CSIRO Publishing

Australian Journal of Agricultural Research

VOLUME 53, 2002

A journal for the publication of original contributions towards the understanding of an agricultural system

All enquiries and manuscripts should be directed to:

Australian Journal of Agricultural Research
CSIRO Publishing
PO Box 1139 (150 Oxford St)
Collingwood, Vic. 3066, Australia

Telephone: +61 3 9662 7628
Fax: +61 3 9662 7611
Email: publishing.ajar@csiro.au

Published by CSIRO Publishing
for CSIRO and the Australian Academy of Science

www.publish.csiro.au/journals/ajar

Aust. J. Agric. Res., 2002, **53**, 379–389

Using crop simulation to generate genotype by environment interaction effects for sorghum in water-limited environments

Scott C. Chapman[A], *Mark Cooper*[B], *and Graeme L. Hammer*[C]

[A]Corresponding author; CSIRO Plant Industry, Long Pocket Laboratories, 120 Meiers Rd, Indooroopilly, Qld 4068, Australia; email: scott.chapman@csiro.au

[B]School of Land and Food Sciences, The University of Queensland, Brisbane, Qld 4072, Australia; present address: Pioneer Hi-Bred, Johnston, Iowa, USA.

[C]Agricultural Production Systems Research Unit, Queensland Department of Primary Industries, PO Box 102, Toowoomba, Qld 4350, Australia.

Abstract. Multi-environment trials (METs) used to evaluate breeding lines vary in the number of years that they sample. We used a cropping systems model to simulate the target population of environments (TPE) for 6 locations over 108 years for 54 'near-isolines' of sorghum in north-eastern Australia. For a single reference genotype, each of 547 trials was clustered into 1 of 3 'drought environment types' (DETs) based on a seasonal water stress index. Within sequential METs of 2 years duration, the frequencies of these drought patterns often differed substantially from those derived for the entire TPE. This was reflected in variation in the mean yield of the reference genotype. For the TPE and for 2-year METs, restricted maximum likelihood methods were used to estimate components of genotypic and genotype by environment variance. These also varied substantially, although not in direct correlation with frequency of occurrence of different DETs over a 2-year period. Combined analysis over different numbers of seasons demonstrated the expected improvement in the correlation between MET estimates of genotype performance and the overall genotype averages as the number of seasons in the MET was increased.

Additional keywords: modelling, cultivar evaluation, genotypic variance, *Sorghum bicolour*.

Introduction

In crop breeding programs, testing of lines or crosses takes place over 1 or more years and locations. During the early stages of the breeding process, large numbers of potential cultivars (generating high genotypic variance) may be tested in only 1 year and in 1 or 2 locations. Characteristically, the number of locations and years of testing increases as the process of selection reduces the number of lines under test and the potential cultivars move closer to commercial release. The number of locations evaluated in these multi-environment trials (METs) may range from fewer than 10 in small markets such as summer crops in Australia to hundreds in large markets such as maize in the USA. Although new hybrid cultivars may have been adequately tested for tolerance to biotic stresses (through artificial screening), a general consequence of the short time frame for field testing is that there is inadequate exposure of the hybrids to the range of climatic conditions that a released cultivar will encounter. In the north-eastern grain cropping area of Australia, these aspects of testing for environmental adaptation are of particular concern due to substantial inter-annual variation in summer rainfall, which is greatly influenced in 1–30-year cycles by the interactions between the southern oceans and the atmosphere (El Niño Southern Oscillation (ENSO) effects) (Nicholls 1988).

The abiotic and biotic stresses over many years and locations in the geographic mandate area of a plant breeding program comprise a complex 'target population of environments' (TPE) (Comstock 1977). The distribution of the stresses experienced in the sampling of the TPE complicates breeding programs by creating a substantial genotype by environment interaction (G × E) in any series of cultivar performance tests. This has been demonstrated for sorghum (*Sorghum bicolour*) in Australia (Chapman *et al.* 2000*a*), among other crops. The unpredictability of the climate and its generally poor correlation with locations imply that it is difficult to choose a cultivar for a location based on the future expectations of weather. Hence, released cultivars must usually be selected to have a degree of 'broad adaptation' to the environments.

 10.1071/AR01070 0004-9409/02/040379

Chapman *et al.* (2000*b*) used a sorghum simulation model to quantify the temporal and spatial variation in the occurrence of different drought environment types (DETs) as determined by seasonal patterns of water stress. The frequency of occurrence of different DETs at different locations was partially associated with measures of the effects of locations on the performance of hybrids in evaluation trials. Although that study showed that there may be value in producing genotypes with slightly different specific adaptations for particular locations/soil types in the longer term, it would be difficult to identify these genotypes given the time period that would have to be sampled by METs in Australian conditions.

There are many options for plant breeders in the establishment of the germplasm pool and methods of evaluating and selecting improved genotypes. To deal with the complexity of directing the accumulation of favourable gene networks into improved hybrids, researchers are using simulation tools to complement experimentation. Cooper *et al.* (1999) have reviewed their application in plant breeding programs. Podlich and Cooper (1998) developed QU-GENE for this type of application and have demonstrated the importance of sampling the range of environments in the TPE, and the value of weighting data from METs (as samples of the TPE) in accordance with their expected frequency in the TPE. Podlich *et al.* (1999) reported an advantage in breeding progress of the weighted method of up to 20%, particularly when the size of the MET was small and the cross-over component of the genotype by environment interaction was substantial.

Obviously, plant breeders cannot ever evaluate germplasm in the full range of environments that may occur in the target geographical mandate. QU-GENE is a quantitative genetics simulation platform that enables genotype performance to be quantified in terms of the statistical interaction of gene effects. An additional simulation technology (dynamic biophysical crop simulation) allows us to define values for trait parameters and estimate their effects on yield as mediated by the growth of the crop in response to the soil and climate conditions encountered. This requires a reliable crop model that incorporates the direct and indirect effects of target traits. Using a cropping system model (APSIM; McCown *et al.* 1996) with a sorghum module developed for this purpose, Hammer *et al.* (1996) simulated the yields of a set of 24 sorghum near-isolines in 3 locations over 10 years. The statistical attributes of these data, including the actual and relative sizes of the genotypic and G × E interaction variance components generated for yield, were comparable with that observed in actual trials.

In extending the approach of Hammer *et al.* (1996), we have here used the cropping system model to evaluate the expected genotype performance in a much larger sample of the TPE than is ever possible experimentally, by sampling the entire period of reliable weather record (>100 years) for representative locations in the Australian sorghum cropping region. Using subsets of these data, we can then interpret the effectiveness of the sampling of the TPE over the small numbers of consecutive years available to plant breeders in conducting METs. We appreciate that simulation requires some major simplifications about crop responses to the environment, but models do allow us to interpret the long-term and spatial effects of environment, where experimentation is simply not possible. The main objective of this paper is to outline the basic dataset and interpret the variation in environment and genetic interaction with environment for several physiological parameters affecting dryland adaptation of sorghum. A simple correlation analysis determines the relationship between long-term yield estimates and shorter term samples (1–6 years) that are available to practical plant breeders. In other papers (e.g. Cooper *et al.* 2001; Chapman *et al.* 2002), we are extending the methodology to simulate entire 'adaptation landscapes' of gene–environment effects and 'searching' these populations using a simulation model of the plant breeding process, QU-GENE (Podlich and Cooper 1998).

