

SHORT GENOME REPORT

Open Access

Genome sequence of the *Lotus spp.* microsymbiont *Mesorhizobium loti* strain NZP2037

Simon Kelly¹, John Sullivan¹, Clive Ronson¹, Rui Tian², Lambert Bräu³, Karen Davenport⁴, Hajnalka Daligault⁴, Tracy Erkkila⁴, Lynne Goodwin⁴, Wei Gu⁴, Christine Munk⁴, Hazuki Teshima⁴, Yan Xu⁴, Patrick Chain⁴, Tanja Woyke⁵, Konstantinos Liolios⁵, Amrita Pati⁵, Konstantinos Mavromatis⁶, Victor Markowitz⁶, Natalia Ivanova⁵, Nikos Kyrpides^{5,7} and Wayne Reeve^{2*}

Abstract

Mesorhizobium loti strain NZP2037 was isolated in 1961 in Palmerston North, New Zealand from a *Lotus divaricatus* root nodule. Compared to most other *M. loti* strains, it has a broad host range and is one of very few *M. loti* strains able to form effective nodules on the agriculturally important legume *Lotus pedunculatus*. NZP2037 is an aerobic, Gram negative, non-spore-forming rod. This report reveals that the genome of *M. loti* strain NZP2037 does not harbor any plasmids and contains a single scaffold of size 7,462,792 bp which encodes 7,318 protein-coding genes and 70 RNA-only encoding genes. This rhizobial genome is one of 100 sequenced as part of the DOE Joint Genome Institute 2010 *Genomic Encyclopedia for Bacteria and Archaea-Root Nodule Bacteria* (GEBA-RNB) project.

Keywords: Root-nodule bacteria, Nitrogen fixation, Symbiosis, *Alphaproteobacteria*

Introduction

Mesorhizobium loti strain NZP2037 (ICMP1326) was isolated in 1961 from a root nodule off a *Lotus divaricatus* plant growing near Palmerston North airport, New Zealand [1]. Strain NZP2037 is distinguished from most other strains of *M. loti* by its broad host range (see below), including the ability to form effective nodules on the agriculturally important legume *Lotus pedunculatus* (syn. *L. uliginosus*) [2]. Most *M. loti* strains, including the type strain NZP2213, are only able to induce uninfected nodule primordia on this host [2,3].

The ability of *M. loti* strains to form effective nodules on *L. pedunculatus* was correlated with their 'in vitro' sensitivity to flavolans (condensed tannins) present in high concentration in the roots of this legume [4]. The resistance of *M. loti* strain NZP2037 to flavolans from *L. pedunculatus* was associated with the presence of a strain-specific polysaccharide component in the outer cell membrane complex of the bacterium [5]. However the genes required for the synthesis of this flavolan-binding polysaccharide have not been identified and

whether the polysaccharide is necessary for nodulation of *L. pedunculatus* has not been established.

Nodulation and nitrogen fixation genes in *Mesorhizobium loti* strains are encoded on the chromosome on acquired genetic elements termed symbiosis islands [6]. The sequence of the strain NZP2037 symbiosis island was recently reported and it was found that it was split into two regions of 528 kb and 5 kb as the result of a large-scale genome rearrangement [7]. This observation is confirmed by the whole-genome sequence reported in this paper. The Nod factor produced by NZP2037 contains an extra carbamoyl group at its non-reducing end compared to that produced by most other *M. loti* strains [8] and the NZP2037 symbiosis island contains a *nodU* gene that is likely responsible for this modification [7]. The symbiosis island was also found to contain *nodFEGA* genes absent from *M. loti* strain R7A that may lead to the incorporation of unsaturated fatty acid moieties on the Nod factor [7]. Whether these genes contribute to the broad host range of strain NZP2037 has not been reported.

The broad host range of NZP2037 was exploited by Hotter and Scott [9] to show that rhizobial exopolysaccharide was required for the formation of infected nodules on the indeterminate host *Leucaena leucocephala* but not on the determinate nodulating host *L. pedunculatus*. This

* Correspondence: W.Reeve@murdoch.edu.au

²Centre for Rhizobium Studies, Murdoch University, Murdoch, Perth, Australia
Full list of author information is available at the end of the article

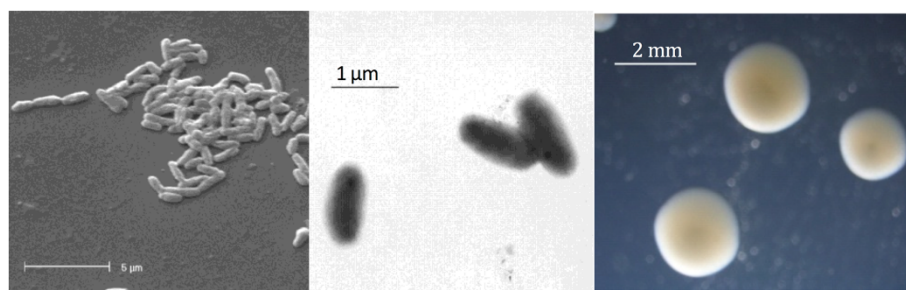


Figure 1 Images of *Mesorhizobium loti* strain NZP2037 using scanning (left) and transmission (center) electron microscopy and the appearance of colony morphology on 1/2LA (right).

