

# Improving the accuracy of selection for late maturity $\alpha$ -amylase in wheat using multi-phase designs

D. G. Butler<sup>A,D</sup>, M. K. Tan<sup>B</sup>, and B. R. Cullis<sup>C</sup>

<sup>A</sup>Primary Industries & Fisheries, Department of Employment, Economic Development & Innovation, 203 Tor Street, Toowoomba, Qld 4350, Australia.

<sup>B</sup>NSW Department of Primary Industries, Elizabeth Macarthur Agricultural Institute, Private Mail Bag 8, Camden, NSW 2570, Australia.

<sup>C</sup>NSW Department of Primary Industries, Wagga Wagga Agricultural Institute, Private Mail Bag, Wagga Wagga, NSW 2650, Australia.

<sup>D</sup>Corresponding author. Email: david.butler@deedi.qld.gov.au

**Abstract.** The assessment of grain defect traits is assuming greater importance in wheat germplasm selection. Late maturity  $\alpha$ -amylase is one such characteristic that renders wheat unsuitable for high value end products, even though the grain may appear sound. Phenotyping defect traits typically involves a multi-phase process, where genetic material for assay has been affected by non-genetic sources of variation in one or more previous linked stages of experimentation or preparation. The adoption of appropriate statistical design and analysis methods in these situations is, however, not widespread. Substantial sources of non-genetic variation were identified in the analysis of a designed experiment to measure late maturity  $\alpha$ -amylase expression, indicating the potential for improved selection decisions. A simulation study based on these results suggests that significant gains over current methods in the accuracy of phenotyping this grain defect can be achieved with sound multi-phase statistical design and analysis techniques. Although restricted in scope, the simulation also indicates that a considerable increase in estimated heritability could be expected from the proposed methodology.

## Introduction

The objective of plant breeding is to provide new varieties with improved performance for several traits of economic importance. A key trait in this regard for cereal crops is grain yield. Varietal performance for grain yield is assessed *via* a series of field trials (known as multi-environment trials, or METs), the design and analysis of which have received substantial attention in the literature. Increasingly, grain quality and defect traits are having a more significant role in the selection process. The trait we focus on here is late maturity  $\alpha$ -amylase (LMA), a genetic defect of wheat, which can render grain unsuitable for many higher value end products. The frequency of lines in wheat breeding programs which express LMA appears reasonably high, and hence LMA has become an important trait that is now routinely assessed in all wheat breeding programs in Australia.

In contrast to grain yield, statistical design and analysis methods specific to the measurement of LMA expression have been far less extensively researched. Like many grain quality traits in wheat, phenotyping for LMA involves a complex process. The most common procedure is based on the protocol developed by Mrva and Mares (2001), hereafter referred to as the MM protocol. Briefly, their approach involves growing plants either in a glasshouse, field trial, or microclimate room. Spikes are tagged at anthesis and subsequently, primary tillers are either cut and placed in a refrigerated waterbath in a controlled-environment room, or more recently whole plants are simply

moved *in situ*, maintaining the temperature at 10–12°C with 12 h of light and 12 h of darkness. After 8–10 days, plants are then transferred to a warm glasshouse or returned to the microclimate room for ripening. At harvest ripeness, grains are sampled at random from the main shoot tiller of each plant. The MM protocol assesses LMA by placing a composite sample of 5 half grains (the half which is distal to the embryo) in tubes on a 96-place tray rack, processing, and an aliquot from each tube is then placed on the corresponding position of an antibody coated plate for ELISA. Lastly, absorbance readings (optical density, OD) are then taken from the plate. A high value of OD indicates the presence of high isoelectric point  $\alpha$ -amylase and therefore high LMA.

The MM protocol for phenotyping LMA is an example of a multi-phase experiment. The definition of a multi-phase experiment is an experiment which involves several time periods, and has observational units which are completely different to those from the preceding phases. The simplest multi-phase experiment involves 2 phases, and these were first introduced by McIntyre (1955). A very common 2-phase experiment in agriculture involves a field phase and a laboratory phase. In the field phase, a field trial is carried out on field (experimental) units and then the grain from that experiment is taken in to the laboratory where the experiment involves laboratory units. Phenotyping for LMA using the MM protocol has at least 4 phases, most of which are unrandomised

and linked in the sense that no additional experimental design is used as we move from phase  $i$  to phase  $i + 1$ ,  $i = 1, 2, 3$ .

Smith *et al.* (2006) present a brief review of the design and analysis of multi-phase experiments in plant breeding. They introduce the notion of a  $p/q$ -rep design, which is a design for 2-phase experiments aimed at reducing the number of observational units to a manageable and economic level, but still allowing estimation of the major sources of non-genetic variation arising from each phase through the fitting of a (complex) linear mixed model. They illustrate the approach using milling yield in wheat and present evidence to suggest that substantial gains in response to selection for milling yield may be achieved by routine deployment of these designs in wheat breeding programs. Despite this work, there has been little adoption of this technology in Australia or overseas. There was sporadic adoption in some wheat and barley breeding programs but the rate of adoption has decreased with the privatisation of wheat and barley breeding in Australia.

