Research

Prediction and overview of the RpoN-regulon in closely related species of the *Rhizobiales*

Bruno Dombrecht*, Kathleen Marchal[†], Jos Vanderleyden* and Jan Michiels*

Addresses: *Centre of Microbial and Plant Genetics, Katholieke Universiteit Leuven, 3001 Heverlee, Belgium. [†]ESAT-SCD, Katholieke Universiteit Leuven, 3001 Heverlee, Belgium.

Correspondence: Jan Michiels. E-mail: jan.michiels@agr.kuleuven.ac.be

Abstract

Background: In the rhizobia, a group of symbiotic Gram-negative soil bacteria, RpoN (σ^{54}, σ^N, NtrA) is best known as the sigma factor enabling transcription of the nitrogen fixation genes. Recent reports, however, demonstrate the involvement of RpoN in other symbiotic functions, although no large-scale effort has yet been undertaken to unravel the RpoN-regulon in rhizobia. We screened two complete rhizobial genomes (*Mesorhizobium loti*, *Sinorhizobium meliloti*) and four symbiotic regions (*Rhizobium etli*, *Rhizobium* sp. NGR234, *Bradyrhizobium japonicum*, *M. loti*) for the presence of the highly conserved RpoN-binding sites. A comparison was also made with two closely related non-symbiotic members of the *Rhizobiales* (*Agrobacterium tumefaciens*, *Brucella melitensis*).

Results: A highly specific weight-matrix-based screening method was applied to predict members of the RpoN-regulon, which were stored in a highly annotated and manually curated dataset. Possible enhancer-binding proteins (EBPs) controlling the expression of RpoN-dependent genes were predicted with a profile hidden Markov model.

Conclusions: The methodology used to predict RpoN-binding sites proved highly effective as nearly all known RpoN-controlled genes were identified. In addition, many new RpoN-dependent functions were found. The dependency of several of these diverse functions on RpoN seems species-specific. Around 30% of the identified genes are hypothetical. Rhizobia appear to have recruited RpoN for symbiotic processes, whereas the role of RpoN in *A. tumefaciens* and *B. melitensis* remains largely to be elucidated. All species screened possess at least one uncharacterized EBP as well as the usual ones. Lastly, RpoN could significantly broaden its working range by direct interfering with the binding of regulatory proteins to the promoter DNA.

Background

The bacterial alternative sigma factor RpoN recognizes and binds a -24/-12-type promoter with the following consensus sequence: 5′-TG**G**CACG-N$_4$-TTG**C**W-3′ (the bold G and C are situated at position -24 and -12 relative to the transcription start site, respectively) [1]. Subsequently, the core DNA-dependent RNA polymerase (core-RNAP) binds to the RpoN-DNA secondary complex to form a stable closed promoter complex. The closed promoter complex is unable to initiate transcription by itself. For this, melting of the

double-stranded (ds) DNA within the closed complex is required [2]. This is accomplished by the nucleotide hydrolysis activity of an activator or enhancer-binding protein (EBP). EBPs bind to enhancer sites situated 100 base-pairs (bp) or more upstream of the transcription initiation site. Each EBP is controlled by its own signal transduction pathway, thereby responding to different conditions [3-5]. As there is almost no leaky expression in the absence of EBPs, the expression of different RpoN-dependent genes is tightly regulated by the different EBPs.

RpoN is also known as σ^{54} (from the 54 kDa molecular weight of the *Escherichia coli* polypeptide), σ^N (this sigma factor was initially discovered for its requirement for the expression of nitrogen metabolism genes) and NtrA or GlnF (names now not in common use). Because the members of this protein family vary considerably in molecular weight, the designation σ^{54} cannot be correctly applied to all of them. Furthermore, we use the name RpoN instead of σ^N as the link with the encoding gene *rpoN* is more obvious. As some bacteria (for example, *Bradyrhizobium japonicum*, *Rhizobium etli* and *Mesorhizobium loti*) have two copies of the gene, the respective proteins can be easily distinguished as $RpoN_1$ and $RpoN_2$.

RpoN helps initiating the transcription of genes encoding proteins for very diverse functions in a broad range of bacteria [5-7]. The processes controlled by RpoN are not essential for cell survival and growth under favorable conditions, with the exception of *Myxococcus xanthus* [8]. The most widespread RpoN-regulated function in bacteria is the assimilation of ammonia [6,7]. The expression of genes coding for glutamine synthetase, an ammonium transporter and the PII proteins is controlled by the NtrC EBP and RpoN. In the enteric nitrogen regulation (Ntr) paradigm, NtrC is activated via phosphorylation by NtrB. The PII protein, a signal transducer, stimulates dephosphorylation of NtrC-P under conditions of nitrogen excess (reflected by a high glutamine/2-ketoglutarate ratio), whereas it does not affect the NtrC phosphorylation status under nitrogen-limiting conditions [9]. Consequently, under low nitrogen conditions, NtrC-P activates transcription of the NtrC-RpoN regulon.

Species of the genera *Allorhizobium*, *Azorhizobium*, *Bradyrhizobium*, *Mesorhizobium*, *Rhizobium* and *Sinorhizobium* are generally referred to as rhizobia. Rhizobial *rpoN* mutants are deficient in symbiotic nitrogen fixation [10-14]. RpoN ensures the transcription of most of the nitrogen fixation genes (*nif/fix*) whose gene products constitute the nitrogenase complex and accessory proteins [15,16]. However, several other symbiosis-related genes are reported to be RpoN-dependent [17-27]. So far, no large-scale effort has been undertaken to unravel the RpoN-regulon in rhizobia. The recent publication of several rhizobial genomes and symbiotic regions is a good opportunity to identify RpoN-regulated genes in rhizobia. To obtain a good view on what

(symbiotic) functions are controlled by RpoN, we carried out an *in silico* analysis on the presence of -24/-12-type promoters in the symbiotic regions of *R. etli* CFN42, *Rhizobium* sp. NGR234, *B. japonicum* USDA110 and *M. loti* R7A [28-30] and in the genomes of *M. loti* MAFF303099 and *Sinorhizobium meliloti* 1021 [31,32]. Two closely related non-symbiotic species belonging to the *Rhizobiales* order of the α-proteobacteria, namely *Agrobacterium tumefaciens* C58 and *Brucella melitensis* 16M [33,34], were also included. These are non-symbiotic plant and animal pathogens, respectively. To date, there is only one report on RpoN-dependent functions in *A. tumefaciens* [35] whereas no information whatsoever is available on the RpoN-regulon of *B. melitensis*. The possible RpoN-dependent genes predicted by the screening were complemented with data from the literature and classified according to the function of the encoded proteins.

Results and discussion
In silico identification of potential RpoN-dependent promoters
The upstream intergenic sequences were extracted from the genomes of *M. loti* MAFF303099, *S. meliloti* 1021, *A. tumefaciens* C58 and *B. melitensis* M16 and the symbiotic regions of *R. etli* CFN42, *Rhizobium* sp. NGR234, *B. japonicum* USDA110 and *M. loti* R7A (see Materials and methods) [28-34]. The upper strand of these sequences was scored against a weight matrix using PATSER (see Materials and methods). Positive matches (possible RpoN-binding sites or -24/-12-type promoters) were classified according to the functions of the encoded gene products (see Additional data files, pages 8,9). To ensure a sound functional description, the predicted proteins were individually screened against the protein databases of the National Center for Biotechnology Information (NCBI) using BLASTP.

