TÜBİTAK

Turkish Journal of Agriculture and Forestry

http://journals.tubitak.gov.tr/agriculture/

Research Article

Turk J Agric For
(2022) 46: 90-103

doi:10.3906/tar-2104-37

Genome-wide association studies (GWAS) revealed a genetic basis associated with floral traits in potato germplasm

Muhammad Abu Bakar ZIA[1,2*], Ufuk DEMİREL[1], Muhammad Azhar NADEEM[3], Fawad ALI[4], Ahmad DAWOOD[5], Muhammad IJAZ[2], Mehmet Emin ÇALIŞKAN[1]

[1] Department of Agricultural Genetic Engineering, Faculty of Agricultural Sciences and Technologies, Niğde Ömer Halisdemir University, Niğde, Turkey

[2]College of Agriculture, Bahauddin Zakariya University, Bahadur Campus Layyah, Pakistan

[3] Faculty of Agricultural Sciences and Technologies, Sivas University of Science and Technology, Sivas, Turkey

[4]Department of Plant Sciences, Quaid-I-Azam University, Islamabad, Pakistan

[5]Department of Agronomy, Bahauddin Zakariya University, Multan, Pakistan

Received: 14.04.2021 • **Accepted/Published Online:** 04.01.2022 • **Final Version:** 09.02.2022

Abstract: Potato is an important noncereal staple crop serving as a source of food for a large number of the world's population. Genome-wide association study (GWAS) analysis has become a useful tool to uncover the genetic basis of important plant traits by revealing significant association with the trait of interest. The present study aimed to explore the phenotypic diversity and to identify the genetic basis associated with important floral traits. A total of 237 tetraploid potato genotypes were used as plant material, and field experiments were conducted according to augmented block design for two consecutive years (2016, 2017). Analysis of variance for the studied floral traits reflected highly significant genotypic effects. Mean data for both years revealed the significant variation for pistil length (5.53 to 9.92 mm), stamen length (6.04 to 9.26 mm), and pistil length above stamen (1.31 to 4.47 mm). Pearson's correlation analysis reflected a highly significant and positive correlation of pistil length with stamen length (r = 0.42) and pistil length above stamen (r = 0.28). Principal component analysis was performed, and the first two PCs were considered accounting for a total of 81.2% variation. Constellation plot divided the studied potato panel into two main populations on the basis of stamen and pistil length. A total of 12,720SNP markers were used for the marker-trait association, and a total of 15 markers were found significantly associated with studied traits across both years. Identifying the same markers across both years helped in the validation of the obtained marker-trait associations. The identified significant markers reflected some of the putative candidate genes that might be beneficial in potato breeding programs. To the best of our knowledge, this is the first study identifying the genetic basis of important floral traits and might be helpful to the scientific community interested in potato marker-assisted breeding in these traits.

Key words: Solanum tuberosum, GWAS, pistil length, stamen length, solCAP

1. Introduction

The world is confronted with increasing population, food insecurity, and climate change. To feed the world's population growing at an unprecedented rate is becoming an immense challenge. The population of the world is predicted to catch the figure of 10 billion by the year 2050 and there is a need to upsurge the world food production from 60 to 110% (Godfray et al., 2010; Tilman et al., 2011). Fast-growing world population has put immense pressure on agricultural crop production system (Kastner et al., 2012; Dempewolf et al., 2014; Khoury et al., 2014). Ultimately, these threats have paved the attention of the world scientific community to implement advanced breeding tools that have a significant impact on the improvement of world food security. Keeping in view, there is a need to harness genetic diversity through collection and characterization of germplasm for the investigation of novel variations for breeding perspectives (Karık et al., 2019; Barut et al., 2020; Yildiz et al., 2020; Nadeem 2021; Shimira et al., 2021;).

Potato (*Solanum tuberosum* L.) is among the vital members of *Solanaceae* and ranked 4th among food crops after wheat, rice, and corn, respectively (Carputo et al., 2004). It is known to possess a potential of offering increased and sufficient food for the world's population (FAOSTAT, 2017). Potato is an auto-tetraploid plant that reproduces by asexual means (Knapp, 2008; Bali et al., 2018). It exhibits vast distinctions in its ploidy level that

* Correspondence: abziathebest@gmail.com

varies from diploid to hexaploid (Spooner et al., 2005). Generally, cultivated potato is tetraploid in nature (2n = 4x = 48) having a haploid genome size of around 844 Mbp (Potato Genome Sequencing Consortium 2011). Cultivated potatoes originated from the wild species that were extensively cultivated in Andean region (Peru, Northern Argentina, and Bolivia) (Jansky and Spooner, 2018). The domestication of cultivated potatoes occurred 8000 to 10,000 years ago from diploid wild species (2n = 2x = 24), and, over time, it became the most important pillar of food security around the world (Pearsall, 2008). Potato was cultivated on an area of above 173 million hectares with the production of around 370 million tonnes around the world in 2019. In Turkey, during 2019, potato was cultivated on an area of 1.4 million hectares with the production of 4.9 million tonnes (FAOSTAT, 2021)

GWAS and QTL mapping are vital and authentic methodologies for the detection of genome and genetic analysis. These methodologies are used to detect the genetic makeup of the traits using genotypic and phenotypic data of the population under study (Nadeem et al., 2018). GWAS is highly preferred over QTL mapping due to the presence of high recombination rate in the studied population, less time-consuming, and wide range efficacy of the markers identified in genome-wide association mapping (Korte and Farlow, 2013; Nadeem et al., 2018). This approach has been successfully applied for the investigation of marker-trait association in many crops (Raggi et al., 2019; Raman et al., 2019; Ali et al., 2020; Nadeem et al., 2021; Paudel et al., 2021).

In any crop plant, floral organs are observed and measured as combined ones. Flowers with various organs mostly act in a combined fashion to attract pollinators, donations, and receipt of pollens (Edwards and Weinig, 2011). It is very clear that flowers are mainly responsible for reproduction, the size, and shape of the floral organs are regularly coordinated, and the genotypic correlation among floral traits are strongly maintained within plant populations (Conner and Via, 1993; Oneil and Schmitt, 1993; Carr and Fenster, 1994; Conner and Sterling, 1995; Juenger et al., 2000; van Kleunen and Ritland, 2004; Ashman and Majetic, 2006; Brock and Weinig, 2007). Natural selection in floral phenotypes by pollinators is known to be one of the main reasons for the evolution of floral traits correlation (Berg, 1960). It is also known that covariation of the floral organs may arise from a common developmental or genetic architecture, i.e., the association among floral organs may occur due to pleiotropic genes that affect variation in numerous traits (Lande, 1980). Flowering traits play an important role as they define the genotype of the off-springs (Arteaga et al., 2015). Among these traits, pistil length, and stamen length are very important as these traits participate in producing the desired outcomes (Scot et al., 2004). Stamen produces the pollen grains, which are necessary for the reproduction process (Scot et al., 2004). Pistil length is another vital component of floral traits that determine the selfing and crossing nature of the plants (Fernández et al., 2009). To the best of our knowledge, this is the pioneering study exploring the genetic basis of important floral traits, including pistil length (PL), stamen length (SL), and pistil length above stamen (PLAS) of potato via genome-wide association study.

