

Drone-imaging assisted genome-wide association studies reveal key quantitative trait loci for emergence and late blight resistance in tetraploid potato

Aalborg, Trine; Christiansen, Mathias Timmermann; Romé, Hélène; Chawade, Aakash; Bagge, Merethe; Nielsen, Kåre Lehmann

Published in:
Crop Science

DOI (link to publication from Publisher):
[10.1002/csc2.70261](https://doi.org/10.1002/csc2.70261)

Creative Commons License
CC BY 4.0

Publication date:
2026

Document Version
Publisher's PDF, also known as Version of record

[Link to publication from Aalborg University](#)

Citation for published version (APA):

Aalborg, T., Christiansen, M. T., Romé, H., Chawade, A., Bagge, M., & Nielsen, K. L. (2026). Drone-imaging assisted genome-wide association studies reveal key quantitative trait loci for emergence and late blight resistance in tetraploid potato. *Crop Science*, 66(2), Article e70261. <https://doi.org/10.1002/csc2.70261>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

Take down policy

If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

ORIGINAL ARTICLE

Crop Breeding & Genetics

Drone-imaging assisted genome-wide association studies reveal key quantitative trait loci for emergence and late blight resistance in tetraploid potato

Trine Aalborg¹  | **Mathias Timmermann Christiansen**² | **Hélène Romé**³ 
Aakash Chawade^{4,5}  | **Merethe Bagge**² | **Kåre Lehmann Nielsen**^{1,6} 

¹Department of Chemistry and Bioscience, Aalborg University, Aalborg, Denmark

²Danespo A/S, Give, Denmark

³Center for Quantitative Genetics and Genomics, Aarhus University, Aarhus, Denmark

⁴Department of Plant Breeding, Swedish University of Agricultural Sciences, Alnarp, Sweden

⁵Phenoyard AB, Lund, Sweden⁶KMC Amba, Brande, Denmark

Correspondence

Trine Aalborg, Department of Chemistry and Bioscience, Aalborg University, Aalborg, Denmark.

Email: traa@bio.aau.dk

Assigned to Associate Editor Laura Shannon.

Present address

Hélène Romé, SEGES Innovation P/S,
Aarhus, Denmark.

Funding information

NordGen, Grant/Award Number: project SustainPotato II; Danish Ministry of Food, Agriculture and Fisheries (GUDP), Grant/Award Number: project RESPOT (34009-20-1643)

Abstract

Solanum tuberosum L. (potato) is the world's most important vegetable crop, and developing improved cultivars is paramount for global food security. The efficacy of the genomic prediction models that accelerate breeding and genome-wide association studies (GWAS) depends on large, high-quality phenotypic datasets, which are often associated with inhibitory large costs for manual phenotyping. Recent advances in drone-based phenotyping offer an affordable, high-throughput phenotyping approach. This study evaluated the reliability of novel drone-based phenotyping for GWAS in tetraploid potato. Image-derived data on plant emergence and late blight resistance were collected using a drone-mounted red-green-blue camera across separate field trials over two to three growing seasons. The dataset comprised 6262 phenotypic observations of emergence and 6972 observations of late blight resistance. A total of 2576 clones in an emergence panel and 1520 clones in a late blight resistance panel were genotyped for 8731 markers using a custom single nucleotide polymorphism array. GWAS yielded five significant quantitative trait loci (QTLs) for late blight resistance—three coinciding with known *Rpi*-gene clusters and *StCDF1*—

Abbreviations: AUDPC, area under the disease progress curve; BLUE, best linear unbiased estimator; BSA, bulk-segregant analysis; CDF, cumulative distribution function; CI, confidence interval; GRM, genomic relationship matrix; GWAS, genome-wide association studies; LB, late blight resistance; LD, linkage disequilibrium; LOCO, leave-one-chromosome-out; M2EA, Mavic 2 Enterprise Advanced; MAF, minor allele frequency; MAS, marker-assisted selection; PC, principal component; PCA, principal component analysis; QTL, quantitative trait locus; RGB, red-green-blue; s.e., standard error; SNP, single nucleotide polymorphism; UAV, unmanned aerial vehicle; WEC, whole-exome capture.

This is an open access article under the terms of the [Creative Commons Attribution](#) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2026 The Author(s). *Crop Science* published by Wiley Periodicals LLC on behalf of *Crop Science* Society of America.

and four significant minor-effect QTLs underpinning emergence. Rediscovery of known late blight loci supports the reliability of the approach. Flanking regions of the emergence QTLs include candidate genes encoding proteins involved in plant growth and development, stress response, compound trafficking, and nutrient assimilation. However, 58%–76% of genes in the flanking regions lack functional annotation, underscoring the need for transcriptomic studies of early plant development. We demonstrate that integrating remote phenotyping and GWAS can yield robust results in tetraploid potatoes, overcoming the past limitation of phenotyping costs.

Plain Language Summary

The performance of genomics-based models applied in plant breeding relies on large, high-quality genotypic and phenotypic datasets. However, large-scale manual phenotyping is associated with inhibitory large costs. Recent advances in automated remote phenotyping based on image analysis offer accurate, high-throughput phenotyping. While the accuracy of remote phenotyping data has previously been benchmarked for several traits in potato, it has had limited prior application in genomics modeling. In this study, we integrate novel drone-based phenotypic data in genome-wide association studies (GWAS) of two traits in tetraploid potato to validate its application in genomics-based models. Our results demonstrate that remote phenotyping data enable reliable GWAS for late blight resistance and emergence in potato.

1 | INTRODUCTION

We are currently facing a global food crisis, with global food demand projected to increase with 56% by 2050 compared to 2010 (Van Dijk et al., 2021). Meeting this demand through intensified agricultural inputs is increasingly constrained by environmental and health considerations, including restrictions on fertilizer and pesticide use, a shortage of freshwater for irrigation, and a projected loss of arable farmlands (Sishodia et al., 2020). These pressures necessitate the rapid development of high-yielding, resource-efficient, and stress-resilient crop cultivars (Lenaerts et al., 2019). The crop potato, *Solanum tuberosum* L., is a key target in this context, as it is the world's most important non-cereal food crop and is a major staple crop across diverse agroecological regions (Devaux et al., 2014; FAOSTAT, 2024; Jennings et al., 2020).

Modern breeding approaches provide a powerful framework to accelerate genetic gain in tetraploid potato. Marker-assisted selection (MAS), informed by genome-wide association studies (GWAS), enables early selection of superior parent clones based on genetic markers linked to trait-associated loci (Collard & Mackill, 2008; Lenaerts et al., 2019; Yuan et al., 2024). Identification of causative allele variants affecting trait phenotype, through, for example, GWAS, can also be used to improve elite cultivars through genome editing

techniques (Carlsen et al., 2021, 2024; Johansen et al., 2019; Nadakuduti et al., 2018).

In GWAS, detection power is radically improved with larger collections of paired genotypic/phenotypic data (Uffelmann et al., 2021). The increased availability, automation, and substantial reduction in costs of genotyping that have followed in the wake of the genomics era mean that genotyping is no longer the bottleneck in the genomic modeling of agronomic plant traits. Rather, it is now (again) the cumbersome phenotyping that constitutes the limiting factor (Spindel et al., 2018). The rapid collection of genomic data of crop plants increases the demand for faster collection of trait phenotypic data across genotyped lines (Jeon et al., 2023; Yang & Zhai, 2022). Modern plant breeding is still largely dependent on traditional phenotyping methods, including visual grading in the field by specialists over the course of a growing season, plant measures, as well as destructive analytical measurements on a high number of harvested samples—this is time-consuming and expensive (Chang-Brahim et al., 2024; Lizarazo et al., 2023; Rodríguez et al., 2021; Spindel et al., 2018), and thus constitutes a limiting factor for the rapid development of novel, improved crop cultivars. Furthermore, manual measurements are subject to some degree of observer bias and noise, affecting measurement precision (Rodríguez et al., 2021; Sugiura et al., 2016).

Methods of high-throughput phenotyping have been in rapid development over the past decade (Jeon et al., 2023). The recent advancement in the implementation of remote phenotyping techniques offers a nondestructive, accurate, automated, and efficient high-throughput alternative to the traditional in-field manual phenotyping (Chang-Brahim et al., 2024). The use of remote sensing technologies in phenotyping has great potential for plant breeding and research applications, as it is possible to collect extensive amounts of visual data in a time-efficient manner and relatively inexpensively compared to traditional methods. It also digitalizes the phenotypic data, which were previously reliant on manual, and hence more error-prone, recording and transcription into usable formats (Cobb et al., 2019). The remote sensing methods include several categories based on the sensor platform (e.g., satellite, unmanned aerial vehicle [UAV], or ground-based instruments), and the type of sensor attached to the platform, such as red-green-blue (RGB), multispectral, hyperspectral, thermal, or light detection and ranging (LiDAR) (Abdullah et al., 2023). Compared with other remote sensing platforms, UAVs allow lower operational costs and less dependence on weather conditions than satellites, with increased throughput and data complexity compared to ground-based sensing platforms (Y. Shi et al., 2016). With recent advances in sensor technology, remote sensor platforms, and imaging technology, image-based remote sensing is being successfully implemented in field trials for the precise estimation of crop phenotypes (Chang-Brahim et al., 2024; Sishodia et al., 2020; Sun et al., 2022). Also in potatoes, a diverse range of remote sensing technologies has been applied in the assessment of disease incidence and phenotyping of agronomic traits, including late blight and bacterial wilt resistance, yield, canopy cover, and emergence (Duarte-Carvajalino et al., 2018; Franceschini et al., 2019; B. Li et al., 2019, 2020; Liu et al., 2022; Lizarazo et al., 2023; Njane et al., 2023; Rodríguez et al., 2021; Sankaran et al., 2017; Y. Shi et al., 2022; Sugiura et al., 2016).

Disease detection and monitoring are of particular importance for potatoes, since late blight, caused by the oomycete *Phytophthora infestans*, is one of the most destructive crop diseases in the world (Haas et al., 2009). While early and accurate detection is important for management by fungicide application (Duarte-Carvajalino et al., 2018; Y. Shi et al., 2022), accurate measures of disease resistance phenotypes will also be of particular importance for genomic breeding efforts, supporting both genomic selection models, selection during breeding, as well as association studies (Hickey et al., 2017). Disease symptoms mainly appear on the leaves, facilitating relatively facile detection from the plant canopy by remote sensing techniques (Sun et al., 2022). Late blight assessment in potatoes has been implemented with UAV-based remote sensing by RGB (Sugiura et al., 2016), multispectral (Duarte-Carvajalino et al., 2018; Rodríguez

Core Ideas

- Drone-imaging based phenotyping produced reliable results in genome-wide association studies (GWAS) of potato late blight resistance and emergence.
- Quantitative trait loci (QTLs) identified using remote phenotyping data aligned with known trait QTLs and trait-associated genes.
- Data structure and flight protocol variation highlight new challenges for integration of remote phenotyping data in GWAS.
- The integration of remote-imaging phenomics with genomics represents a methodological advance for potato breeding.

et al., 2021), and hyperspectral (Franceschini et al., 2019; Y. Shi et al., 2022) cameras. Quantification of late blight from electromagnetic spectral data relies on differentiating healthy and infected tissue and estimating the degree of damage. Leaf chlorophyll content can be used as a proxy since it is a general indicator of plant stress, including disease (Sishodia et al., 2020).

Plant emergence is a less explored physiological trait in potatoes, but it influences field management as well as cultivar evaluation (B. Li et al., 2019, 2020). With aggravating climate changes causing more frequent incidents of extreme weather events, it is becoming increasingly relevant, since early emergence of the canopy and root system reduces the need for weed control, and established plants are more resilient to early-season drought and flooding, and can promote field draining following floods. Additionally, the aboveground biomass is an important physiological indicator of crop nutritional status and overall growth (Liu et al., 2022). Total tuber yield is positively correlated with a long growth period, and canopy cover and size influence sunlight interception and hence overall photosynthetic rate (Njane et al., 2023). Rapid formation of the canopy also promotes the conservation of precipitation for rain-fed fields, and established root systems can improve irrigation efficiency. Canopy emergence (here measured as Day 0–20 of growth following sprout emergence from soil) and morphology are variable across cultivars (B. Li et al., 2020), and the identification of the markers underpinning the genetic variance of this complex trait, facilitated by accurate aerial phenotyping, may assist efficient breeding for this trait. Previous studies have assessed potato emergence by aboveground biomass or canopy cover using RGB (B. Li et al., 2019, 2020), multispectral (Njane et al., 2023; Sankaran et al., 2017), and hyperspectral (B. Li et al., 2020; Liu et al., 2022) imaging.

In this study, we present the results of GWAS of tetraploid potatoes using phenotypic data on plant emergence and late

blight resistance (LB) collected by an aerial drone using an RGB camera at distinct time points across the growing season. Separate field trials were established for emergence and LB phenotyping, using two distinct, but overlapping panels of commercial cultivars and elite breeding material from Danespo A/S for the trials of each trait. The emergence panel included 2576 genotyped lines, and the LB panel spanned 1520 genotyped lines. Using graded scores of plant emergence and LB computed by the commercial provider Phenoyard AB from RGB image data, we confirm three previously known quantitative trait loci (QTLs) for LB and identify two additional LB minor QTLs and four minor QTLs for emergence.

2 | MATERIALS AND METHODS

Statistical analyses were performed using R Statistical Software (v4.4.1) (R Core Team, 2024) in RStudio (v2024.04.2+764) (Posit Team, 2024). Graphics were generated using base R or the ggplot2 package (v3.5.1) in R (Wickham, 2016), unless otherwise stated.

2.1 | Plant material and field trial design

2.1.1 | Plant material

A population of commercial cultivars and elite breeding material, called the RESPOT population, was established at Danespo A/S in Give, Denmark. It was created as part of their breeding population and is based on commercial cultivars and advanced breeding material from the Danespo A/S breeding pool. It also included progeny of sexual hybridizations of the elite parental material. The genotyped proportion of the RESPOT population consisted of 5866 clones in total. Two subpopulations were evaluated in separate field trials for UAV remote phenotyping for LB and plant emergence, respectively. These panels will henceforth be referred to as the LB panel and the emergence panel. The LB panel consisted of 1520 genotyped clones that were analyzed in LB field trials in three consecutive years. The LB panel clones were chosen based on the possibility of qualitative resistance through pedigree. Only materials older than Year 3 in the breeding program were evaluated for LB. The emergence panel consisted of 2576 genotyped clones that were evaluated for plant emergence in two consecutive years. The emergence panel consisted of almost all generations in the breeding program (except the Year 1 material). A total of 882 genotyped clones were included in both the LB and emergence panels and were hence evaluated in both field trials.

Our experiment monitors a breeding population (RESPOT), and this had several implications for our

experimental design. The population was subject to ongoing selection concurrently with field trials, resulting in clones being removed from trials across the experimental period, and early year material being introduced in Year 2 or Year 3 of the trials. This meant that 1007 out of 1520 clones in the LB trial and 1907 out of 2576 clones in the emergence trial were measured in only 1 year. Lastly, phenotyping trials were conducted in parallel with propagation (on dedicated fields). This means that a proportion of the clones did not have sufficient plant material for multiple within-year replicates (220 clones in 2022 and eight clones in 2023 of the LB trial; 928 clones in 2023 and 727 clones in 2024 of the emergence trial). Therefore, the power to precisely determine phenotypic scores for each clone is limited, whereas the power to identify underlying genetic factors influencing the trait phenotypes is retained, assuming a shared pool of allele variants in the breeding population. Supporting Information File S1 lists the number of replicates of each clone in each trial. While the described lack of across- and within-year clone replicates of course reduces the robustness of the experiment, we have instead been able to screen a very large, and genetically diverse population, ranging from commercial cultivars to early year breeding offspring. This allows our study to overcome an important bias in only screening high-performance clones that have been subject to extensive selection, which would underestimate the true genetic diversity influencing trait phenotype.

