

# Genomewide prediction to target russet formation in apple

Ashley A. Powell<sup>1,2\*</sup> , Sarah A. Kostick<sup>1</sup>, Rex Bernardo<sup>3</sup> and James J. Luby<sup>1</sup>

<sup>1</sup> Department of Horticultural Science, University of Minnesota, Saint Paul, MN 55108, USA

<sup>2</sup> Department of Veterinary and Biomedical Sciences, University of Minnesota, Saint Paul, MN 55108, USA

<sup>3</sup> Department of Agronomy and Plant Genetics, University of Minnesota, Saint Paul, MN 55108, USA

\* Corresponding author, E-mail: [powel532@umn.edu](mailto:powel532@umn.edu)

## Abstract

Russet formation in apple (*Malus domestica* Borkh.) is a superficial skin disorder that detracts from fruit appearance and is likely controlled by many small-effect quantitative trait loci (QTLs). Genomewide prediction has been reported to be an effective breeding approach when targeting highly quantitative traits in apple. Our objective was to investigate the utility of genomewide prediction for russet formation within an apple breeding program. Germplasm included 1,009 unselected offspring from 13 full-sib families derived from 14 breeding parents. 'Honeycrisp' and 'Minneiska', two breeding parents prone to moderate levels of russet, were highly represented. High-quality single nucleotide polymorphism data (947 SNPs) and three years of shoulder and lenticel russet formation data were leveraged in this study. Moderate predictive abilities ( $r = 0.28$ – $0.35$ ) were observed across training-testing set scenarios and models. In this germplasm, the inclusion of previously detected QTLs as fixed effects in the model did not have significant effects on predictive abilities. Postdiction (retrospective) analyses demonstrated that genomewide predictions and phenotypic observations agreed for 54% of advanced selections. Genomewide prediction is a promising approach when targeting russet formation, a trait that cannot be phenotypically observed in offspring in apple breeding programs until they are past their juvenile phase.

**Citation:** Powell AA, Kostick SA, Bernardo R, Luby JJ. 2024. Genomewide prediction to target russet formation in apple. *Fruit Research* 4: e023 <https://doi.org/10.48130/frures-0024-0016>

## Introduction

Appearance is an important attribute of fruit quality for fresh eating apples (*Malus domestica* Borkh.). Poor appearance can result in fruit being downgraded from the fresh-eating market to the processed market, which can result in a three-to-four-fold loss of wholesale income<sup>[1]</sup>. Russet formation a superficial skin disorder, affects the cuticle layer and forms a brown, corky patch that can be localized or spread across the fruit. In full russet cultivars of apples, russet is not considered a flaw and research has shown full russet fruit to have higher concentrations of sugar<sup>[2]</sup>, which can be explained by the increased levels of water loss in fruit with higher levels of russet<sup>[3]</sup>. In apple cultivars that only have partial russet, low levels of russet formation are typically tolerated but high levels have been associated with increased levels of fruit shrivel in storage and predisposition to cracking, usually around the stem cavity and shoulder area<sup>[3,4]</sup>. Field observations have shown high levels of russet formation on the lenticels are associated with higher levels of infected lenticels. Infected lenticels can result in discolored lenticels and can also reduce storability, though there is a lack of research on the topic in the literature<sup>[4]</sup>.

Phenotyping partial russet formation can be challenging due to the strong influence of abiotic and biotic environmental conditions on russet formation and the challenges associated with subjective visual ratings. Several genomic regions associated with russet formation have been previously reported in apple<sup>[5–12]</sup> with most having small effects or being characterized in a relatively narrow set of germplasm. One study using a full russet cultivar as a parent reported a large effect of QTL on

linkage group 12<sup>[7]</sup>. The use of a full russet cultivar as a parent suggests that for full or extreme levels of russet formation, there is a single large-effect genomic region underlying russet formation. For partial russet formation, a quantitative trait, russet formation levels are likely controlled by many small-effect loci throughout the genome.

Genomewide selection has been reported as an effective breeding approach in many crops, including for quantitative traits in apple<sup>[9,10,12–16]</sup>. Unlike marker-assisted selection (MAS), which relies on a few markers associated with QTLs, genomewide selection uses hundreds to thousands of markers to predict performance (breeding values). Genomewide prediction models use both the phenotypic and genotypic data of one set of germplasm (a training set) to train a model that can predict performance using only genotypic data of another set of germplasm (a testing or validation set). Model accuracy, quantified as predictive ability, is assessed by calculating the correlation between observed and predicted values. Genomewide prediction models are most accurate when a large training set is used, the testing and training sets are closely related, the trait has a high heritability and marker coverage is sufficient across the genome<sup>[17,18]</sup>.