Determination of an efficient design for a multi-phase experiment requires knowledge of the relative magnitude of the non-genetic sources of variation. For example, non-genetic variation for milling yield in wheat appears to be mostly due to laboratory variation (Smith *et al.* 2006), and so an efficient multi-phase design would require sufficient replication of field plots in the laboratory. On the other hand, Cullis *et al.* (2003) show that for most barley malting traits, the non-genetic variation is dominated by field variation, implying that sufficient replication of varieties from the field be maintained in the laboratory phase of the experiment.

It is well accepted that selection for LMA expression is an imperfect process, meaning that the phenotype can often be a poor predictor (in a statistical sense) of the genotype. Non-genetic influences on LMA expression include temperature during grain development (Mrva and Mares 2001) and hence, at a macro-level, LMA expression would exhibit genotype  $\times$  environment interaction. Furthermore, on a micro-level, variation in growth and development in glasshouses may also significantly affect LMA expression. Mrva and Mares (2001) therefore recommend using healthy non-diseased plants grown in optimal conditions.

There appears to have been no study undertaken to assess the relative magnitude of the non-genetic sources of variation in designed experiments using the MM protocol. Since phenotyping for LMA involves at least 4 linked processes (strictly not phases), non-genetic variation may arise from variation between plots (or pots) within a line, between tillers within a plot (or pot) within a line, between seeds within a tiller, and between OD values measured (hypothetically) on the same seed. In this paper we propose, as a simple extension to the MM protocol, an approach to measuring LMA expression based on multi-phase design principles and the analysis methods in Smith *et al.* (2006). We illustrate the proposed method with data from an LMA experiment and, using these results, undertake a simulation study to examine the relative accuracy of the proposed method to a method similar to the one in use.

## Materials and methods

This manuscript presents a detailed description of the design and analysis methods but limited detail on the (biological)

experimental materials and methods. A complete account of the latter can be found in Tan *et al.* (2010).

### Glasshouse design

An experiment to measure LMA expression was conducted in 2007–08 at Cobbity in New South Wales. In total, 188 doubled haploid (DH) lines derived from a cross between an advanced breeding line WW1842 and Whistler were used in this experiment. In total, 440 pots were arranged in 2 microclimate rooms with 220 in each room. The rooms were further subdivided into 2 blocks and 2 sides within each block. This resulted in 4 subblocks of 55 pots, each arranged in an  $11 \times 5$  rectangular array.

Parents and DH lines were then randomised to the pots in a restricted manner. First, replication varied between lines and parents (hereafter simply lines). There was a total of  $m_1 = 130$  lines with 2 replicates, and  $m_2 = 60$  lines with 3 replicates. Lines were assigned to pots so that each room contained either one (160) or two (30) pots of each line, randomised so that each side within each block contained only one pot of each of 55 lines. Statistically, therefore, the factor Room is resolvable and the model term Room : Block : Side is binary with respect to the factor Lines, where each term in this font represents either factors or model terms, where the latter are explained in more detail below.

Four plants were grown in each pot and at anthesis spikes from healthy plants were tagged and 26–28 days after anthesis the plants from each pot within an *induction cohort* were transferred to a cool-temperature room maintained at 10–12°C with 12 h of light and 12 h of darkness. After 8–10 days the plants were then returned to their original position in the microclimate rooms until the plants reached harvest ripeness. The term *induction cohort* is used to define the group of pots that were moved into the cool-temperature room on the same day. This is a natural blocking factor in the sense that it is likely to explain variation in the expression of LMA, but there is no design aspect to the allocation of pots to each cohort. The experimental units for the cool-temperature room are the same as the experimental units from phase 1, although there is a re-assignment and re-positioning of the units. The actual position within the cool-temperature room was not recorded and so we are not able to account for variation within the cool-temperature room.

Induction dates varied from 11/01/2008 to 28/01/2008 and tillers from a total of 425 pots were deemed sufficiently healthy to produce a reliable result. For the final phase of the experiment we aimed to assay ~20 grains from each pot, which meant that ~5 grains from each primary tiller from each plant would be used. In total there were only 1375 tillers (out of a potential 1700) which were used. The distribution of healthy grains (within those tillers which had grains) varied from 1 to 62 grains, with a median of 13. This presented challenges for the allocation of grains to the ELISA plates according to some statistical design, described below.

### Plate design

Each ELISA plate had 96 wells arranged in a 12 column  $\times$  8 row array to which the seeds from the 425 healthy pots were ultimately assigned. Using as many healthy seeds as possible, to a maximum of 20, yields over 7000 experimental units (wells) requiring a minimum of 73 plates. The number of seeds per pot actually used

in the final design varied from 1 to 22, with 25% of pots having less than 19 seeds. With the additional requirement of at least one *blank* (actually used for positive controls (Spica) in the final assay) per plate, the seeds for assay were ultimately distributed over 75 plates.

