

ARTICLE

Open Access

Turbocharging introgression breeding of perennial fruit crops: a case study on apple

Satish Kumar¹, Elena Hilario², Cecilia H. Deng ² and Claire Molloy¹

Abstract

The allelic diversity of primitive germplasm of fruit crops provides a useful resource for introgressing novel genes to meet consumer preferences and environmental challenges. Pre-breeding facilitates the identification of novel genetic variation in the primitive germplasm and expedite its utilisation in cultivar breeding programmes. Several generations of pre-breeding could be required to minimise linkage drag from the donor parent and to maximise the genomic content of the recipient parent. In this study we investigated the potential of genomic selection (GS) as a tool for rapid background selection of parents for the successive generation. A diverse set of 274 accessions was genotyped using random-tag genotyping-by-sequencing, and phenotyped for eight fruit quality traits. The relationship between 'own phenotypes' of 274 accessions and their general combining ability (GCA) was also examined. Trait heritability influenced the strength of correspondence between own phenotype and the GCA. The average (across eight traits) accuracy of predicting own phenotype was 0.70, and the correlations between genomic-predicted own phenotype and GCA were similar to the observed correlations. Our results suggest that genome-assisted parental selection (GAPS) is a credible alternative to phenotypic parental selection, so could help reduce the generation interval to allow faster accumulation of favourable alleles from donor and recipient parents.

Introduction

Growing demand for nutrient-rich foods, increasing biotic and abiotic risks in fruit production, and evolving consumer preferences require development of novel cultivars. The narrow genetic base of advanced genetic material generally inhibits introduction of new traits, hence utilising wild or semi-wild gene pools as sources of genetic variation is a common breeding practice. Introduction of novel genes/traits from exotic germplasm have been achieved through two main strategies¹: firstly, pre-breeding of primitive/exotic germplasm using recurrent selection, then using enriched germplasm as a source for crossing with elite lines; and secondly, backcrossing or pseudo-backcrossing (PBC) for introgression of known major genes from the donor into elite lines or commercial

cultivars. The first approach increases the frequency of desirable genes and can help develop 'pre-bridging' germplasm, which can later be used as a pollen source for making crosses with elite lines².

Pre-breeding of fruit crops germplasm using recurrent selection is very rare³, so most efforts have focussed on introgression of major genes (e.g. disease resistance) using a PBC strategy where foreground selection is carried out using marker assisted selection (MAS) and the resistant seedlings are then planted in the orchard for fruit quality and other trait evaluation^{4,5}. PBC often results in 'linkage drag' (i.e. linkage between the gene of interest and undesirable genes that are inherited together from the donor parent), so repeated PBC with elite parents is applied to minimise the non-desirable genomic content of the donor parent in selected seedlings. In addition to the foreground MAS for major genes, a marker-based background selection (which involves selection of only those resistant seedlings that have inherited the least proportion

Correspondence: Satish Kumar (Satish.Kumar@plantandfood.co.nz)

¹The New Zealand Institute for Plant and Food Research Limited, Hawkes Bay Research Centre, Havelock North, New Zealand

²The New Zealand Institute for Plant and Food Research Limited, Mount Albert Research Centre, Auckland, New Zealand

© The Author(s) 2020



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

of donor genome) can also be incorporated in a PBC breeding strategy^{6–9}.

The breeding cycle length for outbred perennial fruit crops is long (for example, about 7 years for apple and pear), so pre-breeding generally takes decades to develop parents suitable for crossing with elite lines. Agrotechnical (e.g. growth chamber) methods have been suggested to reduce the juvenility period^{8,10}. Alternatively, a transgenic early flowering approach for the introgression of fire-blight resistance from the wild apple 'Evereste', showed that five breeding cycles could be completed within seven years⁹. However, transgenic plants are still regulated in many countries, so this strategy cannot be used for commercial cultivar breeding at the moment.

Genomic selection (GS) has been shown to be an effective tool to reduce the cycle length for pre-breeding^{2,11,12} and commercial cultivar breeding¹³. In GS-based parental breeding schemes, deleterious variation can be deselected to increase favourable allele frequencies irrespective of their origin, thus minimising linkage drag^{2,12}. Numerous traits of interest could potentially be controlled by multiple genes, so GS would be ideal for introgressing traits of higher genetic complexity relatively faster than conventional breeding^{14–16}.