Four previous genomewide prediction studies have included partial russet formation on apple<sup>[9,10,12,14]</sup>. Average predictive abilities for genomewide prediction models for russet formation ranged from  $-0.06$  to  $0.82$ <sup>[9,10,12,14]</sup>, with most studies reporting moderate predictive abilities of approximately 0.3. Previous studies did not include germplasm derived from 'Honeycrisp', an important US apple cultivar and breeding

parent<sup>[19–24]</sup>. Partial russet formation has been observed in both the ancestors and progeny of 'Honeycrisp', including its offspring, 'Minneiska' (Sweetango® apple), in conducive environmental conditions. Previous research by Powell et al. has identified russet formation QTLs in 'Honeycrisp'-derived germplasm<sup>[8]</sup> which can enable the investigation of genomewide prediction with russet formation QTLs as fixed effects and their impact on predictive abilities. Genomewide prediction models utilizing fixed effects have not been studied in relation to russet formation.

In this study, we investigated the utility of genomewide prediction for partial shoulder and lenticel russet formation in 'Honeycrisp'-derived breeding germplasm within the context of the UMN apple breeding program. Both prediction and postdiction (retrospective analysis) were used to determine how useful genomewide prediction models would be in a breeding program setting. Here it was hypothesized that: 1) use of genomewide prediction for russet formation in apple would result in moderate predictive abilities as seen in other studies; 2) inclusion of previously detected russet formation QTLs as fixed effects in genomewide prediction models would result in higher predictive abilities; and 3) in a postdiction analysis, the use of genomewide prediction would identify advanced selections prone to russet formation.

## Materials and methods

### Germplasm and genotypic data

Offspring from 13 pedigree-connected full-sib families ( $n = 1,009$  offspring) were evaluated for russet formation levels. The 13 families were comprised of two half-sib families. Nine families had 'Honeycrisp' as a common parent and the other four families had 'Minneiska', an offspring of 'Honeycrisp', as the common parent. The second parents of each cross comprised 12 individuals, either a cultivar or advanced selection, from a diverse set of parents used in the UMN apple breeding program.

Individuals were genotyped using the International RosBREED single nucleotide polymorphism (SNP) Consortium 8K Illumina Infinium® array v1<sup>[25]</sup> or the Illumina Infinium 20K array<sup>[26]</sup> and SNPs common to both arrays were used for analysis ( $n = 2,213$ ). The germplasm and genotyping were previously described in detail by Powell et al.<sup>[8]</sup>. Further SNP curation was done via SNP-QC, a software that (1) imputes missing marker data and (2) removes markers that are monomorphic, redundant, have too many missing values, or have a low minor allele frequency<sup>[27]</sup>. Parameters for SNP-QC are reported in Supplemental Table S1. The total marker number was reduced from 2,213 to 947 markers with an average of 56 markers per linkage group (Supplemental Table S1).

For the postdiction analyses, an additional 163 advanced selections were used. Genotyping for these selections followed the same methods described in Powell et al.<sup>[8]</sup>. SNP-QC was rerun to include both the training set and advanced selections and the total markers used in postdiction analyses were 974.

### Phenotypic data

Phenotypic russet formation data on the 1,009 unselected seedlings, previously described in Powell et al.<sup>[8]</sup>, were used in genomewide prediction analyses. For this paper, we have summarized the results here and family summaries for

shoulder and lenticel russet formation are summarized in Supplemental Tables S2 and S3. Visual ratings were collected across three years, 2017–2019, using a 1–10 scale with each ordinal value associated with a bin of 10% (i.e., 1 = 1%–10% area affected by russet formation). Year effect had been observed<sup>[8]</sup> and best linear unbiased predictions (BLUPs) were calculated and used as phenotypic values in genomewide prediction analyses. BLUPs are referred to throughout this paper, and the study data was collected and calculated in, as 'russet ratings'. Shoulder russet ratings captured the proportion of fruit shoulder surface area, which ranged from the edge of the stem cavity to the body of the fruit that was affected by russet. Lenticel russet ratings estimated the proportion of lenticels affected by russet. Average rating of 2.0 for shoulder ratings ( $sd = 1.6$ , range 1.0–10.0) and 4.4 for lenticel ratings ( $sd = 2.5$ , range 1.0–10.0) were observed. Russet formation data are described in detail by Powell et al.<sup>[8]</sup>.

In this study, the severity of lenticel russet formation was also investigated. If an individual had severe lenticel russet in any year, meaning russet formation on the lenticel exceeded the immediate area of and around the affected lenticel, a value of one was added to the lenticel russet rating to capture an individual's susceptibility to developing severe lenticel russet and moving them up a russet rating category. Seventy-one individuals had severe lenticel russet in at least one year of data collection and their lenticel russet ratings were adjusted.

