

Trait physiology and crop modelling as a framework to link phenotypic complexity to underlying genetic systems

Graeme L. Hammer^{A,B,E}, Scott Chapman^C, Erik van Oosterom^A, and Dean W. Podlich^D

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

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

^CCSIRO Plant Industry, Queensland Bioscience Precinct, 306 Carmody Rd, St Lucia, Qld 4067, Australia.

^D7250 NW 62nd Ave, PO Box 552, Pioneer Hi-Bred International Inc., Johnston, IA 50131, USA.

^ECorresponding author. Email: g.hammer@uq.edu.au

Abstract. New tools derived from advances in molecular biology have not been widely adopted in plant breeding for complex traits because of the inability to connect information at gene level to the phenotype in a manner that is useful for selection. In this study, we explored whether physiological dissection and integrative modelling of complex traits could link phenotype complexity to underlying genetic systems in a way that enhanced the power of molecular breeding strategies. A crop and breeding system simulation study on sorghum, which involved variation in 4 key adaptive traits—phenology, osmotic adjustment, transpiration efficiency, stay-green—and a broad range of production environments in north-eastern Australia, was used. The full matrix of simulated phenotypes, which consisted of 547 location–season combinations and 4235 genotypic expression states, was analysed for genetic and environmental effects. The analysis was conducted in stages assuming gradually increased understanding of gene-to-phenotype relationships, which would arise from physiological dissection and modelling. It was found that environmental characterisation and physiological knowledge helped to explain and unravel gene and environment context dependencies in the data. Based on the analyses of gene effects, a range of marker-assisted selection breeding strategies was simulated. It was shown that the inclusion of knowledge resulting from trait physiology and modelling generated an enhanced rate of yield advance over cycles of selection. This occurred because the knowledge associated with component trait physiology and extrapolation to the target population of environments by modelling removed confounding effects associated with environment and gene context dependencies for the markers used. Developing and implementing this gene-to-phenotype capability in crop improvement requires enhanced attention to phenotyping, ecophysiological modelling, and validation studies to test the stability of candidate genetic regions.

Additional keywords: gene-to-phenotype modelling, complex traits, molecular breeding, virtual plants.

Introduction

The enhanced ability to undertake genome-scale molecular biology and accumulate associated data has not been matched by an improved ability to design and engineer improved complex traits in crop plants. Molecular biology has delivered commercial successes in enabling crop plants to better resist pests and tolerate herbicides and, more recently, in improving product quality. These achievements have been made via single gene transformations that scale well from molecular expression to plant level response (Somerville and Somerville 1999; Cahill and Schmidt 2005). Although there is increasing interest in molecular breeding for more complex crop growth and development traits, successful approaches

based on major genes are not readily adapted to more complex quantitative traits (Snape 2004). The challenge remains as to how to manipulate more complex growth and development traits associated with crop adaptation and yield.

The inability to connect information at gene level to the expressed phenotype in a manner that is useful for selection and plant breeding has restricted adoption of molecular approaches in plant breeding (Mifflin 2000). Enhanced capabilities in genotyping have not been matched by development of enhanced capabilities in phenotyping or by development of enhanced approaches to link genotype and phenotype (Campos *et al.* 2004). The situation is particularly complicated for complex growth and development traits, as

they are associated with genes interacting in networks so that gene-to-phenotype relationships are not straightforward, as shown in rice by Li *et al.* (2001) and Luo *et al.* (2001). The functional consequences for the organism, which arise from the interplay of the organism–environment system with these gene networks, occur at a higher level of biological organisation. Hence, gene \times gene and gene \times environment interactions have major influences on the expressed phenotype. Such context dependencies provide a major limitation for molecular breeding (Podlich *et al.* 2004).

Integrating across biological levels of organisation using a gene-to-phenotype modelling approach may present a way forward, but it also presents a major challenge (Chapman *et al.* 2002b; Cooper *et al.* 2002). The notion of a virtual or *in silico* plant suitable for this purpose has occupied the thinking of molecular biologists (Minorsky 2003) and whole-plant physiologists (Hammer *et al.* 2002, 2004; Tardieu 2003; Yin *et al.* 2004). The former suggest a path integrating from the gene and gene function level up to the organism phenotype (i.e. ‘bottom-up’), whereas the latter begins with the phenotype and uses physiological dissection to drive towards the molecular genomic level (i.e. ‘top-down’). We have argued elsewhere (Hammer *et al.* 2004) for informed dialectic across levels of biological organisation to enhance progress in systems biology and in developing the form of *in silico* plant most useful to crop improvement. We suggest that the bottom-up approach is likely to suffer from an inability to deal with complexity generated by gene and environment context dependencies, especially for complex growth and development traits. Although some recent gene network models [e.g. galactose metabolism in yeast (Peccoud *et al.* 2004); transition to flowering in *Arabidopsis* (Welch *et al.* 2003)] show promise for this approach (see Welch *et al.* 2005), they are based on extensive underpinning research that has characterised relevant networks and pathways, and they still relate only to components of whole-organism function (i.e. they do not deal with complex integrated traits such as yield).

Crop models with generic approaches to underlying physiological processes (e.g. Wang *et al.* 2002) underpin the top-down approach to gene-to-phenotype modelling (Hammer *et al.* 2002). Crop simulation models have captured much of the understanding of plant growth and development processes generated over nearly 40 years of plant systems research (Sinclair and Seligman 1996). They provide a dynamic framework for the physiological dissection of complex growth and development traits. However, initial attempts (White and Hoogenboom 1996; Yin *et al.* 2003) to use agronomic crop models for this purpose by optimising a range of model coefficients so that the model best fitted observed phenotypic variation among sets of genotypes, had limited success. The modest predictive capabilities found highlighted the need to better understand the

physiological basis of the genetic variation involved via studies with controlled genetic backgrounds before seeking such predictive capability across diverse material. The studies also showed the reliance of the approach on the validity with which the crop model architecture and associated coefficients captured and integrated the physiological basis of the genetic variation.

Tardieu (2003) demonstrated that these limitations could be overcome in recent studies that used an ecophysiological model of plant water use. The parameters of control equations in the simple, but physiologically robust, component model were conclusively linked to genetic variation by their ability to predict the behaviour of transformed plants. The model parameters thus represented coordinated genotypic responses that quantified a ‘meta-mechanism’ at a higher level of biological organisation. In a similar manner, Reymond *et al.* (2003) and Tardieu *et al.* (2005) combined quantitative trait locus (QTL) analysis with an ecophysiological model of the response of maize leaf elongation rate to temperature and water deficit by conducting the QTL analysis on the model parameters. They were able to validate the QTLs by successfully predicting the elongation rates of new lines in the mapping population using their QTL profile to determine relevant parameters for use in the component model for leaf elongation. It remains to be seen whether this demonstrated gene-to-phenotype capability of modelling at organ or component scale can be successfully applied at the organism scale. It seems clear that crop models that better capture plant function and control will be required for this task (Hammer *et al.* 2002; Yin *et al.* 2004). However, the question still remains as to the merit of this approach in relation to impact on selection and efficacy in plant breeding – the yardstick noted in the introduction to this series (Cooper and Hammer 2005).

In this study we explore the question of whether using a crop growth and development modelling framework can link phenotypic complexity to underlying genetic systems in a way that enhances the power of molecular breeding strategies. We approach this question by using a crop and breeding system simulation study on sorghum to consider the potential value to marker-assisted selection (MAS) of the understanding and predictive power that might be generated by physiological dissection and integrative modelling of complex traits. The steps in this approach are:

- (1) linking genetic variation in the key adaptive traits considered to their physiological determinants;
- (2) simulating crop phenotypes and classifying production environments associated with the range of potential genotypes and the target population of environments (TPE);
- (3) analysing genetic effects in the simulated phenotypic data, assuming differing levels of knowledge of the underlying gene-to-phenotype relationships;

- (4) simulating effectiveness of MAS breeding strategies, assuming differing levels of knowledge and integrative modelling of the underlying gene-to-phenotype relationships.

We discuss the key issues associated with developing and implementing this gene-to-phenotype capability in crop improvement.

