Identification and functional analysis of recent IS transposition events in rhizobia

Ezequiel G. Mogro[1], Walter O. Draghi[1], Antonio Lagares[1] and Mauricio J. Lozano[1*]

Abstract

Rhizobia are alpha- and beta- Proteobacteria that, through the establishment of symbiotic interactions with leguminous plants, are able to fix atmospheric nitrogen as ammonium. The successful establishment of a symbiotic interaction is highly dependent on the availability of nitrogen sources in the soil, and on the specific rhizobia strain. Insertion sequences (ISs) are simple transposable genetic elements that can move to different locations within the host genome and are known to play an important evolutionary role, contributing to genome plasticity by acting as recombination hot-spots, and disrupting coding and regulatory sequences. Disruption of coding sequences may have occurred either in a common ancestor of the species or more recently. By means of IScomapare, we identified Differentially Located ISs (DLISs) in nearly related rhizobial strains of the genera *Bradyrhizobium*, *Mesorhizobium*, *Rhizobium* and *Sinorhizobium*. Our results revealed that recent IS transposition could have a role in adaptation by enabling the activation and inactivation of genes that could dynamically affect the competition and survival of rhizobia in the rhizosphere.

Abbreviated summary

Using ISCompare we identified Differentially Located Insertion Sequences (DLISs) in a diversity of rhizobial strains of agronomical relevance. We found that DLISs tend to be inserted more frequently within intergenic regions. However we found a considerable proportion in the neighborhood of promoter regions and within coding sequences. Although a small number of DLISs were found in each strain, these results suggest that gene disruption by DLISs could play a role in environmental adaptation.
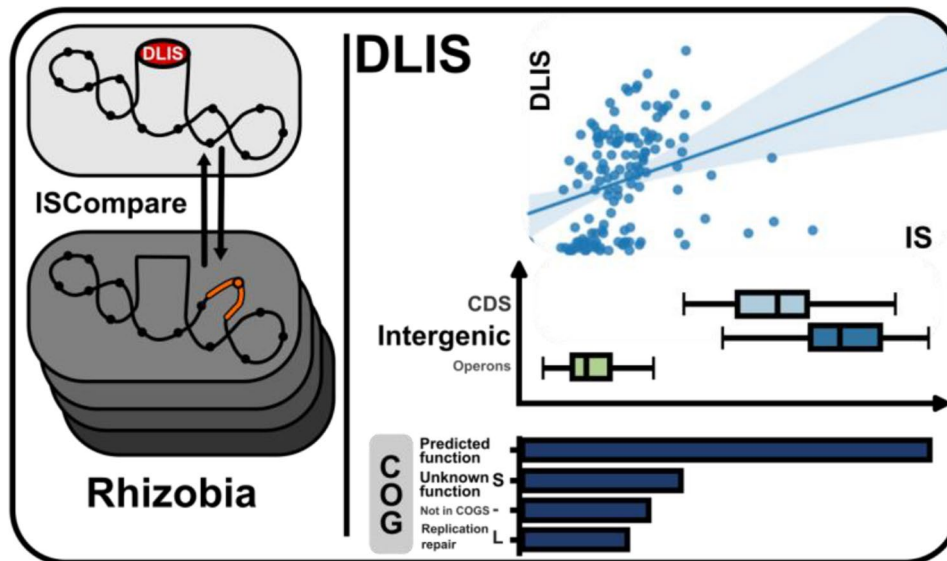
*Correspondence:
Mauricio J. Lozano
maurijlozano@gmail.com; mjlozano@biol.unlp.edu.ar

Full list of author information is available at the end of the article

Graphical Abstract

Introduction

Rhizobia are alpha- and beta- proteobacteria that, through the establishment of symbiotic interactions with leguminous plants, are able to fix atmospheric nitrogen as ammonium. As a result of this interaction, rhizobia induce the development of specialized root structures, the nodules, where they differentiate into bacteroids with the ability to fix atmospheric N_2, providing the plant with a source of nitrogen in exchange for photosynthetic carbon, mostly in the form of dicarboxylic acids [1–5]. Rhizobia are thus of particular interest due to their potential to improve crop yields and reduce the need for the use of synthetic nitrogen fertilizers [6].

The successful establishment of a symbiotic interaction between rhizobia and legumes is highly dependent on the availability of nitrogen sources in the soil, and on the specific strain of infecting rhizobia, being some strains highly competitive for the colonization of plant roots but poor N_2 fixers [7]. The symbiotic space (i.e. the number of nodules) is limited by the host to meet only its nitrogen needs and is therefore accessed only by the most competent rhizobia. The colonization of the rhizosphere –the region of soil under the influence of plant roots– is the earliest step in the symbiotic process, and is critical during the selection of the rhizobial strains that will end up within the nodules [8, 9]. The rhizosphere is a complex and dynamic environment where the interaction between plants, microorganisms, and soil particles takes place, thus playing a crucial role in plant growth and health [10]. Nevertheless, it is not solely inhabited by rhizobia, but also by a diverse community of bacteria, fungi, and other microorganisms that compete for nutrients, but also for the access to the host plant. The colonization of the rhizosphere is therefore a complex process that involves both plant-microbe interactions and competition with other rhizospheric organisms, and is influenced by various factors, including its growth rate, motility, and production of enzymes and other metabolites [9, 11]. Understanding the mechanisms and factors that influence rhizobia ability to colonize the rhizosphere can provide some insight into the dynamics of the rhizobia-legume symbiosis, thus providing valuable information to try improving positive microbe-plant interactions.

Rhizobial genomes are very plastic, and are usually constituted by a chromosome, chromids –replicons of large size, difficult to distinguish from plasmids–, and several plasmids, some of them of large size –megaplasmids– and some smaller, generally considered as part of the accessory genome [3, 12, 13]. Most of the genes required for the establishment of the symbiotic interaction with legume plants are located in megaplasmids, or in certain cases, in chromosomal integrative and conjugative elements (ICEs) [3], and there is evidence suggesting that they actively mobilize within rhizobial populations [14]. Furthermore, rhizobial genomes contain a great number of insertion sequences (IS) [15–17].

Insertion sequences (ISs) are simple transposable genetic elements that can move to different locations within the bacterial genomes [18]. ISs are known to play an important evolutionary role, contributing to genome

plasticity by acting as recombination hot-spots, and disrupting coding and regulatory sequences [18–21]. In rhizobia, genome architecture has been shown to be under constant modification, with replicons frequently being cointegrated and excised [22]. Moreover, the role of tandemly repeated ISs as drivers of genomic recombination events was demonstrated in artificial evolution experiments [23]. ISs can also modify the expression of genes, being well-known the mutation of *Sinorhizobium meliloti* 1021 *expR* –a LuxR family transcriptional regulator that controls the expression of the symbiotically active exopolysaccharide (EPS) EPS II– by the insertion of a ISRm2011-1 insertion sequence [24]. Beside this example, rhizobial genomes possess large amounts of partial and most likely inactive transposases suggesting that some IS insertions could have taken place in a distant ancestor. The role of recent transposition events in rhizobia has not been thoroughly explored thus far.

In this study, we used ISCompare [25] software to identify ISs that have changed their location in nearly related rhizobial strains, and analyzed the genes disrupted by these differentially located ISs (DLISs). Through this study, we have gained a better understanding of the impact of recent IS transposition events in the rhizobial lifestyle.

Materials and methods
Identification of differentially located ISs with ISCompare
ISCompare was used to compare the genomes of several rhizobial strains from the following genera and species: *Bradyrhizobium diazoefficiens*, *Bradyrhizobium japonicum*, *Mesorhizobium ciceri*, *Mesorhizobium loti*, *Rhizobium etli*, *Rhizobium leguminosarum*, and *Sinorhizobium meliloti*. Genome sequences were obtained from the National Center for Biotechnology Information (NCBI) assembly database [26]. The accession numbers for the sequences used in this study are compiled in Table S1.A. The sequences were downloaded in FASTA, CDS FASTA, and GBFF formats. The sequences used as IS database for ISCompare were downloaded from https://github.com/thanhleviet/Isfinder-sequences.

ISCompare [25] is a tool used to find differential located ISs between two closely related strains. For each of the analyzed species we made pairwise comparisons using a selected strain as reference and all the complete closed genomes available as of 5/Jan/2023 as target. Only the results classified by ISCompare as DLIS and with a complete IS match, were analyzed in this study. The software was run with the following parameters: E-value=1e-10, minLength=50, ISdiff=50, scaffoldDiff=20, minAlnLength=50, surroundingLen=500, surroundingLen2=500, shift=0. The total number of DLISs located in accessory replicons –plasmids, megaplasmids– or in the main replicon –chromosome– was determined. DLIS counts were normalized by the species-average size of the main and accessory replicons –number of DLISs per 100,000 base pairs– to account for the difference in their genome sizes. For DLISs located within genomic symbiotic islands, the counts were normalized by their respective size. The statistical analysis was done with python scripts using scipy [27]. For the comparison of proportions Chi2 test was used summing the counts over all the strains within each species.

