Theor Appl Genet (2007) 115:819–835
DOI 10.1007/s00122-007-0611-4

ORIGINAL PAPER

Global adaptation patterns of Australian and CIMMYT spring bread wheat

Ky L. Mathews · Scott C. Chapman · Richard Trethowan · Wolfgang Pfeiffer · Maarten van Ginkel · Jose Crossa · Thomas Payne · Ian DeLacy · Paul N. Fox · Mark Cooper

Received: 24 February 2007 / Accepted: 8 July 2007 / Published online: 1 September 2007

Abstract The International Adaptation Trial (IAT) is a special purpose nursery designed to investigate the genotype-by-environment interactions and worldwide adaptation for grain yield of Australian and CIMMYT spring bread wheat (*Triticum aestivum* L.) and durum wheat (*T. turgidum* L. var. *durum*). The IAT contains lines representing Australian and CIMMYT wheat breeding programs and was distributed to 91 countries between 2000 and 2004. Yield data of 41 reference lines from 106 trials were analysed. A multiplicative mixed model accounted for trial variance heterogeneity and inter-trial correlations characteristic of multi-environment trials. A factor analytic model explained 48% of the genetic variance for the reference lines. Pedigree information was then incorporated to partition the genetic line effects into additive and non-additive components. This model explained 67 and 56% of the additive by environment and non-additive by environment genetic variances, respectively. Australian and CIMMYT germplasm showed good adaptation to their respective target production environments. In general, Australian lines performed well in south and west Australia, South America, southern Africa, Iran and high latitude European and Canadian locations. CIMMYT lines

Communicated by F. van Eeuwijk.

Electronic supplementary material The online version of this article (doi:10.1007/s00122-007-0611-4) contains supplementary material, which is available to authorized users.

Ky L. Mathews (✉) · I. DeLacy · M. Cooper
The School of Land, Crop and Food Sciences,
The University of Queensland, St. Lucia,
QLD 4072, Australia
e-mail: ky.mathews@csiro.au

S. C. Chapman
CSIRO Plant Industry, Queensland Biosciences Precinct,
306 Carmody Rd, St. Lucia, QLD 4067, Australia
e-mail: scott.chapman@csiro.au

R. Trethowan · W. Pfeiffer · M. van Ginkel ·
J. Crossa · T. Payne
International Maize and Wheat Improvement Center
(CIMMYT), Apdo. Postal 6-641, 06600 México D.F, Mexico

P. N. Fox
ACIAR, GPO Box 1571, Canberra, ACT 2601, Australia

Present Address:
K. L. Mathews
CSIRO Plant Industry, Queensland Biosciences Precinct,
306 Carmody Rd, St. Lucia, QLD 4067, Australia

Present Address:
R. Trethowan
Plant Breeding Institute, The University of Sydney,
PMB 11, Camden, NSW 2570, Australia

Present Address:
W. Pfeiffer
HarvestPlus, c/o CIAT, A.A. 6713 Cali, Colombia

Present Address:
M. van Ginkel
Department of Primary Industries, Horsham,
VIC 3401, Australia

Present Address:
M. Cooper
Pioneer Hi-Bred International Inc,
P.O. Box 552, Johnston, IA 50131, USA

performed well at CIMMYT's key yield testing location in Mexico (CIANO), north-eastern Australia, the Indo-Gangetic plains, West Asia North Africa and locations in Europe and Canada. Maturity explained some of the global adaptation patterns. In general, southern Australian germplasm were later maturing than CIMMYT material. While CIANO continues to provide adapted lines to northern Australia, selecting for yield among later maturing CIMMYT material in CIANO may identify lines adapted to southern and western Australian environments.

Introduction

For much of its history, wheat (*Triticum* sp.) breeding in Australia has been regionally based. Diverse and variable environments over the area sown to wheat [latitude range of 22°S (sub-tropical) to 38°S (temperate)] has led to specific adaptation of regional gene pools and genetic diversity (Brennan and Fox 1998; O'Brien et al. 2001; Brennan and Quade 2006). The introduction of CIMMYT (International Maize and Wheat Improvement Center) semi-dwarf wheats into Australia in the early 1960s contributed to improved yields in these environments and significantly increased the diversity of the Australian germplasm pool (O'Brien et al. 2001; Parker et al. 2002). Regular introductions from CIMMYT have maintained and broadened the genetic base of disease resistance, and between 1973 and 1993 were shown to have increased the genetic diversity of Australian germplasm (Brennan and Fox 1998). The CIMMYT contribution to Australian-grown varieties was approximately 20% by area in 2003, having reached a maximum of 35% in 1990 (Brennan and Quade 2006).

While CIMMYT germplasm is commonly utilised in Australia for specific traits, such as disease resistance, there has been a potential divergence in the adaptation of Australian and CIMMYT gene pools over the past 30 years (Brennan and Quade 2006). Different environmental conditions between CIMMYT's major breeding and testing locations (Ciudad Obregón, north-western Mexico, 27°20′N and 38 m above sea level and Toluca in the central Mexican highlands, 19°16′N and 2,640 m above sea level) provide different selection pressures that result in high disease pressures and selection for low sensitivity of flowering time to photoperiod. Australian environments in the south and west have chemically and physically hostile soils (Sadras et al. 2003) whilst rainfall patterns vary considerably across the wheat growing regions (Stephens and Lyons 1998a). In contrast, the CIMMYT locations have relatively non-toxic soils and predictable rainfall patterns; semi-arid conditions in Obregón (average annual rainfall 270 mm, supplemented by irrigation) and high rainfall in Toluca (average annual rainfall 800 mm).

With the recent restructuring and increased commercialisation of Australian wheat breeding, it is timely to consider how to best utilise CIMMYT germplasm in Australia, especially as several new breeding programs are developing a national focus, and a need for broad-scale evaluations. Conversely, where specific adaptation has been developed in Australian germplasm, such as adaptation to soil toxicities, it may be of use in CIMMYT's public breeding program, and by breeding programs in other countries.

Cooper and Woodruff (1993) showed that indirect selection, based on yield performance in CIMMYT international trials could be used to identify potential CIMMYT germplasm for environments in Australia's northern production region. The results of the International Adaptation Trial (IAT) reported in this study provide an opportunity to verify this in newer germplasm and to begin to determine the underlying reasons for the observed adaptation patterns. The IAT contains both 'probe' and 'reference' lines. Probe lines allow investigation of specific environmental factors, such as response to soil toxicity. Reference lines represent a sample of the germplasm base from a breeding program and are used to investigate relationships among environments (Cooper and Fox 1996). Global adaptation patterns were investigated in the present study for a set of modern Australian and CIMMYT reference lines in the IAT; compared with Cooper and Woodruff (1993) who investigated the adaptation of CIMMYT nurseries only.

In previous analyses of the IAT, the value of the semi-dwarfing alleles was demonstrated, especially where trial yield was >ca. 3 t ha^{-1}, by comparing the performance of near-isolines for the *Rht-B1b* (*Rht1*) and *Rht-D1b* (*Rht2*) (Mathews et al. 2006; Chapman et al. 2007). Another important trait, genotype maturity, is often used to manage adaptation to variation in the time of sowing and avoidance of frost exposure at flowering, and has been considered in Australian environments to be a major contributor to G × E for yield (Cullis et al. 1996; Frensham et al. 1998). It is also an important trait to consider when investigating the underlying causes of global adaptation patterns and so was incorporated into the present study.

Plant breeders frequently use specific mating designs to estimate the general combining ability (which includes both additive and additive × additive components) of lines, and in hybrid crops to interpret changes in this, and specific combining ability (non-additive: which includes both dominance and components of all sources of epistatic effects) with environment (de la Vega and Chapman 2006). Advances in statistical methodology allow inclusion of pedigree information to model the genetic effects in multi-environment trials (METs) as is common practice in animal breeding (Crossa et al. 2006; de la Vega and Chapman 2006; Oakey et al. 2006, 2007; Burgueño et al. 2007). The pedigree information, in the form of an additive relationship

matrix, partitions the overall genetic effect into additive and non-additive components. In particular, the estimation of additive genetic effects allows wheat breeders to better identify parents when developing new germplasm with adaptation to subsets of the test environments.

A key assumption of METs, such as the IAT, is that differences in the adaptive patterns between the genotypes can be investigated through comprehensive analysis of genotype by environment interactions (G × E) revealed for important traits; e.g., grain yield, phenology, disease resistance and tolerance to abiotic stresses. Many 'fixed effect' models have been used, such as additive main effects and multiplicative interaction models (AMMI), and the shifted multiplicative (SHMM) and site regression (SREG) models of Crossa and Cornelius (2002), e.g. Trethowan et al. (2001), Lillemo et al. (2004) and others. These models are generally used in conjunction with pattern analysis, a combination of ordination and clustering analyses. Pattern analysis techniques use the genotype performance data from METs to discriminate among environments and to conduct a complementary analysis of genotype broad and specific adaptation (Peterson and Pfeiffer 1989; Cooper and DeLacy 1994; DeLacy et al. 1996). More recently, linear mixed effect models have been used and allow detailed modelling of intra- and inter-trial correlations of unbalanced datasets while including random multiplicative terms to explicitly model the structure of the G × E (Piepho 1997, 1998; Eeuwijk et al. 2001; Smith et al. 2001b; Cullis et al. 2003). It is instructive to see whether these models can perform on large datasets that comprise substantial imbalance in genotypes, replication and experimental design.

This paper examines the grain yield adaptation of representative Australian and CIMMYT germplasm for a large sample of global spring wheat production environments. The objective was to investigate the G × E patterns for grain yield in the IAT using multiplicative mixed models and to use the pedigree relationship matrix to partition the additive and non-additive genetic line effects, and investigate their interaction with environment. In addition, flowering data were used to investigate possible causes for the observed adaptive patterns.

