Journal of Ecology 2007
95, 8–16

ESSAY REVIEW

An ecologist's guide to ecogenomics

N. J. OUBORG and W. H. VRIEZEN*

*Department of Molecular Ecology, Radboud University Nijmegen, Toernooiveld 1, 6526ED Nijmegen, the Netherlands and *Department of Plant Cell Biology, Radboud University Nijmegen, Toernooiveld 1, 6526ED Nijmegen, the Netherlands*

Summary

1 Currently, plant ecologists are increasingly adopting approaches and techniques from molecular biology. The new field of ecogenomics aims at understanding the mechanistic basis for adaptation and phenotypic variation by using genomic techniques to investigate the mechanistic and evolutionary basis of species interactions, and focuses on identifying the genes affected by evolution.

2 While the entire toolbox of genomics is only available for model species such as *Arabidopsis thaliana*, we describe the options open to ecologists interested in pursuing an ecogenomics research program on ecologically relevant traits or phenomena in non-model species, for which part of the genomic toolbox may be currently unavailable. In these non-model species, a viable ecogenomics research program is possible with relatively modest effort.

3 Four challenges to further development of ecogenomics are described and discussed: (i) the ecogenomic study of non-model species; (ii) reconciliation of experimental languages of ecology and evolutionary biology with molecular biology; (iii) development of specific ecogenomic data analysis tool; and (iv) adoption of a multidisciplinary cooperative research culture.

4 An important task for ecologists is to provide the necessary ecological input (the 'eco' part) to ecogenomics.

Key-words: adoptation, evolutionary ecology, gene expression, microarray, sequence analysis

Journal of Ecology (2007) **95**, 8–16
doi: 10.1111/j.1365-2745.2006.01197.x

Introduction

Modern plant ecology is a highly multidisciplinary science, with input from fields including biochemistry, biophysics, mathematics and statistics, physiology and systems biology. From the early 1990s, plant ecologists increasingly adopted approaches and techniques from molecular biology. The use of molecular markers to target the variability of the DNA of individuals has become an inextricable part of various subdisciplines of plant ecology, especially plant population biology. Most of the marker techniques available so far, including microsatellite-PCR, AFLP and ISSR (see Glossary for definition of terms), aim to find out the variability in non-coding, and therefore supposedly selectively neutral parts of the genome. The integration of these techniques has advanced plant ecology far beyond what was possible before the introduction of these methods. Molecular markers are now routinely used in the measurement of seed and pollen dispersal (Ouborg *et al*. 1999), the study of breeding systems (e.g. Touzet *et al*. 2004), the assessment of genetic diversity between and within plant populations (e.g. Eckstein *et al*. 2006), the identification of species (viz. DNA barcoding, e.g. Chase *et al*. 2005), the elucidation of phylogenetic relationships between plant species (e.g. Soltis *et al*. 2004), the study of kinship (e.g. Slavov *et al*. 2005) and the survey of phylogeographical patterns (e.g. Magri *et al*. 2006).

One area where molecular markers have had far less impact is in the study of adaptation. As markers are supposed to be selectively neutral (perhaps with the exception of those markers that are closely genetically linked to functional genes, such as some single nucleotide polymorphisms (SNPs), their variation by definition says little about the effects of selection. The desire to search for markers for traits under selection has frequently been stated (e.g. Van Tienderen *et al*. 2002). We are now at the start of a second wave of molecular biological approaches entering plant ecology. Increasingly,

Correspondance: N.J. Ouborg (tel. + 31243652470; e-mail: j.ouborg@science.ru.nl).

© 2006 The Authors
Journal compilation
© 2006 British
Ecological Society

techniques and approaches originating from genomic research are being incorporated in to plant ecology, so that the focus is shifting from measuring temporal and spatial demographic processes by using neutral markers, to a detailed understanding of the function and variation of functional genes in an ecological context. This approach, which has been named 'eco-genomics' (Feder & Mitchell-Olds 2003), integrates the disciplines of ecology and molecular biology. Ecology is the study of the interactions between individuals, populations and communities and their biotic and abiotic environment. The realm of molecular biology lies at the subcellular organizational level, where the functioning of the genome of organisms is studied. Ecogenomics integrates these two disciplines by trying to capture the interaction between environment and phenotype, using mechanistic, genomic processes as the basis for the explanation.

The aims of ecogenomics

The first aim of ecogenomics is to elucidate the mechanisms of adaptation. Ecological and evolutionary research frequently demonstrates that species, populations and individuals are adapted to their own environment, i.e. the fitness of individuals is higher in their home environment than in other environments. Plant ecology, especially, has a longstanding tradition of studying this adaptation, as individual plants have no other option than to cope with the environment they grow in. Evolutionary biology studies adaptation from the perspective of the balance between selection and other evolutionary forces (more specifically, drift and migration) that may lead to adaptation. The incorporation of genomic tools now allows us to open the black box between environment and phenotype, and to try to understand the mechanisms behind adaptation.

The second important target of ecogenomics is to investigate the mechanistic causes of phenotypic variation. This goal aligns with the first: whereas the first goal focuses on the causes of fitness variation, the second goal concentrates on the causes of phenotypic variation. Together these two objectives will deliver the basis for a mechanistic understanding of evolution.

A derived third goal is to identify the genetic targets of selection. Selection does act on phenotypes, but the evolutionary consequences occur at the level of genes. Incorporating genomic tools may allow identification of the genes that are actually evolving (Lee & Mitchell-Olds 2006).

Fourth, understanding the mechanisms that lead from environmental cues to phenotypic response will create the potential to understand more fully complex evolutionary phenomena, such as the origin of genetic trade-offs and the basis of phenotypic plasticity.

Finally, the genomic toolbox offers completely new ways to design ecological experiments (Jackson *et al.* 2002). Not only does the toolbox allow ecological and evolutionary processes to be examined in unprecedented molecular detail, the tools also allow new ways of experimental manipulation. See Snoeren *et al.* (2006) for an overview of this use of ecogenomics in the study of multi-species interactions.

