Patron: Her Majesty The Queen

ROTHAMSTED RESEARCH

Rothamsted Research
Harpenden, Herts, AL5 2JQ

Telephone: +44 (0)1582 763133
Web: http://www.rothamsted.ac.uk/

Rothamsted Repository Download

A - Papers appearing in refereed journals

Mauchline, T. H., Hayat, R., Roberts, R., Powers, S. J. and Hirsch, P. R. 2014. Assessment of core and accessory genetic variation in Rhizobium leguminosarum symbiovar trifolii strains from diverse locations and host plants using PCR-based methods. *Letters in Applied Microbiology.* 59 (2), pp. 238-246.

The publisher's version can be accessed at:

- https://dx.doi.org/10.1111/lam.12270

The output can be accessed at: https://repository.rothamsted.ac.uk/item/8qz00.

© Please contact library@rothamsted.ac.uk for copyright queries.

12/03/2019 11:17 repository.rothamsted.ac.uk library@rothamsted.ac.uk

Rothamsted Research is a Company Limited by Guarantee
Registered Office: as above. Registered in England No. 2393175.
Registered Charity No. 802038. VAT No. 197 4201 51.
Founded in 1843 by John Bennet Lawes.

Letters in Applied Microbiology

Letters in Applied Microbiology ISSN 0266-8254

ORIGINAL ARTICLE

Assessment of core and accessory genetic variation in *Rhizobium leguminosarum* symbiovar *trifolii* strains from diverse locations and host plants using PCR-based methods

T.H. Mauchline[1], R. Hayat[1,2], R. Roberts[1,3], S.J. Powers[1] and P.R. Hirsch[1]

1 Rothamsted Research, Harpenden, UK
2 PMAS Arid Agriculture University, Rawalpindi, Pakistan
3 University of Reading, Reading, UK

Significance and Impact of the Study: The nitrogen-fixing symbiosis between *Rhizobium leguminosarum* and host legumes has been recognized as a key part of sustainable agriculture for many years; this study provides new tools to study rhizobial biogeography which will be invaluable for extending the cultivation of legumes and indicating whether or not inoculation is necessary.

Keywords
bacterial chromosome, ERIC PCR, gyrB, nodA, symbiotic plasmids.

Correspondence
Tim Mauchline, Department of AgroEcology, Rothamsted Research, Harpenden, Hertfordshire AL5 2JQ, UK.
E-mail: tim.mauchline@rothamsted.ac.uk

2014/0271: received 10 February 2014, revised 4 April 2014 and accepted 7 April 2014

doi:10.1111/lam.12270

Abstract

The nitrogen-fixing symbiosis between *Rhizobium leguminosarum* and host legumes is recognized as a key part of sustainable agriculture. A culture collection containing rhizobia isolated from legumes of economic importance in the UK and worldwide, maintained at Rothamsted Research for many years, provided material for this study. We aimed to develop and validate efficient molecular diagnostics to investigate whether the host plant or geographical location had a greater influence on the genetic diversity of rhizobial isolates, and the extent to which the core bacterial genome and the accessory symbiosis genes located on plasmids were affected. To achieve this, core housekeeping genes and those involved in symbiosis interactions were sequenced and compared with genome-sequenced strains in the public domain. Results showed that some *Rh. leguminosarum* symbiovar *trifolii* strains nodulating clovers and *Rh. leguminosarum* sv. *viciae* strains nodulating peas and vicias shared identical housekeeping genes, clover nodule isolates from the same location could have divergent symbiosis genes, and others isolated on different continents could be very similar. This illustrates the likely co-migration of rhizobia and their legume hosts when crops are planted in new areas and indicates that selective pressure may arise from both local conditions and crop host genotypes.

Introduction

Nitrogen-fixing symbioses between root nodule-inducing bacteria and leguminous plants are becoming increasingly important in agricultural systems, as oil and chemical fertilizer costs increase. Clover is important in UK pastures, providing nitrogen to grasses and improving livestock nutrition. The clover symbiont *Rhizobium leguminosarum* symbiovar *trifolii* is ubiquitous in UK soil and the common European clovers are normally nodulated by 'effective' nitrogen-fixing rhizobia (Hirsch 1996; Macdonald *et al.* 2011). By contrast in Australia, clovers were imported to improve grassland and often required inoculation with compatible strains of *Rh. leguminosarum* sv. *trifolii* (Hutton 1968; Yates *et al.* 2011) although rhizobia may also travel on seeds of the host plant. The transcontinental history of clovers means that rhizobia re-isolated from root nodules are not necessarily 'native' to the soil from which the roots were collected. The species *Rh. leguminosarum* contains three symbiovars that nodulate clovers, peas and vicias or phaseolus beans (sv. *trifolii*, sv. *viciae* and sv. *phaseoli*, respectively). Host

specificity is determined by *nod* genes carried on one of several large plasmids which may be transferred to other *Rh. leguminosarum* strains, resulting in a mobile accessory genome that can be equivalent in size to the bacterial chromosome (Hirsch 1996).

The common chromosomal lineage of the species *Rh. leguminosarum* and host-specific selection of symbiosis genes in the various symbiovars results in the development of communities containing individual isolates with closely related core genomes but divergent plasmid differing in host specificity and conversely isolates with divergent chromosome but closely related symbiosis plasmids (Laguerre *et al.* 1996; Tian *et al.* 2010; Provorov *et al.* 2012). There is a considerable variation in host specificity within each symbiovar, not all rhizobial isolates that can nodulate a host will be 'effective' (able to fix N), and some will be more efficient than others. Thus, there is a need to identify appropriate rhizobial inoculants when legumes are grown far from their centres of evolutionary origin (Howieson *et al.* 2005). Identification methods for nodule isolates used in the past gave variable results (serology and phenotyping) or were time-consuming (plasmid profiling) and had limited discrimination (Hirsch 1996). They have been superseded by PCR-based DNA fingerprinting methods (Depret *et al.* 2004; Muresu *et al.* 2005; Provorov *et al.* 2012); the recent advent of next-generation sequencing facilitates amplicon sequence comparison of environmental isolates (Shams *et al.* 2013).