The RESPOT clones were assigned to market groups. The market groups span four overall groups: crisps, French fry, fresh market, and starch processing. Market groups are assigned by Danespo A/S based on end use products, initial manual phenotyping, and parental pedigree. The market groups constitute separate breeding programs at Danespo A/S.

2.1.2 | LB field trials

The LB field trials were planted in Give, Denmark, in late May/early June in 2022–2024. The trials were planted with 75 cm between each row and 33 cm between each plant. The distance between each plot was 99 cm in the years 2022–2023 and 132 cm in 2024. The increased inter-plot distance was applied to facilitate plot gridding in the drone image analysis software. Each plot included three plants. Randomized replicates were included in the trials each year. Of the 1520 genotyped clones that were evaluated in LB field trials, 127 had replicates in all 3 years, 386 had replicates in 2 of 3 years, and 1007 were only included in a single trial year. Within-year replicates ranged from 1 to 8 (83 and 53 for two cultivars), as some of the plant material was propagated concurrently with field trials or was too young to enter earlier trials. Field management, including irrigation, fertilization, pest and disease control, and weeding, was performed in accordance with local

management and standard potato production practices, except omission of late blight control agent application postinoculation of late blight spreader rows. The LB trials were planned to minimize confounding with senescence—to decouple the phenotypes, plants were planted late and infected early in the season.

Spreader rows inoculated with a spore suspension of *P. infestans* were used to ensure even infection pressure across the field trial. The spreader rows consisted of the highly susceptible cultivar Bintje and a collection of other plant materials with varying levels of susceptibility. A mixture of susceptibility of the spreader row plants was used to prolong the infection pressure across the growth season following the die-off of the Bintje plants. Each plot in the spreader rows included one Bintje tuber and two tubers from the collection of susceptible cultivars. The first and last rows of each trial lane were spreader rows, and the first and last plots of each row were spreader plots. Spreader rows were placed systematically following every two rows of trial material in a repeating pattern for each lane: one spreader row, two rows of trial material, in repetition, and finally terminating in a spreader row (Figure S1).

Inoculation with *P. infestans* spore suspension was performed using a back sprayer in mid-July in each trial year. The inoculum consisted of a spore suspension of the late blight population endemic in the Danespo A/S fields in the preceding year. Each year, 10 samples were collected in the field during a late blight outbreak and stored on Bintje tuber slices. The cultures were maintained with regular subculturing onto new Bintje tuber slices. The regular subculturing continued until 14 days before inoculation of the LB field trial. At this point, the stored *P. infestans* population was multiplied by cultivating each late blight sample on approximately 10 plates with 40–50 Bintje tuber slices on each plate. On the day of inoculation, spores were harvested from each plate by washing the mycelium with cold, demineralized water. The spore suspensions were pooled and incubated at 4°C for at least 3–4 h. The time of inoculation depended on weather conditions but was generally performed in the afternoon. It was postponed until evening if the weather was warm and dry, to ensure a better establishment of the pathogen. In the 2022 field trial, preventative treatment was applied from sprouting until inoculation. No preventative late blight treatment was applied in the 2023–2024 field trials. LB was evaluated in drone flight missions at four to six temporal points across the growth season prior to harvest.

2.1.3 | Emergence field trials

The emergence field trials were planted in 2023–2024 in Give, Denmark. They were planted 75 cm between each row and 33 cm between each plant. Each plot included seven plants,

and the distance between each plot was 132 cm. The field trials were planted from early May until early June.

The emergence panel field trials were grouped according to market subgroups, each subgroup having a dedicated trial, delimited by border rows, in the same field (Figure S2). Some starch market subgroup trials (also some crisps subgroups in 2023) were located on a different field than the remaining market groups, as the starch market group cultivars generally have varying maturity and require different nutrient treatments compared to the remaining market groups. The first and last rows of each subgroup trial lane were guard rows. Guard plots of three tubers were planted before the first plot and after the last plot of each row of the subgroup trial lanes. Of the 2576 genotyped clones that were evaluated in emergence field trials, 669 had replicates in both years, and 1907 were only included in a single trial year. Within-year replicates generally ranged from one to eight (34 clones had more), as some of the plant material was propagated concurrently with field trials or was too young to enter earlier trials. Field management, including irrigation, fertilization, pest and disease control, and weeding, was performed in accordance with local management and standard potato production practices. Emergence was evaluated in drone flight missions at two to three temporal points early in the growth season.

2.2 | UAV image acquisition and field data collection

A Mavic 2 Enterprise Advanced (M2EA) drone was used to collect image field data for automated remote phenotyping using a 48MP RGB camera capable of 32x digital zoom. The camera was mounted on a three-axis gimbal for stabilization. The M2EA drone is a quadcopter with four rotors. The drone was equipped with a real-time-kinematic module, attaching each drone image with few-centimeter accuracy location data. Hence, the sensor platform operated independently of ground control points. The applied flight software was DJI Pilot, installed on the remote DJI Smart Controller of the M2EA.

The LB field trial flight protocols included at least four flights across the growth season (Table 1). The first flight was performed when approximately 50% of the plants were heavily infected with late blight, based on manual inspection. Before the first flying mission, the plants were inspected for symptoms of early blight (*Alternaria solani* causative agent) by manual field inspection. This was to prevent misidentification of late blight during phenotyping. The emergence field trial flight protocols included two to three flights across the growing season (Table 1). The first flight was performed when approximately one-third of the trialed plants had emerged, based on manual inspection. The consecutive flight(s) were performed with a spacing of at least 2 days. All flight

TABLE 1 Drone flight dates of the emergence and late blight resistance field trials.

Emergence					
Flight	2023		2024		
	Field 1	Field 2	Field 1	Field 2	
				Lane 2	Lane 4
1	June 1	May 26	June 3	May 26	June 27
2	June 4	May 30	June 5	May 30	July 3
3	June 7	June 2	June 7	–	–

Late blight resistance			
Flight	2022	2023	2024
	Field 1	Field 1	Field 1
1	July 18	July 12	July 18
2	August 5	July 31	July 25
3	August 9	August 2	July 29
4	August 11	August 3	July 31
5	–	August 4	–
6	–	August 11	–

protocols were performed at maximal wind speeds of 10.5 m/s and drone heights of 20 m.

Phenotyping of both emergence and LB from drone-based image data was performed by the external provider Phenoyard AB, Lund, Sweden. Following the collection of raw image data during each flight mission, the image data were directly uploaded to the Phenoyard AB web application for image analysis. No preprocessing was performed.

2.3 | UAV image processing and phenotype extraction

The image analysis was performed as a commercial service by Phenoyard AB. The analysis included image alignment, generation of a dense point cloud, construction of a digital surface model, and the production of georeferenced orthomosaic. This orthomosaic served as the base layer for all subsequent analyses. The plot boundaries were generated as a shapefile and overlaid onto the orthomosaic. Within each rectangular plot, individual rows of potato plants were identified and segmented. Late blight severity and emergence were predicted for each plot in each field trial using respective machine learning based models. The models were developed and trained on independent, curated datasets not containing any imagery from the experimental fields analyzed in this study. The training dataset comprised UAV-captured RGB imagery of potato canopies from multiple field trials conducted across different geographic locations and growing seasons (years). These multilocation, multiyear datasets were designed to include a wide range of phenotypic variability, canopy structures, and lighting conditions to improve model robustness. Both the training

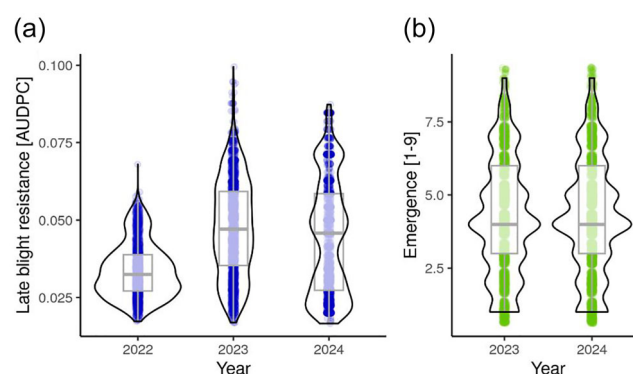


FIGURE 1 (a) Distribution of late blight resistance phenotypes computed as the relative area under the disease progress curve (AUDPC) in the late blight resistance (LB) panel and (b) of the plant emergence phenotypes in the emergence panel.

data and the models were regularly updated throughout the trial years. Hence, the likelihood of bias in data collection (if any) is minimal. For emergence, the key parameters of phenotype computation were total plot biomass, single plant biomass, plant uniformity across plots, and the number of plants in the plot. LB was assessed through estimation of the leaf chlorophyll content based on disparate reflectance profiles of healthy and stressed (e.g., diseased) plants in the visible electromagnetic spectrum (Sishodia et al., 2020). Once the image analysis was completed by the external provider, the computed phenotypic data were outputted as a .csv format file. As all plants in all field trials were photographed, there was no missing phenotypic data.

The output was two discrete phenotypes of emergence and LB, both evaluated on 1–9 scales. Discrete phenotype scales were used to ensure compatibility with the existing manual grading scales for LB and developmental phenotypes at Danespo A/S, that is, allowing integration of manual and drone data in the breeding program. For emergence, 9 denotes *early emergence*, and 1 denotes *late emergence*. The emergence scale was split into three temporal points where Grades 9–7 were given at the earliest flight, subsequently 6–4, and 3–1 in the progressively later flights. For Field 2, 2024 data, with only two flights, Grades 1–6 were given based on the first flight. Each plot was only scored once. For LB, 1 was *not susceptible*, 9 was *completely susceptible*, and a score of 10 denoted dead plants. The obtained time series of discrete LB phenotypes computed for the LB panel was converted into a single relative area under the disease progress curve (AUDPC) value (relative to the total monitoring period) for each observed plot using the trapezoid method with an adjustment for the total length of the monitoring period using the `audpc` function of the `agricolae` package (v1.3.7) (de Mendiburu & Yaseen, 2020). The phenotype distributions presented differing variances and means for each year (Figure 1a; Table 2). To account for year-to-year varia-

TABLE 2 Number of observations (obs.), number of unique clones observed, range of recorded phenotypic value, mean, standard deviation (SD), and coefficient of variance for phenotypic data in each recorded year. Emergence was measured on a discrete scale, and late blight resistance was computed as a relative area under the disease progress curve (AUDPC).

Phenotype	Year	No. of obs.	No. of clones	Range	Mean	SD	Coefficient of variance (%)
Emergence	2023	2923	1654	1–9	4.21	1.96	46.6
Emergence	2024	3339	1591	1–9	4.27	1.96	45.9
Late blight resistance	2022	1734	676	0.0173–0.0679	0.0337	0.00873	25.9
Late blight resistance	2023	2342	735	0.0168–0.0995	0.0474	0.0115	32.7
Late blight resistance	2024	2896	749	0.0165–0.0873	0.0445	0.0186	41.8

tion in LB, the relative AUDPC values were mean-centered and scaled to unit variance (*z*-score transformation) for each year separately, such that each year had a mean of zero and a standard deviation of one. This standardization was performed independently within each year to remove differences in disease pressure and environmental conditions across years while preserving relative differences among genotypes.

To assess agreement between manual and remote phenotyping of LB, Kendall's tau-b and weighted Cohen's kappa were calculated using the psych package (v2.6.1) for a pilot trial prepared according to previous descriptions. The phenotypic data were collected and matched for two time points in the 2021 growing season, namely August 4 and August 11. Both methods use the same scale 1–9, where 1 is *not susceptible* and 9 is *completely susceptible*. The score 10 for dead plants was unique to the drone phenotypes and was excluded in this analysis. This data included 1310 observations of 535 clones (one or three replicates of each clone), 157 of which were also included in the field trial experiments.

2.4 | Genotyping

Genotyping of the RESPOT population was performed across 5 years in yearly batches using a custom 20k Affymetrix chip. The entire RESPOT population of 5866 breeding clones was genotyped using 22,224 biallelic single nucleotide polymorphism (SNP) positions anchored in the potato DM v6.1 reference genome (Pham et al., 2020). Preprocessing and quality control of the genotype data were performed batch-wise according to ThermoFisher's recommended best practices. The batch-wise preprocessing of the data resulted in a subset of missing/erroneous SNP positions being discarded across years (batches). Therefore, a filtering step was implemented during data merging, retaining only SNPs with high-quality genotypes across all batches (9426 SNPs). Subsequently, monomorphic SNPs were removed, and the biallelic SNPs were filtered using a missing rate of $\leq 50\%$ per SNP, a missing rate of $\leq 70\%$ per clone, and a minor allele frequency (MAF) of $\geq 1\%$, generating a final set of 8731 SNPs. A marker density plot of the final set of 8731 biallelic SNPs was

created using the CMplot package (v4.5.1) in R (LiLin-Yin, 2023). Individuals were filtered to an individual call rate of 0.4—this did not remove any individuals from the analysis. However, five clones had been genotyped in more than one batch. For each of these clones, only the genotype batch with the lowest missing rate was kept.

The overall missing data rate of the genotypes for the final set of 8731 SNPs of the 5866 clones was 11.6%. Imputation of these data points was conducted using a random forest approach with 50 trees using the missRanger R package (v2.6.1) (Mayer, 2024). Imputation accuracy was assessed by calculating the Pearson correlation between the imputed genotypes after introducing a proportion of missing values to the genotype data and the imputation of the true missing values before introducing missing values. As the observed missing rate was 11.6%, proportions of 10% and 15% missing values were introduced randomly. The process was repeated 10 times, and the results are provided in Table S1. The standard error (s.e.) was computed as the standard deviation divided by the square root of the number of replicates.

2.5 | Population structure analysis

Population structure was characterized for the emergence and LB panels separately by principal component analysis (PCA) of their respective genomic relationship matrices (GRMs). The AGHmatrix package (v2.1.4) (Amadeu et al., 2023) was used to compute the GRMs based on the VanRaden method 1 with adjustment for tetraploidy (VanRaden, 2008). As the imputed entries contained non-integer values, the genotype dosages were converted to allele ratios, and the ratio option was applied. PCA was performed using the prcomp function of the built-in stats R package (R Core Team, 2024).

2.6 | Genome-wide association studies

GWAS was performed for each trait. Initially, phenotypic observations were converted to best linear unbiased estimators (BLUEs) for each line. The potato line was modeled as a fixed effect (all other effects were random) and the BLUE

was computed for each line by the `mmer` function of the `sommer` package (v4.3.5) in R (Covarrubias-Pazaran, 2016), using the following model for both the emergence and LB panels (Equation 1):

$$\mathbf{y} = \mu + \mathbf{X}\mathbf{b} + \mathbf{Z}_1\mathbf{f} + \mathbf{Z}_2\mathbf{s} + \mathbf{e} \quad (1)$$

where $\mathbf{y}$ is a vector of observed phenotypes (emergence panel) or standardized within-year relative AUDPCs (LB panel); μ is the population mean; $\mathbf{X}$ is a design matrix for the fixed effect; $\mathbf{b}$ is the vector of BLUEs for each line modeled as a fixed effect; $\mathbf{Z}_n$ ($n = 1, 2$) are incidence matrices for each of the random effects; $\mathbf{f}$ is a vector of environmental effects with $\mathbf{f} \sim N(0, \mathbf{I}\sigma_f^2)$, where $\mathbf{I}$ is an identity matrix and σ_f^2 is the variance due to uncorrelated environmental effects—for the emergence panel the environmental effect covers year $\times$ subtrial $\times$ field $\times$ lane, and for the LB panel, it covers year $\times$ field $\times$ lane as the LB panel was not planted with subtrial partitioning in the field layout; $\mathbf{s}$ is a vector of spatial effects, following a multivariate normal density with $\mathbf{s} \sim (0, \mathbf{S}\sigma_s^2)$, where $\mathbf{S}$ is a spatial relationship matrix and σ_s^2 is the spatial effect variance. The $\mathbf{e}$ is a vector of random residuals with $\mathbf{e} \sim N(0, \mathbf{I}\sigma_e^2)$, where σ_e^2 is the residual variance.

