

Relationship between geographical origin, seed size and genetic diversity in faba bean (*Vicia faba* L.) as revealed by SSR markers

Şurhan Göl¹ · Sami Doğanlar¹ · Anne Frary¹

Received: 2 January 2017 / Accepted: 7 May 2017 / Published online: 11 May 2017
© Springer-Verlag Berlin Heidelberg 2017

Abstract Faba bean (*Vicia faba* L.) is an important legume species because of its high protein and starch content. Broad bean can be grown in different climatic conditions and is an ideal rotation crop because of the nitrogen fixing bacteria in its roots. In this work, 255 faba bean germplasm accessions were characterized using 32 SSR primers which yielded 302 polymorphic fragments. According to the results, faba bean individuals were divided into two main groups based on the neighbor-joining algorithm ($r = 0.91$) with some clustering based on geographical origin as well as seed size. Population structure was also determined and agreed with the dendrogram analysis in splitting the accessions into two subpopulations. Analysis of molecular variance (AMOVA) revealed high levels of within population genetic variation. Genetic similarity and geographical proximity were related with separation of European accessions from African and Asian ones. Interestingly, there was no significant difference between landrace (38%) and cultivar (40%) diversity indicating that genetic variability has not yet been lost due to breeding. A total of 44 genetically well-characterized faba bean individuals were selected for a core collection to be further examined for yield and nutritional traits.

Keywords Broad bean · Fava bean · SSRs · Genetic diversity · Population structure · Core set

Introduction

Faba bean (*Vicia faba* L.), was domesticated in the Levant where archaeological evidence of its cultivation dates to the 10th millennium B.P. (Caracuta et al. 2015). Faba bean cultivation spread to Anatolia and then Europe via the Mediterranean coast and China via Mesopotamia (Cubero 2011). Due to the high protein and starch content of its seeds, this legume has a vital role in the human diet especially in the Mediterranean region, the Middle East, North Africa, much of Asia, and Latin America (Torres et al. 2010). In addition, faba bean has essential vitamins and many antioxidants. Faba bean also harbors high amounts of vicine and convicine, medicinally important components (Ray and Georges 2010) which have anti-nutritional effects. Faba bean's value as a crop is enhanced by its ability to grow in different climates, on different types of soil and to fix nitrogen in its roots. These are essential features for sustainable agriculture on unfavorable land (Arbaoui and Link 2008; Alghamdi et al. 2012a). Faba bean is a facultative cross-pollinated species with $2n = 2x = 12$ chromosomes and the largest described genome size among legumes, approximately 13,000 Mb (Johnston et al. 1999). This is 25 times larger than the genome of the model legume *Medicago truncatula* (Rispaill et al. 2010). During its dispersal, faba bean acquired specific adaptations and was subjected to selection for agronomic traits related to plant architecture and seed size, weight and shape. Based on seed characters, the species can be divided into four groups: *major*, *equina*, *minor*,

Communicated by S. Hohmann.

Electronic supplementary material The online version of this article (doi:10.1007/s00438-017-1326-0) contains supplementary material, which is available to authorized users.

✉ Anne Frary
annefrary@iyte.edu.tr

¹ Department of Molecular Biology and Genetics, Izmir Institute of Technology, Urla, Izmir 35430, Turkey

and *paucijuga* (Cubero 1974). The largest seeds (*major*) emerged in South Mediterranean countries and China, whereas medium-seeded types (*equina*) are mainly cultivated throughout the Middle East, North Africa, and Australia. Small seeds (*minor* and *paucijuga*) are mostly found in Ethiopia and are favored in North European agriculture (Duc 1997).

Plant genetic resources are essential for future food security. These resources provide materials to produce cultivars with new and improved features for changing circumstances including climate change and new disease outbreaks. Although a large number of diverse landraces were cultivated before modern plant breeding, in many places uniform cultivars have taken the place of traditional varieties and landraces. In addition, factors like disease, environmental pollution, and change in dietary habits can negatively affect crop genetic diversity. According to the Rural Advancement Foundation International, about 97% of US Department of Agriculture listed germplasm has been lost in the last 80 years (Upadhyaya et al. 2010). Although the potential value of the rich gene and allele content of wild species and landraces is recognized by many breeders, it is often difficult to use such resources because of a lack of solid information about the genetic value in germplasm collections. A major problem for germplasm evaluation, conservation, regeneration, duplication, and documentation (Van Hintum et al. 2000) is the large size of many germplasm collections. One way to solve this problem is to establish a core set of plant material which represents all of the diversity in the collection in a reasonable number of individuals.

Various molecular genetic marker systems have been used to assess the genetic diversity and relationships between accessions in faba bean collections such as restriction fragment length polymorphism (RFLP) (Torres et al. 1993), random amplified polymorphic DNA (RAPD) (Torres et al. 1993; Link et al. 1995), amplified fragment length polymorphism (AFLP) (Zeid et al. 2003), inter-simple sequence repeat (ISSR) (Terzopoulos and Bebeli 2008), simple sequence repeat (SSR) (Oliveira et al. 2016), and single nucleotide polymorphism (SNP) markers (Kaur et al. 2014; Sallam et al. 2016a, b). However, this work is still very limited when compared to other crop species. Thus, the aim of the present work was to characterize 255 faba bean accessions representing all areas of the world for their molecular genetic diversity. To this end, 32 simple sequence repeat (SSR) markers yielding 302 polymorphic fragments were assayed on the faba bean material allowing determination of their genetic diversity and population structure as well as selection of a core set of germplasm for use by breeders.

