Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Improving Genomic Prediction of Late Blight Resistance in Potato by Incorporating Indirect Genetic Effects and Foliage Maturity–Corrected Breeding Values

Sagar Kafle

sagar.kafle@qgg.au.dk

Aarhus University

Thinh Tuan Chu

Aarhus University

Merethe Bagge

Danespo A/S

Mathias Timmermann Christiansen

Danespo A/S

Mila Tost

Aarhus University

Hermann Bürstmayr

BOKU University

Just Jensen

Aarhus University

Research Article

Keywords: Genomic Selection, Late Blight Resistance, Indirect Genetic Effects, Potato Breeding, Spatial Effects

Posted Date: November 9th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11242334/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Improving Genomic Prediction of Late Blight Resistance in Potato by Incorporating Indirect Genetic Effects and Foliage Maturity–Corrected Breeding Values

Sagar Kafle[12*], Thinh Tuan Chu[1], Merethe Bagge[2], Mathias Timmermann Christiansen[2], Mila Tost[1], Hermann Bürstmayr[3], Just Jensen [1*]

[1]Center for Quantitative Genetics and Genomics, Aarhus University, Aarhus, Denmark

[2]Danespo A/S, Dyrskuevej 15, 7323 Give. Denmark

[3]Institute of Crop Breeding and Genomics, University of Natural Resources and Life Sciences Vienna, Konrad-Lorenz-Str. 20, 3430 Tulln, Austria

*sagar.kafle@qgg.au.dk 0000-0002-7965-0158

*just.jensen@au.dk 0000-0003-3291-8468

Abstract

Late blight of potato (*Phytophthora infestans*) remains a persistent threat to global potato production. Due to the pathogen's ability to rapidly overcome race-specific resistance genes, modern potato breeding must focus on durable quantitative resistance, including the use of genomic selection approaches. In this study, several genomic prediction models for late blight resistance were evaluated, differing in their treatment of spatial effects and incorporation of indirect genetic effects (IGEs) from neighbors. The optimal model was subsequently extended to a bivariate formulation to derive maturity-corrected genomic breeding values, effectively removing the confounding influence of foliage maturity on late blight resistance. Our results showed that additive genetic variance for late blight resistance accounted for a substantial proportion of total variance ($h^2 = 0.32 \pm 0.02$) and IGEs from neighboring plots were significant based on a standard likelihood ratio test. Although IGEs explained only a small portion of total variances, their inclusion markedly improved model prediction accuracy. We also observed a relatively high genetic correlation between foliage maturity and late blight score (0.56 ± 0.09), indicating that late-maturing lines tend to exhibit higher levels of durable resistance to late blight. After correcting for the confounding influence of foliage maturity, the narrow-sense heritability of the residual breeding value remained moderate ($h^2 = 0.26$), suggesting that meaningful genetic gain through selection is still possible. From a breeder's perspective, selection based on total late blight resistance could unintentionally favor late-maturing genotypes, due to the positive genetic correlation between two traits. However, with residual breeding value, it is possible to get a response to selection for late blight resistance without affecting foliage maturity type.

Keywords: Genomic Selection, Late Blight Resistance, Indirect Genetic Effects, Potato Breeding, Spatial Effects

Key message: Genomic selection improves the accuracy of selection for late blight resistance, and maturity-corrected breeding values enable resistance selection without unintentionally favoring late maturity.

Introduction

Since its emergence in the mid-19th century in the northeastern United States (Lyons 1964; Austin Bourke and M 1964; Saffer et al. 2024), late blight of potato (*Phytophthora infestans*) has remained a persistent challenge to global potato production, largely because the pathogen can rapidly overcome race-specific Rpi (Resistance to *Phytophthora infestans*) genes (Collins et al. 1999; Fry 2008; Blossei et al. 2022). Although over 70 Rpi genes have been identified and mapped across potato chromosomes, their lack of durability has shifted the breeding strategy towards quantitative resistance, which utilizes the cumulative effects of both major and minor Quantitative Trait Loci (QTLs) to provide durable but partial protection (Fry 2008; Haverkort et al. 2016; Stefańczyk et al. 2020). In a meta-analysis by Danan et al. (2011), 144 QTLs for late blight resistance were grouped into 24 meta-QTLs, which were then found to be distributed across all potato chromosomes. Of these, 25 % were associated with Resistance Gene Analogues (RGA), while 50% overlapped with regions previously identified as Durable Resistance Loci (DRLs). Genomic Selection (GS) effectively leverages these numerous minor-effects genes (DRLs) that individually might confer a low level of resistance but, when combined, contribute to more stable and long-lasting resistance (McGee et al. 2025).

Originally developed in context of animal breeding, GS quickly found its way into crop breeding and has become a state-of-the-art method for selection (for a review see Jannink et al. 2010). As such, GS has now been established in potato breeding, and has shown promising results for yield, disease resistance and quality traits (Slater et al. 2016; Stich and Van Inghelandt 2018; Endelman et al. 2018; Yusuf et al. 2024). In a GS program, breeders use field trial data, in combination with genomic information, to obtain unbiased estimates of genomic breeding values for candidate lines. Spatial and neighboring effects are common in such breeding trials, which if unaccounted for can lead to biases in estimation. While traditional row-column models are standard for spatial correction, their effectiveness is reduced in larger or irregular field designs. As an alternative, the random moving-mean model, which assumes random effects of surrounding plots on a target plot within a predefined grid in a field could offer a solution (Xu et al. 2022). Further, practical constraints in breeding trials, often restrict a fully randomized field layout; for instance, closely related lines (family lines) or lines targeting the same markets are frequently planted in proximity. Such an arrangement can introduce bias in estimation of breeding values (e.g., areas planted with susceptible family lines might experience higher inoculum load potentially biasing the estimation of direct genetic effects (Pélissier et al. 2023). These kinds of effects are referred to as Indirect Genetic Effects (IGEs), which describe the influence of an individual's genotype in modifying the phenotype of neighboring plants. Such effects have been documented in work of Bijma (2014); Frank (2007) and Muir (2005). While IGEs have been explored in the context of general disease susceptibility (Pélissier et al. 2023), their specific importance for genomic prediction of disease resistance remains uninvestigated.

Further complication in the genetic analysis of late blight resistance lies in the association of durable late blight resistance with foliage maturity. It has been observed that late-maturing lines tend to exhibit higher levels of resistance (Umaerus 1959; Collins et al. 1999; Oberhagemann et al. 1999; Visker et al. 2004). Both physiological and genetic causes have been proposed in different studies. Physiologically, a short photoperiod has been proposed to link early foliage maturity and late blight susceptibility (Umaerus 1959; Visker et al. 2004). Genetically, the association could be attributed to pleiotropic effects or tight linkage between loci controlling foliage maturity and late blight resistance (Collins et al. 1999; Oberhagemann et al. 1999; Visker et al. 2004; Bormann et al. 2004; Sliwka et al. 2007). Because late maturity is not a desirable trait in many potato breeding programs, separating maturity effects from true resistance remains a critical challenge. Several studies have corrected late blight scores for maturity at a phenotypic level (Bormann et al. 2004, Xue et al. 2019), but addressing it at a genetic level could better disentangle late blight resistance from the confounding influence of maturity.

The objective of this study is twofold: (1) to estimate genetic parameters for late blight resistance and identify an optimal genomic prediction model by evaluating a random moving-mean approach combined with IGEs from neighbors and comparing its performance with the standard row–column method for field-trial correction; and (2), to

use the optimal model identified above to derive maturity-corrected genomic breeding values for late blight resistance by removing the confounding influence of foliage maturity.

Materials and Methods

Plant Material and Experimentation

Phenotypic and genomic information for lines tested in this study was provided by Danespo A/S. These lines were developed and are maintained by Danespo A/S as part of their standard breeding program. Lines included in this study are being developed for 13 different market segments, which broadly represent four primary end-use categories (starch, table, fries, and crisp). A complete list of these market segments is provided in supplementary table S2. In this paper, the terms “clones”, “varieties” and “lines” are used interchangeably to refer to genetically unique individuals.

The late blight trials were conducted in Give, Denmark, over a three-year period from 2022 to 2024. The field design remained largely consistent over the three years of testing, with only minor adjustments. Each year, the entire field was divided into large blocks (lanes), each consisting of 18 testing rows. The total number of blocks (lanes) varied annually, depending on the number of genotypes under evaluation and the number of replicates included in the field trials (see Table 1). The row-to-row distance was maintained at 75 cm, while the plant-to-plant distance was kept at 33 cm. The distance between plots was maintained at 99 cm in the year 2022 and 2023, and this was increased to 132 cm in 2024. Each plot consisted of plants grown from three different tubers of a line. Generally, the number of replicates per genotype varied by testing generation, with early-generation lines generally having fewer replicates. Agronomic management included fertilization at a rate of 120 kg N/ha and standard pre-planting weed management. Additionally, mechanical weed control was carried out in the 2024 growth season.

Table 1 Description of the field design and lines evaluated in late blight trials conducted during 2022–2024.

	2024	**2023**	**2022**	**Total**
Records	4000	2970	3078	10048
Lines tested	1212	975	1717	3197
Blocks (lanes)	3	2	2	7
Rows/block	18	18	18	54
Columns/block	84	110	86	280

The entire perimeter of the field was planted with susceptible spreader lines. Moreover, spreader rows were introduced after every two test rows throughout the field, with the first and last row of each block (lane) also designated as spreader rows (see supplemental Figure S1). Similarly, within each row, the first and last plots were planted with spreader plants. Each spreader row was cultivated with one Bintje cultivar tuber and two tubers of other susceptible cultivars from the Danespo A/S planting materials to promote infection. Spreader rows were inoculated with a spore suspension of *Phytophthora infestans*. A detailed description of the disease inoculation and infection procedures is provided in (Aalborg et al. 2025b). No preventive late blight treatments were applied in the 2023 and 2024 field trials. However, in the 2022 growing season, preventive treatment was applied from the time of sprouting until inoculation of the pathogen.

