RESEARCH ARTICLE

Rhizobium leguminosarum is the symbiont of lentils in the Middle East and Europe but not in Bangladesh

M. Harun-or Rashid[1,2], Javier Gonzalez[1], J. Peter W. Young[3] & Michael Wink[1]

[1]Department of Biology, Institute of Pharmacy and Molecular Biotechnology, Heidelberg University, Heidelberg, Germany; [2]Soil Microbiology Laboratory, Soil Science Division, Bangladesh Institute of Nuclear Agriculture (BINA), Mymensingh, Bangladesh; and [3]Department of Biology, University of York, York, UK

Correspondence: Michael Wink, Department of Biology, Institute of Pharmacy and Molecular Biotechnology, Heidelberg University, Im Neuenheimer Feld 364, 69120 Heidelberg, Germany.
Tel.: +49 6221 544847;
fax: +49 6221 544884;
e-mail: wink@uni-hd.de

Received 15 March 2013; revised 5 August 2013; accepted 7 August 2013.
Final version published online 19 September 2013.

DOI: 10.1111/1574-6941.12190

Editor: Kornelia Smalla

Keywords
rhizobia; legume; lentil greges; speciation; recombination.

Abstract

Lentil is the oldest of the crops that have been domesticated in the Fertile Crescent and spread to other regions during the Bronze Age, making it an ideal model to study the evolution of rhizobia associated with crop legumes. Housekeeping and nodulation genes of lentil-nodulating rhizobia from the region where lentil originated (Turkey and Syria) and regions to which lentil was introduced later (Germany and Bangladesh) were analyzed to determine their genetic diversity, population structure, and taxonomic position. There are four different lineages of rhizobia associated with lentil nodulation, of which three are new and endemic to Bangladesh, while Mediterranean and Central European lentil symbionts belong to the *Rhizobium leguminosarum* lineage. The endemic lentil grex *pilosae* may have played a significant role in the origin of these new lineages in Bangladesh. The presence of *R. leguminosarum* with lentil at the center of origin and in countries where lentil was introduced later suggests that *R. leguminosarum* is the original symbiont of lentil. Lentil seeds may have played a significant role in the initial dispersal of this *Rhizobium* species within the Middle East and on to other countries. Nodulation gene sequences revealed a high similarity to those of symbiovar *viciae*.

Introduction

Nitrogen is an essential nutrient for all living organisms and necessary for high crop yield and plant quality in agriculture, but only prokaryotes can convert atmospheric nitrogen into forms that are usable to plants. Rhizobia are nitrogen-fixing soil bacteria that are able to enter a mutual symbiosis with leguminous plants that fully or partially satisfy the nitrogen demand of the host plant. During the infection process, rhizobia produce a number of host-specific factors, and thus, it has been assumed that rhizobia have coevolved with their host plants (Perret *et al.*, 2000; Martinez-Romero, 2009). *Rhizobium leguminosarum* is a cosmopolitan and well-studied species in the genus *Rhizobium*. The name *R. leguminosarum* was first proposed by Frank (1889) for all nodule-forming bacteria, and the species currently has three biovars that differ in their host plant specificity (Jordan, 1984).

Lentil (*Lens culinaris*) is the oldest crop that was domesticated in the Fertile Crescent around 9000 years ago (Zohary & Hopf, 2000; Toklu *et al.*, 2009) and remains popular worldwide for human nutrition and for soil fertility management (Sarker & Erskine, 2006; Sonnante *et al.*, 2009). The region of origin encompasses southeastern Turkey and northern Syria, including the sources of the rivers Tigris and Euphrates (Lev-Yadun *et al.*, 2000) and cultivation spread to Cyprus and, via the Danube, to Europe around 7000 years ago. In Georgia, lentil was cultivated 4000–5000 years ago and was transported to the Indian subcontinent around 2000–2500 years ago (Erskine, 1997; Sonnante *et al.*, 2009). Rhizobia can be carried on the testa of seeds (Perez-Ramírez *et al.*, 1998), allowing them to disperse to different geographical regions along with the seeds.

Nucleotide sequences of the 16S rRNA genes are widely used as genetic markers for bacterial classification but, for a more precise identification and description of closely related bacterial species, multilocus sequence analysis (MLSA) using different protein-coding genes has become the preferred method (Ludwig & Klenk, 2005; Konstantinidis *et al.*, 2006; Martens *et al.*, 2008). Phylogenies

FEMS MICROBIOLOGY ECOLOGY

inferred from chromosomal genes and plasmid-encoded symbiosis genes of rhizobia are frequently found to be incongruent (Sprent, 1994; Laguerre *et al.*, 1996; Young & Haukka, 1996). Recombination occurs frequently in bacteria and plays an important role in the evolution of bacteria (Vinuesa *et al.*, 2005; Bailly *et al.*, 2006; den Bakker *et al.*, 2008; Tian *et al.*, 2012).

The diversity of rhizobia from pea, faba bean, and vetches has been studied previously (Laguerre *et al.*, 1996; Mutch & Young, 2004; Hou *et al.*, 2009; and many others). It has been concluded that *R. leguminosarum* is the main nodulating species in the tribe Vicieae (Tian *et al.*, 2010; and references therein), although related species such as *R. pisi* and *R. fabae* have also been described (Ramírez-Bahena *et al.*, 2008; Tian *et al.*, 2008). In contrast, there are relatively few studies on rhizobia that nodulate lentil (Hynes & O'Connell, 1990; Moawad & Beck, 1991; Laguerre *et al.*, 1992a, b; Geniaux & Amarger, 1993; Materon *et al.*, 1995; Rashid *et al.*, 2009), and we found three distinct species-level lineages related to *R. etli and R. phaseoli* (López-Guerrero *et al.*, 2012) in Bangladesh, while *R. leguminosarum* was absent (Rashid *et al.*, 2012). It is therefore important to examine lentil symbionts from other geographic regions to establish whether lentils are exceptional from other legumes of the tribe Vicieae in having different symbionts. We compared DNA sequences of rhizobia isolated from three countries where the grex *pilosae* is absent (Barulina, 1930; Sarker & Erskine, 2006) with those of isolates isolated previously (Rashid *et al.*, 2012) from field-grown *pilosae* (A. Sarker, M. Rahman and M.A. Samad, pers. commun.) lentils in Bangladesh. The aim of this study was (i) to explore the genetic diversity and identity of lentil-nodulating rhizobia, (ii) to evaluate the levels of genetic diversity and population structure of these bacteria from different geographical locations.

Materials and methods

Soil samples, plant growth, and nodule separation

Rhizobia were isolated from nodules of lentils (variety BINA-3, grex *pilosae*) grown in potted soil under glasshouse conditions and from field-grown lentils (variety Anicia, small green lentil, originally imported from France, W. Mammel, pers. commun.). Field-grown nodules were collected from Lauterach, Baden-Württemberg, Germany. Soil samples were collected from nine locations in Germany, one in Turkey, and two locations in Syria (Supporting Information, Table S1). All soil samples were collected from cultivated soils, except for one that had been collected from a forest in Germany. Soil samples were kept well separated and processed for growing plants under glasshouse conditions within 3–12 days of collection. About 2.5–3.0 kg of soil was transferred to a surface-sterilized plastic pot to grow lentils. One pot per locality was used to grow 2–3 lentil plants for 5 weeks.

Surface-sterilized (1 min in 70% ethanol and 3–5 min in 3% NaOCl) and pregerminated (48 h on 1% water agar) lentil seeds were then placed on potted soil. After germination, a maximum of three plants were cultivated for 5 weeks. Plants were irrigated alternately (i.e. water, then N-free seedling solution, then water, etc.) with sterile water and nitrogen-free seedling solution when needed. After 5 weeks, plants were uprooted carefully, the roots washed with water, dried with tissue paper, and then preserved on silica gel until further processing.

Isolation of Bacteria

For rhizobial isolation, we selected 2–11 pink nodules from lentils of each pot representing a single location, and a maximum of five pink nodules/plant from field-grown lentil. We isolated and purified rhizobia from the selected nodules following standard protocols using CRYEMA (yeast extract mannitol agar medium with congo red) described in Somasegaran & Hoben (1994). Each colony was purified by repeated streaking on CRYEMA medium and preserved at −80 °C with 25% glycerol and at 4 °C on agar slants for further study.

Determination of rhizobial population from collected soil

The number of rhizobial cells in 1 g of collected soil sample was determined following standard protocols described by Brockwell (1963). Fivefold serial dilution with three replicates for each soil sample and four replicates for each dilution were used for plant inoculation. Seeds were surface-sterilized and germinated on agar plates, and the seedlings were transferred to growth medium as described in Rashid *et al.* (2012). Harvested plants were scored for the presence or absence of nodules. The rhizobial population density of soil samples was determined following the MPN table (Brockwell, 1963; and references therein).

