Published July 28, 2016

RESEARCH

Molecular Marker Information in the Analysis of Multi-Environment Trials Helps Differentiate Superior Genotypes from Promising Parents

M. Gabriela Borgognone,* David G. Butler, Francis C. Ogbonnaya, and M. Fernanda Dreccer

ABSTRACT

The statistical analysis of multi-environment trial data aims to provide reliable and accurate predictions of genotype performance across the target environments and information on specific performance from the interaction of genotypes with the environments. Genetic gain can be achieved faster when selections are based on predictions from a model that accounts for the relationships among genotypes rather than from a model that assumes unrelated genotypes. Yield and plant height data from 37 international wheat trials were analyzed using a linear mixed model that accounted for relationships among the genotypes via a genomic relationship matrix $\mathcal{G}$ derived from 2487 polymorphic DArT molecular markers for 197 genotypes. The elements of this matrix reflect the actual proportion of the parts of the genome surveyed that is identical by state between pairs of individuals, and including it into the model resulted in generally lower average prediction error variances of individual trials in the analyses. Partitioning the total genetic effects into additive and residual non-additive genetic effects has familiar interpretations for plant breeders and facilitates exploring genotype by environment interactions for additive and total effects. This interpretation is still possible with the form of $\mathcal{G}$ used in this paper. This method of analysis could be readily implemented to accelerate genetic gain by plant breeding programs that have molecular markers for the genotypes under study.

M.G. Borgognone and D.G. Butler, Dep. of Agriculture and Fisheries, Leslie Research Facility, 13 Holberton Street, Toowoomba, QLD 4350, Australia; F.C. Ogbonnaya, International Center for Agricultural Research in the Dry Areas (ICARDA), Beirut, Lebanon, and Grains Research and Development Corporation, Level 4 East Building, 4 National Circuit, Barton, ACT 2600, Australia; M.F. Dreccer, CSIRO Agriculture, Cooper Laboratory, PO Box 863, Univ. of Queensland, Warrego Highway, Gatton, QLD 4343, Australia. Received 7 Mar. 2016. Accepted 10 June 2016. Assigned to Associate Editor Jesse Poland. *Corresponding author (Gabriela.Borgognone@daf.qld.gov.au).

Abbreviations: **A**, additive relationship matrix; E-BLUPs, Empirical Best Linear Unbiased Predictions; EM, Expectation–Maximization; FA, factor analytic; $\mathcal{G}$, genomic relationship matrix; G × E, genotype by environment interaction; IRR, irrigated; MET, multi-environment trials; PEV, prediction error variance.

Plant breeding programs aim to develop new crop varieties with improved characteristics over existing ones available to farmers. Appropriate methods of selection are crucial to successfully identify the best possible parents for crossing, and to choose the best performing genotypes to progress in the program. Both objectives generally involve evaluating the genotypes under consideration in series of designed experiments across several locations and years.

Multi-environment trials (MET) are series of field experiments where genotypes grow under different conditions (environments) determined by combinations of year, location and sometimes management factors (Smith et al., 2001). The statistical analysis of MET data aims to provide reliable and accurate predictions of genotype performance across the target environments and information on specific performance from the interaction of genotypes with the environments (G × E) (Smith et al., 2001).

Published in Crop Sci. 56:2612–2628 (2016).
doi: 10.2135/cropsci2016.03.0151

Linear mixed models are widely used for the analysis of MET data. The model described by Smith et al. (2001) accounts for within trial design effects and spatial variability, between trial residual variance heterogeneity, and adopts a multiplicative factor analytic (FA) model for the G × E effects. The FA model is preferred in the MET setting because it allows for genetic variance heterogeneity between trials, different genetic correlation across trials and can have greater predictive accuracy than alternative models when there is substantial G × E (Kelly et al., 2007). This improved predictive accuracy positively impacts selection gains (Kelly et al., 2007). It is commonly assumed that genotype effects are independent, however, in breeding programs individuals are likely to be related.

The FA model may be extended to include a relationship matrix reflecting the expected degree of co-ancestry among the genotypes under study (Oakey et al., 2007). This relationship matrix can be derived from the pedigrees of the individuals (Henderson, 1976) and is usually referred to as the numerator or additive relationship matrix, denoted by **A**. The elements in **A** reflect the predicted proportion of the genome that is identical by descent between two individuals (Hayes et al., 2009). This is a more realistic representation of the relationship among individuals than assuming them to be independent.

Additive genetic effects (or breeding values) are useful when selecting parents for crossing, while non-additive effects can be responsible for specific adaptation (Oakey et al., 2006). For the additive component, the **A** matrix models the covariance between relatives, and the identity matrix is used for the non-additive component (Oakey et al., 2006). The total genetic effects (i.e., the sum of the additive and non-additive genetic effects) predict the overall performance of genotypes in the target environments and are used in selecting genotypes for promotion in the breeding program and eventual release to industry as varieties. However, all selections made under the FA model that assumes unrelated genotypes are based on the total genotypic effects.

A limitation of the **A** matrix is that it is based on expected relationships between individuals given their pedigree. However, in plant populations where selection of genotypes is undertaken over several generations, relationships between individuals can be far from what is expected. For example, full siblings could have a greater proportion of the genome in common than expected and this proportion could also be greater with one parent than the other (Oakey et al., 2006). Additionally, pedigrees based on crossing history are not always known, are uncertain or incomplete.

A relationship matrix for the genotypes under study can also be derived from their molecular marker information. The elements of this matrix reflect the proportion of the parts of the genome surveyed that is identical by state between pairs of individuals. Genotypes identical by state share the same genetic region, although it cannot be assumed that they inherited this region from a common ancestor (Jordan et al., 2005). This matrix is often referred to as the realized or genomic relationship matrix, denoted by $\boldsymbol{G}$.

Several methods for obtaining $\boldsymbol{G}$ have been reported (see VanRaden, 2008 for three examples). The first method proposed by VanRaden (2008) derives a zero-centered and scaled matrix to calculate the genomic relationships. This form of the $\boldsymbol{G}$ matrix is analogous to the additive relationship matrix **A** and can be used instead of **A** in the linear mixed model formulation. The advantage of using this form of $\boldsymbol{G}$ is that the genetic effects can still be partitioned into additive and non-additive components, facilitating the interpretation of results and the selection process for plant breeders.

Most of the MET data analyses using an FA model for parent and varietal selection reported in the literature assume independent genotype effects (e.g., Smith et al., 2007; Dreccer et al., 2007; Stefanova and Buirchell, 2010), while some assume genotypes related through the **A** matrix (e.g., Beeck et al., 2010; Hardner et al., 2012). MET models for genomic selection incorporate some form of the genetic relationship matrix. The objective of these models is to achieve accuracy in the prediction of genetic values. For example, both Burgueño et al. (2012) and Dawson et al. (2013) fitted models for G × E using an FA model that incorporated a $\boldsymbol{G}$ matrix. Their goal was to obtain genomic predictions, that is, to predict the performance of newly developed genotypes for which phenotypic data is not available and to predict the performance of genotypes across environments when not all genotypes are evaluated in all environments (Burgueño et al., 2012). There seems to be no applications that use the FA model for the analysis of MET data where correlated genotype effects are represented by the matrix $\boldsymbol{G}$ that, with the objective of making selections, differentiate between total and additive genetic effects, and thoroughly explore the G × E patterns. The aim of the present paper is to show, through the analysis of an international wheat MET dataset, how to incorporate the $\boldsymbol{G}$ matrix in the FA model and how to interpret the results of the analysis to make more robust selection decisions. The phenotypic traits *yield* and *plant height* were chosen for illustration due to their contrast in heritability and genetic complexity as well as their agronomic and commercial importance. This method of analysis could be readily implemented by plant breeding programs that have access to molecular markers for the genotypes under study.

MATERIALS AND METHODS

Trials and Traits

The MET dataset consists of 37 international wheat trials grown across 4 yr (2009–2012) in 17 locations spread across six countries (Australia, Ethiopia, Lebanon, Mexico, Morocco, and Syria). The trials were part of a Generation Challenge Program project (http://www.generationcp.org/) where locations were chosen as representative of different drought mega-environments characterized by summer, winter, or uniform rainfall distribution following Hodson and White (2007) and Dreccer et al. (2013). Most trials were planted at the conventional sowing time for the location and under rainfed conditions. For contrast, four trials were planted late and 10 trials were irrigated. Trial codes are composed of the last two digits of the year, two letters to identify the country, two or three letters for the location, and a suffix (IRR or Late) to indicate if the trial was irrigated and/or had late sowing (Table 1).

Composition and physical layout of 37 candidate trials for the MET analyses are summarized in Table 1 along with the trial means for yield (t ha^{-1}) and plant height (cm). One trial was excluded from the analysis of yield (09AU_Gat_IRR_Late). This was a late sown trial, with a short growing season that resulted in a narrow range of flowering dates and similar yield expression, and therefore its estimated genetic variance was zero. Five trials were excluded from the analysis of plant height because they had no data for this trait.

Table 1. Summary of the 37 candidate trials for multi-environment trials (MET) analyses of yield and plant height. Trials with no mean for a trait were excluded from the MET analysis for that specific trait.