Phenotypic data collection

To obtain disease score measurements, drone-based field imaging (Mavic 2-Enterprise Advanced- 48MP RGB camera system) was conducted on 4 to 6 flight missions across each growing season. Typically, the first drone flight mission was initiated a specific days after inoculation, depending upon progression of infection (i.e. when approximately 50% of the plants exhibited visible symptoms of late blight) and prevailing weather conditions (see Table S1). To ensure that the infection was due to late blight and not early blight, manual inspection was carried out

before the first drone flight attempt. Flights were conducted at a maximal wind speed of 10.5 m/s, with the drone flying at a height of 20 meters.

Captured raw images were processed via Phenoyard AB (2023), Lund, Sweden, which provided a quantitative assessment of late blight severity for each plot. Disease severity was estimated by analyzing leaf chlorophyll content, which differs between healthy and infected plants, using variations in reflectance within the visible spectrum (Gitelson, Gritz, and Merzlyak, 2003). Each plot was assigned a disease score from 1 (most susceptible) to 9 (highly resistance), according to the internal disease scoring protocol developed by Danespo A/S. Repeated disease score measurements from flight missions was then used to estimate Area Under the Disease Progress Curve (AUDPC), following the approach defined in Jeger & Viljanen-Rollinson (2001) and Shaner (1977).

Foliage maturity data were obtained from routine breeding trials, rather than from disease trials, as the presence of late blight symptoms can interfere with reliable assessment of foliage maturity. Foliage maturity was scored in the field manually using a 1-9 scale, where 9 indicates advanced senescence (i.e early dying off). Assessments were conducted at three time points during the growing season and the scale is split accordingly. During the first assessment, which was conducted when few clones start dying off, the upper end of scales (score 9, 8, and 7) were used. The second assessment occurred once a larger proportion of clones had progressed into later stage of senescence, at which point middle portion of the scale (scores 6, 5, and 4) were used. A final assessment was carried out later in the season to capture the most delayed clones, and the lower end of the scale (scores 3,2, and 1) was applied. Clones exhibiting no signs of senescence at the third assessment were assigned a score of 1. The time between each assessment was approximately 3 days.

Genomic and pedigree information

Genotyping of the clones used in this study was carried out using a custom 20k Affymetrix chip by the French-based company Florimond Desprez, Chapelle-en-Pévèle. The clones were genotyped using 22,224 biallelic SNP positions anchored in the potato DM v 6.1 reference genome. After following filtering based on Thermo Fischer's guidelines, a total of 9,426 SNPs were retained that had a high-quality genotype. Monomorphic SNPs were removed, and biallelic SNPs were kept using a missing data rate of ≤ 50 % and a Minor Allele Frequency (MAF) of ≥ 1 %, generating a final set of 8,731 SNPs. Imputation of the missing SNP data was carried out using a random forest approach with 50 trees in the missRanger R package (v2.6.1). Details of the genotyping workflow and protocols can be found in Aalborg et al. (2025).

The total size of the pedigree records available included 85,481 lines. Custom R code was used to trace the pedigree for the lines under study (3,197). More details on the completeness of the pedigree can be obtained from Yousaf et al. (2025).

Adjustment for heterogeneity of variance

In multi-environmental trials, components of variance (additive genetic and residual variance) often vary across different test environments, and that ignoring this heterogeneity can bias predictions. As such, a flexible approach is to treat the performance in each environment as a separate trait and run a multivariate model. In such models, higher genetic correlation between different test environments means that genotype × environment (GE) interactions result in varying the absolute difference between evaluations of candidates, whereas in the case of lower genetic correlation, there might be changes in ranking of genotypes across environments. Unlike heterogeneity of genetic variance which requires running a multivariate analysis, we can account for heterogeneity of residual variance with constant additive genetic variance by running a generalized linear mixed model. However, if the heterogeneity of variance occurs across all the tested parameters (heterogeneity on a phenotypic level), variance stabilization across all tested environments could be appropriate to ensure that individuals are compared on the same scale.

In our study, an initial univariate model of three years' data revealed violations of the homogeneity of residuals assumptions. Subsequently, a multi-environmental (ME) trivariate model was fitted, where late blight score

measured at different years of testing were parameterized as independent traits. Within this ME framework, residual covariances across years were constrained to zero, because the residuals in different years are independent. The ME model yielded genetic correlations of 0.76-0.87 across different years of testing, indicating minimal re-ranking of individuals. However, the heterogeneity affected all variance parameters in the model, i.e. the heterogeneity was basically on the phenotypic level. Given that variance components for all tested random effects varied substantially across years of testing, we implemented a variance stabilizing transformation. This approach scaled the phenotypic records from each environment to a common phenotypic variance, using variances of year 2022 as the reference environment. As detailed in the supplementary file ('Section: Heterogeneity of variance'), this transformation is predicated on two key assumptions: first, that genetic correlations between environments are high (little re-ranking between environments), and second, that heritability remains constant across years even as individual variance components differ.

Statistical models

The dataset used in this study comprised 3,197 potato lines, for which pedigree information was available. However, SNP marker data were available for 1,897 lines only. This study was conducted in two stages, which comprised (1) identification of an optimal genomic prediction model for late blight resistance, and (2) estimation of residual breeding value for late blight resistance after adjustment for foliage maturity.

To identify the most appropriate model for late blight resistance, we compared three alternative approaches, (1) conventional Row-Column Model, which accounted for spatial variation in field trials using row, column and block effects (we refer to this as 'RC'), (2) Random Moving mean Model (referred to as 'RMMM'), that accounted for spatial effects by incorporating the surrounding plots of a target plot directly as a random effect during model specification. The idea is that in the absence of spatial gradients, the local mean around a target plot should approximate zero, whereas systematic deviations would indicate localized environmental variation. The third model extended the RMMM by inclusion of IGEs from the neighbors, we referred to this model as 'RMMM + IGE'.

Fixed effects included year and market segment (groups), and all the random effects described below were significant based on standard likelihood ratio tests. Across all three models, the direct genetic component was partitioned into additive and non-additive (specific combining ability; sca) effects. The environmental component included spatial effects and residuals. Genotype-by-year (G×Y) interactions were modeled for additive genetic components. However, the three approaches differ primarily in the way spatial variation was modeled in field trials and in the extended model (RMMM+IGEs), in the incorporation of IGEs.

Following the variance stabilization procedure, a univariate model was fitted on the variance stabilized phenotypic data, for subsequent analysis. Model specifications are given as follows;
Row column model ('RC'):

$$\boldsymbol{y}^{*} = \boldsymbol{X\beta} + \boldsymbol{Z_1 g} + \boldsymbol{Z_2 (gy)_g} + \boldsymbol{Z_3 sca} + \boldsymbol{Z_4 b} + \boldsymbol{Z_5 r} + \boldsymbol{Z_6 c} + \boldsymbol{e} \qquad (1)$$

Random Moving Mean Model ('RMMM'):

$$\boldsymbol{y}^{*} = \boldsymbol{X\beta} + \boldsymbol{Z_1 g} + \boldsymbol{Z_2 (gy)_g} + \boldsymbol{Z_3 sca} + \sum_{i=1}^{9} \boldsymbol{Z_i s} + \boldsymbol{e} \qquad (2)$$

Random Moving Mean Model + Indirect Genetic Effects ('RMMM+IGE'):

$$\boldsymbol{y}^{*} = \boldsymbol{X\beta} + \boldsymbol{Z_1 g} + \boldsymbol{Z_2 (gy)_g} + \boldsymbol{Z_3 sca} + \sum_{i=1}^{9} \boldsymbol{Z_i s} + \sum_{j=1}^{8} \boldsymbol{Z_j n_g} + \boldsymbol{e} \qquad (3)$$

Where, $\mathbf{y}^{*}$, $\mathbf{X}$, $\boldsymbol{\beta}$ are vectors for adjusted AUDPC, design matrix for fixed effects, and fixed effects respectively. $\mathbf{g}$, $(\mathbf{gy})_{\mathbf{g}}$, $\mathbf{sca}$, $\mathbf{b}$, $\mathbf{r}$, $\mathbf{c}$, $\mathbf{s}$, $\mathbf{n_g}$ and $\mathbf{e}$ represents vectors of direct additive genetic effects, additive component of G×Y interaction, specific combining ability for parental combinations, row, block and column effects (RC model), spatial effects (RMMM model), IGEs of neighbor lines and residuals. $\mathbf{Z_1}$, $\mathbf{Z_2}$, $\mathbf{Z_3}$, $\mathbf{Z_4}$, $\mathbf{Z_5}$, $\mathbf{Z_6}$, $\mathbf{Z_i}$, $\mathbf{Z_j}$ are the corresponding conformable design matrix allocating phenotypic records to $\mathbf{g}$, $\mathbf{(gy)_g}$, $\mathbf{sca}$, $\mathbf{b}$,$\mathbf{r}$,$\mathbf{c}$,$\mathbf{s}$, and $\mathbf{n_g}$ respectively.

For the above random terms, variance are distributed as $\mathbf{g} \sim N(0, \mathbf{A}\sigma_g^2)$, $\mathbf{gy_g} \sim N(0, \mathbf{A_{gy}}\sigma^2_{gy_g})$, $\mathbf{sca} \sim N(0, \mathbf{I}\sigma^2_{sca})$,$\mathbf{s} \sim N(0, \mathbf{I}\sigma_s^2)$, $\mathbf{b} \sim N(0, \mathbf{I}\sigma_b^2)$, $\mathbf{r} \sim N(0, \mathbf{I}\sigma_r^2)$, $\mathbf{c} \sim N(0, \mathbf{I}\sigma_c^2)$, and $\mathbf{e} \sim N(0, \mathbf{I}\sigma_e^2)$. For RMM+IGEs, direct additive effects and indirect genetic effects follow a multivariate normal with $\begin{bmatrix}\mathbf{g}\\ \mathbf{n_g}\end{bmatrix} \sim MVN(0, \mathbf{A}\otimes G_o)$, where $G_o = \begin{bmatrix}\sigma_g^2 & \sigma_{g.n_g}\\ \sigma_{g.n_g} & \sigma^2_{n_g}\end{bmatrix}$.

