Monitoring symbiotic gene expression in *Rhizobium* sp. NGR234

X. Perret and W.J. Broughton

LBMPS, Université de Genève, 1 ch. de l'Impératrice, 1292 Chambésy, Genève, Switzerland.

ABSTRACT

Rhizobium sp. NGR234 has the largest capacity to nodulate legumes of any soil bacterium known. Its genome comprises three replicons: a 536 kb symbiotic plasmid, a c.a. 2,000 kb mega-plasmid, and a 3,700 kb chromosome. Sequencing the 536 kb plasmid, pNGR234*a* showed that apart from *nodE*, *nodG*, and *nodPQ*, most Nod-factor biosynthetic genes are plasmid-borne, and regulated by a number of transcriptional regulators of the LysR family. Of these, NodD1 is of primary importance. It is activated by a variety of phenolic compounds found in the rhizospheres of plants, and the NodD1-flavonoid complex enhances expression of *nod*-box controlled genes. A second part of the broad host-range of NGR234 can be attributed to the 16 *nod*-genes which are responsible for the elaboration of a large family (>80 members) of lipo-oligosaccharidic Nod-factors. A further 10 genes (*rhcC1* to *rhcV*) encode functions necessary for the elaboration of the bacterial type III protein secretion system (TTSS). The TTSS is an essential component of pathogenicity in both animal (e.g. *Salmonella, Shigella, Yersinia*) and plant (*Erwinia, Pseudomonas*, *Ralstonia*, and *Xanthomonas*) pathogens. Secretion of at least two proteins is controlled by the TTSS of NGR234 and occurs after flavonoid induction. Inactivation of the TTSS by insertional mutagenesis blocks the secretion of both proteins, and strongly affects the ability of NGR234 to nodulate a variety of tropical legumes including *Pachyrhizus tuberosus* and *Tephrosia vogelii*. High resolution transcription analysis of pNGR234*a* shows that many more genes than the known *nod*, *nol*, *rhc*, *nif* and *fix* loci are transcribed under symbiotic conditions. This, together with the large number of newly discovered symbiotic promoters suggests that many more genes than previously thought are necessary for a functional symbiosis.

Introduction

Symbioses between legumes and *Azo*(*Brady*)(*Sino*)*Rhizobium* contribute more fixed nitrogen to the global pool than any other association of diazotrophs. Our prokaryotic model for analysis of symbiotic determinants uses the broad host-range *Rhizobium* sp. NGR234 which nodulates more than 110 genera of legumes as well as the non-legume *Parasponia andersonii* [9]. Most of NGR234's symbiotic determinants are carried by the 536 kb plasmid, pNGR234*a* [1,2,6]. Using techniques developed for guanosine/cytosine rich genomes [3], we were able to sequence a set of overlapping cosmids [7] and thus rapidly derive the complete nucleotide sequence of pNGR234*a* [4]. This replicon contains 416 putative open reading frames (ORFs), a third of which show no significant similarity to any known protein.

Microbial Biosystems: New Frontiers
Proceedings of the 8th International Symposium on Microbial Ecology
Bell CR, Brylinsky M, Johnson-Green P (eds)
Atlantic Canada Society for Microbial Ecology, Halifax, Canada, 1999.

Analysis of these 416 ORFs by traditional methods (e.g. insertional mutagenesis) is obviously impractical since screening for altered phenotypes on 110 genera of legumes is almost impossible. Instead, we developed methods to monitor differential gene expression under conditions that NGR234 could encounter during symbiosis. Using specific primer-pairs, 441 PCR fragments containing all the predicted ORFs as well as intergenic regions were amplified. After separation on agarose gels, these fragments were transferred to filters, and hybridised against radioactively labelled RNA derived from either liquid cultures of NGR234 or from nodulated plants. To mimic the early and late stages of root-hair infection, liquid cultures of NGR234 were grown with or without an inducer of *nod*-gene expression (daidzein) for 1 and 24 h, respectively. Genes expressed within the plant were identified by labelling RNA taken from plants forming determinate- (*Vigna unguiculata*) or indeterminate-nodules (*Cajanus cajan*) [8].