Materials and methods

This simulation experiment and its analysis required several steps that are detailed in following sections and are summarised here. Firstly, a standard 'reference' genotype (SSSS, see text below) was simulated in an opportunity cropping system for 108 years at 6 locations (648 potential 'trials') across the sorghum production region of Australia. After elimination of 'failed' trials, the stress encountered by the reference genotype was used to define the pattern of drought stress for each trial. Data from each trial were then analysed in 2 ways: (*1*) in combinations of locations within a year (a single year MET: MET1); and (*2*) in combinations of locations across several years (i.e. in METs of 2, 3, 4, and 6 years). Analyses applied to the trial data included computation of the variance components for genotypic and genotype by environment interaction effects with simulated experimental error variances.

Cropping system simulation

All simulations were done with version 1.5 of the 'Sorg' crop module developed within the APSIM cropping systems simulation model (McCown *et al.* 1996, website www.apsru.gov.au). APSIM, developed by a team of about 30 scientists, uses weather data on a daily time step to interact with a specified soil profile and simulate the soil and plant processes associated with water and nitrogen during fallow and in-crop states. All parameter initialisation and management files for APSIM are external to the model and can be easily updated as new information becomes available. The sorghum module has undergone extensive development using data from more than 40 experiments to enhance its capacity to realistically simulate the interactions among physiological processes that occur when key parameters have been varied (Hammer *et al.* 2001). The model has been particularly designed to account for the genetic variation in 4 traits observed in experiments in the region: flowering time (PH), transpiration efficiency (TE), osmotic adjustment (OA), and stay-green (SG). As part of the simulation of maximum leaf area (Hammer *et al.* 1993), the model requires the potential number of tillers (dependent on location and time of year) to be input. All other parameters are 'generic'. Published descriptions of the model are given

Table 1. Characteristics of the planting windows and the soil types at each location

Region	Planting window	Tillers per main culm	Locations (lat./long.)	Soil type (depth)	Water holding capacity (mm)	No. of seasons planted
Central Queensland (CQ)	1 Nov., 31 Jan.	0.25	Biloela (24.4S, 150.52E)	Black Vertisol (0.8 m)	Medium (170)	101
			Emerald (23.57S, 148.18E)	Black Vertisol (0.8 m)	Low (120)	196
Southern Queensland (SQ)	1 Oct., 15 Nov.	0.5	Dalby (27.17S, 151.27E)	Black Vertisol (1.5 m)	High (250)	102
	16 Nov., 15 Jan.	0.25	Miles (26.67S, 150.18E)	Grey Vertisol (0.8 m)	Medium (170)	95
Northern NSW (NNSW)	1 Oct., 15 Nov.	0.5	Gunnedah (30.98S, 150.25E)	Black Vertisol (1.5 m)	High (250)	88
	16 Nov., 15 Dec.	0.25	Moree (29.47S, 149.85E)	Grey Vertisol (0.8 m)	Low (170)	83
	16 Dec., 1 Jan.	0.0				

Table 2. Values used for different 'expression states' (levels) of the four physiological traits in the crop simulation

Trait name	Model variable description and units	Value in allele combination		
		Lower (L)	Standard (S)	Upper (U)
Transpiration efficiency (TE)	Transpiration efficiency coefficient — the efficiency with which transpired water is used during the assimilation of dry matter (MPa)	0.008	0.009	0.010
Phenology (PH)	Thermal time from the end of juvenile stage to panicle initiation (degree-days)	90	115	140
Osmotic adjustment (OA)	Growth amount required per grain set (g/grain) and fraction of stem biomass available for retranslocation (%)	n.a.	0.00083 20%	0.00075 36%
Stay-green (SG)	Specific leaf nitrogen (SLN), the target SLN of new leaf (g N/m^2 leaf)	1.35	1.5	1.65

n.a., Not applicable.

by Hammer and Muchow (1994) and Hammer *et al.* (1996, 2001), although recent improvements have adopted concepts of 'emergent' properties (Hammer 1998) in seeking more realistic simulation of genetic variation in traits via underlying physiological functionality.

Weather records were obtained for 6 locations across the sorghum growing region: 2 stations in each of the 3 major areas (Table 1). Using representative soil types for each location, we simulated a continuous summer sorghum–winter fallow cropping system with planting only taking place if criteria for sowing were met within a planting 'window'. For each region (Table 1), the planting window and potential tiller numbers were defined and the crop was planted at 50 000 plants/ha. In all cases, the sowing rule for planting within the window was the same: the soil had to have accumulated 80 mm of water, and the field needed to have received 25 mm of rain within a 4-day period. Initially, the continuous system simulation was done using the SSSS (standard) genotype to define the soil moisture conditions at planting and the planting dates for each multi-environment trial. The growth and yield of all other genotypes were then simulated using these starting conditions for each season.

Genotype simulation

The public sorghum breeding program in Australia, based in the Queensland Department of Primary Industries (QDPI), has focussed on the development of midge resistance and on the stay-green trait to improve yields under drought (Henzell 1992). Staff at QDPI, The University of Queensland, and CSIRO Plant Industry (formerly Tropical Agriculture) have investigated these, and other traits, with the aim of designing breeding strategies to improve adaptation of sorghum. The research has established the approximate range of genetic variation for the 4 physiological traits mentioned above (Table 2). In this paper we aimed to represent trials of a factorial combination of the extreme (U, upper; L, lower) and 'standard' (S) gene expression states for these traits. Whilst these exact genotypes may not (yet) exist in the breeding population, the trial attempts to simulate the potentially available range (see below).

The 54 genotypes used in this paper have been designated in alphabetical order of the values for the 4 traits, e.g. LLSL is genotype number 1 (and has the 'lower' values for the traits TE, PH, and SG, and the 'standard' value for OA) and UUUU is genotype 54 (and has the 'upper' values for each trait). These gene definitions have been simplified from the set used in the simulation of the whole population of genotypes (>4000) that incorporates more complex genetic architecture (Chapman *et al.* 2002). To further simplify discussion here, we assume that the S and U effects are simple additive gene effects (2 and 3 genes, respectively) beyond the base level established by the expression state at level L (1 gene).

Starting parameters in the sorghum model were set to reflect known genetic variation in the underlying physiological traits. Genetic variation in transpiration efficiency (TE) was simulated by either decreasing (L) or increasing (U) the standard transpiration efficiency coefficient for sorghum of 0.009 MPa by about 10% (Table 2), a range that has been observed in studies over sorghum genotypes in well-watered and water-limited situations (Hammer *et al.* 1997; Mortlock and Hammer 1999). The TE coefficient represents the product of the atmospheric vapour pressure deficit (vpd) and the observed transpiration efficiency of the plant (g dry matter/g transpired water). Hence, resulting transpiration efficiency computed on any day will be influenced by both environment (vpd) and genotype (TE coefficient level).

