Cloning and structure of group C1 O antigen (rfb gene cluster) from *Salmonella enterica* serovar *montevideo*

SANG J. LEE, LAJWANT K. ROMANA and PETER R. REEVES*

Department of Microbiology, The University of Sydney, NSW 2006, Australia

(Received 28 May 1991; revised 18 October 1991; accepted 28 October 1991)

The *Salmonella enterica* group C1 O antigen structure has a Man-Man-Man-GlcNAc backbone with a glucose branch, which differs from the *S. enterica* group B O antigen structure which has a Man-Rha-Gal backbone with abequose as side-chain. We have cloned the group C1 rfb (O antigen) gene cluster from serovar *montevideo* strain M40, using a low-copy-number cosmid vector. The restriction map of the group C1 (M40) rfb gene cluster was compared with that of group B strain LT2 by Southern hybridization and restriction enzyme analysis. The results indicate that the flanking genes are very similar in the two strains, but there is no detectable similarity in the rfb regions. We localized the mannose pathway genes rfbM and rfbK and one of the genes, rfbK, shows considerably similarity to cpsG of strain LT2, suggesting that part of the mannose pathway in the group C1 rfb cluster is derived from a gene of the M antigen (cps) cluster. The M antigen, which forms a capsule, is comprised of four sugars, including fucose. The biosynthetic pathway of GDP-fucose has steps in common with the GDP-mannose pathway, and the rfbK gene cluster has isogenes of rfbK and rfbM, presumably as part of a fucose pathway. We discuss the structure and possible evolution of the group C1 rfb gene cluster.

Introduction

Lipopolysaccharide is a characteristic component of the outer membrane of Gram-negative bacteria. A typical lipopolysaccharide consists of three separate parts; an inner hydrophobic lipid A region, a central hydrophilic core oligosaccharide and an outer hydrophilic O antigen. The O antigen comprises repeats of a specific oligosaccharide (O unit) which carries antigenic specificity and in *Salmonella enterica* shows a high level of structural variation (Kauffman, 1966; Lüderitz *et al.*, 1968). The major genes specifying synthesis of the O antigen are located in the rfb gene cluster (Sanderson & Roth, 1989). Several of the well-characterized O units (groups A, B, D and E) have an identical Man-Rha-Gal backbone. They differ in the presence and nature of a dideoxyhexose side-chain sugar (Mäkelä & Stocker, 1984).

We have previously reported the cloning and sequencing of the rfb gene cluster from serovars (sv.) *typhimurium* (strain LT2) (Brahmbhatt *et al.*, 1988; Wyk & Reeves, 1989; Jiang *et al.*, 1991), *typhi* and *paratyphi* (Verma *et al.*, 1988; Verma & Reeves, 1989; Liu *et al.*, 1991), and *anatum* (Wang *et al.*, 1992), representative of groups B, D, A and E respectively. The genes for the rhamnose and mannose biosynthesis pathways and for the galactose transferase are very similar in all four strains. In groups A, B and D there are substantial differences in regions concerned with the dideoxyhexose side-chain sugar synthesis (Brahmbhatt *et al.*, 1988; Jiang *et al.*, 1991; Verma *et al.*, 1988; Verma & Reeves, 1989; Wyk & Reeves, 1989) and in group E1 the dideoxyhexose pathway region is replaced by a shorter segment (Wang *et al.*, 1992).

Not all *S. enterica* O antigens have the Man-Rha-Gal backbone and, for example, the group C1 O antigen consists of four mannose residues, one N-acetylglucosamine residue and a glucose side branch (Fuller & Staub, 1968; Lindberg *et al.*, 1988).

In this paper we report the cloning and analysis of the rfb gene cluster from group C1 *S. enterica* strain M40 (sv. *montevideo*) and compare it with the group B and other rfb gene clusters. Note that the strains of different *Salmonella* groups used in this study are not only members of the same species (*S. enterica*) but also of the same subspecies (Le Minor & Popoff, 1987; Brenner, 1978), and the variations among these rfb gene clusters should be seen in this light. The old species names are retained as serovar names.

* Author for correspondence. Tel. 2 692 2536; fax 2 692 4571.

