Enhancing the accuracy of genetic values of tested early-stage potato clones utilizing genotypic, pedigree, and drone-retrieved spectral information.

Alessio Maggiorelli, Kathrin Thelen, Vanessa Prigge, Julien Bruckmüller, Katja Muders, Stefanie Hartje, Benjamin Stich

Article - Version of Record

Suggested Citation:
Maggiorelli, A., Thelen, K., Prigge, V., Bruckmüller, J., Muders, K., Hartje, S., & Stich, B. (2025). Enhancing the accuracy of genetic values of tested early-stage potato clones utilizing genotypic, pedigree, and drone-retrieved spectral information. The Plant Phenome Journal , 8(1), Article e70053.
https://doi.org/10.1002/ppj2.70053

Received: 19 May 2025 | Accepted: 26 October 2025
DOI: 10.1002/ppj2.70053

The Plant Phenome Journal OPEN ACCESS

ORIGINAL ARTICLE

Special Section: Phenomic Prediction in Plant Breeding

Enhancing the accuracy of genetic values of tested early-stage potato clones utilizing genotypic, pedigree, and drone-retrieved spectral information.

Alessio Maggiorelli[1,2] | **Kathrin Thelen**[1,3] | **Vanessa Prigge**[4] | **Julien Bruckmüller**[4] | **Katja Muders**[5] | **Stefanie Hartje**[6] | **Benjamin Stich**[1,2,3]

[1]Julius Kühn-Institut (JKI), Institute for Breeding Research on Agricultural Crops, Sanitz, Germany
[2]Institute of Quantitative Genetics and Genomics of Plants (QGGP), Heinrich-Heine-University, Düsseldorf, Germany
[3]Faculty of Agricultural and Environmental Sciences, University of Rostock, Rostock, Germany
[4]SaKa Pflanzenzucht GmbH & Co. KG, Windeby, Germany
[5]Nordring-Kartoffelzucht- und Vermehrungs-GmbH, Sanitz, Germany
[6]EUROPLANT Innovation GmbH & Co. KG, Lüneburg, Germany

Correspondence
Benjamin Stich, Julius Kühn-Institut (JKI), Institute for Breeding Research on Agricultural Crops, Rudolf-Schick-Platz 3a, Sanitz, 18190, Germany.
Email: benjamin.stich@julius-kuehn.de

Assigned to Associate Editor Valerio Hoyos-Villegas.

Funding information
Fachagentur Nachwachsende Rohstoffe, Grant/Award Numbers: 22011818, 2222NR078A

Abstract

The first generations of most clone breeding programs do not allow multi-environment trials due to the lack of sufficient reproductive organs. We integrated the phenotypic records of such an early timepoint, called the A clone stage in potato (*Solanum tuberosum* L.) breeding, into a genomic prediction framework and analyzed (i) if different relationship information can enhance the accuracy of the breeding values from such unreplicated trials, and (ii) whether a weighted combination of these enhanced breeding values with predicted breeding values calculated through genomic prediction can further improve accuracy. Compared to the breeding values derived without considering any relationship information among the clones, enhanced breeding values achieved a higher accuracy for 7–14 traits, depending on the relationship matrix used (multispectral, genetic, pedigree, or a combination). Combining these enhanced breeding values with predicted values from genomic prediction via a heritability index resulted in a weak increase of accuracy (0.5%) for five traits. However, a grid search to find the optimal combination of weights revealed more traits (8–12) with a higher mean increase of accuracy (0.5%–3.8%). Our findings show that the accuracies of unreplicated clones can be increased with easy-to-collect multispectral and pedigree data. The magnitude of this accuracy increase was negatively

---

Abbreviations: AEM, adjusted entry mean; BBCH, biologische bundesanstalt für land- und forstwirtschaft, bundessortenamt und chemische industrie; BLUP, best linear unbiased prediction; EBV, enhanced breeding value; GBLUP, genomic best linear unbiased prediction; GS, genomic selection; GP, genomic prediction; HBV, holistic breeding value; MET, multi-environment trial; NIRS, near-infrared spectroscopy; PS, phenomic selection; SNP, single nucleotide polymorphism; TS, training set; TV, true genetic value; UAV, unmanned aerial vehicle.

correlated with the per plot heritabilities, which makes this approach effective for lowly heritable traits. We see potential for this approach in the early stages of potato breeding programs, like the seedling or single hill stage, where the exact assessment of complex traits is impossible due to low heritability.

Plain Language Summary

This study focuses on ways to improve the accuracy of determining the performance of new potato varieties using different types of information, such as family relatedness information, genetic, and drone image data. The goal was to evaluate if these methods could help in early stages of potato breeding programs where each potato variety is tested only superficially. The results showed that while the current testing methods already gave fairly accurate results, adding extra information improved accuracy further. This was especially true for traits that are less influenced by genetics. Different data types worked better for different traits, but in general, a mix of family relatedness information and drone images was enough to produce good results. Overall, this extra information helps breeders to make accurate selection decisions and we see potential that these methods could be applied in even earlier breeding program stages where the testing quality is usually very low.

1 | INTRODUCTION

A key challenge in developing improved varieties in potato (*Solanum tuberosum* L.) breeding programs is the time investment required to test the performance of new variety candidates that were created by crossing two or more clones (Jansky & Spooner, 2018). This is mainly because potato is a clonally propagated species with a low propagation coefficient. Therefore, a variety candidate undergoes at least three seasons of multiplication until it can be tested in a multi-environment trial (MET) which is, so far, the common method in potato breeding programs to obtain a robust prediction of the performance of a clone across environments (DeLacy et al., 1996). Furthermore, destructively measured traits like chips, French fry color, or bruising tolerance put additional pressure on tuber availability for testing at an acceptable heritability level. One possible way to approach this challenge is genomic prediction (GP) (Meuwissen et al., 2001).

GP is a method that predicts the performance of genotypes by capitalizing on all genetic markers simultaneously, thus, eliminating the need to phenotype. The basis of GP is often a regression model where a precalculated variance-covariance structure, also called relationship matrix, is specified for the genetic effect. A GP model is trained via a training set (TS) with known geno- and phenotypes. Once it is trained, predictions can be made for untested genotypes with a certain accuracy as long as genetic marker data are available to calculate a genetic relationship matrix between the untested genotypes and a training population. Selection based on the predictions calculated via GP is called genomic selection (GS). The potential of GS to select on predicted rather than assessed phenotypes has been shown first for livestock breeding (García-Ruiz et al., 2016; Hayes & Goddard, 2010; Legarra et al., 2008). After that, GS was also quickly tested and applied in plant breeding with promising results for several crops important for food production like wheat (*Triticum aestivum*)(e.g., Daetwyler et al., 2014; He et al., 2016; Juliana et al., 2020), soybean (*Glycine max* L. merr) (e.g., Matei et al., 2018; Stewart-Brown et al., 2019), and maize (*Zea mays* L.) (e.g., Dos Santos et al., 2016; Lohithaswa et al., 2022; Zhao et al., 2012). For potato breeding, GS also showed a high potential for various predicted traits (Enciso-Rodriguez et al., 2018; Habyarimana et al., 2017; Stich & Van Inghelandt, 2018). In addition, various TS compositions have been evaluated (Sverrisdóttir et al., 2018) and the question of how to optimally implement GS in real potato breeding programs has been investigated with the help of computer simulations (Wu et al., 2023).

Since less time has to be invested into testing, GS can help to increase genetic gain by reducing the cycle length of a breeding program. The drawback lies in a selection accuracy, which can be much lower than one depending on the trait. However, besides the above-described classical use of GS, its concept can also be applied to improve the accuracy of already phenotyped genotypes (Endelman et al., 2014). In this second approach, breeding values are calculated by integrating unreplicated phenotypic records from preliminary yield trials into the GP framework which can result in higher prediction accuracies compared to the first use of GS while the need to phenotype is not eliminated. Therefore, this second approach represents a middle way between GS and classical testing through METs. Michel et al. (2017) applied this concept by using phenotypic records from preliminary yield trials of winter wheat to estimate their breeding values and reported higher prediction accuracies than GP alone. Terraillon et al. (2022) also used the same concept to show in a simulation study that GP models can enhance the estimation of genotypic values in unreplicated trials in general. However, the suitability of this approach, namely, to enhance breeding value estimations by incorporating unreplicated early-stage phenotypes, has not been evaluated in the context of clone breeding programs so far.

Although genotyping costs have been reduced dramatically in the last decades (Knol et al., 2016), GS remains expensive for large-scale breeding applications. This is especially true in the early stages of potato breeding programs, which feature very large numbers of variety candidates (Wu et al., 2023). Rincent et al. (2018) proposed a promising alternative to GS called phenomic selection (PS), which circumvents the genotyping bottleneck. In PS, predictions are based on absorbance or reflectance spectra of the genotypes instead of genetic marker profiles. In wheat, Rincent et al. (2018) reported higher or similar accuracies compared to GS depending on the predicted trait by utilizing spectra obtained through lab-based near-infrared spectroscopy (NIRS). Since then, several studies have highlighted the possibility of using unmanned aerial vehicles (UAVs) for spectra acquisition in phenomic prediction with promising results (Galán et al., 2020; Krause et al., 2020; Maggiorelli et al., 2024; Sun et al., 2020). The use of UAVs means a significant reduction in work effort and cost compared to lab-based NIRS and even more so compared to the genotyping that would be required for GS.

After PS was first proposed, several studies investigated how genomic and phenomic predictors could be combined and what effect the integration of both information sources has on predictions. Krause et al. (2019) showed that combining molecular markers with hyperspectral reflectance data in multi-kernel models can increase prediction accuracies in wheat compared to single-kernel approaches, depending on the prediction scenario. Galán et al. (2020) reported similar results for hybrid rye by combining genomic and hyperspectral reflectance-derived relationship matrices in multi-kernel and bivariate models using plant height fitted as a secondary trait. For potato, an increase in predictive ability was reported for 20 out of 22 analyzed traits by combining multispectral reflectance data and genotypic markers (Maggiorelli et al., 2024). These findings were reported in the context of the classical use of GS and PS, which is to predict untested genotypes. However, the question remains open whether spectral phenotypes and their combination with genetic markers can also be used to improve the accurate estimation of genetic values from, for example, early yield trials, which is the second use case of predictive breeding approaches.

The objectives of this study were to (i) investigate whether the accuracy of early-stage potato clone breeding values can be improved by integrating their phenotypic records in combination with genomic, phenomic, pedigree data, or a combination of these in prediction models, and (ii) combine the enhanced breeding values (EBVs) from the first objective with predicted breeding values from GP and assess the effect on breeding value accuracy.

