Article:
Zhang, X., Li, F., Young, J.P.W. et al. (2002) What does a bacterial genome sequence represent? Mis-assignment of MAFF 303099 to the genospecies Mesorhizobium loti. Microbiology. pp. 3330-3331. ISSN: 1465-2080

sMMO genes are expressed in heterologous hosts. We have recently identified a *groEL* gene located 5′ of the *mmoX* gene cluster in *M. trichosporium* OB3b which, when mutated, results in a mutant with an sMMO-minus phenotype (Murrell *et al*., unpublished observations). This *groEL*, which is not present on the sMMO gene cluster constructs used in experiments by Wood and colleagues (1, 2), may be essential for the correct assembly of the sMMO or sMMO regulatory proteins, which could account for the high level expression of sMMO in the homologous host since this *groEL* is present both on the chromosome of *M. trichosporium* OB3b and on the expression plasmid used (3).

J. Colin Murrell

Department of Biological Sciences, University of Warwick, Coventry, CV4 7AL, UK

Tel: +44 24 76523553. Fax: +44 24 76523568. e-mail: cmurrell@bio.warwick.ac.uk

1. Jahng, D. & Wood, T. K. (1994). Trichloroethylene and chloroform degradation by a recombinant pseudomonad expressing soluble methane monooxygenase from *Methylosinus trichosporium* OB3b. *Appl Environ Microbiol* **60**, 2473–2482.

2. Jahng, D., Kim, C. S., Hanson, R. S. & Wood, T. K. (1996). Optimization of trichloroethylene degradation using soluble methane monooxygenase of *Methylosinus trichosporium* OB3b expressed in recombinant bacteria. *Biotechnol Bioeng* **51**, 349–359.

3. Lloyd, J. S., Finch, R., Dalton, H. & Murrell, J. C. (1999). Homologous expression of soluble methane monooxygenase genes in *Methylosinus trichosporium* OB3b. *Microbiology* **145**, 461–470.

What does a bacterial genome sequence represent? Misassignment of MAFF 303099 to the genospecies *Mesorhizobium loti*

Doolittle (1) recently discussed the potential and limitations of bacterial genome sequencing, and emphasized the importance of lateral gene transfers in bacterial adaptation, e.g. to become pathogens. The transferable genes, or 'accessory genome', should comprise genes that are advantageous intermittently and are not uniformly distributed among individuals of a species, though they may be shared between species. The important phenotypes that they encode are often of economic interest to medicine, e.g. pathogenicity islands and antibiotic resistance, and agriculture, e.g. symbiotic nitrogen fixation by rhizobia in association with plant roots. Such traits are sometimes used to identify bacteria, but this can lead to an 'identity crisis'. This is illustrated by the largest genome that Doolittle (1) discusses, that of a

