MPMI Vol. 8, No. 1, 1995, pp. 172–176

Research Note

GmN56, a Novel Nodule-Specific cDNA from Soybean Root Nodules Encodes a Protein Homologous to Isopropylmalate Synthase and Homocitrate Synthase

Hiroshi Kouchi and Shingo Hata

National Institute of Agrobiological Resources, Tsukuba, Ibaraki 305, Japan
Received 14 June 1994. Accepted for 14 September 1994.

We have isolated a novel nodule-specific cDNA clone, GmN56, from soybean root nodules. The expression of GmN56 was induced almost concomitantly with the onset of nitrogen fixation, together with leghemoglobin and other late nodulin genes. *In situ* hybridization studies demonstrated the localization of GmN56 mRNA in the bacterial infected cells of mature nodules. The predicted amino acid sequence of the GmN56 protein exhibits significant homology to those of LeuA (isopropylmalate synthase) of several microorganisms and NifV (putative homocitrate synthase) of nitrogen-fixing bacteria, suggesting that GmN56 encodes an enzyme catalyzing a reaction involving acetyl-CoA and α-keto acid as substrates.

Additional keywords: Glycine max, nodulins, nitrogen fixation, symbiosis.

The symbiotic interaction between legume plants and rhizobia leads to the formation of root nodules in which the intracellular bacteria are capable of fixing atmospheric nitrogen. This interaction is accompanied by the specific expression of sets of both plant and bacterial genes. The products of plant genes induced exclusively in nodules are termed 'nodulins' (Van Kammen 1984). Nodulins are classified into early and late nodulins according to the time of appearance of their transcripts during the nodule developmental process. The early nodulin genes are expressed at early stages of nodule formation and are considered to be involved in the bacterial infection process and/or nodule morphogenesis (Nap and Bisseling 1990). On the other hand, late nodulins, which appear after completion of nodule organogenesis and just before the onset of nitrogen-fixation activity, play essential roles in symbiotic functioning, i.e., nitrogen fixation and related processes. They have been further divided into metabolic and structural nodulins, according to defined or predicted functions and their intracellular localization (Verma *et al.* 1989).

We previously isolated a set of nodulin cDNAs from soybean nodules at different developmental stages and described the characterization of several early nodulin genes (Kouchi and Hata 1993). While screening for nodule-specific genes by means of a subtractive cDNA hybridization technique, we also identified several novel late nodulin cDNAs. In this re-

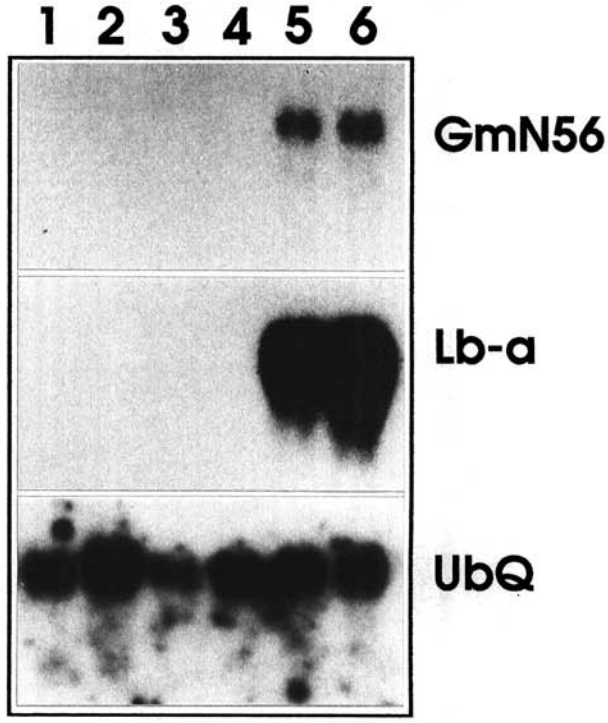

Fig. 1. Accumulation of GmN56 mRNA during nodule development. Poly(A⁺) RNA (1 μg) isolated at various times during nodule development was separated by denaturing 1% agarose gel electrophoresis, blotted onto nitrocellulose membranes, and then probed with cDNAs of GmN56, leghemoglobin (Lb-a) and ubiquitin (UbQ, as a control), using standard procedures (Sambrook *et al.* 1989). Lanes (1) uninfected roots; (2) 4-day-old nodulated root segments; (3) 8-day; (4) 10-day; (5) 13-day; (6) 17-day-old nodules.

