A multi-parent faba bean (*Vicia faba* L.) population for future genomic studies

Hamid Khazaei[1]*, Frederick L. Stoddard[2], Randy W. Purves[1], Albert Vandenberg[1]

[1] *Department of Plant Sciences, University of Saskatchewan, 51 Campus Drive, Saskatoon, S7N 5A8, Canada*

[2] *Department of Agricultural Sciences, Viikki Plant Science Centre, University of Helsinki, 00014, Helsinki, Finland*

*Corresponding author: H. Khazaei, ORCID ID: 0000-0002-5202-8764, contact: hamid.khazaei@gmail.com, or call +1(306) 966-5859.

Published in Plant Genetic Resources: Characterization and Utilization (2018) 16: 419-423. Doi 10.1017/S1479262118000242

Abstract

Faba bean (*Vicia faba* L.) is a valuable grain legume and a staple protein crop in many countries. Its large and complex genome requires novel approaches for its genetic dissection, such as the development of a multi-parent advanced generation intercross (MAGIC) population. Here we introduce a MAGIC population developed from four founders (ILB 938/2, Disco/2, IG 114476 and IG 13238). The selection of parental lines was based on geographic (Ecuador, France, Bangladesh and China), genetic, and phenotypic diversity. The parental lines were inbred and then genotyped using 875 SNP (single nucleotide polymorphism) markers. Based on molecular data the parents had high homozygosity and high genetic distance among them. The population segregates for several important traits such as seed morphology, seed chemistry, phenology, plant architecture, drought response, yield and its components, and resistance to *Botrytis fabae*. The dynamics of the population were examined by using segregation patterns of simply inherited Mendelian traits such as stipule spot pigmentation (SSP) and flower colour at different generations. All 1200 four-way cross F1 plants had pigmented flowers and stipule spots. The segregation ratios for white flower colour (single gene, *zt2*) fit 7:1, 13:3 and 25:7 at F2, F3 and F4 generations, respectively, and the segregation ratio of SSP (two recessive unlinked genes, *ssp1* and *ssp2*) fit 49:15 and 169:87 at the F2 and F3 generations, respectively, demonstrating unbiased generation advance. We will subject the F5 generation of this population to a high-throughput SNP array and make it available for further phenotyping and genotyping.

Keywords: faba bean; multi-parent population; MAGIC, phenotyping; pigmentation

Introduction

Faba bean (*Vicia faba* L.) is a high protein, nutritious grain legume crop. Its mixed breeding system and unknown wild progenitor along with its large and complex genome [2n=2x=12, haploid genome size ~ 13 Gbp (Soltis et al., 2003)] are important challenges for faba bean genetic and breeding studies (O'Sullivan and Angra, 2016).

Bi-parental crosses used to develop F2, RIL (recombinant inbred line) and BC (backcross) populations for genetic studies have been used to map QTL (quantitative trait loci) for genetic traits of interest in faba bean (Khazaei, 2014; Webb et al., 2016). However, these types of genetic populations are limited by their genetic recombination level and only two alleles are present at any locus (Singh and Singh, 2015). In contrast, multi-parent advanced generation intercross (MAGIC) populations present multiple alleles, resulting in increased recombination and mapping resolution for complex traits (Cavanagh et al., 2008). The multi-parent nature of MAGIC populations increases the intercrossing and shuffling of the genome, which makes them better suited for genetic studies, particularly for faba bean which has limited genomic resources. MAGIC populations also provide deeper knowledge on genomic structure and improve pre-breeding resources (Huang et al., 2015). They are powerful tools for identifying alleles and loci responsible for economically important traits, especially with the recent availability of affordable high-throughput genotyping platforms and advances in statistical methods of data analysis (Meng et al., 2016).

MAGIC populations have been developed and analysed in various plant species (*reviewed in* Huang et al., 2015). In faba bean, a multi-parent population derived from 11 European winter bean collection was employed to identify genomic regions governing frost adaptation, using 189 individuals and 156 SNP (single nucleotide polymorphism) markers (Sallam and Martsch, 2015).