The ecogenomic toolbox

The number of tools that genomics research offers ecology is large and still expanding. It is impossible to fully describe all existing tools in detail here. Overviews can be found elsewhere (Gibson 2002; Thomas & Klaper 2004; Seifert *et al.* 2005; Ranz & Machado 2006; Van Straalen & Roelofs 2006). Three broad categories of techniques can be distinguished: (i) methods concerned with sequence variation; (ii) methods concerned with the expression of genes; and (iii) methods trying to couple genes to phenotype.

(I) SEQUENCE VARIATION

First, there are the techniques that are directed at describing DNA-sequences. DNA-sequencing, i.e. determining the order of bases in stretches of a DNA molecule, is already a fully incorporated technique in ecology, used in, for instance, species identification and phylogenetic and phylogeographical inferences. New techniques are available that target sequence variation in genes or gene-promoter regions. Most prominent is the use of SNPs, that form excellent markers for sequence variation in coding regions (Lee *et al.* 2004).

(II) GENE EXPRESSION ANALYSIS (TRANSCRIPT PROFILING)

The second category of techniques is directed at measuring the transcriptional expression of genes as a function of environment, genotype, internal location in the organism, developmental stage or time. Various techniques are available that differ in the amount of a priori knowledge required, cost and in the number of samples that can be processed per unit time.

Most methods to determine gene expression are based on measurement of the concentration of gene transcripts in a particular tissue at a particular moment. Differences in transcript levels between samples are used to identify genes that are involved in a certain process that plays a role in adaptive behaviour. When considerable sequence information is available, either through genome or cDNA (expressed sequence tags (ESTs)) sequencing, microarrays can be made to assess global patterns of gene expression variation (transcriptome analysis). Alternatively, cDNA-AFLP (Amplified Fragment Length Polymorphisms), a PCR-based screening method, can be used for transcriptome analysis if no sequence data are available.

Microarrays are glass or quartz surfaces covered with minute samples of DNA and are fabricated by high-speed robotics. There are two main types of microarray technology that differ in the property of the DNA (see below).

© 2006 The Authors
Journal compilation
© 2006 British Ecological Society,

cDNA microarray (DNA chips)

In this type, samples of probe DNA, which can be PCR product, cDNA or genomic DNA are immobilized on a glass surface and hybridized to a set of targets. These targets are synthesized from mRNA and the resulting cDNA is labelled with a fluorescent dye (Schena *et al*. 1995). Typically, one RNA sample is labelled with green Cyanine 3-dNTP (Cy3) and the other with red Cyanine 5-dNTP (Cy5). The two labelled RNA populations are hybridized to one glass slide and scanned using a fluorescent imager. Differences in mRNA concentration between the two samples will become evident as different colours of the spot, on a scale from green (Cy3 > Cy5) via yellow (Cy3 = Cy5) to red (Cy5 > Cy3). These colour differences indicate up- or down-regulation of the particular gene in one sample compared to the other.

Oligonucleotide microarray

The second type of microarrays are oligonucleotide chips, which are glass slide arrays that are manufactured using a photolithographic technique and combinatorial chemistry, allowing oligonucleotide DNA samples to be synthesized directly onto a quartz wafer. The oligonucleotide probes range in length from 25 to 60 or more nucleotides, and 'standard' chips for a number of organisms (including *Arabidopsis thaliana*, barley, rice, soybean and tomato) are now commercially available (e.g. Affymetrix, Agilent and Amersham). Oligonucleotide chips are designed to reject targets that are not identical. This level of specificity is essential when measuring the expression of two very similar gene sequences. Oligonucleotide chips are more sensitive and in general give better gene-hybridization signals than cDNA chips. However, the synthesis of these microarrays requires knowledge of the DNA sequences to be spotted. Sequencing projects of plant genomes and expressed sequence tags (ESTs) are therefore the basis of the production of oligonucleotide chips. In addition, these types of chips have to be manufactured by a specialized company whereas cDNA microarrays are often made in-house.

cDNA-AFLP

An alternative method for genome-wide expression analysis is cDNA-AFLP, which does not require DNA sequence data, and has been used successfully in many different organisms (Breyne & Zabeau 2001; examples are: *Petunia* spp.). (Cnudde *et al*. 2003), *Sorghum* (Pring & Tang 2004) and cucumber (Bae *et al*. 2006). cDNA-AFLP is a PCR-based method to accurately determine gene expression profiles by quantitative analysis of band intensities (Breyne *et al*. 2003; Fig. 1). Furthermore, the sensitivity and specificity of the method allows the detection of poorly expressed genes and the determination of subtle differences in transcriptional activity. cDNA-AFLP starts with cDNA synthesis from mRNA. The cDNA is then digested with two different restriction enzymes, and adapters are ligated to the ends of the fragments. Only fragments that are digested by both restriction enzymes and thus have different adapters at the end are amplified in a PCR. In the following amplification steps, a mixture of cDNAs is fractionated into smaller subsets by selective PCR amplification using primers based on the adapters that contain one or more extra selective nucleotides. One of the selective primers is radioactively labelled, which makes it possible to visualize the fragments on a high-resolution gel (Fig. 1). Fragments with interesting intensity patterns can be cut from the gel and the DNA can be isolated and sequenced.

Thus, microarray and cDNA-AFLP are both methods to identify genes based on global expression patterns. Another technique, Subtractive Hybridization PCR (SSH PCR), is a cDNA subtraction method that is based on the enrichment of genes specifically expressed in one sample and the exclusion of genes expressed in both sample tissues that are to be compared. This method is therefore not suitable for the identification of genes that are (slightly) modulated in expression, but can be valuable for the isolation of genes that are activated in a certain situation (Siebert *et al*. 1995).

