Ordered Cosmid Library of the *Mesorhizobium loti* MAFF303099 Genome for Systematic Gene Disruption and Complementation Analysis

Yoshiyuki Hattori 1, Hirofumi Omori 1, Masaki Hanyu 1, Noriko Kaseda 1, Elina Mishima 1, Takakazu Kaneko 2, Satoshi Tabata 2 and Kazuhiko Saeki 1,3

1 Department of Biology, Graduate School of Science, Osaka University, 560-0043 Japan
2 Kazusa DNA Research Institute, Kisarazu, 292-0812 Japan
3 Corresponding author: E-mail, ksaeki@bio.sci.osaka-u.ac.jp; Fax, +81-6-6850-5425.

For effective exploitation of the genome sequence information of *Lotus* microsymbiont, *Mesorhizobium loti* MAFF303099, to discover gene functions, we have constructed an ordered and mutually overlapping cosmid library using an IncP broad host-range vector. The library consisted of 480 clones to cover approximately 99.6% of the genome with average insert size and overlap of 26.9 and 11.1 kbp, respectively. The genome of *M. loti* consists of a single chromosome and two plasmids. The chromosome (7,036,071 bp) was covered 99.68% by 445 clones with four gaps, although two clones were unstable in *E. coli*. The larger plasmid pMLa (351,911 bp) was completely covered by 23 clones, while the smaller pMLb (208,315 bp) was covered 98.85% by 12 clones with two gaps. We have also made ancillary plasmids to facilitate the construction of deletion mutants using derivatives of the library clones. As a pilot experiment to uncover regions which contain novel symbiotic genes, 13 deletion mutants were constructed to lack in total 180.5 kbp of the genome. All the mutants formed apparently normal nodules and supported symbiosis, however, one mutant that lacked a 5.3 kbp chromosomal region, 4,551,930–4,557,222, did not produce normal exopolysaccharides as judged by fluorescence on medium containing Calcofluor. The results supported the effectiveness of the approach to detect gene functions.

**Keywords:** Exopolysaccharide — Genome library — *Mesorhizobium loti* — Nitrogen fixation — Symbiosis.

**Abbreviation:** ORF, open reading frame.

### Introduction

Bacteria of the family *Rhizobiaceae* and host legume plants establish symbiosis in that the microsymbionts convert atmospheric dinitrogen (N₂) to ammonia (NH₃) which is successively utilized by the plant symbiont. To establish of nitrogen-fixing symbiosis, both symbionts induce new developmental programs which are initiated by the primary signal exchange of flavonoids and nod factors (Denarie et al. 1996, Broughton et al. 2000, Perret et al. 2000). This exchange is probably followed by further and alternate signal exchanges leading to the organogenesis of nodules and differentiation of the microsymbionts to bacteroids (Denarie et al. 1996, Niner and Hirsch 1998), nonetheless, information is sketchy on these later steps despite their importance. Bacterial genes, such as nod, fix, nif and exo, are apparently responsible for the establishment of functional symbiosis, however, there is increasing evidence that other genes also influence the establishment and maintenance of nitrogen-fixing symbiosis (Niner and Hirsch 1998).

