BMC Plant Biology

RESEARCH

Open Access

Genome-wide association study identifies candidate genes for glycoalkaloid biosynthesis in tetraploid potato (*Solanum tuberosum* L.) tubers

Xiaocheng Zhou[1], Jianlong Yuan[1], Lixiang Cheng[1], Lulu Xia[1], Zhensan Tang[1] and Feng Zhang[1*]

Abstract

Background Steroidal glycoalkaloids (SGAs), derived from cholesterol, act as natural defenses against pathogens and pests. In cultivated potatoes, α-solanine and α-chaconine are the primary SGAs, distributed throughout the plant, with their biosynthesis mechanisms differing across various tissues. The variation in SGAs content between the cortex and perimedullary zone reflects tissue-specific metabolic regulation in potato tuber. Higher SGAs levels in the cortex may enhance defense against external threats. This spatial distribution provides a theoretical basis for breeding strategies aimed at balancing resistance and food quality by regulating SGAs accumulation in specific tissues of potato tubers. Excessive levels of SGAs in potato tubers can compromise both their quality and edibility. Additionally, SGAs exhibit pharmacological properties, including anti-protozoal, antibacterial, antiviral, anti-tumor, and anti-inflammatory effects.

Results This study conducted genome-wide association study (GWAS) on SGAs content in the cortex and perimedullary zone of 117 diverse potato germplasm accessions, utilizing 22,983,689 high-quality SNPs. Candidate genes were subjected to analyses of stability, pleiotropy, GO and KEGG enrichment, and haplotype profiling. Twelve candidate genes associated with SGAs biosynthesis in potato tubers were identified, encoding UDP-glycosyltransferase superfamily proteins (*Soltu.DM.11G005750*, *Soltu.DM.11G005760*, *Soltu.DM.11G005770*, *Soltu.DM.11G005820*), fatty acid hydroxylase superfamily proteins (*Soltu.DM.01G029600*, *Soltu.DM.01G029610*, *Soltu.DM.01G029620*, *Soltu.DM.01G029640*, *Soltu.DM.01G029650*, *Soltu.DM.10G008360*), alkaline/neutral invertase (*Soltu.DM.11G006090*), and pleiotropic drug resistance (*Soltu.DM.11G006080*).

Conclusions This study provides a theoretical basis for elucidating the genetic mechanisms underlying SGAs biosynthesis in potatoes and will facilitate the breeding of new potato varieties.

Keywords Potato, SGAs, GWAS, Candidate genes

Background

Potato (*Solanum tuberosum* L.) is a perennial herbaceous dicot plant belonging to the *Solanaceae* family and the *Solanum* genus. It is an important crop that functions both as a staple food and an economic crop, playing a crucial role in ensuring food security. Potato tubers contain starch, protein, vitamins, amino acids, steroidal glycoalkaloids (SGAs), polyphenols, ascorbic

*Correspondence:
Feng Zhang
zhangf@gsau.edu.cn
[1] State Key Laboratory of Aridland Crop Science, College of Agronomy, Gansu Agricultural University, Lanzhou 730070, China

acid, carotenoids, and tocopherols [1]. SGAs are toxic nitrogen-containing secondary metabolites. Regulatory thresholds vary: the European Food Safety Authority (EFSA) recommends a maximum of 100 mg/kg fresh weight (FW), while the U.S. Food and Drug Administration (FDA) considers levels above 200 mg/kg FW or 1000 mg/kg dry weight (DW)potentially harmful [2]. When the intake reaches 2–3 mg/kg body weight (BW), it becomes harmful to humans, and exceeding 3 mg/kg BW can be lethal [3]. Excessive intake of SGAs can negatively affect the nervous and digestive systems [4]. Over 80 types of SGAs have been identified in potatoes, with α-solanine and α-chaconine accounting for more than 95% of the total SGAs in cultivated varieties [5–7]. SGAs are present in most tissues of potato plants, but their levels vary significantly across organs, tissues, and developmental stages. The biosynthesis of SGAs mainly occurs in young and actively growing tissues [6]. SGAs in potato tubers are primarily concentrated within the 1.5 mm-thick cortex, approximately 100 cell layers thick [8]. SGAs biosynthesis is influenced by both environmental factors and intrinsic genetic regulation [9, 10].

The biosynthesis of SGAs in potatoes is controlled by multiple genes. *HMGR*, *SQS* [11], and *SSR2* [12] are key genes involved in synthesizing the SGAs precursor, cholesterol, and indirectly affect SGAs biosynthesis. Previous studies has identified 10 genes involved in the biosynthesis of SGAs from cholesterol, including *GAME1* (*SGT1*), *GAME2* (*SGT3*), *GAME4* (*PGA4*), *GAME6*, *GAME7* (*PGA2*), *GAME8* (*PGA1*), *GAME11* (*16DOX*), *GAME12*, *GAME17* and *GAME18*, which are organized as clusters on chromosomes 7 and 12 [13]. Additionally, several transcription factors, including GAME9 [14], COI1 [15], TCP14 [16], HY5 [17] and MYB113 [18], regulate the biosynthesis of SGAs in potatoes. Regulatory genes for SGAs biosynthesis have also been identified in other *Solanum* plants, such as *GAME25* [19, 20], *S5αR* [21], *GAME31* [22], and *GAME5* [23] in tomato (*Solanum lycopersicum* L.) and *GAME15* [24] in black nightshade (*Solanum nigrum* L.).

Compared to traditional linkage analysis, Genome-wide association study (GWAS) eliminates the need for genetic population construction. It facilitates high-precision gene mapping with numerous molecular markers and identifies multiple allelic variations within a population, particularly in complex agronomic traits. Currently, GWAS is a key approach for identifying candidate genes associated with important agronomic and quality traits [25]. GWAS and domestication selection analysis of cultivated and wild potato populations identified *ERF9*, a domestication gene that negatively regulates SGAs metabolism in the potato pith and is located within the *ERF* gene cluster on chromosome 1 [26]. GWAS was performed on two tetraploid potato populations, identifying the candidate genes *Sga1.1*, *Sga3.1*, *Sga5.1*, *Sga7.1*, *Sga11.1*, *Sgr7.1,* and *Sgr8.1* associated with SGAs biosynthesis in tuber flesh. These genes located on chromosomes 1, 3, 5, 7, 8, and 11. Among them, *Sga1.1*, *Sga7.1*, *Sgr7.1*, and *Sgr8.1* co-localized with *GAME9*, *GAME6/GAME11*, *SGT1*, and *SGT2* [27]. However, these studies were conducted on whole potato tubers without distinguishing between the cortex and perimedullary zone, which differ significantly in secondary metabolites. A finer tissue classification of tuber is required for more precise analysis. There are marked differences in the expression levels of SGAs biosynthesis-related genes between cortex and perimedullary zone of potato tubers, owing to their distinct tissue-specific structures and functions. As the outer defensive tissue, the cortex generally exhibits higher SGAs biosynthesis activity and gene expression levels than that in perimedullary zone [26]. Consequently, the directly identifying differentially expressed genes in individual tissue based on GWAS results may be confounded by tissue-specific expression patterns, reducing the accuracy and reproducibility of gene discovery. In contrast, focusing on candidate genes commonly expressed in both tissues helps eliminate such biases, thereby improving the precision and reliability of key gene identification. This strategy lays a stronger foundation for subsequent functional validation and breeding applications.

This study utilized high-quality SNP markers obtained from whole-genome resequencing of 117 tetraploid potato cultivars and SGAs content data from the cortex and perimedullary zone tissues of tubers collected at two locations over one year to conduct a GWAS. This study aims to identify SNP loci and candidate genes significantly linked to SGAs biosynthesis, seeking novel genes from the cortex and perimedullary zone regions to further elucidate the genetic regulatory mechanisms of SGAs biosynthesis, thereby provide a reference for molecular breeding in potatoes.

Materials and methods

Plant materials and field trials

A set of 117 potato accessions, including cultivars and advanced lines, was used, including 83 advance lines from the International Potato Center (CIP), nine accessions from North America, nine from Europe, and sixteen Chinese cultivars (lines). In 2023, the materials were cultivated at two locations, Shandan (SD) and Weiyuan (WY). SD is located at 35°59′N, 101°13′E, 2,041 masl, and has average annual temperature of 8.5°C with average annual precipitation of 187 mm. WY is located at 35°02′N, 104°07′E, 2,450 masl, and has average annual temperature of 4.7°C with average annual precipitation

of 540 mm. Tubers were planted at the end of April and harvested at the end of September. The field trials at both locations followed a randomized block design using four replicates. With each cultivar (line) allocated four plots. Each plot was planted with 5 seed pieces, with 25 cm in-row spacing and 50 cm between rows. Seed tubers of all cultivars were planted in double rows on single ridges. Cultivation and field management practices were uniformly applied across all plots to ensure consistency. After harvest, mature, undamaged tubers without pest infestations and without tuber greening were selected for further experiments and data measurement.

SGAs content determination

Four tubers from each variety were selected, washed thoroughly, and dried. The skin was removed, and the tubers were sliced into 0.8 mm thick sections. Each slice was divided into cortex (periderm to vascular ring) and perimedullary zone (vascular ring to pith) tissue based on the vascular ring boundary. The samples were rapidly frozen in liquid nitrogen and stored at −80°C. The frozen samples were then freeze-dried using a SCIENTZ In-situ Freeze Dryer, ground into powder, and packaged for long-term storage at −80°C.

500 mg of freeze-dried powder was weighed into a 15 ml centrifuge tube, and 8 ml extraction solution (80% methanol + 0.1% formic acid) was added, and repeated three times per sample. The mixture was vortexed using a constant temperature shaker. The sample was then placed in an ultrasonic cleaner and subjected to ultrasonic treatment for 1 h at 25°C, with vortexing every 15 min. The sample was then centrifuged at 25°C and 10,000 rpm for 10 min in a desktop high-speed refrigerated centrifuge. The supernatant was collected (1 ml) and filtered through a 0.22 μm organic filter prior to analysis [26].

