



**Full Length Article**

## Genetic Diversity and Structure Analysis of a Worldwide Collection of Faba bean (*Vicia faba*) Genotypes using ISSR Markers

Ahmed A. Qahtan, Abdulrahman A. Al-Atar, Eslam M. Abdel-Salam\*, Mohamed A. El-Sheikh, Abdel-Rhman Z. Gaafar and Mohammad Faisal

Department of Botany and Microbiology, College of Sciences, King Saud University, P.O. Box 2455, Riyadh 11451, Saudi Arabia

\*For correspondence: eabdelsalam@ksu.edu.sa; eabdelsalam@hotmail.com

Received 31 August 2020; Accepted 23 December 2020; Published 25 January 2021

### Abstract

The success of breeding programs depends on the extent of genetic variability. Inter-simple sequence repeats (ISSR) have been widely utilized in investigations, including the characterization of many plant species genetically. This research aimed to examine both the genetic diversity and relationships of 92 faba bean (*Vicia faba* L.) genotypes from different geographical areas using ISSR markers. Eleven ISSR primers generated a total of 189 repeatable amplified bands, of which 109 were polymorphic. Values of polymorphism information content (PIC) and gene diversity averaged 0.3484 and 0.1438 and ranged 0.089–0.715 and 0.0742–0.2065, respectively. The studied accessions of faba bean plant differentiated into four main clusters, prevalently based on geographical origin through UPGMA clustering analysis and principal component analysis (PCA), deriving four major groupings based on pedigree and origin relationships. The STRUCTURE software analysis results were significantly aligned with the PCA and showed five main clusters; each one represents one continent. AMOVA showed high variation and differentiation among nations from different continents. The discrimination power of ISSR markers obtained in this study suggests that they could be used to examine the diversity of faba bean genotypes efficiently and precisely and encourage targeted crossing strategies. © 2021 Friends Science Publishers

**Keywords:** ISSR; Faba bean; Gene diversity; Genetic differentiation; Polymorphism; Breeding

### Introduction

Faba bean (*Vicia faba* L.  $2n = 12$ ) is an annual herbal plant that belongs to the Fabaceae family and is considered one of the most important legume crops used in human and animal nutrition because of its high seed protein content (25%–30%) and nutritional value (Crépon *et al.* 2010; Hou *et al.* 2015). This crop is widely cultivated globally, where the main producer of beans in the world is China (34.5%), followed by Ethiopia, Australia, France and Morocco (FAOSTAT 2016). According to FAOSTAT (2016), the total cultivated area of faba beans reached approximately 2.15 million hectares yielding more than 4.14 million tons. Faba bean production ranked seventh among all legume crops worldwide. However, in recent decades, the production has noticeably decreased, mainly due to biotic (diseases, weeds, and insects) and abiotic (drought, waterlogging, and marginal management) conditions (El-Fouly 1982). In the last decade, faba bean production was remarked with the lowest yield growth rate among all legume crops, which mainly attributed to different biotic and abiotic stresses (Maalouf *et al.* 2019). The most efficient way to overcome the stress that limits faba bean production

is to generate varieties appropriate to the cultivating conditions of faba beans. In years, faba bean breeders have succeeded in improving disease and pest resistance, quality, yield, and agronomic performance by utilizing excellent germplasm assets. For example, Sallam *et al.* (2016) identified new quantitative trait loci (QTL) associated with frost tolerance in faba bean using QTL mapping and GWAS analysis. Kaur *et al.* (2014) detected polymorphism in 45 worldwide faba bean accessions using 657 single nucleotide polymorphism markers and concluded that heterozygosity among accessions was high because of the allogamous nature of the species. However, because of the frequent use of a small number of elite germplasm lines, genetic diversity in faba bean is regularly diminished. Therefore, some studies indicated the declining genetic diversity of faba beans (Yadav *et al.* 2017). For example, the faba bean genetic diversity in China (the largest producer of this crop) was found to be very narrow as based on the expressed sequence tag marker results of different accessions from Europe and China (Gong *et al.* 2011). Similarly, in the recent 50 years, the genetic diversity of faba beans in Tunisia was reported to be declining (Babay *et al.* 2020). Hence, enriching the faba bean genetic diversity becomes

critical to the development process of new varieties characterized by tolerance/resistance to abiotic and biotic stresses (Alghamdi *et al.* 2015, 2017; Ammar *et al.* 2015). Recently, Nurmansyah *et al.* (2020) examined the genetic diversity and population structure of 120 faba bean samples, including 116 mutants, using amplified fragment length polymorphism (AFLP) and found that these samples have narrow genetic diversity, and the mutant did not affect the genetic distances between samples.

The existence of genetic diversity to characterize and understand the variation among plant resources is a precondition in the development of an effective plant breeding program (Upadhyaya *et al.* 2011). Previously, typical genetic diversity studies depended on agronomic and morphological traits (Ferguson and Robertson 1999). However, these traits are not sufficient to cover the whole genome diversity, sensitive to the change of environmental conditions and dependent on the developmental stage (Lee 2006). Consequently, it is important to identify the degree of diversity and relation among accessions through its characterization at the molecular level to support agromorphological diversity characterization.

Molecular markers are not affected by environmental conditions, can be implemented quickly, and are safe for selecting remarkable agricultural traits. Various DNA markers, such as inter-simple sequence repeats (ISSR), have been developed and successfully used to characterize genetic diversity in various crop plants. Outstanding progress has been made throughout previous investigations to identify the genetic similarities or variation for the breeding programs of faba bean; RAPD (Alghamdi 2008; Yahia *et al.* 2014; Backouchi *et al.* 2015), ISSR (Alghamdi *et al.* 2011; Abdel-Razzak *et al.* 2012; Salazar-Laureles *et al.* 2015), AFLP (Ammar *et al.* 2015; Akash *et al.* 2017; Alghamdi *et al.* 2017), SSR (Akash *et al.* 2017; Göl *et al.* 2017; Rebaa *et al.* 2017), and SRAP (Alghamdi *et al.* 2012; Ammar *et al.* 2015; Alghamdi *et al.* 2017). However, previous studies examining the worldwide genetic diversity and relationships between different accessions are lacking. Such studies are critical for the development of new varieties and utilization of genetic diversity resources available all over the world. Therefore, the current study was performed to examine the genetic variability, the intra- and inter-genotypic relationships, and population structure of 92 faba bean accessions originated from 43 countries within five continents based on ISSR markers. The results of the current study will surely help in identifying and selecting genetically contrasting accessions to assist in broadening the germplasm base for future faba breeding programs.