Materials and methods

Plant material and DNA isolation

Faba bean seeds were supplied by the Centre for Genetic Resources (CGN, Wageningen, Netherlands), Aegean Agricultural Research Institute (AARI, Izmir, Turkey), Nordic Gene Bank (NGB, Alnarp, Sweden), Jeffrey Paull at the University of Adelaide (Australia), Fred Stoddard at University of Helsinki (Finland) and International Center for Agricultural Research in the Dry Areas (ICARDA, Cairo, Egypt). In total, 255 accessions were used consisting of 151 landraces and 104 cultivars (Supplemental Table 1). The material represented 30 different countries. Turkey was the most common origin with 99 accessions included in the study. In terms of seed size, the 255 individuals were classified into three categories: 99 (39%) large, 85 (33%) medium and 53 (21%) small-seeded individuals with 18 (7%) samples having insufficient seed to determine size. Five seeds from each accession were planted and grown in the greenhouse at Izmir Institute of Technology, Turkey. Young and light green faba bean leaves were collected and used for genomic DNA extraction. Genomic DNA extraction was done with the Wizard Magnetic 96 Plant System (Promega Corp., Madison, WI, USA) and the Beckman Coulter Biomek NX Workstation (Beckman Coulter, Brea, CA, USA) in conformity with the manufacturer's instructions. After extraction, DNA samples were dissolved and homogenized in TE buffer and quantity and quality of the DNA samples were measured with a Nanodrop ND-1000 spectrophotometer (Thermo Scientific™, Vantaa, Finland). DNA samples were stored at -20°C .

SSR analysis

A total of 32 SSR primer pairs (GBSSR-VF and VfG) were selected for analysis in the faba bean accessions (Table 1). Primer information for the 13 GBSSR-VF and 19 VfG primers was obtained from Suresh et al. (2013) and Zeid et al. (2009). These primers were designed from genic and genomic regions of the faba bean genome, respectively.

Polymerase chain reaction (PCR) mixtures were prepared with the following components in 20 μl reaction mixtures for GBSSR-VF: 2 μl 10X PCR buffer (50 mM KCl, 10 mM Tris-HCl, 1.5 mM MgCl_2 , pH: 8.3), 1 μl MgCl_2 (1.5 mM), 0.75 μl dNTP (0.2 mM), 0.75 μl forward and 0.75 μl reverse primers (10 pmol), 0.75 μl Taq polymerase (0.25 U), 13 μl sterile double-distilled water, and 1 μl DNA ($\sim 10\text{ ng}/\mu\text{l}$). PCR conditions were

Table 1 SSR primers used for the molecular genetic analysis, polymorphism ratios and gene diversity

Primer name	Polymorphic/amplified bands	Gene diversity ^a
GBSSR-VF-8	10/10	0.28 ± 0.04
GBSSR-VF-19	4/4	0.22 ± 0.06
GBSSR-VF-20	5/5	0.33 ± 0.07
GBSSR-VF-22	7/7	0.23 ± 0.05
GBSSR-VF-52	7/7	0.36 ± 0.05
GBSSR-VF-113	3/4	0.05 ± 0.01
GBSSR-VF-115	5/5	0.24 ± 0.05
GBSSR-VF-131	7/7	0.15 ± 0.04
GBSSR-VF-149	7/7	0.27 ± 0.05
GBSSR-VF-153	3/3	0.33 ± 0.07
GBSSR-VF-154	4/4	0.17 ± 0.04
GBSSR-VF-159	2/2	0.02 ± 0.01
GBSSR-VF-164	2/2	0.19 ± 0.09
VfG 1	18/18	0.18 ± 0.04
VfG 3	9/9	0.17 ± 0.06
VfG 4	3/3	0.14 ± 0.04
VfG 9	13/13	0.22 ± 0.04
VfG 10	13/13	0.40 ± 0.03
VfG 11	11/11	0.30 ± 0.03
VfG 13	10/10	0.24 ± 0.06
VfG 15	12/12	0.22 ± 0.04
VfG 19	12/12	0.16 ± 0.03
VfG 27	15/15	0.21 ± 0.04
VfG 31	15/16	0.11 ± 0.02
VfG 34	10/11	0.17 ± 0.04
VfG 41	16/16	0.26 ± 0.04
VfG 44	13/13	0.21 ± 0.03
VfG 47	15/15	0.18 ± 0.03
VfG 53	16/16	0.16 ± 0.03
VfG 67	12/12	0.25 ± 0.03
VfG 69	12/12	0.26 ± 0.05
VfG 87	11/11	0.29 ± 0.04

^a For each primer, average gene diversity ± standard error is presented

optimized as: initial step of 10 min at 94 °C for denaturation, 35 cycles with 30 s at 95 °C, 30 s at 55–57 °C as annealing temperature (depending on primer pair), 30 s at 72 °C for extension, and a final extension step of 10 min at 72 °C. Amplification of DNA was performed with BIO-RAD Thermal Cycler™ (BIO-RAD, California, USA). PCR mixtures were prepared with the following components in 20 µl reaction mixtures for VfG: 2 µl 10X PCR buffer (50 mM KCl, 10 mM Tris–HCl, 1.5 mM MgCl₂, pH: 8.3), 1.2 µl MgCl₂ (1.5 mM), 0.5 µl dNTP

