

*Article*

# The Genetic Basis of Resistance to Late Blight in Tetraploid Potatoes

Hélène Romé <sup>1</sup>, Merethe Bagge <sup>2</sup>, Just Jensen <sup>1,\*</sup>

<sup>1</sup> Center for Quantitative Genetics and Genomics, Aarhus University, Aarhus 8000, Denmark; hero@seges.dk (HR)

<sup>2</sup> DANESPO A/S, Dyrskuevej 15, Give 7323, Danmark; mba@danespo.com (MB)

\* Correspondence: Just Jensen, Email: Just.Jensen@qgg.au.dk; Tel.: +45-4082-1680

---

## ABSTRACT

Late blight is a disease responsible for major losses to the potato industry and breeding new resistant varieties is a sustainable way to fight late blight and reduce the need for chemical treatments. This study aimed at understanding the genetic basis of late blight resistance in a commercial population of potatoes. Historical records from a large breeding program were used. Different statistical models were evaluated to determine the contribution of additive genetic variance, non-additive genetic variance, and genotype-by-decade of testing interaction variance to the total phenotypic variance. GWAS were conducted to identify genomic regions associated with late blight resistance and the part of the additive genetic variance explained by these regions were estimated. Breeding values for resistance toward late blight were predicted using four different modeling approaches. In the two last models the most significant SNP of significant genomic regions were included as fixed effects. Predictive ability was assessed using five-fold cross-validation by correlating corrected phenotypes or breeding values predicted using all data with those predicted without the phenotypes in the validation populations. Results showed that late blight resistance was highly heritable and that the contribution of the genotype-by-decade interaction was significant, indicating differential resistance toward different late blight strains dominating in each decade. Three genomic regions significantly affecting late blight resistance explained only 13% of the total additive genetic variance in resistance. Genomic breeding values (GEBV) showed high predictive ability suggesting that genetic or genomic selection could be an effective tool to develop new late blight resistant varieties. A strategy combining genomic selection and including knowledge on genomic regions associated with resistance is advantageous.

---

## Open Access

Received: 26 Nov 2025

Accepted: 21 Jan 2026

Published: 26 Jan 2026

Copyright © 2026 by the author. Licensee Hapres, London, United Kingdom. This is an open access article distributed under the terms and conditions of Creative Commons Attribution 4.0 International License.

---

**KEYWORDS:** *Phytophthora infestans*; tetraploid potatoes; resistance genes; host pathogen coevolution; genomic selection

---

## ABBREVIATIONS

LB, Late Blight; Rgene, gene of resistance; *Pi*, *Phytophthora infestans*; AVR, avirulence gene.

## INTRODUCTION

Late blight (LB) is one of the most devastating diseases in potatoes, tomatoes, and tobacco. It is caused by an oomycete called *Phytophthora infestans* (*Pi*), which can attack all parts of the plant. An infection by this pathogen can result in complete destruction of the plant in a few days. *Pi* originated from Peru and Mexico and was most likely introduced to Europe in the mid XIXth century along with contaminated potatoes imported from North America in order to fight another disease. Subsequently, LB was partially the cause of the famine in Europe in the XIXth century, that in particular hit Ireland [1]. From this point, massive efforts were conducted to fight LB, going from the development of fungicides, management recommendations, and development of resistant lines through screening for natural resistance in established potato varieties as well as in wild potatoes [1]. The consequences of LB and of the control measures taken were in 2009 estimated to cost €9 billion worldwide [2]. This cost included the loss due to LB contamination, the cost of the fungicides and many other factors such as the societal cost, including for example the impact on human health. While having an important cost, the use of fungicides is widely preponderant to fight LB. In 2016 in the Netherlands, the amounts of fungicides used against LB represented 50% of the total amount of pesticides used [3]. Besides their economic and societal costs, with time, some fungicides became inefficient as *Pi* evolved to become resistant to the treatments. One of the latest examples is the LB genotype EU43 which is resistant to mandipropamid [4].

Another way to fight LB is by establishing new genetic resistance. Following the mid-XIXth century crisis, intensive work was done especially in the early XXth century to identify wild potatoes which were resistant to LB and the genes responsible for this resistance [5]. It is assumed that the resistance of a host to a pathogen relies on the interaction of a gene of resistance (Rgene) expressed in the host and an avirulence gene (AVR gene) expressed in the pathogen [6]. Those two genes presumably co-evolved simultaneously and, therefore, the plants and the pathogens carrying the matching pairs of Rgenes and AVR genes originate from the same geographic regions. This phenomenon is the reason why primary sources of resistance to *Pi* were searched for in wild populations of *Solanum* spp. in South America and Mexico. Over 70 genes of resistance to *Pi* has now been recognized as identified in wild *Solanum* species and as well as some in cultivated potatoes *Solanum tuberosum* [7,8]. This has led to the development of varieties showing higher resistance. However, here again, because of its ability to evolve, *Pi* can escape the resistance

conveyed by incorporated Rgenes. Indeed, the introgression of eleven dominant resistant genes first identified in *Solanum demissum* L. (R1–R11) did not lead to long term resistance to *Pi* [9]. Moreover, while some resistance genes such as *Rpi-blb3* [10] or the *Rpi2* locus [11] confer resistance toward a large spectrum of pathotypes while others such as *R3a*, *R3b*, *R6* and *R7* [12,13] confer a resistance only to specific pathotypes. One of the strategies proposed to slow down the ability of the pathogen to escape resistance is to stack several Rgenes in one cultivar [14]. However, such R-genes need a broad spectrum of resistance in order to ensure long term resistance. Introgression of two or three resistance genes through traditional breeding is a lengthy process. It took half a century to get two new varieties of *S. tuberosum* carrying *Rpi-blb2* coming initially from *S. bulbocastanum* [3]. The difficulty in introgression also depends on the endosperm balance number (EBN) that determines the ease of crossing related potato species. Different methods such as hybrid breeding and genetic engineering could be used to stack Rgenes more efficiently in cultivated potato. Some studies have used such technologies and were able to stack up to 4 Rgenes into *S. tuberosum* [3,15–17], leading to plants being highly resistant to LB. In theory, if a single clone carried the advantageous allele for four R-genes, it would be nearly impossible for the pathogen to overcome this resistance [18]. However, this relies on prerequisites such as that the R-genes should be 100% effective, which most likely is unrealistic. Moreover, currently, the deployment on the European market of varieties developed using genetic engineering is limited. We should also consider if resistance should only rely on the Rgenes or whether genetic variation in the polygenic background is of importance for resistance to LB. In a recent review McGee et al. [19] argued that use of identified R-Genes needs to be combined with quantitative disease resistance, often based on many minor genes, because resistance due to specific genes tends to vane over time due to strong selection on the disease agents.