Determination of soil pH

Air-dried soil samples (200 g) were first ground and then sieved (2 mm) to remove large particles. From the sieved sample, 10 g was used to determine the pH. From each locality, the soil pH was measured using 0.01 mM $CaCl_2$ following the protocol ISO 2006 (International Standard Organization; www.iso.org) with a HANNA pH meter (HI 98150).

Nodulation and cross-inoculation tests

All isolates were tested for nodule formation with lentils under growth chamber conditions following standard protocol (Somasegaran & Hoben, 1994). A randomly selected set of 30 isolates was used for cross-inoculation tests with *Lathyrus sativus* and *Pisum sativum*. Lentil (Teller 'Linsen' from Müller's Mühle, commercial grade, from Germany), grass pea (Binamasur-1 from Bangladesh), and pea (variety unknown, commercial grade from Germany) were surface-sterilized using 70% ethanol (1 min) and 3% NaOCl (3–5 min). After sterilization, seeds were washed with sterile water (six times), imbibed in sterile water for 4 h, and then allowed to germinate on 1% water agar for 48 h. Germinated seeds were then transferred to glass tubes (32 mm × 170 mm) containing Fåhreus (1957) medium. Glass tubes containing germinated seed were wrapped with aluminum foil and placed in a growth chamber set to 25 °C temperature with 14-h light/10-h dark cycles for 5 weeks. There were three replications for each isolate, and uninoculated plants were kept as a control.

DNA isolation, PCR, and DNA sequencing

Bacterial isolates were grown at 28 °C for 24–36 h in tryptone-yeast (TY) medium (Beringer, 1974). DNA was extracted following Chen & Kuo (1993) and dissolved in TE buffer. DNA concentration and purity were measured by UV spectrophotometry. PCR amplifications were carried out with about 50 ng of purified DNA. Primer sequences and PCR conditions for the sequencing of four housekeeping genes (16S rRNA gene, *recA*, *atpD*, and *glnII*) and two nodulation genes (*nodC* and *nodD*) are provided in Table S1. For sequencing, PCR products were precipitated by addition of one volume of 4 M ammonium acetate with 10 volumes of 100% ice-cold ethanol and pelleted by centrifugation (16 060 ***g*** for 30 min at 4 °C). The pellet was washed with 70% ethanol and dried at 65 °C for 10 min. The purified PCR products were dissolved in ultrapure HPLC-grade water before sequencing. Sequencing was performed using an ABI 3730 automated capillary sequencer (Applied Biosystems) with the ABI Prism Big Dye Terminator Cycle Sequencing Ready Reaction Kit version 3.1 (carried out by STARSEQ GmbH, Mainz, Germany). Sequences generated in this study are deposited in GenBank under the accession numbers KC679411–KC679680 (Table S4).

Phylogenetic analyses

Sequences were aligned using CLUSTALW in BIOEDIT version 7.1.3 (Hall, 1999) with manual adjustment. We reconstructed phylogenetic trees using the neighbor-joining (NJ) and maximum likelihood (ML) algorithms in MEGA version 5 (Tamura *et al.*, 2011). The general time reversible (GTR) model of sequence evolution with gamma distribution was used in ML analysis. Bootstrap support for each node was evaluated with 1000 replicates. Protein-coded housekeeping gene trees (*recA*, *atpD*, *gln II,* and the tree from their concatenated sequence) were rooted with sequences from *Rhizobium yuanmingense* and 16S rRNA gene tree was rooted with *Mesorhizobium loti*. Uncorrected genetic distances (*p*-distance) between different phylogenetic sublineages and the type strain of *R. leguminosarum* were estimated in MEGA version 5 (Tamura *et al.*, 2011). Neighbor-net (Bryant & Moulton, 2004) analyses were conducted with the program SPLITSTREE4 version 4.12.3 (Huson & Bryant, 2006). Genes were analyzed separately and together in a combined data set.

Population genetic analyses

Parameters such as recombination events and gene flow were measured with DNASP version 5.10.01 (Rozas *et al.*, 2010). The population structure was evaluated with STRUCTURE version 2.3.3 (Pritchard *et al.*, 2000; Falush *et al.*, 2003) from housekeeping gene sequences. The most likely number of clusters (K = 1–10) was determined under an 'admixture' model, 20 000 'burn-in' and 50 000 sampling iterations following the procedure described by Evanno *et al.* (2005), which gave the highest peak against eight assumed populations. With an estimated K = 8, five extra long runs of 20 000 'burn-in' and 100 000 sampling iterations were performed. Hierarchical analysis of molecular variance (AMOVA; Excoffier *et al.*, 1992) was conducted in Arlequin version 3.5 (Excoffier *et al.*, 2005).

Recombination and mutation analyses

Levels of recombination, mutation rates, and 50% majority rule consensus trees (with and without recombination) were estimated from housekeeping genes using CLONALFRAME version 1.1 (Didelot & Falush, 2007). Three independent runs were performed with a 100 000 'burn-in' and 300 000 sampling iterations. Satisfactory MCMC convergence was judged following the criterion of (Gelman & Rubin, 1992).

We made use of three approaches to estimate the level of recombination in samples from Germany, Turkey, and Syria: (i) minimal intragenic recombination events (Rm) were detected and compared with expected values of coalescence simulations based on 10 000 genealogy replications at 95% confidence level (Hudson *et al.*, 1992; Rozas *et al.*, 2010) analyzing single genes and the combined data set in DNASP version 5.10.1 (Rozas *et al.*, 2010) (ii) the

Shimodaira–Hasegawa (S-H) test (Shimodaira & Hasegawa, 1999) was performed to compare ML tree topologies for phylogenetic congruence as implemented in TREE-PUZZLE version 5.2 (Schmidt *et al.*, 2003); and (iii) recombination rates were determined by the relative impact of recombination as compared with point mutations in the genetic diversification of the lineages (*r*/*m* proportion; Guttman & Dykhuizen, 1994) and the relative frequency of the occurrence of recombination as compared with point mutation in the history of the lineage (ρ/θ proportion; Milkman & Bridges, 1990); these analyses were carried out in CLONALFRAME version 1.1 as described before.

Results

Bacterial isolation, soil pH, and rhizobial population density

A total of 98 rhizobial colonies were isolated from lentil nodules representing 12 localities in three countries: Germany ($N = 78$), Turkey ($N = 12$), and Syria ($N = 8$), and different genes were sequenced from 58 isolates (Tables S1 and S2). In addition, we included seven previously isolated strains from Bangladesh (Rashid *et al.*, 2012) in the analyses for species delineation and comparison. A single colony was isolated from each of the selected nodules. Isolates identifier and the corresponding localities are documented in Table S1. Soil pH ranged around neutral (pH 6.5–7.4), with the exception of one forest soil sample from Heidebuckelweg, Heidelberg that presented an acidic pH of 4.8 (Table S1). Rhizobial population density varied across different localities in Germany, from 114 cells g^{-1} soil in Bürstadt (Hessen) to 2.18×10^3 cells g^{-1} soil in Ostrach (Baden-Württemberg; Table S1).

Nodulation, cross-inoculation, and symbiotic effectiveness test

Nodulation efficiency test showed that all 98 isolates were able to form nodules with lentil within 3–4 weeks after inoculation under growth chamber conditions. In a cross-inoculation test, a set of 30 randomly selected isolates were able to form nodules with both *Lathyrus sativus* and *Pisum sativum* under the same growth conditions. All isolates produced dark pink nodules, and plant leaves were darker green compared to the noninoculated controls, demonstrating that all isolates were symbiotically effective.

Phylogenetic analyses based on housekeeping gene sequences

We amplified the 16S rRNA gene (about 1500 bp length) and obtained sequences of about 1100–1350 bp from 38 rhizobial isolates originating from three different countries. BLAST searches indicated high similarities (99–100%) to *R. leguminosarum*. We recovered very similar topologies using different tree reconstruction methods, that is, NJ and ML. Phylogenetic analyses based on 16S rRNA gene sequences revealed that all isolates from the three different geographical origins were closely related to *R. leguminosarum* but separate from the Bangladeshi isolates (Supporting Information, Fig. S1).