The spatial effect was defined (Raffo et al., 2022) as the combination of nine spatial subcomponents, where one subcomponent is the spatial effect for the square centered on the plot of the observation (i.e. the target plot) and the eight remaining subcomponents are the spatial effects for the square centered on the eight plots surrounding the target plot. Virtual plots were introduced over trial borders and into empty spaces of the grid to control border effects and ensure complete surrounding plots of all target plots. The spatial effects were identified by a unique ID given by the plot row (X) and column (Y) coordinates. The spatial correlation matrix $\mathbf{S}$ was computed for each lane of the late blight trials and for each subgroup lane in the emergence field trial (emergence field trials layouts were planted according to subgroups [Figure S2]), accounting for trial layout, using Equation (2):

$$\mathbf{S} = \frac{\mathbf{X}_1\mathbf{X}_1'}{\text{tr}(\mathbf{X}_1\mathbf{X}_1')/n} \quad (2)$$

$\mathbf{X}_1$ is an $m \times q$ indicator matrix with $m = \text{no. observations}$ (i.e., no. of real plots) and $q = \text{no. real} + \text{virtual plots}$. The indicator matrix relates observations to spatial effects in $\mathbf{S}$. tr is the trace (sum of diagonal elements) and n is the total no. of plots. This definition of spatial effect is not based on any assumptions on gradient or specific patterns across the field, allowing for spatial heterogeneity in any direction to accommodate the generally random fluctuations in spatial variance observed across a field (Raffo et al., 2022). The spatial matrices of the emergence and LB panel field trials were combined

for each trait. Heatmaps of a subset of the combined $\mathbf{S}$ matrices for the emergence and LB field trials are shown in Figures S3 and S4.

GWAS was performed by additive or simplex dominance single-marker regression using the `GWASpoly` package (v2.13) (Rosyara et al., 2016) based on the computed BLUEs. The following linear mixed model was used for GWAS of each panel (Equation (3)):

$$\mathbf{b} = \mu + \mathbf{W}\mathbf{c} + \mathbf{X}_i\beta_i + \mathbf{g} + \mathbf{e} \quad (3)$$

$\mathbf{b}$ is the vector of calculated BLUEs; μ is the population mean; $\mathbf{W}$ is a design matrix for the fixed effects and $\mathbf{c}$ is a vector of fixed effects including principal components (PCs) of the respective panel GRM (and genotypes for QTL-associated SNPs included as cofactors for one LB model and one emergence model); $\mathbf{X}_i$ is the vector of SNP genotype at the i th position; β_i is the corresponding genetic effect of the i th SNP; $\mathbf{g}$ is a vector of random genomic breeding values with $\mathbf{g} \sim N(0, \mathbf{G}\sigma_g^2)$, where $\mathbf{G}$ is the GRM of the respective panel and σ_g^2 is the genetic variance, where $\mathbf{e}$ is a vector of random residuals with $\mathbf{e} \sim N(0, \mathbf{I}\sigma_e^2)$ and σ_e^2 is the residual variance. A new $\mathbf{G}$ -matrix was computed for each chromosome using the `GWASpoly` package with leave-one-chromosome-out (LOCO) cross-validation. Based on population structure analysis of the panel GRMs, PCs of the respective $\mathbf{G}$ -matrices were used to correct for population structure for the LB panel (three PCs) and the emergence panel (two PCs) by including these as covariates in the model.

Additional models were also tested for each market group for each trait, utilizing the equations described above (Equations 1–3). These were based on genotypic data for individuals within each market group. New GRMs were calculated for each market group and used to calculate the random genomic breeding values, and PCs of the market group-based $\mathbf{G}$ -matrices (three PCs for the LB panel, two PCs for the emergence panel) were included as fixed effects in the respective models. LOCO cross-validation was applied. For the full emergence panel, an additional GWAS was also performed without including a spatial effect (i.e., excluding the $\mathbf{Z}_2\mathbf{s}$ random effect in Equation 1).

Modern cultivars are expected to have some proportion of recurrent genetic variation in their pedigree due to inbreeding. The relatively low genomic inbreeding coefficients of the emergence panel (0.0055) and of the LB panel (0.0071) likely reflect the extreme heterozygosity of tetraploid potato (The Potato Genome Sequencing Consortium, 2011). The genomic inbreeding coefficients were calculated as $(\text{diag}(\mathbf{G}) - 1)/(\text{ploidy} - 1)$, where $\text{diag}(\mathbf{G})$ is the mean diagonal element of the GRM (Gallais, 2003). To correct the p -values for any overdispersion arising from stratified populations (Devlin et al., 2001) and/or other structures in the data not accounted

for in the applied models, we used the genomic inflation factor, λ_{gc} (Hinrichs et al., 2009). Genomic inflation factors were calculated for each model of each trait using Equation (5):

$$\lambda_{gc} = \frac{\text{median}(\chi^2)}{Q_{\chi^2}^{-1}(0.5, 1)},$$

$$\chi^2 = Q_{\chi^2}^{-1}(P, 1) \quad (5)$$

The genomic inflation factor, λ_{gc} , was computed as the median value of the χ^2 -statistics of the SNPs, converted from each p -value, P , using the inverse cumulative distribution function (CDF) of the χ^2 -distribution, $Q_{\chi^2}^{-1}$, with 1 degree of freedom, divided by the expected median, assuming no association between the SNP and the trait phenotype, that is, the χ^2 -statistic of the 50th percentile. For models with $\lambda_{gc} > 1$, indicating systematic effects not captured by the model, a correction was performed by dividing the SNP χ^2 -statistics by λ_{gc} and the resulting values were converted to corrected p -values using the CDF of the χ^2 -distribution with 1 degree of freedom (Equation 6):

$$P_{\text{corrected}} = 1 - Q_{\chi^2} \left(\left(\frac{\chi^2}{\lambda_{gc}} \right), 1 \right) \quad (6)$$

Bonferroni correction was used to control false positive associations with a false discovery rate of $p < 0.05/N$, where 0.05 is the significance threshold and N is the number of markers tested. Q - Q and Manhattan plots were generated using the qqman package (v0.1.9) (Turner, 2018).

The genetic variance explained by the most significant SNP of each QTL was calculated using Equation (7) (Park et al., 2010):

$$\sigma_{\text{SNP}_i}^2 = 2 \cdot \beta_i^2 \cdot f_i (1 - f_i) \quad (7)$$

where β_i is the effect size of the i th SNP and f_i is the MAF of the i th SNP. This was used to calculate the fraction of phenotypic and additive genetic variance explained by each of the most significant QTL SNPs using the total phenotypic (σ_p^2) and genetic variances (σ_g^2) explained by the model.

The narrow sense additive genetic heritability for each trait was computed as the ratio of additive genetic (σ_a^2) to total phenotypic variance (σ_p^2) (Equation 8):

$$h^2 = \frac{\sigma_a^2}{\sigma_p^2}, \quad \sigma_p^2 = \sigma_a^2 + \frac{\sigma_e^2}{n} \quad (8)$$

where $\frac{\sigma_e^2}{n}$ is the residual variance divided by the minimum number of clone replicates, that is, $n = 1$. This overestimates

the error term for the clones with replicates but provides the most conservative heritability estimate, corresponding to the plot heritability.

2.7 | Candidate gene mapping

Candidate gene mapping was performed for each QTL by exploring the gene annotation of the v6.1 reference genome (Pham et al., 2020) in regions 500 kb upstream and 500 kb downstream of the most significant SNP of the QTL. A region of 1 Mb total window size was applied based on the estimated average linkage disequilibrium (LD) block size of 0.6–1.5 Mb in tetraploid potato (Vos et al., 2017).

3 | RESULTS

3.1 | Phenotyping

Based on a pilot trial of LB scored using both manual grading and remote phenotyping in 2021 at two time points (August 4 and August 11) during the inoculated late blight epidemic, we found good agreement between the two phenotyping approaches. Significant Kendall's tau-b correlations were 0.7 (p -value $< 2.2 \times 10^{-6}$) for August 4 and 0.67 (p -value $< 2.2 \times 10^{-6}$) for August 11 and weighted Cohen's kappa of 0.77 (95% confidence interval [CI]: [0.74, 0.79], s.e. = 0.014) and 0.68 (95% CI: [0.65, 0.7], s.e. = 0.013), respectively. Emergence cannot be compared to a single equivalent manual phenotype, as it is computed as a combination of plant count, plant biomass (canopy cover), and plant uniformity.

For the LB experimental field trials, a total of 6972 observations of 1520 genotyped clones were recorded across three consecutive years (Table 2; Figure 1). We observed year-to-year variation in LB based on mean and standard deviation across all clones. Using 16 commercial cultivars with at least four replicates per LB trial (i.e., year) with known expected LB profiles, we also observed this trend (Figure S5). While there was overall good concordance between the measured resistance by remote phenotyping and the expected resistance level (benchmarking the reliability of the remote phenotyping approach), not all cultivars displayed uniform resistance across all years—some were found to be more susceptible in 2024, others increasingly resistant. This pattern is consistent with *Rpi*-gene-based resistance, which will be sensitive to the disparate inoculum used across years. This year-to-year variation was accounted for through within-year standardization prior to GWAS analysis.

A total of 6262 observations were recorded for 2576 genotyped clones in the emergence panel field trials across two consecutive years. The phenotypic variation was

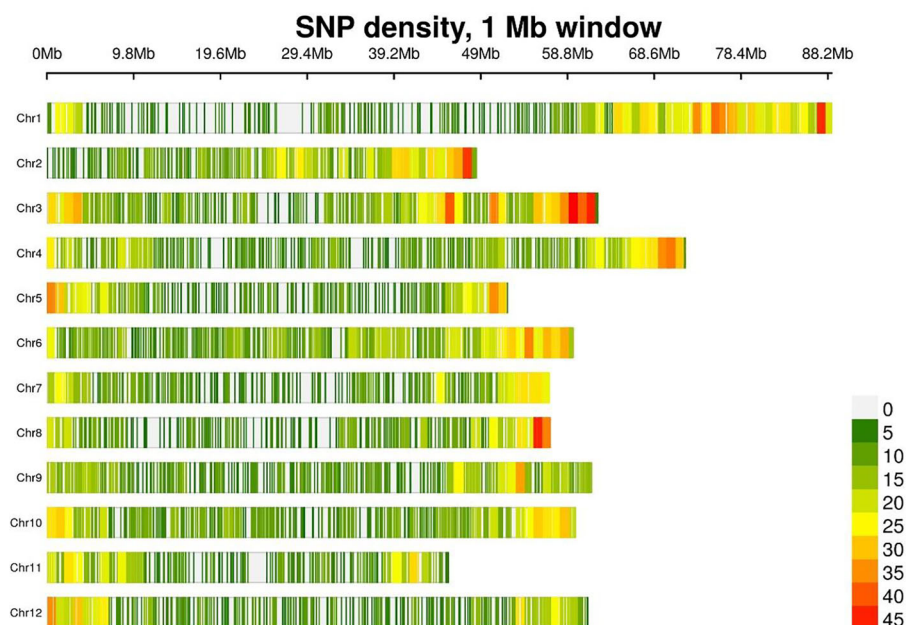


FIGURE 2 Heatmap of marker density in 1 Mb windows for each of the 12 chromosomes. The color gradient denotes marker count.

similar across the 2 years for this phenotype when comparing full trials. Looking at 16 commercial cultivars with several replicates across both years, we found that emergence estimates varied little for the same clone within the same trial year, but there was some evidence of across-year variation (Figure S6). No additional data processing was performed prior to GWAS analysis of this trait.

The starch market segment had significantly lower mean relative AUDPC values compared to the other market segments in all years (Figures S7 and S8; Table S2). This showed that the starch market segment was associated with superior disease resistance (measured as leaf chlorophyll content) across the growing season compared to the remaining market groups. The starch potato market group had experienced the greatest selection for disease resistance during breeding at Danespo A/S. However, the starch market group also had a higher variance in senescence compared to other market groups and generally senesced later. As late senescing plants retain high chlorophyll content longer (Domínguez & Cejudo, 2021; Hörtensteiner, 2009), some of the effect on chlorophyll content for this market group might also be due to confounding of these phenotypes.

3.2 | Genotyping and imputation

The 8731 markers included in the association analyses were distributed across all 12 chromosomes, and high marker density regions are consistent with gene-dense apocentromeric regions and repressed meiotic recombination in the gene-sparse pericentromeric regions (Pham et al., 2020) (Figure 2).

A study of SNP reduction by LD pruning from Selga et al. (2020) demonstrated that a marker density of 5000 markers was sufficient to stagnate QTL discovery (i.e., no loss of information) in GWAS of four traits in tetraploid potato, including LB. Based on this, our marker density is expected to fully capture all trait QTLs.

The imputation accuracy of the missing genotype data entries was assessed by masking 10% and 15% of entries (by introducing artificial missing entries into the genotype data), and the correlation between imputed genotype and called genotype for these entries was used as a measure of imputation accuracy. While 10% and 15% missing entries were introduced in the genotypes, the total number of missing entries effectively introduced in the full RESPOT population genotype data was on average 8.84% and 13.26%, respectively. This is because some of the introduced missing entries corresponded to real missing values. The obtained mean imputation accuracies were 93.48% (s.e. = 0.01) and 93.18% (s.e. = 0.01) (Table S1). These high correlations demonstrate very good concordance between real and imputed genotypes, indicating that imputation of missing entries did not introduce a bias.

