Published August 17, 2015

RESEARCH

Appraisal of Test Location and Variety Performance for the Selection of Tuber Yield in a Potato Breeding Program

M. F. Paget,* L. A. Apiolaza, J. A. D. Anderson, R. A. Genet, and P. A. Alspach

ABSTRACT

Differences in trait responses of genotypes across environments, or genotype × environment interactions (G × E), hinder the progress of genetic improvement. Characterization of these effects helps to determine breeding strategies and improve resource allocation in a cultivar development program. This study used historical multienvironment trial (MET) data (34 trials in five locations) for the analysis of marketable yield of advanced selections in a New Zealand potato (*Solanum tuberosum* L.) breeding program. A factor analytic (FA) model was used for the analysis of these MET data. Contrasts based on the environmental loadings were observed between the program's main trial locations in the North Island (Pukekohe) and the South Island (Lincoln), indicating that these locations optimized differentiation between genotypes in terms of G × E effects. Genetic correlation estimates between trial environments were mostly moderately high (>0.5) to high (>0.8) and ranged from zero to positive with a maximum coefficient of 0.97, suggesting that quantitative (rescaling) rather than qualitative (crossover) G × E effects were of greater importance. A number of newly developed varieties were shown to have higher genetic yield potential than older and established commercial cultivars but did not necessarily show better yield stability over the locations tested.

M.F. Paget and R.A. Genet, New Zealand Institute for Plant & Food Research Limited, Lincoln, Christchurch 8140, New Zealand; L.A. Apiolaza, School of Forestry, Univ. of Canterbury, Christchurch, New Zealand; J.A.D. Anderson, New Zealand Institute for Plant & Food Research Limited, Pukekohe 2676, New Zealand; P.A. Alspach, New Zealand Institute for Plant & Food Research Limited, Motueka 7198, New Zealand. Received 26 Nov. 2014. Accepted 27 Mar. 2015. *Corresponding author (mark.paget@plantandfood.co.nz).

Abbreviations: AMMI, additive main effects and multiplicative interaction; BLUE, best linear unbiased estimator; BLUP, best linear unbiased predictor; FA, factor analytic; G × E, genotype × environment interaction; G × M, genotype × management interaction; LIN, Lincoln; MET, multienvironment trial; MME, mixed-model equation; MW, Manawatu; OHA, Ohakune; PEV, prediction error variances; PFR, New Zealand Institute for Plant & Food Research Ltd.; PK, Pukekohe; RCB, randomized complete-block; REML, residual maximum likelihood; WAI, Waikato.

When testing selection candidates over multiple environments, uncertainty in the estimates of genotype values increases with the magnitude of G × E. This increases the difficulty of identifying superior genotypes and compromises genetic progress from selection (e.g., Annicchiarico, 2002; Bos and Caligari, 2008; DeLacy et al., 1996a,b). A better understanding of G×E effects within a MET testing regime allows a reevaluation of resource allocation and selection strategy in a breeding program. The type and extent of G × E is of particular interest to plant breeders as the characterization of environments will help, in part, to define selection strategies. For example, measures of quantitative G × E (heterogeneity of variance or the scale

Published in Crop Sci. 55:1957–1968 (2015).
doi: 10.2135/cropsci2014.11.0801

change of genotypes) between test locations may help to determine that some locations offer little extra information in terms of differentiating genotypes, that is, a certain degree of environment duplication is present, demanding scrutiny of its opportunity cost. Consideration should therefore be given for such locations to be dropped from the testing schedule. Alternatively, the presence of qualitative G × E (crossover interaction, or the rank change of genotypes) may determine that separate breeding programs for subsets of locations are necessary to select for specific adaptation (e.g., Atlin et al., 2000).

There are numerous statistical approaches to model G × E effects in plant breeding, which have been extensively reviewed by various authors (e.g., Crossa et al., 2010; DeLacy et al., 1996a; Fox et al., 1997; van Eeuwijk et al., 2005). In general, these methods are based on univariate or multivariate methods that vary in their degree of complexity and the information that they provide. Over recent years, flexible multivariate multiplicative methods have found favor, including the additive main effects and multiplicative interaction (AMMI) model (Crossa et al., 1991; Gauch and Zobel, 1988). This approach carries out singular value decomposition on the matrix of the two-way table of G × E effects, whereby each is modeled as the product of a genotypic score and an environmental loading. Additional multiplicative (bilinear) terms are considered if they improve model fit. The AMMI model is classified as one of several types of general linear-bilinear model. More recently, a multiplicative mixed-modeling approach using factor analysis, which is considered as another class of linear-bilinear model and a mixed-model analogy of the AMMI fixed-effect model (Piepho, 1997, 1998; Smith et al., 2001, 2005), has been used to evaluate MET data. Heavy attrition of breeding lines at each stage of a MET series of breeding trials is typical of plant breeding programs and the incomplete nature of such data is better dealt with by residual maximum likelihood (REML)-based procedures. Further, there has been growing trend amongst crop breeders, following their animal- and tree-breeding counterparts, to treat genotypes as random effects, at least in the early stages of trials. Genotype value predictions are shrunk toward the mean to allow for the uncertainty surrounding the distribution of random effects and there is greater flexibility in analyzes with, for example, inclusion of a coefficient of coancestry matrix to take account of relationships among genotypes (e.g., Crossa et al., 2006; Oakey et al., 2006; Piepho et al., 2008; Smith et al., 2005).

Potato crops are known to show variability in seasonal yields over both regional and field scales (e.g., Po et al., 2010; Redulla et al., 2002), suggesting the need for extensive MET evaluation of selection candidates in breeding programs. Studies into G × E in potato breeding studies have generally been restricted to a limited number of advanced clones and cultivars, (e.g., Affleck et al., 2008; Cotes et al., 2002; Tai and Coleman, 1999). Typically, breeders require information on a larger number of genotypes for inference of performance and stability to help to make more informed selections earlier in a breeding program, and there is a desire to distribute clones across multiple locations as early as possible (Haynes et al., 2012; McCann et al., 2012). The potato breeding program of the New Zealand Institute for Plant & Food Research Ltd. (PFR) targets the selection of genotypes that perform well across all major potato production regions, that is, those that are broadly adapted within a New Zealand context. The multivariate analysis of MET data provides an opportunity to assess the extent and type of G × E present in historic potato yield trials, which may go some way to guide resource allocation for METs and the testing strategy for genotype selection in future by evaluating environments as well as genotypes.

The study takes a mixed linear-bilinear modeling approach to measure G × E effects and stability of genotypic responses across the major potato production regions of New Zealand for potato yield. It uses data collected from a series of historic yield trials, comprising advanced selections, over a number of year–location combinations (environments). A FA model is used to measure the relationships between genotype performance and environments and to characterize environments. The yield performance and the stability of recent advanced selections from the PFR breeding program are also compared with those of older established cultivars that are currently and, in some cases, widely grown in New Zealand. The study aims to evaluate test locations that are used for the selection of broadly adapted cultivars to improve selection efficiency. Potato varieties are also evaluated to assess the genetic progress of tuber yield improvement in the New Zealand potato breeding program.

MATERIALS AND METHODS

Data

The data for study were collected from breeding trials of the PFR potato genetic improvement program. Trials were performed between the years 1999 and 2005 (harvest years) at five sites that represent the major potato-producing regions in New Zealand (Fig. 1): Pukekohe, South Auckland (37° 12′ S 174° 57′ E, 141 m asl); Matamata, Waikato (37°48′ S, 175°45′ E, 53 m asl); Ohakune, Central (39°24′ S, 175°24′ E, 741 m asl), Palmerston North, Manawatu (40°21′ S,175°37′ E, 30 m asl), and Lincoln, Canterbury (43°39′ S, 172°28′ E, 14 m asl). For interest, a summary of basic climate data (rain, temperature, and humidity) for the main crop growing season (November through March) at each location is provided in Table 1. At the time of writing, New Zealand regional climate data can be accessed via http://cliflo.niwa.co.nz, and the national soil database can be found at http://smap.landcareresearch.co.nz. Pukekohe and Lincoln are the main potato research sites and, along with Manawatu, can be categorized as PFR on-station trials, as trials are all managed on PFR research farms. Waikato and Ohakune are off-station

Figure 1. Location of the New Zealand Institute for Plant & Food Research Ltd. potato breeding trials in New Zealand. Main research sites are at Pukekohe and Lincoln (solid triangles).

trials, as these are managed within a commercial potato crop. The 34 trials, synonymous with environments (year–location combinations), were clonal stages four and five (C4 and C5 respectively) of main crop tuber yield trials. Target planting dates, harvest dates, and the number of genotypes entered into each trial are shown in Table 2. The C4 trials were only performed in Pukekohe (PK), Manawatu (MW), and Lincoln (LIN). Selected genotypes from the C4 stage were entered into further main crop trials at stage C5 that also included locations Waikato (WAI) and Ohakune (OHA) as well as PK, MW, and LIN. For each season in the Waikato region, there were two trials: an early trial and a late trial. The late trial represented the regional practice of the winter harvesting of potatoes with the crop maintained in the ground for approximately 200 d.