Phylogenetic analyses based on three individual housekeeping genes, *recA* (415 bp from 58 isolates s), *atpD* (472 bp from 55 isolates), and *glnII* (598 bp from 57 isolates; Figs S2–S4), and their concatenated sequence (Fig. 1) recovered five sublineages (IVa–IVe) with high bootstrap support (70–90%). However, the recovered tree topologies differed in phylogenetic analyses, that is, some isolates changed their positions between different sublineages depending on the analyzed gene, revealing phylogenetic incongruence among the loci (Tables S2 and S8). Defined by high bootstrap support, two new sublineages (IVb and IVc) differ from sublineages of *R. leguminosarum* described previously from various hosts and geographic regions (Tian *et al.*, 2010). Phylogenetic analyses and genetic distances of three protein-coding genes (Valverde *et al.*, 2006) with respect to the type strain of *R. leguminosarum* (Fig. 1, Figs S2–S4 and Table S5) suggest that all isolates belong to *R. leguminosarum*.

Species delineation and recombination visualization using neighbor-net analysis

Neighbor-network analysis based on the combined data set (*recA-atpD-glnII*) showed a reticulate structure, in which we identified five sublineages (Fig. 2). Two sublineages (IVb and IVc) are distinguishable from other described *R. leguminosarum* sublineages (Tian *et al.*, 2010). In neighbor-network analysis, by a long edge, lentil isolates from Bangladesh (lineages I, II, and III) differed significantly from German, Turkish, and Syrian isolates, and they did not form any reticulate structure among themselves (Fig. S5). The isolates GLR7, GLR45, TLR14, and TLR10, which lay outside the main phylogenetic clusters (Fig. 1), had unique positions in the network with a high level of reticulation, indicating that they are potentially recombinant for one or more genes.

Genetic diversity analyses using STRUCTURE

With five long runs in STRUCTURE, we determined the optimal number of clusters K to be 8 and found admixture among populations (Fig. 3a and b). However, we obtained admixed structures in isolates GLR7, GLR23, GLR31, GLR33, GLR40, GLR45, GLR67, GLR74, TLR7,

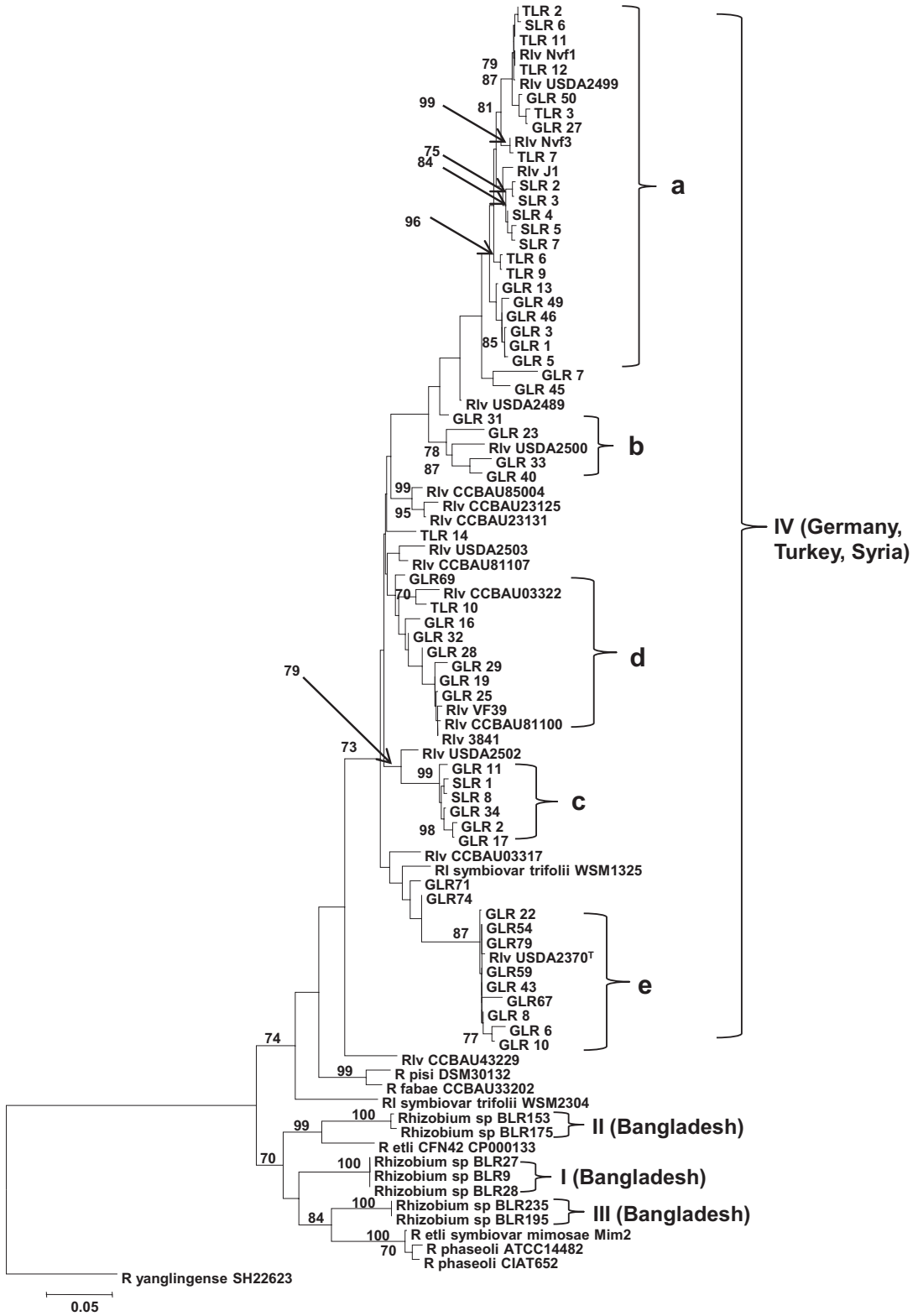

Fig. 1. ML tree from concatenated partial sequences of *recA-atpD-glnII* genes. Bootstrap values shown next to the corresponding branch when > 70% (1000 replicates). GLR, German lentil rhizobia; TLR, Turkish lentil rhizobia; SLR, Syrian lentil rhizobia; R, *Rhizobium*; Rlv, *R. leguminosarum* symbiovar *viciae*; I–IV, lineages; a–e; sublineages within lineage IV (*Rhizobium leguminosarum*).

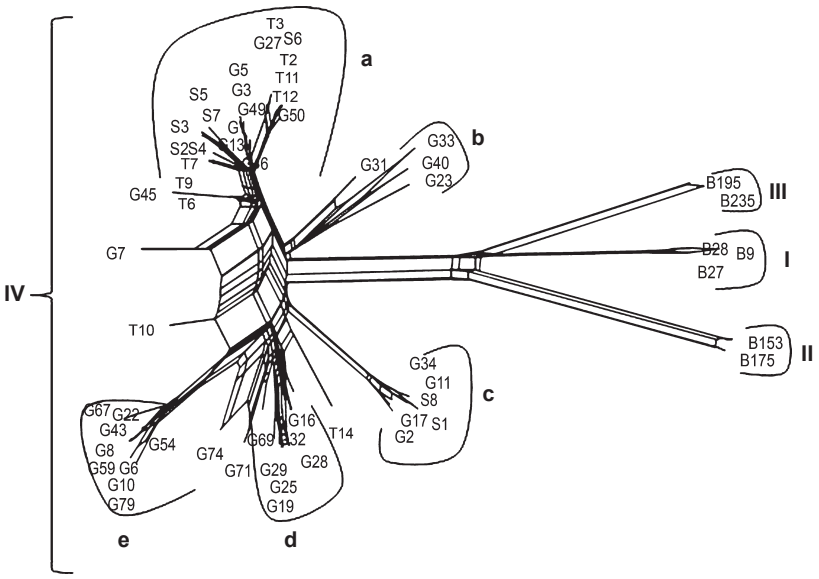

Fig. 2. Split graph from neighbor-network analysis based on concatenated sequence of *recA-atpD-glnII* genes of lentil symbionts. B, Bangladeshi lentil rhizobia; G, German lentil rhizobia; T, Turkish lentil rhizobia; S, Syrian lentil rhizobia; I–IV, lineages; a–e, sublineages within lineage IV (*Rhizobium leguminosarum*).

TLR10, and TLR14. In contrast, the new lineages from Bangladesh apparently did not show any admixture (Fig. 3a and b).

Detection of minimum recombination events and gene flow using DnaSP

Minimum recombination events found in single gene sequences and concatenated sequences of three genes are shown in Table 1. The three housekeeping genes, *recA*, *atpD,* and *glnII*, reveal 12, 14, and 16 recombination events, respectively, and concatenated data showed 43 recombination events. High values (Nm = 4.34) of gene flow were found between Turkish and Syrian isolates, along with nonsignificant K_{ST} values (0.039), while low values (Nm = 1.91) for the same parameters were found between Germany vs. Turkey and Syria (Table S6).

