Mathematical and Computational Evolutionary Biology (2013)

Olivier Gascuel[1] and Tanja Stadler[2]

[1]*Institut de Biologie Computationnelle (IBC), LIRMM, UMR5506, CNRS and Université de Montpellier, France and* [2]*Department of Biosystems Science and Engineering, Eidgenössische Technische Hochschule Zürich, 4058 Basel, Switzerland*
Correspondence to be sent to: Institut de Biologie Computationnelle (IBC), LIRMM, UMR5506, CNRS and Université de Montpellier, CC 05016, Campus St. Priest, Bât. 5, 860 rue de St. Priest, 34095 Montpellier cédex 5, France
E-mail: gascuel@lirmm.fr

This special issue of *Systematic Biology* contains review articles, contributed by keynote speakers after the fifth edition of the "Mathematical and Computational Evolutionary Biology" conference (MCEB; see web sites below) conference, held in 2013. We started in 2003 (under a slightly different name: "Mathematics of Evolution and Phylogeny") at the Mathematics Research Center Henri Poincaré, Paris, with well-known speakers like Walter Fitch and Joe Felsenstein. We had the feeling that the considerable efforts of the keynote speakers to synthesize and present their research in an extensive but understanding manner would deserve to be known by a large audience. Two books containing review articles corresponding to keynote presentations were published, following the 2003 and 2005 conference edition (Gascuel 2005; Gascuel and Steel 2007).

We then moved in 2008 to Montpellier, south of France, into a lovely, remote place, the "Hameau de l'Etoile", which conference attendees will surely remember. In 2012, we decided to start running the conference annually. Indeed, the mathematical and computational tools and concepts now form an essential basis for modern 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 such methodological works, by bringing together researchers from various disciplines: mathematics, computer science, phylogenetics, population genetics, and genomics. Every year we have 8–10 keynote speakers who introduce a field of research and discuss their own work in this field during a long talk (1h30). The afternoons are for short presentations and posters, with plenty of time for discussions. The number of attendees is limited (∼60) in order to favor small group interaction.

Moreover, we have a different focus every year, with emphasis on topical application domains. In 2013 the focus was given to the applications of mathematical and computational evolutionary biology to medicine, with regard to human and cancer genomics, microbiomes, genetic disease, and epidemics. General concepts, models, methods, and algorithms were also presented and discussed, just as during the previous conference editions. Applications of evolutionary biology in medicine and human health already have a long history (e.g., Stearns and Koella 2008), but are becoming highly topical with the advent of next generation sequencing data. Evolutionary tools and concepts are essential to decipher these data in a medical context, as illustrated by Gil McVean during his MCEB 2013 talk "Dissecting the genetic contribution to human disease" (Abecasis et al. 2012). We also had several presentations on the recent progress in molecular epidemiology highlighting how phylodynamic tools can be used to understand the spread of epidemics based on pathogen genetic sequencing data (Drummond et al. 2005; Volz et al. 2009; Rasmussen et al. 2011; Stadler et al. 2013). Such methods are currently being extensively used to understand the spread of Ebola in West Africa (Gire et al. 2014; Stadler et al. 2014; Volz and Pond 2014).

This special issue contains three review articles: on cancer, on the human microbiome, and on gene tree—species tree inference. *Systematic Biology*, with its strong history in attracting, handling and disseminating mathematically inclined articles incorporating phylogenetic methods, is the perfect choice for the publication of such papers. The review "Cancer evolution: mathematical models and computational inference", by Beerenwinkel, Schwarz, Gerstung, and Markowetz, is grounded on the well-known fact that cancer is an evolutionary process characterized by the accumulation of mutations in the genomes of affected cells. A large spectrum of recent approaches are reviewed, to model the initiation and progression of cancer using population dynamics concepts, to reveal the evolutionary relationships between tumor sub-clones using phylogenetic methods, and to describe dependencies among mutations using probabilistic graphical models. The review "Phylogenetics and the human microbiome" by Matsen deals with the analysis of the very large sequence data sets obtained from metagenomic studies of our microbiome. It shows that much of the analytical work for these sequences involves phylogenetics, at least indirectly, and reviews the field, its methods, and the current challenges from

a phylogenetic perspective. This special issue also contains an article by Szöllősi, Tannier, Daubin, and Boussau on "The inference of gene trees with species trees", a typical subject of *Systematic Biology* (and of MCEB, outside of our annual focus). This article reviews the various models used to describe the relationships between gene trees and species trees. These models account for gene duplication and loss, transfer, and incomplete lineage sorting, and may consider several types of events together. A number of approaches are described to reconstruct gene trees when a species tree is known, and vice-versa, and also to jointly infer gene trees and species trees. Simulations and empirical studies show that such sophisticated methods generally improve gene tree reconstruction.

