

The cause and consequences of natural variation: the genome era takes off!

Editorial overview

Juliette de Meaux and Maarten Koornneef

Current Opinion in Plant Biology 2008, 11:99–102

Available online 11th March 2008

1369-5266/\$ – see front matter

© 2008 Elsevier Ltd. All rights reserved.

DOI [10.1016/j.pbi.2008.02.006](https://doi.org/10.1016/j.pbi.2008.02.006)

Juliette de Meaux

Department of Plant Breeding and Genetics,
Max Planck Institute for Plant Breeding
Research, Carl-von-Linné weg 10, 50829
Cologne, Germany
e-mail: koornnee@mpiz-koeln.mpg.de

Juliette de Meaux is seeking to determine how variation in central physiological processes impacts the evolutionary success of *Arabidopsis thaliana* and its close relatives. Her research emphasis is on the relevance of regulatory versus protein-coding variation for the evolution of these processes. Her current work is dedicated to dissecting the adaptive relevance of molecular variation affecting either major life-history transitions, such as the timing of germination, or crucial stress-response systems, such as basal immunity. This requires the accurate characterization of the functional changes associated with nucleotide variation found within and between species.

Maarten Koornneef^{1,2}

¹ Department of Plant Breeding and Genetics,
Max Planck Institute for Plant Breeding
Research, Carl-von-Linné weg 10, 50829
Cologne, Germany

² The Laboratory of Genetics, Wageningen
University, Arboretumlaan 4, NL-6703 BD
Wageningen, The Netherlands
e-mail: koornnee@mpiz-koeln.mpg.de

Maarten Koornneef is interested in the use of genetics to unravel the molecular and biochemical nature of complex traits in plants. In addition to mutant approaches, natural variation within the species *A. thaliana* is being used by the application of quantitative and molecular genetics. For this, novel methods and resources are being developed. Many of the traits that are studied (e.g. seed dormancy, plant performance) have relevance both in agriculture and nature and also require answering questions in population and ecological genetics.

Introduction

Modern biology was born from the systematic description of natural variation, which eventually led Carl Linnaeus (1707–1778) to establish a formal system for naming species. Soon though, the urge to understand the mechanisms of life, both at the biochemical, population, and species levels, required the focus on the restricted set of systems that were experimentally most tractable. This approach has led to the enormous progress made along the last century, in which *Arabidopsis* research reached a climax. Now, it becomes possible to step beyond model-organism-based science: the tools are there and the fascinating variety of physiological, ecological, and evolutionary conundrums posed by natural variation both within and among species is within reach.

In this section, the contributing authors each stress the latest development in the multifaceted study of natural variation. The era that is opening now was paved by the completion of the first whole-genome sequences. The ever-growing number of such data for multiple individuals and/or species, as well as the increased accessibility to high-throughput genotyping and phenotyping, promises a broader perspective for both our understanding of the origin of novelties in life as well as our mastering of biological function of applied importance. Indeed, for the first, molecular predictions can now be verified, while for the second, common mechanism can begin to be disentangled from genotype-specific behaviors in plant biology.

Cause for variation within and between species

Three reviews depict different aspects of the cause for variation within and between species. Soon, the community will be immersed in molecular data on natural variation both within and between species. These species differ in their history, their biology and their ecological range, and are thus submitted to contrasted demographic regimes. Lawton-Rauh [1] describes how demographic processes influence natural variation. The elucidation of these processes is a prerequisite to the efficient study of natural variation.

In addition to demographic effects, Darwinian selection also influences patterns of diversity. Identifying the recent targets of selection in genomes is a theoretical challenge with enormous application potential as it reveals the genetic basis of fitness in natural conditions. Extensive and cost-effective access to genotyping information at large genomic scales has enabled the systematic scrutiny of genomes for genes influencing fitness variation. The review by Tenaillon and Tiffin [2] states on ever more sophisticated approaches to detect genes subject to recent selective sweeps. The application of these approaches to crop genomes promises to illuminate the

subtleties aspects of domestication and excavate genes of potential economical importance in future plant breeding [3,4].

Although positive selection depletes variation, other mechanisms are known to maintain high levels of variation. In this section, Schierup and Vekemans [5] focus on a system that maintains amazing variability over an extraordinary time scale: the self-incompatibility gene complex. Its study in outcrossing *Arabidopsis* relatives offers a textbook example of experimentally verified theoretical predictions, whereas the loss of self-incompatibility in *Arabidopsis thaliana* does not seem to conform to any simple scenario. Now, the impact of this system on further aspects of the speciation process, such as inter-specific introgression, unravels the profound consequences of these systems on natural variation (Castric *et al.*, unpublished data).

Dissecting the molecular genetics of complex traits

Keurentjes *et al.* [6] illustrate the endeavors to link networks of coregulated genes to the phenotypes they underpin. Reaching an unprecedented level of molecular detail, these approaches use a combination of multiparallel high-throughput characterization of diversity at the gene-expression, proteomic, or metabolic levels. An excellent illustration of the success of such approaches was provided by a previous paper, which based on co-expression analysis, successfully linked expression variation to phenotype and network reconstruction [7]. Here, this paper is brought into the larger context of its research field. Multiparallel 'omics' science progresses rapidly and is about to generate munificent databases that will be crucial to the dissection of natural variation at the molecular level.