The different selective pressures that shape the core bacterial genome compared with that of the accessory genome (which includes conjugative plasmids and symbiosis genes) could lead to the generation of types that are particularly adapted to either local soil or host plant variants. Plasmids may differentially influence survival in soil, host plant and nonhost rhizospheres (Moënne-Loccoz and Weaver 1995; Clark *et al.* 2002). Rearrangement of segments of the genome and transfer of entire plasmids lead to multiple variants: substantial sequence diversity was apparent in both the core and accessory genomes of *Rh. leguminosarum* sv. *trifolii* isolated from clover nodules within a 1-m^2 plot in Poland (Mazur *et al.* 2011). Diversity may be driven by both soil and host plants: variation in *Rh. leguminosarum* sv. *trifolii* nodulating recently introduced ruderal clover on Srednii island, Russia, was associated with soil type, whereas variation in *Rh. leguminosarum* sv. *viciae* nodulating native vetches was associated with host-plant populations (Provorov *et al.* 2012). Additionally, horizontal gene transfer events can be a source of diversity. For example, Rashid *et al.* 2014 found evidence that nodulation genes had transferred between *Rhizobium leguminosarum* lineages.

The sequencing of housekeeping and nodulation genes along with PCR-based DNA fingerprinting methods has superseded the traditional identification methods. As such, the purpose of this study was to develop a robust and rapid PCR-based methodology for taxonomic resolution of *Rh. leguminosarum* sv. *trifollii*. We tested three methods against a cosmopolitan collection of *Rh. leguminosarum* sv *trifolii* isolates maintained at Rothamsted Research for many years, some isolated in the UK, others from diverse locations worldwide. We aimed to investigate the biogeography of *Rh. leguminosarum* sv. *trifolii* and to establish the extent to which the site and clover species of origin influence the bacterial chromosome and symbiosis plasmid. The approaches included sequencing and comparing an example of a gene from both the core and accessory genomes: *gyrB* is an essential housekeeping gene encoding DNA gyrase located on the bacterial chromosome; *nodA* is located on the symbiosis plasmid, being part of the accessory genome, and has a role in determining host specificity. We assessed the validity of *nodA* and *gyrB* as taxonomic markers for this group. In addition, DNA fingerprinting with enterobacterial repetitive intergenic consensus (ERIC) PCR (Versalovic *et al.* 1991) was used to give an overview of the entire genome. The survey has been extended by including *gyrB* and *nodA* gene sequences from other *Rh. leguminosarum* sv. *trifolii* and sv. *viciae* strains that are now available online.

Results and discussion

Primer design

The *gyrB*-based primers gyrB f3 and gyrB R5 were designed from alignments of rhizobia *gyrB* genes available in the public databases, in the Align X program (part of the Vector NTI suite), to amplify a range of rhizobia. As a combination, they facilitate amplification of a fragment of the *gyrB* gene in representatives of the *Rhizobium* spp. used in this study, as well as *Sinorhizobium* spp. (*S. meliloti* 1021 and 2011 tested in this study) and *Mesorhizobium* spp. (*M. loti* 1055 and 3209 tested in this study). The *nodA*-based primers were designed to amplify the entire gene; they were used successfully on *Rh. leguminosarum* strains, and *in silico* analysis indicates that they should also facilitate PCR amplification of this gene in *M. loti* and *S. meliloti*.

16S rRNA gene phylogeny of isolates

The 16S rRNA gene from six *Rh. leguminosarum* sv. *trifolii* isolates held in the Rothamsted collection (RCR32, RCR46, RCR59, RCR221, RCR226 and RCR227) was sequenced and compared with 12 *Rh. leguminosarum* sv. *trifolii* (CB782, CC278f, CC283b, SRD1565, SRD1943, TA1, WSM597, WSM1325, WSM1689, WSM2012,

WSM2297, WSM2304) and four *Rh. leguminosarum* sv. *viciae* accessions (3841 and 248, TOM, Vf39) in the NCBI database. Most sequences were identical and differed from the type strain by one base in the 956 b segment compared, isolate RCR221 differed by 2 b and 248 and SRD1565 (which were identical) differed by 3 b from 3841.

Sequence analysis of *gyrB*

Data for *gyrB* was obtained by sequencing the gene in *Rh. leguminosarum* accessions at Rothamsted, or extracting sequence from the NCBI database of genome-sequenced bacteria. Overall, phylogenetic analysis of the 1432 b *gyrB* fragment (Fig. 1) revealed that this gene is useful for discriminating *Rh. leguminosarum* strains. The analysis included 28 examples of *Rh. leguminosarum* sv. *trifolii*, three of *Rh. leguminosarum* sv. *phaseoli* and two of *Rh. leguminosarum* sv. *viciae* (see Table 1 for strain details). Interestingly, most appear to be closely related independent of host plant, however, some *Rh. leguminosarum* sv. *trifolii* strains from Uruguay, Australia, Ethiopia and South Africa appear to group separately from the main group which also contains examples of both *Rh. leguminosarum* sv. *viciae* and sv. *phaseoli*. This indicates that this gene is useful for discriminating *Rh. leguminosarum* isolates, but is unreliable for the separation of strains based on their sv. designation.

