

Preface to the special issue: ecological and evolutionary genomics of populations in nature

CAROL EUNMI LEE* and THOMAS MITCHELL-OLDS†

**Wisconsin Institute of Rapid Evolution, Department of Zoology, University of Wisconsin, Madison, 430 Lincoln Drive, Wisconsin 53706*, †*Department of Biology, Duke University, Box 90338, Durham, North Carolina 27708-0338, USA*

This Special Issue focuses on the application of genomic approaches for addressing central questions in ecology and evolutionary biology using populations in nature. Most of these studies focus on non-model systems, as many key questions in ecology and evolutionary biology cannot be addressed solely using model systems (Gracey & Cossins 2003). While model systems have been instrumental in yielding fundamental insights into genetic and evolutionary mechanisms, they represent only a tiny fraction of organismal diversity. Model organisms tend to have short generation times, cosmopolitan geographic distributions, and generalist strategies under diverse environmental conditions (such as the laboratory). As such, studies using model systems tend to exclude locally adapted species and rare endemics, parasites and symbionts, and species with complex life histories or long generation times.

While applying genomic approaches to non-model systems is still quite challenging, the dichotomy between model and non-model systems is diminishing as genomic techniques become more readily available. Techniques such as high-throughput SNP discovery, custom cDNA microarrays, and inexpensive whole-genome sequencing will serve to make genomic approaches widespread and routine for many non-model systems. In addition, it is possible now to extend information on gene function from model organisms to diverse organisms in nature.

This Special Issue represents early attempts to apply genomic approaches to wild populations, with studies that address questions regarding local adaptation, ecological speciation, species interactions, co-evolution, and genetic variation. In order to elucidate mechanisms of adaptation, the ultimate goal of many of these studies is to identify the actual genetic targets of selection. Identifying the actual targets of selection is necessary because not all instances of phenotypic evolution are a reflection of adaptation (Gould *et al.* 1979), but might reflect changes resulting from linkage (hitch-hiking) or pleiotropic effects (Cork & Purugganan 2004). Identifying genetic targets of selection is inherently difficult, as specific traits are controlled by relatively tiny regions within vast genomes. In addition, adaptation might actually involve selection at the genome-

wide scale, when most studies have focused on effects of one or a few genes.

Genomic approaches, especially used in combination, are particularly useful for identifying genetic targets of selection (and therefore genetic mechanisms of adaptation). For example, comparative genomic expression profiles, such as those obtained using cDNA microarrays, are a potential first step for identifying an initial pool of candidate genes that might be implicated in adaptation. This approach is particularly useful in non-model systems for which whole genome sequences are unavailable (Gracey & Cossins 2003). However, only a subset of genes that show evolutionary shifts in gene expression will constitute the actual targets of selection, as many of the shifts will have resulted from *trans*-regulatory changes (i.e. changes elsewhere in the genome), and some will have evolved due to linkage (hitch-hiking). Thus, gene expression profiles could be used in conjunction with methods that link functional traits with candidate genes, such as association and QTL-mapping studies.

The first part of this Special Issue includes seven studies that employ cDNA microarray (or macroarray) expression profiles to address questions regarding sources of variation (Whitehead & Crawford 2006), ecological speciation and local adaptation (Derome *et al.* 2006; Knight *et al.* 2006; Lai *et al.* 2006), and species interactions, such as symbiosis (Moran & Degan 2006), parasitism (Wang *et al.* 2006), and herbivory (Ralph *et al.* 2006). Whitehead and Crawford's paper is a comprehensive review of the extent and sources of variation in gene expression within and among taxa uncovered by microarray analysis, and an examination of evolutionary processes that account for that variation. They find that microarray-based measures of gene expression are precise given appropriate experimental design. Most notably, they find a nonlinear relationship between expression differences and evolutionary divergence among taxa due to functional constraint (Whitehead & Crawford 2006).