*Lotus japonicus* and *Mesorhizobium loti* represent one of the most promising model systems of legume–rhizobium symbiosis, especially as a model that forms determinant-type gIboIar nodules (Hayashi et al. 2000). The host legume *L. japonicus* is suited to both modern and classical genetics with its characteristics, such as small genome size, a short generation cycle, self-compatibility, genetic transformation and transposon tagging (Handberg and Stougaard 1992, Stiller et al. 1997). The significance of *L. japonicus* has been exemplified by the recent identification of plant genes such as the genes *NIN* and *SYMRK* that is responsible for the nodule organogenesis program (Schauer et al. 1999, Stracke et al. 2002). The usefulness of this plant is strengthened by accumulation of various mutants (Imaiumi-Anraku et al. 1997, Szczyglowski et al. 1998, Senoo et al. 2000, Wopereis et al. 2000, Kawaguchi et al. 2002) and by large-scale analyses of its cDNAs and genome (Asamizu et al. 2000, Sato et al. 2001, Nakamura et al. 2002). The microsymbiont *M. loti* has been less intensively studied compared to other major rhizobial species such as *Sinorhizobium meliloti* (Galibert et al. 2001) and *Bradyrhizobium japonicum* (Göttfert et al. 2001), and most of the strains studied, such as, NZP2037, NZP2037, NZP2035 and NZP2238, had been isolated from the nodule of *Lotus corniculatus* or *Lotus tenuis* (Jarvis et al. 1982). This situation has changed by the determination of the whole genome nucleotide sequence of the *M. loti* strain MAFF303099 (Kaneko et al. 2000a, Kaneko et al. 2000b). The *M. loti* nod, fix and nif genes are situated in a so called ‘symbiosis island’ of about 500 kbp long (Kaneko et al. 2000a, Kaneko et al. 2000b, Sullivan and Ronson 1998, Sullivan et al. 2002), but there may be other genes essential for
symbiosis but located outside the island such as the inv (invasion) deficient gene (Ward et al. 1989). Since MAFF303099 was originally isolated from nodules of L. japonicus (for review, Saeki and Kouchi 2000), this strain is the most suitable strain to uncover the still unidentified genes important for nitrogen-fixing symbiosis. However, the strain MAFF303099 has been even less studied than other M. loti strains and its molecular genetic resource is very limited except the genome sequence and gene annotations deduced from homology and/or similarity. In order to fully exploit the sequence data to discover novel gene functions, a well-organized DNA segment library of MAFF303099 genome may be quite useful in the same way as the Kohara and Perret libraries have been used, respectively, to study the whole genome of Escherichia coli and the plasmid genome of Rhizobium species NGR224 (Kohara et al. 1987, Perret et al. 1991). The expected library should consist of a complete set of mutually overlapping plasmid clones that covers the entire genome of M. loti. Each clone of the library should replicate both in E. coli and in M. loti, and be equipped with an electronic data set for its detailed characteristics to be utilized for functional analysis. Such a library would be a valuable genome resource not only for molecular genetic analysis of M. loti MAFF303099 but also for comparative studies using other M. loti strains and other organisms.

Here, we describe construction of an ordered cosm id library of M. loti MAFF303099 genome and database to meet the above requisites. We also describe the construction of plasmids to assist systematic genome deletion-insertion mutagenesis to search for novel symbiotic genes and report initial results of the mutagenesis study.

**Results**

**Construction of initial cosm id library**

The initial library construction was performed as a part of a whole-genome sequencing project. As a cosmid vector for construction of the ordered library, we made pKS800 which is a derivative of the vector pLAFR1 (Friedman et al. 1982) with its sole EcoRI site replaced with a multiloning site of NheI/ PacI/BamHI/Pmel/XhoI. This vector like its parent has an IncPα replication origin and is stably maintained as low copy plasmid in both E. coli and M. loti, though its relatively large size of more than 20 kbp limits the size of an insert to about 20–30 kbp using the λ phage packaging process.

Partial Sau3A1 digest of total DNA from M. loti MAFF303099 was size-fractionated and ligated to the BamHI site of pKS800 to make about 4,000 clones. E. coli strains harboring individual cosm id clones were maintained in 96-well microtiter plates: glycerol stocks totalling 41 plates were stored at –80°C. Although we found some empty vectors in the stock, more than 3,000 clones contained inserts of the expected 20–30 kbp. Therefore, the initial library has at least 10 equivalents of the approximately 7 Mbp genome of M. loti MAFF303099. Statistically, this collection should cover over 99% of the MAFF303099 genome.

**Ordering cosm id clones**

After DNA minipreparation of the clones, sequences of both ends of the insert were determined and fed to the Phred/Phrap program to extract candidates to cover the entire genome. About 370 clones were chosen as a start set for a mutually overlapping ordered library. The end sequences were also subjected to BLAST search against the whole genome sequence of M. loti MAFF303099 to determine positions, sizes and integrity of the inserts. Since the assignment of every open reading frame (ORF) in MAFF303099 had been done (Kaneko et al. 2000a, Kaneko et al. 2000b), attempts were made to select additional clones to cover every ORF in at least one clone. Similarly, efforts were made to select clones to encompass every potential operon in at least one cosm id clone by covering close and consecutive ORFs in the same strand. Selection was also carried out to make an overlap of adjacent clones be longer than 2.1 kbp, if possible, taking into account that the average gene density was one gene in every 1,042 bp in MAFF303099 (Kaneko et al. 2000a, Kaneko et al. 2000b). By these criteria, about 160 candidates were added.