This study uses simulated phenotype data from a previous gene-to-phenotype modelling and simulation study at crop scale for sorghum (Chapman *et al.* 2002a, 2002b, 2003; Cooper *et al.* 2002) to address this more general question. In those studies we focussed on how gene \times gene and gene \times environment interactions that arise from the underlying determinants of complex crop growth and development traits would challenge molecular breeding. The dynamic biophysical cropping systems model (APSIM; Keating *et al.* 2003) was used to simulate sorghum phenotypes for a broad range of water-limited production environments in Australia, based on assumed levels of variation in 15 genes controlling 4 adaptive traits. The number of genes involved in controlling each specific trait was determined from best available evidence. ‘Virtual genotypes’ were created by deriving combinations of ‘expression states’ for each trait. The expression states were determined from the number of positive alleles (i.e. alleles that increased trait phenotype) in the set of genes associated with each trait. The expression states were linked with crop model coefficients that quantified their effects on the specific traits. By simulating a range of virtual genotypes over a range of production environments, a database of simulated phenotypes was generated. The database of simulated phenotypes was linked to the QU-GENE breeding system simulation platform (Podlich and Cooper 1998) to explore effects of cycles of selection on yield gain for a range of selection scenarios. We use the same database and linkage to QU-GENE to simulate the MAS breeding strategies considered in this study.

Linking genetic variation in adaptive traits to physiological determinants

A robust physiological determinants framework for crop growth and development provides a means to analytically dissect phenotypic variation to aid understanding of underlying causes while simultaneously providing a means to predict emergent phenotypic consequences by integrating effects of variation in component factors and processes. Crop modelling based around the continued improvement of initial framework concepts (e.g. Charles-Edwards 1982; Passioura 1983; Tanner and Sinclair 1983; Monteith 1986, 1988; Sinclair and Horie 1989) has now evolved to a level where the simultaneous pursuit of explanation and prediction at the whole-crop level is possible (Hammer *et al.* 2002). The yield of a determinate grain crop is determined by the

number of grains set and the size they are able to achieve. Both factors are strongly determined by rate of crop growth, but at differing stages of development: the former around flowering and the latter after flowering. Grain size is also influenced by the retranslocation of pre-anthesis assimilate. The timing of flowering is controlled by interplay of genetics and environment, particularly temperature and photoperiod. The rate of biomass accumulation relates to ability to capture light or water and the efficiency with which either can be used. Both depend on the nature of canopy development, which is influenced by crop development (via maximum leaf number), management (via density), temperature, and genetics (e.g. tillering, branching, architecture). Capture of water depends on the nature of the soil profile, its rate of exploration by roots, and their ability to extract water. The efficiency of water use is affected by the aerial environment via the influence of vapour pressure deficit on water flux. The efficiency of radiation use is affected by leaf nitrogen status, which in turn is related to the availability of nitrogen to the crop. The interplay of the developmental timetable for organ growth with the availability of assimilate, water, and nitrogen determines crop water and nitrogen status and growth and development patterns. The balances between demand for, and supply of, these factors among organs underpin organism controls of allocation and growth. Appropriately specifying and quantifying the response and control equations of this general framework are critical in effective crop modelling. This is akin to modelling plant hormone action without modelling the hormones (de Wit and Penning de Vries 1983). The specifications provide both the basis to identify and quantify key differences among lines (genotypes) in field studies, and the ability to predict emergent consequences of putative genetic variation on the phenotype.

In this study, we focussed on the physiology and genetics of 4 key adaptive traits for sorghum—phenology, osmotic adjustment, transpiration efficiency, stay-green—and placed the known genetic variation in the context of the physiological determinants framework of our crop models that had evolved progressively as insights improved (Chapman *et al.* 1993; Hammer 1998; Hammer *et al.* 2001; Wang *et al.* 2002).

Phenology

Differences in the rate of development among sorghum genotypes are known to relate to differing responses to temperature and photoperiod (Hammer *et al.* 1989; Major *et al.* 1990). Genetic variation in phenology can be predicted by quantifying these responses with photo-thermal models (e.g. Craufurd *et al.* 1999). Differences in duration before floral initiation will generate differences in number of leaves produced, with consequent effects on canopy leaf-area development (Muchow and Carberry 1990; Hammer *et al.* 1993) and thus, patterns of water use through the crop cycle.

A major drawback of this photo-thermal framework is that it does not account properly for some of the effects of temperature \times photoperiod interactions on phenology that have been observed for sorghum. For example, the photo-thermal model predicts that thermal time to anthesis and final leaf number of a hybrid are independent of temperature under a given photoperiod, yet temperature effects, independent of photoperiod, have been reported (Caddel and Weibel 1971; Major *et al.* 1990; Morgan *et al.* 2002). Consistent with this, phenology is accelerated under natural asynchrony between thermoperiod and photoperiod (i.e. temperature increases after transition from dark to light), but is slowed under 'unnatural' asynchrony (Morgan *et al.* 1987; Ellis *et al.* 1997).

There is some potential to capture these interactions using a gene network model, as illustrated by Welch *et al.* (2003). They could generate genotype \times temperature interactions for timing of transition to flowering in *Arabidopsis thaliana*, using a gene network model. Transition to flowering in *Arabidopsis* is determined by several converging pathways, including a facultative long-day and a photoperiod-independent autonomous pathway, which primarily responds to temperature (Blázquez 2000). Photoperiod \times temperature interactions arise as a consequence of genes affecting both of these pathways (e.g. those affecting light receptors). Capturing these molecular interactions dynamically at the phenotypic level, however, requires appropriate quantitative knowledge to scale upwards from the molecular level.

In this study we used the photo-thermal models implemented in crop models for many species (Wang *et al.* 2002) and varied the thermal time required to reach floral initiation to simulate genetic differences. Hart *et al.* (2001) found 3 major QTLs for maturity in sorghum, so we assumed that 3 genes with 2 alleles per locus, acting in a simple additive manner, gave rise to the genetic variation, which ranged from 90 to 140 degree-days. This range was consistent with the range of maturity and leaf number known to occur in locally adapted sorghum hybrids (Hammer *et al.* 1989). Seven expression states spread uniformly over this range were associated with the number of positive alleles present over the 3 loci (i.e. 0–6 positive alleles; see Chapman *et al.* 2003 for details).

Osmotic adjustment

The active accumulation of osmolyte compounds, usually called osmotic adjustment (OA), results in a decrease of cell osmotic potential and thus in maintenance of water absorption and cell turgor pressure, which might contribute to sustaining physiological processes and improving yield in water-limited environments (Ludlow and Muchow 1990). A recent study of the physiological mode of action of OA in high and low OA sorghum lines (Snell 2004) found that, under specific water-limited conditions, high OA lines had

a greater ability to set grain and retranslocate carbohydrate from the stem to grain during grain filling. This effect may have resulted from enhanced maintenance of metabolic activity during grain set and filling. This result was consistent with earlier findings of Ludlow *et al.* (1990). However, it is known that the specific environmental circumstances conferring any advantage occur infrequently and are low-yielding situations, so that the overall value of this trait is low (Hammer *et al.* 1999; Snell 2004). Although this finding is consistent with recent questioning of the value of OA to crop yield under drought (Serraj and Sinclair 2002), it remains relevant to this study because of these genotype \times environment interactions.

In this study we generated genetic variation in OA by increasing the potential to set grain and the ability to retranslocate stem biomass to grain under specific moisture limitation conditions. The increase in grain number was generated by reducing the amount of crop biomass growth required between floral initiation and flowering to produce an individual grain from 0.00083 to 0.00075 g/grain. Enhanced remobilisation of assimilate from stem during grain filling was generated by increasing the fraction of stem biomass at flowering that was potentially available for retranslocation from 20% to 36%. Both mechanisms were only invoked under circumstances when the crop demand for moisture could not be met by the supply ability of the soil–root system. Although this range was consistent with results of field studies (Snell 2004), there remains some uncertainty concerning the degree of water limitation required to initiate the changes. Based on findings of Basnayake *et al.* (1995) we assumed that 2 genes with 2 alleles per locus, acting in a simple additive manner, gave rise to the genetic variation. Five expression states spread uniformly over this range were associated with the number of positive alleles present over the 2 loci (i.e. 0–4 positive alleles; see Chapman *et al.* 2003 for details).

