Genetic Diversity and Population Structure of Faba Bean (*Vicia faba*) from Turkey

Şurhan Göl, Sami Doğanlar, and Anne Frary[1]

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

Additional index words. broad bean, core set, fava bean, genetic diversity, microsatellites, SSRs, Turkey, *Vicia faba*

Abstract. Faba bean (*Vicia faba* L.) is an important food and feed legume because of the nutritional value of its seed protein and starch content, good biomass, and high efficiency nitrogen fixation. This study analyzed the molecular genetic diversity and population structure of 101 Turkish faba bean accessions using simple sequence repeat (SSR) markers. A total of 32 SSR markers yielded 281 loci of which 277 (98.6%) were polymorphic. Mean dissimilarity between genotypes was 0.355, a relatively high value which was expected given faba bean's large genome and limited breeding history. According to the results, faba bean individuals grouped into three main clusters based on both distance matrix (neighbor-joining algorithm) and model-based (population structure) clustering analyses. Clustering was not correlated with seed size or origin within Turkey. Six cultivars were included in the analysis and showed high genetic diversity compared with the landraces, as expected, given the fact that at least some of the cultivars were bred by hybridization. A total of 47 individuals were selected for the core collection to represent the diversity of Turkish faba bean germplasm. This core set encompasses material adapted to all growing regions and should be a priority for morphological characterization.

Faba bean (*Vicia faba*) is an important nutritional source for human food and animal feed (Duc, 1997). Faba bean seeds are rich in protein, starch, cellulose, and minerals (Hacıseferoğulları et al., 2003) and are consumed in large quantities in developing countries such as China, Turkey, Egypt, Ethiopia, and Central America (Łabuda, 2012). *V. faba* is not only a good alternative to animal proteins, but also an attractive product because of its low cost, long storage life, and easy transportation (Hacıseferoğulları et al., 2003). Moreover, it is a perfect rotational crop because the nitrogen fixing bacteria in its roots help to enhance soil productivity. Indeed, faba bean is one of the most efficient temperate legumes in terms of nitrogen fixation (Phillips, 1980). Faba bean is the fourth most widely grown pulse crop in Turkey (Pekşen et al., 2006).

Although its wild progenitor is unknown, *V. faba* seems to have been first domesticated in the Levant where archaeological evidence of its cultivation dates to the 10th millennium BP (Caracuta et al., 2015). Faba bean cultivation spread to Anatolia and then Europe via the Mediterranean coast and to India and China via Mesopotamia (Cubero, 2011). The crop was later introduced to Latin America by the Spanish. During its dispersal, faba bean gained specific adaptations to different environments and uses. These adaptations are reflected in plant architecture and seed size, weight, and shape (Alghamdi et al., 2012). Accordingly, faba bean is separated into four groups based on seed size: *major*, *equina*, *minor*, and *paucijuga* (Cubero, 1974). The largest seeded type (*major*) is generally found in south Mediterranean countries and China. The medium-sized type (*equina*) emerged in the Middle East, north Africa, and Australia, whereas small-seeded types (*minor* and *paucijuga*) are prevalent in Ethiopia and northern Europe (Duc, 1997).

Crop germplasm provides the raw material to produce improved cultivars to meet the world's future food and climate requirements. Germplasm and landrace diversity has arisen due to soil, climatic, cropping, and use differences among regions and farmers (Ernesto, 1992). However, genetically variable traditional varieties are being replaced by a few uniform cultivars which pose a threat to intraspecies diversity and future crop sustainability. Fortunately, faba bean germplasm has been stored in seed banks with more than 38,000 accessions represented in collections worldwide (Duc et al., 2010). This germplasm is a wealth of material for future breeding efforts. Ancient faba bean remains have been found in northwest Syria and Turkey (Tanno and Willcox, 2006); therefore, this region is expected to house highly diverse germplasm. Turkish material has been collected by the Aegean Agricultural Research Institute

(AARI) in Izmir, Turkey; however, it has not yet been fully characterized for its morphological properties and molecular diversity. Thus, it is very difficult for breeders to know how they can take advantage of these genetic resources. To address this problem, the genetic diversity and population structure of 101 Turkish faba bean accessions were analyzed using SSR markers. In addition, a core collection was selected based on genetic diversity.

Materials and Methods

Plant material and DNA isolation. A total of 95 landraces and 6 cultivars were used as plant material (Supplemental Table 1). The majority of the accessions (88) were obtained from AARI. These accessions included 84 landraces and 4 Turkish cultivars: Kıtık2003 (individual 98), Eresen87 (individual 99), Filiz99 (individual 100), and Salkım (individual 101) and represented all of AARI's publicly available material. The landraces were collected throughout Turkey and were mostly large-seeded (71%) with fewer medium (22%) and very few small-seeded (5%) types (Supplemental Table 1). The remaining 13 accessions were obtained from the Center for Genetic Resources (CGN), Wageningen, The Netherlands and included two cultivars: Turkeye (individual 2) and Haba Alemeda (individual 9).