## Materials and Methods

### Plant material

A total of 92 faba bean accessions were used in this study

(Table 1), of which 89 were obtained from the United States Department of Agriculture, Agricultural Research Service (USDA, ARS), Western Regional Plant Introduction Station, Pullman, Washington, United States, while the remaining three accessions were obtained from College of Agriculture, King Saud University, Riyadh, Saudi Arabia. The country of origin and phenotypic characteristics of each accession are shown in supplementary Table (S1).

### DNA extraction

Genomic DNA was extracted from 3-week-old leaves from each faba bean accession using a modified cetyltrimethylammonium bromide (CTAB) method (Khan *et al.* 2007). Genomic DNA was quantified on nanodrop 8000- spectrophotometer at 260/280 nm, while the integrity was further confirmed on 1% agarose gel electrophoresis.

### ISSR amplifications

ISSR primers were obtained from Metabion International AG, Germany. Suitability and reproducibility for genetic diversity of forty-one ISSR were examined. However, 11 primers only showed reproducible and unambiguous bands. The annealing temperature of each primer was optimized based on its melting temperature. Primer sequences and melting temperatures are shown in the supplementary Table (S2). PCR amplification was performed for the extracted DNA from each genotype after dilution to 50 ng/ $\mu$ L. PCR reaction volume was adjusted to 25- $\mu$ L using PCR Master mix (GoTaq® Green Master Mix, 2X, Promega, USA) on a T100™ PCR Thermal Cycler (Bio-Rad, USA). Electrophoresis at 5 V/cm of the obtained PCR products was performed using 1.3% agarose gel supplemented with 0.5  $\mu$ g/mL ethidium bromide in 1 $\times$  TBE buffer. Pictures of the separated PCR products in agarose gels were captured utilizing a Syngene bio-imaging gel documentation system (Syngene, USA).

### ISSR banding patterns analysis

The bands generated via electrophoresis separation for each genotype using ISSR primers were used to assess the genetic diversity based on the pattern of these bands, whether they are present or absent in each lane. On this basis, absent bands were scored a "0" while present ones were scored a "1". This scoring system was applied for the most distinct and prominent bands only. The scoring matrix produced by this method was analyzed using POPGENE 1.32 (Yeh *et al.* 2010) to investigate the genetic diversity among the studied genotypes based on several parameters, *e.g.*, Nei's gene diversity index (He), population diversity (H<sub>S</sub>), the Shannon index (I), estimation of gene flow (Nm), inter-population differentiation (G<sub>ST</sub>), total gene diversity (H<sub>T</sub>) within populations and percentage of polymorphic loci (PPL) (Nei 1973). Partitioning of the genetic variation

**Table 1:** Faba bean accessions from geographical origins worldwide

Origin (continent)	No. of accessions	No. of countries	Origin (country)
Asia	32	15	Saudi Arabia (3), Yemen (2), Syria (2), Lebanon (2), Jordan (2), Iraq (2), China (2), India (2), Iran (2), Japan (2), Kyrgyzstan (2), Pakistan (3), Taiwan (2), Nepal (2), Australia (2).
Africa	16	7	Egypt (4), Sudan (2), Tunisia (2), Morocco (2), Kenya (2), South Africa (2), Ethiopia (2).
Europe	26	13	Turkey (2), Cyprus (2), Finland (2), France (2), Germany (2), Greece (2), Sweden (2), United Kingdom (2), Spain (2), Poland (2), Italy (2), Hungary (2), Belgium (2).
North America	8	3	United States (4), Canada (2), Mexico (2).
South America	10	5	Ecuador (2), Chile (2), Colombia (2), Bolivia (2), Peru (2).

within and among the different studied population was examined via non-parametric analysis of variance (AMOVA) using GenAEx 6.503 software (Peakall and Smouse 2006, 2012). Population genomic analyses were performed via GenAEx through running principal component analysis (PCA) using the haploid binary genetic distance matrix generated from highly confident ISSR to reveal the genetic relationships for the geographic region tested.

### Population structure analysis

STRUCTURE 2.3.4 (Pritchard *et al.* 2000) is a software used for studying population structure using multi-locus genotype data. Structure analysis was performed for the 92 studied faba bean accessions using principal component and Bayesian model-based analyses. First, the optimum number of clusters (K) for whole data was determined via running 10 testing runs of STRUCTURE setting K from 1 to 10. In this stage, correlated allele frequencies and admixture model were assumed in each run, consisting of a burn-in period of 10,000 and Markov Chain Monte Carlo (MCMC) replicates of 100,000. Second, after estimating the optimum K, STRUCTURE software was used again to perform 10 runs setting K from 1 to 5. Nevertheless, each run consisted of a burn-in period of 100,000, and MCMC replicates of 750,000 with correlated allele frequencies and the admixture model assumed. Finally, the most likely K was chosen using the STRUCTURE HARVESTER website (Earl and vonHoldt 2012) depending on the  $\Delta K$  method (Evanno *et al.* 2005) and the plateau criterion (Pritchard *et al.* 2000). The run with the highest likelihood estimates to assign cluster proportions to individuals was used for further representation based on the output. DISTRUCT 1.1 software (Rosenberg 2004) was used for graphical representation of population structure using the output files of STRUCTURE software. The estimated membership cut-off value was determined to 0.8, and all the members with values less than that cut-off were assigned to the mixed group. Afterward, data were fragmented based on estimated membership values, and STRUCTURE analysis was performed again to reveal a lower level of population structure. The same procedure described above was applied again for each fragment of the dataset, starting with testing runs and ending with a graphical presentation using DISTRUCT. Nevertheless, K settings were different for the fragments of the dataset based on the testing runs results.

