

Advances in faba bean genetics and genomics

Article

Accepted Version

O'Sullivan, D. M. and Angra, D. (2016) Advances in faba bean genetics and genomics. *Frontiers in Genetics*, 7. 150. ISSN 1664-8021 doi: <https://doi.org/10.3389/fgene.2016.00150>
Available at <https://centaur.reading.ac.uk/67667/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

Published version at: <http://dx.doi.org/10.3389/fgene.2016.00150>

To link to this article DOI: <http://dx.doi.org/10.3389/fgene.2016.00150>

Publisher: Frontiers Media

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

1 **Advances in faba bean genetics and genomics**

2 **Running title:** Faba bean genetics and genomics

3 Donal M. O'Sullivan, Deepti Angra.
4 School of Agriculture, Policy and Development,
5 University of Reading,
6 PO Box 237,
7 Whiteknights,
8 Reading,
9 RG6 6AR,
10 United Kingdom.
11 d.m.osullivan@reading.ac.uk

12

13 **Abstract:**

14 *Vicia faba* L, is a globally important grain legume whose main centres of diversity are the Fertile
15 Crescent and Mediterranean basin. Because of its small number (six) of exceptionally large and easily
16 observed chromosomes it became a model species for plant cytogenetics the 70s and 80s. It is
17 somewhat ironic therefore, that the emergence of more genomically tractable model plant species
18 such as *Arabidopsis* and *Medicago* coincided with a marked decline in genome research on the
19 formerly favoured plant cytogenetic model. Thus, as ever higher density molecular marker coverage
20 and dense genetic and even complete genome sequence maps of key crop and model species emerged
21 through the 1990s and early 2000s, genetic and genome knowledge of *Vicia faba* lagged far behind
22 other grain legumes such as soybean, common bean and pea.

23 However, cheap sequencing technologies have stimulated the production of deep transcriptome
24 coverage from several tissue types and numerous distinct cultivars in recent years. This has permitted
25 the reconstruction of the faba bean meta-transcriptome and has fuelled development of extensive sets
26 of Simple Sequence Repeat and Single Nucleotide Polymorphism (SNP) markers.

27 Genetics of faba bean stretches back to the 1930s, but it was not until 1993 that DNA markers were
28 used to construct genetic maps. A series of Random Amplified Polymorphic DNA-based genetic
29 studies mainly targeted at quantitative loci underlying resistance to a series of biotic and abiotic
30 stresses were conducted during the 1990's and early 2000s. More recently, SNP-based genetic maps
31 have permitted chromosome intervals of interest to be aligned to collinear segments of sequenced
32 legume genomes such as the model legume *Medicago truncatula*, which in turn opens up the
33 possibility for hypotheses on gene content, order and function to be translated from model to crop.
34 Some examples of where knowledge of gene content and function have already been productively
35 exploited are discussed.

36 The bottleneck in associating genes and their functions has therefore moved from locating gene
37 candidates to validating their function and the last part of this review covers mutagenesis and genetic
38 transformation, two complementary routes to validating gene function and unlocking novel trait
39 variation for the improvement of this important grain legume.

40

41 **Keywords:** *Vicia faba*; transcriptome; single nucleotide polymorphism; genetic linkage map; synteny.

42 Introduction

43 Origins and importance of *Vicia faba* in world agriculture

44 *Vicia faba* L.(*Vf*) or faba bean is a grain legume of great importance in world agriculture due to its
45 high yield potential compared to alternative grain legumes, its ability to fix nitrogen through
46 symbiosis with *Rhizobium leguminosarum* in its root nodules, but most crucially for its role as a staple
47 dietary protein source in North African and Middle Eastern cultures. The species is thought to have
48 been domesticated in the Eastern Mediterranean region, perhaps somewhere between Afghanistan and
49 the Eastern Mediterranean (van de Wouw et al., 2001), but no extant wild relative has yet been found
50 although new species closely related to *Vf* have been found in recent decades (Maxted, 1993).
51 Nonetheless, there is great variability within the domesticated genepool, with the major centre of
52 diversity centred around the Mediterranean basin and secondary centres of diversity in the Nile
53 Valley, South America, and Central and Eastern Asia (Duc et al., 2010), providing much untapped
54 potential to breeders. This diversity of forms is exemplified by the botanical classification of ‘*major*’
55 (large-seeded or ‘broad’ bean), ‘*equina*’ (mid-sized or ‘horse’ bean) and ‘*minor*’ (small, rounded
56 seed) types. *Vf* is one of the earliest domesticated Old World agricultural crops, with credible
57 archaeobotanical evidence linking it to the pre-pottery Neolithic period around 10,000 BP in a site in
58 North-West Syria (Tanno and Willcox, 2006). However, again in keeping with its great diversity and
59 adaptability, it has a long history of utilisation outside the centre of origin in diverse agroecological
60 settings from the Boreal and Atlantic Maritime climates to arid and sub-tropical regions. The most
61 current available statistics for dry beans show it ranking 7th in terms of global production of grain
62 legumes with production in 2013 of 1.17 million tonnes (FAOSTAT, 2013), though this probably
63 grossly underestimates the nutritional and societal importance of the species as a majority of faba
64 bean worldwide is cultivated and consumed locally by subsistence farmers. Furthermore, FAOSTAT
65 figures relate to dry beans and there are no reliable global data for large-seeded broad bean types
66 picked green and consumed fresh. China is the leading producer, followed currently by Ethiopia,
67 Australia, France, Egypt, Morocco, Sudan and the UK, though rank order amongst the latter seven
68 countries has historically been somewhat fluid. Despite the strong cultural attachment of
69 Mediterranean basin and Middle Eastern populations to consumption of faba bean, for a variety of
70 reasons addressed below, many of these countries have become net importers.

71 Challenges facing faba bean

72 Yield stability

73 Faba bean is generally speaking the most productive of grain legumes in environments where rainfall
74 is not limiting or in irrigated conditions, and can indeed be a highly profitable crop, especially if the
75 economic benefits of biologically fixed nitrogen and enhanced weed and disease control in subsequent
76 crops are considered (Preissel et al., 2015). However, in common with many legumes, its yields are
77 relatively unstable, which is thought to be one important reason underlying the low inclusion rate of
78 leguminous crops in European agriculture in particular (Cernay et al., 2015). This inherent yield
79 instability is thought to be due, in part, to its apparently profligate flowering habit whereby many
80 flowers spread over many nodes are produced containing far more ovules than the capacity of the
81 plant to fertilize and fill those potential seed sites, resulting in a high but variable rate of flower and
82 fertilized ovule abortion. Optimal yield production in faba bean is also dependent on symbiosis with
83 *Rhizobium leguminosarum* biovar *viciae* to produce nitrogen-fixing root nodules as well as on the
84 pollination services of wild bee populations to ensure both optimal seed set and outcrossing rates.
85 Pollinator insufficiency has been shown recently in the UK to explain up to 64% yield loss (Nayak et
86 al., 2015) while a large N-fixation response to rhizobial inoculation in the field has been shown even

87 in the presence of high natural rhizobial populations (Denton et al., 2013). On top of these variable
88 symbiotic or mutualistic interactions are layered a host of abiotic and biotic stresses, combinations of
89 which will exact further environment-dependent losses. The many specific pests and pathogens of
90 faba bean include fungi (*Ascochyta fabae* – ascochyta blight, *Botrytis fabae* – chocolate spot,
91 *Uromyces fabae* – bean rust, *Peronospora viciae* f.sp. *faba* - downy mildew), insects (*Aphis fabae* –
92 black bean aphid, *Bruchus rufimanus* – the bruchid seed beetle), nematodes (*Ditylenchus gigas* – stem
93 nematode) and parasitic plants (*Orobanche crenata* and *Orobanche foetida* – broomrape) to name just
94 a few of the most prominent. Detailing the fascinating coevolutionary struggles between these diverse
95 organisms and their common host is beyond the scope of this review and interested readers can find
96 further details in one of many recent reviews of biotic stresses on faba bean (Sillero et al., 2010
97 Stoddard et al., 2010; Pérez-de-Luque et al., 2010). Chief amongst abiotic stresses are heat and
98 drought, which play all the more important a role in yield determination because of the sensitivity of
99 flower fertility to even mild or transient levels of stress. Fertility of developed flowers has recently
100 been shown to be highly sensitive to heat stress, though total yield losses are mitigated by the shift of
101 yield production to higher nodes formed after a stress event has passed (Bishop et al., 2016b). And a
102 further study showed that bee pollination can significantly mitigate damage to fertility caused by heat
103 stress (Bishop et al., 2016a). Cold/frost tolerance is an important trait for winter beans (Sallam et al.,
104 2015), and tolerance to salinity soils is pertinent in semi-arid regions (Tavakkoli et al., 2012). Again,
105 abiotic stress resistance is a wide topic which has been adequately covered in recent reviews (Patrick
106 and Stoddard, 2010 Link et al., 2010, Khan et al., 2010).

