MOLECULAR GENETIC ANALYSIS IN FABA BEAN (*Vicia faba* L.)

A Thesis Submitted to
the Graduate School of Engineering and Sciences of
İzmir Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in Molecular Biology and Genetics

by
Mazen Ali ABUZAYED

July 2019
İZMİR

We approve the thesis of **Mazen A S ABUZAYED**

Examining Committee Members

Prof. Dr. Sami DOGANLAR
Molecular Biology and Genetics, İzmir Institute of Technology

Prof. Dr. Hüseyin Çağlar KARAKAYA
Molecular Biology and Genetics, İzmir Institute of Technology

Prof. Dr. Hülya İLBİ
Department of Horticultural, Ege University

Assoc. Prof. Dr. Gülistan MEŞE ÖZÇİVİCİ
Molecular Biology and Genetics, İzmir Institute of Technology

Assoc. Prof. Dr. Zeynel DALKILIÇ
Aydın Adnan Menderes University,
Faculty of Agriculture, Department of Horticulture

12 July 2019

Prof. Dr. Sami DOĞANLAR
Supervisor, Molecular Biology and Genetics,
İzmir Institute of Technology

Prof. Dr. Volkan SEYRANTEPE
Head of Department of Molecular
Biology and Genetics

Prof. Dr. Aysun SOFUOĞLU
Dean of the Graduate School of
Engineering and Sciences

ACKNOWLEDGMENTS

I would like to thank my supervisor Prof. Dr. Sami DOĞANLAR for giving me the opportunity to study in his research lab, and for his support and encouragement. I would like to thank Prof. Dr. Anne FRARY for her patience, understanding and endless support while analyzing data and writing this thesis and other manuscripts.

This study was supported by grant 0424.STZ.2013-2 from The Republic of Turkey's Ministry of Science, Industry and Technology to Prof. Dr. Anne FRARY with contributions from Polen Seed Co. Additional support was received from the Scientific and Technological Research Council of Turkey grant 113E326 to Dr. Jens ALLMER.

I would like to thank Ali UNCU and Ayse OZGER UNCU for SSR validation by sequencing. Also, I would like to thank Ibrahim ÇELIK, for helping in data analysis. I would like to thank Neslihan TEK EKEN and Nergiz GURBUZ for helping in HPLC analysis.

I would like to thank the Plant Molecular Genetics Lab members for their kindness and cooperation.

ABSTRACT

MOLECULAR GENETIC ANALYSIS IN FABA BEAN (*Vicia faba* L.)

Faba bean (*Vicia faba* L.) is an important food legume crop with a huge genome. In this study, we used Next Generation Sequencing (NGS) technology for development of genomic simple sequence repeat (SSR) markers. A total of 14,027,500 sequence reads were obtained comprising 4,208 Mb. From these reads, 56,063 contigs were assembled (16,367 Mb) and 2138 SSRs were identified. Mono and dinucleotides were the most abundant, accounting for 57.5% and 20.9% of all SSR repeats, respectively. A total of 430 primer pairs were designed from contigs larger than 350 nucleotides and 50 primers pairs were tested for validation of SSR locus amplification. Nearly all (96%) of the markers were found to produce clear amplicons and to be reproducible. Thirty-nine SSR markers were then applied to 46 faba bean accessions from worldwide origins, resulting in 161 alleles with 87.5% polymorphism, and an average of 4.1 alleles per marker. Gene diversity (GD) of the markers ranged from 0.00 to 0.48 with an average of 0.27. Testing of the markers showed that they were useful in determining genetic relationships and population structure in faba bean accessions. In addition, 26 morphological and seven biochemical (phenolics content, flavonoids, protein, L-DOPA, tannins, vicine and convicine) characters of 61 landraces and 53 faba bean cultivars were analyzed. There was high diversity for the studied characters among the accessions. Association mapping for these morphological and biochemical characters with 59 SSR markers (442 fragments) was conducted using a general linear model based on the Q matrix. As a result, 48 significant loci were detected for 22 morphological characters, and 26 loci were detected for six biochemical traits. The range of LD (r^2) was from 0.09 to 0.18, and from 0.06 to 0.13 for morphological and biochemical associations, respectively. This study can help breeding programs in selection and improvement of faba bean production.

ÖZET

BAKLADA (*Vicia faba* L.) MOLEKÜLER GENETIK ANALİZLER

Bakla (*Vicia faba* L.) çok büyük genomu ile önemli bir baklagil türüdür. Bu çalışmada, genomik basit dizi tekrarı (BDT) markörlerinin geliştirilmesi için İleri Jenerasyon Dizi Analizi Teknolojisi kullanılmıştır. 4.208 Mb'lık genomu kapsayacak şekilde toplam 14.027.500 dizi okunmuştur. Bu okumalardan, 56.063 kontig elde edilmiş olup (16.367 Mb), 2138 adet BDT belirlenmiştir. Tüm BDT'ları içerisinde %57.5 ve %20.9 ile mono- ve di- nükleotid tekrarları, beklenildiği gibi en çok elde edilen tekrar dizileri olmuştur. 350 nükleotidden büyük kontigler için toplam 430 primer çifti tasarlanmış olup, bu primerlerin 50 tanesi BDT lokuslarının çoğaltımının test edilmesinde kullanılmıştır. Bu markörlerin neredeyse hepsi (%96) temiz fragmentler çoğaltmıştır ve tekrar çoğaltabilme özelliğine sahiptir. 39 adet BDT markörü, dünya genelini kapsayan 46 bakla genotipine uygulanmış, sonuçta %87.5 polimorfizm ile 161 adet alel elde edilmiş olup, markör başına ortalama 4.1 alel elde edilmiştir. Markörlerin genetik çeşitlilik (GÇ) değerlerinin ortalaması 0.27 olup, 0.00 ile 0.48 arasında değişiklik göstermiştir. Markörlerin test edilmesi, bakla genotiplerinin genetik ilişkilerinin ve popülasyon yapısının belirlenmesi açısından yararlı olmuştur. Ek olarak, 26 morfolojik ve yedi adet biyokimyasal (fenolik içeriği, flavonoidler, protein, L-DOPA, tanninler, vicine ve convicine) karakter, 61 yerel tür ve 53 bakla çeşidinde analiz edilmiştir. Çalışılan karakterlerde çok yüksek çeşitlilik tespit edilmiştir. Bu morfolojik ve biyokimyasal karakterlerin, 59 BDT markörü ile ilişkilendirme haritası Q matrise bağlı genel doğrusal model kullanılarak oluşturulmuştur. Sonuç olarak, 22 morfolojik karakter için 48 önemli lokus tespit edilirken, 6 biyokimyasal özellik için ise 26 lokus ilişkili bulunmuştur. Bağlantı eşitsizliği (linkage disequilibrium – LD – r^2) değerleri morfolojik özellikler için 0.09 ile 0.18 ve biyokimyasal ilişkiler için 0.06 ve 0.13 aralığında bulunmuştur. Bu çalışma, bakla çeşitlerinin ıslah programlarına ve bakla üretiminin arttırılmasında yardımcı olabilecektir.

To my wife and my children

TABLE OF CONTENTS

LIST OF THE FIGURES

Figure **Page**

Figure **Page**

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1. Faba Bean Origin, Distribution and Production

Faba bean (*Vicia faba* L.) is sometimes referred to as horse, broad or field bean (Duc, 1997). It is an annual herbaceous species and a member of the Fabaceae family, the third largest plant family with 19,000 species and 727 genera (Lewis *et al.*, 2005). Faba bean is one of the most important legume crops and is believed to have originated in the Near East and cultivation started early in Neolithic times, 8.000 B.C. (Karkanis *et al*., 2018; Torres *et al.*, 2006; Cubero, 1974). Faba bean is considered as a partially cross-pollinated species, and the rate of out-crossing ranges from 4 to 84% (Bond and Poulsen, 1983). However, its origin and domestication are still debated because no wild progenitor has been discovered yet, or this progenitor may be extinct. Interspecific hybridization with other *Vicia* species has failed (Duc, 1997).

Faba bean is planted in warm-temperate and subtropical countries in the winter and in northern latitudes in the spring (DUC, 1997). There are four subspecies according to seed size: *V. faba major* has large seeds and is grown in China and South Mediterranean countries; *V. faba equina* has medium seeds and is grown in North Africa and Middle East; *V. faba minor* has small seeds and is grown in Ethiopia and North Europe and *V. f. paucijuga* which is a primitive form found from Afghanistan to India (Cubero, 1974). Production of faba bean is concentrated in nine major agro-ecological regions: the Nile valley, Mediterranean basin, Central Asia, East Asia, Ethiopia, Oceania, Latin America, North America and northern Europe (Torres *et al.*, 2006). According to FAOStat (2019) data, world production of faba bean was 4.84 million tonnes from nearly 2.5 million ha in 2017. China accounted for 37.3% of production followed by Ethiopia, Australia, United Kingdom, Germany, France, and Egypt with 19.2, 7.7, 6.2, 3.9, 3.9, and 2.3% of world production, respectively. Figure 1.1 shows faba bean producers by region. Yield

varies greatly among regions with the greatest yield in Europe (3.0 tonnes/ha), followed by Asia, Africa, Australia, and the Americas with 2.0, 1.8, 1.6, 1.2 tonnes/ha, respectively.

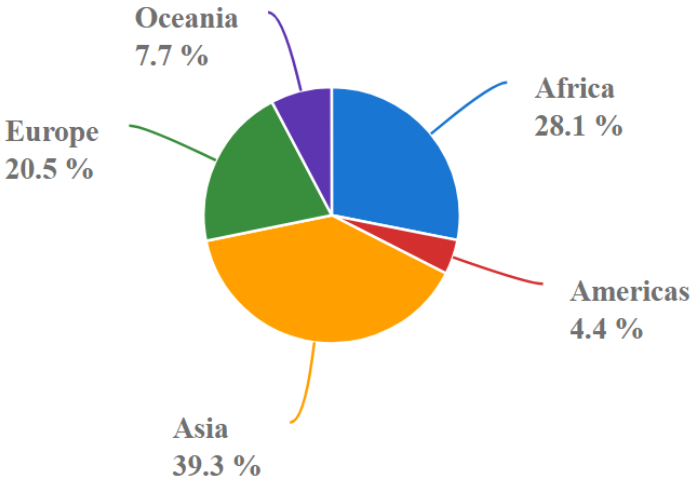

Figure 1.1. World production of faba bean by region.
(Source: FAOStat. May 14, 2019)

According to FAOstat data for world production from 1994 to 2017, faba bean production has increased since 2007 (Figure 1.2). This increase may be due to selection in faba bean and development of new cultivars with high seed yield by breeding programs.

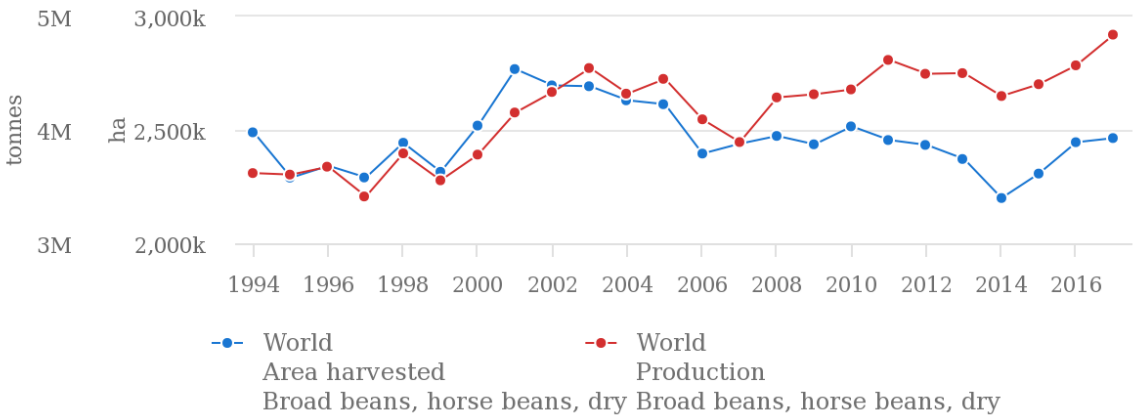

Figure 1.2. Production and area harvested of dry faba bean in the World (Total) 1994-2017.
(Source: FAOStat. May 14, 2019)

1.2. Faba Bean in Turkey

Faba bean is one of the most important legumes in Turkey (Baloch *et al.*, 2014). Turkey ranked 24th for faba bean production in 2017 with 14.7 thousand tonnes cultivated on 5.3 thousand ha, and yield was 2.8 tonnes/ha (FAOStat, 2019). Statistics (Figure 1.3) show a clear decrease (71.7%) in faba bean production since 1994, although there was an increase in yield. This decreased in faba bean production in Turkey is attributed to the heavy demand for the three legumes that are produced in the largest quantities: chickpeas, lentils, and dry beans with 470, 430, and 239 thousand tonnes, respectively. In the last decade, few varieties of faba bean have been developed within Turkey and these cultivars were introduced from International Center for Agricultural Research in the Dry Areas (ICARDA) breeding lines (Karaköy *et al.*, 2014) or from different countries that also prefer large-seeded types (Baloch *et al.*, 2014). The faba bean landraces in Turkey will disappear in the near future due to their replacement by these newly introduced cultivars (Karaköy *et al.*, 2018). Despite the importance of faba bean in the world, only a few breeding studies were conducted in Turkey (Karaköy *et al.*, 2018; Inci and Toker, 2016; Sozen and Karadavut, (2016); Baloch *et al.*, 2014; Karaköy *et al.*, 2014; Inci and Toker, 2011).

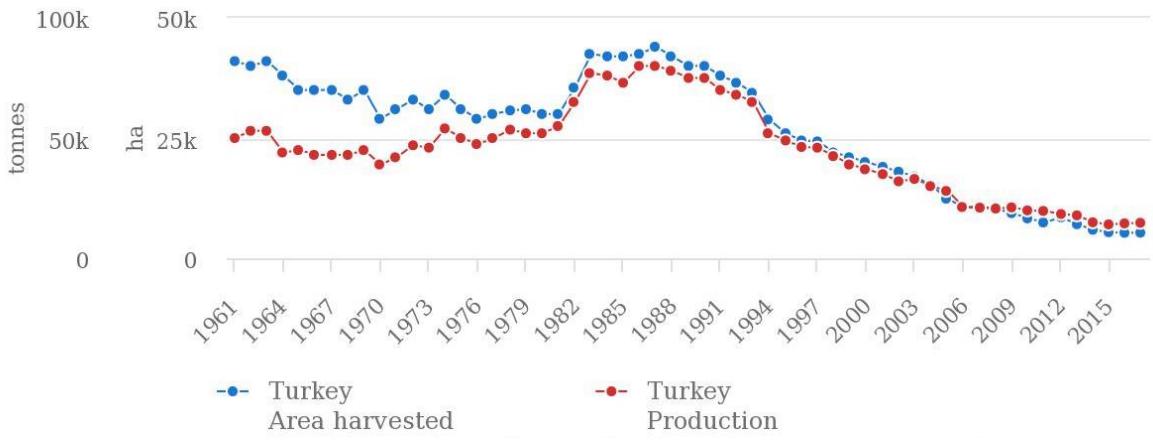

Figure 1.3. Production and area harvested of dry faba bean in Turkey 1961 – 2017.

(Source: FAOStat. May 14, 2019)

1.3. Importance of Faba Bean

Faba bean is widely used in human food, its seeds are consumed dried, fresh, or canned, and also as livestock feed for poultry, pigs and horses. Faba bean has nutritional and anti-nutritional Properties.

1.3.1. Nutritional Properties

This plant also has a role in crop rotation because of its symbiosis with nitrogen-fixing bacteria (DUC, 1997). Faba bean is an excellent source of protein and starch with protein content ranging between 22.0 to 38.2% of seed dry matter, depending on genotype. Therefore, it is considered as a cheap protein in many developing countries where people cannot afford to buy meat (Alghamdi *et al.*, 2012). Total carbohydrate content of faba bean is 57.3%, with average starch content of 47%, faba bean seeds contain 12% fiber and 1.2–4.0% lipids (Baginsky *et al.*, 2013). Faba bean is also an important source for vitamins and a good source of energy with 320 kcal/100 g (Ofuya and Akhidue, 2005). Faba bean is a rich source of folates (B vitamins) with 148 mg/100 g, which have roles in DNA synthesis and repair, amino acid metabolism and prevention of anaemia by helping in red blood cell production (Singh, 2018). Faba bean is rich in mineral elements. Karaköy *et al.* (2018) reported the average nitrogen, phosphorous, potassium, magnesium, calcium contents as 6.40, 0.56, 1.51, 0.35, and 0.62%, respectively. According to the same source, faba bean also contains 17.6, 42.7, 83.8, and 24.0 mg per kg of copper, zinc, iron and manganese, respectively. Faba bean has phenolic compounds including phenolic acids such as gallic acid, flavonoids, tannins, stillbenes and lignins. Flavonoids are very important compounds because they have antioxidant, antiinflammatory, antiviral, anticancer and antiatherosclerotic effects (Nijveldt *et al.*, 2001). Faba bean contains L-3,4-dihydroxyphenylalanine (L-DOPA) a molecule that is used for Parkinson's disease (PD) treatment. PD is the second most common neurodegenerative disease in elderly people resulting in disability due to an imbalance between acetylcholine and dopamine in the brain. L-DOPA can be used for PD treatment due to its ability to cross the barrier between the blood and brain. In the brain, L-DOPA is converted to dopamine to compensate for dopamine deficiency (Topal and Bozoğlu, 2016; Oviedo-Silva *et al.*,

2018). Ramya and Thaakur (2007) reported that faba bean dry seeds contain 0.07% L-DOPA.

1.3.2. Anti-nutritional Properties

Faba bean is a suitable food for diabetics and may help prevent heart disease and reduce levels of blood glucose due to its chemical composition (Baginsky *et al.*, 2013). However, some compounds in the seed can restrict the value of faba bean as a protein source.

1.3.2.1 Vicine and Convicine

Vicine and convicine (VC) are thermostable, glucosidic aminopyrimidine derivatives which accumulate at maturity in the cotyledons of faba bean during seed development (Khamassi *et al.*, 2013). The seed content of VC ranges from 6 to 14 g/kg in wild type, however, VC-free genotypes contain only 5–10% of this amount (Duc *et al.*, 1999). VC is hydrolyzed by β-glucosidase, resulting in the aglycones divicine and isouramil which cause oxidation of glutathione in red blood cells. This can result in haemolysis of red blood cells (favism) in individuals with glucose-6-phosphate dehydrogenase (G6PD) deficiency such that glutathione cannot be regenerated at normal levels. This results in abnormal red blood cells that can be recognized by macrophages in the immune system. Favism can be severe and sometimes fatal with 400 million individuals under risk worldwide (Khamassi *et al.*, 2013). β-glucosidase activity is low in young seeds, at a maximum in ripe seeds and drops again in older seeds. This enzyme is inactivated by seed drying, cooking or by acids similar to adult gastric juice. In animals, a diet containing 50% faba beans reduced absorption of Mn and Zn; and if the diet contained only faba beans, it resulted in impairment in growth, liver weight, and muscle mass in growing rats (Crépon *et al.*, 2010). Duc *et al.* (1989) found a Greek landrace (mutant) with low VC after screening a large number of faba bean accessions. This led to the release of several cultivars in Europe with low VC. The pathway of VC biosynthesis is not known, so there are no identified enzymes or sequenced genes that can be used in understanding these compounds (Purves *et al.*, 2018; Khazaei *et al.*, 2015). There were many studies on VC as an anti-nutritional component of faba bean. Lattanzio *et al.* (1983) measured VC in dry seeds of 10 faba bean varieties using HPLC, and reported averages of 0.58 and 0.36% dry weight for vicine and convicine, respectively. Wang and Ueberschär (1990) analyzed VC

content of 20 varieties by HPLC, average vicine concentration was 6.7 mg and convicine content was 2.3 mg for 1 g dry seeds. Ramsay and Griffiths (1996) measured VC using HPLC in many stages in V. faba and V. narbonensis and illustrated that accumulation of VC was in the cotyledons and that leaf laminae contained no VC. Burbano *et al.* (1995) studied variation of VC during pod development by HPLC with the highest amount found in fresh green cotyledons (80% moisture). Levels gradually declined to a constant level when the seed was dry (40% moisture). Khamassi *et al.* (2013) analyzed 57 accessions by HPLC and a β-glucosidase sensitivity assay was used as further proof for identification of VC content, the range for vicine was from 0.01 to 0.95 and from 0.01 to 0.54% dry weight for convicine.

1.3.2.2. Tannins

Tannins are polyphenol compounds located in the hull of faba bean seed and are considered as one of the most important antinutritional factors (ANFs) (Gutierrez *et al.*, 2007). Tannins interfere with digestive enzyme activity by forming complexes with nutrients resulting in reduced digestibility. Tannins are also responsible for an astringent taste and can be eliminated by dehulling (Marquardt *et al.*, 1983). Tannins concentration ranges between 0.1 to 10.4 g/kg dry matter of the seed (Duc *et al.*, 1999). Wang and Ueberschär, (1990) estimated condensed tannins which ranged from 7.1–149.7 mg catechin equivalent g^{-1} shell. Faba bean with high tannins has black spotted, pink or red flowers while low tannin faba beans have white flowers (Crépon *et al.*, 2010). Phenolic content of faba bean differed significantly among genotypes which correlated with tannin content (Oomah *et al.*, 2011).

Anthocyanins and condensed tannins (proanthocyanidins) protect faba bean against insect attack, microbial pathogens and UV radiation. These compounds are derived from the same precursors, anthocyanidins. Catechin and epicatechin are the building blocks of condensed tannins. Twenty genes were characterized for condensed tannins biosynthesis using the legume model *Medicago truncatula*. Figure 1.4 shows the biosynthesis pathway of condensed tannins and anthocyanin (Li *et al.*, 2016). Absence of tannins is controlled by two recessive genes zt-1 (*V. faba major* L. origin) and zt2 (*V. faba minor* L. origin) and mutation in either of these genes blocks the synthesis of anthocyanins resulting in white-flowered plants (Picard, 1976).

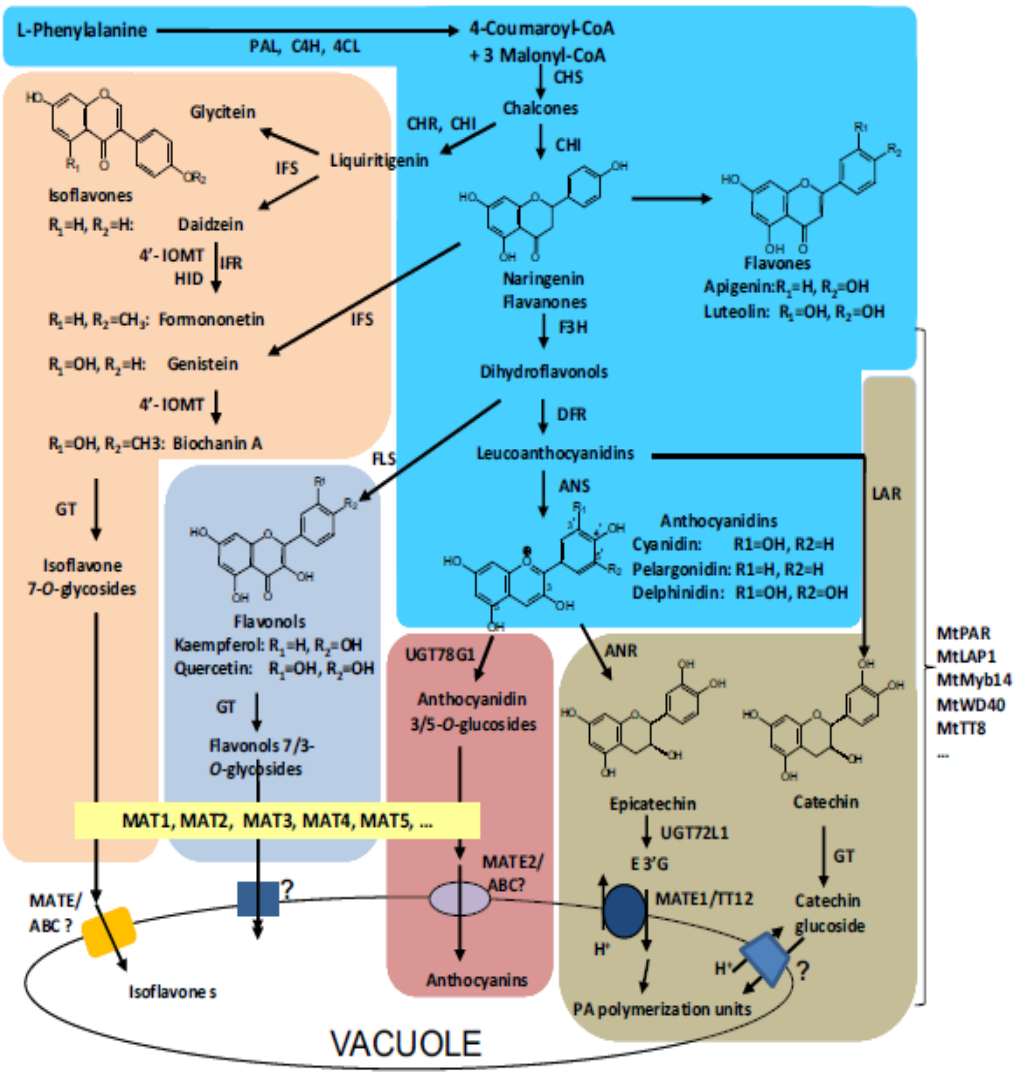

Figure 1.4. Scheme of proanthocyanidin (Condensed tannins and anthocyanins) biosynthesis in *M. truncatula*. (Source: Li *et al*., 2016)

1.3.2.3. Phytic Acid

Phytates are ANFs present in the cotyledons and seed coats of faba beans (Griffiths, 1983). Faba beans are a good source of dietary minerals, such as phosphorus, potassium, calcium, sulfur and iron. However, more than 40–60% of the phosphorus present is unavailable as phytates. Phytic acid in legumes has been reported as reducing the bioavailability of minerals (Deshpande and Cheryan, 1984). Phytic acid content and

phytase activity vary significantly among genotypes (Oomah *et al.*, 2011). Thus relatively low phytic acid content and high phytase activity of low-tannin faba bean genotypes are beneficial and essential traits for their use in human and animal nutrition.

In addition, oligosaccharides (raffinose, stachyose and verbascose) are considered as ANFs and are involved in flatulence production. Consequently, the presence of these sugars in faba beans is a constraint on its utilization (Concepcion *et al.*, 1998).

1.4. Genetic Resources and Breeding

More than 43,695 faba bean accessions are found in 37 germplasm collections throughout the world. ICARDA holds 21% of these accessions followed by 10% at the Institute of Crop Germplasm Resources of the Chinese Academy of Agricultural Sciences, and 6% at the Australian Temperate Field Crops Collection (Khazaei, 2014). There are 373 accessions at Aegean Agricultural Research Institute (AARI) in Izmir, Turkey. Germplasm collections are sources of faba bean diversity which can be leveraged to increase yield and improve other agro-morphological traits. The value of genetic resource collections is greatly enhanced when accessions are evaluated for agronomic, yield, quality, biotic and abiotic stress traits (Duc *et al.*, 2010) as this knowledge helps guide the breeder in selecting parents and potential cultivars.

Faba bean can be grown in different climatic conditions. However, many biotic and abiotic stresses can adversely affect production. Heat, drought, and frost are the most common abiotic factors that limit faba bean production. The optimum temperature for faba bean growth is around 22°C. Faba bean is considered as a cool-season crop and tolerant of chilling between 0 and 10°C. Most faba bean genotypes are tolerant of temperatures down to –5°C for a short time. Other abiotic factors that affect faba bean production include mineral toxicities, salinity and waterlogging (Torres *et al.*, 2012). Many biotic stresses can also affect the growth of faba bean and reduce its production. These factors include some fungal pathogens such as Ascochyta blight (*Ascochyta fabae*), rust (*Uromyces viciae-fabae*), chocolate spot (*Botrytis fabae*), downy mildew (*Peronospora viciae*), and foot rots (*Fusarium* spp.) (Torres *et al.*, 2006). Also, broomrape (*Orobanche crenata*), an aggressive parasitic angiosperm, causes 50% to 80% production losses in faba bean fields (Gressel *et al.*, 2004). The main aim of breeding is

increasing yield, however, low yield stability is partly attributed to the allogamous nature of faba bean and diseases (Torres *et al.*, 2012).

1.5. Faba Bean Genomic Tools and SSR Markers

Faba bean is a diploid with 2*n* = 12 chromosomes and the largest known genome (13,000 Mb) among legumes (Sato *et al.*, 2010; Ellwood *et al.*, 2008). The faba bean genome is about 26 times larger than the legume model species, *M. truncatula*, Figure 1.5 shows a comparison between faba bean and other crop and model plants. Diversity analysis of faba bean genotypes using molecular markers is just beginning (Duc *et al.*, 2010). To study genetic diversity and population structure, as well as for genetic mapping of various traits in breeding programs, we need appropriate molecular marker systems. Several types of molecular markers have been used to characterize and elucidate genetic diversity in faba bean, including Restriction Fragment Length Polymorphism (RFLP) markers (Helsper *et al.*, 1993) (Van de Ven *et al.*, 1990), Random Amplified Polymorphic DNA (RAPD) markers (Link *et al.*, 1995), Amplified Fragment Length Polymorphism (AFLP) markers (Gresta *et al.*, 2010; Zong *et al.*, 2009; Zeid *et al.*, 2003; Terzopoulos and Bebeli, 2008; Wang *et al.*, 2012), Sequence-Specific Amplification Polymorphism (SSAP) markers (Ouji *et al.*, 2012), and Single Nucleotide Polymorphism (SNP) markers (Kaur *et al.*, 2014).

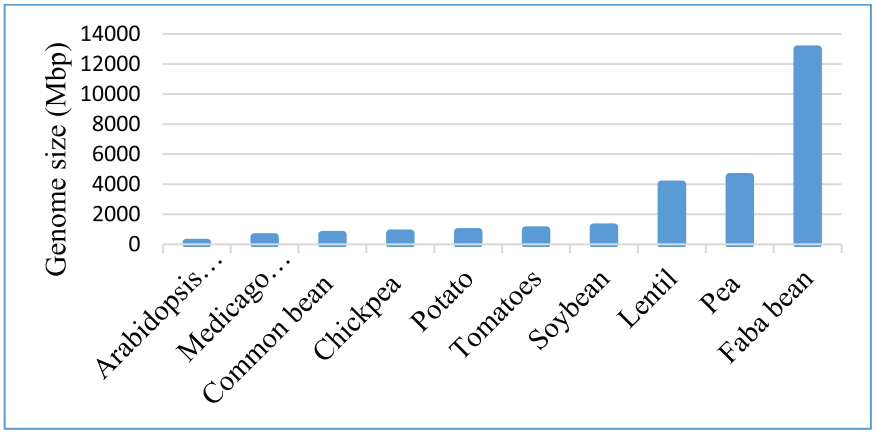

Figure 1.5. Comparison between faba bean genome size and with some other plants.

Among these markers are also simple sequence repeats (SSRs), these are also known as microsatelllites and were first discovered by Hamada *et al.* (1982). SSRs are short tandem repeat motifs which have frequent occurrence in all prokaryotic and eukaryotic genomes (Zane *et al.,* 2002). SSRs occur in coding and non-coding regions and are more informative than other molecular markers such as RAPD, RFLP, and AFLP markers due to the number of different alleles that may exist at a given SSR locus and their ease of use. SSRs are PCR-based, highly abundant in plant genomes, multiallelic, codominant, reproducible and transferable among related species. Because of these characteristics, SSRs are used to study genetic diversity, gene flow, evolutionary studies, infer infraspecific genetic relations, marker-assisted selection, and genetic mapping (Vieira *et al.,* 2016; Senan *et al.,* 2014; Akash and Myers, 2012; Zalapa *et al.,* 2012).

1.5.1 Development of SSR Markers

There are many methods to develop SSR markers. SSR repeats should be identified along with flanking nucleotide sequence to design primers for a given species. The first protocol for de novo isolation of SSR markers was described by Rassmann *et al.* (1991), who used colony hybridization with SSR probes to identify SSR-containing clones, however, it turned out to be laborious and expensive. To date, most faba bean SSR markers were developed by mining of public databases for genic SSRs, cDNA sequencing or SSR-enriched library methods (Pozarkova *et al.,* 2002; Zeid *et al.,* 2009; Gong *et al.,* 2010, 2011; Ma *et al.,* 2011; Akash and Mayers, 2012; Kaur *et al.,* 2012; Yang *et al.,* 2012; Suresh *et al.,* 2013; El-Rodeny *et al.,* 2014). All of these studies (Table 1.1) resulted in the development of approximately 33,117 SSR markers with most (86%) of these from the work of Yang *et al.* (2012). A total of 1133 of the identified markers were used to study genetic diversity in faba bean with 62% (707) of the markers showing polymorphism. This number of validated SSR markers is insufficient because faba bean has a very large genome.

Next-generation sequencing (NGS) methods are high-throughput technologies that can produce millions of short sequences in parallel and are easier, faster and more economical than Sanger sequencing (Shendure and Hanlee, 2008). NGS, along with bioinformatics tools, can be used for large-scale, rapid development of genome-wide and

gene-based SSR loci (Abdelkrim *et al.,* 2009). Two major NGS technologies have been used for development of molecular markers, 454 sequencing (pyrosequencing) and Illumina (sequencing by synthesis). Illumina sequencing can generate 2.5 to 20 Gb per run, while 454 sequencing generates 350 – 720 Mb. and the cost per Mb read by Illumina is cheaper than the 454 sequencing method (Zalapa *et al.,* 2012). Although Illumina sequencing has not yet been used to identify SSRs in faba bean, two studies used 454 sequencing techniques to develop SSR markers for this crop. In the first one, Yang *et al.,* (2012) developed novel SSRs from faba bean using NGS 454 sequencing, from a mixed genome of 247 spring and winter sown faba bean genotypes. They designed 28,503 primers. Of these, 150 were selected for validation and they found that 94 primer pairs were polymorphic among 32 faba bean genotypes and that the number of alleles ranged from 2 to 8 per locus. In the second study, Suresh *et al.* (2013) developed novel transcript sequence cDNA-SSRs in faba bean using 454 pyrosequencing. They designed 240 primers, and found that only 55 primer pairs were polymorphic. They applied the SSRs on 32 faba bean accessions and the number of alleles ranged from 2 to 15 per locus. These new markers are valuable for quantitative trait loci (QTL) mapping, constructing genetic linkage maps and marker-assisted trait selection in faba bean breeding efforts.

Table 1.1. Simple sequence repeat markers developed in faba bean

Reference	Total number of developed markers	Number of tested markers	Number of polymorphic markers	Polymorphism
Pozarkova *et al.,* 2002	391	25	7	28%
Zeid *et al.,* 2009	73	73	54	74%
Gong *et al.,* 2010,	11	11	11	100%
Gong *et al.,* 2011	107	11	11	100%
Ma *et al.,* 2011	21	21	21	100%
Akash *et al.,* 2012	117	34	31	91%

(cont. on next page)

Table 1.1 (cont.)

Kaur *et al.,* 2012;	802	81	39	48%
Yang *et al.,* 2012	28503	150	94	63%
Suresh *et al.,* 2013	1729	240	55	30%
ElRodeny *et al.,* 2014	1363	487	384	79%
Total	33,117	1133	707	

1.6. Genome Mapping

There are many studies in faba bean to identify markers linked to traits. Torres *et al.* (2006) reviewed eight linkage map studies in faba bean that mapped fungal disease and broomrape (*Orobanche crenata*) resistance loci. These studies allowed the identification of genes and QTLs controlling resistance to some of these diseases, however, all were based almost exclusively on dominant markers, primarily RAPD markers.

