Identification of soybean *Bradyrhizobium* strains used in commercial inoculants in Brazil by MALDI-TOF mass spectrometry

Lucas Rolim¹ · Thaís Ribeiro Santiago² · Fábio Bueno dos Reis Junior³ · Ieda de Carvalho Mendes³ · Helson Mario Martins do Vale¹ · Mariangela Hungria⁴ · Luciano Paulino Silva^{1,2}

Received: 23 February 2019 / Accepted: 8 June 2019 / Published online: 25 June 2019
© Sociedade Brasileira de Microbiologia 2019

Abstract

Biological nitrogen fixation (BNF) with the soybean crop probably represents the major sustainable technology worldwide, saving billions of dollars in N fertilizers and decreasing water pollution and the emission of greenhouse gases. Accordingly, the identification of strains occupying nodules under field conditions represents a critical step in studies that are aimed at guaranteeing increased BNF contribution. Current methods of identification are mostly based on serology, or on DNA profiles. However, the production of antibodies is restricted to few laboratories, and to obtain DNA profiles of hundreds of isolates is costly and time-consuming. Conversely, the matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) MS technique might represent a golden opportunity for replacing serological and DNA-based methods. However, MALDI-TOF databases of environmental microorganisms are still limited, and, most importantly, there are concerns about the discrimination of protein profiles at the strain level. In this study, we investigated four soybean rhizobial strains carried in commercial inoculants used in over 35 million hectares in Brazil and also in other countries of South America and Africa. A supplementary MALDI-TOF database with the protein profiles of these rhizobial strains was built and allowed the identification of unique profiles statistically supported by multivariate analysis and neural networks. To test this new database, the nodule occupancy by *Bradyrhizobium* strains in symbiosis with soybean was characterized in a field experiment and the results were compared with serotyping of bacteria by immuno-agglutination. The results obtained by both techniques were highly correlated and confirmed the viability of using the MALDI-TOF MS technique to effectively distinguish bacteria at the strain level.

Keywords *Bradyrhizobium elkanii* · *Bradyrhizobium japonicum* · *Bradyrhizobium diazoefficiens* · Biotyper · ClinPro Tools

Responsible Editor: Luiz Roesch

The authors declare that this work has not been published previously and it is not under consideration for publication elsewhere. Its publication is approved by all authors, and it will not be published elsewhere, in English or in any other language, including electronically, without the written consent of the copyright holder.

✉ Luciano Paulino Silva
luciano.paulino@embrapa.br

Lucas Rolim
roliml@gmail.com

Thaís Ribeiro Santiago
tatasantiago@gmail.com

Fábio Bueno dos Reis Junior
fabio.reis@embrapa.br

Ieda de Carvalho Mendes
ieda.mendes@embrapa.br

Helson Mario Martins do Vale
helson@unb.br

Mariangela Hungria
mariangela.hungria@embrapa.br

¹ Universidade de Brasília (UnB), Brasília, Distrito Federal 70910-900, Brazil

² Embrapa Recursos Genéticos e Biotecnologia, C.P. 02372, Brasília, Distrito Federal 70770-917, Brazil

³ Embrapa Cerrados, C.P. 08223, Planaltina, Distrito Federal 73310-970, Brazil

⁴ Embrapa Soja, C.P. 231, Londrina, Paraná 86001-970, Brazil

Introduction

Soybean (*Glycine max* L. Merr.) is a key crop for the Brazilian economy, which has highly benefited from *Bradyrhizobium* spp. strain selection programs performed for over 5 decades [4, 13]. As soybean is an exotic crop in Brazil, inoculation is imperative when cropped for the first time, and it is noteworthy that reinoculation every year also brings benefits, with average increases of 8% in grain yield; similar situations are found in Argentina and other countries of South America [13].

Some of the main challenges in field experiments performed worldwide with inoculation of legumes are the detection, identification, and monitoring of the inoculant strains occupying the nodules. For decades, the traditional method for strain identification in soybean nodules from field trials has been the evaluation of serological properties (e.g., [9, 17–19, 33, 36]). However, the production of antibodies is challenging and restricted to few laboratories; in addition, there are increasing bioethical concerns about animal sacrifice, demanding the replacement of the method [1]. Serological methods also encompass limitations, such as cross-reactions and problems with antiserum stability, which could lead to false-negative results [29]. In addition, Mpeperek and Wollum [21] reported that even nodules formed by a single occupant with distinctive serology may elicit multiple serological reactions. These methodological issues limit the effectiveness of serology in monitoring legume-rhizobia symbioses.

New possibilities have arisen with the advances in DNA-based techniques, such as the identification of DNA fingerprints by repetitive element palindromic PCR (rep-PCR) [12], by multilocus sequence typing (MLST) [32], and by the use of genetic markers [25], among others. However, these methodologies are labor-consuming and cost-limited when hundreds of isolates have to be studied.

Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) is an example of a high-throughput method that allows the comparison of protein fingerprints obtained from different microbial cells based on reference spectral data, by using various algorithms integrated into commercially available systems. In recent years, this approach has been increasingly used for phenotyping and identifying several species of microorganisms (e.g., [37, 38]), including rhizobia [8, 27, 40]. It is a simple technique with few refinements in sample preparation, allowing the identification of a large number of microorganisms even in samples derived from complex mixtures [16]. However, one limitation relies on a robust validation of the applicability of environmental microorganisms, as the great majority of the studies and databases available are related to pathogenic human bacteria and fungi (e.g., [7, 15, 26, 39]).