3.3 | Population structure analysis

The LB and emergence panels are subpopulations of the RESPOT population, sharing 882 clones but otherwise including genetically distinct plant material. The market group composition and population structure of both panels are shown in Figure 3. Market groups are assigned by Danespo A/S according to the commercial market group of

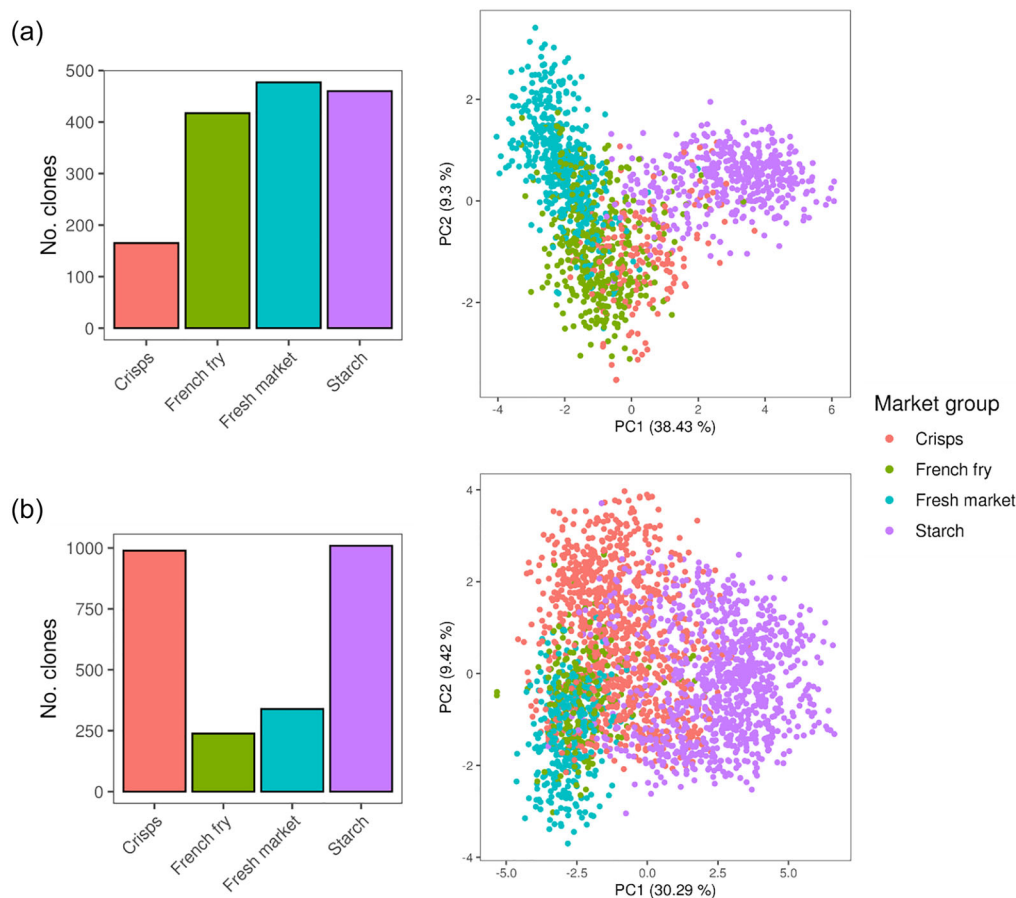


FIGURE 3 (a) The population of market groups of the genotyped late blight resistance (LB) panel (left) and principal component analysis of the genomic relationship matrix of 8731 single nucleotide polymorphism (SNP) markers for the 1520 genotyped clones in the LB colored by market group (right). Principal component 1 is plotted against principal component 2, explaining 38.43% and 9.3% of the total explained variance, respectively. (b) The population of market groups of the genotyped emergence panel (left) and principal component analysis of the genomic relationship matrix of 8731 SNP markers for the 2576 genotyped clones in the emergence colored by market group (right). Principal component 1 is plotted against principal component 2, explaining 30.29% and 9.42% of the total explained variance, respectively (Figures S9 and S10).

established cultivars and based on end use product, pedigree, and manual phenotypic assessment of the breeding material. Market grouping is relevant as market groups constitute distinct breeding programs within Danespo A/S, which is expected to give rise to inter-market group population structure. Furthermore, subgroupings exist within each market group, for example, the fresh market group includes both the Scandinavian fresh market, export fresh market to Mediterranean Europe, and colored flesh. This can give rise to some population structure within market groups, but it is expected to be minor compared to any inter-market group structure. While the LB panel is composed of similar proportions of French fry (28%), fresh market (31%) and starch (30%) potatoes, and some crisp (11%) potatoes (Figure 3a), the larger emergence panel is most heavily populated by starch (39%) and crisp (38%) potatoes, including only 9% and 13% French fry and fresh market group clones, respectively (Figure 3b).

PCA of the GRMs of the emergence and LB panels confirmed the presence of substantial population structure related

to market groups in both panels (Figure 3; Figures S11 and S12). In both panels, PC1, explaining 38.43% of the total explained variance of the LB panel and 30.29% in the emergence panel, separates the starch group from the French fry and fresh market groups, while crisps are not completely divergent from either of those genetic groupings. Meanwhile, for the LB panel, PC2, explaining roughly a quarter of the genetic variance captured by PC1, separates the fresh market group from the French fry and crisp groups. The inter-market group structure is consistent with the market group-specific breeding programs at Danespo A/S, but we also observe shared alleles across the breeding programs, manifesting as overlaps across market groups. For example, along PC1 the crisp market group overlaps with both the French fry and starch market groups. This is likely a result of recurrent parents in the pedigrees of the market groups.

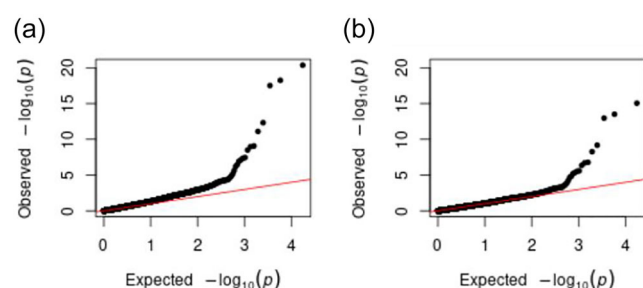
The detected population structure was included in the GWAS model, both through the G-matrix and by including PCs 1–2 as cofactors in the statistical model. PC3 was also

TABLE 3 Quantitative trait loci (QTLs) significantly associated with late blight resistance based on an additive model.

QTL	Chromosome	Position (bp)	<i>p</i> -value	MAF	$\sigma^2_{\text{SNP}_{\text{pheno}}}$ (%)	$\sigma^2_{\text{SNP}_{\text{genetic}}}$ (%)
QTL1	4	6,324,916	3.032×10^{-6}	0.078	1.117	1.863
QTL2	5	4,785,270	8.833×10^{-16}	0.243	3.711	6.189
QTL3	9	55,983,609	6.668×10^{-10}	0.047	1.502	2.505
QTL4	7	52,686,044	4.199×10^{-6}	0.124	0.857	1.439

 $h^2 = 0.60$

Abbreviations: MAF, minor allele frequency; SNP, single nucleotide polymorphism.

**FIGURE 4** *Q*-*Q* plots of observed against expected $-\log_{10}(p\text{-values})$, assuming no association between markers and trait for late blight using the additive model. (a) Before correction with the genomic inflation factor. (b) After correction.

included as a cofactor in the GWAS models of LB, as it captured a genotyping batch effect in the LB panel. The same effect was captured by PC2 for the emergence panel. The respective PCs account for a substantial 9.42% of the total explained variance in the emergence panel and 4.48% in the LB panel (Figures S13 and S14).

3.4 | GWAS of late blight resistance

The *Q*-*Q* plots of observed versus expected $\log_{10}(p\text{-values})$ show that the *p*-values using the additive model are inflated (Figure 4a). The additive model had a genomic inflation factor (λ_{gc}) of 1.37, while the dominant model λ_{gc} ranged from 1.25 for the alternative allele to 1.01 for the reference allele. The *p*-values of each model were corrected using their respective λ_{gc} , and the corrected *p*-values of the dominant allele models were subsequently combined. Following correction, significant associations were identified (Figure 4b).

The Manhattan plot of adjusted *p*-values of the additive GWAS model of LB shows three significant trait QTLs on chromosomes 4, 5, and 9, referred to as QTL1, QTL2, and QTL3, respectively (Figure 5a; Table 3; Supporting Information File S2). QTL1 consisted of a single SNP with a MAF of 0.078. This rare allele explained 1.117% of the phenotypic variance. QTL2 (*p*-value 8.833×10^{-16}) explained 3.711% of the phenotypic variance and had a MAF of 0.243, that is, it constitutes a relatively common variant in the LB panel. QTL3

(*p*-value 6.668×10^{-10}) is underpinned by a rare allele with a MAF of 0.047. As a consequence of the low MAF, the latter SNP explains only 1.502% of the total phenotypic variance. Inclusion of the observed LB QTLs as fixed effects in the GWAS model (i.e., including the genotype of the most significant SNP of a QTL as a fixed effect iteratively until the signal disappeared) yielded an additional single-SNP minor QTL (QTL4) on the South arm of chromosome 7 (Figure 5b; Table 3; Supporting Information File S2). It is a relatively common allele (MAF of 0.124) contributing to less than 1% of the phenotypic variance.

Because LB is partly conferred by dominant resistance genes, known as *Rpi*-genes, a simplex dominance GWAS model was tested in addition to the more general additive model, to possibly capture additional dominant loci not significantly associated using the additive model. The dominant model also identified QTL2-3, and QTL1 was just below the significance threshold (Figure 5c). Using a dominant model slightly shifts the most significant SNP of QTL2 to another SNP approximately 52 kb downstream. However, both SNPs are included in QTL2 in both models (Supporting Information File S2).

3.5 | GWAS of emergence

The *Q*-*Q* plots of observed versus expected $\log_{10}(p\text{-values})$ show that the *p*-values are highly inflated (Figure 6a), presenting with a λ_{gc} of 1.82. As population structure and genotyping batch effects have been extensively accounted for in the model, this likely reflects systematic structure in the phenotypic data not captured by the environmental or spatial effects. Regardless, following conservative correction of the *p*-values using the λ_{gc} , significant associations were identified (Figure 6b).

Manhattan plots of the corrected *p*-values of emergence are shown in Figure 7. Three QTLs with significant associations were identified in total, one on chromosome 2 (*p*-value = 1.496×10^{-9}) and two on chromosome 5 (*p*-values = 5.062×10^{-6} and 4.033×10^{-6}) (Figure 7a; Table 4; Supporting Information File S2) called QTL5, QTL6, and QTL7, respectively. Based on the most significant SNP of each QTL, QTL5 explains 1.437% of the phenotypic variance

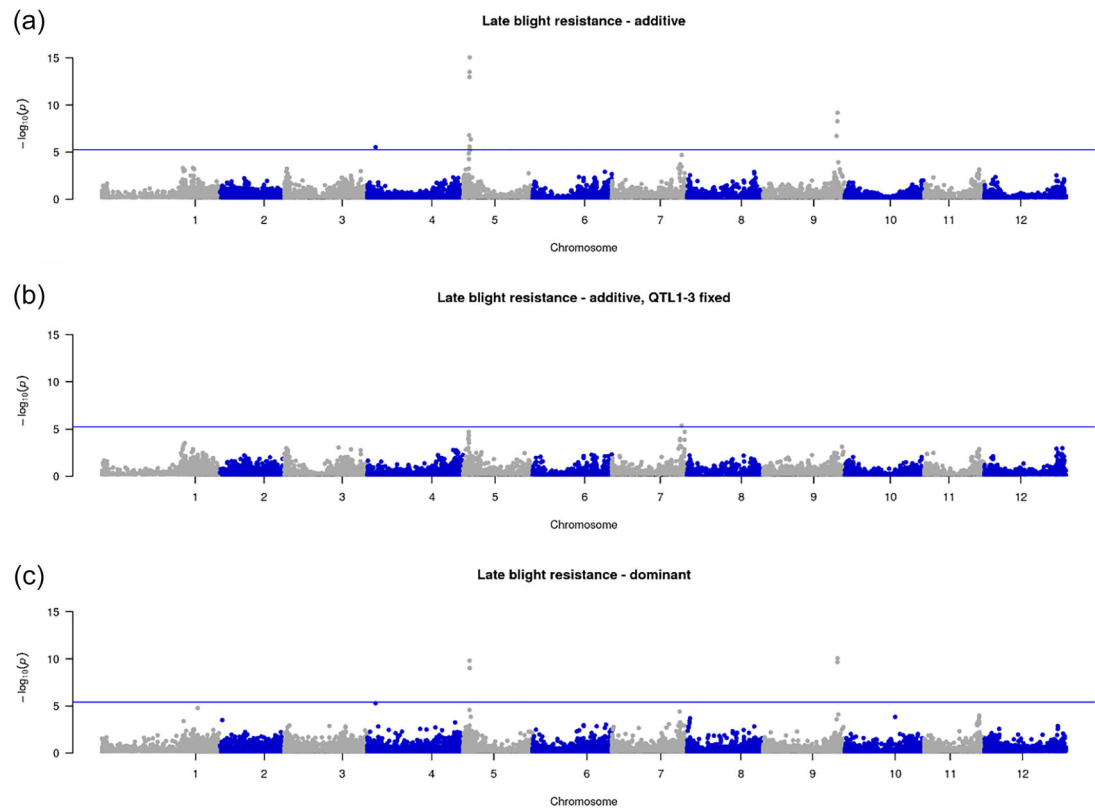


FIGURE 5 Manhattan plots of corrected $-\log_{10}(p\text{-values})$ obtained by genome-wide association studies (GWAS) of late blight resistance. (a) Late blight resistance using an additive model, (b) late blight resistance using an additive model with QTL1-3 included as fixed effects, and (c) late blight resistance using a simplex dominant model (combined data for reference and alternative allele). The Bonferroni correction significance threshold is shown as a blue horizontal line.

TABLE 4 Quantitative trait loci (QTLs) significantly associated with plant emergence.

QTL	Chromosome	Position (bp)	<i>p</i> -value	MAF	$\sigma^2_{\text{SNP}_{\text{pheno}}}$ (%)	$\sigma^2_{\text{SNP}_{\text{genetic}}}$ (%)
QTL5	2	26,729,700	1.496×10^{-9}	0.251	1.437	3.289
QTL6	5	11,507,032	5.062×10^{-6}	0.381	0.929	2.127
QTL7	5	37,078,168	4.033×10^{-6}	0.200	1.314	3.001
QTL8	3	46,471,107	3.634×10^{-6}	0.176	0.421	0.964

$h^2 = 0.44$

Abbreviations: MAF, minor allele frequency; SNP, single nucleotide polymorphism.

for emergence, QTL6 explains 0.929%, and QTL7 explains 1.314%. We also noted a flattened peak below the significance threshold on chromosome 3. Inclusion of the most significant SNPs for QTLs 5 and 6 as covariates rendered this a significant trait QTL (QTL8) ($p\text{-value} = 3.634 \times 10^{-6}$) (Figure 7b), explaining 0.421% of the phenotypic variance for emergence (Table 4). As inclusion of QTL6 as a covariate removes both itself and QTL7, these are likely linked. Based on minor allele frequencies, the alleles underpinning the phenotypic variation are relatively common in the RESPOT emergence panel for all QTLs.

The phenotypic effects exerted by QTLs 5–7 on emergence appear to function by additive transmission, based on the phenotypic distribution as a function of allele dosage (Figure 8). For QTL5, increased dosage of the alternative allele is associated with progressively earlier emerging plants. Comparatively, for QTL6 and QTL7, the alternative allele contributes to later emergence in cultivars by an additive inheritance mechanism. Interestingly, it is the heterozygous states that are most commonly encountered for QTL6 in the emergence panel, indicating that selection has favored a mixed allele profile for this particular QTL, which is associated

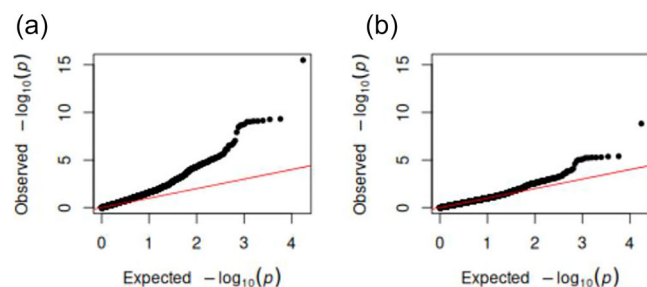


FIGURE 6 *Q-Q* plots of observed against expected $-\log_{10}(p\text{-values})$, assuming no association between markers and trait for emergence. (a) Before correction with the genomic inflation factor. (b) After correction.

with an intermediate emergence relative to either homozygous state. For QTL8, the alternative allele is associated with earlier emergence, but the effect is small compared to those imposed by the other emergence QTLs.