Corresponding author: H. Kouchi.

Present address of S. Hata: Faculty of Agriculture, Kyoto University, Kyoto 606, Japan.

The nucleotide sequence of GmN56 has been submitted to the GSDB, DDBJ, EMBL, and NCBI nucleotide sequence databases with the accession number, D38015.

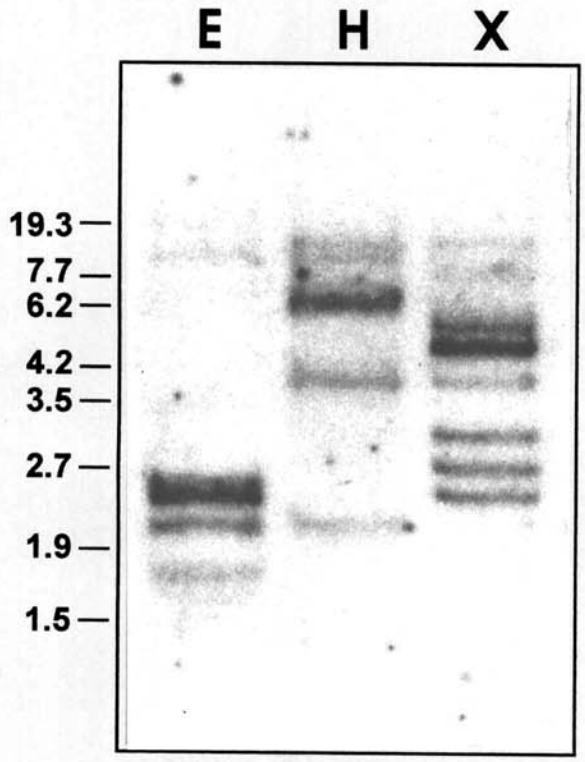

Fig. 2. Southern blot analysis of soybean genomic DNA with GmN56 cDNA as a probe. Soybean DNA (10 µg) was digested to completion with *Eco*RI (E), *Hin*dIII (H) or *Xba*I (X), subjected to 0.8% agarose gel electrophoresis, blotted onto a nitrocellulose membrane, and then hybridized with ^{32}P-labeled GmN56 cDNA. Genomic DNA was isolated from etiolated soybean hypocotyls as described (Rogers and Bendich 1988).

port, we describe the sequence analysis, temporal and spatial expression, and some functional implications of one of these cDNA clones, designated as GmN56.

The strategy for the isolation of nodule-specific cDNAs was described previously (Kouchi and Hata 1993). In brief, ^{32}P-labeled first-strand cDNA was synthesized from nodule poly(A)$^+$ RNA and then subjected to two cycles of subtractive hybridization with a large excess of uninfected root poly(A)$^+$ RNA. Unhybridized cDNA was selected by hydroxylapatite column chromatography and used to screen nodule cDNA libraries. By this procedure, GmN56 was isolated from the nodules of soybean plants (*Glycine max* L. 'Akisengoku') 21 days after sowing and inoculation of *Bradyrhizobium japonicum* (strain A1017).

Figure 1 shows the time course of accumulation of GmN56 mRNA, together with that of leghemoglobin (Lb) during nodule development. Under our conditions of plant growth, nodules were visible as small bumps 6–7 days after sowing and inoculation, and nitrogen fixation (acetylene reduction) activity first appeared at day 13 (Kouchi *et al.* 1989). GmN56 mRNA (approximately 2.4 kb) was first detected at day 13, at the same time as leghemoglobin mRNA. Thus, GmN56 has been assigned as a late nodulin gene. GmN56 mRNA was not detected in either uninfected roots, stems, or leaves of soybean seedlings (data not shown). GmN56 message is much rarer than those of Lb and the peribacteroid membrane-related nodulin genes tested, as judged by the comparison of the intensities of hybridization signals. Soybean genomic DNA was digested to completion with restriction enzymes and then subjected to Southern blot analysis using the entire GmN56 cDNA as a probe (Fig. 2). Two or more major hybridization bands and multiple weaker bands were observed even under conditions of high stringency, suggesting that GmN56 comprises a small gene family. The GmN56 cDNA probe did not