Molecular markers have also been developed to assist in selection for agronomic traits such as low VC content. Thus, Gutierrez *et al.* (2006) crossed a high VC content line (Vf6) with a VC free genotype [line 1268, a VC- mutant allele which reduced VC content by 10–20 fold (Duc *et al.,* 1989)], and analyzed the F2 population (n=130). Two RAPD markers were linked to the allele VC- and a linkage map was constructed. Khazaei *et al.* (2015) used 210 F5 plants from a cross between Melodie/2 (low VC) and ILB 938/2 (normal VC). The plants were genotyped with 188 SNPs and the study resulted in the identification of a strong single QTL for VC concentration on chromosome 1. Tannins can also be removed by mutation in one of two independent genes named zt1 and zt2. Therefore, tannin free genotypes of faba beans can be obtained by breeding and selection of recessive alleles of one of the two genes. Gutierrez *et al.* (2007) analyzed an F2 population (n=90) from crossing the Vf6 genotype (high tannin, colored flower) and a zt2 line. They identified five RAPD markers that showed a clear association with zt2.

Study of plant characteristics has also been done using faba bean morphological descriptors (Anishettty *et al.,* 1985) including seed size, color and shape; pod length;

number of seeds per pod; plant height; flower color and other traits. Avila *et al.* (2005) constructed a linkage mapping using an F2 population (29H x Vf136), which was the first step towards the identification of molecular markers for agronomic traits (floral characters, yield distribution and yield characters). A total of 15 QTLs were detected for 8 out of 10 traits using RAPD and isozyme markers. In another study, comparative mapping suggested the conservation of one of the faba bean genomic regions controlling the character days to flowering in five other legume species (alfalfa, lotus, pea, lupine, and chickpea). Additional syntenic co-localizations of QTLs controlling pod length and number of seeds per pod between faba bean and Lotus *japonicus* are likely (Cruz-Izquierdo *et al.,* 2012). Ma *et al.* (2013) used SSRs for analysis of 129 F2 individuals from crossing a large-seeded Chinese variety 91825 and small-seeded K1563 resulting in a linkage map with 128 SSR on 15 linkage groups.

1.7. Association Mapping

Linkage and association mapping are the two main methods for locating genes or QTLs in a genome. Genetic linkage mapping frequently uses a biparental segregating population. Its main limitations are that only two alleles at any given locus can be studied in bi-parental crosses of diploids and low mapping resolution (Flint-Garcia *et al.,* 2003). In contrast, association mapping, also known as linkage disequilibrium (LD) mapping, utilizes historic recombination events within a population of unrelated individuals for identification of genetic markers for qualitative or quantitative traits using statistical tools. Association mapping gives improved mapping resolution, reduced research time and the potential to discover a greater number of beneficial alleles (Yu and Buckler, 2006). Association mapping was first reported in maize (Remington *et al.,* 2001) and recently has been widely used in many plants (Huang and Han, 2014). However, only a few experiments have been conducted using association mapping for morphological traits in faba bean (Ali *et al.,* 2016; Sallam *et al.,* 2016a; Sallam *et al.,* 2016b). Ali *et al.,* 2016 conducted association mapping in 189 faba bean lines using 175 SNP and 1147 AFLP markers and many putative QTLs were detected for drought and freezing stress. In another three studies by Sallam and co-workers (Sallam *et al.,* 2016a; Sallam *et al.,* 2016b; Sallam and Martsch, 2015), SNP (156) markers that were generated from a legume model (*M. truncatula*) were used to conduct association mapping in 189 faba bean

accessions. Many putative QTLs were detected for frost tolerance, some morphological, physiological and yield traits.

1.8. Aim of the Study

The first aim of this research was to develop genomic SSR markers in faba bean using next generation sequencing (Illumina Mi-Seq Technology). The SSR markers were tested in faba bean and polymorphic SSRs were applied to 46 diverse genotypes of faba bean to study genetic diversity and population structure. Secondly, 26 morphological traits (14 quantitative and 12 qualitative) were measured for 114 genetically diverse genotypes (including 61 landraces and 53 cultivars) collected from different locations in the world over two growing seasons. In addition, the chemical composition of dry seeds including, total phenolic compounds, flavonoids, protein, L-DOPA, tannins, vicine, and convicine were analyzed in the faba bean genotypes. Finally genotypic data of the 114 genotypes were used for association mapping of the characterized metabolic and morphological traits. In this way, the number and individual phenotypic effects of loci controlling the traits were determined. In addition, markers tightly linked to the traits of interest were identified to allow marker-assisted breeding of faba bean.

CHAPTER 2

MATERIALS AND METHODS

2.1. SSR Development

For SSR development, the Turkish cultivar (Filiz -99) were selected for sequencing using NGS Technology. The SSRs were detected and the primers were designed, then the population structure and genet diversity were studied using 46 faba bean accessions.

2.1.1. Plant Material

A total of 46 faba bean accessions from 17 countries were used as plant material (Table 2.1). Eight accessions were from the Netherlands Gene Bank (NGB); 18 from the Centre for Genetic Resources, the Netherlands (CGN); 11 from the Aegean Agricultural Research Institute (AARI, Turkey); five from the University of Adelaide (Australia) and four from the International Center for Agricultural Research in the Dry Areas (ICARDA, Syria). Seeds were planted and grown in the growth chamber at 24 to 25˚C, 16 h photoperiod.

Table 2.1. Faba bean genotypes used for validation of genomic SSR markers and study genetic diversity

No.	Sample Name	Origin	No.	Sample Name	Origin
1	NGB8642	Finland	24	CGN10385	Turkey
2	NGB1547.1	Finland	25	CGN10374	Syria
3	Mikko	Finland	26	CGN10325	Syria
4	Witkiem manida	Germany	27	TR23018	Turkey

(cont. on next page)

Table 2.1 (cont.)

5	Ukko	Germany	28	TR31590	Turkey
6	Kontu	Germany	29	TR33140	Turkey
7	NGB1542.1	Finland	30	TR37255	Turkey
8	NGB1548.2	Finland	31	TR44876	Turkey
9	CGN7874	Spain	32	TR44928	Turkey
10	CGN15563	Syria	33	TR49380	Turkey
11	CGN15619	Egypt	34	TR53748	Turkey
12	CGN13485	Pakistan	35	TR61267	Turkey
13	CGN13464	UK	36	Ascot	Australia
14	CGN10391	Egypt	37	Manafest	Australia
15	CGN7826	Greece	38	Fiord	Australia
16	CGN7716	Italy	39	Fiesta	Australia
17	CGN7844	Jordan	40	Aquadulce	Australia
18	CGN7781	Netherland	41	Filiz-99	Turkey
19	CGN15641	Netherland	42	Salkım	Turkey
20	CGN10382	Turkey	43	26139	.Colombia
21	CGN10371	Algeria	44	26145	Egypt
22	CGN18892	Netherland	45	ILB938/2	Unknown
23	CGN07875	India	46	Melodie/2	France

2.1.2. DNA Sequencing

For SSR identification, faba bean cultivar Filiz-99 was provided by AARI. Total genomic DNA was extracted using the Wizard Magnetic 96 Plant System (Promega Corp., Madison, WI, USA) and the Beckman Coulter Biomek NX Workstation. Sequencing of genomic DNA was done by next generation sequencing (Illumina Mi-Seq Technology) by the Biotechnology Center at the University of Wisconsin-Madison, USA (https://www.biotech.wisc.edu/). This technology produced 300 nucleotide long, paired-end reads. Raw data and further information, can be found at the SRA database of NCBI (SRA id: SRP076364).

2.1.3. Data Pre-Processing

Adapter and linker sequences were removed from reads with cutadapt (Martin, 2011) version 1.8.3 software using default settings. Any trimmed reads smaller than 50 nucleotides were removed from the dataset since they disrupt mapping and assembly processes. Reads were then mapped with Bowtie version 2.1.0 (Langmead and Salzberg, 2012) software using default settings against the human genome to remove possible human contaminants. Contamination may occur during DNA extraction or next generation sequencing and possible contaminant reads were excluded from the dataset.

2.1.4. Sequence Assembly

ABySS version 1.3.6 (Simpson *et al.,* 2009), a de novo, parallel, paired-end sequence assembler, was employed for sequence assembly. More than 100 runs were performed with different settings such as changing kmer (all possible substrings of length k contained in reads) and required number of reads to make a contig. Assembly quality was based on various parameters, such as the weighted median of contig lengths (N50), a commonly used measure. The best assembly was identified according to the N50 value, assembly nucleotide length (closeness to the estimated size of the *V. faba* genome), length of largest contig and contig number. The settings that were finally chosen to create contigs were: kmer (k=25) and number of reads (n=2).

2.1.5. SSR Detection, Annotation and Primer Design

Only contigs larger than 200 nucleotides were analyzed for SSR detection using an in-house tool SiSeer (http://bioinformatics.iyte.edu.tr/index.php?n=Softwares.SiSeeR). The minimum number of repeats required to identify perfect SSRs was 10 for mononucleotides, four for dinucleotides, and three for motifs comprised of three or more nucleotides. Identified SSR sequences were extracted with their genomic context (padded with 100 nucleotides) and were converted to FASTA formatted sequences. These queries were searched against the Uniprot non-redundant plant protein database (http://www.uniprot.org/uniprot/?query=taxonomy%3A%22Viridiplantae+%5B33090%

5D%22&sort=score) (Taxonomy = Viridiplante) with BLASTX version 2.2.30. Primer design was performed on contigs larger than 350 nucleotides with Primer3 (primer_core) version 2.3.6 (Koressaar and Remm, 2007) using default parameters (Primer task=generic, primer optimum size=20, primer minimum size=18, primer maximum size=24, primer product size=100-300, primer minimum TM=50, primer maximum TM=60, and primer optimum TM=55).

2.1.6. Validation of Genomic SSR Markers

For SSR validation and to ensure that the expected SSRs were amplified by the primers, seeds were planted and grown in the growth chamber at 24 to 25˚C, 16 h photoperiod. Total DNA from the youngest leaves was extracted using CTAB extraction buffer according to Doyle and Doyle (1990). DNA quality was checked by agarose gel electrophoresis and quantification was done by spectrophotometer (Thermo Scientific, Multiskan GO). DNAs from two faba bean samples (ILB938/2 and Melodie/2) were used as template for PCR with four of the SSR markers. Amplified products were analysed using the dye-terminator sequencing method. First, PCR products were purified with the DNA Clean and Concentrator – 5 Kit (Zymo Research) and were used as template in sequencing reactions prepared using GenomeLab DTCS-Quick Start Kit (Beckman Coulter). The thermal cycling conditions were 30 cycles of 96˚C for 20 sec, 50˚C for 20 sec, 60˚C for 4 min. ZR DNA Sequencing Clean-up Kit (Zymo Research) was used for purification of the reaction mixture for each SSR amplicon which was then resuspended in 30 µL sample loading solution (Beckman Coulter) and run on a Beckman CEQ8800 capillary electrophoresis device using the LFR-c method (injection voltage 2.0 kV for 10–15 sec, separation temperature 60˚C, separation voltage 7.4 kV, separation time 45 min).

2.1.7. SSR Amplification

Fifty primer pairs of the developed markers were applied on 46 faba bean accessions (Table 2.1) and SSR amplification were carried out in a final volume of 20 µl and contained 30 ng DNA, 1X PCR buffer, 1.5 mM MgCl2, 0.2 mM dNTPs, 1 pmol forward

and reverse primers, 1 U Taq Polymerase. PCR conditions were 95 ˚C for 4 min for one cycle, followed by 35 cycles of 45 sec at 95˚C for denaturation, 1 min at 55˚C for annealing and 1 min at 72˚C for extension, the final extension cycle was at 72˚C for 5 min. PCR reactions were performed in a Veriti 96-Well Thermal Cycler (Applied Biosystems). PCR products were separated using a capillary electrophoresis instrument (Fragment Analyzer Automated CE System, Advanced Analytical) using the DNF-900 dsDNA Reagent Kit (Advanced Analytical). SSR alleles were visualized and scored using PROSize 2.0 software version 1.2.1.1 (Advanced Analytical).

2.1.8. SSR Data Analysis

SSR alleles were scored for presence (1) or absence (0). Codominant scoring was not possible because most of the markers amplified more than two fragments and alleles could not be identified. Fragments that were observed at less than 10% (low frequency) in faba bean accessions were excluded from all analyses because such products can be unreliable. Gene diversity (Nei, 1973) was calculated depending on the frequency of the allele for each SSR marker and the calculations were performed with the GDdom online (http://plantmolgen.iyte.edu.tr/GDdom/) computer program (Abuzayed *et al.*, 2017) using the formula of Roldan-Ruiz *et al.* (2000):
$GD_i = 2f_i (1 - f_i)$. Where GD_i is the gene diversity of marker 'i', f_i is the frequency of the amplified allele (band presence), and $1 - f_i$ is the frequency of the null allele.

2.1.9. Population Structure Analysis

Marker data were used to infer population structure of the 46 faba bean accessions with the Structure computer program (Pritchard *et al.*, 2000), models with 1 to 10 subpopulations (K) were tested for 20 iterations. Burn-in period was 100,000 and number of Monte Carlo Markov Chain repeats was 500,000. Structure Harvester computer program (Earl and vonHolt, 2012) was used to calculate ΔK values for each model based on posterior probabilities. The model with the highest ΔK was selected as the best. Inferred ancestry threshold was set as ≥ 0.70 to assign the accessions to subpopulations. Accessions with lower probabilities were assigned to the admixed group.

2.1.10. Genetic Diversity Analysis

To study genetic diversity, the binary presence/absence data were used to generate a dissimilarity matrix using the Dice coefficient as implemented by Darwin5 computer program (http://darwin.cirad.fr/product.php). The distance data were used to construct a dendrogram of the 46 faba bean accessions using unweighted neighbor-joining method.

2.2. Morphological Characterization

For morphological characterization, 114 faba bean genotypes (61 landraces and 53 cultivars) were selected from different sources as described in Table 2.2.

2.2.1. Plant Material

Seeds for the 114 genotypes (Table 2.2) were grown in the field in Gulbahce, Izmir, on the west coast of Turkey with coordinates (38˚19’ 47.1’’ N and 26˚38’ 27.8’’ E) during two successive growing seasons (2016/17 and 2017/18). The location has average annual precipitation of 697 mm and average annual temperature of 17.8°C (Turkish State Meteorological Service). Fifteen seeds were sown on the last week of October in 2016 and 2017, respectively. The spacing between rows was 1 m with 0.30 cm within rows. The plants were rain-fed except when conditions were dry. No fertilizers were used during the growing season. The final dry pods harvest was on the last week of May in both 2017 and 2018.

Table 2.2. Faba bean landraces and cultivars used in this study.

Genotype	CV or L	Origin	Genotype	CV or L	Origin
CGN07697	CV	Unknown	CGN18860	CV	Netherlands
CGN07699	CV	Unknown	CGN18862	CV	Netherlands

(cont. on next page)

Table 2.2 (cont.).

CGN07717	L	Ethiopia	CGN18867	CV	Unknown
CGN07719	L	Italy	CGN18867.1	CV	Unknown
CGN07721	L	Unknown	CGN18878	CV	Germany
CGN07723	CV	Netherlands	CGN18892	CV	Netherlands
CGN07725	L	Unknown	CGN18893	CV	Belgium
CGN07728	L	Unknown	CGN18905	CV	Austria
CGN07730	CV	Czech Rep.	CGN18906	CV	UK
CGN07730.1	CV	Czech Rep.	CGN18909	CV	Germany
CGN07733	L	Iraq	CGN18920	CV	Finland
CGN07734	L	Ethiopia	CGN18923	CV	France
CGN07735	L	Ethiopia	CGN18927	CV	Ethiopia
CGN07736	CV	Italy	CGN18934	CV	UK
CGN07738	L	Egypt	CGN18941	CV	Germany
CGN07739	L	Egypt	CGN18942	CV	China
CGN07740	L	Egypt	CGN19979	CV	Spain
CGN07751	L	Unknown	CGN19981	CV	Italy
CGN07757	L	Greece	CGN19982	CV	Netherlands
CGN07766	L	Unknown	CGN19985	CV	unknown
CGN07781	L	Netherlands	CGN19986	CV	Netherlands
CGN07826	L	Greece	CGN19987	CV	Turkey
CGN07827	CV	Ethiopia	CGN19987.1	CV	Turkey
CGN07827.1	L	Ethiopia	CGN19993	CV	Netherlands
CGN07836	L	Ethiopia	NGB1540.1	L	Finland
CGN07842	CV	Unknown	NGB1542.1	L	Finland
CGN07843	L	Italy	NGB1547.1	L	Finland
CGN07844	L	Jordan	NGB1550.1	L	Finland
CGN07871	L	Afghanistan	NGB20019.2	L	Finland
CGN07933	L	Afghanistan	NGB8640	L	Finland
CGN07934	L	Ethiopia	NGB8642	L	Finland
CGN07938	L	Afghanistan	NGB8643	L	Finland
CGN10320	L	Turkey	TR12540	L	Turkey
CGN10322	L	Turkey	TR23018	L	Turkey
CGN10330	L	Ethiopia	TR28096	L	Turkey
CGN10362	L	Turkey	TR31590	L	Turkey
CGN10371	L	Algeria	TR31912	L	Turkey
CGN10382	L	Turkey	TR44862	L	Turkey
CGN10382.1	L	Turkey	TR53947	L	Turkey
CGN10384	L	Turkey	TR71255	L	Turkey

(cont. on next page)

Table 2.2 (cont.).

CGN10385	L	Turkey	TR75421	L	Turkey
CGN10387	L	Turkey	Aquadulce	CV	Australia
CGN10391	L	Egypt	Ascot	CV	Australia
CGN12320	L	Greece	Eresen87	CV	Turkey
CGN12321	L	Greece	Farah	CV	Australia
CGN13445	CV	Germany	Fiesta	CV	Australia
CGN13464	CV	UK	Filiz99	CV	Turkey
CGN13487	L	Pakistan	Fiord	CV	Australia
CGN13518	CV	Netherlands	Icarus	CV	Australia
CGN15556	L	Ethiopia	Kontu	CV	Germany
CGN15615	L	Ethiopia	Manafest	CV	Australia
CGN15620	L	Ethiopia	Melodie/2	CV	France
CGN15621	CV	unknown	Mikko	CV	Finland
CGN15639	CV	Australia	Nura	CV	Australia
CGN15641	CV	Netherlands	PBA Rana	CV	Australia
CGN15644	L	Spain	Salkim	CV	Turkey
CGN18856	L	Jordan	Witkiem Manita	CV	Finland

CV: Cultivar

L: Landrace

2.2.2. Morphological Analysis

Five plants were randomly selected from each genotype in each season to evaluate the 26 morphological characters (14 quantitative and 12 qualitative) as indicated in Table 2.3. The quantitative characters were leaflets per leaf, number of stems per plant, plant height (cm), days to flowering, number of flowers per inflorescence, pod length (cm), number of pods per node, total number of pods per plant, maximum number of ovules per pod, number of seeds per pod, seeds fertilized (%), fresh seed water content (%), 100-seed weight (g) and dry seed yield per plant (g). The qualitative characters were leaflet shape, stipule spot pigmentation, stem pigmentation at flowering time, mature stem color, intensity of petal streaks, wing petal color, pod angle, pod shape, pod surface reflectance, seed shape, seed coat color and hilum color. Characterization was done according to the methods and scales of the faba bean descriptors of the International Board for Plant Genetic Resources (Anishettty, 1985).

Table 2.3. Morphological characteristics examined in the faba landraces and cultivars over two seasons.

Trait	Scale	Description
Quantitative characteristics		
Leaflets per leaf	number	Mean for 5 leaves (1 from 5 separate plants) observed on fully expanded leaves at the median flowering node
Stems per plant	number	Mean from 5 representative plants in late flowering
Plant height	cm	Measured at near maturity from 5 plants.
Days to flowering	days	Days from sowing to 50% of plants in flower
Flowers per inflorescence	number	Mean # per raceme from 2 intermediate nodes on 5 plants
Pod length	cm	Mean of 5 pods
Number of pods per node	number	Mean on the 2nd pod-bearing node of 5 plants
Total pods per plant	number	Mean of 5 plants
Maximum number of ovules per pod	number	Mean of 5 pods for which fertilized and unfertilized ovules were counted
Number of seeds per pod	number	Mean of 5 pods
Seeds fertilized	%	Mean of 5 pods and calculated from maximum number of ovules per pod and number of seeds per pod
Water content of fresh seeds	%	Five fresh seeds were weighed, freeze-dried for 72 h and water loss calculated
100-seed weight	g	From 100 randomly chosen dry seeds
Seed yield	g	Mean for dry seeds of 5 plants
Qualitative characteristics		
Leaflet shape	1,2,3	Narrow, intermediate or rounded

(cont. on next page)

Table 2.3 (cont.).

Stipule spot pigmentation	0,+	Absent or present
Stem pigmentation at flowering	0,3,5,7, x	Absent, weak, intermediate, strong or mixed
Mature stem color	1,2,	Light or dark
Intensity of petal streaks	0,3,5,7	Streaks on standard petal: none, slight, moderate or intense
Wing petal color	1,2,3	Uniformly white, uniformly colored or spotted,
Pod angle	1,2,3,x	Attitude at maturity. Erect, horizontal, pendent or mixed (on second pod-bearing node)
Pod shape	1,2,	Sub-cylindrical, flattened
Pod surface reflectance	1,2,3	On tender pods: matte or glossy
Seed shape	1,2,3,x	Flattened, angular, roun, or mixed
Seed coat color	1,2,3,4, 3,7,9	Within one month after harvest: black, dark brown, light brown, light green, dark green, violet or white
Hilum color	1,2,x	Black, colorless or mixed

2.3. Biochemical Analysis

In this section the chemical composition of dry seeds including, total phenolic compounds, flavonoids, protein, L-DOPA, tannins, vicine, and convicine were analyzed.

2.3.1. Sample Preparation and Extraction

Dried faba bean seeds were collected from 114 genotypes during two successive growing seasons (2016/17 and 2017/18). The seeds for each season were ground to fine powder using a grinder (Knife Mill GRINDOMIX GM 200). The powder was sieved through a sieve-mesh. Extractions for dry seeds were done in 1 % methanolic HCl solvent with 1 g seed per 10 ml solvent left overnight with shaking (250 rpm) and then the tubes were

centrifuged at 1500 rpm for 10 minutes, the supernatants were filtered with whatman No.1 filter paper. Chemical analyses of total phenols, flavonoids, protein, and tannins in seeds of the 114 faba bean genotypes were performed in triplicate, and two replicates were used to measure vicine, convicine, and L-DOPA. Averages were calculated for each year and for both years combined.

2.3.2. Total Phenolic Content

The total phenolic content was determined using Folin-Ciocalteu assay as reported in Elfalleh *et al.* (2009). Spectrophotometric reads of absorbance were at 765 nm. A calibration curve for gallic acid was constructed using 20 to 300 mg/l. Total phenolic contents were converted into milligrams gallic acid equivalents (GAE) per 100 g of dry weight (DW) for the faba bean seed extract.

2.3.3. Tannins Content

Tannins were measured as described by Price *et al*. (1978) using a vanillin colorimetric method where tannins react with vanillin to produce a pink color that is measured spectrophotometrically at 500 nm. A calibration curve was constructed using 10 to 700 mg/L catechine and the tannins concentration was expressed as catechin equivalents (CE) in mg per 100 g DW.

2.3.4. Flavonoids Content

Total flavonoids were measured as reported by Nasri *et al.* (2011). This method is based on the formation of a complex flavonoid-aluminum, having maximum absorbance at 510 nm that can be read spectrophotometrically. A calibration curve was established using rutin with concentrations between 0.01 to 1 mg/ml and the results expressed as rutin equivalents (RE) in mg per 100 g DW.

2.3.5. Protein Content

Total protein was extracted according to the Jones *et al.* (1989) method. One ml 0.1 N NaOH was added to 50 mg dry seed powder and was mixed for 30 sec using vortex mixer and left for 30 min at room temperature. The samples were remixed for 30 sec and centrifuged for 5 min at 10,000 g. The supernatant was used for protein quantitation according to the Bradford (1976) method based on the interaction between protein and Coomassie Brillant Blue G250 (CBBG-250) in acid conditions. One hundred milligrams CBBG-250 (Sigma-Adrich Co.) was dissolved in 50 ml of 95% ethanol and 100 ml 85% phosphoric acid was added. The standard calibration curve was constructed using bovine serum albumin (BSA), linearization of the Bradford calibration curve was calculated by measuring absorbance at 595 and 450 nm (Zor and Selinger, 1996).

2.3.6. Vicine and Convicine and L-DOPA Contents

Water was used for extraction following the method of Pulkkinen *et al.* (2015). Two ml water was added to 0.1 g seed powder, vortex-mixed for 2 min and centrifuged 10 min at 13000 rpm. The supernatant was transferred to a new tube and 12 µl HCL (1M) for protein precipitation was added and vortexed for 30 sec. The samples were centrifuged 5 min at 13000 rpm, and the supernatant was filtered by 0.45 nm disc filter. Two replicates were used to measure VC and L-DOPA using HPLC following the Lattanzio *et al.* (1982) method to separate extracts on a reversed-phase column (C18) with UV detection at 280 nm. Inertsil ODS 4 column (250 x 4.6 mm x 5µm) was used for extract separation. Twenty µl extract was used for injection, water was used as mobile phase at a flow of 0.7 ml/min. L-DOPA was used as a reference standard and measurement of vicine and convicine was according to Khamassi *et al.* (2013). Total of VC was calculated by summing the values of vicine and convicine.

2.4. Statistical Analysis

Averages were calculated for each year and for both years combined. The quantitative characters were analyzed by mean, coefficient of variation (CV) and range. Principal

component analysis (PCA) was used to reduce the quantitative variables into a smaller set of artificial variables. The percentage distributions for qualitative characters were calculated using descriptive statistics. The Spearman's rho correlation coefficient was used to study the relationships between variables. All calculations were made using SPSS (18.0) and OriginPro (2017 version) statistical software.

2.5. Association Mapping

DNA was extracted from young ground faba bean leaves using a CTAB method according to Doyle and Doyle (1990). The integrity of DNA was assessed by gel electrophoresis. DNA was quantified by spectrophotometer (Thermo Scientific, Multiskan GO). A total of 59 SSR markers was used for genotyping the 114 genotypes. The markers were from different sources: 26 FbgSSR markers (Abuzayed *et al*., 2017), 14 GBSSR-VF markers (Suresh *et al.*, 2013) and 19 VfGSSR markers (Zeid *et al.*, 2009). The GBSSR-VF and VfGSSR markers were amplified as described by Göl *et al.* (2017). Primer sequences and annealing temperatures are in Table 2.4. SSR amplification for each FbgSSR primer pair was performed as described in section 2.1.7 of the material and methods. Each SSR band was scored as present (1) or absent (0) for each primer pair in 114 genotypes to generate binary data. These data were used for determination of population structure with the Structure 2.2.3 program (Pritchard *et al.*, 2000) as described in section 2.1.9 of materials and methods.

The data generated for the SSR markers were used for association mapping of 26 morphological characters using four models. The general linear model (GLM), GLM based on the Q matrix, mixed linear model (MLM) model, and MLM based on kinship and Q matrix were tested. Before conducting association mapping analysis, minor SSR alleles (below 10% frequency) were removed. Linkage disequilibrium (LD) values (r^2 and *P*-values) between SSR loci were calculated using TASSEL v2.1 software (Bradbury *et al.*, 2007). SSR loci with P values less than 0.001 were considered to be associated with a given trait and the r^2 value for the marker was used to estimate the QTL effect. Q-values were calculated to measure significance in terms of the false discovery rate for the whole set of *P*-values with a threshold Q value of 0.2 (Storey and Tibshirani, 2003).

Table 2.4. SSRs markers used in association mapping.

Primer name	Forward primer sequences (5'-3')	Reverse primer sequences (5'-3')	Annealing temperature (°C)
FbgSSR26	GGTTGTGTCACTTTTCTTGG	AATAAGACCTTAACTTTATTAACC	55
FbgSSR29	ACTTCCAAAAATTTCAGAATCTC	CCCAACTGAAGAAAAGGGTA	55
FbgSSR30	TCCAAAAATTTCAGAATCTCCA	CCCAACTGAAGAAAAGGGTA	55
FbgSSR37	ATGCACGTTACAAGACATTG	CTTTCCTCGCAAAAGGATTG	55
FbgSSR140	TTCAAATGTAAACAGGCGTG	ACCGTTGAGAGTAAAAGGAA	55
FbgSSR198	TGAGACAAATCAGCATTCCA	GCATTTGCATTCACATTTGG	55
FbgSSR229	TTCTAGAATTGGTGCTCCTG	TGCTTGAATATTGAGAGAAGT	55
FbgSSR293	TGAGTGGAGATCTGCTAAGA	AGCAATTGCATTCTAAAGCC	55
FbgSSR309	GAACTATGAAGAGCAGCAGT	AGTTGTTTACATGGACGTGT	55
FbgSSR322	AAGGTGGTGGTGATTCAATT	ATTTTATCTTGCCCATGGGT	55
FbgSSR375	TTCAACCGGTAAAGAGAAGG	ACCAAAACTCTGATGGTGAA	55
FbgSSR382	TGAGAAAGTTGAGTGACTGG	ACCTTTGATAAATTGGAATAGA	55
FbgSSR451	GAACGACTTGAGAGAGAGTC	TTTTAAACCCTAAGGACGGG	55
FbgSSR518	AGTTCTCAAAGCGTTCTTCT	GCTTGTATATTGTGTGAAGTCT	55
FbgSSR520	GCTTGCAAGTAAGTGTGTTT	GAAAGGTTGTGGTTGATTGG	55
FbgSSR525	GGACACATCTCAATCATCCA	ACACATCTCTTGTTACAGCA	55
FbgSSR545	TGAATTCTCTTCTCACGTGG	CGAGTCAATTTGCACAAACT	55
FbgSSR563	TTTATGAATTGGCGTTGTGG	AACAAAACTCACCTTTCAATT	55
FbgSSR564	TCCCTTTTGCTTGTTTATGA	CCTCCGTGTTATCAAACAGT	55
FbgSSR604	CGTTTTGGCTCATAATGCTT	TTTTAGCCATGTACTGTGCT	55
FbgSSR619	TATTTTAGTGGCCAGATGCA	TGGAGAGGTGTTTCAACAAA	55
FbgSSR631	AATGTGATAAGCGCAACATG	TTGGTATTTATCGCTTGTCT	55
FbgSSR633	CTCCAAAACCAGAGTCTGTT	TTTATCTGTAGAGGCATCGC	55
FbgSSR643	GGCAAAAGATGGAGTCCTTA	TAATTTTTGGGCATTGGGAC	55
FbgSSR663	ACTCGAAATCCATCAAGCAT	GCTTTGTGCACCAACAATAT	55
FbgSSR675	ATTGGGGAACTGCCTAATTC	GCAATTTATCAAACACTTGGTG	55
GBSSR-VF19	TCCATCAACCTCAAATCCA	CCGTACTTGTCCACGGAA	55
GBSSR-VF20	TCCACCAAGTCCACCTGA	AATAAGGGCGCAGGAGAG	55
GBSSR-VF22	CGAAGCCTCCTCCTCTTC	CAAGTGGCCGTTTTTCAA	57
GBSSR-VF52	GGTTTCTTGTCCAAATAAGACG	TGCGATTCTGGAAATTGG	56
GBSSR-VF113	TGGTGGTGCTTCTTTCCA	TGGTGAGCTTGGAACTGC	55
GBSSR-VF115	TGCTGCTTTTCCAACCAT	GTGCATGCCATAACAAAA	55
GBSSR-VF119	GTGGCCTGTACTGGTGGA	ACTCGTTGGGGCTAGGAA	57
GBSSR-VF131	CCGTACTAAATGAAGCCTTT	GGCAATCAAGTCCGGTAA	55
GBSSR-VF149	ACGACATGGTGATGAATCCT	ACGTGACCGAGTGACGAC	57
GBSSR-VF153	TCCCGACGCTACTTCTCA	CCGAGATCTGCAAACAGC	55
GBSSR-VF154	ACACCAATGTTTTTGCGG	TCCTGACTTTGCTGAGGC	55
GBSSR-VF159	GTGCCATCATCCTCGAAA	CAGCTGCTAGGTTGCCTG	57
GBSSR-VF164	ACCATTTGGCCTGTTCCT	CAAGGAGGGTTGTTTACGA	55
VfGSSR1	TTTCAGCAAACTAGAACCAATC	GGCATTCAGTTTTTACCTTGTA	50

(cont. on next page)

Table 2.4 (cont.).

VfGSSR3	TTCTTTGGTCCTCTCTCTATC	GCACTGTTGTTGCTGATACAA	51
VfGSSR4	AAGGGGAGGGCATAACAGAA	AATCCGCAAGGGTCTTCTT	52
VfGSSR9	GGTTTTGAATAGAAATGCAA	AAGATGTGTCAATATTGTTTT	50
VfGSSR10	ACCAAAACGCGCACTTATCA	AAGAGAGAGAAGAGAGCTTC	50
VfGSSR11	GCAAAAGGAGAGCAAGGGAA	CGAAAGAGGGGGACATTTTGT	52
VfGSSR13	GGTTGGGATCTTTTAGGTTGAA	TGGCCTTATATCCGTCCAAT	51
VfGSSR15	TCGATAGGGTTTCAGATTGA	GATGTTGACGGTGGTGTTT	50
VfGSSR19	AGCGATGGTGCTCATGCTTA	TCTCTCACGGAATCACATCTTT	52
VfGSSR27	CCCAAAAAGAGACGAACTGTAT	AGGGTTCATACGTTTGGCTT	51
VfGSSR31	ATAAGAGAGAACGAGGGAGAA	TTATGGTGGGACGTCTTACAT	51
VfGSSR34	GCACTCGAAGGAATTAATTTT	GAACAGTTGTTTCGTGTCGTA	51
VfGSSR41	AGCCCATGGTTCAAATGCAA	GCAGTCATGCCACTGCTTA	51
VfGSSR44	GATGTTGTTGGTGTTGTTTA	CAATTAGGAGCAAAATCAGA	50
VfGSSR47	CGATTGTTTGCAGAGGAGATA	ACAGAGAGGGACAGAGAGAA	52
VfGSSR53	GGTTCATGAAAAGAGGTTAG	CATTTTCCGTTCTCTCTCTA	50
VfGSSR67	GTTCATCAAGCACCAATCTAAAC	TCAATTTGGTTTATCTCTCTCTCT	52
VfGSSR69	ATTGGGGAGGATGAAGGTT	TTCCATTTTCCGTTCTCTCT	50
VfGSSR87	AGGGCCAGCGTGATCCAATA	TGGGTTGGGATCTTTTGGTTG	53

CHAPTER 3

RESULTS

3.1. SSR Development

This section is divided into three subsections, sequence assembly and simple sequence repeat identification, primer design and ssr validation, finally the SSR markers were tested in faba bean and polymorphic SSRs were applied to 46 diverse genotypes of faba bean to study genetic diversity and population structure.

3.1.1. Sequence Assembly and Simple Sequence Repeat Identification

Illumina sequencing of faba bean Filiz-99 resulted in 14,027,500 sequence reads comprising more than 4,208 Mb. Removal of adaptor and linker from sequence reads resulted in 4,182 Mb with an average cleaned sequence length of 298 nucleotides (nt). Any sequences larger than 200 nt were assumed to be a contig. As a result, 56,063 contigs were retrieved which encompassed 16.37 Mb, representing 0.13 % of the genome (Table 3.1). Contigs were mined for SSRs, resulting in the identification of 2138 SSRs. SSR length ranged between 6 and 32 nt, with an average length of 13 nt. Among the 2138 SSRs identified, mononucleotide repeats were the most abundant, representing 57.5 % of all SSRs. Dinucleotide repeats were the second most common type and accounted for 20.9 % of all SSRs. Trinucleotides represented 6.5 % of SSRs (Figure 3.1). The most common motifs were the A/T repeat (98.9 %) for mononucleotides, and AG/CT (26.6 %) for dinucleotides (Table 3.2), followed by GA/TC (23.2 %). Among trinucleotides, the most frequent repeats were AAT/ATT repeats were the most abundant (23.1 %) followed by ATA/TAT repeats (19.0 %).