3.6 | Candidate gene mapping

Candidate gene mapping yielded candidate gene lists of 119 gene annotations for QTL1, 102 gene annotations for QTL2, 57 genes for QTL3, 109 genes for QTL4, 102 entries for QTL5, 61 candidate genes for QTL6, 12 entries for QTL7, and lastly 104 entries for QTL8 (Supporting Information S3). For six out of the nine QTLs, genes potentially acting in pathways linked to the phenotype could be identified. Further studies are necessary to determine the true causative gene alleles.

4 | DISCUSSION

4.1 | Phenotyping

There was good agreement between the manual and remote phenotyping results in the pilot trial of LB, but not a perfect correlation. The two approaches are probing disparate features of the LB trait. Whereas manual scoring is based on expert evaluation of late blight infection, for example, lesions, remote scoring is based on leaf chlorophyll content. Hence, a correlation could not necessarily be expected, but for this phenotype we observed good agreement.

We noted some year-to-year variation in disease progression rate and/or augmented disease susceptibility—both in the phenotypic distribution across each trial, and across control cultivars with replicates in all trials. This can reflect some variation in phenotyping, but as several cultivars have consistent observations across years, we speculate that a $G \times E$ effect is more likely, for example, arising from environmental factors favoring *P. infestans* pathogenicity. Weather records (Danish Meteorological Institute, <https://www.dmi.dk/>) (Table S3) showed that precipitation across the cultivation period from

May to July varied across the trial years, which is one factor that can greatly influence the growth conditions for the late blight pathogen (Bagchi et al., 2024). Additionally, the pathogenicity of the isolate used for each inoculation, that is, the epidemic isolate of Danespo A/S fields in the previous year, could have been markedly different in pathogenicity across trial years. Lastly, the clones trialed in any given year might be characterized by lower genetic resistance compared to those screened in other years, as the LB trial had limited across-year replicates of several clones. However, as we also observe year-to-year variation for the same cultivars, this is not the main cause of the effect. Standardization of the LB phenotypic data is expected to have sufficiently controlled for year-to-year phenotypic variation in our analysis.

For emergence, we similarly noted accurate within-year phenotyping of the same cultivars across replicates, indicating robust remote phenotyping. However, there was notable variation across the 2 years for some clones, which could indicate a substantial $G \times E$ effect of this complex developmental trait and/or some across-year variation in phenotyping. This highlights a limitation of our study, in having some clones included in only one trial year, as there were not sufficient replicates to estimate a $G \times E$ effect.

4.2 | Population structure and market groups

The population structure analysis showed that the market groups, demarcated into separate breeding populations at Danespo A/S, constitute genetically distinct base populations with some shared alleles likely due to recurrent parents in their pedigrees. This was to be expected based on the disparate phenotypic profiles that characterize the end use products of each market group, and the distinct subclusters formed for each market group clearly demonstrate the market group-focused breeding strategies applied at Danespo A/S.

It is plausible that each market group is associated with different trait genetic underpinnings, for example, different *Rpi*-gene introgressions found in the pedigree of progenitors unique to some or single market groups. This could be highly relevant to assess for, for example, the starch market group, which has experienced the longest targeted selection of LB of the four market groups. To explore this, we performed GWAS analysis for both LB and emergence within each market group using the same model as for each of the full panels (including correction for environmental effects, spatial effects, and population structure—the population structure is based on a GRM of each market group in the panel) (Figures S15–S22; Table S4; Supporting Information File S2). Interestingly, we identified QTL signals on different chromosomes in different market groups for both traits. For the emergence phenotype, QTL5 was reproduced using the crisps market group alone. Using the starch market group, a broad signal

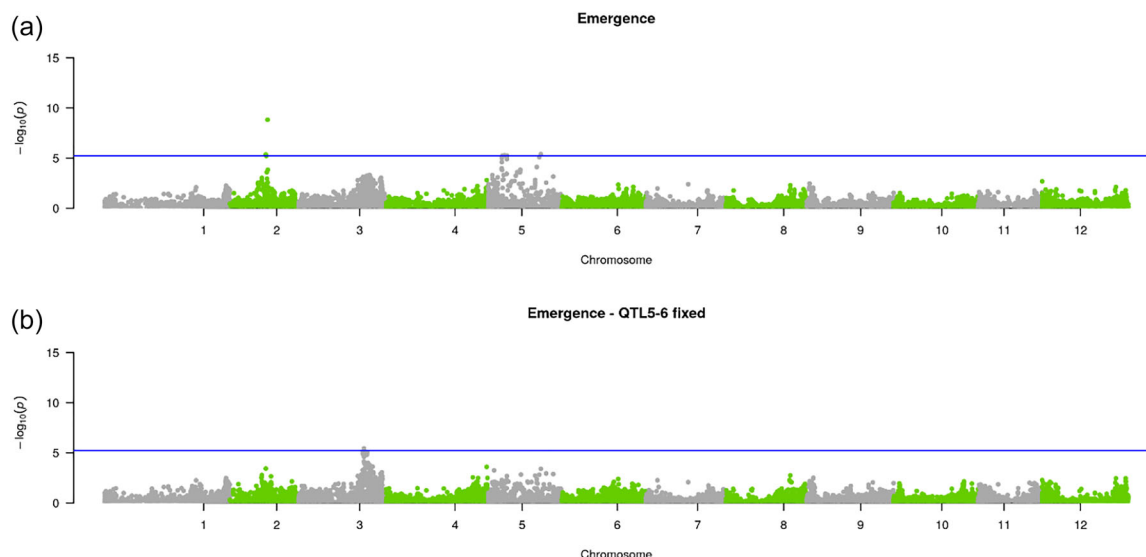


FIGURE 7 Manhattan plot of corrected $-\log_{10}(p\text{-values})$ obtained by genome-wide association studies (GWAS) of emergence (a) using an additive model and (b) using an additive model with quantitative trait loci (QTL)5–6 included as fixed effects. The Bonferroni correction significance threshold is shown as a blue horizontal line.

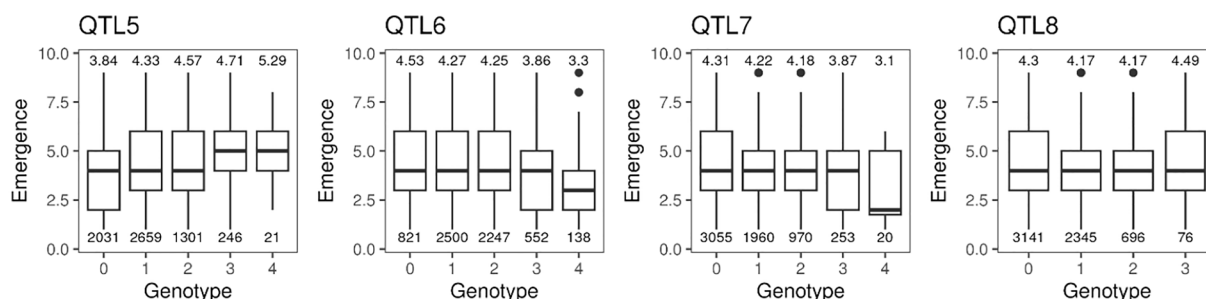


FIGURE 8 Boxplots of emergence phenotype as a function of allele dosage for quantitative trait loci (QTL)5–8 across all clones in the emergence panel. Imputed genotypes were discretized by rounding to the closest integer. The mean phenotypic value is shown above the boxplots of each genotypic state, and the number of phenotypic observations for each genotypic state is shown below.

surrounding QTL6 for emergence—from 9.4 to 13.1 Mb—could be identified. For LB, the fresh market group alone reproduced QTL2 and QTL3. Notably, using the French fry market group alone yielded a single significant marker on chromosome 2, at 4.2 Mb, not encountered in the other LB panel models (QTL9). QTL9 is, however, not located in the vicinity of a known resistance gene cluster (Supporting Information File S2). With larger genotyped subpopulations, there might be merit in exploring trait associations across the entire population, with correction for population structure arising from the distinct market group breeding subpopulations, as well as within each market group.

4.3 | Late blight resistance GWAS

GWAS identified four QTLs associated with LB in the RESPOT population, several of which (QTL1–3) colocal-

ize with known *Rpi*-gene clusters in the DM v6.1 reference genome (Supporting Information S3).

QTL1 is located directly between two clusters of resistance genes on chromosome 4 in the reference genome (Supporting Information S3). Chromosome 4 encodes several active *Rpi*-genes, including homologues *Rpi-blb3*, and *Rpi-abpt* (Lokossou et al., 2009), as well as *Rpi-R2*, *Rpi-demf1*, *Rpi-edn1.1-1.2*, *Rpi-hjt1.1-1.3*, *Rpi-mcd1*, *Rpi-snk1.1-1.2*, and *Rpi-amr3* (Paluchowska et al., 2022). Particularly, the *Rpi-R2* on the short arm of chromosome 4 (X. Li et al., 1998) is of interest, as cultivars such as Gloria that harbor this *Rpi*-gene are included in the RESPOT pedigree. Resistance through *Rpi-R2* is quantitative, providing a delayed onset of infection by virulent *P. infestans* isolates, with responses varying depending on the genetic background of the isolate carrying the *Rpi-R2* virulence factor (Pilet et al., 2005).

QTL3 also colocalized with known *Rpi*-gene clusters. Chromosome 9 hosts several known active, qualitative resis-

tance genes, including *Rpi-R8*, *Rpi-R9a*, *Rpi-mcql*, *Rpi-nrs1*, *Rpi-ver1*, *Rpi-dlc1*, and the three *Rpi-vnt1* homologues (Haverkort et al., 2016; Paluchowska et al., 2022). The RESPOT parental population includes Sarpo Mira, which harbors *Rpi-R8* (Paluchowska et al., 2022; Vossen et al., 2016), and also cultivar Alouette, with a lineage back to the progenitors of *Rpi-vnt1* introgressions (Paluchowska et al., 2022).

QTL2 colocalizes with the *Rpi-R1* introgression from *Solanum demissum* in the 5.76–6.20 Mb region on chromosome 5 (based on the Cooperation 88 genome assembly [Bao et al., 2022]) (Leonards-Schippers et al., 1992; Bao et al., 2022; Paluchowska et al., 2022). This commonly deployed *Rpi-R1* resistance gene is present in several RESPOT parent cultivars including Bzura, Cara, Innovator, Spunta, and Stirling (Paluchowska et al., 2022). The qualitative genetic resistance conferred by this *Rpi*-gene is considered overcome by *P. infestans* (Coomber et al., 2024). However, it is possible that *Rpi-R1* confers a quantitative resistance profile as a gene segment of the chromosome 5 *S. demissum* *Rpi-R1* introgression has been associated with quantitative contributions to LB (Gebhardt et al., 2004). While the *Rpi-R1* resistance gene is unlikely to contribute to durable resistance against the currently endemic late blight strains by single gene deployment, it could be a component in a multigene strategy of combined qualitative and quantitative resistance to generate a complex and enhanced resistance profile by *Rpi*-gene pyramiding in future cultivars (Haverkort et al., 2016; Pilet et al., 2005; Rietman et al., 2012). An alternative hypothesis that cannot be excluded is that there is a novel cryptic *Rpi*-gene in close proximity to *Rpi-R1* with retained activity that has not previously been identified due to confounding of the locus with *Rpi-R1* through tight linkage.

However, QTL2 might also be a result of confounded phenotypes. Measuring the chlorophyll content directly assesses plant stress at all stages of plant development (Maxwell & Johnson, 2000). As leaves experience chlorophyll loss through chloroplast degradation during senescence (Domínguez & Cejudo, 2021; Hörtensteiner, 2009), the LB and senescence phenotypes are confounded through the general measure of chlorophyll content. Another probable confounder arises through a pleiotropic effect of the genes controlling the onset of senescence, as plants experience a deterioration of overall host resistance during senescence, increasing late blight susceptibility in this developmental phase (Kloosterman et al., 2013; L. Shi et al., 2024). The QTL2 location is in relatively close proximity of the *CYCLING DOF FACTOR1* (*StCDF1*) gene (PGSC003DMG400018408 [v4.03]/Soltu.DM.05G005140 [v6.1] from 4,485,531 to 4,488,495 on chromosome 5 [Pham et al., 2020; Zagorščak et al., 2024]) that influences potato plant senescence (Kloosterman et al., 2013). Alleles encoding truncated *StCDF1* gene products confer early vine maturity and long-day tuberization

by modulating the mediation between the circadian clock and the tuberization signal (L. Shi et al., 2024). The alleles are dominant but also exert their effect in a dosage-dependent manner, accelerating both the onset and progression of senescence (Hoopes et al., 2022). The putative detection of a significant *StCDF1* QTL using this phenotype could be indicative of a genetic correlation between senescence and LB. However, the experimental trials were designed to uncouple a pleiotropic effect between LB and senescence as much as possible. Further elucidation of the true genetic variation underpinning QTL2 is required.

A general consideration of the phenotypic variance explained by *Rpi*-gene QTLs is that the active *Rpi*-genes are expected to confer a great phenotypic effect on each clone, but their rarity leads to a relatively low population-wide phenotypic effect. Additionally, in the RESPOT population, based on the Danespo A/S breeding population, *Rpi*-genes are generally deployed as a single *Rpi*-gene in each cultivar. Hence, there might be an additional additive phenotypic effect conferred by the *Rpi*-genes if they were to be deployed as multi-*Rpi*-gene cassettes in single cultivars, similar to the high field resistance of Sarpo Mira with five *Rpi*-genes, including the quantitative *Rpi-Smira-2* (Paluchowska et al., 2022). Increased resistance conferred by a gene stack compared to single deployed *Rpi*-genes was also demonstrated for transgenic lines with *Rpi-blb1*, *Rpi-blb2*, and *Rpi-vnt1.1* in the Victoria cultivar (Ghislain et al., 2019). The detection of multiple putative *Rpi*-genes in the Danespo A/S breeding population demonstrates that pyramiding of multiple *Rpi*-genes is possible through breeding.

QTL4 is unlikely to be constituted of an *Rpi*-gene, as it is not located in a known *Rpi*-gene cluster (Supporting Information File S3). There are very few *Rpi*-genes encoded on chromosome 7 (i.e., *Rpi1*, *Rpi2*, and *Rpi-mch1*) (Paluchowska et al., 2022), and none of them are harbored by known cultivars in the RESPOT pedigree. Additionally, it is a relatively common allele in the RESPOT population with a small phenotypic effect. While it could be a false positive association, it is also possible that it is underpinned by novel genetic variation that influences overall host resistance or leaf chlorophyll content. For example, the candidate region includes genes encoding proteins involved in the glutathione pathway, detoxification, hormone biosynthesis, and signal transduction pathways influencing root growth and drought tolerance, and leaf cell division that could influence overall plant growth and vigor, as well as biotic and abiotic stress tolerance (Hasi & Kakimoto, 2022; He et al., 2018; Hernández Estévez & Rodríguez Hernández, 2020; Lo et al., 2008; Zhao et al., 2020).

The late blight QTLs are a benchmark for the applied remote phenotyping approach, as we identify QTLs consistent with known active *Rpi*-gene clusters that can confer quantitative or qualitative resistance to *P. infestans*. The identified

QTLs 2–3 also reproduce results obtained by Romé et al. (2026) in a GWAS of manually scored LB phenotypes on the genotyped RESPOT clones. Based on these benchmark results, we show that the QTLs found for LB are reliable, indicating that our results for emergence based on the remote phenotyping approach are also reliable.

