The symbiosis genes of pea

LaRue, T.A.
and
Weeden, N.F.

Boyce Thompson Institute for Plant Research
Ithaca, NY 14853-1801,USA
Department of Horticultural Sciences
Cornell University, Geneva NY 14456, USA

The ability to form a symbiosis with nitrogen fixing bacteria occurs in 10 families of dicotyledonous plants. The best known of these of course is the Leguminosae; approximately 85% of legumes can form nodules (1). In agronomy, the nitrogen fixing ability of soybean, lucerne and clover are of major research interest because of the great areas planted with these crops. However, the species most often studied for basic knowledge is the pea, which has served as an experimental plant from the earliest investigations. The early work is reviewed in detail by Fred et al. (10).

Boussingault in 1838 provided the first quantitative data that clover and peas increased their nitrogen content from the air, but he suspected it was an absorption of ammonia. Atwater in 1884 had data indicating that the N increment of peas was from fixation of nitrogen gas ("against the best evidence and opinion at the present time"). However, he could not fathom the means. It was in 1886, that Hellriegel clearly demonstrated with pea that N fixation required a symbiosis with microbes in the root nodules. Use of organisms isolated from pea nodules was central when, in 1889, Prazmowski showed that the symbiotic associations were species-specific.

The bacteria that form a symbiosis with legumes are called rhizobia. At the present time, the strains that nodulate *Pisum*, *Lens* and *Vicia* are classified as *Rhizobium leguminosarum* biovar *viciae*. In this species, about 20% of the DNA is in large plasmids. On one of these, the *sym* plasmid, are found most of the known genes involved in symbiosis.

Nodule development

As microscopic methods have improved, nodule development on pea has been repeatedly described (2, 24, 28). Briefly, competent rhizobia induce the root hairs to curl ("shepherd's crook") and a "thread" of cellulosic material forms in the hair. The rhizobia enter this infection thread and multiply as the thread elongates. After passing from the base of the epidermal cell into the outer cortex, the thread branches. As the branches penetrate towards the inner cortex, anticlinal cell divisions occur in advance of it, and a nodule primordium is established adjacent to the pericycle near the xylem pole. The thread, branching further, penetrates some of the dividing cells, releasing bacteria from the tips of the threads, where they are surrounded by a peribacteroid membrane of plant origin. The infected plant cells enlarge, the bacteria multiply, nitrogen fixing enzymes are formed in the bacteria and leghemoglobin appears in the cytoplasm of the infected cells.

In rhizobia, the nitrogen fixing enzyme nitrogenase, containing Fe and Mo, is coded for by *nif* genes, which are homologous to the *nif* genes found in free living nitrogen fixing bacteria. In addition, mutational analysis of *R. leguminosarum* has identified *nod* and *fix* genes, which are involved in nodule formation or function, respectively. Some of these genes are host-specific, while others (the "common" genes) are conserved and occur in all species and biovars of rhizobia. The symbiosis also may be delayed by mutations in *exo* genes, which are involved in production of the exopolysaccharide of the bacteria (2).

The rhizobia bind to the root. There is evidence that the specificity of binding is due to the plant's lectin. A gene for a pea lectin was introduced into *Trifolium repens* with *Agrobacterium rhizogenes* as a vector. The resulting transformed clover roots could then be nodulated by *R. leguminosarum* bv. *viciae* (4). Binding may be to the polysaccharide capsule of the rhizobium, but solid evidence for this is lacking. An alternate theory is that the bridge is by a calcium requiring binding protein produced by the rhizobia (29).

The *nod* genes are induced by compounds secreted by the roots. In pea, these may include naringenin (34), apigenin-7-0-glucoside, eriodictyol, 7-hydroxyflavone, 3, 5, 7, 3' tetra hydroxy-4' methoxy flavone and 7, 3' hydroxy-4'methoxyflavone. The *nod* genes *D*, *ABC*, *EF* and *L* are induced to form a tetrasaccharide of D-glucosamine on which a hydroxy and 3 amino groups are acetylated and another amino group is acylated by a C 18:4 unsaturated fatty acid. This compound causes root hair curling and cortical cell divisions in roots (30).

