Plant, Cell and Environment (2005) **28**, 2–20

Gene function beyond the single trait: natural variation, gene effects, and evolutionary ecology in *Arabidopsis thaliana*

S. J. TONSOR[1], C. ALONSO-BLANCO[2,3] & M. KOORNNEEF[4,5]

[1]Department of Biological Sciences, University of Pittsburgh, 4249 Fifth Avenue, Pittsburgh PA 15260, USA, [2]Departamento de Biotecnología, Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria, Carretera de A Coruña, km 7, 28040, Madrid, Spain, [3]Departamento de Genética Molecular de Plantas, Centro Nacional de Biotecnología, CSIC, Campus de la Universidad Autónoma de Madrid, Cantoblanco, 28049 Madrid, Spain, [4]Laboratory of Genetics, Wageningen University, Arboretumlaan 4, 6703 BD Wageningen, The Netherlands and [5]Max Planck Institute for Plant Breeding Research, Carl-von-Linné-Weg 10, D-50892 Cologne, Germany

ABSTRACT

The purpose of plant functional genomics is to describe the patterns of gene expression and internal plant function underlying the ecological functions that sustain plant growth and reproduction. Plants function as integrated systems in which metabolic and developmental pathways draw on common resource pools and respond to a relatively small number of signal/response systems. Plants are also integrated with their environment, exchanging energy and matter with their surroundings and are consequently sensitive to changes in energy and resource fluxes. These two levels of integration complicate the description of gene function. Internal integration results in single genes often affecting multiple characteristics (pleiotropy) and interacting with multiple other genes (epistasis). Integration with the external environment leads to gene expression and the genes' phenotypic effects varying across environmental backgrounds (gene–environment interaction). An accurate description of the function of all genes requires an augmentation, already underway, of the study of isolated developmental and metabolic pathways to a more integrated approach involving the study of genetic effects across scales of variation usually regarded as the purview of ecological and evolutionary research. Since the evolution of gene function also depends on this complex of gene effects, progress in evolutionary genetics will also require understanding the nature of gene interactions and pleiotropy and the constraints and patterns they impose on adaptive evolution. Studying gene function in the context of the integrated organism is a major challenge, best met by developing co-ordinated research efforts in model systems. This review highlights natural variation in *A. thaliana* as a system for understanding integrated gene function in an ecological and evolutionary context. The current state of this research integration in *A. thaliana* is described by summarizing relevant approaches, current knowledge, and some potentially fruitful future studies. By introducing some of the fundamental questions of ecological and evolutionary research, experimental approaches and systems that can reveal new facets of gene function and gene effect are also described. A glossary is included in the Appendix.

Key-words: *Arabidopsis thaliana*; ecology; ecological genetics; evolutionary biology; evolutionary genomics; gene function; model organisms; quantitative trait locus; recombinant inbred lines.

INTRODUCTION

Functional genomics is greatly enriched when integrated with evolutionary genetics by enhanced use of natural variation in model systems such as *Arabidopsis thaliana*. Functional genomics has successfully identified functional roles for less than 10% of the 26 800 genes in the *A. thaliana* genome (Østergaard & Yanofsky 2004). This has been accomplished largely by gene function knockout through laboratory mutagenesis, followed by observation of altered phenotypes associated with specific pathways under standard laboratory or greenhouse conditions (Fig. 1a). Through observation of phenotypic variants associated with knockout mutants, together with clever use of a variety of molecular and classical genetic approaches, clues as to both the function and genomic location of the genes responsible for observed phenotypic variants can be gained. This process can ultimately lead to sequencing and cloning of the gene. This approach has focused largely on elucidating the genes involved in specific pathways thought to be of particularly high value. While the knockout mutants approach will remain an important, perhaps the most important, approach to gene discovery, additional approaches using natural variation are growing in their contribution. The use of natural variation is currently contributing to gene discovery, to fully describing the nature of gene function, to the understanding the genetic architecture of complex traits, and to understanding both internal integration of metabolism and development and integration of internal function with the external environment.

Despite the powerful knowledge gained from the laboratory study of specific pathways using knockout mutants, three observations point to looming limitations imposed

Correspondence: Stephen J. Tonsor. E-mail: tonsor@pitt.edu

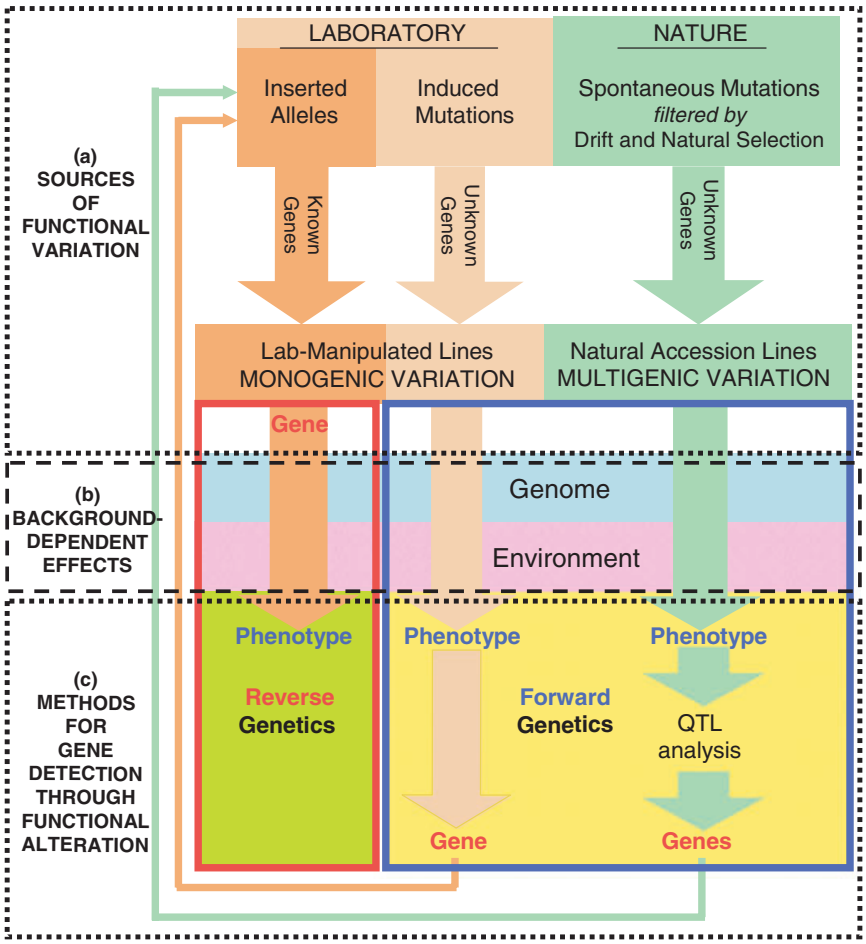

Figure 1. Flowchart illustrating approaches to gene function investigation in an evolutionary/ecological perspective. Gene discovery and function are elucidated through the comparison of genetic variants for loci controlling the trait of interest. Two sources of genetic variation are used to study of genetic architecture of complex phenotypes viewing the organism from the outside and attempting to derive the key attributes of the phenotype's genetic basis (forward genetics). (a) The traditional approach has been through the artificial generation of knockout mutations. This approach consists in comparing 'wild-type' and mutant lines differing in single genes (depicted in light brown). (b) In addition, spontaneous natural mutations influencing the trait can be found through the survey of natural variation. Genetically based differences in phenotype between two naturally occurring strains are multigenic, having arisen through the accumulation of multiple mutations through natural selection and genetic drift (these elements of the flow chart are delineated in green). Multi-generational crosses involving the two lines allow the identification of QTL (quantitative trait loci). This is followed by the identification of (some of the) the genes through NIL characterization and cloning. (c) Gene effects are typically dependent on both genetic and environmental backgrounds, indicated in blue and lavender. (d) Once the genes are identified, a refinement of the description of gene function is achieved by studying the phenotype from the inside out (reverse genetics). A more detailed examination of the identified gene's effects in the context of various genetic backgrounds can be accomplished through the insertion of the variant allele in a number of genetic backgrounds, indicated in darker brown. Quantification of the influence of environmental variation on gene effects can be accomplished by assaying those effects in a variety of natural and artificial environments.

by this approach. First, Bouché & Bouchez (2001) observe that a large proportion of known knockouts in *A. thaliana* have no obvious phenotypic effect. Mutagenic functional knockout of single genes therefore will not reveal the function for all genes. Second, the effect of a knockout on the phenotype can also be highly context-dependent, where both the genetic and environmental background can magnify or erase the effect. A third problem, resulting from the use of a limited range of phenotypic observations, is that a gene can have effects on multiple apparently unrelated characteristics. Many of these effects are indirect or well downstream from the pathway position of the gene product, calling for care in equating function with effect.

These three observations lead to the major foci of this paper: first, it is important to clarify the distinction between gene function and gene effect. Second, the effects of a single gene can be diverse and highly dependent on context. These complications of multiple effects and multiple dependencies are central areas of study in evolutionary and ecological genetics, and can greatly enhance our understanding of the integration of the phenotype and full

discovery of gene function. Third, exploration of gene function in the integrated plant calls for model systems, coordination across laboratories with complementary skills, and a well-developed set of tools for studying natural variation in gene function from a broad array of approaches. Natural variation in *A. thaliana* is the best-developed plant system for this kind of research.

Gene function and gene effect: beyond single-trait analysis

Gene function is used in both broad and narrow senses. Gene function in the broad sense is the effect of the gene on the phenotype[G.*]. Gene function in the narrowest sense is the biochemical or structural role played by the gene product. A full understanding of the role played by each of the genes in the plant genome will require study at three levels beyond the single gene's effect on a single trait: (1) the study of pleiotropy by the simultaneous analysis of multiple traits; (2) the study of variation in gene effects due to gene–genome interactions; and (3) the study of variation in gene effects due to gene–environment interactions. Although these are typically treated as noise in functional genomics, they can provide valuable additional insight into gene function.

Since gene 'function' in the broad sense can shift with genetic background and environment, describing a gene's 'function' in terms of phenotypic effect (e.g. gene 'for flowering time') can be problematic. We therefore distinguish between gene *function* and gene *effect*. We suggest that the description of a gene's *function* be constrained to the biochemistry of the gene product (e.g. an enzyme catalyses a particular reaction, or a chaperone assists in protein folding and maintenance of tertiary structure), and gene *effect* be used to describe the effect of a gene on the phenotype. The gene effect is really the ecology of the gene, namely the interaction between the gene product and its milieu, determining the effect of the gene on the whole organism and its performance.

The rapid rise in knowledge of the genome and ever more facile techniques in molecular genetics are increasingly being applied to studies of the genetic architecture of ecologically important traits. This work is revealing a complex pattern of relationship between gene function and gene effects. For both functional genomics in the narrow sense and for evolutionary and ecological genetics, pleiotropy[G.*], epistasis[G.*], and gene–environment interaction[G.*] are necessary and critical elements of study. Plant research is now entering an era in which these topics can finally be approached by combining the tools of molecular genetics, quantitative genetics, and evolutionary ecology. Key to this endeavour is the development of model systems in which natural variation in both environment and genome is used to examine the integration of both genetic and ecological function. In this paper, we review progress to date in the development of natural variation in *A. thaliana* as a study system for functional ecological and evolutionary genomics.

The advantages of studying natural variation in *A. thaliana*

Through the use of natural variation, additional functional variation can be discovered that both differs from and compliments information from laboratory-assayed, laboratory-generated mutants. Three aspects of natural variation are pertinent: natural genetic variation in a focal gene itself, natural variation in the genetic background, and natural variation in the environment in which the gene is expressed.