Table 1 Classification and general features of *Mesorhizobium loti* strain NZP2037 according to the MIGS recommendations [18,19]

MIGS ID	Property	Term	Evidence code
	Current classification	Domain <i>Bacteria</i>	TAS [19]
		Phylum <i>Proteobacteria</i>	TAS [20]
		Class <i>Alphaproteobacteria</i>	TAS [21]
		Order <i>Rhizobiales</i>	TAS [22,23]
		Family <i>Phyllobacteriaceae</i>	TAS [23,24]
		Genus <i>Mesorhizobium</i>	TAS [16]
		Species <i>Mesorhizobium loti</i>	TAS [3]
		Strain NZP2037	TAS [3]
	Gram stain	Negative	IDA
	Cell shape	Rod	IDA
	Motility	Motile	IDA
	Sporulation	Non-sporulating	NAS
	Temperature range	Mesophile	NAS
	Optimum temperature	28°C	NAS
	Salinity	Unknown	NAS
MIGS-22	Oxygen requirement	Aerobic	TAS [3]
	Carbon source	Various	TAS [16,25]
	Energy source	chemoorganotroph	TAS [16,25]
MIGS-6	Habitat	Soil, root nodule, host	TAS [3]
MIGS-15	Biotic relationship	Free living, Symbiotic	TAS [3]
MIGS-14	Pathogenicity	None	NAS
	Biosafety level	1	TAS [26]
	Isolation	Root nodule of <i>Lotus divaricatus</i>	TAS [27]
MIGS-4	Geographic location	Adjacent Palmerston North Airport, NZ	TAS [1]
MIGS-5	Nodule collection date	1961	TAS [1]
MIGS-4.1	Latitude	-40.1914	TAS [1]
MIGS-4.2	Longitude	175.3701	TAS [1]
MIGS-4.3	Depth	5 cm	IDA
MIGS-4.4	Altitude	46 meters	IDA

Evidence codes – IDA: Inferred from Direct Assay; TAS: Traceable Author Statement (i.e., a direct report exists in the literature); NAS: Non-traceable Author Statement (i.e., not directly observed for the living, isolated sample, but based on a generally accepted property for the species, or anecdotal evidence). These evidence codes are from the Gene Ontology project [28].