Distribution of ISs in rhizobia
A custom python script was used to count the genes and pseudogenes annotated as transposase or being part of an insertion sequence. In order to obtain this information, genbank files were processed with Biopython [28] and if the words 'transposase' or 'insertion seq' were found in a feature qualifier 'product', a transposase was counted. Pseudogenes were identified based on the presence of a 'pseudo' feature qualifier. In addition, the replicon information –transposon counts per replicon– was saved for further analysis.

DLISs functional analysis
To determine whether an IS was located within an intergenic region that could correspond to a promoter, the following analysis was done. First, the two closest CDS to a DLIS insertion site were found using the "closest" command from bedtools software [29]. When the insertion site was located within 150 nucleotides upstream of a CDS, the DLIS was considered to be interrupting a probable promoter region. This cutoff value was selected based on the average location of transcriptional start sites (TSS) in *S. meliloti* at -68 nucleotides [30] from the start codon −5′-UTRs longer than 100 nt were found in 1,041 genes−, and the fact that promoter sequences are usually in the −10 and −35 positions from the TSS. To identify DLISs interrupting a probable operon, the two closest CDS to the insertion site were determined with bedtools "closest" command, and the orientation and a distance between these CDS was analyzed. A possible operon was considered if the two CDS were in the same orientation, and separated by a maximum of 100 nucleotides [31, 32].

In the case of DLISs inserted within coding sequences, a functional classification was made using COG (Cluster of orthologous groups database) [33]. COGs were assigned to each protein using the eggNOG-mapper web server with the eggNOG 5 database [34].

Figures were plotted using python matplotlib [35] and seaborn [36]. Whenever it was required, final figures were edited with Inkscape v 1.3 (The Inkscape team, https://inkscape.org/).

RESULTS
Analysis of ISs and DLISs distribution in rhizobia
The genomes of all *Bradyrhizobium diazoefficiens*, *Bradyrhizobium japonicum*, *Mesorhizobium ciceri*,
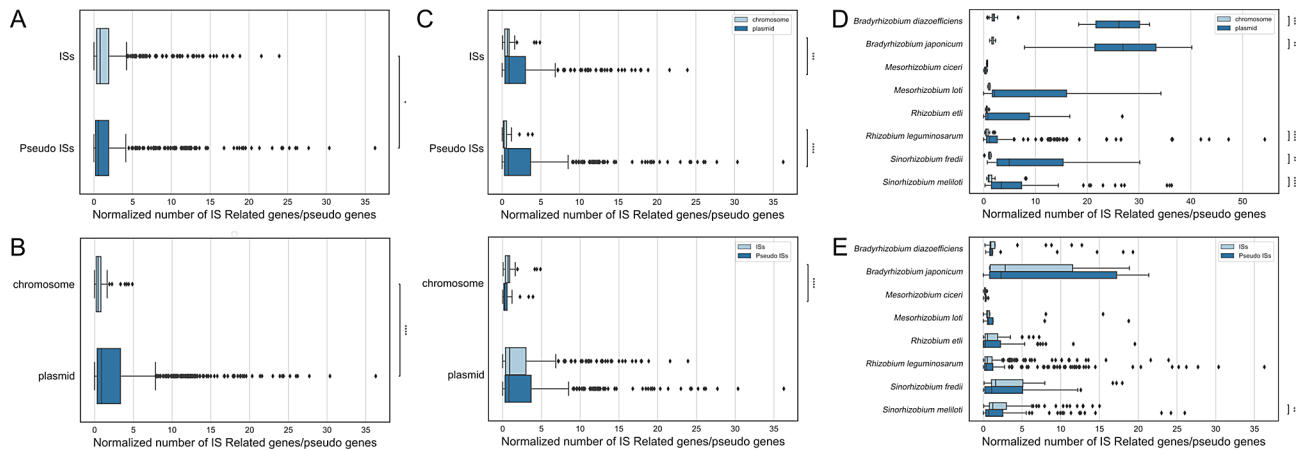
Fig. 1 Distribution of transposase and IS related genes. **A**. Normalized IS and pseudo IS counts. **B**. Normalized number of IS elements by replicon. **C**. Normalized numbers of ISs and pseudo ISs by replicon. **D**. Normalized number of IS elements by species and replicon. **E**. Normalized number of ISs and pseudo ISs by species. The number of transposases and IS related genes was estimated from the genbank files annotation using custom python scripts that looked for the terms 'transposase' and 'insertion seq' in the product descriptions of coding sequences and pseudo genes. Normalized counts were calculated as ISs / 100,000 bp. Significance was determined using the Mann-Whitney test and the normalized counts (*: 1.00e-02 < p <= 5.00e-02; **: 1.00e-03 < p <= 1.00e-02; ***: 1.00e-04 < p <= 1.00e-03; ****: p <= 1.00e-04)

Mesorhizobium loti, *Rhizobium etli*, *Rhizobium legumi-nosarum*, and *Sinorhizobium meliloti* strains containing a complete assembly level were downloaded from NCBI Assembly database [26] (accessed 5/Jan/2023) and pair-wise comparisons to identify DLISs were made using ISCompare (Table S1.B). ISCompare identifies DLISs by searching all the ISs in a query genome, extract-ing their flanking genomic DNA sequences, and look-ing for them in a target genome. A DLIS is reported when those sequences are found uninterrupted by an IS in the target genome. Next the reference and tar-get genomes are exchanged and the process is repeated [25]. For each species, a specific strain was used as ref-erence for all the pairwise comparisons. In order to esti-mate the proportion of active ISs –the relation between the number of DLISs and the total number of ISs in the pair of genomes– the total number of ISs in a determined genome was estimated as the number of transposases plus the number of other genes annotated with functions related to insertion sequences (Fig. 1, Table S1.C). The estimated numbers of total ISs were in accordance with those previously reported [12, 15–17]. A recurrent obser-vation is that chromosomal replicons tend to have lower number of total ISs whereas plasmids have a considerably higher amount (Fig. 1.B, Table S1.E). This does not seem to be the case for DLISs (Fig. 2.B., Table S2) which were detected with comparable frequency in plasmids (less but not significantly so) and chromosomal replicons. Of the rhizobia studied here, *Sinorhizobium* and *Bradyrhi-zobium* were the ones that showed the greatest number of DLISs (Fig. 2.A) which agrees with the fact that they also present the highest number of ISs in their genomes (Fig. 1.D-E, Table S1.C). Further, they also presented the

highest proportion of ISs that have recently changed their location in the genome –active ISs–, calculated as the number of DLISs / (total ISs on reference strain+total ISs on target strain) (Fig. 3.A). Finally, a weak but sig-nificant correlation between the number of DLISs and the total number of ISs was observed (Fig. 3.B, $R^2=0.11$; p-value$=4.552\times10^{-5}$).

Next, the distribution of DLISs insertion sites within genes, intergenic regions and possible operons was deter-mined. Most DLISs presented insertion sites within intergenic regions (ca. 60%, Fig. 4.A, Table S3.A), and the lowest proportion was observed for DLISs inserted within putative operons. Both *Mesorhizobium* spe-cies analyzed and *R. etli* presented a higher proportion of DLISs inserted within coding sequences (Fig. 4.B-C), although they also presented the lowest numbers of DLISs. Remarkably, half of the intergenic DLISs have the potential to interrupt promoter regions (Fig. 4.C).

For genera presenting genomic symbiotic islands (GSI) —*Bradyrhizobium* and *Mesorhizobium* [16, 37, 38]— we searched for DLISs that were inserted within the GSI in the reference genome, while absent in the corresponding target genome. We found 6 DLISs in *B. japonicum* USDA6, of which two were intragenic and four were intergenic, one of the latter inserted near a promoter. Also the number of normalized DLIS was higher in the GSI (0.452±0.245) than in the chromosome (0.076±0.036) (Table S2.C), although no significant dif-ferences were found (Wilcoxon test, *p*=0.0625).