Natural genetic variation is often more subtle than laboratory-generated knock-out mutations and can therefore elucidate mechanisms of pathway control and cross-pathway linkages that may not be detectable with complete pathway knock-outs, especially in cases in which knockouts are lethal. Unlike laboratory-generated mutations, which are examined for their effects within a few generations of their occurrence, many tens to millions of generations have passed between the time most naturally occurring mutations arise and the time of their accession for study in functional genomics. Consequently, mutations underlying natural variation have experienced 'filtering' by natural selection[G.*]. Gene variation that exists in high frequency in nature is therefore more likely to be adaptive than laboratory-generated mutations. Natural mutations may also be followed by subsequent mutations that ameliorate negative pleiotropic[G.*] effects of earlier mutations (Lenski 1988; Smith & Macnair 1998) and these will be favoured by natural selection. Viewing mutations *in situ*, that is, in natural genomes that have evolutionarily accommodated those mutations, gives a richer view of integrated gene function and a more complete picture of the genetic architecture of plant function. This is especially true when we recombine the co-evolved background of a natural accession and a second genotype, as is done in recombinant inbred line (RIL) populations. RIL populations allow separate examination of the effects of multiple small regions of the genome that each influence a phenotypic trait[G.*].

Gene function investigations that employ multiple environments make it clear that new genes can be identified in new environmental conditions (e.g. Weinig *et al*. 2002). In addition to providing expanded techniques for understanding gene function at the organismal level, natural variation can provide unique information on gene function at the higher biological levels of organization such as populations or species, the subjects of ecology and evolution. Ultimately, gene function makes best sense when it is understood in the context of the living plant, namely in light of the gene's impact on survival and reproduction in the face of the challenges posed by the ambient environment.

Natural variation: increasing the toolset for investigating gene function

The tools used to analyse natural variation in *A. thaliana* build upon the analysis of the induced single gene mutants (Fig. 1). Analysis with induced single gene mutants employs two approaches, sometimes referred to as forward and

reverse genetic analysis. In forward analysis, altered phenotypes are used to infer the functional role of the mutation, and in combination with a number of additional assays and approaches, deduce the precise location of the gene (see Koornneef, Alonso-Blanco & Vreugdenhil 2004; Østergaard & Yanofsky 2004 for reviews of methods). The final step in gene function identification is the sequencing and cloning of the gene. Once the gene sequence is known or suspected, mutant variants of the gene can be inserted into 'wild-type' genotypes or well-described laboratory lines. This reverse genetic approach allows the effects of the gene to be determined with certainty.

The same approaches can be used with natural variation (Fig. 1). However, differences between wild accessions are nearly always under multigenic control and naturally varying environmental factors have large influences on phenotypic expression of gene effects, increasing the complexity of forward analyses. In general, gene mapping and gene cloning steps therefore require additional and often more complex analyses when working with natural variation. For the estimation of the (minimum) number of loci involved, their map positions and the size of their contribution to the genetic differences, QTL (quantitative trait locus) mapping procedures are applied (Sax 1923; Kao, Zeng & Teasdale 1999). In QTL mapping, trait values are tested for association with each of a large set of marker genotypes in which the markers are evenly distributed throughout the genome. Immortal mapping populations such as recombinant inbred lines (RIL) are very efficient for QTL analysis, since each line is genotyped for the marker loci only once, and multiple genetically identical individuals from each line can be grown in a limitless series of experiments [Note: mutations will accumulate over multiple generations of line culture]. Detailed descriptions of the general genetic and molecular techniques of QTL analysis and QTL cloning and specific approaches for *A. thaliana* are available elsewhere (Alonso-Blanco & Koornneef 2000; Yano 2001; Koornneef *et al*. 2004). Despite the increased complexity involved in the use of natural variation and QTL-based approaches, gene function and the evolution of genetic architecture will be understood much more fully as a result of their use (Purugganan & Gibson 2003).

Attributes of a model system for gene function

Arabidopsis thaliana (L.) Heyhn. is a small weedy herb in the mustard family (Brassicaceae or Cruciferae). The species occurs naturally nearly throughout the northern hemisphere at ruderal sites and at a variety of naturally open and/or disturbed sites. Although constrained largely to sites in which competition is minimal, it grows across an otherwise broad variety of site conditions. *Arabidopsis thaliana* has been found from sea level up to high in the Himalayas and from northern Scandinavia to Africa, including the Cape Verde Islands at 16° latitude, and the mountains of Tanzania and Kenya (Hoffmann 2002), as well as in northern North America where it was probably introduced from Europe. Thus *A. thaliana* is considered a cosmopolitan generalist species.

Arabidopsis thaliana was adopted as a model plant for a number of reasons that are relevant for investigations of the evolutionary ecology of gene function (Meinke *et al*. 1998). A large research community has generated an impressive body of knowledge on development and molecular biology of the species. The complete genome sequence is a major resource (AGI 2000). Efficient functional analysis is facilitated by the ease of genetic transformation, the availability of full genome microarrays, and the large collection of gene knock-out lines (Meinke *et al*. 2003). In addition extensive field-based collections are used to generate genetically diverse mapping populations that have been or are currently being characterized, and are available through a web-based distribution system (http://www.arabidopsis.org).

THE STUDY OF GENE EFFECT IN AN EVOLUTIONARY CONTEXT: FUNCTIONAL PLANT EVOLUTION

Gene functions evolve, but natural selection acts on gene effects

Understanding the evolution of gene function requires understanding gene function, phene[G,*] function, and the relationship between them. One might then think that enumeration and description of all of the genes of *A. thaliana* would finally enable an understanding of evolutionary change in natural populations. However, while the genes are the units of evolution, the phenotype is the unit of selection. The phenotype is the interface between the genetic programme and the environment, and the success of the phenotype determines the fate of the genes it contains. To the extent that there is a genetic basis to the phenotypic variation, the phenotypic distribution in the next generation will differ from the previous generation due to the functionally driven differences in survival and reproduction of genes that affect function, mediated through the function of the phenotypes in which the genes are found.

Gene effects are random variables, influenced by developmental noise and the environment

If genotypes mapped unequivocally onto phenotypes; that is, if a specific genotype always produced a specific phenotype, we could just skip consideration of the phenotype and make genotypes the sole focus in evolutionary studies. However, the phenotype is the outcome of the execution of the inherited, dynamic physiological/developmental programme that resides in the genome. The outcome is only partly predictable, because the programme's effects on survival and reproduction depend on chance and the environment as well as on the conditional expression of specific genes within the genome.

Chance plays a role in the effect of a gene in any instantiation of the genotype, even at the scale of transcript number per cell (Blake *et al*. 2003). The huge literature on fluctuating asymmetry, developmental instability and environmentally induced phenotypic variation leaves no doubt that noise is inherent in the mapping of genotypes on phenotypes, leading to only approximate developmental and metabolic homeostasis. Because of noise, gene effects have to be estimated with statistical descriptors, even when the underlying genes are known, especially for complex traits. Thus confidence in the magnitude of gene effects requires replication and statistical descriptors of confidence.

In addition to developmental noise, there is often a strong environmental context-dependency to gene effects (discussed in more detail below). The genetic effects that we estimate are therefore often only accurate under the conditions in which we measure them, unless we also have a clear predictive description (i.e. a forcing function) of the mechanisms causing environmental dependency (Lynch & Walsh 1998).

The variable nature of gene effects has been evident to quantitative geneticists for more than 80 years. Fisher, Wright and the other early statistical geneticists (see Provine 1971) showed how a statistical descriptor of heritability[G,*]- [i.e. at the simplest level, the breeding value of a gene and its variance (Fisher 1930)], allowed plant breeding and evolutionary studies to proceed despite noise in the mapping of genes on phenes, and despite profound ignorance of the 'black box' of molecular genetics and physiology. This approach has been fantastically successful in improving crop yields and food supplies through selective plant and animal breeding. The power of the approach derives from the fact that the statistical descriptions of selection and inheritance closely mimic the way the environment 'sees' the genes, noise and all. However, statistical genetic predictions of gene effects and evolutionary trajectories applied to the natural world or to long-term evolutionary dynamics have been inaccurate (see Roff 1997). The most important inaccuracies result from the same issues discussed above: lack of constancy in genetic effects across genetic and environmental backgrounds, and the extent of pleiotropy and its constancy. In addition, quantitative genetics has suffered from ignorance of underlying genetic, developmental and metabolic mechanisms, and of how variation in genetic architecture[G,*] is generated; that is, the effects of the natural mutational accumulation of variation on the distribution of genotypes available for the evolutionary process. Applying molecular biological tools to the description of the genetic architecture of complex traits and following the effects of genes through the plant's physiological processes and environmental interactions will open the black box, increasing understanding of real natural complexity in gene effects on the phenotype, and their evolution.

The identification of the genetic basis of key physiological and developmental traits is currently underway in a number of laboratories and creates a very exciting system for the study of evolutionary ecology. As knowledge of both the genetic architecture of complex traits and of gene effects matures, we will be able to address some fundamental and long-standing questions in evolutionary biology. These same questions will become increasingly important in addressing the full complexity of gene function.

Next, we describe existing information from studies of natural variation in *A thaliana*, and what it tells us about both gene function and evolutionary ecology. We first address current issues in evolutionary genetics that are also important understanding the evolution of genetic architecture and gene function, and progress toward addressing those issues using natural variation in *A. thaliana.* We then discuss key interactions of *A. thaliana* with its environment, reviewing what is known and suggesting fruitful future approaches. Finally, we discuss the limitations of natural variation in *A. thaliana* as a model system, and the potential for overcoming those limitations.

GENETIC ARCHITECTURE AND THE COMPLEXITIES OF GENE EFFECTS

Pleiotropy: simultaneous analysis of multiple traits

Pleiotropy, the influence of a single gene on the phenotypic value of more than one trait, is both importance to evolution and poorly understood. Pleiotropy creates genetic correlations between traits, and genetic correlations will alter or constrain the evolutionary response to selection (Lande & Arnold 1983; Houle 1994). However, the extent, causes, and evolutionary malleability of pleiotropy are poorly understood.

Pleiotropic effects can be quite complex. For instance the genes *FRIGIDA* (*FRI*) (Johanson *et al*. 2000) and *FLOWERING LOCUS C* (*FLC*) (Sheldon *et al*. 2000) were isolated from natural accessions on the basis of their effect on flowering time. Analysis of genetic lines differing in *FRI* and *FLC* alleles suggests that these loci have major effects on both flowering time and water use efficiency (McKay, Richards & Mitchell-Olds 2003). Comparative QTL mapping suggests that *FRI* and *FLC* affect nitrogen content (Loudet *et al*. 2003a). *FLC* has also been shown to influence circadian leaf movements (Swarrup *et al*. 1999). Another flowering time gene, *CRY02*, has been shown to also affect fruit length, ovule number per fruit and percentage unfertilized ovules (El-Assal *et al*. 2004). These genes '*for* flowering time' illustrate the difficulty of assigning a purpose to a gene based on phenotypic observations of single traits.

Unexpected pleiotropy is sometimes discovered as a result of a gene's effects on two different processes, resulting in one gene with two names. For instance abscissic acid (ABA)-related mutants of *A. thaliana* have appeared in many screens such as the *wiggum* and *era* mutants (Ziegelhoffer, Medrano & Meyerowitz 2000) the first being selected on the basis of the floral phenotype and the other on the basis of ABA responses. Unexpected pleiotropic effects can be due to previously unknown relationships

between pathways. These pleiotropic links can suggest future research aimed at gene function, physiology and development.

The extent and nature of pleiotropy is gene-dependent, as illustrated by seed pigmentation genes *TT1, TT2, TT1*2 and *BAN*. These genes are only expressed in one cell layer of the developing embryo (Debeaujon *et al*. 2003). By contrast, the biosynthesis genes *TT3* and *TT4* control pigmentation in all other tissues. Additional genes such as *TTG1* and *TTG2* influence not only pigmentation but also a variety of traits including trichome and root hair formation. Therefore, a given process can have genes specifically limited to that process, whereas other genes affect a multitude of additional processes. We can expect that genes will exhibit a diverse array of pleiotropic effects, from virtually none to fairly ubiquitous influences on the plant's metabolism, development and life history.

Under what circumstances might we expect a gene to exhibit pleiotropy? When genes regulate the important signal molecules such as plant hormones and photoreceptors, multiple pleiotropic effects are likely a *priori*. Pleiotropy can also result from shared substrate requirements for two pathways when the substrate is limiting. Limited shared substrate availability can result in genetically based changes in sink strength in one pathway altering flux through a co-limited pathway. Thus a gene whose function is in one pathway has pleiotropic effects on a second pathway. Increasing the availability of the shared substrate sufficiently would eliminate this pleiotropic effect, illustrating the potentially complex interplay between environmental and genetic effects, and again emphasizing the sometimes-slippery nature of the function/effect relationship.