0001-6951 © 1992 SGM
Methods

Bacterial strains and plasmids. The strains used are listed in Table 1. A series of plasmids carrying DNA spanning the entire rfb gene cluster of S. enterica LT2 (Brahmbhatt et al., 1988) (Fig. 1) was used in this study. All plasmids made in this study are shown in Table 1, or in Figs 1, 2 or 3, with details of construction given in the text where appropriate.

Media and materials. Bacteria were grown in nutrient broth containing (per litre of deionized water): yeast extract, 5 g; peptone, 10 g; NaCl, 5 g. Nutrient agar (NA) plates contained 1.2% (w/v) bacteriological agar in nutrient broth. LBK broth (Rosenberg et al., 1985) was used for growing strains for in vivo packaging. Media were supplemented with appropriate antibiotics [ampicillin (Amp) 25 pg ml−1; streptomycin (Str) 10 pg ml−1]. All enzymes and radiolabeled compounds used in this study were purchased as described previously (Brahmbhatt et al., 1988). Sodium gluconate was obtained from Sigma and S. enterica factor 07 antiserum was purchased from Difco.

DNA techniques. The methods used for DNA preparation, agarose gel electrophoresis, nick translation of DNA, in vitro DNA hybridization, colony hybridization, autoradiography, ligation, recovery of DNA from agarose gels and bacterial transformation were those described by Maniatis et al. (1982). The high and low stringency conditions used in hybridization experiments are described by Howley et al. (1979). EcoRI-cut SPPl DNA (Ratcliff et al., 1979), alone or together with uncut λ linearized DNA, was used as a molecular size standard.

Cosmid cloning. Sau3A-partially-digested chromosomal DNA from strain M40 was prepared by the method of Maniatis et al. (1982) and mixed with BamHI-digested, low-copy-number, cosmid vector pPR691 (Jiang et al., 1987) with a 3:1 ratio of insert:vector. The mixture was ligated and then packaged in vitro and transduced using the DNA packaging kit from Boehringer into strain LE392. The clones were selected after plating out on NA plates containing Kan, Spec and Str.

Serology. The O7 antigen epitope of serogroup C1 was detected by slide agglutination of cells as described by Leinonen (1985).

Gluconate dehydrogenase test. Sodium gluconate was added to minimal agar (Davis & Mingoli, 1950) as carbon source instead of glucose. A gnd-deletion strain RW231 was used as host (Nasoff & Wolf, 1980). Only clones containing a functional gnd gene will grow on these plates.

Chromosomal DNA hybridizations. For dot hybridization, chromosomal DNA was spotted onto a nitrocellulose filter as described by Beltz et al. (1983) and hybridized as described by Maniatis et al. (1982). For Southern hybridization the DNA was digested with EcoRI and PstI, electrophoresed and blotted, and then hybridized with nick-translated cosmid.

Transduction. Cosmids were transferred to S. enterica by transduction with phage λ, using strain SMR10 (Rosenberg et al., 1985), and the method of Jacobs et al. (1986) for in vivo packaging of the cosmid. pPR1042 was first transferred to strain SMR10 by transformation, and pPR756 by transduction using low-titre packaged cosmid obtained by phage λ infection of strain LE392 carrying pPR756. Induction of SMR10 gave much better yields of packaged cosmid than λ infection and this was necessary for subsequent transduction of S. enterica.

LPS analysis. LPS was extracted from 10 ml of overnight culture. Cells were pelleted and washed in 10 ml 30 mM-Tris/HCl, pH 8.1, then resuspended in 0.2 ml 20% (v/v) sucrose in 30 mM-Tris/HCl, pH 8.1. Following the addition of 20 μl of lysozyme (2 mg ml−1 in 0.1 M-EDTA, pH 7.3), the mixture was incubated for 30 min on ice before freezing at −70 °C. After 30 min, the mixture was thawed and 3 ml 3 mM-EDTA was added followed by sonication four times for 15 s each while sitting

