

INTEGRATED GENETIC ANALYSIS FOR POTATO IMPROVEMENT

R.J. KERR¹, G.W. DUTKOWSKI¹, L. LI², T.A. MCRAE³, R. NOVY⁴, B.
SCHNEIDER⁴, and B. TIER³

SUMMARY

The POTATOPLAN project aims to adapt comprehensive genetic evaluation systems used in forest tree and livestock breeding to potatoes. These systems use Best Linear Unbiased Prediction (BLUP) to incorporate all measurements from all relatives and correlated traits to best predict the additive and total genetic values of all genotypes for all traits. This enables optimal selection of genotypes for breeding and deployment as any genotype at any stage of testing can be compared with any other, including commercial varieties. The method can remove biases due to natural and artificial environmental variation while accounting for GxE, genetic trends over time due to selection, differences in data amount and quality, and varying sampling procedures. The models have been adapted to potatoes by estimating additive genetic effects using an additive relationship matrix that recognises the polyploid nature of most potato varieties. Family merit score counts (number of selected progeny) are converted into a binary selection trait for the genotypes in the family, with a value of 1 for selected named progeny and 0 for unselected progeny. This trait is used to predict the value of the selected progeny for all traits that contribute to the merit score through the genetic correlations determined on an experimental basis. Measurements of genotype performance using different plot or sample sizes is accounted for by using weighted analysis with the individual plant or standard sample size error variance as the reference point. These approaches should allow the more widespread adoption of BLUP methods in potato and other crop breeding. POTATOPLAN also includes a comprehensive database of pedigree, measurements, trial designs, genetic and other model parameters, and resultant genetic values to make them useable for breeders and growers.

Key words: *Solanum tuberosum*, genetic evaluation, POTATOPLAN.

INTRODUCTION

Genetic evaluation methods used in forest tree and livestock breeding emphasise the prediction of additive effects or breeding values. Breeding values are used to guide selection of new parents for mating and recurrent selection is fundamental for genetic improvement in forestry and livestock species. In forestry significant non-additive effects are added on to the breeding value in order to guide selection of material for deployment. The aim of this study is to assess the relevance of these analytical methods to potato breeding and cultivar development programs.

Why these methods have not yet made any impact in potatoes is attributable to the unique nature of potato breeding and the biology of the species. The cultivated potato is an autotetraploid and often significant proportions of the observed genetic variance in quantitative traits is assumed to be non-additive due di-, tri- and tetra-genic interactions. A parental selection and crossing plan places emphasis on the creation of a large, genetically diverse progeny generation. Favourable non-additive

¹ PlantPlan Genetics, C/- School Plant Science, UTAS, Hobart, Australia

² Animal Genetics and Breeding Unit, UNE, Armidale, NSW, Australia

³ Southern Tree Breeding Association, PO Box 1811, Mount Gambier, SA, Australia

⁴ Small Grains and Potato Germplasm Research Unit, Aberdeen, Idaho, USA

effects observed in the cloned progeny generation can be captured through clonal propagation. Parents are chosen for producing certain cross types, often based on their own phenotypic performance. A succession of clonal generations is used to cull down to a small list of potential new cultivars from among the tens of thousands of progeny. Earlier clonal generations are often visually assessed, while later clonal generations employ objective tests and genotypes may be ranked on clonal phenotypic means.

Why adapt methods that have been designed for use in diploid species, in which breeding and selection have much different emphases? Tarn *et al.* (1992) noted that general combining abilities (GCA) (breeding values divided by two) are relevant in potato breeding, because the ratio of GCA variance to specific combining ability (SCA) variance is medium to high for most economically important traits. They suggest that full-sib seedling progeny test information is necessary for prediction of GCA. The current study proposes that generation of seedling progeny for performance recording, though helpful, is not absolutely necessary for accurate prediction of GCAs. It is our hypothesis that sufficient information exists in routine breeding and selection programs, and that this information has yet to be properly exploited.

The key to predicting a parent's GCA is having a substantial number of performance recorded progeny, via a substantial number of mates. It is unlikely that sufficient information exists within one non-seedling progeny generation. By the time important objective data comes to hand in later clonal generations, many progeny and numerous families have been discarded. However, if information can be collated into a single analysis across clonal generations within and across progeny generations, and perhaps even across programs, accurate prediction of GCA can be achieved in most cases. Central to genetic evaluation in forestry and livestock is the integration of program wide data and pedigree. The statistical technique used is called Best Linear Unbiased Prediction (BLUP), which uses relationship matrices to account for all known additive, and possibly non-additive, relationships among individuals. It is these matrices that allow information across trials (cohorts of progeny) and across time (generations) to be combined into a single integrated analysis. An integrated analysis will also increase the accuracy of predicting the additive and non-additive components of the genetic value for a genotype.

An important feature of BLUP methodology is its ability to exploit knowledge on the genetic architecture of the species, namely the genetic correlations among traits, as well as the ratios of additive to non-additive variance for each trait. Tarn *et al.* (1992) noted that many traits influence the profitability of a potato production system. Hence a multivariate approach to predicting genetic values is advantageous if estimates of genetic correlations among traits can be obtained.

In order to demonstrate and assess the appropriateness of BLUP methods and integrated analysis in potato breeding we have used data and pedigree from two progeny generations belonging to the Agricultural Research Service (ARS) Small Grains and Potato Germplasm Research Unit located in Aberdeen, Idaho. The ARS program services the fresh market, frozen French fry, crisp and dehydration industries with objectives that include: improving tuber qualities to increase long term storability and processing recovery; incorporating combined genetic resistance to major field and stored product diseases and pests; and improving production efficiency. The program

is fairly typical in that a large seedling generation is grown from true seeds (botanical) in the glass house each year and glass house tubers from each seedling are planted in the field the following year as single hill plots. In the subsequent field year, single hill selections are planted in unreplicated plots consisting of 12 plants (hills) of the selected genotype. Genotypes selected from the 12 hills plots then are advanced the following field year into yield trials consisting of 24 hill plots replicated twice in a randomized complete block design (RCB). Further years of replicated trials follow before final selections advance to Tri-State and Western Region Potato Variety trials. A modified TREEPLAN® system (Kerr et al. 2001) (named POTATOPLAN) and its associated data management system were used for the study.

