Phyton-
International Journal of
Experimental Botany

Tech Science Press

DOI: 10.32604/phyton.2024.048992

ARTICLE

Genome-Wide Analysis for Yield-Related Agronomic and Biochemical Traits of Chinese and Bangladeshi Grass Pea Genotypes Using SSR Markers

Md. Mosiur Rahman[1,2], Md. Ruhul Quddus[3], Quanle Xu[4], Muhammad Malek Hossain[2], Rong Liu[1], Mengwei Li[1], Xin Yan[1], Guan Li[1], Yishan Ji[1], Chenyu Wang[1], Ashutosh Sarker[5], Tao Yang[1], Xuxiao Zong[1], Md. Monoar Hossain[6], Saleh Alfarraj[7], Mohammad Javed Ansari[8], Sagar Maitra[9,*] and Akbar Hossain[10,*]

[1]Institute of Crop Sciences, Chinese Academy of Agricultural Sciences (CAAS), Beijing, 100081, China

[2]Pulses Research Center, Bangladesh Agricultural Research Institute (BARI), Gazipur, 1701, Bangladesh

[3]Hybrid Rice Division, Bangladesh Rice Research Institute (BRRI), Gazipur, 1701, Bangladesh

[4]Northwest Agriculture & Forestry University, Yangling, 712199, China

[5]International Center for Agricultural Research in the Dry Areas (ICARDA), New Delhi, 110012, India

[6]Wheat Breeding Division, Bangladesh Wheat and Maize Research Institute, Dinajpur, 5200, Bangladesh

[7]Zoology Department, College of Science, King Saud University, Riyadh, 11451, Saudi Arabia

[8]Department of Botany, Hindu College Moradabad (Mahatma Jyotiba Phule Rohilkhand University, Bareilly), Moradabad, 244001, India

[9]Department of Agronomy and Agroforestry, Centurion University of Technology and Management, Odisha, 761211, India

[10]Soil Science Division, Bangladesh Wheat and Maize Research Institute, Dinajpur, 5200, Bangladesh

[*]Corresponding Authors: Sagar Maitra. Email: sagar.maitra@cutm.ac.in; Akbar Hossain. Email: akbar.hossain@bwmri.gov.bd

Received: 23 December 2023 Accepted: 30 April 2024 Published: 30 August 2024

ABSTRACT

Grass pea (*Lathyrus sativus* L.) is an imperative food crop cultured in dryland agricultural ecology. It is a vital source of dietary protein to millions of populaces living in low-income countries in South-East Asia and Africa. This study highlights the improvement of genomic properties and their application in marker-trait relationships for 17 yield-related characters in 400 grass pea genotypes from China and Bangladesh. These characters were assessed via 56 polymorphic markers using general linear model (GLM) (P+G+Q) and mixed linear model (MLM) (P+G+Q+K) in the tassel software based on the linkage disequilibrium and population structure analysis. Population structure analysis showed two major groups and one admixed group in the populace. Statistically significant loci pairs of linkage disequilibrium (LD) mean value (D′) was 0.479. A total of 99 and 61 marker-trait associations in GLM and MLM models allied to the 17 traits were accepted at a 5% level of significance. Among these markers, 21 markers were associated with more than one trait; 12 marker-trait associations passed the Bonferroni correction threshold. Both models found six markers C41936, C39067, C34100, C47146, C47638, and C43047 significantly associated with days to maturity, flower color, plant height, and seed per pod were detected in the Hebei and Liaoyang location ($p \leq 0.01$), and the interpretation rate (R^2 value) 11.2% to 43.6%. Conferring to the consequences, the association analysis methodology may operative system for quantitative, qualitative, and biochemical traits related to gene position mapping and support breeders in improving novel approaches for advancing the grass pea quality.

KEYWORDS

Grass pea; dryland agriculture; genome-wide association; yield contributing traits; SSR markers

1 Introduction

Grass pea (*Lathyrus sativus* L.) is a vital pulse crop of profitable importance in Bangladesh, India, Pakistan, Nepal, and Ethiopia [1]. It is also cultured in Central, South, and Eastern Europe, West Asia, and North Africa [2]. *Grass pea* is mostly self-pollinated and diploid with 2n = 14, although the outcrossing rate is 2.2% [3]. Nutritional security, domestic food, and profitable earnings are possible through reinforcement and extension of the cropping patterns of legumes [4].

The *Lathyrus* genus has an unlimited implication as food and silage but is neglected. High protein content, erosion control, nitrogen fixation, and insect and disease resistance are extra advantages of *Lathyrus* species [5]. *Lathyrus* cultivation is deteriorating, putting the species at risk of genetic loss [6]. *Lathyrus* is a mineral source for cattle in Mediterranean climates [7,8]. Grass pea produces more seeds than other legumes, particularly in dry conditions, and does not shatter their pods. Farmers harvest the plant for seed production over feed [9]. Despite many advantages, consuming an anti-nutritional issue, β-N-oxalyl-L-α, β-diamino propionic acid (β-ODAP) in leaves of the plant with seeds is that reasons of a neurological disorder that consequences in stable paralysis of the lower limbs in individuals [10,11].

The main goals of *grass pea* breeding by reduce β-ODAP content and increase the combination of yield components such as seeds pod^{-1}, pod plant$^{-1,}$ and forage yield varieties of grass pea [2]. The discovery of molecular markers associated with low diamino propionic acid (ODAP) allele's transfer of this trait into locally adapted germplasm is a promising method for grass pea breeding [12].

Molecular markers provide various advantages over traditional phenotyping processes in plant materials regardless of environmental impacts [13]. Inter simple sequence repeat marker (ISSR) can be used in genome mapping, evolutionary biology, and genetic diversity for high polymorphism [14]. Expressed sequence tag-based simple sequence repeat (EST-SSR) markers may be used for proportional mapping and genetic linkage maps in a variety of species [15]. A single marker has specific advantages, but when combined, they produce a broad application in the assessment of population structure, genetic differences, and aided selection for crop development [14,16,17]. Some studies have revealed that both EST-SSR and ISSR markers are largely used to advance the genetic linkage maps [18].

Mostly, self-pollinated crop shows a high level of linkage disequilibrium (LD) and cross-pollinated crops show low LD because they will have a greater recombination degree and the linkage disequilibrium breakdown between gene loci [19]. High linkage disequilibrium performs small mapping resolution, whereas Association mapping in populations with low linkage disequilibrium needs many markers [20]. The non-sampling non-random association of alleles among linked or unlinked loci raised by linkage disequilibrium is based on Relationship mapping to detect genetic regions related to agronomic traits [19,21,22]. The population structure of germplasm pools is also essential for relations between genetic and useful diversity and is suitable for association studies [23]. As a result, population structure is involved as an effect in the model of association analysis [18,24].

Most of the traits such as yield, agronomy, quality, and resistance belong to quantitative traits measured by several genes from segregating populations in the crop. The discovery of quantitative trait loci (QTL) with slight contributions to phenotypic traits and atmospheric sensitivity is challenging [25]. The natural populations are taken as the experimental tools, identifying several alleles on the identical locus and

directing the single loci. Linkage analysis detects the allele regulatory an object trait in advance; association analysis modifies the object of the gene rapidly [26].

The correlation coefficient of path analysis has a direct and indirect effect on yield per plant and yield-related traits. Path analysis exposes whether the association of the traits with yield is due to their direct effect or is a consequence of their indirect effect via other traits. Path analysis documents the division of the correlation coefficient between its components or mechanisms [27]. Path analysis is equitable by assessing the direct effect of one variable on another and also separates the correlation for its mechanisms [28]. The purposes of this study were (a) to evaluate the population structure and the genetic diversity of the 400 grass pea germplasm, (b) to detect SSR markers related to the studied traits, and (c) to evaluate the comprehensive effects of allelic arrangement for breeding visions.

2 Materials and Methods

2.1 Plant Materials

A total of 400 grass pea genotypes with 200 accessions from the Center for Crop Germplasm Resources, Chinese Academy of Agricultural Sciences (CAAS), China Which collected from China, Bangladesh, Afghanistan, Armenia, Georgia, Azerbaijan, Syria, Spain, Germany, Ukraine, Russia, Italy, Czechoslovakia, Netherland, Slovakia, France, Algeria, Tunisia, Tanzania and 200 accessions from Plant Genetic Resources Center, Bangladesh Agricultural Research Institute (BARI), Gazipur which collected from different regions of Bangladesh (Tables 1 and S1).

Table 1: Geographical origin-based distribution of 400 accessions from annual *Lathyrus sativus* L. species

Origin	Origin of country	Amount of accessions	Longitude (°E)	Latitude (°N)
Asia	China	22	35.861	104.195
	Bangladesh	202	23.685	90.356
	Afghanistan	2	33.939	67.710
	Armenia	3	40.069	45.038
	Georgia	1	32.165	82.900
	Azerbaijan	1	40.143	47.576
	Syria	80	34.802	38.996
Europe	Spain	2	40.463 * (Longitude °W)	3.749
	Germany	2	51.165	10.451
	Ukraine	3	48.379	31.165
	Russia	70	61.524	105.318
	Italy	2	41.871	12.567
	Czechoslovakia	1	49.817	15.473
	Netherlands	1	52.132	5.291
	Slovakia	1	48.669	19.699
	France	1	46.227	2.213
Africa	Algeria	1	28.033	1.659
	Tunisia	1	33.886	9.537
	Tanzania	4	6.369	34.888

Note: * For 'Spain', Longitude °W, but other countries Longitude °E.

2.2 Procedures of Phenotypic Data, ODAP Detection, and Molecular Characterization

The 400 grass pea accessions (G1–G400) were planted in both locations. About 200 Bangladeshi accessions (G201–G400) were damaged by cold weather but the 200 Chinese accessions (G1–G200) were sustained in Liaoyang province in 2019. The study was designated for two years and two locations but we cannot include here 2nd-year data due to the pandemic situation of COVID-19. Only one year of the 200 Chinese accessions (9G1–G200) in both locations results were pooled. Each accession was planted in a single row with 15–20 plants. The plant-to-plant distance is 5–10 cm and row to row distance is 40 cm.

Five plants from each accession were randomly selected and surveyed for recording the observations for 11phenotypic traits for days to 1st flowering, days to 50% flowering (50% DF), flower color (FC) scoring 1 to 11 [29]; where, flower with white colour took 1 days, blue took 3, pink took 7, violet took 10 days to maturity (DM). Plant height (PH), Primary branches (PB), Pods per plant (PP), Seeds per pod (SP), Yield per plant (YP), 100 seed weight (100 SW).

In the case of seed coat color (SC) scoring 1 to 10 [29]; where, grey is scored in 3, grey mottled is scorded in 9, green mottled scorded in 10 for an average of five plants in both locations.