Differentiation between geographical groups

AMOVA of the concatenated sequence of protein-coding genes (*recA-atpD-gln II*) from samples from Germany, Turkey, and Syria indicated the existence of significant differences between geographical regions (Table S6), although the percentage of variation remained low among populations (11–20%) compared to the variation found within populations (80–90%). However, there were no significant differences between Turkish and Syrian isolates ($P > 0.05$; Table S6). Overall, AMOVAs showed that German samples differed significantly from Turkish and Syrian samples, with Turkish samples the most different (Table S6).

Relative impact of recombination and point mutation

We determined the relative effect of recombination vs. point mutations from *recA*, *atpD,* and *gln II* gene sequences, which were used to make concatenated sequence. The r/m was 1.52 and the ρ/θ was 0.269, suggesting a greater importance of recombination over mutation for explaining the observed genetic diversity (Table 2). The reconstructed dendrograms from ClonalFrame analyses revealed much shorter times to the most recent common ancestor (TMRCA) when considering recombination (TMRCA = 0.220 (0.130–0.352), Fig. 4, Table S7) than when assuming no recombination (TMRCA = 1.110 (0.342–3.073) Fig. S6, Table S7). The tree topology also differed between these phylograms. For example, the position of the sublineage e and polytomies in the sublineages a and e (Fig. 4) are well resolved assuming no recombination (Fig. S6). However, we obtained more polytomies in the dendrogram when considering recombination (Fig. 4).

Symbiotic gene analyses

The maximum likelihood analyses of nodulation gene sequences recovered five groups (based on *nodC* from 38

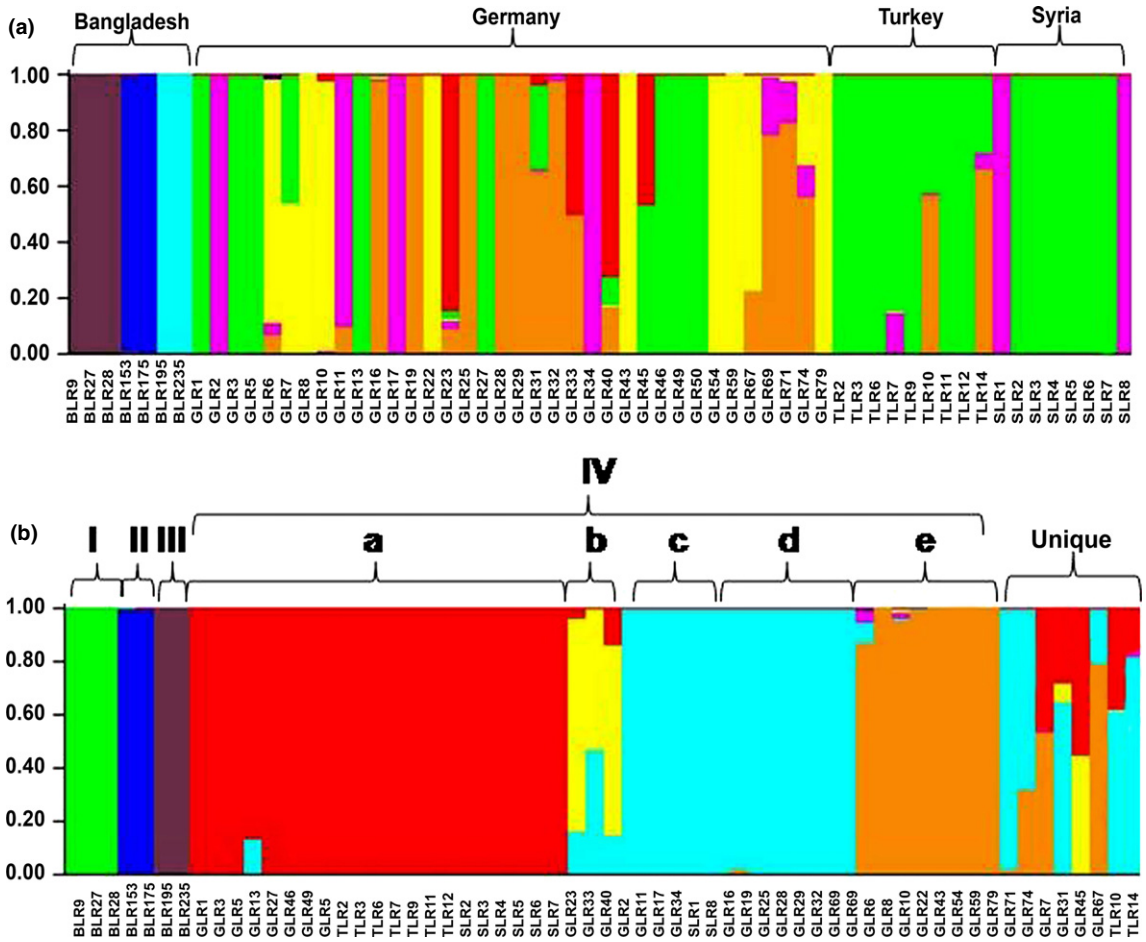

Fig. 3. Ancestry of different lineages and sublineages as inferred by STRUCTURE assuming $K = 8$ ancestral populations. Individual strains are marked on the horizontal axis, and their corresponding column is filled with color according to the inferred proportion, which was derived from one of the ancestral populations. Bar plot (a) showing strains ordered by geographical position, and bar plot (b) showing strains arranged by lineages and sublineages. GLR, German lentil rhizobia; TLR, Turkish lentil rhizobia; SLR, Syrian lentil rhizobia; I–III, lineages; a–e, sublineages within lineage IV (*Rhizobium leguminosarum*).

Table 1. Minimum recombination events on different genes

				Coalescence simulation		
Gene	*N*	Gene size (bp)	Rm	Rm average	Confidence interval	*P* < observed Rm
recA	58	415	12	8.65	5–13	0.945
atpD	55	472	14	3.65	1–6	1.00
glnII	57	598	16	5.83	2–10	1.00
Combined	53	1485	43	17.02	10–24	1.00

Rm, minimal intragenic recombination events; *N*, number of samples; bp, base pair.

isolates and *nodD* from 23 isolates) from Germany, Turkey, and Syria (Fig. 5 and Fig. S7). In *nodC* gene analysis, group A contained lentil isolates from Germany, Turkey, and Syria. This group clusters with previously described strains isolated from Peru, Spain, and United Kingdom from different members of the legume tribe Vicieae. This group includes the isolate BLR195, which was isolated from lentils in Bangladesh and belongs to lineage III, suggesting a clear case of horizontal transfer of nodulation genes between lineage III and IV. In same tree (*nodC*), group B and D contained isolates from Germany only, while groups C and E contained isolates from Turkey (except GLR10) and Syria, respectively. Among seven groups (A–E from Germany, Turkey, and Syria, and

Table 2. Recombination effect inferred by ClonalFrame (confidence intervals are shown between parentheses)

Run	*R*	*r/m*	ρ/θ	θ
1	26 (12–47)	1.479 (0.833–2.330)	0.252 (0.112–0.471)	109 (71–160)
2	31 (14–57)	1.611 (0.908–2.544)	0.294 (0.125–0.550)	109 (72–160)
3	26 (12–46)	1.490 (0.855–2.408)	0.262 (0.121–0.502)	101 (67–146)
Average	28 (13–50)	1.520 (0.865–2.428)	0.269 (0.119–0.507)	106 (70–155)

R, recombination rate; *r/m*, relative impact of recombination as compared with point mutation; ρ/θ, relative frequency of the occurrence of recombination as compared with point mutation; θ, mutational rate.