MATERIALS AND METHODS

For integrated analysis a thorough and extensive preparation is needed. This preparation involves: collating possibly many hundreds of thousands of data points from multiple sources; defining an extensive pedigree that is able to link together all the data points; understanding and structuring the genetic architecture of the species. Each of these processes will be described in turn.

Trials and measurements

Measurements recorded from six trials from two series were used in the study: single hill trials in years 2004 and 2005; 12 hill trials in 2005 and 2006; 24 hill trials in 2006 and 2007. We use the word “series” to denote the succession of clonal trials stemming from each progeny generation. Each series is denoted as series A (starting from 2004) or series B (starting from 2005). Each trial is denoted using the series and the plot size, as in A1, A12 and A24.

Table 1 provides a description of each trait measured and which trials the traits were measured in. Measurements were not balanced across trials. Vine and some disease resistance traits were measured in the multiple hill trials. Traits such as percentage of tubers meeting US Number 1 grade specification and breeder’s merit score were recorded only in 24 hill trials. In single hill trials the only information recorded was the number of sibs selected per family to advance to the next stage (12 hill trials). This selection fraction was translated into a binary selection score for each genotype planted in single hill trials. The strategy is to regard this trait as a surrogate trait for the actual traits on which selection decisions were based.

Table 1. Measurements recorded in trials that were included in the POTATOPLAN analysis

Measured trait	Explanation	Abbreviation	Measured in trials
Number selected	The number of sibs per family selected	SELECTED	A1, B1
Vine maturity	Vine maturity on a 1 (very early; 100% dead vine) to 9 (very late) scale	VINE_MAT	A12, A24, B12, B24
Vine size	Size of the potato vine on a 1 (very small) to 9 (very large) scale	VINE_SIZE	A12, A24, B12, B24
Vert	Disease rating of Verticillium Wilt from 0 (none) to 9 (severe)	VERT	A12, A24, B12, B24
Early Blight	Disease rating from 0 (none) to 9 (severe)	EB	A12, A24, B12, B24
Yield	Pounds of seed harvested	YIELD	A12, A24, B12, B24
Specific gravity	Specific gravity shown in abbreviated form, e.g. 81 would signify a specific gravity of 1.081	SG	A24, B12, B24
Color 40 F	Fry color of samples stored at 40 degrees Fahrenheit for about 14 weeks. Fry scores are based on USDA color standards of 0 (very light) to 4 (very dark). Color less than 2 is acceptable by industry standards.	COL_40	A12, A24, B12, B24
Color 45 F	Fry color of samples stored at 45 degrees Fahrenheit for about 14 weeks	COL_45	A12, A24, B12, B24
%SE 40 F	Percent of sampled fries stored at 40 degrees F for about 14 weeks that have sugar ends. Sugar ends is a darkening in the fry color at the tips of French fries stemming from sugar accumulation in the outer perimeter of the tuber flesh.	SE_40	A12, A24, B12, B24
%SE 45 F	Percent of sampled fries stored at 45 degrees F for about 14 weeks that have sugar ends.	SE_40	A12, A24, B12, B24
LB resistance (foliar)	Material with potential late blight resistance evaluated in a higher disease pressure area. Scores indicate a foliar disease rating from 0 (none) to 9 (severe)	LB_FOLIAR	B12
LB resistance (tuber)	Late blight resistance scores indicating a percent infection in tubers	LB_TUBER	B12
Corky Ringspot Resistance percent infection	Material with potential corky ringspot resistance evaluated in a higher disease pressure area. Scores indicate percentage of 20 tuber sample showing symptoms.	CRR_PCENT	B12
Corky Ringspot Resistance severity	Scores indicate severity from 1 (none) to 5 (severe)	CRR_SEVERITY	B12
% US No. 1	Summed weight of tubers, which individually are greater than 4 oz and which do not have external defects, expressed as a percentage of total weight	US_NO1	A24, B24
% Hollow heart	The percentage of tubers over 12 oz that have hollow heart	HH_PCENT	A24
Merit Score	Breeder's merit score	MERIT_SCORE	A24, B24

Pedigree Definition

Definition of the pedigree was based on identifying families that were tested in single hill trials in years 2004 and 2005 and material that was added subsequently in multi-hill field trials over years 2005 to 2007, which was either check material or material sourced from other Tri-State programs. When POTATOPLAN is executed ancestors of all parents with tested progeny are retrieved by recursively querying pedigree records in the database and which extend as far back as 1992. Table 2 summarises the breakdown of parents by generation and source. In total 1506 parents were identified in the complete pedigree, of which 558 have progeny tested in trials. Nine parents, which are well known commercial cultivars, were also used as checks. A majority of parents (1211) have unknown ancestors, hence are allocated to the base generation (generation 1). This highlights the need to access pedigree records prior to 1992. Many of the base generation parents were not bred locally in Aberdeen, which emphasises the need to recover ancestral information from other programs or from public databases.

Table 2. Break down of parents by generation and source.

Generation	Total	“A” parents	Non “A” parents	Parents with progeny tested	Parents with clones tested
1	1211	594	617	427	8
2	256	254	2	115	1
3	39	39	0	16	0
Total	1506	887	619	558	9

Accuracy of genetic evaluation is influenced by the degree of genetic connectedness between trials in which individuals are measured in. Table 3 summarises the numbers of parents in each trial and the numbers of common parents used between trials. Of note is the large numbers of common parents used between trials of different series. It is unlikely lack of connectedness will be an issue for integrated analysis. Table 4 summarises the numbers of genotypes in each trial and the numbers of common genotypes used between trials.