107 **Quality**

108 Meeting the protein demand of a growing global population represents a challenge not only from the
109 yield perspective but also from the quality perspective. As yield potential continues to increase
110 through the efforts of breeders, it is important for the protein density of the crop to be maintained, if
111 not improved, from typical values of around 30% dry weight. The well-established relative deficiency
112 of faba bean in essential sulphur-containing amino acids, cysteine and methionine, can of course be
113 balanced in food and feed processing, but amelioration of the balance of essential amino acids would
114 be a valuable breeding target which has not yet been addressed (Multari et al., 2015). Finally, in order
115 for the nutritive value of faba protein to be maximized and nitrogen- and phosphorous-rich wastes in
116 the food chain minimized, serious research is needed to identify and remove anti-nutritional factors
117 that inhibit the normal digestion of starch and protein in the gut (amylase and protease inhibitors), that
118 cause oxidative stress (e.g. vicine) or that sequester key nutrients (e.g. phytic acid). Much research on
119 faba bean quality has focussed on two key anti-nutritional factors: tannins and vicine/convicine.
120 Vicine and convicine are particularly intriguing as the metabolized derivatives of these pyrimidine
121 glucosides cause a serious and potentially fatal condition known as favism in genetically predisposed
122 humans and are found only in a handful of *Vicia* spp, the most notable of which is faba bean (Ray and
123 Georges, 2010). Seed coat tannins and vicine-convicine are naturally high in a large majority of
124 current varieties (e.g. Khamassi et al., 2013) and have been shown to lower protein digestibility and
125 energy content in a variety of animal feeding studies. The nutritional quality of faba bean and status of
126 research on anti-nutritional factors have been thoroughly reviewed by Crepon et al., (2010).

127 Against this backdrop of high yield potential and high utility in increasing the sustainability of
128 nitrogen cycling in agro-ecosystems and security of protein supply in the food chain, it is both timely
129 and appropriate to review how genetic and genomics are already contributing to the accelerated
130 breeding of high yielding, climate-resilient and nutritious faba beans and how we can expect these
131 technologies to be exploited in the future.

132 Genomics

133 Transcriptomes

134 The first major contribution to systematic faba bean transcriptome knowledge was the release of
135 approximately 5,000 Expressed Sequence Tags (EST) from developing embryos of a broad bean
136 variety ‘Windsor’ described by (Ray and Georges, 2010). This study provided a useful snapshot of the
137 functional classification and relative expression level of the more abundant transcripts from the
138 embryo transcriptome in the early to middle stages of its development, even though restricted to one
139 genotype and one tissue. Kaur et al., (2012) and Yang et al., (2012) undertook 454 sequencing of the
140 transcriptome specifically to underpin SSR discovery and this significantly increased the volume of
141 transcriptome data available although in both cases from mixed genotypes. Subsequently, studies
142 began to encompass multiple, separate inbred genotypes (Webb et al., 2016; Ocaña et al., 2015) as
143 well as multiple tissues and genotypes (Ray et al., 2015). The deepest transcriptome coverage yet
144 produced has come from Illumina sequencing of a library of mixed tissues enriched with embryo
145 transfer cells from variety ‘Fiord’ (Arun-Chinnappa and McCurdy, 2015; Zhang et al., 2015). At the
146 time of writing, transcriptomes from nine specifically identified single genotypes and a selection of
147 tissues including whole seedling, root, shoot, leaf, seed coat and embryo were available (summarised
148 in **Table 1**), which offers for the first time to the faba bean research community the possibility to
149 conduct crude electronic Northern analyses and to mine genotypic variants from multiple genetic
150 backgrounds.

151 SNPs

152 The extreme paucity of *Vf* sequence of any description in public databases, lamented by previous
153 reviewers e.g. Gnanasambandam et al., (2012), meant that very small numbers of SNPs had been
154 discovered prior to 2014. The first genetic linkage map of faba bean to explicitly target gene-based
155 polymorphisms was reported by Ellwood et al., (2008), who adopted a strategy of cross-species
156 amplification of conserved orthologues in order to identify polymorphic intron-targeted markers,
157 which were implemented at first either as Cleaved Amplified Polymorphic Sequence (CAPS) or
158 Single Nucleotide Primer Extension (SNuPE) assays. Later, many of these polymorphic intron
159 sequences were converted from CAPS/SNuPE to Kompetitive Allele Specific PCR (KASP) format
160 (Cottage et al., 2012), which made the exploitation of this first suite of SNPs more accessible. The
161 advent of RNA-Seq datasets fuelled the next wave of SNP development. Kaur et al., (2014) designed
162 an iSelect assay based on 768 SNPs, of which 551 were placed on a genetic map generated from a
163 RIL population of the cross Icarus x Ascot. More recently, Webb et al., (2016) reported design of
164 individual KASP assays for 845 SNPs mined from alignment of assembled transcriptomes of ‘Albus’
165 and ‘BPL10’ inbred lines; of these, 653 were successfully mapped. The burgeoning transcriptome
166 datasets described in the previous section permit ever greater numbers of SNPs to be called, for
167 example Ray et al., (2015) reported 5,300 unique variants where alternate alleles could be found in
168 one more of the discovery genotypes SSNS-1, A01155, or CDC Fatima and Ocaña et al., (2015)
169 reported 39,060 SNP and 3,669 InDel polymorphisms in their analysis of transcriptomes of INRA-
170 29H and *Vf*136, though these larger SNP sets remain to be validated as working SNP assays.

171 We have taken this SNP mining exercise further by recently conducting a meta-transcriptome
172 assembly using all *Vf* transcriptome datasets available on 1st October 2015 and scoring mapped reads
173 from individual genotypes for SNP variants. This SNP mining exercise, detailed results of which will
174 be presented elsewhere, resulted in identification so far of more than 320,000 SNPs, of which 69,407
175 have a ‘Freebayes’ SNP quality score >200. Our analysis of variant content of the *VC* interval
176 provides an illustration of how SNP mining can be useful in practice. The *VC* gene was recently

177 mapped to a 4.6 cM interval on the tip of *Vf* chromosome I between SNP markers orthologous to
178 Medtr2g007220 and Medtr2g010740 (Khazaei et al., 2015). As this interval lies within a larger region
179 of largely uninterrupted collinearity with Mt2, it is reasonable to hypothesize that gene order and
180 content within this interval is largely conserved in *Vf*. According to our preliminary analysis, of the
181 308 predicted Medicago (Mt4.0) genes within the interval Medtr2g007220-Medtr2g010740, there is at
182 least partial transcript coverage in the *Vf* meta-transcriptome for 179, and of these, 146 contain a total
183 of 840 high confidence SNPs. **Figure 1** shows a typical informative alignment for a gene towards the
184 centre of this interval. Over the 1534 bp contig, which covers the full coding region predicted in
185 Medicago together with some 5' and 3'-UTR sequence, there are 16 varietal *Vf* SNPs. We are
186 currently in the process of validating a number of SNPs mined from genes predicted by synteny to
187 belong to the VC region, ordering them genetically and using them to further delineate the VC interval
188 (Angra & O'Sullivan, in preparation).

189 **EST-SSRs and genomic SSRs**

190 The other category of molecular marker which has received much attention from the faba bean
191 research community was Expressed Sequence Tag - Simple Sequence Repeats (EST-SSRs). As
192 inherently highly discriminatory, co-dominant markers embedded in genic locations which can be
193 readily associated with orthologous positions in related species, EST-SSRs are an attractive category
194 of marker and offer the advantage that repeats can be mined from a single genetic background with
195 relatively low sequence coverage, so were in the first wave of marker sets to be developed from
196 transcriptome data. Initially modest numbers of EST-SSRs reported by Gong et al., (2011); Ma et al.,
197 (2011) were followed by greater numbers of candidate and validated EST-SSRs (El-Rodeny et al.,
198 2014; Kaur et al., 2012). Although highly portable and informative, EST-SSRs, as with all types of
199 SSR marker, can be difficult to score in a fully automated fashion, and are inherently less amenable
200 than SNPs and INDELs to conversion to high throughput parallel assay formats.

201 In contrast to EST-SSRs, genomic SSRs anchored in non-coding genomic sequence may be difficult
202 to validate as locus-specific assays due to the prevalence of complex, repetitive DNA sequences
203 outside coding regions, which constrain the design of locus-specific primers. Furthermore, most
204 genomic SSRs which fall outside conserved coding regions cannot be used to make syntenic bridges
205 to better characterised relatives. Nonetheless, genomic SSRs, mined from SSR-enriched genomic
206 DNA libraries, have been successfully developed by Zeid et al., (2009) and Yang et al., (2012). **Table**
207 **2** summarises key molecular marker sets which have been developed to date for faba bean.

208 **Genetics**

209 Much progress has been made in the power and possibilities of genetic analysis in faba bean since the
210 pioneering work of Adela Erith, working at the University of Reading in the 1920s, on inheritance of
211 colour, size and form of seeds and other traits (Erith, 1930). Erith's seminal paper showed Mendelian
212 inheritance of genes controlling readily observable (but nonetheless important) characters including
213 hilum colour, seed and flower colour and height. However, many years were to pass before the genes
214 underlying these or any other traits have been identified. Between then and now, *Vf* genetics has
215 passed through distinct phases which will be described below in chronological order.