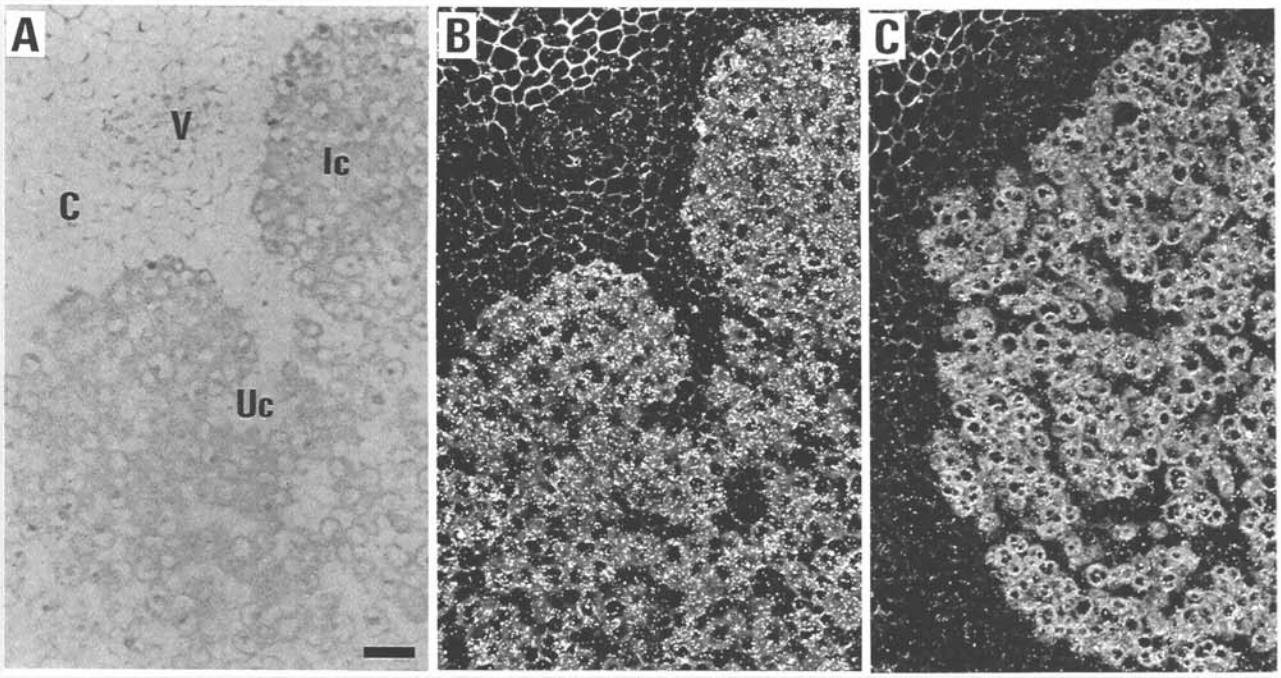

Fig. 3. Localization of GmN56 mRNA in soybean nodules (16-day-old). Paraffin-embedded sections (7 µm thick) were probed with ^{35}S-labeled antisense RNA of GmN56 according to the *in situ* hybridization procedure (Yang *et al.* 1993). Bright- **(A)** and dark- **(B)** field micrographs are shown. Hybridization to the sense probe is also shown **(C)**. Ic, infected cell; Uc, uninfected cell; C, cortex; V, vascular bundle. Bars represent 100 µm.

