



HAL
open science

Special Issue: Mathematical and Computational Evolutionary Biology-2015

David Bryant, Olivier Gascuel

► **To cite this version:**

David Bryant, Olivier Gascuel. Special Issue: Mathematical and Computational Evolutionary Biology-2015. *Systematic Biology*, 2017, 66 (1), pp.1 - 2. 10.1093/sysbio/syw111 . pasteur-02405199

HAL Id: pasteur-02405199

<https://pasteur.hal.science/pasteur-02405199v1>

Submitted on 11 Dec 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Editorial

Syst. Biol. 66(1):1–2, 2017

© The Author(s) 2017. Published by Oxford University Press, on behalf of the Society of Systematic Biologists.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com
DOI:10.1093/sysbio/syw111

Special Issue: Mathematical and Computational Evolutionary Biology—2015

DAVID BRYANT¹ AND OLIVIER GASCUEL^{2,3,*}

¹Department of Mathematics and Statistics, University of Otago, PO Box 56, Dunedin 9054 - New Zealand; ²Unité Bioinformatique Evolutive C3BI, USR 3756 Institut Pasteur - CNRS Paris, France and ³Institut de Biologie Computationnelle, LIRMM, UMR 5506, Université de Montpellier - CNRS Montpellier, France

*Correspondence to be sent to: Institut de Biologie Computationnelle, LIRMM, UMR 5506, Université de Montpellier - CNRS, Montpellier, France; E-mail: olivier.gascuel@pasteur.fr.

This special issue of *Systematic Biology* contains review articles contributed by keynote speakers at the “Mathematical and Computational Evolutionary Biology” conference (MCEB; see web sites below), held in 2015. This is the seventh conference in this series, the first being held in Paris in 2003 with keynote addresses from such notables as Joe Felsenstein, Ziheng Yang, and Walter Fitch. Oxford University Press published two books based on the 2003 and 2005 Paris meetings (Gascuel 2005; Gascuel and Steel 2007).

Mathematical and computational tools have become essential for evolutionary studies. The rise of sound statistical and combinatorial approaches in evolutionary and genome biology has offered considerable improvements beyond the original *ad hoc* methods, providing new methods and algorithms to handle ever-growing data sets. However, such advances require an increasingly sophisticated mathematical treatment of the problems at hand, as well as faster algorithms and computers in order to answer important biological questions.

The aim of MCEB is to contribute to the development of this methodology by bringing together researchers from various disciplines: mathematics, computer science, phylogenetics, population genetics, and genomics. Interdisciplinarity is hard to do right, but we think we have struck on a formula which works well: beautiful locations; plenty of time for discussion and collaboration; full length (75 min) keynote lectures where speakers were given strict instructions to devote a lot of time to background and review.

The 2015 conference was held on Porquerolles, a small island just south of Marseilles. The focus of the conference was “New Data, New Questions, New Methods.” New generation sequencing techniques have multiplied not just the amount, but also the types of genetic data produced (e.g., SNPs, RNA-seq, RAD-seq, ancient DNA, Hi-C, single-cell sequencing, etc), giving rise to new questions, and new methodologies to answer them. These methodologies are often cross-disciplinary, with applications to diverse research topics such as epigenetics, gene expression, genome evolution, cancer,

development, and epidemiology. General concepts, models, methods, and algorithms were also presented and discussed, just as during the previous conference editions. Along these lines, we had short talks ranging from evolution of root-knot nematodes to inference of population sizes, from codon substitution models to viral evolution and gene order variation.

This special issue contains five review articles, all based on keynote lectures delivered at the conference. *Systematic Biology*, with its strong history in attracting, handling, and disseminating mathematically inclined articles in these fields, is the perfect choice for the publication of such papers.

The review “Evolutionary Patterns and Processes: Lessons from Ancient DNA” by Ludovic Orlando’s team, is emblematic of our yearly theme on new data. It shows that thanks to recent development of high-throughput sequencing and sophisticated molecular techniques, ancient DNA research has now come of age, merging together approaches in phylogenomics, population genomics, epigenomics, and metagenomics. Ancient DNA provides direct access to the most important dimension in evolution—time, opening a wealth of fundamental evolutionary processes to be studied at unprecedented resolution.