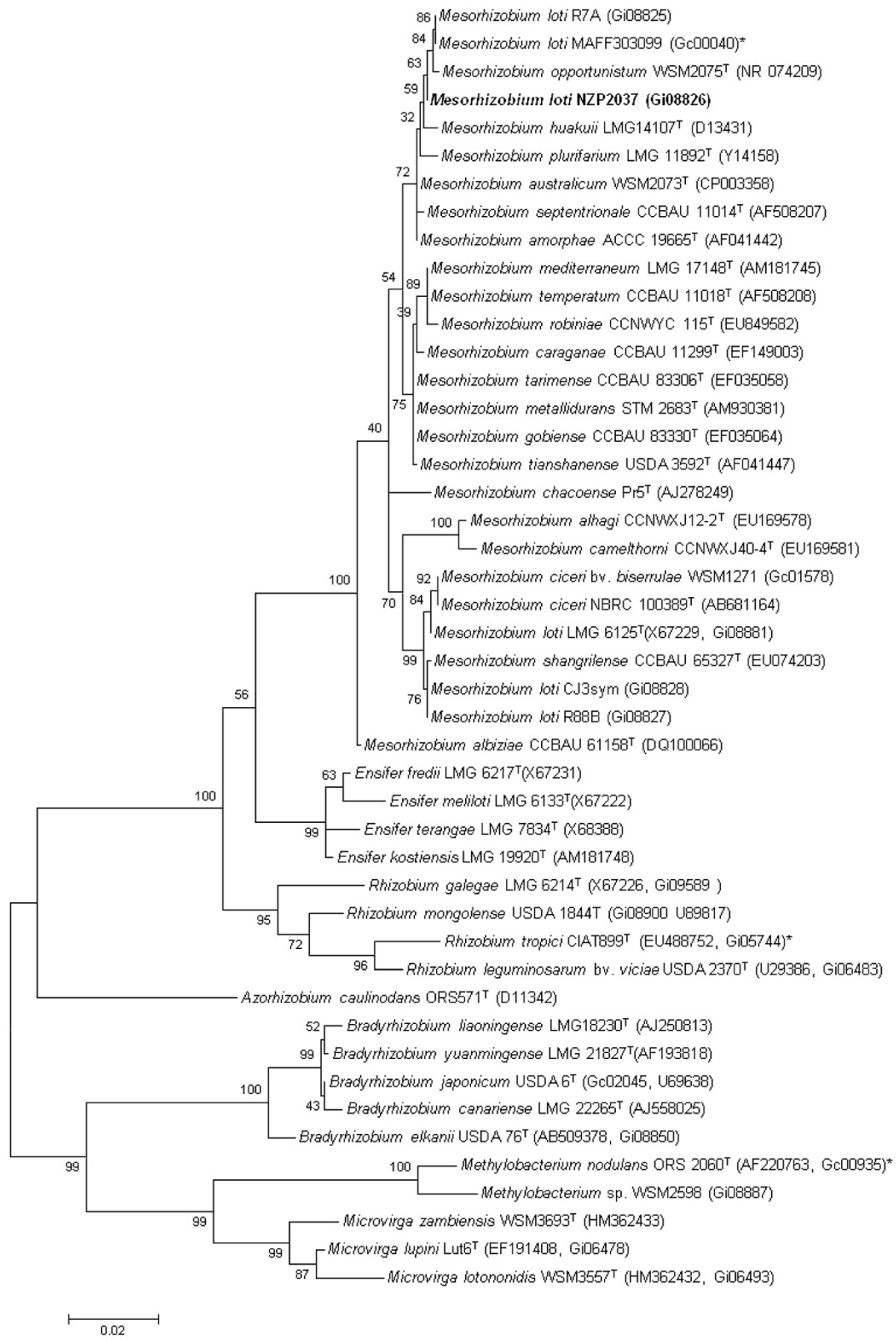


Figure 2 (See legend on next page.)

(See figure on previous page.)

Figure 2 Phylogenetic tree showing the relationships of *Mesorhizobium loti* NZP2037 with other root nodule bacteria based on aligned sequences of the 16S rRNA gene (1,290 bp internal region). All sites were informative and there were no gap-containing sites. Phylogenetic analyses were performed using MEGA [29], version 5. The tree was built using the Maximum-Likelihood method with the General Time Reversible model [30]. Bootstrap analysis [31] with 500 replicates was performed to assess the support of the clusters. Type strains are indicated with a superscript T. Brackets after the strain name contain a DNA database accession number and/or a GOLD ID (beginning with the prefix G) for a sequencing project registered in GOLD [32]. Published genomes are indicated with an asterisk.

observation supported suggestions that acidic EPS is required for effective nodulation of indeterminate but not determinate nodulating legumes (reviewed by [10]). However recent work by Kelly *et al.* using *M. loti* strain R7A showed that certain rhizobial exopolysaccharide mutants including *exoU* mutants induced only uninfected nodules on *L. corniculatus*, supporting a role for exopolysaccharide in determinate nodulation [11]. Interestingly, *exoU* mutants of NZP2037 form effective nodules on *L. corniculatus* [12], again suggesting that NZP2037 may produce a strain-specific surface polysaccharide that plays a symbiotic role.

Here we present a summary classification and a set of general features for *M. loti* strain NZP2037 together with the description of the complete genome sequence and annotation.

Classification and general features

Mesorhizobium loti strain NZP2037 is in the order *Rhizobiales* of the class *Alphaproteobacteria*. Cells are described as non-sporulating, Gram-negative, non-encapsulated, rods. The rod-shaped form varies in size with dimensions of 0.5-0.75 µm in width and 1.25-1.5 µm in length (Figure 1 left and center). They are

Table 2 Genome sequencing project information for *Mesorhizobium loti* NZP2037

MIGS ID	Property	Term
MIGS-31	Finishing quality	High-quality-draft
MIGS-28	Libraries used	Illumina Standard (short PE) and CLIP (long PE) libraries
MIGS-29	Sequencing platforms	Illumina HiSeq2000 technology
MIGS-31.2	Sequencing coverage	Illumina: 509x
MIGS-30	Assemblers	Velvet version 1.1.05; Allpaths-LG version r39750 phrap, version 4.24
MIGS-32	Gene calling method	Prodigal 1.4, GenePRIMP
	Genbank accession	AQZP00000000
	Genbank Registration Date	September 16, 2013
	GOLD ID	Gi08826
	NCBI project ID	81803
	Database: IMG	2517572076
	Project relevance	Symbiotic nitrogen fixation, agriculture

moderately fast growing, forming 2 mm diameter colonies within 5 days and have a mean generation time of approximately 6 h when grown in TY broth at 28°C [13]. Colonies on G/RDM agar [14] and half strength Lupin Agar (½LA) [15] are opaque, slightly domed, mucoid with smooth margins (Figure 1 right).

Strains of this organism are able to tolerate a pH range between 4 and 10. Carbon source utilization and fatty acid profiles of *M. loti* have been described previously [3,16,17]. Minimum Information about the Genome Sequence (MIGS) is provided in Table 1.

Figure 2 shows the phylogenetic neighborhood of *M. loti* strain NZP2037 in a 16S rRNA gene sequence based tree. This strain has 99.7% (1,363/1,367 bp) 16S rRNA gene sequence identity to *M. loti* MAFF303099 (GOLD ID: Gc00040) and 99.6% sequence identity (1,362/1,397 bp) to *M. opportunistum* WSM2075 (GOLD ID: Gc01853).