Table 3. Numbers of parents by trial (**bold**) and numbers of parents in common between trials

	A1	A12	A24	B1	B12	B24
A1	223	207	65	136	124	43
A12	207	330	76	140	141	43
A24	65	76	78	51	49	26
B1	136	140	51	291	256	66
B12	124	141	49	256	371	69
B24	43	43	26	66	69	73

Table 4. Numbers of genotypes by trial (**bold**) and numbers of genotypes in common between trials

	A1	A12	A24	B1	B12	B24
A1	80251	1365	51	0	0	2
A12	1365	1969	70	0	6	9
A24	51	70	88	0	5	9
B1	0	0	0	72698	1590	65
B12	0	6	5	1590	1940	71
B24	2	9	9	65	71	78

Genetic architecture

Estimates of additive and SCA variances for each trait, and genetic and environmental correlations among traits were obtained using restricted maximum likelihood (REML) methodology implemented in the ASReml software package (Gilmour *et al.*, 2006). ASReml fits an additive genetic relationship matrix appropriate for a diploid species. An additive relationship matrix appropriate for a tetraploid species is the same, if inbreeding in the parents and double reduction is considered non-existent. Given that only a shallow pedigree was available at the time of completing the variance component estimation process and that little is known about the value for the double reduction parameter, it would make very little difference between assuming the correct relationship matrix and the matrix that ASReml computes by default. An identity relationship matrix was assumed for SCA effects.

The process of estimating the genetic architecture begins with separate, univariate analyses of individual traits, within a trial. A multivariate analysis involving all traits recorded in a trial is reached progressively, i.e. from univariate to bivariate to tri-variate etc. A consensus value for a genetic correlation between traits is obtained by averaging individual trial estimates. Multivariate, multi-trial analyses were used to estimate genetic correlations between traits that are recorded just in one trial. Consensus inter-trait additive genetic and SCA correlations, which were used in the integrated analysis, are shown in Tables 5 and 6, respectively.

Reports on estimates of genetic correlations between traits for potatoes are scarce in the literature. The few that exist support the current findings. Pereira *et al.* (1994) found a favourable, positive genetic correlation between more acceptable colour in chips and specific gravity. The correlation in Table 5 is negative, which is favourable because lower numbers for COL_40 and COL_45 denote a more acceptable colour. Specific gravity has been shown to be highly correlated with starch and glucose content (Jakuczun and Zimnoch-Guzowska, 1994). Sinha *et al.* (2008) found a negative correlation between glucose content and color as scored using the USDA system. The high additive genetic correlation (0.58) between VERT and EB agrees with the general observation that these diseases often occur concurrently (Pavek and Corsini, 1994). Both diseases result in early dying, i.e. lower values for VINE_MAT, and reduced yields. Goth and Haynes (2000) conclusively demonstrate the adverse effects of verticillium wilt on yield, tuber size distribution and specific gravity. The current study detected a significant, moderate, negative correlation between US_NO1 and VINE_SIZE. Deliberate vine killing is practised in industry to improve tuber size distribution. The correlations estimated in the current study

suggest that genotypes which are prone to early death, and have smaller vines will naturally produce better tuber size distributions.

The diagonals in Table 5 indicate unit additive genetic variances for all traits. This is achieved by dividing the trait measurements in each trial by the square root of the within trial estimate of the additive genetic variance for the trait. This strategy is used to remove scale effects associated with each trial. Use of trait specific error variances in POTATOPLAN implies that narrow sense heritabilities can change from trial to trial. This strategy results in more weight given to information in trials with lower environmental influences. The diagonals in Table 6 indicate SCA to additive genetic variance ratios, which for 15 of the 18 traits analysed, are less than unity. This result, indicating the prevalence of additive effects among quantitative traits important in the potato, confirms findings reached by Tarn *et al.* (1992). In particular they noted high GCA/SCA ratios for color of potato chips scored under various processing regimes. They cite Tai and Hodgson (1975) and Malcolmson and Killick (1980) who found resistance to late blight was mainly due to GCA effects of the parents. The results in this study suggest that SCA effects are important for corky ringspot resistance and success fraction of progeny in a cross at the initial selection stage.

Table 5. Consensus inter-trait additive genetic correlations obtained using ASReml and used in POTATOPLAN

	SELECTED	VINE_MAT	VINE_SIZE	VERT	EB	YIELD	SG	COL_40F	COL_45F	SE_40F	SE_45F	LB_FOLIAR	LB_TUBER	CRR_TUBER	CRR_SEVERITY	US_NO1	HH_PCEN	MERT_SCORE
SELECTED	1.00	0.00	0.00	0.00	0.00	-0.25	-0.20	-0.20	-0.20	0.14	0.20	0.00	-0.15	-0.30	-0.30	0.00	0.00	0.00
VINE_MAT	0.00	1.00	0.67	-0.83	-0.62	0.21	0.07	-0.11	-0.12	0.00	0.12	0.00	0.00	0.00	0.00	0.00	0.00	0.00
VINE_SIZE	0.00	0.67	1.00	-0.67	-0.41	0.37	0.18	-0.18	-0.19	0.00	0.07	0.00	0.42	0.00	0.00	-0.28	0.00	0.00
VERT	0.00	-0.83	-0.67	1.00	0.58	-0.15	-0.15	0.17	0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
EB	0.00	-0.62	-0.41	0.58	1.00	0.15	0.00	0.00	0.00	0.00	-0.49	0.00	0.00	0.00	0.00	0.00	0.00	0.00
YIELD	-0.25	0.21	0.37	-0.15	0.15	1.00	0.17	0.00	0.00	0.00	-0.21	0.00	0.00	0.00	0.00	0.00	-0.28	0.14
SG	-0.20	0.07	0.18	-0.15	0.00	0.17	1.00	-0.41	-0.45	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
COL_40F	-0.20	-0.11	-0.18	0.17	0.00	0.00	-0.41	1.00	0.88	-0.28	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
COL_45F	-0.20	-0.12	-0.19	0.14	0.00	0.00	-0.45	0.88	1.00	-0.14	-0.13	0.00	0.00	0.00	0.00	-0.14	0.00	0.00
SE_40F	0.14	0.00	0.00	0.00	0.00	0.00	0.00	-0.28	-0.14	1.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00
SE_45F	0.20	0.12	0.07	0.00	-0.49	-0.21	0.00	0.00	-0.13	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LB_FOLIAR	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00
LB_TUBER	-0.15	0.00	0.42	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00
CRR_TUBER	-0.30	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00
CRR_SEVERITY	-0.30	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00
US_NO1	0.00	0.00	-0.28	0.00	0.00	0.00	0.00	0.00	-0.14	0.02	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.23
HH_PCEN	0.00	0.00	0.00	0.00	0.00	-0.28	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00
MERT_SCORE	0.00	0.00	0.00	0.00	0.00	0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.23	0.00	1.00