Sequence analysis of *nodA*

The 464 b sequence from *nodA* separated according to symbiovar, the six *Rh. leguminosarum* sv. *viciae* isolates forming separate clusters. As such, it is apparent that accessory genes such as *nodA* can be used to discriminate *Rh. leguminosarum* strains according to symbiovar, this is not the case for *gyrB*. They were excluded from the final analysis to focus on relationships within sv. *trifolii* (Fig. 2). These showed some clustering that matched the *gyrB* groups and some that did not. For example, Caucasian and Russian *T. ambiguum* isolates RCR213 and CC283b had identical *nodA* sequences, whereas the northern UK perennial clover isolates (RCR3, RCR10, RCR35, RCR49) grouped together with RCR59, as was the case for *gyrB,* but the Australian annual clover isolate RCR227 did not, having <95% *nodA* sequence identity. The Uruguayan sv. *trifolii* strains separate from the other strains, as they did with their equivalent *gyrB* sequences; however, the Australian isolate RCR212 is no longer clustered with these strains, but is rather positioned in the main grouping with this gene. Additionally, the African strains remain represented in a separate group.

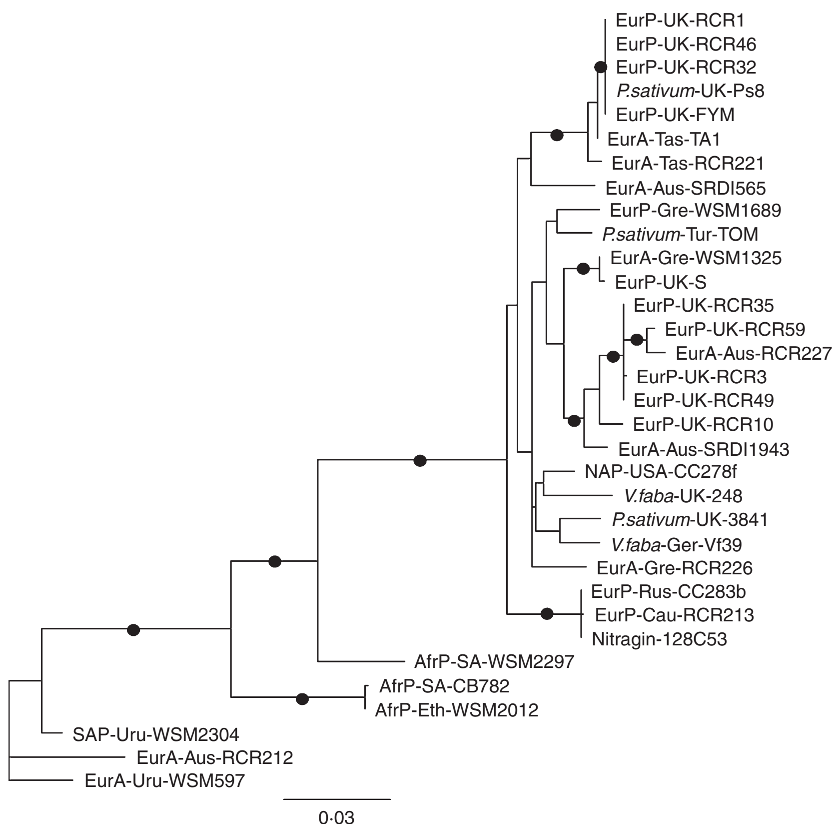

Figure 1 Phylogenetic tree generated by PAUP* of *Rhizobium leguminosarum* based on *gyrB* sequences. Names start with plant from which strain was isolated (EurP: European perennial clover; EurA: European annual clover; SAP: South American perennial clover; NAP: North American perennial clover; AfrP: African perennial clover) followed by region of origin (Aus: Australia; Tas: Tasmania; Gre: Greece; Tur: Turkey; Uru: Uruguay; Eth: Ethiopia; SA: South Africa) and name as listed in Table 1. Bootstrap consensus of branching less than 70% unless indicated: ● 90–100%; ○ 80–90%.

Table 1 *Rhizobium leguminosarum* strains

Strain	Synonym	Source	Original host	Effective on *Trifolium repens*	References
Rhizobium leguminosarum* symbiovar *trifolii					
RCR1		Hertfordshire, UK	*T. repens*	+	1
RCR3		Cumbria, UK	*T. repens*	+	1
RCR10		Dumbartonshire, UK	*T. repens*	+	1
RCR32		Aberystwyth, UK	*T. repens*	+	1, 2
RCR35		Orkney, UK	*T. repens*	+	1
RCR46		Bedfordshire, UK	*T. glomeratum*	+	1, 2, 33
RCR49		Scotland, UK	*T. repens*	+	1
RCR59		Hertfordshire, UK	*T. pratense*	+	1
RCR212	SU297/3	NSW, Australia	*T. incarnatum*	+	1
RCR213	Ta 12	Caucasus,	*T. ambiguum*	+	1
RCR221	TA1	Tasmania	*T. subterraneum*	+	1, 2, 3, 4
RCR226	CC2480a	N. Greece	*T. subterraneum*	+	1, 2, 3
RCR227	WU95, WSM170	W. Australia	*T. subterraneum*	+	1, 2, 3, 5
FYM		Bedfordshire, UK	*T. repens*	+	6
S		Bedfordshire, UK	*T. repens*	−	6
WSM597		Uruguay	*T. pallidum*	?	7
WSM2304		Uruguay	*T. polymorphum*	Not tested	5, 7, 8, 10,11
CB782	Group K	Kenya	*T. sempilosum*	−	7, 8
WSM2012	MAR1468	Ethiopia	*T. rueppellianum*	Not tested	7, 8
WSM2297		South Africa	*T. africanum*	Not tested	7, 9
SRDI943	V2-2	NSW, Australia	*T. subterraneum*	Not tested	7, 10
WSM1689		Greece	*T. uniflorum*	−	7, 8
CC278f		Colorado, USA	*T. nanum*	Not tested	7, 8
CC283b	WSM46	Russia	*T. ambiguum*	−	7, 8
SRDI565	NJ-8	NSW, Australia	*T. subterraneum*	Not tested	7, 10
TA1		Tasmania	*T. subterraneum*	+	7, 8
WSM1325		Greece	Annual clover	Not tested	5, 7, 12
Rhizobium leguminosarum* symbiovar *viciae effective on *Pisum, Vicia,* noninfective on *Trifolium spp.*					
3841	300str^r	Peterborough, UK	*Pisum sativum*		7, 13
JI248		Norwich, UK	*Vicia faba*		7, 13
Vf39		Bielefeld, Germany	*Vicia faba*		7, 14
128C53	1684 is str^r rif^r	Nitragin, USA	Not reported		7, 15
TOM		Turkey	*Pisum sativum*		7, 16
Ps8		York, UK	*Pisum sativum*		7