Both cDNA-AFLP and SSH often lead to the identification of many genes that might have a function in the studied process and are based on expression differences of few samples. The same process can now be studied under other conditions or in other species/genotypes using the identified set of genes. This can be achieved by generating 'dedicated' microarrays, containing only the candidate ESTs from the cDNA-AFLP or SSH procedure, making possible fast analysis of all process-specific genes in other plants.

qRT-PCR

Differences in expression of specific sequences are often validated by another method of analysis such as reverse transcriptase (RT)-PCR. qRT-PCR stands for quantitative Real-Time Reverse Transcription PCR and is an advancement of the basic PCR technique. Through the use of a fluorescent detection system, the starting amount of nucleic acid in the reaction can be quantified (Heid *et al*. 1996). Quantification is achieved by measuring the increase in fluorescence during the exponential phase of PCR. In general, the results of microarrays and qRT-PCR correlate strongly (although notable exceptions frequently occur) (Dallas *et al*. 2005). qRT-PCR is less technically demanding but can analyse only relatively small numbers of genes per unit time, as compared with a microarray. However, new techniques are becoming available, such as microfluidic cards that contain 384 miniaturized, real-time PCR assays (www.appliedbiosystems.com), which considerably increase the throughput capacity of qRT-PCR.

(III) GENE FUNCTION ANALYSIS

The third category of techniques aims to identify gene function. Once genes of interest have been identified

© 2006 The Authors
Journal compilation
© 2006 British
Ecological Society,

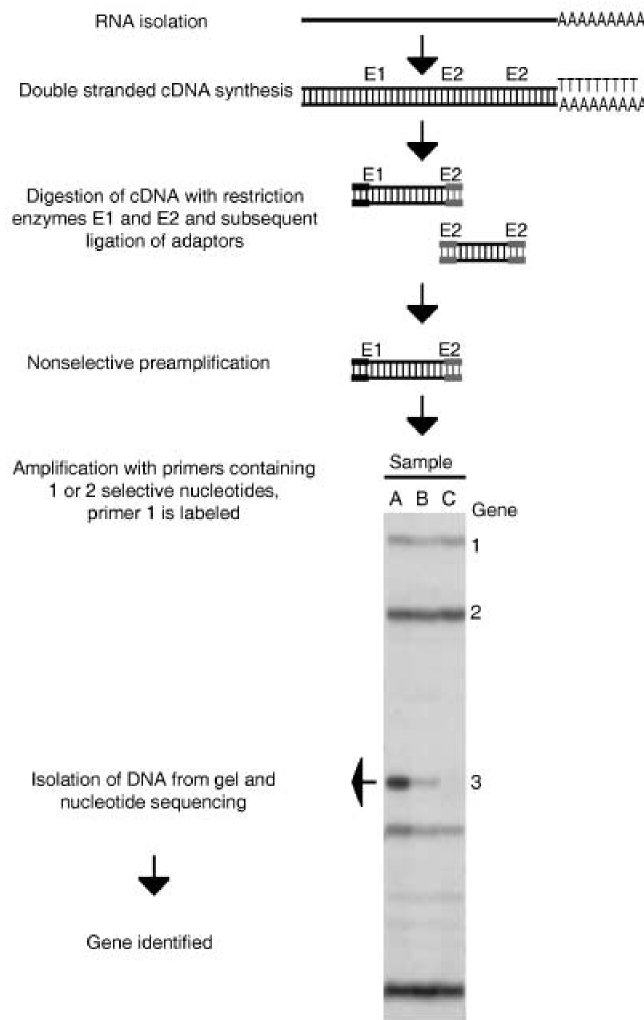

Fig. 1 Scheme describing the cDNA-AFLP procedure. cDNA is synthesized from mRNA and unique restriction fragments are processed from the different transcripts. In this example, quantification of the intensity of the obtained fragments demonstrate that the transcript level of gene 3 is much higher in sample A than in B or C. The fragment can be isolated from the gel and its nucleotide sequence leads to the corresponding gene. E1 and E2 refer to the cleavage sites of two different restriction enzymes.

their function often still has to be determined. Genes can have low sequence homology to other known genes. But even if there is considerable homology with a known gene, its function in the process of interest might be unclear. Modification of the expression of such a plant gene can lead to a changed phenotype and shed more light on its function. This can be achieved by several different approaches; a gene can be disrupted (mutation) so that it can not produce a functional protein, its regulation can be modulated (silencing/activation), or a new gene can be added to the plant (modification).

Mutations can be introduced in various ways: chemically, with radiation, or biologically with transposons; different methods leading to different types of mutation with different efficiencies (reviewed by Alonso & Ecker 2006). For instance, EMS (EthylMethane Sulphonate) is a mutagen that introduces single basepair mutations. These SNPs in a target gene can be identified by TILLING (Targeted Induced Local Lesions in Genomes; Till *et al*. 2003). The technique can be adapted for use in a high-throughput facility, and has been successfully applied to several plant species, such as *Arabidopsis thaliana* and *Brassica oleracea* (Colbert *et al*. 2001). TILLING is also used for detecting multiple types of natural polymorphisms in natural populations, a variant called Ecotilling. Ecotilling can be used to effectively discover natural polymorphisms in a large set of individuals of any plant species (Comai *et al*. 2004).

Instead of creating a mutant, the function of genes can also be studied in a procedure called gene silencing. One way to efficiently create gene knockout phenotypes is a system for virus-induced gene silencing (VIGS; Robertson 2004). VIGS is a form of RNA-mediated gene silencing, which involves the production of large amounts of dsRNA that activates a host defence mechanism resulting in the degradation of mRNAs with homology to the sequences within the dsRNA. In VIGS, certain RNA viruses are used to produce dsRNAs that trigger the silencing mechanism. The virus is engineered to contain sequences from a plant gene of interest,

© 2006 The Authors
Journal compilation
© 2006 British Ecological Society,

resulting in the degradation of the mRNA from that gene as well. This gene-silencing mechanism, RNA interference (RNAi), also known as post-transcriptional gene silencing (PTGS) or co-suppression (Fire *et al*. 1998; Waterhouse & Wang 2001) seems to have evolved as a defence mechanism against viruses and endogenous transposons (Robertson 2004).

The third method to study gene function is to introduce foreign DNA into the genome of the plant species. This can be achieved by different methods. Most common is the use of *Agrobacterium tumefaciens* to introduce a vector carrying a DNA-construct that either can silence or activate the gene of interest. *Agrobacterium tumefaciens* also mediates the transfer of DNA to the plant for VIGS and RNAi. Unfortunately, some species are resistant to *A. tumefaciens* and other plant-associated bacteria such as *Rhizobium* spp. which might severely hamper the generation of transgenic plants. The effect of genetically modified gene expression in a plant on its phenotype can best be tested under natural conditions in the field. A major problem with GMOs is, however, that it is unknown what will happen when they escape from test fields. As it cannot be excluded that they might outcompete wild organisms or interbreed with wild relatives and produce hybrids, transgenic plants can only be tested under controlled conditions. Moreover, many traits are not defined by only one gene and are linked to many genes on different loci.