(0.2 mM), 1.5 µl forward and 1.5 µl reverse primers (10 pmol), 0.5 µl Taq polymerase (0.25 U), 11.3 µl sterile double-distilled water, and 1.5 µl DNA (~15 ng/µl). PCR conditions were optimized as: initial step of 4 min at 94 °C for denaturation, 30 cycles with 45 s at 95 °C, 1 min at 50–53 °C annealing temperature (depending on primer pair), 1 min at 72 °C for extension, and a final extension step of 7 min at 72 °C. Amplification of DNA was performed with the same instrument as above. After DNA amplification, DNA fragments were separated via Fragment Analyzer™ Automated CE System (Advanced Analytical Technologies, Inc., Ames, USA) using DNF-900 dsDNA Reagent Kit (Advanced Analytical) according to the manufacturer's instructions. SSR alleles were visualized and scored for presence/absence with PROsize 2.0 software version 1.2.1.1 (Advanced Analytical) and visual inspection. Average gene diversity (GD, also known as polymorphism information content or PIC) (Roldán-Ruiz et al. 2000) was calculated for each of the 32 markers using the GDdom program (Abuzayed et al. 2016).

Population structure and diversity analyses

Dissimilarity Analysis and Representation for Windows (DARwin5) software was used to determine genetic diversity of the accessions (Perrier and Jacquemoud-Collet 2006). Allelic data were used to calculate a distance matrix using Dice's coefficient and to perform clustering analysis with the unweighted neighbor-joining algorithm. The program STRUCTURE 2.2.3 was used to determine the number of subpopulations that best represented population structure of the faba bean accessions (Pritchard et al. 2000). The analysis in STRUCTURE was run with burn-in period of 100,000 and 500,000 MCMC replications. Subpopulation numbers (*K*) from 2 to 10 were tested with 20 iterations for each value of *K*. An ad hoc statistic introduced by Evanno et al. (2005) was used to determine the correct estimated number of subpopulations with STRUCTURE HARVESTER online program (Earl and Von Holdt 2012). GenAlEx 6.5 software (Peakall and Smouse 2012) was used to determine genetic diversity parameters of faba bean accessions and to perform analysis of molecular variance (AMOVA).

Core set

The core set (Kim et al. 2007) was constructed by PowerCore (v.1.0) which uses a heuristic search based on the SSR marker data.

Results

SSR analysis

A total of 255 accessions were assayed with 32 SSR primer pairs for determination of allelic polymorphism. Almost all markers (91%) were polymorphic for the tested faba beans with the exception of three markers: GBSSR-VF-113, VfG 31, and VfG 34. Overall, 305 loci were detected in the 255 individuals (Table 1) with an average of 8 alleles per locus. On the whole, 302 (99%) of 305 loci were polymorphic for all of the individuals. Marker VfG 1 amplified the most fragments (18) while the fewest polymorphic fragments (two) were amplified by GBSSR-VF-159 and GBSSR-VF-164. The genetic diversity (GD, also known as polymorphism information content, PIC) of the markers ranged from 0.02 to 0.40 (the maximum possible GD is 0.50). Average GDs for the GBSSR-VF and VfG primers were 0.22 and 0.19, respectively.

Genetic diversity and population structure

The Dice coefficient was used to determine genetic diversity of the 255 accessions (Supplemental Excel file). Maximum genetic dissimilarity was calculated as 0.59 (59%) for genotypes 255 (Disco/2) and 23 (CGN15563). Minimum genetic dissimilarity was 0.17 (17%) for genotypes 241 (FIORD) and 239 (NURA). The 255 units generated 32,385 distance values with a mean of 0.39. Mantel test results indicated a high correlation ($r = 0.91$) between the dissimilarity matrix and dendrogram. Unweighted neighbor-joining analysis resulted in a dendrogram with two clusters and one outgroup (Fig. 1). Group A and Group B had 78 and 176 individuals, respectively. Genotype 89 (CGN10382) was placed as an outgroup by the analysis.

Population structure analysis showed that the faba bean accessions fell into two subpopulations (Fig. 2) with a relatively high ΔK value at $K = 2$ (Supplemental Fig. 1). Standard deviation (SD) for each value of K is also essential to determine the correct number of subpopulations and the SD results supported the choice of $K = 2$ (Supplemental Figs. 2, 3). A cut-off value of 70% was used to group genotypes. Thus, genotypes with an identity value under 70% probability of belonging to a given subpopulation were considered as intermixed (Supplemental Table 1). Based on this, subpopulation 1 had 87 accessions (34.1% of all accessions), subpopulation 2 had 158 accessions (61.9% of accessions) and 10

