GENETIC MAPPING OF PEA \textit{(Pisum sativum L.)} GENES INVOLVED IN SYMBIOSIS (review)

V.A. ZHUKOV\textsuperscript{1}, O.Yu. SHTARK\textsuperscript{1}, T.A. NEMANKIN\textsuperscript{1}, 2, A.A. KRYUKOV\textsuperscript{1}, A.Yu. BORISOV\textsuperscript{1}, I.A. TIKHONOVICh\textsuperscript{1}, 3

\textsuperscript{1}All-Russia Research Institute for Agricultural Microbiology, Federal Agency of Scientific Organizations, 3, sh. Podolskogo, St. Petersburg, 196608 Russia, e-mail zhukovf01@yahoo.com;
\textsuperscript{2}Biocad Joint Stock Co., 34-A, ul. Svyazi, pos. Strelna, St. Petersburg, 198515 Russia;
\textsuperscript{3}Saint Petersburg State University, 7/9, Universitetskaya nab., St. Petersburg, 199034 Russia (ORCID: Zhukov V.A. orcid.org/0000-0002-2411-9191, Borisov A.Y. orcid.org/0000-0001-9834-7368, Tikhonovich I.A. orcid.org/0000-0001-8968-854X)

Acknowledgements: Authors wish to thank Dr. E.S. Ovchinnikova (University of Sydney, Australia) for assistance in experiments with pea molecular markers.

Supported by Russian Science Foundation, grant № 14-24-00135

Received December 14, 2015

Abstract

This article presents a review of current data on genetic mapping of pea \textit{(Pisum sativum L.)} genes participating in development and regulation of arbuscular-mycorrhizal and legume-Rhizobial symbioses. By means of mutual analysis several regulatory symbiotic genes (\textit{Sym}-genes) were identified in model and crop legumes, particularly, among them more than 40 pea \textit{Sym}-genes. Some of them are already cloned and sequenced, and structural and functional similarity was demonstrated for orthologous \textit{Sym}-genes in different legume species. The functions of these genes are diverse and include the control of perception of the microsymbiont’s signal molecules, activation of the signal cascade (which is common for both legume-rhizobial and arbuscular-mycorrhizal symbioses), and consequent transcriptional changes in root cortex. To identify the sequence of mutated pea genes, an approach is used that is based on comparative genetic mapping and search for candidate gene in the genome of closely related legume plant barrel medic (\textit{Medicago truncatula} Gaertn.). The web site www.phytozome.net (D.M. Goodstein et al., 2012) presents the current state of the barrel medic’s genome sequencing in the form of genome browser, which facilitates the search for homologous genes and the sequence analysis of candidate genes. Significant similarity of pea and barrel medic genomes allows development of gene-based molecular markers, comparison of obtained pea genetic map with \textit{M. truncatula} genome, and pea gene cloning after finding mutations with similar phenotypic manifestation. Currently, most of pea \textit{Sym}-genes are mapped in genome; that resulted in identification of the sequences of 14 symbiotic genes. In particular, authors of the present review were able to sequence the pea genes \textit{Sym35} — the homologue of \textit{NIN} of \textit{Lotus japonicus} (Regel.) K. Larsen (A.Y. Borisov et al., 2003), \textit{Sym37} — the homologue of \textit{NFR1} of \textit{L. japonicus} (V.A. Zhukov et al., 2008), \textit{Sym33} — the homologue of \textit{IPD3} of barrel medic (E. Ovchinnikova et al., 2011), \textit{Cochleata} — the homologue of \textit{NOOT} of barrel medic (J.M. Couzigou et al., 2012). In recent years, considering the development of modern technologies of Next Generation Sequencing and massive genotyping, an avalanche of data on mapping pea gene-based data is being accumulated. Saturation of pea genetic map with markers, undoubtedly, will facilitate mapping of symbiotic genes and identification of their sequences; this will help to broaden the understanding of how the system of genes, which control interactions with beneficial soil microorganisms, functions in pea.

Keywords: legumes, legume-rhizobial symbiosis, arbuscular mycorrhiza, symbiotic plant genes, genetic mapping, synteny, gene-based molecular markers

Mutualistic interactions with microorganisms are essential for evolution and a successful life strategy of terrestrial plants [1, 2]. Benefits acquired by plants as a result of symbiosis are a facilitated uptake of minerals from soil and the enhanced tolerance to infections and abiotic stressors [3, 4]. Arbuscular mycorrhiza (AM), an interaction between plant roots and \textit{Glomeromyces} fungi, is the most common endosymbiosis typical of 80-90 % land plants. AM provides
nutrient assimilation (mainly phosphorus and nitrogen-containing substances) from soil [4]. The symbiosis between legume Fabaceae plants and rhizobia (Alphaproteobacteria, Betaproteobacteria and Gammaproteobacteria) is best known among the nitrogen-fixing plant–microbe symbioses. In legume-rhizobium symbiosis (LRS) the specialized structures (nodules) colonized by bacteria that fix atmospheric nitrogen are developed on plant roots [3]. AM and LRS have a lot of similarities, pointing to their evolutionary relationship [5, 6]. High genetic and metabolic integration of the partners is characteristic of both symbioses. AM and LRS development involves certain molecular and cellular events which are clearly interrelated among partners and accompanied by a constant exchange of signals and metabolites [6, 7].