Developmental rate in sorghum can be predicted via known responses to temperature and photoperiod (Hammer *et al.* 1989). To reflect genetic variation in phenology (PH) the thermal duration of the developmental phase prior to floral initiation was either decreased (L) or increased (U) by 25 degree-days (Table 2). Differences in duration prior to floral initiation will generate differences in number of leaves

produced (Hammer and Muchow 1994) with consequent effects on canopy leaf area development (Carberry *et al.* 1993; Hammer *et al.* 1993) and, thus, patterns of water use through the crop cycle.

For osmotic adjustment (OA), genetic variation was simulated by increasing potential grain number and the ability to retranslocate stem biomass to grain under moisture limitation. Although it is not directly apparent from the trait description, this effect arises from a physiological analysis of mode of action of OA in high and low OA sorghum lines (P. Snell, PhD Thesis, unpubl. data; under review) as summarised by Hammer *et al.* (1999). They found that, at the crop level, high OA lines (selected on OA based on screening for OA in stressed plants) had a greater ability to set grain and re-translocate carbohydrate from the stem to grain. The genetic increase in grain number (resulting from the jump from S to U) was simulated by reducing the amount of crop biomass growth required between floral initiation and flowering to produce an individual grain (Table 2). Enhanced remobilisation of assimilate from stem during grain filling was simulated by increasing the fraction of stem biomass at flowering that was potentially available for re-translocation. 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. Only standard (S) and enhanced (U) levels of OA were simulated as we have no evidence to suggest the existence of more detrimental OA effects.

Genetic variation in the 'staygreen' trait (SG) was simulated by modifying the target specific leaf nitrogen (SLN g N/m^2 leaf area) of new leaf (Table 2). This mechanism was reported by Borrell *et al.* (2000) from physiological studies on mode of action of SG in hybrids from a cross of parents with high and low levels of SG. Increasing the target SLN (U) allows increased N uptake during canopy development. Subsequently, during grain-filling, depletion of N from leaves is delayed, resulting in stay-green. Whereas this mechanism appears to explain at least one form of SG (Borrell *et al.* 2000), other mechanisms are also likely to exist.

Classification of patterns of drought environment types (*DETs*)

The definition and classification of DETs for the SSSS genotype have been described elsewhere (Chapman *et al.* 2000*b*). Briefly, the crop simulation model produces a daily 'stress index', which is termed relative transpiration (RT), and is the ratio of water supply to water demand. The RT was averaged over each 100 degree-days ['thermal time weeks' (TTweeks)] of the season to produce 11 values over the season. Thus, a matrix of 11 columns (TTweeks) by 565 rows (trials) was produced for pattern analysis.

Clustering was used to group the trials from all locations and seasons into 3 DETs according to how similar the RT pattern was across the season (Muchow *et al.* 1996). Firstly, a proximity matrix was created by calculating squared Euclidean distances among all of the trials. A hierarchical agglomerative cluster analysis (Ward's method) was then used to group the locations. Analyses of variance determined how much of the variation was explained at each level of grouping. At the 3 group level, the seasonal RT values were averaged over members within each group to give a sequence of RT against TTWeek for the DET group. Frequency tables were constructed to analyse the occurrence of these DET groups at different locations. All analyses were undertaken using a library of S-Plus (MathSoft Inc., Seattle) functions developed for this purpose (K. Basford, School of Land and Food, The University of Queensland; software available from http://www.uq.edu.au/~agkbasfo/index.html).

Analysis of simulated yields and effect of MET sampling

Simulated yields were obtained for each combination of genotype, location, and year in which the planting criteria were met. In some cases, genotypes failed (due to drought) and a zero yield was simulated. To reduce the degree of imbalance in the data, any trials where more than one-third of the genotypes failed were deemed 'trial failures' and were deleted from the analysis. In other cases, genotypes with zero yields retained those values.

The data were coded by genotype, trial (environment), location, and year. Additional factors defined the METs that a plant breeding program may utilise to test a set of hybrids for release. Beginning with year 1891, the METs were defined as being of $k = 1$ (MET1), 2 (MET2), 3 (MET3), 4 (MET4), or 6 (MET6) years in length, to result in 108, 54, 36, 27, and 18 sequential METs for the entire 108-year record. Genotype means were averaged over trials within different METs of different lengths. Correlations [$n = 54$, signif. $r = 0.27$ ($P < 0.05$) or 0.35 ($P < 0.01$)] were computed between these means and the genotype means averaged across all of the trials.

The analyses described above were repeated after re-processing the data to include experimental error (micro-environment) effects. In previous studies of actual data (Chapman *et al.* 2000*a*), we have determined that the coefficient of variance (CV) for grain yield in Queensland and NSW was approximately 12% of the trial mean yield for trials (i) of 3 replications. This CV was used to compute an average error variance ($\bar{s}^2$) for the entire dataset of 0.194 t^2/ha with the estimated value for individual trials ($\hat{s}_i^2$) ranging from 0.006 to 0.835 t^2/ha, reflecting the range of mean yields obtained. To take into account the variable effect of the error variances for each simulated trial, a yield weight (w_i) was calculated as used by Cullis *et al.* (1996):

$$w_i = r_i \bullet \bar{s}^2 / \hat{s}_i^2$$

where r_i is the number of replicates in trial i.

In about 20 of the trials, the simulated yields were such that the trial essentially 'failed' (<0.2 t/ha), resulting in extremely low error variances and high weights. For this reason, the w_i values were restricted to a maximum of 70, based on analysis of Chapman *et al.* (2000*a*), where there was a maximum w_i of 62.5 from 168 sorghum hybrid trials over 17 years of testing.

To obtain estimates of genotypic (σ^2_g) and genotype by environment (σ^2_{ge}) variance including experimental error effects, the simulated yield data were analysed using a residual maximum likelihood method (Patterson and Thompson 1975) implemented by the ASREML software (Gilmour *et al.* 1998). Environments (e_j) were considered fixed (as they encompassed the 'entire' TPE) and genotypes (g_i) and genotype by environment interaction effects were considered random on the basis that the genotypes should represent the 'potential' range of genetic variation for the traits of interest. The method was used to estimate σ^2_g and σ^2_{ge} over all locations and years, as well as best linear unbiased predictors (BLUPs) for the genotypes and a predicted value for each genotype by trial combination. Additional variance component estimates ($\sigma^2_{g(k)}$ and $\sigma^2_{ge(k)}$) were computed for different METs (k) within a series of METs, i.e. within MET1, MET2, etc. To evaluate the effects of the number of years of sampling of the TPE, the correlations applied to the raw data were repeated using the predicted values for genotype by trial combinations, i.e. the predicted values averaged for all trials were correlated against those for individual METs within any of the 5 sequences of METs.