Five seeds from each accession were planted and grown in soil containing perlite and natural fertilizer in the greenhouse at Izmir Institute of Technology, Izmir, Turkey. Genomic DNA was extracted from the youngest and the lightest green leaves using a bulk harvest of the five individuals of each accession. Extraction was done using the Wizard Magnetic 96 Plant System (Promega Corp., Madison, WI) with the Biomek NX Workstation (Beckman Coulter, Brea, CA) according to the manufacturer's instructions.

SSR analysis. A total of 32 SSR primer pairs [GBSSR-VF (Suresh et al., 2013) and VfG (Zeid et al., 2009)] were selected for analysis on the Turkish faba bean population. These primers were designed from genic and genomic regions of the faba bean genome, respectively. Polymerase chain reaction (PCR) for GBSSR-VF primers was performed with the following components in a total volume of 20:2 µL 10× PCR buffer (50 mM KCl, 10 mM tris-HCl, 1.5 mM MgCl$_2$, and 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 (≈10 ng/µL). To amplify DNA, PCR conditions were optimized as follows: one step of 10 min at 94 °C for denaturation, 35 cycles of 30 s at 95 °C, 30 s at 55–57 °C annealing temperature (depending on primer pair), 30 s at 72 °C for extension, and a final extension step of 10 min at 72 °C. PCR for VfG primers was performed in a total volume of 20:2 µL 10× PCR buffer, 1.2 µL MgCl$_2$ (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). VfG primer PCR conditions were optimized as follows: one step of 4 min at 94 °C for denaturation, 30 cycles with 45 s at 95 °C, 1 min at 50–53 °C annealing (depending on primer pair), 1 min at 72 °C for extension, and a final extension step of 7 min at 72 °C. Amplification products were then separated on a Fragment Analyzer (Advanced Analytical, Ankeny, IA) capillary electrophoresis system using the DNF-900 dsDNA Reagent Kit (Advanced Analytical) according to the manufacturer's instructions. SSR alleles were visualized and scored using PROSize 2.0 software version 1.2.1.1 (Advanced Analytical).

Diversity and population structure analysis. Allelic data obtained from fragment analysis were scored according to band presence/absence. Average gene diversity (GD) (Roldán-Ruiz et al., 2000) was calculated for each of the 32 markers using an in-house program (http://plantmolgen.iyte.edu.tr/GDdom/). The allelic data were also used for calculation of a genetic diversity distance matrix using Dice's coefficient and clustering analysis was done with the unweighted neighbor-joining algorithm. These analyses were done with the DARwin5 (Dissimilarity Analysis and Representation for Windows) software program (http://darwin.cirad.fr/product.php).

To detect the correct number of subpopulations explaining population structure, Structure 2.2.3 was used as a clustering program (Pritchard et al., 2000). The analysis was run with burn-in periods of 100,000 and 500,000 MCMC (Markov Chain Monte Carlo) replications and the ad hoc statistic (Evanno et al., 2005) was used to determine the correct estimated number of clusters with Structure Harvester online program (Earl and Von Holdt, 2012). Subpopulation number (K) was tested from 2 to 10 with 20 iterations for each group. Subpopulations were determined according to a threshold of ≥0.70 inferred ancestries. Genotypes with an identity value under this threshold were considered to be intermixed.

Core collection. PowerCore (v.1.0) was used for selection of a core collection (Kim et al., 2007). This program uses a heuristic search for establishing a core set based on the SSR marker data.

Results and Discussion

SSR analysis. Allelic polymorphisms were determined using 32 SSR primer pairs for the 101 Turkish faba bean accessions. All markers except GBSSR-VF-20, GBSSR-VF-113, VfG 31, and VfG 34 were polymorphic for the population. Markers VfG1, VfG41, and VfG53 had the most polymorphic bands with 15 amplified fragments for each marker. Markers GBSSR-VF-159 and GBSSR-VF-164 had the fewest polymorphic fragments, only two. Overall, the markers produced 281 loci of which 277 (98.6%) were polymorphic for the 101 individuals. Thus, each SSR

primer combination gave an average of nine polymorphic fragments. As calculated using a scale of 0 to 0.5, the average gene diversity value was 0.20. Average GD was 0.22 for the GBSSR-VF primers and 0.19 for VfG