Breeding activities are highly boosted with the advent of molecular markers and sequencing technologies. It is highly proposed to implement codominant molecular markers as compared to dominant markers, as they are more robust and highly informative in nature (Nadeem et al., 2018). Single nucleotide polymorphism (SNPs) are codominant markers and known as the markers of choice and have a great deal of utilization in several marker-trait association studies in potato (Stich et al., 2013; Mosquera et al., 2016; Vos et al., 2017; Zia et al., 2020; Yousaf et al., 2021). Potato SolCAP Array has been developed with its enormous significance in genotyping assays (Felcher et al., 2012). With the development of high-density genotyping SolCAP array techniques, genotyping of potato germplasm up to 64K SNPs can be performed at a considerably low price. SolCAP potato array technology is being predominantly used in potatoes nowadays, and many of scientists have published their work using this technique. The adaptation of diploid linkage maps with potato genome sequencing and their sketches were described by Felcher et al. (2012). Firstly, it was used in European potato germplasm in which the population structure (Q) and LD of eight diploid potato clones and 36 different tetraploid cultivars were studied (Stich et al., 2013). Berdugo-Cely et al. (2017) evaluated 809 different Colombian species for different plant traits in potatoes. More recently, Zia et al. (2020) evaluated 237 diverse potato genotypes for important agronomic and morphological traits in potatoes. Yousaf et al. (2021) conducted GWAS analysis to reveal novel genomic regions controlling root and stolon traits in potatoes. Among 192 genotypes, around 50 different genomic regions were detected in root and stolon traits (Yousaf et al., 2021). Wilson et al. (2021) and Parra-Galindo et al. (2021) have recently conducted GWAS analysis in their studies and identified some genomic regions significantly associated with the studied traits. As is obvious from the above discussion, understanding the floral traits is very important for potato improvement. Keeping in view, the current investigation aimed to identify the genetic basis associated with floral traits like PL, SL, and PLAS in a collection of 237 tetraploid potato genotypes by implementing genome-wide association analysis using SolCAP SNP 12k array information.

2. Materials and methods

2.1. Plant materials

A total of 237 tetraploid potato genotypes from diverse sources were utilized during current exploration (detailed information about studied germplasm can be obtained from Zia et al. 2020). The sources of the tested potato genotypes are as follows: International Potato Center (CIP) Peru (194 genotypes), Hungary (10 genotypes), Japan (2 genotypes), progressive breeding lines from Niğde Ömer Halisdemir University Turkey's potato breeding program (26 genotypes) as well as check cultivars (5 genotypes) being used in Turkey, respectively, i.e., Desiree, Agria, Melody, Granola, and Russet Burbank.

2.2. Field experiment and phenotypic data evaluation

The two field experiments were conducted at Experimental Farm of Potato Research Institute Niğde (37° 54′ 57″ N, 34° 41′ 37″ E), Turkey between 2016 and 2017 to derive phenotypic data. Genotypes were planted in Augmented Block Design (Petersen, 1985) keeping row to row and plant to plant distance at 75 and 30 cm, respectively. Augmented design is preferred to be utilized to evaluate a higher number of crop germplasm/genotypes in limited resources like land, labor cost, etc. A total of 20 tubers of each genotype were planted in a row length of 6 m. Normal potato production practices were followed during the growing period in both years. The climatic conditions including maximum, minimum, and average temperature are presented in Figure 1. Data was recorded for important floral traits like PL, SL, and PLAS, respectively. A total of 10 plants were selected to observe the floral traits. Pistil length, stamen length, and pistil length above stamen in flowers of the selected plants were recorded with the help of a digital compass. The observed data on floral traits were averaged to obtain a single mean for the purpose of analysis.

2.3. Genomic DNA extraction and genotyping

DNA extraction was performed utilizing Gene JET Plant Genomic DNA Purification mini kit by Thermo scientific, conferring the supplier's instructions. Agarose gel (1%) was used to determine the quality and quantity of the extracted DNA. Furthermore, BioSpec-nano (Shimadzu) spectrophotometer was used for the confirmation of purification in extracted DNA samples. SNP markers (Infinium 12K potato SNP array) were used for genotyping purpose. Illumina HiScan SQ system was used to read the array. The genotypes were assigned to each locus using the software Genome Studio; diploid markers obtained from 12K SolCAP Potato genotyping array were used.

A total of 12,720 markers were obtained from SolCAP 12K array. All the monomorphic markers and the markers that could not be called were discarded. SNPs with 20% missing data were also discarded due to lack of required information. Minor allele frequency (MAF) threshold of 5% was used to the selected SNP's before performing the maker-trait association analysis. Overall, the filtration process discarded 4527 markers (35.58%), while the remaining markers (8,193) were used to perform various genetic analyses.

2.4. Statistical analysis

All data were subjected to analysis of variance (ANOVA) of each individual location using SAS software version 9.3 (SAS Institute, 2010 The SAS System for Windows. Cary (NC): SAS Inst.) using a general linear model (GLM) (Gomez and Gomez, 1984). Mean, maximum, minimum, Pearson's correlation coefficient test among the studied floral traits, and principal component analysis were performed using the statistical software XLSTAT (www.xlstat.com). The cluster constellation plot for 237 tetraploid potato genotypes was constructed by JMP 14.1.0 statistical software (2018, SAS Institute Inc., Cary, NC, USA).

2.5. Population structure and association mapping

Population structure was undertaken as described by Zia et al. (2020). The level of significance among the identified markers was obtained at $p \leq 0.05$ utilizing Genstat v18 (VSN International) software. Best linear unbiased estimator (BLUE) was used for the calculation of correlation coefficients. Genstat 18.1 was used to calculate the BLUE values using mean phenotypic data of the field experiments conducted during 2016 and 2017. Achenbach et al. (2009) described that BLUE is the coefficient assumed by the ordinary least squares estimator. Based on allele frequencies, for all SNPs with D', $r^{2,}$ or chi-square, LD was estimated (Achenbach et al., 2009; Hill and Robertson, 1968). LD values of polymorphic sites were plotted on *X* and *Y*-axis. Two sets of marker sites were assessed by each cell with a specific color, which codes the manifestation of significant LD. The colored bar code express the significance threshold level. To indicate the genetic distance, the pairwise LD values (r^2) were designed, and the nonlinear logarithmic regression curve of r^2 was expressed by internal trend line. Linked markers were used for the calculation of pairwise correlation of LD decay with significant LD ($p < 0.001$). The threshold for LD decay was considered below $r^2 = 0.2$. Mixed linear model (MLM) under TASSEL software (Bradbury et al., 2007) was used to estimate the marker-trait association (Zhang et al., 2010). Quantile-Quantile plot (qq plot) was drawn to logically check the correction of phenotypic data. TASSEL software was used for studying association mapping of the given data (Bradbury et al., 2007). P = 1x10^-4 i.e. –log10 (0.0001) was used for a standard value of Bonferroni correction by following the previously exercised method used by different researchers (Kullo et al., 2014; Yuan et al., 2019).

For genome-wide association studies (GWASs), a regular mixed linear model (MLM) was implemented in