## Results

### ISSR polymorphism

A preliminary experiment with a bulk of ten various accessions from different countries was performed to screen a total of 41 ISSR primers. This preliminary analysis showed that only 11 primers produced and amplified polymorphic, reproducible, and unambiguous bands (Table 2). Therefore, these 11 primers were selected to characterize further the 92 different genotypes examined in this study. Further analysis using selected ISSR primers produced 189 repeatable and unambiguous bands, of which 109 bands were polymorphic, and thus the calculated average of polymorphic fragments for each primer was 10. The average percentage of polymorphic bands obtained for all genotypes was 56.75%, and the values ranged from 40.0% to 72.2%. The minimum scored fragment size was 100 bp, and the maximum was 1,800 bp. The average PIC value for each primer was 0.3484, and the different PIC values ranged from 0.089 to 0.715. The clearest patterns and best polymorphism were obtained using the dinucleotide repeats (CA)<sub>8</sub>RG (primer FB18) and (AC)<sub>8</sub>TG (primer FB46).

### Genetic diversity

Table 3 shows the different genetic diversity parameters calculated for each geographical location (continent). The PPL for germplasm from different continents ranged from 21.88% (South America) to 82.81% (Asia). On the other hand, Nei's gene diversity (He) of germplasm from different geographical locations ranged from 0.0900 (North America) to 0.2065 (Asia). Similarly, Shannon's information index (I) ranged from 0.1125 (South America) to 0.3233 (Asia). It is evident that the Asian group had the highest genetic diversity, while the South American germplasm group had the lowest genetic diversity. Moreover, African and European accessions had significantly similar genetic diversity parameters indicating high genetic diversity and possible relation between these groups.

Values of Nei's gene diversity index, Shannon's information index, and PPL for the germplasm from different countries were 0.116, 0.172, and 53.282%, respectively. The coefficient of gene differentiation (GST) is considered one of the most reliable gene differentiation methods. In this study, GST values for germplasm from different countries were 0.256. Total diversity and

**Table 2:** Amplification result and polymorphism of ISSR primers used in this study

Primer	Sequence	Tm	Bands No.	Polymorphic bands No.	Polymorphic bands (%)	PIC
1	(AG) <sub>8</sub> T	50	15	6	40.0	0.5157
2	(GA) <sub>8</sub> A	50	18	11	61.1	0.4173
3	(AG) <sub>8</sub> YA	53	22	15	68.2	0.1137
4	(AC) <sub>8</sub> YG	55	14	6	42.9	0.6116
5	(CA) <sub>8</sub> RG	55	18	13	72.2	0.3184
6	(GAG) <sub>4</sub> RC	48	17	9	52.9	0.1075
7	(AG) <sub>8</sub> GC	56	15	9	60.0	0.7151
8	(AC) <sub>8</sub> AG	54	16	8	50.0	0.2148
9	(AC) <sub>8</sub> CG	56	18	10	55.6	0.0898
10	(AC) <sub>8</sub> GA	54	16	9	56.3	0.4167
11	(AC) <sub>8</sub> TG	54	20	13	65.0	0.3122

Tm: Annealing temperature; PIC: Polymorphism information content

**Table 3:** Statistical analysis of genetic diversity in each geographical location using the ISSR marker (standard deviations in parentheses)

Geographical origin	# of Polymorphic Loci	PPL	He	I
Asia	106	82.81	0.2065 (0.1828)	0.3233 (0.2506)
Africa	80	62.50	0.1765 (0.1825)	0.2747 (0.2604)
Europe	82	64.06	0.1716 (0.1831)	0.2677 (0.2600)
North America	45	35.16	0.0900 (0.1461)	0.1444 (0.2198)
South America	28	21.88	0.0742 (0.1518)	0.1125 (0.2233)
Average	68.2	53.282	0.14376	0.22452

PPL: percentage of polymorphic loci; He: Nei's gene diversity; I: Shannon's information index

**Table 4:** Distribution of genetic variability within and between faba bean countries and regions, measured by AMOVA analysis of the ISSR data

	Source of variance	Df	Sum of square	Variance component	% total of variance	Significance
Groups based on geographical origin	Variance among regions	4	362.182	4.045	27%	$P \leq 0.001$
	Variance among countries	38	723.013	7.209	48%	
	Variance within countries	49	185.750	3.791	25%	
	Total	91	1270.946	15.044	100%	

**Table 5:** Nei's unbiased measures of genetic identity and distance between regions through the ISSR analysis

Region *	Asia	Africa	Europe	North America	South America
Asia	****	0.9512	0.9182	0.8667	0.8499
Africa	0.05	****	0.9321	0.8736	0.8654
Europe	0.0854	0.0704	****	0.9034	0.9103
North America	0.143	0.1352	0.1015	****	0.897
South America	0.1626	0.1446	0.094	0.1087	****

\*Nei's genetic identity (above diagonal) and genetic distance (below diagonal)

population diversity among different countries were 0.116 and 0.086, respectively. Nei's unbiased genetic distance analysis for all accessions revealed that Kyrgyzstani (Kg 1 - Kg 2), Canadian (Ca 1 - Ca 2) and Bolivian and Peruvian (Bo 1 - Pe 1) germplasm had the lowest genetic distance (0.0237). On the other hand, several accessions showed the maximum genetic distance (0.4212) including, for example, Japanese (Jp 2) with South African (Za 2) and Belgian (Be

2), Ecuadorian with (Ec 2) Iraqi (Iq 2) and Taiwanese (Tw 2), and Iraqi (Iq 1) with Ethiopian (Et 2), Swedish (Se 2), and Ecuadorian (Ec 1). The full matrix of genetic identity and distances between the 92 studied accessions are shown in the supplementary excel sheet (supplementary Excel S1).

