

Preface to Special Issue: Complex traits and plant breeding—can we understand the complexities of gene-to-phenotype relationships and use such knowledge to enhance plant breeding outcomes?

Mark Cooper^{A,D} and Graeme L. Hammer^{B,C}

^APioneer Hi-Bred International Inc., 7250 N. W. 62nd Avenue, PO Box 552, Johnston, IA 50131, USA.

^BAgricultural Production Systems Research Unit, School of Land and Food Sciences,
The University of Queensland, Brisbane, Qld 4072, Australia.

^CAgricultural Production Systems Research Unit, Queensland Department of Primary Industries and Fisheries,
Toowoomba, Qld 4350, Australia.

^DCorresponding author. Email: Mark.Cooper@pioneer.com

Additional keywords: modelling, prediction, epistasis, pleiotropy, interaction.

Collective wisdom: If the door will not open after continual attempts at pushing, try pulling it!

Some would argue that the answer to the question we pose in the title of this introductory paper is already known, and is yes. For some traits, e.g. resistance to particular pests and diseases, there is no doubt that we have used detailed knowledge of the genetics that underlies trait phenotypic variation to enhance the performance of the product outcomes from plant breeding (e.g. Cahill and Schmidt 2004). However, we also recognise that plant breeding is fundamentally an undertaking in multi-trait improvement and that there are many important traits where we do not yet have a sufficient knowledge of the causal genetic variation to enable similar approaches. Therefore, we use these initial successes as encouragement and to argue for a continuation of research on other traits.

The papers included in this Special Issue of the Australian Journal of Agricultural Research are focussed around two related questions that impinge on our ambitions and abilities to enable knowledge-based approaches to molecular enhanced breeding: (1) the feasibility of, and (2) appropriate strategies for, constructing predictive gene-to-phenotype models of complex traits. Although it can be argued that the traits manipulated in plant breeding programs range from genetically simple to complex, most of the important traits that have preoccupied the field of quantitative genetics and plant breeders are towards the complex end of this continuum. Plant breeders have consistently demonstrated significant capacity to make desirable changes to many of the complex yield, quality, and agronomic traits of crop plants that are considered important for sustainable production in their respective target agricultural systems. Most often studies conducted to quantify genetic progress from breeding have focussed on the key endpoint traits, such as the improvement of grain yield and associated changes in other traits (e.g.

Fehr 1984; Duvick *et al.* 2004). Although several of these retrospective studies have found a tendency for increased yield to arise from change in biomass partitioning rather than increased total biomass production, many of the genetic and physiological details underpinning the improvements in yield varied among breeding programs and with the period of investigation (Duvick *et al.* 2004). Further, their effect in the intended target environments depended on the crop and on features of the environments within the target production system (Allard 1999). It is widely recognised that realising much of the genetic improvements from breeding is conditional on the use of appropriate agronomic management practices. The interplay between genotype, management, and environment is critical in realising improvements in crop performance (Cooper and Hammer 1996; Cooper *et al.* 2002; Yin *et al.* 2004). The ubiquitous nature of these interactions, and the degree to which they are considered in breeding programs, have contributed to an ongoing debate about the extent of the outcomes from breeding that are realised by different groups of farmers in the diversity of the global target production systems.

Today, as we consider molecular breeding strategies for complex traits, it is important to recall that for the majority of the 20th Century, formal plant breeding was conducted by direct selection on the phenotypes of the traits of interest, within populations of genotypes that contained appropriate sources of genetic variation for those traits, without knowledge of the gene-to-phenotype details of the genetic architecture of the traits. The details of the genetic variation and the effects of selection on the genetic composition of populations of elite breeding lines were not well understood. In most cases, successful breeding

was an outcome from the combination of experienced breeders working with relevant germplasm and applying proven breeding methods. Despite this lack of scientific knowledge of the detail of the changes at the genetic level, progress was made (Niebur *et al.* 2004). Throughout the course of the 20th Century, research across several disciplines provided advances in our understanding of some aspects of genetic variation and gene-to-phenotype associations for traits. Building on the foundational work by Fisher, Wright, and Haldane, quantitative genetics developed as a subdiscipline of genetics and was used to study the properties of genes in populations and the expected influence of inferred genetic models on trait phenotypes (Falconer 1960; Mather and Jinks 1971; Comstock 1996; Falconer and Mackay 1996; Kearsley and Pooni 1996; Lynch and Walsh 1998). Some of the predictions made from this theoretical framework matched observation and some did not (Coors 1999; Duvick *et al.* 2004). The formal training of the cohorts of plant breeders that emerged in the second half of the 20th Century included at least an introduction to quantitative genetic theory and its applications to plant breeding (Hallauer and Miranda 1988; Comstock 1996). The combination of the theoretical framework and trained practitioners resulted in the application of quantitative genetics to problems in plant breeding and as a basis for design and comparison of breeding strategies (Hanson and Robinson 1963; Coors 1999). In many respects the classical models of quantitative genetics were successful in providing a framework around which plant breeders could begin to question the nature of progress from selection, even if they did not enlighten the biological detail of the gene-to-phenotype relationships that were changed. A notable success was the large body of research that has contributed to the design and application of hybrid breeding methods in crops (Comstock *et al.* 1949; Cooper *et al.* 2004; Duvick *et al.* 2004; Niebur *et al.* 2004).

