The WbaK acetyltransferase of *Salmonella enterica* group E gives insights into O antigen evolution

Yaoqin Hong,1 Katarzyna A. Duda,2 Monica M. Cunneen,1 Otto Holst2 and Peter R. Reeves1

1School of Molecular Bioscience, University of Sydney, New South Wales 2006, Australia
2Division of Structural Biochemistry, Research Center Borstel, Leibniz-Center for Medicine and Biosciences, Airway Research Center North (ARCN), German Center for Lung Research (DZL), D-23845 Borstel, Germany

O antigens are polysaccharides consisting of repeat units of three to eight sugars, generally assembled by genes in a discrete O antigen gene cluster. *Salmonella enterica* produces 46 forms of O antigen, and most of the variation is determined by genes in the gene cluster. However, in some cases the structures are modified by enzymes encoded outside of the gene cluster, and several such modifications have been reported for *Salmonella enterica* group E, some with the genes on bacteriophages and one gene at a distant chromosomal site. We identified the enzyme, WbaK, that is responsible for O-acetylating the subgroup E1 O antigen, and found that the gene is located just downstream of the gene cluster as currently known. The *wbaK* gene appears to have been imported by a recombination event that also replaced the last 37 bp of the *wbaP* gene, indicating that homologous recombination was involved. Some of the group E strains we studied must have the original gene cluster, as they lack *wbaK* and the sequence downstream of *wbaP* is very similar to that in several other *S. enterica* O antigen gene clusters. In effect the gene cluster was extended by one gene in subgroup E1. It appears that a function that is usually encoded by a gene outside of the gene cluster has been added to the gene cluster, in this case giving an example of how such gene clusters can evolve.

INTRODUCTION

Most Gram-negative bacteria have O antigen as part of their cell surface lipopolysaccharide (LPS). Of these, the O antigen is in most cases heteropolymeric, i.e. composed of repeats of three to eight sugar residues called O units. However, there are also homopolymeric O antigens present in some forms of LPS. The heteropolymeric structures are enormously diverse, with over 186 forms in *Escherichia coli* (including *Shigella* strains) and 46 forms in *Salmonella enterica* (Feng et al., 2004; Guibourdenche et al., 2010).

All *S. enterica* O antigens, except for O54, are synthesized by the Wzx/Wzy-dependent pathway (Valvano, 2003). The genes associated with the synthesis are generally grouped as an O antigen gene cluster, which in *S. enterica* is flanked by the *galF* and *gnd* genes. The synthesis begins with O unit assembly on a lipid carrier, undecaprenyl pyrophosphate, which is embedded in the cytoplasmic face of the inner membrane. These assembled O units are translocated to the periplasmic face by the Wzx translocase, and polymerized by the Wzy polymerase, with the modal length determined by the Wzz co-polymerase (Valvano, 2003). Polymeric O antigens of varied chain length are transferred from undecaprenyl pyrophosphate to lipid A/core by the WaaL ligase (Valvano, 2003), and the complete LPS molecules are exported to the outer membrane (Silhavy et al., 2010).

Most *S. enterica* O antigens that are synthesized by the Wzx/Wzy-dependent pathway utilize either N-acetylgalcosamine or N-acetyl-galactosamine as their initial O unit sugar (Valvano, 2003); the exceptions are groups A, B, C2–C3, D1, D2, D3 and E, which use galactose (Gal). These groups, except for C2–C3, have the same O unit main chain, as illustrated for groups D2 and E in Fig. 1. Correspondingly, these groups have rather similar gene clusters (Fig. 1).