4.4 | Emergence GWAS

The initial potato plant growth following tuber dormancy break (the vegetative growth stage) is not readily delimited into discrete phases, as the regulation of sprout initiation, elongation, emergence, and subsequent stem extension and canopy formation is controlled by a complex network of interconnected signaling pathways (Qu et al., 2024; Thornton, 2020). For this reason, we have included prior association studies on these related phenotypes in our review of known genetic associations to emergence, as only a few previous studies have analyzed the genetic underpinnings of potato sprouting and emergence. In our study, emergence was considered the first 0–20 days following emergence from the soil.

Several small-effect QTLs have previously been found for emergence on chromosomes 1, 3, 5, 8, 10, and 12 by multi-trait QTL linkage analysis, explaining between 1.4% and 6.7% of the observed phenotypic variance. The emergence QTLs were found to be pleiotropic with other developmental and agronomic traits, including the onset of growth, duration of flowering, the onset of senescence, tuber count, and overall yield (Hurtado-Lopez et al., 2015). Several unlinked QTLs for tuber sprout growth have been identified on chromosomes 2, 3, 4, 5, and 10 using a combination of whole-exome capture (WEC) and bulk-segregant analysis (BSA). The QTLs displayed either depressing or promoting effects on the phenotypic output (Sharma et al., 2021), similar to the opposite phenotypic effects of the QTL5 and QTL8 compared to QTL6 and QTL7 alternative alleles observed in this study. These studies confirm that potato emergence is a highly complex, polygenic trait. However, the narrow-sense heritability of 0.44 also indicates a breeding potential for this trait.

Our results contribute to the sparse knowledge of emergence QTLs in crop potatoes. The chromosome 2 QTL identified between 24.4 and 25.5 Mb with a WEC-BSA approach of high and low sprout growth rate by Sharma et al. (2021) largely coincides with QTL5 at 26.7 Mb identified in this study. Interestingly, Hurtado-Lopez et al. (2015) found significant emergence QTL effects on either arm of chromosome 5, similar to our results, where we observe a signal on the North arm of chromosome 5 at 11.5 Mb and one on the South arm of chromosome 5 at 37.1 Mb (QTL6-7). The WEC-BSA analysis of Sharma et al. (2021) only found one chromosome 5 QTL for sprout growth between 43.4 and 43.9 Mb, that is,

on the South arm of the chromosome, >6 Mb downstream of QTL7.

Using the WEC-BSA approach, Sharma et al. (2021) also identified an emergence QTL on chromosome 3 located at 42.2–42.5 Mb. This was somewhat proximal to our QTL8, peaking at 46.5 Mb.

Based on our emergence QTL locations, the created lists of candidate genes span 102, 61, 12, and 104 candidate genes for QTL5, QTL6, QTL7, and QTL8, respectively. For all QTLs, the lists include several genes annotated with DNA-binding domains that could be involved in transcriptional regulation as well as protein kinases and G-coupled receptors, possibly involved in signaling pathways. In addition to these entries, a subset of strong candidates (Supporting Information File S3) was identified based on annotation. For QTL5, these included senescence-associated genes, genes encoding proteins involved in plant stress tolerance and detoxification, for example, sulfiredoxin, proteins involved in intra- and intercellular trafficking, for example, phloem proteins 2-A1 and 2-A10, proteins involved in plant growth, for example, through brassinosteroid-promoted growth, protein subunits of photosystem I, development- and cell death-domain proteins, and effector proteins in siRNA-based gene silencing pathways (Coll-Garcia et al., 2004; Hernández Estévez & Rodríguez Hernández, 2020; Jia et al., 2019; Klahre et al., 1998; Liao et al., 2020; Tenhaken et al., 2005; Tripathy et al., 2021; Xian et al., 2006). While senescence is likely to be somewhat correlated with emergence, as they both govern complex plant developmental processes, we do not believe that QTL6 is linked to the maturity locus (*StCDF1*), as QTL6 is >6 Mb downstream of the locus and as we could accurately detect a QTL around the maturity locus (4.4–5.6 Mb, QTL2) for LB, known to be linked to maturity through the measure of leaf chlorophyll.

QTL7 is located in a relatively gene-sparse region compared to the other emergence QTLs, but several strong candidate genes surround the QTL, including a purple acid phosphatase gene. These proteins are involved in phosphorus assimilation from soil and remobilization within the plant tissue, which have also been coupled with developmental processes, including root development and senescence (Bhadouria & Giri, 2021). Candidate genes also included plant chitinases, involved in biotic and abiotic stress responses, phytoene desaturases that are part of light-harvesting photosynthetic carotenoid pigment biosynthesis, and AGAMOUS-like subfamily MADS-box transcription factor proteins that are key regulators of flower development (Dreni & Kater, 2014; Koschmieder et al., 2017; Vaghela et al., 2022). While several annotations for QTL8 candidate genes were at family level, candidate genes encoded proteins involved in compound transport, hormone signaling, signal transduction, and protein family members involved in plant development and biotic/abiotic stress response, for exam-

ple, a homeobox and a trithorax-related protein (Denancé et al., 2014; Hundertmark & Hinch, 2008; Liu et al., 2018; Mahajan et al., 2016; Stortenbeker & Bemer, 2019; Vigers et al., 1992; Xu et al., 2019). For QTL6, the majority of the genes were, however, uncharacterized beyond protein family or domain ontology, or could not be annotated (67%), hindering the identification of credible candidate genes based on annotation alone. For QTL5, QTL7, and QTL8 as well, 64%, 58%, and 76% of gene annotations were elusive, respectively. In summary, flanking regions of the emergence QTLs include candidate genes encoding proteins involved in plant growth and development, stress response, compound trafficking, and nutrient assimilation, which could plausibly influence this genetically complex developmental phenotype. However, the extensive lack of functional gene annotation for the emergence QTLs (58%–76%) presents a bottleneck in candidate gene identification and highlights the need for transcriptomics studies for early development stages, including emergence and sprouting, to facilitate identification of causative gene variation. Further research is needed regarding both the further delimitation of the emergence QTLs and the characterization of the genes likely to contribute to this complex phenotype.

4.5 | Aerial phenotyping in potato breeding and genomics

Aerial phenotyping using UAVs and high-resolution sensors promises great advantages in potato agriculture and breeding, as it allows high-throughput, cost-effective phenotyping of field trials and eliminates subjective biases that can be introduced in manual phenotyping (Yang & Zhai, 2022). By outsourcing the translation of image data into phenotypes to a commercial provider, this solution also demands a minimum of specialized expertise from the breeder. The real benefit of remote sensing is in scaling to the large number of plots found in commercial breeding programs. While simple, qualitative assessment (e.g., LB) on a 1–9 scale can be done in approximately 10 sec./plot, breeding programs routinely have thousands of plots to phenotype, rendering the total effort substantial, as phenotyping timescales linearly with plot number and number of traits to phenotype. In contrast, drone-based phenotyping time is largely independent of the number of plots, and completely independent of the number of traits, as the same image data is used to compute all traits of a trial. Instead, phenotyping time is dominated by a fixed preparation time per mission, assuming that subsequent image processing and interpretation are automated. This feature is key for the high-throughput phenotyping attainable using remote sensing.

Our results demonstrate that reliable phenotypes based on canopy characteristics can be used in quantitative genetic analysis of potato breeding populations, and project opportu-

nities of combining remote phenotyping data in both GWAS and genomic prediction of these and other plant-associated traits. Indeed, the high/moderate narrow-sense heritabilities of 0.60 for LB and 0.44 for emergence indicate that both traits are underscored by additive genetic variance that can be utilized in genomic prediction models for the traits based on the remote phenotyping data. This will facilitate the evaluation of much larger training populations in both GWAS and genomic prediction, contributing to improved model performance on the subject breeding populations (Desta & Ortiz, 2014).

However, we also note that particularly the GWAS of emergence indicated the presence of substantial data structuring that cannot readily be accounted for through the incorporation of terms capturing variation due to permanent environmental effects related to field layout, local spatial heterogeneity, or genetic population structure. Partitioning of the analysis into market groups indicated that the data structuring is somewhat retained only in the crisps and starch market group data based on genomic inflation factors of >1.40 (Table S4). The observed structure is, hence, likely not due to some inter-market group genetic population structure not captured by the model. It is also unlikely to have resulted from variability in seed tuber size, as this would, if at all, have resulted in systematic structure in all market groups. Instead, it could be related to the phenotypic data of these two market groups, which are compounded upon combination. The starch and crisps market groups were evaluated for emergence in two fields in 2023. In 2024, only starch group clones were trialed across two separate fields. Meanwhile, the remaining market groups were isolated in one field each year. Hence, the starch and crisp market groups have the most elaborate permanent environmental effects. However, the year $\times$ field $\times$ subtrial $\times$ lane covariate should account for such an effect. The inflation could also be related to the flight protocol, as different fields were assessed on different days and across the highest number of separate flight missions (Table 1). These flight protocols can introduce systematic variation across flights—strategies to properly test and account for such novel data structure need to be outlined to facilitate data integration. Alternatively, the inflation might be related to the high variance of senescence within the starch market group, which might also produce a broad emergence profile not accurately captured across the two to three flights. Nonetheless, we note that with the applied conservative correction of the observed genomic inflation, we may lose power of detection, but significant trait QTLs using aerial drone-based phenotyping data for the canopy-based traits LB and emergence are indeed reliable.

A particular challenge, however, remains for large-scale application of automated aerial phenotyping of root and tuber crops, including potatoes, as the underground roots and tubers cannot be assessed directly by aerial measures. Hence, some agronomic traits, such as tuber starch content, skin finish, and flesh color of potatoes will rely on either (1) lab-based

assessment following harvest of samples, or (2) evaluation by proxies of correlated traits associated with the plant canopy, as has previously been done for potato crop yield (B. Li et al., 2020). Further research is needed to establish reliable plant-associated phenotypes with significant correlation to tuber yield, physiology, and quality traits. Automatization of tuber quality trait phenotyping, such as tuber flesh and skin color, dry matter content, moisture, bruising, sprouting, and post-frying crisps and French fry color indices (sugar content), can also be achieved postharvest without extensive sample preparation using hyperspectral imaging detection of physical and chemical tuber characteristics of harvested tuber samples (Peraza-Alemán et al., 2024).

5 | CONCLUSIONS

Using an aerial phenotyping approach based on RGB image data collected using a drone UAV, we were able to identify a total of five QTLs associated with LB in potatoes and four QTLs influencing potato plant emergence. Three of the LB QTLs coincided with known *Rpi*-gene clusters in the potato genome, effectively validating the remote phenotyping approach for genomics-based analysis. The emergence QTLs also largely coincided with recently identified QTLs for this complexly governed plant trait, further delimiting the candidate gene regions of these loci—however, the underlying causative genetic variation of these QTLs that influence the complex, polygenic emergence trait remains unknown. The identified markers can be directly implemented in MAS models for these traits and further exploration of the causative variants underpinning each trait QTL can allude to future genome editing targets for crop optimization. This study demonstrates the power of remote phenotyping in the elucidation of the genetic variation underscoring established and novel plant-associated agronomic traits. The remote phenotyping technologies have the power to revolutionize modern breeding programs by providing access to unprecedented large-scale amounts of affordable, high-quality phenotypic data.

AUTHOR CONTRIBUTIONS

Trine Aalborg: Conceptualization; formal analysis; investigation; methodology; visualization; writing—original draft; writing—review and editing. **Mathias Timmermann Christiansen:** Data curation; methodology; writing—original draft; writing—review and editing. **Hélène Romé:** Formal analysis; methodology; writing—original draft; writing—review and editing. **Aakash Chawade:** Data curation; funding acquisition; writing—review and editing. **Merethe Bagge:** Conceptualization; data curation; funding acquisition; project administration; writing—review and editing. **Kåre Lehmann**

Nielsen: Conceptualization; funding acquisition; methodology; supervision; writing—review and editing.

ACKNOWLEDGMENTS

The authors have nothing to report.

CONFLICT OF INTEREST STATEMENT

M.T.C. and M.B. are employed by Danespo A/S, Denmark. H.R. is currently employed by SEGES Innovation P/S, Denmark. A.C. is partially employed by Phenoyard AB. K.L.N. is partially employed by KMC Amba, Denmark. The remaining authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The phenotypes, genotypes, and metadata used in the study are available at Zenodo repository at: <https://doi.org/10.5281/zenodo.17251323>. Due to proprietary rights of the applied custom SNP array and proprietary rights of the advanced breeding clones of Danespo A/S, the clone names and SNP positions are masked.

ORCID

Trine Aalborg  <https://orcid.org/0000-0002-5657-0091>

Hélène Romé  <https://orcid.org/0000-0002-0714-5961>

Aakash Chawade  <https://orcid.org/0000-0002-6500-4139>

Kåre Lehmann Nielsen  <https://orcid.org/0000-0002-5598-4401>

REFERENCES

- Abdullah, H. M., Mohana, N. T., Khan, B. M., Ahmed, S. M., Hossain, M., Islam, K. S., Redoy, M. H., Ferdush, J., Bhuiyan, M. A. H. B., Hossain, M. M., & Ahamed, T. (2023). Present and future scopes and challenges of plant pest and disease (P&D) monitoring: Remote sensing, image processing, and artificial intelligence perspectives. *Remote Sensing Applications: Society and Environment*, 32, 100996. <https://doi.org/10.1016/J.RSASE.2023.100996>
- Amadeu, R. R., Garcia, A. A. F., Munoz, P. R., & Ferrão, L. F. V. (2023). AGHmatrix: Genetic relationship matrices in R. *Bioinformatics*, 39, btad445. <https://doi.org/10.1093/BIOINFORMATICS/BTAD445>
- Bagchi, P., Sawicka, B., Stamenkovic, Z., Marković, D., & Bhattacharjee, D. (2024). Potato late blight outbreak: A study on advanced classification models based on meteorological data. *Sensors*, 24, 7864. <https://doi.org/10.3390/S24237864>
- Bao, Z., Li, C., Li, G., Wang, P., Peng, Z., Cheng, L., Li, H., Zhang, Z., Li, Y., Huang, W., Ye, M., Dong, D., Cheng, Z., VanderZaag, P., Jacobsen, E., Bachem, C. W. B., Dong, S., Zhang, C., Huang, S., & Zhou, Q. (2022). Genome architecture and tetrasomic inheritance of autotetraploid potato. *Molecular Plant*, 15, 1211–1226. <https://doi.org/10.1016/J.MOLP.2022.06.009>
- Bhadouria, J., & Giri, J. (2021). Purple acid phosphatases: Roles in phosphate utilization and new emerging functions. *Plant Cell Reports*, 41, 33–51. <https://doi.org/10.1007/S00299-021-02773-7>
- Carlsen, F. M., Johansen, I. E., Yang, Z., Liu, Y., Westberg, I. N., Kieu, N. P., Jørgensen, B., Lenman, M., Andreasson, E., Nielsen, K. L., Blennow, A., & Petersen, B. L. (2021). Strategies for efficient

