

Evolutionary Genomics

■ Yuval Itan

Human Genetics of Infectious Diseases, The Rockefeller University, New York, NY, USA.

■ Pascale Gerbault

Research Department of Genetics, Evolution and Environment, University College London, London, UK.

■ Gur Pines

Department of Chemical and Biological Engineering, University of Colorado at Boulder, Boulder, CO, USA.

Supplement Aims and Scope

This supplement is intended to focus on evolutionary genomics. *Evolutionary Bioinformatics* aims to provide researchers working in this complex, quickly developing field with online, open access to highly relevant scholarly articles by leading international researchers. In a field where the literature is ever-expanding, researchers increasingly need access to up-to-date, high quality

scholarly articles on areas of specific contemporary interest. This supplement aims to address this by presenting high-quality articles that allow readers to distinguish the signal from the noise. The editor in chief hopes that through this effort, practitioners and researchers will be aided in finding answers to some of the most complex and pressing issues of our time.

Evolutionary genomics covers a wide range of subjects investigating the evolution of species' genomes.¹ This supplement presents various models of evolution, from bacteria² and plants³ to humans.⁴⁻⁷ Indeed, the similarities in genetic codes across all types of organisms allow for comparisons of DNA sequences between and within species.⁸ Interpretation of these comparisons can nonetheless be challenging because of the diversity of approaches available and the way genotypes are determined, ie genotyping or sequencing techniques used, and characteristics of the markers or loci of interest.⁵

The extent to which observed changes in DNA sequences are adaptive has always been central to the understanding of evolution and is reflected here by the number of papers addressing the effect of natural selection on observed patterns of diversity.^{3,4,8} Furthermore, other processes, including genetic drift, gene overlap, migration, mutation and recombination interact with selection to shape patterns of diversity.⁶ These interactions can lead to counterintuitive outcomes. For example, the highest frequencies of the allele associated with lactase persistence in Europe are found in the North-West. This

may lead to the assumption that the allele arose and diffused from there. However, computer simulations have shown that the joint effect of positive selection and allele-surfing with the spread of farmers from the Near East about 8,000 years ago could lead to a similar pattern.^{9,10} Computational models continue to be insightful analytic tools to understand the interplay between demographic and adaptation processes.^{3,4,9}

Recent advances in high throughput genomics techniques provide better coverage of organism populations, and genomes. Consequently, the field of evolutionary genetics that was mainly model-based in the 20th century has more recently been driven by big data to become the relatively new field of evolutionary genomics. Genome-wide surveys of populations have proven power to detect associations between genetic variants and diseases or complex traits and to investigate for the effects of selection.¹¹ They show that evolutionary genetics is more relevant to practical problems than ever before, such as plant improvement³ and human health,⁷ where evolutionary insights aid in the identification of patients' disease-causing alleles and genes.^{12,13} A single change in the DNA can have multiple consequences at the organism level.¹⁴ Studies that



will approach these various aspects of a change in the DNA, including genomic context, functional effect, and evolutionary modeling (ways an adaptive change increases in frequency in a population), will provide a better understanding of the adaptability of organisms and species.¹⁵

Evolutionary genomics could also be applied to identify potentially beneficial genetic variants from the single gene to the whole genome scale, for human needs such as for the production of commodity chemicals and pharmaceuticals. Such genomic approaches are mainly directed at two extremes: (1) focusing on a single or few genomic loci to thoroughly investigate responses to localized edits, such as the recombineering technique used mainly in *E. coli*,^{16,17} and (2) an approach termed Adaptive Laboratory Evolution (ALE) in which cells are being challenged with an environmental stress and adapt to regain fitness.^{18,19} Between these two extremes, one using prior knowledge and rational design and the other a completely unbiased ALE approach, other approaches emerged that try to incorporate the best of each method. Such approaches, such as MAGE and TRMR,^{20,21} combine rational design with multiple genome-wide edits. Recent advances in genome engineering tools such as the CRISPR-CAS9 system^{22,23} hold promise for even more precise and systematic editing technologies, which will allow rational whole genome surveys for specific traits at the single nucleotide resolution.