A method that aims to associate phenotypic traits with regions in the DNA is Quantitative Trait Locus (QTL) mapping. QTL mapping is a way to search for correlations between a particular molecular marker and a particular 'quantitative' trait. A QTL is defined as a genetic locus where functionally different alleles segregate and cause significant effects on a quantitative trait. The results of QTL mapping studies together with the information about a specific trait in different environments can lead to the identification of genes that contribute to natural variation. For species from which the genome has been sequenced, it is possible to locate a QTL on the sequence and then directly look for candidate genes (Price 2006).

The ecologist's options

Ecological research focuses on ecologically interesting species in their natural environment. The complete set of tools described above is only available for a limited number of genomic model species, with *Arabidopsis thaliana* being the best known example. These species have become model species of genomics because of their suitable properties: small genomes, short generation times and easy maintenance in the glasshouse. They are not chosen because of their specific ecology, and they certainly do not cover all life-history strategies and habitats. The question then is: if one is interested in pursuing an ecogenomic approach on an ecological phenomenon that is not covered by a genomic model species, what are the options?

One answer to this is to use close relatives of the model species, where the tools might still be applicable. Many follow this option by focusing on close relatives of *Arabidopsis thaliana*, such as *A. lyrata* (e.g. Heidel *et al*. 2006) and *A. halleri* (e.g. Filatov *et al*. 2006).

The other answer is to use the ecological phenomenon as the central point of departure, and work with whatever tools happen to be available. This second strategy is gaining popularity. Tools for an increasing number of non-model species are becoming available. A quick and incomplete screening of the literature shows that ecogenomic work, using at least one of the described tools, has been performed in legumes (Domoney *et al*. 2006), *Brassica rapa* (e.g. Suwabe *et al*. 2006), *Thlaspi caerulescens* (e.g. Rigola *et al*. 2006), *Triticum aestivum* (e.g. Endo *et al*. 2006; Ogihara 2006), *Populus* spp. (e.g. Sterky *et al*. 2004; Cronk 2005; Gilchrist *et al*. 2006), *Senecio* spp. (Hegarty *et al*. 2005), *Boechera holboellii* (Knight *et al*. 2006) and *Solanum nigrum* (Mueller *et al*. 2005). Web-resources for an increasing number of species are available (e.g. *Arabidopsis*: http://www.arabidopsis.org/; *Populus* http://www.ornl.gov/sci/ipgc/; *Medicago* http://www.noble.org/medicago/index.html; and *Mimulus* http://www.genome.clemson.edu/projects/mimulus/; an overview of plant resources is available at http://www.tigr.org/plantProjects.shtml). Whole-genome sequencing projects, which form the basis for the development of most other tools, are on their way for an increasing number of plant species.

But what if one is interested in a 'new' species for which no specific genomic tools are yet available? The 'problem' with non-model species and an ecogenomic approach is that no, or very little, sequence data are available. Thus, either an extensive sequence program must be started, which may in most cases be unrealistic given time and money constraints, and the large genome sizes involved, or ecogenomic research is started without detailed sequence data being available at the start. Although obtaining extensive sequence data will be very important in any ecogenomic project, starting without sequence data being available does represent a viable tactic for ecologists.

Two options are available, that may complement each other (Fig. 2). The first option (option 1; Fig. 2) is to use cDNA-AFLP (or SSH) to identify genes that are differentially expressed in two well-defined (sets of) samples that are differentiated in the phenotypic trait or environmental factor under study. The resulting differentially expressed ESTs can be designated as candidates for regulation of the phenotypic phenomenon under study, and can be further studied in several ways. They can be spotted on a dedicated microarray (cDNA chip), which can then be used to screen larger numbers of samples (option 1A; Fig. 2). Alternatively, the ESTs can be sequenced. Typically in the order of a few hundred ESTs emerge from the procedure, and these can be sequenced relatively quickly. qRT-PCR primers for the most interesting ones can be designed and samples can be screened with these primers in a qRT-PCR procedure (option 1B; Fig. 2).

© 2006 The Authors
Journal compilation
© 2006 British
Ecological Society,

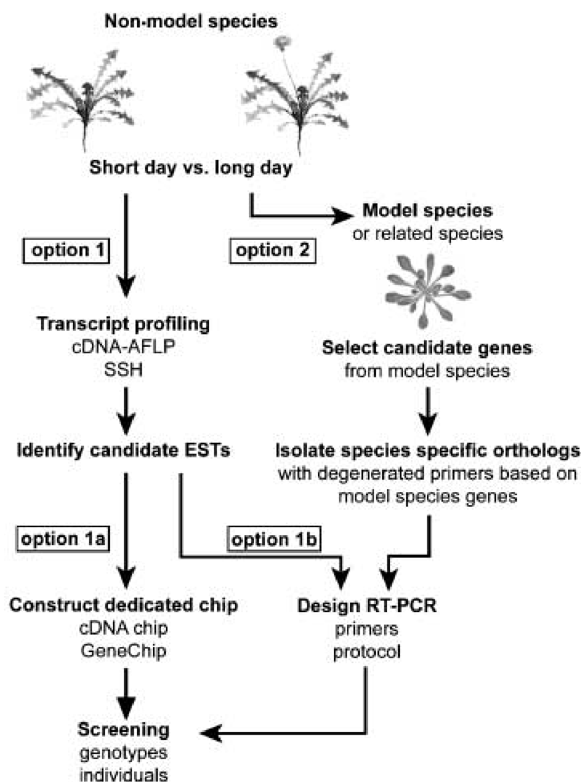

Fig. 2 A schematic outline of the options open for an ecogenomic approach to the study of non-model plant species. Starting point is two (sets of) phenotypes, in this example as response to day length (short vs. long). Options are further explained in the text.