primers. Thus, there was no significant difference in GD results for genic vs. genomic SSR markers. The highest GD was 0.39 for GBSSR-VF-153, whereas the lowest value was 0.02 for GBSSR-VF-159 (Supplemental
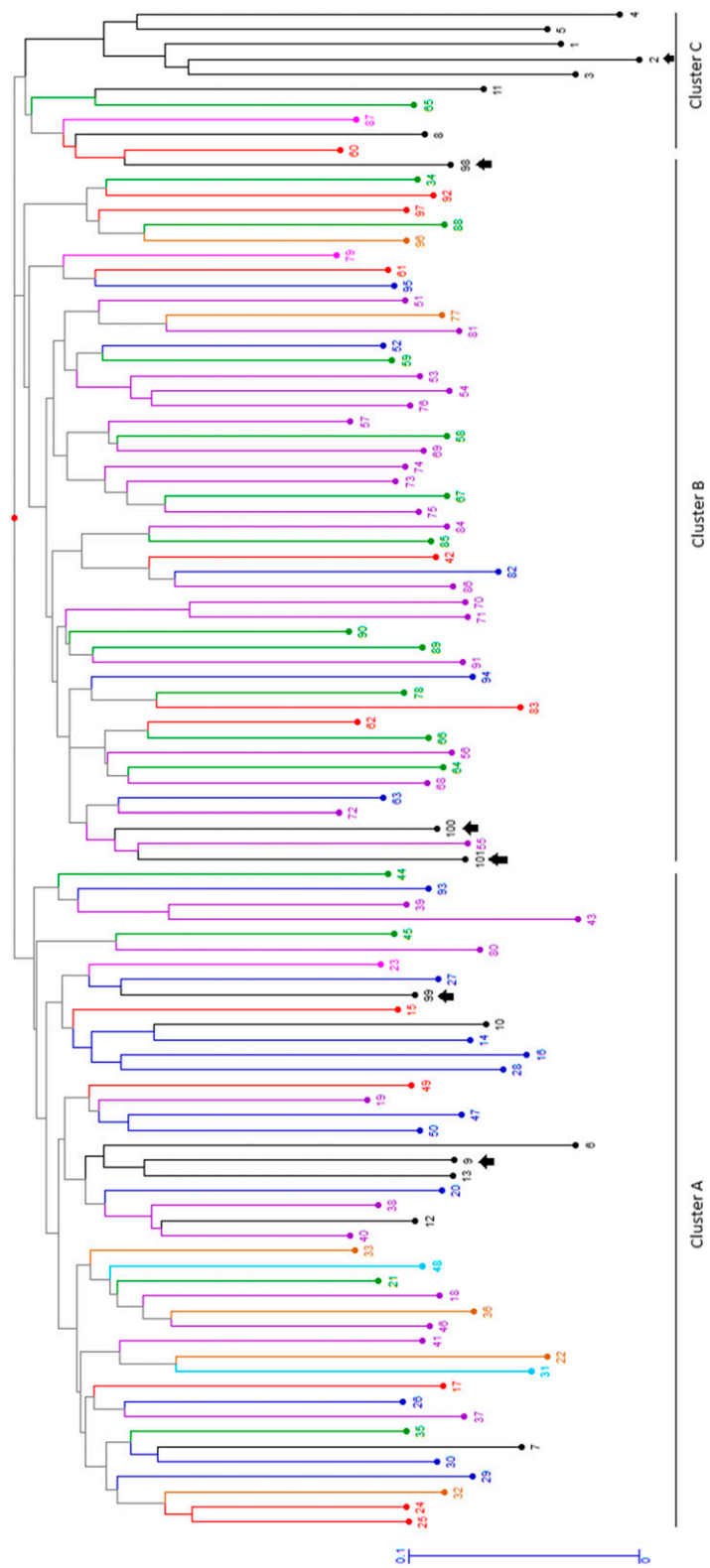

Fig. 1. Dendrogram showing genetic diversity of the faba bean genotypes. Dark blue, green, pink, brown, purple, red, and light blue represent Aegean, Black Sea, Central Anatolia, Eastern Anatolia, Marmara, Mediterranean, and southeastern regions, respectively. All other genotypes are in black. Positions of cultivars are indicated with arrows.

Table 2). Previous work reported an average GD of 0.69 for GBSSR-VF primers (Suresh et al., 2013) and 0.60 for VfG primers (Zeid et al., 2009) using a scale of 0 to 1. Although the Turkish germplasm values are lower, it is important to note that the previous work used only 32 and 10 accessions, respectively, and that many of these were collected from diverse locations throughout the world.

Diversity and population structure analysis. The 101 individuals were analyzed for their genetic diversity and a dendrogram was drawn using the Dice coefficient and unweighted neighbor-joining algorithm (Fig. 1). The maximum genetic dissimilarity between genotypes was 0.541 (54%) for genotypes 7 (CGN10362) and 2 (CGN19987). The minimum genetic dissimilarity was 0.184 (18%) for genotypes 40 (TR33561) and 38 (TR33517). The 101 units gave 5050 distance values with a mean of 0.355. Thus, the Turkish faba bean germplasm is quite diverse with very few genotypes that are highly similar. These results were expected because faba bean has a very large genome. In addition, it has been subjected to limited breeding efforts because it is not an economically important crop in many countries. Moreover, faba bean is at least partially cross-pollinated (Gnanasambandam et al., 2012), a mating strategy that maintains higher polymorphism than strict self-pollination.