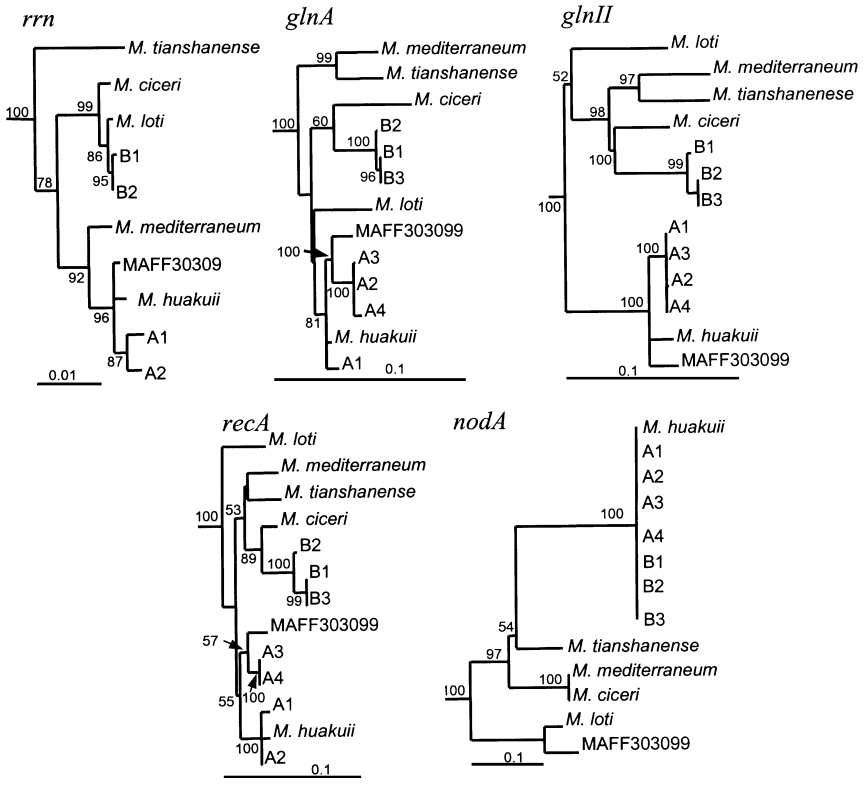

Fig. 1. Phylogenetic trees showing evolutionary relationships among *Mesorhizobium* strains for five partial gene sequences: 16S *rrn* (606–1420 bp, gaps ignored in pair-wise comparisons), *glnA* (936 bp), *glnII* (828 bp), *recA* (440 bp) and *nodA* (470 bp). Sequences labelled A1–4 (strains HN8A1, ZJ15B7, JX2B5 and JS5A15) and B1–3 (HB5A4, HN14A16 and HN15B23) are from isolates that nodulate *Astragalus sinicus* (7), MAFF 303099 indicates sequences taken from the published genome sequence. These sequences were amplified and sequenced according to the methods described in and compared to the corresponding sequences of *Mesorhizobium* type strains described in: ref. 2 for 16S *rrn* and *recA*; ref. 5 for *glnA* and *glnII* and ref. 7 for *nodA*. All trees were calculated using PAUP* (version 4.0b10; http://paup.csit.fsu.edu/) using the neighbour-joining algorithm with the HKY model. Trees were rooted using the corresponding *Sinorhizobium meliloti* sequence (not shown on trees). Bootstrap percentages shown on each branch were estimated from 1000 replicates. New sequences for the *A. sinicus* isolates have been deposited in the databases under accession numbers AJ459585–AJ459605.

strain of rhizobia (MAFF 303099) isolated from a *Lotus corniculatus* root nodule. The strain was identified as *Mesorhizobium loti* but we suggest, based on phylogeny, that it belongs to a different species: *Mesorhizobium huakuii* biovar *loti*. Host-range variants in rhizobia are generally called 'biovars', in which the chromosome type (basic genome) defines the species and the symbiosis type (accessory genome) the biovar. MAFF 303099 differs considerably from the *M. loti* type strain at four chromosomal loci (*rrn*, *glnA*, *glnII*, *recA*; Fig. 1) and more resembles the type strain of *M. huakuii* (every bacterial species is defined by similarity to a single representative, the 'type strain').

The affiliation with *M. huakuii* is particularly striking for glutamine synthetase II (*glnII*), as MAFF 303099 shares a distinctive sequence thought to have been acquired by the ancestor of the *M. huakuii* lineage through lateral gene transfer from a *Rhizobium*-like species (5). We found this same *glnII* signature in four strains representing the dominant (> 96%) chromosomal type (6) among 204 isolates from *Astragalus sinicus*, the typical host of *M. huakuii* (type A1–4 in Fig. 1). Three strains representing the remaining genotypic diversity belong to a different lineage (type B1–3 in Fig. 1). We conclude that CCBAU 2609^T is truly representative of *M. huakuii*, and that MAFF 303099 belongs to this species.

In contrast, MAFF 303099 has symbiosis genes (represented here by *nodA*) that are most similar to those of other *Lotus* symbionts (Fig. 1). Symbiosis genes are mobile, being usually found on plasmids or transmissible genetic islands. All *A. sinicus* symbionts screened to date have identical *nodA* sequences (7), indicating relatively recent (in evolutionary time) transfer of their symbiosis genes between species. More directly, Sullivan *et al*. (4) detected transfer of the genes for nodulation of *Lotus* from an inoculant strain into several different chromosomal backgrounds within four years.