Predicting breeding and SCA values

Breeding and total genetic values for genotypes and SCA values for families are predicted by POTATOPLAN using data and pedigree across 18 traits, six trials and three generations. POTATOPLAN uses customised software for solving very large systems of mixed model equations. A detailed description of analytical method was presented in a previous conference (Dutkowski *et al.*, 2006). More recent features include the correct additive relationship matrix for autopolyploid species, and adjustment of the error variance according to the degree of clonal replication in each trial. Thus in this study 12 hill plots have larger error variances on average than 24 hill plots. It is possible that 24 hill plots are heavily rogued resulting in fewer than 12 plants.

An integrated BLUP analysis results in a list of breeding values (EBV) for every genotype in the pedigree, which includes parents and ancestors which are not tested. This implies that tested genotypes can be compared across trials with any established cultivar, provided the latter is linked in through the pedigree. Currently the method can only predict SCA effects for those families with tested progeny and predict total genetic values (EGV) for tested genotypes. The EGV comprises the SCA effect of the family it belongs to and a Mendelian non-additive sampling component, the variance of which can be set to include di-, tri-, tetra-genic and even epistatic sampling. An estimate of the Mendelian non-additive sampling variance must be provided, either by direct input, or by setting it at some level relative to the SCA variance. For this study it was set at three times the SCA variance. In the near future relationship matrices describing the non-additive genetic structure in the pedigree will be included, thus providing the means to predict SCA and non-additive effects for non-tested material.

In order to highlight the advantages of an integrated BLUP analysis, smaller analyses, in conjunction with the complete analysis that included all 18 traits and all six trials were carried out. For example a single trait, single trial analysis was compared with a single trait, multiple trial analyses in order to highlight the advantages of using data across trials, without the confounding benefits of using information on correlated traits. EBV from the smaller analyses were also checked with EBV derived using ASREML and were found to be the same.

RESULTS

In non-experimental field trials such as the ones used in this study, it is normal for genotypes to be compared on the basis of clonal phenotypic means (CPM). Figure 1 shows scattergrams of EBV and EGV versus CPM. The top left plot shows EBV computed from a single trial, single trait analysis for YIELD in the B24 trial against the CPM, while the plot on the top right shows EGV. Genotypes are marked using their clone designation, or cultivar name if it exists. The bottom plots show a similar comparison but for CRR_SEVERITY measured in B12. Generally EBV and EGV align well with CPM. EGV should align more tightly because they are more indicative of the phenotype. EBV and EGV are normally distributed around zero. CRR_SEVERITY is a categorical type trait. However EBV and EGV for this trait have a continuous distribution, because information from relatives is combined with a genotype's observation to compute its

value. For each genotype the amount and quality of relatives' information varies, thus producing the continuous range shown

