Plant Science 215–216 (2014) 180–189

Contents lists available at ScienceDirect

Plant Science

journal homepage: www.elsevier.com/locate/plantsci

# Molecular polymorphism related to flowering trait variation in a *Phaseolus vulgaris* L. collection

Lorenzo Raggi[a], Carlo Tissi[a], Andrea Mazzucato[b], Valeria Negri[a,*]

[a] Dipartimento di Biologia Applicata, Università degli Studi di Perugia, 06121 Perugia, Italy
[b] Dipartimento di Scienze e Tecnologie per l'Agricoltura, le Foreste, la Natura e l'Energia, Università degli Studi della Tuscia, 01100 Viterbo, Italy

ARTICLE INFO

*Article history:*
Received 5 September 2013
Received in revised form 31 October 2013
Accepted 5 November 2013
Available online 12 November 2013

*Keywords:*
*Phaseolus vulgaris* L.
Flowering time
Marker-trait association
Candidate genes
*VRN1* gene

ABSTRACT

The aim of this study was to investigate the flowering variation and the molecular polymorphism in key regulatory genes that control flowering in a *Phaseolus vulgaris* L. collection of 94 accessions from Europe and the Americas. The analysis of variance revealed that the difference in days-to-flowering between accessions was significant, with European accessions characterized by flowering precocity. Population structure analysis corroborated previous data on the genetic distinction between the Andean and Mesoamerican gene pools. A low level of admixture was detected. Genomic sequences of 15 gene fragments were obtained. About 7.0 kb per accession were sequenced and a total of 48 nucleotide substitutions identified. A Mixed Linear Model analysis, including population structure and kinship, was used to identify marker-trait associations. Haplotype tagging single nucleotide polymorphisms (htSNPs) associated with the studied traits were detected: in *PvVRN1* and *PvPHYB* with days-to-flowering, in *PvMYB29* with number of flower buds per inflorescence and in *PvTFL1z* and *PvFCA* with inflorescence length. The two genes associated with days-to-flowering control belong to the photoperiod and vernalization pathways. In particular, the *PvVRN1* gene appears to play an important role in regulating the adaptation process of common bean.



## 1. Introduction

The timing of the transition to flowering is critical to reproductive success and crop production in plants. In recent years, considerable progress has been made in the molecular dissection of the flowering pathway, especially in *Arabidopsis* [1–4] and rice [5–7] where many of the key regulatory genes have been identified and functionally characterized. Mutant analyses, genetic and molecular interaction studies have allowed many of the genes involved in flowering control to be assigned to distinct regulatory inductive pathways: the so-called vernalization, photoperiod, gibberellin and autonomous pathways. Since genetic pathways controlling plant reproduction are well-conserved in higher plants, the study of reproductive mechanisms in non-model species has been facilitated [8,9] and a growing number of related genes have been identified in crops. However, few studies have considered the diversity of flowering time genes in legumes and its relationship to adaptation traits when plants were introduced into new environments [10,11].

Common bean (*Phaseolus vulgaris* L.) originated in the tropical and sub-tropical regions of the American continent. The wild forms of this species, which are distributed over an area extending from northern Mexico (approx. 30° N) to northwestern Argentina (approx. 35° S), at altitudes ranging from 500 to 2000 m asl., typically show photoperiod-sensitivity for flowering. Wild common bean only flowers under short-day photoperiods (day length less than 10–12 h) which allows seed production to be completed before the onset of the dry season. Common bean germplasm underwent a strong selection pressure for photo-insensitivity when introduced into Europe, or other northern latitudes, where the warm growing season coincides with long-day photoperiods [12]. Gniffke [13] studied photoperiod sensitivity in some common bean cultivars, in relation to their origin, and demonstrated that those from temperate regions had a higher frequency of photoperiod-insensitive genotypes compared with those collected in Latin America. In addition, studies conducted at the International Center for Tropical Agriculture (CIAT) demonstrated that genotypes with a type II growth habit (*i.e.* indeterminate bush with erect branches) are mainly daylength-neutral, while those that have a type IV growth habit (*i.e.* indeterminate with semi-climbing main stem and branches) are highly daylength-sensitive [14].

*Abbreviations:* AM, Association Mapping; BLASTn, Basic local alignment search tool nucleotide; dd, days; *dof*, duration of flowering; *dtf*, days-to-flowering; *fbi*, flower buds per inflorescence; *fbs*, flower bud size; FDR, false discovery rate; *gwh*, growth habit; htSNP, haplotype tagging SNP; *inl*, inflorescence length; *lat*, latitude of origin; MLM, Mixed Linear Model; *nms*, nodes on the main stem.

* Corresponding author. Tel.: +39 075 5856218; fax: +39 075 5856224.
*E-mail addresses:* lorenzo.raggi@gmail.com (L. Raggi), tissicarlo@gmail.com (C. Tissi), mazz@unitus.it (A. Mazzucato), valeria.negri@unipg.it (V. Negri).

0168-9452/$ – see front matter © 2013 Elsevier Ireland Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.plantsci.2013.11.001

Temperature also shows an important relationship with photoperiod sensitivity in common bean. It has been observed that genotypes from Colombia (typically grown in the cool highlands) are more photoperiod-sensitive in comparison with those from Venezuela (grown in warmer sites at lower altitudes) even though the latitudes of the two countries are similar [14]. All these factors make common bean a good system for investigating the effects of selection on candidate genes that control flowering time in relation to adaptation.