Transpiration efficiency

The amount of biomass produced by sorghum plants per unit of water transpired, i.e. transpiration efficiency (TE), is known to vary among genotypes (Donatelli *et al.* 1992; Henderson *et al.* 1996; Hammer *et al.* 1997). It has also been shown that these genetic differences are maintained under water limitation (Mortlock and Hammer 2000). TE is inversely proportional to vapour pressure deficit of the atmosphere and once normalised for this effect the TE coefficient for sorghum is accepted as 9 Pa (Tanner and Sinclair 1983). In this study we generated genetic variation in TE by allowing the standard TE coefficient for sorghum to vary from 8 to 10 Pa. This range was consistent with the variation observed in experimental studies (Hammer *et al.* 1997; Mortlock and Hammer 2000). No QTLs or genes are known for transpiration efficiency in sorghum. We assumed that 5 genes were involved, with the expectation

that it may be complex as it represents an integrated measure at whole-plant level, i.e. genetic variation in TE could be related to several underlying causal factors. We assumed that 2 alleles per locus, acting in a simple additive manner, gave rise to the genetic variation. Eleven expression states spread uniformly over the range in TE coefficient were associated with the number of positive alleles present over the 5 loci (i.e. 0–10 positive alleles; see Chapman *et al.* 2003 for details).

Stay-green

Stay-green is the ability of leaves to retain their integrity (i.e. greenness) during the grain-filling period. In addition to cosmetic forms of stay-green (Thomas and Howarth 2000), genetic variation in stay-green has been related to improved yield under water-limited conditions in sorghum (Borrell *et al.* 2000b), which is the focus in this study. Three mechanisms of functional expression of stay-green have been identified (Borrell *et al.* 2000a; Thomas and Howarth 2000): delayed onset of senescence, reduced rate of senescence, and increased leaf area index (LAI) at anthesis. In sorghum, the expression of stay-green during grain filling can be viewed as a consequence of the balance between demand for nitrogen (N) by the grains and supply of N through soil N-uptake and translocation from vegetative plant parts, including stems and leaves. Leaf-N translocation occurs if grain-N demand cannot be met through soil-N uptake and stem-N translocation (van Oosterom *et al.* 2006b).

The onset of leaf senescence can be delayed by increased soil-N uptake during grain filling, as observed under terminal drought stress by Borrell and Hammer (2000). Such increased N-uptake could be associated with either increased water uptake (i.e. transpiration, T) or increased transpiration efficiency (TE) (A. K. Borrell, pers. comm., 2004) as at least one of these is required to explain enhanced biomass accumulation and yield of stay-green types. Another mechanism to delay onset of leaf senescence is through increased availability of stem-N for translocation, although this mechanism might compromise the leaf-N status if increased stem-N is not matched by increased total N-uptake.

Increased specific leaf nitrogen (SLN) or leaf area at anthesis can also reduce the rate of leaf senescence. Once leaf-N translocation starts, the amount of leaf area that needs to senesce in order to meet the grain N demand declines with improving leaf-N status (van Oosterom *et al.* 2006b). Genotypic variation in maximum SLN under optimum conditions has been observed for sorghum, and was associated with differences in leaf size (van Oosterom *et al.* 2006a). Increased leaf area at anthesis can be achieved through increased partitioning of dry matter and N to the leaves (Borrell and Hammer 2000), although there are also indications that increased LAI of stay-green hybrids is due to inability of other hybrids to compensate

for their smaller leaf area on the main shoot through tillering, particularly if resource availability is limited (van Oosterom *et al.* 2006a).

This framework can explain genotypic differences and effects of genotype \times environment interaction on the expression of stay-green. Phenotypic expression of stay-green becomes an emergent consequence of the interplay of differences in underlying traits such as leaf size, leaf SLN, dry matter partitioning, N uptake, and possibly transpiration or transpiration efficiency.

In this study we varied the target SLN of new leaf over the range 1.35–1.65 gN/m² leaf area to generate genetic variation in stay-green. This range was consistent with expectations from experimental studies (Borrell and Hammer 2000). Tao *et al.* (2000) identified 5 QTLs associated with stay-green in local sorghum germplasm, so we assumed that 5 genes with 2 alleles per locus, acting in a simple additive manner, gave rise to the genetic variation. Eleven expression states spread uniformly over this range were associated with the number of positive alleles present over the 5 loci (i.e. 0–10 positive alleles; see Chapman *et al.* 2003 for details). Although the physiological framework presented above suggests that variation in stay-green is associated with other underlying drivers, they were not included in this initial study. The interaction with variation in transpiration efficiency will be incorporated via the separate treatment of that trait. It is possible that combinations of the 5 known QTLs for stay-green relate to differing underlying mechanisms, but insufficient is known at this stage to proceed further in this regard.

Simulating phenotypes and classifying production environments

Attributes of production environments and phenotypes arising from specific trait combinations were generated by simulation. We conducted the simulation studies using the sorghum module of the APSIM modelling platform (Keating *et al.* 2003), which uses the generic framework outlined by Wang *et al.* (2002). The sorghum module has undergone extensive development to enhance its capacity to realistically simulate the interactions among physiological processes (Hammer *et al.* 2001). Recent improvements have adopted concepts of 'emergent' properties (Hammer 1998) in seeking more realistic simulation of genetic variation in traits via underlying physiological functionality.

Environment types were characterised by quantifying the degree of crop water limitation simulated throughout the crop cycle. Development and growth of a reference genotype were simulated for a large sample (547) of location–season combinations and the resulting seasonal patterns of water limitation were clustered into like types (Chapman *et al.* 2002a). Three distinct patterns were identified: mild terminal stress (MTS), severe terminal stress (STS), and mid-season stress (MSS) (Fig. 1). The MTS environment

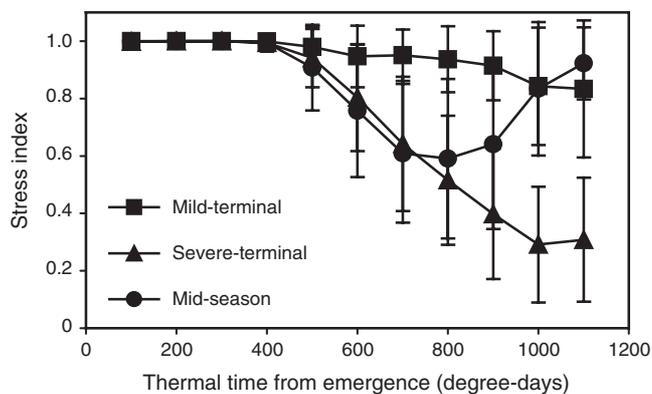


Fig. 1. Patterns of water limitation throughout the growing season associated with the 3 environment types identified from the simulation and cluster analysis of sorghum production environments in NE Australia (after Chapman *et al.* 2002a). The stress index is the ratio of water supply to crop demand. The lines show the mean stress index values for the location–season combinations making up each group and the vertical bars show the standard deviation. Mean anthesis date for the reference genotype used was 732 degree-days.

type occurred in 37% of location–season combinations and represented situations where little or no water limitation was experienced until well after anthesis. In contrast, the STS type (35% occurrence) reflected early onset of water limitation that became increasingly severe as the crop cycle progressed. The MSS type (28% occurrence) was associated with early onset of water limitation that was relieved during the grain-filling period.

The effects on average simulated yield, generated by differing combinations of positive alleles for the 4 traits incorporated, varied with environment type (Fig. 2). The yield outcomes for any specific allele combination represent the emergent consequence of the perturbation of the functional plant–environment dynamics contained in the model that is associated with the changes in specifications of response and control equations. As expected, the average yield was greater in MTS environments. It was notable, however, that although combining all positive alleles for phenology (late flowering) resulted in higher average yield in MTS environments, the opposite occurred in STS environments (Fig. 2). Hence, a clear genotype \times environment interaction was generated. Although this interaction is perhaps expected, it nonetheless highlights the point of the potential for considerable confounding in the absence of the environment classification. Chapman *et al.* (2000) illustrated how the sequence of environment types would confound the expression of 2 genotypes. The physiological basis of the interaction relates to the additional number of leaves and hence, greater canopy leaf area, generated with later maturity. This causes greater demand for water, which is detrimental in STS environments where water becomes limiting early in the growth cycle (Fig. 1). In contrast, in MTS environments, where water