Creating lines resistant to LB through breeding is a possibility to avoid crop destruction by the disease and reduce the current reliance on pesticides. This is also reflected in the general strategy for the International Potato Center (CIP) [20]. Heritability estimates for LB resistance reported in literature ranged from 0.31 to 0.85 [21–25], which suggests that selection for LB resistance can be very effective. One of the challenges for selecting for LB resistance has been to be able to improve resistance without negatively impacting other traits of economic importance. One example is maturity, for which it has been observed that cultivars selected for LB resistance showed late maturity. It has been shown that QTLs for blight resistance and QTLs for maturity were found at same position [26,27].

Traditional genetic improvement in potatoes has been slow due to the need for extensive testing in multiple environments and the low multiplication rate due to the vegetative formation of clones. However, the genetic constitution of clones is fixed after the first crossing generation

and this, combined with extensive genotyping in early multiplication generations, can lead to genomic predictions of line performance before large scale multi environmental trials (METs) can be conducted. If genotyping is sufficiently cost effective this can lead to a remarkably high selection intensity and short generation intervals.

The lack of persistent resistance toward *Pi* as discussed above is due to constant development in the disease agent. The Euroblight network monitors the rapid evolution of *Pi* across the main potato growing regions in Europe. Clearly there are major time-dependent changes in the dominant *Pi* populations. Yearly reports on results of this monitoring can be found at <https://agro.au.dk/forskning/internationale-platforme/euroblight> (accessed on 20 July 2023) [28]. Long term testing of resistance will enable monitoring development in resistance over time and to estimate the genetic correlation between resistance measured at separate times. This will indicate the combined effect of evolution of the disease agent and the effects of potential selection for resistance for LB in potato. To our knowledge such analysis has not been attempted before.

In the period investigated the breeding program mostly was based on sequential phenotypic selection on disease traits in early clonal generations, and later selection on maturity, yield, and quality traits. If genetic correlations between early recorded traits and later yields traits are negative this may lead to limited genetic progress because advances in early selection steps are counteracted by selection in later steps due to negative relationships between early and late recorded traits.

The aims of our study were to estimate the heritability for LB resistance in a Danish potato population and search of quantitative trait loci with large effects in order characterize the genetic architecture of LB resistance in potato and to develop genomic based selection criteria for LB resistance, where both identified genome areas carrying disease resistance and quantitative disease resistance based on many minor genes are exploited to increase resistance to LB. In addition, we aim to evaluate the genetic development of LB resistance over longer time periods.

## MATERIAL AND METHODS

### Plant Material

The tested plants were all parts of the routine breeding program run by Danespo A/S. Early in the breeding program wild relatives were crossed into the population and in more recent years markers have been used for early selection and genes stacking. Historically, selection in the population has primarily been based on phenotypic records from inoculated field trials. The same testing procedure for LB resistance has been used throughout the period investigated.

Every year new parents were selected and crossed. Seeds were grown per family in the green house and at an early stage moved to individual pots. A single tuber from each individual pot was grown in the field for replication. In following years tubers from individuals of interest were tested for LB resistance. There were no clones that did not stem from the routine breeding program. The criteria for selecting parents have changed over the years but is not expected to influence results reported here. The plant material stem from a total of 36 years of crossing new parents. In early years only a few progenies of selected parents were tested but in later years a large group of clones were included (Table 1). This reflects a gradual increase in the importance of LB resistance in the breeding program.

**Table 1.** Description of the phenotypic data \*.

	Number of Observations	Number of Lines	Mean	SD
All Years	3580	1204	4.91	2.31
1980s	15	4	3.80	0.86
1990s	53	22	4.85	2.06
2000s	241	87	5.30	2.70
2010s	698	474	5.62	2.51
2020s	2573	924	4.69	2.18

\* The number of observations, number of lines with records, the average value of the raw record of LB resistance and its standard deviation by decades and all over the years.

### Phenotyping

The resistance to *Pi* was phenotyped by exposing the plants to an isolate of the disease agent collected in Denmark the year preceding the resistance test. Tests for disease resistance were performed for 36 consecutive years and, therefore, different pathotypes of *Pi* were used. However, no systematic characterization of the pathotype was carried out, but it must be expected that the dominant pathotype has changed over the years. (Euroblight, <https://agro.au.dk/forskning/internationale-platforme/euroblight>) (accessed on 20 July 2023) [28]. The test trial was designed by alternating one infector row and two test rows. The infector row was a mix of Bintje, which is a highly susceptible variety and was used across all the 36 years of test, and 2 other susceptible varieties that varied over time. On all four sides of the trial, there was an infector row. Each plot consisted of three plants with a distance of 33 cm, and each plot was 75 cm away from each other. The resistance to LB was assessed by attributing a grade to each plot, constituting three plants, based on foliage observation. The grades ranged from 1 (dead plant) to 9 (no sign of infection). A grade of 2 meant that only 0.1% to 4% of the green leaves remained, and the stems were fully attacked. For a grade of 3, 5% to 14% of the leaves were still green and the stems were also attacked. When 15% to 39% of the leaves were green and the stems were attacked, it was given a grade 4. Grade 5 meant that 40% to 59% of green leaves persisted, with the stems still being affected. If 60% to 70% of the leaves were green but

the stems remained untouched, it was given a grade of 6. Grade 7 was given when only 5 to 10 leaves showed infection, and finally, grade 8 signified just 1 to 2 leaves were infected. This definition corresponds to scoring for resistance toward LB which is common among breeders. Papers focusing on epidemiology tend more often to use a reversed scale and then denoting it susceptibility. A full description of the scores is given in [29].

### Genotyping

Sample DNA from 1628 varieties were genotyped using a custom 10k Affymetrix chip (10647 SNP). Genotyping was done over two batches, with five lines being genotyped in both batches. Fifteen SNP markers showing inconsistent genotypes in more than three of the five lines genotyped in both years were discarded. Then, only one genotype record was kept for further analyses of the duplicated genotypes. Next, 1108 SNP were excluded for having a call rate lower than 50%. No individuals were excluded based on an individual call rate lower than 40%. At the end, the genotype dataset used for analysis consisted of 1628 individuals genotyped for 9524 SNP. Finally, missing values were imputed using a random forest approach [30] with 50 trees using the missRanger R package [31].

The genomic map used to find the location of significant SNPs is the map corresponding to the DM v6.1 assembly [32]. The SNPs were mapped on the twelve chromosomes of potato, and a group of SNPs that had an unassigned position on the genome.

### Models for Variance Component Estimation

Different univariate models were evaluated to find the best fitting model.

