$  and  $2m_2 - 1 \geq 0$ .

Thus, all hypotheses are satisfied if  $m_j \geq \frac{1}{2}$  for  $j = 1, 2$  holds and this clearly ensures that the system (1.1) forms a dynamical system.

Incidence data has been utilized to model the distribution of the data and estimate the parameters of the infectious disease model. Assimilation of data with mathematical models of infectious disease outbreaks is able to generate insight from limited data [29]. In [27] a negative binomial model was used to fit malaria incidence rates as a function of year and location. We refer the readers to [46] for more detailed discussion on logistic growth models. Logistic regression is an example of a binary regression where the covariates say a vector  $\mathbf{X}$ , are related to the probability of the occurrence of the event of interest,  $Prob(Y = 1 | X)$  through a logistic function  $Prob(Y = 1 | X) = \frac{1}{1 + Ce^{-X\beta}}$  as link. A logistic regression model was developed for predicting the dynamics of simple epidemics [40] based on the generalization of fitting a logistic curve [34]. The rate of spread of an infectious disease is proportional to the number of infected individuals  $X$  constrained by the total population  $S$ . The dynamics of the spread is modeled as

$$\frac{dX}{dt} = kX(S - X). \quad (1.8)$$

Upon further simplifications we write the solution of (1.8) as

$$X(t) = \frac{S}{e^{-(Sk_t + c_1)}} \text{ that is } X = \frac{S}{1 + Ae^{-skt}}$$

where  $A = e^{-c_1}$ . Utilizing incidence data for different diseases we estimate the parameters of the Model (1.8) in a later section.

## 1.3 Informatics

Several evolutionary computing methods have been used widely for the analysis of infectious data to predict the vital characteristics of transmission dynamics of the disease. For example, genetic algorithms and support vector machine (SVM) were employed in [13, 51], multiple linear regression in [22], artificial neural networks are employed in [49, 39]. In this section we present the statistical and data mining methods employed in this chapter for predictive modeling of infectious diseases.

### 1.3.1 *Multiple Linear Regression*

Multiple linear regression (MLR) models the relationship between two or more independent variables (predictors) and a dependent variable (predictand) by fitting a linear equation to observed data. Every value of the independent variable  $x$  is associated with a value of the dependent variable  $y$ . The MLR fit is based on

minimization of errors between the observed and predicted in least square sense. The MLR model for a given set of  $n$  observations is expressed as  $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} + \epsilon_i$  for  $i = 1, 2, \dots, n$ .

The observed values for  $y$  vary about their means  $\mu_y$ , resulting in a variation parameter  $\epsilon_i$  in the MLR model. We can express the model in terms of

$$\text{Observed data} = \text{FIT} + \text{RESIDUAL},$$

where the *FIT* represents the expression  $\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p$  and *RESIDUAL* represents the deviations of the observed values  $y$  from their mean.

### 1.3.2 K-Means Clustering

The k-means clustering works on the expectation of maximization algorithm to find the centers of natural clusters in the data. It assumes that the object attributes form a vector space. The objective is to minimize total intra-cluster variance, or, the squared error function

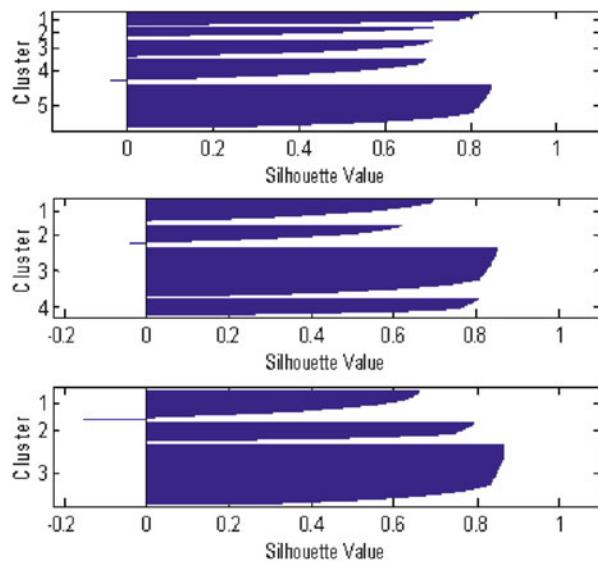
$$V = \sum_{i=1}^k \sum_{x_j \in S_i} [x_j - \mu_i]^2$$

where there are  $k$  clusters  $S_i$ ,  $i = 1, 2, \dots, k$ , and  $\mu_i$  is the centroid or mean point of all the points  $x_j \in S_i$ . The initial number of clusters is specified as three based on the visual inspection of the data sets, for generating clusters from the data using k-means clustering. K-means clustering is a parametric technique where it is required to provide the number of clusters as a parameter. To know how well-separated the resulting clusters are, a Silhouette plot is constructed using cluster indices output from k-means. The Silhouette plot displays a measure of how close each point in one cluster is to points in the neighboring clusters. This measure ranges from  $+1$ , indicating points that are very distant from neighboring clusters, through  $0$ , indicating points that are not distinctly in one cluster or another, to  $-1$ , indicating points that are probably assigned to the wrong cluster. From Fig. 1.1 it is derived that the optimal number of clusters is three as Silhouette plots of four and five clusters show very low Silhouette values.

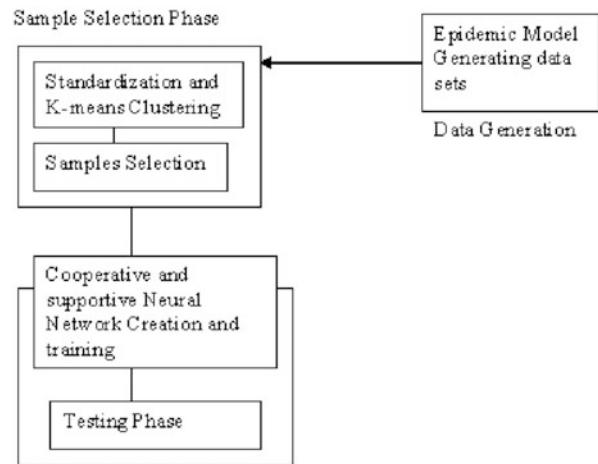
### 1.3.3 Cooperative and Supportive Neural Networks

Cooperation and support among neurons existing in a network which contribute to their collective capabilities and distributed operations constitutes a cooperative and supportive neural networks (CSNN). The activation dynamics of CSNN and their properties are discussed in [44]. A learning algorithm with a k-mean preprocessing has been utilized in [39] to estimate the rate of spread of epidemics.

**Fig. 1.1** Silhouette plot with five, four, and three clusters generated using k-means clustering



**Fig. 1.2** Operating mechanism for the proposed learning paradigm for cooperative networks



The operating mechanism for training a neural network is shown in Fig. 1.2. The neural network architecture is depicted in Fig. 1.3. Algorithm 1 provides the mechanism for building and training a CSNN with preprocessing.

---

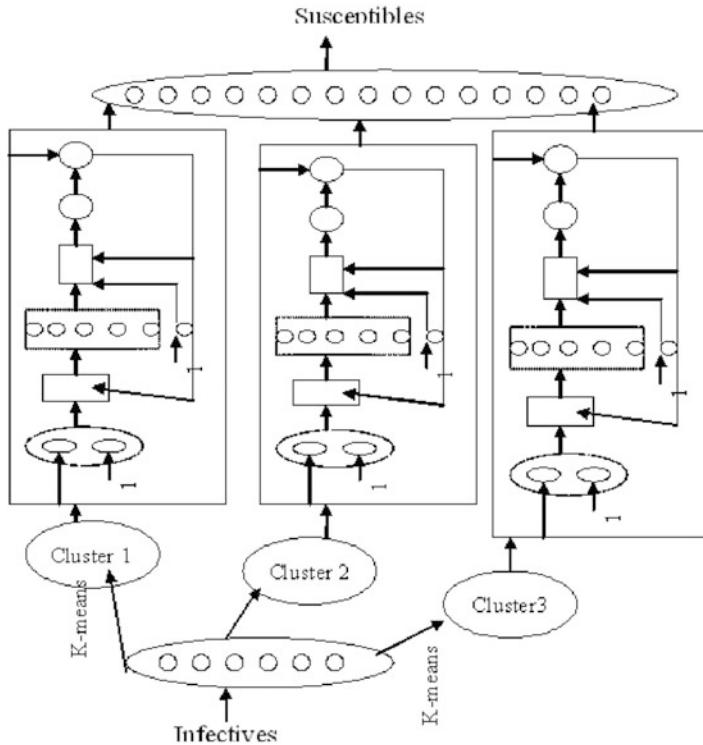
**Algorithm 1:** Algorithm for building a CSNN

---

**Input:**

Data set  $S$  containing susceptible and infective population measured over a time interval.

(continued)



**Fig. 1.3** Cooperative and supportive multi-layer feed-forward neural network architecture for computing susceptible for a given infective population

---

**Algorithm 1:** (continued)

---

**Output:** Trained CSNN for estimating susceptible population.

**Algorithm**

- (1) Identify the number of clusters  $n$  from the data set using the Silhouette plots.
  - (2) Generate a data set  $P_j$  from data set  $S$  where  $j = 1, \dots, n$  denotes the number of clusters using k-means clustering.
  - (3) Split each data set in  $P_j$  into training and validation data sets.
  - (4) Corresponding to each of the data set  $P_j$  construct a neural network having configuration 1-5-1 for estimating the number of susceptible individuals.
  - (5) Using a conjugate gradient descent algorithm train each neural network.
  - (6) For each test record, the cluster centers are matched with the corresponding data record and the related neural network in CSNN is activated to obtain the required estimation.
  - (7) Return CSNN
  - (8) END
-

### 1.3.4 Ensemble of Neural Networks

A finite collection of neural networks employed for constructing a solution for a given task is known as an ensemble of neural networks [20]. Each neural network in the ensemble learns a different pattern from the same task, thereby improving the predictive capabilities of the classifiers. In general, a neural network ensemble consists of two steps, namely (1) training a number of component neural networks and (2) combining the predictions of each component. The ensemble neural networks have been applied in different fields such as recognition of face [18, 53, 54], infectious diseases [41], optical character recognition [52], medical diagnosis [7, 11, 43], remote sensing image classification [19], and seismic signals classification [30]. The procedure for building an ENN for estimating the parameters of the infectious disease model is presented in Algorithm 2.

#### 1.3.4.1 New ENN Algorithm

---

##### Algorithm 2: The NENN Algorithm

---

**Input:** Data set  $S(m, n)$  where  $m$  and  $n$  are number of records and attributes, respectively

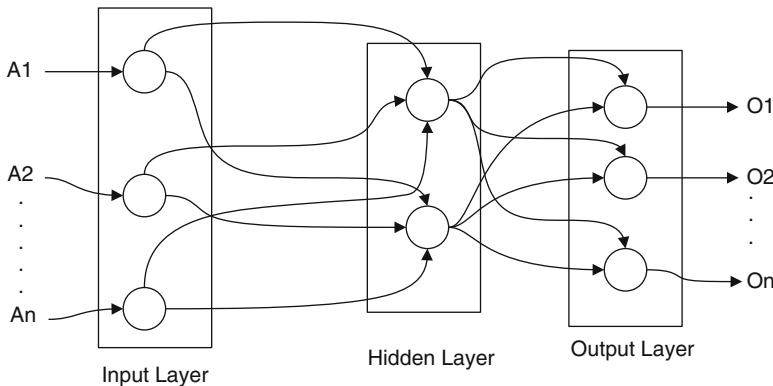
**Output:** Ensemble neural network classifier  $E$  consisting of trained multi-layered perceptron neural network classifiers.

#### Algorithm

- (1) Generate a data set  $S'$  from data set  $S$  where  $S' \subseteq S$ .
- (2) Identify and collect all records in a data set  $S'$  and split them into training and validation data sets using a  $k$  fold cross validation procedure. Denote the training and testing data set by  $T_k$  and  $V_k$ , respectively.
- (3) Consider a record in the training data pertaining to a particular cross fold  $T_1 \in T_k$ .
- (4) Initialize an Ensemble Neural Network  $E \leftarrow \emptyset$ 
  - (a) Construct a multi-layered feed-forward neural network  $N$  and train using the Levenberg-Marquardt back-propagation algorithm and data set  $T_1$
  - (b) Add the neural network classifier constructed built in 3 having an input layer, a hidden layer, and an output layer to the Ensemble initialized in 3  $E \leftarrow E \cup N$
  - (c) Consider a record  $V_1$  from the validation set  $V$ 
    - (i) simulate the outputs from the neural network built in step 3 using the  $V_1$  as inputs
    - (ii) train the ensemble neural network built in 3 using the Levenberg–Marquardt back-propagation algorithm

---

(continued)



**Fig. 1.4** A multi-layer perceptron architecture

---

**Algorithm 2:** (continued)

---

- (iii) Compute mean squared error (MSE)  $M_j$  of the ensemble neural network
  - (d) Repeat the steps in 3 for each of the validation data elements in the set  $V$
  - (e) Compute the average  $M_{avg} = \frac{\sum_{j=1}^k M_j}{k}$  of the ensemble neural network classifier  $E$
  - (f) Retain the network classifier built in 3 into the ensemble  $E$  if there is an improved performance over the previously built ensemble neural network.
  - (5) Repeat the steps in 3 for each of the training sets in  $T$ .
  - (6) RETURN  $E$
  - (7) END
- 

Figure 1.4 shows a multi-layer perceptron used as a component in the ENN architecture. A novel ENN architecture is shown in Fig. 1.5. The architecture is capable for learning both classification and regression tasks.

The verification and validation procedure for the ENN is given in Algorithm 3.

#### 1.3.4.2 Verification and Validation of Neural Networks

---

**Algorithm 3:** Algorithm for Verification and Validation of Neural Networks

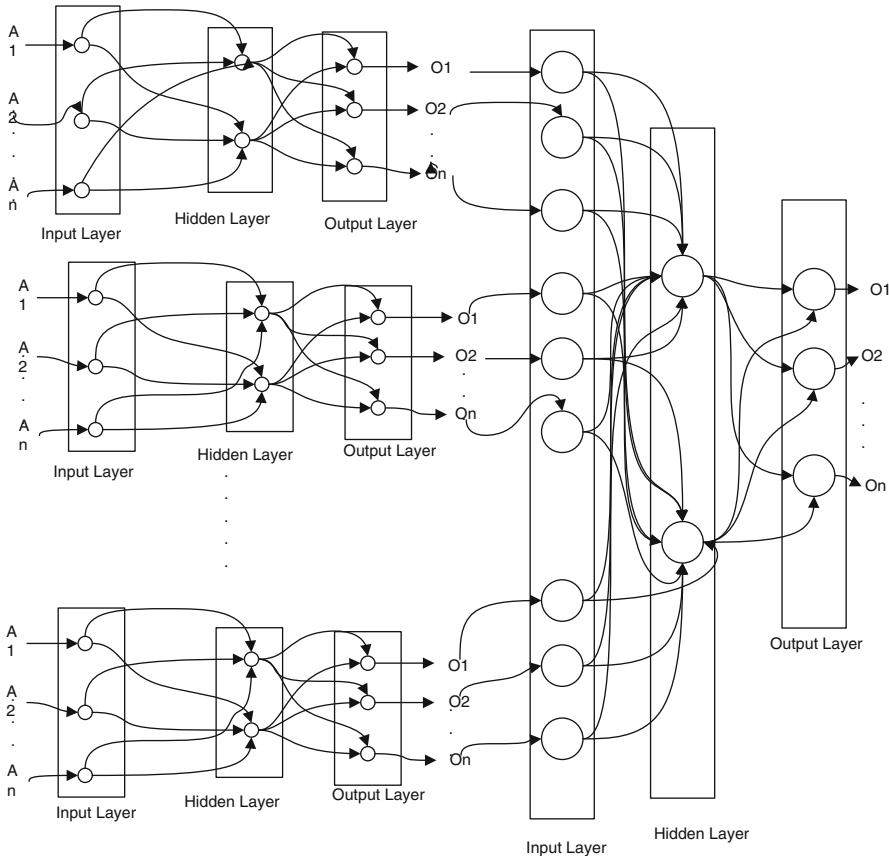
---

Input:

- (a) Data set  $S(m, n)$  where  $m$  and  $n$  are number of records and attributes, respectively

---

(continued)



**Fig. 1.5** A novel ensemble neural network architecture-RNENN

---

**Algorithm 3:** (continued)

---

- (b) Ensemble neural network  $E$  consisting of trained multi-layered perceptron neural network

**Output:** Average MSE  $M_{avg}$  of the ensemble neural network  $E$ .

Algorithm

- (1) Generate a data set  $S'$  from data set  $S$  where  $S' \subseteq S$
- (2) Identify and collect all records in the data set  $S'$  and split them into training and validation data sets using a  $k$  fold cross validation procedure. Denote the training and testing data sets by  $T_k$  and  $V_k$ , respectively.

---

(continued)

**Table 1.1** Wilcoxon sign rank statistics for paired samples

Method	Rank sums (+, -)	Test statistics	Critical value	p-value
Bagging	105.0,0,0	0.0	22	0.00012
Stacking	105.0,0,0	0.0	22	0.00012
C4.5	105.0,0,0	0.0	22	0.00012
MLP	105.0,0,0	0.0	22	0.00012
Logistic	104.0,1,0	1.0	22.0	0.00024
RBF	105.0,0,0	0.0	22	0.00012
BayesNet	105.0,0,0	0.0	22	0.00012
Naive Bayes	105.0,0,0	0.0	22	0.00012

---

**Algorithm 3:** (continued)

---

- (3) Consider a record in the validation set pertaining to a particular cross fold  $V_1 \in V_k$
  - (4) Measure the outputs from the ensemble neural network classifier using  $V_1$  as an input.
  - (5) Compute the MSE  $M_j$  of the ensemble neural network classifier
  - (6) Repeat steps (3)–(5) for each element in the set  $V_k$
  - (7) Compute the average MSE  $M_{avg} = \frac{\sum_{j=1}^k M_j}{k}$  for  $k$  cross fold validation data sets
  - (8) RETURN  $M_{avg}$
  - (9) END
- 

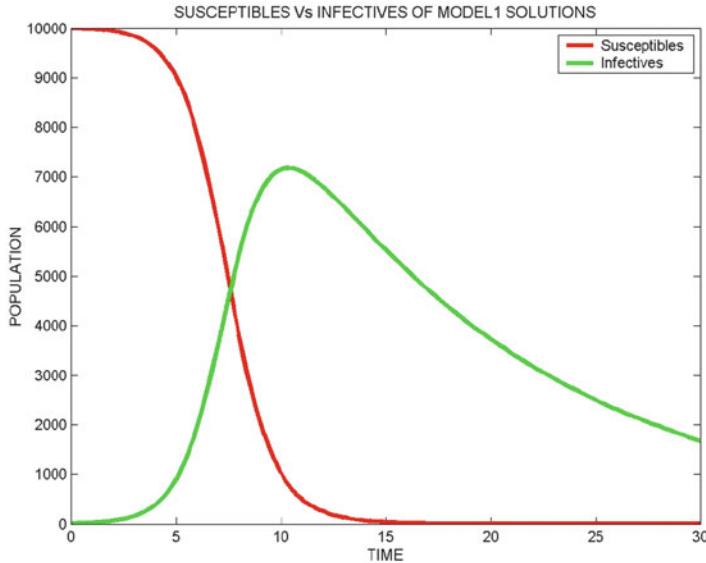
The ENN architecture is employed as a classifier on different bench mark data sets in [42]. The results are compared with other ensemble methods using Wilcoxon tests shown in Table 1.1. The ENN architecture has been employed for regression problems in [41].

## 1.4 Experiments and Results

In this section we discuss the methodologies employed in simulations and estimations of parameters using informatics.

### 1.4.1 Simulations on Infectious Disease Models

The first part of this section deals with the results of the Model (1.1) simulated for various admissible values of the parameters  $m_1$  and  $m_2$ . Model (1.1) considers a situation in which sublinearity between the susceptible and infective populations is not taken into account and with fixed vaccination effort.



**Fig. 1.6** Interactions between susceptible and infective population using Model 1

#### 1.4.1.1 A Bilinear Incidence Model

This example deals with the case in which  $m_1 = m_2 = 1$ ,  $u=10$ ,  $\beta = 0.0001$ ,  $\gamma = 0.8$ , susceptible population = 10,000, and infective population = 10. This corresponds to the model studied in [1], given by

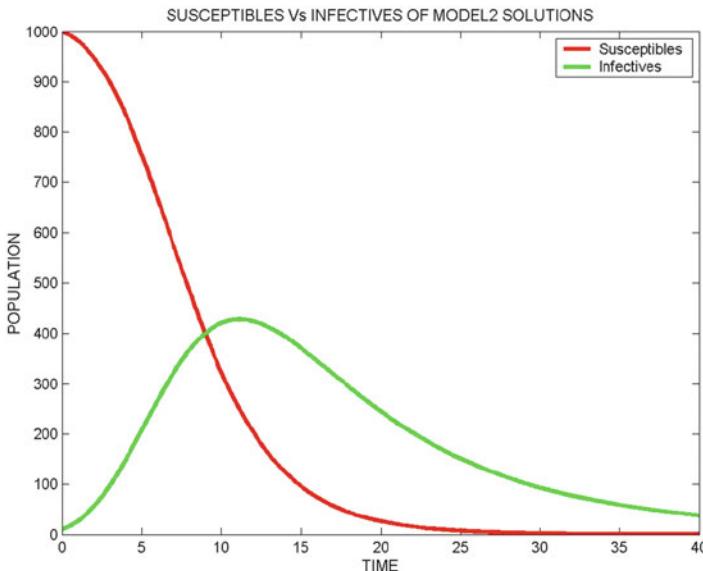
$$\begin{aligned} x'_1 &= -\beta x_1 x_2 - u \\ x'_2 &= \beta x_1 x_2 - \gamma x_2 \end{aligned} \quad (1.9)$$

Figure 1.6 shows interactions between the susceptible and the infective populations generated using Matlab<sup>©</sup> software.

#### 1.4.1.2 Model with Sublinear Interactions

$$\begin{aligned} x'_1 &= -\beta x_1^{m_1} x_2^{m_2} - S(x_1) \\ x'_2 &= \beta x_1^{m_1} x_2^{m_2} - \gamma P(x_2) \end{aligned} \quad (1.10)$$

This model is an improvement over the bilinear incidence model in the sense that the vaccination effort is based on the infection rate and is dynamic in character. Also, the interactions between the populations are not necessarily linear. This model has the following parameters  $m_1 = 0.8$ ,  $m_2 = 0.7$ ,  $S(x) = (x_1^{0.4}/(\text{vaccination effort} + x_1^{0.4}))$ ,  $P(x) = x_1^{1.2}$ ,  $\beta = 0.01$ ,  $\gamma = 0.04$ , initial susceptible = 1,000, and initial infective = 10.



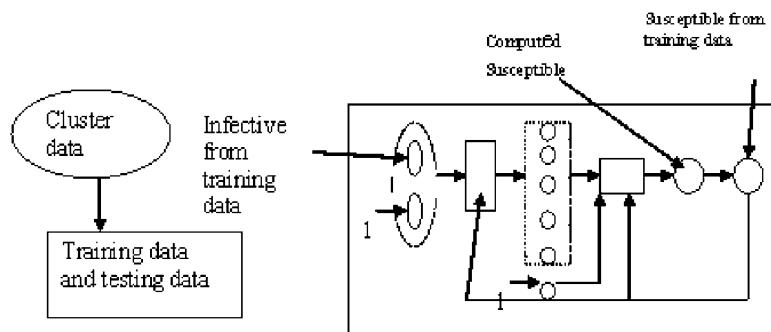
**Fig. 1.7** Interactions between susceptible and infective individuals during the spread of the epidemic using numerical simulation in Matlab for Model (2)

The interactions between the susceptible and infective populations are shown in Fig. 1.7 which has been generated using Matlab<sup>®</sup> software. The vaccination effort has rendered decline in the infective population well ahead, as observed in Model (1.4.1.1). This clearly establishes that the vaccination effort given by our Model (1.4.1.2) is more realistic.

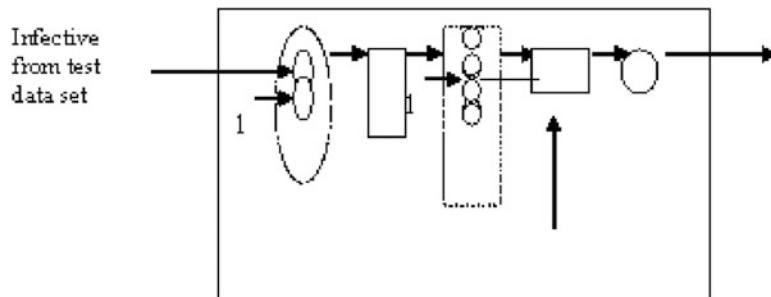
#### 1.4.2 Estimation of Transmission Rate Using CSNN

Initially the entire susceptible and infective populations are standardized to mean zero and standard deviation one. Using k-means clustering algorithm the data sets are clustered. The set of susceptible and infective populations in each cluster is then partitioned into training data and testing data. The training data is given as input to the neural network. When the training is complete the testing data is given as input and outputs are computed.

The three-layer neural network architecture with one input, five hidden neurons, and one output is considered for training. Each layer has one bias node. The tan hyperbolic function is used as a threshold function for each of the neurons. The network is trained using the back-propagation learning algorithm with momentum, which has been found effective. The procedure and the network architecture designed are shown in Fig. 1.8. The data used for training is generated from the mathematical model given in Eq. (1.1) through numerical solutions using Matlab<sup>®</sup>.



**Fig. 1.8** Neural network architecture for training



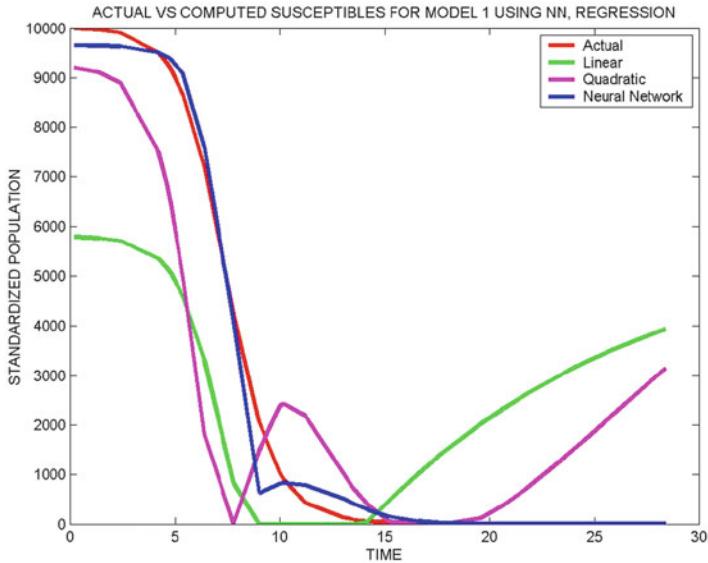
**Fig. 1.9** Testing procedure for the estimation of the susceptible given the infective as input

The network is trained to predict the number of susceptible population given the number of infective population to the input neuron. The purpose of predicting susceptible is to estimate the rate of spread of the epidemic. During the training phase the error generated due to differences in the predicted and actual susceptible populations is propagated in the backward direction and weights are adjusted. The network is trained once the mean square error reaches a defined required value. When the training is complete the infective in the testing data set is given as input and the susceptible are computed. The structure of the trained network to compute the susceptible population for a given infective population is shown in Fig. 1.9.

The training and testing are done for each of the clusters and the susceptible are computed for the test data sets for all the clusters.

#### 1.4.2.1 Neural Network Training

We have applied statistical methods such as regression analysis for Model 1 and compared the performance of the neural network with the statistical methods. Neural network architecture with configuration 1-5-1 is designed. When k-means clustering is applied to the input data three clusters are obtained. Each individually



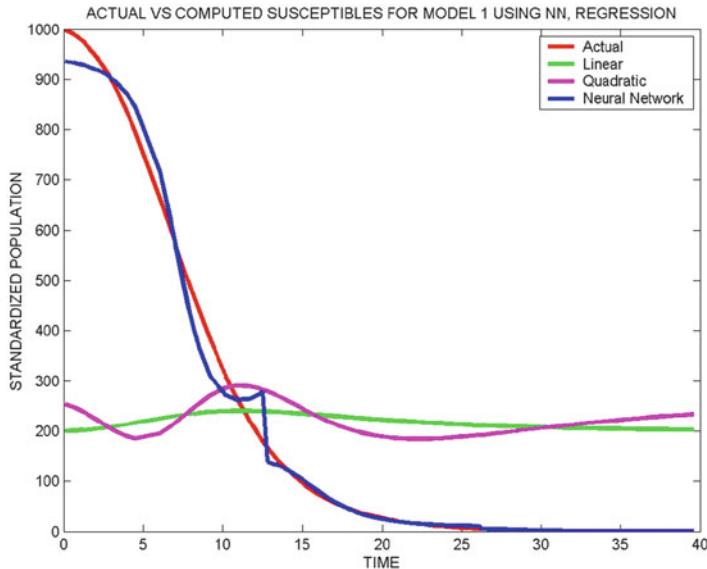
**Fig. 1.10** Susceptible individuals computed from neural network and regression analysis giving infective as inputs for Model 1

clustered data is given as input to the three neural networks for training using back-propagation with momentum learning algorithm. Also, a linear and quadratic regression analysis is carried out and the results are given in Fig. 1.10. Figure 1.10 shows the actual susceptible and the susceptible estimated from the neural network and regression analysis. It may be seen that the estimation by the neural network is better when compared with that obtained by using regression method.

Neural network architecture model described in Sect. 1.3.3 is considered for training the network. Linear and quadratic regression analysis is carried out on this data and the results are plotted in Fig. 1.11. From Fig. 1.11 it may be observed that the neural network estimation of susceptible is better than those obtained by the regression method.

#### 1.4.2.2 Comparison of Estimated Transmission Rates

In this section, we employ a neural network architecture and estimate the rate of spread of the epidemic and compare the same with the actual calculated rate. The rate is the product  $\beta x_1^{m_1} x_2^{m_2}$ .



**Fig. 1.11** Susceptible individuals computed from neural network and regression analysis with infective population as input for Model 2

### Model with Bilinear Incidence

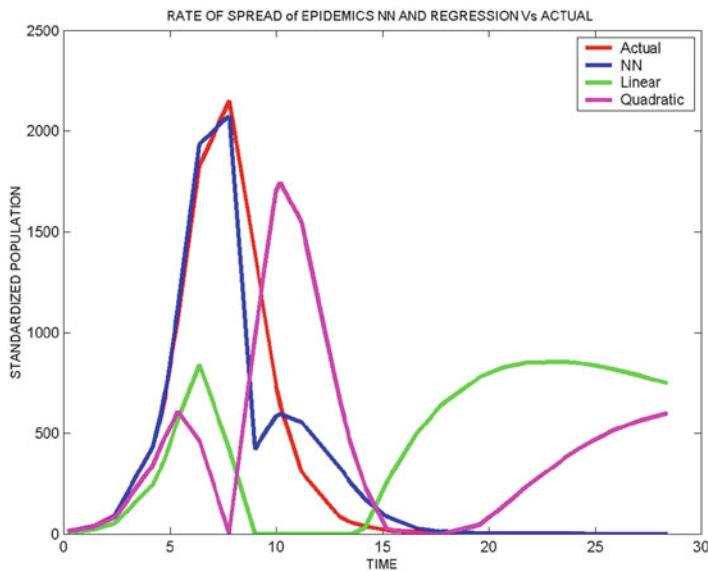
This example shows the rate calculated from a neural network and the actual calculated rate ( $\beta x_1^{m_1} x_2^{m_2}$ ) are in good agreement than that obtained from polynomial regression. Simulations are conducted with the following parameter values  $m_1 = m_2 = 0$ ,  $\beta = 0.001$ .

Figure 1.12 shows that the rate estimated by the neural network is very close to the actual rate compared to the one estimated by using the regression methods.

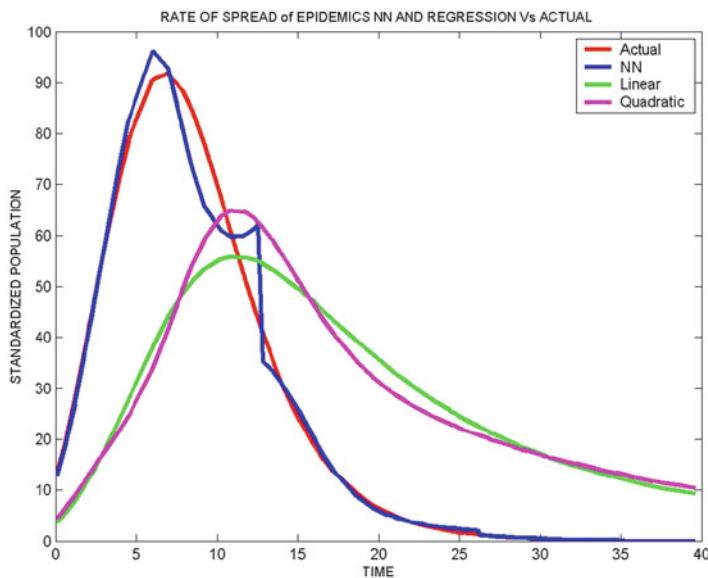
### Model with Sublinear Interactions

In this example Model 2 is considered with mutual interference parameters and the rate of spread of the epidemic is calculated from the neural network and also the polynomial regression method for the following values of the parameters:  $m_1 = 0.8$ ,  $m_2 = 0.7$ ,  $\beta = 0.01$ .

From Fig. 1.13 it is observed that the results obtained through the neural network are closer to the actual results than those obtained by using the polynomial regression analysis.



**Fig. 1.12** Estimated rate from neural network, regression analysis, and actual calculated rate of spread of epidemics Model 1



**Fig. 1.13** Estimated rate from neural network, regression analysis, and actual calculated rate of spread of epidemics Model 2

**Table 1.2** Performance of individual NN in the ENN created from each of the cross folds

NN	MSE	Slope	Y intercept	Regression R-value
1	0.031084	0.745238	0.090212	0.816524
3	0.030431	0.747684	0.094584	0.796353
4	0.029207	0.751159	0.089058	0.877135
8	0.028659	0.751337	0.090126	0.872514
10	0.029132	0.771177	0.082494	0.905537

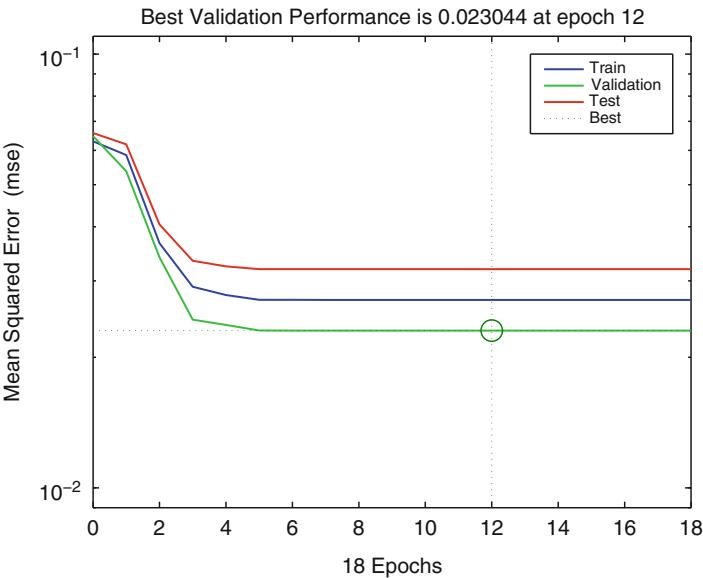
**Table 1.3** Validation of ENN using ten fold cross validation

Fold	R-value	MSE
1	0.869159	0.030800
2	0.863576	0.031113
3	0.863494	0.031573
4	0.866678	0.031119
5	0.868970	0.030366
6	0.868909	0.030844
7	0.865487	0.031410
8	0.869981	0.030334
9	0.870395	0.030167
10	0.863315	0.031335

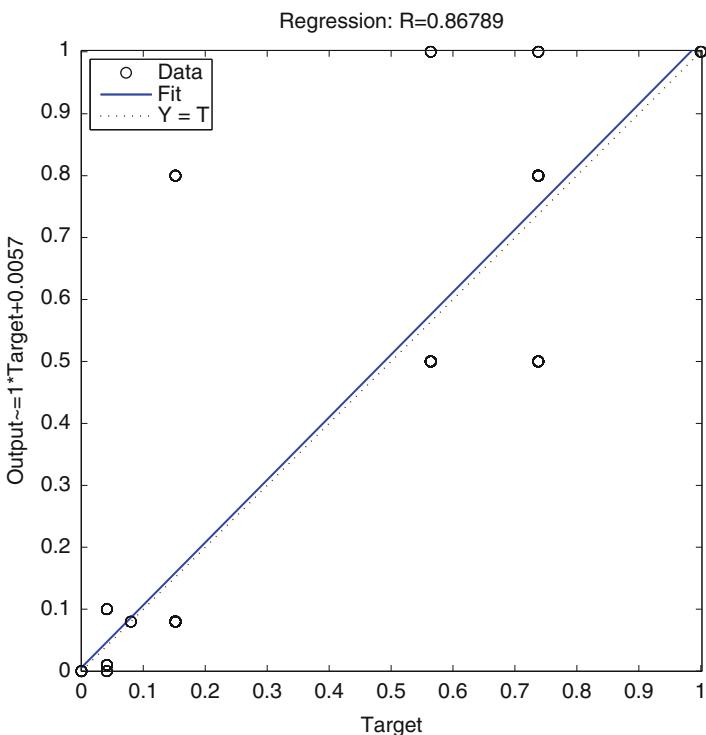
#### 1.4.3 *Estimation of Parameters of the Mathematical Model Using ENN*

The data set obtained from the simulation discussed in Sect. 1.4 is utilized to build an ENN. The data sets are divided into training and validation using a holdout method of cross validation given in Matlab<sup>©</sup> in with 30 % allotted for validation. The training data is utilized to create training and testing data for each of the cross folds. The training and testing data sets comprise of incidence data (infectives) in the input and parameters  $\beta$ ,  $\gamma$ ,  $m_1$ ,  $m_2$  in the output.

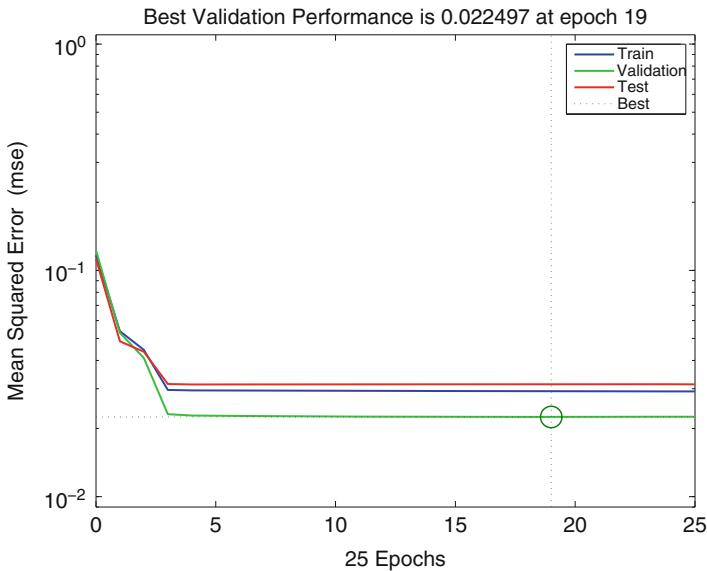
The training and building of ENN is done utilizing the Algorithm 2. Table 1.2 shows the performance of the NN built over each of the cross folds. To have a robust ensemble, the NN models whose R-squared value is greater than or equal to 0.85 are only included. The validation of ENN using is carried out using the Algorithm 3. Table 1.3 gives the R-value and the mean squared error (MSE) of the ENN for each of the validation data provided in the cross fold. The authors in [37] obtained an average R-value from  $10 \times 10$  fold cross validation as 0. 866996 and MSE of 0. 030906. Figures 1.14–1.23 show the performance and regression plots of NN included in the ENN.



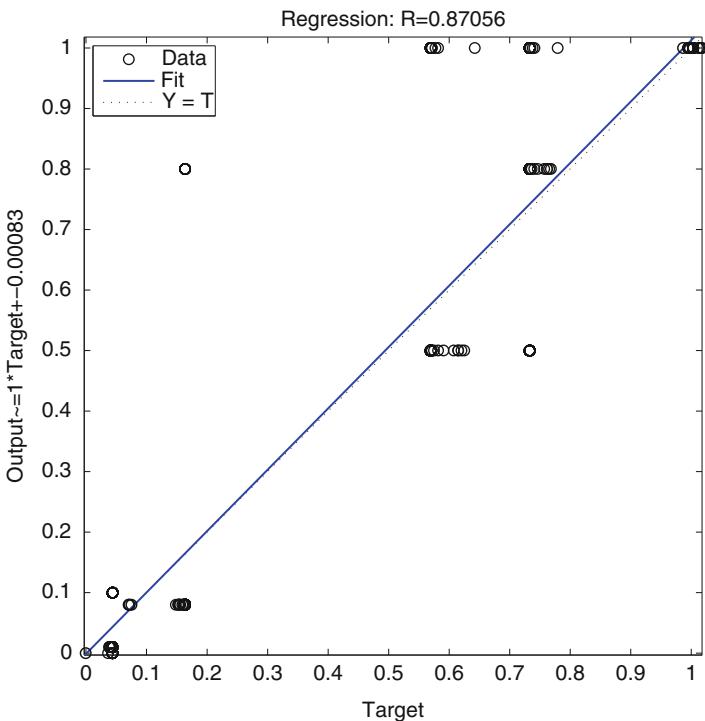
**Fig. 1.14** Performance of NN built over data in Crossfold 1



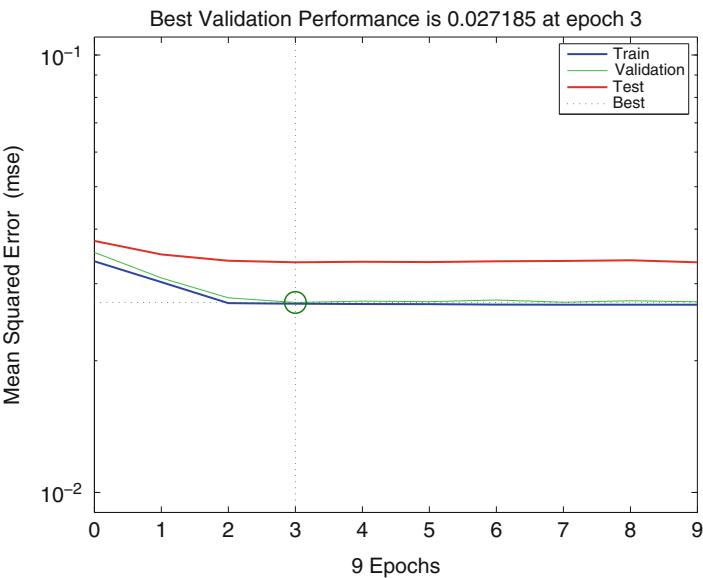
**Fig. 1.15** Regression plot of NN built over data in Crossfold 1



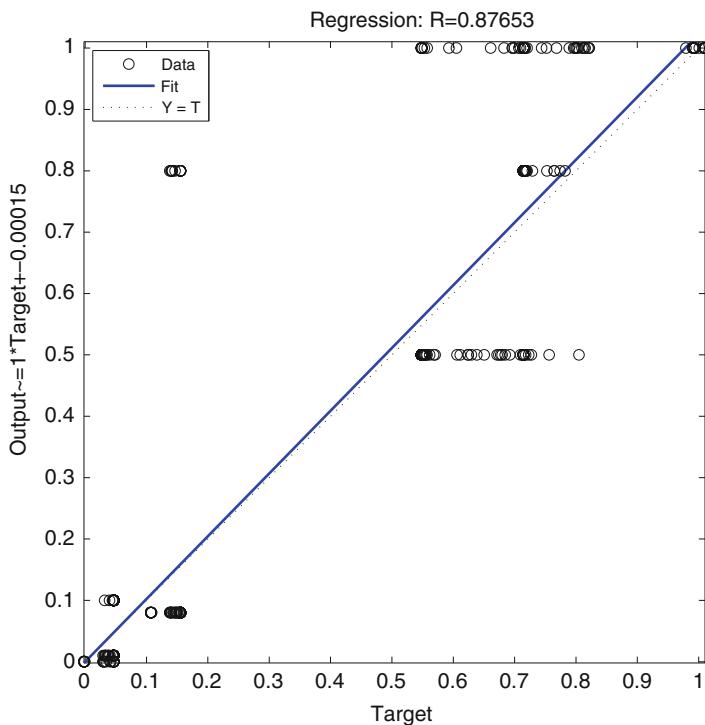
**Fig. 1.16** Performance of NN built over data in Crossfold 3



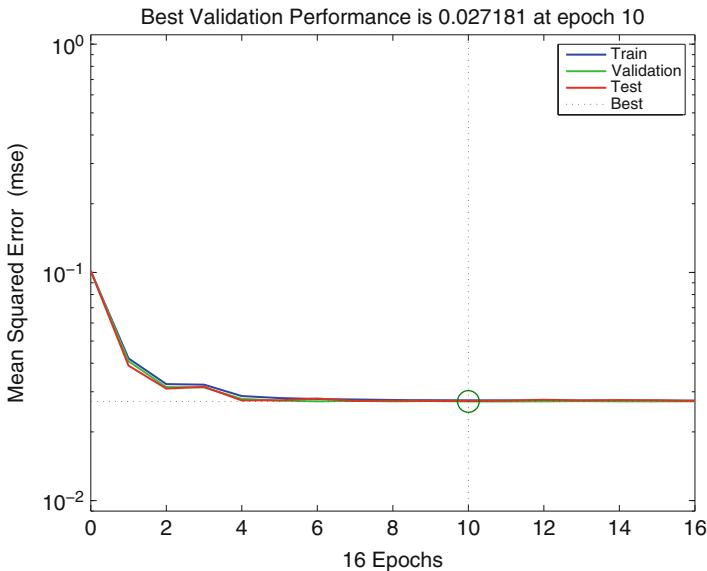
**Fig. 1.17** Regression plot of NN built over data in Crossfold 3



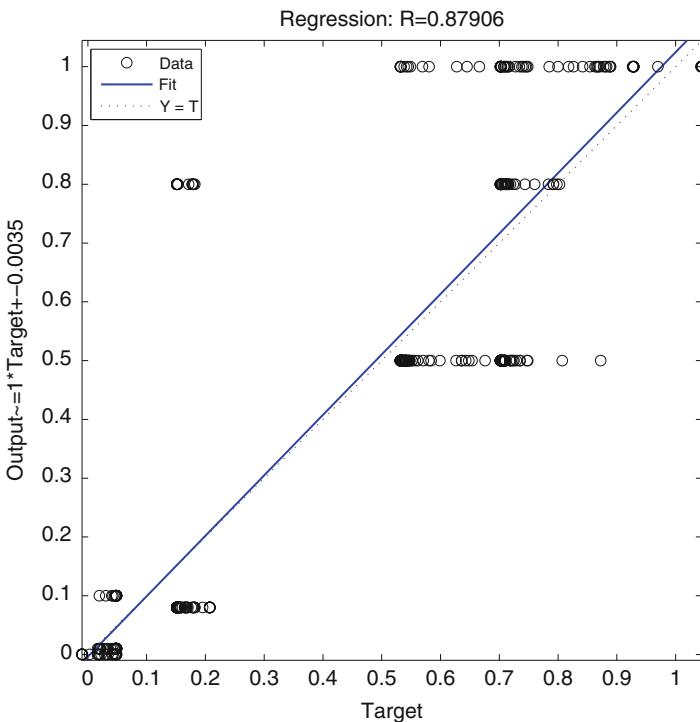
**Fig. 1.18** Performance of NN built over data in Crossfold 4



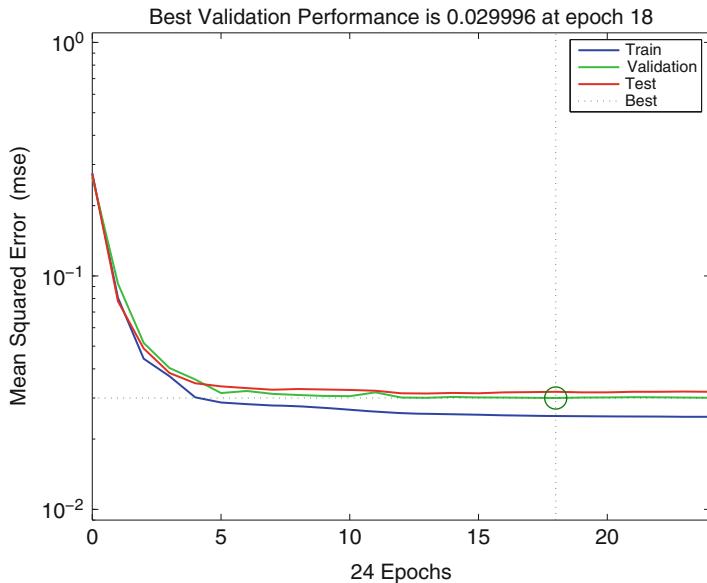
**Fig. 1.19** Regression plot of NN built over data in Crossfold 4



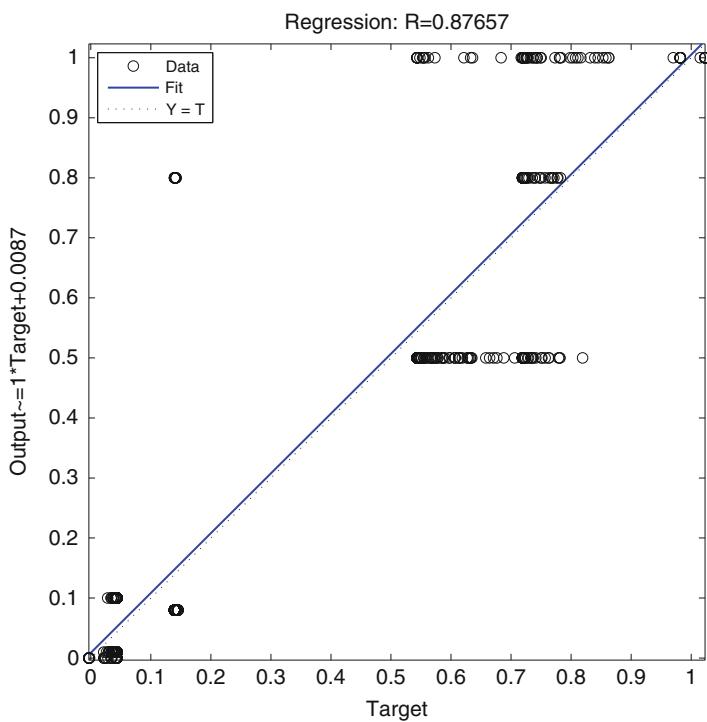
**Fig. 1.20** Performance of NN built over data in Crossfold 8



**Fig. 1.21** Regression plot of NN built over data in Crossfold 8



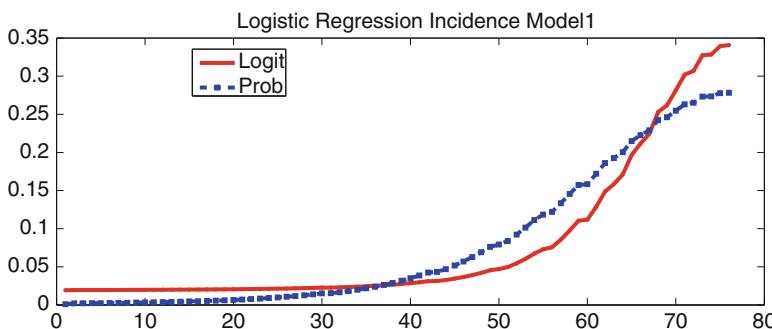
**Fig. 1.22** Performance of NN built over data in Crossfold 10



**Fig. 1.23** Regression plot of NN built over data in Crossfold 10

**Table 1.4** Parameters of the Logistic Regression Function

Case	B1	B2	Model
case (i)	-3.9413	0.0012	$Y = \frac{1}{1 + e^{-(3.9413+X*0.0012)}}$
case (ii)	-2.6241	0.0005	$Y = \frac{1}{1 + e^{-(2.6241+X*0.0005)}}$
case (iii)	-2.9121	0.0007	$Y = \frac{1}{1 + e^{-(2.9121+X*0.0007)}}$
case (iv)	-2.8192	0.0006	$Y = \frac{1}{1 + e^{-(2.8192+X*0.0006)}}$

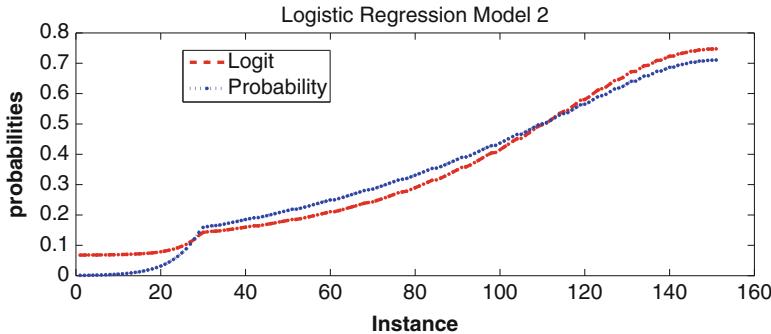
**Fig. 1.24** Estimated probabilities using logistic regression and actual probabilities using Model 1

#### 1.4.4 Mathematical Model from Incidence Data

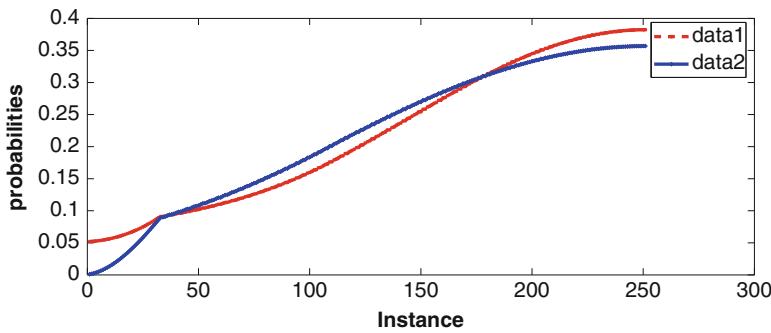
The data required for simulation is generated by numerically solving the equations of the model Eq. (1.3) under different cases given in Table 1.4. The data set has attributes population,  $m_1$ ,  $m_2$ ,  $\beta$ ,  $\gamma$ , type (susceptible, infective). The mathematical model described in chapter Eq. (1.1) is considered for generating the continuous data sets for analysis. We consider the following cases for simulations

- case (i)  $\beta = 0.01$ ,  $\gamma = 0.8$ ,  $p = 1$ ,  $q = 0.5$
- case (ii)  $\beta = 0.0001$ ,  $\gamma = 0.05$ ,  $p = 1$ ,  $q = 1$
- case (iii)  $\beta = 0.1$ ,  $\gamma = 0.08$ ,  $p = 0.5$ ,  $q = 0.5$
- case (iv)  $\beta = 0.01$ ,  $\gamma = 0.08$ ,  $p = 0.8$ ,  $q = 0.5$

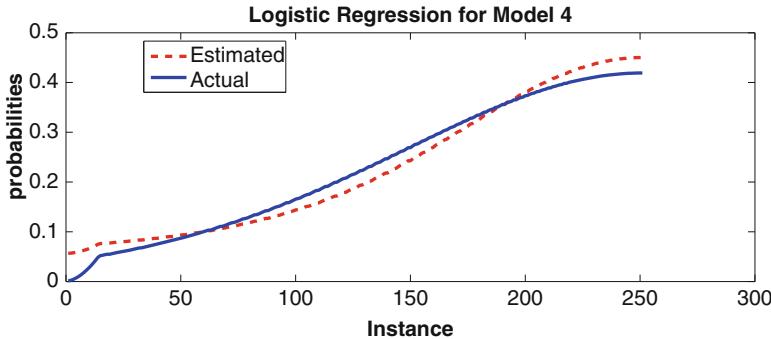
Given the total susceptible population ( $N$ ), number of infective ( $X$ ), probability of infection ( $Y = \frac{X}{N}$ ) for each of the above cases are computed. These computations are used in estimating the parameters of logistic function utilizing generalized linear models [33, 35] implemented in Matlab<sup>®</sup>. We then obtain the mathematical model as a logistic differential equation for each of the above cases. We have



**Fig. 1.25** Estimated probabilities using logistic regression and actual probabilities using Model 2



**Fig. 1.26** Estimated probabilities using logistic regression and actual probabilities using Model 3



**Fig. 1.27** Estimated probabilities using logistic regression and actual probabilities using Model 4

utilized a binomial distribution with logit link function to obtain the parameters of the logistic function

$$B = \text{glm fit}(X, [\text{Yones}(N, 1)]', \text{'binomial', 'link', 'logit'}) \quad (1.11)$$

**Table 1.5** Comparison of actual probabilities and logistic regression model

Case	R <sup>2</sup>	RMSE
case (i)	0.9217	0.02712
case (ii)	0.9732	0.03675
case (iii)	0.9663	0.02035
case (iv)	0.9719	0.02218

The parameters obtained for the above cases are shown in Table 1.4. The estimated probabilities using logistic regression and actual probabilities for all cases are shown in Figs. 1.24–1.27.

A linear fit is done between the estimated probabilities and actual probabilities to obtain the R-squared and root mean squared error values (RMSE). Table 1.5 shows the evaluation of the logistic regression for building the model for estimating the rate of spread of the disease.

## 1.5 Discussion

In this chapter, we have introduced the readers to two different, yet very useful ways of building mathematical models for the control of infectious diseases. The first approach develops models based on the underlying physical or biological principles, while the second is completely novel and it rests on incidence data of the disease phenomenon. Both these approaches have clear advantages and merits and the central idea in these methodologies is to provide answers to several key questions of concern to public health administrators. The modeling activity has been highlighted by describing the several hidden intricate processes involved. Informatics, as is well known, generates new knowledge that is useful in formulating solutions for fundamental questions such as disease transmission, early outbreak of epidemics, and disease surveillance (from public health perspective), by developing new software tools to create that knowledge. Accordingly this chapter deals with a summary of existing knowledge besides presenting some new algorithms that rely upon artificial intelligence and soft computing tools. However, there exist some other modeling approaches that dwell on statistical methods. It is hoped that this chapter would stimulate new research that would be of interest to mathematicians, biologists, doctors, and other scientists working in related areas.

**Acknowledgments** This research is supported by the Foundation for Scientific Research and Technological Innovation (FSRTI)—A Constituent Division of Sri Vadrevu Seshagiri Rao Memorial Charitable Trust, Hyderabad - 500035, India.

## References

1. Antoniou GE, Mentzelopoulou S (1995) Neural networks: an application to the epidemics. In: *Proceedings Neural, Parallel and Scientific Computations* 1:18–21
2. Bailey N (1975) *The mathematical theory of infectious diseases and its applications*. Griffin, London
3. Boccara N, Cheong K (1992) Automata network sir models for the spread of infectious diseases in populations of moving individuals. *J Phys A: Math Gen* 25:2447–2461
4. Brauer F, Castillo-Chavez C (2001) *Mathematical models in population biology and epidemiology*. Springer, Berlin
5. Buehler J, Hopkins R, Overhage J, Sosin D, Tong V (2004) Framework for evaluating public health surveillance systems for early detection of outbreaks: recommendations from the cdc working group. *Morbidity and mortality weekly report* 53(RR05):1–11
6. Chan EH, Sahai V, Conrad C, Brownstein JS (2011) Using web search query data to monitor dengue epidemics: a new model for neglected tropical disease surveillance. *PLoS Negl Trop Dis* 5(5):e1206
7. Cho SB, Won HH (2007) Cancer classification using ensemble of neural networks with multiple significant gene subsets. *Appl Intelligence* 26(3):243–250
8. Christakis NA, Fowler JH (2010) Social network sensors for early detection of contagious outbreaks. *PLoS ONE* 5(9):e12948
9. Coelho FC, Codeco CT, Gomes MGM (2011) A bayesian framework for parameter estimation in dynamical models. *PLoS ONE* 6(5):e19616
10. Coppel W (1965) *Stability and asymptotic behaviour of differential equations*. D. C. Heath, Boston
11. Das R, Turkoglu I, Sengur A (2009) Diagnosis of valvular heart disease through neural networks ensembles. *Comput Methods Programs Biomed* 93(2):185–191
12. Diekmann O, Heesterbeek J (2000) *Mathematical epidemiology of infectious disease*. Wiley, New York
13. Dong-bin Z, Jian-qiang Y (2006) Analysis of infectious disease data based on evolutionary computation. *Intelligence and Security Informatics. Lect Notes Comput Sci* 3917:179–180
14. Erbe L, Freedman HI (1985) Modeling persistence and mutual interference among subpopulations of ecological communities. *Bull Math Biol* 47:295–304
15. Freedman HI (1979) Stability analysis of a predator-prey system with mutual interference and density-dependant death rates. *Bull Math Biol* 41:67–78
16. Freedman HI, Sree Hari Rao V (1983) The tradeoff between mutual interference and time lags in predator-prey systems. *Bull Math Biol* 45:991–1004
17. Gross L (2006) A new model for predicting outbreaks of west nile virus. *PLoS Biol* 4(4):e101
18. Gutta S, Huang J, Takacs JHB, Wechsler H (1996) Face recognition using ensembles of networks. In: *Proceedings of the 13th International Conference on Pattern Recognition*, Springer, pp 50–54
19. Han M, Zhu X, Yao W (2012) Remote sensing image classification based on neural network ensemble algorithm. *Neurocomputing* 78(1):133–138
20. Hansen L (1990) Neural network ensembles. *IEEE Trans Pattern Anal Mach Intell* 12 (10):993–1001
21. Hartman P (1964) *Ordinary differential equations*. Wiley, New York
22. Holtgrave DR, Crosby RA (2004) Social determinants of tuberculosis case rates in the united states. *Am J Prev Med* 26(2):159–162
23. Hooten MB, Anderson J, Waller LA (2010) Assessing north american influenza dynamics with a statistical sirs model. *Spat Spatiotemporal Epidemiol* 1(2–3):177–185
24. Jiang X, Cooper G (2010) A bayesian spatio-temporal method for disease outbreak detection. *J Am Med Inform Assoc* 17(4):462–471

25. Johnson GD, Eidson M, Schmit K, Ellis A, Kulldorff M (2006) Geographic prediction of human onset of west nile virus using dead crow clusters: an evaluation of year 2002 data in new york state. *Am J Epidemiol* 163(2):171–180
26. Julian K, Eidson M, Kipp A, Weiss E, Petersen L, Miller J (2002) Early season crow mortality as a sentinel for west nile virus disease in humans, northeastern united states. *Vector Borne Zoonotic Dis* 2:145–155
27. Kakchapat S, Ardkae J (2011) Modeling of malaria incidence in nepal. *J Res Health Sci* 11 (1):7–13
28. Kuo HI, Lu CL, Tseng WC, Li HA (2009) A spatiotemporal statistical model of the risk factors of human cases of h5n1 avian influenza in south-east asian countries and china. *Public Health* 123(2):188–193
29. LaDew S, Glass G, Hobbs N, Latimer A, RS O (2011) Data-model fusion to better understand emerging pathogens and improve infectious disease forecasting. *Ecol Appl* 21(5):1443–14460
30. Liu TY, Li GZ, Liu Y, Wu GF, Wang W (2006) Estimation of the future earthquake situation by using neural networks ensemble. In: *Proceedings of the third international conference on advances in neural networks*, vol Part III. Springer, Berlin, Heidelberg, ISNN'06, pp 1231–1236
31. Marc C, Jean-Franois G, Pejman R (2006) Dynamics of infectious diseases and pulse vaccination: teasing apart the embedded resonance effects. *Physica D* 223:26–35
32. Mc Clamroch N (1980) State models of dynamic systems: a case study approach. Springer, Berlin, Heidelberg, New York
33. McCullagh P, Nelder JA (1989) Generalized linear models, 2nd edn. Chapman & Hall, London
34. Nelder JA (1961) The fitting of a generalization of the logistic curve. *Biometrics* 17(1):89–110
35. Nelder JA, Wedderburn RWM (1972) Generalized linear models. *J Roy Stat Soc A Gen* 135:370–384
36. Nemytskii V, Stepanov V (1960) Qualitative theory of differential equations. Princeton University Press, Princeton
37. Phan QM, Mark AS, Chris J, Nigel F, Birgit S (2011) Spatiotemporal analyses of highly pathogenic avian influenza h5n1 outbreaks in the Mekong River Delta, Vietnam, 2009. *Spat Spatiotemporal Epidemiol* 2(1):49–57
38. Shujing G, Lansun C, Juan J, Angela T (2006) Analysis of a delayed epidemic model with pulse vaccination and saturation incidence. *Vaccine* 24:6037–6045
39. Sree Hari Rao V, Naresh Kumar M (2010) Estimation of the parameters of an infectious disease model using neural networks. *Nonlinear Anal: Real World Appl* 11(3):1810–1818
40. Sree Hari Rao V, Naresh Kumar M (2012a) Incidence data modeling using logistic regression for infectious diseases, preprint
41. Sree Hari Rao V, Naresh Kumar M (2012b) A new ensemble neural network architecture for estimating parameters of infectious disease, preprint
42. Sree Hari Rao V, Naresh Kumar M (2012c) Novel algorithms for constructing an ensemble of neural networks, preprint
43. Sree Hari Rao V, Naresh Kumar M (2012d) A novel ensemble of neural network classifier for computer aided medical diagnosis of dengue, preprint
44. Sree Hari Rao V, Raja Sekhara Rao P (2007) Cooperative and supportive neural networks. *Phys Lett A* 371:101–110
45. Sree Hari Rao V, Venkata Ratnam K (2004) Multi parameter dynamic optimization algorithms and application to a problem of bioinformatics relating to the spread of an epidemic. *Electron Model* 24:105–116
46. Tsoularis A (2001) Analysis of logistic growth models. *Res Lett Inform Math Sci* 2:23–46
47. Wagner M, Tsui F, Espino J, Dato V, Sittig D, Caruana R, McGinnis L, Deerfield D, Druzdzel M, Fridsma D (2001) The emerging science of very early detection of disease outbreaks. *J Public Health Manag Pract* 7(6):51–59
48. Waltman P (1974) Deterministic threshold models in the theory of epidemics. Springer, Heidelberg

49. Wenzhe C, Pinqiang D, Yonglu C, Dingning C, Zhengyi J (2012) Bp neural network model for early diagnosis of kawasaki disease. *Adv Mater Res* 468–471:723–726
50. Woolhouse M (2011) How to make predictions about future infectious disease risks. *Philos T Roy Soc B: Biological Sciences* 366(1573):2045–2054
51. Xiuju F, Liew C, Soh H, Lee G, Hung T, Ng LC (2007) Time-series infectious disease data analysis using svm and genetic algorithm. In: IEEE congress on evolutionary computation
52. Zhang P, Bui TD, Suen CY (2007) A novel cascade ensemble classifier system with a high recognition performance on handwritten digits. *Pattern Recogn* 40(12):3415–3429
53. Zuo F, de With PHN (2005) Fast face detection using a cascade of neural network ensembles. In: Proceedings of the 7th international conference on advanced concepts for intelligent vision systems. Springer, Berlin, Heidelberg, ACIVS’05, pp 26–34
54. Zuo F, de With PHN (2008) Cascaded face detection using neural network ensembles. *EURASIP J Adv Signal Proc* 2008:736508

# Chapter 2

## Percolation Methods for SEIR Epidemics on Graphs

Alberto Gandolfi

### 2.1 Introduction

Epidemic mathematical modelling has seen an enormous growth [32, 44, 4, 8], from the first deterministic models [50] to stochastic ones [3, 2, 20, 6], based on its ability to capture more and more features of real epidemics, building on the continuous development of mathematical techniques and computer power.

At about the same time percolation theory [77, 52, 42] has moved from simple [23] to more and more elaborate models [9] and detailed questions [27] providing a mathematically more developed theory than that of epidemics [83, 17].

Both percolation and most epidemic models are concerned with the spatial features of a random subset of some network, with the fundamental difference that while the spread of infectious diseases consists of a dynamical process, the mathematical theory of percolation is concerned with a static random object. Percolation is thus easier in principle and only particularly simple epidemic models are directly translated into percolation [41, 55]; however, it is becoming more and more clear over the years that there is in fact a more elaborate [73] and still mostly uncovered interplay between models in the two areas. This review plans to discuss some of the main features of this relation and to hint at some possible future developments.

There are a number of directions in which such interplay can be effective: as mentioned, in some simple cases or for some purposes, it is possible to disregard time in modelling an epidemic and be left with a percolation network. On other occasions, percolation-based observables can bound or approximate quantities of interest in an epidemic model. Again, it can be that the random set of infected individuals in an epidemic is itself modelled by the random set of a percolation

---

A. Gandolfi (✉)

Dipartimento di Matematicae Informatica U. Dini, Università di Firenze,  
Viale Morgagni 67/A, 50134 Firenze, Italy  
e-mail: [gandolfi@math.unifi.it](mailto:gandolfi@math.unifi.it)

process. Finally, since the networks on which epidemics take place can often be taken as random, the networks themselves could be built or analyzed by a suitably defined percolation process.

We review some examples in all these directions: the first three directions are explored in the first part of this review, while the second part deals with the last direction. The final part introduces a new epidemic model which incorporates most of the interesting features of epidemic modelling discussed in the other parts and which, due to its relation with various percolation models, is still suitable for explicit mathematical analysis. For simplicity and clarity we start the review from a specific stochastic model for an SEIR epidemic as discussed in the recent review [46], to which we refer for motivations and further details. For each example, we include a very simple calculation of the probability that one initial infected individual never transmits the infection in a network with four individuals and exponentially distributed random infectious times; the simple examples can be used to appreciate details and implementation difficulties; even at this simple level, models exhibit substantial differences and complications.

## 2.2 SEIR Epidemic Models and Percolation

### 2.2.1 Deterministic Models

Mathematical modelling of epidemics started with Daniel Bernoulli in the XIX century, and one of the major advances has been the 1927 Kermack–McKendrick model [50]. This describes an epidemic by the functions  $s(t)$ ,  $i(t)$  and  $r(t)$  which denote the fractions of susceptible, infected, and recovered individuals at time  $t$ . Such functions must satisfy the condition  $s(t) + i(t) + r(t) = 1$  and their evolution is described by the differential equations:

$$\begin{aligned} s'(t) &= -\lambda s(t)i(t) \\ i'(t) &= \lambda s(t)i(t) - \gamma i(t) \\ r'(t) &= \gamma i(t) \end{aligned}$$

with starting points  $s(0) = 1 - \epsilon$ ,  $i(0) = \epsilon$  and  $r(0) = 0$ , where  $\lambda$  and  $\gamma$  are two parameters [20]. If  $\lambda s(0) > \gamma$ , then initially  $i(t)$  is increasing (and only later decreasing) so that an outbreak is starting; otherwise, if  $\lambda s(0) < \gamma$ , then  $i(t)$  decreases from the very start and infectives disappear steadily. The epidemic threshold is expressed in terms of the basic reproduction number  $\mathbf{R}_0 = \lambda/\gamma$ : for small  $\epsilon$ , outbreak occurs at about  $\mathbf{R}_0 = 1$ . From the equations,  $\mathbf{R}_0$  can be interpreted as the number of secondary infections from an infected individual. The final outbreak size of the epidemics  $r(\infty)$  satisfies

$$1 - r(\infty) = (1 - \epsilon)e^{-\mathbf{R}_0 r(\infty)};$$

and one can see that  $r(\infty)$  is then strictly positive only for  $\mathbf{R}_0 > 1$  [20].

The Kermack–McKendrick model has many, often unrealistic, assumptions; in particular, the transmission of infectious diseases occurs at random. This led to stochastic models, on which we focus in the rest of the presentation.

### 2.2.2 General Stochastic SEIR Epidemic Model

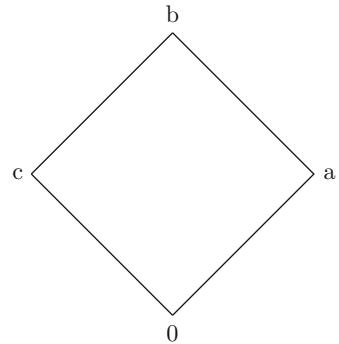
To introduce stochastic models, we start from a set  $V$  of individuals and a list of neighboring relations; it makes sense to take  $V$  finite, but properties which hold for large  $V$  are sometimes better understood in an infinite setting, so  $V$  can be countable as well. We thus have a reference network  $G = (V, \mathcal{E})$ , where  $\mathcal{E} \subseteq \{(v_1, v_2), v_i \in V\}$  is the set of directed neighboring relations indicating that individual  $v_1$  can possibly infect individual  $v_2$ .

In a general stochastic SEIR model each individual  $i \in V$  is in one of the four states: S (susceptible), E (exposed in the latent period), I (infected), and R (recovered or removed, thus immune). An individual  $i \in V$  who is initially in state S is exposed at some random time  $t_i$ , where  $t_i = \infty$  if  $i$  is never infected, moving to state E. After some random time  $e_i$  the individual  $i$  becomes infected, moving to state I, and after another random time  $h_i$  the individual moves to state R to signal perpetual recovery. The latent period  $e_i$  is a.s. finite and is identically 0 in the SIR model. The infectious period  $h_i$  is strictly positive and a.s. finite.

When  $i$  is in state I,  $j$  is in state S, and  $(i, j) \in \mathcal{E}$  to signal network transmissibility, there is a random time  $\tau_{i,j}^*$  such that at time  $t_j = t_i + e_i + \tau_{i,j}^*$  the disease is transmitted to  $j$  who is thus exposed. Clearly,  $\tau_{i,j}^* \in (0, h_i]$  or  $\tau_{i,j}^* = \infty$  to indicate that there has been no transmission.

The starting state of the individuals is generally taken to be S for all but one or a few who start in state I, and the model is generally run by subsequent state changes: it can happen that a few individuals get exposed, indicating a minor outbreak, or that eventually a large proportion of individuals is infected and ends up in state R, signalling a major outbreak. More specifically, we have four mutually disjoint subsets  $S_t, E_t, I_t, R_t \subseteq V$  which evolve for  $t \in [0, \infty)$ . In the initial state  $E_0 = R_0 = \emptyset$  with  $S_0$  and  $I_0$  nonempty and often  $|I_0| = 1$ . Eventually, the evolution will reach a final stable state  $S_\infty, E_\infty = \emptyset, I_\infty = \emptyset, R_\infty$ . The model is thus identified by the reference graph  $G$  and the distributions  $F_i^E(e)$ ,  $F_i^H(h)$  and  $F_{i,j}^*(\tau|h)$  of  $e_i$ ,  $h_i$  and  $\tau_{i,j}^*$ , respectively, where the last one indicates a conditional distribution given the value of  $h$  [46]. The major issue in epidemic analysis is then concerned with the proportion of  $R_\infty$  which for large networks is generally either very small or considerably large; a large outbreak is defined as the case  $R_\infty \gg 0$ , for instance equal to a positive fraction of  $|V|$ . Other features are related to the parameters of the above distributions, such as the parameters range in which a large outbreak can take place, the probability of such an outbreak, the expected final size of the infected population and the size distribution, the transmission probabilities to a

**Fig. 2.1** The elementary network used for some simple examples



given individual, the size distribution of infected group in case of no major outbreak, and the time evolution.

Many epidemic models can be retrieved from the general stochastic SEIR model by appropriate choices of the network  $G$  and of the distributions which identify the model. Although the model is complicated by the simultaneous presence of a large number of random variables, its analysis can be quite often simplified. In the opposite direction, it will be convenient to add additional random variables to capture more features of real epidemics, as we discuss later. The next section starts with some simplifications.

Here is the first of a series of simple examples. In all examples  $V = \{0, a, b, c\}$  and  $I_0 = \{0\}$ , and we compute the probability  $Prob(R_\infty = \{0\})$  that the epidemics does not spread from 0. The only relevant infectious period is that of 0, which we take to have rate  $\mu$ :  $F_0^H(h) = 1 - e^{-\mu h}$ . After the neighboring relations  $\mathcal{E}$  of the network are determined, for each  $i, j \in V$ ,  $(i, j) \in \mathcal{E}$ , the transmission rate once  $i$  is infected is some  $\beta_{i,j}$  so  $F_{i,j}^*(\tau|h) = 1 - e^{-\beta_{i,j}h\tau}$ . One can easily see that the latent periods  $e_i$ 's do not play any role in the calculations.

*Example 1 (General stochastic SEIR epidemics on a network Fig. 2.1).* Let  $\mathcal{E} = \{\{0, a\}, \{0, b\}, \{a, c\}, \{b, c\}\}$ . Then

$$Prob(R_\infty = \{0\}) = \int_0^\infty \mu e^{-\mu h} e^{-h(\beta_{0,a} + \beta_{0,b})} dh = \frac{\mu}{\mu + \beta_{0,a} + \beta_{0,b}}$$

### 2.2.3 Time Evolution

Although epidemics have a time development which is clearly of interest and could be studied from the model [14], many issues can be more easily studied by reducing or eliminating the time dependence of the model. We start our simplification from here.

Already, by the definition of the general stochastic SEIR model one can run the model in two equivalent ways: “on the fly,” which mimics the evolution of the epidemics and amounts to determining the value of the random variables when needed; or “a priori” assigning all the values of the random variables before determining the epidemics and then using values if and when needed [46]. In the “a priori” method one samples for each individual  $i$  the variables  $e_i$  and  $h_i$  from their respective distributions, and then for each ordered pair  $(i, j)$  one samples  $\tau_{i,j}^*$  from the conditional distribution  $F_{i,j}^*(\tau | h)$  for the already sampled  $h = h_i$ .

The a priori sampling suggests that, if we aim at disregarding time, all we really care is whether  $\tau_{i,j}^* \leq h_i$  or  $\tau_{i,j}^* = \infty$ , in other words whether there has been transmission of the disease from  $i$  to  $j$  or not. In fact, if  $\tau_{i,j}^* \leq h_i$  and  $i$  has been infected, then the infection reaches  $j$ : either it reaches it from the infectious contact with  $i$ , or it has already reached it from a previous infectious contact. In any case,  $t_j < \infty$  and the spread can continue from  $j$ . Such information is sufficient to determine all the features related to  $R_\infty$ , although not those related to  $R_t$  and the  $t_i$ ’s.

So we can introduce the random probability  $p(i, j, h_i) = \text{Prob}(\tau_{i,j}^* \neq \infty)$  and discard both the random times  $\tau_{i,j}^*$  and  $e_i$  from the model. After this, the only remaining random variables are the  $h_i$ ’s.

*Example 2 (Timeless general stochastic SEIR epidemics on a network).* Let  $\mathcal{E}$  be as in Example 1. The only difference is the introduction of  $p(i, j, h_i) = 1 - e^{-\beta_{i,j} h_i}$  so

$$\text{Prob}(R_\infty = \{0\}) = \int_0^\infty \mu e^{-\mu h} (1 - p(0, a, h))(1 - p(0, b, h)) dh = \frac{\mu}{\mu + \beta_{0,a} + \beta_{0,b}}$$

If the infectious periods are constant  $h_i \equiv h$  and the probabilities  $p(i, j, h_i) = p(i, j, h) = p(i, j) = p$  do not depend on  $i$  and  $j$ , time can be discretized and we get the classical Reed–Frost model [8]. The simplification is enormous, and unrealistic, but clearly one can get more explicit results [57, 43]. The model is then equivalent to the most elementary percolation process.

### 2.2.4 Bernoulli Percolation

Mathematical percolation models have been devised as extremely simplified models of some random medium still capable of capturing features like sudden global changes as consequence of particular microscopic small changes. For instance, one can think of the passage of fluid in a porous medium: mathematical percolation would model the medium as a random subset of some regular networks with edges open to the fluid independently selected: with the increasing density of open edges the system remains impermeable till it suddenly becomes almost completely permeable.

In more formal mathematical detail, the *standard Bernoulli bond percolation* with parameter  $p \in [0, 1]$  on a graph  $G = (V, \mathcal{E})$  starts from independently and

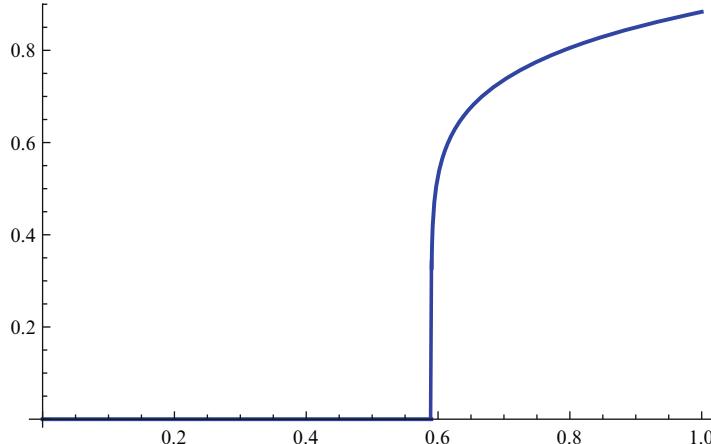
**Fig. 2.2** A percolation cluster in  $\mathbb{Z}^2$



randomly selecting each edge  $e \in \mathcal{E}$  with probability  $p$  and then studying the behavior of the generated random subset formed by the vertices  $V$  and the selected edges. One often focuses on maximal connected components of the random subgraph called clusters. In *site percolation*, instead, the vertices are randomly selected and the random subgraph is formed by the selected vertices and the edges incident to them. Bond percolation is used to model porous media, interactions, spread of fire, and so on, while site percolation can model randomly distributed particles [23], mixed materials, etc. [52, 77]. The qualitative properties of the two models are generally the same if the dimension of the space in which they are embedded is the same, but quantitative features vary.

Although the definition is in general terms, the theory becomes interesting when  $G$  is embedded in some geometric space, for instance is the  $d$ -dimensional integer lattice or some other two-dimensional lattice (Fig. 2.2). In fact, the relative simplicity of the process allows to deal with the constraints of a spatial structure, a relevant feature also for epidemics which is ignored in the simplest models. This is the main reason why percolation theory can be useful in studying epidemics.

When  $G$  is infinite, one key question in percolation theory is whether there are, with positive probability, infinite clusters. One of the first results in percolation theory is that, depending on the graph, there exists a critical  $p_c \in (0, 1)$  such that for  $p < p_c$  there are no infinite clusters and for  $p > p_c$  there is an infinite cluster containing a fixed given vertex (the origin, for instance) with positive probability: in this case we say that percolation occurs. In  $\mathbb{Z}^2$ ,  $p_c = 1/2$  [51] for bond percolation and it is an unknown value estimated around 0.59 for site



**Fig. 2.3** Two-dimensional site percolation probability

percolation [42]. Moreover, the transition is sharp, in the sense that the probability that the origin is in an infinite cluster grows very rapidly to one (with a jump in such a function or in its derivative) (Fig. 2.3).

*Example 3 (Bernoulli percolation on a network).* Let  $\mathcal{E}$  be as in Example 1. In Bernoulli percolation with parameter  $p$

$$\text{Prob}(\text{ the origin } 0 \text{ is isolated }) = (1 - p)^2.$$

Other quantities studied in percolation theory have a corresponding meaning in the epidemic model. The measure of the development of the epidemics by the successive generations, which is to say that for each infected site we record the minimum number of infections needed to reach it, is called chemical distance and the diffusion of the outbreak measured by the chemical distance growths linearly with time instead of exponentially as it does in the general epidemic model. In particular, a shape theorem holds, in the sense that denoting by  $A_n$  the set of vertices infected using at most  $n$  infections, there exists a deterministic set  $A \subseteq \mathbb{Z}^d$  such that  $A_n/n \rightarrow A$  with probability one (see [40]). In addition, if percolation occurs the infinite cluster is unique, and, on the other hand, the sets of infected individuals in the absence of a big outbreak, which is finite clusters in the subcritical regime, have an exponentially small size [42].

In percolation theory, critical exponents describe features like the shape of the percolation probability, or the distribution of the size of the cluster of the origin when  $p = p_c$ . The exponents originate by the fact that it is believed that several quantities are polynomial, with a critical exponent as power. For instance, the probability  $\theta_p$  that the origin is in an infinite cluster is believed to behave like  $\theta_p \approx (p - p_c)^\beta$ . The value of the critical exponents is believed to depend only on essential features of the graph, like dimensionality, but not on local details, defining

some universality classes of models which share the same critical exponents [17]. A similar theory is also expected to hold for epidemic models: [65, 41, 28] argue that epidemics are governed by the “dynamics percolation” universality class.

A mathematically very challenging theory has been developed to describe the asymptotic behavior of large clusters at the critical point  $p_c$  [27, 83], showing that several quantities and random subsets become conformal invariant in the scaling limit (i.e., the network mesh going to zero). Results so far hold only for very specific two-dimensional graphs [76, 26].

Clearly, one can consider random edges up to some bounded length: this is called *short-range Bernoulli percolation* and basically no new phenomena appear.

### 2.2.5 Reed–Frost Model

In the general stochastic SEIR model, if  $p(i, j, h_i) = p(i, j)$ , then each directed edge between pairs of individuals is generated independently since the only dependencies between different transmissions from a vertex  $i$  stem from the common parameter  $h_i$ . In the Reed–Frost model  $p(i, j) = p \in [0, 1]$ . Transmissions are then independent and identically distributed (i.i.d.) random variables, and the “a priori” realization is exactly the Bernoulli percolation model with parameter  $p$ : in fact, the presence of an edge  $\{i, j\}$  in the percolation configuration corresponds to the transmission through that edge when the first of  $i$  and  $j$  is exposed to the infection; if none is ever exposed, then the presence of  $\{i, j\}$  can be disregarded, and if  $\{i, j\}$  is not present and the other individual is later exposed, then transmission through  $\{i, j\}$  is in any case irrelevant to the future course of the epidemics.

We then have a complete transcription: the set of infected individuals is the percolation cluster of the origin, the distribution of small epidemics is the distribution of subcritical clusters, the critical value for epidemics is the percolation critical point, the probability of outbreak is the probability of an infinite percolation cluster, the final size of the epidemic is the density of the infinite cluster, and the probability of being infected is the probability of being in the cluster of the origin.

Although this is an unrealistic simplification, it can serve as a benchmark or for testing plausible features of the model. When we later discuss realistic networks, the direct translation of the Reed–Frost epidemic model into Bernoulli percolation can be used to get a first insight into the features of some classes of networks.

*Example 4 (Comparison of general stochastic SEIR epidemics with percolation).* Let  $\mathcal{E}$  be as in Example 1 and the general SEIR epidemics. It is possible to obtain a Bernoulli percolation, or, equivalently, a Reed–Frost epidemic model, by averaging the infectious periods  $\bar{h} \equiv \int_0^\infty h \mu e^{\mu h} dh = 1/\mu$  and then taking  $p(\bar{h}) = 1 - e^{-\frac{\beta_{0,a} + \beta_{0,b}}{2} \frac{1}{\mu}}$ .

If the transmission rates  $\beta_{i,j} = \beta$  do not depend on the individuals, then one can also take the average transmission probability  $\bar{p} = \bar{p}(0, a) = \bar{p}(0, b) = \int_0^\infty$

$(1 - e^{-\beta})\mu e^{\mu h} dh = \frac{\beta}{\mu + \beta}$  and consider transmissions to be independent, as in Example 3.

Then

$$\begin{aligned} \text{Prob}(R_\infty = \{0\}) &= (1 - p(\bar{h}))^2 = e^{-(\beta_{0,a} + \beta_{0,b})\frac{1}{\mu}}, \\ \text{Prob}'(R_\infty = \{0\}) &= \left(1 - \frac{\beta}{\mu + \beta}\right)^2 = \left(\frac{\mu}{\mu + \beta}\right)^2. \end{aligned}$$

Since  $e^{-(\beta_{0,a} + \beta_{0,b})\frac{1}{\mu}} < \left(\frac{\mu}{\mu + 2\beta}\right)$  and  $\left(\frac{\mu}{\mu + \beta}\right)^2 < \left(\frac{\mu}{\mu + 2\beta}\right)$  for small  $\mu$ , percolation with average infectious periods (see Sect. 2.2.10 below for a development of this bound) or that with average transmission probabilities overestimate the outbreak probability computed in Example 1. None of the two reproduces the exact quantity.

### 2.2.6 Basic Reproduction Number

With the translation of Reed–Frost epidemic into percolation one begins to discover how the spread of infectious diseases takes place between individuals dispersed in a space with some geometrical features.

The first discovery is that the value  $\mathbf{R}_0 = 1$  of the basic reproduction number does not, in general, identify the threshold. For instance, in  $\mathbb{Z}^2$ , the percolation critical point is  $p_c = 1/2$ , but at  $p = 1/2$  we have  $\mathbf{R}_0 = 2$ . The reason is that saturation occurs quite rapidly and infected individuals are surrounded by already recovered individuals, so that a larger value of  $\mathbf{R}_0$  is needed. On the other hand, in many epidemics the value of  $\mathbf{R}_0$  greatly exceeds 1.

Since the basic reproduction number  $\mathbf{R}_0$  is often available in concrete situations, it can be used to estimate the model parameters from which the occurrence of percolation can be inferred, often by means of simulations. In the Reed–Frost model on  $\mathbb{Z}^d$ ,  $p = \mathbf{R}_0/4$  and thus a large epidemic occurs if  $\mathbf{R}_0/4 > p_c = 1/2$ .

### 2.2.7 Fixed Infectious Periods

When the infectious periods  $h_i \equiv h$  are constant and the probabilities  $p(i, j, h_i) = p(i, j)$  then it is still possible to formulate a percolation network equivalent to the general stochastic SEIR model. First generate i.i.d. random variables  $p(i, j)$ 's for each ordered pair  $(i, j)$  of individuals. Suppose that the epidemics starts at some vertex (the origin, for instance) and let  $\{i, j\} \in E$ ; once again, if  $i$  is the first to be exposed to the infectious disease between those at  $i$  and  $j$ , then use the realization of  $p(i, j)$ ; if the first exposed is  $j$ , then use the realization of  $p(j, i)$ , and if none among  $i$  and  $j$  is infected, then take one of the realizations of  $p(i, j)$  or  $p(j, i)$  at random. As

we argued before, possible transmission from the second individual exposed has no relevance to the course of the epidemic. In this construction, the edges have been taken independently, thus realizing a (nonhomogeneous) percolation process. Notice that this occurs in spite of the fact that the  $p(i, j)$ 's depend on the directed edges: this is because we really care only about the first infected in the pair of neighbors and on whether it transmits the disease or not; the second infected and its attempt to transmit are irrelevant in the SIR model. Independence derives from the fact that the  $p(i, j)$ 's are independent for different edges, and the infectious period is constant. Then, in the SIR epidemic with constant infection times also we have a complete transcription in terms of percolation.

