

## EDITORIAL

# Next-generation sequencing propels environmental genomics to the front line of research

*Heredity* (2015) **114**, 429–430; doi:10.1038/hdy.2015.23

So far in the 21st century, next-generation sequencing (NGS) has been the major technological breakthrough in biology, promoting the emergence of a new multi-disciplinary approach: environmental genomics (Ungerer *et al.*, 2008; Landry and Aubin-Horth, 2014). Environmental genomics captures the structure and dynamics of a range of molecular components, including DNA and RNA in (meta) genomes and (meta)transcriptomes. In combination with other high-throughput approaches, called (meta)omics, it addresses an integrated understanding of complex biological systems at various organization levels (individual, population and community) and in different dimensions of time and space. NGS allows us to delve deeper to uncover biological systems, permitting measurement of species richness (Taberlet *et al.*, 2012), abundance, and hence how species can interact when sharing and competing for resources. The cryptic, as well as the rare but active fraction of biodiversity are now accessible to study in all environments (Scheffers *et al.*, 2012; Hugoni *et al.*, 2013).

Understanding the evolutionary history of organisms, as well as identifying genes or genomic structures that are involved in adaptive evolution (that is, in response to natural selection) can be answered by NGS approaches. The identification of important functional variants requires the integration of multiple approaches including molecular population genetics, phylogenetics and formal genetics (Feder and Mitchell-Olds, 2003). NGS technologies allow us to address evolutionary and population issues hitherto unthinkable, and even those considered part of the field of theoretical biology (Savolainen *et al.*, 2013).

This special issue, coming from biennial symposia launched by the French national network on Environmental Genomics (French National Network on Environmental Genomics (<http://www.cnrs.fr/inee/recherche/actionsincitatives-RTP-Genoenvironnementale.htm>; 2011, 2013, 2015), exemplifies the use of NGS in prokaryotes and eukaryotes, from single organisms to complex communities living in marine, terrestrial and atmospheric environments. It comprises three review articles and eight original papers reporting scientific and technical advances in genomics to decipher biological diversity and species delineation, functional genomics, as well as genome evolution through the study of environmental selection and strategies for extracting relevant biological knowledge in non-model species.

Schlotterer *et al.* (2014) revisit the ‘evolve and resequence’ approach to highlight genes and SNPs under selection through sequencing of pools of individuals. Experimental evolution has long been used as a traditional approach to uncover selective processes at the origin of population polymorphisms. The challenge is to increase the power and accuracy in the detection of genomic polymorphisms, reducing false positives and at a limited cost. The authors discuss optimal experimental and analytical strategies using Pool-Seq method through

various biological examples, comparing this method with other genome-wide association methods.

Genome-wide evolutionary analyses are the focus of the two following papers investigating biological diversity at the species level. Peyretailade *et al.* (2014) review the role of the obligate intracellular parasites, that is, Microsporidia fungi, in the food web at various trophic levels. Using different omics tools, the authors evaluate their diversity and abundance, and discuss their impact on ecosystem functioning. The study of Pante *et al.* (2014) investigates the Rad-seq method to delineate species within a rapidly diverged group for which very little genomic information is available. It focuses on deep-sea octocorals, *Chrysogorgia*, known to be widely distributed and to exhibit conflicting patterns between morphological and molecular data. By comparing their results with those on cetacean, beetles, swordtail and platyfish, the authors pointed out limitations of Rad-seq for phylogenetic reconstructions, which may be taxon dependent and may interfere with molecular clocks.

The three next papers investigate the structure, dynamics and functioning of bacterial populations and communities in different environments. An *et al.* (2014) highlight the changes caused by bacterial communities transported along with the dust in atmospheric sandstorms, over planetary distances. The authors hypothesize that the origin of the sandstorms may be explained by variations among sites with regards to wind sources and trajectories. They outline the potential contribution of atmospheric sandstorm bacterial communities in the sand deposited by the sandstorm that may drastically alter human health, in conjunction with other transported mineral or chemical pollutants. Terrat *et al.* (2014) compare an existing DNA fingerprint method (ARISA) and NGS metabarcoding to evaluate the taxa–area relationships among microbial communities in soil. The authors conclude that NGS metabarcoding provides a very accurate evaluation of soil microbial community assemblies in broad-scale biogeographic studies. Kwaziborski *et al.* (2015) combine genomics and transcriptomics to investigate bacterial (*Pectobacterium–Rhodococcus*) interactions within the plant environment. The authors analyze how chemical communication among bacteria may coordinate or alter gene expression of members of plant microbiotes. This work sheds new light on mechanisms that may be further developed as biocontrol agents in the course of plant protection against pathogens.