The concentration of α-solanine and α-chaconine was measured using a triple quadrupole mass spectrometer coupled with liquid chromatography (LC–MS). The chromatographic column used was a C18 column (100 mm × 2.1 mm, 1.8 μm), with mobile phase A comprising 100% methanol and mobile phase B containing 0.1% formic acid in water. The flow rate was set to 0.3 ml/min, with a column temperature of 35.0°C and UV detection at 200–210 nm using a 2998 PDA detector. The injection volume was 10 μl. Calibration curves from standards of α-solanine and α-chaconine were used to quantify the content (mg/kg) based on peak area. The content of α-solanine and α-chaconine was reported as mg/kg (PPM).

Phenotypic data analysis

Descriptive statistical analyses, including mean, standard deviation, and coefficient of variation (CV), were performed using Microsoft Office Excel 2016. Kolmogorov–Smirnov (Shapiro–Wilk tests) and homogeneity (Levene tests) of variance were assessed using SPSS 26.0. Analysis of variance (ANOVA) was performed using the least significant difference (LSD) method ($\alpha = 0.05$) to determine significance. Pearson correlation coefficients were calculated, and significance was evaluated using two-tailed tests, charts were plotted using Origin 2022, and ANOVA for genotype (G), environment (E), and genotype × environment interaction (G × E) as well as broad-sense heritability (H^2) for each trait were calculated using Genstat 21.0.

Genome-wide association study

The mixed linear model (MLM) in EMMAX software was used for the association analysis, where population structure (Q) was treated as a fixed effect and relatedness (K) as a random effect. This model reduces the influence of false positives from population structure and accounts for relatedness and population structure to improve result accuracy. Based on previously published potato GWAS studies, SNPs significantly associated with the target traits were identified using a significance threshold of -lg(P-value) > 6 [28, 29]. Using the double-haploid material DM1-3516 R44 as the reference genome, 117 tetraploid potato accessions were subjected to 150 bp paired-end sequencing on the Illumina HiSeq X Ten platform, with an average depth of 30 ×. Reads were aligned using BWA, and SNPs were filtered using GATK by excluding those with a minor allele frequency < 10% or a missing rate > 20%. In total, 22,983,689 high-quality SNPs were retained for subsequent GWAS analysis of tuber SGAs biosynthesis [30]. The model is as follows:

$$y = X\alpha + Z\beta + W\mu + e$$

Where: y: phenotypic value; X: fixed effect matrix; α: fixed effect parameter; Z: SNP matrix; β: SNP effect; W: random effect matrix; μ: predicted random individual effect; e: random residual.

Candidate gene mining

Candidate genes were functionally annotated based on their gene IDs in the Potato Genome Database (http://spuddb.uga.edu/). Subsequently, candidate genes were analyzed for stability (simultaneously regulating the biosynthesis of SGAs in both the cortex and perimedullary zone of potato tubers), pleiotropy (simultaneously regulating the synthesis of both α-solanine and α-chaconine in the potato tuber), GO and KEGG enrichment, and haplotype analysis of significant loci. Candidate gene selection was primarily based on tissue stability and pleiotropy, and further supported by GO annotation, KEGG pathway enrichment, and haplotype analysis. Genes

identified across two criteria were comprehensively compared to determine key candidates.

Results

Descriptive statistics and G × E effect variance analysis

In this study, the content of α-solanine and α-chaconine in the cortex and perimedullary zone of 117 potato tubers were measured (Table S1). The results showed that the α-solanine content in the tuber cortex ranged from 18.95 to 1123.49 mg/kg DW with coefficients of variation ranging from 90.79% to 92.68% and the α-chaconine content ranged from 34.22 to 974.28 mg/kg DW with coefficients of variation ranging from 57.94% to 66.79%. In the tuber perimedullary zone, α-solanine ranged from 6.29 to 868.72 mg/kg DW with coefficients of variation ranging from 140.40% to 144.86% and α-chaconine ranged from 13.93 to 678.23 mg/kg DW with coefficients of variation ranging from 99.36% to 102.10% (Table 1).

The genotype (E), environmental (G), and genotype × environment interaction (G × E) for the content of α-solanine α-chaconine in the cortex and perimedullary zone of potato tubers, showed extremely significant

Table 1 Descriptive statistical analysis of SGAs content in tubers from two locations and two tissue parts

Trait	parts	Environment	Max	Min	Mean ± SD	CV
α-Solanine (mg/kg DW)	cortex	SD	1123.49	19.87	164.36 ± 152.33	92.68%
		WY	934.95	18.95	148.15 ± 134.52	90.79%
	perimedullary zone	SD	868.72	11.44	67.33 ± 94.54	140.40%
		WY	635.28	6.29	53.89 ± 78.07	144.86%
α-Chaconine (mg/kg DW)	cortex	SD	974.28	34.22	214.39 ± 143.20	66.79%
		WY	639.54	40.57	173.67 ± 100.63	57.94%
	perimedullary zone	SD	678.23	13.93	86.54 ± 85.99	99.36%
		WY	453.96	14.04	62.78 ± 64.10	102.10%

Table 2 Analysis of variance for SGAs content in the cortex and perimedullary zone of potato tubers

Traits	Source of variation	Sum of square	Mean square	F-test	P-value	Percentage of sum of squares (%)	H^2
Cortex-(α-Solanine)	G	12,250,000.00	106,600.00	1535.29	< 0.001	84.81%	82.74%
	E	46,790.00	46,790.00	674.17	< 0.001	0.32%	
	G × E	2,115,000.00	18,390.00	264.96	< 0.001	14.64%	
	Residual	31,930.00	69.40				
	Total variation	14,443,720.00					
Cortex-(α-Chaconine)	G	8,629,000.00	75,030.00	964.28	< 0.001	78.57%	90.88%
	E	296,300.00	296,300.00	3807.73	< 0.001	2.70%	
	G × E	2,022,000.00	17,580.00	225.97	< 0.001	18.41%	
	Residual	35,790.00	77.81				
	Total variation	10,983,090.00					
Perimedullary Zone-(α-Solanine)	G	4,791,820.81	41,668.01	1638.03	< 0.001	90.89%	76.50%
	E	31,890.53	31,890.53	1253.66	< 0.001	0.60%	
	G × E	436,834.66	3798.56	149.33	< 0.001	8.29%	
	Residual	11,701.45	25.44				
	Total variation	5,272,247.45					
Perimedullary Zone-(α-Chaconine)	G	3,507,741.10	30,502.10	993.67	< 0.001	85.30%	86.01%
	E	100,295.69	100,295.69	3267.35	< 0.001	2.44%	
	G × E	490,029.53	4261.13	138.82	< 0.001	11.92%	
	Residual	14,120.32	30.70				
	Total variation	4,112,186.64					

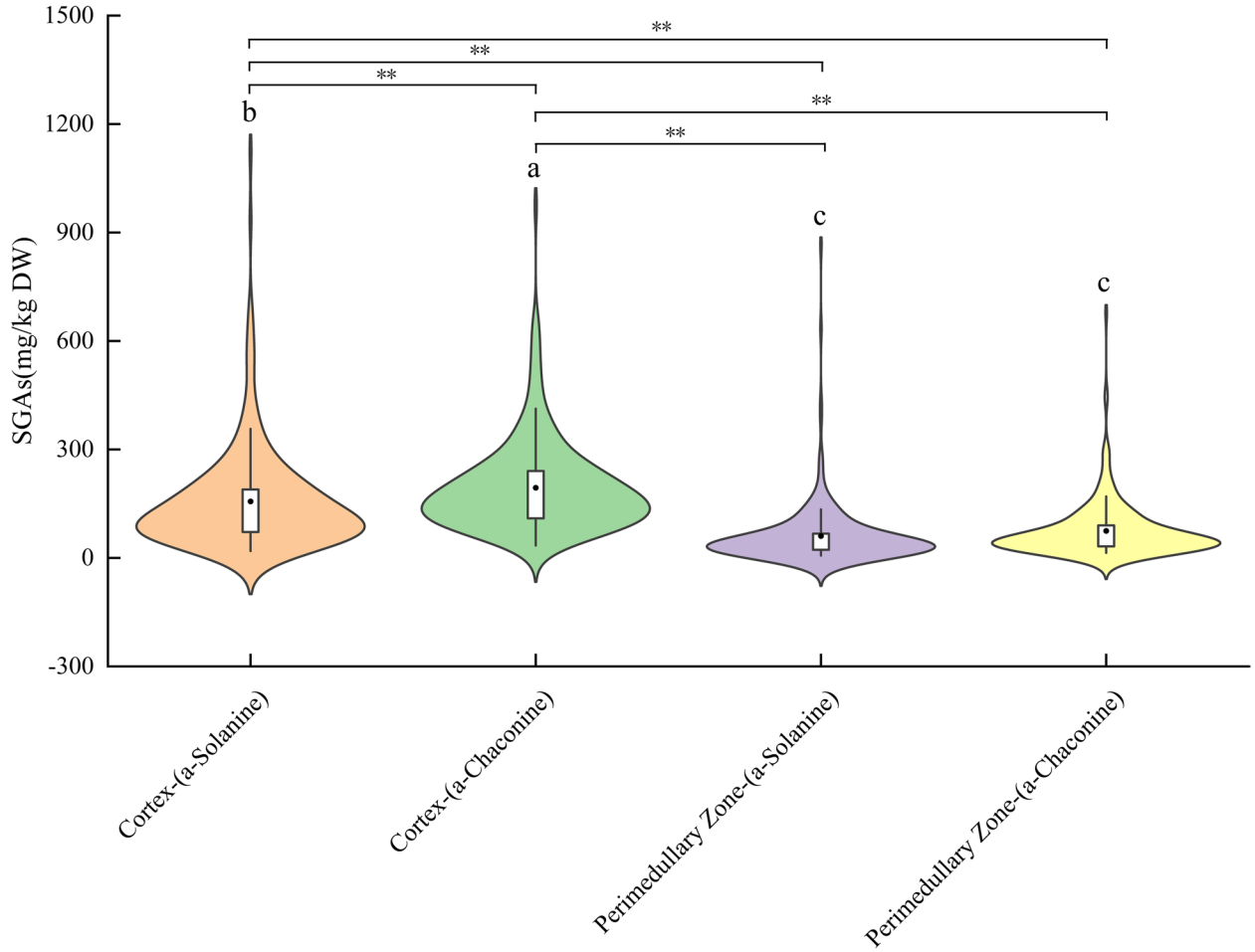

Fig. 1 Significance analysis of SGAs content in the cortex and perimedullary zone of potato tubers. Data are represented as average values with the SD of three independent biological replicates. Significant difference (**$P < 0.01$). The same below