Group E differs from other Gal-initiated groups in the lack of a side branch, such as tyvelose as seen in group D2 (Fig. 1). It was at one time subdivided into groups E1, E2, E3 and E4, but these can now be seen as subgroups of group E, as they have the same genes in the gene cluster, and the
same main chain in the O unit. The E1 O unit has antigen factors 3 and 10. Factor 3 is reported to be due to the β-Man-(1→4)-Rha linkage in the O antigen main chain (Uchida et al., 1963), and this is also found in both E2 and E3 (Robbins & Uchida, 1962). Factor 10 corresponds to an O-acetyl group on carbon 6 of the Gal residue (Uchida et al., 1963). ε15 is the phage that converts subgroup E1 cells to subgroup E2. This phage codes for three O antigen related proteins, one that suppresses expression of the O-acetyltransferase, another for inhibiting the E1 ε-(1→6) O antigen polymerase, and also an alternative polymerase that gives a β-linkage (Kropinski et al., 2007). This gives the antigenic formula O3.15, with factor 15 due to presence of the β-(1→6) polymerization linkage. Another phage, ε34, only colonizes E2 strains, converting them to E3 with antigenic formula O3.15,34, O34 being due to glucosylation of C-4 of the Gal residues, i.e. E3 carries both ε15 and ε34 prophages (Grimont & Weill, 2007). The E4 O antigen has a Glc residue on C-6 of the Gal residues that is due to an oafC gene that maps close to purE (Mäkelä & Stocker, 1984).

The D2 and E O antigens differ only by the presence of a side-branch tyvelose residue covalently attached to the Man residue in D2 (Hellerqvist et al., 1970). Note that the D2 gene cluster is proposed to have arisen by a recombination event between groups D1 and E (Xiang et al., 1994), and part of the gene cluster is very similar to that of group E (Fig. 1). The D2 and E gene clusters differ only in the presence of tyvelose synthesis genes and a tyvelose glycosyltransferase gene in D2, and having different forms of the wex gene. An additional gene, wbaK, previously known as orf17.4, is found 170 bp downstream of the wbaP gene of the published group E1, E2 and E3 sequences (X60666.1, KC688885.1, FR775224.1, NZ_AFCV01000775.1 and NZ_AFCM01000803.1), which is not found elsewhere in public sequence databases.

In silico analysis predicts that WbaK is an acetyltransferase (CAA43082.1, CYB96288.1, ZP_12181337.1 and ZP_12133661.1). Most acetyltransferases for bacterial surface polysaccharide are membrane-associated proteins, of which the majority are members of the Acyl_transf_3 Pfam family (Punta et al., 2012). These include OafA (S. enterica LT2), Oac (Shigella flexneri prophage Sf6) and WcJE (Streptococcus pneumoniae serotype 9V). However, in some cases these acetyltransferase genes are found to encode soluble proteins and are members of the Hexapep family, for example, wbbJ and wbaL from E. coli K-12 O16 and S. enterica group C2 respectively. WbaK is a typical Acyl_transf_3 protein that has nine putative transmembrane segments, and has a conserved protein domain (COG3594) commonly associated with either acyl- or acetyltransferase function at the N-terminal end of the protein. Here, we show that WbaK is an acetyltransferase of the Acyl_transf_3 family responsible for the O-acetyl group on the Gal residues of the group E1 O antigen. We also present evidence indicating that the wbaK gene was introduced by recombination, and propose that the E4 strains have the ancestral form of the Salmonella group E gene cluster.

**METHODS**

**Bacterial strains, media and growth conditions.** The bacterial strains used in this study are listed in Table I. Bacteria for genetic and phenotypic analysis were grown at 37°C in nutrient broth. Where appropriate, growth media were supplemented with 25 μg ml⁻¹ ampicillin or kanamycin.

**Targeted gene replacements.** The wbaK gene was replaced by a kan gene as previously described (Datsenko & Wanner, 2000). Briefly, cells carrying pKD46 were grown at 30°C until an OD₆₀₀ of approximately 0.4 was reached, and the lambda recombinase genes were induced with 1-aminobase (10 mM) for 40 min. Cells were washed with ice-cold 10% glycerol and used for electrottransformation. Successful transformants were selected using kanamycin, and verified by sequencing (Australian Genome Research Facility). See Table S1 (available in Microbiology Online) for oligonucleotides used in gene replacements.

**LPS preparation for SDS-PAGE and visualization by silver staining and immunoblot.** LPS was prepared as described previously (McGrath & Osborn, 1991) but with modifications suggested in a recent study (Hong et al., 2012). Samples were loaded on 13% acrylamide gel for Tricine SDS-PAGE (Brown et al., 1991). LPS either was visualized by silver staining as described (Brown et al., 1991) or was transferred to a nitrocellulose membrane for immunoblot analysis (Sambrook et al., 1989). Membranes for immunoblotting were first
reacted with anti-Salmonella O10 antibody (Difco). Anti-rabbit IgG conjugated with horseradish peroxidase (Pierce) was used for the secondary reaction. The immunoblots were visualized with 0.3 % 4-chloro-1-naphthol, 0.073 % H2O2, and 10 % methanol in 1× Tris buffered saline with Tween 20 (Sambrook et al., 1989).

