Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Plant Biology

From genome studies to agricultural biotechnology: closing the gap between basic plant science and applied agriculture

Editorial overview

Douglas R Cook and Rajeev K Varshney

Current Opinion in Plant Biology 2010, **13**:115–118

1369-5266/$ – see front matter

DOI 10.1016/j.pbi.2010.03.003

Douglas R Cook

Department of Plant Pathology, University of California, Davis 95616, CA, USA
e-mail: drcook@ucdavis.edu

Doug Cook's research has a dual focus on legume genomics and symbiotic nitrogen fixation. His research interests in legume genomics have evolved from the development of *Medicago truncatula* as a model system for legume biology in the 1990s, to a bifurcated focus involving the evolution of crop legume genomes and community genomics of adaptive phenotypes in natural populations of *M. truncatula*. His work on genome evolution spans two time frames, ~55 millions of years to the last common ancestor of all major legume crops, and the past 10 thousand years through the domestication bottleneck. A byproduct of these studies is genome-scale data sets that interconnect the genomes of crop and model legumes, which in turn enable translation from model species to applied outcomes. Much of his work on legume genomics is focused on India and sub-Saharan Africa, where legumes have an especially important role as sources of human nutrition and soil fertility.

Rajeev K Varshney[1,2]

[1] Centre of Excellence in Genomics, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru 502324, AP, India
[2] Genomics towards Gene Discovery Subprogramme, Generation Challenge Programme, c/o CIMMYT, Int APDO Postal 6-641, 06600 Mexico, D.F., Mexico
e-mail: r.k.varshney@cgiar.org

Rajeev Varshney has an interest in structural, functional, and comparative genomics. While working on wheat and barley for about a decade, he contributed to the development of a large number of genome resources including ESTs, SSRs, SNPs, genetic maps, comparative maps, and markers associated with traits of interest to breeders. At ICRISAT his group works in strategic collaboration with several partners around the world toward developing large-scale genomic resources, for example, millions of short transcript reads, thousands of SSR and SNP markers and development of the first or comprehensive genetic maps. A primary focus is the identification of markers associated with drought tolerance in so-called 'orphan legume crops' of the semi-arid tropics. He is collaborating with breeders for using these genomic resources in molecular breeding applications to develop superior varieties of legumes with enhanced drought tolerance in India and sub-Saharan Africa.

Background

In the broad field of plant biology, access to genome technologies and genome-scale data sets is driving a convergence of subdisciplines that have been historically separate. For example, the next generation of ecologists and crop geneticists are just as likely to be concerned about protein interaction networks that underlie complex phenotypes as they are about the more abstract genetic and evolutionary processes from which such phenotypes derive. Then what will be the difference between an ecologist and a crop geneticist, or a protein biochemist, and a molecular breeder? Genomics is the primary driver of this unification in the biological sciences, though the unification is admittedly still in its infancy. The factors driving this convergence are many, but at a practical level they include first, more comprehensive and better annotated genome data sets; second, a transition from the situation where genome technologies were used to generate data sets accompanied by limited understanding, to a situation where genome technologies are tools for hypothesis-driven research and the resulting data sets are synthesized to yield new understandings; third, rapid decreases in the cost, combined with staggering increases to the scale and accuracy of genome analyses, especially in the areas of nucleic acid (re)-sequencing and genotyping technologies; and fourth, increased computational capacity and new statistical approaches to manage and analyze increasingly large and complex genome data sets.

The continuum from basic to applied science is one axis of the disciplinary convergence mentioned above. Our concept in organizing this issue of *Current Opinion in Plant Biology* was to invite a series of articles spanning the gap from basic genomics and genome characterization, to the application of genome-related biotechnologies to agriculture, especially in the developing world. Thus, a subtheme in this issue, perhaps more as an emergent property of the combined articles than of any one article in particular, is the 'translation' of basic genome sciences to crop improvement. On the basic side we have invited authors whose expertise and perspectives range from analysis of genetic and genomic process in natural populations, through domestication and its consequences for the genomes of crop species, to detailed characterization of 'omic' networks that underlie complex phenotypes. On the applied side the collected articles encompass statistically intensive areas of research, such as crop modeling and association genetics, to the relatively applied topic of conservation and utilization of crop genetic diversity, through a consideration of the challenges and opportunities of crop biotechnology in the developing world.