Studies on ecology of rhizobia, including those that monitor nodule occupancy under field conditions, investigate shifts

in rhizobial population with changes in soil management, and determine diversity, among others, could be highly speeded up by MALDI-TOF MS. However, for that, one main limitation is to investigate if MALDI-TOF MS also shows an acceptable accuracy at the strain level.

The aim of this study was to verify the feasibility of obtaining MALDI-TOF MS-specific profiles that could assign different rhizobia strains. For that, we constructed a supplementary MALDI-TOF MS database and investigated four different strains of *Bradyrhizobium* spp. (SEMIA 587, SEMIA 5019 (29W), SEMIA 5079 (CPAC 15), and SEMIA 5080 (CPAC 7)) used in commercial inoculants for the soybean crop in Brazil and nowadays established in more than 35 million hectares in Brazil, as well as in other countries of South America and Africa. Additionally, this new supplementary MALDI-TOF MS database was tested for the evaluation of nodule occupancy by the *Bradyrhizobium* strains in symbiosis with soybean in a field experiment, and the results were compared with serotyping of nodule bacteria using immuno-agglutination.

Materials and methods

Bradyrhizobium strains

The elite strains of *B. elkanii* SEMIA 587 and SEMIA 5019, *B. japonicum* SEMIA 5079, and *B. diazoefficiens* SEMIA 5080 used in this study are deposited at the Embrapa Cerrados Collection of Multifunctional Microorganisms (WFCC Culture Collection # 776). These strains have been used in commercial inoculants since the 1980s (SEMIA 587 and SEMIA 5019) and 1990s (SEMIA 5079 and SEMIA 5080) [23, 24]. In addition, strains SEMIA 587, SEMIA 5019, SEMIA 5079, and SEMIA 5080 are associated with the serogroups SEMIA 587, 29W, USDA 123, and CB 1809, respectively [2].

Preparation of samples for MALDI-TOF MS

Bradyrhizobium spp. strains were grown on TY solid medium (Sigma-Aldrich, St. Louis, Missouri, USA). After 4 days of incubation at 28 °C, individual colonies were transferred to 1.5-mL microcentrifuge tubes with a sterilized toothpick and homogenized with 300 µL of sterilized deionized water. Subsequently, 700 µL of absolute ethanol (Sigma-Aldrich, St. Louis, Missouri, USA) was added to the mixture, and the samples were centrifuged at 15,500g for 2 min. After this, the supernatant was discarded and the pellet was dried at room temperature for 1 h. In sequence, 30 µL of formic acid 70% (Sigma-Aldrich, St. Louis, Missouri, USA) and 30 µL of acetonitrile (Sigma-Aldrich, St. Louis, Missouri, USA) were added, followed by vigorous homogenization and then

centrifugation at 15,500g for 2 min. In the next step, 1 µL of the supernatant was placed onto spots of a stainless steel MALDI target plate type MSP 96 (Bruker Daltonics GmbH, Bremen, Germany) and allowed to air-dry at room temperature. After, 1 µL of alpha-cyano-4-hydroxycinnamic acid matrix (Sigma-Aldrich, St. Louis, Missouri, USA), prepared with 250 µL of acetonitrile (Sigma-Aldrich, St. Louis, Missouri, USA), 200 µL of type I water (Milli-Q® Advantage A10 Water Purification System, Millipore, Burlington, Massachusetts, USA), and 50 µL of 3% trifluoroacetic acid (Sigma-Aldrich, St. Louis, Missouri, USA), was applied to each spot and air-dried at room temperature prior to the analysis in the mass spectrometer. A total of 24 individual extractions were obtained for each strain and deposited in the MALDI target plates in two replicates.

MALDI-TOF MS

A microFlex LRF MALDI-TOF MS (Bruker Daltonics GmbH, Bremen, Germany) was used for all mass spectral acquisitions. The instrument employs a nitrogen laser at 337 nm, and it was applied at an intensity of 20–65% in the positive linear mode. The acquisition rate in the linear mode used was 20 Hz with the mass/charge ratios detected at the range of m/z 2000–20,000. The laser was shot 240 times split into six steps of 40 shots that start from different target positions. We accumulated these shots only when there were ionized regions of the samples aiming to obtain a final spectrum. The acquisitions were carried out in an automatic acquisition mode to scan in each spot through large spirals, and the peaks were detected using the centroid algorithm according to the manufacturer's default settings (signal/noise ratio, 2; minimum intensity limit, 600; maximum number of peaks, 300). The equipment was calibrated using as standard an *Escherichia coli* protein extract (Bruker Daltonics GmbH, Bremen, Germany).

