Annual Plant Reviews (2009) **35**, 331–349
doi: 10.1002/9781119312994.apr0383
http://onlinelibrary.wiley.com

Chapter 12

PERSPECTIVES ON ECOLOGICAL AND EVOLUTIONARY SYSTEMS BIOLOGY

Christina L. Richards,[1] Yoshie Hanzawa,[1] Manpreet S. Katari,[1] Ian M. Ehrenreich,[1,2] Kathleen E. Engelmann[3] and Michael D. Purugganan[1]

[1] *Department of Biology and Center for Genomics and Systems Biology, New York University, New York, NY, USA*
[2] *Department of Genetics, North Carolina State University, Raleigh, NC, USA*
[3] *University of Bridgeport, Bridgeport, CT, USA*

See colour figures at the end of this chapter.

Abstract: Understanding the emergent properties inherent to genome function requires an integrated approach of data from all levels of biology. Molecular biology data alone does not describe the complex interacting functions of organisms, while studies at the level of ecological communities and ecosystems have provided little insight into the molecular underpinnings of adaptation. Merging ecology and evolution into systems biology allows researchers to exploit a wealth of genomic information by incorporating the natural phenotypic, genetic and epigenetic diversity of model systems as well as their diverse ecologies and evolutionary histories. Here, we suggest that systems biology could more fully address the question of how organisms respond to environment if studies incorporated real field settings or experimental manipulation of relevant environmental factors. In addition, although the application of genomic approaches to non-model systems has been slow, we highlight some of the significant progress that has been made. Ecological and evolutionary systems biology will lead to a much more sophisticated understanding of the origins and functions of biological diversity, and serve as a critical component in deciphering how organisms respond to complex environments.

Keywords: *Arabidopsis thaliana*; *Caenorhabditis elegans*; *Drosophila* species; ecological genomics; epigenetics; ecological transcriptome; experimental design; flowering time network; *Fundulus heteroclitus*; natural environment; non-model systems

Originally published in *Annual Plant Reviews*, Volume 35: *Plant Systems Biology*.
First Edition. Edited by Gloria M. Coruzzi and Rodrigo A. Gutiérrez.

12.1 Emergent properties of systems biology, ecology and evolution

Understanding biological diversity is a complex question that motivates studies from the molecular level through the level of ecosystems, and at all intervening levels of biological organization. As genomics and systems biology begin to mature, work that synthesizes ecology and evolution with systems biology can begin to explore how these multiple levels of biology contribute to diversity. Systems biology approaches will help to elucidate the complex interplay among different types of biological data (molecular, regulatory, physiological, phenotypic and environmental) and are likely to facilitate unprecedented advances in our understanding of how organisms respond to the environment over both short and long timescales. Ultimately, these approaches can be expanded to more thoroughly address many classic ecological and evolutionary questions about competition, adaptation, invasive species interactions and even global climate change.

It would be difficult to discuss comprehensively in one chapter all the facets that bring ecology, evolution and systems biology together. It is possible, however, to visit the landscape where these disciplines intersect and enrich our understanding of biological processes. In this chapter, we explore some aspects of the possible interfaces between systems biology and ecology, and point out opportunities for further exploration for a new generation of ecological and evolutionary systems biologists. Using examples from several model organisms, particularly *Arabidopsis thaliana*, we illustrate the synergistic interface between systems biology, ecology and evolution.

12.2 Complex environments and ecological systems biology

Organisms in the real world are continuously exposed to multiple environmental signals and must respond appropriately to the dynamic conditions found in nature. Temperature, photoperiod and resource availability are just some key environmental conditions that cue organismal responses. There have been significant advances in dissecting how these and other environmental signals are translated by the organism to appropriate gene expression levels that may ultimately determine phenotypes (Pigliucci, 1996; Schlichting and Smith, 2002); however, with very few exceptions, these studies have been carried out in homogenous environments in controlled laboratory conditions. The natural world, in contrast, is anything but controlled. Dynamic environmental signals can fluctuate temporally during an individual's life cycle and can change spatially according to climate, with varying degrees of predictability. Complex natural environments are the norm, and it is in this context that developmental pathways and physiological states actually function and evolve. It is unclear how these fluctuating signals interact with each other and with an organism's genotype to fine-tune phenotypic

response, but clearly the next step in this area of research is to determine how gene expression is modulated in the wild.

Understanding how genes and functional genetic networks are regulated in natural ecological settings in the midst of fluctuating environmental signals and between genetically distinct individuals remain key issues that lie at the intersection of ecology, evolution and systems biology. Although molecular studies have provided data on a wide variety of functionally important genes in various organisms, this data remains largely insufficient to describe the complex interactions that underlie most biological processes, thus highlighting the limitations of reductionist approaches (Mazzochi, 2008). Similarly, studies at the level of populations, ecological communities and ecosystems rarely address the molecular underpinnings of organismal variation and adaptation. Systems biology provides tools and perspectives that can examine the emergent properties inherent in genome function and allow an integrated approach of data from all biological levels (O'Malley and Dupré, 2005; Bonneau *et al.*, 2007; Bruggeman and Westerhoff, 2007).

12.3 Gene networks and the ecological transcriptome

Although the goal of biology is to understand how organismal phenotypes and the genetic architecture of various traits have evolved in nature, gene expression in the wild is poorly understood. This gap in our knowledge is unfortunate, since data on the characteristics of gene expression under natural conditions is a necessary step if we are to determine how genes are regulated in the dynamic, complex environments observed in natural field conditions. It is likely that laboratory expression experiments across several treatments, even with attempts to integrate them, will be unable to mimic how gene expression behaves in natural settings. Microarray technologies allow us to measure the levels and patterns of gene expression in the dynamic ecological settings of field environments, and systems biology will help interpret this massive amount of data so that we can begin to more fully understand the ecological transcriptome.

There are a large number of global gene expression studies in *A. thaliana* (e.g. Harmer *et al.*, 2000; Birnbaum *et al.*, 2003; Schmid *et al.*, 2005) as well as other plant species (e.g. Alba *et al.*, 2005; Li *et al.*, 2006; Starker *et al.*, 2006; Swanson-Wagner *et al.*, 2006). However, most of these studies were undertaken in controlled laboratory conditions, and we could only find five global gene expression studies of wild plant species (as opposed to crop plants) in field conditions (Miyazaki *et al.*, 2004; Taylor *et al.*, 2005; Yang and Loopstra, 2005; Ainsworth *et al.*, 2006; Schmidt and Baldwin, 2006). This limited number of reports clearly demonstrates that there are significant transcriptional differences between controlled and field growth conditions.

One of the earliest field microarray studies was on *Solanum nigrum* using a small gene array. This study showed that methyl jasmonate (MeJa) elicitation in competitive environments resulted in different patterns of transcriptional

changes between greenhouse and field-grown plants (Schmidt and Baldwin, 2006). MeJa-treated plants consistently showed differences in transcript levels for genes involved in oxylipin signalling and primary metabolism. However, there were a larger number of downregulated genes in greenhouse-grown compared with field-grown plants, which suggests that the latter are less sensitive to MeJa elicitation. This decreased downregulation of gene transcripts in field plants is thought to be the result of greater exposure to abiotic stresses in the wild (Schmidt and Baldwin, 2006), and clearly demonstrates the importance of measuring response in natural environments.

Another series of studies examined gene transcription responses to increased CO_2 levels in free air CO_2 enrichment (FACE) in *Populus* (Taylor *et al.*, 2005), soybean (Ainsworth *et al.*, 2006) and *A. thaliana* (Miyazaki *et al.*, 2004). Following 6 years of exposure of a *Populus* clone to elevated CO_2 levels, gene expression response depended on the developmental age of the leaves, and ~50 transcripts differed significantly between different CO_2 environments (Taylor *et al.*, 2005). A similar result was also reported in soybean (Ainsworth *et al.*, 2006). The *A. thaliana* study examined both CO_2 and ozone exposure. Microarrays with probes for ~26 000 DNA elements were used to compare transcripts from plants grown in growth chambers to those grown in FACE rings under ambient field conditions. In this study, most changes in gene expression were observed between growth chamber and ambient field conditions rather than between atmospheric treatments. Greater than 1000 transcripts were either up or downregulated between controlled versus field ambient conditions compared with high versus low CO_2 or ozone levels. There was a preponderance of genes associated with general defence reactions, secondary metabolism, redox control, energy provision, protein turnover, signalling and transcription (Miyazaki *et al.*, 2004).