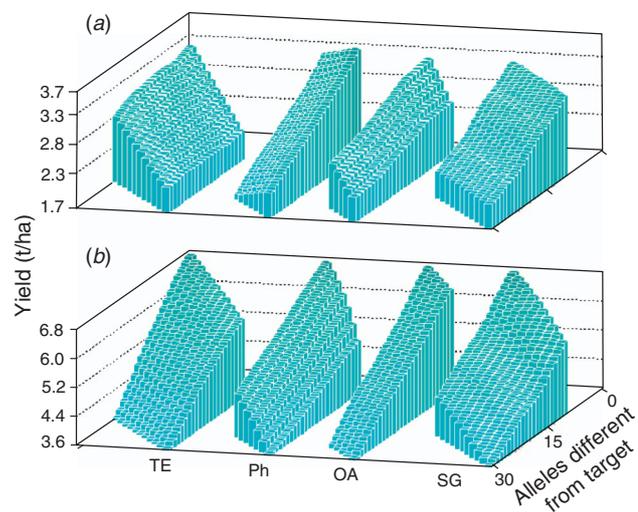


Fig. 2. Average simulated yield in all (a) severe terminal stress and (b) mild terminal stress environments for all genotype combinations. The target genotype is defined as that genotype containing all the positive alleles for all traits (i.e. those alleles that increase the numerical value of the trait phenotype). For each trait, the columns (running from the front to the back of the figure) depict average yields associated with a specific number of positive alleles for that trait (11 columns for TE, 7 for phenology, 5 for osmotic adjustment, 11 for stay-green) and decreasing degree of difference from the target genotype due to increasing contributions of positive alleles from the other 3 traits. For each trait, the column to the left relates to genotypes with the maximum number of positive alleles associated with that trait. The column to the right for each group has no positive alleles for that trait.

is not so limiting, the additional canopy leaf area is able to generate increased biomass accumulation and yield via enhanced light capture. There was also an interaction for the OA trait. Accumulating all positive alleles for OA generated some effect on yield in STS environments, but there was little effect in MTS environments. This reflects the need for specific environmental conditions for effective expression of this trait.

The value of combinations of traits varied with environment type (Fig. 2). In STS environments, positive alleles for TE were associated with higher yield on their own, whereas in MTS environments, their value only became evident when combined with positive alleles for other traits. This result reflects the over-riding value of more efficient use of water in generating biomass in water-limited environments. This trait delays the onset of water limitation as the available water resource is diminished less rapidly. Although not directly linked to the stay-green (SG) trait in this study, the TE trait is also likely to generate a stay-green phenotype by delaying the onset of senescence. This mechanism is of less importance in environments where water limitation is not as prevalent (MTS). Conversely, positive alleles for SG had little value in STS environments until combined with other traits but

had greater individual effect in MTS environments. This reflects the mechanism based on N dynamics used to underpin this trait in this study. The additional N in the SG type generates more value in situations where water is not often limiting and the enhanced accumulation of biomass, associated with delayed senescence and higher radiation use efficiency (RUE), can progress unimpeded. This result suggests that some other physiological mechanism (e.g. TE) is more likely responsible for causing stay-green in water-limited environments.

The simulated average phenotypic effects among environment types mask the large variability among individual environments (547 location–season combinations) and genotypes (4235 expression states). In reality, only a sample of both is usually available and the environment and gene context dependencies present in any sample hinder interpretation and progress in crop improvement. In this study, we have generated the full set of combinations (individual environments \times expression states) and use this to gain a more realistic appraisal of dealing with complex traits in variable environments. In essence, the simulated phenotypes characterise the performance or adaptation landscape that confronts a plant-breeding program (see Cooper *et al.* 2005).

Analysing genetic effects

To determine consequences of increasing physiological understanding on insight into genetic effects the complete data matrix of simulated phenotypes (4235 expression states, derived from allelic variation at 15 loci, for each of 547 environments, derived from location–season combinations) was subjected to a range of analyses (Fig. 3). In the first instance, a conventional quantitative genetics analysis of yield variance components (Comstock and Moll 1963; Cooper and DeLacy 1994) indicated near equal amounts of G and $G \times E$ interaction effects (top panel, Fig. 3). This reflects a common outcome faced in plant breeding programs when phenotypic evaluation is all that is available to guide selection. In this case though, the additive effects incorporated in relating allelic variation to variables influencing response and control equations in the crop model had generated the significant $G \times E$ interaction at the level of grain yield. Next, an analysis of gene effects on yield for the 15 genes involved indicated the degree to which each gene influenced the phenotypic variation for yield (second panel, Fig. 3). In this analysis the average effect size of the positive allele was calculated for each of the 15 genes across the entire dataset. It indicated what might be found from a QTL analysis with good statistical power (i.e. adequate population size and marker density over a large number of environments). The result showed that the genes had varying levels of effect on the phenotypic variation. This reflects the situation where molecular information can assist in identifying key genomic regions associated with

yield but the ecophysiological basis of the associations is unknown. Next, the typing of environments, based on using the crop model as a virtual entry in each trial (as discussed above), was incorporated in the analysis of gene effects. It was immediately clear that average gene effects varied substantially among environments (third panel, Fig. 3). In fact, genes with strong positive effects overall and in MTS and MSS environments had negative average effects in STS environments. The environment typing was able to better resolve effects of particular genes within a general type of environment, i.e. it started to unravel the $G \times E$ interaction for yield in a way that improved the potential value of molecular information. Finally, the association of genes with their broad physiological basis was incorporated in the analysis (bottom panel, Fig. 3). The average yield effects varied substantially among traits in the different environment types. The grouping of the genes based on their physiological association provides an even stronger basis to understand and use the $G \times E$ interaction observed at the yield level. It becomes clear that in this case all negative yield effects in STS environments are associated with phenology, whereas major positive yield effects in that environment type are associated with TE. Such additional information would underpin a focus on specific molecular information (e.g. candidate genes or QTLs) in particular environments (e.g. drought coincident with the development stage of flowering). It also provides a means to target relevant phenotyping in specific environment types.

While environmental characterisation and physiological knowledge help to explain and unravel gene and environment context dependencies, the analysis of average gene effects on yield across all environments masks some of the key effects of environmental variability. Although there is greater variability of gene effects on yield among environment types than within them, the gene effects still vary considerably within an environment type (Fig. 4). For many of the trait–environment type combinations, gene effects associated with individual environments can change sign. For phenology, the average gene effect is negative in STS and positive in both MSS and MTS environment types (Fig. 3). However, for individual environments within these groupings, gene effects can be positive in some STS environments and negative in some MSS and MTS environments. This inconsistency can cause major problems in the detection of QTLs and their effective use in molecular breeding. The result highlights the effect of environmental variability on QTL detection, which will be discussed further in the next section.

Investigating the power of breeding strategies

We simulated a marker-assisted selection (MAS) breeding strategy based on the analyses of gene effects presented above by linking the performance landscape of simulated phenotypes to the QU-GENE breeding system simulation