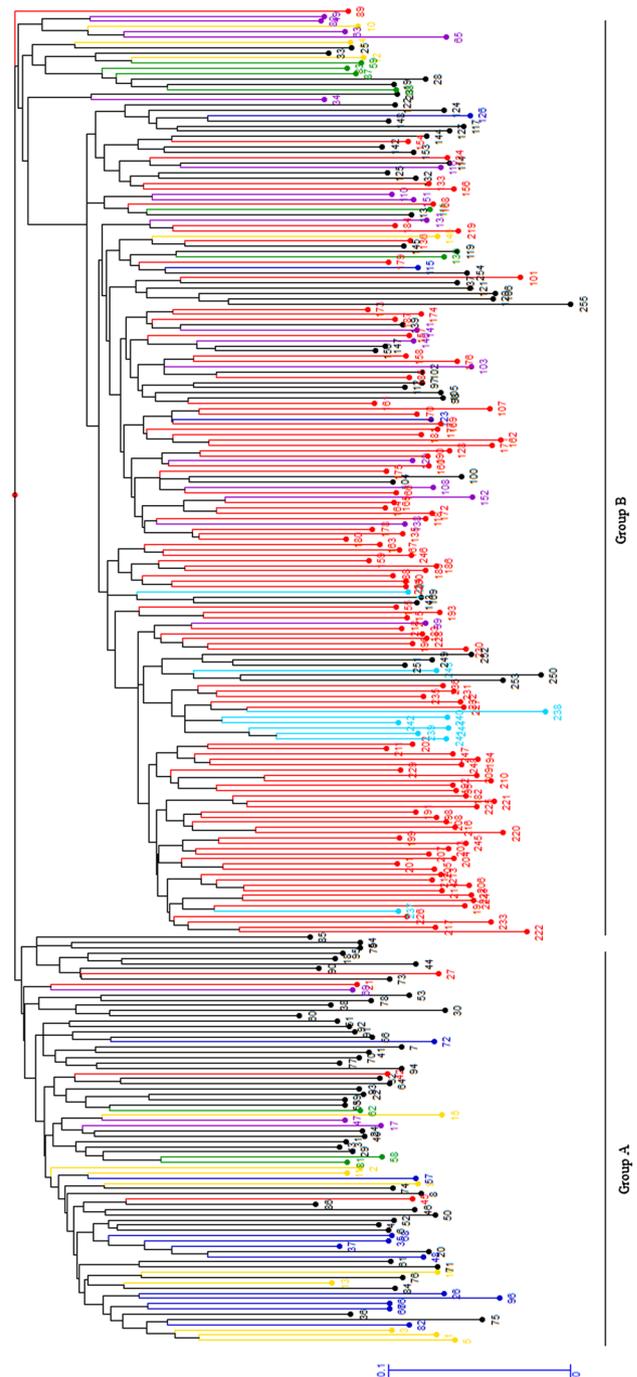


Fig. 1 Genetic diversity of the faba bean genotypes. Regions with the most abundant accessions are color coded: Turkey (red), Netherlands (dark blue), Australia (light blue), Finland (yellow), Afghanistan (green), and Ethiopia (purple). All other genotypes are in black

accessions (3.9%) were intermixed. Genetic diversity (h) and Shannon's information index (I) for subpopulation 1, subpopulation 2 and the intermixed germplasm



Fig. 2 The bar plot for $K = 2$. *Red* and *green* represent subpopulations 1 and 2, respectively. X-axis displays the individuals in the populations and the y-axis shows the percentage identity of belonging to each subpopulation

ranged from 0.15 to 0.22 and from 0.24 to 0.35, respectively. Subpopulation 1 had the highest genetic diversity while the intermixed group had the lowest level of diversity (Table 2). AMOVA was used to estimate the variance among and within populations. The results indicated that the majority of genetic variation in faba bean was due to a remarkable degree of within population variation (90%, Table 3). Only 10% of the genetic variation was attributed to differences among populations.

Core set

A core set was constructed with a total of 44 individuals selected among the 255 faba bean accessions (Table 4). The core set represents the maximum genetic diversity present in all individuals with a minimum number of genotypes. Minimum genetic dissimilarity of the core set was detected between genotypes 237 (Ascot) and 245 (Kitik2003) and was 0.27. The maximum genetic dissimilarity was 0.59 which was between genotypes 23

(CGN15563) and 255 (Disco/2). The mean dissimilarity of the core set was 0.41.

Discussion

Abundant faba bean genetic resources are available for cultivation of this crop in different environments and for further improvement via breeding (Liu and Hou 2010). However, effective leverage of these resources requires knowledge of the extent and patterns of faba bean genetic diversity. Such knowledge is crucial for establishing and managing germplasm collections and ensuring crop development programs (Gwak 2008; Ma et al. 2009). In this study, we determined the genetic diversity of 255 faba bean accessions from throughout the world.

SSR polymorphism

SSR markers are useful for studying genetic diversity because they are highly polymorphic and multi-allelic. In this study, SSR markers produced a sufficient number of polymorphic bands to study faba bean genetic diversity. The highest GD was 0.39 for VfG 10, whereas the lowest value was 0.02 for GBSSR-VF-159 (Table 1). Previous work reported an average GD of 0.69 for 55 GBSSR-VF primer pairs among 32 faba bean accessions (Suresh et al. 2013) and 0.60 for 54 VfG primers (Zeid et al. 2009) among 10 Egyptian and Spanish faba bean accessions. Oliveira et al. (2016) also reported that GD ranged between 0.07 and 0.66 for 28 GBSSR-VF primers among 53 *V. faba*, 2 *V. johannis* and 7 *V. narbonensis* accessions from

Table 2 Analysis of genetic diversity for the two subpopulations and intermixed group of faba bean accessions from worldwide geographical origins

Population	Number of individual	Ne ^a	H ^a	I ^a	Number of private bands
Subpopulation 1	87	1.36 ± 0.02	0.22 ± 0.01	0.35 ± 0.01	14
Subpopulation 2	158	1.30 ± 0.02	0.19 ± 0.01	0.31 ± 0.01	11
Intermixed	10	1.25 ± 0.02	0.15 ± 0.01	0.24 ± 0.02	0

Ne no. of effective alleles = $1/(p^2 + q^2)$, I Shannon's information index = $-1 \times (p \times \ln(p) + q \times \ln(q))$, h diversity = $1 - (p^2 + q^2)$

^a Standard error

Table 3 Analysis of molecular variance (AMOVA) using SSR data for the two subpopulation and intermixed group

Source	DF	SS	MS	Estimated variation	% Variation	p value [†]
Among populations	2.00	649.65	324.83	4.47	10	<0.0001
Within populations	252.00	10344.09	41.05	41.05	90	<0.0001
Total	254.00	10993.74		45.52	100	<0.0001