Trials at all North Island locations (PK, MW, WAI, and OHA) were based on Latinized row–column designs of varying size with 20 to 200 genotypes replicated twice (C4 trials) or three times (C5 trials). Each genotype occurred once, at most, in both rows and columns across a trial of rectangular plots. A typical plot was made up of 12 tubers in total, planted in a six by two arrangement. The South Island trials (LIN) were randomized complete-block (RCB) designs typically of 200 to 300 genotypes replicated three times. Plots were made up of 12 tubers in total, planted in a 12 by one arrangement.

Each plot was harvested, and yield was recorded as marketable tuber yield after undersized (<80 g) and defective tubers had been removed. Defective tubers, for example, may have secondary or abnormal growths, rot, or excessive greening. Plot yield was converted to metric tonnes per hectare for analyzes of marketable tuber yield. Marketable yield, as described and hereon in referred to simply as tuber yield, is usually considered to be the total economic yield, as there is often no economic value attributed to undersized but otherwise sound tubers. Although there were a total of 1619 genotypes represented in the data, many were lost after only 2 yr of testing through the discarding of unsuitable candidates. Genotypes of particular interest were tested in at least four locations over 3 yr and were made up of both New Zealand and international cultivars as well as advanced clonal selections. Many of the international cultivars are popular commercial cultivars widely grown for fresh and processing production in New Zealand.

For clarification, the term variety in this paper is used generically and can refer to both clonal selections and cultivars. It is also used synonymously with genotype. The term cultivar is used to describe a variety that has been officially named and commercially released and is or has been previously protected under plant variety rights.

Table 1. Seasonal climate data (averages for November through March) from 1998 to 2005†.

Trial location	Rain	Temperature	Total degree days	Relative humidity
	mm	°C	10°C	%
Pukekohe (PK)	452	17.6	1145	85
Manawatu (MW)	393	16.4	966	78
Ohakune (OHA)‡	–	–	–	–
Waikato (WAI)	303	16.9	1027	86
Lincoln (LIN)	286	15.3	689	77

† Data retrieved from http://cliflo.niwa.co.nz (accessed 18 Mar. 2015).

‡ Data not available.

Statistical Model

For illustration, the general form of the linear mixed model for the jth trial (environment) was: $\mathbf{y}_j = \mathbf{X}_j m_j + \mathbf{Z}_{1j}\mathbf{b}_j + \mathbf{Z}_{2j}\mathbf{g}_j + \mathbf{e}_j$, where $\mathbf{y}_j$ is the vector of yield observations, m_j denotes the fixed effects of trial means, $\mathbf{b}_j \sim \left(0, \mathbf{I}\sigma^2_{\mathrm{b}_j}\right)$ and $\mathbf{g}_j \sim \left(0, \mathbf{I}\sigma^2_{\mathrm{g}_j}\right)$ are vectors of random (nongenetic) design factors and genetic effects, respectively, $\mathbf{e}_j \sim \left(0, \mathbf{I}\sigma^2_{\mathrm{e}_j}\right)$ is the vector of random error terms, and $\mathbf{I}$ is an identity matrix. The terms $\mathbf{X}_j$, $\mathbf{Z}_{1j}$, and $\mathbf{Z}_{2j}$ are known incidence matrices of zeros and ones that relate the phenotypic observations to their corresponding vectors. The nongenetic factors were trial blocking factors and included the rows and columns of the incomplete block designs (PK, MW, WAI, and OHA) and the complete blocks of the RCB designs (LIN). For MET analysis, the mixed-model equations (MME) were constructed to analyze the vector of observations for the 34 trials ($\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_{34}$) tested from 1999 to 2005. The joint distribution of the random terms was assumed to follow a multivariate normal distribution with means and covariances defined by the following:

$$\begin{bmatrix}\mathbf{b}\\ \mathbf{a}\\ \mathbf{e}\end{bmatrix} \sim \mathbf{N}\left(\begin{bmatrix}\mathbf{0}\\ \mathbf{0}\\ \mathbf{0}\end{bmatrix}, \begin{bmatrix}\mathbf{B}_0 \otimes \mathbf{I}_\mathbf{b} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{G}_0 \otimes \mathbf{I}_\mathbf{g} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{R} \otimes \mathbf{I}_\mathbf{e}\end{bmatrix}\right)$$

where $\mathbf{0}$ are null matrices, $\mathbf{B}_0$, $\mathbf{G}_0$ and $\mathbf{R}$ are covariance matrices for design factors (row and column or block), genetic, and residual

Table 2. Summary of target potato planting and harvest dates (1999–2005), canopy days, days from planting to harvest, and number of lines tested per trial. All figures are approximate.

Trial location	Planting date	Canopy days†	Harvest date	Days to harvest	Clones per trial
Pukekohe (PK)	1 November	140	20 May	200	100–200
Manawatu (MW)	25 October	140	10 April	170	40–120
Ohakune (OHA)	10 November	140	1 June	200	20–30
Waikato E-L (WAI)‡	1 Oct., 10 Nov.	120, 140	1 March, 1 June	150, 200	20–30
Lincoln (LIN)	10 October	130	10 April	180	200–300

† Canopy days are the number of days from planting to canopy loss (by natural senescence, chemical desiccation, or mechanical means).

‡ E, early trial; L, late trial.

effects, respectively, and $\otimes$ is the direct (Kronecker) product. The matrix $\mathbf{B}_0$ is a diagonal matrix of (nongenetic) scaled identity matrices, the variance structure of plot error effects $\mathbf{R}$ are assumed to be block diagonal, and $\mathbf{I}$ are identity matrices. The assumption was that the variance matrix for the genotype effects has the separable form $\mathbf{G}_g = \mathbf{G}_0 \otimes \mathbf{I}$, where $\mathbf{G}_0$ is the matrix of genetic variances and covariances between environments and $\mathbf{I}$ is an identity matrix. The term $\mathbf{I}$ can be replaced by $\mathbf{A}$ with the coefficient of coancestry between genotypes as elements (the numerator relationship matrix) as shown, for example, by Oakey et al. (2006), Crossa et al. (2006), and Kelly et al. (2009).

The form of the genetic variance matrix, $\mathbf{G}_0$, fitted was factor analytic (**FA*k***) (Piepho, 1998; Smith et al., 2001), allowing for heterogeneity of genetic variance and covariance between environments. The loadings (as the leading principal components) and residuals, or specific variances, are given by: $\mathbf{G}_0 = \mathbf{\Lambda}\mathbf{\Lambda}' + \varphi$; where $\mathbf{\Lambda}$ is a ($j \times k$) matrix of environmental loadings ($\mathbf{\Phi}_{jk}$) and φ is a ($j \times j$) diagonal matrix of specific variances:

$$\mathbf{\Lambda} = \begin{bmatrix} \mathbf{\Phi}_{11} & \mathbf{\Phi}_{12} & \cdots & \mathbf{\Phi}_{1k} \\ \mathbf{\Phi}_{21} & \mathbf{\Phi}_{22} & \cdots & \mathbf{\Phi}_{1k} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{\Phi}_{j1} & \mathbf{\Phi}_{j2} & \cdots & \mathbf{\Phi}_{jk} \end{bmatrix}$$

This is considered as a (random) regression of genotype and G × E on (k) latent covariables (the environmental loadings) with a separate slope for each genotype (the genotype scores) and separate or common intercepts, depending on whether genotype main effects and G × E are combined or fitted separately. There was no distinction between genotype main effects and G × E effects in the present analysis.