The current paper introduces a large faba bean population derived from intercrossing four inbred lines that are genetically contrasting for important agro-morphological traits including drought adaptation, seed morphology, seed chemistry and disease resistance. The segregation patterns of Mendelian traits such as pigmentation of stipule spots and flowers at different generations are presented.

Experimental details

Parental line selection

The four faba bean inbred lines used to create the MAGIC population and their morphological and biochemical characteristics of interest are shown in Table 1. More information on potential traits for phenotyping including physiological traits and seed morphology and seed chemistry is presented in Table S1. Each parental line was selfed for three generations to reach a higher level of homozygosity before F1 crosses were made.

Disco/2 is a selection from the zero-tannin cultivar Disco carrying the *zt2* gene (Picard 1976; Gutierrez et al., 2008) and low vicine-convicine (*vc*$^-$ gene, Duc et al., 1989; confirmed in Khazaei et al., 2017 and Purves et al., 2018). Disco was released in France as a source of combined zero tannin and low v-c (FEVITA type, Duc et al., 2004).

ILB 938/2 is an Ecuadorian faba bean accession with high water use efficiency (Link et al., 1999; Khazaei et al., 2013; Khazaei et al., 2014b) and resistance to chocolate spot (*Botrytis fabae*, Maalouf et al., 2016). ILB 938/2 also carries a gene that decouples pigmentation in flowers from pigmentation in stipules (Khazaei et al., 2014a).

IG 114476 is a *Paucijuga*-type faba bean with small seeds (1000 seed weight of ~ 150 g), short stature, and highly branched main stem growth habit (personal observation).

IG 13238 is an early maturing faba bean accession (Stoddard et al., 2016) from China.

Hybridization

Bi-parental crosses (Disco/2 × ILB 938/2) and (IG 114476 × IG 132238) were made, and then a large number of four-way crosses [(Disco/2 × ILB 938/2) × (IG 114476 × IG 132238)] were made in the insect-proof greenhouse of the Department of Agricultural Sciences, the

University of Helsinki, Finland in 2014 (Figure S1). One thousand two hundred four-way F1 plants were grown and generations were advanced by single seed descent to the F2, F3 and F4 in the controlled environment conditions facility of the University of Saskatchewan, Canada. In all generations, the growing conditions were maintained in a similar manner. The photoperiod of the growth unit was set to 14 / 10 h (light / dark). The photosynthetic photon flux density was ~ 300 μmol m^{-2} s^{-1} at the plant canopy level. The temperature was adjusted to 21 °C light / 15 °C dark.

In all three generations, the flower colour and stipule spot pigmentation (SSP) phenotypes were recorded for each plant during population development. The goodness-of-fit of observed segregation ratio to expected ratios for both flower colour and SSP were determined using the standard chi-squared (χ^2) test (R statistics, R Development Core Team, 2016). The expected number of offspring was obtained from the hypothesis of one (flower colour) and two (SSP) unlinked recessive genes.

Parental lines genotyping

The four inbred parental lines were genotyped using 875 SNP markers developed by Webb et al. (2016). DNA was extracted from a seedling leaf of a single plant of each parent lines. Nine discs (diameter 5 mm) from healthy, newly expanded leaves of the parental lines were shipped to the LGC Genomics laboratory (LGC Genomics, Beverly, MA, USA) for genomic DNA extraction and SNP genotyping according to the manufacturer's instructions.

Results and Discussion

The four inbred lines used as parents were highly homozygous. ILB 938/2 and IG 114476 showed the highest homozygosity among parental lines (Table 1). The genotyping call on parental lines is presented in Table S2. Polymorphic SNPs were 24% when only two bi-parental crosses were considered, but about 47% of SNP markers were polymorphic among lines from four-way crosses. The haplotype diversity of parental lines over 678 mapped SNP markers from the consensus map (Webb et al., 2016) is presented in Table S3. The Medtr2g009270 gene, collinear with the low v-c gene (Khazaei et al., 2017), faba bean flanking markers on chromosome 1 both could distinguish Disco/2 as low v-c (Table S3, parental lines phenotyping for v-c was presented Table S1).