Functional distribution of genes interrupted by DLISs

To analyze the cellular functions affected by DLISs inser-tions, we recovered the amino acid sequences of the
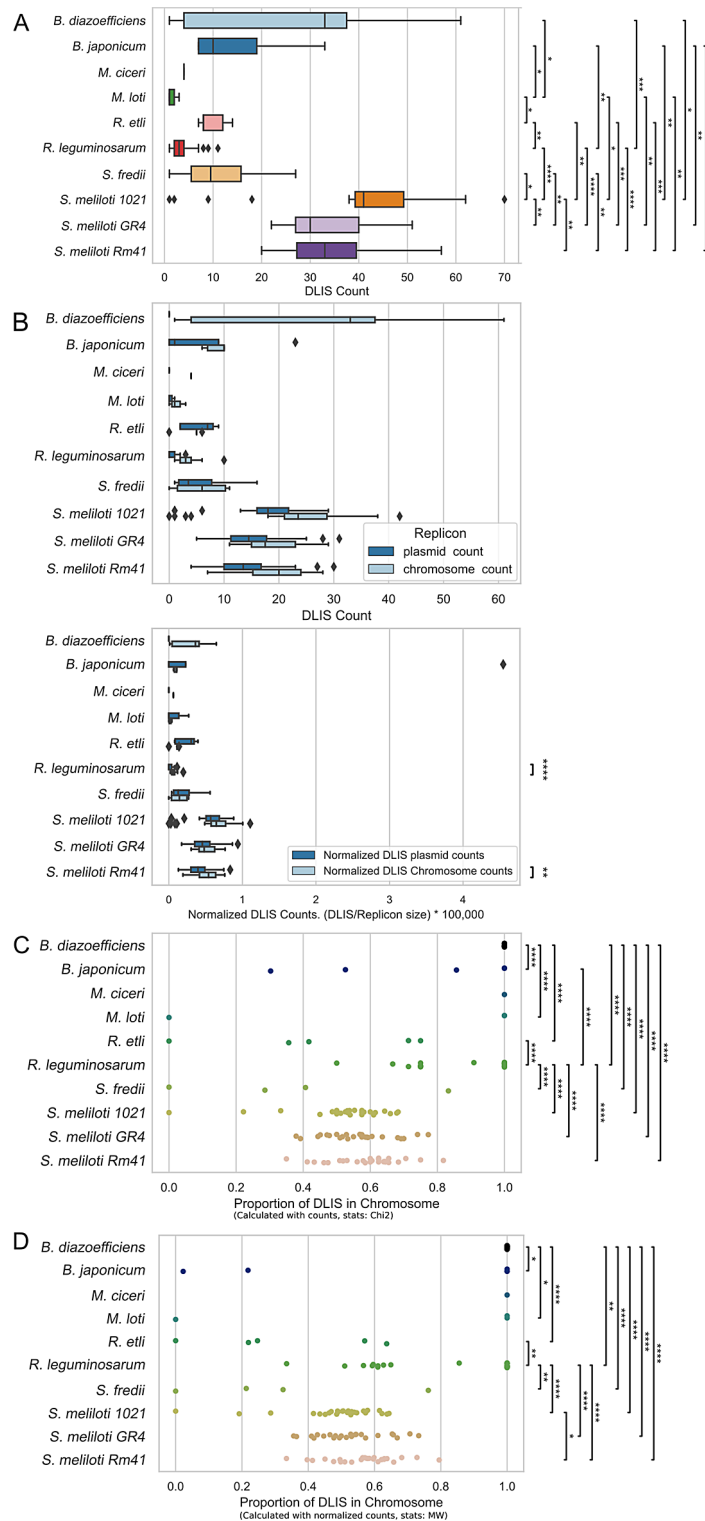
Fig. 2 Distribution of DLISs. Chromosomes vs. plasmids. **A**. Total DLIS counts. **B**. DLIS counts and normalized DLIS counts discriminated by rhizobia and replicon type. **C**. Proportions of DLISs in chromosomal replicons, calculated from DLIS counts. **D**. Proportions of DLISs in chromosomal replicons, calculated from DLIS normalized counts. DLISs were identified with ISCompare, and those with a full match for an insertion sequence and confidently identified as DLIS were analyzed. Significance was determined either using Chi2 test and the total DLIS counts in each species, or the Mann-Whitney test and the normalized DLIS counts (DLIS counts / 100,000 bp). *: 1.00e-02 < p < = 5.00e-02; **: 1.00e-03 < p < = 1.00e-02; ***: 1.00e-04 < p < = 1.00e-03; ****: p < = 1.00e-04
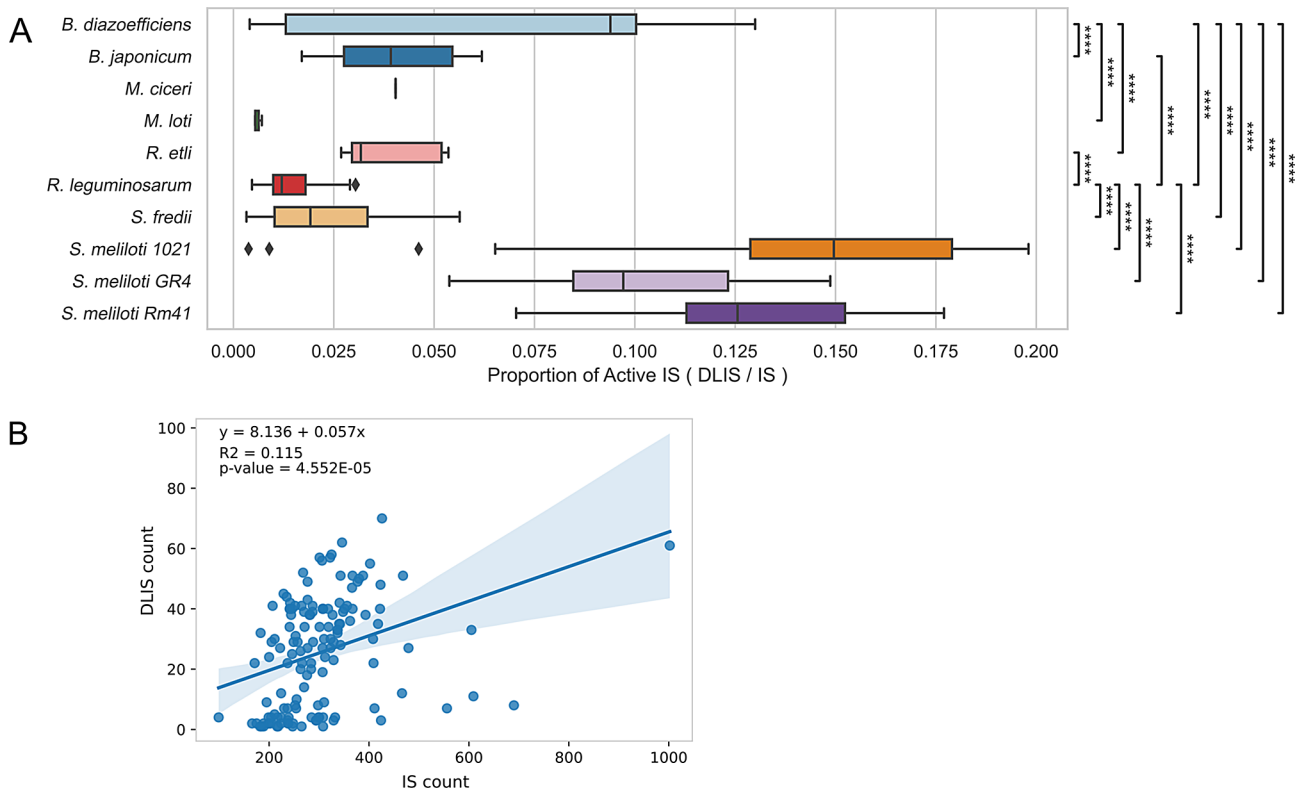
Fig. 3 Proportion of recently active ISs. (**A**) Proportion of active ISs. The proportion of recently active ISs was estimated as the relation of DLISs over the total ISs count. Significance was determined using the Mann-Whitney test and the proportions of active ISs (DLISs / ISs). *: $1.00e-02 < p <= 5.00e-02$; **: $1.00e-03 < p <= 1.00e-02$; ***: $1.00e-04 < p <= 1.00e-03$; ****: $p <= 1.00e-04$. (**B**) Pearson correlation of IS counts and DLIS counts
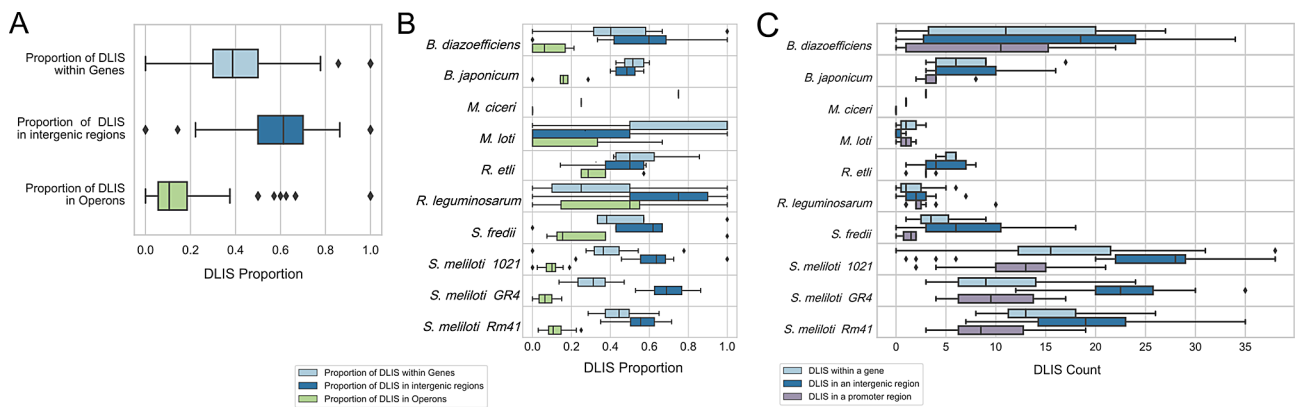