These experiments point out the large-scale differences in expression patterns between growth chamber versus field conditions, but also demonstrate the feasibility of assaying for global transcript abundance levels in field-grown plants using microarray technologies. Still, these studies only deliver a list of possible genes involved in response to these environmental treatments. At best, this approach has identified some novel genes involved in response to different environments (Taylor *et al.*, 2005). Systems biology approaches allow researchers to go beyond the Venn diagram descriptions of differential gene expression (Li *et al.*, 2006; Kammenga *et al.*, 2007; Lee *et al.*, 2008) to formulate dynamic networks that can incorporate changes over time and environments (Bonneau *et al.*, 2007).

One important component of the ecological transcriptome that has largely been overlooked in ecology and evolution is epigenetic effects. Epigenetic effects are the subject of intense study in genomics, and several studies have begun to develop comprehensive maps of epigenetic marks (Vaughn *et al.*, 2007; Zhang *et al.*, 2007; Zilberman *et al.*, 2007; Cokus *et al.*, 2008; Zhang *et al.*, 2008). DNA methylation is the most well-understood epigenetic mark and the few existing studies of natural variation in genome methylation

report epigenome polymorphisms in *Gossypium* (Keyte *et al.*, 2006), *Arabidopsis* (Cervera *et al.*, 2002; Riddle and Richards, 2002, 2005), *Oryza* (Ashikawa, 2001; Wang *et al.*, 2004), *Pisum* (Knox and Ellis, 2001) and *Spartina* (Salmon *et al.*, 2005).

There is ever-increasing evidence that heritable variation in ecologically relevant traits can be generated through a suite of epigenetic mechanisms that can alter phenotypes even in the absence of genetic variation (Grant-Downton and Dickinson, 2005, 2006; Jablonka and Lamb, 2005; Rapp and Wendel, 2005; Richards, 2006). Molecular studies using methylation sensitive markers in *Triticum* (Sherman and Talbert, 2002) and *Arabidopsis* (Burn *et al.*, 1993) show that external temperature can change methylation patterns that induce early flowering time. Similarly, one study found a nearly 10% reduction in methylation in induced early flowering lines of *Linim usitatissimum* compared to closely related lines with normal flowering time (Fields *et al.*, 2005). Furthermore, studies in *Arabidopsis*, *Brassica*, *Oryza*, *Spartina* and *Triticum* reveal that methylation patterns can be radically altered by hybridization or polyploidization (Chen and Pikaard, 1997; Comai *et al.*, 2000; Liu *et al.*, 2001; Shaked *et al.*, 2001; Madlung *et al.*, 2002; Salmon *et al.*, 2005).

Epigenetic variation appears to be common in plants and is therefore likely to have effects that are visible to natural selection (reviewed by Rapp and Wendel, 2005; Grant-Downton and Dickinson, 2006). Recent studies also indicate that in some cases, environmentally induced epigenetic changes may be inherited by future generations (Richards, 2006; Whitelaw and Whitelaw, 2006; Bossdorf *et al.*, 2008). In addition, because epigenetic processes are an important component of hybridization and polyploidization events, they may play a key role in speciation and the biology of many invasive species through these processes (Ellstrand and Schierenbeck, 2000; Liu and Wendel, 2003; Rapp and Wendel, 2005; Salmon *et al.*, 2005; Chen and Ni, 2006; Bossdorf *et al.*, 2008). To date, very few ecology and evolution studies have considered the importance of epigenetic effects and most have not gone beyond documenting that they exist (Rapp and Wendel, 2005; Bossdorf *et al.*, 2008). With the appropriate experimental design (Bossdorf *et al.*, 2008), systems biology may be the best context with which to disentangle the contributions of environment, genotype and epigenotype to phenotypic variation.

12.4 Analysis of systems biology data: the role of ecological and evolutionary methods

Ecological and evolutionary information may contribute to systems biology through enhanced functional genome annotation. Placing genes and molecular networks into functional ecological and evolutionary contexts exposes the behaviour of the network in a realistic setting. Expression studies can be rendered even more informative if they incorporate experimental

manipulations or treatments of ecologically relevant factors like light, water, nutrients or salt levels or even whole field environments (Fig. 12.1). A large percentage of genes in the genome either encode hypothetical proteins or are of unknown function; ecological and evolutionary studies may be especially useful in proving the basis for expanding the annotation of those genes in the genome. Bioinformatic approaches combine the genome-wide data collected in these types of experiments with previously verified genetic interactions to assemble putative gene interaction networks. Such gene interaction networks may suggest novel associations and functions of genes, and how these relationships change across environments of interest. An important component of the systems biology approach is to use these findings to generate new hypotheses and new experimental designs. Iterating this process can refine the model of these network interactions.

Experimental ecology also provides a long history of developing methods in experimental design and analysis to identify the relevant contributions of environmental factors to organismal response (Bailey, 1981; Sokal and Rohlf, 1995; Scheiner and Gurevitch, 2001). Based on the constraints of the experimental setup and organisms, this may require split plot or repeated-measures designs and include randomization and blocking to control for spatial or systemic contributions to variation, and avoid confounding these elements with variables of interest. Typically, these studies then use analysis of variance (ANOVA) to disentangle the contribution of multiple environmental variables or experimental treatments to variation in ecologically important traits (Sokal and Rohlf, 1995; Scheiner and Gurevitch, 2001). The general model for ANOVA allows for examination of several independent variables by defining the relationship between traits and any number of independent variables, and mixed-model ANOVAs allow for testing of random effects like population and genotype (Littell *et al.*, 2006) within this context.

These approaches have been adapted to the interpretation of gene expression and gene specific modelling of microarray data in particular by fitting a global normalization model incorporating all of the genes, and then running a separate ANOVA for each gene (Wolfinger *et al.*, 2001; Aryoles and Gibson, 2006). In addition, using ANOVA allows for ecological genomic studies to test whether the pattern of gene expression variation is correlated with environmental or ecological variables like temperature, precipitation levels, soil moisture, daylength, photosynthetically active radiation (PAR), age, herbivory and disease status.

12.5 The ecological and evolutionary context of model organisms: the example of *Arabidopsis* and beyond

Much of the advance in molecular and developmental biology in the last few decades owes to the focus of researchers on a handful of model organisms as subjects of experimental study. There has been relatively little attention paid

to the natural history of these model species, but that has changed in the last few years as investigators begin to understand the natural contexts in which these organisms live. There has been increasing efforts to use these model species to address fundamental questions in ecology – to exploit the standing natural variation observed in these species and to study their development and behaviour in their natural environments.

The need for a systems biological approach in the study of ecological dynamics can be illustrated by considering recent work with *A. thaliana* (L.) Heynh (Pigliucci, 1998; Koornneef *et al.*, 2004; Shimizu and Purugganan, 2005). This species is a weedy annual plant, occupying disturbed habitats such as the margins of agricultural fields as well as natural ruderal environments. The native range of *A. thaliana* covers Eurasia and Northern Africa, and it is naturalized widely in the world, including in North America and Japan (Hoffmann, 2002). Evolutionary analysis of a set of genome-wide markers suggests that the current species range, which includes most of Eurasia, and parts of North Africa and North America, is the result of the expansion of the species ~17 000 years ago from two glacial refugia in the Iberian Peninsula and Asia (Sharbel *et al.*, 2000). The presence of *A. thaliana* in North America is a recent phenomenon, which is likely due to the migration of Europeans to the continent over the last 300–400 years.