Table 3.1. Preprocessing and assembly statistics for the faba bean genomic sequences.

Parameter	Raw Sequence	Cleaned Sequences	Contigs
Total number of sequences	14,027,500	14,024,398	56,063
Minimum sequence length (nt)	300	25	200
Maximum sequence length (nt)	300	300	39,301
Average sequence length (nt)	300	298	288.645
Total number of bases	4,208,250,000	4,182,709,546	16,367,617

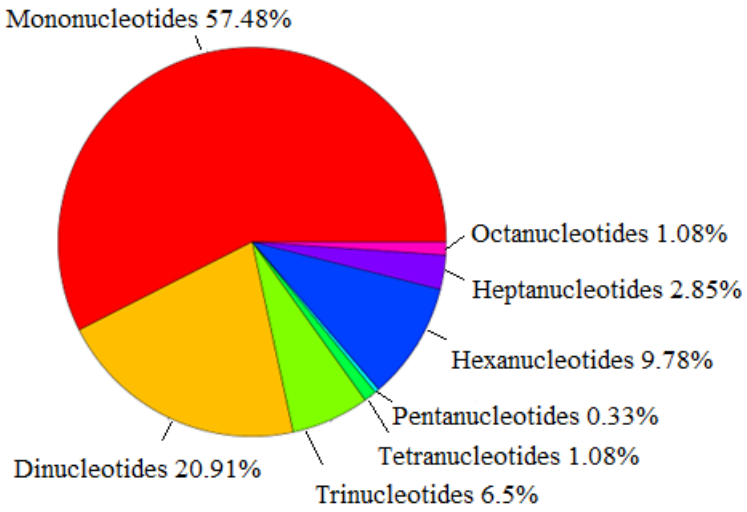

Figure 3.1. Pie chart of simple sequence repeat types in faba bean genome.

Table 3.2. Most abundant simple sequence repeat (SSR) motifs* in faba bean genomic sequence.

SSR Motif	Number of SSRs	Motif Frequency (%)
A/T	1216	98.9
TA	85	21.2
AT	77	17.3
GA/TC	103	23.2

(cont. on next page)

Table 3.2 (cont.).

AG/CT	119	26.6
AAT/ATT	28	23.1
ATA/TAT	23	19.0
TTA/TAA	18	14.8
TTC/GAA	13	10.7
AAAT/ATTT	4	26.6
AATA/TATT	4	26.6
ATAA/TTAT	3	13.3

*Motif frequencies are relative to SSR type. Only motif with a frequency ≥ 10% and number of SSRs > 1 are listed

3.1.2. Primer Design and SSR Validation

Of the 2138 SSRs identified in contigs, 430 met the requirements for primer design. We tested 50 of the designed primers for their amplification efficiency on two faba bean genotypes (ILB938/2 and Melodie/2). Primer sequences and repeat motifs are found in Table 3.3. Of these primers, 48 (96 %) amplified products. For SSR validation and to ensure that the expected SSRs were amplified by the primers, two faba bean samples were used for amplification of the SSR regions followed by sequencing. Genotype ILB938/2 was used as a DNA template for markers FbgSSR26 (TTGAAT), FbgSSR37 (AG), and FbgSSR309 (AG) while Melodie/2 was used as a DNA template for FbgSSR293 (ACCAAA) (Table 3.3). All four sequences contained the expected motifs (Figure 3.2 and Appendix A), proving that these primers amplified the expected SSRs.

Table 3.3. Simple sequence repeat (SSR) markers used for the molecular genetic analysis of faba bean.

SSR primers	Sequence	Repeat motif	Number of polymorphic alleles/total alleles	Gene diversity* (GD)
FbgSSR11-F	GAGTGAGGACAAATCAAGGT	(TTTCTG/CAGAAA)$_3$	0/1	0
FbgSSR11-R	AGGCAAACCTCTTGTTACAA			

(cont. on next page)

Table 3.3 (cont.).

FbgSSR26-F	GGTTGTGTCACTTTTCTTGG	(ATTCAA/TTGAAT)$_3$	1/2	0.12 ± 0.12
FbgSSR26-R	AATAAGACCTTAACTTTATTAACC			
FbgSSR29-F	ACTTCCAAAAATTTCAGAATCTC	(AAATTG/CAATTT)$_4$	7/7	0.35 ± 0.06
FbgSSR29-R	CCCAACTGAAGAAAAGGGTA			
FbgSSR30-F	TCCAAAAATTTCAGAATCTCCA	(AAATTG/CAATTT)$_4$	9/10	0.29 ± 0.05
FbgSSR30-R	CCCAACTGAAGAAAAGGGTA			
FbgSSR37-F	ATGCACGTTACAAGACATTG	(AG/CT)$_9$	9/9	0.38 ± 0.04
FbgSSR37-R	CTTTCCTCGCAAAAGGATTG			
FbgSSR109-F	CATGTCTCCTCACCATTTCA	(ATTG/CAAT)$_5$	0/1	0
FbgSSR109-R	TGTAGCGGAACTCAAATGAA			
FbgSSR140-F	TTCAAATGTAAACAGGCGTG	(AC/GT)$_6$	7/7	0.31 ± 0.05
FbgSSR140-R	ACCGTTGAGAGTAAAAGGAA			
FbgSSR198-F	TGAGACAAATCAGCATTCCA	(AGTTTTGA/TCAAAACT)$_3$	1/1	0
FbgSSR198-R	GCATTTGCATTCACATTTGG			
FbgSSR229-F	TTCTAGAATTGGTGCTCCTG	(TC/GA)$_7$	5/5	0.39 ± 0.04
FbgSSR229-R	TGCTTGAATATTGAGAGAAGT			
FbgSSR293-F	TGAGTGGAGATCTGCTAAGA	(ACCAAA/TTTGGT)$_3$	4/4	0.40 ± 0.03
FbgSSR293-R	AGCAATTGCATTCTAAAGCC			
FbgSSR306-F	CCACTCATTACCTTGAACCA	(GA/TC)$_8$	5/7	0.20 ± 0.06
FbgSSR306-R	CAACATCATCAGAAGCAACC			
FbgSSR309-F	GAACTATGAAGAGCAGCAGT	(CT/AG)$_9$	5/5	0.42 ± 0.06
FbgSSR309-R	AGTTGTTTACATGGACGTGT			
FbgSSR319-F	CTTCCGTCTTCTTTCCGTAT	(TGCAAG/CTTGCA)$_3$	0/1	0
FbgSSR319-R	ATAACTAATAGCAGCACCGG			
FbgSSR322-F	AAGGTGGTGGTGATTCAATT	(AAAATG/CATTTT)$_3$	5/5	0.26 ± 0.07
FbgSSR322-R	ATTTTATCTTGCCCATGGGT			
FbgSSR375-F	TTCAACCGGTAAAGAGAAGG	(CTTAGG/CCTAAG)$_3$	0/2	0
FbgSSR375-R	ACCAAAACTCTGATGGTGAA			
FbgSSR382-F	TGAGAAAGTTGAGTGACTGG	(GAATTG/CAATTC)$_3$	4/4	0.30 ± 0.05
FbgSSR382-R	ACCTTTGATAAATTGGAATAGA			
FbgSSR443-F	AAAACATCAATTTTGACTCAT	(TATTTAT/ATAAATA)$_3$	2/3	0.33 ± 0.16
FbgSSR443-R	TGAAGCAAATAAAATAACAGCAAG			
FbgSSR444-F	GCACCTGGCAAAATGATTTA	(AATTCTG/CAGAATT)$_3$	0/1	0
FbgSSR444-R	GCGTTTCAGCATTTTCAAAC			
FbgSSR451-F	GAACGACTTGAGAGAGAGTC	(TC/GA)$_6$	7/7	0.38 ± 0.06
FbgSSR451-R	TTTTAAACCCTAAGGACGGG			
FbgSSR518-F	AGTTCTCAAAGCGTTCTTCT	(AT/AT)$_6$	4/4	0.38 ± 0.03
FbgSSR518-R	GCTTGTATATTGTGTGAAGTCT			

(cont. on next page)

Table 3.3 (cont.).

FbgSSR520-F	GCTTGCAAGTAAGTGTGTTT	$(AG/CT)_6$	7/7	0.33 ± 0.04
FbgSSR520-R	GAAAGGTTGTGGTTGATTGG			
FbgSSR525-F	GGACACATCTCAATCATCCA	$(CAGTCA/TGACTG)_3$	4/5	0.23 ± 0.08
FbgSSR525-R	ACACATCTCTTGTTACAGCA			
FbgSSR545-F	TGAATTCTCTTCTCACGTGG	$(TC/GA)_6$	2/2	0.48 ± 0.01
FbgSSR545-R	CGAGTCAATTTGCACAAACT			
FbgSSR563-F	TTTATGAATTGGCGTTGTGG	$(CT/AG)_7$	9/9	0.32 ± 0.03
FbgSSR563-R	AACAAAACTCACCTTTCAATT			
FbgSSR564-F	TCCCTTTTGCTTGTTTATGA	$(CAT/ATG)_6$	2/3	0.19 ± 0.11
FbgSSR564-R	CCTCCGTGTTATCAAACAGT			
FbgSSR566-F	GCAAGAAGCAACATCCATTT	$(AC/GT)_6$	0/1	0
FbgSSR566-R	TTGCTTCAATCCTTCGAAGA			
FbgSSR599-F	TGTTTGGGACCTTTCTTTGA	$(TC/GA)_7$	0/1	0
FbgSSR599-R	GCAAGTCACCATCAAACAAA			
FbgSSR604-F	CGTTTTGGCTCATAATGCTT	$(TTCCTC/GAGGAA)_3$	4/4	0.33 ± 0.02
FbgSSR604-R	TTTTAGCCATGTACTGTGCT			
FbgSSR617-F	ATAGATGCCTCTCTCCATGT	$(GA/TC)_6$	0/1	0
FbgSSR617-R	GAAGGAGGACTAGACTGACT			
FbgSSR619-F	TATTTTAGTGGCCAGATGCA	$(TA/TA)_6$	4/5	0.34 ± 0.09
FbgSSR619-R	TGGAGAGGTGTTTCAACAAA			
FbgSSR623-F	AAAACCCATTTCTGGTACGA	$(AAACTA/TAGTTT)_3$	0/1	0
FbgSSR623-R	AGACAACCAACGTCGAATAA			
FbgSSR631-F	AATGTGATAAGCGCAACATG	$(ACTCTCA/TGAGAGT)_3$	2/2	0.23 ± 0.04
FbgSSR631-R	TTGGTATTTATCGCTTGTCT			
FbgSSR633-F	CTCCAAAACCAGAGTCTGTT	$(TCATCG/CGATGA)_3$	2/2	0.29 ± 0.09
FbgSSR633-R	TTTATCTGTAGAGGCATCGC			
FbgSSR643-F	GGCAAAAGATGGAGTCCTTA	$(ACAAAACT/AGTTTTGT)_3$	3/3	0.19 ± 0.07
FbgSSR643-R	TAATTTTTGGGCATTGGGAC			
FbgSSR663-F	ACTCGAAATCCATCAAGCAT	$(AG/CT)_6$	8/8	0.36 ± 0.04
FbgSSR663-R	GCTTTGTGCACCAACAATAT			
FbgSSR675-F	ATTGGGGAACTGCCTAATTC	$(GA/TC)_7$	6/6	0.40 ± 0.04
FbgSSR675-R	GCAATTTATCAAACACTTGGTG			
FbgSSR679-F	TGGATTGCATGCATGGTATA	$(TAGT/ACTA)_5$	4/5	0.31 ± 0.08
FbgSSR679-R	TCCAAAAGTCAGCTTGATGA			
FbgSSR695-F	GTTCTGTAAACACTAGGGCA	$(GA/TC)_8$	9/9	0.41 ± 0.04
FbgSSR695-R	TGTTGACGGTGATTTGTTTG			
FbgSSR734-F	CTCTTCTACAACGTCCCAAA	$(GTTGGT/ACCAAC)_3$	0/1	0
FbgSSR734-R	TCTCCTGAAAACGGTAAAGG			

*For each marker, average gene diversity ± standard error is presented

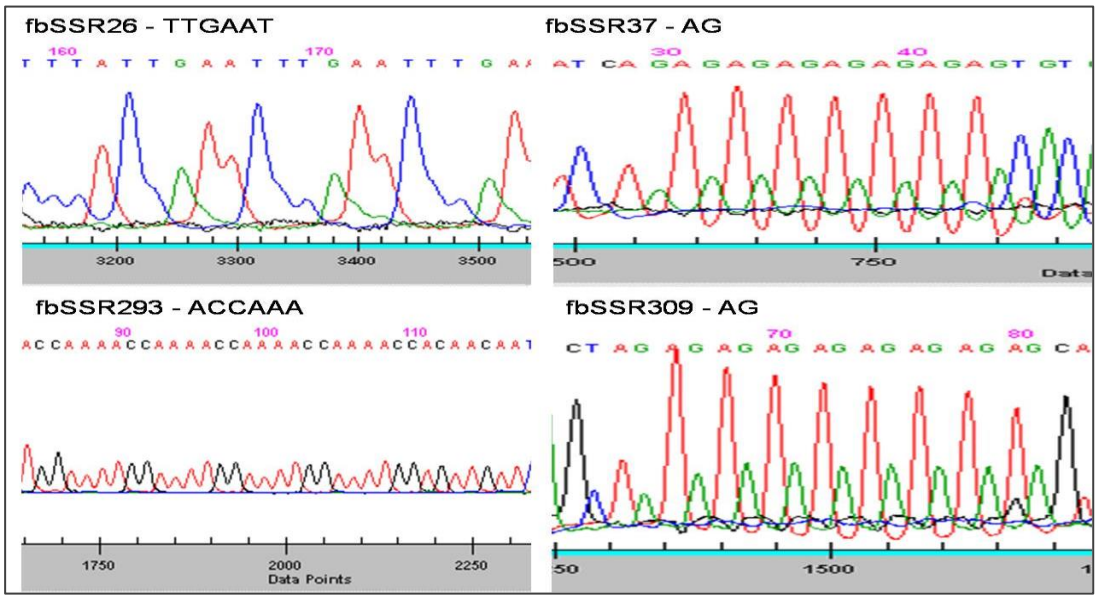

Figure 3.2. Sequencing electropherograms displaying the four simple sequence repeat motifs within the PCR amplicons.

3.1.3. Gene Diversity, Population Structure and Genetic Diversity Analyses Using SSR Markers

A total of 39 genomic SSR markers which showed clear amplification were applied to 46 faba bean genotypes (Figure 3.3). Turkey was represented by the most genotypes (13) with 5 genotypes each from Finland and Australia. Germany, the Netherlands, Syria and Egypt were represented by 3 genotypes with the remaining 11 genotypes from other countries (Table 3.4). Thirty-one of 39 markers (79 %) were polymorphic. The SSR primers generated 161 alleles, 141 (87.6 %) of which were polymorphic (Table 3.3). The average number of amplified fragments per genomic SSR marker was 4.1, with a range of 1–10 alleles. The average gene diversity value (also called polymorphism information content, PIC) of the markers (based on a calculation that ranges from 0.0 to 0.5) was 0.27, with the highest value calculated for FbgSSR545 (0.48 ± 0.01). The lowest value was zero for monomorphic markers (Table 3.3).

The SSR data were used to study the population structure and genetic diversity of the 46 faba bean accessions. Population structure assigned the genotypes to two

subpopulations (Table 3.4; Figure 3.4), the first subpopulation included 18 genotypes, the second included 16 genotypes while 12 genotypes were assigned as admixed. All genotypes from Germany and Finland were assigned to subpopulation 1. Australian genotypes were assigned to subpopulation 2 while the Turkish genotypes were distributed between subpopulations 1, 2 and admixed.

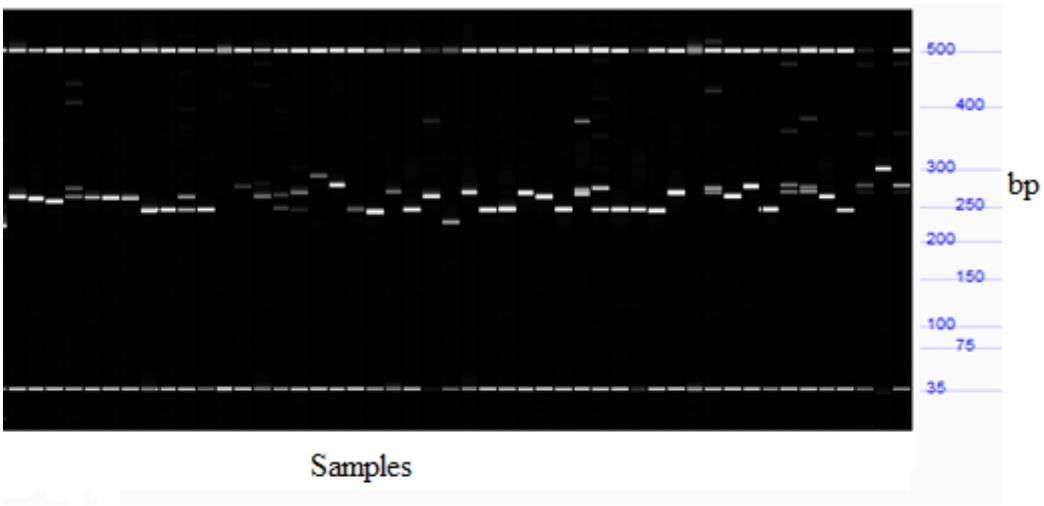

Figure 3.3. Example of PCR fragments using capillary electrophoresis of 46 faba bean accessions by FbgSSR293 primer.

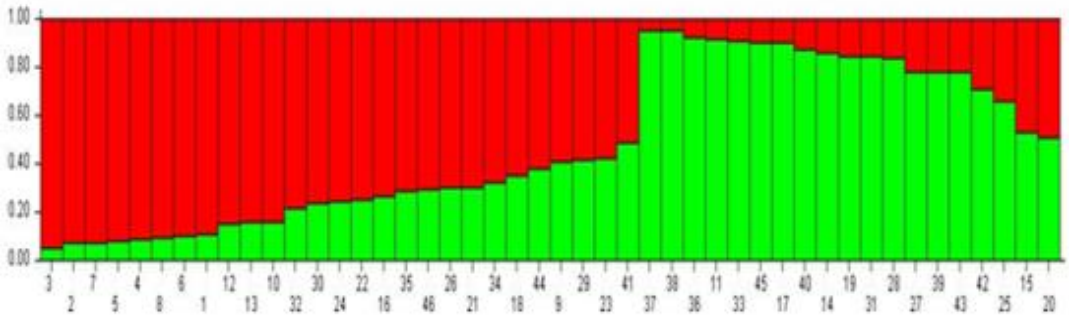

Figure 3.4. Bar plot showing genetic structure of 46 faba bean genotypes with K = 2. Each genotype is represented by a vertical bar which is partitioned according to estimated membership in the two clusters. Cluster A fraction is shown in red and cluster B fraction in green.

Table 3.4. Faba bean genotypes used in the study. Cluster assignments of 46 faba bean genotypes according to Structure and DARwin analyses.

No.	Sample Name	Origin	Inferred ancestry subpopulation		Subpopulation assignment*	Cluster**
			1	2		
1	NGB8642	Finland	0.887	0.113	1	A
2	NGB1547.1	Finland	0.928	0.072	1	A
3	Mikko	Finland	0.944	0.056	1	A
4	Witkiem manida	Germany	0.911	0.089	1	A
5	Ukko	Germany	0.915	0.085	1	A
6	Kontu	Germany	0.899	0.101	1	A
7	NGB1542.1	Finland	0.924	0.076	1	A
8	NGB1548.2	Finland	0.901	0.099	1	A
9	CGN7874	Spain	0.591	0.409	Admixed	A
10	CGN15563	Syria	0.838	0.162	1	A
11	CGN15619	Egypt	0.084	0.916	2	B
12	CGN13485	Pakistan	0.850	0.150	1	C
13	CGN13464	UK	0.841	0.159	1	A
14	CGN10391	Egypt	0.140	0.860	2	B
15	CGN7826	Greece	0.470	0.530	Admixed	B
16	CGN7716	Italy	0.734	0.266	1	A
17	CGN7844	Jordan	0.097	0.903	2	B
18	CGN7781	Netherland	0.647	0.353	Admixed	A
19	CGN15641	Netherland	0.153	0.847	2	B
20	CGN10382	Turkey	0.486	0.514	Admixed	A
21	CGN10371	Algeria	0.694	0.306	Admixed	B
22	CGN18892	Netherland	0.750	0.250	1	A
23	CGN07875	India	0.572	0.428	Admixed	C
24	CGN10385	Turkey	0.757	0.243	1	A
25	CGN10374	Syria	0.340	0.660	Admixed	A
26	CGN10325	Syria	0.698	0.302	Admixed	A
27	TR23018	Turkey	0.216	0.784	2	C
28	TR31590	Turkey	0.158	0.842	2	B
29	TR33140	Turkey	0.583	0.417	Admixed	C
30	TR37255	Turkey	0.763	0.237	1	A
31	TR44876	Turkey	0.153	0.847	2	C
32	TR44928	Turkey	0.785	0.215	1	A
33	TR49380	Turkey	0.088	0.912	2	B

(cont. on next page)

Table 3.4 (cont.).

34	TR53748	Turkey	0.673	0.327	Admixed	A
35	TR61267	Turkey	0.713	0.287	1	A
36	Ascot	Australia	0.076	0.924	2	A
37	Manafest	Australia	0.043	0.957	2	B
38	Fiord	Australia	0.050	0.950	2	B
39	Fiesta	Australia	0.220	0.780	2	A
40	Aquadulce	Australia	0.128	0.872	2	C
41	Filiz-99	Turkey	0.512	0.488	Admixed	B
42	Salkım	Turkey	0.291	0.709	2	C
43	26139	.Colombia	0.220	0.780	2	B
44	26145	Egypt	0.615	0.385	Admixed	A
45	ILB938/2	Unknown	0.096	0.904	2	B
46	Melodie/2	France	0.702	0.298	1	A

a Genotypes were assigned to subpopulations based on the proportion of inferred ancestry with a threshold of ≥ 0.70

b Cluster assignments based on the neighbor-joining dendrogram

A dendrogram was drawn using the Dice coefficient and the unweighted neighbor-joining algorithm (Figure 3.5). Average pairwise dissimilarity among the 46 faba bean accessions was 0.29, with the highest value, 0.46 (54 % similarity), calculated between accessions from Turkey (TR37255) and Greece (CGN07826). The lowest dissimilarity was 0.16 (84 % similarity) calculated between Australian accessions (Manafest and Fiord). The faba bean accessions grouped into three clusters (A, B, and C) in the dendrogram (Figure 3.5). Cluster A included 26 accessions and dissimilarity ranged from 0.16 to 0.40 with an average was 0.28. Cluster A had four subclusters, one of these subclusters had all the accessions from Finland and two accessions from Germany. Cluster B had 13 accessions and cluster C had 7 accessions. The Turkish accessions were distributed to all three clusters A, B, C (Table 3.4; Figure 3.5). Dendrogram and population structure analyses showed high correspondence. All of subpopulation 1 coincided with cluster A, except one accession from Pakistan which grouped in cluster C. Also, all of subpopulation 2 coincided with cluster B, except for four accessions which grouped in cluster C (Table 3.4).

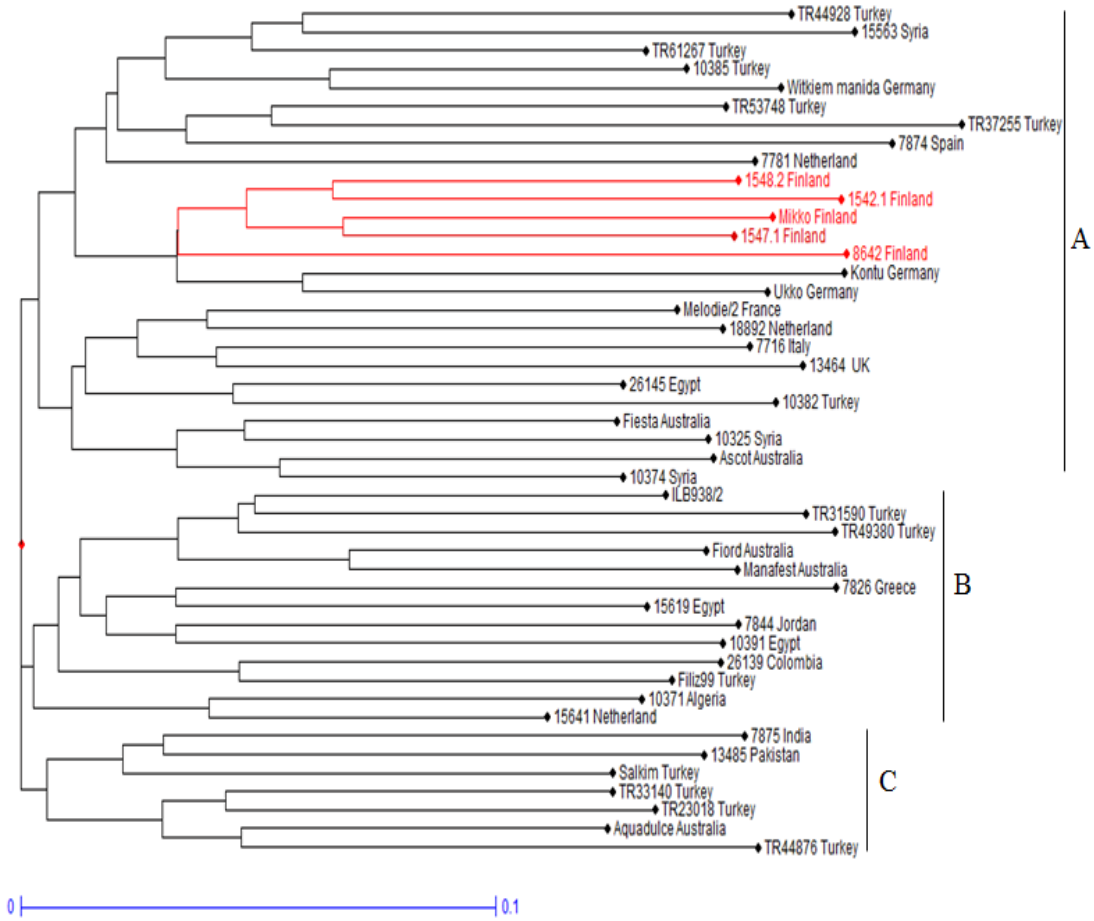

Figure 3.5. Unweighted neighborjoining dendrogram of 46 faba bean accessions based on 161 simple sequence repeat (SSR) alleles. Accession names and origins are provided.

3.2. Morphological Trait Diversity

The morphological traits (quantitative and qualitative) of 114 faba bean genotypes (61 landraces and 53 cultivars) were characterized.

3.2.1 Quantitative Characteristics

Fourteen quantitative characters were assessed in the 114 faba bean genotypes over two seasons (Appendix B). Means, ranges and CVs were calculated for all of the material and for the 61 landraces and 53 cultivars separately (Table 3.5). Three vegetative growth traits

were examined: number of leaflets per leaf, number of stems per plant and plant height. Leaflets per leaf was observed on fully expanded leaves at the median flowering node. Leaflet number ranged from 4.2 to 6.4 with average of 5.4 leaflets for the landraces. Similar variation and values in leaflet number were seen in the cultivars: 3.9 to 6.6 with an average of 5.5 leaflets per leaf. There was more variation in stem number per plant. Two Turkish landraces had the most stems: CGN10382 and TR44862 with 11.5 and 8 stems, respectively. CGN7719 from Italy and the French cultivar Melodie/2 had the fewest stems, 2.6 and 2.7, respectively. Overall, the 114 genotypes averaged 4.8 stems per plant. Among all of the genotypes, CGN7730 (Czech Republic) and CGN07781 (Netherlands) were the tallest with heights of 155.5 and 148.5 cm, respectively. The shortest were NGB8640 (Finland) and TR71255 (Turkey) which were approximately 40% shorter than the overall average height with values of 62.8 and 68.9 cm, respectively. Cultivars and landraces had average heights of 108.1 and 105.6 cm respectively.

Two quantitative flowering traits were assessed: days to 50% flowering and number of flowers per inflorescence. Days to flowering for all genotypes ranged from 59.0 to 132.0 d with an average of 91.7 d. NGB1547.1 from Finland was the earliest flowering accession and reached 50% flowering 59 days after seed sowing. On average, the cultivars flowered slightly, but not significantly, later than the landraces, 94.4 and 89.4 d, respectively. The number of flowers per inflorescence ranged from 2.6 in the Finnish cultivar Mikko to 8.1 in the Czech cultivar CGN7730.1 with average value of 5.0 flowers per inflorescence over all genotypes. The average for landraces and cultivars were 4.9 and 5.2, respectively, an insignificant difference.

Three pod quantitative characters were measured: pod length, number per node and number per plant. Overall average pod length was 9.7 cm with the shortest pods in CGN07751 (unknown origin) and CGN13487 (Pakistan): 6.0 and 6.2 cm, respectively. The longest pods were observed in the Turkish cultivar CGN19987 (17.3 cm) followed by the Spanish cultivar CGN19979 with 15.7 cm pods. The cultivars' pods were significantly (t test, $P < 0.0001$) longer than landraces with average value of 10.5 and 9 cm, respectively. The number of pods per node varied only slightly from 1.0 to 2.4 with an average of 1.4 pods. The total number of pods per plant for landraces was slightly more than cultivars with an overall average of 27.2 pods. Wide variation was observed for this trait in both landraces and cultivars (CVs of 58.4 and 43.5%, respectively). The landraces CGN7827 (Ethiopia) and CGN10382 (Turkey) had the most pods per plant with 92.6 and

Table 3.5. Quantitative morphological traits for the 114 faba bean genotypes (61 landraces and 53 cultivars). Means, CV, and range are given for landraces and cultivars.

	Landraces			Cultivars			Landraces and cultivars	
Trait	Mean ± SE	CV (%)	Range	Mean ± SE	CV (%)	Range	Mean ± SE	CV (%)
Leaflets per leaf	5.4 ± 0.1	9.1	4.2–6.4	5.5 ± 0.1	10.9	3.9–6.6	5.5 ± 0.1	10.1
Stems per plant	4.8 ± 0.2	31.2	2.6–11.5	4.7 ± 0.1	22.3	2.7–7.2	4.8 ± 0.1	27.5
Plant height, cm	105.6 ± 2.4	17.7	62.8–148.5	108.1 ± 2.7	17.9	73.4–155.5	106.8 ± 1.8	17.8
Days to flowering	89.4 ± 2.8	24.5	59–132	94.4 ± 2.4	18.3	62.5–132	91.7 ± 1.9	21.8
Flowers per inflorescence	4.9 ± 0.1	18.9	3–7	5.2 ± 0.1	19.3	2.6–8.1	5 ± 0.1	19.3
Pod length, cm	9 ± 0.2	17.9	6–12.5	10.5 ± 0.3	22.9	7.3–17.3	9.7 ± 0.2	22.3
Number of pods per node	1.4 ± 0.03	18.5	1–2.1	1.4 ± 0.04	22.9	1–2.4	1.4 ± 0.03	20.7
Number of total pods per plant	28.4 ± 2.1	58.4	5–92.6	25.7 ± 1.5	43.5	7.8–63.9	27.2 ± 1.3	53.0
Maximum ovules per pod	3.4 ± 0.1	11.6	2.7–4.4	3.7 ± 0.1	14.1	2.8–6.1	3.6 ± 0.1	13.6
Number of seeds per pod	3.3 ± 0.1	11.6	2.5–4.2	3.5 ± 0.1	12.5	2.7–4.5	3.4 ± 0.04	12.4
Seeds fertilized, %	95.7 ± 0.6	4.6	79.2–100	94.3 ± 0.9	6.8	69.3–100	95 ± 0.5	5.7
Water content of fresh seeds, %	72.1 ± 0.7	7.0	57.9–82.2	73.3 ± 0.5	5.1	65.1–81.1	72.7 ± 0.4	6.3
100-seed weight, g	72.9 ± 3.1	33.4	30.8–130.6	83.1 ± 5	43.8	29.4–174.6	77.7 ± 2.9	39.9
Seed yield, g	55.2 ± 4.3	60.4	12.3–158	52.4 ± 3.6	50.6	17.7–174.1	53.9 ± 2.8	56.4

77.5 pods, respectively. CGN10385 (Turkey) and CGN7719 (Italy) had the fewest pods: 5.0 and 7.1, respectively. Notably, CGN7719 also had very few stems perhaps explaining its lack of pod-set.

Six ovule/seed traits were examined. The average maximum number of ovules per pod was 3.6 with a range of 2.7 to 6.1 ovules over all 114 genotypes. There was a significant difference (t test, $P < 0.0005$) between landraces and cultivars with averages of 3.4 and 3.7 ovules per pod, respectively. Number of seeds per pod had high values of 4.5 and 4.4 for CGN19979 (Spain) and CGN18862 (Netherlands) cultivars, respectively, both of which had significantly longer pods than average. The fewest seeds per pod were produced by CGN07734 (Ethiopia) and TR44862 (Turkey) landraces with 2.5 to 2.7 seeds per pod, respectively. Interestingly, TR44862 was one of the two landraces that produced the most stems, therefore, seed production may have been sacrificed for vegetative growth. On average, the genotypes produced 3.4 seeds per pod with a significant difference (t test, $P < 0.01$) between landraces and cultivars (3.3 and 3.5 seeds per pod, respectively). Percent fertilization (development of the ovules into mature seeds) varied from 69.3 to 100.0% with an average of 95% and no significant difference between landraces and cultivars was observed. The average water content of fresh seeds was 72.7% and ranged from 57.9 to 82.2% with similar averages for landraces and cultivars. Seed weight was determined for a bulk sample of 100 seeds and showed considerable variability in the accessions (CV 39.9%). CGN19985 (unknown origin) had the greatest 100-seed weight (174.6 g) followed by three Turkish cultivars: Eresen87, Filiz99, and Salkim (approximately 156 g). Interestingly, cultivars also had the lowest weights: Mikko (29.4 g) and Kontu (30.5 g). Average 100-seed weight was 77.7 g. The majority of genotypes (63.1%) could be classified as medium-seeded with 100-seed weight ranging from 51.1 to 95.4 g. Large-seeded types made up 21.9% of genotypes with values from 100.5 to 174.6 g. The third group (14.9%) consisted of small-seeded individuals with values from 29.4 to 48.6 g. The 100-seed weight of cultivars averaged 83.1 g heavier than that of landraces (72.9 g). Dry seed yield was the most variable quantitative trait with more variation in the landraces (CV 60.4%) than cultivars (CV 50.6%). Both groups had average dry seed yield of approximately 54 g per plant. CGN18906 (UK) and CGN7827 (Ethiopia) had the highest yields with 174.1 and 158.0 g dry seeds per plant, respectively. This was not an unexpected result for CGN7827 as this genotype also had the most pods per plant. The Finnish landraces NGB1542.1 and NGB8642 had the lowest yields with 12.3 and 14.0 g dry seeds per plant, respectively.

