

Association analysis of grain yield and yield components in Quinoa (*Chenopodium quinoa* Willd.) using SSR markers

Ebrahim Sourilaki

University of Guilan

Babak Rabiei

rabiei@guilan.ac.ir

University of Guilan

Hassan Marashi

University of Ferdowsi

Vahid Jokarfard

University of Guilan

Andreas Börner

Institute of Plant Genetics and Crop Plant Research

Article

Keywords: Association mapping, General linear model, Grain yield and yield components, Marker-trait association, Multiple linear model

Posted Date: November 28th, 2023

DOI: <https://doi.org/10.21203/rs.3.rs-3621441/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Scientific Reports on March 12th, 2024. See the published version at <https://doi.org/10.1038/s41598-024-56587-0>.

Abstract

Quinoa is a pseudocereal crop that is considered as a valuable crop to provide food and nutrition security worldwide due to its high nutritional value and resistance to environmental stresses. However, no considerable genetic and especially molecular researches has been done on this plant and the genetic basis of many traits is unknown. In this study, the genetic and molecular diversity of 60 quinoa accessions was assessed using agronomically important traits related to grain yield as well as microsatellite (SSR) markers, and informative markers linked to the studied traits were identified using association mapping. The results showed that most of the studied traits had a relatively high diversity, but grain saponin and protein content showed the highest diversity. High diversity was also observed in all SSR markers, but KAAT023, KAAT027, KAAT036, and KCAA014 showed the highest values for most of the diversity indices and can be introduced as the informative markers to assess genetic diversity in quinoa. Population structure analysis showed that the studied population probably includes two subclusters, so that out of 60 quinoa accessions, 29 (48%) and 23 (38%) accessions were assigned to the first and second subclusters, respectively, and eight (13%) accessions were considered as the mixed genotypes. The results of association mapping using general linear model (GLM) identified 39 significant marker-trait associations (MTAs), while 36 significant MTAs were observed in multiple linear model (MLM). The five markers KAAT030, KCAA022, KGA003, KCAA107, and KAAT023 were significantly associated with more than one trait using both the GLM and MLM models and can therefore be suggested for mapping studies in quinoa populations. The findings of the present study provide useful information about the genetic basis of the studied traits and can be used to breed quinoa populations.

1. Introduction

Quinoa (*Chenopodium quinoa* Willd.) is an allotetraploid ($2n = 4x = 36$) that shows amphidiploid inheritance for most qualitative traits¹⁹. It is one of the most important crops in the South American regions³. It is also compatible with desert conditions, warm and dry, frosty and cold climates and can grow in the area with 40–88% moisture and survives 38 to -4 °C temperature^{4,5,17,21}.

Iran is a vast country with varied climatic conditions in which a large part of it is dry. Soil in most parts of Iran has limitations in term of texture and physical and chemical properties¹⁸. Central desert of Iran is prone to farming, but no plant can grow due to water deficit. In such conditions, quinoa cultivation can be developed due to low water requirement and contribute to employment in these areas. Since quinoa is rich in protein, it is an outstanding alternative for rice and can even be used in mixed with rice.

Yield and yield related traits are the most important traits which are controlled by the quantitative trait loci (QTLs), and influenced by the environment²⁹. Breeding yield through conventional methods has slow results due to poor understanding of the genetic control of yield. Molecular markers especially simple sequence repeats or microsatellites (SSRs) are helpful tools for identifying the genomic regions associated with yield and it's related traits and can facilitate the conventional breeding programs using indirect selection by marker-assisted selection (MAS) method²². Microsatellites are one of the greatest

achievements of biologists among all kinds of molecular markers in identifying genetic diversity inter- and intra-species. Codominant inheritance, extension on the genome, high level of polymorphism, reproducibility, and simplicity make them effective and useful for many applications⁴⁸. Previous studies has been shown that microsatellites are highly polymorphic due to the variability in the number of sequence repeats and therefore have been used as important markers in genetic mapping and population studies in many organisms³². Recently, QTL mapping has been used to identify the gene loci involved in the morphological and physiological traits as well as to explore their inheritance and the nature of gene action¹⁰. In the last decade, the application of molecular markers along with accurate phenotyping has led to the identification of several QTLs for improving the grain yield⁴⁹.

Linkage mapping and association analysis are among important quantitative genetic methods that have been used in order to know the relationships between a genotype and a specific phenotype. Although the success of linkage mapping in identifying the QTLs has been proven, given that the detected QTL is more than a few centimorgans and contains hundreds of genes, identifying the suitable candidate QTL is problem^{15,26}. Several generations of selfing to provide the mapping populations such as recombinant inbred lines through controlled crosses is another limitation of the linkage mapping³⁴. In contrast, association mapping as an alternative method to identify the relationship between a marker and a trait has many advantages compared to QTL mapping, including the use of natural populations, increasing QTL resolution and increasing allelic coverage³⁶. This method is more compatible with germplasm and genetic diversity, and it allows the mapping of several traits simultaneously. Therefore, there is no need to create two-parental populations for each trait which results in extra costs for genotypic and phenotypic evaluation. It is also widely used in human and animal genetics, where creating a large segregated population is impossible¹¹.

The first genetic linkage map in quinoa was constructed based on 230 AFLP, 19 SSR, and 6 RAPD markers³². This genetic map consisted of 35 linkage groups with a total length of 1020 cM, and an average distance of 4 cM between markers. Their results showed that SSR markers indicated higher heterozygosity than AFLP and RAPD markers, which can be used in genetic mapping and quinoa breeding programs³². Jarvis et al.²⁰ used 216 SSR markers with an average heterozygosity of 57.0 to construct the linkage map of quinoa. Their map included 38 linkage groups with a total length of 913 cM²⁰. Rodriguez and Isla⁴² evaluated and compared the genetic diversity of quinoa genotypes using AFLP markers and 20 morphological traits. They prepared a similarity tree diagram based on AFLP markers and showed that it is possible to identify molecular similarities and differences between quinoa genotypes that might be associated with important morphological traits such as grain color, panicle color, phenology and geographic distribution⁴². EL-Harty et al.¹² also evaluated the genetic diversity of 32 quinoa genotypes based on 17 qualitative and 11 quantitative traits under the Saudi Arabia conditions and separated the studied genotypes into four different major groups. The results of molecular analysis also showed that the genotypes originating from South America had higher genetic diversity based on

various indices, including the average number of alleles, Shannon's index, Nei's genetic diversity and polymorphic percentage¹².

Although the association mapping method has been used to identify desirable alleles for different traits in different plants^{25,27,44,53}, there are relatively few reports in quinoa. especially for the genotypes used in this study.

In this study, a diverse set of quinoa germplasm was assessed using some important characteristics related to yield and yield components as well as microsatellite markers. The objectives of the current study were to evaluate the genetic diversity in germplasm based on morphological traits and molecular markers and to identify informative markers associated with yield and yield components by association analysis.

2. Materials and methods

2.1 Plant materials and phenotypic assessment

A collection of 60 quinoa accessions was used in this study. All genotypes were obtained from the IPK Gene Bank, Leibniz Institute of Plant Genetics and Crop Plant Research, Germany. The origin and the accession number of the studied genotypes are presented in Table 1. The phenotypic evaluations were carried out at the experimental fields in Kuhdasht (47°36' E, 33°32' N, and altitude of 1195 m) and Poldokhtar (47°42' E, 33°09' N, and altitude of 673 m) counties, Lorestan province, Iran, in two cropping years, 2019–2020 and 2020–2021. To perform the experiments, a randomized complete block design with three replications was used. Seeds of each genotype were planted in three rows of 10 m at a depth of 2 cm with a distance of 40 cm between rows and 3–5 cm between plants on the rows. To measure the studied traits, a random sample of 15 plants from each plot was selected. The measured traits included number of days to flowering, number of days to physiological maturity, plant height (cm), panicle length (cm), stem diameter (cm), grain saponin percentage, grain protein percentage, 1000-grain weight (g), grain yield (kg.ha⁻¹), and harvest index.