From both locations, we collected the leaves from the middle part of the stems while the plants were flowering and 2–3 gm of fresh leaf samples from 5 plants per accession dried out for 20 min at 110°C packed with aluminum foil and dry at room temperature before loading at −20˚C in falcon tubes for measuring six biochemical traits such as β-ODAP fresh ($\mu g g^{-1}$), α-ODAP fresh ($\mu g g^{-1}$), Total-ODAP fresh ($\mu g g^{-1}$), β-ODAP dry ($\mu g g^{-1}$), α-ODAP dry ($\mu g g^{-1}$) and Total-ODAP dry ($\mu g g^{-1}$) through the HPLC method. We used $NaHCO_3$, FDNB (2, 4-dinitrofluorobenzene), BPS (phosphate buffer solution), and 100 mM NaAc-HAc (Sodium Acetate Buffer solution) for the preparation of ODAP standard. After preparation of ODAP standard stock solution of (not ginseng element) 5 mg/mL, prepared by $NaHCO_3$ 100 µL and 10 mg/mL FDNB solution, carry out the derivatization reaction in a constant temperature water bath at 60°C, take out (In a dry bath at 60°C for 30 min), cool in a room temperature. We used a 100 µL accurately pipette and gradually diluted with 0.5 mol/L $NaHCO_3$ solution to 317.5, 158.75, 79.38, 39.69, 19.84, 9.92, 4.96 µg/mL. The series of standard solutions were added 800 µL of phosphate buffer solution filtered with a 0.45 µm organic phase and measured. Add 4.0 mL of $NaHCO_3$ solution to 0.1g of dried grass pea leaf powder, incubate at 4°C for 5 min on a shaker, and centrifuge at 2000 r/min at 4°C for 10 min. Supernatant carefully handled 1 mL pipette of the into a 2 mL Eppendorf in a centrifuge tube. Centrifuge at 12,000 r/min for 10 min at room temperature. Take 0.1 mL of the supernatant in a 2 mL centrifuge tube to prepare the derivatization reaction. Add 100 µL FDNB to 100 µL sample, and bath in 60°C water bath for 30 min. Then add 0.8 mL of phosphate buffer. Dilute to 1 mL for later use. Then centrifuge at 12,000 rpm (Revolutions per minute) for 5 min, draw 700 µL of the supernatant, filter, and load the sample (Note: In the calculation, after calculating the corresponding β-ODAP concentration, multiply the reaction volume by 1 mL to calculate the total concentration). At the same time, derivatize the ODAP standard product to make the standard curve. Equilibrium of the chromatographic column respectively put acetonitrile, filtered deionized water, and 17% acetonitrile (use 0.1M, pH4.4 NaAc-) in the three storage bottles (A, B, C) of the chromatograph (HAc preparation). The column/needle was washed with 90% acetonitrile for 10 min, and then gradually adjusted the ratio of acetonitrile to water (online mixing of the instrument), so that the ratio of acetonitrile reached 20% (at least 20 min to complete), and finally equilibrated with 17% acetonitrile for 30 min until the detection signal baseline was level. Straight, the column pressure is constant.

We used well-distributed SSRs; the legume genome database will find primers sequence (5′-3′), PCR product size (bp), and repeat motif of the SSRs markers (Table 2). Plants were grown in a net house and leaf samples were collected from 5 random young seedlings (20–30 days) and mixed. Then genomic

DNA was extracted from the mixed leaf of 5 random young seedlings of each accession exploiting the Cetyl Trimethylammonium Bromide (CTAB) method [30,31]. Polymerase Chain Reactions (PCR) were conducted at a volume of 20 µL reaction comprising 10 µL 2x TagPCR Master Mix (Hooseen, Beijing, China), 2 µL primer, 3 µL of genomic DNA (30 ng) and dd H_2O 5.0 µL. On the K960 Thermal Cycler (Jingle, Hangzhou, China), microsatellite loci were enlarged with the following cycle: 5 min of initial denaturation at 95°C; 35 cycles of (denaturation for 30 s at 95°C, 30 s at annealing temperature, 45 s prolongation at 72°C) and final elongation at 72°C for 10 min. Gel electrophoresis was conducted using 8% non-denaturing polyacrylamide gel with 280 Volts and 50 watts and visualized through 0.1% silver nitrate staining (Fig. S1). Gel documentation was done using a BIO-RAD Gel Doc XR + machine. SSR results were scored according to the band size using AlphaEaseFC4.0 (Alpha Inotech, San Leandro, CA, USA).

Table 2: Properties of 56 polymorphic SSR markers used in this study (FP = Forward primer, RP = Reverse primer, Ta = Annealing temperature)

Markers	Repeat motif	Primer sequence (5′-3′)	Major allele size (bp)	Ta/°C
c34700_g1_i3	(T)10	FP-ACCAAAGGATGCAGGGTCTA RP-TAGTCGTGGTGTCGTGGTGT	299	54
c34887_g1_i1	(T)11aatttac(T)10	FP-TGGAGGACGAGCAACAATAA RP-TGTTGTTGATGGAAACAAATGA	132	54
c36504_g1_i3	(T)11	FP-CACACACCATTACGCACACA RP-TGGTGTCGTGGTCGTAGGTA	252	54
c34633_g1_i1	(T)10	FP-ATCGTAAACCGTGAGGGTCA RP-AAGCTTGTGGTGGCTACTGC	352	54
c31994_g1_i1	(T)12	FP-CACAACCAACGCCAATACAG RP-CCGTAGTACCGCGCTTATTC	165	54
c45717_g1_i1	(A)14	FP-TTTGTGTCACAGCCCTGTTT RP-CATGTTGGCTGCAAGTTTGT	195	52
c47533_g1_i6	(T)10	FP-GCAACAACAAATGCAACATC RP-TGTTGTTACTGCTGCTGCTCT	110	52
c42976_g1_i1	(T)11	FP-GACCTCGAGGGACATTAGCA RP-CAAAGAAAGAGAAAGGACACAA	325	52
c35999_g1_i1	(T)11	FP-TGTCTGGTGTGTGTGGTGTG RP-CGACACGTACGCAACGAC	196	52
c41936_g3_i2	(T)10	FP-CACCACCATAACCACCTCCT RP-ATGCGATTGAAGGGATGAAC	364	52
c39067_g2_i1	(T)10	FP-TTCAGATGCAGGTGGTTCAG RP-AACGGTGCGACTCTTGCTAT	291	52
c47694_g1_i2	(TG)6	FP-CACACCCTCAGGTCCTCAAT RP-ATGGCACAAAATTTCCCAAA	159	52
c39234_g1_i1	(T)12	FP-CCACTTCCACCTTTGACCAC RP-GGAGATCTGATGCAACCCTT	193	52

(Continued)

Table 2 (continued)

Markers	Repeat motif	Primer sequence (5′-3′)	Major allele size (bp)	Ta/°C
c46949_g2_i1	(AG)7	FP-TGATTTGCATTGGTTGCACT RP-GCTCCGTATGTTAAGTCTTTCAA	125	52
c47146_g1_i2	(A)14	FP-CGAGAAACAGCCTTTACCGT RP-GGTTTTTCGAATCCCCAAAT	221	52
c46049_g2_i1	(TG)6	FP-CCAAGGAAGCAAGGCTTTTT RP-TTACAATGGTCAGGCAAGCA	102	52
c34957_g2_i1	(CT)8	FP-GGCTTCCAAGAACAAAGCTG RP-TTACACCAACACATTTCAATGAC	200	52
c34100_g1_i1	(CT)8	FP-TGGTGTGGACAAGCTTTTTG RP-GAGCCTTGATCCCAATGAAC	172	52
c37441_g1_i2	(TGG)5	FP-TGGTCAAACTTTCAATGGCA RP-TAAAAACATAGCTGCGGGCT	224	52
c75340_g1_i1	(AG)8	FP-GCGGTGATGGTTGTCTTTTT RP-CACGGTATTCCACAAATATGC	101	52
c47441_g1_i7	(AT)6	FP-CACCAAAAACCTCTCAAACCA RP-TGAGTGAGAGTGAAATGCGG	223	52
c38070_g1_i2	(ATA)5	FP-CTGGCACCATAGGGTCAGTT RP-CGCGCATACATACAAAGCAG	195	52
c43047_g1_i1	(CT)7	FP-ATTTTGTTGTCAAATTGTCTTGTTA RP-CTAATCACAGATGCGCTCCA	233	52
c38894_g1_i1	(GAT)6	FP-CCAAAGTCCCTTTGCATTGT RP-GCCTTCTAAAGCCTTTGCCT	209	52
c31592_g1_i2	(CTA)7	FP-GTGGATTTGCTTTGGGATGT RP-TTCTTGACCCATCACGTTTG	160	52
c38694_g1_i1	(CT)6	FP-GCAGCAACAAGAATCCCAAT RP-TCACAGCCAGAACAAATCAGA	227	52
c47638_g1_i3	(TGG)5	FP-TATTTTGCTCAAAGACGGGG RP-ACACAGGTCGTTCTCCACAA	207	52
c36717_g1_i1	(AAG)5	FP-TGTCTTTACCGGCCTCTGTT RP-CTACCCTACAAGCCTGCTGC	187	52
c33181_g1_i1	(GGA)5	FP-TTCTGAAGATTGTTGCTGCG RP-CGTTCGTCTGGAGTTCCACT	333	52
c37493_g1_i1	(T)11	FP-TCCCTGTATTCATTTTGTTTTCA RP -TTCCATTGATGATGAGGGGT	271	52
c44073_g1_i3	(TTC)6	FP-CCCTTCAAACTTCAAAACCAA RP-AGGAAGGAAAGTTGGTCGGT	245	52
c43025_g1_i3	(TGTT)5	FP-TCCGTAGCGAATCAAGTGTT RP-TTGGCGCATATGTTGGAGTA	264	52

(Continued)

Table 2 (continued)

Markers	Repeat motif	Primer sequence (5′-3′)	Major allele size (bp)	Ta/°C
c39130_g2_i1	(GAAGAG)10	FP-TAGAAACTTGCACGCACCAC RP-TAGAAACTTGCACGCACCAC	362	52
c43652_g1_i5	(T)10	FP-TTTCTTTTTCATTTTTCTCCTTAAA RP-TGCAATAATTTGGGGAAAGG	247	52
c41895_g1_i1	(T)10	FP-CGTCGGTGACTAGGGAGAAC RP-AGAGTTGCCGGAGAGTGAAA	338	52
c43144_g1_i1	(GTT)5	FP-TGTGCCCATTCAACAAACAT RP-CGAGAAGAACGAGAAGTGGG	302	52
c35761_g1_i2	(ACT)5	FP-ACAGGTTTCCGAAGCATACG RP-CAAGTTCAAACTTCGACGCA	390	52
c43223_g1_i7	(CAT)5	FP-GGGTTTGAGGAGTTTGGACA RP-TCCTCTTCATCTTGCGGTCT	316	52
c45378_g1_i1	(ACACA)5	FP-GAGAAAAATAACCACCGCCA RP-CACACAGCAACACGTCCTCT	262	52
c37339_g1_i3	(T)12	FP-TTCGTGTGCAAAACGTTCAT RP-GATTTCCTGATTGCTCCCAA	150–170	52
c45586_g1_i2	(CT)8cacaccaac tcaaaacacaaca cctaaaattttcc agcaaaataag(T)12	FP-CCACCAAATTTCCCTTTTTG RP-GTACGAGAGGTTGACTTTTGTTT	239	52
G213	(GT)9c (GT)7	FP-TGTTGTTGGGAATTTCGTGA RP-CCAAGGCGTGAGCTATCTTC	230–270	52
G15624	(AAC)11	FP-GGTGCAGTGCTTGAAGATGA RPP-TTAATGTCCGACGAAACGAA	130–150	52
G17922	(CCA)5	FP-ATGGCTGAGGAGCTTTTT RP-TCACTTCCGGAATTCTCACC	140–150	52
G18078	(TGT)8	FP-CGACAGTTGCGACCAGTCTA RP-GATTCGGGATTTTTGGGTT	200–210	52
G205	(GT)7gcgtgtgc ctgcgtctctgcgag tgcgtgc(GT)6	FP-CACCACATCCACACACACCT RP-CCAGAGTTGTGAAAGTGCGA	157–165	52
G15709	(CAT)5	FP-CACCAAAAGAGAAGGACAAGG RP-GGTTGATTAGCCTTAGGGGG	180–220	52
G19207	(AAG)5	FP-GTTTTGGGGTTTCCCATTTT RP-CACATCCAAACCTTTCAGCA	230–250	52
G6	(AAC)12	FP-TTTGACGATGAATGGGATGA RP-AATTTGCGCGGTTAAACAAC	170–180	52
G33	(AC)6	FP-ATTTTTCAACGGATTGCAGG RP-TCGCAAGTGCACAACACATA	130–150	52

