

Genome-Wide Association Analysis of Late Blight Resistance Traits in Potato Germplasm Resources

Fang Wang

Academy of Agriculture and Forestry Science, Qinghai University

Meiling Zou

Hainan University

Long Zhao

Academy of Agriculture and Forestry Science, Qinghai University <https://orcid.org/0000-0002-1499-6713>

Huaqing Li

Academy of Agriculture and Forestry Science, Qinghai University

Zhiqiang Xia (✉ zqxia@hainanu.edu.cn)

Hainan University, Haikou

Jian Wang

Academy of Agriculture and Forestry, Qinghai University

Research article

Keywords: Potato, Late blight, GWAS, Population structure, Genetic diversity

Posted Date: December 1st, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-113697/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background: Potatoes are dicotyledonous plants of the genus *Solanum*, family Solanaceae, and contain large amounts of starch, proteins, and trace elements required by the human. Potato late blight is the main disease hindering potato production. In this study, *Phytophthora infestans* were used to quantify late blight resistance in 284 germplasm resources, and resistance genes were mined through genome-wide association analysis.

Results: The results showed that among the 284 potato germplasm resources, 37 showed immunity, 15 were highly resistant to late blight, 30 were moderately resistant to late blight, 107 were moderately susceptible to late blight, and 95 were highly susceptible to late blight. Through screening and filtering, 22,489 high-quality single-nucleotide polymorphisms (SNPs) and indels were obtained. Through population structure analysis and principal-component analysis, 284 germplasm resources were divided into eight subgroups, which was consistent with the results of the phylogenetic tree analysis. The genetic diversity index of the 284 potato germplasm resources was 0.2161, and the differentiation index of each subgroup was 0.0251-0.1489. A mixed linear model was built to perform an association analysis on the diameter of the lesions identified from isolated leaves of potato affected by late blight. The genes within 100 kb of both sides of the obtained significant SNP loci were searched and functionally annotated, and 18 candidate genes were obtained. Twenty-two candidate genes were obtained from the association analysis of disease resistance grade.

Conclusions: 284 potato germplasm resources were used to identify for *Phytophthora infestans* resistance. The potato germplasm resources were divided into 8 subgroups by population structure analysis, and the main differentiation among subgroups was moderate. Candidate genes were mined by genome-wide association analysis. The results of this study provides the foundation for the genetic improvement of potato varieties resistant to late blight.

Background

Potato (*Solanum tuberosum* L.) is one of the most important food crops in the world, its total output ranking third in the world [1]. Potato late blight is a fungal disease caused by *Phytophthora infestans*, and it is also the most serious disease in potato production [2]. The spores and sporangia of *Phytophthora infestans* can be spread widely and quickly through water and air currents and infect plant tissues within 7 to 10 days [3, 4]. It causes dead leaves and rotten tubers, which seriously affect the yield and quality of potatoes [5]. Outbreaks of late blight have caused more and more genotypes to become susceptible to late blight, and it often leads to the decline or loss of resistance to late blight in resistant varieties in production. Therefore, pathogens causing late blight are considered to have the highest evolutionary potential, the greatest risk, and the most ease at overcoming a single or a few disease resistance genes [6].

In traditional agriculture, the use of protective agents or systemic fungicides is the main method of controlling late blight [7]. However, pesticides not only pollute the environment but also accelerate the development of resistance and the mutation of pathogens. Transgenic technology can quickly obtain disease-resistant varieties, but the adoption rate of genetically modified crops is still very low. In addition, the application of transgenic technology to food security has always been controversial [8].

The concept of genome-wide association study (GWAS) was proposed by Risch (1996) [9]. Hansen (2001) first reported the application of a GWAS to the growth properties of wild beet [10]. Followed by Aranzana (2005)

verified the feasibility of GWAS in *Arabidopsis thaliana* by studying its flowering time and pathogen resistance [11]. This was the first complete genome-wide association analysis in plants. With the advancement of statistical models and sequencing technologies, GWAS has gradually become an important tool for analysing the genetic basis of complex traits in plants. GWAS has been done in rice [12, 13], corn [14], rape [15], sorghum [16], barley [17], and *Arabidopsis* [18], leading to the identification of multiple genes related to important agronomic traits. GWAS uses linkage disequilibrium (LD), that is, two loci in the genome with more or less statistical association in their sequences due to recombination history, to link genotype with phenotype. GWAS overcomes several shortcomings of quantitative trait locus mapping in that GWAS provides higher resolution, is less time-consuming, requires fewer resources, and gives more consideration to allelic diversity [19, 20, 21].

LD means that in a specific population, the probability of co-occurrence of a specific allele located at a certain locus and an allele of another locus is greater than that of two alleles randomly distributed in the population. It determines the accuracy of the association analysis and the number and density of molecular markers used. Researchers can use the LD decay distance to calculate the number of markers required for GWAS, then calculate the detection efficiency and accuracy of the GWAS: GWAS marker amount = genome size/LD decay distance. The selection of the population (the selection of germplasm materials), the hierarchical structure of the population, and the genetic drift of the population affect the accuracy of the association analysis.

With the advancement of high-throughput sequencing technology, GWAS has become an important means to identify the relationship between the target phenotypic traits and genetic markers or genes, as well as to detect allele loci that have a strong role in controlling the related traits [22]. In this study, a GWAS was performed on the late blight resistance results of 284 isolated samples of potato leaves to evaluate the genetic diversity, genetic structure, and genetic relationships within the experimental population and to screen the candidate genes related to late blight resistance of potato, thus laying the foundation for genetic methods to improve late blight resistance in potato.

Results

Genotype analysis of potato population

All SNPs and indels were filtered for minor allele frequency (MAF) > 0.05 and Hardy–Weinberg equilibrium P value > 0.001, and 20,382 high-quality SNPs and 2,107 indels were obtained. Annotation of the high-quality SNPs and indels showed that 18,683 (83.07%) were located in intergenic regions; 3,806 (16.92%) were located in the gene regions of the genome, of which 951 were located in un-transcribed regions, 2,796 were located in introns, and only 1,682 SNPs were in coding regions. In the coding regions, 771 SNPs produced silent mutations and 911 SNPs produced missense mutations, for ratio of 1.18 (Table 1).

Table 1
Summary of single-nucleotide polymorphisms and indels

Totals SNPs and indels	intergenic	untranslated region		intron	coding sequence			nonsyn/synratio
		3' UTR	5' UTR		Total	missense	synonymous	
22,489	18,683	538	413	2,796	1,682	771	911	1.18

Identification and analysis of the late blight resistance of potato leaves *in vitro*

Among the 284 materials tested, 37 germplasm resources were asymptomatic or had only necrotic spots at the inoculation site, so they had a disease severity grade of 0; 15 germplasm resources had a lesion diameter (d) \leq 5 mm, no chlorotic halo at the edge of the lesion, and thus a disease severity grade of 1; 30 germplasm resources had $5 \text{ mm} < d \leq 10 \text{ mm}$, the lesion spot was water-soaked with chlorotic halo on the edge, and so the disease severity grade was 2; 107 germplasm resources had $10 \text{ mm} < d \leq 20 \text{ mm}$, a water-soaked lesion with a chlorotic halo on the edge, and so a disease severity grade of 3; 95 germplasm resources had $d > 20 \text{ mm}$, a lesion spot covered with a uniformly thick layer of mildew, and thus a disease severity grade of 4 (Table S1).

Analysis of population structure

The ADMIXTURE software was used to analyse 22,489 high-quality SNPs and Indels, the largest cluster subgroup value (K) was assumed to be each integer from 1–12, and the cross-validation (CV) error of each K value was calculated (Fig. 1A). When K was 1–4, the CV error gradually increased. When K was greater than 4, the CV error dropped rapidly to a nadir at $K = 8$, and for $K > 8$ it gradually increased. Therefore, $K = 8$ was optimal, that is, the entire potato population was divided into eight subgroups.

PCA was performed using all high-quality SNPs and indels. The calculation and analysis process were performed by R software (Fig. 1C). After the analysis was completed, plots were generated by R. For plotting, the eight subgroups inferred by ADMIXTURE software were used for grouping. The results showed that the eight subgroups could basically be distinguished on the PC1 axis, and the clustering results were consistent with the population structure division. Potatoes are native to the Andes of South America, and the history of artificial cultivation can be traced back to southern Peru from 8,000 to 5,000 BC.

According to the Q value of each material in these eight subgroups, each material was classified into the subgroup with the largest Q value (Fig. 1B). Subgroup 1 to Subgroup 8 had 11, 33, 25, 16, 12, 30, 51, and 86 germplasm resources, respectively. The distribution of the eight subgroups showed difference on the PC1 axis, and the clustering results were consistent with the population structure division. The eight subgroups of potato could not all be clustered together on the phylogenetic tree.

(A) The population structure of 284 potato materials was analysed by using ADMIXTURE software. CV error was calculated when $K = 1–12$. (B) When $K = 8$. In this population structure, each individual is represented by a line with eight different colours. According to the proportion of colours, which subgroup the variety belongs to can be inferred. (C) PCA was performed on all 284 potato samples with high-quality polymorphic loci. Each dot represents a sample.

The neighbour joining method was used to construct a phylogenetic tree, and the tree diagram was drawn with iTOL software to explore the genetic relationships between the 284 potato germplasms. Overall, the clustering results were consistent with the division of the population structure: subgroup 1, subgroup 2, and subgroup 6

clustered together well, while samples of other subgroups could be clustered together, and there was certain cross-over between samples (Fig. 2).

Genetic diversity revealed by SNP markers

According to the 22,489 high-quality SNPs and indels data, the genetic diversity (π) of all 284 potato germplasm resource was 0.2161, and the genetic diversity index of eight subgroups was between 0.1638 and 0.2502. Among them, subgroup 8 had the lowest genetic diversity index (0.1638), and subgroup 6 had the highest (0.2502) (Table 2). These data show that there was rich genetic diversity in the 284 potato germplasm resources.

Table 2
Summary of genetic diversity (π)

Subgroup	1	2	3	4	5	6	7	8	Total
π	0.2378	0.2387	0.2353	0.2007	0.2080	0.2502	0.2130	0.1638	0.2161

The population pairwise F -statistics (F_{ST}), a measure of population differentiation, was used to evaluate the degree of difference between subgroups of the 284 potato germplasm resource (Table 3). It was found that the F_{ST} among the subgroups was between 0.0251–0.1489, and the subgroup 1 and the subgroup 8 had the highest F_{ST} (0.1489), and subgroup 3 and subgroup 7 had the lowest F_{ST} (0.0251). Subgroup 2 and Subgroup 3, Subgroup 2 and Subgroup 7, Subgroup 3 and Subgroup 7, Subgroup 3 and Subgroup 8, and Subgroup 7 and Subgroup 8 were relatively weakly differentiated, and their genetic relationships were relatively close, whereas there was a moderate degree of differentiation between other subgroups.

Table 3
Summary of population pairwise F -statistics (F_{ST})

Subgroup	1	2	3	4	5	6	7	8
1								
2	0.0920							
3	0.0881	0.0285						
4	0.1442	0.0602	0.0714					
5	0.1310	0.0802	0.0719	0.1228				
6	0.0737	0.0642	0.0661	0.1170	0.0936			
7	0.1076	0.0325	0.0251	0.0611	0.0726	0.0869		
8	0.1489	0.0600	0.0479	0.0972	0.1098	0.1051	0.0298	

GWAS analysis of late blight resistance of potato germplasms

With the obtained SNPs and indels, the mixed linear model analysis was used to perform correlation analysis on the lesion spot diameter identified on the isolated leaves affected by late blight (Fig. 3) and disease resistance grade (Fig. 4). For the lesion diameters, $P < 4.4 \times 10^{-7}$ was set as the threshold to screen significant loci, and 18

candidate genes were obtained after annotating genes located at or near the significant loci. For the disease resistance grade, $P < 5 \times 10^{-2}$ was set as the threshold to determine the significant loci, and 22 candidate genes were obtained after annotating genes located at or near the significant loci.