+ effective N-fixing nodulation; − nodules ineffective; ? unknown; (1) Dye (1981); (2) Wang *et al.* (1986); (3) Kumarasinghe and Nutman (1979); (4) Hepper and Lee (1979); (5) Yates *et al.* (2008); (6) Hirsch *et al.* (1993); (7) JGI GOLD http://www.genomesonline.org; (8) Howieson *et al.* (2005); (9) J. Howieson personal Communication; (10) Melino *et al.* (2012); (11) Reeve *et al.* (2010b); (12) Reeve *et al.* (2010a); (13) Hirsch (1979); (14) Priefer (1989); (15) Phillips and Torrey (1972); (16) Lie (1978).

Comparison of an 881 b sequence from *nodD* extracted from the GOLD database showed a similar clustering for African and South American isolates, which fell into distinct groups, the North American isolate CC287f more similar to those from Uruguay than those from Europe, Australia and the Caucasus. It would be interesting to see whether a distinct group of US isolates would be formed with more examples from this region.

Utility of *gyrB* and *nodA* genes for investigating *Rhizobium leguminosarum* symbiovar *trifolii* phylogeny

The *nodA* and *gyrB* sequences show clear differences between *Rh. leguminosarum* sv. *trifolii* strains that can be linked to biogeography. We compared strain discrimination based on these genes with a genomic average nucleotide identity (ANI) scheme devised by Zhang *et al.* (2012). The ANI scheme, based on the genes *SMc00019*, *truA* and *thrA*, has been applied to a wide range of rhizobia (Zhang *et al.* 2012) and has been suggested as a replacement for classical DNA-DNA hybridization methodology. We constructed phylogenies from *SMc0019*, *truA*, *thrA*, *nodA* and *gyrB* sequence data available in the public databases for 11 *Rh. leguminosarum* sv. *trifolii* isolates (identified in bold in Table 1). We found that trees based on *SMc0019*, *truA* and *thrA* as well as a concatenate of all three genes were partially congruent with each other as well as with trees based on the *nodA* or *gyrB* (data not

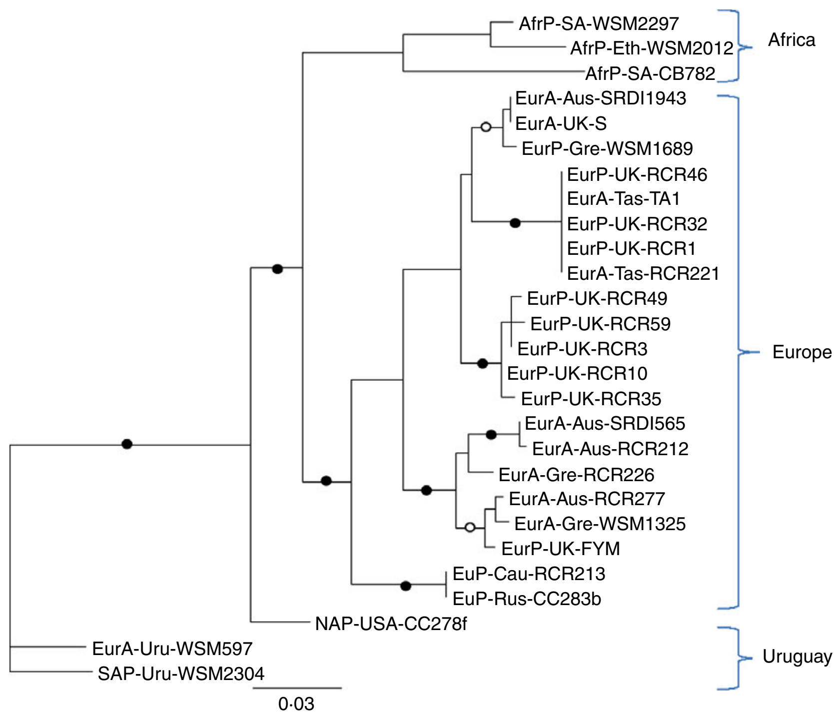

Figure 2 Phylogenetic tree generated by PAUP* of *Rhizobium leguminosarum* bv. *trifolii* based on *nodA* sequences. Names start with plant from which strain was isolated (EurP: European perennial clover; EurA: European annual clover; SAP: South American perennial clover; NAP: North American perennial clover; AfrP: African perennial clover) followed by region of origin (Aus: Australia; Tas: Tasmania; Gre- Greece; Tur: Turkey; Uru: Uruguay; Eth: Ethiopia; SA: South Africa) and name as listed in Table 1. Bootstrap consensus of branching less than 70% unless indicated: ● 90–100%; ○ 80–90%.

shown). Additionally, all trees resolved the species into three main groupings based on biogeography (i.e. Uruguay, Africa and Europe). However, the pairings within the European group were found to vary between each gene as well as with the three-gene concatenation suggested by Zhang *et al.* (2012) in their ANI scheme (data not shown). As such, we conclude that the use of the core genome-located *gyrB* as well as the accessory genome-located *nodA*, in common with *truA*, *thrA* and *SMc0019*, is suitable for broad-scale phylogeny of this species, although the use of several suitable genes in multilocus sequence typing or ANI schemes will likely provide higher resolution discrimination.