To make matters more complicated, it is clear that pleiotropic effects themselves evolve. Ancient genes have been 'recruited' for new functions (Schmid *et al*. 2003) through amplification of a pleiotropic effect. Evolution can also lead to damping of pleiotropic effects. Mutations that are advantageous in their effects on one attribute can have strongly negative pleiotropic effects through other attributes, as with the *de novo* evolution of phage resistance (Burch & Chao 1999). The local genome is then likely to evolve ameliorative effects at other loci (Burch & Chao 1999; Whitlock & Otto 1999; Moore, Rozen & Lenski 2000).

No empirical systematic investigations of the distribution of pleiotropic effects for natural mutations exist for plants. The prediction of the extent and nature of pleiotropic effects for any particular gene remains poor, partly due to a lack of knowledge of the organization of gene networks in the organism (but see Hirai *et al*. 2004). Extending studies of modularity (Murren, Pendleton & Pigliucci 2002, Wagner 1996) with descriptions of mutational correlations in the organization of metabolism and development will greatly aid understanding of the generation of genetic variation through mutation and the constraints that pleiotropy imposes on adaptive evolution. Ultimately, categorization of a plant's traits into modules, or clusters of correlated traits, would greatly facilitate the study of the genetics of functional integration. This would be particularly valuable if indicator traits could be identified for each cluster. These should be traits for which there is low measurement effort, high repeatability and high correlation with other traits in the cluster. Natural accessions, NILs and mutants could then be screened for a fuller range of effects, allowing an exploration of the extent to which the correlations between traits observed in the $P^{G.*}$ and $G^{G.*}$ matrices result from patterns of correlated selection (hence environmental influences) versus patterns of mutational effects among traits (metabolic and developmental linkage).

Genetic context dependency in gene effects

Evolutionary genetic theory is largely based on additive$^{G.*}$, that is, non-interactive, effects of genes, because of the more simple pattern of inheritance of additive effects [and because even interacting genes often contribute primarily to additive genetic variance$^{G.*}$ (see Whitlock *et al*. 1995)]. Since gene combinations are shuffled during the recombinational events of sexual reproduction, the interaction effect is not inherited intact by the offspring, but only as function of the relative frequencies in the mating population of the alleles at interacting loci, making the evolution of interacting genes mathematically complex and thus far not fully explored theoretically. For at least 75 years, evolutionary geneticists have argued about the importance of gene interactions in the evolutionary process, with a paucity of theoretical or empirical information for resolution of the conflicts (Whitlock *et al*. 1995; Wade & Goodnight 1998; Coyne, Barton & Turelli 2000; Whitlock & Phillips 2000; Wolf, Brodie & Wade 2000). Because interacting genes can participate in qualitatively different patterns of adaptive evolution compared to purely additive genes, this remains one of the most important areas of evolutionary genetic research.

Interactions between genes are of two types, interlocus (epistasis) and intralocus (dominance). Epistasis$^{G.*}$ (itself of many types; see Phillips 1998) occurs when the effect of a certain allele of a gene at a locus 'A' depends on the alleles of another gene present at locus 'B'. Interaction between genes of the same locus is referred to as dominance. In *A. thaliana* dominance effects among natural alleles have only been addressed in a few flowering time analyses such as Koornneef *et al*. (1994); Lee *et al*. (1994), Alonso-Blanco *et al*. (1998). Consequently even less is known about dominance among QTL alleles than for epistasis. No published studies address epistasis or dominance in a general systematic way in *A. thaliana*. However, it is evident that epistasis is common as deduced from QTL mapping analyses. Most *A. thaliana* QTL studies that have tested for interactions between QTL have found epistasis, with amounts varying considerably among populations and analyses (Borevitz *et al*. 2002; Kliebenstein *et al*. 2002; Kobayashi & Koyama 2002; Loudet *et al*. 2003b).

Overall, gene interactive effects are generally underestimated for several reasons (Whitlock *et al*. 1995). The number of individuals of each of the necessary two-locus genotypes required for detecting epistasis is often low in

QTL studies using RIL populations. The presence of alleles of large effect at other loci can also interfere with the detection of the two-locus interactive effect. Finally, the multiple comparisons necessary to search for gene interactions requires that the significance threshold LOD value be increased, making detection of gene interactions even more difficult.

Comparison of gene–gene interactions among Arabidopsis RIL mapping populations will provide valuable insight into the nature of genetic variation. However, direct comparison of interactive effects between RIL populations are difficult, because differences in experimental environments, experimental design and analytical techniques cannot always be separated from differences in the actual genetic architecture. Natural populations may exhibit differentiated[G,*] co-adapted[G,*] genomes, in light of the high levels of selfing and low movement between populations observed in *A. thaliana.* It will be important to determine the extent to which the functional properties of an allelic variant either 'travel with' the allele as it is substituted into the diverse genetic backgrounds of other populations, or constitute a gene–genome interaction (see similar work by Wade 1984 in *Tribolium*). This can be tested by introgressing alleles into multiple backgrounds to develop NILs that then can be used to estimate additive and epistatic effects of allele substitutions and estimating the background-dependent variance in allelic effects (breeding value).

Environmental context-dependency in gene effects

Genotype–environment interaction (abbreviated G × E) is defined as a difference between genotypes in response to changes in environment (Lynch & Walsh 1998). Genotypes can differ in trait sensitivities to environmental variation, and thus in their responses to the environmental variable (Schlichting & Pigliucci 1998). The magnitude of G × E indicates the ability of a species or population to evolve an adaptive response to environmental variation. Although clearly of major importance in the diversification and adaptation of plants, there is considerable controversy regarding the genetic basis of G × E. QTL mapping in a number of experimental RIL populations detected G × E interactions and identified the QTL involved (QTL–environment interaction). For instance, many QTL influencing root characteristics (Rauh, Basten & Buckler 2002) shoot growth (Loudet *et al.* 2003b) and water and anion contents (Loudet *et al.* 2003a) are limited or modified in their effects depending on the available source of nitrogen. Gene effects on flowering time are also specific to particular controlled light environments differing in photoperiod (Maloof *et al.* 2001), or in other aspects of the environment (Weinig, Stinchcombe & Schmitt 2003).

The significance of gene effects on most traits can only be understood in the environment in which the genome in question evolved. Hoffman (2002) describes the range of environmental conditions in which *A. thaliana* has been collected. The standard growth conditions in laboratory and greenhouse *A. thaliana* studies fall outside or at the very edge of this natural range. Since gene–environment interaction is commonplace, one should be cautious when interpreting the meaning of the effects discovered in the laboratory. This is especially true with more complex and integrative traits. Adjusting the growth conditions to more closely match those in which *A. thaliana* evolved will improve this situation. A number of studies using natural accessions grown in ecologically contrasting sites to examine gene–environment interactions and the environmental dependency of QTL effects are appearing in print (e.g. Weinig *et al.* 2002, 2003). These studies yield much valuable insight but one cannot make adaptive interpretations regarding the evolution of the QTL unless the studies include the sites from which the accessions originated. In general, not enough is known about the sites of origin of current *A. thaliana* stock centres accessions to make this possible. New collections are in development with exact and accessible collection locations. The development of new genetic resources (RIL and NIL populations) from those well-documented accessions will definitely benefit our understanding of gene function in an ecological/evolutionary perspective.

The origin of variation: natural mutations and the distribution of their effects

A vast catalogue of mapped mutations is available for *A. thaliana* (Meinke *et al.* 2003). However, we have very little knowledge of the prevalent types of spontaneous mutations in natural plant populations, their relative rates of occurrence, the distribution of their phenotypic effects, or their importance in generating diversity relative to gene duplication-related processes or epigenetic modification. As a result, we use the distribution of artificial mutation effects to predict the characteristics of natural mutation accumulation.

We do have estimates of spontaneous mutational rates per locus of *A. thaliana*, which range from 1×10^{-4} to 1×10^{-6}, similar to estimates from other organisms (Klekowski 1992; Shaw, Byers & Darmo 2000). In addition, a mutation accumulation population of 120 lines has been developed starting from a single common Columbia (Col) ancestor (*Col*umbia is a originally wild-collected genetic line that is widely used in *Arabidopsis* laboratories). This population now includes 30 generations (R. Shaw, personal comm. to S.J.T.). By the sixteenth generation the lines had significantly diverged. The rate of divergence per generation appears to be non-linear, suggesting interactions among accumulating mutations. Surprisingly, the distribution of mutational effects on seed production at generation 16 was centred approximately on zero, meaning that mutations of positive and negative effect were about equally likely (Shaw *et al.* 2000). These results contrast with previous work on spontaneous mutations in *A. thaliana* (Schulz, Lynch & Willis 1999) that estimated an average negative mutational effect resulting in a 1% decline in fitness per generation. These contrasting results reflect the lack of a

general understanding of the distribution of naturally accumulating mutational effects (Shaw, Geyer & Shaw 2002). Questions of particular interest include: how much do the effects of a mutation depend on genetic and environmental background? What are the molecular bases of the prevalent kinds of spontaneous mutations in natural populations? How much does the distribution of effects depend on the kind of mutation? How much does it depend on the type of gene being mutated? How well are the effects of spontaneous natural mutations mimicked by laboratory-generated mutations of various kinds? What is the matrix of mutational effects (the M matrix^{G,*}), and how well does the M matrix predict the genetic variance–covariance ($G^{G,*}$) matrix? Does the G matrix reflect adaptation to the P matrix^{G,*} (the phenotypic variance–covariance matrix, the only matrix 'seen' by selection), and the vector of selection coefficients or more closely reflect the structure of the M matrix? These fundamental questions are now within reach by combining ecological, statistical, biochemical and molecular genetic approaches.

THE STUDY OF GENE EFFECT IN AN INTEGRATED CONTEXT: FUNCTIONAL PLANT ECOLOGY

Genetic effects on the properties of the individual are the lowest level of ecological functional gene effects, and it is at this most fundamental level that we have most knowledge. However, an ecological perspective on gene function extends the effects of the gene's functional properties up the ecological hierarchy, recognizing that any genetically based differences between populations are also gene effects, as are the characteristics of species within communities and ultimately the biological properties of ecosystems. This framework can provide mechanistic bases and genetic understanding of critically important ecological processes for the first time.

The preponderance of ecological work in *A. thaliana* is concentrated at the level of the individual and the interaction of the individual plant with individuals of other species that are trophically involved with *A. thaliana*. As functional ecological genomics matures, it may be possible to identify a suite of traits that together can account for most of the variation in fitness among populations. The areas discussed below represent the beginning of such a list, by identifying critical interactions between *A. thaliana* and its environment and accumulating information on the genetic basis of variations in those interactions. We also suggest some investigative approaches for the most readily addressed ecological interactions and issues in functional ecological genomics.

The individual is the interface between the genome and the environment

The individual plant is the basic unit of population biology. The individual's attributes result from a complex interplay between the effects of gene expression on metabolism and development, and feedback on the individual's ability to function in its environment. The individual's fitness (the number of offspring contributed to the next generation) depends on the success of metabolism and development in capturing and managing available energy and matter.

Understanding functional adaptation^{G,*} requires that we describe the relationship between functional variation and fitness. This must be done in a way that ensures: (1) all causes of fitness differences, anticipated and unanticipated, are integrated in the measured fitnesses of genetic variants; and (2) the description is made in the environmental background in which the organism evolved (Pigliucci 2003). Assessing the integrated fitness consequences is necessitated by the frequent occurrence of trade-offs between negative and positive fitness effects among correlated traits. These trade-offs are often unanticipated, and a mutation's positive fitness effects on a focal trait can be offset or even swamped by strong opposing fitness effects on correlated traits. Measuring fitness in the natural un-manipulated environment ensures a realistic estimate of the measured trait and its fitness consequences.

