

## NEWS AND VIEWS PERSPECTIVE

### Ecological annotation of genes and genomes through ecological genomics

CHRISTIAN R. LANDRY\* and  
NADIA AUBIN-HORTH†

\*Département de Biochimie, Faculté de Médecine, †Département de Sciences Biologiques, Université de Montréal, CP 6128, Succursale Centre Ville, Montréal, Québec, Canada H3C 3J7

#### Abstract

Ecological genomics is a research field that aims to determine how a genome or a population of genomes interacts with its environment across ecological and evolutionary timescales. This matter was the central theme of the symposium on Ecological Genomics that took place at the First meeting of the Canadian Society for Ecology and Evolution, held at the University of Toronto in May 2007. Through their research on a diverse array of organisms, the various speakers illustrated how ecology and evolution benefit from genomics, and indirectly how genomics can benefit from evolutionary ecology.

#### Analysis of QTL in natural settings to test long-standing evolutionary questions

There is a large body of theory on adaptation of populations to new environmental conditions. However, what type of genes, how many loci and what types of mutations and of what effects are involved in shaping adaptive traits in natural populations are still largely unknown. The threespine stickleback (*Gasterosteus aculeatus*) is an important species in ecology and evolution for our understanding of fundamental processes such as speciation and adaptive radiation. It is also becoming an important model in genomics, with resources that include a linkage map, a sequenced genome on the way and identified quantitative trait loci (QTL) and specific segregating nucleotide substitutions associated with adaptive phenotypes. As presented by Sean Rogers (University of British Columbia), these bodies of knowledge can now be combined to tackle current models of adaptation (Orr 2005). By dissecting the genetic architecture of adaptive divergence of freshwater populations from a marine ancestor, Rogers and his colleague Dolph Schluter hope to be able to understand how many mutational steps and of what size this small fish has taken in the recent past to climb the adaptive fitness hills.

Correspondence: Christian R. Landry, Fax: 514-343-3205; E-mail: clandry@post.harvard.edu; Nadia Aubin-Horth, Fax: 514-343-2293; E-mail: n.aubin-horth@umontreal.ca

### Life cycle variation through the lens of genomics

Variation in life histories is a central question in evolutionary biology. Yet, how this variation is encoded in genomes is still mostly unknown. Nadia Aubin-Horth (Université de Montréal) takes advantage of genomics in an ecological model species, Atlantic salmon (*Salmo salar*), sampled in its natural habitat, to explore this question. Aubin-Horth and colleagues showed how life cycle variations are related to large-scale gene expression changes in the brain, which integrates external and internal signals and orchestrates many phenotypic changes (Aubin-Horth *et al.* 2005). She first compared males that matured at a very small size without leaving freshwater to males of the same age that do not. She then compared individuals that migrate to the ocean the following year to individuals that stay in fresh water. By combining genomic information on each life-cycle stage, she revealed that a particular set of genes was recruited in both early maturation and migration. This illustrates how a given set of genes is re-used to build the different stages of this complex life cycle, shedding light on how life-history trade-offs unravel at the molecular level (Roff 2007).

### Genomics of ecological speciation

Recent and ongoing speciation of North American lake whitefish (*Coregonus clupeaformis*) is a prime example of how ecology is driving phenotypic evolution. This freshwater salmonid has diverged several times into two forms in lakes created at the retreat of the continental ice sheet. After having discovered the morphological and behavioural traits that characterized the dwarf and normal ecotypes through population genetics, Louis Bernatchez (Université Laval) sought to understand the genomic architecture that differentiates them. His team developed linkage maps and identified QTL that underlie the adaptive changes between these two emerging species. They found evidence that the loci associated with these QTL were distributed in non-neutral fashion in nature, which confirms adaptive changes on a genome-wide scale (Rogers & Bernatchez 2007). They also tested whether the transcriptome of these two forms might have evolved in parallel in different lakes. Using microarrays developed for a closely related salmonid, their work uncovered parallel changes in expression level of several key genes involved in energy metabolism, such as ATP synthase C, which is expressed at higher levels in muscles of the dwarf ecotype and is likely associated with increased swimming activity (Derome *et al.* 2006). This confirmed that changes in muscle physiology and energetics covary with their morphological and behavioural differentiation. Bernatchez's investigations illustrate perfectly how genomic resources for a closely related, anchor species can be used to answer key evolutionary questions. This *tour de force* is a reference in how ecology and genomics can meet to illuminate evolution.