Supplementary database

The representative main spectrum profile (MSP) of each strain was obtained using the average of up to 48 spectra originated from the 24 individual procedures of protein extraction replicated twice in the MALDI target plates. From these 48 spectra, a minimum of 40 were used to generate the list of m/z and intensity signals, which make up the reference spectrum. The addition of standard spectra was performed using the tool “creating mass spectral profiles” (mass spectral profile (MSP)) available in the MALDI Biotyper software version 3.0 and library version 3.3.1.2 updated with 5989 entries (Bruker Daltonics GmbH, Bremen, Germany) which comprises the default parameters stated by the manufacturer (minimum frequency required for peak selection, 25%; maximum number of desired peaks to form the MSP, 70).

Spectral visualization and statistical analysis

Mass spectra were acquired and processed using the software FlexAnalysis version 3.3 (Bruker Daltonics GmbH, Bremen, Germany). Data analysis started with a pretreatment of the raw data, subtracting baseline and normalizing a set of spectra by peak alignment using prominent peaks, and a peak picking procedure. Subsequently, the mass spectra were opened in the ClinPro Tools software version 2.2 (Bruker Daltonics GmbH, Bremen, Germany) for statistical analysis. This software was used for the comparison and inspection of mass spectra, as well as for the identification of specific peaks in order to further discriminate the analyzed strains.

The spectra obtained for each strain at a range of m/z 2000 to 20,000 were analyzed using the ClinPro Tools. Moreover, to validate each statistical analysis, at least 40 spectra of each strain were analyzed with the software and automatically recalibrated with a specific algorithm that is already available in the ClinPro Tools. As reported in the manufacturer's standard procedure, the recalibration function is capable to reduce possible molecular mass shifts that occurred during multiple measurements of the same sample.

After the recalibration process, the software automatically created an average spectrum for all replicates of each strain. The average spectra obtained were compared with each other, and a report of peaks was generated. In this case, only the presence/absence of peaks was taken into account; their absolute intensity was not considered at this moment. Further, the statistical analysis of the data sets was based on multivariate statistics centered on the principal component analysis (PCA), and the results were represented by a 3D dot plot, a peak distribution 2D view, and a conventional PCA ordination plot, which were automatically generated by the software. The PCA reduces the variables of the complex set of data, generating a set of new variables called principal components (PCs) [14]. After the PCA, an analysis of neural networks using the genetic algorithm was also applied towards comparisons [3].

Evaluation of nodule occupancy by *Bradyrhizobium* strains in symbiosis with soybean in a field experiment

A field experiment was conducted at the Embrapa Cerrados experimental station, located close to the city of Planaltina, Federal District, Brazil (15° 35' 30" S and 47° 42' 0" W, at an altitude of 1175 m). The experiment was established in November 2008 in an Oxisol cultivated with soybean, after the removal of its original vegetation (Cerradão). Soil analysis (2014/2015 crop season) is presented in Table 1. Before sowing, the area was fertilized according to the standard recommendation: 500 kg ha⁻¹ of N-P-K (0-20-20); other macro- and micronutrients were

Table 1 Soil chemical properties in the study area. Samples were collected from the 0- to 20-cm depth before the 2014/2015 crop season

pH (H ₂ O)	H + Al cmol _c dm ⁻³	Al ³⁺ dm ⁻³	Ca	Mg	K mg kg ⁻¹	P mg kg ⁻¹	Soil organic C g kg ⁻¹
5.7	5.7	0.1	2.2	0.8	70.0	9.8	18.6

supplied whenever necessary, according to the soil analysis. The experiment was conducted in a complete randomized block design with four replications, in 17 × 4 m plots, consisting of nine rows spaced by 50 cm. In the first year (2008/2009 crop season), seeds were inoculated with *B. japonicum* SEMIA 5079 or *B. diazoefficiens* SEMIA 5080. In the following crop seasons (2009/2010, 2010/2011, 2011/2012, 2012/2013, 2013/2014, and 2014/2015), the plots originally inoculated with SEMIA 5079 or SEMIA 5080 were subdivided into three subplots of 5 × 4 m spaced by 1 m, where the soybean started to be cultivated under three treatments: (a) without inoculation, (b) yearly inoculation with the strain SEMIA 5079, and (c) yearly inoculation with the strain SEMIA 5080, in a total of 24 subplots (two original plots × three subplots × four repetitions). In all years, the inoculation was performed in order to obtain a concentration of approximately 1,000,000 cells/seed.

During the soybean flowering stage, in the 2014/2015 crop season, a total of six plants were collected per subplot. Twenty four bacteria were isolated according to Hungria [11] from nodules randomly selected among the plants collected from each subplot. A total of 576 isolates were analyzed and identified according to their protein mass spectra obtained by the MALDI-TOF MS technique.

Preparation of field samples for MALDI-TOF MS followed the same steps described before. The mass spectra obtained for each isolate were compared with all the standard spectra present in the MALDI Biotyper database and also with the new supplementary database previously generated with the four *Bradyrhizobium* spp. strains. The identification by comparison is based on the mass/charge ratio (*m/z*) of the ions and the relative intensity and frequency of each peak of the corresponding mass spectrum. The output of this analysis consisted of a table (report) in xml format with the species/strains with the highest similarity value (log score) ranked in the first position, with this log score value ranging from 0 to 3, following an ascending scale of similarity.

Percentages for nodule occupancy by each strain were determined by their distribution among the isolates from each subplot. For this purpose, only isolates identified as one of the four *Bradyrhizobium* strains used in the soybean inoculants were considered.