DF degree of freedom, SS sum of squares, MS mean of squares

[†] With 9999 data permutations

Table 4 Core set selected from the 255 accessions

Genotype number	Genotype name	Genotype number	Genotype name
5	Mikko	132	CGN07875
6	Witkiem Manita	134	CGN10385
8	Kontu	143	CGN10374
10	NGB1547.1	144	CGN10325
18	CGN07874	154	TR12123
20	CGN07699	162	TR23018
23	CGN15563	171	TR31590
24	CGN15619	176	TR33140
27	CGN19987	184	TR37255
29	CGN13485	193	TR44876
30	CGN13464	195	TR44928
40	CGN10391	200	TR49380
44	CGN10347	211	TR53748
46	CGN07826	219	TR61267
53	CGN07716	237	Ascot
54	CGN07844	245	Kıtık2003
57	CGN07781	246	Eresen87
76	CGN07751	249	BPL710
82	CGN15641	252	Giza4
89	CGN10382	253	ILB938/2
100	CGN10371	254	Aurora/2
119	CGN19977	255	Disco/2

Portugal, Spain and Morocco. All of these previous studies used a GD/PIC calculation on a scale of 0–1, thus, the current work found similar levels of GD as previous studies. Work in opium poppy demonstrated that genomic SSRs are more polymorphic than EST-SSRs (Selale et al. 2013; Celik et al. 2014) and may be more useful for assaying functional diversity in germplasm collections (Varshney et al. 2005). However, in our study, there was no significant difference between GD results for genic (GBSSR-VF) and genomic (VfG) SSR markers indicating that both are useful for revealing genetic diversity in this crop.

Genetic diversity and population structure

The dendrogram and population structure analyses supported each other in both analyses divided the accessions into two clusters/subpopulations. Although the analyses are based on different methods (model and distance-based, respectively), both showed similar clustering in the faba bean population. Results for Group A and subpopulation 1 were approximately 98% identical, while Group B and subpopulation 2 were about 90% identical. AMOVA results showed that the genetic differentiation was overwhelmingly due to within population variation. Our data are consistent with previous studies in faba bean which

revealed more genetic variation and differentiation within populations as compared to among populations using ISSR (Terzopoulos and Bebeli 2008; Wang et al. 2012) isozyme (Sonnante et al. 1997; Oujii et al. 2011) and SSR markers (Oliveira et al. 2016). Similar results have been seen in other legumes such as soybean (*Glycine max*) (Jin et al. 2006) and *Lathyrus* (Belaid et al. 2006), as well as in olive (Belaj et al. 2007), apple (Zhang et al. 2012) and lettuce (Rauscher and Simko 2013). The high level of genetic variability observed within populations of such species is most likely due to their partial or completely allogamous reproductive systems. Indeed, cross-pollination of faba bean is known to be facilitated by insect pollinators like bees (Bond and Kirby 1999).

When comparing geographical genetic diversity among faba bean accessions, regional aggregation was observed (Fig. 1). Most of the Turkish varieties supplied by AARI showed clustering with each other and with Australian accessions. These results suggest that the Australian accessions may have originated in Turkey. Individuals originating from the Netherlands also grouped together. In contrast, accessions from Finland did not cluster tightly, but did fall into the same main cluster, Group A. Ethiopian accessions also did not show tight clustering. Interestingly, accessions from Afghanistan were seen in both Groups A and B. In addition to regional clustering, the dendrogram indicated a general continental aggregation of the genotypes. Individuals originating from European countries including the Netherlands, UK, Germany, Belgium, Greece, Spain, and Finland were clustered in Group A. Individuals originating from Asian and African countries including Turkey, Afghanistan, Syria, Lebanon, Algeria, Ethiopia, Egypt, Jordan, and China were found in Group B. Accessions from Australia were also found in Group B, again suggesting that these individuals originated from Asia. Zeid et al. (2003) found no clear grouping of faba bean genotypes according to geographical origin; however, Alghamdi et al. (2012b) demonstrated association between molecular diversity and geographic origin. In agreement with our results, Kaur et al. (2014) observed clustering of Australian cultivars using SNP markers.

Seed size and clustering

The dendrogram results were analyzed to detect correlation between seed size and clustering. This type of analysis is essential because there are four subspecies (major, equina, minor, and paucijuga) of faba bean and this separation is based on origin and seed size. The pairwise dissimilarity of large-seeded accessions ranged between 0.18 and 0.57, with an average of 0.38 (data not shown). Medium-seeded individuals had minimum and maximum pairwise dissimilarity values of 0.21 and 0.58, respectively. The average