Analysis of 16S rRNA gene sequences supports the classification of all the strains used as belonging to the same species, *Rh. leguminosarum,* in agreement with other studies (Rogel *et al.* 2011). This is confirmed by the grouping of variants of the core housekeeping gene *gyrB*, where identical sequences were found in different symbiovars. Analysis of *gyrB and nodA* also indicates 'European' and 'non-European' grouping of *Rh. leguminosarum* sv. *trifolii,* with most Australian isolates and the North American isolate falling into the 'European' group, possibly indicating a history of shared clover seed stocks and rhizobial inoculant strains. The RCR culture collection was once a source of material that was distributed worldwide, which could explain, for example, the similarity in strains isolated from Australia (RCR227) and the UK (RCR59). It also acquired strains from other collections and researchers and it is intriguing to discover that the Caucasian isolate RCR213, acquired prior to 1976, has identical *nodA* and very similar *gyrB* sequences to Russian isolate CC283b held in Australia (Howieson *et al.* 2005). It is possible that they are derivatives of the same original strain although this cannot be verified.

It is interesting that the *nodA* phylogeny supports *gyrB* in most respects with distinct South American, African and Caucasian groups although there are some differences in grouping within the European/Australian isolates. The similarity between the two Uruguayan isolates is interesting as WSM597 came from a European annual clover nodule and WSM2304 from a South American perennial clover. However, the results overall indicate that *Rh. leguminosarum* sv. *trifolii* is a cosmopolitan species, relationships being determined in part by the site of isolation of the strain and in part by the regional origin of the host plant species. European annual and perennial clovers originated in the Eastern Mediterranean and Caucasus although many are now considered 'native' flora in Western Europe. The major differences in host specificity may be between this group, and the African and South American species (Howieson *et al.* 2005).

ERIC profiles

ERIC profiles of *Rh. leguminosarum* sv. *trifolii* and sv. *viciae* strains in the Rothamsted Culture Collection were compared (Fig. 3). They resolved in two main

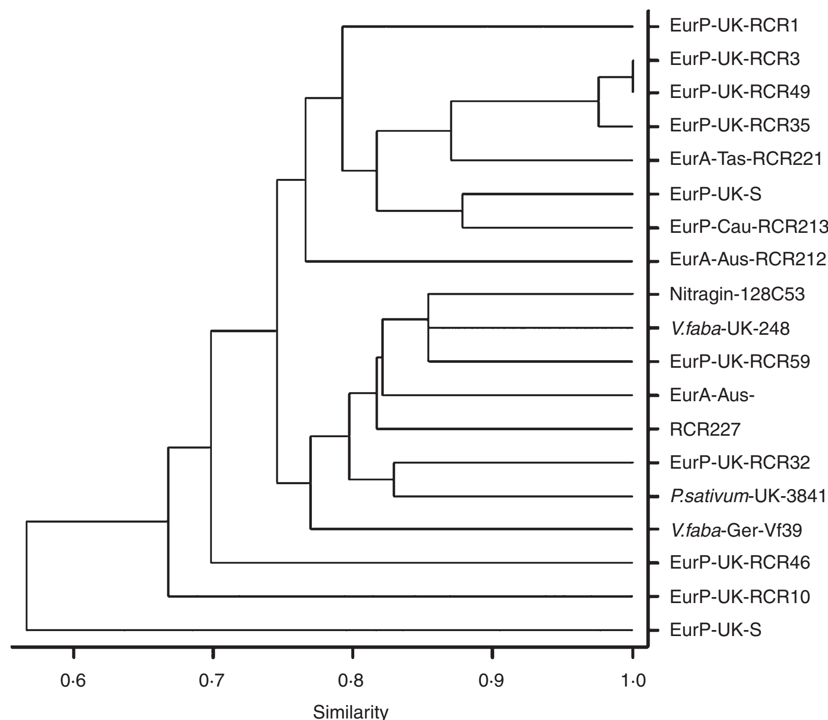

Figure 3 Cluster analysis of enterobacterial repetitive intergenic consensus profiles. Names start with plant from which strain was isolated (EurP: European perennial clover; EurA: European annual clover; SAP: South American perennial clover; NAP: North American perennial clover; AfrP: African perennial clover) followed by region of origin (Aus: Australia; Tas: Tasmania; Gre: Greece; Tur: Turkey; Uru: Uruguay; Eth: Ethiopia; SA: South Africa) and name as listed in Table 1.

groups each with 0·75 similarity. One group contained RCR3, RCR35 and RCR49 (0·95 similarity) which had almost identical *gyrB* and *nodA* sequences; four other *Rh. leguminosarum* sv. *trifolii* isolates fell into this group. The other group contained both *Rh. leguminosarum* sv. *trifolii* and sv. *viceae* and included RCR212 which had been shown to have a substantially different *gyrB* sequence. The *Rh. leguminosarum* sv. *trifolii* outliers (RCR46, RCR10, S) had only 0·70 similarity to either main groups. Identical *gyrB* or *nodA* sequences did not necessarily determine ERIC similarity: RCR32 and RCR1 fell into different groups, with RCR46 as an outlier. The ERIC profiles reveal differences in the overall arrangement of the core and accessory genomes that are not predicted by single gene markers *gyrB* and *nodA*. Previous studies involving some of the strains investigated in this work confirm this. For example, *Rh. leguminosarum* sv. *trifolii* isolates RCR32, RCR46, RCR221/TA1, RCR226, RR227/WU95, FYM and S had different plasmid profiles (Hirsch *et al.* 1987, 1993), although RCR32, RCR46 and RCR221 shared very similar *gyrB* and *nodA* sequences.