Estimated breeding values (EBV) are commonly used to rank individuals in a breeding programme. An individual's own phenotype (e.g. phenotypic records on a seedling) is one indicator of its EBV. Alternatively, EBV of an individual/seedling can be obtained from the performance of its progeny. In pre-breeding programmes of perennial fruit crops, own phenotype of seedlings is generally used for selection of parents for the next generation. However, own phenotype may not be a reliable indicator of EBV, especially for low-heritability traits, which could lead to sub-optimal parental selection¹⁷. Given the long time it takes to obtain progeny-based performance, fruit breeders continue with this current practice, accepting the known risks. Schlathölter et al.⁹ proposed using GS to identify parents for the next generation, but the efficacy of genome-assisted parental selection (GAPS) has not yet been fully established for pre-breeding programmes of perennial fruit crops. GAPS would help shorten the parental breeding cycle, so the main objectives of this study were to evaluate: (1) the correlation between seedlings own phenotype and progeny-test performance (i.e. general combining ability; GCA); (2) the accuracy of GAPS to predict own phenotype and GCA.

Material and methods

Plant material and phenotypes

A recurrent selection based pre-breeding programme was initiated during the early 1990s when about 35,000 apple seeds representing 500 open-pollinated families were imported from different countries into New

Zealand¹⁸. An open-pollinated (OP) family represented seeds (which were pollinated by random pollen) harvested from a female parent. There was no passport data available on female accessions, and majority of these accessions belong to the domesticated species *Malus × domestica*. Seedlings were grown on their own roots, and after phenotypic evaluation for adaptation and fruit quality traits, a subset of c. 500 seedlings representing genetic diversity of the first-generation population, were identified. These 500 selections were propagated onto 'M9' rootstock and planted in duplicate at the Havelock North research orchard of Plant & Food Research for further assessment and long-term conservation. Nearly 350 (70%) of these accessions were used as parents to create a second generation, with each parent being involved, on average, in two crosses. The second generation comprised c. 10,000 seedlings; of these about 200 accessions (20%) were used as parents to create a third generation³. Outstanding accessions from this pre-breeding programme have been used in Plant & Food Research's cultivar/parental development programmes.

A set of 274 accessions was chosen for this study, taking completeness of own phenotypic data as well as progeny-test data into consideration. Own phenotypes of each accession were recorded, and their GCA performance was obtained based on the phenotypes of their progenies. Because of the time it takes for progeny testing, own phenotypes and progeny performance were observed over different years. Six fruits from each plant were assessed over two consecutive years. Fruit were harvested when judged mature based on a change in skin background colour, and when the starch pattern index was between 2 and 3 using a scale from 0 (full starch) to 7 (no starch). Six fruits from each plant were harvested each fruiting season over 2 consecutive years, and were stored for 42 days at 0.5 °C, then a further 1 day at 20 °C before evaluation. The phenotypes used for this study included fruit weight (AVFW; g), russet coverage (RUSS; 0–9 scale), fruit firmness (FIRM; kg/cm³), soluble solids content (SSC; %), crispness (CRISP; 0–9 scale), juiciness (JUIC; 0–9 scale), astringency (ASTR; 0–9 scale) and titratable acidity (TA; percentage of malic acid in fruit juice). The average value of six fruits were used to represent each phenotype of each accession. Further details of harvesting and assessment protocol were reported previously³.

DNA extraction

Total genomic DNA was extracted from 274 accessions using the Qiagen DNeasy Plant mini kit with the following modifications: Three leaf discs (diameter 6 mm) were obtained with a hand-held 1 Hole Plier Punch sterilised by dipping in ethanol for about 1 min, followed by distilled water for ~15 s, then dried out with lint-free tissues. The leaf discs were placed in a 1.5-mL micro-centrifuge tube

containing lysis buffer and RNase A. The leaf sample was ground by hand with a sterile plastic pestle until most of the tissue was disintegrated. The sample was processed as instructed by the vendor, except that the final elution of DNA was done twice, with 1 h of incubation for each elution event. Both eluates were combined and quality assessed on a Fragment Analyzer (Genomic DNA High Sensitivity kit). DNA preps with a Genomic DNA Quality number equal or larger than 6.4 were considered acceptable (threshold value set to 10,000 bp). The average concentration was 15 ng/μL (1.4 ng/μL minimum, 63.2 ng/μL maximum).

Random tagged genotyping-by-sequencing library preparation

As a test to verify that the observed range of DNA concentration would produce random-tagged DNA for library preparation, a set of 64 preps representing such range was amplified by PEP-PCR (Primer extension pre-amplification) and enriched by TD-PCR (touch down) as described by Hilario et al.¹⁹. After confirming that enough amplicon template could be obtained from this range of DNA concentrations, all DNA samples were processed and random tagged genotyping-by-sequencing (rtGBS) libraries prepared using the restriction enzyme BamHI. This enzyme was chosen after screening 10 restriction enzymes bioinformatically and later experimentally, according to the calculation sheet provided by the Gene Pool at the University of Edinburgh (<http://genomics.ed.ac.uk/>) by Mark Blaxter's group, and uploaded in protocols.io (<https://www.protocols.io/view/genotyping-by-sequencing-gbs-library-protocols-kzmcx46/abstract>).