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Za} + \mathbf{e} \quad (1)$$

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Za} + \mathbf{Wp} + \mathbf{e} \quad (2)$$

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Za} + \mathbf{Egxd} + \mathbf{e} \quad (3)$$

Where **X**, **Z**, **W** and **E** are the incidence matrices for fixed effects, the additive genetic effect, the non-additive genetic effect and the genotype by decade of testing interaction effect, respectively. The vector **b** of the fixed effects included market segment, genotyping batch, and year of the challenge test. The vectors **a**, **p**, **gxd** and **e** are vectors of the additive genetic effect, the non-additive genetic effect, the genotype by decade interaction effect and the residual environmental effect, respectively. For the interaction, the decades were built on the year of the challenge test and so we had five decades; 1980s, 1990s, 2000s, 2010s and 2020s. The effect of **p** was left out of model (3) because the corresponding variance component converged toward zero. The distributional assumptions were  $\mathbf{a} \sim N[0, \mathbf{A}\sigma_a^2]$ ,  $\mathbf{p} \sim N[0, \mathbf{I}\sigma_p^2]$ , and  $\mathbf{gxd} \sim N[0, \mathbf{I}\sigma_{gxd}^2]$ . Where **A** was the pedigree-

based relationship matrix for all clones, and  $\mathbf{I}$  is an identity matrix of size corresponding to the number of individuals in  $\mathbf{p}$  and the number of genotype by decade interaction effects for  $\mathbf{gxd}$ . The pedigree-based relationship matrix was computed using the “AGHmatrix” package in R [33] assuming a ploidy of four and a probability of occurrence of double reductions of 0.005.

A multivariate (MT) model was used to estimate the (co)variance components for the three decades, for which more than 100 observations were available. In this case, there are three phenotypes,  $\mathbf{LB}_{00}$ ,  $\mathbf{LB}_{10}$  and  $\mathbf{LB}_{20}$  which are the phenotypes of the resistance to LB for tests done in 2000s, 2010s and 2020s, respectively. Only the phenotypes of those decades were kept because the few observations available for the early decades lead to convergence issues.

$$\begin{bmatrix} \mathbf{LB}_{00} \\ \mathbf{LB}_{10} \\ \mathbf{LB}_{20} \end{bmatrix} = \begin{bmatrix} \mathbf{X}_{00} & 0 & 0 \\ 0 & \mathbf{X}_{10} & 0 \\ 0 & 0 & \mathbf{X}_{20} \end{bmatrix} \begin{bmatrix} \mathbf{b}_{00} \\ \mathbf{b}_{10} \\ \mathbf{b}_{20} \end{bmatrix} + \begin{bmatrix} \mathbf{Z}_{00} & 0 & 0 \\ 0 & \mathbf{Z}_{10} & 0 \\ 0 & 0 & \mathbf{Z}_{20} \end{bmatrix} \begin{bmatrix} \mathbf{a}_{00} \\ \mathbf{a}_{10} \\ \mathbf{a}_{20} \end{bmatrix} + \mathbf{e} \quad (4)$$

Where  $\mathbf{X}_{00}$ ,  $\mathbf{X}_{10}$ ,  $\mathbf{X}_{20}$ ,  $\mathbf{Z}_{00}$ ,  $\mathbf{Z}_{10}$  and  $\mathbf{Z}_{20}$  are the incidence matrices for the fixed effect and the additive genetic effect.  $\mathbf{b}$  is the vector of the fixed effect, which are market segment, the genotyping batch, and the year of the challenge test. The vectors  $\mathbf{a}_{xx}$ , and  $\mathbf{e}$  were the additive genetic effect and the residual environmental effect, respectively,

$$\text{with } \begin{bmatrix} \mathbf{a}_{00} \\ \mathbf{a}_{10} \\ \mathbf{a}_{20} \end{bmatrix} \sim \text{MVN} \left[ \mathbf{0}, \mathbf{A} \otimes \begin{pmatrix} \sigma_{a_{00}}^2 & \sigma_{a_{00}a_{10}} & \sigma_{a_{00}a_{20}} \\ \sigma_{a_{00}a_{10}} & \sigma_{a_{10}}^2 & \sigma_{a_{10}a_{20}} \\ \sigma_{a_{00}a_{20}} & \sigma_{a_{10}a_{20}} & \sigma_{a_{20}}^2 \end{pmatrix} \right],$$

The (co)variance components in all models were estimated by REML using the AI-REML procedure in the DMU software package [34].

## Heritability Estimates

Narrow sense heritability ( $h^2$ ) estimates were computed, as the estimated additive genetic variance divided by the estimated phenotypic variance and presented in tables as relative variance due to additive genetic effects. Similarly, the proportion of non-additive genetic variance was presented as the relative genetic variance due to non-additive effects. The broad sense heritability ( $H^2$ ) is then the sum of the relative variance due to additive and due to non-additive genetic effects. The estimated phenotypic variance was computed as the sum of all the estimated variance components for each model. Therefore, both narrow sense and broad sense heritability refer to an individual measurement from a plot consisting of three plants and do not refer to the heritability of variety means because the amount of replication varies.

## Genome Wide Association Study (GWAS)

The effect of each SNP (from 1 to 9524) was estimated using single marker regression using a SNP-by-SNP approach (step 1). If at least one significant SNP was found on the genome, the genotype of the most

significant SNP, referred to as the Top-SNP, was added to the model as a fixed effect (step 2), and GWAS were run again.

$$\text{step 1} \Rightarrow \mathbf{y} = \mathbf{Xb} + \beta^* \mathbf{SNP}_i + \mathbf{Za} + \mathbf{Egxd} + \mathbf{e} \quad (5)$$

$$\text{step 2} \Rightarrow \mathbf{y} = \mathbf{Xb} + \beta^* \mathbf{Top-SNP} + \beta^* \mathbf{SNP}_i + \mathbf{Za} + \mathbf{Egxd} + \mathbf{e} \quad (6)$$

The first step of model 5 is like model 3. The only difference is the addition of a regression on an individual SNP.  $\beta$  is the effect of  $\mathbf{SNP}_i$ . In step 2, the genotype of the most significant SNP identified across the genome in the first step was added as a fixed effect to the model (**Top-SNP**). This process was repeated until no more significant SNP were identified. The significance threshold was set at  $0.05/9524 = 5.25 \times 10^{-6}$  based on the Bonferroni correction.