The following two papers present methodologies that will be useful in studies dealing with the characterization of adaptive genes in evolutionary ecology. Ferrandiz-Rovira *et al.* (2015) present methodologies for reliably genotyping large numbers of individuals, for example in the case of parental relationships or population genomics that may be used in any genetic system and in any organism. Genotyping was obtained using NGS at four MHC loci

in wild Alpine marmots. The limitations, according to genome characteristics (for example, high copy number variation), loci of interest and commercial suitable platforms, are discussed. Gouin *et al.* (2014) rescue unmapped reads generally discarded from NGS analyses. The authors highlight the potential loss of information that these may bring. This approach is particularly important in non-model species or organisms whose reference genome is of poor quality and where the level of divergence between the reads and the reference genome is relatively high. Using re-sequencing approaches and an original pipeline, this study discovered variation in symbiotic composition during ecological speciation in Aphids among individual host genomes that are linked to the origin of the biotypes from which they arise.

The following three papers highlight the benefits of different NGS methods, and their combination, to understand the evolutionary response of organisms facing environmental changes. Porcelli *et al.* (2015) provide a review of emerging patterns of putative adaptive genomic responses, using the case of heat stress in metazoans. Thanks to a combination of both genomic and transcriptomic information, they uncover specific classes of genes under selection, from the regulation of gene transcription to mRNA translation. Future directions are discussed, and include combining RNA-seq with either proteomics or metabolomics to elucidate genetic networks plus regulatory and metabolic responses during adaptation to different thermal environments. The impact of genetic architecture on species evolution and adaptation is investigated in the following studies. Using standard linkage mapping methods after RAD-seq, Huber *et al.* (2015) examine species diversification linked to mimicry polymorphism in a clade of South American *Heliconius* species. Thanks to a combined set of QTL methods, from multivariate quantitative genetics to fine-mapping candidate genes, the authors confirmed the complex genomic architecture of wing patterning in this genus and explored the evolution of genetic architecture within a larger comparative framework. Lopes *et al.* (2015) investigate ecological mechanisms involved in the avoidance of competition between closely related sympatric species through a DNA metabarcoding diet analysis. Hi-seq sequencing was performed using markers characterizing the majority of the plant richness, as well as the mollusks and arthropods from the original habitat of two parapatric species of subterranean rodents living in the southern Brazilian coastal plain, and then compare with total DNA extraction from feces and stool samples. The results confirm the herbivorous status of the species and suggested a dietary partitioning in the sympatric region.

Global change, coming from natural phenomena and human activities, can drastically impact biological systems at unprecedented speeds. The diversity of NGS approaches, together with bioinformatics developments, open new areas that go more deeply in the characterization of biological systems, ranging from complex regulatory networks within communities up to metabolomics network within a single cell. Knowing what is there, what functions are performed and how regulatory pathways are activated, are challenging questions for which environmental genomics can provide important insights.

#### CONFLICT OF INTEREST

The authors declare no conflict of interest.

#### ACKNOWLEDGEMENTS

We thank Pr. M DuBow (Université Paris-Sud, Orsay) for critical reading of the editorial, as well as CNRS for supporting the RTP Génomique Environnementale (2011–2013) and GDR3692 Génomique Environnementale (2015–2018). Finally, they would also like to particularly thank Michael Bruford for his help and his patience during the submission and editing process.