However, the drawback to the measurement of gene effects in the natural environment is that we can only conjecture about the specific environmental effector that causes the fitness effect, since two environments will differ in a number of ways, only one or some of which are causing the differences in effects. Certainty about causes behind function and fitness variation requires that additional experiments are conducted in which researchers control and manipulate both the environment and the phenotypic/genetic variation present. This approach is illustrated by the elegant experiments of Külheim, Ågren & Jansson (2002), exploring the function of the photosynthetic feedback de-excitation pathways (qE-type non-photochemical quenching) in an ecological context. Külheim *et al*. (2002) compared two laboratory-generated mutants, *npq1* and *npq4*, with the original 'wild-type' natural genotype (the Columbia strain). By growing the two mutants and Col in the field they demonstrated a clear fitness decrease associated with both mutations. By also comparing the genotypes in specific independent light manipulations, they demonstrated that protection against rapid onset of high light was the cause of the fitness advantage of the de-excitation pathway, rather than protection against high light *per se*. Similarly, the use of well-characterized laboratory mutants and 'wild-type' natural genotypes in the context of fitness differences is allowing the elucidation of adaptive gene function with regard to shade responses at ecological and evolutionary levels (Ballaré & Scopel 1997; Schmitt, Dudley & Pigliucci 1999; Munir *et al*. 2001). In the next section we describe recent ecological genomic work using natural variation in *A. thaliana*, classified in two general ecological topics: life history variation, and plant responses to abiotic and biotic factors.

The mechanistic basis of life-history variation

Life history studies describe the individual trait variation that determines demographic variation at the population

level, and its causes. Demography is an emergent characteristic of the population, influenced by each individual's life cycle events, such as key developmental shifts. Life history traits are defined as attributes of the individual plant that, when measured across a population, determine the elements of the life history transition matrix. *Arabidopsis thaliana* life history traits include the key timing traits such as timing of germination, timing of the onset of flowering and timing of senescence. Life history traits also include components of fecundity, such as flower number, seed number and seed size. The timing and conditions in which the plant allocates meristems and energy to key life history processes plays a critical role in determining rates of population growth or decline. Thus understanding the timing of life history events, their genetic control and fitness consequences, is a central challenge to ecology.

Seed dormancy and germination

Natural selection is thought to optimize germination timing in the face of environmentally imposed trade-offs. Increased growth and reproduction occurs when germination occurs early in benign early seasons while increased mortality can result when germination occurs early in stressful early seasons. Natural selection on the timing of germination varies in magnitude and sign across locations and years for other winter annuals (Kalisz 1986), and one might expect similar variation in selection among sites and years for *A. thaliana*. The timing of germination is under strong natural selection in *A. thaliana*, since manipulation of the germination timing showed that spring-germinating seeds died before reproducing, whereas autumn-germinating seeds performed best under the conditions of the experiment (Donohue 2002). Germination timing is complex and highly variable among naturally collected genotypes of *A. thaliana* and can be strongly affected by environmental factors such as low temperatures and light (Nordborg *et al*. 2002). All this variation makes possible comparisons of germination-related life history differences within a species and suggests why some populations are spring annual, while most are winter annuals. In addition, the ecological genetics of dormancy genes is further complicated by maternal variation, which can play a strong but environment-dependent role in seed dormancy/germination timing (Munir *et al*. 2001).

Seed dormancy properties can have a strong influence on fitness. If two genotypes have equal fecundity but one germinates immediately and one lays dormant for multiple seasons, the non-dormant genotype has the largest influence on population growth, since its generation time is shorter. However cross-season dormancy can decrease the risk of population extinction due to catastrophic conditions in some years. Thus we might expect an evolutionary optimization of dormancy in natural populations. Seed dormancy is influenced by at least seven QTLs such as those called *d*elay *o*f *g*ermination (*DOG*) genes (Van der Schaar *et al*. 1997; Alonso-Blanco *et al*. 2003). The production of NILs incorporating single and multiple dormancy alleles such as those described for the *DOG* genes (Alonso-Blanco *et al*. 2003) in the background genomes of two or more natural populations differing in selection regimes could be used to provide powerful data on the fitness effects and adaptive significance of this complex and agriculturally important trait.

Flowering time variation

Flowering time variation is clearly under stabilizing selection in many species because of a complex set of trade-offs involving flowering time. In *A. thaliana*, the time of bolting is genetically correlated with the number of rosette leaves. Genes conferring early flowering can decrease the primary productivity of the individual plant, resulting in lower leaf number and earlier seed production, but lower lifetime seed number (Mitchell-Olds 1996). However, when resource availability decreases (or stress increases) during the growing season, earlier flowering may be advantageous because only the early flowering plants are able to produce seed. The optimal timing will be determined by the rate of decrease of resource availability (or increase in stress) through time, the increase in early seed production and the decrease in total reproductive value^{G,*} associated with a given decrease in the time to onset of flowering. This optimum will clearly vary among environments that differ in the seasonality of resource availability and stress. Therefore, it is not surprising that natural populations of *A. thaliana* vary in the timing of the onset of flowering. The functional genomics of flowering time has been reviewed elsewhere (Simpson & Dean 2002; Mouradov, Cremer & Coupland 2002; Koornneef *et al*. 2004), and natural variation has been a key element of flowering time research (Koornneef *et al*. 2004). Two of the genes accounting for natural variation, *FRIGIDA* (*FRI*) (Johanson *et al*. 2000) and Flowering Locus C (*FLC*) (Michaels & Amasino 1999; Sheldon *et al*. 2000) have been particularly well studied. Significant variation in flowering time among 38 naturally collected genotypes was associated with disruptions in the *FRI* gene's open reading frame (ORF) (Johanson *et al*. 2000). Early flowering genotypes exhibited one or two deletions in the *FRI* ORF. The authors concluded that multiple loss-of-function mutations have provided the mechanistic basis for the evolution of early flowering populations. Le Corre, Roux & Reboud (2002) further examined sequence variation in *FRI*, comparing 25 natural collections from western Europe. They showed that the pattern of variation in synonymous and non-synonymous substitutions supports the hypothesis that there has been local evolution of flowering time, and that this evolution is concentrated in certain functional parts of the gene. Furthermore, two different alleles with reduced function of *FLC* have been identified, generated by the insertion of transposon-related sequences in the first intron (Michaels *et al*. 2003). Thus, it is shown that the early flowering behaviour of many Arabidopsis accessions is derived from later flowering ancestor accessions by loss of function of *FRI* and/or *FLC*.

In addition to *FRI* and *FLC*, numerous QTL have been identified in mapping analyses of seven different crosses between natural accessions. Some of them involve a late flowering and an early flowering accession (Clarke *et al*. 1995; Kuittinen, Sillanpää & Savolainen 1997), and identified *FRI* and *FLC* as epistatic loci. In contrast, other crosses involve two early accessions and these also identified several QTL as well as gene–gene interactions (Kowalski *et al*. 1994; Jansen *et al*. 1995; Loudet *et al*. 2003, Koornneef *et al*. 2004). Several populations have been analysed in multiple-controlled environments differing in light and temperature treatments (Stratton 1998). In all cases, significant interaction between the QTL effects and the environments have been found, showing a striking environmental dependency of QTL effects on flowering time. Weinig *et al*. (2002) planted the Col × Ler RILs in two field locations in 2 years and found large shifts in the contributions of QTL to flowering time variation among sites and seasons.

In each study cited above a small number of QTL of large effect occur, and there appears to be some overlap in the large-effect QTL across studies. Natural knockout mutations really do play a role in adaptive flowering time evolution, since much of the documented *FRI* variation is attributable to knockouts. These studies demonstrate that a large number of genes (at least 14) conditionally influence flowering time, depending on genetic and environmental background. The functional identities of the QTL of minor effect are virtually unknown. Natural accessions of *A. thaliana* vary in the genetic mechanisms adapting them to local selection regimes through modifications in flowering times (see Mouradov *et al*. 2002 for a discussion). Identification of the genes and description of the functions of their gene products will vastly expand our understanding of the mechanisms underlying variation in this important life history trait, its plastic responses, and its evolution.

The mechanistic bases of plant responses to abiotic factors

Stress can be defined in multiple conflicting ways (Hoffmann & Parsons 1991, 1997). Here we define a stress as an environmental factor that, at the levels experienced, reduces plant fitness compared to other levels of the same factor. Abiotic stress responses have been the main focus of plant stress research with *A. thaliana*, with light, temperature (Sung *et al*. 2003), drought and salt (Zhu 2002) receiving the greatest attention. In addition studies of other abiotic factors such as atmospheric CO_2 and soil metal stress are currently underway. Although progress has been made in identifying genes associated with responses to each of these stresses, very little is known about natural variation in stress responses in *A. thaliana.* To illustrate the ongoing work on the mechanistic basis of the responses to abiotic factors, below we describe some of the responses that are being studied.

Light quality responses

Arabidopsis responds to the ratio of red (R) and far-red (FR) light (which is affected by the proportion of direct sunlight versus indirect light filtered by leaf canopies) through the phytochrome photoreceptors, whose signals affect the expression of hundreds of genes (Tepperman *et al*. 2001). These phytochrome responses include architectural changes that influence plant height by increasing elongation and decreasing branching, and constitute the shade avoidance response (Smith 1995). The adaptive nature of this plastic response to neighbor shading has been demonstrated (Ballaré & Scopel 1997; Pigliucci 1999; Schmitt *et al*. 1999; Ballaré 2001). By testing more than 150 natural accessions of *A. thaliana* substantial variation for the R : FR responses as well as for the blue light responses among the accessions has been documented (Maloof *et al*. 2001; Botto & Smith 2002). The shade avoidance response appears to be genetically independent of the control of flowering time associated with R : FR ratio (Botto & Smith 2002). QTL mapping analyses (Borevitz *et al*. 2002; Botto *et al*. 2003) have identified several loci associated with hypocotyl elongation and cotyledon unfolding responses to different light qualities, indicating that variation for shade avoidance and other light responses are under multigenic control. These studies indicate that natural variation for hypocotyl responses is determined in some cases by variation in the phytochrome A (*PhyA*) gene, while variation for cotyledon unfolding involves the blue light photoreceptor gene *CRY2* (Botto *et al*. 2003). Several QTL co-locate with other light signalling genes but several do not. Isolation and sequencing of the genes are needed to illuminate the functional basis of their involvement.

Responses to novel atmospheres

Because of human-induced changes in atmospheric chemistry, responses to changing atmosphere have been a major research focus for more than a quarter-century (Sharma *et al*. 1979). Despite this, there has been very little work on the genetic basis of response variation, and the associated potential for adaptation to changing atmospheres. Several studies have been conducted in *Arabidopsis* showing the existence of genetic variation among natural populations for different responses to atmospheric CO_2 variation. Natural accessions sampled from the species' elevational range exhibit variation in dry weight gain in response to atmospheric CO_2 concentrations and respond to selection driven by atmospheric CO_2 content (Ward & Strain 1997, 2000). Other Arabidopsis accessions from throughout the species' geographic range exhibit variation for other responses to a doubling of atmospheric CO2, such as N composition, flowering time, mass allocation and reproduction, and the correlations among the traits shift significantly between atmospheres, suggesting shifts in the genetic architecture of the underlying traits (Tonsor, in review). In another study, 48 natural accessions of *Arabidopsis* were shown to exhibit nearly as much variation in stomatal density responses to

elevated CO_2 as did a broad comparison among species (Woodward, Lake & Quick 2002). Overall these studies indicate that responses to CO_2 concentration are strongly dependent on the environment, in particular to drought stress (Woodward *et al*. 2002) and N concentration [mutant studies: (Sun *et al*. 2002), natural accessions: (Zhang & Lechowicz 1995)] but still the mechanistic basis of this variation is poorly understood.