*Example 5 (Inhomogeneous Bernoulli percolation on a network).* Let  $E$  be as in Example 1. With  $p(i, j)$  as above we have

$$\text{Prob}(\text{the origin } 0 \text{ is isolated}) = (1 - p(0, a))(1 - p(0, b)).$$

### 2.2.8 Plague Among Great Gerbils in Kazakhstan

A more direct relation to percolation has been discovered recently in the interpretation of the threshold for the plague epidemic among great gerbils in Kazakhstan: in [33] data about a plague epidemic (infection with *Yersinia pestis*) among great gerbils in Kazakhstan has been analyzed. Great gerbils build a burrow system with a small group living in each burrow and data recorded by monitoring in Kazakhstan shows a threshold related to the abundance (fraction of occupied burrows). In [33] a realistic network is taken, with appropriate occupied burrow spatial density, and then an SIR epidemic is simulated by assuming that one burrow is infected and that at each infected site at exponential times an attempted infection occurs; the destination of the infectious attempt has been based on the study of great gerbil movements which are responsible for carrying fleas around: the destination burrow of the infectious attempt is then determined with a random distribution based on the distance from the infected burrow. Outbreak is defined as the spread of infection at some distance. As the observation distance is increased, the probability of outbreak goes from linearly increasing in the burrow occupation rate to a discontinuous function jumping at some critical value. The simulations indicate a threshold at burrow occupation rate  $p \approx 0.31$ , while the collected data indicated a threshold at a 95 % confidence interval of about (0.287, 0.373). At the same time the basic reproduction number at the threshold has been estimated as  $R_0 \approx 1.5$ . Thus it appears that a percolation phase transition is taking place even among a non-static population.

Notice that the explanatory model consists of some modification of the standard percolation models: burrows are not placed at the vertices of a regular two-dimensional network; in addition, both site variables (in terms of occupied burrows) and bond variables (in terms of transmissions) are present; and, finally,

connections are not only between closeby vertices, but can also reach over a long distance. This last remark brings us to long-range percolation, which is discussed in a later section.

### 2.2.9 *Epidemic Percolation*

Back to the general stochastic SEIR model: it can be given a graphical representation using directed edges. Once the random variables  $h_i$  and  $p(i, j, h_i)$  are determined, for each  $i$  and  $j$  in  $V$  if  $(i, j) \in \mathcal{E}$  we select the directed edge  $(i, j)$  with probability  $p(i, j, h_i)$ . Then, starting from the set  $I_0$  of initially infected individuals we determine  $R_\infty$  by including  $I_0$  and all the individuals who can be reached from  $I_0$  following the directed edges selected above.

For a given  $I_0$ , the random set  $R_\infty$  incorporates all of the relevant information about the epidemics except time evolution. Such a random set is in fact a percolation process, named epidemics percolation [46]. Differently from the standard percolation processes, the edges are directed, but more importantly, dependent.

The directionality of the model implies that it has to be analyzed by considering the outcomponent and incomponent of each individual  $i$ , which is to say, the sets of individuals which are reachable from  $i$  in the forward direction, and the set of individuals from which  $i$  can be reached [46].

The dependence comes from the common parameter  $h_i$  used to determine the  $p(i, j, h_i)$ 's with common  $i$  [55], as in Example 2. The selection of the directed edges  $(i, j)$  has, however, only a local dependence: in fact, if  $i \neq i'$ , then  $(i, j)$  and  $(i', j)$  are independently selected. Such condition is called one-dependence, since there is dependence only up to distance one [39]. One-dependence implies that if  $\beta_{i, j} = \beta$  for all pairs of individuals, then there is a critical  $\beta_c$  such that for  $\beta < \beta_c$  the probability  $P_\beta$  of a big (infinite) outbreak is zero, while  $P_\beta > 0$  for  $\beta > \beta_c$  [81, 9]. Bernoulli percolation with parameter equal to the average transmission probabilities overestimate the large outbreak probability [54], as we have seen in Example 4.

From a general point of view, the fact that we now relate epidemics to directed percolation suggests that this is in a different universality class than that related to ordinary percolation.

### 2.2.10 *Extra Variables: Infectivity and Susceptibility*

Additional features of the spread of infectious diseases can be incorporated in the stochastic SEIR model by additional individual-based random variables. Infectivity  $W_i$  and susceptibility  $\bar{W}_i$  are introduced in [59] as i.i.d. random variables, not

necessarily independent for fixed  $i$ , and letting  $p(i, j) = p(W_i, \bar{W}_j)$  be the transmission probability from  $i$  to  $j$  once the first is infected and the second susceptible.

This model deals with different aspects of the infectious transmission, but in some instances it is equivalent to the general stochastic SEIR model: if  $p(i, j, h_i) = p(h_i)$  depends on  $i$  only through  $h_i$  and does not depend on  $j$ , then the SEIR model is equivalent to the Meester–Trapman model with  $W_i = h_i$ ,  $\bar{W}_i = 1$  and the same function  $p$ .

The main result of [59] is that under general conditions several features of epidemics such as the probability of a large outbreak and the expected final size of the outbreak are bounded by the corresponding percolation quantities (probability of percolation and expected cluster size) from above by a Bernoulli bond percolation model (obtained by letting  $p(x, y) = x y$ ,  $W_i \equiv W$ ,  $\bar{W}_i \equiv \bar{W}$ ,  $W\bar{W} = p$ , with  $W$  and  $\bar{W}$  constants) and from below by a Bernoulli site percolation model (obtained by letting  $p(x, y) = x y$  and  $p = P(W_i = \bar{W}_i = 1) = 1 - P(W_i = \bar{W}_i = 0)$ ); we indicate the Bernoulli bond percolation model by  $P_p^{bond}$  and the Bernoulli site percolation model by  $P_p^{site}$ . In particular, suppose that for some random variables  $W$  and  $\bar{W}$ ,  $W_i \equiv W$  and  $\bar{W}_i \equiv \bar{W}$ , then if  $\Xi$  is a collection of hoppable paths (i.e., a collection closed under the operation of switching paths at crossing points and which can be suitably approximated by the first parts of the paths), then the event  $C^\Xi$  that for at least one path in  $\Xi$  all of its edges are present satisfies the following [59]. Let  $(W, \bar{W})$  be a random vector and let  $p(x, y) = p(x y)$  be such that  $p(z)$  is increasing and concave. Then, for every

$$p \geq p(\max[E(W\bar{W}), E(W)E(\bar{W})]),$$

we have

$$P(C^\Xi) \leq P_p^{bond}(C^\Xi).$$

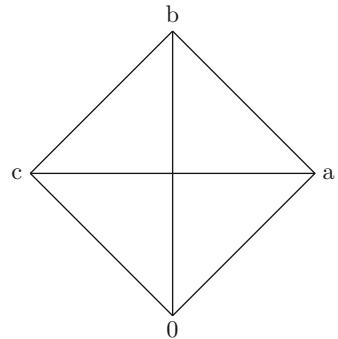
A similar result shows that if  $p \leq E(W\bar{W})$  then  $P(C^\Xi) \geq P_p^{site}(C^\Xi)$ . These results allow us to estimate the probability of a large outbreak, and hence the epidemics critical threshold, and the probability that a given vertex will ever be infected, by corresponding quantities in the percolation model, which are generally relatively easier to compute or estimate.

## 2.3 Random Networks and Percolation

### 2.3.1 Random Graphs

So far we have paid little attention to the network structure. The Kermack–McKendrick model can be recovered from the general SEIR epidemic model when  $G$  is the complete graph. Models on complete graph  $\mathcal{E} = V^2$  are called mass-action or mean-field models.

**Fig. 2.4** The complete network



*Example 6 (Kermack–McKendrick model [50] Fig. 2.4).* Let  $\mathcal{E} = V \times V$  and let the transmission rates be constant. In this simple example we omit the usual rescaling by the inverse of the population size; we then have  $\beta_{i,j} \equiv \beta$ . Then

$$\text{Prob}(R_\infty = \{0\}) = \int_0^\infty \mu e^{-\mu h} (1 - p(h))^3 dh = \frac{\mu}{\mu + 3\beta}.$$

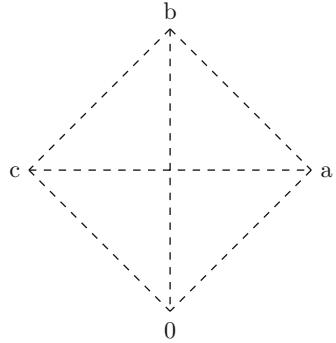
As we mentioned, mass-action models do not capture the inhomogeneity in the number of individuals that can be potentially infected by a given individual: this is better described with a reference network, in particular if this is random.

The classic example of a random network [5] is the Bernoulli random graph, usually called Erdős–Renyi [36, 37]. Starting from  $n$  vertices the random graph  $G(n, p)$  is generated by connecting each pair of vertices independently with probability  $p_n$ . Depending on the value of  $p_n$  there are different asymptotic behaviors of the random graph as  $n$  diverges and one important transition is called the emergence of the giant component. If  $np_n < 1$ , then a graph in  $G(n, p_n)$  will almost surely have no connected components of size larger than  $O(\log n)$ . If  $np_n = 1$ , then a graph in  $G(n, p_n)$  will almost surely have largest component whose size is of order  $n^{2/3}$ . If  $np_n$  tends to a constant  $c > 1$ , then a graph in  $G(n, p_n)$  will almost surely have a unique giant component containing a positive fraction of the vertices and no other component will contain more than  $O(\log n)$  vertices.

Using the random graph  $G$  as a reference graph for the standard stochastic SEIR model, a large outbreak is possible only for  $np_n \geq 1$ . Due to the independence of the edges in the random graph, this model actually produces just a change of transmission probability. This is seen by a joint construction of epidemic evolution and the random graph [63]. With the usual initial sets  $S_0$  and  $I_0$ , the evolution is split into successive discretized generations identified by  $S_k, I_k, R_k$ . Each individual in  $I_k$  makes an independent infectious attempt on each individual in  $S_k$  which is successful with probability  $p_n p(h_i)$ ; if the infection is transmitted by at least one successful attempt, then the target individual enters  $I_{k+1}$ , else it goes into  $S_{k+1}$ . Finally,  $R_{k+1} = I_k \cup R_k$  and  $S_{k+1} = S_k \setminus I_{k+1}$ .

The timeless stochastic SEIR epidemic evolution is correctly reconstructed by this procedure, and one can retrieve the random graph as well. For any edge  $\{i, j\}$ , if

**Fig. 2.5** The basic random network; dashed lines are random



$i, j \notin R_\infty$ , then the edge is taken with probability  $p$ ; otherwise, if the infectious attempt through  $\{i, j\}$  has been successful (in whichever direction it was attempted), then  $\{i, j\}$  is taken; if the unique infection attempt was not successful (and was attempted from  $i$ ), then  $\{i, j\}$  is taken with probability  $\frac{p_n(1-p(h_i))}{1-p_n p(h_i)}$  [63]. Then, any edge through which a transmission has been attempted or realized is taken with probability  $p_n p(h_i) + (1 - p_n p(h_i)) \frac{p_n(1-p(h_i))}{1-p_n p(h_i)} = p$  and all edges are selected independently.

Asymptotics for large  $n$  are discussed in [64] and [63], where it is shown that in this case if  $p = p_n = \alpha/n$  for some  $\alpha > 0$  and the transmission rates are rescaled by the average number  $np_n - 1$  of neighbors in the random graph so that  $\beta_{i,j} \equiv \beta / (np_n - 1)$ , then the basic reproduction number is  $\mathbf{R}_0 = \alpha \int_0^\infty \mu e^{-\mu h} e^{-\frac{\beta}{np_n-1} h} dh$ . Note that  $\mathbf{R}_0$  is the way to identify the epidemic threshold since these are mass-action models.

*Example 7 (SEIR on a random graph Fig. 2.5).* The network edges in  $\mathcal{E}$  are chosen randomly with uniform probability  $p_4$ ; transmission probabilities should be rescaled by the factor  $4p_4 - 1$  but we avoid it in this example to simplify comparison with the other examples. Then  $p(i, j, h) = 1 - e^{-\beta h}$ . Hence

$$\begin{aligned} \text{Prob}(R_\infty = \{0\}) &= \int_0^\infty \mu e^{-\mu h} \sum_{k=0}^3 \binom{3}{k} p_4^k (1-p_4)^{3-k} (1-p(h))^k dh \\ &= \int_0^\infty \mu e^{-\mu h} (1-p_4 p(h))^3 dh \\ &= \sum_{k=0}^3 \binom{3}{k} p_4^k (1-p_4)^{3-k} \int_0^\infty \mu e^{-\mu h} e^{-\beta k h} dh \\ &= \sum_{k=0}^3 \binom{3}{k} p_4^k (1-p_4)^{3-k} \frac{\mu}{\mu + \beta k}. \end{aligned}$$

In this case, the basic reproduction number satisfies  $\mathbf{R}_0 = \frac{\alpha \mu}{\beta + \mu}$ .

### 2.3.2 Small-World Networks

In spite of the interesting developments of the last section, random graphs are not satisfactory from some points of view. Starting from a paper by Watts and Strogatz [82] new indicators have been used to classify graphs and new classes of networks have been discovered. One indicator is the average point-to-point distance  $L$  on the graph, obtained by averaging, over all pairs of vertices on the graph, their graph distance on the graph. Another is the clustering coefficient  $C$ , obtained by considering the set  $A_x$  of neighbors of a vertex  $x$  and the fraction  $C_x$  of edges between pairs in  $A_x$  compared to the total possible:  $C$  is the average of  $C_x$ . Regular lattices have a big clustering coefficient and short average distance, while random graphs of the Erdős–Renyi type have a small clustering coefficient and short average distance. However, it has been noticed that real networks have a small average distance, but a large clustering coefficient. Thus several examples have been built of graphs exhibiting these features: Watts and Strogatz obtained the first example by starting from an ordered set of  $n$  vertices, connecting all vertices with indices closer than a certain distance  $k$  and then randomly rewiring the edges one after the other with some probability  $\beta \in [0, 1]$ ; rewiring consists of placing the second end point of an edge uniformly on the remaining vertices not creating loops or double wiring. For  $\beta = 0$  the lattice is regular and for  $\beta = 1$  it is of Erdős–Renyi type. For  $N > > k > > \log N$  and intermediate values of  $\beta$  clustering is still close to that of regular lattice and average distance is close to that of random graphs.

*Example 8 (SEIR on a Watt–Strogatz small-world network).* Starting with the network of Example 1, the network is realized by randomly rewiring with probability  $\delta$ , using the order  $0, a, c, b$ , the edge which connects the vertex to the successive one and replacing it uniformly where possible. Calculations are quite long and the result is

$$\begin{aligned}
 \text{Prob}(R_\infty = \{0\}) &= \int_0^\infty \mu e^{-\mu h} ((1 - p(h)) \left( \frac{1}{2}(1 - \delta)^3 \delta + \frac{10}{12}(1 - \delta)\delta^3 + \frac{3}{8}\delta^4 \right) \\
 &\quad + (1 - p(h))^2 \left( (1 - \delta)^4 + 3(1 - \delta)^3 \delta + \frac{61}{12}(1 - \delta)^2 \delta^2 \right. \\
 &\quad \left. + \frac{59}{24}(1 - \delta)\delta^3 + \frac{9}{16}\delta^4 \right) \\
 &\quad + (1 - p(h))^3 \left( \frac{1}{2}(1 - \delta)^3 \delta + \frac{11}{12}(1 - \delta)^2 \delta^2 \right. \\
 &\quad \left. + \frac{17}{24}(1 - \delta)\delta^3 + \frac{1}{16}\delta^4 \right) dh \\
 &= \frac{\mu}{48(\beta + \mu)(2\beta + \mu)(3\beta + \mu)} \\
 &\quad \times (144\beta^2 + 48\beta^2\delta - 188\beta^2\delta^2 + 318\beta^2\delta^3 - 127\beta^2\delta^4 \\
 &\quad + 192\beta\mu - 44\beta\delta^2\mu + 94\beta\delta^3\mu - 35\beta\delta^4\mu + 48\mu^2).
 \end{aligned}$$

The first models of small-world graphs are regular enough that one can find explicit expressions for the critical values of the parameters at which a large outbreak takes place [67].

How average distance, clustering, or other features affect the value of the epidemic threshold is not very clear. In general, exact results and numerical analysis seem to suggest that clustering reduces the critical point, although it does not have a clear effect on the final size of the epidemic [21]. These conclusions are, however, challenged by several other features, for instance, correlation of in and out degrees: [62] investigates a small-size network and notices that for such a network to have a lower epidemic threshold than other network structures, there needs to be a positive correlation between the number of links to and from nodes. When this correlation is negative the epidemic threshold for small-size networks can be higher [62] concludes that clustering does not necessarily have an influence on the epidemic threshold if connectance is kept constant and that analyses of the influence of the clustering on the epidemic threshold in directed networks can also be spurious if they do not consider simultaneously the effect of the correlation coefficient between in- and out-degree. Finally, it is observed in [60] that clustering by itself does not seem to reduce the epidemic threshold, but that this is the effect of assortativity, the tendency for nodes to contact nodes of similar degree, which is generally associated with clustering.

### 2.3.3 *Scale-Free Networks*

At the same time, another feature has been considered: the degree distribution of the vertices of the graph decays exponentially in Erdős–Renyi random graphs and in WS but appears to have a power law decay in some realistic network, such as the Internet. Random graphs with this property are called scale-free [24]. A possible construction is by preferential attachment, a procedure in which edges are randomly selected from one vertex with a probability favoring a second vertex with already higher degree [13, 1, 38, 22].

Epidemics has been widely studied in scale-free networks and the surprising feature that simulations and approximations have suggested at first is that the epidemic threshold for a large outbreak is zero [61, 48, 49, 86, 13]. Consequences would be impressive: in a large such network a virus can spread and create a large epidemic no matter how small the individual-to-individual infection probability is.

The issue has not been fully clarified, however. [85] introduces a new scale-free network, which is built recursively by substitutions and studies it asymptotically for a large number of iterations. In such a limit there is a nonzero critical value  $\lambda_c$  for the appearance of a giant component (i.e., for the epidemic outbreak). Another nonzero critical point for a spatial scale-free network is in [72].

Here is an example of preferential attachment with just four individuals even if such a small size induces an excessive oversimplification.

*Example 9 (SEIR on a preferential attachment network).* Take a random graph  $G' = (V', \mathcal{E})$  with probability  $\delta$  of each edge in  $V' = \{a, b, c\}$ ; then attach 0 to each individual  $i \in V'$  with probability 0 if  $i$  is isolated in  $G'$ ,  $1/2$  if  $i$  has one neighbor in  $G'$ , and 1 if  $i$  has two neighbors in  $G'$ . From such a random network  $G$  an SEIR epidemic with probability  $p(h) = 1 - e^{-\beta h}$  of transmission is run starting from 0. Then

$$\begin{aligned} \text{Prob}(R_\infty = \{0\}) &= \int_0^\infty \mu e^{-\mu h} \left( (1-\delta)^3 + \frac{3}{4}(1-\delta)^2\delta + \frac{3}{8}(1-\delta)\delta^2 \right. \\ &\quad + (1-p(h)) \left( \frac{3}{2}(1-\delta)^2\delta + \frac{9}{8}(1-\delta)\delta^2 \right) \\ &\quad + (1-p(h))^2 \left( \frac{3}{4}(1-\delta)^2\delta + \frac{9}{8}(1-\delta)\delta^2 \right) \\ &\quad \left. + (1-p(h))^3 \left( \frac{3}{8}(1-\delta)\delta^2 + \delta^3 \right) \right) dh \\ &= \frac{1}{4(\beta + \mu)(2\beta + \mu)(3\beta + \mu)} \\ &\quad \times (\beta^3(24 - 54\delta + 45\delta^2) + 2\beta^2\mu(22 - 27\delta + 21\delta^2 - 12\delta^3) \\ &\quad - 3\beta\mu^2(-8 + 4\delta - 5\delta^2 + 5\delta^3) + \mu^3(4 + 3\delta^2 - 2\delta^3)). \end{aligned}$$

### 2.3.4 Spatial Networks and Long-Range Percolation

As they are mostly mass-action models, random graphs, small-world, and scale-free networks do not capture the geographical dispersion of individuals, often in two- or three-dimensional environments, which was a key ingredient of most percolation models. The simplest percolation models, on the other hand, are not scale-free nor have the appropriate clustering coefficient or average distance.

Models which combine dimensionality and nonlocal connections are those named spatial networks: individuals are positioned within a given area (or volume) and two individuals are connected with a probability that depends on their distance. By changing the distribution of individuals or the connection probabilities, it is possible to generate a wide variety of networks, from highly clustered lattices to small-world arrangements to globally connected random networks [35, 71, 45]. They are, however, generally not scale-free, a limitation we address later.

The percolation counterpart of spatial networks is long-range percolation, which is generally defined as a random graph which has  $V = \mathbb{Z}^d$  as vertex set and an edge between two vertices at distance  $r$  with probability  $p_{(r)}$ , all edges being selected independently. In general,  $p_{(r)}$  only depends on the distance, for instance

$$p(r) = 1 - e^{-\lambda r}, \quad \text{with} \quad \lambda(r) \approx qr^{-s} \quad (2.1)$$

for some  $s > 0$  and some  $q$  which could be a constant or a slowly varying function of  $r$ . In long-range percolation the probability of occurrence of an infinite cluster follows a particular pattern depending on  $s$ . In dimension  $d$ , if  $s < d$  there is an infinite percolation cluster with probability one, while if  $s > d$  there is a critical value for  $q$ : at small  $q$  there is no percolation, but this does occur for large  $q$  [19]. Such behavior occurs even in dimension  $d = 1$ , in which case the ordinary Bernoulli model exhibits no percolation, even if the nontrivial critical point for  $q$  appears only for  $s \in (1, 2)$  see [74].

Clearly long-range percolation can be considered on other reference networks: an example on a hierarchical lattice is described, for instance, in [53].

Reed-Frost epidemics on the long-range percolation network are studied in [79] and [80]: this gives us the occasion to discuss the role of  $\mathbf{R}_0$ . We have already seen that in spatial graphs the condition  $\mathbf{R}_0 = 1$  no longer identifies the onset of a large outbreak. An epidemic model based on a long-range percolation network could be a good place to try an alternative definition of  $\mathbf{R}_0$ . As an attempt, the set  $B_k$  of vertices at graphical distance  $k$  from the origin is considered in [80], together with the quantities

$$\mathbf{R}_*^{(1)} = \liminf_{k \rightarrow \infty} E(|B_k|)^{1/k}, \quad \mathbf{R}_*^{(2)} = \limsup_{k \rightarrow \infty} E(|B_k|)^{1/k}$$

(using  $\mathbf{R}_*$  if the two coincide). It is then shown that for  $\lambda(r) \approx r^{-\beta}$  the following occurs: for  $\beta < d$  or  $\beta = d$  and some other conditions then  $\mathbf{R}_* = \infty$ ; for  $\beta = d$  and some other conditions then  $1 < \mathbf{R}_*^{(1)} \leq \mathbf{R}_*^{(2)} < \infty$ ; finally, for  $\beta > d$  then  $\mathbf{R}_* = 1$  (see also [18]). This could be a first indicator of the epidemic transition to a large outbreak, although it is not able to identify a transition from a small to a large outbreak if  $\beta \in (d, 2d)$ . The issue is thus still unresolved and for the time being one more easily identifies the onset of a large outbreak by indicating the value of  $\mathbf{R}_0$  at which this occurs.

*Example 10 (SEIR epidemics on long-range percolation).* The edges  $\{0, a\}$  and  $\{0, b\}$  are in  $\mathcal{E}$  with probability  $p_{(1)} = q/1^2 = q$  and  $\{0, c\} \in \mathcal{E}$  with probability  $p_{(2)} = q/2^2 = q/4$ , all independent of one another. With  $p(i, j, h) = p(h) = 1 - e^{-\beta h}$  we have

$$\begin{aligned} \text{Prob}(R_\infty = \{0\}) &= \int_0^\infty \mu e^{-\mu h} \sum_{k=0}^2 \binom{2}{k} q^k (1-q)^{2-k} \left( \frac{q}{4} (1-p(h)) \right)^{k+1} \\ &\quad + \left( 1 - \frac{q}{4} \right) (1-p(h))^k dh \\ &= \sum_{k=0}^2 \binom{2}{k} q^k (1-q)^{2-k} \left( \frac{q}{4} \frac{\mu}{\mu + \beta(k+1)} + \left( 1 - \frac{q}{4} \right) \frac{\mu}{\mu + \beta k} \right). \end{aligned}$$

As we mentioned, spatial networks and long-range percolation are, however, generally not scale-free: this led Yukich [84] to introduce another class of networks in which a random number is selected for each vertex and then two vertices are connected if their distance is less than both the random numbers. The idea is appealing, but does not seem directly suitable to study epidemics: we shall exploit a similar direction after introducing modularity.

### 2.3.5 Modular Networks with Communities

It has become clear in the last few years that there is an even more relevant feature of real populations, in particular human ones: the modular or community structure, in which individuals have contacts through different communities to which they belong and, what is more relevant, contacts and thus infectious rates depend on the community [69].

The first step in this direction is to consider, next to usual contacts, a family or household structure in which infections occur at a faster rate [15, 16, 10]. In particular, [11] and [12] study a random network with  $m$  households each with  $n$  individuals and global random interactions. The spread of infection has then two rates, one for the global and one for the local connections. The epidemic threshold is determined by an extension of the basic reproduction number: suppose there is an initial infected individual 0 and let  $C_0$  be the number of individuals, not in its household, directly infected by 0,  $T$  the number of those eventually infected within the household, and  $C_1$  those directly infected by any individual different from 0 in its household. Then the onset of a large outbreak depends on whether  $\mathbf{R}_* > 1$  or  $\mathbf{R}_* < 1$ , with  $\mathbf{R}_* = E(C_0) + E(T)E(C_1)$ . This is to be contrasted with the usual  $\mathbf{R}_0$  which is between  $E(C_0)$  and  $E(C_0) + E(T)$ .

*Example 11 (SEIR epidemics with households).* Assume that  $\{0, a, b\}$  are in the same household and their connecting edges are automatically in  $\mathcal{E}$ . The individual in  $c$  is outside the household and there is some probability  $p_{(c)}$  that  $\{0, c\} \in \mathcal{E}$ . The infection rate  $\beta_H$  within the household is higher than the infection rate  $\beta_{(c)}$  between 0 and  $c$

$$\begin{aligned} \text{Prob}(R_\infty = \{0\}) &= \int_0^\infty \mu e^{-\mu h} ((1 - p_{(c)})(1 - p(0, a, h))(1 - p(0, b, h)) \\ &\quad + p_{(c)}(1 - p(0, a, h))(1 - p(0, b, h))(1 - p(0, c, h))) dh \\ &= \int_0^\infty \mu e^{-\mu h} \left( (1 - p_{(c)}) e^{-2\beta_H} + p_{(c)} e^{-(2\beta_H + \beta_{(c)})} \right) dh \\ &= (1 - p_{(c)}) \frac{\mu}{\mu + 2\beta_H} + p_{(c)} \frac{\mu}{\mu + 2\beta_H + \beta_{(c)}} \end{aligned}$$

In another direction, several papers have investigated the modular structure of established networks to evaluate the effects on epidemics: [66] presents a class of new divisive algorithms for the discovery of community structure in networks and gives references to other existing methods; [65] shows that the modularity can be expressed in terms of the eigenvectors of a characteristic matrix for the network, named modularity matrix, and by this expression generates a spectral algorithm for community detection (see also [31] and [70]).

Modular effects on epidemics are directly studied in [58] which reports indications, by exact formulas and simulations, that the epidemic threshold decreases with community structures. [30] shows that internal links within the communities are less responsible of the spreading of a disease with respect to external ones. In a slightly different model, dynamical behaviors of epidemics on scale-free networks with community structure is considered in [78] where an SIR model is studied, with communities and movement of individuals away from infected communities.

An interesting network with modular as well as hierarchical structure is proposed in [68]. It starts from a modular network consisting of  $m$  modules, each containing  $n$  nodes. The idea is to take a rich structure within each module, and then analyze its modification under the hierarchical construction. Within each module the connectivity (i.e., the probability of a link between any pair of nodes) is taken to be some  $\rho_1$ , while the connectivity between modules is  $\rho_2 = \rho \rho_1$ , with  $\rho \in [0, 1]$ . The hierarchy is introduced by adding another set of  $m$  modules (each having  $n$  nodes) with the same  $\rho_1$  and  $\rho_2$ . The nodes belonging to these two different sets of modules are then connected, but with a probability  $\rho_3 = \rho^2 \rho_1$ . The resulting network has  $2nm$  nodes and  $l = 2$  hierarchical levels. To increase the number of hierarchical levels to  $l = 3$ , a similar network is added with  $2nm$  nodes to the existing network and, as above, links between these two networks occur with a probability  $\rho_4 = \rho^3 \rho_1$ . Thus, to get a network with  $l = h$  hierarchical levels, the above procedure is repeated  $h - 1$  times. The final network contains  $M = 2(h - 1)m$  number of modules. No study of epidemics has been carried out on the network, and, in any case, the random network has no clear spatial features.

## 2.4 A New Class of Epidemic Models and Their Percolation Analysis

### 2.4.1 *Epidemics on Modular Spatial Hierarchical Scale-Free Networks*

Gathering several of the previous considerations we propose here a new class of random networks called SLEM which is likely to capture important features of epidemics in human communities: networks in the class have spatial dimensionality, modularity, are built with a hierarchical structure, and can be made scale-free by a suitable choice of the parameters. The class is likely to be flexible enough to

model several different situations but is also likely to be regular enough that the analysis of most examples will not be too difficult. Here is a general description of the class, followed by one detailed example.

To construct the SHEM random network, let the vertex set of the graph be  $V = \mathbb{Z}^d$  or any regular finite or infinite lattice. The set of edges  $\mathcal{E}$  contains all the nearest neighbor edges of  $V$  plus a random set of edges defined as follows. Consider a probability distribution  $\nu$  on  $H = \{0, 1\}^{\mathbb{N}}$ , and for each  $i \in V$  an independent realization  $\eta(i) \in H$ . Next, for each  $k \in \{0, 1, \dots\}$  partition  $V$  into blocks  $B_r(k)$  of level  $k$ ; although easier to picture, blocks are not necessarily hypercubes. Then connect two vertices  $i, j \in V$  if there exists a level  $k$  and a block  $B_r(k)$  such that  $i, j \in B_r(k)$  and  $\eta_k(i) = \eta_k(j) = 1$ ; multiple edges can be removed or not, according to the desired model. Clearly, the number of levels is finite in any realistic example, with the largest block being the entire population, but can be taken infinite for mathematical convenience.

Vertices  $i \in V$  represent individuals, and blocks  $B_r(k)$  represent communities. For instance, blocks at level 1 may represent families, at level 2 classes or workplaces, at level 3 schools or companies. At higher levels, blocks may represent special communities which are accessed by only few individuals selected out of a large potential set of members: sports teams, political bodies, company boards, etc. By appropriate selection of the forms of the blocks and of the distribution  $\nu$  one can reproduce the community distributions of human populations of any size. Edges in the graph represent the closeness of individuals and the opportunity of infectious contacts in the case of epidemic. Nearest-neighbor edges in  $\mathbb{Z}^d$  represent the pairs of individuals which are extremely close by, whether or not they are in the same family; note that in  $d = 2$  the number of such close relations is 4. Edges within communities, which connect all members of the community, represent the opportunity of disease transmission within the community itself.

Once the random network  $G = (V, \mathcal{E})$  is constructed one can study or simulate a diffusion of an infectious disease on  $G$ , along the edges both between nearest neighbors and within the communities. Once an individual is infected, transmission to neighbors in the random network takes place at some rate, which can be different for different pairs: it actually makes sense to have community-dependent rates of transmission to the susceptible individual belonging to the same community. A natural assumption is that transmission between a pair of individuals is attempted independently in each shared community, increasing the transmission rate if many communities are shared. It makes sense, also, that the rate at which transmission takes place decreases with the size of the community, as direct contacts become more rare. All of these rates may even depend on infectivity and susceptibility.

To simplify things in this presentation, we restrict ourselves now to a specific form of the distribution  $\nu$ : we fix a nonnegative integer distribution  $P$  with probabilities  $p_k$ ,  $k \in \{0, 1, \dots\}$ , and for each vertex  $i \in V$  consider an independent realization  $X_i$  with distribution  $P$ ;  $\eta_k(i) = 1$  if  $X_i \geq k$  and else  $\eta_k(i) = 0$ . This creates a strong and unnatural dependence between the different communities containing a given individual, but it makes the mathematical analysis much easier, while still capturing some features of human communities.

*Example 12 (SEIR epidemics in the hierarchical random network with communities).* Assume that the transmission probabilities  $p(0, i, h) = p(h)$  are independent of  $i$  but are higher within the household, and that  $X_i$  satisfy  $P(X_i = 0) = 2/3 = 1 - P(X_i \geq 1)$ .  $a$  and  $b$  are also the nearest neighbors of 0. Then

$$\begin{aligned} \text{Prob}(R_\infty = \{0\}) &= \frac{2}{3} \int_0^\infty \mu e^{-\mu h} ((1 - p_H(h))^2 dh \\ &\quad + \frac{1}{3} \sum_{r=0}^3 \binom{3}{r} \int_0^\infty \mu e^{-\mu h} \left(\frac{1}{3}\right)^r \left(\frac{2}{3}\right)^{3-r} ((1 - p_H(h))^2 ((1 - p(h))^r dh \\ &= \frac{2}{3} \frac{\mu}{\mu + 2\beta_H} + \frac{1}{3} \sum_{r=0}^3 3r \left(\frac{1}{3}\right)^r \left(\frac{2}{3}\right)^{3-r} \frac{\mu}{\mu + 2\beta_H + r\beta} \end{aligned}$$

The random network proposed in this section is reminiscent of the one introduced in [84], described above at the end of Sect. 2.3.4, which is obtained as a limit case (see below), but captures more features of the spread of infectious diseases. Similar ideas without spatial dimensions appear in [25] and [75]. Note that the random values  $X_i$ 's assigned to each individual to construct the graph play a different role from that assigned to infectivity and susceptibility in [59]. The key distinguishing element is that in the proposed model the variable  $X_i$ 's are used both to generate the random network through the blocks and to determine the intensity of the interactions via the number of shared blocks.

#### 2.4.2 Percolation Analysis of a Specific Case

Here is a specific example within the above class, which is further studied in detail in [29]. Consider dimension  $d = 2$ ,  $p_k = (\alpha - 1)\alpha^{-(k+1)}$ ,  $k = 0, 1, \dots$ , and nested blocks  $\{B_{r,s}(k)\}$ ,  $r, s \in \mathbb{Z}$ , of size  $2^k$  of the form

$$B_{r,s}(k) = \{(i_1, i_2) \in \mathbb{Z}^2 \mid r2^k \leq i_1 \leq (r+1)2^k - 1, s2^k \leq i_2 \leq (s+1)2^k - 1\}.$$

A random network  $G$  is then created as described above, by removing multiple edges and adding all nearest neighbor edges in  $\mathbb{Z}^2$ . The rigidity of the blocks does not allow to reproduce the community distribution for the low level blocks, such as the household distribution <http://www.statistics.gov.uk/census/>. But it correctly reproduces asymptotic features, such as the distribution of community sizes: this is in general polynomial, with a fraction of communities of size  $m$  of the order of  $c/m^5$  [56]. In fact, the average number of members of a community at level  $k$  is  $\frac{2^{2k}}{\alpha^k}$ , and thus a community has size  $m$  at level  $k(m) = \frac{\log m}{\log 4 - \log \alpha}$ . If we consider a

population of  $2^{2\bar{k}}$ , there are  $\frac{2^{2\bar{k}}}{2^{2k(m)}} = \frac{2^{2\bar{k}}}{m^{\frac{2\log 2}{\log 4 - \log \alpha}}} \approx c/m^{4.8}$  communities of size  $m$ . The average vertex degree  $\bar{d}$  of the origin satisfies

$$\bar{d} = 4 + \sum_{k=2}^{\infty} \frac{1}{\alpha^{2k}} (2^{2k} - 2^{2(k-1)}) = 4 + \frac{3}{\alpha^2 - 4}$$

for  $\alpha = 3$  we have  $\bar{d} \approx 4.6$ . For  $\gamma - 1 = \frac{\log_2 \alpha}{2 - \log_2 \alpha}$  the degree distribution satisfies  $P(d \geq k) \approx ck^{-(\gamma-1)}$  for some constant  $c > 0$ , hence  $P(d = k) \approx ck^{-\gamma}$  and the random network is scale-free for  $2 \leq \alpha \leq 2^{\frac{4}{3}}$ . For  $\alpha = 3$  we have  $P(d = k) \approx c/k^{2.81}$  [29].

To perform a simple analysis of an epidemics on the random network  $G$ , we take a Reed–Frost model, i.e. the SEIR epidemic model with constant infectious periods and transmission rates not depending on the individual. This amounts to a Bernoulli percolation process on  $G$ . Transmission probabilities, however, depend on the community shared by two individuals. There is a transmission probability  $p$  between nearest neighbors, so that at the level of the nearest neighbor graph a standard Bernoulli percolation takes place. Then within each community  $B_{r,s}(k)$  we assume that the probability of transmission is  $p\rho^k$ ,  $\rho \in [0, 1]$ , with transmissions independent for the various individuals and for the different communities even between the same individuals. Therefore, once one of  $i$  or  $j$  is infected and the other is susceptible, then the probability of transmission is  $p(i,j, X_i, X_j) = 1 - \prod_{k=0}^{\infty} (1 - p\rho^k \mathbb{I}_{\{\{i,j\} \ni r,s; X_i, X_j \geq k \text{ and } i,j \in B_{r,s}(k)\}})$ .

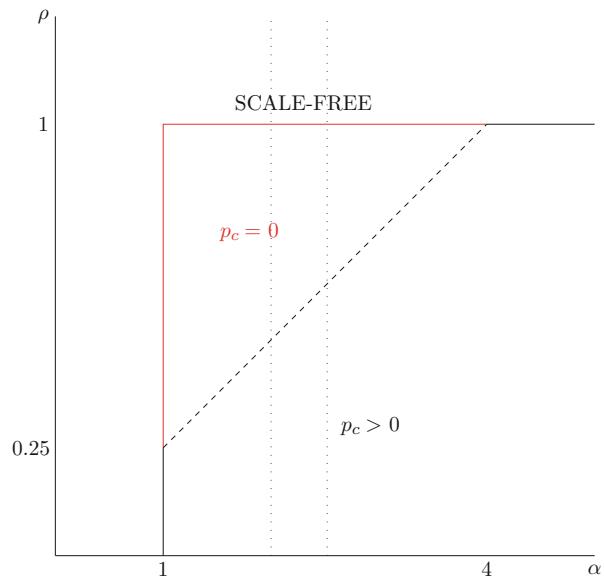
In the limit  $\alpha \rightarrow 1$ ,  $p_k \rightarrow 0$  and  $P(X \geq k) \rightarrow 1$  for all  $k$ : in practice  $X_i \equiv \infty$ , each individual belongs to all communities and if  $m \approx \log_2 r(i, j)$ , where  $r(i, j)$  is the distance between  $i$  and  $j$ , is the linear size of the first block to which  $i$  and  $j$  belong

$$\begin{aligned} p(i,j) &= 1 - \prod_{k=m}^{\infty} (1 - p\rho^k) \\ &\approx 1 - e^{-\sum_{k=m}^{\infty} p\rho^k} \\ &\approx c p \rho^m \approx c p \rho^{\log_2 r(i,j)} \approx c p r(i,j)^{\log p / \log 2} \end{aligned}$$

for some constant  $c > 0$ . Therefore, the model is close to the long-range percolation model on  $\mathbb{Z}^2$  with  $s = -\log \rho / \log 2$  and  $q = cp$ . Percolation, and hence a large outbreak in the Reed–Frost epidemics, occurs at all values of  $p$  if  $s = -\log \rho / \log 2 < 2$ , i.e.  $\rho > 1/4$ , while there is a nontrivial transition in  $p$  between small and large outbreak if  $\rho < 1/2^d = 1/4$ : the model is studied in detail in [53, 34, 7]. It is now interesting to investigate if this behavior occurs also in the scale-free parameter region.

It turns out that for  $\alpha > 4$  the nearest neighbor transmissions take over and thus the model is basically equivalent to the short-range Bernoulli percolation discussed at the end of Sect. 2.2.4 [29]. In the limit  $\alpha \rightarrow \infty$  only nearest neighbor edges

**Fig. 2.6**  $\alpha - \rho$  phase diagram of the hierarchical modular spatial random network in dimension 2



remain, so that we obtain the standard Bernoulli bond percolation in  $\mathbb{Z}^2$ , with the exact critical point  $p_c = 1/2$  (see Sect. 2.2.4). Also for  $\rho = 0$  one retrieves the standard Bernoulli bond percolation in  $\mathbb{Z}^2$ , as no transmission can occur at any higher level. For  $\rho = 1$  the model has some independent interest: two vertices  $i$  and  $j$  are connected if and only if  $X_i, X_j \geq k(i, j)$  where  $k(i, j)$  is the size of the smallest block containing them. The model resembles that of Yukich (see end of Sect. 2.3.4), but with some differences due to the block structure. Thus the model presented here interpolates between long-range and short-range percolation, and in part of the interpolating region it is scale-free.

For  $1 < \alpha < 4$  the behavior is similar to that of long-range percolation: if  $\rho < \alpha/4$  there is a nontrivial transition from small to large outbreak, while for  $\rho > \alpha/4$  the large outbreak occurs for all values of  $p$ : a detailed proof involves several technicalities [29].

As this behavior takes place also in the scale-free region it clarifies the claim that in scale-free networks there could be no small outbreak at all transmission rates: this is indeed the case but only if the decrease in transmission rates in large communities is too slow. These calculations suggest various lines of intervention, such as reducing transmission rates in large communities or vaccinating their more prominent members (who would also belong to even larger communities).

These results are so far only at qualitative level but once more appropriate structures of low level blocks and more realistic transmission mechanisms are introduced, they can be used for quantitative evaluations of epidemics (Fig. 2.6).

## 2.5 Conclusions

We have seen that there is a rich and nontrivial interplay between epidemic and percolation models; such interplay takes advantage of suitable transformations which eliminate time-dependent random variables and may provide exact or approximate relations. On the other hand, it is desirable to introduce new individually based random variables to describe features like the susceptibility of individuals or the structure of the network, and percolation can often encompass such new sources of random behavior.

In particular, we have proposed a new family of random networks which incorporate spatial features, modularity and scale-free characteristics using individual-based random variables. Various types of percolation appear as limiting cases of the new model. The study of the additional features has been carried out, however, only for one specific example, and it would be interesting to study the general stochastic SEIR model on such random networks and analyze the corresponding percolation models. The new model has also been presented in rather general terms, and one needs to develop statistical procedures like model selection, network determination, parameter estimation, simulation and prediction.

The above considerations support the view that the interplay between percolation and epidemics is still at an early stage.

## References

1. Albert R, Jeong H, Barabasi A-L (1999) Diameter of the World Wide Web. *Nature* 401:130
2. Allen LJS (2008) An introduction to stochastic epidemic models. Summer School on Mathematical Modeling of Infectious Diseases, University of Alberta Lecture Notes
3. Allen LJS, Fienberg SE, Holland PW (2008) An introduction to stochastic epidemic models. Springer, Berlin
4. Anderson RM, May RM (1991) Infectious diseases of humans. Oxford University Press, Oxford
5. Andersson H (1999) Epidemic models and social networks. *Math Scientist* 24:128
6. Andersson H, Britton T (2000) Stochastic epidemic models and their statistical analysis. Springer Lecture Notes in Statistics. Springer, New York
7. Athreya SR, Swart JM (2011) Survival of contact processes on the hierarchical group. Preprint. arXiv:0808.3732v3
8. Bailey NTJ (1975) The mathematical theory of infectious diseases and its applications, 2nd edn. Griffin, London
9. Balister PN, Bollobas B (2005) Continuum percolation in the square and the disk. *Random Struct Algor* 26:392–403
10. Ball FG, Mollison D, Scalia-Tomba G (1997) Epidemics with two levels of mixing. *Ann Appl Probab* 7(1):46–89
11. Ball F, Sirl D, Trapman P (2009) Threshold behaviour and final outcome of an epidemic on a random network with household structure. *Adv Appl Probab* 41:765–796
12. Ball F, Sirl D, Trapman P (2010) Analysis of a stochastic SIR epidemic on a random network incorporating household structure. *Math Biosci* 224(2):53–73

13. Barabási A-L, Albert R (1999) Emergence of scaling in random networks. *Science* 286:509–512
14. Barbour AD (1975) The duration of the closed stochastic epidemic. *Biometrika* 62:477–482
15. Bartoszynski R (1972/73) On a certain model of an epidemic. *Zastos Mat* 13:139–151
16. Becker NG, Dietz K (1995) The effect of household distribution on transmission and control of highly infectious diseases. *Math Biosci* 127:207–219
17. Beffara V, Sidoravicius V (2006) Percolation. *Encyclopedia of mathematical physics*, vol 4. Elsevier, Amsterdam, pp 2120–2126
18. Benjamini I, Berger N (2001) The diameter of long-range percolation clusters on finite cycles. *Random structures and algorithms* 19:102–111
19. Berger N (2002) Transience, recurrence and critical behavior for long-range percolation. *Commun Math Phys* 226:531–558
20. Britton T (2005) Stochastic epidemic models: a survey. Cambridge University Press, New York.
21. Britton T, Deijfen M, Lagers AN, Lindholm M (2008) Epidemics on random graphs with tunable clustering. *J Appl Probab* 45(3):743–756
22. Britton T, Janson S, Martin-Löf A (2007) Graphs with specified degree distributions, simple epidemics, and local vaccination strategies. *Adv in Appl Probab* 39(4):922–948
23. Broadbent S, Hammersley J (1957) Percolation processes I. Crystals and mazes. *Proc Cambridge Philos* 53:629–641
24. Caldarelli G (2007) Scale-Free networks complex webs in nature and technology. Oxford University Press, Oxford
25. Caldarelli G, Capocci A, De Los Rios P, Munoz MA (2002) Scale-Free networks from varying vertex intrinsic fitness. *Phys Rev Lett* 89:258702
26. Camia F, Newman CM (2006) Two-dimensional critical percolation: the full scaling limit. *Comm Math Phys* 268(1):1–38
27. Cardy J (2008) Conformal field theory and statistical mechanics. Exact methods in low-dimensional, statistical physics and quantum computing. *Les Houches Summer School Lectures*.
28. Cardy JL, Grassberger P (1985) Epidemic models and percolation. *J Phys A-Math Gen* 18:L267–L271
29. Cecconi L, Gandolfi A (2011) SIR epidemics on a scale-free spatial nested modular network. *arXiv:1107.1532*
30. Chu X, Guan J, Zhang Z, Zhou S (2009) Epidemic spreading in weighted scale-free networks with community structure. *J Stat Mech-Theory E* 2009(07):P07043
31. Clauset A, Newman MEJ, Moore C (2004) Finding community structure in very large networks. *Phys Rev E* 70(6):066111
32. Daley DJ, Gani J (2001) Epidemic modelling: an introduction. Cambridge University Press, Cambridge, UK
33. Davis S, Trapman P, Leirs H, Begon M, Heesterbeek JAP (2008) The abundance threshold for plague as a critical percolation phenomenon. *Nature* 454:634–637
34. Dawson D, Gorostiza L (2011) Percolation in an ultrametric space. Preprint. *arXiv:1006.4400v2*
35. Eames KTD, Keeling MJ (2002) Modeling dynamic and network heterogeneities in the spread of sexually transmitted diseases. *Proc Natl Acad Sci USA* 99:13330–13335
36. Erdős P, Rényi A (1959) On random graphs, I. *Publicationes Mathematicae (Debrecen)* 6:290–297
37. Erdős P, Rényi A (1960) The evolution of random graphs. *Magyar Tud Akad Mat Kut Int Klizleményei* 5:17–61
38. Eriksen KA, Hornquist M (2001) Scale-free growing networks imply linear preferential attachment. *Phys Rev E* 65(1):017102
39. Gandolfi A, Keane M, De Valk V (1989) Extremal two-correlations of two-valued stationary one-dependent processes. *J Probab Theory Rel* 80:475–480

40. Garet O, Marchand R (2004) Asymptotic shape for the chemical distance and first-passage percolation in random environment. *ESAIM: Probab Statist* 8:169–199
41. Grassberger P (1983) On the critical behaviour of the general epidemic process and dynamical percolation. *Math Biosci* 63:157–172
42. Grimmett GR (1999) Percolation. vol. 321 of *Grundlehren der Mathematischen Wissenschaften*, 2nd edn. Springer, Berlin
43. Gutfraind A (2010) Monotonic and non-monotonic epidemiological models on networks. Preprint. arXiv:1005.3470v2
44. Hethcote HW (2000) The mathematics of infectious diseases. *J Soc Ind Appl Math* 42:599–653
45. Keeling MJ (2005) Implications of network structure for epidemic dynamics. *Theor Popul Biol* 67:1–8
46. Kenah E, Miller JC (2011) Epidemic percolation networks, epidemic outcomes, and interventions. *Interdiscip Perspect Infect Dis* 2011:1–13
47. Kenah E, Robins JM (2007) Second look at the spread of epidemics on networks. *Phys Rev E* 76(3):036113
48. Kephart JO, Sorkin GB, Chess DM et al (1997) Fighting computer viruses. *Sci Am* 277:56–61
49. Kephart JO, White SR, Chess DM (1993) Computers and epidemiology. *IEEE Spectr* 30:20–26
50. Kermack W, McKendrick A (1927) A contribution to the mathematical theory of epidemics. *Proc R Soc London A* 115:700–721
51. Kesten H (1980) The critical probability of bond percolation on the square lattice equals 1/2. *Comm Math Phys* 74:41–59
52. Kesten H (1982) Percolation theory for mathematicians. *Progress in Probability and Statistics*, vol. 2, Birkhauser, Boston
53. Koval V, Meester R, Trapman P (2011) Long-range percolation on the hierarchical lattice. Preprint. arXiv:1004.1251v1
54. Kuulasmaa K (1982) The spatial general epidemic and locally dependent random graphs. *Appl Probab* 19:745–758
55. Kuulasmaa K, Zachary S (1984) On spatial general epidemics and bond percolation processes. *J Appl Prob* 21(4):911–914
56. Lancichinetti A, Kivela M, Saramaki J, Fortunato S (2010) Characterizing the community structure of complex networks. *PLoS One* 5:e11976
57. Lefèvre C, Picard P (1990) A non-standard family of polynomials and the final size distribution of Reed-Frost epidemic processes. *Adv Appl Prob* 22:25–48
58. Liu ZH, Hu BB (2005) Epidemic spreading in community networks. *Europhys Lett* 72:315
59. Meester R, Trapman P (2010) Bounding basic characteristics of spatial epidemics with a new percolation model. Preprint.
60. Miller J (2007) Predicting the size and probability of epidemics in populations with heterogeneous infectiousness and susceptibility. *Phys Rev E* 76 010101(R)
61. Moreno Y, Gömez JB, Pacheco AF (2003) Epidemic incidence in correlated complex networks. *Phys Rev E* 68(3):035103
62. Moslonka-Lefebvre M, Pautasso M, Jeger MJ (2009) Disease spread in small-size directed networks: epidemic threshold, correlation between links to and from nodes, and clustering. *J Theor Biol* 260(3):402–411
63. Neal P (2003) SIR epidemics on a bernoulli random graph. *J Appl Probab* 40(3):779–782
64. Neal P, Martin-Löf A (1986) Symmetric sampling procedures, general epidemic processes and their threshold limit theorems. *J Appl Probab* 23(2):265–282
65. Newman MEJ (2006) Modularity and community structure in networks. *Proc Natl Acad Sci USA* 103:8577
66. Newman MEJ, Girvan M (2004) Finding and evaluating community structure in networks. *Phys Rev E* 69(2):026113
67. Newman MEJ, Watts DJ (1999) Scaling and percolation in the small-world network model. *Phys Rev E* 60:7332–7342

68. Pan RK, Sinha S (2008) Modular networks with hierarchical organization: the dynamical implications of complex structure. *Pramana: J Phys* 71(2008):331–340
69. Pellis L, Ferguson NM, Fraser C (2011) Epidemic growth rate and household reproduction number in communities of households, schools and workplaces. *J Math Biol* 63(4):691–734
70. Radicchi F, Castellano C, Cecconi F, Loreto V, Parisi D (2004) Defining and identifying communities in networks. *Proc Natl Acad Sci USA* 101(9):2658–2663
71. Read JM, Keeling MJ (2003) Disease evolution on networks: the role of contact structure. *Proc R Soc B* 270:699–708
72. Sander LM, Warren CP, Sokolov IM (2003) Epidemics, disorder, and percolation. *Physica A* 325(1):1–8
73. Sander LM, Warren CP, Sokolov IM, Simon C, Koopman J (2002) Percolation on heterogeneous networks as a model for epidemics. *Math Biosci* 180:293–305
74. Schulman LS (1983) Long range percolation in one dimension. *J Phys A Lett* 16:L639–L641
75. Servedio VDP, Buttà P, Caldarelli G (2004) Vertex intrinsic fitness: how to produce arbitrary scale-free networks. *Phys Rev E* 70(5):056126
76. Smirnov S (2005) Critical percolation and conformal invariance. In: XIVth International Congress on Mathematical Physics. World Scientific Publishing, Hackensack, pp 99–112
77. Stauffer D, Aharony A (1994) Introduction to percolation theory, 2nd edn. Taylor and Francis, London
78. Suna HJ, Gaoa ZY (2007) *Physica A: Statistical Mechanics and its Applications* 381:491–496
79. Tan Z-J, Zou X-W, Jin Z-Z (2000) Percolation with long-range correlations for epidemic spreading. *Phys Rev E* 62:8409–8412
80. Trapman P (2010) The growth of the infinite long-range percolation cluster. *Ann Prob* 38(4):1583–1608
81. Van den Berg J, Grimmett GR, Schinazi RB (1998) Dependent random graphs and spatial epidemics. *Ann Appl Probab* 8(2):317–336
82. Watts DJ, Strogatz SH (1998) Collective dynamics of “small-world” networks. *Nature* 393(6684):440–442
83. Werner W (2004) Random planar curves and Schramm-Loewner evolutions. In: Lectures on Probability Theory and Statistics. Lecture Notes in Mathematics, vol 1840. Springer, Heidelberg, pp 107–195
84. Yukich, JE (2006) Ultra-small scale-free geometric networks. *J Appl Probab* 43:665–677
85. Zhang Z, Zhou S, Zou T, Chen L, Guan J (2009) Different thresholds of bond percolation in scale-free networks with identical degree sequence. *Phys Rev E* 79(3):031110
86. Zhou T, Fu ZQ, Wang BH (2006) Epidemic dynamics on complex networks. *Prog Nat Sci* 16(5):452–457

# Chapter 3

## Dynamics of Tuberculosis in a Developing Country: Nigeria as a Case Study

Daniel Okuonghae and Andrei Korobeinikov

### 3.1 Introduction

#### 3.1.1 *Background*

Tuberculosis is one of the diseases that is currently considered re-emerging [18]: it was assumed that TB is on its way “out” in many countries, until the number of TB cases began to increase in the late 1980s [29]. The causes behind the observed increase of active TB cases are the source of many studies; it is believed that the growth of human population increased the recurrence of TB outbreaks leading to its current overwhelming high level of endemicity in many countries [5, 8, 16]. It is estimated that a third of the world’s population is currently infected with tuberculosis; of the 1.7 billion of the people estimated to be infected with TB, 1.3 billion live in developing countries [31]. Despite the development of effective therapy, tuberculosis continues to cause high mortality in humans, and especially in developing countries, with sub-Saharan Africa having the highest incidence in per capita rate [20, 23]. The total amount of infected includes 8–12 millions of the active cases and three millions of deaths annually [20, 27].

Tuberculosis (TB) is an infectious disease that is caused by the bacteria *Mycobacterium tuberculosis* (*M. tuberculosis*). Tuberculosis transmission to humans probably occurred from domesticated animals in early agricultural settlements. The infection is spread by droplets from a cough or sneeze of an infected person; tubercle bacillus carried by such droplets lives in the air for a short period

---

D. Okuonghae (✉)

Department of Mathematics, University of Benin, PMB 1154, Benin City, Edo State, Nigeria  
e-mail: [danny.okuonghae@corpus-christi.oxon.org](mailto:danny.okuonghae@corpus-christi.oxon.org); [daniel.okuonghae@uniben.edu](mailto:daniel.okuonghae@uniben.edu)

A. Korobeinikov

Centre de Recerca Matematica, Campus de Bellaterra, Edifici C, 08193 Bellaterra,  
Barcelona, Spain  
e-mail: [akorobeinikov@crm.cat](mailto:akorobeinikov@crm.cat)

(about 2 h) [29]. The lungs is the most common site of infection, but the *Mycobacteria* may spread to form non-pulmonary tuberculosis involving distant sites and other organs. The causative agents of tuberculosis were first isolated by Koch in 1882 as *tubercle bacilli*. The organism belongs to a group of pathogenic and saprophytic bacteria generally known as *Mycobacteria* species [15]. Most infections in humans result in an asymptomatic latent infection. The infection typically stays dormant or inactive for very long, for years in many cases. However, in some cases it can reactivate: about one in ten latent infections eventually progresses to active disease, which, if left untreated, kills more than 50 % of its victims.

A primary infection by *M. tuberculosis* may occur at any age, but a high incidence level of the disease in a population generally leads to earlier occurrence of a primary infection [15]. Post-primary tuberculosis may result from (1) exogenous super infection or reinfection; (2) direct progression of a primary lesion; (3) haematogenous spread; and (4) reactivation of a dormant primary lesion [15]. A number of factors predispose to the occurrence of tuberculosis in a community. These include poor personal and environmental sanitation and over-crowding. Susceptibility to TB in individuals is also increased by such factors as larger infecting dose, immunocompromised individuals, diabetes, age and sex, use of immunosuppressive drugs, malnutrition, alcohol/drug abuse and addiction, postgastrectomy, heredity and, in some cases, race [15]. A general belief is that occasional contacts with a TB-active case (an infectious individual) rarely lead to the infection, and that the most secondary cases are the result of prolonged and sustained close contacts with a primary case. Nevertheless, under the right conditions a single person with active TB could infect many other people [8]. The case of a teacher-librarian with active TB, who infected the children in his/her classroom but not the children who visited the library [26, 29], supports the assertion of difference between casual and close contacts for there to be a transmission of the disease.

Latently infected individuals (inactive TB) become infectious (active TB) after a variable (typically, long) latency period, which duration ranges from months to decades. The majority of the infected individuals never progress to the active TB state. In contrast, an average duration of the infectious period is relatively short (a few months), and it becomes shorter in developed countries due to availability of effective treatment. The individuals having latent TB infection are neither clinically ill nor capable of transmitting TB [21]. However, at older age the immunity could wane, and those who have a latent TB may be at risk of developing active TB as a consequence of either exogenous reinfection (that is, acquiring a new infectious doses due to continuous exposure to another infectious individual) or endogenous reactivation of the pre-existing dormant infection [28, 32].

Tuberculosis is diagnosed definitively by identifying the causative organism (*M. tuberculosis*) in a clinical sample (for example, sputum or pus). The microscopic examination of sputum for acid-fast bacilli is a simple and rapid test that is commonly used to provide a presumptive diagnosis of infectious tuberculosis. When this test is not possible, a probable but although sometimes inconclusive diagnosis may be made using imaging (X-rays or scans) or a tuberculin skin test (Mantoux test). While the patients with tuberculosis with sputum smears negative

for acid-fast bacilli are less infectious than those with positive smears, both theoretical and empirical evidence suggest that they are still able to transmit *M. tuberculosis* [3]. In general, the tuberculosis tests are used to identify infected persons, susceptible groups, and to determine the pattern of immunity in human populations for the purposes of prevention and control of the disease [15].

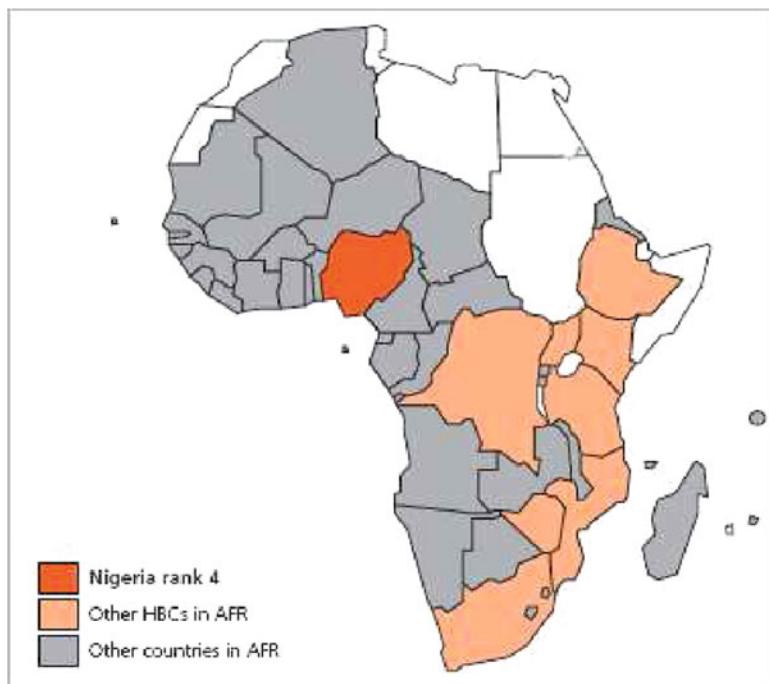
Effective antibiotic treatment with *streptomycin* was introduced in 1946, but preventive therapy was not feasible until the introduction of *isoniazid* in 1952 [2]. Treatment of TB is lengthy and has side effects [8]. Instead of the short course of antibiotics that is typically used to treat other bacterial infections, TB requires considerably longer periods of treatment, around 6–24 months, to entirely eliminate mycobacteria from the body. The drugs must be taken religiously for this period, and the lack of compliance with this drug treatment is a very serious problem that can lead to a relapse of the disease and to the development of antibiotic-resistant TB. The latter is one of the most serious public health problems that society is facing today [8].

A TB vaccine called BCG (Bacillus of Calmette and Guerin) has been available for many decades. BCG vaccine is made of a live, weakened strain of *Mycobacterium bovis* (a cousin of *Mycobacterium tuberculosis*). The vaccine is cheap and remains the only vaccine available against tuberculosis till the date [1]. In Nigeria, the BCG vaccine is usually given to the new born children as a part of the vaccination program of the Federal Government. The vaccine is essential for the children who have a negative tuberculin test and who are continuously exposed but cannot be separated from the adults who are untreated or ineffectively treated for TB [1].

The control and prevention of tuberculosis involves a laboratory identification of the organism from the sputum, culture and sensitivity tests to establish the antibiotic sensitivity pattern, and chemotherapeutic treatment guided by the laboratory observations [15]. Early diagnosis and treatment is essential for the control and prevention of spread. Control measures might include chemoprophylaxis, control of animal tuberculosis, improvement of housing conditions in terms of ventilation and avoidance of overcrowding, health education of family members and the community on the cause of the disease and their possible roles towards the patient, and provision and use of necessary physiotherapy and rehabilitation facilities [15].

### **3.1.2 Direct Observation Therapy Strategy in Nigeria**

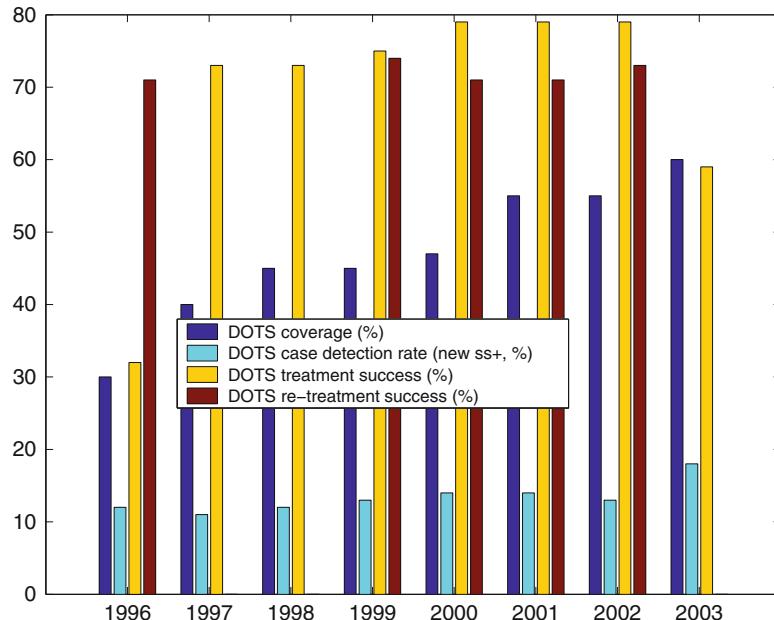
In 1993, alarmed by the rising number of deaths and increasing infection rates, the World Health Organization (WHO) declared TB as a global emergency. A global control strategy adopted by the WHO with the goal of reducing the number of active TB cases is the Direct Observation Therapy Strategy (DOTS). The problem is that the non-adherence to treatment of TB leads to a resurgence of resistance strains of TB, making it difficult to cure. In response, the DOTS evolved as a strategy that makes it compulsory for patients to complete their treatment. The DOTS program involves the use of a nurse or a surrogate who delivers and observes the patients



**Fig. 3.1** Tuberculosis burden in the high-burden countries (HBC) in WHO Africa Region (AFR) (from WHO [35])

taking all the doses of their drugs, rather than relying on the patients taking the drugs unsupervised [10]. The patients may either come to a health facility (the clinic-based DOT) [10, 13], or be visited at the location where the patients are found, e.g. at work, home or shelter (the community-based DOT) [11].

Nigeria has been ranked the fourth among the 22 countries designated by the WHO as high-burden countries (HBC) for TB (see Fig. 3.1). Nigeria is also said to have the highest number of new TB cases in Africa [30, 35] having about 300,000 estimated TB cases recorded each year; these result in 30,000 deaths annually. Moreover, the total of the notified cases for TB of all forms increased from 46,473 in 2003 to 59,493 in 2004. According to WHO [35], the detection of smear-positive cases tripled between 1996 and 2004. In 2002, Nigeria had nearly 368,000 new TB cases. Of these, 159,000 were the pulmonary sputum smear-positive (SS+) cases. In 2004, TB mortality was revealed to be 82/100, 000, and 1.7 % of the new cases were found to be multi-drug resistant [35]. At the end of 2005, 66,848 cases of TB had been notified, of which 55 % were SS+. The public health burden posed by TB is becoming increasingly heavy, as the country's HIV epidemic unfolds [22]. In 2002, the HIV infection rate among adult TB patients was estimated at 27 %. A 2004 study by the Federal Ministry of Health showed that the prevalence of TB for every 100,000 is about 531, with about 30 % of these co-infected with HIV [22]. The WHO estimates that about 27 % of Nigeria's TB patients are HIV-positive.



**Fig. 3.2** DOTS implementation in Nigeria. Source WHO [35]

**Table 3.1** DOTS implementation in Nigeria; from WHO [35]

DOTS level of (%)	1996	1997	1998	1999	2000	2001	2002	2003
Coverage	30	40	45	45	47	55	55	60
Case detection rate	12	11	12	13	14	14	13	18
Treatment success	32	73	73	75	79	79	79	59
Re-treatment success	71	—	—	74	71	71	73	—

The introduction of Direct Observation Therapy Strategy (DOTS) as a means of effecting the proper treatment of tuberculosis did not improve significantly the situation in Nigeria. Figure 3.2 shows the level of DOTS implementation and surveillance in Nigeria, whereas Table 3.1 shows the implementations of DOTS in Nigeria from 1996 to 2003.

DOTS seems to be highly effective in a promoting successful treatment: a comparison of the self-treatment versus various forms of DOTS shows that the completion of treatment is significantly higher when the treatment is supervised [10, 12]. Although a program like the DOTS is essential for reducing the TB relapse rate and preventing the emergence of drug-resistant strains, its actual impact on the control of tuberculosis transmission is not clear [2, 4]. In this chapter we explore some aspects of implementation of the DOTS. To provide an insight into the dynamics of TB under the DOTS, we use simple mathematical models that were formulated having the real-life Nigerian situation in mind.

## 3.2 DOTS with Imperfect Screening Coverage

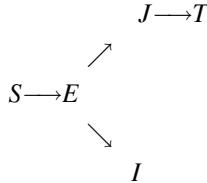
### 3.2.1 Model

The Direct Observation Therapy Strategy (DOTS) was proposed and implemented in a number of developing countries as a practical strategy for bringing down the level of tuberculosis. However, the effectiveness of DOTS depends on a number of other factors, and the most crucial of these is probably the level of detection. The problem is that currently the screening does not cover an entire population, and that the level of detection is far from 100 %. This leads us to the question as to how the level of detection affects the general trends for TB. Another practically relevant question is the minimal level of detection rate that must be achieved in order to eradicate TB, or at least to minimize the level of incidence. To address these issues and explore the overall effect of DOTS on the dynamics of tuberculosis, we consider in this section mathematical models for the spread of TB which take cognizance of the level of detection.

A typical TB spread model is usually referred to as the *SEIT* (Susceptible—Exposed—Infected—Treated) model. According to this model, the susceptible individuals are assumed to be infected via direct contact with the infectious. After an instance of infection, an individual enters the exposed class that corresponds to the latent state of the disease. Then the infection progresses to the active stage, and then after treatment and recovery the individual finally moves to the treated class. The corresponding transfer diagram for this model is

$$S \longrightarrow E \longrightarrow I \longrightarrow T.$$

To incorporate the imperfect detection into the *SEIT* model framework, we assume that there are two distinctive infectious classes, namely the detected infectious individuals and the undetected infectious individuals. We assume that the probability of detection is  $\omega$ , and that when the latency ends and TB progresses into the active stage, the individual is either detected and enters the detected compartment  $J$ , or remains undetected and enters the undetected class  $I$ . The detected individuals are treated and can eventually recover and enter the treated class  $T$ ; the undetected individuals receive no treatment and remain infectious until death. Accordingly, we assumed that  $S(t)$ ,  $E(t)$ ,  $I(t)$ ,  $J(t)$ ,  $T(t)$  are the fractions of the susceptible individuals, the exposed (latent) individuals, the undetected and the detected infectious individuals, and the treated/recovered individuals in the population at time  $t$ , respectively, and that  $S + E + I + J + T = 1$  is the total population. The corresponding transfer diagram is



Let  $\Lambda$  be the recruitment rate into the susceptible class; the recruitment could include immigrants and/or newborns; either of these is assumed to be uninfected and susceptible. Let  $\mu$  be the per capita natural mortality rate; then  $1/\mu$  is an average lifespan of healthy individuals. Let  $d$  be the TB-induced mortality rate with  $d_1$  being the mortality rate for undetected cases and  $d_2$  is that for the detected TB cases. We assume that the population is homogenously mixed and everybody is equally likely to be infected by an infectious individual in the case of contact, and that the transmission of infection occurs with a bilinear incidence rate. We denote  $\alpha_1, \alpha_2 > 0$  the per capita incidence rates for the susceptible class, and  $\gamma_1, \gamma_2 \geq 0$  the incidence rates for the treated class, respectively. Here, for both rates, the subindex 1 is for the contacts with undetected infectious and the subindex 2 is for those with the detected.

It was long observed that TB exhibits strong dose-dependent effects. For example, that is clearly seen in the above-mentioned case of a teacher–librarian. In the framework of a *SEIJT* model, this means dependency (non-linear, in general) of the incidence rate on the number of the infectious individuals. Alternatively, this dose-dependency could be incorporated into a model in the form of the exogenous reinfection. To incorporate this dose-dependency into the model, we assume that duration of the latent state inversely depends on the infective dose. We assume that  $\beta_1, \beta_2 \geq 0$  are the transmission rate for the latent class due to exogenous reinfection by the undetected and detected infectious individuals, respectively, and that  $k$  is the rate of progression from the latent state to active tuberculosis in the absence of exogenous reinfection. For the sake of simplicity, we assume now that the dose-dependency directly affects the length of the latent state. We assume that a length of the latent period is  $(\sigma \cdot f(I, J))^{-1}$ , where  $\sigma^{-1} > 0$  is the average length of the latency when both,  $I$  and  $J$ , are very small (that is,  $I, J \rightarrow 0$ ), and  $f(I, J)$  is a positive non-decreasing function such that  $f(0, 0) = 1$ . For simplicity we assume that  $f(I, J) = 1 + \delta_1 I + \delta_2 J$ . One can expect that  $k \leq \sigma$  holds, as  $\sigma$  includes both the recoveries and deaths of the exposed individuals. If  $\omega$  is the probability of detection, then  $(1 - \omega)$  is the probability that the case remains undetected and enters the  $I$ -class. The exposed individuals that develop active TB are either detected and enter the  $J$ -class at the rate  $\omega k(1 + \delta_1 I + \delta_2 J)$ , or are not detected and enter the  $I$ -class at the rate  $(1 - \omega)k(1 + \delta_1 I + \delta_2 J)$ , or recover and return into the class of the susceptibles at the rate  $r_0$ ; or they die at the rate  $\mu$ .

We assume that all patients under the DOTS (that is the detected infectious individuals) complete their treatment, which is not necessarily successful. Let  $r_2$  be the recovery rate for the detected infectious individuals under the DOTS scheme.

**Table 3.2** Notation and definition of the variables and parameters

Symbols	Description/Explanation
$S$	(Susceptible) individuals not infected and susceptible to infection
$E$	(Exposed) individuals who are infected but not infectious (latent)
$I, J$	(Infectious) individuals who are able to infect others
$T$	(Treated) individuals treated from latent or active TB infection
$\Lambda$	Recruitment rate
$\alpha_i, \beta_i, \gamma_i$	Transmission rates
$\mu$	Per-capita natural mortality rate
$d_1, d_2$	Per-capita excess death rate due to tuberculosis
$k$	Per-capita progression rate
$r_0$	Per-capita treatment rate for latently infected
$r_1, r_2$	Per-capita treatment rate for actively infected
$q$	Fraction of infectious treated under DOTS
$p$	Fraction of infectious not treated under DOTS
$n_1$	Fraction of infectious treated under DOTS that were successful
$n_2$	Fraction of infectious treated under non-DOTS that were successful
$m$	Fraction of failed treatments that were successfully re-treated under DOTS

Let  $q$  be the probability that the treatment is successful, while  $p = 1 - q$  is the probability that the treatment is unsuccessful and that the individual returns to the latent state. The patients, whose cases are not detected, are not treated and either die at the rate  $d_1$ , or self-cure and revert to the latent state at the rate  $r_1$  after an average of 2 years [6, 7, 33]. We also assume that the individuals in the latent class can be recovered/treated at a rate  $r_0$ . These assumptions and the flow diagram lead to the following system of non-linear ordinary differential equations:

$$\dot{S} = \Lambda - \alpha_1 S I - \alpha_2 S J - \mu S, \quad (3.1a)$$

$$\dot{E} = \alpha_1 S I + \alpha_2 S J + \gamma_1 T I + \gamma_2 T J + r_1 I + p r_2 J - \sigma(1 + \beta_1 I + \beta_2 J) E, \quad (3.1b)$$

$$\dot{I} = (1 - \omega)k(1 + \beta_1 I + \beta_2 J)E - cI, \quad (3.1c)$$

$$\dot{J} = \omega k(1 + \beta_1 I + \beta_2 J)E - bJ, \quad (3.1d)$$

$$\dot{T} = r_0 E + q r_2 J - \gamma_1 T I - \gamma_2 T J - \mu T. \quad (3.1e)$$

Here  $\sigma \geq k$ ,  $b = \mu + d_2 + r_2$  and  $c = \mu + d_1 + r_1$ . The variables and parameters in the model are summarized in Table 3.2.

For a *SEIT* model, it is usually assumed that the transmission rate for the treated is lower than that for the susceptibles, that is  $0 \leq \gamma \leq \alpha$ . For instance, for sexually transmitted diseases this may reflect an additional protection, which can be a consequence of education. However, clinical observations give no evidence that the treatment provides any additional protection against tuberculosis. That is, for TB the transmission coefficients for the susceptibles and the treated are equal or have negligible difference. However, if we accordingly assume that  $\gamma_i = \alpha_i$ , then in

the framework of this model the treated and the susceptibles are equivalent, and hence the *SEIJT* model is equivalent to a *SEIJS* model

$$\begin{aligned}\dot{S} &= \Lambda - \alpha_1 SI - \alpha_2 SJ - \mu S + qr_2 J + r_0 E, \\ \dot{E} &= \alpha_1 SI + \alpha_2 SJ + r_1 I + pr_2 J - \sigma(1 + \beta_1 I + \beta_2 J)E, \\ \dot{I} &= (1 - \omega)k(1 + \beta_1 I + \beta_2 J)E - cI, \\ \dot{J} &= \omega k(1 + \beta_1 I + \beta_2 J)E - bJ.\end{aligned}\tag{3.2}$$

### 3.2.2 Equilibrium States and the Basic Reproduction Number

Model (3.2), as well as the model (3.1), always has a disease-free equilibrium state  $E_0$  with the coordinates  $S_0 = \Lambda/\mu, E_0 = I_0 = J_0 (= T_0) = 0$ . Apart from this equilibrium, both models can also have positive equilibria. If a model has a positive equilibrium state and this equilibrium state is stable, then the disease can persist in the population endemically near this equilibrium level. If the equilibrium state is asymptotically stable, then all the subpopulations tend to the corresponding equilibrium levels. Besides, if the disease-free equilibrium state is stable, then the population can remain disease-free indefinitely (or as long as the conditions for the stability of the equilibrium state are valid). Moreover, if the disease-free equilibrium state is asymptotically stable, then there is a critical level of the infectious population such that the disease cannot be sustained in the population when the number of infectious individuals is below this critical level, and eventually disappears. As long as the conditions for this asymptotic stability are valid, the population remains disease-free, even if the disease will be occasionally re-introduced by migrants or travellers. The most favourable situation is, however, when the disease-free equilibrium is globally asymptotically stable; in this case the disease fades out from the population for any initial level of infection.

The ultimate objective of the DOTS is elimination of tuberculosis, or at least reduction of the number of cases to an “acceptable” level. In the terms of mathematical modeling, this implies that the ultimate aim of DOTS is to bring the system in a situation when (1) there is no stable positive equilibria and (2) the disease-free equilibrium is stable. This consideration motivates the study of the asymptotic properties of the equilibrium states: our further goal is finding out whether the aforesaid situation exists and establishing the conditions which lead to its actual realization. These conditions may serve as a guidance for the health authorities and the clinical management.

The asymptotic properties of the *SEIJT* and *SEIJS* models and the existence of the endemic (positive) equilibrium states depend on the basic reproduction number  $R_0$ , which is usually defined as an *expected number of secondary cases produced, in a completely susceptible population, by a typical infective individual* [14]. It is remarkable that the basic reproduction numbers for models (3.1) and (3.2) are the

same. The basic reproduction number can be calculated using the next generation matrix approach [34]. This approach is summarized thus:

1. Rearrange the order of compartments in such a way that the infected classes ( $E$ ,  $I$  and  $J$  for this model) would come before the noninfected classes ( $S$  in this case).
2. Define  $\mathcal{F}_i$  to be the rate of appearance of new infections in compartment  $i$ .
3. Let  $\mathcal{V}_i$  be the difference between the rate of removal from compartment  $i$  and the entry rate into compartment  $i$  by all means other than infection.

For model (3.2), these yield

$$\mathcal{F} = \begin{pmatrix} \alpha_1 SI + \alpha_2 SJ \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

and

$$\mathcal{V} = \begin{pmatrix} \sigma(1 + \beta_1 I + \beta_2 J)E - r_1 I - pr_2 J \\ cI - (1 - \omega)k(1 + \beta_1 I + \beta_2 J)E \\ bJ - \omega k(1 + \beta_1 I + \beta_2 J)E \\ \alpha_1 SI + \alpha_2 SJ + \mu S_\Lambda - qr_2 J - r_0 E \end{pmatrix}.$$

The transmission matrices  $F$  and  $V$  are defined as

$$F = \frac{\partial \mathcal{F}_i}{\partial x_j}(E_0), \quad V = \frac{\partial \mathcal{V}_i}{\partial x_j}(E_0),$$

where  $x_1 = E$ ,  $x_2 = I$ ,  $x_3 = J$ , and  $E_0$  is the disease-free equilibrium state. For system (3.2) these yield

$$F = \begin{pmatrix} 0 & \alpha_1 \frac{\Lambda}{\mu} & \alpha_2 \frac{\Lambda}{\mu} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

and

$$V = \begin{pmatrix} \sigma & -r_1 & -pr_2 \\ -(1 - \omega)k & c & 0 \\ -\omega k & 0 & b \end{pmatrix}.$$

The basic reproduction number  $R_0$  is the maximum eigenvalue (the spectral radius) of matrix  $F V^{-1}$  [34].

Applying this approach, we find that for models (3.1) and (3.2) the basic reproduction number is

$$R_0 = R_{01} + R_{02}, \quad (3.3)$$

where  $R_{01} = h Q_1$  and  $R_{02} = h Q_2$ , with

$$h = k \left/ \left( \sigma - \frac{k(1-\omega)r_1}{c} - \frac{kp\omega r_2}{b} \right) \right.$$

and

$$Q_1 = (1-\omega) \frac{\alpha_1 \Lambda}{\mu c}, \quad Q_2 = \omega \frac{\alpha_2 \Lambda}{\mu b}.$$

We have to note that  $h$  has an important biological meaning: it is the fraction of the infected individuals who develop active TB.

### 3.2.3 Properties of Equilibria

The properties of the disease-free equilibria are defined by the following theorem:

**Theorem 2.1 (Okuonghae and Korobeinikov [25]).** *If  $R_0 \leq 1$  for the model (3.2), then the infection-free equilibrium state  $E_0 = (\Lambda/\mu, 0, 0, 0)$  is the only non-negative equilibrium state of this model, and it is globally asymptotically stable.*

*Proof.* We note that the region

$$\mathcal{D} = \{(S, E, I, J) \in \mathbb{R}_{\geq 0}^4 \mid S + E + I + J \leq \Lambda/\mu\}$$

is a positively invariant set and a global attractor of this system. That is, any trajectory initiated anywhere in the non-negative region  $\mathbb{R}_{\geq 0}^4$  of the phase space eventually enters the region  $\mathcal{D}$  and remains there. To prove this, we first note that the non-negative region  $\mathbb{R}_{\geq 0}^4 = S, E, I, J \geq 0$  is an invariant set of system (3.2), and consider a Lyapunov function  $W(S, E, I, J) = S + E + I + J$ . This function satisfies

$$\begin{aligned} \frac{dW}{dt} &= \frac{d}{dt}(S + E + I + J) \\ &= \Lambda - \mu S + r_0 E - (\sigma - k)(1 + \beta_1 I + \beta_2 J)E - (\mu + d_1)I - (\mu + d_2)J \\ &< \Lambda - \mu S - (\mu + d_1)I - (\mu + d_2)J. \end{aligned}$$

That is,  $\frac{dW}{dt} < -d_1 I - d_2 J < 0$  holds for all  $(S, E, I, J) \in \{\mathbb{R}_{\geq 0}^4 \setminus \mathcal{D}\}$ , and hence, by the virtue of Lyapunov–LaSalle asymptotic stability theorem [19], the set  $\mathcal{D}$  is a global attractor and an invariant set of system (3.2).

This result also implies that the system has no equilibrium states outside of region  $\mathcal{D}$ . Indeed, the derivative of a Lyapunov function is equal to zero at an equilibrium state, whereas  $\frac{dW}{dt} < 0$  strictly for all  $(S, E, I, J) \in \{\mathbb{R}_{\geq 0}^4 \setminus \mathcal{D}\}$ .

It is sufficient now to prove the global stability of  $Q_0$  in  $\mathcal{D}$ . We consider a Lyapunov function

$$U(S, E, I, J) = S - S_0 \ln S + E + BI + CJ,$$

where  $B = (\alpha_1 S_0 + r_1)/c$  and  $C = (\alpha_2 S_0 + pr_2)/b$ . This function satisfies

$$\begin{aligned} \frac{dU}{dt} &= \Lambda - \alpha_1 SI - \alpha_2 SJ - \mu S + qr_2 J + r_0 E \\ &\quad - \Lambda \frac{S_0}{S} + \alpha_1 S_0 I + \alpha_2 S_0 J + \mu S_0 - qr_2 J \frac{S_0}{S} - r_0 E \frac{S_0}{S} \\ &\quad + \alpha_1 SI + \alpha_2 SJ + r_1 I + pr_2 J - (\sigma + \beta_1 I + \beta_2 J)E \\ &\quad + B((1 - \omega)(k + \beta_1 I + \beta_2 J)E - cI) \\ &\quad + C(\omega(k + \beta_1 I + \beta_2 J)E - bJ) \\ &= \Lambda \left(2 - \frac{S_0}{S} - \frac{S}{S_0}\right) + qr_2 J \left(1 - \frac{S_0}{S}\right) + r_0 E \left(1 - \frac{S_0}{S}\right) \\ &\quad + (\alpha_1 S_0 + r_1 - Bc)I + (\alpha_2 S_0 + pr_2 - Cb)J - (\mu + r_0)E \\ &\quad + (B(1 - \omega)k + C\omega k - \sigma)(1 + \beta_1 I + \beta_2 J)E. \end{aligned}$$

Here,

$$2 - \frac{S_0}{S} - \frac{S}{S_0} = \left(1 - \frac{S_0}{S}\right) \left(1 - \frac{S}{S_0}\right) \leq 0$$

for all  $S \geq 0$ . Furthermore,  $\alpha_1 S_0 + r_1 = Bc$  and  $\alpha_2 S_0 + r_2 = Cb$  by definition of  $B$  and  $C$ , and

$$\begin{aligned} (1 - \omega)Bk + \omega Ck - \sigma &= (1 - \omega) \frac{\alpha_1}{c} \frac{\Lambda}{\mu} k + \omega \frac{\alpha_2}{b} \frac{\Lambda}{\mu} k - \left(\sigma - (1 - \omega)k \frac{r_1}{c} - \omega k \frac{pr_2}{b}\right) \\ &= \frac{1}{h} (Q_1 h + Q_2 h - 1) = \frac{1}{h} (R_0 - 1). \end{aligned}$$

That is,  $R_0 \leq 1$  ensures that  $\frac{dU}{dt} \leq 0$  holds for all  $(S, E, I, J) \in \mathcal{D}$ , and the equality  $\frac{dU}{dt} = 0$  holds only at  $E_0$ . Therefore, by Lyapunov asymptotic stability theorem [19],  $R_0 \leq 0$  ensures that the disease-free equilibrium state  $E_0$  is globally asymptotically stable in  $\mathbb{R}_{\geq 0}^4$ .

The uniqueness of the equilibrium state  $E_0$  in  $\mathbb{R}_{\geq 0}^4$  follows from the fact that  $E_0$  is the only point in  $\mathbb{R}_{\geq 0}^4$  where the derivatives of both Lyapunov functions, namely  $W$  and  $U$ , are equal to zero.

This completes the proof.

**Table 3.3** Relationship between the fraction of detected patients under DOTS, the number of secondary infections, and the basic reproduction number  $R_0$

$Q$	$\omega$	$R_0$
10	0.38	1.5
9.7	0.48	1.455
8.9	0.58	1.335
8.8	0.61	1.32
7.0	0.76	1.05
6.7	0.78	1.005

### 3.2.4 Elimination of Infection and Reduction of the Number of Incidence

This theorem implies that if  $R_0 \leq 1$  holds, then for model (3.2), for any initial level of infection, the disease eventually fades out from the population. This result gives us the condition for the critical detection rate  $\omega_c$ , which has to be maintained to eliminate the disease. Indeed, it is easy to see that  $R_0 \leq 1$  holds when

$$\omega \geq \omega_c = \frac{\frac{\alpha_1 \Lambda / \mu + r_1}{c} - \frac{\sigma}{k}}{\frac{\alpha_1 \Lambda / \mu + r_1}{c} - \frac{\alpha_2 \Lambda / \mu + pr_2}{b}},$$

where  $c = \mu + d_1 + r_1$  and  $b = \mu + d_2 + r_2$ . The transmission coefficients  $\alpha_1$  and  $\alpha_2$  can be estimated by matching them with  $Q_1$  and  $Q_2$ , respectively, and with the per capita number of the secondary infections produced by the active TB cases  $Q = Q_1 + Q_2$ . Table 3.3 summarizes results for a range of scenarios corresponding to different values of  $Q$  together with other parameter values. The effect of the variation of the detection rate  $\omega$  on the basic reproduction number is clearly visible from Table 3.3. As  $\omega$  increases,  $R_0$  visibly decreases, showing that maintaining the detection level  $\omega$  is crucial to control tuberculosis. Table 3.3 also shows the values of the basic reproduction numbers for different values of  $Q$ .

These results show that an increase in the detection level  $\omega$  reduces the value of the basic reproduction number  $R_0$ . Moreover, these results also show that the current rate of 21 % case detection in Nigeria is insufficient, and, unless effort is made to improve it, the current low detection rate would effectively diminish the efforts of the DOTS program. The vast number of undetected cases results not only in an increase of the active TB incidence; there also is a huge pool of latent individuals. This is hardly surprising considering the high number of secondary infections caused by the undetected individuals. In fact, results in Table 3.3 indicate that the task of bringing down the incidence of tuberculosis in Nigeria may require maintaining the case detection rate higher than 70 %, which is recommended by the WHO.