(Continued)

Table 2 (continued)

Markers	Repeat motif	Primer sequence (5′-3′)	Major allele size (bp)	Ta/°C
G26	(AC)16	FP-ATGTAGGCGTTACTGGACGC RP-AATCTCCGATTTGAAACCCC	230–250	52
G61	(AC)8	FP-CCTGGTATGGCTATTGAGGC RP-CCCGATTTTGATGTTTTACACC	170–200	52
G76	(ACA)6	FP-AAACTACCAAAAACGTTCCACA RP-TGGAGACGATGATGAATGGA	230–250	52
DY396423	(GT)8	FP-TTGTGGGGCTTGTTACACTG RP-CAACAACAGCATAAATACCCCTTT	160–180	53
MtBA32F05	(AG)5	FP-TCACAACTATGCAACAAAAGTG RP-G TGGGTCGGTGAATTTTCTGT	239	56
Ls989	(GT)8	FP-GGGCTTGTTACACTGATATGT RP-AACAGCATAAATACCCCTTTT	138–152 (7)	55

2.3 Statistical Analysis

To determine the genetic distance using the Power marker software 3.25 [32]. To estimate population genetic structure using the STRUCTURE V2.3.4 [33] based on an admixture model, where the K value was set to 1 to 10. The run was repeated 4 times for each K, K = 2 by way of best value based on LnP (D) with the method distinct by Evanno et al. [34]. The method regulates the suitable K value and estimates the Q parameter. The kinship coefficient (kinship) is generated by TASSEL V4.3 software. Path analysis was conducted using R studio 1.4.1717. Marker-trait associations were designed through general linear model (GLM) [35] and mixed linear model (MLM) association assessment including Q (structure likeness) + K (kinship) conditions into the TASSEL V4.3 software package. The p values of the marker linked with the QTL were controlled by multiple analysis rectification through regulation of the false detection rate. The number of permutation runs in GLM was set to 10000. Significant MTAs were declared up to $\alpha = 5\%$ [36]. The data density and normality, data variability parameters as well as correlation analysis of 11 phenotypic traits and 6 biochemical traits use the STAR V2.01 software (Fig. S2).

3 Results

3.1 Genotypic Data-Based Population Genetic Structure Analysis of Grass Pea

198 SSR markers were randomly screened to validate polymorphism first 28% of them were polymorphic which means 56 polymorphic markers were used to calculate the Polymorphism Information Content (PIC) of the markers and genetic diversity of the 400 grass pea germplasm but the location of markers in the chromosome is unknown (Table S2). Evanno's ΔK and LnP (D) explain two genetically different populaces (i.e., K = 2) based on delta K standards. The population structure showed that the total population was separated into two main groups and one admixed collection (Fig. 1). The Red color Pop1 group exposed 48.25% (193 accessions) which come from Bangladesh. The Green color Pop2 group displayed 47.75% (191 accessions) coming from diverse environmental areas (China, Bangladesh, Afghanistan, Armenia, Georgia, Azerbaijan, Syria, Spain, Germany, Ukraine, Russia, Italy, Czechoslovakia, Netherland, Slovakia, France, Algeria, Tunisia, Tanzania) which delivered from Chinese Academy of Agricultural Sciences (CAAS), China and also revealed their genetic correlation is very adjacent with Chinese accessions. Moreover, Pop1 and Pop2 are distinct from each other, Admixture

containing 4.00% (16 accessions) between Pop1 and Pop2, among them, nine genotypes from Pop1 and seven accessions from Pop2. PCA and PCoA also supported the population structure analysis (Figs. 1 and S3).

Figure 1: Population structure of 400 grass pea germplasm collected across the world (K = 2) where Pop1 (Red) belongs to 48.25%, Pop2 (Green) belongs to 47.75% and Admixed group belongs to 4.00%

3.2 Phenotypic Variation of Yield-Related Agronomic and Biochemical Traits

We observed large variations in studied traits among the grass pea accessions in our study. A total of 400 accessions were divided into two parts; Chinese accessions G1–G200 and Bangladeshi accessions G201–400 which showed highly significant ($p < 0.0001$) variances occurred with 11 yields related agronomic traits and 6 biochemical traits separately grown at Hebei and Liaoyang province in 2019. The phenotypic data (Table 3) of 17 traits (11 phenotypic and 6 biochemical traits) in two environments were studied for each accession and were employed for association study.

Table 3: Mean values, ranges, and coefficient of variation for 17 agronomic traits (11 phenotypic traits and 6 biochemical traits) measured among 400 grass pea genotypes in different geographic diversity panels (DGDP) grown at the Hebei and Liaoyang station, CAAS, China in 2019

DGDP {n = 400 accessions (G1–G400) in Hebei and only 200 accessions (G1–G200) in Liaoyang Province}						
Traits	Locations	Mean ± SE	Min	Max	STD	CV
Days to 1st flowering	Hebei (G1–G200)	104.1 ± 0.178	100	108	2.508	2.409
	Hebei (G201–G400)	104.14 ± 0.163	100	108	2.301	2.210
	Liaoyang (G1–G200)	62 ± 0.120	57	65	1.701	2.744
Days to 50% flowering	Hebei (G1–G200)	110.47 ± 0.179	105	114	2.520	2.281
	Hebei (G201–G400)	109.93 ± 0.180	105	116	2.549	2.319
	Liaoyang (G1–G200)	63.298 ± 0.115	58	67	1.624	2.565
Days to maturity	Hebei (G1–G200)	153.10 ± 0.237	150	184	3.347	2.186
	Hebei (G201–G400)	169.47 ± 0.332	165	178	4.695	2.770
	Liaoyang (G1–G200)	103.212 ± 0.144	98	108	2.024	1.961
Plant height (cm)	Hebei (G1–G200)	34.30 ± 0.372	18	57	5.238	15.270
	Hebei (G201–G400)	85.115 ± 1.220	43	133	17.253	20.270
	Liaoyang (G1–G200)	27.566 ± 0.669	10	70	9.410	34.137
Primary branches	Hebei (G1–G200)	4.419 ± 0.055	3	7	0.781	17.676
	Hebei (G201–G400)	7.029 ± 0.088	4	12	1.243	17.684
	Liaoyang (G1–G200)	1.949 ± 0.082	0	6	1.148	58.872

(Continued)

Table 3 (continued)

DGDP {n = 400 accessions (G1–G400) in Hebei and only 200 accessions (G1–G200) in Liaoyang Province}						
Traits	Locations	Mean ± SE	Min	Max	STD	CV
Flower color	Hebei (G1–G200)	3.010 ± 0.022	1	7	0.318	10.579
	Hebei (G201–G400)	4.215 ± 0.195	1	7	2.764	65.564
	Liaoyang (G1–G200)	9.753 ± 0.092	3	10	1.296	13.289
Pods per plant	Hebei (G1–G200)	32.338 ± 0.707	10	63	9.960	30.800
	Hebei (G201–G400)	49.95 ± 1.078	17	113	15.251	30.533
	Liaoyang (G1–G200)	9.091 ± 0.343	1	25	4.827	53.099
Seed per pod	Hebei (G1–G200)	3.236 ± 0.053	0.807	8.929	0.755	23.333
	Hebei (G201–G400)	2.189 ± 0.036	1.116	5.31	0.512	23.392
	Liaoyang (G1–G200	3.589 ± 0.068	1	9	0.960	26.749
Yield per plant (gm)	Hebei (G1–G200)	0.674 ± 0.007	0.460	1.057	0.098	14.662
	Hebei (G201–G400)	1.585 ± 0.037	0.614	2.889	0.524	33.037
	Liaoyang (G1–G200	1.774 ± 0.068	0.29	5.15	0.950	53.560
100 Seed weight (gm)	Hebei (G1–G200)	3.370 ± 0.035	2.302	5.285	0.494	14.662
	Hebei (G201–G400)	7.926 ± 0.185	3.072	14.445	2.619	33.037
	Liaoyang (G1–G200)	3.910 ± 0.037	2	6	0.526	13.452
Seed coat color	Hebei (G1–G200)	9.373 ± 0.034	9	10	0.485	5.174
	Hebei (G201–G400)	9.345 ± 0.034	9	10	0.477	5.100
	Liaoyang (G1–G200)	3.101 ± 0.053	3	10	0.740	8.135
β-ODAP fresh ($\mu g g^{-1}$)	Hebei (G1–G200)	4500.017 ± 82.00	770.068	8161.252	1150.96	25.576
	Hebei (G201–G400)	1630.124 ± 35.324	621.936	3553.428	499.562	30.464
α-ODAP fresh ($\mu g g^{-1}$)	Hebei (G1–G200)	1770.565 ± 38.07	177.180	3298.411	534.492	30.184
	Hebei (G201–G400)	704.123 ± 19.203	164.414	2662.817	271.574	38.569
Total-ODAP fresh ($\mu g g^{-1}$)	Hebei (G1–G200)	6270.582 ± 117.41	1036.892	10,885.75	1647.971	26.280
	Hebei (G201–G400)	2334.247 ± 51.877	804.564	4990.49	733.656	31.430
β-ODAP dry ($\mu g g^{-1}$)	Hebei (G1–G200)	17338.05 ± 312.76	2526.104	29,362.95	4389.899	25.319
	Hebei (G201–G400)	9405.455 ± 196.241	3648.347	19,697.59	2775.263	29.507
α-ODAP dry ($\mu g g^{-1}$)	Hebei (G1–G200)	6807.919 ± 141.82	791.662	12,922.06	1990.583	29.239
	Hebei (G201–G400)	4076.185 ± 107.892	937.030	13,234.99	1525,828	37.433
Total-ODAP dry ($\mu g g^{-1}$)	Hebei (G1–G200)	24145.97 ± 444.20	3690.962	41,051.58	6234.764	25.821
	Hebei (G201–G400)	13481.64 ± 292.52	4585.378	27,935.32	4122.762	30.581

Note: Flower color range 1 to 11, here 1 = white, 3 = blue, 7 = pink, 10 = violet; Seed coat color range 1 to 10, here 3 = grey, 9 = grey mottled, 10 = green mottled; STD means standard deviation; CV means Coefficient of Variation.