The potential for trend (including autocorrelation) within plates, the total number of experimental units, and the degree of imbalance in *treatment* replication all but eliminated the use of any standard design generator. Butler *et al.* (2008) demonstrate the role that nearest neighbour properties play in first-order autoregressive processes and derive a computationally efficient approximate optimality criterion that can be used to give an optimal allocation of treatments to experimental units, or in this case, pots to wells. The criterion actually used was an extension of Chan and Eccleston (2003) using row and column neighbours only. This criterion can be considered a first-order approximation to that in Butler *et al.* (2008), with the additional properties that it is faster to compute and has a known theoretical lower bound.

Consider a single  $8 \times 12$  plate to which seeds from  $v$  pots are to be allocated in some optimal way, and define  $v \times v$  neighbour concurrence matrices  $N_r$  and  $N_c$  in the row and column directions, respectively. The  $(i, j)$ th elements of these matrices ( $n_{r_{ij}}$ ,  $n_{c_{ij}}$ ) contain the number of times pots  $i$  and  $j$  occur as row or column neighbours, respectively. Nearest neighbour balance requires that the diagonal elements of  $N_r$  and  $N_c$  be zero (i.e. no self neighbours) and the off-diagonal elements be equal or differ by at most 1. Based on this we can define an optimality criterion:

$$\mathcal{N}^{opt} = \min \left[ \mathcal{N} = \sum_{i < j} (n_{r_{ij}}^2 + n_{c_{ij}}^2) \mid \text{trace}(N_r + N_c) = 0 \right]$$

that minimises the sums of squares of the off-diagonal elements of  $N_r$  and  $N_c$  subject to there being no self adjacencies.

Minimising the criterion  $\mathcal{N}$  is an approximation to minimising the average pair-wise variance between treatments (Butler *et al.* 2008), and is used in an iterative framework to arrive at an optimal design. Specifically, an initial design is repeatedly permuted by pair-wise treatment swaps and  $\mathcal{N}$  calculated at each step; the final design is accepted when  $\mathcal{N}^{opt}$  reaches the lower bound. The procedure is easily extended to multiple plates, possibly arranged in *resolvable* sets.

The initial layout for the LMA experiment was established as follows: 15 *super-blocks* (replicates) were defined, each of which contained 5 plates, which could also be considered as 5 (incomplete) blocks. Notionally, each replicate appears as a 40 row  $\times$  12 column array, where plates 1, 2, . . . , 5 occupy rows 1–8, 9–16, . . . , 33–40, respectively. The replicates are resolvable, as are the plates (within replicates) with respect to those pots with multiple seeds within replicates. Pots were allocated in repeats of 5 seeds (one for each plate, subject to available seed) within replicates in a cyclic fashion across replicates in a  $p$ -rep design (Cullis *et al.* 2006). For example, pots 1–12 appear in replicate 1 no less than 5 times, at least once in each plate, while the remaining pots appear just once, and are randomly distributed across the 5 plates; 5 blanks were also included, one per plate. Those pots replicated in replicate 1 appear just once in most of the other 14 replicates to achieve the target of  $\sim 20$  repeats. In replicate

2, pots 13–26 (plus blanks) are replicated 5 times while all other pots are replicated once, and so on. Some additional seeds of previously duplicated pots were included in some replicates to complete the layout, accounting for those pots with limited healthy seed.

The final randomised design was determined iteratively, minimising the overall criterion:

$$\mathcal{N}_o = \sum_{i=1}^{15} \mathcal{N}_i$$

where  $\mathcal{N}_i$  is the optimality criterion for the  $i$ th *super-block*. The initial design described above was permuted through a sequence of pair-wise treatment (pots) swaps, subject to the restrictions that treatments were not exchanged between replicates or between plates within replicates, establishing these as *pseudo strata* in the design. At each iteration,  $\mathcal{N}_o$  was calculated and compared with a theoretical lower bound. In addition, only designs binary with respect to replicate rows and columns were accepted, giving some protection against inefficiencies in the presence of plate row and plate column effects within replicates.

### Analysis

Fitting of the linear mixed model involves residual maximum likelihood (REML) estimation of the variance parameters. Given these estimates, we obtain empirical best linear unbiased estimates (E-BLUEs) of the fixed effects, and empirical best linear unbiased predictions (E-BLUPs) of the random effects. All analyses presented here have been conducted using the mixed model software ASReml-R (Butler *et al.* 2009).

Linear mixed models can be succinctly represented in text (and in software such as ASReml-R) by extending the symbolic model formulae of Wilkinson and Rogers (1973). The full linear mixed model has at least 2 *formula* objects, one to specify the fixed effects, the other to specify the random effects. These formulae contain model terms separated by a '+' operator.