Integrity and stability of each candidate clone was assessed by restriction analysis of DNA preparations from multiple independent colonies obtained after streaking E. coli glycerol stock on LB-agar, because usefulness of a genomic library might be impaired by rearrangement of cloned segments that can occur both in vivo and in vitro. In cases of in vivo rearrangement, agarose gel electrophoresis patterns of digested DNA preparation disagree with the calculated patterns based on the whole genome nucleotide sequence. Restriction analysis with at least three different enzymes indicated that less than 1% of the candidate clones gave different patterns, therefore in vivo segment rearrangement was a minor event in the collection. The in vivo rearrangement clones, if found, were omitted from the final library and substituted by other redrafted candidate(s). However, we could not omit two unstable clones that often became shortened in E. coli and showed unexpected restriction patterns. The clones encompassed a region containing minC, minD and minE homologs, which probably encode proteins required for proper cell division (Rothfield et al. 1999). Simultaneous detailed inspection of end sequences revealed that approximately 2% of the clones contained a very short Sau3A1 fragment of 10–200 bp together with expected long fragments of more than 20 kbp. The short inserts were probably by in vitro rearrangement derived from co-ligation of more than one insert DNA to the vector DNA. We attempted to omit all the in vitro rearrangement clones, however, six such clones remained in final library since we could not find appropriate substitutes.

**Properties of the ordered library**

The established ordered library consisted of 480 clones...
(A) Chromosome-1

Fig. 1
An ordered cosmid library of *Mesorhizobium loti*
Fig. 1

An ordered cosmid library of *Mesorhizobium loti*
An ordered cosmid library of *Mesorhizobium loti*
Chromosome-5

Fig. 1
An ordered cosmid library of \textit{Mesorhizobium loti}

Fig. 1
An ordered cosmid library of *Mesorhizobium loti* 1550 with an average insert size of 26.9 kbp and average mutual overlap of 11.1 kbp as summarized in Fig. 1 and Table 1. The library covered more than 99.6% of the total *M. loti* MAFF303099 genome. Information on each clone can be obtained from links in the web sites at http://www.kazusa.or.jp/rhizobase/, Rhizobase, and http://miya.bio.sci.osaka-u.ac.jp/.

The main chromosome (7,036,071 bp) was covered about 99.68% by 445 clones (number 1.1 to 346.1) with four gaps of about 6.35, 1.27, 8.87 and 6.21 kbp; each gap at the RhizoBase coordinates 693,017–699,370, 1,237,817–1,239,086, 2,999,182–3,008,054 and 6,796,231–6,802,438. There were very short overlaps of 4 and 73 bp at coordinates 115,232–115,235 and 3,411,941–3,412,013, respectively, and relatively short overlaps of 1,793, 1,967 and 1,976 at coordinates 3,784,698–3,786,490, 5,217,382–5,219,348 and 6,507,872–6,509,847, respectively. All the other overlaps were longer than 2.1 kbp, which is twice of average gene density in MAFF303099 (Kaneko et al. 2000a, Kaneko et al. 2000b). There were two unstable clones, numbers 81 and 81.1, that encompass the region at coordinate 1,582,000–1,610,000. These clones displayed a strong tendency to become shortened, probably by in vivo rearrangement in *E. coli*. In addition, six in vitro rearrangement clones remained, in which a small DNA fragment of 23–240 bp had co-ligated with a major DNA fragment of more than 25 kbp. The sequences of the co-ligated fragments in these clones will be also provided at the web site described above.

The “symbiosis island” at coordinate 4,644,792–5,255,766 was completely covered by 38 clones with no rearrangement clones. The larger plasmid pMLa (351,911 bp) was completely covered by 23 clones (number 1001.1 to 1017.1) with a minimum overlap of 3.76 kbp, whereas, the smaller pMLb (208,315 bp) was covered about 98.85% by 12 clones (number 2001 to 2013) with two gaps. The gaps for pMLb coverage were 977 and 1,419 bp at coordinate 82,871–83,483 and 197,316–198,734, respectively. There were relatively short overlaps for pMLb of 1,609, 653 and 1,722 at coordinates 17,551–19,159, 40,375–41,027 and 162,064–163,785, respectively, but other overlaps were longer than 4.29 kbp. There were no in vitro rearrangement clones for the two plasmids.