Genetic analysis

Two approaches have been made in studying the host genes involved in symbiosis. One is to identify proteins (nodulins) whose concentration is enhanced in developing or functional nodules, compared with uninoculated roots. The other method is to characterise genes identified in naturally occurring or induced mutants which have an altered symbiosis.

The organ-specific proteins which appear first are the "early nodulins" (Table 1), which were identified by differential screening of a pea nodule cDNA library. The protein ENOD 12 (28) is induced in root hairs by cell-free culture filtrate of appropriate rhizobia. It is found also in cells containing the infection thread, in cells in advance of the infection thread, and in the developing nodule. The protein is rich in hydroxyproline, and may contribute to the cell wall reformation required for the infection thread.

ENOD5 is also hydroxyproline rich, but is similar to the arabinogalactans (28). Its distribution suggests that it is part of the plasma membrane of the infection thread.

ENOD2 which is also rich in hydroxyproline is found in the nodule parenchyma. This layer of cells is thought to control the flux of oxygen into the nodule (12, 34).

Table 1. Nodulins and nodule enhanced proteins in *P. sativum*

Designation	Site or role in nodule	Ref.
Leghemoglobins	Cytosol of infected cells; aid in diffusion of O_2	12
GS_{N1}, GS_{N2}	Glutamine synthetases. Incorporation of newly fixed NH_3	31
ENod2	Repeating pentapeptides, hydroxy proline rich. Nodule parenchyma	34
ENod5	Hydroxyproline-rich cell wall protein; infection thread	23
ENod12	Arabinogalactan-like; may precede infection thread	23
ENod3, ENod14	Transient expression. Cysteine clusters suggest metal binding. May be involved in transport of Fe or Mo to bacteroids	23
Nod40'	Appears before leghemoglobin	12
N40, N68	Appear at same time as leghemoglobin	12
N21	Appears late in development, after leghemoglobin	12

Leghemoglobin, which in legumes is coded for by a multi-gene family, appears just before nitrogen fixation is detectable. A feature of all legume nodules is that they contain enhanced amounts of a form of glutamine synthetase (GS_n), different from the isoforms found in cytosol (GS_c) or plastids (GS_p). At present, the pea is unique in having two nearly identical forms of GS_n (31). These "twins" differ by only a few amino acids (32).

Some naturally occurring symbiosis variants are known for pea, but the genes involved have seldom been backcrossed into standard lines for experimental studies. Fortunately, it has proven easy to obtain induced mutants with altered symbiotic properties. Our group (17), Jacobsen (26), and Duc (5) have made mutants of the cultivars Sparkle, Rondo and Frisson, respectively. We are now testing these three collections for allelism. These mutants include peas with excessive nodule numbers (nod^{++}), no or few nodules (nod^{-}), or nodules which do not adequately fix nitrogen (nod^{+}, fix^{-}).

Table 2. Genes of *P. sativum* involved in nodule formation or function.

Locus	Phenotype and defining line	Ref.
nod-1; nod-2	High nodule number. 'Parvus'	11
nod-3	nod^{++}. Rondo nod3	26
c	Controls nodule number. 'Afghanistan III'	25
Sym-1	Nodulation requires > 26°C. 'Iran'	21
sym-2	Strain specific; infected by rhizobial strain TOM bearing nod X. 'Afghanistan'	15, 16
sym-3	nod^{-} fix^{-} (unconfirmed). 'Afghanistan'	15
Sym-4	Strain specific nodulation; *P. humile* JI261	21
sym-5	nod^{-}. Mutational 'hot spot'. E2, R88	17
sym-6	Ineffective nodules with strain F_{13}. 'Afghanistan'	21
sym-7	nod^{-}. E69	17
sym-8	nod^{-}. R25	22
sym-9	nod^{-}. R72	22
sym-10	nod^{-}. N15	17
sym-11	nod^{-}. N24	20
sym-12	nod^{-}. K5	26
sym-13	$nod^{+}fix^{-}$. E135; P58 (g)	5, 19
sym-14	nod^{-}. E135N	19
sym-15	Few nodules, short lateral roots. E151	17
sym-16	Few nodules, short lateral roots. R50	17
sym-17	Few nodules; dwarf; short thick roots. R72	20
sym-18	Strain specific nodulation. E54	17
sym-19	nod^{-}. NEU5 K24	17,26
sym-20	nod^{-}. R80	17
sym-21	Few nodules; short lateral roots. E132	17
Sym-22	Few or no nodules. *P. humile* JI1794	
sym-23	$nod^{+}fix^{-}$. P59	5
sym-24	$nod^{+}fix^{-}$. P60	5
sym-25	$nod^{+}fix^{-}$. P61	5
sym-26	$nod^{+}fix^{-}$. P63	5
sym-27	$nod^{+}fix^{-}$. P12	5
sym-28	nod^{++}. Short internodes and stem deformation. 190 F	5
	$nod^{+}fix^{-}$. $FN_1$27	27
brz	Few nodules; bronze leaves. E107	14, 18