Results

Sampling of environment types

From 648 attempted 'trials' or environments over locations and years, 565 trials met planting criteria and 30510 genotype/trial combinations were simulated. Another 18 trials were deemed to have 'failed' (see text above), leaving

Table 3. Number of planted seasons from 108 possible at each location (648 in total) classified into each of three season types characterised in Fig. 1

Numbers in parentheses are the percentage of planted seasons at a location (or in all locations) that were grouped into the same season type. For right-hand column, numbers in parentheses are percentage of total seasons

Location	Group			Number of trials planted
	Mild terminal	Severe terminal	Mid-season	
Biloela	46 (46%)	27 (27%)	27 (27%)	100 (93%)
Emerald	29 (31%)	33 (35%)	33 (35%)	95 (88%)
Dalby	58 (58%)	18 (18%)	24 (24%)	100 (93%)
Miles	36 (39%)	36 (39%)	21 (23%)	93 (86%)
Gunnedah	28 (34%)	34 (41%)	21 (25%)	83 (77%)
Moree	6 (8%)	42 (55%)	28 (37%)	76 (70%)
Total	203 (37%)	190 (35%)	154 (28%)	547 (84%)

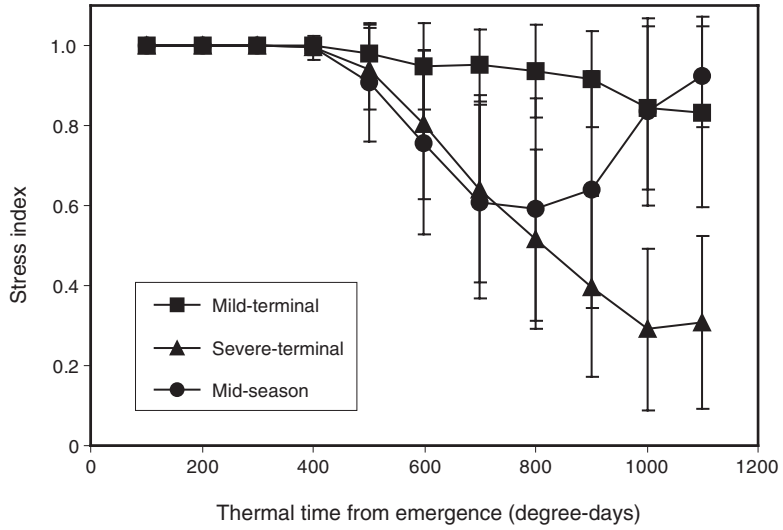

Fig. 1. For 3 groups (drought environment types), the mean stress index (ratio of water supply to demand) across the season for the reference (SSSS) genotype. Lines show the mean values of the stress index (with bars representing standard deviations) for all members of each environment type over 11 thermal time periods during the season. Mean anthesis date was 732 degree-days.

a total of 547 trials (84% of the potential number of locations and years) (Table 3). In the two Queensland regions, fewer intended trials were successful on the shallower soil within each region (Emerald and Miles). In northern NSW, fewer than 80% of trials were a 'success', with a greater number of failures in the shallower soil (Moree).

Fig. 1 shows the stress patterns that were determined for 3 DET groups clustered from all locations and seasons for the entire dataset, based on the reference genotype (SSSS). In all cases, stress as measured by the relative supply and demand for water by the crop was minimal (*c*. 1.0) until about 400 degree-days after planting. Beyond that point, where leaf area has been well established, the ability of the crop to maintain a balance between supply and demand for water decreased within all 3 groups of stress types. The mean thermal time to anthesis for this genotype was 732 degree-days, with a standard deviation of 18 degree-days. In one DET (termed 'mild terminal' stress), the mean stress index remained higher than 0.8 throughout the period of grain filling (>732 degree-days), and resulted in a mean yield for genotype SSSS of 5.15 t/ha. A pattern of decreasing stress index (i.e. continuously increasing stress) was experienced in another DET ('severe terminal' stress) where mean yield was 2.35 t/ha. With a mean yield of 3.54 t/ha, the pattern of stress observed in the 'mid-season' stress DET was similar to the severe terminal DET until 700 degree-days when the stress began to be relieved through mid to late grain-fill.

Across all locations and years (i.e. for the TPE), the frequencies of the 2 terminal DETs were similar and slightly greater than that for the mid-season DET (Table 3). Over the whole record, the frequencies at Emerald, Gunnedah, and

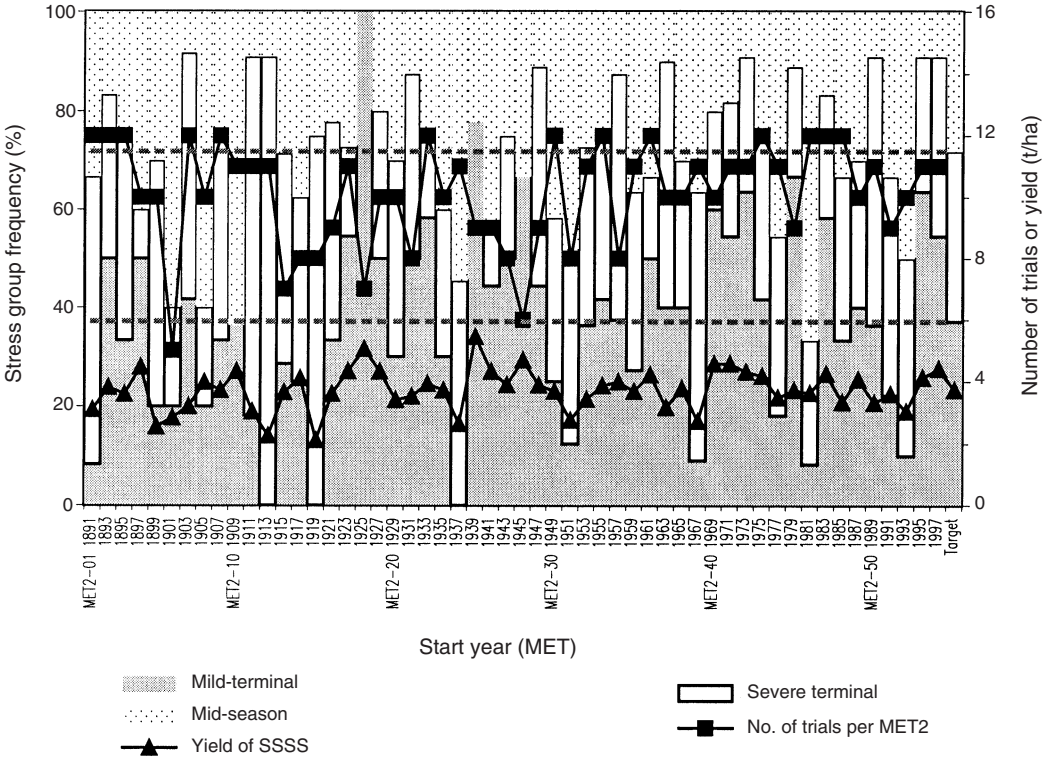

Fig. 2. Frequencies of the different drought environment types (Fig. 1) across 6 locations within a series of 54 two-year multi-environment trials (METs) and across all seasons (target, same as the last row of Table 3). The lines (squares) indicate the number of 'successful' trials in each MET (out of a possible 2 years and 6 locations) and the simulated grain yield (triangles) of the reference (SSSS) genotype averaged within each MET and for all trials (target).

Miles were closest to the frequencies over all locations, whereas the mild terminal DET was over-represented at Biloela and Dalby and under-represented at Moree.