This study observed that the grass pea cultivars from different environmental areas have a vast variation in each experiment site. Among the 10 agronomic traits the higher coefficient of variation in 1st DF (2.744%), 50% DF (2.565%), PH (34.137%), PB (58.872%), PP (53.099%), SP (26.749%) and YP (53.560%) was discovered in Liaoyang location than Hebei location among the Chinese accessions (G1–G200). DM (2.770%), FC (65.564%), and 100 SW (33.037%) are the higher coefficient of variation in Bangladeshi accessions (G201–G400) than Chinese accessions in the Hebei location. The phenotypic coefficient of variance was highest in FC (65.564%) and lowest in 50% DF (2.565%) among the agronomic traits. In a qualitative trait, SC (8.135%) was the higher coefficient of variation observed in Liaoyang than Hebei location (both accessions). Among the biochemical traits the higher coefficient of variation was observed in β ODAP fresh (30.464%), β ODAP dry (29.507%), α ODAP fresh (38.569%), α ODAP dry (37.433%), T ODAP fresh (31.430%) and T ODAP dry 30.581%) of Bangladeshi accessions than Chinese accessions at Hebei location. The highest coefficient of variation was observed in α ODAP fresh (38.569%), and the lowest in β ODAP dry (29.507%). It is important to further study seventeen traits among multiple environments over several years.

3.3 Correlation and Path Coefficient

Correlation analysis specified that there are significant relations among the considered traits (Table 4). In the Hebei location for genotypes G1–G200, strong positive correlations were found between YPP and 100 SW; among the six biochemical traits (correlation range 0.71 to 0.98) and between 50% DF and 1st DF. SC showed a strong negative correlation with YPP, 100 SW, and SC and FC showed a strong negative correlation with YPP, 100 SW, and SC. SC showed a moderate positive correlation with 100 SW and YPP, between DM and PH, and finally 50% DF and 1st DF showed a moderate positive correlation with DM, PH, YPP, and 100 SW, respectively. A moderate negative correlation was found between 50% DF, 1st DF, and SP and FC at $p \leq 0.05$ (Fig. 2a).

Table 4: Path coefficients and correlation where β-ODAP dry and β-ODAP fresh are dependent variables and 1st DF, 50% DF, PH, PB, FC, PP, SP, YPP, SC are predictor variables

Predictor variables	Correlation with β-ODAP dry	Direct effect (D)	Indirect effect (I)									Total indirect effect (I)	Total effect (D+I)
			1st DF	50% DF	PH	PB	FC	PP	SP	YPP	SC		
1st DF	0.059	0.016		0.123	−0.039	−0.010	0.033	0.004	−0.007	−0.060	−0.002	0.043	0.059
50% DF	0.119	0.130	0.015		−0.018	−0.003	0.042	0.003	−0.003	−0.046	−0.002	−0.011	0.119
PH	−0.641	−0.277	0.002	0.008		−0.125	−0.014	0.025	−0.039	−0.222	0.000	−0.364	−0.641
PB	−0.589	−0.164	0.001	0.002	−0.211		−0.009	0.029	−0.033	−0.203	−0.001	−0.425	−0.589
FC	−0.190	−0.124	−0.004	−0.044	−0.031	−0.012		0.002	−0.003	0.022	0.006	−0.066	−0.190
PP	−0.408	0.043	0.002	0.009	−0.162	−0.110	−0.005		−0.026	−0.158	0.000	−0.451	−0.408
SP	0.492	0.062	−0.002	−0.006	0.174	0.086	0.006	−0.018		0.188	0.002	0.430	0.492
YPP	−0.604	−0.287	0.003	0.021	−0.214	−0.116	0.009	0.024	−0.041		−0.003	−0.317	−0.604
SC	−0.023	−0.019	0.002	0.012	0.000	−0.005	0.042	0.001	−0.005	−0.050		−0.004	−0.023
Predictor variables	Correlation with β-ODAP fresh	Direct effect (D)	Indirect effect (I)									Total indirect effect (I)	Total effect (D+I)
			1st DF	50% DF	PH	PB	FC	PP	SP	YPP	SC		
1st DF	0.003	−0.125		0.219	−0.059	−0.009	0.032	0.006	−0.003	−0.060	0.003	0.128	0.003
50% DF	0.083	0.231	−0.118		−0.027	−0.003	0.041	0.004	−0.001	−0.047	0.003	−0.148	0.083
PH	−0.770	−0.426	−0.017	0.015		−0.119	−0.014	0.033	−0.019	−0.223	0.000	−0.344	−0.770

(Continued)

Table 4 (continued)

Predictor variables	Correlation with β-ODAP fresh	Direct effect (D)	Indirect effect (I)									Total indirect effect (I)	Total effect (D+I)
			1st DF	50% DF	PH	PB	FC	PP	SP	YPP	SC		
PB	−0.675	−0.156	−0.007	0.004	−0.325		−0.009	0.037	−0.016	−0.204	0.001	−0.519	−0.675
FC	−0.213	−0.121	0.034	−0.079	−0.048	−0.012		0.002	−0.002	0.022	−0.010	−0.092	−0.213
PP	−0.470	0.056	−0.013	0.017	−0.249	−0.104	−0.005		−0.013	−0.159	0.001	−0.526	−0.470
SP	0.552	0.030	0.014	−0.011	0.267	0.082	0.006	−0.024		0.189	−0.003	0.522	0.552
YPP	−0.692	−0.288	−0.026	0.037	−0.330	−0.110	0.009	0.031	−0.020		0.005	−0.404	−0.692
SC	0.021	0.029	−0.013	0.022	−0.001	−0.005	0.041	0.001	−0.003	−0.050		−0.008	0.021

Note: Residual effect (R^2) = 0.499 in β-ODAP dry and Residual effect in β-ODAP fresh (R^2) = 0.3313, whereas, 1st DF = Days to 1st flowering, 50% DF = 50% days to flowering, PH = Plant height, PB = Primary branches, FC = Flower color, PP = Pods per plant, SP = Seeds per pod, YPP = Yield per plant, SC = Seed color. NB: Days to maturity (DM) and 100 seed weight (SW) traits were discarded from this analysis due to multicollinearity problem.

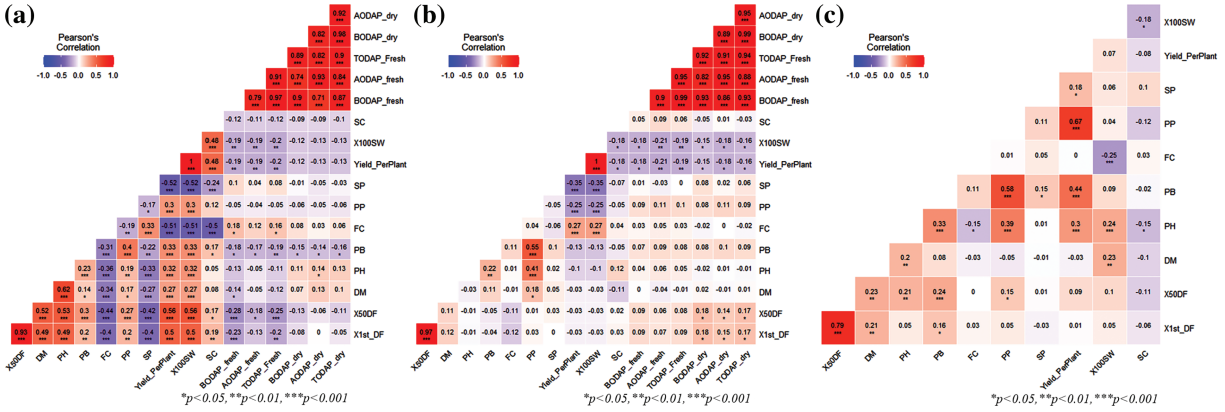

Figure 2: The correlation of 17 traits (10 agronomic traits, 1 qualitative trait, and 6 biochemical traits) in two environments among (a) Hebei (G201–G400), (b) Hebei (G1–G200), (c) Liaoyang (G1–G200)

In the Hebei location for genotypes G200–G401, strong positive correlations were found between YPP and 100 SW; between 50% DF and 1st DF, and among the six biochemical traits (correlation range 0.82 to 0.99). There is no strong negative correlation. A moderate positive correlation was found between PB and PP; and between PH and PP at a 5% significance level (Fig. 2b). In the Liaoyang location for genotypes G1–G200, there was no strong positive correlation. A moderate positive correlation was found between 50% DF and 1st DF; and between PB and PP. Most of the agronomic traits showed a low positive correlation to each other but only SC showed a negative correlation with 1st DF, 50% DF, DM, PH, PB, PP, YPP, and 100 SW at $p \leq 0.05$ (Fig. 2c).

In the case of β-ODAP dry, Table 4 shows that days to 1st flowering (1st DF), 50% days to flowering (50% DF), PP and SP exerted a positive direct effect on β-ODAP dry (0.016, 0.130, 0.043 and 0.062, respectively), whereas PH, PB, FC, YPP and SC had a negative direct effect on β-ODAP dry (−0.277, −0.164, −0.124, −0.287 and −0.019, respectively) but the total effect of 1st DF, 50% DF and SP were positive. The highest positive indirect effects on β-ODAP dry were observed with SP (0.188, 0.174) through YPP and PH, respectively. The highest negative indirect effects on β-ODAP dry were observed with PH (−0.222) through YPP and YPP (−0.214) through PH. The residual effect explains how best the independent variables account for the variability of the dependent variable (β-ODAP dry) and its value is 0.499.

In the case of β-ODAP fresh, Table 4 shows that 50% DF, PP, SP, and SC exerted a positive direct effect on β-ODAP fresh (0.231, 0.056, 0.030 and 0.029, respectively), whereas 1st DF, PG, PB, FC, YPP, and SC had a negative direct effect on β-ODAP fresh (−0.125, −0.426, −0.156, −0.121 and −0.288, respectively) but the total effect was positive on 1st DF (0.003), 50% DF (0.083), SP (0.552) and SC (0.021) and negative on PH (−0.777), PB (−0.675), FC (−0.213), PP (−0.470) and YPP (−0.692). The highest positive indirect effects on β-ODAP fresh were observed with SP (0.267, 0.189) through PH and YPP, respectively. The highest negative indirect effects on β-ODAP fresh were observed with PH (−0.223) through YPP and 50% DF (−0.118) through 1st DF. For the dependent variable, the β-ODAP fresh residual value is 0.3313.

3.4 Linkage Disequilibrium Analysis

The distribution diagrams of the genome-wide linkage disequilibrium (LD) attenuation were drawn in TASSEL 4.3 based on R^2 and D′ from the results of the LD and the genetic distance. The basis and premise of association analysis is the between gene linkage disequilibrium of the diversity of object traits and gene (locus) polymorphism to find marker loci with the purposes of specific genes strictly associated with phenotypic differences. In total, 1526 arrangements of 56 pairs of SSR primers were found, of which the arrangements with $R^2 \geq 0.1$ were considered for 93.18%. The probability of $p \leq 0.01$ was sustained and there remained were unbalanced combinations in the pair (Fig. 3), where the R^2 value ranged from 0.020–0.696 and D′ ranged from 0.143–0.862. The 56 pairs of SSR primers nominated in this study have linkage imbalance in 400 grass pea germplasm from the consequence of analysis, which can be associated with considerable traits.