In our analysis, the number of false-positive matches is estimated to be very low. The use of PATSER in combination with a strong weight matrix is a preferred method to identify true binding sites [36]. The weight matrix used here was based on a set of 186 characterized RpoN-binding sites of 44 different bacterial species (Table 1) [1]. The lower threshold for the scores was chosen such that only matches strongly resembling the consensus were retained (see Materials and methods). This high stringency allows for a high specificity at the expense of the sensitivity. Indeed, some reports mention the presence of very poorly conserved -24/-12-type promoters that appear to be functional to some degree (see Additional data files, pages 8,9). Our stringent procedure most probably misses these sites. Therefore, the retained matches might represent an underestimation of the actual number of active -24/-12-type promoters. Second, only intergenic sequences were considered for the screening, as all known -24/-12-type promoters are situated in intergenic regions [7]. Moreover, the matches all have the correct orientation, as the upper strand

Table 1

Weight matrix based on 186 characterized -24/-12-type promoters [1]

	T	G	G* -24	C	A	C	G	N	N	N	N	T	T	G	C* -12	W
A	12	2	0	12	139	11	55	51	46	44	38	13	4	1	9	76
C	14	0	0	147	23	122	17	48	64	42	62	22	18	2	173	5
G	10	184	186	6	18	10	103	69	36	35	43	15	10	181	1	17
T	150	0	0	21	6	43	11	18	40	65	43	136	154	2	3	88

*Position relative to transcription start site.

of the intergenic sequences was used (for matches on the lower strand, see 'Additional control mechanisms involving RpoN'). The median of the distances from the -12 (C) position - the cytidylate residue on position -12 relative to the transcription initiation site - of a match to the start codon of the downstream coding sequence (CDS) varies from 74 to 156 bp (Table 2). This is somewhat different from the situation in *E. coli*, where the average distance amounts to 50 bp [7]. Although roughly 75% of the matches are within 200 bp upstream of the start codon, matches with distances over 1,000 bp were also retained (Table 2). A good example to justify this is the case of the mapped promoter of the *B. japonicum* gene *fixB*, which is situated 720 bp upstream of the coding region [37]. As is the case with other *in silico* prediction methods, the results of our analysis may have been biased by the approach used.

A good test case to evaluate the reliability of our predictions is to compare the matches with experimental data. In our laboratory we confirmed the RpoN-dependent expression of several *R. etli* genes with predicted RpoN-binding sites (see Additional data files, pages 3,4,6,7,12,13). The products of

Table 2

Comparison of different members of the *Rhizobiales*

	Re	NGR	Bj	R7A	Ml	Sm	At	Bm
Genome								
Upper strand matches*					82	43	31	12
RpoN controlled gene products					184	62	69	24
Size (Mb)					7.59	6.69	5.67	3.29
Matches/Mb					10.80	6.43	5.47	3.65
GC content (%)					62.4	62.1	58.1	57
Symbiotic region†								
Upper strand matches*	24	23	31	25	24	15		
RpoN controlled gene products	52	60	87	55	58	30		
Size (Mb)	0.37	0.54	0.41	0.50	0.61	1.35		
Matches/Mb	64.86	42.59	75.61	50.00	39.34	11.11		
GC content (%)	57.3	58.5	59.1	59.3	59.7	60.4		
Intergenic GC content (%)‡	55.0	55.9	57.3	58.0	57.0	56.4	52.0	49.9
Lower strand matches*	7	11	21	16	67¥	32¥	23¥	12¥
-12 (C) distance to start codon (bp)§	91 (231)	88 (149)	156 (214)	156 (288)	102 (209)	97 (210)	74 (142)	123 (345)
Possible EBPs¶	1	2	2	3	8	7	6	3

*Matches (possible RpoN-binding sites or -24/-12-type promoters) were predicted using PATSER (see Materials and methods). †p42d (Re); pNGR234*a* (NGR); 410 kb region on chromosome (Bj); 502 kb symbiotic island on chromosome (R7A); 610 kb symbiotic island on chromosome (Ml); pSymA (Sm). ‡The GC-content of the intergenic (non-coding) sequences was calculated with FramePlot 2.3.2 [68]. §The distribution of the -12 (C) position on the intergenic sequences is biased towards the corresponding downstream CDS (this distribution is visualized by means of a boxplot in the Additional data files, page 29). Therefore the median and third quartile (between brackets) were used as indicative parameters for these distances. ¶Possible EBPs were predicted using HMMER 2.2g (see Materials and methods) (see also Table 3). ¥The upstream intergenic region of the *rpoN* gene is included. At, *Agrobacterium tumefaciens* C58; Bm, *Brucella melitensis* 16M; Bj, *Bradyrhizobium japonicum* USDA110; Ml, *Mesorhizobium loti* MAFF303099; NGR, *Rhizobium* sp. NGR234; R7A, *M. loti* R7A; Re, *Rhizobium etli* CFN42; Sm, *Sinorhizobium meliloti* 1021.

these genes are involved in a wide variety of functions, such as nitrogen fixation (*nifH* and *iscN-nifUS*) [38,39], oxidative stress and gene regulation (*spxA-rpoN₂*) [13,40], and transport (*yp104-103* and *yp100*) [41] (G. Dirix, M. Van Guyse and J.M., unpublished results). For the genes on plasmid pNGR234a of *Rhizobium* sp. NGR234, an extensive expression analysis is available as well as an annotation of RpoN-dependent promoters [42]. Our screening confirmed the 15 previously annotated RpoN-dependent promoters on pNGR234a (controlling 50 CDSs). All these promoters are upregulated in bacteroids, resembling the NifA-RpoN-dependent symbiosis-specific expression pattern of the nitrogen fixation genes (see Additional data files). In addition to these promoters, eight other matches were found, controlling ten CDSs. Five out of these eight are highly expressed in bacteroids (see Additional data files), indicating that they are probably functional. In the symbiotic region of *B. japonicum*, 31 matches were found (controlling 87 CDSs) (Table 2), which were also previously annotated as RpoN-dependent promoters [29]. Thirteen of these promoters were shown to be functional and controlled by NifA and RpoN (see Additional data files). Furthermore, the fact that most - if not all - common RpoN-controlled genes were picked up for all species screened, including those already characterized in rhizobia, gives additional support to the methodology used. Common RpoN-dependent functions include assimilation of ammonium, transport of C_4-dicarboxylates and nitrogen fixation (see Additional data files, pages 1-7, 10).

Functional classification of RpoN-dependent genes in the *Rhizobiales*

The matches (also referred to as possible RpoN-binding sites or possible -24/-12-type promoters) obtained from the *in silico* search were complemented with literature data on RpoN-regulated genes in rhizobia, and classified according to protein function (see Additional data files).

For the four genomes screened, matches were found upstream of the genes whose products are involved in the uptake and assimilation of ammonia: *glnK-amtB*, *glnBA* and *glnII* (see Additional data files, pages 1,2). A similar regulation and gene organization is found in other α-proteobacteria [9]. *B. melitensis* differs slightly in that *amtB* constitutes a monocistronic unit and the *glnII* ortholog was not identified.

Possible RpoN-binding sites were found upstream of many nitrogen fixation genes in the six symbiotic, nitrogen-fixing species (*R. etli*, *Rhizobium* sp. NGR234, *B. japonicum*, *M. loti* MAF303099 and R7A and *S. meliloti*). Most rhizobial nitrogen fixation genes are controlled by RpoN and the NifA EBP (see Additional data files, pages 2-7). The latter protein is activated under microaerobic conditions and is highly active in bacteroids [15,16].

A possible RpoN-binding site is also present in the promoter region of *dctA*, encoding a C_4-dicarboxylate transporter, in all species screened, except for *B. melitensis*, where no such gene is present, and *R. etli*, where it is not present on the symbiotic plasmid (see Additional data files, page 10). In *M. loti* MAFF303099, -24/-12-type promoters were identified in front of the two C_4-dicarboxylate transporter genes. Rhizobial *dctA* mutants are deficient in nitrogen fixation. The expression of *dctA* is controlled under free-living conditions by RpoN-DctD, whereas in bacteroids, the DctD EBP is (partially) replaced by another unknown EBP. The DctD activator protein is activated by the DctB sensor protein that senses the presence of C_4-dicarboxylates together with DctA [16].