Heat responses

Heat stress responses are perhaps the best-studied stress responses. Heat shock proteins (HSPs), which are proteins transcriptionally induced upon exposure to moderate heat, function as molecular chaperones that protect the plant from high temperatures. Over-expression of HSP101 provides protection against high temperatures in *A. thaliana* (Queitsch *et al.* 2000) and appears to play a central role in acquired high-temperature tolerance (Hong & Vierling 2000). However, heat shock responses are genetically complex and involve a whole suite of expression responses affecting many traits in the organism. Arabidopsis natural accessions differ in their germination responses to heat (E. Vierling, personal comm.). Clerkx *et al*. (2004) found that three QTL explained 30% of germination variation following high-temperature exposure. Novel experiments using a RIL population and making use of pharmacological knockout of HSP90 resulted in large increases in variation among RIL lines in a broad variety of traits (Queitsch, Sangster & Lindquist 2002), suggesting that the level of expression of heat shock chaperones may modulate the phenotypically expressed level of genetic variation available for natural selection, with major implications for the evolutionary genetic response to heat stress (Knight & Mitchell-Olds 2002). Selection experiments in field-grown populations in which HSP-manipulated versus un-manipulated populations are compared for evolutionary responses would increase understanding of plant evolutionary responses to extreme environments. Heat shock response may elicit different responses in allocation and life history in different parts of the range of *A. thaliana*. The onset of high temperatures is likely to have significantly different consequences in the moist early summer of northern Europe versus the end of the wet season in parts of Spain, south-western Asia and northern Africa. Comparing whole-lifetime effects of heat shock for a variety of natural accessions could therefore provide a fascinating picture of regional adaptations to heat as a signal of seasonal change.

Generalized stress responses

In evolutionary ecology, there is longstanding interest in the possibility of generalized stress responses (Hoffman & Parsons 1991) and in the evolution of general stress tolerant versus stress intolerant plants (Hodgson *et al*. 1999; Stanton, Roy & Thiede 2000). Many of the gene products expressed in response to high temperature such as the small HSPs (sHSPs) are also expressed in response to other stresses, or appear to be effective in providing tolerance to multiple stresses (Sun & Verbruggen 2002). Small HSPs are a diverse group of proteins encoded by 19 open reading frames in *A. thaliana*. A number of sHSPs are expressed in response to water stress in sunflowers (Almoguera, Coca & Jordano 1996; Coca *et al*. 1996). A subset of sHSPs is induced by oxidative stress (Pla *et al*. 1998; Lee 2000; Sun *et al*. 2001). Others are induced by cold stress, heavy metals and ozone (reviewed by Sun & Verbruggen 2002). There is also accumulating evidence of common reliance on ABA and/or phospholipid signalling under multiple stresses (Zhu 2002; Narusaka *et al*. 2003). Although the evolution of a generalized stress response appears possible based on genetic expression data, thus far the evidence for it in whole-plant ecological studies is equivocal (Stanton *et al*. 2000). Analysis of multiple responses in Arabidopsis is identifying response regulators involved in several abiotic factors (Oono *et al*. 2003; Shinozaki, Yamaguchi-Shinozaki & Seki 2003). The combination of these studies with the identification of the genes accounting for the natural variation in *A. thaliana* will provide the lacking mechanistic basis of stress responses.

The plant–biota interface: plant population biology and community ecology

Many organisms that coexist with *A. thaliana* influence biomass accumulation and retention and in turn influence the fitness of both *A. thaliana* and the species that influence it. The application of genomics in the context of the interactions of *A. thaliana* with its neighbours will provide an important and largely lacking mechanistic basis for understanding plant community ecology. This work is only beginning and three areas of current work are highlighted here. These constitute the three interactions most ecologists would identify as the major biotic factors limiting the distribution and abundance of plant populations: interactions with pathogens, with herbivores and with competitors.

Plant interactions with pathogens

Interactions with pathogens are largely controlled by gene-for–gene interactions; a specific allele of a host resistance gene confers resistance to a specific virulence allele of the pathogen (Flor 1956). Much of the work on pathogen resistance in *Arabidopsis* has been done in relation the pathogen *Pseudomonas syringae*. *Pseudomonas syringae* is a pathogenic bacterium that infects natural populations of *A. thaliana* (Jakob *et al*. 2002), and induces a gene-for-gene virulence/resistance system. The *A. thaliana Rps2* gene was one of the first *Arabidopsis* resistance genes identified molecularly (Mindrinos *et al*. 1994) and, like R-genes in other plants (Jones & Jones 1997), possesses a leucine-rich repeat region that is thought to directly interact with the virulence gene product of *P. syringae* or related factors. *RPS2* shows high levels of polymorphism, with high rates of amino acid substitution per polymorphism and apparent diversifying selection (Caicedo, Schaal & Kunkel 1999).

Rps2 sequence variation in 27 natural accessions of *A. thaliana* provides evidence of natural selection maintaining variation in *Rps2* alleles and confirms the previously noted deep bifurcation in the gene's phylogeny: two major resistance/susceptibility clades appear to have been maintained by selection (Mauricio *et al*. 2003). Most resistance alleles are of ancient origin in *Arabidopsis* and persist over tens of millions of years (Stahl *et al*. 1999). Sequence analyses of natural variation of R genes show that they have rapid evolutionary rates consistent with a host–pathogen resistance/virulence arms race (Bergelson *et al*. 2001). Widespread polymorphism in both plant host resistance genes and pathogen (a) virulence genes has been classically explained as a putative trade-off between resistance to one virulence allele and susceptibility to others. As the pathogen evolves new virulence alleles, selection is thought to favour the spread of newly effective alternatives to the previously prevalent resistance alleles, thus establishing an evolutionary arms race in which new virulence and resistance genes replace old ones (see Holub 2001 for review). Current theories of polymorphism maintained through balancing selection assume a trade-off between resistance benefits and costs of the resistance system. Strong costs have been associated with the *Arabidopsis Rps2* resistance mechanism (Tian *et al*. 2003). Patterns of DNA sequence variation of R-genes in natural accessions of *A. thaliana* are consistent with a 'trench warfare' rather than 'red queen' version of the arms race, in which ancient alleles wax and wane in prevalence with shifts in pathogen virulence–allele frequencies.

Plant interactions with herbivores

Secondary chemistry plays a key role in protection against herbivores and pathogens (Kliebenstein 2004). Responses to insect herbivory are complex, and plant defences treat distinct herbivores in very different ways. A suite of major insect herbivores of *A. thaliana* has been identified. *Trichoplusia ni* and *Spodoptera littoralis* (generalists) and *Plutella xylostella* (a specialist) have been used extensively in studies of response to herbivory. Several QTL mapping studies in experimental populations derived from *A. thaliana* natural accessions have partially identified the genetic basis of herbivore resistance. Variation at the TASTY locus influences *T. ni* feeding preferences (Jander *et al*. 2001). The myrosinase-glucosinilate defense system (Bones & Rossiter 1996), partly induced by methyl jasmonate, participates in natural variation in insect resistance (Kliebenstein *et al*. 2001a, b, 2002). Several loci have been mapped in two RIL populations. One locus is associated with resistance to both *P. xylostella* and *T. ni*, while some QTL are specific to one or the other herbivore (Kliebenstein *et al*. 2002). Generalist resistance is associated with the myrosinase–glucosinilate characteristics of the RILs, while the resistance to the specialist herbivore *P. xylostella* is not. These results are in agreement with evolutionary theories suggesting that the specialist herbivores may overcome host plant chemical defences, while the generalist will be sensitive to them (Kliebenstein *et al*. 2002). *Arabidopsis thaliana* resistance to mammal (rabbit) herbivory has been recently analysed under natural environments, and is under multigenic control. However, the genes involved are distinct from those associated with resistance to insect herbivory (Weinig *et al*. 2003). Together, these studies show that the resistance to any particular herbivore is genetically complex and resistance to multiple herbivores appears to require multiple genetic systems.

Both pathogen resistance and herbivory resistance have intriguing trade-offs. For instance, ethylene signalling, involved in pathogen resistance, reduces resistance to the generalist *S. littoralis*, but does not affect resistance to the specialist *P. xylostella* (Stotz *et al*. 2000). In addition, salicylic acid (SA) signalling induced by ethylene, increases susceptibility to *P. xylostella* and to *T. ni* (Cui *et al*. 2002). There is cross talk between methyl jasmonate-induced insect resistance, necro-trophic bacterial pathogen and biotrophic bacterial pathogen resistance pathways (Traw *et al*. 2003). Field studies using natural variation and tracing the effects of the multiple genes involved will be an important step in understanding the complex evolutionary trajectories that will result from exposure to multiple herbivores and pathogens.

Plant interactions with competitors: resource uptake and use

Resource-based models of interspecific interaction address species coexistence and community composition on the basis of population-level relative resource draw-down, uptake and use efficiencies (Chase & Leibold 2003). This approach has explanatory power for both species interactions and diversity/stability/productivity relationships (Tilman & Lehman 2001), fuelling interest in the evolutionary genetics of community composition (Knight & Chase 2003). In spite of a long-standing and burgeoning literature on population- and community-level studies on resource-based competition, there has been little investigation of the link between variation in individual plant resource uptake/use properties and the outcome of competitive interactions.

Natural variation in *A. thaliana* is an excellent system for exploring these links. The mechanisms behind community-level questions can be initially explored in intraspecific comparisons in which specific physiological systems can be manipulated against a common genetic background. Genetic variation among natural accessions has been demonstrated in nitrogen uptake through both passive and active mechanisms (Rauh *et al*. 2002), nitrogen use efficiency (Loudet 2003b), properties related to water use efficiency (Loudet *et al*. 2003a, McKay *et al*. 2003) and photosynthetic properties and partitioning into above- and below-ground functions (S. Tonsor, unpublished). Study of the link between variation in resource uptake and use and competitive interactions has just started. Natural accessions of *A. thaliana* perform best with genetically identical competitors at present CO_2 concentrations, but with genetically diverse competitors at twice present day CO_2 concentra-

tions (Andalo, Goldringer & Godelle 2001). Brautigam & Tonsor (2002) showed that nitrogen uptake influenced competitive performance with other genotypes of *A. thaliana* under conditions differing in the available nitrogen and light intensity, while nitrogen use efficiency was a significant correlate with competitive performance only when nitrogen was limiting. Knowledge of the genetic architecture, physiology and biochemistry of resource uptake and use can provide insight into the mechanisms defining resource-based competition, as well as illuminating the constraints involved in the evolution of competitive ability.

CURRENT LIMITATIONS, NEEDS AND CAVEATS OF ARABIDOPSIS AS AN ECOLOGICAL AND EVOLUTIONARY STUDY SYSTEM

Limitations and needs

Although the momentum behind ecological and evolutionary genetic research with *Arabidopsis* is building at a rapid rate, some challenges remain.

1 Detailed local site information is lacking for most *A. thaliana* natural accessions, with only approximate locations recorded for most publicly available accessions. Therefore the detailed knowledge of the ecology of *A. thaliana* and the relationship between phenotype, genotype and native environmental attributes are unknown. This limits the ability to conduct common garden and reciprocal transplant experiments, a mainstay of functional ecology. A new set of accessions from well-characterized sites from throughout the native range of *A. thaliana* is needed. This is perhaps the most important step in making *A. thaliana* accessible as a model system for evolutionary ecology. The investment of effort to develop a standard set of environmental measures at the sites of new accessions will pay off in the long run. We suggest the following information be recorded: accurate GPS coordinates, elevation, slope, aspect, standard soil chemical and physical properties, percentage canopy closure, leaf area index, dominant plant species, directions for site access, site history and ownership.

2 Evolutionary and ecological functional genomics studies require a combination of observations under both natural field conditions and controlled conditions in the laboratory and greenhouse. As previously discussed, the results of laboratory and greenhouse studies will be most accurately interpretable if conditions mimic the natural environment closely (except of course where the plant is being purposefully perturbed). The establishment of a realistic range of standard growth conditions that mirror field conditions would have a large payoff in the long run.

3 Although RIL populations established from crosses between natural accessions have already provided exciting insight into natural variation in genetic architecture, a general picture of variation in gene function and gene effects cannot be obtained until many experimental segregating populations from a diverse array of natural accessions are available. The establishment of RIL populations and other experimental segregating marker populations founded from crosses between ecologically and geographically disparate crosses will be particularly valuable.Multiple experimental segregating populations (such as RIL populations) are needed, as well as other approaches to assessing genetic variation. A single experimental segregating population such as a RIL population can identify very valuable new alleles involved in plant responses. However it tells only a little about the evolutionary properties of the trait. QTL software typically calculates a 'heritability' estimate from RIL measurements, but this is a different quantity from the conventional narrow-sense heritability of evolutionary genetics. The conventional narrow-sense heritability in natural populations estimates the proportion of the population's phenotypic variance that is additive genetic, and is predicated on the assumption that you have adequately sampled the distribution of additive genetic effects within the population. In contrast, the heritability from an experimental segregating population typically estimates the genetic proportion of the total phenotypic variance from a cross between only *two* genotypes from *different* populations. The traditional heritability is based on random mating among many potentially diverse genotypes in a single population. Being nearly completely selfing, *A. thaliana* is anything but randomly mating, and as a result has low genetic variation within populations with most of the genetic diversity distributed among populations. Consequently, analyses of many experimental segregating populations will be necessary to assess genetic variation in *A. thaliana.* Second, an appropriate mathematical description of the evolutionary response to selection in *Arabidopsis* will differ from the conventional random-breeding model. Before we can describe expected responses to selection, it will be necessary to develop a metric that plays the same role in the highly selfing and widely dispersed *A. thaliana* that the narrow-sense heritability plays for populations with random breeding in ideal closed populations.