Current projections suggest that the global population is likely to plateau at some 9 billion people by the middle of this century [1]. Of these, more than

1 billion people are estimated to lack sufficient dietary energy, the vast majority of whom will live in the developing world [2]. With this ever-increasing human population and amidst the fear of shrinking resources in terms of cultivatable area, irrigation resources, newly emerging pathogens and pests, stagnant yields, etc. a primary challenge to plant biology in the coming years will be to deliver improved agricultural genotypes that can feed the ever-growing population. These new genotypes will likely take greater advantage of genetic diversity within crop germplasm collections, in wild progenitor species, and in taxa whose evolutionary distance precludes genetic crosses as a means of gene introduction — this all adds up to a situation where plant biology, especially genomics and biotechnology, will play a leading role. Moreover, a one-size-fits-all approach is unlikely to be appropriate, as a diversity of phenotypes will be necessary to continue to meet the demands of regional environments, economic and social constraints, and cultural preferences. As David Baulcombe pointed out in his recent *Science* editorial [3], massive increases to total food production will be required and the need will be most severe within our children's lifetimes. The broader issues of food security, agricultural sustainability, and public acceptance of crop biotechnology, all include challenges that will need to be addressed by plant biologists, and to which the subjects covered by these collected articles are relevant.

Genome studies and molecular genetics — enhancing the knowledge base

One of the most enabling features of current genomics is the ability to characterize genome-scale variation at the level of individual species, bringing the power of genomics to bare on population-scale data sets in the natural environment as well as in germplasm collections that underlie the world's staple food crops. Literally overnight, this capacity has created alliances between the historically separate disciplines of crop genetics, evolution and ecology researchers, and genome scientists. The alliance, so-to-speak, is built around the common need to tackle the significant challenges of experimental design, data management, statistical analysis, and data interpretation. One of the major opportunities in such studies is to identify and validate gene–phenotype associations that pervade large sets of germplasm, including both natural populations and collections of domesticated accessions. These themes and subthemes recur in several articles within this issue.

Perhaps one of the most exciting convergences is that between the community of scientists studying ecology and population genetics of natural populations, with the genomic community where methods for high-throughput discovery of genetic variation are becoming routine. Ecology and population genetics are theory-rich disciplines that study aspects of genetic and phenotypic variation with the goal of deducing the nature and origins of natural adaptations. Because their study systems occur in the natural world, these scientists have developed experimental and statistical approaches that embrace rather than avoid complexity, for example, environmental, spatial and genetic heterogeneity. When these approaches are combined with methods to genotype at ultra-high densities, involving potentially every base in each test subject, it is increasingly possible to harness the power of historical recombination and thus identify the genetic basis of even complex adaptations. Friesen and von Wettberg provide an overview of the approaches used in this field of 'ecological' or 'community' genomics, and they review examples of recent successes in deducing the molecular bases of natural adaptations.

The molecular underpinnings of many plant phenotypes, including agronomic traits, involve complex interactions that integrate among regulatory circuits, across developmental space, and between distinct genome fractions (e.g. the transcriptome, the proteome, the metabolome, and the epigenome). In their article on 'Omics meet networks ...', Moreno-Risueno and colleagues review recent advances in systems biology approaches that aim to integrate these large and diverse data sets, and they describe novel insights that are modifying how we conceptualize genetic interactions. Of relevance to the goal of translating basic genome sciences to agriculture, the ability to deduce key determinants of genomic network function in model systems, such as *Arabidopsis*, may inform studies in lesser characterized crop species. For example, principles and experimental approaches learned in the model systems may enable related studies in crop species, or the model systems may directly contribute genes that are candidates for crop improvement by means of technologies such as genetic transformation. In a related article, Urano *et al.* review recent developments in the genomic networks of plant abiotic stress. Abiotic stress is one of the primary constraints to agricultural productivity, and the networks for abiotic stress perception and response often link directly to pathways that determine growth, reproduction, nutrient acquisition, etc. A key challenge for the future will be to integrate knowledge of systems biology at the scale of populations, so that the contribution of genomic diversity to phenotypic plasticity can be more effectively modeled and used rationally in crop improvement programs. Such integration is beyond current capacities, but in the mean time the availability of genome-scale data sets in key crop species such as maize and soybean, is enabling researchers to test the conservation of regulatory interactions identified in model species.