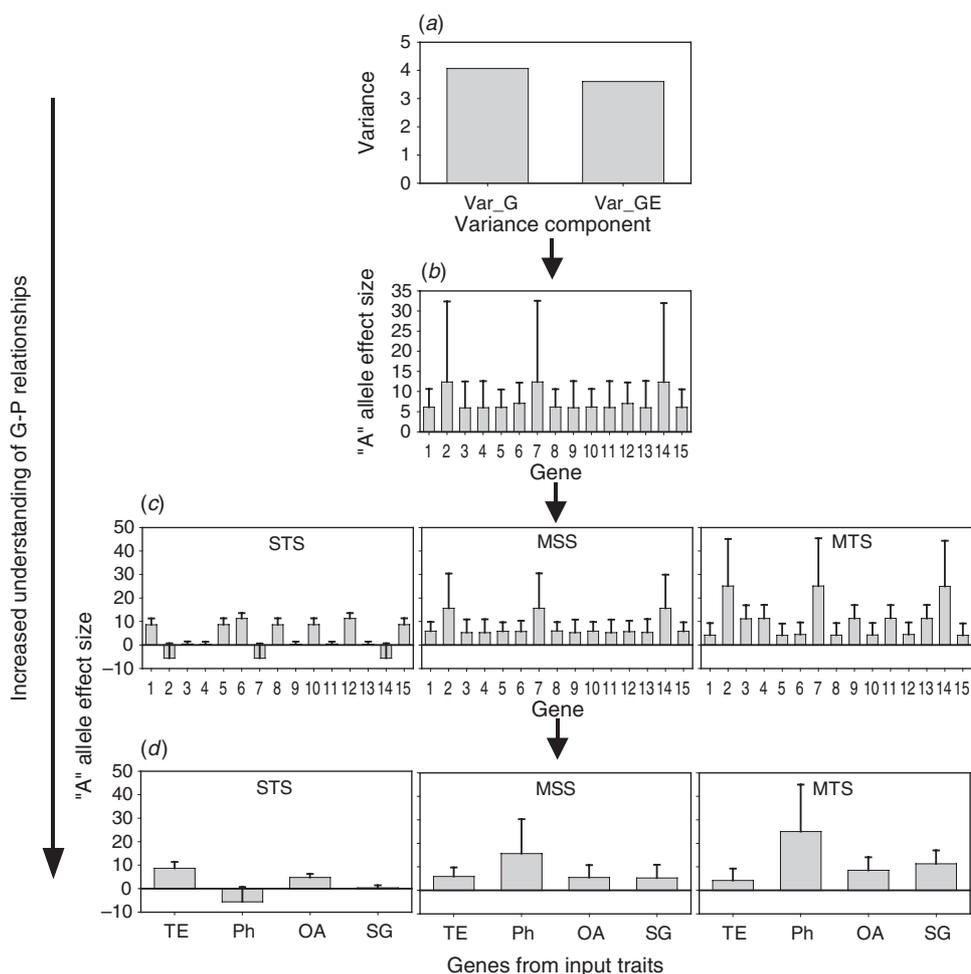


Fig. 3. Analysis of genetic and environmental effects with increasing understanding of gene-to-phenotype relationships. The top panel shows variance components analysis from conventional phenotypic analysis on yield for the whole dataset. The second panel extends this to analysis of gene effects for yield, similar to QTL analysis. The bars represent average yield effects across all environments and the line indicates the standard deviation of effect size. The third panel extends this to analysis of gene yield effects by environment type, which must be defined by simulation. The lower panel extends this by grouping gene yield effects by underpinning physiological traits, which requires enhanced knowledge of trait physiology and genetics. In all cases, gene yield effects are defined in relation to the positive allele for increasing trait value. A negative allele yield effect indicates that the opposite allele is defined as favourable for yield.

platform (Podlich and Cooper 1998; Cooper *et al.* 2005). The effects of all the genes were estimated at the start of the breeding process, and it was assumed that markers close to the genes were available. Marker scores were allocated based on the contribution of gene effects to yield in a single environment, i.e. a QTL analysis was conducted using yield outcomes in each of the environments in the target population of environments (TPE, Fig. 4), and MAS was then conducted based on that QTL analysis. Hence, the number of scenarios simulated was the same as the number of environments used in the study. MAS was implemented in a manner such that an equal amount of weighting was given to the genotypic and phenotypic information in the random environments sampled

from the TPE in each cycle of testing and selection in the breeding process.

There was a wide divergence in rate of yield gain in the TPE with breeding cycle, depending on the environment chosen for the QTL analysis (Fig. 5). This result reflected the fact that particular genes only demonstrated their greater value in particular environments, as noted above (Figs 3 and 4). Hence, selection of environments for definition of QTLs assumes considerable importance as the presence of undefined QTL \times environment interaction in the QTL detection phase can undermine the value of MAS. In particular, the yield gains associated with QTL analyses in STS environments were generally the lowest.

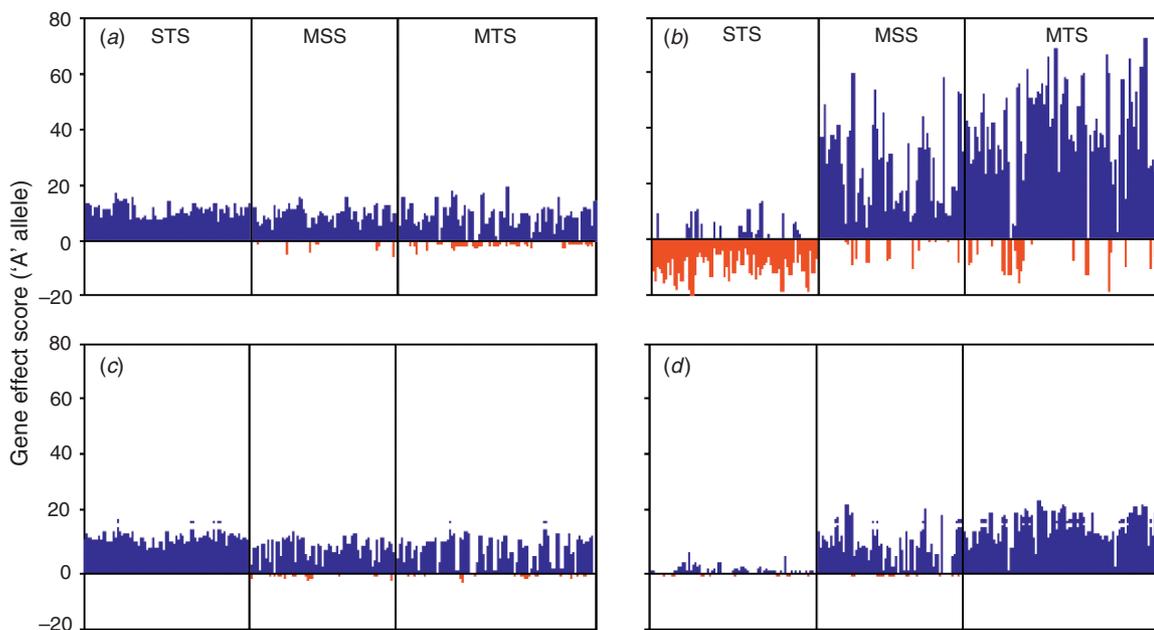


Fig. 4. Analysis of gene effects for individual environments in the 3 environment types (severe terminal stress, STS; mid-season stress, MSS; mild terminal stress, MTS) for genes associated with (a) transpiration efficiency, (b) phenology, (c) osmotic adjustment, and (d) stay-green. Blue bars represent positive effects (of the positive allele) and red bars represent negative effects.

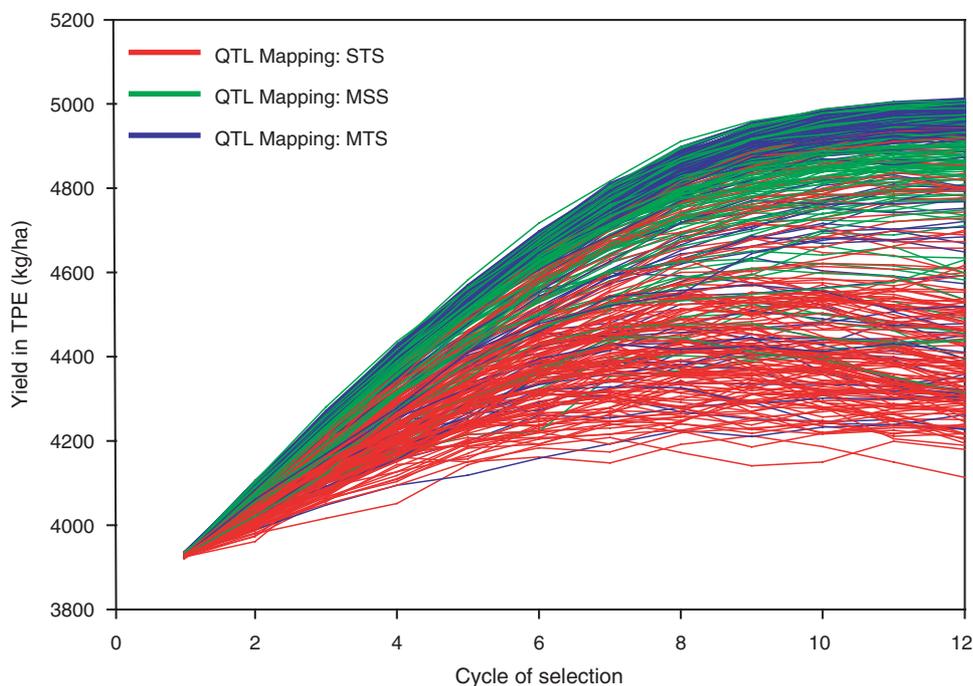


Fig. 5. Average yield in the target population of environments (TPE) over 12 cycles of selection for a range of marker-assisted selection scenarios. Each trajectory represents results using a QTL analysis based on a single environment (i.e. one location–season combination out of the 547 total combinations considered). All 547 environments in the TPE were used to derive the range of scenarios examined. Trajectories are coloured depending on the environment type of the location–season combination used for the underpinning QTL analysis.