Table 1. Bacterial strains, plasmids and phage

<table>
<thead>
<tr>
<th>Strain</th>
<th>Relevant characteristics</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Escherichia coli*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C600 (P334)</td>
<td>F− λ− thi-1 leuB6 lacY1 tonA21 supE44</td>
<td>Maniatis et al. (1982)</td>
</tr>
<tr>
<td>LE392 (P3743)</td>
<td>F− λ− hsdR514 supE44 supF38 lacY1 galK2 galT22 metB1 trpR55</td>
<td>Maniatis et al. (1982)</td>
</tr>
<tr>
<td>SMR10 (P3833)</td>
<td>E. coli C-1a: F− λ cos2 ΔB xis1 red3 gamam210 clt857 min5 Sam7)</td>
<td>Rosenberg et al. (1985)</td>
</tr>
<tr>
<td>Sd874, CGSC 5935† (P4052)</td>
<td>F− λ− lacZ2286 trp49 Δ(sbcB-rfb)86 upp12 relAI lacZ50</td>
<td>Neuhard &amp; Thomassen (1976)</td>
</tr>
<tr>
<td>Salmonella enterica</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LBS010 (P9120)</td>
<td>LT2 derivative: leu hsdL(r−r+m*) galE trpD2 rpsL120 ilc452 metE551 metA22 hsdA(r−r+m) hasB(r−r+m)</td>
<td>Bullas et al. (1983)</td>
</tr>
<tr>
<td></td>
<td>sv. montevido</td>
<td>IMVS†</td>
</tr>
<tr>
<td>M40</td>
<td>SH162 carrying pAM117</td>
<td>This study</td>
</tr>
<tr>
<td>P9236</td>
<td>SH162 carrying pAM117</td>
<td>P. H. Mäkelä, National Health Institute, Helsinki, Finland</td>
</tr>
<tr>
<td>SH1612</td>
<td>sv. montevido</td>
<td></td>
</tr>
<tr>
<td>Plasmids</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pPR691</td>
<td>Cosmid replicon, pSC101</td>
<td>Jiang et al. (1987)</td>
</tr>
<tr>
<td>pPR714</td>
<td>pPR691 carrying group B rfb gene cluster</td>
<td>Jiang et al. (1991)</td>
</tr>
<tr>
<td>pUC18</td>
<td>Replicon, pMB1</td>
<td>Vieira &amp; Messing (1982)</td>
</tr>
<tr>
<td>pAM117</td>
<td>pPR322 carrying E. coli lamB gene</td>
<td>Heine et al. (1988)</td>
</tr>
<tr>
<td>Phage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CS145 (P1150)§</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>F− λ*(lec857 S7) thi trpR</td>
<td></td>
</tr>
</tbody>
</table>

* All E. coli strains are K12 derivatives except SMR10.
† Obtained from Coli Genetics Stock Center, Dept. of Biology, Yale University, New Haven, Conn., U.S.A.
‡ Obtained from Institute of Medical and Veterinary Science, Adelaide, Australia.
§ Obtained from Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA.
on ice. Cellular debris was removed by centrifugation at 2.5 × 10^3 g for 15 min and the supernatant centrifuged at 2.6 × 10^4 g for 1 h. The resulting pellet was resuspended in 50 μl of sample buffer (Lugtenberg et al., 1975); 10–20 μl of this preparation was boiled for 5 min and 1 μl of proteinase K (25 μg ml^-1) added before loading onto SDS-PAGE (18% resolving gel and 8% stacking gel, containing 0.2% SDS). Silver staining was done according to Tsai & Frasch (1982).

**Enzyme assays.** An overnight culture (100 ml) of cells in nutrient broth was washed with 0.9% NaCl and resuspended in Tris/MgCl\(_2\)/EDTA (0.05/0.01/0.001 M, pH 7.5). The suspension was sonicated for 2.5 min at 10 kHz (Branson Sonifier B15) while sitting on ice and centrifuged at 2.7 × 10^4 g for 40 min at 4 °C. The supernatant was then used for enzyme analysis. Phosphomannomutase (rfbK) activity was assayed as described by Nikaido et al. (1966) and guanosine diphosphomannose (GDP-Man) pyrophosphorylase (rfbM) activity was measured by the method of Munch-Petersen (1962).