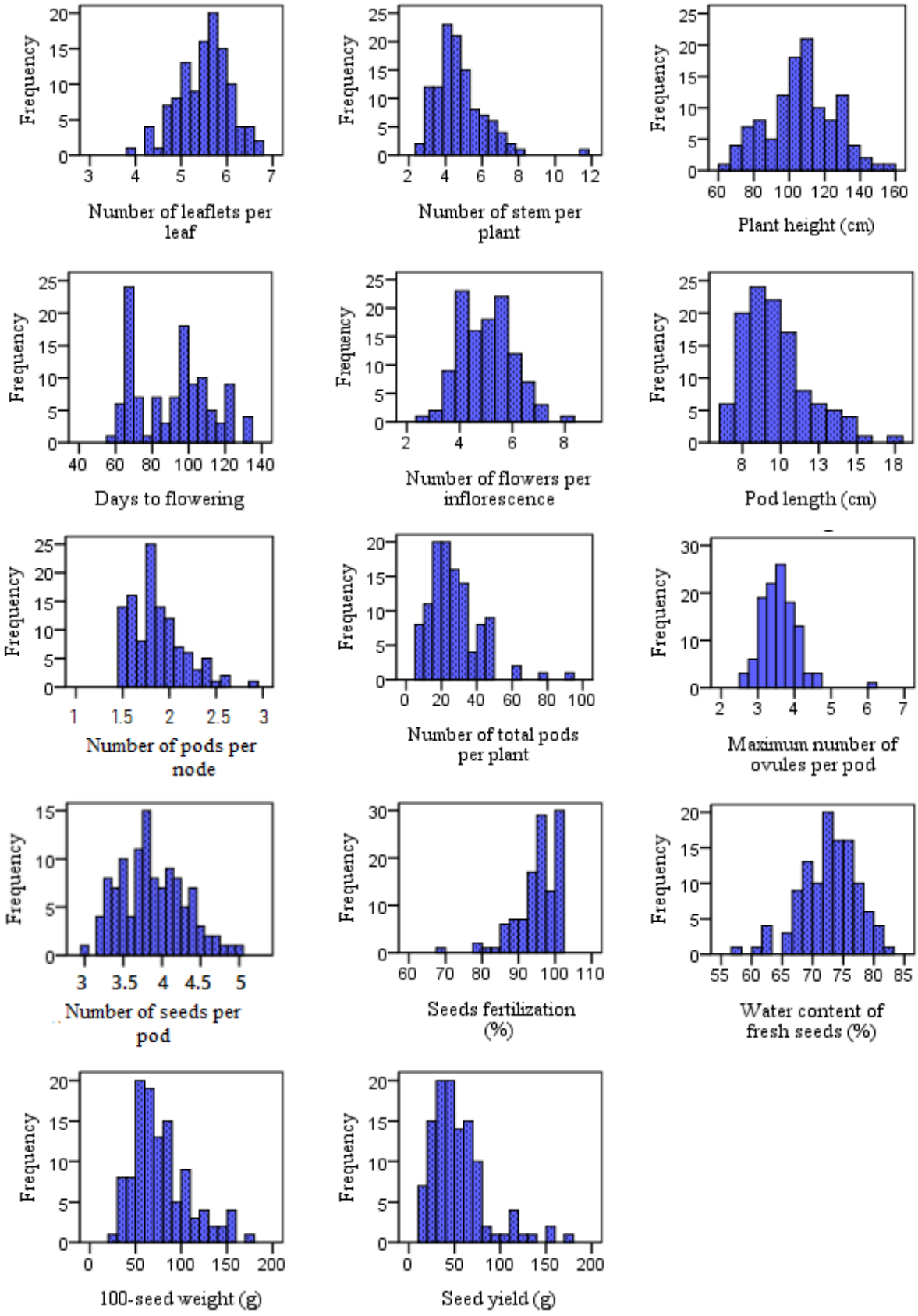

Figure 3.6. Histograms of 14 quantitative traits within 114 faba bean genotypes.

3.2.2. Qualitative Characteristics

The data for 12 qualitative traits were separately analyzed for the 61 landraces and 53 cultivars over two years (Table 3.6; Appendix B). Four vegetative traits were examined: leaflet shape, stipule spot pigmentation, stem pigmentation at flowering and mature stem color. Intermediate leaflet shape was most common for both landraces and cultivars accounting for 70.4 and 66.0% of the genotypes, respectively. The landraces and cultivars had similar percentages of round leaflets. Fewer of the landraces had narrow leaflets (14.8%) as compared to the cultivars (18.9%). All landraces had brown spots on the stipule, while 9.4% of the cultivars lacked this brown pigment. Stem pigmentation indicates the presence of anthocyanin and was variable in the population. Most genotypes had intermediate pigmentation (49.2% for landraces and 46.2% for cultivars). Most of the other accessions had weak or strong stem pigmentation (approximately 25% each) with only a few cultivars completely lacking pigmentation. Mature stem color was similar for both landraces and cultivars with the majority of genotypes having dark color (82.0 and 84.9% respectively).

Table 3.6. Qualitative morphological traits in the faba bean genotypes. Percentages in each trait class are given for landraces and cultivars.

Trait	Class: description	landraces (%)	Cultivars (%)	Landraces and Cultivars (%)
Leaflet shape	Narrow	14.8	18.9	16.7
	Intermediate	70.4	66.0	68.4
	Round	14.8	15.1	14.9
Stipule spot pigmentation	Absent	0.0	9.4	4.4
	Present	100.0	90.6	95.6
Stem pigmentation at flowering	Absent	0.0	7.5	3.5
	Weak	26.2	21.2	23.9
	Intermediate	49.2	46.2	47.8
	Strong	24.6	25.0	24.8
Mature stem color	Light	18.0	15.1	16.7
	Dark	82.0	84.9	83.3
Intensity of petal streaks	No streaks	0.0	7.7	3.5
	Slight	27.9	13.5	21.2
	Moderate	41.0	42.3	41.6
	Intense	31.1	36.5	33.6
Wing petal color	Uniform white	0.0	9.4	4.4
	Spotted	100.0	90.6	95.6
Pod angle	Erect	85.2	67.9	77.2
	Horizontal	13.1	22.6	17.5

(cont. on next page)

Table 3.6 (cont.).

	Pendent	1.6	9.4	5.3
Pod shape	Sub-cylindrical	68.9	67.9	67.5
	Flattened	31.1	32.1	32.5
Pod surface reflectance	Matte	27.9	26.4	27.2
	Glossy	72.1	73.6	72.8
Seed shape	Flattened	50.8	50.9	50.9
	Round	49.2	49.1	49.1
Seed coat color	Black	6.6	1.9	4.6
	Dark brown	3.3	3.8	3.7
	Light brown	52.4	52.8	53.7
	Light green	23.0	26.4	25.9
	Dark green	1.6	0.0	0.9
	Violet	6.6	1.9	4.6
	White	0.0	9.4	6.5
	Mixed	6.6	3.8	5.3
Hilum color	Black	86.9	75.5	81.6
	Colorless	4.9	9.4	7.0
	Mixed	8.2	15.1	11.4

Two flower characteristics were assessed. The intensity of streaks on the standard petal was most often moderate followed by intense for both landraces and cultivars. Only cultivars had unstreaked petals (7.7%). All of the landraces had spotted wing petals; however, 9.4% of the cultivars had uniformly white wing petals.

Three pod traits were examined: pod angle, shape and surface reflectance. Pod angle attitude at maturity on the second pod-bearing node was determined. The majority had erect pods: 85.2 and 67.9% for landraces and cultivars, respectively. Horizontal pods were observed in 13.1% of landraces and 22.6% of cultivars. Only a few landraces and cultivars had pendent pods (1.6 and 9.4%, respectively). Pod shape was divided into two groups (sub-cylindrical and flattened) with similar distributions in the landraces and cultivars. The majority of genotypes (approximately 68%) had sub-cylindrical pods. Most genotypes had pods with glossy surface reflectance with averages of 72.1 and 73.6% for landraces and cultivars, respectively. The remaining plants had matte pods.

Seeds were characterized for three qualitative traits: shape, color and hilum color. Seed shape of the landraces and cultivars was equally divided between flattened and round. All of the landraces and cultivars with 100-seed weight above 88.8 g had flattened shape. There was high diversity among faba bean genotypes for seed coat color (Figure 3.7. A-G). The predominant color for landraces and cultivars was light brown (52.4 and 52.8%, respectively), followed by light green (23.0 and 26.4%, respectively) with only white seeds seen in cultivars. Indeed, white was the third most common seed color in cultivars (9.4%). Landraces had equal proportions of black, violet, and mixed seeds

(6.6%). The majority of landraces and cultivars had black hila: 86.9, and 75.5%, respectively (Figure 3.7. H). Seeds with colorless hilum were seen in both landraces and cultivars (4.9 and 9.4% respectively). All white seeds had colorless hila and white wing petals.

A: light brown B: Dark brown C: Light green D: Dark green

E: Black F: Violet G: White H: Hilum color

Figure 3.7. A-G, Seed coat color diversity of 114 faba bean genotypes. H: seed with black hilum (a) and colorless hilum (b).

3.2.3. Correlations between Characters

Spearman correlation coefficients between the 26 morphological characters were computed for the two-season averages (Appendix B). The highest positive significant correlation was between stipule spot pigmentation and wing petal color (r = 1.00) followed by maximum number of ovules and seeds per pod (r = 0.88). Hilum color had the same negative significant correlation (r = -0.78) with both stipule spot pigmentation and wing petal color. Also, a strong negative correlation was detected between 100-seed weight and seed shape such that larger seeds were more often flattened and not round (r = - 0.72). There were significant positive correlations between pod length and many characters: 100-seed weight, pod angle, number of ovules per pod, seed coat color, and pod shape (r = 0.69 to 0.44). Pod length was also significantly negatively correlated with seed shape (r = -0.57) and pods per node (r = -0.43). Plant dry seed yield was significantly and positively associated with pod number, plant height and stem number (r = 0.61 to 0.33). One hundred-seed weight was negatively correlated with pods per node and pods

per plant (r = -0.55 and -0.36, respectively) and also was positively correlated with pod angle, pod shape and seed coat color (r = 0.54 to 0.35). A negative correlation was seen between pod angle and seed shape (r = -0.49). A moderate positive correlation was detected between stem pigmentation at flowering and petal streak intensity (r = 0.46).

3.2.4. Principle Component Analysis

Principal component analysis reduced the 14 quantitative morphological variables into six components which accounted for 77.48% of the variance in the data. (Table 3.7). Pod length and 100-seed weight were the greatest positive contributors to the first component, while pods per node had the strongest negative association. The second component was most positively correlated to maximum ovules per pod and number of seeds per pod. Stem per plant, total pods per plant, and seed yield were positive contributors to the third component. All the components are listed in Table 3.7 with Eigen values, variance (%), and Cumulative of variance (%).

Table 3.7. Principal component analysis of quantitative traits. Eigen values are given for the six principal component (PC) axes.

Variable	Component axes					
	PC-1	PC-2	PC-3	PC-4	PC-5	PC-6
Leaflets per leaf	0.195	-0.036	0.118	-0.475	0.010	0.598
Stems per plant	0.117	-0.291	0.637	-0.190	0.368	-0.054
Plant Height	0.076	0.150	0.400	0.606	0.256	0.146
Days to flowering	-0.265	0.030	-0.060	-0.056	0.775	-0.187
Flowers per inflorescence	-0.076	0.111	-0.095	0.165	-0.004	0.840
Pod length	0.807	0.440	0.050	0.005	-0.014	-0.063
Number of pods per node	-0.706	0.023	0.094	0.078	0.094	-0.268
Number of total pods per plant	-0.424	-0.071	0.838	0.106	-0.093	-0.006
Maximum ovules per pod	0.151	0.947	-0.027	-0.214	0.018	0.059
Number of seeds per pod	0.041	0.925	-0.046	0.255	-0.004	0.054
Seed fertilized	-0.205	-0.043	0.009	0.798	-0.002	-0.010
Water content of fresh seeds	0.227	-0.009	0.082	0.225	0.786	0.209
100-seed weight	0.899	-0.008	0.080	-0.193	0.071	-0.153
Seed yield	0.187	0.100	0.873	0.143	-0.069	-0.026
Eigen values	2.91	2.25	2.01	1.39	1.23	1.05
% of variance	20.79	16.10	14.37	9.93	8.80	7.49
Cumulative % of variance	20.79	36.89	51.27	61.19	70.00	77.48

3.3. Biochemical Characterization

Seven biochemical characters (phenolic acids, flavonoids, protein, L-DOPA, tannins, vicine and convicine content) were assessed in the 114 faba bean genotypes over two seasons and all measurements are listed in Appendix C. Means, ranges and CVs were calculated for all of the material and for the 61 landraces and 53 cultivars separately (Table 3.8).

Table 3.8. Chemical characterization for the 114 faba bean genotypes. Means, % CV, and range are given for landraces and cultivars.

Trait	Landraces			Cultivars			Landraces and cultivars	
	Mean ± SE	CV (%)	Range	Mean ± SE	CV (%)	Range	Mean ± SE	CV (%)
Phenolics (mg GAE /100 g DW)	389.2 ± 6.9	13.8	280.2–559.2	383.2 ± 6.9	17.3	261.1–554.9	386.4 ± 6.9	15.5
Flavonoids (mg RE/100 g DW)	221.6 ± 5.5	19.5	120.2–306.6	217.4 ± 7.0	23.5	130–343.2	219.6 ± 4.4	21.4
Protein (%)	26.4 ± 0.7	20.0	16.8–40.8	25.4 ± 0.5	14.7	18.8–35.7	26.0 ± 0.4	18.0
L-DOPA (mg/100g DW)	46.9 ± 2.6	43.0	9.2–100.9	46.4 ± 2.6	44.1	18.6–113.4	46.6 ± 2.6	43.5
Tannin (mg CE/100 g DW)	155.7 ± 7.2	36.3	61.5–396.1	140.9 ± 7.2	41.8	11.9–277.1	146.6 ± 7.2	36.3
Vicine (%)	0.36 ± 0.01	20.2	0.15–0.58	0.33 ± 0.01	23.9	0.02–0.66	0.35 ± 0.01	20.6
Convicine (%)	0.21 ± 0.01	28.6	0.09–0.39	0.21 ± 0.01	28.8	0.01–0.34	0.21 ± 0.01	27.1
Total VC (%)	0.58 ± 0.01	17.6	0.34–0.91	0.5 ± 0.01	19.5	0.03–0.94	0.54 ± 0.01	17.3

3.3.1. Total Phenolics Content

Phenolic compounds represent a large group of molecules with a variety of functions in plant growth, development, and defense. In this study the Folin Ciocalteu method was used to estimate phenolic compound content in dry seeds. Gallic acid was used (20–300 mg/L) to obtain a calibration curve (Appendix C). Total phenolics contents values of the 114 faba bean genotypes are given in Appendix C. Contents ranged from 261.1 for CGN19993 (Netherlands) to 559.2 mg GAE/100 g for CGN10382.1 (Turkey) with an average of 386.4 mg GAE/100 g DW (Table 3.8). No significant difference was detected between landraces and cultivars. The frequency distribution for the trait shows an approximately normal distribution (Figure 3.8)

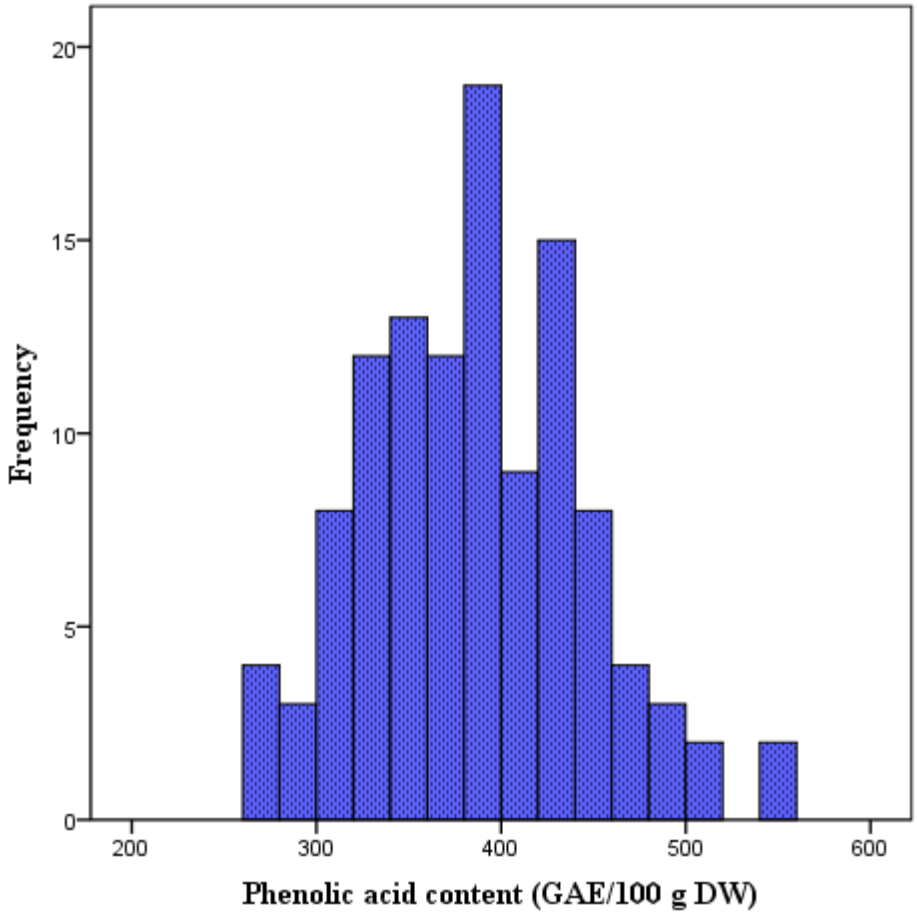

Figure 3.8. Frequency distribution of phenolics content for 114 faba bean genotypes.

3.3.2. Flavonoids Content

Flavonoids content of the accessions was measured with spectrophotometry using rutin (0.01 to 1.0 mg/ml) for construction of a calibration curve (Appendix C). The total flavonoids ranged from 120.2 for CGN10362 (Turkey) to 343.2 for CGN7730.1 (Czech Rep.) with an average of 219.6 mg RE/100 g dry weight. No significant difference was observed between landraces and cultivars. All concentration values are given in Appendix C and Figure 3.9 shows the frequency distribution for the genotypes.

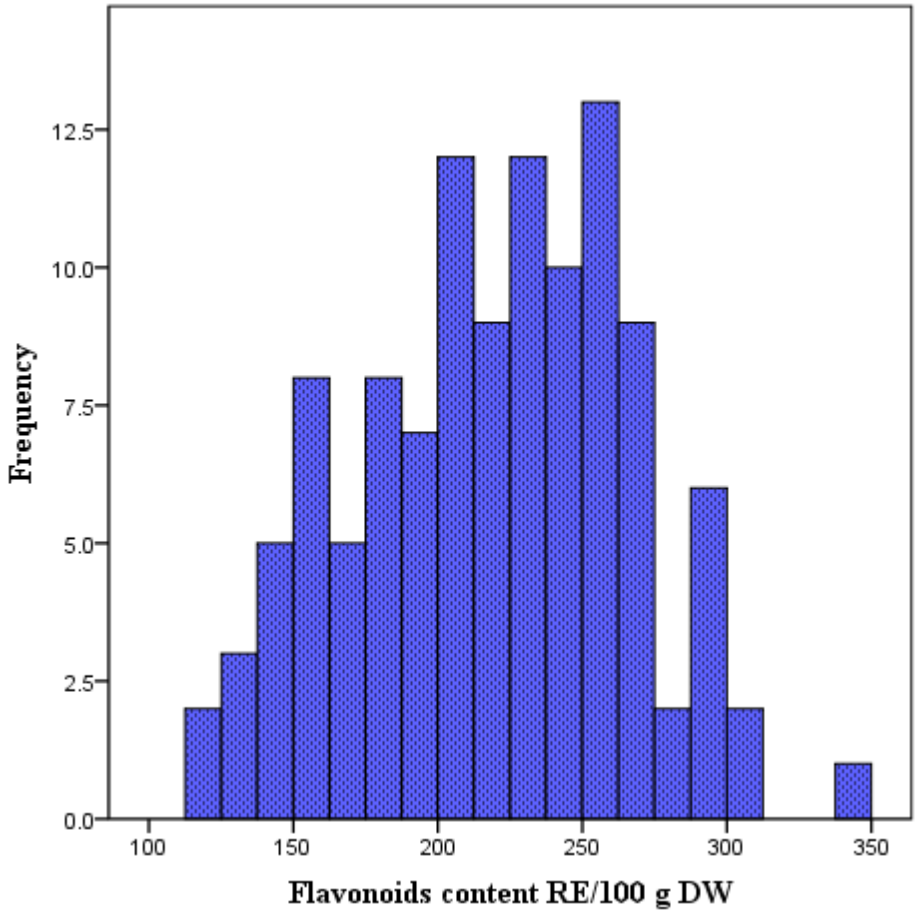

Figure 3.9. Frequency distribution of flavonoids content for 114 faba bean genotypes.

3.3.3. Protein Content

To determine faba bean seed protein content, a BSA calibration curve with linearization was constructed (Appendix C). Among the 114 faba bean genotypes evaluated in this study, the range of dry seed total protein content was 16.8% for CGN10362 (Turkey) to 40.8% for NGB20019.2 (Finland) with an average of 26.0% with no significant difference between landraces and cultivars. Protein content had more variation in the landraces (CV 20.0%) than cultivars (CV 14.7%). All protein concentration values are presented in Appendix C. The frequency distribution (Figure 3.10) shows that very few accessions had extremely high levels of protein.

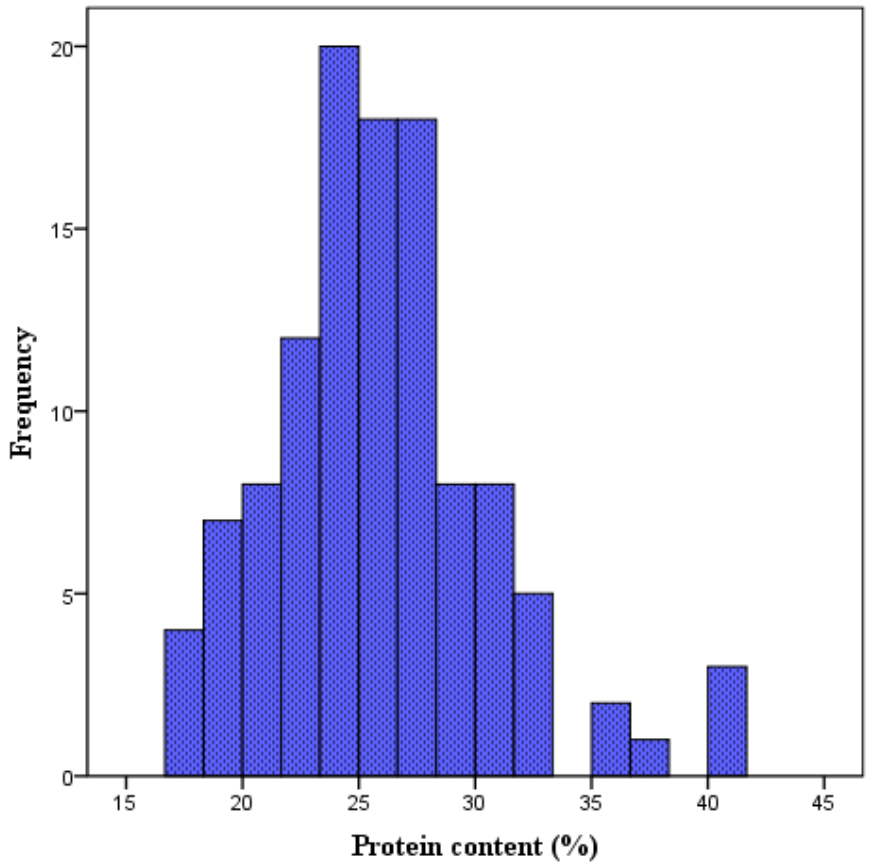

Figure 3.10. Frequency distribution of protein content for 114 faba bean genotypes.

3.3.4. L-DOPA Content

L-DOPA had the highest variation for the biochemical traits (CV = 43.5%). L-DOPA concentration ranged from 9.2 to 113.4 for CGN07751 (Unknown) and CGN18909 (Germany) genotypes, respectively, with an average of 46.6 mg/100 g of seed dry weight. No significant difference was seen in the L-DOPA contents of landraces versus cultivars. All the values are listed in Appendix C. Figure 3.11 shows the frequency distribution of L-DOPA concentration values. The majority of genotypes (74%) had contents ranging from 20 to 60 mg/100 g dry weight.

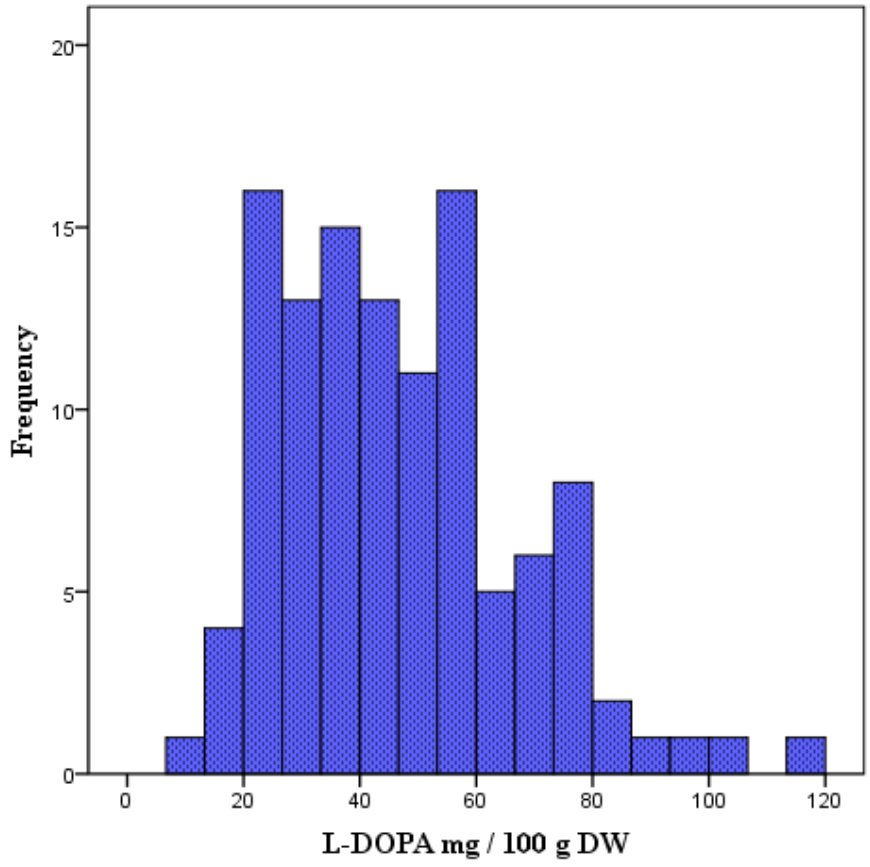

Figure 3.11. Frequency distribution of L-DOPA content for 114 faba bean genotypes.

3.3.5. Tannins Content

A calibration curve (10 to 700 mg/L catechine) was constructed (Appendix C) and tannins concentration was calculated as catechine equivalents. The range in values for this trait was between 11.9 for CGN15641 (Netherlands) to 277.1 for CGN18893 (Belgium) with an average of 146.6 mg CE/100 g DW. CGN18893 also had high phenolics content (493.7 mg GAE/100 g DW). No significant difference was observed between landraces and cultivars. There was a strong significant positive correlation between phenolic acids and tannins concentrations (r= 0.66). Also, tannins had a significant negative correlation (r = -0.30) with both stipule spot pigmentation and wing petal color. All the tannin concentrations are listed in Appendix C. Figure 3.12 shows the frequency distribution of tannins which indicates that most individuals had moderate levels of tannins between 80 and 220 mg CE/100 g DW.

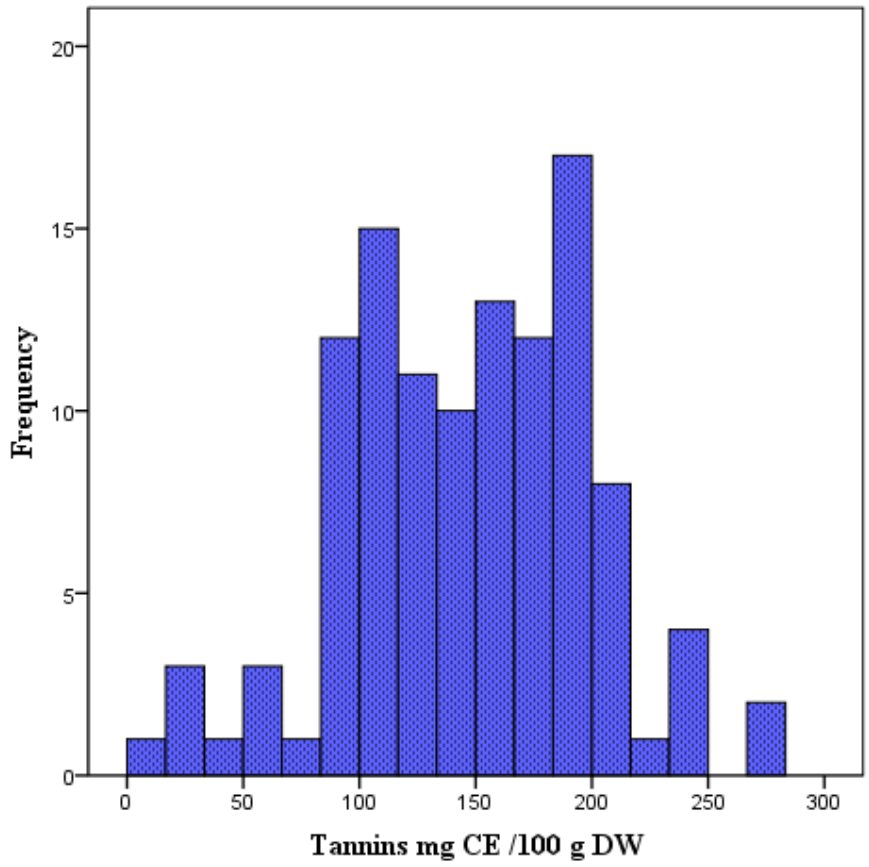

Figure 3.12. Frequency distribution of tannins content for 114 faba bean genotypes.

3.3.6. Vicine and Convicine Contents

Vicine concentration ranged from 0.02 to 0.66 with an average of 0.35% of seed dry weight. There was a significant difference (t test, $P < 0.02$) between landraces and cultivars with averages of 0.36 and 0.33% vicine, respectively. The average content for convicine was 0.21 and ranged from 0.01 to 0.39% of seed dry weight. Total concentration of vicine and convicine was calculated and ranged from 0.03 to 0.94 for Melodie/2 (France) and CGN18905 (Austria), respectively, with an average of 0.56% (Appendix C). Figure 3.13 shows the frequency distribution of total vicine and convicine and indicates that most genotypes had moderate levels of these anti-nutritional compounds.

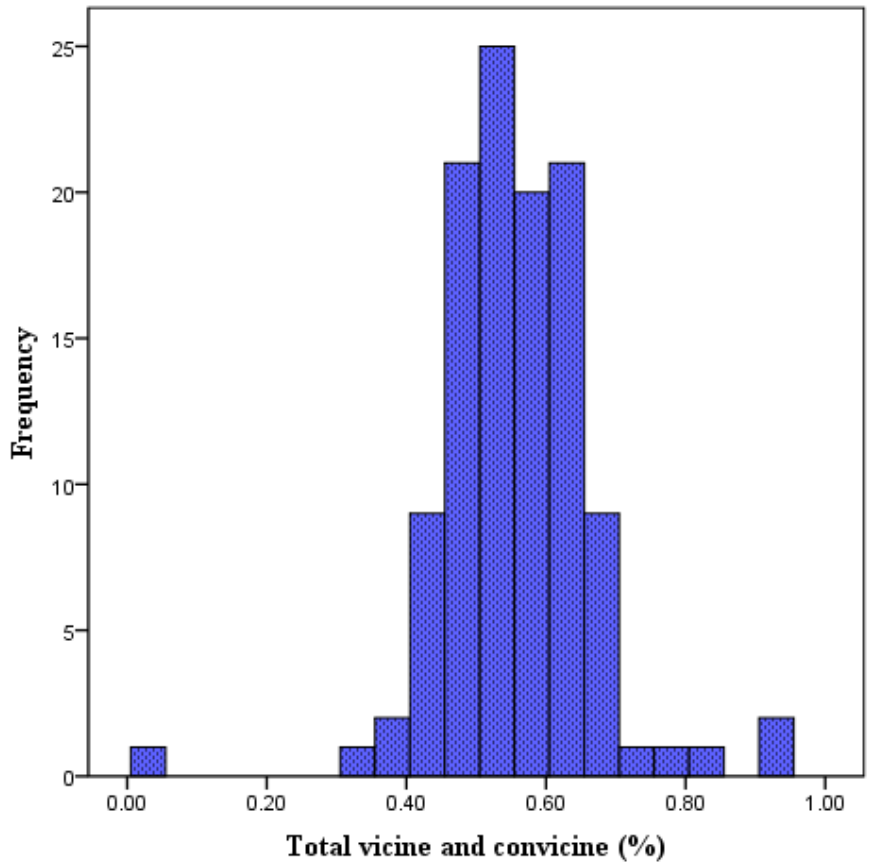

Figure 3.13. Frequency distribution of total vicine and convicine concentration for 114 faba bean genotypes.

3.4. Association Mapping

The 59 SSR primer pairs generated 442 polymorphic fragments for the 114 faba bean accessions. These polymorphic alleles were used to localize the QTLs underlying the 26 morphological and seven biochemical traits using association mapping analysis. Different association mapping models [GLM, GLM (Q), MLM (K), and MLM (K+Q)] were used to compare the proportion of significant results (Table 3.9). The GLM (Q) model had the highest proportion of significant results among the four association models (15.3%) and was used for association mapping of morphological and biochemical traits. A total of 11,440 marker pair combinations were tested and, of these, 285 (2.5 %) associations were at a significance level of *P* value ≤ 0.01. LD values (r^2) ranged from 0.05 to 0.18. Only the markers with a *P* value ≤ 0.001 for morphological and ≤ 0. 01 for biochemical traits are presented here.

Table 3.9. Association models tested to determine the best model for association analysis.

Model	π0 (%) *	π1 (%) **
GLM	85.8	14.2
GLM (Q)	84.7	15.3
MLM (K)	97.2	2.8
MLM (K++Q)	95.7	4.3

* Overall proportion of true null hypotheses (FDR)
** Proportion of significant results

3.4.1. Morphological Traits Association Mapping

Mapping identified 48 (0.4%) significant SSR marker associations with 22 of the 26 characters at a significance level of P value ≤ 0.001 (Table 3.10). Of these SSR marker alleles, 12 were associated to five of the seven vegetative traits. No SSR marker alleles were identified for leaflets per leaf and stem pigmentation at flowering time. Seven SSR marker alleles were associated with three of the four inflorescence traits. Flowers per

inflorescence was without any associations. Twenty SSR marker alleles were associated with eight of the nine pod traits. No association was found for pod angle. Nine SSR marker alleles were linked with the six seed traits. Most of the QTLs had moderate LD values of approximately 0.10. Marker associations that explained higher phenotypic variation are highlighted below.

Vegetative traits: Two QTLs were identified for leaflet shape with LD values of approximately 0.10. Four marker alleles were associated with stipule spot pigmentation and all explained moderate amounts of variation. Three QTLs were identified for number of stems per plant with the most significant locus, VfGSSR41-138 having an LD value of 0.18. Two loci were detected for stem pigmentation at maturity and only one QTL was identified for plant height.

Inflorescence traits: Days to flowering was associated with one SSR marker allele (VfGSSR87-323) and was the most statistically significant with the highest LD value (0.15) for the inflorescence traits. Two QTLs were identified for intensity of petal streaks and four loci were associated with wing petal color. Each of these petal coloration QTLs had moderate LD values.