Theorem 2.1 describes the global properties of model (3.2) for  $R_0 \leq 1$  and states that maintaining  $R_0 \leq 1$  is sufficient for eradication of the disease: the global asymptotic stability implies that for any positive initial values  $S(0) > 0$  and  $I$

$(0) \geq 0$  the solution eventually approaches the infection free equilibrium state  $E_0$ . Using the same model, it can also be proved that this condition is not only sufficient but is also necessary. Indeed, the condition  $R_0 > 1$  implies that a single infected individual, who is introduced into entirely susceptible population, produces on average more than one infected individual in the next generation. One may expect that in this case the disease successfully invades and persists in a population. Standard linear analysis shows that for both models, namely *SEIJT* model (3.1) and *SEIJS* model (3.2), when the basic reproduction rate  $R_0$  is greater than one, then the disease-free equilibrium is unstable. In mathematical terms this means that a supercritical bifurcation occurs when  $R_0 = 1$ : the disease-free equilibrium state, which is asymptotically stable for all  $R_0 \leq 1$ , loses its stability, and a stable positive equilibrium state appears. It can be rigorously proven that this positive equilibrium state is globally asymptotically stable (and hence all solutions converge to this equilibrium state). However, in this chapter we prefer to demonstrate the necessity of maintaining  $R_0 \leq 1$  for elimination of the disease using simple and transparent arguments.

Minimization of the incidence of tuberculosis requires that the population of all infected classes decrease, that is we require that  $\frac{dE}{dt} < 0$ ,  $\frac{dI_1}{dt} < 0$  and  $\frac{dI_2}{dt} < 0$  hold. It is easy to see that for model (3.1) these inequalities hold when

$$\alpha_1 SI + \alpha_2 SJ + \gamma_1 TI + \gamma_2 TJ + r_1 I + pr_2 J - \sigma(1 + \beta_1 I + \beta_2 J)E < 0, \quad (3.4)$$

$$(1 - \omega)k(1 + \beta_1 I + \beta_2 J)E - cI < 0, \quad (3.5)$$

and

$$\omega k(1 + \beta_1 I + \beta_2 J)E - bJ < 0 \quad (3.6)$$

hold. Substituting (3.5) and (3.6) into (3.4) yields

$$\frac{\alpha_1}{c}S + \frac{\gamma_1}{c}T + \frac{r_1}{c} - \frac{\sigma}{k} < \left( \frac{\alpha_1}{c}S + \frac{\gamma_1}{c}T + \frac{r_1}{c} - \frac{\alpha_2}{b}S - \frac{\gamma_2}{b}T - \frac{pr_2}{b} \right) \omega.$$

Solving this inequality with respect to  $\omega$  gives

$$\omega > \omega_r = \frac{\frac{\alpha_1}{c}S + \frac{\gamma_1}{c}T + \frac{r_1}{c} - \frac{\sigma}{k}}{\frac{\alpha_1}{c}S + \frac{\gamma_1}{c}T + \frac{r_1}{c} - \frac{\alpha_2}{b}S - \frac{\gamma_2}{b}T - \frac{pr_2}{b}}$$

It is easy to see that  $\omega_r < \omega_c$  for all  $S < S_0 = \Lambda/\mu$ , and that  $\omega_r$  tends to  $\omega_c$  as  $S$  tends to the disease-free equilibrium level  $S_0$ . Furthermore,  $\omega_r = \omega_c$  holds for  $S = S_0$ . This actually implies the necessity of condition  $R_0 \leq 1$  for ultimate elimination of the disease. However, this result also implies a practical and realistic strategy for reduction of the incidence. If, for whatever reason, the condition for  $\omega_c$  cannot be met strictly in practice, the health officials at least may attempt to fulfill the condition for reduction maintaining  $\omega \geq \omega_r$ , and consider the complete elimination as a future objective.

### 3.3 The Level of Coverage and the Treatment Success Rate

The treatment for tuberculosis is not always successful and does not necessarily assume full recovery. It frequently occurs that after treatment a case of active tuberculosis turns into latent state, which can progress to the active state thereafter. This is particularly relevant when patients tend to neglect treatment, or just are not persistent enough and neglect to take the medication in due course. The aim of this section is to explore how the level of coverage by the DOTS, and the failure of treatment can affect the efforts to control tuberculosis.

We consider the following *SEIT* model of tuberculosis [24]:

$$\frac{dS}{dt} = \Lambda - \beta_1 SI - \mu S, \quad (3.7a)$$

$$\begin{aligned} \frac{dE}{dt} = & (\beta_1 S + \beta_2 T)I - (k + \mu + r_0)E \\ & + (1 - m)(1 - n_1)qr_1 I + (1 - n_2)pr_1 I, \end{aligned} \quad (3.7b)$$

$$\frac{dI}{dt} = kE - (d + \mu + r_1)I, \quad (3.7c)$$

$$\frac{dT}{dt} = r_0 E + n_1 qr_1 I + n_2 pr_1 I - \mu T - \beta_2 TI + mr_1 q(1 - n_1)I. \quad (3.7d)$$

Here  $0 \leq n_1 \leq 1$ ,  $0 \leq n_2 \leq 1$ ,  $0 \leq m \leq 1$  are the fractions of the cases that were treated successfully under DOTS, treated successfully under a non-DOTS regime, and those who were successfully re-treated under DOTS after their initial treatment failed;  $S + E + I + T = 1$  is the total population. This model is mathematical and epidemiologically well posed. Indeed, it is easy to see that under the flow described by Eq. (3.7), the region  $\mathcal{D} = \{(S, E, I, T) \in \mathbb{R}_{\geq 0}^4; S + E + I + T \leq \Lambda/\mu\}$  is a positively invariant set of the model. That is, each solution initiated in  $\mathbb{R}_+^4$  eventually approaches  $\mathcal{D}$  and remains there. Using the next generation matrix approach [34] (as explained in the previous section), the basic reproduction number for this model is

$$R_0 = fQ, \quad (3.8)$$

where

$$f = \frac{k}{(k + \mu + r_0) - k(q(1 - m)(1 - n_1)r_1 + p(1 - n_2)r_1)/\gamma},$$

and

$$Q = \frac{\beta_1}{\gamma}, \quad \gamma = (d + \mu + r_1).$$

The disease-free equilibrium of model (3.7) has coordinates  $S_0 = \Lambda/\mu$ ,  $E_0 = 0$ ,  $I_0 = 0$  and  $T_0 = 0$ . For  $R_0 > 1$ , there also exists an endemic equilibrium state.

To determine the behaviour of the model near the disease-free equilibrium state, we consider linearization of the system. Evaluating the Jacobian of system (3.7) at the disease-free equilibrium, we have that

$$J_0 = \begin{pmatrix} -\mu & 0 & -\beta_1 & 0 \\ 0 & -(k + \mu + r_0) & \beta_1 + (1 - m)(1 - n_1)qr_1 + pr_1(1 - n_2) & 0 \\ 0 & k & -\gamma & 0 \\ 0 & r_0 & n_1qr_1 + pr_1n_2 + mr_1q(1 - n_1) & -\mu \end{pmatrix} \quad (3.9)$$

The eigenvalues  $\lambda_i$  (where  $i = 1, \dots, 4$ ) of the linearized system satisfy the characteristic equation

$$(\mu + \lambda)^2 \cdot \det(Z - \lambda I) = 0, \quad (3.10)$$

where  $I$  is the identity matrix, and

$$Z = \begin{pmatrix} -(k + \mu + r_0) & \beta_1 + (1 - m)(1 - n_1)qr_1 + p(1 - n_2)r_1 \\ k & -\gamma \end{pmatrix}. \quad (3.11)$$

The disease-free equilibrium state is locally asymptotically stable when all four of the eigenvalues of the linearized matrix are negative. Clearly  $\lambda_1 < 0$  and  $\lambda_2 < 0$  must hold to satisfy  $(\mu + \lambda)^2 = 0$ . Matrix  $Z$  has both its eigenvalues negative only when its trace  $\text{tr}(Z)$  is negative and the determinant  $\det(Z)$  is positive. Here,  $\text{tr}(Z) = -\gamma - k - \mu - r_0 < 0$  holds since all parameters are positive. Furthermore,

$$\begin{aligned} \det(Z) &= \gamma(k + \mu + r_0) - k(\beta_1 + (1 - m)(1 - n_1)qr_1 + p(1 - n_2)r_1) \\ &= \gamma(k + \mu + r_0) - k((1 - m)(1 - n_1)qr_1 + p(1 - n_2)r_1)(1 - R_0), \end{aligned}$$

and hence  $\det(Z) > 0$  holds when  $R_0 < 1$ ,  $\det(Z) = 0$  when  $R_0 = 1$ , and  $\det(Z) < 0$  when  $R_0 > 1$ . Therefore, the disease-free equilibrium state is locally asymptotically stable when  $R_0 < 1$ . At  $R_0 = 1$  a supercritical bifurcation occurs, and the equilibrium state reverses its stability.

The direct Lyapunov method enables us to obtain a stronger result and to establish the sufficient conditions for the global asymptotic stability of the equilibrium state.

**Theorem 3.1.** *The disease-free equilibrium state of system (3.7) is globally asymptotically stable if  $R_0 \leq 1$ .*

The proof of this theorem is similar to proof of Theorem 2.1, and we omit it.

In order to evaluate the impact of the treatment success rate on the tuberculosis dynamics, we apply the same approach that was used in the previous section for the

detection rate. To minimize the incidence of tuberculosis, we require that the population sizes of all infected classes decrease, that is  $\frac{dE}{dt} < 0$  and  $\frac{dI}{dt} < 0$  holds. Combining these inequalities and carrying out the algebraic manipulations in the same way as it was done in Sect. 3.2, we obtain the following condition with respect to  $q$ :

$$q > \frac{\beta_1 S + \beta_2 T + (1 - n_2)r_1 - \gamma(k + \mu + r_0)/k}{(1 - n_2)r_1 - (1 - m)(1 - n_1)r_1} = q_r \quad (3.12)$$

Here  $n_2$  and  $m$  or  $n_1$  cannot be equal to 1 at the same time.

For the eradication of tuberculosis,  $\mathcal{R}_0 < 1$  must hold. In terms of the fraction of the infectious individuals undergoing treatment under DOTS,  $q$ , this condition can be written as

$$q > \frac{\beta_1 + (1 - n_2)r_1 - \gamma(k + \mu + r_0)/k}{(1 - n_2)r_1 - (1 - m)(1 - n_1)r_1} = q_c \quad (3.13)$$

It is easy to see that  $q_r < q_c$  when  $S < S_0$ , and that  $q_r = q_c$  holds when  $S = S_0$ . Therefore, in the framework of this model, the condition  $\mathcal{R}_0 < 1$  is not only sufficient but is also necessary for eradication of tuberculosis. For the disease control, if the condition  $q > q_c$  cannot be achieved in real-life practice, the health authority may instead attempt to keep the condition  $q > q_r$ , which provides a steady decrease of the incidence. The complete elimination can be considered as a future objective.

Note that the models discussed in this chapter did not explicitly include time delays at each stage of transition, even though in reality there are delays in the time spent in the different classes or compartments. Actually, the models, implicitly, have delays in the different classes, depending on the values of some crucial parameters. For instance, in model (3.2), we assumed that recently infected (or reinfected) individuals will spend, at most (approximately),  $\frac{1}{\sigma}$  time in the latent class  $E$ , where  $\sigma$  is made up of the progression rate,  $k$ , the recovery rate,  $r_0$  and death,  $\mu$ . The mean infectious period for individuals in the  $I$  and  $J$  classes was seen to be  $\frac{1}{c}$  time and  $\frac{1}{d}$  time, respectively, where  $c = \mu + d_1 + r_1$  and  $b = \mu + d_2 + r_2$ .

Moreover, Castillo-Chavez and Feng in [9] concluded that the addition of a distributed delay to model the long and variable periods of latency (the rate of removal from the infected class) did not alter the qualitative dynamics of the TB model they proposed when compared to a model with exponentially distributed latency or infectious periods. Also in [17], the introduction of a distribution of the latency period did not change the qualitative dynamics of the TB model proposed; a forward bifurcation characterized the dynamics of the delay model as well as the model with exponentially distributed latency period (for a non explicit delay model).

### 3.4 Conclusion

The Direct Observation Therapy Strategy was proposed by the WHO as a practical strategy of decreasing the level of tuberculosis incidence and preventing the emergence of drug-resistant strains. However, the effective implementation of the DOTS in real-life practice can be affected by a number of factors, and in particular by an imperfection of the detection and treatment. A problem is that in many developing countries it is hardly possible to achieve a complete or near complete screening, and hence the level of detection is far from being ideal. Furthermore, even for the detected cases it is not always possible to provide treatment under the DOTS. The goal of this chapter was to explore how the incompleteness of detection and treatment affects the overall implementation of the strategy.

In order to address these issues of apparent practical importance, we employ simple mathematical models. In this chapter, we examined the effects of DOTS on the dynamics of tuberculosis vis a vis the fraction of infectious individuals who were detected and were undergoing treatment under DOTS. We conducted a qualitative analysis of the models, studying the existence of steady states and establishing the conditions for their stability. The analysis showed that for the models considered, the disease-free equilibrium state exists and is globally asymptotically stable, if the fraction of active cases which are detected and receiving treatment exceeds a certain critical level. These critical values, among other things, depend on the number of new infections, the re-infection rates and the probability of treatment to fail. This analysis leads to practical recommendations for the reduction of the level of incidence of tuberculosis. We have to note that the results of the numerical simulations remarkably confirm the qualitative analysis. One conclusion that can be withdrawn from this analysis is that the overall efficacy of the DOTS crucially depends on the success of the detection. Increasing the fraction of detected and successfully treated leads to a reduction on the value of the basic reproduction number, and hence the number of infectious individuals, and by extension, the number of latent infections.

The basic reproduction number  $R_0$  often serves as a criterion that completely determines a disease behaviour in a population. The condition for this quantity to be less than unity implies the ultimate elimination of the disease. Our analysis demonstrates that this result can be achieved provided that fractions of the detected and the treated infectious individuals under DOTS exceed a certain critical level, which is determined in terms of the epidemiological parameters. If for whatever reason this critical value cannot be reached, then there is a possibility of an epidemic, which could end up in settling at an endemic equilibrium level. The analysis also leads to a conclusion that when maintaining the condition  $R_0 < 1$  for a prolonged time is practically impossible, then a weaker condition, which leads to a gradual reduction of the number of cases, can be applied instead. It should be remembered, however, that these weaker condition does not guarantee the total extermination of the disease, and hence this condition should be applied only as a temporary measure.

In conclusion, we would like to note that the models, which we formulated and analysed in this chapter, can be considered simplistic and even naïve, and lacking many features of real-life infections (such as delays and non-linear functional responses, etc.). The authors have to stress that it was their intention to use the models that are as simple as possible in order to demonstrate these ideas and methods. It is obvious, however, that the same methods and ideas can be immediately applied to more sophisticated models. Nonetheless, while application of these ideas to more advanced models requires considerably more complicated mathematical technique, the principal conclusions that are drawn in this chapter would still be valid.

## References

1. Ale A, Oyedele A, Falola F, Adepegba A (2007) Infants at risk amid scarcity of BCG vaccine. *The Punch Newspaper* 47(1376):2
2. Aparicio JP, Hernandez JC (2006) Preventive treatment of tuberculosis through contact tracing. In: Gumel Abba B (editor-in-chief), Castillo-Chavez Carlos (ed), Mickens Ronald E (ed), Dominic P. Clemence (ed) *Mathematical studies on human disease dynamics: emerging paradigms and challenges*. AMS contemporary mathematics series, vol 410. American Mathematical Society pp 17–29
3. Behr MA, Warren SA, Salamon H, Hopewell PC, Ponce de Leon A, Small PM (1999) Transmission of *Mycobacterium tuberculosis* from patients smear-negative for acid-fast bacilli. *Lancet* 353:444–449
4. Bishai WR, Graham NM, Harrington S, Pope DS, Hooper N, Astemborski J, Sheely L, Vlahov D, Glass GE, Chaisson RE (1998) Molecular and geographic patterns of tuberculosis transmission after 15 years of directly observed therapy. *J Am Med Assoc* 280:1679–1684
5. Blower SM, Small PM, Hopewell PC (1996) Control strategies for tuberculosis epidemics: new models for old problems. *Science* 273(5274):497–500
6. Borgdorff MW (2004) New measurable indicator for tuberculosis case detection. *Emerg Infect Dis* 10(9):1523–1528
7. Borgdorff MW, Floyd K, Broekmans JF (2002) Interventions to reduce tuberculosis mortality and transmission in low and middle-income countries. *Bull World Health Organ* 80:217–227
8. Castillo-Chavez C, Feng Z (1997) To treat or not to treat: the case of tuberculosis. *J Math Biol* 35:629–656
9. Castillo-Chavez C, Feng Z (1998) Mathematical models for the disease dynamics of tuberculosis. In: Arino O, Kimmel M (eds) *Proceedings of the fourth international conference on mathematical population dynamics*, World Scientific, Singapore, 1998
10. Chan ED, Iseman MD (2002) Current medical treatment for tuberculosis. *Brit Med J* 325:1282–1286
11. Chaulk CP, Friedman M, Dunning R (2000) Modeling the epidemiology and economics of directly observed therapy in Baltimore. *Int J Tuberc Lung Dis* 4:201–207
12. Chaulk CP, Kazandjian VA (1998) Directly observed therapy for treatment completion of pulmonary tuberculosis: consensus statement of the public health tuberculosis guidelines panel. *J Am Med Assoc* 279:943–948
13. Cohn DL, Catlin BJ, Peterson KL, Judson FN, Sbarbaro JA (1990) A 62-dose, 6-month therapy for pulmonary and extrapulmonary tuberculosis: a twice-weekly, directly observed, and cost-effective regimen. *Ann Intern Med* 112:407–415
14. Diekmann O, Heesterbeek JA, Metz JAJ (1990) On the definition and the computation of the basic reproductive ratio,  $R_0$  in models of infectious diseases in heterogeneous populations. *J Math Biol* 28:365–382

15. Ezenwa AO (1985) Community health and safety in the tropics. Safety Sciences, Lagos, Nigeria
16. Feng Z, Castillo-Chavez C, Capurro AF (2000) A model for tuberculosis with exogenous reinfection. *Theor Pop Biol* 57:235
17. Feng, Z., Huang W, Castillo-Chavez C (2001) On the role of variable latent periods in mathematical models for tuberculosis. *J Dyn Differ Equ* 13:425–452
18. Hethcote HW (2000) The mathematics of infectious diseases. *SIAM Rev* 42(4):599–653
19. La Salle J, Lefschetz S (1961) Stability by Liapunov's direct method. Academic, New York
20. Magombedze G, Garira W, Mwenje E (2006) Modelling the human immune response mechanisms to *Mycobacterium tuberculosis* infection in the lungs. *Math Biosci Eng* 3(4):661–682
21. Miller B (1993) Preventive therapy for tuberculosis. *Med Clin N Am* 77:1263–1275
22. Muanya C Tuberculosis as a global emergency. *The Guardian Newspaper*, April 24, 2004, p 17
23. North RJ, Yu-Jin J (2004) Immunity to tuberculosis. *Annu Rev Immunol* 22:599–623
24. Okuonghae D (2008) Modelling the dynamics of tuberculosis in Nigeria. *J Nig Assoc Math Phys* 12(1):417–430
25. Okuonghae D, Korobeinikov A (2007) Dynamics of tuberculosis: the effect of Direct Observation Therapy Strategy (DOTS) in Nigeria. *Math Modelling Nat Phenomena: Epidemiol* 2 (1):113–128
26. Raffalii J, Sepkowitz KA, Armstrong D (1996) Community-based outbreaks of tuberculosis. *Arch Int Med* 156:1053
27. Schluger NW, Rom WN (1998) The host immune response to tuberculosis. *Am J Respir Crit Care Med* 157:679–691
28. Smith PG, Moss AR (1994) Epidemiology of tuberculosis. In: Bloom BR (ed) *Tuberculosis: pathogenesis, protection, and control*. ASM Press, Washington, pp 47–59
29. Song B, Castillo-Chavez C, Aparicio JP (2000) Tuberculosis models with fast and slow dynamics: the role of close and casual contacts. *Math Biosci* 180:187–205
30. Soyinka A Nigeria ranks fourth among TB high-burden countries. *The Punch Newspaper*, January 15, 2007, p 3
31. Ssematimba A, Mugisha JYT, Luboobi LS (2005) Mathematical models for the dynamics of tuberculosis in density-dependent populations: the case of Internally Displaced Peoples Camps (IDPCs) in Uganda. *J Math Stat* 1(3):217–224
32. Styblo K (1991) Selected papers: epidemiology of tuberculosis. *Roy Neth Tuberc Assoc* 24:55–62
33. Styblo K, Bumgarner JR (1991) Tuberculosis can be controlled with existing technologies: evidence. *Tuberculosis Surveillance Research Unit*, The Hague, pp 60–72
34. van den Driessche P, Watmough J (2002) Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Math Biosci* 180:29–48
35. WHO (2006) Global tuberculosis control. WHO Report, Geneva

# Chapter 4

## Two-Component Signalling Systems of *M. tuberculosis*: Regulators of Pathogenicity and More

Ruchi Agrawal, Vignesh H. Narayan, and Deepak Kumar Saini

### 4.1 Introduction

Tuberculosis is one of the oldest infectious diseases known to mankind. It is cited in ancient texts by many interesting but dreadful names including “white plague,” “consumption,” and “Yakshma.” The disease is characterized by slow “wasting” of internal systems in the body, primarily respiratory and immune systems hence the name “consumption,” which literally means to consume or use from inside. Clinically, the disease manifestations include persistent, productive and bloody cough, breathlessness and chest pain in case of pulmonary infection. In almost 25 % of cases the infection manifests in extrapulmonary locations, such as bones, central nervous system, peripheral nervous system and genitourinary system. The systemic symptoms, which are manifested in all forms of tuberculosis irrespective of site of infection, include weight loss, chills, night sweat, loss of appetite and fever. The causative agent is disseminated by aerosol droplets released in the air by infected individuals by coughing, sneezing or even speaking. This form of the dissemination is very lethal as a very small bacillary dose is enough to elicit infection and the bacilli can stay viable in air for extended durations of time.

In the case of pulmonary tuberculosis the disease diagnosis is performed by staining of sputum samples for the presence of acid-fast bacilli and a chest X-ray. In most Western countries, a skin test called tuberculin skin test (Mantoux test) is performed as a routine prognostic test. This test is performed to detect the presence of anti-mycobacterial antibodies in the body which mounts an immune response to mycobacterial purified protein derivative (PPD) leading to the development of an “induration” at the site of injection. A positive test is indicative of a resident mycobacterial infection. For treatment, the infected individuals are subjected to chemotherapy consisting of a

---

R. Agrawal • V.H. Narayan • D.K. Saini (✉)

Department of Molecular Reproduction, Development and Genetics, Indian Institute of Science, Sir CV Raman Avenue, Bangalore 560012, India  
e-mail: [deepak@mrdg.iisc.ernet.in](mailto:deepak@mrdg.iisc.ernet.in)

combination of multiple antibiotics such as rifampicin, isoniazid, ethambutol and pyrazinamide and other secondary line of antibiotics if needed. The treatment duration for tuberculosis is generally long, 6–9 months on account of complex architecture and low metabolism of the tuberculosis bacilli (discussed below). Such a long treatment regimen is the major reason of the patient non-compliance for therapy and primary cause of emergence of drug-resistant tuberculosis called as MDR (multi-drug resistant—resistance to at least rifampin and isoniazid), XDR (extensively drug resistant—MDR resistance plus resistance to a fluoroquinolone and an aminoglycoside) and CDR (completely drug resistant) bacillary forms.

## 4.2 Tuberculosis Aetiology: *Mycobacterium tuberculosis*

Tuberculosis is caused by Gram +ve, slow growing, aerobic, non-motile, lipid-rich and acid-fast bacillus, *Mycobacterium tuberculosis*. The name was coined by Robert Koch to describe the “fungus-like” bacterium (Myco-bacterium) grown from lung specimens of tuberculosis patients [1]. The thick outer layer of *M. tuberculosis* is made of complex sugars and lipids, making it virtually impregnable to a large number of disinfectants and evaporation, allowing the bacteria to survive for extended durations even in dry state. The waxy outer layer is mainly composed of mycolic acids, responsible for the formation of caseous granulomas, a typical feature of tuberculosis infection. *M. tuberculosis* is primarily an intracellular pathogen which resides inside the host cells and evades the immune clearance process. Typical intracellular pathogens upon entry into a host cell are degraded by fusion of phagosome (which carries the pathogen) with the lysosomes, which contain hydrolytic enzymes. The formation of phago-lysosome is hence considered essential for pathogen clearing. The tuberculosis bacilli interfere with the formation of phagolysosome, allowing their sustenance in the endosome/phagosome of the infected cells for long durations [2]. This prolonged maintenance of *M. tuberculosis* in the host cells has been termed as latency, as during this stage the bacilli do not replicate and lie dormant for extended periods of time (discussed below). The presence of resident dormant *M. tuberculosis* bacilli in the infected organ is indicated by the presence of caseation. Progression of latent tuberculosis to an active disease is accompanied by necrosis of these caseous granulomas, which leads to spilling of infective bacilli in the lungs. At this stage the extracellular *M. tuberculosis* is liberated free in the airway facilitating spread of the infection [3].

## 4.3 Diagnosis of *M. tuberculosis*

An individual suspected to be infected, in general is empirically diagnosed by the physician based on physical symptoms followed by a chest X-ray. Conventional confirmatory methods to diagnose tuberculosis involve culturing sputum, fluid or biopsy samples obtained from the patient and staining for the presence of acid-fast

bacilli. Though culturing provides clear identification, it usually takes weeks for these bacilli to grow, and a diagnosis based purely on acid-fast staining is marred with issues related to low sensitivity and false-positives. With these limitations in mind, rapid and sensitive nucleic acid amplification-based detection techniques, and more recently, the QuantiFERON-TB Gold assay has been introduced.

In QuantiFERON assay, IFN- $\gamma$  levels in blood samples are measured by ELISA after stimulation with specific peptides based on *M. tuberculosis* proteins ESAT-6 (early secretory antigenic target-6), CFP-10 (culture filtrate protein-10) and TB7.7 [4]. Nucleic acid-based detection techniques aim to discriminate between *M. tuberculosis* and other acid-fast staining bacteria using genetic regions that are unique to the pathogen. A commonly used test is the restriction fragment length polymorphism analysis of an *M. tuberculosis*-specific transposon known as IS6110 [5]. Since different *M. tuberculosis* strains have different copy numbers of IS6110 and in different regions, this test can be used to distinguish between members of the *Mycobacterium tuberculosis* complex (MTBC). However, recent studies have demonstrated that this insertion sequence is not exclusive to MTBC, as a similar sequence has been found in the non-pathogenic *Mycobacterium smegmatis* [6]. PCR of the virulence determining gene *devR* is also used to identify pathogenic *M. tuberculosis*, and has been found to be equivalent in sensitivity to IS6110 PCR [7]. Other PCR assays such as the Cobas TaqMan MTB from Roche Diagnostics, Switzerland uses the amplification of the 1,500 bp long 16s rRNA encoding gene of *M. tuberculosis* to detect the presence of the pathogen [8]. Detection of the target DNA is achieved by cleavage of the fluorescent dye-based oligonucleotide probes in real time. More recently, the Cepheid Xpert MTB/RIF assay (Cepheid AB, Sweden) has been introduced, which uses nested real-time PCR to amplify a 192 bp segment of the *M. tuberculosis* complex *rpoB* gene using five overlapping fluorescent probes that span the 81 bp rifampin-resistance-determining region [9]. The Xpert assay has been shown to outperform the TaqMan MTB assay with sensitivity and specificity of 95 % and 100 %, respectively, as compared to 78 % and 98 % for TaqMan [8]. However, the sensitivity for extrapulmonary specimens and other samples which have paucity in pathogen load is variable. The challenge still remains to obtain efficient lysis of the mycobacterial cell and eliminate false-positives and false-negatives.

#### **4.4 Virulence and Pathogenicity of *M. tuberculosis*: The Issue with Latency**

Besides the clinical manifestations of tuberculosis mentioned above, infection with the bacilli is also characterized by presence of dormant bacilli in the body, which are virulent enough to generate full blown disease on reactivation. These dormant bacilli can stay alive up to 40–50 years after infection, residing in granulomas found predominantly in lungs, liver and spleen. Due to the low metabolic state of these bacilli, they have also been termed as “latent bacilli” and the process which allows survival of bacilli with low metabolic rate is termed as “latency” [10, 11]. The issue

of the presence of “dormant” bacilli in infected individuals is controversial as independent investigations have differed on the nature of these bacilli, as to whether they are latent or persistent. It has been suggested that resident *M. tuberculosis* bacilli in the system can be a mix of both slow-growing and dormant bacilli and their fate is decided by the milieu of the tissue of residence [12]. Based on this, a model has been recently proposed, which suggest that the bacilli are constantly reinfecting the epithelial cells. This model is based on observations that (1) latent bacilli can be killed with isoniazid and (2) bacilli survive for years in the lung epithelium which is frequently recycled. The reinfection process requires the presence of metabolically active bacilli, thus making them susceptible to isoniazid and it also explains the prolonged presence of bacilli in the epithelium [13]. Contrary to this, in a more recent human cohort study of infected individuals wherein the bacilli had reactivated after a gap of 20–30 years or more, molecular epidemiological data alongside genomics approaches established that *M. tuberculosis* does not replicate during latency and truly lies dormant [14].

To understand latency, a number of in vitro as well as in vivo models for *M. tuberculosis* growth and infection have been established. One of the more popular in vitro models, called “Wayne dormancy model” is based on acclimatizing cultured bacilli to microaerophilic/hypoxic conditions which is considered similar to the granulomatous environment [10, 15]. Genetic studies have revealed many genes/proteins whose expression is altered in bacilli grown in these atypical conditions. The candidates include genes involved in stress responses, molecular chaperones, metabolic enzymes and a number of signal transduction and regulatory proteins [16]. Similarly, other in vitro dormancy models based on nutrient starvation [17, 18], nitric oxide stress [19] and vitamin C treatment [20] have been developed and have yielded similar results. Despite all this, our present knowledge about the dormancy program of *M. tuberculosis* is still very limited, indicating that a highly complex regulatory process is perhaps involved in establishing mycobacterial dormancy.

#### 4.5 Insights into *M. tuberculosis* Pathogenicity Revealed by Whole Genome Sequence

The complete genome sequence of a virulent strain of *M. tuberculosis*, H37Rv was generated in 1998 [21]. *M. tuberculosis* genome is 4.4 Mb long and has the potential to encode ~4,000 proteins. The sequence revealed a number of anticipated as well as unanticipated facets of the tubercle bacilli. Preponderance of genes involved in fatty acid biosynthesis and alternate metabolism was not surprising keeping in mind the complexity of its metabolic niche and the lipids present on its cell wall. But the presence of a limited number of two-component signalling systems was a major surprise. Two-component systems (TCS) encode the most basic form of extracellular signal sensing and responding systems and are extensively present in all prokaryotes and even in some lower eukaryotes [22]. In *E. coli* a total of 65 genes encoding TCS proteins are present [23] and a similar number are found

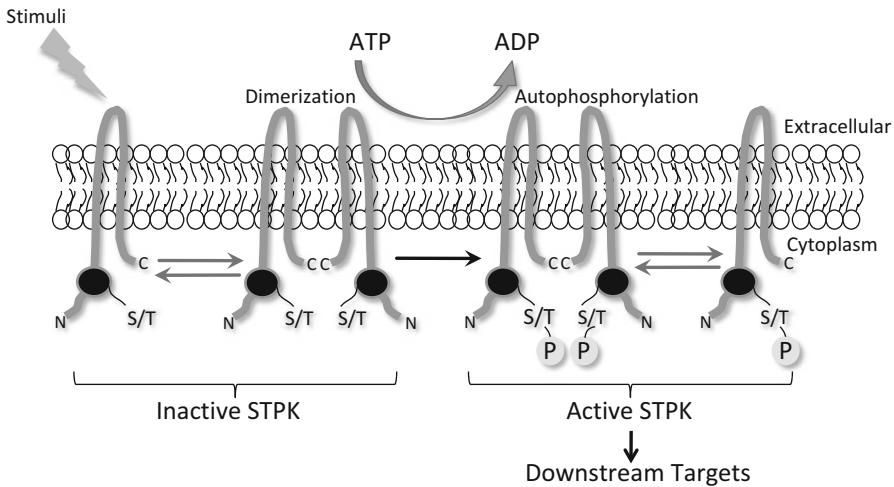
in *Bacillus subtilis* [24], whereas only 31 genes are present in the genome of *M. tuberculosis* [21]. This lack of sensory systems indicated towards existence of alternate forms of signalling mechanisms in *M. tuberculosis* and the genome analysis did reveal the presence of eukaryotic like serine–threonine protein kinases [25, 26]. The genome of *M. tuberculosis* (H37Rv strain) encodes 11 complete pairs of two-component systems, 6 orphan response regulator proteins, 2 orphan sensor kinases and 11 Ser/Thr protein kinases. Most of these proteins have been extensively characterized at both biochemical and genetic levels.

Since then a number of other pathogenic as well as non-pathogenic strains of *Mycobacterium spp.* have been sequenced, largely towards the goal of identifying and studying the genes which makes *M. tuberculosis* such a potent pathogen [27]. Homologues of *M. tuberculosis* like *M. avium*, *M. bovis* and *M. marinum*, which cause tuberculosis in birds, cattle and fish, respectively, have similar pathogenic manifestations and their genomes are also more than 99.9 % similar to that of *M. tuberculosis* [28–32]. To identify the reasons for virulence of *M. tuberculosis*, a non-virulent strain of *M. tuberculosis*, H37Ra was also sequenced. H37Ra is derived from the same lineage as its virulent counterpart, H37Rv. Comparative genome sequence analysis of H37Ra, H37Rv and another highly virulent strain, CDD1551 revealed that overall H37Ra contains 130 “H37Ra-specific” genetic changes. The changes include 39 insertions, 15 deletions and 76 single nucleotide variations, which are responsible for attenuation of the virulent strain, H37Rv. Based on small number of changes which were observed it is speculated that perhaps it is the change in regulatory and metabolic genes which has led to altered pathogenicity [33].

More recently *Mycobacterium tuberculosis* comparative genome sequencing project has been initiated by Broad Institute of Harvard along with MIT (<http://www.broadinstitute.org>). In this project 8 clinical strains of *M. tuberculosis* (virulent) are undergoing whole genome sequencing to understand the molecular basis of their pathogenicity, transmission and drug resistance.

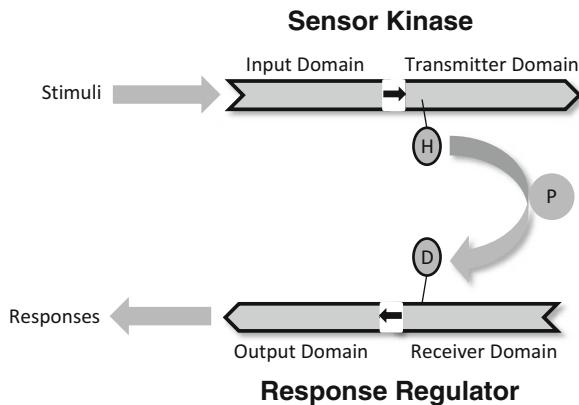
## 4.6 Signalling Systems in *M. tuberculosis*

As mentioned above, the genome sequence of *M. tuberculosis* revealed only a handful of two-component systems. It is perhaps not entirely speculative to assume that the lack of two-component signalling systems in *M. tuberculosis* genome have been compensated by the presence of eukaryotic-like serine–threonine protein kinases (STPKs). In a canonical STPK pathway, the extracellular stimulus induces dimerization of a serine–threonine protein kinase leading to *trans*-autophosphorylation at a conserved Ser/Thr residue. The phosphorylated protein then initiates an intracellular phosphorylation cascade which in turn regulates function of many proteins including transcription factors, which alter the gene expression profile of the cell (Fig. 4.1). Unlike this, in TCSs, the extracellular stimulus is detected by the sensor kinase protein, which induces its autophosphorylation at a conserved histidine residue. The phosphorylated protein in turn interacts with a response regulator protein to



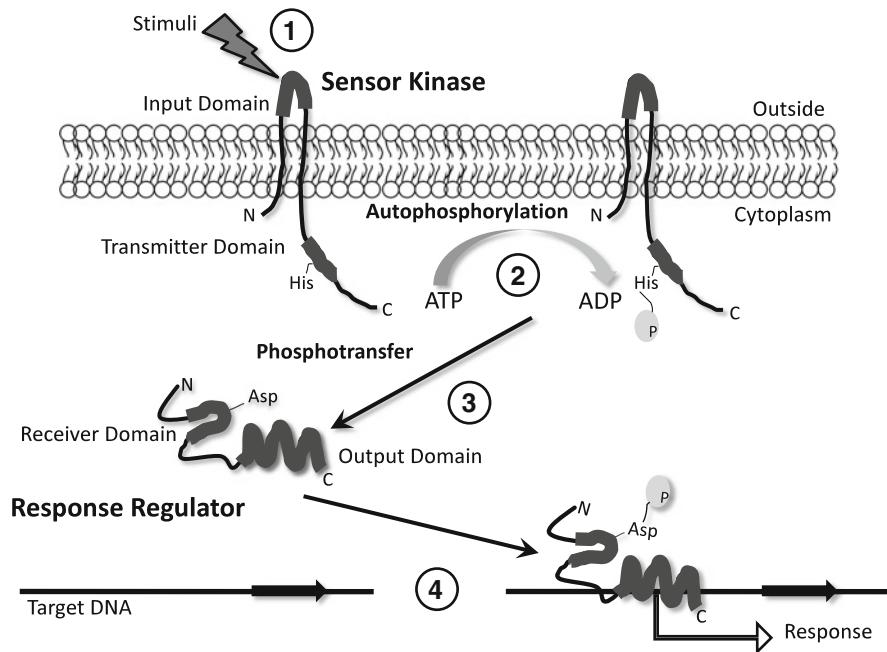
**Fig. 4.1** Canonical signalling pathway of serine–threonine protein kinases

**Fig. 4.2** Simplified diagram of signalling by two-component systems



transfer the phosphoryl group by an “energy-independent” phosphotransfer mechanism. In this reaction, the phosphoryl group from the phosphorylated histidine residue of the sensor kinase is directly transferred to the aspartate residue of the response regulator protein (Fig. 4.2). The alteration in the phosphorylation status of the response regulator protein modulates its DNA-binding ability consequently regulating the expression of target genes [34, 35].

Based on these steps, the major difference in STPKs and TCSs is that the alteration in gene expression is the only mechanism of bringing about changes in cellular physiology by TCSs, whereas STPKs can also modulate the activity of existing cellular proteins by altering their phosphorylation status. By virtue of these differences, it is expected that STPKs can respond faster than TCSs when a change in external milieu is detected and they regulate a more complex network of genes.



**Fig. 4.3** Canonical signalling pathway of two-component signal transduction systems. Steps 1–4 represent distinct stages of this pathway; these steps are also used as target reactions for identifying modalities against two-component systems for anti-microbial drug discovery

#### 4.6.1 Two-Component Signal Transduction Systems: An Overview

Most prokaryotic organisms use two-component signalling systems to sense and respond to changes in a variety of stimuli including light, oxygen, voltage, nutrients, chemicals, ions, and stress signals, which elicit changes in various physiological processes such as chemotaxis, virulence, growth and antibiotic resistance. The “two” components of this system are the *sensor kinase protein* which senses the change in the milieu and the *response regulator protein* which in turn modulates the cellular response by bringing about changes in the target gene expression. Genetically, most TCS genes are present as operons and cognate pairs are always co-expressed [22] allowing coordination of the input and output signals. The following steps are involved in a canonical signal transduction pathway of a basic two-component system:

1. The first step involves stimulus detection by the sensor kinase through its N-terminal sensory domain (step 1, Fig. 4.3). The “sensory” domain is variable across sensor kinases and detects changes in levels of various environmental factors [36].
2. The stimulus detection triggers phosphorylation of the sensor kinase at a conserved histidine residue in the catalytic domain. The phosphorylation reaction is autocatalysed, perhaps by a dimerization event which facilitates recruitment of ATP and a divalent ion (generally  $Mg^{2+}$ ) at its catalytic core (step 2, Fig. 4.3).

The sensor protein acts as a kinase (ATPase) in this step and hydrolyses ATP to yield ADP and phosphate which is covalently attached to the histidine residue by a phosphoramidate bond.

3. In the next step, phosphorylated sensor kinase interacts with the response regulator protein and transfers the phosphoryl group from its histidine to a conserved aspartate residue at the N-terminal receiver domain of the response regulator protein (step 3, Fig. 4.3) to generate an energy rich acyl phosphate.
4. The change in the phosphorylation status of the response regulator protein switches the affinity of its C-terminal DNA-binding domain to the target DNA (step 4, Fig. 4.3). The phosphorylation switch thus allows modulation of the gene expression in bacteria in response to the changes in its environment.

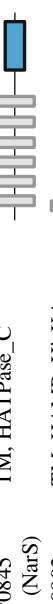
The sensor kinase and response regulator proteins contain many conserved motifs/domains categorized in various functional classes. The organization of conserved domains provides these proteins modular architecture wherein the individual domains retain biological activity in the absence of other domains. This organization is similar to most signalling proteins, where activity of the catalytic domain is regulated by a distinct regulatory domain. This modular architecture also facilitates characterization of catalytic and regulatory domains independently, without interference from the other domain [22].

In a typical sensor kinase the N-terminal module contains transmembrane domain/s which facilitates localization of these proteins to the plasma membrane along with the “sensory” domains. In mycobacterial sensor kinases, sensory domains such as GAF, PAS and HAMP have been reported. The C-terminal domain is catalytic and is characterized by the presence of ATPase/kinase activity. This kinase domain is involved in ATP binding and phosphorylation of the conserved histidine residue, to which phosphoryl group is attached after hydrolysis of ATP. The kinase activity of the C-terminal domain is modulated by activation of the N-terminal sensory domain by the stimulus making the system sensitive to alterations in the stimulus. Table 4.1 lists the domains and their location based on sequence analysis for all the sensor kinase proteins of *M. tuberculosis*.

Similarly, in response regulator proteins, the N-terminal receiver domain is catalytic, which receives the phosphate from the phosphorylated sensor kinase. The domain is marked by the presence of a conserved aspartate residue which accepts phosphate from the histidine residue of sensor kinase. This domain also contains additional conserved aspartate residues, which bind to divalent metal ions like  $Mg^{2+}$ , essential for catalyzing the phosphotransfer reaction. The C-terminus generally has transcription regulatory domains having DNA-binding activity. The phosphorylation status of the receiver domain alters the DNA-binding ability of this domain by virtue of structure perturbation, allowing response regulator proteins to act as transcriptional “switches.” Table 4.2 lists the domains/motifs and their locations predicted by sequence analysis of all the response regulator proteins of *M. tuberculosis*.

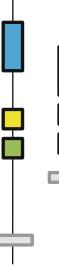
As mentioned previously, the genome of *M. tuberculosis* H37Rv encodes 11 complete pairs of TCSs, 2 orphan sensor kinases and 6 orphan response regulators. Orphan TCS are those proteins for which the cognate sensor kinase or response regulator protein is not co-expressed or present as an operon. These orphan proteins

Table 4.1 Domains, motifs and their locations in sensor kinase proteins of *M. tuberculosis*

Sensor kinase	Domains/motifs <sup>b</sup> [36]	Domain organization	Histidine position	Stimulus (if known) <sup>c</sup>
Rv0490 (SenX3)	TM, HisKA, HATPase_C		167 <sup>a</sup>	Redox state, phosphate [81]
Rv600c	HATPase_C		NA	Hypoxia?
Rv601c	HAMP		131 <sup>a</sup>	Hypoxia?
Rv0758 (PhoR)	TM, HAMP, HisKA, HATPase_C		259	Mg <sup>2+</sup> [64]
Rv0845 (NarS)	TM, HATPase_C		241	?
Rv0902c (PrrB)	TM, HAMP, HisKA, HATPase_C		240	Host-cell interaction, entry into macrophage [44]
Rv0982 (MprB)	TM, HAMP, HisKA, HATPase_C		249 <sup>a</sup>	Nutrient starvation, hypoxia [66]
Rv1028c (KdpD)	TM, HisKA, HATPase_C		642	Nutrient starvation [18]
Rv1032c (TrcS)	TM, HisKA, HATPase_C		287	?
Rv3132c (DevS)	GAF, HATPase_C		395 <sup>a</sup>	Hypoxia, nitric oxide, carbon monoxide, ascorbic acid [50]

(continued)

Table 4.1 (continued)

Sensor kinase	Domains/motifs <sup>b</sup> [36]	Domain organization	Histidine position	Stimulus (if known) <sup>c</sup>
Rv3245c (MtrB)	TM, HAMP, HisKA, HATPase_C		305	Iron [42]
Rv3764c (TcrY)	TM, HAMP, HisKA, HATPase_C		256 <sup>a</sup>	Iron limitation [84]
Rv2027c (DosT)	GAF, HATPase_C		392 <sup>a</sup>	Hypoxia, nitric oxide, carbon monoxide, ascorbic acid [50]
Rv3220c (Pdtas)	HATPase_C		303	?

The sequence for all the sensor kinase proteins was retrieved from the genome sequence of *M. tuberculosis* H37Rv and analysed on SMART server (Simple modular architecture research tool)

<sup>a</sup>Experimentally validated

<sup>b</sup>Motif location is based on predicted position across entire length of protein

<sup>c</sup>On the basis of target gene induction or direct activation

Key: Grey, transmembrane domains (TM); yellow, histidine kinase domain (HisKA); blue, Histidine kinase-type ATPase catalytic domain (HATPase\_C); green, HAMP (Histidine kinases, Adenylyl cyclases, Methyl-binding proteins, Phosphatases) domain; and golden, GAF (cGMP phosphodiesterase, Adenylyl cyclase, FhA) domain

**Table 4.2** Domains, motifs and their location in the response regulator proteins of *M. tuberculosis*

Response regulator	Domains/motifs	Domain organization <sup>a</sup>	Aspartate position	Target genes <sup>b</sup>
Rv0491 (RegX3)	Rec, Trans_reg_C		52 <sup>c</sup>	<i>ald, cydAB, gltA1, psfS</i> [81]
Rv0602c (TcrA)	Rec, Trans_reg_C		73 <sup>c</sup>	–
Rv0757 (PhoP)	Rec, Trans_reg_C		71	<i>ESAT-6, CFP-10, devRS</i> [71]
Rv0844c (NarL)	Rec, HTH		61	–
Rv0903c (PrrA)	Rec, Trans_reg_C		58	–
Rv0981 (MprA)	Rec, Trans_reg_C		48 <sup>c</sup>	<i>sigE, sigB, pepD</i> [77]
Rv1027c (KdpE)	Rec, Trans_reg_C		52	–
Rv1033c (TrcR)	Rec, Trans_reg_C		82	<i>Rv1057</i> [87]
Rv3133c (DevR)	Rec, HTH		54 <sup>c</sup>	<i>Rv2623, narK2, Rv2031c-Rv2028c, tgs1, Rv2628, Rv1997, Rv1733c, hspX</i> [50]
Rv3246c (MtrA)	Rec, Trans_reg_C		53	<i>iniB, kgtP, fbpB, Rv0674, dnaA</i> [43]
Rv3765c (TcrX)	Rec, Trans_reg_C		54 <sup>c</sup>	–
Rv0195 Orphan	HTH		–	–
Rv0260c Orphan	Trans_reg_C		–	–
Rv0818 Orphan	Trans_reg_C		–	–
Rv1626 (PdtA-R)	Rec		65	–
Rv2884 Orphan	Trans_reg_C		–	–
Rv3143 Orphan	Rec		64	–

The sequence for all the response regulator proteins was retrieved from the genome sequence of *M. tuberculosis* H37Rv and analysed on SMART server (Simple modular architecture research tool)

<sup>a</sup>Motif location is based on predicted position across entire length of protein

<sup>b</sup>Identified by direct DNA-binding assays

<sup>c</sup>Experimentally validated

Key: Purple, receiver domain (Rec); red, transcription regulation domain (Trans\_reg\_C); blue, helix-turn-helix DNA-binding domain (HTH)

can communicate with other orphan proteins or they can also communicate with other proteins already paired with their respective partner proteins.

Tables 4.1 and 4.2 lists the conserved features of all the sensor kinases and response regulator proteins encoded by the genome of *M. tuberculosis* H37Rv. Table 4.1 also lists the conserved histidine residue in the sensor kinase protein which is either predicted or has been experimentally shown to undergo phosphorylation and the stimulus which is detected by them to activate the pathway. Similarly Table 4.2 also lists the conserved aspartate residue in the response regulator protein which is either predicted or has been experimentally shown to undergo phosphorylation and the target genes to which direct binding of the response regulator have been demonstrated.

## 4.7 Role of TCS in Mycobacterial Pathogenicity

Based on the regulatory role of TCSs in cellular physiology, it is anticipated that they will also play important roles in mycobacterial pathogenicity. Investigations about the role of individual TCS proteins have demonstrated their essentiality in virulence of *M. tuberculosis*. The role of a few of the TCSs of *M. tuberculosis* in its virulence was examined by Parish et al. [37] and four of the six TCS examined had an effect on the bacterial virulence. Similarly, in vivo expression of response regulator genes in a macrophage infection model were examined and mycobacterial regulatory genes which were induced on infection in host cells were identified [38]. Based on these observations and other studies involving TCSs in mycobacterial survival and virulence, we have classified them in three basic categories:

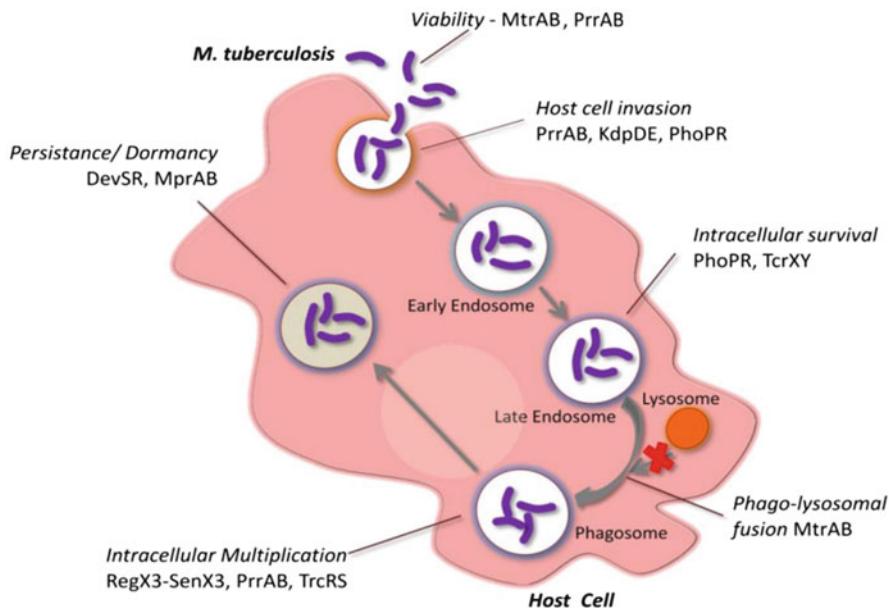
### 4.7.1 Two-Component Systems Essential for Mycobacterial Survival

MtrAB and PrrAB are the only TCSs found till now which are essential for cellular viability [39, 40]. The response regulator *mtrA* and *prrAB* operon could be knocked out only when an episomal copy of the gene was present. Surprisingly, in vivo studies revealed that the overexpression of MtrA protein prevented the multiplication of *M. tuberculosis* in macrophages, mouse liver and spleen on account of impaired ability to block phagolysosome formation [41]. The levels of MtrA have also been shown to be induced twofold in high-iron conditions [42] and *dnaA*, *iniB*, *kgtP* have been identified as gene targets [43].

The *prrAB* was identified from a transposon mutagenesis library of *M. tuberculosis* which showed defect in early phase of infection in bone marrow derived macrophages [44]. The transcripts for this TCS are also induced after infection into macrophages [45]. Based on these findings it has been suggested that PrrAB system is required for adaptation in the early phases of intracellular growth as well [44, 46]. The system has been characterized at the biochemical level [46, 47] but phosphorylation sites for PrrA as well as PrrB, signals recognized by PrrB and the target genes of PrrA are yet to be known.

### 4.7.2 Two-Component Systems Essential for Mycobacterial Virulence

In a comprehensive study to evaluate the role of TCSs in mycobacterial virulence, Parish et al. knocked out 6 of the 11 TCSs and analysed the impact of individual strains on virulence using an immunodeficient SCID mouse infection model. They found that deletion of four TCSs viz. *devRS*, *tcrXY*, *trcRS* and *kdpDE* resulted in hypervirulence phenotypes [37]. Similar to this, other studies have shown the involvement of PrrAB,



**Fig. 4.4** Involvement of two-component systems of *M. tuberculosis* in its pathogenicity during various stages of infection of a host cell. The diagram is based on known route taken by *M. tuberculosis* bacilli after infection into host cells. The role of various two-component systems is based on previous reports

PhoPR, SenX3-RegX3 and MprAB in various stages of mycobacterial pathogenesis. Figure 4.4 lists the involvement of various TCSs in mycobacterium physiology using a simplified *in vivo* infection pathway in a host cell. The two-component systems which play a role in mycobacterial pathogenicity are described below:

#### 4.7.2.1 DevRS (DosRS) and DevR-DosT (DosRT) Systems

These are the most well-characterized two-component systems of *M. tuberculosis* and were named *dev* on the basis of their differential expression in virulence or *dos* for dormancy survival. The genes encoding the DevRS two-component system were found to be differentially expressed between a virulent, H37Rv and H37Ra, an avirulent strain [48, 49]. Later studies showed that the system is also critical for dormancy survival of *M. tuberculosis*, hence the name *dos* was coined by an independent group of investigators [16, 50]. The response regulator protein, DevR is common to two sensor kinases DevS (DosS) and Rv2027c (DosT), both of which can phosphorylate DevR protein [51, 52]. This crosstalk allows both sensor kinases to modulate the same genetic network regulated via the DevR response regulator protein. A number of extracellular stresses which serve as signals for these sensors have been reported including oxygen and hypoxia [17, 53–55], redox [56], nitric

oxide [57], heme [19, 56, 58], nutrient starvation [18] and vitamin C [20], which are sensed by two conserved GAF domains present in both the sensor kinase proteins [55]. Since these triggers are primarily present during mycobacterial latency, these TCSs are considered as operators of the dormancy regulon in *M. tuberculosis* [10, 59].

Consequently, these proteins were also found to be critical for mycobacterial virulence [37, 49] and survival during in vivo latency in granulomas [60]. It has also been postulated that both sensor kinases detect same environmental cue/s but with differential sensitivities thereby responding in a stepwise manner [20, 59]. More than 100 genes are targets of this TCS including stress response genes such as *Rv2626c*,  $\alpha$ -crystallin and *narK2* [16]. The proteins of this TCS have been crystallized, and the ligand-binding site has been characterized using the structure information [56, 58, 61]. Due to importance of these proteins in mycobacterial virulence, they have also been projected as targets for anti-tuberculosis drug discovery [62, 63].

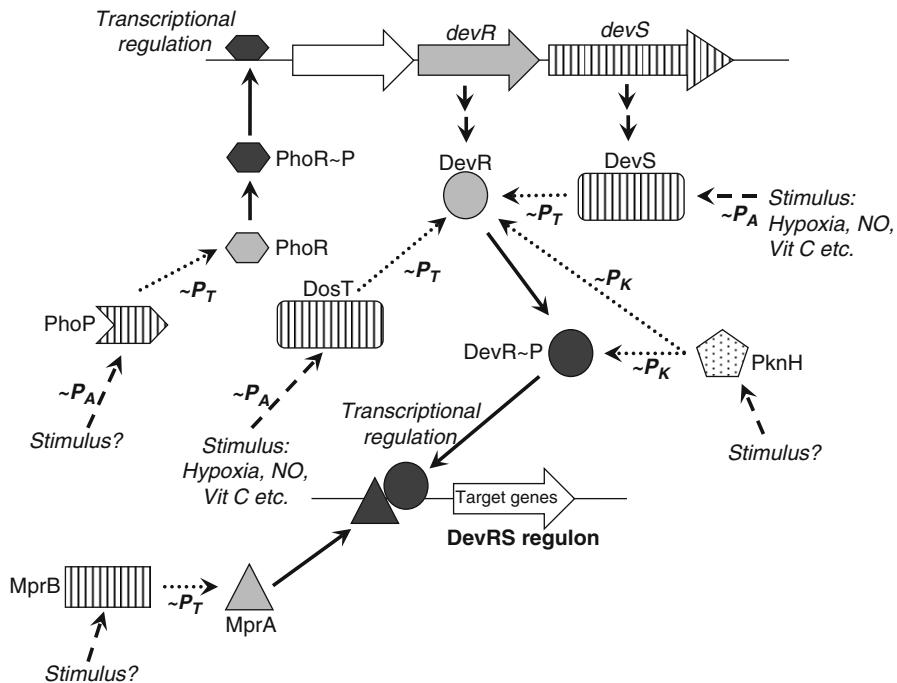
Another interesting aspect of this system includes its regulation by other mycobacterial signalling proteins such as PhoP [64] and the serine threonine kinase, PknH [65], providing another layer of regulation of these proteins, which is not yet determined for other TCS proteins. Some of the genes of the DevRS regulon were also found to be co-regulated by the MprAB TCS [66]. Furthermore, in agreement with the above mentioned roles of these TCSs, the DevRS system was also found to be naturally mutated and constitutively upregulated in highly pathogenic Beijing strain of *M. tuberculosis* [67].

#### 4.7.2.2 PhoPR System

Early studies demonstrated the importance of this system in mycobacterial virulence when inactivation of the *phoPR* TCS resulted in significant attenuation of bacillary growth in infected cells and mice models [68, 69]. It is predicted that this system detects changes in the levels of extracellular Mg<sup>2+</sup> ions. More recently, genome sequencing and comparison of avirulent strain *M. tuberculosis* H37Ra with its virulent counterpart *M. tuberculosis* H37Rv revealed the presence of a single nucleotide variation in the gene encoding PhoP response regulator protein [33, 70]. PhoP of H37Ra contains S219L substitution which is present in the highly conserved  $\alpha$ 3 helix of its DNA-binding domain. This change is predicted to alter the expression of downstream genes critical for mycobacterial pathogenicity such as ESAT-6 and CFP-10 [64, 70–72]. The system has also been shown to regulate expression of other molecules such as manLAM (mannose-capped lipoarabinomannans), some trehaloses, sulpholipids and DevRS TCS which are essential for mycobacterial pathogenicity [64, 73]. Overall, this system is considered critical for mycobacterial virulence and is also a potential therapeutic target.

#### 4.7.2.3 MprAB System

Initial studies to identify regulatory elements whose expression is critical for entrance and maintenance of persistent infection involved in mycobacterial



**Fig. 4.5** Regulation of DevR signalling by other signalling pathways of *M. tuberculosis*. Modulation of expression of target genes of DevRS regulon by environmental stimuli detected by DevRS, MprAB and PhoPR TCS and PknH STPK. Grey response regulator proteins in basal state, black phosphorylated response regulator proteins, stripes sensor kinase proteins, dotted lines phosphorylation-mediated change, dashed lines stimulus,  $\sim P_A$  autophosphorylation reaction,  $\sim P_T$  phosphotransfer reaction and  $\sim P_K$  represents a kinase-mediated phosphorylation reaction

virulence led to identification of a mycobacterial persistence regulator, later termed as *mprA* [38]. Later, the genes were found to be induced during infection in mice [74], nutrient starvation and hypoxia [75]. The system has been characterized at the biochemical level [76], and *pepD*, *sigE* and *sigB* genes have been identified as its target genes [75, 77]. In a study using *mprA* knockout strain, expression of a large number of genes of DevRS regulon were found to be modulated [66] suggesting another instance of cross-regulation of signalling networks (Fig. 4.5).

#### 4.7.2.4 SenX3-RegX3

This was the first two-component system of *M. tuberculosis* to be characterized at the biochemical and genetic level [78]. Subsequent functional studies demonstrated that this system is involved in mycobacterial virulence in both immunodeficient and

immunocompetent mice [79]. More recently, in the SenX3 sensor kinase, presence of a PAS domain (which senses changes in levels of oxygen and redox) was predicted using structure-based homology modelling approaches. Concomitantly, it was also established that the two-component system is critical for virulence of the bacilli [80] and it senses both redox state and phosphate levels [81].

#### 4.7.2.5 KdpDE System

Alteration in ionic concentration of monovalent and divalent ions is considered as a potent cue for bacteria to know the niche they are residing in. The Kdp system senses changes in levels of  $K^+$ , in host cells while the extracellular levels are 5 mM, intracellular levels are significantly higher at 140 mM. This change may allow the bacteria to know when it shifts from extracellular to intracellular environment. This ionic variation induced change in osmotic stress is detected by KdpD which in turn regulates the expression of *kdpFABC* operon which encodes a  $K^+$  transport system. The deletion of *kdpDE* results in hypervirulent phenotype in SCID mice [37], but no other evidence exists for its role in mycobacterial virulence. Biochemically, it has been demonstrated that the sensory domain of KdpD forms ternary complexes with two membrane lipoproteins, LprJ and LprF, hinting that perhaps they serve as accessories to the ligand-binding domain [82]. Biochemically the system is still uncharacterized and no downstream gene targets are known.

#### 4.7.2.6 TcrXY System

Inactivation of this TCS shows hypervirulent phenotype in the SCID mouse model of infection [37]. The proteins of this TCS have been characterized at biochemical level but no target gene for the same is known till date [83]. The system is induced under limited iron availability in culture [84] and *tcrX* has been shown to be transiently expressed in human macrophages [85].

#### 4.7.2.7 TrcRS System

Similar to *tcrXY* TCS, inactivation of this system also shows hypervirulent phenotype in the SCID mouse model [37]. These proteins have been characterized at biochemical level [86] but amino acid residues which undergo phosphorylation in both TrcS and TrcR are yet to be identified. Studies using microarray have compared the *M. tuberculosis* wild type and *trcS* mutant and found out approximately 50 genes which were regulated differentially between these strains. Intracellularly expressed Rv1057 has been identified as its target gene to which direct physical binding has been demonstrated [87, 88].

#### 4.7.3 ***Two-Component Systems Having No Role in Mycobacterial Pathogenicity or Two-Component Systems with Uncharacterized Roles***

For NarLS, PdtaRS and TcrA two-component systems, no association with virulence of *M. tuberculosis* is known till date. Three protein member Rv600c/Rv601c-TcrA TCS is extensively characterized at biochemical level [89–91], while Rv0600c has been implicated in sensing the environmental signal which mediate ATP hydrolysis to transfer the phosphoryl moiety to a conserved His at Rv0601c protein, from which finally the phosphoryl group is transferred to the response regulator protein, TcrA [91]. The environmental signal to which Rv0600c responds, is still yet to be defined. For NarLS TCS, only the crystal structure for NarL response regulator protein is available [92]. Phosphorylation studies for NarLS have not been performed yet. In the SCID mouse infection studies, no effect was observed on the growth of bacilli, for a  $\Delta$ narL mutant and a pdtaS mutant [37]. PdtaS is an orphan sensor kinase, which is biochemically active and capable of phosphorylating an orphan response regulator protein, Rv1626. Rv1626 is a phosphorylation dependent transcription antitermination regulator, termed as PdtaR [93, 94].

### 4.8 **Integration of Signalling Systems in *M. tuberculosis***

In classical two-component signalling there is strict specificity of communication between a sensor kinase and its cognate response regulator protein, but exceptions to this rule have been observed. The cross communication between a sensor kinase and a “non-cognate” response regulator protein can occur as long as they are structurally similar to their cognate protein. Genome wide phosphotransfer profiling performed for all the TCSs of *Caulobacter crescentus* [95, 96] and *E. coli* [97] also suggests towards the presence of specificity in the sensor kinase and response regulator protein interactions [98], but at the same time the interactions were not completely exclusive and many instances of crosstalk were observed [95, 97].

For *M. tuberculosis*, the non-cognate pair communication has been detected between DevR response regulator protein and an orphan sensor kinase, DosT (Rv2027c) suggesting that perhaps it may occur for other TCSs of *M. tuberculosis* [51]. PdtaR is an orphan response regulator and its interacting partner PdtaS was identified by homology modelling using STRING program. This two-component system (in addition to four other cognate two-component systems—*regX3-senX3*, *prrA-prrB*, *mprA-mprB*, *mtrA-mtrB*) is found to be present in all the mycobacterium genomes sequenced till date [93]. Besides these non-cognate interactions, no other cross-communication pathway has been identified in *M. tuberculosis*.

Unlike direct “contact”-dependent communication of two proteins from different signalling networks as mentioned above, transcriptional modulation of one

**Table 4.3** Role of all the two-component systems of *M. tuberculosis*

Two-component system		Modulating stimulus	Physiological role	Link to other pathways
Sensor kinase	Response regulator			
SenX3	RegX3	Redox state, phosphate	Growth and virulence	–
Rv0600c/ Rv0601c	TcrA	Hypoxia?	–	–
PhoR	PhoP	Mg <sup>2+</sup>	Complex lipid biosynthesis and Virulence	Regulates DevRS operon
NarS	NarL	?	–	–
PrrB	PrrA	Macrophage entry, nitrogen limitation	Survival, growth and virulence, intracellular growth	–
MprB	MprA	Nutrient starvation, hypoxia	Maintenance of persistence	Regulates DevRS regulon
KdpD	KdpE	Nutrient starvation	Growth and virulence	–
TrcS	TrcR	?	–	–
DevS	DevR	Hypoxia, nitric oxide, carbon monoxide, ascorbic acid	Virulence and dormancy	PhoPR and MprAB TCS; PknH STPK
MtrB	MtrA	Iron	Survival and resistance within host	–
TcrY	TcrX	Iron limitation	Virulence	–
DosT (Orphan)	DevR	Hypoxia, nitric oxide, carbon monoxide, ascorbic acid	Virulence and dormancy	–
PdtaS (Orphan)	PdtaR (Orphan)	?	–	–
–	Rv0195 (Orphan)	?	–	–
–	Rv0260c (Orphan)	?	–	–
–	Rv0818 (Orphan)	?	–	–
–	Rv2884 (Orphan)	?	–	–
–	Rv3143 (Orphan)	?	–	–

The cellular cues (triggers), the physiological role shown by gene knock-down approaches (if known) and the link of the TCS with any other signalling pathways is indicated

signalling pathway by another is relatively more common. In *M. tuberculosis* it has been observed that the PhoPR TCS directly regulates the expression of the dormancy associated *devRS* operon [64]. Similarly, other signalling proteins such as PknH and MprAB modulate the activity of DevR protein and DevRS regulon respectively [65, 66]. Table 4.3 lists the features of all TCSs of *M. tuberculosis* in terms of extracellular stimulus which modulate them, their physiological role along

with details of other signalling pathways which communicate with them. Figure 4.5 shows integration of various signalling networks with DevRS pathway and how multiple signals can modulate the DevRS regulon.

## 4.9 Two-Component Systems as a Drug Target

### 4.9.1 *Current Anti-tubercular Therapy*

Keeping in mind the nature of tuberculosis infection and the prolonged therapy needed for eradicating the infection, World Health Organization (WHO) recommends a DOTS (directly observed therapy short course) treatment for drug-susceptible (DS)-TB. It involves an initial administration of isoniazid, rifamycin, pyrazinamide and ethambutol for the first 2 months followed by isoniazid and rifamycin for the next 4 months. For MDR-TB, treatment includes a combination of eight to ten drugs with therapies lasting up to 18–24 months, whereas, the treatment options for XDR-TB are very limited.

### 4.9.2 *Limitations of Current Therapies*

Long duration and complexity of current therapies which result in patient non-compliance is one of the major limitations of the current therapeutic regimen. Adverse effects of the drugs used for the treatment also lead to non-adherence. Consequently, the increasing incidences of MDR, XDR and TDR (totally drug resistant) strains and co-infection of TB and HIV have become a major problem. BCG is the only vaccine available for tuberculosis, and appears most effective against childhood forms of disease, but is inefficient in preventing adult pulmonary tuberculosis.

### 4.9.3 *Targets Under Study for New Drug Development*

Attempts are continuously being made to combat tuberculosis by developing new drug regimens or formulations which can overcome the limitations of the current therapy. The thrust areas focused on making the treatment effective include, shortening the duration of the treatment, reducing the drug toxicity, demonstrable activity against MDR/XDR-TB, minimal interaction with medications for HIV and other co-medications and most importantly, to have potential to treat latent TB.

The fluoroquinolones including moxifloxacin and gatifloxacin are being investigated in clinical trials to shorten TB treatment [99, 100]. Similarly,

nitroimidazopyrans, diarylquinolines and Diamine-SQ109 are currently in phase I/II clinical trials [101–103] as potential anti-TB drugs. Towards the objectives of developing good vaccines, few candidate mycobacterial proteins including **MVA85A**, **rBCG30**, **72F fusion protein** and **ESAT6-Ag85b fusion protein** are in the stage of testing.

#### 4.9.4 TCS as Potential Target

Owing to the roles played by TCSs in *M. tuberculosis* survival and physiology and the absence of TCSs in higher eukaryotes, these systems *per se* are attractive targets for therapeutic intervention against pathogenic bacteria. The observed attenuation of *phoP* and/or *phoR* mutant strains has led to investigations of the potential utility of an *M. tuberculosis* *phoP* mutant (SO2 strain, a derivative of clinical isolate MT103) as a possible live-vaccine candidate [104]. Similarly, an independent bioscreening approach targeting DevRS TCS, led to isolation of “compound 10,” which can “lock” DevR in an inactive conformation [62]. This “locking” prevents it from binding target DNAs and renders it incapable of inducing the dormancy regulon, supposedly critical for latency responses, thereby making DevR a novel drug target against “dormant” bacilli.

Different steps of the TCS pathway which can be used as target for drug discovery are mentioned as steps 1–4 in Fig. 4.3 and continuous efforts are being made to discover new TCS inhibitors. It is anticipated that further optimization and pharmacological evolution of “compound 10” may lead to development of a potent drug against dormant bacteria. Furthermore, a combinatorial approach of combining molecules targeting bacterial components involved in both active and latent stages, may lead to the development of the therapy which would be effective against active as well as latent bacilli.

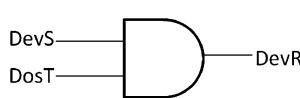
#### 4.9.5 Future of Drug Therapy and Impact of Latency

The current anti bacterial drugs target enzymes or metabolic pathways involved in biosynthesis of cell wall, DNA, RNA or proteins. These drugs show maximum activity against dividing cells and are not effective against latent bacilli which have low metabolic rates. No new drugs for tuberculosis treatment have been introduced for past 40 years, though many lead compounds have reached clinical evaluation stages. As mentioned above, since all the existing treatments have long therapeutic regimen, the major goals of any new drug against tuberculosis mandates a shorter treatment regimen and higher efficacy [105].

## 4.10 Mathematical Modelling of Two-Component Signalling

Today mathematics is playing an increasingly important role in biology. From modelling currents in ion channels and neural circuits to networks of cellular processes and molecular interactions, mathematics can provide insights that biology alone cannot decipher. Most complex biological systems are modelled using a reductionist approach. First, a simple network of interactions between all known and measurable components of the system is sketched followed by the formation of equations to describe the kinetics of various interactions and transitions from one state to another. The use of logic circuits is a further simplification where inputs and outputs can be denoted as existing in two states alone, eliminating the need to consider gradients in stimulus and response. Though, not all interactions in signalling pathways are strictly one-to-one, highlighting the existence of a complex circuitry where a single response is modulated by more than one upstream component.

Using the help of well-known two-component systems of *M. tuberculosis*, DevR-DevS and DosT-DevR, we have modelled these relatively simple systems consisting of three interacting partners using logic gates, enabling a theoretical prediction of the output given input parameters. As demonstrated experimentally, DevR can be phosphorylated by both DevS and DosT [51, 52]. Furthermore, the sensor kinases (SK) also have dephosphorylation activity towards phosphorylated response regulator (RR) protein, thus, RR phosphorylation by one activated SK may be countered by the RR dephosphorylation activity of the inactive one. It can thus be predicted that at unimolecular concentrations and assuming equal affinities of both SKs, DevS and DosT for RR, DevR; DevR phosphorylation will occur only if both DevS and DosT are activated by their respective stimuli. This is a classic example of an AND gate where a positive output is obtained only when both input sources are active due to the antagonistic nature of their “off” and “on” states. Figure 4.6 depicts DevR-DevS-DosT circuitry and its corresponding truth table. This model of an AND gate can then be tested in an *in vitro* or *in vivo* scenario after



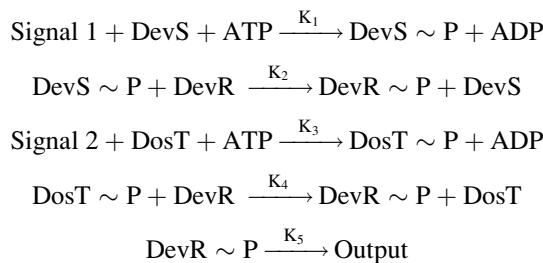
	DevS (Input 1)	DosT (Input 2)	DevR (Output)
	+	+	+
	+	-	-
	-	+	-
	-	-	-

**Fig. 4.6** DevS-DevR-DosT interaction network shown in the form of a flow chart, a logic gate and a truth table. The table is based on assumption that all the three catalytic activities, autophosphorylation, phosphotransfer and dephosphorylation of the SKs, DevS and DosT have comparable kinetic affinities. The +ve inputs in the truth table refers to presence of stimuli and +ve output refers activation of downstream genes, while -ve indicates the absence of stimulus and no gene activation respectively

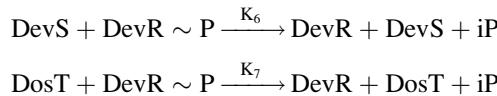
the gradual addition of layers of complexity to the network such as relative rates of phosphorylation and dephosphorylation of the SKs and the RR, basal expression levels and abundance of the interacting proteins, the effect of variation in physiological conditions such as salt concentrations, pH, temperature and stress and the possibility of non-specific interaction partners in the cellular milieu which incidentally are known to modulate these two-component systems.

Further, the DevS-DevR-DosT network can also be modelled as a set of equations, with variables acting as parameters that determine the response. Such equations are usually in the form of transition kinetics with interaction affinities defined by association and dissociation rate constants as well as rate constants of reactions.

Reactions that favour a positive response are the following:



Reactions that favour a negative response are the following:



Where  $\text{DevS} \sim \text{P}$ ,  $\text{DosT} \sim \text{P}$  and  $\text{DevR} \sim \text{P}$  are the phosphorylated forms of DevS, DosT and DevR proteins, respectively.

The positive output is generated by  $\text{DevR} \sim \text{P}$  only if

$$K_1 + K_2 + K_3 + K_4 + K_5 > K_6 + K_7;$$

Adding parameters such as the rate of autodephosphorylation of DevR, makes the reaction network more complex, however, these parameters also provide greater physiological relevance to the mathematical model.

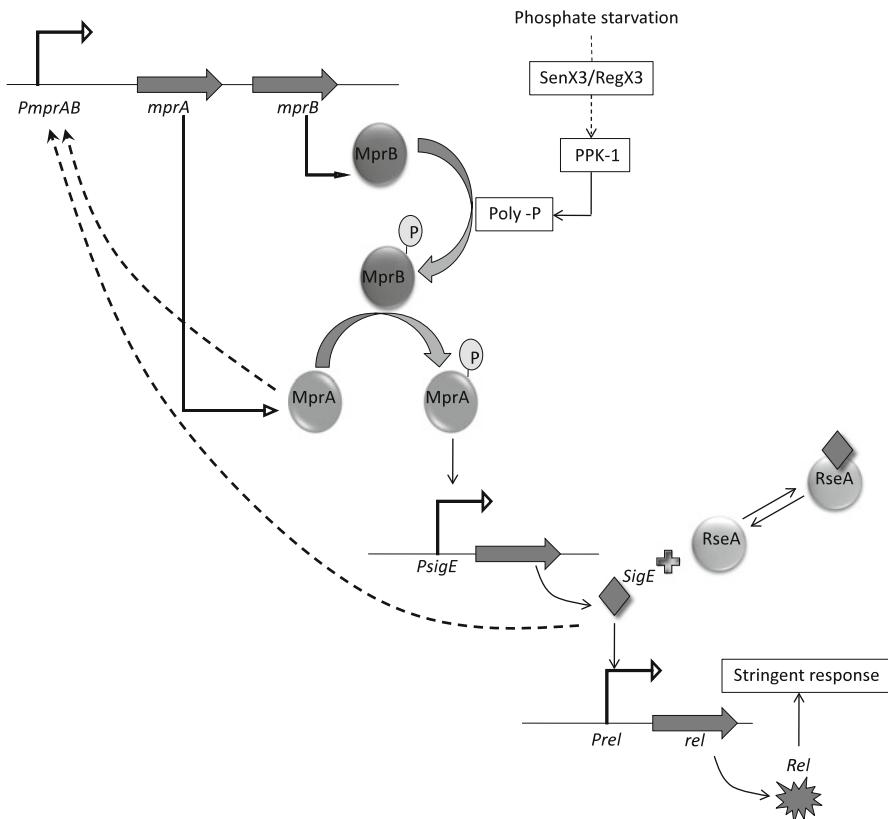
From this model, we hypothesize that DevR activation requires positive signals from both DevS and DosT and that DosT, through its phosphatase activity could act as a potent inhibitor of response to sub-optimal stimuli. A study by Honaker et al. [59] has proposed a role for DosT as the initiator of the hypoxic response, where they have found that DosT induces DevS through the DevRS regulon, followed by a fall in the activity of DosT while DevS maintains the response at peak levels. However, the mechanism by which the hypoxic response is switched off and the roles of inactive DosT and DevS as sinks which deactivate DevR are yet

to be elucidated. Although this system has been extensively characterized at the molecular level, there are some observations, including one mentioned above, which are yet to be explained. For example, clinical isolates of *M. tuberculosis* from Colombia have recently shown to harbour a 3.6 kb deletion resulting in a loss of the DevR regulon genes [106]. The strain appears to have retained its virulence despite the lack of this apparently virulence determining gene segment. Contrastingly, the multi-drug resistant W-Beijing strain overexpresses the DosR regulon, a characteristic which has lent it hypervirulent phenotype [107]. These evidences point to an underlying complexity in the DevR regulon network as well as the possibility that there may be other factors that influence virulence, thereby making the DevRS system extremely interesting. Comprehensive mathematical models however have not emerged due in part to the high degree of complexity in the DevR-DosT-DevS system with multiple interaction partners and bidirectional transfers of the phosphate current.

Conversely, the MprAB system, which regulates the stress response network in mycobacteria has been extensively modelled due to the simple one-to-one nature of its interaction and the lack of non-cognate components adding to the network. Modelling of the MprAB system has yielded intriguing insights into the functioning of bistable systems and the role of noise and positive feedback in the maintenance of such signalling networks. It has been shown that besides the MprA/MprB two-component system (TCS), the stress response network is also regulated by alternative sigma factor  $\sigma$ E, its anti-sigma factor RseA and enzyme polyphosphate kinase 1 (Ppk1) [108].  $\sigma$ E regulates the expression of the stringent response regulator, RelA and Ppk1 is involved in synthesis of polyphosphate required for the phosphorylation of MprB [109].

The mechanisms controlling bistability in this stress response network have been analysed using mathematical modelling methods [110]. Using a decoupling approach the signalling network was split into transcriptional and post-translational modules which facilitated a reduction in the dimensionality of the hitherto dynamical system. The transcriptional module was built using a set of equations describing the basal and activated transcription states of the promoter. The translational module involved equations for translation reactions, mRNA degradation and protein degradation rates. Post-translational modules included phosphorylation and dephosphorylation reactions of the MprAB system as well as the  $\sigma$ E-RseA interaction. The authors found that bistability was a product of interactions at the protein, transcription and translation levels. MprA/MprB positive autoregulation as well as  $\sigma$ E activation were insufficient to maintain bistability while addition of the post translational regulation of  $\sigma$ E by RseA allowed for the maintenance of bistability.

Sureka et al. reported bimodal expression of the stringent response regulator Rel in a population [108]. Using rate equations for transcriptional and translational processes involved in the system, steady-state expressions were found to be valid over a wide parameter range. This bimodality was shown to be a product of gene expression noise. Moreover, evidence for hysteresis in *rel* expression was observed, which is an indicator of bistability. Hysteresis occurs because the stimulus required



**Fig. 4.7** Genetic and biochemical pathways which have been shown to regulate the persistence response in *M. tuberculosis*. These pathways have formed the basis of various mathematical modelling studies

to change from the low to the high state is greater than the stimulus required to reverse the transition. Thus gene expression noise involving positive feedback is an important parameter regulating persistence, which can be linked to phenotypic heterogeneity arising in a genetically identical bacterial cell population with “persisters” forming a subpopulation of the total cell population.

Figure 4.7 shows the important components of the MprAB dependent stress or stringent response pathway. Here, a positive feedback loop is active in the network as the production of MprA~P enhances MprA synthesis in stress conditions, acting as a “persistence regulator.” The *mprAB* operon has a basal level of gene expression [75, 76] irrespective of the positive feedback. Downstream of *mprAB* operon, the phosphorylation status of the MprA protein regulates the expression of the alternate sigma factor *SigE*, which in turn regulates the transcription of *RelA*, the universal stress response regulator. The second feedback loop involves regulation of the levels of MprAB through *SigE* thus bringing in the second layer of regulation.

It is obvious from these evidences that mathematical approaches such as those highlighted above will play a key role in defining and remodelling the approaches taken to combat tuberculosis and other diseases.

## 4.11 Future of Two-Component System Signalling Research

Classical methods of studying two-component systems have relied on utilizing biochemical and genetic methods viz. protein characterization and gene inactivation. These methods have been very successful in elucidating the role of various TCSs in bacterial physiology including adaptation to atypical conditions such as pathogenesis. Though gene inactivation studies have facilitated better understanding of role of TCS proteins in bacterial virulence, the dynamics of activity of these proteins *in vivo* is still poorly understood. Identity and nature of extracellular signals which activate two-component systems is another area which warrants investigation. Overall the following lacunae exist in our current knowledge about mycobacterial two-component systems.

1. Evidences to show that phosphorylation of sensor kinase and response regulator proteins happen *in vivo* are very limited [111]. Due to the labile nature of the phosphorylated histidine and aspartate residues, isolation of native phosphorylated proteins has not been easy [112, 113]. Alternate approaches to study these events are required, for e.g., chemotactic two-component system of *E. coli*, CheY-CheZ has been studied using *in vivo* imaging methods [114–116]. Similar approaches are needed for understanding TCSs of *M. tuberculosis*. Besides this, evidences for *in vivo* dimerization, interaction of sensor kinase and response regulator protein are also not available.
2. The extracellular cues which are detected by different sensor kinases to activate various response regulator proteins are still unknown. The lack of this information prevents us from understanding the true biological role of any two-component system [36].
3. *In vivo* interaction of TCS proteins with STPKs is also not very well characterized. The only known evidence is that of PknH regulating DevR protein activity [65]. Similar modulation is feasible for other signalling proteins, including sensor kinases, but no information is currently available.
4. Information about dynamic changes in the expression levels of various signalling proteins across different strains of *M. tuberculosis* is not available. It has been observed that in highly virulent Beijing strain, DevRS system is constitutively active [67]. Similar instances for other TCSs are not known and require further investigation.
5. No information about relationship of epidemiology of various strains and the dynamics of signalling pathways in them is currently available.

To summarize, the field of signalling in mycobacteria is still in its infancy and more ground needs to be covered before an understanding of mycobacterial pathogenesis and the role of two-component signalling systems in it can be completely understood.

**Acknowledgements** Supported by financial aid from IISc (Indian Institute of Science); CSIR (Council for Scientific and Industrial Research, New Delhi) and Department of Biotechnology (DBT) to D.K.S. laboratory and a Junior Research Fellowship to R.A. from CSIR.