Fig. 4 Distribution of DLISs in coding and intergenic regions. **A**. Proportion of DLISs inserted within genes, intergenic regions and putative operons. **B**. Proportion of DLISs inserted within genes, intergenic regions and putative operons discriminated by rhizobial species. **C**. Absolute count numbers of DLISs within genes, intergenic regions, and putative promoter regions. The count of the promoter regions is a subset of the counts for intergenic regions. When a DLIS was inserted within a short intergenic region between two divergent genes it was counted twice since it could be affecting the transcription of both genes. The location of the DLIS was determined using custom python scripts and the bedtools software as described in material and methods. Significance was determined using the Mann-Whitney test and the normalized counts

uninterrupted version of the proteins from the genomes that didn't have the DLIS. All these proteins were annotated with COGs using the eggNOG-mapper web server [34]. We analyzed all the genes with an inserted DLIS in the target genomes (the uninterrupted copy located in the reference genomes, Table S3.B). A total of 510 of

the identified DLISs were inserted within genes, of which 280 (55%) presented a genomic product annotation that was not related to transposases or hypothetical proteins. The overall results showed that the top COG categories –those with the greatest number of DLISs– were S (Function unknown), K (Transcription), L (Replication,
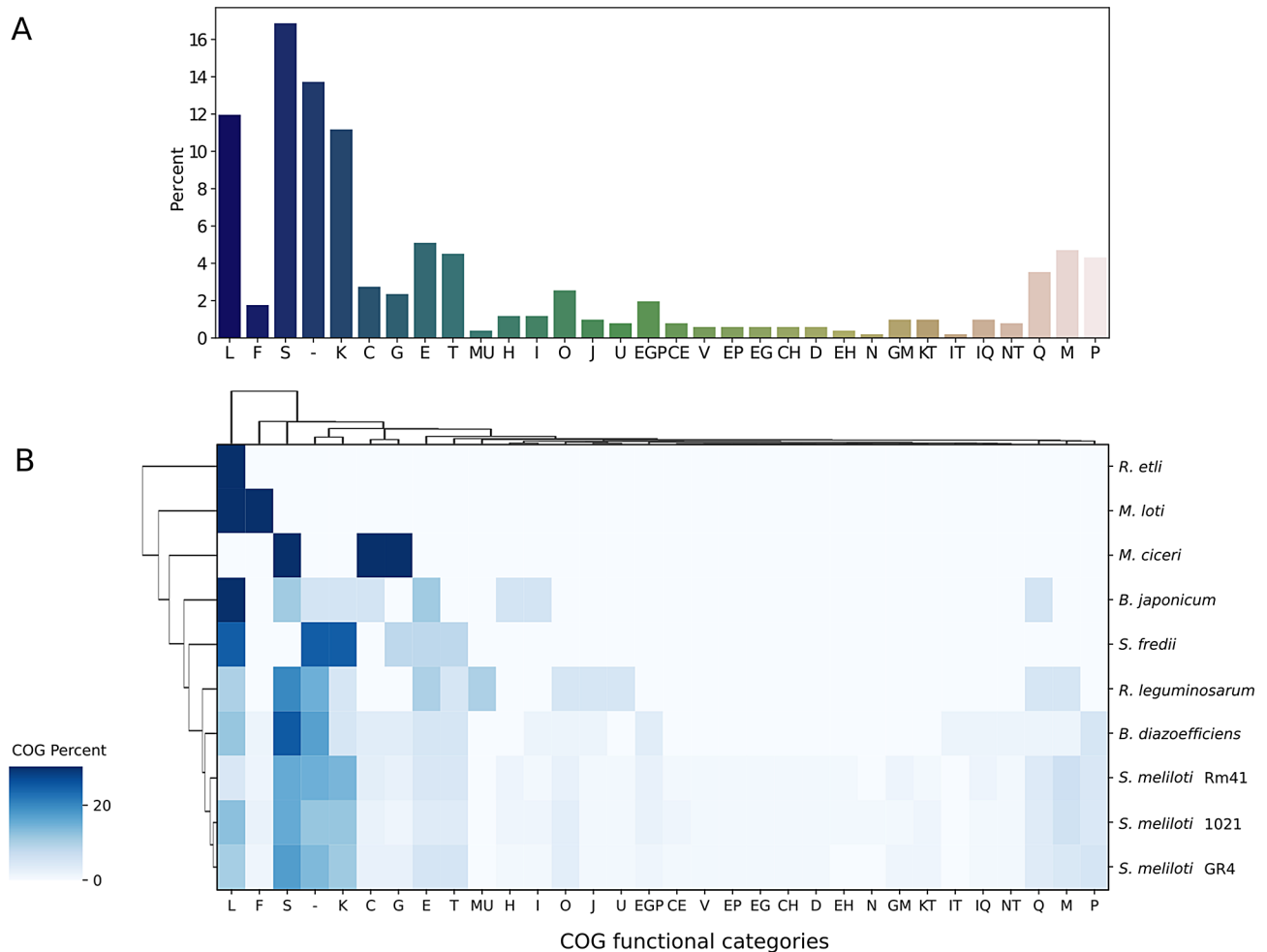
Fig. 5 Distribution of COG functional categories. COGs functional categories for the genes interrupted by a DLIS were assigned with eggNOG-mapper. **A**. Distribution of DLIS interrupted genes by COG functional category. **B**. Heatmap of COG distribution per reference strain. COG categorías are as follows. -: Not assigned; **A**: RNA processing and modification; **B**: Chromatin Structure and dynamics; **C**: Energy production and conversion; **D**: Cell cycle control and mitosis; **E**: Amino Acid metabolism and transport, **F**: Nucleotide metabolism and transport, **G**: Carbohydrate metabolism and transport; **H**: Coenzyme metabolism; **I**: Lipid metabolism; **J**: Translation; **K**: Transcription; **L**: Replication and repair; **M**: Cell wall/membrane/envelope biogenesis; **N**: Cell motility; **O**: Post-translational modification, protein turnover, chaperone functions; **P**: Inorganic ion transport and metabolism; **Q**: Secondary metabolites biosynthesis, transport and catabolism; **R**: General Functional Prediction only; **S**: Function Unknown; **T**: Signal Transduction; **U**: Intracellular trafficking and secretion; **V**: Defense mechanisms: **W**: Extracellular structures: X: Mobilome: prophages, transposons; **Y**: Nuclear structure; **Z**: Cytoskeleton

recombination and repair), E (Amino Acid metabolism and transport), M (Cell wall/membrane/envelope biogenesis), T(Signal Transduction), P (Inorganic ion transport and metabolism) and Q (Secondary metabolites biosynthesis, transport and catabolism) (Fig. 5. Table 1, Table S3.C). L, S and K were also the principal categories in most rhizobia.