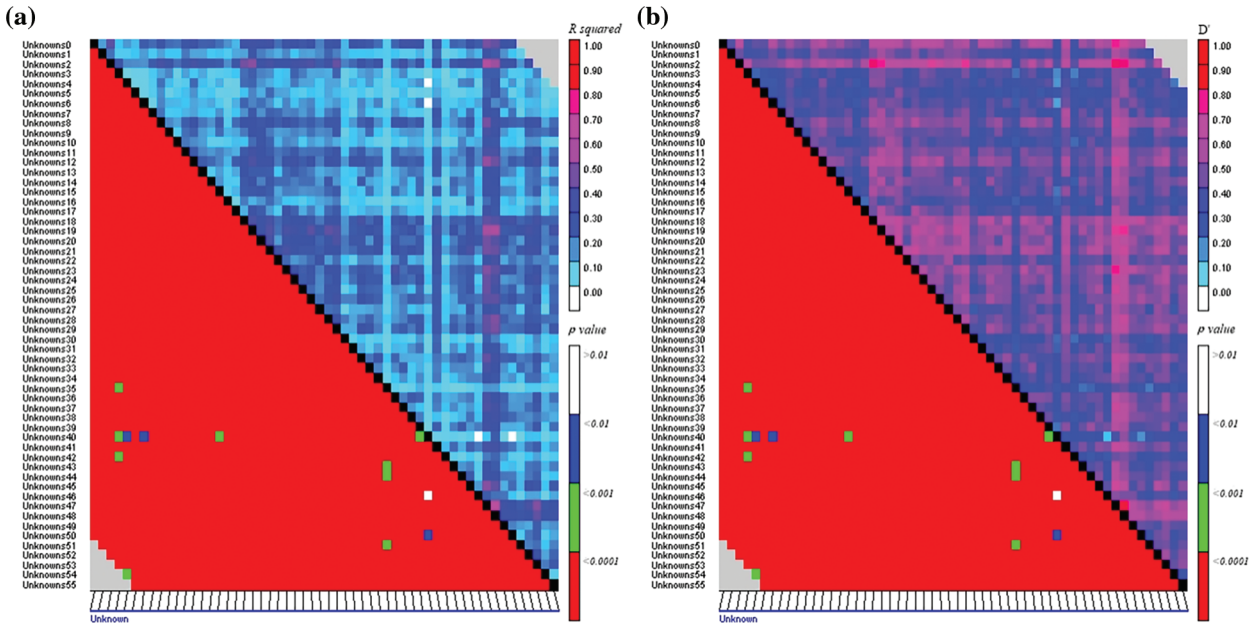

Figure 3: Linkage disequilibrium (LD) between SSR markers showing with (a) R^2 and *p* value (b) D′ and *p* value

3.5 Association Studies of SSR Marker with Yield-Related Agronomic and Biochemical Traits

The interrelated Q value and Q+K value of grass pea resources were used as covariates in the GLM_Q and MLM_Q+K models correspondingly, and the association studies of the 56 SSR molecular markers with yield-related agronomic traits, qualitative traits, and biochemical traits are executed to detect the associated

markers and set on their interpretation rate. Of the marker-trait associations identified for individual trials, 99 MTAs (Marker trait associations) as listed in the GLM model and 61 MTAs as listed in the MLM model (Table 5) were selected based on their *p* value and occurrence in multiple trials. The sequences of 56 polymorphic SSR markers listed and diversity statistics were provided in Tables 2 and S2.

Table 5: Determination of marker-trait association using GLM and MLM model with phenotypic and biochemical data of two locations in 2019

Trait	Location		Hebei Province (G1–G200)			Hebei Province (G201–G400)				Liaoyang Province (G1–G200)			
	Model	Marker	Marker F	Marker P	Marker R^2	Marker	Marker F	Marker P	Marker R^2	Marker	Marker F	Marker P	Marker R^2
1st DF	GLM	C39067	3.041	0.018	0.058	C38694	3.109	0.028	0.046	G18078	4.019	0.008	0.058
1st DF	GLM	Ls989	3.122	0.046	0.031	C31592	2.399	0.039	0.059	C41895	2.812	0.027	0.055
1st DF	GLM									G15709	2.746	0.030	0.054
1st DF	MLM	C38694	2.746	0.044	0.046	C39067	3.138	0.016	0.072	G18078	3.510	0.016	0.053
1st DF	MLM									C41895	2.717	0.031	0.055
1st DF	MLM									G15709	2.606	0.037	0.053
50% DF	GLM	C39067	3.138	0.016	0.061					C43144	2.382	0.030	0.070
50% DF	GLM	C47533	3.062	0.029	0.045					C34887	3.393	0.036	0.034
50% DF	GLM	C39130	2.483	0.033	0.060					C43652	2.522	0.042	0.050
50% DF	GLM									G213	2.758	0.044	0.041
50% DF	MLM					C39067	3.274	0.013	0.074	C43144	2.319	0.035	0.071
α-ODAP dry	GLM	G213	4.447	0.005	0.064	G33	2.853	0.011	0.082				
α-ODAP dry	GLM					C39234	2.844	0.025	0.055				
α-ODAP dry	GLM					Ls989	3.742	0.025	0.037				
α-ODAP dry	MLM	G33	3.116	0.006	0.094	G213	3.220	0.024	0.056				
α-ODAP fresh	GLM	G213	4.282	0.006	0.062	C39234	3.246	0.013	0.063				
α-ODAP fresh	GLM					Ls989	4.295	0.015	0.042				
α-ODAP fresh	MLM	G33	2.332	0.034	0.071	G213	3.058	0.030	0.055				
α-ODAP fresh	MLM	Ls989	3.362	0.037	0.034								
α-ODAP fresh	MLM	C39234	2.546	0.041	0.051								
β-ODAP dry	GLM	G213	3.326	0.021	0.048	G33	2.924	0.009	0.084				
β-ODAP dry	GLM					C47638	2.957	0.021	0.057				
β-ODAP dry	GLM					C31592	2.578	0.028	0.063				
β-ODAP dry	GLM					C42976	2.669	0.034	0.052				
β-ODAP fresh	GLM	G213	3.424	0.018	0.050	C47638	3.524	0.008	0.068				
β-ODAP fresh	GLM					C31592	2.747	0.020	0.066				

(Continued)

Table 5 (continued)

Trait	Location		Hebei Province (G1–G200)			Hebei Province (G201–G400)				Liaoyang Province (G1–G200)			
	Model	Marker	Marker F	Marker P	Marker R^2	Marker	Marker F	Marker P	Marker R^2	Marker	Marker F	Marker P	Marker R^2
β-ODAP fresh	GLM					C34100	2.769	0.029	0.054				
β-ODAP fresh	GLM					C43144	2.248	0.040	0.065				
β-ODAP fresh	MLM	C47638	3.637	0.007	0.073								
β-ODAP fresh	MLM	C31592	2.337	0.043	0.059								
100 SW	GLM									C46049	2.652	0.035	0.052
100 SW	GLM									DY396423	2.676	0.048	0.040
100 SW	MLM									DY396423	2.838	0.039	0.043
100 SW	MLM									C46049	2.494	0.044	0.051
DM	GLM	C41936	49.777	0.000*	0.436	C34100	7.807	0.000*	0.138	C34887	4.453	0.013	0.043
DM	GLM	C39067	37.000	0.000*	0.435					G76	2.580	0.039	0.050
DM	GLM	C42976	2.779	0.019	0.068					C47638	2.287	0.048	0.055
DM	GLM									G19207	3.090	0.048	0.030
DM	MLM	C34100	4.725	0.001	0.095	C41936	29.657	0.000	0.454	C34887	4.073	0.019	0.040
DM	MLM					C39067	22.075	0.000	0.450	G19207	3.816	0.024	0.038
DM	MLM					C42976	2.564	0.029	0.065	G76	2.589	0.038	0.051
FC	GLM	C47146	8.661	0.000*	0.120	C46049	3.737	0.012	0.053	C47638	6.340	0.000*	0.142
FC	GLM	C38070	4.038	0.002	0.075					G33	2.553	0.021	0.075
FC	GLM	C46049	4.072	0.003	0.061					Ls989	3.188	0.043	0.032
FC	GLM	C41936	4.431	0.005	0.051								
FC	GLM	C36504	3.541	0.008	0.054								
FC	GLM	G15624	2.855	0.016	0.055								
FC	GLM	C34887	3.978	0.020	0.031								
FC	GLM	DY396423	3.309	0.021	0.038								
FC	GLM	C41895	2.875	0.024	0.044								
FC	GLM	G205	2.636	0.025	0.051								
FC	GLM	C34100	2.844	0.025	0.044								
FC	MLM	C46049	3.633	0.014	0.054	C45586	5.125	0.001	0.102	C47638	5.580	0.000*	0.142
FC	MLM					C43652	3.346	0.011	0.067	G33	2.437	0.027	0.075
FC	MLM					C38070	2.994	0.013	0.075	Ls989	3.118	0.046	0.032
PB	GLM	C43144	2.730	0.014	0.078	MtBA32F05	2.621	0.036	0.048	C47694	4.741	0.003	0.068
PB	GLM	C43047	2.489	0.033	0.060	C36717	2.399	0.039	0.055	C37441	3.617	0.014	0.053
PB	GLM	C46949	2.406	0.038	0.058	C35999	2.495	0.044	0.046	G61	2.561	0.021	0.075
PB	GLM	MtBA32F05	2.321	0.045	0.056					C36504	2.727	0.031	0.054
PB	GLM	C31592	2.290	0.047	0.056								
PB	GLM	C37441	2.676	0.048	0.039								
PB	MLM					C43144	2.593	0.019	0.078	C47694	4.411	0.005	0.067
PB	MLM					C43047	2.398	0.039	0.060	G61	2.596	0.019	0.079
PB	MLM					C46949	2.323	0.045	0.058	C37441	3.375	0.019	0.052
PB	MLM									C36504	2.505	0.044	0.051
PH	GLM	C34633	2.861	0.038	0.042	C34100	9.909	0.000*	0.169	C41936	3.425	0.018	0.051
PH	GLM	G76	2.469	0.046	0.049					C45586	2.993	0.020	0.059
PH	MLM	C34100	5.400	0.000	0.109	C34633	2.858	0.038	0.044	C45586	3.161	0.015	0.064

(Continued)

Table 5 (continued)

Trait	Location		Hebei Province (G1–G200)			Hebei Province (G201–G400)				Liaoyang Province (G1–G200)			
	Model	Marker	Marker F	Marker P	Marker R^2	Marker	Marker F	Marker P	Marker R^2	Marker	Marker F	Marker P	Marker R^2
PH	MLM									C41936	2.985	0.032	0.046
PP	GLM	C42976	2.468	0.034	0.061	C39234	4.493	0.002	0.081	C46949	2.312	0.046	0.057
PP	GLM					C34700	3.039	0.019	0.057				
PP	GLM					C37339	2.972	0.033	0.042				
PP	GLM					C35761	2.328	0.044	0.055				
PP	MLM	C39234	4.174	0.003	0.084	C42976	2.440	0.036	0.062				
PP	MLM	C37339	3.693	0.013	0.056								
PP	MLM	C34700	2.835	0.026	0.057								
PP	MLM	G76	2.851	0.039	0.043								
SC	GLM	C46049	2.989	0.020	0.059					C34887	4.323	0.015	0.043
SC	GLM									C45378	2.837	0.026	0.056
SC	MLM					C46049	2.670	0.034	0.059	C34887	4.181	0.017	0.043
SC	MLM									C45378	2.732	0.030	0.056
SP	GLM	C47638	16.711	0.000*	0.300	C43047	2.980	0.020	0.057	G61	2.975	0.008	0.085
SP	GLM	C43047	4.891	0.000*	0.112	C46949	2.358	0.042	0.057	C46949	2.504	0.032	0.061
SP	GLM	C42976	2.923	0.014	0.070	C38894	2.727	0.045	0.040				
SP	GLM	C36717	2.903	0.015	0.070								
SP	GLM	C39234	2.858	0.025	0.055								
SP	GLM	Ls989	3.600	0.029	0.035								
SP	GLM	C43025	3.012	0.031	0.044								
SP	GLM	C41895	2.504	0.044	0.049								
SP	MLM					C47638	11.053	0.000	0.277	G61	2.603	0.019	0.079
SP	MLM					C43047	3.633	0.004	0.091	C46949	2.278	0.048	0.057
SP	MLM					C43025	3.012	0.031	0.045	C33181	2.662	0.049	0.040
SP	MLM					C36717	2.465	0.034	0.062				
SP	MLM					C39234	2.490	0.045	0.050				
T-ODAP dry	GLM	G213	3.843	0.011	0.055	G33	3.143	0.006	0.089				
T-ODAP dry	GLM					C47638	2.613	0.037	0.051				
T-ODAP dry	GLM					C43223	2.523	0.042	0.049				
T-ODAP dry	MLM	G33	2.969	0.009	0.090								
T-ODAP dry	MLM	C47638	2.525	0.042	0.051								
T-ODAP dry	MLM	C43223	2.449	0.048	0.049								
T-ODAP Fresh	GLM	G213	3.862	0.010	0.056	C47638	3.190	0.014	0.062				
T-ODAP Fresh	GLM					G33	2.210	0.044	0.065				
T-ODAP Fresh	GLM					C31592	2.311	0.046	0.056				