### Ecological genomics of laboratory model organisms

*Saccharomyces cerevisiae*, the wine yeast, has for a long time been considered a domesticated species, but surveys of its natural isolates associated with insects, fruits and trees revealed the species' tremendous variation in nature. Given the wealth of knowledge on its molecular and cellular biology, this species offers unprecedented opportunities to ask fundamental questions about the molecular bases of phenotypic variation in ecologically relevant settings. Christian Landry (Harvard University) used yeasts isolated from the wild and DNA microarrays to test whether natural strains show genetic variation in their transcriptional responses to environmental perturbations (Landry *et al.* 2006). Landry and colleagues showed that not only was gene expression level variable among strains, but the extent of variation depended on the growth conditions. Furthermore, genes that showed genetic variation for gene expression level were particularly enriched for genes that are known to be nonessential for growth in rich laboratory conditions. Landry's work underscores the need to study genes and genomes of model organisms in an ecological context to understand the functional significance of gene expression variation.

### Genetic perturbations and responses to natural conditions

Trees are usually seen as slow-growing plants that are not amenable to genetics or experimental biology. Poplar (*Populus trichocarpa*) is an exception to this rule. With a sequenced genome size only four times larger than that of the model plant *Arabidopsis*, the ability to perform nonhomologous mutagenesis and clonal propagation, as well as fast growth in controlled conditions, *P. trichocarpa* is a perfect ecological genomics model tree. Sharon Regan (Queen's University) is using these qualities to uncover the molecular basis of phenotypic variation for traits that are important for our understanding of both natural variation and potential targets for bioenergy breeding programmes. Although most genomic screens in model organisms are done with gene knockouts, Regan has opted for an approach where she can induce gene expression at random positions in the genome. With hundreds of over-expression lines in hand, her laboratory is screening traits known for ecologically important variation, such as pubescence, lignin content and branching patterns, with the aim to identify genes that have important functions in natural settings.

### Environmental metagenomics

The development of high throughput sequencing has enabled the rise of the new field of metagenomics, which seeks to characterize microbial communities at the sequence level without the need for isolation and cultivation, by sampling them directly in their natural habitat. We know next to nothing about the diversity of RNA viruses in the ocean, despite the fact that they contribute as much biomass as millions of blue whales would. By extracting RNA molecules from ocean

water off the coast of British Columbia and sequencing their reversed transcribed DNA products, Andrew Lang (Memorial University) discovered unprecedented diversity of RNA viruses (Culley *et al.* 2006). There was such a diversity of sequences in the DNA libraries derived from this sea water that only one-third to one-quarter of their sample showed significant similarity to sequences in major databases. He also found geographical variation in this diversity. Using computational methods developed for whole-genome shotgun assembly, Lang and colleagues assembled whole viral genomes that revealed novel evolutionary lineages whose biology is unknown. Lang's work and that of many others (Venter *et al.* 2004) is launching a new era of natural history by studying the invisible and unculturable marine microcosm that likely contains many more surprises, both in terms of species and molecular diversity.

