Archives of Microbiology

Comparative proteome analysis of Acidaminococcus intestini supports a relationship between outer membrane biogenesis in Negativicutes and Proteobacteria

Manuscript Number: AOMI-D-14-00009

Full Title: Comparative proteome analysis of Acidaminococcus intestini supports a relationship between outer membrane biogenesis in Negativicutes and Proteobacteria

Article Type: Short Communication

Abstract: The presence of bona fide outer membranes in members of the class Negativicutes is anomalous as phylogenetic analyses place this class within the phylum Firmicutes. To explore the relationships of a representative member of Negativicutes, we have performed a whole proteome BLAST analysis of Acidaminococcus intestini, which indicates that a substantial proportion (7%) of the A. intestini proteome is closely related to sequences from members of the phylum Proteobacteria. In addition we have identified key proteins involved in outer membrane biogenesis in A. intestini. This work highlights the need for further studies to define the relationships and evolutionary history of the Negativicutes.

Corresponding Author: Radhey S. Gupta, Ph.D.
McMaster
Hamilton, Ontariou CANADA

Corresponding Author's Institution: McMaster

First Author: Chantal Campbell

First Author Secondary Information:

Order of Authors: Chantal Campbell
Iain C. Sutcliffe
Radhey S. Gupta, Ph.D.

Order of Authors Secondary Information:

Author Comments: Dear Professor Stackebrandt,

I am herewith submitting a manuscript entitled "Comparative proteome analysis of Acidaminococcus intestini supports a relationship between outer membrane biogenesis in Negativicutes and Proteobacteria", in the form of a short communication, for your consideration for publication in the Archives of Microbiology. The work describes here was carried out in collaboration with Prof. Iain C. Sutcliffe of the Northumbria University (UK) and it report analysis of genome sequence data to understand the origin of outer cell membrane is some atypical Gram-negative bacteria. The results described here provide important insights in this regard. We believe the data presented is well suited to Archives of Microbiology as your journal considers manuscripts that report analysis of 'mining' of data' if new information, interpretation, or hypotheses emerge. The manuscript has been formatted to match the journal's short communication format. We hope that this work will be considered suitable for publication in Archives of Microbiology and look forward to receiving your decision soon.

Sincerely yours,

Prof Radhey Gupta
on behalf of the authors
| Suggested Reviewers: | Dr. Paul Lawson  
paul.lawson@ou.edu  
expert in the taxonomy of clostridia and relatives |
|---------------------|-------------------------------------------------------------------------------------------------|
|                     | Prof. Brian Hedlund  
brian.hedlund@unlv.edu  
expert in genomics who has used genomic data to explore cell envelope characteristics |
|                     | Dr. Damien Devos  
damien.devos@cos.uni-heidelberg.de  
expert in the use of genomic data to explore cell envelope characteristics |
Comparative proteome analysis of Acidaminococcus intestini supports a relationship between outer membrane biogenesis in Negativicutes and Proteobacteria.

Chantal Campbell¹, Iain C Sutcliffe² and Radhey S. Gupta*¹

¹Department of Biochemistry and Biomedical Sciences, McMaster University, Hamilton, Ontario, Canada L8N 3Z5

²Faculty of Health and Life Sciences, Northumbria University, Newcastle upon Tyne NE1 8ST, UK

*Corresponding Author:
Phone: (905) 525-9140 ext. 22639
Fax: (905) 522-9033
Email: gupta@mcmaster.ca
Abstract

The presence of bona fide outer membranes in members of the class Negativicutes is anomalous as phylogenetic analyses place this class within the phylum Firmicutes. To explore the relationships of a representative member of Negativicutes, we have performed a whole proteome BLAST analysis of Acidaminococcus intestini, which indicates that a substantial proportion (7%) of the A. intestini proteome is closely related to sequences from members of the phylum Proteobacteria. In addition we have identified key proteins involved in outer membrane biogenesis in A. intestini. This work highlights the need for further studies to define the relationships and evolutionary history of the Negativicutes.