In these environments, the favoured alleles for phenology were opposite to those required for yield advance in the TPE, i.e. for those situations, the QTL analysis would result in selection pressure for the wrong alleles and lead to a suboptimal response to selection in the TPE. Although the use of a single environment for QTL definition might be considered extreme, and the phenology \times environment interaction as not unexpected, the example highlights what might occur in situations where the underlying context of the QTL effects on yield is not known.

The knowledge associated with trait physiology and modelling the association of QTLs with component traits and yield clarifies the QTL \times environment issues that confound this outcome with MAS. Such information enhances confidence in QTLs either by improved awareness of the importance of the environment type used in their definition (i.e. environment context dependency) or by improved awareness of association to specific traits (i.e. gene context dependency) and the relative importance of those traits in the TPE. The potential value and impact of use of this knowledge in breeding was examined by simulating cycles of MAS assuming increasing levels of knowledge associated with the markers (Fig. 6). In the first instance, it was assumed only that markers were detected and used in MAS with equal weighting, i.e. their relative weightings and importance in specific environments were not known. Second, use of the markers in MAS was weighted by the effect sizes found in the genetic analysis associated with their detection. In this case, the weightings would vary depending on the contribution of gene effects to yield in the single experiment (i.e. environment) used for detection (as for Fig. 5). Lastly, the markers were weighted based on the relative value of

their associated trait in the TPE. This requires sufficient physiological understanding to link markers to associated physiological traits and to underpin a quantitative modelling capability to predict consequences of variation in these component physiological traits in the TPE.

There are significant differences in rate of advance in average yield in the TPE with cycles of selection among these 3 MAS scenarios (Fig. 6). The first 2 scenarios require no physiological understanding and integrative modelling. On average, the use of weighted marker selection gave a slight advantage over MAS without weighting. This occurred despite the large variability in weightings associated with individual environments (Fig. 5). The responses shown in Fig. 6 represent yield advance averaged over the full set of individual simulations. When physiologically weighted marker selection was used, the rate of average yield advance in the TPE was significantly enhanced and reached a higher level. This occurred because the knowledge associated with component trait physiology and extrapolation to the TPE by modelling removed confounding effects associated with environment and gene context dependencies for the markers used.

Can trait physiology and modelling add value to plant breeding?

Although the underlying genetic controls assumed for the traits used in this analysis were very simple, the results nonetheless demonstrated that trait physiology and integrative modelling could add value to plant breeding by unravelling environment and gene context dependencies that cause inefficiencies in MAS. The success of molecular breeding relies on an effective prediction of phenotypic variation based on allelic variation. Current approaches to MAS for complex traits rely heavily on the use of statistical approaches that are based on linear models (see discussion by van Eeuwijk *et al.* 2005). Their predictive power is poor when interactions among genes and/or environments (i.e. context dependencies) are important. The added value from a trait physiology and modelling framework arises because consequences of these interactions on the resultant phenotype are an emergent property of the framework. Hence, predictive power and the effective implementation of MAS are enhanced.

To realise the added value of a trait physiology and modelling framework will require its effective integration into plant-breeding programs. Better characterising production environments and the more effective use of molecular markers by defining intrinsically stable QTLs associated with complex traits offer realistic initial targets. Beyond this, there are opportunities for guiding gene discovery and for improved evaluation of potential of specific transgenics. Environment characterisation requires improved attention to soil and climate conditions encountered in breeding trials so that models can be used as virtual

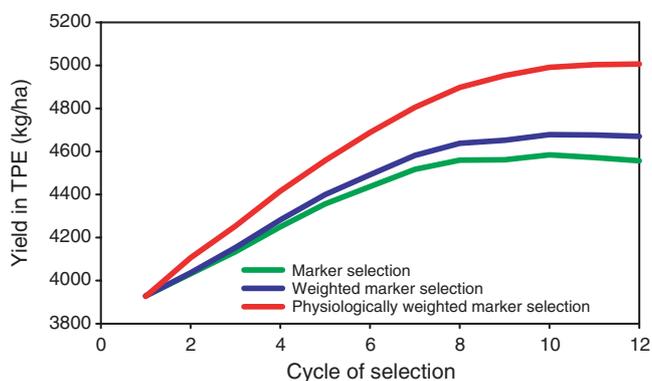


Fig. 6. Average yield in the target population of environments (TPE) over 12 cycles of selection for 3 marker-assisted selection scenarios. The trajectories for marker selection and weighted marker selection represent the average result over individual breeding system simulations based on a QTL analysis from each single environment. The trajectory for physiologically weighted marker selection represents the average result over simulations where markers have been assigned to physiological traits and marker weights have been derived from the simulated value of that trait in the TPE.

entries. This facilitates weighting of particular trials relative to their importance in unravelling genotype \times environment interactions and their representation in the TPE (Chapman *et al.* 2000). However, to define stable QTLs requires far more attention to phenotyping than is normal practice in breeding programs. It is likely that specialist studies facilitating more in-depth physiological dissection among lines of interest will be required. This could be combined with broader screening in breeding populations using relevant managed environments and selection indices that reflect the underpinning physiological basis for trait variation (Campos *et al.* 2004). It will be necessary to test the stability of QTLs identified in this way using validation studies based on their predictive capability. Reymond *et al.* (2003) report an example of this approach. They used an ecophysiological model to identify QTLs for control of leaf growth in maize and then validated their stability by using the QTLs to predict responses of other genotypes. The challenge is now to determine whether this approach can be effective at the whole-plant/crop level for more complex traits.

To be useful, the physiological frameworks used for trait dissection and modelling at whole-plant/crop level must realistically capture the functional basis of the genetic variation for complex traits of interest. Most existing crop models, which were constructed to deal mostly with agronomic issues, are not well structured in this regard, as found by Dingkuhn (1996) for carbon and nitrogen partitioning, by Jeuffroy *et al.* (2002) for capture and use of nitrogen, and as noted by Yin *et al.* (2004) in their recent review. The crop physiological modes of action of the complex trait must be understood and quantified and the crop model must be sufficiently detailed to simulate the consequences on growth and development generated by the interaction of those modes of action with the environment (Hammer *et al.* 1996). Appropriately specifying and quantifying the response and control equations are critical in effective crop modelling. The control equations most likely reflect the basis of metabolic signalling in plants and thus provide focal points for links to underlying genetic systems (Hammer *et al.* 2002). Tardieu (2003) demonstrated this concept using an ecophysiological model of plant water use. He identified stable meta-mechanisms at the plant level that reflected the parameterisation of the response and control equations of the model. Hammer *et al.* (2004) argue that these meta-mechanisms provide the bridge across levels of biological organisation that will link molecular biology and crop improvement via *in silico* plant technologies.

There remains some uncertainty as to the degree of residual aggregation in meta-mechanisms in crop models that will be most effective for linking phenotype complexity to underlying genetic systems. Modelling concepts framed around source–sink balance among organs and regulation of the supply–demand dynamics for carbon, nitrogen, and water provide a pathway forwards (Dingkuhn 1996) but

their implementation may become highly parameterised (e.g. Drouet and Pages 2003). We suggest that a key feature of the development of effective models in this domain will be the retention of simplicity, while simultaneously improving the rigour and generality of dealing with functional control. Our operating hypothesis is that physiologically sound but simple whole-crop models will provide the balance between capturing process understanding and the predictive utility needed to effectively link phenotype to genotype. The meta-mechanisms identified using such models will remain some distance away from gene complexes on the scale of biological organisation, but by combining with advanced statistical quantitative genetics approaches (e.g. Podlich *et al.* 2004; van Eeuwijk *et al.* 2005) they should provide sufficient unravelling of environment and gene context dependencies to have significant effect on breeding strategies.

There is increasing advocacy for the application of crop physiological knowledge and integrative modelling in breeding for complex traits (Hammer *et al.* 2002, 2004; Campos *et al.* 2004; Edmeades *et al.* 2004; Yin *et al.* 2004). The results of this crop and breeding system simulation study support this view as the analysis indicates potential rates of yield improvement significantly greater than obtained through established selection systems. The simplifying assumptions used here have undoubtedly generated significant departure from the complexities of the real situation. The example represents one possible point on the complexity continuum of the genetic state-space discussed by Cooper *et al.* (2005). However, the promise of using trait physiology and crop modelling to link phenotypic complexity to underlying genetic systems is clearly evident.