differences ($P < 0.001$) (Table 2). While α-solanine and α-chaconine content in the perimedullary zone showed no significant difference, all other traits exhibited highly significant pairwise differences ($P < 0.01$) (Fig. 1). The genotype effects sum of squares for α-solanine and α-chaconine in the potato tuber cortex and perimedullary zone accounted for 84.81%, 78.57%, 90.89%, and 85.30% of the total variance. The environmental effect sum of squares accounted for 0.32%, 2.70%, 0.60%, and 2.44% of the total variation sum of squares, while the genotype ×environment interaction effect sum of squares accounted for 14.64%, 18.41%, 8.29%, and 11.92%. Furthermore, the broad-sense heritability for α-solanine and α-chaconine in the cortex and perimedullary zone tissues was calculated, with values of 82.74%, 90.88%, 76.50%, and 86.01% (Table 2), indicating that SGAs exhibit stable genetic traits, being primarily governed by genetic factors with minimal environmental influence.

Correlation analysis

To examine the relationships among traits, Pearson correlation analysis was conducted on all traits to identify significant correlations (Fig. 2). A phenotypic diversity and correlation analysis of α-solanine and α-chaconine content in 117 potato varieties (lines) from two locations and two parts of the tuber was conducted. The results showed that the α-solanine and α-chaconine content in the cortex of potato tubers at the SD and WY followed a normal distribution, while the α-solanine and α-chaconine content in the perimedullary zone of potato tubers at two locations followed a skewed normal distribution, with the peak shifted to the left. The α-solanine and α-chaconine content between the two parts of the potato tuber at the two locations showed extremely significant positive correlations ($P < 0.01$) for all pairs. The correlation coefficients for α-solanine and α-chaconine content in the same part and location of the potato tuber at the same site were 0.91, 0.91, 0.93,

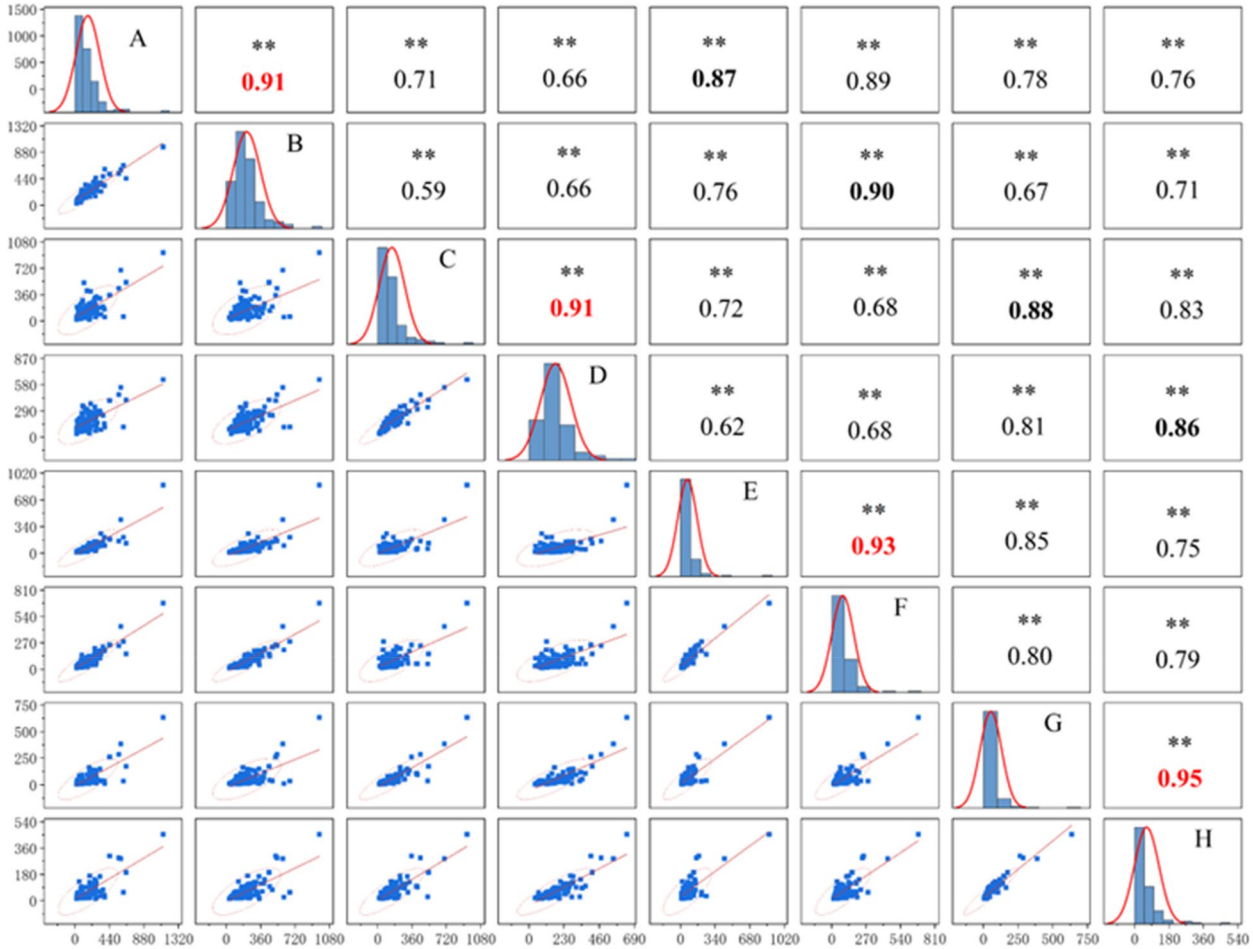

Fig. 2 Correlation analysis of SGAs content between the cortex and perimedullary zone potato tubers A: SD-Cortex-(α-Solanine); B: SD-Cortex-(α-Chaconine); C: WY-Cortex-(α-Solanine); D: WY-Cortex-(α-Chaconine); E: SD-Perimedullary Zone-(α-Solanine); F: SD-Perimedullary Zone-(α-Chaconine); G: WY-Perimedullary Zone-(α-Solanine); H: WY-Perimedullary Zone-(α-Chaconine)

and 0.95. For the same substances in different parts at the same locations, the correlation coefficients were 0.89, 0.90, 0.88, and 0.86.

GWAS of SGAs in potato tubers

Based on the phenotypic data of α-solanine and α-chaconine content and the genotype data from re-sequencing, a GWAS analysis of α-solanine and α-chaconine content was conducted for 117 tetraploid potato samples from two locations and two parts (Table S2). We used the mixed linear model (MLM) implemented in EMMAX, accounting for both population structure (Q) and kinship (K), to minimize false positives. Manhattan and QQ plots were generated. In the QQ plot, the SNP distribution closely followed the expected null distribution, indicating effective correction for population structure and relatedness, with a low false positive rate and high reliability of the SNP loci. In the Manhattan plot, significant SNP loci were identified when the association threshold -lg(P) (P-value) was greater than or equal to 6 (Fig. 3).

In the potato tuber cortex at the SD location, thirty-three SNP loci significantly associated with α-solanine biosynthesis were identified, distributed across nine chromosomes, except for chr07, chr08, and chr12. Among these, the chr05-27836136 showed the highest -lg(P) value of 8.13 (Fig. 3A). Similarly, twenty-seven SNP loci significantly associated with α-chaconine biosynthesis were identified, distributed across chr01, chr04, chr05, chr06, chr09, and chr11. The chr11 contained the most significant SNPs (9), with chr11-5,781,356 having the highest -lg(P) value of 7.53 (Fig. 3B). In the potato tuber cortex at the WY location, fifty-eight SNP loci significantly related to α-solanine biosynthesis were located on chr01, chr05, chr08, chr09, chr11, and chr12. The chr08-49707410 had the highest -lg(P) value of 8.07 (Fig. 3C). Twelve SNP loci significantly associated with α-chaconine biosynthesis were identified on