- gene editing in protoplasts of *Solanum tuberosum* theme: Determining gRNA efficiency design by utilizing protoplast (research). *Frontiers in Genome Editing*, 3, 795644. <https://dx.doi.org/10.3389/fgeed.2021.795644>
- Carlsen, F. M., Westberg, I., Johansen, I. E., Andreasson, E., & Petersen, B. L. (2024). Strategies and protocols for optimized genome editing in potato. *The CRISPR Journal*, 8, 37–50. <https://doi.org/10.1089/CRISPR.2024.0068>
- Chang-Brahim, I., Koppensteiner, L. J., Beltrame, L., Bodner, G., Saranti, A., Salzinger, J., Fanta-Jende, P., Sulzbachner, C., Bruckmüller, F., Trognitz, F., Samad-Zamini, M., Zechner, E., Holzinger, A., & Molin, E. M. (2024). Reviewing the essential roles of remote phenotyping, GWAS and explainable AI in practical marker-assisted selection for drought-tolerant winter wheat breeding. *Frontiers in Plant Science*, 15, 1319938. <https://doi.org/10.3389/FPLS.2024.1319938/BIBTEX>
- Cobb, J. N., Juma, R. U., Biswas, P. S., Arbelaez, J. D., Rutkoski, J., Atlin, G., Hagen, T., Quinn, M., & Ng, E. H. (2019). Enhancing the rate of genetic gain in public-sector plant breeding programs: Lessons from the breeder's equation. *Theoretical and Applied Genetics*, 132, 627–645. <https://doi.org/10.1007/S00122-019-03317-0>
- Collard, B. C. Y., & Mackill, D. J. (2008). Marker-assisted selection: An approach for precision plant breeding in the twenty-first century. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363, 557–572. <https://doi.org/10.1098/rstb.2007.2170>
- Coll-Garcia, D., Mazuch, J., Altmann, T., & Müssig, C. (2004). EXORDIUM regulates brassinosteroid-responsive genes. *FEBS Letters*, 563, 82–86. [https://doi.org/10.1016/S0014-5793\(04\)00255-8](https://doi.org/10.1016/S0014-5793(04)00255-8)
- Coomber, A., Saville, A., & Ristaino, J. B. (2024). Evolution of *Phytophthora infestans* on its potato host since the Irish potato famine. *Nature Communications*, 15, 1–12. <https://doi.org/10.1038/s41467-024-50749-4>
- Covarrubias-Pazarán, G. (2016). Genome-assisted prediction of quantitative traits using the R package sommer. *PLoS ONE*, 11, e0156744. <https://doi.org/10.1371/JOURNAL.PONE.0156744>
- de Mendiburu, F., & Yaseen, M. (2020). *agricolae: Statistical procedures for agricultural research*. CRAN.
- Denancé, N., Szurek, B., & Noël, L. D. (2014). Emerging functions of nodulin-like proteins in non-nodulating plant species. *Plant and Cell Physiology*, 55, 469–474. <https://doi.org/10.1093/PCP/PCT198>
- Desta, Z. A., & Ortiz, R. (2014). Genomic selection: Genome-wide prediction in plant improvement. *Trends in Plant Science*, 19, 592–601. <https://doi.org/10.1016/j.tplants.2014.05.006>
- Devaux, A., Kromann, P., & Ortiz, O. (2014). Potatoes for sustainable global food security. *Potato Research*, 57, 185–199. <https://doi.org/10.1007/S11540-014-9265-1>
- Devlin, B., Roeder, K., & Wasserman, L. (2001). Genomic control, a new approach to genetic-based association studies. *Theoretical Population Biology*, 60, 155–166. <https://doi.org/10.1006/TPBI.2001.1542>
- Domínguez, F., & Cejudo, F. J. (2021). Chloroplast dismantling in leaf senescence. *Journal of Experimental Botany*, 72, 5905–5918. <https://doi.org/10.1093/JXB/ERAB200>
- Dreni, L., & Kater, M. M. (2014). MADS reloaded: Evolution of the AGAMOUS subfamily genes. *New Phytologist*, 201, 717–732. <https://doi.org/10.1111/NPH.12555>
- Duarte-Carvajalino, J. M., Alzate, D. F., Ramirez, A. A., Santa-Sepulveda, J. D., Fajardo-Rojas, A. E., & Soto-Suárez, M. (2018). Evaluating late blight severity in potato crops using unmanned aerial vehicles and machine learning algorithms. *Remote Sensing*, 10, 1513. <https://doi.org/10.3390/RS10101513>
- FAOSTAT. (2024). *Food and agriculture organization of the United Nations Statistics Division*. <https://www.fao.org/faostat/en/#data/QCL>
- Franceschini, M. H. D., Bartholomeus, H., van Apeldoorn, D. F., Suomalainen, J., & Kooistra, L. (2019). Feasibility of unmanned aerial vehicle optical imagery for early detection and severity assessment of late blight in potato. *Remote Sensing*, 11, 224. <https://doi.org/10.3390/RS11030224>
- Gallais, A. (2003). *Quantitative genetics and breeding methods in autopolyploid plants* (1st ed.). INRA.
- Gebhardt, C., Ballvora, A., Walkemeier, B., Oberhagemann, P., & Schüler, K. (2004). Assessing genetic potential in germplasm collections of crop plants by marker-trait association: A case study for potatoes with quantitative variation of resistance to late blight and maturity type. *Molecular Breeding*, 13, 93–102.
- Ghislain, M., Byarugaba, A. A., Magembe, E., Njoroge, A., Rivera, C., Román, M. L., Tovar, J. C., Gamboa, S., Forbes, G. A., Kreuze, J. F., Barekye, A., & Kiggundu, A. (2019). Stacking three late blight resistance genes from wild species directly into African highland potato varieties confers complete field resistance to local blight races. *Plant Biotechnology Journal*, 17, 1119–1129. <https://doi.org/10.1111/PBI.13042>
- Haas, B. J., Kamoun, S., Zody, M. C., Jiang, R. H. Y., Handsaker, R. E., Cano, L. M., Grabherr, M., Kodira, C. D., Raffaele, S., Torto-Alalibo, T., Bozkurt, T. O., Ah-Fong, A. M. V., Alvarado, L., Anderson, V. L., Armstrong, M. R., Avrova, A., Baxter, L., Beynon, J., Boevink, P. C., ... Nusbaum, C. (2009). Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. *Nature*, 461, 393–398. <https://doi.org/10.1038/nature08358>
- Hasi, Q., & Kakimoto, T. (2022). ROP interactive partners are involved in the control of cell division patterns in Arabidopsis leaves. *Plant and Cell Physiology*, 63, 1130–1139. <https://doi.org/10.1093/PCP/PCAC089>
- Haverkort, A. J., Boonekamp, P. M., Hutten, R., Jacobsen, E., Lotz, L. A. P., Kessel, G. J. T., Vossen, J. H., & Visser, R. G. F. (2016). Durable late blight resistance in potato through dynamic varieties obtained by cisgenesis: Scientific and societal advances in the DuRPh project. *Potato Research*, 59, 35–66. <https://doi.org/10.1007/s11540-015-9312-6>
- He, R., Zhuang, Y., Cai, Y., Agüero, C. B., Liu, S., Wu, J., Deng, S., Walker, M. A., Lu, J., & Zhang, Y. (2018). Overexpression of 9-cis-epoxycarotenoid dioxygenase cisgene in grapevine increases drought tolerance and results in pleiotropic effects. *Frontiers in Plant Science*, 9, 360507. <https://dx.doi.org/10.3389/fpls.2018.00970>
- Hernández Estévez, I., & Rodríguez Hernández, M. (2020). Plant glutathione S-transferases: An overview. *Plant Gene*, 23, 100233. <https://doi.org/10.1016/J.PLGENE.2020.100233>
- Hickey, J. M., Chiurugwi, T., Mackay, I., & Powell, W. (2017). Genomic prediction unifies animal and plant breeding programs to form platforms for biological discovery. *Nature Genetics*, 49, 1297–1303. <https://doi.org/10.1038/NG.3920>
- Hinrichs, A. L., Larkin, E. K., & Suarez, B. K. (2009). Population stratification and patterns of linkage disequilibrium. *Genetic Epidemiology*, 33, S88–S92. <https://doi.org/10.1002/gepi.20478>
- Hoopes, G., Meng, X., Hamilton, J. P., Achakkagari, S. R., de Alves Freitas Guesdes, F., Bolger, M. E., Coombs, J. J., Esselink, D., Kaiser,

- N. R., Kodde, L., Kyriakidou, M., Lavrijssen, B., van Lieshout, N., Shereda, R., Tuttle, H. K., Vaillancourt, B., Wood, J. C., de Boer, J. M., Bornowski, N., ... Finkers, R. (2022). Phased, chromosome-scale genome assemblies of tetraploid potato reveal a complex genome, transcriptome, and predicted proteome landscape underpinning genetic diversity. *Molecular Plant*, 15, 520–536. <https://doi.org/10.1016/j.molp.2022.01.003>
- Hörtensteiner, S. (2009). Stay-green regulates chlorophyll and chlorophyll-binding protein degradation during senescence. *Trends in Plant Science*, 14, 155–162. <https://doi.org/10.1016/j.tplants.2009.01.002>
- Hundertmark, M., & Hinch, D. K. (2008). LEA (late embryogenesis abundant) proteins and their encoding genes in *Arabidopsis thaliana*. *BMC Genomics*, 9, 1–22. <https://doi.org/10.1186/1471-2164-9-118>
- Hurtado-Lopez, P. X., Tessema, B. B., Schnabel, S. K., Maliepaard, C., Van der Linden, C. G., Eilers, P. H. C., Jansen, J., van Eeuwijk, F. A., & Visser, R. G. F. (2015). Understanding the genetic basis of potato development using a multi-trait QTL analysis. *Euphytica*, 204, 229–241. <https://doi.org/10.1007/S10681-015-1431-2/TABLES/2>
- Jennings, S. A., Koehler, A. K., Nicklin, K. J., Deva, C., Sait, S. M., & Challinor, A. J. (2020). Global potato yields increase under climate change with adaptation and CO₂ fertilisation. *Frontiers in Sustainable Food Systems*, 4, 248. <https://doi.org/10.3389/fsufs.2020.519324>
- Jeon, D., Kang, Y., Lee, S., Choi, S., Sung, Y., Lee, T. H., & Kim, C. (2023). Digitalizing breeding in plants: A new trend of next-generation breeding based on genomic prediction. *Frontiers in Plant Science*, 14, 1092584. <https://doi.org/10.3389/FPLS.2023.1092584/BIBTEX>
- Jia, Q., Kong, D., Li, Q., Sun, S., Song, J., Zhu, Y., Liang, K., Ke, Q., Lin, W., & Huang, J. (2019). The function of inositol phosphatases in plant tolerance to abiotic stress. *International Journal of Molecular Sciences*, 20, 3999. <https://doi.org/10.3390/IJMS20163999>
- Johansen, I. E., Liu, Y., Jørgensen, B., Bennett, E. P., Andreasson, E., Nielsen, K. L., Blennow, A., & Petersen, B. L. (2019). High efficacy full allelic CRISPR/Cas9 gene editing in tetraploid potato. *Scientific Reports*, 9(1), 1–7. <https://doi.org/10.1038/s41598-019-54126-w>
- Klahre, U., Noguchi, T., Fujioka, S., Takatsuto, S., Yokota, T., Nomura, T., Yoshida, S., & Chua, N. H. (1998). The Arabidopsis DIMINUTO/DWARF1 gene encodes a protein involved in steroid synthesis. *The Plant Cell*, 10, 1677–1690. <https://doi.org/10.1105/TPC.10.10.1677>
- Kloosterman, B., Abelenda, J. A., Gomez, M. D. M. C., Oortwijn, M., De Boer, J. M., Kowitzanich, K., Horvath, B. M., Van Eck, H. J., Smaczniak, C., Prat, S., Visser, R. G. F., & Bachem, C. W. B. (2013). Naturally occurring allele diversity allows potato cultivation in northern latitudes. *Nature*, 495(7440), 246–250. <https://doi.org/10.1038/nature11912>
- Koschmieder, J., Fehling-Kaschek, M., Schaub, P., Ghisla, S., Brausemann, A., Timmer, J., & Beyer, P. (2017). Plant-type phytoene desaturase: Functional evaluation of structural implications. *PLoS ONE*, 12, e0187628. <https://doi.org/10.1371/JOURNAL.PONE.0187628>
- Lenaerts, B., Collard, B. C. Y., & Demont, M. (2019). Review: Improving global food security through accelerated plant breeding. *Plant Science*, 287, 110207. <https://doi.org/10.1016/j.plantsci.2019.110207>
- Leonards-Schippers, C., Gieffers, W., Salamini, F., & Gebhardt, C. (1992). The R1 gene conferring race-specific resistance to *Phytophthora infestans* in potato is located on potato chromosome V. *Molecular & General Genetics*, 233, 278–283. <https://dx.doi.org/10.1007/BF00587589>
- Li, B., Xu, X., Han, J., Zhang, L., Bian, C., Jin, L., & Liu, J. (2019). The estimation of crop emergence in potatoes by UAV RGB imagery. *Plant Methods*, 15, 1–13. <https://doi.org/10.1186/S13007-019-0399-7/FIGURES/8>
- Li, B., Xu, X., Zhang, L., Han, J., Bian, C., Li, G., Liu, J., & Jin, L. (2020). Above-ground biomass estimation and yield prediction in potato by using UAV-based RGB and hyperspectral imaging. *ISPRS Journal of Photogrammetry and Remote Sensing*, 162, 161–172. <https://doi.org/10.1016/j.isprsjprs.2020.02.013>
- Li, X., Van Eck, H. J., Rouppe Van Der Voort, J. N. A. M., Huigen, D. J., Stam, P., & Jacobsen, E. (1998). Autotetraploids and genetic mapping using common AFLP markers: The R2 allele conferring resistance to *Phytophthora infestans* mapped on potato chromosome 4. *Theoretical and Applied Genetics*, 96, 1121–1128. <https://doi.org/10.1007/S001220050847>
- Liao, Z., Hodén, K. P., Singh, R. K., & Dixelius, C. (2020). Genome-wide identification of Argonautes in Solanaceae with emphasis on potato. *Scientific Reports*, 10(1), 1–10. <https://doi.org/10.1038/s41598-020-77593-y>
- LiLin-Yin (2023). CMplot: Circle Manhattan Plot (R Package Version 4.4.1) [Computer software]. CRAN.
- Liu, Y., Feng, H., Yue, J., Fan, Y., Jin, X., Zhao, Y., Song, X., Long, H., & Yang, G. (2022). Estimation of potato above-ground biomass using UAV-based hyperspectral images and machine-learning regression. *Remote Sensing*, 14, 5449. <https://doi.org/10.3390/RS14215449>
- Liu, Y., Zhang, A., Yin, H., Meng, Q., Yu, X., Huang, S., Wang, J., Ahmad, R., Liu, B., & Xu, Z. Y. (2018). Trithorax-group proteins ARABIDOPSIS TRITHORAX4 (ATX4) and ATX5 function in abscisic acid and dehydration stress responses. *New Phytologist*, 217, 1582–1597. <https://doi.org/10.1111/NPH.14933/bib>>
- Lizarazo, I., Rodriguez, J. L., Cristancho, O., Olaya, F., Duarte, M., & Prieto, F. (2023). Identification of symptoms related to potato Verticillium wilt from UAV-based multispectral imagery using an ensemble of gradient boosting machines. *Smart Agricultural Technology*, 3, 100138. <https://doi.org/10.1016/J.ATECH.2022.100138>
- Lo, S. F., Yang, S. Y., Chen, K. T., Hsing, Y. I., Zeevaart, J. A. D., Chen, L. J., & Yu, S. M. (2008). A novel class of gibberellin 2-oxidases control semidwarfism, tillering, and root development in rice. *The Plant Cell*, 20, 2603. <https://doi.org/10.1105/TPC.108.060913>
- Lokossou, A. A., Park, T.-H., Van Arkel, G., Arens, M., Ruyter-Spira, C., Morales, J., Whisson, S. C., Birch, P. R. J., Visser, R. G. F., Jacobsen, E., & Van Der Vossen, E. A. G. (2009). 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. *Molecular Plant-Microbe Interactions*, 22, 630–641. <https://dx.doi.org/10.1094/MPMI-22-6-0630>
- Mahajan, A. S., Kondhare, K. R., Rajabhoj, M. P., Kumar, A., Ghate, T., Ravindran, N., Habib, F., Siddappa, S., & Banerjee, A. K. (2016). Regulation, overexpression, and target gene identification of Potato Homeobox 15 (POTH15)—A class-I KNOX gene in potato. *Journal of Experimental Botany*, 67, 4255–4272. <https://doi.org/10.1093/JXB/ERW205>
- Maxwell, K., & Johnson, G. N. (2000). Chlorophyll fluorescence—A practical guide. *Journal of Experimental Botany*, 51, 659–668. <https://doi.org/10.1093/JEXBOT/51.345.659>
- Mayer, M. (2024). *missRanger: Fast imputation of missing values*. CRAN. <https://doi.org/10.32614/CRAN.PACKAGE.MISSRANGER>