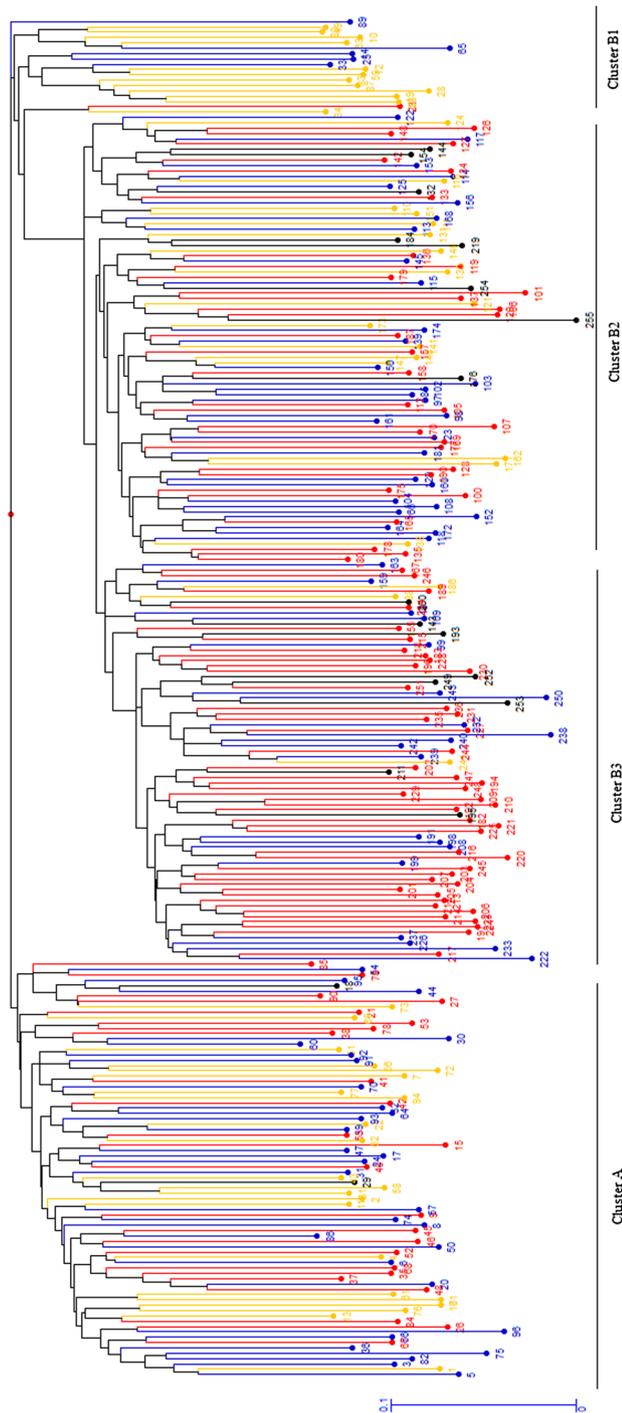


Fig. 3 Genetic diversity of the faba bean genotypes with seed size highlighted. Red, blue and yellow represent large, medium and small-seeded genotypes, respectively

pairwise dissimilarity of medium-seeded individuals was 0.39. Accessions of small-seeded individuals had values that ranged between 0.22 and 0.58, with an average of 0.40. Cluster A contained mostly small and medium-seeded individuals which is typical for individuals originating from

Northern Europe, Ethiopia, Afghanistan, and the Middle East (Fig. 3). In the same way, Cluster B1 contained mostly small-seeded accessions from Ethiopia and Afghanistan. Cluster B3 was composed of medium and large-seeded individuals originating from Turkey and Australia. Large-seeded types are typical of Turkey and other Mediterranean countries. The fact that the Australian genotypes fell into this cluster is not surprising because those genotypes are all cultivars, therefore, they may have been bred for large seeds. In comparison, Cluster B2 was composed of genotypes of all seed sizes and the genotypes had diverse geographical origins. Zeid et al. (2003) did not observe a correlation between faba bean seed size and clustering. Although there was a general correlation between seed size and origin in our study, definite conclusions cannot be made as the subspecies identities of the faba bean genotypes used in this work are not known.

Genetic dissimilarity of landraces versus cultivars

Genetic diversity analysis was conducted for the landraces (151 individuals) and cultivars (104 individuals), separately. Within the landraces, minimum GD was detected between genotypes 178 (TR33517) and 180 (TR33561) (0.18). The maximum genetic dissimilarity was 0.58 (58%) between genotypes 162 (TR23018) and 96 (CGN7948). The mean diversity (GD) for the landraces was 0.38. Within cultivars, minimum genetic dissimilarity was observed between genotypes 241 (FIORD) and 239 (NURA) (0.17). The maximum was between genotypes 5 (Mikko) and 253 (ILB93812) (0.59). The mean GD for the cultivars was 0.40. The diversity values (h) for landraces and cultivars gave a similar result indicating that cultivars were slightly, but not significantly more diverse than landraces. Shannon's information index for landraces and cultivars ranged from 0.30 to 0.36, respectively (Table 5).

In general, it is believed that cultivars have less genetic diversity than landraces; however, our results showed similar levels of diversity in each type of faba bean. Faba bean has a very large genome (13,000 Mb) and has not undergone extensive breeding efforts. In addition, although the crop is only partially allogamous, cross-pollination can be facilitated by bees (Aouar-Sadli et al. 2008). All of these factors may contribute to the maintenance of genetic diversity in cultivars as well as landraces.

Core set

A total of 44 genetically well-characterized faba bean individuals were selected for a core collection (Table 4). These accessions encompass the maximum genetic diversity present in all of the accessions used in this study (17.2% of the entire set). While core collections often include only 10%

Table 5 Analysis of genetic diversity for cultivars and landraces of faba bean accessions from worldwide geographical origins

Population	Number of individuals	Ne ^a	H ^a	I ^a	Number of private bands
Cultivars	104	1.36 ± 0.02	0.23 ± 0.01	0.36 ± 0.01	28
Landraces	151	1.29 ± 0.02	0.19 ± 0.01	0.30 ± 0.01	17

Ne no. of effective alleles = $1/(p^2 + q^2)$, I Shannon's information index = $-1 \times (p \times \ln(p) + q \times \ln(q))$, h diversity = $1 - (p^2 + q^2)$

^a Standard error

of the entire collection (Brown 1989), a larger core set was chosen because of high genetic diversity. Similar to this result, Turkish melon (*Cucumis melo*), sesame (*Sesamum indicum*) and opium poppy (*Papaver somniferum*) germplasm core sets represented similar proportions of the entire collections (10 to 20%) (Frery et al. 2013, 2015; unpublished data). The core collection should be further examined for yield, nutritional, disease, and drought resistance traits. This information can then be used with the genetic diversity data to develop new faba bean varieties.