A. thaliana displays a wide range of ecological relationships, including within- and between-species interactions and adaptations to abiotic environments. It responds physiologically and developmentally to a variety of environmental cues, including light, daylength, vernalization, nutrient and water levels (reviewed by Pigliucci, 1998; Koornneef *et al*., 2004; Shimizu and Purugganan, 2005), can be infected by a wide array of bacterial and fungal pathogens, and is preyed upon by many insect herbivores (Kliebenstein *et al.*, 2002). Despite the role of *A. thaliana* as a model plant system, remarkably little is known about the phenotypic range and performance of this species in the wild. Most of our knowledge of this organism is in the artificial environment of the laboratory, and there is mounting evidence that the behaviour of this organism can differ substantially in the wild. A few field studies of *A. thaliana* have begun to shed light on the ecological genetics of this organism. Early investigations have examined selection in this species at short spatial scales in field conditions (Stratton and Bennington, 1996), and documented selection costs for trichomes as a defence against herbivores (Mauricio and Rausher, 1997). Other studies have looked at the seasonal germination timing in the field (Donohue *et al.*, 2005), fitness costs of *R* disease resistance gene polymorphisms (Tian *et al.*, 2003) and the role of epistasis in fitness-related traits (Malmberg *et al.*, 2005).

Some of the most detailed field studies in *A. thaliana* have focused on the genetic architecture of flowering time, which is arguably one of the most important traits for the ecology and evolution of plant species. A quantitative trait locus (QTL) mapping study for date of bolting (the transition from vegetative to reproductive growth) was undertaken in natural seasonal field

environments in Rhode Island and North Carolina (Weinig *et al.*, 2002). This study revealed that photoperiod-specific QTLs found in controlled conditions were undetectable in natural environments, while several QTLs with major effects on flowering time in one or more field environments were undetectable under controlled environment conditions (Weinig *et al.*, 2002). Candidate gene association studies, while not definitive, have suggested that common allelic variation at the flowering time genes *CRY2* (Olsen *et al.*, 2004), *FRI* (Caicedo *et al.*, 2004; Stinchcombe *et al.*, 2004) and *FLC* (Caicedo *et al.*, 2004) is associated with flowering time diversity in natural field conditions. Altogether, these QTL and association mapping studies suggest that the genetic architecture of this life history transition differs significantly between laboratory and natural environments.

While *A. thaliana* has become a robust ecological and evolutionary model system, particularly for field studies, other model genomic organisms offer additional advantages. *Drosophila melanogaster*, a cosmopolitan species, has a long history in the study of evolutionary genetics, and has some of the most advanced tools available for evolutionary studies. In particular, recent advances include the complete genome sequences of 12 closely related species and genome-wide polymorphism data for multiple species (Clark *et al.*, 2007; Stark *et al.*, 2007). Similarly, *Caenorhabditis elegans*, a soil nematode is now emerging as a key model species for quantitative genomics (Li *et al.*, 2006), and systems biology (Gunsalus *et al.*, 2005) (see Chapter 3) could be used to potentially integrate the large amount of data on this organism at the ecological level. There are also a large number of bacterial and fungal systems that can be exploited to study microbial ecology (Whitaker and Banfield, 2006; Wilmes and Bond, 2006; Xu, 2006; Bonneau *et al.*, 2007).

In addition to a wealth of information on model systems, the development of genomic technologies has expanded the range of possible wild species that can be used in ecological and evolutionary systems biology. For example, genome sequencing has been completed in the deciduous forest tree species of *Populus* (Jansson and Douglas, 2007) and the monkey flower *Mimulus guttatus* (Wu *et al.*, 2008), and tools for use in ecological and evolutionary studies in these and other species are becoming available. The completed sequencing of 12 genomes of *Drosophila* species (Clark *et al.*, 2007; Stark *et al.*, 2007) and the near completion of the *A. thaliana* relatives *A. lyrata* and *Capsella rubella* offer opportunities to significantly expand evolutionary systems biology studies beyond model systems as we decipher the commonalities and differences between closely related species. In addition, researchers can increasingly exploit tools and data in species that are closely related to the model systems with genomic resources. For example, microarray chips designed for model taxa can be used for transcriptome studies in non-model species because those probes that do not hybridize can be identified and left out of an analysis (Slotte *et al.*, 2007; Travers *et al.*, 2007). Horvath *et al.* (2003) used this approach with *A. thaliana* microarrays to analyze gene expression in several distant species, including leafy spurge and poplar. With these and other on-going efforts, such

as the increasing availability of massively parallel sequencing machines, it is now possible that systems biology can be studied in a large number of relevant taxa (Bonneau *et al.*, 2007).

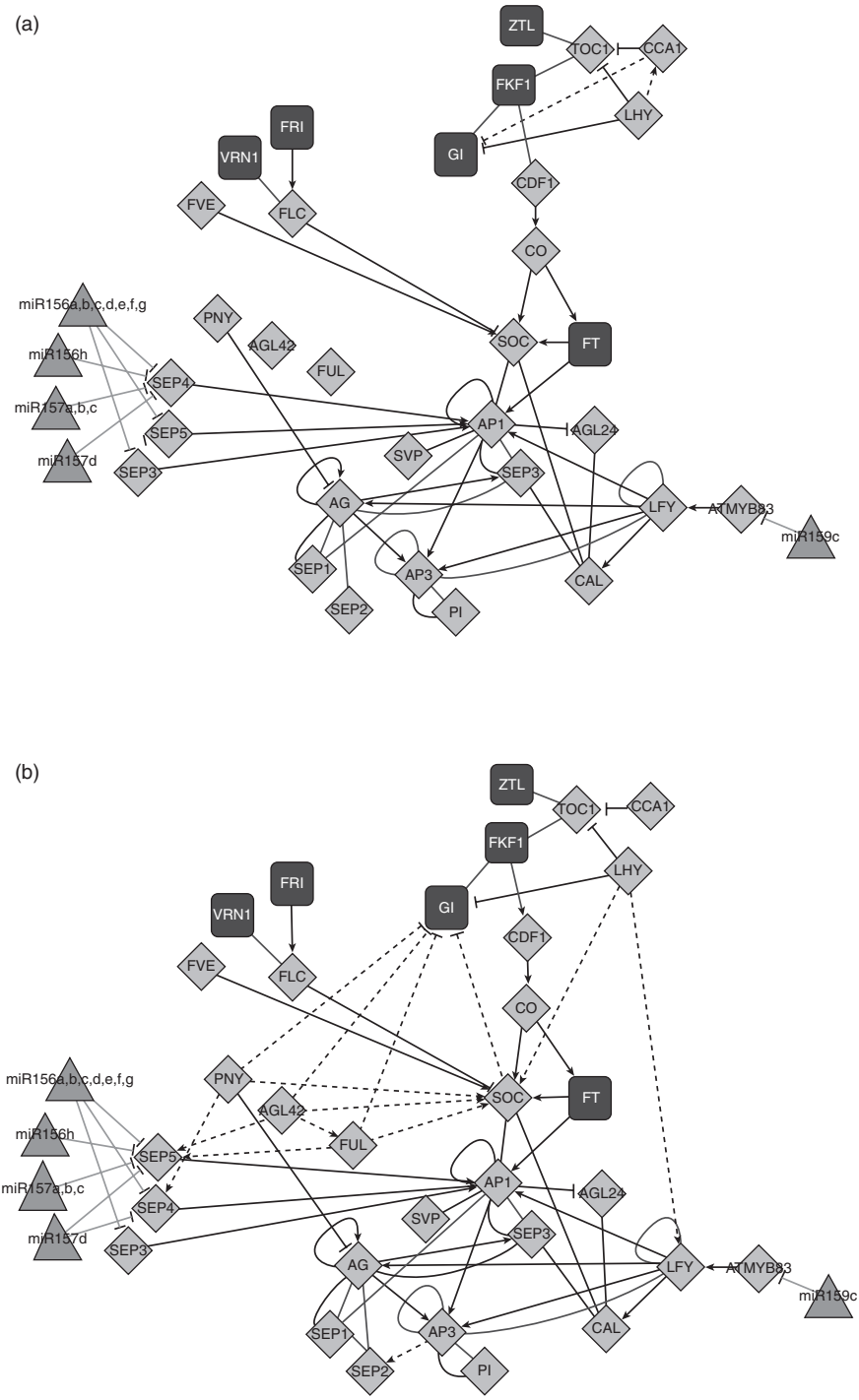
12.6 Natural variation in genomes and gene networks

The genetic variation extant within species and the phenotypic variation that may result from it are key components that underlie species diversification and adaptation. There have been concerted efforts to understand the extent of molecular variation between individuals and populations, and there has been a recent resurgence in interest in mapping natural genetic variants that may be responsible for natural phenotypic variation. These include genome mapping techniques that scan genomes and map trait loci, such as methods for QTL mapping (Mackay, 2001) and linkage disequilibrium/association mapping (Cardon and Abecasis, 2003).