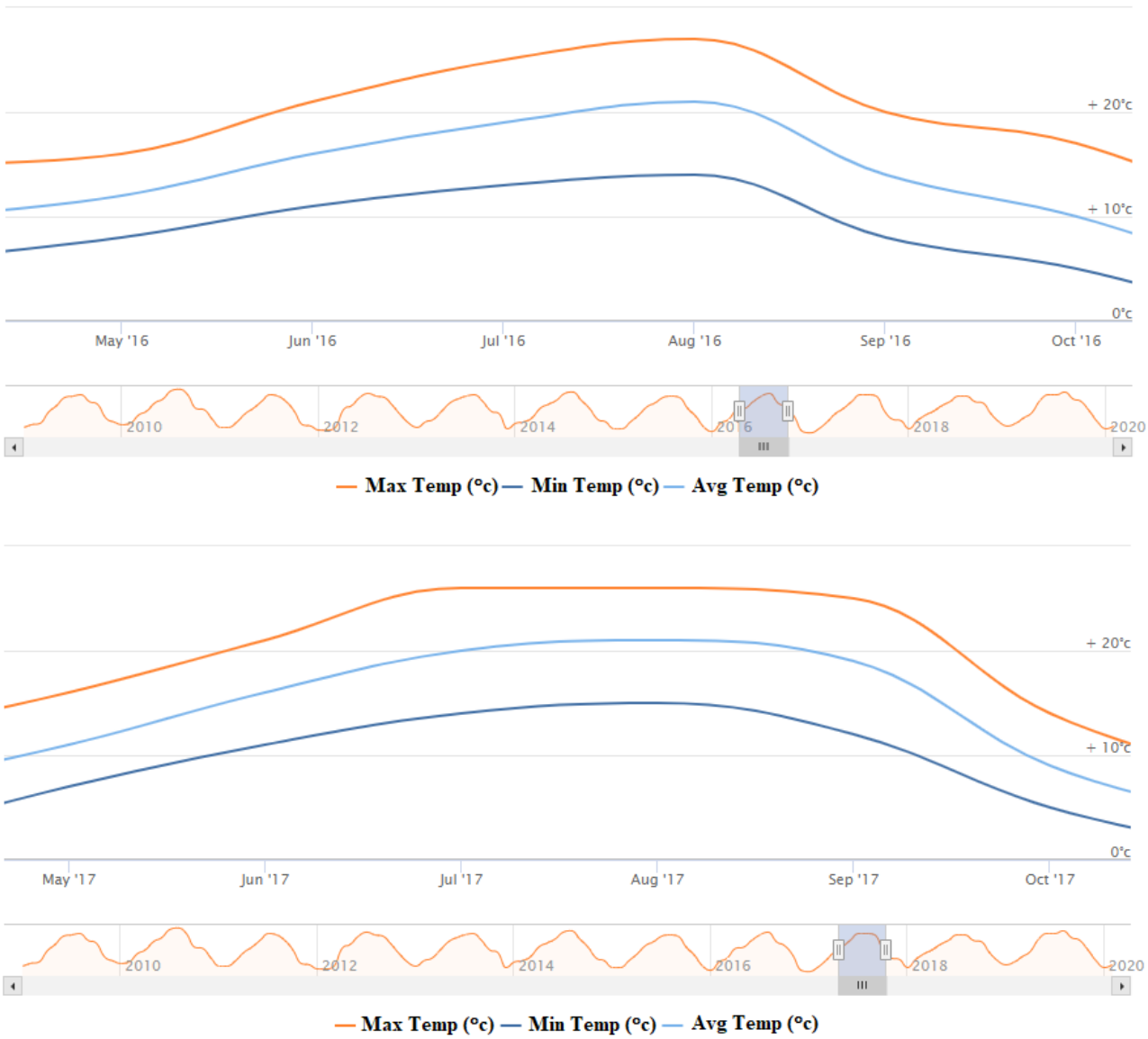

Figure 1. The climatic conditions during sowing years 2016 (a) and 2017 (b) including maximum, minimum, and average temperature in Niğde, Turkey.

which 5 Q groups and 2PCs were incorporated into the model. The equation used for MLM is given below.

$y = X\beta + Zu + e$

Where y = vector of the BLUEs of each genotype, β = vector containing fixed effects of the genetic marker and the intercept, u = is the vector of genotypic random additive genetic effects, X and Z = corresponding design matrices, e = is the vector of residuals in the model; the residual (e) and genetic (u) effects were assumed to be random, normally distributed, and independent.

3. Results and discussion

3.1. Floral traits performance

Analysis of variance (ANOVA) showed that genotypes, as well as check cultivars, had highly significant ($p < 0.001$) differences for all studied traits (Table 1). These results were found in line with the previous research studies, which also confirmed that the genotypic effect is the source of variations (Anoumaa et al., 2017; Tessema et al., 2020). Mean, maximum, minimum, and standard deviation for all studied traits are provided in Table 2. PL ranged from

Table 1. Summary of analysis of variance for studied floral traits in 237 potato genotypes.

ANOVA for PL							
Source	Df	Sum Sq	Mean Sq	F value		Pr(>F)	
Genotypes	256	194.59	0.7601	2.15E+28	<	2.20E-16	***
Control	4	22.518	5.6296	1.59E+29	<	2.20E-16	***
Residuals	24	78.67	1.24				
ANOVA for SL							
Source	Df	Sum Sq	Mean Sq	F value		Pr(>F)	
Genotypes	256	355.42	1.3884	4.75E+28	<	2.20E-16	***
Control	4	6.34	1.5838	5.42E+28	<	2.20E-16	***
Residuals	24	105.9	2.1				
ANOVA for PLAS							
Source	Df	Sum Sq	Mean Sq	F value		Pr(>F)	
Genotypes	256	219.783	0.8585	2.00E+28	<	2.20E-16	***
Control	4	20.121	5.0303	1.17E+29	<	2.20E-16	***
Residuals	24	98.6	3.9				

***Significant at $p < 0.001$

Table 2. Summary of mean, maximum and minimum for studied potato floral traits.

Variable	Minimum	Maximum	Mean	Std. deviation
PL2016	5.957	12.335	8.673	1.075
PL2017	3.390	9.980	6.119	0.897
Overall	5.533	9.921	7.396	0.736
Variable	Minimum	Maximum	Mean	Std. deviation
SL2016	4.537	8.397	6.687	0.779
SL2017	6.160	13.040	8.941	1.038
Overall	6.048	9.268	7.814	0.665
Variable	Minimum	Maximum	Mean	Std. deviation
PLAS2016	0.755	4.724	2.242	0.682
PLAS2017	0.950	6.090	2.833	0.813
Overall	1.310	4.473	2.537	0.527

8.76 mm to 12.33 mm with a mean value of 8.67 during 2016. Minimum and maximum PL was obtained by genotype CIP- 397077.16 and CIP- 395196.4, respectively (Table 2). During 2017, PL was ranged from 3.39 to 9.98 with a mean value of 6.11. Potato genotypes CIP-37 and CIP- 391585.179 reflected minimum and maximum PL in the year 2017. PL across both years (2016 and 2017) was ranged from 5.53 mm to 9.92 mm with a mean value of 7.39. Maximum and minimum PL was observed for genotypes CIP- 302499.30 and CIP- 391585.179, respectively. Presence of high PL in this study might be due to inclusion of diverse and increased germplasm number as compared to earlier studies. Mean SL during 2016 was 6.68, while Genotype-30 and Genotype-179 reflected minimum (4.53) and maximum (8.39) SL. During 2017, minimum and maximum SL was 6.16 and 13.04 showed by Genotype-161 and Genotype-126, while mean SL for whole germplasm was 8.94. Mean SL across both years (2016 and 2017) was 7.81, while minimum and maximum SL was 6.04 and 9.26 reflected by Genotype-161 and Genotype-126. Presence of high SL in this study might be due to inclusion of diverse and big germplasm population as compared to

earlier studies. Floral trait PLAS during 2016 was ranged from 0.75 (Genotype-165) to 4.72 (Genotype-125) with a mean value of 2.24. PLAS during 2017 was ranged from 0.95 (Genotype-22) to 6.09 (Genotype-126) with a mean value of 2.83. PLAS across both years (2016 and 2017) was ranged from 1.31 (Genotype-32) to 4.47 (Genotype-126) with a mean value of 2.53. Presence of increased PLAS in this study might be due to inclusion of diverse and large number of germplasms as compared to earlier studies.

Frequency distribution was performed for the studied floral traits using mean data of both years and revealed normal distribution (Figure 2). Pearson's correlation was also performed, and PL reflected highly significant and positive correlation with SL (r = 0.42) and PLAS (r = 0.28) (Figure 2). Correlation analysis is mainly applied to investigate the association among the traits, and the evaluated information can be best used for crop improvement by indirect selection of the components (Ali et al., 2020; Nadeem et al., 2020). As highly significant and positive correlation of SL was observed with PL and PLAS, the existence of correlation among studied traits can be due to genetic linkage or epistatic effects among various genes (Ozer et al., 2010).