Proteins with a classification in categories S and (-) corresponded majorly to hypothetical and conserved hypothetical proteins, or to proteins with unknown function. However, 70% of the proteins in the S category were annotated in the genome with a more complete function description (putative cellulose biosynthesis, chitinase, dihydroxyacetone kinase and nitrile hydratase activities, among others; Table S3.B). The L category

contained DLISs that in most cases were inserted within genes encoding for transposases, integrases and reverse transcriptases, with only a few DLISs inserted in genes annotated as involved in DNA repair, nucleases, DNA polymerases, recombinases and methylases. For the K category, the proteins interrupted by DLISs corresponded to different families of transcriptional regulators, acetyltransferase domain containing proteins, and proteins involved in plasmid partition. Particularly, NifA and NodD1 presented DLISs insertions in *S. meliloti* AK76 and *B. japonicum* USDA 6, respectively. The *nifA* gene of *S. meliloti* –a regulatory nitrogen fixation gene required for the induction of several key *nif* and *fix* genes [39]– presented an inserted DLIS in strain AK76. NifA null mutants induce white nodules in the roots of the

Mogro *et al. Mobile DNA*   (2024) 15:17

Page 8 of 14

Table 1 List of genes interrupted by a DLIS and annotated with a gene name in the reference strain

Gene Name	Product	Organism
acaP	Carbonic anhydrase	S. meliloti Ak57
acsA1	Acetyl-coenzyme A synthetase 1 – hypothetical protein	S. meliloti RCAM1115
afuA	ABC-type Fe3 + transport system, periplasmic component	S. meliloti HM006
aglR	HTH-type transcriptional regulator AglR	S. meliloti Ak57
aspB	Aspartate/tyrosine/aromatic aminotransferase	S. meliloti 1021
corA	magnesium and cobalt transport protein CorA	S. meliloti GR4
cydB	cytochrome d ubiquinol oxidase subunit II	B. diazoefficiens NK6
degP4	Trypsin-like serine protease, periplasmic, C-terminal PDZ domain	S. meliloti USDA1106
deoR1	transcriptional regulator protein, Putative sugar-binding domain	S. meliloti USDA1157
dinB	DNA polymerase IV 2, error-prone DNA polymerase	S. meliloti HM006
dnaB	putative replicative DNA helicase	S. meliloti GR4
eutB	Ethanolamine ammonia lyase large subunit	S. meliloti RMO17
fhs	formate–tetrahydrofolate ligase	M. loti 582
flaR	Adenylate kinase; ATPase AAA, CRISPR/Cas system-associated protein Cas3	S. meliloti RU11_001, Rm41
gabD	Succinate-semialdehyde dehydrogenase [NADP(+)] GabD	S. meliloti AK83
glcF	putative glycolate oxidase iron-sulfur subunit	S. meliloti GR4
groS	co-chaperone GroES	B. diazoefficiens F07S3
gshA	glutathione synthetase	B. japonicum CC829
idnO	Gluconate 5-dehydrogenase	S. meliloti M162
kdgK	2-dehydro-3-deoxygluconokinase, pfkB family carbohydrate kinase	M. ciceri R30
kdsD	arabinose-5-phosphate isomerase, SIS family, GutQ KpsF subfamily	S. meliloti KH35c
lpsB2	polysaccharide biosynthesis protein CapD	S. meliloti 1021
lpxK	tetraacyldisaccharide 4'-kinase	S. meliloti 1021, GR4
luxR	transcriptional regulator, autoinducer-binding domain-containing protein	S. meliloti RU11_001
macA	ABC-type family efflux transporter MFP subunit	S. meliloti Ak57
metB	Cystathionine gamma-synthase	S. meliloti KH46
modC	Molybdenum import ATP-binding protein ModC	S. meliloti AK83
msmE	Multiple sugar ABC transporter substrate-binding protein	S. fredii CCBAU_45436
nhaA	Na+/H+ antiporter	S. meliloti KH35c
nifA	Nif-specific regulatory protein	S. meliloti AK76
nirB	Nitrite reductase [NAD(P)H]	S. meliloti USDA1106
nodD1	LysR family transcriptional regulator	B. japonicum USDA_6
norG	Transcriptional regulator with HTH domain and aminotransferase MocR family domain	S. meliloti S35m
nthA	nitrile hydratase subunit alpha	S. meliloti Rm41
parB	ParB/RepB/Spo0J family partition protein	R. leguminosarum 31B
pepT	peptidase T	R. leguminosarum Vaf-108
pstS	phosphate ABC transporter substrate-binding protein PstS	B. diazoefficiens NK6
psuG	putative enzyme involved in pigment biosynthesis, pseudouridine degradation pathway	S. meliloti RCAM1115
pyk	pyruvate kinase	B. diazoefficiens NK6

Table 1 (continued)

Gene Name	Product	Organism
rbn	ribonuclease BN (exoribonuclease), putative membrane protein	S. meliloti M162
rhcU	EscU/YscU/HrcU familyT3SS apparatus switch protein	B. diazoefficiens USDA_110
rhiA	RhiA, rhizosphere-expressed protein	R. leguminosarum Vaf10
rkpI	Capsular polysaccharide biosynthesis protein rkpI, Sulfatase	S. meliloti USDA1157
sbcD	exonuclease SbcCD subunit D	B. diazoefficiens USDA_110
sqr	NAD(P)/FAD-dependent oxidoreductase	B. diazoefficiens H12S4
surF	surfeit locus 1 family protein, SURF1-like	S. meliloti USDA1021
tauZ	UPF0324 membrane protein, TIGR00698 family, Putative sulfate exporter	S. meliloti AK21, GR4
terC	transmembrane protein	S. meliloti KH46
tyrS	tyrosine–tRNA ligase	R. leguminosarum RCAM2802
usg	Usg protein, hypothetical protein	S. meliloti KH35c, AK21
virB6	VirB6 type IV secretion protein	S. meliloti USDA1157
ybaL	Inner membrane protein ybaL, monovalent cation proton antiporter 2	S. meliloti AK83
yceA	Putative rhodanese-related sulfurtransferase	S. meliloti RU11_001
ykuE	Calcineurin-like phosphoesterase, hypothetical protein	S. meliloti KH46
yqaA	hypothetical membrane protein	S. meliloti 1021, Rm41
ytfG	NmrA-like family, hypothetical protein	S. meliloti SM11
zapE	cell division protein ZapE, AFG1-like ATPase	B. diazoefficiens XF7

host plant, have reduced swarming ability, and present lower levels of acyl-homoserine lactones and of extracellular proteins [40]. In *B. japonicum* USDA 6, NodD1, the positive regulatory protein of the *nodYABC* operon [41], was found to be interrupted by a DLIS. In that case, a shorter ORF is annotated in the genome, and it could be possible that its product is still functional.

Several proteins in the M (Cell wall/membrane/envelope biogenesis) COG category were interrupted by DLISs in *S. meliloti strains*, including three enzymes involved in the biosynthesis of lipopolysaccharide (LPS) (Tetraacyldisaccharide 4'-kinase, LpxK; Arabinose-5-phosphate isomerase, KdsD; and the polysaccharide biosynthesis protein LpsB2), a capsular polysaccharide biosynthesis export transmembrane protein (RkpI), and a choline-glycine betaine transporter. LpxK is involved in the biosynthesis of lipopolysaccharide (LPS), catalysing the sixth step in the lipid A synthesis [42], while LpsB2 is required for O-antigen biosynthesis and its mutants presented reduced motility, grew faster than the parental strain, and were more sensitive to maize benzoxazinones and polymyxin B [43, 44]. KdsD is required for the biosynthesis of 3-deoxy-d-manno-octulosonic acid (Kdo), a key sugar in the core region of LPS [45]. RkpI is a transmembrane protein involved in capsular polysaccharide (KPS) biosynthesis [46]. Both LPS and KPS have been reported to have a role in symbiosis, KPS being relevant in the early steps of the infection process, while LPS has been shown to be important in the later stages of the nodulation processes [47, 48]. In addition, the gene encoding for MacA, a membrane fusion protein of an ABC-Type efflux transporter (a type 1 secretion system that transports diverse molecules including antibiotics and peptides across the inner and outer membranes) [49, 50], was interrupted in *S. meliloti* Ak57.

In the E category (Amino Acid metabolism and transport), insertions were within genes encoding different enzymes annotated as arginine/lysine/ornithine decarboxylase, aspartate/tyrosine/aromatic aminotransferase (AspB), choline dehydrogenase, ethanolamine ammonia lyase (EutB), and phosphoribosyl anthranilate isomerase (Usg). AspB catalyses the transfer of an alpha amino group from aromatic amino acids, but also from aspartate, to different substrates. It was reported that it contributed, under high levels of exogenous tryptophan, to the biosynthesis of indole-acetic acid (IAA), and to nitrogen scavenging under nitrogen deprivation [51]. Also related to the IAA metabolism, *nthA* a gene encoding for nitrile hydratase subunit alpha was found interrupted by a DLIS in *S. meliloti* RM41 [52].