Materials and methods

Data set

Sixty bread wheat (*Triticum aestivum* L.) and 20 durum wheat (*Triticum turgidum* L. var. *durum*) lines, primarily of either CIMMYT or Australian origin were chosen by CIMMYT and Australian breeders to represent their breeding programs for drought adaptation characteristics, ability to contrast for soil borne problems (abiotic and biotic) and for differential agronomic traits to assist in environmental characterisation by bioassay. In the Australian grown trials, in 2001–2003, seed quarantine availability resulted in only 59 lines being in common with the CIMMYT distributed IAT. In this paper, a subset of 21 Australian and 20 CIMMYT 'reference' lines considered to represent the existing spring bread wheat germplasm base of the respective breeding programs were of interest (Table 1).

The IAT was distributed globally by CIMMYT in 2000–2004 to representative spring wheat production regions. Grain yield (t ha^{-1}) data from 183 trials (116 locations) between 2000 and 2004 were returned. Participants in the IAT were requested to apply fungicide, but to otherwise employ local agronomic management practices. The dataset was filtered for disease incidence (either reported or clearly detectable through analysis of probe genotypes) and data quality; retaining yield data from 106 trials representing 74 locations (Fig. 1). Eleven of the 106 trials were 'managed environment' trials grown at different irrigation levels at CIMMYT's drought evaluation site, Centro de Investigaciones Agricolas del Noroeste (CIANO), in north-western Mexico (27°20′N, 109°56′W, 38 m above sea level); 32 trials were grown across the wheat production regions of Australia; 26 in Asia and the remainder distributed across Europe, North and South America, the Middle East and Africa. In this study, trials not grown in Australia or at CIANO are referred to as 'international' trials.

Flowering data as days from sowing to heading (DTH) were returned for 73 of the 106 trials, and in-crop weather data were available to calculate thermal time for days to heading (TTDH) for 51 trials. Daily meteorological data were obtained from global surface summary of day (GSOD) dataset maintained by the National Climate Data Centre (http://www.ncdc.noaa.gov) for trial locations other than in Australia or at CIANO; from SILO at the Bureau of Meteorology Australia for Australian locations (http://www.bom.gov.au/silo), and from the CIANO meteorological station for trials at CIANO, Mexico. For the international locations, GSOD meteorological stations selected were within a 100 km radius and 100 m altitude of the trial location. In most cases (65%), the meteorological stations were within 40 km of the trials.

Four types of experimental designs were used by collaborators (the first three by those receiving CIMMYT distributed trials and the last by all Australian grown trials). They were: (i) split-plot two-replicate α-lattice designs, with the split within replicate for bread and durum wheat types for 2000–2002 and 2004 (11 trials at CIANO, which also had row-column spatial information); (ii) separate two-replicate α-lattice designs for bread and durum wheat (durum wheats were not always grown adjacent to the bread wheat block and hence were treated separately for all trials grown internationally in 2001, 2002 and 2004);

Table 1 Australian and CIMMYT reference genotypes (ranked by yield) in the International Adaptation Trial

GenCode	Name	Pedigree	Breeding Program[c]	Yield BLUPs (t/ha)	TTH[d] BLUPs (°C)	Maturity Class	Pedigree Origin	COP[e] with	
								ATTILA	SILVERSTAR
ATTLA	ATTILA	ND/VG9144//KAL/BB/3/YACO/3/VEERY #5	CIMMYT	4.138	1,261	Mid	CIMMYT_Veery	1.0000	0.1685
HXL	HXL7573/2*BAU	HXL7573/2*BAU	CIMMYT	4.106	1,295	Late	CIMMYT_Old	0.1165	0.1491
URESK	URES/JUN//KAUZ	URES/JUN//KAUZ	CIMMYT	4.086	1,271	Mid	CIMMYT_Veery	0.3398	0.2054
CETTIA	CETTIA	CIANO T 79/PARULA//CHILERO	CIMMYT	4.068	1,209	Early	CIMMYT_Old	0.2114	0.1938
PSTOR	PASTOR	PFAU/SERI M 82//BOBWHITE	CIMMYT	4.001	1,286	Late	CIMMYT_Veery	0.2930	0.1881
KAUZ	KAUZ DWARF	PVN/5*SUPER KAUZ	CIMMYT	3.985	1,244	Mid	CIMMYT_Veery	0.3366	0.1880
SSERI	SUPER SERI #1	SERI*4//AGA/6*YR/3/SERI M 82	CIMMYT	3.926	1,294	Late	CIMMYT_Veery	0.5626	0.1850
INQLB	INQALAB 91	WL 711/CROW	CIMMYT	3.919	1,200	Early	CIMMYT_Old	0.1927	0.1846
PSTRO	PASTOR*2/OPATA	PASTOR*2/OPATA M 85	CIMMYT	3.908	1,234	Mid	CIMMYT_Veery	0.2694	0.1864
NESSR	NESSER DWARF	PVN/5*NESSER	CIMMYT	3.866	1,221	Early	AUS_CIMMYT	0.0885	0.0942
KENDY	KENNEDY	SERI M 82/HARTOG	LRC	3.861	1,248	Mid	CIMMYT_Veery	0.4075	0.2741
TUI	TUI	HERMOSILLO M 77/SAPSUCKER//VEERY	CIMMYT	3.846	1,200	Early	CIMMYT_Veery	0.2440	0.1721
JANZ	JANZ	3AG3/4*CONDOR//COOK	LRC	3.820	1,292	Late	AUS_North&East	0.0910	0.4587
SITTA	SITTA	MINIVET/VEERY #5	CIMMYT	3.819	1,234	Mid	CIMMYT_Veery	0.3816	0.2234
CHLPR	CHIL/PRL	CHILERO/PARULA	CIMMYT	3.812	1,195	Early	CIMMYT_Old	0.2121	0.1912
JUNB	JUN/BOMB	JUNCO/BUCK OMBU	CIMMYT	3.812	1,245	Mid	CIMMYT_Old	0.1655	0.2204
WESTA	WESTONIA	SPICA/TIMGALEN//TOSCA/3/CRANBROOK//JACUP*2/BOBWHITE	WADA	3.809	1,179	Early	AUS_CIMMYT	0.1065	0.1189
SLVSR	SILVERSTAR	PAVON S/TM56	WW	3.803	1,152	Early	AUS_North&East	0.1685	1.0000
GLVEZ	GALVEZ DWARF	PVN/5*GALVEZ S 87	CIMMYT	3.775	1,218	Early	CIMMYT_Old	0.2009	0.2767
PRLSV	PRL/SARA//TSI/VEE#5	PRL/SARA//TSI/VEE#5	CIMMYT	3.768	1,272	Mid	CIMMYT_Veery	0.2581	0.1660
PVON	PAVON DWARF	PAVON F 76//PVN*2/SERI/2*PAVON F 76	CIMMYT	3.740	1,254	Mid	CIMMYT_Old	0.2540	0.3591
WW425	WW425	CARIANCA 422/ANAHUAC F 75	WW	3.703	1,276	Late	AUS_CIMMYT	0.0835	0.0855
VLCAN	VULCAN	CONDOR/PITIC 62//CONDOR S	SPS	3.652	1,253	Mid	AUS_North&East	0.1052	0.4416
DMDBD	DIAMONDBIRD[a]	VICAM S 71//CIANO F 67/SIETE CERROS T 66/3/KALYANSONA BLUEBIRD	WW	3.647	1,305	Late	CIMMYT_Old	0.2546	0.2395
PRFDL	PROINTA FEDERAL[b]	SWM1703/CM15856-6M-5Y-500M	CIMMYT	3.638	1,265	Mid	CIMMYT_Veery	0.1932	0.1885
SUNCO	SUNCO	SUN9E-27*4/3AG14//WAGGA WAGGA 15/3/3*COOK	SUN	3.635	1,311	Late	AUS_North&East	0.0907	0.3999
DLRBD	DOLLARBIRD	WREN/GABOTO//KALYAN BLUEBIRD	WW	3.569	1,215	Early	CIMMYT_Old	0.2236	0.2065
EXLBR	EXCALIBUR	RAC177 (SR26)/UNICULM//PITIC S/GLAIVE	RAC	3.553	1,296	Late	AUS_CIMMYT	0.0623	0.0796

Table 1 continued

GenCode	Name	Pedigree	Breeding Program[c]	Yield BLUPs (t/ha)	TTH[d] BLUPs (°C)	Maturity Class	Pedigree Origin	COP[e] with	
								ATTILA	SILVERSTAR
CNDOW	CNDOW	CNDO/R143//ENTE/MEXI_2/3/AEGILOPS SQUARROSA (TAUS)/4/WEAVER	CIMMYT	3.541	1,275	Late	AUS_CIMMYT	0.0713	0.0696
KRCHF	KRICHAUFF	WARIQUAM//KLOKA/PITIC 62/3/WARIMEK/HALBERD/4/3AG3AROONA	WARI	3.539	1,237	Mid	AUS_South	0.0650	0.1628
TRDNT	TRIDENT	VPM1/5*COOK//4*SPEAR	RAC	3.533	1,284	Late	AUS_CIMMYT	0.0533	0.0847
BRNGA	BARUNGA	HALBERD/AROONA//3*SCHOMBURGK/3/2*MOLINEUX	WARI	3.533	1,234	Mid	AUS_South	0.0655	0.2069
BTSCH	BT-SCHOMBURGK	HALBERD/AROONA//3*SCHOMBURGK	WARI	3.532	1,216	Early	AUS_South	0.0599	0.2098
HRTOG	HARTOG[a]	VICAM S 71//CIANO F 67/SIETE CERROS T 66/3/KALYANSONA BLUEBIRD	LRC	3.526	1,238	Mid	CIMMYT_Old	0.2438	0.3648
SONLK	SONALIKA	II53.388/AN//YT54/N10B/3/LR/4/B4946.A.4.18.2.1Y/Y53//3*Y50	CIMMYT	3.515	1,113	Very early	AUS_CIMMYT	0.0811	0.1483
GROKE	GOROKE	TM56*2/4AUSEN21//77-702D	VIDA	3.514	1,267	Late	AUS_North&East	0.0816	0.4548
SUNVL	SUNVALE	COOK*2/VPM1//3*COOK	SUN	3.431	1,356	Very late	AUS_North&East	0.0876	0.4040
SNLIN	SUNLIN	SUNELG*2//SUNECA*3/VPM1	SUN	3.397	1,287	Late	AUS_CIMMYT	0.0419	0.0534
YRWGL	YR10-WARIGAL	WARIGAL*3//AROONA/MORO	WARI	3.374	1,293	Late	AUS_South	0.0619	0.2236
FRAME	FRAME	MOLINEUX/3*DAGGER	WARI	3.365	1,349	Very late	AUS_CIMMYT	0.0529	0.0806
CDOUX	CADOUX	CENTRIFEN/GAMENYA (F3)//GAMENYA/3/JACUP	WADA	3.186	1,304	Late	AUS_CIMMYT	0.0765	0.0879