Year	Country	Location	Management†	Trial code‡	Rows	Columns	Replicates	Genotypes	Mean yield	Mean height
							No.		t ha^{-1}	cm
2009	Australia	Gatton	Conv., rainfed	09AU_Gat	18	12	3	72	5.2	86
2009	Australia	Gatton	Conv., irrigated	09AU_Gat_IRR	18	8	2	72	6.3	97
2009	Australia	Gatton	Late, irrigated	09AU_Gat_IRR_Late	18	8	2	72		82
2009	Ethiopia	Dera	Conv., rainfed	09ET_De	30	20	3	200	0.4	41
2009	Ethiopia	Melkassa	Conv., rainfed	09ET_Mel	30	20	3	200	0.8	59
2009	Mexico	CIANO	Conv., irrigated	09MX_CIA_IRR	32	3	2	47	4.1	91
2010	Australia	Gatton	Conv., rainfed	10AU_Gat	16	18	2	131	5.9	94
2010	Australia	Gatton	Conv., irrigated	10AU_Gat_IRR	16	18	2	128	5.9	92
2010	Australia	Gatton	Late, irrigated	10AU_Gat_IRR_Late	16	9	2	72	5.0	75
2010	Australia	Leeton	Conv., rainfed	10AU_Lee	16	9	2	72	3.7	
2010	Australia	Leeton	Conv., irrigated	10AU_Lee_IRR	16	9	2	72	4.0	75
2010	Australia	Temora	Conv., rainfed	10AU_Tem	16	9	2	72	5.4	
2010	Ethiopia	Dera	Conv., rainfed	10ET_De	10	40	2	200	2.3	75
2010	Ethiopia	Kulumsa	Conv., rainfed	10ET_Kul	10	40	2	200	3.6	87
2010	Ethiopia	Melkassa	Conv., rainfed	10ET_Mel	10	40	2	200	2.3	89
2010	Lebanon	Terbol	Conv., rainfed	10LE_Ter	13	31	2	200	3.1	90
2010	Morocco	Sidi el Aidi	Conv., rainfed	10MO_SEA	20	30	3	200	1.8	86
2010	Morocco	Sidi el Aidi	Conv., irrigated	10MO_SEA_IRR	20	30	3	200	1.9	87
2010	Mexico	CIANO	Conv., rainfed	10MX_CIA	16	12	4	47	3.0	73
2010	Mexico	CIANO	Conv., irrigated	10MX_CIA_IRR	12	8	2	47	6.5	101
2010	Syria	Malkiya	Conv., rainfed	10SY_Mal	12	50	3	200	2.0	
2010	Syria	Tel-Hadya	Conv., rainfed	10SY_THa	50	12	3	200	3.0	82
2010	Syria	Tel Tair	Conv., rainfed	10SY_TTa	12	50	3	200	0.4	
2011	Australia	Temora	Conv., rainfed	11AU_Tem	16	8	2	64	3.9	90
2011	Australia	Yanco	Conv., rainfed	11AU_Yan	16	8	2	64	5.5	99
2011	Lebanon	Kfr-Dhan	Conv., rainfed	11LE_KDa	9	44	2	200	4.0	99
2011	Lebanon	Terbol	Conv., rainfed	11LE_Ter	9	44	2	200	4.9	104
2011	Morocco	Marchouch	Conv., rainfed	11MO_Mar	20	20	2	200	3.2	
2011	Morocco	Sidi el Aidi	Conv., rainfed	11MO_SEA	20	20	2	200	4.5	98
2011	Morocco	Sidi el Aidi	Conv., irrigated	11MO_SEA_IRR	20	20	2	200	4.2	98
2011	Mexico	CIANO	Conv., rainfed	11MX_CIA	8	12	2	47	1.2	57
2011	Mexico	CIANO	Conv., irrigated	11MX_CIA_IRR	12	8	2	47	5.5	91
2011	Syria	Breda	Conv., rainfed	11SY_Bre	33	12	2	200	2.8	57
2011	Syria	Malkiya	Conv., rainfed	11SY_Mal	23	18	2	200	4.9	93
2011	Syria	Tel-Hadya	Conv., rainfed	11SY_THa	40	10	2	200	3.3	73
2012	Morocco	Sidi el Aidi	Late, rainfed	12MO_SEA_Late	20	20	2	200	2.2	77
2012	Morocco	Zemamra	Late, rainfed	12MO_Zem_Late	20	20	2	200	2.5	

† Conventional vs. late time of sowing for the location.

‡ Trial code acronym composed of two digits for the year, two letters for the country, two or three letters for the location, and the suffix _IRR to indicate if the trial was irrigated or _Late for late sown. Most trials were planted at the conventional sowing time for the location, under rainfed conditions (not indicated in the trial code).

Germplasm

A total of 243 cultivars, recombinant inbred lines, and experimental lines (collectively referred to as 'genotypes' in this paper) were included in the MET. From these, 211 were elite lines from drought nurseries from the International Center for Agricultural Research in the Dry Areas (ICARDA) and local checks, and 32 were lines from different origin contrasting for traits putatively involved in drought tolerance. The genotypes spanned a range of maturity groups. Genotype concurrence among trials was generally robust, with pairs of trials having between 46 and 200 genotypes in common. Two exceptions are the 2011 Australian trials (11AU_Tem and 11AU_Yan, Table 1). These trials have low concurrences with 12 other trials: four genotypes in common with all 2009 Australian and all Mexican trials, seven with the 2010 Australian trials in the same locations, and eight with 10AU_Gat_IRR_Late.

Genotyping

The germplasm panel was planted in the greenhouse. Young leaves were harvested from five plants per genotype, frozen in liquid nitrogen and stored at −80°C prior to DNA extraction. Genomic DNA was extracted according to Ogbonnaya et al. (2001). A 100 ng μL^{-1} DNA from each sample was used for whole-genome profiling using DArT markers by Triticarte Pty Ltd, Yarralumla, Australia (http://www.triticarte.com.au/). A final number of 2487 polymorphic markers with a quality parameter and a call rate greater than 80% and minor allele frequency >5% was retained from the 3051 initial markers. Of the 2487 polymorphic DArT markers, 2007 markers were of known map locations and the marker order and position information were obtained from the integrated consensus DArT genetic map of wheat (Detering et al., 2010). The marker data had low proportions of missing values: 90 markers had >10% missing values and only 5 markers had between 20 and 35% missing values. DArT markers were available for 197 of the 243 genotypes. The lines not genotyped were the sets contrasting for traits and some of the ICARDA lines that were discontinued after assessment during the initial evaluation and/or multiplication.

Linear Mixed Model

A MET dataset consists of trait data collected for each plot in t trials, where m genotypes have been grown (not necessarily all genotypes in all trials). The n_i plots for the ith trial are assumed laid out in a rectangular array of r_i rows by c_i columns, so that $n_i = r_i \times c_i$, for $i = 1, \ldots, t$. The model for a univariate analysis of the $N = \sum_{i=1}^{t} n_i$ response vector $\mathbf{y}$ is

$$\mathbf{y} = \mathbf{X}\tau + \mathbf{Z}_g\mathbf{u}_g + \mathbf{Z}_p\mathbf{u}_p + \mathbf{e}$$

where τ is a vector of fixed effects with associated design matrix $\mathbf{X}$, $\mathbf{u}_g$ is a $mt \times 1$ vector of random genetic effects for each genotype in each trial with associated design matrix $\mathbf{Z}_g$, $\mathbf{u}_p$ is a vector of random non-genetic (or peripheral) effects with associated design matrix $\mathbf{Z}_p$, and $\mathbf{e}$ is the vector of combined residuals for individual trials. In the simplest case the vector τ comprises an overall mean (intercept) for the trials but may include other effects such as specific trend terms. The vector $\mathbf{u}_p$ includes non-genetic factors such as trial-specific blocking and any other design factors. All trials used randomized complete block designs with plots arranged in a rectangular grid indexed by column and row. The preferred approach was to include a random block term in the model for each trial to account for this design factor. Some trials showed global trends and required linear row or linear column effects, and others required random column or row effects due to the presence of extraneous variation. Relatively few trials needed a smoothing spline term (Verbyla et al., 1999) along either the row or column dimension. The column and row correlations of the stationary spatial process were modelled by an autoregressive model of first order. For details on this approach regarding the analysis of individual trials, see Gilmour et al. (1997) and Stefanova et al. (2009).

The vectors of random effects $\mathbf{u}_g$, $\mathbf{u}_p$, and $\mathbf{e}$ are assumed pairwise independent and distributed as Gaussian with zero means. The variance matrix of $\mathbf{u}_p$ is given by $\mathbf{G}_p = \oplus_{l=1}^{b}\mathbf{G}_{pl}$, where b is the number of sub-vectors in $\mathbf{u}_p$, which are assumed mutually independent, and $\oplus$ is the direct sum operator. Each $\mathbf{G}_{pl}$ is assumed to be a scaled identity matrix. The variance matrix of $\mathbf{e}$ is given by $\mathbf{R} = \oplus_{i=1}^{t}\mathbf{R}_i$, a block-diagonal matrix with errors from individual trials, which are assumed mutually independent. In the general case, each $\mathbf{R}_i$ is assumed to be a function of some vector of variance parameters Φ_i, that is $\mathrm{var}(\mathbf{e}_i) = \mathbf{R}_i(\Phi_i)$. In agricultural field experiments, a separable first-order autoregressive process provides a plausible model for the residuals (Gilmour et al., 1997). In this case, $\Phi_i = (\Phi_{ic}, \Phi_{ir}, \sigma_i^2)'$ and $\mathbf{R}_i(\Phi_i) = \sigma_i^2\mathbf{R}_{ic}(\Phi_{ic}) \otimes \mathbf{R}_{ir}(\Phi_{ir})$ where $\mathbf{R}_{ic}$ and $\mathbf{R}_{ir}$ are the (parameterised) correlation matrices for the column and row dimensions, respectively, of the ith trial, and $\otimes$ is the Kronecker product.

The model for $\mathbf{u}_g$ considered in this paper is $\mathbf{u}_g = \mathbf{u}_a + \mathbf{u}_e$ where $\mathbf{u}_a$ and $\mathbf{u}_e$ represent the additive and non-additive (or residual) genetic effects, respectively; $\mathbf{u}_a$ and $\mathbf{u}_e$ are assumed independent. In hybrid crops, non-additive effects may reflect both dominance and epistasis, while in self-pollinated or inbred crops such as wheat, non-additive effects likely reflect epistatic interactions because inbreeding will mostly eliminate dominance effects (Oakey et al., 2006).

Assuming that G × E effects are represented by a simple two-way structure and that the variance matrix for genotype effects in individual trials has a separable form, then $\mathrm{var}(\mathbf{u}_g) = \mathbf{G}_t \otimes \mathbf{G}_v$. The $t \times t$ positive definite variance matrix $\mathbf{G}_t$ has diagonal elements representing the genetic variances for individual trials (environments) and off-diagonal elements representing genetic covariances between pairs of environments; $\mathbf{G}_v$ reflects the relationship between genotypes and can take different forms. Smith et al. (2001) assumed unrelated genotypes and therefore $\mathbf{G}_v = \mathbf{I}_m$. Oakey et al. (2006, 2007) assumed genotypes related according to the additive relationship matrix $\mathbf{A}$ and therefore $\mathrm{var}(\mathbf{u}_g) = \mathbf{G}_a \otimes \mathbf{A} + \mathbf{G}_e \otimes \mathbf{I}_m$, where $\mathbf{G}_a$ and $\mathbf{G}_e$ are the additive and non-additive genetic variance matrices across trials.

In this paper $\mathbf{A}$ is replaced by $\boldsymbol{\mathcal{G}}$, a genomic relationship matrix calculated from molecular markers for the genotypes as proposed by VanRaden (2008) which ensures that interpretation of the predicted genetic effects remains unchanged. The method to obtain the genomic relationship matrix is explained in detail in the following section.

The FA model based on h factors, denoted FAh, applied to the genotype effects in each trial is $\mathbf{u}_g = (\boldsymbol{\Lambda} \otimes \mathbf{I}_m)\boldsymbol{f} + \boldsymbol{\delta}$, where $\boldsymbol{\Lambda}$ is a matrix of environmental loadings, $\boldsymbol{f}$ contains scores for each genotype, and $\boldsymbol{\delta}$ is the vector of residuals (or lack of fit) for the model (Smith et al., 2001). Consequently, var($\mathbf{u}_j$) is $\mathbf{G}_j = \boldsymbol{\Lambda}_j\boldsymbol{\Lambda}_j' + \boldsymbol{\Psi}_j$, with j = a, e. In this context, $\boldsymbol{\Lambda}_j\boldsymbol{\Lambda}_j'$ is referred to as the regression part of the model and $\boldsymbol{\Psi}_j$ is referred to as the lack of fit of the model and is a diagonal matrix with specific variances for each trial, reflecting the specific G × E from the trials not accounted for by the regression part of the model.

Not all genotypes in all trials had marker information. Additional independent trial-specific model terms (having simple variance components) were included for those trials where genotypes without marker information were present. Relationships among them, if any, were ignored.