The Munch-Petersen GDP-Man pyrophosphorylase assay did not work well for strains carrying both rfbK and gnd, and an alternative assay was developed which depended on the presence of the rfbK gene. This assay was based on the phosphomannomutase assay. The reaction mix contained 50 μmol Tris/HC1 (pH 7.5), 5 μmol MgSO\(_4\), 1 μmol cysteine, 5 μmol NaF, 1 μmol NADP, 1 μmol glucose-6-phosphate dehydrogenase (ZWF), 0.2 μmol GDP-Man and 1 μmol pyrophosphate in a total volume of 1 ml. GDP-Man pyrophosphorylase converts GDP-Man to Man-1-P and the phosphomannomutase present in the extract then converts this to Man-6-P, and the reaction then continues as for the phosphomannomutase assay.

Both GDP-Man pyrophosphorylase and phosphomannomutase reactions were recorded at 340 nm at 25 °C for 0–30 min. Enzyme activities were calculated from the initial linear rates of cofactor reduction after subtraction of endogenous activities. Endogenous activities were measured in assays without α-D(+)mannose 1-phosphate or pyrophosphate for phosphomannomutase and GDP-Man pyrophosphorylase activities, respectively. The protein concentrations of extracts were estimated by the Lowry method. One unit of enzyme activity was defined as that which reduced 1 μmol NADP min^-1.

**Results**

**Cloning the rfb region of strain M40**

Plasmid pPR465 carries group B DNA from positions -2.16 to 5.34 of strain LT2 (Brahmbhatt et al., 1988) (Fig. 1), mostly upstream of the start of the rfb gene cluster at position 3.98. A preliminary experiment showed that pPR465 hybridized with chromosomal DNA from strain M40 as detected by dot hybridization (data not shown), and it was therefore used as a radioactive probe to screen by colony hybridization a cosmid bank obtained from strain M40. Of the 1000 colonies tested, 30 presumptive positives were selected for further analysis. Fourteen different EcoRI restriction patterns were obtained (data not shown), and a representative clone of each was tested with O7 antiserum by slide
agglutination. Five were positive and designated pPR755, pPR756, pPR757, pPR758 and pPR759 (Fig. 2). The five clones have related restriction patterns.

The gnd gene is known to be located between his and rfb in S. enterica strain LT2 (Sanderson & Hurley, 1987) downstream of rfb. The five clones were transferred into strain RW231 (Table 1) by transformation and the resulting strain tested positive for agglutination with 07 antiserum. The region from -7-41 to a presumptive rfb region of strain LT2 is shown on top. The other cosmids are aligned at the conserved EcoRI site at position 0-0 of strain LT2.

An E. coli lamB gene was introduced into S. enterica group C1 rfb deletion strain SH1612, to give strain P9236, which now has the $\lambda$ receptor and can be used for $\lambda$ transduction: this allowed pPR756 to be transferred into strain P9236 by transduction and the resulting strain tested positive for agglutination with 07 antiserum. Strain P9236 has a deletion of its rfb region and we conclude that pPR756 carries and expresses the complete rfb cluster of group C1.

To confirm that no rearrangement had occurred during cloning, cosmid pPR756 and M40 chromosomal DNA were digested with EcoRI and probed with pPR756 as described in Methods. All fragments observed were identical in cosmid and chromosomal DNA except the two which contained vector DNA in the cosmid (data not shown).

Cosmid clones expressing the group C1 O antigen of strain M40 were found to be unstable in the his-rfb delete group C1 strain P9236, although relatively stable in E. coli K12 including the his-rfb delete strain S$\phi$874. This is probably due to the toxic effects of a gene(s) within the cloned DNA.

Mapping the cosmid clones

Cosmid pPR756 was partially digested with BgII (it cuts the insert DNA and not the vector pPR691), religated, transferred into C600 and colonies selected on NA plates containing Kan, Spec and Str. Four different size cosmids, pPR941, pPR1039, pPR1040 and pPR1041, were obtained from the colonies, which had one or more of the four BgII fragments of the insert deleted (Fig. 1).