Future directions in evolutionary genomics may apply biological distance metrics alternative to sequence similarity, such as the human gene connectome,^{24,25} gene-level quantitative accumulated mutational damage and intolerance, and gene-specific mutation impact predictions to complement traditional selective pressure approaches.^{26–28} In the era of high throughput data, efficient computational methodologies are essential.²⁹ A strong interdisciplinary effort will facilitate solving the various challenges in today's evolutionary genomics.

REFERENCES

- Ridley M. *Evolution*. Blackwell; 2003.
- Lin K. Phylogenomic analysis of closely related bacterial genomes: Combining the features of both overlapping genes and collinear genomic regions. *Evolutionary Bioinformatics*. 2016, In Press.
- Allaby R. An individual based model to explore selection in plant domestication. *Evolutionary Bioinformatics*. 2016, In Press.
- Currat M, Gerbault P, Sanchez-Mazas A. SELECTOR: a program to simulate genetic lineages under selection in a spatially-explicit population framework. *Evolutionary Bioinformatics*. 2016, In Press.
- Nunes M. Using unformat and Gene[rate] to analyse data with ambiguities in population genetics. *Evolutionary Bioinformatics*. 2016, In Press.
- Lopez S, Van Dorp L, Hellenthal G. Genetic studies supporting the human migration out of Africa and waves of dispersal to other continents. *Evolutionary Bioinformatics*. 2016, In Press.
- Mezzavilla M, Gasparini P. Patterns of genetic variation in BMI and obesity-related genes across the Silk Road. *Evolutionary Bioinformatics*. 2016, In Press.
- Diekmann Y, Pereira-Leal JB. Be aware of phylogeny: Gene tree affects inference of sites under episodic positive selection. *Evolutionary Bioinformatics*. 2016, In Press.
- Itan Y, Powell A, Beaumont MA, Burger J, Thomas MG. The origins of lactase persistence in Europe. *PLoS Comput Biol*. 2009;5:e1000491.
- Gerbault P, Liebert A, Itan Y et al. Evolution of lactase persistence: an example of human niche construction. *Philos Trans R Soc Lond B Biol Sci*. 2011;366:863–77.
- Bigham AW, Mao X, Mei R et al. Identifying positive selection candidate loci for high-altitude adaptation in Andean populations. *Hum Genomics*. 2009;4:79–90.
- Casanova JL, Conley ME, Seligman SJ, Abel L, Notarangelo LD: Guidelines for genetic studies in single patients: lessons from primary immunodeficiencies. *J Exp Med*. 2014;211:2137–49.
- Goldstein DB, Allen A, Keebler J et al. Sequencing studies in human genetics: design and interpretation. *Nat Rev Genet*. 2013;14:460–70.
- Kamberov YG, Wang S, Tan J et al. Modeling Recent Human Evolution in Mice by Expression of a Selected EDAR Variant. *Cell*. 2013;152:691–702.
- Eldholm V, Norheim G, von der Lippe B et al. Evolution of extensively drug-resistant *Mycobacterium tuberculosis* from a susceptible ancestor in a single patient. *Genome Biol*. 2014;15:490.
- Court DL, Sawitzke JA, Thomason LC. Genetic engineering using homologous recombination. *Annu Rev Genet*. 2002;36:361–88.
- Pines G, Freed EF, Winkler JD, Gill RT. Bacterial Recombineering: Genome Engineering via Phage-Based Homologous Recombination. *ACS Synth Biol*. 2015;4:1176–85.
- Tenaillon O, Rodriguez-Verdugo A, Gaut RL et al. The molecular diversity of adaptive convergence. *Science*. 2012;335:457–61.
- Toprak E, Veres A, Michel JB, Chait R, Hartl DL, Kishony R. Evolutionary paths to antibiotic resistance under dynamically sustained drug selection. *Nat Genet*. 2012;44:101–5.
- Wang HH, Isaacs FJ, Carr PA et al. Programming cells by multiplex genome engineering and accelerated evolution. *Nature*. 2009;460:894–8.
- Freed EF, Winkler JD, Weiss SJ et al. Genome-Wide Tuning of Protein Expression Levels to Rapidly Engineer Microbial Traits. *ACS Synth Biol*. 2015;4:1244–53.
- Jiang W, Bikard D, Cox D, Zhang F, Marraffini LA. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat Biotechnol*. 2013;31:233–9.
- Pines G, Pines A, Garst AD, Zeitoun RI, Lynch SA, Gill RT. Codon compression algorithms for saturation mutagenesis. *ACS Synth Biol*. 2015;4:604–14.
- Itan Y, Zhang SY, Vogt G et al. The human gene connectome as a map of short cuts for morbid allele discovery. *Proc Natl Acad Sci U S A*. 2013;110:5558–63.
- Itan Y, Mazel M, Mazel B et al. HGCS: an online tool for prioritizing disease-causing gene variants by biological distance. *BMC Genomics*. 2014;15:256.
- Itan Y, Shang L, Boisson B et al. The human gene damage index as a gene-level approach to prioritizing exome variants. *Proc Natl Acad Sci U S A*. 2015;112:13615–20.
- Petrovski S, Wang Q, Heinzen EL, Allen AS, Goldstein DB. Genic intolerance to functional variation and the interpretation of personal genomes. *PLoS Genet*. 2013;9:e1003709.
- Itan Y, Shang L, Boisson B et al. The mutation significance cutoff: gene-level thresholds for variant predictions. *Nat Methods*. 2016;13:109–10.
- Shendure J, Ji H. Next-generation DNA sequencing. *Nat Biotechnol*. 2008;26:1135–45.