During symbiotic nitrogen fixation, hydrogen is produced as an obligate byproduct of the nitrogenase reaction. Some rhizobia possess a hydrogen-uptake (Hup) system that re-oxidizes the hydrogen, a process that is thought to increase the efficiency of nitrogen fixation. This is, for instance, the case in *B. japonicum*, where the -24/-12-type promoters controlling these genes were confirmed in the screening (see Additional data files, pages 7,8). In *B. japonicum*, HoxA-RpoN is required for *hupSL* expression under free-living conditions and FixK$_2$ under symbiotic conditions, whereas in *Rhizobium leguminosarum*, expression of *hup* genes is only induced in bacteroids and depends on NifA-RpoN [21,43,44].

As well as these previously well-characterized RpoN-dependent functions, several other groups of genes were found in the *in silico* screening. Potential target genes preceded by a possible -24/-12-type promoter belong to the following functional categories: nodulation; transport; detoxification; gene regulation; cell envelope; amino-acid biosynthesis; energy and DNA metabolism; translation; integration and recombination; and various others (see Additional data files, pages 8-28).

An interesting observation is the large number of hypothetical CDSs preceded by a possible RpoN-binding site (see Additional data files, pages 21-28). Of the 593 gene products identified in the screening, 22% have no function attributed, but do have homologs in other species. Proteins encoded by orphan genes, unique to one species, make up 9% of the total. Some of these hypothetical CDSs having a possible -24/-12-type promoter were picked up in several genomes: *atu1754-smc01000-mll0186*, *smc00999-mll0185*, *mll9007-y4nG*, *mll9408-y4nH*, *orf196-y4vH*, *orf538-y4vI*, *orf354-y4vJ*, *id79-mlr5912-y4vQ-msi288-yp001*, *yp003-y4wP-yp021-yp099*, *mlr5819-y4wI*, *mll5852-msi351*, *mlr5886-msi324*, *mlr5914-msi286* and *msr5928-msi276*. The majority of these hypothetical genes in *Rhizobium* sp. NGR234 are highly expressed in bacteroids [42], suggesting a possible involvement of the putative gene products in the symbiosis.

Another interesting class of possible RpoN-regulated genes encodes proteins catalyzing detoxification reactions (see Additional data files, pages 12,13). During symbiosis, rhizobia are confronted with toxic compounds and oxidative stress in the form of reactive oxygen species [45]. Effective

protection against these harmful agents would thus enhance the symbiotic performance of the rhizobia. Two good candidate proteins to fulfill such a function were identified in this screening: a peroxiredoxin and a cytochrome P450. Peroxiredoxins reduce peroxides to the alcohol and water. The gene encoding the symbiosis-specific peroxiredoxin SpxA is highly expressed in bacteroids of several rhizobial species in a NifA-RpoN-dependent way [13,40,42]. Cytochrome P450s catalyze the monoxygenation of a broad spectrum of substrates, including toxic compounds. Several genes encoding cytochrome P450s have a possible RpoN-binding site in their promoter region. One subclass of these P450s is conserved among several rhizobial species and is highly expressed in bacteroids [27,42]. Expression of the *Rhizobium* sp. BR816 homolog depends on NifA-RpoN and another, so far unidentified, factor [27]. Furthermore, two matches were found upstream of *B. japonicum nrgA* and *nrgBC*. The expression of these genes, possibly involved in detoxification reactions, was shown to be controlled by NifA and RpoN [46].

Transport systems have an important role at different stages of the rhizobium-legume symbiosis [41,47-49]. Many transporter-encoding gene loci are preceded by a possible -24/-12-type promoter (see Additional data files, pages 9-12). The majority of these transporters are of the ABC type, which is consistent with the fact that this type is the most abundant transporter in *Rhizobiaceae* family [33]. The RpoN-regulated transporters might be involved in symbiotic (as, for instance, DctA) or non-symbiotic (in *A. tumefaciens*) processes.

Many regulatory genes have a possible RpoN-binding site in their promoter region (see Additional data files, pages 13,14). The RpoN regulon is known to expand its working range by controlling the expression of other regulatory genes [6,7]. The induction of the *y4qH* gene (preceded by a possible RpoN-binding site and encoding a LuxR-type transcriptional regulator) in bacteroids of *Rhizobium* sp. NGR234 illustrates how RpoN might indirectly control the expression of genes involved in symbiosis [42]. Other examples include RpoN dependency of *R. leguminosarum fnrN*, or the autoregulation of *R. etli rpoN$_2$* [13,14]. The latter form of control also appears to be conserved in *M. loti* MAFF303099 and R7A, where a possible -24/-12-type promoter was identified upstream of the second copy of *rpoN* (see Additional data files, page 7).

The gene products of the rhizobial nodulation genes (*nol*, *nod*, *noe*) synthesize and export a class of signal molecules, the so-called Nod factors, involved in the early steps of the symbiotic interaction with leguminous plants. Although the transcription of most nodulation genes depends on the NodD regulator protein in the presence of a suitable plant-root-derived flavonoid, several reports mention an RpoN dependency (to some extent) of some of the nodulation genes ([22-24,26,50] and see Additional data files, pages

8,9). However, the screening did not reveal any matches upstream of rhizobial nodulation genes. Because of the high stringency of the screening, several promoters, such as those of *Rhizobium* sp. NGR234 *nodD$_2$* and *S. meliloti nodD$_3$*, may have been missed. It thus seems that less conserved -24/-12-type promoters are functional to some degree. On the other hand, such poorly conserved promoters may be the exception rather than the rule. Nevertheless, the -24/-12-type promoter of the *B. japonicum nfeC* gene (involved in nodule formation) was confirmed in the screening.

The other functional categories (see Additional data files, page 14-21) include, among others, cell envelope, amino-acid biosynthesis, energy metabolism, DNA metabolism, translation, and integration and recombination. Unlike the functions discussed above, these categories comprise many different proteins from different bacterial species. It appears that the RpoN dependency of the encoding genes in these categories is not conserved in the *Rhizobiales*. This might reflect a species-specific recruitment of RpoN for gene transcription.

Signals controlling RpoN-dependent gene expression in the *Rhizobiales*

What signals control the activity of EBPs in species of the *Rhizobiales*? An *in silico* screening for EBPs in the proteomes of members of the *Rhizobiales* (see Materials and methods) revealed that the number of candidate EBPs is proportional to the genome size (Table 2). *M. loti*, *S. meliloti*, *A. tumefaciens* and *B. melitensis* all have roughly 1 EBP per million bases (Mb) (8, 7, 6 and 3, respectively), a number differing from a previous estimate (13, 12 and 9, respectively; the *B. melitensis* genome was not published at the time) [33]. However, our predicted EBPs scored significantly better (E-value 10^{-25} or less) with the Pfam Sigma54_activat domain HMM motif than the best of the rejected proteins (E-value 10^{-5} or greater), giving confidence to our prediction. Each of the candidate EBPs was subjected to a BLASTP search (default settings) against the NCBI protein databases and classified according to homology (Table 3). A similar analysis of *Pseudomonas aeruginosa*, *Salmonella typhimurium* and *E. coli* proteomes revealed that these bacteria have a higher number of possible EBPs (22, 15 and 13, respectively) than the species of the *Rhizobiales* (data not shown).

All four genomes screened have NtrC and NtrX orthologs (Table 3), which are both nitrogen-responsive EBPs [9]. Rhizobial NtrC-RpoN-regulated genes are downregulated early in bacteroid differentiation [51,52]. This downregulation seems to be necessary for an effective symbiosis as ectopic expression of *R. etli amtB* alters the bacteroid differentiation process and the ability to enter the host cells [53].