Plasmids pPR941, pPR1039, pPR1040, pPR1041 and pPR756 were digested with several restriction enzymes including EcoRI. This enabled an EcoRI map for each of the five representative cosmids to be aligned (Fig. 2). A cluster of four EcoRI sites is conserved between strains LT2 and M40 and as the right-hand site of this cluster was used as the zero position for LT2, this was also done for M40.

To complete the partial restriction map of the presumptive rfb region, the BgII fragment from positions -4-39 to 6-66, the EcoRI fragment from positions 0-00 to 10-39 and the SalI fragment from positions 10-13 to 17-76 of pPR756 were each cloned into pUC18. Double enzyme digests were then employed in appropriate combinations to confirm the order of smaller fragments. A detailed restriction map from -16-78 to 17-76 is shown in Fig. 1.

The region from -7-41 to a ClaI site at 3-98 is highly conserved, with all sites except for the ClaI site itself, (recognizable as GTGAT at positions 3974 to 3979 in the LT2 sequence; Jiang et al., 1991), being present in both strains. The region upstream of -7-41 was not studied in detail although small differences were evident, but downstream from 3-98, there was no similarity detectable for some distance. Thus the similarity between group B and C1 rfb region stops just before the start of first essential rfbB gene in group B (rfbB at position 4-1).

Comparison of the rfb region of strains M40 and LT2 by Southern hybridization

Cosmid pPR756 and strain M40 chromosomal DNA were probed in Southern blots using various segments of the rfb region of strain LT2 (Fig. 1). The inserts of pPR685 (positions -9-61 to -4-39), pPR570 (positions -4-39 to -2-16) and pPR465 (positions -2-16 to 5-34) are contiguous in strain LT2 and they probe three contiguous fragments in pPR756, and on the chromosome of strain M40, confirming the alignment upstream of rfb. The fragment that hybridized with pPR465 is, however, shorter, being only 6-14 kb in length. pPR685, in addition to probing the M40 eps region upstream of rfb, also hybridized with two regions on pPR756 (and the
Salmonella group C1 rfb gene cluster

M40 chromosome) from positions 2-93 to 3-98 and 10-13 to 11-81, and the significance of this is discussed below.

pPR341 and pPR301, which together cover most of the LT2 rfb gene cluster from positions 5-34 to 21-34, did not hybridize at all with M40 chromosomal DNA or pPR756 even at low stringency. pPR1105 (positions 21-34 to 23-03, downstream of rfb and including the gnd gene of LT2) hybridized with a fragment near the right-hand end of pPR756. The EcoRI–PstI (M40 positions 12-57 to 14.26) fragment was the same size as the corresponding LT2 fragment, indicating that the non-similarity ends at position 12-57 in strain M40, just before the gnd gene.

These results show that orf2.80 and gnd, which flank the rfb region in LT2, are both present in M40. orf2.80 ends at 3.72 and gnd starts at 13.15, leaving 9.45 kb of DNA between them which has no evident similarity to the rfb region of LT2.

rfb function

To define the minimum region required for rfb function of M40, a series of deletion plasmids was made: pPR756 was partially digested with Sau3A, ligated and transformed into strain LE392. Transformants were selected on NA plates containing Kan, Spec and Str. pPR1042 (Fig. 3), which had DNA from approximate position −15.0 to a PvuII site at position 2-13 deleted, was transferred into S. enterica rfb deletion strain P9236 by transduction and into E. coli rfb deletion strain Sφ874 by transformation: the resulting strains agglutinated with O7 antiserum. pPR1042 was partially digested with PstI, religated and transformed into strain Sφ874. The transformants (pPR1273, pPR1274, pPR1275 and pPR1276) thus obtained had one or more of the PstI fragments deleted (Fig. 3). All failed to agglutinate with O7 antiserum. These experiments show that the complete M40 rfb region is present in plasmid pPR1042, and that all or part of the DNA on each of three PstI fragments covering positions 2-97 to 14-26, a total of 11-29 kb, is essential for rfb function. This region (from positions 2-97 to 14-26), includes the region which replaces the rfb cluster of strain LT2.