[a] Hartog is an Australian release from the CIMMYT cross Pavon; Diamondbird is from the same cross but selected in Australian environments

[b] Prointa Federal is a selection out of the cross called BOBWHITE, the pedigree of this cross is given

[c] *WW* Wagga Wagga Agricultural Research Centre; *RAC* Roseworthy Agricultural College; *VIDA* Victorian Institute for Dryland Agriculture; *WADA* Western Australian Department of Agriculture; *SUN* Sydney University Narrabri; *SPS* Sunprime Seeds Ltd, *WARI* Waite Agricultural Research Institute; *LRC* Leslie Research Centre

[d] Average thermal time to heading (T_{base} = 0°C)

[e] Coefficient of parentage

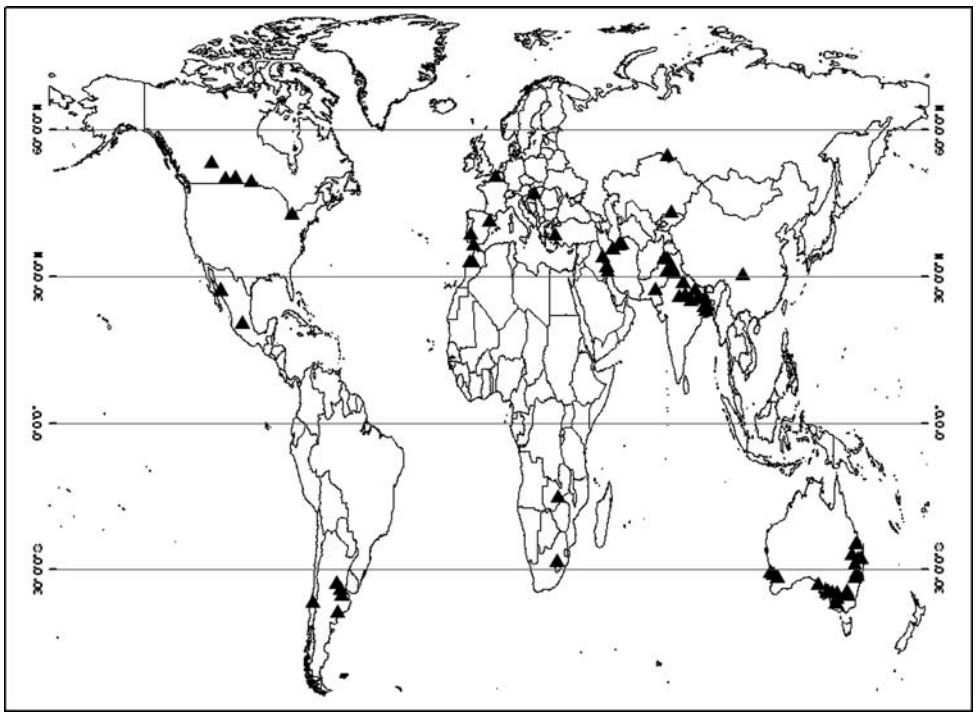

Fig. 1 Global distribution of the International Adaptation Trial 2001–2004

(iii) two-replicate α-lattice design with no split for wheat type (CIMMYT distributed trials in 2003); and (iv) α-row-column designs with spatial information, 28% of the Australian trials were augmented check single replicate designs, the remainder had two replicates. An environment is defined to be a location by year combination. Thus, for designs (i), (iii) and (iv) an environment is a trial, but for design (ii) an environment has the bread and durum wheat trials nested within. Therefore, in the statistical model defined below there are both environment and trial effects.

Statistical analyses

Statistical model

This section describes the process of fitting a multiplicative mixed model (*standard* model) for analysing multi-environment trials. The *standard* model is extended to include the relationship matrix, **A**, thus incorporating pedigree information to model the genetic effects (*pedigree* model). All analyses were performed applying the ASREML software (Gilmour et al. 2006), Version 2.0[1].

Standard model

A multiplicative mixed model was fitted to the raw plot data in a one-stage analysis. However, a one-stage analysis across environments first requires the best design parameters to be determined for each individual trial with genotypes fitted as random (Smith et al. 2001a). For trials where the row and column information was available (Australian and CIANO trials) the residual structure was modelled with an auto-regressive process of order one (AR1) in both the row and column directions. Best spatial models were fitted to these trials following Gilmour et al. (1997). The 'trial-specific' effects, e.g. fixed linear row/column or random row/column, and the design factors, e.g. replicate and incomplete blocks within replicate (both random), were included for each trial, when they made a significant improvement to the base model (based on Wald *F* statistics for fixed effects, log-likelihood ratio tests for random effects). The generalised heritability, h_G^2 was calculated for each trial following Cullis et al. (2006), that is,

$$h_G^2 = 1 - \frac{\text{PEV}}{2\sigma_g^2}, \qquad (1)$$

where PEV, is the predicted error variance, or average variance of the difference, and σ_g^2 is the genotypic variance. The former was obtained from the predict statement in ASREML. The generalised heritability is the proportion of total phenotypic variance explained by the genotypic component and can be used to calculate the expected genetic gain, which Cullis et al. (2006) have shown is well correlated to realized genetic gain.

For environments where the bread and durum wheat blocks were sown as adjacent trials, a single error variance was estimated and trials were nested within each

[1] VSN International Ltd., Hemel Hempstead HP1 1ES, UK, http://www.vsni.co.uk

environment. Although the durum wheats were not of specific interest in this analysis, these trials were retained to minimise the degree of imbalance in the dataset and to allow fitting of design and spatial effects (see the description of trial designs above). The genotype effect structure was separated into two random effects: *reference:* the 41 reference lines with pedigree information, and *other:* the remaining mixture of local checks and probe and durum wheat lines of either the Australian or CIMMYT breeding programs.

The mixed linear model for m genotypes ($i = 1, 2, 3,\ldots, m$) and p ($j = 1, 2, 3,\ldots, p$) environments, and combining the genotype main effects (G) and the G × E effects (GGE, say) is

$$\mathbf{y} = \mathbf{X}\tau + \mathbf{Z_g g} + \mathbf{Z_u u} + \eta \qquad (2)$$

where $\mathbf{y}$ is the $(n \times 1)$ data vector of the response variable across p environments with N_j plots per environment j; $\boldsymbol{\tau}$ is a $(t \times 1)$ vector of fixed effects, including environment main effects and trial-specific effects and $\mathbf{X}$ the associated $(n \times t)$ design matrix of 0s and 1s that relates $\mathbf{y}$ with the fixed effects; $\mathbf{u}$ is a vector of length b which contains subvectors $\mathbf{u}_i$ for each ith random effect with design matrix $\mathbf{Z_u}$ which can be partitioned conformably as $[\mathbf{Z}_{u_1}\ldots\mathbf{Z}_{u_b}]$ and corresponding variance–covariance matrices G_{u_i}. The subvectors are assumed mutually independent with variance $\sigma_i^2\mathbf{I}_{b_i}$. The trial-specific effects in $\boldsymbol{\tau}$ and $\mathbf{u}$ describe the best spatial model for each trial including extraneous field variation and experimental design based terms such as blocking factors (Gilmour et al. 1997). A subvector $\mathbf{u_g}$ for the subset of lines in *other*, that is, the lines that were not modelled using the pedigree information, is also included. The associated variance–covariance matrix $\boldsymbol{G_g} = \mathrm{var}(\mathbf{u_g})$ was modelled as described below for var($\mathbf{g}$). Vector $\mathbf{g}$ is a vector of m random genotype effects for each p environment; thus if some genotypes are not in an environment there will be zero columns in design matrix $\mathbf{Z_g}$ and hence unbalanced data are managed. Vector $\boldsymbol{\eta}$ is the vector of residuals for each observation. The variance–covariance matrix of $\mathbf{g}$ that combines the main effect of genotypes and G × E can be represented by the separable variance structure $\mathrm{var}(\mathbf{g}) = \boldsymbol{G_e} \otimes \boldsymbol{G_v}$, where $\boldsymbol{G_e}$ and $\boldsymbol{G_v}$ are the symmetric $p \times p$ environment and $m \times m$ genotype component matrices, respectively. $\boldsymbol{G_e}$ is the environment genetic variance–covariance matrix and in the *standard* model we assume $\boldsymbol{G_v} = \boldsymbol{I_m}$, i.e. no pedigree structure. In the *standard* model, the variance–covariance matrices of *reference* [var($\mathbf{g}$)] and *other* [var($\mathbf{u_g}$)] were modelled in the same way. For the purposes of the following description of modelling these variance–covariance matrices, $\mathbf{g}$ can represent either the *reference* or *other* genotype subsets.

The dataset was unbalanced for genotypes, with substantial heterogeneity of variance and covariance between trials. Plant breeders typically do not model this variance–covariance structure and often fit a G + G × E model with a homogenous variance and covariance, corresponding to the compound symmetry (CS) model. This model is included for comparative purposes. However, since var($\mathbf{g}$) is not modelled, the $\mathbf{A}$ matrix was not fitted to the CS model. The environment genetic variance–covariance matrix $\boldsymbol{G_e}$ was first modelled using a diagonal (DIAG) structure which allows different genetic variances to be fitted for each environment but assumes a between-trial correlation of zero. This is equivalent to fitting each trial individually. The variance estimates from this model were used as initial starting values for modelling $\boldsymbol{G_e}$ using a factor analytic (FA) structure. Factor analytic models are residual maximum likelihood (REML)-based multiplicative mixed models which allow the variance–covariance matrix of the random effects to be modelled (Piepho 1997, 1998; Eeuwijk et al. 2001; Smith et al. 2001b).