Symbiotaxonomy

Like most other *M. loti* strains including the type strain NZP2213, strain NZP2037 forms effective nodules on *Lotus corniculatus*, *L. tenuis*, *L. japonicus*, *L. burtii*, *L. krylovii*, *L. filicaulis* and *L. schoelleri* [2,33]. However, it also forms nitrogen-fixing nodules on several hosts that strain NZP2213 only induces uninfected nodules on. These hosts include *Lotus pedunculatus*, *L. angustissimus*,

Table 3 Genome statistics for *Mesorhizobium loti* NZP2037

Attribute	Value	% of total
Genome size (bp)	7,462,792	100.00
DNA coding region (bp)	6,448,323	86.41
DNA G + C content (bp)	4,683,660	62.76
Number of scaffolds	1	
Number of contigs	5	
Total genes	7,388	100.00
RNA genes	70	0.95
rRNA operons	1*	
Protein-coding genes	7,318	99.05
Genes with function prediction	5,982	80.97
Genes assigned to COGs	5,882	79.62
Genes assigned Pfam domains	6,121	82.85
Genes with signal peptides	654	8.85
Genes coding transmembrane proteins	1,735	23.48

*1 copy of 5S, 2 copies of 16S and 1 copy of 23S rRNA genes.

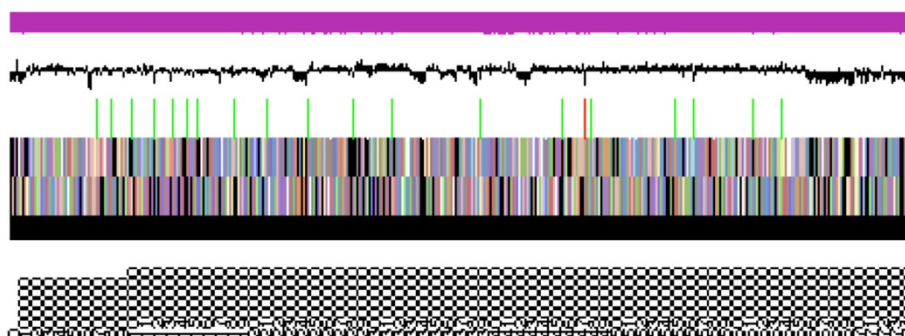


Figure 3 Graphical map of the single scaffold of *Mesorhizobium loti* NZP2037. From bottom to the top: Genes on forward strand (color by COG categories as denoted by the IMG platform), Genes on reverse strand (color by COG categories), RNA genes (tRNAs green, sRNAs red, other RNAs black), GC content, GC skew.

L. subbiflorus, *Leuceana leucocephala*, *Carmichaelia flagelliformis*, *Ornithopus sativus* and *Clianthus puniceus* [33].

Genome sequencing and annotation information

Genome project history

This organism was selected for sequencing on the basis of its environmental and agricultural relevance to issues in global carbon cycling, alternative energy production, and biogeochemical importance, and is part of the Community Sequencing Program at the U.S. Department of Energy, Joint Genome Institute (JGI) for projects of relevance to agency missions. The genome project is deposited in the Genomes OnLine Database [32] and a high-quality-draft genome sequence in IMG. Sequencing, finishing and annotation were performed by the JGI. A summary of the project information is shown in Table 2.

Growth conditions and DNA isolation

M. loti strain NZP2037 was grown to mid logarithmic phase in TY rich medium [34] on a gyratory shaker at 28°C at 250 rpm. DNA was isolated from 60 mL of cells using a CTAB (Cetyl trimethyl ammonium bromide) bacterial genomic DNA isolation method [35].

Genome sequencing and assembly

The draft genome of *M. loti* NZP2037 was generated at the DOE Joint Genome Institute (JGI) using Illumina technology [36]. For this genome, we constructed and sequenced an Illumina short-insert paired-end library with an average insert size of 270 bp which generated 9,401,642 reads and an Illumina long-insert paired-end library with an average insert size of 3047.66 +/- 2184.11 bp which generated 16,067,290 reads totaling 3,820 Mbp of Illumina data. (unpublished, Feng Chen). All general aspects of library construction and sequencing performed at the JGI can be found at the JGI website [37].

The initial draft assembly contained 13 contigs in 6 scaffolds. The initial draft data was assembled with

Table 4 Number of protein coding genes of *Mesorhizobium loti* NZP2037 associated with the general COG functional categories