Serotyping of nodule bacteria using immuno-agglutination

The antisera used in this work, against strains SEMIA 5019 (serogroup 29 W), SEMIA 587 (serogroup SEMIA 587), SEMIA 5079 (serogroup USDA 123), and SEMIA 5080 (serogroup CB 1809), were previously prepared, tested, and validated by Mendes et al. [19]. Serotyping of nodule-occupying strains was done by immuno-agglutination [35]. Recovery percentages for serogroups 29W, SEMIA 587, CB 1809, and USDA 123 were determined by serotyping 50 nodules selected at random from each subplot.

MALDI-TOF MS and immuno-agglutination comparison

Data obtained using MALDI-TOF MS and immuno-agglutination were submitted to Pearson's correlation using R software version 3.4.3 (R Foundation for Statistical Computing, Vienna, Austria). In addition, standard deviations of the nodule occupancy means (four repetitions) obtained using both methods were calculated.

Results

Construction of a supplementary database using MALDI-TOF MS applied to the identification of *Bradyrhizobium* spp. strains

Mass spectra successfully acquired from the 24 biological replicates of each strain were obtained in the *m/z* 2000 to 20,000 range (Fig. 1). The mass spectra of each strain showed a large number of peaks, some visually coincident and others not, but the high amount of data impaired some direct evaluation. Thus, PCA was used to analyze and rationalize this large number of peaks in order to quantitatively validate some possible similarities/dissimilarities between the strains.

Principal component analysis as a tool to validate differentiation of libraries

To support and highlight any possible differences between the ion profiles of the elite strains, all spectra were exported to the ClinPro Tools software followed by automatic comparison of multiple spectra through PCA.

Figure 2 shows the PCA by using multiple representation styles, such as peak distribution 2D view and 3D dot plot window. There was no overlap between strains SEMIA 5019 and SEMIA 587, consequently allowing an unambiguous differentiation. The differentiation among these two strains and SEMIA 5079 and SEMIA 5080 was also clear. However, using these same tests, considering the group

Fig. 1 Representative standard mass spectra of the four elite strains obtained from the average of 24 distinct acquisitions

variance range for 29 W and 587, there was an overlap between strains SEMIA 5079 and SEMIA 5080; therefore, these findings required an individual analysis step solely for these two strains, considering only the variance between their

standard spectra, which finally allowed their unequivocal differentiation (Fig. 3).

According to the variance of individual peaks, it is possible to conclude that all the libraries were statistically different,

Fig. 2 Principal component analysis (ClinPro Tools) of all spectra obtained from the elite strains. Results represented by a 3D dot plot window (a), a peak distribution 2D view (b), and a conventional PCA ordination plot (c). In a and b, SEMIA 5079 is represented by yellow,

SEMIA 5080 by red, SEMIA 587 by blue, and SEMIA 5019 by green dots. In c, SEMIA 5079, SEMIA 5080, SEMIA 587, and SEMIA 5019 are indicated by symbols $\diamond$, $\times$, $\square$, and $\circ$, respectively. The joint confidence ellipses encompass 95% of the data

Fig. 3 Principal component analysis (ClinPro Tools) of all spectra acquired from strains SEMIA 5079 and SEMIA 5080. Results represented by a 3D dot plot window (**a**), a peak distribution 2D view (**b**), and a conventional PCA ordination plot (**c**). In **a** and **b**, SEMIA 5079

is represented by red and SEMIA 5080 by green dots. In **c**, SEMIA 5079 and SEMIA 5080 are indicated by symbols $\times$ and $\circ$, respectively. The joint confidence ellipses encompass 95% of the data

therefore allowing the identification of the bacteria at the strain level, something beyond the major functionality of the Biolytper method.

Neural network analysis for the identification of determinants and mismatched peaks in the differentiation of libraries

In order to evaluate the level of discrimination between the four strains, an analysis of neural networks was applied.

Five remarkably discriminant peaks were identified with the observation of all libraries together, with major peaks between m/z 2000 and 12,000 (GA 100%; CV 97.85%) (Fig. 4). Similar profiles were also observed when comparing only strains SEMIA 5079 and SEMIA 5080, also with five discriminating peaks in the range of m/z 2000 and 8000 (GA 100%; CV 98.35%) (Fig. 5).

These results statistically corroborated the possibility of strain differentiation associated with the ion profiles related

Fig. 4 Neural network graphic acquired from the four investigated strains, in the range of m/z 2000 to 18,000. The total average spectrum is in gray color. The red spectrum represents the class average spectrum that is calculated within the spectral recalibration workflow. In the analysis of neural networks, the matching peaks are represented by blue bars and the five differential peaks/regions represented by red bars

Fig. 5 Neural network graphic acquired from strains SEMIA 5080 and SEMIA 5079 in the range of m/z 2000 to 18,000. The total average spectrum is in gray color. The red spectrum represents the class average spectrum that is calculated within the spectral recalibration workflow. In the analysis of neural networks, the matching peaks are represented by blue bars and the five differential peaks/regions represented by red bars

to the spectral library inserted in the MALDI-TOF MS supplementary database.