2.2 DNA extraction and genotypic assessment

Genomic DNA was extracted from the young and fresh leaves at 2–4 leaf stage (25 days old) according to the CTAB method⁴⁵. The quality and quantity of the extracted DNA were evaluated by electrophoresis on a 1% agarose gel and spectrophotometer (LAMBDA 1050UV). A total of fourthy microsatellite (SSR) markers were used for molecular analysis and amplification of quinoa genomic DNA. The quality and quantity of extracted DNA on 1% agarose gel and spectrophotometric method was determined. A total of 40 SSR markers were selected using previous studies^{20,31} (Table 2).

The amplification (PCR) reaction was carried out in a reaction solution of 20 µl containing 2 µl of PCR buffer (10X), 0.96 µl of MgCl₂ (50 mM), 1.2 µl of dNTPs (2 mM), 0.8 µl of forward and reverse primers each of 10 pmol concentration, 1.2 µl of Taq DNA Polymerase (1 U/µl), 4 µl of the template DNA (70 ng/

μl), and 9.06 μl of ddH₂O¹⁴. The PCR amplification was carried out in a thermal cycler (PTC 100, M.J. Research, Inc.) at the Genomics Laboratory of Faculty of Agricultural Sciences, University of Guilan, Iran. The thermal cycle protocol included initial denaturation at 94°C for 5 min, followed by 30 cycles of 94°C for 1 min (denaturation), 55°C for 1 min (annealing), 72°C for 1 min (extension), and final extension at 72°C for 5 min¹⁴. A total of 3 μl PCR products were denatured and separated by vertical electrophoresis system on 6% polyacrylamide denaturing gels, and electrophoretic bands were revealed using silver staining⁴¹.

Table 1
Characteristics of the studied quinoa genotypes

Row	Genotype	Code	Source	seed color	No.	Genotype	Code	Source	seed color
1	CHEN67	D2190	Peru	Brown	31	CHEN196	D9407	Chile	Yellow
2	CHEN68	D2191	Peru	Golden-brown	32	CHEN199	D9409	Peru	Bright-white
3	CHEN71	D2196	Chile	Light brown	33	CHEN202	D9413	Peru	White
4	CHEN83	D2194	Bolivia	Bright-white	34	CHEN301	D9789	Chile	White
5	CHEN84	D2195	Bolivia	White	35	CHEN204	D9415	Chile	Golden
6	CHEN89	D5078	Bolivia	Bright	36	CHEN205	D9416	Chile	White
7	CHEN90	D5079	Chile	White	37	CHEN206	D9417	Chile	Golden
8	CHEN91	D5081	Bolivia	Golden	38	CHEN207	D9418	Chile	Bright
9	CHEN115	D9316	Bolivia	White	39	CHEN209	D9420	Chile	Bright-white
10	CHEN119	D9319	Bolivia	Whitish-yellow	40	CHEN210	D9421	Chile	White
11	CHEN121	D9336	Chile	Yellow	41	CHEN212	D9426	Chile	Golden
12	CHEN123	D9428	Peru	White	42	CHEN214	D9429	Peru	White
13	CHEN126	D9339	Peru	Bright	43	CHEN215	D9730	Peru	Bright
14	CHEN128	D9320	Chile	Whitish-yellow	44	CHEN216	D9431	Peru	White
15	CHEN133	D9361	Bolivia	Yellow	45	CHEN217	D9432	Chile	Bright
16	CHEN146	D9374	Bolivia	Bright-white	46	CHEN218	D9434	Chile	Whitish-yellow
17	CHEN151	D9382	Chile	White	47	CHEN220	D9439	Peru	Yellow
18	CHEN154	D9385	Peru	White	48	CHEN223	D9442	Chile	Bright-white
19	CHEN156	D9390	Chile	Golden	49	CHEN225	D9443	Peru	White
20	CHEN159	D9376	Bolivia	Brown	50	CHEN255	D9502	Chile	White
21	CHEN167	D9346	Chile	Yellow	51	CHEN268	D9548	Chile	Bright-white

Row	Genotype	Code	Source	seed color	No.	Genotype	Code	Source	seed color
22	CHEN171	D9350	Chile	Bright-white	52	CHEN270	D9558	Chile	Yellow
23	CHEN172	D9351	Peru	White	53	CHEN297	D9786	Chile	Bright-white
24	CHEN179	D9358	Chile	White	54	CHEN299	D9788	Chile	White
25	CHEN182	D9392	Peru	White	55	CHEN328	D9803	Peru	White
26	CHEN189	D9400	Peru	White	56	CHEN364	D9855	Chile	Whitish-yellow
27	CHEN191	D9402	Chile	Bright	57	CHEN371	D9862	Chile	Yellow
28	CHEN193	D9404	Chile	White	58	CHEN390	D9878	Peru	Bright-white
29	CHEN194	D9405	Chile	Bright	59	CHEN391	D9879	Chile	White
30	CHEN195	D9406	Peru	Whitish-yellow	60	CHEN392	D9880	Peru	White

- All genotypes were obtained through personal correspondence with Professor Andreas Brenner from the Leibniz Institute of Plant Genetics and Crop Plant Research, Department of Genebank, Germany. (Attachment file 1)
- all experimental research and field studies on plants, including the collection of plant material, complied with relevant institutional, national, and international guidelines and legislation.

Table 2
Microsatellite (SSR) markers used for molecular evaluation of quinoa genotypes in this study

Marker name	Forward sequence	Reverse sequence	Annealing temperature	Expected PCR product size
KAAT001	tggctatatcatatgcgtaatgtg	gggctcagattgtatctcgac	59°	176
KAAT006	tctgcaggatcggtaacctt	ttgtatctcggttcccact	54°	171
KAAT007	aggtacaggcgcaaggatac	cggtagcatagcacagaacg	55°	197
KAAT023	agattgtatctcggctttcca	cacttcattgtattgcatttagga	51°	225
KAAT024	cctaatagccacggtttccta	ccgctgaatagacacccagt	57°	199
KAAT025	gagtgggagcccagattgta	agcaaagtaaatttcaacaaagca	53°	160
KAAT026	cggagtcagatgggtctggt	tcaagtgcagctcaatcacc	60°	179
KAAT027	tttaaactttattgacccttgaaa	ggatgctattgcattgctga	57°	192
KAAT030	tcaaatatgtgtggaccactctaag	ccaatttctgtaaattgattgactt	53°	203
KAAT031	agagaccaatgccgtaga	gttcgctatagctagaggagtgg	51°	205
KAAT033	tgccaactgacgagacaaag	gcgggagctcatatcttcac	62°	208
KAAT036	ggcagcgatcgtgaaata	gggacccaaattgtatctcg	63°	254
KAAT040	gcatgagtggtaatggagga	cttgaaggagcagtattattcac	67°	166
KAAT047	tctcggttccctactaatttcttg	tttatgcagcaagggttgtaaa	51°	196
KAAT050	tcatgcctaggatcttgcttt	tcgtatacggactaaattgtccac	57°	158
KCAA006	ttgagcaggatgatgtggag	ttggagaaacataccttgttgg	55°	165
KCAA014	gaatttgcatgcccttcatt	ccgccctcgctactatgat	53°	176
KCAA015	tggttgagggcaaacatacc	tgagggtgaagaggaggatg	58°	204
KCAA019	gtagttgggcggatgtgtct	gcgactgagctagcaggttt	56°	163
KCAA022	ccaattgcatgctcctcatt	aatgcaaacatgggaggaga	60°	153
KCAA106	atatggaagtcggccaacag	gcatgctcatcatttgttgc	63°	148
KCAA107	caccagaaccctcgatctaca	tggttactgttgttgttaatttg	67°	264
KGA002	aaagaacgcataccttccaat	aacctagccaacactccctaaa	56°	198
KGA003	attgccgacaatgaacgaat	atgtaaatggcatgtccaac	61°	158
KGA006	aaacaaattctatcattcggttagg	gccaacgagcctgatgtaa	66°	168