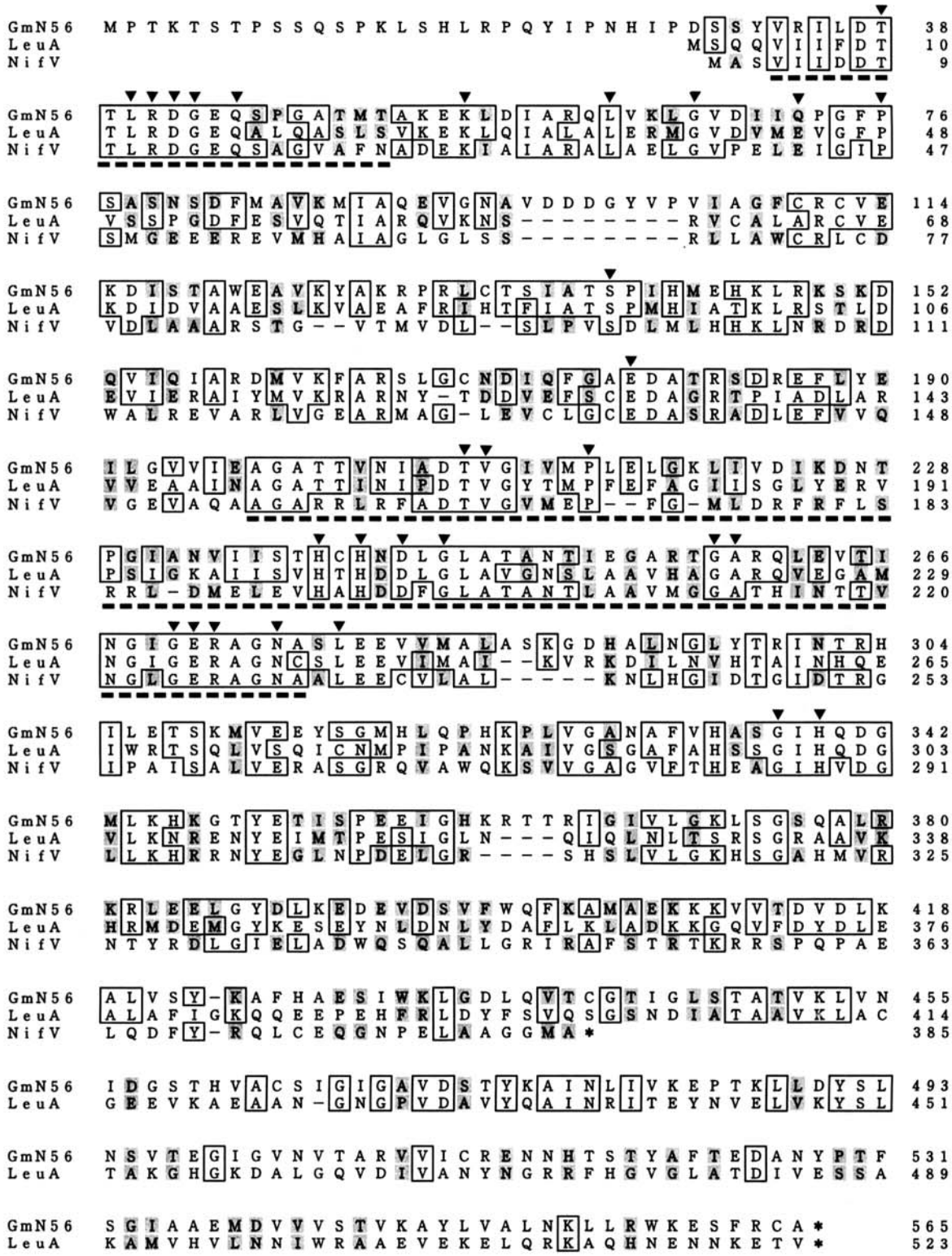

Fig. 4. Comparison of the amino acid sequences of the GmN56 protein (deduced), isopropylmalate synthase of *E. coli* (LeuA), and the NifV protein of *A. vinelandii* (NifV). Identical amino acids are boxed and similar amino acids are shaded. The gaps indicated by dashes were introduced to optimize the alignment. Conserved amino acids between the NifVs of *A. vinelandii, A. chroococcum,* and *K. pneumoniae* and LeuAs of *S. typhimurium* and *S. cerevisiae* are indicated by arrowheads (Evans *et al.* 1991). Two regions of high similarity between the NifV and LeuA proteins are shown by dashed underlines (Meijer and Tabita 1992). The nucleotide sequence of GmN56 was determined on both strands by the dideoxy-chain termination method.

hybridize to DNA of Bradyrhizobium (results not shown).