The vector of random genotypic effects representing the combined genotype and G × E effects can be written as

$$\mathbf{g} = (\Lambda \otimes I_m)f + \delta$$

where Λ's are loadings for each environment j, f contains the scores for each genotype i and δ is the residual term for the multiplicative model. The variance matrix for the combined genotype and G × E effects is given by

$$\mathrm{var}(\mathbf{g}) = (\Lambda\Lambda' + \psi) \otimes I_m,$$

where $\boldsymbol{\psi}$ is a diagonal matrix of the p environment specific variances. This accommodates the variance–covariance heterogeneity among trials. Since the genotype main effect is not fitted in this model, the diagonal of var($\mathbf{g}$), σ^2_{GGE} say, is a vector (length p) of the combined genotypic and genotype by environment variance for each environment j. When only one factor is considered, $k = 1$, the model has one multiplicative term and is denoted as FA(1), for $k = 2$ FA(2) has two multiplicative components, etc. When $k > 1$, linear constraints are imposed on the loadings to ensure a unique solution, such that, for $k = 2$, $\lambda_{j2} = 0$ (Smith et al. 2001b).

The factor loadings, Λ, and scores, f, from an FA(2) model were used to produce a factor analytic biplot, whose interpretation is equivalent to the biplots obtained from principal components analysis as described elsewhere (Cooper and DeLacy 1994; DeLacy et al. 1996; Yan and Hunt 2002; de la Vega and Chapman 2006; Mathews et al. 2006). First, the constrained loadings were rotated to obtain a principal component representation such that $\Lambda^* = \Lambda\Lambda'\Lambda$ (Smith et al. 2001b). The rotated loadings were then scaled by the genetic variance for each environment, and the genotype scores, f, scaled by their maximum such that

Table 2 Summary of the environment genetic variance structure fitted for each of the genetic components in each model

Model	q[b]	AIC[c]	Reference		Other
			G_a[d]	G_e or G_i	G_g
0	742	4,803.55	–	CS	CS
1	951	1,445.00	–	DIAG	DIAG
2	1,163	1,374.58	–	XFA1	XFA1
3	984	366.72		XFA2	XFA2
4	984	6,883.74		XFA2 (47)	XFA2
5	989	3,511.10	DIAG	DIAG	DIAG
6	989	3375.10	DIAG (47)[e]	DIAG (97)	DIAG
7	1,036	2,684.42	XFA1 (47)	DIAG (97)	DIAG
8	1,183	2,136.94	XFA1 (47)	XFA1 (97)	DIAG
9	1,289	1,021.66	XFA1 (47)	XFA1 (97)	XFA1
10	1,336	968.80	XFA2 (47)	XFA1 (97)	XFA1
11	1,433	715.86	XFA2 (47)	XFA2 (97)	XFA1
12[a]	1,539	0.00	XFA2 (47)	XFA2 (97)	XFA2

[a] Final model

[b] q number of variance parameters fitted

[c] Akaike Information Criteria (Akaike 1974). AIC are relative to Model 12, so that positive values indicate the AIC is higher than Model 12

[d] G_a: additive variance matrix, G_i: non-additive variance matrix, G_e: genetic variance matrix (standard model), G_g: genetic variance matrix for *other* lines with no pedigree information

[e] (Sites) number of sites fitted (if not specified all sites fitted)

biplot axis scales were from –1 to 1. An environment vector was produced by drawing a line between the origin and the factor loading co-ordinates for that environment. The length of a vector represents the proportion of genetic variance explained by the two factors for that environment; and the cosine of the angle between any two environment vectors is the genetic correlation estimated from the two factors. The genotype scores were plotted as points using the GenCodes as labels (Table 1). To interpret the effect of a genotype in a particular environment, draw a perpendicular line between the genotype and environment vector of interest. The distance from the origin to the intersection with the perpendicular line is a measure of the proportion of genetic variance contributed by that genotype to that environment. The variance explained for each factor k was calculated as

$$\%\text{var explained} = \text{mean}\left(\frac{1}{\sigma^2_{\text{GGE}}}\text{diag}(\lambda_k \lambda'_k)\right) * 100. \qquad (3)$$

Since a genotype main effect was not fitted in these models, the first factor explains the maximum amount of genetic covariance between environments, and the second the next largest amount, orthogonal to the first.

Pedigree model

Oakey et al. (2006) described a methodology to partition additive and non-additive line effects from multi-environment plant breeding datasets, and that method is followed here, albeit for a much larger set of environments. For the 41 reference lines of interest both the pedigree information and replication of lines within and across environments were available. Pedigree information was available in the form of a coefficient of parentage (COP) matrix and, since the raw plot data were fitted directly, was replicated in the analysis. Therefore, the $(m \times 1)$ vector of genetic line effects $\mathbf{g}$ can be partitioned into a vector of additive line effects, $\mathbf{a}$, and a vector of non-additive line effects, $\mathbf{i}$, such that $\mathbf{g} = \mathbf{a} + \mathbf{i}$. The non-additive components, dominance and epistasis cannot be distinguished by this method. In homozygous inbred lines used here, there is no variance among genotypes attributed to dominance effects so that the non-additive effects are interpreted to consist only of additive by additive interaction (epistasis). The mixed model in (2) can then be written as

$$\mathbf{y} = \mathbf{X}\tau + \mathbf{Z}_g\mathbf{a} + \mathbf{Z}_\mathbf{g}\mathbf{i} + \mathbf{Z}_u\mathbf{u} + \eta \qquad (4)$$

where $\mathbf{X}\boldsymbol{\tau}$, $\mathbf{Z}_\mathbf{u}\mathbf{u}$, $\boldsymbol{\eta}$ and $\mathbf{Z}_\mathbf{g}$ are defined as above. The $(m \times 1)$ vector of non-additive effects $\mathbf{i}$ for the m lines with pedigree information has distribution, $\mathbf{I} \sim N(\mathbf{0},\sigma_i^2\mathbf{I}_m)$.

The $(m \times 1)$ vector of additive effects $\mathbf{a}$ of the m lines with pedigree information has distribution, $\mathbf{a} \sim N(\mathbf{0}, \sigma_a^2\mathbf{A})$, where $\mathbf{A}$ is the $(m \times m)$ known additive relationship matrix.

The additive relationship matrix can be obtained directly from the COP matrix (Henderson 1976). Each COP estimates the expected percentage of alleles identical by descent at loci within a given reference population. The COP between any two lines is an estimate of the expected genome-wide inbreeding coefficient of their offspring. A refinement of this calculation for inbred crops assumes that each line is completely homozygous, that lines without common parentage are unrelated and that parents contribute equally to the offspring, despite inbreeding and selection (St Martin 1982). A value close to 1 indicates two lines are closely related while non-related lines have a value of 0. The COP matrix among the 41 lines was calculated using the International Crop Information System (McLaren et al. 2004), and was then multiplied by two to obtain the additive relationship matrix, $\mathbf{A}$. The use of the $\mathbf{A}$ matrix assumes that the reference (or base) population from which the genotypes arise has not undergone selection. This is the cumulative effect of the following assumptions from quantitative genetic theory: (i) the genotypes arise from the same base population, (ii) the genotypes in this base population are unrelated, and (iii) the base population is in linkage phase equilibrium. As these assumptions are rarely achieved in

practical plant breeding the **A** matrix can introduce bias in the analysis. However, the use of the **A** matrix is justified when it is considered as a measure of similarity, to estimate the genetic distances between genotypes. This is the use intended here. In such situations, the primary motivation for including the **A** matrix in the mixed model is to account for the different degrees of expected coancestry among the genotypes included in the study. For this application it is assumed, as a first approximation, that the additive genetic covariance shows a linear decay with genetic distance. These issues are investigated and discussed in Piepho et al. (2007). For the purposes of the analysis of the IAT data set it is assumed that the base population is the active pool of spring bread wheat lines that can be accessed in Australian and CIMMYT breeding programs. From the perspective of a wheat breeder using germplasm from Australian and CIMMYT breeding programs, this is a relevant base population. However, this is not the classical reference population of quantitative genetics and it can be questioned whether all of the genetic assumptions will hold. In this analysis, making inferences to the assumed base population is not the primary motivation for including the **A** matrix and should only be considered with caution. Rather, the **A** matrix is utilised as a practical, approximate method to account for coancestry.

The decomposition of **g** into additive, **a**, and non-additive, **i**, genetic effects leads to $\text{var}(\mathbf{g}) = \boldsymbol{G}_e \otimes \boldsymbol{G}_v$, where $\boldsymbol{G}_v$ is no longer the identity matrix, $\boldsymbol{I}_m$. Instead

$$\text{var}(\mathbf{g}) = \mathbf{G}_a \otimes \mathbf{A} + \mathbf{G}_i \otimes \mathbf{I}_m. \quad (5)$$

The additive and non-additive variance–covariance matrices for environments are $\boldsymbol{G}_a$ and $\boldsymbol{G}_i$ respectively: Model 5 in Oakey et al. (2007). The additive and non-additive interaction with environment are referred to as A × E and I × E, respectively.

Factor analytic biplots for the A × E and I × E components were produced in the same way as described for the G × E component. Providing biplots for these two components allows a visual dissection of the G × E patterns so that the underlying genetic basis for adaptation can be explored. The A × E biplot allows genotypes with large additive variance and broad adaptation to be identified as potential parents. I × E biplots allows genotypes with specific adaptation to an environment to be identified, but if their I × E component is large (far from the origin) then they may be better identified for variety selections, rather than potential parents.

Maturity class

In the IAT dataset, relative maturity, as estimated by DTH did not significantly explain the variability observed for yield (data not shown). Across trials, there was a large variation in latitude (38°29′S to 55°6′N) and average maximum temperature (8.69–32.1°C). Therefore, the DTH were converted to thermal degree days for the 51 trials where both DTH and in-crop daily weather data were available:

$$\text{TTH} = \sum \left(\frac{T_{\min} + T_{\max}}{2} - T_{\text{base}} \right), \quad (6)$$

where $T_{\text{base}} = 0$°C and the summation is across the days from sowing to heading.