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 new for *Systematic Biology*, will spread and be continued in the coming years.

Conference Web Sites

(all last accessed November 2014)

http://www.lirmm.fr/~guindon/ihp/
http://www.lirmm.fr/MEP05/
http://www.lirmm.fr/MIEP08/
http://www.lirmm.fr/mceb2012/
http://www.lirmm.fr/mceb2013/
http://www.lirmm.fr/mceb2014/
http://www.lirmm.fr/mceb2015/

References

Abecasis G.R., Auton A., Brooks L.D., DePristo M.A., Durbin R.M., Handsaker R.E., Kang H.M., Marth G.T., McVean G.A. 2012. An integrated map of genetic variation from 1,092 human genomes. Nature 491(7422):56–65.

Drummond A.J., Rambaut A., Shapiro B., Pybus O.G. 2005. Bayesian coalescent inference of past population dynamics from molecular sequences. Mol. Biol. Evol. 22(5):1185–1192.

Gascuel O., editor. 2005. Mathematics and evolution and phylogeny. Oxford, UK: Oxford University Press.

Gascuel O., Steel M., editors. 2007. Reconstructing evolution: new mathematical and computational advances. Oxford, UK: Oxford University Press.

Gire S.K., Goba A., Andersen K.G., Sealfon R.S., Park D.J., Kanneh L., Jalloh S., Momoh M., Fullah M., Dudas G., Wohl S., Moses L.M., Yozwiak N.L., Winnicki S., Matranga C.B., Malboeuf C.M., Qu J., Gladden A.D., Schaffner S.F., Yang X., Jiang P.-P., Nekoui M., Colubri A., Ruth Coomber M., Fonnie M., Moigboi A., Gbakie M., Kamara F.K., Tucker V., Konuwa E., Saffa S., Sellu J., Azziz Jalloh A., Kovoma A., Koninga J., Mustapha I., Kargbo K., Foday M., Yillah M., Kanneh F., Robert W., Massally J.L. B., Chapman S.B., Bochicchio J., Murphy C., Nusbaum C., Young S., Birren B.W., Grant D.S., Scheiffelin J.S., Lander E.S., Happi C., Gevao S.M., Gnirke A., Rambaut A., Garry R.F., Humarr Khan S., Sabeti P.C. 2014. Genomic surveillance elucidates Ebola virus origin and transmission during the 2014 outbreak. Science 345(6202):1369–1372.

Rasmussen D.A., Ratmann O., Koelle K. 2011. Inference for nonlinear epidemiological models using genealogies and time series. PLoS Comput. Biol. 7(8):e1002136.

Stadler T., Kühnert D., Bonhoeffer S., Drummond A. J. 2013. Birt–death skyline plot reveals temporal changes of epidemic spread in HIV and hepatitis C virus (HCV). Proc. Natl Acad. Sci. U. S. A. 110(1): 228–233.

Stadler T., Kühnert D., Rasmussen D.A., du Plessis L. 2014. Insights into the Early epidemic spread of Ebola in Sierra Leone provided by viral sequence data. 1st ed. PLoS Currents: Outbreaks. 6 October.

Stearns S.C., Koella J.K. 2008. Evolution in health and disease. 2nd ed. Oxford, UK: Oxford University Press.

Volz E., Pond S.L.K. 2014. Phylodynamic analysis of Ebola virus in the 2014 Sierra Leone Epidemic. 1st ed. PLoS Currents: Outbreaks. 24 October.

Volz E.M., Pond S.L.K., Ward M.J., Brown A.J.L., Frost S.D. 2009. Phylodynamics of infectious disease epidemics. Genetics 183(4):1421–1430.