A rotation is applied to the matrices of genotypic scores and environmental loadings to obtain a principal component solution as a more useful interpretation (Smith et al., 2001). For genotypes, when there is no distinction between genotype main effects and G × E effects (genotype effects nested within environments), the first score factor mainly represents genotype performance and the second score factor illustrates genotype stability (Stefanova and Buirchell, 2010), which is considered as a dynamic stability if yield performance in each environment is is parallel to the mean response of the tested set of genotypes (Annicchiarico, 2002). For more detailed information on the statistical method, Smith et al. (2001) provide a comprehensive account of the use of genetic variance structures in the analysis of crop trial data, while Cullis et al. (2010) demonstrate the application of FA models to MET data. The analyzes of the data were undertaken using R (R Development Core Team, 2012) with the mixed models fitted using ASReml-R (Butler et al., 2009).

The best linear unbiased estimators (BLUEs) of the fixed effects, $\hat{m}$, and best linear unbiased predictors (BLUPs) of the random effects ($\hat{b}$ and $\hat{g}$) were obtained from the solutions to the MME (e.g., Lynch and Walsh, 1998; Smith et al., 2001, 2005). Variance components are unknown and were estimated from the data using REML (Patterson and Thompson, 1971). Empirical genetic values are therefore a result of applying variance components in the MME that are estimated from the data in an iterative process, so providing empirical BLUEs of fixed effects and empirical BLUPs of random effects. The 95% confidence intervals of BLUPs of genotype values were calculated from the prediction error variances (PEV), with the PEV obtained from the inverse of the coefficient matrix of the mixed-model equations for random genotype (variety) effects.

From the results of FA modeling, a heatmap was used to illustrate the genetic relationships and G × E across the trials. First- and second-factor environmental loadings were plotted on the correlation scale as a uniplot to group environments according to their genetic correlations. This is described by Cullis et al. (2010), offering greater clarity than biplots as genotype scores and environmental loadings are plotted on separate graphs, which is more favorable for plant breeding trials that typically deal with large numbers of genotypes and environments. The interpretation of the environmental loadings, that is, the directions and projections of the vectors, illustrated in such plots from a FA2 model is analogous to AMMI and other linear-bilinear models with two components (e.g., Fox et al., 1997; Yang et al., 2009). The squared length of a vector for an environment indicates the proportion of genetic variance modelled for that environment by the two factors and its particular relationships with another vector of an environment is the cosine of the angle between the two vectors, providing the genetic correlation arising from the two factors (Smith et al., 2001). Standard errors of the environmental loadings were obtained by jackknifing; each environment was deleted in turn, data reanalyzed, and standard errors obtained from the resampled results.

RESULTS

The mean tuber yields for plots in trials ranged from 26 (LIN-C4-05) to 70 t ha^{-1} (WAI-C5-03L), with a maximum plot yield of 139 t ha^{-1} (LIN-C4-02) (Fig. 2). Lincoln trials were routinely irrigated, which probably resulted in a greater opportunity for genotypes to better express their yield potential, as their mean yields were consistently high

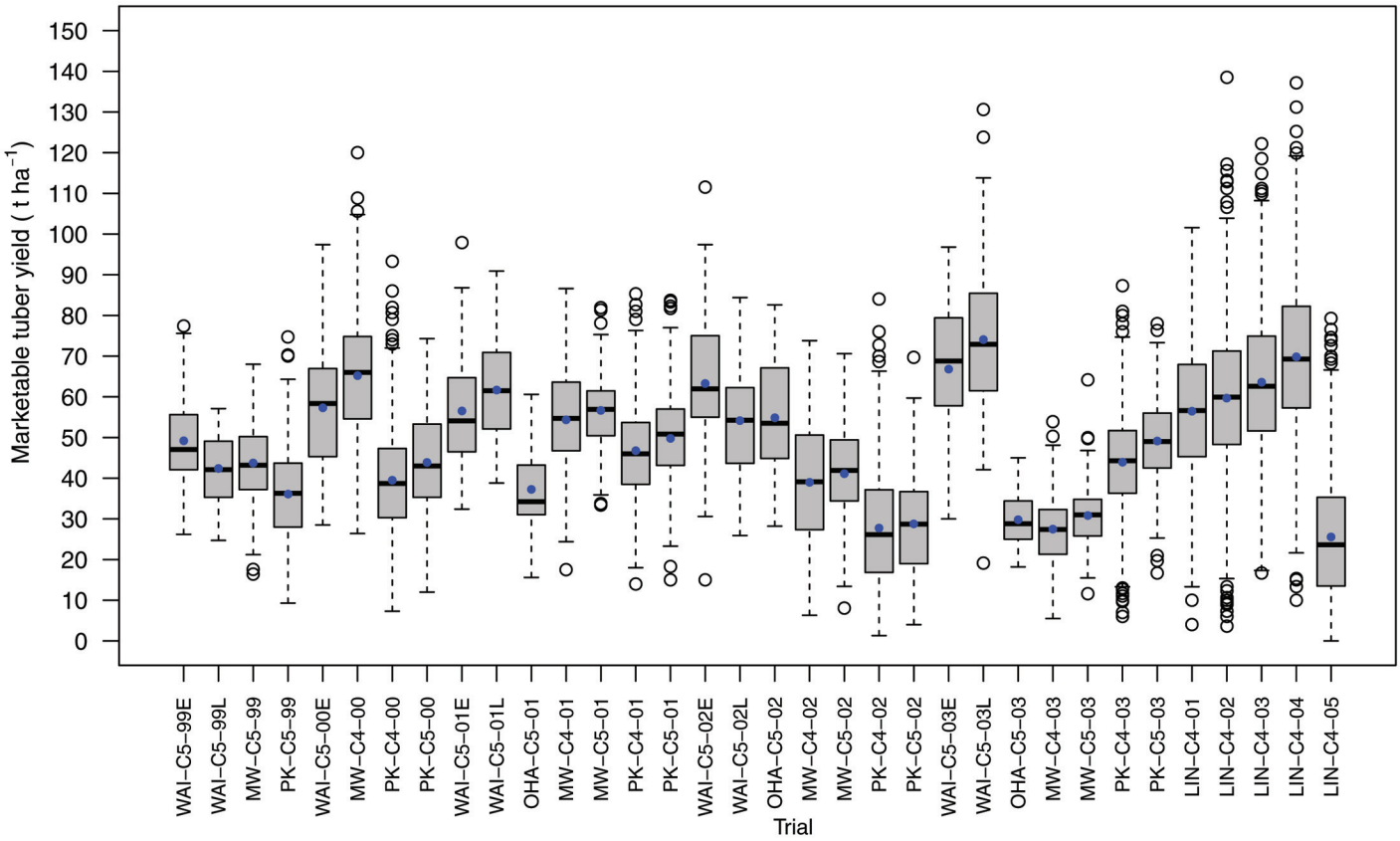

Figure 2. Boxplots of (marketable) potato tuber yield in tonnes per hectare from 34 trials. The prefixes WAI, MW, PK, OHA, and LIN refer to Waikato, Manawatu, Pukekohe, Ohakune, and Lincoln locations, respectively. Trials are grouped by island (North, South) and then by year and location. Mean yields are indicated by the filled circles.

(56–70 kg ha^{-1}) for the years 2001 to 2004. Phenotypic standard deviations for these four LIN trials ranged from 16 to 19 compared with a range of mostly 10 to 14 for other trials, with some exceptions falling outside this range (data not shown). The low mean yield (26 t ha^{-1}) and high coefficient of variation (59) for LIN-C4-05 were probably due to water logging that was reported for a period of the growing season for this trial (plot yield range of 0–79 t ha^{-1}). A larger variation in mean yield (across years) was observed for PK trials than LIN, ranging from 28 to 50 t ha^{-1}. For each location that trialled both C4 and C5 stages (PK, MW, and LIN), observed mean yields (and usually their standard deviations) were reasonably similar within each location-year (but not across years within locations). Although similar, there was a general pattern of observed mean yields increasing and variation decreasing from stages C4 to C5 for trials grown in the same location–year combination, reflecting the selection of varieties for tuber yield.