Indeed, a mature understanding of physiological mechanisms is fundamental to a number of economical and environmental issues. In this endeavor, *Arabidopsis* research will maintain the lead and describe with ever-greater accuracy the networks in which genes cooperate to produce the phenotype. A deeper knowledge of the molecular basis of phenotypic variation, in turn, should facilitate the transfer of molecular knowledge to more or less related species.

The detailed analysis of heavy-metal tolerance in *Arabidopsis halleri* provides one of the most advanced instances of how one can take advantage of relatedness to *A. thaliana* to elucidate both the history and molecular basis of physiological adaptations to new environments. The review by Pauwels *et al.* [8] shows how geographical patterns can be used to reconstruct current and past adaptive processes and to identify the most relevant genotypes to be investigated at the molecular level.

The management of pest and pathogen diseases in plants represents a similar challenge. Plant diseases are causing phenomenal economical losses in the world and this situation will probably be exacerbated by climate change [9]. In the model species *A. thaliana*, plant defense systems are very well described but naturally occurring interactions are poorly understood (see [10,11] for some attempt). In their review, Salvaudon *et al.* [12] recapitulate advances at the molecular and ecological levels and illustrate how ecological and gene-based approaches are still like two ships passing in the dark. The authors propose to focus on quantitative variation against species with a demonstrated ecological impact. Indeed, selection pressure imposed by pathogens is local and will predominantly impact defense systems that are locally variable. Studies focusing on major genes may mask this diversity, thereby obscuring the effect of selection. A coherent analysis of complex traits such as fitness in the presence of disease requires ecologically realistic approaches.

An increasingly diverse array of plant model systems

Over the past decades, plant molecular biology was dominated by research in the model species *A. thaliana*. Rice, the first fully sequenced grass genome, is now best positioned as a major model for dissecting the relationship between gene function, the domestication process, and the most complex of all traits: yield. In their review, Doi *et al.* [13] describe the tools available in rice research to elucidate the cause of variation at the nucleotide level. In addition, their review provides a good illustration of how basic and applied research can be synergistic. The elucidation of genes causing reproductive barriers will illuminate both our understanding of speciation and the management of germplasm in this crop. Furthermore, the identification of domestication genes reveals a predominant role of mutations causing loss of function in the response to artificial selection.

Progress in comparative genomics will also improve our understanding of tree phenotypic variation and its cause. The review by Neale and Ingvarsson [14] sheds light on the recent advances and promising paths taken in tree research. Although traditionally seen as fairly untractable systems, resources for quantifying variation in natural environments are exceptionally abundant in these species. In addition, the concomitant advances made in the angiosperm genus *Populus* and in various gymnosperms will facilitate the discovery of the genetic basis for common physiological properties relevant for adaptation.

The way technological advances allow research to progress from model species into a diverse array of species with fundamentally different ecology is also described in a thrilling review by Karrenberg and Widmer [15]. Taking

some of the most important ecological challenges that can only be partially (if at all!) addressed in *A. thaliana* (ranging from adaptation to edaphic and climatic conditions to plant reproduction or biotic interactions), [Karrenberg and Widmer](#) summarize recent advances in the dissection of the molecular basis of natural variation in various model species for ecology. Their review describes the methods currently used to characterize ecologically relevant variation, whose accuracy promises to increase steadily in the future, if ever more talented molecular biologists, ecologists, and evolutionary biologists set their heart on this challenge.

Cross-talks and future developments

Along this section of Current Opinion in Plant Biology, several aspects are recurrently mentioned, pointing to areas that will presumably progress particularly rapidly along the next years. The question of the role of gene duplication in phenotypic evolution and adaptation is mentioned in two reviews [2,15]. Indeed, with the availability of whole-genome sequences, duplicate gene families can be examined with such detail that each duplication event can be dated [16]. Simultaneously, the characterization of quantitative variation is being used to answer an ever more diverse range of questions: to accurately characterize complex traits [6], to understand how phenotypic evolution is constrained or facilitated by genetic architecture [12], or simply to identify candidate adaptive genes [13–15]. Quantitative trait mapping, using ‘omic’ approaches when possible, and more and more frequently complemented by association studies, will be the necessary path to the accurate elucidation of fitness variation in nature [1,6,12–15].

The analysis of variation for gene regulation is also a recurrent topic on which much hope is placed to identify the molecular basis of ecologically relevant variation [6,8,15]. Indeed, means to study expression variation at a large scale can now be developed in a few years (or even months) as EST libraries can be used to build custom arrays at acceptable costs. The accurate detection of genetically or environmentally relevant expression variation provides an attractive solution to accelerate gene and pathway discoveries in processes that have so far remained locked in untractable organisms [17,18].