Conclusions

Use of a crop growth and development modelling framework can link phenotype complexity to underlying genetic systems and indicates ways to enhance the power of molecular breeding strategies. Such a framework facilitates meaningful characterisation of production environments and physiological dissection of complex traits. Both aspects aid in identifying intrinsically stable QTLs by reducing the incidences of undefined environment and gene context dependencies, which inhibit the utility of traditional statistical approaches used in molecular breeding. Implementing this gene-to-phenotype capability in crop improvement will require enhanced attention to phenotyping, developing the robust ecophysiological modelling framework required, and conducting the validation studies needed to test the stability of QTLs identified.

Acknowledgments

An earlier version of this paper was published in the Proceedings of the 4th International Crop Science Congress, held in Brisbane, 26 September–1 October 2004. We thank the Congress organisers for permission to publish this

updated manuscript as part of this series, which was based on presentations made to a symposium forming part of that Congress. We also thank Mark Cooper for invaluable discussions on framing the ideas and analyses reported in this paper, and Jeremy Lecoeur for many useful suggestions on an earlier version of this manuscript.

References

- Basnayake J, Cooper M, Ludlow MM, Henzell RG, Snell PJ (1995) Inheritance of osmotic adjustment in three grain sorghum crosses. *Theoretical and Applied Genetics* **90**, 675–682. doi: 10.1007/BF00222133
- Blázquez M (2000) Flower development pathways. *Journal of Cell Science* **113**, 3547–3548.
- Borrell AK, Hammer GL (2000) Nitrogen dynamics and the physiological basis of stay-green in sorghum. *Crop Science* **40**, 1295–1307.
- Borrell AK, Hammer GL, Douglas ACL (2000a) Does maintaining green leaf area in sorghum improve yield under drought? I. Leaf growth and senescence. *Crop Science* **40**, 1026–1037.
- Borrell AK, Hammer GL, Henzell RG (2000b) Does maintaining green leaf area in sorghum improve yield under drought? 2. Dry matter production and yield. *Crop Science* **40**, 1037–1048.
- Caddel JL, Weibel DE (1971) Effect of photoperiod and temperature on the development of sorghum. *Agronomy Journal* **63**, 799–803.
- Cahill DJ, Schmidt DH (2005) Use of marker assisted selection in a product development breeding program. In 'New directions for a diverse planet'. Proceedings of the 4th International Crop Science Congress', 26 Sept.–1 Oct. 2004, Brisbane, Australia. (Published on CD ROM. Web site: www.cropsscience.org.au)
- Campos H, Cooper M, Habben JE, Edmeades GO, Schussler JR (2004) Improving drought tolerance in maize: a view from industry. *Field Crops Research* **90**, 19–34. doi: 10.1016/j.fcr.2004.07.003
- Chapman SC, Cooper M, Hammer GL (2002a) Using crop simulation to generate genotype by environment interaction effects for sorghum in water-limited environments. *Australian Journal of Agricultural Research* **53**, 379–389. doi: 10.1071/AR01070
- Chapman SC, Cooper M, Hammer GL, Butler DG (2000) Genotype by environment interactions affecting grain sorghum. II. Frequencies of different seasonal patterns of drought stresses are related to location effects on hybrid yields. *Australian Journal of Agricultural Research* **51**, 209–221. doi: 10.1071/AR99021
- Chapman SC, Cooper M, Podlich D, Hammer GL (2003) Evaluating plant breeding strategies by simulating gene action and dryland environment effects. *Agronomy Journal* **95**, 99–113.
- Chapman SC, Hammer GL, Meinke HM (1993) A sunflower simulation model: I. Model development. *Agronomy Journal* **85**, 725–734.
- Chapman SC, Hammer GL, Podlich DW, Cooper M (2002b) Linking bio-physical and genetic models to integrate physiology, molecular biology and plant breeding. In 'Quantitative genetics, genomics, and plant breeding'. (Ed. M Kang) pp. 167–187. (CAB International: Wallingford, UK)
- Charles-Edwards DA (1982) 'Physiological determinants of crop growth.' (Academic Press: Sydney, NSW)
- Comstock RE, Moll RH (1963) Genotype–environment interactions. In 'Statistical genetics and plant breeding'. Publication 982. (Eds WD Hanson, HF Robinson) pp. 164–196. (National Academy of Sciences – National Research Council: Washington, DC)
- 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, DeLacy IH (1994) Relationships among analytical methods used to study genotypic variation and genotype-by-environment interaction in plant breeding multi-environment experiments. *Theoretical and Applied Genetics* **88**, 561–572. doi: 10.1007/BF01240919
- Cooper M, Hammer GL (2005) Complex traits and plant breeding: can we understand the complexities of gene-to-phenotype relationships and use such knowledge to enhance plant breeding outcomes? *Australian Journal of Agricultural Research* **56**, 869–872.
- Cooper M, Podlich DW, Smith OS (2005) Gene-to-phenotype models and complex trait genetics. *Australian Journal of Agricultural Research* **56**, 895–918.
- Craufurd PQ, Mahalakshmi V, Bidinger FR, Mukuru SZ, Chantreau J, *et al.* (1999) Adaptation of sorghum: characterisation of genotypic flowering responses to temperature and photoperiod. *Theoretical and Applied Genetics* **99**, 900–911. doi: 10.1007/s001220051311
- Dingkuhn M (1996) Modelling concepts for the phenotypic plasticity of dry matter and nitrogen partitioning in rice. *Agricultural Systems* **52**, 383–397. doi: 10.1016/0308-521X(95)00078-J
- Donatelli M, Hammer GL, Vanderlip RL (1992) Genotype and water limitation effects on phenology, growth, and transpiration efficiency in grain sorghum. *Crop Science* **32**, 781–786.
- Drouet J-L, Pages L (2003) GRAAL: a model of Growth, Architecture, and carbon Allocation during the vegetative phase of the whole maize plant—model description and parameterisation. *Ecological Modelling* **165**, 147–173. doi: 10.1016/S0304-3800(03)00072-3
- Edmeades GO, McMaster GS, White JW, Campos H (2004) Genomics and the physiologist: bridging the gap between genes and crop response. *Field Crops Research* **90**, 5–18. doi: 10.1016/j.fcr.2004.07.002
- 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.
- Ellis RH, Qi A, Craufurd PQ, Summerfield RJ, Roberts EH (1997) Effects of photoperiod, temperature and asynchrony between thermoperiod and photoperiod on development to panicle initiation in sorghum. *Annals of Botany* **79**, 169–178. doi: 10.1006/anbo.1996.0328
- Hammer GL (1998) Crop modelling: current status and opportunities to advance. *Acta Horticulturae* **456**, 27–36.
- Hammer GL, Butler D, Muchow RC, Meinke H (1996) Integrating physiological understanding and plant breeding via crop modelling and optimization. In 'Plant adaptation and crop improvement'. (Eds M Cooper, GL Hammer) pp. 419–441. (CAB International, ICRISAT & IRRI: Wallingford, UK)
- Hammer GL, Carberry PS, Muchow RC (1993) Modelling genotypic and environmental control of leaf area dynamics in grain sorghum. I. Whole plant level. *Field Crops Research* **33**, 293–310. doi: 10.1016/0378-4290(93)90087-4
- Hammer GL, Chapman SC, Snell P (1999) Crop simulation modelling to improve selection efficiency in plant breeding programs. In 'Proceedings 9th Assembly Wheat Breeding Society of Australia, Toowoomba, 27 Sept.–1 Oct. 1999'. (Eds P Williamson *et al.*) pp. 79–85. (Wheat Breeding Society of Australia: Toowoomba)
- Hammer GL, Farquhar GD, Broad IJ (1997) On the extent of genetic variation for transpiration efficiency in sorghum. *Australian Journal of Agricultural Research* **48**, 649–655. doi: 10.1071/A96111
- 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