Marker name	Forward sequence	Reverse sequence	Annealing temperature	Expected PCR product size
KGA041	tttgggtgcaaattgtgttca	ttccaagaccaaaccctctc	57°	234
KGA042	ttggtagtgggtaagagaacctg	ctccctccagccacataatc	51°	193
KGA047	gcagtgcattgaatttgaca	gaagctggcaccttatacatgc	55°	179
KGA048	acgtcgaggatggctaggt	ccaacaatcatcatcaataccc	58°	207
KGA053	aaatttctgcctctgtgcaa	ctcaaacttctgcctcctga	53°	186
KGA054	tgttgattgataatatgtaattggtgga	cattcataacagcgagagatgg	58°	199
KGA055	cccaaccaccaaacttaca	gaaaggaaagtgtattgcaaagaa	56°	181
KGA056	gactaacggtgtccaaactgc	ccttctgcattacaccgtca	60°	190
KGA059	ataaccactccgatggcaaa	cagccacctggcagttaga	63°	171
KGA109	accttgaaccacaccgaaac	tcgctgctcatcaccatatt	51°	161
KGA111	aatggtaaacagaccagactagca	tgggttcatttagtagaatcaagg	57°	149
KGA114	tgttgagtgcgctttaatgg	aataggtgtagccgcgtagg	53°	178
KGA116	ccttccttctctacgctctcc	tgggacccaaatctttcatag	62°	173
KGA117	gctttgtagacacctgtcatgg	ccactccgatgataaagttagaatg	63°	193
KGA118	gctgtgtttgacccatgttg	caaccacagcaaagggtgtga	67°	183

2.3 Data analysis

For phenotypic data, descriptive statistics including minimum, maximum, mean, standard deviation, and coefficient of variation were calculated. The molecular data were manually scored using the binary coding system, (1) for presence and (0) for absence of each specific allele in all studied genotypes according to molecular weight using DNA size marker 100 bp Fermentas. The number of observed and polymorph alleles for each SSR locus in all quinoa genotypes were calculated. The effective number of alleles, gene diversity index and Shanon's diversity index were evaluated as described by Kimura and Crow²⁴, Nei³⁵ and Lewontin²⁸ respectively. The polymorphic information content (PIC) index, indicating the ability of each marker for distinguishing the genotypes³⁷, was also calculated as expected heterozygosity based on the studies of Botstein et al.⁷. All genetic diversity indices were calculated using the PopGene software ver. 1.32⁵⁰, and the Power Marker software ver. 3.25³⁰. Cluster analysis using the neighbor joining method was used to group and determine the differences and/or the similarities between genotypes.

To perform association mapping, the effective population structure of the studied genotypes into separate subpopulations and identification of mixed genotypes was first performed using Bayesian method by Structure software⁴⁰. This method attributes each of the genotypes to a hypothetical subpopulation with a probability level, so that the degree of linkage disequilibrium in each subpopulation is minimum and the gametic stage equilibrium is maximum. The initial hypothetical subpopulation values were considered from 1 to 10 and to increase the accuracy, ten replications were performed for each subpopulation. Therefore, the admixture model and the allelic frequency independence with 10,000 burn-in and 10,000 Markov Chain Monte Carlo (MCMC) models, were used to obtain the maximum likelihood curve⁶. Structure software calculates a Qst matrix for each K value (actual number of subpopulations), which includes estimating the probability coefficients of membership of each genotype in each subpopulation. In the resulting plot, when the membership percentage of a genotype to a cluster is greater than or equal to 0.7, the genotype is assigned to that cluster, but if the membership percentage is less than 0.7, it is defined as a mixed genotype⁴⁷. The actual number of subpopulations (K) was determined based on ΔK statistic using Evano et al.¹³ method.

General linear model (GLM) and mixed linear model (MLM) were used to identify informative and linked markers with the evaluated traits using TASSEL 3.0 software⁸. In the general linear model, the population structure coefficient matrix (Q matrix) is only used to determine the relationship between markers and traits, while in the mixed linear model, kinship relationships matrix (K matrix) along with Q matrix (K + Q matrices) are used to avoid false correlations between markers-traits².

3. Results and discussion

3.1 Descriptive statistics

Descriptive statistics of the studied traits in 60 quinoa genotypes are presented in Table 3. The results showed that most of the studied traits had a relatively high diversity. The highest diversity was observed in grain saponin and grain protein with 18.54% and 16.24%, respectively. Since in association mapping, the relationship between phenotypic diversity of the studied traits with polymorphisms in the genome (genotypic diversity) are assessed and markers related to the target traits are identified, so the mapping results will be more accurate by increasing the diversity of the studied traits. This diversity is the results of genes recombinations throughout the evolutionary history in natural populations, and in fact all meiotic events during the evolutionary history of the organism are considered in association mapping⁵⁴. Therefore, the existence of high diversity in the studied population in this research can be provide more reliable and creditable results for association mapping.

Table 3
Discriptive statistics of the studied traits in 60 quinoa genotypes

Trait	Minimum	Maximum	Range	Mean $\pm$ SD	CV (%)
Days to flowering	65.28	103.23	38.05	83.76 $\pm$ 1.78	8.34
Days to maturity	105.31	148.23	42.92	131.32 $\pm$ 3.24	11.21
Plant height (cm)	77.34	132.16	54.82	81.76 $\pm$ 8.25	10.27
Panicle length (cm)	12.24	38.16	25.92	29.3 $\pm$ 0.26	7.64
Grain length (mm)	1.43	2.58	1.15	1.75 $\pm$ 0.09	12.76
Grain saponin (%)	0.07	1.91	1.84	0.97 $\pm$ 0.08	18.54
Grain protein (%)	10.21	23.34	13.13	18.22 $\pm$ 0.57	16.24
1000-grain weight (g)	1.55	3.96	2.41	2.43 $\pm$ 0.45	13.24
Grain yield (kg.ha ⁻¹)	2166.3	4208.3	2041.7	3285.44 $\pm$ 2.06	15.39
Harvest index	28.19	58.23	30.04	34 $\pm$ 2.06	11.29

3.2 Genetic diversity of the studied population based on SSR markers

The genetic diversity indices of SSR markers in the sixty studied quinoa genotypes are shown in Table 4. A total of 140 scorable alleles with an average of 3.5 alleles were amplified by 40 SSR markers, of which 136 alleles (97%) were polymorph. Among the studied markers, KAAT027, KAAT036, KAAT040, KCAA014, KGA042, KGA055 and KGA059 with five alleles showed the highest number of polymorph alleles. The effective number of alleles varied from 1.08 (marker KAAT006) to 3.72 and 3.75 (markers KAAT036 and KAAT027, respectively), with an average of 1.81 alleles. The polymorphic information content (PIC) ranged from 0.23 (marker KAAT007) to 0.89 (marker KAAT023) with an average of 0.60. The average values of Shannon's diversity index and Nei's gene diversity index were 0.54 and 0.44 per each SSR locus, respectively. Marker KAAT027 indicated the highest values for both Shannon's (0.82) and Nei's gene (0.80) diversity indices and marker KAAT007 showed the lowest values (0.15 and 0.09, respectively). In total, the relatively high diversity were observed for most of the studied markers, which indicates the appropriate selection and high efficiency of the selected markers in this study. Therefore, markers KAAT023, KAAT027, KAAT036, and KCAA014 showed the highest values for most of the diversity indices in this research and can be recommended as the highest appropriate markers to evaluate genetic diversity in quinoa. On the other hand, it seems that this marker type to be the most informative for assessing genetic diversity in quinoa and for discriminating among the varieties.