Pod traits: Pod length, pod shape, pod surface reflectance, and number of seeds per pod were each associated with two SSR loci. Only one of these QTLs, VfGSSR15-181 associated with pod seed number, had an LD value that was slightly higher (LD = 0.15) than the others which explained moderate amounts of phenotypic variation. Numbers of pods per node, pods per plant, ovules per pod and seeds fertilized each had three marker associations. The most significant association was between VfGSSR1-263 and seeds fertilized which had an LD value of 0.15.

Seed traits: Water content of fresh seeds, seed shape, 100-seed weight, and seed yield had one QTL each. Each of these loci explained moderate amounts of phenotypic variation. Among the seed traits, coat color was linked to the greatest number of markers (three). FbgSSR451-295 had the most significant association with a Q-value of 0.003 and LD value of 0.18. Two QTLs were detected for hilum color, the more significant was linked to FbgSSR545-213 and had a LD value of 0.17.

Table 3.10. Faba bean SSR markers associated with morphological traits using GLM Q model**.**

Trait	Locus	LD value (r^2)	*P*-value ≤ 0.001	Q-value ≤ 0.2
Vegetative				
Leaflet shape	VfGSSR13-172	0.11	0.000596	0.172
	VfGSSR10-216	0.09	0.000857	0.172
Stipule spot pigmentation	VfGSSR34-221	0.11	0.000488	0.084
	FbgSSR451-295	0.11	0.000600	0.084
	FbgSSR293-276	0.09	0.001000	0.098
	VfGSSR9-156	0.09	0.001000	0.098
Stem per plant	VfGSSR41-138	0.18	0.000001	**0.0002**
	FbgSSR140-274	0.10	0.000339	0.068
	FbgSSR30-167	0.09	0.000710	0.095
Mature stem color	GBSSRVF8-292	0.10	0.000271	0.120
	VfGSSR1-280	0.09	0.000945	0.120
Plant height	VfGSSR87-183	0.12	0.000171	0.058
Inflorescences				
Days to flowering	VfGSSR87-323	0.15	0.000027	**0.008**
Intensity of petal streaks	GBSSRVF153-245	0.10	0.000805	0.136
	FbgSSR30-118	0.10	0.000945	0.136
Wing petal color	VfGSSR34-221	0.11	0.000488	0.084
	FbgSSR451-295	0.11	0.000600	0.084
	FbgSSR293-276	0.09	0.001000	0.098
	VfGSSR9-156	0.09	0.001000	0.098
Pod				
Pod length	VfGSSR34-221	0.10	0.000634	0.138
	VfGSSR31-150	0.09	0.001000	0.138
Pod shape	VfGSSR9-168	0.10	0.000488	0.100
	FbgSSR309-306	0.10	0.000641	0.100
Pod surface reflectance	VfGSSR9-124	0.12	0.000163	0.065
	VfGSSR19-134	0.11	0.000876	0.176
Number of pods per node	VfGSSR47-327	0.13	0.000098	**0.039**
	VfGSSR31-150	0.10	0.000481	0.069
	VfGSSR27-195	0.11	0.000516	0.069
Number of total pods per plant	VfGSSR1-280	0.12	0.000248	0.099
	VfGSSR31-211	0.10	0.000559	0.112

(cont. on next page)

Table 3.10 (cont.).

	VfGSSR10-239	0.09	0.000887	0.119
Maximum number of ovules per pod	VfGSSR19-256	0.13	0.000138	**0.030**
	VfGSSR15-181	0.13	0.000179	**0.030**
	VfGSSR10-196	0.11	0.000264	**0.030**
Number of seeds per pod	VfGSSR15-181	0.14	0.000095	0.167
	VfGSSR10-196	0.12	0.000099	0.167
Seeds fertilized	VfGSSR1-263	0.15	0.000031	**0.011**
	FbgSSR675-262	0.12	0.000376	0.067
	VfGSSR47-135	0.10	0.000819	0.097
Seed				
Water content of fresh seeds	VfGSSR47-218	0.14	0.000075	**0.025**
Seed shape	VfGSSR15-204	0.11	0.000536	0.214
Seed coat color	FbgSSR451-295	0.18	0.000009	**0.003**
	VfGSSR34-221	0.12	0.000365	0.069
	VfGSSR87-310	0.11	0.000608	0.076
Hilum color	FbgSSR545-213	0.17	0.000034	**0.007**
	FbgSSR293-276	0.11	0.001000	0.118
100-seed weight	VfGSSR31-150	0.09	0.000819	0.200
Seed yield	FbgSSR564-265	0.11	0.000627	0.090

Markers with multiple trait associations: QTLs linked to allele VfGSSR34-221 were detected for four traits: stipule spot pigmentation, pod length, wing petal color, and seed coat color. VfGSSR31-150 was associated with QTLs for pod length, number of pods per node, and 100-seed weight. Another allele of VfGSSR31, VfGSSR31-211, detected a QTL for total pods per plant. VfGSSR10-196 was linked to loci for the pod traits, maximum number of ovules per pod and number of seeds per pod. Two additional alleles of VfGSSR10 had associations with leaflet shape and total pods per plant. VfGSSR9-156 was associated with QTLs for stipule spot pigmentation and wing petal color. VfGSSR9 was also linked to pod shape and surface reflectance. Three associations were observed for individual alleles of FbgSSR293 and FbgSSR451. Both FbgSSR293-276 and FbgSSR451-295 alleles were associated with stipule spot pigmentation and wing petal color and seed coat color.

3.4.2. Biochemical Traits Association Mapping

Mapping identified 26 significant SSR marker associations with six of the seven biochemical traits at a significance level of *P* value ≤ 0.01 (Table 3.11). No QTLs were identified for tannins. The LD level (r^2) ranged from 0.05 to 0.13. Two QTLs were identified for phenolic content. One QTL was associated with flavonoids content. Seven loci were associated with protein traits with LD values ranging from 0.06 to 0.09. Five SSR marker alleles were linked with L-DOPA content. For this trait, GBSSRVF154-285 had the highest LD value (0.11). Vicine content was associated with seven marker alleles; two of these, GBSSRVF131-242 and GBSSRVF153-231 were the most statistically significant and had LD values of 0.13 and 0.12, respectively. Four QTLs were detected for convicine content. None of the vicine and convicine QTLs co-localized.

Table 3.11. Faba bean SSR markers associated with chemical traits using GLM Q model**.**

Trait	Locus	LD value (r^2)	*P*-value ≤ 0.01	Q-value
Phenol	VfGSSR69-313	0.09	0.0018	0.70
	FbgSSR675-249	0.06	0.0089	0.87
Flavonoids	FbgSSR564-265	0.07	0.0034	0.88
Protein	VfGSSR9-168	0.09	0.0015	0.36
	VfGSSR1-218	0.09	0.0022	0.36
	VfGSSR10-216	0.07	0.0039	0.36
	GBSSRVF115-192	0.07	0.0056	0.36
	FbgSSR293-242	0.07	0.0067	0.36
	GBSSRVF52-364	0.06	0.0072	0.36
	VfGSSR11-279	0.07	0.0073	0.36
L-DOPA	GBSSRVF154-285	0.11	0.0003	0.13
	VfGSSR3-149	0.07	0.0036	0.54
	FbgSSR520-361	0.07	0.0049	0.54
	FbgSSR309-285	0.06	0.0071	0.54
	GBSSRVF154-265	0.07	0.0073	0.54
Vicine	GBSSRVF131-242	0.13	0.00008	0.03
	GBSSRVF153-231	0.12	0.00021	0.04

(cont. on next page)

Table 3.11 (cont.).

	VfGSSR67-148	0.09	0.00200	0.26
	GBSSRVF131-254	0.08	0.00300	0.27
	VfGSSR1-354	0.08	0.00350	0.27
	GBSSRVF153-225	0.07	0.00730	0.42
	VfGSSR41-207	0.06	0.00760	0.42
Convicine	VfGSSR11-210	0.11	0.00056	0.20
	VfGSSR67-187	0.08	0.00400	0.62
	FbgSSR309-310	0.06	0.00850	0.62
	FbgSSR140-286	0.06	0.00890	0.62

CHAPTER 4

DISCUSSION

4.1. SSR Development

Faba bean is an important food legume crop with a huge genome. Development of genetic markers for faba bean is important to study diversity and for molecular breeding. In this study, we used Next Generation Sequencing (NGS) technology for the development of genomic simple sequence repeat (SSR) markers.

4.1.1. SSR Markers Developed by NGS

NGS has become a common method for discovering SSRs in plants because it can be easily performed on non-model organisms. Moreover, it is rapid and more cost-effective than traditional SSR development methods and Sanger sequencing (Zalapa *et al.*, 2012), and allows sequencing of millions of bases. In this study, NGS did not provide good genome coverage (0.13 %), but was sufficient for the development of 2138 nonredundant SSR markers for the faba bean genome, allowing detection of one SSR marker every 7.6 kb (on average) in the 16.37 Mb of sequenced contigs. Cardle *et al*. (2000) reported one SSR every 6.8 Kb in genomic DNA of many plants, and one SSR every 6.04 kb for Arabidopsis genomic DNA compared to 14 kb for ESTs. Akash and Mayers (2012) developed mono-, di-, tri- and tetranucleotide EST-SSR markers from publicly available faba bean ESTs and reported one SSR repeat every 6.13 kb. When only the same types of repeats (mono-, di-, tri- and tetranucleotide) are compared, we found one SSR repeat every 8.9 Kb. El-Rodeny *et al.* (2014) examined di-, tri-, and tetranucleotides derived from faba bean ESTs with one SSR every 34.4 kb, while we observed one genomic SSR every 26.9 kb when mononucleotides were excluded. The variable frequency of genic and genomic SSRs may reflect a difference in their distribution in coding sequences compared to the entire genome. In addition, Leclercq *et al.* (2007) reported that the variable number

of genomic SSRs identified in genomes is due to the algorithm tools used for mining and their parameter settings which can determine how many SSRs are detected. Among the selected genomic SSRs, mononucleotide repeats were the most abundant (57.5%) and outnumbered di and trinucleotide repeats. Dinucleotide repeats (20.9 %) were more abundant than trinucleotide repeats (6.5%). This result agrees with the fact that mono- and dinucleotide repeats outnumber trinucleotide repeats in eukaryotic intergenic and intron regions (Toth *et al.,* 2000). Reports of genic SSR development in faba bean demonstrated that trinucleotides were most abundant in coding regions. For example, Akash and Mayers (2012) identified 38% trinucleotide repeats followed by mononucleotides (36%) and dinucleotides (22%). El-Rodeny *et al.* (2014) also reported that trinucleotides were the most abundant and accounted for 72.7% of SSRs, followed by dinucleotides (21.9%). The motif length frequency differences between genomic and genic SSRs is most probably due to selection pressure on genic SSRs which reduces the fixation of mutations leading to frameshifts.

Among mononucleotides, A/T repeats (Table 3.2) were the most frequent (98.9%), agreeing with the observation that the most common SSR repeats in plants are A/T (Cardle *et al.,* 2000). Among dinucleotide repeats, AG/CT was most frequently observed (26.6%), followed by GA/TC (23.2%) which is in agreement with El-Rodeny *et al.* (2014), who reported that AG/CT was the most frequent genic dinucleotide (57.4%) in faba bean. Gong *et al*. (2011) reported that AG/CT and GA/TC were the most common (33.3%) genic dinucleotides while we observed that GA/TC (23.2%) was the second most common genomic dinucleotide. Among trinucleotides, AAT/ATT and ATA/TAT repeats were the most abundant accounting for 23.1 and 19.0%, respectively. Cordoba *et al.* (2010) also reported that ATA/TAT repeats were the most abundant (46%) in common bean while Cardle *et al.* (2000) reported that AAT/ATT was the most common trinucleotide in other plants.

4.1.2. Population Structure and Genetic Diversity Assessment

A total of 39 genomic SSR markers were selected based on their amplification efficiency and were applied to 46 faba bean accessions from 17 countries. All markers produced clear, reproducible fragments. Rare alleles were excluded from genetic analysis. The number of alleles ranged from 1 to 10 with an average of 4.1 alleles. Abid *et al.* (2015)

reported up to 10 alleles with an average of 5.9 alleles per locus when genetic diversity of 46 faba bean accessions was analyzed using 17 SSR markers. In our study, population structure analysis assigned faba bean accessions into two distinct subpopulations and a group of admixed accessions. A dendrogram was constructed to understand the genetic relationship of faba bean accessions from different origins. In general, accessions from the same origin did not form exclusive clusters. Turkish accessions were distributed through all three clusters reflecting the genetic diversity of these accessions. This result was expected because most Turkish faba beans were introduced from different countries and have started to replace the few original Turkish accessions (Baloch *et al.*, 2014). In contrast, all Finnish accessions grouped in one subcluster indicating the narrow genetic basis of these accessions. Population structure analysis coincided with dendrogram clustering with a few exceptions.

4.2. Morphological Trait Diversity and Association Mapping

Tens of thousands of faba bean accessions are available in germplasm collections throughout the world. Morphological characterization can enrich the data and association mapping can generate markers that can used in breeding programs. As far as we know, this is the first study that used SSR markers for genome wide association analysis of faba bean morphological traits.

4.2.1. Morphological Evaluation

Although faba bean is a seed crop, analysis of its vegetative traits is important as such traits as leaflet and stem number help determine photosynthetic capacity which, in turn, affects yield. Most of the faba bean accessions had intermediate leaflet shape and this trait was moderately and positively correlated with plant height, seed yield, pods per plant and seeds per pod. Similar results were obtained in soybean by other researchers who found that intermediate leaflet plants had more pods per plant and higher yield than those with either ovate or narrow leaflets (Krisnawati and Adie, 2017; Dinkins *et al.*, 2002). Fine mapping in soybean by Jeong *et al.* (2011) indicated that the genomic regions controlling leaflet shape and seed per pod mapped to a 0.7 cM interval in soybean. Such genetic linkage may explain the correlation between these two traits and suggests that selection for intermediate types may be useful for yield improvement. Our association mapping

results revealed two loci for leaflet shape without any co-localization with yield parameters traits. The discrepancy may be due to the fact that the faba bean genome is more complex and 11-fold larger than the soybean genome, thus, additional markers should be mapped in the genome.

The faba bean genotypes examined in this work had an average of 5.5 leaflets per leaf, in agreement with the work of Terzopoulos *et al.* (2003) who examined Greek accessions and found that leaflet number ranged from 4 to 8 with an average of 5.5 leaflets per leaf. However, when measured in the same way, South Tunisian accessions had only 3.5 leaflets per leaf (Yahia *et al.*, 2012). Duc (1997) reported that faba bean leaflet number increased from the bottom to the top of the stem. Because the average plant height of our material was approximately twice that of the Tunisian accessions, the difference in leaflet number may result from plant height variability. No QTLs were identified for this trait.

While the landraces and cultivars had the same average for stems per plant (4.8), there was variation when compared with other studies, Yahia *et al.* (2012) and Terzopoulos *et al.* (2003) reported averages of 3.4 and 3.9 stems per plant, respectively. This variation may be due to the effect of environmental conditions on the trait which has low broad sense heritability (17.6%) when measured by Alan and Geren (2007). The most stems per plant were for the two Turkish genotypes and may reflect their adaptation to the growth conditions used in this work.

Plant height is a yield parameter. Kumar *et al. (*2018) reported average height as 117.8 cm which agreed with our result (106.8 cm). However, Al Barri and Shtaya (2013) that reported an average of 76.8 cm when their accessions were grown without fertilizer or irrigation. This variation may be explained by Della (1988) who reported that significant differences in plant height occur depending on irrigation. Also, Hegab *et al.* (2014) found that faba bean height was significantly reduced as sowing date was delayed.

Days to flowering is an inflorescence trait, and this trait is controlled by major environmental factors such as ambient temperature and photoperiod (Seaton *et al.*, 2015). Torres *et al.* 2011 reported that faba bean can tolerate chilling from 0 to 10°C and can tolerate brief exposure to temperatures down to -5°C. Although Izmir has a moderate climate and the coldest month in the year is January with temperature between 5-11°C with rare snow fall (World Weather and Climate Information), the 2016/17 season was exceptional as temperatures fell below zero (-3°C) and there was snow in the beginning of January. The faba bean survived and overcame these harsh conditions but there was a significant difference (t test, $P < 0.0001$) for days to flowering between the 2 seasons

which was longer with 99 days in 2016/17 compared to 84 days in 2017/18 season. Açikgöz (2001) confirmed the low temperatures delay flowering in *Vicia* genus.

In our study, 4 and 5 flowers per inflorescence were predominant (68%). The same number of flowers per inflorescence was observed when Polignano *et al.* (1999) collected 1565 accessions from 39 countries. Although producing flowers is the first step towards yield, not all the flowers become pods. Also, not all the ovules of the pod (here, 5%) develop into mature seeds. While the number of flowers is important for yield, the color of the flowers can be used as a morphological marker for tannin content because tannins and anthocyanin are derived from the same precursors (Li *et al.,* 2016). Mutation in this pathway results in faba bean with free or low tannin and with white flowers and absence of both stipule color and stem pigmentation. In this study, the genotypes that had these morphological characters also had low tannins.

There was variation in pod length and averages agreed with Yahia *et al.* (2012) who reported an average pod length of 9.4 cm. Sharifi (2015) reported that broad sense heritability was the highest for pod length, 98%, indicating that environmental factors had a small effect. However, Gezahegn and Tesfaye (2017) reported that pod length and seeds per pod of faba bean were significant influenced by inter- and intra-row spacing. In this work, there were significant positive correlation between pod length and 100- seed weight ($r = 0.69$) which agreed with Al Barri and Shtaya (2013) and Musallam *et al.* 2004 who reported high correlations ($r = 0.88$ and 0.74, respectively) indicating the importance of pod length to increasing seed yield.

In faba bean, 100-seed weight, number of seeds per pod, and total pods per plant are the primary component of yield. In this study, the highest 100-seed weight was for CGN19985 (unknown origin) followed by the three Turkish cultivars, Eresen87, Filiz99, and Salkim confirming the fact that large seeds are preferred in Turkey as reported by Karaköy *et al.* (2014) and Alan and Geren (2007). The Finnish Mikko cultivar had the lowest 100-seed weight (29.4 g) in agreement with Pulli and Vestbero (1981) who reported 22 g for Mikko. Despite its low seed yield, this cultivar is cultivated in Finland due to its extreme earliness in maturation (Hovinen, 1988). Gezahegn and Tesfaye (2017) determined that 100-seed weight was not affected by inter- and intra-row spacing. Also, Toker (2004) reported that the broad sense heritability for 100-seed weight was 99%. In the present study, 100-seed weight showed variation such that the cultivars had 10.2% heavier weight than landraces indicating the impact of selection and breeding.

4.2.2. Morphological Association Mapping

The number of morphological traits (26) and genotypes (114) evaluated make this one of the largest association mapping studies performed in faba bean to date. There was variation among morphological traits such as vegetative, inflorescence, pod and seed characters and this variation was required to perform association mapping. A total of 11,440 (0.42 %) marker pairs showed significant LD ($P \leq 0.001$). LD can occur from false positive association between the trait and markers due to the fact that faba bean is partially cross-pollinated, and the rate of out-crossing ranges from 4 to 84% (Bond and Poulsen, 1983). To correct for false positive associations, the GLM model was used with Q matrix of population structure generated by the Structure 2.2.3 program (Pritchard *et al.*, 2000). In this study, the range of LD values (r^2) was from 0.09 to 0.18. In an association mapping study in faba bean using 156 SNP marker conducted by Sallam *et al*. (2016b), 12, 9, 2, and 1 QTLs were detected for seed yield, days to flowering, plant height, and 100-seed weigh respectively, the LD value range was from 0.03 to 0.09 which are lower than detected in our work. Ersoz *et al.* (2007) explained that the levels of LD are highly variable across the genome due to variation in the rate of recombination and selection. Frary *et al.* (2019) suggested that the small LD values of the QTL are typical of truly quantitative traits that controlled by multiple genes and that these minor QTLs are responsible for a small amount of the total phenotypic variation in the trait.

In this study, a total of 48 SSR markers were associated with morphological traits at the $P \leq 0.001$ significance level. Yield and its components are the most importance traits in breeding programs. There was one QTL for seed yield and a total of 6 QTL associations (12.5% of the total number) were for yield primary components (number of total pods per plant, number of seeds per pod and 100-seed weight). Seven loci (14.6%) impacted secondary yield components (stem per plant, plant height and number of pods per node).

Co-localization of marker alleles for different traits may indicate that a single locus influences multiple traits (pleiotropy) or that the region contains linked genes. In this study, flower stipule and wing petal color co-localized to VfGSSR34-221, FbgSSR451-295, FbgSSR293-276 and VfGSSR9-156 (Table 3.10). Also, two of these

four alleles (VfGSSR34-221 and FbgSSR451-295) were associated with seed coat color. Interestingly, VfGSSR34-221 was also co-localized with pod length with a significant positive correlation between pod length and seed coat color. Since pod length had correlations with the most yield parameters and especially with 100-seed weight (Appendix B), this means that seed coat color can be used as a marker for yield. Tiryaki *et al.* (2016) found a weak negative correlation among seed coat color and seed yield parameters in vetch (*Vicia sativa* L.). Picard (1976) found that the tannin-free characteristic was controlled by two complementary recessive genes with a pleiotropic effect on plant flower and seed coat color in faba bean.

In the present study, the most associations (33 of 48) were from AG repeat SSR markers. Association of AG repeats to the most traits was expected because AG is the most frequent dinucleotide in faba bean. Gong *et al.* (2011) observed 33.3% AG repeats. El-Rodeny *et al.* 2014 reported that AG was the most frequent dinucleotide (57.4%) in faba bean, we also found an AG repeat frequency of 26.6%.

4.3. Biochemical Analysis and Association Mapping

Phenolic compounds represent a large group of molecules with a variety of functions in plant growth, development, and defense. Phenolic compounds including flavonoids are considered as natural source antioxidants (Chaieb *et al.* 2011). There was variation for phenolic content and flavonoids in the 144 genotypes (15.5 and 21.4% respectively), No significant difference was observed between landraces and cultivars. Chaieb *et al*. (2011) reported that phenolic content and flavonoids are affected by genetic factors as well as the growing stage and degree of maturity. Association mapping resulted in three different QTLs for these traits: two for phenol and one for flavonoids. Faba bean protein is considered as an alternative to animal proteins in poor countries. The protein content of faba bean ranged from 20 to 40% (Torres *et al.*, 2012). The results here showed that only 4.4% of 114 genotypes had protein contents more than 35%.

L–DOPA is the precursor of the neurotransmitter dopamine. Of the biochemical traits, L–DOPA had the highest variation in the 114 genotypes (CV, 43.5%) with an average of 46.6 mg/100 g DW (0.05% L-DOPA). This result agreed with Ramya and Thaakur (2007) who reported 0.07% content for dry seeds. L-DOPA was associated with five QTLs without any co-localization with other traits.

Tannins, vicine and convicine are considered as ANFs. Tannin is the main ANF which reduces protein digestibility. All the genotypes that showed a small amount of tannin (< 60 mg CE/100 g DW, 5.3% of accessions) were cultivars indicating the effect of selection and breeding. Although zero tannin is preferred, such genotypes may be more susceptible to soil borne diseases (Helsper *et al.*, 1993).

No QTLs were identified for tannins in this work. Gutierrez *et al.* (2007) identified five RAPD markers associated with low tannins (zt2 gene). VC can cause favism in individuals with glucose-6-phosphate dehydrogenase (G6PD) deficiency. Here, the VC contents agree with Duc *et al*. (1989) when a collection of 918 genotypes was assayed. The variations for vicine and convicine were from 0.22 to 1.01%, and from 0.07 to 0.96% DW, respectively. No relationship was found between VC content and protein or tannins contents. Gutierrez *et al.* 2006 reported two RAPD markers linked to VC gene. Here, two of 7 SSRs alleles were associated to vicine with highly significant *P* values of 0.00008 and 0.0002 and Q-values of 0.03 and 0.04, respectively. These QTLs may be good markers for vicine content in faba bean breeding.

CHAPTER 5

CONCLUSION

This is the first study aimed at development of genomic SSR markers for faba bean using NGS. Sequencing and mining of the faba bean genome allowed identification of 2138 SSR markers. A subset of these SSR markers was used to test efficiency of PCR amplification and study genetic diversity and population structure within faba bean. Thus, the markers were found to be a useful tool for further studies of genetic diversity and population structure and in genetic mapping and breeding of faba bean. The morphological diversity between 114 genotypes for two seasons was analyzed and association mapping conducted using SSR markers. These morphological characters showed variation for important yield parameters. 100-seed weight is an important yield character; CGN19985 had the highest value followed by three Turkish accessions: Eresen87, Filiz99 and Salkim, respectively. Also CGN19987 had the longest pods followed by CGN19979, Eresen87 and CGN19985. These two characters (100-seed weight, pod length) are important and have high broad sense heritability and are not affected by environmental factors. The association mapping resulted in 48 loci associated with 22 of 26 morphological characters and 26 loci associated with six of seven chemical traits. This study produced SSR markers, phenotypic data and QTLs with linked markers that can potentially help in breeding improved faba bean cultivars for both morphological and biochemical traits.

REFERENCES

Abdelkrim, J. et al. (2009) 'Fast, cost-effective development of species-specific microsatellite markers by genomic sequencing', BioTechniques, 46(3), pp. 185–192. doi: 10.2144/000113084.

Abid, G. et al. (2015) 'Genetic Relationship and Diversity Analysis of Faba Bean (*Vicia faba* L. var. Minor) Genetic Resources Using Morphological and Microsatellite Molecular Markers', Plant Molecular Biology Reporter, 33(6), pp. 1755–1767. doi: 10.1007/s11105-015-0871-0.

Abuzayed, M. et al. (2017) 'GDdom: An Online Tool for Calculation of Dominant Marker Gene Diversity', Biochemical Genetics, 55(2), pp. 155–157. doi: 10.1007/s10528-016-9779-0.

Abuzayed, M. A. et al. (2017) 'Development of genomic simple sequence repeat markers in faba bean by next-generation sequencing', Plant Molecular Biology Reporter. Plant Molecular Biology Reporter, 35(1), pp. 61–71. doi: 10.1007/s11105-016-1003-1.

Açikgöz, E. (2001) 'Yem bitkileri', Uluda Üniversitesi Güçlendirme Vakf Yayn No: 182. VPABursa 58, pp. 584 (in Turkish)

Akash, M. W. and Myers, G. O. (2012) 'The development of faba bean expressed sequence tag-simple sequence repeats (EST-SSRs) and their validity in diversity analysis', Plant Breeding, 131(4), pp. 522–530. doi: 10.1111/j.1439-0523.2012.01969.x.

Alan, O. and Geren, H. (2007) 'Evaluation of heritability and correlation for seed yield and yield components in faba bean (*Vicia faba* L.)', Journal of Agronomy, 6(3), pp. 484–487. doi: 10.3923/ja.2007.484.487.

Al Barri, T. and J Y Shtaya, M. (2013) 'Phenotypic Characterization of Faba Bean (*Vicia faba* L.) Landraces Grown in Palestine', Journal of Agricultural Science, 5(2), pp. 110–117. doi: 10.5539/jas.v5n2p110.

Alghamdi, S. S. et al. (2012) 'Faba bean genomics: Current status and future prospects', Euphytica, 186(3), pp. 609–624. doi: 10.1007/s10681-012-0658-4.

Ali, M. B. M. et al. (2016) ‘Association analyses to genetically improve drought and freezing tolerance of faba bean (*Vicia faba* L.)’, Crop Science, 56(3), pp. 1036–1048. doi: 10.2135/cropsci2015.08.0503.

Ouji, A. et al. (2012) ‘Genetic diversity of faba bean (*Vicia* faba L.) populations revealed by sequence specific amplified polymorphism (SSAP) markers’, African Journal of Biotechnology, 11(9), pp. 2162–2168. doi: 10.5897/ajb11.2991.

Anishettty, N. M. et al. (1985) ‘Faba bean descriptor’, AGPG:IBPGR, 85, pp. 116

Avila, C. M. et al. (2005) ‘QTL detection for agronomic traits in faba bean (*Vicia faba* L.)’, Agriculturae Conspectus Scientificus, 70(3), pp. 65–73. Available at: http://www.scopus.com/inward/record.url?eid=2-s2.0-33645676455&partnerID=tZOtx3y1.

Baginsky, C. et al. (2013) ‘Phenolic compound composition in immature seeds of fava bean (Vicia *faba* L.) varieties cultivated in Chile’, Journal of Food Composition and Analysis, 31(1), pp. 1–6. doi: 10.1016/j.jfca.2013.02.003.

Baloch, F. S. et al. (2014) ‘Variation of some seed mineral contents in open pollinated faba bean (*Vicia faba* L.) landraces from Turkey’, Turkish Journal of Agriculture and Forestry, 38(5), pp. 591–602. doi: 10.3906/tar-1311-31.

Bond, D. A. and Poulsen, M. H. (1983) ‘Pollination. In: Hebblethwaite PD (ed) The faba bean (*Vicia faba* L.)’, Butterworths, London, UK, pp. 77–101

Bradbury, P. J. et al. (2007) ‘TASSEL: Software for association mapping of complex traits in diverse samples’, Bioinformatics, 23(19), pp. 2633–2635. doi: 10.1093/bioinformatics/btm308.

Bradford, M. M. (1976) ‘A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding’, Analytical Biochemistry, 72, pp. 248–254.

Burbano, C. et al. (1995) ‘Variation of favism-inducing factors (vicine, convicine and L-DOPA) during pod development in *Vicia faba* L.’, Plant Foods for Human Nutrition, 47(3), pp. 265–274. doi: 10.1007/BF01088335.

Cardle, L. et al. (2000) ‘Computational and experimental characterization of physically clustered simple sequence repeats in plants’, Genetics 156(2), pp. 847–854.

Chaieb, N. et al. (2011) 'Polyphenolscontent and antioxidant capacity of thirteen faba bean (*Vicia faba* L.) genotypes cultivated inTunisia', Food Research International, 44(4), pp. 970-977.

Concepcion, V. et al. (1998) 'Nutrients and antinutritional factors in faba beans as affected by processing', Z Lebensm Unters Forsch A 207, pp. 140–145

Cordoba, J. M. et al. (2010) 'Integration of physical and genetic maps of common bean through BAC-derived microsatellite markers', BMC Genomics 11:436. doi:10.1186/1471-2164-11-436

Crépon, K. et al. (2010) 'Nutritional value of faba bean (*Vicia faba* L.) seeds for feed and food', Field Crops Research, 115(3), pp. 329–339. doi: 10.1016/j.fcr.2009.09.016.

Cruz-Izquierdo, S. et al. (2012) 'Comparative genomics to bridge *Vicia faba* with model and closely-related legume species: Stability of QTLs for flowering and yield-related traits', Theoretical and Applied Genetics, 125(8), pp. 1767–1782. doi: 10.1007/s00122-012-1952-1.

Cubero, J. I. (1974) 'On the evolution of *Vicia faba* L.', Theor Appl Genet 45(2) , pp. 47–51.

Della, A. (1988) 'Characteristics and variation of Cyprus faba bean germplasm', FABIS Newsletter (21), pp. 9-12.

Deshpande, S. and Cheryan, M. (1984) 'Effect of phytic acid, divalent cations and their interactions on α-amylase activity', J Food Sci 49, pp. 516–519

Dinkins, R. D. et al (2002) 'Expression of the narrow leaflet gene for yield and agronomic traits in soybean', J Heredity 93: 346-351.

Doyle, J. J. and Doyle J. E. (1990) 'Isolation of plant DNA from fresh plant tissue', Focus (12), pp. 13–15.

Duc, G. et al. (1989) 'Search of genetic variability for vicine and convicine content in *Vicia faba* L: a 1st report of a gene which codes for nearly zero vicine and zero-convicine contents' , Recent Advances of Research in Antinutritional Factors in Legume Seeds, pp. 305-313.

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

DUC, G. et al. (1999) ‘Genetic variability for feeding value of faba bean seeds (*Vicia faba*): comparative chemical composition of isogenics involving zero-tannin and zero-vicine genes’, The Journal of Agricultural Science, 133(2), pp. 185-196.

Duc, G. et al. (2010) ‘Diversity maintenance and use of *Vicia faba* L. genetic resources’, Field Crops Research, 115(3), pp. 270–278. doi: 10.1016/j.fcr.2008.10.003.

Earl, D. A. and vonHolt, B. M. (2012) ‘Structure Harvester: a website and program for visualizing structure output and implementing the Evanno method’, Conserv Genet Resour 4, pp. 359–361. doi :10.1007/s12686-011-9548-7

El-Rodeny, W. et al. (2014) ‘Development of EST-SSR markers and construction of a linkage map in faba bean (*Vicia faba*)’, Breeding Science, 64(3), pp. 252–263. doi: 10.1270/jsbbs.64.252.

Elfalleh, W. et al. (2009) ‘Physico-chemical properties and DPPH-ABTS scavenging activity of some local pomegranate (*Punica granatum*) ecotypes’, International Journal of Food Sciences and Nutrition, 60(SUPPL. 2), pp. 197–210. doi: 10.1080/09637480903067037.

Ellwood, S. R. et al. (2008) ‘Construction of a comparative genetic map in faba bean (*Vicia faba* L.); conservation of genome structure with Lens culinaris’, BMC Genomics, 9, pp. 1–11. doi: 10.1186/1471-2164-9-380.

Ersoz, E. S. et al. (2007) ‘Genomics-Assisted Crop Improvement’, (December). doi: 10.1007/978-1-4020-6295-7.

Flint-Garcia, S. A., Thornsberry, J. M. and Buckler, E. S. (2003) ‘Structure of Linkage Disequilibrium in Plants’, Annual Review of Plant Biology, 54(1), pp. 357–374. doi: 10.1146/annurev.arplant.54.031902.134907.

Frary, A. et al. (2019) ‘Analysis of European hazelnut (*Corylus avellana*) reveals loci for cultivar improvement and the effects of domestication and selection on nut and kernel traits’, Molecular Genetics and Genomics, 294(2), pp. 519–527. doi: 10.1007/s00438-018-1527-1.

Gezahegn, A. M. and Tesfaye, K. (2017) ‘Optimum inter and intra row spacing for faba bean production under Fluvisols’, 4, pp. 10–19.

Göl, Ş. Doğanlar, S. and Frary, A. (2017) ‘Relationship between geographical origin, seed size and genetic diversity in faba bean (*Vicia faba* L.) as revealed by SSR

markers', Molecular Genetics and Genomics, 292(5), pp. 991–999. doi: 10.1007/s00438-017-1326-0.

Gong, Y. et al. (2010) 'Generation and characterization of 11 novel EST derived microsatellites from *Vicia faba* (Fabaceae)', American Journal of Botany, 97(7), pp. 69–71. doi: 10.3732/ajb.1000166.

Gong, Y. et al. (2011) 'Genetic Diversity Analysis of Faba Bean (*Vicia faba* L.) Based on EST-SSR Markers', Agricultural Sciences in China. Chinese Academy of Agricultural Sciences, 10(6), pp. 838–844. doi: 10.1016/S1671-2927(11)60069-2.

Gressel, J. et al. (2004) 'Major heretofore intractable biotic constraints to African food security that may be amenable to novel biotechnological solutions', Crop Protec 23, pp. 661–689.

Gresta, F. et al. (2010) 'A study of variability in the Sicilian faba bean landrace "Larga di Leonforte"', Genetic Resources and Crop Evolution, 57(4), pp. 523–531. doi: 10.1007/s10722-009-9490-7.