Core Ideas

- Genetic, pedigree, or multispectral relationship information enhance selection accuracies of tested potato clones in unreplicated trials.
- Easy-to-collect pedigree and multispectral data obtained via drone are sufficient to enhance accuracy for most potato traits.
- Accuracy gains from added relationship information are highest for traits with low heritability and, thus, have their highest potential for application in early potato breeding program stages, like the seedling or single hill stage.

2 | MATERIALS AND METHODS

2.1 | Plant material and experimental design

Our study was based on 988 tetraploid *S. tuberosum* L. clones provided and grown by three German breeding companies, namely SaKa (SaKa Pflanzenzucht GmbH & Co. KG), Norika (Nordring- Kartoffelzucht- und Vermehrungs- GmbH), and EUROPLANT (EUROPLANT Innovation GmbH & Co), here referred to as the PotatoTools population. The material consisted of 980 clones at the A clone level, which represents the first stage in a typical potato breeding program in which more than one tuber can be tested but where the number of tubers per clone is still insufficient to test the material in METs. The A clones were randomly selected from all the available material of the respective breeding programs representing four market purpose groups, namely, crisp production (339), French fries production (247), table potato (246), and starch production (148). The A clones have underwent two stages of selection after crossing and were selected by each breeding company from the prior stage (i.e., single hill stage) to be representative of the diversity available at that stage of the breeding program for the respective market purpose

TABLE 1 Family structure of the analyzed clones.

Breeder	Number of full-sib families	Number of clones per family: Min	Max	Mean
EUROPLANT	47	1	23	6.30
Norika	17	10	25	17.65
SaKa	99	1	29	3.88

group. In total across all breeders, the 980 A clones belonged to 163 full-sib families with an average of six clones per family (Table 1). The other eight clones were elite potato cultivars used as comparative checks across environments. These were selected to represent different maturity groups and all market-purpose groups.

Each breeder grew their material in one breeder-specific location in 2019 and then in two breeder-specific locations in the years 2020 and 2021 resulting in 15 different year × location combinations, referred to hereafter as environments (Table 2; Figure S1). Due to the limited number of available tubers per clone harvested from the prior breeding stage (i.e., single hill stage) the clones planted in 2019 were tested with a lower number of plants per plot than in the subsequent years (Table 2). We refer to this environment as the true A clone environment. In a typical potato breeding program, only a subset of the A clones would have been selected to be tested in the B clone stage in more than one location the year after. However, in the PotatoTools project, all clones were retained and evaluated again in the two subsequent years. This means that all A clones were tested in multiple environments, which is usually not the case in a conventional potato breeding program.

The clones of our study were grown in augmented row–column designs with one replicate per clone per environment, except for the eight checks that were present once per block (Table 2). The SaKa experiments at both environments in 2021 were separated into two trials representing two different maturity groups: extra early + early (with three blocks) and medium early + medium late (with five blocks).

Plants were phenotyped by the respective breeder during the growing season and after harvest for 14 traits that were comparable across all environments and breeders (Table S1). These traits included developmental criteria (emergence, foliage development, and maturity), measured traits (yield, starch content, and polyphenol oxidase activity), tuber-specific traits (shape, eye depth, skin type, etc.), as well as disease symptoms (rhizoctonia symptoms and scab symptoms) assessed without any artificial inoculation (Table S1). In detail, yield was measured as kilograms per plot and adjusted for the number of emerged plants per plot. All experiments were conducted using local agronomic management practices.

2.2 | Multispectral data

Multispectral data were obtained by overflight with an UAV of model XR6 Hexacopter from Air6 Systems. The drone was equipped with a Tetracam MicroMCA camera with six channels, two of which measured absorbance at near-infrared level, and took pictures at an altitude of approximately 100 ms above the field. Raw images were (i) calibrated and coregistered with PixelWrench (Tetracam Inc.), (ii) photogrammetrically evaluated with Metashape (Agisoft), and finally, (iii) the plot values were statistically assigned with MiniGIS 2.0 (geokonzept GmbH). Absorbances were measured per plot where the outer 20- to 30-cm width were not included to minimize edge and soil effects. For 2020 and 2021, different cameras of the same model but with different available channels were used, resulting in multispectral data with five overlapping and two different channels across both years (Maggiorelli et al., 2024). Due to a malfunction, the 670 nm channel was not assessed in 2021.

The UAV was used to take pictures on three flight dates in both years. The flight dates were chosen at specific phenological growth stages according to the Biologische Bundesanstalt für Land- und Forstwirtschaft, Bundessortenamt und CHemische Industrie (BBCH) scale (Meier et al., 2018), where the two medium-early maturing check varieties, Agria and Verdi, were chosen as references. The first flight took place at approximately BBCH scale stage 31 (main stem elongation stage with beginning crop cover), the second at approximately stage 65 (first inflorescence flowering stage), and the last flight was performed at approximately stage 91 (beginning senescence) in both years. Multispectral data were gathered and preprocessed by geo-concept GmbH. We used the mean reflectances per plot for each channel and flight date combination for further analyses. Spectral data were scaled, centered, and evaluated for outliers via principal component analysis. We identified 2.8% and 2.6% outliers in the spectra of Windeby 2020 and Windeby 2021, respectively, and set them to missing value. Removed outliers were median imputed.

2.3 | Statistical analyses

2.3.1 | Phenotypic data analyses

Phenotypic data were analyzed for SaKa and across all three breeders, separately. In both locations in 2021, the SaKa genetic material was divided into two distinct trials, namely, early and medium late maturing clones as described above. A trial effect was estimated using the check varieties present in each block as a fixed effect in a linear model (cf. Maggiorelli et al., 2024). Phenotypic observations of the checks and entries were then adjusted for each trait by subtracting the

TABLE 2 Environments (i.e., year–location combinations used for phenotypic evaluation) in the PotatoTools project.

Environment	No. of entries	No. of families	No. of blocks	No. of plants per plot
Europlant 2019 Kaltenberg	299	46	4	10
Europlant 2020 Kaltenberg	297	47	4	16
Europlant 2020 Böhlendorf	287	46	2	16
Europlant 2021 Kaltenberg	300	48	4	16
Europlant 2021 Böhlendorf	300	48	1	16
Norika 2019 Groß Lüsewitz	300	17	2	9
Norika 2020 Groß Lüsewitz	300	17	4	18
Norika 2020 Mehringen	300	17	3	20
Norika 2021 Groß Lüsewitz	297	17	4	18
Norika 2021 Mehringen	300	17	2	20
Saka 2019 Windeby	458	107	8	10
Saka 2020 Windeby	387	99	8	16
Saka 2020 Gransebieth	387	99	8	16
Saka 2021 Windeby	387	99	8	16
Saka 2021 Gransebieth	387	99	8	16

trial effect from each observation in the corresponding trial and environment and were used for all subsequent analyses. After this correction, we analyzed the SaKa data using the following model:

$$Y_{gcrbe} = \mu + G_g + E_e + (GE)_{ge} + B_{be} + R_{re} + C_{ce} + \epsilon_{gcrbe}, \tag{1}$$

where Y_{gcrbe} is the trial-adjusted phenotypic observation of the gth clone in the bth incomplete block, rth row, and cth column, nested in the eth environment; G_g is the effect of the gth clone; E_e is the random effect of the eth environment; $(GE)_{ge}$ is the random interaction effect between the gth clone and the eth environment; B_{be}, R_{re}, and C_{ce} are the random effects of the bth incomplete block, the rth row and cth column, which were all nested in the eth environment; and ϵ_{gcrbe} is the residual error term. The column effect C was only included in the final model, if σ_c^2 was significant as tested by a restricted likelihood-ratio test for the corresponding trait (Crainiceanu & Ruppert, 2004). Outliers were removed based on visual inspections of quantile–quantile normal as well as residuals versus fitted values plots.

We computed the broad-sense heritability on a per plot basis across all SaKa environments as follows:

$$H^2_{\text{plot}} = \frac{\sigma^2_G}{\sigma^2_G + \sigma^2_{GE} + \sigma^2_e}, \tag{2}$$

where σ^2_G is the genotypic, σ^2_{GE} is the genotype × environment, and σ^2_e is the error variance taken from model (1).

To analyze the phenotypic data across all breeders, we used a modified version of model (1) that estimated the genotype–environment interaction effect for the checks as described by Thelen et al. (2025):

$$Y_{gcrbe} = \mu + G_g + E_e + c_g(GE_{ge}) + B_{be} + R_{rbe} + C_{cbe} + \epsilon_{gcrbe}, \tag{3}$$

where, c_g is a dummy variable separating checks with $c_g = 1$ from entries with $c_g = 0$. Except for G_g, all effects were regarded as random. Because of the different naming conventions across breeders and blocks, we nested row and column effects in blocks for this model. Based on this analysis for each trait, records with a standardized residual value greater than 3.5 were considered outliers and were removed from the dataset (Thelen et al., 2025). In the next step, a correction for the check-based block effect was realized as described above for the trial effect, in case of a significant ($\alpha = 0.05$) likelihood ratio test in model (3). The corrected trait values were used for all further analyses.

We calculated four different adjusted entry means (AEMs) for each trait by setting G_g as fixed in either model (1) or model (3) but using different combinations of environments. First, AEMs were calculated across all SaKa environments with model (1) excluding the true A clone environment, that is, SaKa 2019. These AEMs were considered as the true genetic values (TVs) of the SaKa clones. We correlated the breeding values calculated in subsequent analyses with the TVs and referred to the resulting correlation in the following as accuracy.

Furthermore, AEMs including all breeders were calculated with model (3) across (i) all 15 environments, (ii) all environments excluding the true A clone environment of SaKa, and (iii) all EUROPLANT and Norika environments as well as the true A clone environment of SaKa.

In a sub-analysis, adjusted check means were calculated using all possible combinations of environments across all breeders with model 1. Then, we correlated the adjusted check means based on all 15 environments with the respective means of the adjusted check means based on 1–14 environments to investigate the effect of the number of considered environments on the reliability of AEMs as breeding values.

Phenotypic data were preprocessed and analyzed using R version 4.3.2 (R Core Team, 2023) and linear mixed effect models were fitted with the R package lme4 (Bates et al., 2015). Significance of effects of both models can be seen in Tables S2 and S3.