Table 2 lists pea genes involved in the symbiosis which have been defined by inheritance tests. Allelism tests for mutants without gene designations are still in progress. About 30 genes have already been identified; this number approximates the number of *nod* and *fix* genes known in the microsymbiont.

In most cases, the phenotype is independent of the strain of *R. leguminosarum* tested. However, some genes condition an altered specificity, *sym-2*, originally described in the

primitive line 'Afghanistan' (15, 21), results in nodulation by only a few strains of *R. leguminosarum.* Most of these strains carry a gene *nod* X on the *sym* plasmid, a gene not found in non-infective strains (3). To date, this is the only example of a "gene for gene" effect in nodule formation. Loci *Sym-4* and *sym-6* in primitive lines and induced mutations at *sym-18* also condition restricted strain specificity, but the rhizobial genes involved are not yet studied.

One of the factors in the host determining *Rhizobium* strain specificity is the lectins produced in roots (4). However, Lu et al. (unpublished) have demonstrated that two of the strain-specificity mutants in pea (*sym-2* and *sym-18*) do not reflect mutations in the lectin genes but map to different chromosomal locations. The strain-specificity mutations may still involve lectins by way of genes producing post-translational modifications. Alternatively, strain-specificity may be influenced by other factors as yet not identified.

The pea controls nodule number, in part, by aborting infection. The biochemical nature of this auto-regulation is unknown, but some mutants now available may provide useful experimental material, for these mutants hypernodulate. Our laboratory has been unable to obtain high nodule numbers on the cv. 'Parvus', so the existence of *nod-1* and *nod-2* is unconfirmed. The *nod-3* mutant of Rondo (27) and the 190F line (*sym-28*) of Frisson (5) both have higher than normal nodule numbers. Allelism tests have not been reported, but the phenotypes are different. Unlike *nod-3,* 190F is pleiotropic; the shoot has short internodes. Grafting experiments indicate that the hypernodulation of *nod-28* is controlled by the genotype of the shoot (5) while that of *nod-3* is controlled by the genotype of the root (27).

We have only begun to characterise the many nod$^-$ mutants. The large number indicates that it should be possible to "dissect" nodule formation by blocking it at many stages. R25 (*sym-8*) does not react at all to the presence of rhizobia; lacking even root hair curling. R72 (*sym-9*) shows root hair swelling in the presence of *R. leguminosarum,* but no curling or infection. These two mutants are obviously blocked at the earliest steps of infection (22).

A mutational "hot spot" must occur at *sym-5,* for seven independently isolated mutants were obtained at this locus (17). Except for forming few or no nodules, these mutants appear normal. There are the normal number of infections, but nodule primordia are seldom formed in advance of the infection threads (13). The *sym-5* mutants may involve an increased sensitivity to the hormone ethylene, for primordia initiation and nodule formation are increased by lowered root temperature (which lowers ethylene formation) (9) or by exogenous inhibitors of ethylene formation or action (8).