### Gene networks in adaptations

Evolutionary and ecological studies have long treated organisms as black boxes, due to the lack of information on their genes and gene networks. However, we can now take advantage of genetics and genomics information from model organisms and study these species in ecological settings. The ecological relevance of phenotypic plasticity of *Arabidopsis thaliana* to temperature and lighting conditions has been investigated thoroughly. However, how nucleotide variation at key genes maps to variation in phenotypes is currently unknown. For example, cold weather makes plants such as *Arabidopsis* competent to flower through a process called vernalization. The *Flowering Locus C (FLC)* gene plays a central role in this response as it inhibits flowering and is repressed by vernalization. John Stinchcombe (University of Toronto) and colleagues tested whether allelic variation at this locus was associated with vernalization sensitivity (Stinchcombe *et al.* 2005). Their results show a cline in vernalization sensitivity along the latitudinal gradient of origin of the 21 European accessions surveyed. Additionally, one *FLC* allele exhibited greater variance in phenotype, suggesting that other polymorphic loci may be associated with this important ecological trait. This reveals even greater complexity than expected in the molecular regulatory network, once again showing how studying genomes in an ecological context is a crucial step in the comprehension of gene functions.

### Ecological annotation of genes and genomes

From this series of talks, it becomes evident that the fields of ecology and evolution benefit greatly from the genomic tools developed to study laboratory model organisms. What is less obvious is that by investigating the roles of genes, gene variants and gene combinations in natural or seminatural settings, ecological genomics contributes to fundamental genomics by providing a second order of annotation, i.e. an *ecological* annotation. This is particularly evident for model organisms. Although most of what we know from eukaryotic cell biology is derived from work on *Saccharomyces cerevisiae*, a significant fraction of its genome (15%) remains to be functionally described. This lack of knowledge in great part

stems from the fact that gene function is environment-specific. For this reason, yeast geneticists now recognize the need to study genes in their ecological context (Pena-Castillo & Hughes 2007). This undoubtedly extends to all model organisms. Ecological genomics shows how connections between fields of research that have been traditionally separated can answer fundamental questions that neither field could have offered on its own.

## References

- Aubin-Horth N, Landry CR, Letcher BH, Hofmann HA (2005) Alternative life histories shape brain gene expression profiles in males of the same population. *Proceedings of the Royal Society of London, Series B, Biological Sciences*, **272**, 1655–1662.
- Culley AI, Lang AS, Suttle CA (2006) Metagenomic analysis of coastal RNA virus communities. *Science*, **312**, 1795–1798.
- Derome N, Duchesne P, Bernatchez L (2006) Parallelism in gene transcription among sympatric lake whitefish (*Coregonus clupeaformis* Mitchell) ecotypes. *Molecular Ecology*, **15**, 1239–1249.
- Landry CR, Oh J, Hartl DL, Cavalieri D (2006) Genome-wide scan reveals that genetic variation for transcriptional plasticity in yeast is biased towards multi-copy and dispensable genes. *Gene*, **366**, 343–351.
- Orr HA (2005) The genetic theory of adaptation: a brief history. *Nature Reviews Genetics*, **6**, 119–127.
- Pena-Castillo L, Hughes TR (2007) Why are there still over 1000 uncharacterized yeast genes? *Genetics*, **176**, 7–14.
- Roff DA (2007) Contributions of genomics to life-history theory. *Nature Reviews Genetics*, **8**, 116–125.
- Rogers S, Bernatchez L (2007) The genetic architecture of ecological speciation and the association with signatures of selection in natural lake whitefish (*Coregonus* sp. Salmonidae) species pairs. *Molecular Biology and Evolution*, **24**, 1423–1438.
- Stinchcombe J, Caicedo A, Hopkins R *et al.* (2005) Vernalization sensitivity in *Arabidopsis thaliana* (Brassicaceae): the effects of latitude and FLC variation. *American Journal of Botany*, **92**, 1701–1707.
- Venter JC, Remington K, Heidelberg JF *et al.* (2004) Environmental genome shotgun sequencing of the Sargasso Sea. *Science*, **304**, 66–74.

---

Christian Landry is a postdoctoral fellow at Université de Montréal. His research focuses on the evolution of genes, genomes and gene networks in model organisms. Nadia Aubin-Horth is an assistant professor at Université de Montréal. Her laboratory uses an integrative approach combining functional genomics, endocrinology, evolution and behavioural biology to study the molecular bases of variation in behaviour and life cycles.

---