Code	Value	% age	COG category
J	205	3.14	Translation, ribosomal structure and biogenesis
A	0	0.00	RNA processing and modification
K	603	9.24	Transcription
L	242	3.71	Replication, recombination and repair
B	7	0.11	Chromatin structure and dynamics
D	34	0.52	Cell cycle control, mitosis and meiosis
Y	0	0.00	Nuclear structure
V	82	1.26	Defense mechanisms
T	262	4.01	Signal transduction mechanisms
M	320	4.90	Cell wall/membrane biogenesis
N	51	0.78	Cell motility
Z	2	0.03	Cytoskeleton
W	1	0.02	Extracellular structures
U	142	2.17	Intracellular trafficking and secretion
O	197	3.02	Posttranslational modification, protein turnover, chaperones
C	381	5.84	Energy production conversion
G	624	9.56	Carbohydrate transport and metabolism
E	798	12.22	Amino acid transport metabolism
F	95	1.46	Nucleotide transport and metabolism
H	242	3.71	Coenzyme transport and metabolism
I	275	4.21	Lipid transport and metabolism
P	267	4.09	Inorganic ion transport and metabolism
Q	213	3.26	Secondary metabolite biosynthesis, transport and catabolism
R	811	12.42	General function prediction only
S	675	10.34	Function unknown
-	1,506	20.38	Not in COGS

Allpaths, version 39750, and the consensus was computationally shredded into 10 Kbp overlapping fake reads (shreds). The Illumina draft data was also assembled with Velvet [38], version 1.1.05, and the consensus sequences were computationally shredded into 1.5 Kbp overlapping fake reads (shreds). The Illumina draft data was assembled again with Velvet using the shreds from the first Velvet assembly to guide the next assembly. The consensus from the second VELVET assembly was shredded into 1.5 Kbp overlapping fake reads. The fake reads from the Allpaths assembly and both Velvet assemblies and a subset of the Illumina CLIP paired-end reads were assembled using parallel phrap, version 4.24 (High Performance Software, LLC). Possible mis-assemblies were corrected with manual editing in Consed [38-41]. Gap closure was accomplished using repeat resolution software (Wei Gu, unpublished), and sequencing of bridging PCR fragments with Sanger technology. The total ("estimated size" for unfinished) size of the genome is 7.5 Mbp and the final assembly is based on 3,820 Mbp of Illumina draft data, which provides an average 509× coverage of the genome.

Genome annotation

Genes were identified using Prodigal [42] as part of the DOE-JGI genome annotation pipeline, followed by a round of manual curation using the JGI GenePrimp pipeline [43]. The predicted CDSs were translated and used to search the National Center for Biotechnology Information (NCBI) nonredundant database, UniProt, TIGRFam, Pfam, PRIAM, KEGG, COG, and InterPro databases. These data sources were combined to assert a product description for each predicted protein. Non-coding genes and miscellaneous features were predicted using tRNAscan-SE [44], RNAMer [45], Rfam [46], TMHMM [47], and SignalP [48]. Additional gene prediction analyses and functional annotation were performed within the Integrated Microbial Genomes (IMG-ER) platform [49,50].

Genome properties

The genome is 7,462,792 nucleotides with 62.76% GC content (Table 3 and Figure 3) and is comprised of a single scaffold and no plasmids. From a total of 7,388 genes, 7,318 were protein encoding and 70 RNA-only encoding genes. Within the genome, 286 pseudogenes were also identified. The majority of genes (80.97%) were assigned a putative function while the remaining genes were annotated as hypothetical. The distribution of genes into COGs functional categories is presented in Table 4.

Conclusion

The *M. loti* NZP2037 genome consists of a single chromosome of 7.46 Mb predicted to encode 7,388 genes. The sequencing was completed to the stage where a single scaffold comprising 5 contigs was obtained. NZP2037

differs from other well-characterised *M. loti* strains in that it is able to form effective nodules on the host *L. pedunculatus* (*syn. L. uliginosus*) [2]. The molecular basis of this extended host range remains unknown; however NZP2307 carries additional *nod* genes (*nodU*, *nodFEG* and a second copy of *nodA*) not found in other well-characterised *M. loti* strains such as MAFF303099 and R7A [7]. Preliminary studies suggest it may also produce some different surface polysaccharides to R7A [11,12].

Previously it was demonstrated that NZP2037 contains a transmissible plasmid of 240 MDa (approximately 360 kb) designated pRlo22037a [25]. Strain PN4010, a plasmid-cured derivative of NZP2037, showed enhanced levels of nitrogen fixation and competitiveness on *Lotus pedunculatus* versus the wild-type. Reintroduction of the plasmid into PN4010 returned the strain to the wild-type phenotype [51]. A type IV secretion system consisting of a *trb* gene cluster (Locus tags 7041-7051 coordinates 70104004-7113626) and *traG* (locus tag 6995 coordinates 7068484-7070472) highly similar (80-98% amino acid identity) to that of the *M. loti* strain MAFF303099 pMlb plasmid are located at the end of the scaffold. This finding and comparison of the genome sequence with that of *M. loti* strains R7A and MAFF303099 suggests that the right end of the single large scaffold may in fact be a large plasmid.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

JS and CR supplied the strain and background information for this project and contributed to the assembly of the manuscript with WR, TR supplied DNA to JGI and performed all imaging, WR coordinated the project and all other authors were involved in either sequencing the genome and/or editing the paper. All authors read and approved the final manuscript.