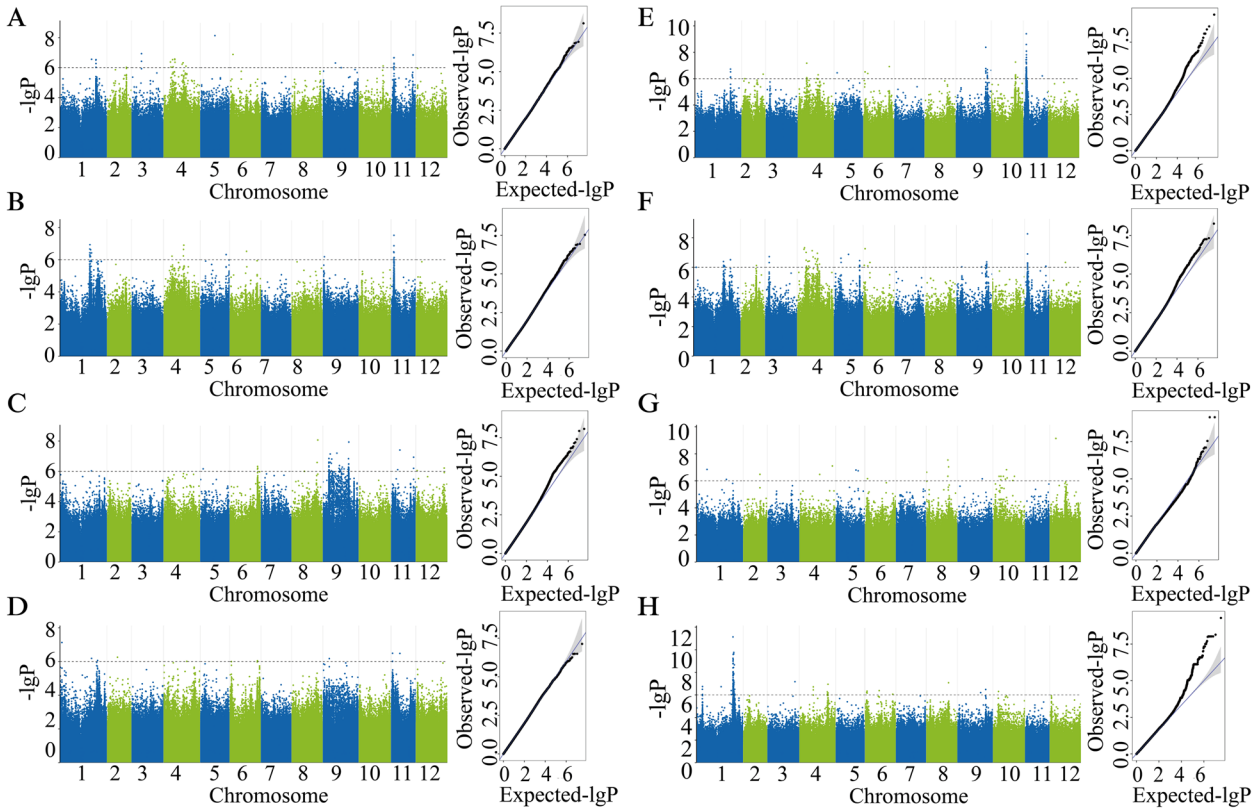

Fig. 3 Manhattan plots and QQ plots for the association analysis of SGAs content in the cortex and perimedullary zoneof potato tubers. **A**: SD-Cortex-(α-Solanine); **B**: SD-Cortex-(α-Chaconine); **C**: WY-Cortex-(α-Solanine); **D**: WY-Cortex-(α-Chaconine); **E**: SD-Perimedullary Zone-(α-Solanine); **F**: SD-Perimedullary Zone-(α-Chaconine); **G**: WY-Perimedullary Zone-(α-Solanine); **H**: WY-Perimedullary Zone-(α-Chaconine)

chr01, chr02, chr04, chr06, chr09, and chr11. The chr01-3678669 had the highest -lg(P) value of 7.13 (Fig. 3D). In the potato tuber perimedullary zone at the SD location, one hundred and forty-nine SNP loci significantly associated with α-solanine biosynthesis were identified across eight chromosomes, except for chr03, chr07, chr08, and chr11. The highest number of significant SNPs (105) was detected on chr11, with chr11-5,751,239 exhibiting the highest -lg(P) value of 9.40 (Fig. 3E). Sixty-eight SNP loci significantly associated with α-chaconine biosynthesis were identified on nine chromosomes, except for chr07, chr08, and chr10. The chr11-5,751,239 had the highest -lg(P) value of 8.26 (Fig. 3F). In the potato tuber perimedullary zone at the WY location, twenty-eight SNP loci significantly associated with α-solanine biosynthesis were identified across nine chromosomes, except for chr03, chr07, and chr11. The chr12-12,156,173 had the highest -lg(P) value of 9.14 (Fig. 3G). Four hundred and sixty-four SNP loci significantly associated with α-chaconine biosynthesis were identified across chr01, chr03, chr04, chr06, chr08, chr09, and chr10. The chr01 contained the most significant SNPs (444), with chr01-69377399 having the highest -lg(P) value of 11.14 (Fig. 3H).

In the potato tuber cortex, ninety-one significant SNP loci related to α-solanine biosynthesis and 39 significant SNP loci related to α-chaconine biosynthesis were detected; in the perimedullary zone, one hundred and seventeen significant SNP loci related to α-solanine biosynthesis and 532 significant SNP loci related to α-chaconine biosynthesis were identified. A greater number of significant SNP loci were identified on chr01 and chr11, and these SNPs exhibited clustered on the chromosomes. Several key structural genes involved in SGAs biosynthesis are enriched on these two chromosomes. Analysis of the stability loci revealed seven significant loci that control α-solanine biosynthesis (chr11-5,880,845, chr11-5,751,239, chr01-67823758, chr11-5,915,915, chr11-5,947,099, chr11-5,947,121 and chr11-5,773,430) and eight significant loci that control α-chaconine biosynthesis in both the potato tuber cortex and perimedullary zone (chr04-38347564, chr11-5,751,239, chr11-5,773,430, chr01-58246362, chr05-49012001, chr11-5,954,884, chr11-5,961,340 and chr01-71006745). In the potato tuber cortex, five significant loci jointly regulate α-solanine and α-chaconine biosynthesis(chr11-5,751,239, chr11-5,773,430, chr11-5,947,099, chr11-5,947,121 and chr11-6,880,345); in the perimedullary zone, eighteen significant loci jointly regulate both α-solanine and α-chaconine

biosynthesis(chr06-3686091, chr06-3686092, chr11-5,751,239, chr11-5,773,430, chr09-57277891, chr11-5,892,207, chr11-5,949,487, chr11-5,936,370, chr11-5,915,915, chr11-5,946,860, chr09-57280131, chr09-57278883, chr11-5,880,845, chr06-8815285, chr09-57279370, chr11-5,954,884, chr11-5,961,340 and chr09-57214860). These significant SNP loci that control both α-solanine and α-chaconine biosynthesis exhibit pleiotropy. Comparative analysis of stability and pleiotropy loci revealed that the loci chr11-5,751,239 and chr11-5,773,430 exhibit both stability and pleiotropy, providing reference for the identification of candidate genes involved in SGAs biosynthesis.

Candidate gene identification for SGAs in potato tubers

Candidate gene analysis for stability and pleiotropy

Using the gene IDs corresponding to significant SNPs identified through GWAS, functional annotation of candidate genes was performed via the Potato Genome Database (http://spuddb.uga.edu/). A total of 497 candidate genes potentially involved in SGAs biosynthesis were identified, and their biological functions were subsequently annotated. In the potato tuber cortex and perimedullary zone, two hundred and eighteen genes are associated with α-solanine biosynthesis, and 242 genes are associated with α-chaconine biosynthesis. Among them, fifteen genes simultaneously regulate α-solanine biosynthesis in both the potato tuber cortex and perimedullary zone, and 18 genes simultaneously regulate α-chaconine biosynthesis in both regions. Functional annotation of these 33 stability genes revealed that 6 genes are involved in the biosynthesis of both α-solanine and α-chaconine in the potato tuber cortex and perimedullary zone. These 6 genes exhibit both stability and pleiotropy and encode UDP-glycosyltransferase superfamily proteins (*Soltu.DM.11G005750*, *Soltu.DM.11G005760*, *Soltu.DM.11G005770*, *Soltu.DM.11G005820*), alkaline/neutral invertase (*Soltu.DM.11G006090*), and pleiotropic drug resistance (*Soltu.DM.11G006080*). In addition, eight candidate genes simultaneously regulate α-solanine and α-chaconine biosynthesis in the potato tuber cortex, while 21 candidate genes regulate α-solanine and α-chaconine biosynthesis in the potato tuber perimedullary zone. Functional annotation of these 29 pleiotropy genes revealed that 6 genes simultaneously regulate both α-solanine and α-chaconine biosynthesis in the potato tuber cortex and perimedullary zone. These 6 genes exhibit both pleiotropy and stability and encode UDP-glycosyltransferase superfamily proteins (*Soltu.DM.11G005750*, *Soltu.DM.11G005760*, *Soltu.DM.11G005770*, *Soltu.DM.11G005820*), alkaline/neutral invertase (*Soltu.DM.11G006090*), and pleiotropic drug resistance (*Soltu.DM.11G006080*).

Go and KEGG enrichment analysis of candidate genes

Perform Go and KEGG enrichment analysis on the genes identified by GWAS that are significantly associated with SGAs biosynthesis, to further identify the candidate genes involved in SGAs biosynthesis. The GO annotation revealed that these genes are mainly involved in biological processes (BP), cellular components (CC), and molecular functions (MF), with each category annotated with 11, 7, and 6 sub-functions. In GO analysis, genes were predominantly annotated under metabolic processes (16.14%), cellular processes (12.16%), and biological regulation (5.24%) within the BP category. In terms of CC, the genes were mainly annotated to membrane part, cell, and cell part, accounting for 2.73%, 2.52%, and 2.52% (Fig. 4A). In MF, the genes were mainly annotated to binding, catalytic activity, and enzyme regulator activity, accounting for 21.80%, 17.19%, and 1.05%. GO enrichment analysis revealed a total of 30 enriched GO terms, with significantly enriched terms including the lipid biosynthetic process (GO:0008610), lipid binding (GO:0008289),'de novo'pyrimidine nucleobase biosynthetic process (GO:0006207), regulation of transcription, DNA-templated (GO:0006355), UDP-glycosyltransferase activity (GO:0008194), nucleoside metabolic process (GO:0009116), base-excision repair (GO:0006284), 4 iron, 4 sulfur cluster binding (GO:0051539), lipid transport (GO:0006869), palmitoyltransferase activity (GO:0016409), DNA-binding transcription factor activity (GO:0003700), methyltransferase activity (GO:0008168), DNA repair (GO:0006281), ATPase activity (GO:0016887), protein kinase CK2 complex (GO:0005956), nucleotidase activity (GO:0008252), protein kinase regulator activity (GO:0019887), and translation initiation factor binding (GO:0031369) (Fig. 4B). Lipid biosynthetic process and UDP-glycosyltransferase activity are closely associated with steroid biosynthesis and glycosylation, highlighting their potential roles in SGAs metabolic pathways.