Potato genome-wide LD

The LD-decay curve was obtained by analysing the LD of all 284 potato germplasm resources using the 22,489 SNP markers obtained from the whole genome (Fig. 5). The results show that LD decreased as the physical distance between SNPs increased. When taking the coefficient of determination $r^2 = 0.1$ as the decay threshold, within this population the decay distances of subgroup 3 and subgroup 8 were approximately 0.9 kb, those of subgroup 5 and subgroup 7 were approximately 1 kb, those of subgroup 2 and subgroup 4 were approximately 1.1 kb, that of subgroup 1 and subgroup 6 were approximately 1.2 kb, and that of the entire population was approximately 0.9 kb. All of these are far lower than the LD-decay distances of cultivated rice (123 kb), cultivated soybean (133 kb), cultivated corn (30 kb), and cultivated cassava (8 kb) and slightly smaller than that of the maize inbred line population (1.5 kb) [23].

Candidate genes

Genome-wide association analysis of disease severity (lesion spot diameter)

We annotated the genes located in the significant loci and within 100 kb around these loci, and 18 candidate genes were found. The detailed information is shown in Table 4. Among these genes, PGSC0003DMG400028682 encodes chitinase 1, which is related to immune response; PGSC0003DMG400036902 encodes a glycine-rich protein involved in the defence ability of the cell wall; and PGSC0003DMG400027651, PGSC0003DMG401013892, and PGSC0003DMG400015828 encode ethylene receptor 2, ethylene response transcription factor 7, and ERF transcription factor 5, which are all related to the response to ethylene. Ethylene can coordinate with jasmonic acid during invasion of pathogens. Protein kinases such as those encoded by PGSC0003DMG400000043, PGSC0003DMG400016913, PGSC0003DMG400044850, PGSC0003DMG400007634, and PGSC0003DMG400006739 can catalyse protein phosphorylation and are regulatory factors mediating signal transduction in response to external stimuli. PGSC0003DMG400031455 and PGSC0003DMG400047159 encode late blight resistance proteins, and PGSC0003DMG400031878 encode NBS-LRR resistance proteins, which are directly related to potato late blight resistance. Among them, NBS-LRR proteins are the most important late blight resistance proteins. PGSC0003DMG400023057 and PGSC0003DMG400016899 are leucine-rich repeat sequence family proteins that may be involved in the specific recognition of pathogenic effectors.

Table 4
Candidate genes of significant association markers

CHR	Position	Candidate gene	Start	End	description
1	72247021	PGSC0003DMG400000043	72233966	72243677	Protein kinase
3	17084761	PGSC0003DMG400016913	17041192	17046193	3-phosphoinositide dependent protein kinase 1
		PGSC0003DMG400031455	17182186	17186379	Late blight resistance protein
5	5568522	PGSC0003DMG400023057	5564035	5567888	Leucine Rich Repeat family protein
	23042227	PGSC0003DMG400016899	23031367	23033031	Leucine Rich Repeat family protein
		PGSC0003DMG400016900	23089073	23090433	Gene of unknown function
6	935787	PGSC0003DMG400031878	920425	927075	NBS-LRR resistance protein
	17021861	PGSC0003DMG400006739	16991723	17001511	Serine/threonine-protein phosphatase PP1 isozyme 3
		PGSC0003DMG400043784	17040350	17041875	Gene of unknown function
	18227724	PGSC0003DMG400044850	18176603	18183507	Calcium-dependent protein kinase
7	51816228	PGSC0003DMG400027651	51816086	51822785	Ethylene receptor 2
8	11363802	PGSC0003DMG400047159	11343551	11343883	Late blight resistance protein
	37714191	PGSC0003DMG400007634	37744509	37747990	Serine/threonine protein kinase
11	6786565	PGSC0003DMG400028682	6788959	6789991	Chitinase 1
12	6421567	PGSC0003DMG401013892	6424671	6425243	Ethylene-responsive transcription factor 7
	12912464	PGSC0003DMG400036902	12954240	12957322	Glycine rich protein 2
	40498866	PGSC0003DMG400015828	40428885	40430635	ERF transcription factor 5
		PGSC0003DMG400046173	40523269	40524198	Gene of unknown function

Genome-wide association analysis of disease severity grade

By annotating genes in the significant loci and nearby, 22 candidate genes were found (Table 5). The finding of the five genes PGSC0003DMG400027651, PGSC0003DMG400000043, PGSC0003DMG400006739, PGSC0003DMG400047159, and PGSC0003DMG400028682 was consistent with the results of the genome-wide analysis of lesion diameter. PGSC0003DMG400025989 encode PGSC0003DMG400026048 are ERF1, and PGSC0003DMG400026821, PGSC0003DMG400027651, and PGSC0003DMG400036493 encode ethylene response transcription factor 4, ethylene receptor 2, and AP2/ERF domain-containing transcription factors, respectively, which are related to the response of ethylene. The six genes PGSC0003DMG400023584, PGSC0003DMG400019737, PGSC0003DMG400005532, PGSC0003DMG400008506,

PGSC0003DMG400033667, and PGSC0003DMG400016323 encode serine/threonine protein kinases. PGSC0003DMG400033661 encodes a mitogen-activated protein kinase. PGSC0003DMG400023346 encodes a serine/threonine protein phosphatase. Plant protein kinases can catalyse protein phosphorylation, and protein phosphorylation is the main method of signal transduction. PGSC0003DMG400019926 is a plant resistance protein that is directly involved in potato late blight resistance. PGSC0003DMG400047228, PGSC0003DMG400041609, PGSC0003DMG400042169, and PGSC0003DMG400033671 encode unknown proteins.

Table 5
candidate genes of significant association markers

CHR	Position	Candidate gene	Start	End	description
1	69527286	PGSC0003DMG400025989	69503694	69504447	ERF1
		PGSC0003DMG400026048	69548531	69549211	ERF1
	72247021	PGSC0003DMG400000043	72233966	72243677	Protein kinase
3	44871940	PGSC0003DMG400036493	44865282	44866451	AP2/ERF domain containing transcription factor
4	8095786	PGSC0003DMG400023584	8093725	8096358	Serine/threonine protein kinase
	9276044	PGSC0003DMG400019737	9274796	9280465	Serine/threonine protein kinase
	20429483	PGSC0003DMG400005532	20327311	20330083	Serine/threonine protein kinase
		PGSC0003DMG400047228	20522006	20522377	Gene of unknown function
5	16350485	PGSC0003DMG400041609	16320408	16320749	Gene of unknown function
		PGSC0003DMG400008506	16418030	16428188	Receptor serine threonine protein kinase
	31701650	PGSC0003DMG400033661	31687765	31689237	Mitogen-activated protein kinase
		PGSC0003DMG400042169	31960558	31964367	Gene of unknown function
	32532025	PGSC0003DMG400019926	32513930	32514640	Plant resistance protein
	51527765	PGSC0003DMG400023346	51518931	51524033	Serine/threonine protein phosphatase
6	7094249	PGSC0003DMG400033671	7082039	7082791	Gene of unknown function
		PGSC0003DMG400033667	7164679	7168294	Serine/threonine protein kinase
	17031409	PGSC0003DMG400006739	16991723	17001511	Serine/threonine-protein phosphatase PP1 isozyme 3
	39855777	PGSC0003DMG400016323	39852449	39854311	Serine/threonine protein kinase
7	49532127	PGSC0003DMG400026821	49531657	49532565	Ethylene-responsive transcription factor 4
	51816228	PGSC0003DMG400027651	51816086	51822785	Ethylene receptor 2
8	11363730	PGSC0003DMG400047159	11343551	11343883	Late blight resistance protein
11	6786565	PGSC0003DMG400028682	6788959	6789991	Chitinase 1

Expression patterns of candidate genes

Four candidate genes were randomly selected to verify their gene expression patterns in late blight-resistant varieties A1, CIP10-1, and 0422 – 19 and susceptible varieties D8, UK7, and FAVORITA by qRT-PCR. The green leaves around the plaque, after inoculation with *Phytophthora infestans* in vitro for five days, and leaves without inoculated were used. Then their relative expression levels were calculated (Fig. 6). The results showed that most of the candidate genes were up-regulated after inoculation. The expression levels of PGSC0003DMG400028682 and PGSC0003DMG401013892 in resistant varieties were higher than those in susceptible varieties, while the expression levels of PGSC0003DMG400000043 in susceptible variety UK7 were higher than those in resistant variety 0422 – 19, and PGSC0003DMG400006739 in susceptible variety Favorita was higher than that in resistant variety Favorita Sexual variety A1. The results showed that the resistance to late blight was a quantitative trait controlled by multiple genes.

Discussion

In our study, we evaluated the resistance to *Phytophthora infestans* in 284 Potato Germplasms. Of these, 37 Potato Germplasms showed the immunity. Population structure analysis was used to assign the 284 potato germplasms into 8 subpopulations. The high-quality SNP markers revealed the moderate level in the main differentiation among the eight subpopulations. At the same time, genome-wide association analysis was carried out to identify the significant loci for late blight resistance traits in germplasm resources and screen candidate genes.

Identification and analysis of late blight resistance of Germplasm Resources

In recent years, the plethora of reports presented the identification of Potato Resistance to late blight; however, there are some differences in the identification results. The mixed strains selected in this study have yet to be determined at physiologic race level. With our studies for many years in the laboratory (Ref or unpublished data), we discovered that the resistance of potato leaves to *Phytophthora infestans* in vitro was basically consistent with that at the adult stage, and therefore, the identification results of detached leaves could represent the resistance at the adult stage. Feiwuruita is a high susceptible late blight variety, which is again consistent with our result [24]. It is speculated that identification of late blight resistant germplasm resources and discovery of durable resistant materials will be an important issue in breeding the high-quality late blight resistant potato varieties in near future. In this study, 37 materials showed immunity and 15 germplasms conferred highly resistance to late blight disease which provided the valuable materials for late blight resistant potato breeding.

Analysis of potato population structure and genetic diversity

In this study, 284 potato germplasm resources were used for the association analysis of resistance to late blight, 137 of which were from the International Potato Center (CIP). By testing genetic markers for Hardy-Weinberg equilibrium and suballelic frequency filtering, 22,489 high-quality polymorphic loci were obtained. The stratification of population structure and the uneven distribution of alleles are important reasons for false associations between genotypes and traits [25]. The population structure analysis of 284 potato germplasm resources in this study found that when $K = 8$, the CV error value was the smallest. Therefore, the 284 potato germplasm resources were divided into eight subgroups. The genetic diversity index of the potato population was 0.2161, indicating that there was abundant genetic diversity. The F_{ST} among subgroups was mostly

between 0.05 and 0.15, and the differentiation between subgroups was mostly moderate, indicating that there was a certain degree of differentiation in germplasm resources, but the degree of differentiation was not high.