M. loti, *S. meliloti* and *A. tumefaciens* all have two DctD EBP paralogs (Table 3). DctD is activated by the DctB sensor protein, which senses C$_4$-dicarboxylates in collaboration with the DctA C$_4$-dicarboxylate transporter. The second *dctD*

Table 3

Enhancer binding proteins (EBPs) of different members of the *Rhizobiales*

	Re	NGR	Bj	R7A	Ml	Sm	At	Bm
NtrC					Mlr0389 (14021374)	SMc01043 (15074392)	Atu1446 (17739866)	Bmel0866 (17987149)
NtrX					Mlr0400 (14021376)	SMc01045 (15074394)	Atu1448 (17739869)	Bmel0868 (17987151)
DctD-1				Msi357 (20804186)	Mlr5842 (14025622)	SMb20613 (15141413)	Atu3296 (17741681)	
DctD-2					Mlr7239 (14026791)	SMb21200 (15140751)	Atu0114 (17738434)	
NifA	RE1SP0000735 (21492735)	Y4uN (16519979)	ID736 (12620732)	Msi346 (20804175) Msi361 (20804190)	Mll5857 (14025634) Mll5837 (14025617)	SMa0815 (14523538)		
EBP-1					Mll3001 (14023386)	SMc04011 (15075819)	Atu3764 (17742191)	Bmell0011 (17988355)
EBP-2						SMb20102 (15139974)	Atu4741 (17743247)	
EBP-3		Y4pA (16519894)	ID407 (12620599)		Mll9384 (14028438)			

EBPs were predicted using HMMER 2.2g (see Materials and methods). The species acronyms are as in Table 2. Orthologs (having at least 30% amino-acid identity over the full length of the polypeptide) are situated on the same row (GenBank GI number). EBP-n: uncharacterized EBPs.

copy might encode the so-called alternative symbiotic activator, ensuring *dctA* expression in *S. meliloti* and *R. leguminosarum dctD* mutants during symbiosis [16].

All nitrogen-fixing species possess a NifA homolog - *M. loti* MAFF303099 and R7A even have two (Table 3). Rhizobial NifA proteins are activated by the low oxygen tension in the nodules, which is the key signal controlling expression of the nitrogen fixation genes and a prerequisite for nitrogen fixation [54]. NifA activates RpoN-dependent transcription of many rhizobial genes, including nitrogen fixation genes (see Additional data files). For the NifA-RpoN-dependent expression of some genes, microaerobiosis alone is not sufficient to equal the expression levels observed in bacteroids (see Additional data files, pages 2,4,6,7,13,24) [27,39]. An additional signal appears to control the optimal functioning of NifA.

The presence of additional uncharacterized EBPs suggests that other signals control RpoN-dependent gene expression in the *Rhizobiales* (Table 3). These signals could control expression of genes involved in symbiosis (*M. loti* and *S. meliloti*) or other, so far uncharacterized, functions (*A. tumefaciens* and *B. melitensis*). One common signal might control the activity of EBP-1, present in the four genomes (Table 3).

Comparison between symbiotic and non-symbiotic *Rhizobiales*
Globally, a higher number of possible RpoN-binding sites per megabases of genome is observed in symbiotic members of the *Rhizobiales* than in non-symbiotic ones (Table 2). This ratio is even higher in the symbiotic regions and that might reflect the recruitment of NifA-RpoN-dependent promoters, which are highly upregulated in bacteroids, for late symbiotic functions. RpoN is not essential for bacterial survival, but controls genes whose products have a wide range of different functions [6]. Apparently, there has been a selection in bacteroids for the NifA-RpoN-type promoter, not only for nitrogen fixation but for other symbiotic functions as well. This is illustrated by the high number of non-*nif*/*fix* genes on pNGR234*a* preceded by a -24/-12-type promoter and preferentially expressed in bacteroids (see Additional data files) [42]. The symbiotic island of *M. loti* R7A is about 109 kb smaller than that of MAFF303099 and the two areas share only a conserved backbone of 248 kb [30]. This distinction seems not to be reflected by the RpoN-regulon, as the majority of possible RpoN-dependent loci (75%) on the two islands is conserved (see Additional data files).

Although an *A. tumefaciens rpoN* mutant is not affected in tumorigenesis, expression of *vir* genes, chemotaxis or flagellum formation, it does not grow on nitrate as sole nitrogen source nor on C_4-dicarboxylates as sole carbon source, and grows only very poorly on arginine as sole nitrogen source [35]. We were able to identify possible RpoN-binding sites in the promoter regions of the *dctA* gene, the *glnK-amtB*, *glnBA*, *glnII* and *nrtABC-nirBD* loci (see Additional data files, pages 1,2,10,12,16), providing genetic evidence for the previously observed phenotypes. As could be expected, no

virulence, tumorigenesis, chemotaxis or flagellar genes were picked up in screening, although a good match was found upstream of *motB*, which encodes a flagellar motor protein. Strikingly, in *Agrobacterium* the total number of gene products whose genes are preceded by a possible -24/-12-type promoter is approximately equal to that of *Sinorhizobium* (Table 2). *S. meliloti* and *A. tumefaciens* are closely related at the level of protein identity as well as nucleotide colinearity and conservation of gene order [33]. This relatedness appears to be somewhat conserved at the regulatory level, as *S. meliloti* and *A. tumefaciens* share seven predicted RpoN-binding sites upstream of the same gene loci: *dctA*, *glnK-amtB*, *glnBA*, *glnII*, *nrtABC*, *atu1754-smc01000* and *atu4739-smb20103* (see Additional data files, pages 1,2,10,12,18,21,24). The relatively high number of positive matches in *A. tumefaciens* and the lack of tumorigenesis-related phenotype in the *rpoN* mutant hints to other, so far unidentified, RpoN-controlled functions. Indicative of this is the number (22) of possible RpoN-binding sites upstream of hypothetical genes in *A. tumefaciens* (see Additional data files, pages 21,26). Of the predicted proteins of *A. tumefaciens*, 16% are not present in *M. loti* or *S. meliloti* and 89% of these are hypothetical gene products [33], leaving room for genetic diversity between rhizobia and *A. tumefaciens*.

B. melitensis is an intracellular pathogen of humans and animals, and is closely related to rhizobia and *A. tumefaciens* [55]. Only 12 matches resulted from the *in silico* analysis for possible RpoN-binding sites (Table 2), two of which were found in the upstream region of *amtB* and *glnBA* (see Additional data files, pages 1,2). At present, no report of a *B. melitensis rpoN* mutant is available. A good candidate function for RpoN to regulate in *B. melitensis* would be pathogenesis, as is the case for several other pathogens [6]. Six possible -24/-12-type promoters were found in the *B. melitensis* genome upstream of hypothetical CDSs (see Additional data files, pages 21,22,26).

E. coli was estimated to have around 30 -24/-12-type promoters, controlling mainly functions related to nitrogen metabolism [7]. Although there is some overlap of RpoN-controlled functions between *E. coli* and members of the *Rhizobiales* (for example, ammonium assimilation and hydrogen oxidation), RpoN seems to control a more diverse set of functions in the latter. This appears to somewhat contradict the observation that *E. coli* possesses a higher number of EBPs (see 'Signals controlling RpoN-dependent gene expression in the *Rhizobiales*').

Additional control mechanisms involving RpoN
Expression studies of *rpoN* in wild-type and *rpoN*⁻ backgrounds revealed that RpoN negatively controls its own expression in *B. japonicum*, *Acinetobacter calcoaceticus*, *R. etli* and *R. leguminosarum* bv. *viciae* [14,56-58]. The presence of a putative oppositely oriented RpoN-binding site on the template strand of *rpoN* promoter regions, overlapping

with the -10 promoter region and the transcription start site, was proposed [58]. It was stated that, as RpoN is able to bind to -24/-12-type promoters in the absence of core-RNAP, the negative autoregulation of the *rpoN* genes occurs by direct interference of RpoN with the binding of σ⁷⁰-holo-RNAP complex to the -10 promoter region of the *rpoN* gene. Site-directed mutagenesis of the highly conserved GG to TT in the putative RpoN-binding site of the *rpoN* promoter relieved the negative autoregulation, giving strong support to the above hypothesis [14].