Coupling systems biology with population variation can provide an evolutionary and ecological context to the large number of global gene expression studies by utilizing multiple sources of information across a diversity of genotypes. One area of interest, for example, is whether key points in a network are responsible for obvious phenotypic differences between genotypes. This can be illustrated by recent studies on variation between different accessions of *A. thaliana* in the important flowering time genes. The network of genes in the flowering time pathway in *A. thaliana* is one of the best-studied regulatory pathways that control a key life history trait in plants (Mouradov *et al.*, 2002; Simpson and Dean, 2002). Two genes in the vernalization pathway in particular, *FRI* and *FLC* are believed to play an essential role in determining response to prolonged cold temperature (Johanson *et al.*, 2000; Weinig *et al.*, 2002; Olsen *et al.*, 2004). *FLC*, which represses flowering, is regulated by *FRI*. An active *FRI-FLC* pathway results in late flowering, whereas an inactive *FRI-FLC* pathway results in early flowering.

Through controlled crosses, researchers have developed nearly isogenic lines (NIL) of *A. thaliana,* which are genetically identical except for one target locus. For example, two of these lines with a functional *FRI* allele have either a non-functional or a functional *FLC* allele (Lee *et al.*, 1993; Michaels and Amasino, 1999). To develop these lines, a functional *FRI* allele from the wild accession Sf2 was introgressed into Col-0, which carries an active *FLC*, resulting in an active *FRI-FLC* pathway (Lee *et al.*, 1993). The inactive *FRI-FLC* pathway was obtained by introducing a loss of function mutation *flc-3* (Michaels and Amasino, 1999).

Figure 12.2 demonstrates how this change in functionality of *FLC* dramatically alters the network of genes involved in the flowering time pathway. These networks are constructed with the multi-network programme encompassed in a software platform for plant systems biology called virtual plant (Gutiérrez *et al.*, 2005) using a publicly available microarray data set


(a)
ZTL
TOC1
CCA1
FKF1
LHY
GI
FRI
VRN1
CDF1
FVE
FLC
CO
miR156a,b,c,d,e,f,g
PNY
AGL42
SOC
FT
miR156h
FUL
SEP4
miR157a,b,c
SEP5
AP1
AGL24
SVP
SEP3
miR157d
SEP3
AG
LFY
ATMYB33
miR159c
SEP1
AP3
CAL
SEP2
PI
(b)
ZTL
TOC1
CCA1
FKF1
LHY
FRI
GI
VRN1
CDF1
FVE
FLC
CO
miR156a,b,c,d,e,f,g
PNY
SOC
FT
AGL42
miR156h
FUL
SEP5
AP1
miR157a,b,c
SEP4
AGL24
SVP
SEP3
miR157d
SEP3
AG
LFY
ATMYB33
miR159c
SEP1
AP3
CAL
SEP2
PI

on the early stages of transition to flowering (day 0 to day 7 after induction) (Schmid *et al.*, 2003). The multi-network tool embodied in the virtual plant platform incorporates database information on interactions derived from protein–protein interaction databases (BIND, DIP), protein–DNA interactions (Transfac), miRNA:RNA interactions, coexpression data and literature-based interactions predicted by a text mining tool called GeneWays, as described in Gutierrez *et al.* (2007). The resulting interaction network of flowering time genes shown in Fig. 12.2, shows that losing the functionality of *FLC* results in the addition of 14 edges and loss of 2 edges (denoted by two dashed edges in Fig. 12.2 a and 14 dashed edges in Fig. 12.2b). The loss of functionality of *FLC* also results in the connection of two additional nodes, AGL42 and FUL (compare Figs. 12.2 a and 12.2b). This relatively simple example of using a multi-network approach to depict the changes in interactions among flowering genes allows us to explore the effect of variation in functional genes and their interactions. This approach can then lead to novel biological questions that can be explored experimentally, like addressing the roles of *AGL42* and *FUL* in active *FRI-FLC* plants.

The variability in organismal transcriptomes that arises from natural genetic variants has been documented in a number of different species, including *D. melanogaster* (Jin *et al.*, 2001; Hsieh *et al.*, 2007), yeast (Townsend *et al.*, 2003), *A. thaliana* (Chen *et al.*, 2005; Juenger *et al.*, 2006; Keurentjes *et al.*, 2007; West *et al.*, 2007), *C. elegans* (Li *et al.*, 2006, 2007; Lee *et al.*, 2008) and the fish *Fundulus heteroclitus* (Oleksiak *et al.*, 2002; Whitehead and Crawford, 2006a, b; Burnett *et al.*, 2007). In one of the first studies examining expression variation, microarray analysis in *F. heteroclitus*, showed substantial levels of transcript variation within and between populations (Oleksiak *et al.*, 2002). A study in *D. melanogaster* completed in North Carolina and California populations also showed a bimodal (or at times multimodal) distribution of transcript abundance within populations, indicating variation segregates for gene expression levels (Hsieh *et al.*, 2007).

Several recent studies attempt to map genes that underlie variation in gene expression patterns across the entire genome (as reviewed by Gibson and Weir, 2005; Rockman and Kruglyak, 2007). These expression QTL (eQTL) studies have provided insights into the genetic architecture of natural

Figure 12.2 Comparison of the flowering gene networks of plants carrying active and inactive *FRI-FLC* pathway. A strong *FRI* allele from a wild accession Sf2 was introgressed into Col-0 that carries an active *FLC*, resulting in an active *FRI-FLC* pathway. The inactive *FRI-FLC* pathway was obtained by introducing a loss of function mutation *flc-3*. (a) The network with active *FRI-FLC* pathway (*FRI* Sf2, *FLC*). (b) The network with inactive *FRI-FLC* pathway (*FRI* Sf2, *flc-3*). The diamond nodes show transcription factors, the triangle nodes are micro RNAs and the round rectangle nodes are target genes. The black edges with arrowheads suggest transcriptional activation and the black edges with T-ends are transcriptional repression. The grey edges with no end marker are protein–protein interactions. The edges that are present in one network, but missing in the other, are indicated by dashed lines.

expression variation. A recent study in *A. thaliana* using a recombinant inbred line mapping population between the Bay-0 and Sha accessions reveal that approximately 1/3 of the differentially expressed genes were controlled by loci in *cis* (West *et al.*, 2007). Moreover, eQTLs that were regulated in *trans* map to several hotspots that regulated hundreds to thousands of transcript differences, suggesting that the genetic control of transcript level between individuals can be complex. Interestingly, almost all of the >36 000 eQTLs identified were associated with small phenotypic effects. Moreover, by combining eQTL methods with selection of candidate regulatory genes, one can reconstruct regulatory gene networks that are associated with natural variation in expression and other phenotypic traits (Keurentjes *et al.*, 2007).