Analysis of candidate genes

Late blight is a serious disease worldwide, threatening the potato industry and food security. Late blight resistance potato breeding is very important in China and abroad, and breeding to have multiple disease resistance genes in the same variety is an important strategy to prevent late blight. So far, 11 broad-spectrum resistance genes (R1-R11) have been discovered, and these 11 major R genes have been successfully located on the potato genetic map [26, 27]. Combinations of different disease resistance genes can provide ideal late blight resistance [28]. Therefore, optimizing the known disease resistance gene combinations, making full use of R disease resistance genes with broad-spectrum resistance characteristics, and discovering new broad-spectrum and longer-lasting disease resistance genes from the abundant potato resources are the effective means for the cultivation of resistant potato varieties in the future [29]. The abundant wild resources of potato are the source of R genes. Potato contains many genes encoding cytosolic nucleotide binding site–leucine-rich repeat (NBS-LRR) resistance proteins [30]. The R gene can be introduced into cultivated varieties by crossing conventional varieties with wild species containing this gene, thus helping cultivated varieties to achieve durable resistance to late blight [31]. In the abundant potato resources, there are many unknown late blight resistance genes waiting to be discovered.

In this study, 18 and 22 candidate genes, respectively, were found by the whole-genome association analysis of the diameter of the lesion spots and the disease severity grade identified from isolated potato leaves affected by late blight *in vitro*. Among them, the candidate gene encoding chitinase 1 gene may be involved in the immune response, and chitinase can improve the resistance of plants to fungi [32, 33]. The whole-genome association analysis of lesion diameter and disease resistance grade found that three and five candidate genes were associated with ethylene response, respectively. Ethylene can coordinate with jasmonic acid to regulate and function together during pathogen invasion [34]. The correlation analysis of lesion diameter found that two late blight resistance proteins and one NBS-LRR resistance protein may be directly related to potato late blight resistance. NBS-LRR proteins are the most important late blight resistance proteins and are characterized by the same conserved structure containing an N-terminal leucine zipper or coiled coil, a nucleotide-binding site, and leucine-rich repeats. The cloned anti–late blight genes contain several highly conserved structures, such as a phosphate-binding domain (Ploop), a kinase-2 group, or a GLPL group, in the NBS region [35, 36]. This study provides genetic resources for follow-up research and lays the foundation for the genetic improvement of potato resistance to late blight.

Conclusions

The result showed that 37 potato germplasm resources were immuned to *Phytophthora infestans* and 15 were highly resistant to *Phytophthora infestans*. Eight subgroups were categorized according to population structure analysis. There was abundant genetic diversity among the subpopulations mainly with moderate differentiation. The whole genome association analysis revealed that there were 18 candidate genes in the lesion diameter and 22 candidate genes in disease resistance grade.

Methods

Sample collection

The 284 potato germplasm resources used for this GWAS were collected from the Qinghai Plateau Potato Experimental Station, including 103 from China, one from Australia, three from Belarus, five from Canada, five from the United Kingdom, 133 from the international Potato Center (CIP), two from Israel, one from the Netherlands, five from Russia, two from the United States, and 24 from unknown sources. The library label of these samples were recorded in Table S1. All materials were stored in Academy of Agriculture and Forestry Sciences, Qinghai University (E101°, N36°). No permission was required in collecting the plants. In this study, Fang Wang was responsible for the planting and identification of these samples.

Quantification of resistance to late blight in potato germplasm resources

Preparation of microbe mixture

The "Northwest Mixed Microbes" containing mixture of pathogenic microbes causing late blight was used for inoculation. The northwest mixture was composed of five strains of "1352", "1353", "1835", "1837" and "1839", all of which were presented by Professor Jiehua Zhu of Hebei Agricultural University. The surface of the susceptible potato pieces was sterilized with a 95% alcohol flame, cut into slices approximately 0.5 cm thick, and placed in a petri dish covered with moist filter paper. After the potato slices were inoculated with the pathogen, they were placed in an 18 °C incubator and incubated in the dark for the recovery and amplification of the microbes. After the sporangia were produced in large quantities (approximately 3 to 5 days), the sporangia were collected, put into sterile water, filtered through a nylon mesh to remove the hyphae, and made into a sporangia suspension, which was placed in a refrigerator at 4 °C for 2 to 3 h. After the zoospores were released, the sporangia and empty sporangia shells were removed by filtering through a steel mesh with a pore size of 12 µm, and then the concentration was adjusted to 5×10^4 /mL with a haemocytometer for future use.

Indoor inoculation and identification

The varieties to be tested were sown in the experimental field shed of Qinghai Plateau Potato Experimental Station (E101°, N36°). Each variety was sown in one row. The growth temperature was between 18 °C and 25 °C. Natural light and timely fertilization, watering, and weeding were provided. When the seedlings grew to have 7–10 compound leaves, the leaves were picked for *in vitro* inoculation and identification indoors.

Two layers of sterilized filter paper were spread in a humidified transparent box, and a certain amount of sterile water was added for moisturizing. The leaves of the potato variety to be tested were placed on the filter paper with the back of the leaves facing upwards, and a micropipette was used to draw up 15 µL of the zoospore suspension and add the liquid dropwise onto the back of each leaf. The humidified box was placed in an 18 °C incubator with light for humidified culture (8 h in darkness and 16 h in light every day). The disease condition was observed 5 days after inoculation, and data for statistical analysis were collected. Ten leaves were inoculated for each treatment.

Indoor investigation methods

Five days after inoculation, the different disease severity levels were determined according to the diameter of the lesion, and the diameter of the lesion was measured by the cross method. The diameter of the lesion = $1/2 \times$

(length + width) (unit: mm). The specific standards are shown in Table 6.

Table 6
Appraisal criteria for resistance of potato leaves in vitro against late blight

Resistant type	Grade of disease	Disease degree
Immune	0	There were no symptoms or necrotic spots on the inoculated sites.
Highly resistance	1	The diameter of the lesion (d): $d \leq 5$ mm; there was no chlorotic halo at the edge of the lesion; there was no mould layer on the lesion or there was little mould layer on the edge of the lesion.
Moderately resistance	2	The diameter of the lesion: $5 \text{ mm} \leq d \leq 10$ mm; the lesion is water immersed with chlorotic halo on the edge; uniform mold layer is produced on the lesion, but the mold layer is sparse.
Moderately Susceptible	3	The diameter of the lesion: $10 \text{ mm} \leq d \leq 20$ mm; the lesion was water soaked, with chlorotic halo on the edge; uniform mold layer was produced on the lesion, but the mold layer was thick.
Highly susceptible	4	The diameter of the lesion: $d > 20$ mm; uniform and dense mould layer appeared on the lesion.

DNA preparation and sequencing

The improved cetyl trimethylammonium bromide (CTAB) method [37] was used to extract genomic DNA from potato leaves. After detection and quantification of the concentration by 1% agarose gel electrophoresis, the working DNA solution was diluted to 100 ng/ μ l and stored at -20 °C. The amplified-fragment SNP and methylation (AFSM) approach [38] was then used to construct EcoRI-MspI and EcoRI-HpaII libraries of the 284 potato DNA samples. After the monoclonal detection met the requirements, the EcoRI-MspI and EcoRI-HpaII libraries were mixed into one library at a ratio of 1:1, and HiSeq 2500 was used to perform paired-end 150-bp sequencing on the constructed sequencing library.

SNP calling and annotation

We used a Perl script (<http://afsmseq.sourceforge.net/>) to filter the original sequencing data, count the total number of reads obtained from sequencing, assign the reads to each individual based on the barcodes designed using the AFSM technology, and count the number of reads in each individual. Bowtie2 software [39] was used to align the optimized sequencing reads to the potato DM reference genome (<http://solanaceae.plantbiology.msu.edu/>), and SAMtools [40] and VCFtools (<http://vcftools.sourceforge.net/>) were used to detect SNP and indel loci. Based on the potato DM reference genome v4.03, the snpEff software [41] was used to identify the mutation locations (intergenic region, untranslated region/UTR, gene upstream region, or gene downstream region), mutation types (synonymous, missense, frameshift, and non-frameshift), and annotate them at the same time.

Analysis of population structure and genetic diversity

We first used PHYLIP (<http://evolution.genetics.washington.edu/phylip.html>) to calculate the genetic distance matrix of the sample. We then used the Notepad ++ software to save the genetic distance matrix file in a suitable format. A phylogenetic tree was constructed using the neighbour joining method. After generating the tree file,

iTOL (<https://itol.embl.de/>) was used to draw the phylogenetic tree diagram. GCTA software was used to conduct principal-component analysis (PCA) of the potato population materials with the detected SNPs as inputs [42]. R software was then used to calculate the vector of each principal component and draw the PCA scatter plot. In addition, ADMIXTURE software [43] was used to analyse the population structure and estimate the optimal number of population subgroups. PLINK software [44] was used to adjust the input file format for ADMIXTURE software, and then we input the file. The K value range of the number of subgroups was set to 1–12, and the appropriate value of K for the number of subgroups was determined according to the obtained cross-validation error value. The genetic composition coefficient (Q) of each material in each subgroup was used to construct the population-genetic structure matrix. VCFtools software (<https://vcftools.github.io/index.html>) was used to calculate the genetic diversity (π) and population pairwise F -statistics (F_{ST}) [45]. According to Wright, when F_{ST} is equal to 0 or 1, it indicates that there is no differentiation or complete differentiation between subgroups, respectively. If $0 < F_{ST} < 0.05$, $0.05 \leq F_{ST} < 0.15$, $0.15 \leq F_{ST} < 0.25$, or $0.25 \leq F_{ST} < 1$, this indicates that the subgroups have weak, medium, strong, or very strong genetic differentiation, respectively [46]. In the entire group and each subgroup (determined by the population structure), the r^2 value was used to determine the genome-wide LD through pairwise comparisons between 22,489 SNP markers.

LD analysis

In the entire population and each subgroup (inferred using ADMIXTURE), the value of r^2 was used to evaluate the LD relationship between each pair of polymorphic sites throughout the genome, and the value of r^2 was calculated by PopLDdecay software [47] for high-quality SNPs after filtering. The genetic distance was sorted from small to large, and then the average value of LD r^2 in the segment was calculated to draw a scatter lot with a smooth curve. The genetic distance interval where the curve intersected with the straight line representing a non-collinear r^2 of 99% was the LD decay distance.

Association analysis

In this study, data from 22,489 high-quality SNPs and indels were typed to perform a genome-wide association analysis on the severity (lesion diameter) and resistance grade of this population. We used the compressed mixed linear model of TASSEL 5.0 software [48] for correlation analysis. The kinship matrix was obtained with the same software, and the threshold for the significance of the severity was set at $1/n$, where n represents the number of all SNPs. The threshold for the significance of resistance grade was set at 0.05. Inputting the physical location of the SNP in the potato genome and its P value, the qqman package of R software was used to draw the Manhattan plot. SAMtools was used to manually verify regions that were significantly correlated with the reordered read results of the potato reference genome PGSC_DM_v4.03.

Candidate gene screening

Based on potato SNP annotation and LD decay, and according to the functional annotation of the loci, the genes where the loci were located were used as the candidate genes. If a locus was located upstream and downstream of other genes at the same time, the upstream and downstream genes were also used as candidate genes. If a locus was located in the intergenic region, the upstream and downstream genes that were closest to the locus were used as candidate genes. A mixed linear model was used to perform association analysis on the traits of late blight.

qRT-PCR verification

Primers were designed based on the coding region sequence (CDS) of the candidate genes, and the actin gene was the internal reference (Table 7). A fluorescence reverse transcription kit (TaKaRa) was used to generate cDNA using 500 ng RNA as the template. A fluorescence quantitative RT-PCR kit (TaKaRa) was used to perform quantitative real-time polymerase chain reaction (qRT-PCR) with the thermal cycling programme of 95 °C 30 s followed by 40 cycles of 95 °C 5 s and 60 °C 30 s. Excel 2016 was used to sort and analyse the gene expression fluorescence qRT-PCR (Quantitative real-time polymerase chain reaction) data, and the $2^{-\Delta\Delta Ct}$ method was used to calculate the relative expression levels.