4 The previous point regarding the assessment of natural genetic variation in *A. thaliana* raises the question of what represents an evolutionary unit in *A. thaliana* in the wild. More information about the spatial scale of recombination and genetic structure in the wild is required. The application of spatial structure models [e.g. 'two-gener' analysis (Smouse, Dyer, Westfall & Sork 2001)], allows the assessment of genetic structure through both seed and pollen-mediated gene flow, and spatial autocorrelation analysis for the assessment of the decay in genotypic correlation with distance. This approach would greatly assist in developing a realistic model for the inheritance of genetic variation and the prediction of response to selection in *A. thaliana*. Work in other species (Podolsky & Holtsford 1995; Kittelson & Maron 2001) indicates that the molecular genetic data based on neutral variation within and among populations are poor gauges of

the degree of differentiation between populations in quantitative traits. Thus, information on geographic patterns of variation in both kinds of traits for *A. thaliana* would be valuable.

5 Finally, ecological and evolutionary genetic studies require very large sample sizes. Fortunately, with *A. thaliana* the ability to measure large numbers of plants for ecologically important traits in a short period of time reduces one of the crucial constraints in evolution and ecology studies. The development of high-throughput systems for outdoor and controlled-environment field genotyping and phenotyping, including physiological, developmental and metabolic whole-plant parameters, will greatly speed the use of *A. thaliana* as a model system in ecology and evolution.

Caveats

Regardless of the potential of *A. thaliana* for ecological and evolutionary studies, certain limitations remain.

1 Despite the broad geographic and ecological distribution of this generalist species, it grows in a limited range of ecological environments. Physiological and allocational variation within the species may not exist. Rather the variation could be due to variation in timing of life history events relative to the timing of local stresses.

2 *Arabidopsis* is virtually completely selfing, so it is a poor fit to population–evolutionary genetic models. This limits its appeal as a model, even if an appropriate analogue to the narrow-sense heritability of an ideal random-breeding closed population can be developed. Research on closely related outcrossers such as *A. lyrata* may largely obviate this limitation.

3 The species has an apparent star phylogeny, which suggests recent evolution. While recent evolution is not in itself problematic, a more complex intraspecific phylogeny would allow comparative studies within a phylogenetic context to be undertaken. Since this is becoming a tool of major importance in evolutionary ecology, a star phylogeny might be considered problematic in that phylogeographic and phylogenetic approaches may not yield as much information as they would in a phylogeny with branches rooted at various ages thus providing a variety of times since divergence and degrees of relationship (current investigations may modify this view of the phylogeny). Clearly *A. thaliana* cannot be used as a model to answer all questions in ecology and evolutionary biology. However, using the tools and knowledge developed in *A. thaliana* in studies of its congeners and closely related genera (Mitchell-Olds 2001) can provide a phylogenetic context to efficiently answer those questions in a diverse range of ecological and evolutionary situations. Current research with *A. thaliana* is producing a vast knowledge of gene function, gene effect as well as a suite of tools that will allow technology and knowledge transfer to a broad array of plants functioning in diverse ecosystems and roles.

Concluding remarks

Prodigious knowledge at the intra-organismal level is being accumulated in the wild species *A. thaliana* on its genome, physiology and development [more than 2000 publications per year (Somerville & Koornneef 2002]. In contrast, our mechanistic understanding of ecology and evolution is quite incomplete. Without expansion of this mechanistic understanding, we face frustrating limitations in the construction of an explanatory and theoretical framework for the causes of functional diversity, the evolutionary optimization of plant function, and its organization into communities and ecosystems. This intra-organismal information provides an ideal background to address ecological and evolutionary questions using *A. thaliana* as a model organism. Functional genomics, especially when used in the context of natural variation, can be a very powerful tool for addressing ecological and evolutionary questions, which in turn will have strong paybacks for understanding function and effects for all genes.

Progress in understanding the mechanisms underlying evolutionary and ecological processes requires that ecologists, evolutionary and molecular biologists enter into investigations lying outside their traditional areas of study. This will be accomplished by developing collaborative relationships among the various research areas, the combination of efforts reaping a reciprocal benefit. For instance the techniques and equipment currently available to work at the intra-organismal level (genes, metabolism, development and physiology) will be efficiently brought into an ecological and evolutionary perspective. In addition, as plant function is explored at increasingly integrated levels (at the whole-plant and plant–environment interface), function necessarily becomes less deterministic and more like a random variable. Understanding plant function will require the implementation of statistical approaches used in ecology and evolution. We expect that in the next decade we will witness an explosion of ecological and evolutionary studies using *A. thaliana*, which will illuminate classical eco-evo questions and raise new ones. It will also provide major input into the understanding of gene function in the integrated organism.

REFERENCES

Arabidopsis Genome Initiative (AGI). (2000) Analysis of the genome of the flowering plant *Arabidopsis thaliana*. *Nature* **408,** 796–815.

Almoguera C., Coca M.A. & Jordano J. (1996) *Plant Molecular Biology* **4,** 1093–1099.

Alonso-Blanco C. & Koornneef M. (2000) Naturally occurring variation in *Arabidopsis:* an underexploited resource for plant genetics. *Trends in Plant Science* **5,** 22–29.

Alonso-Blanco C., Bentsink L., Hanhart C.J., Vries H.B.E. & Koornneef M. (2003) Analysis of natural allelic variation at seed dormancy loci of *Arabidopsis thaliana*. *Genetics* **164,** 711–729.

Alonso-Blanco C., El-Assal E.-D., Coupland G. & Koornneef M. (1998) Analysis of natural allelic variation at flowering time loci in the Landsberg *erecta* and Cape Verde Islands ecotypes of *Arabidopsis thaliana*. *Genetics* **149,** 749–764.

Andalo C., Goldringer I. & Godelle B. (2001) Inter- and Intragenotypic competition under elevated carbon dioxide in *Arabidopsis thaliana*. *Ecology* **82,** 157–164.

Ballaré C.L. (2001) Arabidopsis mutants and other model systems in plant physiological ecology. *Trends in Plant Science* **6,** 1360–1385.

Ballaré C.L. & Scopel A.L. (1997) Phytochrome signalling in plant canopies: testing its population-level implications with photoreceptor mutants of Arabidopsis. *Functional Ecology* **11,** 441–450.

Bergelson J., Kreitman M., Stahl E.A. & Tian D. (2001) Evolutionary dynamics of plant *R*-genes. *Science* **292,** 2281–2285.

Blake W.J., Kaern M., Cantor C.R. & Collins J.J. (2003) Noise in eukaryotic gene expression. *Nature* **422,** 633–637.

Bones A.M. & Rossiter J.T. (1996) The myrosinase-glucosinolate sytem, its organization and biochemistry. *Physiologia Plantarum* **97,** 194–208.

Borevitz J.O., Maloof J.N., Lutes J., *et al.* (2002) Quantitative trait loci controlling light and hormone response in two accessions of *Arabidopsis thaliana*. *Genetics* **160,** 683–696.

Botto J.F., Alonso-Blanco C., Garzaron I., Sanchez R.A. & Casal J.J. (2003) The Cape Verde Islands allele of cryptochrome 2 enhances cotyledon unfolding in the absence of blue light in *Arabidopsis*. *Plant Physiology* **133,** 1547–1556.

Botto J.F. & Smith H. (2002) Differential genetic variation in adaptive strategies to a common environmental signal in Arabidopsis accessions: phytochrome-mediated shade avoidance. *Plant, Cell and Environment* **25,** 53–63.

Bouché N. & Bouchez D. (2001) Arabidopsis gene knockout: phenotypes wanted. *Current Opinion in Plant Biology* **4,** 111–117.

Brautigam P.S. & Tonsor S.J. (2002) Mechanisms of resource competition in *Arabidopsis thaliana.* In *Ecological Society of America Annual Meetings*. Abstract reference no. BRA–1015–606907. http://abstracts.co.allenpress.com/pwcb/esa2002/document/?ID=17582

Burch C.L. & Chao L. (1999) Evolution by small steps and rugged landscapes in the RNA virus ϕ6. *Genetics* **151,** 921–927.

Caicedo A.L., Schaal B.A. & Kunkel B.N. (1999) Diversity and molecular evolution of the *Rps2* gene in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the USA* **96,** 302–306.

Chase J.M. & Leibold M.A. (2003) *The Ecological Niche*. University of Chicago Press, Chicago, IL, USA.

Clarke J.H., Mithen R., Brown J.K.M. & Dean C. (1995) QTL analysis of flwoering time in *Arabidopsis thaliana*. *Molecular and General Genetics* **242,** 81–89.

Clerkx E.J.M., El-Lithy M.E., Vierling E., Ruys G.J., Blankestijn-De Vries H., Groot S.P.C., Vreugdenhil D. & Koornneef M. (2004) Analysis of natural allelic variation of *Arabidopsis* seed quality traits between the accessions Landsberg *erecta* and Shakdara, using a new recombinant inbred line population. *Plant Physiology* **135,** 1–12.

Coca M.A., Almoguera C., Thomas T.L. & Jordano J. (1996) *Plant Molecular Biology* **31,** 863–866.

Coyne J.A., Barton N.H. & Turelli M. (2000) Is Wright's shifting balance process important in evolution? *Evolution* **54,** 306–317.

Cui J.P., Jander G., Racki L.R., Kim P.D., Piercec N.E. & Ausubel F.M. (2002) Signals involved in *Arabidopsis* resistance to *Trichoplusia ni* caterpillars induced by virulent and avirulent strains of the phytopathogen *Pseudomonas syringae*. *Plant Physiology* **129,** 551–564.

Debeaujon I., Nesi N., Perez P., Devic M., Grandjean O., Caboche M. & Lepeniec L. (2003) Proanthocyanidin-accumulating cells in *Arabidopsis* testa: regulation of differentiation and role in seed development. *Plant Cell* **15,** 2514–2531.

Donohue K. (2002) Germination timing influences natural selection on life-history characters in *Arabidopsis thaliana*. *Ecology* **83,** 1006–1016.

El-Assal S.E.D., Alsonso-Blanco C., Peeeters A.J.M., Wagemaker C., Weller J.L. & Koornneef M. (2004) The role of cryptochrome 2 in flowering in *Arabidopsis*. *Plant Physiology* **134,** 539–539.

Fisher R.A. (1930) *The Genetical Theory of Natural Selection*. Clarendon Press, Oxford, UK.

Flor H.H. (1956) The complementary gene systems in flax and flax rust. *Advances in Genetics* **8,** 29–54.

Hirai M.Y., Yano M., Goodenowe D.B., Kanaya S., Kimura T., Awazuhara M., Arita M., Fujiwara T. & Saito K. (2004) Integration of transcriptomics and metabolomics for understanding of global responses to nutritional stresses in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the USA* **101,** 10205–10210.

Hodgson J.G., Wilson P.J., Hunt R., Grime J.P. & Thompson K. (1999) Allocating C-S-R plant functional types: a soft approach to a hard problem. *Oikos* **85,** 282–294.

Hoffmann A.A. (2002) Biogeography of *Arabidopsis thaliana*. *Journal of BioGeography* **29,** 125–134.

Hoffmann A.A. & Parsons P.A. (1991) *Evolutionary Genetics of Environmental Stress*. Oxford University Press, Oxford UK.

Hoffmann M.H. & Parsons P.A. (1997) *Extreme Environmental Change and Evolution*. Cambridge University Press, Cambridge, UK.