A Mantel test gave a high correlation ($r = 0.90$) between the dissimilarity matrix and dendrogram. According to the unweighted neighbor-joining analysis, three clusters were obtained. Cluster A had 44 individuals, Cluster B had 46 individuals, and Cluster C had 11 individuals (Supplemental Table 1; Fig. 1). The accessions did not cluster according to seed size (data not shown) indicating that there is no genetic basis for the classification of faba bean by size (i.e., *major*, *equina*, *minor*, and *paucijuga*). Faba bean individuals originating from different locations of Turkey also did not show tight clustering according to geographic region (Fig. 1). Previous studies with different types of molecular markers revealed positive correlations between faba bean molecular diversity and geographical origin (Link et al., 1995; Terzopoulos and Bebeli, 2008; Zeid et al., 2003; Zong et al., 2009). However, these analyses used accessions from broader geographic distributions and did not limit themselves to a single country's germplasm. Average GD was calculated for each region that was represented by at least 10 accessions/cultivars (Table 1). The Marmara region had maximum average GD; however, this value was not significantly higher than those obtained for other regions. Compared with the landraces from different regions, cultivars had high GD and average pairwise dissimilarity. Turkeye (individual 2), Haba Alemeda (individual 9), Kıtık2003 (individual 98), Eresen87 (individual 99), Filiz99 (individual 100), and Salkım (individual 101) belonged to clusters C, A, C, A, B, and B, respectively (Fig. 1). Thus, the cultivars were a good sample of the diversity present in Turkish faba bean germplasm. This also suggested that, at least in these six cultivars, the gene pool has not yet been narrowed by breeding.

According to population structure analysis, the faba bean accessions fell into three subpopulations (Fig. 2) with a relatively high ΔK value at $K = 3$ (Supplemental Fig. 1). The SD for each value of K is also important when deciding the correct number of subpopulations and these values supported the choice of $K = 3$ (Supplemental Fig. 1). Subpopulation 1 had 33 individuals, subpopulation 2 had 34 individuals, subpopulation 3 had 6 individuals, and the intermixed group had 28 individuals (Supplemental Table 1). There was a 94% overlap of the accessions belonging to subpopulation 1 and cluster A of the dendrogram, a 94% overlap in subpopulation 2 and cluster B, and an 83% overlap between subpopulation 3 and cluster C. These results indicate a high level of concordance between the results of the two clustering methods which use different algorithms. Interestingly, all four of the Turkish cultivars were admixed. In general, these results agree with their pedigrees as three of the four cultivars (Filiz99, Kitik2003, and Salkim) were developed by hybridization between faba bean genotypes including material from the International Center for Agricultural Research in the Dry Areas.

Core collection. A total of 47 individuals were selected for the core collection (Supplemental Table 3). These individuals represent the maximum genetic diversity present in all of the accessions used in this study in the minimum number of genotypes (46.5% of the entire set). Although core collections are often chosen to include 10% of the entire collection (Brown, 1989), the PowerCore program uses an algorithm that selects the set with highest marker diversity with no limitation on size of the core set (Kim et al., 2007). Thus, the relatively large size of the core set was due to the high genetic diversity of the entire collection. In contrast, the program selected much smaller core sets (10% to 20%, respectively) from similar-sized collections of Turkish melon (*Cucumis melo*), sesame (*Sesamum indicum*), and opium poppy (*Papaver somniferum*), germplasm which had less overall molecular genetic diversity (Frary et al., 2013, 2015; unpublished data).

The core set was analyzed for its genetic diversity using the Dice's coefficient and unweighted neighbor-joining algorithm. The maximum genetic dissimilarity between genotypes was 0.54 (54%) for genotypes 3 (CGN10321) and 6 (CGN10384). The minimum genetic dissimilarity was 0.21 (21%) for genotypes 63 (TR49386) and 72 (TR53748). The 47 units gave 861 distance values with a mean value of 0.37. We also analyzed GD for the core set. Maximum GD was determined as 0.41, whereas minimum GD was determined as 0.24 (Table 1). Average GD and pairwise dissimilarity of the core set were slightly higher than for all accessions, cultivars only, and the regional groupings. Thus, it was clear that the core set preserved the genetic diversity present in the Turkish faba bean germplasm collection. In addition, the core set contained genotypes collected from every region in Turkey where faba bean is grown. Such coverage is essential because Turkey has diverse climatic regions and adaptation to growing region is an important consideration in breeding. Although origin was not used as a criterion for core selection, genotypes from the Marmara and Aegean regions were most prevalent (30% and 15% of the core, respectively) and these are the areas where most of Turkey's faba bean production occurs (Karbuz et al., 2008). Individuals in the core collection should have priority for morphological characterization and for use in breeding studies to develop new varieties.