The total percentage variance accounted for by FA2 was a satisfactory 75%, an increase of 11% from the FA1 variance model. Starting values from the results of FA2 were used for the attempted fit of FA3, but convergence failed. The FA2 model was a reasonable fit for many of the trials and first latent variables were all positive (Table 3). Model fit, however, was particularly poor for trials PK-C4-01 and PK-C5-02. Model fit was also compromised, albeit to a lesser degree, for trials PK-C4-02, MW-C5-99, MW-C5-02, MW-C5-03, WAI-C5-99L, WAI-C5-01E, WAI-C5-02E, WAI-C5-03E, OHA-C5-03, and LIN-C4-05, as indicated by the reasonably poor percentage of variance that was accounted for by the FA2 model.

The relationship between pairs of environments is represented by the cosine of the angle between two environment vectors (Fig. 3). In general, environment vectors with a similar direction, as plotted from the origin approximate a high and positive correlation. As dissimilarity in the direction of two environment vectors increases, then the correlation between the two environments decreases, and those with opposite directions approximate a high and negative correlation of G × E effects (Fox et al., 1997). The span of vectors in Fig. 3, between PK-C5-01 (15) and OHA-C5-02 (18) and therefore encompassing all trials (not drawn), was subtended by an angle of approximately 90°, therefore indicating a pairwise correlation between these two trials (and a minimum correlation obtained from all pairwise combinations) of about zero.

As a general rule, the uniplot (Fig. 3) illustrated that trials were more likely to be clustered by location rather than by year. Most trials at the two main PFR research sites (PK and LIN) were delineated from one another,

Table 3. Trial numbers and identities with estimates of rotated environment loadings (first and second [scaled] latent variables, ℓ_1 and ℓ_2, respectively) and the percentage of variance accounted for (%*V*) by the first latent variable (ℓ_1) and the first and second latent variables ($\ell_1 + \ell_2$).

Trial no.†	Trial‡	ℓ_1	ℓ_2	%*V* (ℓ_1)	%*V* ($\ell_1 + \ell_2$)
1	WAI-C5-99E	0.67	−0.61	44	81
2	WAI-C5-99L	0.66	−0.03	44	44
3	MW-C5-99	0.70	−0.03	49	49
4	PK-C5-99	0.82	0.31	67	76
5	WAI-C5-00E	0.96	−0.29	92	100
6	MW-C4-00	0.84	0.09	70	71
7	PK-C4-00	0.68	0.43	46	65
8	PK-C5-00	0.75	0.43	56	75
9	WAI-C5-01E	0.79	0.01	62	62
10	WAI-C5-01L	0.93	0.28	87	95
11	OHA-C5-01	0.87	−0.50	75	100
12	MW-C4-01	0.95	−0.08	90	90
13	MW-C5-01	0.86	0.10	74	74
14	PK-C4-01	0.49	0.34	24	36
15	PK-C5-01	0.72	0.50	52	77
16	WAI-C5-02E	0.59	−0.50	35	61
17	WAI-C5-02L	0.84	0.51	71	97
18	OHA-C5-02	0.57	−0.82	33	100
19	MW-C4-02	0.83	0.26	68	75
20	MW-C5-02	0.72	0.32	52	62
21	PK-C4-02	0.65	0.38	43	57
22	PK-C5-02	0.24	−0.08	6	7
23	WAI-C5-03E	0.74	0.30	55	63
24	WAI-C5-03L	0.93	0.11	86	87
25	OHA-C5-03	0.76	0.10	58	59
26	MW-C4-03	0.88	−0.25	77	84
27	MW-C5-03	0.65	−0.19	42	46
28	PK-C4-03	0.89	0.02	78	79
29	PK-C5-03	0.75	−0.29	56	65
30	LIN-C4-01	0.86	−0.31	75	84
31	LIN-C4-02	0.80	−0.34	64	75
32	LIN-C4-03	0.84	−0.25	70	76
33	LIN-C4-04	0.82	−0.29	67	76
34	LIN-C4-05	0.71	−0.20	51	55

† Trial numbers correspond to those presented in Fig. 3a.

‡ E, early trial; L, late trial.

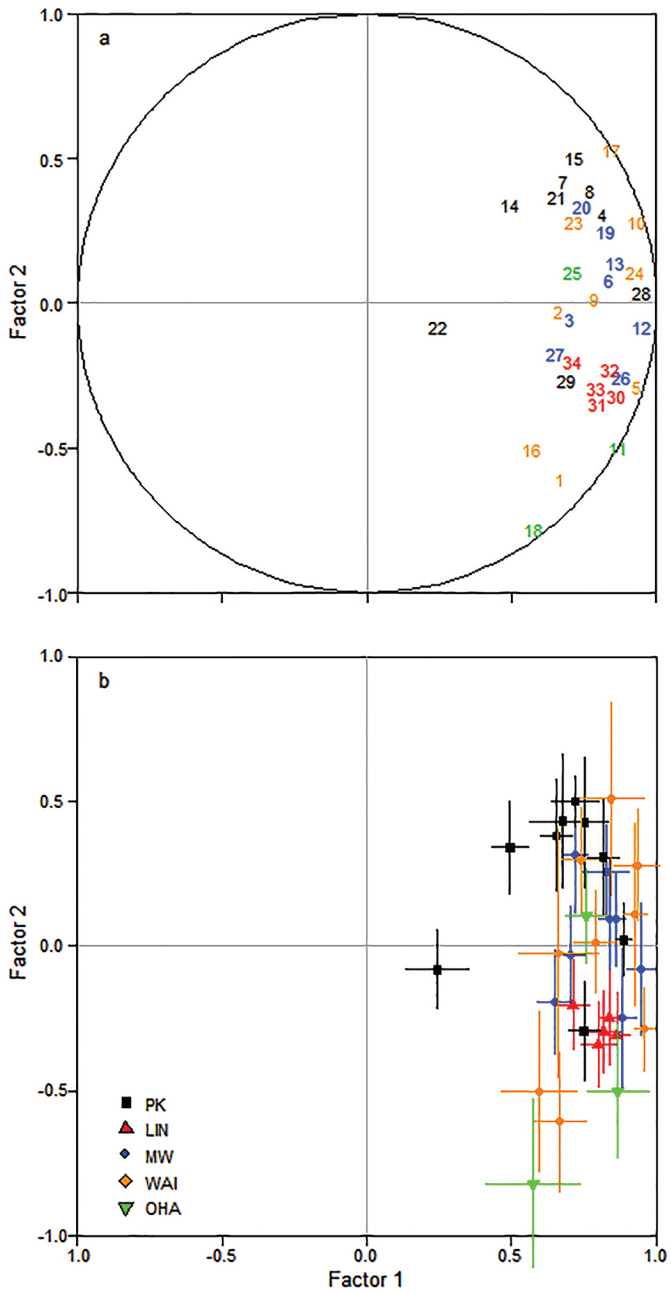

Figure 3. Environment uniplot of the genetic effect for potato tuber yield. Factor 1 and Factor 2 represent the rotated environment loadings on a correlation scale: (a) trials (given as trial numbers and are found with their corresponding trial identities in Table 3) and; (b) trials with trial locations represented as symbols and standards errors for the environmental loadings. The prefixes WAI, MW, PK, OHA, and LIN refer to Waikato, Manawatu, Pukekohe, Ohakune, and Lincoln locations, respectively.

and each tended to group together. The LIN trials in particular were grouped together closely and had negative second latent variables. The PK trials were mostly grouped together (with the exception of three trials) and, in contrast to those of LIN trials, second latent variables were mostly positive. For trials PK-C5-02 and PK-C4-01, yields were poor, and the model accounted for only 7 and 36% of their total variation, respectively (Table 3). Large specific variances were also found for PK-C5-02 and (to a lesser extent) for PK-C4-01. These results indicated that the interpretation of pairwise genetic correlations inferred from the uniplot involving these two trials may be unreliable (Cullis et al., 2010). The MW trials were not so tightly clustered together but fell between the two main PK and LIN clusters, which reflected the geographic (latitudinal) location of these three on-station trials (Fig. 1). For off-station trials (WAI and OHA), the clustering of locations was not so easy to discern compared with that of on-station trials (PK. MW, and LIN) because of greater variability in second latent variables for WAI and OHA than LIN and PK. The three OHA trials, for instance, were widely dispersed with