LPS extraction and purification for structural analysis. For structural analysis, LPS was extracted from delipidated, dry bacterial masses of strains M31 and M2888 utilizing the hot phenol/water procedure (Westphal & Jann, 1965). The crude LPS were further purified by enzymic treatment with DNase and RNase (37 °C, 2 h) and Proteinase K (56 °C, 1 h). The yields were: M31, 39.6 mg, 10.7 % of dry bacterial mass; M2888, 57.7 mg, 11.6%. Purified LPS was hydrolysed with 1% acetic acid for 1 h at 100 °C, then ultracentrifuged (105 000 g, 1 h, 4 °C) in order to remove lipid A. The crude O antigens were further fractionated on a column (45 × 2.5 cm) of Sephadex G-50 in 0.05 M pyridinium acetate buffer, pH 4.5, yielding amounts of 3.3 (M31, 0.9 % of bacterial dry mass) and 0.6 (M2888, 0.12 %) mg.

NMR spectroscopy. NMR spectroscopy experiments were carried out after H–2H exchange on samples in 99.9 % 2H2O. 1D (1H, 13C) and 2D homonuclear (1H,1H) correlation spectroscopy (COSY), total correlation spectroscopy (TOCSY), and rotating-frame Overhauser effect spectroscopy (ROESY), as well as heteronuclear (1H,13C) single quantum correlation–distortionless enhancement by polarization transfer (HSQC-DEPT) and heteronuclear multiple bond correlation (HMBC) experiments were recorded at 27 °C with a Bruker DRX Avance 700 MHz spectrometer (operating frequencies 700.75 MHz for 1H NMR, 176.2 MHz for 13C NMR), equipped with a 5 mm CPQCI multinuclear-inverse cryo-probehead with a z gradient, and applying standard Bruker software. COSY, TOCSY, and ROESY experiments were recorded using datasets (t1×t2) of 4096 × 512 points, COSY with 4 scans, and TOCSY, ROESY with 8 scans. The TOCSY experiments were carried out in the phase-sensitive mode with mixing times of 60 ms, and the ROESY experiments, of 300 ms. HMBC spectra were adjusted to long range proton carbon coupling constant of 10 Hz. Chemical shifts were reported relative to an internal standard of acetone (δH 2.225, δC 31.45).

RESULTS

The WbaK protein modifies group E O antigen

We had difficulty in the genetic manipulation of the S. enterica group E1 serovar (sv.) Anatum strain (M32) which was used in previous studies on the group E gene cluster (McConnell et al., 2001; Wang et al., 1992), and instead used an E1 sv. Amsterdam strain (M31). The wbaK gene was replaced by a kan cassette to give a ΔwbaK mutant, M2888.

LPS was extracted and run on an SDS-PAGE gel. Both strains, M31 (Fig. 2a, lanes 1 and 3) and M2888 (lanes 2 and 4), have a typical LPS pattern with lipid A-core associated with a variable numbers of O unit repeats, and shows the characteristic modal length regulation. However, for the ΔwbaK mutant, M2888, we observed a change in the pattern of the O antigen ladder, as the bands for LPS with O antigen (lanes 2 and 4) run consistently faster than the corresponding bands of the wild-type (lanes 1 and 3), with the core+11 band of the mutant running at the same rate as core+10 of the wild-type. The LPS modal repeat region also migrated faster for the mutant. We showed that this observed difference was not due to any difference in

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype description</th>
<th>Source/reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>M31</td>
<td>S. enterica group E1 serovar Amsterdam</td>
<td>Institute of Medical and Veterinary Science, Australia</td>
</tr>
<tr>
<td>M32</td>
<td>S. enterica group E1 serovar Anatum</td>
<td>Wang et al. (1992)</td>
</tr>
<tr>
<td>M2888</td>
<td>ΔwbaK::kan</td>
<td>This study</td>
</tr>
<tr>
<td>M2900</td>
<td>S. enterica group D2 sv. Strasbourg</td>
<td>Hong et al. (2012)</td>
</tr>
</tbody>
</table>