The *in silico* screening revealed the presence of oppositely oriented possible RpoN-binding sites upstream of the *rpoN* genes of *M. loti*, *S. meliloti*, *A. tumefaciens* and *B. melitensis* (Table 2). A comparison with the *rpoN* coding sequences of *R. etli*, *R. leguminosarum*, *Rhizobium* sp. NGR234, *B. japonicum*, *M. loti* and *S. meliloti* revealed that the *rpoN* genes of *A. tumefaciens* and *B. melitensis* were incorrectly annotated. Their coding sequences should be 63 bp longer and shorter, respectively. An alignment of the *rpoN* promoter regions of these species shows the strong conservation of these promoters (Figure 1). In addition, the screening of the lower strand of the intergenic sequences revealed a slightly lower number of matches than that of the correctly oriented RpoN-binding sites (Table 2). It is thus not unconceivable that RpoN could alter the expression of these genes in a way similar to its own autoregulation, that is, by interference with the binding of the holo-RNAP or a regulatory protein to the promoter. This would significantly broaden the working domain of RpoN.

Conclusions
A highly specific *in silico* screening method was applied to predict members of the RpoN-regulon in eight different species of the *Rhizobiales*. The matches obtained were individually checked and classified according to protein function, resulting in a highly annotated and manually curated dataset. This dataset was complemented with available literature data on members of the RpoN-regulon in *Rhizobiales*. In addition, a screening was carried out to identify possible EBPs controlling the expression of RpoN-dependent genes. Together, these data serve as a source of exhaustive information on the (possible) roles of RpoN in symbiotic and non-symbiotic processes.

RpoN-binding sites were found upstream of genes involved in common RpoN-dependent functions, such as assimilation of ammonium and uptake of C_4-dicarboxylic acids. The symbiotic members of the *Rhizobiales* seem to have recruited RpoN for the expression of nitrogen fixation and other symbiotic genes. This is illustrated by the high number of possible RpoN-binding sites in the symbiotic regions of these bacteria. Other RpoN-dependent symbiotic functions might include detoxification and transport or might be controlled indirectly through other regulatory proteins. Whereas an

Figure 1

Manual alignment of *rpoN* promoter sequences. At (*Agrobacterium tumefaciens*, GI: 17738659); Bj (*Bradyrhizobium japonicum*, GI: 152137); Bm (*Brucella melitensis*, GI: 17983821); Ml (*Mesorhizobium loti*, GI: 14023393); NGR (*Rhizobium* sp. NGR234, GI: 152431); Re (*Rhizobium etli*, GI: 1046228), Rl (*Rhizobium leguminosarum* bv. *viciae*, GI: 5759116), Sm (*Sinorhizobium meliloti*, GI: 152389). Upper line, -35 and -10 consensus sequences of *Escherichia coli* and the transcription start site (*) as determined for *S. meliloti* [69]. Lower line, consensus sequence of -24/-12-type promoter [1]. Nucleotides are shaded in black (100% conserved) or gray (75% conserved). The numbers represent the distance (in bp) from the end of the alignment to the start codon of the downstream *rpoN* gene.

A. tumefaciens rpoN mutant only displays common RpoN-dependent phenotypes, the relatively high number of possible RpoN-binding sites present in its genome points to several other, yet unidentified, RpoN-dependent functions. So far, no reports are present on RpoN-dependent phenotypes in *B. melitensis*. This animal pathogen has a significantly lower number of possible RpoN-binding sites than the other members of the *Rhizobiales*. *B. melitensis* RpoN might be required for infection of the host organism, as is the case for other pathogens. Furthermore, the species screened seem to have recruited RpoN independently, in a species-specific manner, for the transcription of different gene sets. The high percentage of hypothetical conserved and non-conserved CDSs preceded by a possible RpoN-binding site opens up ample opportunities for future research. Several uncharacterized EBPs were identified besides the 'classic' EBPs such as NtrC, NtrX, NifA and DctD. This implies that signals other than nitrogen, oxygen and C_4-dicarboxylates control the expression of RpoN-dependent genes in species of the *Rhizobiales*. Identification of these signals will give better insight into yet uncharacterized RpoN-dependent functions. Finally, a similar number of possible RpoN-binding sites were found on the lower strand of the upstream intergenic sequences. RpoN might thus significantly extend its working domain by blocking the binding of transcription regulatory factors. Such is the case, for instance, in the negative autoregulation of *rpoN*.

Although much consideration was given in our analysis to the design and, to some extent, the experimental validation of the approach, experimental confirmation will ultimately be required to establish the biological meaning of the predicted -24/-12-type promoters, as is the case with all computer predictions.

In conclusion, a highly efficient method was applied to predict the RpoN-regulon of different members of the *Rhizobiales*

group. The same approach might be used for the prediction of RpoN-dependent genes in other bacterial species.

Materials and methods
Retrieval of intergenic sequences
Complete DNA sequences from *A. tumefaciens* C58 (circular chromosome: NC_003304, linear chromosome NC_003305, plasmid AT: NC_003306, plasmid Ti: NC_003308), *B. japonicum* USDA110 (symbiotic chromosomal region: AF322012 and AF322013), *B. melitensis* 16M (chromosome I: NC_003317, chromosome II: NC_003318), *M. loti* MAFF303099 (chromosome: NC_002678, plasmid pMla: NC_002679, plasmid pMlb: NC_002682), *M. loti* R7A (symbiotic island: AL672111), *Rhizobium* sp. NGR234 (plasmid pNGR234*a*: NC_000914), *R. etli* CFN42 (plasmid p42d: NC_004041) and *S. meliloti* 1021 (chromosome: NC_003047, plasmid pSymA: NC_003037, plasmid pSymB: NC_003078) were extracted from GenBank. Intergenic sequences were identified by automatically parsing the corresponding GenBank files using the modules of INCLUsive [59,60]. An intergenic region is defined as the non-coding region between two genes. Intergenic regions smaller than 10 nucleotides were discarded, as the corresponding genes are likely to belong to an operon.

Prediction of possible RpoN-binding sites
The intergenic regions were screened with the PATSER module of the Regulatory Sequence Analysis Tools (RSAT) [61-63] for the presence of the -24/-12 promoter consensus sequence. PATSER scores N-mers (in this case 16-mers) from a sequence against a given weight matrix. A set of 186 RpoN-dependent promoters from different bacterial species [1] was used to generate the weight matrix (Table 1). Initially, this matrix was trained against 67 -24/-12-type promoters with a known transcriptional start site. From the

distribution of these scores (Figure 2), it was decided to retain all matches with a score higher than or equal to the fifth percentile (8.9). PATSER was run with the GC content of the intergenic sequences as a measure for the *a priori* probabilities of the nucleotides. The intergenic GC content differs markedly from the total GC content of the genomes (Table 2).

Prediction of possible EBPs

An estimate of the number of possible EBPs in the respective proteomes was obtained by looking for the presence of the Pfam Sigma54_activat domain [64]. Sigma54_activat is a conserved domain present in EBPs that is involved in the ATP-dependent interaction with RpoN. The predicted proteome sequences of *A. tumefaciens*, *B. melitensis*, *S. meliloti* and *M. loti* MAFF303099 were downloaded from Proteome Analysis @ EBI [65] and the protein sequences of the symbiotic regions of *B. japonicum*, *M. loti* R7A, *Rhizobium* sp. NGR234 and *R. etli* were obtained from the NCBI protein database [66]. The protein sequences were subsequently queried with the PF00158 HMM motif using HMMER 2.2g [67]. Matches with an E-value less than or equal to 10^{-25} were retained.