### Table 1 Summary of genome coverage by ordered cosmid library of *M. loti* MAFF303099

<table>
<thead>
<tr>
<th>Genome</th>
<th>Genome size (bp)</th>
<th>Number of clones</th>
<th>Average size (bp)</th>
<th>Average overlap (bp)</th>
<th>Coverage (%)</th>
<th>Number of gaps</th>
<th>Number of unstable clones</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>7,596,297</td>
<td>480</td>
<td>26,922</td>
<td>11,096</td>
<td>99.67</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>Chromosome</td>
<td>7,036,071</td>
<td>445</td>
<td>27,003</td>
<td>11,192</td>
<td>99.68</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>pMLa</td>
<td>351,911</td>
<td>23</td>
<td>26,520</td>
<td>11,038</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>pMLb</td>
<td>208,315</td>
<td>12</td>
<td>25,147</td>
<td>8,284</td>
<td>98.85</td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>
An ordered cosmid library of *Mesorhizobium loti* 1551

Use of the library to make deletion mutants

One of the main uses of the ordered library is to employ it for systematic detection of still unidentified genes important for symbiotic nitrogen fixation (a scheme depicted in Fig. 2). Presence of such gene(s) in a particular genome region may be quickly surveyed by construction of a series of deletion mutants that lack defined regions of the genome. A defined deletion locus can be made as a derivative, disruption plasmid, of an ordered clone by inserting a drug-resistant gene after digestion with an appropriate restriction enzyme. The deletion locus in the disruption plasmid is then exchanged with the genomic authentic locus by homologous recombination in the presence of incompatible “substitution” plasmid with incPβ replication origin, since the cosmid vector pKS800 has an IncPβ replication origin. We made two “substitution” plasmids pKS807 and pKS808 by inserting levan sucrase-kanamycin/neomycin phosphotransferase genes cassette, sacB-neo, and levan sucrase-spectinomycin adenylyltransferase genes cassette, sacB-aadA, into the plasmid R751 (Meyer and Shapiro 1980). Once a deletion mutant was obtained, either “substitution” plasmid could be excluded from the mutant by selecting sucrose resistance for their loss from the cells.

As a pilot experiment to make deletion mutants, we chose 10 ordered clones (Table 2). The clones were digested by one of the enzymes, *Bam*HI, *Eco*RI, *Hin*III, *Kpn*I, *Pst*I and *Sac*I, that do not cut pKS800, and inserted with the ΔGmR gene cassette which is a gentamicin acetyltransferase gene (*accC1*) surrounded by translation and transcription terminators at both ends (Ω). The resultant 13 plasmids with deletion-insertion loci (listed in Table 2) were transferred to *M. loti* MAFF303099 by conjugation via *E. coli* hosts. Then, either pKS807 (or pKS808) was further transferred into the resultant transconjugants to force recombination. Deletion candidates were selected on tryptone yeast extract medium supplemented with CaCl₂ (TY medium) by gentamicin resistance for retaining the deletion-insertion locus, by tetracyclin sensitivity for losing the cosmid vector portion of the deletion-insertion locus, by kanamycin (or spectinomycin) resistance for acquiring the “substitution” plasmid pKS807 (or pKS808) and by phosphomycin for *M. loti*. The deletion of the genome region was verified by Southern hybridization analysis (a portion of data is shown in Fig. 3B).
Table 2  Summary of *M. loti* genome deletion mutants

<table>
<thead>
<tr>
<th>Strain</th>
<th>Cosmid clone utilized</th>
<th>Enzyme utilized for deletion</th>
<th>Delected genomic region</th>
<th>Number of ORFs deleted</th>
<th>Symbiotic phenotype</th>
<th>Production of Calcofluor-binding exopolysaccharides</th>
<th>Growth on minimum medium</th>
</tr>
</thead>
<tbody>
<tr>
<td>DM001K</td>
<td>1</td>
<td>KpnI</td>
<td>32,255–53,598</td>
<td>25</td>
<td>Nod+, Fix+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>DM003P</td>
<td>3</td>
<td>PstI</td>
<td>79,480–95,371</td>
<td>18</td>
<td>Nod+, Fix+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>DM004B</td>
<td>4</td>
<td>BamHI</td>
<td>88,784–113,475</td>
<td>13</td>
<td>Nod+, Fix+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>DM006H</td>
<td>6</td>
<td>HindIII</td>
<td>123,031–135,561</td>
<td>16</td>
<td>Nod+, Fix+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>DM007P</td>
<td>7</td>
<td>PstI</td>
<td>136,769–155,221</td>
<td>15</td>
<td>Nod+, Fix+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>DM007E</td>
<td>7</td>
<td>EcoRI</td>
<td>137,587–154,870</td>
<td>14</td>
<td>Nod+, Fix+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>DM196B</td>
<td>196</td>
<td>BamHI</td>
<td>3,927,855–3,931,648</td>
<td>2</td>
<td>Nod+, Fix+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>DM223Bp</td>
<td>223</td>
<td>BamHI</td>
<td>4,551,930–4,557,222</td>
<td>7</td>
<td>Nod+, Fix+</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>DM229E</td>
<td>229</td>
<td>EcoRI</td>
<td>4,685,751–4,713,835</td>
<td>26</td>
<td>Nod+, Fix+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>DM229H</td>
<td>229</td>
<td>HindIII</td>
<td>4,689,502–4,696,262</td>
<td>7</td>
<td>Nod+, Fix+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>DM230B</td>
<td>230</td>
<td>BamHI</td>
<td>4,711,350–4,736,948</td>
<td>22</td>
<td>Nod+, Fix+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>DM230H</td>
<td>230</td>
<td>HindIII</td>
<td>4,711,542–4,732,603</td>
<td>18</td>
<td>Nod+, Fix+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>DM324K</td>
<td>324</td>
<td>KpnI</td>
<td>6,636,803–6,656,843</td>
<td>17</td>
<td>Nod+, Fix+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