In this species, several major loci and Quantitative Trait Loci (QTL) have been reported for flowering time, photoperiod sensitivity and growth habit regulation [15 and refs. therein]. In particular, the *Ppd* (photoperiodic-induced delay in flowering) locus is linked to the *fin* locus that, with a second yet unnamed locus, appears to be responsible for determinacy [16–18]. Kwak and collaborators [15] reported that one of the *Arabidopsis Terminal Flower 1* (*TFL1*) homologues in common bean, *PvTFL1y*, co-segregated with the *fin* locus in a recombinant inbred line population from a cross between a wild indeterminate genotype and a domesticated determinate genotype. The results of an extensive analysis of the origin and diversity of the *PvTFL1y* mutation suggested that determinate types derived from indeterminate types [19]. Finally, homologues of *LEAFY* (*LFY*) and *APETALA1* (*AP1*) from several model legume species were isolated and characterized [8,20]. Their functions in the legume lateral secondary inflorescences architecture seem to be similar to those in the *Arabidopsis* inflorescence [21].

To effectively utilize natural genetic variation in plant breeding, knowledge of correlations between phenotype and genotype is of crucial importance. In recent years, Association Mapping (AM) has become an important tool to search for genes that are responsible for quantitative trait variation [22–24]. This approach relies on the detection of linkage disequilibrium (LD) between genetic markers and a phenotype of interest and exploits the recombination events accumulated over many generations to increase the accuracy of detected associations. In particular, candidate gene association studies specifically target genes with functions known to be involved in the control of a particular trait. With this method it is possible to look at variation directly in the genes rather than at anonymous markers. Candidate gene AM was successfully used to identify the genes underlying flowering time QTLs in *Arabidopsis* [25–27] and in important crop species [23,28–32]. While the use of the genome-wide scan approach has increased in the scientific community, the candidate gene approach is still widely used. The increasing number of known genes that control flowering time and flower development makes these complex traits particularly attractive for association studies.

The aim of this study was to investigate the flowering time variation and the molecular polymorphism in key regulatory genes that control flowering time in a *P. vulgaris* collection. Associations between morphological and molecular data were tested. The genetic structure and relatedness of the collection were taken into consideration in order to reduce the number of spurious associations.

## 2. Materials and methods

### 2.1. Plant material and DNA extraction

The collection was kindly provided by the "Scuola di Scienze Agrarie, Forestali, Alimentari ed Ambientali" (Università degli Studi della Basilicata, Potenza, Italy). It is composed of five wild accessions, three cultivars and 86 landraces (45 from Europe and 41 from the Americas) for a total of 94 accessions. Each accession was multiplied in isolation for seven generations starting from one seed each multiplication year. For each accession, the latitude of origin (*lat*) was retrieved from the Nordic Gene Bank (NGB, http://www.nordgen.org/ngb/) and from CIAT (ciat.cgiar.org) databases. A complete list of the studied materials with details about their origin, biological and cultivation status and phaseolin type is reported in Supplementary Table 1.

DNA was isolated from one plant per each accession using a TissueLyser II (Qiagen) and the DNeasy 96 Plant Kit (Qiagen) according to the protocols provided by the manufacturer. The same plant was used for morphological characterization and was reproduced in isolation, as described above, in 2011. One plant from its progeny was used for repeating morphological analyses in 2012.

### 2.2. Morphological characterization of the collection

Morphological characterization of accessions was carried out on a single plant basis in a nursery for two consecutive years (2011 and 2012). The nursery, with an automatic irrigation system, was located in Perugia Italy, (latitude 43°06′09.56″ N, longitude 12°23′40.71″ E, elevation 443 m asl.). Accessions were sown in pots (40 cm diameter), using a slightly acid soil mixture of clay loam (30%) and potting-compost (70%), at the end of April. For each accession, the following traits related to flowering were recorded on the first three inflorescences: inflorescence length (*inl*, mm, at the opening of the first flower), flower bud size (*fbs*, as a score of 3, 5, and 7 for small, medium and large flowers, respectively, referring to the oldest unopened flower) and flower buds per inflorescence (*fbi*, number). In addition, days-to-flowering (*dtf*, days from sowing to the opening of the first flower) and duration of flowering (*dof*, days from the opening of the first flower to the end of flowering) were recorded. The plant architecture traits recorded were: growth habit (*gwh*, 1 = determinate bush, 2 = indeterminate bush with erect branches, 3 = indeterminate bush with prostrate branches, 4 = indeterminate with semi-climbing main stem and branches) and nodes on the main stem (*nms*, number of nodes before the first inflorescence). At the end of September, the plants that had not yet flowered were moved into a climate-controlled greenhouse with an automatic irrigation system. A *dtf* of 162 was assigned to the individuals that had not flowered by the end of the experiment (162 dd from sowing). Zhao et al. [25] reported that this procedure did not affect the association analysis results. The range of variation for the quantitative traits scored in the two years was depicted by box plots.

Spearman's rank correlation coefficients between the recorded traits were calculated using the arithmetic means over years. In addition, in order to test the reliability of morphological characterization, the Spearman's rank correlation coefficients between 2011 and 2012 trait data were calculated using the SAS software (SAS Institute Inc, 2004). In order to delineate clusters of morphologically similar individuals, a principal component analysis (PCA) was carried out. The effects of genotype ($i$) and environment ($j$) on each morphological trait were investigated by analysis of variance (ANOVA), after standardization of data ($\log_{10}$), and following the model $Y_{ij} = \mu + \alpha_i + \beta_j$ (where $\mu$ is the general mean, $\alpha_i$ is the genotypic main effect $\beta_j$ the environmental main effect). In the ANOVA, each year of experimentation was considered as a different environment. In addition, the effect of geographical origin on *dtf* (Europe *vs* America) was also investigated using one-way ANOVA procedure. All analyses were conducted using XLSTAT (Addinsoft).