216 **Classical genetic analysis**

217 The period 1930-1993 was the era of molecular marker-free genetic analysis. During this period and
218 beyond, useful foundations were laid showing simple inheritance patterns for a variety of heritable
219 traits. As mentioned already, Erith, (1930) found single dominant or semi-dominant genes explaining

220 flower, seed and hilum colour as well as plant height, though there was no well-developed locus
221 nomenclature in this early work. Faba bean received much attention in the 1960s to the early 1990s as
222 a model for cytogenetics due to its uncommonly large haploid genome size of 13 Gb (Soltis et al.,
223 2003) and modest haploid chromosomes number ($n=6$), which made for large, readily observable
224 chromosomes. For a time, the study of chromosome breakage response to a range of physical and
225 chemical agents became quite fashionable (e.g. Menke et al., 2000; Milan and Upadhyay, 2007;
226 Rybaczek et al., 2008; Sobita and Bhagirath, 2005) and *Vf* was one of the early crop species whose
227 chromosomes were found to be amenable to flow sorting (Kovarova et al., 2007). The more direct
228 relevance of cytogenetics to modern molecular genetics was the identification of asynaptic mutants
229 (Sjödin, 1970), on which a series of trisomic stocks were founded. Genetic analyses of crosses
230 involving trisomic parents allowed genetic markers to be assigned to physical chromosomes (Patto et
231 al., 1999). The definitive work on pre-molecular faba bean genetics was the monumental work of
232 Sjödin, 1971 to collect, induce, classify and cross mutants and spontaneous variants in a wide variety
233 of characters of interest. Although such comprehensive surveys of genetic variation have not since
234 been carried out, induced and spontaneous mutants affecting a variety of traits have been described in
235 this period. For instance, a gene controlling symbiosis, *sym-1*, was described by Duc and Picard,
236 (1986). Duc et al., (1999) reported a recessive allele of the *il* locus conferring a green cotyledon; this
237 has a similar phenotype as the green cotyledon allele of the *Sc* locus, mapped by Khazaei et al., (2014)
238 to chromosome IV.

239 Work on genetics was not confined to visible morphological phenotypes either. Ramsay, (1997)
240 showed Mendelian inheritance of a major effect seed dormancy gene, denoted *doz*, detected by scoring
241 the timing of seed germination semi-quantitatively in a set of segregating Recombinant Inbred Lines
242 (RILs). Unfortunately, few of the populations generated and characterized in these early genetic
243 studies have been maintained to the present day, though the trait sources (or at least sources carrying
244 analogous phenotypic states) have for the most part. However, quantitative traits, especially those
245 under oligogenic control, were always going to require a molecular mapping approach so that medium
246 or even small size effects could be placed within a sufficiently dense phenotype-independent marker
247 framework.

248 **Molecular genetics**

249 A combination of isozymes and RAPD markers formed the basis of the first *Vf* molecular marker map
250 (Torres et al., 1993), consisting of 66 markers arranged in eleven linkage groups, in which,
251 interestingly, the first indications of synteny between faba bean and other *Fabaceae* were noted.

252 From this point on, *Vf* genetic studies followed a standard pattern. Typically biparental RIL or F_2
253 populations generally consisting of no more than 200 progeny lines were genotyped with non-
254 sequence based markers, chiefly RAPD. The highlights of this phase of faba bean genetics were a
255 series of RAPD-based QTL studies targeting resistance to pathogens and parasites – for instance, to
256 ascochyta (Avila et al., 2004; Diaz-Ruiz et al., 2009), rust (Avila et al., 2003), and *Orobanche* (Diaz-
257 Ruiz et al., 2010; Gutierrez et al., 2013). This early period of integration of molecular marker
258 technology in genetic mapping studies has been reviewed by Torres et al., (2010) and
259 Gnanasambandam et al., (2012).

260 In parallel with the acceleration in growth of faba bean sequence and marker datasets, there has been a
261 correspondingly encouraging increase in the density and utility of gene-based genetic maps. The first
262 gene-based genetic map of *Vicia faba*, composed in the main of 127 co-dominant, portable, Intron-
263 Targeted Amplified Polymorphism (ITAP) markers, was that of Ellwood et al., (2008). A linkage map
264 comprising 128 EST-SSR markers was produced by Ma et al., (2013) and a 552-locus map

265 comprising loci generated from 235 faba bean-derived EST-SSRs and 162 markers derived by cross-
266 amplification from red and white clover was reported recently by El-Rodeny et al., (2014). The
267 consensus map of Satovic et al., (2013) was the first where a majority of markers mapped to just six
268 chromosomally assigned linkage groups, though a minority of these were gene-based, transportable,
269 co-dominant markers. More recently, 551 SNPs and 71 SSRs were combined in Kaur et al., (2014),
270 though not all of the 12 linkage groups reported could be definitively assigned to one of the six
271 physical chromosomes. The densest SNP coverage available in a fully physically anchored consensus
272 linkage map to date is that reported by Webb et al., (2016). This combined 34 SNP markers
273 discovered in *Vf6/Vf27* backgrounds by Ellwood et al., (2008) and converted to KASP format by
274 Cottage et al., (2012) with 653 new ‘Albus’ x ‘BPL10’ SNPs into a single 687-locus consensus map
275 with all markers mapping to just six linkage groups each of which could be assigned to a physical
276 chromosome.

277 **Description of Synteny**

278 The great benefit of the progress in mapping ever greater numbers of sequence-based markers is that
279 the by now well-established conservation of gene order amongst related legume genomes could be
280 used to anchor genetic maps from unsequenced legumes (in this instance faba bean) to the *Medicago*
281 *truncatula* (*Mt*) genome. The first sequence-based genetic map of faba bean which allowed the global
282 pattern of *Vf-Mt* synteny to be observed was that of Ellwood et al., (2008). Although the number of
283 markers (127) was modest, a clear picture of extensive macrosynteny emerged. El-Rodeny et al.,
284 (2014) and Kaur et al., (2014) elaborated this picture with a greater marker density, but the clearest
285 and most complete picture of the extent of macrosynteny between *Vicia faba* and *Medicago*
286 *truncatula* comes from the Webb et al., (2016) study. All 653 newly discovered and mapped SNP
287 markers in this latter study were selected following highly conservative filtering for single copy *Vf*
288 sequences with a clear best reciprocal BLAST hit relationship with a single copy gene in *M.*
289 *truncatula*. Reflecting these stringent marker design criteria, the name of the each new Webb et al.,
290 (2016) *Vf* marker carries explicit reference to its presumed *Mt* orthologue. Combined with the fact that
291 each linkage group in the Webb et al., (2016) consensus map corresponds to a full physical *Vf*
292 chromosomes meant that for the first time all six *Vf* chromosomes could be aligned to the *Mt*
293 sequence without spurious interruptions to macrosynteny pattern caused by lack of marker coverage.

294 **Exploitation of synteny**

295 The hallmark of ‘pre-synteny’ genetic studies in all species, *Vf* being no exception, was that the
296 reporting of a given gene or QTL in proximity of a certain molecular marker was quite frequently the
297 end of the story as there was no method to target markers to a region of interest without screening vast
298 numbers of anonymous markers such as AFLPs or RAPDs.

299 The year 2012 marked a new departure for faba bean with sequence-based and synteny-anchored
300 marker maps being applied to flowering time QTL (Cruz-Izquierdo et al., 2012). Likewise, detection
301 of ascochyta resistance QTL by Kaur et al., (2014) permitted localization of QTL in regions with clear
302 colinearity to fully sequenced model genomes. Khazaei et al., (2014) studying stomatal traits and
303 Khazaei et al., (2015) mapping vicine-convicine content were able to do likewise. In theory now, a
304 given QTL would be associated with an interval whose gene content and even order could be
305 predicted by exploiting synteny with *Medicago* and marker development efficiently targeted to this
306 predicted gene content.

307 As outlined in the previous sections, much of this progress has stemmed from the exploitation of
308 sequencing and cost-effective genotyping technologies to achieve reasonably dense coverage of all six

309 *Vf* chromosomes with sequence-based molecular markers that allow the *Vf* gene-based genetic map to
310 be confidently aligned in large part with fully sequenced reference genomes such as Medicago
311 (Young et al., 2011) or soybean (Schmutz et al., 2010).