## References

1. Harries AD (2008) Robert Koch and the discovery of the tubercle bacillus: the challenge of HIV and tuberculosis 125 years later. *Int J Tuberc Lung Dis* 12(3):241–249
2. Barker LP et al (1997) Differential trafficking of live and dead *Mycobacterium marinum* organisms in macrophages. *Infect Immun* 65(4):1497–1504
3. Russell DG et al (2009) Foamy macrophages and the progression of the human tuberculosis granuloma. *Nat Immunol* 10(9):943–948
4. Mori T et al (2004) Specific detection of tuberculosis infection: an interferon-gamma-based assay using new antigens. *Am J Respir Crit Care Med* 170(1):59–64
5. McEvoy CR et al (2007) The role of IS6110 in the evolution of *Mycobacterium tuberculosis*. *Tuberculosis (Edinb)* 87(5):393–404
6. Coros A, DeConno E, Derbyshire KM (2008) IS6110, a *Mycobacterium tuberculosis* complex-specific insertion sequence, is also present in the genome of *Mycobacterium smegmatis*, suggestive of lateral gene transfer among mycobacterial species. *J Bacteriol* 190(9):3408–3410
7. Haldar S et al (2007) Simplified detection of *Mycobacterium tuberculosis* in sputum using smear microscopy and PCR with molecular beacons. *J Med Microbiol* 56(Pt 10):1356–1362
8. Causse M et al (2011) Comparison of two molecular methods for rapid diagnosis of extrapulmonary tuberculosis. *J Clin Microbiol* 49(8):3065–3067
9. Miller MB et al (2011) Performance of Xpert MTB/RIF RUO assay and IS6110 real-time PCR for *Mycobacterium tuberculosis* detection in clinical samples. *J Clin Microbiol* 49 (10):3458–3462
10. Wayne LG (1994) Dormancy of *Mycobacterium tuberculosis* and latency of disease. *Eur J Clin Microbiol Infect Dis* 13(11):908–914
11. Wayne LG, Lin KY (1982) Glyoxylate metabolism and adaptation of *Mycobacterium tuberculosis* to survival under anaerobic conditions. *Infect Immun* 37(3):1042–1049
12. Cardona PJ, Ruiz-Manzano J (2004) On the nature of *Mycobacterium tuberculosis*-latent bacilli. *Eur Respir J* 24(6):1044–1051
13. Cardona PJ (2009) A dynamic reinfection hypothesis of latent tuberculosis infection. *Infection* 37(2):80–86
14. Yang Z et al (2011) How dormant is *Mycobacterium tuberculosis* during latency? A study integrating genomics and molecular epidemiology. *Infect Genet Evol* 11(5):1164–1167
15. Wayne LG, Sohaskey CD (2001) Nonreplicating persistence of mycobacterium tuberculosis. *Annu Rev Microbiol* 55:139–163
16. Sherman DR et al (2001) Regulation of the *Mycobacterium tuberculosis* hypoxic response gene encoding alpha-crystallin. *Proc Natl Acad Sci U S A* 98(13):7534–7539
17. Kendall SL et al (2004) The *Mycobacterium tuberculosis* dosRS two-component system is induced by multiple stresses. *Tuberculosis (Edinb)* 84(3–4):247–255

18. Betts JC et al (2002) Evaluation of a nutrient starvation model of *Mycobacterium tuberculosis* persistence by gene and protein expression profiling. *Mol Microbiol* 43(3):717–731
19. Voskuil MI et al (2003) Inhibition of respiration by nitric oxide induces a *Mycobacterium tuberculosis* dormancy program. *J Exp Med* 198(5):705–713
20. Taneja NK et al (2010) *Mycobacterium tuberculosis* transcriptional adaptation, growth arrest and dormancy phenotype development is triggered by vitamin C. *PLoS One* 5(5):e10860
21. Cole ST et al (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* 393(6685):537–544
22. Stock AM, Robinson VL, Goudreau PN (2000) Two-component signal transduction. *Annu Rev Biochem* 69:183–215
23. Mizuno T (1997) Compilation of all genes encoding two-component phosphotransfer signal transducers in the genome of *Escherichia coli*. *DNA Res* 4(2):161–168
24. Ogura M, Tanaka T (2002) Recent progress in *Bacillus subtilis* two-component regulation. *Front Biosci* 7:d1815–d1824
25. Chao J et al (2010) Protein kinase and phosphatase signaling in *Mycobacterium tuberculosis* physiology and pathogenesis. *Biochim Biophys Acta* 1804(3):620–627
26. Alber T (2009) Signaling mechanisms of the *Mycobacterium tuberculosis* receptor Ser/Thr protein kinases. *Curr Opin Struct Biol* 19(6):650–657
27. Brosch R et al (2000) Comparative genomics of the mycobacteria. *Int J Med Microbiol* 290 (2):143–152
28. Stinear TP et al (2008) Insights from the complete genome sequence of *Mycobacterium marinum* on the evolution of *Mycobacterium tuberculosis*. *Genome Res* 18(5):729–741
29. Wynne JW et al (2010) Resequencing the *Mycobacterium avium* subsp. *paratuberculosis* K10 genome: improved annotation and revised genome sequence. *J Bacteriol* 192(23):6319–6320
30. Gordon SV et al (2001) Genomics of *Mycobacterium bovis*. *Tuberculosis (Edinb)* 81 (1–2):157–163
31. Garnier T et al (2003) The complete genome sequence of *Mycobacterium bovis*. *Proc Natl Acad Sci U S A* 100(13):7877–7882
32. Li L et al (2005) The complete genome sequence of *Mycobacterium avium* subspecies *paratuberculosis*. *Proc Natl Acad Sci U S A* 102(35):12344–12349
33. Zheng H et al (2008) Genetic basis of virulence attenuation revealed by comparative genomic analysis of *Mycobacterium tuberculosis* strain H37Ra versus H37Rv. *PLoS One* 3(6):e2375
34. Goudreau PN, Stock AM (1998) Signal transduction in bacteria: molecular mechanisms of stimulus-response coupling. *Curr Opin Microbiol* 1(2):160–169
35. West AH, Stock AM (2001) Histidine kinases and response regulator proteins in two-component signaling systems. *Trends Biochem Sci* 26(6):369–376
36. Mascher T, Helmann JD, Unden G (2006) Stimulus perception in bacterial signal-transducing histidine kinases. *Microbiol Mol Biol Rev* 70(4):910–938
37. Parish T et al (2003) Deletion of two-component regulatory systems increases the virulence of *Mycobacterium tuberculosis*. *Infect Immun* 71(3):1134–1140
38. Zahrt TC, Deretic V (2001) *Mycobacterium tuberculosis* signal transduction system required for persistent infections. *Proc Natl Acad Sci U S A* 98(22):12706–12711
39. Zahrt TC, Deretic V (2000) An essential two-component signal transduction system in *Mycobacterium tuberculosis*. *J Bacteriol* 182(13):3832–3838
40. Haydel SE et al (2012) The *prrAB* two-component system is essential for *Mycobacterium tuberculosis* viability and is induced under nitrogen-limiting conditions. *J Bacteriol* 194 (2):354–361
41. Fol M et al (2006) Modulation of *Mycobacterium tuberculosis* proliferation by MtrA, an essential two-component response regulator. *Mol Microbiol* 60(3):643–657
42. Rodriguez GM et al (2002) *ideR*, an essential gene in *mycobacterium tuberculosis*: role of *ideR* in iron-dependent gene expression, iron metabolism, and oxidative stress response. *Infect Immun* 70(7):3371–3381

43. Li Y et al (2010) The characterization of conserved binding motifs and potential target genes for *M. tuberculosis* MtrAB reveals a link between the two-component system and the drug resistance of *M. smegmatis*. *BMC Microbiol* 10:242
44. Ewann F et al (2002) Transient requirement of the PrrA-PrrB two-component system for early intracellular multiplication of *Mycobacterium tuberculosis*. *Infect Immun* 70 (5):2256–2263
45. Graham JE, Clark-Curtiss JE (1999) Identification of *Mycobacterium tuberculosis* RNAs synthesized in response to phagocytosis by human macrophages by selective capture of transcribed sequences (SCOTS). *Proc Natl Acad Sci U S A* 96(20):11554–11559
46. Ewann F, Locht C, Supply P (2004) Intracellular autoregulation of the *Mycobacterium tuberculosis* PrrA response regulator. *Microbiology* 150(Pt 1):241–246
47. Nowak E et al (2006) Structural and functional aspects of the sensor histidine kinase PrrB from *Mycobacterium tuberculosis*. *Structure* 14(2):275–285
48. Dasgupta N et al (2000) Characterization of a two-component system, devR-devS, of *Mycobacterium tuberculosis*. *Tuber Lung Dis* 80(3):141–159
49. Malhotra V et al (2004) Disruption of response regulator gene, devR, leads to attenuation in virulence of *Mycobacterium tuberculosis*. *FEMS Microbiol Lett* 231(2):237–245
50. Park HD et al (2003) Rv3133c/dosR is a transcription factor that mediates the hypoxic response of *Mycobacterium tuberculosis*. *Mol Microbiol* 48(3):833–843
51. Saini DK, Malhotra V, Tyagi JS (2004) Cross talk between DevS sensor kinase homologue, Rv2027c, and DevR response regulator of *Mycobacterium tuberculosis*. *FEBS Lett* 565 (1–3):75–80
52. Saini DK et al (2004) DevR-DevS is a bona fide two-component system of *Mycobacterium tuberculosis* that is hypoxia-responsive in the absence of the DNA-binding domain of DevR. *Microbiology* 150(Pt 4):865–875
53. Wisedchaisri G et al (2005) Structures of *Mycobacterium tuberculosis* DosR and DosR-DNA complex involved in gene activation during adaptation to hypoxic latency. *J Mol Biol* 354 (3):630–641
54. Roberts DM et al (2004) Two sensor kinases contribute to the hypoxic response of *Mycobacterium tuberculosis*. *J Biol Chem* 279(22):23082–23087
55. Sousa EH et al (2007) DosT and DevS are oxygen-switched kinases in *Mycobacterium tuberculosis*. *Protein Sci* 16(8):1708–1719
56. Cho HY et al (2009) Structural insight into the heme-based redox sensing by DosS from *Mycobacterium tuberculosis*. *J Biol Chem* 284(19):13057–13067
57. Lee JM et al (2008) O2- and NO-sensing mechanism through the DevSR two-component system in *Mycobacterium smegmatis*. *J Bacteriol* 190(20):6795–6804
58. Ioanoviciu A et al (2009) DevS oxy complex stability identifies this heme protein as a gas sensor in *Mycobacterium tuberculosis* dormancy. *Biochemistry* 48(25):5839–5848
59. Honaker RW et al (2009) Unique roles of DosT and DosS in DosR regulon induction and *Mycobacterium tuberculosis* dormancy. *Infect Immun* 77(8):3258–3263
60. Converse PJ et al (2009) Role of the dosR-dosS two-component regulatory system in *Mycobacterium tuberculosis* virulence in three animal models. *Infect Immun* 77 (3):1230–1237
61. Wisedchaisri G et al (2008) Crystal structures of the response regulator DosR from *Mycobacterium tuberculosis* suggest a helix rearrangement mechanism for phosphorylation activation. *J Mol Biol* 378(1):227–242
62. Gupta RK et al (2009) Structure-based design of DevR inhibitor active against nonreplicating *Mycobacterium tuberculosis*. *J Med Chem* 52(20):6324–6334
63. Saini DK, Tyagi JS (2005) High-throughput microplate phosphorylation assays based on DevR-DevS/Rv2027c 2-component signal transduction pathway to screen for novel antitubercular compounds. *J Biomol Screen* 10(3):215–224
64. Gonzalo-Asensio J et al (2008) PhoP: a missing piece in the intricate puzzle of *Mycobacterium tuberculosis* virulence. *PLoS One* 3(10):e3496

65. Chao JD et al (2010) Convergence of Ser/Thr and two-component signaling to coordinate expression of the dormancy regulon in *Mycobacterium tuberculosis*. *J Biol Chem* 285 (38):29239–29246
66. Pang X et al (2007) Evidence for complex interactions of stress-associated regulons in an *mprAB* deletion mutant of *Mycobacterium tuberculosis*. *Microbiology* 153(Pt 4):1229–1242
67. Fallow A, Domenech P, Reed MB (2010) Strains of the East Asian (W/Beijing) lineage of *Mycobacterium tuberculosis* are *DosS/DosT-DosR* two-component regulatory system natural mutants. *J Bacteriol* 192(8):2228–2238
68. Walters SB et al (2006) The *Mycobacterium tuberculosis* *PhoPR* two-component system regulates genes essential for virulence and complex lipid biosynthesis. *Mol Microbiol* 60 (2):312–330
69. Perez E et al (2001) An essential role for *phoP* in *Mycobacterium tuberculosis* virulence. *Mol Microbiol* 41(1):179–187
70. Lee JS et al (2008) Mutation in the transcriptional regulator *PhoP* contributes to avirulence of *Mycobacterium tuberculosis* H37Ra strain. *Cell Host Microbe* 3(2):97–103
71. Frigui W et al (2008) Control of *M. tuberculosis* ESAT-6 secretion and specific T cell recognition by *PhoP*. *PLoS Pathog* 4(2):e33
72. Ryndak M, Wang S, Smith I (2008) *PhoP*, a key player in *Mycobacterium tuberculosis* virulence. *Trends Microbiol* 16(11):528–534
73. Gonzalo Asensio J et al (2006) The virulence-associated two-component *PhoP-PhoR* system controls the biosynthesis of polyketide-derived lipids in *Mycobacterium tuberculosis*. *J Biol Chem* 281(3):1313–1316
74. Talaat AM et al (2004) The temporal expression profile of *Mycobacterium tuberculosis* infection in mice. *Proc Natl Acad Sci U S A* 101(13):4602–4607
75. He H et al (2006) *MprAB* is a stress-responsive two-component system that directly regulates expression of sigma factors *SigB* and *SigE* in *Mycobacterium tuberculosis*. *J Bacteriol* 188 (6):2134–2143
76. Zahrt TC et al (2003) Functional analysis of the *Mycobacterium tuberculosis* *MprAB* two-component signal transduction system. *Infect Immun* 71(12):6962–6970
77. He H, Zahrt TC (2005) Identification and characterization of a regulatory sequence recognized by *Mycobacterium tuberculosis* persistence regulator *MprA*. *J Bacteriol* 187 (1):202–212
78. Himpens S, Locht C, Supply P (2000) Molecular characterization of the mycobacterial *SenX3-RegX3* two-component system: evidence for autoregulation. *Microbiology* 146 (Pt 12):3091–3098
79. Parish T et al (2003) The *senX3-regX3* two-component regulatory system of *Mycobacterium tuberculosis* is required for virulence. *Microbiology* 149(Pt 6):1423–1435
80. Rickman L et al (2004) A two-component signal transduction system with a PAS domain-containing sensor is required for virulence of *Mycobacterium tuberculosis* in mice. *Biochem Biophys Res Commun* 314(1):259–267
81. Roberts G et al (2011) Control of *CydB* and *GltA1* expression by the *SenX3 RegX3* two component regulatory system of *Mycobacterium tuberculosis*. *PLoS One* 6(6):e21090
82. Singh A et al (2006) Dissecting virulence pathways of *Mycobacterium tuberculosis* through protein–protein association. *Proc Natl Acad Sci U S A* 103(30):11346–11351
83. Bhattacharya M, Biswas A, Das AK (2010) Interaction analysis of *TcrX/Y* two component system from *Mycobacterium tuberculosis*. *Biochimie* 92(3):263–272
84. Bacon J et al (2007) Lipid composition and transcriptional response of *Mycobacterium tuberculosis* grown under iron-limitation in continuous culture: identification of a novel wax ester. *Microbiology* 153(Pt 5):1435–1444
85. Haydel SE, Clark-Curtiss JE (2004) Global expression analysis of two-component system regulator genes during *Mycobacterium tuberculosis* growth in human macrophages. *FEMS Microbiol Lett* 236(2):341–347

86. Haydel SE, Dunlap NE, Benjamin WH Jr (1999) In vitro evidence of two-component system phosphorylation between the *Mycobacterium tuberculosis* TrcR/TrcS proteins. *Microb Pathog* 26(4):195–206
87. Haydel SE, Clark-Curtiss JE (2006) The *Mycobacterium tuberculosis* TrcR response regulator represses transcription of the intracellularly expressed Rv1057 gene, encoding a seven-bladed beta-propeller. *J Bacteriol* 188(1):150–159
88. Haydel SE et al (2002) Expression, autoregulation, and DNA binding properties of the *Mycobacterium tuberculosis* TrcR response regulator. *J Bacteriol* 184(8):2192–2203
89. Shrivastava R, Ghosh AK, Das AK (2009) Intra- and intermolecular domain interactions among novel two-component system proteins coded by Rv0600c, Rv0601c and Rv0602c of *Mycobacterium tuberculosis*. *Microbiology* 155(Pt 3):772–779
90. Shrivastava R, Ghosh AK, Das AK (2007) Probing the nucleotide binding and phosphorylation by the histidine kinase of a novel three-protein two-component system from *Mycobacterium tuberculosis*. *FEBS Lett* 581(9):1903–1909
91. Shrivastava R et al (2006) Functional insights from the molecular modelling of a novel two-component system. *Biochem Biophys Res Commun* 344(4):1327–1333
92. Schnell R, Agren D, Schneider G (2008) 1.9 Å structure of the signal receiver domain of the putative response regulator NarL from *Mycobacterium tuberculosis*. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 64(Pt 12):1096–1100
93. Morth JP et al (2005) A novel two-component system found in *Mycobacterium tuberculosis*. *FEBS Lett* 579(19):4145–4148
94. Morth JP et al (2004) The crystal and solution structure of a putative transcriptional antiterminator from *Mycobacterium tuberculosis*. *Structure* 12(9):1595–1605
95. Laub MT, Biondi EG, Skerker JM (2007) Phosphotransfer profiling: systematic mapping of two-component signal transduction pathways and phosphorelays. *Methods Enzymol* 423:531–548
96. Biondi EG et al (2006) A phosphorelay system controls stalk biogenesis during cell cycle progression in *Caulobacter crescentus*. *Mol Microbiol* 59(2):386–401
97. Yamamoto K et al (2005) Functional characterization in vitro of all two-component signal transduction systems from *Escherichia coli*. *J Biol Chem* 280(2):1448–1456
98. Skerker JM et al (2005) Two-component signal transduction pathways regulating growth and cell cycle progression in a bacterium: a system-level analysis. *PLoS Biol* 3(10):e334
99. Gillespie SH, Billington O (1999) Activity of moxifloxacin against mycobacteria. *J Antimicrob Chemother* 44(3):393–395
100. Fung-Tomic J et al (2000) In vitro antibacterial spectrum of a new broad-spectrum 8-methoxy fluoroquinolone, gatifloxacin. *J Antimicrob Chemother* 45(4):437–446
101. Stover CK et al (2000) A small-molecule nitroimidazopyran drug candidate for the treatment of tuberculosis. *Nature* 405(6789):962–966
102. Andries K et al (2005) A diarylquinoline drug active on the ATP synthase of *Mycobacterium tuberculosis*. *Science* 307(5707):223–227
103. Jia L et al (2005) Simultaneous estimation of pharmacokinetic properties in mice of three anti-tubercular ethambutol analogs obtained from combinatorial lead optimization. *J Pharm Biomed Anal* 37(4):793–799
104. Aguilar D et al (2007) Immunological responses and protective immunity against tuberculosis conferred by vaccination of Balb/C mice with the attenuated *Mycobacterium tuberculosis* (phoP) SO2 strain. *Clin Exp Immunol* 147(2):330–338
105. Gardner CA, Acharya T, Pablos-Mendez A (2005) The global alliance for tuberculosis drug development—accomplishments and future directions. *Clin Chest Med* 26(2):341–347, vii
106. Isaza JP et al (2012) Whole genome shotgun sequencing of one Colombian clinical isolate of *Mycobacterium tuberculosis* reveals DosR regulon gene deletions. *FEMS Microbiol Lett* 330 (2):113–120

107. Reed MB et al (2007) The W-Beijing lineage of *Mycobacterium tuberculosis* overproduces triglycerides and has the DosR dormancy regulon constitutively upregulated. *J Bacteriol* 189(7):2583–2589
108. Sureka K et al (2008) Positive feedback and noise activate the stringent response regulator rel in mycobacteria. *PLoS One* 3(3):e1771
109. Sureka K et al (2007) Polyphosphate kinase is involved in stress-induced mprAB-sigE-rel signalling in mycobacteria. *Mol Microbiol* 65(2):261–276
110. Tiwari A et al (2010) The interplay of multiple feedback loops with post-translational kinetics results in bistability of mycobacterial stress response. *Phys Biol* 7(3):036005
111. Forst S et al (1990) In vivo phosphorylation of OmpR, the transcription activator of the *ompF* and *ompC* genes in *Escherichia coli*. *J Bacteriol* 172(6):3473–3477
112. Hohenester UM et al (2010) Stepchild phosphohistidine: acid-labile phosphorylation becomes accessible by functional proteomics. *Anal Bioanal Chem* 397(8):3209–3212
113. Klumpp S, Kriegstein J (2002) Phosphorylation and dephosphorylation of histidine residues in proteins. *Eur J Biochem* 269(4):1067–1071
114. Kentner D, Sourjik V (2006) Spatial organization of the bacterial chemotaxis system. *Curr Opin Microbiol* 9(6):619–624
115. Kentner D, Sourjik V (2009) Dynamic map of protein interactions in the *Escherichia coli* chemotaxis pathway. *Mol Syst Biol* 5:238
116. Kentner D, Sourjik V (2010) Use of fluorescence microscopy to study intracellular signaling in bacteria. *Annu Rev Microbiol* 64:373–390

## Chapter 5

# *Mycobacterium tuberculosis*: Evolution, Host–Pathogen Interactions, and Implications for Tuberculosis Control

Marcos Burgos

### 5.1 Introduction

*Mycobacterium tuberculosis* continues to be a major cause of morbidity and mortality in the world. One-third of the world population is estimated to be infected with tuberculosis resulting in 8.8 million new cases and 1.4 million deaths associated to tuberculosis in 2011 (WHO-fact sheet-2011). The emergence of multidrug-resistant (MDR) and other forms of extensive-drug-resistant (XDR) tuberculosis in many parts of the world is threatening to send us back to an era when tuberculosis was an untreatable disease. Despite the fast pace of progress in the field of tuberculosis research in the last decade, the implementation of promising discoveries continues to be a challenge. In most countries, National Tuberculosis Programs are still using smear microscopy, a 120-year-old technology, to diagnose tuberculosis and treatment regimens that have not changed in the past 40 years. Furthermore, we still have important gaps in our knowledge of the basic biology of *M. tuberculosis*, as it pertains to pathogen and human interactions. We still do not know the extent of genetic diversity in tuberculosis bacteria, nor do we understand the implications of this diversity in terms of virulence, vaccine, and drug development. Nevertheless, there is fascinating new research into the global diversity of tuberculosis strains and its association to ancient human migrations out of Africa, as well as to more recent migration patterns of humans in the last 500 years. These most recent insights into the diversity of tuberculosis and human coevolution are promising a more far-reaching understanding of the biology of tuberculosis with potential payoffs for the eventual elimination of tuberculosis.

---

M. Burgos (✉)

Division of Infectious Diseases, Department of Medicine, New Mexico  
Veterans Healthcare System, Albuquerque, NM, USA

Division of Infectious Diseases, University of New Mexico Health Science Center,  
Albuquerque, NM, USA  
e-mail: [mburgos@salud.unm.edu](mailto:mburgos@salud.unm.edu)

## 5.2 Pathogenesis of *M. tuberculosis*

The pathogenesis of *M. tuberculosis* involves a three-step basic process: transmission of bacteria, establishment of infection, and progression to disease. *M. tuberculosis* is transmitted through aerosols, whereby host cough-airborne particles of 1–5 µm in diameter penetrate deep into the lung reaching the alveolar sacs where the initial infection is established. Once acquired by the pulmonary route, *M. tuberculosis* is engulfed by alveolar macrophages where the pathogen (TB) manipulates its host cell to enable its survival and hijacks the normal house-keeping functions of the cell to control the surrounding tissues [1, 2]. Once inside the cell, the intracellular bacillary load increases, T cells are recruited to the site of infection, where they attract and activate more macrophages in an attempt to eradicate the pathogen and, in turn, these arriving cells also get infected. This stimulation of the immune response leads to the generation of a pulmonary parenchymal focus or a granulomatous lesion that walls off the bacterium from further immune detection [2, 3]. The granuloma has a central caseous necrotic core that contains by-products of mycobacteria and of dead host cells. In this environment, it is speculated, host cells and mycobacteria interact by a pathogen-mediated process [1]. The hallmark of TB pathology is the intracellular immune subversion and control of the external environment by *M. tuberculosis* that produces the granuloma and whereby infection is established [1, 2, 4]. Ninety percent of infected individuals can successfully wall off the bacilli in these granulomas, leading to latent infection without clinical signs of disease. It is estimated that in 5 % of individuals, TB infection will progress from latent to active disease within 2 years after the initial infection, and in an additional 5 % of individuals active disease will develop at some later point in their lives [5]. In most of the infected cases, TB disease presents as a pulmonary form, with classic radiographic findings of upper lobe infiltrates and/or cavitation. Aerolisation of tubercle bacilli from smear positive patients and subsequent transmission to other susceptible host represents an essential step in the infection cycle of *M. tuberculosis*. Pulmonary involvement is the primary site of disease in 80–85 % of cases and other extrapulmonary sites of disease involve 15–20 % of the cases [5]. Extrapulmonary disease from an evolutionary perspective is a less virulent form of disease, since it cannot be aerosolized, and it is a dead end for the transmission cycle of *M. tuberculosis*. The rate of progression to active disease is dramatically increased among persons who are coinfected with HIV. Extrapulmonary TB disease is more common in the setting of HIV infection. Commonly reported extrapulmonary sites of disease are the lymph nodes, pleura, bones, genitourinary system, and the central nervous system; although virtually any other organ system can be involved with *M. tuberculosis* disease.

The distinctive feature of *M. tuberculosis* is its persistence in the human host requiring inter-host transmission [4]. This persistence is often expressed as virulence or fitness, the capacity of the mycobacterium to spread, survive, and reproduce inside macrophages and cause significant pulmonary disease to allow transmission. Once in the host, the pathogen's priority is to modulate immune

elimination, permitting it to enter a latent state for years. For example, recently, Comas and colleagues [6] showed that human T cell epitopes of MTBC are highly conserved. Therefore, it seems immune recognition works for the benefit of establishing infection. However, this ability to remain dormant in tissues without progression to disease can be modified by host factors such as HIV infection, malignancy, hemodialysis, malnutrition, diabetes, and the use of immunosuppressive drugs [3]. Some aspects of the life cycle and virulence of *M. tuberculosis* can be estimated with the use of molecular epidemiological studies by measuring the number of secondary infections caused by an index case [7], the propensity to cause cavitary disease and the ability to disseminate or to cause both pulmonary and extrapulmonary disease [8, 9]. The use of in vivo and in vitro models can also help quantify the virulence of *M. tuberculosis*. For example in mouse model studies, virulence is measured as the time of growth in inoculum size required to kill the host and by quantifying the number of colony forming units (CFUs) in the lung homogenates and by measuring the type of host's immune defense mechanism [10, 11]. Many efforts have been done in trying to evaluate pathogenicity differences between isolates of *M. tuberculosis*. However, interpreting virulence studies is challenging since these investigations often evaluate one of many aspects of the life cycle of tuberculosis.

### 5.3 *M. tuberculosis* Complex

The *M. tuberculosis* complex (MTBC) concept dates back to the identification of two closely related but distinct species of slowly growing mycobacteria: *M. tuberculosis* and *Mycobacterium bovis* [12–14]. Today, the *M. tuberculosis* complex forms a group of seven species of slowly growing mycobacteria. These species are noteworthy for their lack of genetic variation, 99.9 % similarity at the nucleotide level and identical 16 rRNA sequence [12, 13, 15]. Furthermore, MTBC members have a clonal population structure and have little evidence of genetic recombination between members of the MTBC [13, 14, 16]. A clonal population structure facilitates phylogenetic analysis of MTBC based on variation at the genomic level. This genetic variation is exploited by finding single nucleotide polymorphisms (SNPs), deletions or insertions of small or large pieces of DNA and by identifying the recombination or replication of variable repeats in the genome [16, 17].

Individual members of MTBC can be distinguished by their pathogenesis, host tropism, and disease presentation [18, 19]. *M. tuberculosis* and *Mycobacterium africanum* are uniquely obligate human pathogens with no other known animal or environmental reservoirs and these pathogens propagate via human to human transmission by establishing infection mostly in the lungs. On the other hand *M. bovis* affects humans and cattle. *M. bovis* disease is still observed in mostly rural areas of developing countries where people consume raw milk, the most common vehicle of transmission. In addition, *M. bovis* is not usually transmitted

between humans since this illness primarily manifests as extra-pulmonary disease. Other members of MTBC include *Mycobacterium caprae*, *Mycobacterium microti*, *Mycobacterium pinnipedii*, and *Mycobacterium canettii* [2, 13]. *M. caprae* affects goats and sheep, *M. microti* affects voles, and *M. pinnipedii* affects seals and sea lions [13]. In the case of *M. canettii*, which was first described in the 1960s, it has been isolated from fewer than 60 patients from the horn of Africa [20]. Inexplicably, this organism has no evidence of human-to-human transmission and there is as of yet no evidence of an animal or environmental reservoir [21]. Occasionally members of MTBC have been isolated in other mammals, but the disease does not appear to be maintained and transmitted. Smith and colleagues [13] proposed to refer to these host-adapted forms of mycobacteria as ecotypes of *M. tuberculosis* rather than as different species because of their close genetic identity with similar pathological presentations. Newly recognized additions to the MTBC include the *Dassie bacillus* [22] and the recently identified *Mycobacterium mungi* [23], which are the closest relatives of *M. africanum* within the *M. tuberculosis* complex. The *Dassie bacillus* has been isolated from Dassies, or Rock Hyrax, in South Africa [22], and *M. mungi* causes disease in troops of banded mongoose in northern Botswana [23, 24].

Pathogens that exhibit substantially low DNA sequence diversity are referred to as genetically monomorphic [14]. MTBC members are a prominent example of such pathogens. Until recently, genetically monomorphic pathogens such as *M. tuberculosis* have been difficult to analyze by standard genotyping methods because genetic diversity is limited in this pathogen [14, 25]. More recently, with the introduction of PCR-based genotyping methods based on Clustered Regulatory Short Palindromic Repeats (CRISPR) and Variable Number Tandem Repeats (VNTR), analyzing genetically monomorphic bacteria has become less limited [17, 26, 27]. These techniques have become the gold standard for routine genotyping of *M. tuberculosis* [14, 28] and they have been applied very successfully to address a variety of epidemiological and transmission dynamics questions [29]. However, Comas and colleagues [14] found that quantitative analysis of spoligotyping and VNTR-based genotyping methods revealed significant amounts of homoplasy and substantial propensity for convergent evolution. Consequently, using these methods to define deep phylogenetic groupings in MTBC could result in misleading discoveries [14, 30]. Here, we will review some of the most important genotyping methods for epidemiological and phylogenetic studies in *M. tuberculosis*.

## 5.4 Past and Current Tools of Molecular Epidemiology

Before the introduction of molecular typing methods, it was not possible to distinguish between individual strains of *M. tuberculosis* and drug susceptibility patterns were often used as a proxy for relatedness of the isolates. It was after the introductions and application of molecular typing methods, such as restriction

fragment length polymorphism (RFLP) analysis based on the insertion sequence IS6110, in the 1990s that new temporal and spatial dimension was brought to the study of transmission of tuberculosis. A new field of research called molecular epidemiology emerged with the use of genotyping markers and for the first time we developed an understanding of the transmission dynamics of tuberculosis that classical epidemiology by itself could not provide [29, 31–33]. In the 1990s IS6110 RFLP was the most extensively studied and widely used typing method for distinguishing among MTBC isolates. The number of copies of IS6110 and their positions in the *M. tuberculosis* chromosome were sufficient to generate a genotype or fingerprint to distinguish between strains.

Discrimination between isolates of *M. tuberculosis* is of utmost importance in epidemiological and evolutionary studies. Genotyping tools are now used routinely in most developed countries by Tuberculosis Control Programs in their efforts to do contact investigations, to study relapse and re-infection and to study transmission dynamics of tuberculosis in a community [29, 32, 34]. In general with the use of classical epidemiology and genotyping tools we can do molecular epidemiology studies to help interpret the possible spatial and temporal mode of the spread of tuberculosis, as well as risk factors associated with chains of transmission. In addition, genotyping of *M. tuberculosis* isolates can be used to aid contact investigations and support case-finding [30]. Furthermore, these tools can help detect false positive cultures or laboratory cross contamination which can prevent unnecessary treatment of individuals [31]. Genotyping studies distinguish strain-to-strain differences in terms of the geographic distribution and incidence of cases in a community. Studies of large international collections of MTBC strains identified the first genotype families or lineages among MTBC [35, 36]. These new insights allow us to study important aspects of epidemics, pathogenicity, and phylogeographic distribution of strains. In addition, we have started to study the microbial-host evolution of tuberculosis [15, 37, 38]. Below, we will review the most important genotyping tools for molecular epidemiology and phylogeny studies.

### 5.4.1 RFLP

Restriction fragment length polymorphism typing (RFLP) is a technique in which bacterial DNA is cleaved by a restriction enzyme, and the DNA restriction fragments are separated by gel electrophoresis. Thereafter, the Southern blot technique is used to transfer the separated DNA restriction fragments to a membrane. A labeled DNA probe is applied to the membrane in a hybridization assay, complementary to repetitive sequences of interest [31]. One often-applied probe contains an IS6110-specific sequence. IS6110 is a repetitive, mobile insertion sequence element of 1.35 kb. The IS6110 element can duplicate and integrate into new genomic sites, however, not always in a random manner [39–41]. The number of IS6110 copies varies from zero to about 25 per strain and with the variation in insertion sites IS6110 typing can yield thousands and thousands of different banding patterns. From the

early 1990s on, RFLP with IS6110 was the gold standard in *M. tuberculosis* typing. The application of this method allowed international studies and has brought significant new insights in the epidemiology of TB [34, 36, 41–45]. However, IS6110 RFLP suffers from problems in reproducibility and it requires specialized software to analyze and compare results. In addition, it requires users to grow the mycobacteria in order to obtain abundant quantities of DNA to perform this method [31, 45].

### 5.4.2 Spoligotyping

The use of DNA RFLP analysis to differentiate between strains of *M. tuberculosis* is slowed down by the need to culture this organism. On the other hand, spoligotyping, is a polymerase chain reaction (PCR)-based method, which requires very small amounts of DNA, which can be obtained without resorting to bacterial culture. Spoligotyping interrogates a Direct Repeat sequence comprising a repetitive 36-base-pair element separated by short nonrepetitive sequences [41, 46]. Atypical strains of mycobacteria have been analyzed using spoligotyping and do not give a signal, indicating the specificity of this technique for *M. tuberculosis* [47].

Spoligotyping has many advantages. The ability to perform this typing method directly on sputum samples makes it very useful in clinical situations [46, 48]. This technique has been demonstrated to be helpful when discriminating between isolates of *M. tuberculosis* with few IS6110 bands [47]. Another additional advantage of this secondary marker technique is that it is economical, easy to perform, and a rapid means of typing the *M. tuberculosis* complex [31]. While DNA RFLP typing is time-consuming, it is fast and it is used in many mycobacterium laboratories as an initial screening step before applying a secondary technique of greater discriminatory power [41, 45]. The wide use of spoligotyping is demonstrated by a 39,609 spoligotype patterns collection in an international database from more than 121 countries [49]. The application of spoligotyping for phylogenetic studies is limited by its considerable propensity for convergent evolution [14]. However, spoligotyping can still be informative for population molecular epidemiology studies and for identification of important *M. tuberculosis* families. For example, a characteristic loss of 34 spacers caused by a deletion of a genomic region known as RD207 allows the identification of the “Beijing” lineage of MTBC [50].

### 5.4.3 VNTR Typing

Mycobacterial interspersed repetitive units (MIRUs) represent a PCR-based typing system that identifies novel polymorphic loci, which are independent of existing techniques such as IS6110 typing [51–54]. They are a specific class of variable number of tandem repeats (VNTR) that have been identified at 41 different loci

in the genome of *M. tuberculosis*. Each comprises strings of short repetitive sequences. The number of repeats at different loci varies between strains. PCR amplification across each MIRU generates fragments of different sizes from different strains. The determination of the tandem repeat sizes can be performed in an agarose gel-based electrophoresis method or on DNA sequencing machines. Moreover, the results are expressed as numerical codes and are therefore very easy to compare and exchange [44, 52]. Among different sets of MIRU-VNTR loci described for typing *M. tuberculosis* isolates, a new international standard typing method system based on 24 loci was proposed in 2006, by a group of international institutions [17]. The discriminatory power seems to be comparable to IS6110 RFLP, while VNTR typing is faster and more economical than RFLP typing [52]. The introduction of this genotyping method replaced IS6110-RFLP as the gold standard and allowed VNTR patterns to be compared in international databases. Because of the ease of use and the improved performance, VNTR typing has been integrated in TB control programs on an international scale in the United States, Europe, and other parts of the world [30, 44]. A global epidemiological database is available that has led to insights into the geographical distribution and evolution of *M. tuberculosis* [26, 44]. However, the usefulness of VNTR for phylogenetic studies of *M. tuberculosis* remains questionable. Comas and colleagues recently showed that although VNTR performs better than spoligotyping it shows only limited congruence to wide genome sequencing phylogenies [14].

#### 5.4.4 Regions of Difference

Large sequence polymorphisms (LSP) or regions of difference (RD) are particularly attractive approaches for studying the phylogeny of *M. tuberculosis* strains as these methods simultaneously provide information about the biological basis of the phenotype [31]. Because these deletions rarely occur independently at exactly the same chromosomal locus, they can be seen as unique and irreversible genetic events. The number and distribution of these deletions provides a genomic signature, which can be used for constructing robust phylogenetic relationships. For example, with the use of these regions of difference we now define the phylogeny and geographic distribution of six important *M. tuberculosis* lineages [55] that will be reviewed in the geographic structure of MTBC section. When these deletions disrupt coding regions of the genome, the loss of specific genes could influence important pathogenic characteristics of the strains. For instance, although these RD could be slightly deleterious to the pathogen, some deletions are likely to offer short-term advantages of escape from the host immune system or confer strong advantages, such as antibiotic resistance, or promote transmission [55]. Even though RDs are very useful in defining phylogenetic groupings, these markers offer insufficient discriminatory power for routine molecular epidemiological investigation because the genetic distance between two groups of strains cannot be determined solely on the basis of absence and presence of RDs [14].

### 5.4.5 SNP and MLSA/MLST Typing

Historically molecular epidemiology studies were limited by the use of genotyping markers that were not able to accurately estimate the overall level of chromosomal relationships among tuberculosis isolates [39, 42, 46]. This problem was to an extent solved by the development of genome wide analysis of synonymous polymorphism (sSNPs) [56]. sSNPs became ideal markers for genotyping of MTBC because they represent unique genetic events and show virtually no convergent evolution or homoplasy [14, 37]. More specifically, sSNPs act as an unambiguous marker due to the low genetic diversity and the absence of horizontal gene transfer within *M. tuberculosis* [14, 27, 38, 57, 58]. Thus, because SNPs are evolutionary neutral they provide a powerful strategy for large-scale analysis of population genetic studies examining phylogenetic relationships among strains of tuberculosis.

Different techniques that allow the detection of SNPs in high-throughput assays have been developed and are used to distinguish between bacterial strains in phylogenetic studies [56, 59]. However, SNPs suffer from some disadvantages, i.e. that the estimation of genetic distances is impossible and linear phylogenies are commonly derived [2, 14, 58]. Moreover, depending on the specificity of the selected SNPs for the tested strains, a low discriminatory power might be achieved due to the possibility of phylogenetic discovery bias. In addition, the determination of SNPs requires either a two-step process (PCR and sequencing) or a specialized system. In summary, while SNPs are well suited for defining deeply rooted phylogenetic relationships, they are of limited value for routine molecular typing.

Because, *M. tuberculosis* harbors little DNA sequence variation a large number of genes need to be sequenced to obtain a reliable phylogeny. Multilocus sequence typing consists of sequencing either SNPs (MLST) or entire gene sequences (MLSA) for comparative analysis. For example, Hershberg and colleagues used MLSA of a large set of coding sequence of 89 genes from a global collection of MTBC strains to reveal that the human-adapted members of MTBC are more genetically diverse than generally recognized at that point [27]. These authors found that MLSA allows investigators to identify true phylogenetic relationships, and to discover SNPs that can be used as powerful genotyping markers. Moreover, Comas and colleagues [14] found that the phylogenies derived from MLST data established true phylogenetic relationships as compared to spoligotyping and MIRU-VNTR data sets which are characterized by significant amounts of homoplasy or convergence evolution. However, although MLSA data is very useful in defining phylogenetic groupings, these markers do not have sufficient discriminatory power for routine molecular epidemiological investigations. In addition this technique could be time-consuming and impractical for use in routine laboratories.

### 5.4.6 Whole Genome Sequencing

The emergence of whole genome sequencing (WGS) as a research tool in the tuberculosis field has allowed unprecedented discoveries of the genetic diversity in *M. tuberculosis*. WGS gives us a snapshot of the entire variations of the bacterial genome. In particular, WGS is useful when applied to pathogens characterized by a low amount of genetic diversity, as monomorphic species such as MTBC. Another advantage of WGS is that phylogenetic relationships are based on a large amount of data which allow one to reconstruct reliable deep-phylogenies even of closely related species such as tuberculosis [6, 15, 58]. Moreover, WGS comparative studies of MTBC strains have allowed the reconstruction of robust phylogenies demonstrating a hidden genetic variation and wide geographic distribution of this pathogen [60]. In addition, Niemann and colleagues [61] used WGS to compare a drug-susceptible and a MDR isolate of a rapidly spreading *M. tuberculosis* Beijing genotype clone from Uzbekistan and found that despite having identical DNA fingerprint patterns these strains had substantial genomic diversity. These findings support the use of WSG for refining the genetic resolution in the context of mixed infections, reinfections, and microevolution processes occurring during transmission because traditional molecular epidemiology methods are limited in capturing genetic variations and some aspects of transmission dynamics can be misinterpreted. Furthermore WSG results have challenged the “identical fingerprint equal same strain” dogma in the interpretation of failures, relapses, and reinfections, which could be critical in the interpretation of clinical trials. For example, misinterpretation of failures, relapses, and reinfections may erroneously suggest poor performance of a drug in a clinical trial and premature withdrawing of the experimental drug from the market with adverse public health implications in tuberculosis [60].

Furthermore the genetic variability resolution provided by WGS has allowed researchers to investigate within host evolution of tuberculosis and challenge long standing assumptions about the pathogenesis of tuberculosis. For example, Saunders and colleagues [62] used WGS on serial sputum isolates to characterize genetic diversity in a single patient infected with a drug-susceptible strain and in whom drug resistance had evolved through stepwise development to multiple drug resistance. Interestingly, the authors only found mutations conferring drug resistance, with no additional mutations found in nonrepetitive regions. These results demonstrate that *M. tuberculosis* is considerably stable genetically within the host, as compared to the greater variability seen in in vitro studies of evolution of drug resistance, or in models of hypermutability [63]. In other words, the mutations seen in this study are more consistent with bacterial models in which resistance is the product of the bacillary load rather than hypermutation models, where exposure to antibiotics is thought to generate a hypermutable state in *M. tuberculosis* [64].

The high resolution of WGS has also been informative in outbreak situations of tuberculosis where genotyping and contact tracing alone did not capture the true dynamics of an outbreak. For example, in Vancouver, British Columbia a recent

outbreak of *M. tuberculosis* was detected and, by standard typing methods, it was defined as a single clonal outbreak with the same MIRU/VNTR patterns and non revealing contact tracing source [65]. To further investigate the heterogeneity of the strains involved, the authors sequenced the genomes of 36 isolates that had an identical MIRU-VNTR pattern. WGS revealed two branches of distinct lineages, indicating two separate clusters of cases. Further contact investigation analysis with additional information from the cluster of cases helped to identify the transmission chain of events. This study further demonstrates the depth of genetic variation provided by the use of WGS in the transmission dynamics of tuberculosis and further outlines the limitations of classical genotyping methods in investigations of the transmission dynamics of tuberculosis.

WGS has been proposed as a gold standard for strain typing in *M. tuberculosis* because of its superior discriminatory power in strain typing and in its ability for discerning deep-phylogenetic associations [14]. Currently WGS cannot be used in routine settings because the richness and the abundant amount of information produced by genome sequencing causes a bottleneck in terms of data storage, analysis, and interpretation of results. Nevertheless, the cost of WGS continues to drop and it has become more feasible to consider WGS as a primary tool for genotyping *M. tuberculosis* isolates. Based on the rapid progress made in genotyping in the last decade, it is likely that in the foreseeable future, WSG can become a one-stop shop “gold standard” tool for *M. tuberculosis* virulence, phylogenetic and molecular epidemiology studies.

## 5.5 A Genotyping Approach for Phylogenetic and Epidemiological Studies

The suitability of a genotyping approach is dependent on the research question and the setting of the study. In molecular epidemiology studies, it is important that the genotyping markers be highly discriminatory. For phylogeny and population genetic analyses, the markers should be phylogenetically robust with low homoplasy and minimal rates of convergent evolution [14]. Deep phylogenetic information might be of little relevance for molecular epidemiology but unambiguous groupings of MTBC strains are essential in research questions that pertain to virulence and the evolution of tuberculosis [14, 58]. For example, interpreting the evolutionary structure and spread of bacterial pathogens requires robust strain assignment [14, 25]. A robust phylogenetic framework of the genetic population structure of MTBC could help discover functional biological features among lineages [66].

Additionally, the ideal DNA genotyping marker for the study of TB transmission dynamics is one that is polymorphic enough to distinguish among unrelated isolates and stable enough to detect isolates that are related [8]. In addition the selected methodology should be simple, affordable with rapid turnaround time and the

results should be expressed in a standard format that can be shared and compared among different laboratories. To date, no single genetic marker is equally suitable for molecular epidemiology and phylogenetic studies.

Comas and colleagues [14] propose a genotyping approach with existing methods in which strain lineages are first identified by SNPs for phylogenetic studies and further delineation of individual strains within each lineage for epidemiological purposes is achieved with VNTR markers. In addition, because WGS has the potential to replace current genotyping laboratory methods in the future a WSG strategy should be entertained where feasible. A WSG strategy would generate increasingly accurate data for evolutionary and epidemiological studies as well as more robust classification of strains during clinical or experimental association studies. Future implementation of genotyping methodologies at a global level will allow us to make national and international comparisons of MTBC strains with the potential for further refining our understanding of the evolutionary structure of MTBC and the biology of this formidable species.

## 5.6 Molecular Epidemiology Lessons for Tuberculosis Control

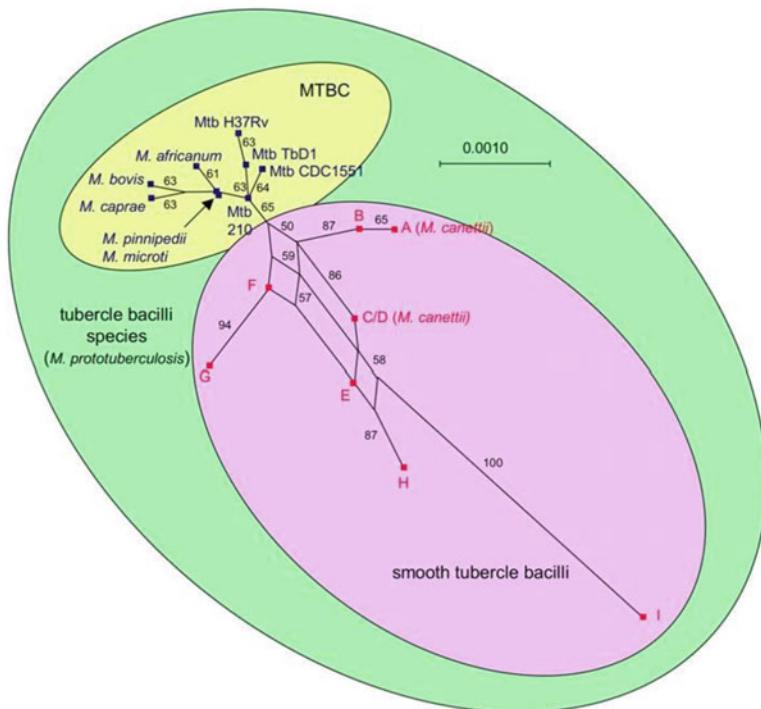
In the last two decades, molecular epidemiological studies generated important new insights into the transmission dynamics of tuberculosis, mostly in developed countries [34, 36, 41–45]. For example, before the availability of more reliable molecular epidemiological tools, the belief was that 10 % of patients developed *M. tuberculosis* disease as a result of recent transmission. However, population-based molecular epidemiological studies in San Francisco demonstrated that the rate of recent infection was as much as 33 % which was much higher than the estimated 10 % predicted by traditional epidemiological studies [42]. In San Francisco, for example one patient generated 6 % of new cases of tuberculosis in a year due to transmission. Similar studies in New York and Amsterdam and other cities demonstrated recent transmission of up to 40 % [67]. In addition, these studies found that recent transmission was associated with persons of lower socio-economic groups, native ethnic minorities, and HIV coinfection [31, 68, 69]. Furthermore, these studies highlighted the importance of tuberculosis control efforts in interrupting transmission, especially amongst groups at high risk. In a meta analysis of pooled data from molecular epidemiology studies, alcohol abuse, injection drug use, and homelessness, all characteristics of marginalized populations, were found to be consistently significant in populations of low TB incidence [33]. Therefore the use of genotyping tools has become important for tuberculosis control programs to assess ongoing transmission of infection and determine its contribution to disease burden. The implementation of molecular epidemiology tools is now an integrative part of tuberculosis control programs in many developed countries at a local and national level. However, more research is needed to

understand TB transmission dynamics in high-burden countries. Additionally, the focus of molecular epidemiology studies so far has largely been on tracing outbreaks of closely related strains in high-income countries. Nevertheless, with the use of more informative tools, a microbial evolutionary perspective is beginning to emerge with a global overview of the genetic diversity and phylogeography of *M. tuberculosis*. In the following sections, we will review the most important research developments in the population structure and host-pathogen coevolution of *M. tuberculosis*.

## 5.7 Origin of Tuberculosis

The origins of *M. tuberculosis*, a uniquely human pathogen, until very recently were hidden somewhere in antiquity, in the distant coevolution of man and this pathogen. However, these origins have begun to be unraveled through breakthroughs in the study of human and pathogen evolution. Current fossil record evidence indicates that fully modern humans emerged in sub-Saharan Africa about 19,500 years ago [70]. New genetic evidence indicates that the dispersal out of Africa occurred closer to 60,000–50,000 years ago and by around 35,000 years ago, modern humans were present in Europe, Asian, and Australia [20]. The presence of all major lineages of MTBC in Africa, including the unique presence of *M. canettii*, a smooth tubercle bacilli, which shares a remote common ancestor with MTBC and whose origin is the Horn of Africa, led to the hypothesis that tuberculosis originated in Africa. For example, Gutierrez and colleagues proposed that *M. canettii* and the other smooth TB bacilli represent a group of organisms likely to contain the ancestor of MTBC [71]. These authors based their findings on DNA sequence data of six housekeeping genes which date the origin of what they have referred to as “*Mycobacterium prototuberculosis*” to approximately 3 million years (Fig. 5.1). Not surprisingly, this finding is controversial as it suggests that the tuberculosis progenitor predated the origin of modern humans [2]. Having said that, the earliest evidence of tuberculosis in early humans comes from a cranial frontal bone of a 500,000-year-old *Homo erectus* fossil excavated in Turkey [72]. However, this paleomicrobiological connection still needs to be confirmed.

An African origin for human tuberculosis is also consistent with what we know about human evolution. During hunter-gatherer times, human populations remained small and geographically scattered, which could have favored within-family transmission and stable host-pathogen associations or low virulent organisms [73]. In other words, in this environment, the initial adaptation for a longer latency period in human tuberculosis occurred as a mode of microbial survival due to the low density of the host populations available throughout hunter-gatherer times [4]. Thus, the small effective population size of the hunter-gatherer stage of human evolution was a bottleneck for highly virulent human tuberculosis, selecting pathogens that could be transmitted decades later after the initial infection. With the larger population sizes that have developed since the rise of agriculture, evolution

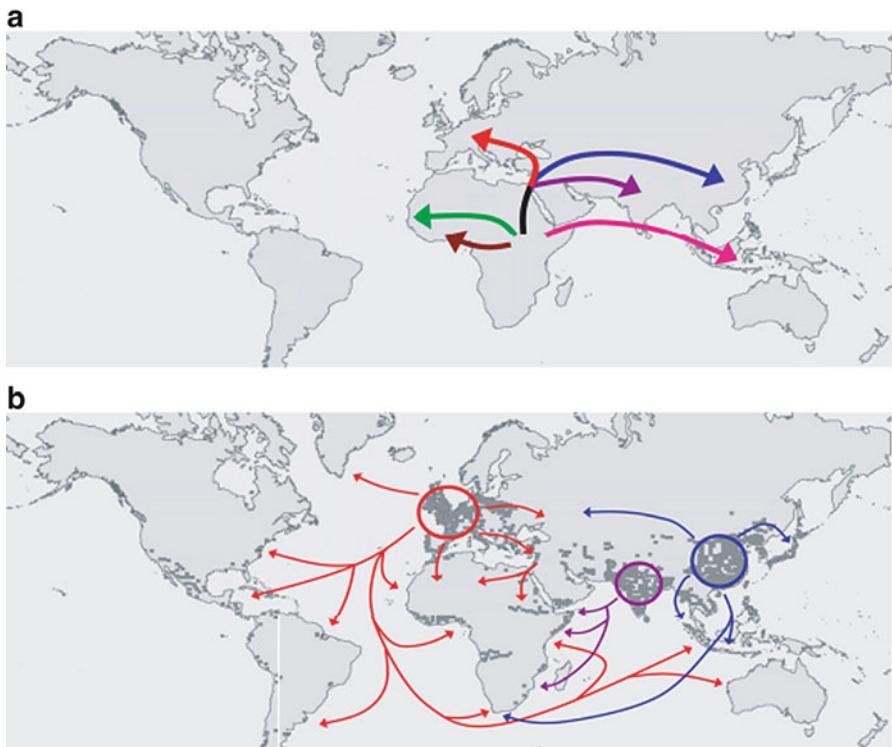


**Fig. 5.1** Phylogenetic associations between “smooth” tubercle bacilli and classical MTBC. Multilocus sequence data of six housekeeping genes show that the smooth tubercle bacilli are more genetically diverse than MTBC. The phylogenetic relationships among members of the smooth tubercle bacilli suggest ongoing horizontal gene exchange, which appears to be absent in MTBC. Adapted from Gutierrez et al. [71], and Gagneux [20]

favored more virulent human-adapted tuberculosis pathogens. Therefore the phylogeographic distribution of MTBC we see today [27] reflects the ancient *M. tuberculosis* human migration out of Africa around 50,000 years ago with more stable host–pathogen associations and less-virulent MTBC lineages. It is not until the more recent population explosion that we see the emergence of more virulent lineages of MTBC, similar to the phylogeographic structure that has been proposed for *Helicobacter pylori* [74].

## 5.8 Out-of-and-Back-to-Africa Scenario

The first evidence for a strong phylogeographical association between *M. tuberculosis* and its human host comes from the preferential sympatric combinations of MTBC and particular ethnic groups [75]. However, it was not until the analysis by Hershberg et al. of DNA sequences from 89 genes in 108 isolates of MTBC that the first robust study of genetic distances between strains provided



**Fig. 5.2** “Out-of-and-back-to-Africa” scenario for the evolutionary history of human adapted *M. tuberculosis*. (a) MTBC originated in Africa and some lineages accompanied the out-of-Africa migrations of modern humans. (b) The three evolutionary “modern” MTBC lineages scattered throughout Europe, India and China, and expanded as a consequence of the sharp increases in human populations in these regions starting a few centuries ago. These lineages then spread throughout the world via human migration. Colored arrows correspond to the six main human-adapted MTBC lineages shown in Fig. 5.3. Adapted from Hershberg et al. [27]. A review of the evolutionary scenario proposed in this figure can be found in Gagneux et al. [20]

strong evidence for a new evolutionary scenario of the origin of MTBC [27]. This scenario proposes that anatomically modern humans were infected with *M. tuberculosis* prior to their migration out of Africa and that *M. tuberculosis* has evolved and remained intimately associated with human populations ever since. Following modern human migration patterns, ancient *M. tuberculosis* underwent population expansion and diversification in Africa and out of Africa, into South of India and Southeast Asia (Fig. 5.2) [27, 70]. For example, Hershberg and colleagues see the distribution of the ancient Lineage 1 around the Indian Ocean as the earliest spread of modern humans out of Africa around 50,000 years ago [27, 70]. A similar scenario has been proposed for *H. pylori* [74]. Subsequently, the modern MTBC lineages 2, 3, and 4, spread from Africa with human migration into Europe and Asia, followed much later by the global spread to the Americas and return to Africa through waves of trade, conquest, and travel [27].

## 5.9 The Spread of Human TB to Animals?

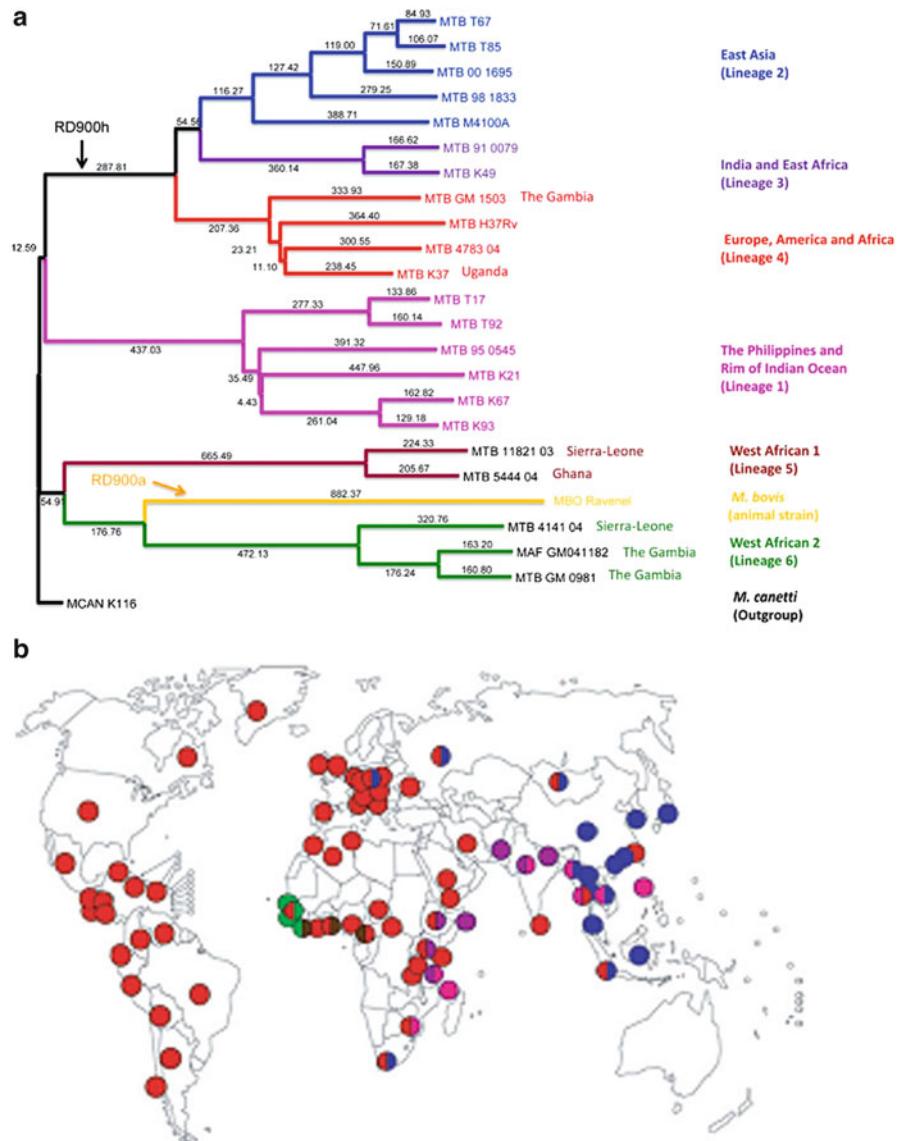
There is archeological evidence in the Neolithic period of skeletal tuberculosis with abnormalities of the spine. The best and most abundant documentation comes from Egypt, where some skeletons demonstrate typical deformities of extrapulmonary tuberculous disease. This form of disease, which is also commonly found in humans with *M. bovis*, led to the compelling view that *M. bovis* spread to humans from animals when man first started farming and domesticating animals sometime during the Neolithic [76]. This hypothesis was reinforced by the ecological understanding of *M. bovis* as a generalist, having a broader host animal range beyond cattle, and *M. tuberculosis* as a specialist, whose range is mostly in humans. From the Neolithic period to modern times, it was believed that *M. bovis* evolved to the more specialized human pathogen, *M. tuberculosis*. This elegant evolutionary hypothesis has been revised drastically in the last 12 years and we now believe the opposite is true. There is indisputable evidence from comparative genomics and population genetic studies that *M. tuberculosis* species in animals evolved from lineages of *M. tuberculosis* that cause disease in humans [18, 77]. There is also evidence that animal-derived members of the MTBC clustered together, while human-derived members of the MTBC phylogeny represent different human populations [27]. In addition, other studies demonstrate that *M. bovis* has undergone numerous deletions from its genome relative to *M. tuberculosis* and that its genome is 60,000 base pairs smaller than the genome of *M. tuberculosis* [78, 79]. Since *M. tuberculosis* is a clonal organism that is not able to repair genomic deletions by recombination, it is assumed that *M. bovis* originated from a familial species with a larger genome [2, 20]. In addition, paleomicrobiological evidence supports comparative genomics findings that *M. tuberculosis* preceded *M. bovis*. For example, genotyping of ancient DNA from a 17,000-year-old sample from a bison and from numerous human samples from between 9,000 to 10,000 years ago demonstrate the presence of human-associated *M. tuberculosis* and *M. africanum*, but not animal-associated *M. bovis* [80]. Consequently, with confidence we can now say that *M. tuberculosis* preceded *M. bovis* in the evolutionary hierarchy, not the other way around. Interestingly comparative genomic analyses have revealed a high degree of genetic diversity in *M. bovis* and the species are comprised of several ecotypes, each of which is adapted to particular animal host species [81].

## 5.10 The Phylogeographic Structure of MTBC

Human-adapted members of the *M. tuberculosis* complex (MTBC) until recently have been assumed to be essentially genetically identical [19]. However, more recent comparative sequence studies of global human strains from different regions of the world found a greater genetic variation, which afterwards has been exploited for a more robust strain classification [82]. This genetic variation was studied with

informative LSP markers in a global sample of 875 *M. tuberculosis* clinical isolates from 80 countries to show that all isolates belonged to one of the six major lineages [75]. This study defined a robust phylogeny for MTBC and the results were congruent with results from other studies that reveal a particular geographic distribution of six main strain lineages of *M. tuberculosis* and *M. africanum* [37, 49, 75, 83, 84]. However, all of these studies had important limitations because the actual phylogenetic distances and relative genetic diversities within and between mycobacterial lineages had not been determined [2, 25, 82]. In addition some of these studies used genotyping markers such as spoligotyping and MIRU/VNTR which are prone to convergent evolution and are limited in their application for phylogenetic and population genetic analysis [14]. Also, some of these studies used a small number of informative SNPs obtained from few genome sequences of MTBC strains that were available at the time [37, 83]. Limited numbers of SNPs from analysis of few genomes generate a phylogenetic discovery bias and misinterpretation of strain genetic associations [25]. Although Gagneux and colleagues were able to use genomic deletions to classify strains unambiguously, LSP-based phylogenetic studies are also difficult to interpret and deletions do not correlate with phylogenetic distances and therefore do not indicate how closely related one strain is to another [2, 82].

More recently, Hershberg and colleagues [27] investigated the genetic distances between strains and produced the most well defined phylogeny of the MTBC using in-depth analyses of a large set of coding sequence data from a global collection of MTBC isolates. The authors investigated the genetic diversity within MTBC using seven megabases of DNA sequence data from a representative collection 99 human-adapted strains and seven strains from animal-adapted ecotypes [27]. This study found strong evidence for a clonal population structure of MTBC, without evidence of ongoing horizontal gene transfer. In addition, this study revealed that the human-adapted members of MTBC are more genetically diverse than generally recognized and that this diversity can be linked to human demographic and migratory events. Based on their findings Hershberg and colleagues proposed an evolutionary scenario of human-adapted lineages of MTBC that can be divided into two major groupings: ancient and modern. In a subsequent study, Comas and colleagues [6] used a total of 9,037 unique SNPs to derive a genome-wide phylogeny of 22 strains representative of the global diversity of the *M. tuberculosis* complex (MTBC). The authors, further defined six major lineages into the current nomenclature that we use today: Lineage 1 (The Philippines, Rim of Indian Ocean), Lineage 2 (East-Asian), Lineage 3 (India–East Africa), Lineage 4 (Euro–American, Africa), Lineage 5 (West Africa 1), and Lineage 6 (West Africa 2), more commonly known as *M. africanum* (Fig. 5.3) In the new proposed phylogeny, Lineage 5 and 6 (*M. africanum*) are the most ancestral groups. These findings are also supported by the resent WGS of lineage 6 (West African type 2) which was found to contain a unique sequence, RD900, that was independently lost during the evolution to the modern *M. tuberculosis* group (Lineage, 2, 3, and 4) and the lineage leading to the MTBC associated with animal strains (Fig. 5.3) [24]. These studies strongly suggest that *M. africanum* represent the most ancient



**Fig. 5.3** The global population structure and geographical distribution of MTBC. (a) Global phylogeny of human-adapted MTBC is based on analysis of 22 whole-genome sequences by Gagneux et al. [75], and the whole genome sequence data of *M. africanum* GM041182 [24]. Numbers in branches refer to the corresponding number of SNPs inferred. Lineage names are according to Hersbergh et al. [27], and lineage numbers are according to Comas et al. [6]. Adapted from Comas et al. [6], and Bentley et al. [24]. (b) Global distribution of the six main human-adapted MTBC lineages. Colored dots represent the dominant MTBC lineages by countries. A review of the nomenclature and comparison with other typing systems can be found in Coscolla and Gagneux [85] and Gagneux [20]

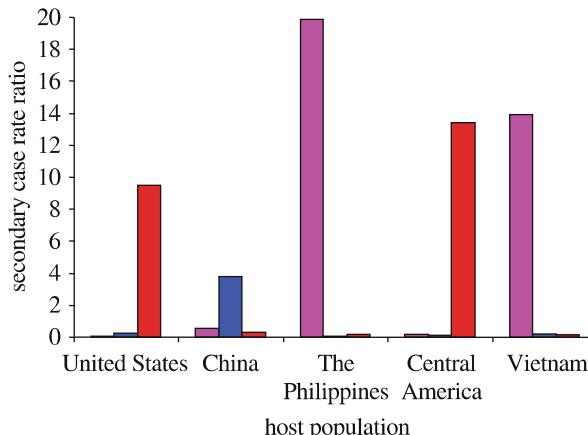
lineages and since these lineages are indigenous to Africa, this finding reinforces the notion that human MTBC originated in Africa (Fig. 5.2) [27, 75, 86].

Historically in microbiology there has been a problem in distinguishing genotype from phenotype in bacteria due to the difficulty in defining meaningful biological and genetic boundaries within related species and subspecies. Due to this difficulty, there is no widely accepted agreement of a “species concept” for bacteria [25]. The lack of classification consensus between species and subspecies applies to MTBC, where there exists no agreement for a taxonomic scheme. This lack of robust classification among strain grouping in MTBC is partly responsible for the usual discordant results and adds confusion in interpreting studies of genotype-phenotype associations. The numerical nomenclature proposed by Comas and colleagues [6] for the six major lineages of MTBC help establish important genetic boundaries (Fig. 5.3). Recently, Coscolla and colleagues [85] used this nomenclature of MTBC to review nearly 100 published reports of strain virulence and immunogenicity in experimental models. We will use this numerical nomenclature combined with some group names commonly used in the past to help summarize the genotypic and phenotypic differences of two important lineages of MTBC and their associations to clinical outcomes.

## 5.11 MTBC Virulence and Host–Pathogen Coevolution

Given the long road in the origins of MTBC in Africa and its subsequent spread around the world, the question remains; are there signs of host pathogen coevolution? In regards to this issue, more evidence supports a population structure of MTBC lineages with adaptation to different human populations [37, 49, 83, 86]. As proposed by the evolutionary scenario described by Hershberg et al. [27], and modified by Comas et al. [6], the three “modern lineages,” 2, 3, and 4 are thought to be more successful than the other “ancient lineages,” lineages 1, 5, and 6, due to the population explosion over the last two centuries in Europe and Asia (Fig. 5.3b). That is, in ecological terms, the population explosion favored the emergence of more virulent MTBC lineages because access to a larger susceptible human population allows adaptation to higher microbial virulence and shorter latency periods, as it is suggested to have occurred with the modern lineages of MTBC [4, 20, 87].

Portervin and colleagues [88] studied virulence differences in modern lineages of *M. tuberculosis* in human monocyte-derived macrophages. The authors showed that the macrophages infected with modern strains produced less pro-inflammatory cytokines than those infected with ancient strains. Interestingly, previous work had showed a lower pro-inflammatory immune response profile in modern MTBC strains with higher virulence profiles in animal models [10]. In a molecular epidemiology study in the Gambia, where ancient lineages of *M. africanum* are indigenous (West African Lineage 1 and West African Lineage 2), de Jong and colleagues [89] followed the latency period of tuberculosis among HIV negative household contacts of active tuberculosis cases. The authors found that, although



**Fig. 5.4** Preferential sympatric transmission of MTBC lineages in San Francisco. In each human population, the number of secondary cases caused by sympatric MTBC lineages was significantly higher compared with the secondary cases caused by allopatric lineages. Sympatric and allopatric lineages were defined based on their usual phylogeographic association shown in Fig. 5.2. Red, lineage 4 or Euro-American lineage; Blue, lineage 2 or East-Asian lineage; Pink, lineage 1 or Indo-Oceanic lineage. Adapted from Gagneux et al. [75] and Gagneux et al. [20]

there were no differences in exposure to modern versus ancient MTBC strains, contacts of modern MTBC index cases were more likely to develop active tuberculosis when compared to contacts exposed to index cases with ancient MTBC disease. This latter study supports the adaptation of modern lineages to shorter latency periods. Overall, these studies provide additional support for the existence of host–pathogen adaptation with the evolution to higher virulence and shorter latency periods in modern MTBC strains.

Further evidence for coevolution in human tuberculosis derives from the global population structure of *M. tuberculosis*. Gagneux and colleagues [75] found that the six major lineages defined by phylogeographic associations were more predisposed to having a sympatric transmission than an allopatric transmission. In other words, modern strains of East-Asian lineage 2 were more likely to transmit to Chinese patients; Indo-Oceanic Lineage 1 strains were more likely to transmit to patients from the Philippines and Euro-American lineage 4 strains were more likely to transmit among U.S. born patients (Fig. 5.4). Curiously, U.S.-born tuberculosis patients of non-Chinese and non-Filipino ethnicity who were infected with allopatric strains, belonging to lineages 2 and 3, were more likely to be HIV-positive or homeless. This finding suggests that, although these lineages are less adapted to transmit and cause disease in fully immunocompetent members of allopatric human populations, they can do so in the context of impaired host resistance. Taken together, these observations suggest that particular lineages of *M. tuberculosis* might be adapted to specific human populations and maladapted to others, supporting variable host–pathogen coevolution. However, formal proof for such an adaptation remains elusive because of the difficulty in separating biological factors from the social determinants of disease [15].

### 5.11.1 Lineage 2, East Asian, “Beijing/W” Strain an Important Cause of Tuberculosis in the World

Lineage 2, East Asian, Beijing/W strains have been extensively investigated due to their increased ability to spread and cause disease [9, 11]. Lineage 2 was first identified by IS6110-RFLP and spoligotyping of *M. tuberculosis* strains from the Beijing area in China in 1992–1994 and thus earned the name Beijing strains [36, 90]. Lineage 2 strains are found in East Asia, Eastern Europe and also parts of Western Europe, African, and Latin American countries [11, 36]. This lineage is estimated to cause 50 % of all the tuberculosis cases in Asia [66]. In Vietnam, 53 % of 563 isolates from new TB cases, belonged to lineage 2 [91], and in Hong Kong, 71 % of 500 randomly selected *M. tuberculosis* isolates belong to Lineage 2 [92]. A high proportion of Lineage 2 strains are found in the Beijing city area [93]. Due to these findings it is also assumed Lineage 2 may have originated and spread from East Asia. For instance, Lineage 2 is thought to have its origins in North Central China more than 1,000 years ago [93]. Its subsequent spread into most parts of Asia, Europe, and Africa appear to have been driven by human migration. In the Western Cape region of South Africa, strains belonging to Lineage 2 are thought to have been introduced approximately 400 years ago following the sea trade route from East Asia to Europe [9].

The explosive spread of Lineage 2 is best described by a molecular epidemiology study of *M. tuberculosis* on the island of Gran Canaria, Spain. This genotype was initially introduced by an immigrant from Africa and over a period of 4 years after its introduction, this lineage became the most common strain on the island [43]. A similar situation occurred in New York City, where Lineage 2 strains, at one point, accounted for 25 % of all MDR-TB cases in the USA after their association with nosocomial transmission among HIV-infected individuals during the early 1990s [35]. In the Western Cape in South Africa, molecular epidemiological studies demonstrated that Lineage 2 contributes up to 30 % of the disease burden and that this number appears to be increasing exponentially [94]. For example, among children in Capetown, South Africa, Lineage 2 increased from 13 to 33 % from 2000 to 2003 [94]. Lineage 2 has also been responsible for numerous disease outbreaks throughout the world, many of which have been drug resistant [35]. In some other areas of the world, this lineage has also been associated with cases of tuberculosis in young individuals [94, 95]. Furthermore, the most frequently observed genotype associated with *M. tuberculosis* disease globally is Lineage 2 strains [36, 49].

Some of the most common factors cited for the global success of Lineage 2 include increased migration of people around the world, selective pressure to overcome the protective BCG vaccine-induced immunity and propensity for development of drug resistance and increased transmission [9]. However, the results from different studies on these putative host and biological properties of this lineage appear to be inconclusive due to discordant findings [95]. These conflicting results could represent our inability to further desegregate the population structure of

MTBC into more defined sublineages. In regards to this latter point, some progress has been made. Schurch and colleagues [96] recently established a phylogeny of the *M. tuberculosis* Beijing genotype family based on genome-wide SNPs. The authors showed that the modern Lineage 2 resulted from a recent clonal expansion and they identified 61 SNPs that further subdivided Lineage 2 into 27 sublineages. Further genomic studies with full GWS of more strains of Lineage 2 could improve the resolution of the population structure of the MTBC. These potential findings are not unexpected since phylogeographical associations to strain sublineages have been found in studies of diseases such as plague, *Mycobacterium lepra*, and buruli ulcer [97–99]. These studies highlight the importance of ongoing research into the genetic mechanisms underlying the phenotypic and genotypic characteristics of *M. tuberculosis*. Recently Hanekom and colleagues [9] reviewed the molecular and phenotypic characteristics of lineage 2 East Asian strain, emphasizing the virulence, transmissibility, and apparent association with the development of drug resistance.

### 5.11.2 *M. africanum*: An Important Cause of Human Tuberculosis in West Africa

Half the burden of human tuberculosis in West Africa is caused by *Mycobacterium africanum* [100]. There are two ancient lineages of *M. africanum*: *M. africanum* Lineage 5 West African type 1, common to Eastern West Africa, and Lineage 6 *M. africanum* West African type 2, common to Western West Africa. These two types of *M. africanum* are distinct subspecies within the *M. tuberculosis* complex. *M. africanum* West African type 2 is phylogenetically closer to the animal strains like *M. bovis* [18, 27, 101]. The proximity of *M. africanum* West African type 2 to the animal isolates on the phylogenetic tree of the MTBC raises the possibility of an animal reservoir for this lineage (Fig. 5.3). Phylogenetically, the *Dassie bacillus* [22] and the recently identified *Mycobacterium mungi* [23] are the closest relatives of *M. africanum* within the *M. tuberculosis* complex. However, no candidate animal reservoir has been detected for *M. africanum* despite extensive surveillance among cattle, sheep, pigs, and goats in the West African countries [24, 100]. Today, *M. africanum* has been infrequently identified in humans in areas outside of West Africa [100].

Lineage 6 *M. africanum* type 2, shows distinct clinical presentations when compared to the other human-associated lineages of *M. tuberculosis*. Patients with this lineage tend to be older, HIV infected, severely malnourished, and have more pulmonary infiltrates by chest-X-ray despite similar duration of disease when compared to patients with *M. tuberculosis* [100]. Interestingly, in this study, the authors observed similar transmission rates of *M. africanum* and *M. tuberculosis* among household contacts. However, contacts exposed to *M. africanum* were significantly less predisposed to progress to active disease than contacts exposed

to *M. tuberculosis* [89]. Remarkably, despite massive migrations of West Africans to the Americas through slavery, *M. africanum* did not establish itself in the Americas because it was probably outcompeted by *M. tuberculosis* due to its lower rate of progression to disease [100]. Parenthetically, lower progression to disease of *M. africanum* makes this organism a potential model for understanding containment of progression from infection to disease. In addition the recent whole genome sequence of *M. africanum* identified genomic differences between *M. africanum* West African 2 and the other strains of the MTBC which could explain the apparent attenuated characteristic of *M. africanum* [24]. Surprisingly, a recent study described in Ghana a host genetic polymorphism (IRGM-261TT) that is associated with protection against Euro-American *M. tuberculosis*, yet not against *M. africanum* [102]. This result raises the possibility that the increased frequency of this polymorphism may have contributed to the establishment of *M. africanum* as a pathogen in West African populations. Taken together, these findings demonstrate the importance of host-pathogen factors in the variability in progression to tuberculosis disease in humans. These findings have important implications for our understanding of tuberculosis epidemics and the development of vaccines and treatments for latent tuberculosis infection. In the following section, we will review the tuberculosis control factors that are influencing the emergence of drug resistance and the microbial evolutionary perspectives that are shaping the biology and epidemiology of drug-resistant tuberculosis.

## 5.12 Tuberculosis Control and the Emergence of Drug Resistance

The legacy of poor tuberculosis control and the growing tuberculosis-HIV coepidemic continue to fuel the emergence of tuberculosis and drug resistance, which in turn threatens efforts of achieving the United Nations millennium goal of eliminating tuberculosis by the year 2050 [103]. For instance, in the past 3 decades the world has observed the evolution of strains from drug susceptible to Multidrug-resistant tuberculosis (MDR-TB) to extensively drug-resistant (XDR) tuberculosis. More recently the emergence of totally drug-resistant strains of *M. tuberculosis* (TDR-TB) have been reported in India and Iran [104, 105]. Resistance to at least isoniazid and rifampicin, the most effective drugs against tuberculosis, is defined as MDR-TB. In turn, MDR-TB plus resistance to a fluoroquinolone and at least one second-line injectable agent is defined as XDR-TB. Despite the emergence of virtually untreatable disease, no new drugs have been introduced for the treatment of tuberculosis in the last 50 years [82].

The greatest tragedy that can happen to a patient with tuberculosis is evolution of MDR or XDR because the chances of cure will be drastically reduced and treatment costs can overwhelm local and National Tuberculosis Programs. For instance, the cost of drugs alone for treating the average MDR-TB patient is 50–200 times higher

than for treating drug-susceptible TB patients. In addition to the overall cost of drugs, the cost for case management of MDR-TB patients has been found to be more than ten times higher than that for drug-susceptible patients ([www.who.int/tb/challenges/xdr/faqs](http://www.who.int/tb/challenges/xdr/faqs)). These forms of drug-resistant tuberculosis are associated with poor cure rates because treatment requires prolonged duration of therapy for up to 24 months with second line drugs that are less effective and more toxic [106]. Because cure depends on the extent of drug resistance, the lack of success of treatment for XDR-TB is much smaller than for MDR-TB. In addition, worldwide only a small proportion of these cases are detected and treated given that many low- and middle-income countries lack financial resources and capacity for second line drug susceptibility testing, required to detect MDR/XDR-TB. In relevance to this issue, only two of 27 high- burden MDR-TB countries routinely test for resistance to second-line drugs [107]. Currently, there are no strategies for treating XDR-TB patients and many of these continue to reside in their communities until their death. This latter situation poses the greatest challenge to tuberculosis control by allowing the opportunity of transmission of highly resistant bacteria in many communities.

Prior exposure to anti-TB drugs is a well-established risk factor for drug resistance. Sequential global surveys have shown that the prevalence of drug-resistant TB is influenced by the quality of TB control programs and by local epidemiologic circumstances [108–111]. In this context, evolution from drug-susceptible to drug-resistant organisms usually emerges due to poor tuberculosis control practices. For instance, lack of access or delay in drug susceptibility testing, inadequate treatment regimens, unsupervised treatment, interrupted drug supplies, and lack of adherence to treatment contributes to the emergence of drug resistance. In addition, a patient's comorbidities such as HIV can influence the pharmacodynamic properties of TB drugs and increase the chance of drug resistance development [112, 113]. As a consequence of poor management, step-wise accumulation of chromosomal resistance mutations occurs, with the progressive accrual of mutations from single drug resistance to the emergence of MDR and XDR organisms. This process of evolution to drug-resistant organisms can occur in individual patients (secondary resistance) or as a result of direct transmission of drug-resistant strains (primary resistance) with subsequent accumulation of additional drug resistance by further transmission and incorrect treatment. Thus, the evolution of MDR and XDR tuberculosis is not only a catastrophe for individual patients, but also for other patients by the further spread and amplification of drug-resistant tuberculosis in the community.

In this context of poor tuberculosis control, it is not surprising that the World Health Organization (WHO) surveillance data of multidrug and extensively drug-resistant tuberculosis reported the highest rates of resistance ever documented [114]. For example, among all incident TB cases globally, 3.6 % are estimated to have MDR-TB. In 2008, an estimated 500,000 cases of MDR- TB emerged globally and MDR disease caused an estimated 150, 000 deaths. The highest incidence of MDR and XDR tuberculosis are found in countries with a history of poor tuberculosis control. For example, almost 50 % of MDR-TB cases worldwide are estimated

to occur in China and India. In China alone, 100,000 MDR-TB cases are emerging annually. China has reported 5.7 % MDR-TB among new cases and 25.6 % among those previously treated. In Russian data from 12 sites, reported proportions indicate up to 28 % MDR-TB among new TB cases. Eight countries reported XDR-TB in more than 10 % of MDR-TB cases and a prevalence of up to 23 % of XDR-TB among MDR patients in South Korea [110]. To date, a cumulative total of 77 countries have confirmed patients with XDR-TB [110, 115]. This alarming epidemic of highly resistant tuberculosis highlights the importance of understanding the tuberculosis control forces shaping the evolution and spread of drug-resistant tuberculosis.

The increasing rates of MDR and XDR tuberculosis indicate persistently poor TB control practices [107]. A particular challenge is the current evidence of extensive transmission of primary XDR tuberculosis [110]. A case in point is South Africa, where HIV infection has been a strong risk factor for primary XDR tuberculosis as described in a highly publicized outbreak in Tugela Ferry [109]. In a more recent follow up study in South Africa, Gandhi and colleagues [109] found that the majority of patients with primary XDR disease were HIV positive and had been admitted to a hospital and had experienced prolonged exposure to XDR tuberculosis patients. In addition they documented with a molecular epidemiology study, multiple generations of XDR tuberculosis transmission over time [116]. Moreover the XDR tuberculosis strains had identical resistance mutations against primary and second line drugs, supporting the notion that this was an outbreak of a highly resistant XDR strain. In addition the patients in this study were linked epidemiologically to the same hospitals [116]. However, this is not the only study in the context of an outbreak in South Africa; there is a broader epidemic of XDR tuberculosis occurring in South Africa. For instance, from 2001 to 2007 in KwaZulu Natal the numbers of cases have grown exponentially to a total of 792 XDR tuberculosis cases. This rapid geographic dissemination strongly suggests primary transmission for the XDR epidemic in South Africa. This hypothesis is supported by identical whole genomic sequencing of the Beijing/lineage 2 strains of the patients in these outbreaks [117]. Of interest, this Beijing strain appears to be part of a broad multi-institutional epidemic of XDR tuberculosis that was first detected in Tugela Ferry back in 2006 [109].

Drug-resistant tuberculosis was recognized as a significant public health threat in high-income countries in the early 1990s following many nosocomial outbreaks of MDR-TB due to poor infection control practices [35, 68, 118–120]. The current XDR tuberculosis epidemic with almost incurable forms of disease demonstrates the need to prevent further emergence of drug resistance by strengthening tuberculosis programs, identifying patients with drug resistance earlier on and ensuring that patients complete their tuberculosis treatment under directly observed treatment programs [121]. However, as per prior experience to mitigate the transmission of drug resistance, other control strategies must include efforts to treat patients with effective therapy and also strengthening infection control practices to prevent the spread of MDR and XDR tuberculosis [122]. In addition, in settings with a high incidence of HIV-TB coinfections, diagnosing HIV early and treating HIV-infected

patients with antiretrovirals could result in reduced progression of latent disease to active disease [123, 124]. However, although it is well known that treatment of latent tuberculosis infection reduces the risk of active TB in HIV positive individuals, this strategy will likely not work with individuals infected with MDR and XDR tuberculosis.

### 5.13 Bacterial Fitness and Compensatory Evolution in Drug Resistance

Although poor tuberculosis control practices are indisputably an important factor in the selection and transmission of drug-resistant strains, these practices do not sufficiently explain the current global distribution of drug-resistant TB [107, 114]. Thus, we continue to debate whether the current problem of drug-resistant TB is primarily attributable to direct transmission of drug-resistant strains or to development of resistance during the treatment of individual patients [125]. In terms of primary resistance, there are serious reservations regarding the extent of the dissemination of MDR and XDR tuberculosis strains in areas of hot spots of drug resistance [126, 127]. For example, experimental data and mathematical models suggest that the reduction in bacterial fitness imposed by antimicrobial resistance could influence the frequency of drug-resistant TB in a population [128, 129]. These considerations are important in order to determine whether drug-resistant and drug-susceptible strains of *M. tuberculosis* have an equal likelihood of spreading in the community. For instance, a scenario in which drug-resistant and drug-susceptible strains are equally transmissible will lead to the eventual emergence and establishment of MDR and XDR strains of *M. tuberculosis* due to increased difficulties in treating patients with drug-resistant pathogens. Furthermore, these patients will remain infectious for prolonged periods of time.

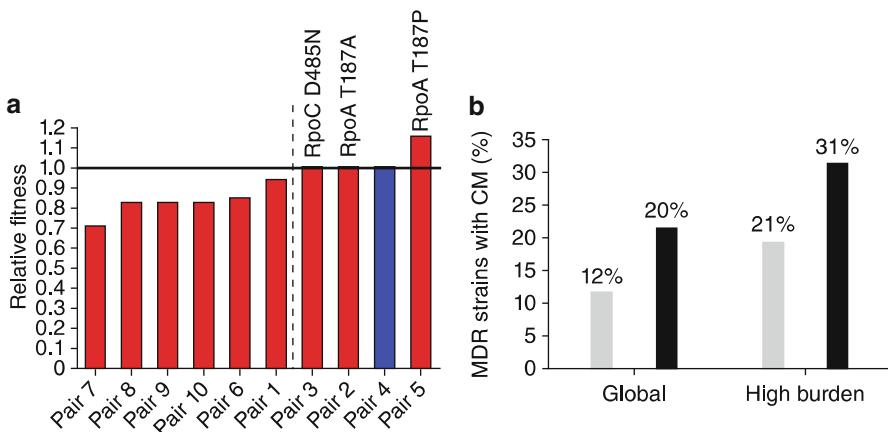
Our current understanding of the emergence of drug resistance lacks a microbial evolution perspective that takes into account the evolutionary forces that drive the emergence of MDR and XDR strains of *M. tuberculosis* [130]. For instance, the relative fitness of drug-resistant strains of *M. tuberculosis* can be influenced by the specific drug resistance-conferring mutation and strain genetic background [131]. Additionally, compensatory evolution, which has been shown to mitigate the fitness cost associated with drug resistance in other bacteria, is also an important factor in the emergence and spread of drug-resistant *M. tuberculosis* [132]. In this context, we will proceed to review some of the most important molecular mechanisms and evolutionary forces that drive drug resistance in *M. tuberculosis*.

One of the major obstacles to understanding the molecular basis of tuberculosis resistance is differentiating between primary drug resistance-conferring mutations and secondary compensatory mutations that arise to counteract the detrimental effects on bacterial physiology. For instance, isoniazid is a prodrug that is activated by the product of the *katG* gene, a catalase-peroxidase protein, leading to the

interference of mycolic acid synthesis [133, 134]. The peroxidase function of *katG* has an antioxidant effect which protects the bacillus against oxidative stress. Mutations in the *katG* gene that reduce this function result in resistance to isoniazid and make *M. tuberculosis* more susceptible to oxidative stress [134]. As with other infectious bacteria, *M. tuberculosis* develops compensatory mutations that mitigate the fitness cost imposed by the acquisition of resistance [135]. For example, some clinical strains with *katG* mutations also contain mutations in the promoter of the *ahpC* gene that result in overexpression of *ahpC*. This overexpression appears to help the microorganism withstand the oxidative stress previously controlled by a well-functioning *katG* gene [136].

Other studies have evaluated point mutations for their virulence propensity. For example, the isoniazid mutation most commonly seen in clinical isolates is the point mutation Ser315 Thr in *katG*. This variant in a laboratory mutant manifested reduced capacity for prodrug activation while still retaining some of its catalase-peroxidase function [137], and retained virulence in a mouse model [138]. There is limited data on the pathogenesis of MTB isoniazid-resistant phenotypes with alterations in non-*katG* genes, and these studies have yielded inconsistent results regarding virulence. For example, Wilson and colleagues [139] found that virulence was reduced in both susceptible and resistant strains even if the resistant strains included mutations in non-*katG* genes. In contrast, Heym and colleagues [140] found that restoration of virulence was not seen in an isoniazid-resistant *M. tuberculosis* strain that over-expressed *ahpC* (a putative compensatory mechanism to the initial resistant mutation). The discrepancy between these two studies could be explained by differences in the mutational profile of laboratory-induced resistant strains and strains derived from actual patients. For instance, Gagneux and Colleagues [131] used a population-based molecular epidemiological approach to measure the relative transmission of isoniazid-resistant strains of *M. tuberculosis* harboring different *katG* mutations in TB patients in San Francisco during a 9-year period. The authors found that only strains with the *katG* S315T mutation or mutations outside of *katG* were associated with successful transmission. In contrast, other strains with isoniazid resistance-conferring mutations likely to diminished *katG* activity did not cause one secondary case during the study period. These results demonstrate that different isoniazid drug resistance-conferring mutations can have heterogeneous effects on strain fitness and transmission.

In the case of rifampin, resistance mutations developed during treatment through the selection of bacterial variants in the *rpoB* gene. To study the fitness influence of in vitro-derived and clinically derived rifampin-resistance mutations on *M. tuberculosis*, Gagneux and colleagues performed competitive fitness experiments in drug-susceptible and drug-resistant paired strains [141]. The authors found that in these *rpoB* laboratory-derived mutants, rifampin resistance was universally associated with a competitive fitness cost. Additionally, they found, the specific resistance mutation and the lineage background of the strains influenced this fitness defect. For instance, the *rpoB* S531L mutant had higher fitness in both strain backgrounds, while the *rpoB* H526D mutant had a different fitness defect in the two different lineages. Interestingly, when the authors compared clinical strains of



**Fig. 5.5** Experimental and clinical relevance of presumed compensatory mutations. **(a)** Experimental competitive fitness of ten clinical isolates that acquired rifampin resistance over the course of treatment compared to their susceptible counterparts. The amino acid changes encoded by high probability of compensatory mutations are indicated in the pair in which they were identified. Bar colors indicate strain lineage (blue, lineage 2; red, lineage 4). **(b)** Percentage of MDR strains with putative compensatory mutations in *rpoA* or *rpoC*. Gray bars, the percentage of strains carrying high-probability mutations; black bars, strains carrying any putative compensatory mutation. Data for a global collection of strains and for regions of Abkhazia/Georgia, Uzbekistan, and Kazakhstan with high MDR TB burden are shown. Adapted from Comas et al. [132]. A review of the relevance of compensatory mutations can be found in a review by Borrell and Gagneux [20]

*M. tuberculosis* that had acquired rifampin resistance during patient treatment, they found that four out of five strains with the *rpoB* S531L mutations had no fitness defects when compared to the drug-susceptible counterparts. Moreover, the *rpoB* S531L mutation associated with the lowest fitness cost in the laboratory and no fitness cost in some clinical strains is also the most frequent rifampin resistance mutation among clinical isolates globally.

An intriguing question is whether rifampin-resistant strains also developed compensatory mutations. In order to address this issue, Comas and colleagues [132] used WGS of laboratory and clinically derived strains. The authors demonstrated that the acquisition over time of particular mutations in *rpoA* and *rpoC* genes in rifampin-resistant *M. tuberculosis* strains, results in the emergence of MDR strains with higher fitness (Fig. 5.5). In other words, these mutations acted as compensatory alterations that mitigated the fitness cost associated with resistance mutations in *rpoB*, the target of rifampin. Moreover, the authors demonstrated that these compensatory mutations were present in more than 30 % of MDR clinical isolates from high-burden countries. The role of compensatory mutations in *rpoA* and *rpoC* have also been shown to exist in *Salmonella enterica*, where these mutations mitigate the impaired fitness due to rifampin resistance-conferring mutations in the *rpoB* gene [142]. A follow up study in South Africa in an area of high incidence of tuberculosis confirmed the role of *rpoC*-compensatory mutations and

strains with different genetic background in the transmission of MDR and XDR tuberculosis [143]. These studies illustrate the complex evolutionary interactions between drug resistance-conferring mutations, compensatory mutations, and the different strain lineages in *M. tuberculosis*.

From the above studies one can conclude that isoniazid, rifampin, and MDR strains are heterogeneous in terms of both resistance mutations and virulence. The findings that compensatory evolution is selecting for more fit strains of MDR and XDR tuberculosis is alarming because in modeling studies of drug resistance, a small population of relatively fit MDR and XDR strains can eventually outcompete both the drug-susceptible strains and the less fit resistant strains [128, 129]. In other words, the public health implication of the emergence of drug resistance is that in the absence of specific efforts to limit transmission, MDR and XDR tuberculosis can reach epidemic proportions as is the case in some countries. In the United States substantial human and financial resources were invested in combating the epidemic of drug resistance in the 1990s [121]. We need a similar response worldwide to reverse the spread of untreatable forms of *M. tuberculosis*. At this juncture in time, we invest now or continue to pay later.

## 5.14 Conclusion

Molecular epidemiology tools play an integrative part in tuberculosis control programs in many developed countries at a local and national level. However, despite these new advances, tuberculosis continues to remain a devastating infectious disease, disproportionately impacting the world's poorest countries. Other technologies based on the knowledge of complete genome sequences are helping advance our understanding of *M. tuberculosis* in all its complexity—from its origins in Africa to its increasing genetic variation that parallels historical patterns of human evolution and migration. These new insights into the biology of *M. tuberculosis* have revolutionized the molecular and genetic epidemiology of this pathogen. From the new phylogeographic MTBC framework, it is essential to investigate the clinical implications of strain variation. In particular, the possible consequences of host–pathogen coevolution and the impact of MTBC diversity on immune responses for the development of new diagnostic test, drugs, and vaccines. Developing new TB diagnostic tests, shorter treatment regimens, and a highly effective vaccine may be the only ways to eliminate tuberculosis in the future. The remaining challenge, however, is to complement the basic principles of tuberculosis control with these new tools on a global scale. Regardless of the instruments we develop, improving the odds of eliminating tuberculosis by 2050 will rest on the social, institutional, and epidemiological context in which we apply these tools.