More generally, the questions of functional variation and evolutionary change are knocking at the door of most research fields aiming at the discovery of new genes for improved control of economically important functions, from the molecular basis of speciation in rice [13] to the physiology of heavy-metal tolerance [8]. In addition, the understanding of defense mechanisms and their efficiency against pathogens, symbionts, and/or abiotic stresses requires reconciling multiple time, spatial but also molecular and physiological scales [8,12,15]. Soon, it will be possible to efficiently address

the question of how plants coordinate in nature the demands imposed by multiple biotic and abiotic challenges requiring the joint regulation of plastic responses and life history. For example, the extent to which recently derived annuality in *A. thaliana* explains the architecture of disease defense in this species is an exciting challenge to be addressed. These advances will be provided both by the availability of multiple genomes in diverse sets of closely related species, and by a detailed understanding of the dynamics of physiologically and adaptively relevant variation within species. It is good news that ecological approaches are being increasingly developed in model organisms while in nonmodel organisms scientists strive to dissect their phenotypic properties at the molecular level.

References

1. Lawton-Rauh A: **Demographic processes shaping genetic variation.** *Curr Opin Plant Biol* 2008, **11**:103–109.
2. Tenaillon MI, Tiffin PL: **The quest for adaptive evolution: a theoretical challenge in a maze of data.** *Curr Opin Plant Biol* 2008, **11**:110–115.
3. Caicedo AL, Williamson SH, Hernandez RD, Boyko A, Fledel-Alon A, York TL, Polato NR, Olsen KM, Nielsen R, McCouch SR *et al.*: **Genome-wide patterns of nucleotide polymorphism in domesticated rice.** *PLoS Genet* 2007, **3**:1745–1756.
4. Yamasaki M, Tenaillon MI, Bi IV, Schroeder SG, Sanchez-Villeda H, Doebley JF, Gaut BS, McMullen MD: **A large-scale screen for artificial selection in maize identifies candidate agronomic loci for domestication and crop improvement.** *Plant Cell* 2005, **17**:2859–2872.
5. Schierup MH, Vekemans X: **Genomic consequences of selection on self-incompatibility genes.** *Curr Opin Plant Biol* 2008, **11**:116–122.
6. Keurentjes JJB, Koornneef M, Vreugdenhil D: **Quantitative genetics in the age of omics.** *Curr Opin Plant Biol* 2008, **11**:123–128.
7. Keurentjes JJB, Fu JY, Terpstra IR, Garcia JM, van den Ackerveken G, Snoek LB, Peeters AJM, Vreugdenhil D, Koornneef M, Jansen RC: **Regulatory network construction in Arabidopsis by using genome-wide gene expression quantitative trait loci.** *Proc Natl Acad Sci U S A* 2007, **104**:1708–1713.
8. Pauwels M, Roosens N, Frérot H, Saumitou-Laprade P: **When population genetics serves genomics: putting adaptation back in a spatial and historical context.** *Curr Opin Plant Biol* 2008, **11**:129–134.
9. Garrett KA, Dendy SP, Frank EE, Rouse MN, Travers SE: **Climate change effects on plant disease: genomes to ecosystems.** *Annu Rev Phytopathol* 2006, **44**:489–509.
10. Goss EM, Bergelson J: **Fitness consequences of infection of Arabidopsis thaliana with its natural bacterial pathogen Pseudomonas viridiflava.** *Oecologia* 2007, **152**:71–81.
11. Allen RL, Bittner-Eddy PD, Grenville-Briggs LJ, Meitz JC, Rehmany AP, Rose LE, Beynon JL: **Host-parasite coevolutionary conflict between Arabidopsis and downy mildew.** *Science* 2004, **306**:1957–1960.
12. Salvaudon L, Giraud T, Shykoff J: **Genetic diversity in natural populations: a fundamental component of plant–microbe interactions.** *Curr Opin Plant Biol* 2008, **11**:135–143.
13. Doi K, Yasui H, Yoshimura A: **Genetic variation in rice.** *Curr Opin Plant Biol* 2008, **11**:144–148.
14. Neale DB, Ingvarsson PK: **Population, quantitative and comparative genomics of adaptation in forest trees.** *Curr Opin Plant Biol* 2008, **11**:149–155.

15. Karrenberg S, Widmer A: **Ecologically relevant genetic variation.** *Curr Opin Plant Biol* 2008, **11**:156-162.
16. Vieira FG, Sanchez-Gracia A, Rozas J: **Comparative genomic analysis of the odorant-binding protein family in 12 *Drosophila* genomes: purifying selection and birth-and-death evolution.** *Genome Biol* 2007, **8**:R235 231-216.
17. Shiu SH, Borevitz JO: **The next generation of microarray research: applications in evolutionary and ecological genomics.** *Heredity* 2008, **100**:141-149.
18. Fay JC, Wittkopp PJ: **Evaluating the role of natural selection in the evolution of gene regulation.** *Heredity* 2008, **100**: 191-199.