KEGG annotation analysis revealed that these genes were predominantly associated with metabolism (78.33%), genetic information processing (23.33%), environmental information processing (10.00%), organismal systems (3.33%), cellular processes (3.33%), and human diseases (1.67%) (Fig. 5A). KEGG enrichment analysis showed that candidate genes were significantly enriched in biosynthesis and metabolic pathways. Among these, the pathway for cutin, suberine, and wax biosynthesis (KO00073) showing the highest enrichment. Other pathways with lower enrichment significance included purine metabolism (KO00230), pyrimidine metabolism (KO00240), alanine, aspartate, and glutamate metabolism (KO00250), alpha-linolenic acid metabolism (KO00592),

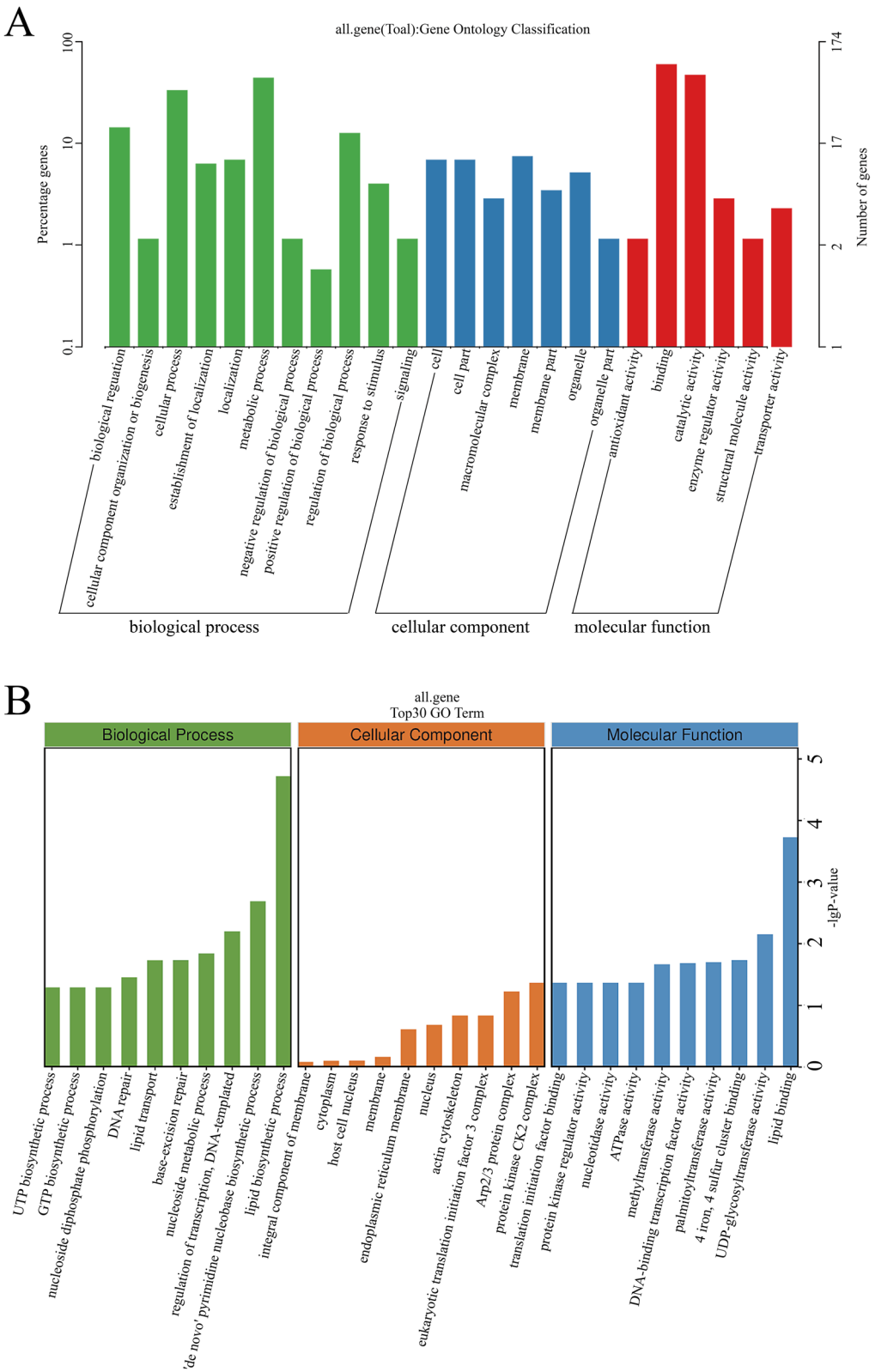

Fig. 4 GO enrichment analysis of candidate genes related to SGAs content in potato tubers

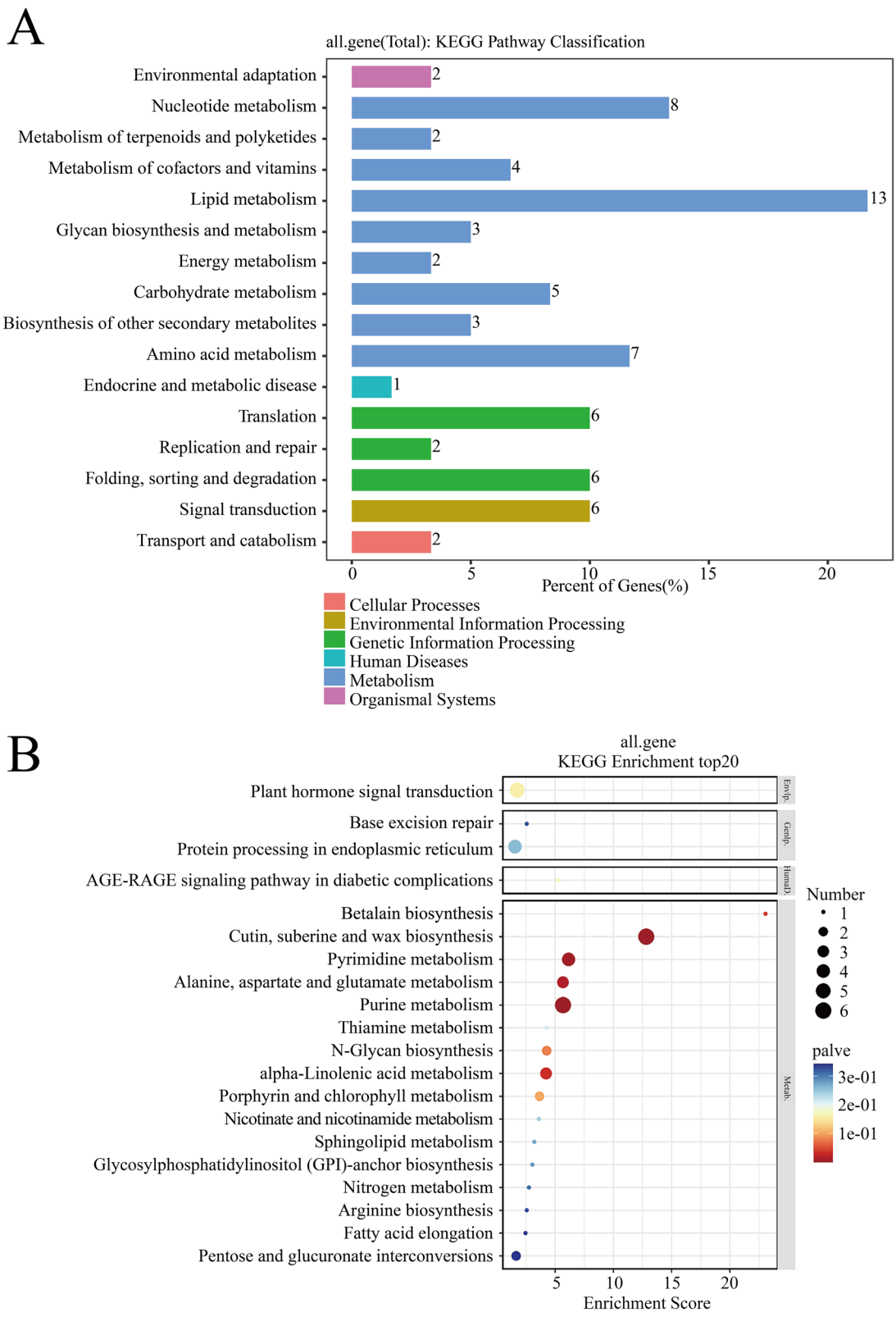

Fig. 5 KEGG enrichment analysis of candidate genes related to SGAs content in potato tubers

and betalain biosynthesis (KO00965) (Fig. 5B). Cutin, suberine, and wax biosynthesis is closely associated with steroid biosynthesis and glycosylation, highlighting their potential roles in SGAs metabolic pathways.

Significant loci haplotype analysis

By analyzing the Manhattan and QQ plots of the traits SD-Cortex-(α-Solanine) and WY-Cortex-(α-Chaconine), two highly associated loci were identified, both of which had SNPs distributed nearby. Based on the linkage disequilibrium between SNPs, the SNPs surrounding the loci showed similar signal intensities. LD block analysis of the chr01-69377399 locus (WY-Cortex-(α-Chaconine)) was performed by calculating the correlation of all SNPs within a 20 kb upstream and downstream candidate region. Six genes were identified, three of which (*Soltu.DM.01G029600*, *Soltu.DM.01G029620*, *Soltu.DM.01G029640*) are likely involved in SGAs biosynthesis, encoding fatty acid hydroxylase superfamily proteins (Fig. 6A C). Haplotype analysis revealed that 17 accessions carried 'CT", 98 carried "CC", and accessions with "CT" exhibited significantly higher tuber SGAs levels compared to those with "CC" (Fig. 6E).

LD block analysis of chr11-5,751,239 (SD-Cortex-(α-Solanine)) was conducted by correlating all SNPs within 20 kb upstream and downstream of the locus. Seven genes were identified, and functional annotation suggested that four genes (*Soltu.DM.11G005750*, *Soltu.DM.11G005760*, *Soltu.DM.11G005770*, and *Soltu.DM.11G005780*) might be involved in SGAs biosynthesis, encoding UDP-glycosyltransferase superfamily proteins (Fig. 6B D). Haplotype analysis revealed that 2 accessions carried "AA", 35 carried "AC", and 78 carried "CC". Accessions with "AA" and "AC" had significantly higher tuber SGAs levels than those with "CC", but no significant difference was observed between "AA" and "AC" (Fig. 6F).