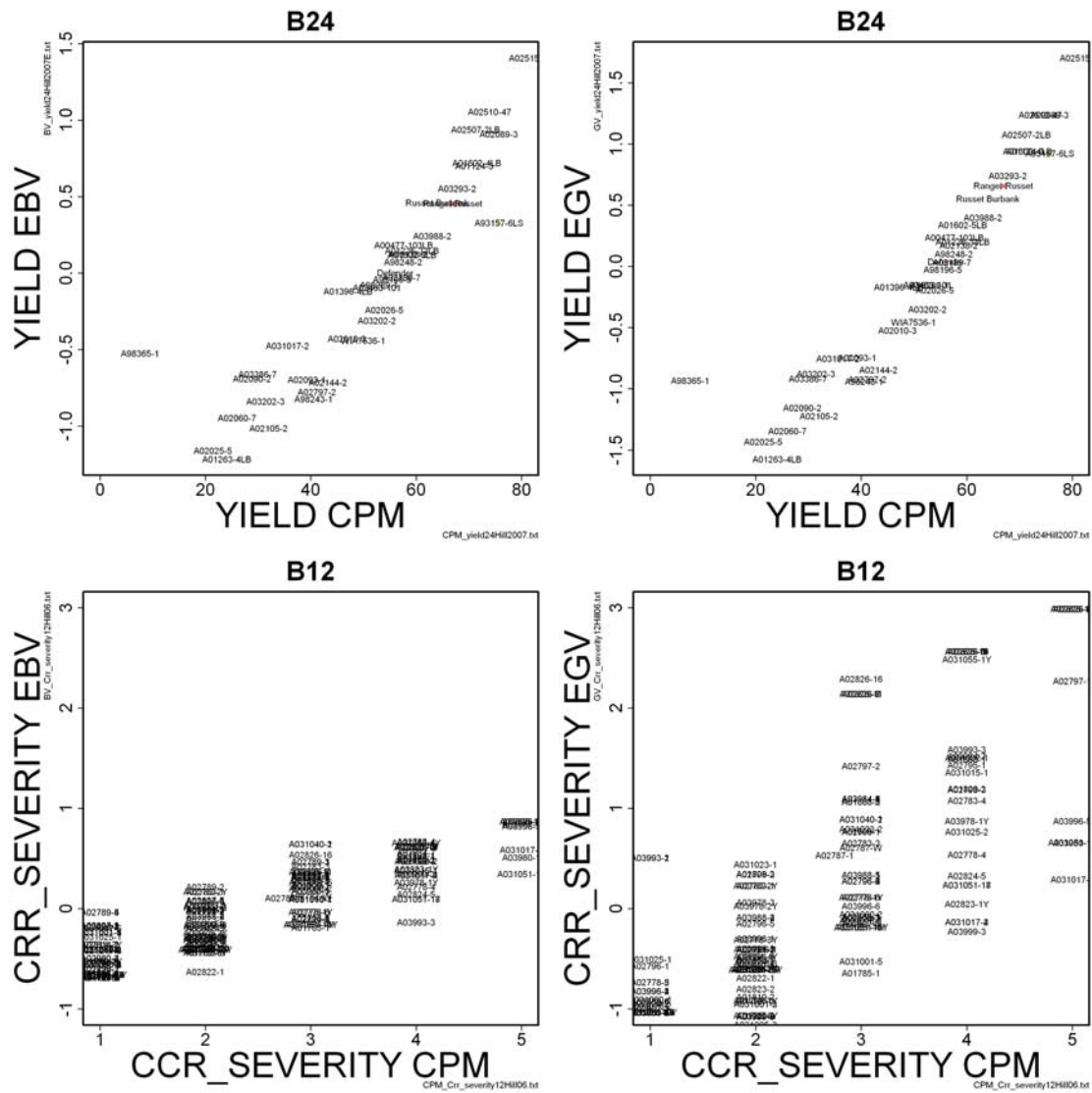


Figure 1. Comparison between predicted genetic values (EBV, EGV) and clonal phenotypic means (CPM) for YIELD and CRR_SEVERITY. Genotype IDs are shown.

Integrated analysis means combining data from all sources including different trials. Five genotypes with clonal information in three trials (A12, A24 and B24) were identified and used to illustrate the effects of combining trial data. The plot on the top left hand side of Figure 2 shows their EBV, which are computed using only information in the B24 trial, plotted against their CPM for yield per plant. Again, breeding values align

well with phenotypes. There is one clone (A01263-4LB) that has a very poor yield and a correspondingly low, negative breeding value. The plot on the right hand side shows EBV computed using information from the three trials. Interestingly A01263-4LB now has a positive breeding value. This is indicative of this genotype showing reasonable clonal performance for yield in two of the three trials and highlights an important point. Breeding (and total genetic) values, computed using information from multiple trials (different selections stages, different locations, different years), will not correlate with clonal performance measured in a single trial, as well as when the genetic values are computed in a single trial analysis. The bottom two plots demonstrate how rankings based on CPM for yield per plant can change from trial to trial. The bottom left compares B24 with A24 and bottom right compares B24 with A12. A major systematic effect is noticeable in A12 causing overall yield per plant to be much lower than the other two trials. Animal breeders recognised there were “herd” or “flock” effects when they first started to combine data across populations and would pre-adjust the data for systematic environmental effects. This statistical strategy is often referred to as BLP. The danger is that the pre-adjustment is not sufficiently accurate and genetic values will be biased because residual systematic effects confound the genetic effects. BLUP was invented to simultaneously estimate systematic environmental and genetic effects.