Principal component analysis (PCA) was performed, and the first two PCs were considered to account for a total of 81.2% variation (Figure 3). SL and PL were found to be the key variation contributors in the 1st PC, while PLAS was the key variation contributor in the 2nd PC. PCA can be helpful to investigate the important plant traits that can utilize to characterize the variation among experimental

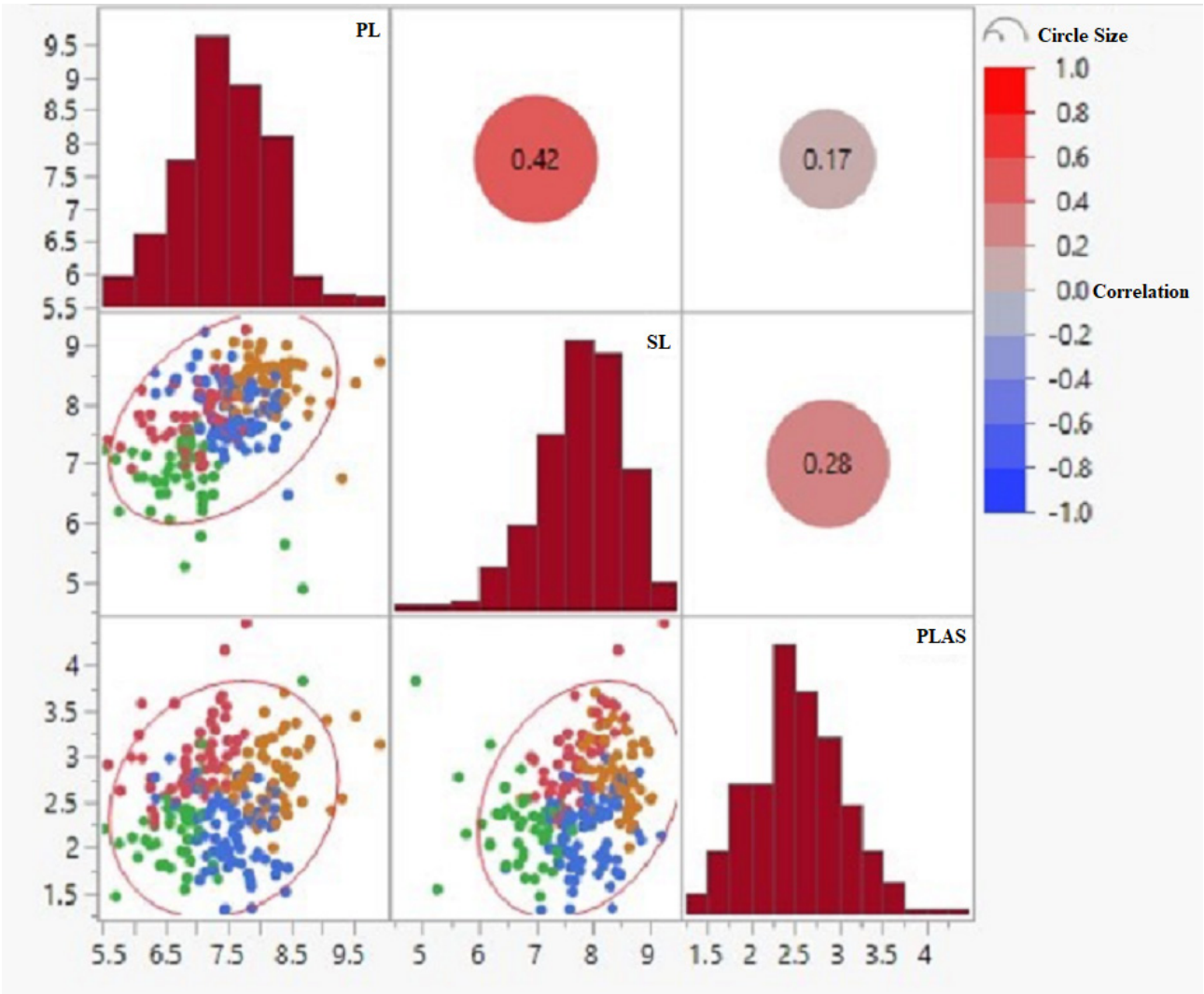

Figure 2. The whole figure shows frequency distribution and Pearson's correlation for studied traits in tetraploid potato germplasm.

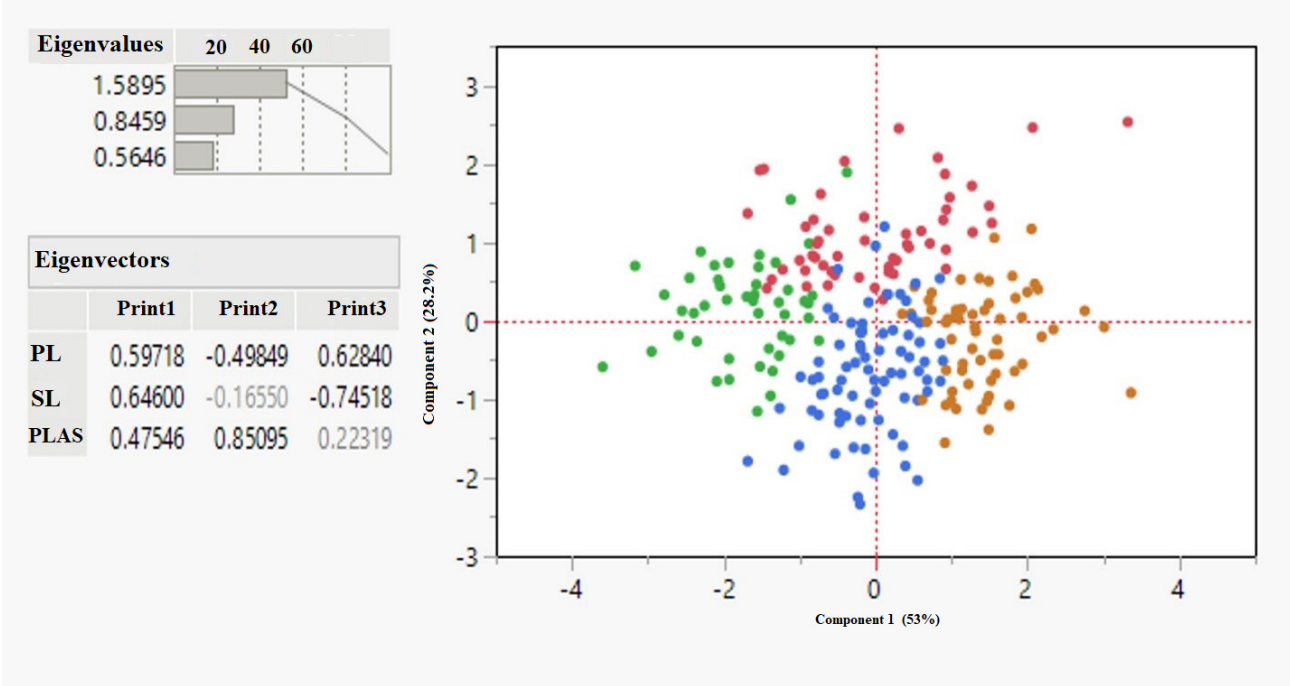

Figure 3. Principal component analysis for tetraploid potato germplasm.

materials (Chakravorty et al., 2013). According to Ali et al. (2020), traits consistently adding variation in each PC governed by genes can be helpful during selection to develop desirable cultivars.

The implemented constellation plot divided 237 tetraploid potato genotypes into two main populations on the basis of SL and PL (Figure 4). Population A was found bigger than population B by clustering a total of 138 genotypes. Population A was further clustered into A1 and A2 subpopulations. Subpopulation A1 clustered a total of 77 genotypes, while subpopulation A2 clustered 61 genotypes. Population B clustered a total of 99 genotypes and subdivided into subpopulations B1 and B2, respectively. Subpopulation B1 contains 53 genotypes, while subpopulation B2 clustered 46 genotypes. Correlation analysis was performed to identify relationship among the studied floral traits (Figure 4). Positive and significant correlation was observed among the three floral traits. Davis (2001) supported our result as he observed a positive and significant correlation between the floral traits, i.e., anther length and ovary length.

