Interface Focus

rsfs.royalsocietypublishing.org

Introduction

Cite this article: Morozov A. 2013 Modelling biological evolution: recent progress, current challenges and future direction. Interface Focus 3: 20130054.
http://dx.doi.org/10.1098/rsfs.2013.0054

One contribution of 11 to a Theme Issue 'Modelling biological evolution: recent progress, current challenges and future direction'.

Subject Areas:
bioinformatics

Keywords:
decision-making, quantitative genetics, biological information, biodiversity, adaptive dynamics, optimal control

Author for correspondence:
Andrew Morozov
e-mail: am379@leicester.ac.uk

Modelling biological evolution: recent progress, current challenges and future direction

Andrew Morozov

Department of Mathematics, University of Leicester, University Road, Leicester LE1 7RH, UK

Mathematical modelling is widely recognized as a powerful and convenient theoretical tool for investigating various aspects of biological evolution and explaining the existing genetic complexity of the real world. It is increasingly apparent that understanding the key mechanisms involved in the processes of species biodiversity, natural selection and inheritance, patterns of animal behaviour and coevolution of species in complex ecological systems is simply impossible by means of laboratory experiments and field observations alone. Mathematical models are so important because they provide wide-ranging exploration of the problem without a need for experiments with biological systems—which are usually expensive, often require long time and can be potentially dangerous. However, as the number of theoretical works on modelling biological evolution is constantly accelerating each year as different mathematical frameworks and various aspects of evolutionary problems are considered, it is often hard to avoid getting lost in such an immense flux of publications. The aim of this issue of *Interface Focus* is to provide a useful guide to important recent findings in some key areas in modelling biological evolution, to refine the existing challenges and to outline possible future directions. In particular, the following topics are addressed here by world-leading experts in the modelling of evolution: (i) the origins of biodiversity observed in ecosystems and communities; (ii) evolution of decision-making by animals and the optimal strategy of populations; (iii) links between evolutionary and ecological processes across different time scales; (iv) quantification of biological information in evolutionary models; and (v) linking theoretical models with empirical data. Most of the works presented here are in fact contributed papers from the international conference 'Modelling Biological Evolution' (MBE 2013), which took place in Leicester, UK, in May 2013 and brought together theoreticians and empirical evolutionary biologists with the main aim of creating debates and productive discussions between them. Finally, we should emphasize that the individual papers in this issue are not limited to only one of the topics mentioned above, but often lie at the interface of them.

Three of the contributions within this issue focus on revealing the mechanisms of high biodiversity and species creation [1–3]. Yamaguchi & Iwasa [1] consider a particularly important case of speciation called allopatric speciation, which is caused by accumulating genetic divergence owing to geographical isolation of populations (e.g. located on different islands). Allopatric speciation has always been considered as a major process in species creation [4] (since the revolutionary work of Charles Darwin) and various models have already been developed [5,6]. The main particularity of the work by Yamaguchi and Iwasa is that the authors consider the more realistic case where the populations are not completely isolated from each other but are related by rare migration. By using different modelling approaches for different numbers of isolated populations, it is shown that species creation can be maximized at an intermediate rate of migration between the sites/islands, a conclusion which is found to be in good agreement with some available empirical data [7].

Royal Society Publishing
Informing the science of the future

Beckett & Williams [2] consider another mechanism of diversification and speciation as a result of coevolution in multi-species host–parasite systems. As a study case, they investigate evolution of biodiversity in a community of bacteria and phages in the ocean. Mathematically, their model set is a simple ODE-based multi-strain chemostat model, where parasites compete over their hosts. The strength of infection by the phage depends on genetic matching with the host bacteria via 'the lock-and-key' mechanism. In the model, negative frequency-dependent selection—maintained by 'kill-the-winner' dynamics—eventually results in diversification of both bacteria and phages. The authors consider the emerging interaction network, that is, the map of 'who infects whom', analyse its multi-scale structure and finally compare the theoretical results to empirical datasets recently obtained from samples in the Atlantic Ocean [8]. They argue that the 'nested-modular' multi-scale interaction network obtained from the real data [8] can be easily reproduced in their simple model, and thus prove the coevolutionary origin of the high diversity observed in natural bacteria–phage communities in the ocean.