*Symbiotic phenotype was accessed by the slant agar method. Nod+ indicates that the strain formed nodule on *L. japonicus* B-129 roots within 14 d after infection. In addition, the formed nodules were indistinguishable, by naked eye observation, from those induced by wild-type MAFF303099, Fix+ indicates the infected *L. japonicus* plants grew as fast as those infected by wild-type MAFF303099; the mutant grown on RDM agar medium, however, it did not fluoresce under the same conditions.

The obtained mutants were analyzed for the capacity to establish nitrogen-fixing symbiosis with *L. japonicus* B-129, to produce exopolysaccharides that bind to Calcofluor (Leigh et al. 1985) and to grow on rhizobium-defined medium (RDM) that contained glucose and ammonium as sole carbon and nitrogen sources, respectively (Ronson et al. 1987) (Table 2). The capacity to produce exopolysaccharides was examined, because loss of the capacity (Exo-) often affects symbiotic phenotype of rhizobia: for example, most Exo- mutants of *S. meliloti* are reported to form non-nitrogen-fixing ineffective nodules on alfalfa (Leigh et al. 1985) and Exo- mutants of *M. loti* strain NZP2037 are reported to form effective nodules on *Lotus pedunculatus* but not on *Leucaena leucocephala* (Hotter and Scott 1991). The capacity to grow on the minimal medium RDM was examined, because auxotrophic mutants of *S. meliloti* are frequently reported to show reduced symbiotic capacity such as low in vivo nitrogenase activity (Kerppola and Kahn 1988). All the mutants, except one mutant named DM223Bp, formed effective nodules to support the growth of plant hosts (Nod+ and Fix+ phenotype), fluoresced on TY-agar medium containing Calcofluor (Fluorescent Brightener 28) and grew on RDM as fast as the wild-type MAFF303099. The mutant DM223Bp like the other 12 mutants displayed Nod- and Fix- phenotype and grew on the RDM medium, however, it did not fluoresce on TY-agar medium containing Calcofluor indicating that the mutant has some defects in exopolysaccharide production (Fig. 3C). This mutant DM223Bp was derived from a mutant locus made by partial BamHI digestion of the number 223 cosmid and lacked a 5.3 kbp chromosomal region at coordinate 53,286–75,186 (Fig. 3A, B). The deleted region spanned seven ORFs and a tRNA gene for glycine (trnG-CCC) (Kaneko et al. 2000a, Kaneko et al. 2000b) (Table 3). The non-fluorescence phenotype on Calcofluor medium was fully complemented by the number 223 cosmid to the extent indistinguishable from wild-type MAFF303099 (Fig. 3), confirming that the deletion impaired at least one gene for normal production of exopolysaccharides.

**Discussion**

The availability of the complete genome sequence of *M. loti* strain MAFF303099 has significantly enhanced our understanding of *M. loti*. Comparative understanding is further strengthened by the recent determination of the sequences of the complete genome *S. meliloti* strain 1021 (Galibert et al. 2001), symbiotic gene regions of *M. loti* strain R7A (Sullivan et al. 2002), *Rhizobium* species NGR234 (Freiberg et al. 1997) and *B. japonicum* strain USDA110 (Göttfert et al. 2001). To fully exploit the wealth of sequence information, however, solid genetic material for *M. loti* MAFF303099 has been lim-
An ordered cosmid library of *Mesorhizobium loti*

Here, we have constructed a canonical ordered cosmid library that covers approximately 99.6% of the MAFF303099 genome using a broad host range vector replicable in *M. loti*. This defined library is intended primarily to be a reliable resource, by which almost any genome DNA would be easily accessible, and thus suitable starting material for systematic gene disruption and complementation study. The constructed library contained two unstable clones, number 81 and 81.1, which corresponded to the region containing *minC*, *minD* and *minE* homologs at coordinate 1,605,044–1,606,980. Each of the three *min* homologs is a sole ortholog in the *M. loti* MAFF303099 genome (Kaneko et al. 2000a, Kaneko et al.