Several studies had previously reported the high PIC values and polymorphism for microsatellite markers in quinoa populations^{19,20,31,52}. Maughan et al.³² used microsatellite markers to study quinoa germplasm with the aim of introducing high reproducible and informative microsatellite markers. A total of 1276

gene loci with ATG, ATT, and CA repeats in the genomic library were sequenced, and the microsatellite markers were designed for 397 gene loci. The number of observed alleles in each gene locus was between 2–13 alleles with an average of 4 alleles in each locus, and the heterozygosity of markers was 57% on average. They observed high heterozygosity in 67 markers and introduced these markers as appropriate for linkage mapping as well as use in quinoa breeding programs³². Jarvis et al.²⁰ also used 216 SSR markers to construct quinoa linkage map. The average heterozygosity value of these markers was calculated 0.57. They constructed the first linkage map of quinoa with 216 SSR markers in a recombinant inbred lines (RILs) population consisting of 38 linkage groups with total length of 913 cM²⁰.

Table 4
Diversity indices calculated for the studied SSR markers in 60 quinoa genotypes

SSR marker	Number of observed alleles	Number of polymorph alleles	Effective number of alleles	Polymorphic information content	Shannon's diversity index	Nei's gene diversity index
KAAT001	4	4	2.16	0.56	0.58	0.45
KAAT006	2	2	1.08	0.32	0.21	0.14
KAAT007	2	2	1.15	0.23	0.15	0.09
KAAT023	3	3	2.33	0.89	0.74	0.70
KAAT024	3	3	1.62	0.76	0.63	0.58
KAAT025	3	3	2.12	0.72	0.69	0.65
KAAT026	4	4	2.35	0.46	0.35	0.29
KAAT027	5	5	3.75	0.85	0.82	0.80
KAAT030	2	2	1.11	0.81	0.65	0.59
KAAT031	3	3	1.59	0.76	0.61	0.57
KAAT033	2	2	1.13	0.73	0.65	0.60
KAAT036	5	5	3.72	0.85	0.78	0.75
KAAT040	5	4	3.13	0.77	0.63	0.59
KAAT047	4	4	2.33	0.43	0.35	0.29
KAAT050	2	2	1.13	0.79	0.69	0.62
KCAA006	2	2	1.18	0.85	0.74	0.68
KCAA014	5	5	2.64	0.88	0.79	0.74
KCAA015	2	2	1.24	0.83	0.67	0.63
KCAA019	2	2	1.12	0.79	0.69	0.65
KCAA022	3	3	1.38	0.59	0.46	0.32
KCAA106	4	4	1.61	0.48	0.38	0.35
KCAA107	3	3	1.73	0.58	0.54	0.46
KGA002	3	2	1.23	0.49	0.40	0.28
KGA003	4	4	1.51	0.56	0.43	0.37
KGA006	3	3	1.29	0.64	0.35	0.28
KGA041	4	4	2.16	0.62	0.55	0.46

SSR marker	Number of observed alleles	Number of polymorph alleles	Effective number of alleles	Polymorphic information content	Shannon's diversity index	Nei's gene diversity index
KGA042	5	5	1.95	0.33	0.63	0.41
KGA047	4	4	1.32	0.48	0.37	0.27
KGA048	3	3	1.71	0.44	0.43	0.36
KGA053	4	4	1.81	0.59	0.59	0.51
KGA054	4	3	1.64	0.61	0.59	0.53
KGA055	5	5	2.35	0.68	0.71	0.55
KGA056	3	3	1.69	0.39	0.33	0.30
KGA059	5	5	1.61	0.51	0.57	0.42
KGA109	4	4	1.51	0.55	0.36	0.23
KGA111	3	2	1.33	0.44	0.47	0.37
KGA114	5	5	2.48	0.61	0.65	0.51
KGA116	3	3	1.57	0.46	0.48	0.38
KGA117	4	4	1.55	0.54	0.46	0.38
KGA118	4	4	1.39	0.39	0.44	0.35
Mean	3.5	3.4	1.81	0.60	0.54	0.44

3.3 Population structure

In the genetic studies, population structure which shows the relationships of individuals inter- and intra-population, provides a perspective on the evolutionary relationships between individuals in a population. In ideal association mapping, the studied population should not have a specific structure and not be structurally divided into different subpopulations. If the effect of population structure and kinship relationships on association analysis are not considered, false positive results will be obtained^{9,39}. To identify the population structure in this research, the studied quinoa genotypes were first grouped by cluster analysis using Neighbor-Joining method (Fig. 1). The results showed that 60 quinoa genotypes could be divided into two subpopulations with 38 and 22 genotypes, respectively. Furthermore, to analyze the population structure and determine the actual number of subpopulations, the STRUCTURE software ver. 4.0 was used and the optimal number of subpopulations was calculated based on the change of ΔK versus K values. The results indicated that the number of possible subpopulations in the studied quinoa genotypes is equal to two subpopulations (Fig. 2). Therefore, K = 2 was considered as the optimal number of subpopulations to analyze the population structure and calculate the membership coefficient of individuals in each cluster (Q-matrix). As shown in Fig. 3, amongst 60 genotypes studied in the current

research, 29 (48%) and 23 (38%) genotypes belonged to the first and second subpopulations, respectively, and eight genotypes (13%) which had a membership coefficient of less than 0.7, were considered as mixed genotypes. El-Harty et al.¹² studied the population structure of 32 quinoa genotypes using STRUCTURE software and identified two possible clusters ($K = 2$), with 13 genotypes (41%) belonging to the first cluster and 8 genotypes (25%) belonging to the second cluster, and 11 genotypes (34%) also having mixed structure.

In addition to the population structure, the linkage disequilibrium (LD) is another factor affecting association mapping. Coinheritance of two or more loci on a chromosome within a genetic region is called genetic linkage¹⁴. Linkage equilibrium is the random combinations of alleles at various loci where observed haplotype frequencies agree with the predicted haplotype frequencies in a population, but LD is a non-random association of alleles at various loci and describes non equal haplotype frequencies¹. Linkage disequilibrium values between the studied marker pairs in the present study along with their significant probability levels are shown in Fig. 4. Several demographic and genetic factors are affecting LD significantly, out of which mutation and recombination are the key factors². Population structure, new mutations, autogamy, genetic drift, epistasis, genomic rearrangement, selection and kinship significantly increase LD, while higher recombination and mutation rates, repetitive mutations, gene conversion and outcrossing significantly decrease LD¹⁶. Nevertheless, significant LD is often observed between distant loci located even on different chromosomes causing spurious associations in association mapping studies. It should be noted that selection of a population with higher or lower LD levels depends on the objective of mapping study¹⁴. The extensive level of LD (long stretched or genetically unlinked LD) reduces the number of markers required for marker-trait association (MTAs) but lowers the mapping resolution, whereas less extensive or short stretched LD needs relatively more number of markers to identify a gene but increase mapping resolution¹.

3.4 Association analysis using GLM and MLM models

To identify the linked molecular markers with the measured traits in quinoa genotypes, association analysis were carried out by TASSEL 3.0 software using the general linear model (GLM) based on the membership coefficient matrix (Q-matrix) as well as the mixed linear model (MLM) based on the membership coefficient and kinship relationship matrices (Q + K matrices).

The results of GLM model showed that 39 markers had significant associations with the evaluated traits, of which 25 and 14 associations were significant at 5% and 1% probability levels, respectively (Table 5). Grain yield and 1000-grain weight with six and five significant associations with SSR markers, respectively, had the most number of significant MTA, while for days to flowering, days to maturity, saponin percentage and protein percentage, only three significant MTAs were identified. The phenotypic variation justified by each marker ranged from 12–33%, so that 9 and 24 markers explained only 12% of the phenotypic variation of panicle length and harvest index, while 1 accounted for 33% of the phenotypic variation of grain yield. Also, a number of markers were significantly associated with more than one trait, including 9 (day to maturity and plant height), 36 (plant height and panicle length), 20 (panicle length and

grain length), 22 (grain length and 1000-grain weight), 40 (saponin percentage and harvest index) and 4 (1000-grain weight and grain yield), and especially 24 (with panicle length, saponin percentage, grain yield and harvest index).