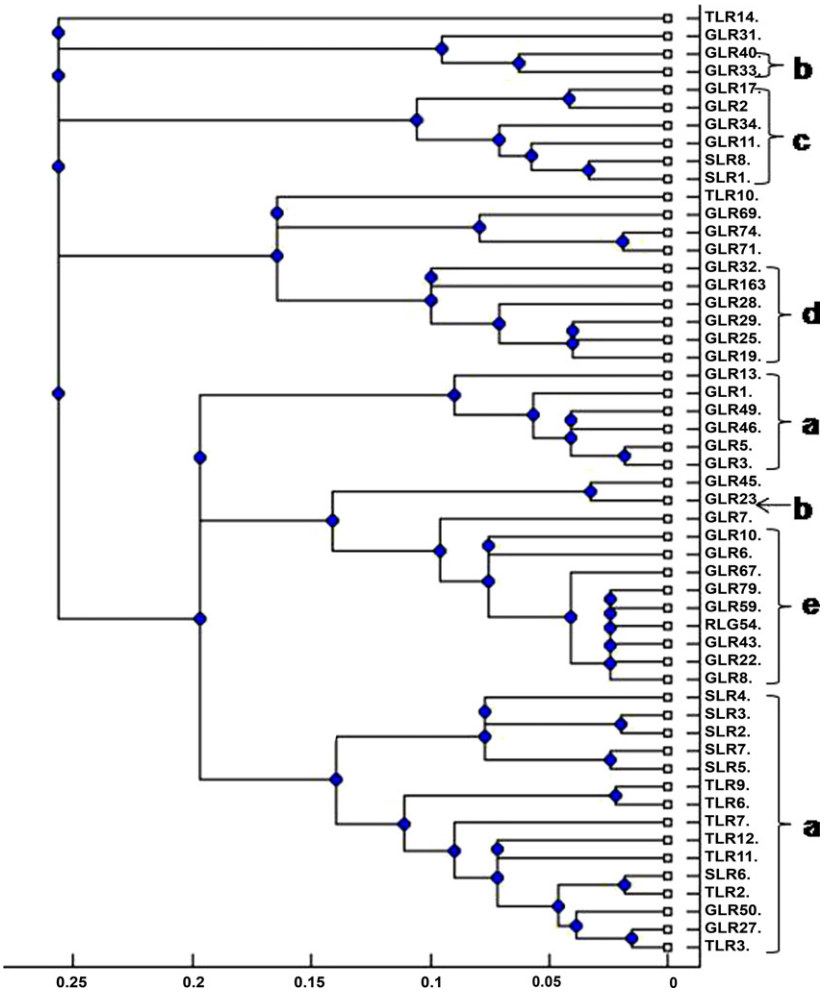

Fig. 4. Majority rule consensus tree (50%) inferred from concatenated partial sequence of *recA-atpD-glnII* genes using CLONALFRAME allowing recombination. The scale indicates the time in coalescent units. GLR, German lentil rhizobia; TLR, Turkish lentil rhizobia; SLR, Syrian lentil rhizobia; a–e, sublineages within lineage IV (*Rhizobium leguminosarum*).

groups I–II from Bangladesh), the group C from Turkey and the group II from Bangladesh differed considerably from existing *nodC* sequences of *R. leguminosarum* symbiovar *viciae* strains. Moreover, two isolates with unique *nodC* sequences, GLR2 and SLR2, showed significant differences to existing strains.

We sequenced *nodD* gene from 24 isolates (11 from Germany, six from Turkey, six from Syria, and one from Bangladesh) to compare with the *nodC* gene analysis to determine whether they were congruent or not. The reconstructed ML tree from the *nodD* gene sequences showed similar tree topologies (except for the isolates

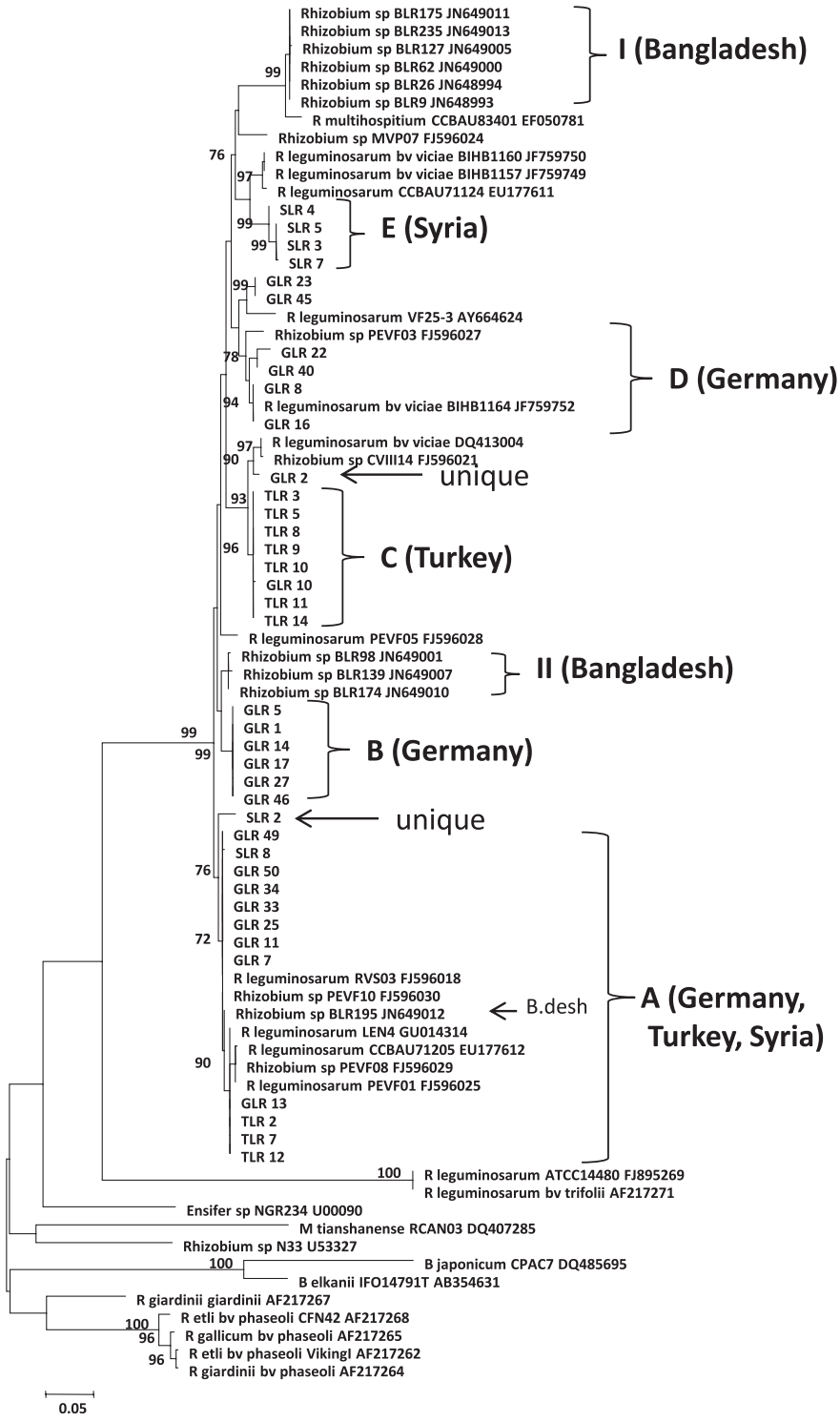

Fig. 5. Unrooted ML tree from partial *nodC* gene sequences. Bootstrap value shown next to the corresponding branch when > 70% (1000 replicates). GLR, German lentil rhizobia; TLR, Turkish lentil rhizobia; SLR, Syrian lentil rhizobia; BLR, Bangladeshi lentil rhizobia; Rlv, *Rhizobium leguminosarum* symbiovar *viciae*; Rl, *R. leguminosarum*; A–E, nodulation gene group from German Turkish and Syrian isolates; I–II, nodulation gene group from Bangladeshi isolates.

GLR17 and SLR 4) to the one based on *nodC* gene sequences (Fig. S5). The exception of GLR17 and SLR4 isolates may arise from internal rearrangement of the nod region by recombination. Group A corresponds to the previously described *nodD* type II/*nodD* type g (Laguerre *et al.*, 2003; Mutch & Young, 2004; Tian *et al.*, 2010 and reference therein) from faba bean rhizobia from different geographical locations (Jordan, Spain, Canada, and UK). Group B showed similarity to the previously described *nodD* type III from faba bean rhizobia from Europe (France and UK) and China (Tian *et al.*, 2010). Although Turkish isolates form a separate group (group C) in the *nodC* gene tree, this group showed similarity to a previously described *nodD* type I from the Middle East and China, suggesting rearrangement of *nodD* gene within the nod region by recombination. Although the *nodC* group from Syria (E) was close to previously described strains, in the *nodD* gene tree, this group formed a strong separate group from existing *nodD* groups. The isolate GLR17 had an identical sequence to a distinct strain found in France from pea rhizobia. Although Bangladeshi isolates formed a strongly separate group in *nodC* gene, members of this group were very close to the previously described *nodD* type IV (Tian *et al.*, 2010) in *nodD* gene. In terms of topology, the nodulation gene trees differed from trees based on chromosomal genes, suggesting horizontal/lateral gene transfer events of nodulation genes among different lineages and sublineages.