Expressing DTH in thermal time accounts for the main effects of temperature on development, and in spring wheats (with low vernalisation requirement) will generally emphasise genotypic differences in photoperiod response, but should also capture minor vernalisation effects.

A G × E matrix of the best linear unbiased predictors (BLUPs) for TTH was produced from the *standard* model (i.e. **A** matrix was not modelled) as described above for yield; five of the 51 trials were removed due to lack of genetic variance. The clustering and ordination methods of pattern analysis (Cooper et al. 1996; DeLacy et al. 1996) were applied using R version 2.4 (R Development Core Team 2007) to this matrix which was standardised by columns (environments) with squared Euclidean distance as the proximity measure and Ward's incremental sums of squares method as the fusion criterion in an agglomerative hierarchical clustering procedure. This resulted in five classes (Table 1) which explained 91% of the genotypic variation for TTH and assisted in interpreting the G × E, A × E and I × E patterns.

Results

Standard model

The trial mean yield varied from 0.42 to 8.69t ha^{-1}, the generalised heritability, h^2_G, calculated for each trial, ranged from 0.13 to 0.97 and the trial genetic variances, σ^2_g, from 0.004 to 1.566 for yield. (Fig. 2). There was no relationship between mean yield and heritability ($R^2 = 0.016$, $P > 0.05$). The (G × E)/G ratio from the typical G + G × E model for yield where the trial variance–covariance matrix was not modelled was 3.8 (Model 0, Table 2), indicating a large G × E for this dataset. A table of the results for individual environments is available online in an electronic supplement (Table S1).

The mixed models were developed incrementally from the *standard* (**g**) to the *pedigree* (**a** + **i**) model for the 106 trials (Table 2). For the *standard model*, where $\boldsymbol{G}_v = \boldsymbol{I}_m$ was assumed, a diagonal (DIAG, Model 1) variance–covariance

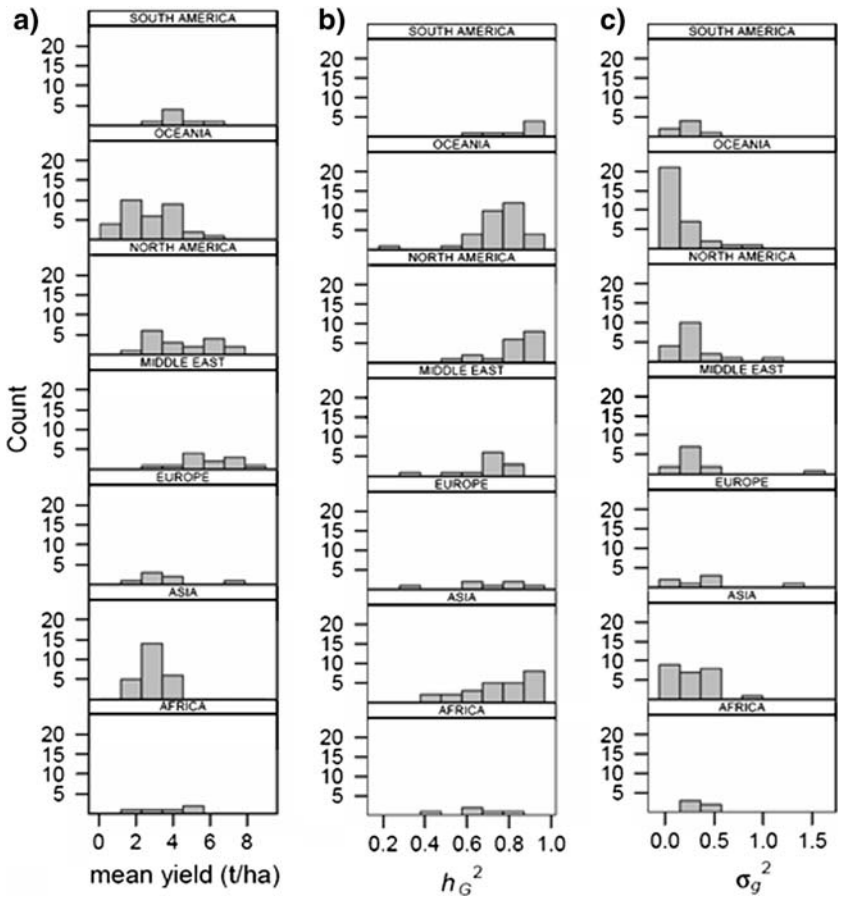

Fig. 2 Distribution of environment mean yield (**a**) generalised heritability (**b**) and (**c**) genetic variance for the 106 individual environments by region. North America includes the 11 CIANO trials

structure was first used to model the $\boldsymbol{G}_e$ matrix of both the *reference* (**g**) and *other* ($\mathbf{u}_g$) random genotype effects. The resulting parameters were used to initialise the FA(1) model and, subsequently, the FA(1) parameters initialised the FA(2) model. The Akaike information criterion (AIC) was used to compare non-nested models (Akaike 1974) with the lowest AIC being found when using the FA(2) model for both var(**g**) (*reference*) and var($\mathbf{u}_g$) (*other*) matrices (Model 3, Table 2). The average percentage of genotypic variance explained by the FA(2) in these two components was 48% for the *reference* and 66% for *other* genotype effects. The AIC for the G + G × E model (Model 0, Table 2) was much larger than Models 1–3 where the variance–covariance structure was modelled and the A matrix was not (Table 2), indicating a poor fit to the data for this model.

The biplot (Fig. 3) shows the G × E pattern from the FA(2) model (Model 3, Table 2) of the 41 reference lines in the IAT, explaining 30.1 and 17.6% of the variance in factors 1 and 2, respectively. While the FA(2) model explains almost 50% of the genetic variance, there are components of the variation that have not been captured in the biplot and prudence is recommended in the interpretation. Thirteen of 18 trials in south and western Australian were clustered on the left-hand side and were not well correlated with the CIANO (CIMMYT's research station) environments on the right-hand side (see electronic supplement Table S1 for factor loadings). However, they were well correlated with trials from South Africa, Argentina, Iran and high latitude locations in Canada and Hungary. Seven of the 10 trials in the northern region of Australia were well correlated with the CIANO environments. Factor 1 (horizontal axis) clearly separated most of the genotypes into the Australian (left-hand side) and CIMMYT (right-hand side) germplasm groups. Thus, for this set of germplasm, the Australian lines are well adapted to southern and western Australian environments and CIMMYT lines well adapted to CIANO, northern Australian and international locations in India, Bangladesh, Iran, northern Africa and Argentina.

The Australian lines, Westonia and Silverstar, and to a lesser extent Janz were broadly adapted to Australian environments, while most other Australian lines were generally best adapted to the regions where they were bred. Kennedy, an Australian line with direct CIMMYT co-ancestry, was the only Australian line that showed moderate broad adaptation to all environments. In keeping with CIMMYT's major breeding objective, a number of CIMMYT lines, Cettia, Attila, Kauz and HXL7573/2*Bau were broadly adapted to most environments. However, in general, the CIMMYT lines were not well adapted to the

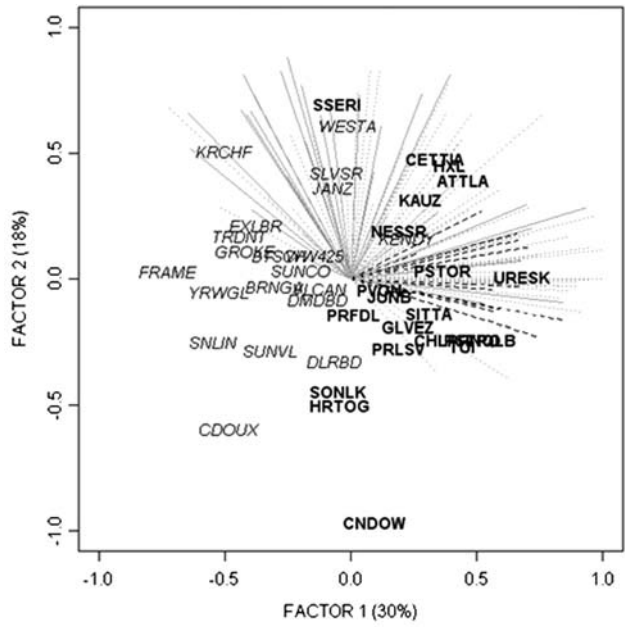

Fig. 3 Factor analytic, $k = 2$, biplot of IAT yield data, Model 3 from Table 2. Australian lines (*ITALIC*), CIMMYT lines (**BOLD**). Vector representations are Australian sites *solid lines*; CIANO *dashed lines*; international sites *dotted lines*

southern and western environments of Australia. Choosing Silverstar and Attila as examples of broadly adapted lines from Australia and CIMMYT, respectively, their co-ancestry with the other reference lines was investigated (Table 1). The COP values indicate that Silverstar is most related to northern and eastern released Australian lines whilst Attila is most related to CIMMYT 'Veery' wheats (COP > 0.35), and the COP between these two lines is 0.1685 (Table 1). The full COP matrix is available online in an electronic supplement (Table S2).

Pedigree model

The relationship matrix, **A**, was fitted to determine how much of the divergence in adaptation patterns seen in the G × E biplot (Fig. 3) could be explained by the pedigree structure of these 41 lines. From the diagonal model (Model 5, Table 2) of the additive and non-additive components, **a** and **i**, 47 and 97 of the 106 environments had non-null effects, respectively. There were no environments where both the additive and non-additive effects were null. A FA(2) *standard* (**g**) model (Model 4, Table 2) applied to the 47 environments with non-null additive genetic effects showed the same patterns as in the G × E biplot (Fig. 3), and explained a similar proportion of G × E, 47%, as the model with all environments (Model 3, Table 2). Therefore, it was appropriate to continue modelling and interpreting the additive and non-additive components of these 47 environments, while the remaining 59 environments were retained in the analysis dataset and modelled without the **A** matrix. The average yield of the 47 environments with significant additive variance was 0.5 t ha^{-1} greater than those with null additive genetic variance. These environment subsets were used in the models 5 to 12, Table 2.