Table 7
qRT-PCR primers for the selected genes

Gene ID	Forward primer (5'–3')	Reverse primer (5'–3')
Action	AGATGCTTACGCTGGATGGAATGC	TTCCGGTGTGGTTGGATTCTGTTC
PGSC0003DMG400000043	AGCAGCTCAAGCACAGAACTCTTC	CACGCTCCACCTCCAATTCATCTC
PGSC0003DMG400028682	TGGGCATCAACATTAACCCCAA	CCAACCAGACAGGCTAGCCA
PGSC0003DMG401013892	CATGGGGTCGATTCTCTGCTGAG	CCCTCCGCTGTATCAAATGTACCC
PGSC0003DMG400006739	CTTACCACGTCCAAGTCCATCC	GCCCCACCCCTTTACATCTTTACC

Abbreviations

SNPs: Single-nucleotide polymorphisms; GWAS: genome-wide association study; LD: linkage disequilibrium; MAF: minor allele frequency; CV: cross-validation; FST: F-statistics; CIP: International Potato Center; CTAB: cetyl trimethylammonium bromide; AFSM: amplified-fragment SNP and methylation; CDS: coding region sequence; qRT-PCR: quantitative real-time polymerase chain reaction.

Declarations

Acknowledgements

We are thankful to Professor Jiehua Zhu of Hebei Agricultural University for providing the strain of *Phytophthora infestans*. We would like to thank Qinghai University and Hainan University for their supports.

Authors' contributions

FW, ZX and JW conceived and designed the experiments. FW, LZ and HL conducted the experiment work. ZX, MZ and LZ analyzed the experimental results. FW and MZ wrote the manuscript. ZX and JW reviewed, and contributed to improve it. FW reviewed the last version of the manuscript. All authors read and approved the final manuscript.

Funding

This work was supported by the National Natural Science Foundation of China (NO. 31660417), the Modern Agricultural Industry Technology System (NO. CARS-9), the Central Public-interest Scientific Institution Basal Research Fund for Chinese Academy of Tropical Agricultural Sciences (No. 1630052019022). Fang Wang, director of the funded project, mainly provides ideas and 284 potato germplasm resources. The funding body did not play any role in the design of the study nor in the collection, analysis and interpretation of data, nor in the writing of the manuscript.

Availability of data and materials

All data generated or analysed during this study are included in this published article and its additional files.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹ Academy of Agriculture and Forestry Sciences, Qinghai University, Xining 810016, China; ² National Key Laboratory of Sanjiangyuan Ecology and Plateau Agriculture and Animal Husbandry, Qinghai University, Xining 810016, China; ³ Key Laboratory of Qinghai-Tibet Plateau Biotechnology Ministry of Education, Xining 810016, China; ⁴ Qinghai Provincial Key Laboratory of Potato Breeding, Xining 810016, China; ⁵ Hainan University, Haikou 570228, China; ⁶ The Institute of Tropical Biosciences and Biotechnology (ITBB), CATAS, Haikou 571101, China.

References

1. Paul RJB, Glenn B, Brian F, Eleanor MG, Ingo H, John TJ, Ankush P, Mark AT, Lesley T, Ian KT. Crops that feed the world 8: Potato: are the trends of increased global production sustainable? *Food Security*. 2012;4(4):477-508.
2. Fry WE, Birch PRJ, Judelson HS, Grunwald NJ, Danies G, Everts KL, Gevens AJ, Gugino BK, Johnson DA, Johnson SB, McGrath MT, Myers KL, Ristaino JB, Roberts PD, Secor G, Smart CD. Five reasons to consider *Phytophthora infestans* a reemerging pathogen. *Phytopathology*. 2002;105(7):966-981.
3. Zwankhuizen MJ, Zadoks JC. *Phytophthora infestans*'s 10-year truce with Holland: a long-term analysis of potato late-blight epidemics in the Netherlands. *Plant Pathology*. 2002;51:413-423.

4. Enciso-Rodriguez F, Douches D, Lopez-Cruz M. Genomic selection for late blight and common scab resistance in tetraploid Potato (*Solanum tuberosum*). G3-GENES GENOMES GENETICS. 2008;8(7):2471-2481.
5. Wu EJ, Yang LN, Zhu W, Chen XM, Sheng LP, Zhen JS. Diverse mechanisms shape the evolution of virulence factors in the potato late blight pathogen *Phytophthora infestans* sampled from China. Scientific Reports. 2016;6:26182.
6. McDonald BA, Linde C. Pathogen population genetics, evolutionary potential, and durable resistance. Annu Rev Phytopathol. 2002;40(1):349-379.
7. Maurice S, Montes MS, Nielsen BJ, Bdker, L, Martin MD, Jnck CG, Kjller R, Rosendahl S. Population genomics of an outbreak of the potato late blight pathogen, *Phytophthora infestans*, reveals both clonality and high genotypic diversity. Molecular Plant Pathology. 2009;20(8):1134-1146.
8. Dibden J, Gibbs D, Cocklin C. Framing GM crops as a food security solution. Journal of Rural Studies. 2013;29:59-70.
9. Risch N, Merikangas K. The future of genetic studies of complex human diseases. Science. 1996;273(5281):1516-1517.
10. Hansen M, Kraft T, Ganestam S, Sll T, Nilsson NO. Linkage disequilibrium mapping of the bolting gene in sea beet using AFLP markers. Genetical Research. 2001;77(1):61-66.
11. Aranzana MJ, Kim S, Zhao K, Bakker E, Horton M, Jakob K, Lister C, Molitor J, Shindo C, Tang C, Toomajian C, Traw B, Zheng H, Bergelson J, Dean C, Marjoram P, Nordborg M, Genome-wide association mapping in *Arabidopsis* identifies previously known flowering time and pathogen resistance genes. PLoS Genetics, 2005;1(5):e60.
12. Wu JH, Feng FJ, Lian XM, Teng XY, Wei HB, Yu HH, Xie WB, Yan M, Fan PQ, Li Y. Genome-wide association study (GWAS) of mesocotyl elongation based on resequencing approach in rice. BMC Plant Biology. 2015;15(1):218.
13. Zhou Y, Tang H, Cheng MP, Dankwa KO, Chen ZX, Li ZY, Gao S, Liu YX, Jiang QT, Lan XJ, Pu ZE, Wei ZM, Zheng YZ, Hickey LT, Wang JR. Genome-wide association study for pre-harvest sprouting resistance in a large germplasm collection of Chinese wheat landraces. Frontiers in Plant Science. 2017;8(93):401.
14. Wang M, Yan JB, Zhao JR. Genome-wide association study (GWAS) of resistance to head smut in maize. Plant Science. 2012;196:125-131.
15. Hatzig SV, Frisch M, Breuer F, Nesi N, Ducournau S, Wagner MH, Leckband G, Abbadi A, Snowdon RJ. Genome-wide association mapping unravels the genetic control of seed germination and vigor in Brassica napus. Frontiers in Plant Science. 2015;6:221.
16. Morris GP, Ramu P, Deshpande SP, Hash CT, Shah T, Upadhyaya HD, Riera-Lizarazu O, Brown PJ, Acharya CB, Mitchell SE, Harriman J, Glaubitz JC, Buckler ES, Kresovich S. Population genomic and genome-wide association studies of agroclimatic traits in sorghum. Proceedings of the National Academy of Sciences. 2013;110(2):453-458.
17. Alqudah AM, Sharma R, Pasam RK, Graner A, Kilian B, Schnurbusch T. Genetic dissection of photoperiod response based on GWAS of pre-anthesis phase duration in spring barley. PLoS One. 2014;9(11):113-120.
18. Si W, Saleh A, Ivaro C.I, Corina MF, Marek M, Rik K, Joost BK, Alisdair RF, Lothar W, Yariv B. Combined use of genome-wide association data and correlation networks unravels key regulators of primary metabolism in

- Arabidopsis thaliana*[J]. PLoS Genetics. 2016;12(10):1-36.
19. Nordborg M, Weigel D. Next-generation genetics in plants. Nature. 2008;456:720–723.
 20. Zhu C, Gore M, Buckler E S, Yu JM. Status and prospects of association mapping in plants. Plant Genome. 2008;1:5-20.
 21. Korte A, Farlow A. The advantages and limitations of trait analysis with GWAS: a review. Plant Methods. 2013;9:9.
 22. Pradhan SK, Barik SR, Sahoo A, Mohapatra S, Nayak DK, Mahender A, Meher J, Anandan A, Pandit E. Population structure, genetic diversity and molecular marker trait association analysis for high temperature stress tolerance in rice. PLoS One. 2016;11(8): e0160027.
 23. Zhang SK, Chen X, Lu C, Ye JQ, Zou ML, Lu KD, Feng SB, Pei JL, Liu C, Zhou XC, Ma PA, Li ZG, Liu CJ, Liao Q, Xia ZQ, Wang WQ. (2018) Genome-Wide association studies of 11 agronomic traits in cassava (*Manihot esculenta* Crantz). Frontiers in Plant Science, 2018;9:503.
 24. Biao LI, Fei LI, Zun-Guo L. Control efficacy of frowncide and ranman against potato late blight. Guizhou Agricultural ences. 2011, 39(12):133-135.
 25. Flint-Garcia SA, Thornsberry JM, Buckler ES. Structure of linkage disequilibrium in plants. Annual review of plant biology. 2003;54:357-374.
 26. Li X, Eck HJV, Voort JNAMR, Huigen DJ, Stam P, Jacobsen E. 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. 1998;96(8):1121-1128.
 27. Ballvora A, Ercolano MR, Weiss J, Meksem K, Bormann CA, Oberhagemann P, Salamini F, Gebhardt C. The R1 gene for potato resistance to late blight (*Phytophthora infestans*) belongs to the leucine zipper/NBS/LRR class of plant resistance genes. The Plant Journal. 2002;30(3):361-371.
 28. Rietman H, Bijsterbosch G, Cano LM, Lee HR, Vossen JH, Jacobsen E, Visser RGF, Kamoun S, Vleeshouwers VGAA. Qualitative and quantitative late blight resistance in the potato cultivar sarpomir is determined by the perception of five distinct RXLR effectors. Molecular Plant-Microbe Interactions. 2012;25(7):910-919.
 29. Haverkort AJ, Hillier JG. Cool farm tool - potato: model description and performance of four production systems. Potato Research. 2011;54:355-369.
 30. Jupe F, Pritchard L, Etherington GJ, Mackenzie K, Cock PJ, Wright F, Sharma SK, Bolser D, Bryan GJ, Jones JD, Hein I. Identification and localisation of the NBS-LRR gene family within the potato genome. BMC Genomics. 2012;13:75.
 31. Halterman DA, Kramer L C, Wielgus S, Jiming J. Performance of transgenic potato containing the late blight resistance gene RB. Plant Disease. 2008;92(3):339-343.
 32. Tabei Y, Kitade S, Nishizawa Y, Kikuchi N, Kayano T, Hibi T, Akutsu K. Transgenic cucumber plants harboring a rice chitinase gene exhibit enhanced resistance to gray mold (*Botrytis cinerea*). Plant Cell Reports. 1998;17(3):159-164.
 33. Hiroshi A, Yoko N, Shigeru A. Enhanced Resistance against a Fungal Pathogen *Sphaerotheca humuli* in Transgenic Strawberry Expressing a Rice Chitinase Gene. Plant Biotechnology. 1997;14(3):145-149.
 34. Lieselotte1 DB, Monica H, David DV. Connecting growth and defense: The emerging roles of brassinosteroids and gibberellins in plant innate immunity. Molecular Plant. 2014;7(6):943-959.