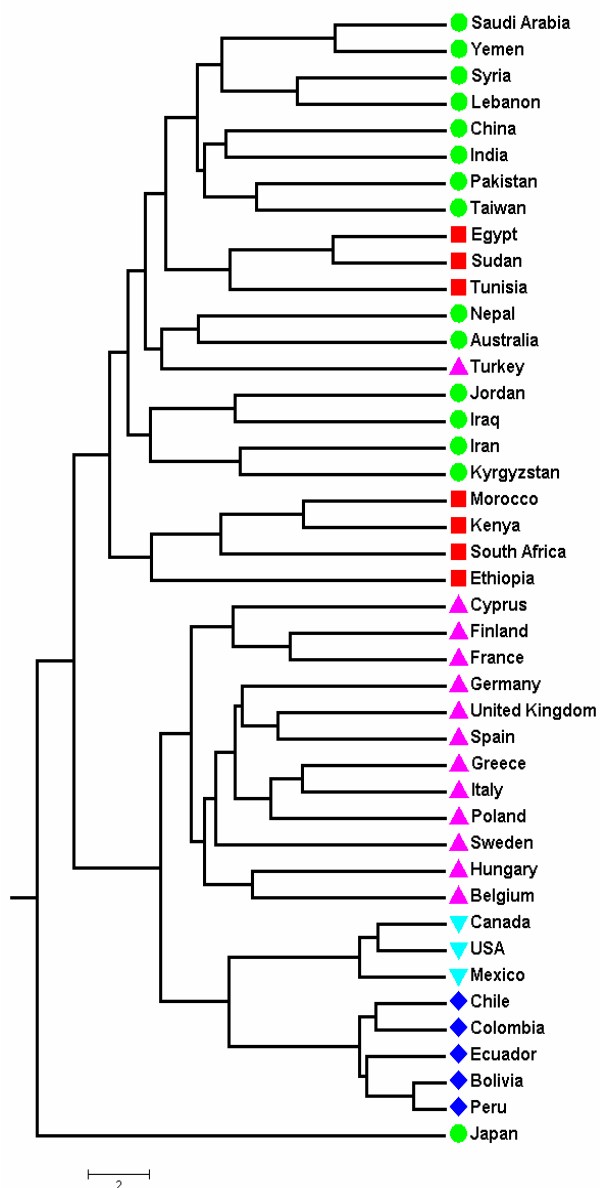
AMOVA analysis was performed to validate the  $G_{ST}$  values obtained. This analysis showed that the majority of genetic variation (48%) between different studied germplasm was due to genetic variation among different countries (Table 4). However, only 27% of the variation was due to variance among different geographical regions (continents), and 25% of the variation could be attributed to variance within countries.

### Genetic relationships between different continents

ISSR allele frequencies were used to calculate Nei's unbiased genetic identity measures and distances between different geographical regions (Table 5). The obtained results showed that the studied faba bean germplasm had a genetic identity ranging between 0.8499 (Asia and South America) and 0.9512 (Asia and Africa). Furthermore, pairwise genetic distances between the studied geographical locations ranged from 0.05 (Asia and Africa) and 0.1626 (Asia and South America).

### Clustering analysis

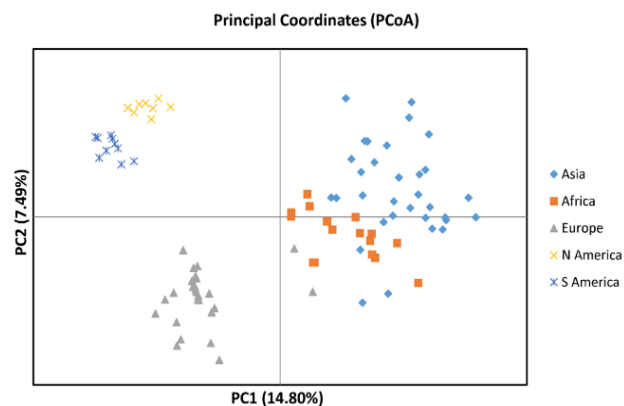
A dendrogram (Fig. 1) was generated based on the UPGMA method for analysis of clusters *via* POPGEN 1.32 and the studied accessions were separated/clustered into four main groups, where the first group included all Asian and some African accessions, the second group included the remaining African accessions, the third group included all European. The fourth group included accessions from South and North America. The Americas faba bean germplasms were clearly separated from others in UPGMA clustering analysis based on ISSR molecular marker data. It was found that faba bean accessions originated in Africa had close genetic relationships with those that originated in Asia. The obtained dendrogram showed that the Japanese accessions had the furthest relationship with the rest of the Asian accessions. Nevertheless, the Peruvian and Bolivian accessions, both from South America, had the closest relationships. Similarly, the largest genetic distance between studied accessions was observed in the Turkish accessions with the rest of European accessions. On the other hand, there was a significant genetic relationship between Greek and Italian accessions. Generally, accessions that originated in Africa showed a close genetic relationship with those that originated in Europe. Analysis of grouped accessions from each continent showed that accessions originated in Africa and Asia was clustered together. The outcomes of this study showed the faba bean resources from all continents are firmly linked to their geographical origin.



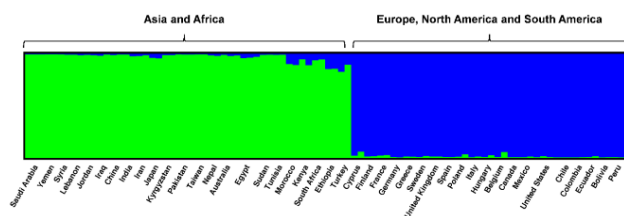
**Fig. 1:** UPGMA dendrogram of faba bean germplasms from geographical origins worldwide based on haploid binary genetic distance matrix generated from highly confident ISSR. Shapes and colors vary based on geographical origin of each accession: Asian (green circles), African (red squares), European (purple triangles), North American (cyan flipped triangle), and South American (blue rhombus)

### PCA analysis

The genetic relationship between individual accessions was further investigated via PCA after labeling each studied accession based on its geographical origin. As shown in Fig. 2, accessions originated in Africa, and Asia showed a strong relationship with their geographical origins. Accessions from Asia, Europe, and the two Americas were quite distinct from each other.