The final model, Model 12, had a FA(2) variance–covariance structure fitted to both the $\boldsymbol{G_a}$ (additive) and $\boldsymbol{G_i}$ (non-additive) components for the *reference* genotypes, and also to the var($\mathbf{u_g}$) for the *other* genotypes. This model fits significantly better than the G + G × E model with a compound symmetry variance–covariance (Model 0, Table 2) typically utilised by plant breeding programs. From Model 12, the additive by environment (A × E) patterns and non-additive by environment (I × E) patterns allowed exploration of the underlying genetic causes in the observed G × E patterns. The A × E biplot (Fig. 4a) again showed the specific adaptation of Australian and CIMMYT germplasm to their respective environments, and also emphasised the diversity of those environments for this component of variance. The average additive genetic variance explained by an FA(2) model was 66.8%.

The northern Australian environments, which lie on the bottom right hand side of the biplot were correlated with trials grown at CIANO, Mexico, under full irrigation and drought in 2001, and with international trials in Pakistan, Pergamino Argentina, Spain, Morocco and Patanga India. Lines that were well adapted to these environments were CIMMYT lines Sonalika, Cndo/R143//Ente/Mexi_2/3/Aegilops Squarrosa (Taus)/4/Weaver (CNDOW), Inqalab 91, Prointa Federal and Chilero/Parula (CHLPR). With the exception of Cndo/R143//Ente/Mexi_2/3/Aegilops Squarrosa (Taus)/4/Weaver, which is derived from a cross with hexaploid synthetic wheat, these can all be described as 'older' CIMMYT materials, and also earlier maturing. The more recently released CIMMYT material such as Attila, Pastor, Ures/Jun//Kauz, which can be considered 'Veery-derived' wheats showed best adaptation in the set of CIANO trials, mostly grown in 2002, and locations in the Indo-Gangetic Plains, West Asia North Africa, Canada and Turkey. They were negatively correlated with the southern Australian environments and were not well correlated with the CIANO, 2001 and northern Australian environments. In general, the 2002 CIANO trials had shorter crop lengths than the 2001 trials at CIANO. This was due to a warmer vegetative stage in 2002 (minimum temperature 10.2°C versus 8.6°C in 2001 and maximum temperature 26.9°C versus 27.6°C) and almost twice the rainfall (32.3 versus 17.2 mm in 2001) during grain filling. For example, there was a 21-day shorter crop length in the irrigated conventional tillage trial grown at CIANO in 2002 compared to 2001, so that this crop matured into a wetter than normal environment, but with less time for tillering during

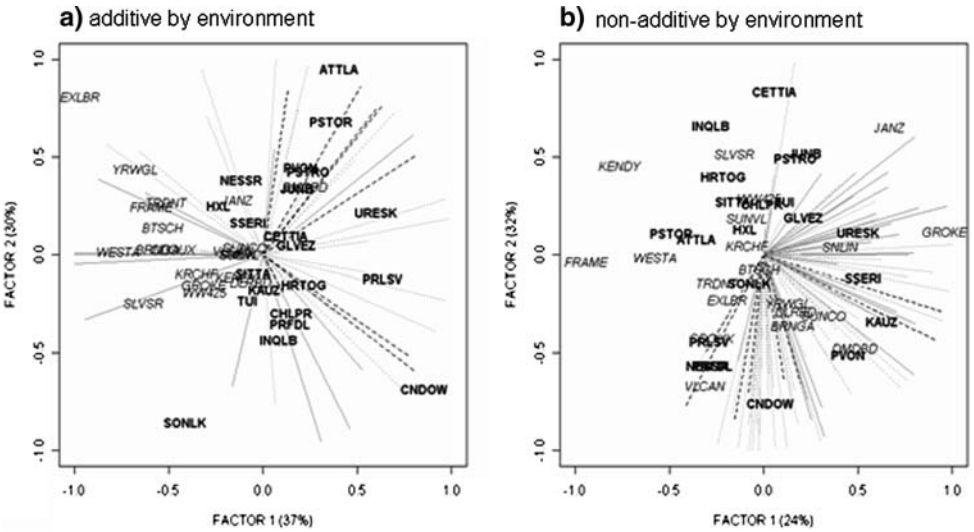

Fig. 4 Factor analytic, $k = 2$, biplots for additive (**a**) and non-additive (**b**) genetic effects. Model 12 from Table 2. Australian lines (*ITALIC*), CIMMYT lines (**BOLD**). Vector representations are Australian sites *solid lines*; CIANO *dashed lines*; international sites *dotted lines*

pre-anthesis development. The Australian trial which correlated with these environments was grown in southern New South Wales at Wagga Wagga in 2003. The three Western Australian trials in this subset were represented in each of the three groups described.

All of the Australian lines, except Diamondbird and Hartog (direct CIMMYT derivatives, Table 1), showed either adaptation to the southern Australian and correlated international environments, for which they were bred or were near the origin (indicating either 'average' adaptation or poor modelling of variation for the line effects in the biplot).

For the non-additive by environment interaction (I × E) the pattern of the environment vectors (Fig. 4b) was similar to that in the G × E biplot (Fig. 3). However, the distribution of the lines across the I × E space was different from that observed in both the G × E and A × E patterns with substantial overlap of the Australian and CIMMYT germplasm sets. The percentage variance explained in the I × E biplot was 56%, suggesting, as in the G × E biplot (Fig. 3), that there are unexplained environmental effects contributing to the variation.

Maturity class

The maturity classes determined from the *standard* model on thermal time to heading (TTH) were used to interpret the A × E patterns. The cluster analysis of TTH resulted in five maturity classes described as very late, late, mid, early and very early (Table 1). The five maturity class means (measured in thermal time) were plotted against the average minimum temperature in the vegetative stage (sowing to 100°C days before flowering) for each environment (Fig. 5). The majority of Australian lines were classified as having late maturity and CIMMYT lines as early and mid maturing lines. The late and very late lines showed a classic vernalisation response, requiring some cold conditions to induce early flowering, relative to warm environments. The late and very late Australian lines (Excalibur, Frame, Sunvale, Trident) are all known to contain vernalisation requiring (vrn) alleles (personal communication, H. Kuchel and M. Brougham 2006), although the exact alleles of all the lines used here are not known. In addition, lines such as Sunvale may also contain photoperiod sensitive alleles. The mid maturing lines had a similar TTH in all environments, and so in warm environments flowered at a substantially earlier date than the late/very late lines. Sonalika was the only member of the very early class and was specifically included in the IAT as a probe genotype for earliness. It was one of the early 'green-revolution' wheats: selected to avoid heat during grain-filling in the Indian plains and released in India in 1967 it became one of the most widely grown lines in the irrigation areas of India, Pakistan, Nepal and Bangladesh (Ortiz and Mowbray 2007). Both the very early and early maturity classes had a decreasing trend with increasing average minimum temperature which is the opposite to a normal vernalisation response, and which could result from other gene combinations related to photoperiod and/or the length of the juvenile (non-responsive) stage. These responses could not be separated from the apparent vernalisation response (Fig. 5).

In the G × E biplot (Fig. 3), the majority of late and very late lines and three Australian early maturing lines (Westonia, Silverstar and BT-Schomburgk) were clustered with Australian and correlated international environments on the left-hand side, whilst the remaining early and most of the mid maturing lines were found on the right-hand side with CIANO, northern Australia and correlated international

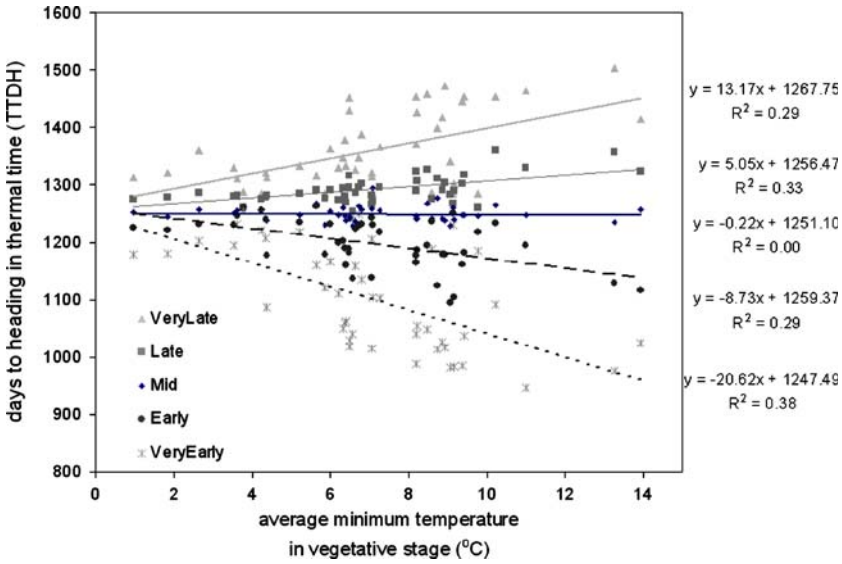

Fig. 5 Days to heading (in thermal time) for maturity class levels versus average minimum temperature (°C) in the vegetative stage

environments. Similar patterns were seen in the A × E biplot (Fig. 4a), although there was slightly more overlap between the early and late maturity groups. In contrast, the I × E biplot (Fig. 4b) displayed a more obvious distinction between the early and late maturity classes. The general trend was from early to mid to late maturity ranging from the top left hand corner to the bottom right hand corner of the biplot. The use of these biplots by plant breeders is discussed further below.