Two quality control steps were implemented: the first after random-tagging the DNA, and the second after amplifying the rtGBS library. Any samples that did not produce a library were repeated. Once all samples produced a library, five pools were prepared by combining a maximum of 56 libraries per pool. The pools were concentrated and quantified by fluorescence before and after size selection using a 3% agarose gel (250–650 bp). The sized fragments were assessed by gel electrophoresis, quantified by fluorescence and sent to Australian Genome Research Facility (AGRF), Brisbane, for sequencing. Each pool of rtGBS libraries was sequenced in one lane, on the Illumina HiSeq2000 platform, in single-end mode.

Sequencing data analysis and variant calling

Sequencing data were downloaded from AGRF and data integrity confirmed with md5sum. Data quality was checked with FastQC-0.11.2. Barcode ligation proficiency and enzyme digestion efficiency were verified with fastq-multx in ea-utils.1.1.2-537 package in two steps: sample barcodes (sequence trimmed after de-multiplexing) and BamHI residue sequence (not-trimmed), both with a

minimum phred score of Q20. For genotyping, we used the TASSEL3 pipeline²⁰. First of all, 64-base long unique tags were extracted from the compressed fastq file for each pool through 'FastqToTagCountPlugin' with a minimum count of 3. Tags from the five pools were merged with 'MergeMultipleTagCountPlugin' to create the overall tag count table, then converted to Fastq format with 'TagCountToFastqPlugin'. Afterwards tag sequences were mapped to *Malus_x_domestica.v1.0*-primary genome²¹ using Bowtie2²² with parameters '--very-sensitive --end-to-end' in multi-thread mode. The Sequence Alignment Map (SAM) file was converted to Tags on Physical Map (topm) format using 'SAMConverterPlugin' followed by 'ModifyTBTHDF5Plugin' to achieve HDF5 format for rapid accessing of large datasets. In the meantime, Tags by Taxa (TBT) data in HDF5 format were created from all the raw sequencing files together with the tag count table using the SeqToTBTHDF5Plugin. To speed up variant calling, we ran the TagsToSNPByAlignmentPlugin for each pool against each *Malus* chromosome in parallel. Single nucleotide polymorphisms (SNPs) called on the same chromosome from different pools were merged with MergeDuplicateSNPsPlugin, while duplicate samples were combined with MergeIdenticalTaxaPlugin. The combined SNPs were filtered using GBSHapMapFiltersPlugin with a minimum site coverage (-mnScov) 0.9, minimum taxa coverage (-mnTCov) 0.8 and minimum minor allele frequency (-mnMAF) 0.05. The missing SNP genotypes were imputed using BEAGLE²³ through the R package Synbreed²⁴.

Analysis of own phenotypes

As the measurements of own phenotypes of 274 accessions were repeated over two successive years, the following individual-tree mixed linear model (MLM) accounting for permanent environment effect was used for each trait:^{13,25}

$$y = Xb + Za + Zp + e \quad (1)$$

where y is the vector of observations, b is the vector of fixed effects (overall mean, harvest year); a , p and e are vectors of random additive genetic effects, permanent environment effects, and residual effects, respectively. The matrix X is the incidence matrix for the fixed effects and Z is the incidence matrix relating observations to seedlings. The associated variances with the random effects, a , p and e were σ_a^2 , σ_p^2 and σ_e^2 , respectively. Pedigree records were used to derive additive genetic relationships among all 274 accessions, and then ASReml software²⁶ was used to obtain best linear unbiased prediction of additive effects (BLUP-BV) of each accession for each trait. These BLUP-BVs were later used for GS model development.