- Hammer GL, van Oosterom EJ, Chapman SC, McLean G (2001) The economic theory of water and nitrogen dynamics and management in field crops. In 'Proceedings 4th Australian Sorghum Conference'. Kooralbyn, Qld, 5–8 Feb. 2001. (Eds AK Borrell, RG Henzell) (CD ROM Format. Range Media Pty Ltd. ISBN 0-7242-2163-8)
- Hammer GL, Sinclair TR, Chapman S, van Oosterom E (2004) On systems thinking, systems biology and the *in silico* plant. *Plant Physiology* **134**, 909–911. doi: 10.1104/pp.103.034827
- Hammer GL, Vanderlip RL, Gibson G, Wade LJ, Henzell RG, Younger DR, Warren J, Dale AB (1989) Genotype by environment interaction in grain sorghum II. Effects of temperature and photoperiod on ontogeny. *Crop Science* **29**, 376–384.
- Hart GE, Schertz KF, Peng Y, Syed NH (2001) Genetic mapping of *Sorghum bicolor* (L.) Moench QTLs that control variation in tillering and other morphological characters. *Theoretical and Applied Genetics* **103**, 1232–1242. doi: 10.1007/s001220100582
- Henderson SA, von Caemmerer S, Farquhar GD, Wade LJ, Hammer GL (1996) Correlation between carbon isotope discrimination and transpiration efficiency in lines of the C₄ species *Sorghum bicolor* in the glasshouse and the field. *Australian Journal of Plant Physiology* **25**, 111–123.
- Jeuffroy MH, Ney B, Ourry A (2002) Integrated physiological and agronomic modelling of N capture and use within the plant. *Journal of Experimental Botany* **53**, 809–823. doi: 10.1093/jexbot/53.370.809
- Keating BA, Carberry PS, Hammer GL, Probert ME, Robertson MJ, *et al.* (2003) An overview of APSIM, a model designed for farming systems simulation. *European Journal of Agronomy* **18**, 267–288. doi: 10.1016/S1161-0301(02)00108-9
- Li ZK, Luo LJ, Mei HW, Wang DL, Shu QY, *et al.* (2001) Overdominant epistatic loci are the primary genetic effects of inbreeding depression and heterosis in rice. I. Biomass and grain yield. *Genetics* **158**, 1737–1753.
- Ludlow MM, Muchow RC (1990) A critical evaluation of traits for improving crop yields in water-limited environments. *Advances in Agronomy* **47**, 107–153.
- Ludlow MM, Santamaria JM, Fukai S (1990) Contribution of osmotic adjustment to grain yield in *Sorghum bicolor* (L.) Moench under water-limited conditions. II. Water stress after anthesis. *Australian Journal of Agricultural Research* **41**, 67–78. doi: 10.1071/AR9900067
- Luo LJ, Li ZK, Mei HW, Shu QY, Tabien R, Zhong DB, Ying CS, Stansel JW, Khush GS, Paterson AH (2001) Overdominant epistatic loci are the primary genetic effects of inbreeding depression and heterosis in rice. II. Grain yield components. *Genetics* **158**, 1755–1771.
- Major DJ, Rood SB, Miller FR (1990) Temperature and photoperiod effects mediated by the sorghum maturity genes. *Crop Science* **30**, 305–310.
- Mifflin B (2000) Crop improvement in the 21st Century. *Journal of Experimental Botany* **51**, 1–8. doi: 10.1093/jexbot/51.342.1
- Minorsky PV (2003) Achieving the *in silico* plant: systems biology and the future of plant biological research. *Plant Physiology* **132**, 404–409. doi: 10.1104/pp.900076
- Monteith JL (1986) How do crops manipulate water-supply and demand? *Transactions of the Royal Society of London A* **316**, 245–259.
- Monteith JL (1988) Does transpiration limit the growth of vegetation or vice-versa? *Journal of Hydrology* **100**, 57–68. doi: 10.1016/0022-1694(88)90181-3
- Morgan PW, Finlayson SA, Childs KL, Mullet JE, Rooney WL (2002) Opportunities to improve adaptability and yield in grasses: lessons from sorghum. *Crop Science* **42**, 1791–1799.
- Morgan PW, Guy LW, Pao CI (1987) Genetic regulation of development in *Sorghum bicolor*. III. Asynchrony of thermoperiods with photoperiods promotes floral initiation. *Plant Physiology* **83**, 448–450.
- Mortlock MY, Hammer GL (2000) Genotype and water limitation effects on transpiration efficiency in sorghum. *Journal of Crop Production* **2**, 265–286. doi: 10.1300/J144v02n02_11
- Muchow RC, Carberry PS (1990) Phenology and leaf area development in a tropical grain sorghum. *Field Crops Research* **23**, 221–237. doi: 10.1016/0378-4290(90)90056-H
- van Oosterom EJ, Hammer GL, Borrell AK, Chapman SC, Broad IJ (2005a) Functional dynamics of the nitrogen balance of sorghum. I. N-balance during pre-anthesis period. *Field Crops Research* (In press).
- van Oosterom EJ, Hammer GL, Chapman SC, Borrell AK, Broad IJ (2005b) Functional dynamics of the nitrogen balance of sorghum. II. N-balance during grain filling. *Field Crops Research* (In press).
- Passioura JB (1983) Roots and drought resistance. *Agricultural Water Management* **7**, 265–280. doi: 10.1016/0378-3774(83)90089-6
- Peccoud J, Vander Velden K, Podlich D, Winkler C, Arthur L, Cooper M (2004) The selective values of alleles in a molecular network model are context dependent. *Genetics* **166**, 1715–1725. doi: 10.1534/genetics.166.4.1715
- Podlich D, Cooper M (1998) QU-GENE: a simulation platform for quantitative analysis of genetic models. *Bioinformatics* **14**, 632–653. doi: 10.1093/bioinformatics/14.7.632
- Podlich DW, Winkler CR, Cooper M (2004) Mapping as you go: an effective approach for marker-assisted selection of complex traits. *Crop Science* **44**, 1560–1571.
- 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
- Serraj R, Sinclair TR (2002) Osmolyte accumulation: can it really help increase crop yield under drought conditions? *Plant, Cell and Environment* **25**, 333–341. doi: 10.1046/j.1365-3040.2002.00754.x
- Sinclair TR, Horie T (1989) Leaf nitrogen, photosynthesis, and crop radiation use efficiency: a review. *Crop Science* **29**, 90–98.
- Sinclair TR, Seligman NG (1996) Crop modelling: from infancy to maturity. *Agronomy Journal* **88**, 698–704.
- Snape J (2004) Challenges of integrating conventional breeding and biotechnology: a personal view! In 'New directions for a diverse planet. Proceedings of the 4th International Crop Science Congress, 26 Sept.–1 Oct. 2004, Brisbane, Australia'. (Published on CD ROM. Web site: www.cropscience.org.au)
- Snell P (2004) The contribution of osmotic adjustment to grain yield of sorghum in dryland production environments. PhD thesis, The University of Queensland, Brisbane, Australia.
- Somerville C, Somerville S (1999) Plant functional genomics. *Science* **285**, 380–383. doi: 10.1126/science.285.5426.380
- Tanner CB, Sinclair TR (1983) Efficient water use in crop production: research or re-search?. In 'Limitations to efficient water use in crop production'. (Eds HM Taylor, WR Jordan, TR Sinclair) pp. 1–27. (American Society of Agronomy: Madison, WI)
- Tao YZ, Henzell RG, Jordan DR, Butler DG, Kelly AM, McIntyre CL (2000) Identification of genomic regions associated with staygreen in sorghum by testing RILs in multiple environments. *Theoretical and Applied Genetics* **100**, 1225–1232. doi: 10.1007/s001220051428
- 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.
- Thomas H, Howarth CJ (2000) Five ways to stay green. *Journal of Experimental Botany* **51**, 329–337. doi: 10.1093/jexbot/51.suppl_1.329
- Wang E, Robertson MJ, Hammer GL, Carberry PS, Holzworth D, Meinke H, Chapman SC, Hargreaves JNG, Huth NI, McLean G (2002) Development of a generic crop model template in the cropping system model APSIM. *European Journal of Agronomy* **18**, 121–140. doi: 10.1016/S1161-0301(02)00100-4
- 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.
- Welch SM, Roe JL, Dong Z (2003) A genetic neural network model for flowering time control in *Arabidopsis thaliana*. *Agronomy Journal* **95**, 71–81.
- White JW, Hoogenboom G (1996) Simulating effects of genes for physiological traits in a process-oriented crop model. *Agronomy Journal* **88**, 416–422.
- de Wit CT, Penning de Vries FWT (1983) Crop growth models without hormones. *Netherlands Journal of Agricultural Science* **31**, 313–323.
- Yin X, Stam P, Kropff MJ, Schapendonk Ad HCM (2003) Crop modelling, QTL mapping, and their complimentary role in plant breeding. *Agronomy Journal* **95**, 90–98.
- 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