Acknowledgements

This work was performed under the auspices of the US Department of Energy Office of Science, Biological and Environmental Research Program, and by the University of California, Lawrence Berkeley National Laboratory under contract No. DE-AC02-05CH11231, Lawrence Livermore National Laboratory under Contract No. DE-AC52-07NA27344, and Los Alamos National Laboratory under contract No. DE-AC02-06NA25396.

Author details

¹Department of Microbiology and Immunology, University of Otago, Dunedin, New Zealand. ²Centre for Rhizobium Studies, Murdoch University, Murdoch, Perth, Australia. ³School of Life and Environmental Sciences, Deakin University, Deakin, Victoria, Australia. ⁴Los Alamos National Laboratory, Bioscience Division, Los Alamos, NM, USA. ⁵DOE Joint Genome Institute, Walnut Creek, CA, USA. ⁶Biological Data Management and Technology Center, Lawrence Berkeley National Laboratory, Berkeley, CA, USA. ⁷Department of Biological Sciences, King Abdulaziz University, Jeddah, Saudi Arabia.

Received: 13 June 2014 Accepted: 16 June 2014

Published: 8 December 2014

References

1. Young JM, Fletcher MJ. *International Collection of Micro-organisms from Plants Catalogue*. 3rd ed. Landcare Research; 1997. http://scd.landcareresearch.co.nz/Specimen/ICMP_1326