A keynote lecture by Asger Hobolth formed the basis of the paper “Statistical Inference in the Wright–Fisher Model Using Allele Frequency Data” by Paula Tataru and co-authors. This is a concise and comprehensive review of the many variants, extensions, and approximations of the Wright–Fisher model. The model forms the basis of much of population genetics, and the authors have brought together a substantial amount of literature into one framework. The authors examine each of the standard extensions to the model, summarizing what we know about the distribution of allele frequencies, approximations based on diffusions, and moment-based techniques. The review includes new comparative studies of the accuracy and computational efficiency of the different approaches.

Philippe Lemey gave a lecture reporting on the remarkable recent progress being made in the area

of phylodynamics, reviewed in “Emerging Concepts of Data Integration in Pathogen Phylodynamics” by Guy Baele and co-authors. The paper shows how these tools permit reconstruction of the evolutionary and ecological dynamics of important pathogens. A key development in this area has been the design of techniques to integrate sequence information with a wide range of alternative data sources, often within a Bayesian framework. The analyses described incorporate geographic data, evolutionary covariates, trait data, and ancestral samples. The power of these new approaches is demonstrated using a wide and diverse range of examples and case studies, from HIV-1 and influenza to bat rabies.

A keynote by Jukka Corander formed the basis of the review article “Fundamentals and Recent Developments in Approximate Bayesian Computation” by Jarno Lintusaari and co-authors, which deals with the last side of our theme: new methods. The current models developed in phylogenetics and evolutionary biology are often so complex that the likelihood function to compute the “probability” of the data being analyzed cannot be computed. We then refer to simulations to obtain models producing data that resemble the observed data. A number of methods to conduct such “Approximate Bayesian Computations” or ABC have been developed in the recent years, with successful applications in population genetics and evolutionary biology. This article explains the fundamentals of ABC, reviews the classical algorithms, and highlights recent developments.

In “The Shape of Phylogenetic Treespace,” Katherine St. John reviews significant new developments in our mathematical understanding of phylogenetic trees and the mathematical space they live in. Trees are awkward objects to work with mathematically or statistically, and any major new developments in the area of phylogenetic methods are going to rely on improved understanding of treespace. This review provides an excellent introduction for anyone, mathematician or biologist, wanting to understand this field. All of the main metrics (distance measures) on trees are reviewed, as are the main selection criteria and search algorithms.

We hope that these review papers, containing a number of references and providing detailed descriptions of broad fields, will be useful to a large number of *Systematic Biology* readers and beyond, and that they will attract new, mathematically inclined researchers in these exciting domains. We also hope that this form of publication (review papers online), which is relatively new for *Systematic Biology* [but see Gascuel and Stadler 2015, this special issue following MCEB 2013 (focus on health and medicine) contains three review articles totalizing >20,000 downloads to date], will spread and be continued in the coming years.

CONFERENCE WEB SITES

2003: <http://www.lirmm.fr/~guindon/ihp/>
 2005: <http://www.lirmm.fr/MEP05/>
 2008: <http://www.lirmm.fr/MIEP08/>
 2012: <http://www.lirmm.fr/mceb2012/>
 2013: <http://www.lirmm.fr/mceb2013/>
 2014: <http://www.lirmm.fr/mceb2014/>
 2015: <http://www.lirmm.fr/mceb2015/>
 2016: <http://www.lirmm.fr/mceb2016/>
 2017: <http://www.lirmm.fr/mceb2017/>

ACKNOWLEDGMENTS

Thanks to our sponsors: IBC, Labex NUMEV and CeMEB, GDR BIM, LIRMM, CNRS, INRA and Université de Montpellier, and to Caroline Benoist and Sylvain Milanese for all their help in organizing the conference.

REFERENCES

- Gascuel O., editor. 2005. Mathematics and evolution and phylogeny. Oxford: Oxford University Press.
 Gascuel O., Steel M., editors. 2007. Reconstructing evolution: new mathematical and computational advances. Oxford: Oxford University Press.
 Gascuel O., Stadler T. 2015. Mathematical and computational evolutionary biology – 2013. *Syst. Biol.* 64(1):1–2.