The interaction of environment with natural genetic variation can provide a glimpse into the basis for expression (and ultimately phenotypic) plasticity in a systems biology context. An experiment with eQTL mapping in *C. elegans* under various growth temperatures demonstrated that nearly 60% of 308 *trans*-acting eQTLs showed a significant eQTL-by-environment interaction, while only 8% of 188 *cis*-acting genes showed an eQTL-by-environment interaction (Li *et al.*, 2007). This suggests that genetic differences in gene expression plasticity are largely regulated in *trans* such that expression variation in groups of genes are driven by individual loci.

Although, these studies were done under controlled laboratory conditions, the transcriptome can differ between genetically distinct *A. thaliana* accessions from different parts of the species range when grown under field conditions. Using microarray technology, the expression levels in the Bay-0 and Sha accessions were assayed at the four leaf seedling stage in a field site in Long Island, New York, in the fall of 2006 (K. Engelmann, D. Nielsen and M. Purugganan, unpublished data). The Bay-0 accession originally came from a fallow field in Bayreuth, Germany (50.0°N, 11.6°E at an altitude of 300–400 m) while Sha is from a mountainous site at Shahdara, Tajikistan (39.3°N and 68.3°E at an altitude of 3300–3400 m).

The correlation in gene expression between the two accessions is high ($r = 0.96$). But the distribution of *p*-values for differences between the two accessions in the field environment reveals that numerous genes are differentially expressed between the two accessions (Fig. 12.3). From this analysis, 401 genes display significant expression differences between Bay-0 and Sha in the field. Two hundred fifty-five (64%) of these genes are transcription factors or metabolic enzymes, including zinc-finger, homeodomain, bZIP, WRKY-type and myb-like transcription factors, the giberellin response factor RGA1, several auxin-induced proteins and fructose metabolism enzymes. One hundred forty-six genes (36%) encode hypothetical proteins or proteins of unknown function, and studies on their differential expression in ecological field environments may provide clues for further functional annotation of this class of genes. Which of these differences in gene expression, if any, account for the observed accession-specific differences in flowering time under field conditions, remains to be explored. These studies in the wild can also provide an

ecological context for understanding the real-world functions of genes and genetic networks, and permit an elucidation of the ecological transcriptome.

12.7 The future of ecological and evolutionary systems biology

At one level, systems biology combines genome-level interaction maps with dynamic modelling at the sub-genome level, where specified inputs and outputs allow the identification of key regulatory components or parameters of the system. Considering the complexities of developing these predictive models, systems biology must develop through close collaboration between experimental, computational and theoretical approaches. Albert and Assmann (see Chapter 1) point out that biochemical reactions within and between cells take place on timescales spanning several orders of magnitude (Papin *et al.*, 2005) and that these timescales are modulated by molecules or complexes and their interactions as well as the environmental conditions (Han *et al.*, 2004; Balázsi *et al.*, 2005). In addition to the dynamic changes in the state of network nodes, the characteristics of biological networks are shaped by dynamic events whose impact occurs on ecological and evolutionary timescales. As Albert and Asmann suggest, we argue that integration of ecological, evolutionary and epigenetic characteristics with transcriptional, metabolic and signal transduction networks is the 'final frontier' of systems biology.

The overarching goal at the intersection of systems biology, ecology and evolutionary biology is to evaluate whether the properties of biological networks as we depict them reflect reality at all levels of biology. Model systems with their full battery of genomic information, combined with the great amount of phenotypic, genetic and epigenetic variation they harbour, can provide a great deal of power to investigate how organisms are able to respond to their ecological milieu. The presence of genotype-by-environment interactions illustrate that the rapidly escalating amount of genomic data and tools applied to model systems in controlled conditions must also reflect natural variation between individuals, populations and species and the importance of understanding how molecular networks behave under real-world ecological conditions.

In addition, recent ecological genomic studies demonstrate that these approaches can be applied not just to well-known model species but to an even broader array of ecologically important organisms. Systems biology should be a critical component of fleshing out how organisms are able to respond to complex environments. This will require effort from all levels of the biological sciences, but should lead to a much more sophisticated understanding of the origins and functions of biological diversity. Moreover, in this world of changing environments, including the global climate, systems biology merged with ecology and evolution may provide predictive insights into adaptive responses of organisms to the future.

Acknowledgements

The authors would like to thank Aviv Madar, Gloria Coruzzi and Alexis Cruikshank for helpful comments and editing. This work was supported by an NSF FIBR grant (EF-0425759), an NSF Plant Genome Research Programme grant (DBI-0319553) and the Guggenheim Foundation.

References

Ainsworth, E.A., Rogers, A., Vodkin, L.O., Walter, A. and Schurr, U. (2006) The effects of elevated CO2 concentration on soybean gene expression. An analysis of growing and mature leaves. *Plant Physiol* **142**, 135–147.

Alba, R., Payton, P., Fei, Z.J., McQuinn, R., Debbie, P., Martin, G.B., *et al.* (2005) Transcriptome and selected metabolite analyses reveal multiple points of ethylene control during tomato fruit development. *Plant Cell* **17**, 2954–2965.

Aryoles, J.F. and Gibson, G. (2006) Analysis of variance of microarray data. *Meth Enzym* **411**, 214–223.

Ashikawa, I. (2001) Surveying CpG methylation at 5′-CCGG in the genomes of rice cultivars. *Plant Mol Biol* **45**, 31–39.

Bailey, R.A. (1981) A unified approach to design of experiments. *J R Stat Soc Ser A* **144** (2), 214–223.

Balázsi, G., Barabási, A.L. and Oltvai, Z.N. (2005) Topological units of environmental signal processing in the transcriptional regulatory network of Escherichia coli. *Proc Natl Acad Sci USA* **102**, 7841–7846.

Birnbaum, K., Shasha, D.E., Wang, J.Y., Jung, J.W., Lambert, G.M., Galbraith, D.W., *et al.* (2003) A gene expression map of the *Arabidopsis* root. *Science* **302**, 1956–1960.

Bonneau, R., Facciotti, M.T., Reiss, D.J., Schmid, A.K., Pan, M., Kaur, A., *et al.* (2007) A predictive model for transcriptional control of physiology in a free living cell. *Cell* **131**, 1354–1365.

Bossdorf, O., Richards, C.L. and Pigliucci, M. (2008) Epigenetics for ecologists. *Ecol Lett* **11**, 106–115.

Bruggeman, F.J. and Westerhoff, H.V. (2007) The nature of systems biology. *Trends Microbiol* **15**, 45–50.

Burn, J.E., Bagnall, D.J., Metzger, J.D., Dennis, E.S. and Peacock, W.J. (1993) DNA methylation, vernalization, and the initiation of flowering. *Proc Natl Acad Sci USA* **90**, 287–291.

Burnett, K.G., Bain, L.J., Baldwin, W.S., Callard, G.V., Cohen, S., Di Giulio, R.T., *et al.* (2007) Fundulus as the premier teleost model in environmental biology: opportunities for new insights using genomics. *Comp Biochem Physiol Part D* **2**, 257–286.

Caicedo, A.L., Stinchcombe, J., Schmitt, J. and Purugganan, M.D. (2004) Epistatic interaction between the *Arabidopsis* FRI and FLC flowering time genes establishes a latitudinal cline in a life history trait. *Proc Natl Acad Sci USA* **101**, 15670–15675.

Cardon, L.R. and Abecasis, G.R. (2003) Using haplotype blocks to map human complex trait loci. *Trends Genet* **19**, 135–140.

Cervera, M.-T., Ruiz-Garcia, L. and Martinez-Zapater, J. (2002) Analysis of DNA methylation in *Arabidopsis thaliana* based on methylation-sensitive AFLP markers. *Mol Genet Genom* **268**, 543–552.

Chen, W.Q.J., Chang, S.H., Hudson, M.E., Kwan, W.K., Li, J.Q., Estes, B., *et al.* (2005) Contribution of transcriptional regulation to natural variations in *Arabidopsis*. *Genome Biol* **6** (4), Art. No. R32.