## References

1. Russell DG (2011) *Mycobacterium tuberculosis* and the intimate discourse of a chronic infection. *Immunol Rev* 240(1):252–268
2. Smith NH, Hewinson RG, Kremer K, Brosch R, Gordon SV (2009) Myths and misconceptions: the origin and evolution of *Mycobacterium tuberculosis*. *Nat Rev Microbiol* 7 (7):537–544
3. Iseman MD (2000) A clinician's guide to tuberculosis, 1st edn. Lippincott Williams & Wilkins, Philadelphia
4. Blaser MJ, Kirschner D (2007) The equilibria that allow bacterial persistence in human hosts. *Nature* 449(7164):843–849
5. Small PM, Fujiwara PI (2001) Management of tuberculosis in the United States. *N Engl J Med* 345(3):189–200
6. Comas I, Chakravarti J, Small PM, Galagan J, Niemann S, Kremer K et al (2010) Human T cell epitopes of *Mycobacterium tuberculosis* are evolutionarily hyperconserved. *Nat Genet* 42(6):498–503
7. Burgos M, DeRiemer K, Small PM, Hopewell PC, Daley CL (2003) Effect of drug resistance on the generation of secondary cases of tuberculosis. *J Infect Dis* 188(12):1878–1884
8. Flores L, Kato-Maeda M, Metcalfe JZ (2011) Genotyping of *Mycobacterium tuberculosis*: application in epidemiologic studies. *Future Microbiol* 6(2):203–216
9. Hanekom M, Gey van Pittius NC, McEvoy C, Victor TC, Van Helden PD, Warren RM (2011) *Mycobacterium tuberculosis* Beijing genotype: a template for success. *Tuberculosis* 91 (6):510–523
10. Reed MB, Domenech P, Manca C, Su H, Barczak AK, Kreiswirth BN et al (2004) A glycolipid of hypervirulent tuberculosis strains that inhibits the innate immune response. *Nature* 431(7004):84–87
11. Hanekom M, Van der Spuy GD, Streicher E, Ndabambi SL, McEvoy CRE, Kidd M et al (2007) A recently evolved sublineage of the *Mycobacterium tuberculosis* Beijing strain family is associated with an increased ability to spread and cause disease. *J Clin Microbiol* 45 (5):1483–1490
12. Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D et al (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* 393 (6685):537–544
13. Smith NH, Kremer K, Inwald J, Dale J, Driscoll JR, Gordon SV et al (2006) Ecotypes of the *Mycobacterium tuberculosis* complex. *J Theor Biol* 239(2):220–225
14. Comas I, Homolka S, Niemann S, Gagneux S (2009) Genotyping of genetically monomorphic bacteria: DNA sequencing in *Mycobacterium tuberculosis* highlights the limitations of current methodologies. *PLoS ONE* 4(11):e7815
15. Comas I, Gagneux S (2009) The past and future of tuberculosis research. *PLoS Pathog* 5(10): e1000600
16. Van Belkum A (2003) High-throughput epidemiologic typing in clinical microbiology. *Clin Microbiol Infect* 9(2):86–100
17. Supply P, Allix C, Lesjean S, Cardoso-Oelemann M, Rüsch-Gerdes S, Willery E et al (2006) Proposal for standardization of optimized mycobacterial interspersed repetitive unit-variable-number tandem repeat typing of *Mycobacterium tuberculosis*. *J Clin Microbiol* 44 (12):4498–4510
18. Brosch R, Gordon SV, Marmiesse M, Brodin P, Buchrieser C, Eiglmeier K et al (2002) A new evolutionary scenario for the *Mycobacterium tuberculosis* complex. *Proc Natl Acad Sci U S A* 99(6):3684–3689
19. Sreevatsan S, Pan X, Stockbauer KE, Connell ND, Kreiswirth BN, Whittam TS et al (1997) Restricted structural gene polymorphism in the *Mycobacterium tuberculosis* complex indicates evolutionarily recent global dissemination. *Proc Natl Acad Sci U S A* 94 (18):9869–9874

20. Gagneux S (2012) Host-pathogen coevolution in human tuberculosis. *Philos Trans R Soc Lond B Biol Sci* 367(1590):850–859
21. Fabre M, Hauck Y, Soler C, Koeck J-L, Van Ingen J, Van Soolingen D et al (2010) Molecular characteristics of “Mycobacterium canettii” the smooth Mycobacterium tuberculosis bacilli. *Infect Genet Evol* 10(8):1165–1173
22. Huard RC, Fabre M, De Haas P, Claudio Oliveira Lazzarini L, Van Soolingen D, Cousins D et al (2006) Novel genetic polymorphisms that further delineate the phylogeny of the Mycobacterium tuberculosis complex. *J Bacteriol* 188(12):4271–4287
23. Alexander KA, Laver PN, Michel AL, Williams M, Van Helden PD, Warren RM et al (2010) Novel Mycobacterium tuberculosis complex pathogen, *M. mungi*. *Emerg Infect Dis* 16 (8):1296–1299
24. Bentley SD, Comas I, Bryant JM, Walker D, Smith NH, Harris SR et al (2012) The genome of *Mycobacterium africanum* West African 2 reveals a lineage-specific locus and genome erosion common to the *M. tuberculosis* complex. *PLoS Negl Trop Dis* 6(2):e1552
25. Achtman M (2008) Evolution, population structure, and phylogeography of genetically monomorphic bacterial pathogens. *Annu Rev Microbiol* 62:53–70
26. Weniger T, Krawczyk J, Supply P, Harmsen D, Niemann S (2012) Online tools for polyphasic analysis of *Mycobacterium tuberculosis* complex genotyping data: now and next. *Infect Genet Evol* 12(4):748–754
27. Hershberg R, Lipatov M, Small PM, Sheffer H, Niemann S, Homolka S et al (2008) High functional diversity in *Mycobacterium tuberculosis* driven by genetic drift and human demography. *PLoS Biol* 6(12):e311
28. Kato-Maeda M, Metcalf JZ, Flores L (2011) Genotyping of *Mycobacterium tuberculosis*: application in epidemiologic studies. *Future Microbiol* 6(2):203–216
29. Van Soolingen D, Kremer K (2009) [Findings and ongoing research in the molecular epidemiology of tuberculosis]. *Kekkaku* 84(2):83–89
30. Schüürch AC, Van Soolingen D (2011) DNA fingerprinting of *Mycobacterium tuberculosis*: from phage typing to whole-genome sequencing. *Infect Genet Evol*. <http://www.ncbi.nlm.nih.gov/pubmed/22067515>. Accessed 20 Feb 2012
31. Burgos MV, Pym AS (2002) Molecular epidemiology of tuberculosis. *Eur Respir J* 20 (36 suppl):54s–65s
32. Mathema B, Kurepina NE, Bifani PJ, Kreiswirth BN (2006) Molecular epidemiology of tuberculosis: current insights. *Clin Microbiol Rev* 19(4):658–685
33. Fok A, Numata Y, Schulzer M, FitzGerald MJ (2008) Risk factors for clustering of tuberculosis cases: a systematic review of population-based molecular epidemiology studies. *Int J Tuberc Lung Dis* 12(5):480–492
34. van Soolingen D, Borgdorff MW, de Haas PEW, Sebek MMGG, Veen J, Dessens M et al (1999) Molecular epidemiology of tuberculosis in the Netherlands: A nationwide study from 1993 through 1997. *J Infect Dis* 180(3):726–736
35. Bifani PJ (1996) Origin and interstate spread of a New York City multidrug-resistant *Mycobacterium tuberculosis* clone family. *JAMA* 275(6):452–457
36. Glynn JR, Whiteley J, Bifani PJ, Kremer K, Van Soolingen D (2002) Worldwide occurrence of Beijing/W strains of *Mycobacterium tuberculosis*: a systematic review. *Emerg Infect Dis* 8 (8):843
37. Filliol I, Motiwala AS, Cavatore M, Qi W, Hazbón MH, Bobadilla del Valle M et al (2006) Global phylogeny of *Mycobacterium tuberculosis* based on single nucleotide polymorphism (SNP) analysis: insights into tuberculosis evolution, phylogenetic accuracy of other DNA fingerprinting systems, and recommendations for a minimal standard SNP set. *J Bacteriol* 188 (2):759
38. Hirsh AE, Tsolaki AG, DeRiemer K, Feldman MW, Small PM (2004) Stable association between strains of *Mycobacterium tuberculosis* and their human host populations. *PNAS* 101 (14):4871–4876

39. Van Soolingen D, De Haas PE, Hermans PW, Groenen PM, Van Embden JD (1993) Comparison of various repetitive DNA elements as genetic markers for strain differentiation and epidemiology of *Mycobacterium tuberculosis*. *J Clin Microbiol* 31(8):1987–1995
40. Van Embden JD, Van Soolingen D, Small PM, Hermans PW (1992) Genetic markers for the epidemiology of tuberculosis. *Res Microbiol* 143(4):385–391
41. Kremer K, Van Soolingen D, Frothingham R, Haas WH, Hermans PW, Martín C et al (1999) Comparison of methods based on different molecular epidemiological markers for typing of *Mycobacterium tuberculosis* complex strains: interlaboratory study of discriminatory power and reproducibility. *J Clin Microbiol* 37(8):2607–2618
42. Small PM, Hopewell PC, Singh SP, Paz A, Parsonnet J, Ruston DC et al (1994) The epidemiology of tuberculosis in San Francisco—a population-based study using conventional and molecular methods. *N Engl J Med* 330(24):1703–1709
43. Caminero JA, Pena MJ, Campos-Herrero MI, Rodríguez JC, García I, Cabrera P et al (2001) Epidemiological evidence of the spread of a *Mycobacterium tuberculosis* strain of the Beijing genotype on Gran Canaria Island. *Am J Respir Crit Care Med* 164(7):1165–1170
44. de Beer JL, Kremer K, Ködmön C, Supply P, van Soolingen D (2012) First worldwide proficiency study on variable-number tandem-repeat typing of *Mycobacterium tuberculosis* complex strains. *J Clin Microbiol* 50(3):662–669
45. Van Soolingen D (2001) Molecular epidemiology of tuberculosis and other mycobacterial infections: main methodologies and achievements. *J Intern Med* 249(1):1–26
46. Kamerbeek J, Schouls L, Kolk A, Van Agterveld M, Van Soolingen D, Kuijper S et al (1997) Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. *J Clin Microbiol* 35(4):907–914
47. Goguet de la Salmonière YO, Li HM, Torrea G, Bunschoten A, Van Embden J, Gicquel B (1997) Evaluation of spoligotyping in a study of the transmission of *Mycobacterium tuberculosis*. *J Clin Microbiol* 35(9):2210–2214
48. Molhuizen HO, Bunschoten AE, Schouls LM, Van Embden JD (1998) Rapid detection and simultaneous strain differentiation of *Mycobacterium tuberculosis* complex bacteria by spoligotyping. *Methods Mol Biol* 101:381–394
49. Brudey K, Driscoll JR, Rigouts L, Prodinger WM, Gori A, Al-Hajj SA et al (2006) *Mycobacterium tuberculosis* complex genetic diversity: mining the fourth international spoligotyping database (SpolDB4) for classification, population genetics and epidemiology. *BMC Microbiol* 6:23
50. Tsolaki AG, Gagneux S, Pym AS, Goguet de la Salmoniere Y-OL, Kreiswirth BN, Van Soolingen D et al (2005) Genomic deletions classify the Beijing/W strains as a distinct genetic lineage of *Mycobacterium tuberculosis*. *J Clin Microbiol* 43(7):3185–3191
51. Skuce RA, McCorry TP, McCarroll JF, Roring SMM, Scott AN, Brittain D et al (2002) Discrimination of *Mycobacterium tuberculosis* complex bacteria using novel VNTR-PCR targets. *Microbiology* 148(Pt 2):519–528
52. Allix-Béguec C, Harmsen D, Weniger T, Supply P, Niemann S (2008) Evaluation and strategy for use of MIRU-VNTRplus, a multifunctional database for online analysis of genotyping data and phylogenetic identification of *Mycobacterium tuberculosis* complex isolates. *J Clin Microbiol* 46(8):2692–2699
53. Mazars E, Lesjean S, Banuls AL, Gilbert M, Vincent V, Gicquel B et al (2001) High-resolution minisatellite-based typing as a portable approach to global analysis of *Mycobacterium tuberculosis* molecular epidemiology. *Proc Natl Acad Sci* 98(4):1901
54. Supply P, Lesjean S, Savine E, Kremer K, Van Soolingen D, Locht C (2001) Automated high-throughput genotyping for study of global epidemiology of *Mycobacterium tuberculosis* based on mycobacterial interspersed repetitive units. *J Clin Microbiol* 39(10):3563–3571
55. Tsolaki AG, Hirsh AE, DeRiemer K, Enciso JA, Wong MZ, Hannan M et al (2004) Functional and evolutionary genomics of *Mycobacterium tuberculosis*: insights from genomic deletions in 100 strains. *PNAS* 101(14):4865–4870

56. Gutacker MM, Smoot JC, Migliaccio CAL, Ricklefs SM, Hua S, Cousins DV et al (2002) Genome-wide analysis of synonymous single nucleotide polymorphisms in *Mycobacterium tuberculosis* complex organisms: resolution of genetic relationships among closely related microbial STRAINS. *Genetics* 162(4):1533–1543
57. Supply P, Warren RM, Bañuls A-L, Lesjean S, Van Der Spuy GD, Lewis L-A et al (2003) Linkage disequilibrium between minisatellite loci supports clonal evolution of *Mycobacterium tuberculosis* in a high tuberculosis incidence area. *Mol Microbiol* 47(2):529–538
58. Schürch AC, Van Soolingen D (2012) DNA fingerprinting of *Mycobacterium tuberculosis*: from phage typing to whole-genome sequencing. *Infect Genet Evol* 12(4):602–609
59. Allard D, Whittam TS, Murray MB, Cave MD, Hazbon MH, Dix K et al (2003) Modeling bacterial evolution with comparative-genome-based marker systems: application to *Mycobacterium tuberculosis* evolution and pathogenesis. *J Bacteriol* 185(11):3392
60. Ford C, Yusim K, Ioerger T, Feng S, Chase M, Greene M et al (2012) *Mycobacterium tuberculosis*–heterogeneity revealed through whole genome sequencing. *Tuberculosis (Edinb)* 92(3):194–201
61. Niemann S, Köser CU, Gagneux S, Plinke C, Homolka S, Bignell H et al (2009) Genomic diversity among drug sensitive and multidrug resistant isolates of *Mycobacterium tuberculosis* with identical DNA fingerprints. *PLoS ONE* 4(10):e7407
62. Saunders NJ, Trivedi UH, Thomson ML, Doig C, Laurenson IF, Blaxter ML (2011) Deep resequencing of serial sputum isolates of *Mycobacterium tuberculosis* during therapeutic failure due to poor compliance reveals stepwise mutation of key resistance genes on an otherwise stable genetic background. *J Infect* 62(3):212–217
63. Boshoff HIM, Reed MB, Barry CE III, Mizrahi V (2003) DnaE2 polymerase contributes to in vivo survival and the emergence of drug resistance in *Mycobacterium tuberculosis*. *Cell* 113(2):183–193
64. Dos Vultos T, Mestre O, Tonjum T, Gicquel B (2009) DNA repair in *Mycobacterium tuberculosis* revisited. *FEMS Microbiol Rev* 33(3):471–487
65. Gardy JL, Johnston JC, Sui SJH, Cook VJ, Shah L, Brodkin E et al (2011) Whole-genome sequencing and social-network analysis of a tuberculosis outbreak. *N Engl J Med* 364 (8):730–739
66. Chen Y-Y, Chang J-R, Huang W-F, Kuo S-C, Su I-J, Sun J-R et al (2012) Genetic diversity of the *Mycobacterium tuberculosis* Beijing family based on SNP and VNTR typing profiles in Asian countries. *PLoS ONE* 7(7):e39792
67. van Deutekom H, Gerritsen JJJ, van Soolingen D, van Ameijden EJC, van Embden JDA, Coutinho RA (1997) A molecular epidemiological approach to studying the transmission of tuberculosis in Amsterdam. *Clin Infect Dis* 25(5):1071–1077
68. Daley CL, Small PM, Schecter GF, Schoolnik GK, McAdam RA, Jacobs WR Jr et al (1992) An outbreak of tuberculosis with accelerated progression among persons infected with the human immunodeficiency virus. An analysis using restriction-fragment-length polymorphisms. *N Engl J Med* 326(4):231–235
69. Jasmer RM, Hahn JA, Small PM, Daley CL, Behr MA, Moss AR et al (1999) A molecular epidemiologic analysis of tuberculosis trends in San Francisco, 1991–1997. *Ann Intern Med* 130(12):971–978
70. Goebel T (2007) The missing years for modern humans. *Science* 315(5809):194–196
71. Gutierrez MC, Brisson S, Brosch R, Fabre M, Omaïs B, Marmiesse M et al. (2005) Ancient origin and gene mosaicism of the progenitor of *Mycobacterium tuberculosis*. *PLoS Pathog* 1 (1). <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1238740/>. Accessed 24 Sep 2012
72. Kappelman J, Alçıçek MC, Kazancı N, Schultz M, Ozkul M, Sen S (2008) First *Homo erectus* from Turkey and implications for migrations into temperate Eurasia. *Am J Phys Anthropol* 135(1):110–116
73. Barnes E (2007) Diseases and human evolution. UNM Press, New Mexico

74. Linz B, Balloux F, Moodley Y, Manica A, Liu H, Roumagnac P et al (2007) An African origin for the intimate association between humans and *Helicobacter pylori*. *Nature* 445 (7130):915–918
75. Gagneux S, DeRiemer K, Van T, Kato-Maeda M, de Jong BC, Narayanan S et al (2006) Variable host-pathogen compatibility in *Mycobacterium tuberculosis*. *PNAS* 103 (8):2869–2873
76. Hewinson RG, Vordermeier HM, Smith NH, Gordon SV (2006) Recent advances in our knowledge of *Mycobacterium bovis*: a feeling for the organism. *Vet Microbiol* 112 (2–4):127–139
77. Mostowy S, Inwald J, Gordon S, Martin C, Warren R, Kremer K et al (2005) Revisiting the evolution of *Mycobacterium bovis*. *J Bacteriol* 187(18):6386–6395
78. Garnier T, Eiglmeier K, Camus J-C, Medina N, Mansoor H, Pryor M et al (2003) The complete genome sequence of *Mycobacterium bovis*. *Proc Natl Acad Sci U S A* 100 (13):7877–7882
79. Brosch R, Pym AS, Gordon SV, Cole ST (2001) The evolution of mycobacterial pathogenicity: clues from comparative genomics. *Trends Microbiol* 9(9):452–458
80. Djelouadji Z, Raoult D, Drancourt M (2011) Palaeogenomics of *Mycobacterium tuberculosis*: epidemic bursts with a degrading genome. *Lancet Infect Dis* 11(8):641–650
81. Smith NH (2012) The global distribution and phylogeography of *Mycobacterium bovis* clonal complexes. *Infect Genet Evol* 12(4):857–865
82. Gagneux S, Small PM (2007) Global phylogeography of *Mycobacterium tuberculosis* and implications for tuberculosis product development. *Lancet Infect Dis* 7(5):328–337
83. Gutacker MM, Mathema B, Soini H, Shashkina E, Kreiswirth BN, Graviss EA et al (2006) Single-nucleotide polymorphism-based population genetic analysis of *Mycobacterium tuberculosis* strains from 4 geographic sites. *J Infect Dis* 193(1):121–128
84. Wirth T, Hildebrand F, Allix-Béguec C, Wölbeling F, Kubica T, Kremer K et al (2008) Origin, spread and demography of the *Mycobacterium tuberculosis* complex. *PLoS Pathog* 4 (9):e1000160
85. Coscolla M, Gagneux S (2010) Does *M. tuberculosis* genomic diversity explain disease diversity? *Drug Discov Today Dis Mech* 7(1):e43–e59
86. Reed MB, Gagneux S, Deriemer K, Small PM, Barry CE (2007) The W-Beijing lineage of *Mycobacterium tuberculosis* overproduces triglycerides and has the DosR dormancy regulon constitutively upregulated. *J Bacteriol* 189(7):2583–2589
87. Comas I, Gagneux S (2011) A role for systems epidemiology in tuberculosis research. *Trends Microbiol* 19(10):492–500
88. Portevin D, Gagneux S, Comas I, Young D (2011) Human macrophage responses to clinical isolates from the *Mycobacterium tuberculosis* complex discriminate between ancient and modern lineages. *PLoS Pathog* 7(3):e1001307
89. De Jong BC, Hill PC, Aiken A, Awine T, Antonio M, Adetifa IM et al (2008) Progression to active tuberculosis, but not transmission, varies by *Mycobacterium tuberculosis* lineage in The Gambia. *J Infect Dis* 198(7):1037–1043
90. Van Soolingen D, Qian L, De Haas PE, Douglas JT, Traore H, Portaels F et al (1995) Predominance of a single genotype of *Mycobacterium tuberculosis* in countries of East Asia. *J Clin Microbiol* 33(12):3234–3238
91. Anh DD, Borgdorff MW, Van LN, Lan NT, Van Gorkom T, Kremer K et al (2000) *Mycobacterium tuberculosis* Beijing genotype emerging in Vietnam. *Emerg Infect Dis* 6 (3):302–305
92. Chan MY, Borgdorff M, Yip CW, De Haas PEW, Wong WS, Kam KM et al. (2001) Seventy percent of the *Mycobacterium tuberculosis* isolates in Hong Kong represent the Beijing genotype. *Epidemiol Infect* 127(01). [http://www.journals.cambridge.org/abstract\\_S0950268801005659](http://www.journals.cambridge.org/abstract_S0950268801005659). Accessed 30 Oct 2012
93. Mokrousov I (2008) Genetic geography of *Mycobacterium tuberculosis* Beijing genotype: a multifacet mirror of human history? *Infect Genet Evol* 8(6):777–785

94. Cowley D, Govender D, February B, Wolfe M, Steyn L, Evans J et al (2008) Recent and rapid emergence of W-Beijing strains of *Mycobacterium tuberculosis* in Cape Town, South Africa. *Clin Infect Dis* 47(10):1252–1259
95. Parwati I, Van Crevel R, Van Soolingen D (2010) Possible underlying mechanisms for successful emergence of the *Mycobacterium tuberculosis* Beijing genotype strains. *Lancet Infect Dis* 10(2):103–111
96. Schürch AC, Kremer K, Hendriks ACA, Freyee B, McEvoy CRE, Van Crevel R et al (2011) SNP/RD typing of *Mycobacterium tuberculosis* Beijing strains reveals local and worldwide disseminated clonal complexes. *PLoS ONE* 6(12):e28365
97. Röltgen K, Qi W, Ruf M-T, Mensah-Quainoo E, Pidot SJ, Seemann T et al (2010) Single nucleotide polymorphism typing of *Mycobacterium ulcerans* reveals focal transmission of buruli ulcer in a highly endemic region of Ghana. *PLoS Negl Trop Dis* 4(7):e751
98. Morelli G, Song Y, Mazzoni CJ, Eppinger M, Roumagnac P, Wagner DM et al (2010) *Yersinia pestis* genome sequencing identifies patterns of global phylogenetic diversity. *Nat Genet* 42(12):1140–1143
99. Monot M, Honoré N, Garnier T, Zidane N, Sherifi D, Paniz-Mondolfi A et al (2009) Comparative genomic and phylogeographic analysis of *Mycobacterium leprae*. *Nat Genet* 41(12):1282–1289
100. De Jong BC, Antonio M, Gagneux S (2010) *Mycobacterium africanum*—review of an important cause of human tuberculosis in West Africa. *PLoS Negl Trop Dis* 4(9):e744
101. Mostowy S, Cousins D, Brinkman J, Aranaz A, Behr MA (2002) Genomic deletions suggest a phylogeny for the *Mycobacterium tuberculosis* complex. *J Infect Dis* 186(1):74–80
102. Intemann CD, Thye T, Niemann S, Browne ENL, Amanu Chinbuah M, Enimil A et al. (2009) Autophagy gene variant IRGM – 261T contributes to protection from tuberculosis caused by *Mycobacterium tuberculosis* but not by *M. africanum* strains. *PLoS Pathog* 5 (9). <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2735778/>. Accessed 17 Oct 2012
103. Jamison DTDT, Breman JGJG, Measham ARAR, Alleyne GG, Claeson MM, Evans DBDB et al (eds) (2006) Disease control priorities in developing countries, 2nd edn. World Bank, Washington (DC)
104. Velayati AA, Masjedi MR, Farnia P, Tabarsi P, Ghanavi J, ZiaZarifi AH et al (2009) Emergence of new forms of totally drug-resistant tuberculosis bacilli super extensively drug-resistant tuberculosis or totally drug-resistant strains in Iran. *Chest* 136(2):420–425
105. Udwadia ZF, Amale RA, Ajbani KK, Rodrigues C (2012) Totally drug-resistant tuberculosis in India. *Clin Infect Dis* 54(4):579–581
106. Burgos M, Gonzalez LC, Paz EA, Gournis E, Kawamura LM, Schechter G et al (2005) Treatment of multidrug-resistant tuberculosis in San Francisco: an outpatient-based approach. *Clin Infect Dis* 40(7):968–975
107. Zignol M, van Gemert W, Falzon D, Sismanidis C, Glaziou P, Floyd K, Raviglione M (2012) Surveillance of anti-tuberculosis drug resistance in the world: an updated analysis, 2007–2010. *Bull WHO* 90:111
108. Pablos-Méndez A, Raviglione MC, Laszlo A, Binkin N, Rieder HL, Bustreo F et al (1998) Global surveillance for antituberculosis-drug resistance, 1994–1997. *N Engl J Med* 338 (23):1641–1649
109. Gandhi NR, Moll A, Sturm AW, Pawinska R, Govender T, Laloo U et al (2006) Extensively drug-resistant tuberculosis as a cause of death in patients co-infected with tuberculosis and HIV in a rural area of South Africa. *Lancet* 368(9547):1575–1580
110. Dalton T, Cegielski P, Akksilp S, Asencios L, Caoili JC, Cho S-N et al (2012) Prevalence of and risk factors for resistance to second-line drugs in people with multidrug-resistant tuberculosis in eight countries: a prospective cohort study. *Lancet*;380(9851):1406–17
111. Shah NPR (2008) Extensively drug-resistant tuberculosis in the united states, 1993–2007. *JAMA* 300(18):2153–2160
112. McIlheron H, Rustomjee R, Vahedi M, Mthiyane T, Denti P, Connolly C et al (2012) Reduced antituberculosis drug concentrations in HIV-infected patients who are men or have low

- weight: implications for international dosing guidelines. *Antimicrob Agents Chemother* 56(6):3232–3238
113. Mukinda FK, Theron D, Van der Spuy GD, Jacobson KR, Roscher M, Streicher EM et al (2012) Rise in rifampicin-monoresistant tuberculosis in Western Cape, South Africa. *Int J Tuberc Lung Dis* 16(2):196–202
  114. WHO (2010) Multidrug and extensively drug-resistant TB (M/XDR-TB). World Health Organization, Geneva
  115. WHO (2012) Surveillance of drug resistance in tuberculosis. WHO. [http://www.who.int/tb/publications/mdr\\_surveillance/en/index.html](http://www.who.int/tb/publications/mdr_surveillance/en/index.html). Accessed 21 Dec 2012
  116. Gandhi NR, Weissman D, Moodley P, Ramathal M, Elson I, Kreiswirth BN et al (2013) Nosocomial transmission of extensively drug-resistant tuberculosis in a rural hospital in South Africa. *J Infect Dis* 207(1):9–17
  117. Ioerger TR, Koo S, No E-G, Chen X, Larsen MH, Jacobs WR et al (2009) Genome analysis of multi- and extensively-drug-resistant tuberculosis from KwaZulu-Natal, South Africa. *PLoS ONE* 4(11):e7778
  118. McElroy PD, Sterling TR, Driver CR, Kreiswirth B, Woodley CL, Cronin WA et al (2002) Use of DNA fingerprinting to investigate a multiyear, multistate tuberculosis outbreak. *Emerg Infect Dis* 8(11):1252–1256
  119. Moss AR, Alland D, Telzak E, Hewlett D Jr, Sharp V, Chiliade P et al (1997) A city-wide outbreak of a multiple-drug-resistant strain of *Mycobacterium tuberculosis* in New York. *Int J Tuberc Lung Dis* 1(2):115–121
  120. Frieden TR, Sherman LF, Maw KL, Fujiwara PI, Crawford JT, Nivin B et al (1996) A multi-institutional outbreak of highly drug-resistant tuberculosis: epidemiology and clinical outcomes. *JAMA* 276(15):1229–1235
  121. Frieden TR, Fujiwara PI, Washko RM, Hamburg MA (1995) Tuberculosis in New York City—turning the tide. *N Engl J Med* 333(4):229–233
  122. Stroud LA, Tokars JI, Grieco MH, Crawford JT, Culver DH, Edlin BR et al (1995) Evaluation of infection control measures in preventing the nosocomial transmission of multidrug-resistant *Mycobacterium tuberculosis* in a New York City hospital. *Infect Control Hosp Epidemiol* 16(3):141–147
  123. Lawn SD, Wood R, De Cock KM, Kranzer K, Lewis JJ, Churchyard GJ (2010) Antiretrovirals and isoniazid preventive therapy in the prevention of HIV-associated tuberculosis in settings with limited health-care resources. *Lancet Infect Dis* 10(7):489–498
  124. Akolo C, Adetifa I, Shepperd S, Volmink J (2010) Treatment of latent tuberculosis infection in HIV infected persons. *Cochrane Database Syst Rev* (1):CD000171
  125. Borrell S, Gagneux S (2009) Infectiousness, reproductive fitness and evolution of drug-resistant *Mycobacterium tuberculosis*. *Int J Tuberc Lung Dis* 13(12):1456–1466
  126. Dye C, Watt CJ, Bleed DM, Hosseini SM, Raviglione MC (2005) Evolution of tuberculosis control and prospects for reducing tuberculosis incidence, prevalence, and deaths globally. *JAMA* 293(22):2767
  127. Dye C, Williams BG (2010) The population dynamics and control of tuberculosis. *Science* 328(5980):856–861
  128. Dye C, Williams BG, Espinal MA, Raviglione MC (2002) Erasing the world's slow stain: strategies to beat multidrug-resistant tuberculosis. *Science* 295(5562):2042–2046
  129. Cohen T, Murray M (2004) Modeling epidemics of multidrug-resistant *M. tuberculosis* of heterogeneous fitness. *Nat Med* 10(10):1117–1121
  130. Borrell S, Gagneux S (2011) Strain diversity, epistasis and the evolution of drug resistance in *Mycobacterium tuberculosis*. *Clin Microbiol Infect* 17(6):815–820
  131. Gagneux S, Burgos MV, DeRiemer K, Encisco A, Muñoz S, Hopewell PC et al (2006) Impact of bacterial genetics on the transmission of isoniazid-resistant *Mycobacterium tuberculosis*. *PLoS Pathog* 2(6):e61
  132. Comas I, Borrell S, Roetzer A, Rose G, Malla B, Kato-Maeda M et al (2011) Whole-genome sequencing of rifampicin-resistant *M. tuberculosis* strains identifies compensatory mutations in RNA polymerase. *Nat Genet* 44(1):106–110

133. Zhang Y, Heym B, Allen B, Young D, Cole S (1992) The catalase-peroxidase gene and isoniazid resistance of *Mycobacterium tuberculosis*. *Nature* 358(6387):591–593
134. Timmins GS, Deretic V (2006) Mechanisms of action of isoniazid. *Mol Microbiol* 62 (5):1220–1227
135. Maisnier-Patin S, Andersson DI (2004) Adaptation to the deleterious effects of antimicrobial drug resistance mutations by compensatory evolution. *Res Microbiol* 155(5):360–369
136. Sherman DR, Mdluli K, Hickey MJ, Arain TM, Morris SL, Barry CE III et al (1996) Compensatory *ahpC* Gene expression in isoniazid-resistant *Mycobacterium tuberculosis*. *Science* 272(5268):1641–1643
137. Rouse DA, DeVito JA, Li Z, Byer H, Morris SL (1996) Site-directed mutagenesis of the *katG* gene of *Mycobacterium tuberculosis*-, effects on catalase- peroxidase activities and isoniazid resistance. *Mol Microbiol* 22(3):583–592
138. Pym AS, Domenech P, Honoré N, Song J, Deretic V, Cole ST (2001) Regulation of catalase-peroxidase (KatG) expression, isoniazid sensitivity and virulence by *furA* of *Mycobacterium tuberculosis*. *Mol Microbiol* 40(4):879–889
139. Wilson T, de Lisle GW, Marcinkeviciene JA, Blanchard JS, Collins DM (1998) Antisense RNA to *ahpC*, an oxidative stress defence gene involved in isoniazid resistance, indicates that *AhpC* of *Mycobacterium bovis* has virulence properties. *Microbiology* 144(10):2687–2695
140. Heym B, Stavropoulos E, Honoré N, Domenech P, Saint-Jeanis B, Wilson TM et al (1997) Effects of over expression of the alkyl hydroperoxide reductase *AhpC* on the virulence and isoniazid resistance of *Mycobacterium tuberculosis*. *Infect Immun* 65(4):1395–1401
141. Gagneux S, Long CD, Small PM, Van T, Schoolnik GK, Bohannan BJM (2006) The competitive cost of antibiotic resistance in *Mycobacterium tuberculosis*. *Science* 312 (5782):1944–1946
142. Brandis G, Wrande M, Liljas L, Hughes D (2012) Fitness-compensatory mutations in rifampicin-resistant RNA polymerase. *Mol Microbiol* 85(1):142–151
143. De Vos M, Muller B, Borrell S et al (2012) Putative compensatory mutations in the *rpoC* gene of rifampin-resistant *mycobacterium tuberculosis* are associated with ongoing transmission. *Antimicrob Agents Chemother* 57(2):827–832. doi:[10.1128/AAC.01541-12](https://doi.org/10.1128/AAC.01541-12)

## Chapter 6

# Trends in HIV Transmission According to Differences in Numbers of Sexual Partnerships Among Men Who Have Sex with Men in China

Lei Zhang, Eric P.F. Chow, and David P. Wilson

## 6.1 Background

### 6.1.1 HIV Epidemic in China

The development of the HIV epidemic in China can be divided into four distinct phases [1–3]. The first phase (1985–1988) was marked by China's first diagnosis of an AIDS case in Beijing in 1985, followed by 22 cases in seven other provinces, including 18 foreigners and 4 Chinese residents, and all were infected by HIV-contaminated blood products [4, 5]. During the second phase (1989–1993), 146 HIV/AIDS cases were diagnosed among injecting drug users (IDUs) in Yunnan. The first HIV case among MSM (men who have sex with men) case was also diagnosed in Beijing in 1989, but HIV transmission among MSM remained low [6–8]. The third phase (1994–2000) covered the large expansion in the numbers of infections among IDU and the epidemic started to spread from Yunnan to other Chinese provinces. By the mid-1990s it was clear that the epidemic had severely afflicted the population of IDUs, and by the late 1990s HIV infections had been reported in all 31 provinces, autonomous regions, and municipalities, with drug users accounting for 60–70 % of reported infections [9]. The fourth phase (2001–present) was marked by a series of events and interventions that responded to HIV on a national level [1, 2]. China released two consecutive 5-year action plans to limit the spread of HIV/AIDS, in 2001 and 2006, respectively [10, 11]. Several national prevention and treatment programs such as “China CARES (China Comprehensive AIDS Response)” and “Four Frees, One Care” policy were established in response to the rapid emerging HIV epidemic in China.

By the end of 2009, it was estimated that 740,000 (0.56–0.92 million) people in China were living with HIV [12], out of a total population of 1.3 billion. In contrast

---

L. Zhang (✉) • E.P.F. Chow • D.P. Wilson

The Kirby Institute, The University of New South Wales, Sydney, NSW 2052, Australia  
e-mail: [lizhang@kirby.unsw.edu.au](mailto:lizhang@kirby.unsw.edu.au)

to other regions around the world, China has a relatively low HIV prevalence of 0.057 % (0.042–0.071 %) among its general population [2, 12]. However, the HIV epidemics in China was rapidly expanding within specific at-risk groups [12, 13] with large geographical variations in mode of transmission. This followed the typical pattern of HIV epidemics in Asia whereby HIV first spread among IDU and followed by transmission among sex workers [14]. However, sexual transmission, especially homosexual transmission, has become the primary mode of transmission in recent years [12, 15]. The latest statistics revealed that the proportion of new HIV annual infections due to male-to-male sexual intercourse has increased from 12 % in 2007 to 33 % in 2009 [16]. The estimated national HIV prevalence among MSM has tripled, from 1.4 to 5.3 %, during 2001–2009 [17]. However, the HIV epidemic among MSM varies across different regions of China [17]; it has been shown that Southwest region has disproportionately high HIV disease burden (i.e., 10.9 % prevalence rate in Chongqing, 11.1 % in Sichuan, 11.8 % in Kunming [18, 19]) compared with other parts of China. In addition, fast-increasing temporal trends were observed in all regions of China, such as Beijing (from 1.5 % in 2004 to 5.9 % in 2008) [18, 20–22], Guangdong (from 1.8 % in 2003 to 5.2 % in 2008) [18, 23], Liaoning (from 1.0 % in 2003 to 8.7 % in 2009) [18, 24, 25], Shanghai (from 1.5 % in 2005 to 7.5 % in 2008) [18, 26] and Zhejiang (from 0.0 % in 2006 to 5.1 % in 2008) [18, 27]. Accumulating evidence indicates that MSM in China, as in most settings, have a much higher chance to be infected with HIV than other population groups. A recent meta-analysis study showed that the odds of being infected with HIV among MSM was 19.3 (95 % CI, 18.8–19.8) times higher than that among the general population in low- and middle-income countries [28]; in particular, the corresponding ratio is 45.1 (39.6–51.4) in China [28]. The rapid transmission of HIV and high odds ratios among Chinese MSM were likely associated with their high-risk sexual behaviors.

### **6.1.2 *MSM Sexual Behaviors***

Currently, there are approximately 5–10 million homosexual men in China [29] and of which approximately 95 % had male sexual partner in past 6 months [25]. In China, sexual partners among MSM can be categorized into three main types: (1) regular; (2) noncommercial casual and (3) commercial. Regular partners include those who have a stable relationship and do not involve any exchange of money and sex (i.e., boyfriends). Noncommercial casual partners refer to MSM partners who engage in a brief and temporary relationship and have no exchange of sex for money (i.e., “419”—a codeword for “for one night”). Commercial partners are MSM who sell or buy sex from another male. It is well known that multiple sexual partners can increase the risk of HIV infection, but high-risk sexual behaviors such as unprotected penile–anal sexual intercourse can be a major risk factor for HIV transmission [30]. It has been reported that in the United States unprotected sexual intercourse among regular MSM partnerships contribute to the majority of new HIV

infections [31]. On contrast, little is known about how sexual intercourse among different sexual partnerships among MSM had contributed to new HIV incidence cases in China. Besides, previous studies showed that consistent and proper use of condoms may reduce HIV transmission by approximately 85–95 % [32–36]. In China, temporal increasing trends in consistent condom use among MSM have been reported [37]. However, whether similar increasing trends in condom usage among each specific type of sexual partnership among MSM remains unclear. Therefore, an investigation and estimate of HIV incidence among MSM, taking into consideration the pattern of sexual behaviors and condom usage in different partnerships could provide important insights into the routes and trends of HIV transmission among Chinese MSM.

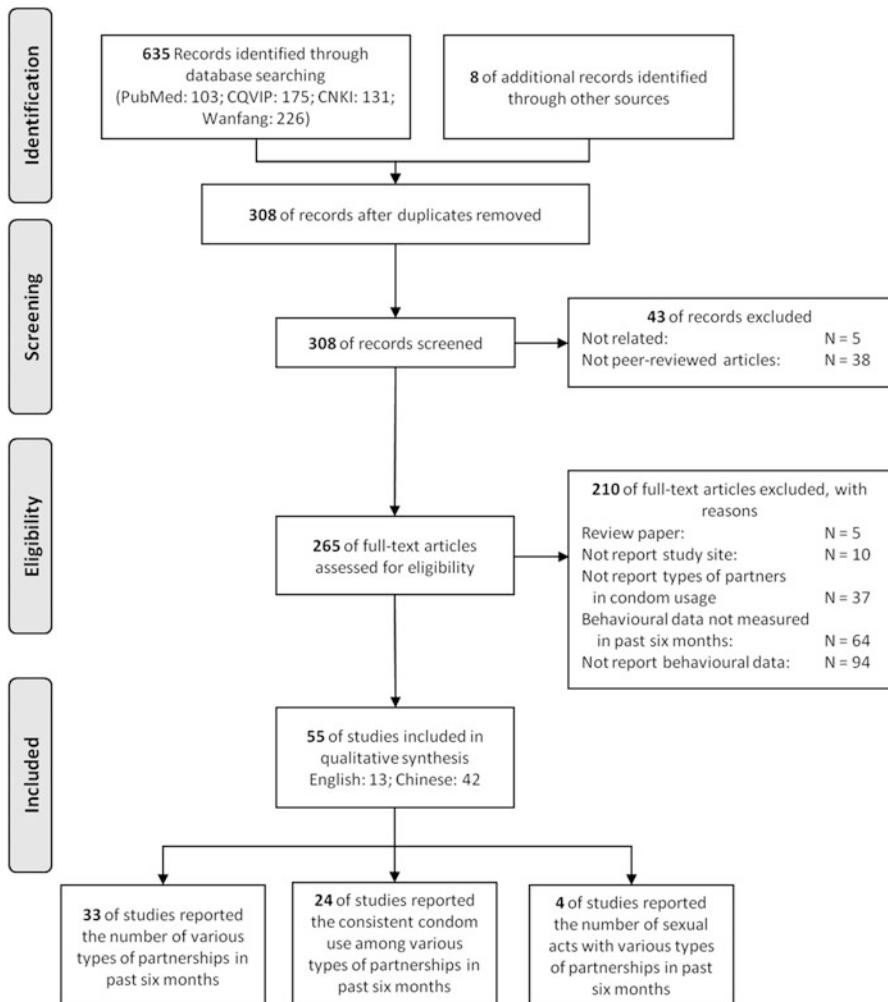
Studies specifically reporting sexual behaviors and condom usage among Chinese MSM according to their partnership types are scattered. Through a comprehensive review on published literature, we collated all available behavioral data from these studies to describe the overall behavioral patterns of Chinese MSM and temporal trends. We further developed a mathematical optimization routine to reconcile the distributions of sexual behavioral data over the last decade and infer the heterogeneous distributions of behavioral patterns that exist among the population of Chinese MSM. Based on these distributions and a transmission risk equation, HIV incidence rates associated with each type of partnership were estimated. We then further discussed the risk of HIV transmission among each sexual partnership type and its implications to HIV control and prevention among Chinese MSM.

## 6.2 Method

### 6.2.1 *Systematic Review of Published Literature*

A systematic literature search for published peer-reviewed studies was performed by two independent investigators (EPFC, LZ) from both English and Chinese electronic databases, including English: PubMed, Chinese: Wanfang Data, China National Knowledge Infrastructure (CNKI), and Chinese Scientific Journals Full text Database (CQVIP), from 2001 to May 2011. The keywords searched in the databases included (“human immunodeficiency virus” OR “HIV” OR “Acquired immune deficiency syndrome” OR “AIDS”) OR (“homosexual” OR “MSM” OR “men who have sex with men” OR “gay”) OR (“behaviors” OR “attitudes” OR “high-risk behaviors” OR “condom” OR “partners”) AND (“China”). The references in the relevant articles were searched manually. We limited our search to articles published in Chinese and English languages only.

We applied the following inclusion and exclusion criteria. Studies were included if they: (1) reported the percentage of consistent condom use in past 6 months with any types of male partners (i.e., regular, noncommercial casual, or commercial); or



**Fig. 6.1** Flow chart for selection of studies with number of articles (N)

(2) reported the number of different types of sexual partners in the past 6 months (i.e., regular, noncommercial casual, commercial, or any); or (3) reported the number of sexual acts in the past 6 months with different types of male partners (i.e., regular, noncommercial casual, commercial, or any); through a peer-reviewed publication of cohort or cross-sectional study. We excluded studies in which the target study population is specifically money boys or MSM drug users. Review papers, non-peer-reviewed theses, local reports, conference abstracts, and presentations were excluded from this study. Studies which did not report study location, time periods, and sample size were also excluded. If the study was duplicated in the databases, studies published in Chinese or published latest were excluded from this review. The procedure of literature filtering was shown as a flowchart in Fig. 6.1.

For all eligible studies, we extracted the following information about the study: first author and year of publication, study period, study location, age of MSM, recruitment method, sampling method, sample size, and types of male sexual partners (regular, noncommercial casual, commercial, or any) (Tables 6.1, 6.2, and 6.3). However, we included studies only with at least 2 data points on the distribution of number of partners or acts.

### ***6.2.2 Fittings and Optimization of MSM Sexual Behaviors***

The optimization methodology has been previously published [38]. Briefly, log-normal distributions represent probability distributions of a random variable whose logarithmic values are normally distributed and have been accepted as a good approximation to the distribution of sexual partners ( $N_i \sim LN(\mu_i, \delta_i)$ ) and acts ( $S_i \sim LN(\mu_i, \delta_i)$ ) of human populations [39]. It has been widely accepted and applied in numerous studies [38, 40–42]. We hence fitted a cumulative lognormal distribution function to the cumulative numbers of sexual partners (or acts) for each distribution dataset from collated studies. For each fitted distribution, mean and median and their 95 % confidence intervals were calculated (Tables 6.1 and 6.2). We investigated the temporal trend in numbers of sexual partners and acts of MSM during 2002–2010 using the Spearman nonparametric correlation test on the calculated mean (and median) and the study year ( $p < 0.05$  indicates statistically significant trend).

Since behavioral data were collected from studies implemented in different geographical location and time with various study methods, heterogeneity is expected to be substantial. We therefore employed an optimization approach to reconcile the available sexual behavioral data; that is, the distribution of the number of sexual partners and acts. Four types of sexual partnerships were investigated in the published literature, namely, regular, noncommercial casual, commercial, and overall partnerships in the past 6 months. We made the following assumptions: (1) if a study does not specify any sexual type, the reported distribution is assumed to represent the overall distribution; (2) the distribution of the overall number of sexual partners (or acts) is the sum of the corresponding lognormal distributions from the separate types of sexual partners, which are also approximated by lognormal distributions. Notably, this approximation of additive distributions enables us to estimate the distribution of the number of sexual partners (or acts) in any type of partnerships even in scenarios where empirical data is completely absent (Fig. 6.2), as long as the overall partners (or acts) distributions are available. The lognormal distribution assumption is also important for the subsequent optimization exercise, where it fits empirical data to the distribution without arbitrarily imputing missing data; (3) noncommercial casual partnerships are commonly “one-off,” suggesting similar distributions of numbers of casual sexual partners and acts. In such cases, large uncertainties for absent distribution are included in the optimization

**Table 6.1** Studies reporting number of male sexual partners among men who have sex with men in China

First author, published year	Study design							Number of different sexual partners in past 6 months mean/median		
		Study period	Study location	Region	Age range (mean)	Method of recruitment	Method of sampling	Regular partners	Noncommercial casual partners	Commercial partners
Choi KH, 2004 [65]	09/2001-01/2002	Beijing	North	18-69 (27)	MSM-identified bars, clubs, parks, and bathhouses	Convenience sampling	N/A	N/A	N/A	Mean: 3.91 Median: 2.69
Mao M, 2008 [66]	2003	Chengdu & Nanchong, Sichuan	Southwest	16-76 (28.1)	Gay bars, hotlines, Web sites, parks	N/A	N/A	N/A	N/A	Mean: 5.36 Median: 2.31
He Q, 2005 [67]	04/2003-05/2003	Guangzhou, Guangdong	South Central	18-55	Gay-oriented activity venue, Internet	Peer-referral, Snowball sampling, gay Web site advertisement	Mean: 0.49 Median: 0.25	N/A	N/A	N/A
Yang HT, 2006 [68]	05/2004-07/2004	liangsu	East	18-56 (28.0)	Gay-oriented activity venue	N/A	Mean: 1.15 Median: 1.05	N/A	N/A	Mean: 6.54 Median: 2.98
Ma X, 2007 [20]	2004	Beijing	North	≥18	MSM community	Respondent-driven sampling	N/A	N/A	N/A	Mean: 3.13 Median: 1.67
Li X, 2006 [69]	2004	Xi'an	Northwest	N/A (28.0)	Gay-oriented activity venue	N/A	N/A	N/A	N/A	Mean: 4.87 Median: 2.09
Zhang BC, 2007 [70]	2004	6 cities (Chongqing, Shenyang, Dalian, Qingdao, Nanjing, Xi'an)	N/A	15-72 (27.6)	Gay bars and volunteer activity venue	Snowball sampling	N/A	N/A	N/A	Mean: 3.36 Median: 1.93
Choi KH, 2007 [26]	09/2004-06/2005	Shanghai	East	18-56 (28.0)	MSM venues	Respondent-driven sampling	N/A	N/A	N/A	Mean: 2.97 Median: 1.39
Xu J, 2007 [71]	03/2005-12/2005	Hefei, Anhui	East	18-42 (23.3)	N/A	Respondent-driven sampling, peer-referral, gay Web site advertisement	Mean: 0.77 Median: 0.72	N/A	N/A	Mean: 0.99 Median: 0.95

Li X, 2008 [72]	06/2005–11/ 2005	Beijing	North	17–54	Web site advertisement, outreach in MSM clubs, bars, parks, and bathhouses	N/A	N/A	N/A	Mean: 4.61 Median: 1.81
Zhu Jl, 2007 [73]	06/2005–12/ 2005	Hefei, Anhui	East	18–29 (20.4)	Gay Web sites	Respondent-driven sampling, peer-referral	N/A	N/A	Mean: 1.49 Median: 0.96
Zhang BC, 2008 [25]	2005–2006	9 cities (Shanghai, Nanjing, Harbin, Shenyang, Wuhan, Zhengzhou, Chongqing, Chengdu, Xi'an)		13–78 (29.1)	Gay-oriented activity venue	Snowball sampling	N/A	Mean: 4.42 Median: 1.95	N/A Mean: 6.60 Median: 2.57
Sun ZX, 2007 [74]	2006	Zhejiang	East	≥15	Participants in a gay-oriented activity	Gay Web site advertisement	N/A	N/A	N/A Mean: 1.77 Median: 1.99
Choi J, 2007 [75]	2006	Nanchang, Jiangxi	East	16–50 (25.3)	Gay bars	N/A	N/A	N/A	N/A Mean: 1.28 Median: 0.67
Ma X, 2007 [20]	2006	Beijing	North	≥16	MSM community	Respondent-driven sampling	N/A	N/A	N/A Mean: 6.25 Median: 3.09
Li N, 2007 [76]	2006	Henan	South Central	17–68 (28.3)	National HIV/AIDS sentinel surveillance	N/A	N/A	N/A	N/A Mean: 4.72 Median: 2.56
Fu LJ, 2007 [27]	02/2006–12/ 2006	Shaoxing, Zhejiang	East	19–41 (26.0)	Peer-referral	Snowball sampling	N/A	N/A	N/A Mean: 1.28 Median: 1.02
Ma J, 2007 [77]	04/2006–10/ 2006	Tianjin	North	N/A (25.2)	Gay Web sites	STD group: Mean: 1.38	N/A	N/A	Non-STD group: Mean: 1.21 Median: 0.49 STD group: Mean: 4.66 Median: 2.98 Non-STD group: Mean: 2.36 Median: 1.23

(continued)

Table 6.1 (continued)

First author, published year	Study design	Number of different sexual partners in past 6 months mean/median						
		Study period	Study location	Region	Age range (mean)	Method of recruitment	Method of sampling	Any male partners
Zhang D, 2007 [78]	05/2006–08/2006	N/A	N/A	≥18	Gay Web sites	N/A	N/A	N/A
Ouyang L, 2008 [79]	06/2006–09/2006	Chongqing	Southwest	16–60 (30.2)	MSM bars, bathhouse	Snowball sampling	N/A	Mean: 1.29
Chen SH, 2008 [80]	07/2006–08/2006	Nanning, Guangxi	South Central	18–45 (26.2)	N/A	Peer-referral, Web site	N/A	Mean: 2.73
Feng L, 2009 [81]	07/2006–09/2006	Chongqing	Southwest	N/A	MSM activity venues (clubs, bars, bathhouses, outdoor cruising areas), community outreach, peer-recruitment, and Web-based recruitment	Venue-based and cruising area-based convenience sampling, Ibid	N/A	Mean: 1.72
Zhou J, 2008 [82]	07/2007–09/2007	Ibid	Ibid	Ibid	Ibid	Ibid	N/A	Mean: 0.33
Liu H, 2009 [83]	09/2006–12/2006	Guiyang, Guizhou	Southwest	15–49 (24.0)	Gay venues, Web sites, hotlines	N/A	N/A	Mean: 1.10
Zhou SJ, 2008 [84]	2007	Shenzhen, Guangdong	South Central	18–45	MSM venues (sauna, bar, public park)	Respondent-driven Sampling	N/A	Mean: 1.34
Wang Y, 2008 [25]	01/2007	Chongqing	Southwest	≥16	Gay bathrooms, bars, activity centers	N/A	N/A	Mean: 0.88
Feng Y, 2010 [85]	03/2007–06/2007	Mianyang, Sichuan	Southwest	16–57 (24.8)	MSM community	Respondent-driven Sampling	Mean: 0.78 Median: 0.48	Mean: 3.19
					MSM community	Snowball sampling	Mean: 1.52 Median: 0.20	Mean: 1.65
						N/A	Mean: 0.73	Mean: 1.95
						N/A	Mean: 2.72	Mean: 1.34
						N/A	N/A	Mean: 0.55
						N/A	N/A	N/A

Zhao HP, 2009 [86]	01/2008–10/ 2008	Harbin, Heilongjiang	Northeast	17–52 (25.0)	MSM who came to voluntary counseling and testing clinics	N/A	N/A	N/A	N/A	Mean: 4.46 Median: 1.96
Feng LG, 2010 [87]	02/2008–06/ 2008	Chongqing	Southwest	N/A	Web site, bar bars, clubs, parks, bathhouse	Snowball sampling	N/A	N/A	N/A	Mean: 2.19 Median: 1.11
Xu J, 2010 [19]	03/2008–07/ 2008	Beijing, Harbin, Zhengzhou, Chengdu	N/A	N/A (29.2)	N/A	Snowball sampling	N/A	N/A	N/A	Mean: 2.90 Median: 1.81
Liang L, 2009 [88]	04/2008–06/ 2008	N/A	N/A	N/A	MSM who had vol- untary counse- ling and testing services	Snowball sampling	N/A	N/A	N/A	Mean: 2.32 Median: 1.53
Wang ZJ, 2010 [89]	05/2008–06/ 2009	Yangzhou, Jiangsu	East	18–78 (33.5)	Web site advertisement	N/A	N/A	N/A	N/A	Mean: 3.28 Median: 1.10
Wei C, 2011 [90]	2008	Jinan, Shandong	East	N/A	MSM volunteer group	Respondent- driven sampling	N/A	N/A	N/A	Mean: 2.73 Median: 1.62
Xu J, 2011 [91]	04/2008–01/ 2009	Liaoning	Northeast	MSM community	N/A	N/A	N/A	N/A	N/A	Mean: 1.64 Median: 0.81

*Note:* Spearman correlation tests were performed to investigate the associations between median/mean of the number of sexual partners and the year of study for each sexual type. For studies spanned more than 1 year, the mid-point of the time interval was used. The correlation results are the following: regular partnership, median vs. year ( $p = 0.95$ ,  $r = 0.05$ ), mean vs. year ( $p = 0.52$ ,  $r = 0.36$ ); noncommercial casual partnership, median vs. year ( $p = 0.92$ ,  $r = -0.21$ ), mean vs. year ( $p = 0.75$ ,  $r = 0.32$ ); commercial partnership: insufficient data for correlation tests; overall partnership, median vs. year ( $p = 0.12$ ,  $r = -0.30$ ), mean vs. year ( $p = 0.20$ ,  $r = -0.25$ ).

**Table 6.2** Studies reporting number of acts with male partners among men who have sex with men in China

First author, published year	Study design	Study location	Region	Age range (mean)	Method of recruitment	Method of sampling	Number of acts with different types of partnerships in past 6 months mean/median <sup>a</sup>			
							Regular partners	Noncommercial casual partners	Commercial partners	Any male partners
Cai WD, 2005 [92]	08/2004–11/2004	Shenzhen	South Central	17–58 (25.9)	Gay bars, fitness centers	N/A	N/A	N/A	N/A	Mean: 11.02 Median: 3.65
Yang HT, 2006 [68]	05/2004–07/2004	Jiangsu	East	18–56 (28.0)	Gay-oriented activity venue	N/A	N/A	N/A	N/A	Mean: 2.38 Median: 1.51
Zhang X, 2007 [93]	01/2005–12/2006	Beijing	North	18–55 (26.1)	VCT clinic at Chaoyang District CDC	N/A	N/A	N/A	N/A	<i>Insetive:</i> Mean: 4.90 Median: 4.81 <i>Receptive:</i> Mean: 5.02 Median: 4.77
Xiao Y, 2009 [94]	07/2006–09/2007	Chongqing	Southwest	18–68 (27.7)	Gay-oriented volunteer workgroups, gay bars, bathhouses, clubs, saunas, gay Web sites	Multiplier methods: venue-based recruitment, Internet advertisements, community outreach, and snowball sampling	N/A	N/A	N/A	Mean: 1.45 Median: 1.29

There is insufficient data to perform Spearman correlation for median/mean number of sexual acts and the year of study

**Table 6.3** Studies reporting the consistent condom use in past 6 months among men who have sex with men with different types of male sexual partners in China

First author, published year	Study period	Study location	Region	Age range (mean)	Method of recruitment	Consistent condom use with different types of partnerships in past 6 months, n/N (%)		
						MSM venues (28.1)	Snowball sampling	Regular partners
Lan Y, 2004 [95]	09/2003–12/2003	Sichuan	Southwest	N/A	MSM venues (bars, bath-houses, Web hotlines, Web sites, parks)	53/336 (15.8 %)	53/325 (16.3 %)	32/99 (32.3 %)
Yang HT, 2006 [68]	05/2004–07/2004	Jiangsu	East	18–56 (28.0)	Gay-oriented activity venue	N/A	N/A	17/28 (60.7 %)
He Q, 2008 [96]	2004	Chengdu, Sichuan	Southwest	17–62 (26.1)	MSM community	Snowball sampling	N/A	10/44 (22.7 %)
	2005	Ibid	Ibid	17–49 (25.4)	Ibid	Ibid	N/A	24/41 (58.5 %)
	2006	Ibid	Ibid	16–58 (25.8)	Ibid	Ibid	N/A	17/36 (47.1 %)
	2007	Ibid	Ibid	16–74 (27.7)	Ibid	Ibid	N/A	19/34 (55.9 %)
Tian XB, 2006 [97]	06/2005–09/2005	Nanchong, Sichuan	Southwest	16–65	Gay-identified venues, Web hotlines, Web sites	Snowball and convenience sampling	4/84 (4.8 %)	0.9 % (1/109)
Xing JM, 2007 [98]	10/2005–12/2005	Changsha, Hunan	South Central	14–63 (29.8)	Gay bars, bath house, public toilet	N/A	N/A	18/36 (50.0 %)
Chen SC, 2007 [99]	07/2005–09/2005	Hangzhou, Zhejiang	East	18–70 (28.0)	Gay-identified venues, Web sites	Snowball sampling	N/A	N/A
Lau JT, 2006 [47]	02/2004–08/2005	Kunming, Yunnan	Southwest	15–75	Gay-identified venues, Web site	Multiple sampling (Convenience, snowball)	N/A	331/730 (45.3 %)
								90/130 (69.2 %)

(continued)

Table 6.3 (continued)

First author, published year	Study period	Study design			Consistent condom use with different types of partnerships in past 6 months, n/N (%)				
		Study location	Region	Age range (mean)	Method of recruitment	Method of sampling	Regular partners	Noncommercial casual partners	Commercial partners
Cai CF, 2008 [100]	06/2006–12/2006	Zhejiang	East	16–46 (26.6)	Gay-identified venues	N/A	N/A	N/A	12/18 (66.7 %)
Wu J, 2008 [101]	10/2006–12/2006	Shanghai	East	18–54 (26.6)	Gay-identified venues	N/A	18/80 (22.5 %)	71/144 (49.3 %)	N/A
Wang HL, 2008 [47]	08/2006–10/2006	Shenyang, Liaoning	Northeast	18–60	N/A	N/A	69/212 (32.5 %)	98/218 (45.0 %)	95/188 (50.5 %)
Duan YW, 2010 [102]	12/2006–02/2007	Mianyang, Sichuan	Southwest	≥ 18	Peer-referral	N/A	34/137 (24.6 %)	23/75 (30.8 %)	13/33 (39.3 %)
China Global Fund AIDS Program Round 5, 2007 [103]	Ibid	Yibin, Sichuan	Ibid	Ibid	Ibid	N/A	43/142 (30.3 %)	33/77 (43.2 %)	23/58 (40.4 %)
Liu LY, 2008 [104]	2006	Chongqing	Southwest	18–65	MSM venues	Snowball sampling	N/A	N/A	27/44 (61.4 %)
Xiao Y, 2009 [94]	07/2006–09/2007	Heilongjiang, Chongqing	Northeast	N/A (29.0)	Gay-identified venues	Ibid	N/A	N/A	17/30 (56.7 %)
			Southwest	18–68 (27.7)	Gay-oriented volunteer	Multiplier methods: workgroups, gay bars, bath-houses, clubs, saunas, gay Web sites	N/A	55/202 (27.2 %)	144/202 (71.5 %)

Feng Y, 2010 [85]	03/2007–06/2007	Chengdu, Sichuan	Southwest	16.8–44.5	MSM community	Snowball sampling	84/358 (23.5 %)	149/386 (38.6 %)	N/A
Wang Y, 2008 [105]	12/2006–01/2007	Mianyang, Sichuan	Southwest	16–57 (24.8)	MSM network	Respondent-driven Sampling	34/137 (24.8 %)	23/76 (30.3 %)	13/33 (39.4 %)
Zeng H, 2011 [106]	07/2007–08/2007	Chongqing	Southwest	15–65 (26.4)	Web site, MSM venues	Snowball sampling	N/A	N/A	38/50 (76.0 %)
Cai YM, 2009 [107]	10/2007–12/2007	Shenzhen, Guangdong	South Central	18–45	Gay-identified venues	Respondent-driven Sampling	N/A	N/A	25/33 (75.8 %)
Liu H, 2009 [83]	2007	Shenzhen, Guangdong	South Central	18–45	MSM venues (sauna, bar, public park)	Respondent-driven Sampling	N/A	N/A	21/26 (80.8 %)
Wen XQ, 2010 [108]	04/2008–06/2008	Guilin, Guangxi	South Central	N/A	MSM venues	Snowball sampling	N/A	N/A	4/13 (30.8 %)
Wang M, 2009 [109]	2008	Kunming, Yunnan	Southwest	16–80	MSM venues, Web sites	Random sampling	N/A	192/306 (62.7 %)	N/A
Xu JJ, 2011 [91]	04/2008–01/2009	Liaoning	Northeast	≥18	MSM community	N/A	180/371 (48.5 %)	254/436 (58.3 %)	N/A
Gong CT, 2010 [110]	10/2008–06/2009	Quanzhou, Fujian	East	19–46 (22.0)	MSM venues	Snowball sampling	N/A	N/A	2/5 (40 %)
Ji GP, 2010 [111]	2008 1st round	Anhui	East	18–72 (26.7)	N/A	N/A	N/A	N/A	11/23 (47.8 %)
	2008 2nd round	Ibid	Ibid	18–72 (24.8)	N/A	N/A	N/A	N/A	11/21 (52.4 %)
	2009	Ibid	Ibid	18–72 (26.0)	N/A	N/A	N/A	N/A	21/35 (60.0 %)

Note: Linear regressions were used to investigate the temporal trend of condom usage during 2002–2010 for each sexual type. For studies spanned more than 1 year, the mid-point of the time interval was used. The slopes of the linear regressions are the following: 5.49 % (95 % CI, 1.33–9.64 %), 8.29 (95 % CI, 2.59–13.99 %), and 7.02 % (95 % CI, 0.70–13.33 %) for regular, casual, and commercial partners, respectively. All slopes are significantly greater than zero

procedure. The optimization routine minimizes the difference between the optimized fits and the actual datasets with the error function:

$$E = \sum_i (N_i - D_i)^2,$$

where,  $N_i \sim LN(\mu_i, \delta_i)$  denotes curves of lognormal distributions of sexual partners with each sexual type  $i$ , mean  $\mu_i$  and variance  $\delta_i$ , and  $D_i$  denotes a randomly generated bootstrapped dataset of the original datasets for the corresponding sexual type. Thus the error function represents the sum of differences between the optimized distribution curves and bootstrapped data points. Bootstrapping is performed by resampling on the original datasets to generate a new set of datasets with identical dimensions as the original datasets for each simulation. The weighted resampling process is designed in a way that the probability a dataset being resampled is proportional to its sample size. This ensures studies with larger sample size weights proportionally during optimization. The optimization routine is repeated 100 times and the best 50 fits are selected to provide uncertainty bounds. This optimization route was performed on a linear regression model of the temporal trend of condom usage with estimated 95 % confidence intervals. Uncertainty bounds are reported as 95 % confidence intervals in parentheses. The mean number of sexual partners and acts, together with their 95 % confidence intervals, are hence calculated numerically based on the fitted lognormal distribution with the analytical expression  $e^{\mu+\sigma^2/2}$ .

The distribution of 6-monthly incidence rates of HIV infection for each sexual partnership type is calculated based on the distributions of number of sexual partners ( $N_i \sim LN(\mu_i, \delta_i)$ ), number of sexual acts ( $S_i \sim LN(\mu_i, \delta_i)$ ), and condom usage ( $C_i \sim N(\mu_i, \delta_i)$ ) in the past 6 months, according to the following mathematical expression (which is a relatively standard force of infection term in transmission modeling, using a Bernoulli equation):

$$\begin{aligned} \text{HIV incidence rate} \\ \text{in sexual partnership} \\ \text{with type } i \\ \widehat{I_i(t)} &= p(t)S_i \left( 1 - \underbrace{(1 - (1 - \varepsilon)\beta)^{N_i(t) \cdot C_i(t) / S_i(t)}}_{\text{Transmission probability}} \right) \\ &\quad \underbrace{\left( 1 - \beta \right)^{N_i(t) \cdot (1 - C_i(t)) / S_i(t)}}_{\text{Transmission probability}} \end{aligned}$$

The estimated annual HIV incidence among MSM is hence the sum of incidence contributed by each sexual type. Here,  $p(t)$  denotes the HIV prevalence among Chinese MSM during 2001–2010, the prevalence data is extracted from a recent publication that synthesized the available HIV prevalence data in China during 2001–2010 in a meta-analysis [43]. This study reported that HIV prevalence among Chinese MSM increased from 1.3 % (95 % CI: 0.8–2.1 %) prior to 2004 to 2.4 % (95 % CI: 1.7–3.2 %) during 2005–2006 and then to 4.7 % (95 % CI: 3.9–5.6 %) after 2007. The per-act infectivity,  $\beta$ , for homosexual penile-anal sex was assumed to be 1–1.5 % [30], whereas the condom efficacy,  $\varepsilon$ , was assumed to be 85–95 % [32–36]. Biologically, the expression  $(1 - (1 - \varepsilon)\beta)^{N_i(t) \cdot C_i(t) / S_i(t)}$  refers to the

transmission probability of HIV infection per protected penile-anal act, whereas  $N_i(t) \cdot (1 - C_i(t)) / S_i(t)$  indicates the probability per un-protected penile-anal act. The product of these two expressions hence represents the overall HIV transmission probability of a MSM population with  $N_i$  sexual partners and  $C_i$  condom use on average. As HIV incidence is also proportional to the temporal trend of HIV disease burden  $p(t)$  in the population and the average number of sexual acts  $S_i$ , we derive above expression. Based on the optimization routine, each simulation will generate a unique set of distribution of number of sexual partners and acts and condom usage. The calculation of the above Bernoulli equation takes into account of these distributions and other biological parameters. This results in a distribution of HIV incidence. The calculation is repeated for 50 times to estimate the confidence intervals of incidence distribution. Annual HIV incidence was calculated by multiplying the 6-monthly incidence estimates by two. We define incidence rates above 20 cases per 1,000 persons as high incidence rates and use this threshold to assess the temporal trend of HIV epidemic trend among Chinese MSM.

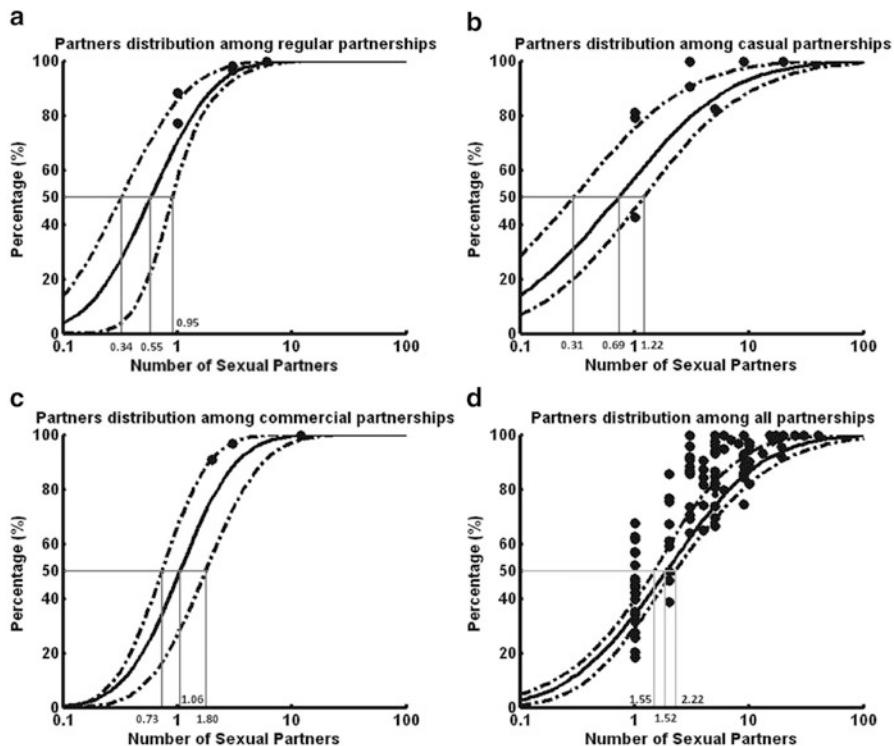
## 6.3 Results

### 6.3.1 Literature Search and Selection

We identified 635 articles from one English and three Chinese electronic databases (103 in PubMed, 175 in CQVIP, 131 in CNKI, and 226 in Wanfang); we also identified 8 extra articles from the reference lists of the selected articles. After screening the titles of articles we excluded 335 articles due to duplication of articles from multiple databases. We screened the abstracts of the remaining 308 articles, following which 43 articles were excluded because five were unrelated and 38 were non peer-reviewed. The remaining 265 articles were eligible for full-text screening; we further excluded 210 articles (five were review papers, ten did not report the study site, 37 did not report the types of partners in condom usage, 64 did not report the MSM behavioral data in past 6 months and 94 did not report any behavioral data). Finally, 55 articles (13 in English and 42 in Chinese) were included in qualitative synthesis (33 studies reported the numbers of partners, 24 studies reported the consistent condom use among different types of partnerships and 4 studies reported the number of sexual acts with different types of partnerships). The research strategy and selection of articles are shown schematically in Fig. 6.1.

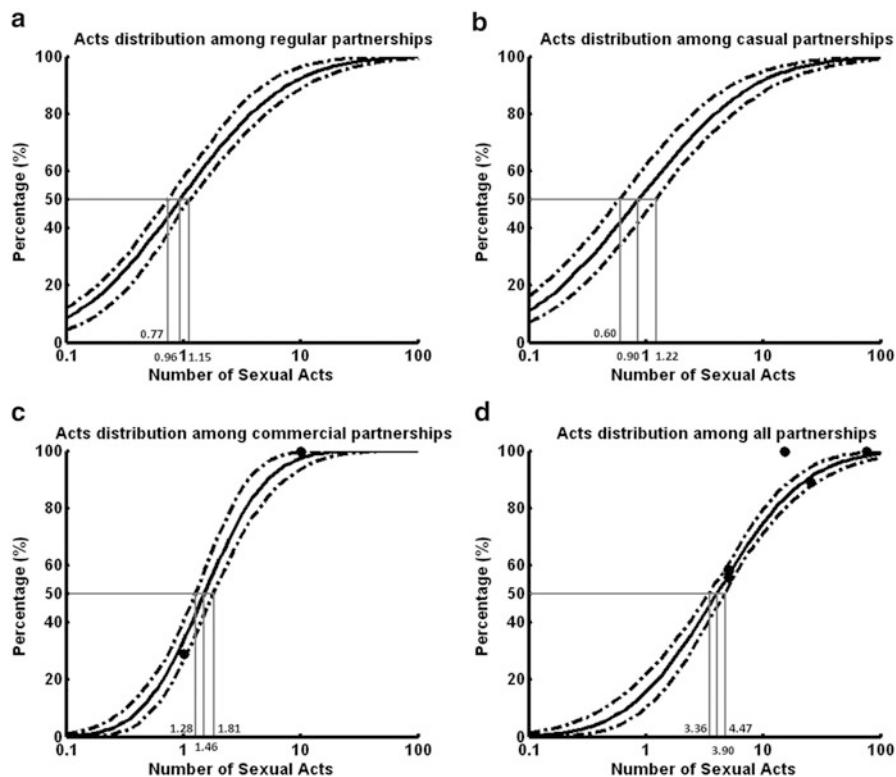
### 6.3.2 Sexual Behaviors Among Chinese MSM

A lognormal distribution function was used to fit partner distributions of individual studies. In this part of the analysis, each distribution simulation results in a pair of



**Fig. 6.2** The distribution the number of (a) regular, (b) casual, (c) commercial, and (d) all sexual partners among Chinese MSM in the past 6 months during 2002–2010

mean and median values (Table 6.1). Spearman correlation tests were performed on the mean and median of number of sexual partners (and acts) with their studied years whenever data were sufficient, but none of the correlations were found to be significant (Tables 6.1, 6.2, and 6.3). In contrast, condom usage levels among MSM were shown to be significantly increasing over time among all sexual types (Fig. 6.3). The annual rates of condom usage increased significantly at rates of 5.49 % (95 % CI, 1.33–9.64 %), 8.29 (95 % CI, 2.59–13.99), 7.02 % (95 % CI, 0.70–13.33 %) for regular, casual, and commercial partners respectively during 2002–2010 (Fig. 6.4). Since distributions in numbers of sexual partners and acts did not vary temporally, this enabled us to pool relevant datasets from individual studies without considering their study years. An optimization procedure was then employed to reconcile all available data on distributions from different types of sexual partnerships. Figures 6.2 and 6.3 reveal that in the past 6 months, the mean number of regular, casual, and commercial sexual partners for an average Chinese MSM are 0.96 (95 % CI, 0.59–1.18), 3.75 (95 % CI, 1.72–6.25), and 1.61 (95 % CI, 0.97–2.78) respectively (Table 6.4), whose sum was comparable with the overall number of sexual partners 6.17 (95 % CI, 3.45–9.96). Further, the mean number of sexual acts for these partnerships were 4.16 (95 % CI, 2.87–5.63), 5.33 (95 % CI,

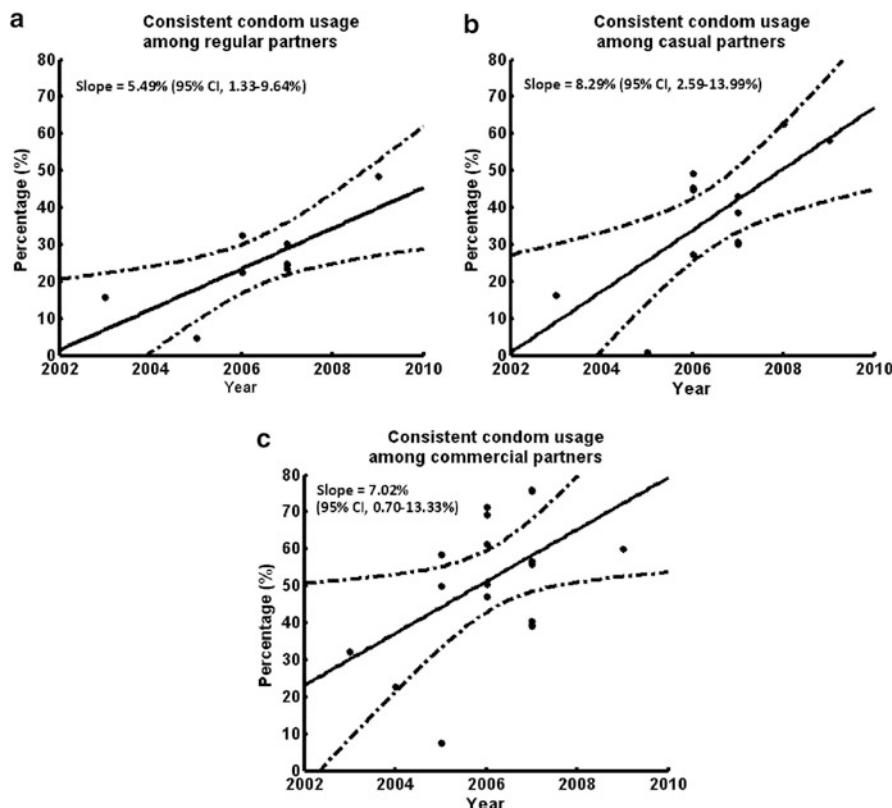


**Fig. 6.3** Distribution of number of acts among (a) regular, (b) casual, (c) commercial, and (d) all partners among Chinese MSM in past 6 months during 2002–2010

2.84–7.28), and 2.38 (95 % CI, 1.77–3.70) respectively in the same period, which corresponded to a total number of acts of 11.87 (95 % CI, 8.87–15.25) (Table 6.4). These indicate that an average Chinese MSM participates in approximately 4.33 (95 % CI, 2.81–6.46), 1.42 (95 % CI, 0.62–3.08), 1.48 (95 % CI, 0.79–3.30) acts with each of his regular, casual, and commercial partners in the past 6 months. Notably, all mean values represented the average behaviors of the entire MSM population. The diverse patterns of sexual behavior among MSM were summarized by the distributions in Figs. 6.2 and 6.3.

### 6.3.3 Estimated Annual HIV Incidence

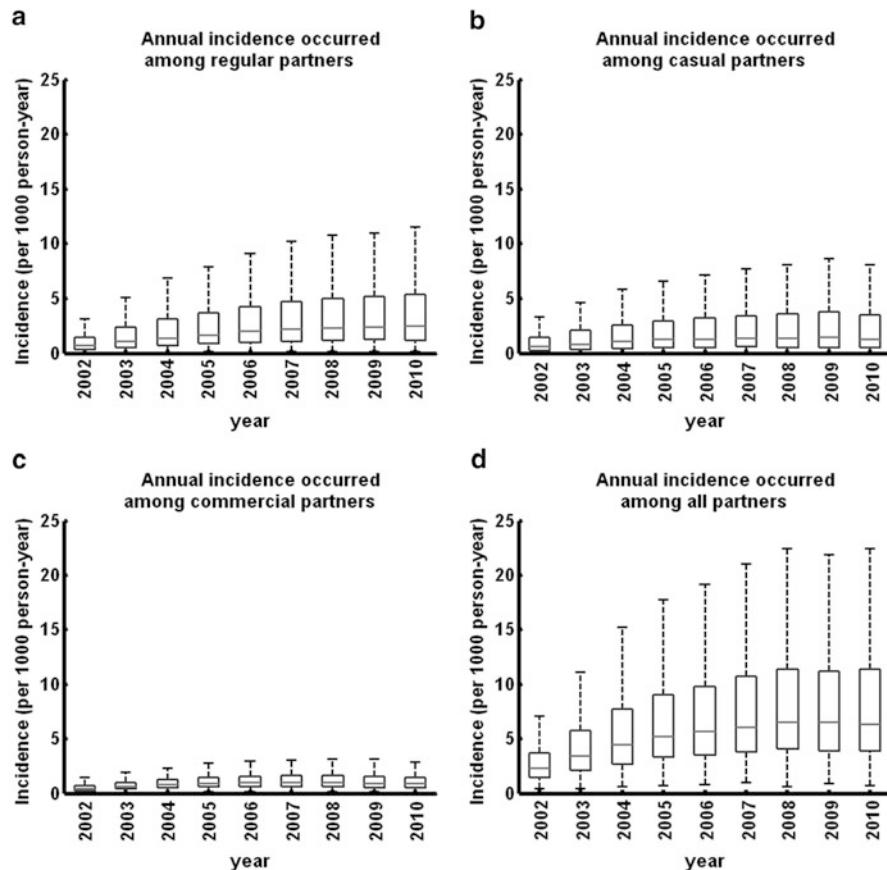
The distribution of annual incidence rates was calculated based on the distributions of number of sexual partners, number of sexual acts, and condom usages in partnership types. Our results showed that HIV incidences in all sexual types follow



**Fig. 6.4** Consistent condom usage among (a) regular, (b) casual, (c) commercial partners in past 6 months among Chinese MSM during 2002–2010

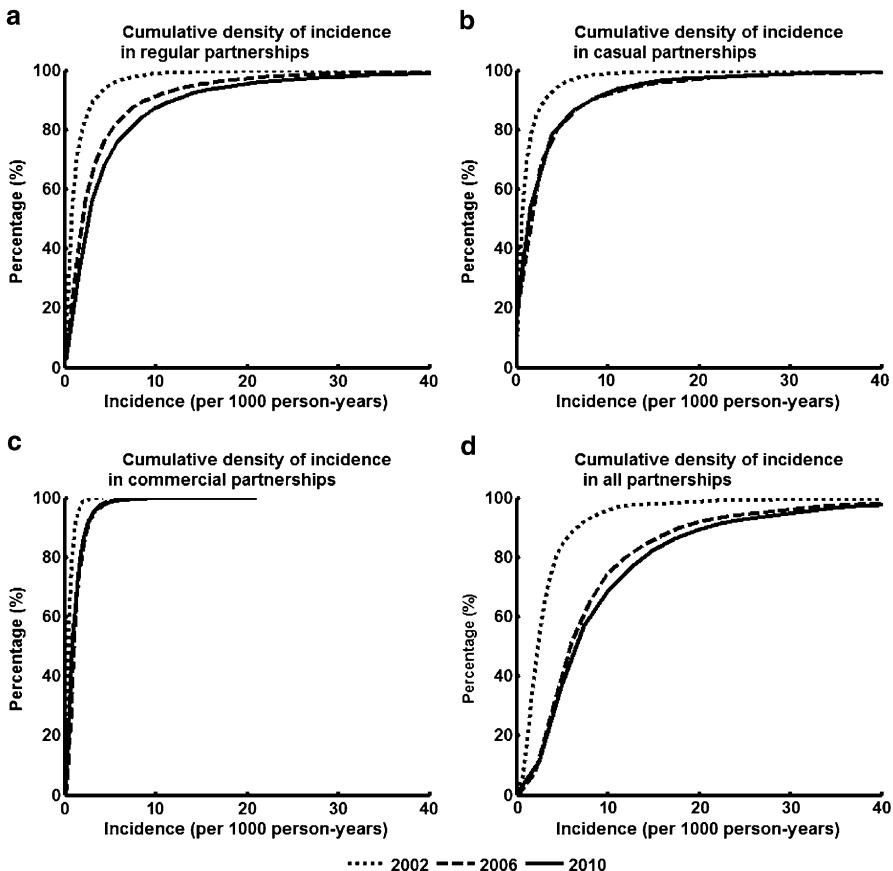
**Table 6.4** Summary of key results on the behavioral distribution of number of sexual partners and acts among MSM in China

	Optimization fittings	
	Mean	Median
Number of partners in the past 6 months		
Regular partners	0.96 (0.59–1.18)	0.55 (0.34–0.95)
Casual partners	3.75 (1.72–6.25)	0.69 (0.31–1.22)
Commercial partners	1.61 (0.97–2.78)	1.06 (0.73–1.80)
All partners	6.17 (3.45–9.96)	1.86 (1.52–2.22)
Number of sexual acts in the past 6 months with		
Regular partners	4.16 (2.87–5.63)	0.96 (0.77–1.15)
Casual partners	5.33 (2.84–7.28)	0.90 (0.66–1.22)
Commercial partners	2.38 (1.77–3.70)	1.46 (1.28–1.81)
All partners	11.87 (8.87–15.25)	3.90 (3.36–4.47)



**Fig. 6.5** Annual HIV incidence occurred among (a) regular, (b) casual, (c) commercial, and (d) all partners among Chinese MSM during 2002–2010

an increasing trend during the period 2002–2010 (Fig. 6.5). Annual HIV incidence among casual and commercial partners gradually increased from 0.44 (95 % CI, 0.39–0.51) to 0.67 (95 % CI, 0.62–0.72) per 1,000 person-years (pys) in 2002 to 1.13 (95 % CI, 1.05–1.21) and 1.49 (95 % CI, 1.43–1.56) per 100 pys in 2006 respectively, then both stabilized at relatively constant level during the period 2006–2009. In comparison, annual incidence among regular MSM partners maintained a steady increasing trend from 0.46 (95 % CI, 0.44–0.49) per 1,000 pys in 2002 to 1.66 (95 % CI, 1.60–1.72) per 1,000 pys in 2010 (Fig. 6.5). Accordingly, the overall incidence increased 2.6-fold from 2.23 (95 % CI, 2.11–2.44) to 5.69 (95 % CI, 5.37–6.00) per 1,000 pys during the same period. Figure 6.6 demonstrated the temporal shifts of the HIV incidence profile. It was estimated that 3.87 % of susceptible Chinese MSM had a HIV incidence rate over 10 per 1,000 pys in 2002, and this percentage quickly increased to 29.19 % in 2006 then 31.28 % in 2010. If a threshold in HIV incidence

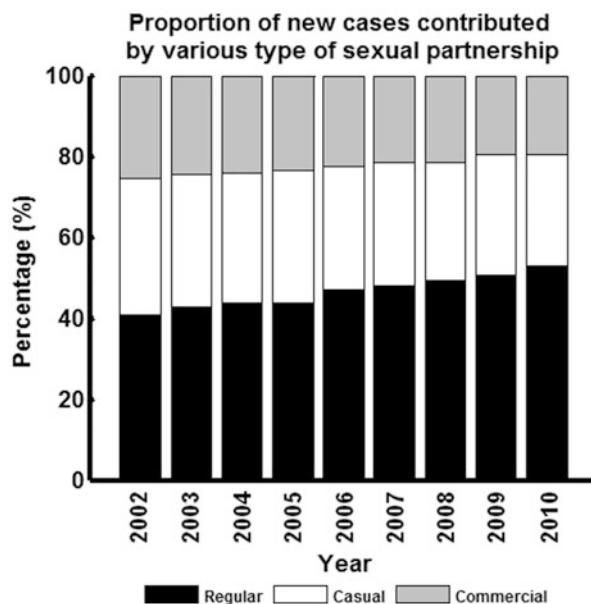


**Fig. 6.6** Cumulative density of incidence among (a) regular, (b) casual, (c) commercial, and (d) all partners

of 20 per 1,000 pys is considered, then we estimate a sharp increase from 0.23 % in 2002 to 8.61 % in 2006 and 11.06 % in 2010 (Fig. 6.6d).

Notably, the composition of new infections from different types of sexual partnerships changed substantially over time. In 2002, the proportion of new infections attributable to regular, casual, and commercial partnerships were 32, 35, and 33 % respectively, with each type contributing approximately equally. In comparison, the profile changed to 43, 32, and 25 %, respectively in 2010. Infections attributable to regular partnerships started to dominate, with an overall increase of 11 %, whereas the proportion attributable to commercial partnerships reduced by 8 % over the decade (Fig. 6.7).

**Fig. 6.7** Composition of new HIV infections contributed by high-risk sexual behaviors in various types of partnerships from 2002 to 2010



## 6.4 Discussion

### 6.4.1 MSM Sexual Behaviors

In this study we presented, for the first time, a rigorous analysis on the distribution of number of sexual partners and acts among Chinese MSM. Based on the estimates from the optimization we investigated the temporal trend of annual HIV incidence among MSM in China from 2002 to 2010. Here, we define, in a statistical sense, an “average” Chinese MSM to be an individual who possesses behavioral characteristics that are exactly the means of the behavioral distributions. Our analysis showed that on average a Chinese MSM has 0.96 (95 % CI, 0.59–1.18), 3.75 (95 % CI, 1.72–6.25), and 1.61 (95 % CI, 0.97–2.78) regular, casual, and commercial partners respectively in the past 6 months. This confirms a previous finding that noncommercial casual partners represent the majority of MSM partnerships [44]. Consistently, the majority of sexual acts were casual (mean 5.33, 95 % CI, 2.84–7.28). This was similar to the number of regular acts 4.16 (95 CI%, 2.87–5.63) but substantially higher than the number of commercial acts 2.38 (95 % CI, 1.77–3.70). However, in terms of sexual acts per partnership, our results indicate that casual and commercial acts were mostly once-off, and frequency of sexual activities per regular partnership is as low as 4.33 (95 % CI, 2.81–6.46) acts in the past 6 months. To our best knowledge, this is the first time that the numbers of sexual acts per partnership were estimated according to their sexual types in

Chinese settings for MSM. Our analysis also reported significant increases in condom usage among all types of MSM sexual partnerships. This is consistent with latest findings reported from other studies [37, 45–47].

Sexual behaviors among Chinese MSM are more conservative in comparison with MSM in developed countries. In the United States, the median numbers of regular and casual partners of MSM in the past 12 months are reported to be 1 and 3 respectively [48], which are more than twice the respective numbers (regular, 0.55 (95 % CI, 0.34–0.95); casual 0.69 (95 % CI, 0.31–1.22)) in the past 6 months in China. In Australia, the average number of overall partnerships in the past 6 months have been reported to be 12–14 during the period 1999–2006 [49], almost twofold the corresponding finding of 6.17 (95 % CI, 3.45–9.96) in China. In addition, the average number of anal intercourse acts per regular partner in the past 6 months is 42–62 (1.6–2.4 acts/week) in the US [49, 50], whereas the number of acts per casual partnership is 26–52 (1–2 acts/week [49–51]) in Australia. These are five to ten times higher than our estimated number of acts among MSM in China. The more conservative sexual behaviors of Chinese MSM are likely a result of traditional familial values, societal expectations, and self moral perception of homosexual acts among MSM in China [52–54].