**Fig. 2:** Two-dimension principal components analysis of faba bean accessions from geographical origins worldwide



**Fig. 3:** Population structure of 92 faba bean accessions using STRUCTURE and DISTRUCT software

### Population structure analysis

In this study, STRUCTURE software was used to infer the population structure of the 92 studied faba bean accessions. The  $\Delta K$  and plateau criterion methods were adopted to identify the number of clusters (K). Analyzing the whole data sets reveals that the data has two clusters (the highest value was for K=2). All accessions from Europe, North America, and South America (n = 42) were separated in one cluster; however, the remaining 50 African and Asian accessions were separated in the other cluster (Fig. 3). The cluster of Europe and the Americas contained accessions from Cyprus (2), Finland (2), France (2), Germany (2), Greece (2), Sweden (2), UK (2), Spain (2), Poland (2), Italy (2), Hungary (2), Belgium (2), Canada (2), Mexico (2), USA (4), Chile (2), Colombia (2), Ecuador (2), Bolivia (2) and Peru (2). On the other hand, Africa and Asia's cluster contained accessions from Saudi Arabia (3), Yemen (2), Syria (2), Lebanon (2), Jordan (2), Iraq (2), China (2), India (2), Iran (2), Japan (2), Kyrgyzstan (2), Pakistan (3), Taiwan (2), Nepal (2), Australia (2), Egypt (4), Sudan (2), Tunisia (2), Morocco (2), Kenya (2), South Africa (2), Ethiopia (2) and, surprisingly, Turkey (2).

To reveal the lower levels of population structure, all the accessions of each cluster were analyzed independently using STRUCTURE software. Regarding the cluster of Europe and the Americas (with 42 accessions), STRUCTURE analysis revealed 3 different clusters. All

accessions from European countries (Cyprus, Finland, France, Germany, Greece, Sweden, U.K., Spain, Poland, Italy, Hungary, Belgium) were clustered together. Similarly, accessions from North America (*i.e.*, Canada, Mexico, USA) were clustered together. All the South American (Chile, Colombia, Ecuador, Bolivia, Peru) accessions were clustered together; however, Columbian accessions showed the lowest estimated membership value (0.813), which is still beyond the specified cut-off value.

On the other hand, Asian and African accessions showed clear separation into two clusters. All accessions from African countries (Egypt, Sudan, Tunisia, Morocco, Kenya, South Africa, Ethiopia) were separated into one cluster. In contrast, all the accessions from Asian countries (Saudi Arabia, Yemen, Syria, Lebanon, Jordan, Iraq, China, India, Iran, Japan, Kyrgyzstan, Pakistan, Taiwan, Nepal) were separated in the other cluster. Furthermore, Turkish and Australian accessions were separated in the Asian cluster.

The results obtained using STRUCTURE analysis were significantly aligned with the PCA analysis, which showed five main clusters; each one represents one continent with all the accessions from countries in each continent gathered in one cluster (Fig. 2). Furthermore, PCA showed that Turkish and Australian accessions were separated within the Asian cluster.

## Discussion

In the current study, we analyzed the genetic diversity and relationship of 92 faba bean genotypes grown in different ecotypes distributed in different geographical regions (continents) worldwide. Development of high yield or tolerant crops depends mainly upon the availability of plant genetic resources that could be utilized in different breeding programs aiming to produce plants with superior characteristics (Yadav *et al.* 2017). Providing access to a wide range of genetic diversity significantly enhances plant breeders' capacity to develop new elite cultivars (Galluzzi *et al.* 2020). However, genetic diversity is continuously diminished because of either natural process including domestication and dispersal or breeding programs that aims to selection of specific characteristics (Louwaars 2018). Nevertheless, diversity of genetic resources could be enriched in any crop by identification and detection of new polymorphic bands or new alleles *via* genetic diversity studies (Nurmansyah *et al.* 2020). The obtained results in the current study showed that Asian, European, and African germplasm characterized by the highest genetic diversity and the largest number of effective alleles among all other geographical locations, despite the small size samples of the European ( $n = 26$ ) and African ( $n = 16$ ) groups as compared to the Asian group ( $n = 32$ ). The groups obtained from North America and South America showed the lowest genetic diversity than other groups. This could be attributed to the size of the sample as it was very small (North America:  $n = 8$ , South America:  $n = 10$ ). It is well

established that larger sample sizes retain higher number of polymorphic markers (Franco-Duran *et al.* 2019). Furthermore, the Asian and European accessions represented 15 and 13 different countries, respectively, which adds to the geographical range of the studied accessions and increases genetic diversity. Different growing regions of these accessions have different climate conditions, potentially adding more to the selection of genetic uniqueness and differentiation for the gene pool of faba beans. On the other hand, genotypes from North America and South America represented 3 and 5 countries, respectively, limiting the genetic selection and the genetic diversity among the studied accessions.

ISSR marker analysis was applied in the current study to compare different faba bean genotypes from different geographical areas using locations worldwide. A similar genetic diversity between accessions originated in Africa and Europe was observed in terms of PPL, Nei's gene diversity, and Shannon's information index. However, genetic diversity was less in genotypes of North America and South American groups. These results are in accordance with the results reported earlier by Zong *et al.* (2010). The genetic diversity parameters showed high values (*e.g.*, PPL = 53.282%) among species. These results could be attributed to the collection of different species from different geographical locations (continents) worldwide and thus had significant genetic diversity. Geographical barriers between different countries lead to separation between germplasm and thus higher genetic diversity observed in germplasm collected from different countries/continents (Duc *et al.* 2010; Zong *et al.* 2010; Galluzzi *et al.* 2020). The results of the ISSR marker were further confirmed by AMOVA analysis. The obtained results showed that around 48% of the variation between the studied genotypes could be attributed to variance among countries.  $G_{ST}$  values for germplasm from different countries in this study was 0.256. In the studies of population genetic diversity,  $G_{ST}$  values are considered high if exceeded 0.15 (Nei 1978). The values obtained in this study indicate a great genetic variation between accessions from different geographical locations, confirming the results of other analysis and the results reported earlier (Alghamdi *et al.* 2011; Abdel-Razzak *et al.* 2012; Ammar *et al.* 2015; Akash *et al.* 2017; Alghamdi *et al.* 2017).