This population combines important key genes for faba bean breeding such as drought response, flowering date, maturity date, height, and seed morphology (seed size and seed coat colour) and seed chemistry (v-c, L-DOPA, folate and tannin, *see* Table S1). These traits make the population suitable for further development of genetic studies and for pre-breeding genetic sources. It also offers the opportunity to develop individuals with the combinations of superior alleles that directly contribute to breeding programs.

The dynamic in the population was examined by morphological markers (pigmentations) on individuals in each generation. All of the 1200 four-way F1 plants had pigmented stipules and flowers (Table 2). The F2 segregation ratio for *zt2* confirmed a monogenic Mendelian inheritance (7 coloured:1 colourless), with wild-type flower being dominant to white flower. The F2 segregation ratio for *ssp* fit 49:15 (pigmented:colourless) at the F2 generation, which fits a model of double-recessive complementary behaviour between two unlinked genes in a four-way cross. In the F3 generation, the segregation

pattern fit segregation ratios of 13:3 for *zt2* and 169:87 for *ssp* (Table 2). In the F4 generation, the ratios fit 25:7 for the *zt2* gene.

The four parental lines were chosen from different agro-ecological zones, providing a basis for genetic analysis of agro-ecological adaptation. They were representative of four diverse germplasm pools (Mediterranean, subtropical, warm temperate and cool temperate climates) from three continents, from high and low latitudes and altitudes, and representative of both food and feed types. For example, IG 114476 is from the sub-tropical savannah of Bangladesh, ILB 938/2 is from Mediterranean adaptation zone and Disco/2 is from the temperate climate of northern latitudes. IG 132238 is representative of Chinese germplasm which is considered to be distinct from all other faba bean germplasm (Duc et al., 2010). The UPGMA cluster analysis results confirmed that the four parental lines are genetically distinct (Figure S2). In the winter faba bean population (Sallam and Martsch, 2015), 11 founders from UK, France and Germany were distributed in three main groups, but this was not based on their country of origin.

The segregation of flower colour and stipule pigmentation as conferred by the *zt2* and *ssp* genes confirmed the dynamics of the population (Gutierrez et al., 2008; Khazaei et al., 2014a). Phenotyping for more complex traits such as yield components, physiological and biochemical traits may also be considered when genotyping data become available for this population. Development of molecular markers is an important step toward map-based genetic analysis and can help dissect genotype-phenotype relationships and build genome sequencing assemblies. The current population will be genotyped using 50 K SNP array in near future which will provide a valuable genetic tool for collaborative faba bean research. The plant materials will be available to others researchers worldwide based on available seed stored in our inventory.

Acknowledgments

The authors acknowledge funding from Emil Aaltonen Foundation (Emil Aaltosen Säätiö), the University of Helsinki, Finland, the Agriculture Development Fund (Government of Saskatchewan, Canada) and the NORFAB project (Protein for the Northern Hemisphere) funded by Innovation Fund Denmark. HK also thanks Markku Tykkyläinen, Sini Lindstrom and Sanna Peltola, technical assistants of the glasshouse of the University of Helsinki for their kind assistance during the experiments in Finland, and Allison Sackville and Michael Miller at the University of Saskatchewan, Canada. Many thanks to Professor Wolfgang Link (Georg-August-University, Göttingen, Germany) for providing the seeds of ILB 938/2 and Disco/2 and Dr. Kenneth Street for providing ICARDA germplasm accessions.

References

Cavanagh C, Morell M, Mackay I and Powell W (2008) From mutations to MAGIC: resources for gene discovery, validation and delivery in crop plants. *Current Opinion in Plant Biology* 11: 215–221.