Discussion

This paper applied multiplicative mixed models to examine the global grain yield adaptation patterns of Australian and CIMMYT germplasm. Consistent with previous studies of yield performance across international spring wheat production environments, large G × E interactions were identified (ca. four times the genetic variance). Sivapalan et al. (2003) reported similar findings when considering West Asia North Africa (WANA) and Australian environments, i.e. that Australian and CIMMYT genotypes showed substantial specific adaptation to their respective target environments. Some of the lines common to both studies showed similar responses. Attila, for example, was the top-performer across all environment types in both studies and Nesser, although not the highest-yielding performer in the IAT, was also broadly adapted in both studies. Other lines, such as Pastor (Pfau/Seri//Bow in Sivapalan et al. (2003)) and Kauz performed differently between the studies. In the WANA/Australia study, Pastor was recommended as having wide adaptation across all environments, whereas in the IAT it was better suited to the northern Australian and CIANO environments, than the southern and western Australian environments. The reverse interpretation was made for Kauz. These types of differences are not unexpected between studies with different reference genotype sets and environmental samples (Cooper and DeLacy 1994). While the WANA/Australia study was focused on adaptation for Australian breeding regions with the WANA, the IAT adopted a broader and larger sampling of both Australian and international environments, including the northern region of Australia where CIMMYT germplasm had historically been well adapted.

The current analysis showed that application of multiplicative mixed models is feasible for large, complex datasets and could be more widely adopted in the analysis of METs. These models accommodate unbalanced sets of genotypes and allow a more appropriate modelling of the G × E variance structure compared with the typical practice of using a compound symmetry model which assumes a common variance–covariance structure. For this dataset, the factor analytic mixed model allowed identification of the observable genotypic and environmental differences which contributed to these G × E adaptation patterns. The standard multiplicative mixed model was extended to incorporate the pedigree relationship matrix using the methods described in (Oakey et al. 2006, 2007). This allowed the genetic line effects to be partitioned into additive and non-additive components which have utility in selecting lines as parents for broad or specific adaptation. Fifty-nine of the 106 trials in this analysis had a null additive genetic variance. An investigation of environmental factors, such as rainfall and crop length did not indicate common environmental constraints which may have caused these null additive genetic variances. The relatively small number of lines modelled using pedigree information (41) in this study may not allow for precise

estimation of additive genetic variance. Estimation of the additive and non-additive genetic variances with changing population size is a topic worthy of further investigation.

The difference in patterns observed between the additive by environment interaction (A × E) and non-additive by environment (I × E) patterns has consequences for selection of lines as parents or for release as best commercial lines. The selection of best (stable) lines as parents in some environments is different from selection of the best commercial lines because they show different patterns of response across environments. The genotypes at the extremes of the A × E biplot (Fig. 4a) are likely to be better candidates as parents for use in the environments with which they are associated in the plot. For example, Attila, Pastor, Sonalika, Ures/Jun//Kauz and Cndo/R143//Ente/Mexi_2/3/Aegilops Squarrosa (Taus)/4/Weaver would be good candidates for parents at CIANO and correlated environments while Australian lines: Excalibur, Silverstar, Westonia and YR10-Warigal may be (and are) chosen for south and western Australian and correlated environments. Breeders interested in developing lines adapted to both CIANO and southern Australia might also make crosses between these two germplasm groups. In the I × E plot (Fig. 4b), extreme genotypes show that a greater proportion of their genetic performance is related to specific gene combinations for the environments that they are correlated with. These include Janz and Goroke in southern Australian environments and Cndo/R143//Ente/Mexi_2/3/Aegilops Squarrosa (Taus)/4/Weaver, Kauz and Pavon in CIANO and northern Australia.

The different A × E and I × E patterns for genotype groups with similar pedigrees could provide further insight into the realised breeding value of lines. For example, several 'Cook' type wheats (Vulcan, Janz, Goroke) all have low values for A × E and high values for I × E. This could result from two factors: (i) it is difficult to find good gene combinations for broad adaptation in this background; (ii) strong selection for other traits (e.g. prime hard quality) could result in bias in the relationship matrix, i.e. due to selection they are likely to be more related to each other than is indicated by the COP. Oakey et al. (2006) and Crossa et al. (2006) have discussed other potential limitations in applying the pedigree matrix to these types of data. In addition to the issue of bias introduced by selection, they discussed the mismatch of the identical COP values between full-siblings when it is highly unlikely that their genotypes are identical. Ideally, the method would be applied to data from specific mating designs so that estimates of additive and non-additive effects could be compared between this method and conventional methods. The implications of a higher proportion of non-additive relative to additive effects on breeding efficiency are also significant: recombining specific gene combinations that confer a particular phenotype is much more difficult than selection for traits that are largely under additive control. Significant non-additive genetic effects have been observed in populations derived from crosses between the CIMMYT line Seri M 82 and northern Australian released lines Banks and Hartog (Peake 2003). Hence, it was recommended that large numbers of progeny (> 250) from crosses with high non-additive effects be sampled to maximise response to selection (Peake 2003). Recently, the methods used here were extended to model the additive × additive and additive × additive × environment effects, however, the non-additive effects were not modelled (Burgueño et al. 2007). By modelling both the additive and additive × additive effects, a better approximation to the general combining ability is possible.

CIMMYT's wheat breeding program has been built around a system of shuttle breeding between CIANO, located in the desert of north-western Mexico (27°N, 38 m above sea level) and the highlands of central Mexico (19°N, 2,640 m above sea level). This has been a successful system as germplasm is selected under two diverse environments and is cycled quickly (Braun et al. 1996). One of the key outcomes of this alternating selection system is a general lack of photoperiod sensitivity in the derived materials and, due to short turn-around times between seasons, a relatively low vernalisation requirement (Ortiz and Mowbray 2007). Of the four CIMMYT lines classed as 'late' in the IAT, HXL7573/2*Bau is derived from a cross between a CIMMYT and a Chinese wheat from high latitudes, while Cndo/R143//Ente/Mexi_2/3/Aegilops Squarrosa (Taus)/4/Weaver is derived from a cross with hexaploid synthetic wheat and a late maturity type. The remaining two CIMMYT lines in the 'late' maturity class were Pastor and Super Seri #1, which both contain the rye translocation on the short arm of chromosome 1B (1BL/1RS) and performed well at most Australian locations. Australian production environments range in average crop lengths of 130 to 160 days in the northern region and of 170 to more than 200 days in the southern region, while western environments were represented in each of these groups. Sowing time for wheat in Australia is highly dependent on rainfall occurrence between April and June (Stephens and Lyons 1998b). In the northern region, significant rainfall in March/April will trigger sowing in April/May. However, if rainfall is delayed, sowing may occur in June or July. To avoid frost at flowering while ensuring that grain-filling is completed before summer, mid to late maturing varieties are preferred for early sowing and early lines are preferred for later sowing. In the southern and western environments, later maturing varieties perform best in the longer season environments (Riffkin et al. 2003), but in general, the southern and western lines tend to be later maturing that the northern lines.

The maturity classes derived from cluster analyses using days to heading measured in thermal time assisted in interpretation of the adaptation patterns in the biplots (Figs. 3, 4), although, it is clear that phenology is not the only driver for the difference in adaptation between Australian and CIMMYT germplasm. The maturity class responses against temperature showed some possible vernalisation differences among the five thermal time classes that could be investigated further (Fig. 5).

Crop season length appeared to be a major differentiating factor between northern and southern Australian environments and influenced the adaptation of CIMMYT derived materials in southern and Western Australia. If selection for yield among later maturing lines is conducted in Mexico, rather than selection of the highest yielding materials regardless of maturity class, then materials better adapted to southern Australia may be identified more frequently. Additionally, soil abiotic and biotic constraints differ among northern and southern Australia sites and CIMMYT's CIANO yield testing site. Soils in northern Australia and CIANO are similar in structure as are their biotic constraints (Doyle et al. 1987; Nicol and Ortiz-Monasterio 2000; Bell and Eagles 2003); whereas soils in southern Australia tend to be more 'hostile' to crop root growth.

While Australian and CIMMYT germplasm were well adapted to their respective target breeding regions there is potential for the higher yielding broadly adapted CIMMYT germplasm to have greater impact in Australia and for Australian germplasm to be used for adaptation to international regions with similar environmental constraints. Future research incorporating environmental characterisation would assist in interpreting the G × E patterns observed here and the selection of appropriate germplasm, either Australian or CIMMYT. For example, the following observations can be made with reference to the mega-environment (ME) classification system devised by CIMMYT, and updated periodically (Rajaram et al. 1994; Rajaram and Van Ginkel 2001; Trethowan et al. 2005). International environments that correlated with southern and western Australian environments in this study tended to be from ME9 (low rainfall, moderately cold, facultative growth habit). Locations from ME1 (low rainfall/irrigated temperate, spring growth habit) and ME5 (high rainfall/irrigated, hot, spring growth habit) locations correlated well with CIANO and northern Australian environments. Therefore, Australian wheat breeders working in southern and Western Australia may derive greater benefit by sourcing high yielding lines from ME9 environments rather than targetting the traditional CIMMYT spring wheat nurseries. In contrast, the correlation of northern Australian environments with key CIMMYT selection environments, both CIANO and international spring wheat environments, is consistent with previous studies and highlights the potential for indirect selection based on the results of CIMMYT METs (Cooper et al. 1993).

In summary we recommend that (i) multiplicative mixed models be considered for routine analysis of multi-environment trials and have shown that this is possible for a large, unbalanced dataset; (ii) pedigree information be included to quantify additive and non-additive effects and provide extended knowledge for selection of parental lines. The analysis of the IAT has demonstrated that Australian and CIMMYT germplasm are well adapted to their respective target environments and these environments are diverse. Further, maturity, controlled by photoperiod response, vernalisation requirement and agronomic factors (e.g. irrigation and fertiliser), explains some of the adaptation differences between Australian and CIMMYT germplasm and their adaptation to their diverse environments.

Acknowledgments We gratefully acknowledge Arthur Gilmour and Helena Oakey for their input in the implementation of the statistical models in ASReml and Sandra Micallef for providing COP information. We thank the many international collaborators who grew the IAT, especially the following Australian wheat breeders for their support (in alphabetical order): R. Eastwood, S. Jefferies, M. Lu, P. Martin, G. Rebetzke, A. Scott, J. Sheppard, R. Shorter, P. Wilson and R. Wilson. We also appreciate the Queensland Department of Primary Industries for the seed distribution of the IAT in Australia from 2001 to 2004 and financial support from the Australian Grains Research and Development Corporation (projects UQ123, CSP00068 and CSP00076).