The analysis of the data is dictated largely by the randomisation processes which were used in each phase of the experiment. The linear mixed model for the LMA experiment is therefore given by

```
fixed = od ~ 1,
random = ~ ID + GoP + Plate + Room + Room : Block +
Room : Block : Side + Induction + Pot + Pot : Tiller +
Plate : PlateRow + Plate : PlateCol
rcov = ~ units
```

where the ':' operator forms interactions between terms, and factor names (e.g. GoP denoting *Group of Plates*) are capitalised to distinguish them from variates (e.g. od). The *fixed* formula has 3 components: the dependent variable od, the '~' symbol, meaning *is modelled by*, and the terms in the fixed part of the model (where 1 represents the overall mean). The *random* formula specifies those terms which are to be fitted as random effects, and the *rcov* formula specifies the variance model for the residuals. In this case we assume that the residuals are independent and follow a Gaussian distribution with mean zero and variance  $\sigma^2$ . The factors have obvious meaning, in the sense that Room, Block, Side, Pot, and Tiller define the blocking structure

and plants within pot with (2, 2, 2, 425, 4) levels respectively for the glasshouse phase of the experiment. The blocking structure for the ELISA phase of the experiment is described by the factors *GoP*, *Plate*, *PlateRow*, and *PlateCol*, each with (15, 75, 8, 12) levels, respectively. Note that defining *Plate* with 75 levels obviates the need to fit the model term  $GoP : Plate$ . Lastly, *Induction* is included as a blocking factor with 14 levels for the cool-room phase of the experiment.

We assume that the random effects for each of the model terms in the *random* formula are independent and follow a Gaussian distribution with mean zero and variance denoted by  $\gamma_j \sigma^2$  where  $j = 1, \dots, 11$  indexes the model terms in the order presented in the *random* formula. If the aim of the analysis was to estimate variation between DH lines alone, as a precursor to QTL identification, say, then we would define an additional factor with 3 levels, taking value 1 and 2 for the parents and 3 for the DH lines, and include this factor in the linear mixed model as a fixed effect. The focus of this analysis, however, is the estimation of the non-genetic sources of variation with a view to exploring the efficiency of LMA screening based on the MM protocol, and for simplicity we chose to omit this term.

#### *Relative efficiency of MM protocol and multi-phase designs*

The performance of the multi-phase design approach used here relative to a protocol similar to the MM protocol was assessed using a simulation study for a case similar to that described above. An experiment with 110 genotypes was considered, where the glasshouse experiment comprised 220 pots, each pot containing 2 plants. In each pot it was assumed that grain was harvested from one primary tiller, and that for each genotype, there were 10, 10, 20, and 40 grains for each of the 4 pot  $\times$  tiller combinations (i.e. 2 pots per genotype  $\times$  2 plants per pot). For simplicity, the glasshouse design was considered a randomised block design with 2 complete blocks, and induction cool-room effects were assumed negligible (consistent with the results discussed below).

For the ELISA plate phase of the simulated experiment, the following two approaches to design and analysis were considered.

**Multi-phase design:** Two grains were sampled for each genotype, pot, and plant combination, giving a total of 8 grains per genotype. The design protocol described above was followed by defining 440 phase 2 samples (110 genotypes by 2 pots by 2 tillers) and assigning these to 10 ELISA plates according to a resolvable nested row-column design, except that only columns 1–11 were used on each plate for compatibility with the MM design described below. This design required a total of 10 plates, with plates 1–5 being duplicate 1 and plates 6–10 duplicate 2. Rows 1–4 and rows 5–8 in column 12 were assigned to blank or positive controls, respectively.

**MM-like protocol:** The actual MM protocol uses 8 wells for each genotype, with each well being a composite sample of 5 grains. It was not possible to directly compare the multi-phase approach with the MM protocol as the data at hand were based on single seed samples; it is most likely that the sampling distribution of composite samples would differ from that described later for single seeds. The design and analysis

approaches used for the MM protocol on simulated data were therefore based on the characteristics estimated from the analysis of the LMA (single seed) experiment.

In total, 8 grains per genotype were sampled, but either completely random sampling or stratified sampling was used, the latter ensuring that each plant within each pot was sampled twice. The 8 wells in columns 1–11 on plate 1 were assigned the 8 grains for genotypes 1–11, 8 wells in columns 1–11 on plate 2 to the 8 grains for genotypes 12–22, and so on. Rows 1–4 and rows 5–8 in column 12 were assigned to blank or positive controls, respectively.

In total, 400 datasets were simulated according to the following scheme. For those plate wells assigned to genotypes, genetic and non-genetic effects were included for terms for which the percentage of variation was greater than 1% from the LMA experiment analysis.

The symbolic representation of the model for the simulation experiment is given by:

```
fixed = od ~ 1,
random = ~ ID + GoP + Plate + Pot + Pot : Tiller +
Plate : PlateRow + Plate : PlateCol
rcov = ~units
```

noting that this model is consistent with the randomisation processes used in the design of the simulation experiment.