Characterizing the processes of genome evolution is an important objective for basic genome scientists, while the outcome of genome evolution defines the structure and complexity of agricultural species. Genomic changes occurring over time scales of tens of millions of years can be deduced from the comparison of multiple plant genomes within key taxa. Of great interest are the grasses and legumes, not only because each group contains several plants with sequenced genomes, but also because the parallel domestication of grasses and legumes was key to the development of agriculture: beans and maize in the Americas; soybean and rice in Asia; and grain legumes (e.g. chickpea and lentil) and small grain grasses (e.g. wheat and barley) in the Middle East. Devos and Sato *et al.* review recent developments in the areas of grass genome evolution and legume genome structure, respectively. Genome comparisons are yielding insight into the factors that have contributed to the structure of modern-day genomes, including the impact of genome rearrangements and retrotransposons on genome plasticity in grasses; moreover, the promise of comparative genomics for translating basic science to applied outcomes is being pursued in both grasses and legumes.

One outcome of genome comparisons is the recognition that all plant genomes are derived from a series of ancestral polyploidy events. In fact, polyploidy is likely to have punctuated key evolutionary advances in a wide range of taxa. When polyploidization is relatively recent, as in the case of certain *Brassica* species and also in cotton and soybean, or when polyploid genomes are synthesized in the laboratory, it is possible to analyze the ways in which gene and genome duplication impacts gene function. Duplicated genes can experience diverse trajectories of change and selection, leading to gene loss, as well as a range of subfunctionalization and neofunctionalization outcomes. The impacts of these processes on the remodeling of gene expression are especially profound and they can have correspondingly profound impacts on organismal phenotypes. In their article, Jackson and Chen review the topic of genome duplication and expression plasticity.

From the perspective of agriculture, one of the most relevant evolutionary pathways is that of domestication, a process that has occurred in parallel in every crop lineage within the last ~10 thousand years. Tang *et al.* review our current understanding of plant domestication, placing key phenotypic changes within the context of humankind's transition to an agricultural lifestyle. Of particular interest, they consider the ways in which new genome technologies are impacting the study of plant domestication. They speculate on the potential to apply these tools to the domestication of a new generation of crop species, and they remind us that domestication is an ongoing process in which humans continually reinvent the function of plant genomes.

Plant biotechnology — applying genomic research to crop improvement

Although similar genes and pathways are likely to underlie agronomic traits in crop species and adaptations in natural populations, crop geneticists have tended to approach genetic analysis from a very different perspective relative to their colleagues studying ecology and population genetics. This is in part because the genetic structure of crop species is nothing like a natural population. But equally important is the fact that until recently the primary goals of these scientists have been different — crop geneticists have been less interested in the origins of genotype and phenotype variation, and more focused on the outcome, namely crop improvement.

A logical starting point for crop genetics is the management and analysis of genetic resources. However, although a wealth of crop germplasm is available worldwide, with about 6 million accessions held in over 1400 gene banks, breeders have tapped less than 1% of these collections for crop improvement. Glaszmann and colleagues advocate an increased focus on the selection of 'core reference sets' and subsequent characterization of their genomic and phenotypic attributes; especially important will be the use of next-generation sequencing technologies and high-throughput genotyping technologies to understand and leverage genetic diversity. Broadening the use of 'core reference sets' will facilitate material sharing within the crop genetics and breeding community and will enable plant breeders to create novel plant gene combinations and select crop varieties more suited to the needs of diverse agricultural systems.

The value of describing genetic diversity in germplasm collections increases in direct proportion to our ability to correlate genotypic and phenotypic variation. Toward this end, association mapping, also called linkage disequilibrium (LD) mapping, holds great promise. Association genetics deals with the analysis of statistical associations between genotypes, usually individual SNPs or haplotypes, determined in a collection of individuals, and the traits (phenotypes) of the same individuals. Antoni Rafalski discusses the potential and application of association genetics in crop improvement, especially in the context of improved genotyping and DNA sequencing methodologies. Rafalski emphasizes the accompanying need for precise phenotyping and the use of appropriate experimental design and statistical approaches.