Table 1 The overall average and standard deviation (SD) of various phenotypes (AVFW, RUSS, AVFF, SSC, CRISP, JUICE, ASTR, TA) of parental and progeny populations

Trait (unit)	Own phenotype		Progeny phenotype	
	Average	SD	Average	SD
AVFW (g)	136	90	112	65
RUSS (0–9)	2.5	0.99	2.4	0.96
AVFF (kg/cm ³)	7.8	2.5	6.8	2.3
SSC (°Brix)	15.4	1.8	14.2	2.1
CRISP (0–9)	3.5	1	3.3	1.2
JUIC (0–9)	2.6	1.2	3.1	1.4
ASTR (0–9)	0.9	1.1	1	1.2
TA (%)	0.83	0.37	0.72	0.51

AVFW fruit weight, RUSS russet, AVFF fruit firmness, SSC soluble solids content, CRISP crispness, JUICE juiciness, ASTR astringency, TA titratable acidity

Analysis of progeny test data

For each trait, the GCA of all 274 accessions was obtained from phenotypes of their offspring. Individual-tree MLM (Eq. 1) was used for this analysis, and additive genetic relationships among offspring, derived from their pedigree records, were taken into account. Estimates of narrow-sense heritability (h^2) of each trait were obtained as the ratios of additive variance (σ_a^2) to the total phenotypic variance ($= \sigma_a^2 + \sigma_p^2 + \sigma_e^2$). The correspondence between BLUP-BV (derived from own phenotype) and GCA (derived from progeny phenotypes) was estimated as the product-moment correlation between these two measures.

Genomic prediction and validation

BLUP-BV, obtained from Eq. 1 using own phenotypes of 274 accessions, were used as ‘phenotypes’ for the development and validation of genomic predictions as:

$$y = \mu \mathbf{1}_n + Zg + \varepsilon \quad (2)$$

where \mathbf{y} is a vector (274×1) of BLUP-BVs for a given trait; μ is an intercept, $\mathbf{1}_n$ is a vector of 1 s; \mathbf{Z} is a design matrix relating observations to accessions, \mathbf{g} is a vector of genomic breeding values with a normal distribution; $\mathbf{g} \sim N(0, \mathbf{G}\sigma_a^2)$, where \mathbf{G} is a realised genomic relationship matrix derived using genome-wide SNPs²⁷.

For each trait, the accuracy of genomic prediction was assessed through an 11-fold cross-validation. Records were randomly partitioned into 11 subsets of 25 records each (except one subset with 24 records). The records in one subset were in-turn set to missing values and were predicted using the model developed from the remaining 10 subsets, until phenotypes in all subsets were predicted.

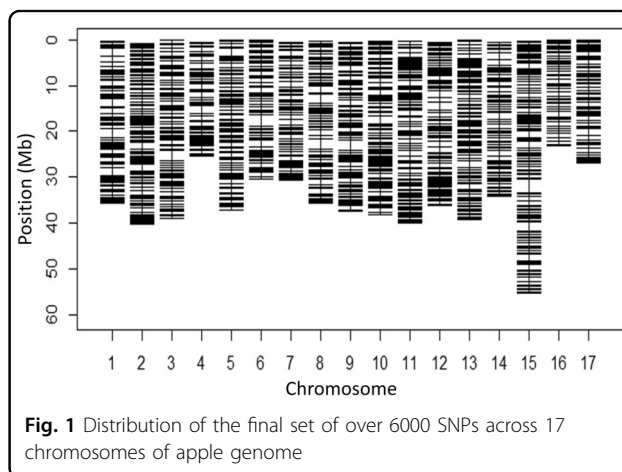


Fig. 1 Distribution of the final set of over 6000 SNPs across 17 chromosomes of apple genome

This 11-fold cross-validation scheme was repeated 20 times, and the prediction accuracy was calculated as the correlation between observed BLUP-BV and predicted genomic breeding value (GEBV) in the validation set. In addition, correlations between GEBV and GCA were also obtained in each validation set. Equation 2 was implemented using R package BGLR²⁸.

Results

About 5000 offspring (derived from cross-pollination) and 274 parents were phenotyped: the distribution of phenotypes in these two groups is presented in Supplementary Fig. S1. The average fruit weight (AVFW) in parental population was higher than in the progeny population (136 versus 112 g), and fruit firmness (FIRM), soluble solids content (measured as °Brix) and titratable acidity (TA) were slightly higher in the parental population (Supplementary Fig. S1; Table 1). The average values of all other traits were similar in both populations. Except for AVFW, phenotypic variability was similar or higher in the progeny population than in the parental phenotypes (Table 1).

Initially, about 50 K SNPs were obtained following the filtering criteria described in the Methods section. SNPs were subsequently discarded using these criteria: more than 10% missing data; minor allele frequency (MAF) < 5%; and SNPs separated by one bp. As a result, a total of about 6400 high-quality SNPs, varying from 246 SNPs on LG16 to 594 SNPs on LG15, were retained (Fig. 1). Overall the final missing data rate was about 4%, and the missing SNP genotypes were imputed before further analyses. Using these SNPs, the pair-wise genomic relationship among our 274 accessions varied between 0 and 0.7 with an average relationship coefficient of 0.22 (Fig. 2). Within-accession heterozygosity (the proportion of SNP loci that were heterozygous) ranged from 0.11 to 0.44 among 274 accessions, with an average of 0.23.