$\mathbf{A}$ is the numerator relationship matrix (NRM). $\mathbf{A_{gy}}$ is derived from the same (co)variance structure as the A matrix but is specifically used to model genetic relationships between lines within each year of testing, as mentioned in (Ortiz et al. 2022) $\mathbf{A_{gy}} = \begin{bmatrix}\mathbf{A_{2022}} & 0 & 0\\ 0 & \mathbf{A_{2023}} & 0\\ 0 & 0 & \mathbf{A_{2024}}\end{bmatrix}$ (see fig S2).

Since heritability and variance component estimates obtained using H-matrix (single-step analysis combining pedigree and genomic information) could be biased due to selective genotyping (Wang et al. 2020). Variance components and thus heritability estimates were obtained from a pedigree-based Numerator Relationship Matrix (A) matrix. Model above used 'A', so-called ABLUP, for estimating variance components. The estimated parameters from ABLUP model were subsequently used within the HBLUP framework for genomic prediction, where NRM matrix was replaced by the H matrix. NRM was constructed following Kerr et al. (2012). The Genomic Relationship Matrix (GRM) was computed using the formulation originally proposed for diploids by VanRaden (VanRaden, 2008), and subsequently adapted for polyploid in Ashraf et al. (2016). For single-step analyses, we followed an approach described by Martini et al. (2018). All of these implementations were carried out using the "AGHmatrix (v2.1.4)" R package (Amadeu et al. 2023).

Multi trait model (MT): The RMMM+ IGEs model was extended to MT model, by considering late blight and foliage maturity in a bivariate analysis.

$$\begin{bmatrix}\boldsymbol{y_1^*}\\ \boldsymbol{y_2}\end{bmatrix} = \begin{bmatrix}\boldsymbol{X_1} & 0\\ 0 & \boldsymbol{X_2}\end{bmatrix}\begin{bmatrix}\boldsymbol{\beta_1}\\ \boldsymbol{\beta_2}\end{bmatrix} + \begin{bmatrix}\boldsymbol{Z_{11}} & 0\\ 0 & \boldsymbol{Z_{21}}\end{bmatrix}\begin{bmatrix}\boldsymbol{g_1}\\ \boldsymbol{g_2}\end{bmatrix} + \begin{bmatrix}0\\ \boldsymbol{Z_{21}}\end{bmatrix}[\boldsymbol{l}] + \begin{bmatrix}\boldsymbol{Z_{12}} & 0\\ 0 & \boldsymbol{Z_{22}}\end{bmatrix}\begin{bmatrix}(\boldsymbol{gy})_{g_1}\\ (\boldsymbol{gy})_{g_2}\end{bmatrix} + \begin{bmatrix}0\\ \boldsymbol{Z_{22}}\end{bmatrix}[(\boldsymbol{gy})_l] +$$

$$\begin{bmatrix}\boldsymbol{Z_{13}}\\ 0\end{bmatrix}[\boldsymbol{sca}] + \begin{bmatrix}\sum_{i=1}^{9}\boldsymbol{Z_i}\\ 0\end{bmatrix}[\boldsymbol{s}] + \begin{bmatrix}\sum_{j=1}^{8}\boldsymbol{Z_j}\\ 0\end{bmatrix}[\boldsymbol{n_{g_1}}] + \begin{bmatrix}\boldsymbol{e_1}\\ \boldsymbol{e_2}\end{bmatrix} \quad (4)$$

Where, $\mathbf{y_1^*}$ and $\mathbf{y_2}$ are records for variance adjusted AUDPC and foliage maturity respectively. The terms $\boldsymbol{\beta_1}, \mathbf{X_1}, \mathbf{g_1}, \mathbf{sca}, (\mathbf{gy})_{\mathbf{g_1}}, \mathbf{n_{g_1}}, \mathbf{e_1}$ follows the definitions provided in Eq. (3), while $\boldsymbol{\beta_2}, \mathbf{X_2}, \mathbf{g_2}, (\mathbf{gy})_{\mathbf{g_2}}, and\ \mathbf{e_2}$ denotes the equivalent component for foliage maturity. Unique to foliage maturity, $\mathbf{l}\ and\ (\mathbf{gy})_{\mathbf{l}}$ represent the non-additive line effects and line component of G×Y interaction.

For above random terms, variance are distributed as $\begin{bmatrix}\mathbf{g_1}\\ \mathbf{n_{g_1}}\\ \mathbf{g_2}\end{bmatrix} \sim MVN(0, \mathbf{A}\otimes\mathbf{G_o})$, where $\mathbf{G_o} = \begin{bmatrix}\sigma^2_{g_1} & \sigma_{g_1.n_{g_1}} & \sigma_{g_1.g_2}\\ \sigma_{g_1.n_{g_1}} & \sigma^2_{n_{g_1}} & \sigma_{g_2.n_{g_1}}\\ \sigma_{g_1.g_2} & \sigma_{g_2.n_{g_1}} & \sigma^2_{g_2}\end{bmatrix}$, $\mathbf{gy_{g_1}} \sim N(0, \mathbf{A_{gy_1}}\sigma^2_{gy_{g_1}})$, $\mathbf{sca} \sim N(0, \mathbf{I}\sigma^2_{sca})$, $\mathbf{s} \sim N(0, \mathbf{I}\sigma^2_s)$, $\mathbf{l} \sim N(0, \mathbf{I}\sigma^2_l)$, $\mathbf{gy_{g_2}} \sim N(0, \mathbf{A_{gy_2}}\sigma^2_{gy_{g_2}})$, $\mathbf{gy_l} \sim N(0, \mathbf{I}\sigma^2_{gy_l})$, $\mathbf{e_1} \sim N(0, \mathbf{I}\sigma^2_{e_1})$ and $\mathbf{e_2} \sim N(0, \mathbf{I}\sigma^2_{e_2})$.

It should be noted that foliage maturity and late blight scores were recorded from different field trials. Therefore, due to statistical independence, it is assumed that their error variance does not share any covariance; the same assumption also applied to genotype-by-year (G×Y) interactions.

Residual breeding value for late blight conditional on the foliage maturity were obtained as in (Strathe et al. 2014): $\boldsymbol{g_{re} = g_1 - b.g_2}$; where, $b = \frac{\sigma_{g1.g2}}{\sigma^2_{g_2}}$; where, $\mathbf{g_{re}}$ is the residual breeding value for late blight; $\mathbf{g_1}$ is the breeding value for late blight, $\mathbf{g_2}$is the breeding value for foliage maturity and b is the genetic regression of foliage maturity on late blight score.

Variance components and heritability

The estimation of variance component was carried out using the Average-Information Restricted Maximum Likelihood method (AI-REML), as implemented in the DMU software package (Madsen et al. 2014). Total variance ($\widehat{\sigma^2_y}$) is estimated using ABLUP as $\widehat{\sigma^2_y} = \bar{G}\widehat{\sigma^2_g} + \widehat{\sigma^2_{(gy)_g}} + \widehat{\sigma^2_{sca}} + \widehat{\sigma^2_b} + \widehat{\sigma^2_r} + \widehat{\sigma^2_c} + \widehat{\sigma^2_e}$ (RC); $\widehat{\sigma^2_y} = \bar{G}\widehat{\sigma^2_g} + \widehat{\sigma^2_{(gy)_l}} + \widehat{\sigma^2_{sca}} + 9\widehat{\sigma^2_s} + \widehat{\sigma^2_e}$ (RMMM); $\widehat{\sigma^2_y} = \bar{G}\widehat{\sigma^2_g} + \widehat{\sigma^2_{(gy)_l}} + \widehat{\sigma^2_{sca}} + 9\widehat{\sigma^2_s} + 8\bar{G}\widehat{\sigma^2_{n_g}} + \widehat{\sigma^2_e}$ (RMMM+IGEs).

Where, $\bar{G}$ is the average of diagonal values in the relationship matrix and reflects the average level of inbreeding among the tested lines. $\widehat{\sigma^2_g}, \widehat{\sigma^2_{(gy)_g}}, \widehat{\sigma^2_{sca}}, \widehat{\sigma^2_s}, \widehat{\sigma^2_{n_g}}, \widehat{\sigma^2_b}, \widehat{\sigma^2_r}, \widehat{\sigma^2_c}$ and $\widehat{\sigma^2_e}$ represents additive genetic variance, genetic components of G×Y variance, specific combining ability variance, spatial variance (RMMM), IGEs of neighbor variance, block, row and column variance (RC model), and residual variance. $\widehat{\sigma^2_s}$ is the spatial variance due to the effect of a single surrounding plot on to target plot, and since each target plot is influenced by eight neighbors and itself, the total spatial component is approximated as $9\widehat{\sigma^2_s}$. Similarly, $\widehat{\sigma^2_{n_g}}$ captures the indirect effects from a single neighboring line. As with the spatial effects, the total variance due to IGEs of neighbors is calculated as the sum of the contribution from the eight neighbors ($8\bar{G}\widehat{\sigma^2_{n_g}}$).

In the MT model, the total phenotypic variance for late blight was estimated as $\widehat{\sigma^2_{y_1}} = \bar{G}\widehat{\sigma^2_{g_1}} + +\widehat{\sigma^2_{(gy)_{g_1}}} + \widehat{\sigma^2_{sca}} + 9\widehat{\sigma^2_s} + 8\widehat{\sigma^2_{n_{g_1}}} + \widehat{\sigma^2_{e_1}}$; where, $\widehat{\sigma^2_{g_1}}, \widehat{\sigma^2_{(gy)_{g_1}}}, \widehat{\sigma^2_{sca}}, \widehat{\sigma^2_s}, \widehat{\sigma^2_{n_{g_1}}}$, and $\widehat{\sigma^2_{e_1}}$ are same as defined above. To estimate the variance of the residual breeding value for late blight conditional on foliage maturity, the additive genetic variance of late blight was adjusted as: $\sigma^2_{g_1|\,g_2} = \bar{G}\widehat{\sigma^2_{g_1}} - \sigma_{g_1.g_2}\left(\sigma^2_{g_2}\right)^{-1}\sigma_{g_1.g_2}$;where, $\sigma^2_{g_1|\,g_2}$is additive genetic variance of late blight conditional on foliage maturity, $\sigma^2_{g_1}$ and $\sigma^2_{g_2}$ are the additive genetic variance for late blight resistance and foliage maturity, $\sigma_{g_1.g_2}$is genetic covariance between the two traits.