Acknowledgments This study was supported by grant 0424.STZ.2013-2 from The Republic of Turkey's Ministry of Science, Industry and Technology with contributions from Polen Seed Co. We are grateful to Dr. Jeff Paull from The University of Adelaide and Dr. Fred Stoddard from University of Helsinki for providing faba bean seed samples.

Author contributions statement ŞG did laboratory experiments, performed data analysis and drafted manuscript; SD designed the study; AF designed the study and revised manuscript.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical standards statement This article does not contain any studies with human participants or animals performed by any of the authors.

References

Abuzayed M, El-Dabba N, Frery A, Doganlar S (2016) GDDom: an online tool for calculation of dominant marker gene diversity. *Biochem Genet* 55(2):155–157

Alghamdi SS, Migdadi HM, Ammar MH, Paull JG, Siddique KHM (2012a) Faba bean genomics: current status and future prospects. *Euphytica* 186(3):609–624

Alghamdi SS, Al-Faif SA, Migdadi HM, Khan MA, El-Harty EH, Ammar M (2012b) Molecular diversity assessment using sequence related amplified polymorphism (SRAP) markers in *Vicia Faba* L. *Int J Mol Sci* 13(12):16457–16471

Aouar-Sadli M, Louadi K, Doum SE (2008) Pollination of the broad bean (*Vicia faba* L. var. major) (Fabaceae) by wild bees and honey bees (Hymenoptera: Apoidea) and its impact on the seed production in the Tizi-Ouzou area (Algeria). *Afr J Agric Res* 3(4):266–272

Arbaoui M, Link W (2008) Effect of hardening on frost tolerance and fatty acid composition of leaves and stems of a set of faba bean (*Vicia faba* L.) genotypes. *Euphytica* 162(2):211–219

Belaid Y, Chtourou-Chorbel N, Marrakchi M, Trifi-Farah N (2006) Genetic diversity within and between populations of *Lathyrus genus* (Fabaceae) revealed by ISSR markers. *Genet Resour Crop Ev* 53:1413–1418

Belaj A, Muñoz-Diez C, Baldoni L, Porceddu A, Barranco D, Satovic Z (2007) Genetic diversity and population structure of wild olives from the north-western Mediterranean assessed by SSR markers. *Ann Bot* 100:449–458

Bond DD, Kirby EJ (1999) *Anthophora plumipes* (Hymenoptera: Anthophoridae) as a pollinator of broad bean (*Vicia faba* major). *J Apic Res* 38:199–204

Brown AHD (1989) Core Collections: a practical approach to genetic resources management. *Genome* 31(2):818–824

Caracuta V, Barzilai O, Khalailiy H, Milevski I, Paz Y, Vardi J, Regev L, Boaretto E (2015) The onset of faba bean farming in the Southern Levant. *Sci Rep* 5:14370

Celik I, Gultekin V, Allmer J, Doganlar S, Frery A (2014) Development of genomic simple sequence repeat markers in opium poppy by next-generation sequencing. *Mol Breeding* 34:323–334

Cubero JI (1974) Evolution of *Vicia faba* L. *Theor Appl Genet* 45(2):47–51

Cubero JI (2011) The faba bean: a historic perspective. *Grain Legumes* 56:5–7

Duc G (1997) Faba bean (*Vicia faba* L.). *Field Crop Res* 53(1–3):99–109

Earl DA, Von Holdt BM (2012) Structure harvester: a website and program for visualizing Structure output and implementing the Evanno method. *Conserv Genet Resour* 4:359–361

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. *Mol Ecol* 14(8):2611–2620

Frery A, Şiğva HÖ, Tan A, Taşkın T, İnal A, Mutlu S, Haytaoğlu M, Doğanlar S (2013) Molecular genetic diversity in the Turkish national melon collection and selection of a preliminary core set. *J Am Soc Hortic Sci* 138(1):50–56

Frery A, Tekin P, Celik I, Furat S, Uzun B, Doganlar S (2015) Morphological and molecular diversity in Turkish sesame germplasm and core set selection. *Crop Sci* 55(2):702–711

Gwak JG (2008) Molecular diversity assessment and population structure analysis in Mung bean, *Vigna radiata* (L.) Wilczek. Dissertation, Seoul National University

Jin Y, He TH, Lu BR (2006) Genetic spatial clustering: significant implications for conservation of wild soybean (*Glycine soja*: Fabaceae). *Genetica* 128:41–49

Johnston JS, Bennett MD, Rayburn AL, Galbraith DW, Price HJ (1999) Reference standards for determination of DNA content of plant nuclei. *Am J Bot* 86(5):609–613

Kaur S, Kimber RB, Cogan NO, Materne M, Forster JW, Paull JG (2014) SNP discovery and high-density genetic mapping in