3.2. Marker trait association

3.2.1. Stamen length (SL)

Stamen length is one of the important traits taken under consideration for the GWAS because it plays an important role in determining the pollination pattern of the flower (Scot et al., 2004). A total of 5 SNPs on different chromosomes were identified for stamen length out of 12,720 SNPs used (Table 3). Out of these 5 SNPs, Solcap_snp_c1_9717 and Solcap_snp_c1_9724 were localized on chromosome number 2, solcap_snp_c2_9035, and solcap_snp_c2_9218 were mapped on chromosome number 6 and solcap_snp_c2_444474 was mapped on chromosome number 8 (Figure 5). Li et al. (2001) aimed to investigate the QTL associated with stamen length and ratio of stigma exertion and stated that QTLs controlling stamen length located at the region of C424-G39 (chromosome 2) and C2807-C1263 (chromosome 9), respectively. We also found some significant associations (Solcap_snp_c1_9717 and Solcap_snp_c1_9724) on chromosome 2. We can consider these markers for future studies of stamen length in potatoes. Solcap_snp_c1_9717 and Solcap_snp_c1_9724 catalyze the hydrolysis of ATP coupled with the transport of calcium (https://www.uniprot.org/uniprot/Q13510). Solcap_snp_c2_9035 helps in transcription regulation, while solcap_snp_c2_9218 is a member of the GTP -cyclohydrolase family of enzymes. Guanosine triphosphate cyclohydrolase (GTPCH) is part of the folate and biopterin biosynthesis pathways. It is responsible for the hydrolysis of guanosine triphosphate (GTP) to form 7,8-dihydroneopterin triphosphate, and the SNP that was found to be associated on chromosome number 8 (solcap_snp_c2_44474) and the protein associated with this SNP is 3β-Hydroxysteroid dehydrogenase (3β-HSD) and is also suggested to be involved in cardenolide biosynthesis (Seidel et al., 1990).

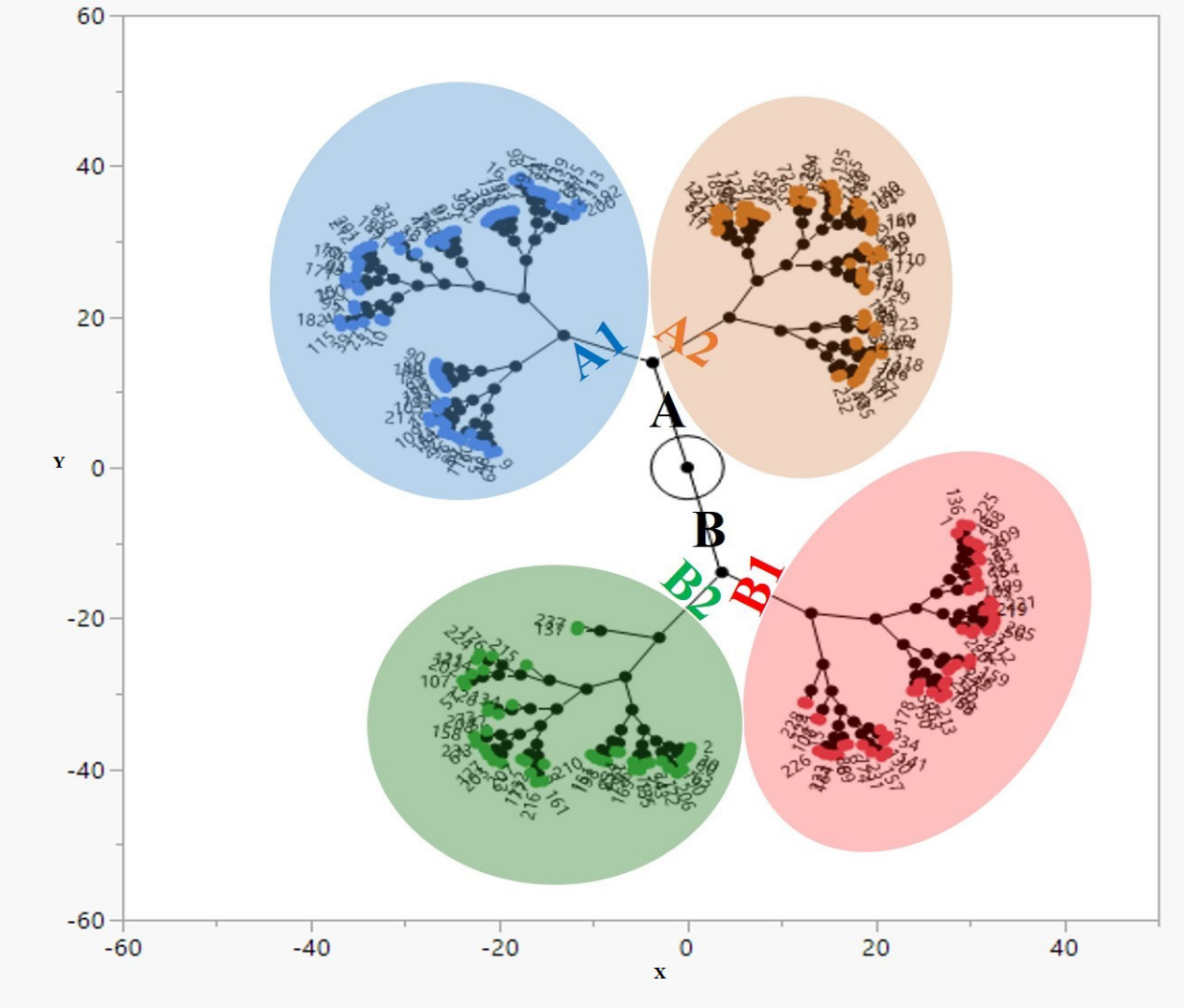

Figure 4. A constellation plot analysis shows the clustering of tetraploid potato germplasm panel having 237 potato lines.

3.2.2. Pistil length (PL)

Pistil length is important to determine the selfing and crossing nature of the plants (Fernández et al., 2009). If the length of pistil is more, then the chances of selfing minimize and vice versa. Pistil length in potato flower was associated with 5 SNPs during 2016, 2017, and BLUE (Table 3). The identified SNPs like solcap_snp_c2_43285 was localized on chromosome number 1. Similarly, SNPs solcap_snp_c2_42381, solcap_snp_c2_8513 and solcap_snp_c2_23030 were localized on chromosome number 5, while SNP solcap_snp_c2_34608 was localized on chromosome number 8. (Figure 6). Solcap_snp_c2_43285 is actually hypoxia-responsive family protein and is responsible for mitochondrial respiratory chain complex IV assembly. Solcap_snp_c2_42381 is a At3g63230 gene that responses to starvation. Solcap_snp_c2_8513 is a kinase enzyme that modifies other proteins by chemically adding phosphate groups to them (phosphorylation). Phosphorylation usually results in a functional change of the target protein (substrate) by changing enzyme activity, cellular location, or association with other proteins. Solcap_snp_c2_23030 is pentatricopeptide repeat- containing protein and is associated with multiple functions like chloroplast mRNA processing, chloroplast organization as well as regulation of translation. Solcap_snp_c2_34608 is an uncharacterized protein Ycf49.

3.2.3. Pistil length above stamen (PLAS)

Genome-wide association analysis identified 5 SNPs for pistil length above stamen present on different chromosomes (Table 3). Among these 5 SNPs, solcap_

Table 3. Markers associated with various flowering traits in potato germplasm under study.