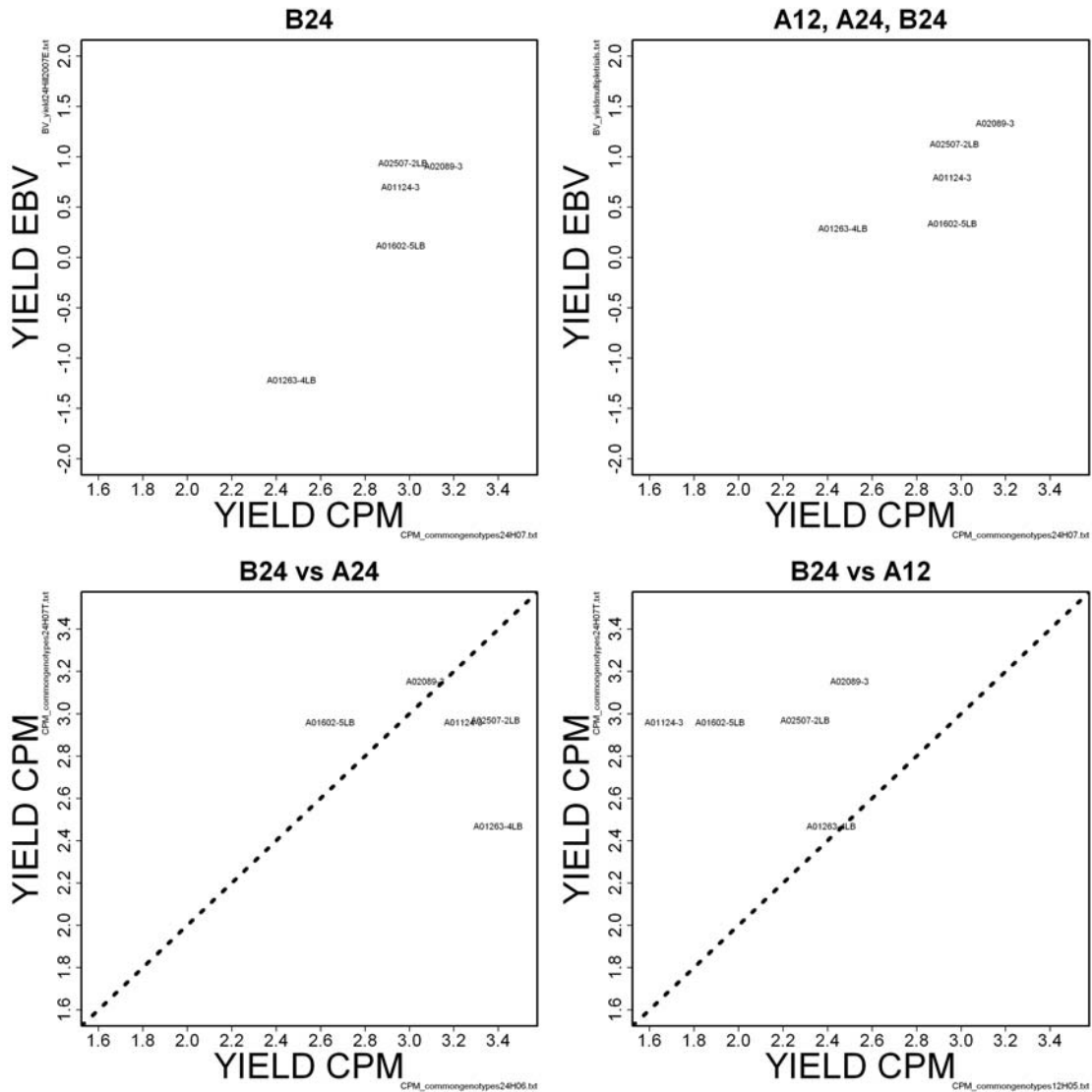


Figure 2. Comparison between EBV and CPM for yield using information on a single trial (top left) and multiple trials (top right) and comparison between CPM measured in trials B24 versus A24 (bottom left) and in trials B24 and A12.

Sources from which to combine data also include information on relatives. Prior to the adoption of BLUP, plant and animal breeders would often formulate standard selection indices to incorporate relatives' data, where the weights in the index often did not change from individual to individual. This assumes the type and amount of familial information was the same. BLUP has been described as allowing customised selection indices for every individual, because for every individual the type and amount of familial information will change. The use of the additive relationship matrix in BLUP allows this customisation to occur and enables breeding values to be computed for individuals

without observations. To illustrate, seven genotypes were isolated from the A12 trial with varying amounts of information (see Table 7).

Table 7. Parents of material tested in A12 with differing amounts of information

Clone	Number of mates in A12	Number of half-sib progeny in A12	Measured in trials	Number of mates across 4 trials	Number of half-sib progeny across 4 trials
Ranger Russet	18	132	A12, B12, A24, B24	42	163
Premier Russet	18	95	A24, B24	61	287
Russet Burbank	2	26	A12, B12, A24, B24	2	26
Russet Norkotah	3	4	A12, B12, A24	4	6
NorDonna	1	1	B12	2	3
Yukon Gold	1	1	A24	1	1
AC93026-9RU	1	5		8	39

In trial A12 there are 18 families where Ranger Russet is a parent and the total number of cloned progeny tested from these 18 families is 132. It is also directly measured in this trial. In contrast there are two genotypes (NorDonna and Yukon Gold) with only one cloned progeny each and no direct measurements in A12. The clone Premier Russet has plenty of progeny, but no direct measurements in A12. Parent AC93026-9RU has few tested progeny in A12, but across 4 trials has 39. Figure 3 displays the EBVs for these seven genotypes estimated using only information in A12 (x-axis) and when combining information from A12, B12, A24 and B24 (y-axis). The first point to make is that again genotypes re-rank when more information is included. It is noticeable that NorDonna and Yukon Gold are the top-ranked genotypes of the seven for vine size, when using only information from A12, but then drop in rank when more information is obtained, which included their direct measurements. This highlights the inaccuracy of the prediction when there is no observation recorded directly for the individual, and the individual has minimal information from relatives (in this case a single progeny). Of note is the placement of Russet Burbank: with more information it becomes the second top ranked variety of the seven. AC93026-9RU is an average ranked parent with only five tested progeny, but with 39 tested progeny from eight mates its ranking improves dramatically. The bar chart in Figure 3 shows the accuracy of EBV prediction. An accuracy of one denotes the estimated value is perfectly correlated to the true value. Accuracy is very low when a parent has only one tested progeny (NorDonna and Yukon Gold, light gray bar), but improves dramatically when the parent in question is clonally tested and this information becomes available (dark gray bar). There is little advantage in testing over 100 progeny in terms of accuracy (e.g. Premier Russet). It is interesting to compare Premier Russet with Ranger Russet. Both genotypes have numerous progeny tested in A12, but only Ranger Russet has itself been tested as a clone. However this makes little difference in terms of accuracy. Thus clonal tests are not necessary to predict breeding value when sufficient progeny information exists. On the basis of a single trial AC93026-9RU has too few progeny for accurate prediction of its breeding value. Across four trials it has more progeny (39 from eight mates), but still no direct measurement. The

accuracy has improved, but not to the same degree as other genotypes which also have direct clonal measurements.

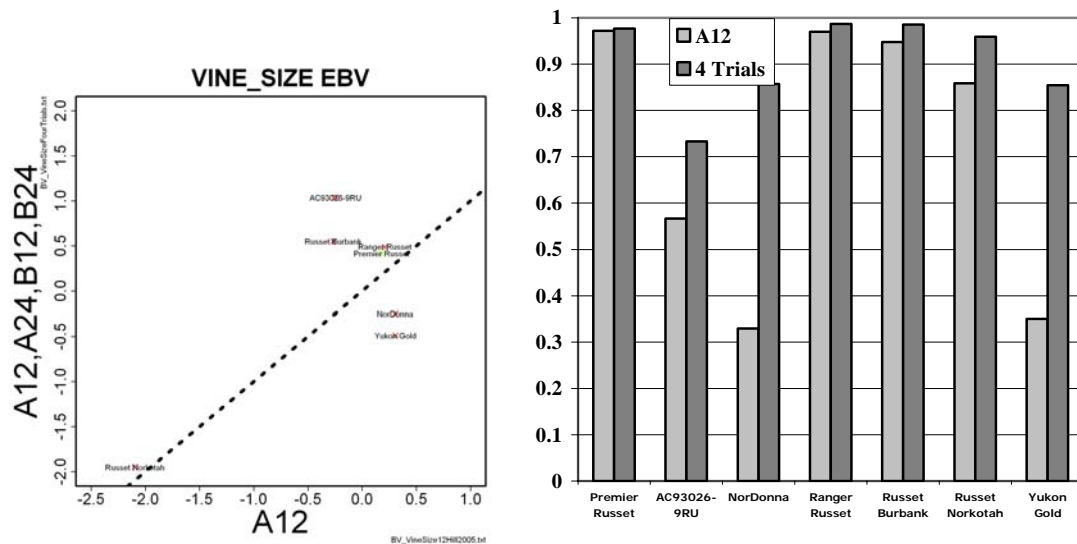


Figure 3. Comparison between VINE_SIZE EBV, when computed using information from a single trial (A12), and when computed using information from four trials (including A12). The bar chart on the right shows the accuracy of EBV prediction in both cases.