The American faba bean accessions were clearly separated from other continents' accessions, in light of PCA and UPGMA grouping analysis. The gene pools from different regions with genetic uniqueness and differentiation might be related to reproductive separation because of wide geographic isolation and non-overlapping natural adaptation (Rajaram 1999; Zong *et al.* 2009). Polignano *et al.* (1999) studied the phenotypic variations among different faba bean germplasm and found that geographical origin has a significant relation to phenotypic variation. Moreover, the genetic diversity of Ethiopian and Afghani faba bean accessions were evaluated based on quantitative characters with an obvious relationship between geographic origin and

phenotypic differences (Polignano *et al.* 1993).

In the current study, population structure analysis using STRUCTURE software revealed the same results as the PCA. Several studies indicated PCA and Bayesian model-based STRUCTURE analysis showed the same results in different plants, *e.g.*, Persian walnut (Mohsenipour *et al.* 2010; Bernard *et al.* 2018), coconut (Odong *et al.* 2011), soybean (Torres *et al.* 2015), sweet cherry (Campoy *et al.* 2016), maize (Vigouroux *et al.* 2008), coffee (Cubry *et al.* 2013) and apples (Liang *et al.* 2015).

Moreover, numerous vital regional-based variations associated with various morphological trait combinations were detected in a collection of faba bean germplasm pure lines (Robertson and El-Sherbeeney 1991) and likewise the case for some traits relating to the tolerance of biotic and abiotic stresses (Duc *et al.* 2010). These morphological data showed extraordinary differentiation from different regions of faba bean genotypes. The current study results confirmed that the genetic differentiation between faba bean genotypes has a clear connection with their geographical origin. Consequently, divergent natural selection and reproductive isolation emerging from wide geographic separation might be a significant explanation behind the faba bean gene pools differentiation from different regions. These results will surely help in development of strategic breeding programs via providing the required information regarding the available genetic resources could be utilized.

Calculated Nei's genetic identity based on ISSR allele frequencies showed a great identity between Asian, African, and European genotypes. The pairwise distances between the genotypes belonging to these geographical locations were lower than the distances between them, and the genotypes belong to North America or South America. Moreover, the results obtained from UPGMA analysis based on the continent that each country belongs to showed that North America and South America groups were significantly isolated from all the other geographical locations. There was a close genetic relationship between Asian, African, and European genotypes. Therefore, further breeding studies examining the potential application of crosses between Asian genotypes and other genetic materials collected from different continents with higher diversity including African and European genotypes may have an extraordinary importance in production of new cultivars or improvement of existing ones. In a previous report, the genetic diversity of 97 elite faba bean accessions from North Africa, Asia, Northern, and Southern Europe were examined using AFLP primers (Zeid *et al.* 2003). The clustering analysis based on PCA and Jaccard's similarity coefficient showed that Asian lines were grouped, while North African and South European lines were grouped. Therefore, it was believed that faba bean mainly originated from the Middle or the Near East and then expanded to Central and Northwestern Europe and finally to North Africa moving to Asia (Zong *et al.* 2009, 2010). Additionally, some similarity was distinguished among

North African and European germplasm in molecular data, which affirmed the spread courses (Cubero 1974). The results of the current study shed light on the global genetic diversity of faba bean germplasm and thus paved the way for future breeding programs aiming to utilize such diversity for production of or improvement of faba bean cultivars.

## Conclusion

This study provided a comprehensive genetic diversity and population structure analysis among 92 faba bean accessions collected from different geographical using ISSR markers. The results obtained in the current study indicated that the genetic diversity and structure of the faba bean populations worldwide are significantly connected to their ecological distribution and geographical origin, useful for germplasm management and future breeding programs. However, there is a need to conduct more extensive studies and examinations of agronomical and morphological characters to affirm the distinct differentiation of gene pools between faba bean accessions from different parts of the world.

## Acknowledgments

The authors are thankful to the Research Supporting Project number (RSP-2020/86), King Saud University, Riyadh, Saudi Arabia.

## Author Contributions

Ahmed A. Qahtan: Conceptualization, Methodology, investigation, software, formal analysis, writing - original draft; Abdulrahman A. Al-Atar: Methodology, validation, funding acquisition, supervision, writing - review and editing; Eslam M. Abdel-Salam: Conceptualization, methodology, investigation, software, formal analysis, writing - original draft; Mohamed A. El-Sheikh: Methodology, validation, writing - review and editing; Abdel-Rahman Z. Gaafar: Conceptualization, methodology, formal analysis, writing - original draft; Mohammad Faisal: Methodology, formal analysis, writing - review and editing.