### Additive Genetic Variance Explained by Significant Chromosome Regions

A QTL with significant effect on LB resistance was defined as a SNP or a group of SNPs capturing the same effects on LB. The group of SNPs were defined using model (6) in the two-step procedure of the GWAS. Each group of SNPs were represented by only the most significant SNP, referred as the Top-SNP. Model (5) was rerun with all Top-SNP included as fixed effects to estimate the remaining additive genetic variance. The relative amount of additive genetic variance explained by the chromosome region were then expressed as the reduction in additive genetic variance from model (6) compared with the estimate obtained from model (3). All reductions were expressed as percent reduction from additive genetic variance estimated in model (3). A similar procedure was used to estimate the amount of additive genetic variance explained by all significant SNPs across the genome

### Link between the Genotype at the Top-SNP and the Phenotype

To show the effect of the Top-SNP on LB resistance, we collected the genotypes of all individuals for each of Top-SNP identified. The phenotypic distribution for each genotype was then visualized in a box plot with subplots for each number of the alternative allele for each QTL identified.

### Prediction of Breeding Values

Four different breeding values were predicted for each clone. EBV-PBLUP is the breeding value predicted using model (3) where a pedigree-based relationship matrix was used. EBV-GBLUP is the breeding value estimated using GBLUP. In this case the model used model (3), but a genomic relationship matrix was used instead of a pedigree relationship matrix. The genomic relationship matrix was built using the “AGHmatrix” package in R as well [33] accounting for ploidy. The variance components used to compute the EBV-PBLUP and the EBV-GBLUP were the ones estimated using model (3). To evaluate the advantage of including QTL for

prediction, breeding values were also estimated using an extended model (6) where all the Top-SNP identified were added as a fixed regression to the model. This extended model was used both with the pedigree and the genomic relationship matrix, leading to EBV-QTL-PBLUP and EBV-QTL-GBLUP.

### Cross Validation of Models

To assess accuracy of the four different predicted breeding values, a five-fold cross-validation strategy was used. The population was randomly divided into five folds, so that each clone was present in one and only one fold. When a fold was used as the validation population, all the phenotypes corresponding to the clones of this fold were masked. The breeding values of the clones of the given fold used as validation population were then predicted using the phenotypes from the other four folds. We will refer to this breeding value as  $EBV_{reduced}$ . Since four different models were tested, there were four different  $EBV_{reduced}$  ( $EBV-PBLUP_{reduced}$ ,  $EBV-GBLUP_{reduced}$ ,  $EBV-QTL-BLUP_{reduced}$  and  $EBV-QTL-GBLUP_{reduced}$ ) predicted for each fold. All folds in turn were used as validation population so that in the end all clones had breeding values predicted based on information from the other folds. Breeding values were also estimated using the whole population and we will refer to those breeding values as  $EBV-PBLUP_{full}$ ,  $EBV-GBLUP_{full}$ ,  $EBV-QTL-BLUP_{full}$  and  $EBV-QTL-GBLUP_{full}$ .

The accuracy of prediction was computed as the correlation between the  $EBV_{reduced}$  and (1) the average phenotype for each clone corrected for the fixed effect ( $y_c$ ) and (2) the  $EBV_{full}$ . The dispersion of the breeding values was computed as the slope of the regression of  $y_c$  or  $EBV_{full}$  on  $EBV_{reduced}$ . The correlations were computed in a multivariate analysis of variance approach using a model including year of crossing to account for potential genetic trends.

The correlations and regressions were computed for each fold and the results presented are the average values across the five folds. The standard errors have been computed as the standard deviation divided by the square root of the number of folds, hence five.

## RESULTS

### Data Description

The number of observations recorded, the number of lines assessed, the average value, and standard deviation of the observations are given in Table 1 for each decade of testing and for the whole period (All years). Phenotypes on LB resistance were recorded in (parts of) five successive decades. In the 1980s and 1990s few records were collected on a small number of varieties with a considerable increase in the amount to testing in recent years. The average number of replicated plots per variety was 2.9 and on average each clone was tested in 1.9 years. The grade of

resistance for the observation done in the 1980s ranged from 2 to 5, which explains the small standard deviation compared to the other decades.

### Variance Components and Population Genetic Parameters

The estimated variance components from models (1), (2), and (3) are presented in Table 2. In model (2), the variance explained by the non-additive genetic effect (**p**) was not significantly different from 0. In the model 3, a large genotype per decade interaction (**gxd**) effect was observed. Adding the interaction in the model (going from model (1) to model (3)), led to a decrease in the estimated additive genetic variance and most importantly to a decrease in the residual variance. Indeed, for model (3), almost all the phenotypic variance is explained by the additive genetic effect and the **gxd** effect. Moreover, the  $-2\text{Log}(L)$  was very much smaller for the model including the genotype by decade interaction effect than for model (1) where the interaction effect was omitted. Model (3) was initially run including the non-additive genetic effect (**p**), but the corresponding variance component converged toward zero and was, therefore, removed from the model. The relative variance components for the additive genetic effect presented in Table 2 corresponds to narrow sense heritability for a record on a single plot with three plants. The sum of relative contributions for additive genetic variance and non-additive genetic variance in model (2) corresponds to an estimate of broad sense heritability, again related to an individual three plant plot.

**Table 2.** Variance components estimated using three different models \*.

	Model a_e (1)		Model a_p_e (2)			Model a_gxd_e (3)		
	a	e	a	p	e	a	gxd	e
Estimates	3.36	0.86	3.09	0.18	0.86	2.56	0.86	0.64
SE	0.17	0.03	0.25	0.13	0.02	0.19	0.09	0.02
Relative	0.80	0.20	0.75	0.04	0.21	0.63	0.21	0.16
SE	0.01	0.01	0.04	0.03	0.01	0.03	0.02	0.01
$-2\text{Log}(L)$	5582		5578			5311		

\* a is the additive genetic variance, p is the non-additive genetic effect, gxd is the genotype by decade effect and e is the residual variance. Relative is the part of total phenotypic variance explained.  $-2\text{Log}(L)$  is the  $-2\log$  of the likelihood.

The estimated variance components for LB resistance assessed in different decades from model (4) are presented in Table 3. The phenotypes of the 1980s and 1990s were not included because only few records were available. The estimates of the additive variance and of the residual variance are different across the decades showing a clear decline from first to last decade. The lowest heritability estimate was observed for LB10, corresponding to the resistance challenge done in the 2010s. In contrast, the heritability estimates for LB00 and LB20 were both at 0.86.

**Table 3.** Variance components and SE estimated in each decade of challenge \*.

	LB00		LB10		LB20	
	a	e	a	e	a	e
Estimates	5.87	0.97	3.28	1.45	3.01	0.50
SE	0.94	0.12	0.31	0.13	0.16	0.02
Relative	0.86	0.14	0.69	0.31	0.86	0.14
SE	0.03	0.03	0.03	0.03	0.01	0.01

\* LB00, LB10, LB20 corresponds to the challenges to LB done in the 2000s, 2010s and 2020s respectively. a is the additive genetic variance and e is the residual variance. Relative is the part of total phenotypic variance explained.