Tasmanian subterranean clover isolate TA1 that had been acquired by the Rothamsted collection before 1976 and maintained as RCR221 illustrates that DNA sequences change over time. The sequence of *nodA* is identical, but there are 9 b changes in *gyrB*. This discrepancy may reflect the relatively longer sequence of *gyrB*, three times that of *nodA* and indicates a rate of approx. 0·2% nucleotide substitutions per decade.

Overall, the study has illustrated that genes indicative of the core and accessory genomes in *Rh. leguminosarum* can provide detailed information on the relationship between strains, allowing comparison of a historical culture collection with recently sequenced genomes and illustrating that both environment and host plant are likely to influence diversity. We show that, *gyrB*, an example of a core housekeeping gene is useful for discrimination of *Rhizobium leguminosarum* strains, regardless of their sv. designation. However, *nodA* part of the accessory genome can be used to discriminate isolates according to their sv. designation as well as biogeography at the intra-sv. level. Additionally, we show that examples of individual core and accessory genome loci are suitable for intra-species diversity of *Rh. leguminosarum* sv. *trifolii,* and compare well with more sophisticated approaches based on multiple loci for broad-scale genotyping. The tools to study rhizobial biogeography provided in this study will be useful for aiding the expansion of legume cultivation and can be used to indicate whether or not rhizobial inoculation is necessary.

Materials and methods

Rothamsted Rhizobium collection

The Rothamsted Collection of Rhizobium (RCR) was accrued over many years and contains rhizobial isolates suitable for different host legumes. Although the

collection is no longer actively curated, lyophilised cultures have been maintained for the past 25 years. Cultures were revived on yeast extract mannitol agar (YEMA) or tryptone yeast (TY) agar at 28°C for 1 week and tested for effective nodulation of *Trifolium repens* (Hirsch and Skinner 1992). Any *Rh. leguminosarum* sv. *trifolii* accessions unable to nodulate were excluded from the study, but six strains of *Rh. leguminosarum* sv. *viciae* in our collection, for which full genome sequence data are available, were included. These are listed in Table 1 along with the *Rh. leguminosarum* sv. *trifolii* strains not held in our collection but for which genome data are available. DNA was extracted from cultures grown in TY broth for subsequent analysis using the Gram negative extraction method with the GenElute Bacterial Genomic DNA Kit (Sigma-Aldrich).

PCR

PCR on 10 ng *Rh. leguminosarum* DNA was performed with primers listed in Table 2. The PCR mixtures (25 μl) contained 1 μl of sample (diluted up to 10-fold in water or undiluted), each primer at a concentration of 0·1 μmol l^{-1}, 1x PCR buffer (1·5 mmol l^{-1} Mg^{2+}; Bioline, London, UK), 1 mmol l^{-1} $MgCl_2$, each dNTP (Fermentas Life Sciences, Burlington, Canada) at a concentration of 0·2 mmol l^{-1}, and 1 U of Biotaq polymerase (Bioline). The thermo-cycling conditions were as follows for *gyrB*, *nodA* and 16S rRNA gene fragments: 95°C 1 min, followed by 30 cycles of 94°C 1 min, 60°C 1 min, 72°C 1 min and a final extension step of 72°C for 5 min. For ERIC PCR, identical conditions were used except that annealing was at 52°C. Finally, 5 μl of each PCR sample was subjected to electrophoresis for *gyrB*, *nodA* and 16S rRNA gene fragments, and 10 μl for ERIC PCR products. 1% agarose gels were run at 90V for 60 min with 1-kb ladder (Fermentas) and stained with ethidium bromide (0·5 μg ml^{-1}) to determine whether PCR amplification had been successful. The remaining PCR product for *gyrB*, *nodA* and 16S rRNA gene fragments were PCR purified (Qiagen, Venlo, Netherlands) prior to sequencing.

DNA sequencing

Sequencing of PCR products was performed by Eurofins MWG/Operon (Germany) using a PCR product concentration of 5 ng μl^{-1} for the *nodA* gene, and 10 ng μl^{-1} for both *gyrB* and 16S rRNA gene fragments. The primer NodA 3f was used for the *nodA* gene, and the *gyrB* gene fragment was sequenced from both ends with primers gyrB 3f and gyrB 5r as was the 16S rRNA gene fragment with primers Muyzer F and 1389r. Sequences retrieved were tested with the BLASTN algorithm using default settings to ensure their identity as *Rhizobium leguminosarum* (Altschul *et al.* 1990).

DNA sequence alignment

PCR product DNA sequence pairs for the *gyrB* and 16S rRNA gene fragments were used to create '*in silico*' molecules in the Vector Nti suite (Invitrogen, Carlsbad, CA).