Table 1. Bacterial strains used in this study

Fig. 2. LPS expression in S. enterica group E strain M31 and ΔwbaK mutants. Samples were separated by 13 % Tricine-SDS-PAGE and stained by silver nitrate. The asterisk (*) indicates the position of LPS with 10 O unit repeats. The modal repeat regions for each LPS are indicated by boxes. (a) The effect of wbaK mutation on LPS mobility. (b) Mixing of LPS from M31 and the ΔwbaK mutant gave a distorted ladder pattern (the region is indicated within the arrows). Note that the brightness was adjusted in (b) using Neo-image (http://www.neomaging.cn/en) to give a more even contrast for low molecular mass and high molecular mass bands. NA, not applicable.
running conditions in different lanes, by running a mixed sample (Fig. 2b). Individual LPS from the wild-type and the \( \Delta wbaK \) mutant gave clear banding patterns up to about core +12 (Fig. 2b, lanes 1 and 3). However the mixed sample, containing half of the loadings for each respective LPS, lost the banding pattern for LPS with 6–8 O units, but the bands for core +1 to core +3, and for core +9 to core +12, were well resolved. Also, bands for core +9 to 12 O units for the wild-type were aligned with the bands with core +9 to 12 O units in the mutant. This difference in mobility suggests that the LPS of \( \Delta wbaK \) mutant (M2888) is missing a structural component present in the wild-type. As this modification did not impair O antigen polymerization, it is likely that WbaK modifies the group E O unit backbone, with loss of O-acetylation of the Gal residues being the obvious candidate.

## O-Acetylation of Gal is mediated by WbaK

We undertook 1D and 2D NMR experiments and performed complete assignment of \(^1\text{H}\) and \(^{13}\text{C}\) resonances (Table 2) of the O antigen isolated from wild-type (M31) LPS by combining the information obtained from COSY, TOCSY, ROESY, as well as HSQC-DEPT and HMBC experiments. The obtained structure was identical to that from \( S.\ entera\) group E1 sv. Anatum (L’vov et al., 1989). By comparing \(^1\text{H}\) NMR profiles (Fig. 3), we were able to detect an O-acetylated Gal species in the wild-type O antigen that was absent in the mutant. In detail, the \(^1\text{H}\) NMR spectrum of the O antigen of \( \Delta wbaK \) lacked the high field signal at 2.12 p.p.m. originating from the CH3 group of O-acetyl, as well as three low field shifted signals due to proton deshielding at 4.17 p.p.m., 4.23 p.p.m. and 4.31 p.p.m. (H5, H6a, H6b of the Gal\(^6\)OAc, respectively). This confirms that WbaK is the transferase responsible for O-acetylation of Gal residues in the group E O antigen.

### O10 specific sera do not bind de-O-acetylated LPS

To investigate the role of the O-acetyl group in serology, we performed an immunoblot analysis with anti-O10 antibody using LPS extracted from the group E1 wild-type strains (M32 and M31) and the M31 \( \Delta wbaK \) mutant (M2888). We also included LPS from M2900 which is a D2 strain that produces group E-like O antigen (Hong et al., 2012). It is known that the D2 O antigen is not O-acetylated, so the O antigen produced by strain M2900 should be identical to that of the \( \Delta wbaK \) mutant.

Note that the same sample preparations and loadings were used for Fig. 4(a) and 4(b). Although the loading for \( \Delta wbaK \) (M2888) LPS (Fig. 4a, lane 3) was equivalent to that of the wild-type strains (Fig. 4a, lanes 1 and 2), the mutant LPS gave almost no immunoblot reaction (Fig. 4b, lane 3) when compared with the wild-type LPS samples (lanes 1 and 2). LPS extracted from M2900 also gave a comparably weak reaction with the anti-O10 antibody (Fig. 4a, lane 4), as indicated in Fig. 4(a). We also tested this antiserum on LPS extracted from both E2 and E4 which have an identical O unit backbone structure to E1 but no O-acetyl moiety. Both LPS were unable to bind the anti-O10 antibody used and gave a negative reaction (Fig. S1).