312 The exploitation of macrosynteny between crop and model genomes has a number of benefits. Firstly,
313 once a trait has been mapped to a genetic interval which aligns well to a segment of a model
314 sequenced genome, knowledge of gene function in the model species can be translated back to target
315 crop species. In an instance illustrated in **Figure 2**, *ZTI* (controlling flower pigmentation as well as
316 seed coat tannins) was mapped to the *Vf*_Mt3g092810_001 – *Vf*_Mt3g094760_001 interval, with clear
317 synteny to a portion of Mt3. Perusal of the annotation information for the syntenic interval in
318 Medicago revealed a logical biological candidate in the form of the *Transparent Testa Glabra 1*
319 (*TTG1*) WD40 transcription factor (Medtr3g092840), which had previously been shown in *Mt* to
320 determine flower colour, and in the follow-up of this hypothesis, a deletion in the recessive allele
321 sequence was found to plausibly explain the *zt* (unpigmented) phenotype (Webb et al., 2016). This co-
322 called translational genomics approach leverages prior investment in model species biology and fast-
323 tracks causative allele identification. Secondly, quite independently of translation of biological
324 information, molecular marker targeting to specific regions of interest is possible using the syntenic
325 framework. This is relevant in situations where the target interval is large or initial candidate
326 functional information is absent and is illustrated in **Figure 2** by the case of the *VC* locus. The
327 published interval in which this gene maps is collinear with part of Mt2 (Khazaei et al., 2015) and as
328 discussed earlier, SNPs discovered in sequences orthologous to Medicago genes in the syntenic
329 interval can be used selectively and cost-effectively in further high-resolution mapping of the locus.
330 Thirdly, this syntenic framework can in principle be used to reverse map genes in *Vf*. Here, we use the
331 example of the *Vf* *TERMINAL FLOWER 1* (*VtTFL1*) gene. Avila et al., (2007) examined the
332 translational hypothesis that an orthologue of the *TFL1* gene controls determinacy of flowering in *Vf*
333 as it does in Arabidopsis, soybean and numerous other legume and non-legume species by showing
334 correlation of the determinate type with a diagnostic non-synonymous substitution in a conserved
335 residue of the coding region across a diverse panel of determinate and indeterminate types. They did
336 not, however, genetically map *VtTFL1*. Once again, clear macro-colinearity between the *Vf* region
337 corresponding to the *Mt* region harbouring *MtTFL1* suggests a working hypothesis (which remains to
338 be proven) that the published *VtTFL1* sequence should map to the long arm of *Vf* chromosome 1.

339 **Genomic Diversity**

340 Another arena in which genomics is playing a crucial role is in the description and rational
341 exploitation of diverse, but often poorly characterised, genetic resources. Despite the lack of a known
342 wild progenitor species, the domesticated *Vf* genepool, as alluded to in the introduction, is extremely
343 diverse. This diversity is manifest in the plethora of morphological forms and adaptations to diverse
344 agroecological settings. However, full exploitation of this diversity for faba bean improvement
345 necessitates efficient methods for quantifying and mapping genomic diversity. In parallel with the
346 development of SSAP, AFLP, SSR and SNP marker types, a succession of studies have used
347 increasingly powerful molecular marker sets to quantify and map genomic diversity.

348 One of the first surveys of genomic diversity in diverse *Vf* germplasm was carried out by Sanz et al.,
349 (2007) using retrotransposon-based Sequence-Specific Amplified Polymorphism (SSAP) markers.
350 This study found relatively little fine genetic structure and the first molecular evidence that ‘*major*’,
351 ‘*minor*’ and ‘*equina*’ lines do not form distinct clades, but rather are completely dispersed across the
352 *Vf* phylogeny. Due to the low numbers of *Vf* lines (n=20), it was not possible to draw any conclusions

353 on geographic partitioning of diversity, though it was noted that samples from the better-represented
354 countries in the study were well dispersed across the phylogenetic tree. These themes were
355 recapitulated by Zeid et al., (2003) using AFLP markers to genotype 79 inbred lines from Europe,
356 North Africa and Asia. Again, ‘major’ and ‘minor’ botanical types were clearly shown not to be
357 genetically distinct, though here for the first time was a suggestion that the (n=8) lines of Asian origin
358 formed a genetically differentiated group. The AFLP-based study of Zong et al., (2009) took this
359 further with a study of large numbers of Chinese landraces (n=204) from the Chinese Academy of
360 Agricultural Sciences (CAAS) germplasm collection, mainly winter types. They found the mainland
361 Chinese winter germplasm to be completely distinct from other Asian, African and European diverse
362 lines included for comparison, in line with the long history of cultivation of winter-type faba bean in
363 relatively isolated mountainous regions of China. Kwon et al. (2010), using Targeted Region
364 Amplified Polymorphism (TRAP) markers, also observed Chinese landraces drawn from the United
365 States Department of Agriculture (USDA) germplasm collection (n=107) to cluster completely
366 separately from 30 comparator lines from Asia and Europe. However, the largest study of faba bean
367 genomic diversity to date has been reported by Wang et al., (2012a), whose large sample (n=802)
368 confirmed not only the genetic distinctness of Chinese germplasm from African, European and other
369 Asian germplasm, but for the first time, convincingly showed differentiation amongst Chinese
370 provinces and between winter and spring ecotypes. SNP genotyping platforms have been more
371 recently used on small samples of diversity (Cottage et al., 2012; Kaur et al., 2014) and the prospect
372 of wider deployment of SNP markers on larger panels of diverse material promises the possible
373 foundations for future *Vf* genome-wide association studies.

374 **Functional genomics**

375 **Mutagenesis**

376 The potential of mutagenesis as a tool for breeding was amply demonstrated by Sjödin, (1971),
377 reviewing not only observations made in the extensive mutagenesis programme run by Swedish
378 breeders Svalov-Weibull in the ‘Primus’ genetic background, but also a series of spontaneous mutants
379 reported by a host of previous researchers. The only other mutagenesis programme reported in the
380 literature was 23 years later, when the isolation of five nodulation mutants in a screen of 20,000 M₂
381 EMS-mutagenised lines in the cv. ‘Ascott’ background was reported by Duc, (1995).

382 Interest in mutagenesis has undergone something of a resurgence in recent years. The summary of
383 ongoing public sector programmes provided in **Table 4**, collated from a series of personal
384 communications, shows a number of target traits that correspond to specific mutations not likely to be
385 found in nature. For example, the control of the devastating parasitic weed, *Orobanche crenata*,
386 referred to in an earlier section, could be effectively controlled, together with a range of other
387 troublesome broad-leaved weeds, if faba bean varieties with target-site mutations that render them
388 insensitive to particular actives were developed Gressel, (2009). Examples are the Ser653 and Ala205
389 mutations in the AcetoLactate Synthase target of imidazolinone and amidosulfuron families of
390 herbicide actives, which do not occur in nature, but have been documented to occur at low frequency
391 under strong selection pressure in the field, and there are examples in numerous crop species of
392 targeted isolation of induced mutants (reviewed in Tan et al., 2005). At the time of writing, a number
393 of imazapyr resistance mutations have been identified by Mao et al., (2014) and are undergoing
394 further characterization. As a further example of the ongoing mutagenesis programmes listed in Table
395 4, **Figure 3** shows a selection of phenotypes observed in an M₂ X-ray and EMS populations grown
396 during 2015 at the University of Reading.

397 Whilst the primary driver for contemporary mutagenesis programmes remains the identification of
398 novel induced variants for direct incorporation in active breeding programmes, a broader motivation
399 can be ascribed to some of the recent activity. Until efficient genetic transformation systems have
400 been adopted and made available as a service, reverse genetics in the form of TILLing (Till et al.,
401 2003) may be one of the most amenable techniques for validation of candidate genes identified via
402 synteny-based approaches, especially if combined with powerful re-sequencing techniques to identify
403 the causative mutations in a single step (Wang et al., 2012b).

404 **Genetic transformation/genome editing**

405 Genetic modification represents both a research tool, permitting testing of hypotheses on gene
406 function by over-, mis-expression or knockdown/knockout studies and an outlet for genetic research
407 in generation of targeted phenotypic modifications based on knowledge of gene function. Stable
408 germline transformation of *Vf* using *in vitro* regeneration of *Agrobacterium*-infiltrated (non-
409 meristematic) internode stem segments was first reported by Böttinger et al., (2001). Adopting a
410 somewhat different strategy, Hanafy et al., (2005) infiltrated excised (meristematic) embryo axes with
411 *Agrobacterium* and successfully recovered stable transgenic lines. Both methods, however, reported
412 low primary transformation efficiencies and relied on micro-grafting of putative transgenic shoot
413 material onto non-transgenic roots, a slow and highly manual process. Hanafy et al., (2013) later
414 reported abiotic stress resistance phenotypes of *Vf* transgenic lines overexpressing potato PR10a using
415 their previous methods. This remains to our knowledge the sole successful demonstration to date of
416 the feasibility of a biotechnological approach to *Vf* improvement. In the absence of a robust and
417 efficient transformation method, some attention has been devoted to the task of decreasing generation
418 time using tissue-culture based embryo rescue, with some success (Mobini et al., 2015). However, the
419 search for an efficient transformation method has been recently taken up by other groups (e.g.
420 Abdelwahd et al., (2014).

421 The prospects afforded by new insights into the phenotypic effects of allelic variation and the more
422 refined biotechnological possibilities afforded by rapidly maturing genome editing technologies (Gaj
423 et al., 2013) could potentially stimulate renewed interest in genetic transformation. An example of a
424 game-changing product which could readily be generated using even a medium efficiency
425 transformation system would be herbicide resistance obtained by directed mutagenesis of endogenous
426 herbicide target genes e.g. introduction of heterologous glyphosate resistance of bacterial origin.