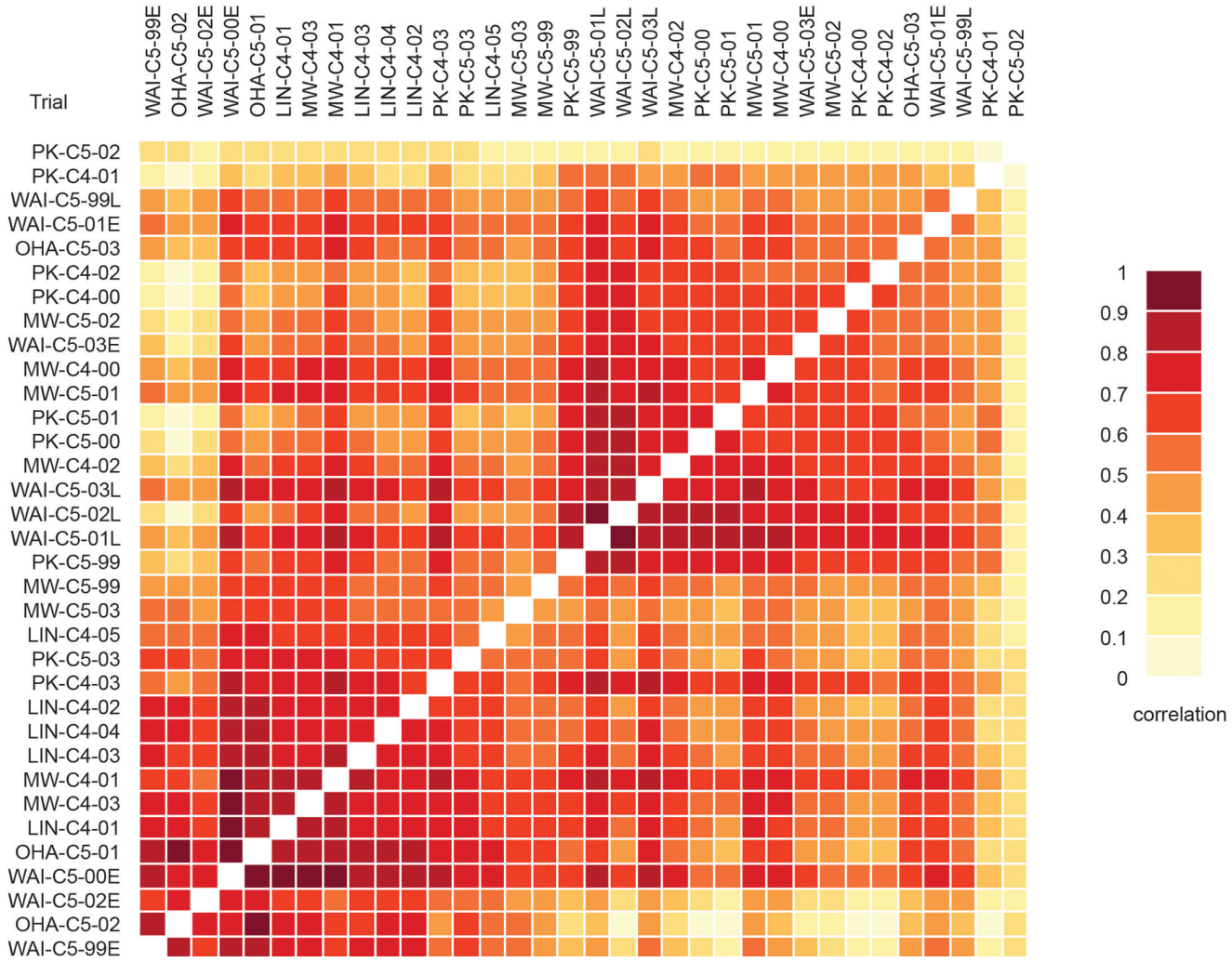

Figure 4. Heatmap of the genetic correlation estimates of potato tuber yield from 34 trials between 1999 and 2005. The prefixes WAI, MW, PK, OHA, and LIN refer to Waikato, Manawatu, Pukekohe, Ohakune, and Lincoln locations, respectively.

a large range of positive to negative latent variables. There may also have been a contrast between the WAI early trial (positive) and WAI late trial (negative).

The genetic correlation matrix is illustrated by the heatmap shown in Fig. 4. Genetic correlations ranged from 0 to 0.97, with two groups of trials displaying particularly strong correlations. This pattern is also reflected, to a large degree, by the uniplot in Fig. 3. The reduced correlations between PK and LIN in most years are visually represented, as are the low correlations between both LIN and PK and some OHA and WAI trials. The heatmap also illustrates the low correlations between PK-C5-02 (and PK-C4-01) and all (or most) other trials, which is a further indication that the uniplot may not be a reliable means to infer pairwise genetic correlations involving these particular trials.

The predicted genotype values and stability measures for a number of advanced clones and cultivars are shown in Fig. 5. High-yielding, advanced clones that have been developed by the PFR breeding program can be identified (Fig. 5, 6), for example, Moonlight, Allura, Summer Delight, but these selections did not necessarily demonstrate greater stability than recently imported cultivars. There is evidence to suggest there has been genetic improvement for potato yield in New Zealand previous to 1970 based on the comparison of more recently developed varieties with old cultivars that are still widely grown (Fig. 6). In 2011, the cultivars included in this analysis made up approximately 75 to 80% of the seed tuber production area (with ~33 cultivars accounting for the remainder), which indicates their importance as commercially grown cultivars in New Zealand. The pre-1970 cultivars (shown in Fig. 6) contributed to ~20% of the total seed tuber growing area, while three cultivars, namely Agria, Nadine, and Moonlight, accounted for ~40%. The international (imported) cultivars Russet Burbank, Desiree, Draga, Nadine, Ranger Russet, Laura, and Agria, together, contributed a large proportion of the total potato production area (~40%), but the predicted mean yield for most of these cultivars was below the observed mean commercial yield of 46 t ha^{-1} recorded from 2001 to 2010 (FAOSTAT, 2013). Varieties have been developed in New Zealand that are both high yielding and acceptable in terms of cooking and