- faba bean (*Vicia faba* L.) permits identification of QTLs for ascochyta blight resistance. *Plant Sci* 217:47–55
- Kim KW, Chung HK, Cho GT, Ma KH, Chandrabalan D, Gwag JG, Kim TS, Cho EG, Park YJ (2007) PowerCore: a program applying the advanced M strategy with a heuristic search for establishing allele mining sets. *Bioinformatics* 23:2155–2162
- Link W, Dixkens C, Singh M, Schwall M, Melchinger AE (1995) Genetic diversity in European and Mediterranean faba bean germ plasm revealed by RAPD markers. *Theor Appl Genet* 90(1):27–32
- Liu Y, Hou W (2010) Genetic diversity of faba bean germplasms in Qinghai and core germplasm identified based on AFLP analysis. *Legume Genomics and Genet* 1
- Ma KH, Kim NS, Lee GA, Lee SY, Lee JK, Yi JY, Park YJ, Kim TS, Gwag JG, Kwon SJ (2009) Development of SSR markers for studies of diversity in the genus *Fagopyrum*. *Theor Appl Genet* 119:1247–1254
- Oliveira HR, Tomás D, Silva M, Lopes S, Viegas W, Veloso MM (2016) Genetic diversity and population structure in *Vicia faba* L. landraces and wild related species assessed by nuclear SSRs. *PLoS ONE* 11(5):e0154801
- Ouji A, Suso MJ, Rouaïssi M, Abdellaoui R, El Gazzah M (2011) Genetic diversity of nine faba bean (*Vicia faba* L.) populations revealed by isozyme markers. *Genes Genom* 33:31–38
- Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* 28:2537–2539
- Perrier X, Jacquemoud-Collet JP (2006) DARwin software. <http://darwin.cirad.fr/>. Accessed 25 October 2016
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945–959
- Rauscher G, Simko I (2013) Development of genomic SSR markers for fingerprinting lettuce (*Lactuca sativa* L.) cultivars and mapping genes. *BMC Plant Biol* 13:1–11
- Ray H, Georges F (2010) A genomic approach to nutritional, pharmacological and genetic issues of faba bean (*Vicia faba*): prospects for genetic modifications. *GM Crops* 1(2):99–106
- Rispail N, Kaló P, Kiss GB, Ellis THN, Gallardo K, Thompson RD, Prats E, Larrainzar E, Ladrera R, González EM, Arrese-Igor C, Ferguson BJ, Gresshoff PM, Rubiales D (2010) Model legumes contribute to faba bean breeding. *Field Crop Res* 115(3):253–269
- Roldán-Ruiz I, Dendauw J, Van Bockstaele E, Depicker A, De Loose M (2000) AFLP markers reveal high polymorphic rates in ryegrasses (*Lolium* spp.). *Mol Breeding* 6(2):125–134
- Sallam A, Arbaoui M, El-Esawi M, Abshire N, Martsch R (2016a) Identification and verification of QTL associated with frost tolerance using linkage mapping and GWAS in winter faba bean. *Front Plant Sci*. doi:10.3389/fpls.2016.01098
- Sallam A, Dhanapal AP, Liu S (2016b) Association mapping of winter hardiness and yield traits in faba bean (*Vicia faba* L.). *Crop Pasture Sci* 67(1):55–68
- Selale H, Celik I, Gultekin V, Allmer J, Doganlar S, Frary A (2013) Development of EST-SSR markers for diversity and breeding studies in opium poppy (*Papaver somniferum* L.). *Plant Breed* 132:344–351
- Sonnante G, Piergiovanni AR, Pignone D (1997) Distribution of allozyme variation in *Vicia benghalensis* (Leguminosae) suggests the existence of two gene pools. *Plant Syst Evol* 207:99–110
- Suresh S, Park JH, Cho GT, Lee HS, Baek HJ, Lee SY, Chung JW (2013) Development and molecular characterization of 55 novel polymorphic cDNA-SSR markers in faba bean (*Vicia faba* L.) using 454 pyrosequencing. *Molecules* 18(2):1844–1856
- Terzopoulos PJ, Bebeli PJ (2008) Genetic diversity analysis of Mediterranean faba bean (*Vicia faba* L.) with ISSR markers. *Field Crop Res* 108(1):39–44
- Torres AM, Weeden NF, Martin A (1993) Linkage among isozyme, RFLP and RAPD markers in *Vicia faba*. *Theor Appl Genet* 85(8):937–945
- Torres AM, Avila CM, Gutierrez N, Palomino C, Moreno MT, Cubero JI (2010) Marker assisted selection in faba bean (*Vicia faba* L.). *Field Crop Res* 115(3):243–252
- Upadhyaya HD, Yadav D, Dronavalli N, Gowda CLL, Singh S (2010) Mini core germplasm collections for infusing genetic diversity in plant breeding programs. *Electron J Plant Breed* 1(4):1294–1309
- Van Hintum TJJ, Brown AHD, Spillane C, Hodgkin T (2000) Core collections of plant genetic resources. International Plant Genetic Resources Institute, Rome
- Varshney RK, Graner A, Sorrells ME (2005) Genic microsatellite markers in plants: features and applications. *Trends Biotech* 23:48–55
- Wang HF, Zong XX, Guan JP, Yang T, Sun XL, Ma Y, Redden R (2012) Genetic diversity and relationship of global faba bean (*Vicia faba* L.) germplasm revealed by ISSR markers. *Theor Appl Genet* 124(5):789–797
- Zeid M, Schön CC, Link W (2003) Genetic diversity in recent elite faba bean lines using AFLP markers. *Theor Appl Genet* 107(7):1304–1314
- Zeid M, Mitchell S, Link W, Carter M, Nawar A, Fulton T, Kresovich S (2009) Simple sequence repeats (SSRs) in faba bean: new loci from Orobanche resistant cultivar “Giza 402”. *Plant Breeding* 128(2):149–155
- Zhang Q, Li J, Zhao Y, Korban SS, Han Y (2012) Evaluation of genetic diversity in Chinese wild apple species along with apple cultivars using SSR markers. *Plant Mol Biol Rep* 30:539–546