In their work, Best & Hoyle [3] also investigate the emergence of biodiversity in a host–parasite system as a result of evolution. Firstly, the authors provide a detailed review of processes which can potentially cause evolutionary branching of parasite virulence in host–parasite interactions: superinfection, density-dependent host mortality, preferential predation on infected host, etc. They then suggest an interesting novel mechanism by which disruptive evolution of the parasite virulence arises as a consequence of a limited immune range of the host; in particular, they relax the key assumption made in the early models that immune hosts are perfectly protected from all future parasite strains. Using a standard adaptive dynamics approach [9], they show that in the case where the immune range of the host is not very large, there can be a series of branching events in the model creating new viral strains. As their model is quite general, the authors argue that the reported mechanism may be a major explanatory factor for the high diversity of parasites in vertebrate host populations observed in nature.

The detailed review of Alizon [10] continues the investigation of the evolution of virulence in host–parasite systems as presented in [3], by considering the situation with multiple infections (a host can be simultaneously infected by different parasites), which is typically the case in both wildlife and humans [11,12]. Evolution of virulence in complex systems with multiple infections is currently widely studied, because in such systems selective pressure can act in a completely different way compared with a single infection case. For instance, Alizon demonstrates that superinfection can result in evolutionary branching of virulence [10]. In the review, various models of co-infection are discussed, together with their underlying biological assumptions. It is shown as well that, surprisingly, some of these models are not suitable for use in evolutionary epidemiology as they show a frequency-dependent advantage for neutral mutants, which is clearly a model artefact. The author suggests that the recently developed Price equation framework [13] provides a straightforward way of incorporating high genetic diversity of parasites into the well-developed framework of 'classical' epidemiological models.

When studying population dynamics of interacting species in nature, it is often crucial to have a clear understanding of the various strategies that those populations use to optimize their fitness, thus enhancing their persistence and survival rate. Classical examples of such situations are the various strategies of vertical migration used by herbivorous zooplankton in the water column [14] and the complex strategies of switching between an active state and an energy-saving physiological state (the diapause) taken by some invertebrates [15]. The standard theoretical approach to determine the best strategy for population persistence is implementation of game theory and optimal control theory: the optimal strategy is determined by maximizing a certain functional related to fitness. Staňková *et al.* [16] use these approaches (based on the Hamilton–Jacobi–Bellman equations of dynamic programming) to find the optimal strategy for allocating time between the active state and the diapause state in a generic predator–prey model, together with a case study of the interaction between predator and prey mites on fruit trees. The authors construct a set of models of varying complexity and compare the predicted strategies with empirical observations. They conclude that the model which does not take into account the energy dynamics of the organisms yields contradictory predictions, because it shows multiple entering/exiting the diapause state by the mites, which is not observed in reality [16].

In her study, Conradt [17] addresses another biologically relevant scenario in which animals must choose an optimal strategy. In this case, the decision is shared by all animals in the group, i.e. there is collective decision-making. In particular, the author considers the situation where the collective decision of social animals within a group involves elements of conflict: the ultimate decision can clash with the preferred decision of some individuals. Recently, the same author showed that conflicting interests of individuals eventually make the collective decision more accurate [18], and this paper addresses the generality of this finding. By constructing a fairly simple model and computing the probabilities of group decisions, the author concludes that for animals without personal preferences, the probability of making a correct collective decision is higher. Moreover, Conradt [17] argues that in the situation where a decision is crucial for the survival of the whole group, animals should 'avoid' having their preferences for false positives compared to those for false negatives. Finally, the model shows that individual knowledge of the average success rate of a decision option can be quite misleading because it negatively influences the decision accuracy.