Narrow-sense heritability (h^2), based on progeny test data, varied from 0.15 (for CRISP and JUIC) to 0.50

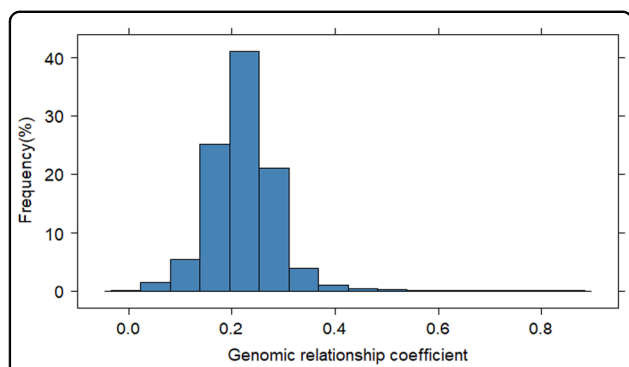


Fig. 2 Distribution of genomic coefficient of relationships among 274 apple accessions

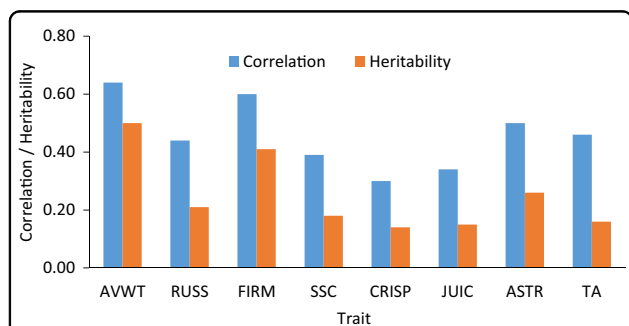


Fig. 3 Correlation between own phenotype and progeny-based performance, and narrow-sense heritability for various apple fruit traits (AVWT fruit weight, RUSS russet, FIRM firmness, SSC soluble solids content, CRISP crispness, JUIC juiciness, ASTR astringency, TA titratable acidity)

(AVFW). After AVFW, fruit firmness was the second most heritable (0.41) trait, followed by ASTR (0.26) and RUSS (0.21) (Fig. 3). Estimated correlation between own performance (BLUP-BV) and progeny-based performance (GCA) was 0.60 or higher for AVWT and FIRM, and varied between 0.30 and 0.50 for all other traits (Fig. 3). Further investigation of data presented in Fig. 3 showed that, across all traits, the relationship between own phenotype and GCA was highly correlated (0.94) with the square-root of trait heritability.

Accuracy of genomic predictions of own phenotype (i.e. correlation between GEBV and BLUP-BV) is shown in Fig. 4. The average prediction accuracy was higher than 0.70 for AVWT, RUSS, FIRM and ASTR; and varied between 0.60 and 0.70 for CRISP, JUIC, SSC and TA (Fig. 4). Small circles outside a box-and-whisker plot represent those values which is less than first-quartile (Q1) or greater than third-quartile (Q3) by more than 1.5 times the interquartile range (Q3–Q1).

The average correlation between predicted own performance (GEBV) and GCA was higher than 0.50 for AVWT, FIRM and ASTR; and between 0.30 and 0.50 for

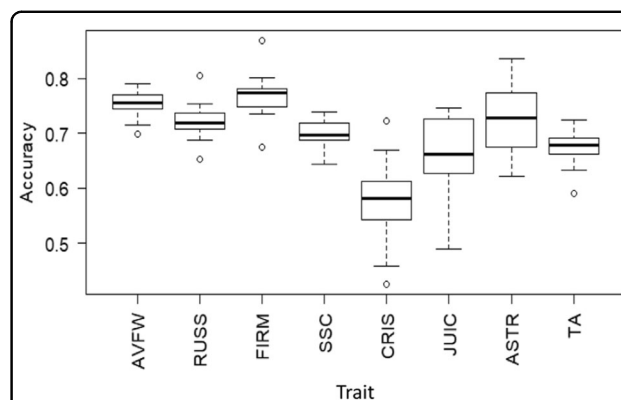


Fig. 4 Genomic predicted accuracy of own phenotypes, and correlation between predicted/realised own phenotype and general combining ability (GCA) for various apple fruit traits (AVWT fruit weight, RUSS russet, FIRM firmness, SSC soluble solids content, CRISP crispness, JUIC juiciness, ASTR astringency, TA titratable acidity)