Holub E.B. (2001) The arms race is ancient history in *Arabidopsis*, the wildflower. *Nature Reviews Genetics* **2,** 516–527.

Hong S.W. & Vierling E. (2000) Mutants of *Arabidopsis thaliana* defective in the acquisition of tolerance to high temperature stress. *Proceedings of the National Academy of Sciences of the USA* **97,** 4392–4397.

Houle D. (1994) Genetic covariance of fitness correlates: What genetic correlations are made of and why it matters. *Evolution* **45,** 630–648.

Jakob K., Goss E.M., Van Araki H.T., Kreitman M. *et al.* (2002) Pseudomonas viridflava *and* P. syringae- *natural pathogens of* Arabidopsis thaliana. *Molecular Plant–Microbe Interactions* **15,** 1195–1203.

Jander G., Cui J.P., Nhan B., Peirce N.E. & Ausubel F.M. (2001) The tasty locus on chromosome 1 of *Arabidopsis* affects feeding of the insect herbivore *Trichoplusia ni*. *Plant Physiology* **126,** 890–898.

Jansen R.C., Van Ooijen J.W., Stam P., Lister C. & Dean C. (1995) Genotpe by environment interaction in genetic mapping of multiple quantitative trait loci. *Theoretical and Applied Genetics* **91,** 33–37.

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

Jones D.A. & Jones J.D.G. (1997) The role of leucine-rich repeat proteins in plant defenses. *Advances in Botanical Research* **24,** 89–167.

Kalisz S. (1986) Variable selection on the timing of germination in *Collinsia verna* (Scrophulariaceae). *Evolution* **40,** 479–491.

Kao C.H., Zeng Z.B. & Teasdale R.D. (1999) Multiple interval mapping for quantitative trait loci. *Genetics* **152,** 1203–1216.

Kittelson P.M. & Maron J.L. (2001) Fine-scale genetically-based differentiation of life-history traits in the perennial shrub Lupinus arboreus. *Evolution* **55,** 2429–2438.

Klekowski E.J.J. (1992) Review: mutation rates in diploid annuals-are they immutable? *International Journal of Plant Science* **153,** 462–465.

Kliebenstein D.J. (2004) Secondary metabolites and plant/environment interactions: a view through *Arabidopsis thaliana* tinged glasses. *Plant, Cell and Environment* **27,** 675.

Kliebenstein D.J., Kroymann P., Brown A., Figuth D., Pedersen D., Gershenzon J. & Mitchell-Olds T. (2001a) Genetic control of natural variration in *Arabidopsis* glucosinolate accumulation. *Plant Physiology* **126,** 811–825.

Kliebenstein D.J., Lambrix V.M., Reichelt M., Gershenzon J. & Mitchell-Olds T. (2001b) Gene duplication in the diversitfication of secondary metabolism: Tandem 2-oxoglutarate-dependent dioxygenases control flucosiinolate biosynthesis in *Arabidopsis*. *Plant Cell* **13,** 681–693.

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

Knight T.M. & Chase J.M. (2003) Community genetics: toward a synthesis. *Ecology* **84,** 580–582.

Knight C.A. & Mitchell-Olds T. (2002) Chaperones as buffering agents? *Science* **296,** 2348–2349.

Kobayashi Y. & Koyama H. (2002) QTL analysis of Al tolerance in recombinant inbred lines of *Arabidopsis thaliana*. *Plant and Cell Physiology* **43,** 1526–1533.

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

Koornneef M., Blankestijn-De Vries H., Hanhart C.J., Soppe W.J.J. & Peeters A.J.M. (1994) The phenotype of some late-flowering mutants is enhanced by a locus on chormosome 5 that is not effective in the Landsberg *erecta* wild-type. *Plant Journal* **6,** 911–919.

Kowalski S.P., Lan T.H., Feldmann K.A. & Paterson A.H. (1994) QTL mapping of naturally-occuring variation in flowering time of *Araboidopsis thaliana*. *Molecular and General Genetics* **245,** 548–555.

Kuittinen H., Sillanpää M.J. & Savolainen O. (1997) Genetic basis of adaptation: flowering time in *Arabidopsis thaliana*. *Theoretical and Applied Genetics* **95,** 573–583.

Külheim C., Ågren J. & Jansson S. (2002) Rapid regulation of light harvesting and plant fitness in the field. *Science* **297,** 91–93.

Lande R. & Arnold S.J. (1983) The measurement of selection on correlated traits. *Evolution* **37,** 1210–1226.

Le Corre V., Roux F. & Reboud X. (2002) DNA polymorphism at the *FRIGIDA* gene in *Arabidopsis thaliana:* extensive nonsynonymous vairation is consistent with local selection for flowering time. *Molecular Biology and Evolution* **19,** 1261–1271.

Lee I., Aukerman M.J., Lohman K.N., Michaels S.D., Weaver L.M. *et al.* (1994) Isolation of *LUMINIDEPENDENS*: a gene involved in the control of flowering time in *Arabidopsis*. *Plant Cell* **6,** 75–83.

Lenski R.E. (1988) Experimental studies of pleiotropy and epistasis in *Escherichia coli*.2. Compensation for maladaptive effects associated with resistance to virus-T4. *Evolution* **42,** 433–440.

Loudet O., Chaillou S., Krapp A. & Daniel-Vedele F. (2003a) Quantitative trait loci analysis of water and anion contents in interaction with nitrogen availability in *Arabidopsis thaliana*. *Genetics* **163,** 711–722.

Loudet O., Chaillou S., Merigout P., Talbotec J. & Daniel-Vedele F. (2003b) Quantitative trait loci analysis of nitrogen use efficiency in *Arabidopsis*. *Genetics* **131,** 345–358.

Lynch M. & Walsh B. (1998) *Genetics and Analysis of Quantitative Traits*. Sinauer Associates, Sunderland, MA, USA.

Maloof J.N., Borevitz J.O., Dabi T., *et al.* (2001) Natural variation in light sensitivity of *Arabidopsis*. *Nature Genetics* **29,** 441–446.

Mauricio R., Stahl E.A., Korves T., Tian D., Kreitman M. & Bergelson J. (2003) Natural selection for polymorphism in the disease resistance gene *Rps2* of *Arabidopsis thaliana*. *Genetics* **163,** 735–746.

McKay J.K., Richards J.H. & Mitchell-Olds T. (2003) Genetics of drought adaptation in *Arabidopsis thaliana*: I. pleiotropy contributes to genetic correlations among ecological traits. *Molecular Ecology* **12,** 1137–1151.

Meinke D.W., Cherry J.M., Dean C., Rounsley S.D. & Koornneef M. (1998) *Arabidopsis thaliana*: a model plant for genome analysis. *Science* **282,** 662–663.

Meinke D.W., Meinke L.K., Showalter T.C., Schissel A.M., Mueller L.A. & Tzafrir I. (2003) A sequence-based map of Arabidopsis genes with mutant phenotypes. *Plant Physiology* **131,** 409–418.

Michaels S.D. & 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.

Michaels S.D., Hey Y.H., Scortecci K.C. & Amasino R.M. (2003) Attenuation of FLOWERING LOCUS C activity as a mechanism for the evolution of summer-annual flowering behavior in *Arabidopsis*. *Proceedings of the National Academy of Sciences of the USA* **100,** 10102–10107.

Mindrinos M., Katagiri F., Yu G.L. & Ausubel F.M. (1994) The *A. thaliana* disease resitance gene *RPS-2* encodes a protein containing a nucleotide-binding site and leucine-rich repeats. *Cell* **78,** 1089–1099.

Mitchell-Olds T. (1996) Genetic constraints on life history evolution: quantitative trait loci influencing growth and flowering in *Arabidopsis thaliana*. *Evolution* **50,** 1849–1858.

Mitchell-Olds T. (2001) *Arabidopsos thaliana and* its wild relatives: a model system for ecology and evolution. *Trends in Ecology and Evolution* **16,** 693–700.

Moore F.B.-G., Rozen D.E. & Lenski R.E. (2000) Pervasive compensation for deleterious mutations in experimental populations of *E. coli*. *Proceedings of the Royal Society of London B* **267,** 515–522.

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

Munir J., Dorn L.A., Donohue K. & Schmitt J. (2001) The effect of maternal photperiod on seasonal dormancy in *Arabidopsis thaliana*. *American Journal of Botany* **88,** 1240–1247.

Murren C.J., Pendleton N. & Pigliucci M. (2002) Evolution of phenotypic integration in Brassica. *American Journal of Botany* **89,** 655–663.

Narusaka Y., Nakashima K., Shinwari Z.K., *et al.* (2003) Interaction between two *cis*-acting elements, ABRE and DRE, in ABA-dependent expression of *Arabidopsis rd29A* gene in response to dehydration and high-salinity stresses. *Plant Journal* **34,** 137–148.

Nordborg M., Borevitz J.O., Bergelson J., *et al.* (2002) The extent of linkage disequuilibrium in *Arabidopsis thaliana*. *Nature Genetics* **30,** 190–193.

Oono Y., Seki M., Nanjo T., *et al.* (2003) Monitoring expression profiles of *Arabidopsis* gene expressionduring rehydration process after dehydration using *ca.* 7000 full-length cDNA microarray. *Plant Journal* **34,** 868–887.

Østergaard L. & Yanofsky M.F. (2004) Establishing gene function by mutagenesis in *Arabidopsis thaliana*. *Plant Journal* **39,** 682–696.

Phillips P.C. (1998) The language of gene interaction. *Genetics* **149,** 1167–1171.

Pigliucci M. (2003) Selection in a model system: ecological genetics of flowering time in *Arabidopsis thaliana*. *Ecology* **84,** 1700–1712.

Pigliucci M.S. & Schmitt J. (1999) Genes affecting phenotypic plasticity in *Arabidopsis:* pleiotropic effects and reproductive fitness of photmorphogenic mutants. *Journal of Evolutionary Biology* **12,** 551–562.

Pla M., Huguet G., Verdaguer D., Puigderrajols P., Llompart B., Nadal A. & Molinas M. (1998) Stress proteins co-expressed in suberized and lignified cells and in apical meristems. *Plant Science* **139,** 49–57.

Podolsky R.H. & Holtsford T.P. (1995) Population structure of mophological traits in *Clarkia dudleyana* 1. Comparison of F_{ST} between allozymes and morphological traits. *Genetics* **140,** 733–744.

Provine W. (1971) *The Origins of Population Genetics*. University of Chicago Press, Chicago, IL, USA.

Purugganan M. & Gibson G. (2003) Merging ecology, molecular evolution and functional genetics. *Molecular Ecology* **12,** 1109–1112.

Queitsch C., Hong S., Vierling E. & Linquist S. (2000) Heat shock protein 101 plays a crucial role in thermal protection in *Arabidopsis thaliana*. *Plant Cell* **12,** 479–492.

Queitsch C., Sangster T.A. & Lindquist S. (2002) Hsp90 as a capacitor of phenotypic variation. *Nature* **417,** 618–624.

Rauh B.L., Basten C. & Buckler E.S. (2002) Quantitative trait loci analysis of growth response to varying nitrogen sources in *Arabidopsis thaliana*. *Theoretical and Applied Genetics* **104,** 743–750.

Roff D. (1997) *Evolutionary Quantitative Genetics*. Chapman & Hall, London, UK.

Sax K. (1923) The association of size differences with seed-coat pattern and pigmentation in *Phaseolus vulgaris*. *Genetics* **8,** 552–560.

Schlichting C.D. & Pigliucci M. (1998) *Phenotypic Evolution: a Reaction Norm Perspective*. Sinauer Associates, Sunderland, MA, USA.

Schmid K.J., Sorensen T.R., Stracke R., Torjek O., Altmann T., Mitchell-Olds T. & Weisshaar B. (2003) Large-scale identification and analysis of genome-wide single-nucleotide polymorphisms for mapping in *Arabidopsis thaliana*. *Genome Research* **13,** 1250–1257.

Schmitt J., Dudley S.A. & Pigliucci M. (1999) Manipulative approaches to testing adaptive plasticity: Phytochrome-mediated shade-avoidance responses in plants. *American Naturalist* **154,** S43–S54.

Schulz S.T., Lynch M. & Willis J.H. (1999) Spontaneous deleterious mutation in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the USA* **96,** 11393–11398.