### Genomewide prediction

Genomewide predictions for shoulder and lenticel russet ratings were computed through models implemented via ridge-regression best linear unbiased prediction (RRBLUP) using the software, RRBLUP2<sup>[27]</sup>. For shoulder russet ratings, two genomewide prediction models were tested: (1) all SNPs included as random effects and (2) most SNPs were included as random effects while SNPs at previously detected russet formation QTL regions<sup>[8]</sup> were included as fixed effects. Significant SNPs within the QTL regions were identified and chosen using the output generated by RRBLUP2 for the random effect models. RRBLUP2 generates the effect each SNP has for every prediction model. SNPs within the QTL region with the largest absolute SNP effect value were chosen. For both LG2 and LG6, the SNP chosen had the largest or were tied for the largest absolute SNP effect value within the QTL region for all random effect prediction models. An unpublished modified version of RRBLUP2, which enables the inclusion of fixed effects, was utilized to compute the model with fixed effects. The two SNPs chosen to be fixed effects were called in the parameter files and the software removed them before calculating the random effects. As no large-effect QTLs associated with lenticel russet formation have been reported<sup>[8]</sup>, a model that included all SNPs as random effects was used for genomewide predictions of lenticel russet.

For each genomewide prediction model, a single untested family was used as the testing set. This procedure replicated scenarios often seen in the UMN apple breeding program where half-siblings or more distantly related material are the only available related germplasm for phenotyping. Three different types of training sets were used: (1) All = all breeding germplasm minus the test family; (2) Honeycrisp = all 'Honeycrisp'-derived families minus the test family; and (3) Minneiska = all 'Minneiska'-derived families minus the test family. Each

family was used as a testing set in at least two training set scenarios. Each family was used as a testing set for 'All' and 'Honeycrisp' training sets. The four Minneiska-derived families were also used in a genomewide prediction model where each family was used as a testing set using the other 'Minneiska'-derived families as a training set. Predictive abilities were quantified as the Pearson correlation coefficient ( $r$ ) between observed and predicted russet ratings in the testing set. A complete list of all genomewide prediction models examined with a summary of testing and training sets is described in [Supplemental Table S4](#).

### Postdiction

Retrospective analyses (i.e., postdiction) were conducted to determine if genomewide prediction would have enabled identification of advanced selections prone to russet formation in University of Minnesota's apple breeding program in comparison to traditional phenotypic selection. Russet ratings were predicted for 163 advanced selections using all unselected offspring ( $n = 1,009$ ) as the training set. Field notes on advanced selections were composed of an overall score (0–9) and/or notes on russet formation which were then translated into an estimated shoulder russet rating (1–10).

Assessment of postdiction analyses was conducted by comparing 2021 and 2022 breeder field notes (i.e., consensus ratings from breeder personnel) to predicted values for a subset of advanced selections with ( $n = 76$ ). Due to the nature of the postdiction analysis (e.g., different years for data collection, different assessment methods for russet formation, etc.), comparisons were made on a broad basis by independently categorizing individuals as having 'high' or 'low' levels of russet formation. Categorization of 'high' and 'low' was conducted using the mean of each population as the 'high'/'low' threshold. The mean russet rating for the 76 advanced selections based on breeder field notes was 2.2 (range = 1.0–7.0) and therefore 2.2 was used as the threshold to categorize advanced selections as having 'high' or 'low' levels of observed russet formation ([Supplemental Table S5](#)). Predicted values for advanced selections were also independently used to categorize advanced selections as having 'high' or 'low' levels of predicted russet formation based on the average predicted russet rating. Mean russet ratings for genomewide prediction values, using the All germplasm training set, was 2.0 (range = 0.1–4.0) and therefore 2.0 was used as the threshold to categorize individuals as having 'high' or 'low' levels of predicted russet formation. To assess if genomewide prediction would have affected the accurate categorization of russet formation levels, a chi-square test was performed. The null hypothesis was that genomewide prediction used retrospectively would not have had an impact on categorization accuracy and individuals would have been evenly categorized into 'high' and 'low' bins regardless of phenotypic values. Lenticel russet formation was not investigated as lenticel russet was not specifically considered in the breeder field evaluations and no postdiction analyses could be done.

### Assessment of models for breeding purposes

Sensitivity analyses for both prediction and postdiction were conducted to assess the effects of culling thresholds and genomewide prediction models on culling intensity and accuracy. Accuracy was quantified as the proportion of individuals correctly classified in 'keep'/'cull' categories that are

dependent upon culling thresholds. Two types of misclassifications were calculated and referred to as Type A and Type B errors. A Type A error described the event when an individual was culled based on predicted values but would have been kept based upon phenotypic selection. A Type B error was when an individual was kept based on predicted values but would have been culled based on phenotypic selection. Culling thresholds for shoulder russet formation were assessed using the University of Minnesota's apple breeding program standards (maximum allowance of shoulder russet formation ~20%) with two less stringent categories also assessed for cases of multiple trait selection in which flexibility of threshold is required to accommodate higher priority traits. Culling thresholds tested for shoulder russet formation were 2.0, 2.5, and 3.0. Since this is the first investigation of lenticel russet formation, culling thresholds evenly distributed across the phenotyping range was used. Culling thresholds of 2.5, 5.0, and 7.5 for lenticel russet formation were tested.