Note: 1st DF = Days to first flowering, 50% DF = Days to 50% flowering, DM = Days to maturity, PH = Plant height, PB = Primary branches, FC = Flower color, PP = Pod per plant, SP = Seeds per pod, SC= Seed coat color, YP = Yield per plant, 100 SW = 100 seed weight and T-ODAP = Total ODAP. *Passed against Bonferroni correction test [Threshold = −log (0.05/No. of markers)].

Based on the genotype data, phenotypic data, and the Q-matrix from population structure results, a general linear model and mixed linear model were used to analyze the marker-trait associations. Four hundred genotypes (G1–G200) and (G201–G400) were used to reveal the marker-trait association. Among the 56 SSR markers, 5, 5, and 23 markers were significantly associated at <0.1%, <1%, and <0.05% levels, respectively for 11 traits in the GLM model. Maximum 43.6% PVE was found for C41936. In the MLM model, 13 and 5 markers were significantly associated at <1% and <0.05% levels, respectively for 7 traits. A maximum of 45.37% PVE was found for the same marker C41936 (Table 5).

3.6 Yield-Related Agronomic Traits

We observed that the results of GLM_Q and MLM_Q+K study all yield-related agronomic traits are significantly associated in Hebei and Liaoyang locations among the genotypes G1–G200 and G201–G400 at $p \leq 0.05$ (Table 3). For days to 1st flowering, the five markers in the MLM analysis outcomes are the same as those identified in the GLM outcomes, and these five markers C39067, C38694, G18078, C41895, and G15709 were identified in both models having PVE values ranging from 4.6% to 7.2%. For days to 50% flowering the two markers in the MLM analysis outcomes are the same as those identified in the GLM outcomes, and these two markers C39067 and C43144 were identified in both models having PVE values ranging from 6.1% to 7.4%. For days to maturity, the seven markers in the MLM analysis outcomes are the same as those identified in the GLM outcomes, and these seven markers C34100, C41936, C39067, C42976, C34887, G19207, and G76 were identified in both models having PVE value ranged 3.8% to 9.5%. For plant height, the four markers in the MLM analysis outcomes are the same as those identified in the GLM outcomes, and these four markers C34100, C34633, C45586, and C41936 were identified in both models having PVE values ranging from 4.4% to 10.9%.

For primary branches, the seven markers in the MLM analysis outcomes are the same as those identified in the GLM outcomes, and these seven markers C43144, C43047, C46949, C47694, G61, C37441, and C36504 were identified in both models having PVE value ranged 5.1% to 7.9%. For flower color, the seven markers in the MLM analysis outcomes are the same as those identified in the GLM outcomes, and these seven markers C46049, C45586, C43652, C47638, G33, C38070, and Ls989 were identified in both models having PVE value ranged 0.1% to 14.2%. For pods plant^{-1} the four markers in the MLM analysis outcomes are the same as those identified in the GLM outcomes, and these four markers C39234, C37339, C34700, and C42976 were identified in both models having PVE values ranging from 4.2% to 8.4%. For seeds per pod, the six markers in the MLM analysis outcomes are the same as those identified in the GLM outcomes, and these six markers C47638, C43047, C43025, C36717, C39234, and C46949 were identified in both models having PVE values ranged 4.4% to 27.7%. For seed color, the three markers in the MLM analysis outcomes are the same as those identified in the GLM outcomes, and these three markers C46049, C34887, and C45378 were identified in both models having PVE values ranging from 4.3% to 5.9%. For yield per plant, no marker-trait association was found in this analysis against yield per plant. For 100 seed weight, the two markers in the MLM analysis outcomes are the same as those identified in the GLM outcomes, and these two markers DY396423 and C46049 were identified in both models having PVE values ranging from 4.0% to 5.2% (Table 5).

3.7 Biochemical Traits

Six biochemical traits showed MTAs in Hebei province for genotypes G1–G200 and genotypes G201–G400 but in Liaoyang province, data are not available here. We observed that the results of GLM_Q and MLM_Q+K studies all biochemical traits are significantly associated in Hebei location among the genotypes G1–G200 and G201–G400 at $p \leq 0.05$. For α-ODAP dry the two markers in the MLM analysis outcomes are the same as those identified in the GLM outcomes, and these two markers G33 and G213 were identified in both models having PVE values ranging from 5.6% to 9.4%. For β-ODAP dry

total of five marker traits were associated, β-ODAP dry was significantly associated with the G213 marker having PVE 4.8% among the genotypes G1–G200 and with three markers G33, C47638 and C43223 having PVE ranged from 4.9% to 8.9% among the genotypes G201–G400. There is no marker shown in the MLM and GLM outcomes. For total-ODAP dry the three markers in the MLM analysis outcomes are the same as those identified in the GLM outcomes, and these three markers G33, C47638, and C43223 were identified in both models having PVE values ranging from 5.1% to 9.0%. For α-ODAP fresh the three markers in the MLM analysis outcomes are the same as those identified in the GLM outcomes, and these three markers Ls989, C39234, and G213 were identified in both models having PVE values ranging from 3.4% to 5.5%. For β-ODAP fresh the two markers in the MLM analysis outcomes are the same as those identified in the GLM outcomes, and these two markers C47638 and C31592 were identified in both models having PVE values ranging from 5.9% to 7.3%. For total-ODAP fresh was significantly associated with the G213 marker having PVE 5.6% among the genotypes G1–G200 and with three markers G33, C47638 and C31592 having PVE ranged from 5.6% to 6.5% among the genotypes G201–G400. There is no marker shown in the MLM analysis outcomes are the same as those identified in the GLM outcomes.

Among the 56 polymorphic SSR markers, 21 markers were associated with more than one trait. G213 is associated with seven traits viz., 50% DF, α-ODAP dry, α-ODAP fresh, β-ODAP dry, β-ODAP fresh, T-ODAP dry, T-ODAP fresh and C47638 also associated with seven traits namely β-ODAP dry, β-ODAP fresh, DM, FC, SP, T-ODAP dry, T-ODAP fresh. Three markers associated with 5 traits such as Ls989 associated with 1st DF, α-ODAP dry, α-ODAP fresh, FC, SP whereas, C31592 was related with 1st DF, β-ODAP dry, β-ODAP fresh, PB, T-ODAP fresh and G33 showed association with α-ODAP dry, β-ODAP dry, FC, T-ODAP dry, T-ODAP Fresh. Four markers are associated with four traits such as C34887 (50% DF, DM, FC, SC), C39234 (α-ODAP dry, α-ODAP fresh, PP, SP), C42976 (β-ODAP dry, DM, PP, SP) and C34100 (β-ODAP fresh, DM, FC, PH). Twelve markers are linked with two to three traits.

4 Discussion

4.1 Population Structure Analysis

Assessment of population genomic structure is a precondition for genome-wide association studies because population structure is generally responsible for false associations [21]. A practically exact population structure may clue to greater genetic variances between groups, supplementary genetic relationships in every group as well as mostly decrease the deficiency in association analysis [37]. Therefore, the precision of the association study rests on whether the population structure was suitable [38]. Earlier, a reporter evaluated genotype clusters in cotton using Q-matrix [20]. Some reporters such as cotton [39,40], rice [41], and grass pea [42] evaluated the population structure by STRUCTURE software to indicate K values and population structure with K conforming to the highest structural level. The whole population was divided into two major populations and one admixed population, which was equitable to remove the false association things in the analysis of association studies. The diversity and kinship between two groups are related to geographical locations. The grass pea (*Lathyrus sativus*) found in Bangladesh formed a completely separate group indicating that human activities did not influence the diversity in Bangladesh but other locations' diversity was influenced by human activities. That means Bangladesh might be the center of origin of *Lathyrus sativus* which is supported by the previous finding of Smartt [43].

4.2 Phenotypic Variation of 17 Traits

As a dry areas crop grass pea accessions revealed the mean of the eleven yield-related phenotypic traits against the Chinese genotypes (G1–G200) and Bangladeshi genotypes (G201–G400) in Hebei location was higher than Chinese genotypes (G1–G200) in Liaoyang location but the Coefficient variation (CV%) of most of the yield-related phenotypic traits against the Chinese genotypes (G1–G200) in Liaoyang province was

higher than Hebei province among the same genotypes (G1–G200) but three traits such days to maturity, flower color and 100 seed weight were showed the higher coefficient of variation in Bangladeshi accessions (G201–G400) than Chinese accessions in Hebei location. Interestingly the qualitative trait seed color showed a higher coefficient of variation of Chinese genotypes (G1–G200) in Liaoyang than the Hebei location against genotypes (G1–G200) and genotypes (G201–G400). The former study showed that grass pea has a large difference among different phenotypic traits dependent on genotypes and ecosystem and also they have observed the same finding among the different phenotypic traits [5,44–46]. The mean of six biochemical traits of Chinese genotypes (G1–G200) showed higher than Bangladeshi genotypes (G201–G400) in the Hebei location but the Coefficient variation (CV%) of biochemical traits Bangladeshi genotypes (G201–G400) were higher than Chinese genotypes (G1–G200) in Hebei location. Previous studies reported that the effect of genotypes on protein or ODAP content is very important because it changes highly depending on environmental factors (soil and ecosystem) and they have observed the same finding among the biochemical traits [44].

4.3 Correlation and Path Coefficient

The correlation studies explain only the nature and amount of link of yield-related agronomic traits but it does not deliver the specifics of direct and indirect effects. Correlation studies (Fig. 2) specified that yield per plant showed a moderate to strongly positive correlation with traits such as 100 seed weight, seed color, pods per plant, plant height, days to 1st flowering, and days to 50% flowering. Some traits were found to be strongly negative with yield per plant that is seed per pod and flower color. These correlations of similar findings were reported by Singh et al. [47] and Ratna et al. [48] reported that yield per plant was negatively correlated with plant height. We also found that six biochemical traits showed a strong positive correlation among the genotypes G1–G200 and G201–G400 in the Hebei location and correlation range (0.71 to 0.98). Interestingly, we observed that ODAP contents were significantly low and a negative correlation showed with all agronomic traits (Fig. 2). This outcome can be hopeful to advance new varieties with low ODAP content and high-yielding varieties and convey a good message was caused through the toxin reduction and this idea also should be taken consideration with the future breeding prospect [49].