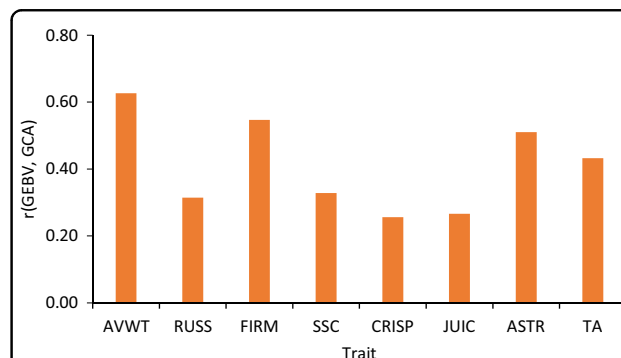


Fig. 5 Correlation between own predicted performance (GEBV) and general combining ability (GCA) for various apple fruit traits (AVWT fruit weight, RUSS russet, FIRM firmness, SSC soluble solids content, CRISP crispness, JUIC juiciness, ASTR astringency, TA titratable acidity)

RUSS, SSC and TA; and between 0.25 and 0.30 for CRISP and JUIC (Fig. 5), very similar to the correlation between observed performance (BLUP-BV) and GCA (Fig. 3).

Discussion

A key goal of creating pre-bridging germplasm is to evaluate and exploit the available genetic variability using recurrent selection cycles, to generate advanced genetic material suitable as parents for cultivar development populations². In contrast, the focus of traditional PBC schemes is to introduce monogenic traits such as disease resistances from donor accessions, and to reduce the unwanted share of wild/primitive genome through multiple PBC cycles. Phenotypic screening for background selection can be slow and expensive. To fast-track introgression, MAS for the donor gene (i.e. foreground selection) could be followed by marker-based background selection to identify seedlings with high genomic similarity to the elite lines^{6–8,29}.

Genomic-similarity based selection could ensure the high recovery of genomic content of the recipient lines when using a high-density genotyping platform. However, in most pre-breeding programmes, the major goal is not necessarily to capture the recipient's genomic background, but to build on some specific economically important traits. High genomic similarity may not necessarily deliver high phenotypic similarity to the elite parent, hence GS models, derived using genotype and phenotypes, would provide a more efficient and targeted approach for background selection. Thus, GS would lead to desired genetic improvement of target traits relatively faster than using genomic-similarity based selection³⁰.

GS would be a more efficient strategy for background selection, especially for complex traits, as favourable alleles from the donor and recipient lines can be targeted across the whole genome. Accuracy of GS studies in apple and pear cultivar breeding were encouraging^{13,31–33}, but the current applications of GS in perennial fruit crops are generally aimed at skipping Stage 1 seedling evaluation so that candidates for Stage 2 evaluation can be identified based on GEBV—hence reducing the cultivar development timeline by at least five years¹³. Although, the candidates selected for Stage 2 testing could also be used as parents of the next generation, it is suspected that their worthiness as parents could be compromised mainly because of recombination events.

The results that relationship between own phenotype and GCA was highly correlated with the square-root of trait heritability, suggested that heritability expresses the reliability of phenotypic value as a guide to the breeding value. For this reason, the determination of heritability of breeding target traits is one of the first objectives in breeding programmes. The accuracy of predicting phenotypes of our 274 accessions was 0.58 for the least heritable trait (CRISP) and 0.75 for the most heritable trait (AVWT). Trait heritability, genetic architecture of the trait, and genetic relatedness between the training and validation sets, are among the key factors influencing the accuracy of genomic predictions^{34–36}. Relatively smaller training population sizes and lower average pairwise genomic relatedness would have contributed to lower prediction accuracies than those reported earlier for the same traits¹³.

This study showed that the degree of correspondence between own phenotype and progeny-based performance (i.e. GCA) was strongly influenced by the heritability i.e. for low-heritability traits, own phenotype could be a poor indicator of GCA²⁵. Nevertheless, pre-breeding programmes of perennial fruit crops (e.g. apple and pear) use seedlings' own phenotypes for background selection of parents of the next generation. Our results suggest that parental selection based on genomic prediction of own phenotypes would be similarly accurate as this conventional selection. Hence, GS approach would obviate the

need for costly and time-consuming phenotypic screening for background selection in pre-breeding programmes. Linkage disequilibrium between pairs of loci would decay over successive backcross cycles, which could require recalibration of GS model every couple of generations.