Keywords: Acidaminococcus; Clostridia; lipopolysaccharide; Negativicutes; phylogeny.
Bacterial cells exhibit one of two major cell envelope architectures, either monoderm (i.e. a single cytoplasmic membrane (e.g. most *Firmicutes* and *Actinobacteria*) or diderm (i.e. a plasma membrane and a lipid outer membrane e.g. *Proteobacteria*) (Gupta 2011; Sutcliffe 2010). At the phylum level, it appears that most phyla are typically diderm and that within the typically monoderm phyla there are some important diderm exceptions (Sutcliffe 2010). An intriguing example of this is the presence of members of the class *Negativicutes* within the phylum *Firmicutes* (Marchandin et al 2010). Members of this class appear to have typical diderm cell envelopes, notably with an outer membrane based on lipopolysaccharide (Mavromatis et al. 2009; Sutcliffe 2010; Tocheva et al., 2011). In this regard it is notable that some members of the class *Clostridia* (e.g. *Halothermothrix orenii*) also exhibit diderm lipopolysaccharide-based cell envelopes. The relationship between the class *Clostridia* and the class *Negativicutes* has yet to be fully resolved; although the status of the latter class has recently been questioned by Yutin and Galperin (2013), other analyses (Segata et al. 2013; Gupta et al., unpublished) support the integrity of the *Negativicutes* taxon.

We are interested in further investigating the basis of outer membrane biogenesis in *Negativicutes*. Thus to explore the relationships between a representative *Negativicute* and members of other taxa, BLAST (Altschul et al.1997) searches were conducted on all proteins found in the *Acidaminococcus intestini* RyC-MR95 genome (D’Auria et al. 2011). The sources (species level) of the first three ‘hits’ from the BLAST search that were not members of *Negativicutes* and had expect values of less than $10^{-5}$ were recorded. The phylum of each top hit (or in the case of *Firmicutes*, the class for each top hit) was also recorded. The frequency of each top hit phylum/class was tallied to determine which phyla/classes were most
related to the *Negativicutes* with respect to the proteins analysed. Proteins that did not have a non-*Negativicutes* hit or that had an insignificant top hit (i.e. expect [E] values >10^{-5}) were excluded from the tally. As a control, the analysis was repeated using all proteins encoded in the *Erysipelothrix rhusiopathiae* genome (Ogawa et al. 2011) as this monoderm species is representative of an independent class (*Erysipelotrichia*) within the *Firmicutes*.

Only the top hit from each BLAST search was taken into account when determining the closest relatives to the *A. intestini* proteins (although the 2nd and 3rd hits typically showed similar patterns). 2027 out of the 2400 proteins were used due to the fact that 373 of the proteins did not have significant first hits (E>10^{-5}) or did not have any hits that were from non-*Negativicutes*. Hits from members of the class *Clostridia* represented approximately 68% of top relatives to the proteins, with members of the class *Bacilli* the second most frequent top hit, representing approximately 11.5% of the top relatives (Fig. 1). Notably, the third most frequent top hit (7%, 142 proteins) was to sequences from members of the Gram-negative phylum *Proteobacteria* (Fig. 1). Overall, 8.6% of the *A. intestini* proteins have closest homologues encoded by members of diderm phyla. In contrast, for the control analysis with 1257 *E. rhusiopathiae* proteins, only 1.4% of the top hits were from members of *Proteobacteria* and a total of 2.9% hits from members of diderm phyla. Thus, hits to *Proteobacteria* sequences are 5-times more frequent for an *A. intestini* query than for the *Erysipelothrix* control.