Path coefficient analysis (Table 4) is an arithmetical method to divide the correlation coefficients into direct and indirect effects of independent variables on dependent variables. The study of path analysis specified that total correlation coefficients with β-ODAP fresh and dry as well as the outcome of partitioning these correlations interested in indirect and direct effects contributions through other variables. From Correlation Studies, we found that the correlation coefficient between β-ODAP (fresh and dry) and the predictor variables is almost equal to its total effect. Therefore, correlation explains the true relationship and a direct selection through these traits will be effective. However, when the positive direct and indirect effects were added to the negative direct and indirect effects for traits, the sum of direct and indirect effects of the studied traits was positive and negative which means days to 1st flowering (1st DF), days to 50% flowering (50% DF), seed pod^{-1} (SP) and seed color (SC) have positively influenced the β-ODAP dry and β-ODAP fresh using the total effect. But the remaining traits plant height (PH), primary branches (PB), flower color (FC), pod plant^{-1} (PP), and yield plant^{-1} (YPP) were negatively influenced in both cases. Similar findings regarding plant height (PH) and Pod plant^{-1} (PP) were shown by Yang [50] and also showed plant height negatively and Pod plant^{-1} positively correlated with seed yield at the genotypic level. Lambein et al. [49] reported that the correlation of β-ODAP content with yield and yield-correlated traits such as 1000 seed weight (SW) and Pods plant^{-1} (PP) were significant and negative. So, the negative and significant correlation between β-ODAP and yield should be taken into consideration with this aspect by the breeders.

4.4 Linkage Disequilibrium

The preliminary element of all LD statistics is the variance among the practical and predictable haplotype incidences at polymorphic loci, and the scientific formulas for calculations may be established [21]. Concisely, LD is considered pairwise between two polymorphic positions; and the most commonly used LD procedures are R^2 and D′. The interval of both parameters varies from 0 to 1. But our study showed R^2 value ranged from 0.020–0.696 and D′ ranged from 0.143–0.862, and high LD showed less mapping resolution for less amount of marker as a self-pollinated crop, whereas Association mapping in populaces with low LD needs a high amount of markers [20].

4.5 Association Studies for Yield-Related Agronomic and Biochemical Traits

Different previous studies have reported that the Q value of population structure analysis can effectively improve the reliability of the association analysis [51]. In association analysis, the MLM model is more appropriate than the GLM model [52,53]. The MLM model not only reflects the impact of the Q value in the population structure but also reflects the K value of genetic association that interrupts the association analysis. A total of 99 and 61 MTAs (Marker trait associations) were shown in GLM and MLM models, respectively. In the MLM analysis outcomes, an entire of 5 markers related to agronomic traits were identified in total accessions ($p \leq 0.01$), and phenotypic variance explained (PVE) by these MTAs varied from 10.21% to 45.37%. In the MLM analysis outcomes, an entire of seven markers related to biochemical traits (ODAP content) were identified in total accessions ($p \leq 0.01$), and phenotypic variance explained (PVE) by these MTAs varied from 3.4% to 9.4%. In the future, these markers may be helpful for marker-assisted molecular breeding of grass pea considerable traits. Nevertheless, using a varied germplasm pool, widely distributed markers, multi-year phenotypic data, and multi-environment might reduce MTA analysis error. The 56 markers alleles ($p \leq 0.05$) linked with 17 traits used for GWAS in this study were paralleled to other reported QTLs in grass peas. An earlier study has specified that marker-based gene pyramiding is an active policy for marker-aided selection [54]. In the current study, 21 SSR markers with promising alleles were associated with more than one trait in grass pea and might be useful for the advancement of grass pea accessions in future breeding platforms.

Linkage disequilibrium between a locus and a marker that consults or inspires a phenotypic trait is the foundation for association mapping. The amount of recombination examined differs significantly between classic linkage mapping and association mapping techniques of QTL finding. The likelihood of identifying meaningful marker-trait relationships over small genetic distances improved dramatically when diverse genotypes were used. In our study, among the 56 polymorphic SSR markers, 21 markers were associated with more than one trait in GLM and MLM models among the studied genotypes. In our study, using GLM and MLM models and passing in Bonferroni correction test we observed that the results of GLM and MLM model, a total of 6 markers C41936, C39067, C34100, C47146, C47638, and C43047 associated with our studies considerable traits days to maturity, flower color, plant height and seed per pod were detected in the Hebei and Liaoyang location ($p \leq 0.01$), and the interpretation rate (R^2 value) 11.2% to 43.6% (Table 3). The markers IPL-102 and IPL-65 showed significant association with the number of days to flowering and days to maturity having high R^2 values in both GLM and MLM model analyses reported in a study [51]. So, these markers may be the genetic cause for the relationship between the pleiotropic properties of genes and the traits.

5 Conclusion

The selected six SSR primers, i.e., C41936, C39067, C34100, C47146, C47638, and C43047 might be used in our considerable traits and pooled via marker-assisted selection to advance accessions for excellent yield superiority. The biochemical traits (ODAP content) are low; and negative correlation with all agronomic traits in this study. The evidence of this study delivered the phenotypic variation of yield-

related traits, population structure, and elite alleles stimulated us to take a further study to recommend an exhaustive outline for applying the studied outcomes in the platform for future genetics and in grass pea breeding. So it is a good message for low ODAP and high-yielding grass pea varieties. Low ODAP content high-yielding germplasms might be nominated as parents in the breeding platform based on the association outcomes and it will establish the base for new detailed association mapping lessons.

Acknowledgement: Authors are thankful to the Researchers Supporting Project Number (RSP2025R7) King Saud University, Riyadh, Saudi Arabia.

Funding Statement: We acknowledge the financial support from the Protection and Utilization of Crop Germplasm Resources project from the Ministry of Agriculture and Rural Affairs of China (2019NWB036-07), China Agriculture Research System of MOF and MARA-Food Legumes (CARS-08), National Infrastructure for Crop Germplasm Resources Project from the Ministry of Science and Technology of China (NICGR2019), Agricultural Science and Technology Innovation Program (ASTIP) in CAAS and Bangladesh-Second Phase of the National Agricultural Technology Program-Phase II Project, Bangladesh Agricultural Research Council (BARC), Bangladesh (P149553). This project was also supported by Researchers Supporting Project Number (RSP2025R7), King Saud University, Riyadh, Saudi Arabia.

Author Contributions: Conceptualization, Md. Mosiur Rahman, Md. Ruhul Quddus, Quanle Xu, Muhammad Malek Hossain, Rong Liu, Mengwei Li, Xin Yan, Guan Li, Yishan Ji, Chenyu Wang, Saleh Alfarraj, Tao Yang, Xuxiao Zong and Muhammad Malek Hossain; methodology, Md. Mosiur Rahman, Md. Ruhul Quddus, Quanle Xu, Muhammad Malek Hossain, Rong Liu, Mengwei Li, Xin Yan, Guan Li, Yishan Ji, Chenyu Wang, Saleh Alfarraj, Tao Yang, Xuxiao Zong and Muhammad Malek Hossain; software, Md. Mosiur Rahman, Akbar Hossain, Saleh Alfarraj, Sagar Maitra and Sagar Maitra, Mohammad Javed Ansari; validation, Md. Mosiur Rahman, Md. Ruhul Quddus, Quanle Xu, Muhammad Malek Hossain, Rong Liu, Mengwei Li, Xin Yan, Guan Li, Yishan Ji, Chenyu Wang, Saleh Alfarraj, Tao Yang, Xuxiao Zong and Muhammad Malek Hossain; formal analysis, Quanle Xu, Md. Mosiur Rahman, Saleh Alfarraj, Akbar Hossain, Sagar Maitra and Sagar Maitra, Mohammad Javed Ansari; investigation, Md. Mosiur Rahman, Md. Ruhul Quddus, Quanle Xu, Muhammad Malek Hossain, Rong Liu, Mengwei Li, Xin Yan, Guan Li, Yishan Ji, Chenyu Wang, Ashutosh SarkerS, Tao Yang, Xuxiao Zong and Muhammad Malek Hossain; resources, Quanle Xu, Md. Mosiur Rahman, Ashutosh Sarker, AH and Sagar Maitra, Mohammad Javed Ansari; data curation, Quanle Xu, Md. Mosiur Rahman, Ashutosh Sarker, Akbar Hossain, Sagar Maitra and Saleh Alfarraj,Mohammad Javed Ansari; writing—original draft preparation, Md. Mosiur Rahman, Md. Ruhul Quddus, Quanle Xu, Muhammad Malek Hossain, Rong Liu, Mengwei Li, Xin Yan, Guan Li, Yishan Ji, Chenyu Wang, Ashutosh Sarker, Tao Yang, Xuxiao Zong and Muhammad Malek Hossain; writing—review and editing, Quanle Xu, Md. Mosiur Rahman, Ashutosh Sarker, Akbar Hossain and Sagar Maitra, Mohammad Javed Ansari; visualization, Quanle Xu, Md. Mosiur Rahman, Ashutosh Sarker, Akbar Hossain, Sagar Maitra and Sagar Maitra, Mohammad Javed Ansari; supervision, Quanle Xu, Akbar Hossain, Saleh Alfarraj, Mohammad Javed Ansari, Sagar Maitra and Ashutosh Sarker; project administration,Quanle Xu, Md. Mosiur Rahman, Ashutosh Sarker, Akbar Hossain, Sagar Maitra and Mohammad Javed Ansari; funding acquisition, Quanle Xu, Md. Mosiur Rahman, Ashutosh Sarker, Akbar Hossain, Sagar Maitra and Sagar Maitra, Mohammad Javed Ansari. All authors reviewed the results and approved the final version of the manuscript.

Availability of Data and Materials: The datasets generated during and/or analyzed during the current study are available from the corresponding authors upon reasonable request.

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Supplementary Materials: The supplementary material is available online at https://doi.org/10.32604/phyton.2024.048992.