Evaluation of nodule occupancy by *Bradyrhizobium* strains in symbiosis with soybean in a field experiment

Table 2 is a compilation of the nodule occupancy (%) data taking into account the four *Bradyrhizobium* strains used in soybean inoculants and measured by immuno-agglutination and MALDI-TOF MS.

Nodule occupancy by SEMIA 5080 was generally higher in treatments where it was reinoculated annually, from the second year onwards, ranging from 52 to 55% when measured by immuno-agglutination and from 33 to 41% when evaluated by MALDI-TOF MS (Table 2). After 6 years, this strain was

not detected in nodules by MALDI-TOF MS in treatments where it had not been reintroduced by inoculation. On the other hand, immuno-agglutination showed that nodule occupancy by this strain was reduced from 52 to 35% when there was no reinoculation at all or to 11% when it was replaced by SEMIA 5079. Using immuno-agglutination, SEMIA 5080 was found in nodules from treatments where SEMIA 5079 was inoculated only in the first year and in treatments where SEMIA 5079 was reinoculated every year, presenting 27% and 12% of nodule occupancy, respectively.

SEMIA 5079 also showed higher nodule occupancy where it was reinoculated annually, from 61 to 62% when evaluated by immuno-agglutination and from 74 to 95% when evaluated by MALDI-TOF MS. In the treatments where the strain was introduced only in the first year, nodule occupancy was high when there was no introduction of another strain, at 46% when

Table 2 Inoculation with different strains and its effect on nodule occupancy evaluated by the immuno-agglutination and MALDI-TOF MS methods (means of 4 repetitions $\pm$ standard deviation)

Inoculated strain		Nodular occupation 6th year (%)							
		Immuno-agglutination				MALDI-TOF MS			
1st year	2nd year on	SEMIA 5080	SEMIA 5079	SEMIA 587	SEMIA 5019	SEMIA 5080	SEMIA 5079	SEMIA 587	SEMIA 5019
SEMIA 5080	SEMIA 5080	52 $\pm$ 12	18 $\pm$ 14	22 $\pm$ 6	7 $\pm$ 6	33 $\pm$ 2	43 $\pm$ 7	21 $\pm$ 2	3 $\pm$ 1
SEMIA 5080	Non-inoculated	35 $\pm$ 12	36 $\pm$ 25	20 $\pm$ 6	9 $\pm$ 7	0 $\pm$ 0	45 $\pm$ 3	55 $\pm$ 6	0 $\pm$ 0
SEMIA 5080	SEMIA 5079	11 $\pm$ 6	61 $\pm$ 2	16 $\pm$ 9	12 $\pm$ 8	0 $\pm$ 0	74 $\pm$ 4	20 $\pm$ 2	6 $\pm$ 1
SEMIA 5079	SEMIA 5079	12 $\pm$ 3	62 $\pm$ 7	18 $\pm$ 8	8 $\pm$ 3	0 $\pm$ 0	95 $\pm$ 8	5 $\pm$ 2	0 $\pm$ 0
SEMIA 5079	Non-inoculated	27 $\pm$ 4	46 $\pm$ 5	15 $\pm$ 8	12 $\pm$ 4	0 $\pm$ 0	86 $\pm$ 5	14 $\pm$ 3	0 $\pm$ 0
SEMIA 5079	SEMIA 5080	55 $\pm$ 7	15 $\pm$ 8	26 $\pm$ 7	4 $\pm$ 2	41 $\pm$ 3	18 $\pm$ 2	41 $\pm$ 2	0 $\pm$ 0

evaluated by immuno-agglutination and 86% when evaluated by MALDI-TOF MS. When the strain SEMIA 5079 was replaced by SEMIA 5080, its nodule occupancy decreased to 15% when evaluated by immuno-agglutination and to 18% when evaluated by MALDI-TOF MS.

SEMIA 587 and SEMIA 5019 presented percentages of nodule occupancy of 15 to 26% and 4 to 12%, respectively, when evaluated by immuno-agglutination. When determined by MALDI-TOF MS, the percentages by these strains ranged from 5 to 55% and from zero to 6%, respectively.

The decrease in nodule occupancy evaluated by immuno-agglutination when SEMIA 5080 was not reinoculated was of 33%. When it was replaced by SEMIA 5079, there was a reduction of 79% in its nodule occupancy. When evaluated by MALDI-TOF MS, these decreases were of 100% in both cases. In the case of SEMIA 5079, there was a reduction on nodule occupancy measured by immuno-agglutination of 26% when it was not reinoculated and of 76% when it was replaced by SEMIA 5080. When evaluated by MALDI-TOF, these decreases were of 9% and 81%, respectively (Table 2). These results highlight the higher persistence capacity of SEMIA 5079 when compared with SEMIA 5080, as reported before [19].

Despite the differences in the percentage of nodule occupancy when evaluated by immuno-agglutination and MALDI-TOF MS, Pearson's correlation coefficients between the two data set showed a correlation with an R -squared value (R^2) equal to 0.78, highly significant (p value of 8.18×10^{-6}), proving that both methodologies are capable to lead to similar findings.