2.3.2 | Estimation of breeding values

Breeding values for the SaKa clones of the true A clone environment were estimated by calculating best linear unbiased predictions (BLUPs) using the following model:

$$Y_{gb} = \mu + c_y(G_g) + c_n(G_g) + B_b + \epsilon_{gb}, \tag{4}$$

where Y_{gb} is the phenotypic observation of clone g in block b; μ is the general mean; c_y and c_n are dummy variables separating the checks and entries, respectively; G_g is the effect of the gth clone; and B_b is the effect of the bth block. As breeding values were calculated for all clones except the checks, all effects except for $c_y(G_g)$ were treated as random. Calculating breeding values this way allowed specification of variance-covariance structures for the entries with $G_g \sim N(0, \boldsymbol{X}\sigma_g^2)$, where σ_g^2 is the genotypic variance of the entries and $\mathbf{X}$ is the relationship matrix. If an identity matrix ($\mathbf{I}$) was used as relationship matrix $\mathbf{X}$, we referred to the resulting values in the following as breeding values. In contrast, EBVs were calculated by capitalizing on other sources of data to derive $\mathbf{X}$.

The calculation of the spectral and genomic relationship matrix was described in detail by Maggiorelli et al. (2024). In summary, the multispectral relationship matrix ($\boldsymbol{M}$) was derived from spectral profiles of SaKa Windeby 2020 and SaKa Windeby 2021 and calculated as $\boldsymbol{M} = \frac{\boldsymbol{SS'}}{m_s}$, where $\mathbf{S'}$ is the transpose of $\mathbf{S}$ which was the spectra matrix of dimensions $n \times m_s$, where n is the number of entries (384) and m_s is the number of flight date and channel combinations (7 channels × 3 flight dates × 2 environments = 42). Spectra of both years were used simultaneously by joining the profiles of both years together for $\mathbf{S}$. The genomic relationship matrix $\mathbf{G}$ was calculated as $\boldsymbol{G} = \frac{\boldsymbol{ZZ'}}{m_g}$, where $\mathbf{Z}$ is the matrix of single nucleotide polymorphisms (SNPs) with the dimensions $n \times m_g$, where m_g is the number of SNPs (210,000) (Baig et al., 2025). Lastly, the pedigree relationship matrix $\mathbf{A}$ was calculated with the R package AGHmatrix (Amadeu et al., 2023) with the function for the additive pedigree relationship matrix for autotetraploids according to Kerr et al. (2012). For this, missing and partially missing parent information were included in the pedigree data. The pedigree information ranged back one generation for five clones, two generations for 162 clones, three generations for 160 clones, four generations for 42 clones, and five generations for 15 clones. Each of the above-described relationship matrices was centered and scaled to a sample variance of 1 (Kang et al., 2010) to ensure comparability between the three different relationship sources.

To investigate the usefulness of borrowing relationship information in cases with low data quality, we performed computer simulations for which we considered the traits starch content and yield, and added a normally distributed error to the phenotypic observations such that lower per plot heritabilities were reached. The standard error required for these simulations was derived by solving equation (2) for σ_e.

In addition to relationship matrices derived from one source of information, we also investigated EBVs capitalizing on spectral, genomic, and pedigree relationships simultaneously. These mixed relationship matrices $\mathbf{C}_{\mathbf{max}}$ were calculated as follows:

$$\boldsymbol{C}_{\mathbf{max}} = \boldsymbol{M} \times w_M + \boldsymbol{G} \times w_G + \boldsymbol{A} \times w_A, \tag{5}$$

where w_M, w_G, and w_A are weights between 0 and 1 and $w_M + w_G + w_A = 1$. The weights varied in steps of 0.01. In total, 5151 combinations were evaluated for each trait.

2.3.3 | Calculation of holistic breeding values

In breeding programs in which GS is applied, the possibility exists to refine the estimation of genetic values for tested A clones further by combining the above-explained EBVs with the genomic best linear unbiased predictions (GBLUPs) from GS based on an existing prediction model. We refer to the resulting combination as holistic breeding value (HBV).

GBLUPs were calculated as follows:

$$Y_u = \mu + U_u + \epsilon_u, \tag{6}$$

where Y_u is the AEM of clone u and U is the random genetic effect of the uth clone. U is independent of ϵ_u with $U \sim N(0, \boldsymbol{G}\sigma_U^2)$. Here, σ_U^2 is the variance of the genetic effects and $\mathbf{G}$ is the additive genomic relationship matrix. We calculated GBLUPs under three different prediction scenarios by considering different TSs that resulted in high, medium, and low prediction accuracies. This way, HBVs can be compared under varying prerequisites for GP.

In the first and pessimistic approach, we used the geno- and phenotypes of the other breeders in the PotatoTools project, namely EUROPLANT and Norika, to predict the true A clone phenotypes of SaKa genotypes. For this, we used the AEMs across all 15 environments derived from model (3) of the EUROPLANT and Norika clones as the TS. In the second and optimistic approach, we additionally included all SaKa genotypes as well as all phenotypes of SaKa except of the true A clone environment in the TS. This was realized by using the AEMs calculated across every environment excluding the true A clone environment from model (3) as TS. Our third and realistic approach assumed that the true A clone phenotypes were available and incorporated them in the TS together with the data from EUROPLANT and Norika. Here, the TS consisted of the AEMs calculated across all EUROPLANT and Norika environments as well as the true A clone environment of SaKa.

Finally, HBVs were calculated by combining GBLUPs and EBVs in two ways: first, through a heritability index, and second, through a grid search. For the heritability index, we computed genomic heritabilities via shrinkage parameter λ analogous to Hofheinz et al. (2012) and Michel et al. (2017) with:

$$h_{\lambda}^{2} = \frac{1}{\sigma_e^2/\sigma_g^2 + 1}, \tag{7}$$

where σ_g^2 and σ_e^2 were taken from model (4) or model (6) to compute the heritabilities for the EBVs ($h^2_{\lambda\text{EBV}}$) or the GBLUPs ($h^2_{\lambda\text{GBLUP}}$), respectively. The HBVs were then calculated for each trait as follows:

$$\text{HBV} = \text{EBV}' \times h^2_{\lambda\text{EBV}} + \text{GBLUP}' \times h^2_{\lambda\text{GBLUP}}, \tag{8}$$

where EBV′ and GBLUP′ are the scaled vectors of EBVs and GBLUPs.

For the grid search, we modified equation (8) by substituting $h_{\lambda\text{EBV2}}$ and $h^2_{\lambda\text{GBLUP}}$ with the variable weights w_{EBV} and w_{GBLUP}. These were values between 0 and 1 and $w_{\text{EBV}} + w_{\text{GBLUP}} = 1$. The weights varied in steps of 0.001 resulting in 1001 weight combinations. We conducted a grid search for each combination of (i) trait, (ii) GBLUP prediction scenario (optimistic, pessimistic, and realistic), and (iii) method of calculating EBVs (**G**, **A**, **M**, C_{max}). The accuracy of HBVs and EBVs was assessed via correlation with the TVs.

3 | RESULTS

3.1 | Enhancing breeding values

We started calculating breeding values of the clones in the true A clone environment using the identity matrix (**I**), that is, without exploiting information from related individuals. The accuracies largely depended on the trait and ranged from 0.190 (shape short axis) to 0.870 (shape long axis) (Table 3). These accuracies were strongly correlated (0.90) with the per plot heritabilities of clones in the true A clone environment. In the next step, we calculated EBVs by using the multispectral (**M**), genetic (**G**), or pedigree (**A**) relationship matrix as the genetic variance-covariance matrix. Compared to **I**, a higher accuracy was achieved for seven traits with **M**, 12 traits with **G**, and 11 traits with **A**. A closer look at these relationship matrices revealed that **G** and **A** had similar structure ($\text{cor}(\boldsymbol{G}, \boldsymbol{A}) = 0.74$), while **M** showed highly different patterns of similarities ($\text{cor}(\boldsymbol{M}, \boldsymbol{G}) = 0.10$; $\text{cor}(\boldsymbol{M}, \boldsymbol{A}) = 0.11$) (Figure 1).

The mixed relationship matrix $\mathbf{C_{max}}$ which was determined via grid search for each trait achieved higher accuracy for nine traits compared to **I**, **M**, **G**, or **A** alone. All of these were achieved via a two-way combination of relationship matrices, three of which were through a combination of **G** and **A**, although these were highly correlated (Figure 1). Interestingly, $\mathbf{C_{max}}$ never combined all three available relationship matrices. However, the accuracy of breeding values for five traits, namely, emergence, foliage development, shape long axis, shape short axis, and skin type, was the same between $\mathbf{C_{max}}$ and one of the individual matrices (Table 3).

For each trait and relationship matrix, we calculated the increase in accuracy compared to the accuracy achieved with **I** (Table 3), which were denoted as $\delta_{\boldsymbol{M}}$, $\delta_{\boldsymbol{G}}$, $\delta_{\boldsymbol{A}}$, and $\delta_{\boldsymbol{C}_{\text{max}}}$, respectively. The highest mean increase across all traits was achieved with EBVs calculated with mixed relationship matrices ($\delta_{\boldsymbol{C}_{\text{max}}} = +6.2\%$), followed by pedigree ($\delta_{\boldsymbol{A}} = +3.2\%$), genomic ($\delta_{\boldsymbol{G}} = +2.4\%$), and multispectral ($\delta_{\boldsymbol{M}} = +1.2\%$) relationship matrices. Correlating the δ of every trait for each relationship matrix with the per plot heritabilities of the traits resulted in low to very low correlations for the pedigree ($\text{cor}(\delta_{\boldsymbol{A}}, h^2_{\text{plot}}) = -0.06$), genomic ($\text{cor}(\delta_{\boldsymbol{G}}, h^2_{\text{plot}}) = 0.01$), and multispectral ($\text{cor}(\delta_{\boldsymbol{M}}, h^2_{\text{plot}}) = 0.20$) relationship matrix and a moderately low negative correlation with the mixed relationship matrix ($\text{cor}(\delta_{\boldsymbol{C}_{\text{max}}}, h^2_{\text{plot}}) = -0.31$). However, after filtering out the disease traits rhizoctonia and scab symptoms, which were assessed in environments without artificial disease inoculation and, thus, have to be interpreted with care, the correlations changed to −0.52, −0.15, 0.07, and −0.58, respectively, showcasing that $\mathbf{C_{max}}$, **A** and, to a lesser extent, **G** were able to improve the accuracy of breeding values of less heritable traits in our case.