## References

- Abdel-Razzak H, A Alfrmawy, H Ibrahim, AA El-Hanafy (2012). Genetic diversity in faba bean (*Vicia faba* L.) using inter-simple sequence repeat (ISSR) markers and protein analysis. *Life Sci J* 9:497-503
- Akash MW, W Al-Awaida, A Ateyyeh, H Al-Debei, M Saleh, A Zatimeh, N Salameh, M Alawin, SM Hasan (2017). Exploring genetic variations in faba bean (*Vicia faba* L.) accessions using newly developed EST-SSR markers. *Pak J Bot* 49:667-672
- Alghamdi SS (2008). Varietal identification and genetic purity assessment of F1 hybrid seeds using RAPD markers in faba bean (*Vicia faba* L.). In: *VI International Symposium on In Vitro Culture and Horticultural Breeding* 829, pp:269-274, Brisbane, Australia
- Alghamdi SS, SA Al-Faifi, HM Migdadi, SL Al-Rowaili, EH El-Harty, M Farooq (2017). Genetic diversity and field performance of mung bean, faba bean and lentil genotypes in the Kingdom of Saudi Arabia. *Intl J Agric Biol* 19:689-696



- Alghamdi SS, S Al-Faifi, H Migdadi, M Khan, E EL-Harty, M Ammar (2012). Molecular diversity assessment using sequence related amplified polymorphism (SRAP) markers in *Vicia faba* L. *Intl J Mol Sci* 13:16457–16471
- Alghamdi SS, SA Al-Faifi, HM Migdadi, MH Ammar, KHM Siddique (2011). Inter-simple sequence repeat (ISSR)-based diversity assessment among faba bean genotypes. *Crop Past Sci* 62:755–760
- Alghamdi SS, Al-Shameri AM, Migdadi HM, Ammar MH, El-Harty EH, Khan MA, Farooq M (2015) Physiological and molecular characterization of faba bean (*Vicia faba* L.) genotypes for adaptation to drought stress. *J Agron Crop Sci* 201:401–409
- Ammar MH, F Anwar, EH El-Harty, HM Migdadi, SM Abdel-Khalik, SA Al-Faifi1, M Farooq, SS Alghamdi (2015) Physiological and yield responses of faba bean (*Vicia faba* L.) to drought stress in managed and open field environments. *J Agron Crop Sci* 201:280–287
- Ammar MH, SS Alghamdi, HM Migdadi, MA Khan, EH El-Harty, SA Al-Faifi (2015). Assessment of genetic diversity among faba bean genotypes using agro-morphological and molecular markers. *Saudii J Biol Sci* 22:340–350
- Babay E, K Khamassi, W Sabetta, MM Miazzi, C Montemurro, D Pignone, D Danzi, MM Finetti-Sialer, G Mangini (2020). Serendipitous *in situ* conservation of faba bean landraces in Tunisia: A case study. *Genes* 11; Article 236
- Backouchi I, M Aouida, N Khemiri, M Jebara (2015). Genetic diversity in Tunisian populations of faba bean (*Vicia faba* L.) based on morphological traits and molecular markers. *Genet Mol Res* 14:7587–7596
- Bernard A, T Barreneche, F Lheureux, E Dirlwanger (2018). Analysis of genetic diversity and structure in a worldwide walnut (*Juglans regia* L.) germplasm using SSR markers. *PLoS One* 13; Article e0208021
- Campoy JA, E Lerigoleur-Balsemin, H Christmann, R Beauvieux, N Girollet, J Quero-García, E Dirlwanger, T Barreneche (2016). Genetic diversity, linkage disequilibrium, population structure and construction of a core collection of *Prunus avium* L. landraces and bred cultivars. *BMC Plant Biol* 16; Article 49
- Crépon K, P Marget, C Peyronnet, B Carrouée, P Aresé, G Duc (2010). Nutritional value of faba bean (*Vicia faba* L.) seeds for feed and food. *Field Crops Res* 115:329–339
- CuberoJI(1974). On the evolution of *Vicia faba* L. *Theor Appl Genet* 45:47–51
- Cubry P, F De Bellis, D Pot, P Musoli, T Leroy (2013). Global analysis of *Coffea canephora* Pierre ex Froehner (Rubiaceae) from the Guineo-Congolese region reveals impacts from climatic refuges and migration effects. *Genet Resour Crop Evol* 60:483–501
- Duc G, S Bao, M Baum, B Redden, M Sadiki, MJ Suso, M Vishniakova, X Zong (2010). Diversity maintenance and use of *Vicia faba* L. genetic resources. *Field Crops Res* 115:270–278
- Earl DA, BM vonHoldt (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour* 4:359–361
- El-Fouly MM (1982). Flower and Pod Drop. In: *Faba Bean Improvement: Proceedings of the Faba Bean Conference held in Cairo, Egypt*, pp:177–184, March 7–11, 1981. Hawtin G, C Webb (Eds.). Springer, Dordrecht, Netherlands
- Evanno G, S Regnaut, J Goudet (2005). Detecting the number of clusters of individuals using the software structure: A simulation study. *Mol Ecol* 14:2611–2620
- FAOSTAT (2016). *FAOSTAT statistics database*. FAO, Rome, Italy
- Ferguson ME, LD Robertson (1999). Morphological and phenological variation in the wild relatives of lentil. *Genet Resour Crop Evol* 46:3–12
- Franco-Duran J, J Crossa, J Chen, SJ Hearne (2019). The impact of sample selection strategies on genetic diversity and representativeness in germplasm bank collections. *BMC Plant Biol* 19; Article 520
- Galluzzi G, A Seyoum, M Halewood, I López Noriega, EW Welch (2020). the role of genetic resources in breeding for climate change: The case of public breeding programmes in eighteen developing countries. *Plants* 9; Article 1129
- Göl Ş, S Doğanlar, A Frary (2017). Relationship between geographical origin, seed size and genetic diversity in faba bean (*Vicia faba* L.) as revealed by SSR markers. *Mol Genet Genomics* 292:991–999
- Gong YM, SC Xu, WH Mao, ZY Li, QZ Hu, GW Zhang, J Ding (2011). Genetic diversity analysis of faba bean (*Vicia faba* L.) based on EST-SSR markers. *Agric Sci Chin* 10:838–844
- Hou W, X Zhang, J Shi, Y Liu (2015). Genetic diversity analysis of faba bean (*Vicia faba* L.) germplasms using sodium dodecyl sulfate-polyacrylamide gel electrophoresis. *Genet Mol Res* 14:13945–13953
- Kaur S, NOI Cogan, JW Forster, JG Paull (2014). Assessment of genetic diversity in faba bean based on single nucleotide polymorphism. *Diversity* 6:88–101
- Khan S, MI Qureshi, T Alam, M Abdin (2007). Protocol for isolation of genomic DNA from dry and fresh roots of medicinal plants suitable for RAPD and restriction digestion. *Afr J Biotechnol* 6:175–178
- Lee M (2006). The Phenotypic and Genotypic Eras of Plant Breeding. In: *Plant Breeding: The Arnel R. Hallauer International Symposium*, pp:213–218. Lamkey KR, M Lee (Eds.). Blackwell Publishing Ltd, Oxford, UK
- Liang W, L Dondini, PD Franceschi, R Paris, S Sansavini, S Tartarini (2015). Genetic diversity, population structure and construction of a core collection of apple cultivars from Italian germplasm. *Plant Mol Biol Rep* 33:458–473
- Louwaars NP (2018). Plant breeding and diversity: A troubled relationship? *Euphytica* 214; Article 114
- Maalouf F, J Hu, DM O'Sullivan, X Zong, A Hamwieh, S Kumar, M Baum (2019). Breeding and genomics status in faba bean (*Vicia faba*). *Plant Breed* 138:465–473
- Mohsenipoor S, K Vahdati, R Amiri, MR Mozaffari (2010). Study of the genetic structure and gene flow in Persian walnut (*Juglans regia* L.) using SSR markers. *Acta Hort* 861:133–142
- Nei M (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89:583-590
- Nei M (1973). Analysis of gene diversity in subdivided populations. *Proc Natl Acad Sci USA* 70:3321–3323
- Nurmansyah, SS Alghamdi, HM Migdadi, MA Khan, M Afzal (2020). AFLP-based analysis of variation and population structure in mutagenesis induced faba bean. *Diversity* 12; Article 303
- Odong TL, JV Heerwaarden, J Jansen, TJLV Hintum, FAV Eeuwijk (2011). Determination of genetic structure of germplasm collections: Are traditional hierarchical clustering methods appropriate for molecular marker data? *Theor Appl Genet* 123:195–205
- Peakall R, PE Smouse (2012). GENALEX 6.5: Genetic analysis in Excel. Population genetic software for teaching and research—an update. *Mol Ecol Notes* 28:2537–2539
- Peakall R, PE Smouse (2006). GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes* 6:288–295
- Polignano GB, E Alba, P Ugenti, G Scippa (1999). Geographical patterns of variation in Bari faba bean germplasm collection. *Genet Resour Crop Evol* 46:183–192
- Polignano GB, P Ugenti, G Scippa (1993). The pattern of genetic diversity in faba bean collections from Ethiopia and Afghanistan. *Genet Resour Crop Evol* 40:71–75
- Pritchard JK, M Stephens, P Donnelly (2000). Inference of population structure using multilocus genotype data. *Genetics* 155:945-959
- Rajaram S (1999). Approaches for breaching yield stagnation in wheat. *Genome* 42:629–634
- Rebaa F, G Abid, M Aouida, S Abdelkarim, I Aroua, Y Muhovski, J-P Baudoin, M M'Hamdi, K Sassi, M Jebara (2017). Genetic variability in Tunisian populations of faba bean (*Vicia faba* L. var. major) assessed by morphological and SSR markers. *Physiol Mol Biol Plants* 23:397–409
- Robertson LD, M El-Sherbeeny (1991). Distribution of discretely scored descriptors in a pure line faba bean (*Vicia faba* L.) germplasm collection. *Euphytica* 57:83–92
- Rosenberg NA (2004). distruct: a program for the graphical display of population structure. *Mol Ecol Notes* 4:137–138
- Salazar-Laureles ME, DJ Pérez-López, A González-Huerta, LM Vázquez-García, E Valadez-Moctezuma (2015). Genetic variability analysis of faba bean accessions using Inter-simple sequence repeat (ISSR) markers. *Chil J Agric Res* 75:122–130