## The \( wbaK \) gene might be acquired by lateral gene transfer

The O antigen gene cluster sequences available for the set of Gal-initiated groups all lack the \( wbaK \) gene, except for subgroup E1, E2 and E3 strains. The \( wbaK \) gene in the E1, E2 and E3 gene clusters is flanked by the same non-coding sequences. We found that these sequences have no homology to the intergenic regions of other Gal-initiated groups including E4. There are two E4 sequences (KC688884.1 and NZ_CAGQ01000043.1) and these resemble the majority of the Gal-initiated gene clusters in the lack of the \( wbaK \) gene and its flanking DNA. That is, the sequence between \( wbaP \) and \( gnd \) in group A, B, C2–C3, D1, D2, D3 and subgroup E4 strains differs from that in subgroup E1, E2 and E3 strains. The divergence extends 37 bp into the \( wbaP \) gene, which is highly conserved in all known Gal-initiated \( Salmonella \) groups, except for nine conserved substitutions (Fig. S2) in the \( wbaP \) gene in E1 (FR775224.1, X60666.1, KC688885.1), E2 (NZ_AFCV01000775.1) and E3 (NZ_AFCM01000803.1) gene clusters. These substitutions correlate with the presence

### Table 2. \(^1\text{H}\) and \(^{13}\text{C}\) NMR chemical shifts (\( \delta, \text{ppm} \)) of the O antigen of wild-type E1 strain (M31). Spectra were recorded at 27 °C in \(^2\text{H}_2\text{O} \) relative to internal acetone (\( \delta_\text{H} 2.225; \delta_\text{C} 31.45 \))

<table>
<thead>
<tr>
<th>Sugar residue</th>
<th>1</th>
<th>( \delta_{\text{C},\text{H}} )</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6a</th>
<th>6b</th>
<th>Ac CH(_3)</th>
<th>Ac CO</th>
</tr>
</thead>
<tbody>
<tr>
<td>A 3-( \alpha )-d-Galp(^6)OAc</td>
<td>(^1\text{H})</td>
<td>5.02</td>
<td>–</td>
<td>3.93</td>
<td>3.94</td>
<td>4.10</td>
<td>4.17/4.06</td>
<td>4.23/3.74</td>
<td>4.31/3.74</td>
<td>2.12</td>
</tr>
<tr>
<td>(50%)/3-( \alpha )-d-Galp</td>
<td>(^{13}\text{C})</td>
<td>99.44</td>
<td>173</td>
<td>68.71</td>
<td>78.70</td>
<td>70.23</td>
<td>69.80/70.23</td>
<td>65.19/62.30</td>
<td>21.5</td>
<td>175.26</td>
</tr>
<tr>
<td>B 4-( \alpha )-L-Rhap</td>
<td>(^1\text{H})</td>
<td>5.02</td>
<td>–</td>
<td>4.05</td>
<td>3.98</td>
<td>3.68</td>
<td>3.89</td>
<td>1.34</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>(^{13}\text{C})</td>
<td>103.52</td>
<td>174</td>
<td>71.56</td>
<td>71.31</td>
<td>81.33</td>
<td>69.15</td>
<td>18.34</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>C 6-( \beta )-d-Manp</td>
<td>(^1\text{H})</td>
<td>4.89</td>
<td>–</td>
<td>4.09</td>
<td>3.64</td>
<td>3.76</td>
<td>3.52</td>
<td>3.74</td>
<td>4.02</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>(^{13}\text{C})</td>
<td>102.40</td>
<td>161</td>
<td>71.86</td>
<td>74.54</td>
<td>67.51</td>
<td>75.59</td>
<td>67.09</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>
of \textit{wbaK}. We propose that the \textit{wbaK} gene, together with its flanking sequences and at least 37 bp of the \textit{wbaP} gene, was acquired by lateral gene transfer in a recombination event (Fig. S3), and this is strongly supported by the substitutions in \textit{wbaP}.

**DISCUSSION**

We show that WbaK is the transferase for the O-acetyl residue in the \textit{S. enterica} group E subgroup E1 O antigen. Deletion of the \textit{wbaK} gene blocks addition of the O-acetyl moiety as shown by comparing the NMR profiles for the wild-type and \textit{ΔwbaK} (M2888) LPS, in which the wild-type gave clear signals for an O-acetylated Gal residue that were absent in the mutant. We note that \textit{wbaK} is the last gene in the gene clusters of the Gal-initiated O antigens to be assigned a function.