3.2 | Holistic breeding values

In the next step, EBVs were combined with GBLUPs estimated through GP to form HBVs. We started by combining both values through a heritability index and considered three different scenarios (optimistic, realistic, and pessimistic) to calculate GBLUPs for this purpose. The median accuracies of GBLUP across all traits were 0.972 for the optimistic, 0.615 for the realistic, and 0.218 for the pessimistic prediction

TABLE 3 Broad-sense heritability on a per plot basis (H^2_{plot}) and accuracy of breeding values for each trait. Breeding values were either calculated using the identity matrix (**I**) or by capitalizing on different relationship matrices resulting in enhanced breeding values (EBVs). EBVs were calculated for each trait with multispectral (**M**), genomic (**G**), pedigree (**A**) relationship matrices, or via a weighting of all three that resulted in a combined relationship matrix $\mathbf{C}_{\text{max}}$ which yielded the highest accuracy. Accuracy is defined as the correlation of the respective breeding values to the true values. The latter is defined as the adjusted means across SaKa 2020 and 2021 environments. w_M, w_G, and w_A refer to the weights of the multispectral, genetic, and pedigree relationship matrix used for $\mathbf{C}_{\text{max}}$.

		Relationship matrix					**Weight for C_{max}**		
Trait	H^2_{plot}	I	M	G	A	C_{max}	w_M	w_G	w_A
Emergence	0.51	0.625	0.641	0.629	0.627	0.641	100	0	0
Eye depth	0.43	0.701	0.701	0.721	0.725	0.725	0	8	92
Foliage development	0.45	0.626	0.644	0.627	0.626	0.644	100	0	0
General impression	0.33	0.413	0.404	0.440	0.452	0.454	9	0	91
Maturity	0.60	0.601	0.642	0.607	0.609	0.642	97	0	3
Polyphenol oxidase activity	0.57	0.681	0.677	0.689	0.690	0.690	0	24	76
Rhizoctonia symptoms	0.49	0.713	0.712	0.702	0.701	0.716	67	0	33
Scab symptoms	0.20	0.243	0.233	0.237	0.182	0.243	0	0	0
Shape long axis	0.78	0.870	0.869	0.887	0.884	0.887	0	100	0
Shape short axis	0.12	0.190	0.200	0.204	0.256	0.256	0	0	100
Skin type	0.47	0.749	0.751	0.755	0.757	0.757	0	0	100
Starch content	0.80	0.822	0.827	0.838	0.844	0.844	2	0	98
Tuber size	0.42	0.495	0.494	0.549	0.563	0.564	0	9	91
Yield	0.50	0.583	0.611	0.594	0.595	0.612	53	0	47

scenario. We assessed for each trait and scenario, if any HBVs achieved a higher accuracy than the corresponding GBLUPs and EBVs. If this was the case, we referred to this trait as an HBV max trait and computed the increase in accuracy realized with the HBVs compared to either the GBLUP or EBV accuracy, depending on which was higher. We found five HBV max traits (emergence, foliage development, rhizoctonia symptoms, yield, and skin type) in the realistic prediction scenario (Figure 2). In both, the pessimistic and optimistic prediction scenarios, HBVs were consistently less accurate than the GBLUPs and EBVs. The gain in accuracy was marginal for HBVs calculated through a heritability index with a mean increase of 0.5% in the realistic scenario.

To discover the maximum potential of the accuracies of HBVs, we considered a grid search with varying weights for the EBVs and GBLUPs. In general, the GBLUPs were more accurate than the EBVs in the optimistic scenario and vice versa in the pessimistic scenario (Figure 3). The grid search revealed a higher number of HBV max traits than the heritability approach (Figure 3). In detail, the HBVs were more accurate for 12 traits in the optimistic, eight traits in the realistic, and 11 traits in the pessimistic scenario. However, the gain in accuracy compared to GBLUPs and EBVs was marginal in most cases. The mean increase across all traits was 0.6% in the optimistic, 0.5% in the realistic, and 3.8% in the pessimistic scenario (Figure 3). The highest increases were realized in the pessimistic scenario for eye depth (+8.3%), shape short axis (+20.5%), and polyphenol oxidase activity (+3.7%).

The optimal weight of the EBVs in the HBVs varied depending on the considered prediction scenario. EBVs were weighted less in the optimistic (mean weight across traits = 0.161), and more in the pessimistic prediction scenario (mean weight across traits = 0.940). The realistic prediction scenario, in which GBLUPs and EBVs were similarly accurate (Figure 3) had a mean optimal weight of 0.821 for the EBVs. If $\mathbf{C}_{\text{max}}$ was used to calculate EBVs for HBVs, the resulting HBV max accuracies were always the highest possible; however, the differences in accuracy of $\mathbf{C}_{\text{max}}$ and the other relationship matrices (**A**, **M**, **G**, **I**) were small: the mean accuracy across traits and prediction scenarios was 0.774 for $\mathbf{C}_{\text{max}}$, 0.768 for **A**, 0.760 for **M**, and 0.753 for **G**. Notably, the realistic prediction scenario showed the strongest differences in accuracy depending on the relationship matrix used for the calculation of the EBV. Here, $\mathbf{C}_{\text{max}}$ had the highest mean accuracy with 0.662 and **G** the lowest with 0.600.

4 | DISCUSSION

4.1 | Increasing the accuracies of breeding values in unreplicated trials

In this study, we exploited the relationship information assessed from four different sources (pedigree, multispectral, genomic, or a combination) to enhance the accuracy of potato clone breeding values from unreplicated trials. The resulting

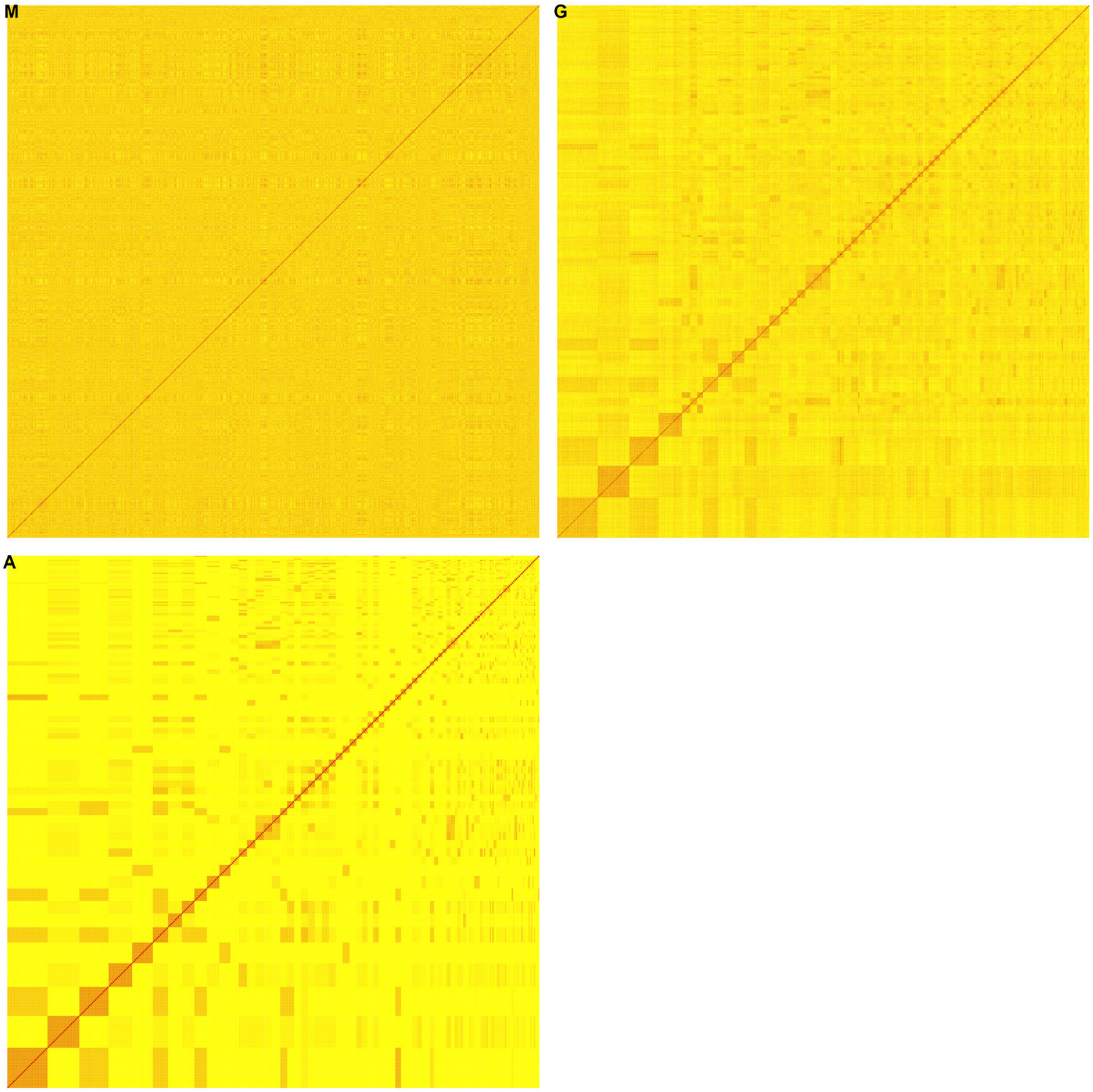

FIGURE 1 Heat maps of the three different relationship matrices used to estimate enhanced breeding values (top = multispectral [**M**], middle = genomic [**G**], bottom = pedigree [**A**]). Each heatmap is ordered according to the family structure of the germplasm whereby the clones are arranged on the *x* and *y* axes in same order (red = high similarity, yellow = low similarity).

EBVs were correlated with the TVs to estimate accuracy. However, the TV of a genotype is unknown in experimental data. Thus, we correlated the EBVs with the adjusted means that were calculated across the observations from the following years consisting of four environments in total. The upside of such an approach is that EBVs can be validated based on experimental data. The downside is that due to the genotype × environment interaction as well as experimental error, the adjusted means show a correlation <1 to the unknown TVs. However, we argue that the deviation from the true accuracies is small as analysis of the number of incorporated environments to calculate breeding values for the checks, which are present in all 15 environments, revealed very high accuracies even when only two environments were considered (Figure S2) as well as diminishing further gains in accuracy after three considered environments. Furthermore, we observed a very high correlation (0.98) between the AEMs across all five SaKa environments and the AEMs across all SaKa environments except the true A clone environment. Additionally, only a minimal gain in heritability from four to five testing environments without replicates has been reported in the context of the optimal design of METs for potato by Ortiz et al. (2023). Therefore, we are confident that our procedure to approximate the TV as the adjusted means across four SaKa environments is a sufficiently reliable proxy to assess the accuracy of breeding values.