Evolutionary shifts in gene expression profiles could be used to explore genetic targets involved in local adaptation and ecological speciation. Lai and colleagues describe the

creation of a custom cDNA microarray for annual sunflowers and its use to elucidate patterns of gene expression in *Helianthus annuus*, *Helianthus petiolaris*, and the homoploid hybrid species *Helianthus deserticola*. Results show that transport-related proteins are over-represented among differentially regulated genes, perhaps indicating an important role in ecological divergence of the hybrid species (Lai *et al.* 2006). Knight and colleagues examine gene expression and the genetics of local adaptation for drought tolerance in a wild relative of *Arabidopsis thaliana*, *Boechera holboellii*. They find genetic differentiation for transpirational water loss, with greater water-use efficiency in drought-adapted genotypes. In addition, cDNA-AFLPs are used to identify candidate genes that are differentially expressed in response to plant water stress (Knight *et al.* 2006). Insights into evolutionary constraints could be greatly enhanced by examining parallelisms in genetic mechanisms across independent evolutionary events (Hoekstra & Price 2004). Derome and colleagues test the hypothesis that parallel phenotypic divergence in two lakes between sympatric dwarf and normal whitefish (*Coregonus clupeaformis*) ecotypes is accompanied by parallel divergence in gene expression. Consistent with phenotypic differences between these forms, they find parallelism in differential expression of genes belonging to functional groups involved in energy metabolism and regulation of muscle contraction (Derome *et al.* 2006).

Gene expression profiles could also be used to probe evolutionary adaptations that arise from species interactions. Although symbionts have historically been experimentally intractable because they cannot be cultivated, complete genome sequences of some species have recently made them amenable to comparative gene expression analysis. Moran and Degnan's review discusses the functional genomics of the symbiont *Buchnera*. Expression studies reveal that the *Buchnera* genome has mostly lost regulatory mechanisms, and consists essentially of an abbreviated *E. coli* genome that can function as an obligate symbiont of insects (Moran, Degnan 2006). Wang and colleagues apply a subtractive cDNA macroarray analysis as a first step to examine *Mycoplasma* parasite-induced gene expression changes in the house finch (*Carpodacus mexicanus*). The emerging technology of cDNA macroarray hybridization offers the possibility for providing a rapid, high-throughput, and relatively inexpensive method to screen a cDNA library (Wang *et al.* 2006). Ralph and coworkers report large-scale studies of gene expression in hybrid poplar (*Populus trichocarpa* × *deltoides*) in response to herbivory by forest tent caterpillars (*Malacosoma disstria*). Hundreds of up- and down-regulated genes were identified, including known defensive loci, as well as new candidate genes that may improve understanding of plant defence against insect herbivores (Ralph *et al.* 2006).

The second part of this issue includes studies that apply gene-mapping approaches, such as quantitative trait locus (QTL) mapping, to link functional traits with regions of the genome (QTL). As genetic maps do not exist for most species, to explore the genetics of fitness-related traits in wild passerine birds Dawson and colleagues produce a linkage map of 196 microsatellites by BLASTing flanking sequence against the recently assembled chicken genome (Dawson *et al.* 2006). Garner and Latta use QTL mapping in *Avena barbata* to examine the role of genotype–environment interaction ($G \times E$) in local adaptation. Measuring either phenotypic traits or overall fitness, there is little evidence that natural selection favours alternate QTL alleles in different environments, such that $G \times E$ is unlikely to contribute to local adaptation among the field environments (Gardner & Latta 2006). Doust and Kellogg examine plant morphology and branching in a QTL-mapping cross between weedy green millet (*Setaria viridis*) and domesticated foxtail millet (*Setaria italica*) grasses. Likely candidate genes underlying the QTL include *teosinte branched1* and *barren stalk*. This study illustrates the use of comparative genomic techniques to leverage genetic information from model plants (such as maize and rice) to address evolutionary and ecological questions in related non-model species (Doust & Kellogg 2006). As stated above, gene expression analyses could be performed in conjunction with a QTL approach, to determine whether differentially expressed genes and functional traits map to the same QTL. Juenger and colleagues combine whole-genome expression profiling with near-isogenic line mapping (i.e. expression QTL mapping) in order to examine genetic differences that may contribute to water-use efficiency in *A. thaliana*. They find 25 differentially expressed genes in a chromosomal region of a physiological QTL that controls increased transpiration and reduced water-use efficiency at the whole-plant level (Juenger *et al.* 2006).