Discussion

During potato plant growth, tuber SGAs content is affected by diverse growth conditions. Environmental factors such as high temperature, lower altitudes, drought, low phosphorus and potassium fertilizers, and high nitrogen fertilizer promote SGAs formation and accumulation to defend against stress [31]. The higher rainfall and average annual temperature at the SD, combined with its lower altitude compared to WY, likely account for the higher SGAs content at SD. The SGAs content in the cortex exceeds that in the perimedullary zone region because SGAs primarily localize within the 1.5 mm cortex layer of the tuber, with biosynthesis occurring in phelloderm cells. The higher α-chaconine content compared to α-solanine in most genotypes may result from competitive inhibition during their biosynthesis. The biosynthesis of α-solanine and α-chaconine within the same region shows a strong correlation, and the biosynthesis of α-solanine across different regions and α-chaconine across different regions also demonstrates a certain degree of strong correlation. α-Solanine and α-chaconine biosynthesis is influenced by G, E, and G× E, showing extensive genotypic variation. In this study, the broad-sense heritability of tuber SGAs content ranged from 76.50% to 90.88%, consistent with previous findings of 86%–89% [32], suggesting that SGAs biosynthesis in potato tubers is minimally affected by the environment and primarily influenced by genotype.

Potato SGAs biosynthesis is a complex trait controlled by multiple genes. 497 candidate genes were identified, and a comparison with previously reported genes involved in potato SGAs biosynthesis confirmed three consistent genes: *GAME2* (*Soltu.DM.07G014160*), *GAME6* (*Soltu.DM.07G014190*), and *GAME9* (*Soltu.DM.01G031000*). This confirms the feasibility of conducting GWAS on SGAs biosynthesis in potato tubers using autotetraploid potato cultivars. Six candidate genes were identified by analyzing SNP loci exhibiting both tissue stability and pleiotropy, along with their associated candidate genes. The six candidate genes encode UDP-glycosyltransferase superfamily proteins (*Soltu.DM.11G005750*, *Soltu.DM.11G005760*, *Soltu.DM.11G005770*, *Soltu.DM.11G005820*), alkaline/neutral invertase (*Soltu.DM.11G006090*), and pleiotropic drug resistance (*Soltu.DM.11G006080*).

GO enrichment analysis identified 30 GO terms, with lipid biosynthesis (*Soltu.DM.01G029600*, *Soltu.DM.01G029620*, *Soltu.DM.01G029640*, *Soltu.DM.01G029650*, *Soltu.DM.10G008360*) and UDP-glycosyltransferase activity (*Soltu.DM.11G005440*, *Soltu.DM.11G005710*, *Soltu.DM.11G005760*, *Soltu.DM.11G005770*, *Soltu.DM.11G005840*, *Soltu.DM.11G005980*, *Soltu.DM.11G006000*, *Soltu.DM.11G018080*) being significantly enriched and related to SGAs biosynthesis. KEGG enrichment analysis indicated that candidate genes are mainly involved in metabolic processes, with the most significant enrichment observed in Cutin, suberine and wax biosynthesis (*Soltu.DM.01G029600*, *Soltu.DM.01G029610*, *Soltu.DM.01G029620*, *Soltu.DM.01G029640*, *Soltu.DM.01G029650*, *Soltu.DM.10G008360*). Haplotype analysis identified seven genes, with *Soltu.DM.11G005750*、*Soltu.DM.11G005760*、*Soltu.DM.11G005770*、*Soltu.DM.11G005780* encode UDP-glycosyltransferase superfamily proteins, while *Soltu.DM.01G029600*、*Soltu.DM.01G029620*、*Soltu.DM.01G029640* primarily encodes fatty acid hydroxylase superfamily proteins.

Through the analysis of stability and pleiotropy SNP loci, genes, GO and KEGG enrichment analyses, and haplotype

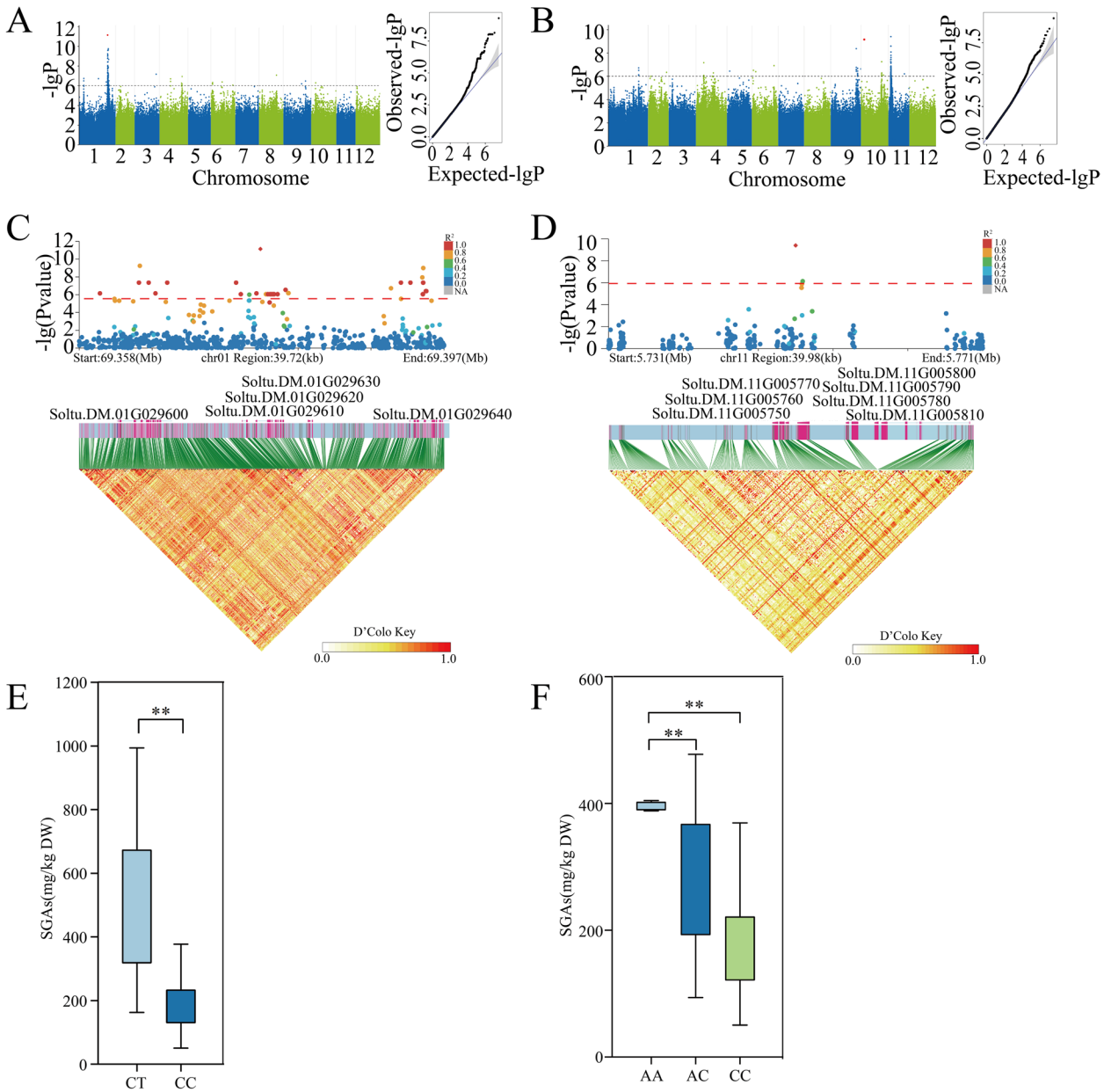

Fig. 6 Linkage disequilibrium heatmaps, and haplotype analysis of two candidate SNP regions. In the Manhattan plot, red dots indicate SNP loci selected for haplotype analysis. **A**, **C** and **B**, **D** display Manhattan plots and the linkage disequilibrium for chr01-69377399 and chr11-5,751,239, with dots above the threshold line representing significant SNP markers, the color intensity of the blocks correlates with the R2 values, indicating the degree of LD. **E** and **F** are boxplots showing the variance in phenotype between different haplotypes within the two SNP regions. The significance of differences between different haplotypes was indicated by "**" at 0.01 level by ANOVA

analysis of significant loci, a total of 12 candidate genes associated with the biosynthesis of potato tuber SGAs were identified. These genes were classified into four categories according to their functional annotations: UDP-glycosyltransferase superfamily proteins (*Soltu.DM.11G005750*, *Soltu.DM.11G005760*, *Soltu.DM.11G005770*, *Soltu.DM.11G005820*), fatty acid hydroxylase superfamily proteins (*Soltu.DM.01G029600*, *Soltu.DM.01G029610*, *Soltu.DM.01G029620*, *Soltu.DM.01G029640*, *Soltu.DM.01G029650*, *Soltu.DM.10G008360*), alkaline/neutral invertase (*Soltu.DM.11G006090*), and pleiotropic drug resistance (*Soltu.DM.11G006080*).

UDP-glycosyltransferases catalyze the glycosylation of plant secondary metabolites and play an important

role in plant defense responses [33]. The biosynthesis of potato SGAs from solanidine to α-solanine and α-chaconine involves the action of three UDP-glycosyltransferases, which sequentially attach glucose (SGT2) [34], rhamnose (SGT3) [35], and galactose (SGT1) [36] to solanidine. The potato SGAs biosynthetic pathway is still not completely elucidated and requires one rhamnosyltransferase and glucosyltransferase that remained to be identified to date [37]. In this study, we identified four novel genes encoding UDP-glycosyltransferase superfamily proteins that are potentially involved in SGAs biosynthesis. These genes are hypothesized to cooperate with the previously characterized *SGT1*, *SGT2*, and *SGT3* to modify the SGAs core scaffold by regulating the sequence and number of sugar moieties on the oligosaccharide chain, thereby modulating the structural diversity and toxicity of SGAs.