Genomic Relationship Matrix Based on Molecular Markers

The genomic relationship matrix was constructed based on the molecular markers of the 197 genotypes, which were arranged in a matrix of genotypes by markers, $\mathbf{M}^{197\times2487}$. The elements of $\mathbf{M}$ are 0 (indicating absence of the dominant marker), 1 (indicating presence of the dominant marker, either heterozygous or homozygous), or a missing value in cases where the marker score could not be determined. The calculation of the $\boldsymbol{\mathcal{G}}$ matrix used in this paper involves two steps: the first one centers the matrix of allele effects around zero and the second one scales the matrix so it becomes analogous in interpretation to the additive relationship matrix $\mathbf{A}$. The formula for $\boldsymbol{\mathcal{G}}$ below is that proposed by VanRaden (2008) for an unselected, outbred population; Endelman and Jannink (2012) derived the same formula from an identical-by-state approach, justifying its use with any population when the number of markers is much larger than the number of genotypes. Let the matrix $\mathbf{P}^{197\times2487}$ contain elements $2(p_k - 0.5)$ in the kth column, where p_k is the frequency of the dominant allele at marker k, that is, the sum of 1's in marker k divided by the number of non-missing values for that marker. Then, $\boldsymbol{\Pi} = \mathbf{M} - \mathbf{P}$ is the zero-centered matrix of allele effects. A divisor equal to twice the sum of the variances of the markers scales $\boldsymbol{\Pi}\boldsymbol{\Pi}'$, so that

$$\boldsymbol{\mathcal{G}} = \frac{\boldsymbol{\Pi}\boldsymbol{\Pi}'}{2\sum p_k(1-p_k)}$$

The $\boldsymbol{\mathcal{G}}$ matrix was obtained from the *A.mat* function in the R (R Core Team, 2014) package rrBLUP (Endelman, 2011). This function provides two methods for imputation of missing marker values: using the marker mean and using the Expectation–Maximization (EM) algorithm based on the multivariate normal distribution, which is a useful approximation in the context of predicting breeding values with a genomic relationship matrix (Poland et al., 2012). In a simulation study, Poland et al. (2012) found that imputing with the marker mean had higher error than using the EM algorithm but no significant differences with respect to genomic prediction accuracy were found. However, when using the mean for imputation significant bias in the estimation of breeding values was found (there was a trend to underestimate the true breeding values) but no significant bias was detected when using EM algorithm. For these reasons we chose the approach based on the EM algorithm.

The marker scores must be converted to {–1, 0, 1} to be used in rrBLUP. The original 1's in $\mathbf{M}$ were preserved, indicating presence of the dominant marker, and the original 0's were changed to –1's, indicating absence of the dominant marker. No score was changed to 0 in this context.

Software and Estimation

Variance parameters in the linear mixed model were estimated by residual maximum likelihood (REML) (Patterson and Thompson, 1971). Empirical Best Linear Unbiased Estimates (E-BLUEs) for fixed effects, and Empirical Best Linear Unbiased Predictions (E-BLUPs) for random effects were obtained using ASReml 4 (Gilmour et al., 2015). Commonly, genomic relationship matrices based on molecular markers are not positive definite and cannot be inverted. Note that even though the zero-centered $\boldsymbol{\mathcal{G}}$ matrix is singular and therefore non-invertible, ASReml 4 can use it when fitting the linear mixed model. This feature of ASReml 4 was the main reason to choose this software. As an example, the code for the last model fitted to plant height is given in the Appendix.

Post-Processing of Results

Heatmaps and cluster analysis were applied to the additive and total genetic correlation matrices to identify environment groups where genotypes performed similarly (Cullis et al., 2010).

Cluster analysis was performed using the *agnes* agglomerative hierarchical algorithm from the package cluster in R (Maechler et al., 2014), as described by Cullis et al. (2010). The clustering algorithm uses average linkage and was applied to the dissimilarity matrix for the genetic effects, obtained by subtracting the estimated genetic correlation matrix from $\mathbf{1}_{197}$. Clusters for additive and total genetic effects were determined using 0.5 as the cut-off point on the dissimilarity scale of the dendrogram to ensure that the average genetic correlation among trials in any given cluster is at least 0.5. Each cluster represented a different environment group.

The interpretation of the environment groups generated by cluster analysis when the $\boldsymbol{\mathcal{G}}$ matrix is included in the MET analysis contains an extra layer of useful information because both additive and total genetic effects can be investigated. Environment groups for total genetic effects could differ from those for additive genetic effects. For low-heritability traits such as yield, trials tend to have a reasonable proportion of non-additive genetic variance. When this non-additive genetic variance is well accounted for by the model, trials may belong to different environment groups for additive and total genetic effects. Thus, the best candidate lines for release to industry may differ from the best candidate parents for future crosses. We propose a novel way for the interpretation of the environment groups through two scenarios introduced below.

In Scenario I, trials belong to the same environment group for both additive and total genetic effects when most of the pattern in the genotype rankings is due to additive effects. This happens when the proportion of additive genetic variance in the trials is high, well explained by the model, and trials have either (i) a low proportion of non-additive genetic variance or (ii) a higher proportion of non-additive genetic variance but

poorly explained by the model. Best potential parents and best candidates for release to industry will be in better agreement in type (i) than in type (ii) environment groups.

In Scenario II, trials belong to different environment groups for additive and total genetic effects when non-additive effects change the pattern in the genotype rankings. This happens when the proportion of additive genetic variance in the trials is moderate, relatively well explained by the model, and trials also have a moderate proportion of non-additive variance well explained by the model. Consequently, the top yielding genotypes and best potential parents will generally not overlap.

Models Fitted and Selection Indices

The FA models were fitted in a (nested) hierarchical sequence and the log-likelihood ratio test was used to assess the significance of the improvement in fit after each higher-order model was fitted. The initial model assumed no covariance between trials for both additive and non-additive genetic effects. Subsequent models added one FA factor at a time for the genetic effects. The initial parameter estimates for any given FA model were derived from the preceding lower order model.

The final model for yield was the highest order FA model that could be fitted that had a significant improvement with respect to the previous model. Higher order FA models for yield failed to converge, which could indicate a flat likelihood surface in the region of the solution (Section 15.5 of the ASreml Manual, Gilmour et al., 2015). The final model for plant height was the last one with a significant ($P < 0.001$) improvement according to the log-likelihood ratio test.

An indicator of the goodness of fit of the final FA model is the percentage of genetic variance it explains out of the total genetic variance. This indicator, called the percentage of variance accounted for (%vaf), is defined as

$$\%\text{vaf} = 100[\text{diag}(\boldsymbol{\Lambda}\boldsymbol{\Lambda}')/\text{diag}(\boldsymbol{\Lambda}\boldsymbol{\Lambda}' + \boldsymbol{\Psi})]$$

where $\boldsymbol{\Lambda}$ is the matrix of environmental loadings, and $\boldsymbol{\Psi}$ is the diagonal matrix with specific variances for each trial. Additionally, the FA model that assumes independent genotype effects was also fitted to both traits.

The prediction error variance for each trial (PEV_i) is defined as

$$\text{PEV}_i = \frac{1}{m}\sum\nolimits_{j=1}^{m}\left[(\mathbf{u}_{gij} - \tilde{\mathbf{u}}_{gij}) - E(\mathbf{u}_{gij} - \tilde{\mathbf{u}}_{gij})\right]^2$$

for $i = 1, \ldots, t$, that is, the variance of the difference between the true genetic effects for each genotype and the predicted genetic effects (E-BLUPs). For comparative purposes, estimated PEVs for each trial were obtained from the fitted models as the square of the average standard error of difference.

Genotypes for the different environment groups can be selected using selection indices for the additive and total genetic effects. The simplest index, as used in this paper, gives equal weights to all trials that belong to a particular environment group. Therefore, the selection index for additive (total) genetic effects is the average of the additive (total) E-BLUPs across all trials in the environment group. For a summary of weighted and unweighted selection indices see Oakey et al. (2007).

RESULTS

Yield: Additive and Total Effects

Yield is a very complex trait and remains a major objective of most breeding programs worldwide. The MET analysis for yield included 36 of the 37 trials. The model of best fit to the data included three factors (FA3) for the additive genetic variance matrix and two factors (FA2) for the non-additive genetic variance matrix.

Additive Effects

Four environment groups were identified for the additive effects (A1.yld–A4.yld, Table 2; Fig. 1a), and are described in terms of trial mean yield and country of origin for illustration purposes. All environment groups contained the 197 genotypes.

Environment group A1.yld comprised mainly high yielding Australian trials (5–6.3 t ha^{-1}, Table 1) plus 11MO_Mar and 11SY_Mal (Table 2). All trials except for 11SY_Mal had 56 to 99% additive genetic variance, with a high percentage of the variance accounted for (%vaf) by the model (66–100%vaf, Table 2). Trial 11SY_Mal is a high yielding trial (4.9 t ha^{-1}, Table 1) with only 22% additive genetic variance and the lowest genetic correlations with all other trials in the group (Fig. 1a). Knowledge of the specific trial would determine if it should be included in this environment group. The average additive genetic correlation among trials was 0.65.

Environment group A2.yld was formed by trials from Australia, Lebanon, Mexico, and Syria with the widest range in trial mean yield of all groups (2.8–6.5 t ha^{-1}, Table 1). In general, trials in this group had lower percentages of additive genetic variance than those in A1.yld (Table 2). This group had the highest average additive genetic correlation among trials (0.72).

Environment group A3.yld included only Moroccan trials. The four Sidi el Aidi trials had >50% additive genetic variance, well explained by the model (>75%vaf, Table 2). 12MO_ZEM_Late had only 35% additive genetic variance, of which 59% was explained by the model (Table 2), and was mainly correlated with the 2010 Sidi el Aidi trials (Fig. 1a). Trial mean yields ranged 1.8–4.5 t ha^{-1} and average additive genetic correlation among trials was 0.70.

Environment group A4.yld, formed mainly by trials from Ethiopia, also included 10MX_CIA. Mean yields of trials in this group were extremely variable, with the 2009 Ethiopian trials yielding 0.3 and 0.8 t ha^{-1} and the remaining four trials yielding 2.3 to 3.6 t ha^{-1} (Table 1). The average additive genetic correlation among trials in this environment group was the lowest of all groups (0.62).

Correlations between environment groups ranged from low-positive to negative (Fig. 1a). The top 20 genotypes to be considered as potential parents generally differed between environment groups. Most pairs of groups had a maximum of four genotypes in common in

Table 2. Yield MET results. For each trial, estimated total genetic and residual variances, percentages of the total genetic variance that are additive and non-additive (these two columns add to 100%), percentages of additive and non-additive genetic variances explained by the regression parts of the final model (FA3 for additive effects and FA2 for the non-additive effects), and environment group to which each trial belongs for additive and total effects.