The breeding values that were calculated for the true A clone environment based on the identity matrix (**I**) revealed accuracies >0.5 for 10 out of 14 traits (Table 3). As expected, these accuracies were highly correlated (0.9) with the per plot

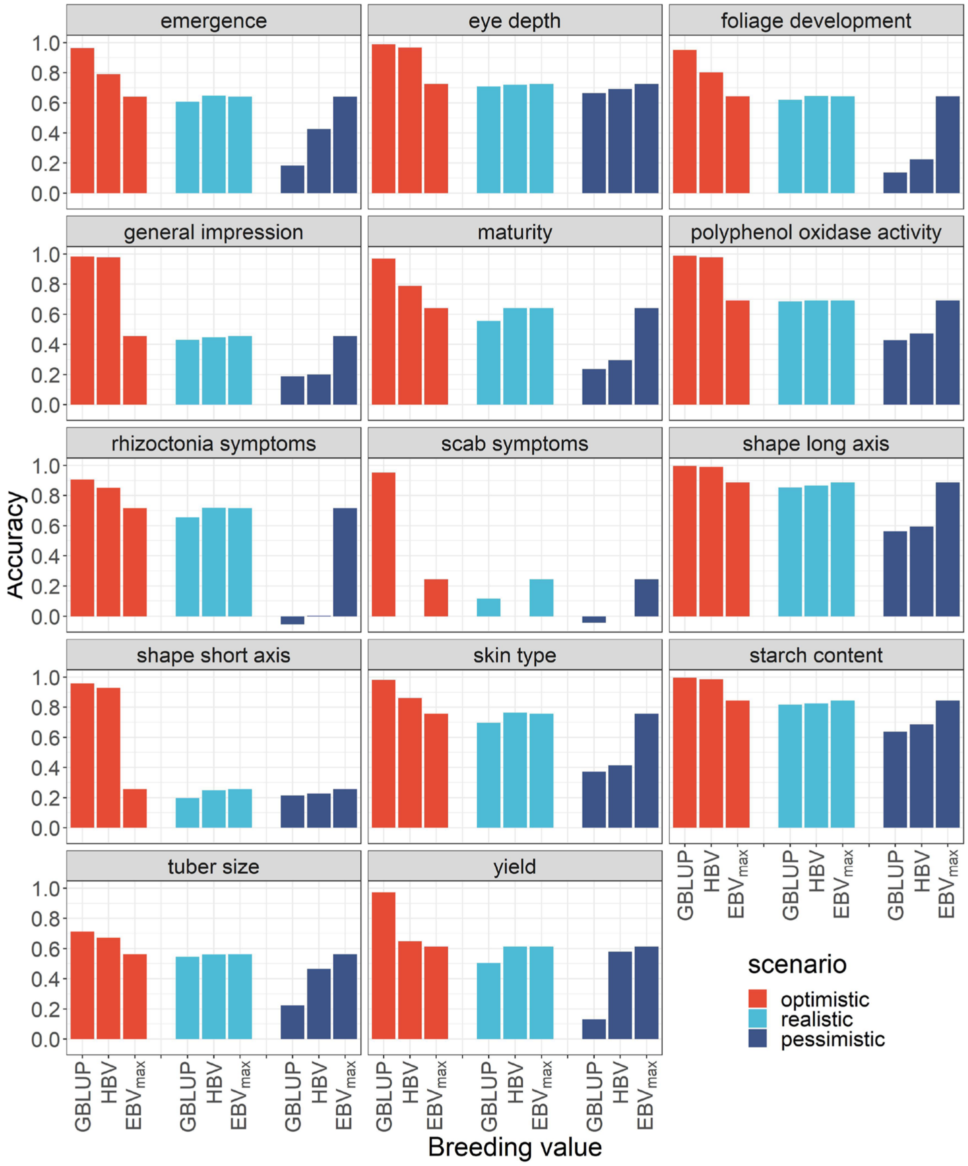

FIGURE 2 Accuracy of three different breeding values (genomic best linear unbiased prediction [GBLUP], holistic breeding value [HBV], enhanced breeding value [EBV$_{max}$]) of the SaKa clones in the true A clone environment. Accuracy was defined as the correlation of the breeding value to the true genetic value (TV), where the latter is defined as the adjusted entry mean across SaKa 2020 and 2021 environments. GBLUP were breeding values of SaKa clones estimated with genomic prediction. EBV$_{max}$ utilized a combination of spectral, genetic, and pedigree relationship matrices together with the phenotypic records of the clones in the true A clone environment to calculate breeding values that maximized accuracy per trait. HBV combined both breeding values via heritability index where the heritability was calculated via shrinkage parameter (λ). We considered three scenarios to calculate GBLUPs which results in three different HBVs. The optimistic approach utilized adjusted entry means across all Norika and EUROPLANT environments as well as SaKa 2020 and 2021 environments to train the GBLUP model. The realistic approach utilized adjusted entry means across all Norika and EUROPLANT environments, as well as the true A clone environment of SaKa. Adjusted means for the pessimistic approach were calculated only from all Norika and EUROPLANT environments.

heritabilities of the traits. The high accuracies with **I** illustrate that A clone breeding values are moderately to highly accurate even in a single environment without replicated entries or any exploitation of relationship information for most traits. One explanation for the high accuracies of the classical BLUP values is the high per plot heritabilities with values between 0.12 and 0.8 found in this study (Table 3). Compared to other widely grown crops, the per plot heritability of tuber yield of potato was particularly high (Laidig et al., 2008), which makes selection in unreplicated trials effective (Ticona-Benavente & Silva Filho, 2015). However, despite the already high accuracies that were observed for the classical BLUP calculated based on the **I** matrix, even higher accuracies were achieved for 12 out of 14 traits through EBVs, that is, when relationship

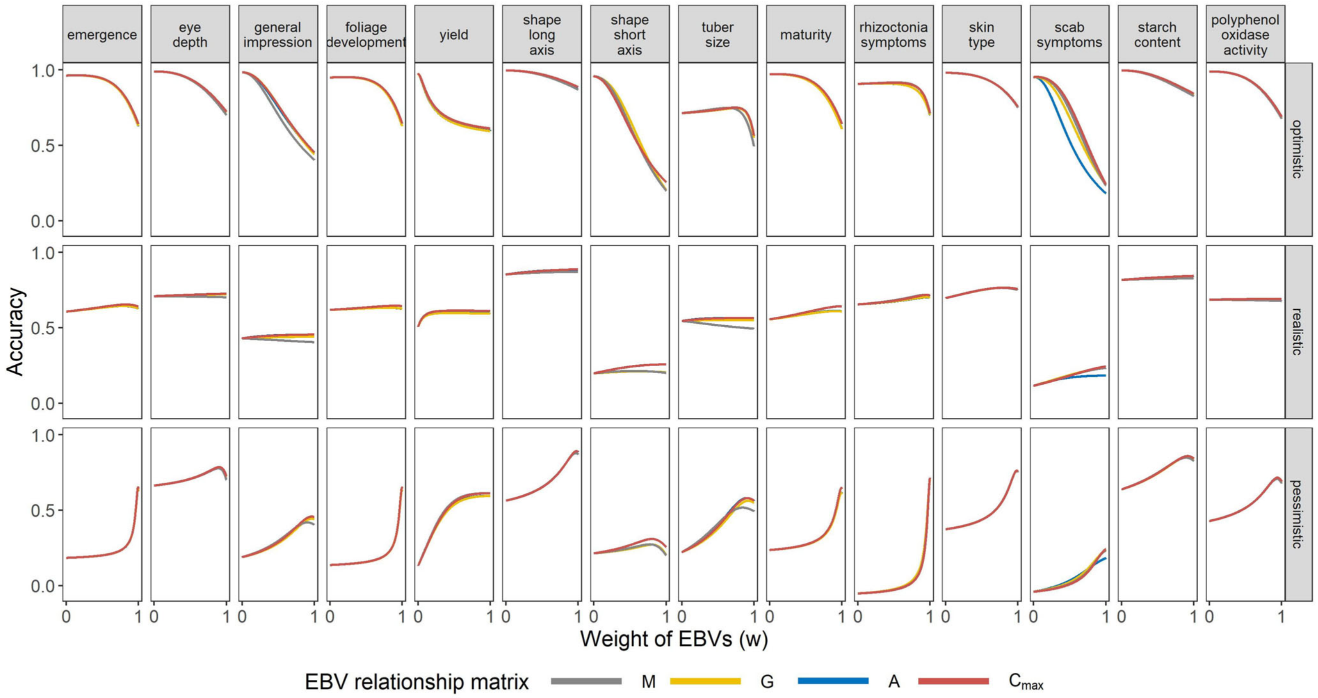

FIGURE 3 Accuracy of holistic breeding values (HBVs) depending on the weight of enhanced breeding values (EBVs). Accuracy is defined as the correlation of the breeding value to the true genetic value (TV), where the latter is defined as the adjusted mean across SaKa 2020 and 2021 environments. HBVs were computed for each trait as a weighted combination of EBVs(right-hand side of each panel) and breeding values calculated with genomic prediction (genomic best linear unbiased prediction [GBLUP])(left-hand side of each panel). EBVs were calculated for each trait with multispectral (**M**), genomic (**G**), pedigree (**A**) relationship matrices, or via a weighting of all three that resulted in a combined relationship matrix $\mathbf{C}_{\mathbf{max}}$ which yielded the highest accuracy. Three scenarios, were considered to calculate GBLUPs with genomic prediction while EBVs stayed the same across all scenarios. The optimistic approach utilized adjusted entry means across all Norika and EUROPLANT environments, as well as SaKa 2020 & 2021 environments to train the GBLUP model. The realistic approach utilized adjusted entry means across all Norika and EUROPLANT environments, as well as the true A clone environment of SaKa. Adjusted means for the pessimistic approach were calculated only from all Norika and all EUROPLANT environments.

information in the form of a genetic, multispectral, pedigree, or a combined relationship matrix is exploited in addition (Table 3). Notably, the before-mentioned increase in accuracy of EBVs versus BLUPs was higher for lowly heritable traits like general impression and shape short axis compared to traits of higher heritability (Table 3). The reason for this observation is that borrowing information from related clones to improve the accuracy of a BLUP is more important for lowly heritable traits compared to traits where a much higher proportion of the differences among the clones is due to genetics, that is, they are highly heritable. These findings illustrate the attractiveness of using EBVs for selection in clone breeding.