- Sallam A, M Arbaoui, M El-Esawi, N Abshire, R Martsch (2016). Identification and verification of QTL associated with frost tolerance using linkage mapping and GWAS in winter faba bean. *Front Plant Sci* 7; Article 1098
- Torres AR, AK Grunvald, TB Martins, MA Santos, NG Lemos, LAS Silva, M Hungria (2015). Genetic structure and diversity of a soybean germplasm considering biological nitrogen fixation and protein content. *Sci Agric* 72:47–52
- Upadhyaya HD, SL Dwivedi, M Ambrose, N Ellis, J Berger, P Smýkal, D Debouck, GDuc, DDumet, A Flavell, SK Sharma, N Mallikarjuna, CLL Gowda (2011). Legume genetic resources: Management, diversity assessment, and utilization in crop improvement. *Euphytica* 180:27–47
- Vigouroux Y, JC Glaubitz, Y Matsuoka, MM Goodman, JG Sánchez, J Doebley (2008). Population structure and genetic diversity of New World maize races assessed by DNA microsatellites. *Amer J Bot* 95:1240–1253
- Yadav S, N Verma, A Singh, N Singh, S Rana, S Ranga, K Kumar (2017). Diversity and development in fababean. *Legum Res* 40:618–623
- Yahia Y, H Hannachi, A Ferchichi (2014). Genetic diversity of *Vicia faba* L. based on random amplified polymorphic DNA and simple sequence repeat markers. *Acta Bot Gallica* 161:151–158
- Yeh F, R Yang, T Boyle (2010). Microsoft windows-based freeware for population genetic analysis, Popgene version 1.31, 1999. Available at: <http://www.ualberta.ca/~fyeh/>. Acesso em 1:12
- Zeid M, CC Schön, W Link (2003). Genetic diversity in recent elite faba bean lines using AFLP markers. *Theor Appl Genet* 107:1304–1314
- Zong X, J Ren, J Guan, S Wang, Q Liu, J Paull, R Redden (2010). Molecular variation among Chinese and global germplasm in spring faba bean areas. *Plant Breed* 129:508–513
- Zong X, X Liu, J Guan, S Wang, Q Liu, JG Paull, R Redden (2009). Molecular variation among Chinese and global winter faba bean germplasm. *Theor Appl Genet* 118:971–978