To compare genomewide selection to marker-assisted selection for shoulder russet formation, marker-assisted selection was included in the above sensitivity test. In this study, marker-assisted selection for russet formation was based on alleles at QTLs identified in Powell et al.<sup>[8]</sup>. Powell et al.<sup>[8]</sup> reported that the presence of two LOW alleles was associated with lower shoulder russet formation, regardless of the other alleles present<sup>[8]</sup>. Therefore, individuals with fewer than two LOW alleles across the LG2 and LG6 QTLs would have been culled based on marker-assisted selection.

## Results

### Genomewide prediction

Moderate predictive abilities (mean  $r = 0.28$ – $0.35$ ) were estimated for both shoulder and lenticel russet formation across models and training sets ([Table 1](#), [Supplemental Tables S2 & S3](#)).

#### Random effects model

Mean predictive abilities for shoulder russet formation using the random effects model were 0.33 for the All training set, 0.30 for the Honeycrisp training set, and 0.28 for the Minneiska training set ([Table 1](#), [Supplemental Table S2](#)). When using the All germplasm training set, the test families with the highest predictive abilities were 'Honeycrisp' × MN1964 ( $r = 0.60$ ) and 'Honeycrisp' × 'Jonafree' ( $r = 0.58$ ) whereas the test families with the lowest predictive abilities were 'Honeycrisp' × AA44 ( $r = 0.12$ ), 'Honeycrisp' × MN1915 ( $r = 0.15$ ), 'Honeycrisp' × MN1836 ( $r = 0.18$ ), and 'Minneiska' × 'Wildung' ( $r = 0.18$ ). Predictive abilities for 'Honeycrisp'-derived families were consistent between training sets and each of the training sets predictive abilities were within 0.05 of the All training set predictive abilities per family. 'Minneiska'-derived families tended to have lower predictive abilities when using the Honeycrisp germplasm training set compared to using the All or Minneiska germplasm training set ([Supplemental Table S2](#)).

Predictive abilities for lenticel russet formation were similar across training sets and were 0.34 using the All training set, 0.30 for the Honeycrisp training set, and 0.35 for the Minneiska training set ([Table 1](#), [Supplemental Table S3](#)). When using the All germplasm training set, the families with the highest predictive abilities were 'Minneiska' × MN1702 ( $r = 0.53$ ),

**Table 1.** Summary of genomewide prediction for shoulder and dual lenticel russet ratings.

Trait <sup>a</sup>	Training set <sup>b</sup>	Fixed effects <sup>c</sup>	Mean predictive ability (r) <sup>d</sup>	SD <sup>e</sup>	Minimum <sup>f</sup>	Maximum <sup>g</sup>
Shoulder	All families	No	0.33	0.15	0.12	0.60
	Honeycrisp-derived families		0.30	0.15	0.11	0.57
	Minneiska-derived families		0.28	0.15	0.09	0.46
	All families	Yes	0.33	0.15	0.10	0.62
	Honeycrisp-derived families		0.31	0.15	0.09	0.58
	Minneiska-derived families		0.35	0.18	0.08	0.48
Lenticel	All families	No	0.34	0.10	0.23	0.53
	Honeycrisp-derived families		0.30	0.11	0.11	0.47
	Minneiska-derived families		0.35	0.10	0.26	0.46

<sup>a</sup> Shoulder russet ratings and the dual lenticel russet ratings; <sup>b</sup> Three training sets were used: (1) all breeding germplasm minus untested family, (2) all 'Honeycrisp'-derived families minus untested family, and (3) all 'Minneiska'-derived families minus untested family; <sup>c</sup> Fixed effect model included two SNPs at the previously reported LG2 and LG6 shoulder russet QTLs<sup>[6]</sup> as fixed effects; <sup>d</sup> Correlation between observed and predicted values using Pearson correlation averaged across testing set families; <sup>e</sup> Standard deviation of r all runs for each trait-training-model set; <sup>f</sup> Minimum r within each trait-training-model set; <sup>g</sup> Maximum r within each trait-training-model set.

'Honeycrisp' × MN1836 ( $r = 0.49$ ), and 'Honeycrisp' × 'Minnewashta' ( $r = 0.47$ ). The families with the lowest predictive abilities were 'Honeycrisp' × 'WA 2' ( $r = 0.23$ ), 'Honeycrisp' × MN1702 ( $r = 0.25$ ), 'Honeycrisp' × MN1764 ( $r = 0.26$ ), and 'Minneiska' × MN1965 ( $r = 0.26$ ).