Discussion

Genetic diversity of chromosomal genes and species delineation of lentil rhizobia

We sequenced three protein-coding housekeeping genes (*recA*, *atpD,* and *glnII*) that contain valuable information for determining biogeographic patterns (Palys *et al.*, 1997; Lan & Reeves, 2001; Vinuesa *et al.*, 2005) and identified five sublineages (IVa–IVe) among the 58 isolates from three countries (Turkey, Syria, and Germany). All sublineages belonged to *R. leguminosarum,* and two sublineages (IVb & IVc) did not show any great similarity to the sublineages that were described earlier within *R. leguminosarum* (Tian *et al.*, 2012). We found phylogenetic incongruence in reconstructed trees from chromosomal genes, suggesting an influence of recombination on housekeeping genes. Consistent with this, results from different analyses (S-H test, recombination analyses, estimation of TMRCA) showed a substantial level of recombination among sublineages (Fitch, 1997; Shimodaira & Hasegawa, 1999; Bryant & Moulton, 2004; Tian *et al.*, 2010).

Network analysis allow conflicting or alternative phylogenetic histories to account for ambiguities caused by recombination, hybridization, gene conversion, and gene transfer (Fitch, 1997). From this analysis, we obtained a clear reticulate structure among different sublineages from three different countries, but a long edge between Bangladeshi isolates, and samples from the other three countries. In other words, network analysis showed a clear difference between lineage IV and the others by a long edge, suggesting that lineage IV belongs to separate species (Bailly *et al.*, 2006). Lentil-nodulating rhizobia from Bangladesh were well separated by long edges from *R. leguminosarum* (those nodulating lentils in the Mediterranean region and Central Europe) and always formed three lineages distinct from closely related *R. etli* and *R. phaseoli* with no incongruence in phylogenetic analyses. This is evidence that the Bangladeshi samples belong to separate, currently undescribed, species. Phylogenetic analyses (ML and BI) of three housekeeping genes and their concatenated sequence found three distinct species-level lineages in Bangladesh without any incongruence. These lineages also showed substantial differences, not only from closely related species, but also among themselves. Moreover, these lineages also showed phenotypic differences from closely related species (Rashid *et al.*, 2012). Based on protein-coding genes (*recA*, *atpD* and *glnII*), lineage IV is genetically similar to the *R. leguminosarum* type strain (> 94%) and forms a reticulate structure with *R. leguminosarum* in network analysis suggesting (Bailly *et al.*, 2006; Valverde *et al.*, 2006; Santillana *et al.*, 2008) that these sublineages belong to *R. leguminosarum*.

Origin and distribution of new lineages in Bangladesh are influenced by symbiosis with lentil grex *pilosae*

For cultivated lentils, Barulina (1930) proposed six geographical groups, or greges, *viz. europeae*, *asiaticae*, *intermediae*, *subspontanea*, *aethiopicae,* and *pilosae* (Cubero, 1981; Cubero *et al.*, 2009; and references therein). Interestingly, among the six greges of cultivated lentil, three groups are restricted to very specific areas. For instance, *pilosae* is endemic to the Indian subcontinent, *aethiopicae* to Ethiopia and Yemen, and *subspontanea* to Afghanistan. This study and others on lentil rhizobia (Hynes & O'Connell, 1990; Moawad & Beck, 1991; Laguerre *et al.*, 1992a, b; Geniaux & Amarger, 1993; Materon *et al.*, 1995) show clearly that *R. leguminosarum* is the main symbiont of all lentil greges except *pilosae* in the Indian subcontinent.

Although Eastern Turkey and Northern Syria are the area of lentil domestication (Barulina, 1930; Ladizinsky, 1979; Lev-Yadun *et al.*, 2000; Zohary & Hopf, 2000; Cubero *et al.*, 2009), the Himalaya-Hindu Kush may corresponds to the center of origin for small seeded *microsperma* lentils because of the presence of a higher

proportion of endemic varieties in this region. *Pilosae* has a strong pubescence, which is absent in other lentils (Cubero *et al.*, 2009), and a flowering asynchrony, and has not overlapped with any other greges of lentil during the history of domestication and cultivation (Barulina, 1930; Erskine *et al.*, 1994). This grex may also have specific genetic characteristics like nod factor receptors that allow for a successful symbiosis with the new *Rhizobium* species or lineages found in Bangladesh. Thus, *pilosae* and their symbionts may have coevolved in the Indian subcontinent. It is therefore possible that we identified new lineages (Rashid *et al.*, 2012) of rhizobia from Bangladesh due to the significant influence of the *pilosae* grex on their symbiotic partners.

Rhizobium leguminosarum is the original symbiont of lentils

In agreements with other studies (Moawad & Beck, 1991; Laguerre *et al.*, 1992a, b; Geniaux & Amarger, 1993; Moawad *et al.*, 1998; Tegegn, 2006), we found *R. leguminosarum* in the center of origin of lentil and in countries where lentil had been introduced (e.g. Germany). We not only found the same *R. leguminosarum* species but also the same chromosomal genotype (e.g. sublineage IVa) in three different countries. Hence, it could be assumed that *R. leguminosarum* is the original symbiont of lentils. The dispersal of rhizobia with legume seeds is a well-accepted hypothesis (Perez-Ramírez *et al.*, 1998; Aguilar *et al.*, 2004; Álvarez-Martinez *et al.*, 2009). This mode is considered to be the most important among the indirect ways of rhizobium dispersal (Hirsch, 1996; and references therein). Lentil is the oldest crop, has remained popular from ancient times until now, and is found all over the world. Thus, it could be assumed that lentil seeds might have played a significant role in initial dispersal of *R. leguminosarum* symbiovar *viciae* to different countries. However, of the three distinct lineages, one lineage (lineage I) is found all over Bangladesh (Rashid *et al.*, 2012), suggesting that it may have been distributed with *pilosae* seeds.

Phylogenetic incongruence between chromosomal and nodulation genes

Rhizobial species diversity should be described not only based on genetic markers located on chromosomes but also on plasmid-borne nodulation genes (Graham *et al.*, 1991; Amarger *et al.*, 1997; Wang *et al.*, 1999; Laguerre *et al.*, 2001; Silva *et al.*, 2005). The *nodC* and *nodD* genes determine host range, host promiscuity, and the relationships between host plants and rhizobia (Laguerre *et al.*, 2001; Zeze *et al.*, 2001; Iglesias *et al.*, 2007). In our study, phylogenetic analyses based on nodulation genes confirm that these isolates belong to symbiovar *viciae,* and that they have *nodC* and *nodD* genotypes which have previously been described from Europe, Middle East, and China. We also detected new groups within this symbiovar. There was no congruence between chromosomal and nodulation genotypes, which may be due to frequent lateral transfer of nodulation genes between different rhizobial chromosomal genotypes (Sprent, 1994; Young & Haukka, 1996). However, phylogenetic analyses based on nodulation genes (*nodC* and *nodD*) revealed similar tree topologies correlated mostly with ecological regions. Probably, as a consequence of both host and soil microhabitats (Sprent, 1994), we detected new nodulation genotypes (or groups) that were supported by both nodulation genes. These results support the hypothesis that plasmid-borne characters in bacteria change rapidly according to their adaptation to particular environment. Nonetheless, nodulation genes from the Europe and the Middle East did not show great similarities to the isolates from Bangladesh, suggesting that the latter have an independent origin on the Indian subcontinent. A similar hypothesis has been proposed for the origin of *R. etli* (Aguilar *et al.*, 2004).

Conclusions

By analyzing lentil-nodulating rhizobia from four countries on two different continents, we found four different lineages of rhizobia, of which three are new. These three new lineages of rhizobia are endemic to Bangladesh and have probably coevolved with the lentil grex *pilosae*. The presence of common genotypes of *R. leguminosarum* with lentil in different countries suggests that *R. leguminosarum* is the original symbiont of lentil. Additional research is needed to further examine the genetic diversity and population structure of lentil-nodulating rhizobia in different geographical regions.

Acknowledgment

Erasmus Mundus Mobility with Asia (EMMA) provided a PhD scholarship to MHR, and additional funding was provided by the DAAD-Stipendien-und Betreuungsprogramm (STIBET) grant, Heidelberg University. We are grateful to Dr. Martin Krehenbrink for fruitful discussion while writing this manuscript. We thank Dr. Holger Schäfer, Heidi Staudter, Annika Heinemann, Vanessa Erbe, Johannes Reiner, Beate Waibel, Razan Hamoud, Shirin Hamoud, Rasha Abou Aleinein, Dr. Tamer Albayrak, Dr. Sabine Grübe, Woldae Mamel, and Markus Santhosh Braun for their help in collecting samples, and Hedwig Sauer-Gürth and Heidi Staudter (IPMB) and Michael Schilbach (COS) for technical assistance. This work has been performed according to German laws.