References

1. Soren KR, Konda AK, Gangwar P, Tiwari VA, Shanmugavadivel PS, Parihar AK, et al. Development of SSR markers and association studies of markers with phenology and yield-related traits in grass pea (*Lathyrus sativus*). Crop Pasture Sci. 2020;71(8):768–75.
2. Rathi D, Chakraborty S, Chakraborty N. Grasspea, a critical recruit among neglected and underutilized legumes, for tapping genomic resources. Curr Plant Biol. 2021;26:100200.
3. Kumar S, Gupta P, Barpete S, Choukri H, Maalouf F, Sarkar A. Chapter 12-Grass pea. In: The beans and the peas. From orphan to mainstream crops. Woodhead Publishing; 2021. p. 273–87. doi:10.1016/B978-0-12-821450-3.00005-6.
4. Affrifah NS, Uebersax MA, Amin S. Nutritional significance, value-added applications, and consumer perceptions of food legumes: a review. Legume Sci. 2023;5(4):e192.
5. Lassoued S, Giosafatto CV, Mariniello L, Neila TF. Morphological characterization and in vitro digestibility of seven *Lathyrus sativus* (grass pea) accessions originating from Eurasia, Africa, and Canada. Eur Food Res Technol. 2023;249(9):2419–32.
6. Aci MM, Lupini A, Badagliacca G, Mauceri A, Lo Presti E, Preiti G. Genetic diversity among Lathyrus ssp. based on agronomic traits and molecular markers. Agronomy. 2020;10(8):1182.
7. Vaezi B, Mohtashami R, Jozian A, Mirzaei A. Evaluation of genotype× environment interaction and stability analysis of grain and forage yield of grass pea (*Lathyrus sativa* L.) genotypes. J Crop Breed. 2023;15(45):183–93.
8. Chowdhury MA, Slinkard AE. Genetic diversity in grasspea (*Lathyrus sativus* L.). Gen Resour Crop Evol. 2000;47:163–9.
9. Arslan M. Importance of grass pea (*Lathyrus sativus* L.) and bitter vetch (*Vicia ervilia* L.) as promising legumes against of global climate change. J Adnan Menderes Univ Agric Fac. 2019;16(1):97–104.
10. Edwards A, Njaci I, Sarkar A, Jiang Z, Kaithakottil GG, Moore C, et al. Genomics and biochemical analyses reveal a metabolon key to β-L-ODAP biosynthesis in Lathyrus sativus. Nature Commun. 2023;14(1):876. doi:10.1038/s41467-023-36503-2.
11. Rajendran K, Sarker A, Singh M, Abd El-Moneim AM, Nakkoul H. Variation for seed protein and ODAP content in grass pea (*Lathyrus sativus* L.) germplasm collections. Ind J Gen Plant Breed. 2019;79(02):438–43.
12. Hanbury CD, Siddique KHM, Galwey NW, Cocks PS. Genotype-environment interaction for seed yield and ODAP concentration of *Lathyrus sativus* L. and *L. cicera* L. in Mediterranean-type environments. Euphytica. 1999;110(1):45–60.
13. Kumar Yadav V, Radhakrishna A, Mohan Das M, Singh T, Yadav S, Sharma P, et al. Deciphering genetic diversity in grass pea (*Lathyrus sativus* L.) collections using agronomic and forage quality traits and SSR markers. J Agric Sci Technol. 2022;24(6):1429–42.
14. Sen Gupta D, Barpete S, Kumar J, Kumar S. Breeding for better grain quality in Lathyrus. In: Gupta DS, Gupta S, Kumar J, editors. Breeding for enhanced nutrition and bio-active compounds in food legumes. Cham: Springer; 2021. p. 131–56. doi:10.1007/978-3-030-59215-8_6.
15. Das A, Parihar AK, Barpete S, Kumar S, Gupta S. Current perspectives on reducing the β-ODAP content and improving potential agronomic traits in grass pea (*Lathyrus sativus* L.). Front Plant Sci. 2021;12:703275.
16. Lioi L, Galasso I. Development of genomic simple sequence repeat markers from an enriched genomic library of grass pea (*Lathyrus sativus* L.). Plant Breed. 2013;132(6):649–53. doi:10.1111/pbr.12093.
17. Arslan M, Basak M, Aksu E, Uzun B, Yol E. Genotyping of low β-ODAP grass pea (*Lathyrus sativus* L.) germplasm with EST-SSR markers. Brazilian Archiv Biol Technol. 2020;63:e20190150.

18. Björn B, Paulo MJ, Kowitwanich K, Sengers M, Visser RGF, van Eck HJ. Population structure and linkage disequilibrium unravelled in tetraploid potato. Theor Appl Genet. 2010;121(6):1151–70.
19. Gaut BC, Long AD. The lowdown on linkage disequilibrium. Plant cell. 2003;15(7):1502–6.
20. Gonçalves L, Rubiales D, Lourenço M, Patto MC. Exploring grass pea (*Lathyrus sativus* L.) genetic diversity in Mediterranean changing climate conditions. Eur J Agron. 2024;156(2):127142. doi:10.1016/j.eja.2024.127142.
21. Beji S, Fontaine V, Devaux R, Thomas M, Negro SS, Bahrman N, et al. Genome-wide association study identifies favorable SNP alleles and candidate genes for frost tolerance in pea. BMC Genomics. 2020;21:1–21.
22. Nicolas SD, Péros JP, Lacombe T, Launay A, Le Paslier MC, Bérard A, et al. Genetic diversity, linkage disequilibrium and power of a large grapevine (*Vitis vinifera* L.) diversity panel newly designed for association studies. BMC Plant Biol. 2016;16:1–9.
23. Crosta M, Romani M, Nazzicari N, Ferrari B, Annicchiarico P. Genomic prediction and allele mining of agronomic and morphological traits in pea (*Pisum sativum*) germplasm collections. Front Plant Sci. 2023;14:1320506.
24. Gawenda I, Schröder-Lorenz A, Debener T. Markers for ornamental traits in Phalaenopsis orchids: population structure, linkage disequilibrium and association mapping. Mol Breed. 2012;30(1):305–16.
25. Kosev V, Vasileva V. Genetic analysis of quantitative traits of grass pea (*Lathyrus sativus* L.) genotypes. Genetika. 2019;51(2):571–84.
26. Mekonen DA, Abraham A, Oselebe H, Afiukwa C, Abebe TD, Zemene A, et al. Estimation of ODAP contents and heritability of quantitative traits in grass pea (*Lathyrus sativus* L.) accessions from North-Western Ethiopia. Cogent Food Agric. 2024;10(1):2319174.
27. Srivastava A, Sharma A, Singh T, Kumar R. Correlation coefficient and path coefficient in field pea (*Pisum sativum* L.). Int J Curr Microbiol Appl Sci. 2018;7(2):549–53.
28. Bahadur L, Devi B. Estimation of correlation and path analysis coefficient among yield and yield attribution trails of field pea (*Pisum sativum* L.). J Pharmaco Phytochem. 2021;10(1):1696–9.
29. IPGRI. Descriptors for *Lathyrus* spp. International Plant Genetic Resource Resources Institute (IPGRI). Rome, Italy, 2000. p. 60.
30. Dellaporta S, Wood J, Hick J. A plant DNA mini-preparation: version 2. Plant Molecular Biology Report. 1983;1:19–21.
31. Rana MM, Aycan M, Takamatsu T, Kaneko K, Mitsui T, Itoh K. Optimized nuclear pellet method for extracting next-generation sequencing quality genomic DNA from fresh leaf tissue. Methods Protocols. 2019;2(2):54.
32. Liu KJ, Muse SV. Power Marker: an integrated analysis environment for genetic marker analysis. Bioinformatics. 2005;21(9):2128–9.
33. Li Z, Cheng F, Lan S, Guo J, Liu W, Li X, et al. Investigation of genetic diversity and epidemiological characteristics of Pasteurella multocida isolates from poultry in southwest China by population structure, multi-locus sequence typing and virulence-associated gene profile analysis. J Vet Med Sci. 2018;80(6):921–9.
34. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software structure: a simulation study. Mol Ecol. 2005;14(8):2611–20.
35. Bradbury PJ, Zhang Z, Kroon DE Casstevens TM, Ramdoss Y, Buckler ES, et al. TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics. 2007;23(19):2633–5.
36. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J Royal Stat Soc. 1995;57(1):289–300.
37. Mohamed A, García-Martínez S, Carbonell P, José Ruiz J, Loumerem M. Genetic diversity assessment of Spanish and some endangered Tunisian pea (*Pisum sativum* L.) accessions based on microsatellite markers (SSRs). Chem Biodiver. 2023;20(5):e202201033.
38. Gumede MT, Gerrano AS, Amelework AB, Modi AT. Analysis of genetic diversity and population structure of cowpea (*Vigna unguiculata* (L.) Walp) genotypes using Single Nucleotide Polymorphism markers. Plants. 2022;11(24):3480.
39. Cai CP, Ye WX, Zhang TZ, Guo WZ. Association analysis of fiber quality traits and exploration of elite alleles in Upland cotton cultivars/accessions (*Gossypium hirsutum* L.). J Integ Plant Biol. 2014;56(1):51–62.

40. Qin HD, Chen M, Yi XD, Bie S, Zhang C, Zhang YC. Identification of associated SSR markers for yield component and fiber quality traits based on frame map and upland cotton collections. PLoS One. 2015;10(1): e0118073.
41. Islam MZ, Khalequzzaman M, Prince MFRK, Siddique MA, Rashid ESMH, Ahmed MSU. Diversity and population structure of red rice germplasm in Bangladesh. PLoS One. 2018;13(5):e0196096.
42. Wang F, Yang T, Burlyaeva M, Li L, Jiang J, Fang L, et al. Genetic diversity of grass pea and its relatives species revealed by SSRs markers. PLoS One. 2015;10(3):e0118542.
43. Tripathi K, Gore PG, Singh M, Pamarthi RK, Mehra R, Gayacharan C. Legume genetic resources: status and opportunities for sustainability. In: Hasanuzzaman M, editor. Legume crops-prospects, production and uses. 2020. p. 182. doi:10.5772/intechopen.91777.
44. Basaran U, Acar Z, Karacan M, Onar AN. Variation and correlation of morpho-agronomic traits and biochemical contents (protein and β-ODAP) in Turkish grass pea (*Lathyrus sativus* L.) landraces. Turk J Field Crops. 2013;18(2):166–73.
45. Oten M. Evaluation of genetic variability of some local grass pea (*Lathyrus sativus* L.) genotypes using different statistical analysis. Legume Res. 2023;46(8):967–72.
46. Rybinski W, Szot B, Rusinek R. Estimation of morphological traits and mechanical properties of grass pea seeds (*Lathyrus sativus* L.) origination from EU countries. Int Agrophys. 2008;22(3):261–75.
47. Singh S, Verma V, Singh B, Sharma VR, Kumar M. Genetic variability, heritability and genetic advance studies in pea (*Pisum sativum* L.) for quantitative characters. Ind J Agric Res. 2019;53(5):542–7.
48. Ratna M, Begum S, Husna A, Dey SR, Hossain MS. Correlation and path coefficient analysis in Basmati rice. Bangladesh J Agric Res. 2015;40(1):153–61.
49. Lambein F, Travella S, Kuo YH, Van Montagu M, Heijde M. Grass pea (*Lathyrus sativus* L.): orphan crop, nutraceutical or just plain food? Planta. 2019;250:821–38.
50. Yang G, Yang Y, Guan Y, Xu Z, Wang J, Yun Y, et al. Genetic diversity of shanlan upland rice (*Oryza sativa* L.) and association analysis of SSR markers linked to agronomic traits. Biomed Res Int. 2021;7588652, 11. doi:10.1155/2021/7588652.
51. Price AL, Zaitlen NA, Reich D, Pattersson N. New approaches to population stratification in genome-wide association studies. Nat Rev Genet. 2010;11(7):459–63. doi:10.1038/nrg2813.
52. Zhang Z, Ersoz E, Lai CQ, Todhunter RJ, Towari HK, Gore MA, et al. Mixed linear model approach adapted for genome-wide association studies. Nat Genet. 2010;42(4):355–60. doi:10.1038/ng.546.
53. Sacco A, Matteo AD, Lombardi N, Trotta B, Punzo A. Quantitative varietal loci pyramiding for fruit quality traits in tomato. Mol Breed. 2013;31(1):217–22. doi:10.1007/s11032-012-9763-2.
54. Sari H, Eker T, Tosun HS, Mutlu N, Celik I, Toker C. Mapping QTLs for super-earliness and agro-morphological traits in RILs population derived from interspecific crosses between *Pisum sativum* × *P. fulvum*. Curr Iss Mol Biol. 2023;45(1):663–76.