Other insertions that could be relevant for rhizobial fitness were observed in proteins belonging to the categories T (Signal Transduction) and P (Inorganic ion transport and metabolism). In particular DLISs insertions were observed in adenylate/guanylate cyclases, in two component systems (both in histidine kinases and response regulators), and in several Na^+/H^+, Mg^{2+}/Co^{2+}, Fe^{3+} and K^+ transport proteins. Among these, the genes encoding for a magnesium and cobalt transport protein (CorA), a putative molybdenum transport ATP-binding ABC transporter (ModC), and a Na^+/H^+ antiporter (NhaA) presented DLISs insertions in *S. meliloti*. CorA was shown to be important for growth on glucose at 21% O_2 [53], while ModC is required for nitrogen fixation on limiting levels of molybdate and was reported to be the high-affinity molybdate transporter in *B. diazoefficiens* [54]. In the case of NhaA, it was shown to be induced under acid conditions, and it is likely that it functions by expelling protons toward the periplasm [55]. Also in the P category, *pstS*, a part of the *pstSCAB* phosphate-specific transport operon that functions as high-affinity phosphate transporter [56], was found to be interrupted in *B. diazoefficiens* NK6.

In the G category (Carbohydrate metabolism and transport) DLISs were found mainly in ABC transporters including a ribose/xylose/arabinose/galactoside transporter, and a branched-chain amino acid transport system. Finally, in the C category (energy production and conversion), genes encoding for proteins annotated as succinate-semialdehyde dehydrogenase (GabD), nitrite reductase (NirB), cytochrome *bd* terminal oxidase, and pyruvate ferredoxin/flavodoxin oxidoreductase activities presented DLISs insertions. NirB was shown to be involved in the nitrate assimilatory pathway and to participate indirectly in NO synthesis, possibly contributing with the denitrification pathway [57], and thus, to the generation of greenhouse gases. GabD homologs are required for growth on γ-aminobutyrate (GABA) as the sole nitrogen source [58].

Other genes that were found interrupted by DLISs and could affect rhizobial fitness were classified in minority COG categories. Among them, DLISs were found interrupting *acsA1* (acyl-coenzyme A synthetase, an enzyme required for growth using acetate as carbon source) [59] in *S. meliloti* strain RCAM1115; *rhiA* (the first gene of the rhizosphere-expressed genes operon, that was reported to influence nodulation) [60] in *R. leguminosarum* Vaf10; GshA (glutathione synthetase, involved in the biosynthesis of glutathione, which has been shown to be important for growth under defined conditions, and to play an important role in symbiosis) [61–63] in *B. japonicum* CC829; and finally, Carbonic anhydrase (CA) (an enzyme involved in the interconversion of carbon dioxide and bicarbonate, which was was hypothesized to help in the protonation of extracellular NH3 facilitating its diffusion and transport to the plant tissues) [64] in *S. meliloti* AK57.

Discussion

Rhizobial plasmids are known to have a larger number of ISs than chromosomes and it has been shown that they can generate gene disruptions or act as mediators of homologous recombination leading to genomic rearrangements [22]. Surprisingly, we found a greater number of DLISs in chromosomal replicons (Fig. 2.C-D). DLISs are ISs that are inserted into highly conserved regions of a bacterial genome while absent in the homologous region of another genome under comparison, and could therefore be broadly considered as recently active ISs. That DLISs tend to be more abundant in chromosomal replicons than in plasmids might indicate that a higher activity of ISs is present in the chromosomes. However, this result might also be a consequence of the inherent difficulty to confidently identify DLISs in plasmids, since they are less conserved, have a significantly higher number of pseudo ISs than chromosomal replicons (Fig. 1.C), and usually show a higher frequency of recombination events. In this work we only considered the DLISs identified by ISCompare where complete blast hits to ISs in the library were found, disregarding partial ISs which could be related to ancestral insertions, new ISs –absent in the IS database used– or recombination events, and thus the actual number of DLISs in plasmids may have been underestimated. We also observed a weak correlation of the number of identified DLISs with the Average Nucleotide Identity between the strains under comparison (Figure S1), with strains with ANI under 95% presenting in general less than 10 DLISs, and even less located in plasmids. This was expected since the plasmids from those strains are less conserved, making the identification of DLISs using ISCompare even more difficult.

Of the analyzed rhizobia genera, *Sinorhizobium* and *Bradyrhizobium* presented the highest numbers of ISs and DLISs, and also the highest proportions of active ISs (number of DLISs in relation to the total number of ISs in the reference and target strains). Furthermore, we found a weak but significant correlation (r=0.115, p-value=4.552×10^{-5}, Fig. 3.B) between the number of ISs and the number of DLISs. Thus, a rough estimate of the importance of DLISs in genome dynamics could be obtained from the number of ISs which are easier to quantify using bioinformatic methods. A key finding was that most rhizobial DLISs were found to be inserted within intergenic regions (Fig. 4), and almost half of them were far from the 5'-ends of coding sequences. Nevertheless, it has to be also considered that intergenic regions could encode for small ORFs or RNAs which have not yet been annotated. Furthermore, many DLISs were located in the neighborhood of promoter regions (Fig. 4.C, Figure S2). These insertions could have a role in environmental adaptation by affecting transcription.

However a more precise analysis taking into account the predicted promoters, transcriptional start sites, and the presence of transcriptional terminators within the DLISs will be required to accurately assess the real number of DLISs insertions that could have an effect on transcription. In addition, our results also suggest that only a low proportion of the identified DLISs are expected to generate polar mutations by interrupting putative operons (Fig. 4.A-B), and thus knock-out several ORFs with a single transposition event.

Regarding the type of interrupted genes, approximately half of them corresponded to insertion-sequence-related genes and hypothetical proteins, mostly in COG categories L (replication and repair), S (unknown function) and - (not assigned). Using the Fitness Browser database (https://fit.genomics.lbl.gov/, accessed 8/8/2024) we analyzed the phenotype of transposon insertion mutants in all *S. meliloti* 1021 genes that were interrupted by a DLISs insertion in other *S. meliloti* strains (Table S3.B, *S. meliloti* 1021), and found that in in more than 80 growth conditions only 6 genes had a clearly affected phenotype (Table S3.D). These results together with the small number of DLISs per genome (an average of 25 DLISs per comparison, Table S3.A, Fig. 2.A) may indicate that, in rhizobia, mutations produced by DLISs may be of less importance as a mechanism for environmental adaptation than their role as recombination hotspots, that can ultimately lead to gene gain or loss. Nevertheless, certain strains presented DLISs inserted within genes reported to have a role in the symbiotic interaction with the host plant, or to be required for growth in certain conditions. In particular, we identified DLISs inserted in genes corresponding to enzymes involved in the biosynthesis of lipopolysaccharide (LPS), capsular polysaccharide (KPS), on the nitrate assimilatory pathway, and on several regulators, among others. The impact of these insertions in the symbiotic phenotype remains still unknown. Moreover, a homologous of *Smc03123* (putative transcriptional regulator), *a* gene from *S. meliloti* 2011 involved in the competence for rhizosphere colonization at 3 days post-inoculation [9] was interrupted by a DLIS in *S. meliloti* USDA1157.

Competition in the rhizosphere requires rhizobia to be able to use diverse compounds as carbon and nitrogen sources. Possessing a wide repertoire of metabolic pathways and transporters, enable rhizobia to thrive and compete for the nutrients exuded to the rhizosphere by the plant host. Dynamic inactivation of genes by DLISs could play an important role, granting a faster growth capability. However, it could also be detrimental, affecting posterior steps that occur during the interaction with the plant host, or under changing environmental conditions.

Conclusion

In this work we analyzed recent transposition events in rhizobia by identifying the DLISs by means of ISCompare. Our results revealed that recent IS transposition events could play a role in adaptation either by affecting transcription (e.g. by DLISs inserted in the proximity of promoter regions) or through the disruption of coding sequences. This hypothesis is supported by the majority of DLISs being inserted near to possible promoters and transcriptional regulator regions, and by the fact that half of the intragenic DLISs were inserted within genes annotated with at least a rough functional prediction. It should be also noted that hypothetical proteins could have yet undiscovered functions that might prove to be relevant for rhizobial lifestyle either in the soil or in their symbiotic relation with legume plants. However, we found that in *S. meliloti* most of the genes with a DLIS appear to be non-essential. This makes evolutionary sense, since as a consequence of purifying selection DLISs will tend to be fixated more frequently in silent genomic regions. A secondary conclusion is that ISs might have a more important role in adaptation by acting as mediators of homologous recombination, allowing the dynamic interchange of information between chromosomal and plasmid replicons, and between different rhizobial species. To further characterize the role of DLISs in rhizobia, experimental evolution of single rhizobial isolates should be carried on. Through this type of experiment a clearer picture of the dynamics of IS transposition and recombination would probably be achieved.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13100-024-00327-8.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

Supplementary Material 4

Supplementary Material 5

Author contributions
M.J.L: Conceptualization, Methodology, Investigation, Bioinformatic analysis, Writing (original draft and review & editing). E.G. M: Methodology, Investigation, Bioinformatic analysis, Writing (original draft, review & editing). W.O.D: Writing (review & editing), Funding acquisition. A.L: Writing (original draft, review & editing), Funding acquisition. All authors reviewed the manuscript.