Trial code†	Total genetic variance	Residual variance	Additive genetic variance	Non-additive genetic variance	%vaf‡ FA3 additive	%vaf‡ FA2 non-additive	Environment group§ (*n*)	
							Additive effects (A*n*.yld)	Total effects (T*n*.yld)
			%					
10AU_Gat	0.730	0.560	99	1	84	100	1	1
10AU_Gat_IRR	0.824	0.685	97	3	80	100	1	1
10AU_Tem	0.265	0.244	79	21	80	100	1	1
09AU_Gat_IRR	0.196	0.449	75	25	100	23	1	1
10AU_Gat_IRR_Late	0.418	0.433	56	44	100	11	1	1
11MO_Mar	0.285	0.495	60	40	66	100	1	2
11SY_Mal	0.249	0.620	22	78	88	48	1	2
10MX_CIA_IRR	0.738	0.298	53	47	100	100	2	2
11MX_CIA_IRR	0.550	0.231	53	47	85	89	2	2
11LE_KDa	0.106	0.176	99	1	67	100	2	4
09AU_Gat	0.177	0.198	75	25	100	100	2	4
11AU_Yan	0.088	0.161	57	43	69	32	2	4
10SY_THa	0.017	0.056	57	43	93	30	2	4
10AU_Lee	0.087	0.104	53	47	92	55	2	4
10AU_Lee_IRR	0.083	0.097	33	67	100	36	2	4
11SY_Bre	0.021	0.052	32	68	68	37	2	4
09MX_CIA_IRR	0.290	0.106	93	7	93	100	2	6
11LE_Ter	0.146	0.369	53	47	57	48	2	–
10MO_SEA_IRR	0.137	0.416	97	3	100	100	3	3
10MO_SEA	0.093	0.462	84	16	100	100	3	3
11MO_SEA_IRR	0.555	1.288	68	32	87	19	3	3
11MO_SEA	0.397	1.152	50	50	79	23	3	3
12MO_Zem_Late	0.255	0.917	35	65	59	21	3	–
09ET_De	0.012	0.019	70	30	93	34	4	5
09ET_Mel	0.037	0.083	77	23	79	100	4	5
10ET_Mel	0.219	0.179	71	29	82	21	4	6
10MX_CIA	0.239	0.078	86	14	56	100	4	6
10ET_De	0.066	0.250	90	10	85	100	4	6
10ET_Kul	0.225	0.305	93	7	74	100	4	6
10LE_Ter	0.142	0.219	39	61	33	45	–	–
10SY_Mal	0.389	0.197	53	47	23	6	–	–
10SY_TTa	0.002	0.014	58	42	46	100	–	–
11AU_Tem	0.081	0.272	54	46	30	48	–	–
11MX_CIA	0.078	0.053	99	1	32	100	–	–
11SY_THa	0.042	0.053	66	34	36	9	–	–
12MO_SEA_Late	0.264	0.376	38	62	18	27	–	–

† Trial code acronym composed of two digits for the year, two letters for the country, two or three letters for the location, and the suffix _IRRI to indicate if the trial was irrigated or _Late for late sown. Most trials were planted at the conventional sowing time for the location, under rainfed conditions (not indicated in the trial code).

‡ Percentage of variance accounted for by the factors included in the model.

§ Trials that remained ungrouped are indicated with "–".

the top 20; only A2.yld and A3.yld had ten (i.e., half of them) in common.

Total Effects

Six environment groups were identified for the total genetic effects (T1.yld–T6.yld Table 2; Fig. 1b), and illustrate the two scenarios for interpretation proposed earlier. Not all of the 197 genotypes were represented in each of the environment groups: T1.yld and T6.yld had 100 and 193 of the 197 genotypes, respectively.

In Scenario I, trials belong to the same environment group for both additive and total genetic effects when most of the pattern in the genotype rankings is due to additive effects. In this situation, additive and total effects tend to be similar and top yielders and best candidate parents tend to agree, as is the case in environment groups T1.yld, T3.yld, T4.yld and T6.yld.

Environment group T1.yld included five of the seven trials that belonged to A1.yld, specifically high yielding Australian trials (5–6.3 t ha^{-1}, Table 1) mainly located at Gatton

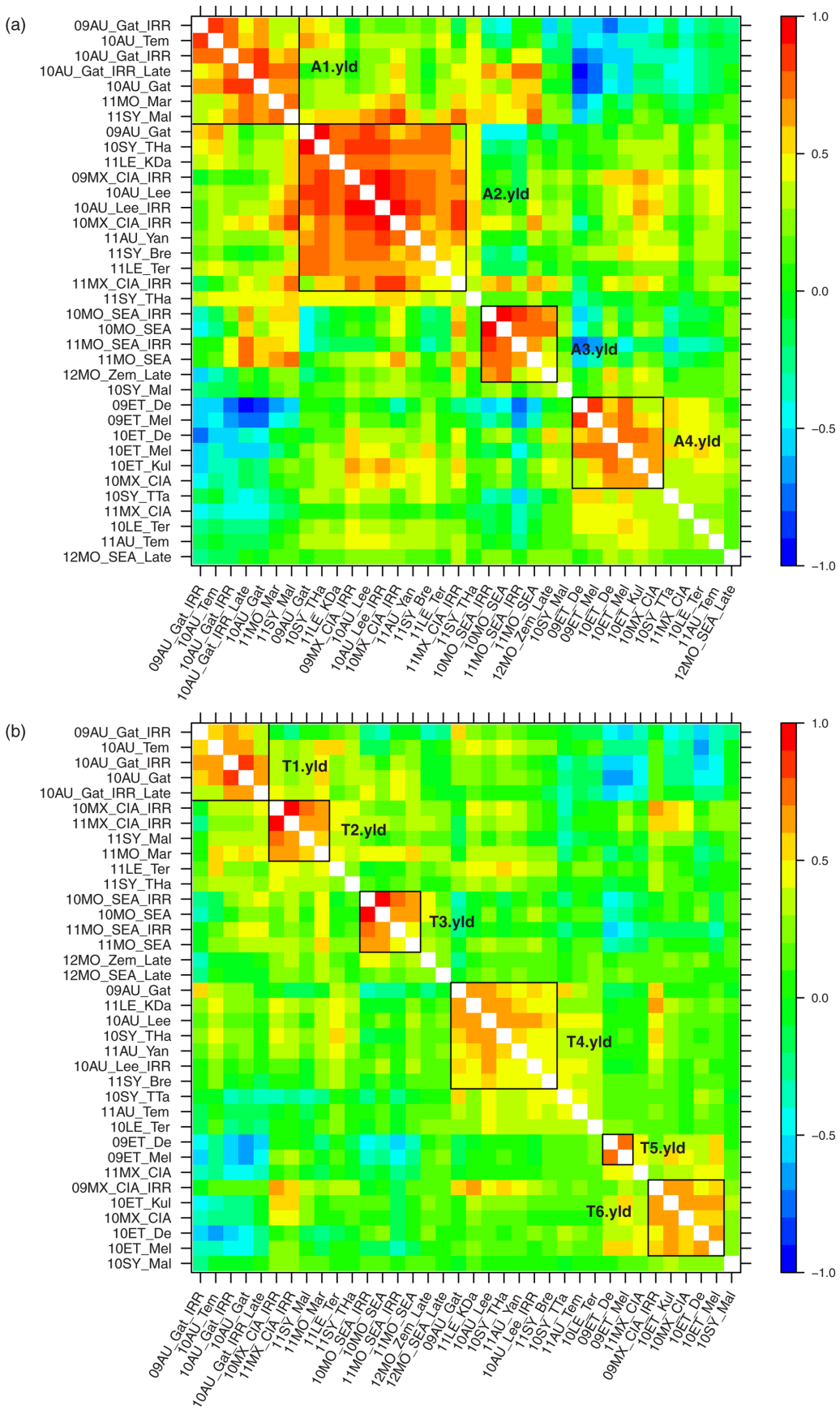

Fig. 1. Heatmaps of the (a) additive and (b) total genetic correlation matrices for yield. The scale on the right side of each figure indicates the magnitude and direction of the genetic correlations. Red-orange indicates strong positive genetic correlation (i.e., strong agreement in genotype rankings) between pairs of trials. Dark shades of blue indicate strong negative genetic correlation (i.e., strong disagreement in genotype rankings) between pairs of trials. Yellow, pale blue, and green indicate weak or no correlation between the trials.

(Table 2). Except for 10AU_Gat_IRR_Late, all trials had mainly additive genetic variance (75–99%) well explained by the model (80–100%vaf). Trial 10AU_Gat_IRR_Late had a higher proportion of non-additive variance (44%) but poorly explained by the model (11%vaf). For this environment group, the top 20 yielders coincided with the top 20 candidate parents for future crosses.

Environment group T3.yld contained the four Morocco trials from Sidi el Aidi, which were also in A3.yld. The 2010 trials yielded 1.8 and 1.9 t ha^{-1}, while the 2011 trials yielded 4.2 and 4.5 t ha^{-1} (Table 1). All trials had 50–97% additive genetic variance well accounted by the model (79–100%vaf) (Table 2). The two 2011 trials had >30% non-additive genetic variance but poorly accounted for by the model (19 and 23%vaf, Table 2). For this environment group, 17 of the 20 top yielders are also the best candidate parents for future crosses (Fig. 2).

Environment group T4.yld included seven of the 11 trials in A2.yld (Table 2). Trial mean yields range from 2.8 to 5.5 t ha^{-1} (Table 1). Trials in this environment group generally had moderate to high non-additive genetic variance (25–68%), partially accounted for by the model (30–55%vaf). In this case, eight of the top 20 selections are different for additive and total effects. The first and second ranked genotypes as potential parents are ranked 5th and 28th, respectively, for yield and the first and second top yielders ranked 5th and 7th, respectively, as potential parents. This is an example of Scenario I, case (ii): even though the trials remain in the same environment groups for additive and total effects there is more disagreement in the selections than for A1.yld and T1.yld or for A3.yld and T3.yld.

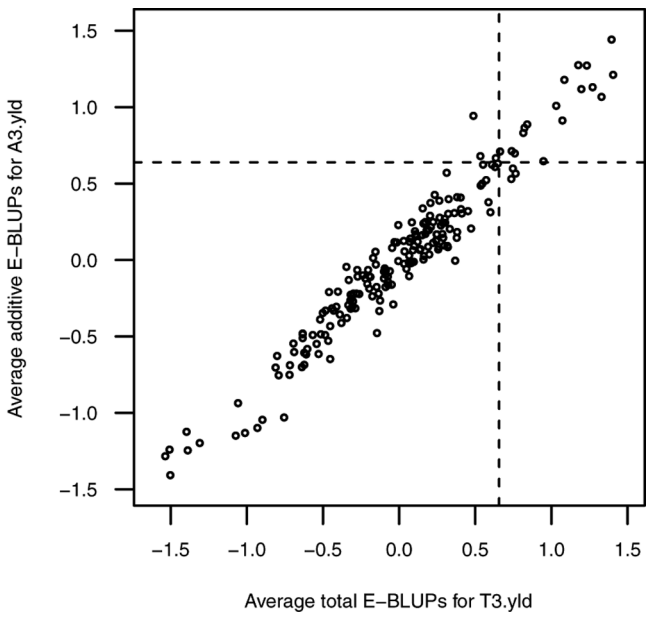

Fig. 2. Average additive E-BLUPs for A3.yld and average total E-BLUPs for T3.yld. These environment groups contain the four Moroccan Sidi el Aidi trials. Points to the right of the vertical dashed line are the 20 top-yielding genotypes, that is, potential varieties to release to industry. Points above the horizontal dashed line are the top 20 candidate parents for future crosses. Seventeen selections coincide.