The increases in accuracies realized through EBVs compared to the BLUPs with the **I** matrix, however, were low on average (+6.2% with $\mathbf{C}_{\mathbf{max}}$) across all traits. This is probably a direct reflection of the high per plot heritabilities observed and, thus, the high quality of the true A clone environment phenotypic data. To test this hypothesis, we simulated EBVs of a lower per plot heritability, by adding a normally distributed error to the phenotypic observations (Table S3). We found that the lower the heritability was, the higher the accuracy of EBVs compared to the BLUPs, which were calculated based on **I** (Figure S3). Simply put, the less the quality of the phenotypic data was the more worthwhile it got to borrow relationship information to enhance accuracy. Since EBV accuracies were as high or higher than the breeding values calculated with **I** and since potato breeders can only indirectly infer the heritability of the trials from replicated checks, we recommend the use of EBVs for selection. This is because selection based on EBVs resembles a robust method with a high selection accuracy within the potential of the respective trait.

Depending on the trait, the different types of relationship information lead to different increases in accuracy compared to the classical BLUP (Table 3). In general, the EBV accuracies achieved with **G** and **A** were similar with mean accuracies across all traits of 0.606 with **G** and 0.608 with **A**. This can partly be explained by the high correlation (0.74) of these two relationship matrices (Figure 1). The similar accuracies between **G** and **A** were surprising considering that higher accuracies are frequently reported with **G** compared to **A** for diploids (Hayes et al., 2009; Velazco et al., 2019) as well as autotetraploids (de Bem Oliveira et al., 2019). This is because **G** describes observed rather than expected relationships, thus, accounting for Mendelian sampling. We speculate that the reason for the nonsignificant difference in performance of **A** compared to **G** in this study was due to the consideration of genetic material related by pedigree, which did not yet undergo strong selection. Given the similar accuracies of **A** and **G** found in this study, the selection based on EBVs derived from **A** instead of **G** seems more attractive for potato breeders because of the attached costs to derive **G**. Only two traits, namely shape short axis and scab symptoms deviated from that trend in that one relationship matrix achieved a significantly higher accuracy (>0.05) than the other. The superiority of **G** versus **A** observed for the accuracy of scab symptoms might be explained by the trait's likely oligogenic nature that is well explained by a only few markers (Yuan et al., 2020). However, for the opposite trend observed for shape short axis, we have no plausible explanation.

While **A** and **G** had the highest positive impact on the accuracy of EBVs for eye depth, general impression, shape, tuber size, and starch content, the multispectral relationship matrix **M** revealed the highest accuracies for yield and plant developmental traits such as emergence, foliage development, and maturity (Table 3). This is in accordance with results of Maggiorelli et al. (2024) where the performance of spectral profiles was compared with SNP markers in the context of a prediction scenario and not a scenario to improve the accuracy of phenotyping as was examined in the current study. The benefit of using spectral information to predict potato yield was also recently shown by Yusuf et al. (2025). Both studies also underlined the large increase of the predictive ability of potato yield if both, genomic and phenomic data, are capitalized simultaneously. However, in contrast to Maggiorelli et al. (2024) and Yusuf et al. (2025), a combination of relationship matrices ($\mathbf{C}_{\mathbf{max}}$) did not lead to a significant increase in accuracy for yield compared to the highest performing single relationship matrix in our study (Table 3). The reason for this could be the aforementioned high per plot heritability of the true A clone environment.

Comparing the similarity patterns of the relationship matrices used to calculate EBVs (Figure 1) revealed large differences between **A** or **G**, and **M**. While **A** and **G** reflect an expected and observed genetic relationship, respectively, **M** is representing a relationship based on the composition of endophenotypes in the tissue captured by the camera at the timepoint of measurement. As such, it is not surprising that developmental traits like maturity, emergence, and foliage development were better predicted with **M** compared to **G** or **A**, since the differences between the clones can be observed by eye and, thus, were captured in the multispectral data. Several studies analyzing phenomic prediction (Galán et al., 2020; Krause et al., 2019; Maggiorelli et al., 2024; Rincent et al., 2018; Yusuf et al., 2025) showed that these spectral patterns do capture genetic variance to varying degrees and, thus, can be used to infer genetic similarities even though one can argue that relationship information between genotypes captured in spectral relationship matrices is not based purely on genetic relationship. In a sense, **M** adds a further layer of information and a snapshot of the present state of the plant compared to **G** or **A**, since **M** entails genetic similarities between the clones after gene expression. Therefore, this extra layer of information should especially aid in the prediction of complex traits like shown in this study for yield which was estimated more accurately with **M** rather than **G** or **A** (Table 3). This trend was also reported by Zhu et al. (2022) who showed for triticale that NIRS-based phenomic prediction outperformed GP for traits with a complex genetic architecture. In contrast, traits like eye depth and tuber size were better estimated using **G** or **A** compared to **M**. We hypothesize that this is because the expression of these traits is not well reflected in the endophenotypic composition of the above-ground canopy because these genetically less complex traits are mainly driven by a few loci.

A possible hindrance in terms of EBV accuracy with **M** was that the spectral profiles in our study were not assessed in the true A clone environment itself but due to budget restrictions in the later years (Windeby 2020 and 2021). The literature about phenomic prediction (Krause et al., 2019; Rincent et al., 2018) reports that while predictions across environments are feasible they can be of lower accuracy compared to within environment predictions. This reduction of prediction accuracy seems to be particularly strong for potato (Maggiorelli et al., 2024). However, these findings were reported in the context of predictions. To test if using spectral profiles from the environment of consideration improves breeding value accuracy in cases where phenotypic data are available, we calculated EBVs for Windeby 2020 (Table S4) where multispectral profiles were directly measured. In this case, accuracy was defined as the correlation with the adjusted means across the other four environments including the true A clone environment. Although the accuracies were higher for all relationship matrices in general (given in Table 3 and Table S4), including **I**, the difference in accuracy between the use of **I** and **M** was similar compared to the same difference for the true A clone environment as described above. From this observation, we conclude that the accuracy reported in our study for the **M** approach is not underestimated by the use of spectral data from an environment other than the true A clone environment. In general, the accuracies observed for the breeding values calculated based on the spectral data **M** were only slightly lower across all traits than those of **G** and **A**. Therefore, **M** represents a good alternative, especially if such information can be assessed cheap and on a large number of clones, for example, using drones.

4.2 | Combining EBVs and GBLUPs

GP is becoming the state-of-the-art method of predictive breeding in many crops (Wartha & Lorenz, 2021). Promising results for the utilization of GP have also been reported for potato breeding programs (Slater et al., 2016; Stich & Van Inghelandt, 2018; Wu et al., 2023). In breeding programs, where GP is already applied, the breeders accumulate genotypic data over time, which helps to design a suitable TS for predictions as well as to adjust the used relationship matrix. For wheat, Michel et al. (2017) showed that breeders can capitalize on this accumulated genomic data not only via GP but also by combining GBLUPs with EBVs to further increase the accuracy of tested genotypes. In this way, relationship information of the genotypes of interest as well as of related genotypes is utilized. Similar to Michel et al. (2017), we combined EBVs and GBLUPs into HBVs via a heritability index but also explored the full potential of this approach for potato via a grid search of the weights of both breeding values.

As GBLUP accuracies benefit from closely related entries in training and test set (Habier et al., 2010; Hickey et al., 2014) as well as frequently updated reference populations (Podlich et al., 2004), a good solution to train the model and calculate GBLUPs for this purpose in practical breeding programs would be to use reference material from earlier cycles. However, since no pheno- and genotyped material from earlier breeding cycles was available for our study from SaKa, we calculated GBLUPs using the information from Norika and EUROPLANT as references. This adds genetically unrelated individuals in the training population which has been shown to reduce the prediction accuracy (Clark et al., 2012; Lorenz & Smith, 2015). Therefore, we classified this scenario as being pessimistic. As the prediction accuracies achieved by this scenario are expected to be lower than what can be realized with a TS of a reasonable size in potato (Stich & Van Inghelandt, 2018), we also considered two more GBLUP scenarios to investigate if accuracies can be increased using HBVs based on different starting conditions for GP. The optimistic scenario trains the model based on all environments except the true A clone environment itself which includes the other four SaKa environments based on the same clones. As the difference in AEMs across all five SaKa environments and across all but the true A clone environment is marginal, the model is close to predicting its own TS. This is also reflected by the very high GBLUP accuracies for most traits in this prediction scenario (Figures 2 and 3). Lastly, the realistic scenario, where the true A clone environment clones are used in the TS, is expected to be similar to a practical data situation except that EUROPLANT and Norika material instead of additional SaKa material were included.

For most traits, we observed no improvement in accuracy for HBVs compared to GBLUPs or EBVs when using a heritability index independent of which GBLUP scenario was considered. Furthermore, for traits where a higher accuracy was achieved with HBVs compared to EBVs or GBLUPs, the increase was small (0.08% –2.3 %)(Figure 2). These findings are in contrast to what Michel et al. (2017) reported for yield and protein content in wheat where the same heritability index led to an increase of 18% compared to KBLUP which is similar to EBVs in our study. The explanation for this discrepancy could be that the accuracy of phenotypic selection on grain yield reported by Michel et al. (2017), was considerably lower (0.21) compared to what we observed for potato tuber yield (0.58)(Table 3). This finding is in agreement with results of Laidig et al. (2008) who reported lower per plot heritabilities for wheat than for potato. This difference in the ability to assess genetic values in unreplicated trials accurately likely explains the discrepancies of accuracy increases realized by additionally considering relationship data in the estimation of HBVs.

For none of the examined traits, we observed that the **G** matrix used to calculate the EBV was selected to calculate the HBVs with the highest accuracy. Instead, compared across all prediction scenarios, we found that **M** was most beneficial for HBV accuracies for emergence, foliage development, yield, maturity, rhizoctonia, and scab symptoms while **A** was more beneficial for HBV accuracies for eye depth, general impression, shape long axis, shape short axis, tuber size, polyphenol oxidase activity, skin type, and starch content. Therefore, potato breeders could rely solely on collecting spectral profiles and considering pedigree information to enhance selection based on HBVs for the traits analyzed in this study. Furthermore, for HBVs, we found no significant increase in accuracy if $\mathbf{C}_{\mathbf{max}}$ was used to derive EBVs compared to the best single relationship matrix, thus, simplifying the process to derive EBVs, as only specific relationship information is needed for specific traits.

While the heritability index to weight EBVs versus GBLUPs resulted for very few traits in improvements for HBVs versus EBVs/GBLUPs (Figure 2), the grid search revealed many more combinations of trait and prediction scenarios in which accuracies increased by combining GBLUPs and EBVs (Figure 3). Notably, even in the pessimistic prediction scenario, accuracies were improved by combining both breeding values for 11 out of 14 traits which shows that it is principally possible to even use strictly unrelated breeding material to derive GBLUPs which are then combined with BLUPs to HBVs with elevated accuracy compared to GBLUPs and EBVs alone. However, our grid search did not reveal a clear pattern on how to arrive at the individual weights to achieve the HBVs such that they are higher than EBVs and GBLUPs. Therefore, despite the promising results, this approach warrants further research before it can be used in practical potato breeding programs.