A combination of genome editing (GE) techniques (e.g. CRISPR-Cas) and GS has been shown to significantly enhance the response to selection in animal breeding schemes³⁷, and a GE-GS combination was also proposed for pre-breeding schemes to develop climate-smart crops³⁸. GE is still a controversial technique, so for the foreseeable future GS-based pre-breeding schemes could improve efficiency in introgressing favourable polygenic traits of donors into high-quality recipient lines whilst also minimizing linkage drag^{2,11}.

In parental breeding programmes of various fruit crops, MAS is commonly used to screen seedlings for monogenic traits (e.g. disease resistance) and then selection for fruit quality traits is carried out based on field phenotypes. Results from this study showed that GS is a credible alternative to phenotypic selection, so for parental breeding of perennial fruit crops we suggest a two-stage selection might be employed—where MAS is used for foreground selection for monogenic traits and GS is used for background selection for fast recovery of the best genomic content of the donor and recipient parents. Prospective parents selected based on MAS and GS could then be propagated in fast-growth chambers to facilitate early transition from the juvenile to the adult phase. Such turbocharged regimes could reduce the breeding cycle of some fruit crops (e.g. apple, pear) to two years, compared with the seven years of a conventional scheme.

Acknowledgements

This Research was supported by The New Zealand Institute for Plant and Food Research Limited's Strategic Science Investment Fund (SSIF) from the New Zealand Ministry of Business, Innovation and Employment (MBIE). Constructive feedback from Dr. Paul Datson and Dr. Luis Gea is thankfully acknowledged.

Conflict of interest

The authors declare that they have no conflict of interest.

Supplementary Information accompanies this paper at (<https://doi.org/10.1038/s41438-020-0270-z>).

Received: 3 November 2019 Revised: 13 February 2020 Accepted: 18 February 2020

Published online: 01 April 2020

References

- Combs, E. & Bernardo, R. Genomewide selection to introgress semidwarf maize germplasm into US Corn Belt inbreds. *Crop Sci.* **53**, 1427–1436 (2013).
- Gorjanc, G., Jenko, J., Hearne, S. J. & Hickey, J. M. Initiating maize pre-breeding programs using genomic selection to harness polygenic variation from landrace populations. *BMC Genom.* **17**, 30 (2016).
- Kumar, S., Volz, R. K., Alspach, P. A. & Bus, V. G. M. Development of a recurrent apple-breeding programme in New Zealand: a synthesis of results, and a proposed revised breeding strategy. *Euphytica* **173**, 207–222 (2010).