DNA sequences from cultures and those extracted from the GOLD database (http://www.genomesonline.org/cgi-bin/GOLD/Search.cgi) were aligned using MUSCLE and phylogenetic trees based on Maximum Likelihood were constructed using the PAUP* module in the bioinformatics package GENEIOUS 6.1.4. The model 'GTR + I + G' was chosen automatically by the Model test function.

The tree topology was supported by bootstrapping 1000 times; branches with support >70% were considered robust. DNA sequences were submitted to GenBank and accession numbers were obtained (Table S1).

Cluster analysis of ERIC profile data

The ERIC profiles from 15 *Rh. leguminosarum* sv. *trifolii* and four sv. *viciae* strains were viewed for presence/absence of 41 detectable bands. Next, the GenStat statistical system (GenStat 2010 Twelfth edition, ©Lawes Agricultural Trust (Rothamsted Research), VSN International Ltd., Hemel Hempstead, UK) was used to construct a similarity matrix using the simple matching measure of similarity. This was used in cluster analysis, incorporating

Table 2 PCR primers used in this study

Primer	Gene	Sequence	Source
nodA 3f	*nodA*	TGGAARMTAWSYTGGGAAAAT	This study
nodA 4r	*nodA*	CCRTTYCGTTCRATCAAYG	This study
gyrB 3f	*gyrB*	ATGTGGTGGAACGAYAGCTA	This study
gyrB 5r	*gyrB*	TCCTGGATRAAKTCGCG	This study
Muyzer F	16S rRNA	CCTACGGGAGGCAGCAG	Muyzer *et al.* (1993)
1389r	16S rRNA	ACGGGCGGTGTGTACAAG	Osborn *et al.* (2000)
ERIC2	–	AAGTAAGTGACTGGGGTGAGCG	Versalovic *et al.* (1991)
R1CIRE	–	CACTTAGGGGTCCTCGAATGTA	Versalovic *et al.* (1991)

the single-link (nearest-neighbour) method of clustering, allowing construction of a dendrogram.

Acknowledgements

Rothamsted Research receives strategic funding from the BBSRC. We also thank Rothamsted International for their support.

Conflict of Interest

The authors declare no conflict of interest.

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. *J Mol Biol* **215**, 403–410.

Clark, I.M., Mendum, T.A. and Hirsch, P.R. (2002) The influence of the symbiotic plasmid pRL1JI on the distribution of GM rhizobia in soil and crop rhizospheres, and implications for gene flow. *Antonie Van Leeuwenhoek* **81**, 607–616.

Depret, G., Houot, S., Allard, M.R., Breuil, M.C., Nouaim, R. and Laguerre, G. (2004) Long-term effects of crop management on *Rhizobium leguminosarum* biovar *viciae* populations. *FEMS Microbiol Ecol* **51**, 87–97.

Dye, M. (1981) *Rothamsted Collection of Rhizobium* Catalogue of Strains. Harpenden, Hertfordshire, UK: Rothamsted Research, p. 48.

Hepper, C.M. and Lee, L. (1979) Nodulation of *Trifolium subterraneum* by *Rhizobium leguminosarum*. *Plant Soil* **51**, 441–445.

Hirsch, P.R. (1979) Plasmid-determined bacteriocin production by *Rhizobium leguminosarum*. *J Gen Microbiol* **113**, 219–228.

Hirsch, P.R. (1996) Population dynamics of indigenous and genetically modified rhizobia in the field. *New Phytol* **133**, 159–171.

Hirsch, P.R. and Skinner, F.A. (1992). The Identification of Rhizobium and Bradyrhizobium. In *Identification Methods in Applied and Environmental Microbiology* ed. Board, R.G., Jones, D. and Skinner, F.A. London: Blackwell, pp. 45–65.

Hirsch, P.R., Wang, C.L. and Beringer, J.C. (1987) Instability in *Rhizobium leguminosarum* strains carrying two symbiotic plasmids. *Heredity* **59**, 310.

Hirsch, P.R., Jones, M.J., McGrath, S.P. and Giller, K.E. (1993) Heavy-metals from past applications of sewage-sludge decrease the genetic diversity of *Rhizobium leguminosarum* biovar *trifolii* populations. *Soil Biol Biochem* **25**, 1485–1490.

Howieson, J.G., Yates, R.J., O'Hara, G.W., Ryder, M. and Real, D. (2005) The interactions of *Rhizobium leguminosarum* biovar *trifolii* in nodulation of annual and perennial *Trifolium* spp. from diverse centres of origin. *Aust J Exp Agr* **45**, 199–207.

Hutton, E.M. (1968) Australia's pasture legumes. *Aust J Agric Res* **34**, 203–218.

Kumarasinghe, R.M.K. and Nutman, P.S. (1979) Influence of temperature on root hair infection of *Trifolium parviflorum* and *Trifolium glomeratum* by root nodule bacteria. 1. Effects of constant root temperature on infection and related aspects of plant development. *J Exp Bot* **30**, 503–515.

Laguerre, G., Mavingui, P., Allard, M.R., Charnay, M.P., Louvrier, P., Mazurier, S.I., Rigottier-Gois, L. and Amarger, N. (1996) Typing of rhizobia by PCR DNA fingerprinting and PCR-restriction fragment length polymorphism analysis of chromosomal and symbiotic gene regions, application to *Rhizobium leguminosarum* and its different biovars. *Appl Environ Microbiol* **62**, 2029–2036.

Lie, T.A. (1978) Symbiotic specialisation in pea plants: the requirement of specific *Rhizobium* strains for peas from Afghanistan. *Ann Appl Biol* **88**, 462–465.

Macdonald, C.A., Clark, I.M., Hirsch, P.R., Zhao, F.J. and McGrath, S.P. (2011) Development of a real-time PCR assay for detection and quantification of *Rhizobium leguminosarum* bacteria and discrimination between different biovars in zinc-contaminated soil. *Appl Environ Microbiol* **77**, 4626–4633.