### Fixed effects

The inclusion of SNPs at shoulder russet formation QTLs as fixed effects in genomewide prediction models often did not result in significantly different predictive abilities (Table 1, Supplemental Table S2). The predictive abilities for shoulder russet formation ranged from 0.08 to 0.62 when russet formation QTL SNPs were included as fixed effects. Predictive abilities of models that included SNPs at QTLs as fixed effects were higher by a maximum of 0.10 ('Honeycrisp' × 'Jonafree') and lower by a maximum of 0.16 ('Minneiska' × 'MN55') compared to the random effects genomewide prediction models. Mean predictive abilities for shoulder russet formation using a fixed effect model were 0.33 using the All training set, 0.31 using the Honeycrisp training set, and 0.35 using the Minneiska training set. Similar to the results for the random effects model, predictive abilities for shoulder russet formation using a fixed effect model for 'Honeycrisp'-derived families were similar between the All training set and Honeycrisp training set. 'Minneiska'-derived families also tended to have lower predictive abilities when using the Honeycrisp germplasm training set when compared to using the Minneiska or All germplasm training set (Supplemental Table S2).

### Postdiction

Of the 76 advanced selections, 43 (57%) had matching categorizations of 'high' or 'low' russet formation across predicted and observed russet ratings. Twenty advanced selections were categorized as having high levels of observed russet ratings based on breeder field notes and 56 were categorized as having low levels of observed russet ratings. Forty-three advanced selections were predicted to have high levels of russet ratings and 31 were predicted to have low russet ratings. In 43 advanced selections, predicted and observed russet formation categorizations were the same. Five advanced selections had prediction categorizations that underestimated the level of russet formation and 18 advanced selections had prediction categorizations that overestimated the level of russet formation (Supplemental Table S5). Chi-square analysis (Supplemental Table S6) rejected the null hypothesis that

genomewide prediction would not have had an impact on categorization accuracy (i.e., that individuals were randomly categorized into 'high' and 'low' bins).

Thirty-four of the 76 advanced selections have been previously used as parents, with 26 of them having low levels of observed russet formation in breeder field notes. Of those 26, 13 parents were predicted to have high levels of russet formation. One of those 14 parents, was released as a cultivar, 'MN80'. While 'MN80' was predicted to have high levels of russet formation, the breeder field notes categorized it as having low levels of russet formation.

### Sensitivity analysis

For shoulder russet formation, marker-assisted selection and genomewide selection resulted in the same number of individuals being culled ( $n = 461$ ) when the culling threshold was 2.0, whereas selection based on phenotypic values would have only culled 302 individuals. The two DNA-based selection methods agreed for 62% of individuals to be culled (Table 2). When compared to marker-assisted selection, genomewide selection showed a reduction in Type A misclassifications at all culling thresholds. Type A misclassifications are individuals that would have been kept based on phenotypic values but were culled based on predicted values (Table 2). Type B misclassifications, or individuals that were kept based upon predicted values but would have been culled based upon phenotypic values, were consistent across (within 1.8%) culling thresholds and prediction platforms. As the culling threshold increased, overall misclassifications decreased. For lenticel russet formation, genomewide selection followed a similar misclassification trend as with shoulder russet formation errors. Since no QTLs were previously reported, marker-assisted selection could not be examined.

Depending on the culling threshold, between 76 and 146 (46%–90%) of advanced selections ( $n = 163$ ) would have been retained based on predicted values for shoulder russet when using genomewide selection. Seventy-three (56%) of the advanced selections ( $n = 130$ ) would have been kept using marker-assisted selection results. For the advanced selections that had QTL haplotype information, marker-assisted selection and genomewide selection were in agreement for 89 individuals. When using genomewide selection for lenticel russet formation, 14–150 (9%–92%) advanced selections would have been kept depending on culling threshold (Table 2).



**Table 2.** Comparison of culling using genomewide selection (GS) at different thresholds and marker-assisted (MAS) for shoulder russet and genomewide selection at different thresholds for lenticel russet formation.

Model – Trait <sup>a</sup>	Testing set <sup>b</sup>	Culling criteria	Cull (n)	Keep (n)	Type A error (%) <sup>c</sup>	Type B error (%) <sup>d</sup>	
GS – Shoulder	Full-sib families	Over 2.0	461	548	26.0	10.2	
		Over 2.5	248	761	15.0	10.8	
		Over 3.0	105	906	7.1	11.5	
MAS – Shoulder		< 2 LOW alleles	461	548	28.2	9.7	
GS – Lenticel		Over 2.5	885	124	22.6	7.1	
		Over 5.0	331	678	17.0	21.1	
GS – Shoulder		Over 7.5	13	996	0.9	14.4	
		Selections	Over 2.0	87	76	–	–
			Over 2.5	47	116	–	–
Over 3.0	17		146	–	–		
MAS – Shoulder*		< 2 LOW alleles	57	73	–	–	
GS – Lenticel		Over 2.5	149	14	–	–	
		Over 5.0	63	100	–	–	
		Over 7.5	13	150	–	–	

<sup>a</sup> Model comparison between genomewide selection (GS) with all markers as random effects using the all training set and marker-assisted selection (MAS); targeted traits were shoulder russet formation and lenticel russet formation; <sup>b</sup> Germplasm used in the testing set; Breeding = all 13 'Honeycrisp'- and 'Minneiska'-derived families were used as a testing family (n = 1,009); Selections = all advanced selections (n = 163), \* Only 130 individuals were assessed for MAS due to unresolved QTL alleles; <sup>c</sup> Type 1 error defined as individuals that were culled based upon predicted values but would have been kept based upon observed values; <sup>d</sup> Type 2 error defined as individuals that were kept based upon predicted values but would have been culled based upon observed values.