Experimental mutagenesis is widely used to elucidate the mechanisms underlying formation and the functioning of mutualistic plant–microbe symbioses. In particular, extensive collections of bean mutants allow to identify several dozen regulatory symbiotic genes [8-11], also called Sym (symbiosis) genes [12]. Many of them have been cloned and sequenced, and various legume species show the similarities in structure and activity of orthologous Sym genes. The functions of these genes vary and comprise the control of reception of the signaling molecules produced by microsymbionts, activation of common signaling cascade for AM and LRS (CSC) and the subsequent changes in transcriptome of the root cortex cells [6, 7, 13].

Garden pea (Pisum sativum L.) belongs to the most important legume crop in Russia and in the world [14]. There are over 100 independent pea mutants defective in LRS (rarely AM) development with 44 Sym genes affected [1, 8, 9, 15-21]. A promising approach to identification of the sequence of mutated pea genes is based on a comparative genetic mapping and search for candidate gene in the genome of closely related legume plant barrel medic Medicago truncatula Gaertn. and Lotus japonicus (Regel.) K. Larsen [8, 22-24].

The purpose of this paper is to overview current data on the genetic mapping of pea Sym genes, including the results obtained in ARRIAM.

Genetic mapping technique. Genetic mapping is based on the classical genetics including determination of linkage groups and estimation of recombination frequency. The distances between genes are measured in map units (m.u.)/Centimorgan (cM) equal to one percent recombinant phenotypes [25]. Genetic mapping involves statistical analysis of massive data on marker allele segregation in a certain generation after the crossing. To this end, at least a dozen of software has been developed, for example, MAPMAKER 3.0, MapL98, JoinMap 4.1, etc. [26-30].

Both morphological and molecular markers (protein- and DNA-based) allow pea gene mapping, however, only the involvement of DNA markers in genetic analysis made it possible to develop large-scale genetic maps with high resolution [31]. Over the past 25 years, various versions of the pea genome genetic maps have been constructed with the use of molecular markers [22, 32-43].

Gene-specific EST (expressed sequence tag) markers identified by amplification of the expressed sequence fragments are considered the most appropriate for genetic mapping pea genes after their detection by experimental mutagenesis. To design primers, the known EST sequence is commonly used. If the restriction site of the endonuclease applied to detect an allelic state of the amplification product coincides with the single nucleotide polymorphism site, the markers are referred to as CAPS (cleaved amplified polymorphic sequences) type markers. Allelic polymorphism of such markers is manifested as different numbers and sizes of the resulting restriction fragments separated by agarose gel electrophoresis [44].
Gene-specific markers in legumes allow to exploit the phenomenon of genomic synteny to transfer knowledge between model (*Lotus japonicus*, alfalfa) and crop legumes including pea [24, 45]. Synteny is referred to as the same order of homologous genes on the chromosomes of different species. This is micro-synteny in case of synteny within a cluster of several closely located genes, and this is macro-synteny if such a pattern is typical of the whole chromosomes or their extended sections.

Research aimed at developing new gene-specific markers and construction of genetic maps for legumes disclosed both macro- and micro-synteny in their genomes [22, 40]. Detection of high synteny among pea and alfalfa *Medicago truncatula* genomes generated a new impetus to the pea genetics development, especially after the successful sequencing of the alfalfa genome. Due to high genome similarity, the knowledge about the alfalfa genes can be used in studying pea genes, developing their gene-specific markers and gene cloning, once homologous mutations (i.e., mutations with similar phenotypic manifestation) found [24]. The Plant Comparative Genomics portal of the Department of Energy’s Joint Genome Institute (USA, http://www.phytozome.net) [46] represents current information on alfalfa genome sequencing using interface of a genome browser that facilitates the search for homologous genes and candidate gene analysis.

**Genetic mapping as a tool for identification of pea symbiotic gene sequences.** Over the past 10 years, the set of more than 100 gene-specific molecular markers has been designed in ARRIAM for mapping pea *Sym*-genes of all linkage groups. In this, we used known ESTs which positions in the genome are described in literature (so-called anchor markers).

Oligonucleotide primers for PCR and amplification of target DNA fragments in different pea lines have been designed using nucleotide sequences of selected ESTs from GenBank (NCBI, http://www.ncbi.nlm.nih.gov). The resulting set has been allowed us to define first or clarify the localization of ten symbiotic genes [20, 47, 48]. The figure shows a generalized diagram summarizing all the data, including those obtained in ARRIAM, on the genetic mapping of pea symbiotic genes [9, 20, 47–49].