The results of MLM model indicated that 36 markers had significant associations with the studied traits, of which 21 associations were significant at 5% and 15 associations at 1% probability levels (Table 5). The MLM model identified fewer marker-trait associations than the GLM model. This discrepancy suggests that the MLM may yield fewer false positive results than the GLM, because the MLM uses the kinship as well as the population structure. The strong population structure and well-differentiated phenotypes among populations might underlie the much smaller number of MTAs identified by MLM³³. As observed in the GLM method, grain yield with six significant MTAs was significantly linked to the highest number of markers and the most tightly linkage was also observed between grain yield and microsatellite marker RM128. Since the MLM model uses Q and K matrices (the membership coefficient and kinship relationship matrices) to map and identify the relationship between markers and traits, so the false positive results are less than the GLM model and seems that the results of this method have more accuracy and validity compared to the GLM model.

Comparison of the results of GLM and MLM models showed that out of 40 MTAs identified by the two methods, 35 associations were significant in both methods and only five associations were different between the two methods (Table 5). There was a significant association between KGA114 and saponin percentage in the MLM method, while it was not significant under the GLM model. In contrast, a significant association was detected between microsatellite markers KGA111, KGA006, KAAT006 and KGA118 with panicle length, grain length, 1000-grain weight and harvest index in the GLM model, respectively, while none of these relationships were not significant under the MLM model. Furthermore, five markers KAAT030, KCAA022, KGA003, KCAA107, and KAAT023 under both GLM and MLM models were significantly associated with more than one trait, which included day to maturity and plant height, panicle length and grain length, panicle length, saponin percentage, grain yield and harvest index, grain length and 1000-grain weight, and 1000-grain weight and grain yield, respectively (Table 5). Therefore, it seems that by selecting superior genotypes based on these markers using marker assisted selection, several traits can simultaneously be improved in the studied population.

Mizuno et al.³³ In order to investigate the population structure and improve the results and reduce the false positive results, they used the MLM model in the investigation of 12 phenotypic traits such as root length, thousand seed weight, plant height and stem diameter in quinoa and reported more accurate results than single linear models K and Q.

Patiranage et al.³⁷ Using linkage mapping method, they investigated the relationship between 10 sets of SNP markers with 17 quantitative and qualitative traits in 310 quinoa genotypes, and while confirming the method used, there was a significant relationship between two sites with 1000 seed weight trait and one site with powdery mildew resistance Poudri and seven loci related to the traits of days to flowering, days to physiological maturity, plant height and Panicle length were reported on the Cq2A chromosome.

Table 5
SSR markers linked to measured traits in the studied quinoa genotypes based
on GLM and MLM models

Trait	SSR marker	GLM		MLM	
		P-value	R ²	P-value	F-value
Days to flowering	KAAT031	0.03	0.18	0.02	7.1
	KAAT025	0.01	0.22	0.02	7.8
	KAAT007	0.01	0.26	0.01	9.3
Days to maturity	KAAT040	0.01	0.25	0.02	7.6
	KAAT027	0.05	0.15	0.01	8.9
	KAAT030	0.03	0.18	0.04	5.3
Plant height (cm)	KCAA106	0.03	0.18	0.03	6.4
	KGA117	0.01	0.24	0.01	8.5
	KGA111	0.02	0.21	0.01	8.1
	KAAT030	0.02	0.20	0.01	7.4
Panicle length (cm)	KCAA022	0.04	0.17	0.01	9.1
	KGA002	0.01	0.22	0.01	8.5
	KGA003	0.02	0.18	0.01	8.3
	KGA111	0.05	0.12	0.08	3.5
Grain length (mm)	KCAA107	0.02	0.20	0.01	7.4
	KGA006	0.04	0.16	0.09	3.6
	KCAA015	0.01	0.25	0.01	7.1
	KCAA022	0.03	0.18	0.03	6.7
Grain saponin (%)	KGA118	0.03	0.18	0.05	4.32
	KGA114	0.07	0.11	0.03	5.23
	KGA041	0.01	0.26	0.01	7.34
	KGA003	0.02	0.20	0.02	6.45
Grain protein (%)	KGA109	0.01	0.27	0.01	6.32
	KGA059	0.02	0.23	0.02	5.51
	KGA116	0.03	0.18	0.04	2.41

1000-grain weight (g)	KCAA107	0.02	0.23	0.01	6.53
	KAAT036	0.01	0.30	0.02	5.74
	KAAT024	0.01	0.28	0.02	5.35
	KAAT023	0.02	0.21	0.05	4.01
	KAAT006	0.05	0.14	0.07	2.93
Grain yield (kg.ha ⁻¹)	KAAT001	0.01	0.33	0.01	9.94
	KAAT023	0.01	0.28	0.02	6.86
	KGA003	0.02	0.27	0.05	3.39
	KAAT047	0.03	0.21	0.03	5.46
	KGA048	0.04	0.17	0.03	5.83
	KCAA006	0.05	0.15	0.02	6.13
Harvest index	KGA054	0.01	0.20	0.02	7.8
	KGA055	0.02	0.17	0.03	7.3
	KGA003	0.03	0.12	0.01	8.5
	KGA118	0.01	0.19	0.09	2.4

Conclusion

In this study, we assessed the genetic diversity of quinoa accessions using some important characteristics related to yield and yield components and microsatellite markers, and identified the population structure and the possible number of subpopulations as well as the significant SSR-trait associations. The results showed that most of the studied traits had a relatively high diversity. A total of 140 scorable alleles with an average of 3.5 alleles per marker were amplified by 40 SSR markers, of which 136 alleles (97%) were polymorph. The relatively high diversity were also observed for most of the studied markers, but KAAT023, KAAT027, KAAT036, and KGA006 showed the highest values for most of the diversity indices in this research and can be recommended as the highest appropriate markers to evaluate genetic diversity in quinoa. The analysis of population structure using cluster analysis with neighbor joining method as well as determination of the actual number of subpopulations by ΔK statistics using STRUCTURE software indicated two possible subpopulations ($K = 2$) in the studied population. Out of 60 studied genotypes in the current research, 29 (48%) genotypes were allocated to the first and 23 (38%) to the second subpopulation, and eight genotypes (13%) were considered as the mixed genotypes. The results of association mapping showed that there were 39 and 36 significant MTAs between the evaluated SSR markers and traits using GLM and MLM models, respectively. Five markers KAAT030, KCAA022, KGA003, KCAA107, and KAAT023 were significantly associated with more than one trait using both GLM and MLM models, which included KAAT030 with day to maturity and plant height, KCAA022

with panicle length and grain length, KGA003 with panicle length, saponin percentage, grain yield and harvest index, KCAA107 with grain length and 1000-grain weight, and KAAT023 with 1000-grain weight and grain yield. The findings of the present study, in addition to demonstrating the effectiveness of association analysis in identifying linked markers with agronomically important traits in quinoa, also provide useful information for further understanding the genetic basis of these traits that will be useful in the future quinoa breeding programs.

Declarations

Conflict of interest

The author declared that they have no conflict of interest.

Funding

This research was funded by the University of Guilan with Grant No. 1635237154.

Author Contribution

All Authors contributed equally to this experiments. E.S.L. performed the experiment and B.R. analyzed the data. E.S.L. wrote the manuscript and B.R. and A.B. revised the manuscript. All authors reviewed and approved the final version of this manuscript.

Acknowledgement

This research was supported by the University of Guilan, Iran. The seeds of all quinoa genotypes were obtained from the IPK Gene Bank, Leibniz Institute of Plant Genetics and Crop Plant Research, Germany.

Data availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

References

1. Al-Maskri, A.Y., Sajjad, M. & Khan, S.H. Association mapping: a step forward to discovering new alleles for crop improvement. *International Journal of Agriculture and Biology*. 14(1), 153–160. <https://doi.org/10.13140/2.1.1925.9524> (2012).