The dashed lines and extreme right-most column in Fig. 2 indicate the average frequencies of occurrence for each of the 3 DETs (i.e. data from the bottom line of Table 3) in the entire TPE. For MET2s (sequential METs of 2 years in length), an average of 10.1 of 12 possible trials (2 years by 6 locations) were 'planted' (Fig. 2), with the fewest (5) being planted in the MET2 that had begun in 1901. In an ideal MET that represented the TPE, all 3 DETs would be sampled in almost equal proportions: DET1 37%; DET2 35%; and DET3 28% (Table 3). In most cases, all 3 DETs (Fig. 1) were sampled at least once within each MET, although the highest proportions recorded for each DET (mild terminal, severe terminal, and mid-season stress) were 100% (in 1925), 91% (1913), and 64% (1909), respectively. In terms of the relative frequencies of the 3 DETs, the MET2 that began in 1953 was the most similar MET to the overall target frequency (Fig. 2). At other extremes, in several MET2s, there were no occurrences of the mild terminal DET (1913, 1919, 1937), and in 1925, 1939, and 1945, no severe terminal DET was sampled. Clearly, the average yield of genotype SSSS for an entire MET2 was depressed when the MET sampled only a small frequency of the mild terminal DET (e.g. 1913, 1919, 1937, 1951, 1967, 1993). The converse applied in trials in 1925 and 1939. The degree to which average yield changes with the sampling by METs of TPE is important in absolute terms, but it is the relative change in genotype performance that most greatly influences plant breeding, i.e. choosing the apparently superior genotype, when its performance is a function of the sampling of the TPE, rather than a real advantage in the entire TPE.

Genotype performance over all years and within METs

The genotype overall mean yields are given in Fig. 3*a*, with the overall average being 3.79 t/ha, and ranging from 2.81 t/ha for genotype LLSL to 4.92 t/ha for genotype UUUU. When ordered by mean yield, the general increase in yield due to gene effects of the different traits is evident in the association between the genotype means and the increasing heights of the columns. However, there are several points where a large change in the mean yield is observed with the same total number of gene expression states, e.g. between the 7th (SLUL) and 8th (LLSU) genotypes in Fig. 3*a*. In this case, compared with the 7th genotype, the 8th genotype had a lower value for the TE (1st letter) and OA (3rd letter) traits and higher values for the stay-green trait

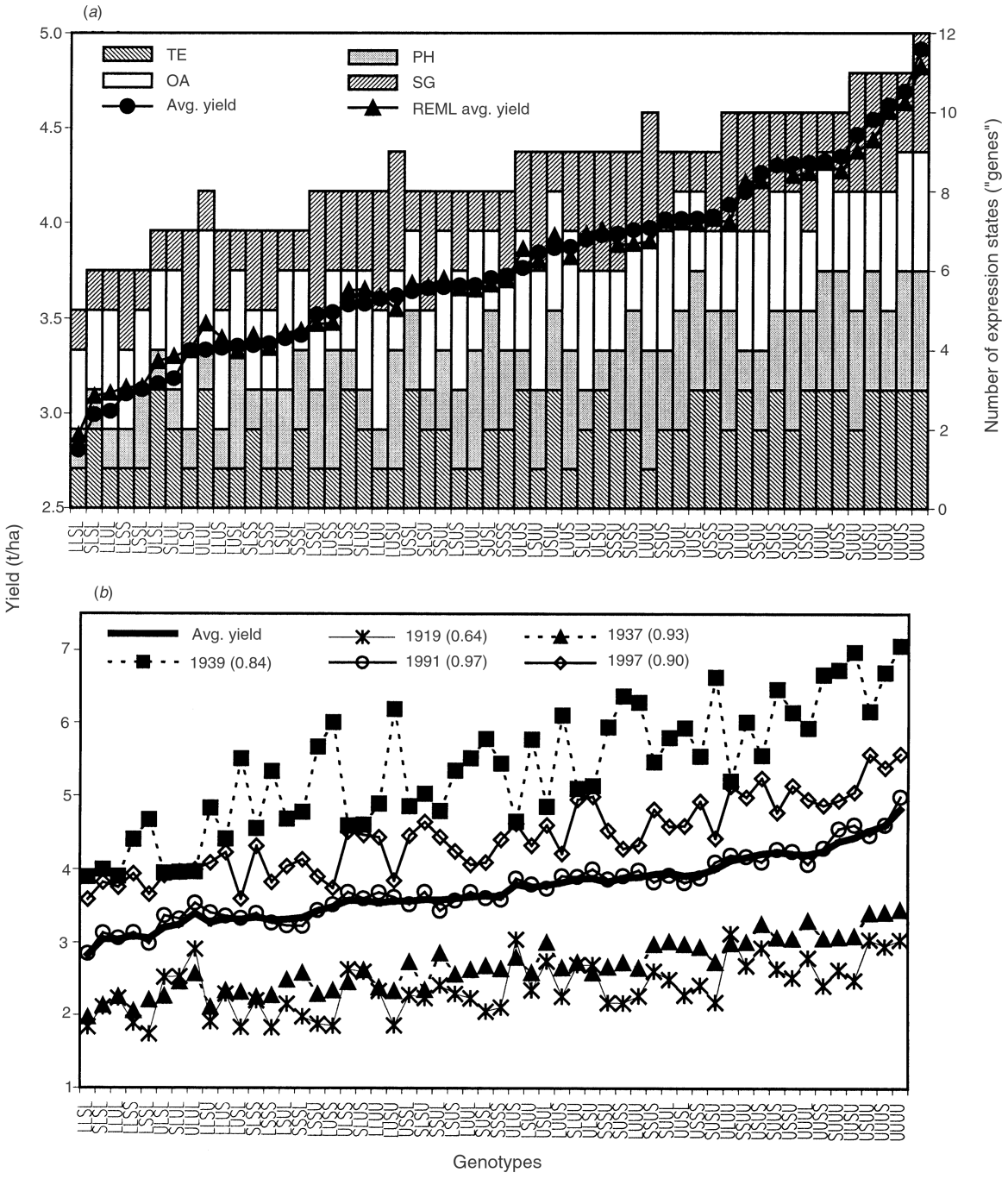

Fig. 3. (*a*) Fifty-four genotypes ranked by their average yields from 547 trials, together with average yields determined after REML analysis to introduce experimental error; and a representation of the 'expression state' of genes (L = 1, S = 2, U = 3) for each of 4 traits (Table 2) making up a genotype. Note that the trait OA has only 2 levels (2 and 3). (*b*) Using the same genotype ranking as in Fig. 3*a*, average genotype yields within 5 different 2-year METs chosen from the 54 METs in Fig. 2.

(4th letter), but had the same early maturity (L value for PH trait). Other large mean yield effects for the same total number of gene effects are evident in the genotype sequences beginning with SUSU and SUUU.

The variation in each of the traits produced similar effects on yield. Genotypes with L expression levels for TE, PH, and SG yielded an average of 0.34, 0.38, and 0.33 t/ha less than the average of the genotypes for the S levels of these traits. The yields for the genotypes with U levels of TE, PH, OA, and SG were 0.29, 0.26, 0.28, and 0.23 t/ha higher than for their respective S level equivalents.