Results

Approximately 60% (247) of the 416 genes and gene-fragments encoded by pNGR234*a* hybridised to radioactively labelled RNA. No transcriptional activity was detected for the remaining 169 under the conditions tested. Undoubtedly, relative differences in turnover of the mRNA, and the presence of functional, duplicated genes elsewhere in the genome, affect the intensity of hybridisation. Addition of daidzein (200 nM) triggered expression of 147 genes (Fig. 1, upper panel), 14 of which belong to three classes of insertion sequences (NGRIS-3, NGRIS-4 and NGRIS-5). Most strongly induced genes are linked to the three *hsn* loci which encode most Nod-factor modifying enzymes [5], and to the TTSS cluster [10].

Computational analyses revealed 19 putative *nod*-box like regulatory sequences, at least six of which are involved in the expression of the known *nod*-loci (see Fig. 1; upper panel). Interestingly, the transcription patterns of downstream ORFs suggest that most of the remaining regulatory elements are probably also functional. Apparent lack of promoter-activity of two putative *nod*-boxes is probably not the result of accumulated mutations however, since major differences between the DNA sequence of these elements and a *nod*-box consensus were not found. In contrast, another two *nod*-boxes seem to have retained regulatory activity despite the absence of properly oriented downstream genes. The observation that one of the three identical, plasmid-borne copies of NGRIS-5 is adjacent to one of these *nod*-boxes implies that a gene was probably separated from its promoter by re-arrangements within the genome. Many flavonoid-inducible genes are not directly under the control of *nod*-box like sequences, however. This points to the existence of alternative or indirect regulatory pathways as in the case of the *rhc* loci, some of which are thought to be activated by y4xI, a *nod*-box controlled regulator of transcription.

Differentiation of rhizobia into nitrogen-fixing bacteroids results in important changes in gene expression (Fig. 1; lower panel). In essence, these involve repression of the majority of flavonoid-inducible transcripts and induction of those required for effective nitrogen fixation. In contrast to the *nod*-genes that are dispersed over pNGR234*a*, *nif*- and *fix*-genes form a 55 kb cluster (*fixU* to y4xE), most of which are regulated by 10 NifA-σ^{54} promoters (Fig. 1; lower panel). Amongst these are putatively new *fix* genes, such as y4wA and y4wB