D Joly<sup>1,2</sup> and D Faure<sup>1,3</sup>

<sup>1</sup>GDR3692 Génomique Environnementale, CNRS, Gif-sur-Yvette Cedex, France; <sup>2</sup>Laboratoire Evolution, Génomes, Comportement, Ecologie UMR9191 CNRS-IRD-Paris-Sud, Gif-sur-Yvette Cedex, France and <sup>3</sup>Institut for Integrative Biology of the Cell, I2BC, UMR9198 CNRS CEA Université Paris-Sud, Saclay Plant Sciences, Gif-sur-Yvette Cedex, France  
E-mail: Dominique.Joly@cnrs-dir.fr

- An S, Sin HH, DuBow MS (2014). Modification of atmospheric sand-associated bacterial communities during Asian sandstorms in China and South Korea. *Heredity* **114**: 460–467.
- Feder ME, Mitchell-Olds T (2003). Evolutionary and ecological functional genomics. *Nat Rev Genet* **4**: 651–657.
- Ferrandiz-Rovira M, Bigot T, Allaine D, Callait-Cardinal MP, Cohas A (2015). Large-scale genotyping of highly polymorphic loci by next generation sequencing: how to overcome the challenges to reliably genotype individuals? *Heredity* **114**: 485–493.
- French National Network on Environmental Genomics (Lyon 2011, Rennes 2013, Montpellier 2015): <http://www.cnrs.fr/inee/recherche/actionsincitatives-RTP-Genoenviromentale.htm>.
- Gouin A, Legeai F, Nouhaud P, Whibley A, Simon JC, Lemaitre C (2014). Whole-genome re-sequencing of non-model organisms: lessons from unmapped reads. *Heredity* **114**: 494–501.
- Huber B, Whibley A, Le Poul Y, Navarro N, Martin A, Baxter S *et al.* (2015). Conservatism and novelty in the genetic architecture of adaptation in *Heliconius* butterflies. *Heredity* **114**: 515–524.
- Hugoni M, Taib N, Debros D, Domaizon I, Jouan Dufournel I, Bronner G *et al.* (2013). Structure of the rare archaeal biosphere and seasonal dynamics of active ecotypes in surface coastal waters. *Proc Natl Acad Sci* **110**: 6004–6009.
- Kwaziborski A, Mondy S, Chong TM, Barbey C, Chan KG, Beury-Cirou A *et al.* (2015). Transcriptome of the quorum-sensing signal-degrading *Rhodococcus erythropolis* responds differentially to virulent and avirulent *Pectobacterium atrosepticum*. *Heredity* **114**: 476–484.
- Landry CR, Aubin-Horth N (2014). *Ecological Genomics: Ecology and the Evolution of Genes and Genomes*, Vol 781. Springer.
- Lopes CM, De Barba M, Boyer F, Mercier C, da Silva Filho PJ, Heidtmann LM *et al.* (2015). DNA metabarcoding diet analysis for species with parapatric vs sympatric distribution: a case study on subterranean rodents. *Heredity* **114**: 525–536.
- Pante E, Abdelkrim J, Viricel A, Gey D, France SC, Boisselier MC *et al.* (2014). Use of RAD sequencing for delimiting species. *Heredity* **114**: 450–459.
- Peyretailade E, Boucher D, Parisot N, Gasc C, Butler R, Pombert JF *et al.* (2014). Exploiting the architecture and the features of the microsporidian genomes to investigate diversity and impact of these parasites on ecosystems. *Heredity* **114**: 441–449.
- Porcelli D, Butlin RK, Gaston KJ, Joly D, Snook RR (2015). The environmental genomics of thermal adaptation. *Heredity* **114**: 502–514.
- Savolainen O, Lascoux M, Merila J (2013). Ecological genomics of local adaptation. *Nat Rev Genet* **14**: 807–820.
- Scheffers BR, Joppa LN, Pimm SL, Laurance WF (2012). What we know and don't know about Earth's missing biodiversity. *Trends Ecol Evol* **27**: 501–510.
- Schlotterer C, Kofler R, Versace E, Tobler R, Franssen SU (2014). Combining experimental evolution with next-generation sequencing: a powerful tool to study adaptation from standing genetic variation. *Heredity* **114**: 431–440.
- Taberlet P, Coissac E, Hajibabaei M, Rieseberg LH (2012). Environmental DNA. *Mol Ecol* **21**: 1789–1793.
- Terrat S, Dequiedt S, Horrigue W, Lelievre M, Cruaud C, Saby NP *et al.* (2014). Improving soil bacterial taxa-area relationships assessment using DNA meta-barcoding. *Heredity* **114**: 468–475.
- Ungerer MC, Johnson LC, Herman MA (2008). Ecological genomics: understanding gene and genome function in the natural environment. *Heredity* **100**: 178–183.