Lateral gene transfer thus shapes the bac-

terial genome on two different time scales. The transfer of *glnII* from *Rhizobium* to *Mesorhizobium* was apparently a single, rare event, since other basic 'housekeeping' genes generally share a consistent phylogeny (2, 5). By contrast, the accessory genome, here represented by symbiosis genes, undergoes detectable transfers within and between species. Accessory DNA makes up 10–25 % of the DNA in the three rhizobial genomes sequenced so far [*M. loti* (NC_002678), *Sinorhizobium meliloti* (NC_003047) and *Agrobacterium tumefaciens* (NC_003305)].

Sequencing an individual bacterial genome will provide a clear picture of the basic genome of the species, but only an arbitrary 'snapshot' of that part of the accessory gene pool that happens to be in the chosen isolate. When genome sequences are available from more than one strain of a species, e.g. *Escherichia coli*, we see clearly that they differ in their complement of accessory genes (3), so that a single strain does not adequately represent the whole species. Characterizing and sequencing more than one member of a species is therefore important, since this will define the common core of genes which defines the species. It will also allow identification of the associated accessory gene pool that provides a species with its adaptations to different niches and determines the variety of key properties such as symbioses or diseases with which the species is associated (1).

Sarah L. Turner,[1]† Xue-Xian Zhang,[2]† Fu-Di Li[2] and J. Peter W. Young[1]

[1]Department of Biology, University of York, PO Box 373, York YO10 5YW, UK.
[2]Department of Microbiology, Huazhong Agricultural University, China.
†Present address: CEH Oxford, Mansfield Road, Oxford OX1 3SR, UK.

Author for correspondence: Sarah L. Turner. Tel: +44 1865 281 630. Fax: +44 1865 281 696. e-mail: sltu@ceh.ac.uk

1. Doolittle, R. F. (2002). Microbial genomes multiply. *Nature* **416**, 697–700.
2. Gaunt, M. J., Turner, S. L., Rigottier-Gois, L., Lloyd-Macgilp, S. A. & Young, J. P. W. (2001). Phylogenies of *atpD* and *recA* support the SSU rRNA-based classification of rhizobia. *Int J Syst Evol Microbiol* **51**, 2037–2048.
3. Perna, N. T., Plunkett, G., Burland, V. & 25 other authors (2001). Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. *Nature* **409**, 529–533.
4. Sullivan, J. T., Eardly, D. B., van Berkum, P. & Ronson, C. W. (1996). Four unnamed species of nonsymbiotic rhizobia isolated from the rhizosphere of *Lotus corniculatus*. *Appl Environ Microbiol* **62**, 2818–2825.
5. Turner, S. L. & Young, J. P. W. (2000). The glutamine synthetases of rhizobia: phylogenetics and evolutionary implications. *Mol Biol Evol* **17**, 309–319.
6. Zhang, X. X., Guo, X. W., Terefework, Z., Cao, Y. Z., Hu, F. R, Lindstrom, K & Li, F. D. (1999). Genetic diversity among rhizobial isolates from field grown *Astragalus sinicus* of Southern China. *Syst Appl Microbiol* **22**, 312–320.
7. Zhang, X. X., Turner, S. L., Guo, X. W., Yang, H. J., Debellé, F., Yang, G. P., Dénarié, J., Young, J. P. W. & Li, F. D. (2000). The common nodulation genes of *Astragalus sinicus* rhizobia are conserved despite chromosomal diversity. *Appl Environ Microbiol* **66**, 2988–2995.

MprF-mediated lysinylation of phospholipids in *Bacillus subtilis* – protection against bacteriocins in terrestrial habitats?

A common strategy for organisms to inhibit bacteria is the production of antimicrobial peptides that damage bacterial membranes and that usually have cationic properties to enable efficient interactions with the anionic polymers in bacterial cell envelopes. Such cationic antimicrobial peptides (CAMPs) are produced by the innate immune systems of humans, animals and plants (e.g. defensins) (7) and many of the bacterial bacteriocins belong to the same class of molecules (3).