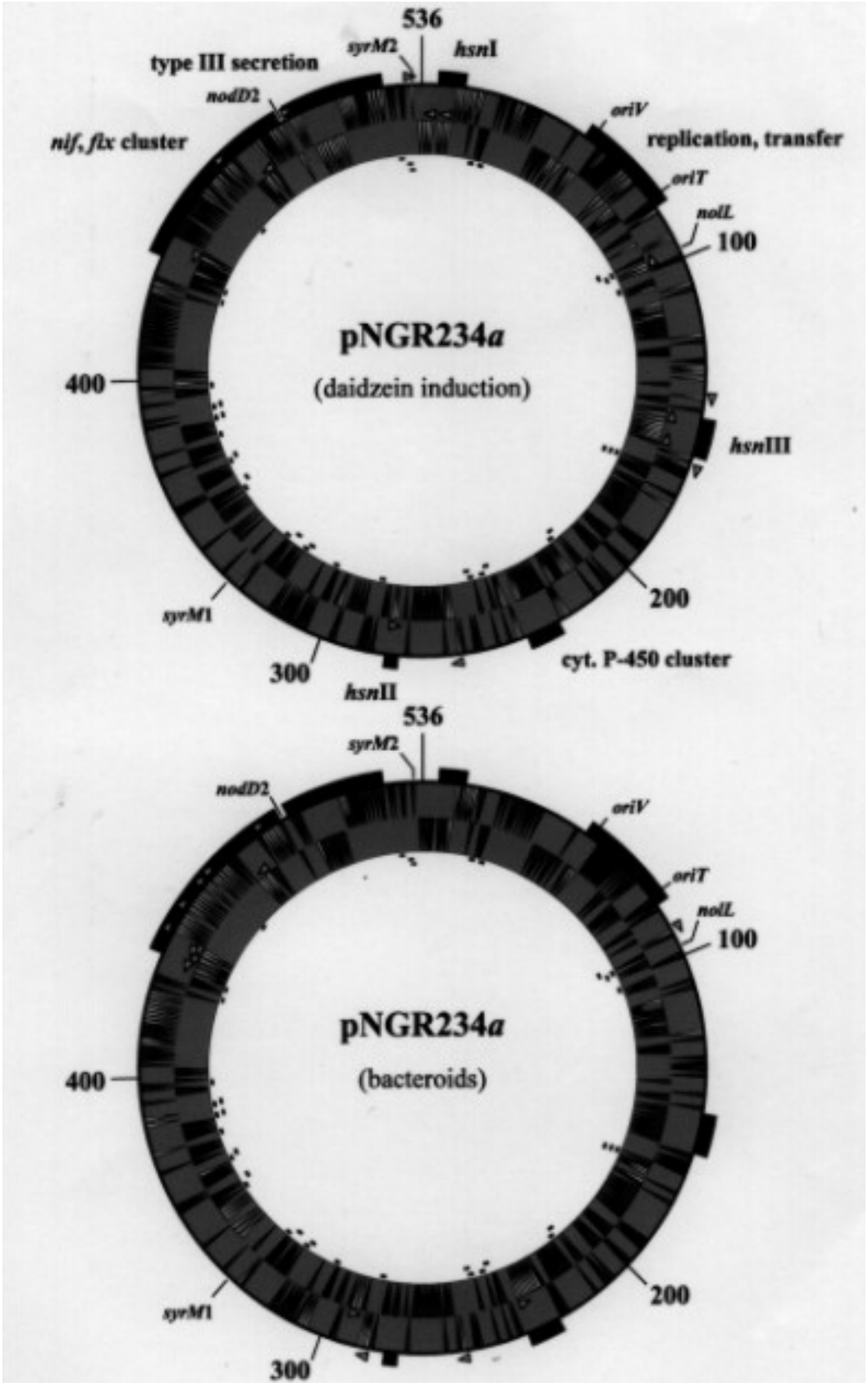
536
syrM2
hsnI
type III secretion
nodD2
oriV
nif, fix cluster
replication, transfer
oriT
nolL
100
pNGR234a
(daidzein induction)
400
hsnIII
200
syrM1
300
cyt. P-450 cluster
hsnII
536
syrM2
nodD2
oriV
oriT
nolL
100
pNGR234a
(bacteroids)
400
200
syrM1
300

Figure 1. Symbiotic transcription map of pNGR234*a*. Upper panel represents those genes which are expressed 1 and 24 h after induction with daidzein. Genes transcribed in nodules of *Vigna unguiculata* are shown in the lower panel. Coloured triangles in the outer and inner concentric circles indicate open reading frames identified on the plus and minus strands, respectively. Unless actively transcribed (marked in yellow), genes are marked in red. For instance, the *hsn* loci are induced by flavonoids (upper panel) but silenced in nodules (lower panel), whereas the cytochrome P450 cluster is expressed in bacteroids. Silent ORFs representing IS-like elements are shown in black, whereas those that are expressed are marked in light blue. Symbiotic promoters identified by sequence analysis are shown as small yellow triangles oriented according to the sense of transcription. Those in the upper panel represent active *nod*-boxes whereas those in the lower diagram represent symbiotic NifA-σ^{54} promoters. Thin concentric black lines represent gene fragments. Positions of major gene clusters are highlighted by thick grey bars.

which encode products that belong to the M16 family of zinc peptidases and are transcribed in nodules.

Discussion

Perhaps because the methods we used failed to detect low-level gene-expression, many predicted putative ORFs are apparently silent. Also, the conditions tested here represent only a small subset of those encountered by NGR234 during its life cycle. Nevertheless, flavonoid-inducible and bacteroid-specific genes represent at least 60% of the 441 ORFs and gene fragments, indicating that pNGR234*a* truly is a symbiotic replicon. In addition, several large intergenic regions seem to encode transcripts induced by flavonoids or the enveloping plant. Analysis of these regions using a new and more sensitive GeneMark matrix compiled from the genes of pNGR234*a* failed to detect new ORFs, however.