2. Pankhurst CE, Craig AS, Jones WT. Effectiveness of *Lotus* root nodules. 1: morphology and flavolan content of nodules formed on *Lotus pedunculatus* by fast-growing *Lotus* rhizobia. *J Exp Bot*. 1979; **30**: 1085-93. <http://dx.doi.org/10.1093/jxb/30.6.1085>
3. Jarvis BDW, Pankhurst CE, Patel JJ. *Rhizobium loti*, a new species of legume root nodule bacteria. *Int J Syst Bacteriol*. 1982; **32**: 378-80. <http://dx.doi.org/10.1099/00207713-32-3-378>
4. Pankhurst CE, Jones WT. Effectiveness of *Lotus* root nodules. 2: relationship between root nodule effectiveness and in vitro sensitivity of fast-growing *Lotus* rhizobia to flavolans. *J Exp Bot* 1979; **30**: 1095-107. <http://dx.doi.org/10.1093/jxb/30.6.1095>
5. Jones WT, Macdonald PE, Jones SD, Pankhurst CE. Peptidoglycan-bound polysaccharide associated with resistance of *Rhizobium loti* strain NZP2037 to *Lotus pedunculatus* root flavolan. *J Gen Microbiol*. 1987; **133**: 2617-29
6. Sullivan JT, Ronson CW. Evolution of rhizobia by acquisition of a 500-kb symbiosis island that integrates into a phe-tRNA gene. *Proc Natl Acad Sci U S A*. 1998; **95**: 5145-9. PubMed <http://dx.doi.org/10.1073/pnas.95.9.5145>
7. Kasai-Maita H, Hirakawa H, Nakamura Y, Kaneko T, Miki K, Maruya J, Okazaki S, Tabata S, Saeki K, Sato S. Commonalities and differences among symbiosis islands of three *Mesorhizobium loti* strains. *Microbes Environ*. 2013; **28**: 275-8. PubMed <http://dx.doi.org/10.1264/jsm.2.ME12201>
8. Lopez-Lara IM, Van den Berg DJ, Thomas-Oates JE, Glushka J, Lugtenberg BJJ, Spaink HP. Structural identification of the lipo-chitin oligosaccharide nodulation signals of *Rhizobium loti*. *Mol Microbiol*. 1995; **15**: 627-38. PubMed <http://dx.doi.org/10.1111/j.1365-2958.1995.tb02372.x>
9. Hotter GS, Scott DB. Exopolysaccharide mutants of *Rhizobium loti* are fully effective on a determinate nodulating host but are ineffective on an indeterminate nodulating host. *J Bacteriol*. 1991; **173**: 851-9. PubMed
10. Gage DJ. Infection and invasion of roots by symbiotic, nitrogen-fixing rhizobia during nodulation of temperate legumes. *Microbiol Mol Biol Rev*. 2004; **68**: 280-300. PubMed <http://dx.doi.org/10.1128/MMBR.68.2.280-300.2004>
11. Kelly SJ, Muszynski A, Kawaharada Y, Hubber AM, Sullivan JT, Sandal N, Carlson RW, Stougaard J, Ronson CW. Conditional requirement for exopolysaccharide in the *Mesorhizobium-Lotus* symbiosis. *Mol Plant Microbe Interact*. 2013; **26**: 319-29. PubMed <http://dx.doi.org/10.1094/MPMI-09-12-0227-R>
12. Kelly SJ. Requirement for Exopolysaccharide in the *Mesorhizobium-Lotus* Symbiosis. Dunedin: University of Otago; 2012: p.259.
13. Sullivan JT, Patrick HN, Lowther WL, Scott DB, Ronson CW. Nodulating strains of *Rhizobium loti* arise through chromosomal symbiotic gene transfer in the environment. *Proc Natl Acad Sci U S A*. 1995; **92**: 8985-9. PubMed <http://dx.doi.org/10.1073/pnas.92.19.8985>
14. Ronson CW, Nixon BT, Albright LM, Ausubel FM. *Rhizobium meliloti ntrA (rpoN)* gene is required for diverse metabolic functions. *J Bacteriol*. 1987; **169**: 2424-31. PubMed
15. Howieson JG, Ewing MA, D'antuono MF. Selection for acid tolerance in *Rhizobium meliloti*. *Plant Soil*. 1988; **105**: 179-88. <http://dx.doi.org/10.1007/BF02376781>
16. Jarvis BDW, Van Berkum P, Chen WX, Nour SM, Fernandez MP, Cleyet-Marel JC, Gillis M. Transfer of *Rhizobium loti*, *Rhizobium huakuii*, *Rhizobium ciceri*, *Rhizobium mediterraneum*, *Rhizobium tianshanense* to *Mesorhizobium gen.nov.* *Int J Syst Evol Microbiol*. 1997; **47**: 895-8
17. Tighe SW, De Lajudie P, Dipietro K, Lindstrom K, Nick G, Jarvis BDW. Analysis of cellular fatty acids and phenotypic relationships of *Agrobacterium*, *Bradyrhizobium*, *Mesorhizobium*, *Rhizobium* and *Sinorhizobium* species using the Sherlock Microbial Identification System. *Int J Syst Evol Microbiol*. 2000; **50**: 787-801. PubMed <http://dx.doi.org/10.1099/00207713-50-2-787>
18. Field D, Garrity G, Gray T, Morrison N, Selengut J, Sterk P, Tatusova T, Thomson N, Allen M, Angiuoli SV, et al. Towards a richer description of our complete collection of genomes and metagenomes "Minimum Information about a Genome Sequence" (MIGS) specification. *Nat Biotechnol*. 2008; **26**: 541-7. PubMed <http://dx.doi.org/10.1038/nbt1360>
19. Woese CR, Kandler O, Wheelis ML. Towards a natural system of organisms: proposal for the domains *Archaea*, *Bacteria*, and *Eucarya*. *Proc Natl Acad Sci U S A*. 1990; **87**: 4576-9. PubMed <http://dx.doi.org/10.1073/pnas.87.12.4576>
20. Garrity GM, Bell JA, Lilburn T. Phylum XIV. *Proteobacteria* phyl. nov. In: *Bergey's Manual of Systematic Bacteriology. Volume 2*. 2nd ed. Garrity GM, Brenner DJ, Krieg NR, Staley JT, editors. New York: Part B, Springer; 2005: p.1.
21. Garrity GM, Bell JA, Lilburn T. Class I. *Alphaproteobacteria* class. In: *Bergey's Manual of Systematic Bacteriology*. 2nd ed. Garrity GM, Brenner DJ, Krieg NR, Staley JT, editors. New York: Springer - Verlag; 2005
22. Kuykendall LD. Order VI. *Rhizobiales* ord. nov. In: *Bergey's Manual of Systematic Bacteriology*. 2nd ed. Garrity GM, Brenner DJ, Krieg NR, Staley JT, editors. New York: Springer - Verlag; 2005. p.324
23. Validation List No. 107. List of new names and new combinations previously effectively, but not validly, published. *Int J Syst Evol Microbiol*. 2006; **56**: 1-6. PubMed <http://dx.doi.org/10.1099/ijs.0.64188-0>
24. Mergaert J, Swings J. Family IV. *Phyllobacteriaceae* fam. nov. In: *Bergey's Manual of Systematic Bacteriology. Volume 2*. 2nd ed. Garrity GM, Brenner DJ, Krieg NR, Staley JT, editors. New York: Part C, Springer; 2005, p.393
25. Pankhurst CE, Broughton WJ, Wieneke U. Transfer of an indigenous plasmid of *Rhizobium loti* to other rhizobia and *Agrobacterium tumefaciens*. *J Gen Microbiol*. 1983; **129**: 2535-43
26. Biological Agents. Technical Rules for Biological Agents. TRBA: p 466. <http://www.baua.de>
27. Crow VL, Jarvis BDW, Greenwood RM. Deoxyribonucleic acid homologies among acid-producing strains of *Rhizobium*. *Int J Syst Bacteriol*. 1981; **31**: 152-72. <http://dx.doi.org/10.1099/00207713-31-2-152>
28. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al. Gene ontology: tool for the unification of biology: the gene ontology consortium. *Nat Genet*. 2000; **25**: 25-29. PubMed <http://dx.doi.org/10.1038/75556>
29. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol*. 2011; **28**: 2731-9. PubMed <http://dx.doi.org/10.1093/molbev/msr121>
30. Nei M, Kumar S. *Molecular Evolution and Phylogenetics*. New York: Oxford University Press; 2000
31. Felsenstein J. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*. 1985; **39**: 783-91. <http://dx.doi.org/10.2307/2408678>
32. Liolios K, Mavromatis K, Tavernarakis N, Kyrpides NC. The Genomes On Line Database (GOLD) in 2007: status of genomic and metagenomic projects and their associated metadata. *Nucleic Acids Res*. 2008; **36**: D475-9. PubMed <http://dx.doi.org/10.1093/nar/gkm884>
33. Pankhurst CE, Hopcroft DH, Jones WT. Comparative morphology and flavolan content of *Rhizobium loti* induced effective and ineffective root nodules on *Lotus* species, *Leuceana leucocephala*, *Carmichaelia flagelliformis*, *Ornithopus sativus*, and *Clianthus puniceus*. *Can J Bot*. 1987; **65**: 2676-85. <http://dx.doi.org/10.1139/b87-358>
34. Beringer JE. R factor transfer in *Rhizobium leguminosarum*. *J Gen Microbiol*. 1974; **84**: 188-98. PubMed <http://dx.doi.org/10.1099/00221287-84-1-188>
35. DOE Joint Genome Institute user homepage. <http://1ofdmq2n8tc36m6i46scovo2e.wengine.netdna-cdn.com/wp-content/uploads/2014/02/JGI-Bacterial-DNA-isolation-CTAB-Protocol-2012.pdf>
36. Bennett S. Solexa Ltd. *Pharmacogenomics*. 2004; **5**: 433-8. PubMed <http://dx.doi.org/10.1517/14622416.5.4.433>
37. DOE Joint Genome Institute. <http://www.jgi.doe.gov>
38. Zerbino DR. Using the Velvet de novo assembler for short-read sequencing technologies. *Curr Protocols Bioinformatics*. 2010; Chapter 11:Unit 11.5
39. Ewing B, Green P. Base-calling of automated sequencer traces using phred. II: error probabilities. *Genome Res*. 1998; **8**: 186-94. PubMed <http://dx.doi.org/10.1101/gr.8.3.175>
40. Ewing B, Hillier L, Wendl MC, Green P. Base-calling of automated sequencer traces using phred. I: accuracy assessment. *Genome Res*. 1998; **8**: 175-85. PubMed <http://dx.doi.org/10.1101/gr.8.3.175>
41. Gordon D, Abajian C, Green P. Consed: a graphical tool for sequence finishing. *Genome Res*. 1998; **8**: 195-202. PubMed <http://dx.doi.org/10.1101/gr.8.3.195>
42. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics*. 2010; **11**: 119. PubMed <http://dx.doi.org/10.1186/1471-2105-11-119>
43. Pati A, Ivanova NN, Mikhailova N, Ovchinnikova G, Hooper SD, Lykidis A, Kyrpides NC. GenePRIMP: a gene prediction improvement pipeline for prokaryotic genomes. *Nat Methods*. 2010; **7**: 455-457. PubMed <http://dx.doi.org/10.1038/nmeth.1457>
44. Lowe TM, Eddy SR. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res*. 1997; **25**: 955-64. PubMed <http://dx.doi.org/10.1093/nar/25.5.0955>