#### 6.4.2 *HIV Incidence Among MSM*

We estimated an average HIV incidence of 5.69 (95 % CI, 5.37–6.00) per 1,000 pys among Chinese MSM in 2010. In contrast, the observed HIV incidence levels in some Chinese regions are 5–15 times higher (26–100/1,000 pys [55–59]). The discrepancy is likely due to multiple reasons. All these empirical studies were performed in Chinese major cities where HIV prevalence among MSM was already known to be high [18, 25, 46]. In addition, MSM participants were selected mostly from MSM venues in these studies, where this MSM subgroup is often more likely to participate in higher-risk sexual behaviors. Our study collects a wide spectrum of studies from different geographical locations based on various methods of recruitment and sampling. Therefore, the estimated average incidence is likely a result of the average behaviors of overall MSM population in China. The fact that the increasing proportion of MSM having an incidence rate exceeding 20 cases per 1,000 pys (0.23 % in 2002 vs. 11.06 % in 2010) suggests a growing proportion of Chinese MSM with high risks of HIV infection. This subpopulation with typically high incidence rate may coincide with the population sampled in the published studies [55–59].

Our analysis did not show significant temporal changes in the numbers of sexual partners and acts for any MSM sexual types. However, condom usages in all MSM sexual types were significantly increased during the past decade. In contrast, the increase in estimated HIV incidence seemed to be counterintuitive. However, the incidence rate increase is likely due to the rapid increase of HIV prevalence among MSM. As HIV prevalence represents an accumulation of infected cases

among the population, susceptible MSM are required to increase their condom usage substantially to avoid becoming infected. Our results argue that despite significant increases in condom usage, the current levels of condom usage were not sufficient to offset the effects of elevated risk of infection as a result of increasing prevalence among the MSM population.

#### ***6.4.3 Implications to Intervention***

Our study has important implications for HIV intervention strategies among MSM in China. Regular partnerships remain the main contributor of new HIV infections throughout the past decade and this trend has been increasing. This is consistent with the fact that the annual increase of condom usage among regular partnerships is lower than among casual and commercial partnerships. Regular partners tend to have more sexual acts per partnership and lower consistent condom usage. To date, there have been no specific harm reduction intervention programs targeting MSM in China. However, this is likely to change in the near future with increasing awareness from both the governmental and scientific communities. Regular partnerships should be the main focus and priority for any future interventions to curb the rapid transmission of HIV. Health resources should be directed towards intervention programs, such as HIV health education and voluntary counseling and testing (VCT), to encourage MSM and their regular sexual partners to be tested for HIV infection on a regular basis. Harm reduction programs that aim to increase condom usage should be implemented to improve safe sexual practices among Chinese MSM. These programs should target MSM venues, online chat-rooms, and entertainment venues to promote condom usages not only with their casual and commercial partners but also among regular partnerships.

#### ***6.4.4 Challenges and Limitations***

Although our results demonstrate a slowly leveling out in the trend of HIV incidence, the HIV epidemic among Chinese MSM will likely be sustained due to trends in sexual behaviors in this population. As the current level of condom usage is insufficient to balance the increase in incidence due to greater pools of infected people with the potential to transmit infection, HIV prevalence will likely continue to increase in the near future to a level where the number of new infected cases is balanced by the number of deaths among HIV-infected MSM. Despite persistent social stigma towards MSM, increasing openness towards homosexual relationships have been observed in Chinese major cities [8]. This indicates that the sexual behavioral profile among Chinese MSM is likely to change to more closely resemble that of the developed countries with greater numbers of sexual partners and acts. In addition, a recent study estimated that 87 % of HIV-infected Chinese

MSM are undiagnosed [60]. Consistent with this, a recent study released by China CDC showed that among the patients who are on antiretroviral therapy (ART) only 7 % were identified as MSM [61], although MSM contributed to approximately one-third of the new HIV cases in 2009 [16]. The low diagnosis and treatment coverage rates are major barriers to HIV control and prevention among Chinese MSM. In the absence of effective scale-up of HIV testing and ART, HIV is likely to continue to spread rapidly among Chinese MSM.

Our study reconciles available MSM behavioral data through an optimization routine. However, it is subjected to several limitations. First, data for the distribution of the number of sexual acts is limited. There were only one available dataset for the distribution of commercial acts and two for overall sex acts, whereas data for regular and casual sex acts were absent. This substantially limited our analysis of the temporal trend in the number of sex acts. The majority of collected studies did not report the number of sex acts as one of their behavioral indicators. This reflects the knowledge gap between qualitative and quantitative HIV research in China, as only few studies were designed in a way that behavioral indicators can be quantified to estimate the risk of HIV transmission. Thus, we noted this important limitation in the empirical data, and restrained the interpretation of modeling results within the scope of the data and its representability of the behaviors of general MSM population in China. Second, the optimization process was performed based on an assumption of lognormal distributions of behavioral indicators. Although this was a valid assumption, the nature of lognormal distribution may lead to overestimates of the actual numbers due to its skewness in large values. Besides, in a study setting, a participant was more likely to provide better account of partners and acts information for specific partnership types, but under-report the overall numbers from all partners. Hence, our optimization exercise may provide higher estimates than the actual empirical data. Third, in the analysis of condom usage, we employed linear regression to estimate the temporal increase of condom usage. A linear assumption is valid when a short study period is considered. As condom usage levels have started to slowly plateau in recent years, the linear regression may overestimate the actual usage rates in later years. For any future studies with more recent data, a nonlinear regression may be more appropriate. Fourth, the current study did not take into consideration a number of factors with significant effects on sexual behaviors and HIV prevalence, including geographical differences, age differences, and education levels. Recent studies have documented that HIV prevalence was much higher in Southwest China than any other parts of the country [17–19, 62]. Whether this is an indication of more risky sexual behaviors among MSM in the Southwest remains unclear. Also, younger MSM have been reported to be more sexually active [63] and MSM with lower education levels may be more prone to high-risk behaviors [64]. Due to limited information in our collected data, we were unable to stratify our analysis according to these factors.

Conclusively, this study investigates the trends HIV incidence rates in relation to their risk sexual behaviors and partnership types among Chinese MSM through a comprehensive literature research on sexual behaviors. With the utilization of a

mathematical optimization approach, this study reconciles the distributions of sexual behavioral data over the last decade and infers the heterogeneous distributions of behavioral patterns among Chinese MSM. Based on the optimization results, we concluded that regular partnership contributes to the most of sexual acts among Chinese MSM. The behavioral data implies an increase in HIV incidence of approximately 2.6-fold from 2.23 (95 % CI, 2.11–2.44) to 5.69 (95 % CI, 5.37–6.00) per 1,000 person-years during 2002–2010. The proportion of new infections attributed to regular partnerships increased by 11 % from 32 to 43 % during the same period. Regular partnerships are the main contributor of new HIV cases among MSM in China, public health intervention strategies are required to increase condom usage and HIV testing rates among regular partners to curb the growing trend HIV incidence.

## References

1. Kartikeyan S, Bharmal R, Tiwari R, Bisen P (2007) HIV and AIDS: basic elements and priorities. Springer, New York
2. Sheng L, Cao WK (2008) HIV/AIDS epidemiology and prevention in China. *Chin Med J (Engl)* 121(13):1230–1236
3. Sun X, Wang N, Li D, Zheng X, Qu S, Wang L, Lu F, Poundstone K (2007) The development of HIV/AIDS surveillance in China. *AIDS* 21(Suppl 8):S33–S38
4. Zeng Y, Fan J, Zhang Q, Wang PC, Tang DJ, Zhon SC, Zheng XW, Liu DP (1986) Detection of antibody to LAV/HTLV-III in sera from hemophiliacs in China. *AIDS Res* 2(Suppl 1): S147–S149
5. Zheng XW (1989) A report on AIDS surveillance from 1985 to 1988 in China. *Zhonghua Liu Xing Bing Xue Za Zhi* 10(2):65–67
6. Zheng X, Tian C, Choi KH, Zhang J, Cheng H, Yang X, Li D, Lin J, Qu S, Sun X et al (1994) Injecting drug use and HIV infection in southwest China. *AIDS* 8(8):1141–1147
7. Sun X, Nan J, Guo Q (1994) AIDS and HIV infection in China. *AIDS* 8(Suppl 2):S55–S59
8. Zhang BC, Chu QS (2005) MSM and HIV/AIDS in China. *Cell Res* 15(11–12):858–864
9. Zhang K, Ma S (2002) Epidemiology of HIV in China (Editorial). *BMJ* 324(6):803–804
10. State Council Office of People's Republic of China (2001) China action plan to contain, prevent and control HIV/AIDS (2001–2005). In: State Council Document, Beijing
11. State Council Office of People's Republic of China (2006) China's action plan (2006–2010) for reducing and preventing the spread of HIV/AIDS In: State Council Document, Beijing
12. Ministry of Health of the People's Republic of China (2010) China 2010 UNGASS Country Progress Report. In: Ministry of Health, Beijing, China
13. Zhu TF, Wang CH, Lin P, He N (2005) High risk populations and HIV-1 infection in China. *Cell Res* 15(11–12):852–857
14. Ruxrungtham K, Brown T, Phanuphak P (2004) HIV/AIDS in Asia. *Lancet* 364(9428):69–82
15. Ministry of Health of the People's Republic of China (2008) UNGASS Country Progress Report P.R. China (2006–2007). In: Ministry of Health, Beijing, China
16. UNGASS (2010) China 2010 UNGASS Country Progress Report (2008–2009)
17. Chow EPF, Wilson DP, Zhang J, Jing J, Zhang L (2011) Human immunodeficiency virus prevalence is increasing among men who have sex with men in China: findings from a review and meta-analysis. *Sex Transm Dis* 38(9):845–857

18. Bao YG, Zhang YH, Zhao JK, Sun JP, Tan HZ (2009) HIV infection and KAP status among men who have sex with men in 14 Chinese cities. *Zhonghua Yu Fang Yi Xue Za Zhi* 43(11):981–983
19. Xu J, Han DL, Liu Z, Ma XY, Wang LL, Xu J, Pang L, Zhang HB, Wu ZY (2010) The prevalence of HIV infection and risk factors among MSM in 4 cities, China. *Chin J Prev Med* 44(11):975–980
20. Ma X, Zhang Q, He X, Sun W, Yue H, Chen S, Raymond HF, Li Y, Xu M, Du H et al (2007) Trends in prevalence of HIV, syphilis, hepatitis C, hepatitis B, and sexual risk behavior among men who have sex with men. Results of 3 consecutive respondent-driven sampling surveys in Beijing, 2004 through 2006. *J Acquir Immune Defic Syndr* 45(5):581–587
21. Zhou ZH, Li SM, Liu YJ, Li QC, Li DL, Zhang Z, Jiang SL, Luo FJ, Shi W, Ruan YH et al (2010) Survey of HIV and syphilis infection and influential factors associated with unprotected anal intercourse in men who have sex with men in Beijing. *China Tropical Med* 10(1):10–12
22. Liu H, Wang N, Zhang Q, Shao Y, Wu T, Liu Y, Jiang HB, Song Y, Zhang L, Yin WY et al (2007) Study of HIV and syphilis infection situation and sexual behavioral characteristics among 416 MSM. *Chin J AIDS STD* 13(3):230–232, 234
23. Lu CG, Yuan F, Shi ZH, Yang JZ, Li XY, Gao L, Lu X, Hu SY (2006) HIV infection survey among MSM. *Chin J Public Health* 2(11):1320–1321
24. Gu Y, Qu P, Xu L, Luo M, Wang X, Gu J, Zhao L, Lu Y-H, Zhou B (2004) Survey of knowledge, attitude, behavior and practice related to STI/HIV among male homosexuality in Shenyang. *Chin J Public Health* 20(5):573–574
25. Zhang B, Li X, Chu Q, Wang N, Wang Z, Zhou S, Tong C, Zhang J, Guan W, Cui Z et al (2008) A survey of HIV/AIDS related behaviors among 2250 MSM in nine major cities of China. *Chin J AIDS STD* 14(6):541–547
26. Choi KH, Ning Z, Gregorich SE, Pan QC (2007) The influence of social and sexual networks in the spread of HIV and syphilis among men who have sex with men in Shanghai, China. *J Acquir Immune Defic Syndr* 45(1):77–84
27. Fu LJ, Fang YR, Guo TY (2007) Investigation of the sexual behaviors among the MSM in Shaoxing City of Zhejiang Province. *Dis Surveill* 22(12):818–819
28. van Griensven F, de Lind van Wijngaarden JW, Baral S, Grulich A (2009) The global epidemic of HIV infection among men who have sex with men. *Curr Opin HIV AIDS* 4 (4):300–307
29. Wong FY, Huang ZJ, Wang W, He N, Marzzurco J, Frangos S, Buchholz ME, Young D, Smith BD (2009) STIs and HIV among men having sex with men in China: a ticking time bomb? *AIDS Educ Prev* 21(5):430–446
30. Jin F, Jansson J, Law M, Prestage GP, Zablotska I, Imrie JCG, Kippax SC, Kaldor JM, Grulich AE, Wilson DP (2010) Per-contact probability of HIV transmission in homosexual men in Sydney in the era of HAART. *AIDS* 24:907–913
31. Sullivan PS, Salazar L, Buchbinder S, Sanchez TH (2009) Estimating the proportion of HIV transmissions from main sex partners among men who have sex with men in five US cities. *AIDS* 23(9):1153–1162
32. Davis KR, Weller SC (1999) The effectiveness of condoms in reducing heterosexual transmission of HIV. *Fam Plann Perspect* 31(6):272–279
33. Weller SC, Davis KR (2002) Condom effectiveness in reducing heterosexual HIV transmission. *Cochrane Database Syst Rev* 1(CD003255)
34. Weller SC (1993) A meta-analysis of condom effectiveness in reducing sexually transmitted HIV. *AIDS* 36(12):1635–1644
35. Pinkerton SD, Abramson PR (1997) Effectiveness of condoms in preventing HIV transmission. *Soc Sci Med* 44(9):1303–1312
36. Fitch JT, Stine C, Hager WD, Mann J, Adam MB, McIlhaney J (2002) Condom effectiveness: factors that influence risk reduction. *Sex Transm Dis* 29(12):811–817

37. Chow EPF, Wilson DP, Zhang L (2012) Patterns of condom use among men who have sex with men in China: a systematic review and meta-analysis. *AIDS Behav* 16(3):653–663
38. Chow EPF, Wilson D, Zhang L (2012) Estimating HIV incidence among female partners of bisexual men in China. *Int J Infect Dis* 16(5):e312–320
39. Kault D (1996) The shape of the distribution of the number of sexual partners. *Stat Med* 15(2):221–230
40. Brisson M, Boily MC, Mâsse BR, Adrien A, Léaune V (1999) Highlights of the sexual activity of the heterosexual population in the province of Quebec. *Sex Transm Infect* 75(5):296–299
41. Zhang L, Chow EPF, Wilson DP (2011) Men who have sex with men in China have relatively low numbers of sexual partners. *Infect Dis Rep* 3(e10):46–51
42. Kault D (2000) Uncertainties in predicting the demographic impact of AIDS. *Theor Popul Biol* 57(4):309–324
43. Zeng G, Xiao Y, Xu P, Feng N, Jin CR, Lu F (2009) Evaluation of effect of community-based HIV/AIDS interventions among men who have sex with men in eighteen cities, China. *Zhonghua Yu Fang Yi Xue Za Zhi* 43(11):977–980
44. Zhang D, Bi P, Lv F, Zhang J, Hiller JE (2007) Changes in HIV prevalence and sexual behavior among men who have sex with men in a northern Chinese city: 2002–2006. *J Infect* 55(5):456–463
45. Lau JT, Wang M, Wong HN, Tsui HY, Jia M, Cheng F, Zhang Y, Su X, Wang N (2008) Prevalence of bisexual behaviors among men who have sex with men (MSM) in China and associations between condom use in MSM and heterosexual behaviors. *Sex Transm Dis* 35(4):406–413
46. Rosenberg ES, Sullivan PS, Dinenno EA, Salazar LF, Sanchez TH (2011) Number of casual male sexual partners and associated factors among men who have sex with men: results from the National HIV Behavioral Surveillance system. *BMC Public Health* 11:189
47. Wilson DP, Hoare A, Regan DG, Law MG (2009) Importance of promoting HIV testing for preventing secondary transmissions: modelling the Australian HIV epidemic among men who have sex with men. *Sex Health* 6(1):19–33
48. Crawford JM, Kippax SC, Mao L, Van de Ven PG, Prestage GP, Grulich AE, Kaldor J (2006) Number of risk acts by relationship status and partner serostatus: findings from the HIM cohort of homosexually active men in Sydney, Australia. *AIDS Behav* 10(3):325–331
49. Mao L, Crawford JM, Hospers HJ, Prestage GP, Grulich AE, Kaldor JM, Kippax SC (2006) “Serosorting” in casual anal sex of HIV-negative gay men is noteworthy and is increasing in Sydney, Australia. *AIDS* 20(8):1204–1206
50. Zhou YR (2006) Homosexuality, seropositivity, and family obligations: perspectives of HIV-infected men who have sex with men in China. *Cult Health Sex* 8(6):487–500
51. Gay Literature from China: In: Search of a happy ending [http://www.iias.nl/iiasn/31/IIASN31\\_27.pdf](http://www.iias.nl/iiasn/31/IIASN31_27.pdf)
52. Liu H, Feng T, Ha T, Cai Y, Liu X, Li J (2011) Chinese culture, homosexuality stigma, social support and condom use: a path analytic model. *Stigma Res Action* 1(1):27–35
53. Zhang M, Chu Z, Wang H, Xu J, Lu C, Shang H (2011) A rapidly increasing incidence of HIV and syphilis among men who have sex with men in a major city of China. *AIDS Res Hum Retroviruses* 27(11):1139–1140
54. Ruan Y, Jia Y, Zhang X, Liang H, Li Q, Yang Y, Li D, Zhou Z, Luo F, Shi W et al (2009) Incidence of HIV-1, syphilis, hepatitis B, and hepatitis C virus infections and predictors associated with retention in a 12-month follow-up study among men who have sex with men in Beijing, China. *J Acquir Immune Defic Syndr* 52(5):604–610
55. Shang H, Xu JJ, Zhang M, Brown K, Reilly K, Wang HL, Hu QH, Ding HB, Chu ZX, Bice T (2010) Syphilis and HIV seroconversion among a 12-month prospective cohort of men who have sex with men in Shenyang, China. *Sex Transm Dis* 37(7):432–439

56. Hao C, Yan H, Yang H, Huan X, Guan W, Xu X, Zhang M, Tang W, Wang N, Gu J et al (2011) The incidence of syphilis, HIV and HCV and associated factors in a cohort of men who have sex with men in Nanjing, China. *Sex Transm Infect* 87(3):199–201
57. Yang H, Hao C, Huan X, Yan H, Guan W, Xu X, Zhang M, Tang W, Wang N, Lau JT (2010) HIV incidence and associated factors in a cohort of men who have sex with men in Nanjing, China. *Sex Transm Dis* 37(4):208–213
58. Chow EPF, Wilson DP, Zhang L (2010) The next era of HIV in China: rapidly spreading epidemics among men who have sex with men. *J Acquir Immune Defic Syndr* 55(4):e32–e33, author reply e33–34
59. Zhang F, Haberer JE, Wang Y, Zhao Y, Ma Y, Zhao D, Yu L, Goosby EP (2007) The Chinese free antiretroviral treatment program: challenges and responses. *AIDS* 21(Suppl 8): S143–S148
60. Xiao Y, Sun J, Li C, Lu F, Allen KL, Vermund SH, Jia Y (2010) Prevalence and correlates of HIV and syphilis infections among men who have sex with men in seven provinces in China with historically low HIV prevalence. *J Acquir Immune Defic Syndr* 53(Suppl 1):S66–S73
61. Wang L, Wang N, Li D, Jia M, Gao X, Qu S, Qin Q, Wang Y, Smith K (2009) The 2007 estimates for people at risk for and living with HIV in China: progress and challenges. *J Acquir Immune Defic Syndr* 50(4):414–418
62. Zhu MQ, Zhang BC, Li XF, Shu TX, Wu SW (2005) High risk behaviors related HIV/AIDS and education variables of men who have sex with men (MSM) in China. *Ningxia Med J* 27 (1):27–29
63. Choi KH, Gibson DR, Han L, Guo Y (2004) High levels of unprotected sex with men and women among men who have sex with men: a potential bridge of HIV transmission in Beijing, China. *AIDS Educ Prev* 16(1):19–30
64. Mao M, Lan YJ, Zhou DL, Gu Y, Zhang JX, Xie YQ, Niu J (2008) Frequency of condom use in anal copulation among homosexual men. *J Prev Med Inf* 24(7):493–495
65. He Q, Wang Y, Lin P, Zhang Z-B, Zhao X-X, Xu H-F (2005) KAP study on AIDS among men who have sex with men in Guangzhou, Guangdong province. *Chin J Dis Control Prev* 9 (2):106–108
66. Yang HT, Ding JP, Chen GH, Li XF, Zhang BC, Guan WH (2006) Behavioral features of men who have sex with men in Jiangsu. *Jiangsu Prev Med* 17(1):1–4
67. Li X, Wang BS, Li XF, Cui M, Zhang BC, Chang WH, Xing AH (2006) Survey on HIV/AIDS-related high-risk behaviors among men who have sex with men in Xi'an. *Chin J Public Health* 22(5):520–521
68. Zhang BC, Zeng Y, Xu H, Li XF, Zhou SJ, Li H, Liao LM, Zhang XM (2007) Study on 1389 men who have sex with men regarding their HIV high-risk behaviors and associated factors in mainland China in 2004. *Zhonghua Liu Xing Bing Xue Za Zhi* 28(1):32–36
69. Xu J, Zhang HB, Zheng YJ, Wang J, Zhu YB, Li ZR, Hu ZW (2007) The prevalence of syphilis and HIV infection among young men who have sex with men in Hefei city. *Chin J Behav Med Sci* 16(3):205–207
70. Li X, Shi W, Li D, Ruan Y, Jia Y, Vermund SH, Zhang X, Wang C, Liu Y, Yu M et al (2008) Predictors of unprotected sex among men who have sex with men in Beijing, China. *Southeast Asian J Trop Med Public Health* 39(1):99–108
71. Zhu JL, Zhang HB, Wu HH (2007) High risk sexual behavior and HIV/STD infection rate among 122 MSM from students. *Chin J AIDS STD* 13(4):350–352
72. Sun ZX, Lin SF, Wen MQ (2007) Investigation on STD and AIDS prevalence for men who have sex with men. *Mod Prev Med* 34(21):4130–4132
73. Choi J, Xu QY, Huang YZ, Wang YN, Liu Y (2007) A survey on AIDS knowledge and sexual behaviors among homosexual men in Nanchang city, Jiangxi Province. *China Med Herald* 4 (17):158–168
74. Li N, Wang Z, Sun GQ, Sun DY (2007) Analysis of HIV/AIDS sentinel surveillance among high risk population in Henan province in 2006. *Chin J AIDS STD* 13(5):427–429

75. Ma J, Guo J (2007) Internet survey of high risk sexual behaviors and sexually transmitted diseases among male homosexuals in Tianjin. *Mod Prev Med* 34(20):3928–3931
76. Yang SJ, Mu HT, Li XQ, Li F, Zheng CJ, Zhang ZZ, Duo LK, Li RL, Shi L, Yang CX (2007) Survey of the knowledge, attitude, and practice on AIDS among MSM population in Urumchi. *Mod Prev Med* 34(20):4624–4625
77. Ouyang L, Ding XB, Zhou C, Lu RR (2008) HIV risk sexual behavior among MSM with different sex orientation in Chongqing. *South China J Prev Med* 34(2):16–19, 23
78. Chen SH, Zhou J, Zhu JJ, Li KF (2008) Analysis on ethological characters of MSM in nanning. *Mod Prev Med* 35(5):902–904
79. Feng L, Ding X, Lu R, Liu J, Sy A, Ouyang L, Pan C, Yi H, Liu H, Xu J et al (2009) High HIV prevalence detected in 2006 and 2007 among men who have sex with men in China's largest municipality: an alarming epidemic in Chongqing, China. *J Acquir Immune Defic Syndr* 52 (1):79–85
80. Zhou J, Zhu JJ, Bin H, Zhang L, Yao M, Zhang W, Zhang W, Zhong JY, You QY, Gao L (2008) A survey of HIV/STD, HBV and HCV infections and risk behaviors among MSM in two central districts of Guiyang city. *Chin J AIDS STD* 14(1):47–48, 51
81. Liu H, Cai Y, Rhodes AG, Hong F (2009) Money boys, HIV risks, and the associations between norms and safer sex: a respondent-driven sampling study in Shenzhen, China. *AIDS Behav* 13(4):652–662
82. Zhou SJ, Pan CB, Meng XR, Chen GQ, Zhang BC (2008) The related research on sexual behavior and HIV infection in MSM of Chongqing. *Chin J Infect Control* 7(6):381–384, 380
83. Feng Y, Wu Z, Detels R, Qin G, Liu L, Wang X, Wang J, Zhang L (2010) HIV/STD prevalence among men who have sex with men in Chengdu, China and associated risk factors for HIV infection. *J Acquir Immune Defic Syndr* 53(Suppl 1):S74–S80
84. Zhao H, Wang G, Liu H (2009) The investigation report of male-to-male contact during the voluntary consultation in Daolin District of Harebin. *Guide China Med* 7(10):181–183
85. Feng LG, Ding XB, Xu J, Ou YL, Xu SM, Zheng JQ, Guo XJ, Yang MF, Liu XP (2010) Study on HIV, syphilis and HCV prevalence and its associated factors among internet MSM comparison to non-internet MSM in Chongqing. *J Trop Med* 10(1):78–82
86. Liang L, Chen ZQ, Miao XF, Li BJ, Bai GY, Zhao HR (2009) Investigation on AIDS knowledge and behaviors among men who have sex with men. *Hebei Med J* 31 (19):2654–2655
87. Wang ZJ, Sun L, Ma XJ (2010) Survey on AIDS/STD risk behaviors and prevalence among men who have sex with men in Guangling, Yangzhou, Jiangsu. *J Prev Med* 21(2):4–7
88. Wei C, Ruan S, Zhao J, Yang H, Zhu Y, Raymond HF (2011) Which Chinese men who have sex with men miss out on HIV testing? *Sex Transm Infect* 87(3):225–228
89. Xu JJ, Reilly KH, Lu CM, Ma N, Zhang M, Chu ZX, Wang JJ, Yun K, Shang H (2011) A cross-sectional study of HIV and syphilis infections among male students who have sex with men (MSM) in northeast China: implications for implementing HIV screening and intervention programs. *BMC Public Health* 11(1):287
90. Cai WD, Fen TJ, Tan JQ, Chen L, Shi XD, Chen PL, Jiang LZ, Tao XY (2005) A survey of the characteristics and STD/HIV infection of homosexuality in Shenzhen. *Mod Prev Med* 32 (4):328–330
91. Zhang X, Wang C, Hengwei W, Li X, Li D, Ruan Y, Shao Y (2007) Risk factors of HIV infection and prevalence of co-infections among men who have sex with men in Beijing, China. *AIDS* 21(Suppl 8):S53–S57
92. Xiao Y, Ding X, Li C, Liu J, Sun J, Jia Y (2009) Prevalence and correlates of HIV and syphilis infections among men who have sex with men in Chongqing Municipality, China. *Sex Transm Dis* 36(10):647–656
93. Lan Y, Gu Y, Wang B, Zhou D, Zhang J (2004) Behavioral features of men who have sex with men. *J Sichuan Univ (Med Sci Edn)* 35(3):372–375
94. He Q, Wu X, Han D, Liang X (2008) HIV infection and risk behavior of men having sex with men in Chengdu City from 2004 to 2007. *J Occup Health Damage* 23(4):222–224

95. Tian XB, Ji YN (2006) Analysis on condom use of men who have sex with men in small and medium city. *Chin J Public Health* 22(11):1316–1317
96. Xing JM, Zhang KL, Chen X, Zheng J (2007) Investigation on HIV/AIDS knowledge and sexual behaviors among men who have sex with men in Hunan province. *Chin J Prev Med* 41 (6):511–513
97. Chen SC, Luo Y, Cheng J, Ding JM, Dai YZ, Xu K, Chen KK, Chen WY, Shi SF (2007) Analysis on results of MSM behavioral surveillance on HIV/AIDS. *Dis Surveill* 22 (3):175–177
98. Cai GF, Ma QQ, Pan XH, Fu LJ, Xu WX, Shan XR, Yang Q (2008) HIV/AIDS related knowledge, attitude, practice and HIV/STD infection among MSM in two cities of Zhejiang province. *China Prev Med* 9(6):482–485
99. Wu J, Chen L, Fan HL, Ruan Y (2008) A survey on the prevalence of HIV-1 and syphilis infection and characteristics of sexual behaviors in MSM (men who have sex with men) living in Shanghai. *J Diagn Concepts Pract* 7(3):296–299
100. Duan YW, Zhang HB, Wang XD, Wang Y, Jiang HY, Wang J (2010) Survey on HIV related risk factors among MSM in Mianyang and Yibin. *Chin J Dis Control Prev* 14(12):1189–1192
101. China Global Fund AIDS Program Round 5 (2007) Chongqing Global Fund AIDS Program—Annual Assessment Report
102. Liu LY (2008) Investigation on HIV/AIDS sexual behavior among men who have sex with men in 2006 in city of Mudanjiang. *Med Ani Prev* 24(7):534–535
103. Wang Y, Zhang HB, Gui ZG, Yang HW, Fan J, Zheng YJ, Liu L (2008) Analysis of behavior of MSM population and the biological monitoring result in Mianyang city. *Mod Prev Med* 35 (19):3780–3783, 3789
104. Zeng H, Ding XB, Xu JW, Lu RR, Zhang L, Zhang ZJ, Wang Y (2011) Status and demand on HIV/AIDS control among men who have sex with men in main districts of Chongqing. *Acad J Sec Mil Med Univ* 32(5):494–499
105. Cai YM, Liu H, Pan P, Hong FC, Feng TJ (2009) Survey of HIV/AIDS related knowledge and high risk behavior among men who have sex with men in Shenzhen using respondent-driven sampling. *South China J Prev Med* 35(1):4–7
106. Wen XQ, Jiang JQ, Zhang ZK, Jiang W (2010) Survey of HIV/AIDS knowledge, high risk behaviors and serology among 315 men who have sex with men in guilin in 2008. *China Trop Med* 10(10):1194–1195
107. Wang M, Deng YH, Dong HY, Duan Y (2009) Behavior surveillance among men who have sex with men in Kunming in 2008. *Soft Sci Health* 23(5):593–598
108. Gong CT, Zhang QH (2010) Results of AIDS monitoring of 252 MSM in Quanzhou City. *China Trop Med* 10(12):1496–1497
109. Ji GP, Xu J, Yao H, He JG, Zhang XP, Li XJ (2010) AIDS knowledge, behavior and HIV prevalence among MSM in Anhui, China. *Anhui J Prev Med* 16(5):335–338

# Chapter 7

## The Impact of *Cryptococcus gattii* with a Focus on the Outbreak in North America

Carla J. Walraven, Maximillian Jahng, Gregory C. Davenport, Hallie Rane, and Samuel A. Lee

### 7.1 Introduction

*Cryptococcus gattii* is an emerging infectious disease with an expanding geographic range that gained increased attention during the Vancouver Island outbreak in 1999 [54]. Cases of infection with *C. gattii* were first reported in regions of Africa and Australia [46, 57]. *C. gattii* has been environmentally isolated from eucalyptus and other trees from tropical and subtropical regions [63, 122]. Prior to the 1999 outbreak, *C. gattii* infections were extremely rare in temperate regions of North America [12, 87]. Subsequently, *C. gattii* has been recognized as a cause of outbreak infections in Vancouver Island, British Columbia, and the Pacific Northwest United States. It also has been seen as a cause of low-level sporadic infections in other parts of North America and worldwide. Interestingly, the genotypes of the outbreak *C. gattii* strains differ from the genotypes of the *C. gattii* strains causing sporadic infections.

*Cryptococcus* is an encapsulated heterobasidiomycetous fungi found ubiquitously in the environment, with spores found in the soil and trees. It is the second most common cause of life-threatening fungal disease in humans and animals worldwide [24]. *C. neoformans* is responsible for an estimated 99 % of human

---

C.J. Walraven • G.C. Davenport  
University of New Mexico Health Science Center, Albuquerque, NM, USA

M. Jahng • H. Rane  
Division of Infectious Diseases, New Mexico Veterans Healthcare System,  
1501 San Pedro SE, Mail Code: 111-J, Albuquerque, NM 87108, USA

S.A. Lee, M.D., Ph.D. (✉)  
Division of Infectious Diseases, New Mexico Veterans Healthcare System,  
1501 San Pedro SE, Mail Code: 111-J, Albuquerque, NM 87108, USA

University of New Mexico Health Science Center, Albuquerque, NM, USA  
e-mail: [SamALee@salud.unm.edu](mailto:SamALee@salud.unm.edu)

cryptococcal infections, predominantly in immunocompromised hosts. In contrast, *C. gattii* is responsible for a small minority of human infections, but many cases have occurred in apparently immunocompetent hosts.

*C. neoformans* and *C. gattii* can cause asymptomatic infection or serious invasive disease involving the lungs, central nervous system (CNS), and, less commonly, other sites of infection. The diagnostic approach to *C. neoformans* and *C. gattii* infections is essentially the same; however, clinicians may not suspect cryptococcosis in an immunocompetent host, since it remains unusual. However, morbidity and mortality due to cryptococcosis can be extremely high, particularly with CNS involvement.

In this chapter, we review the global and molecular epidemiology, taxonomy, microbial pathogenesis and immunology, and clinical considerations of this emerging fungal pathogen. In addition, we discuss environmental modeling of the potential ecological niches of *C. gattii*, and speculative measures for avoidance and control.

## 7.2 Microbiology and Taxonomic Classification

*Cryptococcus* is a polyphyletic genus of basidiomycetous fungi currently composed of approximately 40 species, of which *C. neoformans* and *C. gattii* cause the majority of human disease [50]. Initially cryptococcal infections were believed to be caused by a single anamorphic species, *C. neoformans*. However, antigenic differences in the structure of the polysaccharide capsule revealed four separate serotypes (A, B, C, and D) that are found worldwide. As a result, *C. neoformans* was reclassified into two sibling species, *C. neoformans* (serotypes A and D) and *C. gattii* (serotypes B and C) [86]. The phylogenetic clade comprising these species is referred to as the pathogenic *Cryptococcus* species complex.

Further studies have shown that *C. neoformans* and *C. gattii* are phenotypically and genetically divergent. The *C. neoformans* and *C. gattii* genomes have undergone extensive chromosomal rearrangements and high (~15 %) nucleotide sequence divergence. Divergence time estimates place the *C. neoformans*—*C. gattii* speciation event between 34 and 50 million years ago [38, 86, 115, 159]. The two species have significant morphological differences, including differences in ecology, geographic distribution, pathogenicity, virulence, and resistance to antifungal drugs. Notably, unlike *C. neoformans*, *C. gattii* primarily infects immunocompetent rather than immunocompromised hosts [136].

PCR fingerprinting, amplified fragment length polymorphism analysis, and multilocus sequence typing, among other methods, have identified four primary molecular types within the *C. gattii* population, designated as VGI, VGII, VGIII, and VGIV [8, 106]. Each of the four types is monophyletic, meaning that each type is made up of a clade derived from a single common ancestor [7, 10, 11, 42, 52, 159]. VGI is the predominant molecular type isolated clinically, and VGII is the cause of most *C. gattii* infections in North America. The VGIII and VGIV molecular types are more commonly associated with disease in immunocompromised

hosts [17, 105]. Analyses of the four molecular types have shown extensive genetic variation. Genetic and genomic comparison of *C. gattii* isolates show moderate levels of sequence diversity, and divergence time estimates place the divergence of the four molecular types at approximately 12 million years ago [38, 59, 115]. These lines of evidence indicate that the VG1, VGII, VGIII, and VGIV molecular types are genetically distinct and undergoing divergence, and are therefore designated as discrete varieties of *C. gattii*. Debate is still ongoing as to whether the molecular types comprise separate species [17, 38].

The VGII molecular type can be further divided into three monophyletic subtypes, VGIIa, VGIIb, and VGIIc [52, 59]. These subtypes have emerged recently and are responsible for cryptococcal outbreaks in both humans and animals in Australia, Canada, and the United States [20, 59, 80]. The subtypes seem to be undergoing expansion, and there is evidence for mating and recombination in the global VGII population, but not within or between other *C. gattii* populations [20, 22, 52, 115]. The emergence of these subtypes may reflect an expansion of the *C. gattii* VGII molecular type into novel ecological niches.

### 7.3 Global Epidemiology

The geographic distribution and habitat of *C. neoformans* and *C. gattii* vary considerably. Typically, *C. neoformans* has been associated with soil enriched with nitrogenous waste from bird droppings. Birds most commonly associated with *C. neoformans* include pigeons, turkeys, and chickens; however, reports linking disease to the droppings from pet canaries and parrots have also been described. Disease in birds appears to be rare, indicating they may instead serve as vectors for the fungus.

Cryptococcal infections due to *C. gattii* were first described in Australian aborigines residing in the predominantly rural Northern Territory. Due to the high prevalence of cryptococcal disease in this region, it was postulated that an environmental niche must exist [45]. In 1989, a large-scale environmental sampling of air, soil, and vegetation covering all seasons revealed that the appearance of *C. gattii* strongly coincided with the flowering of *Eucalyptus camaldulensis* (river red gum) trees in the springtime. *C. gattii* was isolated from decaying wood and the hollows of Eucalyptus trees. Unlike *C. neoformans*, soil samples with bird droppings failed to demonstrate the presence of *C. gattii*. Furthermore, the geographic distribution of *E. camaldulensis* trees in Australia appeared to correlate with the epidemiologic distribution of cryptococcal disease found in the region [46]. Further investigations demonstrated that other Eucalyptus species also have an ecological association with *C. gattii*. The extensive exportation of these trees from Australia is conjectured to have contributed to the worldwide spread of *C. gattii* to other countries. Worldwide surveillance studies conducted in the late 1990s indicated *C. gattii* was predominantly distributed in the tropical and subtropical regions with an unusually high prevalence found in Australia, Brazil, Cambodia, Hawaii,

Southern California, Mexico, Paraguay, Thailand, Vietnam, Nepal, and countries in Central Africa. Of the isolates tested, serotype B was about five times more prevalent than serotype C, with the majority of *C. gattii* serotype C being found in Southern California [84]. The diagnoses of human *C. gattii* infections in other parts of the world often prompt an environmental investigation to identify the ecological niche. In some cases, no ecologic association was readily determined, including surveillance from Eucalyptus trees in the region [89, 108, 128, 129].

In 1999, an outbreak of *C. gattii* infections emerged in humans and animals on Vancouver Island and expanded to mainland British Columbia, Canada, and has infected over 200 humans to date. It is the largest reported outbreak of *C. gattii* infections worldwide. Between 1999 and 2007, Vancouver Island had an average annual incidence of 25.1 cases per million, which did not vary by season. An environmental investigation of the British Columbia and Pacific Northwest area revealed that *C. gattii* colonizes numerous species of trees, including fir, maple, and oaks, and has been recovered from soil, water, and air samples [39, 55]. Although it is unknown how *C. gattii* spread to nearby regions, additional cases of *C. gattii* infection have been sporadically reported in California, Idaho, Oregon, and Washington. Of the 60 cases reported to the Centers for Disease Control (CDC) by 2010, 88 % of patients did not have a travel history to the Pacific Northwest or another endemic area [41, 99].

## 7.4 Molecular Epidemiology

Prior to the Vancouver Island outbreak, *C. gattii* was believed to be a tropical/subtropical fungus. However, based on the molecular typing of samples collected from around the world, it is now believed that *C. gattii* encompasses a much larger ecological distribution, which has yet to be fully defined. Of the cryptococcal isolates gathered from samples worldwide, including *C. neoformans*, *C. gattii* accounts for 27 % of the distribution [106].

There are four major genotypes of *C. gattii*: VGII, VGIII, and VGIV. These molecular typings have been used in epidemiological surveillances to characterize relationships between environmental and clinical isolates. Molecular analysis of 2,046 clinical and veterinary cryptococcal isolates obtained from 48 countries revealed that VGII (9 %) is the most common subtype, followed by VGIII (7 %), VGIV (3 %), and VGIV (1 %) [106].

The VGII genotype is widely distributed globally. The VGII genotype also has a wide distribution, including North and South America, Asia, Africa, and Australia. Genotypes VGII and VGIII are associated with human outbreaks in Vancouver Island, British Columbia, the Pacific Northwest in the USA, the northern territory of Australia, and the central province of Papua New Guinea [17]. In Australia, the most commonly found subtype is VGII, whereas VGII is more common in North and South America. Of the 60 Pacific Northwest isolates reported to the CDC between 2004 and 2010, over half were VGII (50 % subtype VGIIa, 32 % subtype VGIIc,

and 10 % subtype VGIIb), 5 % were VGI and 3 % were VGIII. *C. gattii* was rarely isolated from patients or environmental isolates in Europe and parts of Asia, including Japan, China, and Thailand [20, 106, 136]. During the Vancouver Island outbreak, the most common environmental and clinical *C. gattii* subtype was VGIIa in 90–95 % of isolates, with 5–10 % of isolates belonging to VGIIb [106]. In Australia, the VGII genotype has been responsible for infections in eastern states, in the southwest portion of Western Australia, and in the northern territory [32]. These VGII isolates appear to have originated from an original sexual recombination event, followed by rapid clonal expansion [23].

In contrast, genotypes VGIII and VGIV appear to cause infection predominantly in immunocompromised patients, including those with HIV/AIDS, and are mainly found in Africa and the Americas [17]. Overall, VGIII isolates have most commonly been reported from South America, with additional reports for Central America, southern Asia, and North America [138]. VGIII *C. gattii* isolates have been identified as the cause of infection in a cohort of HIV/AIDS patients from Southern California [27]. These VGIII isolates are distributed into two groups, VGIIIa and VGIIIb, suggesting that this lineage has been endemic to Southern California for a prolonged time period. In contrast, the outbreak strains in Vancouver Island and the Pacific Northwest, which are highly clonal in nature, appear to have emerged more recently [18]. The VGIV genotype has been reported in Africa, and infrequently in South America, Mexico, and Asia [66, 138]. In a cohort of HIV/AIDS patients infected with Cryptococcus in Botswana and Malawi, 13.3 and 13.7 %, respectively, were due to *C. gattii*, all of which were genotype VGIV [92]. These strains exhibited little genetic diversity, thus suggesting a common lineage. In HIV/AIDS patients in Botswana, 30 % of cryptococcal isolates were *C. gattii* [140]. Of the numerous additional reports of *C. gattii* infection in Africa among HIV/AIDS patients where molecular typing was performed, all *C. gattii* infections were genotype VGIV [66].

## 7.5 Pathogenesis

### 7.5.1 Microbial Pathogenesis

Cryptococcal yeast cells are spherical to ovoid in shape and surrounded by a characteristic polysaccharide capsule, which plays an important role in human pathogenesis. The capsule is composed of approximately 90 % glucuronoxylmannan (GXM) and to a lesser extent galactoxymannan and other mannoproteins and sugars. The structure and content of the polysaccharide capsule vary in thickness depending on the specific strain and growth conditions. It is this polysaccharide capsule which differentiates the serotypes A, B, C, and D of *Cryptococcus* spp.

In the environment and on nutrient agar, yeasts are capsule-deficient or poorly encapsulated with capsules ranging from 2 to 5  $\mu\text{m}$  in diameter. This size is optimal for facilitating inhalation and deposition of yeast into the alveolar space. At

physiologic temperatures of 37 °C, carbon dioxide concentrations, and in the presence of serum, the polysaccharide capsule thickens in size to 30–80 µm in diameter which protects the yeast from phagocytosis. Both GXM and galactoxymannan suppress the innate immune response [156]; however, the mannoproteins can stimulate cell-mediated immunity. Yeasts that have been phagocytosed can shed their capsule, buffering them from degradative enzymes of the macrophage, and survive intracellularly. Once dormant inside the macrophage, the yeast is transported via the lymphatic system to other sites [2, 49].

In *C. neoformans* infections, an impaired immune system permits hematogenous dissemination of the yeast to other organs, including the CNS where it has a predilection for the meninges. *C. gattii* can also disseminate to other organs, including the CNS; however, infection more commonly occurs in immune competent patients and the exact mechanism by which dissemination occurs has not been fully elucidated.

### 7.5.2 Virulence Factors

*C. neoformans* and *C. gattii* produce a group of well-characterized virulence factors that are critical for pathogenesis in the mammalian host, including capsule production, melanin production, thermotolerance (growth at 37 °C), and extracellular secretion of degradative enzymes.

The capsule of *C. gattii* and *C. neoformans* is one of the most formidable virulence factors produced by these species and affords the yeast protection from neutrophilic phagocytosis during mammalian infection, as well as from predation by soil amoeba in the environment [9, 165]. The polysaccharide capsular components are predominantly glucuronoxylomannan (GXM), and to a lesser extent, galactoxylomannan, which may be altered by the yeast according to its current environment [26, 33, 34, 56, 93, 131, 164]. Triggers for capsular size and polysaccharide ratio changes include CO<sub>2</sub> concentrations, pH, iron availability, as well as the presence of neutralizing antibody and complement [4, 61, 154, 162, 165, 166]. The capsule itself inhibits binding and subsequent lysis by complement components, while capsular polysaccharides also have the ability to decrease cytokine production by dendritic cells, thereby altering the response to this pathogen *in vivo* [77, 95, 160]. Those with active disease have measurable circulating concentrations of GXM in their body fluids, and this systemic distribution of polysaccharide is believed to inhibit leukocyte recruitment to the sites of infection by encouraging L-selectin shedding from the surface of the granulocytes [37]. As an additional barrier to phagocytosis, antigens present on the exterior of the *C. gattii* capsule have been shown to inhibit neutrophil migration, while those found on the exterior of *C. neoformans* encourage phagocyte migration [44].

Both *C. gattii* and *C. neoformans* are capable of melanization, as are nearly all *Cryptococcus* species; melanization is a necessary but not sufficient component in pathogenesis [1, 76, 118, 163]. Melanin production is facilitated by the enzyme

laccase working on a phenolic substrate, such as dopamine in the brain [47, 72, 75, 152]. Melanin production allows the yeast to thwart the effects of oxidative burst while inside host macrophages [94, 152]. Laccase itself also directly offers protection from hydroxyl radicals, as does another yeast metabolite trehalose [94]. In contrast to *C. neoformans*, additional enzymes such as trehalose-6-phosphate synthase and the transcription factor Ste12 $\alpha$  regulate melanin production as well as thermotolerance and other virulence factors such as capsule production [116, 127, 130, 161].

Hyphal formation, which is necessary for mating when an appropriate mating type is encountered, also occurs as a stress response to temperature, nitrogen availability, humidity, and light variations [48, 85, 133]. The pseudohyphal form these environmental cues elicit is requisite to sporulation and the formation of infectious propagules, with the added benefit of making the yeast cell more resistant to predation by soil amoeba [112, 113, 155]. *C. gattii* is also capable of producing fibrils, typically when residing on plant surfaces, but also in order to thwart phagocytosis and destruction by human neutrophils [139].

The main factor that differentiates the pathogenic Cryptococcal species from the nonpathogenic is the ability to grow at physiologic temperature. While these yeasts prefer 30 °C, their ability to survive and thrive at 37 °C allows them to be formidable pathogens in both immunosuppressed and immunocompetent hosts.

While many of the changes outlined in this section are applied globally to the resident yeast population and in response to specific stimuli, e.g., temperature, pH, or nitrogen availability, phenotypic switching usually involves a much smaller subset and is a random event that occurs in less than 1 in 10,000 cells [62]. These changes are spontaneous and reversible, and reversion typically occurs at the same frequency. The majority of *C. gattii* colonies, known as the parent, display the mucoid phenotype, with a minority that transition to the variant smooth colony form, both of which have been isolated in clinical samples from meningoencephalitis patients [62]. The visible colony characteristic changes stem from alterations in the capsule expressed by the yeast, which in turn has a dramatic effect on virulence. In vitro studies have shown that mice inoculated with the smooth variant of *C. gattii* NP1 survive significantly longer than mice injected with the mucoid parent phenotype [62, 77]. This finding appears to be due to the fact that the smooth variant elicits a much more robust immune response; however, only the smooth variant, with its reduced polysaccharide capsule, is capable of passing through the blood-brain barrier. Therefore, each phenotype is suited to a specific niche in the host [77].

Both bacteria and yeast, including the pathogens *C. gattii* and *C. neoformans*, make use of biofilm organization to protect themselves from immune attack and limit the susceptibility and damage from antimicrobial compounds such as flucytosine, nystatin, amphotericin B, and the azoles [62, 126]. As seen with *Staphylococcus aureus* and *Candida albicans*, cryptococcal biofilms have been found in ventriculoperitoneal shunts, catheters, and prosthetic joints [5, 13, 102, 103, 123, 150]. *Cryptococcus* creates the exopolysaccharide matrix needed to support the biofilm by secreting the same polysaccharides found in its capsule. The body is capable of inhibiting biofilm formation with specific antibodies that prevent polysaccharide secretion; however, there is evidence that yeast cells able to

escape macrophages do so as microcolonies potentially implicated in biofilm formation, rather than as individual planktonic cells [3, 104].

Many *C. gattii* virulence factors make it a more effective pathogen than *C. neoformans*, hence its ability to infect immunocompetent hosts. The use of enzymes such as phospholipase allows *C. gattii* to digest cell membrane lipids and lung surfactant, increasing its invasive capability in mammalian hosts [31, 136]. However, it has been shown that protease production by *C. gattii* is generally lower than *C. neoformans*, making *C. gattii* in some respects less effective at tissue invasion and degrading host immune components such as clotting factors, complement, and immunoglobulins [15, 29, 109, 132]. *C. neoformans* and *C. gattii* also produce urease, which is hypothesized to aid in the spread from the lungs to the CNS [120, 145], a hallmark of *C. gattii* infection. Once reaching the CNS, specifically the cerebrospinal fluid, *myo*-inositol phosphate synthase and *myo*-inositol oxygenase allow *C. gattii* to use an isomer of inositol as its sole carbon source, while phosphatidyl 4-kinase is required for the yeast to survive the hostile environment of the CSF [60, 69, 90, 100, 141]. When in the environment the enzyme laccase, mentioned above in melanin production, enables *C. gattii* to digest lignin in decaying wood, which would explain its predilection for trees as its environmental niche [81, 136]. *C. gattii* also produces superoxide dismutase, an enzyme that protects the organism from oxidative assaults employed by many effector cells of the immune system and that has been identified as a critical component for growth at 37 °C [58, 111].

The primary reproduction method employed by *C. gattii*, and yeasts in general, is budding, which asexually produces clones [11]. Sexual reproduction, which is initiated by the secretion of either MF $\alpha$  or Mf $\alpha$  pheromone by their respective mating type in times of reduced food sources, yields a spore chain of parental and recombinant spores [82, 83, 134]. Another technique employed by *C. gattii* during times of starvation, and a significant advantage displayed by *C. gattii* during the Vancouver Island outbreak, is the ability for same-sex mating, or haploid fruiting [52, 53, 91]. Sexual reproduction of yeasts normally requires an  $\mathbf{a}$  and an  $\mathbf{\alpha}$  cell; however, the haploid cells of *C. gattii* are able to fuse, reassort, and recombine as if they were the traditional heterogeneous mating types [53, 91]. Sexual reproduction between either same or opposite sex mating pairs not only expands the genetic diversity of the organism, but is also responsible for producing the infectious propagules believed to be the primary infectious form for *C. gattii* and *C. neoformans*.

### 7.5.3 Immunology and Host Factors

*C. neoformans* has been classically viewed as an opportunistic organism, particularly since the advent of the HIV/AIDS era, although clearly it can infect nonimmunocompromised hosts as well. More recently, *C. gattii* has been viewed as a fungal pathogen of healthy hosts, but cases in HIV/AIDS patients and other immunocompromised patients can occur. The specific differences in host response to infection with *C. neoformans* and *C. gattii* remain to be fully elucidated, as there is a relative paucity of studies of the immune response to *C. gattii* infection.

In addition to contributions from innate, complement-mediated, and humoral immunity, the major host response to cryptococcal infection relies upon cell-mediated immunity. Key components of this response include CD4+ and CD8+ T-lymphocytes, natural killer cells, and activated phagocytes which lead to a Th-1 type immune response and granuloma formation [124]. Although Th-1 type responses are essential for surviving cryptococcal infection in mice, Th-2-mediated cytokines appear to be deleterious [110].

The capsule, which is composed primarily of glucuronoxylomannan (GXM), galactoxylomannan (GalXM), and mannoproteins, has a major influence on host immunity. GXM interacts with macrophages, resulting in reduction in superoxide anion production, decreased pro-inflammatory cytokines including TNF- $\alpha$  and IL-1 $\beta$ , and inhibition of IL-12 (reviewed in [107]). The ability of GXM to stimulate toll-like receptor responses varies considerably between *C. neoformans* and *C. gattii*, including strain-specific differences [51].

*C. gattii* causes granulomatous pulmonary and disseminated disease in mouse inhalation models of infection. In a murine pulmonary model of infection, *C. gattii* infection with representative VGIIa, VGIIb, and VGI strains induced less pulmonary neutrophil infiltration and inflammatory cytokine production than *C. neoformans* reference strain H99 [33]. Similarly, a *C. gattii* strain isolated from New Mexico, which was a VGIII genotype, induced less pulmonary inflammation than *C. neoformans* H99 and *C. gattii* strains representing VGIIa and VGIIb molecular subtypes [151]. Thus, it could be hypothesized that *C. gattii* may be able to avoid triggering an effective Th1-type immune response due to evasion of immune recognition.

*C. gattii* replicates intracellularly within macrophages. The intracellular replication rate within murine macrophages in vitro has been correlated with in vivo virulence in mice [96, 149, 157]. *C. gattii* VGIIa isolates from the US and from the Vancouver outbreak are more virulent in mice than VGIIb isolates from either region [19, 33, 114, 117]. VGIIa isolates replicate more quickly in macrophages in vitro and are more virulent than VGIIb isolates in vivo. Furthermore, intracellular replication rate in murine macrophages correlates very highly with the intracellular replication rate in human macrophages, suggesting that there may be a correlation between *C. gattii* genotype and human virulence and disease [149, 157]. VGIIa isolates obtained worldwide also have increased mean in vitro growth and melanin production, and are more virulent in a nasal inhalation model in BALB/c mice compared to the much less frequently isolated VGIIb isolates [114, 117]. Whereas *C. gattii* VGIIa isolates are responsible for the majority of infections in Vancouver Island, British Columbia, and the Pacific Northwest U.S., a novel genotype, VGIIC, has emerged in Oregon as a major cause of infection. This strain is highly clonal, sexually fertile, and like the VGIIa isolates demonstrates a high intracellular proliferation rate and high degree of virulence in a murine pulmonary infection model [19].

Further studies of the differences in host response to infection with *C. neoformans* compared to *C. gattii* are in progress. Although there is clear overlap in the range of host infection between *C. neoformans* and *C. gattii*, the contrasting ability of *C. gattii* to infect immunocompetent hosts will provide a fascinating window into the mechanisms of fungal pathogenesis and host immunity.

### 7.5.4 Transmission

*C. gattii* is an environmental fungus that is not a normal part of the human commensal microbiota and is considered pathogenic if isolated from a sterile site. Human infections are most commonly acquired from inhalation of spores or desiccated yeast cells from environmental exposure. *C. neoformans* and *C. gattii* both form spores as part of their mating cycle; *C. neoformans* spores, and likely *C. gattii* spores, are infectious in a murine inhalation model [155]. Cryptococcal infection can also present as an isolated cutaneous infection due to traumatic isolation, and cases due to *C. gattii* have been reported in the literature. Cryptococcal infection is not transmitted person to person.

Retrospective data collected from culture confirmed *C. gattii* cases in travelers who took short trips (1–7 days) to the Pacific Northwest, US; these cases were recorded from January 1999 to December 2004. Data for seven patients were available in which the onset of symptoms and estimated travel dates could be correlated to provide an estimation of the incubation period of *C. gattii*. The shortest incubation time was 2 months and the longest time was 11 months, with the average incubation time being 6–7 months [97]. Reactivation of asymptomatic infection, as seen in HIV/AIDS with *C. neoformans*, has also been described in one case report after therapy with high-dose corticosteroids [64].

## 7.6 Clinical Overview

### 7.6.1 Clinical Presentation

Although *C. gattii* is typically thought of as a disease of immunocompetent hosts, recent epidemiological surveys in the Pacific Northwest US and Vancouver Island, British Columbia have indicated that *C. gattii* infection has occurred in a wide variety of patients with modest to overt immunosuppression. From the viewpoint of traditional risk factors, recent oral steroid use, chronic lung disease (e.g., emphysema, COPD), malignancy, smoking, diabetes, and other immunosuppressive conditions have been associated with *C. gattii* infection [67, 98]. Cryptococcosis most commonly presents as either a meningitis or meningoencephalitis, or pulmonary infection. In rare cases, isolated cutaneous infections via traumatic inoculation may occur. In the Pacific Northwest outbreak, most cases of *C. gattii* infection presented with respiratory symptoms and pulmonary cryptococcosis [55, 67]. Similarly, in an epidemiologic study of cryptococcal infections in New Zealand and Australia, the majority of immunocompetent patients who were infected with *C. gattii* presented with pulmonary infections [30]. In contrast, immunocompromised patients, particularly those with HIV/AIDS, usually present with meningitis regardless of cryptococcal species. However, the difference in the site of infection may be more dependent on the immune status of the patient, rather than the species of

*Cryptococcus*, as disseminated disease occurs more frequently in immunocompromised patients [78, 137].

Pulmonary infections due to *Cryptococcus* can range from asymptomatic airway colonization to severe, life-threatening pneumonia. The most common symptoms patients present with are cough, dyspnea, and chest pain, often in conjunction with constitutional symptoms of infection such as fever. Pulmonary cryptococcoses are more common in *C. gattii* infections in immunocompetent patients, in contrast to the more common presentation of extensive and diffuse interstitial or alveolar infiltrates caused by *C. neoformans* infection in immunocompromised patients [30].

Patients with CNS infection due to cryptococcal infection most commonly present with signs and symptoms of subacute meningitis, most commonly, headache. More severe symptoms, such as altered mental status, vision changes, cranial nerve palsies, and seizures, may also occur. Intracranial pressure is often increased in CNS disease and the magnitude of elevation is an indicator of disease severity. CNS infections caused by *C. gattii* in immunocompetent patient is more likely to present with cryptococcoses and hydrocephalus compared to immunocompromised patients with *C. neoformans* infections [30] and may present with more severe neurologic sequelae than those infected with *C. neoformans*, and may require prolonged treatment [137].

Overall, there is a clinical impression that *C. gattii* infection may be more aggressive than infection with *C. neoformans*, although comparative data remain limited. In addition, patients may lack traditional risk factors for *C. neoformans*, and thus *C. gattii* may not be suspected, leading to a more advanced presentation and/or leading to delays or failures in accurate diagnosis and treatment.

### 7.6.2 Diagnosis

The definitive diagnosis of *C. gattii* infection relies upon traditional culture from suspected sites of infection. However, most clinical microbiology laboratories do not differentiate between *C. gattii* and *C. neoformans*. A simple, cost-effective method for differentiating *C. gattii* from *C. neoformans* is to culture the specimen on L-canavanine glycine thymol (CGB) agar, a selective media that turns blue in presence of *C. gattii* [87]. In most cases, clinicians must request specific identification from the clinical microbiology laboratory, which would then be referred to a reference laboratory for speciation. In addition, histopathological examination of affected anatomic sites as part of a diagnostic evaluation can reveal fungal organisms, although this approach does not specifically distinguish *C. gattii* from *C. neoformans*.

Serological analysis of CSF fluid and serum, typically utilizing a serum assay for cryptococcal polysaccharide capsule antigen, is highly sensitive and specific for cryptococcal disease. Three different types of tests are available, enzyme immunoassay (EIA), latex agglutination (LA), and lateral flow assay (LFA). Baseline CSF

antigen titers have positive correlation to the burden of yeast in CSF infections, and high titers in the CSF has been associated with worse outcomes in some studies [14]. While the correlation of titers and yeast burden becomes less clear when treatment is started, serial evaluation of titers during treatment may be useful for detection of treatment failures [28]. These serological assays are not able to distinguish between *C. gattii* and *C. neoformans*.

A slide agglutination test, Crypto Check (Latron Laboratories, Tokyo, Japan), was a commercial serotyping test widely used to distinguish between the different *Cryptococcus* serotypes (*C. neoformans* var. *grubii*—A, *C. gattii*—B and C, *C. neoformans* var. *neoformans*—D, and hybrid A and D serotypes—AD). However, this test is no longer available as of 2004 [73].

Several molecular typing methods have been developed to distinguish between the different *Cryptococcus* species, as well to determine molecular genotypes [17]. These methods include: amplified fragment length polymorphism analysis, random amplified polymorphic DNA, polymerase chain reaction restriction fragment length polymorphism analysis, and multilocus sequence typing. Although these approaches can readily distinguish the different species and molecular genotypes, these methods are generally only available from specialized research laboratories, and are not available in the clinical setting.

### 7.6.3 Treatment

Due to a lack of prospective studies, the optimal treatment of infections caused by *C. gattii* in nonimmunosuppressed (non-HIV, nontransplant) patients has not been definitively established. However, most of the recommendations for the treatment of cryptococcosis due to *C. gattii* can be inferred from studies on treatment of cryptococcosis caused by *C. neoformans* in immunocompromised patients, as outlined in the Infectious Diseases Society of America (IDSA) Cryptococcosis guidelines. However, the IDSA guidelines recommend more extensive radiologic studies and follow-up examinations for *C. gattii* infections due to the higher likelihood of cryptococcomas [125].

The standard treatment of CNS cryptococcal infections is induction therapy with high-dose amphotericin B and flucytosine to quickly reduce yeast burden and minimize complications. After successful induction, consolidation therapy with high-dose fluconazole, followed by a long course of maintenance therapy (6–12 months) with a lower dose of fluconazole, is initiated to completely eradicate the organism and prevent relapse. While the recommended induction period can be as short as 2 weeks, the duration should be guided by severity of infection, presence of neurologic complications, and the patient's response to therapy (e.g., clearance of yeast within 2 weeks of starting induction). In the presence of cryptococcomas, the treatment is typically prolonged, with an induction period of 6 weeks or more, and maintenance period of greater than a year. Surgical removal of cryptococcomas may be necessary, especially if they are

causing severe neurologic complications or if they fail to respond to antifungal therapy alone [125]. Severe pulmonary infections should be treated similarly to treatment of CNS infections. Surgical resection of cryptococcomas may be needed. For mild to moderate pulmonary infections, a prolonged course of fluconazole may be sufficient [125].

There have been several in vitro studies that have examined differences between in vitro susceptibilities of *C. gattii* to different antifungals agents, with varying results. Some studies did not find any differences between *C. gattii* and *C. neoformans* in susceptibility to various azoles, amphotericin B, or flucytosine [143, 144], while others have found *C. gattii* to have elevated MICs compared to *C. neoformans* for fluconazole, other triazoles and flucytosine. Amphotericin MICs were similar between *C. gattii* and *C. neoformans* in these studies [35, 36, 79, 147, 153].

Additionally, differential in vitro susceptibilities within the different *C. gattii* molecular genotypes have been described. In vitro testing of *C. gattii* strains indicated that VGII subtypes exhibited significantly higher geometric mean MICs to fluconazole [35, 65, 70, 135, 146] and flucytosine compared to other *C. gattii* genotypes [35, 64, 65, 71, 146]. Most of these studies found amphotericin B MICs to be similar, except for two studies, which found VGII to have higher MICs compared to other *C. gattii* genotypes [64, 65, 70].

However, the therapeutic implications of elevated in vitro MICs are not clear, and the MIC breakpoints that would correlate in vitro susceptibility with clinical outcomes has not been established by the Clinical and Laboratory Standards Institute (CLSI) or European Committee on Antimicrobial Susceptibility Testing (EUCAST), two of the major clinical laboratory standards groups which determine guidelines on susceptibility testing of microbes.

There are limited outcomes data for *C. gattii* infections, as the data has been limited to retrospective analyses of small groups of patients. In an analysis of *C. gattii* infections in the United States, a majority of patients were hospitalized, and 19 of 57 (33 %) of patients died of or with *C. gattii* infections [67]. Multivariate analysis revealed that age  $\geq 50$  years or recent oral steroid use was positively correlated with death. Surprisingly, CNS symptoms at presentation were negatively correlated with death. In the British Columbia, 19 of 218 (8.7 %) patients died with *C. gattii* infections. Patients who died were more likely to be older, be infected with genotype VIIB, and have CNS disease [55]. Further studies are needed to clarify how comorbidities, type of infection, treatment, and genotype may affect patient outcomes in *C. gattii* infections.

#### **7.6.4 Diagnosis and Treatment in Resource Limited Settings**

Management of severe cryptococcal infections (particularly cryptococcal meningitis) in resource limited areas, such as sub-Saharan Africa and Southeast Asia, may be challenging. Depending on the availability of diagnostic testing, trained medical

and laboratory personnel, medical infrastructure, and pharmaceuticals, the approach to diagnosing and treating patients with severe cryptococcal disease can vary greatly [74, 125].

For diagnostic testing, the World Health Organization (WHO) recommends lumbar puncture and serologic cryptococcal antigen testing for HIV positive patients suspected of having cryptococcal meningitis [158]. LFA has great potential for use in resource limited areas as it requires very little training, does not require refrigeration (unlike the latex agglutination), has the fastest turnaround time of 15 min, and is inexpensive (approximately \$2 USD per test). If serologic tests are not available, microscopic examination with India ink or CSF or blood cultures are also options, although their sensitivity depends on burden of disease. They also require technical training to perform.

Amphotericin B and flucytosine, the first-line induction treatment for cryptococcal meningitis, are expensive and not readily available in most resource limited areas. Even if available, many areas may not have the infrastructure and trained personnel able to administer and monitor amphotericin B therapy. Patients on amphotericin B require close monitoring and management of amphotericin B induced adverse reactions such as nephrotoxicity and potassium wasting [74], (Vanselow et al. 2012), [125].

If available, amphotericin B may be given as monotherapy for induction therapy, followed by fluconazole for consolidation and maintenance therapy [125, 158]. Limited evidence suggests that combination of amphotericin B and high-dose fluconazole for induction may be associated with better outcomes than amphotericin B alone [121].

In areas where amphotericin B is not available, high-dose fluconazole is the mainstay therapy for cryptococcal disease. Fortunately, fluconazole is widely available at no charge through the Diflucan® Partnership Program or for low cost as a generic. High-dose fluconazole is initiated as induction and consolidation for at least 10 weeks, followed maintenance therapy with fluconazole at lower doses [125, 158]. If flucytosine is available, it should be added to high-dose fluconazole during the induction phase. Small studies have demonstrated that the combination may be more fungicidal than fluconazole alone and may be associated with fewer deaths [88, 119].

## 7.7 The Pacific Northwest Outbreak in North America

### 7.7.1 Vancouver Island and British Columbia

The dramatic development of the *C. gattii* outbreak in the temperate regions of the Pacific Northwest region of North America is specifically highlighted from an epidemiologic, molecular genetic, and global health perspective in this section as a case study of an emerging infectious disease. Whereas cases of *C. gattii* had almost exclusively occurred in tropical and subtropical regions, a distinct

outbreak has now occurred in a temperate region, in addition to reports of sporadic cases from other temperate regions. As noted previously, *C. gattii* caused an outbreak of human infections, and domestic and wild veterinary infections, in 1999 in Vancouver Island, in North America [68, 80, 142]. From 1999 to 2007 there has been an estimated mean of 218 cases diagnosed annually in British Columbia, with an average incidence of 5.8 cases/million population, and 25.1 cases/million on Vancouver Island [55]. The majority of these cases (73.9 % of total cases) have been reported from Vancouver Island. Overall, cases on Vancouver Island peaked in 2002, with the number of cases increasing in mainland British Columbia. The vast majority of outbreak strains have been of the VGII genotype [80]. Further genetic analyses of these outbreak strains have identified the VGIIa/major genotype as the predominant genotype responsible for these infections. Of note, genetic analyses revealed that most outbreak isolates were of the  $\alpha$  mating type only [53]. This observation and the identification of an  $\alpha/\alpha$  VGIIa/major diploid *C. gattii* isolate from Vancouver Island has led to the hypothesis that this strain arose from  $\alpha/\alpha$  same-sex mating, creating a fertile, hypervirulent strain with a unique genotype as the primary source of this outbreak [16, 52]. The major outbreak strain is identical to a strain isolated from a patient in Seattle more than 3 decades ago, and to a strain from a *E. camaldulensis* tree in San Francisco approximately 2 decades ago [43]. In contrast, a genotype VGIIb/minor isolate which has contributed to a minority of outbreak cases, is identical to a minority strain, also VGII, in Australia which is highly fertile with genetic evidence of sexual recombination [21, 52]. This VGIIb/minor isolate is less virulent than the VGIIa/major strain in a mouse model [52]. Further dissection of the genetics of the VGIIa and VGIIb outbreak strains will help to provide a precise determination of their evolutionary origin.

### 7.7.2 *Spread to US Pacific Northwest*

Subsequently, it has become evident that the *C. gattii* outbreak has spread the Pacific Northwest region of the United States as early as 2005 [20, 148]. Cases of human and veterinary *C. gattii* infection have been reported in both Washington and Oregon state, most of which are suspected to have been acquired locally and not via travel to other endemic areas [39]. A detailed genetic analysis of 14 human and 8 veterinary isolates of *C. gattii*, obtained during 2006–2008, revealed that nearly all of the Washington state isolates were the major VGIIa genotype identical to the outbreak strain in Vancouver Island and British Columbia. However, the Oregon state isolates included a novel VGIIc genotype. Specifically, 12 of 13 isolates from Washington state were represented by the outbreak major genotype VGIIa, and the remaining isolate was a novel VGIII isolate. The nine isolates from Oregon state were represented by VGIIa/major genotype, VGIIb/minor outbreak genotype, and the novel VGIIc genotype. The precise genetic origins and environmental niche of the VGIIc isolates, which are so far confined to Oregon state, remain to be further elucidated.

### 7.7.3 California

The outbreak strain has now been detected as far south as the northern parts of California, although detailed genetic studies have not been published at this time [16]. It is likely that additional cases of the *C. gattii* outbreak strain will be reported due to expanding geographic range of the fungus. However, the vast majority of *C. gattii* strains reported in California consist of the genotypes VGIIIa and VGIIIb endemic to the southern part of the state [17, 18].

## 7.8 Ecological Modeling of *C. gattii* Infection

Prior to 1999, *C. gattii* was not believed to be endemic to the Pacific Northwest region of North America. Most human cases of infection had been described in tropical or subtropical regions around the world. Since 2008, the *C. gattii* outbreak on Vancouver Island and mainland British Columbia reached the highest endemic incidence reported worldwide, with over 240 human cases and 360 animal cases [6, 99]. Consequently, in 2003, *C. gattii* infections became a reportable disease in British Columbia, heightening awareness of the disease among physicians and veterinarians. Since the initial Vancouver Island outbreak, *C. gattii* infections have been reported from surrounding areas including Washington, Oregon, California, and Idaho [41].

To further describe the environmental and ecological niche of *C. gattii* in the area, a concerted environmental surveillance program was implemented. *C. gattii* samples isolated from both human and animal cases were collected, as were environmental samples from the air, soil, trees, freshwater and seawater of surrounding areas during various seasons to determine transient colonization from ecological reservoirs. Samples were also gathered from inanimate objects such as vehicle tires, wheel wells, and shoe bottoms that had recently traversed through endemic areas on Vancouver Island. Environmental sampling of this magnitude is generally rare; however, since *C. gattii* was deemed an emerging pathogen, an extensive database of *C. gattii* has been compiled by the British Columbia Center for Disease Control to enhance the current knowledge of this pathogen in what is now considered an endemic area [40].

Utilizing the *C. gattii* database, researchers have employed ecological niche modeling to help public health officials determine established residency sites of *C. gattii* and anticipate new areas conducive to the expansion of the geographical range of *C. gattii* [101]. These models take into account topography, climate, vegetation, and the presence or absence of pathogens in the environment to predict areas of similar ecologies that may support the growth of the pathogen. Employing the Genetic Algorithm for Rule-set Prediction (GARP) software on the *C. gattii* database of human and animal cases and environmental sampling, the model was trained using 70 % of the dataset and tested on the remaining 30 % for predictive capabilities. This model predicted that the optimal ecological niche for *C. gattii* was

limited to the central and southeastern coast of Vancouver Island, the Gulf Islands, Sunshine Coast, and Vancouver Lower Mainland. Furthermore, isolated areas of Queen Charlotte Islands, British Columbia central coast and southern interior, and west coast of Vancouver Island were predicted to support the establishment of *C. gattii* [101]. The researchers found that the addition of animal surveillance data to the model greatly improved the predictive abilities of the model and suggested that cases of animal cryptococcosis may serve as an early marker for human disease.

Although these predictive ecologic niche models have their limitations, primarily due to lack of data, they are useful in predicting infections in areas outside those considered to be endemic or currently known to be colonized. These models can be used to alert public officials and healthcare providers to the potential for spread of *C. gattii* in their regions and facilitate early recognition and treatment.

## 7.9 Current and Novel Concepts in Control of Cryptococcal Disease

### 7.9.1 Geographical Influence on Transmission

The geographic influence on the spread of *C. gattii* remains unclear. Cyclical and long term changes in climate, changes in land use, and changes in host factors may all have contributed to the emergence of *C. gattii* in Pacific Northwest. Drier and warmer summers, with milder and wet winters seen in the recent past decades may have likely made parts of the Pacific Northwest more hospitable for *C. gattii* to establish new niches. While speculative, the geographical differences between the Pacific Northwest and other endemic regions worldwide may have given rise to a divergent strain of *C. gattii*. Unlike the non-outbreak strains seen in other parts of the world, the dominant strain seen in the Pacific Northwest is a unique VGIIa strain, which arose from the nonclassical same-sex reproduction between two parents with the  $\alpha$  allele of mating-specific type gene [40]. This unique VGIIa strain may be hypervirulent [19, 33, 114, 117], enabling the spread of disease in the region. Also, unlike the *C. gattii* strains seen in Australia, which are classically associated with Eucalyptus trees, the strains found in Pacific Northwest seem to be more readily able to adapt to broader range of environments. The *C. gattii* strains seen in Pacific Northwest are able to thrive in dry, nutrient poor soil and have been found in many different species of trees, and in water [40].

### 7.9.2 Avoidance

There are currently no established or known measures to substantively reduce the risk of acquiring infection with *C. gattii*, as acquisition appears to be sporadic in nature. In contrast, prevention of *C. neoformans* in patients with HIV/AIDS is

clinically well established through rigorous clinical trials using fluconazole prophylaxis in patients with CD4+ counts of less than 50, and represents current standard-of-care [25]. It is likely that in this highly specific patient population, fluconazole will also help prevent infection with *C. gattii*. From a speculative standpoint, highly immunosuppressed patients may benefit from counseling regarding the epidemiology of *C. gattii*, and perhaps avoidance of areas with a preponderance of trees known to harbor *C. gattii* in endemic areas. However, since *C. gattii* can infect apparently healthy individuals, it is unclear what additional specific avoidance measures can be undertaken. There is no vaccine available.

### 7.9.3 Speculative Measures

From an ecological perspective, commercial tree farming using imported eucalyptus trees may pose a risk of further increasing the environmental distribution of *C. gattii*. Depending on the perceived public health risk, surveillance of colonization of imported trees may be warranted. However, it is clear the *C. gattii* has colonized a large variety of trees in the temperature regions of the Pacific Northwest, and comprehensive control is likely to be extremely difficult if not impossible. In order to better define the environmental niche and geographic range of *C. gattii*, the authors are of the opinion that a much broader surveillance of the distribution of *C. gattii* is warranted, although this would clearly require substantial additional use of limited public health resources. If active environmental surveillance is not economically feasible, a much lower cost approach would be to require mandatory reporting of all human and veterinary cases of *C. gattii* to the Centers for Disease Control.

## 7.10 Conclusions

As infections due to *C. gattii* increase both in the United States and worldwide, epidemiologic surveillance will play an important role in identifying new ecological niches or reservoirs for this emerging fungal pathogen. Since not all clinical and reference laboratories are equipped to differentiate between *C. gattii* and *C. neoformans*, a heightened clinical awareness of the disease will be required to prompt further identification and genotyping to track the global distribution of *C. gattii*.

**Acknowledgments** This manuscript is dedicated to the memory of our beloved colleague and friend, Dr. Gregory Davenport. We would like to thank Dr. Sarah Hardison for helpful suggestions for the manuscript.