Once a narrow set of candidate genes are identified, further testing could be performed using transgenic or knockout approaches (e.g. RNA interference) to examine effects of genes or allelic variants on phenotype. Such approaches have not been widely applied to non-model systems. Another useful genomic approach that is not represented in this volume includes genome-wide scans for signatures of selection, such as selective sweeps (Kim & Nielsen 2004; Nielsen *et al.* 2005). Genomic scans for selection have been applied primarily to humans and *Drosophila* (Glinka *et al.* 2003; Nielsen *et al.* 2005). The limitation of applying this technique to non-model systems is the need for a large number of genetic markers spanning the entire genome. Such genome-wide molecular markers are also essential for other genomic techniques, such as gene-mapping approaches, but are still relatively difficult to obtain without a whole-genome database.

In the final part of the Special Issue, four studies address the topic of genetic variation, including methods for rapid marker (SNP) discovery (Gilchrist *et al.* 2006), effects of genetic variation on community and ecosystem processes (Bangert *et al.* 2006), links between genes and landscape in the wild (Velez & Feder 2006), and variation at functional and neutral loci across a broad geographic range (Bakker *et al.* 2006). These studies constitute important steps towards understanding ecologically relevant polymorphism in natural populations.

Gilchrist and colleagues employ ecotilling to identify single nucleotide polymorphisms (SNPs) in populations of the black cottonwood, *Populus trichocarpa*. This technology allows rapid screening for nucleotide polymorphisms, which can be used for a myriad of genomic techniques (e.g. QTL mapping, genomic scans for selection) and subsequent analyses of molecular diversity (Gilchrist *et al.* 2006). Using a community and ecosystem approach, Bangert and colleagues examine genetic variation in two species of cottonwood trees from the genus *Populus*, which are keystone species influencing arthropod communities in western riparian habitats. They find that genetic variation in poplar provides significant prediction of the community structure of associated arthropods. More generally, polymorphisms in dominant plant species may influence a range of ecological processes at the community and ecosystem level (Bangert *et al.* 2006). Velez and Feder examine a striking colour polymorphism in the luciferase-encoding locus of the Jamaican click beetle, *Pyrophorus plagiophthalmus*, and the evolutionary shift towards longer wavelength. Historical biogeography has played an important role in the modern distribution of colour variants, as well as possible sexual selection or reproductive character displacement (Velez & Feder 2006). Bakker and colleagues examine patterns of genetic variation in *A. thaliana* found in populations from its native range in Eurasia, as well as introduced populations in North America. Polymorphic loci include two pathogen resistance genes as well as other, presumably neutral markers. Substantial levels of genetic variation were found within European populations, with patterns suggesting that most dispersal between populations may be *via* pollen rather than seeds (Bakker *et al.* 2006).

The emerging synthesis of ecological genetics with functional genomics will give us the capacity to address fundamental questions in ecology and evolution at an increasingly mechanistic level, and provide information on interactions between organisms and their environment on a whole organism (and genomic) scale. Mechanistic information from functional genomics aids in illuminating the genetic targets of selection in response to the environment. Selection could act on many functions spanning the genome, and also have cascading effects on a gene network, with subsequent impacts on phenotypic evolution.

Information on genetic mechanisms of adaptation in a system could be extended across taxa, providing hypotheses and predictions on the responses of other organisms to environmental variables. In addition, genomic architecture could have cascading effects on a community or ecosystem, such that natural genomic variation within and among populations could have both profound ecological and evolutionary consequences. Such consequences remain largely unexplored, although current research should bring us closer to addressing such questions in the near future.