Griffiths, D. W. (1983) 'Some anti-nutritive factors in *Vicia faba*', FABIS,Newsletter, 6, pp. 1-3

Gutierrez, N. et al. (2006) 'CAPs markers to assist selection for low vicine and convicine contents in faba bean (*Vicia faba* L.)', Theoretical and Applied Genetics, 114(1), pp. 59–66. doi: 10.1007/s00122-006-0410-3.

Gutierrez, N. et al. (2007) 'Development of SCAR markers linked to a gene controlling absence of tannins in faba bean', Molecular Breeding, 19(4), pp. 305–314. doi: 10.1007/s11032-006-9063-9.

Hamada, H. Petrino, M. G. and Kakunaga, T. (1982) 'A novel repeated element with Z-DNA-forming potential is widely found in evolutionarily diverse eukaryotic genomes', Pnas, 79(21), pp. 6465–6469.

Hegab, A. S. A. et al. (2014) 'Productivity and irrigation requirements of faba-bean in North Delta of Egypt in relation to planting dates', Annals of Agricultural Sciences. Faculty of Agriculture, Ain Shams University, 59(2), pp. 185–193. doi: 10.1016/j.aoas.2014.11.004.

Helsper, J. P. F. G. et al. (1993) 'Antinutritional Factors in Faba Beans (*Vicia faba* L.) As Affected by Breeding toward the Absence of Condensed Tannins', Journal of Agricultural and Food Chemistry, 41(7), pp. 1058–1061. doi: 10.1021/jf00031a008.

Hovinen, S. (1988) ‘Breeding of field bean (*Vicia faba* L.) with early maturity’, Agricultural and Food Science, 60(4), pp. 261-67.

Huang, X. and Han, B. (2014) ‘Natural variations and genome-wide association studies in crop plants’, Annu Rev, Plant Biol (65), pp. 531–551.

Inci, N. E. and Toker, C. (2011) ‘Screening and selection of faba beans (*Vicia faba* L.) for cold tolerance and comparison to wild relatives’, Genet Resour Crop Evol 58, 1169. doi: 10.1007/s10722-010-9649-2

Inci, N. E. and Toker, C. (2016) ‘Yield Criteria in Faba Bean (*Vicia faba* L.)’, International Journal of Environmental Engineering, 3(1), pp. 89-92

Jeong, N. et al. (2011) ‘Fine genetic mapping of the genomic region controlling leaflet shape and number of seeds per pod in the soybean’, Theor Appl Genet, (122), pp. 865-874.doi: 10.1007/s00122-010-1492-5

Jones, C. G. Hare, J. D. and Compton, S. J. (1989) ‘Measuring plant protein with the bradford assay: 1. Evaluation and Standard Method’, Journal of Chemical Ecology, 15, pp. 979–992

Karaköy, T. et al. (2014) ‘Variation for selected morphological and quality-related traits among 178 faba bean landraces collected from Turkey’, Plant Genetic Resources: Characterisation and Utilisation, 12(1), pp. 5–13. doi: 10.1017/S1479262113000208.

Karaköy, T. et al. (2018) ‘Assessment of Micro and Macro Nutrients Contents in the Turkish Faba Bean Germplasm’, “Agriculture for Life, Life for Agriculture” Conference Proceedings, 1(1), pp. 72–78. doi: 10.2478/alife-2018-0011.

Karkanis, A. et al. ‘Faba bean cultivation-revealing novel managing practices for more sustainable and competitive european cropping systems’, Front Plant Sci., 9, 1115. Published 2018 Aug 2. doi:10.3389/fpls.2018.01115

Kaur, S. et al. (2012) ‘Transcriptome sequencing of field pea and faba bean for discovery and validation of SSR genetic markers’, BMC Genomics. BioMed Central Ltd, 13(1), p. 104. doi: 10.1186/1471-2164-13-104.

Kaur, S. et al. (2014) ‘Assessment of genetic diversity in faba bean based on single nucleotide polymorphism’, Diversity, 6(1), pp. 88–101. doi: 10.3390/d6010088.

Khamassi, K. et al. (2013) ‘A baseline study of vicine-convicine levels in faba bean (*Vicia faba* L.) germplasm’, Plant Genetic Resources: Characterisation and Utilisation, 11(3), pp. 250–257. doi: 10.1017/S1479262113000105.

Khazaei, H. (2014) ‘Leaf traits associated with drought adaptation in faba bean (*Vicia faba* L.)’, Available at: https://helda.helsinki.fi/bitstream/handle/10138/135892/lefttrai.pdf?sequence=1.

Khazaei, H. et al. (2015) ‘Flanking SNP markers for vicine–convicine concentration in faba bean (*Vicia faba* L.)’, Molecular Breeding, 35(1). doi: 10.1007/s11032-015-0214-8.

Koressaar, T. Remm, M. (2007) ‘Enhancement and modifications of primer design program Primer3’, Bioinformatics 23(10), pp. 1289–1291. doi:10.1093/bioinformatics/btm091

Krisnawati, M. and Adie, M. M. (2017) ‘The leaflet shape variation from several soybean genotypes in Indonesia’, Biodiversitas 18 (1), pp. 359–364. doi: 10.13057/biodiv/d180147

Kumar, P. and Bishnoi, S. (2018) ‘Genetic variability, heritability and genetic advance for seed yield and other agro-morphological traits in Faba bean (*Vicia faba* L.) genotypes of different origin’, Trends in Biosciences, 10(4), pp. 1246–1248.

Langmead, B. and Salzberg, S. L. (2012) ‘Fast gapped – read alignment with Bowtie2’, Nat Methods 9(4), pp. 357–359. doi:10.1038/nmeth.1923

Lattanzio, V. et al. (1983) ‘Variability of Amino Acids, Protein, Vicine and Convicine in *Vicia faba* (L) Cultivars’, Journal of Food Science, 48(3), pp. 992–993. doi: 10.1111/j.1365-2621.1983.tb14950.x.

Leclercq, S. Rivals, E. and Jarne, P. (2007) ‘Detecting microsatellites within genomes: significant variation among algorithms’, BMC Bioinformatics 8, 125. doi:10.1186/1471-105-8-125

Lewis, G . et al. (2005) ‘Legumes of the world. Royal Botanic Gardens’, Kew Publishing, Richmond, UK, 577 p.

Li, P. et al (2016) ‘Metabolic engineering of proanthocyanidin production byrepressing the isoflavone pathways and redirecting anthocyanidin precursor flux in legume’, Plant Biotechnol J., 14(7), pp. 1604–1618. doi: 10.1111/pbi.12524

Link, W. et al. (1995) ‘Genetic diversity in European and Mediterranean faba bean germplasm revealed by RAPD markers’, Theor. Appl. Genet, 90, pp. 27–32

Ma, Y. et al. (2011) ‘Development and characterization of 21 EST- derived microsatellite markers in *Vicia faba* (fava bean)’, American Journal of Botany, 98(2), pp. 22–24. doi: 10.3732/ajb.1000407.

Ma, Y. et al. (2013) ‘Genetic linkage map of Chinese native variety faba bean (*Vicia faba* L.) based on simple sequence repeat markers’, Plant Breeding, 132(4), pp. 397–400. doi: 10.1111/pbr.12074.

Marquardt, R. R. Muduuli, D. S. and Frohlich, A. A. (1983) ‘Purification and Some Properties of Vicine and Convicine Isolated from Faba Bean (*Vicia faba* L.) Protein Concentrate’, Journal of Agricultural and Food Chemistry, 31(4), pp. 839–844. doi: 10.1021/jf00118a041.

Martin, M. (2011) ‘Cutadapt removes adapter sequences from high-throughput sequencing reads’, EMBnet.journal 17(1), pp. 10–12. doi:http://dx.doi.org/10.14806/ej.17.1.200

Musallam, I. W. Al-karaki, G. N. and Ereifej, K. I. (2004) ‘Ereifej Chemical composition of faba bean genotypes under rainfed and irrigation conditions’, Int. J. Agri. Biol., 6, pp. 359-362.

Nasri, N. et al. (2011) ‘Chemical compounds from Phoenician juniper berries (*Juniperus phoenicea*)’, Natural Product Research, 25(18), pp. 1733–1742. doi: 10.1080/14786419.2010.523827.

Nei, M. (1973) ‘Analysis of gene diversity in subdivided populations’, Proc Natl Acad Sci U S A 70(12), pp. 3321–3323. doi:10.1073/pnas.70.12.3321

Nijveldt, R. J. et al. (2001) ‘Flavonoids: a review of probable mechanisms of action and potential applications’, American Journal of Clinical Nutrition, 74, pp. 418–425.

Ofuya, Z. M. and Akhidue, V. (2005) ‘The Role of Pulses in Human Nutrition: A Review’, Journal of Applied Science & Environment , 9(3), pp. 99–104. doi: 10.4314/jasem.v9i3.17361.

Oomah, B. D. et al. (2011) ‘Phenolics, phytic acid, and phytase in canadian-grown low-tannin faba bean (*Vicia faba* L.) genotypes’, Journal of Agricultural and Food Chemistry, 59(8), pp. 3763–3771. doi: 10.1021/jf200338b.

Oviedo-Silva, C. A. Elso-Freudenberg, M. and Aranda-Bustos, M. (2018) 'L-DOPA trends in different tissues at early stages of *Vicia faba* Growth: effect of tyrosine treatment', Appl. Sci., 8, 2431. doi:10.3390/app8122431

Picard, J. (1976) 'Aperc¸u sur l'he´redite´ du caracte´re absence de tannins dans les graines de fe´verole (*Vicia faba* L.) Ann. Ame´lior', Plantes 26 (1), pp. 101-106.

Polignano G. B. Alba, E. Ugenti, P. and Scippa G. (1999) 'Geographical patterns of variation in Bari faba bean germplasm collection', Genet. Res. Crop Evol. (46), pp. 183-192.

Pozarkova, D. et al. (2002) 'Development and characterization of microsatellite markers from chromosome 1-specific DNA libraries of *Vicia faba*', Biol Plant 45, pp. 337–345.

Price, M. L. Van Scoyoc, S. and Butler, L. G. (1978) 'A critical evaluation of the vanillin reaction as an assay for tannin in sorghum grain', Journal of Agricultural and Food Chemistry, 26, pp. 1214–1218

Pritchard, J. K. Stephens, M. and Donnelly, P. (2000) 'Inference of Population Structure Using Multilocus Genotype Data', doi: 10.1111/j.1471-8286.2007.01758.x.

Pulkkinen, M. et al. (2015) 'Determination of vicine and convicine from faba bean with an optimized high-performance liquid chromatographic method', Food Research International, 76(P1), pp. 168–177. doi: 10.1016/j.foodres.2015.05.031.

Pulli, S. and Vestbero, M. (1981) 'Genetic and management adaptation of field bean (*Vicia faba* L.) in Finland', J. Scient. Agric Soc Finland (53), pp. 328—340.

Purves, R. W. Khazaei, H. and Vandenberg, A. (2018) 'Quantification of vicine and convicine in faba bean seeds using hydrophilic interaction liquid chromatography', Food Chemistry, 240, pp. 1137–1145. doi: 10.1016/j.foodchem.2017.08.040.

Ramsay, G. and Griffiths, D. W. (1996) 'Accumulation of vicine and convicine in *Vicia faba* and V. narbonensis', Phytochemistry, 42(1), pp. 63–67. doi: 10.1016/0031-9422(95)00870-5.

Ramya, K. B. and Thaakur, S. (2007) 'Herbs containing L- Dopa: An update', Anc Sci life, 27(1), pp. 50–5.

Rassmann, K. Schlotterer, C. and Tautz, D. (1991) 'Isolation of simple sequence loci for use in polymerase chain reaction based DNA fingerprinting', Electrophoresis 12, pp. 113-118.

Remington, D. L. et al. (2001) 'Structure of linkage disequilibrium and phenotypic associations in the maize genome', Proc Natl Acad Sci U S A, 98(20), pp. 11479–11484. doi: 10.1073/pnas.201394398.

Roldan-Ruiz, I. et al. (2000) 'AFLP markers reveal high polymorphic rates in ryegrasses (Lolium spp.)', Mol Breeding 6, pp. 125–134. doi:10.1023/A:1009680614564

Sallam, A. et al. (2016a) 'Identification and Verification of QTL Associated with Frost Tolerance Using Linkage Mapping and GWAS in Winter Faba Bean', Frontiers in Plant Science, 7(August), pp. 1–16. doi: 10.3389/fpls.2016.01098.

Sallam, A. Dhanapal, A. P. and Liu, S. (2016b) 'Association mapping of winter hardiness and yield traits in faba bean (*Vicia faba* L.)', Crop & Pasture Science, 67(1), pp. 55–68. doi: 10.1071/CP15200.

Sallam, A. and Martsch, R. (2015) 'Association mapping for frost tolerance using multi-parent advanced generation inter-cross (MAGIC) population in faba bean (*Vicia faba* L.)', Genetica. Springer International Publishing, 143(4), pp. 501–514. doi: 10.1007/s10709-015-9848-z.

Sato, S. Isobe, S. and Tabata, S. (2010) 'Structural analyses of the genomes in legumes', Current Opinion in Plant Biology. Elsevier Ltd, 13(2), pp. 146–152. doi: 10.1016/j.pbi.2009.12.007.

Seaton, D. D. et al. (2015) 'Linked circadian outputs control elongation growth and flowering in response to photoperiod and temperature', Mol. Syst. Biol. (11) pp. 776. doi: 10.15252/msb.20145766

Senan, S. et al. (2014) 'Methods for Development of Microsatellite Markers: An Overview', Notulae Scientia Biologicae, 6(1), pp. 1–13. doi: 10.15835/nsb.6.1.9199.

Sharifi, P. (2015) 'Genetic variation for seed yield and some of agro-morphological traits in faba bean (*Vicia faba* L.) genotypes', Acta agriculturae Slovenica, 105(1), pp. 73–83. doi: 10.14720/aas.2015.105.1.08.

Shendure, J. and Hanlee, J. (2008) 'Next-generation DNA sequencing', Nature Biotechnology 26, pp. 1135–1145.

Simpson, J. T. et al. (2009) 'ABySS: A parallel assembler for short read sequence data', Genome Res 19(6):1117–1123. doi:10.1101/gr.089532.108

Singh, J. (2018) 'Folate Content in Legumes', Biomedical Journal of Scientific & Technical Research, 3(4), pp. 1–6. doi: 10.26717/bjstr.2018.03.000940.

Sozen, O. and Karadavut, U. (2016) 'Determination of morphological and phenological properties of faba beans grown in eastern mediterranean region of turkey', Tarla Bitkileri Merkez Araştırma Enstitüsü Dergisi, 25(2), pp. 209-217. doi: 10.21566/tarbitderg.281344

Storey, J. D. and Tibshirani, R. (2003) 'Statistical significance for genomewide studies', Proc Natl Acad Sci USA (100), pp. 9440–9445. doi: 10.1073/pnas.91.25.12091.

Suresh, S. et al. (2013) 'Development and molecular characterization of 55 novel polymorphic cDNA-SSR markers in faba bean (*Vicia faba* L.) using 454 pyrosequencing', Molecules, 18(2), pp. 1844–1856. doi: 10.3390/molecules18021844.

Terzopoulos, P. J. Kaltsikes, P. J. and Bebeli, P. J. (2003) 'Collection, evaluation and classification of Greek populations of faba bean (*Vicia faba* L.)', Genetic Resources and Crop Evolution, 50(4), pp. 373–381. doi: 10.1023/A:1023962618319.

Terzopoulos, P. J. and Bebeli, P. J. (2008) 'DNA and morphological diversity of selected Greek tomato (*Solanum lycopersicum* L.) landraces', Sci Hortic, 116, pp. 354-361.

Toker, C. (2004) 'Estimates of broad-sense heritability for seed yield and yield criteria in faba bean (*Vicia faba* L.) ', Hereditas. (140), pp. 222-225. doi: 10.1111/j.1601-5223.2004.01780.x

Topal, N. and Bozoğlu, T. (2016) 'Determination of l-dopa (l-3, 4-dihydroxyphenylalanine) content of some faba bean (*Vicia faba* L.) genotypes', Tarım Bilimleri Dergisi, 22(2), pp. 145–151. doi: 10.1501/tarimbil_0000001376.

Torres, A. M. et al. (2006) 'Faba bean breeding for resistance against biotic stresses: Towards application of marker technology', Euphytica, 147(1–2), pp. 67–80. doi: 10.1007/s10681-006-4057-6.

Torres, A. M. et al. (2010) 'Marker-assisted selection in faba bean (*Vicia faba* L.)', Field Crops Research, 115(3), pp. 243–252. doi: 10.1016/j.fcr.2008.12.002.

Torres, A. M. et al. (2011) 'Faba bean', In: Pérez de la Vega M, Torres AM, Cubero JI, Kole C (eds) Genetics, genomics and breeding in crop plants: cool season food legumes. Science Pubs Inc, New Hampshire. pp 50–97

Toth, G. Gaspari, Z. and Jurka, J. (2000) 'Microsatellites in different eukaryotic genomes: survey and analysis', Genome Res 10, pp. 967–981. doi:10.1101/gr.10.7.967

Van de Ven, M. et al. (1990) 'Restriction fragment length polymorphisms as genetic markers in *Vicia*', Heredity, 65, pp. 329–342

Vieira, M. L. et al. (2016) 'Microsatellite markers: what they mean and why they are so useful', Genet Mol Biol., 39(3), pp. 312–328. doi:10.1590/1678-4685-GMB-2016-0027

Wang, P-X. and Ueberschär, K. H. (1990) 'The estimation of vicine, convicine and condensed tannins in 22 varieties of faba beans (*Vicia faba* L.)', Animal Feed Science and Technology 31, pp. 157–165.

Wang, H. et al. (2012) 'Genetic diversity and relationship of global faba bean (*Vicia faba* L.) germplasm revealed by ISSR markers', Theor Appl Genet 124:789–797

Yahia, Y. et al. (2012) 'Analysis of agromorphological diversity of southern Tunisia faba bean (*Vicia faba* L.) germplasm', African Journal of Biotechnology, 11(56), pp. 11913–11924. doi: 10.5897/AJB11.3366.

Yang, T. et al. (2012) 'High-throughput novel microsatellite marker of faba bean via next generation sequencing', BMC Genomics, 13(1). doi: 10.1186/1471-2164-13-602.

Yu, J. and Buckler, E. S. (2006) 'Genetic association mapping and genome organization of maize', Current Opinion in Biotechnology, 17(2), pp. 155–160. doi: 10.1016/j.copbio.2006.02.003.

Zalapa, J. E. et al. (2012) 'Using next-generation sequencing approaches to isolate simple sequence repeat (SSR) loci in the plant sciences', American Journal of Botany, 99(2), pp. 193–208. doi: 10.3732/ajb.1100394.

Zane, L. Bargelloni, L. and Patarnello, T. (2002) 'Strategies for microsatellite isolation: a review', Mol Ecol, 11, pp. 1-16

Zeid, M. Schon, C. and Lin, W. (2003) ‘Genetic diversity in recent elite faba bean lines using AFLP markers. Theor Appl Genet, 107, pp. 1304–1314

Zeid, M. et al. (2009) ‘Simple sequence repeats (SSRs) in faba bean: New loci from Orobanche-resistant cultivar “Giza 402”’, Plant Breeding, 128(2), pp. 149–155. doi: 10.1111/j.1439-0523.2008.01584.x.

Zong, X. et al. (2009) ‘Molecular variation among Chinese and global winter faba bean germplasm’, Theor Appl Genet, 118, pp. 971–978

Zor, T. and Selinger, Z. (1996) ‘Linearization of the Bradford protein assay increases its sensitivity: theoretical and experimental studies’, Anal Biochem, 236, pp 302–308.

APPENDIX A

SSR DEVELOPMENT

Table A.1. Primer details for the 11 SSR markers not given in Table 3.3

SSR primers	Sequence	Repeat motif	Expected size (nt)
FbgSSR138-F	GAAGATGAACAGCATTTCGG	$(TA)_8$	291
FbgSSR138-R	GAGGTGGATTTGGATCTGTT		
FbgSSR263-F	GCAAAAGTTTGTGTTGCAAG	$(CT)_6$	227
FbgSSR263-R	GTGATGATAATTTCGGTCGC		
FbgSSR275-F	ATCGCAGAAAAGAAAACACC	$(GA)_8$	280
FbgSSR275-R	GATCACAGAAAGGAATGGGT		
FbgSSR297-F	ATGATCGGTTCAATTGTTGC	$(AAT)_6$	217
FbgSSR297-R	AATGAATTGCATTGACTTGT		
FbgSSR302-F	ATCATCATCACCTCTAGCCT	$(CT)_8$	283
FbgSSR302-R	CAGTATCAGAGAAACTCACCA		
FbgSSR311-F	TCATTGTTGCAACAGAATAGC	$(AAATGGA)_3$	209
FbgSSR311-R	CAGTTGAATGAAGTTCACCG		
FbgSSR335-F	TATTGATGGGGCCTTAGAGA	$(AT)_6$	237
FbgSSR335-R	ATACATCTCCGGACACACTA		
FbgSSR473-F	AAACACTCTGATCCAACTGG	$(TTCTTGG)_3$	223
FbgSSR473-R	AGATCGCTATTACGGAAACC		
FbgSSR524-F	AATAGCCCCAAACACATCAT	$(TCA)_6$	227
FbgSSR524-R	TTTGGTGGGTCAAATAGGAG		
FbgSSR561-F	GTGGATTATGGCTGTATGGT	$(TA)_6$	202
FbgSSR561-R	TGGAAAGTGTGTCAACAAGA		
FbgSSR740-F	TATGATTTTAGGTGCGGGAG	$(AT)_7$	276
FbgSSR740-R	CTTTCCAAATGAGAAGACGC		

Table A.2. Sequencing results displaying the simple sequence repeat motifs within the PCR amplicons for four SSR markers.

Marker name	PCR amplicon
FbgSSR37- AG	TAGTGAGAGGAAGTTTCAAATAAAATCAGAGAGAGAGAGAGAGT GTGGGAAAGGGAAGTGTCCTAAAACATATTTCTAGATAATTTAAA ATGAGAAATAAGACTTACTAATTTTTTTATGTATAAAATAAGTAGC TCTTGAAACTCCTTTTATAACATATCATCCTACCGATCAAAATATA TACAATGGTTCTGGGTAAACGTAG
FbgSSR26- TTGAAT	ATAAAAAGTTGGGTAATTATGGTCCGTTTGAATAAGAGCGGGTTC CCAAAATCATCACTATATACAGTAAATTACAACGGTGTTTTGAGA CGACGTAACATGATGCAAATTTTAAGTCTGTTCCACTATTTCTGTT TTCTATAATCTATCTACCACCTTTATTGAATTTGAATTTGAA
FbgSSR309- AG	TTTTACCCTAACCAAATGTGAAAATAAACAGGAAAATGTTATAAT TAATATATAGACATGCTAGAGAGAGAGAGAGAGAGCAAAGTTCA AGAATAGTGTGCGATCAGTTTACAGAAGACATGTATAATGCCTGC AATCTTCAGAGTCTACAATGAGGTCTACTGCTTCATGTTTTGTTAT CACCTCATGATGCTGGCCCA
FbgSSR293- ACCAAA	ATAATTCATTCCTAAAAGAAAACAAATAACTATGCATTAACATAA CAAAACAATAAAAAGTCATTCAAAAAGTTATGCAACTACCAAAAC CAAAACCAAAACCAAAACCACAACAATACCACATCACACTAATAT CCATAATTTCTAAACCACAATACCTTCAATCCAGGAGCATAATCCC TACGGTAATAAGCAGGACTCTTAGCAGATCTCCACTCAAAAATTA AANTNTTNTTTNTTTGTNTGTTTNTTGNTTTNTTTTNTTTTGGTTTTT TTNTTTTTTTTTTTNGGGNTTTGNTTTNTTTT

APPENDIX B

MORPHOLOGICAL TRAIT DIVERSITY

Table B.1. Data for the 14 quantitative characteristics of 114 faba bean genotypes for two seasons. Averages and standard errors are provided. The range, average and % CV were calculated from these quantitative trait data.

Name	Number of leaflets per leaf	Number of stem per plant	Plant height (cm)	Days to flowering	Number of flowers per	Pod length (cm)	Number of pods per node	Number of total pods per plant	Maximum of ovules per pod	Number of seeds per pod	Seeds fertilized (%)	Water content of fresh seeds (%)	100-seed weight (g)	Seed yield (g)
CGN07723	5 ±0.3	5 ±0.4	102.4 ±3.8	69.5 ±1.8	5.5 ±0.2	11.4 ±0.1	1.2 ±0.1	18.1 ±2.3	3.9 ±0	3.3 ±0.2	85.6 ±4.7	76 ±4.1	78.2 ±12.5	34.1 ±1.3
CGN07734	5 ±0	4.5 ±0.9	73 ±7.8	98.5 ±18.7	4.5 ±0	6.3 ±1.3	1.4 ±0.3	45.3 ±7.7	2.8 ±0.7	2.5 ±0.8	86.3 ±6	57.9 ±0.9	62.8 ±10	42.3 ±13.4
CGN07781	5.6 ±0	4.3 ±0.5	148.5 ±6	72.5 ±6.7	6.8 ±0	9.8 ±0.2	1.3 ±0.2	41.6 ±4.5	3.1 ±0.1	3 ±0.1	96.7 ±2.4	78.8 ±0.2	76.1 ±2.9	83.8 ±27.9
CGN07836	4.7 ±0.2	4.2 ±0	99.1 ±5.7	67.5 ±3.2	4.3 ±0.7	10.5 ±2.1	1.5 ±0.1	34.8 ±1.7	3.8 ±0.1	3.8 ±0.1	100 ±0	61.4 ±9.9	62.1 ±3.1	57.4 ±12.3
CGN07871	4.8 ±0.2	4.8 ±0.1	127.5 ±5.3	132 ±3.5	3.8 ±0.2	6.8 ±0.3	1.7 ±0.1	46 ±7.6	3.4 ±0.2	3.3 ±0.1	97.2 ±2	76 ±1.3	52.8 ±14.2	50.3 ±5.1
CGN07933	5.2 ±0	5.5 ±1.5	104.5 ±8.1	122.5 ±1.8	6.2 ±0.1	6.2 ±0.7	1.4 ±0.1	26.9 ±8.3	3.2 ±0	3.6 ±0.2	100 ±0	75.3 ±0.3	46.2 ±8.2	27.8 ±4.6
CGN10320	5.4 ±0	6.8 ±0.8	115 ±13.4	106 ±13.4	6.5 ±0.6	10.9 ±0.6	1.3 ±0.2	20.2 ±3.3	3.5 ±0.2	3.1 ±0.5	87.5 ±8.8	79.9 ±0.4	102.3 ±0.1	39.2 ±0.7
CGN10322	4.7 ±0	5.6 ±1.4	93.6 ±26.6	119.5 ±0.4	4.1 ±0.4	8.6 ±0.2	1 ±0	9.3 ±4.5	3.3 ±0.2	3.2 ±0.3	96.7 ±2.4	76.5 ±0.6	77 ±5.8	25.8 ±15.6
CGN10330	4.9 ±0.1	4.6 ±0.1	118.8 ±16.1	66 ±4.2	4.8 ±0	7.5 ±0.8	1.1 ±0.1	42 ±9.2	3.6 ±0	3.3 ±0.2	92.9 ±5.1	70.6 ±3.9	67.1 ±3.2	52.5 ±1.4

(cont. on next page)

Table B.1. (cont.).

CGN10362	6 ±0	4.5 ±0.4	77.4 ±11.1	106 ±13.4	6.2 ±0.6	7.7 ±0.2	1 ±0	12.6 ±1	3.3 ±0.4	2.9 ±0.5	86.7 ±5.7	72.7 ±3.1	60 ±21.2	18 ±3
CGN10371	5.5 ±0.4	4.7 ±2.1	121.3 ±2.3	93 ±7.8	6 ±0	10.2 ±0.4	1.1 ±0.1	46.2 ±2.8	3.8 ±0.1	3.7 ±0.1	97.5 ±1.8	69 ±1.6	89.5 ±4.3	114.1 ±31.7
CGN10382	4.6 ±1.4	7.3 ±1.1	104.3 ±20.7	119 ±0.7	5.5 ±0.2	8 ±0.5	1.4 ±0.3	49.2 ±27.3	3.3 ±0.4	3.3 ±0.4	100 ±0	73.5 ±4.9	67.2 ±6.6	134.6 ±85.3
CGN10382.1	5.7 ±0	11.5 ±2.2	108 ±21.9	122.5 ±1.8	5.1 ±0.3	7.9 ±0.1	1.6 ±0.3	77.5 ±28.6	2.7 ±0	2.7 ±0	100 ±0	77.4 ±3.4	64.2 ±1.7	116.7 ±56.8
CGN10384	6.3 ±0.1	3.9 ±0.1	94.9 ±7.9	66 ±4.2	4.1 ±0.8	11.5 ±1.4	1.3 ±0.1	10.1 ±0.1	3.4 ±0	3.3 ±0.1	97.1 ±2.1	75.8 ±2.2	88.1 ±24.3	25.6 ±11.3
CGN10385	5.3 ±0.1	5.3 ±0.5	99.6 ±24.5	103.5 ±15.2	4.6 ±0.6	9.1 ±0.3	1.3 ±0.1	5 ±7.8	3 ±0	3 ±0	100 ±0	67.7 ±4.8	82.1 ±3.7	41.4 ±24.5
CGN10387	5.8 ±0.6	5.4 ±0.4	130.6 ±11	74.5 ±5.3	5.4 ±0.1	9.5 ±0.5	1.4 ±0.1	27.2 ±0.3	3.8 ±0.3	3.6 ±0.3	94.7 ±0.4	61.9 ±3.2	79.7 ±1	64.1 ±15.8
CGN10391	5.2 ±0	5 ±0.7	75.8 ±7.6	103.5 ±15.2	5 ±0	9.8 ±1.3	1 ±0	8.4 ±0.1	3.5 ±0.4	2.8 ±0.4	79.2 ±4.1	69.6 ±3.7	117.9 ±12.7	21.1 ±2.4
CGN12320	6.4 ±0	4.9 ±1.1	109.1 ±5.6	64 ±5.7	7 ±0	10.1 ±0.3	1 ±0	20 ±4.4	3.8 ±0.3	3.3 ±0.4	86.4 ±2.9	72.7 ±0	82.3 ±11.1	37.4 ±8.1
CGN12321	5.6 ±0	3.5 ±0.4	129.1 ±23.2	110 ±16.3	6.4 ±0.4	9.5 ±0.4	1.3 ±0.1	24.7 ±14.8	3.5 ±0.2	3.5 ±0.2	100 ±0	74.7 ±4.4	89.2 ±0.8	79.9 ±54.2
CGN13445	5.2 ±0	4.3 ±0.8	125.6 ±4	108.5 ±11.7	6.4 ±0	8.3 ±0.1	1.1 ±0.1	25.9 ±10.5	3.8 ±0	3.7 ±0	97.4 ±1.9	76.8 ±2.2	54.6 ±1.9	37.2 ±11.4
CGN13464	6 ±0	5.1 ±1.2	134.5 ±6.7	132 ±3.5	6.6 ±0.5	10 ±0.4	1.2 ±0.1	23.3 ±3.7	3.8 ±0	3.6 ±0	96 ±0.9	79 ±2.9	73.9 ±13.6	43.8 ±8.8
CGN13487	5 ±0	4 ±0.4	105.8 ±2.3	130 ±2.1	4.2 ±0	6.2 ±0.5	1.7 ±0.2	47.5 ±21.8	3.6 ±0.1	3.4 ±0	94.7 ±3.7	74.9 ±0.9	37.4 ±8.8	70 ±45.1
CGN13518	4.6 ±0.3	3.3 ±0.4	130.2 ±30.3	100.5 ±20.2	5.8 ±0.2	11.9 ±1.4	1 ±0	20 ±5.5	4 ±0.2	4 ±0.2	100 ±0	73.3 ±2.4	83.8 ±9.2	48.3 ±17.4
CGN15556	5.5 ±0.1	4.1 ±0.9	77.2 ±9.5	102.5 ±21.6	3.8 ±0.8	8.3 ±0.4	1.2 ±0	25.6 ±8.9	3.6 ±0.1	3.4 ±0.1	94.4 ±0.2	62 ±0.4	69.1 ±12.8	41.5 ±20.1
CGN15615	6.3 ±0	6.2 ±1.6	113 ±4.9	69.5 ±1.8	3.8 ±0	9.1 ±0.5	1.4 ±0.3	25.1 ±0.2	3.5 ±0.1	3.3 ±0.1	94.4 ±3.9	76.8 ±2.3	63 ±9.9	44.9 ±15.2
CGN15620	5.6 ±0	3.2 ±0.3	89.1 ±19.9	69.5 ±1.8	5.2 ±0	10.7 ±2.7	1.1 ±0.1	30.5 ±7.8	3.4 ±0.4	3.2 ±0.3	95 ±3.5	69.2 ±5	75.9 ±4.4	53.9 ±2.4
CGN15621	6.4 ±0	5 ±0.4	98.3 ±3.6	98.5 ±18.7	5.2 ±0	13.6 ±0.5	1.2 ±0	10.5 ±1.5	3.8 ±0.1	3 ±0.6	78.1 ±12	78.5 ±1.7	147.6 ±1.7	39.3 ±6.1
CGN15639	5.6 ±0	6.3 ±0.1	84.7 ±32.8	83 ±7.8	3.9 ±0.1	7.8 ±0.6	1.5 ±0.4	63.9 ±33.2	3.3 ±0.1	3.1 ±0.2	93.8 ±4.4	71.9 ±0	54.4 ±1.3	113.2 ±69.8
CGN15641	6.6 ±0.1	5 ±1.4	104.5 ±2	98.5 ±18.7	5.6 ±0.4	13 ±0	1.4 ±0.2	23.8 ±7	4 ±0	3.4 ±0.5	100 ±0	77.7 ±0.6	67.4 ±4.9	27.3 ±3.1

(cont. on next page)

Table B.1. (cont.).