427 **Conclusions**

428 *Vicia faba* genetics and genomics is now in a much healthier state than it was just a few years ago. We
429 can take heart from the accelerating progress in gene identification and in production of outputs
430 relevant to contemporary methods of breeding. We have seen the transformative effects of
431 transcriptome re-sequencing in opening the doors to high density gene-based marker discovery and
432 mapping. In the near future, we expect that a high throughput SNP chip incorporating many 10's of
433 thousands of genome-wide SNPs discovered from a wide variety of genetic backgrounds will be
434 produced and made available for community use, as well as giving rise to an ultra-high density SNP
435 map in which perhaps as much as half of all *Vf* genes are genetically mapped. In parallel, we should
436 see progress in the application of genotype-by-sequencing methods, both for the unbiased assessment
437 of genetic relationships, breeding applications such as genomic selection and in trait mapping
438 approaches based on bulked segregant analysis. Faba bean's modest but well-linked community of
439 researchers would do well however, to set its sights on a series of even more ambitious targets to
440 enable the community as a whole to elevate its work onto a higher plane of achievement and impact.

441 For example, given the many highly cost effective sequencing and assembly technologies now
442 available, a comprehensive genomic scaffold and haplotype map is surely now within reach. The
443 history of modern agricultural genomics shows the transformative effects of a well-annotated
444 reference genome. Communities of crop researchers who have organised themselves and published
445 strategic roadmaps requiring centralised investment in a professional genome assembly and
446 annotation have captured significant R&D investment and transformed the profile and fate of their
447 communities (rice, wheat, soybean, *Phaseolus*, cowpea). Similarly, on functional genomics platforms:
448 whilst there are no doubt training benefits to having numerous small mutagenesis programmes dotted
449 around the globe, it could be argued that the global community needs one well-funded programme
450 scoped to guarantee saturation mutagenesis, to catalogue the mutations obtained in a public sequence
451 database and to distribute mutant seed to research groups freely on request, an unlikely outcome from
452 nationally funded programmes. Likewise, transformation and genome editing technologies could
453 advantageously be developed and provided as an efficient service from a centralized laboratory.

454 **Tables:**

455

456 **Table 1:** Key *Vf* transcriptome datasets

Bioproject	Genbank reference	Cultivar	Tissue	Reference
PRJNA225873	SRP033593	BPL10	10-d seedling	Webb et al., 2016
PRJNA225881	SRP033121	Albus	10-d seedling	Webb et al., 2016
PRJNA238140	SRX476199	CDC Fatima	6-d root	Ray et al., 2015
	SRX476200	CDC Fatima	6-d shoot	Ray et al., 2015
	SRX476493	CDC Fatima	seed coat	Ray et al., 2015
	SRX476217	SSNS-1	6-d root	Ray et al., 2015
	SRX476220	SSNS-1	6-d shoot	Ray et al., 2015
	SRX475907	A01155	6-d root	Ray et al., 2015
	SRX475873	A01155	6-d shoot	Ray et al., 2015
PRJNA277609	SRP055969	Fiord	Mixed tissues	Arun-Chinnappa and McCurdy, 2015
PRJEB8906	ERP009949	Fiord	Cotyledon epidermis and parenchyma	Zhang et al., 2015
NA	JR964201- JR970413*	Icarus, Ascot	Mixed tissues	Kaur et al., 2012
PRJNA253768	SRP043650	NS	leaves	Suresh et al., 2015
NA	GI:219212932 - GI:219282595	Windsor	2 week old embryo	Ray and Georges, 2010
NA	SRP045955	INRA-29H	leaf	Ocaña et al., 2015
NA	SRP045955	<i>Vf</i> 136	leaf	Ocaña et al., 2015

457

458 NA: Not applicable; NS: Not stated

459 *Only assembled contigs available as TSA

460 **Table 2:** Key molecular marker sets

461

Marker Type	Discovery genotypes	Number of validated polymorphic markers	Number mapped	Reference
EST-SSR	Komasake	647	552	El-Rodeny et al., 2014
genomic SSR	Mixed	90	NA	Yang et al., 2012
EST-SSR	Icarus, Ascot	71	57	Kaur et al., 2014
SNP (KASP)	Albus, BPL10	824	687	Cottage et al., 2012, Webb et al., 2016
SNP (Illumina)	Icarus, Ascot	480	465	Kaur et al., 2014
ITAP	<i>Vf6, Vf27</i>	151	127	Ellwood et al., 2008

462

463

464

465 **Table 3:** Genetic maps

Cross	Number of loci	Number of linkage groups	Average marker interval	Map length (cM)	Reference
Consensus of six F ₂ populations	687	6	2.04 cM	1403.8	Webb et al., 2016
'Nubaria 2' x 'Misr 3' F ₂	552 EST-SSR	6	1.25 cM	687.7	El-Rodeny et al., 2014
Consensus of three RIL populations	729 RAPD, ITAP, SSR, morphological	6	6.31 cM	4602	Satovic et al., 2013
91825 x K1563 F ₂	128 SSR	15	12.4 cM	1587	Ma et al, 2013 ref
Icarus x Ascot F5:6 RILs	522 (57 EST-SSR, 465 SNP)	12	2.33 cM	1216.8	Kaur et al., 2014
Vf6 x Vf27 RILs	127	12	13.27	1685.8	Cottage et al., 2012; Ellwood et al., 2008

466

467

468 **Table 4:** Current/ongoing mutagenesis programmes

Genetic background	Institution/PI	Mutagen	Primary specific targets	No. independent lines in most advanced generation
Hedin/2	Uni Reading/DM O'Sullivan	EMS	Imidizolanone resistance	1,097 M ₂
Hedin/2	Uni Reading/DM O'Sullivan	X-ray	Imidizolanone resistance	985 M ₂
SSNS-1	Uni Saskatchewan/B Vandenberg	NaN ₃	Low phytate	1,500 M ₄
SSNS-1	Uni Saskatchewan/B Vandenberg	EMS	Low phytate	1,500 M ₄
BPL710	ICARDA/F Maalouf	EMS	Glyphosate tolerance	400 M ₆
NA112	ICARDA/F Maalouf	EMS	Herbicide resistance	5,000 M ₂
Nura	Uni Adelaide/J Paull	EMS	Imazapyr resistance	Four imazapyr resistant lines under investigation

469 **Figure Legends:**

470

471 **Figure 1:** Integrative Genomics Viewer (IGV) snapshot of SNPs in the *Vf* orthologue of
472 Medtr2g009080 mined from public transcriptome datasets. The eleven horizontal tracks shown in this
473 view from top to bottom are: the meta-transcriptome contig TR55082:c0_g4_i2 is shown as a
474 schematic, all intravarietal SNP positions; positions where the minor (non-consensus) allele is found
475 in each of the following genotypes: ‘Albus’, ‘Fiord’, ‘SRX641218’, ‘SSN1’, ‘Fatima’, A01155’ and
476 ‘BPL10’, alignment of the reference *Vf* contig TR55082:c0_g4_i2 to the Mt4.0 genome (coding
477 region of Medtr2g009080 in this instance), and in the bottom panel is shown a 3-frame translation of
478 the *Vf* sequence. In this example, there are a total of 16 intravarietal SNPs distributed across the gene -
479 eleven unique to ‘Fiord’, two unique to ‘A001155’ and three minor alleles in common to ‘Fiord’ and
480 ‘A01155’, with the remaining genotypes carrying the unchanged common allele identical to the
481 consensus contig sequence.

482 **Figure 2:** Wheel representation of synteny between pseudomolecule sequence of *Medicago truncatula*
483 chromosomes Mt1-Mt8 and the consensus *Vf* linkage map (*Vf*1-*Vf*6). Syntenic relationships were as
484 described in Webb et al., 2016 and re-drawn for this figure using ‘Circos’ software (Krzywinski et al.,
485 2009). Link lines to *Medicago* orthologues for markers on each *Vf* chromosome are shown in a
486 different colour. The virtually uninterrupted collinearity between the whole of *Vf*3 and the whole of
487 Mt1 is highlighted by representing these orthologous chromosomes in green. To illustrate how useful
488 this syntenic framework is for translational genomics as well as for drawing together knowledge of
489 the genome, three exemplar loci where there is a proven or presumed orthologous relationship are
490 shown as red ticks on the orthologous chromosomes and linked by black lines. These are: the mapped
491 *ZTI* locus and its *Medicago* orthologue *MtWD40-1*, the mapped genetic interval corresponding to the
492 mapped *VC* (vicine-convicine) locus and its corresponding *Medicago* syntenic interval on *Mt*2, and
493 the actual position of the *Medicago TERMINAL FLOWER LOCUS 1 (TFL1)* with a link drawn to the
494 region of *Vf*1 where it is presumed the *Vf* orthologue (*VfTFL1*) is likely to be located.

495 **Figure 3:** Mutant phenotypes observed in X-ray and EMS M₂ lines grown at the University of
496 Reading in 2015. A. disease lesion mimic B. determinate flowering B. curled wing petal C. disease
497 lesion mimic D. reduced flower pigmentation.