References

Aguilar OM, Riva O & Peltzer E (2004) Analysis of *Rhizobium etli* and of its symbiosis with wild *Phaseolus vulgaris* supports coevolution in centers of host diversification. *P Natl Acad Sci USA* **101**: 13548–13553.

Álvarez-Martinez ER, Valverde A, Ramírez-Bahena MH, García-Fraile P, Tejedor C, Mateos PF, Santillana N, Zúñiga D, Peix A & Velázquez E (2009) The analysis of core and symbiotic genes of rhizobia nodulating *Vicia* from different continents reveals their common phylogenetic origin and suggests the distribution of *Rhizobium leguminosarum* strains together with *Vicia* seeds. *Arch Microbiol* **191**: 659–668.

Amarger N, Macheret V & Laguerre G (1997) *Rhizobium gallicum* sp. nov. and *Rhizobium giardinii* sp. nov. from *Phaseolus vulgaris* nodules. *Int J Syst Bacteriol* **47**: 996–1006.

Bailly X, Olivieri I, de Mitas S, Cleyet-Marel J & Béna G (2006) Recombination and selection shape the molecular diversity pattern of nitrogen-fixing *Sinorhizobium* sp. associated to *Medicago*. *Mol Ecol* **15**: 2719–2734.

den Bakker HC, Didelot X, Fortes ED, Nightingale KK & Wiedmann M (2008) Lineage specific recombination rates and microevolution in *Listeria monocytogenes*. *BMC Evol Biol* **8**: 277–289.

Barulina H (1930) Lentil of the U.S.S.R. and of other countries. *Bull Appl Bot Plant Breed Suppl* vol. 1. Leningrad, Vsesoi u znai a akademii a sel'sko khozi a istvennyk, Russia.

Beringer J (1974) R factor transfer in *Rhizobium leguminosarum*. *J Gen Microbiol* **84**: 88–198.

Brockwell J (1963) Accuracy of a plant-infection technique for counting populations of *Rhizobium trifolii*. *Appl Microbiol* **11**: 377–383.

Bryant D & Moulton V (2004) Neighbor-net: an agglomerative method for the construction of phylogenetic networks. *Mol Biol Evol* **21**: 255–265.

Chen WP & Kuo TT (1993) A simple and rapid method for the preparation of gram negative bacterial genomic DNA. *Nucleic Acids Res* **21**: 2260.

Cubero JI (1981) Taxonomy, distribution and evolution of the lentil and its wild relatives. *Lentils* (Witcombe J & Erskine W, eds), pp. 15–38. M.Nijhoff & W. Junk Publishers, Boston, MA.

Cubero JI, de la Vega MP & Fratini R (2009) Origin, taxonomy and domestication. *The lentil: Botany, Production and Uses* (Erskine W, Muehlbauer F, Sarker A & Sharma B, eds), pp. 13–33. CAB International, Oxfordshire, UK.

Didelot X & Falush D (2007) Inference of bacterial microevolution using multilocus sequence data. *Genetics* **175**: 1251–1266.

Erskine W (1997) Lessons for breeders from land races of lentil. *Euphytica* **93**: 107–112.

Erskine W, Smartt J & Muehlbauer FJ (1994) Mimicry of lentil and the domestication of common vetch and grass pea. *Econ Bot* **48**: 326–332.

Evanno G, Regnaut S & Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* **14**: 2611–2620.

Excoffier L, Somouse P & Quattro J (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131**: 479–491.

Excoffier L, Laval G & Schneider S (2005) Arlequin ver. 3.5: an integrated software package for population genetics data analysis. *Evol Bioinform* **1**: 47–50.

Fåhreus G (1957) The infection of clover root hairs by nodule bacteria studied by a simple glass slide technique. *J Gen Microbiol* **16**: 374–381.

Falush D, Stephens M & Pritchard J (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* **164**: 1567–1587.

Fitch W (1997) Networks and viral evolution. *J Mol Evol* **44**: 65–75.

Frank B (1889) *Über die Pilzsymbiose der Leguminosen. Berichte Deutschen Botanischen Gesellschaft*. Verlag von Paul Parey, Berlin.

Gelman A & Rubin DB (1992) Inference from iterative simulation using multiple sequences. *Stat Sci* **7**: 457–511.

Geniaux E & Amarger N (1993) Diversity and stability of plasmid transfer in isolates from a single field population of *Rhizobium leguminosarum* bv. *viciae*. *FEMS Microbiol Ecol* **102**: 251–260.

Graham PH, Sadowsky MJ & Keyser HH (1991) Proposed minimal standards for the description of new genera and species of root- and stem-nodulating bacteria. *Int J Syst Evol Microbiol* **41**: 582–587.

Guttman DS & Dykhuizen DE (1994) Clonal divergence in *Escherichia coli* as a result of recombination, not mutation. *Science* **266**: 1380–1383.

Hall TA (1999) BioEdit: a user friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* **41**: 95–98.

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

Hou BC, Wang ET, Li Y, Jia RZ, Chen WF, Man CX, Sui XH & Chen WX (2009) Rhizobial resource associated with epidemic legumes in Tibet. *Microb Ecol* **57**: 69–81.

Hudson RR, Slatkin M & Maddison WP (1992) Estimation of levels of gene flow from DNA sequence data. *Genetics* **132**: 583–589.

Huson DH & Bryant D (2006) Application of phylogenetic networks in evolutionary studies. *Mol Biol Evol* **23**: 254–267.

Hynes MF & O'Connell MP (1990) Host plant effect on competition among strains of *Rhizobium leguminosarum*. *Can J Microbiol* **36**: 864–869.

Iglesias O, Rivas R, García-Fraile P, Abril A, Mateos PF, Martinez-Molina E & Velázquez E (2007) Genetic characterization of fast growing rhizobia able to nodulate *Prosopis alba* in North Spain. *FEMS Microbiol Lett* **277**: 210–216.

Jordan DC (1984) Family III. *Rhizobiaceae*. *Bergey's Manual of Systematic Bacteriology*, Vol. 1 (Krieg NR & Holt JG, eds), pp. 234–242. Williams and Wilkins, Baltimore.

Konstantinidis KT, Ramette A & Tiedje JM (2006) Toward a more robust assessment of intraspecies diversity, using fewer genetic markers. *Appl Environ Microb* **72**: 7286–7293.

Ladizinsky G (1979) The origin of lentil and its wild genepool. *Euphytica* **28**: 179–187.

Laguerre G, Geniaux E, Marzurier S, Casartelii R & Amarger N (1992a) Conformity and diversity among field isolates of *R. leguminosarum* bv. *viciae*, bv. *trifolii*, and bv. *phaseoli* revealed by DNA hybridization using chromosome and plasmid probes. *Can J Microbiol* **39**: 412–419.

Laguerre G, Mazurier SI & Amarger N (1992b) Plasmid profiles and restriction fragment length polymorphism of *Rhizobium leguminosarum* bv. *viciae* in field populations. *FEMS Microbiol Ecol* **101**: 17–26.

Laguerre G, Mavingui P, Allard MR, Charnay MP, Louvrier P, Mazurier SI, Ligottier-Gois L & 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 Microb* **62**: 2029–2036.

Laguerre G, Nour SM, Macheret V, Sanjuan J, Drouin P & Amarger N (2001) Classification of rhizobia based on *nodC* and *nifH* gene analysis reveals a close phylogenetic relationship among *Phaseolus vulgaris* symbionts. *Microbiology* **147**: 981–993.

Laguerre G, Louvrier P, Allard MR & Amarger N (2003) Compatibility of rhizobial genotypes within natural populations of *Rhizobium leguminosarum* biovar *viciae* for nodulation of host legumes. *Appl Environ Microb* **69**: 2276–2283.

Lan R & Reeves P (2001) When does a clone deserve a name? A perspective on bacterial species based on population genetics. *Trends Microbiol* **9**: 419–424.

Lev-Yadun S, Gopher A & Abbo S (2000) The cradle of agriculture. *Science* **288**: 1602–1603.

López-Guerrero MG, Ormeño-Orrillo E, Velázquez E, Rogel MA, Acosta JL, Gónzalez V, Martínez J & Martínez-Romero E (2012) *Rhizobium etli* taxonomy revised with novel genomic data and analyses. *Syst Appl Microbiol* **35**: 353–358.

Ludwig W & Klenk HP (2005) Overview: a phylogenetic backbone and taxonomic framework for prokaryotic systematic. *Bergey's Manual of Systematic Bacteriology*, Vol. 2 (Brenner DJ, Krieg NR & Staley J, eds), pp. 49–65. Springer, New York, NY.