Nevertheless, although we have made progress in our understanding of the genetic architecture of complex traits, as stated by Axelrod and Cohen (1999) in their book on 'Harnessing Complexity', we often find ourselves in situations where we are working on complex problems and are using an available framework for making predictions that assumes a lot more predictive power than we are likely to have in practice. This is true for many of the complex problems we tackle by plant breeding. Often realised genetic gain is significantly less than predicted gain, even though the relative performance of breeding strategies is, in general, consistent with theoretical prediction (Coors 1999). The gaps that we frequently observe between predictions of expected response to selection in plant breeding and realised response to selection are a classic example of the situation discussed by Axelrod and Cohen (1999).

There has always been much discussion and speculation on the scope for improving our scientific knowledge in biology and our ability to make predictions of gene-to-phenotype relationships by integration of the knowledge and

expertise from different scientific disciplines. This challenge is general and extends well beyond the particular plant breeding problems that are the focus of this special issue (e.g. Ideker *et al.* 2001; Sing *et al.* 2003). If we consult the literature of the 1970s and 1980s it appears that optimism was focussed on anticipated opportunities for efforts that combined plant breeding and physiology. If we consider the 1990s, molecular genetics and genomics became the emphasis in much of this discussion. As we now move into the beginning of the 21st Century there is a high level of enthusiasm for a 'Systems Biology' approach, which to the outsider appears to be a disparate collection of efforts involving integration of knowledge and skills from molecular genetics, computer science, biochemistry, plant physiology, mathematics, and engineering. The one conclusion that can be drawn from any attempt to synthesise and identify trends in this body of literature is that our hopes for improving predictability within the realm of the gene-to-phenotype relationship for important plant traits have focussed on different combinations of disciplines at different times. Each phase of speculation is followed by empirical attempts to realise the ideas in practice, with achievement of variable levels of success that usually end with a phase invoking the need to think in an integrative manner that involves input across disciplines that were not part of the team included in the latest attempt.

Within the broad schema of science discussed by Thomas Kuhn (Kuhn 1996), it is reasonable to consider that we are in the midst of a *paradigm shift* for quantitative genetics and plant breeding. See also the related discussion by Strohman (1997). Up to the late 1980s to early 1990s we could not attain many details of the particulate nature of the genetic architecture of the traits that we studied. Now, at the beginning of the 21st Century we can achieve views of this particulate nature and study the basis of genetic variation from the DNA sequence level to the phenotypes of traits across molecular, cellular, tissue, organ, and physiological process levels of organisation within organisms. As these views emerge there are many surprises as well as confirmations. Thus, the assumed models and the accompanying predictive framework of quantitative genetics are being seriously challenged and in some cases replaced by the widely anticipated gene network models. At this time we have not gained a lot of experience working with the properties of these alternative and more complex gene-to-phenotype models, but it is fair to state that the scientific community is preoccupied with describing them and attempting to understand their properties. The view that is emerging is a much richer characterisation of the genetic basis of variation for the traits that have been the focus for improvement by breeding programs. We are starting to replace what has been historically a qualitative appreciation of the continuum of traits, ranging from simple to complex, with a quantitative description of the genes, gene function, and gene interactions within networks and with environmental variables.

Over the same period as these developments in quantitative genetics were occurring, there was significant sophistication in plant modelling and virtual plant capabilities based on increased knowledge of plant functional biology. During the last quarter of the 20th Century the role of plant/crop modelling reached an accepted level of maturity in relation to supporting scientific investigation and facilitating decision-making by crop managers (Sinclair and Seligman 1996). This was mostly associated with the useful predictive ability of the models in dealing with the interaction between management and environment. However, as the new century emerged, so did the possibilities for using the physiological frameworks contained in crop models to aid in understanding and advancing the genetic regulation of plant performance and plant improvement (Hammer *et al.* 2002). The ability to physiologically dissect and model traits provided an avenue to enhance integration of molecular genetic technologies in crop improvement. Much of the advances in knowledge were fostered by studies on causes of phenotypic variation among lines from breeding populations. This avenue of research was partly stimulated by possibilities arising from molecular genetics. By enhancing the architecture and capability of crop models, we are starting to see an ability to predict complex responses in plant behaviour in a way that can be associated with the underlying genomic variation in growth and development processes (Tardieu 2003). Despite the orders of magnitude of complexity traversed as we move from whole organism to molecular scale, particularly for complex traits, this approach is starting to place physiological understanding in phenotype-to-genotype associations in a manner that also provides predictive ability (Reymond *et al.* 2003).