Sr. No.	SNP	CHR	BP	p Value	R2 Value	Genes Associated	PGSC Loci	genotype	Trait
1	solcap_snp_c2_43285	1	57765672	3.75E-04	0.06276	Hypoxia induced protein conserved region containing protein	PGSC0003DMT400029479	C:T	PL
2	solcap_snp_c2_42381	5	47101740	3.22E-04	0.058485	DUF581 domain containing protein	PGSC0003DMT400056164	A:G	
3	solcap_snp_c2_8513	5	50584053	4.20E-04	0.055798	Kinase	PGSC0003DMT400060442	C:T	
4	solcap_snp_c2_23030	5	4729224	5.98E-04	0.062192	Pentatricopeptide repeat-containing protein	PGSC0003DMT400047445	A:G	
5	solcap_snp_c2_34608	8	53386131	5.30E-05	0.096166	Ycf49	PGSC0003DMT400010086	T:G	
6	solcap_snp_c1_9717	2	22749216	5.53E-04	0.051075	Type IIB calcium ATPase MCA5	PGSC0003DMT400033890	A:G	SL
7	solcap_snp_c1_9724	2	22749951	5.53E-04	0.051075	Type IIB calcium ATPase MCA5	PGSC0003DMT400033890	C:T	
8	solcap_snp_c2_9035	6	57680515	7.98E-04	0.050347	TSI1	PGSC0003DMT400078294	G:A	
9	solcap_snp_c2_9218	6	58219183	9.24E-04	0.05953	GTP cyclohydrolase I	PGSC0003DMT400051801	T:C	
10	solcap_snp_c2_44474	8	49091052	9.53E-04	0.064834	3-beta hydroxysteroid dehydrogenase/isomerase family protein	PGSC0003DMT400038371	C:T	
11	solcap_snp_c2_51111	2	6646692	3.32E-04	0.06013	Conserved gene of unknown function	PGSC0003DMT400054346	A:G	PLAS
2	solcap_snp_c1_8118	2	39940755	5.72E-04	0.068418	Aldehyde dehydrogenase	PGSC0003DMT400009104	T:C	
13	solcap_snp_c2_36061	4	58752313	6.74E-04	0.068034	F-box family protein	PGSC0003DMT400002060	A:G	
14	solcap_snp_c2_45155	7	46903202	9.43E-04	0.050034	Galactosyltransferase family protein	PGSC0003DMT400074517	T:C	
15	solcap_snp_c1_8642	12	3921843	9.56E-06	0.101579	Mitochondrial-processing peptidase subunit alpha	PGSC0003DMT400000784	T:A	

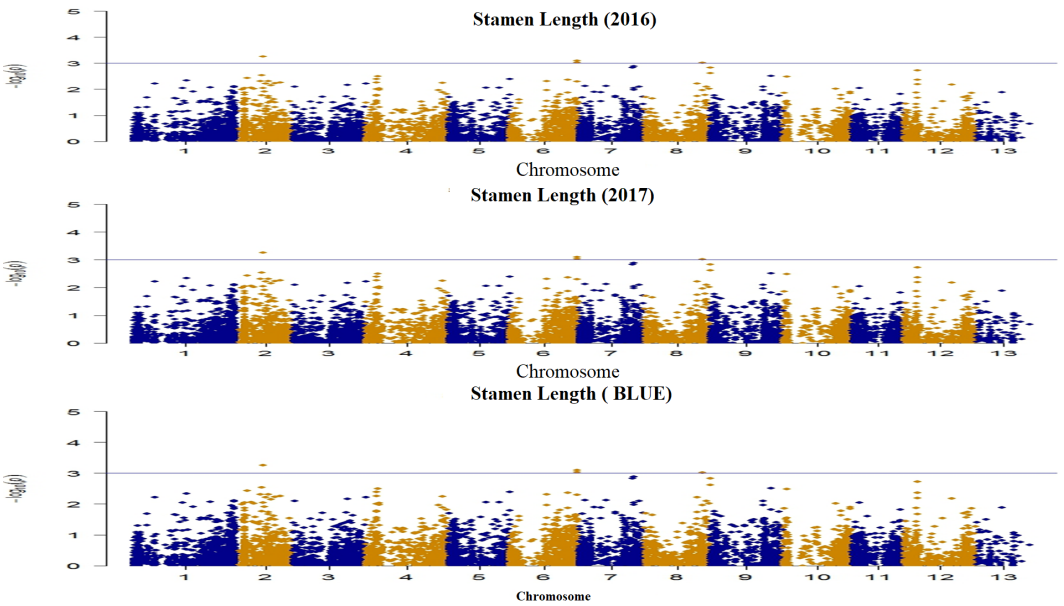

Figure 5. Manhattan plots for the data sets of 2016, 2017, and BLUE for stamen length trait, while 13 represents the unmapped markers.

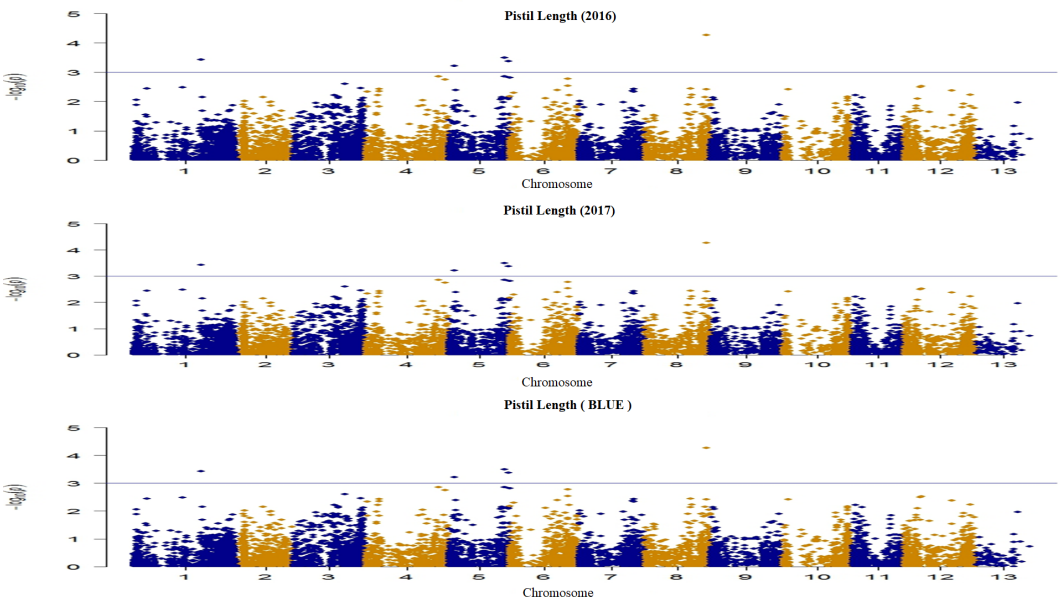

Figure 6. Manhattan plots for the data set of 2016, 2017, and BLUE for Pistil length trait, while 13 represents the unmapped markers.

snp_c2_51111 and solcap_snp_c1_8118 were localized on chromosome number 2, solcap_snp_c2_36061 was mapped on chromosome 4, solcap_snp_c2_45155 was present on chromosome 7, and solcap_snp_c1_8642 was localized on chromosome number 12 (Figure 7). Solcap_snp_c2_51111 is basically conserved gene of unknown function, while solcap_snp_c1_8118 is aldehyde dehydrogenase gene, which is a superfamily of enzymes that detoxify a variety of endogenous and exogenous aldehydes, and are required for the biosynthesis of retinoic

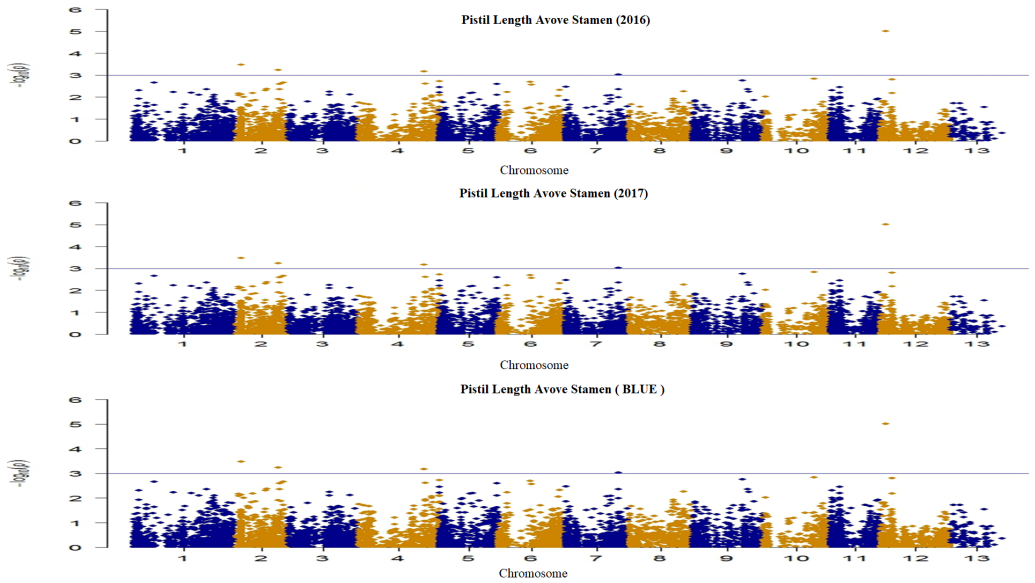

Figure 7. Manhattan plots for the data sets of 2016, 2017, and BLUE for Pistil length above stamen trait, while 13 represents the unmapped markers.

acid (RA) and other molecular regulators of cellular function (Vassalli, 2019). Solcap_snp_c2_36061 is F-box family protein and is involved in many plants vegetative and reproduction growth and development. For example, F-box protein-FOA1 is involved in abscisic acid (ABA) signaling to affect the seed germination (Peng et al., 2012). ACRE189/ACIF1 can regulate cell death and defense when the pathogen is recognized in the tobacco and tomato plant (Ha et al., 2008).