Additional data files

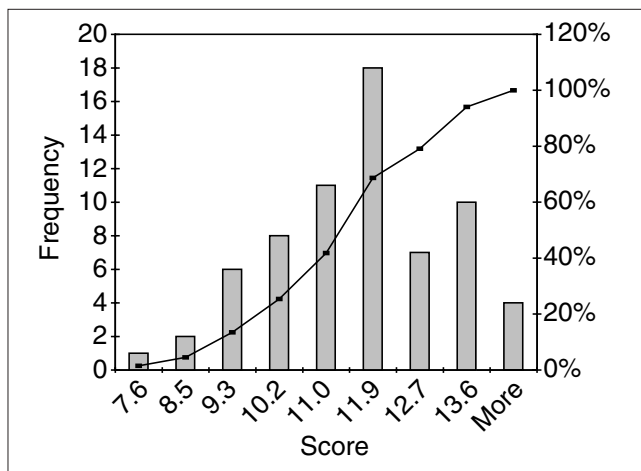An additional data file listing all genes and proteins included in this analysis is available with the online version of the paper. Each protein is accompanied by a functional description, the species from which it comes, GenBank GI number and transcription unit; additional references for each protein are also included in the file. Information on the gene's regulation is provided where available. The data were obtained from an *in silico* analysis (see Materials and methods) and the literature [12-14,17-27,29-34,39-44,46,50,51,56,58,70-88].

References

1. Barrios H, Valderrama B, Morett E: **Compilation and analysis of sigma 54-dependent promoter sequences.** *Nucleic Acids Res* 1999, **27:**4305-4313.
2. Buck M, Gallegos MT, Studholme DJ, Guo YL, Gralla JD: **The bacterial enhancer-dependent sigma(54) (sigma(N)) transcription factor.** *J Bacteriol* 2000, **182:**4129-4136.
3. Kustu S, Santero E, Keener J, Popham D, Weiss D: **Expression of sigma 54 (*ntrA*)-dependent genes is probably united by a common mechanism.** *Microbiol Rev* 1989, **53:**367-376.
4. Morett E, Segovia L: **The sigma 54 bacterial enhancer-binding protein family: mechanism of action and phylogenetic relationship of their functional domains.** *J Bacteriol* 1993, **175:**6067-6074.
5. Shingler V: **Signal sensing by sigma 54-dependent regulators: derepression as a control mechanism.** *Mol Microbiol* 1996, **19:**409-416.
6. Studholme DJ, Buck M: **The biology of enhancer-dependent transcriptional regulation in bacteria: insights from genome sequences.** *FEMS Microbiol Lett* 2000, **186:**1-9.
7. Reitzer L, Schneider BL: **Metabolic context and possible physiological themes of sigma(54)-dependent genes in *Escherichia coli*.** *Microbiol Mol Biol Rev* 2001, **65:**422-444.
8. Keseler IM, Kaiser D: **Sigma(54), a vital protein for *Myxococcus xanthus*.** *Proc Natl Acad Sci USA* 1997, **94:**1979-1984.
9. Arcondéguy T, Jack R, Merrick M: **P-II signal transduction proteins, pivotal players in microbial nitrogen control.** *Microbiol Mol Biol Rev* 2001, **65:**80-105.
10. Ronson CW, Nixon BT, Albright LM, Ausubel FM: ***Rhizobium meliloti ntrA (rpoN)* gene is required for diverse metabolic functions.** *J Bacteriol* 1987, **169:**2424-2431.
11. van Slooten JC, Cervantes E, Broughton WJ, Wong CH, Stanley J: **Sequence and analysis of the *rpoN* sigma factor gene of *Rhizobium* sp. strain NGR234, a primary coregulator of symbiosis.** *J Bacteriol* 1990, **172:**5563-5574.
12. Stigter J, Schneider M, de Bruijn FJ: ***Azorhizobium caulinodans* nitrogen fixation (*nif/fix*) gene regulation: mutagenesis of the *nifA* -24/-12 promoter element, characterization of a *ntrA(rpoN)* gene, and derivation of a model.** *Mol Plant-Microbe Interact* 1993, **6:**238-252.
13. Michiels J, Moris M, Dombrecht B, Verreth C, Vanderleyden J: **Differential regulation of *Rhizobium etli rpoN2* gene expression during symbiosis and free-living growth.** *J Bacteriol* 1998, **180:**3620-3628.
14. Clark SR, Oresnik IJ, Hynes MF: **RpoN of *Rhizobium leguminosarum* bv. *viciae* strain VF39SM plays a central role in FnrN-dependent microaerobic regulation of genes involved in nitrogen fixation.** *Mol Gen Genet* 2001, **264:**623-633.
15. Fischer HM: **Genetic regulation of nitrogen fixation in rhizobia.** *Microbiol Rev* 1994, **58:**352-386.
16. Kaminski PA, Batut J, Boistard P: **A survey of symbiotic nitrogen fixation by rhizobia.** In *The Rhizobiaceae, Molecular Biology of Model Plant-Associated Bacteria*. Edited by Spaink HP, Kondorosi A, Hooykaas PJJ. Dordrecht: Kluwer; 1998: 431-460.
17. Hawkins FK, Johnston AW: **Transcription of a *Rhizobium leguminosarum* biovar *phaseoli* gene needed for melanin synthesis is**

Figure 2
Distribution of PATSER scores (see Material and methods) of 67 -24/-12-type promoters with mapped transcriptional start sites (GenBank GI number: 2979503, 141885, 141892, 38664, 38679, 1769418, 142336, 142378, 142326, 39977, 550310, 408911, 39532, 39516, 39526, 152106, 152283, 152315, 39548, 312974, 12620419, 152100, 152280, 152317, 2316081, 144194, 896457, 144223, 262651, 3493239, 7208421, 262651, 145911, 556890, 41774, 146158, 1004098, 41568, 42538, 149241, 43802, 149252, 149256, 149246, 43857, 43864, 149273, 149275, 150095, 950651, 150993, 490170, 6492432, 151643, 6636054, 46254, 152305, 152230, 340664, 46285, 46324, 46324, 550144, 550144, 664946, 453435, 1649033). Cumulative percentage: black line.

activated by NifA of *Rhizobium* and *Klebsiella pneumoniae*. *Mol Microbiol* 1988, **2:**331-337.

18. Soto MJ, Zorzano A, Mercado-Blanco J, Lepek V, Olivares J, Toro N: **Nucleotide sequence and characterization of *Rhizobium meliloti* nodulation competitiveness genes *nfe*.** *J Mol Biol* 1993, **229:**570-576.

19. Chun JY, Sexton GL, Roth LE, Stacey G: **Identification and characterization of a novel *Bradyrhizobium japonicum* gene involved in host-specific nitrogen fixation.** *J Bacteriol* 1994, **176:**6717-6729.

20. Chun JY, Stacey G: **A *Bradyrhizobium japonicum* gene essential for nodulation competitiveness is differentially regulated from two promoters.** *Mol Plant Microbe Interact* 1994, **7:**248-255.

21. Brito B, Martinez M, Fernandez D, Rey L, Cabrera E, Palacios JM, Imperial J, Ruiz-Argueso T: **Hydrogenase genes from *Rhizobium leguminosarum* bv. *viciae* are controlled by the nitrogen fixation regulatory protein NifA.** *Proc Natl Acad Sci USA* 1997, **94:**6019-6024.

22. Fellay R, Hanin M, Montorzi G, Frey J, Freiberg C, Golinowski W, Staehelin C, Broughton WJ, Jabbouri S: ***nodD2* of *Rhizobium* sp NGR234 is involved in the repression of the *nodABC* operon.** *Mol Microbiol* 1998, **27:**1039-1050.

23. Vlassak KM, de Wilde P, Snoeck C, Luyten E, van Rhijn P, Vanderleyden J: **The *Rhizobium* sp. BR816 *nodD3* gene is regulated by a transcriptional regulator of the AraC/XylS family.** *Mol Gen Genet* 1998, **258:**558-561.

24. Dusha I, Austin S, Dixon R: **The upstream region of the *nodD3* gene of *Sinorhizobium meliloti* carries enhancer sequences for the transcriptional activator NtrC.** *FEMS Microbiol Lett* 1999, **179:**491-499.

25. Garcia-Rodríguez FM, Toro N: ***Sinorhizobium meliloti nfe* (nodulation formation efficiency) genes exhibit temporal and spatial expression patterns similar to those of genes involved in symbiotic nitrogen fixation.** *Mol Plant Microbe Interact* 2000, **13:**583-591.