- Nadakuduti, S. S., Buell, C. R., Voytas, D. F., Starker, C. G., & Douches, D. S. (2018). Genome editing for crop improvement—Applications in clonally propagated polyploids with a focus on potato (*Solanum tuberosum* L.). *Frontiers in Plant Science*, 871, 424117. <https://doi.org/10.3389/fpls.2018.01607>
- Njane, S. N., Tsuda, S., van Marrewijk, B. M., Polder, G., Katayama, K., & Tsuji, H. (2023). Effect of varying UAV height on the precise estimation of potato crop growth. *Frontiers in Plant Science*, 14, 1233349. <https://doi.org/10.3389/FPLS.2023.1233349>
- Paluchowska, P., Śliwka, J., & Yin, Z. (2022). Late blight resistance genes in potato breeding. *Planta*, 255, 127. <https://doi.org/10.1007/S00425-022-03910-6>
- Park, J. H., Wacholder, S., Gail, M. H., Peters, U., Jacobs, K. B., Chanock, S. J., & Chatterjee, N. (2010). Estimation of effect size distribution from genome-wide association studies and implications for future discoveries. *Nature Genetics*, 42, 570–575. <https://doi.org/10.1038/ng.610>
- Peraza-Alemán, C. M., López-Maestresalas, A., Jarén, C., Rubio-Padilla, N., & Arazuri, S. (2024). A systematized review on the applications of hyperspectral imaging for quality control of potatoes. *Potato Research*, 67(4), 1539–1561. <https://doi.org/10.1007/S11540-024-09702-7>
- Pham, G. M., Hamilton, J. P., Wood, J. C., Burke, J. T., Zhao, H., Vaillancourt, B., Ou, S., Jiang, J., & Buell, C. R. (2020). Construction of a chromosome-scale long-read reference genome assembly for potato. *GigaScience*, 9, 1–11. <https://doi.org/10.1093/gigascience/giaa100>
- Pilet, F., Pellé, R., Ellissèche, D., & Andrivon, D. (2005). Efficacy of the R2 resistance gene as a component for the durable management of potato late blight in France. *Plant Pathology*, 54, 723–732. <https://doi.org/10.1111/J.1365-3059.2005.01288.X>
- Posit Team. (2024). *RStudio: Integrated development environment for R*. CRAN.
- Qu, L., Huang, X., Su, X., Zhu, G., Zheng, L., Lin, J., Wang, J., & Xue, H. (2024). Potato: From functional genomics to genetic improvement. *Molecular Horticulture*, 4, 34. <https://doi.org/10.1186/S43897-024-00105-3>
- R Core Team. (2024). *R: A language and environment for statistical computing*. CRAN.
- Raffo, M. A., Sarup, P., Andersen, J. R., Orabi, J., Jahoor, A., & Jensen, J. (2022). Integrating a growth degree-days based reaction norm methodology and multi-trait modeling for genomic prediction in wheat. *Frontiers in Plant Science*, 13, 939448. <https://doi.org/10.3389/FPLS.2022.939448>
- Rietman, H., Bijsterbosch, G., Cano, L. M., Lee, H. R., Vossen, J. H., Jacobsen, E., Visser, R. G. F., Kamoun, S., & Vleeshouwers, V. G. A. A. (2012). Qualitative and quantitative late blight resistance in the potato cultivar Sarpo Mira is determined by the perception of five distinct RXLR effectors. *Molecular Plant-Microbe Interactions*, 25, 910–919. <https://doi.org/10.1094/MPMI-01-12-0010-R>
- Rodríguez, J., Lizarazo, I., Prieto, F., & Angulo-Morales, V. (2021). Assessment of potato late blight from UAV-based multispectral imagery. *Computers and Electronics in Agriculture*, 184, 106061. <https://doi.org/10.1016/J.COMPAG.2021.106061>
- Romé, H., Bagge, M., & Jensen, J. (2026). The genetic basis of resistance to late blight in tetraploid potatoes. *Crop Breeding, Genetics and Genomics*, 8, e260001. <https://doi.org/10.20900/cbagg20260001>
- Rosyara, U. R., De Jong, W. S., Douches, D. S., & Endelman, J. B. (2016). Software for genome-wide association studies in autopolyploids and its application to potato. *The Plant Genome*, 9, plantgenome2015.08.0073. <https://doi.org/10.3835/PLANTGENOME2015.08.0073>
- Sankaran, S., Quirós, J. J., Knowles, N. R., & Knowles, L. O. (2017). High-resolution aerial imaging based estimation of crop emergence in potatoes. *American Journal of Potato Research*, 94, 658–663. <https://doi.org/10.1007/S12230-017-9604-2>
- Selga, C., Koc, A., Chawade, A., & Ortiz, R. (2020). A bioinformatics pipeline to identify a subset of SNPs for genomics-assisted potato breeding. *Plants*, 10, 30. <https://doi.org/10.3390/PLANTS10010030>
- Sharma, S. K., McLean, K., Colgan, R. J., Rees, D., Young, S., Sønderkær, M., Terry, L. A., Turnbull, C., Taylor, M. A., & Bryan, G. J. (2021). Combining conventional QTL analysis and whole-exome capture-based bulk-segregant analysis provides new genetic insights into tuber sprout elongation and dormancy release in a diploid potato population. *Heredity*, 127, 253–265. <https://doi.org/10.1038/s41437-021-00459-0>
- Shi, L., de Biolley, L., Shaikh, M. A., de Vries, M. E., Mittmann, S. U., Visser, R. G. F., Prat, S., & Bachem, C. W. B. (2024). Aging later but faster: How StCDF1 regulates senescence in *Solanum tuberosum*. *New Phytologist*, 242, 2541–2554. <https://doi.org/10.1111/NPH.19525>
- Shi, Y., Han, L., Kleerekoper, A., Chang, S., & Hu, T. (2022). Novel CropdocNet model for automated potato late blight disease detection from unmanned aerial vehicle-based hyperspectral imagery. *Remote Sensing*, 14, 396. <https://doi.org/10.3390/RS14020396>
- Shi, Y., Thomasson, J. A., Murray, S. C., Pugh, N. A., Rooney, W. L., Shafian, S., Rajan, N., Rouze, G., Morgan, C. L. S., Neely, H. L., Rana, A., Bagavathiannan, M. V., Henrickson, J., Bowden, E., Valasek, J., Olsenholler, J., Bishop, M. P., Sheridan, R., Putman, E. B., ... Yang, C. (2016). Unmanned aerial vehicles for high-throughput phenotyping and agronomic research. *PLoS ONE*, 11, e0159781. <https://doi.org/10.1371/JOURNAL.PONE.0159781>
- Sishodia, R. P., Ray, R. L., & Singh, S. K. (2020). Applications of remote sensing in precision agriculture: A review. *Remote Sensing*, 12, 3136. <https://doi.org/10.3390/RS12193136>
- Spindel, J. E., Dahlberg, J., Colgan, M., Hollingsworth, J., Sievert, J., Staggenborg, S. H., Huttmacher, R., Jansson, C., & Vogel, J. P. (2018). Association mapping by aerial drone reveals 213 genetic associations for *Sorghum bicolor* biomass traits under drought. *BMC Genomics*, 19, 1–18. <https://doi.org/10.1186/S12864-018-5055-5>
- Stortenbeker, N., & Bemer, M. (2019). The SAUR gene family: The plant's toolbox for adaptation of growth and development. *Journal of Experimental Botany*, 70, 17–27. <https://doi.org/10.1093/JXB/ERY332>
- Sugiura, R., Tsuda, S., Tamiya, S., Itoh, A., Nishiwaki, K., Murakami, N., Shibuya, Y., Hirafuji, M., & Nuske, S. (2016). Field phenotyping system for the assessment of potato late blight resistance using RGB imagery from an unmanned aerial vehicle. *Biosystems Engineering*, 148, 1–10. <https://doi.org/10.1016/J.BIOSYSTEMSENG.2016.04.010>
- Sun, C., Zhou, J., Ma, Y., Xu, Y., Pan, B., & Zhang, Z. (2022). A review of remote sensing for potato traits characterization in precision agriculture. *Frontiers in Plant Science*, 13, 871859. <https://doi.org/10.3389/FPLS.2022.871859>
- Tenhaken, R., Doerks, T., & Bork, P. (2005). DCD—A novel plant specific domain in proteins involved in development and programmed cell death. *BMC Bioinformatics*, 6, 1–6. <https://doi.org/10.1186/1471-2105-6-169>

- The Potato Genome Sequencing Consortium. (2011). Genome sequence and analysis of the tuber crop potato. *Nature*, 475, 189–195. <https://doi.org/10.1038/nature10158>
- Thornton, M. (2020). Potato growth and development. In J. C. Stark, M. Thornton, & P. Nolte (Eds.), *Potato production systems* (19–33). Springer. <https://doi.org/10.1007/978-3-030-39157-7>
- Tripathy, M. K., Deswal, R., & Sopory, S. K. (2021). Plant RABs: Role in development and in abiotic and biotic stress responses. *Current Genomics*, 22, 26. <https://doi.org/10.2174/1389202922666210114102743>
- Turner, S. D. (2018). qqman: An R package for visualizing GWAS results using Q-Q and Manhattan plots. *The Journal of Open Source Software*, 3, 731. <https://doi.org/10.21105/joss.00731>
- Uffelmann, E., Huang, Q. Q., Munung, N. S., de Vries, J., Okada, Y., Martin, A. R., Martin, H. C., Lappalainen, T., & Posthuma, D. (2021). Genome-wide association studies. *Nature Reviews Methods Primers*, 1(1), 1–21. <https://doi.org/10.1038/s43586-021-00056-9>
- Vaghela, B., Vashi, R., Rajput, K., & Joshi, R. (2022). Plant chitinases and their role in plant defense: A comprehensive review. *Enzyme and Microbial Technology*, 159, 110055. <https://doi.org/10.1016/J.ENZMICTEC.2022.110055>
- Van Dijk, M., Morley, T., Rau, M. L., & Saghai, Y. (2021). A meta-analysis of projected global food demand and population at risk of hunger for the period 2010–2050. *Nature Food*, 2, 494–501. <https://doi.org/10.1038/s43016-021-00322-9>
- VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. *Journal of Dairy Science*, 91, 4414–4423. <https://doi.org/10.3168/JDS.2007-0980>
- Vigers, A. J., Wiedemann, S., Roberts, W. K., Legrand, M., Selitrennikoff, C. P., & Fritig, B. (1992). Thaumatin-like pathogenesis-related proteins are antifungal. *Plant Science*, 83, 155–161. [https://doi.org/10.1016/0168-9452\(92\)90074-V](https://doi.org/10.1016/0168-9452(92)90074-V)
- Vos, P. G., Paulo, M. J., Voorrips, R. E., Visser, R. G. F., Van Eck, H. J., & Van Eeuwijk, F. A. (2017). Evaluation of LD decay and various LD-decay estimators in simulated and SNP-array data of tetraploid potato. *Theoretical and Applied Genetics*, 130, 123–135. <https://doi.org/10.1007/s00122-016-2798-8>
- Vossen, J. H., van Arkel, G., Bergervoet, M., Jo, K. R., Jacobsen, E., & Visser, R. G. F. (2016). The *Solanum demissum* R8 late blight resistance gene is an Sw-5 homologue that has been deployed worldwide in late blight resistant varieties. *Theoretical and Applied Genetics*, 129, 1785–1796. <https://doi.org/10.1007/S00122-016-2740-0>
- Wickham, H. (2016). *Ggplot2: Elegant graphics for data analysis*. Springer-Verlag.
- Xian, P. L., Xue, Y. L., Zhang, J., Zong, L. X., Liu, X., Huan, J. Q., & Dao, W. W. (2006). Molecular and functional characterization of sulfiredoxin homologs from higher plants. *Cell Research*, 16(3), 287–296. <https://doi.org/10.1038/sj.cr.7310036>
- Xu, X., Wan, W., Jiang, G., Xi, Y., Huang, H., Cai, J., Chang, Y., Duan, C. G., Mangrauthia, S. K., Peng, X., Zhu, J. K., & Zhu, G. (2019). Nucleocytoplasmic trafficking of the Arabidopsis WD40 repeat protein XIWI regulates ABI5 stability and abscisic acid responses. *Molecular Plant*, 12, 1598–1611. <https://doi.org/10.1016/J.MOLP.2019.07.001>
- Yang, W., & Zhai, R. (2022). What can aerial phenotyping do and bring to us (breeders)? *The New Phytologist*, 236, 1229. <https://doi.org/10.1111/NPH.18413>
- Yuan, J., Cheng, L., Wang, Y., & Zhang, F. (2024). Genome-wide association studies for key agronomic and quality traits in potato (*Solanum tuberosum* L.). *Agronomy*, 14, 2214. <https://doi.org/10.3390/AGRONOMY14102214>
- Zagorščak, M., Zrimec, J., Bleker, C., Nolte, N., Juteršek, M., Iva Ramšak, Z., Gruden, K., & Petek, M. (2024). Evidence-based unification of potato gene models with the UniTato collaborative genome browser. *Frontiers in Plant Science*, 15, 1352253. <https://doi.org/10.3389/FPLS.2024.1352253>
- Zhao, H., Duan, K. X., Ma, B., Yin, C. C., Hu, Y., Tao, J. J., Huang, Y. H., Cao, W. Q., Chen, H., Yang, C., Zhang, Z. G., He, S. J., Zhang, W. K., Wan, X. Y., Lu, T. G., Chen, S. Y., & Zhang, J. S. (2020). Histidine kinase MHZ1/OsHK1 interacts with ethylene receptors to regulate root growth in rice. *Nature Communications*, 11(1), 1–13. <https://doi.org/10.1038/s41467-020-14313-0>

SUPPORTING INFORMATION

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

How to cite this article: Aalborg, T., Christiansen, M. T., Romé, H., Chawade, A., Bagge, M., & Nielsen, K. L. (2026). Drone-imaging assisted genome-wide association studies reveal key quantitative trait loci for emergence and late blight resistance in tetraploid potato. *Crop Science*, 66, e70261. <https://doi.org/10.1002/csc2.70261>