35. Meyers BC, Dickerman AW, Michelmore RW, Sivaramakrishnan S, Sobral B W, Young ND. Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily. *Plant Journal*. 1999;20(3):317-332.
36. Traut TW. The functions and consensus motifs of nine types of peptide segments that form different types of nucleotide-binding sites. *European Journal of Biochemistry*. 1994;222(1):9-19.
37. Murray MG, Thompson WF. Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Research*. 1980;8(19): 4321-4326.
38. Xia ZQ, Zou ML, Zhang SK, Feng BX, WANG WQ. AFSM sequencing approach: a simple and rapid method for genome-wide SNP and methylation site discovery and genetic mapping. *Scientific Reports*. 2014;4:7300.
39. Langmead B, Salzberg. Fast gapped-read alignment with Bowtie 2. *Nature Methods*. 2012;9(4):357-359.
40. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. The sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25(16):2078-2079.
41. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land S J, Lu XY, Ruden DM. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly*. 2012;6(2):80-92.
42. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: A tool for Genome-wide complex trait analysis. *The American Journal of Human Genetics*. 2011;88(1):76-82.
43. Alexander D H, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*. 2009;19(9):1655-1664.
44. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira M A R, Bender D, Maller J, Sklar P, Bakker PIWD, Daly MJ, Sham PC. PLINK: A tool set for Whole-Genome association and Population-Based linkage analyses. *The American Journal of Human Genetics*. 2007;81(3):559-575.
45. Danecek P, Auton A, Abecasis G, Albers C A, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, McVean G, Durbin R. The variant call format and VCFtools. *Bioinformatics*. 2011;27(15):2156-2158.
46. Dunia P, Ram KB, Ric CHDV, Chris M, Richard V, Guusje B. The patterns of population differentiation in a *Brassica rapa* core collection. *Theoretical and Applied Genetics*. 2011;122(6):1105-1118.
47. Zhang C, Dong SS, Xu JY, He WM, Yang TL. PopLDdecay: a fast and effective tool for linkage disequilibrium decay analysis based on variant call format files. *Bioinformatics*. 2019;35(10):1786-1788.
48. Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics*. 2003;23(19):2633-2635.

Figures

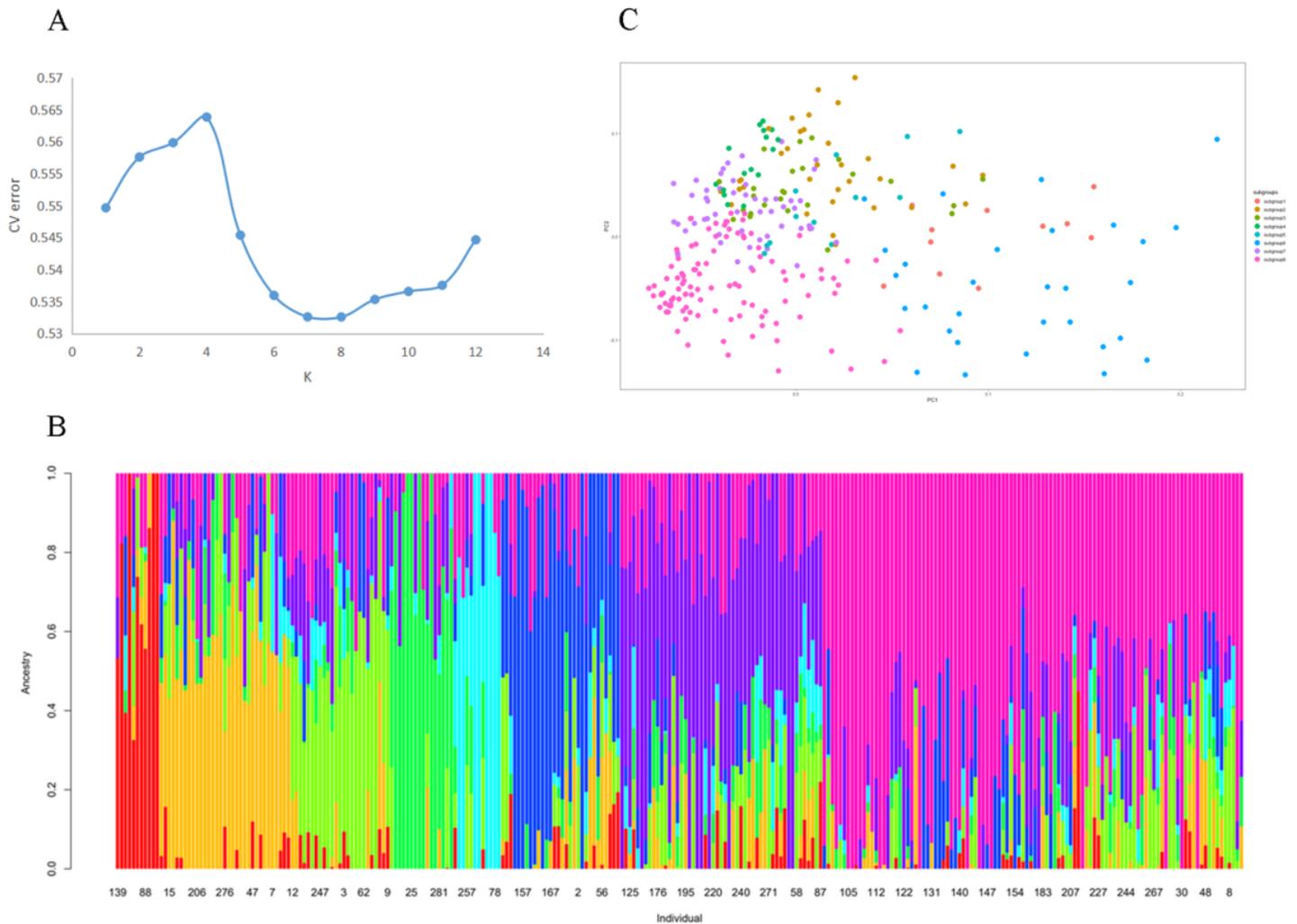


Figure 1

Population-genetic structure of 284 potato cultivars (A) The population structure of 284 potato materials was analysed by using ADMIXTURE software. CV error was calculated when K=1–12. (B) When K=8. In this population structure, each individual is represented by a line with eight different colours. According to the proportion of colours, which subgroup the variety belongs to can be inferred. (C) PCA was performed on all 284 potato samples with high-quality polymorphic loci. Each dot represents a sample.

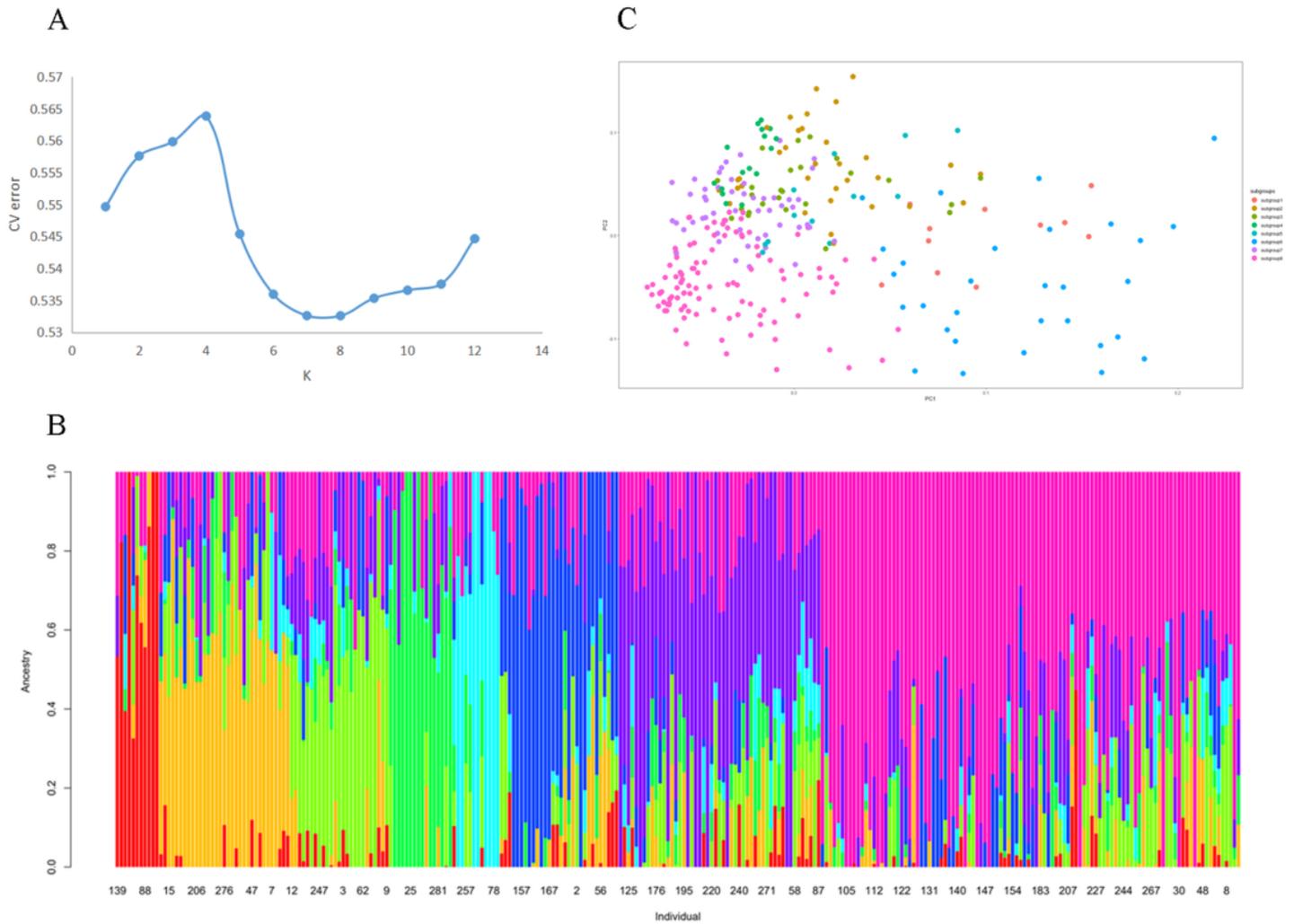


Figure 1

Population-genetic structure of 284 potato cultivars (A) The population structure of 284 potato materials was analysed by using ADMIXTURE software. CV error was calculated when K=1–12. (B) When K=8. In this population structure, each individual is represented by a line with eight different colours. According to the proportion of colours, which subgroup the variety belongs to can be inferred. (C) PCA was performed on all 284 potato samples with high-quality polymorphic loci. Each dot represents a sample.