Fatty acid hydroxylase catalyzes the hydroxylation of long-chain fatty acids. Fatty acids use acetyl-CoA as the starting substrate [38], while SGAs also depend on acetyl-CoA as a key precursor in secondary metabolism [39]. The allocation of acetyl-CoA directly affects the synthesis rates and metabolic fluxes of both pathways. Under environmental stress, plants enhance the synthesis of both fatty acids and SGAs to mitigate adverse conditions. Oxidized lipids, such as jasmonic acid generated by fatty acid hydroxylase activity, serve as signaling molecules that activate disease-resistance-related secondary metabolic pathways [40]. GAME9, a jasmonic acid-responsive transcription factor (JRE4), binds to the promoters of SGAs biosynthetic genes (*GAME4*, *GAME6*, *GAME11*, *GAME12*, *GAME1*, *GAME2*, *SGT2*), thereby upregulating their expression and indirectly affecting SGAs accumulation [14]. The biosynthesis of SGAs is initiated from cholesterol and its derivatives, with hydroxylation being a critical chemical step in the metabolic pathway. Fatty acid hydroxylases are capable of introducing hydroxyl groups into fatty acid chains or steroidal backbones, potentially modifying steroidal precursors such as solanidine and cholestanol. These modifications may enhance the efficiency of subsequent glycosylation reactions catalyzed by glycosyltransferases. In addition to serving as substrates for glycosylation, hydroxylated intermediates may also function as signaling molecules, modulating the expression of other genes involved in SGAs biosynthesis and contributing to the broader regulatory network. Moreover, fatty acid hydroxylases may indirectly influence steroid metabolic flux by modulating lipid metabolism, membrane composition, and associated signaling pathways, thereby affecting the availability and direction of metabolic precursors for SGAs biosynthesis.

Alkaline/neutral invertase (A/N-INV) irreversibly converts sucrose into glucose and fructose [41–43], providing direct carbon skeletons and energy precursors for SGAs. Glucose undergoes glycolysis and the tricarboxylic acid (TCA) cycle to produce acetyl-CoA, a critical precursor for SGAs. Eight A/N-INVs have been identified in potatoes, with *StNINV* expression significantly upregulated under salt, heat, and drought stress [44]. In tomatoes, *SlINV* genes exhibit high expression under light, cold, and drought stress conditions [45]. Under abiotic stress, the transcription factor MYB induces the expression of both *A/N-Inv* [46] genes and SGAs [18]biosynthesis genes (*GAME4*, *GAME6*, *GAME7*, *GAME8*, and *GAME11*), facilitating resource allocation between primary and secondary metabolism. Additionally, increased A/N-INV activity may trigger secondary metabolism-related defense pathways, promoting rapid SGAs accumulation to strengthen plant resistance. Carbon metabolism serves as the fundamental basis for secondary metabolic processes. Alterations in the activity of A/N-INV may influence sugar signaling pathways or cellular energy status, thereby modulating the synthesis of precursors or the overall flux of the SGAs biosynthetic pathway.

Pleiotropic drug resistance (PDR) proteins, a subgroup of the ABC transporter family, utilizes energy from ATP hydrolysis to transport lipids, proteins, steroids, and cellular metabolites across membranes, contributing to secondary metabolite accumulation and stress responses [47, 48]. PDR proteins are hypothesized to transport SGAs from their synthesis sites to vacuoles or the apoplast via an ATP-dependent mechanism, preventing cytoplasmic toxicity and possibly facilitating long-distance transport or inter-tissue redistribution. In rice, overexpression of *OsPDR1* results in higher jasmonic acid levels, stronger growth inhibition, and enhanced disease resistance [49].

PDR proteins mediate the transport of jasmonic acid derivatives to specific target sites, potentially activating SGAs biosynthetic gene expression. PDR proteins may selectively recognize SGAs with specific glycosylation patterns, modulating their transport efficiency and destination, thereby influencing their stability and biological function in plants.

This study is the first to conduct an SGAs-related GWAS analysis based on tissue-specific data, highlighting the metabolic regulatory differences between the cortex and perimedullary zone, and offering a new perspective for understanding the accumulation mechanism of SGAs in different tuber tissues. In addition, the strategy of identifying stable loci and key genes through multidimensional integrative analysis is highly applicable, improving the efficiency and accuracy of candidate gene selection and laying a foundation for further functional validation and precision breeding. The results not only provide theoretical support for improving the genetic

framework of SGAs biosynthesis in potatoes, but also offer potential gene resources and molecular targets for breeding new potato varieties with low SGAs content.

Conclusions

GWAS were conducted on α-solanine and α-chaconine content in the cortex and perimedullary zone of 117 germplasm potato tuber and 22,983,689 high-quality SNP markers. Analysis of stability and pleiotropy for the significant SNPs and candidate genes identified six stable candidate genes. Additionally, based on GO and KEGG enrichment of candidate genes and haplotype analysis of significant loci, six candidate genes associated with potato tuber SGAs biosynthesis were identified. These 12 genes may participate in potato SGAs biosynthesis. We proposed a detailed experimental strategy to validate the tissue-specific functions of the 12 candidate genes in the cortex and perimedullary zone for further studies, including: characterizing the expression patterns of candidate genes in tuber cortex and perimedullary zone using qRT-PCR or RNA-Seq; generating knockout or knockdown mutants to evaluate the role of candidate genes on SGAs accumulation; evaluating associations between candidate genes expression and SGAs accumulation across different tissues or developmental stages.

Abbreviations

SGAs	Steroidal glycoalkaloids
DW	Dry weight
FW	Fresh weight
BW	Body weight
EFSA	European Food Safety Authority
FDA	U.S. Food and Drug Administration
HMGR	3-Hydroxy-3-methylglutaryl CoA reductase
SQS	Squalene synthase
SSR2	Sterol side chain reductase 2
GAME1/SGT1	Sterol alkaloid glycosyltransferase 1
SGT2	Sterol alkaloid glycosyltransferase 2
GAME2/SGT3	Glycosterol rhamnosyltransferase
GAME4/PGA4	A member of the 88D subfamily of cytochrome P450 proteins
GAME6	Cytochrome P450 monooxygenases 72A188
GAME7/PGA2	Cytochrome P450 monooxygenases 72A186
GAME8/PGA1	Cytochrome P450 monooxygenases 72A208
GAME11/16DOX	2-Oxoglutarate dependent dioxygenase
GAME12	Gamma amino butyric acid transaminase 2
GAME25/3βHSD1	3β-Hydroxysteroid dehydrogenase 1
S5αR	Steroid 5αa-reductase
GAME31/23DOX	2-Oxoglutarate-dependent dioxygenase
ERF	Ethylene response factors
GAME 9	Glycoalkaloid metabolism 9
HY5	Elongated hypocotyl 5
COI1	Coronatine insensitive 1
TCP14	Teosinte branched1/Cycloidea/Proliferating cell factor 14
GWAS	Genome-wide association study
MLM	Mixed linear model
BP	Biological process
CC	Cellular component
MF	Molecular function
CIP	International potato center
A/N-INV	Alkaline/neutral invertase
PDR	Pleiotropic drug resistance
SNP	Single nucleotide polymorphism

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12870-025-06766-6.

Additional file 1: Table Sl Phenotypic data of SGAs in the cortex and perimedullary zone of potato tubers from two locations.

Additional file 2: Table S2 SNP loci significantly associated with SGAs biosynthesis in the cortex and perimedullary zone of potato tubers from two locations.

Acknowledgements

We thank State Key Laboratory of Aridland Crop Science for providing experimental materials.

Authors' contributions

F.Z. designed the study. X.Z., L.X. and Z.T. performed the experiments. X.Z., J.Y. and L.C., analyzed the data. X.Z. wrote the manuscript. All authors read and approved the final manuscript.

Funding

This work was supported by State Key Laboratory of Aridland Crop Science (GSCS-2023-Z05), National Natural Science Foundation of China (32360091, 32060046), Gansu Provincial University Industrial Support Project (2023CYZC-44), Gansu Provincial Science Technology Major Project (21ZD11NA002), and Gansu Provincial University Science Research Innovation Platform Major Cultivation Project (2024CXPT-01).

Data availability

The raw data presented in this study are available on request from the corresponding author. The data are not yet publicly available since the project is still ongoing.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 12 March 2025 Accepted: 22 May 2025
Published online: 29 May 2025

References

1. Bagri DS, Upadhyay DC, Jain SK, Upadhyay CP. Biotechnological improvement of nutritional and therapeutic value of cultivated potato. Front Biosci (Schol Ed). 2018;10:217–28.
2. Schrenk D, Bignami M, Bodin L, Chipman JK, Del Mazo J, Hogstrand C, et al. Risk assessment of glycoalkaloids in feed and food, in particular in potatoes and potato-derived products. EFSA J. 2020;18:06222.
3. Aziz A, Randhawa MA, Butt MS, Asghar A, Yasin M, Shibamoto T. Glycoalkaloids (α-chaconine and α-solanine) contents of selected Pakistani potato cultivars and their dietary intake assessment. J Food Sci. 2012;77(3):58–61.
4. Roddick JG. Steroidal glycoalkaloids: nature and consequences of bioactivity. Adv Exp Med Biol. 1996;404:277–95.
5. Shakya R, Navarre DA. LC-MS analysis of solanidane glycoalkaloid diversity among tubers of four wild potato species and three cultivars (*Solanum tumberosum*). J Agric Food Chem. 2008;56:6949–58.
6. Cárdenas PD, Sonawane PD, Heinig U, Bocobza SE, Burdman S, Aharoni A. The bitter side of the nightshades: Genomics drives discovery in Solanaceae steroidal alkaloid metabolism. Phytochemistry. 2015;113:24–32.