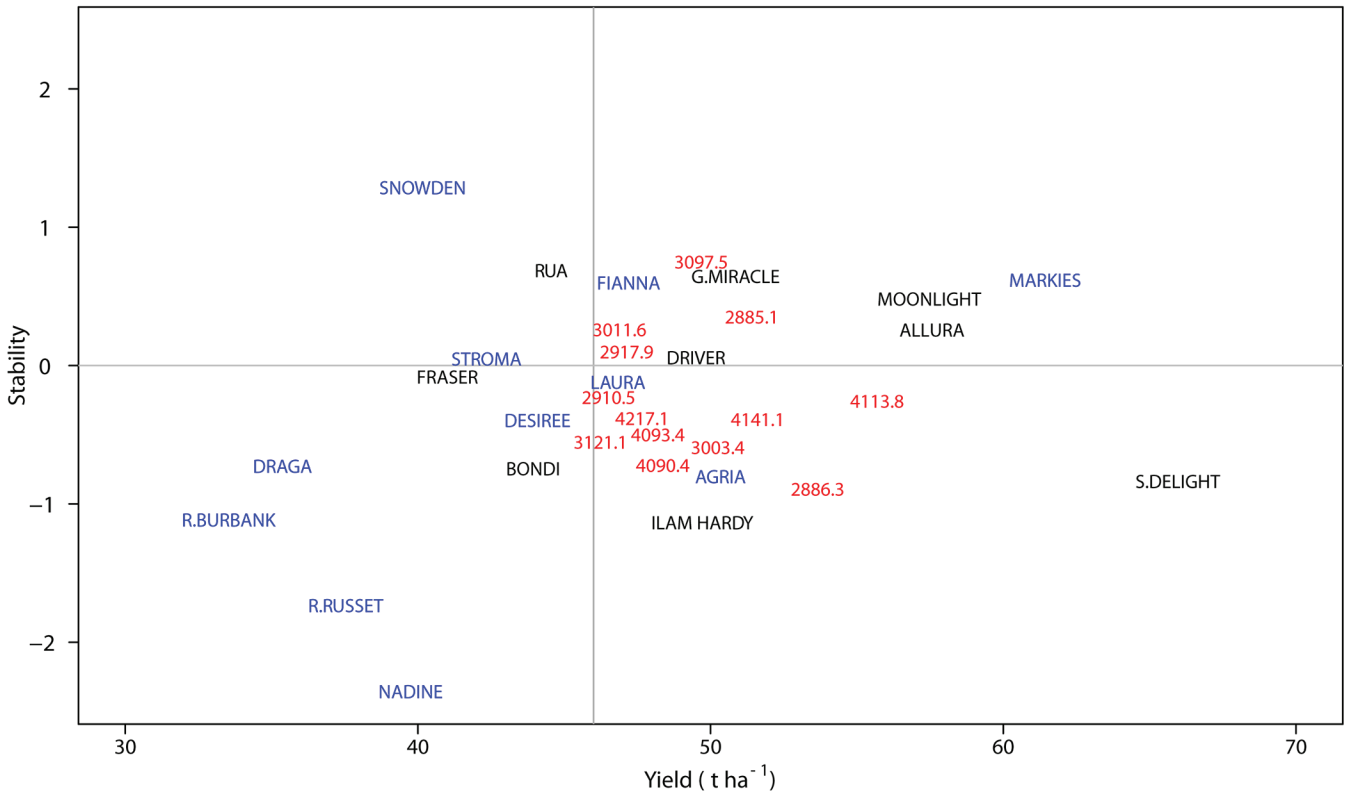

Figure 5. Predicted yields and stability measures (t ha^{-1}) for imported cultivars (blue), New Zealand-bred cultivars (black), and advanced clones (red). Standard errors of stability estimates are not shown for sake of clarity, but ranged from ±0.1 to 0.5. Confidence limits for yield are shown in Fig. 6.

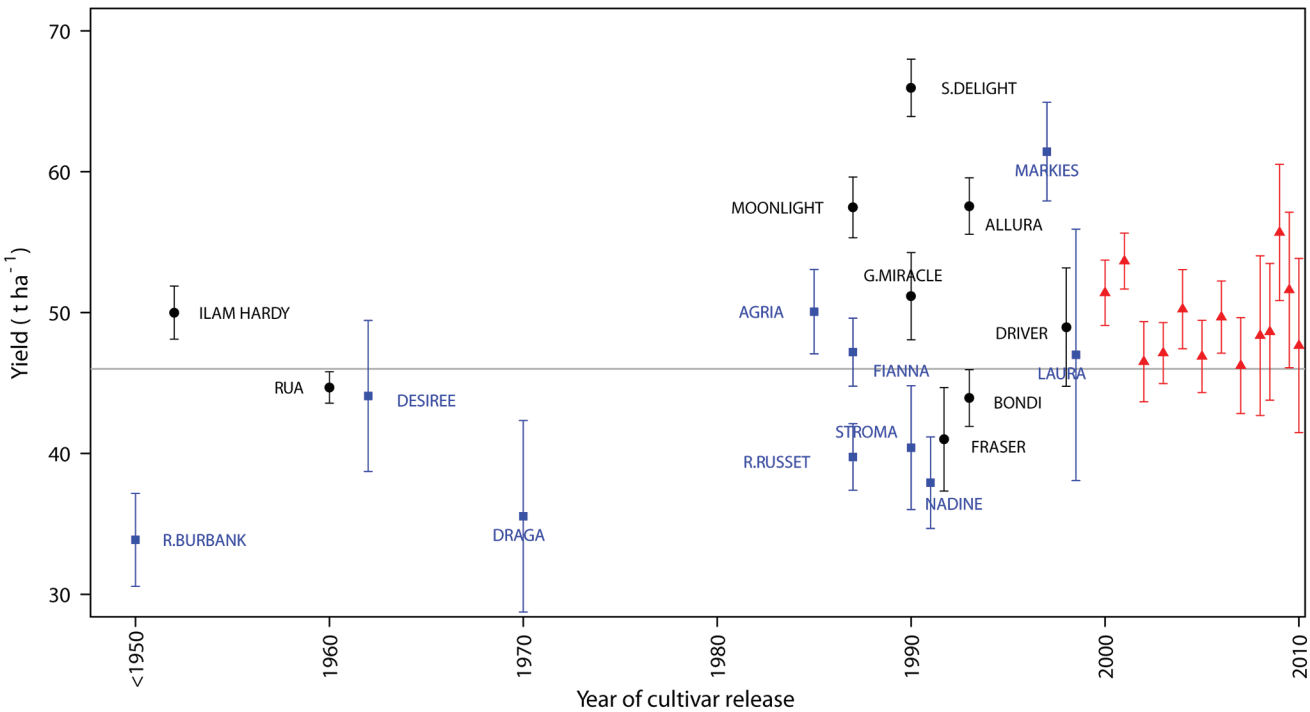

Figure 6. Predicted yields with 95% confidence intervals for imported potato cultivars (blue squares), New Zealand-bred cultivars (black circles), and advanced clones (red triangles). Year of cultivar release is approximate (for advanced clones, this is the year presented for commercial tender) and the horizontal line represents the New Zealand mean seasonal production yield of 46 t ha^{-1} between 2001 and 2010 (FAOSTAT, 2013).

processing quality. Further, many of the advanced clones developed by PFR since 2000 are above this yield threshold, suggesting that, in relation to the production cultivars presented here, the genetic selection for increased tuber yield in a multitrait selection program in New Zealand has been largely successful. However, the genetic improvement of yield of some of the advanced clones tendered for commercial release between 2000 and 2010 has been generally modest and often at or just above the observed mean yield reported over this period. The locally bred cultivars Summer Delight and Allura and the imported cultivar Markies are not widely grown commercially but have been shown to be particularly high yielding under New Zealand test conditions. These results are based on New Zealand test conditions managed for selection purposes and assume that there is no important G × M (genotype × management) interaction.

DISCUSSION

Multilocation Testing in Potato Evaluation Trials

Yield testing in breeding programs is highly resource demanding in terms of land and labor requirements. The retrospective analysis of historical MET data is of interest to plant breeders, as it determines the magnitude and type of G × E effects for traits, which helps to reevaluate breeding strategies. The discrimination of trials, in terms of G × E effects, is also useful, as it can provide plant breeders with information on locations with regard to differentiating genotype performance: "Efficiency in selection necessitates rationalisation of selection locations according to similarity of selection locations in discriminating among the genotypes." (DeLacy et al., 1996b, p. 244). Potato selection in the New Zealand potato breeding program is based on genotypes that are broadly adapted to perform well over all major growing regions in New Zealand. In the present study, the main PFR trial sites, PK and LIN, were identified as two contrasting test locations for the evaluation of tuber yield, and that testing at these two sites (and as early as possible in the selection cycle) is likely to provide the best opportunity to identify broadly adapted clones. Considering the three PFR research farm locations (PK, MW, and LIN), there was a pattern of stratification, in general, across these on-station trials (Fig. 3) perhaps based on latitude (Fig. 1). The MW second-factor environment loadings were more variable than those of LIN, effectively distributed around zero and clustered between PK and LIN (Fig. 3). Genetic correlations between MW trials and other trials were positive and generally moderately high to high (Fig. 4). This may suggest that MW field trials are contributing little extra in terms of discriminating genotypes for tuber yield performance and broad adaptation above and beyond that which is being achieved from comprehensive testing in PK and LIN; resources may be better diverted elsewhere, for example, by improving selection precision from further replication or by increasing selection intensity by evaluating more clones at existing sites. Increasing test locations implies an increased cost of running a breeding program if there is no net benefit in terms of genetic gain or the probability of identifying the best clones.

Commercial potato production in New Zealand is in a temperate maritime climate with some (but not extreme) differences in rainfall, temperature, and humidity. Growing regions such as LIN in the Canterbury region of the South Island, which are prone to long periods of water deficit in the summer months, are routinely irrigated. Production at PK and WAI, on the other hand, has traditionally relied on rainfed production, but in recent years irrigation has been available during extended dry periods. Yields from the regional trials (1999–2005), however, may have been more erratic because of the greater season-to-season variations in precipitation and exposure to cycles of supply and deficit. The potato crop is reported to be very sensitive to soil water condition compared with many other crops, and fluctuating water availability and the timing of availability over the growing season can severely affect total and marketable yield (e.g., Vayda, 1994 and references therein; Walworth and Carling, 2002). Off-station trials (WAI and OHA) behaved more erratically, in general, did not appear to group together in any predictable pattern, and genetic correlations between these trials and others were sometimes low and effectively zero (Fig. 3). To some extent, this may have been a reflection of on-station trials (PUK, MW, and LIN) being more similarly managed within location than the off-station trials of OHA and WAI and, therefore, less affected by management factors that were outside of the breeders' control, for example, planting densities and nutrient management. Interestingly, the mean ratio of genetic to environmental variance was highest for the off-station location of OHA (3.8), with WAI at 1.8 and PUK, MW, and LIN at 2.0, 1.5, and 2.7, respectively.