4.3 | Perspectives on integrating predictive breeding approaches in clone breeding

Our study showed that the accuracies of breeding values derived for A clones from unreplicated trials can be enhanced by exploiting relationship information in different forms even if the collected phenotypic values are highly accurate (Table 3). Thus, this approach can be directly implemented in breeding programs. In addition, we also observed that (i) the increase in accuracy correlated negatively with the per plot heritabilities (−0.58), and (ii) spectral profiles and pedigree information, which are cost-effective to obtain compared to genotype information, were sufficient to enhance accuracy for most traits. Therefore, selection based on EBVs calculated with **M** and **A** is an effective method to ensure high accuracies for unreplicated clones, especially if heritability is low. As such, we speculate that integrating relationship information into breeding values estimation becomes more important for situations where the number of plants per plot is low. A logical conclusion could therefore be that selection based on EBVs enables the possibility to reduce the number of tested tubers per clone in the A clone stage. However, in the specific case of potato breeding this approach is not always practical as the low propagation coefficient of potato results in the necessity to plant every available tuber for multiplication in the early stages.

Following that logic, the more favorable situation for the implementation of selection based on EBVs lies prior to the A clone stage. For potato breeding that would be the seedling and single hill stage where a very large number of clones is considered with only one plant each. Consequently, traits measured in these stages are characterized by lower heritability than in subsequent stages which is why potato breeders usually do not put high selection pressure on assessed traits at such an early timepoint. Borrowing relationship information in the form of spectral profiles or pedigree data could be a highly effective approach to counteract the low heritability of clones in the seedling or single hill stage. However, future research is warranted to unveil the true efficacy of selection based on EBVs at these early timepoints in the potato breeding program.

Although the HBV grid search revealed the potential to achieve higher breeding value accuracies compared to the ones achieved with EBVs, the realized increases in accuracy were marginal (Figure 3). Therefore, selection based on HBVs at the A clone stage can increase selection accuracy but might only be worthwhile in situations where the necessary data (i.e., relationship information from earlier cycles and relationship information as well as phenotypic records of the selection candidates) is already available. Analogous to the EBVs, another interesting point of application can be to select based on HBVs in the single hill or seedling stage for the same reasons as mentioned earlier. However, the practical application of this idea for breeders remains a challenge. A combination of EBVs and GBLUPs through heritability index offers the advantage of a clear weight definition. Our grid search did not reveal a clear path on how to arrive at the individual weights to maximize HBVs. If these weight configurations remain stable across traits and breeders then they could be reused and applied repeatedly, a scenario that warrants further research.

5 | CONCLUSIONS

We integrated genomic, multispectral, and pedigree based relationship information into the estimation of breeding values of unreplicated potato A clones and analyzed the effect on the accuracy of these values. We found that the integration of relationship information increases accuracy for 12 out of 14 traits whereby the margin in accuracy increase was negatively correlated with the heritability of the trait. However, the mean increase in accuracy across all traits was modest ranging from 1.1% to 4.4% compared to the accuracies achieved without incorporation of relationship information. Furthermore, we combined EBVs and GBLUPs through a grid search of the respective weights to form HBVs, which further increased accuracies ranging from 0.5% to 3.8%, depending on the TS for GBLUP. However, these increases in accuracy with HBVs were achieved through specific weight configurations. A heritability index to combine both breeding values was not successful. We found that the different relationship matrices were advantageous for different traits but in general, the multispectral profiles were particularly useful for predicting yield and plant development traits. Combining all three types of relationship information led to maximum accuracies, yet, these were not significantly higher than the best-achieving single relationship matrix. Our findings suggest a promising method to improve breeding value accuracies of unreplicated clones based on easy-to-collect multispectral and pedigree data. We see the most potential for application in early potato breeding program stages, like the seedling or single hill stage, where heritability of the collected phenotypic information is typically low.

AUTHOR CONTRIBUTIONS

Alessio Maggiorelli: Conceptualization; formal analysis; writing—original draft. **Kathrin Thelen**: Data curation; formal analysis; writing—review and editing. **Vanessa Prigge**: Data curation; funding acquisition; resources; writing—review and editing. **Julien Bruckmüller**: Data curation; funding acquisition; resources; writing—review and editing. **Katja Muders**: Data curation; funding acquisition; resources; writing—review and editing. **Stefanie Hartje**: Data curation; funding acquisition; resources; writing—review and editing. **Benjamin Stich**: Conceptualization; funding acquisition;

project administration; supervision; writing—original draft.

ACKNOWLEDGMENTS

We thank geo-konzept GmbH for the collection, curation, and processing of the multispectral data. The funders did not influence the study design; the collection, analysis, and interpretation of data; the writing of the manuscript; and the decision to submit the manuscript for publication.

Open access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST STATEMENT

Vanessa Prigge and Julien Bruckmüller are employed by SaKa Pflanzenzucht GmbH & Co. KG. Katja Muders is employed by Nordring- Kartoffelzucht- und Vermehrungs-GmbH (Norika). Stefanie Hartje is employed by EURO-PLANT Innovation GmbH & Co. The authors have no other relevant financial or non-financial interests to disclose.

DATA AVAILABILITY STATEMENT

The original datasets generated and/or analyzed in the current study are not publicly available due to the material being part of the company secret of the involved breeding companies. However, the data are available in encoded form from the corresponding author upon reasonable request. R scripts for data analysis are available on Github at https://github.com/AlessioMR/P2_scripts.

ORCID

Alessio Maggiorelli https://orcid.org/0009-0000-8764-2572
Julien Bruckmüller https://orcid.org/0000-0002-3900-9833
Benjamin Stich https://orcid.org/0000-0001-6791-8068

REFERENCES

Amadeu, R. R., Garcia, A. A. F., Munoz, P. R., & Ferrão, L. F. V. (2023). Aghmatrix: Genetic relationship matrices in R. *Bioinformatics*, *39*(7), btad445.

Baig, N., Thelen, K., Ayenan, M. A., Hartje, S., Obeng-Hinneh, E., Zgadzaj, R., Renner, J., Muders, K., Truberg, B., Rosen, A., Prigge, V., Bruckmueller, J., Luebeck, J., Van Inghelandt, D., & Stich, B. (2025). An axiom SNP genotyping array for potato: Development, evaluation and applications. *bioRxiv*. https://doi.org/10.1101/2025.08.17.670748

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Bundessortenamt. (2019). *Wertungsprüfrichtlinie kartoffel, überarbeitete richtlinie kartoffel*. https://www.bundessortenamt.de/bsa/media/Files/RILI_4_3_Kartoffel_201902.pdf

Clark, S. A., Hickey, J. M., Daetwyler, H. D., & van der Werf, J. H. (2012). The importance of information on relatives for the prediction of genomic breeding values and the implications for the makeup of reference data sets in livestock breeding schemes. *Genetics Selection Evolution*, *44*, 1–9.

Crainiceanu, C. M., & Ruppert, D. (2004). Likelihood ratio tests in linear mixed models with one variance component. *Journal of the Royal Statistical Society: Series B*, *66*(1), 165–185.

Daetwyler, H. D., Bansal, U. K., Bariana, H. S., Hayden, M. J., & Hayes, B. J. (2014). Genomic prediction for rust resistance in diverse wheat landraces. *Theoretical and Applied Genetics*, *127*, 1795–1803.

deBem Oliveira, I., Resende, M. F., Jr Ferrão, L. F. V., Amadeu, R. R., Endelman, J. B., Kirst, M., Coelho, A. S. G., & Munoz, P. R. (2019). Genomic prediction of autotetraploids; influence of relationship matrices, allele dosage, and continuous genotyping calls in phenotype prediction. *G3: Genes, Genomes, Genetics*, *9*(4), 1189–1198.

DeLacy, I., Basford, K., Cooper, M., Bull, J., & McLaren, C. (1996). Analysis of multi-environment trials—An historical perspective. *Plant Adaptation and Crop Improvement*, *39124*, 39–124.

Dos Santos, J. P. R., Pires, L. P. M., de Castro Vasconcellos, R. C., Pereira, G. S., Von Pinho, R. G., & Balestre, M. (2016). Genomic selection to resistance to *Stenocarpella maydis* in maize lines using dartseq markers. *BMC Genetics*, *17*, 1–10.

Enciso-Rodriguez, F., Douches, D., Lopez-Cruz, M., Coombs, J., & deLos Campos, G. (2018). Genomic selection for late blight and common scab resistance in tetraploid potato (*Solanum tuberosum*). *G3: Genes, Genomes, Genetics*, *8*(7), 2471–2481.

Endelman, J. B., Atlin, G. N., Beyene, Y., Semagn, K., Zhang, X., Sorrells, M. E., & Jannink, J.-L. (2014). Optimal design of preliminary yield trials with genome-wide markers. *Crop Science*, *54*(1), 48–59.

Galán, R. J., Bernal-Vasquez, A.-M., Jebsen, C., Piepho, H.-P., Thorwarth, P., Steffan, P., Gordillo, A., & Miedaner, T. (2020). Integration of genotypic, hyperspectral, and phenotypic data to improve biomass yield prediction in hybrid rye. *Theoretical and Applied Genetics*, *133*(11), 3001–3015.

García-Ruiz, A., Cole, J. B., VanRaden, P. M., Wiggans, G. R., Ruiz-López, F. J., & Van Tassell, C. P. (2016). Changes in genetic selection differentials and generation intervals in US Holstein dairy cattle as a result of genomic selection. *Proceedings of the National Academy of Sciences*, *113*(28), E3995–E4004.

Habier, D., Tetens, J., Seefried, F.-R., Lichtner, P., & Thaller, G. (2010). The impact of genetic relationship information on genomic breeding values in German Holstein cattle. *Genetics Selection Evolution*, *42*, 1–12.

Habyarimana, E., Parisi, B., & Mandolino, G. (2017). Genomic prediction for yields, processing and nutritional quality traits in cultivated potato (*Solanum tuberosum* L.). *Plant Breeding*, *136*(2), 245–252.

Hayes, B. J., & Goddard, M. (2010). Genome-wide association and genomic selection in animal breeding. *Genome*, *53*(11), 876–883.

Hayes, B. J., Visscher, P. M., & Goddard, M. E. (2009). Increased accuracy of artificial selection by using the realized relationship matrix. *Genetics Research*, *91*(1), 47–60.