Table 1. Average genetic diversity and pairwise dissimilarity by region, for cultivars and for total and core sets of germplasm.

| Region/set	Individual no.	Avg GD[z]	Minimum–maximum GD	Avg pairwise dissimilarity	Minimum–maximum pairwise dissimilarity
Aegean	16	0.17 ± 0.05	0–0.40	0.33	0.26–0.43
Black Sea	16	0.17 ± 0.05	0–0.45	0.33	0.26–0.43
Marmara	29	0.19 ± 0.05	0–0.46	0.34	0.18–0.53
Mediterranean	12	0.16 ± 0.06	0–0.35	0.33	0.19–0.45
Cultivars	6	0.20 ± 0.06	0–0.37	0.39	0.29–0.53
Total	101	0.20 ± 0.05	0–0.39	0.36	0.18–0.54
Core set	47	0.22 ± 0.05	0.04–0.37	0.37	0.21–0.54

GD = gene diversity.
[z]For each section, average GD ± SE is presented.

Fig. 2. The bar plot for three subpopulations. Each genotype is represented by a vertical bar along the *x* axis with proportion of membership in each subpopulation given in the *y* axis. Dark grey, light grey, and black bars represent subpopulations 1, 2, and 3, respectively.

Literature Cited

Alghamdi, S.S., H.M. Migdadi, M.H. Ammar, J.G. Paull, and K.H.M. Siddique. 2012. Faba bean genomics: Current status and future prospects. Euphytica 186(3):609–624.

Brown, A.H.D. 1989. Core collections - A practical approach to genetic resources management. Genome/National Research Council Canada 31:818–824.

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

Cubero, J.I. 1974. Evolution of *Vicia Faba* L. Theor. Appl. Genet. 45(2):47–51.

Cubero, J.I. 2011. The faba bean: A historic perspective. Grain Legum. 56:5–7.

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

Duc, G., S. Bao, M. Baum, B. Redden, M. Sadiki, M.J. Suso, M. Vishniakova, and X. Zong. 2010. Diversity maintenance and use of *Vicia faba* L. genetic resources. Field. Crop. Res. 115(3):270–278.

Earl, D.A. and B. vonHoldt. 2012. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv. Genet. Resources 4:359–361.

Ernesto, P. 1992. Maize breeding in the tropics. Crit. Rev. Plant Sci. 9(2):125–154.

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

Gnanasambandam, A., J. Paull, A. Torres, S. Kaur, T. Leonforte, H. Li, X. Zong, T. Yang, and M. Materne. 2012. Impact of molecular technologies on faba bean (*Vicia faba* L.) breeding strategies. Agronomy 2(3):132–166.

Hacıseferoğulları, H., İ. Gezer, Y. Bahtiyarca, and H.O. Mengeş. 2003. Determination of some chemical and physical properties of Sakız faba bean (*Vicia faba* L. Var. *major*). J. Food Eng. 60:475–479.

Karbuz, F., İ. Öztürk, and D.O. Savaş. 2008. Türkiye'de üretilen tarım ürünleri ve ekonomideki yeri. 4 Jan. 2016. <http://www.ito.org.tr/Dokuman/Sektor/1-99.pdf>.

Kim, K.W., H.K. Chung, G.T. Cho, K.H. Ma, D. Chandrabalan, J.G. Gwag, T.S. Kim, E.G. Cho, and Y.J. Park. 2007. PowerCore: A program applying the advanced M strategy with a heuristic search for establishing allele mining sets. Bioinformatics 23:2155–2162.

Link, W.C., M. Dixkens, M. Singh, and A.E.M. Schwall. 1995. Genetic diversity in European and Mediterranean faba bean germ plasm revealed by RAPD markers. Theor. Appl. Genet. 90(1):27–32.

Pekşen, A., E. Pekşen, and C. Artık. 2006. Bazı bakla (*Vicia faba* L.) populasyonlarının bitkisel özellikleri ve taze bakla verimlerinin belirlenmesi. J. Fac. Agr. 21(2):225–230.

Phillips, D.A. 1980. Efficiency of symbiotic nitrogen-fixation in legumes. Annu. Rev. Plant Physiol. Plant Mol. Biol. 31:29–49.

Pritchard, J.K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945–959.

Roldán-Ruiz, I., J. Dendauw, E. Van Bockstaele, A. Depicker, and M. De Loose. 2000. AFLP markers reveal high polymorphic rates in ryegrasses (*Lolium* spp.). Mol. Breed. 6(2):125–134.

Suresh, S., J.H. Park, G.T. Cho, H.S. Lee, H.J. Baek, S.Y. Lee, and J.W. Chung. 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.