Both association mapping and classical linkage mapping studies have been successful in the identification of trait-linked molecular markers as well in the cloning of agronomically important genes. These markers and genes have simultaneously advanced our understanding of molecular mechanisms that underlie traits, and they are increasingly important as tools to develop improved varieties/lines/hybrids. These and related topics are reviewed in two articles in this issue. Kou and Wang summarize recent advances in our understanding of the genes and their protein products that confer quantitative disease resistance. Quantitative disease resistance mechanisms are of interest because they are less well studied than qualitative disease resistance genes, and also because they are often the basis of desirable broad spectrum and durable disease resistance traits. Similarly Ouyang and colleagues provide a comprehensive update on genetic analyses and molecular characterization of hybrid sterility genes in rice. They advocate for investigation of the sequence diversity and geographical distribution of the alleles of various loci at the species level, including wild relatives. This should elucidate the origin of the genes for hybrid sterility and the evolutionary processes behind the establishment of species and subspecies. Such knowledge could eventually facilitate the formulation and development of novel strategies for rice improvement, and could have implications for the improvement of other crop species.

Abiotic stresses such as drought, salinity, and extreme temperature are severe constraints to agriculture production. Such stresses are predicted to be more severe and variable in the coming decades as a consequence of climate change. Crop geneticists and breeders have been engaged to identify the quantitative trait loci (QTLs) for abiotic stress tolerance in several crops. QTL mapping methods for complex traits, however, are challenged by new developments in marker technology, phenotyping platforms, and breeding methods. In meeting these challenges, QTL mapping approaches will need to also acknowledge the central roles of QTL by environment interactions (QEI) and QTL by trait interactions in the expression of complex traits. An article written by van Eeuwijk *et al.* presents an overview of mixed model QTL methodologies that are suitable for many types of populations and that allow predictive modeling of QEI, for both environmental and developmental gradients. They also propose biophysical (crop growth) model simulations as a complement to statistical QTL mapping for the interpretation of the nature of QEI and to investigate better methods for the dissection of complex traits into component traits and their genetic controls. In a related article, Tardieu and Tuberosa discuss the use of phenotyping platforms combined with model-assisted methods to dissect complex traits such as yield and stress tolerance into heritable traits. They describe how the effects of traits on yield are studied using a combination of modeling and field experiments, with the goal of identifying stress scenarios where a given allele has favorable effects.

The application of genome-scale data sets together with high-throughput genotyping technologies and advanced precise phenotyping has led to a range of trait-associated molecular markers and candidate genes that are of interest to breeders in a range of crop species. While the deployment of molecular markers in breeding and the introduction of novel genes through genetic engineering is becoming routine many crops in the private sector, as well as in public research institutions in developed countries, the adoption of molecular breeding methodologies, and genetically engineered crops is far from routine in the developing world. These and related issues are covered in two articles. Ribaut *et al.* discuss bottlenecks that constrain crop improvement in the developing world, including limited human resources and inadequate field infrastructure; they also review opportunities and potential solutions, including the establishment of virtual research and technology platforms for molecular breeding, aided by information and communication technology in developing countries. In a parallel article, Farre *et al.* present the potential of genetic engineering of crops to address many of the world's most challenging and interrelated problems, including hunger, malnutrition, disease, and poverty. They stress that overcoming political and societal barriers is key to realizing the potential of genetic engineering in developing countries.

In summary, genomic data and technologies are impacting essentially all areas of modern plant biology. In the case of agricultural species, recent advances in next-generation sequencing, high-throughput genotyping, and improved bioinformatic tools and statistical approaches, are yielding useful, large-scale genome data sets that increase the opportunity for translation between basic and applied plant science research. As a consequence, knowledge of fundamental biological processes in model species can more readily inform the application of biotechnology tools, such as trait-associated molecular markers and candidate genes, in the development of superior crop varieties. The result is that the gap between basic plant science research and applied agriculture is closing. But we should not be overly satisfied with this convergence, because putting this potential into action will require a redirection of effort and a renewed commitment to meeting the demands of an ever-increasing human population, especially in developing countries.

References

1. Godfray HCJ, Beddington JR, Crute IR, Haddad, Lawrence D, Muir JF, Pretty J, Robinson S, Thomas SM, Toulmin C: **Food security: the challenge of feeding 9 billion people**. *Science* 2010, **327**:812-818.
2. Barrett CB: **Measuring food insecurity**. *Science* 2010, **327**:825-828.
3. Baulcombe D: **Reaping benefits of crop research**. *Science* 2010, **327**:761.