498

499 **References**

500

- 501 Abdelwahd, R., Udupa, S.M., Gaboun, F., Diria, G., Mentag, R., Ibriz, Mohamed and Iraqi, D. (2014)
502 Agrobacterium-mediated transformation of cotyledonary node of *Vicia faba* L. *Rom Agric*
503 *Res* **31**, online first.
- 504 Arun-Chinnappa, K.S. and McCurdy, D.W. (2015) De novo assembly of a genome-wide
505 transcriptome map of *Vicia faba* (L.) for transfer cell research. *Front Plant Sci* **6**.
- 506 Avila, C.M., Atienza, S.G., Moreno, M.T. and Torres, A.M. (2007) Development of a new diagnostic
507 marker for growth habit selection in faba bean (*Vicia faba* L.) breeding. *Theor Appl Genet*
508 **115**, 1075-1082.
- 509 Avila, C.M., Satovic, Z., Sillero, J.C., Rubiales, D., Moreno, M.T. and Torres, A.M. (2004) Isolate
510 and organ-specific QTLs for ascochyta blight resistance in faba bean (*Vicia faba* L.). *Theor*
511 *Appl Genet* **108**, 1071-1078.
- 512 Avila, C.M., Sillero, J.C., Rubiales, D., Moreno, M.T. and Torres, A.M. (2003) Identification of
513 RAPD markers linked to the Uvf-1 gene conferring hypersensitive resistance against rust
514 (*Uromyces viciae-fabae*) in *Vicia faba* L. *Theor Appl Genet* **107**, 353-358.
- 515 Bishop, J., Jones, H.E., Lukac, M. and Potts, S.G. (2016a) Insect pollination reduces yield loss
516 following heat stress in faba bean (*Vicia faba* L.). *Agriculture, Ecosystems & Environment*
517 **220**, 89-96.
- 518 Bishop, J., Potts, S.G. and Jones, H.E. (2016b) Susceptibility of Faba Bean (*Vicia faba* L.) to Heat
519 Stress During Floral Development and Anthesis. *J Agron Crop Sci*, n/a-n/a.
- 520 Böttinger, P., Steinmetz, A., Schieder, O. and Pickardt, T. (2001) Agrobacterium-mediated
521 transformation of *Vicia faba*. *Mol Breeding* **8**, 243-254.
- 522 Cernay, C., Ben-Ari, T., Pelzer, E., Meynard, J.-M. and Makowski, D. (2015) Estimating variability
523 in grain legume yields across Europe and the Americas. *Sci Rep-Uk* **5**, 11171.
- 524 Cottage, A., Gostkiewicz, K., Thomas, J.E., Borrows, R., Torres, A.M. and O'Sullivan, D.M. (2012)
525 Heterozygosity and diversity analysis using mapped single nucleotide polymorphisms in a
526 faba bean inbreeding programme. *Mol Breeding* **30**, 1799-1809.
- 527 Crepon, K., Marget, P., Peyronnet, C., Carrouee, B., Arese, P. and Duc, G. (2010) Nutritional value of
528 faba bean (*Vicia faba* L.) seeds for feed and food. *Field Crop Res* **115**, 329-339.
- 529 Cruz-Izquierdo, S., Avila, C.M., Satovic, Z., Palomino, C., Gutierrez, N., Ellwood, S.R., Phan,
530 H.T.T., Cubero, J.I. and Torres, A.M. (2012) Comparative genomics to bridge *Vicia faba* with
531 model and closely-related legume species: stability of QTLs for flowering and yield-related
532 traits. *Theor Appl Genet* **125**, 1767-1782.
- 533 Denton, M.D., Pearce, D.J. and Peoples, M.B. (2013) Nitrogen contributions from faba bean (*Vicia*
534 *faba* L.) reliant on soil rhizobia or inoculation. *Plant Soil* **365**, 363-374.
- 535 Diaz-Ruiz, R., Satovic, Z., Avila, C.M., Alfaro, C.M., Gutierrez, M.V., Torres, A.M. and Roman, B.
536 (2009) Confirmation of QTLs controlling *Ascochyta fabae* resistance in different generations
537 of faba bean (*Vicia faba* L.). *Crop Pasture Sci* **60**, 353-361.
- 538 Diaz-Ruiz, R., Torres, A.M., Satovic, Z., Gutierrez, M.V., Cubero, J.I. and Roman, B. (2010)
539 Validation of QTLs for *Orobanche crenata* resistance in faba bean (*Vicia faba* L.) across
540 environments and generations. *Theor Appl Genet* **120**, 909-919.
- 541 Duc, G. (1995) Mutagenesis of Faba Bean (*Vicia-Faba* L) and the Identification of 5 Different Genes-
542 Controlling No Nodulation, Ineffective Nodulation or Supernodulation. *Euphytica* **83**, 147-
543 152.
- 544 Duc, G., Bao, S.Y., Baum, M., Redden, B., Sadiki, M., Suso, M.J., Vishniakova, M. and Zong, X.X.
545 (2010) Diversity maintenance and use of *Vicia faba* L. genetic resources. *Field Crop Res* **115**,
546 270-278.
- 547 Duc, G., Moussy, F., Zong, X. and Ding, G. (1999) Single gene mutation for green cotyledons as a
548 marker for the embryonic genotype in faba bean, *Vicia faba*. *Plant Breeding* **118**, 577-578.
- 549 Duc, G. and Picard, J. (1986) Note on the Presence of the Sym-1 Gene in *Vicia-Faba* Hampering Its
550 Symbiosis with *Rhizobium-Leguminosarum*. *Euphytica* **35**, 61-64.

- 551 El-Rodeny, W., Kimura, M., Hirakawa, H., Sabah, A., Shirasawa, K., Sato, S., Tabata, S., Sasamoto,
552 S., Watanabe, A., Kawashima, K., Kato, M., Wada, T., Tsuruoka, H., Takahashi, C., Minami,
553 C., Nanri, K., Nakayama, S., Kohara, M., Yamada, M., Kishida, Y., Fujishiro, T. and Isobe,
554 S. (2014) Development of EST-SSR markers and construction of a linkage map in faba bean
555 (*Vicia faba*). *Breeding Sci* **64**, 252-263.
- 556 Ellwood, S.R., Phan, H.T.T., Jordan, M., Hane, J., Torres, A.M., Avila, C.M., Cruz-Izquierdo, S. and
557 Oliver, R.P. (2008) Construction of a comparative genetic map in faba bean (*Vicia faba* L.);
558 conservation of genome structure with *Lens culinaris*. *Bmc Genomics* **9**.
- 559 Erith, A.G. (1930) The inheritance of colour, size and form of seeds, and of flower colour in *Vicia*
560 *Faba* L. *Genetica* **12**, 477-510.
- 561 Gaj, T., Gersbach, C.A. and Barbas Iii, C.F. (2013) ZFN, TALEN, and CRISPR/Cas-based methods
562 for genome engineering. *Trends in Biotechnology* **31**, 397-405.
- 563 Gnanasambandam, A., Paull, J., Torres, A., Kaur, S., Leonforte, T., Li, H., Zong, X., Yang, T. and
564 Materne, M. (2012) Impact of Molecular Technologies on Faba Bean (*Vicia faba* L.)
565 Breeding Strategies. *Agronomy* **2**, 132-166.
- 566 Gong, Y.M., Xu, S.C., Mao, W.H., Li, Z.Y., Hu, Q.Z., Zhang, G.W. and Ding, J. (2011) Genetic
567 Diversity Analysis of Faba Bean (*Vicia faba* L.) Based on EST-SSR Markers. *Agr Sci China*
568 **10**, 838-844.
- 569 Gressel, J. (2009) Crops with target-site herbicide resistance for Orobanche and Striga control. *Pest*
570 *Manag Sci* **65**, 560-565.
- 571 Gutierrez, N., Palomino, C., Satovic, Z., Ruiz-Rodriguez, M.D., Vitale, S., Gutierrez, M.V., Rubiales,
572 D., Kharrat, M., Amri, M., Emeran, A.A., Cubero, J.I., Atienza, S.G., Torres, A.M. and Avila,
573 C.M. (2013) QTLs for Orobanche spp. resistance in faba bean: identification and validation
574 across different environments. *Mol Breeding* **32**, 909-922.
- 575 Hanafy, M., Pickardt, T., Kiesecker, H. and H.-J., J. (2005) Agrobacterium-mediated transformation
576 of faba bean (*Vicia faba* L.) using embryo axes. *Euphytica* **142**, 227-236.
- 577 Hanafy, M.S., El-Banna, A., Schumacher, H.M., Jacobsen, H.J. and Hassan, F.S. (2013) Enhanced
578 tolerance to drought and salt stresses in transgenic faba bean (*Vicia faba* L.) plants by
579 heterologous expression of the PR10a gene from potato. *Plant Cell Rep* **32**, 663-674.
- 580 Kaur, S., Cogan, N.O.I., Forster, J.W. and Paull, J.G. (2014) Assessment of Genetic Diversity in
581 Faba Bean Based on Single Nucleotide Polymorphism. *Diversity* **6**, 88-101.
- 582 Kaur, S., Kimber, R.B.E., Cogan, N.O.I., Materne, M., Forster, J.W. and Paull, J.G. (2014) SNP
583 discovery and high-density genetic mapping in faba bean (*Vicia faba* L.) permits
584 identification of QTLs for ascochyta blight resistance. *Plant Sci* **217**, 47-55.
- 585 Kaur, S., Pembleton, L.W., Cogan, N.O.I., Savin, K.W., Leonforte, T., Paull, J., Materne, M. and
586 Forster, J.W. (2012) Transcriptome sequencing of field pea and faba bean for discovery and
587 validation of SSR genetic markers. *Bmc Genomics* **13**.
- 588 Khamassi, K., Ben Jeddi, F., Hobbs, D., Irigoyen, J., Stoddard, F., O'Sullivan, D.M. and Jones, H.
589 (2013) A baseline study of vicine-convicine levels in faba bean (*Vicia faba* L.) germplasm.
590 *Plant Genet Resour-C* **11**, 250-257.
- 591 Khan, H.R., Paull, J.G., Siddique, K.H.M. and Stoddard, F.L. (2010) Faba bean breeding for drought-
592 affected environments: A physiological and agronomic perspective. *Field Crop Res* **115**, 279-
593 286.
- 594 Khazaei, H., O'Sullivan, D.M., Jones, H., Pitts, N., Sillanpaa, M.J., Parssinen, P., Manninen, O. and
595 Stoddard, F.L. (2015) Flanking SNP markers for vicine-convicine concentration in faba bean
596 (*Vicia faba* L.). *Mol Breeding* **35**.
- 597 Khazaei, H., O'Sullivan, D.M., Sillanpaa, M.J. and Stoddard, F.L. (2014) Use of synteny to identify
598 candidate genes underlying QTL controlling stomatal traits in faba bean (*Vicia faba* L.).
599 *Theor Appl Genet* **127**, 2371-2385.
- 600 Kovarova, P., Navratilova, A., Macas, J. and Dolezel, J. (2007) Chromosome analysis and sorting in
601 *Vicia sativa* using flow cytometry. *Biol Plantarum* **51**, 43-48.
- 602 Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S.J. and Marra,
603 M.A. (2009) Circos: An information aesthetic for comparative genomics. *Genome research*
604 **19**, 1639-1645.