7. Nahar N, Westerberg E, Arif U, Huchelmann A, OlarteGuasca A, Beste L, et al. Transcript profiling of two potato cultivars during glycoalkaloid-inducing treatments shows differential expression of genes in sterol and glycoalkaloid metabolism. Sci Rep. 2017;7:43268.
8. Ginzberg I, Tokuhisa JG, Veilleux RE. Potato steroidal glycoalkaloids: Biosynthesis and genetic manipulation. Potato Res. 2009;52:1–15.
9. Liu Y, Liu X, Li Y, Pei Y, Jaleel A, Ren M. Potato steroidal glycoalkaloids: properties, biosynthesis, regulation and genetic manipulation. Molecular Horticulture. 2024;4:43.
10. Okamoto H, Ducreux LJM, Allwood JW, Hedley PE, Wright A, Gururajan V, et al. Light regulation of chlorophyll and glycoalkaloid biosynthesis during tuber greening of potato S. tuberosum. Frontiers in Plant Science. 2020;11:753.
11. Ginzberg I, Thippeswamy M, Fogelman E, Demirel U, Mweetwa AM, Tokuhisa J, et al. Induction of potato steroidal glycoalkaloid biosynthetic pathway by overexpression of cDNA encoding primary metabolism HMG-CoA reductase and squalene synthase. Planta. 2012;235:1341–53.
12. Akiyama R, Umemoto N, Mizutani M. Recent advances in steroidal glycoalkaloid biosynthesis in the genus Solanum. Plant Biotechnol (Tokyo). 2023;40:185–91.
13. Itkin M, Heinig U, Tzfadia O, Bhide AJ, Shinde B, Cardenas PD, et al. Biosynthesis of anti-nutritional alkaloids in solanaceous crops is mediated by clustered genes. Science. 2013;341:175–9.
14. Cárdenas PD, Sonawane PD, Pollier J, VandenBossche R, Dewangan V, Weithorn E, et al. GAME9 regulates the biosynthesis of steroidal alkaloids and upstream isoprenoids in the plant mevalonate pathway. Nat Commun. 2016;7:10654.
15. Abdelkareem A, Thagun C, Nakayasu M, Mizutani M, Hashimoto T, Shoji T. Jasmonate-induced biosynthesis of steroidal glycoalkaloids depends on COI1 proteins in tomato. Biochem Biophys Res Commun. 2017;489:206–10.
16. Shen DD, Hua YP, Huang JY, Yu ST, Wu TB, Zhang Y, et al. Multiomic analysis reveals core regulatory mechanisms underlying steroidal glycoalkaloid metabolism in potato tubers. J Agric Food Chem. 2022;70:415–26.
17. Wang C, Bai RQ, Guan JM, Luo JI, He XJ, Chi SY, et al. Promotion of StHY5 in the Synthesis of SGAs during Tuber Turning-green of Potato. Biotechnol Bull. 2024;40:113–22.
18. Liu SX, Cheng YX, Zhao XJ, Wang ES, Liu TT, Zhang HL, et al. The transcription factor StMYB113 regulates light-induced greening by modulating steroidal glycoalkaloid biosynthesis in potatoes (Solanum tuberosum L.). Horticulture Advances. 2024;2:7.
19. Sonawane PD, Heinig U, Panda S, Gilboa NS, Yona M, Kumar SP, et al. Short-chain dehydrogenase/reductase governs steroidal specialized metabolites structural diversity and toxicity in the genus *Solanum*. Proc Natl Acad Sci USA. 2018;115:5419–28.
20. Sonawane PD, Jozwiak A, Panda S, Aharoni A. "Hijacking" core metabolism: a new panache for the evolution of steroidal glycoalkaloids structural diversity. Curr Opin Plant Biol. 2020;55:118–28.
21. Akiyama R, Lee HJ, Nakayasu M, Osakabe K, Osakabe Y, Umemoto N, et al. Characterization of steroid 5α-reductase involved in α-tomatine biosynthesis in tomatoes. Plant Biotechnol. 2019;36:253–63.
22. Cárdenas PD, Sonawane PD, Heinig U, Jozwiak A, Panda S, Abebie B, et al. Pathways to defense metabolites and evading fruit bitterness in genus Solanum evolved through 2-oxoglutarate-dependent dioxygenases. Nat Commun. 2019;10:5169.
23. Nakayasu M, Akiyama R, Kobayashi M, Lee HJ, Kawasaki T, Watanabe B, et al. Identification of α-tomatine 23-hydroxylase involved in the detoxification of a bitter glycoalkaloid. Plant Cell Physiol. 2020;61:21–8.
24. Boccia M, Kessler D, Seibt W, Grabe V, Rodríguez López C E, Grzech D, et al. A scaffold protein manages the biosynthesis of steroidal defense metabolites in plants. Science. 2024;386(6728):1366–72.
25. Yuan JL, Cheng LX, Wang YP, Zhang F. Genome-wide association studies for key agronomic and quality traits in potato (Solanum tuberosum L.). Agronomy. 2024;14:2214.
26. Peng Z. Domestication and inheritance of tuber flesh steroidal glycoalkaloids in diploid potato. Doctoral dissertation. Yunnan Normal University. 2019;21–3.
27. Vos PG, Paulo MJ, Bourke PM, Maliepaard CA, van Eeuwijk FA, Visser RGF, et al. GWAS in tetraploid potato: identification and validation of SNP markers associated with glycoalkaloid content. Mol Breeding. 2022;42:76.
28. Zhang F, Qu L, Gu Y, Xu ZH, Xue HW. Resequencing and genome-wide association studies of autotetraploid potato. Molecular Horticulture. 2022;10:2(1):6.
29. Zhou X, Stephens M. Genome-wide efficient mixed-model analysis for association studies. Nat Genet. 2012;44:821–4.
30. Xia LL. Genome-wide association study and candidate genes mining of phenolic acids in potato. Maser's thesis. Gansu Agricultural University. 2023;30–4.
31. Zhou XC, Cheng LX, Yuan JL, Li WL, Lv T, Yang C. Research progress in the biosynthetic of steroidal glycoalkaloids in Solanum plants. J Plant Genet Resour. 2024;1–22. https://doi.org/10.13430/j.cnki.jpgr.20241025003.
32. Benkeblia N. Potato glycoalkaloids: occurrence, biological activities and extraction for biovalorisation-a review. Int J Food Sci Technol. 2020;55:2305–13.
33. Tiwari P, Sangwan RS, Sangwan NS. Plant secondary metabolism linked glycosyltransferases: An update on expanding knowledge and scopes. Biotechnol Adv. 2016;34:714–39.
34. McCue KF, Allen PV, Shepherd LV, Blake A, Whitworth J, Maccree MM, et al. The primary in vivo steroidal alkaloid glucosyltransferase from potato. Phytochemistry. 2006;67:1590–7.
35. McCue KF, Allen PV, Shepherd LV, Blake A, Maccree MM, Rockhold DR, et al. Potato glycosterol rhamnosyltransferase, the terminal step in triose side-chain biosynthesis. Phytochemistry. 2007;68:327–34.
36. McCue KF, Shepherd LV, Allen PV, Maccree MM, Rockhold DR, Corsini D, et al. Metabolic compensation of steroidal glycoalkaloid biosynthesis in transgenic potato tubers: using reverse genetics to confirm the in vivo enzyme function of a steroidal alkaloid galactosyltransferase $. Plant Sci. 2005;168:267–73.
37. Lucier R, Kamileen MO, Nakamura Y, Serediuk S, Barbole R, Wurlitzer J, et al. Steroidal scaffold decorations in *Solanum* alkaloid biosynthesis. Mol Plant. 2024;17:1236–54.
38. De Bigault Du Granrut A, Cacas JL. How very-long-chain fatty acids could signal stressful conditions in plants? Frontiers in Plant Science. 2016;7:1490.
39. Sawai S, Ohyama K, Yasumoto S, Seki H, Sakuma T, Yamamoto T, et al. Sterol side chain reductase 2 is a key enzyme in the biosynthesis of cholesterol, the common precursor of toxic steroidal glycoalkaloids in potato. Plant Cell. 2014;26:3763–74.
40. Kazan K, Manners JM. Jasmonate signaling: toward an integrated view. Plant Physiol. 2008;146:1459–68.
41. Pan L, Guo Q, Chai S, Cheng Y, Ruan M, Ye Q, et al. Evolutionary conservation and expression patterns of neutral/alkaline invertases in *Solanum*. Biomolecules. 2019;9:763.
42. Yao S, Kodama R, Wang H, Ichii M, Taketa S, Yoshida H. Analysis of the rice SHORT-ROOT5 gene revealed functional diversification of plant neutral/alkaline invertase family. Plant Sci. 2009;176:627–34.
43. Dahro B, Wang F, Peng T, Liu JH. PtrA/NINV, an alkaline/neutral invertase gene of Poncirus trifoliata, confers enhanced tolerance to multiple abiotic stresses by modulating ROS levels and maintaining photosynthetic efficiency. BMC Plant Biol. 2016;16:76.
44. Abbas A, Shah AN, Shah AA, Nadeem MA, Alsaleh A, Javed T, et al. Genome-wide analysis of invertase gene family, and expression profiling under abiotic stress conditions in potato. Biology (Basel). 2022;11:539.
45. Ahiakpa JK, Magdy M, Karikari B, Munir S, Mumtaz MA, Tamim SA, et al. Genome-wide identification and expression profiling of tomato invertase genes indicate their response to stress and phytohormones. J Plant Growth Regul. 2022;41:1481–98.
46. Jin X, Ackah M, Acheampong A, Zhang Q, Wang L, Lin Q, et al. Genome-wide identification of candidate genes associated with heat stress in mulberry (Morus alba L.). Current Issues in Molecular Biology. 2023;45:4151–67.
47. Kang J, Park J, Choi H, Burla B, Kretzschmar T, Lee Y, et al. Plant ABC Transporters. Biochem Biophys Acta. 2011;9:0153.
48. Goossens A, Häkkinen ST, Laakso I, Oksman-Caldentey KM, Inzé D. Secretion of secondary metabolites by ATP-binding cassette transporters in plant cell suspension cultures. Plant Physiol. 2003;131:1161–4.
49. Zhang H, Jing W, Zheng J, Jin Y, Wu D, Cao C, et al. The ATP-binding cassette transporter OsPDR1 regulates plant growth and pathogen resistance by affecting jasmonates biosynthesis in rice. Plant Sci. 2020;298:110582.

Publisher's Note