2. Azizi, H., Aalami, A., Esfahani, M. & Ebadi, A. A. Association and structure analysis of some of rice (*Oryza sativa* L.) genetic resources based on microsatellite markers. *Cereal Research*. 7(1), 1–16. <https://doi.org/10.22124/c.2017.2425> (2017).
3. Bhargava, A., Shukla, S. & Ohri, D. *Chenopodium quinoa*. An Indian perspective. *Industrial Crops and Production*. 23, 73–87(2006).
4. Bhargava, A., Shukla, S., Dixit, B. S., Bannerji, R. & Ohri, D. Variability and genotypex cutting interactions for different nutritional components in *Chenopodium album* L. *Horticultural Science (Prague)*. 33(1), 29–38 (2006).
5. Bhargava, A., Shukla, S., Katiyar, R. S. & Ohri, D. Selection parameters for genetic improvement in *Chenopodium* grain yield in sodic soil. *Journal of Applied Horticulture*. 5(1), 45–48 (2003).
6. Borba, T. C. D. O. et al. Association mapping for yield and grain quality traits in rice (*Oryza sativa* L.). *Genetics and Molecular Biology*. 33, 515–524 (2010).
7. Botstein, D., White, R.L., Skolnick, M. & Davis, R.W. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American journal of human genetics*. 32(3), 314–331(1998).
8. Bradbury, P.J. *et al.* TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics*. 23(19), 2633–2635. <https://doi.org/10.1093/bioinformatics/btm308> (2007).
9. Breseghello, F. & Sorrells, M. Association mapping of kernel size and milling quality in wheat (*Triticum aestivum* L.) cultivars. *Genetics*. 172(2), 1165–1177. <https://doi.org/10.1534/genetics.105.044586> (2006).
10. Degenkolbe, T. *et al.* Identification of drought tolerance markers in a diverse population of rice cultivars by expression and metabolite profiling. *PLoS ONE*. 8(5), e63637. <https://doi.org/10.1371/journal.pone.0063637> (2013).
11. DeWan, A. *et al.* HTRA1 promoter polymorphism in wet age-related macular degeneration. *Science*. 314(5801), 989–992. <https://doi.org/10.1126/science.1133807> (2006).
12. EL-Harty, E.H. *et al.* Morphological and molecular characterization of quinoa genotypes. *Agriculture*. 11(4), 286. <https://doi.org/10.3390/agriculture11040286> (2021).
13. Evanno, G., Regnaut, S. & Goudet, J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*. 14(8), 2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x> (2005).
14. Ghomi, Kh., Rabiei, B., Sabouri, H. & Sabouri, A. Mapping QTLs for traits related to salinity tolerance at seedling stage of rice (*Oryza sativa* L.): An agro-genomics study of an Iranian rice population. *OMICS: A Journal of Integrative Biology*. 17(5), 242–251 (2013).
15. Golicz, A. A., Bayer, P. E. & Edwards, D. Skim-based genotyping by sequencing. *Plant Genotyping: Methods and Protocols*. 257–270 (2015).
16. Gupta, P.K., Rustgi, S. & Kulwal, P.L. Linkage disequilibrium and association studies in higher plants: Present status and future prospects. *Plant molecular biology*. 57(4), 461–485. <https://doi.org/10.1007/s11103-005-0257-z> (2005).

17. Jacobsen, S.E. The world wide potential for quinoa (*Chenopodium quinoa* Willd). Food reviews international. 19(1–2), 167–177 (2003).
18. Jamali, S., Sharifan, H., Hezarjaribi, A. & Sepahvand, N.A. The effect of different levels of salinity on germination and growth indices of two cultivars of quinoa. Journal of Water and Soil Resources Conservation. 6(1), 87–98 (2016). (In Persian with English Abstract).
19. Jarvis, D. *et al.* The genome of *Chenopodium quinoa*. Nature. 542(7641), 307–312. <https://doi.org/10.1038/nature21370> (2017).
20. Jarvis, D., Kopo, O., Jellen, E. & Mallory, M. Simple sequence repeat marker development and genetic mapping in quinoa (*Chenopodium quinoa* Willd). J.Genet. 87(1), 39–51 (2008).
21. Jensen, C. R *et al.* Leaf gas exchange and water relation characteristics of field quinoa (*Chenopodium quinoa* Willd.) during soil drying. European journal of Agronomy. 13(1), 11–25 (2000).
22. Kanagaraj, P. *et al.* Microsatellite markers linked to drought resistance in rice (*Oryza sativa* L.). Current science. 98, 836–839. <http://www.jstor.org/stable/24109855> (2010).
23. Kimura, M. & Crow, J.F. The number of alleles that can be maintained in a finite population. Genetics. 49(4), 725–738. <https://doi.org/10.1093/genetics/49.4.725> (1964).
24. Kimura, M., & Crow, J. F. The number of alleles that can be maintained in a finite population. Genetics. 49(4), 725 (1964).
25. Kover, P. X. *et al.* A multiparent advanced generation inter-cross to fine-map quantitative traits in *Arabidopsis thaliana*. PLoS genetics. 5(7), e1000551 (2009).
26. Kumar, S. *et al.* Meta-QTLs, ortho-MQTLs, and candidate genes for thermotolerance in wheat (*Triticum aestivum* L.). Molecular Breeding. 41, 1–22 (2021).
27. Kumar, V. *et al.* Genome-wide association mapping of salinity tolerance in rice (*Oryza sativa* L.). DNA research. 22(2), 133–145 (2015)..
28. Lewontin, R.C. Testing the theory of natural selection. Nature. 236(5343), 181–182. <https://doi.org/10.1038/236181a0> (1972).
29. Li, X. *et al.* Unraveling the complex trait of harvest index with association mapping in rice (*Oryza sativa* L.). PLoS ONE. 7(1), e29350. <https://doi.org/10.1371/journal.pone.0029350> (2012).
30. Liu, K. & Muse, S.V. PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics. 21(9), 2128–2129. <https://doi.org/10.1093/bioinformatics/bti282> (2005).
31. Mason, S.L. *et al.* Development and use of microsatellite markers for germplasm characterization in quinoa (*Chenopodium quinoa* Willd.). Crop Science. 45(4), 1618–1630. <https://doi.org/10.2135/cropsci2004.0295> (2005).
32. Maughan, P.J *et al.* A genetic linkage map of quinoa (*Chenopodium quinoa*) based on AFLP, RAPD, and SSR markers. Theoretical and Applied Genetics. 109(6), 1188–1195. <https://doi.org/10.1007/s00122-004-1730-9> (2004).
33. Mizuno, N., *et al.* The genotype-dependent phenotypic landscape of quinoa in salt tolerance and key growth traits. DNA Research. 27(4), dsaa022. <https://doi.org/10.1093/dnares/dsaa022> (2020).

34. Myles, S. et al. Association mapping: critical considerations shift from genotyping to experimental design. *The Plant Cell*. 21(8), 2194–2202 (2009).
35. Nei, M. Analysis of gene diversity in subdivided populations. *Proceedings of the national academy of sciences*. 70(12), 3321–3323. <https://doi.org/10.1073/pnas.70.12.3321> (1973).
36. Nordborg, M. & Tavaré, S. Linkage disequilibrium: what history has to tell us. *TRENDS in Genetics*. 18(2), 83–90 (2002).
37. Patiranage, D.S.R *et al*. Genome-wide association study in quinoa reveals selection pattern typical for crops with a short breeding history. *Elife*. 11, e66873. <https://doi.org/10.5061/dryad.zgmsbcc9m> (2020).
38. Powell, W. *et al*. The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Molecular breeding*. 2(3), 225–238. <https://doi.org/10.1007/BF00564200> (1996).
39. Pritchard, J.K. & Donnelly, P. Case-control studies of association in structured or admixed populations. *Theoretical population biology*. 60(3), 227–237. <https://doi.org/10.1006/tpbi.2001.1543> (2001).
40. Pritchard, J.K., Stephens, M. & Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics*. 155(2), 945–959. <https://doi.org/10.1093/genetics/155.2.945> (2000).
41. Rabiei, B., Valizadeh, M., Ghareyazie, B., Moghaddam, M. & Ali, A. J. Identification of QTLs for rice grain size and shape of Iranian cultivars using SSR markers. *Euphytica*. 137, 325–332 (2004).
42. Rodriguez, L.A. & Isla, M.T. Comparative analysis of genetic and morphologic diversity among quinoa accessions (*Chenopodium quinoa* Willd.) of the South of Chile and highland accessions. *African J of Crop Sci*. 6(2), 001–007 (2018).
43. Rodríguez-Peña, R. A. et al. Investigating the genetic diversity and differentiation patterns in the *Penstemon scariosus* species complex under different sample sizes using AFLPs and SSRs. *Conservation Genetics*. 19, 1335–1348 (2018).
44. Roy, J. K., Bandopadhyay, R., Rustgi, S., Balyan, H. S. & Gupta, P. K. Association analysis of agronomically important traits using SSR, SAMPL and AFLP markers in bread wheat. *Current science*. 683–689 (2006).
45. Saghai Maroof, M.A., Soliman, K.M., Jorgensen, A.R. & Allard, R.W. Ribosomal DNA spacer length polymorphisms in barley: Mendelian inheritance, chromosomal location and population dynamics. *Proc Natl Acad Sci USA*. 81, 8014–8018 (1984).
46. Simmonds, N.W. The breeding system of *Chenopodium quinoa*. I. Male sterility. *Heredity*. 27(1), 73–82. <https://doi.org/10.1038/hdy.1971.72> (1971).
47. Spataro, G. *et al*. Genetic diversity and structure of a worldwide collection of *Phaseolus coccineus* L. *Theoretical and applied genetics*. 122(7), 1281–1291. <https://doi.org/10.1007/s00122-011-1530-y> (2011).
48. Tabkhkar, N., Rabiei, B., Samizadeh Lahiji, H. & Hosseini Chaleshtori, M. Genetic variation and association analysis of the SSR markers linked to the major drought-yield QTLs of rice. *Biochemical*