For all the studied models, narrow ($\widehat{h^2}$) and broad($\widehat{H^2}$) sense heritability was estimated as: $\widehat{h^2} = \frac{\bar{G}\widehat{\sigma^2_g}}{\widehat{\sigma^2_y}}$ and $\widehat{H^2} = \frac{\bar{G}\widehat{\sigma^2_g} + \widehat{\sigma^2_{sca}}}{\widehat{\sigma^2_y}}$ respectively. Narrow sense design heritability ($\widehat{h^2_{\bar{y}_c}}$) was estimated as $\widehat{h^2_{\bar{y}_c}} = \frac{\bar{G}\widehat{\sigma^2_g}}{\bar{G}\widehat{\sigma^2_g} + \frac{\widehat{\sigma^2_{(gy)_g}}}{N} + \widehat{\sigma^2_{sca}} + \frac{\widehat{\sigma^2_b}}{n_b} + \frac{\widehat{\sigma^2_r}}{n_b} + \frac{\widehat{\sigma^2_c}}{n_b} + \frac{\widehat{\sigma^2_e}}{n}}$ (RC), $\widehat{h^2_{\bar{y}_c}} = \frac{\bar{G}\widehat{\sigma^2_g}}{\bar{G}\widehat{\sigma^2_g} + \widehat{\sigma^2_l} + \frac{\widehat{\sigma^2_{(gy)_g}}}{N} + \widehat{\sigma^2_{sca}} + 9\widehat{\sigma^2_s} + \frac{\widehat{\sigma^2_e}}{n}}$ (RMMM) , and $\widehat{h^2_{\bar{y}_c}} = \frac{\bar{G}\widehat{\sigma^2_g}}{\bar{G}\widehat{\sigma^2_g} + \frac{\widehat{\sigma^2_{(gy)_g}}}{N} + \widehat{\sigma^2_{sca}} + 9\widehat{\sigma^2_s} + \frac{8\bar{G}\widehat{\sigma^2_{g_n}}}{n} + \frac{\widehat{\sigma^2_e}}{n}}$ (RMMM+IGE and LB+SEN).

For obtaining design heritability for residual breeding value of late blight, the same equation as in LB+SEN was used; however, the genetic variance ($\widehat{\sigma^2_g}$) was replaced by the additive genetic variance of late blight conditional on foliage maturity ($\sigma^2_{g_1|\,g_2}$).

N represents the number of years of testing (*3*), n represents the average number of replications per genotype (*3.14*), and n_b represents the average number of replications per block (*1.4*).

Cross validation and prediction accuracy

The BLUEs and variance components were estimated using the full set of phenotypic data. The BLUE estimates for the fixed effects in the models were used to correct individual plot observations, and then the line average ($\bar{\mathbf{y}}_\mathbf{c}$) was computed after correction for these fixed effects. All models were evaluated using a leave-one-line-out (LOO) cross-validation scheme. In this scheme, the phenotype of a single clone was removed, and the remaining clones served as the training set to predict the masked line's breeding value based on a defined relationship matrix (HBLUP). This process was repeated for all clones, providing a prediction for each line. The prediction accuracy was assessed using Pearson's correlation, calculated between the corrected line averages and the predicted breeding value. This correlation was then scaled by the square root of the design heritability. The scaling provides an estimate of the correlation between the estimated breeding value (EBV) and the underlying true breeding values for lines, when evaluated without their own phenotypic records (Legarra et al. 2008).

For Row Column model, RMMM, and RMMM + IGEs accuracy was obtained as $\mathrm{cor}(\bar{\mathbf{y}}_\mathbf{c}, \hat{\mathbf{g}})/\sqrt{\widehat{h^2_{\bar{y}_c}}}$, whereas, for MT model, prediction accuracy for breeding value of late blight was estimated as $\mathrm{cor}(\bar{\mathbf{y}}_{\mathbf{1c}}, \widehat{\mathbf{g}_\mathbf{1}})/\sqrt{\widehat{\mathbf{h}^2_{\mathbf{1}\bar{\mathbf{y}}_\mathbf{c}}}}$; and accuracy for residual breeding value of late blight score as $\mathrm{cor}(\bar{\mathbf{y}}_{\mathbf{1c}}, \widehat{\mathbf{g}_{\mathbf{re}}})/\sqrt{\widehat{\mathbf{h}^2_{\mathbf{re}\bar{\mathbf{y}}_\mathbf{c}}}}$. Moreover, the regression coefficient of $\bar{\mathbf{y}}_\mathbf{c}$ on $\hat{\mathbf{g}}$ was estimated to obtain the inflation of GEBVs.

Results

Phenotypic description

Descriptive statistics for AUDPC are presented in Fig. 1. A total of 10,048 plot observations with 3,197 unique lines were evaluated for AUDPC across three years of testing. The mean AUDPC was observed to be 65.832 ± 42.12, with the highest average AUDPC recorded in 2023 (108.50 ± 46.7), followed by 2022 (53.48 ± 19.40) and 2024 (43.61 ± 24.10).

Market segments (groups) were defined based on the intended end use of potato lines, including but not limited to starch, table, fries and crisp potatoes. The highest AUDPC was obtained in group 1 (87.90 ± 42.3), followed by group 8 (84.79 ± 47.10), while the lowest AUDPC was observed in group 5 (34.00 ± 28.600) and group 2 (40.66 ± 37.50). The coefficient of variation ranged from 15.76% to 81.68% for AUDPC by group, indicating substantial variability in disease progression across market segments. Early maturing varieties are generally more susceptible to late blight and, therefore, tend to exhibit higher AUDPC values. Detailed information on AUDPC and Variance-stabilized AUDPC by year and market segment is provided in Table S2 and S3.

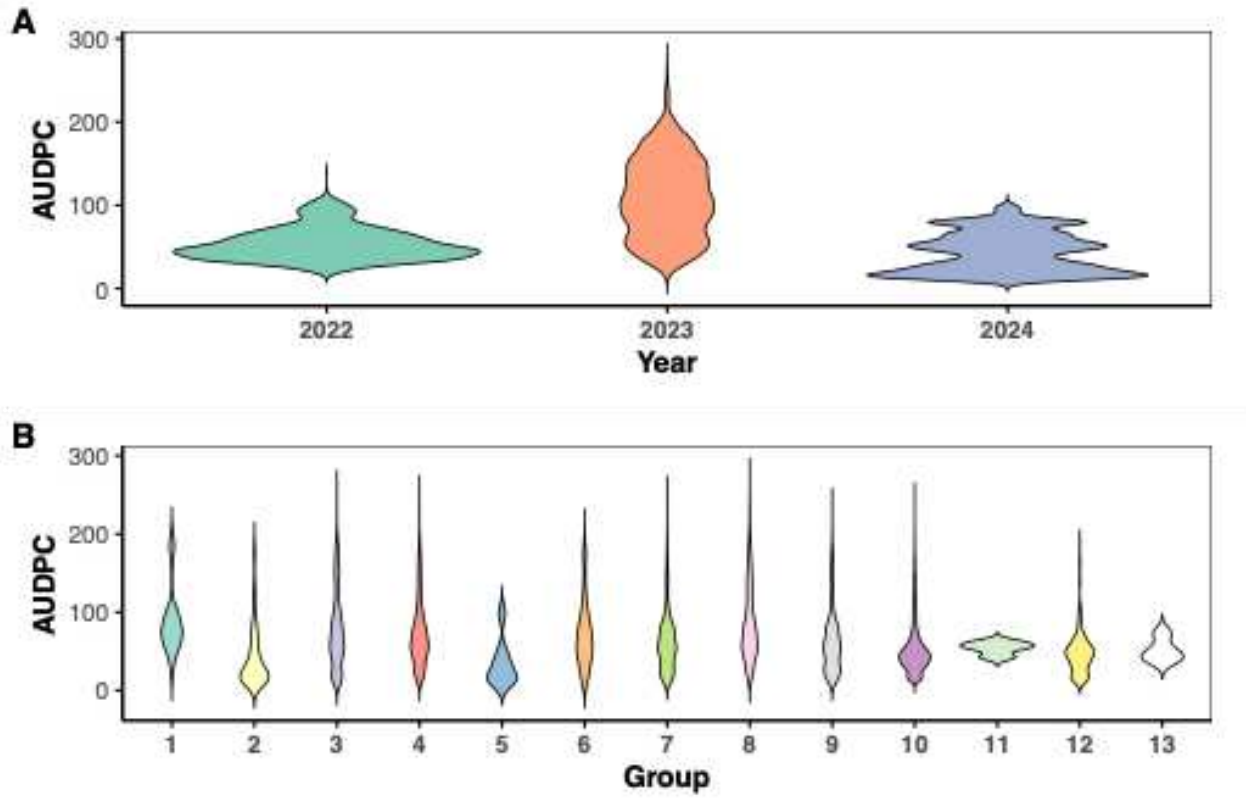


Fig. 1 Density plots showing the distribution of AUDPC values across testing years and market segments

Furthermore, the descriptive statistics for foliage maturity are presented in Supplementary Table S4 and S5. A total of 13,106 records collected over six years were evaluated for foliage maturity. The mean foliage maturity scores were 6.40 ± 1.21. Foliage maturity varied more across market segments (coefficient of variation:11-29%) than across years of testing (coefficient of variation:14-22%).

(Co)variance components

The variance components estimated from the variance-adjusted AUDPC dataset for all tested models are summarized in Figure 2 with details provided in supplementary table S6 and S7. Across different model specifications, the proportion of additive genetic variance ranged from 30% to 34% of the total phenotypic variance. While the total phenotypic variance remained comparable between tested models, notable differences were observed in partitioning of variance components across models.

In the RC model, the additive genetic variance accounted for 30.16% of the total phenotypic variance. The spatial variance, estimated as the sum of block, row and column effects, explained 17.29 % of the total variance. Notably, column variance (8.55% of total phenotypic variance) exceeded the row variance (3.85% of total phenotypic variance), likely reflecting the greater length of columns relative to rows in the experimental layout. In this model, the residual error variance accounted for a relatively large proportion (34.17%) of the total phenotypic variance.