Chen, Z.J. and Ni, Z. (2006) Mechanisms of genomic rearrangements and gene expression changes in plant polyploids. *Bioessays* **28**, 240–252.

Chen, Z.J. and Pikaard, C.S. (1997) Epigenetic silencing of RNA polymeraseI transcription: a role for DNA methylation and histone modification in nucleolar dominance. *Genes Dev* **11**, 2124–2136.

Clark, A.G., Eisen, M.B., Smith, D.R., Bergman, CM., Oliver, B., Markow, T.A., *et al.* (2007) Evolution of genes and genomes on the Drosophila phylogeny. *Nature* **450**, 203–218.

Cokus, S.J., Feng, S., Zhang, X., Chen, Z., Merriman, B., Haudenschild, C.D., *et al.* (2008) Shotgun bisulphite sequencing of the *Arabidopsis*genome reveals DNA methylation patterning. *Nature* **452**, 215–219.

Comai, L., Tyagi, A.P., Holmes-Davis, K.W.R., Reynolds, S.H., Stevens, Y. and Byers, B. (2000) Phenotypic instability and rapid genome silencing in newly formed *Arabidopsis* allotetraploids. *Plant Cell* **12**, 1551–1567.

Donohue, K., Dorn, L., Griffith, C., Kim, E., Aguilera, A., Polisetty, C.R., *et al.* (2005) Environmental and genetic influences on the germination of *Arabidopsis thaliana* in the field. *Evolution* **59**, 740–757.

Ellstrand, N.C. and Schierenbeck, K.A. (2000) Hybridization as a stimulus for the evolution of invasiveness. *Proc Natl Acad Sci USA* **97**, 7043–7050.

Fields, M.A., Schaeffer, S.M., Krech, M.J. and Brown, J.C.L. (2005) DNA hypomethylation in 5-azacytidine-induced early-flowering lines of flax. *Theor App Genet* **111**, 136–149.

Gibson, G. and Weir, B. (2005) The quantitative genetics of transcription. *Trends Genet* **21**, 616–623.

Grant-Downton, R.T. and Dickinson, H.G. (2005) Epigenetics and its implications for plant biology. 1. The epigenetic network in plants. *Ann Bot* **96**, 1143–1164.

Grant-Downton, R.T. and Dickinson, H.G. (2006) Epigenetics and its implications for plant biology. 2. The 'epigenetic epiphany': epigenetics, evolution and beyond. *Ann Bot* **97**, 11–27.

Gunsalus, K.C., Ge, H., Schetter, A.J., Goldberg, D.S., Han, J.D.J., Hao, T., *et al.* (2005) Predictive models of molecular machines involved in *Caenorhabditis elegans* early embryogenesis. *Nature* **436**, 861–865.

Gutierrez, R.A., Lejay, L.V., Dean, A., Chiaromonte, F., Shasha, D.E. and Coruzzi, G.M. (2007) Qualitative network models and genome-wide expression data define carbon/nitrogen-responsive molecular machines in *Arabidopsis*. *Genome Biol* **8**, R7.

Gutiérrez, R.A., Shasha, D.E. and Coruzzi, G.M. (2005) Systems biology for the virtual plant. *Plant Physiol* **138**, 550–554.

Han, J.D., Bertin, N., Hao, T., Goldberg, D.S., Berriz, G.F., Zhang, L.V., *et al.* (2004) Evidence for dynamically organized modularity in the yeast protein-protein interaction network. *Nature* **430**, 88–93.

Harmer, S.L., Hogenesch, L.B., Straume, M., Chang, H.S., Han, B., Zhu, T., *et al.* (2000) Orchestrated transcription of key pathways in *Arabidopsis* by the circadian clock. *Science* **290**, 2110–2113.

Hoffmann, M.H. (2002) Biogeography of *Arabidopsis thaliana* (L.) Heynh. (Brassicaceae). *J Biogeogr* **29**, 125–134.

Horvath, D.P., Schaffer, R., West, M. and Wisman, E. (2003) *Arabidopsis* microarrays identify conserved and differentially expressed genes involved in shoot growth and development from distantly related plant species. *Plant J* **34**, 125–134.

Hsieh, W.P., Passador-Gurgel, G., Stone, E.A. and Gibson, G. (2007) Mixture modeling of transcript abundance classes in natural populations. *Genome Biol* **8** (6), Art. No. R98.

Jablonka, E. and Lamb, M.J. (2005) *Evolution in Four Dimensions* (Cambridge, MA: MIT Press).

Jansson, S. and Douglas, C.J. (2007) *Populus*: a model system for plant biology. *Ann Rev Plant Biol* **58**, 435–458.

Jin, W., Riley, R.M., Wolfinger, R.D., White, K.P., Passador-Gurgel, G. and Gibson, G. (2001) The contributions of sex, genotype and age to transcriptional variance in *Drosophila melanogaster*. *Nat Genet* **29** (4), 389–395.

Johanson, U., West, J., Lister, C., Michaels, S., Amasino, R. and Dean, C. (2000) Molecular analysis of FRIGIDA, a major determinant of natural variation in *Arabidopsis* flowering time. *Science* **290**, 344–347.

Juenger, T.E., Wayne, T., Boles, S., Symonds, V.V., Mckay, J. and Coughlan, S.J. (2006) Natural genetic variation in whole-genome expression in *Arabidopsis thaliana*: the impact of physiological QTL introgression. *Mol Ecol* **15** (5), 1351–1365.

Kammenga, J.E., Herman, M.A., Ouborg, N.J., Johnson, L. and Breitling, R. (2007) Microarray challenges in ecology. *Trends Ecol Evol* **22**, 273–279.

Keurentjes, J.J.B., Fu, J.Y., Terpstra, I.R., Garcia, J.M., Van Den Ackerveken, G., Snoek, L.B., *et al.* (2007) Regulatory network construction in *Arabidopsis* by using genome-wide gene expression quantitative trait loci. *Proc Natl Acad Sci USA* **104** (5), 1708–1713.

Keyte, A.L., Percifield, R., Liu, B. and Wendel, J.F. (2006) Intraspecific DNA methylation polymorphism in cotton (*Gossypium* hirsutum L.). *J Hered* **97**, 444–450.

Kliebenstein, D., Pedersen, D., Barker, B. and Mitchell-Olds, T. (2002) Comparative analysis of quantitative trait loci controlling glucosinolates, myrosinase and insect resistance in *Arabidopsis thaliana*. *Genetics* **161**, 325–332.

Knox, M.R. and Ellis, T.H.N. (2001) Stability and inheritance of methylation states at PstI sites in Pisum. *Mol Genet Genom* **265**, 497–507.

Koornneef, M., Alonso-Blanco, C. and Vreugdenhil, D. (2004) Naturally occurring genetic variation in *Arabidopsis thaliana*. *Ann Rev Plant Biol* **55**, 141–172.

Lee, I., Bleecker, A. and Amasino, R. (1993) Analysis of naturally occurring late flowering in *Arabidopsis thaliana*. *Mol Gen Genet* **237**, 171–176.

Lee, I., Lehner, B., Crombie, C., Wong, W., Fraser, A.G. and Marcotte, E.M. (2008) A single gene network accurately predicts phenotypic effects of gene perturbation in *Caenorhabditis elegans*. *Nat Genet* **40**, 181–188.

Li, L., Wang, X.F., Stolc, V., Li, X.Y., Zhang, D.F., Su, N., *et al.* (2006) Genome-wide transcription analyses in rice using tiling microarrays. *Nat Genet* **38**, 124–129.

Li, Y., Lvarez, O.A.A., Gutteling, E.W., Tijsterman, M., Fu, J.J., Riksen, J.A.G., *et al.* (2007) Mapping determinants of gene expression plasticity by genetical genomics in *C-elegans. PLOS Genet* **2** (12), 2155–2161.

Littell, R.C., Milliken, G.A., Stroup, W.W., Wolfinger, R.D. and Schabenberger, O. (2006) *SAS for Mixed Models* (Cary, NC: SAS Publishing).