The *sym-5* mutants have an altered 66 kD peptide, constitutive to all plant parts (7). It is not yet known if this peptide is the product of *sym-5.*

Duc (5) has identified six non-allelic genes required for nodule function. One of these (g) is allelic to the previously described *sym-13.* Line E135F (*sym-13*) is characterised by small nodules which senesce early (19). The nodules do not have nitrogenase or leghemoglobin.

Table 3. Location of genes in *P. sativum* linkage map.

Gene	Chromosome	Linked loci
sym-2	1	*d; Idh*
sym-5	1	*d; Idh*
sym-7	3	*Lap-1*
sym-8	6	*Arg*
sym-10	1	*I*
sym-11	7	*Skdh*
sym-13	7	*Skdh*
sym-14	2	*Fum*
sym-15	7	*Aldo; Amy-1*
sym-16	5	*Pgd-c*
sym-17	6	*Prx-3*
sym-18	1	*d; Idh*
sym-19	1	*d; Idh*
Sym-22	2	*Fum*
leghemoglobins	1	*d; Idh*
GS_{nI}	6	*Acp-4*
GS_{nII}	7	*Skdh*
lectins	7	*Amy-1; Gal-2*
brz	4	*Was*

An important observation was that some of the nod^- mutants do not form a symbiosis with vesicular arbuscular mycorrhiza (myc^-) (6). This suggests that the two different symbionts must share some process.

Gene mapping

Whereas the *nif*, *nod* and *fix* genes of rhizobia are closely clustered on the *sym* plasmid, the plant genes are widely distributed (Table 3). We have found that each chromosome carries at least one gene involved in nodulation (33). The genes seem randomly distributed with one notable exception. Within a relatively small region (10 cM) of chromosome 1 are found the genes for leghemoglobin and *sym-2, sym-5, sym-18* and *sym-19* (33). *sym* 2, described in the primitive line 'Afghanistan' (15), and *sym-18* condition narrow strain specificity. Mutants at *sym-19* are nod^- and myc^-.

The existence of such a cluster of functionally related genes is unparalleled in higher plant genomes. Gene families, such as the genes encoding the small subunit of ribulose bisphosphate carboxylase, the 45S ribosomal array, or leghemoglobin, are often found in very tightly linked clusters. However, genes related by function, such as those involved in carbohydrate metabolism, chloroplast function, or anthocyanin synthesis are usually widely scattered in the genome. Hence, the *"sym"* cluster on chromosome 1 is particularly intriguing. Initial mapping data place *sym-5* at one end of this cluster, *sym-2, sym-18* and *sym-19* at the other end, and the leghemoglobin genes in the middle. Although *sym-2, sym-18* and *sym-19* have been shown to be non-allelic, recombination between pairs of these loci has not been demonstrated.

The positions of other *sym* loci are not particularly remarkable. The lectin genes map in the vicinity of *sym-11,* but there is no reason to suspect that *sym-11* is a mutation in the expressed lectin gene because there is no change in strain specificity. The recently described dominant low nodule number mutant, *Sym-22,* maps very close to *Fum* (chromosome 2), as does *sym-14.* It is possible that these represent alleles at the same locus, despite one being a dominant mutation and the other recessive. A non-nodulating mutant that also exhibited short lateral roots, *sym-16,* maps near a known mutant, *coh,* displaying short root laterals; however the two mutants complement each other, the hybrid displaying normal roots and a normal nodulation response.

Future Research

The genetic knowledge of nodulation in *P. sativum* is far in advance of that of any other species. This information should be extended and exploited in at least two ways. Firstly, it is desirable to bridge the gap between the *sym* genes and the nodulins; i.e., to identify the products of the *sym* genes on one hand and to subject the nodulins to classical genetic analysis.

Secondly, it may be time to test the utility of *sym* mutants. The genes controlling high nodule number and restricted strain specificity should be applied by pea breeders and agronomists to design cultivars which will nodulate profusely, but only with superior strains of rhizobia which are designed to be compatible with the cultivar.