With the rapid expansion of data and knowledge on the molecular basis of genetic variation and with greater sophistication in quantitative plant functional physiology, we think that it is important to ask: are we in a better position today to understand the complexities of gene-to-phenotype or phenotype-to-genotype associations for traits and to use this knowledge in plant breeding? As a naive starting point we can consider that the two complementary approaches commence at different ends of a continuum of levels of biological organisation, but they both have relevance to this question, which is the core theme of the series of papers in this Special Issue of the Australian Journal of Agricultural Research.

There are 6 papers included in this Special Issue. Walsh (2005) gives an overview of the extensions of the classical framework of quantitative genetics that have been developed to accommodate both additive and non-additive sources of genetic variation in the breeder's predictive framework. Van Eeuwijk *et al.* (2005) discuss the application of molecular marker maps to study the genetic basis of phenotypic responses of genotypes to varying environmental conditions. They use mapping methodology to identify quantitative trait loci (QTL) that influence adaptation to environmental

conditions and that are explanatory of the differential performance of genotypes across varying environmental conditions. Cooper *et al.* (2005) consider theoretical issues and applications of finite locus models that have some flexibility to incorporate the types of non-additive effects suggested by Walsh (2005) and van Eeuwijk *et al.* (2005). The remaining 3 papers provide specific examples of how to deal with the complexity of gene-to-phenotype relationships for traits at the gene network (Welch *et al.* 2005), tissue and organ (Tardieu *et al.* 2005), and multi-trait crop (Hammer *et al.* 2005) levels.

We consider the views discussed in the 6 papers that comprise this Special Issue as a useful starting point for the dialectic that is required for a robust systems biology approach to the study of the genetic architecture of complex traits. We do not wish to constrain the approaches that will emerge from this undertaking by funneling the resulting research through these views, but instead we hope to encourage new ideas and approaches, and thus promote the contents of the Special Issue as an invitation to the necessary dialectic. Regardless of the approaches that are undertaken, it is important to recognise from the outset that their merits will ultimately be judged in terms of the capacity to improve the relationship between predicted and realised response to selection and the performance of the products from the breeding program. We encourage anyone who is working in this area to step back at appropriate times to use these benchmarks to judge the progress that is being made.

References

- Allard RW (1999) 'Principles of plant breeding.' 2nd edn (John Wiley & Sons, Inc.: Brisbane, Qld)
- Axelrod R, Cohen MD (1999) 'Harnessing complexity organizational implications of a scientific frontier.' (The Free Press: New York)
- Cahill DJ, Schmidt DH (2004) Use of marker assisted selection in a product development breeding program. In 'Proceedings of the 4th International Crop Science Congress'. 26 Sept.–1 Oct. 2004, Brisbane, Australia. Invited Symposium Paper: 'New Directions for a Diverse Planet'. (CD-ROM, Web site: www.cropscience.org.au)
- Comstock RE (1996) 'Quantitative genetics with special reference to plant and animal breeding.' (Iowa State University Press: Ames, IA)
- Comstock RE, Robinson HF, Harvey PH (1949) A procedure designed to make maximum use of both general and specific combining ability. *Agronomy Journal* **41**, 360–367.
- Cooper M, Chapman SC, Podlich DW, Hammer GL (2002) The GP problem: quantifying gene-to-phenotype relationships. *In Silico Biology* **2**, 151–164.
- Cooper M, Hammer GL (1996) Synthesis of strategies for crop improvement. In 'Plant adaptation and crop improvement'. (Eds M Cooper, GL Hammer) pp. 591–623. (CAB International, ICRISAT & IRRI: Wallingford, UK)
- Cooper M, Podlich DW, Smith OS (2004) Complex traits and gene to phenotype models. In 'Proceedings of the 4th International Crop Science Congress'. 26 Sept.–1 Oct. 2004, Brisbane, Australia. Invited Symposium Paper: 'New Directions for a Diverse Planet'. (CD-ROM, Web site: www.cropscience.org.au)
- Cooper M, Podlich DW, Smith OS (2005) Gene-to-phenotype models and complex trait genetics. *Australian Journal of Agricultural Research* **56**, 895–918.