Environment group T6.yld contained the three 2010 Ethiopian trials and two Mexican trials (Table 2). All trials except 09MX_CIA_IRR were in A4.yld, whereas 09MX_CIA_IRR was in A2.yld. Except for 10ET_Mel, all trials had low percentage of non-additive genetic variance (7–14%), totally accounted for by the model. Even though 10ET_Mel had 29% non-additive genetic variance, only 21% of it was explained by the model. Although four of the top 20 selections are different for additive and total effects, the top 10 selections coincide.

In Scenario II, trials belong to different environment groups for additive and total genetic effects when non-additive effects change the pattern in the genotype rankings. In this situation additive and total effects tend to differ and top yielders and best candidate parents generally disagree, as is the case in environment groups T2.yld and T5.yld.

Environment group T2.yld contained two trials from A1.yld (11MO_Mar and 11SY_Mal) and two from A2.yld (10MX_CIA_IRR and 11MX_CIA_IRR) (Table 2). Trial mean yields in this group ranged between 3.2 to 6.5 t ha^{-1} (Table 1). These four trials had a high proportion of non-additive genetic variance (40–78%) well accounted for by the model (48–100%vaf) and constituted an environment group that did not line up with any of the environment groups for additive effects. Only two genotypes identified as potential parents in A1.yld and nine in A2.yld coincide with the genotypes identified as best yielders in environment group T2.yld (Fig. 3).

Environment group T5.yld was formed by the two 2009 Ethiopian trials. These trials had the lowest mean yields of all trials (0.3 and 0.8 t ha^{-1}, Table 1) and were included in environment group A4.yld. These two trials formed a separate environment group for total effects because their non-additive effects changed the genotype rankings for total effects (34 and 100%vaf for the non-additive genetic variances). Half of the selections are different for additive and total effects.

In the heatmap (Fig. 1b) there are mainly green and light yellow areas at the intersection of most environment groups, indicating that the top ranking genotypes are mostly different for all groups. The intersections of T5.yld with T1.yld and T3.yld, and of T6.yld with T1.yld contain blue areas indicating crossover G × E and reversed ranking of the genotypes. All pairs of environment groups had a maximum of eight genotypes in common.

Seven trials remained ungrouped for both additive and total effects (Table 2). They had varying percentages of additive variance (38–99%), not well explained by the model (18–46%vaf) and most of them also had <50%vaf for non-additive variance, except for 11MX_CIA and

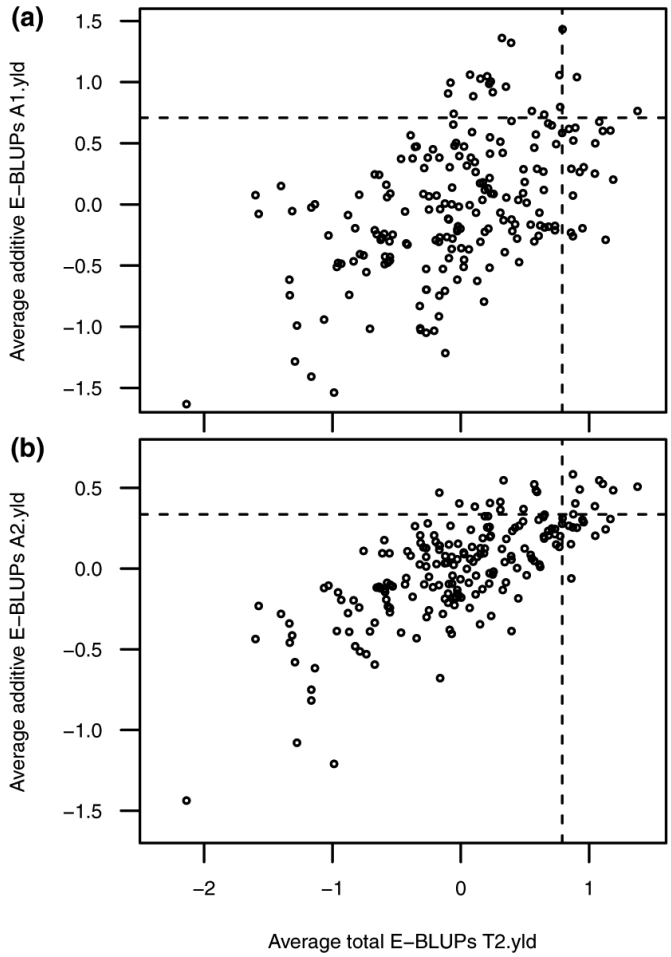

Fig. 3. Average additive E-BLUPs for (a) A1.yld and (b) A2.yld against average total E-BLUPs for T2.yld. Points to the right of the vertical dashed line are the 20 top yielding genotypes, i.e., potential varieties to release to industry. Points above the horizontal dashed line are the top 20 candidate parents for future crosses. Only two and eight selections coincide, respectively.

10SY_TTa. However, 11MX_CIA only had 1% of non-additive variance and 10SY_TTa had the lowest proportion of genetic versus residual variance (17%) of all trials. Reasons for these trials to remain ungrouped could include, for example, untimely rain and management problems. All these ungrouped trials would not be used to make selections for yield. However, they can be useful to identify selections for other traits.

Plant Height: Additive and Total Effects

In contrast to yield, plant height is a highly heritable trait, so trials would be expected to have mostly additive genetic variance. The MET analysis for plant height included 31 of the 37 trials. The model of best fit to the data included two factors (FA2) for each of the additive and the non-additive genetic variance matrices. Most of the trials belong to the same environment group for both additive and total effects (Table 3).

Environment group A1.ht was formed by 25 trials with average additive genetic correlation equal to 0.88. The two trials that had lower additive genetic correlations with the remaining trials in the group (11MX_CIA and 12MO_SEA_Late, Fig. 4a) had only 52 and 48% of the additive genetic variance accounted for by the model, respectively, while all others had 83–100%vaf (Table 3).

Twenty three out of the 25 trials in A1.ht remained together to form T1.ht (Table 3). These trials had 58–100% additive genetic variance highly explained by the model. The only exception was 11MX_CIA with 52%vaf, but all the genetic variance in this trial was additive. Two trials (11MO_SEA_IRR and 11AU_Yan) had 38 and 42% non-additive variance but poorly accounted for by the model (~30%vaf). This is an example of Scenario I.

All other groups formed (A2.ht, A3.ht, T2.ht, and T3.ht) contained trials with a much higher proportion of non-additive than additive variance (51–92%), which was not expected for a highly heritable trait such as plant height. These environment groups very likely lack meaning for selection purposes. As a result, environment groups A1.ht and T1.ht are the most meaningful ones. All top 20 selections coincided for additive and total effects (Fig. 5). In contrast to a similar graph presented for yield (Fig. 2), the points in this Figure are much closer together around a straight line because the pattern in the genotype rankings is controlled by additive effects in a much stronger way than for yield. This is a characteristic of highly heritable traits.

Comparison of MET Analyses Considering Related and Independent Genotypes

To compare the FA model including $\mathcal{G}$ with the FA model that assumes unrelated genotypes, the latter was also fitted to both traits. The models that best described the data were an FA5 for yield (overall 81%vaf) and an FA3 for plant height (overall 90%vaf). For both traits, the majority of trials had a lower average PEV under the FA model including $\mathcal{G}$ (18 out of 36 for yield and 20 out of 31 for plant height; Fig. 6). For yield there were more trials with approximately the same average PEV under both models (15 for yield and 9 for plant height; Fig. 6). For both traits there were only a few trials poorly explained by the FA model using $\mathcal{G}$ and that had a lower average PEV under the FA model that assumes unrelated genotypes (3 for yield and 2 for plant height, respectively).

Comparison of the top 20 yielding genotypes under both models (Fig. 7a) showed that 18 out of the 20 coincided. However, the relativities of the genotype rankings changed. For example, the second and sixth best yielders that would have been selected under the model that assumes unrelated genotypes ranked 13th and 19th, respectively, under the model that assumes them related through $\mathcal{G}$. Similar results were observed for plant height (Fig. 7c), where 17 of the top 20 selections coincided but, for example, the first and seventh ranked under the model that assumes unrelated

Table 3. Plant height MET results. For each trial, estimated total genetic and residual variances, percentages of the total genetic variance that are additive and non-additive (these two columns add to 100%), percentages of additive and non-additive genetic variances explained by the regression parts of the final model (FA2 for both additive and non-additive effects), and environment group to which each trial belongs for additive and total effects.

Trial code†	Total genetic variance	Residual variance	Additive genetic variance	Non-additive genetic variance	%vaf‡ FA2 additive	%vaf‡ FA2 non-additive	Environment group§ (*n*)	
							Additive effects (A*n*.ht)	Total effects (T*n*.ht)
			%					
09AU_Gat_IRR	25	18	100	0	93	–	1	1
09AU_Gat_IRR_Late	34	7	100	0	100	–	1	1
10AU_Lee_IRR	30	7	100	0	88	–	1	1
10ET_Mel	41	38	100	0	95	–	1	1
10MX_CIA	40	8	100	0	86	–	1	1
10MX_CIA_IRR	69	14	100	0	91	–	1	1
11AU_Tem	40	31	100	0	89	–	1	1
11MX_CIA	51	23	100	0	52	–	1	1
11MX_CIA_IRR	52	17	100	0	96	–	1	1
09AU_Gat	33	10	98	2	100	18	1	1
10ET_Kul	51	48	96	4	94	100	1	1
11SY_THa	30	10	94	6	83	100	1	1
11LE_Ter	46	16	89	11	95	100	1	1
10ET_De	42	33	86	14	95	76	1	1
10MO_SEA_IRR	32	52	84	16	99	100	1	1
11MO_SEA	33	39	84	16	96	100	1	1
11SY_Mal	36	55	84	16	88	100	1	1
10MO_SEA	44	45	82	18	99	90	1	1
10LE_Ter	62	30	78	22	98	100	1	1
11LE_KDa	40	21	75	25	96	91	1	1
10SY_THa	41	21	72	28	100	93	1	1
11MO_SEA_IRR	27	63	62	38	100	36	1	1
11AU_Yan	63	8	58	42	100	31	1	1
12MO_SEA_Late	24	73	42	58	48	13	1	–
11SY_Bre	17	15	35	65	100	68	1	2
10AU_Gat_IRR_Late	46	20	49	51	100	8	2	–
10AU_Gat_IRR	95	35	32	68	100	95	2	3
10AU_Gat	82	28	29	71	100	100	2	3
09ET_De	11	31	25	75	48	100	3	2
09MX_CIA_IRR	6	1	8	92	100	15	3	–
09ET_Mel	25	53	41	59	41	79	–	2

† Trial code acronym composed of two digits for the year, two letters for the country, two or three letters for the location, and the suffix _IRR to indicate if the trial was irrigated or _Late for late sown. Most trials were planted at the conventional sowing time for the location, under rainfed conditions (not indicated in the trial code).

‡ Percentage of variance accounted for by the factors included in the model.

§ Trials that remained ungrouped are indicated with "–".

genotypes ranked 3rd and 12th, respectively, under the model that assumes them related through $\mathcal{G}$.