---

1. Allen, O.N. and Allen, E.K. 1981. The Leguminosae. Univ. Wisconsin Press.
2. Brewin, N.J. 1991. Ann. Rev. Cell Biol. 7:191-226.
3. Davis, E.O., Evans, L.J. and Johnston, A.W.B. 1988. Mol. Gen. Genet. 212: 531-535.
4. Diaz, C.L., Melchers, L.S., Hooykaas, P.J.J., Lugtenberg, B.J.J, and Kijne, J.W. 1989. Nature 338: 579-581.
5. Duc, G. and Messager, A. 1989. Plant Sci. 60:207-213.
6. Duc, G.A., Trouvelot, A., Gianinazzi-Pearson, V. and Gianinazzi, S. 1989. Plant Sci. 60:215-222.
7. Fearn, J.C. and LaRue, T.A. 1990. Plant Mol. Biol. 14:207-216.
8. Fearn, J.C. and LaRue, T.A. 1991. Plant Physiol. 96:239-244.

9. Fearn, J.C. and LaRue, T.A. 1991. Plant, Cell and Environ. 14:222-227.
10. Fred, E.B., Baldwin, I.L. and McCoy, E. 1932. Root Nodule Bacteria and Leguminous Plants. Madison.
11. Gelin, O. and Blixt, S. 1964. Agri Hort. Genet. 22:149-159.
12. Govers, F., Gloudemans, T., Moerman, M., van Kammen, A. and Bisseling, T. 1985. EMBO J. 4:861-867.
13. Guinel, F.C. and LaRue, T.A. 1991. Plant Physiol. 97:1206-1211.
14. Guinel, F.C. and LaRue, T.A. 1992. Plant Physiol. 99:515-518.
15. Holl, F.B. 1975. Euphytica 27:767-770.
16. Kneen, B.E. and LaRue, T.A. 1984. Heredity 52:383-389.
17. Kneen, B.E. and LaRue, T.A. 1988. Plant Sci. 58:177-182.
18. Kneen, B.E., LaRue, T.A., Welch, R.M. and Weeden, N.F. 1990. Plant Physiol. 93:717-722.
19. Kneen, B.E., LaRue, T.A., Hirsch, A.M., Smith, C.A. and Weeden, N.F. 1990. Plant Physiol. 94:899-905.
20. Lee, K.H. and LaRue, T.A. 1992. Plant Physiol. 100:1326-1333.
21. Lie, T.A. 1984. Plant Soil 82:415-425.
22. Markwei, C.M. and LaRue, T.A. 1992. Can. J. Microbiol. 38:548-554.
23. Nap, J.-P. and Bisseling, T. 1990. Science 250:948-954.
24. Newcomb, W. 1976. Can. J. Bot. 54:2163-2186.
25. Ohlendorf, H. 1983. Z. Pflanzenzuchtung 91:13-24.
26. Postma, J.G., Jacobsen, E. and Feenstra, W.J. 1988. J. Plant Physiol. 132:424-430.
27. Postma, J.G., Jager, D., Jacobsen, E. and Feenstra, W.J. 1990. Plant Sci. 68:151-161.
28. Scheres, B., van Engelen, F., van der Knaap, E., van der Wiel, C., van Kammen, W. and Bisseling, T. 1990. Plant Cell 2:687-700.
29. Smit, G., Tubbing, D.M.J., Kijne, J.W. and Lugtenberg, B.J.J. 1991. Arch. Microbiol. 155:278-283.
30. Spaink, H.P., Sheeley, T.M., van Brussel, A.A.N., Glushka, J., York, W.S., Tak, T., Geiger, O., Kennedy, E.P., Reinhold, V.N. and Lugtenberg, B.J.J. 1991. Nature 354:125-129.
31. Tingey, S.V., Walker, E.L. and Coruzzi, G.M. 1987. EMBO J. 6:1-9.
32. Walker, E.L. and Coruzzi, G.M. 1989. Plant Physiol. 91:702-708.
33. Weeden, N.F., Kneen, B.E. and LaRue, T.A. 1990. *In* Nitrogen Fixation: Achievements and Objectives. Chapman and Hall, London, pp. 323-330.
34. van der Weil, C., Sheres, B., Franzen, H., van Lierop, M.-J., van Lammeren, A., van Kammen, A. and Bisseling, T. 1990. EMBO J. 9:1-7.