- Coors JG (1999) Selection methodologies and heterosis. In 'The genetics and exploitation of heterosis in crops'. (Eds JG Coors, S Pandey) pp. 225–245. (ASA-CSSA-SSSA: Madison, WI)
- Duvick DN, Smith JSC, Cooper M (2004) Long-term selection in a commercial hybrid maize breeding program. *Plant Breeding Reviews* **24**, 109–151.
- van Eeuwijk FA, Malosetti M, Yin X, Struik PC, Stam P (2005) Statistical models for genotype by environment data: from conventional ANOVA models to eco-physiological QTL models. *Australian Journal of Agricultural Research* **56**, 883–894.
- Falconer DS (1960) 'Introduction to quantitative genetics.' (Oliver and Boyd: Oliver and Boyd: Edinburgh, UK)
- Falconer DS, Mackay TFC (1996) 'Introduction to quantitative genetics.' 4th edn (Longman: Essex, UK)
- Fehr WR (Ed.) (1984) 'Genetic contributions to yield gains of five major crop plants.' CSSA Special Publication No. 7. (Crop Science Society of America: Madison, WI)
- Hallauer AR, Miranda, JB Fo. (1988) 'Quantitative genetics in maize breeding.' 2nd edn (Iowa State University Press: Ames, IA)
- Hammer GL, Chapman S, van Oosterom E, Podlich DW (2005) Trait physiology and crop modelling as a framework to link phenotypic complexity to underlying genetic systems. *Australian Journal of Agricultural Research* **56**, 947–960.
- Hammer GL, Kropff MJ, Sinclair TR, Porter JR (2002) Future contributions of crop modelling—from heuristics and supporting decision-making to understanding genetic regulation and aiding crop improvement. *European Journal of Agronomy* **18**, 15–31. doi: 10.1016/S1161-0301(02)00093-X
- Hanson WD, Robinson HF (1963) 'Statistical genetics and plant breeding.' Publication 982. (National Academy of Sciences – National Research Council: Washington, DC)
- Ideker T, Thorsson V, Ranish JA, Christmas R, Buhler J, Eng JK, Bum-garner R, Goodlett DR, Aebersold R, Hood L (2001) Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. *Science* **292**, 929–934. doi: 10.1126/science.292.5518.929
- Kearsey MJ, Pooni HS (1996) 'The genetical analysis of quantitative traits.' (Chapman and Hall: London)
- Kuhn TS (1996) 'The structure of scientific revolutions.' 3rd edn (The University of Chicago Press: Chicago, IL)
- Lynch M, Walsh B (1998) 'Genetics and analysis of quantitative traits.' (Sinauer Associates, Inc.: Sunderland, MA)
- Mather K, Jinks JL (1971) 'Biometrical genetics.' 2nd edn (Chapman and Hall Ltd: London)
- Niebur WS, Rafalski JA, Smith OS, Cooper M (2004) Applications of genomics technologies to enhance rate of genetic progress for yield of maize within a commercial breeding program. Invited Symposium Paper: 'New Directions for a Diverse Planet'. In 'Proceedings of the 4th International Crop Science Congress'. 26 Sept.–1 Oct. 2004, Brisbane, Australia. (CD-ROM, Web site www.cropscience.org.au)
- Reymond M, Muller B, Leonardi A, Charcosset A, Tardieu F (2003) Combining quantitative trait loci analysis and an ecophysiological model to analyse the genetic variability of the responses of leaf growth to temperature and water deficit. *Plant Physiology* **131**, 664–675. doi: 10.1104/pp.013839
- Sinclair TR, Seligman NG (1996) Crop modelling: from infancy to maturity. *Agronomy Journal* **88**, 698–704.
- Sing CF, Stengård JH, Kardia SLR (2003) Genes, environment, and cardiovascular disease. *Arteriosclerosis, Thrombosis, and Vascular Biology* **23**, 1190–1196. doi: 10.1161/01.ATV.0000075081.51227.86
- Strohman RC (1997) The coming Kuhnian revolution in biology. *Nature Biotechnology* **15**, 194–200. doi: 10.1038/nbt0397-194
- Tardieu F (2003) Virtual plants: modelling as a tool for the genomics of tolerance to water deficit. *Trends in Plant Science* **8**, 9–14. doi: 10.1016/S1360-1385(02)00008-0
- Tardieu F, Reymond M, Muller B, Granier C, Simonneau T, Sadok W, Welcker C (2005) Linking physiological and genetic analyses of the control of leaf growth under changing environmental conditions. *Australian Journal of Agricultural Research* **56**, 937–946.
- Walsh B (2005) The struggle to exploit non-additive variation. *Australian Journal of Agricultural Research* **56**, 873–881.
- Welch SM, Dong Z, Roe JL, Das S (2005) Flowering time control: gene network modelling and the link to quantitative genetics. *Australian Journal of Agricultural Research* **56**, 919–936.
- Yin X, Struik PC, Kropff MJ (2004) Role of crop physiology in predicting gene-to-phenotype relationships. *Trends in Plant Science* **9**, 426–432. doi: 10.1016/j.tplants.2004.07.007

Manuscript received 9 May 2005, accepted 20 June 2005