In the RMMM model, the additive genetic variance explained 32.40 % of the total phenotypic variance. Compared with the RC model, modeling spatial heterogeneity using a random moving mean increased the proportion of variance attributed to spatial effects by 10.91%, with a corresponding reduction in residual variance by 14.53% (Table S6). This indicates that the RMMM model more effectively captured within-field spatial heterogeneity that was not accounted for by the RC model. The remaining variance components were comparable to those estimated under the RC model.

In the RMMM+IGE model, indirect genetic effects (IGEs) from the neighboring plots were incorporated. These IGEs were found to be statistically significant based on the standard likelihood ratio test ($\chi^2 = 55.0, df = 2$). However, the proportion of variance explained by IGEs from neighboring plots was relatively small (1.73% of the total variance). The additive genetic effects explained 34.02% of the total phenotypic variance. Further, the genetic correlation between the direct additive genetic effects of the target line with the IGEs from the neighboring genotypes was estimated at 0.62 ± 0.12. The remaining variance components were comparable to those estimated under the RMMM model.

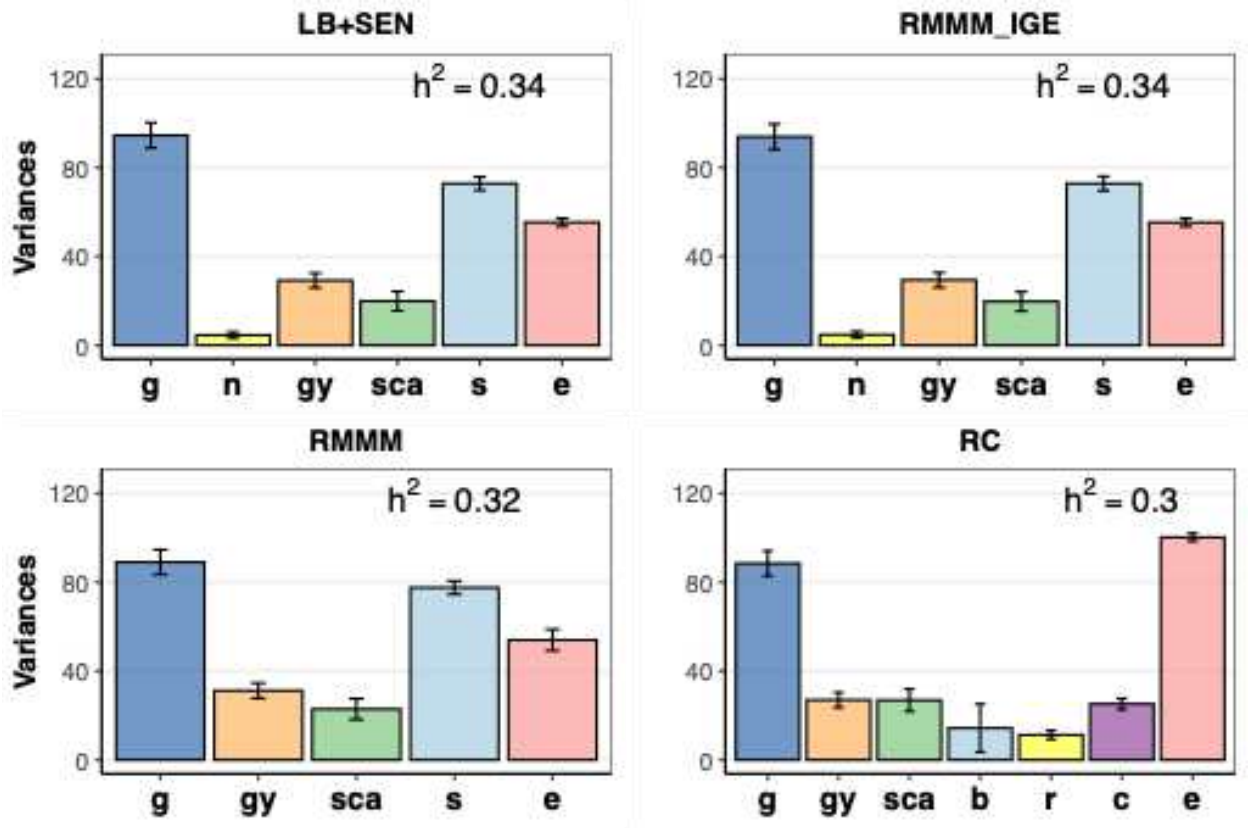


Fig. 2 Estimated variance components for variance adjusted AUDPC of late blight of potato under different models.**g** is additive genetic effect, **gy(g)** is genomic components of G×Y interaction effect, **s** is spatial effects, **n** is IGEs from neighbor plots, **sca** is the specific combining ability, **e** is the residual effects, **b** is the block effect, **r** is the row effects and c is the column effect respectively. h^2 represents the narrow sense heritability estimates of the models (RC: Row-Column Model; RMMM: Random Moving Mean Model; RMMM+IGE: Random Moving Mean Model + Indirect genetic Effect Model, LB+SEN: Bivariate model)

In the LB+SEN (bivariate) model, variance components for late blight were comparable to those obtained from the RMMM+IGE model, as similar partitioning of variance components was expected. However, deriving the residual breeding value to remove genetic effects shared with foliage maturity reduced the additive genetic variance for late blight. Consequently, the additive genetic variance specific to late blight resistance decreased from 34.21% (LB+SEN) to 26.17% (Res_LB+SEN) of the total phenotypic variance.

For foliage maturity, a detailed summary of variance components is provided in Table 2. The additive genetic variance accounted for 8.6 % of the total phenotypic variance. The estimated genetic correlation between foliage maturity and late blight score was 0.56 ± 0.09, indicating that late-maturing lines tend to exhibit higher levels of durable resistance to late blight.

Table 2 Description of the estimated variance components for foliage maturity.

Trait	σ_g^2	σ_l^2	$\sigma_{(gy)g}^2$	$\sigma_{(gy)l}^2$	σ_{error}^2	σ_y^2
Maturity	0.12±0.048 (8.6%)	0.22±0.028 (15.8%)	0.19±0.025 (13.67%)	0.03±0.019 (2.1%)	0.83±0.012 (60.71%)	1.39

(The number in the parenthesis represents contribution of each variance components to total variance)

Heritability

Three types of heritability estimates were calculated in this study: narrow-sense heritability, broad-sense heritability and narrow-sense design heritability. Estimates of different heritability measures for late blight resistance and foliage maturity are given in Table 3.

Table 3 Comparison of heritability estimates for late blight resistance and foliage maturity in potato across different models.

Model	h^2	H^2	$h_{\bar{y}_c}^2$
Late blight			
RC	0.30±0.02	0.39±0.02	0.45±0.03
RMMM	0.32±0.02	0.41±0.02	0.41±0.02
RMMM+IGE	0.34±0.02	0.41±0.02	0.44±0.02
LB+SEN	0.34±0.01	0.41±0.02	0.44±0.02
Res_LB+SEN	0.26±0.01	0.34±0.02	0.35±0.02
Foliage maturity			
LB+SEN	0.09±0.02	0.24±0.01	-

The narrow sense heritability of late blight ranged from 0.30-0.34, with no statistically significant difference among the tested models. Narrow sense heritability for late blight shows moderate additive genetic control. In contrast, the heritability of foliage maturity was substantially lower (0.09), suggesting higher environmental and non-additive genetic variation.

Genomic prediction and cross-validation

Cross-validation of the models was performed using a leave-one-line-out cross-validation. This cross-validation scheme simulates a realistic breeding scenario by assessing how accurately the model can predict the phenotypic value of a line that is missing from a field trial for any reason, including new crosses that are genotyped but not yet tested. The correlation between genomic estimated breeding values and line mean corrected for fixed effects, along with the accuracy of prediction obtained from different models is presented in Figure 3. The results show notable differences in prediction performance between different models.

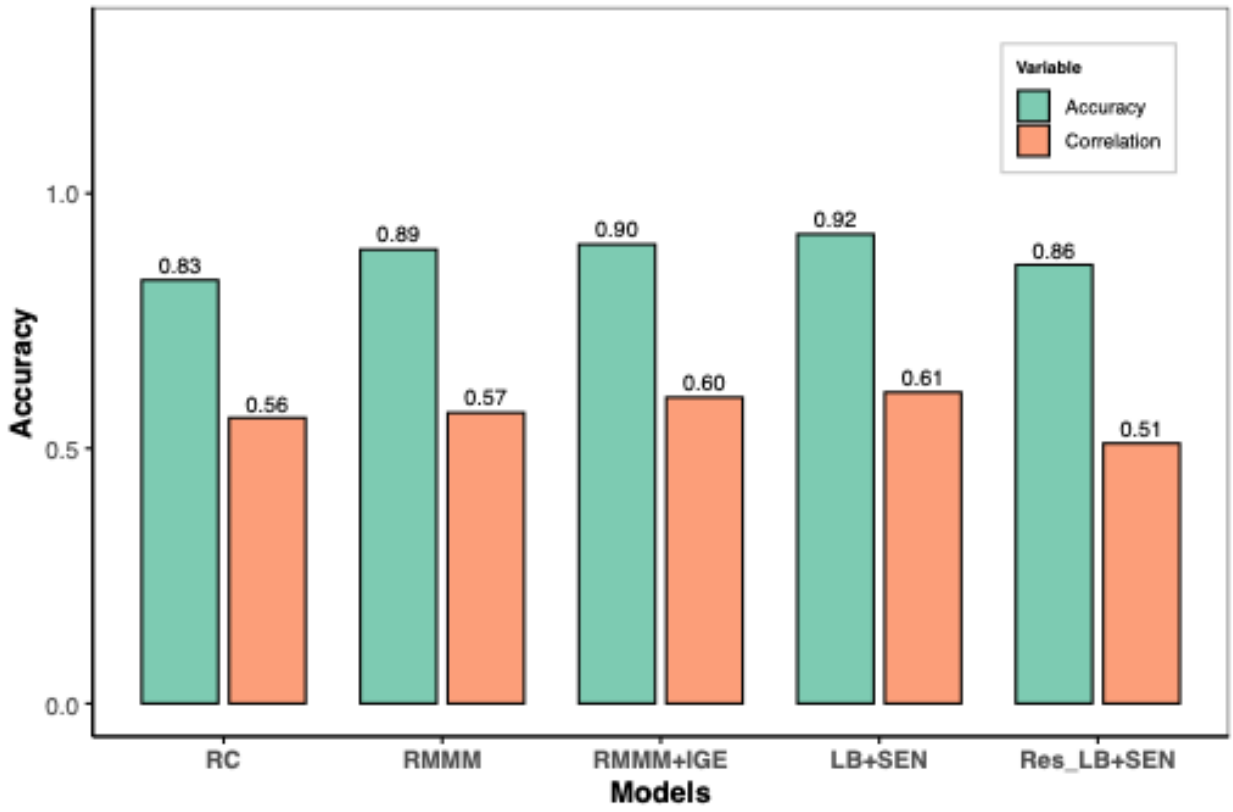


Fig. 3 Accuracy of prediction for different models. The orange bars represent the correlation between line mean corrected for fixed effects, and the GEBVs. The green bar indicates prediction accuracy after scaling the correlation by the square root of heritability (RC: Row-Column Model; RMMM: Random Moving Mean Model; RMMM+IGE: Random Moving Mean Model + Indirect genetic Effect Model, LB+SEN: Bivariate model)