Tree scale: 0.01 ⇐

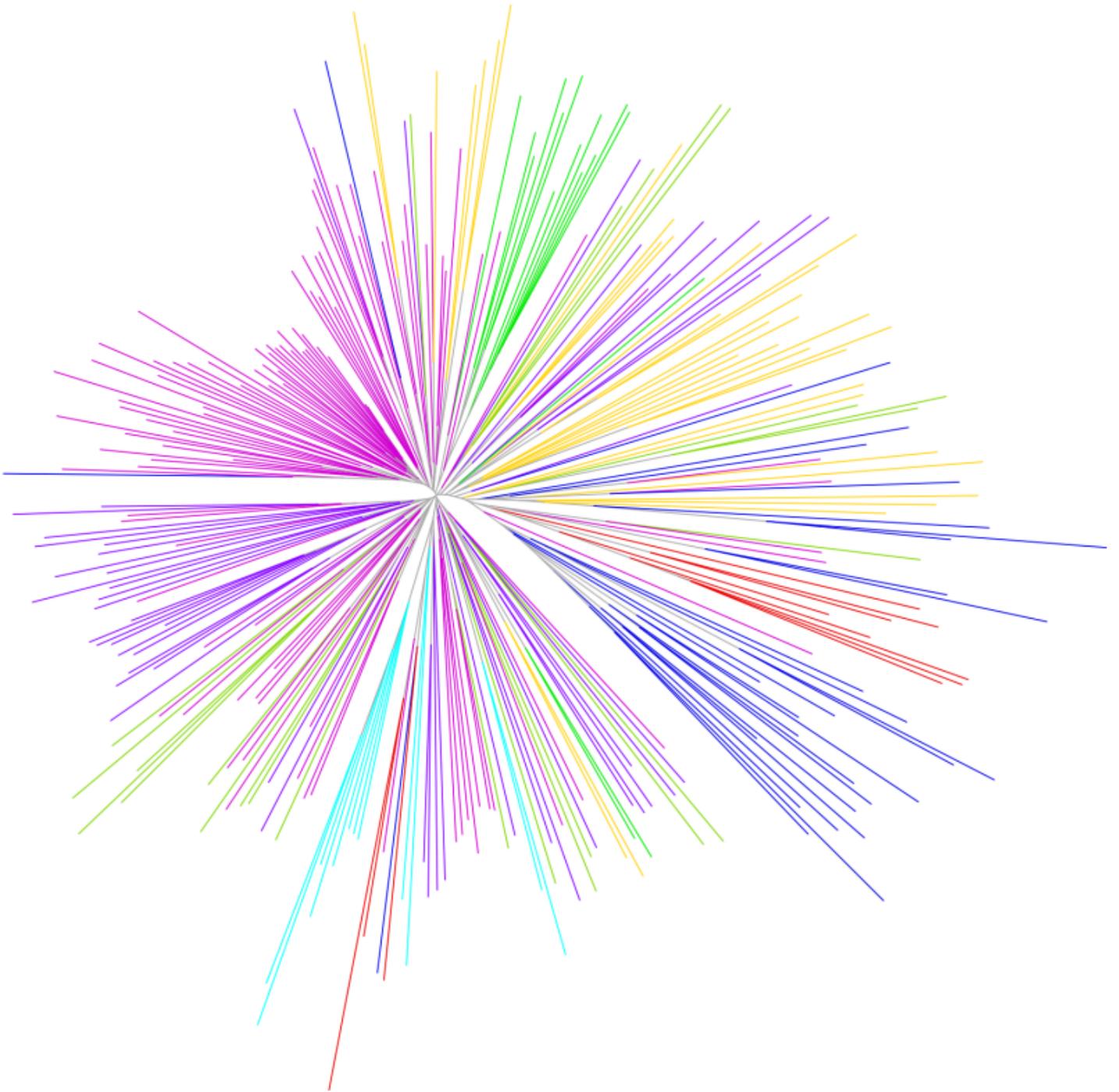


Figure 2

Phylogenetic tree constructed from 284 potato cultivars

Tree scale: 0.01 ⇄

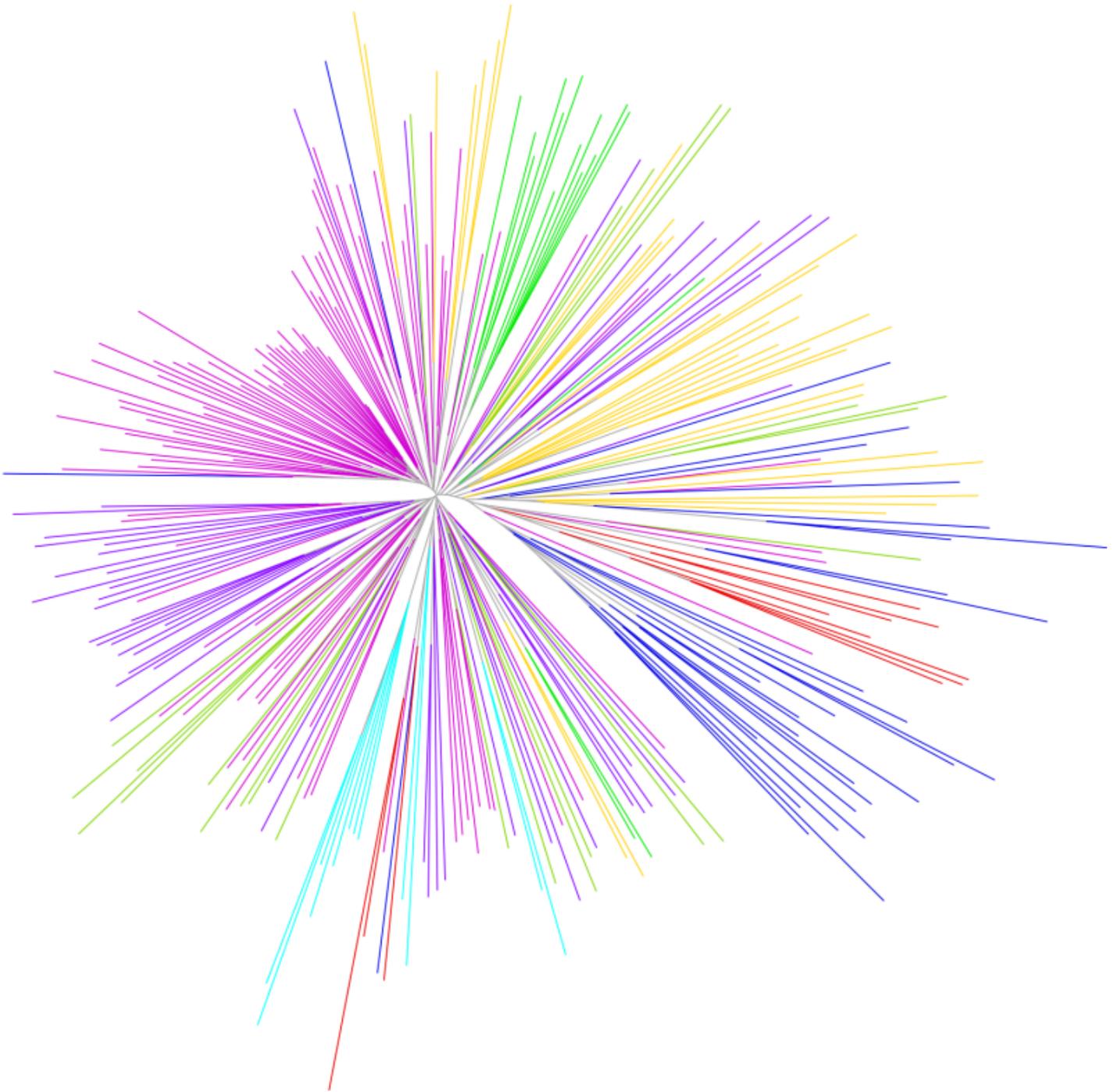


Figure 2

Phylogenetic tree constructed from 284 potato cultivars

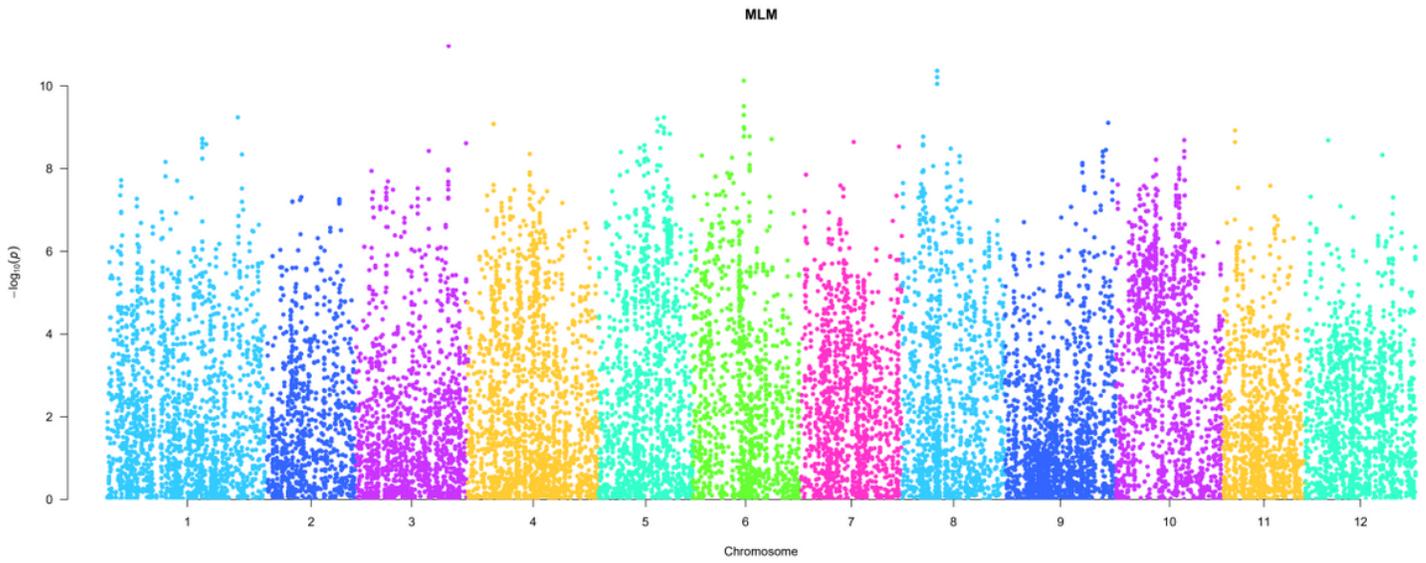


Figure 3

Manhattan plot of the whole-genome association analysis of disease severity

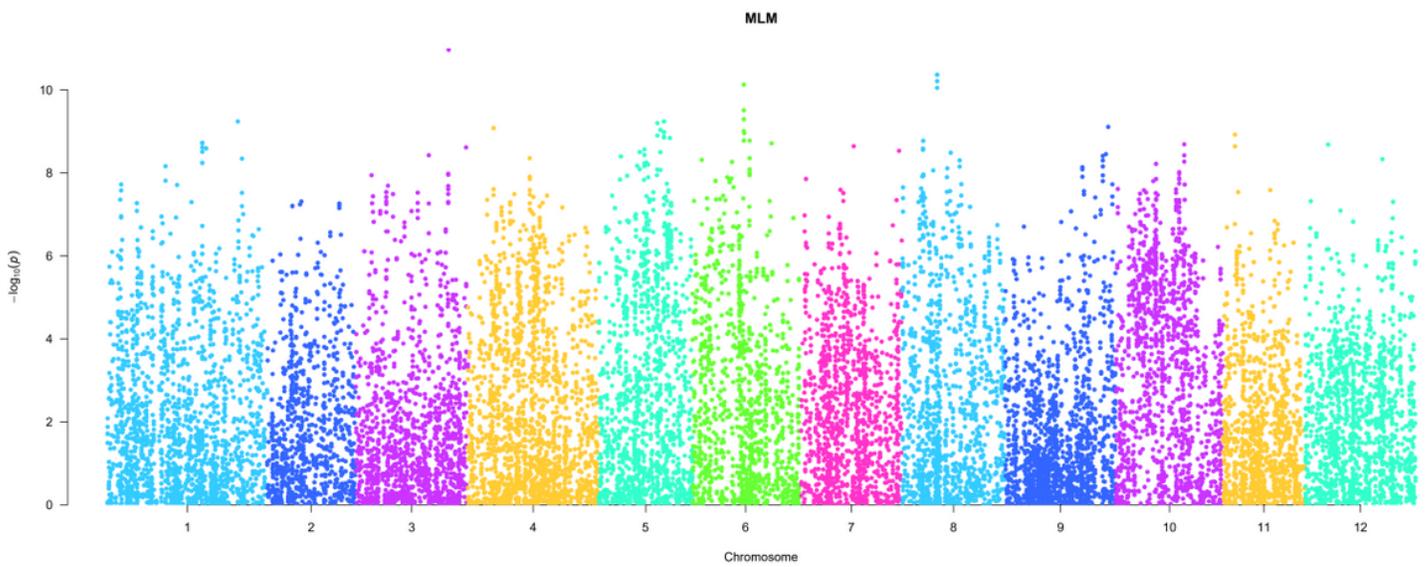


Figure 3