To localize GmN56 gene expression, we performed *in situ* hybridization experiments on nodule sections (Fig. 3). The hybridization signals were weak, probably due to the small quantity of GmN56 mRNA, but signals significantly above the background level were observed in the infected cells. The inner and outer cortices, and uninfected cells in the central part of the nodule showed no significant hybridization signal. Therefore, we conclude that the GmN56 gene expression is restricted to the infected cells.

The GmN56 cDNA clone was sequenced entirely. It has an insert of 2,407 bp and the longest open reading frame consists of 565 amino acids. A protein database search revealed significant homology of the deduced GmN56 protein to the LeuA protein (2-isopropylmalate synthase, EC 4.1.3.12) of several microorganisms, and also to the NifV protein of *Azotobacter* species and *Klebsiella pneumoniae* (Fig. 4). The highest homology (40.6% identical over 478 amino acids) was found to LeuA of *Escherichia coli* (Yura *et al.* 1992). Similar levels of identity were found to LeuA of *Salmonella typhymurium* (Ricca and Calvo 1990), and *Lactococcus lactis* (Godon *et al.* 1992). The homology to the NifV protein of *A. vinelandii* was 33% over 410 amino acids (53%, if conservative amino acids are included).

The sequence homology between the NifV and LeuA proteins has been well documented (Evans *et al.* 1991; Meijer and Tabita 1992). Highly conserved amino acids or regions of high similarity between the NifV and LeuA proteins are also well conserved in the GmN56 protein (Fig. 4). NifV is involved in the synthesis of homocitrate, which is a component of the Fe-Mo cofactor of nitrogenase (Hoover *et al.* 1989), and is postulated to be a homocitrate synthase (EC 4.1.3.21). Although the exact reaction catalyzed by NifV has yet to be determined, it is most likely that NifV catalyzes the condensation of acetyl-CoA and α-ketoglutarate to form homocitrate, based on the analogy to a reaction catalyzed by LeuA that forms isopropylmalate from acetyl-CoA and α-ketoisovalerate. In addition, homocitrate synthase involved in the α-aminoadipate pathway of lysine biosynthesis has been demonstrated to form homocitrate by the condensation of acetyl-CoA and α-ketoglutarate (Jaklitsch and Kubicek 1990). Two regions which are well conserved between the LeuA and NifV proteins are postulated to be responsible for the binding and/or condensation of acetyl-CoA and the α-keto group (Evans *et al.* 1991). Thus, it is very likely that the GmN56 protein catalyzes a reaction involving acetyl-CoA and an α-keto acid as substrates. In this connection, Meijer and Tabita (1992) have suggested that these two well-conserved regions are responsible for the binding of the α-keto acid group rather than the acetyl-CoA, based on the homologies of these two regions with pyruvate carboxylase and oxaloacetate decarboxylase.

GmN56 mRNA is synthesized exclusively in mature nodules. The timing and site of expression of GmN56 suggest that it is involved in a process closely related to the nitrogen fixation activity of the bacteroids. One of the most important physiological features of the infected cell cytoplasm is the production of large amounts of C_4-dicarboxylates, which serve as substrates for bacteroid respiration (Tajima and Kouchi 1990). Thus, taking into account the possible role of GmN56 protein in the catalysis of a reaction involving acetyl-CoA and/or an α-keto acid, it is intriguing to hypothesize that GmN56 has a function related to C_4-dicarboxylate metabolism in the infected cells.

An alternative, and even more intriguing hypothesis from the view of functional symbiosis, might be that the GmN56 protein catalyzes the formation of homocitrate in the plant cytosol. Although neither a structural nor functional homologue of the *nifV* gene has yet been identified in *Bradyrhizobia*, there is no doubt that *Bradyrhizobium* possesses the complete machinery for synthesizing the Fe-Mo cofactor because an intact nitrogenase can be induced in the free-living state under certain conditions. Nevertheless, the additional supply of an essential component of the active center of nitrogenase by the host plant might be necessary to meet the demand of highly efficient and continuing synthesis of the nitrogenase system in symbiotic bacteria. Obviously, further enzymological investigations are now required to examine these hypothesis and to address the question of the function of GmN56.