Solcap_snp_c2_45155 is galactosyltransferase family protein results in the formation of a Lewis a epitope, a trisaccharide (Fuc-alpha-(1->4) [Gal-beta-(1->3)] GlcNAc) characteristic of plant protein N-linked oligosaccharides. Solcap_snp_c1_8642 is mitochondrial-processing peptidase subunit alpha and involved in protein processing and targeting to mitochondrion, mitochondrial calcium ion transmembrane transport as well as in proteolysis.

4. Conclusion

Present study unrevealed the genetic basis controlling important floral traits. Some useful SNPs that could be beneficial in future breeding programs for potato were found. Mean data across both years reflected sufficient amount of phenotypic variability for the studied floral traits. Analysis of variance for the studied floral traits observed highly significant genotypic effects. Correlation analysis revealed highly significant and positive associations among the three floral traits. A total of two main populations (A and B) with further two subpopulations were obtained from constellation plot analysis based on stamen and pistil length. The genome-wide association analysis revealed a total of 15 significant SNPs for stamen length, pistil length, and pistil length above stamen across both years. The identified candidate genes from this study might be useful to the potato breeders. After validation, these markers can be used in future potato breeding programs to accelerate the breeding process as well as to cope with the needs of increasing population.

References

Achenbach U, Paulo J, Ilarionova E, Lubeck J, Strahwald J et al. (2009). Using SNP markers to dissect linkage disequilibrium at a major quantitative trait locus for resistance to the potato cyst nematode Globodera pallida on potato chromosome V. Theoretical and Applied Genetics 118: 619-629.

Ali F, Yilmaz A, Chaudhary HJ, Nadeem MA, Rabbani MA et al. (2020). Investigation of morpho-agronomic performance and selection indices in the international safflower panel for breeding perspectives. Turkish Journal of Agriculture and Forestry 44: 103–120. doi: 10.3906/tar-1902-49

Ali F, Yılmaz A, Nadeem MA, Habyarimana E, Subaşı I et al. (2019). Mobile genomic element diversity in world collection of safflower (*Carthamus tinctorius* L.) panel using iPBS-retrotransposon markers. PloS one 14 (2): e0211985.

Arteaga MC, Bello-Bedoy R, Leon-de la Luz JL, Delgadillo J et al. (2015). Phenotypic variation of flowering and vegetative morphological traits along the distribution for the endemic species Yucca capensis (Agavaceae). Botanical Sciences 93: 765-70.

Ashman TL, Majetic CJ (2006). Genetic constraints on floral evolution: a review and evaluation of patterns. Heredity; 96: 343–352.

Bali S, Patel G, Novy R, Vining K, Brown C et al. (2018). Evaluation of genetic diversity among Russet potato clones and varieties from breeding programs across the United States. PloS One 13 (8): e0201415.

Barut M, Nadeem MA, Karakoy T, Baloch FS (2020). DNA fingerprinting and genetic diversity analysis of world quinoa germplasm using iPBS-retrotransposon marker system. Turkish Journal of Agriculture and Forestry 44 (5): 479-91.

Berdugo-Cely J, Valbuena RI, Sa´nchez-Betancourt E, Barrero LS, Yockteng R (2017). Genetic diversity and association mapping in the Colombian Central Collection of Solanum tuberosum L. Andigenum group using SNPs markers. PLoS ONE 12 (3): e0173039. doi: 10.1371/journal.pone.0173039

Berg RL (1960). The ecological significance of correlation pleiades. Evolution; 14: 171–180.

Bradbury PJ, Zhang ZW, Kroon DE, Casstevens TM, Ramdoss Y et al. (2007). TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 2: 2633-2635. doi: 10.1093/bioinformatics/btm308

Brock MT, Weinig C (2007). Plasticity and environment-specific covariances: an investigation of floral-vegetative and within flower correlations. Evolution 61: 2913–2924.

Carputo D, Aversano R, Frsciante L (2004). Breeding potato for quality traits. In meeting of the physiology section of the European Association for Potato Research 684: 55-64.

Carr DE, Fenster CB (1994). Levels of genetic variation and covariation for Mimulus (Scrophulariaceae) floral traits. Heredity 72: 606–618.

Chakravorty A, Ghosh PD, Sahu PK (2013). Multivariate analysis of phenotypic diversity of landraces of rice of West Bengal. American Journal of Experimental Agriculture 3: 110-123.

Conner J, Via S (1993). Patterns of phenotypic and genetic correlations among morphological and life-history traits in wild radish, Raphanus raphanistrum. Evolution 47: 704–711.

Conner JK, Sterling A (1995). Testing hypotheses of functional relationships--a comparative survey of correlation patterns among floral traits in 5 insect-pollinated plants. American Journal of Botany 1399–1406.

Davis SL (2001). Phenotypic and genetic correlations among floral traits in two species of Thalictrum. Journal of Heredity 92: 361-366.

Dempewolf H, Eastwood RJ, Guarino L, Khoury CK, Müller JV et al. (2014). Adapting Agriculture to Climate Change: A Global Initiative to Collect, Conserve, and Use Crop Wild Relatives, Agroecology and Sustainable Food Systems 38 (4): 369-377. doi: 10.1080/21683565.2013.870629

Edwards CE, Weinig C (2011). The quantitative-genetic and QTL architecture of trait integration and modularity in Brassica rapa across simulated seasonal settings. Heredity 106: 661-677.

FAOSTAT (2017). Food and Agriculture Organization of the United Nations www.faostat.fao.org. Accessed 25 July 2020

FAOSTAT (2021). Food and Agriculture Organization of the United Nations www.faostat.fao.org. Accessed 21 Jan 2021

Felcher KJ, Coombs JJ, Massa AN, Hansey CN, Hamilton JP et al. (2012). Integration of two diploid potato linkage maps with the potato genome sequence. PLoS ONE 7 (4): e36347

Fernández VA, Galetto L, Astegiano J (2009). Influence of flower functionality and pollination system on the pollen size-pistil length relationship. Organisms Diversity and Evolution 9: 75-82.

Godfray H, Charles J (2010). Food security: the challenge of feeding 9 billion people. Science 327: 812-818.

Gomez KA, Gomez AA (1984) Statistical Procedures for Agricultural Research, 2nd ed. John Wiley & Sons, New York (NY).

Ha C, Coombs S, Revill PA, Harding RM, Vu M et al. (2008). Design and application of two novel degenerate primer pairs for the detection and complete genomic characterization of potyviruses. Archives of. Virology 153: 25-36.

Hill WG, Robertson A (1968). Linkage disequilibrium in finite populations. Theoretical and Applied Genetics 38: 226-231.

Jansky SH, Spooner DM (2018). The evolution of potato breeding. Plant Breeding Reviews 41: 169-214.

Juenger T, Purugganan M, Mackay TFC (2000). Quantitative trait loci for floral morphology in Arabidopsis thaliana. Genetics 156: 1379–1392.

Karık Ü, Nadeem MA, Habyarimana E, Ercişli S, Yildiz M et al. (2019). Exploring the genetic diversity and population structure of Turkish laurel germplasm by the iPBS-retrotransposon marker system. Agronomy 9 (10): 647.