Another option (option 2; Fig. 2) is to still make use of information from model species. Although sequences soon diverge between distantly related species to the extent that tools are no longer transferable from one species to another, there is a much higher possibility of using information on gene function between species. There is ample evidence that gene function is relatively conserved, and therefore may be transferred from one species to another. For instance, although the sequence variation in the LEAFY gene is considerable over plant species, it seems to function as a regulator of onset of flowering in all species (e.g. Bomblies *et al*. 2003; Archambault & Bruneau 2004). This makes it possible to use degenerated PCR primers to find the species specific version of the gene, which then can be used in species specific RT-PCR trials.

Thus, even though not all tools are be available for the species of interest, several options are available to start a viable ecogenomic research program.

Challenges for ecogenomics

Even though the field of ecogenomics has great promise, like any other new field it is facing several challenges.

The first challenge is to extend the ecogenomic approach to non-model species in natural environments, to cover a wide variety of ecological conditions and life histories. Even though in the previous section we have mentioned a number of possibilities for taking up this challenge, working along these lines is only just beginning and certainly needs much more input. For plants, an additional challenge is presented by their large genome sizes, and the fact that many plant species are polyploid. This hampers fast acquisition of genome-wide sequence information.

The second challenge is to reconcile the experimental languages of ecology and evolutionary biology with molecular biology. These disciplines come from different traditions. Molecular biologists prefer to work in controlled environments and with homogeneous well-defined genetic material, aiming to remove as much variation as possible. For ecologists, environmental and genotypic variation is their core business, which they try to incorporate in experimental designs rather than controlling for it. An example of the discrepancy between ecological and genomic experimental designs comes from research on the effects of elevated CO_2 on the functioning *Populus* trees. Liu *et al*. (2006) examined variation in ecophysiological responses to CO_2 elevation. They used three CO_2 levels, offspring of three different crosses, and a total of 324 biological replicates. Taylor *et al*. (2005) investigated the effect of elevated CO_2 on the expression of *Populus* genes. They used two CO_2 levels, one genetic clone, and a total of 12 biological replicates. Thus, the level of detail differs by an order of magnitude between these two studies. This makes it difficult to integrate the ecophysiological and transcriptomic data. The challenge to ecogenomics is to fine-tune

© 2006 The Authors
Journal compilation
© 2006 British
Ecological Society,

experimental designs of ecology and molecular biology in order to accomplish true integration of the data that originate from these two fields.

The third challenge lies in how to deal with the ecogenomic data. Using high-throughput tools, such as microarrays, in an ecological design would lead to the expression data of thousands of genes for each of hundreds of individuals. Various sources of variance will influence the data, including methodological, factorial and error sources. Methods will have to be developed to analyse such data sets, and to extract most efficiently and reliably the ecogenomic information contained therein.

The fourth challenge lies in the adoption of a new research culture appropriate for the ecogenomic approach. Ecogenomics is a multidisciplinary discipline and requires close collaboration between ecologists, molecular biologists and bioinformaticians, complemented with – depending on the actual question – input from physiologists, microbiologists, geneticists and evolutionary biologists. As becomes evident from the author list of papers on ecogenomics (e.g. Sterky *et al*. 2004; Domoney *et al*. 2006), international cooperation is also required, both at the levels of expertise and research infrastructure. This requires an open and outreaching attitude towards cooperation and authorship decisions.

Concluding remarks

Ecogenomics, due to the current nature of technical tools and model species, currently seems to be orientated more towards 'genomics' than towards 'eco'. This is understandable given the high technical demands of getting to grips with variation at the molecular level. However, gene function is a meaningless concept when not measured in ecologically relevant environments. Ecologists could, and should, take up the challenge of connecting environmental variation to variation in gene expression and phenotype. In fact, ecogenomic studies are beginning to emerge which address a variety of ecological phenomena, including reproductive biology in plants (Nasrallah *et al*. 2002; Bastow *et al*. 2004), flowering time (Koornneef *et al*. 2004; Stinchcombe *et al*. 2004), herbivory (Voelckel & Baldwin 2004; Ralph *et al*. 2006; Snoeren *et al*. 2006), plant–pathogen interactions (Allen *et al*. 2004), drought resistance (Knight *et al*. 2006), heavy metal tolerance (Rigola *et al*. 2006) and speciation (Hegarty *et al*. 2005). Much of this work is performed using *Arabidopsis* as a model. This has lead to very important advances in our understanding of gene regulation and genomic pathways. However, many ecologically important traits, habitats and biotic and abiotic interactions are not covered by *Arabidopsis*, leaving an important task for ecologists. When focusing on other ecologically interesting species, ecologists may offer an independent, comparative data set, serving as an ecological reference for the work on genomic model species. The expected reward is high: many ecological phenomena have not been investigated yet, and many genes and gene functions have not been discovered. Perhaps most importantly, by understanding the functional and evolutionary mechanisms that lead to phenotypic variation and adaptation, a more thorough understanding of ecological processes and patterns will become possible.

References

Allen, R.L., Bittner-Eddy, P.D., Grenville-Briggs, L.J., Meitz, J.C., Rehmany, A.P., Rose, L.E. & Beynon, J.L. (2004) Host–parasite coevolutionary conflict between *Arabidopsis* and downy mildew. *Science*, **306**, 1957–1960.

Alonso, J.M. & Ecker, J.R. (2006) Moving forward in reverse: genetic technologies to enable genome-wide phenomic screens in *Arabidopsis*. *Nature Reviews Genetics*, **7**, 524–536.

Archambault, A. & Bruneau, A. (2004) Phylogenetic utility of the LEAFY/FLORICAULA gene in the Caesalpinioideae (Leguminosae): gene duplication and a novel insertion. *Systematic Botany*, **29**, 609–626.

Bae, K.M., Kwon, Y.S., Cho, I.H. & Yi, S.I. (2006) Use of cDNA-AFLP for transcript profiling in narrow genetic pools; for example, cucumber (*Cucumis sativus* L.). *Plant Breeding*, **125**, 488–492.