Liu, B., Brubaker, C.L., Mergeai, G., Cronn, R.C. and Wendel, J.F. (2001) Polyploid formation in cotton is not accompanied by rapid genomic changes. *Genome* **44**, 321–330.

Liu, B. and Wendel, J.F. (2003) Epigenetic phenomena and the evolution of plant allopolyploids. *Mol Phylogenet Evol* **29**, 365–379.

Mackay, T.F.C. (2001) The genetic architecture of quantitative traits. *Annu Rev Genet* **35**, 303–339.

Madlung, A., Masuelli, R.W., Watson, B., Reynolds, S.H., Davison, J. and Comai, L. (2002) Remodeling of DNA methylation and phenotypic and transcriptional changes in synthetic *Arabidopsis* allotetraploids. *Plant Physiol* **129**, 733–746.

Malmberg, R.L., Held, S., Waits, A. and Mauricio, R. (2005) Epistasis for fitness-related quantitative traits in *Arabidopsis thaliana* grown in the field and in the greenhouse. *Genetics* **171**, 2013–2027.

Mauricio, R. and Rausher, M.D. (1997) Experimental manipulation of putative selective agents provides evidence for the role of natural enemies in the evolution of plant defense. *Evolution* **51**, 1435–1444.

Mazzochi, F. (2008) Complexity in biology. *EMBO Rep* **9**, 10–14.

Michaels, S.D. and Amasino, R.M. (1999) FLOWERING LOCUS C encodes a novel MADS domain protein that acts as a repressor of flowering. *Plant Cell* **11**, 949–956.

Miyazaki, S., Fredricksen, M., Hollis, K.C., Poroyko, V., Shepley, D., Galbraith, D.W., *et al.* (2004) Transcript expression profiles of *Arabidopsis thaliana* grown under controlled conditions and open-air elevated concentrations of CO2 and of O3. *Field Crops Res* **90**, 47–59.

Mouradov, A., Cremer, F. and Coupland, G. (2002) Control of flowering time: interacting pathways as a basis for diversity. *Plant Cell* **14**, S111–S130.

Oleksiak, M.F., Churchill, G.A. and Crawford, D.L. (2002) Variation in gene expression within and among natural populations. *Nat Genet* **32** (2), 261–266.

Olsen, K.M., Haldorsdottir, S., Stinchcombe, J., Weinig, C., Schmitt, J. and Purugganan, M.D. (2004) Linkage disequilibrium mapping of *Arabidopsis* CRY2 flowering time alleles. *Genetics* **157**, 1361–1369.

O'Malley, M.A. and Dupré, J. (2005) Fundamental issues in systems biology. *Bioessays* **27**, 1270–1276.

Papin, J.A., Hunter, T., Palsson, B.O. and Subramaniam, S. (2005) Reconstruction of cellular signalling networks and analysis of their properties. *Nat Rev Mol Cell Biol* **6**, 99–111.

Pigliucci, M. (1996) How organisms respond to environmental changes: from phenotypes to molecules (and vice versa). *Trends Ecol Evol* **11**, 168–173.

Pigliucci, M. (1998) Ecological and evolutionary genetics of *Arabidopsis*. *Trends Plant Sci* **3**, 485–489.

Rapp, R.A. and Wendel, J.F. (2005) Epigenetics and plant evolution. *New Phytol* **168**, 81–91.

Richards, E.J. (2006) Inherited epigenetic variation – revisiting soft inheritance. *Nat Rev Genet* **7**, 395–401.

Riddle, N.C. and Richards, E.J. (2002) The control of natural variation in cytosine methylation in *Arabidopsis*. *Genetics* **162**, 355–363.

Riddle, N.C. and Richards, E.J. (2005) Genetic variation in epigenetic inheritance of ribosomal RNA gene methylation in *Arabidopsis*. *Plant J* **41**, 524–532.

Rockman, M.V. and Kruglyak, L. (2007) Genetics of global gene expression. *Nat Rev Genet* **7**, 862–872.

Salmon, A., Ainouche, M.L. and Wendel, J.F. (2005) Genetic and epigenetic

consequences of recent hybridization and polyploidy in Spartina (Poaceae). *Mol Ecol* **14**, 1163–1175.
Scheiner, S.M. and Gurevitch, J. (2001) *Design and Analysis of Ecological Experiments* (New York: Chapman and Hall).
Schlichting, C.D. and Smith, H. (2002) Phenotypic plasticity: linking molecular mechanisms with evolutionary outcomes. *Evol Ecol* **16**, 189–211
Schmid, M., Davison, T.S., Henz, S.R., Pape, U.J., Demar, M., Vingron, M., *et al.* (2005) A gene expression map of *Arabidopsis thaliana* development. *Nat Genet* **37**, 501–506
Schmid, M., Uhlenhaut, N.H., Godard, F., Demar, M., Bressan, R., Weigel, D., *et al.* (2003) Dissection of floral induction pathways using global expression analysis. *Development* **130**, 6001–6012.
Schmidt, D.D. and Baldwin, I.T. (2006) Transcriptional responses of *Solanum nigrum* to methyl jasmonate and competition: a glasshouse and field study. *Func Ecol* **20**, 500–508.
Shaked, H., Kashkush, K., Ozkan, H., Feldman, M. and Levy, A.A. (2001) Sequence elimination and cytosine methylation are rapid and reproducible responses of the genome to wide hybridization and allopolyploidy in wheat. *Plant Cell* **13**, 1749–1759.
Sharbel, T.F., Haubold, B. and Mitchell-Olds, T. (2000) Genetic isolation by distance in *Arabidopsis thaliana*: biogeography and postglacial colonization of Europe. *Mol Ecol* **9**, 2109–2118.
Sherman, J.D. and Talbert, L.E. (2002) Vernalization induced changes of the DNA methylation pattern in winter wheat. *Genome* **45**, 253–260.
Shimizu, K. and Purugganan, M.D. (2005) Evolutionary and ecological genomics of *Arabidopsis thaliana*. *Plant Physiol* **138**, 578–584.
Simpson, G.G. and Dean, C. (2002) *Arabidopsis*, the Rosetta stone of flowering time? *Science* **296**, 285–289.
Slotte, T., Holm, K., McIntyre, L.M., Lagercrantz, U. and Lascoux, M. (2007) Differential expression of genes important for adaptation in Capsella bursa-pastoris (Brassicaceae). *Plant Physiol* **145**, 160–173.
Sokal, R.R. and Rohlf, F.J. (1995) *Biometry* (New York: W.H. Freeman).
Stark, A., Lin, M.F., Kheradpour, P., Pedersen, J.S., Parts, L., Carlson, J.W., *et al.* (2007) Discovery of functional elements in 12 *Drosophila* genomes using evolutionary signatures. *Nature* **450**, 219–232.
Starker, C.G., Parra-Colmenares, A.L., Smith, L., Mitra, R.M. and Long, S.R. (2006) Nitrogen fixation mutants of Medicago truncatula fail to support plant and bacterial symbiotic gene expression. *Plant Phsyiol* **140**, 671–680.
Stinchcombe, J.R., Weinig, C., Ungerer, M., Olsen, K.M., Mays, C., Halldorsdottir, S., *et al.* (2004) A latitudinal cline in flowering time in *Arabidopsis thaliana* modulated by the flowering time gene FRIGIDA. *Proc Natl Acad Sci USA* **101**, 4712–4717.
Stratton, D.A. and Bennington, C.C. (1996) Measuring spatial variation in natural selection using randomly-sown seeds of *Arabidopsis thaliana*. *J Evol Biol* **9**, 215–228.
Swanson-Wagner, R.A., Jia, Y., DeCook, R., Borsuk, L.A., Nettleton, D. and Schnable, P.S. (2006) All possible modes of gene action are observed in a global comparison of gene expression in a maize F-1 hybrid and its inbred parents. *Proc Natl Acad Sci USA* **103**, 6805–6810.
Taylor, G., Street, N.R., Tricker, P.J., Sjodin, A., Graham, L., Skogstrom, O., *et al.* (2005) The transcriptome of *Populus* in elevated CO2. *New Phytol* **167**, 143–154.
Tian, D., Traw, M.B., Chen, J.Q., Kreitman, M. and Bergelson, J. (2003) Fitness costs of R-gene-mediated resistance in *Arabidopsis thaliana*. *Nature* **423**, 74–77.