Duc G, Bao SY, Baum M, Redden B, Sadiki M, Suso MJ, Vishniakova M and Zong XX (2010) Diversity maintenance and use of *Vicia faba* L. genetic resources. *Field Crops Research* 115: 270–278.

Duc G, Crepon K, Marget P and Muel F (2004) ''Fevita'' a common name for a new type of faba been cultivar. In: AEP (Eds.), Proceedings of the 5th European Conference on Grain Legumes and ICLGG 2004, 2004/06/07-11, Dijon, France, p. 403.

Duc G, Sixdenier G, Lila M and Furstoss V (1989) Search of genetic variability for vicine and convicine content in *Vicia faba* L. A first report of a gene which codes for nearly zero-vicine and zero-convicine contents. In: Huisman, J., van der Poel AF, Liener IE (Eds.), Recent Advances of Research in Antinutritional Factors in Legume Seeds. Wageningen Pers (Pbs.), pp. 305–313.

Gutierrez N, Avila CM, Moreno MT and Torres AM (2008) Development of SCAR markers linked to *zt-2*, one of the genes controlling absence of tannins in faba bean. *Australian Journal of Agricultural Research* 59: 62–68.

Huang BE, Verbyla KL, Verbyla AP, Raghavan C, Singh VK, Gaur P, Leung H, Varshney RK, and Cavanagh CR (2015) MAGIC populations in crops: current status and future prospects. *Theoretical and Applied Genetics* 128: 999–1017.

Khamassi K, Ben Jeddi F, Hobbs D, Irigoyen J, Stoddard F, O'Sullivan DM and Jones H (2013) A baseline study of vicine–convicine levels in faba bean (*Vicia faba* L.) germplasm. *Plant Genetic Resources: Characterization and Utilization* 11: 250–257.

Khazaei H (2014) Leaf traits associated with drought adaptation in faba bean (*Vicia faba* L.). Doctoral thesis. University of Helsinki, Finland. Available at https://helda.helsinki.fi/handle/10138/135919

Khazaei H, O'Sullivan DM, Sillanpää MJ and Stoddard FL (2014a) Genetic analysis reveals a novel locus in *Vicia faba* decoupling pigmentation in the flower from that in the extra-floral nectaries. *Molecular Breeding* 34: 1507–1513.

Khazaei H, O'Sullivan DM, Sillanpää MJ and Stoddard FL (2014b) Use of synteny to identify candidate genes underlying QTL controlling stomatal traits in faba bean (*Vicia faba* L.). *Theoretical and Applied Genetics* 127: 2371–2385.

Khazaei H, Purves RW, Song M, Stonehouse R, Bett KE, Stoddard FL and Vandenberg A (2017) Development and validation of a robust, breeder-friendly molecular marker for the *vc*$^{-}$ locus in faba bean. *Molecular Breeding* (in press).

Khazaei H, Street K, Bari A and Stoddard FL (2013) Do faba bean (*Vicia faba* L.) accessions from environments with contrasting seasonal moisture availabilities differ in stomatal characteristics and related traits? *Genetic Resources and Crop Evolution* 60: 2343–2357.

Link W, Abdelmula AA, Kittlitz E von, Bruns S, Riemer H and Stelling D (1999) Genotypic variation for drought tolerance in *Vicia faba*. *Plant Breeding* 118: 477–483.

Maalouf F, Ahmed S, Shaaban K, Bassam B, Nawar F, Singh M and Amri A (2016) New faba bean germplasm with multiple resistances to Ascochyta blight, chocolate spot and rust diseases. *Euphytica* 211: 157–167.

Meng L, Guo L, Ponce K, Zhao X and Ye G (2016) Characterization of three rice multiparent advanced generation intercross (MAGIC) populations for quantitative trait loci identification. *The Plant Genome* 9: 1–14. doi: 10.3835/plantgenome2015.10.0109

O'Sullivan DM and Angra D (2016) Advances in faba bean genetics and genomics. *Frontiers in Genetics* 7: 150. doi: 10.3389/fgene.2016.00150

Picard J (1976) Aperçu sur l'héredité du caractère absence de tannins dans les graines de féverole (*Vicia faba* L.). *Annales de l'amélioration des plantes* 26: 101–106.