The genetic correlations between the resistance traits measured across different decades of challenge are presented in Table 4. The correlation ranged from 0.57 to 0.91, with LB00 and LB10 showing the highest correlation. While LB10 and LB20 also showed a high correlation, they may still be considered as distinct traits. The genetic correlation between LB00 and LB20 appeared relatively small compared to the other pairs. This observation corroborates the presence of significant genotype-by-decade interaction.

**Table 4.** Genetic correlations between LB-resistance from different decades of challenge \*.

	LB00	LB10	LB20
LB00	1.00	-	-
LB10	0.91 (0.06)	1.00	-
LB20	0.57 (0.11)	0.75 (0.04)	1.00

\* LB00, LB10, LB20 corresponds to the challenges to LB done in the 2000s, 2010s and 2020s respectively.

### GWAS

Three chromosome regions significantly associated with LB resistance were found and results are summarized in Table 5. The genome regions with significant effects on LB resistance are located on chromosomes 5, 9 and 1.

**Table 5.** QTLs significantly associated with resistance to LB \*.

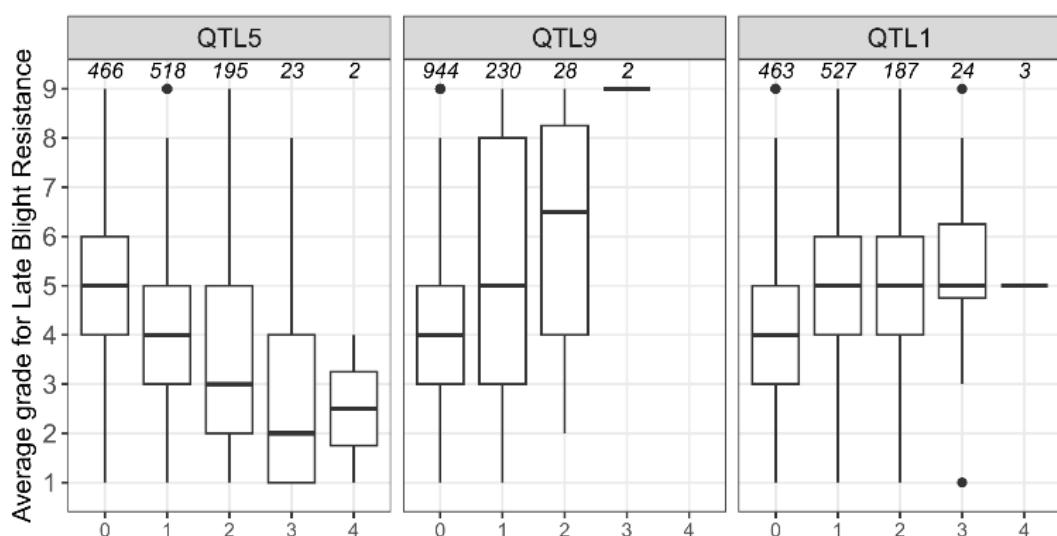
QTL	Chr	Position Top-SNP (bp)	Genetic Variance Explained by Top-SNP (%)	-Log(P)	Top-SNP	MAF
QTL5	5	4726087	5.96	12.84		0.22
QTL9	9	62365725	4.87	8.40		0.06
QTL1	1	68451137	2.34	5.99		0.19

\* Chr is the chromosome, Position Top-SNP is the position of the most significant SNP in the QTL, Genetic variance explained by Top-SNP is the additive genetic variance explained by the Top-SNP, Log(P) Top-SNP is the negative log of the *p*-value of the most significant SNP in the QTL and MAF is the minor allele frequency in our population for the Top-SNP.

TOP-SNP are defined here as the most significant SNP in a genomic region defined as a QTL. We determined that there were 3 chromosome regions with significant effects on LB resistance so there were three TOP SNP, one on chromosome 5, one on chromosome 9, and another SNP on chromosome 1. Even though the QTL on chromosome 5 is known to be linked with maturity [26] we also looked at its effect on LB. Each of the QTL

explained 5.96, 4.87 and 2.34 percent of the total additive genetic variance in LB-resistance, respectively.

The effects of different number of alleles in each of the TOP-SNPs are shown in Figure 1. In the population studied, the number of copies of the minor allele of Top SNP of QTL5 on chromosome 5 confers a negative impact on resistance to LB with a clear additive dosage effect. For Top SNP of QTL9 an increased number of copies of the minor allele leads to increased resistance to LB which also seems to be a dosage effect where having one more of the alternate allele confer higher resistance to LB. The alternative allele of Top SNP of QTL1 seems to confer advantages for resistance to LB, where all heterozygous clones have an increased resistance to LB and the gene, therefore, appear to be dominant. No clones were homozygous for the alternative allele. However, it is complicated to draw conclusion on the effects of the number of alleles on the raw phenotype because of the low frequencies of the genotypes with at least three copies of the minor allele. In this population, the frequency of the beneficial allele was 0.78 for the Top-SNP of QTL5, 0.06 for the Top-SNP of QTL9 and 0.19 for the Top-SNP of QTL1.



**Figure 1.** Boxplot of phenotype at identified QTL \*; \* On the x axis are the genotypes (number of alleles of the minor allele) at the TOP-SNP and on the y axis is the average phenotypic value for all individual genotypes. On the top of each box is the number of clones with this specific genotype.

### Prediction of Breeding Values

The results of cross validation of models for prediction of breeding values are presented in Table 6. The prediction accuracies evaluated as the correlation between  $y_c$  and  $EBV_{reduced}$  ranged from 0.56 to 0.75. Accuracies were similar when using PBLUP, GBLUP or QTL-PBLUP but were markedly higher for the QTL-GBLUP model. The inflation of the distribution of predicted breeding values ranged from 0.91 to 1.32. Overall, for all the models the breeding values were deflated except for the GBLUP where they were inflated.

The correlation between  $EBV_{full}$  and  $EBV_{reduced}$  ranged from 0.70 to 0.90. Accuracies were similar when using PBLUP or QTL-PBLUP but were higher for the models where genomic information was included and especially for the QTL-GBLUP model. The regression indicating dispersion of the predicted breeding values ranged from 0.96 to 1.18.

When associated with the genomic relationship matrix, adding Top-  
SNP of the three QTL identified as fixed regression into the model led to a considerable gain in accuracy. While no gain was observed when PBLUP and QTL-PBLUP model were compared.