## References

1. Alvarado-Ramirez E, Torres-Rodriguez JM, Sellart M, Vidotto V (2008) Laccase activity in *Cryptococcus gattii* strains isolated from goats. Rev Iberoam Micol 25:150–153
2. Alvarez M, Casadevall A (2006) Phagosome extrusion and host-cell survival after *Cryptococcus neoformans* phagocytosis by macrophages. Curr Biol 16(21):2161–2165
3. Alvarez M, Saylor C, Casadevall A (2008) Antibody action after phagocytosis promotes *Cryptococcus neoformans* and *Cryptococcus gattii* macrophage exocytosis with biofilm-like microcolony formation. Cell Microbiol 10:1622–1633
4. Bacon BE, Cherniak R, Kwon-Chung KJ, Jacobson ES (1996) Structure of the 0-deacteylated glucuronoxylomannan from *Cryptococcus neoformans* Cap70 as determined by 2D NMR spectroscopy. Carbohydr Res 283:95–110
5. Banerjee U, Gupta K, Venugopal P (1997) A case of prosthetic valve endocarditis caused by *Cryptococcus neoformans* var. *neoformans*. J Med Vet Mycol 35:139–141
6. Bartlett KH, Kidd SE, Kronstad JW (2007) The emergence of *Cryptococcus gattii* in British Columbia and the Pacific Northwest. Curr Fungal Infect Rep 10(1):108–115
7. Boekhout T, Theelen B, Diaz M, Fell JW, Hop WCJ, Abeln ECA, Dromer F, Meyer W (2001) Hybrid genotypes in the pathogenic yeast *Cryptococcus neoformans*. Microbiology 147:891–907
8. Boekhout T, van Belkum A, Leenders AC, Verbrugh HA, Mukamurangwa P, Swinne D, Schefers WA (1997) Molecular typing of *Cryptococcus neoformans*: taxonomic and epidemiological aspects. Int J Syst Bacteriol 47(2):432–442
9. Bolanos B, Mitchell TG (1989) Phagocytosis of *Cryptococcus neoformans* by rat alveolar macrophages. J Med Vet Mycol 27:203–217
10. Bovers M, Hagen F, Kuramae EE, Boekhout T (2008) Six monophyletic lineages identified within *Cryptococcus neoformans* and *Cryptococcus gattii* by multilocus sequence typing. Fungal Genet Biol 45:400–421
11. Bovers M, Hagen F, Boekhout T (2008) Diversity of the *Cryptococcus neoformans*-*Cryptococcus gattii* species complex. Rev Iberoam Micol 25(1):S4–S12
12. Brandt ME, Hutwagner LC, Klug LA, Baughman WS, Rimland D, Cryptococcal Disease Active Surveillance Group et al (1996) Molecular subtype distribution of *Cryptococcus neoformans* in four areas of the United States. J Clin Microbiol 34(4):912–917
13. Braun DK, Janssen DA, Marcus JR, Kauffman CA (1994) Cryptococcal infection of a prosthetic dialysis fistula. Am J Kidney Dis 24:864–867
14. Brouwer AE, Teparrukkul P, Pinpraphaporn S, Larsen RA, Chierakul W, Peacock S, Day N, White NJ, Harrison TS (2005) Baseline correlation and comparative kinetics of cerebrospinal fluid colony-forming unit counts and antigen titers in cryptococcal meningitis. J Infect Dis 192:681–684
15. Brueske CH (1986) Proteolytic activity of a clinical isolate of *Cryptococcus neoformans*. J Clin Microbiol 23(3):631–633
16. Byrnes EJ, Marr KA (2011) The outbreak of *Cryptococcus gattii* in western North America: epidemiology and clinical issues. Curr Infect Dis Rep 13(3):256–261
17. Byrnes EJ, Bartlett KH, Perfect JR, Heitman J (2011) *Cryptococcus gattii*: an emerging fungal pathogen infecting humans and animals. Microb Infect 13:895–907
18. Byrnes EJ, Li W, Ren P, Lewit Y, Voelz K, Fraser JA, Dietrich FS, May RC, Chaturvedi S, Chaturvedi V, Heitman J (2011) A diverse population of *Cryptococcus gattii* molecular type VGIII in southern Californian HIV/AIDS patients. PLoS Pathog 7(9):e1002205
19. Byrnes EJ, Li W, Lewit Y, Ma H, Voelz K, Ren P, Carter DA, Chaturvedi V, Bildfell RJ, May RC, Heitman J (2010) Emergence and pathogenicity of highly virulent *Cryptococcus gattii* genotypes in the northwest United States. PLoS Pathog 6(4):e1000850
20. Byrnes EJ, Bildfell RJ, Frank SA, Mitchell TG, Marr KA, Heitman J (2009) Molecular evidence that the range of the Vancouver Island outbreak of *Cryptococcus gattii* infection has expanded into the Pacific Northwest in the United States. J Infect Dis 199:1081–1086

21. Campbell LT, Currie BJ, Krockenberger M, Malik R, Meyer W, Heitman J, Carter D (2005) Clonality and recombination in genetically differentiated subgroups of *Cryptococcus gattii*. *Eukaryot Cell* 4(8):1403–1409
22. Campbell LT, Fraser JA, Nichols CB, Dietrich FS, Carter D, Heitman J (2005) Clinical and environmental isolates of *Cryptococcus gattii* from Australia that retain sexual fecundity. *Eukaryot Cell* 4:1410–1419
23. Carricande F, Gilgado F, Arthur I, Ellis D, Malik R, van de Wiele N, Robert V, Currie BJ, Meyer W (2011) Clonality and  $\alpha$ - $\alpha$  recombination in the Australian *Cryptococcus gattii* VGII population—an emerging outbreak in Australia. *PLoS One* 6(2):e16936
24. Casadevall A, Perfect JR (1998) *Cryptococcus neoformans*. ASM Press, Washington, DC, USA
25. Centers for Disease Control and Prevention (2009) Guidelines for prevention and treatment of opportunistic infections in HIV-infected adults and adolescents. *MMWR* 58(RR-4):1–207
26. Charlier C, Chretien F, Baudrimont M, Mordelet E, Lortholary O, Dromer F (2005) Capsule structure changes associated with *Cryptococcus neoformans* crossing of the blood–brain barrier. *Am J Pathol* 166:421–432
27. Chaturvedi S, Dyavaiah M, Larsen RA, Chaturvedi V (2005) *Cryptococcus gattii* in AIDS patients, southern California. *Emerg Infect Dis* 11:1686–1692
28. Chayakulkeeree M, Perfect JR (2006) *Cryptococcosis*. *Infect Dis Clin North Am* 20 (3):507–544, v–vi
29. Chen LC, Blank ES, Casadevall A (1996) Extracellular proteinase activity of *Cryptococcus neoformans*. *Clin Diagn Lab Immunol* 3(5):570–574
30. Chen S, Sorrell T, Nimmo G, Speed B, Currie B, Ellis D, Marriott D, Pfeiffer T, Parr D, Byth K, Australasian Cryptococcal Study Group (2000) Epidemiology and host- and variety-dependent characteristics of infection due to *Cryptococcus neoformans* in Australia and New Zealand. *Clin Infect Dis* 31(2):499–508
31. Chen SCA, Wright LC, Golding JC, Sorrell TC (2000) Purification and characterization of secretory phospholipase B, lysophospholipase and lysophospholipase/transacylase from a virulent strain of the pathogenic fungus *Cryptococcus neoformans*. *Biochem J* 347 (2):431–439
32. Chen SCA, Currie BJ, Campbell HM, Fisher DA, Pfeiffer TJ, Ellis DH, Sorrell TC (1997) *Cryptococcus neoformans* var. *gattii* infection in northern Australia: existence of an environmental source other than known host eucalypts. *Trans R Soc Trop Med Hyg* 91:547–550
33. Cheng PY, Sham A, Kronstad JW (2009) *Cryptococcus gattii* isolates from the British Columbia cryptococcosis outbreak induce less protective inflammation in a murine model of infection than *Cryptococcus neoformans*. *Infect Immun* 77:4284–4294
34. Cherniak R, Sundstrom JB (1994) Polysaccharide antigens of the capsule of *Cryptococcus neoformans*. *Infect Immun* 62(5):1507–1512
35. Chong HS, Dagg R, Malik R, Chen S, Carter D (2010) In vitro susceptibility of the yeast pathogen *Cryptococcus* to fluconazole and other azoles varies with molecular genotype. *J Clin Microbiol* 48(11):4115–4120
36. Chowdhary A, Randhawa HS, Sundar G, Kathuria S, Prakash A, Khan Z, Sun S, Xu J (2011) In vitro antifungal susceptibility profiles and genotypes of 308 clinical and environmental isolates of *Cryptococcus neoformans* var. *grubii* and *Cryptococcus gattii* serotype B from north-western India. *J Med Microbiol* 60(7):961–967
37. Clemons KV, Calich VL, Burger E, Filler SG, Grazziutti M, Murphy J, Roilides E, Campa A, Dias MR, Edwards JE Jr, Fu Y, Fernandes-Bordignon G, Ibrahim A, Katsifa H, Lamaignere CG, Meloni-Bruneri LH, Rex J, Savary CA, Xidieh C (2000) Pathogenesis I: interactions of host cells and fungi. *Med Mycol* 38(Suppl 1):99–111
38. D’Souza CA, Kronstad JW, Taylor G, Warren R, Yuen M, Hu G, Jung WH, Sham A, Kidd SE, Tangen K, Lee N, Zeilmaker T, Sawkins J, McVicker G, Shah S, Gneurre S, Griggs A, Zeng Q, Bartlett K, Li W, Wang X, Heitman J, Stajich JE, Fraser JA, Meyer W, Carter D, Schein J, Krzywinski M, Kwon-Chung KJ, Varma A, Wang J, Brunham R, Fyfe M, Ouellette

- BF, Siddiqui A, Marra M, Jones S, Holt R, Birren BW, Galagan JE, Cuomo CA (2011) Genome variation in *Cryptococcus gattii*, an emerging pathogen of immunocompetent hosts. *mBio* 2(1):e00342–10
39. Datta K, Bartlett KH, Baer R, Byrnes E, Galanis E, Heitman J, Hoang L, Leslie MJ, MacDougall L, Magill SS, Morshed MG, Marr KA et al (2009) Spread of *Cryptococcus gattii* into Pacific Northwest region of the United States. *Emerg Infect Dis* 15(8):1185–1191
40. Datta K, Bartlett KH, Marr KA (2009) *Cryptococcus gattii*: emergence in western north America: exploitation of a novel ecological niche. *Interdiscip Perspect Infect Dis* 2009:176532
41. DeBess E, Cieslak PR, Marsden-Haug N, Goldoft M, Wohrle R, Free C, Dykstra E, Nett RJ, Chiller T, Lockhart SR, Harris J (2010) Emergence of *Cryptococcus gattii*—Pacific Northwest, 2004–2010. *MMWR Morb Mortal Wkly Rep* 59(28):865–868
42. Diaz MR, Boekhout T, Kiesling T, Fell JW (2005) Comparative analysis of the intergenic spacer regions and population structure of the species complex of the pathogenic yeast *Cryptococcus neoformans*. *FEMS Yeast Res* 5:129–1140
43. Diaz MR, Boekhout T, Theelen B, Fell JW (2000) Molecular sequence analyses of the intergenic spacer (IGS) associated with rDNA of the two varieties of the pathogenic yeast, *Cryptococcus neoformans*. *Syst Appl Microbiol* 23(4):535–545
44. Dong ZM, Murphy JW (1995) Effects of the two varieties of *Cryptococcus neoformans* cells and culture filtrate antigens on neutrophil locomotion. *Infect Immun* 63:2632–2644
45. Ellis DH (1987) *Cryptococcus neoformans* var. *gattii* in Australia. *J Clin Microbiol* 25 (2):430–431
46. Ellis DH, Pfeiffer TJ (1990) Natural habitat of *Cryptococcus neoformans* var. *gattii*. *J Clin Microbiol* 28(7):1642–1644
47. Erickson T, Liu L, Gueyikian A, Zhu X, Gibbons J, Williamson PR (2001) Multiple virulence factors of *Cryptococcus neoformans* are dependent on VPH1. *Mol Microbiol* 42:1121–1131
48. Erke KH (1976) Light microscopy of basidia, basidiospores, and nuclei in spores and hyphae of *Filobasidiella neoformans* (*Cryptococcus neoformans*). *J Bacteriol* 128:445–455
49. Feldmesser M, Tucker S, Casadevall A (2001) Intracellular parasitism of macrophages by *Cryptococcus neoformans*. *Trends Microbiol* 9(6):273–278
50. Findley K, Rodriguez-Carres M, Metin B, Kroiss J, Fonseca A, Vilgalys R, Heitman J (2009) Phylogeny and phenotypic characterization of pathogenic cryptococcus species and closely related saprobic taxa in the tremellales. *Eukaryot Cell* 8:353–361
51. Fonseca FL, Nohara LL, Cordero RJ, Frases S, Casadevall A, Almeida IC, Nimrichter L, Rodrigues ML (2010) Immunomodulatory effects of serotype B glucuronoxylomannan from *Cryptococcus gattii* correlate with polysaccharide diameter. *Infect Immun* 78(9):3861–3870
52. Fraser JA, Giles SS, Wenink EC, Geunes-Boyer SG, Wright JR, Diezmann S, Allen A, Stajich JE, Dietrich FS, Perfect JR, Heitman J (2005) Same-sex mating and the origin of the Vancouver Island *Cryptococcus gattii* outbreak. *Nature* 437:1360–1364
53. Fraser JA, Subaran RL, Nichols CB, Heitman J (2003) Recapitulation of the sexual cycle of the primary fungal pathogen *Cryptococcus neoformans* var. *gattii*: implications for an outbreak on Vancouver Island, Canada. *Eukaryot Cell* 2:1036–1045
54. Fyfe M, MacDougall L, Romney M, Starr M, Pearce M, Mak S, Mithani S, Kibsey P (2008) *Cryptococcus gattii* infections on Vancouver Island, British Columbia, Canada: emergence of a tropical fungus in a temperate environment. *Can Commun Dis Rep* 34(6):1–12
55. Galanis E, MacDougall L, Kidd S, Morshed M, British Columbia Cryptococcus gattii Working Group (2010) Epidemiology of *Cryptococcus gattii*, British Columbia, Canada, 1999–2007. *Emerg Infect Dis* 16(2):251–257
56. Garcia-Hermoso D, Dromer F, Janbon G (2004) *Cryptococcus neoformans* capsule structure evolution in vitro and during murine infection. *Infect Immun* 72:3359–3365
57. Gatti F, Eeckels R (1970) An atypical strain of *Cryptococcus neoformans* (San Felice) Vuillemin 1894. I. Description of the disease and of the strain. *Ann Soc Belges Med Trop Parasitol Mycol* 50:689–693

58. Giles SS, Batinic-Haberle I, Perfect JR, Cox GM (2005) *Cryptococcus neoformans* mitochondrial superoxide dismutase: an essential link between antioxidant function and high-temperature growth. *Eukaryot Cell* 4:46–54
59. Gillice JD, Schupp JM, Balajee SA, Harris J, Pearson T, Yan Y, Keim P, DeBess E, Marsden-Haug N, Wohrle R, Engelthaler DM, Lockhart SR (2011) Whole genome sequence analysis of *Cryptococcus gattii* from the Pacific Northwest reveals unexpected diversity. *PLoS One* 6: e28550
60. Goulart L, Silva LK, Chiapello L, Silveira C, Crestani J, Masih D, Vainstein MH (2010) *Cryptococcus neoformans* and *Cryptococcus gattii* genes preferentially expressed during rat macrophage infection. *Med Mycol* 48:932–941
61. Granger DL, Perfect JR, Durack DT (1985) Virulence of *Cryptococcus neoformans*. Regulation of capsule synthesis by carbon dioxide. *J Clin Invest* 76:508–516
62. Gupta G, Fries BC (2010) Variability of phenotypic traits in *Cryptococcus* varieties and species and the resulting implications for pathogenesis. *Future Microbiol* 5(5):775–787
63. Hagen F, Boekhout T (2010) The search for the natural habitat of *Cryptococcus gattii*. *Mycopathologia* 170:209–211
64. Hagen F, van Assen S, Luijckx GJ, Boekhout T, Kampinga GA (2010) Activated dormant *Cryptococcus gattii* infection in a Dutch tourist who visited Vancouver Island (Canada): a molecular epidemiological approach. *Med Mycol* 48(3):528–531
65. Hagen F, Illnait-Zaragozi MT, Bartlett KH, Swinne D, Geertsen E, Klaassen CH, Boekhout T, Meis JF (2010) In vitro antifungal susceptibilities and amplified fragment length polymorphism genotyping of a worldwide collection of 350 clinical, veterinary, and environmental *Cryptococcus gattii* isolates. *Antimicrob Agents Chemother* 54(12):5139–5145
66. Harris J, Lockhart S, Chiller T (2012) *Cryptococcus gattii*: where do we go from here? *Med Mycol* 50:113–129
67. Harris JR, Lockhart SR, Debess E, Marsden-Haug N, Goldoft M, Wohrle R, Lee S, Smelser C, Park B, Chiller T (2011) *Cryptococcus gattii* in the United States: clinical aspects of infection with an emerging pathogen. *Clin Infect Dis* 53(12):1188–1195
68. Hoang LM, Maguire JA, Doyle P, Fyfe M, Roscoe DL (2004) *Cryptococcus neoformans* infections at Vancouver Hospital and Health Sciences Centre (1997–2002): epidemiology, microbiology and histopathology. *J Med Microbiol* 53(9):935–940
69. Hu G, Steen BR, Lian T, Sham AP, Tam N, Tangen KL, Kronstad JW (2007) Transcriptional regulation by protein kinase A in *Cryptococcus neoformans*. *PLoS Pathog* 3(3):e42
70. Iqbal N, DeBess EE, Wohrle R, Sun B, Nett RJ, Ahlquist AM, Chiller T, Lockhart SR, *Cryptococcus gattii* Public Health Working Group (2010) Correlation of genotype and in vitro susceptibilities of *Cryptococcus gattii* strains from the Pacific Northwest of the United States. *J Clin Microbiol* 48(2):539–544
71. Iqbal N, Illnait-Zaragozi MT, Bartlett KH, Swinne D, Geertsen E, Klaassen CH, Boekhout T, Meis JF (2010) In vitro antifungal susceptibilities and amplified fragment length polymorphism genotyping of a worldwide collection of 350 clinical, veterinary, and environmental *Cryptococcus gattii* isolates. *Antimicrob Agents Chemother* 54(12):5139–5145
72. Ikeda R, Shinoda T, Morita T, Jacobson ES (1993) Characterization of a phenol oxidase from *Cryptococcus neoformans* var. *neoformans*. *Microbiol Immunol* 37:759–764
73. Ito-Kuwa S, Nakamura K, Aoki S, Vidotto V (2007) Serotype identification of *Cryptococcus neoformans* by multiplex PCR. *Mycoses* 50(4):277–281
74. Jackson A, Hosseinpour MC (2010) Management of cryptococcal meningitis in sub-saharan Africa. *Curr HIV/AIDS Rep* 7(3):134–142
75. Jacobson ES, Emery HS (1991) Temperature regulation of the cryptococcal phenoloxidase. *J Med Vet Mycol* 29:121–124
76. Jain N, Li L, Hsueh YP, Guerrero A, Heitman J, Goldman DL, Fries BC (2009) Loss of allergen 1 confers a hypervirulent phenotype that resembles mucoid switch variants of *Cryptococcus neoformans*. *Infect Immun* 77(7):128–140

77. Jain N, Li L, McFadden DC, Banarjee U, Wang X, Cook E, Fries BC (2006) Phenotypic switching in a *Cryptococcus neoformans* variety *gattii* strain is associated with changes in virulence and promotes dissemination to the central nervous system. *Infect Immun* 74 (2):896–903
78. Kerkerling TM, Duma RJ, Shadomy S (1981) The evolution of pulmonary cryptococcosis: clinical implications from a study of 41 patients with and without compromising host factors. *Ann Intern Med* 94(5):611–616
79. Khan ZU, Randhawa HS, Chehadeh W, Chowdhary A, Kowshik T, Chandy R (2009) *Cryptococcus neoformans* serotype A and *Cryptococcus gattii* serotype B isolates differ in their susceptibilities to fluconazole and voriconazole. *Int J Antimicrob Agents* 33(6):559–563
80. Kidd SE, Hagen F, Tscharke RL, Huynh M, Bartlett KH, Fyfe M, MacDougall L, Boekhout T, Kwon-Chung KJ, Meyer W (2004) A rare genotype of *Cryptococcus gattii* caused the cryptococcosis outbreak on Vancouver Island (British Columbia, Canada). *Proc Natl Acad Sci* 101:17258–17263
81. Kidd SE, Sorrell TC, Meyer W (2003) Isolation of two molecular types of *Cryptococcus neoformans* var. *gattii* from insect frass. *Med Mycol* 41(2):171–176
82. Kwon-Chung KJ (1980) Nuclear genotypes of spore chains in *Filobasidiella neoformans* (*Cryptococcus neoformans*). *Mycologia* 72:418–422
83. Kwon-Chung KJ (1976) Morphogenesis of *Filobasidiella neoformans*, the sexual state of *Cryptococcus neoformans*. *Mycologia* 68:821–833
84. Kwon-Chung KJ, Bennett JE (1984) Epidemiologic differences between the two varieties of *Cryptococcus neoformans*. *Am J Epidemiol* 120(1):123–130
85. Kwong-Chung KJ, Bennett JE (1978) Distribution of  $\alpha$  and  $\alpha$  mating types of *Cryptococcus neoformans* among natural and clinical isolates. *Am J Epidemiol* 108(4):337–340
86. Kwon-Chung KJ, Boekhout T, Fell JW, Diaz M (2002) Proposal to conserve the name *Cryptococcus gattii* against *C. hondurianus* and *C. bacillisporus* (Basidiomycota, Hymenomycetes, Tremellomycetidae). *Taxon* 51:804–806
87. Kwon-Chung KJ, Polacheck I, Bennett JE (1982) Improved diagnostic medium for separation of *Cryptococcus neoformans* var. *neoformans* (serotypes A and D) and *Cryptococcus neoformans* var. *gattii* (serotypes B and C). *J Clin Microbiol* 15:525–537
88. Larsen RA, Bozzette SA, Jones BE, Haghigiat D, Leal MA, Forthal D, Bauer M, Tilles JG, McCutchan JA, Leedom JM (1994) Fluconazole combined with flucytosine for treatment of cryptococcal meningitis in patients with AIDS. *Clin Infect Dis* 19(4):741–745
89. Laurensen IF, Laloo DG, Naragi S, Seaton RA, Trevett AJ, Matuka A, Kevau IH (1997) *Cryptococcus neoformans* in Papua New Guinea: a common pathogen but an elusive source. *J Med Vet Mycol* 35(6):437–440
90. Lee A, Toffaletti DL, Tenor J, Soderblom EJ, Thompson JW, Moseley MA, Price M, Perfect JR (2010) Survival defects of *Cryptococcus neoformans* mutants exposed to human cerebrospinal fluid result in attenuated virulence in an experimental model of meningitis. *Infect Immun* 78:4213–4225
91. Lin X, Hull CM, Heitman J (2005) Sexual reproduction between partners of the same mating type in *Cryptococcus neoformans*. *Nature* 434(7036):1017–1021
92. Litvintseva AP, Thakur R, Reller LB, Mitchell TG (2005) Prevalence of clinical isolates of *Cryptococcus gattii* serotype C among patients with AIDS in sub-Saharan Africa. *J Infect Dis* 192(5):888–892
93. Littman ML, Tsubura E (1959) Effect of degree of encapsulation upon virulence of *Cryptococcus neoformans*. *Proc Soc Exp Biol Med* 101:773–777
94. Liu L, Tewari RP, Williamson PR (1999) Laccase protects *Cryptococcus neoformans* from antifungal activity of alveolar macrophages. *Infect Immun* 67:6034–6039
95. Lupo P, Chang YC, Kelsall BL, Farber JM, Pietrella D, Vecchiarelli A, Leon F, Kwon-Chung KJ (2008) The presence of capsule in *Cryptococcus neoformans* influences the gene expression profile in dendritic cells during interaction with the fungus. *Infect Immun* 76 (4):1581–1589

96. Ma H, Hagen F, Stekel DJ, Johnston SA, Sionov E, Falk R, Polacheck I, Boekhout T, May RC (2009) The fatal fungal outbreak on Vancouver Island is characterized by enhanced intracellular parasitism driven by mitochondrial regulation. *Proc Natl Acad Sci USA* 106(31):12980–12985
97. MacDougall L, Fyfe M (2006) Emergence of *Cryptococcus gattii* in a novel environment provides clues to its incubation period. *J Clin Microbiol* 44(5):1581–1582
98. MacDougall L, Fyfe M, Romney M, Starr M, Galanis E (2011) Risk factors for *Cryptococcus gattii* infection, British Columbia, Canada. *Emerg Infect Dis* 17(2):193–199
99. MacDougall L, Kidd SE, Galanis E, Mak S, Leslie MJ, Cieslak PR, Kronstad JW, Morshed MG, Bartlett KH (2007) Spread of *Cryptococcus gattii* in British Columbia, Canada, and detection in the Pacific Northwest, USA. *Emerg Infect Dis* 13(1):42–50
100. Mackenzie EA, Klig LS (2008) Computational modeling and in silico analysis of differential regulation of myo-inositol catabolic enzymes in *Cryptococcus neoformans*. *BMC Mol Biol* 9:88. doi:[10.1186/1471-2199-9-88](https://doi.org/10.1186/1471-2199-9-88)
101. Mak S, Klinkenberg B, Bartlett K, Fyfe M (2010) Ecological niche modeling of *Cryptococcus gattii* in British Columbia, Canada. *Environ Health Perspect* 118(5):653–658
102. Martinez LR, Casadevall A (2006) *Cryptococcus neoformans* cells in biofilms are less susceptible than planktonic cells to antimicrobial molecules produced by the innate immune system. *Infect Immun* 74:6118–6123
103. Martinez LR, Casadevall A (2006) Susceptibility of *Cryptococcus neoformans* biofilms to antifungal agents in vitro. *Antimicrob Agents Chemother* 50:1021–1033
104. Martinez LR, Christaki E, Casadevall A (2006) Specific antibody to *Cryptococcus neoformans* glucuronoxylomannan antagonizes antifungal drug action against cryptococcal biofilms in vitro. *J Infect Dis* 194:261–266
105. Meyer W, Trilles L (2010) Genotyping of the *Cryptococcus neoformans/C. gattii* species complex. *Aust Biochem* 41:11–15
106. Meyer W, Gilfado F, Ngamskulrungroj P, Trilles L, Hagen F, Castaneda E, Boekhout T (2011) Molecular typing of the *Cryptococcus neoformans/Cryptococcus gattii* species complex. In: Heitman J, Kozel TJ, Kwon-Chung KJ, Perfect JR, Casadevall A (eds) *Cryptococcus: from human pathogen to model yeast*. ASM Press, Washington, DC, USA
107. Monari C, Bistoni F, Vecchiarelli A (2006) Glucuronoxylomannan exhibits potent immunosuppressive properties. *FEMS Yeast Res* 6(4):537–542
108. Montagna MT, Viviani MA, Pulito A, Aralla C, Tortorano AM, Fiore L, Barbuti S (1997) *Cryptococcus neoformans* var. *gattii* in Italy. Note II. Environmental investigation related to an autochthonous clinical case in Apulia. *J Mycol Med* 7:93–96
109. Muller HE, Sethi KK (1972) Proteolytic activity of *Cryptococcus neoformans* against human plasma proteins. *Med Microbiol Immunol* 158(2):129–134
110. Muller U, Stenzel W, Kohler G, Werner C, Polte T, Hansen G, Schutze N, Straubinger RK, Blessing M, McKenzie AN, Brombacher F, Alber G (2007) IL-13 induces disease-promoting type 2 cytokines, alternatively activated macrophages and allergic inflammation during pulmonary infection of mice with *Cryptococcus neoformans*. *J Immunol* 179:5367–5377
111. Narasipura SD, Chaturvedi V, Chaturvedi S (2005) Characterization of *Cryptococcus neoformans* variety *gattii* SOD2 reveals distinct roles of the two superoxide dismutases in fungal biology and virulence. *Mol Microbiol* 55(6):1782–1800
112. Neilson JB, Frontling RA, Blumer GS (1981) Pseudohyphal forms of *Cryptococcus neoformans*: decreased survival in vivo. *Mycopathologia* 73:57–59
113. Neilson JB, Ivey MH, Bulmer GS (1978) *Cryptococcus neoformans* pseudohyphal forms surviving culture with *Acanthamoeba polyphaga*. *Infect Immun* 20:262–266
114. Ngamskulrungroj P, Serena C, Gilgado F, Malik R, Meyer W (2011) Global VGIIa isolates are of comparable virulence to the major fatal *Cryptococcus gattii* Vancouver Island outbreak genotype. *Clin Microbiol Infect* 17(2):251–258

115. Ngamskulrungroj P, Gilgado F, Faganello J, Litvintseva AP, Leal AL, Tsui KM, Mitchell TG, Vainstein MH, Meyer W (2009) Genetic diversity of the *Cryptococcus* species complex suggests that *Cryptococcus gattii* deserves to have varieties. *PLoS One* 4:e5862
116. Ngamskulrungroj P, Himmelreich U, Breger JA, Wilson C, Chayakulkeeree M, Krockenberger MB, Malik R, Daniel HM, Toffaletti D, Djordjevic JT, Mylonakis E, Meyer W, Perfect JR (2009) The trehalose synthesis pathway is an integral part of the virulence composite for *Cryptococcus gattii*. *Infect Immun* 77:4584–4596
117. Ngamskulrungroj P, Serena C, Gilgado F, Malik R, Meyer W (2011) Global VGIIa isolates are of comparable virulence to the major fatal *Cryptococcus gattii* Vancouver Island outbreak genotype. *Clin Microbiol Infect* 17(2):251–258
118. Nosanchuk JD, Casadevall A (1997) Cellular charge of *Cryptococcus neoformans*: contributions from the capsular polysaccharide, melanin, and monoclonal antibody binding. *Infect Immun* 65:1836–1841
119. Nussbaum JC, Jackson A, Namarika D, Phulusa J, Kenala J, Kanyemba C, Jarvis JN, Jaffar S, Hosseiniour MC, Kamwendo D, van der Horst CM, Harrison TS (2010) Combination flucytosine and high-dose fluconazole compared with fluconazole monotherapy for the treatment of cryptococcal meningitis: a randomized trial in Malawi. *Clin Infect Dis* 50(3):338–344
120. Olszewski MA, Noverr MC, Chen GH, Toews GB, Cox GM, Perfect JR, Huffnagle GB (2004) Urease expression by *Cryptococcus neoformans* promotes microvascular sequestration, thereby enhancing central nervous system invasion. *Am J Pathol* 164(5):1761–1771
121. Pappas PG, Chetchotisakd P, Larsen RA, Manosuthi W, Morris MI, Anekthananon T, Sungkanuparph S, Supparatpinyo K, Nolen TL, Zimmer LO, Kendrick AS, Johnson P, Sobel JD, Filler SG (2009) A phase II randomized trial of amphotericin B alone or combined with fluconazole in the treatment of HIV-associated cryptococcal meningitis. *Clin Infect Dis* 48(12):1775–1783
122. Pfeiffer T, Ellis D (1991) Environmental isolation of *Cryptococcus neoformans gattii* from California. *J Infect Dis* 163:929–930
123. Penk A, Pittrow L (1999) Role of fluconazole in the long-term suppressive therapy of fungal infections in patients with artificial implants. *Mycoses* 42:91–96
124. Perfect JR (2006) *Cryptococcus neoformans*: a sugar-coated killer. In: Heitman J, Filler S, Edwards J, Mitchell A (eds) *Molecular principles of fungal pathogenesis*. ASM Press, Washington, DC
125. Perfect JR, Dismukes WE, Dromer F, Goldman DL, Graybill JR, Hamill RJ, Harrison TS, Larsen RA, Lortholary O, Nguyen MH, Pappas PG, Powderly WG, Singh N, Sobel JD, Sorrell TC (2010) Clinical practice guidelines for the management of cryptococcal disease: 2010 update by the Infectious Diseases Society of America. *Clin Infect Dis* 50(3):291–322
126. Pettit RK, Repp KK, Hazen KC (2010) Temperature affects the susceptibility of *Cryptococcus neoformans* biofilms to antifungal agents. *Med Mycol* 48(2):421–426
127. Petzold EW, Himmelreich U, Mylonakis E, Rude T, Toffaletti D, Cox GM, Miller JL, Perfect JR (2006) Characterization and regulation of the trehalose synthesis pathway and its importance in the pathogenicity of *Cryptococcus neoformans*. *Infect Immun* 74:5877–5887
128. Pfeiffer TJ, Ellis DH (1992) Environmental isolation of *Cryptococcus neoformans* var. *gattii* from *Eucalyptus tereticornis*. *J Med Vet Mycol* 30:407–408
129. Pfeiffer TJ, Ellis DH (1991) Environmental isolation of *Cryptococcus neoformans* var. *gattii* from California. *J Infect Dis* 163:929–930
130. Ren P, Springer DJ, Behr MJ, Samsonoff WA, Chaturvedi S, Chaturvedi V (2006) Transcription factor STE12 $\alpha$  has distinct roles in morphogenesis, virulence, and ecological fitness of the primary pathogenic yeast *Cryptococcus gattii*. *Eukaryot Cell* 5:1065–1080
131. Rivera J, Feldmesser M, Cammer M, Casadevall A (1998) Organ-dependent variation of capsule thickness in *Cryptococcus neoformans* during experimental murine infection. *Infect Immun* 66:5027–5030

132. Ruma-Haynes P, Brownlee AG, Sorrell TC (2000) A rapid method for detecting extracellular proteinase activity in *Cryptococcus neoformans* and a survey of 63 isolates. *J Med Microbiol* 49(8):733–737
133. Rutherford JC, Lin X, Nielsen K, Heitman J (2008) Amt2 permease is required to induce ammonium-responsive invasive growth and mating in *Cryptococcus neoformans*. *Eukaryot Cell* 7:237–246
134. Shen WC, Davidson RC, Cox GM, Heitman J (2002) Pheromones stimulate mating and differentiation via paracrine and autocrine signaling in *Cryptococcus neoformans*. *Eukaryot Cell* 1:366–377
135. Silva DC, Martins MA, Szczesz MW, Bonfetti LX, Matos D, Melhem MS (2012) Susceptibility to antifungal agents and genotypes of Brazilian clinical and environmental *Cryptococcus gattii* strains. *Diagn Microbiol Infect Dis* 72(4):332–339
136. Sorell TC (2001) *Cryptococcus neoformans* variety *gattii*. *Med Mycol* 39:155–168
137. Speed B, Dunt D (1995) Clinical and host differences between infections with the two varieties of *Cryptococcus neoformans*. *Clin Infect Dis* 21(1):28–34
138. Springer DJ, Chaturvedi V (2010) Projecting global occurrence of *Cryptococcus gattii*. *Emerg Infect Dis* 16(1):14–20
139. Springer DJ, Ren P, Raina R, Dong Y, Behr MJ, McEwen BF, Bowser SS, Samsonoff WA, Chaturvedi S, Chaturvedi V (2010) Extracellular fibrils of pathogenic yeast *Cryptococcus gattii* are important in ecological niche colonization and mammalian virulence. *PLoS One* 5: e10978
140. Steele KT, Thakur R, Nthobatsang R, Steenhoff AP, Bisson GP (2010) In-hospital mortality of HIV-infected cryptococcal meningitis patients with *C. gattii* and *C. neoformans* infection in Gaborone, Botswana. *Med Mycol* 48(8):1112–1115
141. Steen BR, Zuyderduyn S, Toffaletti DL, Marra M, Jones SJ, Perfect JR, Kronstad J (2003) *Cryptococcus neoformans* gene expression during experimental cryptococcal meningitis. *Eukaryot Cell* 2:1336–1349
142. Stephen C, Lester S, Black W, Fyfe M, Raverty S (2002) Multispecies outbreak of cryptococcosis on southern Vancouver Island, British Columbia. *Can Vet J* 43(10):792–794
143. Tay ST, Tanty Haryanti T, Ng KP, Rohani MY, Hamimah H (2006) In vitro susceptibilities of Malaysian clinical isolates of *Cryptococcus neoformans* var. *grubii* and *Cryptococcus gattii* to five antifungal drugs. *Mycoses* 49(4):324–330
144. Thompson GR, Wiederhold NP, Fothergill AW, Vallor AC, Wickes BL, Patterson TF (2009) Antifungal susceptibilities among different serotypes of *Cryptococcus gattii* and *Cryptococcus neoformans*. *Antimicrob Agents Chemother* 53(1):309–311
145. Torres-Rodríguez JM, Alvarado-Ramírez E, Gutiérrez-Gallego R (2008) Urease activity in *Cryptococcus neoformans* and *Cryptococcus gattii*. *Rev Iberoam Micol* 25(1):27–31
146. Trilles L, Meyer W, Wanke B, Guarro J, Lazéra M (2012) Correlation of antifungal susceptibility and molecular type within the *Cryptococcus neoformans/C. gattii* species complex. *Med Mycol* 50(3):328–332
147. Trilles L, Fernández-Torres B, Lazéra Mdos S, Wanke B, Guarro J (2004) In vitro antifungal susceptibility of *Cryptococcus gattii*. *J Clin Microbiol* 42(10):4815–4817
148. Upton A, Fraser JA, Kidd SE, Bretz C, Bartlett KH, Heitman J, Marr KA (2007) First contemporary case of human infection with *Cryptococcus gattii* in Puget Sound: evidence for spread of the Vancouver Island outbreak. *J Clin Microbiol* 45(9):3086–3088
149. Voelz K, Lammas DA, May RC (2009) Cytokine signaling regulates the outcome of intracellular macrophage parasitism by *Cryptococcus neoformans*. *Infect Immun* 77 (8):3450–3457
150. Walsh TJ, Schlegel R, Moody MM, Costerton JW, Salcman M (1986) Ventriculoatrial shunt infection due to *Cryptococcus neoformans*: an ultrastructural and quantitative microbiological study. *Neurosurgery* 18:373–375

151. Walraven CJ, Gerstein W, Hardison SE, Wormley F, Lockhart SR, Harris JR, Fothergill A, Wickes B, Gober-Wilcox J, Massie L, Ku TS, Firacative C, Meyer W, Lee SA (2011) Fatal disseminated *Cryptococcus gattii* infection in New Mexico. *PLoS One* 6(12):e28625
152. Williamson PR (1994) Biochemical and molecular characterization of the diphenol oxidase of *Cryptococcus neoformans*: identification as a laccase. *J Bacteriol* 176:656–664
153. Varma A, Kwon-Chung KJ (2010) Heteroresistance of *Cryptococcus gattii* to fluconazole. *Antimicrob Agents Chemother* 54(6):2303–2311
154. Vartivarian SE, Anaissie EJ, Cowart RE, Sprigg HA, Tingler MJ, Jacobson ES (1993) Regulation of cryptococcal capsular polysaccharide by iron. *J Infect Dis* 167:186–190
155. Velagapudi R, Hsueh YP, Geunes-Boyer S, Wright JR, Heitman J (2009) Spores as infectious propagules of *Cryptococcus neoformans*. *Infect Immun* 77:4345–4355
156. Vecchiarelli A, Pericolini E, Gabrielli E, Chow SK, Bistoni F, Cenci E, Casadevall A (2011) *Cryptococcus neoformans* galactoxylomannan is a potent negative immunomodulator, inspiring new approaches in anti-inflammatory immunotherapy. *Immunotherapy* 3(8):997–1005
157. Voelz K, Lammas DA, May RC (2009) Cytokine signaling regulates the outcome of intracellular macrophage parasitism by *Cryptococcus neoformans*. *Infect Immun* 77(8):3450–3457
158. World Health Organization (WHO) (2011) Rapid advice: diagnosis, prevention and management of cryptococcal disease in HIV-infected adults, adolescents and children. [http://whqlibdoc.who.int/publications/2011/9789241502979\\_eng.pdf](http://whqlibdoc.who.int/publications/2011/9789241502979_eng.pdf). Accessed 11/01/2013
159. Xu J, Vilgalys R, Mitchell TG (2000) Multiple gene genealogies reveal recent dispersion and hybridization in the human pathogenic fungus *Cryptococcus neoformans*. *Mol Ecol* 9:1471–1481
160. Young BJ, Kozel TR (1993) Effects of strain variation, serotype, and structural modification on kinetics for activation and binding of C3 to *Cryptococcus neoformans*. *Infect Immunol* 61(7):2966–2972
161. Yue C, Cavallo LM, Alspaugh JA, Wang P, Cox GM, Perfect JR, Heitman J (1999) The STE12 $\alpha$  homolog is required for haploid filamentation but largely dispensable for mating and virulence in *Cryptococcus neoformans*. *Genetics* 153:1601–1615
162. Zaragoza O, Casadevall A (2004) Experimental modulation of capsule size in *Cryptococcus neoformans*. *Biol Proced Online* 6:10–15
163. Zaragoza O, Rodrigues ML, De Jesus M, Frases S, Dadachova E, Casadevall A (2009) The capsule of the fungal pathogen *Cryptococcus neoformans*. *Adv Appl Microbiol* 68(8):133–216
164. Zaragoza O, Telzak A, Bryan RA, Dadachova E, Casadevall A (2006) The polysaccharide capsule of the pathogenic fungus *Cryptococcus neoformans* enlarges by distal growth and is rearranged during budding. *Mol Microbiol* 59:67–83
165. Zaragoza O, Fries BC, Casadevall A (2003) Induction of capsule growth in *Cryptococcus neoformans* by mammalian serum and CO<sub>2</sub>. *Infect Immun* 71:6155–6164
166. Zaragoza O, Taborda CP, Casadevall A (2003) The efficacy of complement-mediated phagocytosis of *Cryptococcus neoformans* is dependent on the location of C3 in the polysaccharide capsule and involves both direct and indirect C3-mediated interactions. *Eur J Immunol* 33:1957–1967

# Chapter 8

## Evaluating the Evolutionary Dynamics of Viral Populations

Lars Steinbrück and Alice Carolyn McHardy

### 8.1 Introduction

Phylogenetic analysis allows the inference of evolutionary relationships from a set of genetic sequences, which may represent a distinct species or a genetic region of individuals of a population. For populations of rapidly evolving organisms, the evolutionary and epidemiological processes may occur on similar timescales. Newly developed analytical methods, known as phylodynamic techniques, allow the joint analysis of the genetic and epidemiological relationships of the underlying data [1, 2]. Based on epidemiological information, such as sampling locations or sampling times, phylodynamic methods enable the geographic migration patterns of individuals of a population to be studied, tracking viral spread across host tissues, searching for genetic sites subject to purifying or positive selection associated with adaptation, dating past evolutionary events, and gaining insights into population-level processes using coalescence analysis. In [3], for example, the migration paths of the highly pathogenic avian influenza A (H5N1) virus across Asia are inferred with a “phylogeographic” approach from genetic sequences and geographic sampling locations. Other studies revealed that chimpanzees serve as a natural reservoir

---

Adapted from: Steinbrück L, McHardy AC (2011) Allele dynamics plots for the study of evolutionary dynamics in viral populations. *Nucleic Acids Res* 39(1):e4, by permission of Oxford University Press.

L. Steinbrück  
Max Planck Institut Informatik, Saarbrücken, Germany

A.C. McHardy (✉)  
Max-Planck Research Group for Computational Genomics and Epidemiology, Max-Planck Institute for Informatics, University Campus E1 4, 66123 Saarbrücken, Germany

Department for Algorithmic Bioinformatics, Heinrich-Heine-University, Universitätsstr. 1, 40225 Düsseldorf, Germany  
e-mail: [mchardy@mpi-inf.mpg.de](mailto:mchardy@mpi-inf.mpg.de)

for pandemic and nonpandemic HIV type 1 [4] based on “phylogeographic” clustering, and identified the epidemic history and geographic source of HIV type 2 based on a molecular clock analysis of dated genetic sequences [5].

We describe a method for analyzing the population-level phylodynamics of a gene, which we call allele dynamics plots (AD-plots). AD-plots combine information from phylogenetic inference and ancestral character state reconstruction with isolate sampling times for the analysis of population-level evolutionary dynamics. Furthermore, we use the AD-plot of a population-level sequence sample to identify the alleles that might be associated with a selective advantage. Based on this, we demonstrate how AD-plots can be used to study evolutionary dynamics and to identify emerging viral strains with the example of two influenza A viruses: the human influenza A (H3N2) and the 2009 swine-origin influenza A (H1N1) viruses.

In research into the evolution of the influenza virus, a method that enables the identification of alleles under selection is to count the number of amino acid changes within a protein at sites under selection, which, in turn, can be identified based on the ratio of nonsynonymous to synonymous mutations (dN/dS) [6]. A recent study suggests, however, that dN/dS ratios may not always be informative with regards to detecting selection within a population. Moreover, the method is lacking in sensitivity when applied to individual sequence sites [7]. A different approach was proposed by Pond et al., who introduced a phylogenetic maximum likelihood test based on a protein evolution model to test for directional evolution at individual sites of an alignment [8, 9]. Further related methods quantify the impact of “key innovations” in species trees, e.g., what would happen if lineages that have acquired a beneficial feature were able to spread faster than others. These methods incorporate clade sizes and shifts in diversification rates identified from the phylogenetic tree based on likelihood estimators in the analysis. For an overview, see [10]. However, these methods were conceived for species-level and not population-level analysis, and to evaluate macro-evolution. The method we describe here does not use dN/dS information and is designed for the analysis of longitudinally sampled population-level sequence data. In this sense, it complements the existing approaches.

### 8.1.1 *Background on Influenza A Viruses*

The influenza virus is a rapidly evolving pathogen that is suited for the application of phylodynamic techniques. The single-stranded negative-sense RNA viruses of the family *Orthomyxoviridae* are a major health risk in modern life, responsible for up to 500,000 deaths annually [11]. Three distinct genera (types A, B, and C) are endemic in the human population. Types B and C evolve slowly and circulate at low levels. However, through rapid evolution of the antibody-binding (epitope) sites of the surface proteins, influenza A continuously evades host immunity from previous infection or vaccination, and regularly causes large epidemics. Influenza A viruses can furthermore be distinguished based on the surface proteins hemagglutinin (HA) and neuraminidase (NA). For type A viruses, 16 known subtypes of HA and 9 of NA occur in various combinations in aquatic birds [12]. In the human population,

influenza A viruses of the subtypes H3N2 and H1N1 currently circulate. Of these, the swine-origin influenza A (H1N1) virus (“swine flu”), which entered the human population in 2009, is currently responsible for the majority of infections [13, 14].

Human influenza A viruses continuously change antigenically in a process known as antigenic drift. This refers to the successive fixation of mutations that affect viral fitness by increasing a virus’ ability to circumvent host immunity and protective antibodies elicited by previously circulating viral variants [6, 15]. Antigenically relevant changes are located mainly in the epitope sites of the viral HA [16–19]. Influenza viruses also have a segmented genome composed of eight distinct segments and can evolve by means of reassortment. In segment reassortment, new viral strains are generated, which can inherit genomic segments from two distinct viruses simultaneously infecting the same host cell. This mechanism can affect antigenic evolution, as segments encoding antigenically novel surface proteins, but which are harbored by viruses with low overall fitness due to other reasons, can thus be transferred into a more favorable genetic context and subsequently rise to predominance [20–25].

Antigenically novel strains of influenza A appear and become predominant in worldwide epidemics on a regular basis, which requires frequent adaptation of the influenza vaccine composition. The World Health Organization (WHO) monitors the genetic and antigenic characteristics of the circulating influenza A virus population and searches for antigenically novel emerging strains in a global surveillance program [26, 27]. The gathered surveillance information, combined with human serological data, is evaluated by a panel of experts. The panel meets twice a year to decide if an update of the vaccine composition for the next winter season for both the Northern and Southern Hemispheres is necessary. This approach results in a well-matched vaccine in most years, and significantly reduces the morbidity and mortality of seasonal influenza epidemics. However, a decreased vaccine efficacy can be caused by a new antigenic variant if it is identified too late to reformulate the vaccine composition.

A large body of work exists on computational studies of influenza A virus evolution. Phylogenetic reconstruction plays a key role here, since it was successfully used to unravel the global migration of human influenza A (H3N2) viruses [28] and to identify East and Southeast Asia as a global evolutionary reservoir of seasonal influenza A (H3N2) viruses [29]. Furthermore, genome-wide phylogenetic analysis of all eight viral segments determined that the evolutionary dynamics of influenza A (H3N2) virus are shaped by a complex interplay between genetic and epidemiological factors, such as mutation, reassortment, natural selection, and gene flow [30].

Besides these analytical studies, further computational methods have been applied to study and predict the evolution of human influenza A (H3N2) viruses. Changes within the hemagglutinin HA1 subunit sequence composition over time were visualized and analyzed by Shih et al. using amino acid frequency diagrams [31]. However, this procedure does not take the underlying evolutionary relationships and structure of the data into account, as isolate sequences and individual sites are treated independently. Plotkin et al. used agglomerative single-linkage clustering on hemagglutinin HA1 genetic sequences for decomposing the data into disjoint clusters, finding that influenza evolution is characterized by a succession of predominant clusters or “swarms” of similar strains [32]. This pattern is also reflected

by a narrow phylogenetic tree topology with one surviving viral lineage over time and a viral diversity that is periodically diminished by selective sweeps of a novel viral strain throughout the population [11, 30]. Analyzing the cluster size–time relation, Plotkin et al. suggested using a representative of the largest cluster as the vaccine strain for the following winter season [32]. Du et al. constructed a co-occurrence network from co-occurring nucleotides across the whole genome [33]. They identified co-occurring inter- and intra-segment changes, and used these co-occurrence modules for sequence clustering. This results in a grouping similar to the structure inferred by phylogenetic reconstruction. Xia et al. used mutual information to identify and visualize co-occurring mutations in a “site transition network” [34]. They also used this network to predict future mutations, resulting in 70 % sensitivity but also in a rather high false positive rate. However, it should be noted that although the term “predicting mutations” may convey that mutations are introduced independently in viral isolates in the following season, the effect that a particular genetic change increases in frequency over two consecutive seasons is often due to a previously low-abundance mutant circulating at higher prevalence.

Most of the above-mentioned studies assess the underlying evolutionary relationships and structure for the population-level sequence sample in some way. However, the standard way to estimate evolutionary relationships is by phylogenetic inference. As described above, Bush et al. identified 18 sites under positive selection by analyzing the ratio of nonsynonymous to synonymous nucleotide substitutions (dN/dS) on the trunk of a phylogenetic tree of hemagglutinin HA1 subunit sequences [6]. They subsequently used these sites to predict the direction of evolution for a phylogenetic tree of influenza A (H3N2) virus HA by identifying the strains within the phylogenetic tree that had the most pronounced evidence for positive selection [35]. However, the dN/dS ratio lacks sensitivity if applied to individual sites, as substantial evidence is required for a site to be considered informative. Not all relevant sites may thus be detectable and, furthermore, the most relevant sites may change over time [15]. In a more recent study, Pond et al. identified nine sites as being under directional selection in the HA segment of the influenza A (H3N2) virus, using a model-based phylogenetic maximum likelihood test. Seven of these sites are not detected with the traditional dN/dS ratio test [9]. Nevertheless, this method depends on the baseline amino acid substitution matrix and failed to identify adaptive sites when applied to dim-light and color vision genes in vertebrates [36].

To analyze the antigenic evolution of influenza A viruses, Smith et al. introduced a novel method known as antigenic cartography, which is based on multidimensional scaling of assay data on hemagglutination inhibition [15, 37]. This technique revealed that antigenic evolution is more clustered than genetic evolution, depending on the antigenic impact of individual amino acid exchanges, and that major changes (cluster jumps) occur every 3–4 years on average [15]. Accordingly, including both antigenic and genetic data within evolutionary models enables the most accurate analysis of influenza A virus evolution. Some studies try to incorporate antigenic data [38–40]; however, because of limited publicly available data, the results have to be approached with caution. To account

for this lack of antigenic information for the respective isolate sequences in our evaluation, we identified all predominant antigenic variants over the analyzed time period based on the genetic changes reported in the literature.

## 8.2 Methods

### 8.2.1 Phylogenetic Inference

HA sequences from 4,913 seasonal human influenza A (H3N2) virus isolates sampled from 1988 to 2008, and from 1,516 swine-origin influenza A (H1N1) virus isolates with exact sampling times (year and month) were downloaded from the influenza virus resource [41] (Tables S1 and S2). Alignments of DNA and protein sequences were created with Muscle [42] and manually curated. Phylogenetic trees were inferred with PhyML v3.0 [43] under the general time reversal GTR + I +  $\Gamma_4$  model, with the frequency of each substitution type, the proportion of invariant sites (I) and the gamma distribution of among-site rate variation, with four rate categories ( $\Gamma_4$ ), estimated from the data. Subsequently, the tree topology and branch lengths of the maximum likelihood tree inferred with PhyML were optimized for 200,000 generations with Garli v0.96b8 [44].

### 8.2.2 Allele Dynamics Plots

We describe AD-plots for visualizing the evolutionary dynamics of a gene in a population and for identifying the alleles that are potentially under directional selection. In a nutshell, AD-plots visualize gene alleles and their frequencies over time and thus enable a detailed analysis of a gene in a population. The basic idea involves the following four steps: (1) Inference of the evolutionary relationships for a sequence sample of a population. (2) Ancestral character state reconstruction and inference of evolutionary intermediates based on the reconstructed evolutionary relationships. (3) Mapping genetic changes to branches of the tree topology and defining the prevalence of distinct alleles of a gene at different points in time. (4) Finally, evaluating how fast new alleles or genetic variants propagate throughout the population.

Population genetics theory posits that in a population of constant size, genetic drift will result in variation in allele frequencies and the continuous fixation of variants even in the absence of selection [45–47]. However, given that selection acts on an allele and confers a fitness advantage to the individual organism, this will allow such alleles to rise faster in frequency than alleles without a selective advantage. Hence, alleles that increase in frequency most rapidly over time are more likely to be subject to directional selection than other alleles. This criterion

can be applied to identify those alleles that might be associated with a selective advantage from AD-plots.

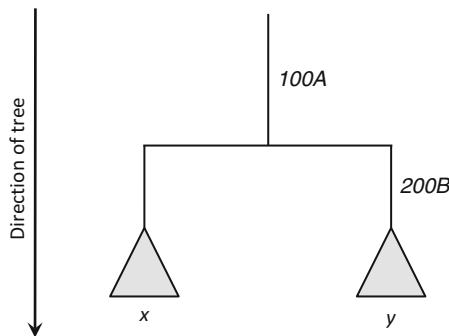
Following the phylogenetic inference of a tree topology using any standard method (maximum likelihood, Neighbor-Joining or a consensus tree constructed from a posterior sample of trees inferred with a Bayesian method [48, 49]), substitution events in the evolutionary history are reconstructed using ancestral character state reconstruction and assigned to individual tree branches. In detail, substitution events are assigned to the tree branches based on the evolutionary intermediates reconstructed as ancestral characters. We use the parsimony method of Fitch et al. [50] for ancestral character state reconstruction; however, in principle, any available method can be applied [51, 52]. In our analysis, we chose the isolate with the earliest sampling date as an out group and used accelerated transformation (AccTran) [51] to resolve ambiguities in character state reconstruction. This procedure results in changes being mapped preferentially closer to the root of the phylogenetic tree.

We define each branch that is associated with a nonempty set of substitutions to represent an individual allele. The number of alleles thus equals the number of branches with nonempty sets of substitutions in the phylogenetic tree. We define the frequency of an allele within a specific period as the ratio of the number of isolates in the subtree of the allele relative to the number of all isolates within the designated period. An allele that occurs later on the path from the root to the most recent isolates includes the substitutions of the alleles that occurred earlier on this path and thus is more specific. Allele frequencies are subsequently adjusted in case multiple related alleles emerge within the same period. Isolates located in the subtrees of a newly defined allele within a period are counted only once for the most closely placed parental allele in the phylogenetic tree. This means that for calculating the allele frequency of all less-specific alleles, isolates that occur in the subtree below the more-specific allele are not considered. Alleles and the relevant substitutions are discussed using the following nomenclature: *allele substitutions \*substitutions of parental alleles from the same period*\* (Fig. 8.1).

### 8.2.3 Construction of AD-Plots for Human Influenza A Viruses

In analyzing the evolution of human influenza A viruses, we are particularly interested in those changes that affect the antigenic properties of a virus. To identify viral variants with increased fitness for propagation through the host population, nonsynonymous genetic changes of hemagglutinin are of particular interest. To this end, we constructed AD-plots from the substitutions for the complete viral HA of the influenza A (H1N1) virus. Secondly, we constructed AD-plots for the seasonal influenza A (H3N2) virus based on the changes in the five epitope regions of HA [16, 17].

**Fig. 8.1** A tree demonstrating the concepts of alleles and allele frequency correction. For allele 100A, only the isolates of subtree  $x$  are counted, whereas for allele 200B \*100A\*, the isolates in subtree  $y$  are considered



**Table 8.1** Antigenically novel viral variants of influenza A (H3N2) that emerged and rose to predominance in worldwide epidemics between 1998 and 2008, and the corresponding substitutions reported in the literature in the five epitope sites of hemagglutinin

Antigenic cluster	Substitutions	Reference
A/Sydney/5/1997 (SY95)	62E, 156Q, 158K, 196A, 276K	[59]
A/Moscow/10/1999 (MO99)	57Q, 137S	[59]
A/Panama/2007/1999 (PA99)	144N, 172E, 192I	[59]
A/Fujian/411/2002 (FU02)	50G, 75Q, 83K, 131T, 155T, 156H, 186G	[60]
A/California/07/2004 (CA04)	145N, 159F, 189N, 226I, 227P	[61]
A/Wisconsin/67/2005 (WI05)	193F	[62]
A/Brisbane/10/2007 (BR07)	50E, 140I	[63]

Note that PA99 is antigenically similar to MO99 and was used as the vaccine candidate strain for MO99 [56]

Influenza infections in the human population show a pattern of seasonality. Peaks of activity occur mainly in the winter months in temperate regions of each hemisphere [53]. We use the standard definitions for the influenza season for the Northern and Southern Hemispheres in our analysis. For the Northern Hemisphere, the influenza season begins on the 1st of October and ends on the 31st of March in the following year. For the Southern Hemisphere, the influenza season begins on the 1st of April and ends on the 30th of September in the same year. For a comparison with the WHO vaccine strain recommendation, we restricted our analysis to sequences sampled up to the end of January for the Northern Hemisphere season and to the end of August for the Southern Hemisphere season, which is when the WHO decides on the vaccine composition.

To identify the alleles corresponding to the viral strains with antigenically novel HA variants, we used the literature to determine the genetic changes reported for every predominant antigenic variant over the analysis period. These appear, on average, every 3.3 years and then predominate worldwide in seasonal epidemics [15]. The changes in these strains for the five HA epitopes are given in Table 8.1.

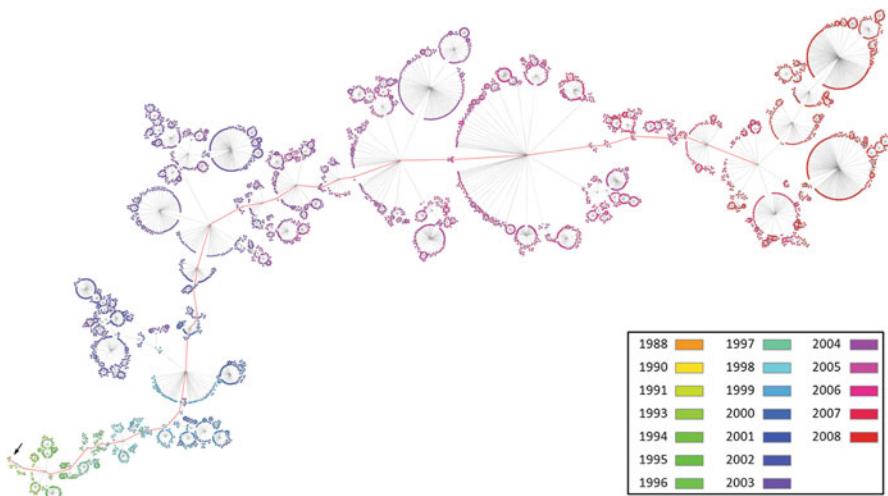
## 8.3 Results

### 8.3.1 Evolutionary Dynamics of Influenza A (H3N2)

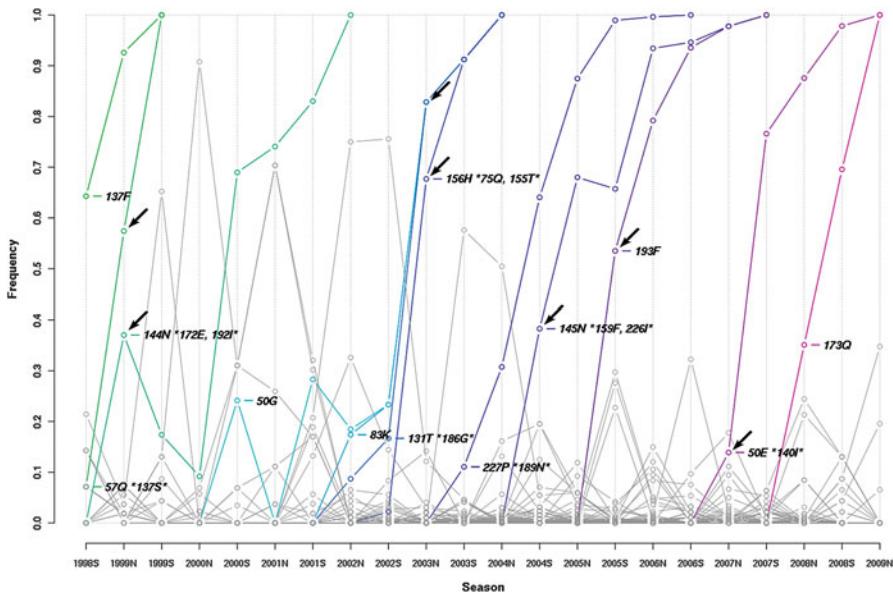
We analyze the evolutionary dynamics of the seasonal influenza A (H3N2) virus with AD-plots generated using a maximum likelihood tree (Fig. 8.2) from available HA sequences. The H3N2 subtype has been circulating since 1968, but here we focus on the time from 1998 until the end of 2008. For this more recent period, there is considerably more sequence data available and the bias of sequences towards isolates with unusual virulence or other atypical properties is reduced [54].

The AD-plot for HA of the human H3N2 virus (Fig. 8.3) shows several alleles that rise to predominance and reach fixation (their frequency in subsequent periods equals one) between 1998 and 2008, such as *57Q* \**137S*\*, *156H* \**75Q*, *155T*\*, and *193F*. Other alleles reach high frequencies and subsequently vanish, such as *160R* in the 1999 Southern season, *273S* in the 2000/01 Northern season or *126D* in the 2003 Southern season. Furthermore, a lot of minor frequency allele variation is evident within each period.

Alleles becoming predominant and rising to fixation in the surviving lineage correspond to substitutions that map to the trunk of the phylogenetic tree of hemagglutinin from the human influenza A (H3N2) virus. Besides such changes, the observable variation of alleles that do not become fixed (gray-colored alleles) is rather high within each time interval in the analyzed sample. Although some alleles



**Fig. 8.2** Maximum likelihood tree topology inferred for 4,913 hemagglutinin sequences of seasonal human influenza A (H3N2). Leaf nodes are *color-coded* according to the sampling dates of the viral isolates. The first sampled isolate, A/Siena/3/1988, is indicated with an arrow. The trunk of the tree (i.e., the path from the root to the most recent clade) is colored in red



**Fig. 8.3** Allele dynamics plot for the major surface protein and antigenic determinant of the seasonal influenza A (H3N2) virus. The Northern and Southern influenza seasons from 1998 to 2008 are shown. Alleles that reach a prevalence of more than 95 % and are subsequently fixed are shown in color; all other alleles are shown in gray. Substitutions are restricted to those that occur in the five epitope regions and are enumerated according to HA1 numbering [86]. Alleles that rise most quickly in frequency and are of interest with respect to vaccine strain selection are indicated by arrows

transiently reach high frequencies, they are only present over a short period. Notably, many of these alleles appear during times when an antigenic variant has been predominant for several years, such as the time from 2000 to 2003, when the A/Panama/2007/1999 (PA99) variant was predominant. In these years, several new alleles with similar antigenic properties, such as *160R* in the 1999 Southern season, *92T* in the 1999/2000 Northern season, *273S* and *50G*, *247C* in the 2000/01 Northern season, and *144D \*186G\** in the 2001/02 Northern season, [55–58] appeared successively and rose to high frequencies without reaching fixation.

Most of the alleles rising to fixation (colored in Fig. 8.3) are associated with substitutions reported in the literature [59–63] for the five distinct strains that represent predominant antigenic variants in the analysis period (Table 8.1). Note that the substitutions of a particular antigenic variant are not necessarily all part of the same allele (i.e., they do not map to the same branch on the trunk of the phylogenetic tree). Instead, they often follow each other in immediate succession in the AD-plot and are located on consecutive trunk branches of the phylogenetic tree. The earliest antigenic variant of the analysis period (PA99) is an exception, in this sense, as a single allele represents multiple substitutions. This reveals the limitations of the dataset for the earlier years, which does not allow the order in

which the PA99 substitutions were acquired by H3N2 to be resolved. For all subsequent antigenic variants, the order of the acquired substitutions is resolved and a set of multiple alleles becoming fixed within an interval are evident from the AD-plot. Thus, the evolutionary path and the order in which these changes were acquired in the evolution of antigenically new strains of H3N2 are revealed in the AD-plot. For instance, for the antigenic variant BR07, which was predominant from 2006 to 2009, the hemagglutinin plot shows that of the two relevant substitutions, 140I was acquired first, followed by 50E.

### 8.3.2 *Identification of Alleles Under Directional Selection in Influenza A (H3N2)*

The AD-plot, which visualizes the changes in frequencies of individual alleles in a sequence sample, enables us to easily identify those alleles that increase in prevalence most rapidly over two consecutive influenza seasons. The corresponding viral strains are likely candidates to be under the influence of directional selection and to have an advantage relative to other alleles. We identified the alleles with the largest increase in frequency between consecutive seasons that do not represent more than 50 % of the sequences in the first season (otherwise they would already be predominant; Table 8.1). Of the strains of the five antigenically distinct predominant variants (MO99/PA99, FU02, CA04, WI05, and BR07), four can be correctly identified by this criterion (Table 8.2). Thus, this measure allows us to use the AD-plots to easily identify the strains that are most relevant when deciding the composition of the influenza A (H3N2) vaccine.

In the 1998/99 Northern season, the allele that scores best is *57Q* \**137S*\*, which represents the MO99 variant that was predominant from the 1999 Southern season to the 2002/03 Northern season [55–58, 64–67]. The allele *144N* \**172E*, *192I*\*, which represents the antigenically very similar strain PA99, ranks second best. In agreement with the AD-plot observations, the WHO also recommended MO99 as the vaccine strain for the 2000 Southern season [55]. As no suitable well-growing candidate strain could be produced, the previously predominant SY97 strain was used in this season for the vaccine. PA99 was subsequently included as a vaccine component starting from 1999/2000 Northern season [56]. Thus, for the SY97-PA99 antigenic cluster transition, the AD-plot allows the timely identification of a suitable strain that is in agreement with the original recommendation of the WHO.

The FU02 variant, which predominated from 2003 to 2004/05 [68–71], is associated with seven distinct substitutions: 50G, 75Q, 83K, 131T, 155T, 156H, and 186G. 155T and 156H define the FU02 antigenic phenotype [72]. In the AD-plot, the seven FU02 substitutions are associated with seven distinct alleles, each with a single substitution. In the 2002/03 Northern season, alleles with the substitutions *131T* \**186G*\* and *156H* \**75Q*, *155T*\* score first and second best, respectively. The best scoring allele for the 2002/03 Northern season lacks the

**Table 8.2** Alleles and their associated antigenic phenotypes with the steepest slopes in the seasons when they are predicted to become predominant

Season	Alleles	Slope	Variant	WHO	Predominant
1998/99 North	<i>57Q</i> * <i>137S</i> * <i>144N</i> * <i>172E</i> , <i>192I</i> *	0.5027 0.3704	MO99 PA99	SY97 [80]	MO99/PA99 [56]
2002 South	<i>155T</i> * <i>75Q</i> * <i>131T</i> * <i>186G</i> * <i>83K</i> <i>50G</i>	0.0833 0.0797 0.0594 0.0485	FU02 FU02 HK02/FU02 HK02/FU02	MO99 [58]	FU02 [68]
2002/03 North	<i>131T</i> * <i>186G</i> * <i>156H</i> * <i>75Q</i> , <i>155T</i> * <i>83K</i> <i>50G</i>	0.6616 0.6546 0.5950 0.5950	FU02 FU02 HK02/FU02 HK02/FU02	FU02 [67]	FU02 [69]
2004 South	<i>145N</i> * <i>159F</i> , <i>226I</i> * <i>227P</i> * <i>189N</i> *	0.3828 0.3331	WE04/CA04 WE04/CA04	WE04 [69]	CA04 [73]
2005 South	<i>193F</i>	0.5350	WI05	CA04 [73]	WI05 [74]
2006/07 North	<i>50E</i> * <i>140I</i> *	0.1389	BR07	WI05 [75]	BR07 [78]

Alleles in one season are ordered by decreasing slope. Further comparisons show the recommended reference strain for the use in the next year's vaccine by the WHO and the predominant antigenic variant in the next year's influenza season for the same hemisphere. Note that A/Hong Kong/1143/2002 (HK02, [50G, 83K, 186G]) is a PA99-like sublineage present before FU02 and A/Wellington/1/2004 (WE04, [159F, 189N, 227P]) was directly replaced by CA04 in 2004/05 Northern season before becoming predominant

relevant substitutions 155T and 156H described for FU02. Here, the frequency indicator does not directly reveal the best candidate strain based on the available data. Antigenic information would probably allow a more detailed analysis. The second high scoring allele would presumably be a good choice as a vaccine strain, as it has other antigenically relevant changes and shows a rapid increase in prevalence during the season. In agreement with this conjecture, the corresponding strain (A/Fujian/411/2002) was recommended by the WHO as the vaccine strain for the 2003/04 Northern season [67]. However, as no suitable well-growing candidate strain could be produced, the MO99/PA99 strain was used for the vaccine. In the 2002 Southern season, the *155T* \**75Q*\* allele ranks first, but the correct allele (*156H* \**75Q*, *155T*), which features all necessary substitutions, increases only little in frequency and is thus not selected.

Interestingly, an additional substitution (186G) found in the highest scoring allele for the 2002/03 Northern season appears independently in another frequent allele in the preceding season. This seems a general aspect of H3N2 evolution—the repeated appearance of the same substitution in multiple different alleles. Often, the respective alleles have different phylogenetic histories, in that they occur in different parts of the tree, and the substitutions are occasionally encoded by different codons. Such repeated changes can either reflect neutral changes at highly variable sequence positions or they can be the result of directional selection against a certain

residue at a given position at this time. The AD-plot allows us to identify such changes easily for further analysis.

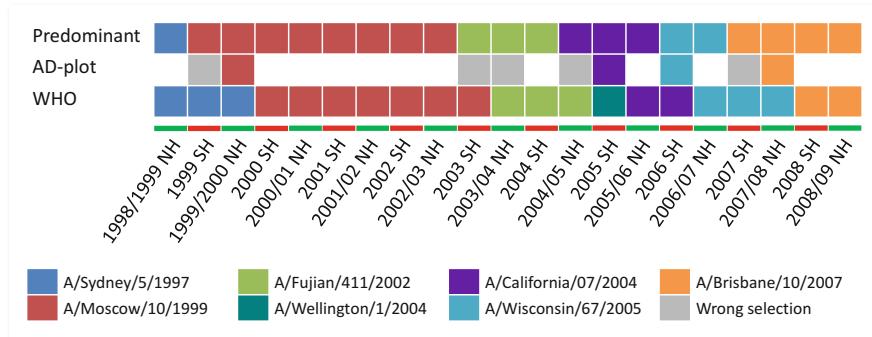
The CA04 variant was predominant from 2004/05 to 2005/06 [73, 74] and was recommended as vaccine strain for the 2005/06 Northern season in the spring of 2005 [71]. The HA allele of this strain scores highest in the 2004 Southern season. Here, the two alleles featuring the substitutions *I45N* \**I59F*, *226I*\*, and *227P* \**I89N*\*, respectively, rank first and second. Both of these alleles contain substitutions of the CA04 variant, but only the top ranking one possesses all relevant substitutions and thus is the correct choice.

The WI05 variant predominated from 2006 to 2006/07 [74, 75] and was recommended one season too late as the vaccine strain for the 2006/07 Northern season [76]. In the 2005 Southern season, the *I93F* allele associated with the WI05 variant scores highest. The second substitution associated with WI05, *225N*, is not evident from this plot, as it is not part of the epitope regions. If nonepitope sites are included in the analysis, both substitutions appear on subsequent branches, corresponding to two consecutive emerging alleles in the plot (data not shown). In this plot, the allele *225N* \**I93F*\* scores highest. The AD-plot thus allows us to identify the WI05 variant from the available data one season before the WHO's official recommendation.

Finally, the antigenic variant BR07, which predominated from 2007 onwards [13, 77–79], scores highest in the 2006/07 Northern season and is represented by an allele with the substitutions *50E* \**140I*\*. A matching strain was recommended for the vaccine of the 2008 Southern season [77]. The AD-plot allows us to identify this emerging variant for the 2007/08 Northern season.

Applying a maximum likelihood test for directional evolution of protein sequences (DEPS) [9] to the HA data of H3N2 from 1988 to 2008 revealed 42 sites in the HA epitopes. Nine of these sites are also under positive selection according to a dN/dS ratio test [8] (data not shown). However, of the 20 epitope sites where changes rise to fixation over the analysis period (Fig. 8.2), only 12 are detected by the DEPS method. This highlights that such rapidly fixed changes cannot all be identified by common selection tests.

Retrospectively, our approach allows the identification of the CA04/WI05 antigenic cluster transition in the 2005 Southern season, 1 year before it rises to predominance in the 2006 season (Fig. 8.4). In all other cases, our method allows us to identify the correct strain one season before the respective antigenic variant becomes predominant: The SY97/MO99 transition is detected in the 1998/99 Northern Hemisphere season, while the MO99 variant became predominant in the 1999 Southern Hemisphere season. The FU02/CA04 transition is predicted in the 2004 Southern Hemisphere season, while CA04 became predominant in the 2004/05 Northern season. Finally, the WI05/BR07 transition is identified in the 2006/07 Northern season, while the BR07 antigenic variant became predominant in the 2007 Southern season. In comparison to the WHO recommendations [13, 14, 55–58, 64–71, 73–80], this approach identifies the newly emerging variants one season earlier. This may be because the WHO tends to be conservative in recommendations, to avoid suggesting an antigenic variant that may never actually rise to

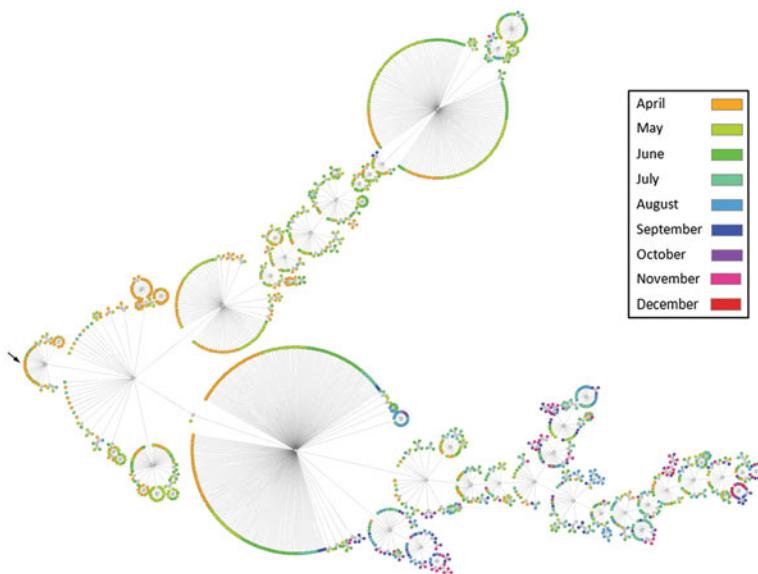


**Fig. 8.4** Comparison of predominant influenza A (H3N2) strains, WHO vaccine strain recommendations and strains identified by AD-plot analysis. For the AD-plot analysis, seasons with antigenic cluster transitions are shown in color. The information shown for the AD-plot and the WHO recommendation represents the selection made 1 year earlier

predominance in the future. However, in general, new variants reach predominance very rapidly, if the time from the first appearance in the available genetic sequences is measured. In all three cases above, the new variant rose to predominance after its first appearance within a single year. Thus, given the available data, predicting this event one year ahead of time would be impossible. Fortunately, in some cases the antigenic changes between successive variants are not that large [15, 37]. For instance, MO99 was antigenically similar to SY97. Thus, even though most isolates sampled in the 1999 Southern season reacted to a higher titer with the ferret antisera raised against MO99 [55], recommending SY97 for the vaccine composition thus did not result in a dramatically lower vaccine efficacy.

### 8.3.3 *Influence of Timing on Antigenic Variant Identification*

Twice a year, in February and September, vaccine strains are recommended for influenza B, influenza A (H3N2), and influenza A (H1N1) to the manufacturers of the seasonal influenza vaccine. This recommendation is made approximately 1 year before the vaccine will be used in the Northern or Southern seasons, respectively [27]. Above, we analyzed the data available only up to that point. If using all available data until the end of the influenza seasons, emerging alleles appear at high frequencies in the respective AD-plot. For example, this happened for the BR07 allele in the 2006/07 Northern Hemisphere season (Fig. 8.3). Previously circulating strains, on the other hand, occur at lower frequencies in comparison, as newly emerging antigenic variants increase in prevalence typically towards the end of a season. This effect is more pronounced for the Northern Hemisphere than for the Southern Hemisphere, possibly because after the vaccine meeting in the Northern

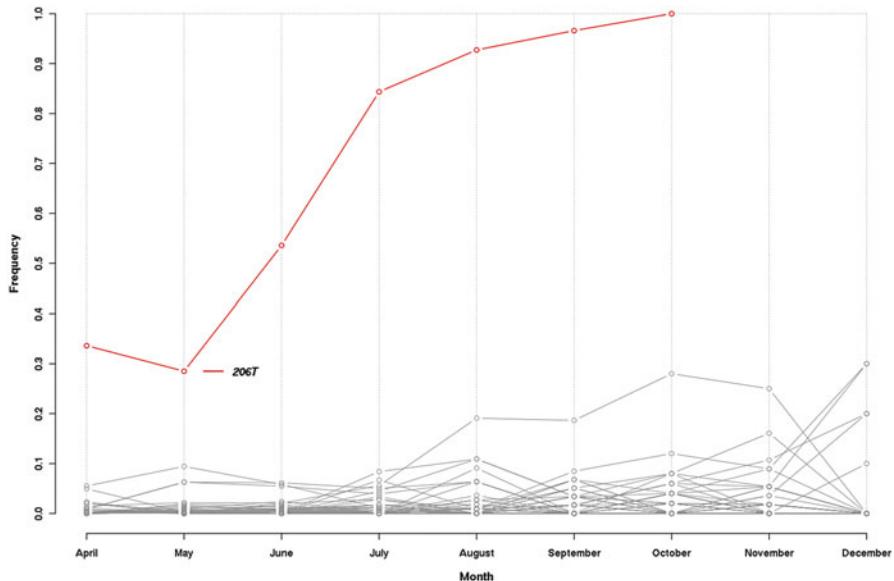


**Fig. 8.5** Maximum likelihood tree topology inferred from 1,516 2009 swine-origin influenza A (H1N1) hemagglutinin sequences. Leaf nodes are *color-coded* according to the sampling dates of the viral isolates. The first sampled isolate, A/California/05/2009, is indicated with an *arrow*

Hemisphere, 2 months of the winter season are still to follow, whereas only 1 month of winter still remains in the Southern Hemisphere. However, overall the picture remains very similar. Based on all available data, all five antigenic variants can be identified based on their rapid increase in prevalence. A noteworthy difference is evident only for the 2002/03 Northern season, where the  $156H *75Q$ ,  $155T^*$  allele of the emerging FU02 antigenic variant now ranks first. In summary, limiting the data to what is available by the time of the WHO vaccine meetings, reduces the frequency of alleles associated with newly emerging variants in the AD-plot, but the ability to identify viral strains that subsequently rise to predominance is preserved in four out of five cases.

### 8.3.4 Evolutionary Dynamics of the Influenza A (H1N1) Virus

We next studied the evolutionary dynamics of the 2009 influenza A (H1N1) virus, using 1,516 available exactly dated HA sequences (Fig. 8.5). The virus has circulated in the human population only since April 2009 [81–83]. Therefore, we have studied the evolutionary dynamics in monthly intervals (Fig. 8.6). As isolate A/California/05/2009 was the only one sampled in March, it was assigned to the 1st of April to avoid errors introduced through the small sample size for March 2009. The



**Fig. 8.6** Allele dynamics plot for the major surface protein and antigenic determinant of the new influenza A (H1N1) based on sequences sampled between April and December of 2009 without allele frequency correction. Alleles that reach a prevalence of more than 95 % and are subsequently fixed are shown in color; all other alleles are shown in gray. Substitutions are enumerated according to H3 HA1 numbering [86]

AD-plots show that one nonsynonymous and another synonymous change become fixed over the analysis period. The corresponding substitutions, T658A (encoding the S206T change (H3 HA1 numbering)) and C1408T (encoding a synonymous substitution for leucine), have already been reported to divide the sequenced isolates into two distinct clusters [84], but have no known antigenic impact [81]. Furthermore, Pan et al. has already reported an increase in allele frequency for the S206T substitution among new H1N1 sequence isolates [85].

Besides these changes, the plot also reveals the existence of several other alleles, which, so far, appear only at low frequencies and did not become fixed until December of 2009. Despite the fact that the data currently is very limited, at this point, the plots do not reveal any alleles or associated substitutions that seem to be on the rise. Thus, based on the available data, the virus currently seems stable in terms of antigenicity, indicating that no update of the vaccine strain for this virus will be required for the 2010/11 season (also reported by the WHO [14]). However, some caution is warranted in this interpretation, as different months are represented very unevenly, with lots of data from April and May of 2009 and much less from the following months.

DEPS analysis of the H1N1 data identifies five sites in HA with evidence for directional evolution. Three of these sites are also predicted to be under positive selection based on a dN/dS ratio test. This includes position 206, where a

nonsynonymous change has become fixed within the analysis period (220 in H1 sequence numbering). This indicates that this site might have been under positive selection and that several further sites could be of relevance for the future evolution of H1N1. However, overall, these results should be taken with care, as the analysis period of 1 year, during which extensive sampling has taken place, is rather short, and the data might be more enriched than samples obtained over longer periods, with many neutral or slightly deleterious mutations.

## 8.4 Discussion

AD-plots provide a simple and easy-to-interpret visualization of the evolutionary dynamics of a gene within a population from a sample of dated genetic sequences. This is particularly helpful for the analysis of large-scale sequence datasets, where a standard visualization such as a phylogenetic tree topology is difficult to interpret manually and does not directly display sampling times. Here, we have applied our method to investigate the evolutionary dynamics of seasonal influenza A H3N2 and H1N1 viruses, for which available sequence data is abundant.

From the AD-plot for influenza A (H3N2), one can easily determine the order in which substitutions of the surviving lineage became fixed over the analysis period, and one can identify the predominant antigenic variants between 1998 and 2008. Furthermore, we propose a novel indicator for directional selection, which allows us to identify the alleles and corresponding substitutions that might have a selective advantage. We demonstrate this approach for identifying future predominant and novel viral strains. With this method, strains for four out of five antigenic phenotype transitions in influenza A (H3N2) evolution can be identified, based on the data available up to the time of the WHO vaccine strain meeting. One limitation for this application is the fact that a particular allele may score best for every time period, with no information on whether it is antigenically similar or different from the current vaccine strain. Hence, antigenic information also has to be considered to decide whether a vaccine update is warranted. In summary, AD-plots enable a sensitive and timely method for detecting emerging viral strains that rise to high frequencies in subsequent seasons. In our analysis, we find that AD-plots permit us to accurately identify those alleles that subsequently rise to predominance and become fixed in the course of viral evolution. In combination with antigenic information on the individual strains, AD-plots thus present a new tool for the detailed analysis of influenza surveillance data that could be used in the selection of strains for the seasonal influenza A virus vaccine.

Secondly, we used AD-plots to analyze the evolutionary dynamics of the 2009 influenza A (H1N1) virus. The AD-plot for this virus reveals several new variants with unique genetic composition that circulate at low levels in the human population and two genetic changes that became fixed in the period from April to December 2009. At this point, the plot does not allow identification of any further

genetic changes that may become fixed in the near future, indicating that the virus currently is evolutionary stable, even though data is limited.

In summary, we present a novel visualization technique for the study of longitudinal population-level sequence samples and for the identification of alleles that are on the rise to predominance. The method allows us to investigate the evolutionary dynamics of rapidly evolving populations, under consideration of the inherent evolutionary relationships and structure of the data. It complements existing methods for detecting sites under directional and positive selection, such as dN/dS ratio tests or DEPS. Note that AD-plots are not limited to the study of influenza A viruses, but can also be applied for the analysis of other fast-evolving populations, such as the intra-host evolution of human immunodeficiency viruses or hepatitis C viruses. Generally, the best results are likely to be obtained if the analyzed sequence sample is representative for a constant-sized population without too much structure (e.g., geographic subdivisions). In this case, variations in frequencies can be taken as estimates for the evolutionary dynamics of the respective population. Finally, while many computational techniques have been applied to predict the evolutionary dynamics of influenza A viruses, our method integrates state-of-the-art phylogenetic inference, ancestral state reconstruction, and a novel indicator of directional selection into the analysis, and thus provides a solution with extensive theoretical support.

## References

1. Grenfell BT, Pybus OG, Gog JR, Wood JLN, Daly JM, Mumford JA, Holmes EC (2004) Unifying the epidemiological and evolutionary dynamics of pathogens. *Science* 303:327–332
2. Pybus OG, Rambaut A (2009) Evolutionary analysis of the dynamics of viral infectious disease. *Nat Rev Genet* 10:540–550
3. Wallace RG, HoDac HM, Lathrop RH, Fitch WM (2007) A statistical phylogeography of influenza A H5N1. *Proc Natl Acad Sci U S A* 104:4473–4478
4. Keele BF, Van Heuverswyn F, Li Y, Bailes E, Takehisa J, Santiago ML, Bibollet-Ruche F, Chen Y, Wain LV, Liegeois F et al (2006) Chimpanzee reservoirs of pandemic and nonpandemic HIV-1. *Science* 313:523–526
5. Lemey P, Pybus OG, Wang B, Saksena NK, Salemi M, Vandamme A-M (2003) Tracing the origin and history of the HIV-2 epidemic. *Proc Natl Acad Sci U S A* 100:6588–6592
6. Bush R (1999) Positive selection on the H3 hemagglutinin gene of human influenza virus. *A. Mol Biol Evol* 16:1457–1465
7. Kryazhimskiy S, Plotkin JB (2008) The population genetics of dN/dS. *PLoS Genet* 4:e1000304
8. Pond SLK, Frost SDW, Muse SV (2005) HyPhy: hypothesis testing using phylogenies. *Bioinformatics* 21:676–679
9. Pond K, Sergei L, Poon AFY, Brown L, Andrew J, Frost SDW (2008) A maximum likelihood method for detecting directional evolution in protein sequences and its application to influenza A virus. *Mol Biol Evol* 25:1809–1824
10. Ricklefs RE (2007) Estimating diversification rates from phylogenetic information. *Trends Ecol Evol* 22:601–610
11. Koelle K, Cobey S, Grenfell B, Pascual M (2006) Epochal evolution shapes the phylodynamics of interpandemic influenza A (H3N2) in humans. *Science* 314:1898–1903

12. Fouchier RAM, Munster V, Wallensten A, Bestebroer TM, Herfst S, Smith D, Rimmelzwaan GF, Olsen B, Osterhaus ADME (2005) Characterization of a novel influenza A virus hemagglutinin subtype (H16) obtained from black-headed gulls. *J Virol* 79:2814–2822
13. WHO (2009) Recommended composition of influenza virus vaccines for use in 2009–2010 influenza season (northern hemisphere winter). *Wkly Epidemiol Rec* 84:65–72
14. WHO (2010) Recommended viruses for influenza vaccines for use in the 2010–2011 northern hemisphere influenza season. *Wkly Epidemiol Rec* 85:81–92
15. Smith DJ, Lapedes AS, de Jong JC, Bestebroer TM, Rimmelzwaan GF, Osterhaus ADME, Fouchier RAM (2004) Mapping the antigenic and genetic evolution of influenza virus. *Science* 305:371–376
16. Wiley D, Wilson I, Skehel J (1981) Structural identification of the antibody-binding sites of Hong Kong influenza haemagglutinin and their involvement in antigenic variation. *Nature* 289:373–378
17. Wiley DC, Skehel JJ (1987) The structure and function of the hemagglutinin membrane glycoprotein of influenza virus. *Annu Rev Biochem* 56:365–394
18. Wilson IA, Cox NJ (1990) Structural basis of immune recognition of influenza virus hemagglutinin. *Annu Rev Immunol* 8:737–771
19. Skehel JJ, Wiley DC (2000) Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. *Annu Rev Biochem* 69:531–569
20. Kuiken T, Holmes EC, McCauley J, Rimmelzwaan GF, Williams CS, Grenfell BT (2006) Host species barriers to influenza virus infections. *Science* 312:394–397
21. Webster R, Bean W, Gorman O, Chambers T, Kawaoka Y (1992) Evolution and ecology of influenza A viruses. *Microbiol Rev* 56:152–179
22. Lowen AC, Palese P (2007) Influenza virus transmission: basic science and implications for the use of antiviral drugs during a pandemic. *Infect Disord Drug Targets* 7:318–328
23. Morens DM, Taubenberger JK, Fauci AS (2009) The persistent legacy of the 1918 influenza virus. *N Engl J Med* 361:225–229
24. Neumann G, Noda T, Kawaoka Y (2009) Emergence and pandemic potential of swine-origin H1N1 influenza virus. *Nature* 459:931–939
25. Zimmer SM, Burke DS (2009) Historical perspective – emergence of influenza A (H1N1) viruses. *N Engl J Med* 361:279–285
26. Cox NJ, Brammer TL, Regnery HL (1994) Influenza: global surveillance for epidemic and pandemic variants. *Eur J Epidemiol* 10:467–470
27. Russell CA, Jones TC, Barr IG, Cox NJ, Garten RJ, Gregory V, Gust ID, Hampson AW, Hay AJ, Hurt AC et al (2008) Influenza vaccine strain selection and recent studies on the global migration of seasonal influenza viruses. *Vaccine* 26:31–34
28. Nelson MI, Simonsen L, Viboud C, Miller MA, Holmes EC, Levin B (2007) Phylogenetic analysis reveals the global migration of seasonal influenza A viruses. *PLoS Pathog* 3:e131
29. Russell CA, Jones TC, Barr IG, Cox NJ, Garten RJ, Gregory V, Gust ID, Hampson AW, Hay AJ, Hurt AC et al (2008) The global circulation of seasonal influenza A (H3N2) viruses. *Science* 320:340–346
30. Rambaut A, Pybus OG, Nelson MI, Viboud C, Taubenberger JK, Holmes EC (2008) The genomic and epidemiological dynamics of human influenza A virus. *Nature* 453:615–619
31. Shih ACC, Hsiao TC, Ho MS, Li WH (2007) Simultaneous amino acid substitutions at antigenic sites drive influenza A hemagglutinin evolution. *Proc Natl Acad Sci U S A* 104:6283–6288
32. Plotkin JB, Dushoff J, Levin SA (2002) Hemagglutinin sequence clusters and the antigenic evolution of influenza A virus. *Proc Natl Acad Sci U S A* 99:6263–6268
33. Du X, Wang Z, Wu A, Song L, Cao Y, Hang H, Jiang T (2008) Networks of genomic co-occurrence capture characteristics of human influenza A (H3N2) evolution. *Genome Res* 18:178–187
34. Xia Z, Jin G, Zhu J, Zhou R (2009) Using a mutual information-based site transition network to map the genetic evolution of influenza A/H3N2 virus. *Bioinformatics* 25:2309–2317

35. Bush RM, Bender CA, Subbarao K, Cox NJ, Fitch WM (1999) Predicting the evolution of human influenza A. *Science* 286:1921–1925
36. Nozawa M, Suzuki Y, Nei M (2009) Reliabilities of identifying positive selection by the branch-site and the site-prediction methods. *Proc Natl Acad Sci U S A* 106:6700–6705
37. Fouchier RAM, Smith DJ (2010) Use of antigenic cartography in vaccine seed strain selection. *Avian Dis* 54:220–223
38. Huang JW, King CC, Yang JM (2009) Co-evolution positions and rules for antigenic variants of human influenza A/H3N2 viruses. *BMC Bioinformatics* 10:S41
39. Lee MS, Chen MC, Liao YC, Hsiung CA (2007) Identifying potential immunodominant positions and predicting antigenic variants of influenza A/H3N2 viruses. *Vaccine* 25:8133–8139
40. Liao YC, Lee MS, Ko CY, Hsiung CA (2008) Bioinformatics models for predicting antigenic variants of influenza A/H3N2 virus. *Bioinformatics* 24:505–512
41. Bao Y, Bolotov P, Dernovoy D, Kiryutin B, Zaslavsky L, Tatusova T, Ostell J, Lipman D (2008) The influenza virus resource at the National Center for Biotechnology Information. *J Virol* 82:596–601
42. Edgar RC (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5:113
43. Guindon S, Gascuel O (2003) A simple, fast, and accurate method to estimate large phylogenies by maximum likelihood. *Syst Biol* 52:696–704
44. Zwickl D (2006) Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. Thesis, The University of Texas at Austin
45. Futuyma DJ (1998) Evolutionary biology, 3rd edn. Sinauer Associates, Sunderland, MA
46. Hein J, Schierup M, Wiuf C (2005) Gene genealogies, variation and evolution: a primer in coalescent theory. Oxford University Press, USA
47. Templeton AR (2006) Population genetics and microevolutionary theory. Wiley, Hoboken, NJ
48. Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17:754–755
49. Drummond A, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* 7:214
50. Fitch WM (1971) Toward defining the course of evolution: minimum change for a specific tree topology. *Syst Zool* 20:406–416
51. Felsenstein J (2004) Inferring phylogenies. Sinauer Associates, Sunderland, MA
52. Pagel M, Meade A, Barker D (2004) Bayesian estimation of ancestral character states on phylogenies. *Syst Biol* 53:673–684
53. Nelson MI, Holmes EC (2007) The evolution of epidemic influenza. *Nat Rev Genet* 8:196–205
54. Ghedin E, Sengamalay NA, Shumway M, Zaborsky J, Feldblyum T, Subbu V, Spiro DJ, Sitz J, Koo H, Bolotov P et al (2005) Large-scale sequencing of human influenza reveals the dynamic nature of viral genome evolution. *Nature* 437:1162–1166
55. WHO (1999) Recommended composition of influenza virus vaccines for use in the 2000 influenza season. *Wkly Epidemiol Rec* 74:321–325
56. WHO (2000) Recommended composition of influenza virus vaccines for use in the 2000–2001 season. *Wkly Epidemiol Rec* 75:61–65
57. WHO (2001) Recommended composition of influenza virus vaccines for use in the 2001–2002 influenza season. *Wkly Epidemiol Rec* 76:58–61
58. WHO (2002) Recommended composition of influenza virus vaccines for use in the 2002–2003 influenza season. *Wkly Epidemiol Rec* 77:62–66
59. Lin Y, Gregory V, Bennett M, Hay A (2004) Recent changes among human influenza viruses. *Virus Res* 103:47–52
60. Hay AJ, Lin YP, Gregory V, Bennett M (2003) WHO collaborating centre for reference and research on influenza, annual report. National Institute for Medical Research, London

61. Hay AJ, Lin YP, Gregory V, Bennet M (2005) WHO collaborating centre for reference and research on influenza, interim report February. National Institute for Medical Research, London
62. Hay AJ, Lin YP, Gregory V, Bennet M (2006) WHO collaborating centre for reference and research on influenza, interim report March. National Institute for Medical Research, London
63. Hay AJ, Daniels R, Lin YP, Xiang Z, Gregory V, Bennet M, Whittaker L (2007) WHO collaborating centre for reference and research on influenza, interim report September. National Institute for Medical Research, London
64. WHO (2000) Recommended composition of influenza virus vaccines for use in the 2001 influenza season. *Wkly Epidemiol Rec* 75:330–333
65. WHO (2001) Recommended composition of influenza virus vaccines for use in the 2002 influenza season. *Wkly Epidemiol Rec* 76:311–314
66. WHO (2002) Recommended composition of influenza virus vaccines for use in the 2003 influenza season. *Wkly Epidemiol Rec* 77:344–348
67. WHO (2003) Recommended composition of influenza virus vaccines for use in the 2003–2004 influenza season. *Wkly Epidemiol Rec* 78:58–62
68. WHO (2003) Recommended composition of influenza virus vaccines for use in the 2004 influenza season. *Wkly Epidemiol Rec* 78:375–379
69. WHO (2004) Recommended composition of influenza virus vaccines for use in the 2004–2005 influenza season. *Wkly Epidemiol Rec* 79:88–92
70. WHO (2004) Recommended composition of influenza virus vaccines for use in the 2005 influenza season. *Wkly Epidemiol Rec* 79:369–373
71. WHO (2005) Recommended composition of influenza virus vaccines for use in the 2005–2006 influenza season. *Wkly Epidemiol Rec* 80:66–71
72. Jin H, Zhou H, Liu H, Chan W, Adhikary L, Mahmood K, Lee MS, Kemble G (2005) Two residues in the hemagglutinin of A/Fujian/411/02-like influenza viruses are responsible for antigenic drift from A/Panama/2007/99. *Virology* 336:113–119
73. WHO (2005) Recommended composition of influenza virus vaccines for use in the 2006 influenza season. *Wkly Epidemiol Rec* 80:342–347
74. WHO (2006) Recommended composition of influenza virus vaccines for use in the 2007 influenza season. *Wkly Epidemiol Rec* 81:390–395
75. WHO (2007) Recommended composition of influenza virus vaccines for use in the 2007–2008 influenza season. *Wkly Epidemiol Rec* 82:69–74
76. WHO (2006) Recommended composition of influenza virus vaccines for use in the 2006–2007 influenza season. *Wkly Epidemiol Rec* 81:82–86
77. WHO (2007) Recommended composition of influenza virus vaccines for use in the 2008 influenza season. *Wkly Epidemiol Rec* 82:351–356
78. WHO (2008) Recommended composition of influenza virus vaccines for use in the 2008–2009 influenza season. *Wkly Epidemiol Rec* 83:81–87
79. WHO (2008) Recommended composition of influenza virus vaccines for use in the 2009 southern hemisphere influenza season. *Wkly Epidemiol Rec* 83:366–372
80. WHO (1999) Recommended composition of influenza virus vaccines for use in the 1999–2000 season. *Wkly Epidemiol Rec* 74:57–61
81. Garten RJ, Davis CT, Russell CA, Shu B, Lindstrom S, Balish A, Sessions WM, Xu X, Skepner E, Deyde V et al (2009) Antigenic and genetic characteristics of swine-origin 2009 A (H1N1) influenza viruses circulating in humans. *Science* 325:197–201
82. Smith G, Vijaykrishna D, Bahl J, Lycett S, Worobey M, Pybus O, Ma S, Cheung C, Raghwan J, Bhatt S et al (2009) Origins and evolutionary genomics of the 2009 swine-origin H1N1 influenza A epidemic. *Nature* 459:1122–1125
83. A Novel Swine-Origin Influenza (H1N1) Virus Investigation Team (2009) Emergence of a novel swine-origin influenza A (H1N1) virus in humans. *N Engl J Med* 360:2605–2615

84. Fereidouni SR, Beer M, Vahlenkamp T, Starick E (2009) Differentiation of two distinct clusters among currently circulating influenza A (H1N1)v viruses, March–September 2009. *Euro Surveill* 14:19409–19411
85. Pan C, Cheung B, Tan S, Li C, Li L, Liu S, Jiang S (2010) Genomic signature and mutation trend analysis of pandemic (H1N1) 2009 influenza A virus. *PLoS One* 5:e9549
86. Nobusawa E, Aoyama T, Kato H, Suzuki Y, Tateno Y, Nakajima K (1991) Comparison of complete amino acid sequences and receptor-binding properties among 13 serotypes of hemagglutinins of influenza A viruses. *Virology* 182:475–485

# Chapter 9

## Modeling the Spread and Outbreak Dynamics of Avian Influenza (H5N1) Virus and Its Possible Control

V. Sree Hari Rao and Ranjit Kumar Upadhyay

### 9.1 Introduction

In the new millennium, the world has seen the emergence of three novel human respiratory viruses: SARS virus (a novel corona virus) in 2003, Avian Influenza virus (H5N1) in 2004 and an international outbreak caused by a new strain of influenza virus 2009 A (H1N1). This novel influenza 2009 A/H1N1 virus contains a combination of swine, avian, and human influenza virus genes. In sharp contrast to H5N1 and SARS viruses which emerged from the Asian continent, H1N1 virus emerged from North America (Mexico). Many important infectious diseases persist on a knife-edge: rapid rates of transmission coupled with brief infectious periods generate boom-and-bust epidemic that court extinction [5]. Such violent epidemic behavior has been observed in measles [14], cholera [33], meningitis [25, 58], and pertussis [49], among others. Several distinct mechanisms have been proposed to explain the spread and outbreak dynamics of avian influenza virus. These examples illustrate the need for understanding alternative spread and reinvasion mechanisms of infectious diseases for effective management and control.

In this chapter, we investigate the spread and outbreak dynamics of low pathogenic avian influenza virus. Avian Influenza, commonly known as Bird Flu, is an epidemic caused by H5N1 Virus, that primarily affects birds such as chickens, wild

---

V. Sree Hari Rao

Foundation for Scientific Research and Technological Innovation, Vadrevu Nilayam,  
13-405, Road No.14, Alakapuri, Hyderabad, Andhra Pradesh 500 035, India  
e-mail: [vshrao@gmail.com](mailto:vshrao@gmail.com)

R.K. Upadhyay (✉)

Department of Applied Mathematics, Indian School of Mines, Dhanbad 826 004, India  
e-mail: [ranjit\\_ism@yahoo.com](mailto:ranjit_ism@yahoo.com)

water birds, ducks, and swans etc. On rare occasions, pigs and humans will also be affected with this virus. Migratory aquatic birds, most notably wild ducks, are the natural reservoir of avian influenza viruses that inhabit the intestines of these birds. Infection in domestic poultry is thought to occur due to contact with these aquatic/wild birds. Fifteen subtypes of influenza virus are known to infect birds, providing a large pool of influenza viruses potentially circulating in bird populations. Avian Influenza is an infection caused by a virus known as *orthomyxoviridae* in virus classification. Influenza virus has only one species in it, which is called influenza A virus. Influenza A infects humans and animals such as birds, pigs, horses, and seals. When it infects birds, it is called avian influenza; when it infects pigs, it is called swine influenza, and so on. Typically, avian influenza viruses occur naturally among birds. But occasionally, avian influenza infects humans who have been in close contact with birds, with a rich pool of genetic and antigenic diversity that often leads to cross-species transmission. Wild birds worldwide carry these viruses in their intestine but usually do not get sick from them. However avian influenza is very contagious among birds and can make some domesticated birds including chickens, ducks, and turkeys very sick and kill them. Infected birds shed influenza viruses in their saliva, nasal secretions, etc. Susceptible birds become infected when they come in contact with the contaminated surfaces. Domesticated birds may become infected with avian influenza viruses through direct contact with infected waterfowl or other infected poultry or through contact with surfaces (such as dirt or cages) or materials (such as water or food) that have been contaminated with the virus [1, 14, 54].