Sources from which to combine data also include information on correlated traits. Figure 4 displays two plots. The plot on the left hand side shows the scattergram of US_NO1 EBV, computed using only single trial (B24) and single trait (US_NO1) information, and the CPM measured in A12. The plot on the right hand side shows the scattergram of EBV computed from a fully integrated analysis (multi-trial, multi-trait) and CPM. The most noticeable difference between the two plots is the wider spread of EBV for US_NO1 computed in the integrated analysis. Because BLUP treats genetic effects as random they are regressed more towards their means when there less information, such as absence of information on correlated traits. US_NO1 was found to have in the current data a moderate, negative, additive genetic correlation (-0.28) with VINE_SIZE. Thus in a multi-trait analysis, data on vine size, which are recorded on the individuals and on their relatives, provide information to predict genetic values for US_NO1. Hence the greater spread of EBV is observed.

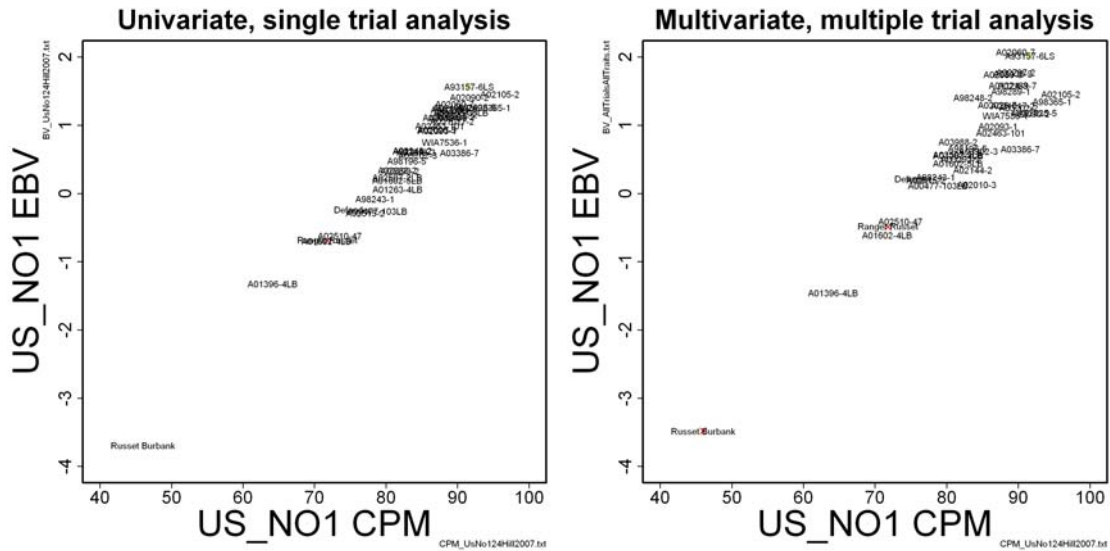


Figure 4. Comparison of clonal phenotypic means (CPM) for US_NO1 and EBV which are computed from a univariate, single trial analysis (left plot) and from a multivariate, multiple trial analysis (right plot)

The above scenario is certainly the case in potato cultivar selection programs. Selection practised at the single hill stage has a high intensity and is usually based on the breeders' perception of how well the progeny fit the overall cross purpose. This perception often aligns with uniformity of tuber shape and size (Bradshaw *et al.*, 1998; Tai, 1975). It is likely that many economically important traits which are recorded later will be affected by this selection step. The current study suggests selection at the single hill stage is correlated to yield, specific gravity, chip color and darkening, and resistance to corky ringspot resistance. Unfortunately, many cultivar selection programs do not record any information at the single hill stage, except noting how many progeny per family have been selected. In POTATOPLAN a strategy has been proposed to treat selection at the single hill stage as a binary trait. This binary trait might act as a surrogate for the actual component traits in selection and help in adjusting the means and variances of important correlated traits. Figure 5 shows scattergrams of EBV, from a bivariate analysis that included SELECTION in the 1 hill trial and the trait of interest recorded in the 12 and 24 hill trials, and from a univariate analysis of the trait of interest recorded in the 24 hill trial. Two traits are shown: SG on the left and YIELD on the right. There is a noticeable shift in the means in the bivariate analyses, in that many of the points are above the one to one correlation line (dashed line). Less noticeable is the narrower spread of EBV on the y axis. For SG the majority of points on the x-axis cover an interval of approximately 2.5 genetic standard deviation units. On the y-axis the interval is approximately 1.5 to 2.0 genetic standard deviation units. This is in accordance with theory. The strategy of using the binary selection trait appears to be working correctly.