ACKNOWLEDGMENTS

We wish to thank Wei-Cai Yang and Ton Bisseling of Agricultural University, Wageningen, The Netherlands, for their suggestions and support in relation to the in situ hybridization experiments. This research was supported by Grants-in-Aid to H. K. from the Ministry of Agriculture, Forestry and Fisheries of Japan (BMP-94-IV-1-1), and to S. H. from the Shorai Foundation for Science and Technology.

LITERATURE CITED

Evans, D. J., Jones, R., Woodley, P. R., Wilborn, J. R., and Robson, R. L. 1991. Nucleotide sequence and genetic analysis of the *Azotobacter chroococcum nifUSVWZM* gene cluster, including a new gene (*nifP*) which encodes a serine acetyltransferase. J. Bacteriol. 173:5457-5469.

Godon, J.-J., Chopin, M.-C., and Ehrlich, S. D. 1992. Branched-chain amino acid biosynthesis genes in *Lactococcus lactis* subsp. *lactis*. J. Bacteriol. 174:6580-6589.

Hoover, T. R., Imperial, J., Ludden, P. W., and Shah, V. K. 1989. Homocitrate is a component of the iron-molybdenum cofactor of nitrogenase. Biochemistry 28:2768-2771.

Jaklitsch, W. M., and Kubicek, C. P. 1990. Homocitrate synthase from *Penicillium chrysogenum*: Localization, purification of the cytosolic isoenzyme, and sensitivity to lysine. Biochem. J. 269:247-253.

Kouchi, H., Tsukamoto, M., and Tajima, S. 1989. Differential expression of nodule-specific (nodulin) genes in the infected, uninfected and cortical cells of soybean (*Glycine max*) root nodules. J. Plant Physiol. 135:608-617.

Kouchi, H., and Hata, S. 1993. Isolation and characterization of novel nodulin cDNAs representing genes expressed at early stages of soybean nodule development. Mol. Gen. Genet. 238:106-119.

Meijer, W. G., and Tabita, F. R. 1992. Isolation and characterization of the *nifUSVW-rpoN* gene cluster from *Rhodobacter sphaeroides*. J. Bacteriol. 174:3855-3866.

Nap, J. P., and Bisseling, T. 1990. Developmental biology of a plant-prokaryote symbiosis: The legume root nodule. Science 250:948-954.

Ricca, E., and Calvo, J. M. 1990. The nucleotide sequence of *leuA* from *Salmonella typhimurium*. Nucleic Acids Res. 18:1290.

Rogers, S. O., and Bendich, A. J. 1988. Extraction of DNA from plant tissues. Page A6/1-A6/11 in: Plant Molecular Biology Manual S. B. Gelvin, R. A. Schilperoort, and D. P. S. Verma, eds. Kluwer Academic Publisher, Dordrecht.

Sambrook, J., Fritsch, E. F., and Maniatis, T. 1989. Molecular Cloning: A Laboratory Manual. 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Tajima, S., and Kouchi, H. 1990. Metabolism of C_4-dicarboxylates and amino acids in soybean nodule bacteroids, and their role in energy supply to nitrogenase. Pages 309-314 in: Nitrogen Fixation: Achievements and Objectives. P. M. Gresshoff, L. Evanth Roth, G. Stacey, and W. E. Newton, eds. Chapman and Hall, New York.

Van Kammen, A. 1984. Suggested nomenclature for plant genes

involved in nodulation and symbiosis. Plant Mol. Biol. Rep. 2:43-45.
Yang, W.-C., Katinakis, P., Hendriks, P., Smolders, A., deVries, F., Spee, J., Van Kammen, A., Bisseling, T., and Franssen, H. 1993. Characterization of GmENOD40, a gene showing novel patterns of cell-specific expression during soybean nodule development. Plant J. 3:573-585.
Yura, T., Mori, H., Nagai, H., Nagata, T., Ishihama, A., Fujita, N., Isono, K., Mizobuchi, K., and Nakata, A. 1992. Systematic sequencing of the *Escherichia coli* genome: Analysis of the 0–2.4 min region. Nucleic Acids Res. 20:3305-3308.