Timing of gene expression correlates well with the presence of *nod*-boxes or NifA-σ^{54} regulatory sequences, suggesting that the search for promoter motifs in large sequences may help identify new symbiotic genes. Generally, *nod*-box dependent genes are flavonoid-inducible and expressed early in the symbiosis while NifA-σ^{54} dependent loci play a role in bacteroids and nitrogen fixation. Interestingly, some loci are apparently controlled by both (e.g. *nod*-box #13 and a NifA-σ^{54} regulatory sequence are immediately upstream of y4vC). By combining sequence and promoter analysis with data from high-resolution transcription maps, it is now possible to define a number of new symbiotic targets that can be inactivated by insertional mutagenesis.

With these methods, it is now possible to measure the impact of various environmental stimuli on the expression of symbiotic genes. Imposition of heat- or cold-shock treatments, adaptation to changes in pH or in osmolarity, growth in micro-aerobic conditions or in the presence of competing *Rhizobium* strains, as well as treatments with root exudates of host- or non-host legumes are all worth studying. Furthermore, the effects of mutations in known transcriptional regulators on gene expression may shed light on the regulatory pathways encoded by pNGR234*a* and their possible interconnections. Ultimately, analyses of this kind may be extended to the complete genome of NGR234, for which an ordered cosmid library has been constructed.

References

1. Broughton WJ, Heycke N, Meyer ZAH, Pankhurst, CE (1984) Plasmid-linked *nif* and "*nod*" genes in fast growing rhizobia that nodulate *Glycine max, Psophocarpus tetragonolobus* and *Vigna unguiculata*. Proc Natl Acad Sci USA 81:3093-3097.

2. Broughton WJ, Wong C-H, Lewin A, Samrey U, Myint H, Meyer ZAH, Dowling DN, Simon, R (1986) Identification of *Rhizobium* plasmid sequences involved in recognition of *Psophocarpus*, *Vigna* and other legumes. J Cell Biol 102:1173-1182.
3. Freiberg C, Perret X., Broughton WJ, Rosenthal A (1996) Sequencing the 500-kb GC-rich symbiotic replicon of *Rhizobium* sp. NGR234 using dye terminators and a "thermostable sequenase": a beginning. Genome Res 6:590-600.
4. Freiberg C, Fellay R, Bairoch A, Broughton WJ, Rosenthal A, Perret, X (1997) Molecular basis of symbiosis between *Rhizobium* and legumes. Nature 387:394-401.
5. Jabbouri S, Relic´ B, Hanin M, Kamalaprija P, Burger U, Promé D, Promé J-C Broughton WJ (1998) *nolO* and *noel* (HsnIII) of *Rhizobium* sp. NGR234 are involved in 3-*O*-carbamoylation and 2-*O*-methylation of Nod-factors. J Biol Chem 279:12047-12055**.**
6. Lewin A, Rosenberg C, Meyer ZAH, Wong C-H, Nelson L, Manen J-F, Stanley J, Dowling DN, Dénarié J, Broughton WJ (1987) Multiple host-specificity loci of the broad host range *Rhizobium* sp. NGR 234 selected using the widely compatible legume *Vigna unguiculata*. Plant Mol Biol 8:447-459.
7. Perret X, Broughton WJ, Brenner S (1991) Canonical ordered cosmid library of the symbiotic plasmid of *Rhizobium* species NGR234. Proc Natl Acad Sci USA 88:1923-1927.
8. Perret X, Freiberg C, Rosenthal A, Broughton WJ, Fellay R. (1998) High resolution transcriptional analysis of the symbiotic replicon of *Rhizobium* sp. NGR234. Mol Microbiol, submitted.
9. Pueppke SG Broughton WJ (1998) *Rhizobium* sp. NGR234 and *R. fredii* USDA257 share exceptionally broad, nested host-ranges. Mol Plant Microbe Interact, in revision.
10. Viprey V, Del Greco A, Golinowski W, Broughton WJ, Perret X (1998) Symbiotic implications of type III protein secretion machinery in *Rhizobium.* Mol Microbiol. 28:1381-1389.