He, S., Schulthess, A. W., Mirdita, V., Zhao, Y., Korzun, V., Bothe, R., Ebmeyer, E., Reif, J. C., & Jiang, Y. (2016). Genomic selection in a commercial winter wheat population. *Theoretical and Applied Genetics*, *129*, 641–651.

Hickey, J. M., Dreisigacker, S., Crossa, J., Hearne, S., Babu, R., Prasanna, B. M., Grondona, M., Zambelli, A., Windhausen, V. S., Mathews, K., & Gorjanc, G. (2014). Evaluation of genomic selection training population designs and genotyping strategies in plant breeding programs using simulation. *Crop Science*, *54*(4), 1476–1488.

Hofheinz, N., Borchardt, D., Weissleder, K., & Frisch, M. (2012). Genome-based prediction of test cross performance in two subsequent breeding cycles. *Theoretical and Applied Genetics*, *125*, 1639–1645.

Jansky, S. H., & Spooner, D. M. (2018). The evolution of potato breeding. *Plant Breeding Reviews*, *41*, 169–214.

Juliana, P., Singh, R. P., Braun, H.-J., Huerta-Espino, J., Crespo-Herrera, L., Govindan, V., Mondal, S., Poland, J., & Shrestha, S. (2020). Genomic selection for grain yield in the CIMMYT wheat breeding program-status and perspectives. *Frontiers in Plant Science*, *11*, 564183.

Kang, H. M., Sul, J. H., Service, S. K., Zaitlen, N. A., Kong, S.-Y, Freimer, N. B., Sabatti, C., & Eskin, E. (2010). Variance component model to account for sample structure in genome-wide association studies. *Nature Genetics*, *42*(4), 348–354.

Kerr, R. J., Li, L., Tier, B., Dutkowski, G. W., & McRae, T. A. (2012). Use of the numerator relationship matrix in genetic analysis of autopolyploid species. *Theoretical and Applied Genetics*, *124*(7), 1271–1282.

Knol, E. F., Nielsen, B., & Knap, P. W. (2016). Genomic selection in commercial pig breeding. *Animal Frontiers*, *6*(1), 15–22.

Krause, M. R., González-Pérez, L., Crossa, J., Pérez-Rodríguez, P., Montesinos-López, O., Singh, R. P., Dreisigacker, S., Poland, J., Rutkoski, J., Sorrells, M., Gore, M. A., & Mondal, S. (2019). Hyperspectral reflectance-derived relationship matrices for genomic prediction of grain yield in wheat. *G3: Genes, Genomes, Genetics*, *9*(4), 1231–1247.

Krause, M. R., Mondal, S., Crossa, J., Singh, R. P., Pinto, F., Haghighattalab, A., Shrestha, S., Rutkoski, J., Gore, M. A., Sorrells, M. E., & Poland, J. (2020). Aerial high-throughput phenotyping enables indirect selection for grain yield at the early generation, seed-limited stages in breeding programs. *Crop Science*, *60*(6), 3096–3114.

Laidig, F., Drobek, T., & Meyer, U. (2008). Genotypic and environmental variability of yield for cultivars from 30 different crops in German official variety trials. *Plant Breeding*, *127*(6), 541–547.

Legarra, A., Robert-Granié, C., Manfredi, E., & Elsen, J.-M. (2008). Performance of genomic selection in mice. *Genetics*, *180*(1), 611–618.

Lohithaswa, H. C., Shreekanth, S. M., Banakara, S. K., Sripathy, K., & Mallikarjuna, M. G. (2022). Genomic selection for enhanced stress tolerance in maize. In *Next-generation plant breeding approaches for stress resilience in cereal crops* (pp. 121–160). Springer.

Lorenz, A. J., & Smith, K. P. (2015). Adding genetically distant individuals to training populations reduces genomic prediction accuracy in barley. *Crop Science*, *55*(6), 2657–2667.

Maggiorelli, A., Baig, N., Prigge, V., Bruckmüller, J., & Stich, B. (2024). Using drone-retrieved multispectral data for phenomic selection in potato breeding. *Theoretical and Applied Genetics*, *137*(3), 70.

Matei, G., Woyann, L. G., Milioli, A. S., de Bem Oliveira, I., Zdziarski, A. D., Zanella, R., Coelho, A. S. G., Finatto, T., & Benin, G. (2018). Genomic selection in soybean: Accuracy and time gain in relation to phenotypic selection. *Molecular Breeding*, *38*, 1–13.

Meier, U. (2018). *Growth stages of mono- and dicotyledonous plants (BBCH Monograph)*. Open Agrar Repositorium. https://doi.org/10.5073/20180906-074619

Meuwissen, T. H., Hayes, B. J., & Goddard, M. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, *157*(4), 1819–1829.

Michel, S., Ametz, C., Gungor, H., Akgöl, B., Epure, D., Grausgruber, H., Löschenberger, F., & Buerstmayr, H. (2017). Genomic assisted selection for enhancing line breeding: Merging genomic and phenotypic selection in winter wheat breeding programs with preliminary yield trials. *Theoretical and Applied Genetics*, *130*, 363–376.

Ortiz, R., Reslow, F., Huicho, J., Vetukuri, R. R., & Crossa, J. (2023). Optimizing multi-environment testing in potato breeding: Using heritability estimates to determine number of replications, sites, and years for field trials. *Euphytica*, *219*(4), 48.

Podlich, D. W., Winkler, C. R., & Cooper, M. (2004). Mapping as you go: An effective approach for marker-assisted selection of complex traits. *Crop Science*, *44*(5), 1560–1571.

R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Rincent, R., Charpentier, J.-P., Faivre-Rampant, P., Paux, E., Le Gouis, J., Bastien, C., & Segura, V. (2018). Phenomic selection is a low-cost and high-throughput method based on indirect predictions: Proof of concept on wheat and poplar. *G3: Genes, Genomes, Genetics*, *8*(12), 3961–3972.

Slater, A. T., Cogan, N. O., Forster, J. W., Hayes, B. J., & Daetwyler, H. D. (2016). Improving genetic gain with genomic selection in autotetraploid potato. *The Plant Genome*, *9*(3), plantgenome2016.02.0021.

Stewart-Brown, B. B., Song, Q., Vaughn, J. N., & Li, Z. (2019). Genomic selection for yield and seed composition traits within an applied soybean breeding program. *G3: Genes, Genomes, Genetics*, *9*(7), 2253–2265.

Stich, B., & Van Inghelandt, D. (2018). Prospects and potential uses of genomic prediction of key performance traits in tetraploid potato. *Frontiers in Plant Science*, *9*, 159.

Sun, C., Feng, L., Zhang, Z., Ma, Y., & Crosby, T. O. (2020). Prediction of end-of-season tuber yield and tuber set in potatoes using in-season UAV-based hyperspectral imagery and machine learning. *Sensors*, *20*(18), 5293.

Sverrisdóttir, E., Sundmark, E. H. R., Johnsen, H. Ø., Kirk, H. G., Asp, T., Janss, L., Bryan, G., & Nielsen, K. L. (2018). The value of expanding the training population to improve genomic selection models in tetraploid potato. *Frontiers in Plant Science*, *9*, 1118.

Terraillon, J., Frisch, M., Falke, K. C., Jaiser, H., Spiller, M., Cselényi, L., Krumnacker, K., Boxberger, S., Habekuß, A., Kopahnke, D., Serfling, A., Ordon, F., & Zenke Philippi, C. (2022). Genomic prediction can provide precise estimates of the genotypic value of barley lines evaluated in unreplicated trials. *Frontiers in Plant Science*, *13*, 735256.

Thelen, K., Prigge, V., Kohlmorgen, A., Muders, K., Truberg, B., Hartje, S., Renner, J., & Stich, B. (2025). Variance and covariance components of agronomic and quality traits assessed in tetraploid potato and their implications on practical breeding. *Frontiers in Plant Science*, *15*, 1505193.

Ticona-Benavente, C. A., & Silva Filho, D. F. D. (2015). Comparison of BLUE and BLUP/REML in the selection of clones and families of potato (*Solanum tuberosum*). *Genetics and Molecular Research*, *14*, 18421–18430.

Tiemens-Hulscher, M., Delleman, J., Eising, J., & Lammerts Van Bueren, E. T. (2013). *Potato breeding: A practical manual for the potato chain*. Aardappelwereld.

Velazco, J. G., Malosetti, M., Hunt, C. H., Mace, E. S., Jordan, D. R., & Van Eeuwijk, F. A. (2019). Combining pedigree and genomic information to improve prediction quality: An example in sorghum. *Theoretical and Applied Genetics*, *132*, 2055–2067.

Wartha, C. A., & Lorenz, A. J. (2021). Implementation of genomic selection in public-sector plant breeding programs: Current status and opportunities. *Crop Breeding and Applied Biotechnology*, *21*, e394621S15.

Wu, P.-Y., Stich, B., Renner, J., Muders, K., Prigge, V., & van Inghelandt, D. (2023). Optimal implementation of genomic selection in clone breeding programs—Exemplified in potato: I. Effect of selection strategy, implementation stage, and selection intensity on short-term genetic gain. *The Plant Genome*, *16*(2), e20327.

Yuan, J., Bizimungu, B., De Koeyer, D., Rosyara, U., Wen, Z., & Lagüe, M. (2020). Genome-wide association study of resistance to potato common scab. *Potato Research*, *63*, 253–266.

Yusuf, M., Meng, X., Stefaniak, T. R., Montesinos-López, O. A., & Shannon, L. M. (2025). Leveraging unmanned aerial vehicle derived multispectral data for improved genomic prediction in potato (*Solanum tuberosum*). *The Plant Genome*, *18*(3), e70082.

Zhao, Y., Gowda, M., Liu, W., Würschum, T., Maurer, H. P., Longin, F. H., Ranc, N., & Reif, J. C. (2012). Accuracy of genomic selection in European maize elite breeding populations. *Theoretical and Applied Genetics*, *124*, 769–776.

Zhu, X., Maurer, H. P., Jenz, M., Hahn, V., Ruckelshausen, A., Leiser, W. L., & Würschum, T. (2022). The performance of phenomic selection depends on the genetic architecture of the target trait. *Theoretical and Applied Genetics*, *135*(2), 653–665.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Maggiorelli, A., Thelen, K., Prigge, V., Bruckmüller, J., Muders, K., Hartje, S., & Stich, B. (2025). Enhancing the accuracy of genetic values of tested early-stage potato clones utilizing genotypic, pedigree, and drone-retrieved spectral information.. *The Plant Phenome Journal*, *8*, e70053. https://doi.org/10.1002/ppj2.70053