## Discussion

This is the first report, to our knowledge, of genomewide prediction targeting russet formation in 'Honeycrisp'-derived germplasm. 'Honeycrisp' is an important breeding parent that is commonly used in major apple breeding programs<sup>[19–21]</sup> throughout the US due to its revolutionizing crisp texture. 'Honeycrisp' also has a unique ancestral background<sup>[24]</sup> which has not been previously studied in the context of russet formation. Moderate predictive abilities (~0.30) for genomewide prediction were estimated with no significant differences between random effect models and models with previously reported QTL included as fixed effects.

### Predictive abilities for russet formation in this study were consistent with other studies

Predictive abilities in this study ( $r = 0.08–0.62$ ) were consistent with other genomewide prediction studies ( $r = 0–0.38$ ) that examined russet formation. The exception was Kumar et al.<sup>[9]</sup> that reported very high predictive abilities ( $r = 0.84–0.96$ ) for genomewide prediction of russet formation. The reasons are unclear for the higher predictive abilities in the Kumar et al.<sup>[9]</sup> study but possible reasons for high predictive abilities include: high levels of relatedness between the training and testing sets, heritability of russet formation in their germplasm, the environment might have been more conducive to russet formation, and the use of full russet parents.

### Predictive abilities varied among and between different training sets

For shoulder russet formation, we typically observed higher predictive abilities in families that had larger standard deviations (i.e., 'Honeycrisp' × MN1764) and generally saw minimal increases in predictive abilities when using the All training set when compared to models that used only half-sibs to predict other half-sibs. Testing families that had smaller standard deviations and ranges in their phenotypic data (i.e., 'Minneiska' × 'Wildung') had in general the lowest predictive abilities. The observation of a smaller phenotypic range producing lower predictive abilities suggests that the model was over-predicting russet formation in families that had low levels of observed

russet. The lack of significant difference in predictive abilities between half-sib (Honeycrisp and Minneiska training sets) and the All training set is most likely due to pedigree connections among the non-'Honeycrisp' and '-Minneiska' parents. Most of the second parents had pedigree connections to other second parents or to 'Honeycrisp' and 'Minneiska'.

For lenticel russet formation, the highest predictive abilities were achieved when using the All training set, most likely due to the increase in training set size. Five families did perform best when using other half-sibs only. These families did not significantly differ from other testing families regarding their phenotypic ranges or standard deviations, suggesting that an increase in relatedness was more impactful on predictive abilities than an increase in training set size for these families.

### Inclusion of fixed effects did not improve predictive ability for shoulder russet formation

Predictive abilities were not significantly different among genomewide prediction models that used all SNPs as random effects and models that included SNPs with fixed effects. This is consistent with previous findings that the inclusion of fixed effects in genomewide prediction models do not significantly improve prediction unless the QTLs have a large effect, accounting for at least 10% of the genetic variation<sup>[28]</sup>. The two QTLs included as fixed effects, when combined, explained only 12% of the phenotypic variation within this population<sup>[8]</sup> and likely did not have large enough effects to result in significantly higher predictive abilities. It is also possible that one SNP per QTL region was insufficient to capture all the variations found within that region. Powell et al.<sup>[8]</sup> used several SNPs across both QTL regions to capture haplotype effects within this germplasm.

### Genomewide prediction could be useful to help breeders identify advanced selections prone to russet

In this study, genomewide prediction identified 28 selections potentially prone to russet. Breeders had not observed significant russet formation on these 28 selections during two years of advanced clonal testing. Possible overestimation of

russet formation by genomewide prediction models might have been due to differences in the rating systems for data collection, lower levels of relatedness between advanced selections and the full-sib training set, or lack of conducive and similar environments between the years of data collection. Another explanation of the assumed overestimation of predicted russet ratings is that the model could be correctly estimating the propensity for high levels of russet formation to occur on these advanced selections. The environmental conditions during phenotypic data collection may have not been conducive for high levels of russet formation and therefore causing a disagreement between predicted and observed values.

### Breeding implications

Russet formation is a complex trait that will continue to be a challenging breeding target until we can accurately assess the susceptibility to russet, or until we can ascertain the true breeding values of the trait. Calculating the true breeding values would require capturing all genetic variation in an environment conducive to russet formation. The moderate predictive abilities observed for both shoulder and lenticel russet formation allow for the utilization of genomewide selection in apple breeding programs.

If genomewide selection for shoulder russet was employed in the University of Minnesota's apple breeding program or other related breeding germplasm with similar distributions of shoulder russet formation, a reduction in misclassification error would be seen when compared to marker-assisted selection using QTLs identified in Powell et al.<sup>[8]</sup> (Table 2). In germplasm more distantly related, genomewide selection could be employed to predict shoulder russet formation but might be less cost-effective than marker-assisted selection unless a breeder is already using genomewide selection for other traits and has genotypic data.