Discussion

Recent research has highlighted the importance of using new techniques, such as MALDI-TOF MS, in ecological and taxonomic large-scale studies. Ferreira et al. [8] extended their MALDI Biotyper 2.0 database to include the type strains of 56 rhizobial species, including the genera *Rhizobium*, *Ensifer*, and *Shinella*. In addition, 35 strains isolated from different sources, previously identified on the basis of their 16S rRNA (*rrs*), *recA*, and *atpD* gene sequences, were used to validate the MALDI-TOF MS as an identification tool for this group of bacteria. In the study of Ferreira et al. [8], all strains were correctly identified at the species level, leading to the conclusion that MALDI-TOF MS may represent an excellent tool for the identification of fast- and slow-growing rhizobia and be applicable to large populations of rhizobial isolates.

Apparently, the definition of species in the *Bradyrhizobium* genus is limited by the low diversity of the 16S rRNA gene sequences, and thus a polyphasic approach is often necessary for phylogeny and taxonomy studies [5, 20]. After the

extension of the MALDI Biotyper 2.0 database to consider *Bradyrhizobium* species, Sánchez-Juanes et al. [27] were able, for the first time, to identify isolates belonging to phylogenetically closely related rhizobia. However, the capability of the MALDI-TOF MS to discriminate strains was not evaluated.

In our study, to discriminate four strains belonging to three different species of *Bradyrhizobium*, we had to improve and optimize the technique with different statistical tools. Different studies confirm the efficiency of MALDI-TOF MS to identify a variety of microorganisms at genus, species, and even subspecies level. Suarez et al. [30] studied the differentiation of *Neisseria meningitidis* strains by MALDI-TOF MS. Although no statistical approach was used to validate their data, it was evident that MALDI-TOF MS allowed a quick typing of *N. meningitidis* isolates. Ziegler et al. [40] showed the importance of a well-defined database for phylogenetic identification of rhizobia. These authors used a different MALDI-TOF MS platform (SARAMIS™) for feeding up the database with spectral profiles of 116 strains, representing the major rhizobial genera. Their approach allowed assessing the effectiveness of MALDI-TOF MS for typing rhizobial strains, providing an accurate and highly resolved analysis of strain diversity at the species and often subspecies level. In another study, Calderaro et al. [3] also showed the possibility of distinguishing viruses at the serogroup level by MALDI-TOF MS, confirming the feasibility of the technique. The authors recommended the statistical validation using the ClinPro Tools program. Following this recommendation, using PCA and neural networks, in the present study, we were able to differentiate the four commercial soybean *Bradyrhizobium* spp. strains.

The statistical approaches used in our study confirmed the feasibility of implementation of the supplementary database for rhizobial identification at the strain level, even with strains that belong to the same species (SEMIA 5019 and SEMIA 587 of *B. elkanii*). Interestingly, the identification of strains that belong to different species, SEMIA 5080 (*B. diazoefficiens*, previously classified as *B. japonicum*, [5]) and SEMIA 5079 (*B. japonicum*), occurred only after an independent statistical comparison between both strains. These findings suggest that for some cases it is not improbable that the protein profiles may differ more between two strains than between some species.

In the field experiment, strain distribution in nodules of different treatments was coincident with findings from other studies, highlighting the importance of yearly reinoculation to keep high nodule occupancy by elite strains such as SEMIA 5080 and the higher competitive and saprophytic abilities of the strain SEMIA 5079 [10, 19, 28, 31, 34]. However, our primary goal with this experiment was to compare the results obtained with MALDI-TOF MS and immuno-agglutination. While the use of MALDI-TOF MS in our study was based on the rhizobial

isolation, the classical and broadly used immuno-agglutination relies on the antigenic characteristics of bacterial mixtures directly extracted from the plant nodules. Nevertheless, the high correlation between both techniques demonstrates that they share similar accuracy and reliability. In fact, advances in the use of MALDI-TOF MS to study nodule bacteria are expected. For example, Ndungu et al. [22] showed that this technique was rapid, affordable, and reliable to identify *Bradyrhizobium* strains directly from cowpea (*Vigna unguiculata* L. Walp.) nodule suspensions, without the need of bacterial isolation.

Besides the correct identification of species, monitoring bacterial strains after their introduction in the soil has always been a major challenge. Among these studies, the identification of strains occupying the legume nodules is the key for the determination of rhizobial competitiveness [6]. The methodologies available today to assess nodule occupancy are time-consuming and expensive or require animal sacrifice. Our results showed that the use of MALDI-TOF MS represents an interesting alternative for solving traditional challenges in studies with legume-rhizobia symbioses, including the evaluation of saprophytic capacity, bacterial root colonization, nodule occupancy, the fate of inoculant microorganisms in the environment, and routine quality control of commercial inoculants.

Conclusions

The MALDI-TOF MS technique in association with statistic methods effectively distinguished the four soybean *Bradyrhizobium* strains (SEMIA 587, SEMIA 5019 (29W), SEMIA 5079 (CPAC 15), and SEMIA 5080 (CPAC 7)) currently used in commercial inoculants in Brazil and other countries of South America and Africa. The supplementary database obtained will allow important advances in ecology and agronomic and industrial studies of the soybean-rhizobia symbiosis.