The highest prediction accuracy was observed for LB+SEN model (0.92), followed by RMMM+IGE (0.90), RMMM (0.89) and RC (0.83) model respectively. Considering IGEs in the RMMM+IGE model improved the prediction performance (0.60) of direct genetic effects (GEBVs) as compared to that of RMMM (0.57) and RC model (0.56). In contrast, only a marginal improvement in correlation was observed for the bivariate LB+SEN model (0.61) relative to the univariate LB model (RMMM+IGE; 0.60), indicating limited benefit from the bivariate formulation in this case.

Prediction accuracy for residual breeding values (0.86) of late blight resistance was markedly reduced relative to that of LB+SEN (0.92). This decrease reflects the removal of genetic information of late blight resistance that is explained by foliage maturity. Furthermore, across all models tested, a marked difference in prediction accuracy was observed between genotyped and non-genotyped lines, as shown in Table 4. As expected, prediction accuracy for genotyped lines was 0.17 to 0.23 points higher than that for non-genotyped lines, reflecting the added value of genomic information in improving prediction performance.

Table 4 Correlation between line mean corrected for fixed effects, and GEBVs, and the accuracy of prediction for genotyped and non-genotyped lines along with inflation of the GEBVs.

Models	**Genotyped**			**Non-Genotyped**		
	Correlation	Accuracy	Inflation	Correlation	Accuracy	Inflation
RC	0.61	0.90	1.16	0.38	0.57	1.22
RMMM	0.62	0.97	1.14	0.40	0.62	1.28
RMMM+IGE	0.63	0.95	1.16	0.46	0.69	1.40

LB+SEN	0.63	0.95	1.15	0.46	0.70	1.38
Res_LB+SEN	0.56	0.94	1.40	0.36	0.61	1.45

As shown in the Fig. 4A, a strong genetic correlation was observed between the GEBVs for late blight resistance and foliage maturity, indicating that these two traits are genetically associated. To account for this relationship, late blight breeding values were corrected based on the genetic regression of late blight score on foliage maturity. After correction, the residual late blight breeding value showed little association with foliage maturity (Fig. 4B). This correction can be interpreted as a maturity-dependent shrinkage of the late blight breeding values, with the magnitude of the correction increasing with the genetic contribution of foliage maturity.

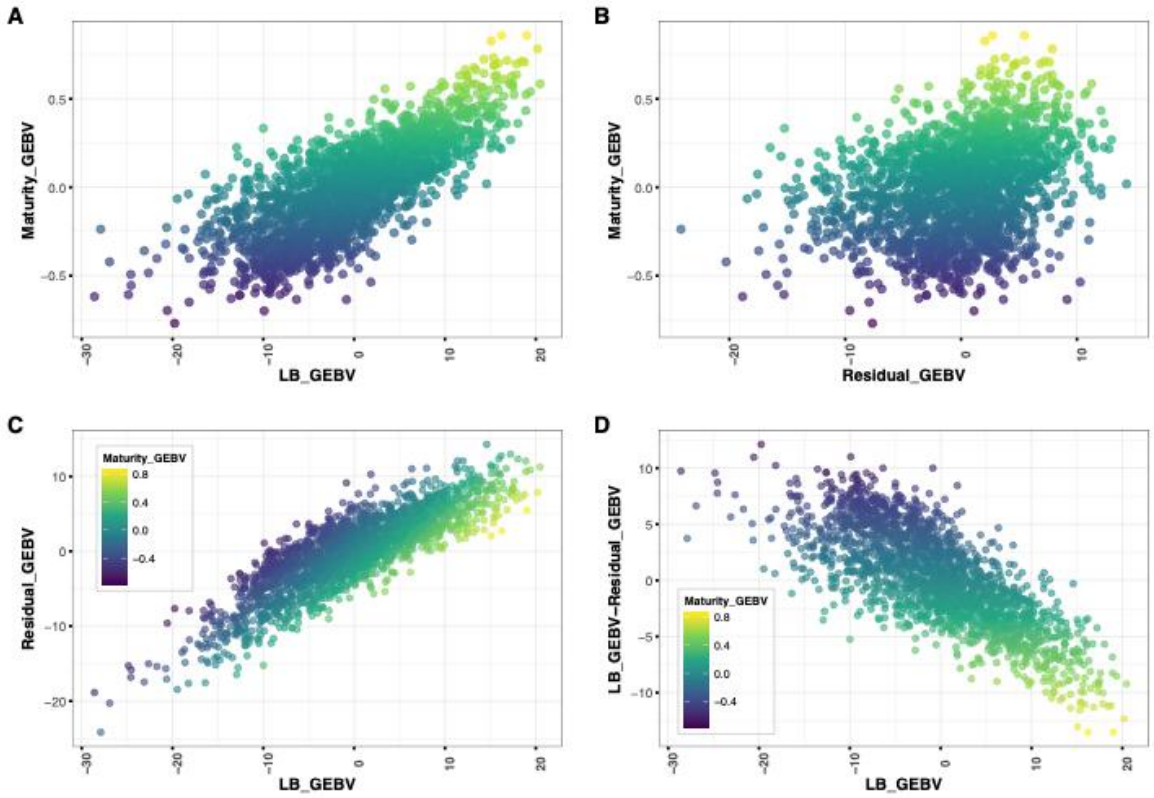


Fig. 4 *A. Scatter plot showing relationship between GEBVs for late blight and foliage maturity (higher values contribute towards susceptibility and early maturity). B. Scatter plot of the residual breeding values for late blight and GEBVs for foliage maturity. C. Scatter plot of residual GEBV versus total GEBV for late blight. For a given level of late blight GEBV, lines with higher maturity GEBV (earlier maturity) tend to have lower residual GEBV (indicating greater resistance), whereas lines with lower maturity GEBV (later maturity) tend to have higher residual GEBV (indicating greater susceptibility). D. Relationship between shrinkage (difference between late blight GEBV and residual breeding value) and late blight GEBV. For lines with higher LB_GEBV (more susceptible), the applied shrinkage reduces the residual breeding value (making them more resistance); this shrinkage is proportional to maturity GEBV. Similarly, for lines with lower LB_GEBV (more resistant), the shrinkage increases the residual breeding value (making them susceptible) and this shrinkage is also proportional to maturity GEBV.*

Discussion

This study aimed to estimate breeding values for late blight resistance in tetraploid potato. First, we compared various genomic prediction models for late blight resistance that differed in their treatment of spatial heterogeneity and Indirect Genetic Effects (IGEs) in order to identify the best performing one. Second, we derived maturity-

corrected breeding values for late blight resistance by removing the confounding genetic influence of foliage maturity.

Spatial variation accounted for a substantial proportion of the total variation (17-32 %) across all the models tested. The RMMM (Random Moving Mean) model was more effective in capturing the spatial trends in the field compared to the RC (Row-Column) model. The RMMM approach is conceptually similar to the moving-means as a covariate model given by Lado et al. (2013), where the phenotype value of a target plot is first regressed on surrounding plots, and the estimated spatial effects are then corrected for before genomic prediction is carried out. However, this approach involves regression on a function of the data in the model, that tend to lead towards downward biases in variance estimates (Guo et al. 2020). Several other approaches have also been proposed to account for spatial variation in field trials, including autoregressive models (Gilmour et al. 1997), moving-means as a covariate (Lado et al. 2013), Bayesian framework (Selle et al. 2019) and smoothing approaches (Verbyla et al. 1999; Rodríguez-Álvarez et al. 2018). However, a comparison of all available methods is beyond the scope of this study. The primary motivation for using the RMMM approach was that, in its absence, IGEs may inadvertently capture spatial trends rather than genetic interaction (results not presented here). Therefore, the combined RMMM + IGE (Random Moving Mean with IGEs) model was implemented to better separate spatial variation from IGEs and obtain more reliable estimates.

In the present study, additive genetic variance for late blight resistance accounted for a substantial proportion (30-34%) of the total phenotypic variance. The narrow-sense plot heritability was estimated at 0.41-0.45, indicating moderate to moderately high heritability. Heritability estimates for late blight resistance reported in the literature vary widely, ranging from approximately 0.31 to 0.86. For example, moderate heritability estimates have been reported by Enciso-Rodriguez et al. (2018;0.29), Solano et al. (2014;0.53–0.58), and Romé et al. (2026;0.63). In contrast, substantially higher estimates have been observed in other studies, including 0.77 (Stich and Van Inghelandt 2018), 0.78–0.79 (Haynes and Christ 1999) and 0.75–0.86 (Sood et al. 2023). The difference in estimates of heritability likely reflects differences in the metrics (plot versus design heritability) used to report them.