It would be informative to the breeder if there were a better understanding of the reasons for the poor genetic correlations between environments. Although the models tested were suitable for the aims of the current study, a disadvantage of the methods presented is that they are statistical rather than descriptive. Descriptive or analytical approaches aim to characterize the response of genotypes and environments in terms of abiotic and biotic factors (Fox et al., 1997), and they may also include physiological or genetical (e.g., quantitative trait loci) information. Descriptive statistical models to analyze G × E data have been reviewed by van Eeuwijk et al. (2005) and Crossa et al. (2010). Such an approach was used in an empirical study by Zhang et al. (2013) to characterize various environments in Australia and the response of seed yield and oil content in canola genotypes. Based on their results, phenology was found to have an influence on the performance

of genotypes in contrasting climates and this information was used to develop a breeding strategy targeted at specific adaptation. Similar approaches could be followed for the evaluation of potato MET data in programs that target specific adaptation in cultivars for a diverse range of climates (e.g., the CIP breeding programs based in Peru).

A recent genomic prediction study in barley by Heslot et al. (2013) used the prediction accuracy, rather than the genetic correlation, between environments to characterize test environments and to improve prediction of genomic breeding values in variety evaluation for the target population of environments. Studies to determine the extent and type of G × E effects found in historic tuber yield trials will help to determine the testing strategy for genomic studies and variety selection for potato breeding programs in future.

Genetic Improvement of Potato Yield and Stability

Douches et al. (1996) found that the genetic yield potential of modern cultivars in the USA had not improved over those of vintage cultivars, which was attributed to a greater focus on tuber (processing) qualities and selection for earlier maturity. Other studies have indicated that the contribution of genetics to improve yields has been small relative to those obtained from developments in agronomic practice (Sneep and Hendriksen, 1979; Walker et al., 2003), which seems to be contrary to reports for other staple crops (e.g., Duvick, 2005; Mackay et al., 2011). The present study has shown that the PFR potato improvement program has made some progress in developing advanced clones (those tendered for commercial release) and commercially released cultivars with marketable tuber yields above those found for established cultivars (Fig. 6) that together make up a large proportion of the current commercial crop in New Zealand (including cultivars Desiree, Nadine, Agria, Russet Burbank, and Ranger Russet). In plant variety terms, some of these cultivars are old. For example, Russet Burbank, a popular french-fry cultivar, dates back to 1908 and Desiree, a popular table cultivar, dates back to 1961 (van Berloo et al., 2007). This, to some degree, illustrates the slow adoption of newer and more productive cultivars, which is a conservatism that is reported to be characteristic of potato production systems, in general, (Tarn et al., 1992; Veilleux and De Jong, 2007; Walker et al., 2003) and is not conducive to advancing genetic progress for tuber yield in potato production.

Point estimates for stability of a number of varieties (cultivars and advanced clones) are given in the present study (Fig. 5) and show that all varieties were relatively stable over the locations tested. Although estimates are likely to be often associated with large standard errors (up to ±0.5 in the present study), particularly for genotypes tested over limited environments, such information is useful for breeders to characterize genotypes. There is no evidence to suggest that New Zealand-bred cultivars or advanced lines are any more or less stable than international cultivars bred offshore. Breeders are more likely to select promising candidates that are more consistent in performance over trials, but this is not explicitly measured in the PFR program and stability is gauged by a genotype's variability in mean performance (relative to common standards) over many trials. From these data, it may be interesting to note that Russet Burbank, which is widely acknowledged to be sensitive to water stress (conditions that rapidly decrease its marketable yield because of the development of deformed tubers), was shown to be more unstable than Desiree (Fig. 5), a cultivar that is reported to be more tolerant of water stress (Vayda, 1994). Desiree showed near-average predicted yields and its stability possibly supports anecdotal reports that yield reliability may have, in part, contributed to its popularity.

With limited resources for testing genotypes, breeders largely have to ignore genotype × management interaction (G × M) effects. Elite potato lines are regularly developed that meet the high expectations that breeders demand but fail to make an impact in a commercial setting. This could, of course, be the result of numerous agronomic, economic, marketing, political, and social factors. However, as Messina et al. (2009) point out, the limitation with field trials is that breeders are searching a restricted set of the large space defined by all combinations of genotypes and target environments. When added to the complication of variable management practices, this expands G × E to an even more complex G × E × M space. Breeders have to contend with G × E and largely ignore the G × M, which is left to agronomists to deal with by identifying best management practices for a small selection of elite cultivars such as planting densities, water management, and nutrient requirements. If this step is neglected, then the potential of new cultivars are often not realized if traditional management practices that apply to older established cultivars are assumed to apply also to new cultivars. Crop yield is a complex trait; to enhance rates of yield improvement and to explore the G × E × M space more effectively, a step change in the understanding of physiological systems and processes and the development of plant simulation models has been proposed as a means to better link the genetic variation of physiological yield determinants with their underpinning genetic systems (e.g., Hammer et al., 2006; Messina et al., 2009).

CONCLUSIONS

The evaluation of MET data from a national potato breeding program identified trial locations that were most suitable for distinguishing the performance of varieties and to select those that were broadly adapted across target production sites. This study also allowed a direct comparison of the yield performance of newly developed varieties with established cultivars that are widely grown

in New Zealand to gauge genetic progress for tuber yield in the breeding program. Selection to improve the genetic potential of tuber yield has resulted in the development and release of new cultivars that are superior to established cultivars, some of which were developed before 1970. Estimates of performance stability enable breeders to further characterize the performance of new varieties over multiple environments using MET data. Analysis of historic potato breeding data using multivariate mixed models can therefore help to guide breeding strategies, monitor genetic progress, and improve resource allocation in cultivar development programs. Further research to relate climatic variables to genotype performance to help to interpret G×E interaction effects should go some way to improving genetic gain and targeting better deployment of specifically adapted potato cultivars. A better understanding of G×E effects will also contribute to the study and application of molecular selection methods for the genetic improvement of this important food staple.

Acknowledgments

We gratefully acknowledge the funding of this study by Potatoes NZ Charitable Trust. We also thank Ron Beatson of The New Zealand Institute of Plant and Food Research Ltd, Alasdair Noble of AgResearch NZ Ltd, Katia Stefanova of The University of Western Australia, and two anonymous reviewers for helpful advice that improved the manuscript.

References

Affleck, I., J.A. Sullivan, R. Tarn, and D.E. Falk. 2008. Genotype by environment interaction effect on yield and quality of potatoes. Can. J. Plant Sci. 88:1099–1107. doi:10.4141/CJPS07207

Annicchiarico, P. 2002. Genotype × environment interaction: Challenges and opportunities for plant breeding and cultivar recommendations. In: P. Annicchiarico, editor, FAO plant production and protection papers 174. Food and Agricultural Organization of the United Nations, Rome. p. 1–132.

Atlin, G.N., R.J. Baker, K.B. McRae, and X. Lu. 2000. Selection response in subdivided target regions. Crop Sci. 40:7–13. doi:10.2135/cropsci2000.4017

Bos, I., and P.D.S. Caligari. 2008. Selection methods in plant breeding. 2nd ed. Springer, Dordrecht.

Butler, D.G., B.R. Cullis, A.R. Gilmour, and B.J. Gogel. 2009. ASReml-R reference manual. Version 3. Queensland Dep. of Primary Industries and Fisheries, Brisbane, QLD.