Data availability
No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
All authors agree with the current submission of the manuscript as it stands.

Competing interests
The authors declare no competing interests.

Author details
[1]Instituto de Biotecnología y Biología Molecular (IBBM), Dep. Ciencias Biológicas - Facultad de Ciencias Exactas, Universidad Nacional de La Plata (UNLP), CONICET CCT-LaPlata. La Plata, Buenos Aires, Argentina

References
1. Jones KM, Kobayashi H, Davies BW, Taga ME, Walker GC. How rhizobial symbionts invade plants: the *Sinorhizobium–Medicago* model. Nat Rev Microbiol [Internet]. 2007;5(8):619–33. 2007/07/17.
2. Oldroyd GED, Murray JD, Poole PS, Downie JA. The rules of engagement in the legume-rhizobial symbiosis. Annu Rev Genet [Internet]. 2011 Jan;45(August):119–44. [cited 2014 Mar 19];.
3. Poole P, Ramachandran V, Terpolilli J, Rhizobia. From saprophytes to endosymbionts. Nat Rev Microbiol [Internet]. 2018;16(5):291–303.
4. Shumilina J, Soboleva A, Abakumov E, Shtark OY, Zhukov VA, Frolov A. Signaling in Legume–Rhizobia Symbiosis. 2023 Dec [cited 2024 Jan 29];24(24):17397.
5. Gourion B, Berrabah F, Ratet P, Stacey G. Rhizobium–legume symbioses: the crucial role of plant immunity. Trends Plant Sci [Internet]. 2015 Mar;20(3):186–94. [cited 2014 Dec 27];.
6. Herridge DF, Peoples MB, Boddey RM. Global inputs of biological nitrogen fixation in agricultural systems. Plant Soil. 2008;311(1–2):1–18.
7. Checcucci A, DiCenzo GC, Bazzicalupo M, Mengoni A. Trade, Diplomacy, and Warfare: the Quest for Elite Rhizobia inoculant strains. Front Microbiol [Internet]. 2017 Nov 9 [cited 2024 Jan 30];8(NOV):2207.
8. Triplett EW, Sadowsky MJ. Genetics of competition for nodulation of legumes. Annu Rev Microbiol [Internet]. 1992/01/01. 1992 Jan 20 [cited 2014 Nov 28];46:399–428.
9. Salas ME, Lozano MJ, López JL, Draghi WO, Serrania J, Torres Tejerizo GA et al. Specificity traits consistent with legume-rhizobia coevolution displayed by *Ensifer meliloti* rhizosphere colonization. Environ Microbiol [Internet]. 2017 Sep 13 [cited 2017 Jun 27];19(9):3423–38.
10. Walker TS, Bais HP, Grotewold E, Vivanco JM. Root exudation and rhizosphere biology. Plant Physiol [Internet]. 2003;132(1):44–51. 2003/05/15.
11. Wheatley RM, Ford BL, Li L, Aroney STN, Knights HE, Ledermann R, S A [Internet]. Lifestyle adaptations of *Rhizobium* from rhizosphere to symbiosis. Proc Natl Acad Sci U. 2020 Sep;22(38):23823–34. [cited 2021 Mar 19];.

12. MacLean AM, Finan TM, Sadowsky MJ. Genomes of the symbiotic nitrogen-fixing bacteria of legumes. Plant Physiol [Internet]. 2007;144(2):615–22. 2007/06/09.

13. Landeta C, Dávalos A, Cevallos MA, Geiger O, Brom S, Romero D. Plasmids with a chromosome-like Role in Rhizobia. J Bacteriol [Internet]. 2011;193(6):1317–26. 2011/01/11.

14. Wardell GE, Hynes MF, Young PJ, Harrison E. Why are rhizobial symbiosis genes mobile? Philos Trans R soc Lond B Biol sci [Internet]. 2022 Jan 17 [cited 2022 Feb 25];377(1842):20200471.

15. Kaneko T. Complete genome structure of the Nitrogen-fixing Symbiotic Bacterium *Mesorhizobium loti*. DNA Res [Internet]. 2000;7(6):331–8.

16. Kaneko T. Complete genomic sequence of Nitrogen-fixing Symbiotic Bacterium *Bradyrhizobium japonicum* USDA110. DNA Res [Internet]. 2002;9(6):189–97.

17. Nelson M, Guhlin J, Epstein B, Tiffin P, Sadowsky MJ. The complete replicons of 16 *Ensifer meliloti* strains offer insights into intra- and inter-replicon gene transfer, transposon-associated loci, and repeat elements. Microb genomics [Internet]. 2018 Apr 19 [cited 2020 Aug 26];4(5):1–11.

18. Siguier P, Gourbeyre E, Chandler M. Bacterial insertion sequences: Their genomic impact and diversity. FEMS Microbiol Rev [Internet]. 2014 Sep 1 [cited 2020 Sep 18];38(5):865–91.

19. Consuegra J, Gaffé J, Lenski RE, Hindré T, Barrick JE, Tenaillon O et al. Insertion-sequence-mediated mutations both promote and constrain evolvability during a long-term experiment with bacteria. Nat Commun [Internet]. 2021 Dec 1 [cited 2021 May 7];12(1):1–12.

20. Vandecraen J, Chandler M, Aertsen A, Van Houdt R. The impact of insertion sequences on bacterial genome plasticity and adaptability [Internet]. Crit Reviews Microbiol Nov. 2017;2:709–30.

21. Iida T, Itakura M, Anda M, Sugawara M, Isawa T, Okubo T et al. H Goodrich-Blair editor 2015 Symbiosis Island shuffling with abundant insertion sequences in the genomes of extra-slow-growing strains of soybean Bradyrhizobia. Appl Environ Microbiol 81 12 4143–54 [Internet].

22. Guo X, Flores M, Mavingui P, Fuentes SI, Hernández G, Dávila G, et al. Natural genomic design in *Sinorhizobium meliloti* : Novel genomic architectures. Genome Res [Internet]. 2003;13(8):1810–7.

23. Arashida H, Odake H, Sugawara M, Noda R, Kakizaki K, Ohkubo S et al. Evolution of rhizobial symbiosis islands through insertion sequence-mediated deletion and duplication. 2021 Jul 16 [cited 2021 Oct 20];16(1).

24. Pellock BJ, Teplitski M, Boinay RP, Bauer WD, Walker GC. A LuxR homolog controls production of symbiotically active extracellular polysaccharide II by *Sinorhizobium meliloti*. J Bacteriol [Internet]. 2002/08/24. 2002;184(18):5067–76.

25. Mogro EG, Ambrosis NM, Lozano MJ. Easy identification of insertion sequence mobilization events in related bacterial strains with ISCompare. G3 Genes|Genomes|Genetics [Internet]. 2021 Aug 7 [cited 2022 May 31];11(8).

26. Kitts PA, Church DM, Thibaud-Nissen F, Choi J, Hem V, Sapojnikov V, et al. Assembly: a resource for assembled genomes at NCBI. Nucleic Acids Res [Internet]. 2016;44(D1):D73–80. [cited 2020 Oct 15];.

27. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat Methods 2020 173 [Internet]. 2020 Feb 3 [cited 2024 Jan 30];17(3):261–72.

28. Cock PJA, Antao T, Chang JT, Chapman BA, Cox CJ, Dalke A et al. Biopython: Freely available Python tools for computational molecular biology and bioinformatics. Bioinformatics [Internet]. 2009 Jun 1 [cited 2020 Sep 18];25(11):1422–3.

29. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinf [Internet]. 2010 Mar 15 [cited 2024 Jan 30];26(6):841–2.

30. Schlüter Jphilip, Reinkensmeier P, Barnett J, Lang MJ, Krol C, Giegerich E. Global mapping of transcription start sites and promoter motifs in the symbiotic α-proteobacterium *Sinorhizobium meliloti* 1021. BMC Genomics [Internet]. 2013 Jan;14(1):156. [cited 2015 Mar 3];.

31. Dam P, Olman V, Harris K, Su Z, Xu Y. Operon prediction using both genome-specific and general genomic information. Nucleic Acids Res [Internet]. 2007 Jan [cited 2024 Aug 9];35(1):288.

32. Salgado H, Moreno-Hagelsieb G, Smith TF, Collado-Vides J. Operons in *Escherichia coli* : genomic analyses and predictions. Proc Natl Acad Sci [Internet]. 2000;97(12):6652–7.