In earlier models of biological evolution, it was generally assumed that evolutionary processes take place on a much slower time scale than that of population dynamics. However, there is a growing body of evidence that evolutionary processes can occur on a similar time scale as population dynamics and that we need to take this fact into account in our models [19,20]. In their contribution, Fussmann & Gonzalez [21] investigate how rapid evolution within a community of competing consumers can create a community evolutionary rescue, i.e. the scenario in which two or more species exhibit rapid evolutionary change as a response to environmental change. This will ensure the recovery of the species, promoting species biodiversity and preventing eventual extinction. The authors consider a simple chemostat model in which two consumers are competing for the same resource and use the methods of quantitative genetics to describe the phenotypic variation of a life-history trait. They also demonstrate that a community evolutionary rescue may occur both simultaneous with or subsequent to the change in the environment, and that alteration in the population cycles within the community can be a possible indicator of the undergoing 'invisible' evolutionary rescue.

When looking for a suitable framework to describe an evolutionary process in a particular biological system, one should keep in mind the major limitations of the framework chosen, including the tacit simplifying assumptions as well as the time-scale limitations. In particular, the adaptive dynamics framework, which is rather popular in the evolutionary modelling literature [9], is often criticized for its key assumptions that mutations are rare and take place via small steps [22]. Moreover, as these two assumptions are somewhat mutually exclusive, adaptive dynamics is sometimes considered to have intrinsic limitations in its ability to describe evolution in real biological systems. In their work, however, Metz & de Kovel [23] argue that this widespread opinion is not necessarily correct and that, for small steps, the separation between the population and the mutational time scales should not necessarily be an issue. In particular, the authors try to extend the formalism of adaptive dynamics from the case of well-mixed clonal populations to more biologically relevant scenarios of Mendelian diploid and haplo-diploid populations. Metz and de Kovel conclude that the canonical equation formalism (after some technical improvement, for details see [23]) applied to Mendelian diploids can provide an even better approximation of the prevalent trait vector change over evolutionary time compared with the classical clonal populations. This becomes possible owing to gene substitutions taking place in parallel, thus impeding a slowing down of evolutionary processes owing to clonal interference.

Despite important recent progress in understanding evolutionary processes using mathematical models, several fundamental challenges still remain to be addressed. Several of them are highlighted in this Theme Issue. For instance, it is well known that biological information plays a pivotal role in evolutionary processes as well as in population dynamics—in fact, the link between ecology and evolution is based on information transfer from one to the other. However, the existing information theory has been developed for the case of objective and discrete information (e.g. using the Shanon entropy). On the other hand, biologically relevant information is often very subjective and of analogue signal type [24], and so there is an urgent need for extension of information theory to be able to cope with real biological objects. This is the main theme of the paper by van Baalen [25] who emphasizes the shortcomings of current information theory for describing ecological and evolutionary processes, and discusses the basic concepts of the emerging theory of biological information. The author argues that a possible starting point could be the use of entropy measures of conditional probabilities (e.g. implementing the so-called Kullback–Leibler divergence), which would help to link the acquisition of information with an eventual increase in fitness. Van Baalen suggests that biological information should be measured in terms of fitness, and in this case evolutionary theory would clearly play a pivotal role in defining such information. The author also argues that when evaluating information based on fitness, we need to be quite careful regarding whose fitness we are considering: that of individuals, clones and symbioses, etc., as different choices can result in completely different information measures.

Another key challenge in evolutionary biology is in linking sophisticated theories to empirical data. This is also a central issue in theoretical ecology [26], but evolutionary processes often take place on larger time scales which seriously impede experimental work and field observations. In their insightful review, Duputié & Massol [27] address the connection between theory and data by considering the evolution of dispersal—a fundamental question in evolutionary ecology [28]. Duputié & Massol [27] first provide a thorough review and classification of the existing types of models for dispersal evolution (e.g. based on selective evolutionary pressures). The authors then discuss experimental testing of various theories on the evolution of dispersal, emphasizing the role of spatial heterogeneity as well as time and spatial scales. Throughout the review, Duputié & Massol [27] highlight the important fact that the field/experimental data often provide information about the observed patterns, but not about the underlying processes. Thus, the data often give us 'information about correlation but not causation' [27]. Finally, some concrete recipes on fulfilling testable predictions of models of evolution of dispersals are provided in the review. Overall, the importance of the link between the theoretical approach and empirical work was stressed in many of the contributions to this Theme Issue, and many included concrete case study-based examples. This emphasizes the need for a more constructive dialogue between theoreticians and experimentalists to enlarge the interdisciplinary collaboration between both groups.