IGEs from neighboring plots were statistically significant based on a likelihood ratio test in the RMMM+IGE model ($\chi^2 = 55.0, df = 2$). The genetic correlation between direct additive genetic effects and IGEs was estimated at 0.62 ± 0.12, indicating a moderately strong positive association. This suggests a synergistic relationship; genotypes with higher susceptibility also tend to exacerbate disease pressure on neighboring plots and vice versa. These findings align with the work of Andrivon et al. (2003) and Connolly et al. (1995), who observed that resistant individuals in heterogenous cropping systems or a varietal mixtures, contributed towards overall disease suppression. While IGEs have been extensively studied in interspecific systems, recent experimental work by Pélissier, Buendia, et al. (2021); Pélissier et al. (2023) demonstrates that these effects also occur within intraspecific mixtures, where the genetic identity of a neighboring lines modulates the disease susceptibility of a focal line -a phenomenon they termed as neighbor modulated susceptibility (NMS). Biologically, IGEs could be explained beyond simple pathogen propagation, including direct signaling between genotypes (chemical or physical signals), modulation of the shared microbiome, changes in resource availability, and neighbor-state dependency (wounded or infection) (for more details, see Pélissier, Buendia, et al. 2021; Pélissier, Violle, et al. 2021). Furthermore, Pélissier et al. (2023) identified specific lines in intraspecific mixtures that consistently increased or reduced susceptibility in neighboring plants. Similarly, breeding values for IGEs derived from genomic prediction model could be used to identify genotypes that suppress disease spread and incorporate them into cultivar mixtures to enhance localized field resistance.

In our study, the correlation between GEBVs and line mean corrected for fixed effects ranged from 0.56 to 0.61, with prediction accuracy from 0.83 to 0.92, suggesting genomic selection can be a viable method for improving durable resistance to late blight. Comparable levels of prediction accuracy have been reported in previous studies, including 0.815–0.864 by Lindqvist-Kreuze et al. (2021), 0.66–0.85 by Stich and Van Inghelandt (2018), while lower levels of accuracy are reported by Sood et al. (2023; 0.41-0.51) and Enciso-Rodriguez et al. (2018; 0.26-0.31).

The observed variation among studies likely reflects differences in prediction models, trait definition, population structure, test materials and metrics used to report predictive performance. In general, the relatively higher level of heritability for late blight trait supports the relatively high genomic prediction accuracies obtained in the study.

Several attempts have been made to account for the correction of foliage maturity when analyzing late blight resistance (Visker et al. 2004; Bormann et al. 2004; Xue et al. 2019). For example, Visker et al. (2004); Xue et al.(2019) used foliage maturity as a fixed regressor to obtain adjusted AUDPC values at the phenotypic level. However, phenotypic correction should be applied with caution, as it accounts for all correlation; even when part of the correlation may result from confounding environmental factors. In contrast, genomic-level correction targets only the genetically confounded component of the correlation.

The estimated genetic correlation between foliage maturity and late blight resistance in our study was 0.56± 0.09, indicating a moderate positive association between two traits. Such a correlation suggests that at least some loci affecting resistance are either genetically linked to, or pleiotropic with, loci controlling foliage maturity. Consistent with this interpretation, several linked loci have previously been reported to influence both traits (Collins et al. 1999; Bradshaw et al. 2004, 2006; Bormann et al. 2004; Costanzo et al. 2005; Danan et al. 2009). However, numerous QTLs are also expected to influence late blight resistance independent of foliage maturity.

The residual breeding value represents the component of additive genetic effects for late blight resistance that remains after removing the genetic covariance with foliage maturity. That is, it represents the additive effects attributable to loci that jointly influence both traits that have been removed (Visker et al. 2004; Bormann et al. 2004). As expected, removing the shared genetic covariance resulted in a reduction in additive genetic variance and, consequently, lower prediction accuracy. The remaining variation captured by the residual breeding value can be attributed to resistance QTLs that are not linked to, or pleiotropic with, foliage maturity loci (Visker et al. 2004). Although the narrow-sense heritability of the residual breeding value is moderate ($h^2 = 0.26$), this can still provide a reasonable response to selection. From a breeder's perspective, selection based on total late blight resistance could unintentionally favor late-maturing genotypes, due to the positive genetic correlation between two traits. However, with residual breeding value, it is possible to obtain a response to selection to late blight without affecting foliage maturity type (Visker et al. 2004; Bormann et al. 2004). In contrast, where late maturity is not an issue, selection on total breeding value might be more advantageous. Thus, the choice between total and residual breeding values should be guided by the breeding objective and target production environments.

References

Aalborg T, Christiansen MathiasT, Romé H, et al (2025) Aerial drone-based phenotyping and GWAS of emergence and late blight resistance in tetraploid potato

Amadeu RR, Garcia AAF, Munoz PR, Ferrão LFV (2023) AGHmatrix: genetic relationship matrices in R. Bioinformatics 39:btad445. https://doi.org/10.1093/bioinformatics/btad445

Andrivon D, Lucas J-M, Ellissèche D (2003) Development of natural late blight epidemics in pure and mixed plots of potato cultivars with different levels of partial resistance. Plant Pathol 52:586–594. https://doi.org/10.1046/j.1365-3059.2003.00882.x

Ashraf BH, Byrne S, Fé D, et al (2016) Estimating genomic heritabilities at the level of family-pool samples of perennial ryegrass using genotyping-by-sequencing. Theor Appl Genet 129:45–52. https://doi.org/10.1007/s00122-015-2607-9

Austin Bourke, M P (1964) Emergence of Potato Blight, 1843–46. Nature 203:805–808. https://doi.org/10.1038/203805a0

Bijma P (2014) The quantitative genetics of indirect genetic effects: a selective review of modelling issues. Heredity 112:61–69. https://doi.org/10.1038/hdy.2013.15

Blossei J, Gäbelein R, Hammann T, Uptmoor R (2022) Late blight resistance in wild potato species—Resources for future potato (Solanum tuberosum) breeding. Plant Breed 141:314–331. https://doi.org/10.1111/pbr.13023

Bormann CA, Rickert AM, Castillo Ruiz RA, et al (2004) Tagging Quantitative Trait Loci for Maturity-Corrected Late Blight Resistance in Tetraploid Potato with PCR-Based Candidate Gene Markers. Mol Plant-Microbe Interactions® 17:1126–1138. https://doi.org/10.1094/MPMI.2004.17.10.1126

Bradshaw JE, Bryan GJ, Hackett CA, et al (2004) Dissection and analysis of quantitative disease resistance in tetraploid potato. Euphytica 137:13–18. https://doi.org/10.1023/B:EUPH.0000040499.19190.fb

Bradshaw JE, Hackett CA, Lowe R, et al (2006) Detection of a quantitative trait locus for both foliage and tuber resistance to late blight [Phytophthora infestans (Mont.) de Bary] on chromosome 4 of a dihaploid potato clone (Solanum tuberosum subsp. tuberosum). Theor Appl Genet 113:943–951. https://doi.org/10.1007/s00122-006-0353-8

Collins A, Milbourne D, Ramsay L, et al (1999) QTL for field resistance to late blight in potato are strongly correlated with maturity and vigour. Mol Breed 5:387–398. https://doi.org/10.1023/A:1009601427062

Connolly T, McNICOL JW, Wastie RL, Stewart HE (1995) Evaluating between-plant interference in field trials for assessing potato genotypes for resistance to late blight. Ann Appl Biol 127:273–282. https://doi.org/10.1111/j.1744-7348.1995.tb06672.x

Costanzo S, Simko I, Christ BJ, Haynes KG (2005) QTL analysis of late blight resistance in a diploid potato family of Solanum phureja x S. stenotomum. TAG Theor Appl Genet Theor Angew Genet 111:609–617. https://doi.org/10.1007/s00122-005-2053-1

Danan S, Chauvin J-E, Caromel B, et al (2009) Major-effect QTLs for stem and foliage resistance to late blight in the wild potato relatives Solanum sparsipilum and S. spegazzinii are mapped to chromosome X. Theor Appl Genet 119:705–719. https://doi.org/10.1007/s00122-009-1081-7

Danan S, Veyrieras J-B, Lefebvre V (2011) Construction of a potato consensus map and QTL meta-analysis offer new insights into the genetic architecture of late blight resistance and plant maturity traits. BMC Plant Biol 11:16. https://doi.org/10.1186/1471-2229-11-16

561 Enciso-Rodriguez F, Douches D, Lopez-Cruz M, et al (2018) Genomic Selection for Late Blight and Common Scab
562 Resistance in Tetraploid Potato (Solanum tuberosum). G3 GenesGenomesGenetics 8:2471–2481.
563 https://doi.org/10.1534/g3.118.200273

564 Endelman JB, Carley CAS, Bethke PC, et al (2018) Genetic Variance Partitioning and Genome-Wide Prediction
565 with Allele Dosage Information in Autotetraploid Potato. Genetics 209:77–87.
566 https://doi.org/10.1534/genetics.118.300685

567 Frank SA (2007) All of life is social. Curr Biol 17:R648–R650. https://doi.org/10.1016/j.cub.2007.06.005

568 Fry W (2008) Phytophthora infestans: the plant (and R gene) destroyer. Mol Plant Pathol 9:385–402.
569 https://doi.org/10.1111/j.1364-3703.2007.00465.x

570 Gilmour AR, Cullis BR, Verbyla AP (1997) Accounting for Natural and Extraneous Variation in the Analysis of
571 Field Experiments. J Agric Biol Environ Stat 2:269–293. https://doi.org/10.2307/1400446

572 Gitelson AA, Gritz † Y, Merzlyak MN (2003) Relationships between leaf chlorophyll content and spectral
573 reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. J Plant
574 Physiol 160:271–282. https://doi.org/10.1078/0176-1617-00887

575 Guo X, Svane SF, Füchtbauer WS, et al (2020) Genomic prediction of yield and root development in wheat under
576 changing water availability. Plant Methods 16:90. https://doi.org/10.1186/s13007-020-00634-0

577 Haverkort AJ, Boonekamp PM, Hutten R, et al (2016) Durable Late Blight Resistance in Potato Through Dynamic
578 Varieties Obtained by Cisgenesis: Scientific and Societal Advances in the DuRPh Project. Potato Res 35–
579 66. https://doi.org/10.1007/s11540-015-9312-6

580 Haynes KG, Christ BJ (1999) Heritability of resistance to foliar late blight in a diploid hybrid potato population of
581 Solanum phureja×Solanum stenotomum. Plant Breed 118:431–434. https://doi.org/10.1046/j.1439-
582 0523.1999.00394.x

583 Jannink J-L, Lorenz AJ, Iwata H (2010) Genomic selection in plant breeding: from theory to practice. Brief Funct
584 Genomics 9:166–177. https://doi.org/10.1093/bfgp/elq001