Acknowledgments The authors thank the Laboratory of Mass Spectrometry of the Embrapa Recursos Genéticos e Biotecnologia for the use of the mass spectrometer and Dr. Juaci Malaquias for his help in the statistics analysis. M. Hungria, I.C. Mendes, and L.P. Silva are research fellows from CNPq.

Funding information This work was supported by Embrapa (02.13.08.001.00.00) and INCT-Plant-Growth Promoting Microorganisms for Agricultural Sustainability and Environmental Responsibility (CNPq 465133/2014-4, Fundação Araucária-STI, CAPES).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

- Andrade A, Pinto SC, Oliveira RS (2002) Animais de Laboratório: criação e experimentação. FIOCRUZ, Rio de Janeiro
- Boddey LH, Hungria M (1997) Phenotypic grouping of Brazilian *Bradyrhizobium* strains which nodulate soybean. Biol Fertil Soils 25:407–415
- Calderaro A, Arcangeletti MC, Rodighiero I, Buttrini M, Gorrini C, Motta F, Gemini D, Medici MC, Chezzi C, De Conto F (2014) Matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry applied to virus identification. Sci Rep 4:1–10
- Chang WS, Lee HI, Hungria M (2015) Soybean production in the Americas. In: Lugtenberg B (ed) Principles of plant-microbe interactions. Springer International Publishing, Basel, pp 393–400
- Delamuta JRR, Ribeiro RA, Ormeño-Orrillo E, Melo IS, Martínez-Romero E, Hungria M (2013) Polyphasic evidence supporting the reclassification of *Bradyrhizobium japonicum* group Ia strains as *Bradyrhizobium diazoefficiens* sp. nov. Int J Syst Evol Microbiol 63:3342–3351
- Dowling DN, Broughton WJ (1986) Competition for nodulation of legumes. Annu Rev Microbiol 40:191–157
- Fenselau C, Demirev PA (2001) Characterization of intact microorganisms by MALDI mass spectrometry. Mass Spectrom Rev 20: 157–171
- Ferreira L, Saánchez-Juanes F, García-Fraile P, Rivas R, Mateos PF, Martínez-Molina E, González-Buitrago JM, Velázquez E (2011) MALDI-TOF mass spectrometry is a fast and reliable platform for identification and ecological studies of species from family Rhizobiaceae. PLoS One 6:e20223. <https://doi.org/10.1371/journal.pone.0020223>
- Ham GE, Frederick LR, Anderson IC (1971) Serogroups of *Rhizobium japonicum* in soybean nodules samples in Iowa. Agron J 63:69–72
- Hungria M, Vargas MAT (1996) Exploring the microbial diversity and soil management practices to optimize the contribution of soil microorganisms to plant nutrition. In: Stacey G, Mullin B, Gresshoff PM (eds). Biology of plant-microbe interactions. International Society of Molecular Plant-Microbe Interactions, St. Paul, 493–496
- Hungria M (1994) Coleta de nódulos e isolamento de rizóbios. In: Hungria M, Araújo RS (eds). Manual de métodos empregados em estudos de microbiologia agrícola. EMBRAPA, Brasília, 157–170
- Hungria M, Campo RJ, Mendes IC (2003) Benefits of inoculation of the common bean (*Phaseolus vulgaris*) crop with efficient and competitive *Rhizobium tropici* strains. Biol Fertil Soils 39:88–93
- Hungria M, Mendes IC (2015) Nitrogen fixation with soybean: the perfect symbiosis? In: De Bruijn FJ (ed) Biological nitrogen fixation vol. 2. John Wiley and Sons Inc., New Jersey, pp 1005–1019
- Jolliffe IT (2002) Principal component analysis, 2nd edition. Springer-Verlag, New York
- Keller PM, Bruderer V, Müller F (2016) Restricted identification of clinical pathogens categorized as biothreats by MALDI-TOF mass spectrometry. J Clin Microbiol 54:816
- Lavigne JP, Espinal P, Messad N, Pantel A, Sotto A (2013) Mass spectrometry: a revolution in clinical microbiology? Clin Chem Lab Med 51:257–270
- Lima SC, Lopes ES, Lemos EGM (1998) Caracterização de rizóbios (*Bradyrhizobium japonicum*) e produtividade de soja. Sci Agric 55:45–69
- McLoughlin TJ, Alt SG, Gonzalez RG, Romero-Severson J (1991) Effects on soybean seed yield by members of 123 serocluster and USDA 110 in the greenhouse and in soils low in indigenous *Bradyrhizobium japonicum*. Can J Microbiol 37:984–988