**Table 6.** Accuracies and inflations of the breeding values predicted using four different models \*.

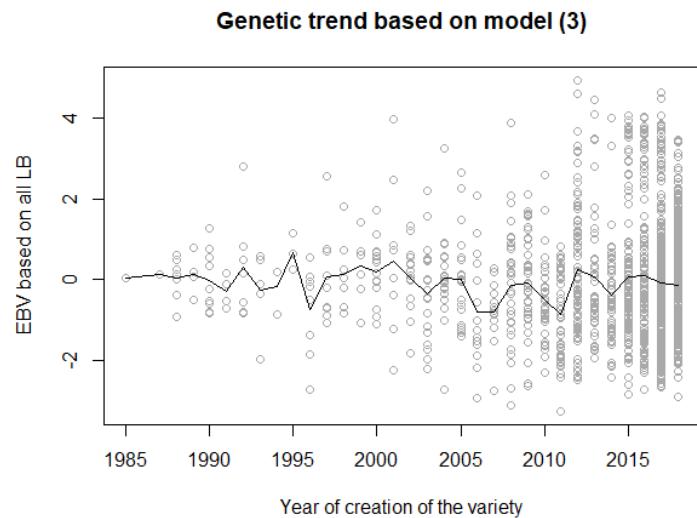
		PBLUP	GBLUP	QTL-PBLUP	QTL-GBLUP
Correlation with $EBV_{reduced}$	$y_c$	0.62 (0.10)	0.58 (0.08)	0.56 (0.11)	0.75 (0.08)
	$EBV_{full}$	0.74 (0.04)	0.82 (0.05)	0.70 (0.09)	0.90 (0.05)
Inflation (Slope of Regression on $EBV_{reduced}$ )	$y_c$	1.28 (0.09)	0.91 (0.04)	1.32 (0.13)	1.25 (0.03)
	$EBV_{full}$	1.18 (0.06)	0.96 (0.02)	1.15 (0.11)	0.99 (0.03)

\*  $y_c$  is the phenotype corrected with the fixed effect and  $ebv_{full}$  is the breeding value estimated with the full dataset. Both were correlated with  $ebv_{reduced}$  which is the breeding value estimated while the phenotype of the validation population was removed. the inflation was computed as the slope of the regression of  $y_c$  or  $ebv_{full}$  on the  $ebv_{reduced}$ .

### Breeding Values and Genetic Trend

Predicted breeding values from model (3), where a genotype by decade interaction effect has been included in the model were used in investigating the genetic trend for LB resistance. Only a few varieties created in the early years of the breeding program had been tested for LB resistance. It is also important to note that some of the varieties presented were included in the test in several years. Indeed, the varieties used as spreader rows have been used across many years, like for example the Bintje variety which has been used over all the 36 years. Moreover, many varieties have related varieties which have been tested as well, leading to more accurate breeding values for these clones.

The estimated genetic trends for LB resistance are shown in Figure 2. No clear trend in the breeding values was observed. It does not seem like there has been any efficient selection for LB resistance as the mean did not change significantly. However, there is an increase in variation of the breeding values with time and we see an increase in number of more resistant (fully or not) varieties from the 2000s with an acceleration from 2010.



**Figure 2.** Genetic trend for LB resistance \*. \* Points are the breeding values estimated using model 3 for each of the varieties. Line is the average breeding value per year of first appearance in the breeding program.

## DISCUSSION

The comparison of the univariate models gave very strong support for model (3), which includes a genotype by decade interaction that accounts for 21 percent of all phenotypic variance in LB resistance. This significant contribution of genotype-by-time interaction to LB resistance variance has been previously reported [24]. The literature reports a wide range of heritability estimates for LB resistance from 0.31 to 0.85 [22–25,35]. Several factors could explain this variability, such as differences in test procedures and trait definition, in the plant material tested, in the pathogens used for inoculation, or in the methodology used for estimating variance components. In this study the estimated narrow sense heritability defined at the single plot level is very high (0.63) providing a solid basis for selection for resistance towards *Pi* based on general polygenic breeding values. However, the considerable amount of genotype by decade interaction suggests that resistance could fluctuate over time, potentially due to evolution in the disease agent or due to selective breeding in the tested varieties. A stronger focus a specific pool of parents may also have contributed to the decrease in genetic variance over time. Similar results were found by [24], who found significant genotype-by-year interaction even when experiments were conducted over a much shorter time span than our study. However, yearly effects tend to be strongly affected by differences in weather from year to year, an effect that is expected to be limited when decades are considered. LB resistance measured across different decades shows very high heritability within each decade, but the genetic correlations between resistance in different decades are significantly lower than unity, particularly between the 2020s and earlier decades. This likely reflects a change in the *Pi* pathogen, rather

than the varieties themselves, since limited selection for LB resistance appears to have occurred in the population. In our study, data collection spanned from the 1980s to 2023, with at least seven different *Pi* pathotypes identified in Denmark between 2006 and 2022 [28]. This means that resistance may reduce or break down due to evolution in *Pi*. However, the model used to predict breeding values includes a genotype by decade interaction and this provides breeding values averaged over all the decades/years involved and this is expected to provide a more general resistance that is not easily circumvented by the *Pi*. The estimate of additive genetic variance and especially the estimate of residual variance reduced over the decades investigated. Genetic and environmental variances decreased over time and genetic correlations between distant decades were significantly lower than unity. Again, this indicates a change in the *Pi* genotypes evolved as there are strong genetic links between clones tested in different decades. However, genetic correlations were sufficient to ensure lasting genetic response for LB resistance. One of the limitations in our study was the scarce amount of data registered in the early years of the breeding program. Acquiring more data per year could help to dissect the genotype by decade interaction, as several factors could have contributed to the interaction. Furthermore, in future testing, a thorough characterization of *Pi* genotypes involved would be of considerable interest, potentially combined with specific testing of resistance toward well characterized *Pi* genotypes.

The absence of genetic trend (Figure 2) implies that selection for LB resistance has been limited during the period studied. As previously discussed, this might result from sequential phenotypic selection and/or because of the use of old material as parents in later years. Resistance development has relied heavily on R-genes, which has been utilized since the introgression of R-genes identified in wild potatoes from South America. However, the R-genes are not always a source of long-lasting resistance [9,36]. For example, the pathotype EU\_41\_A2 acquired the ability to overcome the resistance coming from several R-genes [36]. One proposed strategy has been to accumulate more than one R-genes in a single variety, which is the pyramiding of R-genes approach [14]. In theory, if a single clone carried the beneficial allele for four R-genes, it would become virtually fully resistant to the pathogen targeted [18]. However, this implies for that all the R-genes are 100% effective, which is likely unrealistic [18]. In this study only a few genomic regions affecting LB resistance were identified. This could be because only few resistance genes were present in the population or because they are not effective in the population studied. The latter reason seems most likely as it takes many generations to make a tetraploid population homozygous for a resistance allele.