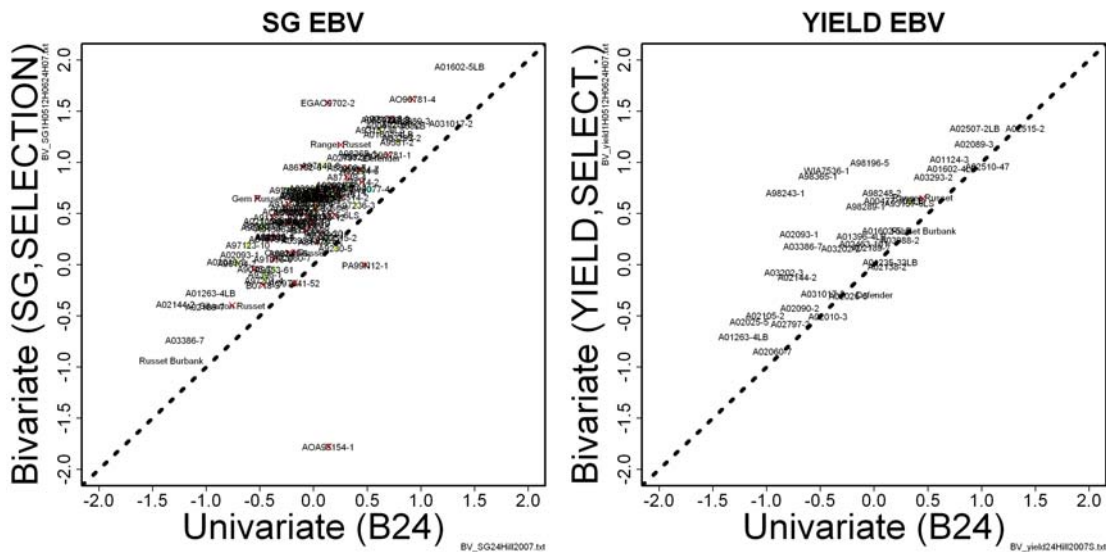


Figure 6. Scattergram of EBV for SG and YIELD computed from a bivariate, multiple trial analysis (y-axis) and univariate, single trial analysis (x-axis). Points on the y-axis have a higher mean and less spread because inclusion of the single hill trial data partly accounts for any selection disequilibrium.

DISCUSSION

This preliminary pilot study has demonstrated that genetic evaluation methodology, which is currently in widespread use in the livestock industries, and emerging as the default methodology used in forestry, has applicability in potato breeding and cultivar selection programs. The breeding values reported can be used to assist decision making in breeding, while total genetic values reported can be used to assist decision making in cultivar selection. Use of genetic values to assist decision making would appear to have greater accuracy and precision than using clonal means derived from a single trial. This is because the method can use and appropriately weight information from multiple sources: different generations, trials; correlated traits; and relatives. Selection of genotypes as parents on the basis of EBV can be done regardless of the testing stage genotypes have reached.

The current study was modest in its scope, in that it included measurements associated with the testing of only two progeny generations. Incorporation of more series of data would help to improve our understanding of the genetic architecture of the species as well as improve the accuracy of prediction in the integrated analysis. Incorporation of data and pedigree from testing stages beyond preliminary yield tests (24 Hill trials) would also be beneficial. Selected material from the Aberdeen 24 hill trials enters Tri-State and Western Region Variety trials. Because these trials are located in other states, the potential for genotype by environment interaction (GxE) to occur becomes an issue. In POTATOPLAN genotype by environment interaction is accommodated by treating different testing sites as different traits (Falconer and MacKay, 1996). However for Tri-

State and Regional trial data to be integrated into the analysis for objective comparison it would also be necessary to integrate other programs' data that is generated prior to the Tri-State testing stage.

The definition of the pedigree in the current study is only preliminary. Pedigree records for the Aberdeen program prior to 1992 have yet to be converted into digital format. Once this has been completed it will be straightforward to incorporate many more generations of ancestors into the analysis. A better pedigree definition will assist in obtaining estimates of genetic parameters describing the population. The narrow genetic base in the cultivated potato (Simmonds, 1969) will have increased average homozygosity levels in advanced generations, thereby reducing variance. Correct knowledge of the genetic variance and a clones inbreeding coefficient is important for determining the size of the Mendelian sampling component of the additive and non-additive genetic value for the clone.

ACKNOWLEDGEMENTS

The financial assistance of AUSINDUSTRY is gratefully acknowledged.

REFERENCES

- Dutkowski, G.W., McRae, T.A., Powell, M.B., Pilbeam, D.J., Joyce, K., Tier, B., and Kerr, R.J. 2006. Benefits from data and pedigree integration in genetic evaluation. In: Proc. 13th Australasian Plant Breeding Conference, April 18-21, 2006. Christchurch, New Zealand. CD-ROM
- Falconer D.S., and MacKay T.F.C. 1996. Introduction to Quantitative Genetics. Longman Scientific and Technical, New York
- Gilmour, A.R., Gogel, B., Cullis, B.R., and Thompson, R. 2006. ASReml, User Guide. Release 2.0. Hemel Hempstead. VSN International Ltd, UK
- Goth, R.W., and Haynes, K.G. 2000. Evaluation of potato genotypes for severity of verticillium wilt, yield and specific gravity in Maine. *Amer. J. Potato Res.* 77:89-94.
- Jakuczun, H. and Zimnoch-Guzowska, E. 2004. Inheritance of glucose content in tubers of diploid potato families. *Amer. J. Potato Res.* 81:359-370
- Pavek, J.J. and Corsini, D.L. 1994. Inheritance of resistance to warm-growing-season fungal diseases, In: J.E. Bradshaw, and G.R. Mackay, eds, *Potato Genetics*, Wallingford, pp.403-410.
- Pereira, A. Da S., Tai, G.C.C., Yada, R.Y., Tarn, T.R., Souza-Machado, V., and Coffin, R.H. 1994. Effect of selection for chip color on some economic traits of potatoes. *Plant Breeding* 113:312-317.
- Sinha, N.K., Cash, J.N., and Chase, R.W. 2008. Differences in sugars, chip color, specific gravity and yield of selected potato cultivars grown in Michigan. *Amer. J. Potato Res.* 69:385-389.
- Simmonds, N.W. 1969. Prospects of potato improvement. Scottish society for research plant breeding report to the annual meeting. Scottish Plant Breeding Station, Pentlandsfield, pp. 18-38.

2009 14th Australasian Plant Breeding Conference (APBC) and
11th Congress of the Society for the Advancement of Breeding Research in Asia and Oceania (SABRAO)
Contemporary Crop Improvement: A Tropical View
Cairns, Queensland, Australia, 10-14 August 2009

Tarn, T.R., Tai, G.C.C, De Jong, H., Murphy, A.M., and Seabrock, J.E.A. 1992. Breeding potatoes for long-day temperate climates. *Plant Breeding Reviews*. 9:217-332.