Data for the wells assigned to blanks were simulated from the regression of the E-BLUEs for the blanks on the sum of the E-BLUPs for the *GoP* and *Plate* effects for each of the plates which contained a blank in the LMA experiment. Additional (residual) variation for these blanks was then added by sampling from an independent realisation from a Gaussian distribution with mean zero and variance 0.014. The value of 0.014 was chosen as this was the estimate of residual (i.e. within plate) variation for a series of ELISA experiments which contained 4 wells of blanks for each of 12 plates (D. Mares, pers. comm.).

Three methods of analysis were applied to each of the simulated datasets: the multi-phase approach (MP), modified MM without stratified sampling (modMM), and modified MM with stratified sampling (modMMS).

## Results

### *Glasshouse phase*

Figure 1 presents the histogram of the range (in days) in induction date between pots within lines. This shows that, perhaps unexpectedly, pots within many lines were induced on different days. Specifically, pots from 31 lines were induced on the same day, 132 on two different days, and 27 on three different days, with the largest range being 11 days. Pots in microclimate room 2 were in general slower to mature.

### *Analysis of LMA experiment*

Due to unavailability of seeds and unexpected mishaps during the final phase of the experiment, 19 blanks were included on 17 ELISA plates. There was at least one well on each plate assigned to a positive control, with two plates having an additional positive control.



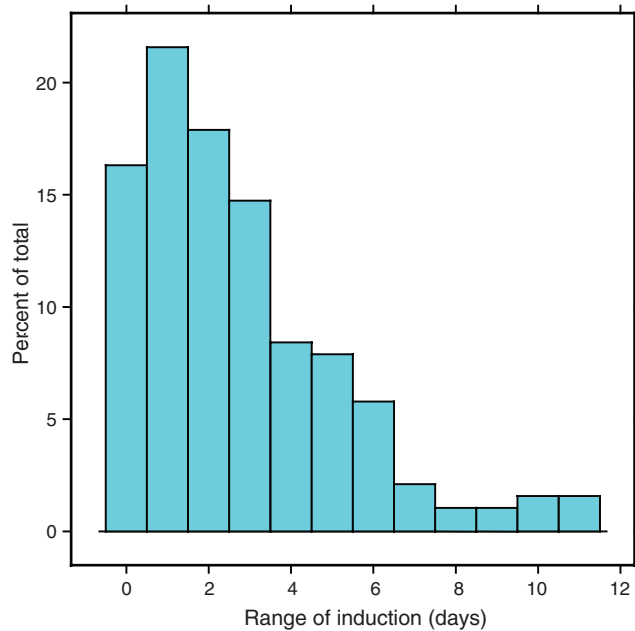


Fig. 1. Histogram of range in induction (days) for each of the DH/lines.

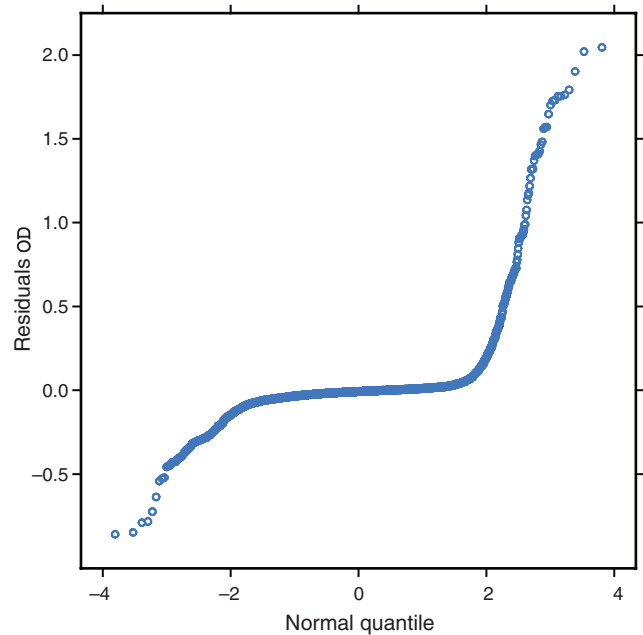


Fig. 2. Q-Q plot for the residuals from the linear mixed model analysis of od.

The range of od values for the blanks was (0.050, 0.077), while for the positive control the range was (1.012, 2.168). This seems typical of the range encountered in similar LMA phenotyping experiments.

Initially the model was fitted to the untransformed od data. Since the data are extensive, we followed McCullagh and Nelder (1994) and undertook detailed examination of systematic departures from the assumptions underlying the model. Figure 2 shows the Q-Q plot (qqmath, R Development Core Team 2008) of the residuals from the initial model. This diagnostic suggests that the assumption of a Gaussian distribution for the residuals is totally untenable. The sample histogram (Fig. 3) of the residuals of 6 lines further illustrates this problem. The 6 lines were selected from the final analysis as the most and least LMA expressive lines. The distribution of residuals for the parental and non-expressing DH lines (WW10897, WW12677, and WW12649) is clumped close to zero, while the distribution of residuals for the expressing DH lines (WW10911, WW12707, and WW10910) is strongly skewed.