4. Bus, V. G. M., Esmenjaud, D., Buck, E. & Laurens, F. in *Genetics and Genomics of the Rosaceae*, Vol. 6 (eds Folta, K. M. & Gardiner, S. E.) Ch. 27 (Springer, 2009).
5. Flachowsky, H. et al. Applying a high-speed breeding technology to apple (*Malus × domestica*) based on transgenic early flowering plants and marker-assisted selection. *N. Phytol.* **192**, 364–377 (2011).
6. Hillel, J. et al. DNA fingerprints applied to gene introgression in breeding programs. *Genetics* **124**, 783–789 (1990).
7. Visscher, P. M. Speed congenics: accelerated genome recovery using genetic markers. *Genet. Res.* **74**, 81–85 (1999).
8. Volz, R. K., Rikkerink, E., Austin, P., Lawrence, T. & Bus, V. G. M. “Fast Breeding” in apple: a strategy to accelerate introgression of new traits into elite germplasm. *Acta Hort.* **814**, 163–168 (2009).
9. Schlathöler, I. et al. Generation of advanced fire blight-resistant apple (*Malus × domestica*) selections of the fifth generation within 7 years of applying the early flowering approach. *Planta* **247**, 1475–1488 (2018).
10. Baumgartner, I. O., Patocchi, A., Franck, L. & Kellerhals, M. Fire blight resistance from ‘Evereste’ and *Malus sieversii* used in breeding for new high quality apple cultivars: strategies and results. *Acta Hort.* **895**, 391–397 (2011).
11. Bernardo, R. Genomewide selection for rapid introgression of exotic germplasm in maize. *Crop Sci.* **49**, 419–425 (2009).
12. Ødegård, J., Sonesson, A. K., Yazdi, M. H. & Meuwissen, T. H. E. Introgression of a major QTL from an inferior into a superior population using genomic selection. *Genet. Sel. Evol.* **41**, 38 (2009).
13. Kumar, S. et al. Genomic selection for fruit quality traits in apple (*Malus × domestica* Borkh.). *PLoS ONE* **7**, e36674 (2012).
14. Yu, X. et al. Genomic prediction contributing to a promising global strategy to turbocharge gene banks. *Nat. Plants* **2**, 16150 (2016).
15. Mascher, M. et al. Genebank genomics bridges the gap between the conservation of crop diversity and plant breeding. *Nat. Genet.* **51**, 1076–1081 (2019).
16. Bernardo, R. Genomewide predictions for backcrossing a quantitative trait from an exotic to an adapted line. *Crop Sci.* **56**, 1067–1075 (2016).
17. Kumar, S. Correlation between clonal means and open-pollinated seedling progeny means and its implications for radiata pine breeding strategy. *Can. J. For. Res.* **36**, 1968–1975 (2006).
18. Noiton, D. & Shelbourne, C. J. A. Quantitative genetics in apple breeding strategy. *Euphytica* **60**, 213–219 (1992).
19. Hilario, E. et al. Random tagging genotyping by sequencing (rtGBS), an unbiased approach to locate restriction enzyme sites across the target genome. *PLoS ONE* **10**, e0143193 (2015).
20. Bradbury, P. J. et al. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* **23**, 2633–2635 (2007).
21. Velasco, R. et al. The genome of the domesticated apple (*Malus × domestica* Borkh.). *Nat. Genet.* **42**, 833–839 (2010).
22. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
23. Browning, S. R. & Browning, B. L. Rapid and accurate haplotype phasing and missing data inference for whole genome association studies using localized haplotype clustering. *Am. J. Hum. Genet.* **81**, 1084–1097 (2007).
24. Wimmer, V., Albrecht, T., Auinger, H. J. & Schön, C. C. Synbreed: a framework for the analysis of genomic prediction data using R. *Bioinformatics* **28**, 2086–2087 (2012).
25. Mrode, R. A. *Linear models for the prediction of animal breeding values*. (CAB Int., 1996).
26. Gilmour, A. R., Cullis, B. R., Harding, S. A. & Thompson, R. *ASReml Update: what’s new in Release 2.00*. (VSN Int. Ltd, Hemel Hempstead, 2006).
27. Van Raden, P. M. Efficient methods to compute genomic predictions. *J. Dairy Sci.* **91**, 4414–4423 (2008).
28. Pérez, P. & de Los Campos, G. Genome-wide regression & prediction with the BGLR statistical package. *Genetics* **206**, 114 (2014).
29. Herzog, E. & Frisch, M. Selection strategies for marker-assisted backcrossing with high-throughput marker systems. *Theor. Appl. Genet.* **123**, 251–260 (2011).
30. Jun-Yan, B. A., Qin, Z. H. & Xiao-Ping, J. I. Comparison of different foreground and background selection methods in marker-assisted introgression. *Acta Genet. Sin.* **33**, 1073–1080 (2006).
31. Muranty, H. et al. Accuracy and responses of genomic selection on key traits in apple breeding. *Hortic. Res.* **2**, 15060 (2015).
32. Minamikawa, M. F. et al. Genome-wide association study and genomic prediction using parental and breeding populations of Japanese pear (*Pyrus pyrifolia* Nakai). *Sci. Rep.* **8**, 11994 (2018).
33. Kumar, S. et al. Marker-trait associations and genomic predictions of inter-specific pear (*Pyrus*) fruit characteristics. *Sci. Rep.* **9**, 9072 (2019).
34. Legarra, A., Aguilar, I. & Misztal, I. A relationship matrix including full pedigree and genomic information. *J. Dairy Sci.* **92**, 4656–4663 (2009).
35. Hayes, B. J., Pryce, J., Chamberlain, A. J., Bowman, P. J. & Goddard, M. E. Genetic architecture of complex traits and accuracy of genomic prediction: coat colour, milk-fat percentage, and type in Holstein cattle as contrasting model traits. *PLoS Genet.* **6**, e1001139 (2010).
36. Kumar, S., Bink, M. C. A. M., Volz, R. K., Bus, V. G. M. & Chagné, D. Towards genomic selection in apple (*Malus × domestica* Borkh.) breeding programmes: prospects, challenges and strategies. *Tree Genet. Genomes* **8**, 1–14 (2012).
37. Jenko, J. et al. Potential of promotion of alleles by genome editing to improve quantitative traits in livestock breeding programs. *Gen. Sel. Evol.* **47**, 55 (2015).
38. Varshney, R. K., Singh, V. K., Kumar, A., Powell, W. & Sorrells, M. E. Can genomics deliver climate-change ready crops? *Curr. Opin. Plant Biol.* **45**, 205–211 (2018).