Townsend, J.P., Cavalieri, D. and Hartl, D.L. (2003) Population genetic variation in genome-wide gene expression. *Mol Biol Evol* **20** (6), 955–963.

Travers, S.E., Smith, M.D., Bai, J., Hulbert, S.H., Leach, J.E., Schnable, P.S., *et al.* (2007) Ecological genomics: making the leap from model systems in the lab to native populations in the field. *Front Ecol Environ* **5**, 19–24.

Vaughn, M.W., Tanurdzic, M., Lippman, Z., Jiang, H., Carrasquillo, R., Rabinowicz, P.D., *et al.* (2007) Epigenetic natural variation in *Arabidopsis thaliana*. *PLoS Biol* **5**, 1617–1629.

Wang, Y.M., Lin, X.Y., Dong, B., Wang, Y.D. and Liu, B. (2004) DNA methylation polymorphism in a set of elite rice cultivars and its possible contribution to inter-cultivar differential gene expression. *Cell Mol Biol Lett* **9**, 543–556.

Weinig, C., Ungerer, M., Dorn, L.A., Kane, N.C., Halldorsdottir, S., Mackay, T.F.C., *et al.* (2002) Novel loci control variation in reproductive timing in *Arabidopsis thaliana* in natural environments. *Genetics* **162**, 1875–1881.

West, M.A.L., Kim, K., Kliebenstein, D.J., van Leeuwen, H., Michelmore, R.W., Doerge, R.W., *et al.* (2007) Global eQTL mapping reveals the complex genetic architecture of transcript-level variation in *Arabidopsis*. *Genetics* **175** (3), 1441–1450.

Whitaker, R.J. and Banfield, J.F. (2006) Population genomics in natural microbial communities. *Trends Ecol Evol* **21** (9), 508–516.

Whitehead, A. and Crawford, D.L. (2006a) Neutral and adaptive variation in gene expression. *Proc Natl Acad Sci USA* **103**, 5425–5430.

Whitehead, A. and Crawford, D.L. (2006b) Variation within and among species in gene expression: raw material for evolution. *Mol Ecol* **15**, 1197–1211.

Whitelaw, N.C. and Whitelaw, E. (2006) How lifetimes shape epigenotype within and across generations. *Hum Mol Genet* **15**, R131–R137.

Wilmes, P. and Bond, P.L. (2006) Metaproteomics: studying functional gene expression in microbial ecosystems. *Trends Microbiol* **14** (2), 92–97.

Wolfinger, R.D., Gibson, G., Wolfinger, E.D., Bennett, L., Hamadeh, H., Bushel, P.D., *et al.* (2001) Assessing gene significance from cDNA microarray expression data via mixed models. *J Comp Biol* **8**, 625–637.

Wu, C.A., Lowry, D.B., Cooley, A.M., Wright, K.M., Lee, Y.W. and Willis, J.H. (2008) *Mimulus* is an emerging model system for the integration of ecological and genomic studies. *Heredity* **100** (2), 220–230.

Xu, J.P. (2006) Microbial ecology in the age of genomics and metagenomics: concepts, tools, and recent advances. *Mol Ecol* **15** (7), 1713–1731.

Yang, S.H. and Loopstra, C.A. (2005) Seasonal variation in gene expression for loblolly pines (*Pinus taeda*) from different geographical regions. *Tree Physiol* **25**, 1063–1073.

Zhang, X., Clarenz, O., Cokus, S., Bernatavichute, Y.V., Pellegrini, M., Goodrich, J., *et al.* (2007) Whole-genome analysis of histone H3 lysine 27 trimethylation in *Arabidopsis*. *PLoS Biol* **5**, e129.

Zhang, X., Shiu, S., Cal, A. and Borevitz, J.O. (2008) Global analysis of genetic, epigenetic and transcriptional polymorphisms in *Arabidopsis thaliana* using whole genome tiling arrays. *PLoS Genet* **4**, e1000032.

Zilberman, D., Gehring, M., Tran, R.K., Ballinger, T. and Henikoff, S. (2007) Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription. *Nat Genet* **39**, 61–69.

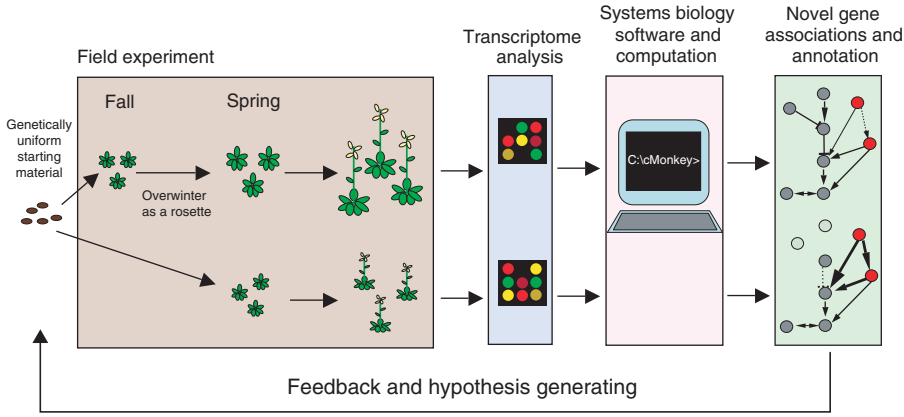


Figure 12.1 Depiction of an experiment that incorporates ecological experimental design with systems biology approaches. First, replicates of the same accession are grown in contrasting environments. Then, tissues of interest are collected at time intervals of interest and analyzed using genome-wide methods, such as microarrays. Bioinformatics can combine these genome-wide data with previously verified genetic interactions to learn putative gene interaction networks (GIN), suggesting novel associations. For example, in the figure, hypothetical novel genes are depicted as red in a known network of genes which are depicted as grey. These findings can then be used to generate new hypotheses and new experimental designs, thus completing one cycle of investigation. Iterating these procedures can lead to an extended and refined model of the GIN. (Adapted from Bossdorf *et al.*, 2008).

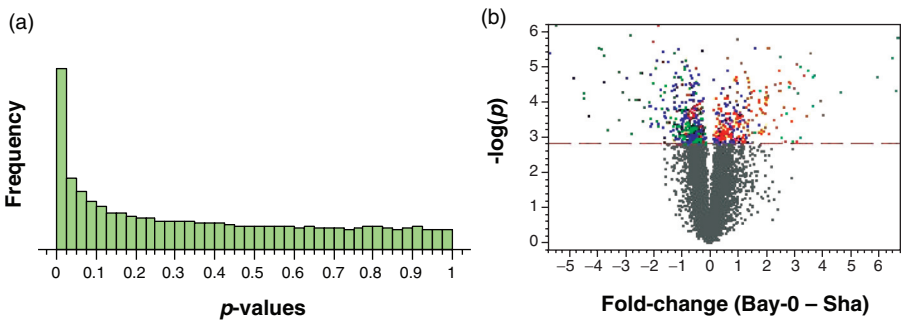


Figure 12.3 Microarray results for field-grown Bay-0 and Sha *A. thaliana* accessions after 3 weeks in the Cold Spring Harbor field site. (a) Distribution of *p*-values for the difference in expression between Bay-0 and Sha genes. (b) Volcano plot that indicates the expression fold-change of Bay-0 versus Sha (x-axis) and the corresponding log-transformed *p*-values for the expression difference (y-axis). The dashed line is the significance threshold with a 5% false discovery rate. The dots above the dashed line (in colour) are those genes that show significant difference in expression levels between Bay-0 and Sha in the field.