A range of transformations from the power family was considered. The scaled inverse cubed transformation given by  $k/od^3$  appeared to provide a reasonable scale for the analysis. Figures 4 and 5 present the Q-Q plot and sample histograms on the transformed scale for the above set of DH lines and parents. The plots indicate that the assumption of normality is reasonable for the transformed optical density, denoted by  $to_d$ .

Table 1 presents a summary of the REML estimates of the variance components  $\hat{\gamma}_j\hat{\sigma}^2$  and residual variance  $\hat{\sigma}^2$  for  $to_d$ . The most striking feature is the substantial amount of explainable non-genetic variation, mostly arising from variation between plates (i.e. the sum of  $GoP$  and  $Plate = 38.9\%$ ), although other model terms associated with the ELISA phase of the experiment also contribute to reducing residual variation. Importantly, the other largest source of variation is associated with tillers between pots.

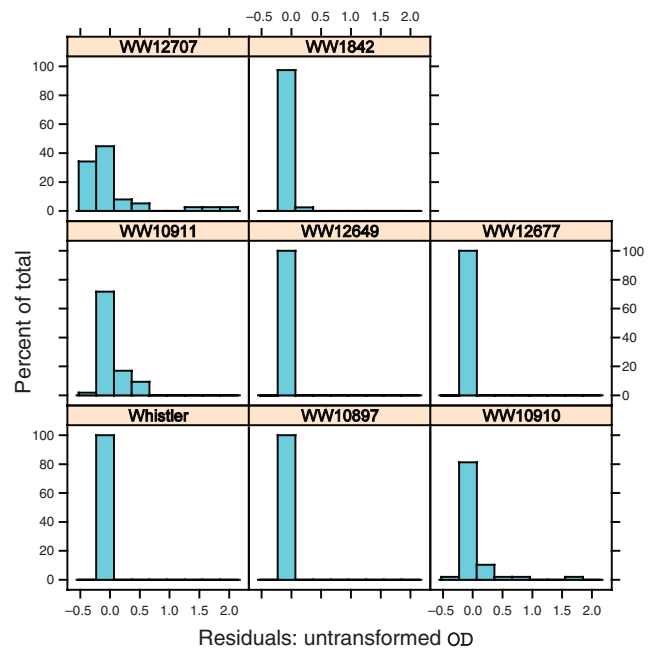


Fig. 3. Sample histograms for six selected DH/lines and parents from the analysis of od.

This implies that seeds from the primary tiller from the same plant tend to be more similar (in terms of LMA expression) than those from primary tillers from different plants. This issue is discussed further below.

Figure 6 presents a plot of the E-BLUPs of the total effects for plate (as the sum of the E-BLUP for  $GoP$  and  $Plate$ ) against plate

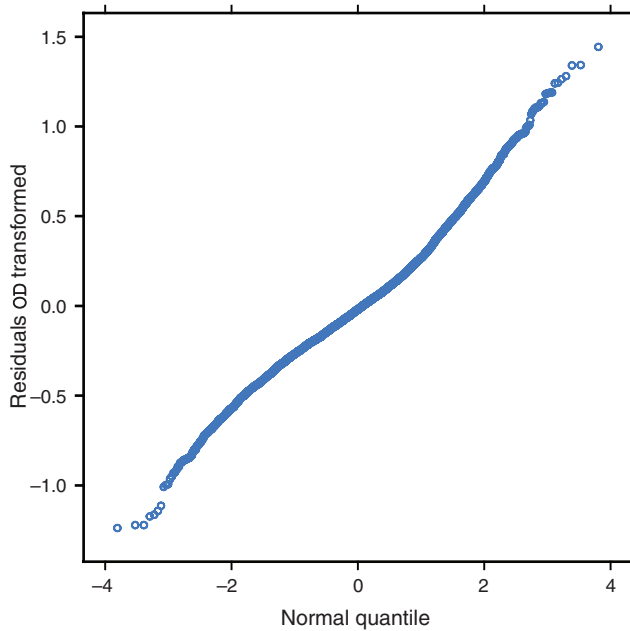


Fig. 4. Q-Q plot for the residuals from the linear mixed model analysis of tod.