Furthermore, the model that includes $\mathcal{G}$ provided predictions for genotype additive effects (or breeding values), allowing more accurate selection of parents for future crosses instead of choosing the best performers as parents, as is the case with the model that assumes unrelated genotype effects. Comparison of the top 20 potential parents that would be chosen under the two models (Fig. 7b and d) showed that for yield and plant height 4 and 5 selections would be different for each trait. Additionally, rankings of the common genotypes were in disagreement. For example, for yield (Fig. 7b) the second and sixth preferred parents under the model that assumes unrelated genotypes ranked 17th and 12th, respectively, for additive effects, and the fifth preferred parent based on additive effects would have been discarded under the model that assumes unrelated genotypes, for which it ranked 54th. Similar results were observed for plant height (Fig. 7d), where for example, the first and eighth preferred parents under the model that assumes unrelated genotypes ranked 4th and 15th for additive effects, and the first preferred for additive effects ranked sixth under model with unrelated genotypes.

DISCUSSION

In this paper we showed how to include a genomic relationship matrix $\mathcal{G}$ that accounts for relationships among genotypes (Oakey et al., 2007) into the FA model for analysis of MET data for the purpose of parental and varietal selection. The relationship matrix typically used in this

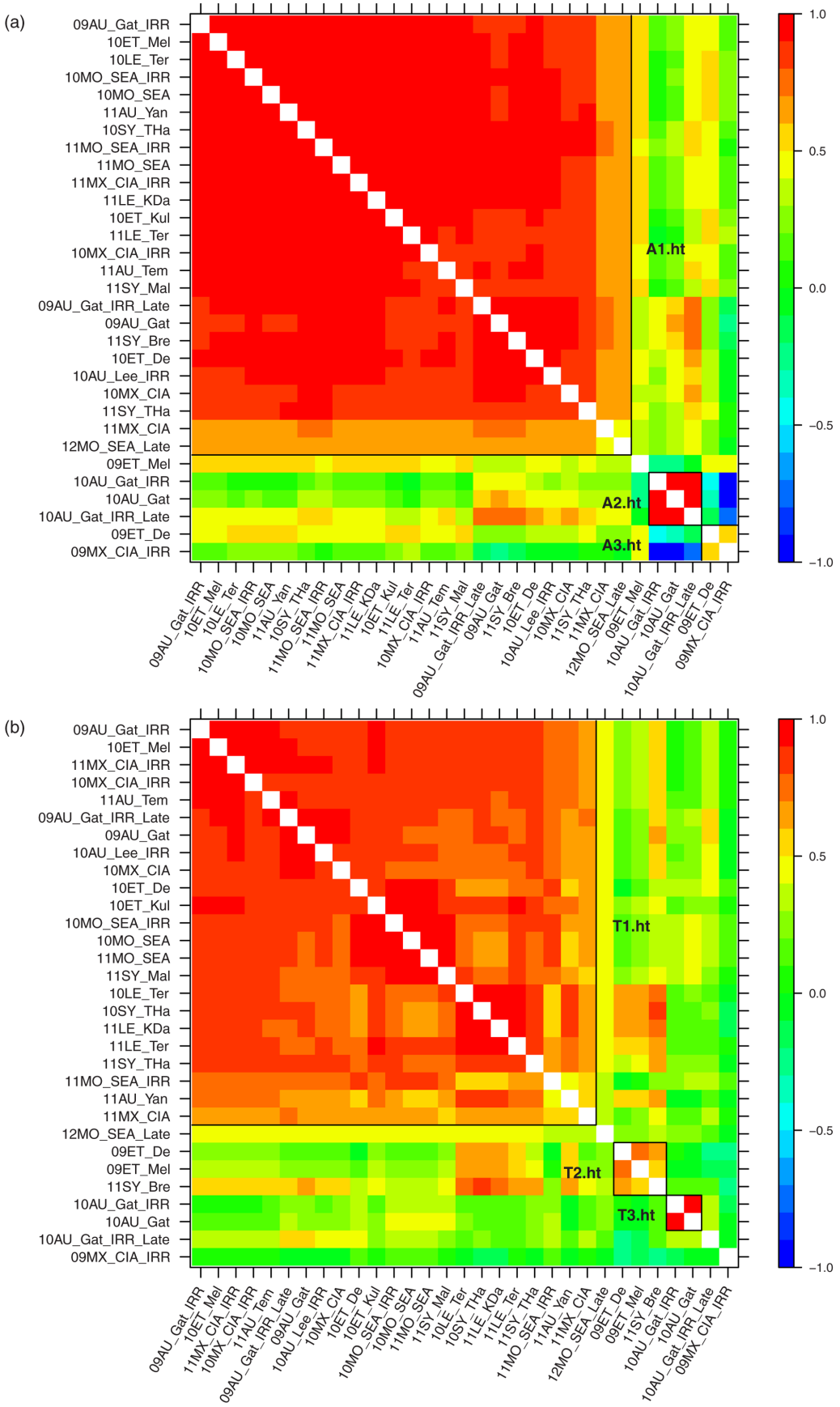

Fig. 4. Heatmaps of the (a) additive and (b) total genetic correlation matrices for plant height. The scale on the right side of each figure indicates the magnitude and direction of the genetic correlations. Red-orange indicates strong positive genetic correlation (i.e., strong agreement in genotype rankings) between pairs of trials. Dark shades of blue indicate strong negative genetic correlation (i.e., strong disagreement in genotype rankings) between pairs of trials. Yellow, pale blue and green indicate weak or no correlation between the trials.

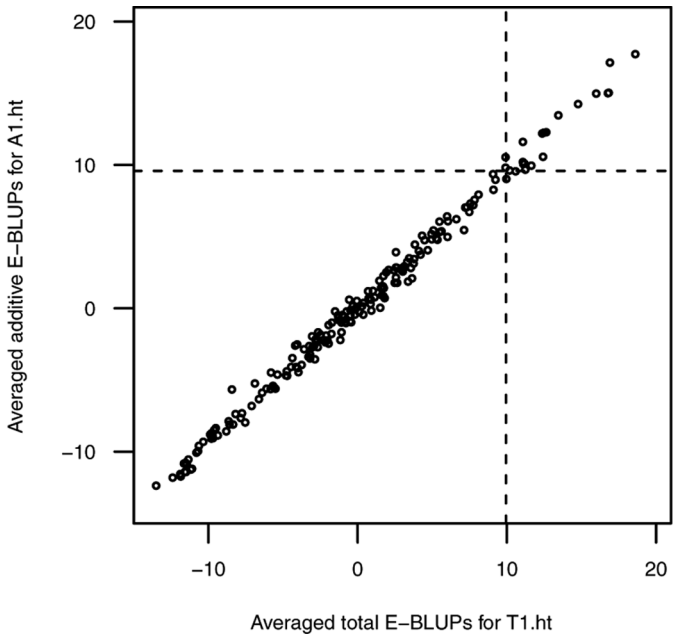

Fig. 5. Average additive E-BLUPs for A1.ht against average total E-BLUPs for T1.ht. Plant height is a highly heritable trait, so total genetic effects are mainly additive genetic effects. The top tallest 20 and candidate parents coincide.

model is the additive genetic relationship matrix **A**, constructed from pedigree information. The construction of **A** assumes that the base (or founder) population from which the genotypes in the MET come from has not undergone selection. This is rarely the case in plant breeding programs (Mathews et al., 2007; Oakey et al., 2006; Piepho et al., 2008). Additionally, **A** reflects the expected proportion of the genome that is identical by descent between two individuals, but relationships between individuals in plant populations where selection of genotypes is undertaken over several generations can be far from what is expected (Oakey et al., 2006). To overcome these problems, DArT molecular markers of the genotypes were used in the construction of $\mathcal{G}$. It should be noted that other types of molecular markers can also be used, such as SNP markers.

The form of $\mathcal{G}$ used in this paper (VanRaden, 2008) is analogous to the **A** matrix and has the advantage that genetic effects can be partitioned in a way that is familiar to plant breeders: additive genetic effects useful when selecting parents for crossing and residual non-additive genetic effects that can be responsible for enhanced or reduced performance of the genotypes in the different environments. The total genetic effects predict the overall performance of genotypes in target environments and are useful when selecting genotypes for release to industry.

Environment groups where genotypes perform similarly were identified for the international wheat MET dataset by applying the multivariate technique of cluster analysis to the additive and total genetic correlation matrices. These environment groups should be interpreted in conjunction with the percentages of additive and

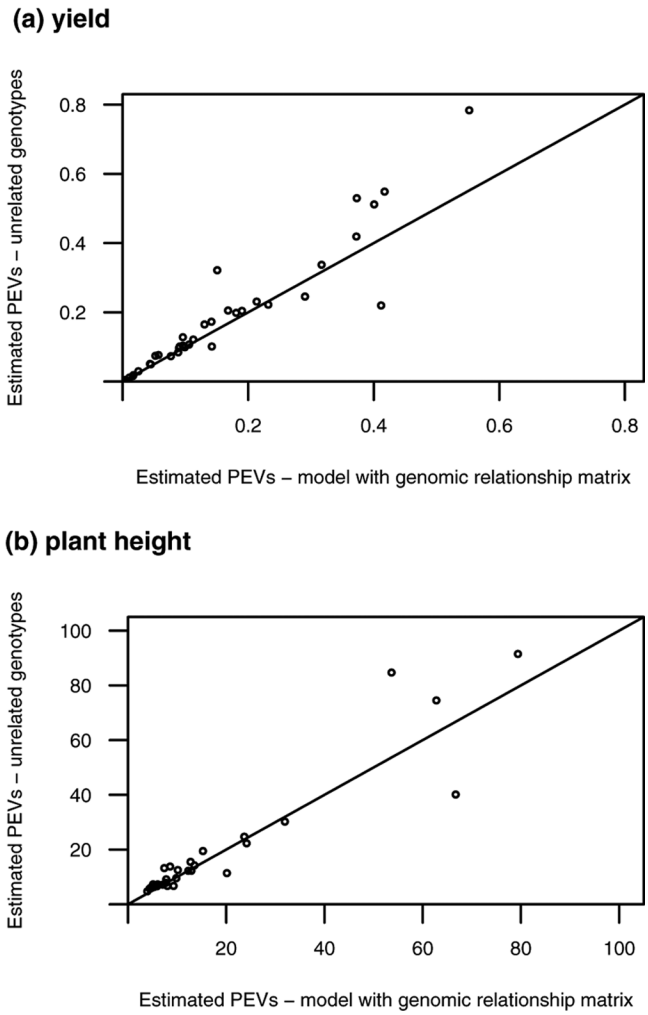

Fig. 6. Estimated prediction error variances (PEVs) for trials included in the (a) yield and (b) plant height MET analyses. The majority of trials had higher PEV under the model that assumes unrelated genotypes (points above the 1:1 line).

non-additive genetic variances for each trial forming the group, and how well the model explains them. Graphical representations of the environment groups on the heatmaps of the genetic correlation matrices facilitated the interpretation by providing the (approximate) magnitudes of the genetic correlations within and between environment groups at a glance. Knowledge about the individual trials that form the specific environment groups together with a thorough understanding of the meaning of the environment group will determine if an environment group should be considered for selection purposes. For example, for plant height two of the three environment groups formed for additive effects as well as two of the three formed for total effects contained only a few trials each. Further examination of the results for those individual trials indicated that those environment groups were not meaningful.