Martens M, Dawyndt P, Coopman R, Gillis M, Vos PD & Willems A (2008) Advantages of multilocus sequence analysis for taxonomic studies: a case study using housekeeping genes in the genus *Ensifer* (including former *Sinorhizobium*). *Int J Syst Evol Microbiol* **58**: 200–214.

Martinez-Romero E (2009) Coevolution in Rhizobium-legume symbiosis? *DNA Cell Biol* **28**: 361–370.

Materon LA, Keatinge JDH, Beck DP, Yurtsever N, Karuc K & Altuntas S (1995) The role of rhizobial biodiversity in legume crop productivity in the West Asian highlands. *Exp Agr* **31**: 485–491.

Milkman R & Bridges MM (1990) Molecular evolution of the *E. coli* chromosome III Clonal frames. *Genetics* **126**: 505–517.

Moawad H & Beck DP (1991) Some characteristics of *Rhizobium leguminosarum* isolates from un-inoculated field-grown lentil. *Soil Biol Biochem* **23**: 933–937.

Moawad H, Badr El-Din SMS & Abdel-Aziz RA (1998) Improvement of biological nitrogen fixation in Egyptian winter legumes through better management of Rhizobium. *Plant Soil* **204**: 95–106.

Mutch LA & Young JPW (2004) Diversity and specificity of *Rhizobium leguminosarum* biovar *viciae* on wild and cultivated legumes. *Mol Ecol* **13**: 2435–2444.

Palys T, Nakamura LK & Cohan FM (1997) Discovery and classification of ecological diversity in the bacterial world: the role of DNA sequence data. *Int J Syst Bacteriol* **47**: 1145–1156.

Perez-Ramírez ON, Rogel MA, Wang E, Castellanos ZJ & Martínez-Romero E (1998) Seeds of *Phaseolus vulgaris* bean carry *Rhizobium etli*. *FEMS Microbiol Ecol* **26**: 289–296.

Perret X, Staehelin C & Broughton WJ (2000) Molecular basis of symbiotic promiscuity. *Microbiol Mol Biol Rev* **64**: 180–201.

Pritchard J, Stephens M & Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959.

Ramírez-Bahena M, García-Fraile P, Peix A, Valverde A, Rivas R, Igual JM, Mateos PF, Martínez-Molina E & Velázquez E (2008) Revision of the taxonomic status of the species *Rhizobium leguminosarum* (Frank 1879) Frank 1889AL, *Rhizobium phaseoli* Dangeard 1926AL and *Rhizobium trifolii* Dangeard 1926AL. *R. trifolii* is a later synonym of *R. leguminosarum*. Reclassification of the strain *R. leguminosarum* DSM 30132 (=NCIMB 11478) as *Rhizobium pisi* sp.nov. *Int J Syst Evol Microbiol* **58**: 2484–2490.

Rashid MH, Sattar MA, Uddin MI & Young JPW (2009) Molecular characterization of symbiotic root nodulating rhizobia isolated from lentil (*Lens culinaris*). *EJEAFChe (ISSN:15794377)* **8**: 602–612.

Rashid MH, Schäfer H, Gonzalez J & Wink M (2012) Genetic diversity of rhizobia nodulating lentil (*Lens culinaris*) in Bangladesh. *Syst Appl Microbiol* **35**: 98–109.

Rozas J, Sanchez-DelBarrio JC, Messeguer X & Rozas R (2010) DnaSP: DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19**: 2496–2497.

Santillana N, Ramírez-Bahena MH, García-Fraile P, Velázquez E & Zuniga D (2008) Phylogenetic diversity based on rrs, *atpD, recA* genes and 16S–23S intergenic sequence analyses of rhizobial strains isolated from *Vicia faba* and *Pisum sativum* in Peru. *Arch Microbiol* **189**: 239–247.

Sarker A & Erskine W (2006) Recent progress in the ancient lentil. *J Agr Sci* **144**: 19–29.

Schmidt HA, Petzold E, Vingron M & von Haeseler A (2003) Molecular phylogenetics: parallelized parameter estimation

and quartet puzzling. *J Parallel Distrib Comput* **63**: 719–727.
Shimodaira H & Hasegawa M (1999) Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol Biol Evol* **16**: 1114–1116.
Silva C, Vinuesa P, Eguiarte L, Souza V & Martínez-Romero E (2005) Evolutionary genetics and biogeographic structure of *Rhizobium gallicum sensu lato*, a widely distributed bacterial symbiont of diverse legumes. *Mol Ecol* **14**: 4033–4050.
Somasegaran P & Hoben HJ (1994) *Handbook for Rhizobia: Methods in Legume-Rhizobium Technology*. Springer, Heidelberg.
Sonnante G, Hammer K & Pignone D (2009) From the cradle of agriculture a handful of lentils: history of domestication. *Rend Lincei* **20**: 2137.
Sprent JI (1994) Evolution and diversity in the legume-rhizobium symbiosis: chaos theory? *Plant Soil* **161**: 1–10.
Tamura K, Peterson D, Peterson N, Stecher G, Nei M & Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* **28**: 2731–2739.
Tegegn ND (2006) *Genetic Diversity and Characterization of Indigenous Rhizobium leguminosarum* biovar viciae Isolates of Cool-season Food Legumes Growth in the Highlands of Ethiopia. Thesis University of Putra, Malaysia.
Tian CF, Wang ET, Wu LJ *et al.* (2008) *Rhizobium fabae* sp. nov., a bacterium that nodulates *Vicia faba*. *Int J Syst Evol Microbiol* **58**: 2871–2875.
Tian CF, Young JPW, Wang ET, Tamimi SM & Chen WX (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.
Tian CF, Zhou YJ, Zhang YM *et al.* (2012) Comparative genomics of rhizobia nodulating soybean suggests extensive recruitment of lineage-specific genes in adaptations. *P Natl Acad Sci USA* **29**: 8629–8634.
Toklu F, Karakoy T, Hakli E, Bicer T, Brandolini A, Kilian B & Özkan H (2009) Genetic variation among lentil (*Lens culinaris* Medik) land races from Southeast Turkey. *Plant Breeding* **128**: 178–186.
Valverde A, Igual JM, Peix A, Cervantes E & Velazquez E (2006) *Rhizobium lusitanum* sp. nov. a bacterium that nodulates *Phaseolus vulgaris*. *Int J Syst Evol Microbiol* **56**: 2631–2637.
Vinuesa P, Silva C, Werner D & Martı′nez-Romero E (2005) Population genetics and phylogenetic inference in bacterial molecular systematics: the roles of migration and recombination in *Bradyrhizobium* species cohesion and delineation. *Mol Phylogenet Evol* **34**: 29–54.
Wang ET, Rogel MA & García-de los Santos A (1999) *Rhizobium etli* bv. mimosae, a novel biovar isolated from *Mimosa affinis*. *Int J Syst Evol Microbiol* **49**: 1479–1491.
Young JPW & Haukka K (1996) Diversity and phylogeny of rhizobia. *New Phytol* **133**: 87–94.
Zeze A, Mutch LA & Young JPW (2001) Direct amplification of *nodD* from community DNA reveals the genetic diversity of *Rhizobium leguminosarum* in soil. *Environ Microbiol* **3**: 363–370.
Zohary D & Hopf M (2000) *Domestication of Plants in the Old World*. Oxford University Press, New York.

Supporting Information

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

Fig. S1. ML tree from partial 16S rRNA sequences.
Fig. S2. ML tree from partial *recA* gene sequences.
Fig. S3. ML tree from partial *atpD* gene sequences.
Fig. S4. ML tree from partial *gln II* gene sequences.
Fig. S5. Split graph from neighbor-network analysis based on concatenated sequence of *recA-atpD-glnII* genes of Lentil symbionts plus previously described strains of *R. leguminosarum. R. etli* and other related species.
Fig. S6. Majority rule consensus tree (50%) inferred from concatenated partial sequence of *recA-atpD-glnII* genes using CLONALFRAME without allowing recombination.
Fig. S7. Un-rooted ML tree from partial *nodD gene* sequences.
Table S1. Isolate numbers, soil pH and rhizobial density for different sample collection localities.
Table S2. Isolates under different lineages and sub-lineages in different analyses.
Table S3. List of primers and PCR conditions.
Table S4. GenBank accession numbers.
Table S5. Genetic distances among different sub-lineages (SL).
Table S6. Hierarchical AMOVA of the genetic structure and gene flow of sub-lineages.
Table S7. Time to the most recent common ancestor (TMRCA) of all samples using CLONALFRAME.
Table S8. Phylogenetic incongruence using Shimodaira-Hasegawa (S-H) test.