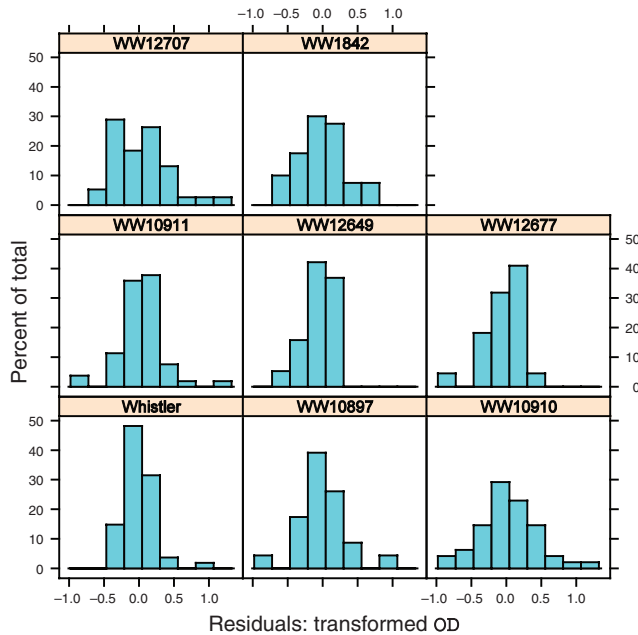


Fig. 5. Sample histograms for six selected DH/lines and parents from the analysis of tod.

number. There is a consistent loss in activity for the plates until plate 28, beyond which the oscillations persist but there is an arrest in this decline. This feature is related to a temperature-dependent decline in activity: the binding steps in the ELISA analysis for the first 28 plates were conducted on the laboratory bench from late summer through to early winter, while this process was conducted

Table 1. Summary of the REML estimates of the variance components for tod from the randomisation-based linear mixed model

| Term                | Component          | Percent |
|---------------------|--------------------|---------|
| ID                  | 0.0354             | 10.6    |
| GoP                 | 0.0539             | 16.2    |
| Plate               | 0.0691             | 20.7    |
| Room                | 0.000 <sup>B</sup> | 0.0     |
| Room : Block        | 0.000 <sup>B</sup> | 0.0     |
| Room : Block : Side | 0.0008             | 0.2     |
| Induction           | 0.0017             | 0.5     |
| Pot                 | 0.0060             | 1.8     |
| Pot : Tiller        | 0.0192             | 5.8     |
| Plate : PlateRow    | 0.0229             | 6.9     |
| Plate : PlateCol    | 0.0103             | 3.1     |
| Residual            | 0.1140             | 34.2    |

<sup>B</sup>Values signify that the REML estimate of the variance component has been constrained on the boundary (i.e. zero).

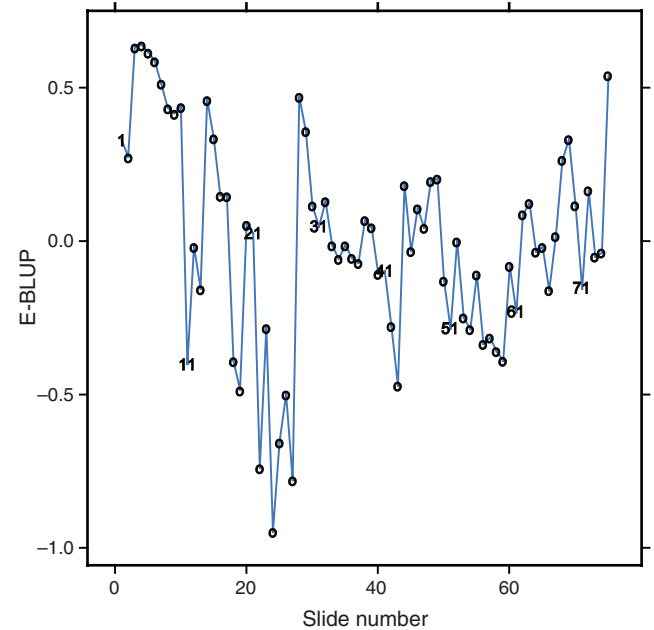


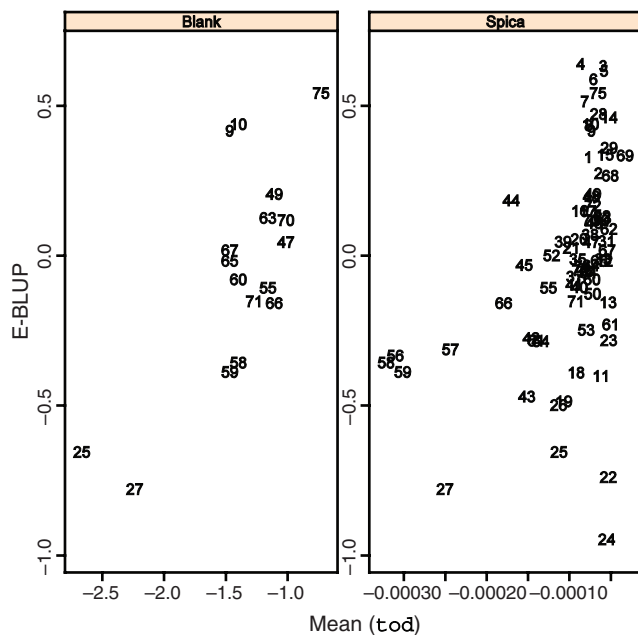
Fig. 6. Plot of E-BLUPs for plates (slides) from the analysis of tod.

in a constant-temperature incubator at 25°C for the remaining plates (slides).