The three QTL identified in this study explained all together 13% of the additive genetic variance for LB resistance. It means that most of the additive genetic variance in LB resistance is due to the polygenic background, i.e., genes with small effects that cannot be individually identified in our data.

One challenge in selection for LB resistance, is that the loci having a positive effect for resistance may lead to a negative response for other traits of interest. This could explain the low allele frequency of the positive allele for LB resistance for QTL9 and QTL1. For example, the pleiotropy associated with a region on chromosome 5, which affects both maturity and LB resistance, is well-documented [26]. To address this, multi-trait evaluation models should be developed to estimate and exploit genetic correlations between traits, and appropriate economic weights should be assigned to account for potential unfavorable correlations between traits of economic importance. Cloning of the genes involved would also aid in the management of LB resistance.

The accuracy of breeding values predicted from pedigree information or from genomic information was very high. This is not surprising due to the high heritability of LB resistance observed and because most varieties will have many relatives that were also tested. Because of this resemblance between the training and the validating population, a high prediction accuracy is expected. The advantage of using a GBLUP model is relatively limited because of the high heritability of LB resistance and because of the cross-validation strategy which also yielded high accuracies with PBLUP. However, in forward prediction of breeding values for new crosses with few or no close relatives tested a greater advantage of GBLUP model is expected. Moreover, we had access to a deep pedigree allowing accurately tracking the history of the population. The correlation between  $EBV_{full}$  and  $EBV_{reduced}$  indicates the relative change in accuracy when a clone is having its own phenotype removed. The drop in accuracy is relatively small especially for the GBLUP model. This means that if a line is genotyped it is possible to accurately predict its LB resistance before testing this line for LB resistance. Enciso-Rodriguez et al (2018) [24] has drawn the same conclusion from their study. This situation provides excellent possibilities for early selection for LB resistance based on such a combined model. Such a model would lead to a stratified testing of new (and older) clones to update the training population and the use of intense selection among all genotyped individuals. Moreover, the QTL-GBLUP model show that while having own phenotype, genomic information does not improve accuracy of prediction for LB resistance, but incorporating extra information on the genetic architecture of LB resistance could be highly relevant. These results must be interpreted with some caution as the QTL were identified on the entire population, which is also the training population, and this may tend to provide too optimistic results, even though the effects were estimated in the training population only. Even though clear genotype-by-year interaction was observed, selection for LB resistance is still expected

to have a lasting effect but continuous testing for LB resistance is needed to identify lack of resistance toward new *Pi* strains.

## CONCLUSIONS

The narrow sense heritability of resistance to LB in potato was large and this shows that results of resistance tests can be used for efficient selection for LB resistance. The resistance changed over time as reflected in a significant interaction between genotype and decade of testing, an interaction that most likely is due to evolution in the *Pi*. However, the genetic correlation between resistance in different time periods is still substantial so that selection effects will be persistent but potentially will reduce over time. Consistent testing of new lines for resistance is therefore needed. A GWAS study identified three genomic regions that affected resistance toward LB. In total these genomic regions explained 13% of the additive genetic variance in LB resistance. I.e., most of the genetic variance in LB resistance is polygenic and obviously all genetic variation should be used in selection for LB resistance. Extending models so that both identified genomic regions and polygenic variation were used in predicting breeding values clearly yielded the most accurate selection criteria.

## DATA AVAILABILITY

The dataset of the study is available from the authors upon reasonable request and will include a confidentiality agreement because the data belongs to an active breeding population.

## AUTHOR CONTRIBUTIONS

Conceptualization, HR and JJ; methodology, HR and JJ; formal analysis, HR; data curation, HR; writing—original draft preparation, HR; writing—review and editing, HR, MB and JJ; project administration, HR; funding acquisition, JJ and MB. All authors have read and agreed to the published version of the manuscript.

## CONFLICTS OF INTEREST

MB are employed by Danespo A/S, but all authors declare that they have no conflicts of interest.

## FUNDING

The project RESBOT was funded by Ministry of Food, Agriculture and Fisheries of Denmark under the Green Development and Demonstrations Program (grant no. 34009-20-1643).

## REFERENCES

1. Zadoks JC. The potato murrain on the European continent and the revolutions of 1848. *Potato Res.* 2008;51(1):5-45.
2. Haverkort AJ, Struik PC, Visser RGF, Jacobsen E. Applied biotechnology to combat late blight in potato caused by *Phytophthora infestans*. *Potato Res.* 2009;52(3):249-64.
3. Haverkort AJ, Boonekamp PM, Hutten R, Jacobsen E, Lotz LAP, Kessel GJT, et al. Durable late blight resistance in potato through dynamic varieties obtained by cisgenesis: Scientific and societal advances in the DuRPh project. *Potato Res.* 2016;59(1):35-66.
4. Abuley IK, Lynott JS, Hansen JG, Cooke DEL, Lees AK. The EU43 genotype of *Phytophthora infestans* displays resistance to mandipropamid. *Plant Pathol.* 2023;72(7):1305-13.
5. Díaz-García G, Enciso-Maldonado GA, Lozoya-Saldaña H. *Solanum demissum* Lindl. in potato breeding. *Revista Chapingo Ser Horticultura.* 2023;29(3):131-48.
6. Flor HH. Current status of the gene-for-gene concept. *Annu Rev Phytopathol.* 1971;9(1):275-96.
7. Tan MA. Genetic mapping and pyramiding of resistance genes in potato. Wageningen (The Netherlands): Wageningen University and Research; 2008.
8. Paluchowska P, Śliwka J, Yin Z. Late blight resistance genes in potato breeding. *Planta.* 2022;255(6):127.
9. Malcolmson JF. Races of *Phytophthora infestans* occurring in Great Britain. *Trans Brit Mycol Soc.* 1969;53(3):417-23.
10. Lokossou AA, Park T-h, van Arkel G, Arens M, Ruyter-Spira C, Morales J, et al. Exploiting knowledge of R/AVR genes to rapidly clone a new LZ-NBS-LRR family of late blight resistance genes from potato linkage group IV. *Mol Plant Microbe Interact.* 2009;22(6):630-41.
11. Yang L, Wang D, Xu Y, Zhao H, Wang L, Cao X, et al. A new resistance gene against potato late blight originating from *Solanum pinnatisectum* located on potato chromosome 7. *Front Plant Sci.* 2017;8:1729.
12. El-Kharbotly A, Palomino-Sánchez C, Salamini F, Jacobsen E, Gebhardt C. R6 and R7 alleles of potato conferring race-specific resistance to *Phytophthora infestans* (Mont.) de Bary identified genetic loci clustering with the R3 locus on chromosome XI. *Theor Appl Genet.* 1996;92(7):880-4.
13. Huang Y, Zhang C, Zhang W, Zhang P, Kang X, Chen W. Variation in the chicken LPIN2 gene and association with performance traits. *Br Poult Sci.* 2015;56(2):175-83.
14. Nelson RR. Stabilizing racial populations of plant pathogens by use of resistance genes. *J Environ Qual.* 1972;1(3):220-7.
15. Rakosy-Tican E, Thieme R, König J, Nachtigall M, Hammann T, Denes T-E, et al. Introgression of two broad-spectrum late blight resistance genes, Rpi-Blb1 and Rpi-Blb3, from *Solanum bulbocastanum* Dun plus race-specific R genes into potato pre-breeding lines. *Front Plant Sci.* 2020;11:699.