References

Akaike H (1974) A new look at statistical model identification. IEEE Trans Autom Control AU 19:716–722

Bell C, Eagles H (2003) Breeding wheat lines for southern Australia according to three climatic mega-environments. "Solutions for a better environment. In: Proceedings of the 11th Australian Agronomy Conference. Society of Agronomy, Geelong, VIC, Australian

Braun H-J, Rajaram S, van Ginkel M (1996) CIMMYT's approach to breeding for wide adaptation. Euphytica 92:175–183

Brennan JP, Fox PN (1998) Impact of CIMMYT varieties on the genetic diversity of wheat in Australia, 1973–1993. Aust J Agric Res 49:175–178

Brennan JP, Quade KJ (2006) Evolving usage of materials from CIMMYT in developing Australian wheat varieties. Aust J Agric Res 57:947–952

Burgueño J, Crossa J, Cornelius PL, Trethowan R, McLaren G, Krishnamachari A (2007) Modeling additive × environment and additive × additive × environment using genetic covariances of relatives of wheat genotypes. Crop Sci 47:311–320

Chapman SC, Mathews KL, Trethowan R, Singh RP (2007) Height and yield relationships between reduced height near-isogenic spring wheats across world wheat production regions. Euphytica. doi:10.1007/s10681-006-9304-3

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. Theor Appl Genet 88:561–572

Cooper M, Fox PN (1996) Environmental characterization based on probe and reference genotypes. In: Cooper M, Hammer GL (eds) Plant adaptation and crop improvement. CAB International, Wallingford, UK, pp 529–547

Cooper M, Woodruff DR (1993) Predicting grain–yield in Australian environments using data from CIMMYT international wheat performance trials. 3. Testing predicted correlated response to selection. Field Crop Res 35:191–204

Cooper M, Delacy IH, Byth DE, Woodruff DR (1993) Predicting grain-yield in Australian environments using data from CIMMYT international wheat performance trials. 2. The application of classification to identify environmental relationships which exploit correlated response to selection. Field Crops Res 32:323–342

Cooper M, DeLacy IH, Basford KE (1996) Relationships among analytical methods used to analyse genotypic adaptation in multi-environment trials. In: Cooper M, Hammer GL (eds) Plant adaptation and crop improvement. CAB International, Wallingford, pp 193–224

Crossa J, Burgueño J, Cornelius PL, McLaren G, Trethowan R, Krishnamachari A (2006) Modeling genotype × environment interaction using additive genetic covariances of relatives for predicting breeding values of wheat genotypes. Crop Sci 46:1722–1733

Crossa J, Cornelius PL (2002) Linear-bilinear models for the analysis of genotype-environment interaction. In: Kang MS (eds) Quantitative genetics, genomics and plant breeding. CAB International, New York, pp 305–322

Cullis BR, Thomson FM, Fisher JA, Gilmour AR, Thompson R (1996) The analysis of the NSW wheat variety database. 2. Variance component estimation. Theor Appl Genet 92:28–39

Cullis BR, Smith AB, Panozzo JF, Lim P (2003) Barley malting quality: are we selecting the best? Aust J Agric Res 54:1261–1275

Cullis BR, Smith AB, Coombes NE (2006) On the design of early generation variety trials with correlated data. J Agric Biol Environ Stat 11:381–393

de la Vega AJ, Chapman SC (2006) Multivariate analyses to display interactions between environment and general or specific combining ability in hybrid crops. Crop Sci 46:957–967

DeLacy IH, Basford KE, Cooper M, Bull JK, McLaren CG (1996) Analysis of multi-environment trials: an historical perspective. In: Cooper M, Hammer GL (eds) Plant adaptation and crop improvement. CAB International, Wallingford, UK, pp 39–124

Doyle AD, McLeod RW, Wong PTW, Hetherington SE, Southwell RJ (1987) Evidence for the involvement of the root lesion nematode *Pratylenchus thornei* in wheat yield decline in northern New South Wales. Aust J Exp Agric 27:563–570

Eeuwijk FAv, Cooper M, DeLacy IH, Ceccarelli S, Grando S (2001) Some vocabulary and grammar for the analysis of multi-environment trials, as applied to the analysis of FPB and PPB trials. Euphytica 122:477–490

Frensham AB, Barr AR, Cullis BR, Pelham SD (1998) A mixed model analysis of 10 years of oat evaluation data: use of agronomic information to explain genotype by environment interaction. Euphytica 99:43–56

Gilmour AR, Cullis BR, Verbyla AP (1997) Accounting for natural and extraneous variation in the analysis of field experiments. J Agric Biol Environ Stat 2:269–293

Gilmour AR, Gogel BJ, Cullis BR, Thompson R (2006) ASReml User Guide Release 2.0 VSN International Ltd, Hemel Hempstead, HP1 1ES, UK

Henderson CR (1976) A simple method for computing the inverse of a numerator relationship matrix used in the prediction of breeding values. Biometrics 32:69–83

Lillemo M, van Ginkel M, Trethowan RM, Hernandez E, Rajaram S (2004) Associations among international CIMMYT bread wheat yield testing locations in high rainfall areas and their implications for wheat breeding. Crop Sci 44:1163–1169

Mathews KL, Chapman SC, Trethowan R, Singh RP, Crossa J, Pfeiffer W, van Ginkel M, DeLacy I (2006) Global adaptation of spring bread and durum wheat lines near-isogenic for major reduced height genes. Crop Sci 46:603–613

McLaren CG, Ramos L, Lopez C, Eusebio W (2004) Applications of the genealogy management system. ICIS Technical Manual

Nicol JM, Ortiz-Monasterio I (2000) Cultivar susceptibility to the root-lesion nematode *Pratylenchus thornei* and wheat yield loss in the state of Sonora, Mexico, and wider implications. Annual Wheat Newsletter, CIMMYT

O'Brien L, Morrell M, Wrigley C, Appels R (2001) Genetic pool of Australian wheats. In: Bonjean AP, William JA (eds) The world wheat book: a history of wheat breeding. Intercept Ltd., London, pp 611–648

Oakey H, Verbyla A, Pitchford W, Cullis B, Kuchel H (2006) Joint modeling of additive and non-additive genetic line effects in single field trials. Theor Appl Genet 113:809–819

Oakey H, Verbyla A, Cullis B, Wei X, Pitchford W (2007) Joint modelling of additive, dominance and epistatic (genetic line) effects in multi-environment trials. Theor Appl Genet 114:1319–1332

Ortiz R, Mowbray D (2007) Dedication: Norman E. Borlaug The humanitarian plant scientist who changed the world. Plant Breed Rev 28:1–37

Parker GD, Fox PN, Langridge P, Chalmers K, Whan B, Ganter PF (2002) Genetic diversity within Australian wheat breeding programs based on molecular and pedigree data. Euphytica 124:293–306

Peake A (2003) Inheritance of grain yield and effect of the 1BL/1RS translocation in three bi-parental wheat (*Triticum aestivum* L.) populations in production environments of north eastern Australia. School of Land and Food Sciences. The University of Queensland, Brisbane

Peterson CJ, Pfeiffer WH (1989) International winter wheat evaluation: relationships among test sites based on cultivar performance. Crop Sci 29:276–282

Piepho H-P (1998) Empirical best linear unbiased prediction in cultivar trials using factor-analytic variance–covariance structures. Theor Appl Genet 97:195–201

Piepho HP (1997) Analyzing genotype-environment data by mixed models with multiplicative terms. Biometrics 53:761–766

Piepho HP, Möhring J, Melchinger AE, Büchse A (2007) BLUP for phenotypic selection in plant breeding and variety testing. Euphyica. doi:10.1007/s10681-007-9449-8

R Development Core Team (2007) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.http://www.R-project.org

Rajaram S, Van Ginkel M (2001) Mexico, 50 years of international wheat breeding. In: Bonjean AP, William JA (eds) The world wheat book: a history of wheat breeding. Intercept Ltd, London, pp 579–608

Rajaram SM, van Ginkel M, Fischer RA (1994) CIMMYT's wheat breeding mega-environments (ME). In: Proceedings of the 8th International Wheat Genetics Symposium, Beijing, China, pp 1101–1106

Riffkin PA, Evans PM, Chin JF, Kearney GA (2003) Early-maturing spring wheat outperforms late-maturing winter wheat in the high rainfall environment of south-western Victoria. Aust J Agric Res 54:193–202

Sadras V, Baldock J, Roget D, Rodriguez D (2003) Measuring and modelling yield and water budget components of wheat crops in coarse-textured soils with chemical constraints. Field Crops Res 84:241–260

Sivapalan S, O'Brien L, Ortiz-Ferrara G, Hollamby GJ, Barclay I, Martin PJ (2003) A comparative study for yield performance and adaptation of some Australian and CIMMYT/ICARDA wheat genotypes grown at selected locations in Australia and the WANA region. Aust J Agric Res 54:91–100

Smith A, Cullis B, Gilmour A (2001a) The analysis of crop variety evaluation data in Australia. Aust N Z J Stat 43:129–145

Smith A, Cullis B, Thompson R (2001b) Analyzing variety by environment data using multiplicative mixed models and adjustments for spatial field trend. Biometrics 57:1138–1147

St Martin SK (1982) Effective population–size for the soybean improvement program in maturity groups 00 to IV. Crop Sci 22:151–152

Stephens DJ, Lyons TJ (1998a) Rainfall-yield relationships across the Australian wheatbelt. Aust J Agric Res 49:211–224

Stephens DJ, Lyons TJ (1998b) Variability and trends in sowing dates across the Australian wheatbelt. Aust J Agric Res 49:1111–1118

Trethowan RM, Crossa J, van Ginkel M, Rajaram S (2001) Relationships among bread wheat international yield testing locations in dry areas. Crop Sci 41:1461–1469

Trethowan RM, Hodson D, Braun HJ, Pfeiffer WH (2005) Wheat breeding environments. In: J. Dubin MAL, Morris ML (ed) Impacts of International Wheat Breeding Research in the Developing World, 1988–2002. CIMMYT, Mexico, D.F., pp 4–11

Yan W, Hunt LA (2002) Biplot analysis of diallel data. Crop Sci 42:21–30