45. Lagesen K, Hallin P, Rodland EA, Staerfeldt HH, Rognes T, Ussery DW. **RNAmmer: consistent and rapid annotation of ribosomal RNA genes.** *Nucleic Acids Res.* 2007; **35**: 3100-8. PubMed <http://dx.doi.org/10.1093/nar/gkm160>
46. Griffiths-Jones S, Bateman A, Marshall M, Khanna A, Eddy SR. **Rfam: an RNA family database.** *Nucleic Acids Res.* 2003; **31**: 439-41. PubMed <http://dx.doi.org/10.1093/nar/gkg006>
47. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. **Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes.** *J Mol Biol.* 2001; **305**: 567-80. PubMed <http://dx.doi.org/10.1006/jmbi.2000.4315>
48. Bendtsen JD, Nielsen H, von Heijne G, Brunak S. **Improved prediction of signal peptides: SignalP 3.0.** *J Mol Biol.* 2004; **340**: 783-95. PubMed <http://dx.doi.org/10.1016/j.jmb.2004.05.028>
49. Markowitz VM, Mavromatis K, Ivanova NN, Chen IM, Chu K, Kyrpides NC. **IMG ER: a system for microbial genome annotation expert review and curation.** *Bioinformatics.* 2009; **25**: 2271-78. PubMed <http://dx.doi.org/10.1093/bioinformatics/btp393>
50. Genomes IM. (*IMG-ER*) platform. <http://img.jgi.doe.gov/er>
51. Pankhurst CE, Macdonald PE, Reeves JM. **Enhanced nitrogen fixation and competitiveness for nodulation of *Lotus pedunculatus* by a plasmid-cured derivative of *Rhizobium loti*.** *J Gen Microbiol.* 1986; **132**: 2321-2328.

doi:10.1186/1944-3277-9-7

Cite this article as: Kelly et al.: Genome sequence of the *Lotus* spp. microsymbiont *Mesorhizobium loti* strain NZP2037. *Standards in Genomic Sciences* 2014 **9**:7.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