- 605 Link, W., Balko, C. and Stoddard, F.L. (2010) Winter hardiness in faba bean: Physiology and
606 breeding. *Field Crop Res* **115**, 287-296.
- 607 Ma, Y., Bao, S.Y., Yang, T., Hu, J.G., Guan, J.P., He, Y.H., Wang, X.J., Wan, Y.L., Sun, X.L., Jiang,
608 J.Y., Gong, C.X. and Zong, X.X. (2013) Genetic linkage map of Chinese native variety faba
609 bean (*Vicia faba* L.) based on simple sequence repeat markers. *Plant Breeding* **132**, 397-400.
- 610 Ma, Y., Yang, T., Guan, J.P., Wang, S.M., Wang, H.F., Sun, X.L. and Zong, X.X. (2011)
611 Development and Characterization of 21 Est-Derived Microsatellite Markers in *Vicia Faba*
612 (*Fava Bean*). *Am J Bot* **98**, E22-E24.
- 613 Mao, D., Paull, J., Oldach, K.H., Preston, C., Yang, S.Y., Davies, P. and McMurray, L. (2014) The
614 development of multiple herbicide tolerances in faba bean (*Vicia faba* L.) through induced
615 mutation. In: *15th Australasian Plant Breeding Conference*. Melbourne, Australia.
- 616 Maxted, N. (1993) Towards a faba bean progenitor! *FABIS Newsletter* **31**.
- 617 Menke, M., Meister, A. and Schubert, I. (2000) N-methyl-N-nitrosourea-induced DNA damage
618 detected by the comet assay in *Vicia faba* nuclei during all interphase stages is not restricted
619 to chromatid aberration hot spots. *Mutagenesis* **15**, 503-506.
- 620 Milan, P.R. and Upadhyay, S. (2007) Impact of food additives on mitotic chromosomes of *Vicia faba*
621 L. *Caryologia* **60**, 309-314.
- 622 Mobini, S.H., Lulsdorf, M., Warkentin, T.D. and Vandenberg, A. (2015) Plant growth regulators
623 improve in vitro flowering and rapid generation advancement in lentil and faba bean. *In Vitro*
624 *CellDevBiol-Plant* **51**, 71-79.
- 625 Multari, S., Stewart, D. and Russell, W.R. (2015) Potential of Fava Bean as Future Protein Supply to
626 Partially Replace Meat Intake in the Human Diet. *Comprehensive Reviews in Food Science*
627 *and Food Safety*, n/a-n/a.
- 628 Nayak, G.K., Roberts, S.P.M., Garratt, M., Breeze, T.D., Tscheulin, T., Harrison-Cripps, J.,
629 Vogiatzakis, I.N., Stirpe, M.T. and Potts, S.G. (2015) Interactive effect of floral abundance
630 and semi-natural habitats on pollinators in field beans (*Vicia faba*). *Agr Ecosyst Environ* **199**,
631 58-66.
- 632 Ocaña, S., Seoane, P., Bautista, R., Palomino, C., Claros, G.M., Torres, A.M. and Madrid, E. (2015)
633 Large-Scale Transcriptome Analysis in Faba Bean (*Vicia faba* L.) under *Ascochyta fabae*
634 Infection. *Plos One* **10**, e0135143.
- 635 Patrick, J.W. and Stoddard, F.L. (2010) Physiology of flowering and grain filling in faba bean. *Field*
636 *Crop Res* **115**, 234-242.
- 637 Patto, M.C.V., Torres, A.M., Koblickova, A., Macas, J. and Cubero, J.I. (1999) Development of a
638 genetic composite map of *Vicia faba* using F-2 populations derived from trisomic plants.
639 *Theor Appl Genet* **98**, 736-743.
- 640 Pérez-de-Luque, A., Eizenberg, H., Grenz, J.H., Sillero, J.C., Ávila, C., Sauerborn, J. and Rubiales, D.
641 (2010) Broomrape management in faba bean. *Field Crop Res* **115**, 319-328.
- 642 Preissel, S., Reckling, M., Schläfke, N. and Zander, P. (2015) Magnitude and farm-economic value of
643 grain legume pre-crop benefits in Europe: A review. *Field Crop Res* **175**, 64-79.
- 644 Ramsay, G. (1997) Inheritance and linkage of a gene for testa-imposed seed dormancy in faba bean
645 (*Vicia faba* L.). *Plant Breeding* **116**, 287-289.
- 646 Ray, H., Bock, C. and Georges, F. (2015) Faba Bean: Transcriptome Analysis from Etiolated Seedling
647 and Developing Seed Coat of Key Cultivars for Synthesis of Proanthocyanidins, Phytate,
648 Raffinose Family Oligosaccharides, Vicine, and Convicine. *Plant Genome-Us* **8**.
- 649 Ray, H. and Georges, F. (2010) A genomic approach to nutritional, pharmacological and genetic
650 issues of faba bean (*Vicia faba*): prospects for genetic modifications. *GM crops* **1**, 99-106.
- 651 Rybaczek, D., Zabka, A., Pastucha, A. and Maszewski, J. (2008) Various chemical agents can induce
652 premature chromosome condensation in *Vicia faba*. *Acta Physiol Plant* **30**, 663-672.
- 653 Sallam, A., Martsch, R. and Moursi, Y. (2015) Genetic variation in morpho-physiological traits
654 associated with frost tolerance in faba bean (*Vicia faba* L.). *Euphytica* **205**, 395-408.
- 655 Sanz, A.M., Gonzalez, S.G., Syed, N.H., Suso, M.J., Saldana, C.C. and Flavell, A.J. (2007) Genetic
656 diversity analysis in *Vicia* species using retrotransposon-based SSAP markers. *Mol Genet*
657 *Genomics* **278**, 433-441.
- 658 Satovic, Z., Avila, C.M., Cruz-Izquierdo, S., Diaz-Ruiz, R., Garcia-Ruiz, G.M., Palomino, C.,
659 Gutierrez, N., Vitale, S., Ocana-Moral, S., Gutierrez, M.V., Cubero, J.I. and Torres, A.M.