CGN15644	5.8 ±0.1	4.7 ±0.2	105 ±5.7	69.5 ±1.8	5.4 ±0.4	7.6 ±0.4	2 ±0.3	31.2 ±6.4	2.8 ±0	2.8 ±0	100 ±0	72.7 ±2.4	67.9 ±5.5	55.2 ±21.2
CGN18856	5 ±0	4.7 ±0.4	122.5 ±21.6	67.5 ±10.3	6.4 ±0.5	12.5 ±0.6	1 ±0	20.8 ±4.4	4.1 ±0.1	3.8 ±0.1	92.6 ±1.9	78.5 ±1	93.6 ±4.4	63.3 ±21.9
CGN18860	4.6 ±0.4	3.7 ±0.2	85.8 ±13.8	122.5 ±1.8	6.1 ±0.6	9.6 ±0.1	1.6 ±0.3	15 ±7.2	3.6 ±0.1	3.5 ±0.2	97.5 ±2.4	76.9 ±1.5	57.4 ±1.3	31.8 ±18
CGN18862	4.9 ±0.5	3.7 ±0.2	103.5 ±17.3	95.5 ±6	4.2 ±1.7	13.4 ±0.4	1 ±0	12.1 ±1.6	4.5 ±0.2	4.4 ±0.3	97.6 ±1.7	77.2 ±0.4	86.4 ±0.6	34.1 ±4.3
CGN18867	6 ±0.6	4.2 ±0.1	129 ±9.2	83 ±7.8	6.1 ±0.1	14.1 ±0.6	1.3 ±0.1	16.3 ±1.4	3.9 ±0.1	3.9 ±0.1	100 ±0	76.5 ±4.2	107.8 ±3	48.9 ±7.3
CGN18867.1	5.1 ±0.6	4.7 ±0.2	109.7 ±19.6	83 ±7.8	5.5 ±0.1	9.9 ±0.2	1 ±0	22.9 ±11.5	4 ±0.6	3.7 ±0.4	93.8 ±4.4	74.5 ±3.7	74.7 ±3.2	71.1 ±41.9
CGN18878	5.3 ±0	3.7 ±0.5	138.5 ±29.3	88 ±4.2	5.2 ±0.1	9.5 ±0.1	1.4 ±0.1	18.3 ±4.9	3.4 ±0	3.3 ±0	95.6 ±1	78.6 ±1.7	83.6 ±4.4	43.3 ±14.8
CGN18892	5 ±0	4.8 ±1.1	123.7 ±21.4	95 ±16.3	5.7 ±0.9	8.3 ±0.2	1.9 ±0.4	25.2 ±15.1	3.8 ±0.2	3.7 ±0.2	98.6 ±1	81.1 ±0.1	37.1 ±0.9	31.2 ±20.9
CGN18893	5.4 ±0.3	3.3 ±0.6	129.7 ±9	114.5 ±15.9	5.5 ±0.3	10.6 ±0.6	1.9 ±0.2	22.1 ±5.7	3.6 ±0.1	3.5 ±0.1	97.4 ±1.9	69.3 ±4.4	58.7 ±4.6	37 ±13.9
CGN18905	6.1 ±0.6	4.2 ±0.1	128.4 ±5.4	67.5 ±3.2	4.9 ±0.1	10.4 ±0.5	1.3 ±0.2	19.6 ±2.6	4.1 ±0.2	3.3 ±0.4	80 ±4.5	65.1 ±4	152.5 ±1.8	33.6 ±16.4
CGN18906	5.9 ±0.3	4.3 ±1.8	102.9 ±27.7	93 ±0.7	5.6 ±0	12.1 ±0.8	1 ±0	34.8 ±22.8	4.3 ±0.1	3.8 ±0.3	88.2 ±5.1	72.1 ±3.3	94.7 ±18.6	174.1 ±117.8
CGN18909	5.6 ±0	3.5 ±0.8	109.4 ±1.8	122.5 ±1.8	5.2 ±0.6	8.2 ±0.1	2.1 ±0.2	19.7 ±1.9	4 ±0	4 ±0	98.8 ±0.9	80.3 ±0.6	54 ±0.7	35.4 ±7.5
CGN18920	5.9 ±0.1	4.4 ±0	103.5 ±5.3	75 ±8.5	4.3 ±0.8	7.3 ±0.7	1.3 ±0.1	33.8 ±9.3	2.8 ±0.1	2.8 ±0.1	100 ±0	75.7 ±4	60.2 ±9.2	33.7 ±7.1
CGN18923	6.6 ±0	4.4 ±0	105.5 ±22.7	66 ±4.2	6.8 ±0.3	8.5 ±0.1	1.6 ±0.3	34.3 ±0.2	3.5 ±0	3.4 ±0	96.1 ±0.8	72.3 ±2.1	39.3 ±0.6	37.9 ±6.6
CGN18927	6 ±0	5.5 ±0.9	120.2 ±3	98.5 ±18.7	6.3 ±0	8.7 ±1	1.1 ±0.1	29.4 ±3.4	4.1 ±0.1	3.9 ±0.1	95.2 ±3.4	71.6 ±3.9	54.4 ±5.8	41.9 ±0.9
CGN18934	3.9 ±0.9	4.1 ±0.1	122.3 ±9.4	122.5 ±1.8	6.9 ±1.3	10.8 ±0.9	1.3 ±0.2	22.6 ±9.9	4.2 ±0.1	4.2 ±0.1	100 ±0	67.2 ±3.1	52.1 ±0.9	49.2 ±26
CGN18941	4.8 ±0	3.8 ±0.1	132.5 ±23.7	108.5 ±11.7	5.5 ±0.9	9.1 ±1.3	1.1 ±0.1	24.3 ±5.9	3.5 ±0.1	3.2 ±0.2	89.9 ±3.2	76.9 ±4.1	79.8 ±7.3	33.5 ±0.9
CGN18942	5.7 ±0.1	5.1 ±0.1	90.8 ±9.8	98.5 ±18.7	5.1 ±0.1	9.6 ±0	1.5 ±0.2	16.2 ±5.5	3 ±0	3 ±0.2	92.3 ±0.2	71.7 ±6.2	82.5 ±4.5	36.1 ±16.4
CGN19979	4.9 ±0.8	4.7 ±0.1	96 ±9.2	69.5 ±1.8	5 ±0.2	15.7 ±0.7	1.3 ±0.1	16.1 ±0.4	4.5 ±0.1	4.5 ±0.1	100 ±0	71.5 ±1	126.6 ±4.9	65.2 ±14.2

(cont. on next page)

Table B.1. (cont.).

CGN19981	6.5 ±0.1	5.6 ±0.6	83.3 ±8.3	98.5 ±18.7	4.3 ±0.2	12.6 ±2.5	1.2 ±0	19.3 ±6.9	6.1 ±1.5	3.9 ±0.4	69.3 ±11.1	75 ±3.7	102.8 ±4	55.4 ±17
CGN19982	4.9 ±0.5	3.3 ±0.5	86 ±19.1	64 ±5.7	5.1 ±0.3	13.5 ±0	1.3 ±0.1	8.9 ±1.9	4.1 ±0.1	3.7 ±0.2	95.1 ±0.1	67.3 ±5.4	84.6 ±21.5	29 ±17.2
CGN19985	5.1 ±0.6	4.3 ±0.8	105 ±4.2	103.5 ±15.2	4.1 ±0.4	14.7 ±0.5	1.3 ±0.1	11.1 ±3.1	3.5 ±0.1	3.2 ±0.1	91.7 ±5.9	73.8 ±3.5	174.6 ±3.6	39.4 ±3.6
CGN19986	6.5 ±0.1	6.3 ±1.9	84 ±2.8	92.5 ±3.9	4.8 ±1.5	11.7 ±0.1	1 ±0	7.8 ±1.6	4.1 ±0.1	3.8 ±0.3	100 ±0	74.4 ±3.6	100.5 ±11.8	17.7 ±6.6
CGN19987	6.2 ±0.1	3.1 ±0.2	80 ±4.9	69.5 ±1.8	3.8 ±0.6	17.3 ±2.5	1.1 ±0.1	14.9 ±0.1	4.6 ±0	4.3 ±0.1	93.5 ±1.5	70.8 ±0.8	141.7 ±20.7	71.5 ±29.3
CGN19987.1	6 ±0	5.9 ±1.2	84.9 ±13.5	98.5 ±18.7	4.6 ±0	9.4 ±0.2	1.3 ±0	34.5 ±12.5	3.4 ±0	3.2 ±0.1	94.1 ±4.2	73 ±3.3	95.4 ±0.3	70.4 ±16
CGN19993	5.7 ±0	6 ±0.7	102.5 ±26.5	62.5 ±6.7	5.9 ±0	9.8 ±0.2	1.3 ±0.2	43.7 ±26	3.4 ±0.1	3.3 ±0.1	97.2 ±2	71.9 ±3.1	72.2 ±6.1	104.8 ±73.1
CGN07697	6.2 ±0.1	3.8 ±0.7	94.5 ±1.8	98.5 ±18.7	4.7 ±0.1	8.2 ±0.1	1.7 ±0.2	43.3 ±15.2	3.5 ±0.1	3.3 ±0	93 ±2.9	68 ±0.8	45.7 ±2	42 ±8
CGN07699	5.6 ±0	4.3 ±0.1	137.5 ±1.1	107.5 ±18	5.9 ±0.2	8.2 ±0.3	1.7 ±0.1	42.9 ±1.1	3.4 ±0	3.4 ±0	100 ±0	73.9 ±1.4	42.7 ±14.2	65.6 ±20.6
CGN07717	5.9 ±0.1	2.9 ±0.2	72 ±9.2	103.5 ±15.2	6 ±0.4	8.5 ±0.2	1.2 ±0.1	8.8 ±0.2	3.8 ±0.1	3.8 ±0.1	100 ±0	69.3 ±1	63.3 ±5.1	14.7 ±4.9
CGN07719	4.8 ±0.6	2.6 ±0.7	108.1 ±23.4	117 ±14.1	4.4 ±0.3	10.4 ±0.2	1.3 ±0.2	7.1 ±3.3	4.2 ±0	4.1 ±0.1	97.6 ±1.7	75.1 ±1.9	70.4 ±0.7	27.9 ±5.3
CGN07721	5.5 ±0.1	5.5 ±0.4	118.6 ±22.9	74.5 ±5.3	4.6 ±0.1	9.9 ±1.1	1.6 ±0	25 ±10.7	2.7 ±0.2	2.7 ±0.2	100 ±0	70.8 ±4.4	94.1 ±10.8	73.1 ±41.6
CGN07725	5.7 ±0.2	3.3 ±0.2	95.3 ±11.8	69.5 ±1.8	4.5 ±0.4	7.6 ±0.4	1.3 ±0	26.1 ±9.3	3 ±0.1	2.8 ±0	93.8 ±4.4	62 ±5	88.2 ±6.3	41.3 ±8.8
CGN07728	5.5 ±0.1	3 ±0	113 ±8.5	122.5 ±1.8	6 ±1.3	8 ±0.1	1.7 ±0.1	41.5 ±14.2	3.8 ±0.1	3.7 ±0.1	97.5 ±1.8	76.5 ±3.7	44.8 ±0.4	63.2 ±26.5
CGN07730	6 ±0	4.2 ±0.6	155.5 ±11	83 ±7.8	4.6 ±0.6	8.9 ±1.1	1 ±0	34.4 ±0.6	3.7 ±0.2	3.7 ±0.2	100 ±0	78 ±0.5	63.1 ±6	65.5 ±7.9
CGN07730.1	5.4 ±0	4.5 ±0.1	144.5 ±19.4	95 ±16.3	8.1 ±0	8.6 ±0	1.4 ±0.3	46.5 ±19.6	4.3 ±0.1	4.1 ±0.1	96.5 ±0.7	71.5 ±4.8	53.1 ±4.1	79.6 ±43.1
CGN07733	5.5 ±0.4	4.6 ±0.2	105.3 ±8.3	103.5 ±15.2	4.4 ±0.2	10.2 ±0.3	1.6 ±0	19.9 ±1.1	3.1 ±0.2	2.9 ±0.1	94.1 ±4.2	73.1 ±4.8	118.6 ±3.4	55.7 ±8.7
CGN07735	4.6 ±0.2	3 ±0.6	110.5 ±0.4	98.5 ±18.7	3 ±0.1	7.3 ±0.5	1.4 ±0.3	22 ±1.8	3.3 ±0.2	3 ±0	91.7 ±5.9	65.3 ±0.8	62.2 ±0	20.4 ±21.1
CGN07736	5.5 ±0	4.1 ±0.5	100.4 ±3.1	106 ±13.4	5.8 ±0	8.6 ±1	1.3 ±0.1	13.1 ±1.9	3.7 ±0.2	3.2 ±0.1	87.5 ±8.8	72.4 ±2.1	61.9 ±2.3	21.6 ±4.4

(cont. on next page)

Table B.1. (cont.).

CGN07738	4.2 ±0.2	4 ±0.8	107.6 ±29.3	83 ±7.8	3.8 ±0	7.3 ±1	1.2 ±0.1	39.4 ±17.8	2.8 ±0.1	2.8 ±0.2	98.1 ±1.4	71.7 ±4	59.5 ±8	62 ±37.1
CGN07739	5.9 ±0.1	4.1 ±0.2	84.3 ±0.7	67.5 ±10.3	4.4 ±0	7.7 ±0.4	1.3 ±0.1	13.9 ±0.1	3.5 ±0.2	3.2 ±0.4	90.6 ±6.6	68.2 ±1.5	68.4 ±2.1	29.5 ±5.6
CGN07740	5.1 ±0.1	7.6 ±1.6	133 ±18.4	83 ±7.8	3.9 ±0.3	7.7 ±0.5	1.6 ±0.1	63.9 ±0.6	3.4 ±0.1	3 ±0	87 ±2.6	73.7 ±2.3	52.2 ±2.4	87.1 ±12.2
CGN07751	5 ±0	4.8 ±0.1	88 ±12.7	122.5 ±1.8	5.2 ±0	6 ±0.1	1.5 ±0.4	20.3 ±2.4	3.3 ±0.1	3.3 ±0.1	100 ±0	69.9 ±5.2	30.8 ±0.7	18.9 ±21.6
CGN07757	5.7 ±0	4.9 ±1.1	94.1 ±7.7	67.5 ±3.2	5.2 ±0.1	9.6 ±0.1	1.8 ±0.1	36 ±14.8	4.4 ±0.1	3.9 ±0.1	88.4 ±4.1	74.1 ±1.2	56.4 ±12.3	54.9 ±22.3
CGN07766	5.7 ±0.2	4.3 ±0.5	117.9 ±11.8	67.5 ±3.2	5.8 ±0.6	8.6 ±0	1.4 ±0.1	15.4 ±1	3 ±0	3.5 ±0.3	100 ±0	76.3 ±1.3	88.8 ±5.9	55.8 ±5.5
CGN07826	4.5 ±0.4	4.8 ±0.1	106.7 ±4.7	64 ±5.7	4.3 ±0.4	12 ±0.2	1.1 ±0.1	19.9 ±7.6	2.9 ±0.1	2.8 ±0.1	96.4 ±2.5	74.3 ±0.3	108.2 ±1.4	51.4 ±23.4
CGN07827	5.1 ±0.1	4.8 ±0.4	144 ±8.5	73.5 ±1.1	4.7 ±0.6	10.1 ±0.5	1.5 ±0.1	92.6 ±37.4	3.8 ±0	3.7 ±0.1	97.4 ±1.9	70.7 ±3.9	52.4 ±0.9	158 ±88.1
CGN07827.1	4.9 ±0.8	4.7 ±0.1	117.8 ±11.7	62.5 ±6.7	5.1 ±0.2	11.2 ±0	1.4 ±0.1	33.1 ±4.6	3.7 ±0.1	3.6 ±0.1	97.2 ±2	70 ±3.5	70.7 ±0.3	64.1 ±22.4
CGN07842	4.3 ±0.2	5.2 ±0.7	106.2 ±29.1	90.5 ±2.5	3.8 ±0.1	8.5 ±1.7	1.9 ±0.1	29.9 ±1.9	3.4 ±0	3.4 ±0	100 ±0	69.4 ±4.3	45.8 ±4.1	39.2 ±12.7
CGN07843	5.6 ±0.2	4.1 ±0.5	133.5 ±14.5	98.5 ±18.7	4.8 ±0.6	10.4 ±0.8	1.5 ±0.1	21.8 ±5.3	4.1 ±0.1	3.9 ±0.1	95.2 ±3.4	72.6 ±3.3	76.8 ±0.4	57.4 ±23.6
CGN07844	4.7 ±0	4 ±0.8	113.7 ±20.3	69.5 ±1.8	5.5 ±0	9.4 ±0.7	1.5 ±0.1	33.3 ±4.7	3.2 ±0.1	3.1 ±0.1	97.1 ±2.1	67.3 ±0.6	58.4 ±4.2	41.4 ±9.5
CGN07934	4.3 ±0.2	4.8 ±0.4	107.5 ±4.6	88 ±4.2	4.3 ±0.1	7.9 ±0.9	2.1 ±0.1	45.4 ±2.8	3.1 ±0.1	2.9 ±0.1	95.2 ±1	74.3 ±3	40.8 ±1.3	45.9 ±9.2
CGN07938	5.3 ±0	3.6 ±0.3	103 ±17	132 ±3.5	4.2 ±0.1	7.8 ±1	1.8 ±0.3	24.6 ±3.8	3.7 ±0.1	3.6 ±0	97.4 ±1.9	71.8 ±1.6	33.1 ±6.3	27.7 ±12.4
NGB1540.1	5.8 ±0	3.8 ±0.1	111.4 ±8.1	69.5 ±1.8	5.6 ±0.6	8.7 ±0.2	1.1 ±0.1	15 ±1	3 ±0	3 ±0	98.3 ±1.2	75.3 ±4.1	58.8 ±0.7	22 ±4.6
NGB1542.1	6 ±0	3.7 ±0.9	76.2 ±10	69.5 ±1.8	7 ±0.2	7.9 ±0.2	1.4 ±0.3	8.1 ±3.1	3.3 ±0.2	3.3 ±0.2	100 ±0	73.1 ±2.7	48.6 ±6.1	12.3 ±6.5
NGB1547.1	5.9 ±0.1	6.5 ±1.3	108.8 ±34.5	59 ±4.2	5.5 ±0.2	10.5 ±0.5	1.3 ±0.2	47.3 ±3.7	3.7 ±0.1	3.6 ±0.2	97.1 ±2	74.2 ±3.2	52.7 ±0.8	59.2 ±8.7
NGB1550.1	5.9 ±0.1	3.1 ±0.4	97 ±19.1	66 ±4.2	3.5 ±0.8	8.3 ±1.9	1.5 ±0.1	23.3 ±4.4	3.5 ±0.4	3.4 ±0.6	95.8 ±6.5	70.7 ±3.7	36.4 ±2.6	27.6 ±9.2
NGB20019.2	5.8 ±0.1	3 ±0.6	101 ±17	64 ±5.7	5.5 ±0.7	8.5 ±0.6	1 ±0	28.1 ±4.5	3.1 ±0.4	3 ±0.3	96.3 ±2.6	66 ±2.5	91.7 ±4.7	69.2 ±30.8

(cont. on next page)

Table B.1. (cont.).

NGB8640	5.1 ±0.6	3.1 ±0.4	62.8 ±13.5	71 ±7.8	5 ±0.4	7.4 ±0.2	1.3 ±0.2	15.2 ±8.6	3.4 ±0	3.3 ±0.1	97.1 ±2.1	69.7 ±5.5	35.8 ±9.6	37.3 ±6.1
NGB8642	5.4 ±0	4.4 ±1.5	72.8 ±10.8	67.5 ±3.2	5.8 ±0.3	6.6 ±0	1.1 ±0.1	11.6 ±0.3	3.4 ±0	3.2 ±0.1	94.1 ±4.2	68.7 ±0.5	51.1 ±3.6	14 ±4.4
NGB8643	5.4 ±0.2	4 ±0.4	89 ±1.4	69 ±9.2	4.1 ±0.6	9.3 ±0.2	1.8 ±0.1	15.4 ±0.7	4.2 ±0.1	4.2 ±0.1	100 ±0	63.2 ±2.1	54.1 ±0.3	23.5 ±8.6
TR12540	5.5 ±0.1	6.5 ±0.6	127.5 ±13.8	88 ±4.2	3.7 ±0.3	12.3 ±0.5	1.1 ±0.1	38.4 ±12.9	3.6 ±0.3	3.5 ±0.2	97.5 ±1.8	75.1 ±0.3	130.6 ±8.6	151.4 ±73.1
TR23018	5.7 ±0.2	5.3 ±0.8	110.3 ±28.5	93 ±7.8	4.1 ±0.8	7.5 ±0.5	1.3 ±0.2	14.6 ±0.4	3.1 ±0.1	3 ±0	95.2 ±1	72.6 ±5.3	58.7 ±3.9	16.3 ±5
TR28096	5.8 ±0.2	5.9 ±0.2	116.2 ±5.8	83 ±7.8	4.9 ±0.4	11.4 ±0.3	1.3 ±0.2	31 ±4.9	3.9 ±0.2	3.9 ±0.2	100 ±0	69 ±4	123.9 ±6.1	123.4 ±42.8
TR31590	5.4 ±0	7 ±0.3	131.5 ±13.1	93 ±7.8	4.2 ±0.5	9.6 ±1.1	1.5 ±0.2	45.8 ±21.6	3.5 ±0.2	3.5 ±0.2	100 ±0	74.5 ±5.5	64.1 ±9.3	119.5 ±73.9
TR31912	5.4 ±0	5.4 ±0.4	109.7 ±0.9	98.5 ±18.7	5 ±0.1	7.1 ±0.4	1.5 ±0.4	18.3 ±1.1	3 ±0.1	2.9 ±0.1	98.4 ±1.1	78.2 ±2.6	84.4 ±1.1	39 ±8.6
TR44862	6.1 ±0.1	8 ±1.7	114.1 ±12.5	122.5 ±1.8	3.7 ±0.9	9.6 ±0.4	1.2 ±0.1	27.7 ±3.6	2.7 ±0.2	2.7 ±0.2	98.3 ±1.2	81 ±0.8	104.7 ±12.8	66.8 ±35.4
TR53947	5.2 ±0	5 ±0.1	111.6 ±9.6	98.5 ±18.7	4.7 ±0	11.1 ±0.8	1.1 ±0.1	19.4 ±3	3.4 ±0.1	3.2 ±0.1	94.1 ±0.2	80.2 ±0.1	122.3 ±11.8	69.2 ±29.9
TR71255	5.5 ±0	3.8 ±0.4	68.9 ±9.8	106 ±13.4	4.2 ±0.8	10.1 ±0	1 ±0	17 ±6.7	3.8 ±0	3.5 ±0.2	92.1 ±5.6	72.9 ±0.5	102.4 ±1	42.7 ±13
TR75421	5.6 ±0	7.6 ±1.4	115.3 ±14.7	112.5 ±14.5	4.2 ±0.3	10.5 ±0.4	1.4 ±0.1	21.4 ±2.8	3.2 ±0.1	3.2 ±0.1	100 ±0	82.2 ±1.8	116.8 ±3.6	64.9 ±5.1
Aquadulce	5.6 ±0	6.7 ±0.1	129.5 ±28.6	93 ±0.7	3.1 ±0	12.1 ±0.4	1.4 ±0	33 ±0	4 ±0.1	4 ±0.1	100 ±0	77.5 ±4.1	101.8 ±6.8	67.8 ±42.4
Ascot	5.7 ±0.1	4.7 ±0.2	73.4 ±13.7	73.5 ±9.5	5.5 ±0.3	8.7 ±0.3	1.5 ±0.2	22 ±8.9	3.3 ±0.1	2.8 ±0.1	86.6 ±7.2	68.5 ±4.2	65.3 ±13.1	46.2 ±28.8
Eresen87	5.4 ±0	4.3 ±0.4	119.8 ±26.3	95.5 ±6	4.1 ±0.1	14.8 ±0.1	1 ±0	16.3 ±5.2	3.9 ±0.1	3.9 ±0.1	100 ±0	76.1 ±2.6	158.4 ±2	94.3 ±40.9
Farah	5.3 ±0.5	5 ±0.7	110.5 ±11	98.5 ±18.7	5.4 ±0.3	8.9 ±0.1	1.5 ±0.4	33.7 ±7.7	3.6 ±0	3.6 ±0	100 ±0	70.1 ±0.6	44.7 ±7.3	43.7 ±1
Fiesta	5.8 ±0	5.8 ±0.7	74.2 ±7.4	122.5 ±1.8	3.9 ±0.4	7.8 ±0.3	1.6 ±0	27.9 ±5.4	3.2 ±0	2.9 ±0.1	90.6 ±2.2	69.7 ±0.2	122.3 ±23.7	58.9 ±21
Filiz99	5.6 ±0	7 ±0.3	125.4 ±14.4	69.5 ±1.8	4.4 ±0	12.6 ±0.2	1.1 ±0.1	22.9 ±0.6	3.2 ±0.1	2.8 ±0.1	88.2 ±8.3	74.6 ±2.3	156.9 ±2.4	71.5 ±30.3
Fiord	5.8 ±0.1	6.6 ±0.7	73.6 ±7.8	102.5 ±21.6	5.4 ±0	8.5 ±1.2	1.9 ±0.1	43.1 ±0.1	3 ±0	3 ±0	100 ±0	67 ±0.6	59.2 ±1.3	53.2 ±14.6

(cont. on next page)

Table B.1. (cont.).

Icarus	5 ±0	6 ±1	95.3 ±11.1	106 ±13.4	4.3 ±0.5	9.9 ±0.2	1.3 ±0.2	26.3 ±11.4	3.2 ±0.3	2.9 ±0.2	90.9 ±1.4	74.1 ±1.8	83.6 ±10.1	30.1 ±2.9
Kontu	5.6 ±0.2	4.9 ±0.6	107.6 ±13.9	73.5 ±9.5	5.5 ±0.1	8 ±1	1.9 ±0.1	40.9 ±19.2	3.7 ±0	3.6 ±0	98.6 ±1	67.2 ±5.7	30.5 ±2.9	45.9 ±27.3
Manafest	6 ±0.1	5.3 ±0.2	96.1 ±7.8	108.5 ±11.7	4.9 ±0.1	10.2 ±0.7	1.3 ±0.1	29 ±0.6	3.2 ±0	2.7 ±0.2	84.4 ±6.6	71.7 ±4.1	102.3 ±16.4	67.1 ±15.6
Melodie/2	5.9 ±0.1	2.7 ±0.1	110.5 ±13.1	112 ±17.7	6.1 ±0.7	9.2 ±0.8	1.3 ±0.2	38 ±8.1	3 ±0.1	2.9 ±0.1	96.9 ±2.2	75.2 ±0.6	73.7 ±3.9	71.9 ±20.6
Mikko	4.3 ±0	4.2 ±0.1	126 ±14.8	95.5 ±6	2.6 ±0.3	7.6 ±0.7	2.4 ±0.2	23.6 ±2.1	3.7 ±0.1	3.6 ±0.1	97.2 ±2	75.7 ±3.9	29.4 ±4.1	21.9 ±9.5
Nura	5 ±0.3	3.8 ±0.1	108 ±10.6	111 ±9.9	4.3 ±0.1	8.6 ±0.8	1.7 ±0.1	30.4 ±9.1	3.3 ±0.1	3.1 ±0.1	93.9 ±0.1	71 ±4.8	65.5 ±10.8	40.5 ±3.6
PBARana	5.9 ±0.1	7.2 ±1.8	106.2 ±4.4	108.5 ±11.7	5 ±0	10.5 ±1.1	1.1 ±0.1	28 ±1.9	3.3 ±0.1	3.2 ±0	97.1 ±2.1	75.7 ±2.2	82 ±2.3	60 ±12.3
Salkim	5.3 ±0	6.7 ±0	109 ±19.6	102.5 ±11	4.2 ±0.5	13.7 ±0.3	1.1 ±0.1	17.3 ±4.9	3.8 ±0.1	3.3 ±0.2	87.5 ±8.8	72.9 ±0.8	154.5 ±12.1	78.8 ±29.1
Witkiem Manida	5.6 ±0	4.1 ±0.2	89.7 ±5.9	66 ±4.2	6.1 ±0.1	14.5 ±0.5	1.1 ±0.1	12.3 ±0.5	3.6 ±0	3.4 ±0.1	94.4 ±3.9	67.7 ±0.8	130.7 ±4.2	41.5 ±6.5

Table B.2. Data for the 12 qualitative characteristics of 114 faba bean genotypes averaged over two seasons.

Name	Leaflet shape	Stipule spot pigmentation	Stem pigmentation at flowering	Mature stem color	Intensity of petal streaks	Wing petal color	Pod angle	Pod shape	Pod surface reflectance	Seed shape	Seed coat color	Hilum color
CGN07723	1	0	0	2	0	1	3	3	2	1	9	2
CGN07734	2	+	3	2	3	3	1	1	1	3	1	1
CGN07781	3	+	3	1	5	3	1	1	2	1	4	1
CGN07836	2	+	3	2	3	3	1	2	1	3	3	1
CGN07871	2	+	3	2	5	3	1	1	2	3	3	1
CGN07933	2	+	3	2	3	3	1	1	2	3	1	1
CGN10320	1	+	3	2	5	3	2	1	2	1	4	1
CGN10322	2	+	5	2	5	3	1	2	2	1	3	1
CGN10330	2	+	5	2	5	3	1	1	1	3	3	1
CGN10362	1	+	7	2	7	3	1	1	1	1	3	x
CGN10371	2	+	7	2	7	3	1	1	1	1	3	1
CGN10382	2	+	5	1	5	3	1	1	1	1	3	1
CGN10382.1	2	+	3	1	5	3	1	1	2	1	x	1
CGN10384	2	+	5	1	5	3	2	2	2	1	4	1
CGN10385	2	+	5	2	5	3	1	2	1	3	3	1
CGN10387	2	+	5	2	3	3	1	1	2	3	3	1
CGN10391	1	+	5	2	5	3	3	1	2	1	3	2
CGN12320	2	+	5	2	3	3	1	1	2	1	3	1
CGN12321	3	+	5	2	3	3	1	1	2	1	4	1
CGN13445	3	+	3	2	7	3	1	1	1	3	3	1
CGN13464	3	+	5	2	5	3	1	1	2	3	3	1
CGN13487	3	+	3	2	5	3	1	1	2	3	3	1
CGN13518	2	+	3	2	7	3	2	1	2	1	3	x
CGN15556	2	+	5	2	3	3	1	1	2	3	3	1
CGN15615	2	+	7	2	7	3	1	2	1	3	3	1
CGN15620	2	+	5	2	5	3	1	2	2	3	3	1
CGN15621	2	+	7	2	5	3	3	2	1	1	4	1
CGN15639	3	+	5	2	5	3	1	2	1	3	3	1
CGN15641	2	0	x	1		1	1	2	2	3	9	2
CGN15644	2	+	3	2	5	3	1	1	2	1	4	1
CGN18856	3	+	3	2	3	3	2	2	2	1	4	x
CGN18860	2	0	0	2	0	1	1	1	2	3	9	2
CGN18862	2	0	0	2	0	1	2	1	2	1	9	2
CGN18867	3	+	5	2	3	3	2	2	2	1	2	1
CGN18867.1	2	+	5	2	3	3	1	1	2	1	3	x
CGN18878	3	+	5	2	5	3	1	1	2	3	3	1
CGN18892	2	+	5	2	3	3	1	1	2	3	3	1
CGN18893	2	+	5	2	7	3	1	2	1	3	3	1
CGN18905	2	0	0	1	0	1	1	2	2	1	9	2
CGN18906	2	+	7	1	7	3	1	2	1	1	4	1
CGN18909	2	+	5	2	5	3	1	1	2	3	2	1

(cont. on next page)

Table B.2. (cont.).

CGN18920	1	+	7	2	5	3	1	1	2	3	3	1
CGN18923	2	+	3	2	5	3	1	1	2	3	4	1
CGN18927	2	+	3	2	3	3	1	1	2	3	4	1
CGN18934	2	+	5	2	5	3	1	1	2	3	3	1
CGN18941	2	+	5	2	3	3	1	1	2	3	3	1
CGN18942	2	+	7	2	7	3	1	1	2	1	4	x
CGN19979	1	+	7	2	3	3	2	2	1	1	7	x
CGN19981	1	+	3	2	5	3	3	1	2	1	3	1
CGN19982	3	+	3	2	7	3	2	2	2	1	3	x
CGN19985	1	+	5	2	5	3	2	2	2	1	4	1
CGN19986	2	+	5	1	7	3	2	1	2	1	4	1
CGN19987	2	+	7	2	7	3	3	2	1	1	4	1
CGN19987.1	1	+	7	2	5	3	1	2	2	1	4	1
CGN19993	2	+	3	2	3	3	2	1	2	1	9	x
CGN07697	2	+	3	2	5	3	1	1	1	3	3	1
CGN07699	2	+	7	2	7	3	1	1	2	3	3	1
CGN07717	2	+	5	2	3	3	1	1	1	1	3	1
CGN07719	2	+	5	2	5	3	1	1	2	3	3	1
CGN07721	3	+	7	1	7	3	2	2	2	1	7	1
CGN07725	1	+	7	2	7	3	1	1	2	3	3	1
CGN07728	3	+	5	1	3	3	1	1	2	3	4	1
CGN07730	3	+	5	2	7	3	1	1	2	3	x	1
CGN07730.1	3	+	3	1	7	3	1	1	2	3	3	1
CGN07733	2	+	5	2	7	3	1	2	2	1	4	1
CGN07735	2	+	3	2	3	3	1	1	1	3	3	1
CGN07736	1	+	7	2	7	3	1	1	2	3	3	1
CGN07738	2	+	5	2	5	3	1	2	1	3	3	1
CGN07739	1	+	5	2	5	3	1	1	2	3	3	1
CGN07740	2	+	3	2	3	3	1	1	2	3	3	1
CGN07751	1	+	5	2	7	3	1	1	2	3	2	1
CGN07757	3	+	3	2	3	3	1	1	2	1	4	1
CGN07766	2	+	5	1	7	3	1	1	1	1	x	1
CGN07826	2	+	5	2	5	3	1	2	2	1	7	1
CGN07827	2	+	3	2	5	3	1	1	2	3	3	1
CGN07827.1	2	+	7	2	3	3	2	1	1	1	5	x
CGN07842	2	+	5	2	5	3	1	1	1	3	3	1
CGN07843	2	+	7	2	7	3	1	2	1	1	3	x
CGN07844	2	+	5	2	5	3	1	2	1	3	3	1
CGN07934	2	+	5	1	3	3	1	1	2	3	3	2
CGN07938	2	+	3	2	3	3	1	1	2	3	9	1
NGB1540.1	3	+	3	2	5	3	1	1	2	3	x	1
NGB1542.1	1	+	5	2	7	3	1	1	2	3	3	1
NGB1547.1	3	+	7	2	7	3	1	1	2	1	7	1
NGB1550.1	2	+	5	2	7	3	1	1	2	3	2	1
NGB20019.2	2	+	7	2	5	3	1	2	2	1	1	1
NGB8640	2	+	7	1	5	3	1	1	2	3	7	1
NGB8642	1	+	7	2	7	3	1	1	2	3	3	1
NGB8643	2	+	5	2	5	3	2	2	2	3	1	1
TR12540	2	+	7	2	7	3	1	2	2	1	3	1
TR23018	2	+	5	1	3	3	1	1	2	3	4	1
TR28096	2	+	5	2	7	3	1	2	1	1	4	1

(cont. on next page)

Table B.2. (cont.).

TR31590	2	+	5	1	5	3	1	1	2	1	3	x
TR31912	2	+	5	2	7	3	1	1	2	1	3	1
TR44862	2	+	5	2	7	3	1	1	2	1	4	1
TR53947	2	+	7	2	7	3	2	2	2	1	4	2
TR71255	1	+	7	2	5	3	2	2	1	1	x	1
TR75421	2	+	7	2	7	3	1	1	2	1	4	1
Aquadulce	2	+	3	2	5	3	2	1	2	1	4	1
Ascot	2	+	7	2	5	3	1	1	2	3	4	1
Eresen87	2	+	5	2	7	3	2	1	2	1	4	1
Farah	2	+	5	2	5	3	1	2	1	1	3	1
Fiesta	2	+	5	2	5	3	1	1	2	1	3	1
Filiz99	2	+	5	2	7	3	2	2	2	1	3	1
Fiord	2	+	7	2	5	3	1	2	2	3	3	1
Icarus	1	+	7	2	7	3	1	1	2	3	x	1
Kontu	2	+	5	1	5	3	1	1	1	3	3	1
Manafest	1	+	5	1	5	3	1	2	2	1	3	1
Melodie/2	2	+	5	2	7	3	1	1	2	3	3	x
Mikko	2	+	3	1	5	3	1	1	1	3	1	1
Nura	2	+	5	2	7	3	1	1	1	1	3	1
PBARana	2	+	7	2	7	3	1	2	1	1	3	1
Salkim	2	+	5	2	5	3	2	1	2	1	4	1
Witkiem Manida	1	+	5	2	7	3	3	1	2	1	4	x

All the traits were assessed as described in Table 2.3.

Table B.3. Correlation coefficient values between morphological traits of faba bean.