- Genetics. 56(4), 356–374. <https://doi.org/10.1007/s10528-018-9849-6> (2018).
49. Vikram, P. *et al.* Drought susceptibility of modern rice varieties: an effect of linkage of drought tolerance with undesirable traits. *Scientific reports*. 5(1), 1–18. <https://doi.org/10.1038/srep14799> (2015).
 50. Ward, S.M. Allotetraploid segregation for single-gene morphological characters in quinoa (*Chenopodium quinoa* Willd.). *Euphytica*. 116(1), 11–16. <https://doi.org/10.1023/A:1004070517808> (2000).
 51. Yeh, F.C. Population genetic analysis of co-dominant and dominant markers and quantitative traits. *Belgian J Bot.* 130, 129–157 (1997).
 52. Zhang, T. *et al.* Development of novel InDel markers and genetic diversity in *Chenopodium quinoa* through whole-genome re-sequencing. *BMC Genomics*. 18(1), 1–15. <https://doi.org/10.1186/s12864-017-4093-8> (2017).
 53. Zhou, J., You, A., Ma, Z., Zhu, L. & He, G. Association analysis of important agronomic traits in japonica rice germplasm. *African journal of Biotechnology*. 11(12), 2957–2970 (2012).
 54. Zhu, C., Gore, M., Buckler, E.S. & Yu, J. Status and prospects of association mapping in plants. *Plant Genome*. 1, 5–20. <https://doi.org/10.3835/plantgenome2008.02.0089> (2008).

Figures

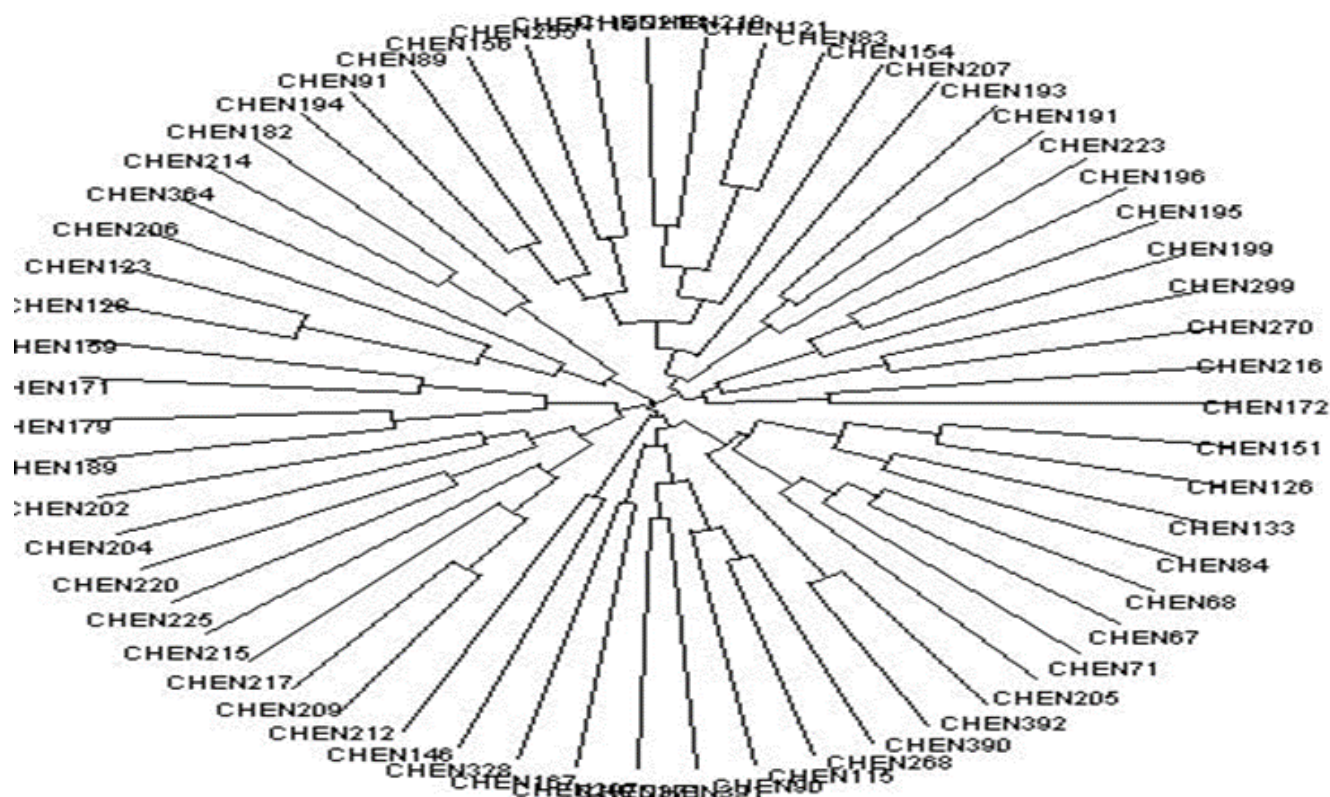


Figure 1

Neighbor joining dendrogram of cluster analysis for grouping 60 quinoa genotypes using 40 SSR markers