16. Su Y, Viquez-Zamora M, den Uil D, Sinnige J, Kruyt H, Vossen J, et al. Introgression of genes for resistance against *Phytophthora infestans* in diploid potato. *Am J Potato Res.* 2020;97(1):33-42.
17. Ghislain M, Byarugaba AA, Magembe E, Njoroge A, Rivera C, Román ML, et al. Stacking three late blight resistance genes from wild species directly into African highland potato varieties confers complete field resistance to local blight races. *Plant Biotechnol J.* 2019;17(6):1119-29.
18. Stam R, McDonald BA. When resistance gene pyramids are not durable—the role of pathogen diversity. *Mol Plant Pathol.* 2018;19(3):521-4.
19. McGee R, Zaleski-Cox M, Jayawardana MA, Tillman BL, Wally O, Esquivel-Garcia L, et al. Breeding for quantitative disease resistance: Case studies, emerging approaches, and exploiting pathogen variation. *Crop Sci.* 2025;65(6):e70202.
20. Center IP. CIP 2030 Strategy. 2024. Available from: <https://hdl.handle.net/10568/151968>. Accessed on 19 Nov 2025.
21. Sood S, Bhardwaj V, Kaushik S, Sharma S. Prediction based on estimated breeding values using genealogy for tuber yield and late blight resistance in auto-tetraploid potato (*Solanum tuberosum* L.). *Helion.* 2020;6(11):e05624.
22. Pajerowska-Mukhtar K, Stich B, Achenbach U, Ballvora A, Lübeck J, Strahwald J, et al. Single nucleotide polymorphisms in the allene oxide synthase 2 gene are associated with field resistance to late blight in populations of tetraploid potato cultivars. *Genetics.* 2009;181(3):1115-27.
23. Haynes KG, Christ BJ. Heritability of resistance to foliar late blight in a diploid hybrid potato population of *Solanum phureja* × *Solanum stenotomum*. *Plant Breed.* 1999;118(5):431-4.
24. Enciso-Rodriguez F, Douches D, Lopez-Cruz M, Coombs J, de los Campos G. Genomic selection for late blight and common scab resistance in tetraploid potato (*Solanum tuberosum*). *G3 Genes Genomes Genet.* 2018;8(7):2471.
25. Solano J, Acuña I, Esnault F, Brabant P. Resistance to *Phytophthora infestans* in *Solanum tuberosum* landraces in Southern Chile. *Trop Plant Pathol.* 2014;39:307-15.
26. Visker M, Keizer L, Van Eck H, Jacobsen E, Colon L, Struik P. Can the QTL for late blight resistance on potato chromosome 5 be attributed to foliage maturity type? *Theor Appl Genet.* 2003;106(2):317-25.
27. Xue W, Haynes KG, Qu X. Characterization of early blight resistance in potato cultivars. *Plant Dis.* 2019;103(4):629-37.
28. About Euroblight [Internet]. 2023. Available from: <https://agro.au.dk/forskning/internationale-platforme/euroblight/about-euroblight>. Accessed on 20 July 2023.
29. Bach HJ. Vejledning i bedømmelser i landsforsøgene 2022. Available from: [https://www.landbrugsinfodk-/media/landbrugsinfo/public/a/8/e/planter\\_kval\\_landsfors\\_regi\\_bedoemmelser-i-landsforsøegenepdf](https://www.landbrugsinfodk-/media/landbrugsinfo/public/a/8/e/planter_kval_landsfors_regi_bedoemmelser-i-landsforsøegenepdf). Accessed on 22. Dec 2025.
30. Tin Kam H. The random subspace method for constructing decision forests. *IEEE Trans Pattern Anal Mach Intell.* 1998;20(8):832-44.

31. Mayer M, Mayer MM. Package ‘missRanger’. R package. 2019. Available from: <https://cran.r-project.org/web/packages/missRanger/index.html>. Accessed on 15 July 2023.
32. Pham GM, Hamilton JP, Wood JC, Burke JT, Zhao H, Vaillancourt B, et al. Construction of a chromosome-scale long-read reference genome assembly for potato. *GigaScience*. 2020;9(9):giaa100.
33. Amadeu RR, Cellon C, Olmstead JW, Garcia AAF, Resende MFR Jr., Muñoz PR. AGHmatrix: R package to construct relationship matrices for autotetraploid and diploid species: A blueberry example. *Plant Genome*. 2016;9(3):plantgenome2016.01.0009.
34. Madsen P, Jensen J. A users guide to DMU. A package for analysing multivariate mixed models, Version 6, Release 5.0. 2010. Available from: [http://dmu.agrsci.dk/dmuv6\\_guide.5.0.pdf](http://dmu.agrsci.dk/dmuv6_guide.5.0.pdf). Accessed on 15 July 2023.
35. Sood S, Bhardwaj V, Bairwa A, Dalamu, Sharma S, Sharma AK, et al. Genome-wide association mapping and genomic prediction for late blight and potato cyst nematode resistance in potato (*Solanum tuberosum* L.). *Front Plant Sci*. 2023;14:1211472.
36. Puidet B, Mabon R, Guibert M, Kiiker R, Soonvald L, Le VH, et al. Examining phenotypic traits contributing to the spread in Northern European potato crops of EU\_41\_A2, a new clonal lineage of *Phytophthora infestans*. *Phytopathology*. 2022;112(2):414-21.

How to cite this article:

Romé H, Bagge M, Jensen J. The genetic basis of resistance to late blight in tetraploid potatoes. *Crop Breed Genet Genom*. 2026;8(1):e260001. <https://doi.org/10.20900/cbgg20260001>.