Cotes, J.M., C.E. Nustez, R. Martinez, and N. Estrada. 2002. Analyzing genotype-by-environment interaction in potato using yield–stability index. Am. J. Potato Res. 79:211–218. doi:10.1007/BF02871937

Crossa, J., J. Burgueño, P.L. Cornelius, G. McLaren, R. Trethowan, and A. Krishnamachari. 2006. Modeling genotype × environment interaction using additive genetic covariances of relatives for predicting breeding values of wheat genotypes. Crop Sci. 46:1722–1733. doi:10.2135/cropsci2005.11-0427

Crossa, J., P.N. Fox, W.H. Pfeiffer, S. Rajaram, and H.G.J. Gauch. 1991. AMMI adjustment for statistical analysis of an international wheat yield trial. Theor. Appl. Genet. 81:27–37. doi:10.1007/BF00226108

Crossa, J., M. Vargas, and A.K. Joshi. 2010. Linear, bilinear, and linear-bilinear fixed and mixed models for analyzing genotype × environment interaction in plant breeding and agronomy. Can. J. Plant Sci. 90:561–574. doi:10.4141/CJPS10003

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

DeLacy, I.H., K.E. Basford, M. Cooper, J. Bull, and C.G. McLaren. 1996a. Analysis of multi-environment trials: An historical perspective. In: M. Cooper and G.L. Hammer, editors, Plant adaptation and crop improvement. CAB International, Wallingford. p. 39–124.

DeLacy, I.H., K.E. Basford, M. Cooper, and P.N. Fox. 1996b. Retrospective analysis of historical data sets from multi-environment trials: Theoretical development. In: M. Cooper and G.L. Hammer, editors, Plant adaptation and crop improvement. CAB International, Wallingford. p. 243–268.

Douches, D.S., D. Maas, K. Jastrzebski, and R.W. Chase. 1996. Assessment of potato breeding progress in the USA over the last century. Crop Sci. 36:1544–1552. doi:10.2135/cropsci1996.0011183X003600060024x

Duvick, D.N. 2005. The contribution of breeding to yield advances in maize (*Zea mays* L.). Adv. Agron. 86:83–145. doi:10.1016/S0065-2113(05)86002-X

FAOSTAT. 2013. Crop production statistics. http://faostat.fao.org/ (accessed 10 Oct. 2014).

Fox, P.N., J. Crossa, and I. Romagosa. 1997. Multi-environment testing and genotype × environment interaction. In: R.A. Kempton and P.N. Fox, editors, Statistical methods for plant variety evaluation. Chapman & Hall, London. p. 117–137.

Gauch, H.G., and R.W. Zobel. 1988. Predictive and postdictive success of statistical analyses of field trials. Theor. Appl. Genet. 76:1–10. doi:10.1007/BF00288824

Hammer, G., M. Cooper, F. Tardieu, S. Welch, B. Walsh, F. van Eeuwijk, et al. 2006. Models for navigating biological complexity in breeding improved crop plants. Trends Plant Sci. 11:587–593. doi:10.1016/j.tplants.2006.10.006

Haynes, K.G., D.M. Gergela, C.M. Hutchinson, G.C. Yencho, M.E. Clough, M.R. Henninger, et al. 2012. Early generation selection at multiple locations may identify potato parents that produce more widely adapted progeny. Euphytica 186:573–583. doi:10.1007/s10681-012-0685-1

Heslot, N., J.L. Jannink, and M.E. Sorrells. 2013. Using genomic prediction to characterize environments and optimize prediction accuracy in applied breeding data. Crop Sci. 53:921–933. doi:10.2135/cropsci2012.07.0420

Kelly, A.M., B.R. Cullis, A.R. Gilmour, J.A. Eccleston, and R. Thompson. 2009. Estimation in a multiplicative mixed model involving a genetic relationship matrix. Genet. Sel. Evol. 41. doi:10.1186/1297-9686-41-33

Lynch, M., and B. Walsh. 1998. Genetics and analysis of quantitative traits. Sinauer Associates, Sunderland.

Mackay, I., A. Horwell, J. Garner, J. White, J. McKee, and H. Philpott. 2011. Reanalyses of the historical series of UK variety trials to quantify the contributions of genetic and environmental factors to trends and variability in yield over time. Theor. Appl. Genet. 122:225–238. doi:10.1007/s00122-010-1438-y

McCann, L.C., P.C. Bethke, M.D. Casler, and P.W. Simon. 2012. Allocation of experimental resources used in potato breeding to minimize the variance of genotype mean chip color and

tuber composition. Crop Sci. 52:1475–1481. doi:10.2135/cropsci2011.07.0392

Messina, C., G. Hammer, Z. Dong, D. Podlich, and M. Cooper. 2009. Modelling crop improvement in a G × E × M framework via gene–trait–phenotype relationships. In: V. Sadras and D. Calderini, editors, Crop physiology: Applications for genetic improvement and agronomy. Academic Press, London. p. 235–265.

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

Patterson, H.D., and R. Thompson. 1971. Recovery of block information when block sizes are unequal. Biometrika 31:100–109.

Piepho, H.P. 1997. Analyzing genotype–environment data by mixed models with multiplicative terms. Biometrics 53:761–767. doi:10.2307/2533976

Piepho, H.P. 1998. Empirical best linear unbiased prediction in cultivar trials using factor-analytic variance–covariance structures. Theor. Appl. Genet. 97:195–201. doi:10.1007/s001220050885

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

Po, E.A., S.S. Snapp, and A. Kravchenko. 2010. Potato yield variability across the landscape. Agron. J. 102:885–894. doi:10.2134/agronj2009.0424

R Development Core Team. 2012. R: A language and environment for statistical computing. R Foundation for Stat. Comput.,Vienna, Austria.

Redulla, C.A., J.R. Davenport, R.G. Evans, M.J. Hattendorf, A.K. Alva, and R.A. Boydston. 2002. Relating potato yield and quality to field scale variability in soil characteristics. Am. J. Potato Res. 79:317–323. doi:10.1007/BF02870168

Smith, A.B., B.R. 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., B.R. Cullis, and R. Thompson. 2005. The analysis of crop cultivar breeding and evaluation trials: An overview of current mixed model approaches. J. Agric. Sci. 143:449–462. doi:10.1017/S0021859605005587

Sneep, J., and A.J.T. Hendriksen. 1979. Plant breeding perspectives: Centennial publication of Koninklijk Kweekbedrijf en Zaadhandel D.J. van der Have, 1879–1979. Centre for Agricultural Publishing and Documentation, Wageningen.

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

Tai, G.C.C., and W.K. Coleman. 1999. Genotype × environment interaction of potato chip colour. Can. J. Plant Sci. 79:433–438. doi:10.4141/P98-109

Tarn, T.R., G.C.C. Tai, H. De Jong, A.M. Murphy, and J.E.A. Seabrook. 1992. Breeding potatoes for long-day, temperate climates. In: J. Janick, editor, Plant Breeding Reviews. John Wiley & Sons, New York. p. 217–332.

van Berloo, R., R.C.B. Hutten, H.J. van Eck, and R.G.F. Visser. 2007. An online potato pedigree database resource. Potato Res. 50:45–57. doi:10.1007/s11540-007-9028-3

van Eeuwijk, F.A., M. Malosetti, X.Y. Yin, P.C. Struik, and P. Stam. 2005. Statistical models for genotype by environment data: From conventional ANOVA models to eco-physiological QTL models. Aust. J. Agric. Res. 56:883–894. doi:10.1071/AR05153

Vayda, M.E. 1994. Environmental stress and its impact on potato yield. In: J.E. Bradshaw and G.R. Mackay, editors, Potato genetics. CAB International, Wallingford. p. 239–261.

Veilleux, R.E., and H. De Jong. 2007. Potato. In: R.J. Singh, editor, Genetic resources, chromosome engineering, and crop improvement. CRC Press, London. p. 17–58.

Walker, T.S., Y.P. Bi, J.H. Li, P.C. Gaur, and E. Grande. 2003. Potato genetic improvement in developing countries and CIP's role in varietal change. In: R.E. Evenson and D. Gollin, editors, Crop variety improvement and its effects on productivity: The impact of international agricultural research. CABI Publishing, Wallingford. p. 315–336.

Walworth, J.L., and D.E. Carling. 2002. Tuber initiation and development in irrigated and non-irrigated potatoes. Am. J. Potato Res. 79:387–395. doi:10.1007/BF02871683

Yang, R.C., J. Crossa, P.L. Cornelius, and J. Burgueño. 2009. Biplot analysis of genotype × environment interaction: Proceed with caution. Crop Sci. 49:1564–1576. doi:10.2135/cropsci2008.11.0665

Zhang, H.P., J.D. Berger, and S.P. Milroy. 2013. Genotype × environment interaction studies highlight the role of phenology in specific adaptation of canola (*Brassica napus*) to contrasting Mediterranean climates. Field Crops Res. 144:77–88. doi:10.1016/j.fcr.2013.01.006