Manhattan plot of the whole-genome association analysis of disease severity

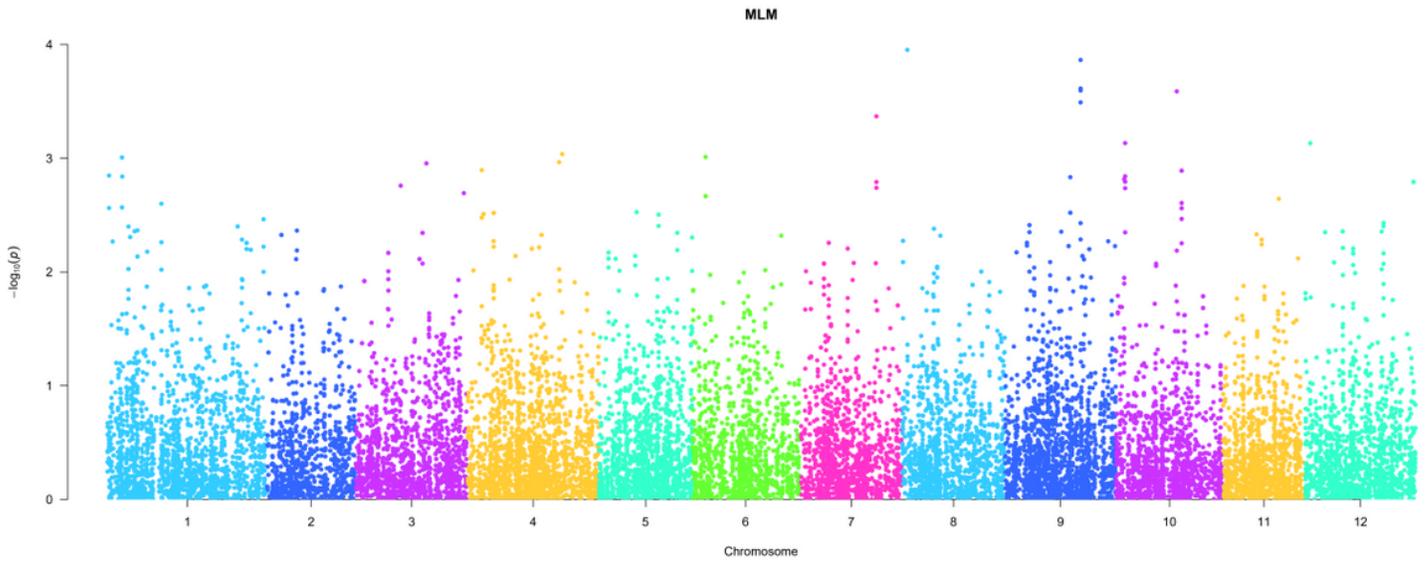


Figure 4

Manhattan plot of the whole-genome association analysis of disease grade

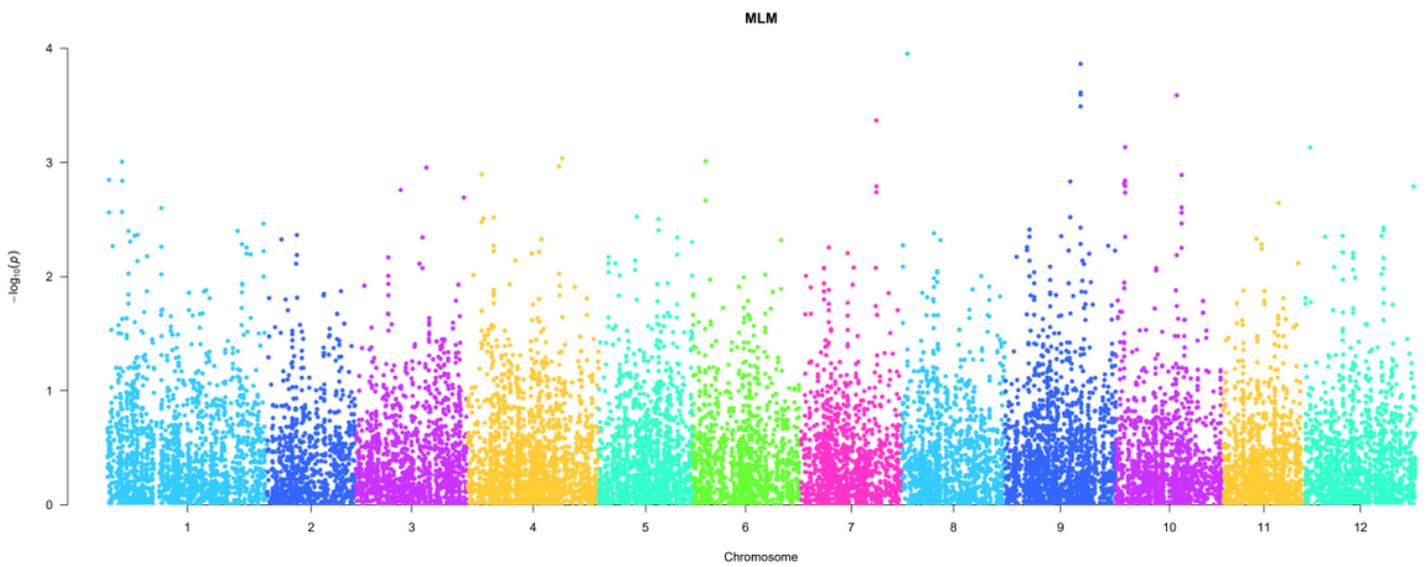


Figure 4

Manhattan plot of the whole-genome association analysis of disease grade

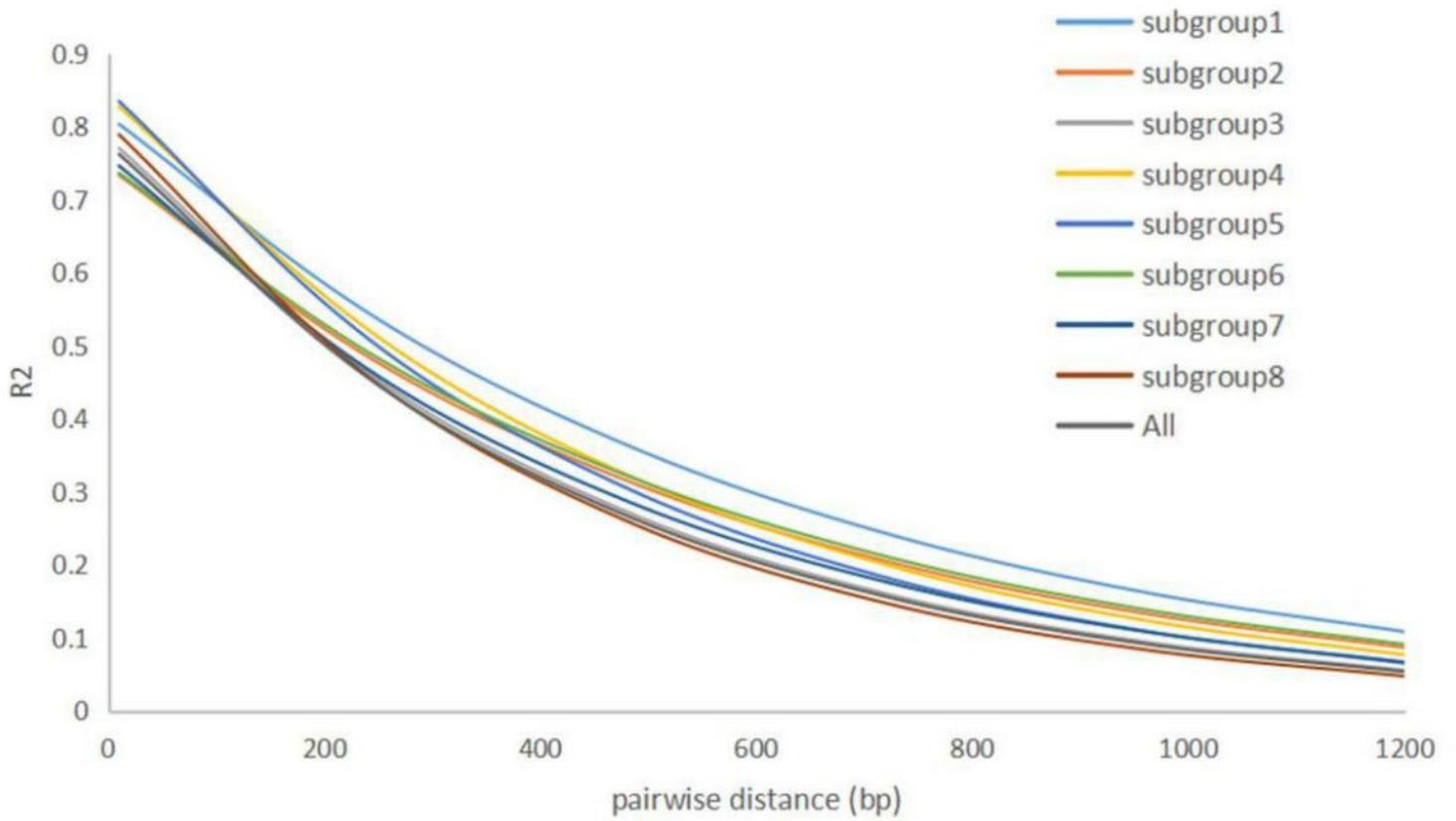


Figure 5

Manhattan plot of the whole-genome association analysis of disease grade