These results were confirmed by SDS-PAGE and silver staining of LPS. LPS was extracted from strain
P9236 containing pPR756 or pPR1042 and from strain Sφ874 containing pPR756, pPR1042, pPR1273, pPR1274, pPR1275 or pPR1276, and analysed by SDS-PAGE and silver staining (Fig. 4). Strains carrying pPR756 or pPR1042 gave a ladder banding pattern typical of LPS, while those carrying pPR1273, pPR1274, pPR1275 and pPR1276 failed to show any LPS on SDS-PAGE.

During the cloning of the M40 rfb gene cluster, it was observed that E. coli K12 strain LE392 carrying functional rfb cosmid clones was resistant to phage λ infection, although the host strain was sensitive. However, the addition of 5 mM-Ca²⁺ to the growth media overcame this resistance. It has been suggested (van Alphen et al., 1978) that Ca²⁺ causes dramatic rearrangements of the lipopolysaccharide and protein of the outer membrane of Gram-negative bacteria, and it seems likely that LPS hinders λ attachment by steric hindrance and that this is overcome in the presence of Ca²⁺.

**Localizing mannose genes in the group C1 rfb region**

We assayed extracts of Sφ874 carrying plasmids with all or part of the rfb region for GDP-Man pyrophosphorylase (rfbM) and phosphomannomutase (rfbK) activity (Fig. 3). pPR1042 confers both functions as expected: pPR1273 lacks both, but the slightly longer clone, pPR943, has rfbM function. Therefore the rfbM gene is located between positions 8-86 and 10-39. pPR1275 and pPR1276 have rfbK function, but pPR1274 does not and the rfbK gene is thus located between positions 10-13 and 11-81.

**Discussion**

We have cloned the rfb gene cluster of S. enterica strain M40 (group C1), using a low-copy-number cosmid vector pPR691. The resulting cosmid pPR756 confers on E. coli K12, including an rfb delete strain Sφ874, and also on a group C1 S. enterica rfb delete strain P9236, the ability to synthesize O7 antigen, indicating that all the genes necessary for group C1 O antigen synthesis and assembly are present on this cosmid clone.

The rfb gene cluster of strain LT2 is 17.2 kb in length (from position 4-1 to position 21-34 of Fig. 2). Southern hybridization and restriction enzyme analysis of the rfb region of strain M40 shows that the flanking genes (orf2-80 and gnd) are the same in both strains but in place of the 17.24 kb rfb cluster of strain LT2, we have 9-45 kb containing the rfb region of strain M40, with no apparent similarity to the rfb cluster of strain LT2.

Strain LT2 has two isogenes for mannose biosynthesis as part of the cps gene cluster (Stevenson et al., 1991); these genes, cpsB and cpsG, are present on pPR685 and Southern hybridization showed that this plasmid hybridized with the rfb gene cluster of strain M40, in addition to showing the presence of a cps region in strain M40 about 7 kb upstream of rfb. pPR685 gives a good hybridization signal with fragments carrying DNA to the right of the SalI site at position 10-13, but hybridizes only poorly with fragments carrying DNA from that site left-ward and does not hybridize at all with fragments carrying DNA left of the ClaI site at position 9-50 up to position 3-98 (data not shown), which locates the part of the rfb cluster with similarity to the cps to the region immediately upstream of position 10-13.

The two M40 rfb genes for mannose biosynthesis, rfbM and rfbK, were localized between approximate positions 8-86 to 10-39 and 10-13 to 11-81 respectively by enzyme assay. It appears that rfbK lies in the region of similarity to the cps region. Examination of the restriction map shows that the maps of cpsG (LT2) and rfbK (M40) can be aligned, there is a conserved SalI site (position 10-13) and, assuming the same gene length, rfbK (M40) runs from positions 9-93 to 11-28. We conclude that cpsG (LT2) and rfbK (M40) are very similar but this does not apply to cpsB (LT2) and rfbM (M40). The region from positions 2-93 to 3-98, just upstream of rfb, also hybridized with pPR685 but did not give any detectable GDP-Man pyrophosphorylase or phosphomannomutase function (Fig. 3) and we cannot speculate as to what this region represents.