Although lenticel russet formation is not well-studied, there have been field observations of increased levels of infected lenticels associated with higher levels of lenticel russet formation, which can affect long-term storability. This study is currently the only study to investigate and report an effective DNA-informed breeding strategy for lenticel russet formation. Further investigations into the association between lenticel russet formation levels and infected lenticels could better inform breeders on optimal thresholds for selection criteria.

An advantage of DNA-informed breeding is the reduction in required phenotyping especially for traits that do not have high-throughput phenotyping methods, such as russet formation. To date, there has been one laboratory study that has reported a method of high-throughput phenotyping for russet formation in Asian pears<sup>[29]</sup>. While both genomewide selection and marker-assisted selection have misclassification errors, the benefit of reduction in phenotyping costs must also be considered. For phenotypic selection on russet formation to occur, each tree must reach fruiting maturity (5–10 years) and have a fruiting season that occurs within a year and/or a location that is conducive to russet formation. With knowledge from genomewide prediction, a breeder could modify testing environments or protocols to confirm susceptibility or, ultimately, deploy a cultivar only in locations with low russet incidence.

While this study focused solely on russet formation traits, breeding programs simultaneously target multiple traits at differing prioritization levels. Within the context of the University of

Minnesota's apple breeding program, russet formation, which is undesirable, has become a trait with increasing importance. Apple breeders with other priorities and breeding targets might view russet formation similarly or as desirable. Moderate predictive abilities reported here highlight the potential utility that genomewide prediction can have for russet formation in apple.

### Author contributions

The authors confirm contribution to the paper as follows: conceptualization: Powell AA, Luby JJ; data curation and formal analyses: Powell AA; funding acquisition: Luby JJ; software: Bernardo R; writing original draft: Powell AA, Kostick SA; review and editing: Powell AA, Kostick SA, Luby JJ, Bernardo R. All authors reviewed the results and approved the final version of the manuscript.

### Data availability

All data generated or analyzed during this study are included in this published article, its supplementary information files, and in <https://doi.org/10.1007/s11295-022-01582-7>.

### Acknowledgments

This research was funded by the Minnesota State Agricultural Experiment Station—University of Minnesota Projects MIN-21-040 and MIN-21-097 and by the United States Department of Agriculture—National Institute of Food and Agriculture Specialty Crop Initiative Projects 2009-51181-05808 and 2014-51181-22378.

### Conflict of interest

The University of Minnesota receives royalty payments related to the 'Honeycrisp', 'Minnewashta', 'Wildung', 'Minneiska', and 'MN55' apple cultivars. James J. Luby, and the University of Minnesota have a royalty interest in these cultivars. These relationships have been reviewed and managed by the University of Minnesota in accordance with its conflicts of interest policies.

**Supplementary information** accompanies this paper at (<https://www.maxapress.com/article/doi/10.48130/frures-0024-0016>)

### Dates

Received 25 May 2023; Revised 21 March 2024; Accepted 12 April 2024; Published online 12 June 2024

### References

1. USDA National Agricultural Statistics Service. 2021. *Noncitrus fruits and nuts 2020 summary*. USDA Report. <https://downloads.usda.library.cornell.edu/usda-esmis/files/zs25x846c/sf269213r/6t054c23t/ncit0521.pdf>
2. Mignard P, Beguería S, Giménez R, Fonti I, Forcada C, Reig G, et al. 2002. Effects of genetics and climate on apple sugars and organic acid profiles. *Agronomy* 12:827
3. Khanal BP, Ikiguni GM, Knoche M. 2019. Russetting partially restores apple skin permeability to water vapour. *Planta* 249:849–60