Although the overall mean yield from the REML analysis was the same as for the raw data (as would be expected), the genotypes underwent re-ranking, especially within groups of similar numbers of gene effects. The raw and REML means were highly correlated over the 54 genotypes (r >0.99), but the re-ranking effect can be seen in the non-alignment of lines for the raw means and the means of the REML predicted data (Fig. 3*a*). Inspection of the data showed that for traits TE and OA there was no difference (<0.01 t/ha) between raw and REML means. However, for PH and SG, the average REML yields for genotypes with the L expression level were 0.065 and 0.040 t/ha higher, respectively, than for the S genotypes. Thus, the re-ranking resulted from an increased estimate of performance of the L level genotypes for PH and SG traits, when experimental error was introduced.

Fig. 3*b* shows the genotype mean yields for several MET2s. The first example is for MET2-1919 in which no mild terminal DET was sampled (see Fig. 2). In this MET, several genotypes that were ranked low over all trials performed quite well, resulting in a relatively poor correlation ($r = 0.63$, $P < 0.01$) between the genotype means for the MET and the overall means.

Two more examples relate to sequences of MET2s. In MET2-1937 the mild terminal DET was again not sampled (Fig. 2), whereas in the following MET2 (1939), the mild terminal DET was sampled 78% of the time. The yields in MET2-1937 were similar ($r = 0.75$, $P < 0.01$) to those obtained in the MET2-1919, which had a similar bias in sampling against the mild terminal DET. In contrast, yields from MET2-1939 were poorly correlated with MET2-1919 ($r = 0.17$). Substantial changes in genotype performance between the consecutive MET2s (1937 and 1939) can be seen in Fig. 3*b*.

Finally, consider 2 MET2s conducted during the 1990s. In MET2-1991, genotype deviations were relatively small and the environment main effect was such that the MET means were similar to those overall. Genotype performance differences were much greater in MET2-1997. For many genotypes, MET2-1997 was a direct contrast with MET2-1939. Although a correlation was still present between these 2 METs ($r = 0.55$, $P < 0.01$), many genotypes (almost all with S or U levels for PH, i.e. not early maturing) that performed well in 1939 were poor performers in MET2-1997.

Variance components and correlations of METs with overall data

For the entire dataset, the sizes of the genotypic and G × E interaction variance components were 0.187 ± 0.037 and 0.223 ± 0.002, respectively. For the MET2 data, the G and G × E components ranged from minimums of 0.015 and 0.074 to maximums of 0.59 and 0.66, respectively (Fig. 4). Averaged across all of the MET2s, G was 110% of G × E. However, in 28 of the 54 MET2s, G was less than G × E, often substantially. In the MET2s discussed above, 1919, 1937, 1939, 1997, the G/G × E ratios were 0.35, 0.56, 0.19, and 0.48, respectively. For each of the MET2s, the correlation coefficients between the genotype means from the MET and the overall genotype means are also given in Fig. 4. Lower correlations were generally, but not exclusively, associated with low G.

Whereas Fig. 4 documents the correlation between MET2 averages and overall averages for the 54 genotypes, Fig. 5 demonstrates the effect of conducting METs that consist of different numbers of years. When the MET was only 1 year, the correlation coefficient ranged from 0.09 to 0.99 with an average of 0.82 (signif. $r = 0.27$ for $P < 0.05$). As the length of the MET was increased from 1 to 6 years, the mean correlation increased to 0.97. For METs beyond 3 years in length, all correlations were greater than 0.8. When the same analysis was applied to the REML genotype by trial estimates, the correlation results were unchanged.

Discussion

In the design of this simulation experiment, we constructed genotypes that differed in the degree in which they express traits that have been demonstrated to have effects on the yield adaptation of sorghum to dryland conditions, particularly through the way in which the growth of the crop is matched with the pattern of water supply through the season. The ranges of trait variation defined were derived from prior experimentation on the pool of sorghum germplasm used in the region. Although a simple genetic model was used (4 traits with up to 3 expression states that were additive in nature), this was sufficient to generate substantial genotypic and G × E interaction effects. Estimated variance components for these 2 effects ranged 4- or 5-fold in METs of 2 years. Although genotype yields from samples of 6 locations over 2 years were often reasonably correlated with means for the entire dataset, there was substantial re-ranking of the genotypes within many of these METs. The results are quite consistent with the types of variation observed in hybrid trials (Chapman *et al.* 2000*a*).

The simulated germplasm pool has a broad base (i.e. high genotypic variance) and it may be argued that the pool is more representative of the state of a breeding population in the early stages of testing, rather than at the stage of cultivar testing. Of the 4 traits used, the TE trait has a dominating effect as it is always expressed, resulting in higher water use efficiency by the crop in any conditions. Although there is indeed measurable variation for TE (Mortlock and Hammer 1999), the bulk of the germplasm in advanced testing exhibits little variation for this trait. For this reason, the data presented here may exaggerate the genetic variance for this trait that is encountered in the advanced stages of the current selection program.

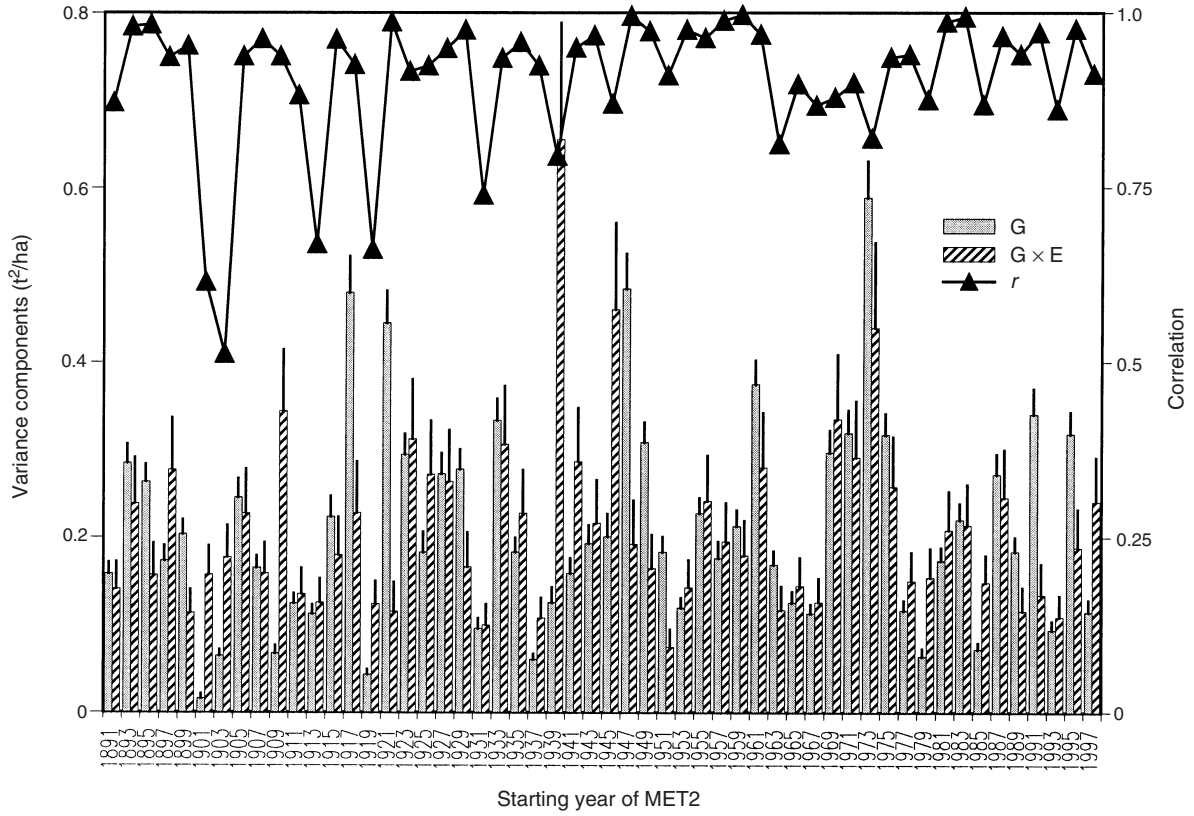

Fig. 4. For a sequence of 54 simulated 2-year METs of 54 genotypes, variance components for genotypic (G) and genotype by environment (G × E) effects and correlations (*r*) between genotype means for the METs and the genotype averages for 547 trials ($n = 54$, $P < 0.05$, signif. $r = 0.27$) (Fig. 3*a*).