An earlier bioinformatic study had predicted \textit{wbaK} to be a non-functional O antigen polymerase (Wang \textit{et al.}, 1992) and some experimental work on the function of WbaK supported this (McConnell \textit{et al.}, 2001). However, that laboratory has more recently obtained experimental results supporting the acetyltransferase model and is now attempting to reconcile its latest findings with the earlier work (M. R. McConnell, personal communication).

The O-acetyl moiety is present only in subgroup E1 (Ewing, 1986). Its absence in subgroups E2 and E3 is explained by the suppressor function carried by phage ε15. E4 strains were thought to not have the O-acetyl moiety because the carbon 6 site is occupied by a Glc residue (Mäkelä & Stocker, 1984). However the finding that two E4 strains have a gene cluster sequence that predates acquisition of the \textit{wbaK} gene indicates that this does not apply in this case.

The group E gene clusters can be divided into two groups, those with or without the \textit{wbaK} gene. It is then likely that the gene cluster in the E4 strains lacking \textit{wbaK} represents the ancestral sequence, with E1 being a derived form due to the acquisition of the \textit{wbaK} gene (Fig. 5). We do not know if there are any E1 (those having epitope O3 only), E2 or E3 strains that have the original (pre-\textit{wbaK}) gene cluster, as the gene is not expressed in the presence of the ε15 phage in E2 and E3. Neither do we know if any E4 strains have the \textit{wbaK} gene, as if the glucosylation of C6 of the Gal residue predominated over O-acetylation it would have no phenotype. It is also possible that a lower level of expression due to competition from the glucosylation might not be detected by routine serology.

The recombination event that brought the \textit{wbaK} gene into group E probably involved homologous recombination in \textit{wbaP} at one end. At the other end there is a sharp break 20 bp upstream of the \textit{gnd} gene, from no homology to identity. We have to remember that what we see now may be very different from the original product of the recombination event, as if the original junction was further into the \textit{wbaP} gene and the chimeric WbaP protein expressed was less effective, then there would be selection for increasing the length of the \textit{Salmonella wbaP} sequence as the gene cluster moved by homologous recombination. It is also possible that the donor species had a \textit{gnd} gene.
not be surprising if over time recombination reduced the amount of donor DNA as discussed above, and in this case selection might also apply to the regulatory region upstream of the gnd gene.

The heterogeneity of O antigens is mostly due to variation within the gene cluster. However there is very little indication of how such variation was generated. The Salmonella Gal-initiated gene clusters have already given us two examples that give insight into gene-cluster evolution, the first being the apparent formation of the D2 gene cluster by combination of approximately equally sized segments from the group B and E gene clusters (Xiang et al., 1994), and the second being the attrition of a wzy gene in the group B gene cluster after its replacement by another wzy gene elsewhere in the chromosome (Curd et al., 1998). Transferases that O-acetylate O antigens are generally found outside of the gene cluster. The two cases that we know of an O antigen acetyltransferase gene being in the O antigen gene cluster are the S. enterica group C2 wbaJ gene and the E. coli K-12 wbbl gene. The wbaJ gene is between two of four glycosyltransferase genes. The glycosyltransferase genes of the Gal-initiated gene clusters are located in inverse order of the biochemical reactions of the encoded proteins (Liu et al., 1995), and the wbaJ gene fits the same pattern (Liu et al., 1995). The wbbl gene is also with other glycosyltransferase genes, which again are in inverse order to their predicted function in assembly, and the predicted, O-acetyltransferase gene also fits into this pattern (Stevenson et al., 1994). The wbaK gene in contrast is at one end of the gene cluster and a putative recombination event can be seen.

We do not know why the transferase gene locations are related to order of functions, but this is presumably a result of selection for some advantage related to protein–protein interactions. Incorporation of a gene into the gene cluster will enable it to move by recombination as a single entity within the species. The presence of the recombination site suggests that wbaK was added relatively recently enabling co-transfer with the other genes, but the location of the other acetyltransferase genes in order of function suggests that wbaK is at a relatively early stage in the incorporation of a gene into an O antigen gene cluster, providing a third example of insight into evolution of O antigen gene clusters.

**ACKNOWLEDGEMENTS**

This work was supported by a University of Sydney Bridging Support Grant (RIMS 2012-00068).

**REFERENCES**


Edited by: M. Holden