Figure 7 presents scatter plots of the E-BLUPs of the total effects of the plate against plate mean values of tod for blank and positive controls, respectively. There is reasonable agreement within the subset of 19 plates with a blank, but little agreement for the positive control for all 75 plates.

*Relative efficiency of MM/protocol and multi-phase designs*

The mean squared correlation (over the 400 simulations) between the true BLUPs and the E-BLUPs for the MP method, and the genotypic mean of the adjusted data for the modMM and modMMS methods, were 0.596, 0.221, and 0.229,



**Fig. 7.** Plot of E-BLUPs for blank and positive controls (Spica) against the plate mean values of tod.

respectively. The squared correlation is analogous to the generalised heritability measure of Oakey *et al.* (2007).

### Discussion

The use of a valid statistical design for LMA phenotyping has clearly demonstrated that there exist substantial non-genetic sources of variation which need to be accounted for if we are to improve the accuracy of phenotyping. The small simulation study has shown that the use of a multi-phase approach to design and analysis is superior to an approach which does not use a valid statistical design and efficient statistical analysis. It could be then inferred that the common practice for LMA testing using the MM protocol may lead to inferior estimates of variety performance and thence poor selection decisions. For single seed assays the accuracy of selection is vastly improved with the use of a multi-phase design, with true replication at all stages and a sound mixed model analysis in which appropriate models are applied at both the genetic and error levels. We have not been able to directly compare the MM protocol as this involves the use of composite samples, although standard statistical arguments would suggest that similar results would apply. Our simulation study is limited in that we have not investigated the relative accuracy for a range of design and variance parameters. Given the substantial improvement, and the *nil* costs involved with implementation of our approach, we see little need for such a

comprehensive study. It would be valuable, however, to undertake a similar study to compare a multi-phase approach with the MM protocol using composited samples.

### Acknowledgments

We gratefully acknowledge the financial support of the Grains Research and Development Corporation of Australia and the Value Added Wheat CRC, Australia (now terminated). We thank Neil Coombes (Wagga Wagga Agricultural Institute) for the elegant experimental design of the glasshouse phase of the experiment, and Alison Smith and Ari Verbyla for helpful discussions.

### References

- Butler DG, Cullis BR, Gilmour AR, Gogel BJ (2009) ASReml-R reference manual, release. 3. Technical Report, Queensland Department of Primary Industries.
- Butler DG, Eccleston JA, Cullis BR (2008) On an approximate optimality criterion for the design of field experiments under spatial dependence. *Australian and New Zealand Journal of Statistics* **50**, 295–307. doi: 10.1111/j.1467-842X.2008.00518.x
- Chan BSP, Eccleston JA (2003) On the construction of nearest-neighbour balanced row-column designs. *Australian and New Zealand Journal of Statistics* **45**, 97–106. doi: 10.1111/1467-842X.00263
- Cullis BR, Smith AB, Coombes NE (2006) On the design of early generation variety trials with correlated data. *Journal of Agricultural, Biological & Environmental Statistics* **11**, 381–393. doi: 10.1198/108571106X154443
- Cullis BR, Smith AB, Panozzo JF, Lim P (2003) Barley malting quality: are we selecting the best? *Australian Journal of Agricultural Research* **54**, 1261–1275. doi: 10.1071/AR02195
- McCullagh P, Nelder JA (1994) 'Generalized linear models.' 2nd edn (Chapman and Hall: London)
- McIntyre GA (1955) Design and analysis of two-phase experiments. *Biometrics* **11**, 324–334. doi: 10.2307/3001770
- Mrva K, Mares DJ (2001) Induction of late maturity  $\alpha$ -amylase in wheat by cool temperature. *Australian Journal of Agricultural Research* **52**, 477–484. doi: 10.1071/AR00097
- Oakey H, Verbyla AP, Cullis BR, Wei X, Pitchford WS (2007) Joint modelling of additive and non-additive (genetic line) effects in multi-environment trials. *Theoretical and Applied Genetics* **114**, 1319–1332. doi: 10.1007/s00122-007-0515-3
- R Development Core Team (2008) 'R: A Language and Environment for Statistical Computing.' (R Foundation for Statistical Computing: Vienna, Austria)
- Smith AB, Lim P, Cullis BR (2006) The design and analysis of multi-phase plant breeding experiments. *Journal of Agricultural Science* **144**, 393–409. doi: 10.1017/S0021859606006319
- Tan MK, Verbyla AP, Cullis BR, Martin P, Milgate AW, Oliver JR (2010) Genetics of late maturity  $\alpha$ -amylase in a doubled haploid wheat population. *Crop & Pasture Science* **61**, (in press).
- Wilkinson GN, Rogers CE (1973) Symbolic description of factorial models for analysis of variance. *Applied Statistics* **22**, 392–399. doi: 10.2307/2346786

Manuscript received 21 April 2009, accepted 21 August 2009