The very weak and even negative correlation between environment groups for yield, both for additive and total effects, highlights that high yielding genotypes and

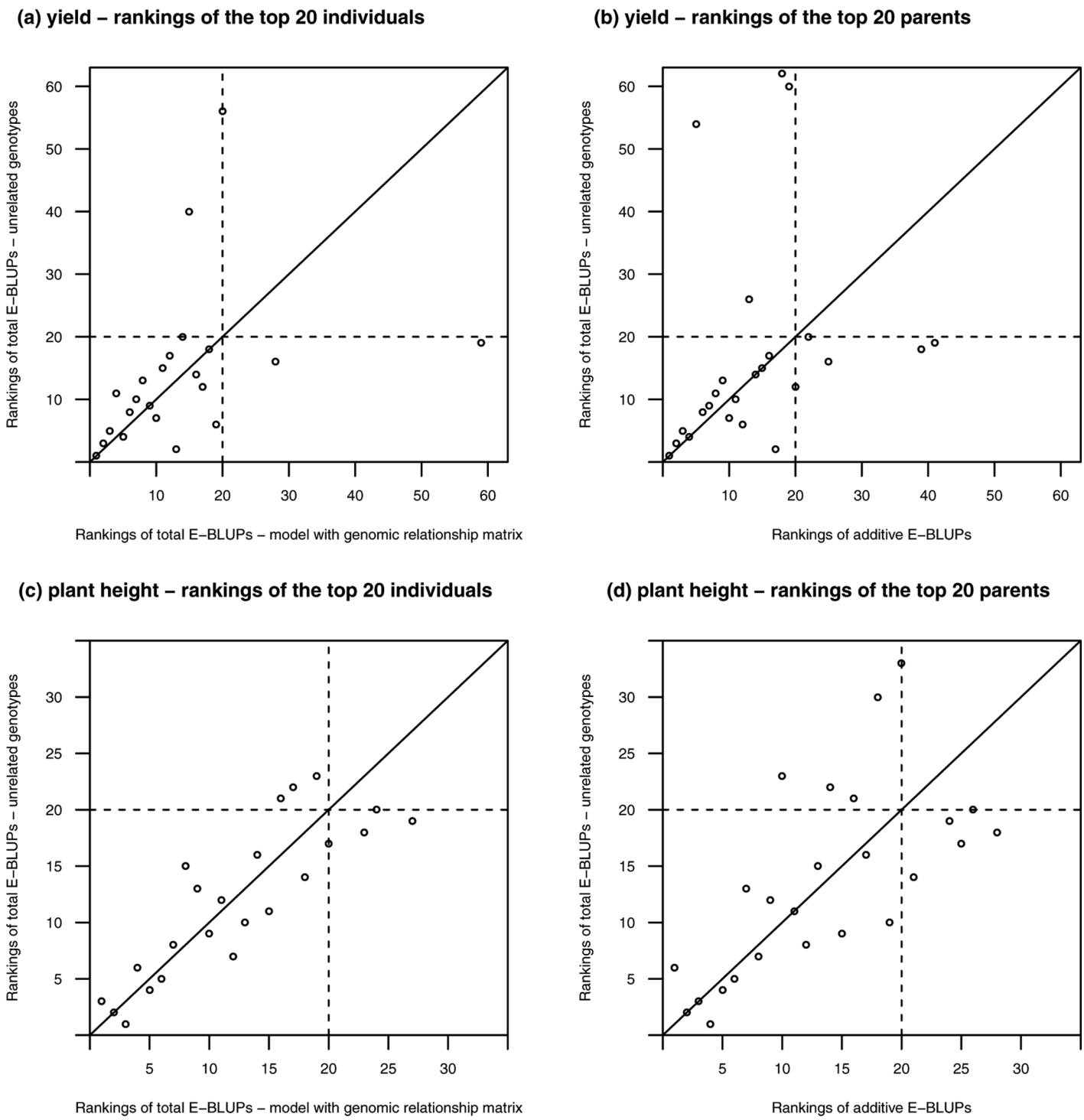

Fig. 7. Rankings of the top 20 individuals and potential parents for yield (a, b) and plant height (c, d). The *y* axis in all graphs presents the rankings for the predicted total effects of the genotypes obtained from the model that assumes independent genotypes. The *x* axis presents the rankings for the predicted total effects (a, c) and the predicted additive effects (b, d) under the model that assumes genotypes related via the genomic relationship matrix. Although the majority of the selections coincide, the relativities of the genotype rankings change.

candidate parents for crossing do not necessarily agree across environment groups, and it may be necessary to make selections within each of them. For plant height, a more heritable trait than yield, most trials were grouped into one main environment group for both additive and total genetic effects, with the top 20 genotypes and top 20 candidate parents for crossing coinciding. It should be noted that breeders do not necessarily look for the 'tallest' genotypes but more commonly those with a height between 70 and 100 cm (Richards, 1992), and data should be looked at from this perspective. When environment groups (nearly) coincide for additive and total genetic effects, the agreement in selections for candidate parents and best performers depends on the percentages of additive genetic variance in the trials that form the environment groups and how well explained they are by the model.

The higher the percentage of additive genetic variance in the trials and the better explained by the model, the higher the coincidence will be.

One of the advantages of including the $\mathcal{G}$ matrix in the analysis of the wheat dataset presented in this paper was predominantly lower average PEVs for total effects of individual trials under the FA model with the $\mathcal{G}$ matrix, compared with those from the FA model that assumed unrelated genotypes. This could be due to the $\mathcal{G}$ matrix describing the underlying distribution in a more realistic way. These results are in agreement with those reported by other authors that included **A** in the FA model: Oakey et al. (2006) reported that models including **A** had lower average PEVs for the genetic effects for 14 individual wheat trials, making it preferable for estimating the genetic value of a genotype; Beeck et al. (2010) and Oakey et al. (2007) reported that the FA model that included **A** also fitted significantly better than the model that ignored pedigrees for canola and sugarcane MET datasets, respectively. When comparing genotype rankings between the model that included $\mathcal{G}$ and the model that assumed unrelated genotypes for our wheat dataset, nearly all top 20 selections coincided. However, the changes in relativities for yield were generally more extreme than for plant height, and this is likely related to the less heritable nature of yield compared to plant height. Additionally, selecting parents based on the model that assumed unrelated genotypes would have ignored highly ranked genotypes identified by the model that included the known genetic relationships.

The comparison of PEVs from FA models using matrices **A** and $\mathcal{G}$ for this wheat dataset was not undertaken because of limitations with **A**. First, the pedigree information was incomplete, a common problem in many breeding programs; and second, fitting the model using an **A** matrix based on information from the available genotypes was computationally expensive, even when using a pedigree of modest depth (four generations) because the number of individuals in the pedigree was several times the number in the data.

In summary, the elements of $\mathcal{G}$ reflect the actual proportion of the parts of the genome surveyed that is identical by state between pairs of individuals under consideration and by including the form of $\mathcal{G}$ proposed by VanRaden (2008) in the FA model the genetic effects can be partitioned in a way familiar to plant breeders. Additionally, for the international wheat MET dataset analyzed, the genotype predictions for the trials generally had lower average PEVs. This method of analysis could be readily implemented to accelerate genetic gain by plant breeding programs that have molecular markers for the genotypes under study.

Acknowledgments

The authors gratefully acknowledge the Generation Challenge Program (Mexico), the Grains Research and Development Corporation (Australia), and the International Centre for Agricultural Research in the Dry Areas (ICARDA) for funding this work. The main author would like to thank Arthur Gilmour for his invaluable help with coding in ASReml 4. FCO and MFD would like to thank Farid Makdis (ICARDA), Hassan Ouabbou (INRA–CRRA), Solomon Gelacha and Firdissa Eticha (EIAR), Tony Condon (CSIRO) and Matthew Reynolds (CIMMYT) for contributing to the project and running field trials. The authors also thank three anonymous reviewers for their helpful comments that significantly improved the manuscript.

APPENDIX

Table A1. Residual maximum likelihood (ASReml) code used for fitting the final model for plant height.

Line	Code
1 †	ASReml code – Plant height MET
2	Trial 31 !A !LL 17
3	replicate !I
4	row !I
5	column !I
6	Genotype 243 !A !LL 60
7	GenoYES 197 !A !LL 60 !L geno_order.txt !LSKIP 1
8	GenoNO 46 !A !LL 20
9	plntht
10	lrow
11	lcol
12	Gmat.grm !skip 1 !PSD
13	alldata.height.csv !CSV !skip 1 !mvinclude !maxit 5
14	!SUBSET IDETrial Trial 3:9,11:12,14:16,19,21:25,28:31
15	plntht ~ mu + Trial + at(Trial,14,30).lcol,
16	!r xfa(Trial,2).giv(GenoYES) +
17	xfa(IDETrial,2).ide(GenoYES)+
18	at(Trial,1:5,7:16,19:25,28:31).GenoNO +
19	at(Trial).replicate + at(Trial,7,9,11,19,28,30,31).column
20	+ at(Trial,4,14,23,30).row,
21	!f mv
22	31 2 2 !NODISPLAY
23	8 column AR 0.2 # 09AU_Gat_IRR 1
24	18 row ID
25	8 column AR 0.2 # 09AU_Gat_IRR_Late 2
26	18 row AR 0.4
27	…
28	20 column AR 0.2 # 12MO_SEA_Late 31
29	20 row AR 0.4
30	xfa(Trial,2).giv(GenoYES) 2
31	xfa(Trial,2) 0 xfa2 !+93 !G31P31PF30P
32	1.7 1.8 1.3 2.4 8.6 2.2 32.9 5.6 39.7 2.7 3.6
33	3.7 2.6 1.2 0.5 0.7 9.3 6.7 0.5 4.1 1.7 1.2
34	3.7 0.1 0.5 2.2 25.6 1.4 4.5 3.5 6.0
35	5.0 6.2 5.8 1.2 2.2 -0.4 1.6 2.9 2.6 5.7 6.2
36	7.5 6.8 7.9 5.8 6.8 8.1 6.3 6.1 6.4 6.6 6.0
37	7.1 4.4 5.8 7.5 5.4 2.6 6.0 5.3 2.4
38	0 30*0.05

Cont'd

Table A1. *Cont'd.*

Line	Code
39	giv(GenoYES) 0 GIV1
40	xfa(IDETrial,2).ide(GenoYES) 2
41	xfa(IDETrial,2) 0 xfa2 !+66 !G22P22PF21P
42	2.2 0.1 2.2 1.4 45.6 26.5 19.2 1.6 0.1 0.4 1.3
43	1.9 0.5 17.2 1.2 0.1 7.0 1.6 2.7 0.1 2.5 11.7 0.5
44	2.5 3.1 -1.0 -0.3 -0.5 -1.6 -1.8 -0.5 3.0 -1.3 -1.7
45	2.9 3.0 2.7 1.4 -1.8 -1.5 2.6 -0.5 1.0 0.5
46	0 21*0.05
47	ide(GenoYES)