Bastow, R., Mylne, J.S., Lister, C., Lippman, Z., Martienssen, R.A. & Dean, C. (2004) Vernalization requires epigenetic silencing of FLC by histone methylation. *Nature*, **427**, 164–167.

Bomblies, K., Wang, R.-L., Ambrose, B.A., Schmidt, R.J., Meeley, R.B. & Doebley, J. (2003) LEAFY homologues zfl1 and zfl2 control inflorescence architecture and flower patterning in maize. *Development*, **130**, 2385–2395.

Breyne, P., Dreesen, R., Cannoot, B., Rombaut, D., Vandepoele, K., Rombauts, S., Vanderhaeghen, R., Inzé, D. & Zabeau, M. (2003) Quantitative cDNA-AFLP analysis for genome-wide expression studies. *Molecular Genetics and Genomics*, **269**, 173–179.

Breyne, P. & Zabeau, M. (2001) Genome-wide expression analysis of plant cell cycle modulated genes. *Current Opinions in Plant Biology*, **4**, 136–142.

Chase, M.W., Salamin, N., Wilkinson, M., Dunwell, J.M., Kesanakurthi, R.P., Haidar, N. & Savolainen, V. (2005) Land plants and DNA barcodes: short-term and long-term goals. *Philosophical Transactions of the Royal Society, Series B*, **360**, 1889–1895.

Cnudde, F., Moretti, C., Porceddu, A., Pezzotti, N. & Gerats, T. (2003) Transcript profiling on developing *Petunia hybrida* floral organs. *Sexual Plant Reproduction*, **16**, 77–85.

Colbert, T., Till, B.J., Tompa, R., Reynolds, S., Steine, M.N., Yeung, A.T., McCallum, C.M., Greene, E.A., Comai, L. & Henikoff, S. (2001) Throughtput screening for induced point mutations. *Plant Physiology*, **126**, 480–484.

Comai, L., Young, K., Till, B.J., Reynolds, S.H., Greene, E.A., Codomo, C.A., Enns, L.C., Johnson, J.E., Burtner, C., Odden, A.R. & Henikoff, S. (2004) Efficient discovery of DNA polymorphisms in natural populations by ecotilling. *The Plant Journal*, **37**, 778–786.

Cronk, Q.C.B. (2005) Plant eco-devo: the potential of poplar as a model organism. *New Phytologist*, **166**, 39–48.

Dallas, P.B., Gottardo, N.G., Firth, M.J., Beesley, A.H., Hoffman, K., Terry, P.A., Freitas, J.R., Boag, J.M., Cummings, A.J. & Kees, U.R. (2005) *Gene Expression Levels Assessed by* Oligonucleotide microarray analysis and quantitative real-time PCR – how well do they correlate? *BMC Genomics*, **6**, Art. No. 59.

Domoney, C., Duc, G., Ellis, T.H.N., Ferrandiz, C., Firnhaber, C., Gallardo, K., Hofer, J., Kopka, J., Kuster, H., Madueno, F., Munier-Jolain, N.G., Mayer, K., Thompson,

© 2006 The Authors
Journal compilation
© 2006 British Ecological Society,

R., Udvardi, M. & Salon, C. (2006) Genetic and genomic analysis of legume flowers and seeds. *Current Opinion in Plant Biology*, **9**, 133–141.

Eckstein, R.L., O'Neill, R.A., Danihelka, J., Otte, A. & Kohler, W. (2006) Genetic structure among and within peripheral and central populations of three endangered floodplain violets. *Molecular Ecology*, **15**, 2367–2379.

Endo, T., Nasuda, S., Kawahara, T., Yasui, Y., Sasakuma, T., Sasanuma, T., Tsujimoto, H., Tanaka, H., Mochida, K., Kawaura, K. & Ogihara, Y. (2006) Wheat as a model plant system for studies of complex genomics of polyploids. *Plant and Cell Physiology*, **47**, S253–S253 (Suppl. S).

Feder, M.E. & Mitchell-Olds, T. (2003) Evolutionary and ecological functional genomics. *Nature Reviews in Genetics*, **4**, 651–657.

Filatov, V., Dowdle, J., Smirnoff, N., Ford-Lloyd, B., Newbury, H.J. & Macnair, M.R. (2006) Comparison of gene expression in segregating families identifies genes and genomic regions involved in a novel adaptation, zinc hyperaccumulation. *Molecular Ecology*, **15**, 3045–3059.

Fire, A., Xu, S.Q., Montgomery, M.K., Kostas, S.A., Driver, S.E. & Mello, C.C. (1998) Potent and specific genetic inference by double-stranded RNA in *Caenorhabditis elegans*. *Nature*, **391**, 806–811.

Gibson, G. (2002) Microarrays in ecology and evolution: a preview. *Molecular Ecology*, **11**, 17–24.

Gilchrist, E.J., Haughn, G.W., Ying, C.C., Otto, S.P., Zhuang, J., Cheung, D., Hamberger, B., Aboutorabi, F., Kalynyak, T., Johnson, L., Bohlmann, J., Ellis, B.E., Douglas, C.J. & Cronk, Q.C.B. (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.

Hegarty, J.H., Jones, J.M., Wlson, I.D., Barker, G.L., Coghill, J.A., Sanchez-Baracaldo, P., Liu, G., Buggs, R.J.A., Abbott, R.J., Edwards, K.J. & Hiscock, S.J. (2005) Development of anonymous cDNA microarrays to study changes to the *Senecio* floral transcriptome during hybrid speciation. *Molecular Ecology*, **14**, 2493–2510.

Heid, C.A., Stevens, J., Livak, K.J. & Williams, P.M. (1996) Real time quantitative PCR. *Genome Research*, **6**, 986–994.

Heidel, A.J., Clauss, M.J., Kroymann, J., Savolainen, O. & Mitchell-Olds, T. (2006) Natural variation in MAM within and between populations of *Arabidopsis lyrata* determines glucosinolate phenotype. *Genetics*, **173**, 1629–1636.

Jackson, R.B., Linder, R., Lynch, M., Purugganan, M., Somerville, S. & Thayer, S.S. (2002) Linking molecular insight and ecological research. *Trends in Ecology and Evolution*, **17**, 409–414.