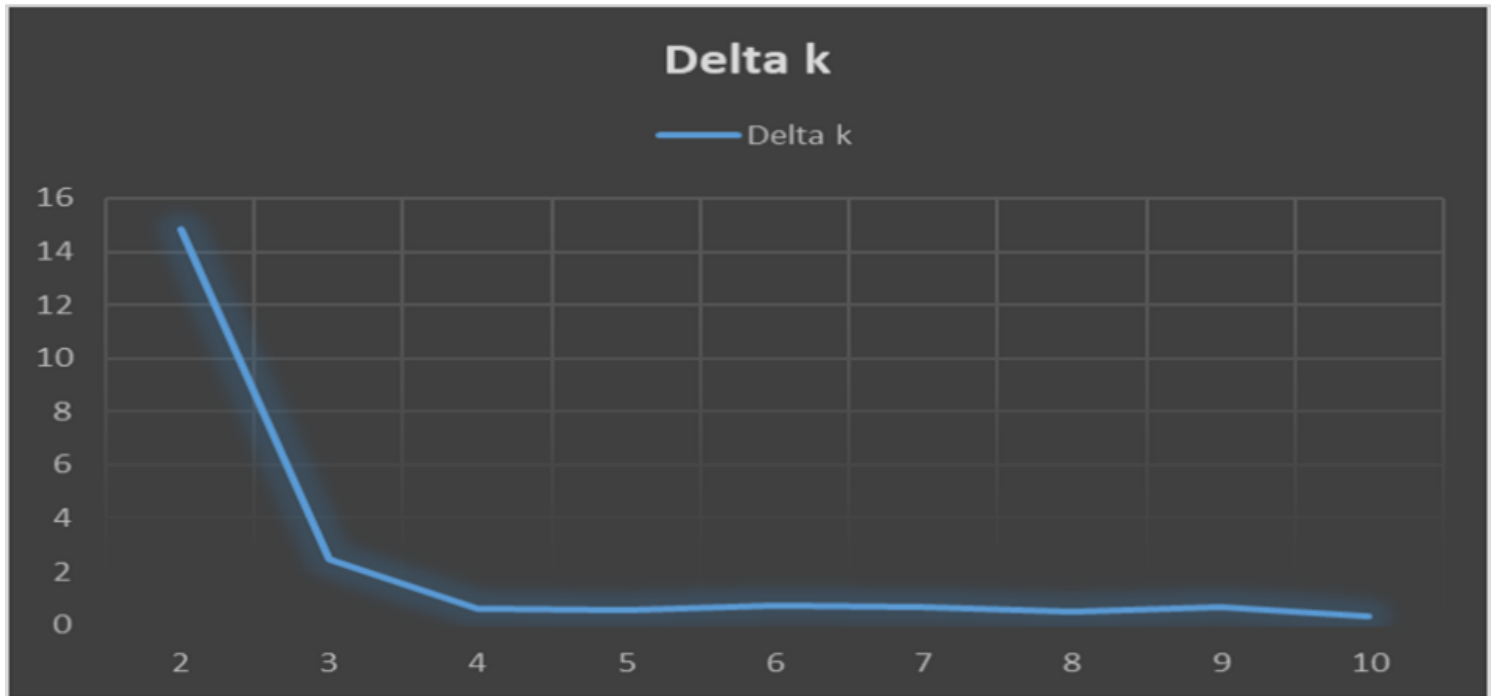


Figure 2

Bilateral charts to determine the number of sub-populations in the studied quinoa genotypes (K=2) based on microsatellite markers

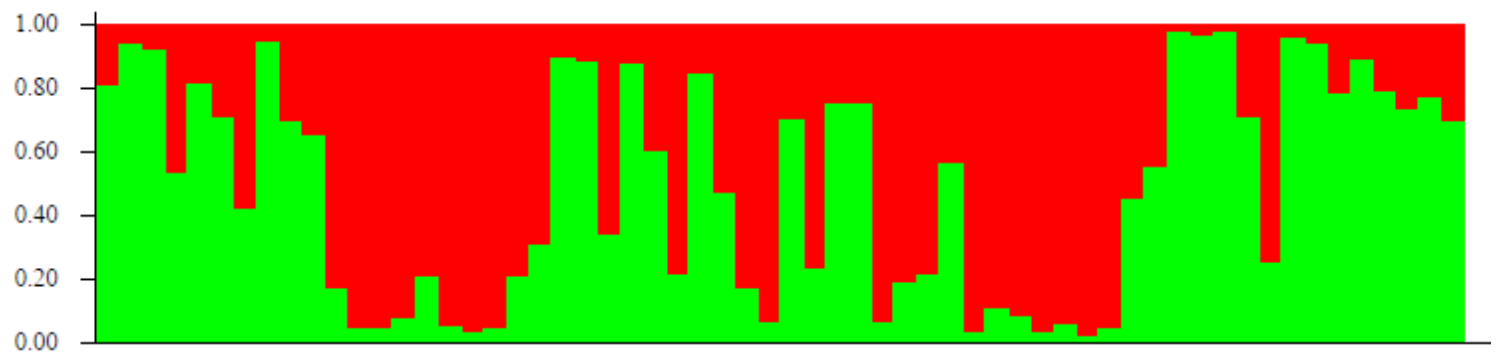


Figure 3

Barplot of structure analysis of the 60 studied quinoa genotypes based on 40 microsatellite markers using Bayesian model. Each color indicate one supopulation or cluster (K=2). Vertical axis show membershipcoefficient of each genotype into each cluster.

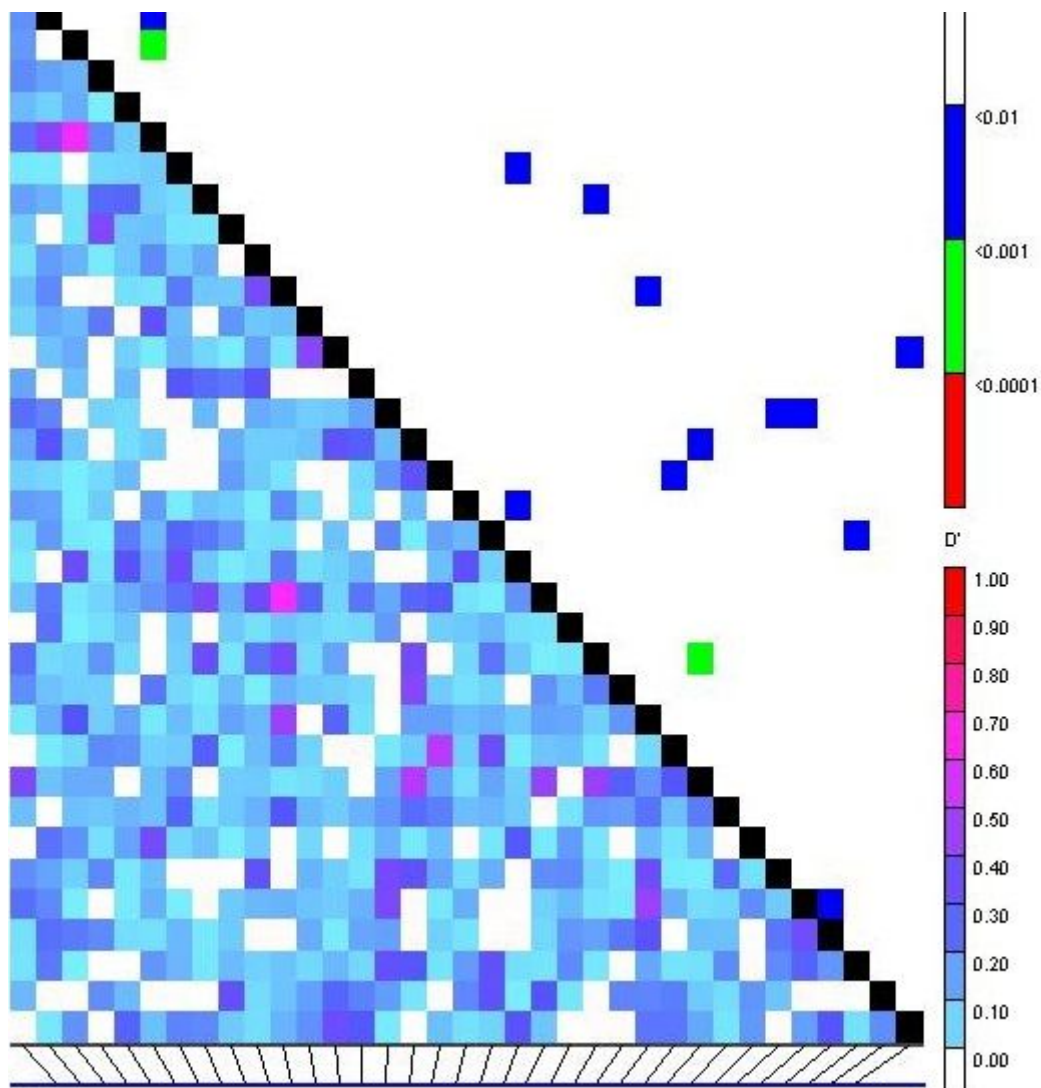


Figure 4

Linkage disequilibrium (LD) plot. Diameter upper and lower indicating linkage disequilibrium and p-value statistics for each pair of marker, respectively.