26. Gao M, D'Haeze W, De Rycke R, Holsters M: **Dual control of the *nodA* operon of *Azorhizobium caulinodans* ORS571 by a nod box and a NifA-sigma54-type promoter.** *Mol Genet Genomics* 2001, **265:**1050-1059.

27. Luyten E, Swinnen E, Vlassak K, Verreth C, Dombrecht B, Vanderleyden J: **Analysis of a symbiosis-specific cytochrome P450 homolog in *Rhizobium* sp. BR816.** *Mol Plant Microbe Interact* 2001, **14:**918-924.

28. Freiberg C, Fellay R, Bairoch A, Broughton WJ, Rosenthal A, Perret X: **Molecular basis of symbiosis between *Rhizobium* and legumes.** *Nature* 1997, **387:**394-401.

29. Göttfert M, Rothlisberger S, Kundig C, Beck C, Marty R, Hennecke H: **Potential symbiosis-specific genes uncovered by sequencing a 410-kilobase DNA region of the *Bradyrhizobium japonicum* chromosome.** *J Bacteriol* 2001, **183:**1405-1412.

30. Sullivan JT, Trzebiatowski JR, Cruickshank RW, Gouzy J, Brown SD, Elliot RM, Fleetwood DJ, McCallum NG, Rossbach U, Stuart GS, *et al.*: **Comparative sequence analysis of the symbiosis island of *Mesorhizobium loti* strain R7A.** *J Bacteriol* 2002, **184:**3086-3095.

31. Kaneko T, Nakamura Y, Sato S, Asamizu E, Kato T, Sasamoto S, Watanabe A, Idesawa K, Ishikawa A, Kawashima K, *et al.*: **Complete genome structure of the nitrogen-fixing symbiotic bacterium *Mesorhizobium loti*.** *DNA Res* 2000, **7:**331-338.

32. Galibert F, Finan TM, Long SR, Puhler A, Abola P, Ampe F, Barloy-Hubler F, Barnett MJ, Becker A, Boistard P, *et al.*: **The composite genome of the legume symbiont *Sinorhizobium meliloti*.** *Science* 2001, **293:**668-672.

33. Wood DW, Setubal JC, Kaul R, Monks DE, Kitajima JP, Okura VK, Zhou Y, Chen L, Wood GE, Almeida NF Jr, *et al.*: **The genome of the natural genetic engineer *Agrobacterium tumefaciens* C58.** *Science* 2001, **294:**2317-2323.

34. DelVecchio VG, Kapatral V, Redkar RJ, Patra G, Mujer C, Los T, Ivanova N, Anderson I, Bhattacharyya A, Lykidis A, *et al.*: **The genome sequence of the facultative intracellular pathogen *Brucella melitensis*.** *Proc Natl Acad Sci USA* 2002, **99:**443-448.

35. Wu ZL, Charles TC, Wang H, Nester EW: **The *ntrA* gene of *Agrobacterium tumefaciens*: identification, cloning, and phenotype of a site-directed mutant.** *J Bacteriol* 1992, **174:**2720-2723.

36. Benítez-Bellón E, Moreno-Hagelsieb G, Collado-Vides J: **Evaluation of thresholds for the detection of binding sites for regulatory proteins in *Escherichia coli* K12 DNA.** *Genome Biol* 2002, **3:**research0013.1-0013.16.

37. Gubler M, Hennecke H: **Regulation of the *fixA* gene and *fixBC* operon in *Bradyrhizobium japonicum*.** *J Bacteriol* 1988, **170:**1205-1214.

38. Michiels J, D'hooghe I, Verretch C, Pelemans H, Vanderleyden J: **Characterization of the *Rhizobium leguminosarum* biovar *phaseoli nifA* gene, a positive regulator of *nif* gene expression.** *Arch Microbiol* 1994, **161:**404-408.

39. Dombrecht B, Tesfay MZ, Verreth C, Heusdens C, Nápoles MC, Vanderleyden J, Michiels J: **The *Rhizobium etli* gene *iscN* is highly expressed in bacteroids and required for nitrogen fixation.** *Mol Gen Genomics*, 2002, **267:**820-828.

40. Dombrecht B: *The complex regulation and role of the* Rhizobium etli *RpoN regulon in the symbiotic interaction with the common bean plant (*Phaseolus vulgaris *L.)*. Leuven: Katholieke Universiteit Leuven; 2001.

41. Michiels J, Dirix G, Vanderleyden J, Xi CW: **Processing and export of peptide pheromones and bacteriocins in gram-negative bacteria.** *Trends Microbiol* 2001, **9:**164-168.

42. Perret X, Freiberg C, Rosenthal A, Broughton WJ, Fellay R: **High-resolution transcriptional analysis of the symbiotic plasmid of *Rhizobium* sp. NGR234.** *Mol Microbiol* 1999, **32:**415-425.

43. Van Soom C, de Wilde P, Vanderleyden J: **HoxA is a transcriptional regulator for expression of the *hup* structural genes in free-living *Bradyrhizobium japonicum*.** *Mol Microbiol* 1997, **23:**967-977.

44. Durmowicz MC, Maier RJ: **The FixK(2) protein is involved in regulation of symbiotic hydrogenase expression in *Bradyrhizobium japonicum*.** *J Bacteriol* 1998, **180:**3253-3256.

45. Becana M, Dalton DA, Moran JF, Iturbe-Ormaetxe I, Matamoros MA, Rubio MC: **Reactive oxygen species and antioxidants in legume nodules.** *Physiol Plant* 2000, **109:** 372-381.

46. Nienaber A, Huber A, Gottfert M, Hennecke H, Fischer HM: **Three new NifA-regulated genes in the *Bradyrhizobium japonicum* symbiotic gene region discovered by competitive DNA-RNA hybridization.** *J Bacteriol* 2000, **182:**1472-1480.

47. Downie JA: **Functions of rhizobial nodulation genes.** In *The Rhizobiaceae, Molecular Biology of Model Plant-Associated Bacteria*. Edited by Spaink HP, Kondorosi A, Hooykaas PJJ. Dordrecht: Kluwer; 1998: 403-416.

48. Rosenblueth M, Hynes MF, Martinez-Romero E: ***Rhizobium tropici teu* genes involved in specific uptake of *Phaseolus vulgaris* bean exudate compounds.** *Mol Gen Genet* 1998, **258:**587-598.

49. Marie C, Broughton WJ, Deakin WJ: ***Rhizobium* type III secretion systems: legume charmers or alarmers?** *Curr Opin Plant Biol* 2001, **4:**336-342.

50. Snoeck C: *Host specificity determinants of* Sinorhizobium *sp. B816 for early signaling during symbiotic interactions*. Leuven: Katholieke Universiteit Leuven; 2001.

51. Taté R, Riccio A, Merrick M, Patriarca EJ: **The *Rhizobium etli amtB* gene coding for an NH_4^+ transporter is down-regulated early during bacteroid differentiation.** *Mol Plant Microbe Interact* 1998, **11:**188-198.

52. Ercolano E, Mirabella R, Merrick M, Chiurazzi M: **The *Rhizobium leguminosarum glnB* gene is down-regulated during symbiosis.** *Mol Gen Genet* 2001, **264:**555-564.

53. Taté R, Cermola M, Riccio A, Iaccarino M, Merrick M, Favre R, Patriarca EJ: **Ectopic expression of the *Rhizobium etli amtB* gene affects the symbiosome differentiation process and nodule development.** *Mol Plant Microbe Interact* 1999, **12:**515-525.

54. Soupène E, Foussard M, Boistard P, Truchet G, Batut J: **Oxygen as a key developmental regulator of *Rhizobium meliloti* N_2-fixation gene expression within the alfalfa root nodule.** *Proc Natl Acad Sci USA* 1995, **92:**3759-3763.

55. Gándara B, Merino AL, Rogel MA, Martinez-Romero E: **Limited genetic diversity of *Brucella* spp.** *J Clin Microbiol* 2001, **39:**235-240.