Purves RW, Khazaei H and Vandenberg A (2018) Quantification of vicine and convicine in faba bean seeds using hydrophilic interaction liquid chromatography. *Food Chemistry* 240: 1137–1145.

R Development Core Team (2016) R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria. ISBN 3–900051–07–0. Available: http://www.R-project.org.

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* 143: 501–514.

Singh BD and Singh AK (2015) Mapping Populations. In: *Marker-Assisted Plant Breeding: Principles and Practices*. New Delhi: Springer India; 2015. pp. 125–150.

Soltis DE, Soltis PS, Bennett MD and Leitch IJ (2003) Evolution of genome size in the angiosperms. *American Journal of Botany* 90: 1596–1603.

Stoddard FL, Khazaei H and Belachew KY (2016) Exploiting germplasm resources for climate change adaptation in faba bean. In: *Applied Mathematics and Omics to Assess Crop Genetic Resources for Climate Change Adaptive Traits*. Bari A, Damania AB, Street K, Mackay M, Dayanandan S (eds). CRC press. ISBN 9781498730136 pp. 207–214.

Webb A, Cottage A, Wood T, Khamassi K, Hobbs D, Gostkiewicz K, White M, Khazaei H, Ali M, Street D, Stoddard FL, Maalouf F, Ogbonnaya F, Link W, Thomas J and O'Sullivan DM (2016) A SNP-based consensus genetic map for synteny-based trait targeting in faba bean (*Vicia faba* L.). *Plant Biotechnology Journal* 14: 177–185.

Table 1. Percent homozygosity, six key phenotypes and origin of four faba bean inbred lines used in this study.

Inbred line	**% Homozygosity**	**Hilum colour**	**Flower colour**	**Seed coat colour**	**Stipule spot pigmentation**	**Vicine-convicine**	**Seed size**	**Flowering / Maturity**	**Origin / Source**
Disco/2	97.8	Colourless	White	Beige	Colourless	Low	Small (400g)	late	INRA[c], France
ILB 938/2 (IG[a] 13987)	99.3	Black	Spotted[b]	Green	Colourless	High	Equina (600g)	intermediate	Ecuador
IG 114476	99.4	Black	Spotted	Dark brown	Spotted	High	Paucijuga (150g)	early	Bangladesh
IG 132238	99.0	Black	Spotted	Beige	Spotted	High	Small-flat (330g)	early	China

[a] *IG*, ICARDA (International Centre for Agricultural Research in the Dry Areas) accession number. More information about accessions can be found at http://www.genesys-pgr.org/

[b] Flowers display a dark spot on the wing petals and also dark vein marking on the standard petal (wild type).

[c] *INRA*, Institut National de la Recherche Agronomique.

Table 2. Observed offspring segregation ratios, values of χ^2 test and corresponding *P* value for studied traits / genes among F1 four-way crosses, F2 and F3 generations.

	Four-way F1 [a]	**F2**				**F3**			
Trait / Gene		**Pigmented**	**Non-pigmented**	**χ^2 (7:1)**	***P***	**Pigmented**	**Non-pigmented**	**χ^2 (13:3)**	***P***
Flower colour [b] (*zt2*)	Spotted	1038	153	0.131	0.718	943	219	0.007	0.933
				χ^2 (49:15)				**χ^2 (169:87)**	
Stipule spot pigmentation [c] (*ssp1* and *ssp2*)	Spotted	892	304	2.614	0.106	750	424	2.377	0.123

[a] 1200 plants.

[b] F4 (Spotted 843 : white 247, χ^2 (25:7): 0.394, $P = 0.530$).

[c] F4 expected ratio is 625:399 (spotted: colourless).