Traits	Leaflets per leaf	Stem per plant	Plant height, cm	Days to flowering	Flowers per inflorescence	Pod length, cm	Number of pods per node	Number of total pods per plant	Maximum ovules per pod	Number of seeds per pod	Seeds fertilized %	Water content of fresh seeds, %	100-seed weight, g
Leaflets per leaf	1.000												
Stem per plant	0.135	1.000											
Plant height	-0.151	0.093	1.000										
Days to flowering	-0.164	0.069	0.057	1.000									
Flowers per inflorescence	0.160	-0.140	0.132	-0.016	1.000								
Pod length	0.111	0.106	0.141	-0.179	0.042	1.000							
Number of pods per node	-0.131	0.018	0.063	0.144	-0.122	-0.427**	1.000						
Number of total pods per plant	-0.042	0.265**	0.324**	0.014	-0.020	-0.300**	0.370**	1.000					
Maximum ovules per pod	0.013	-0.186*	0.133	-0.074	0.139	0.493**	-0.195*	-0.174	1.000				
Number of seeds per pod	-0.061	-0.228*	0.226*	-0.046	0.166	0.350**	-0.051	-0.075	0.876**	1.000			
Seeds fertilized	-0.104	-0.013	0.238*	0.100	0.083	-0.070	0.251**	0.109	-0.033	0.329**	1.000		

(cont. on next page)

Table B.3. (cont.).

Water content of fresh seeds, %	0.058	0.221*	0.379**	0.288**	0.052	0.169	-0.111	-0.075	0.006	0.043	0.130	1.000	
100-seed weight, g	0.201*	0.202*	0.000	-0.074	-0.141	0.691**	-0.551**	-0.358**	0.062	-0.109	-0.242**	0.148	1.000
Seed yield, g	0.045	0.329**	0.380**	-0.055	-0.049	0.263**	-0.029	0.612**	0.005	0.058	0.079	0.074	0.297**
Leaflet shape	-0.018	-0.078	0.455**	-0.052	0.125	0.016	0.138	0.292**	0.110	0.239*	0.237*	0.217*	-0.163
Stipule spot pigmentation	0.023	0.072	0.051	0.015	-0.068	-0.205*	0.041	0.171	-0.232*	-0.089	0.037	-0.133	-0.068
Stem pigmentation at flowering	0.220*	0.077	-0.190*	-0.038	-0.060	0.087	-0.174	-0.118	-0.170	-0.182	-0.019	-0.109	0.221*
Mature stem color	-0.104	-0.088	-0.092	0.045	-0.040	0.065	-0.155	-0.124	0.049	0.033	-0.125	-0.077	0.079
Petal streak intensity	0.219*	0.055	0.029	0.043	-0.043	0.105	-0.109	-0.067	-0.180	-0.113	0.097	0.035	0.170
Wing petal color	0.023	0.072	0.051	0.015	-0.068	-0.205*	0.041	0.171	-0.232*	-0.089	0.037	-0.133	-0.068
Pod angle	-0.011	0.058	-0.122	-0.207*	-0.091	0.625**	-0.349**	-0.417**	0.355**	0.205*	-0.134	0.119	0.538**
Pod shape	0.055	0.055	-0.114	-0.200*	-0.174	0.441**	-0.149	-0.126	0.058	-0.047	-0.070	-0.101	0.375**
Pod surface reflectance	0.116	0.073	0.048	0.020	0.094	0.064	-0.052	-0.123	-0.048	-0.097	-0.033	0.283**	0.081
Seed shape	-0.174	-0.297**	0.024	0.103	0.004	-0.568**	0.373**	0.199*	-0.100	-0.016	0.089	-0.194*	-0.717**
Seed coat color	0.143	0.156	-0.028	-0.167	0.052	0.448**	-0.152	-0.146	0.137	0.022	-0.057	0.210*	0.348**
Hilum color	-0.134	-0.021	-0.078	-0.003	0.047	0.212*	-0.082	-0.187	0.152	-0.024	-0.137	0.147	0.114

** Correlation is significant at the 0.01 level (2-tailed)

* correlation is significant at the 0.05 level (2-tailed)

Table B.3. Cont., Correlation coefficient values between morphological trait averages of faba bean.

Traits	Seed yield, g	Leaflet shape	Stipule spot pigmentation	Stem pigmentation at	Mature stem color	Petal streak intensity	Wing petal color	Pod angle	Pod shape	Pod surface reflectance	Seed shape	Seed coat color	Hilum color
Leaflets per leaf													
Stem per plant													
Plant height													
Days to flowering													
Flowers per inflorescence													
Pod length													
Number of pods per node													
Number of total pods per plant													
Maximum ovules per pod													
Number of seeds per pod													

(cont. on next page)

Table B.3. (cont.).

Seeds fertilized													
Water content of fresh seeds, %													
100-seed weight, g													
Seed yield, g	1.000												
Leaflet shape	0.305**	1.000											
Stipule spot pigmentation	0.209*	0.069	1.000										
Stem pigmentation at flowering	0.102	-0.247**	0.345**	1.000									
Mature stem color	-0.083	-0.140	0.134	0.054	1.000								
Petal streak intensity	0.107	-0.087	0.341**	0.457**	0.026	1.000							
Wing petal color	0.209*	0.069	1.000**	0.345**	0.134	0.341**	1.000						
Pod angle	0.029	-0.159	-0.100	-0.046	0.083	-0.015	-0.100	1.000					
Pod shape	0.144	-0.022	-0.153	0.234*	0.011	0.070	-0.153	0.255**	1.000				
Pod surface reflectance	-0.057	0.014	-0.131	-0.167	-0.009	-0.092	-0.131	0.088	-0.244**	1.000			
Seed shape	-0.329**	0.062	0.039	-0.124	0.063	-0.131	0.039	-0.492**	-0.272**	0.009	1.000		
Seed coat color	0.153	0.041	-0.392**	-0.024	-0.183	-0.129	-0.392**	0.328**	0.103	0.181	-0.397**	1.000	
Hilum color	-0.185	-0.127	-0.778**	-0.220*	-0.151	-0.289**	-0.778**	0.250*	0.130	0.177	-0.100	0.337**	1.000

** Correlation is significant at the 0.01 level (2-tailed)

* correlation is significant at the 0.05 level (2-tailed)

APPENDIX C

BIOCHEMICAL CHARACTERIZATION

Table C.1. Data for seven quantitative biochemical characteristics of 114 faba bean genotypes for two seasons. Averages and standard errors are provided. The range, average and % CV were calculated from these quantitative data.

Name	Phenol mg GAE/100 g DW	Flavonoids mg RE/ 00 g DW	Protein (%)	L-DOPA mg/100g DW	Tannin mg CE/100 g DW	Vicine (%)	Convicine (%)	Total VC (%)
CGN07723	266.9 ± 12	259.8 ± 3.7	24.3 ± 3.7	79.8 ± 1.5	32.6 ± 20.8	0.31	0.16	0.47
CGN07734	280.2 ± 3.2	271 ± 1.6	24.3 ± 1.6	56.4 ± 9.4	93.6 ± 9.4	0.38	0.25	0.63
CGN07781	470.6 ± 57.3	260.6 ± 0.9	24 ± 0.9	41.6 ± 0.4	195.2 ± 40.8	0.32	0.22	0.54
CGN07836	342 ± 9	197.8 ± 0.4	25.8 ± 0.4	25.5 ± 0.2	84.2 ± 6.6	0.34	0.23	0.57
CGN07871	392.9 ± 29.5	219.7 ± 2	24.5 ± 2	79.7 ± 0.5	152.7 ± 20.8	0.37	0.28	0.65
CGN07933	314.3 ± 41.7	187 ± 5.5	30 ± 5.5	25.6 ± 1.1	130.6 ± 45.9	0.29	0.22	0.50
CGN10320	340.3 ± 22.5	157.8 ± 1.4	25.2 ± 1.4	72 ± 0.3	98.5 ± 2.1	0.36	0.21	0.57
CGN10322	397.2 ± 21.2	208 ± 2.5	23.8 ± 2.5	86 ± 0.4	109.4 ± 36.8	0.35	0.27	0.62
CGN10330	412.6 ± 6.7	199.6 ± 1.8	17.7 ± 1.8	22.5 ± 1.4	190.6 ± 41.2	0.33	0.13	0.45
CGN10362	330 ± 27.5	120.2 ± 0.9	16.8 ± 0.9	74 ± 3.1	166.4 ± 10.5	0.24	0.11	0.34
CGN10371	389.3 ± 26	261.5 ± 0.4	26.9 ± 0.4	23.5 ± 0.2	141.1 ± 17.5	0.15	0.36	0.51
CGN10382	361 ± 16.5	231.8 ± 2.7	18.3 ± 2.7	37.7 ± 1.3	83.9 ± 15.9	0.45	0.19	0.65
CGN10382.1	559.2 ± 96.2	271.9 ± 1.5	20.6 ± 1.5	64 ± 0.1	184.7 ± 10.1	0.44	0.19	0.63
CGN10384	394.4 ± 11.7	296.4 ± 1.5	26 ± 1.5	74.4 ± 0.3	174 ± 5.2	0.35	0.26	0.61
CGN10385	431.7 ± 4.9	167.2 ± 8.5	38 ± 8.5	75.9 ± 5.2	213.1 ± 24.8	0.45	0.39	0.84
CGN10387	355 ± 12.5	158.8 ± 0.4	27.8 ± 0.4	33.6 ± 1.6	105.8 ± 15.8	0.49	0.18	0.67
CGN10391	362.5 ± 16.6	255.7 ± 3.6	30.4 ± 3.6	45.8 ± 0.2	102.3 ± 3.1	0.33	0.16	0.49
CGN12320	434.8 ± 28.4	256.8 ± 0.3	21.4 ± 0.3	41.6 ± 0.6	211 ± 51.8	0.32	0.17	0.49
CGN12321	435.7 ± 43.2	218.4 ± 6.8	31.3 ± 6.8	57.3 ± 0	180.3 ± 60.4	0.32	0.21	0.53
CGN13445	435 ± 34.3	292.3 ± 0.9	23 ± 0.9	48.8 ± 0.9	124.1 ± 38.6	0.25	0.20	0.45
CGN13464	480.8 ± 31.1	240.4 ± 0.5	22.3 ± 0.5	44 ± 3	152.8 ± 47.9	0.33	0.22	0.55
CGN13487	395.8 ± 40.6	299.7 ± 1.3	23.4 ± 1.3	52.4 ± 0.5	212.5 ± 24.8	0.42	0.26	0.68
CGN13518	406.3 ± 4.2	242.2 ± 0.4	19.2 ± 0.4	58.8 ± 1.6	141.4 ± 38.8	0.33	0.22	0.55
CGN15556	335.4 ± 10.3	200.1 ± 0.5	26.6 ± 0.5	34.9 ± 0.3	88.8 ± 0.9	0.57	0.22	0.79
CGN15615	353.1 ± 12	257.5 ± 4.5	23.8 ± 4.5	34.1 ± 0.1	156 ± 27.7	0.45	0.19	0.65
CGN15620	400 ± 15.4	196.8 ± 3	30.8 ± 3	31.6 ± 1.3	175.6 ± 14.1	0.38	0.20	0.59
CGN15621	308.6 ± 15.2	220.1 ± 0.9	27.5 ± 0.9	54.4 ± 0	98.2 ± 14.2	0.39	0.28	0.67
CGN15639	385.3 ± 12.8	165.5 ± 0.8	28.1 ± 0.8	26.3 ± 0	154 ± 21.8	0.37	0.23	0.60
CGN15641	280 ± 24.1	295.1 ± 4.3	19.2 ± 4.3	53.8 ± 1	11.9 ± 6.7	0.31	0.26	0.58
CGN15644	414.4 ± 55.4	211.4 ± 1.6	28.9 ± 1.6	34.1 ± 1.3	215.3 ± 7	0.32	0.23	0.55
CGN18856	327.7 ± 44.6	217 ± 1.7	25 ± 1.7	31 ± 0	66.7 ± 23.6	0.44	0.25	0.69
CGN18860	461.7 ± 73.6	267.3 ± 3.1	22.8 ± 3.1	77 ± 3.9	124.1 ± 69.8	0.32	0.23	0.56
CGN18862	323.2 ± 11.7	308.5 ± 3.3	23.6 ± 3.3	22.9 ± 1	22.5 ± 5.3	0.30	0.25	0.56
CGN18867	381.3 ± 28.8	208 ± 0.3	24.9 ± 0.3	38.7 ± 0.9	99.3 ± 5.8	0.35	0.19	0.54
CGN18867.1	352.7 ± 48.6	256 ± 0.5	26.8 ± 0.5	41.9 ± 2.6	95.4 ± 21	0.36	0.20	0.57
CGN18878	424 ± 42.9	263.4 ± 1.2	23.3 ± 1.2	61.6 ± 3.2	220.6 ± 88.3	0.37	0.23	0.60

(cont. on next page)

Table C.1. (cont.).

CGN18892	312.8 ± 5	271.2 ± 1.8	24.8 ± 1.8	81.6 ± 0.4	26 ± 12.7	0.37	0.16	0.53
CGN18893	493.7 ± 7.8	280.8 ± 1.5	27.1 ± 1.5	100.4 ± 4.8	277.1 ± 20.8	0.36	0.13	0.50
CGN18905	308.1 ± 27.2	182.3 ± 2.1	30.1 ± 2.1	24.8 ± 2.7	41.6 ± 7.8	0.66	0.28	0.94
CGN18906	348.5 ± 22.2	152.5 ± 6.3	27.6 ± 6.3	21.5 ± 0.3	126.5 ± 4.4	0.25	0.17	0.41
CGN18909	554.9 ± 37.7	283.4 ± 2.5	26.4 ± 2.5	113.4 ± 0.6	197.5 ± 68.6	0.44	0.14	0.58
CGN18920	411.7 ± 7.3	167.2 ± 0.1	20.1 ± 0.1	57.3 ± 0	173.5 ± 24	0.34	0.22	0.57
CGN18923	513.1 ± 27.6	230.1 ± 1.2	25.7 ± 1.2	42 ± 0.2	249.7 ± 29.6	0.33	0.34	0.67
CGN18927	359.9 ± 8.2	260.8 ± 3.5	19.8 ± 3.5	89.7 ± 1.8	74.5 ± 10.9	0.21	0.15	0.36
CGN18934	428 ± 38.4	210.9 ± 3.8	27.2 ± 3.8	35.2 ± 0.4	170.4 ± 28.1	0.35	0.20	0.55
CGN18941	396.8 ± 41.4	136.4 ± 0	22.5 ± 0	43.9 ± 2.9	187.4 ± 38.1	0.32	0.29	0.61
CGN18942	314.3 ± 21.4	217.1 ± 3.6	28.7 ± 3.6	22.9 ± 0.4	135.1 ± 0.1	0.35	0.14	0.49
CGN19979	359.2 ± 22	184.5 ± 0.2	27.5 ± 0.2	29.4 ± 0.8	105.4 ± 21.4	0.28	0.17	0.45
CGN19981	287.4 ± 7.9	159.9 ± 2.3	21.9 ± 2.3	35.7 ± 0.7	134.2 ± 8.6	0.28	0.20	0.48
CGN19982	450.3 ± 37.7	175.3 ± 2.3	22.8 ± 2.3	52.1 ± 0.9	184.5 ± 19.5	0.34	0.20	0.54
CGN19985	388.4 ± 5.8	140.3 ± 0.2	28 ± 0.2	18.6 ± 0.1	162.7 ± 20.1	0.33	0.12	0.44
CGN19986	415.5 ± 16.1	265.5 ± 0.1	25.5 ± 0.1	60.8 ± 0.1	208.3 ± 51.2	0.28	0.18	0.46
CGN19987	337.4 ± 4.6	228.6 ± 3.7	31.7 ± 3.7	20.5 ± 0.6	194 ± 8.2	0.38	0.17	0.55
CGN19987.1	354.1 ± 12.3	202.3 ± 2.3	30.2 ± 2.3	74.5 ± 3.4	85.8 ± 10	0.28	0.28	0.56
CGN19993	261.1 ± 51.8	195.4 ± 0.4	23.8 ± 0.4	73.2 ± 2.4	175.6 ± 16.9	0.32	0.24	0.56
CGN07697	322.2 ± 7.4	149.2 ± 2.7	25.8 ± 2.7	53.6 ± 5.3	168.7 ± 45.6	0.34	0.30	0.64
CGN07699	351.9 ± 3.9	151.8 ± 3.8	35.7 ± 3.8	46.8 ± 1.6	133.2 ± 16.3	0.34	0.10	0.45
CGN07717	458.3 ± 29.7	209.5 ± 0.7	19.5 ± 0.7	71.4 ± 2	271.4 ± 46.4	0.27	0.23	0.50
CGN07719	422 ± 62.5	264.1 ± 4	25.8 ± 4	70.2 ± 2.4	112.4 ± 25.8	0.44	0.17	0.61
CGN07721	388.4 ± 52.4	160.8 ± 2.3	25 ± 2.3	31.1 ± 3	109.1 ± 13.7	0.38	0.13	0.51
CGN07725	306.4 ± 35.9	123.3 ± 1.3	27.9 ± 1.3	45.9 ± 0.2	110.1 ± 4.8	0.32	0.18	0.50
CGN07728	440.7 ± 21.9	248.8 ± 0	19.9 ± 0	73.3 ± 2.8	193.3 ± 47.2	0.34	0.18	0.52
CGN07730	262.9 ± 11.5	225.2 ± 4.4	32.6 ± 4.4	19.8 ± 1	101.8 ± 21.2	0.26	0.21	0.47
CGN07730.1	443 ± 42.9	343.2 ± 0.9	32 ± 0.9	55.9 ± 1.4	121.7 ± 33.8	0.30	0.23	0.53
CGN07733	446.3 ± 1.1	160.9 ± 0.1	28 ± 0.1	28 ± 0.4	211.1 ± 8.5	0.33	0.23	0.57
CGN07735	437.2 ± 51.6	182.8 ± 0.8	28.5 ± 0.8	28.9 ± 0	156.4 ± 63.1	0.58	0.33	0.91
CGN07736	386 ± 40.7	263.5 ± 7	31.4 ± 7	48.4 ± 0.2	102.5 ± 28.4	0.36	0.21	0.58
CGN07738	329.4 ± 3	234.9 ± 1.7	25.2 ± 1.7	53.3 ± 2.2	95.5 ± 31.1	0.40	0.23	0.63
CGN07739	332.6 ± 12.6	293.7 ± 0.2	27.8 ± 0.2	22.8 ± 1.2	88.2 ± 30.3	0.31	0.18	0.49
CGN07740	369.3 ± 4.6	191.9 ± 1.8	26 ± 1.8	42.9 ± 3.8	145.6 ± 24	0.46	0.20	0.66
CGN07751	338.5 ± 17.2	297.5 ± 0.7	22.3 ± 0.7	9.2 ± 0.4	194.9 ± 19.6	0.33	0.39	0.72
CGN07757	313.3 ± 2.8	306.6 ± 2.4	26.8 ± 2.4	100.9 ± 0.9	109.2 ± 1.2	0.39	0.25	0.65
CGN07766	395.3 ± 16.9	221.4 ± 1.2	26.2 ± 1.2	37 ± 0.3	172.1 ± 61.8	0.27	0.21	0.47
CGN07826	388.5 ± 48	261.6 ± 6.9	32.2 ± 6.9	55.1 ± 2.2	188.1 ± 35.2	0.39	0.15	0.54
CGN07827	426.4 ± 26.9	255.7 ± 1.2	25.1 ± 1.2	40.6 ± 0.6	125.1 ± 16.8	0.32	0.29	0.61
CGN07827.1	318.4 ± 32.1	184.3 ± 11.2	40.6 ± 11.2	34 ± 1.9	119.9 ± 32.9	0.31	0.17	0.48
CGN07842	457.2 ± 13.3	259.6 ± 0.6	22.4 ± 0.6	32.5 ± 0.3	199 ± 4.4	0.34	0.16	0.50
CGN07843	378.2 ± 22.7	160.1 ± 4.4	28 ± 4.4	47.9 ± 0.8	158.1 ± 75.3	0.39	0.09	0.48
CGN07844	358.3 ± 36.4	232.6 ± 12.1	40.6 ± 12.1	44.6 ± 1.3	152.8 ± 44.3	0.36	0.17	0.53
CGN07934	360.5 ± 8.9	186.5 ± 0.2	21.1 ± 0.2	95.7 ± 0.2	141.8 ± 11.7	0.48	0.22	0.70
CGN07938	385.7 ± 22.1	231.7 ± 0.1	19.6 ± 0.1	75.4 ± 0.8	103 ± 4.6	0.29	0.13	0.42
NGB1540.1	422.3 ± 9.2	239.1 ± 2.5	25.8 ± 2.5	54.9 ± 0.5	170.8 ± 56.9	0.32	0.14	0.46
NGB1542.1	495 ± 17.3	223.8 ± 1.1	25.6 ± 1.1	59.1 ± 0.4	160.6 ± 27.2	0.37	0.25	0.62
NGB1547.1	518.2 ± 16.7	269.5 ± 0.2	28 ± 0.2	51.2 ± 1.2	247 ± 18.1	0.35	0.28	0.64
NGB1550.1	478.1 ± 43.9	229.6 ± 0.6	22.5 ± 0.6	70.5 ± 2.3	206.7 ± 68.3	0.44	0.26	0.70
NGB20019.2	427.1 ± 21.7	249.2 ± 11.5	40.8 ± 11.5	33.5 ± 0.3	161 ± 44.5	0.30	0.20	0.50
NGB8640	396.1 ± 4.1	226.7 ± 0.4	26.7 ± 0.4	54.1 ± 0.6	146.5 ± 3.6	0.32	0.23	0.55
NGB8642	395 ± 11.5	244.8 ± 0.7	22.9 ± 0.7	57.3 ± 0.3	119 ± 17.8	0.34	0.14	0.48
NGB8643	458.6 ± 26.9	224.2 ± 0.5	24.1 ± 0.5	16.5 ± 0.2	188.5 ± 44.8	0.30	0.16	0.45
TR12540	389 ± 46.1	199.2 ± 7.4	29.2 ± 7.4	18.8 ± 0.3	136.4 ± 6.1	0.31	0.20	0.51
TR23018	369.3 ± 9.1	210.3 ± 1.1	30.4 ± 1.1	32.8 ± 0.4	204.8 ± 16.9	0.39	0.19	0.58
TR28096	330.9 ± 6.2	244.6 ± 3	31.1 ± 3	32.4 ± 1.8	95 ± 11.2	0.43	0.18	0.61
TR31590	377.7 ± 19.4	206.1 ± 8.2	33 ± 8.2	27.1 ± 0.3	108.8 ± 3.5	0.33	0.18	0.51
TR31912	353.3 ± 25.5	182 ± 4.4	35.5 ± 4.4	64.6 ± 1.4	151.6 ± 16.1	0.29	0.16	0.45

(cont. on next page)

Table C.1. (cont.).

TR44862	346.6 ± 14.4	188.7 ± 0.9	25.6 ± 0.9	32.6 ± 0.5	94.2 ± 3.6	0.36	0.19	0.55
TR53947	346.6 ± 14.4	188.7 ± 0.9	25.6 ± 0.9	32.6 ± 0.5	94.2 ± 3.6	0.36	0.19	0.55
TR71255	379.4 ± 17.6	147.7 ± 4.2	18.3 ± 4.2	25.9 ± 0.8	170.9 ± 49.1	0.34	0.17	0.51
TR75421	374.1 ± 0.8	251.4 ± 3.3	20.9 ± 3.3	34 ± 0.2	198.6 ± 22.8	0.32	0.28	0.60
Aquadulce	409.8 ± 21.7	246.1 ± 1.5	23.4 ± 1.5	52 ± 0.1	171.3 ± 2.9	0.38	0.25	0.64
Ascot	428.6 ± 18.3	145.7 ± 1.1	24.3 ± 1.1	50.1 ± 0.9	199.3 ± 32	0.26	0.22	0.47
Eresen87	435.5 ± 3.2	238.2 ± 2	21.2 ± 2	25.2 ± 1.1	154.2 ± 6.9	0.29	0.21	0.51
Farah	364.5 ± 42.4	249.3 ± 0.7	22.9 ± 0.7	22.8 ± 0.3	134.3 ± 7.2	0.31	0.31	0.62
Fiesta	396.8 ± 12.6	215 ± 0.1	20.2 ± 0.1	29.5 ± 0.3	177.9 ± 30.1	0.26	0.25	0.51
Filiz99	431.6 ± 33.3	231.7 ± 4.3	18.8 ± 4.3	25.6 ± 0	199.5 ± 2.6	0.37	0.20	0.58
Fiord	329.3 ± 9.9	201.3 ± 0.1	27.3 ± 0.1	43.8 ± 0.4	120.8 ± 12	0.32	0.16	0.48
Icarus	447.2 ± 5.5	230.6 ± 2.1	24.4 ± 2.1	36.7 ± 0.7	241.5 ± 14.3	0.31	0.22	0.53
Kontu	469.9 ± 20.5	130 ± 0.7	28.8 ± 0.7	59.2 ± 0	211 ± 2.2	0.37	0.26	0.63
Manafest	420.8 ± 36.2	212.1 ± 3.4	25.5 ± 3.4	33.7 ± 0	125.7 ± 5.6	0.31	0.26	0.56
Melodie/2	407.4 ± 39.1	170.2 ± 3.5	24.4 ± 3.5	27.9 ± 0.5	142.2 ± 49	0.43	0.13	0.55
Mikko	370.6 ± 11.5	163.8 ± 4.3	29.3 ± 4.3	39.9 ± 0.5	101.9 ± 6.9	0.02	0.01	0.03
Nura	415.7 ± 5.7	260.8 ± 8.3	29.6 ± 8.3	49.1 ± 0	190.1 ± 0.1	0.35	0.32	0.67
PBARana	382 ± 13.4	201.4 ± 1.7	24.9 ± 1.7	61.1 ± 2.3	155.1 ± 17.6	0.36	0.26	0.62
Salkim	371.6 ± 4.1	145.8 ± 0.8	23.8 ± 0.8	23.9 ± 0.5	101 ± 20.7	0.25	0.15	0.40
Witkiem Manida	290.7 ± 7.1	235.3 ± 0.2	21.4 ± 0.2	53.4 ± 0.5	56.4 ± 25.7	0.44	0.21	0.65

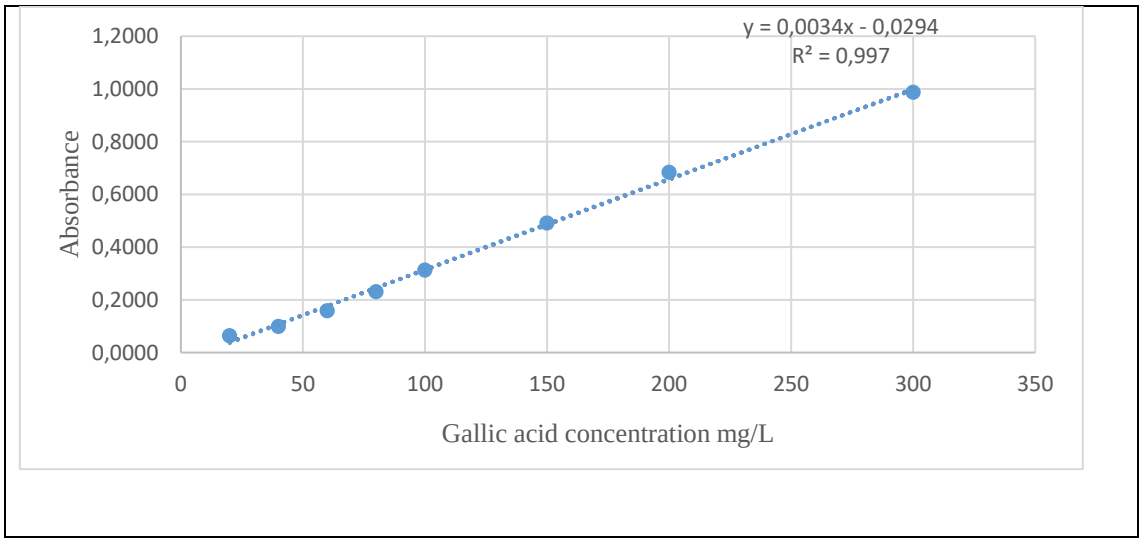

Figure C.1. Calibration curve of gallic acid used to estimate the total phenolic content in faba bean samples

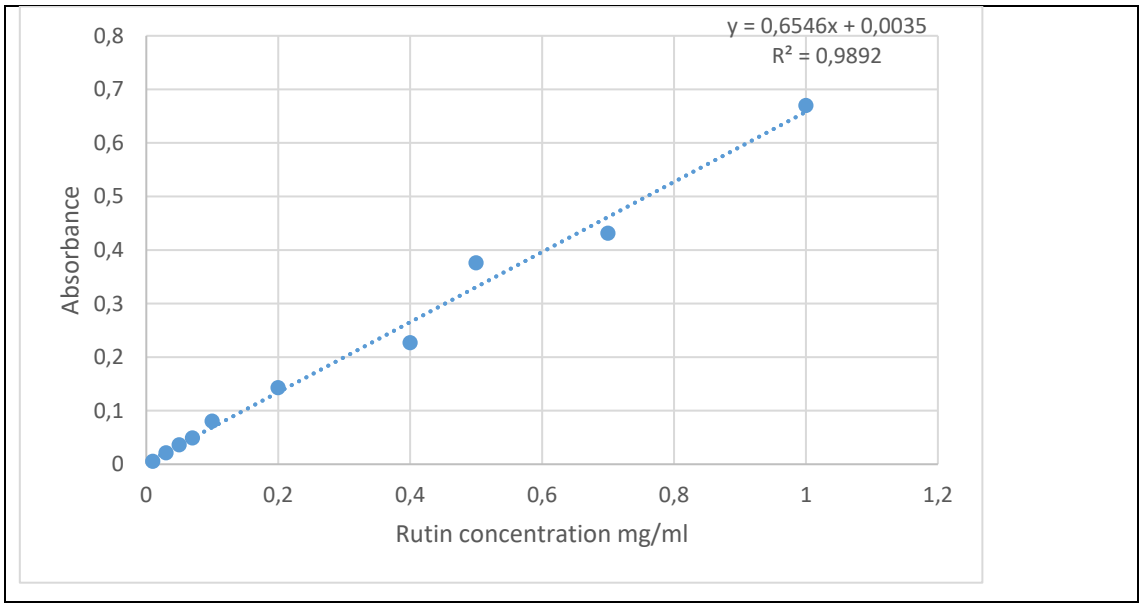

Figure C.2. Calibration curve of rutin used to estimate the flavonoids content of faba bean samples.

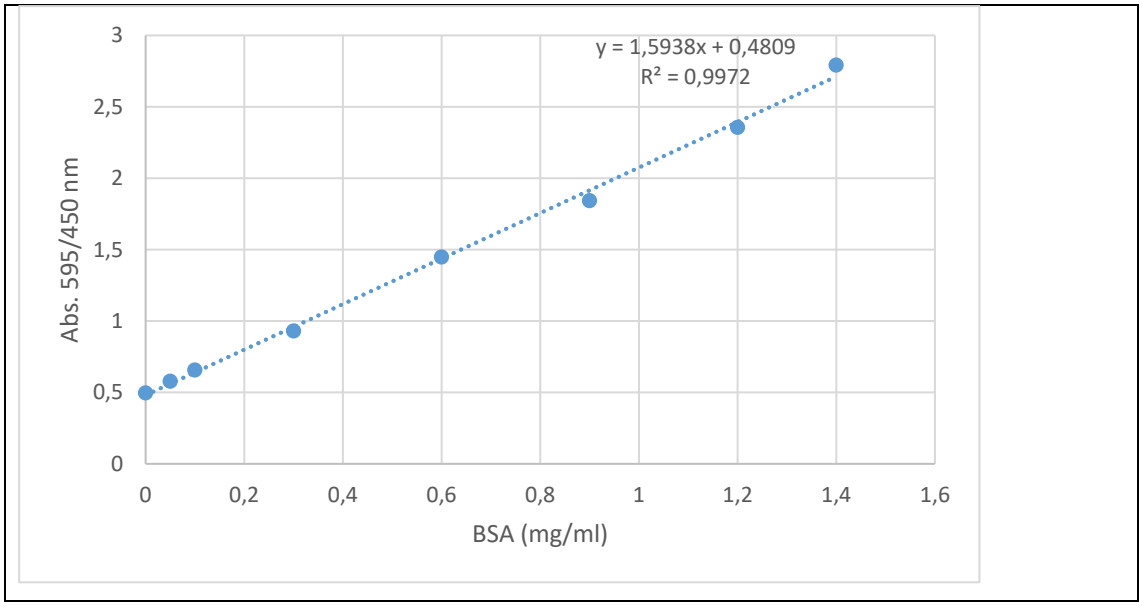

Figure C.3. BSA standard calibration curve with linearization.

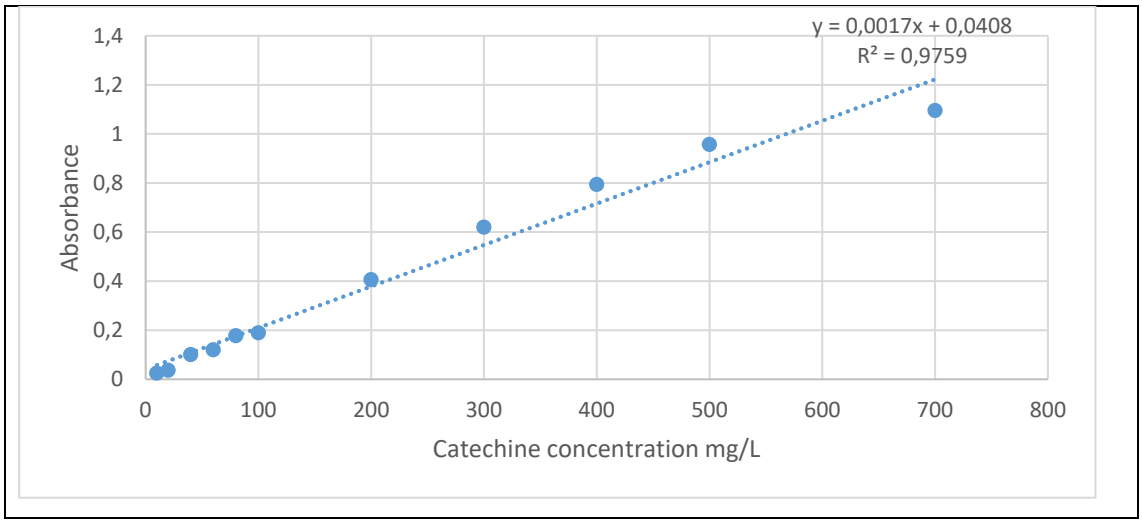

Figure C.4. Calibration curve of catechine used to estimate the tannins content of faba bean samples.

VITA

Mazen Ali ABUZAYED

Date and Place of Birth: 1970 Gaza/ Palestine

E-mail: mazenabuzayed@gmail.com

Education

Ph.D. Molecular Biology and Genetics
2012 - 2019 İzmir Institute of Technology
Molecular Biology and Genetics Department, İzmir, Turkey

MSc. Biological Science
2008 - 2011 Islamic University of Gaza (IUGaza)
Biology and Biotechnology Department, Gaza, Palestine

BSc. Biology - Botany
1990 – 1994 Alfateh University
Botany Department, Libya, Tripoli

Teaching and Professional Experience

Research assistant

2008 - 2011 Islamic University of Gaza (IUGaza)
Biology and Biotechnology Department, Gaza, Palestine

Selected Publications

Abuzayed M, Frary A, Doğanlar S (2018) Genetic diversity of some Palestinian and Turkish olive (*Olea europaea* L.) germplasm determined with SSR markers. IUG Journal of Natural Studies 26(1): 10-17.

Abuzayed M, Göktay M, Allmer J, Doğanlar S, Frary A (2017) Development of genomic simple sequence repeat markers in faba bean by next generation sequencing. Plant Molecular Biology Reporter 35(1): 61-71.

Abuzayed M, ElDabba N, Frary A, Doğanlar S (2017) GDdom: an online tool for calculation of dominant marker gene diversity. Biochemical Genetics 55: 155-157.