The rfb region of strain LT2 is of low G+C content flanked by DNA of typical S. enterica G+C content (Jiang et al., 1991). Although we detected no similarity between the two rfb regions either by restriction mapping or by Southern hybridization, the high number (11) of DraI sites (TTTTAAA) and the low number (1) of SalI sites (GGGCCC) in pPR1042 (data not shown) are comparable to the frequencies in the 0.32 G+C region of strain LT2, and hence the C1 rfb region is probably also generally of low G+C content. However, the mannose genes of the cps gene cluster of strain LT2 have a G+C content of 61% (Stevenson et al., 1991), and thus the rfbK region of M40, which hybridized at high stringency with pPR685, must also be of high G+C content. For the reasons given earlier for groups B (Jiang et al., 1991), D (Verma & Reeves, 1989; Liu et al., 1991), C2 (Brown et al., 1991) and E1 (Wang et al., 1992), one can argue that the evolution of most of the group C1 rfb region occurred within a low G+C content species prior to transfer to S. enterica, but the rfbK gene at least appears to have been derived from the Salmonella cps cluster.

In this study we show that the rfb gene cluster of group C1 shows no similarity by Southern blotting to that of group B. In our previous work on groups A, B, D and E1, we found that genes for comon pathways were in general
very similar or identical. In particular, the rfbK and rfbM genes of the mannose pathway are extremely similar in all four groups. The group C1 mannose pathway genes are not only not closely related to those of the four other groups studied, but the two genes involved seem to have different origins, with one of them being very similar to a group C1 mannose pathway genes of the mannose pathway are extremely similar in all four groups. The group C1 mannose pathway genes are not only not closely related to those of the four other groups studied, but the two genes involved seem to have different origins, with one of them being very similar to a gene in the cps gene cluster.

There are about sixty forms of O antigen recorded for S. enterica (Ewing, 1986). The first four which we studied had rfb gene clusters which appeared to have been acquired by S. enterica by lateral transfer, and to have been assembled from more than one source. The rfb gene clusters of the first four groups studied were clearly related and once one had transferred, the others could readily transfer to S. enterica by homologous recombination in sequences common to all four (Verma & Reeves, 1989; Wang et al., 1992). The group C1 rfb gene cluster, although not related in structure, also appears to have been acquired by lateral transfer, and to have been assembled from more than one source. This finding emphasizes the role of lateral transfer in generating the great diversity of O antigen structure. It is of interest that the group C1 rfb cluster, although not closely related to the other four studied, also appears to have been transferred from a low G+C content species, probably from outside the family Enterobacteriaceae (Jiang et al., 1991), but as yet we have no indication beyond G+C content of the possible origin of the major part of the group C1 rfb gene cluster.

However, in the case of the group C1 rfb gene cluster it appears that after transfer one of the genes (rfbK) was added by duplication of the cpsG gene and translocation of one copy to the rfb cluster, and thus for the first time the source of one of the genes can be identified.

We have also studied the rfb region of a group C2 strain (Brown et al., 1991) and find that although similar in the rhamnose pathway genes to the groups A, B, D and E1, it has no similarity by Southern blotting in the mannose pathway region. Further, Southern blotting shows no similarity between the mannose pathway region of M40 and those of group C2 strains (unpublished data).

We are currently sequencing the rfb region of group C1 strain M40, and that of a group C2 strain, to further define the evolutionary events that took place in their assembly.

This work was supported by a program grant from the Australian Research Council. One of us (S. J. Lee) is grateful to the Department of Microbiology, The University of Sydney for financial assistance. We would like to thank P. H. Mäkelä (National Health Institute, Helsinki, Finland), C. Murray (IMVS, Adelaide), R. E. Wolf Jr, S. M. Rosenberg, J. Neuhard and E. Thomassen for strains, and T. Ferenci for providing us with the lamB plasmid pAM117.

References


Liu, D., Verma, N. K., Romana, L. K. & Reeves, P. R. (1991). Completing the sequence of the rfb region of Salmonella serovars typhi (group D) and paratyphi (group A), and relationships to group B. Journal of Bacteriology 173, 4814–4819.