Sharma R.K., Griffing B. & Scholl R.L. (1979) Variation among races of *Arabidopsis thaliana* (L.) Heynh for survival in limited carbon dioxide. *Theoretical and Applied Genetics* **54,** 11–15.

Shaw R.G., Byers D.L. & Darmo E. (2000) Spontaneous mutational effects on reproductive traits of *Arabidopsis thaliana*. *Genetics* **155,** 369–378.

Shaw F.H., Geyer C.J. & Shaw R.G. (2002) A comprehensive model of mutations affecting fitness and inferences for *Arabidopsis thaliana*. *Evolution* **56,** 453–463.

Sheldon C.C., Rouse D.T., Finnegan E.J., Peacock W.J. & Dennis E.S. (2000) The molecular basis of vernalization: The central role of FLOWERING LOCUS C (FLC). *Proceedings of the National Academy of Sciences of the USA* **97,** 3753–3758.

Shinozaki K., Yamaguchi-Shinozaki K. & Seki M. (2003) Regulatory network of gene expression in the drought and cold stress responses. *Current Opinion in Plant Biology* **6,** 410–417.

Simpson G.G. & Dean C. (2002) Flowering – *Arabidopsis,* the rosetta stone of flowering time? *Science* **296,** 285–289.

Smith H. (1995) Physiological and ecological function within the phytochrome family. *Annual Review of Plant Physiology and Plant Molecular Biology* **46,** 289–315.

Smith S.E. & Macnair M.R. (1998) Hypostatic modifiers cause variation in degree of copper tolerance in *Mimulus guttatus*. *Heredity* **80,** 760–768.

Smouse P.E., Dyer R.J., Westfall R.D. & Sork V.L. (2001) Two-generation analysis of pollen flow across a landscape. I. Male game heterogeneity among females. *Evolution* **55,** 260–271.

Somerville C. & Koornneef M. (2002) A fortunate choice: the history of *Arabidopsis* as a model plant. *Nature Reviews in Genetics* **3,** 883–889.

Stahl E., Dwyer G., Mauricio R., Kreitman M. & Bergelson J. (1999) Dynamics of disease resistance polymoprhism at the RPM1 locus of *Arabidopsis*. *Nature* **400,** 667–671.

Stanton M.L., Roy B.A. & Thiede D.A. (2000) Evolution in stressfull environments. I. Phenotypic variability, phenotypic selection, and response to selection in five distinct environmental stresses. *Evolution* **54,** 93–111.

Stotz H.U., Pittendrigh B.R., Kroymann J., Weniger K., Bauke F.J., Bauke A. & Mitchell-Olds T. (2000) Induced plant defense responses against chewing insects. Ethylene signaling reduces resistance of *Arabidopsis* against Egyptian cotton worm but not diamondback moth. *Plant Physiology* **124,** 1007–1017.

Stratton D. (1998) REaction norm functions and ATL–environment interactions for flowering time in *Arabidopsis thaliana*. *Heredity* **81,** 144–155.

Sun W., van Montague M. & Verbruggen N. (2002) Small heat shock proteins and stress tolerance in plants. *Biochimica et Biophysica Acta* **1577,** 1–9.

Sun W., Bernard C., van de Cotte B., van Montagu M. & Verbruggen N. (2001) At-HSP17.6A, encoding a small heat-shock protein in Arabidopsis, can enhance osmotolerance upon overexpression. *Plant Journal* **27,** 407–415.

Sun J., Gibson K.M., Kiirats O., Okita T.W. & Edwards G.E. (2002) Interactions of nitrate and CO_2 enrichment on growth, carbohydrates and rubisco in *Arabidopsis* starch mutants. Significance of starch and hexose. *Plant Physiology* **130,** 1573–1583.

Sung D.-Y., Kaplan F., Lee K.-J. & Guy C.L. (2003) Acquired tolerance to temperature extremes. *Trends in Plant Science* **8,** 179–187.

Swarrup K., Alonso-Blanco C., Lynn J.R., Michaels S.D., Amasino R.M., Koornneef M. & Millar A.J. (1999) Natural allelic variation identifies new genes in the Arabidopsis circadian system. *Plant Journal* **20,** 67–77.

Tepperman J.M., Zhu T., Chang H.-S., Wang X. & Quail P.H. (2001) Multiple transcription-factor genes are early targets of phytochrome A signalling. *Proceedings of the National Academy of Sciences of the USA* **98,** 9437–9442.

Tian D., Traw M.B., Chen J.Q., Kreitman M. & Bergelson J. (2003) Fitness costs of R-gene-mediated resistance in *Arabidopsis thaliana*. *Nature* **423,** 74–77.

Tilman D. & Lehman C. (2001) Human-caused environmental change: Impacts on plant diversity and evolution. *Proceedings of the National Academy of Sciences of the USA* **98,** 5433–5440.

Tilman D., Reich P.B., Knops J., Wedin D., Mielke T. & Lehman C. (2001) Diversity and productivity in a long-term grassland experiment. *Science* **294,** 843–845.

Traw M.B., Kim J., Enright S.D., Cipollini F. & Bergelson J. (2003) Negative cross-talk between salicylate- andjasmonate-mediated pathways in the Wassilewskija ecotype of *Arabidopsis thaliana*. *Molecular Ecology* **12,** 1125–1135.

Van der Schaar W., Alonso-Blanco C., Léon-Kloosterziel K.M., Jansen R.C., Van Ooijen J.W. & Koornneef M. (1997) QTL analysis of seed dormancy in *Arabidopsis* using recombinant inbred lines and MQM mapping. *Heredity* **79,** 190–200.

Wade M.J. (1984) The effects of genotypic interactions on evolution in structured populations. In *Genetics: New Frontiers, Proceedings of the XV International Congress on Genetics* (eds V.L. Chopra, B.C. Joshi, R.P. Sharma & H.C. Bensal), Oxford and I.B.H. Publishing Co., New Delhi, India, pp. 283–290.

Wade M.J. & Goodnight C.J. (1998) The theories of Fisher and

Wright in the context of metapopulations: when nature does many small experiments. *Evolution* **52,** 1537–1553.

Wagner G.P. (1996) Homologues, natural kinds and the evolution of modularity. *American Zoology* **26,** 36–43.

Ward J.K. & Strain B.R. (1997) The effects of low and elevated CO_2 partial pressure on grwoth and reproduction of *Arabidopsis thaliana* from different elevations. *Plant, Cell and Environment* **20,** 254–260.

Ward J.K., Antonovics J., Thomas R.B. & Strain B.R. (2000) Is atmospheric CO_2 a selective agent on model C_3 annuals? *Oecologia* **123,** 330–341.

Weinig C., Stinchcombe J.R. & Schmitt J. (2003) QTL architecture of resistance and tolerance traits in *Arabidopsis thaliana* in natural environments. *Molecular Ecology* **12,** 1153–1163.

Weinig C., Ungerer M.C., Dorn L.A., Kane N.C., Toyonago Y., Halldorsdottir S.S., Mackay T.F.C., Purugganan M.D. & Schmitt J. (2002) Novel loci control variation in reproductive timing in *Arabidopsis thaliana* in natural environments. *Genetics* **162,** 1875–1884.

Whitlock M.C. & Otto S.P. (1999) The panda and the phage: compensatory mutations and the persistence of small populations. *Trends in Ecology and Evolution* **14,** 295–296.

Whitlock M.C. & Phillips P.C. (2000) The exquisite corpse: a shifting view of the shifting balance. *Trends in Ecology and Evolution* **15,** 347–348.

Whitlock M.C., Phillips P.C., Moore F.B.G. & Tonsor S.J. (1995) Multiple fitness peaks and epistasis. *Annual Review of Ecology and Systematics* **26,** 601–605.

Wolf J.B., Brodie E.D. & Wade M.J. (eds) (2000) *Epistasis and the Evolutionary Process*. Oxford University Press, New York, USA.

Woodward F.I., Lake J.A. & Quick W.P. (2002) Stomatal development and CO_2: ecological consequences. *New Phytologist* **153,** 477–484.

Yano M. (2001) Genetic and molecular dissection of naturally occurring variation. *Current Opinion in Plant Biology* **4,** 130–135.

Zhang J. & Lechowicz M.J. (1995) Response to CO_2 enrichment by two genotypes of *Arabidopsis thaliana* differing in their sensitivity to nutrient availability. *Annals of Botany* **75,** 491–499.

Zhu J.K. (2002) Salt and drought stress signal transduction in plants. *Annual Review of Plant Biology* **53,** 247–273.

Ziegelhoffer E.C. Medrano L.J. & Meyerowitz E.M. (2000) Cloning of the *Arabidopsis* WIGGUM gene identifies a role for farnesylation in meristem development. *Proceedings of the National Academy of Sciences of the USA* **97,** 7633.

Received 20 May 2004; received in revised form 10 September 2004; accepted for publication 27 September 2004

APPENDIX

Glossary of ecological and evolutionary terms used in this paper

Some of the definitions have been simplified to be more approachable (e.g. variance) by omitting technical detail or mathematical formalism.

adaptation	1. A feature that evolved its current state because that state confers higher fitness than past states. 2. The process by which an adaptive feature evolves through the action of natural selection.
adaptive	A feature that confers higher fitness than alternative features.
additive genetic effect	The phenotypic effect of an allele averaged across the distribution of genetic backgrounds of an interbreeding population.
additive genetic variance	The variation in average effects for all alleles in an interbreeding population.
coadapted genome	Evolved changes in allelic composition across multiple loci in response to natural selection on the interactive effects of interlocus allelic combinations.
differentiation	Between-population evolutionary divergence of DNA sequences.
epistasis	Dependency of the effects of the genotype at one locus on the genotype at one or more other loci.
evolution	1. Change in gene frequencies through generational time. 2. Heritable change in mean phenotype through generational time. These definitions are concordant with each other, just viewed at different levels of biological organization.
gene–environment interaction	Dependency of the phenotypic effects of a genotype on the environment in which the phenotype develops.
genetic architecture	All the genetic factors that contribute to a phenotypic trait as well as their magnitudes and interactions.
heritability	1. The extent to which offspring will on average resemble their parents and/or each other, relative to their interbreeding population as whole. 2. (narrow sense) The ratio of additive genetic variance to total phenotypic variance.
introgression	The repeated interbreeding of two lines, usually in the form of backcrossing to one of the lines with the result that a gene or genes from one line are bred into the genetic background of the other line.
linkage disequilibrium	A non–random association between the allelic variants at one locus and those at another, due to physical linkage, non-random mating or post-mating selection.
natural selection	1. (narrow sense of population and quantitative genetics) Covariation between phenotype and fitness within a generation. 2. (broader sense of some evolutionary biologists) An adaptive change in mean phenotype over generational time.
G matrix	For a set of n traits, the G matrix is an $n \times n$ matrix containing the genetic covariances for all possible pairs of traits. It describes the extent to which the deviation from the population mean trait value for any trait i for a group of relatives predicts the deviation from the mean trait value for any second trait j in that group of relatives for $i,j = 1 \ldots n$.
M matrix	For a set of n traits, the M matrix is an $n \times n$ matrix containing the mutational covariances for all possible pairs of traits. It describes the extent to which deviations from the population mean value of any trait i predict the deviations in mean trait value for any second trait j for $i,j = 1 \ldots n$. At least two other kinds of mutational (M) matrices exist.
P matrix	For a set of n traits, the P matrix is an $n \times n$ matrix containing the phenotypic covariances for all possible pairs of traits. It describes the extent to which the mean trait value for any trait i predicts the mean trait value for any second trait j for $i,j = 1 \ldots n$.
phenotype	1. The characters that together comprise the structure and function of the organism. 2. The measurement or state of a single specified character for an individual organism.
pleiotropy	1. An attribute of a gene, such that the gene has effects on multiple phenotypic attributes. 2. The number of traits affected by mutations in a specific gene.
reproductive value	An individual's expected future reproduction.
trait	A measurable attribute of an organism.
transformation	A heritable introduction of exogenous DNA to a cell or organism.
transposon	(or transposable genetic element) is a DNA sequence that is flanked by inverted repeats that can be inserted in a host DNA sequence.