19. Mendes IC, Vargas MAT, Hungria M (2004) Establishment of *Bradyrhizobium japonicum* and *B. elkanii* in a Brazilian Cerrados Oxisol. *Biol Fertil Soils* 40:28–35
20. Menna P, Barcellos FG, Hungria M (2009) Phylogeny and taxonomy of a diverse collection of *Bradyrhizobium* strains based on multilocus sequence analysis of 16S rRNA, ITS, *glnII*, *recA*, *atpD* and *dnaK* genes. *Int J Syst Evol Microbiol* 59:2934–2950
21. Mpeperek S, Wollum AG (1991) Diversity of indigenous *Bradyrhizobium japonicum* in North Carolina soils. *Biol Fertil Soils* 11:121–127
22. Ndungu SM, Messmer MM, Ziegler D, Thuita M, Vanlauwe B, Frossard E, Thonar C (2018) Evaluation of MALDI-TOF mass spectrometry for the competitiveness analysis of selected indigenous cowpea (*Vigna unguiculata* L. Walp.) *Bradyrhizobium* strains from Kenya. *Appl Microbiol Biotechnol* 102:5265–5278
23. Peres JRR, Mendes IC, Suhel AR, Vargas MAT (1993) Eficiência e competitividade de estirpes de rizóbio para a soja em solos de Cerrados. *Rev Bras Cienc Solo* 17:357–363
24. Peres JRR, Vidor C (1980) Seleção de estirpes de *Rhizobium japonicum* e competitividade por sítios de infecção nodular em cultivares de soja. *Agron Sulriograndense* 16:205–219
25. Ramos HJO, Souza EM, Soares-Ramos JRL, Pedrosa FO (2007) Determination of bean nodule occupancy by *Rhizobium tropici* using the double *gfp* and *gusA* genetic markers constitutively expressed from a new broad-host-range vector. *World J Microbiol Biotechnol* 23:713–717
26. Ruelle V, El Moualij B, Zorzi W, Leden TP, Pauw ED (2004) Rapid identification of environmental bacterial strains by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Rapid Commun Mass Spectrom* 18:2013–2019
27. Sánchez-Juanes F, Ferreira L, Alonso de la Vega P, Valverde A, Barrios ML, Rivas R, Mateos PF, Martínez-Molina E, González-Buitrago JM, Trujillo ME, Velázquez E (2013) MALDI-TOF mass spectrometry as a tool for differentiation of *Bradyrhizobium* species: application to the identification of *Lupinus* nodulating strains. *Syst Appl Microbiol* 36:565–571
28. Siqueira AF, Ormeño-Orrillo E, Souza RC, Rodrigues EP, Almeida LGP, Barcellos FG (2014) Comparative genomics of *Bradyrhizobium japonicum* CPAC 15 and *Bradyrhizobium diazoefficiens* CPAC 7: elite model strains for understanding symbiotic performance with soybean. *BMC Genomics* 15:420
29. Somasegaran P, Hoben HJ (1994) Handbook for rhizobia: methods in legume-*Rhizobium* technology. Springer-Verlag, New York
30. Suarez S, Ferroni A, Lotz A, Jolley KA, Guerin P, Leto J, Dauphin B, Jamet A, Maiden MC, Nassif X, Armengaud J (2013) Ribosomal proteins as biomarkers for bacterial identification by mass spectrometry in the clinical microbiology laboratory. *J Microbiol Methods* 94:390–396
31. Torres D, Revalé S, Obando M, Maroniche G, Paris G, Perticari A, Vazquez M, Wisniewski-Dyé F, Martínez-Abarca F, Cassán F (2015) Genome sequence of *Bradyrhizobium japonicum* E109, one of the most agronomically used nitrogen-fixing rhizobacteria in Argentina. *Genome Announc* 3:1566–1614
32. Van Berkum P, Elia P, Song Q, Eardly BD (2012) Development and application of a multilocus sequence analysis method for the identification of genotypes within genus *Bradyrhizobium* and for establishing nodule occupancy of soybean (*Glycine max* L. Merr). *Mol Plant-Microbe Interact* 25:321–220
33. Vargas MAT, Mendes IC, Suhel AR, Peres JRR (1993) Serological distribution of *Bradyrhizobium japonicum* from Brazilian “Cerrados” areas under soybean cultivation. *Rev Microbiol* 24: 239–243
34. Vargas MAT, Mendes IC, Hungria M (2001) Response of field grown bean [*Phaseolus vulgaris* (L)] to *Rhizobium* inoculation and N fertilization in two Cerrados soils. *Biol Fertil Soils* 32:228–233
35. Vincent JM (1970) A manual for the practical study at root-nodule bacteria. Blackwell Scientific Publ, Oxford
36. Weber DF, Keyser HH, Uratsu SL (1989) Serological distribution of *Bradyrhizobium japonicum* from United States soybean production area. *Agron J* 81:786–789
37. Weiser A, Shcneider L, Jung J, Schubert S (2012) MALDI-TOF MS in microbiological diagnostics – identification of microorganisms and beyond (mini review). *Appl Microbiol Biotechnol* 93: 965–974
38. Welker M, Moore ER (2011) Application of whole-cell matrix-assisted laser desorption/ionization time-of-flight mass spectrometry in systematic microbiology. *Syst Appl Microbiol* 34:2–11
39. Williams TL, Andrzejewski D, Lay JO, Musser SM (2003) Experimental factors affecting the quality and reproducibility of MALDI TOF mass spectra obtained from whole bacteria cells. *J Am Soc Mass Spectrom* 14:342–351
40. Ziegler D, Pothier JF, Ardley J, Fossou RK, Pflüger V, Meyer S, Vogel G, Tonolla M, Howieson J, Reeve W, Perret X (2015) Ribosomal protein biomarkers provide root nodule bacterial identification by MALDI-TOF MS. *Appl Microbiol Biotechnol* 99: 5547–5562

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.