4. Winkler A, Athoo T, Knoche M. 2022. Russetting of fruits: etiology and management. *Horticulturae* 8:231
5. Kunihiya M, Moriya S, Abe K, Okada K, Haji T, et al. 2014. Identification of QTLs for fruit quality traits in Japanese apples: QTLs for early ripening are tightly related to preharvest fruit drop. *Breeding Science* 64:240–51
6. Lashbrooke J, Aharoni A, Costa F. 2015. Genome investigation suggests *MdSHN3*, an APETALA2-domain transcription factor gene, to be a positive regulator of apple fruit cuticle formation and an inhibitor of russet development. *Journal of Experimental Botany* 66:6579–89
7. Falginella L, Cipriani G, Monte C, Gregori R, Testolin R, et al. 2015. A major QTL controlling apple skin russetting maps on the linkage group 12 of 'Renetta Grigia di Torriana'. *BMC Plant Biology* 15:150
8. Powell AA, Kostick SA, Howard NP, Luby JJ. 2023. Elucidation and characterization of QTLs for russet formation on apple fruit in 'Honeycrisp'-derived breeding germplasm. *Tree Genetics & Genomes* 19:5
9. Kumar S, Chagné D, Bink MCAM, Volz RK, Whitworth C, et al. 2012. Genomic selection for fruit quality traits in apple (*Malus domestica* Borkh.). *PLoS ONE* 7:e36674
10. Jung M, Keller B, Roth M, Aranzana MJ, Auwerkerken A, et al. 2022. Genetic architecture and genomic predictive ability of apple quantitative traits across environments. *Horticulture Research* 9:uhac028
11. Migicovsky Z, Gardner KM, Money D, Sawler J, Bloom JS, et al. 2016. Genome to phenome mapping in apple using historical data. *The Plant Genome* 9:plantgenome2015.11.0113
12. Minamikawa MF, Kunihiya M, Noshita K, Moriya S, Abe K, et al. 2021. Tracing founder haplotypes of Japanese apple varieties: application in genomic prediction and genome-wide association study. *Horticulture Research* 8:49
13. Muranty H, Troggio M, Sadok IB, Rifai MA, Auwerkerken A, et al. 2015. Accuracy and responses of genomic selection on key traits in apple breeding. *Horticulture Research* 2:15060
14. Cazenave X, Petit B, Lateur M, Nybom H, Sedlak, et al. 2022. Combining genetic resources and elite material populations to improve the accuracy of genomic predictions in apple. *G3 Genes[Genomes]Genetics* 12:jkab420
15. Kostick SA, Bernardo R, Luby JJ. 2023. Genomewide selection for fruit quality traits in apple: breeding insights gained from prediction and postdiction. *Horticulture Research* 10:uhad088
16. Roth M, Muranty H, Di Guardo M, Guerra W, Patochhi A, et al. 2020. Genomic prediction of fruit texture and training population optimization towards the application of genomic selection in apple. *Horticulture Research* 7:148
17. Combs E, Bernardo R. 2013. Accuracy of genomewide selection for different traits with constant population size, heritability, and number of markers. *The Plant Genome* 6:plantgenome2012.11.0030
18. Crossa J, Pérez-Rodríguez O, Cuevas J, Montesinos-López O, Jarquín D, et al. 2017. Genomic selection in plant breeding: methods, models, and perspectives. *Trends in Plant Science* 22:961–75
19. Brown SK, Maloney K. 2011. *US. Apple tree named 'New York 1'.* *US PP22,228 P3.*
20. Bedford DS, Luby JJ. 2008. *US. Apple tree named 'Minneiska'.* *US PP18, 812 P3.*
21. Barritt, B. 2012. *US. Apple tree named 'WA 38'.* *US PP24,210 P3.*
22. Milkovich M. 2023. U. S. apple industry expects 250 million bushels in 2023. *Good Fruit Grower*. [www.goodfruit.com/u-s-apple-industry-predicts-250-million-bushel-crop/?utm\\_source=New+Post+Hort+Show+List&utm\\_campaign=3c7e03ce64-fresh-bites-2023-08-18s&utm\\_medium=email&utm\\_term=0\\_-194d813140-%5BLIST\\_EMAIL\\_ID%5D](http://www.goodfruit.com/u-s-apple-industry-predicts-250-million-bushel-crop/?utm_source=New+Post+Hort+Show+List&utm_campaign=3c7e03ce64-fresh-bites-2023-08-18s&utm_medium=email&utm_term=0_-194d813140-%5BLIST_EMAIL_ID%5D)
23. Wang Y, Çakır M. 2020. Welfare impacts of new demand-enhancing agricultural products: the case of Honeycrisp apples. *Agricultural Economics* 51:445–57
24. Luby JJ, Howard NP, Tillman JR, Bedford DS. 2022. Extended pedigrees of apple cultivars from the University of Minnesota Breeding Program elucidated using SNP array markers. *HortScience* 57:472–77
25. Chagné D, Crowhurst RN, Troggio M, Davey MW, Gilmore B, et al. 2012. Genome-wide SNP detection, validation, and development of an 8K SNP array for apple. *PLoS ONE* 7:e31745
26. Bianco L, Cestaro A, Sargent DJ, Banchi E, Derdek S, et al. 2014. Development and validation of a 20K single nucleotide polymorphism (SNP) whole genome genotyping array for apple (*Malus × domestica* Borkh.). *PLoS ONE* 9:e110377
27. Bernardo R. 2020. Simple software for genomewide prediction, linkage and association mapping, and quality control of marker data. *Crop Science* 60:515
28. Bernardo R. 2014. Genomewide selection when major genes are known. *Crop Science* 54:68–75
29. Abebe AM, Kim Y, Kim J, Kim SL, Baek J. 2023. Image-based high-throughput phenotyping in horticultural crops. *Plants* 12:2061



Copyright: © 2024 by the author(s). Published by Maximum Academic Press, Fayetteville, GA. This article is an open access article distributed under Creative Commons Attribution License (CC BY 4.0), visit <https://creativecommons.org/licenses/by/4.0/>.