Influenza A in humans is mainly a respiratory virus that typically infects cells of the nose and throat, but it can infect lung cells. It spreads when an uninfected person touches contaminated surfaces or inhales viruses coughed or sneezed out by an infected person. The recent avian strain that has been infecting humans is called H5N1 (named for its surface glycoproteins). H5N1 kills a high percentage of people who become infected, but thus far, the virus does not seem to spread well from person to person. Scientists fear that a person already infected with a human influenza A virus may become infected with the avian influenza A. If infected with both viruses, a hybrid avian–human influenza A virus may be generated that has two features making it particularly dangerous for human: avian H and N glycoproteins, to which humans have never been exposed (thus the immune system cannot quickly recognize and control the virus), and RNA from the human virus (enabling the hybrid to spread easily from person to person).

An outbreak of influenza A (H5N1) has been reported in several countries throughout Asia. Cases of Avian influenza A in birds have been confirmed in Cambodia, China, Hong Kong, Indonesia, Japan, Laos, Pakistan, Thailand, and Vietnam. Human cases of avian influenza have been reported in Thailand and Vietnam. During this outbreak investigation, it has been determined that avian influenza is spread from person to person. The current outbreak of avian influenza has prompted the killing of more than 25 million birds in Asia. In general the outbreak of an infectious disease is dependent upon three necessary conditions as the source of an infection, the route of transmission, and the herd susceptibility [32]. Other social

and natural factors also play an important role in the transmission of infection, for example the control measures and the change in temperature. The source of infection that led to the outbreak is not clear. In some researcher's view migratory birds are thought to be carrying the virus [9, 40]. If migratory birds had brought the virus one would have expected outbreaks well before February as bird migrations were over by around November. So we consider in our problem the source of outbreaks of bird flu to be the transportation of infected poultry as globalization has turned the chicken into the world's number one migratory bird species. We have assumed that it is mainly due to the human activities of commerce and trade that spread this epidemic. We acquired some important information about bird flu such as the virus H5N1 is sensitive to temperature changes and the virus survives for shorter time at a high temperature. Also there are many effective control measures to block the virus transmission such as compulsory vaccination, culling of all infected or exposed birds [3]. Indeed, in the absence of such control mechanisms this avian influenza may pose a big threat to global health care [17, 29, 42, 45, 53, 61].

## 9.2 Some Biological Preliminaries

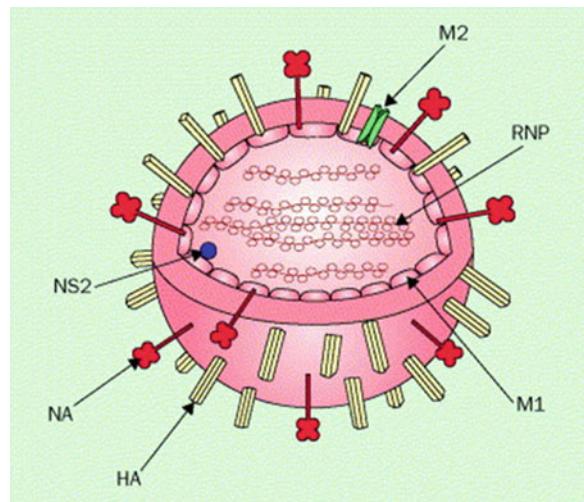
In this section, we present some material that helps one to understand the avian influenza H5N1 virus.

### 9.2.1 *Viral Structure and Taxonomy*

A virus, (an infectious agent that has DNA or RNA and protective covering) is a submicroscopic acellular particle that cannot survive in the absence of a living host cell. It relies on the host cell to replicate and cannot reproduce on its own. Since an antibiotic do not harm a virus, treatment for viral diseases such as flu mainly helps ease the symptoms rather than to kill the viruses. Most viruses cause generally mild diseases like the common cold and some even do not cause any symptom and may go unnoticed, but some cause diseases that can be severe and deadly like Avian influenza, AIDS, SARS, and some form of cancer [57].

The Avian influenza virus (AIV) contains negative (−) sense RNA as genetic material enters the host cell by attachment to the cell surface. Its hemagglutinin binds with the sialic acid present on glycoprotein receptors of the host. After adsorption, it is internalized as an endosome due to the acidic environment of the host cell. Influenza viruses are pleomorphic, mostly spherical or ovoid and filamentous, single-stranded RNA (ssRNA) enveloped viruses with a helical symmetry. They are covered over by lipid/lipoprotein envelope. The viral envelope has lipoprotein membranes that enclose nucleocapsids and nucleoproteins. It is endowed with an inherent capacity for genetic variation and is based on two important features: (1) the presence of a segmented genome, with 8 RNA segments

**Fig. 9.1** Schematic representation of influenza (flu) A virus [27]



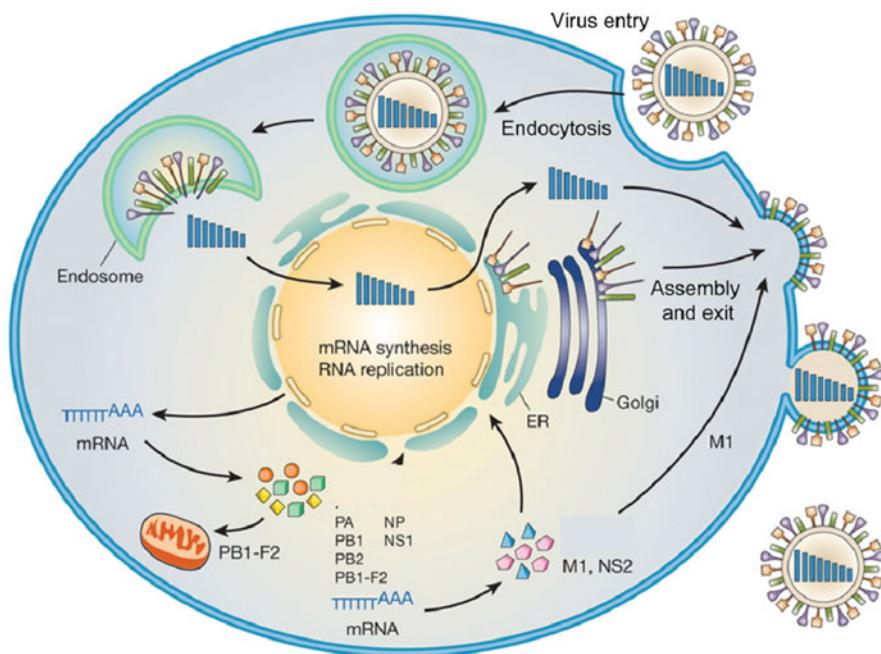
that are genetically independent of each other and (2) a high rate of mutation, especially in the surface hemagglutinin (H) and neuraminidase (N) proteins. The diameter of each enveloped virus ranges from 50–120 nm and filamentous virions are 20 nm in diameter and 200–300 nm long. The genome is in the form of eight negative sense ssRNA fragments. The genome entire length is 12,000–15,000 nucleotides, the largest segment being 23–25 and the smallest being 800–900 nt. The longest RNA strand is closely associated with the nucleoprotein to form helical symmetry (see Fig. 9.1).

There are some 500 distinct spike-like surface proteins of the viral envelope, each projecting 10–14 nm from the surface. There are mainly four types of glycoproteins/antigens:

1. Hemagglutinin (HA) There are 16 types of HA reported.
2. Neuraminidase (NA) There are nine types have been reported. The ratio of HA to NA is about 4–5 to 1.
3. Nucleocapsid protein (NP) It coats the RNA strands.
4. Matrix protein (M) The inner side of the viral envelope is lined by the matrix protein.

In Fig. 9.1, eight ribonucleoprotein segments (RNP) are surrounded by layer of matrix (M1) protein and lipid bilayer taken from host cell at budding. NS2 (NEP) protein is associated with M1. Three viral proteins are incorporated into the lipid bilayer: HA, NA, and M2 protein. HA trimers and NA tetramers form spikes on the surface of the virion. RNP segments contain viral RNA surrounded by nucleoprotein and associated with the polymerase complex [27].

Influenza A replication diagram in presented in Fig. 9.2 and we describe the replication process as follows:



**Fig. 9.2** Influenza A virus replication [4]

The virus loses its envelope in the cytoplasm, and the remaining core moves to the nucleus. Viral proteins are made in the cytoplasm using the host cell's ribosome. In the cell cytoplasm the virus releases its nucleocapsids that further are transported into the nucleus, where mRNA synthesis and replication occurs. Once it enters the nucleus, viral endonuclease snips off. This snipped part of the host mRNA is used as a primer by the virus to synthesize its own mRNA. Next, viral RNA polymerase further extends the primer and makes a complementary (mirror image) plus (+) strand mRNA. Transcription results in eight primary transcripts/pre-mRNA that are further translated in the cytoplasm. The cells treat the viral mRNA like their normal mRNA and use them to make copies of viral proteins. The virus RNA is transcribed to messenger RNA in the nucleus. RNA replication occurs in the nucleus with the help of viral RNA polymerase that was also involved in transcription (see Fig. 9.2). In the same manner, the (+) strand of RNA (e.g., cRNA) is synthesized, and is coated with nucleocapsid proteins soon after it is made. This plus strand is then used as a template to synthesize a new negative RNA strand followed by coating with nucleocapsid proteins. These can further serve as templates for replication, mRNA synthesis, or packaging into virion particles. The (-) strand RNA (e.g., vRNA) are transported into the cytoplasm, where other viral proteins assemble together and are packed into virion particles and, on maturity, buds off from the outer cell membrane and infect new cells.

### 9.2.2 Classification of Influenza Virus

In general the influenza virus or flu virus can be classified into three categories: types A, B, and C which are distinguished by differences in two major internal proteins. Influenza virus type A is the most significant epidemiologically and the most interesting from an ecological and evolutionary stand point, because it is found in a wide variety of bird and mammal species [16, 55] and can undergo major shifts in immunological properties. Type B is largely confined to humans and very little is known about type C. Type A virus is responsible for causing Bird Flu, which was first found in Italy in 1878. Type A virus is further divided into subtypes based on differences in membrane proteins HA and NA, which are the most important targets for the immune system. The notation HhNn is used to refer to the subtype comprising the hth discovered HA proteins and the nth discovered NA protein. The subtype H5N1 virus of type A virus is the main cause of the bird flu [20, 26, 34, 43, 56]. Subtype is further divided into strains; each genetically distinct virus isolated is usually considered to be a separate strain [19, 47]. There are 16 types of HA surface proteins (which are named H1, H2, H3, . . . , H16) and nine types of NA surface proteins (which are named N1, N2, . . . , N9). An influenza virus always has one type of HA surface protein and one NA surface protein and it could be of any combination of H and N, e.g., H5N1, H7N3, H7N7, H5N2, H5N8, and so on [33, 35]. Within these subtypes, some viruses with slightly different nucleotide sequences are present and are classified into strains. The most virulent form so far reported is H5N1 of Influenza virus.

Influenza virus Type A can be divided into two distinct groups on the basis of their ability to cause disease. Highly pathogenic avian influenza (HPAI) can cause up to 100 % mortality in birds [1]. To date, all outbreaks of the highly pathogenic form have been caused by influenza A viruses of subtypes H5 and H7. We will mainly focus on Influenza virus Type A, the most virulent human pathogen and cause of all flu pandemics.

### 9.2.3 Epidemiology and Pathology

Infection with bird flu viruses in domestic poultry causes two main forms of disease that are distinguished by low and high extremes of virulence. The “*low pathogenic*” form may go undetected and usually causes only mild symptoms (such as ruffled feathers). However, the highly pathogenic form spreads more rapidly through flocks of poultry. This form may cause diseases that affect multiple internal organs and has a mortality rate that reaches 90–100 % often within 48 h [15].

The Avian Influenza virus (AIV) refers to Influenza A, found chiefly in birds. Infected birds show clinical symptoms like a sudden drop in egg production, brittle or soft-shelled and even shell-less eggs, congestion, swollen wattles and combs, and swollen skin under the eyes. The risk of human infection from birds is through

coming in close contact with bodily fluids or with contaminated surfaces. Infection can be transmitted from infected bird droppings, saliva, nasal secretions, feces, or blood. These viruses can remain infectious for about one week at human body temperature or a month at 32 °F, and can survive at very low temperatures indefinitely. Symptoms of avian influenza in infected humans are mild fever, myalgia, myositis, and myoglobinuria. However, sore throat, cough, conjunctivitis; some people develop life-threatening complications like respiratory distress syndrome, pneumonia, and multiorgan failure.

In 1997, the first documented infection of humans with an avian influenza virus occurred in Hong Kong. At the same time, the poultry population in Hong Kong was also found to be infected with avian influenza caused by the same pathogenic strain. Studies determined that the infection occurred when the virus jumped directly from birds to humans due to close contact with infected poultry. A pandemic was averted by rapid mass killing/burning of over a million birds—the entire poultry population of Hong Kong [8, 46, 63]. For more than 2 years, the virus has ravaged poultry and caused human illness and death in many Southeast Asian countries and China. In December 2003, a highly pathogenic form of H5N1 caused another outbreak in poultry in South Korea [36]. Another human infection was confirmed in February 2004 when two fatal cases were reported in Hong Kong due to H5N1 [47], followed by 112 cases (57 fatal) from Thailand, Cambodia, Indonesia, and Vietnam. Between April and June 2005, a large number of wild water birds at Qinghai Lake in western China perished after being infected by the virus. During, July–August 2005, outbreaks involving virus were reported from Mongolia, Siberia, and Kazakhstan. The virus reached Turkey, Croatia, Romania, and Greece by October 2005. Ukraine reported outbreaks in November 2005. RnH5N1 viruses also have been isolated from ducks in Southern China [9, 10] and antiviral antibodies have been found in pigs in Vietnam [11]. The virus was infecting chicken and humans in northern Iraq by January 2006. In early February 2006 Nigeria became the first African nation to report the bird flu virus, with an outbreak at a large commercial poultry farm. In February 2006, many European countries, Egypt, and Iran found wild birds infected with H5N1 virus [20, 23, 26, 34, 37, 46]. These cases could be the result of new strains due to reassortant viruses, antigenic shift, or antigenic drift, as explained earlier. People are not immune to these different strains. Generally speaking, an individual has immunity to only those microbes or viruses to which they are earlier exposed. The possibility of dreadful new strains is thus worrying, as people either have no immunity or extremely delayed immunity depending upon the individual's health and age. However, many new harmless strains causing symptomless infections go unnoticed. It was noticed that Spanish flu was most lethal in young adults, who generally are most able to fight off severe infections. One theory why Spanish flu preferentially killed young people is because they are the one with robust and reactive immune systems and therefore were most likely to mount a self-destructive response.

Avian influenza can result in immediate and severe disaster, for example the outbreak in USA in 1983–84 led to destruction of more than 17 million birds at a cost of nearly US\$56 million [32]. Similar case again happened in Hong Kong in

1997–98 [22, 27, 53, 54]. Therefore rapid and effective measures must be taken to stop the spread of epidemics. The most effective measures to prevent the transmission of bird flu are rapid destruction of all infected or exposed birds, proper disposal of carcasses and excrement, quarantining and rigorous disinfectioning of farms and timely use of vaccine [7, 37, 52, 61].

Generally, the virus resides in bird droppings, contaminated soil, and airborne virus. Contaminated equipments, vehicles, food, cages, and clothing like shoes can carry the viruses from farm to farm. Some evidence suggests that flies can also act as mechanical vectors [30]. Wet markets where live birds are sold under crowded and sometimes unsanitary conditions can be another source of spread. These constitute the main cause of the former transmission. Export and import of poultry products are the main cause of the latter transmission, since they can carry the viruses for long distances freely when artificial factors are prevented. Migratory birds can also be a cause of transmission among the countries [9, 41]. Efforts have been made on the study of avian influenza and most of the recent papers focus on topics such as the route of transmission and physiological and biological properties. The bird flu virus of low pathogenicity can mutate into highly pathogenic one after a short time; the virus is sensitive to temperature change (it was found that the virus survives for shorter time at a higher temperature). This kind of influenza is able to transmit to humans under some circumstances; however no sufficient and clear evidences of human-to-human transmission are be found up to now [32].

#### **9.2.4 Recent Efforts in the Control of H5N1**

There are however some studies that aim at finding the possible drug-resistant H5N1 virus are reported in Wang et al. ([60] and the reference there in).

There has been a surge of interest in sensitive specific and rapid detection of avian Influenza virus in recent years and this helps one to find effective diagnosis and disease surveillance. Studies that aim at the detection of viruses by fluorescent DNA barcode-based immunoassay are available in the literature. It has been established that sensitivity of detection is comparable to conventional RT-PCR. For details we refer the readers to Cao et al. ([6] and the references there in).

A study on the first quantification of avian influenza virus in the organs of mute swans which died during the epizootic of H5N1 between January and April 2006 in the Czech Republic has been reported in Rosenbergova et al. [51]. For rapid detection and quantification of avian influenza virus RNA in clinical samples collected from mute swans are utilized to develop the quantitative real time Reverse Transcriptase PCR (qRT-PCR) assay based on a Taq Man Probe.

Though vaccines and antiviral are available that can provide protection from influenza infection, new viral strains emerge continuously due to the plasticity of the influenza genome. An alternative protection methodology that is based on the isolation of a panel of monoclonal antibodies derived from Ig G+ memory B cells of

healthy human subjects that have a capability of recognizing new viral strains is discussed in Grandea III et al. [24].

From bioinformatics point of view to help the researchers develop methods to fight against H5N1 avian flu the following Website <http://www.avian-flu.info> is in vogue since 2004. For more details on this we refer the readers to Liu et al. [39].

## 9.3 Modeling Bird Flu

During the last decade, various mathematical models have been used for infectious diseases in general and for influenza in particular. In case of avian influenza, deterministic models were used for comparing interventions aimed at preventing and controlling influenza pandemics [13, 21], and stochastic model were proposed to model and predict the worldwide spread of pandemic influenza [12, 13]. In this section, we present (1) a deterministic mathematical model which deals with the dynamics of human infection by avian influenza both in birds and in human [18], (2) a one parameter model for spread of H5N1 [28] and (3) the statistical transmission model of bird flu taking into account the factors that affect the epidemic transmission such as source of infection, social and natural factors [59].

### 9.3.1 A Deterministic Mathematical Model

In human, consider SIRS compartmental model that is to say that human susceptible individuals become infectious then removed with temporary immunity after recovery from infection and susceptible when again immunity fades away, in bird population we consider SI compartmental model [18].

Let  $H$  and  $B$  denote the human and bird population sizes respectively. In this model death is proportional to the population size with rate constant  $\mu$  and we assume a constant  $\Omega$  due to births and immigrations. Therefore,  $\frac{dH}{dt} = \Omega - \mu H$ , whereas for bird population we suppose that  $B$  is constant. The human population (respectively bird population) of size  $H$  (resp.  $B$ ) is formed of susceptibles  $S$ , of infective  $I$  and of removed  $R$  (resp.  $S_0$  and  $I_0$ ).  $\beta SI_0/B$  is the human incidence that is, the rate at which susceptible become infective. If the time unit is days, then the incidence is the number of new infection per day. The daily contact rate  $\beta$  is the average number of adequate contacts of a human susceptible with infected birds per day and  $I_0/H$  is the infectious fraction of the population. Time units of weeks, months, or years could also be used. Similarly  $\beta_0 S_0 I_0/B$  is the bird incidence and  $\beta_0$  is the average number of adequate contacts of a bird susceptible with other birds per day. The man life span is taken equal to 25,000 days (68.5 years), and the one for the bird is about 2,500 days (6.8 years).

The model is governed by the following equations:

Human population

$$\begin{aligned}\frac{dS}{dt} &= \Omega - (\mu + \beta I_0/B)S + \delta R, \\ \frac{dI}{dt} &= (\beta I_0/B)S - (\mu + \alpha + \lambda)I, \\ \frac{dR}{dt} &= \lambda I - (\mu + \delta)R, \\ \frac{dH}{dt} &= \Omega - \mu H - \alpha I.\end{aligned}\tag{9.1}$$

Bird population

$$\begin{aligned}\frac{dS_0}{dt} &= \mu_0 B - (\mu_0 + \beta_0 I_0/B)S_0, \\ \frac{dI_0}{dt} &= (\beta_0 I_0/B)S_0 - \mu_0 I_0.\end{aligned}\tag{9.2}$$

The continuity of the right hand side of this system and its derivatives implies that unique solutions exist on maximal interval. Since the solutions approach, enter or stay in a region of attraction  $R$  given by  $\{(S, I, R, S_0, I_0, N) | S \geq 0, S_0 \geq 0, I_0 \geq 0, S + I + R = H \leq \Omega/\mu; S_0 + I_0 = B\}$ , they are eventually bounded and hence exist for  $t \geq 0$ . Therefore, the model is mathematically and epidemiologically well posed.

Introducing the nondimensionalized variables as

$$s = \frac{S}{\Omega/\mu}, \quad i = \frac{I}{\Omega/\mu}, \quad r = \frac{R}{\Omega/\mu}, \quad s_0 = \frac{S_0}{B}, \quad i_0 = \frac{I_0}{B},$$

and with the conditions  $S + I + R = H$  and  $S_0 + I_0 = B$ ,

we obtain  $s + i + r = n$  and  $s_0 = 1 - i_0$ .

The model systems (9.1) and (9.2) become:

$$\begin{aligned}\frac{ds}{dt} &= \mu - (\mu + \beta i_0)s + \delta r, \\ \frac{di}{dt} &= \beta i_0 s - (\mu + \lambda + \alpha)i, \\ \frac{dr}{dt} &= \lambda i - (\mu + \delta)r, \\ \frac{di_0}{dt} &= \beta_0 i_0 (1 - i_0) - \mu_0 i,\end{aligned}\tag{9.3}$$

in the set  $\Omega' = \{(s, i, r, i_0) | s \geq 0, i \geq 0, r \geq 0, s + i + r \leq n \leq 1; 0 \leq i_0 \leq 1\}$ .

Let  $\tilde{R} = \frac{\beta_0}{\mu_0}$ . It has been shown in Derouich and Boutayeb [18] that the model (9.3) admits the trivial equilibrium  $(1, 0, 0, 0)$  if  $\tilde{R} \leq 1$  and this equilibrium is globally asymptotically stable, that is  $\lim_{t \rightarrow \infty} i(t) = 0$  if  $\tilde{R} \leq 1$ . Further the system (9.3) has an endemic equilibrium  $(\bar{s}, \bar{i}, \bar{r}, \bar{i}_0)$  in  $\Omega'$  if  $\tilde{R} > 1$  where

$$\begin{aligned}\bar{s} &= \frac{M\tilde{R}(\mu + \delta)}{\beta(\tilde{R} - 1) \left( M + \frac{\delta}{\mu}(\mu + \alpha) \right) + M\tilde{R}(\mu + \delta)}, \quad \bar{i} = \frac{\beta(\tilde{R} - 1)(\mu + \delta)}{\beta(\tilde{R} - 1) \left( M + \frac{\delta}{\mu}(\mu + \alpha) \right) + M\tilde{R}(\mu + \delta)} \\ \bar{r} &= \frac{\beta\gamma(\tilde{R} - 1)}{\beta(\tilde{R} - 1) \left( M + \frac{\delta}{\mu}(\mu + \alpha) \right) + M\tilde{R}(\mu + \delta)}, \quad \bar{i}_0 = \frac{\mu_0}{\beta_0}(\tilde{R} - 1) \text{ and } M = \mu + \lambda + \alpha.\end{aligned}$$

From the above results it follows clearly that the dynamics of the disease is mainly determined by the average number of adequate contacts of human susceptible with infected birds. For more details we refer the readers to [18].

### 9.3.2 A Discrete Dynamical Model [28]

In Eifert et al. [28], a one component discrete dynamical model for the spread of avian influenza is derived. This model utilizes Lindblad dissipation dynamics [2, 22, 38], for biological rate equation.

It is known in epidemiology that the hygienic stress of the virus ensemble plays an important role in the prevention of this virus. The viruses are damped in their replication rate by hygienic stress and prevention methods, provided these methods are intense. On the other hand if the hygienic stress is minimal, it will stimulate the virus replication.

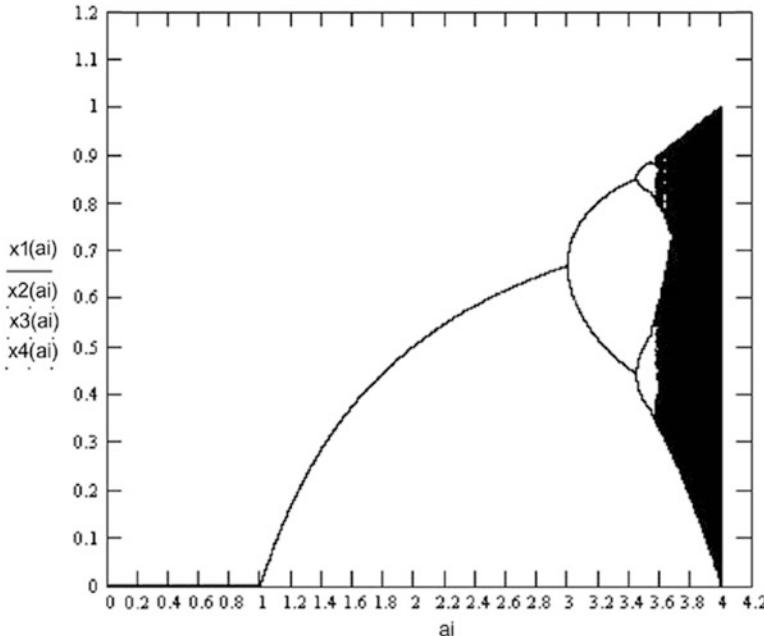
The plain infection model without hygienic stress is given by

$$x_{n+1} = ax_n(1 - x_n). \quad (9.4)$$

where  $x_n$  denotes the relative number of infected being at time step  $n$  and “ $a$ ” is the infection rate. Utilizing the methods suggested in [2, 22, 38], the authors in [28] have modified the model (9.4) incorporating the hygienic stress and assuming a power law relation [44]. The modified model becomes a special case of the following model

$$x_{n+1} = ax_n(1 - x_n) - y^\alpha a^{\beta y} x_n \ln(ax_n), \quad (9.5)$$

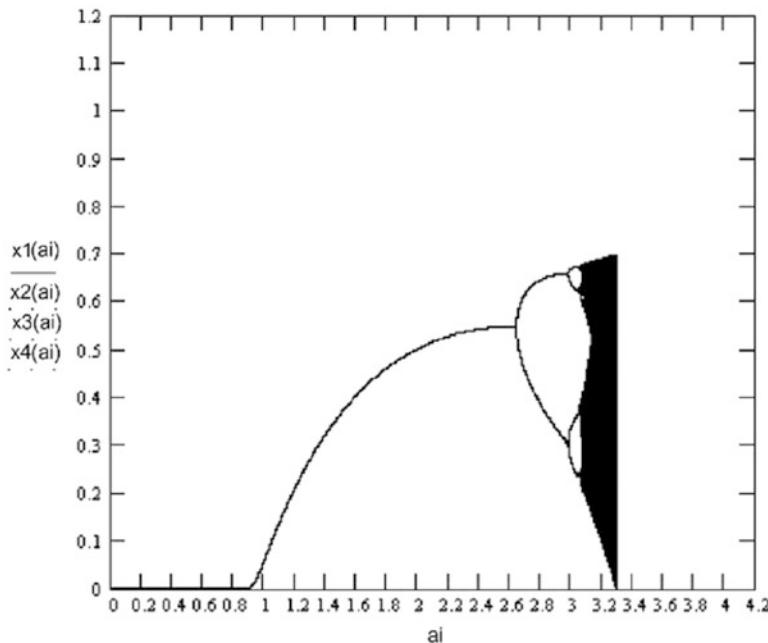
in which  $y$  is a fixed positive number,  $\alpha > 0$ ,  $\beta > 0$  are real positive parameters. The coefficient in front with a power of  $a$ , indicates that the hygienic stress of the



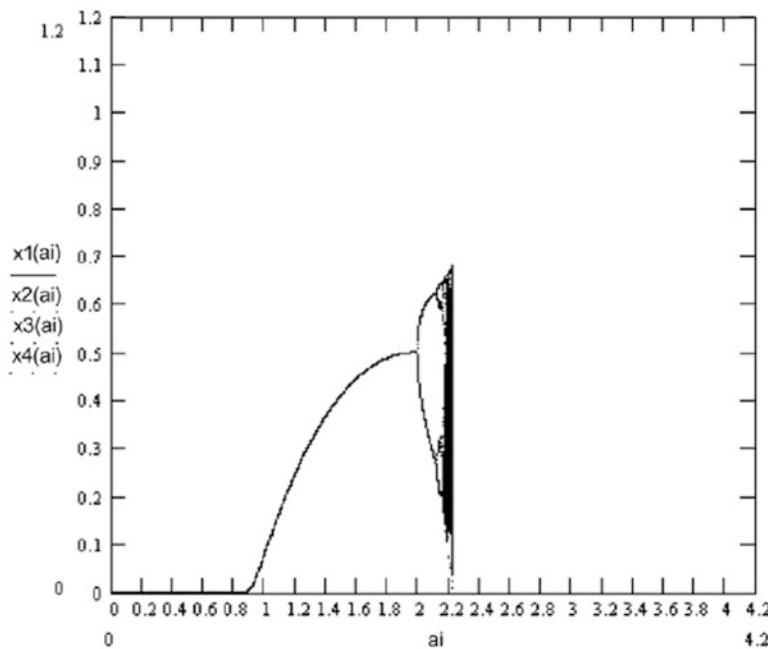
**Fig. 9.3** H5N1 without hygienic stress (May-Feigenbaum scenario for  $y = 0.0$ ) [28]

virus increases with the infection rate, and this power is given by  $\beta y$ . Clearly when the stress coefficient vanishes, model (9.5) should reduce to the plain infection model (9.4). In order that the model (9.5) to reduce to (9.4) in the limiting case, that is, when the stress coefficient vanishes, the term  $y^\alpha$  is introduced. The power law coefficient in front of the term  $x_n \ln (a x_n)$  describes the equation of state for the transport coefficient.

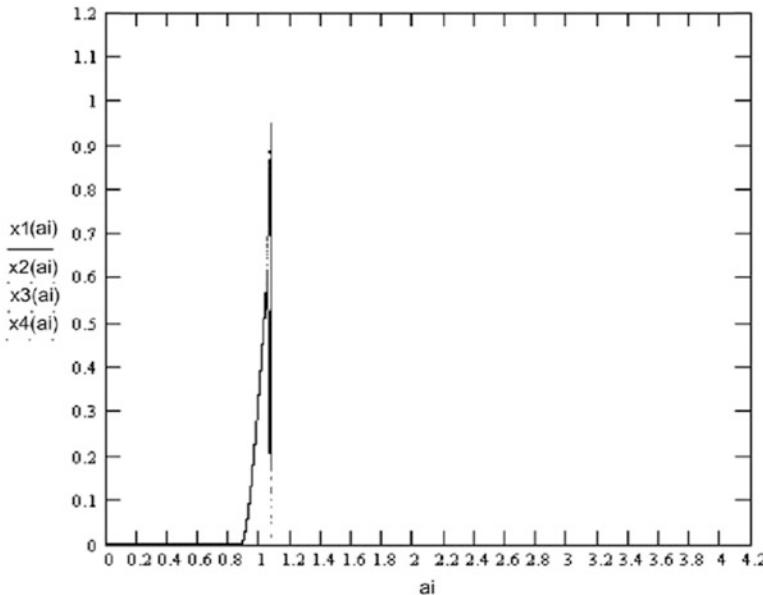
In [28], the authors have considered the model (9.5) with  $\alpha = 1.5$  and  $\beta = 50$  (these numbers are chosen so that the probabilities  $x_n$  stay positive and less than or equal to 1) and performed some numerical calculations. Since the true equation of state for the H5N1 virus is unknown, it is assumed that the hygienic stress coefficient is a monotone function of the infection rate  $a$ . Further model (9.5) reduces to (9.4) when  $y$  vanishes. Computer experiments have revealed that the infection probability is reduced by the hygienic stress and the onset of chaos is earlier as without any hygienic stress. The bifurcation diagrams that explain the spread of avian influenza H5N1 in respect of three specifically chosen stress functions, ( $y = 0.0$ ,  $y = 0.07$ , and  $y = 0.1$ ,  $1.0$ ) are presented in Figs. 9.3, 9.4, 9.5, and 9.6 (Bifurcation diagrams for the spread of avian influenza). For a detailed discussion on these numerics we refer the readers to Eifert et al. [28].



**Fig. 9.4** H5N1 with weak hygienic stress (May-Feigenbaum scenario for  $y = 0.07$ ) [28]



**Fig. 9.5** H5N1 with medium hygienic stress (May-Feigenbaum scenario for  $y = 0.1$ ) [28]



**Fig. 9.6** H5N1 with strong hygienic stress (May-Feigenbaum scenario for  $y = 1.0$ ) [28]

### 9.3.3 *The Statistical Transmission Model*

The mathematical model developed for the recent bird flu predicts that majority of the infection was in wild birds, market birds (includes backyard poultry birds taken to markets for selling) and farm birds, which play an important role in spreading the virus from the epicenter to the nearby centers in the region.

In order to study the patterns of the spread of epidemic, we have made an investigation of outbreaks of the epidemic in 1 week during February 13–18, 2006 till it reached India. On February 18, 2006 the lethal strain of H5N1 virus surfaced in India in the trivial pocket of Navapur in Nandurbar District of Northern Maharashtra. It was the sixth day of our study. On Monday, the first day, outbreaks occurred in five countries of Africa and Eastern Europe namely Nigeria, Greece, Slovenia, Romania, and Bulgaria. On second day bird flu hit four nations of Southeast Asia, central Asia, and Europe namely Indonesia, Iran, Austria, and Germany. On third day, outbreaks took place in two European countries Hungary and Italy. On Thursday, new cases of bird flu were found in three European and African countries Switzerland, Denmark, and Egypt. Reaching France on Friday bird flu at last, on Saturday surfaced in India.

We know the major factors that play an important role in the transmission of bird flu are the way the infected poultry products are transported, air temperature, the control measures (for example, culling the poultry in the infected form, introducing compulsory vaccination to enhance the resistibility of poultry in the noninfected

farms forbidding live birds being sold under crowded and unsanitary conditions), migratory birds, and other infected transportation vehicles (which means the vehicles carry infected poultry or bird dropping or contaminated soil, etc.). Also there are some other factors not considered in our model viz. bird flu transmitting to human beings and virus of low pathogenicity mutating into high pathogenicity after some time, since these elements will not contribute much to the usual transmission of bird flu.

How do these factors affect the transmission? The infected animals are the source of the infection, higher air temperature can drastically cut down the lifetime of the virus, and transportation of infected poultry is the route of transmission. Control measures such as active and effective actions play important role in preventing and destroying the epidemics, which can effectively block the route of transmission of the infection, diminish the source of infection and promote the resistibility of susceptible poultry. So we must take all the major factors into account in the formulation of our transmission model.

These factors may be reflected in the following parameters:

$N(n)$  is the total number of regions of outbreak on the  $n$ th day.

$D(n)$  is the lifetime of the virus regarding the  $n$ th day since the beginning of the epidemic which implicitly corresponds to air temperature.

$I(n)$  is the resistibility of the poultry on the  $n$ th day since the beginning of the epidemic many of the above control measures objectively promote the resistibility of poultry and even human beings.

$f(r)$  is the distribution of the probability that infected poultry products are transported a distance  $r$ .

The following are necessary assumptions for our model [59]:

Let  $P(n, r)$  represent the probability for a new outbreak to take place, then

$$\begin{aligned} P(n, r) &\propto N(n), \\ P(n, r) &\propto D(n), \\ P(n, r) &\propto 1/I(n), \\ P(n, r) &\propto f(r). \end{aligned}$$

Considering above assumptions, we obtain our proposed model for transmission of bird flu as under:

$$P(n, r) = f(r) \cdot R \cdot D(n) \cdot N(n) \cdot \frac{1}{I(n)}, \quad (9.6)$$

where

$$f(r) = r^{-(1+\beta)}.$$

The parameter  $\beta$  is taken close to 0.6 and  $R \in [0,1]$  is a random float number.

### 9.3.3.1 Methodology

In this section, we discuss the methodology that was used for simulation experiments to study the model (9.6). We first discuss, one by one the various parameters taken in our model and explain how these parameters are and finally present the methodology used to predict a new possible outbreak.

1. In our model we have used the random float number  $R$  which we have generated in our program by using random number generator. Why do we need  $R$  in our model? As we know the outbreak of bird flu is a probabilistic instance and not a deterministic one, even though  $f(r)$ ,  $D(n)$ ,  $N(n)$  contribute much to  $P(n,r)$  we cannot definitely assure that there will be a certain outbreak of bird flu in a region but can only say that there is enormous possibility or danger for an outbreak to take place. So an additional random parameter  $R$  is introduced to reflect the uncertainty.
2.  $D(n)$  denotes the lifetime of the virus of bird flu on the  $n$ th day. The H5N1 virus can survive, at cool temperatures in contaminated manure for at least 3 months; in water the virus can survive for up to 30 days at  $0^\circ\text{C}$ ; about four days at  $22^\circ\text{C}$ ; about 3 h at  $56^\circ\text{C}$  and only 30 min at  $60^\circ\text{C}$ .

By means of fitting the above data by curve fitting method, we obtained an approximate formula as follows:

$$D(t) = e^{3.4 - 0.0915 t^{1.1}}, \quad (9.7)$$

where  $t$  represents the air temperature.

In our simulation, we study the epidemic of one week till it reached India that is for a short duration. Hence the temperature change may be taken as a linear approximation regarding the epidemic duration that is

$$t(n) = t_1 + t_2 n, \quad (9.8)$$

in which  $t_1$  and  $t_2$  are two constants that can be determined by fitting the average temperature of the various countries through which the virus reached India. In this model  $t_1 = 0.2$  and  $t_2 = 2.2286$  (started on February 13, 2006). So the relation between the lifetime of the virus and the epidemic duration shall be a compound form of (9.7) and (9.8) as

$$D(n) = e^{3.4 - 0.0915(0.2 + 2.2286n)^{1.1}}. \quad (9.9)$$

3.  $I(n)$  stands for the resistibility of poultry on the  $n$ th day. Obviously the resistibility will increase with the artificial interventions and the control measures. We assume the increase abides by law similar to sigmoid function  $1/(1 + e^{-x})$ . Thus  $I(n)$  assumes the following form

$$I(n) = \frac{B}{1 + (B - 1)e^{-n/C}}. \quad (9.10)$$

Apparently, this is a modified sigmoid form. When  $n = 0$ ,  $I(0) = 1$ , and when  $n = \infty$ ,  $I(n) = B$ ; which indicates the resistibility is impossible to approach a very big number.

4.  $f(r)$  is the distribution of the probability that infected poultry products are transported a distance  $r$ .

$$f(r) = r^{-(1+\beta)}.$$

The basic idea for taking the above-mentioned form of  $f(r)$  stems from Howlett [31]. The researchers at Max Planck Institute for Dynamics and Self Organization, have used the dispersal of dollar bills within the United States as a proxy measurement of human movement (since people cannot be tracked while banknotes travel with the people). They analyzed the data on the peregrinations of more than half-a-million US dollar bills recorded over a 5-year period on an online bill-backing system and given a simple model that only depends upon two parameters. Since the poultry products are also imported and exported very frequently, to various countries, we can assume its transport to be very similar to the human movement. So the probability that infected poultry product is transported from the form may assume a distance  $r$ , for  $r$  larger than 10 km with  $\beta$  close to 0.6. This distribution behaves like a power law. This function decreases as  $r$  grows larger, meaning that transportation of poultry over a long distance is less common than short ones. However it does not decrease, as fast as other common probability distributions, which means that transportation over long distances, are still common enough to have a significant effect. Poultry products make many short journeys, but the occasional long haul ensures that they disperse widely.

5. A threshold value  $S$  is necessary which acts as a criterion: when  $P(n,r)$  is greater than  $S$  there will be an outbreak; otherwise not. However, how to determine the possible number of outbreaks per day denoted by  $K(n)$ ? Intuitively  $K(n)$  shall be in direct proportion to  $N(n)$ , however since only the nearest several outbreaks have notable contribution to the probability of a new outbreak a number of distant outbreaks contribute little, so dependent relation of  $K(n)$  upon  $N(n)$  is of the form

$$K(n) = AN(n)^b, \quad (9.11)$$

in which  $A \geq 1$  and  $0 < b < < 1$ , that is,  $K(n)$  increases slowly with the augment of  $N(n)$ .

Thus the method for predicting an outbreak is explained below.

Suppose the epidemic has begun, we compute the number of actual outbreaks on  $n$ th day. First we generate  $K(n)$  according to (9.11). Then we calculate each  $P(n,r)$

according to (9.6); when  $P \geq S$ , a new outbreak will take place, otherwise not. So, the total number of new outbreaks is always less than  $K(n)$ .

In our model, there are six parameters  $\beta, A, b, B, C, S$  which are all adjustable. So one may argue that too many adjustable parameters may not be an advantage for a “good” model. However, we must analyze independence of these parameters. Here  $S$  is not an independent parameter, rather dependent on  $f(r)$ , hence on  $\beta$ . Parameter  $A$  controls the initial possible number of outbreaks the epidemic may abort if  $A$  is too small and overflow if  $A$  is large. So there shall be a proper intermediate value for  $A$ . The parameter  $b$  denotes the general trend of the outbreaks, the total number of outbreaks will grow too rapidly to be practical if  $b$  is a big number and may be too flat if  $b$  is too small.  $B$  determines the ultimate resistibility, which reflects the final degree of stringency of artificial interventions; the greater  $B$  is the more stringent the interventions are. As of  $C$ , it reflects the average degree of stringency throughout the epidemics; the smaller  $C$  is the more stringent the control measures are. In other words,  $B$  determines the final height of the curve  $I(n)$  and  $C$  controls the shape or the process of  $I(n)$ . Therefore each parameter has a definite meaning and a specific role and has little overlap regarding the role, so the model is reasonable. For conducting simulations we have assumed distances <500 miles as 10 units, <500–1000 miles as 11 units, <1000–1500 miles as 12 units, and <1500–2000 miles as 13 units and so on.

### 9.3.3.2 Numerical Simulation Results

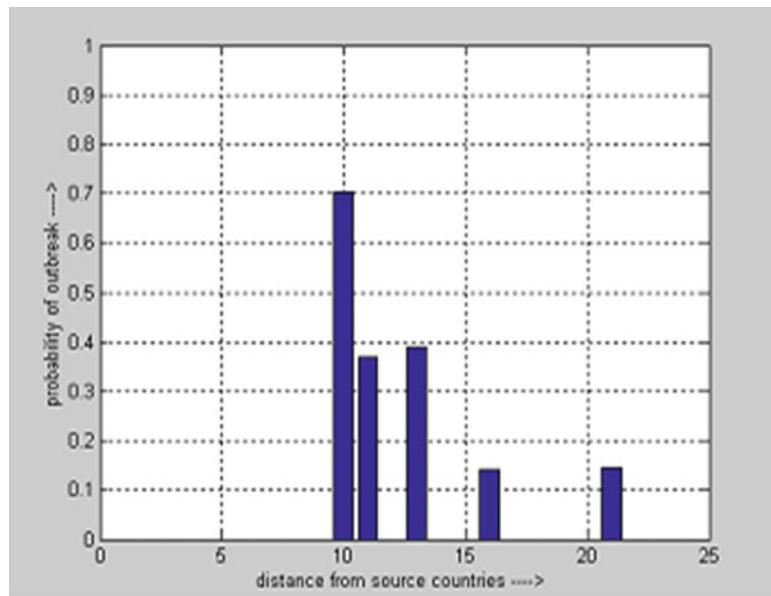
The parameters are initialized as at the beginning:  $N_0 = 6$  refers to the total number of days (for which study is conducted) the epidemic lasts.  $N(1) = 5$  means there are in all five cases of new outbreaks on the first day. The other parameters  $A, b, B, C, S$  will be determined through simulations by comparing with the actual epidemic data.

Results of simulation are shown in Figs. 9.7 and 9.8. The values of the related parameters are:

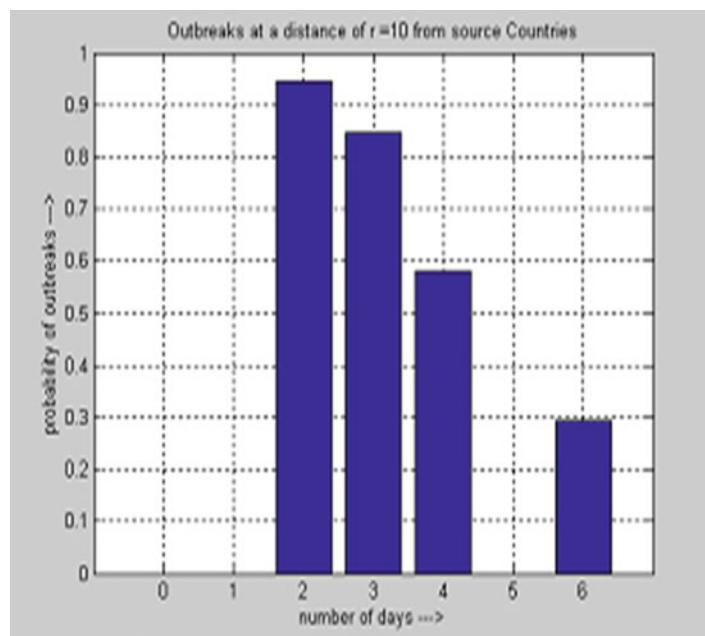
$$\beta \approx 0.6, \quad A = 2, \quad b = 0.33, \quad B = 10, \quad C = 1.75, \quad S = 0.1.$$

It is easy to see that the simulation is roughly in accordance with the actual situation. The parameters related to the artificial interventions include  $B$  and  $C$ . If  $B$  is big enough, few times can the epidemic take place in most simulations; if  $B$  is very small, the epidemic can always happen and most times the overflow can happen. As for  $C$ , if it is small enough, the epidemic hardly occurs, otherwise it can almost happen. Therefore the effective, powerful, and stringent control measures are the key to stop the epidemics.

It should be pointed out that the parameters in the model are independent of each other, since each of them plays an independent role—these parameters have definite meaning, so the result of simulation can hardly coincide with the actual situation of the epidemic if the parameters cannot be adjusted to proper values.



**Fig. 9.7** Graph showing probability of outbreak is inversely proportional to the distance from the source country [59]



**Fig. 9.8** Graph showing the probability of outbreak is inversely proportional to number of days at a fixed distance [59]

**Table 9.1** Probability of outbreaks within a week [59]

Number of days ( $n$ )	Distance from Source ( $r$ )	Probability of outbreaks ( $P$ )	Countries where outbreak occurred
2	10	0.94457699419187	Austria, Germany
	13	0.391542652502286	Iran
	21	0.143943139163055	Indonesia
3	10	0.877733439670344	Italy
	10	0.814356026695878	Hungary
4	10	0.579586050504416	Switzerland
	11	0.333808020423165	Denmark
	11	0.400400629162276	Egypt
5	11	0.376259888883904	France
6	10	0.292578722560819	Bulgaria
	16	0.140831538038502	India

Simulation results are presented in Table 9.1. It is clear from the table that the probability of new outbreak to occur decreases with the increase of distance from the source countries. The probability of occurrence of outbreak in India is very low because the flu virus reaches on the sixth day and also the country is far from the source country. The probability of occurrence of outbreak in Indonesia is also very low even on the second day. This is due to very large distance from the source country and also due to some other factors (like temperature). Figures 9.7 and 9.8 show the pattern that the probability of outbreak to take place follows, depending upon the distance and number of days. These graphs show that the probability of outbreak is inversely proportional to the distance from the source country as well as the day on which the flu virus reaches the country.

## 9.4 Other Control Measures

The first and major control measure should be to take precautions so that an Avian Influenza virus does not cross the “species barrier” and infect human. If it crosses the “species barrier” and infects people, then the medical doctor should be contacted to take the proper medicine and suggestion for precautions. The persons dealing with poultry or poultry products should take the following prevention measures:

1. Practice proper sanitation and good hygiene handling of poultry and poultry products,
2. Wear masks, rubber gloves, and safety glasses,
3. Avoid unnecessary contact with live, sick, or dead birds,
4. Keep hands away from face and mouth when handling birds,
5. Practice proper hand washing, cleaning and disinfection procedures, and avoid touching mouth or eyes. Indirect contact with door knobs, toilet knobs, taps, hand shaking etc., is another possible method of transmission,

6. Wash eggs thoroughly with soap water, rinse, and thoroughly cook at the temperature more than 74 °C, cook chicken until boiling temperature or when cooking temperature exceeds 165 °F (74 °C), because H5N1 viruses are killed at this temperature,
7. Store and prepare raw chicken and eggs separately from other food items to avoid cross contamination,
8. Clean kitchen utensils and surfaces before and after use,
9. There should be routine tests for AIV in poultry and if any positive case is seen, the entire poultry flock and poultry products should be destroyed and should not be consumed by the public at any cost.
10. Report any unusual death or illness of chickens and other birds to the Dead Bird Hotline (888) 551-4636 and report illness among workers in poultry farms to the Health Department.

## 9.5 Discussion and Conclusions

Bird flu is a highly pathogenic epidemic that can result in serious disaster in many areas. Immediate and effective control measures are of great importance in preventing the transmission of avian influenza. So it is challenging to study the problem from various angles, and develop deterministic, discrete and statistical-mathematical transmission models. Using mathematical modeling, Breban et al. [5] have investigated the role of environmental transmission for the pattern and persistence of avian influenza in wild waterfowl and demonstrated that indeed environmental transmission is a fundamental ingredient for the modeling of this epidemic. The persistence mechanism induced by environmental transmission raises novel problems of epidemic control since traditional strategies may prove ineffective in the presence of an environmental viral reservoir [50]. For a global outbreak of influenza to occur, three conditions must be met: a new virus subtype must arise, this subtype must be able to cause serious illness in the human, and it must spread easily from person to person- and continue to do so. The first two conditions have been met with H5N1 human infections. Scientists are working for better understanding of different influenza A strains with the hope of preventing the spread of this virus. One of the reasons for the difficulty in predicting outbreaks is that migratory wild birds, not virus infected human carry the H5N1 virus long distance [62] (one cannot totally rule out the chance of human carrying bird flu for long distances).

Our ongoing investigations relate to the question as how to assess the actual control measures and assign the parameters in the model with proper numerical values. There is a need to strengthen the design of data collection and for advanced experimental facilities to understand the virus spread with respect to all seasons and food consumption patterns. Given the present situation, it is not easy for any government to predict accurately the timing and location of future avian influenza attacks; there is a vide scope to strengthen the databanks and to train relief teams to minimize economic loss, in the event that an outbreak occurs [48]. To complement

these issues with public health side of research activities, the government could consider launching large-scale experimental projects to estimate various rates of spread of H5N1 in different geographical regions, and to come up with a comprehensive approach for effective control of an outbreak. Further modeling work is needed in order to explore the epidemiological dynamics and persistence of avian influenza viruses, with a view to understanding the respective roles of environmental transmission and demographic stochasticity.

**Acknowledgment** The research of the first author is supported by a grant from the Foundation for Scientific Research and Technological Innovation (FSRTI)—A Constituent Division of Sri Vadrevu Seshagiri Rao Memorial Charitable Trust, Hyderabad, 500035, India. This work is supported by University Grants Commission, Govt. of India under grant no. F. No. 42-16/2013 (SR) to Prof. R.K. Upadhyay.

## References

1. Alexander DJ (2000) A review of avian influenza in different bird species. *Vet Microbiol* 74:3–13
2. Alicki R, Messer J (1983) Nonlinear quantum dynamical semigroups for many-body open systems. *J Stat Phys* 32:299–312
3. Anderson RM, May RM (1991) Infectious diseases of humans: dynamics and control. Oxford University Press, Oxford UK
4. Arias CF, Escalera-Zamudio M, Soto-Del Río Mde L, Cobia'n-Guemes AG, Isa P, Lo'pez S (2009) Molecular anatomy of 2009 influenza virus A (H1N1). *Arch Med Res* 40:643–654
5. Breban R, Drake JM, Stallknecht DE, Rohani P (2009) The role of environmental transmission in recurrent avian influenza epidemics. *PLoS Comput Biol* 5(4):1–11
6. Cao C, Dhumpa R, Bang DD, Ghavifekr Z, HØgberg J, Wolff A (2010) Detection of avian influenza virus by fluorescent DNA barcode-based immunoassay with sensitivity comparable to PCR. *Analyst* 135:337–342
7. Centers for Disease Control and Prevention (2004) Cases of influenza A (H5N1)-Thailand, 2004. *MMWR Morb Mortal Wkly Rep* 53:100–103
8. Chan PKS (2002) Outbreak of avian influenza A (H5N1) virus infection in Hong Kong 1997. *Clin Infect Dis* 34:s58–s64
9. Chen H, Smith GJD, Zhang SY, Qin K, Wang J, Li KS (2005) Avian flu: H5N1 virus outbreak in migratory waterfowl. *Nature* 436:191–192
10. Chen H, Deng G, Li Z, Tian G, Li Y, Jiao P (2004) The evolution of H5N1 influenza viruses in ducks in southern China. *Proc Natl Acad Sci U S A* 101:10452–10457
11. Choi YK, Nguyen TD, Ozaki H, Webby RJ, Puthavathana P, Buranathal C (2005) Studies of H5N1 influenza virus infection of pigs by using virus isolated in Vietnam and Thailand in 2004. *J Virol* 79:10821–10825
12. Colizza V, Barrat A, Barthelemy M, Valleron AJ, Vespignani A (2007) Modeling the world wide spread of pandemic influenza: baseline case and containment interactions. *PLoS Med* 4: e13
13. Colizza V, Barrat A, Barthelemy M, Valleron AJ, Vespignani A (2006) The modeling of global epidemics: stochastic dynamics and predictability. *Bull Math Biol* 68:1893–1921
14. Conlan AJK, Grenfell BT (2007) Seasonality and the persistence and invasion of measles. *Proc Biol Sci* 274:1133–1141
15. Cox NJ, Subbarao K (2000) Global epidemiology of influenza: past and present. *Annu Rev Med* 51:407–421

16. Cyranoski D (2005) Bird flu spreads among Java's pigs. *Nature* 435:390–391
17. De Jong JC, Claas EC, Osterhaus AD, Webster RG, Lim WL (1997) A pandemic warning? *Nature* 389:554
18. Derouich M, Boutayeb A (2008) An avian influenza mathematical model. *Appl Math Sci* 2:1749–1760
19. Earn DJD, Rohani P, Bolker BM, Grenfell BT (2000) A simple model for complex dynamical transitions in epidemics. *Science* 287:667–670
20. Ellis TM, Bousfield BR, Bissett LA, Dyrting KC, Luk GS, Tsim ST (2004) Investigation of outbreaks of highly pathogenic H5N1 avian influenza in waterfowl and wild birds in Hong Kong in late 2002. *Avian Pathol* 33:492–505
21. Ferguson NM, Fraser C, Donnelly CA, Ghani AC, Anderson RM (2004) Public health risk from the avian H5N1 influenza epidemic. *Science* 42(4):968–969
22. Gorini V, Frigerio A, Verri M, Kossakowski A, Sudarshan ECG (1978) Properties of quantum Markovian master equations. *Rep Math Phys* 13:149–173
23. Govorkova EA, Rehg JE, Krauss S, Yen HL, Guan Y, Peiris M (2005) Lethality to ferrets of H5N1 influenza viruses isolated from humans and poultry in 2004. *J Virol* 79:2191–2198
24. Grandea AG III, Olsen OA, Cox TC, Renshaw M, Hammond PW, Chan-Hui PY, Mitcham JL, Cieplak W, Stewart SM, Grantham ML, Pekosz A, Kiso M, Shinya K, Hatta M, Kawaoka Y, Moyle M (2010) Human antibodies reveal a protective epitope that is highly conserved among human and nonhuman influenza A viruses. *Proc Natl Acad Sci U S A* 107(28):12658–12663
25. Greenwood B (2006) Pneumococcal meningitis epidemics in Africa. *Clin Infect Dis* 43:701–703
26. Guan Y, Peiris JSM, Lipatov AS, Ellis TM, Dyrting KC, Krauss S (2002) Emergence of multiple genotypes of H5N1 avian influenza viruses in Hong Kong SAR. *Proc Natl Acad Sci U S A* 99:8950–8955
27. Gabareva LV, Kaiser L, Hayden FG (2000) Influenza virus neuraminidase inhibitors. *Lancet* 355(9206):827–835
28. Eifert H-J, Held S, Messer JA (2009) A one-parameter model for the spread of Avian Influenza A/H5N1. *Chaos Soliton Fract* 41(5):2271–2276
29. Hien TT, De Jong M, Farrar J (2004) Avian influenza—a challenge to global health care structures. *N Engl J Med* 351:2363–2365
30. Hope-Simpson RE (1992) The transmission of epidemic influenza. Plenum, New York
31. Howlett R (2006) Travel: fitting the bill. *Nature* 439:402
32. Li J, Ren Q, Chen Xi, Yin J (2004) Study on transmission model of Avian influenza. Proceedings of international conference on Information Acquisition, IEEE, p 54–58
33. King AA, Ionides EL, Pascual M, Bouma MJ (2008) Inapparent infections and cholera dynamics. *Nature* 454:877–880
34. Kuiken T, Rimmelzwaan G, Van RD, Van AG, Baars M, Fouchier R (2004) Avian H5N1 influenza in cats. *Science* 306:241
35. Lee CW, Senne DA, Suarez DL (2004) Effect of vaccine use in the evolution of Mexican lineage H5N2 avian influenza virus. *J Virol* 78:8372–8381
36. Lee CW, Suarez DL, Tumpey TM, Sung HW, Kwon YK, Lee YJ (2005) Characterization of highly pathogenic H5N1 avian influenza A viruses isolated from South Korea. *J Virol* 79:3692–3702
37. Li KS, Guan Y, Wang J, Smith GJ, Xu KM, Duan L (2004) Genesis of a highly pathogenic and potentially pandemic H5N1 influenza virus in eastern Asia. *Nature* 430:209–213
38. Lindblad G (1976) On the generators of quantum dynamical semi groups. *Commun Math Phys* 48:119–130
39. Liu D, Liu QH, Wu LH, Liu B, Wu J, Lao YM, Li XJ, Gao GF, Ma JC (2009) Website for avian flu information and bioinformatics. *Sci China Ser C-Life Sci* 52(5):470–473
40. Liu J, Xiao H, Lei F, Zhu Q, Qin K, Zhang X (2005) Highly pathogenic H5N1 influenza virus infection in migratory birds. *Science* 309:1206
41. Liu M, He S, Walker D, Zhou NN, Perez DR, Mo B (2003) The influenza virus gene pool in a poultry market in south central China. *Virology* 305:267–275

42. Longini IM Jr, Nizam A, Xu S, Ungchusak K, Hanshaoworakul W, Cummings DA (2005) Containing pandemic influenza at the source. *Science* 309:1083–1087
43. Mase M, Tsukamoto K, Imada T, Imai K, Tanimura N, Nakamura K (2005) Characterization of H5N1 influenza A viruses isolated during the 2003–2004 influenza outbreaks in Japan. *Virology* 332:167–176
44. Messer JA (2009) A non-equilibrium phase transition in a dissipative forest model. *Chaos Soliton Fract* 41(5):2456–2462
45. Mills CE, Robins JM, Lipsitch M (2004) Transmissibility of 1918 pandemic influenza. *Nature* 432:904–906
46. Peiris JS, Lai ST, Poon LL, Guan Y, Yam LY, Lim W, Nicholls JM, Yee WK, Yan WW, Cheung MT, Cheng VC, Chan KH, Tsang DN, Yung RW, Ng TK, Yuen KY (2003) Coronavirus as a possible cause of severe acute respiratory syndrome. *Lancet* 361(9366):1319–1325
47. Peiris JS, Yu WC, Leung CW, Cheung CY, Ng WF, Nicholls JM, Ng TK, Chan KH, Lai ST, Lim WL, Yuen KY, Guan Y (2004) Re-emergence of fatal human influenza A subtype H5N1 disease. *Lancet* 363(9409):617–619
48. Rao ASRS (2008) Location of the epicenter of avian bird flu might determine the rapidity of its spread in India. *Curr Sci* 95:313–315
49. Rohani P, Earn D, Greenfell B (2000) Impact of immunisation on pertussis transmission in England & Wales. *Lancet* 355:285–286
50. Rohani P, Breban R, Stallknecht DE, Drake JM (2009) Environmental transmission of low avian influenza viruses and its implications for pathogen invasion. *Proc Natl Acad Sci U S A* 106(25):10365–10369
51. Rosenbergova K, Lany P, Pospisil Z, Kubicek O, Celer V, Molinkova D (2009) Quantification of avian influenza virus in tissues of mute swans using TaqMan real time qRT-PCR. *Vet Med* 54(9):435–443
52. Seo SH, Peiris M, Webster RG (2002) Protective cross-reactive cellular immunity to lethal A/Goose/Guangdong/1/96-like H5N1 influenza virus is correlated with the proportion of pulmonary CD8+ T cells expressing gamma interferon. *J Virol* 76:4886–4890
53. Shortridge KG, Zhou NN, Guan Y, Gao P, Ito T, Kawaoka Y (1998) Characterization of avian H5N1 influenza viruses from poultry in Hong Kong. *Virology* 252:331–342
54. Sims LD, Ellis TM, Liu KK, Dyrting K, Wong H, Peiris M (2003) Avian influenza in Hong Kong 1997–2002. *Avian Dis* 47:832–838
55. Stech J, Xiong X, Scholtissek C, Webster RG (1999) Independence of evolutionary and mutational rates after transmission of avian influenza viruses to swine. *J Virol* 73:1878–1884
56. Sturm-Ramirez KM, Ellis T, Bousfield B, Bissett L, Dyrting K, Rehg JE (2004) Re-emerging H5N1 influenza viruses in Hong Kong in 2002 are highly pathogenic to ducks. *J Virol* 78:4892–4901
57. Suri S (2007) Avian influenza (Bird flu), ProQuest CSA LLC discovery guide, pp 1–11. <http://www.csa.com/discoveryguides/discoveryguides-main.php>
58. Teyssoeu R, Rouzic EML (2007) Meningitis epidemics in Africa: a brief overview. *Vaccine* 25: A3–A7
59. Upadhyay RK, Kumari N, Sree Hari Rao V (2008) Modeling the spread of bird flu and predicting outbreak diversity. *Nonlinear Anal Real World Appl* 9(4):1638–1648
60. Wang SQ, Du QS, Chou KC (2007) Study of drug resistance of chicken influenza A virus (H5N1) from homology-modeled 3D structures of neuraminidases. *Biochem Biophys Res Commun* 354:634–640
61. Webster RG, Bean WJ, Gorman OT, Chambers TM, Kawaoka Y (1992) Evolution and ecology of influenza A viruses. *Microbiol Rev* 56:152–179
62. Webster RG, Hulse D (2005) Controlling avian flu at the source. *Nature* 435:415–417
63. Yuen K, Chan PKS, Peiris M (1998) Clinical features and rapid viral diagnosis of human disease association with avian influenza A (H5N1) virus. *Lancet* 351:467–471

# Index

## A

Allele dynamics plots (AD-plots), human influenza A viruses  
analysis  
    evolutionary dynamics, 220–221  
    population genetics, 209–210  
construction, 210–211  
frequency calculation, 210  
H1N1  
    changes, 219  
    evolution, 218–219  
H3N2  
    antigenic variant BR07, 214  
    consecutive trunk branches,  
        phylogenetic tree, 213  
    fixation, 212–213  
    identification, 214–217  
    maximum likelihood tree, 212  
    surface protein and antigenic  
        determination, 212, 213  
limitation, 220  
phylogenetic  
    inference and reconstruction, 206,  
        208, 210  
    tree, 208, 210, 211  
sensitive and timely method, 220  
substitutions, 210, 211  
visualizing, 209

Avian influenza (H5N1)  
    causing countries, 233  
    classification, influenza virus, 232  
    control, 234–235, 247–248  
    and H1N1, 227  
    infection, 232–233  
    influenza A, humans, 228  
    issues, 247–248

## modeling

    discrete dynamical, 237–240  
    mathematical, 235–237  
    statistical transmission (*see* Statistical  
        transmission model, H5N1)  
outcomes, immediate and severe disaster,  
    233–234, 247  
pathogenic form, 232  
persistence mechanism, 247  
prediction, 247  
reporting, 228  
and SARS, 227  
symptoms, 233  
transmission and outbreak, 227–229,  
    234, 247  
viral structure and taxonomy  
    glycoproteins/antigens, 230  
    helical symmetry, 230  
    lipid/lipoprotein membranes, 229–230  
    negative (–) sense RNA, 229  
    replication process, 230–231  
    ribonucleoprotein segments (RNP), 230  
    single-stranded RNA (ssRNA), 229

## B

Bacillus of Calmette and Guerin (BCG)  
    vaccine, 61

Bernoulli percolation model  
    asymptotic behavior, clusters, 38  
    clusters, 36  
    mathematical, 35  
    measures, epidemics development, 37  
    probability, 37  
    qualitative properties, 36  
    short-range, 38

- Bernoulli percolation model (*cont.*)  
 spatial structure, 36  
 standard, 35–36  
 two-dimensional site, probability, 37
- Bird Flu. *See* Avian influenza (H5N1)
- C**
- Case detection  
*C. gattii* (*see* *Cryptococcus gattii* (*C. gattii*), North America)  
 TB (*see* Tuberculosis (TB), Nigeria)
- C. gattii*. *See* *Cryptococcus gattii* (*C. gattii*), North America
- Chinese MSM sexual behaviors  
 analysis, 167  
 annual HIV incidence, 167  
 contribution, 149  
 data, 170, 171  
 distributions, sexual partnerships, 162, 163  
 diverse, 162, 163  
 fittings and optimization  
   acts, male partners, 156, 162  
   bootstrapping, 160  
   condom usage, 157–158, 162  
   data collection, 151  
   differences, 160  
   implementation, geographical location, 151  
   incidence rates, HIV infection, 160–161  
   lognormal distribution, 160  
   mean and median values, male sexual partner, 151–155, 162  
   probability distributions, 151  
   types, sexual partnerships, 151, 160  
 heterogeneous distributions, 171  
 high-risk and partnerships types, 148, 166–167
- HIV  
 annual incidence, 2002–2010, 165  
 control and prevention, 149, 170  
 incidence, 168–169  
 transmission, 149
- homosexual men, 148
- knowledge gap, 170
- limitation, 170
- literature research, 170–171
- lognormal distribution function, 161
- mean and median values, 162
- regular, commercial and noncommercial casual partners, 148, 171
- reporting, condom usage, 149
- review, published literature, 149
- sexual partners and acts, 162, 164, 167–170
- Spearman correlation tests, 162  
 types, 162, 163  
 United States, 168
- Communities  
 control and prevention, TB, 61  
 DOT, 62  
 modular networks (*see* Modular networks with community)  
 transmission dynamics, tuberculosis, 115  
 XDR-TB, 133
- Cooperative and supportive neural networks (CSNN)  
 activation, 6  
 algorithm, building and training, 7–8  
 architecture, 7–8  
 estimation, transmission rate  
   bilinear incidence, 17–18  
   data, k-means clustering algorithm, 14  
   sublinear interactions, 17  
   testing procedure, 15  
   three-layer architecture, 14  
   training, 15–16  
 operating mechanism, 7
- Cryptococcus gattii* (*C. gattii*), North America  
 British Columbia, 186  
 and *C. neoformans*, 177–178  
 CNS infections, 187  
 control (*see* Prevention and control)  
 diagnosis  
   cost-effective, 187  
   histopathological examination, 187  
   lateral flow assay (LFA), 190  
   L-canavanine glycine thymol (CGB) agar, 187  
   lumbar puncture and serologic cryptococcal antigen testing, 190  
   molecular typing methods, 188  
   serological analysis, 187, 190  
   slide agglutination test, 188  
   traditional culture, 187  
 ecological modeling, 192–193  
 epidemiology  
   geographic distribution and habitat, 179–180  
   molecular, 180–181  
 HIV/AIDS, 186  
 immunology and host factors, 184–185  
 life-threatening fungal disease, 177  
 microbial pathogenesis  
   *C. neoformans* infections, 182  
   CNS, 182  
   nutrient agar and yeast, 181  
   polysaccharide capsule, 181

- microbiology and taxonomic classification, 178–179
- Pacific Northwest outbreak (*see* Pacific Northwest outbreak, *C. gattii*)
- respiratory symptoms and pulmonary cryptococcosis, 186
- transmission, 186
- treatment
- amphotericin B and flucytosine, 189, 190
  - CNS, 189
  - high-dose fluconazole, 190
  - IDSA guidelines, 188
  - induction therapy, CNS infections, 188
  - optimal, 188
  - oral steroid, 189
  - in vitro susceptibilities, 189
- Vancouver Island outbreak, 1999, 177
- virulence factors
- bacteria and yeast, 183–184
  - CNS, 184
  - colony characteristic changes, 183
  - differences, pathogenic and nonpathogenic, 183
  - formidable, 182
  - hyphal formation, 183
  - melanization, 182
  - phospholipase and protease production, 184
  - polysaccharide capsular components, 182
  - sexual reproduction, yeast, 184
  - starvation, 184
- CSNN. *See* Cooperative and supportive neural networks (CSNN)
- D**
- DEPS. *See* Directional evolution of protein sequences (DEPS)
- Directed
- DOTS (*see* Direct Observation Therapy Strategy (DOTS))
  - drug-resistant strains, 133, 135
  - HIV intervention, 169
  - spoligotyping, 116
- Directional evolution of protein sequences (DEPS), 216, 219–221
- Directional selection
- H3N2
    - alleles frequency, consecutive seasons, 214, 215
    - antigenic cluster transition, 216, 217
  - DEPS method, 216
- maximum likelihood test, 216
- observations, AD-plot, 214
- strains, 216, 217
- variants, 214–217
- worldwide epidemics, 1998 and 2008, 211, 214
- Direct Observation Therapy Strategy (DOTS)
- control TB, 63
  - drugs, 62
  - equilibrium states and reproduction number
    - asymptotic properties, SEIJT and SEIJS models, 67–68
    - disease-free, 67
    - guidance, 67
    - mathematical modeling, 67
    - next generation matrix approach, 68–69
  - HIV, 62
  - imperfect screening coverage
    - detected infectious individuals, 65–66
    - effectiveness, 64
    - minimal level detection rate, 64
    - SEIT, 64, 66–67
    - strong dose-dependent effects, 65
    - transmission, 65
    - variables and parameters, 66
  - infection elimination and reduction, 71–72
  - level of coverage and treatment
    - direct Lyapunov method, 74
    - disease-free equilibrium, 74–75
    - failure, 72
    - latency periods, 75
    - SEIT model, 72
  - mathematical models, 63
  - mortality, 62
  - properties, disease-free equilibrium, 69–70
  - self-treatment, 63
  - WHO recommendation, 97
- DOTS. *See* Direct Observation Therapy Strategy (DOTS)
- Drug resistance
- bacterial fitness and compensatory evolution
    - isoniazid, rifampin, and MDR strains, 138
    - katG* gene, 135–136
    - MDR and XDR strains, 135
    - rifampin-resistance mutations, 136
    - Salmonella enterica*, 137
    - tuberculosis resistance, 135
  - tuberculosis control and emergence
    - MDR-TB (*see* Multidrug-resistant tuberculosis (MDR-TB))
    - TDR-TB (*see* Totally drug-resistant strains of *M. tuberculosis* (TDR-TB))

**D**rug target

- current anti-tubercular therapy, 97
- development, 97–98
- impact, latency, 98
- limitations, current therapies, 97
- TCS, potential targets, 98

**E**

ENN. *See* Ensemble of neural networks (ENN)

## Ensemble of neural networks (ENN)

- architecture, RNENN, 10–11
- estimation (*see* Mathematical modelling)
- multi-layer perceptron, 10
- pattern, 9
- performance and regression plots, 11–12
- procedure, 9
- steps, 9
- verification and validation procedure, 10–12

## Entire gene sequences (MLSA), 118

## Epidemics

- bird flu (*see* Avian influenza (H5N1))
- DOTS, 73, 76
- HIV, 62
- HIV, China (*see* Human immunodeficiency virus (HIV), China)
- H3N2 viruses (*see* Human influenza A (H3N2) viruses)
- pathogenicity and phylogeographic distribution, strains, 115
- SEIR (*see* SEIR epidemic and percolation models)
- swine flu (*see* Swine-origin influenza A (H1N1) virus)
- tuberculosis infection, 132

## Extensive-drug-resistant (XDR)

- Beijing strain, 134
- and MDR, 133, 134, 138
- and MDR-TB, 132

**F**

## Fixed infectious periods, 39–40

**G**

Great gerbils, Kazakhstan

- plague epidemic, 40
- SIR epidemic, 40
- standard percolation models, 40

**H**

## Hierarchical

- lattice, 48
- levels, 50
- modular spatial random network, 52, 53
- scale-free networks, 50–52
- structure, modular network, 49–50

## H5N1 hygienic stress, discrete dynamical model

- medium, 238, 239
- and prevention, 237
- strong, 238, 240
- weak, 238, 239
- without stress, 238

## Human immunodeficiency virus (HIV), China

## challenges and limitations

- condom usage level, 157–158, 169
- deaths, 169
- MSM, 169–170
- prevalence, 169
- diagnosis, drug users, 147
- DOTS, 62
- epidemic development, 147
- estimation, annual incidence
  - Chinese MSM, 2002–2010, 165
  - cumulative density, 165–166
  - high-risk sexual behaviors and partnerships types, 166–167
  - new infections, 166
  - rates, 163
  - sexual types, 163
- implications, 169
- literature search and selection, 161
- meta-analysis, 148
- MSM sexual behaviors (*see* Chinese MSM sexual behaviors)
- statistics, 148
- systematic review
  - acts, male partners, 151, 156
  - condom usage, 151, 157–158
  - English and Chinese electronic databases, 149
  - inclusion and exclusion criteria, 149–150
  - literature filtering procedure, 150
  - male sexual partners, 151–155

## Human influenza A (H3N2) viruses

- AD-plots (*see* Allele dynamics plots (AD-plots), human influenza A viruses)
- antigenic cartography, 208
- changes
  - co-occurring inter-and intra-segment, 208

- hemagglutinin HA1 subunit  
sequence, 207
- directional selection, HA segment, 208
- East and Southeast Asia, 207
- evolution, 207
- genetic data, 208–209
- genome-wide phylogenetic analysis, 207
- maximum likelihood tree, AD-plots, 212
- phylogenetic  
inference, 209  
tree, 208  
structure, 208
- I**
- IDSA guidelines. *See* Infectious Diseases Society of America (IDSA) guidelines
- Infectious diseases control  
analogy, ecological problems, 3  
assimilation data, 5  
biological principles, 2  
conditions, epidemiological setting, 3  
contract, epidemic outbreak, 1  
design and strategies, 1  
development, analytical and statistical models, 1  
hypotheses, 4–5  
incidence data, 27  
informatics  
CSNN (*see* Cooperative and supportive neural networks (CSNN))  
ENN (*see* Ensemble of neural networks (ENN))  
k-means clustering, 6  
multiple linear regression, 5–6  
statistical and data mining methods, 5  
transmission, 5
- nonlinear interactions, 3
- prototypes, 3
- qualitative properties, 3
- spreading, susceptible population, 2
- stabilizing process, mutual interference, 3
- stimulation  
bilinear incidence, 13  
outcomes, 12  
sublinear interactions, 13–14  
vaccination efforts, 3
- Infectious Diseases Society of America (IDSA) guidelines, 188
- Influenza A viruses  
AD-plots (*see* Allele dynamics plots (AD-plots), human influenza A viruses)
- antigenic drift, 207
- antigenic variant identification, 217–218
- computational techniques, 221
- dN/dS ratio tests, 221
- evolution, 206
- H1N1 (*see* Swine-origin influenza A (H1N1) virus)
- H3N2 (*see* Human influenza A (H3N2) viruses)
- Orthomyxoviridae*, 206
- phylogenetic techniques, 206
- reconstruction, phylogenetic, 207
- surveillance, 207
- types, 206
- vaccine, 207
- L**
- Long-range percolation, 47–48
- M**
- M. africanum*  
burden, human tuberculosis, 131  
*Dassie bacillus*, 131  
drug-resistant tuberculosis, 132  
host genetic polymorphism, 132  
human-associated lineages, 131  
*Mycobacterium mungi*, 131  
progression, disease, 132
- Mathematical modelling  
bird Flu  
human and bird population, 236  
SI compartmental, 235  
trivial and endemic equilibrium, 237
- cellular processes and molecular interactions, 99
- Chinese MSM, 149
- DevS-DevR-DosT interaction network, 99–101
- DOTS, 63, 73, 76
- ENN parameters, estimation  
analysis and actual calculated rate, 17, 18
- construction, cross folds, 19
- individual performance, 19
- NENN algorithm, 9–10
- performance and regression plots, 19–24
- training and testing data sets, 19
- validation, 19
- genetic and biochemical pathways, 102
- HIV infection, 160
- hysteresis, 101–102

- Mathematical modelling (*cont.*)
- incidence data
    - computations, 25
    - model 1 and 2, 25–27
    - model 3 and 4, 1, 26, 27
    - parameters, logistic function, 25, 26
    - simulations, 25
  - infection elimination and reduction, 71
  - interactions and transitions, 101
  - MprAB system, 101
  - optimization approach, 171
  - post-translational modules, 101
  - SEIR (*see* SEIR epidemic and percolation models)
  - TB, 64
  - transcriptional module, 101
- MDR-TB. *See* Multidrug-resistant tuberculosis (MDR-TB)
- MLSA. *See* Entire gene sequences (MLSA)
- Modular networks with community
- direction, 49
  - effects, 50
  - epidemic threshold, 49
  - hierarchical structure 50
  - interactions, 49
  - SEIR epidemics with households, 49
  - spreading infection, 49
- Molecular epidemiology
- description, 114–115
  - genotype/fingerprint, strains, 115
  - microbial-host evolution, 115
  - RD, 117
  - RFLP, 115–116
  - SNP and MLSA/MLST, 118
  - spatial and temporal mode, 115
  - spoligotyping, 116
  - transmission dynamics, tuberculosis, 115
  - tuberculosis control (*see* Tuberculosis (TB) control)
  - Tuberculosis Control Programs, 115
  - VNTR, 116–117
  - WGS, 119–120
- M. tuberculosis*. *See* *Mycobacterium tuberculosis*
- M. tuberculosis* complex (MTBC)
- allopatric human populations, 129
  - ancient and modern lineages, 128
  - clonal population structure, 126
  - Euro-American and Indo-Oceanic lineage, 129
  - genetically monomorphic pathogens, 114
  - genetic variation, 113, 125–126
  - genome sequences, 126
- genotype-phenotype associations, 128
- global human strains, 125
- global population and distribution, 126–127
- human-adapted lineages, 126
- human monocyte-derived macrophages, 128
- in-depth analyses, 126
- lineage 2, East Asian, “Beijing/W” strain, 130–131
- M. africanum*, 131–132
- M. bovis*, 113–114
- mycobacterial lineages, 126
- phylogeographic associations, 129
- pro-inflammatory cytokines, 128
- robust phylogeny, 126
- slowly growing mycobacteria, 113
- species concept, 128
- Multidrug-resistant tuberculosis (MDR-TB)
- drug-resistant tuberculosis, 134
  - fluoroquinolone, 132
  - lineage 2 strains, 130
  - XDR, 132
- Mycobacterial pathogenicity, TCS
- macrophage infection model, 90
  - mycobacterial survival, 90
  - mycobacterial virulence (*see* Mycobacterial virulence)
- NarLS, PdtaRS and TcrA, 95
- Rv0600c responses, 95
- Mycobacterial virulence
- DevR-DosT (DosRT) systems, 91–92
  - DevRS (DosRS), 91–92
  - infection, host cell, 91
  - KdpDE system, 94
  - MprAB system, 92–93
  - PhoPR system, 92
  - SenX3-RegX3, 93–94
  - TcrXY system, 94
  - TrcRS system, 94
- Mycobacterium tuberculosis*
- acid-fast bacilli and chest X-ray, 79
  - aerosol droplets, 79
  - alveolar macrophages, 112
  - asymptomatic latent infection., 60
  - BCG vaccine, 61
  - capacity, mycobacterium, 112–113
  - causes, 59
  - chemotherapy, 79–80
  - description, 79
  - diagnosis, 60, 80–81
  - distribution, ancient Lineage 5, 124
  - drug-resistant tuberculosis, 80
  - elimination, 61

- extrapulmonary TB disease, 112  
“fungus-like” bacterium, 80  
genetic distances, strains, 123, 124  
genotyping approach, 120–121  
granuloma, 112  
HIV infection, 113  
host–pathogen coevolution, 138  
in vivo and in vitro models, 113  
lungs, 60  
MDR and XDR, 111  
modern human migration patterns, 124  
molecular epidemiology (*see* Molecular epidemiology)  
MTBC (*see* *M. tuberculosis* complex (MTBC))  
mycolic acids, 80  
National Tuberculosis Programs, 111  
phago-lysosome, 80  
PPD, 79  
primary infection, 60  
signalling systems (*see* Signalling systems, *M. tuberculosis*)  
TB infection, 112  
T cells, 112  
TCS, 79  
transmission, 63, 112  
tubercle bacilli, 112  
virulence and pathogenicity, 81–82  
virulence, vaccine and drug development, 111  
WGS, 82–83
- N**  
National Tuberculosis Programs, 111
- P**  
Pacific Northwest outbreak, *C. gattii*  
respiratory symptoms and pulmonary cryptococcoma, 187  
strain, California, 192  
transmission, US Pacific Northwest, 191  
Vancouver Island and British Columbia, 190–191
- Percolation  
Bernoulli (*see* Bernoulli percolation model)  
epidemic (*see* SEIR epidemic and percolation models)  
great gerbils, Kazakhstan, 40–41  
household distribution, 52  
infectious periods and transmission rate, 53  
intervention, 54
- long-range, 53–54  
quantitative evaluations, 54  
and random networks (*see* Random networks and percolation)  
Reed–Frost epidemics, 38–39, 53  
scale-free networks, 50–51, 53  
short-range Bernoulli percolation, 53  
transmission probabilities, 52
- PPD. *See* Purified protein derivative (PPD)
- Prevention and control  
cryptococcal disease  
avoidance, 194  
geographical influence, transmission, 193  
speculative measures, 194
- HIV, 147, 149, 170
- H5N1  
bioinformatics, 235  
conventional RT-PCR, 234  
detection, fluorescent DNA barcode-based immunoassay, 234  
measures, 246–247  
quantification, 234  
vaccines and antiviral, 234–235
- TB, 61
- Purified protein derivative (PPD), 79
- R**  
Random networks and percolation  
graphs  
Bernoulli, 43  
evolution, 43  
Kermack–McKendrick model, 42  
mass-action, 42  
SEIR, 43–44  
standard stochastic, 43  
threshold, 44  
modular/community structure, 49–50  
scale-free, 45  
small-world, 45–46  
spatial and long-range, 47–48
- RD. *See* Regions of difference (RD)
- Reed–Frost model, 38–39
- Regions of difference (RD), 117
- Response regulator  
C-terminal DNA-binding domain, 86  
extracellular cues, 103  
macrophage infection model, 89, 90  
orphan sensor kinase, 95  
phosphorylated protein, 84  
sensor kinase protein, 85
- Restriction fragment length polymorphism (RFLP)

- Restriction fragment length polymorphism (RFLP) (cont.)  
 IS6110 element, 115–116, 130  
 mycobacterium, 116  
 Southern blot technique, 115
- RFLP. *See* Restriction fragment length polymorphism (RFLP)
- S**
- Scale-free networks  
 characteristics, 55  
 with community structure, 49  
 consequences, 46  
 epidemics, 46  
 hierarchical modular spatial random network, 54  
 non-scale-free, 46  
 random graphs, 46  
 SEIR, 47  
 spatial hierarchical, 50–52  
 vertices distribution, 46
- SEIR epidemic and percolation models  
 analysis (*see* Percolation)  
 basic reproduction number, 39  
 Bernoulli, 35–38  
 description, 31  
 deterministic, 32–33  
 directions, 31, 32  
 features, 55  
 fixed infectious periods, 39–40  
 general stochastic, 33–34  
 graphical representation, 41  
 great gerbils, Kazakhstan, 40–41  
 infectivity and susceptibility, 41–42  
 mathematical modelling, 31  
 out and incomponents, 41  
 process, 31–32  
 and random networks (*see* Random networks and percolation)  
 random variables, 55  
 Reed–Frost, 38–39  
 selection, directed edges, 41  
 spatial features, 31  
 time evolution, 33–34
- SEIT model. *See* Susceptible—Exposed—Infected—Treated (SEIT) model
- Serine-threonine protein kinases (STPKs), 84, 103
- Signalling systems, *M. tuberculosis*  
 canonical signalling pathway, 84  
*Caulobacter crescentus* and *E. coli*, 96
- DevR protein and DevRS regulon, 96  
 “energy-independent” phosphotransfer mechanism, 84  
 genome sequence, 83  
 response regulator protein, 84–85  
 STPK, 83  
 STRING program, 95  
 transcriptional modulation, 95–96  
 two-component signal transduction systems  
 (*see* Two-component signal transduction systems (TCS))
- Single nucleotide polymorphisms (SNPs)  
 bacterial strains, 118  
 Beijing genotype family, 131  
 genome sequences, 113, 126  
 strain lineages, 121  
 strains, tuberculosis, 118
- Small-world networks, 45–46
- SNPs. *See* Single nucleotide polymorphisms (SNPs)
- Spatial mode of TB spread, 114–115
- Spatial networks  
 density, 40  
 features, 31  
 and long-range percolation, 47–49  
 networks and long-range percolation, 47–49  
 scale-free networks, 46–47, 50–52  
 structure, 36
- Spoligotyping, 116
- Statistical transmission model, H5N1  
 assumptions, 241  
 methodology, 242–244  
 outbreak, 240  
 outcomes, numerical simulation, 244–247  
 parameters, 241  
 transmission, 240–241
- Stochastic SEIR epidemic model, 33–34
- STPKs. *See* Serine-threonine protein kinases (STPKs)
- Support vector machine (SVM), 5
- Susceptible—Exposed—Infected—Treated (SEIT) model, 64–67, 73
- SVM. *See* Support vector machine (SVM)
- Swine flu. *See* Swine-origin influenza A (H1N1) virus
- Swine-origin influenza A (H1N1) virus  
 AD-plot, 218–219  
 antigenicity, 219  
 DEPS analysis, 219–220  
 evolution, 218–219

**T**

- Totally drug-resistant strains of *M. tuberculosis* (TDR-TB), 132
- Tuberculosis (TB) control
- archeological evidence, 125
  - drug resistance (*see* Drug resistance)
  - genotyping tools, 121
  - host-pathogen coevolution and population structure, 122
  - human and pathogen evolution, 122
  - hunter-gatherer times, human populations, 122
  - M. africanum*, 125
  - M. bovis*, 125
  - M. canettii*, 122
  - “*Mycobacterium prototuberculosis*”, 122, 123
  - phylogeographic distribution, MTBC, 123
  - population-based, 121
  - transmission, 121
- Tuberculosis (TB), Nigeria
- active and inactive, 60
  - BCG vaccine, 61
  - control and prevention, 61
  - diagnosis, 60–61
- DOTS (*see* Direct Observation Therapy Strategy (DOTS))
- infection sites, 59–60
  - late 1980s, 59
  - long term treatment, 61
- M. tuberculosis* (*see* *Mycobacterium tuberculosis*)
- outbreaks, 59
  - side effects, 61
  - streptomycin, 61
  - sub-Saharan Africa, 59
  - susceptibility, 60
  - therapy, 59
  - transmission, 59
  - tubercle bacilli*, 60
- Two-component signal transduction systems (TCS)
- drug target (*see* Drug target)
  - extracellular cues, 103
  - in vivo interaction, TCS proteins, 103
  - mathematical modelling (*see* Mathematical modelling)
  - mycobacterial pathogenicity (*see* Mycobacterial pathogenicity, TCS)
- N-terminal sensory domain, 85–86
- pathogenesis, 103

- phosphorylation reaction, anti-microbial drug, 85
- prokaryotic organisms, 85
- response regulator proteins, 86, 87, 90
- sensor kinase proteins, 86–89
- signalling in mycobacteria, 103–104

Two-component system, 41

**V**

- Variable number tandem repeats (VNTR)
- agarose gel-based electrophoresis method, 117
  - global epidemiological database, 117
  - PCR amplification, 117
  - strain lineages, 121
- Viral populations
- allele dynamics plots (AD-plots), 206
  - DEPS, 221
  - influenza (*see* Influenza A viruses)
  - phylogenetic techniques, 205
  - phylogenetic analysis, 205
  - visualization techniques, 221
- VNTR. *See* Variable number tandem repeats (VNTR)

**W**

- Watt–Strogatz small-world network, 45–46
- Whole genome sequencing (WGS)
- E. coli*, 82–83
  - heterogeneity, 120
  - “H37Ra-specific”, 83
  - hypermutation models, 119
  - “identical fingerprint equal same strain”, 119
- M. avium*, *M. bovis* and *M. marinum*, 83
- MTBC, 119
- pathogenicity, transmission and drug resistance, 83
- rifampin-resistant strains, 137
- tubercle bacilli, 82
- World Health Organization (WHO) on TB
- burden, high-burden countries (HBC), 62
  - declaration, global emergency, 61
  - implementation and surveillance, 63
  - program, 61–62

**X**

- XDR. *See* Extensive-drug-resistant (XDR)