Tanno, K. and G. Willcox. 2006. The origins of cultivation of *Cicer arietinum* L. and *Vicia faba* L.: Early finds from Tell el-Kerkh, north-west Syria, late 10th millennium b.p. Veg. Hist. Archaeobot. 15(3):197–204.

Terzopoulos, P.J. and P.J. Bebeli. 2008. Genetic diversity analysis of Mediterranean faba bean (*Vicia faba* L.) with ISSR markers. Field Crops Res. 108(1):39–44.

Zeid, M., S. Mitchell, W. Link, M. Carter, A. Nawar, T. Fulton, and S. Kresovich. 2009. Simple sequence repeats (SSRs) in faba bean: New loci from Orobanche-resistant cultivar 'Giza 402.' Plant Breed. 128(2):149–155.

Zeid, M., C.-C. Schön, and W. Link. 2003. Genetic diversity in recent elite faba bean lines using AFLP markers. Theor. Appl. Genet. 107(7):1304–1314.

Zong, X., X. Liu, J. Guan, S. Wang, Q. Liu, J.G. Paull, and R. Redden. 2009. Molecular variation among Chinese and global winter faba bean germplasm. Theor. Appl. Genet. 118(5):971–978.

$$DeltaK = mean(|L''(K)|) / sd(L(K))$$
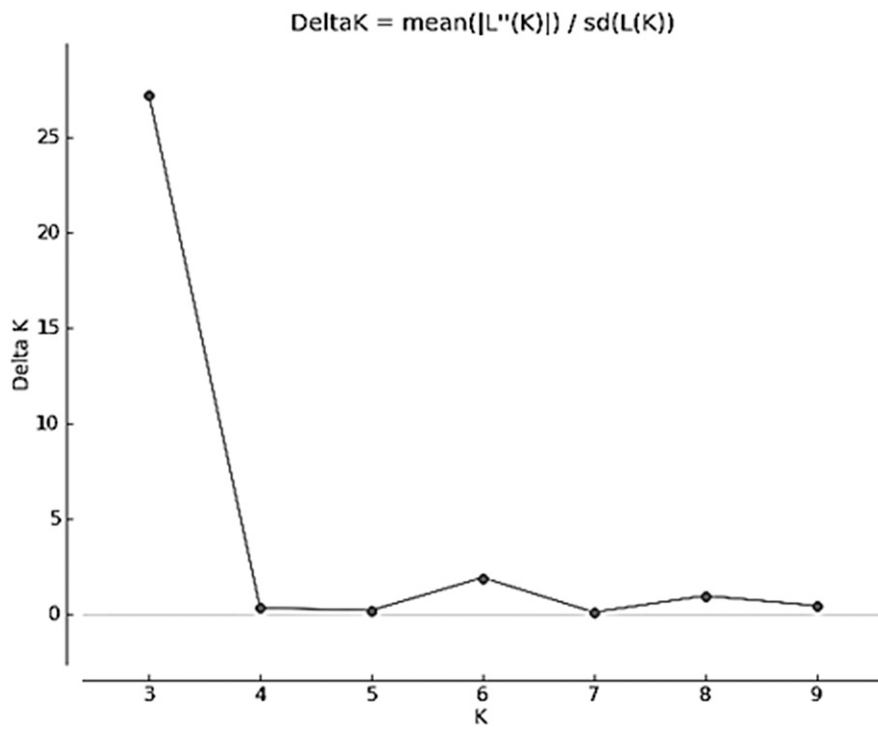
Supplemental Fig. 1. Standard deviation values for each number of subpopulations.

Supplemental Table 1. Plant individuals (*Vicia faba*) used in this study.