---

**Fig. 3** Properties of the *M. loti* genome deletion mutant DM223Bp. (A) Physical map of wild-type MAFF303099, DM223Bp and the insert of cosmid 223. Restriction sites were shown by vertical bars; B, *BamHI*; EcoRI. Position of the tRNA gene *trnG-CCC* is shown by an open arrowhead. Arrows on the wild-type genome indicate ORFs with their Rhizobase numbering; open arrows represent those influenced by genome deletion in DM223Bp. The gentamicin-resistant gene cassette with translation and transcription terminator is shown as Gm. Restriction fragments affected by mutagenesis and visible in Southern analysis (B) are numbered and shown with nucleotide lengths. (B) Southern hybridization analysis of total DNA samples from wild type (W) and DM223Bp (M). Approximately 2 μg of DNA samples were digested by indicated enzymes and probed with the insert of cosmid 223. Numbered bands correspond to the restriction fragments shown in the map (A). (C) Production of Calcofluor-binding exopolysaccharides detected by fluorescence under long-wave UV light: 1, wild type; 2, DM223Bp; 3, DM223Bp harboring the cosmid 223; and 4, DM223Bp harboring the vector pKS800. 5 μl each of over night culture of each strain was spotted on TY-agar containing Calcofluor and grown for 5 d and observed.
Table 3  List of deleted ORFs and tRNA gene in the mutant DM223Bp

<table>
<thead>
<tr>
<th>ORF</th>
<th>Translational Start</th>
<th>End</th>
<th>Product size</th>
<th>Annotation</th>
</tr>
</thead>
<tbody>
<tr>
<td>mlr5660</td>
<td>4,552,174</td>
<td>4,551,287</td>
<td>296</td>
<td>LysR-family transcriptional regulator (transcriptional regulator)</td>
</tr>
<tr>
<td>mlr5661</td>
<td>4,552,752</td>
<td>4,553,420</td>
<td>223</td>
<td>conserved hypothetical protein, with similarity to NlpE1 (AAK11637) in R. etli (nodulation protein Noel)</td>
</tr>
<tr>
<td>mlr5662</td>
<td>4,554,168</td>
<td>4,553,845</td>
<td>108</td>
<td>unknown protein</td>
</tr>
<tr>
<td>mlr5664</td>
<td>4,554,314</td>
<td>4,554,526</td>
<td>71</td>
<td>unknown protein</td>
</tr>
<tr>
<td>mlr5666</td>
<td>4,554,967</td>
<td>4,555,653</td>
<td>229</td>
<td>unknown protein</td>
</tr>
<tr>
<td>mlr5668</td>
<td>4,555,882</td>
<td>4,556,274</td>
<td>131</td>
<td>ArsR-family transcriptional regulator (transcriptional regulator, similar to NolR)</td>
</tr>
<tr>
<td>trnG-CCC</td>
<td>4,556,899</td>
<td>4,556,829</td>
<td>628</td>
<td>transfer RNA for glycine</td>
</tr>
<tr>
<td>mlr5669</td>
<td>4,557,000</td>
<td>4,558,883</td>
<td>628</td>
<td>O-antigen acetylase</td>
</tr>
</tbody>
</table>

*a* ORF numberings from Rhizobase, http://www.kazusa.or.jp/rhizobase/

*b* Annotations are basically according to Rhizobase; however, re-annotation was performed to improve less significant or vague annotation. The improved re-annotation is shown with original annotation in parenthesis.

2000b). Therefore, it is possible that the homologs encode functional products to perturb the normal cell division process of *E. coli* host (Rothfield et al. 1999), though the precise reason for the instability remains to be elucidated. The library also had six gaps of, in total, 25.1 kbp and two very thin overlaps. Since ORFs located these uncloned regions do not seem to encode apparently toxic products for *E. coli*, we are still unable to distinguish if the gaps were caused by biological reason or by statistical reason that the initial number of library clones was not large enough. It remains to be done in future to fill these gaps and to enhance the short overlaps.