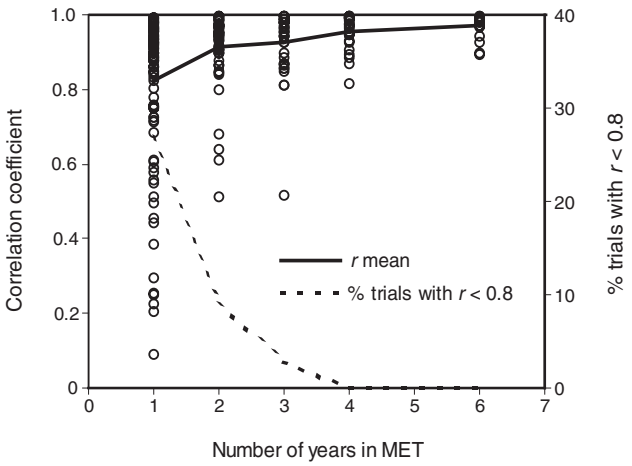

Fig. 5. Correlation coefficients ($n = 53$, $P < 0.05$, signif. $r = 0.27$) between overall genotype means for 547 trials and genotype means estimated from sequential METs consisting of different numbers of years. Lines show for each MET, the mean value of *r* and proportion of *r* values that were less than 0.8.

In each of the 3 regions, the planting criteria were met in fewer than 108 years, especially in the southern-most region (northern NSW) (Table 3) and particularly in the sites with soil of lower water holding capacity (Table 1). The reduced season number in northern NSW is partially a function of the cropping system rules used in the simulation. In the region, growers generally use long fallows (18 months) to allow the soil to accumulate more water prior to planting. By utilising a continuous summer sorghum–winter fallow system in the model, the soil was often not refilled during the winter period, especially in Moree, to the degree it would have been in sites further north. This same explantion holds for the greater number of trial 'failures' in the same sites. The sowing rule also explains why little stress was observed in the first 400 degree-days of the seasons; crops were growing on some stored soil moisture and were able to avoid stress during this period.

The 3 environment types determined by pattern analysis of the seasonal stress indices equated with those defined by Chapman *et al.* (2000*c*) in an analysis of >200 locations across the sorghum growing region. These were a mild terminal stress [equivalent to ET1 in Chapman *et al.* (2000*b*)], a severe terminal stress (ET2) and a mid-season stress that is partially relieved during grain filling (ET3). The 6 locations used in our analysis here were chosen to represent the major locations and soils, with the results indicating that the 3 ETs occur in similar proportions, although ET3 is 7–9% less frequent than the other two (Table 3).

In actual sorghum trials in the region, a correlation has been observed between the experiment error variance and log(yield) of the trial. When this was introduced into the analysis, the most obvious effect was to change the relative performance of the genotypes that differed in expression for the 'phenology' trait. On REML adjusted data, early (L) genotypes performed relatively better overall than they did in the unadjusted data, whereas the late (U) genotypes were relatively poorer overall in the REML adjusted data. This can be explained by the fact that PH(L) genotypes perform best in the more stressed environments that are generally low-yielding and where the effects of the error adjustments were greatest. The REML adjustment scales the data relative to the size of the experiment error variance (DeLacy *et al.* 1996). Hence, in these low yielding environments that were more common in the stressed environments, the yields of all genotypes receive a relatively higher weight (than in other environments) in the REML estimates.

There are two aspects to be considered with respect to improving sampling of the TPE, given the spatial and temporal distribution of environment types. The first is whether repeated use of a location will sample the TPE, and second, to what degree extending METs over years improves sampling of the TPE. In the longer term, Emerald, Gunnedah, and Miles are the most representative of the stress patterns across the entire region. Hence, if attempting to select genotypes for broad adaptation, these sites would be the most appropriate, if testing were feasible over a sufficiently large number of years. However, the use of such 'key sites' would be complicated by the effects of climatic variation on the occurrence of stresses (Fig. 3), as the frequency of stresses within a MET often did not match that of the complete set of environments (this is also the case when only the 3 sites are used, data not shown). In fewer than 5 of the 2-year METs were the frequencies of the 3 environment types within 5% of the target, and in several METs, at least one environment type was not sampled at all (Fig. 3). In these cases, the data from the trials would need to be weighted to indicate the representativeness of the MET as a sample of the TPE (see further discussion below).

In a temporal sense, Figs 3*b* and 4 demonstrate that genotype averages from a 2-year MET are frequently a poor estimate of the long-term performance of individual genotypes. This is despite the generally high correlations with the overall means. Extending the number of years in the MET improves the likelihood that the MET averages will be more representative (Fig. 5). However, extending the METs in time is often not practical, and extensions of the MET in terms of locations are more likely to be acceptable. Alternatives to this scenario are to weight data from each trial. For a single simulated genotype, Chapman *et al.* (2000*b*) classified trials (from 70 years and 6 locations) as being one of several environment types. They found that the mean yield of the genotype in 3-year METs was a better estimate of its long-term mean yield, if the yield from each trial in the MET was first weighted by the long-term frequency of the environment type in which the trial was classified. Using simulation methodology, Podlich and Cooper (1998) demonstrated, for a large number of starting populations and several generations of selection, that genetic advance would be greater using this type of 'weighted selection' than using the trial means only.

The simulations that we have done here can be used in QU-GENE to generate the indirect effects of trait genes on yield as mediated by crop growth and environmental effects. Combining these 2 simulation tools (QU-GENE and APSIM) will enable us to simulate the processes of recombination of the genes determining the traits, evaluation of genotypes in samples of environments from the TPE, and selection strategies. As our understanding of the physiological controls of adaptation improves, the modular structure of APSIM allows us to incorporate these effects into the simulation of crop and environment effects. Similarly, more complex gene action and selection methods can be accommodated by QU-GENE. Obviously, our models are limited by the slow accumulation of this understanding, but provide a reservoir for knowledge. In the experience of Podlich and Cooper (1998), the technology cannot be used for 'cast-iron' recommendations, but rather provides us with a means to eliminate the inefficient ways to operate breeding programs. Readers should be careful in extending our conclusions on the exact temporal nature of the observed G × E interactions and the correlations presented in Figs 4 and 5. This is because our data are based on using a broad genetic base for this sorghum population, which may not reflect the advanced stages of a breeding program when the base population has become genetically narrow.