Genotype (accession)	Individual types	Seed size	Collection location	Subpopulation assignment[z]	Cluster assignment[y]
1 (CGN10322)	Landrace	Large	Turkey	3	C
2 (CGN19987)	Cultivar	Large	Turkey	3	C
3 (CGN10321)	Landrace	Large	Turkey	3	C
4 (CGN10320)	Landrace	Large	Turkey	3	C
5 (CGN10382)	Landrace	Medium	Turkey	3	C
6 (CGN10384)	Landrace	Large	Turkey	Admixed	A
7 (CGN10362)	Landrace	Large	Turkey	1	A
8 (CGN18888)	Landrace	Medium	Turkey	1	C
9 (CGN19987)	Cultivar	Large	Turkey	1	A
10 (CGN10387)	Landrace	Large	Turkey	Admixed	A
11 (CGN10385)	Landrace	Large	Turkey	Admixed	C
12 (CGN10383)	Landrace	Large	Turkey	1	A
13 (CGN10386)	Landrace	Large	Turkey	1	A
14 (TR12123)	Landrace	nd	Aydın	Admixed	A
15 (TR12271)	Landrace	Large	Antalya	Admixed	A
16 (TR12540)	Landrace	Medium	Aydın	3	A
17 (TR12611)	Landrace	Large	İçel	1	A
18 (TR12716)	Landrace	Large	Balıkesir	1	A
19 (TR12767)	Landrace	Medium	Balıkesir	1	A
20 (TR12773)	Landrace	Medium	Manisa	1	A
21 (TR22992)	Landrace	Medium	Giresun	1	A
22 (TR23018)	Landrace	Small	Elazığ	1	A
23 (TR26194)	Landrace	Medium	Konya	1	A
24 (TR26250)	Landrace	Medium	Adana	1	A
25 (TR26265)	Landrace	Large	Hatay	1	A
26 (TR26298)	Landrace	Medium	Muğla	1	A
27 (TR26453)	Landrace	Large	Manisa	Admixed	A
28 (TR26462)	Landrace	Medium	İzmir	1	A
29 (TR26578)	Landrace	Large	İzmir	Admixed	A
30 (TR28096)	Landrace	Large	Muğla	1	A
31 (TR31590)	Landrace	Small	Diyarbakır	Admixed	A
32 (TR31912)	Landrace	Medium	Elazığ	1	A
33 (TR31973)	Landrace	Small	Elazığ	1	A
34 (TR32925)	Landrace	Medium	Artvin	1	B
35 (TR33058)	Landrace	Large	Tokat	1	A
36 (TR33140)	Landrace	nd	Kars	1	A
37 (TR33421)	Landrace	Large	Edirne	1	A
38 (TR33517)	Landrace	Large	Kırklareli	1	A
39 (TR33547)	Landrace	Large	İstanbul	1	A
40 (TR33561)	Landrace	Large	İstanbul	1	A
41 (TR35023)	Landrace	Medium	Bursa	1	A
42 (TR35330)	Landrace	Large	Kahramanmaraş	Admixed	B
43 (TR37041)	Landrace	Large	İstanbul	Admixed	A
44 (TR37200)	Landrace	Large	Sinop	Admixed	A
45 (TR37255)	Landrace	nd	Kastamonu	1	A
46 (TR38142)	Landrace	Medium	Çanakkale	1	A
47 (TR39064)	Landrace	Small	Muğla	1	A
48 (TR40217)	Landrace	Large	Şanlıurfa	1	A
49 (TR40725)	Landrace	Small	İçel	1	A
50 (TR42343)	Landrace	Large	Kütahya	1	A
51 (TR43564)	Landrace	Large	Sakarya	Admixed	B
52 (TR44862)	Landrace	Medium	Manisa	2	B
53 (TR44869)	Landrace	Large	Balıkesir	Admixed	B
54 (TR44876)	Landrace	nd	Balıkesir	Admixed	B
55 (TR44914)	Landrace	Large	Adapazarı	2	B
56 (TR44928)	Landrace	nd	İzmit	Admixed	B
57 (TR44931)	Landrace	Large	İzmit	2	B
58 (TR44941)	Landrace	Large	Bolu	2	B
59 (TR46010)	Landrace	Medium	Trabzon	2	B
60 (TR49377)	Landrace	Medium	Antalya	2	C
61 (TR49380)	Landrace	nd	Hatay	Admixed	B
62 (TR49381)	Landrace	Large	Hatay	2	B
63 (TR49386)	Landrace	Large	İzmir	2	B
64 (TR49387)	Landrace	Large	Ordu	2	B
65 (TR49388)	Landrace	Large	Giresun	Admixed	C
66 (TR51377)	Landrace	Large	Kastamonu	2	B
67 (TR51378)	Landrace	Large	Tokat	2	B
68 (TR53648)	Landrace	Large	Çanakkale	2	B
69 (TR53667)	Landrace	Medium	Çanakkale	2	B
70 (TR53683)	Landrace	Large	Balıkesir	Admixed	B
71 (TR53724)	Landrace	Large	Balıkesir	2	B
72 (TR53748)	Landrace	nd	Tekirdağ	2	B
73 (TR53761)	Landrace	Large	Edirne	2	B
74 (TR53770)	Landrace	Large	Edirne	2	B

Supplemental Table 1. (*Continued*) Plant individuals (*Vicia faba*) used in this study.