The cosmid vector pKS800 used in this study is a derivative of plasmid RK2 and replicates at relatively low copy number, 4 to 7, in *E. coli*; accordingly, the library clones showed rare in vivo rearrangement of cloned DNA fragments indicating its reliability. Because the vector can replicate in *M. loti*, the ordered library clones can be directly used for complementation analysis of mutants and for construction of disruption plasmid as depicted in Fig. 2. This feature, in combination with that our attempt to cover every potential operon in at least one ordered clone if possible, makes the ordered library more readily usable for functional analysis than the older ordered libraries such as the one covering the entire symbiotic plasmid of *Rhizobium* species NGR234 (Perret et al. 1991). In addition, this vector has recognition sites of Paci and PmeI on either side of the BamHI cloning site. These AT-rich 8-base cutters have only three and five recognition sites in the *M. loti* MAFF303099 genome, thus would be useful to map and manipulate a cloned insert, such as to make directed partial deletion. To facilitate gene-disruption experiments, we have also made “substitution” plasmids pKS807 and pKS808 to force homologous recombination by incompatibility between IncPα and IncPB replication origins in *M. loti*. Since the levan sucrase gene sacB in these plasmids confers counter-selectable marker of sucrose sensitivity, their exclusion is easily achieved from constructed mutants. This feature is not only applicable to *M. loti* strains but also to other Gram-negative bacterium for which no general transduction system has been established, unlike the case of *S. meliloti* 1021 for which such phage transduction systems are established (Finan et al. 1984, Martin and Long 1984).

As one of systematic disruption projects to find genes required for symbiotic nitrogen fixation, we proposed a scheme for rapid assessment of the presence of such a gene(s) by making mutants with relatively large deletions (Fig. 2). A drawback of the scheme is that no mutant would be obtained if a target region for deletion contains an essential gene. However, most of essential genes may be avoided by bioinformatic comparison with other well-studied system like *E. coli* (Blattner et al. 1997) and *Bacillus subtilis* (Kunst et al. 1997), and unavailable of a deletion mutant might be a clue to find either a novel essential gene or a combination of genes essential for viability. The results of the pilot deletion experiments indicated that the genes in the deleted 180.4 kbp genomic regions are not essential for growth on both complex TY medium and RDM minimum medium. It is probable that many of *M. loti* genes are not essential for growth on laboratory conditions, since the number of protein-coding ORFs in this organism is as much as 6,752. Our results on the mutant DM223Bp lacking the genome region 4,551,930–4,557,222 revealed that at least one gene impaired by the deletion is required for normal pro-
0.34. The product shows higher similarity (Expect = 2 e-04) with product of nlpe1 just upstream of lpeA that is required for antigenic maturation of lipopolysaccharide in Rhizobium etli (Duelli et al. 2001) and even higher similarity to hypothetical proteins in other non-symbiotic bacteria. Also, the mll5668 product may not be an ortholog of NoIR in M. loti, because another ORF, mll2806, shows much higher similarity (62% identity) to S. melloti NoIR. The mll5668 product may be classified as a member of ArsR-family transcriptional regulators which include several metalloregulatory proteins (Wu and Rosen 1991). It remains to be determined which particular gene is responsible for normal exopolysaccharide production in DM223Bp. Although it is revealed that both mll5661 and mll5688 are not essential for noduleulation and symbiotic nitrogen fixation with L. japonicus B-129, overall results of this study might exemplify the effectiveness of our approach to uncover gene functions in the M. loti genome.

Materials and Methods

Bacterial strains, media and growth conditions

E. coli strain XL1 Blue MR (Stratagene, La Jolla, CA, U.S.A.) was used for initial propagation of the cosmid library, while strain DH10B (Grant et al. 1990) was used for construction of derivatives. E. coli strain HB101 harboring plasmid pRK2013 (Ditta et al. 1980) was used for conjugation helper to transfer cosmid derivatives from E. coli hosts to M. loti strains. M. loti MAFF303099 was from Genetic Resource Center, National Institute of Agrobiological Sciences. This wild-type strain is naturally resistant to phosphomycin. E. coli strains were grown at 37°C in Luria broth or Terrific broth (Sambrook et al. 1989). M. loti strains were grown at 28°C in complex medium, TY, which contained 0.5% (w/v) tryptone, 0.3% (w/v) yeast extract and 0.1% (w/v) CaCl₂·2H₂O at pH 6.8, or the RDM described by Ronson et al. (1987). Antibiotics were used as follows for both E. coli and M. loti: tetracycline (Tc), 5 μg ml⁻¹; gentamicin (Gm), 20 μg ml⁻¹; kanamycin (Km), 50 μg ml⁻¹; trimethoprim (Tm), 100 μg ml⁻¹; and phosphomycin (Pm), 100 μg ml⁻¹.