Knight, A.K., Vogel, H., Kroymann, J., Shumate, A., Witsenboer, H. & Mitchell-Olds, T. (2006) Expression profiling and local adaptation of *Boechera holboelli* populations for water use efficiency across a naturally occurring water stress gradient. *Molecular Ecology*, **15**, 1229–1237.

Koornneef, M., Alonso-Blanco, C. & Vreugdenhil, D. (2004) Naturally occurring genetic variation in *Arabidopsis thaliana*. *Annual Reviews in Plant Biology*, **55**, 141–172.

Lee, C.E. & Mitchell-Olds, T. (2006) Preface to the special issue: ecological and evolutionary genomics of populations in nature. *Molecular Ecology*, **15**, 1193–1196.

Lee, S.H., Walker, D.R., CreganP.B. & Boerma, H.R. (2004) Comparison of four flow cytometric SNP detection assays and their use in plant improvement. *Theoretical and Applied Genetics*, **110**, 167–174.

Liu, N., Dang, Q.-L. & Parker, W.H. (2006) Genetic variation of *Populus tremuloides* in ecophysiological responses to CO_2 elevation. *Canadian Journal of Botany*, **84**, 294–302.

Magri, D., Vendramin, G.G., Comps, B., Dupanloup, I., Geburek, T., Gomory, D., Latalowa, M., Litt, T., Paule, L., Roure, J.M., Tautau, I., van der Knaap, W.O., Petit, R.J. & de Beaulieu, J.L. (2006) A new scenario for the Quaternary history of European beech populations: palaeobotanical evidence and genetic consequences. *New Phytologist*, **171**, 199–221.

Mueller, L.A., Solow, T.H., Taylor, N., Skwarecki, B., Buels, R., Binns, J., Lin, C.W., Wright, M.H., Ahrens, R., Wang, Y., Herbst, E.V., Keyder, E.R., Menda, N., Zamir, D. & Tanksley, S.D. (2005) The SOL genomics network. A comparative resource for Solanaceae biology and beyond. *Plant Physiology*, **138**, 1310–1317.

Nasrallah, M.E., Liu, P. & Nasrallah, J.B. (2002) Generation of self-incompatible *Arabidopsis thaliana* by transfer of two S locus genes from *A. Lyrata Science*, **297**, 247–249.

Ogihara, Y. (2006) Wheat as a model plant system for studies on complex genomics of polyploids. *Plant and Cell Physiology*, **47**, S22–S22 (Suppl. S).

Ouborg, N.J., Piquot, Y. & van Groenendael, J.M. (1999) Population genetics, molecular markers and the study of dispersal in plants. *Journal of Ecology*, **87**, 551–568.

Price, A.H. (2006) Believe it or not, QTLs are accurate! *Trends in Plant Science*, **11**, 213–216.

Pring, D.R. & Tang, H.V. (2004) Transcript profiling of male-fertile and male-sterile *Sorghum* indicates extensive alterations in gene expression during microgametogenesis. *Sexual Plant Reproduction*, **16**, 289–297.

Ralph, S.G., Yueh, H., Friedmann, M., Aeschliman, D., Zeznik, J.A., Nelson, C.C., Butterfield, Y.S.N., Kirkpatrick, R., Liu, J., Jones, S.J.M., Marra, M.A., Douglas, C.J., Ritland, K. & Bohlmann, J. (2006) Conifer defence against insects: microarray gene expression profiling of Sitka spruce (*Picea sitchensis*) induced by mechanical wounding or feeding by spruce budworms (*Choristoneura occidentalis*) or white pine weevils (*Pissodes strobi*) reveals large-scale changes of the host transcriptome. *Plant, Cell and Environment*, **29**, 1545–1570.

Ranz, J.M. & Machado, C.A. (2006) Uncovering evolutionary patterns of gene expression using microarrays. *Trends in Ecology and Evolution*, **21**, 29–37.

Rigola, D., Fiers, M., Vurro, E. & Aarts, M.G.M. (2006) The heavy metal hyperaccumulator *Thlaspi caerulescens* expresses many species-specific genes, as identified by comparative expressed sequence tag analysis. *New Phytologist*, **170**, 753–765.

Robertson, D. (2004) VIGS vectors for gene silencing: many targets, many tools. *Annual Review of Plant Biology*, **55**, 495–519.

Schena, M., Shalon, D., Davis, R. & Brown, P. (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science*, **270**, 467–470.

Seifert, M., Scherf, M., Epple, A. & Werner, T. (2005) Multi-evidence microarray mining. *Trends in Genetics*, **21**, 553–558.

Siebert, P.D., Chenchik, A., Kellog, D.E., Lukyanov, K.A. & Lukyanov, S.A. (1995) An improved PCR method for walking in uncloned genomic DNA. *Nucleic Acids Research*, **23**, 1087–1088.

Slavov, G.T., Howe, G.T., Gyaourova, A.V., Birkes, D.S. & Adams, W.T. (2005) Estimating pollen flow using SSR markers and paternity exclusion: accounting for mistyping. *Molecular Ecology*, **14**, 3109–3121.

Snoeren, T.A.L., de Jong, P.W. & Dicke, M. (2006) Ecogenomic approach to the role of herbivore-induced plant volatiles in community ecology. *Journal of Ecology*, doi: 10.1111/j.0022-0477.2006.01183.x

Soltis, D.E., Albert, V.A., Savolainen, V., Hilu, K., Qiu, Y.L., Chase, M.W., Farris, J.S., Stefanovic, S., Rice, D.W., Palmer, J.D. & Soltis, P.S. (2004) Genome-scale data, angiosperm relationships, and 'ending incongruence': a cautionary tale of phylogenetics. *Trends in Plant Science*, **9**, 477–483.

Sterky, F., Bhalerao, R.R., Unneberg, P., Segerman, B., Nilsson, P., Brunner, A.M., Charbonnel-Campaa, L., Lindvall, J.J., Tandre, K., Strausss, S.H., Sundberg, B., Gustafsson, P., Uhlen, M., Bhalerao, R.P., Nilsson, O., Sandberg, G., Karlsson, J., Lundeberg, J. & Jansson, S. (2004) A *Populus* EST resource

© 2006 The Authors
Journal compilation
© 2006 British Ecological Society,

for plant functional genomics. *Proceedings of the National Academy of Sciences USA*, **101**, 13951–13956.