Finally, an important general theme constantly emphasized in this special issue is the trade-off between model complexity and model tractability/generality in describing evolution in real systems. It is well known that all biological models are strictly speaking incorrect because they only can provide some approximation of complex real systems, but on the other hand making the model unnecessarily too complex and including too many details (even if they are justified by some empirical data) can impede the understanding of generic properties of the system. Thus, finding the right balance between complexity and generality becomes a crucial issue in model construction. A good example demonstrating the existence of such a trade-off is provided in the paper by Yamaguchi & Iwasa [1], where the authors compared a family of models of increasing complexity to describe allopatric speciation, starting from the simple ODE approximation and ending with more complex individual-based modelling. The authors progressively show where each approximation will be sufficient and where a more complex model is needed to be used to correctly describe the system, at least qualitatively. Beckett & Williams [2] elegantly demonstrate that a very simple model of bacteria–phage interaction can result in patterns which are in very good agreement with those observed in the real ocean. On the contrary, Staňková *et al.* [16] show that simplifying a description of a predator–prey system by neglecting the energy level of the individuals within populations can result in erroneous conclusions regarding the optimal strategy of organisms within a population. Thus, their insightful study emphasizes that we need to be very careful about oversimplifying biological systems when constructing our models.

Acknowledgements. I thank the contributors, reviewers and editors for their collaboration in producing this Theme Issue of *Interface Focus*. Professor S. Petrovskii provided several important recommendations which helped me a lot in organizing this issue. Also I am grateful to all the participants and speakers of MBE 2013 (Leicester, UK) for contributing to the success of the conference with plenty of interesting and stimulating discussions, which sometimes even developed into heated debates.