- 660 (2013) A reference consensus genetic map for molecular markers and economically important
661 traits in faba bean (*Vicia faba* L.). *Bmc Genomics* **14**.
- 662 Schmutz, J., Cannon, S.B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., Hyten, D.L., Song, Q.,
663 Thelen, J.J., Cheng, J., Xu, D., Hellsten, U., May, G.D., Yu, Y., Sakurai, T., Umezawa, T.,
664 Bhattacharyya, M.K., Sandhu, D., Valliyodan, B., Lindquist, E., Peto, M., Grant, D., Shu, S.,
665 Goodstein, D., Barry, K., Futrell-Griggs, M., Abernathy, B., Du, J., Tian, Z., Zhu, L., Gill, N.,
666 Joshi, T., Libault, M., Sethuraman, A., Zhang, X.C., Shinozaki, K., Nguyen, H.T., Wing,
667 R.A., Cregan, P., Specht, J., Grimwood, J., Rokhsar, D., Stacey, G., Shoemaker, R.C. and
668 Jackson, S.A. (2010) Genome sequence of the palaeopolyploid soybean. *Nature* **463**, 178-
669 183.
- 670 Sillero, J.C., Villegas-Fernández, A.M., Thomas, J., Rojas-Molina, M.M., Emeran, A.A., Fernández-
671 Aparicio, M. and Rubiales, D. (2010) Faba bean breeding for disease resistance. *Field Crop*
672 *Res* **115**, 297-307.
- 673 Sjödin, J.A.N. (1970) Induced asynaptic mutants in *Vicia faba* L. *Hereditas* **66**, 215-232.
- 674 Sjödin, J.A.N. (1971) Induced morphological variation in *Vicia faba* L. *Hereditas* **67**, 155-179.
- 675 Sobita, K. and Bhagirath, T. (2005) Effects of some medicinal plant extracts on *Vicia faba* root tip
676 chromosomes. *Caryologia* **58**, 255-261.
- 677 Soltis, D.E., Soltis, P.S., Bennett, M.D. and Leitch, I.J. (2003) Evolution of genome size in the
678 angiosperms. *Am J Bot* **90**, 1596-1603.
- 679 Stoddard, F.L., Nicholas, A.H., Rubiales, D., Thomas, J. and Villegas-Fernández, A.M. (2010)
680 Integrated pest management in faba bean. *Field Crop Res* **115**, 308-318.
- 681 Suresh, S., Kim, T.-S., Raveendar, S., Cho, J.-H., Yi, J.-Y., Lee, M.C., Lee, S.-Y., Baek, H.-J., Cho,
682 G.-T. and Chung, J.-W. (2015) Transcriptome characterization and large-scale identification
683 of SSR/SNP markers in symbiotic nitrogen fixation crop faba bean (*Vicia faba* L.). *Turk J*
684 *Agric For* **39**, 459-469.
- 685 Tan, S., Evans, R.R., Dahmer, M.L., Singh, B.K. and Shaner, D.L. (2005) Imidazolinone-tolerant
686 crops: history, current status and future. *Pest Manag Sci* **61**, 246-257.
- 687 Tanno, K. and Willcox, G. (2006) The origins of cultivation of *Cicer arietinum* L. and *Vicia faba* L.:
688 early finds from Tell el-Kerkh, north-west Syria, late 10th millennium BP. *Veg Hist*
689 *Archaeobot* **15**, 197-204.
- 690 Tavakkoli, E., Paull, J., Rengasamy, P. and McDonald, G.K. (2012) Comparing genotypic variation in
691 faba bean (*Vicia faba* L.) in response to salinity in hydroponic and field experiments. *Field*
692 *Crop Res* **127**, 99-108.
- 693 Till, B.J., Reynolds, S.H., Greene, E.A., Codomo, C.A., Enns, L.C., Johnson, J.E., Burtner, C.,
694 Odden, A.R., Young, K., Taylor, N.E., Henikoff, J.G., Comai, L. and Henikoff, S. (2003)
695 Large-scale discovery of induced point mutations with high-throughput TILLING. *Genome*
696 *research* **13**, 524-530.
- 697 Torres, A.M., Avila, C.M., Gutierrez, N., Palomino, C., Moreno, M.T. and Cubero, J.I. (2010)
698 Marker-assisted selection in faba bean (*Vicia faba* L.). *Field Crop Res* **115**, 243-252.
- 699 Torres, A.M., Weeden, N.F. and Martin, A. (1993) Linkage among isozyme, RFLP and RAPD
700 markers in *Vicia faba*. *Theor Appl Genet* **85**, 937-945.
- 701 van de Wouw, M., Enneking, D., Robertson, L.D. and Maxted, N. (2001) Vetches. In: *Plant Genetics*
702 *Resources of Legumes in the Mediterranean* pp. 134-158.
- 703 Wang, H.F., Zong, X.X., Guan, J.P., Yang, T., Sun, X.L., Ma, Y. and Redden, R. (2012a) Genetic
704 diversity and relationship of global faba bean (*Vicia faba* L.) germplasm revealed by ISSR
705 markers. *Theor Appl Genet* **124**, 789-797.
- 706 Wang, T.L., Uauy, C., Robson, F. and Till, B. (2012b) TILLING in extremis. *Plant Biotechnology*
707 *Journal* **10**, 761-772.
- 708 Webb, A., Cottage, A., Wood, T., Khamassi, K., Hobbs, D., Gostkiewicz, K., White, M., Khazaei, H.,
709 Ali, M., Street, D., Duc, G., Stoddard, F.L., Maalouf, F., Ogbonnaya, F.C., Link, W., Thomas,
710 J. and O'Sullivan, D.M. (2016) A SNP-based consensus genetic map for synteny-based trait
711 targeting in faba bean (*Vicia faba* L.). *Plant Biotechnol J* **14**, 177-185.
- 712 Yang, T., Bao, S.Y., Ford, R., Jia, T.J., Guan, J.P., He, Y.H., Sun, X.L., Jiang, J.Y., Hao, J.J., Zhang,
713 X.Y. and Zong, X.X. (2012) High-throughput novel microsatellite marker of faba bean via
714 next generation sequencing. *Bmc Genomics* **13**.

- 715 Young, N.D., Debelle, F., Oldroyd, G.E., Geurts, R., Cannon, S.B., Udvardi, M.K., Benedito, V.A.,
716 Mayer, K.F., Gouzy, J., Schoof, H., Van de Peer, Y., Proost, S., Cook, D.R., Meyers, B.C.,
717 Spannagl, M., Cheung, F., De Mita, S., Krishnakumar, V., Gundlach, H., Zhou, S., Mudge, J.,
718 Bharti, A.K., Murray, J.D., Naoumkina, M.A., Rosen, B., Silverstein, K.A., Tang, H.,
719 Rombauts, S., Zhao, P.X., Zhou, P., Barbe, V., Bardou, P., Bechner, M., Bellec, A., Berger,
720 A., Berges, H., Bidwell, S., Bisseling, T., Choisne, N., Couloux, A., Denny, R., Deshpande,
721 S., Dai, X., Doyle, J.J., Dudez, A.M., Farmer, A.D., Fouteau, S., Franken, C., Gibelin, C.,
722 Gish, J., Goldstein, S., Gonzalez, A.J., Green, P.J., Hallab, A., Hartog, M., Hua, A.,
723 Humphray, S.J., Jeong, D.H., Jing, Y., Jocker, A., Kenton, S.M., Kim, D.J., Klee, K., Lai, H.,
724 Lang, C., Lin, S., Macmil, S.L., Magdelenat, G., Matthews, L., McCorrison, J., Monaghan,
725 E.L., Mun, J.H., Najjar, F.Z., Nicholson, C., Noirot, C., O'Bleness, M., Paule, C.R., Poulain,
726 J., Prion, F., Qin, B., Qu, C., Retzel, E.F., Riddle, C., Sallet, E., Samain, S., Samson, N.,
727 Sanders, I., Saurat, O., Scarpelli, C., Schiex, T., Segurens, B., Severin, A.J., Sherrier, D.J.,
728 Shi, R., Sims, S., Singer, S.R., Sinharoy, S., Sterck, L., Viollet, A., Wang, B.B., Wang, K.,
729 Wang, M., Wang, X., Warfsmann, J., Weissenbach, J., White, D.D., White, J.D., Wiley, G.B.,
730 Wincker, P., Xing, Y., Yang, L., Yao, Z., Ying, F., Zhai, J., Zhou, L., Zuber, A., Denarie, J.,
731 Dixon, R.A., May, G.D., Schwartz, D.C., Rogers, J., Quetier, F., Town, C.D. and Roe, B.A.
732 (2011) The Medicago genome provides insight into the evolution of rhizobial symbioses.
733 *Nature* **480**, 520-524.
- 734 Zeid, M., Mitchell, S., Link, W., Carter, M., Nawar, A., Fulton, T. and Kresovich, S. (2009) Simple
735 sequence repeats (SSRs) in faba bean: new loci from Orobanche-resistant cultivar 'Giza 402'.
736 *Plant Breeding* **128**, 149-155.
- 737 Zeid, M., Schon, C.C. and Link, W. (2003) Genetic diversity in recent elite faba bean lines using
738 AFLP markers. *Theor Appl Genet* **107**, 1304-1314.
- 739 Zhang, H.M., Wheeler, S., Xia, X., Radchuk, R., Weber, H., Offler, C.E. and Patrick, J.W. (2015)
740 Differential transcriptional networks associated with key phases of ingrowth wall construction
741 in trans-differentiating epidermal transfer cells of *Vicia faba* cotyledons. *BMC plant biology*
742 **15**, 103.
- 743 Zong, X.X., Liu, X.J., Guan, J.P., Wang, S.M., Liu, Q.C., Paull, J.G. and Redden, R. (2009)
744 Molecular variation among Chinese and global winter faba bean germplasm. *Theor Appl*
745 *Genet* **118**, 971-978.
- 746
- 747