Stinchcombe, J.R., Weinig, C., Ungerer, M., Olsen, K.M., Mays, C., Halldorsdottiret, S.S., Purugganan, M.D. & Schmitt, J. (2004) A latitudinal cline in flowering time in *Arabidopsis thaliana* modulated by the flowering time gene FRIGIDA. *Proceedings of the National Academy of Sciences USA*, **191**, 4712–4717.

Suwabe, K., Tsukazaki, H., Iketani, H., Hatakeyama, K., Kondo, M., Fulimura, M., Nunome, T., Fukuoka, H., Hirai, M. & Matsumoto, S. (2006) Simple sequence repeat-based comparative genomics between *Brassica rapa* and *Arabidopsis thaliana*: the genetic origin of clubroot disease. *Genetics*, **173**, 309–319.

Taylor, G., Street, N.R., Tricker, P.J., Sjödin, A., Graham, L., Skogstöm, O., Calfapietra, C., Scarascia-Muguozza, G. & Jansson, S. (2005) The transcriptome of *Populus* in elevated CO_2. *New Phytologist*, **167**, 143–154.

Thomas, M.A. & Klaper, R. (2004) Genomics for the ecological toolbox. *Trends in Ecology and Evolution*, **19**, 439–445.

Till, B.J., Reynolds, S.H., Greene, E.A., Codomo, C., Enns, L., Johnson, J., Burtner, C., Odden, A., Young, K., Taylor, N., Henikoff, J.G., Comai, L. & Henikoff, S. (2003) Large-scale discovery of induced point mutations with high-throughput TILLING. *Genome Research*, **3**, 524–530.

Touzet, P., Hueber, N., Burkholz, A., Bornes, S. & Cuguen, J. (2004) Genetic analysis of male sterility restoration in wild cytoplasmic male sterility G of beet. *Theoretical and Applied Genetics*, **109**, 240–247.

Van Straalen, N.M. & Roelofs, D. (2006) *An Introduction to Ecological Genomics*. Oxford University Press, Oxford.

Van Tienderen, P.H., de Haan, A.A., van der Linden, C.G. & Vosman, B. (2002) Biodiversity assessment using markers for ecologically important traits. *Trends in Ecology and Evolution*, **17**, 577–582.

Voelckel, C. & Baldwin, I.T. (2004) Generalist and specialist lepidopteran larvae elicit different transcriptional responses in *Nicotiana attenuate*, which correlates with larval FAC profiles. *Ecology Letters*, **7**, 770–775.

Waterhouse, P. & Wang, M.-B. (2001) Gene silencing as an adaptive defence against viruses. *Nature*, **411**, 834–842.

Glossary

Affymetrix Chip: Oligonucleotide arrays or GeneChips, developed by the Affymetrix Company, are based on hybridization to small, high-density arrays containing tens of thousands of synthetic oligonucleotides.

AFLP: Amplified fragment length polymorphism PCR is a method to detect polymorphisms in DNA. It analyses variations in genomic DNA by comparing fragmented DNA from different genotypes. Fragmentation occurs with restriction enzymes and is sequence dependent. Polymorphisms are visualised by separation of the fragments on a polyacrylamide gel.

cDNA: Complement or copy DNA obtained by reverse transcription of mRNA.

cDNA-AFLP: cDNA-Amplified Fragment Length Polymorphisms is a PCR based genome-wide expression analysis technique used to identify differentially expressed genes.

dsRNA: Double-stranded RNA is RNA with two complementary strands, comparable to the DNA from higher organisms. The genetic material of some viruses exists of dsRNA and can in eukayotic cells after infection initiate the process of RNA interference via the formation of siRNAs (small interfering RNAs).

EST: An expressed sequence tag, the cDNA sequence of the coding region of a gene expressed in a plant. ESTs are produced by sequencing of several hundreds of base pairs from both ends of cDNA clones.

ISSR: Inter-simple sequence repeat analysis is a single primer PCR that amplifies products between two simple sequence repeats to produce reproducible markers useful for the quantification of genetic polymorphism within species.

Microarray: is a hybridization based genome wide expression analysis technique used to identify differentially expressed genes. A microarray is a glass (or other substrate) slide onto which hundreds to thousands of ESTs have been spotted.

PCR: Polymerase chain reaction is a molecular biology technique for enzymatically replicating DNA. The technique allows a small amount of DNA to be amplified exponentially *in vitro* and it can be performed without restrictions on the form of DNA and it can be modified to perform an array of genetic manipulations.

qRT-PCR: Quantitative real-time PCR is a modification of PCR used to measure the quantity of a PCR product during the reaction. It is an indirect method for quantitatively measuring starting amounts of DNA.

QTL: Quantitative Trait Locus analysis maps the genomic location of one or more genes that affect a continuously (rather than discrete) variable phenotypic trait.

RNAi: is a mechanism where the presence of certain fragments of double-stranded RNA (dsRNA) interferes with the expression of a particular gene. Also known as post-transcriptional gene silencing (PTGS) or co-suppression.

RT-PCR: Reverse transcription polymerase chain reaction is a technique for amplifying a defined piece of a ribonucleic acid (RNA) molecule. The RNA strand is first reverse transcribed into its complementary DNA (cDNA), followed by amplification of the resulting DNA using PCR.

SSH: Subtractive Hybridization is a technique that enables comparison of two populations of mRNA and obtains clones of genes that are specifically expressed in one.

TILLING: Targeted Induced Local Lesions in Genomes. A systematic method to identify a specific mutant in a mutagenized plant collection.

Transcript profiling: characterization of the set of genes and their levels of expression on a genomic scale in a particular group of cells, at one specific moment, and under specific experimental conditions.

VIGS: virus-induced gene silencing. Gene silencing triggered by viral vector that encodes a dsRNA (double stranded RNA).

Received 11 August 2006
revision accepted 10 October 2006
Handling Editor: Michael Hutchings

© 2006 The Authors
Journal compilation
© 2006 British Ecological Society,