Genotype (accession)	Individual types	Seed size	Collection location	Subpopulation assignment[z]	Cluster assignment[y]
75 (TR53781)	Landrace	Large	Tekirdağ	2	B
76 (TR53797)	Landrace	Large	Tekirdağ	Admixed	B
77 (TR53947)	Landrace	Large	Kars	2	B
78 (TR53948)	Landrace	Large	Kastamonu	2	B
79 (TR53949)	Landrace	Large	Kayseri	2	B
80 (TR61267)	Landrace	nd	Tekirdağ	Admixed	A
81 (TR61300)	Landrace	Large	Kırklareli	2	B
82 (TR71255)	Landrace	Large	Kütahya	2	B
83 (TR71521)	Landrace	Medium	Adana	2	B
84 (TR71612)	Landrace	Large	Bilecik	2	B
85 (TR74159)	Landrace	Large	Amasya	2	B
86 (TR74184)	Landrace	Large	Bursa	2	B
87 (TR74214)	Landrace	Medium	Karaman	Admixed	B
88 (TR74330)	Landrace	Large	Samsun	Admixed	B
89 (TR74337)	Landrace	Large	Trabzon	2	B
90 (TR75408)	Landrace	Large	Zonguldak	2	B
91 (TR75421)	Landrace	Large	Bilecik	2	B
92 (TR75431)	Landrace	Large	Burdur	2	B
93 (TR77139)	Landrace	Medium	Manisa	2	A
94 (TR77143)	Landrace	Medium	Muğla	2	B
95 (TR78048)	Landrace	Large	Kütahya	Admixed	B
96 (TR79569)	Landrace	Large	Erzincan	Admixed	B
97 (TR80277)	Landrace	Large	Burdur	2	B
98 (Kıtık2003)	Cultivar	Large	Turkey	Admixed	B
99 (Eresen87)	Cultivar	Large	Turkey	Admixed	A
100 (Filiz99)	Cultivar	Large	Turkey	Admixed	B
101 (Salkım)	Cultivar	Large	Turkey	Admixed	B

Subpopulation and cluster assignments are given according to population structure and dendrogram analyses, respectively.
[z]Subpopulation assignment based on the proportion of inferred ancestry with a threshold of ≥ 0.70.
[y]Cluster assignments based on the neighbor-joining dendrogram.

Supplemental Table 2. Fragment polymorphism and GD of the SSR markers applied to 101 faba bean accessions.

Primer name	Total no. of amplified loci/polymorphic loci	GD[z]
GBSSR-VF-8	10/10	0.26 ± 0.03
GBSSR-VF-19	4/4	0.18 ± 0.08
GBSSR-VF-20	5/4	0.33 ± 0.08
GBSSR-VF-22	7/7	0.28 ± 0.04
GBSSR-VF-52	7/7	0.36 ± 0.05
GBSSR-VF-113	4/3	0.06 ± 0.02
GBSSR-VF-115	5/5	0.22 ± 0.05
GBSSR-VF-131	7/7	0.13 ± 0.04
GBSSR-VF-149	7/7	0.25 ± 0.05
GBSSR-VF-153	3/3	0.39 ± 0.04
GBSSR-VF-154	4/4	0.08 ± 0.02
GBSSR-VF-159	2/1	0.02 ± 0.01
GBSSR-VF-164	2/2	0.26 ± 0.11
VfG 1	15/15	0.23 ± 0.04
VfG 3	9/9	0.19 ± 0.05
VfG 4	3/3	0.17 ± 0.04
VfG 9	13/13	0.23 ± 0.04
VfG 10	13/13	0.34 ± 0.04
VfG 11	11/11	0.24 ± 0.04
VfG 13	8/8	0.26 ± 0.06
VfG 15	12/12	0.19 ± 0.03
VfG 19	11/11	0.15 ± 0.03
VfG 27	9/9	0.26 ± 0.05
VfG 31	14/13	0.06 ± 0.02
VfG 34	8/7	0.14 ± 0.03
VfG 41	15/15	0.22 ± 0.04
VfG 44	10/10	0.14 ± 0.04
VfG 47	13/12	0.11 ± 0.03
VfG 53	15/15	0.12 ± 0.03
VfG 67	12/12	0.23 ± 0.03
VfG 69	12/12	0.23 ± 0.05
VfG 87	11/11	0.20 ± 0.04

GD = gene diversity, SSR = simple sequence repeat.
[z]For each primer, average GD $\pm$ SE is presented.

Supplemental Table 3. Individuals selected for the core collection.

Genotype no.	Genotype name
1	CGN10322
2	CGN19987
3	CGN10321
4	CGN10320
5	CGN10382
6	CGN10384
10	CGN10387
11	CGN10385
13	CGN10386
14	TR12123
16	TR12540
19	TR12767
22	TR23018
23	TR26194
25	TR26265
27	TR26453
30	TR28096
31	TR31590
36	TR33140
39	TR33547
43	TR37041
45	TR37255
47	TR39064
54	TR44876
55	TR44914
56	TR44928
60	TR49377
63	TR49386
66	TR51377
67	TR51378
70	TR53683
71	TR53724
72	TR53748
74	TR53770
77	TR53947
78	TR53948
80	TR61267
81	TR61300
86	TR74184
87	TR74214
89	TR74337
91	TR75421
94	TR77143
96	TR79569
98	Kıtık2003
100	Filiz99
101	Salkım