Lead Guest Editor **Dr Yuval Itan**

Dr Yuval Itan is Research Associate of Computational Human Genomics at The Rockefeller University. He completed his PhD at University College London. He now works primarily in correlating human genotypes with disease phenotypes. Dr Itan is the author or co-author of over 30 published papers, has orally presented at 20 conferences, and has been an ad hoc reviewer for prestigious journals such as *Nature Methods*, *PNAS* and *PLoS Computational Biology*.



yitan@rockefeller.edu

<http://lab.rockefeller.edu/casanova/HGC>

<http://lab.rockefeller.edu/casanova/GDI>

<http://lab.rockefeller.edu/casanova/MS>

Guest Editors

GUR PINES

Dr Gur Pines is a Research Associate of Chemical and Biological Engineering at University of Colorado Boulder. He completed his PhD at Weizmann Institute of Science. He now works primarily on directed evolution and on developing tools for bacterial genome editing. Dr Pines is the author or co-author of 10 published papers.



gur.pines@colorado.edu

<https://profiles.ucdenver.edu/display/228447>

PASCALE GERBAULT

Dr Pascale Gerbault is a Research Associate of Human Evolutionary Genetics at University College London (UCL). She completed her PhD at UCL and has previously worked at the University of Geneva, Switzerland, and the University of Montreal, Canada. She now works primarily in evolutionary genetics, with a focus on genetic anthropology and detection of signals of selection on the genome. Dr Gerbault is the author or co-author of ten published peer-reviewed journal papers.



p.gerbault@ucl.ac.uk

<https://www.ucl.ac.uk/mace-lab/people/pascale>

SUPPLEMENT TITLE: Evolutionary Genomics

CITATION: Itan et al. Evolutionary Genomics. *Evolutionary Bioinformatics* 2015:11(S2) 53–55 doi: 10.4137/EBO.S39729.

TYPE: Editorial

FUNDING: Authors disclose no external funding sources.

COMPETING INTERESTS: Authors have disclosed no potential conflicts of interest.

COPYRIGHT: © the authors, publisher and licensee Libertas Academica Limited. This is an open-access article distributed under the terms of the Creative Commons CC-BY-NC3.0 License.

CORRESPONDENCE: yitan@rockefeller.edu

All editorial decisions were made by the independent academic editor. All authors have provided signed confirmation of their compliance with ethical and legal obligations including (but not limited to) use of any copyrighted material, compliance with ICMJE authorship and competing interests disclosure guidelines.