References

- Bakker EG, Stahl EA, Toomajian C *et al.* (2006) Distribution of genetic variation within and among local populations of *Arabidopsis thaliana* over its species-range. *Molecular Ecology*, **15**, 1405–1418.
- Bangert RK, Turek RJ, Rehill B *et al.* (2006) A genetic similarity rule determines arthropod community structure. *Molecular Ecology*, **15**, 1379–1391.
- Cork JM, Purugganan MD (2004) The evolution of molecular genetic pathways and networks. *Bioessays*, **26**, 479–484.
- Dawson DA, Burke T, Hansson B *et al.* (2006) A predicted microsatellite map of the passerine genome based on chicken–passerine sequence similarity. *Molecular Ecology*, **15**, 1299–1320.
- Derome N, Duchesne P, Bernatchez L (2006) Parallelism in gene transcription among sympatric lake whitefish (*Coregonus clupeaformis* Mitchell) ecotypes. *Molecular Ecology*, **15**, 1239–1249.
- Doust AN, Kellogg EA (2006) Effect of genotype and environment on branching in weedy green millet (*Setaria viridis*) and domesticated foxtail millet (*Setaria italica*) (Poaceae). *Molecular Ecology*, **15**, 1335–1349.
- Gardner KM, Latta RG (2006) Identifying loci under selection across contrasting environments in *Avena barbata* using quantitative trait locus mapping. *Molecular Ecology*, **15**, 1321–1333.
- Gilchrist EJ, Haughn GW, Ying CC *et al.* (2006) Use of Ecotilling as an efficient SNP discovery tool to survey genetic variation in wild populations of *Populus trichocarpa*. *Molecular Ecology*, **15**, 1367–1378.
- Glinka S, Ometto L, Mousset S, Stephan W, De Lorenzo D (2003) Demography and natural selection have shaped genetic variation in *Drosophila melanogaster*: a multi-locus approach. *Genetics*, **165**, 1269–1278.
- Gould SJ, Lewontin RC (1979) The spandrels of San Marco and the Panglossian Paradigm: a critique of the adaptationist programme. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **205**, 581–598.
- Gracey AY, Cossins AR (2003) Application of microarray technology in environmental and comparative physiology. *Annual Review of Physiology*, **65**, 231–259.
- Hoekstra HE, Price T (2004) Parallel evolution is in the genes. *Science*, **303**, 1779–1781.
- Juenger TE, Wayne T, Boles S *et al.* (2006) Natural genetic variation in whole-genome expression in *Arabidopsis thaliana*: the impact of physiological QTL introgression. *Molecular Ecology*, **15**, 1351–1365.

- Kim Y, Nielsen R (2004) Linkage disequilibrium as a signature of selective sweeps. *Genetics*, **167**, 1513–1524.
- Knight CA, Vogel H, Kroymann J *et al.* (2006) Expression profiling and local adaptation of *Boechera holboellii* populations for water use efficiency across a naturally occurring water stress gradient. *Molecular Ecology*, **15**, 1229–1237.
- Lai Z, Gross BL, Zou Y, Andrews J, Rieseberg LH (2006) Microarray analysis reveals differential gene expression in hybrid sunflower species. *Molecular Ecology*, **15**, 1213–1227.
- Moran NA, Degnan PH (2006) Functional genomics of *Buchnera* and the ecology of aphid hosts. *Molecular Ecology*, **15**, 1251–1261.
- Nielsen R, Williamson S, Kim Y *et al.* (2005) Genomic scans for selective sweeps using SNP data. *Genome Research*, **15**, 1566–1575.
- Ralph S, Oddy C, Cooper D *et al.* (2006) Genomics of hybrid poplar (*Populus trichocarpa* × *deltoides*) interacting with forest tent caterpillars (*Malacosoma disstria*): normalized and full-length cDNA libraries, expressed sequence tags, and a cDNA microarray for the study of insect-induced defences in poplar. *Molecular Ecology*, **15**, 1275–1297.
- Velez S, Feder JL (2006) Integrating biogeographic and genetic approaches to investigate the history of bioluminescent colour alleles in the Jamaican click beetle, *Pyrophorus plagiophthalmus*. *Molecular Ecology*, **15**, 1393–1404.
- Wang ZS, Farmer K, Hill GE, Edwards SV (2006) A cDNA macroarray approach to parasite-induced gene expression changes in a songbird host: genetic response of house finches to experimental infection by *Mycoplasma gallisepticum*. *Molecular Ecology*, **15**, 1263–1273.
- Whitehead A, Crawford DL (2006) Variation within and among species in gene expression: raw material for evolution. *Molecular Ecology*, **15**, 1197–1211.