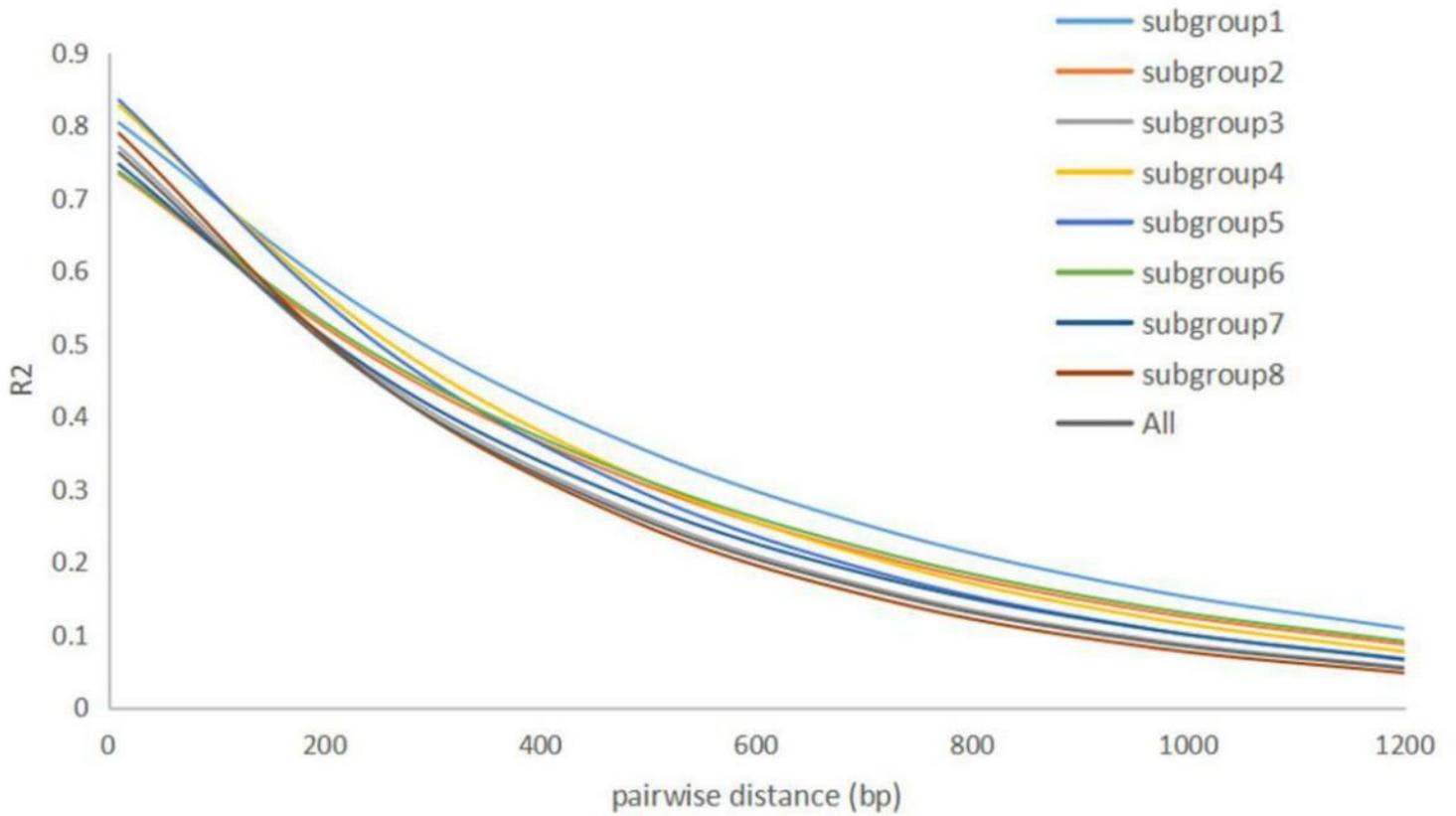


Figure 5

Manhattan plot of the whole-genome association analysis of disease grade

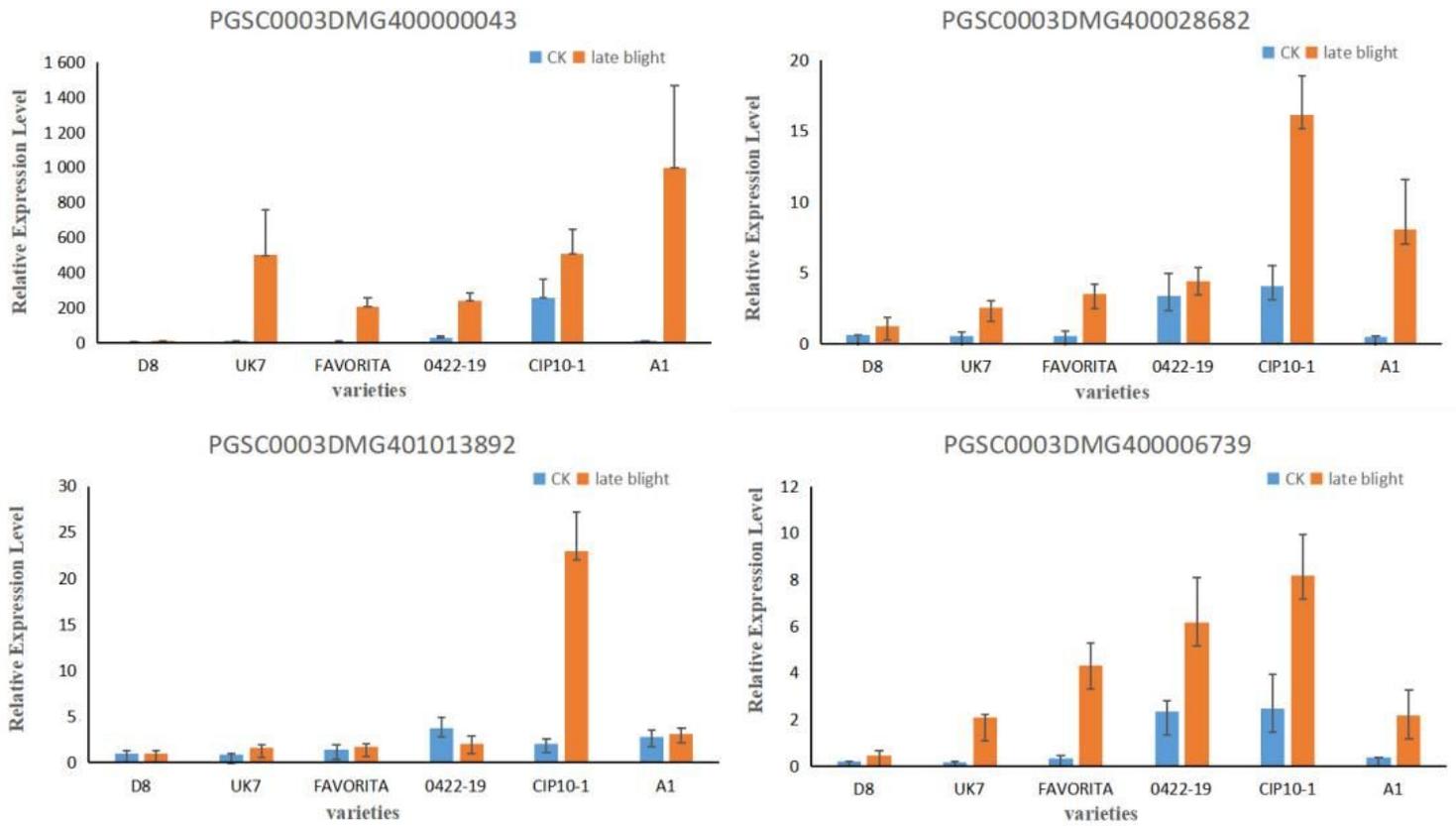


Figure 6

Experimental verification of gene expression levels by qRT-PCR

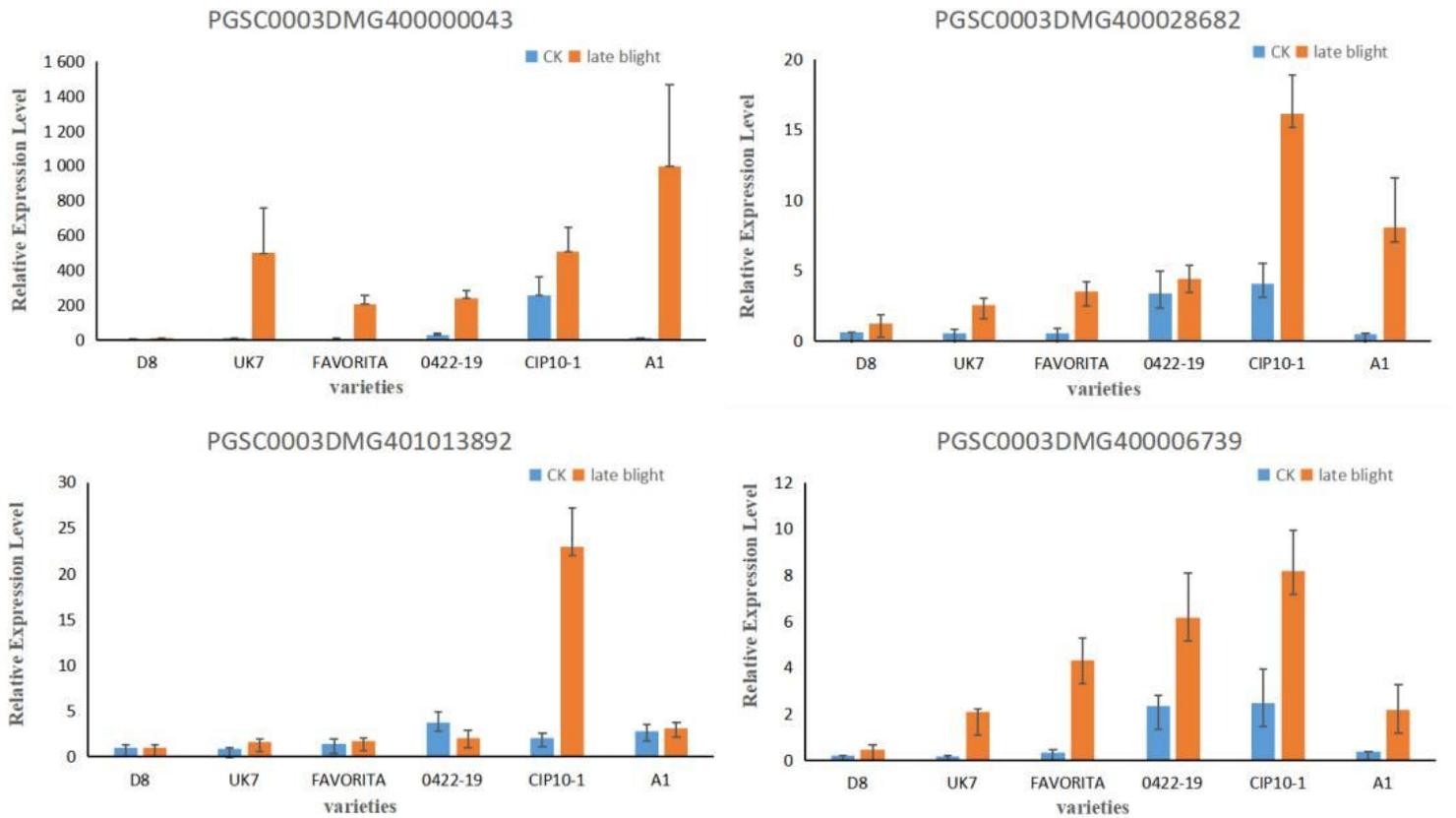


Figure 6

Experimental verification of gene expression levels by qRT-PCR

Supplementary Files

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

- [TableS1S2.pdf](#)
- [TableS1S2.pdf](#)