585 Jeger MJ, Viljanen-Rollinson SLH (2001) The use of the area under the disease-progress curve (AUDPC) to assess
586 quantitative disease resistance in crop cultivars. Theor Appl Genet 102:32–40.
587 https://doi.org/10.1007/s001220051615

588 Kerr RJ, Li L, Tier B, et al (2012) Use of the numerator relationship matrix in genetic analysis of autopolyploid
589 species. TAG Theor Appl Genet Theor Angew Genet 124:1271–1282. https://doi.org/10.1007/s00122-012-
590 1785-y

591 Lado B, Matus I, Rodríguez A, et al (2013) Increased Genomic Prediction Accuracy in Wheat Breeding Through
592 Spatial Adjustment of Field Trial Data. G3 GenesGenomesGenetics 3:2105–2114.
593 https://doi.org/10.1534/g3.113.007807

594 Legarra A, Robert-Granié C, Manfredi E, Elsen J-M (2008) Performance of Genomic Selection in Mice. Genetics
595 180:611–618. https://doi.org/10.1534/genetics.108.088575

596 Lindqvist-Kreuze H, De Boeck B, Unger P, et al (2021) Global multi-environment resistance QTL for foliar late
597 blight resistance in tetraploid potato with tropical adaptation. G3 GenesGenomesGenetics 11:jkab251.
598 https://doi.org/10.1093/g3journal/jkab251

599 Lyons FSL (1964) The great hunger: Ireland, 1845-9. By Cecil Woodham-Smith. , Pp. xii, 510. London Hamish
600 Hamilton. 1962. 30s. Ir Hist Stud 14:77–79. https://doi.org/10.1017/S0021121400020186

Madsen P, Jensen J, Labouriau R, et al (2014) DMU - A Package for Analyzing Multivariate Mixed Models in quantitative Genetics and Genomics: 10th World Congress on Genetics Applied to Livestock Production (WCGALP)

Martini JWR, Schrauf MF, Garcia-Baccino CA, et al (2018) The effect of the H−1 scaling factors τ and ω on the structure of H in the single-step procedure. Genet Sel Evol 50:16. https://doi.org/10.1186/s12711-018-0386-x

McGee R, Zaleski-Cox M, Jayawardana MA, et al (2025) Breeding for quantitative disease resistance: Case studies, emerging approaches, and exploiting pathogen variation. Crop Sci 65:e70202. https://doi.org/10.1002/csc2.70202

Muir WM (2005) Incorporation of Competitive Effects in Forest Tree or Animal Breeding Programs. Genetics 170:1247–1259. https://doi.org/10.1534/genetics.104.035956

Oberhagemann P, Chatot-Balandras C, Schäfer-Pregl R, et al (1999) A genetic analysis of quantitative resistance to late blight in potato: towards marker-assisted selection. Mol Breed 5:399–415. https://doi.org/10.1023/A:1009623212180

Ortiz R, Crossa J, Reslow F, et al (2022) Genome-Based Genotype × Environment Prediction Enhances Potato (Solanum tuberosum L.) Improvement Using Pseudo-Diploid and Polysomic Tetraploid Modeling. Front Plant Sci 13:. https://doi.org/10.3389/fpls.2022.785196

Pélissier R, Ballini E, Temple C, et al (2023) The genetic identity of neighboring plants in intraspecific mixtures modulates disease susceptibility of both wheat and rice. PLOS Biol 21:e3002287. https://doi.org/10.1371/journal.pbio.3002287

Pélissier R, Buendia L, Brousse A, et al (2021a) Plant neighbour-modulated susceptibility to pathogens in intraspecific mixtures. J Exp Bot 72:6570–6580. https://doi.org/10.1093/jxb/erab277

Pélissier R, Violle C, Morel J-B (2021b) Plant immunity: Good fences make good neighbors? Curr Opin Plant Biol 62:102045. https://doi.org/10.1016/j.pbi.2021.102045

Rodríguez-Álvarez MX, Boer MP, Eeuwijk FA van, Eilers PHC (2018) Correcting for spatial heterogeneity in plant breeding experiments with P-splines. Spat Stat 23:52–71. https://doi.org/10.1016/j.spasta.2017.10.003

Romé H, Bagge M, Jensen J (2026) The Genetic Basis of Resistance to Late Blight in Tetraploid Potatoes. Crop Breed Genet Genomics 8:. https://doi.org/10.20900/cbgg20260001

Saffer A, Tateosian L, Saville AC, et al (2024) Reconstructing historic and modern potato late blight outbreaks using text analytics. Sci Rep 14:2523. https://doi.org/10.1038/s41598-024-52870-2

Selle ML, Steinsland I, Hickey JM, Gorjanc G (2019) Flexible modelling of spatial variation in agricultural field trials with the R package INLA. Theor Appl Genet 132:3277–3293. https://doi.org/10.1007/s00122-019-03424-y

Shaner G (1977) The Effect of Nitrogen Fertilization on the Expression of Slow-Mildewing Resistance in Knox Wheat. Phytopathology 77:1051. https://doi.org/10.1094/Phyto-67-1051

Slater AT, Cogan NOI, Forster JW, et al (2016) Improving Genetic Gain with Genomic Selection in Autotetraploid Potato. Plant Genome 9:plantgenome2016.02.0021. https://doi.org/10.3835/plantgenome2016.02.0021

Sliwka J, Jakuczun H, Lebecka R, et al (2007) Tagging QTLs for late blight resistance and plant maturity from diploid wild relatives in a cultivated potato (Solanum tuberosum) background. TAG Theor Appl Genet Theor Angew Genet 115:101–112. https://doi.org/10.1007/s00122-007-0546-9

641 Solano J, Bravo I, Esnault F, Brabant P (2014) Resistance to Phytophthora infestans in Solanum tuberosum
642 landraces in Southern Chile. Trop Plant Pathol 39:307–315. https://doi.org/10.1590/S1982-
643 56762014000400005

644 Sood S, Bhardwaj V, Bairwa A, et al (2023) Genome-wide association mapping and genomic prediction for late
645 blight and potato cyst nematode resistance in potato (Solanum tuberosum L.). Front Plant Sci 14:.
646 https://doi.org/10.3389/fpls.2023.1211472

647 Stefańczyk E, Plich J, Janiszewska M, et al (2020) Marker-assisted pyramiding of potato late blight resistance genes
648 Rpi-rzc1 and Rpi-phu1 on di- and tetraploid levels. Mol Breed 40:89. https://doi.org/10.1007/s11032-020-
649 01169-x

650 Stich B, Van Inghelandt D (2018) Prospects and Potential Uses of Genomic Prediction of Key Performance Traits in
651 Tetraploid Potato. Front Plant Sci 9:. https://doi.org/10.3389/fpls.2018.00159

652 Strathe AB, Mark T, Nielsen B, et al (2014) Deriving Genomic Breeding Values for Residual Feed Intake from
653 Covariance Functions of Random Regression Models: 10th World Congress on Genetics Applied to
654 Livestock Production (WCGALP)

655 Umaerus V (1959) The relationship between peroxidase activity in potato leaves and resistance toPhytophthora
656 infestans. Am Potato J 36:124–131. https://doi.org/10.1007/BF02857274

657 VanRaden PM (2008) Efficient methods to compute genomic predictions. J Dairy Sci 91:4414–4423.
658 https://doi.org/10.3168/jds.2007-0980

659 Verbyla AP, Cullis BR, Kenward MG, Welham SJ (1999) The Analysis of Designed Experiments and Longitudinal
660 Data by Using Smoothing Splines. J R Stat Soc Ser C Appl Stat 48:269–311

661 Visker MHPW, van Raaij HMG, Keizer LCP, et al (2004) Correlation between late blight resistance and foliage
662 maturity type in potato. Euphytica 137:311–323. https://doi.org/10.1023/B:EUPH.0000040451.21852.d8

663 Wang L, Janss LL, Madsen P, et al (2020) Effect of genomic selection and genotyping strategy on estimation of
664 variance components in animal models using different relationship matrices. Genet Sel Evol 52:31.
665 https://doi.org/10.1186/s12711-020-00550-w

666 Xu Z, Cannon SB, Beavis WD (2022) Applying Spatial Statistical Analysis to Ordinal Data for Soybean Iron
667 Deficiency Chlorosis. Agronomy 12:2095. https://doi.org/10.3390/agronomy12092095

668 Xue W, Haynes KG, Qu X (2019) Characterization of Early Blight Resistance in Potato Cultivars. Plant Dis
669 103:629–637. https://doi.org/10.1094/PDIS-05-18-0794-RE

670 Yousaf MF, Tomar V, Romé H, et al (2025) Rate of double reduction and genetic variability in yield, quality, and
671 senescence related traits in tetraploid potato (Solanum tuberosum L.). Front Plant Sci 16:.
672 https://doi.org/10.3389/fpls.2025.1560123

673 Yusuf M, Miller MD, Stefaniak TR, et al (2024) Genomic prediction for potato (Solanum tuberosum) quality traits
674 improved through image analysis. Plant Genome 17:e20507. https://doi.org/10.1002/tpg2.20507

675 Phenoyard (2023). https://phenoyard.com/. Accessed 16 May 2026

676

677

678

Funding:

Open access funding provided by Aarhus Universitet. The author(s) declare that financial support was received for the research and/or publication of this article. This study was supported by the Green Development Program (GUDP) under the Ministry of Food, Agriculture, and Fisheries, Denmark, with research grant no. 34009-20-1643.

Competing Interests:

The authors declare no competing interests

Author Contributions

SK: Data curation, Formal analysis, Investigation, Validation, Visualization, Writing – original draft.

TTC: Validation, Writing – review & editing

MB: Data curation, Funding acquisition, Project administration, Writing – review & editing.

MTC: Data acquisition

MT: Validation, Writing – review & editing.

HB: Project administration, Writing – review & editing

JJ: Funding acquisition, Methodology, Project administration, Supervision, Validation, Visualization, Writing – review & editing

Data Availability

Data is available from the corresponding author upon reasonable request and subject to approval and material transfer agreement with the industry partner.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- KafleetalTAGSupplementary.docx