DNA biochemistry and plasmid construction

DNA manipulations were according to standard procedures (Sambrook et al. 1989). Initial library construction and DNA sequencing were as described previously (Kaneko et al. 2000a). Southern hybridization analysis was carried out using cosmid insert DNA labeled with digoxigenin-UTP as described previously (Saecki et al. 1993).

Fingerprinting of cosmid DNA was performed by inverted field gel electrophoresis using PPI-200 Programmable Power Inverter (MJ Research Inc., Waltham, MA, U.S.A.), after digestion with either EcoRI, BamHI, Nhel, Xhol or HindIII restriction enzymes.

The cosmider vector pKS800 was constructed by inserting a DNA fragment from the second copy of pBR322, which included a 1.2 kbp aadA fragment that had been excised from plasmid pHP45sp (A. J. Prentki and Krisch 1984) with primers SacI-A and 5'-TTT GGA TCC AGT TGC AAT CCA-3' and SacI-B and 5'-TTG GAT CCT TGA CCT GAT-3'. The second crossover PCR was performed with the end primers, Bm-aadA-2 and Bm-aadA-1, and ligated to the solo BglII site of plasmid R751 to make pKS808.

Construction of genome deletion mutants

M. loti mutants were obtained by the procedures depicted in Fig. 2. First, a cosmid clone was digested with a restriction enzyme (when required followed by end-blunting with T4 phage DNA polymerase), then ligated with the 1.6 kbp Gm⁸ gene cassette that had been excised from plasmid pMG01 (Schweizer 1993) by BamHI or SmaI digestion. The resultant plasmid had the M. loti genomic region deleted and replaced with the Gm⁸ gene marker. Second, the deletion–insertion plasmid was transferred into M. loti MAFF303099 by tri-parental conjugation with a helper E. coli harboring plasmid pRK2013 (Ditta et al. 1980). Transconjugants were selected by Tc⁶, Gm⁸ and Pm⁶. Third, into the transconjugants, the “substitution” plasmid pKS807 (or pKS808) was further transferred consecutively from E. coli and selected for Gm⁸, Tm⁶ and Km⁶ (or Spc⁶). Then, the deletion–insertion copy was homogenized by repeated liquid culture in the presence of Gm⁸ and Km⁶ (or Spc⁶). The candidates for genome deletion mutants were selected for Tc⁶. The mutant was repeatedly cultured without Tm and Km (or Spc) and then was spread on a TY agar plate containing 5% (w/v) sucrose and Gm to obtain cells that had lost the “substitution” plasmid. Finally, construction of an expected mutant was confirmed by Southern hybridization with probes of Gm⁸ Cassette and the cosmider insert initially used.

Plant tests

Nodulation was observed by the slant agar method. Slant agar used was 20 ml B&D Nitrogen-free medium (Broughton and Dilworth 1971) with 0.9% (w/v) agar in 24×180 mm glass test tube. Seeds of L. japonicus B-129 were surface sterilized and germinated at 24°C in the dark. After 2–3 d, seedlings were transplanted to the slant agar, then infected with washed M. loti cells as described previously (Suzuki et al. 2001). The tube containing the hypocotyl of the plant sprout was covered partially with aluminum foil for shading the root. Subsequent cultivation was performed under controlled environment using a growth chamber (MLR-350H, Sanyo Ltd, Moriguchi, Japan) for 16 h/8 h day/night cycle at 24°C temperature with 70% humidity.

Detection of exopolysaccharide production

M. loti strains were grown on TY agar medium containing 0.02% (w/v) Calcofluor (Fluorescent brightener 28, Sigma) neutralized with NaOH and fluorescence was visualized under long-wave UV light. Cells were inoculated by either streaking or spotting 5 μl over night culture and grown for 4–6 d at 28°C.

Acknowledgment

This work was in part supported by grant from Program for Promotion of Basic Research Activities for Innovative Bioscience (PRoBRAIN) and from Kazusa DNA Research Institute Foundation. The

Downloaded from https://academic.oup.com/pcp/article-abstract/43/12/1542/1914911/Ordered-Cosmid-Library-of-the-Mesorhizobium-loti by guest on 16 September 2017
authors thank Dr. T. Uchiumi for communicating that M. loti strains are naturally resistant to phosphomycin.

References


An ordered cosmid library of *Mesorhizobium loti*


(Received August 22, 2002; October 2, 2002)