56. Kullik I, Fritsche S, Knobel H, Sanjuan J, Hennecke H, Fischer HM: ***Bradyrhizobium japonicum* has two differentially regulated, functional homologs of the sigma 54 gene (*rpoN*).** *J Bacteriol* 1991, **173:**1125-1138.

57. Ehrt S, Ornston LN, Hillen W: **RpoN (sigma 54) is required for conversion of phenol to catechol in *Acinetobacter calcoaceticus*.** *J Bacteriol* 1994, **176:**3493-3499.

58. Michiels J, Van Soom T, D'hooghe I, Dombrecht B, Benhassine T, de Wilde P, Vanderleyden J: **The *Rhizobium etli rpoN* locus: DNA sequence analysis and phenotypical characterization of *rpoN*, *ptsN*, and *ptsA* mutants.** *J Bacteriol* 1998, **180:**1729-1740.

59. Thijs G, Moreau Y, De Smet F, Mathys J, Lescot M, Rombauts S, Rouze P, De Moor B, Marchal K: **INCLUSive: Integrated Clustering,**

Upstream sequence retrieval and motif Sampling. *Bioinformatics* 2002, **18**:331-332.

60. **Biol@Sista: Software**
[http://www.esat.kuleuven.ac.be/~dna/Biol/Software.html]

61. Hertz GZ, Hartzell III GW, Stormo GD: **Identification of consensus patterns in unaligned DNA sequences known to be functionally regulated.** *Comput Applic Biosci* 1990, **6**:81-92.

62. Hertz GZ, Stormo GD: **Identification of consensus patterns in unaligned DNA and protein sequences: a large-deviation statistical byasis for penalizing gaps.** In *Bioinformatics and Genome Research*. Edited by Lim HA, Cantor CR. Singapore: World Scientific, 1995: 201-216.

63. **RSA-tools** [http://rsat.ulb.ac.be/rsat/]

64. **Pfam: Sigma54_activat**
[http://www.sanger.ac.uk/cgi-bin/Pfam/getacc?PF00158]

65. **Proteome analysis @ EBI**
[http://www.ebi.ac.uk/proteome/index.html]

66. **Entrez-Protein**
[http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Protein]

67. **Sean Eddy's Lab:HMMER 2.2** [http://hmmer.wustl.edu/]

68. **FramePlot 2.3.2** [http://www.nih.go.jp/~jun/cgi-bin/frameplot.pl]

69. Albright LM, Ronson CW, Nixon BT, Ausubel FM: **Identification of a gene linked to *Rhizobium meliloti* ntrA whose product is homologous to a family to ATP-binding proteins.** *J Bacteriol* 1989, **171**:1932-1941.

70. Michel-Reydellet N, Desnoues N, de Zamaroczy M, Elmerich C, Kaminski PA: **Characterisation of the *glnK-amtB* operon and the involvement of AmtB in methylammonium uptake in *Azorhizobium caulinodans*.** *Mol Gen Genet* 1998, **258**:671-677.

71. Michel-Reydellet N, Desnoues N, Elmerich C, Kaminski PA: **Characterization of *Azorhizobium caulinodans* glnB and glnA genes: involvement of the P(II) protein in symbiotic nitrogen fixation.** *J Bacteriol* 1997, **179**:3580-3587.

72. Martin GB, Thomashow MF, Chelm BK: ***Bradyrhizobium japonicum* glnB, a putative nitrogen-regulatory gene, is regulated by NtrC at tandem promoters.** *J Bacteriol* 1989, **171**:5638-5645.

73. Chiurazzi M, Iaccarino M: **Transcriptional analysis of the *glnB-glnA* region of *Rhizobium leguminosarum* biovar *viciae*.** *Mol Microbiol* 1990, **4**:1727-1735.

74. Arcondéguy T, Huez I, Fourment J, Kahn D: **Symbiotic nitrogen fixation does not require adenylylation of glutamine synthetase I in *Rhizobium meliloti*.** *FEMS Microbiol Lett* 1996, **145**:33-40.

75. Martin GB, Chapman KA, Chelm BK: **Role of the *Bradyrhizobium japonicum* ntrC gene product in differential regulation of the glutamine synthetase II gene (*glnII*).** *J Bacteriol* 1988, **170**:5452-5459.

76. Patriarca EJ, Chiurazzi M, Manco G, Riccio A, Lamberti A, De Paolis A, Rossi M, Defez R, Iaccarino M: **Activation of the *Rhizobium leguminosarum* glnII gene by NtrC is dependent on upstream DNA sequences.** *Mol Gen Genet* 1992, **234**:337-345.

77. Szeto WW, Nixon BT, Ronson CW, Ausubel FM: **Identification and characterization of the *Rhizobium meliloti* ntrC gene: *R. meliloti* has separate regulatory pathways for activation of nitrogen fixation genes in free-living and symbiotic cells.** *J Bacteriol* 1987, **169**:1423-1432.

78. Michiels J, Vanderleyden J: **Cloning and sequence of the *Rhizobium leguminosarum* biovar *phaseoli* fixA gene.** *Biochim Biophys Acta* 1993, **1144**:232-233.

79. Hontelez JG, Lankhorst RK, Katinakis P, van den Bos RC, van Kammen A: **Characterization and nucleotide sequence of a novel gene *fixW* upstream of the *fixABC* operon in *Rhizobium leguminosarum*.** *Mol Gen Genet* 1989, **218**:536-544.

80. Bauer E, Kaspar T, Fischer HM, Hennecke H: **Expression of the *fixR-nifA* operon in *Bradyrhizobium japonicum* depends on a new response regulator, RegR.** *J Bacteriol* 1998, **180**:3853-3863.

81. Quinto C, Delavega H, Flores M, Leemans J, Cevallos M A, Pardo M A, Azpiroz R, Girard M D, Calva E, Palacios R: **Nitrogenase reductase: a functional multigene family in *Rhizobium phaseoli*.** *Proc Natl Acad Sci USA* 1985, **82**:1170-1174.

82. Szeto WW, Zimmerman JL, Sundaresan V, Ausubel FM: **A *Rhizobium meliloti* symbiotic regulatory gene.** *Cell* 1984, **36**:1035-1043.

83. Aguilar OM, Reilander H, Arnold W, Puhler A: ***Rhizobium meliloti* nifN (*fixF*) gene is part of an operon regulated by a *nifA*-dependent promoter and codes for a polypeptide homologous to the *nifK* gene product.** *J Bacteriol* 1987, **169**:5393-5400.

84. Reid CJ, Poole PS: **Roles of DctA and DctB in signal detection by the dicarboxylic acid transport system of *Rhizobium leguminosarum*.** *J Bacteriol* 1998, **180**:2660-2669.

85. Jiang J, Gu BH, Albright LM, Nixon BT: **Conservation between coding and regulatory elements of *Rhizobium meliloti* and *Rhizobium leguminosarum* dct genes.** *J Bacteriol* 1989, **171**:5244-5253.

86. Murphy PJ, Heycke N, Trenz SP, Ratet P, de Bruijn FJ, Schell J: **Synthesis of an opine-like compound, a rhizopine, in alfalfa nodules is symbiotically regulated.** *Proc Natl Acad Sci USA* 1988, **85**:9133-9137.

87. Fischer HM, Babst M, Kaspar T, Acuna G, Arigoni F, Hennecke H: **One member of a *groESL*-like chaperonin multigene family in *Bradyrhizobium japonicum* is co-regulated with symbiotic nitrogen fixation genes.** *EMBO J* 1993, **12**:2901-2912.

88. Weidenhaupt M, Fischer HM, Acuna G, Sanjuan J, Hennecke H: **Use of a promoter-probe vector system in the cloning of a new NifA-dependent promoter (*ndp*) from *Bradyrhizobium japonicum*.** *Gene* 1993, **129**:33-40.

comment

reviews

reports

deposited research

refereed research

interactions

information