References

1. Yamaguchi R, Iwasa Y. 2013 First passage time to allopatric speciation. *Interface Focus* **3**, 20130026. (doi:10.1098/rsfs.2013.0026)
2. Beckett ST, Williams HTP. 2013 Coevolutionary diversification creates nested-modular structure in phage–bacteria interaction networks. *Interface Focus* **3**, 20130033. (doi:10.1098/rsfs.2013.0033)
3. Best A, Hoyle A. 2013 A limited host immune range facilitates the creation and maintenance of diversity in parasite virulence. *Interface Focus* **3**, 20130024. (doi:10.1098/rsfs.2013.0024)
4. Rosenzweig ML. 1995 *Species diversity in space and time*. Cambridge, UK: Cambridge University Press.
5. Nei M, Maruyama T, Wu CI. 1983 Models of evolution of reproductive isolation. *Genetics* **103**, 557–579.
6. Orr HA, Orr LH. 1996 Waiting for speciation: the effect of population subdivision on the waiting time to speciation. *Evolution* **50**, 1742–1749. (doi:10.2307/2410732)
7. Avise JC. 2000 *Phylogeography*. Cambridge, MA: Harvard University Press.
8. Flores CO, Valverde S, Weitz JS. 2013 Multi-scale structure and geographic drivers of cross-infection within marine bacteria and phages. *ISME J.* **7**, 520–532. (doi:10.1038/ismej.2012.135)
9. Metz JAJ, Geritz SAH, Mesźena G, Jacobs FJA, Van Heerwaarden JS. 1996 Adaptive dynamics: a geometrical study of the consequences of nearly faithful reproduction. In *Stochastic and spatial structures of dynamical systems* (eds SJ van Strein, SM Verduyn Lunel), pp. 183–231. Amsterdam, The Netherlands: North-Holland.
10. Alizon S. 2013 Co-infection and super-infection models in evolutionary epidemiology. *Interface Focus* **3**, 20130031. (doi:10.1098/rsfs.2013.0031)
11. Telfer S, Lambin X, Birtles R, Beldomenico P, Burthe S, Paterson S, Begon M. 2010 Species interactions in a parasite community drive infection risk in a wildlife population. *Science* **330**, 243–246. (doi:10.1126/science.1190333)
12. Brogden KA, Guthmiller JM, Taylor CE. 2005 Human polymicrobial infections. *Lancet* **365**, 253–255. (doi:10.1016/S0140-6736(05)17745-9)
13. Day T, Gandon S. 2007 Applying population-genetic models in theoretical evolutionary epidemiology. *Ecol. Lett.* **10**, 876–888. (doi:10.1111/j.1461-0248.2007.01091.x)
14. Morozov AY. 2012 Incorporating complex foraging of zooplankton in models: role of micro and mesoscale processes in macroscale patterns. In *Dispersal, individual movement and spatial ecology: a mathematical perspective* (eds M Lewis, P Maini, S Petrovskii), pp. 1–10. New York, NY: Springer.
15. Veerman A. 1985 Diapause. In *Spider mites: their biology, natural enemies and control* (eds W Helle, MW Sabelis). World Crop Pests, vol. 1A, pp. 279–316. Amsterdam, The Netherlands: Elsevier Science Publishers.
16. Staňková K, Abate A, Sabelis MW, Buša J, You L. 2013 Joining or opting out of a Lotka–Volterra game between predators and prey: does the best strategy depend on modeling energy lost and gained? *Interface Focus* **3**, 20130034. (doi:10.1098/rsfs.2013.0034)
17. Conradt L. 2013 Collective animal decisions: preference conflict and decision accuracy. *Interface Focus* **3**, 20130029. (doi:10.1098/rsfs.2013.0029)
18. Conradt L. 2012 Models in animal collective decision-making: information uncertainty and conflicting preferences. *Interface Focus* **2**, 226–240. (doi:10.1098/rsfs.2011.0090)
19. Matthews B *et al*. 2011 Toward an integration of evolutionary biology and ecosystem science. *Ecol. Lett.* **14**, 690–701. (doi:10.1111/j.1461-0248.2011.01627.x)
20. Yoshida T, Jones LE, Ellner SP, Fussmann GF, Hairston Jr NG. 2003 Rapid evolution drives ecological dynamics in a predator–prey system. *Nature* **424**, 303–306. (doi:10.1038/nature01767)
21. Fussmann GF, Gonzalez A. 2013 Evolutionary rescue can maintain an oscillating community undergoing environmental change. *Interface Focus* **3**, 20130036. (doi:10.1098/rsfs.2013.0036)
22. Barton NH, Polechova J. 2005 The limitations of adaptive dynamics as a model of evolution. *J. Evol. Biol.* **19**, 1186–1190. (doi:10.1111/j.1420-9101.2005.00943.x)
23. Metz JAJ, de Kovel CGF. 2013 The canonical equation of adaptive dynamics for Mendelian diploids and haplo-diploids. *Interface Focus* **3**, 20130025. (doi:10.1098/rsfs.2013.0025)
24. Donaldson-Matasci MC, Bergstrom CT, Lachmann M. 2010 The fitness value of information. *Oikos* **119**, 219–230. (doi:10.1111/j.1600-0706.2009.17781.x)
25. van Baalen M. 2013 Biological information: why we need a good measure and the challenges ahead. *Interface Focus* **3**, 20130030. (doi:10.1098/rsfs.2013.0030)
26. Codling E, Dumbrell AJ. 2012 Mathematical and theoretical ecology: linking models with ecological processes. *Interface Focus* **6**, 144–149. (doi:10.1098/rsfs.2012.0008)
27. Duputié A, Massol F. 2013 An empiricist's guide to theoretical predictions on the evolution of dispersal. *Interface Focus* **3**, 20130028. (doi:10.1098/rsfs.2013.0028)
28. Levin SA, Muller-Landau HC, Nathan R, Chave J. 2003 The ecology and evolution of seed dispersal: a theoretical perspective. *Annu. Rev. Ecol. Syst.* **34**, 575–604. (doi:10.1146/annurev.ecolsys.34.011802.132428)