† Explanation of the ASReml code lines: (1) title line; (2) trials in the data file: number, alphanumeric variable, name length; (3–5) replicate, row, and column: all integer variables; 6, genotypes in the data file: number, alphanumeric variable, name length; (7) genotypes that have marker data: number, alphanumeric variable, name length (*geno_order.txt* is a file with the genotypes listed in the same order as in the genomic relationship matrix [GRM] file to ensure that the variable is coded to agree with the order of names in the GRM matrix. If the file has a header row, it needs to be skipped); (8) genotypes without marker data: number, alphanumeric variable, name length; (9) plant height; (10–11) centered rows and columns; (12) *Gmat.grm* is the GRM obtained from the *A.mat* function in rrBLUP, supplied in the sparse format (a free format file with three numbers per line (row, column, value) defining the lower triangle row-wise of the matrix). This file must be sorted column within row, and every diagonal element must be present or will be assumed to be zero. *Gmat.grm* was created with the R code given in Part 2 below. Note that the genomic matrix can also be supplied in the dense format (a file containing the matrix presented lower triangle row-wise, with each row beginning on a new line). The extension *.grm* indicates that it is not an inverse matrix and ASReml will invert it. !PSD allows the matrix to be positive semi definite. This allows ASReml to introduce Lagrangian multipliers to accommodate linear dependencies and rows with zero elements, and allows ASReml to proceed; (13) data file name, the first line contains headings so it must be skipped; (14) creates a subset with trials that have non-additive variance; (15) Trial and linear trends for columns (fixed effects at Trial level); (16) factor analytic (FA) model of order 2 fitted to additive genetic effects (random effects); (17) FA model of order 2 fitted to non-additive genetic effects (random effects); (18) variance components fitted for trials with genotypes without markers (random effects); (19–20) replicate, column and row effects at Trial level (random effects); (21) missing values (fixed terms in a sparse set); (22) variance header line: 31 R structures (one for each trial) are the product of 2 dimensions, there are 2 G structures. !NODISPLAY prevents variograms and residuals plots to be displayed for each trial; (23–29) two lines for each trial with the dimensions for columns and rows and the spatial model fitted to each direction; (30) additive genetic effects model heading, contains two parts; (31) first part: FA model of order 2, with 93 initial values provided (the first 62 obtained from the previous order model (specific variances and FA1 loadings) and the last 31 provided by user; (32–37) initial values for additive genetic effects model; (39) second part: relationship matrix for genotypes is the inverse of the GRM provided; (40) non-additive effects model heading, contains 2 parts; (41) first part: FA model of order 2,with 66 initial values provided (the first 44 obtained from the previous order model (specific variances and FA1 loadings) and the last 22 provided by user; (42–46) initial values for non-additive genetic effects model; (47) second part: there is no relationship matrix for genotypes.

Table A2. R code to write the genomic relationship matrix obtained through rrBLUP in sparse format.†

```
mat2sparse <- function (X, rowNames = dimnames(X)[[1]])
{
    which <- (X != 0 & lower.tri(X, diag = TRUE))
    df <- data.frame(row = t(row(X))[t(which)], col = t(col(X))[t(which)],
        val = t(X)[t(which)])
    if (is.null(rowNames))
        rowNames <- as.character(1:nrow(X))
    attr(df, "rowNames") <- rowNames
    df
}
write.table(mat2sparse(G.mat), 'Gmat.grm',row.names=FALSE)
```

† G.mat is the relationship matrix obtained through the *A.mat* function in rrBLUP.

References

Beeck, C.P., W.A. Cowling, A.B. Smith, and B.R. Cullis. 2010. Analysis of yield and oil from a series of canola breeding trials. Part I. Fitting factor analytic mixed models with pedigree information. Genome 53:992–1001. doi:10.1139/G10-051

Burgueño, J., G. de los Campos, K. Weigel, and J. Crossa. 2012. Genomic prediction of breeding values when modeling genotype by environment interaction using pedigree and dense molecular markers. Crop Sci. 52:707–719. doi:10.2135/cropsci2011.06.0299

Cullis, B.R., A.B. Smith, C.P. Beeck, and W.A. Cowling. 2010. Analysis of yield and oil from a series of canola breeding trials. Part II. Exploring variety by environment interaction using factor analysis. Genome 53:1002–1016. doi:10.1139/G10-080

Dawson, J.C., J.B. Endelman, N. Heslot, J. Crossa, J. Poland, S. Dreisigacker, Y. Manès, M.E. Sorrells, and J.-L. Jannik. 2013. The use of unbalanced historical data for genomic selection in an international wheat breeding program. Field Crops Res. 154:12–22. doi:10.1016/j.fcr.2013.07.020

Detering, F., E. Hunter, G. Uszynski, P. Wenzl, and K. Andrzej. 2010. A consensus genetic map of wheat: Ordering 5000 Wheat DArT markers. 20th ITMI and 2nd WGC Workshop, 1–5 September, Beijing.

Dreccer, A.F., A.G. Borgognone, F.C. Ogbonnaya, R.M. Trethowan, and B. Winter. 2007. CIMMYT-selected derived synthetic bread wheats for rainfed environments: Yield evaluation in Mexico and Australia. Field Crops Res. 100:218–228. doi:10.1016/j.fcr.2006.07.005

Dreccer, M., A. Condon, F. Makdis, L. Barnes, H. Ouabbou, F. Eticha, et al. 2013. Assessment of the relative impact of key physiological traits on adaptation to drought in different mega-environments. InterDrought-IV. Perth, Australia.

Endelman, J.B. 2011. Ridge regression and other kernels for genomic selection with R package rrBLUP. Plant Genome 4:250–255. doi:10.3835/plantgenome2011.08.0024

Endelman, J.B. and J.L. Jannink. 2012. Shrinkage estimation of the realized relationship matrix. G3:Genes Genomes Genetics 2:1405–1413.

Gilmour, A.R., B.R. Cullis, A.P. Verbyla, and A.C. Gleeson. 1997. Accounting for natural and extraneous variation in the analysis of field experiments. J. Agric. Biol. Environ. Stat. 2:269–293. doi:10.2307/1400446

Gilmour, A.R., B.J. Gogel, B.R. Cullis, S.J. Welham, and R. Thompson. 2015. ASReml user guide. Release 4.0 VSN International Ltd, Hemel Hempstead, UK. http://www.vsni.co.uk (verified 20 June 2016).

Hardner, C.M., I.S.E. Bally, and C.L. Wright. 2012. Prediction of breeding values for average fruit weight in mango using a multivariate individual mixed model. Euphytica 186:463–477. doi:10.1007/s10681-012-0639-7

Hayes, B.J., P.M. Visscher, and M.E. Goddard. 2009. Increased accuracy of artificial selection by using the realized relationship matrix. Genet. Res. 91:47–60. doi:10.1017/S0016672308009981

Henderson, C.R. 1976. A simple method for computing the inverse of a numerator relationship matrix used in prediction of breeding values. Biometrics 32:69–83. doi:10.2307/2529339

Hodson, D.P., and J.W. White. 2007. Use of spatial analyses for global characterization of wheat-based production systems. J. Agric. Sci. 145:115–125. doi:10.1017/S0021859607006855

Jordan, D.R., Y.Z. Tao, I.D. Godwin, R.G. Henzell, M. Cooper, and C.L. McIntyre. 2005. Comparison of identity by descent and identity by state for detecting genetic regions under selection in a sorghum pedigree breeding program. Mol. Breed. 14:441–454. doi:10.1007/s11032-005-0901-y

Kelly, A.M., A.B. Smith, J.A. Eccleston, and B.R. Cullis. 2007. The accuracy of varietal selection using factor analytic models for multi-environment plant breeding trials. Crop Sci. 47:1063–1070. doi:10.2135/cropsci2006.08.0540

Maechler, M., P. Rousseeuw, A. Struyf, M. Hubert and K. Hornik. 2014. cluster: Cluster analysis basics and extensions. R package version 1.15.2. R Development Core Team. https://www.r-project.org/ (verified 20 June 2016).

Mathews, K.L., S.C. Chapman, R. Trethowan, W. Pfeiffer, M. van Ginkel, J. Crossa, T. Payne, I. DeLacy, P.N. Fox, and M. Cooper. 2007. Global adaptation patterns of Australian and CIMMYT spring bread wheat. Theor. Appl. Genet. 115:819–835. doi:10.1007/s00122-007-0611-4

Oakey, H., A. Verbyla, W. Pitchford, B. Cullis, and H. Kuchel. 2006. Joint modeling of additive and non-additive genetic line effects in single field trials. Theor. Appl. Genet. 113:809–819. doi:10.1007/s00122-006-0333-z

Oakey, H., A.P. Verbyla, B.R. Cullis, X.M. Wei, and W.S. Pitchford. 2007. Joint modeling of additive and non-additive (genetic line) effects in multi-environment trials. Theor. Appl. Genet. 114:1319–1332. doi:10.1007/s00122-007-0515-3

Ogbonnaya, F.C., S. Seah, A. Delibes, J. Jahier, I. Lopez-Brana, R.F. Eastwood, et al. 2001. Molecular-genetic characterisation of a new nematode resistance gene in wheat. Theor. Appl. Genet. 102:623–629. doi:10.1007/s001220051689

Patterson, H.D., and R. Thompson. 1971. Recovery of inter-block information when block sizes are unequal. Biometrica 58:545–554. doi:10.1093/biomet/58.3.545

Piepho, H.P., J. Möhring, A.E. Melchinger, and A. Büchse. 2008. BLUP for phenotypic selection in plant breeding and variety testing. Euphytica 161:209–228. doi:10.1007/s10681-007-9449-8

Poland, J., J. Endelman, J. Dawson, J. Rutkoski, S.Y. Wu, Y. Manes, et al. 2012. Genomic selection in wheat breeding using genotyping-by-sequencing. Plant Genome 5:103–113. doi:10.3835/plantgenome2012.06.0006

R Core Team. 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Richards, R.A. 1992. The effect of dwarfing genes in spring wheat in dry environments. 1. Agronomic characteristics. Aust. J. Agric. Res. 43:517–527. doi:10.1071/AR9920517

Smith, A., B. Cullis, and R. Thompson. 2001. Analyzing variety by environment data using multiplicative mixed models and adjustments for spatial field trend. Biometrics 57:1138–1147. doi:10.1111/j.0006-341X.2001.01138.x

Smith, A.B., J.K. Stringer, X. Wei, and B.R. Cullis. 2007. Varietal selection for perennial crops where data relate to multiple harvests from a series of field trials. Euphytica 157:253–266. doi:10.1007/s10681-007-9418-2

Stefanova, K.T., and B. Buirchell. 2010. Multiplicative mixed models for genetic gain assessment in lupin breeding. Crop Sci. 50:880–891. doi:10.2135/cropsci2009.07.0402

Stefanova, K.T., A.B. Smith, and B.R. Cullis. 2009. Enhanced diagnostics for the spatial analysis of field trials. J. Agric. Biol. Environ. Stat. 14:392–410. doi:10.1198/jabes.2009.07098

VanRaden, P.M. 2008. Efficient methods to compute genomic predictions. J. Dairy Sci. 91:4414–4423. doi:10.3168/jds.2007-0980

Verbyla, A.P., B.R. Cullis, M.G. Kenward, and S.J. Welham. 1999. The analysis of designed experiments and longitudinal data by using smoothing splines. J. R. Stat. Soc. Ser. C Appl. Stat. 48:269–311. doi:10.1111/1467-9876.00154