Mazur, A., Stasiak, G., Wielbo, J., Kubik-Komar, A., Marek-Kozaczuk, M. and Skorupska, A. (2011) Intragenomic diversity of *Rhizobium leguminosarum* bv. *trifolii* clover nodule isolates. *BMC Microbiol* **11**, 123.

Melino, V.J., Drew, E.A., Ballard, R.A., Reeve, W.G., Thomson, G., White, R.G. and Ohara, G.W. (2012) Identifying abnormalities in symbiotic development between *Trifolium* spp. and *Rhizobium leguminosarum* bv. *trifolii* leading to sub-optimal and ineffective nodule phenotypes. *Ann Bot* **110**, 1559–1572.

Moënne-Loccoz, Y. and Weaver, R.W. (1995) Plasmids influence growth of rhizobia in the rhizosphere of clover. *Soil Biol Biochem* **27**, 1001–1004.

Muresu, R., Sulas, L., Polone, E. and Squartini, A. (2005) PCR primers based on different portions of insertion elements can assist genetic relatedness studies, strain fingerprinting and species identification in rhizobia. *FEMS Microbiol Ecol* **54**, 445–453.

Muyzer, G., de Waal, E.C. and Uitterlinden, A.G. (1993) Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl Environ Microbiol* **59**, 695–700.

Osborn, A.M., Moore, E.R.B. and Timmis, K.N. (2000) An evaluation of terminal-restriction fragment length polymorphism (T-RFLP) analysis for the study of microbial community structure and dynamics. *Environ Microbiol* **2**, 39–50.

Phillips, D.A. and Torrey, J.H. (1972) Plasmids influence growth of rhizobia in the rhizosphere of clover. *Plant Physiol* **49**, 11–15.

Priefer, U.B. (1989) Genes involved in lipopolysaccharide production and symbiosis are clustered on the chromosome of *Rhizobium leguminosarum* biovar *viciae* Vf39. *J Bacteriol* **171**, 6161–6168.

Provorov, N.A., Andronov, E.E., Onishchuk, O.P., Kurchak, O.N. and Chizhevskaya, E.P. (2012) Genetic structure of the introduced and local populations of *Rhizobium leguminosarum* in plant-soil systems. *Microbiology* **81**, 224–232.

Rashid, M.H., Gonzalez, G., Young, P.W. and Wink, M. (2014) *Rhizobium leguminosarum* is the symbiont of lentils in the Middle East and Europe but not in Bangladesh. *FEMS Microbiol Ecol* **87**, 64–77.

Reeve, W., O'Hara, G., Chain, P., Ardley, J., Bräu, L., Nandasena, K., Tiwari, R., Copeland, A. *et al.* (2010a) Complete genome sequence of *Rhizobium leguminosarum* bv. *trifolii* strain WSM1325, an effective microsymbiont of annual Mediterranean clovers. *Stand Genomic Sci* **2**, 347–356.

Reeve, W., O'Hara, G., Chain, P., Ardley, J., Bräu, L., Nandasena, K., Tiwari, R., Copeland, A. *et al.* (2010b) Complete genome sequence of *Rhizobium leguminosarum* bv *trifolii* strain WSM2304, an effective microsymbiont of the South American clover *Trifolium polymorphum*. *Stand Genomic Sci* **2**, 66–76.

Rogel, M.A., Ormeno-Orrillo, E. and Romero, E.M. (2011) Symbiovars in rhizobia reflect bacterial adaptation to legumes. *Syst Appl Microbiol* **34**, 96–104.

Shams, M., Vial, L., Chapulliot, D., Nesme, X. and Lavire, C. (2013) Rapid and accurate species and genomic species identification and exhaustive population diversity assessment of *Agrobacterium* spp. using *recA*-based PCR. *Syst Appl Microbiol* **36**, 351–358.

Tian, C.F., Young, J.P.W., Wang, E.T., Tamimi, S.M. and Chen, W.X. (2010) Population mixing of *Rhizobium leguminosarum* bv. *viciae* nodulating *Vicia faba*: the role of recombination and lateral gene transfer. *FEMS Microbiol Ecol* **73**, 563–576.

Versalovic, J., Koeuth, T. and Lupshi, J.R. (1991) Distribution of repetitive DNA sequences in eubacteria and application to fingerprinting of bacterial genomes. *Nucleic Acids Res* **19**, 6823–6831.

Wang, C.L., Beringer, J.E. and Hirsch, P.R. (1986) Host plant effects on hybrids of *Rhizobium leguminosarum* biovars *viciae* and *trifolii*. *J Gen Microbiol* **132**, 2063–2070.

Yates, R.J., Howieson, J.G., Reeve, W.G., Brau, L., Speijers, J., Nandasena, K.D., Real, E., Sezmis, G.W. *et al.* (2008) Host-strain mediated selection for an effective nitrogen-fixing symbiosis between *Trifolium spp*. and *Rhizobium leguminosarum* biovar *trifolii*. *Soil Biol Biochem* **40**, 822–833.

Yates, R.J., Howieson, J.G., Reeve, W.G. and O'Hara, G.W. (2011) A re-appraisal of the biology and terminology describing rhizobial strain success in nodule occupancy of legumes in agriculture. *Plant Soil* **348**, 255–267.

Zhang, Y.M., Tian, C.F., Sui, X.H., Chen, W.F. and Chen, W.X. (2012) Robust markers reflecting phylogeny and taxonomy of rhizobia. *PLoS ONE* **7**, e44936.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1 Sequence accessions.