The recently described bacterial mechanisms protecting against a wide range of CAMPs may be beneficial in many kinds of environments. Most of these mechanisms involve modulations of the net charge in the bacterial cell envelope to reduce accumulation of cationic peptides (4). Modification of teichoic acids with positively charged D-

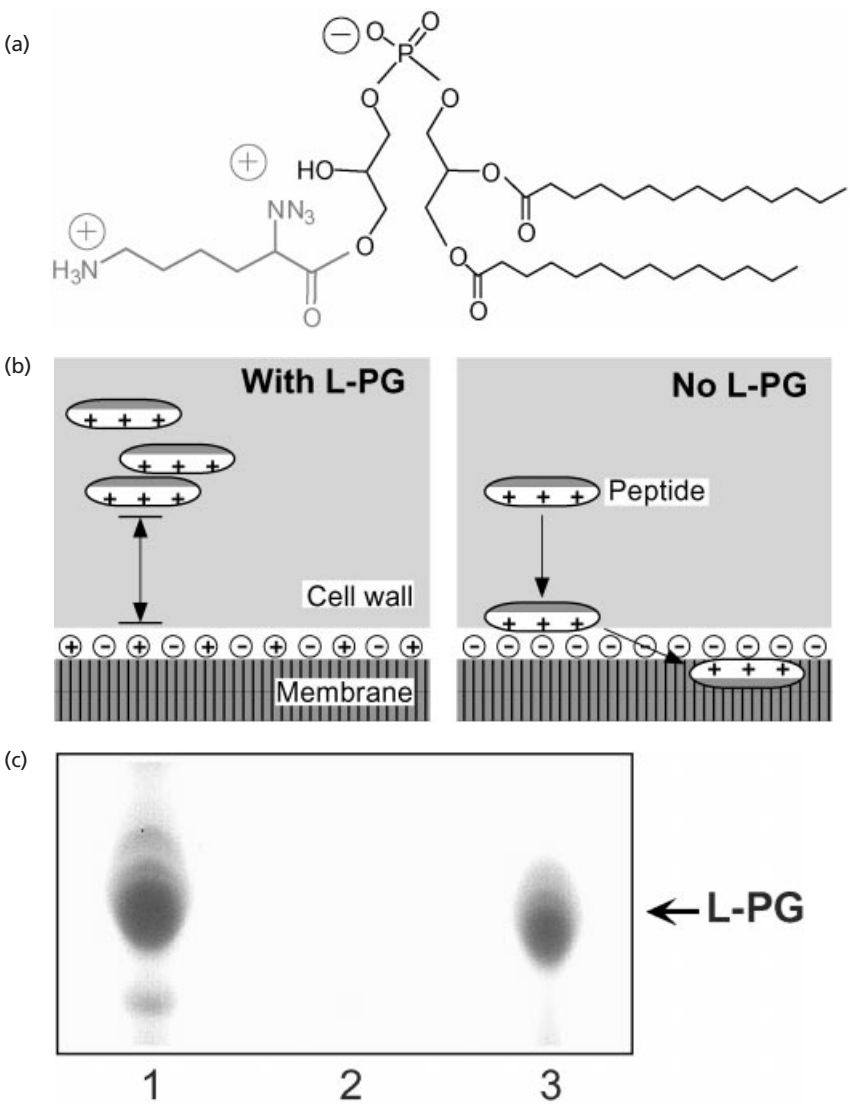

Fig. 1. (a) Structure of lysylphosphatidylglycerol (L-PG). The L-lysine is shown in grey. (b) L-PG-mediated resistance against CAMPs is based on electrostatic repulsion from the bacterial membrane. Modified with permission from (4). (c) Production of L-PG by *Bacillus subtilis* (lane 1) *Staphylococcus aureus* (lane 3) and an *S. aureus mprF* mutant (lane 2). Polar lipids were prepared, separated by TLC and stained with ninhydrin as described recently (5).