Kastner T, Rivas MJI, Koch W, Nonhebel S (2012). Global changes in diets and the consequences for land requirements for food. Proceedings in National. Academy of Sciences 109 (18): 6868-6872. doi: 10.1073/pnas.1117054109

Khoury CK, Bjorkman AD, Dempewolf H, Ramírez-Villegas J, Guarino L et al. (2014). Increasing homogeneity in global food supplies and the implications for food security. Proceedings in National. Academy of Sciences 111: 4001-4006.

Knapp S (2008). Potatoes and poverty. Nature 455: 170-171.

Kullo IJ, Shameer K, Jouni H, Lesnick TG, Pathak J et al. (2014). The ATXN2 SH2B3 locus is associated with peripheral arterial disease: an electronic medical record-based genome-wide association study. Frontiers in Genetics 5: 166. doi: 10.3389/fgene.2014.00166

Lande R (1980). The genetic covariance between characters maintained by pleiotropic mutations. Genetics 94: 203–215.

Li L, Strahwald J, Hofferbert HR, Lubeck J, Tacke E et al. (2001). DNA variation at the invertase locus invGE/GF is associated with tuber quality traits in populations of potato breeding clones. Genetics 170 (2): 813–821

Mosquera T, lvarez MF, Goamez JM, Muktar MS, Paulo MJ et al. (2016). Targeted and untargeted approaches unravel novel candidate genes and diagnostic SNPs for quantitative resistance of the potato (*Solanum tuberosum* L.) to Phytophthora infestans causing the late blight disease. PLoS One 11 (6): e0156254. doi: 10.1371/journal.pone.0156254

Nadeem MA (2021). Deciphering the genetic diversity and population structure of Turkish bread wheat germplasm using iPBS-retrotransposons markers. Molecular Biology Reports 48: 6739–6748. doi: 10.1007/s11033-021-06670-w

Nadeem MA, Habyarimana E Karakoy T, Baloch FS (2021). Genetic dissection of days to flowering via genome-wide association studies in Turkish common bean germplasm. Physiology and Molecular Biology in Plants doi: 10.1007/s12298-021-01029-8.

Nadeem MA, Karaköy T, Yeken MZ, Habyarimana E, Hatipoğlu R et al. (2020). Phenotypic characterization of 183 Turkish common bean accessions for agronomic, trading, and consumer-preferred plant characteristics for breeding purposes. Agronomy 10: 272.

Nadeem MA, Nawaz MA, Shahid MQ, Doğan Y, Comertpay G et al. (2018). DNA molecular markers in plant breeding: current status and recent advancements in genomic selection and genome editing. Biotechnology and Biotechnological Equipments 32: 261-285. doi: 10.1080/13102818.2017.1400401

Oneil P, Schmitt J (1993). Genetic constraints on the independent evolution of male and female reproductive characters in the tristylous plant Lythrum salicaria. Evolution 47: 1457–1471.

Ozer S, Karaköy T, Toklu F, Baloch FS, Kilian B et al. (2010). Nutritional and physicochemical variation in Turkish kabuli chickpea (Cicer arietinum L.) landraces. Euphytica 175: 237–249.

Parra-Galindo MA, Soto-Sedano JC, Mosquera-Vásquez T, Roda F (2021). Pathway-based analysis of anthocyanin diversity in diploid potato. PLoS ONE 16 (4): e0250861. doi: 10.1371/journal.pone.0250861

Paudel D, Dareus R, Rosenwald J, Muñoz-Amatriaín M et al. (2021). Genome-Wide Association Study Reveals Candidate Genes for Flowering Time in Cowpea (Vigna unguiculata [L.] Walp.). Frontiers in Genetics 12: 667038. doi: 10.3389/fgene.2021.667038

Peng J, Yu D, Wang L, Xie M, Yuan C et al. (2012). Arabidopsis F-box gene FOA1 involved in ABA signaling. Science China Life Sciences 55: 497–506. doi: 10.1007/s11427-012-4332-9

Petersen RG (1985). Augmented designs for preliminary yield trials (revised). Rachis 4: 27–32.

Raggi L, Caproni L, Carboni A, Negri V (2019). Genome-wide association study reveals candidate genes for flowering time variation in common bean (Phaseolus vulgaris L.). Frontiers of Plant Science 10: 962. doi: 10.3389/fpls.2019.00962

Raman H, Raman R, Qiu Y (2019). GWAS hints at pleiotropic roles for FLOWERING LOCUS T in flowering time and yield-related traits in canola. BMC Genomics 20: 636. doi: 10.1186/s12864-019-5964-y

Scott RJ, Spielman M, Dickinson HG (2004). Stamen structure and function. The Plant Cell 16:S46-S60.

Seidel S, Kreis W, Reinhard E (1990). Δ5-3β-Hydroxysteroid dehydrogenase/ Δ5-Δ4-ketosteroid isomerase (3βHSD), a possible enzyme of cardiac glycoside biosynthesis, in cell cultures and plants of Digitalis lanata EHRH. Plant Cell Reports 8: 621–624.

Shimira F, Boyaci HF, Çilesiz Y, Nadeem MA, Baloch FS et al. (2021). Exploring the genetic diversity and population structure of scarlet eggplant germplasm from Rwanda through iPBS-retrotransposon markers. Molecular Biology Reports 48: 6323–6333. doi: 10.1007/s11033-021-06626-0

Spooner DM, McLean K, Ramsay G, Waugh R, Bryan GJ (2005). A single domestication for potato based on multi locus amplified fragment length polymorphism genotyping. Proceeding in National Academy of Sciences 102 (41): 14694–14699

Stich B, Urbany C, Hoffmann P, Gebhardt C (2013). Population structure and linkage disequilibrium in diploid and tetraploid potato revealed by genome-wide high-density genotyping using the SolCAP SNP array. Plant Breeding 132 (6): 718-724.

Tilman D, Blazer C, Hill J, Befort BL (2011). Global food demand and the sustainable intensification of agriculture. Proceeding in National Academy of Sciences 108: 20260-20264.

van Kleunen M, Ritland K (2004). Predicting evolution of floral traits associated with mating system in a natural plant population. Journal of Evolutionary Biology 17: 1389–1399.

Vassalli G (2019). Aldehyde Dehydrogenases: Not Just Markers, but Functional Regulators of Stem Cells. Stem Cells International 2019: 1-15 doi: 10.1155/2019/3904645

Vos PG, Paulo MJ, Voorrips RE, Visser RG, van Eck HJ et al (2017). Evaluation of LD decay and various LD decay estimators in simulated and SNP-array data of tetraploid potato. Theoretical and Applied Genetics 130 (1): 123–35.

Wilson S, Zheng C, Maliepaard C, Mulder HA, Visser RGF et al. (2021). Understanding the Effectiveness of Genomic Prediction in Tetraploid Potato. Frontiers in Plant Science 12: 672417. doi: 10.3389/fpls.2021.672417

Yildiz M, Kocak M, Nadeem MA, Cavagnaro P, Barboza K et al. (2020). Genetic diversity analysis in the Turkish pepper germplasm using iPBS retrotransposonbased markers. Turkish Journal of Agriculture and Forestry 44 (1): 1-4.

Yousaf MF, Demirel U, Naeem M, Caliskan ME (2021) Association mapping reveals novel genomic regions controlling some root and stolon traits in tetraploid potato (Solanum tuberosum L.). 3 Biotech 11: 174. doi: 10.1007/s13205-021-02727-6

Yuan Y, Cairns JE, Babu R, Gowda M, Makumbi D et al. (2019). Genome-wide association mapping and genomic prediction analyses reveal the genetic architecture of grain yield and flowering time under drought and heat stress conditions in maize. Frontiers in Plant Sciences 9: 1919. doi: 10.3389/fpls.2018.01919

Zhang Z, Ersoz E, Lai CQ, Todhunter RJ, Tiwari HK et al. (2010). Mixed linear model approach adapted for genome-wide association studies. Nature Genetics 42: 355-360.

Zia MAB, Demirel U, Nadeem MA, Caliskan ME (2020). Genome-wide association study identifies various loci underlying agronomic and morphological traits in diversified potato panel. Physiology and Molecular Biology in Plants 26: 1003–1020.