This paper demonstrates that crop simulation technology can be used to generate and provide insights into the large scale evaluation of genotypes for adaptation to dryland conditions. Although the simulation has been designed with a relatively simple additive genetic model, when examined at the phenotypic level of the yield determined by the trait combinations, it has generated both emergent pleiotropic effects (i.e. effects on the trait and yield) and epistatic effects, whereby some additive gene combinations can generate lower or higher yields than might have been expected. These non-additive effects on yield were a consequence of the interaction of the genotype trait combinations and the environmental conditions encountered in the TPE. The analysis presented here is only of a small dataset and is designed to illustrate how the data can be used. In the future, a much larger dataset will be suitable to test methods of quantitative genetic analysis to determine the additive and epistatic effects of the trait genes on yield. It also functions as a platform to test new statistical methods associated with defining environment and genetic effects.

Acknowledgments

We acknowledge the contributions of staff and scientists in the Queensland Department of Primary Industries (Greg McLean, Andrew Borrell, Enli Wang), CSIRO Sustainable Ecosystems (Erik van Oosteroom), and The University of Queensland (Kevin Micallef) in the development of the new sorghum module and implementing the simulations reported here. Parts of the experimental research supporting the new sorghum module were funded by grants from the Australian Centre for International Agricultural Research (grant CS1/94/968) and the Grains Research Development Corporation (grants DAQ230NR, DAQ370). This work was presented as an oral paper at the International Crops Science Congress, Hamburg, August 2000.

References

Borrell AK, Hammer GL, Henzell RG (2000) Does maintaining green leaf area in sorghum improve yield under drought? II. Dry matter production and yield. *Crop Science* **40**, 1037–1048.

Carberry PS, Muchow RC, Hammer GL (1993) Modelling genotypic and environmental control of leaf area dynamics in grain sorghum. II. Individual leaf level. *Field Crops Research* **33**, 311–328.

Chapman SC, Cooper M, Butler D, Henzell RG (2000*a*) Genotype by environment interactions affecting yield of grain sorghum. I. Characteristics that confound interpretation of hybrid yield. *Australian Journal of Agricultural Research* **51**, 197–207.

Chapman SC, Cooper M, Hammer GL, Butler D (2000*b*) Genotype by environment interactions affecting yield of grain sorghum. II. Frequencies of different seasonal patterns of drought stress are related to location effects on hybrid yields. *Australian Journal of Agricultural Research* **51**, 209–222.

Chapman SC, Cooper M, Podlich DW, Hammer GL (2002) Evaluating plant breeding strategies by simulating gene action and environment effects to predict phenotypes for dryland adaptation. *Agronomy Journal* (Special Issue, Crop Modelling and Genomics; in review)

Chapman SC, Hammer GL, Butler DG, Cooper M (2000*c*) Genotype by environment interactions affecting grain yield of sorghum. III Temporal sequences and spatial patterns in the target population of environments. *Australian Journal of Agricultural Research* **51**, 223–233.

Chapman SC, Hammer GL, Podlich DW, Cooper M (2002) Linking bio-physical and genetic models to integrate physiology, molecular biology and plant breeding. In 'Quantitative genetics, genomics and plant breeding'. (Ed. MS Kang) (In press) (CAB International: Wallingford, UK)

Comstock RE (1977) Quantitative genetics and the design of breeding programs. In 'Proceedings of the International Conference on Quantitative Genetics'. Ames, Iowa, 16–21 August 1976. pp. 705–718. (Iowa State University Press: Ames, IA)

Cooper M, Chapman SC, Podlich DW, Hammer GL (2001) The GP problem: quantifying gene to phenotype relationships. *In-silico Biology* <http://www.bioinfo.de/isb/2001/ (In press)

Cooper M, Podlich DW, Jensen NM, Chapman SC, Hammer GL (1999) Modelling plant breeding programs. *Trends in Agronomy* **2**, 33–64.

Cullis BR, Thomson FM, Fisher JA, Gilmour AR, Thompson R (1996) The analysis of the NSW wheat variety database. I. Modelling trial error variance. *Theoretical and Applied Genetics* **92**, 21–27.

DeLacy IH, Cooper M, Basford KE (1996) Relationships among analytical methods used to study genotype-by-environment interactions and evaluation of their impact on response to selection. In 'Genotype-by-environment interaction: new perspectives'. (Eds MS Kang, HG Gauch) pp. 51–84. (CRC Press: Boca Raton, FL)

Gilmour AR, Cullis BR, Welham SJ, Thompson R (1998) ASREML program user manual. NSW Agriculture Biometrics Bulletin.

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.

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, Sept. 1999. pp. 79–85. (Wheat Breeding Society of Australia: Toowoomba, Qld)

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.

Hammer GL, Muchow RC (1994) Assessing climatic risk to sorghum production in water-limited subtropical environments. I. Development and testing of a simulation model. *Field Crops Research* **36**, 221–234.

Hammer GL, van Oosteroom E, Chapman SC, Mclean G (2001) Supply and demand economics applied to crop growth. In '4th Australian Sorghum Conference, Kooralbyn, Qld, Feb. 2001'. Australian Institute of Agricultural Scientists and Technologists (CD-ROM).

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.

Henzell RG (1992) Grain sorghum breeding in Australia: current status and future prospects. In 'Proceedings of the 2nd Australian Sorghum Conference Vol. 1'. Gatton, 4–6 February 1992. (Eds MA Foale, RG Henzell, PN Vance) pp. 70–82. (Australian Institute of Agricultural Science: Melbourne)

McCown RL, Hammer GL, Hargreaves JNG, Holzworth DP, Freebairn DM (1996) APSIM: a novel software system for model development, model testing, and simulation in agricultural systems research. *Agricultural Systems* **50**, 255–271.

Mortlock MY, Hammer GL (1999) Genotype and water limitation effects on transpiration efficiency in sorghum. *Journal of Crop Production* **2**, 265–286.

Muchow RC, Cooper M, Hammer GL (1996) Characterizing environmental challenges using models. In 'Plant adaptation and crop improvement'. (Eds M Cooper, GL Hammer) pp. 349–364. (CAB International, ICRISAT & IRRI: Wallingford, UK)

Nicholls N (1988) El Niño-Southern Oscillation and rainfall variability. *Journal of Climate* **1**, 418–421.

Patterson HD, Thompson R (1975) Maximum likelihood estimation of components of variance. In 'Proceedings of the 8th International Biometrics Conference'. (Ed. LCA Corsten) pp. 197–207. International Biometric Conference, 1974, Constanta, Romania. (Editura Academiei Republicii Socialiste: Bucharest, Romania)

Podlich DW, Cooper M (1998) QU-GENE: a simulation platform for quantitative analysis of genetic models. *Bioinformatics* **14**, 632–653.

Podlich DW, Cooper M, Basford KE (1999) Computer simulation of a selection strategy to accommodate genotype–environment interactions in a wheat recurrent selection programme. *Plant Breeding* **118**, 17–28.

Manuscript received 16 March 2001, accepted 26 October 2001