![Pea symbiotic gene mapping](image-url)

**Pea (*Pisum sativum* L.) symbiotic (*Sym*) gene mapping:** I-VII — the linkage groups. For *Sym17* (in parentheses) the position should be additionally confirmed [50]. The genes which position has been determined or confirmed in ARRIAM (St. Petersburg, Russia) are marked black and those reported...
by other researchers are marked gray. The relative position of these genes is not defined: Coch—Sym38, Sym16—Sym27, Sym14—Sym22, Sym11—Sym13, Sym15—Sym29.

Genetic mapping of mutants using gene-specific markers allowed us to identify positions of 13 Sym-genes in pea. Many of them were determined simultaneously with that of orthologous genes in other legumes, including model species [20, 48, 51-57]. Generalized scheme to determine order of the pea genes included four steps. These are the gene localization on pea genetic map; search for known genes in the syntenic region of the model legumes (Medicago truncatula or Lotus japonicus); selection of candidate homologous genes for the target pea gene by phenotypic comparison between the mutants in pea and the model legume species; amplification and sequencing the pea gene with the primers designed based on the sequence of a model legume candidate gene [8].

The seven pea genes have been sequenced in ARRIAM independently or in collaboration with world’s leading laboratories. Of them, the pea Sym33, Sym35 and Sym37 genes have been sequenced based on the homology to alfalfa gene IPD3 [56], Lotus japonicus gene NIN [52] and NFRI gene [54], respectively. The Sym33 gene encodes the protein that interacts with Ca$^{2+}$-calmodulin-dependent protein kinase (CaM kinase), the central component of common signaling cascade (CSC), Sym35 gene encodes the LRS-specific CSC-dependent transcription factor, and Sym37 genes encodes the Nod factor (bacterial signal) receptor [6]. The pea Sym40 gene was identified as a homologue of previously known gene EFD, encoding a transcription factor of the ethylene response factor (ERF) group in alfalfa [48], and the mutation corresponding gene Sym41 is a variant (weak allele) of another symbiotic pea gene Sym19, encoding the CSC-involved receptor kinase [20]. The orthologous genes Cochleata and NOOT controlling nodule meristem have been recently identified as homologs of Arabidopsis thaliana (L.) Heyhn. BOP1 and BOP2 genes involved in leaf and floral morphogenesis [57]. Six of these pea genes have been sequenced based on the gene maps except Sym34 with unknown localization. For Sym34 sequencing the homology to the alfalfa NSP1 gene which encodes CSC-dependent transcription factor common to LRS and AM was assumed solely on the basis of the phenotype of the pea mutant [21].

Thus, in pea the nucleotide sequences of 14 Sym-genes are presently identified which is only a third of that identified by mutational analysis, i.e., many Sym-genes have yet to be mapped and sequenced.

Next generation technologies — an outlook. In recent years, with the development of next-generation sequencing and massive genotyping, the information on gene-specific makers in pea is increasingly accumulated. Next generation sequencing (NGS) allows a high speed to read a great number of fragments, which makes it possible to analyze polymorphism of thousands sequences considered as potential markers. Massive genotyping using GoldenGat microchips (Illumina, USA) allows a one-time analysis of thousands designed markers. Due to these approaches, several thousand gene-specific markers have been mapped in pea in the past two years [43].

The whole genome sequencing in P. sativum undoubtedly will facilitate genetic mapping of the symbiotic genes. However, pea, like many other legumes, is still waiting for a chance to be among the species with sequenced genomes [58], while its studies are mostly limited to transcriptome sequencing [59-62].

Thus, the use of gene-specific markers for genetic mapping allows to exploit genome synteny to transfer data obtained for model legumes (Lotus japonicus, alfalfa) to economically significant crops, including pea. Design of new-type "interspecies” markers derived from pairs of orthologous genes from different leguminous plants and saturation of pea genetic map obviously can greatly
assist in the implementation of a complete pea genome assembly based on the next generation sequencing (NGS). In turn, the mapping of symbiotic genes and identification of their sequences will help in better understanding of how the pea genes can control plant interaction with beneficial soil microorganisms.

REFERENCES

17. Jacob L.M., Petrova O.S., Tsyganov V.E., Borisov A.Y., Tikhonovich I.A.


25. Inge-Vechtomov S.G. *Genetika s osnovami selektcii* [Genetics and basic breeding]. St. Petersburg, 2010 (in Russ.).


Ovchinnikova E., Journet E.P., Chabaud M., Cosson V., Ratet P., Duc G., Fedorova E., Liu W., Op den Camp R., Zhukov V., Tikhonovich I., Bo-