33. Galperin MY, Wolf YI, Makarova KS, Alvarez RV, Landsman D, Koonin EV. COG database update: focus on microbial diversity, model organisms, and widespread pathogens. Nucleic Acids Res [Internet]. 2021 Jan 8 [cited 2024 Jan 30];49(D1):D274–81.

34. Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H et al. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. Nucleic Acids Res [Internet]. 2019 Jan 8 [cited 2022 Mar 2];47(D1):D309–14.

35. Hunter JD, Matplotlib. A 2D Graphics Environment. Comput Sci Eng. 2007;9(03):90–5.

36. Waskom ML. seaborn: statistical data visualization. J Open Source Softw [Internet]. 2021 Apr 6 [cited 2024 Jan 30];6(60):3021.

37. Kaneko T, Maita H, Hirakawa H, Uchiike N, Minamisawa K, Watanabe A et al. Complete genome sequence of the soybean Symbiont *Bradyrhizobium japonicum* strain USDA6T. Genes 2011, 2, Pages 763–87 [Internet]. 2011 Oct 28 [cited 2024 Aug 15];2(4):763–87.

38. Kasai-Maita H, Hirakawa H, Nakamura Y, Kaneko T, Miki K, Maruya J, et al. Commonalities and differences among symbiosis islands of three *Mesorhizobium loti* strains. Microbes Environ [Internet]. 2013 Jun;1(2):275–8. [cited 2024 Aug 15];.

39. Agron PG, Ditta GS, Helinski DR. Oxygen regulation of *nifA* transcription in vitro. Proc Natl Acad Sci [Internet]. 1993;90(8):3506–10.

40. Gong Z, Zhu J, Yu G, Zou H. Disruption of nifA Gene influences multiple Cellular processes in *Sinorhizobium meliloti*. J Genet Genomics [Internet]. 2007;34(9):783–9.

41. Stacey G. *Bradyrhizobium japonicum* nodulation genetics. FEMS Microbiol Lett [Internet]. 1995;127(1–2):1–9.

42. Emptage RP, Pemble CW, York JD, Raetz CRH, Zhou P. Mechanistic characterization of the tetraacyldisaccharide-1-phosphate 4'-kinase LpxK involved in lipid a biosynthesis. Biochem [Internet]. 2013 Apr 2 [cited 2024 Feb 8];52(13):2280–90.

43. García-de los Santos A, Brom S. Characterization of two plasmid-borne lps β loci of *Rhizobium etli* required for Lipopolysaccharide Synthesis and for Optimal Interaction with plants. Mol Plant-Microbe Interact [Internet]. 1997;10(7):891–902.

44. Ormeno-Orrillo E, Rosenblueth M, Luyten E, Vanderleyden J, Martinez-Romero E. Mutations in lipopolysaccharide biosynthetic genes impair maize rhizosphere and root colonization of Rhizobium tropici CIAT899. Env Microbiol. 2008;10(5):1271–84. [Internet]. 2008/03/04.

45. Jenkins CH, Scott AE, O'Neill PA, Norville IH, Prior JL, Ireland PM. The Arabinose 5-Phosphate Isomerase KdsD Is Required for Virulence in *Burkholderia pseudomallei*. O'Toole G, editor. J Bacteriol [Internet]. 2023;205(8).

46. Kiss E, Reuhs BL, Kim JS, Kereszt A, Petrovics G, Putnoky P, et al. The rkpGHI and -J genes are involved in capsular polysaccharide production by *Rhizobium meliloti*. J Bacteriol [Internet]. 1997;179(7):2132. [cited 2024 Apr 3];.

47. Becker A, Fraysse N, Sharypova L. Recent advances in studies on structure and symbiosis-related function of rhizobial K-antigens and lipopolysaccharides. Mol Plant Microbe Interact [Internet]. 2005/09/20. 2005 [cited 2020 Dec 17];18(9):899–905.

48. Lagares A, Hozbor DF, Niehaus K, Otero AJ, Lorenzen J, Arnold W et al. Genetic characterization of a *Sinorhizobium meliloti* chromosomal region in lipopolysaccharide biosynthesis. J Bacteriol. 2001/02/07. 2001;183(4):1248–58.

49. Alav I, Kobylka J, Kuth MS, Pos KM, Picard M, Blair JMA, et al. Structure, Assembly, and function of tripartite efflux and type 1 Secretion systems in Gram-negative Bacteria. Chem Rev [Internet]. 2021;121(9):5479–596.

50. Eda S, Mitsui H, Minamisawa K. Involvement of the *smeAB* multidrug efflux pump in resistance to plant antimicrobials and contribution to nodulation competitiveness in *Sinorhizobium meliloti*. Appl Environ Microbiol. 2011;77(9):2855–62.

51. Kittell BL, Helinski DR, Ditta GS. Aromatic aminotransferase activity and indoleacetic acid production in Rhizobium meliloti. J Bacteriol [Internet]. 1989;171(10):5458–66.

52. Liu Y, Jiang X, Guan D, Zhou W, Ma M, Zhao B, et al. Transcriptional analysis of genes involved in competitive nodulation in *Bradyrhizobium diazoefficiens* at the presence of soybean root exudates. Sci Rep [Internet]. 2017;7(1):10946.

53. Wheatley RM, Ramachandran VK, Geddes BA, Perry BJ, Yost CK, Poole PS. Role of O 2 in the Growth of *Rhizobium leguminosarum* bv. viciae 3841 on Glucose and Succinate. Becker A, editor. J Bacteriol [Internet]. 2017;199(1):e00572-16.

54. Cheng G, Karunakaran R, East AK, Poole PS. Multiplicity of Sulfate and Molybdate transporters and their role in Nitrogen fixation in *Rhizobium leguminosarum* Bv. Viciae Rlv3841. Mol Plant-Microbe Interact [Internet]. 2016;29(2):143–52.

55. Guerrero-Castro J, Lozano L, Sohlenkamp C. Dissecting the acid stress response of *Rhizobium tropici* CIAT 899. Front Microbiol [Internet]. 2018;9.

56. Yuan ZC, Zaheer R, Finan TM. Regulation and properties of PstSCAB, a high-affinity, high-velocity phosphate transport system of *Sinorhizobium meliloti*. J Bacteriol [Internet]. 2006/01/24. 2006;188(3):1089–102.

57. Ruiz B, Le Scornet A, Sauviac L, Rémy A, Bruand C, Meilhoc E. The nitrate assimilatory pathway in *Sinorhizobium meliloti*: contribution to NO production. Front Microbiol [Internet]. 2019 [cited 2021 Jun 24];10(JUL).

58. Prell J, Bourdès A, Karunakaran R, Lopez-Gomez M, Poole P. Pathway of γ-Aminobutyrate metabolism in Rhizobium leguminosarum 3841 and its role in Symbiosis. J Bacteriol [Internet]. 2009;191(7):2177–86.

59. Aneja P, Dziak R, Cai GQ, Charles TC. Identification of an Acetoacetyl Coenzyme A synthetase-dependent pathway for utilization of L-(+)-3-Hydroxybutyrate in *Sinorhizobium meliloti*. J Bacteriol [Internet]. 2002;184(6):1571–7.

60. Cubo MT, Economou A, Murphy G, Johnston AW, Downie JA. Molecular characterization and regulation of the rhizosphere-expressed genes rhiABCR that can influence nodulation by *Rhizobium leguminosarum* biovar viciae. J Bacteriol [Internet]. 1992;174(12):4026–35.

61. Harrison J, Muglia CI, Sype G, Van De, Aguilar OM, Puppo A, Frendo P et al. Glutathione plays a fundamental role in growth and symbiotic capacity of *Sinorhizobium meliloti*. J Bacteriol [Internet]. 2004/12/17. 2005;187(1):168–74.

62. Sobrevals L, Müller P, Fabra A, Castro S. Role of glutathione in the growth of *Bradyrhizobium* sp. (peanut microsymbiont) under different environmental stresses and in symbiosis with the host plant. Can J Microbiol [Internet]. 2006;52(7):609–16.

63. Taté R, Cermola M, Riccio A, Diez-Roux G, Patriarca EJ. Glutathione is required by *Rhizobium etli* for glutamine utilization and Symbiotic Effectiveness. Mol Plant-Microbe Interact [Internet]. 2012;25(3):331–40.

64. Kalloniati C, Tsikou D, Lampiri V, Fotelli MN, Rennenberg H, Chatzipavlidis I, et al. Characterization of a *Mesorhizobium loti* α-Type carbonic anhydrase and its role in Symbiotic Nitrogen fixation. J Bacteriol [Internet]. 2009;191(8):2593–600.

Publisher's note