Of the 142 *A. intestini* proteins for which sequences from *Proteobacteria* were the top hits outside of *Negativicutes*, 14 (10%) corresponded to outer membrane function and 10 others (7%) can be linked to LPS biosynthesis (Supplementary Table 1). In addition, 21 (15%) of the 142 proteins are of unknown function. To
further explore the nature of the outer membrane biogenesis pathway in *A. intestini*, we therefore looked for orthologues of key proteins involved in biogenesis and functioning of the *Escherichia coli* outer membrane (Table 1). Clear homologues of all proteins were found encoded in the *A. intestini* genome, with six exceptions. Notably, the outer membrane biogenesis proteins were localised into two loci in the *A. intestini* genome, Acin_0625- Acin_0636 and Acin_1764-Acin_1776 (Table 1). The proteins lacking clear homologues by BLAST analysis include LpxH, a UDP-sugar hydrolase. However, this step in lipid A biosynthesis is bypassed by an alternative step catalysed by LpxI in α-Proteobacteria, many δ-Proteobacteria and some other diderm phyla (Metzget IV and Raetz 2010; Opiyo et al. 2010). Notably an LpxI homologue is encoded by Acin_1764 in the *A. intestini* genome (Table 1). Mavromatis et al. (2009) reported that both *Thermosinus carboxydivorans* (*Negativicutes*) and *H. orenii* (Class Clostridia, order Halanaerobiales) also have a complete lipid A biosynthesis path except for LpxH (Mavromatis et al. 2009) and an LpxI homologue is also encoded in each of these genomes (data not shown). Notably, almost all (11/12) of the *A. intestini* proteins that function in the lipid A pathway (Table 1) have a closest proteobacterial homologue from δ-Proteobacteria (data not shown).

A homologue of LptD (OstA), part of the LPS transfer machinery, was not found in the *A. intestini* genome. However, Acin_0634 is noted to contain OstA domains and resides within an *A. intestine* LPS biosynthesis locus and so may replace LptD; similarly, an LptC homologue was not detected by BLAST analysis but Acin_0633 encodes an LptC (PF06835) family member. Our analysis did not identify a homologue of BamD, an accessory part of the outer membrane assembly machinery, although this component is not uniformly conserved in diderm bacteria.
Homologues of LolA and LolB, which function in the *E. coli* pathway by which lipoproteins are moved to the outer membrane, were not identified but, again, this pathway is not well conserved even within *Proteobacteria* (Sutcliffe et al. 2013).

The above data are consistent with a close relationship between a significant proportion of the proteome (7%) of a representative of *Negativicutes* and the *Proteobacteria*, particularly with regard to cell envelope biogenesis. Importantly, the other phyla of diderm prokaryotes (e.g. *Fusobacteria*, *Synergistetes*) or even diderm members of the class *Clostridia* (i.e. members of the order *Halanaerobiales* such as *H. orenii*), did not show significant numbers of top BLAST hits to the protein queries from the representative *Negativicutes* (Fig. 1; Supplementary Table 2). With regard to the *Negativicutes*, while our results suggest that a large number of genes, particularly those involved in cell envelope biogenesis, are probably laterally acquired from *Proteobacteria*, and δ-Proteobacteria in particular, it is important to recognize that the results of BLAST hits are influenced by numerous factors and they are not always the closest relatives (Koski and Golding, 2001). Hence, to gain further understanding of the origin of the outer membrane in the *Negativicutes*, it will be helpful to carry out additional studies on members of these groups to determine the origin of the proteins related to outer membrane biogenesis.

**Acknowledgements**

The work from McMaster University was supported by a research grant from the Natural Sciences and Engineering Research Council of Canada.
References


Ogawa Y, Ooka T, Shi F, Ogura Y, Nakayama K, Hayashi T, Shimoji Y (2011) The Genome of Erysipelothrix rhusiopathiae, the causative agent of swine erysipelas,


Figure legend:

Figure 1. Top BLAST hits summarising the closest relative (phyla; class within *Firmicutes*) of 2027 signature proteins from the *A. intestinii* genome (A) or 1257 proteins from the *E. rhusiopathiae* genome (B). Phyla/classes that represented less than 1.5% of the hits were placed cumulatively into the ‘Others’ category.
Table 1. Homologue of key outer membrane (OM) biogenesis proteins and representative OM proteins identified in the *A. intestini* genome by BLAST analysis with *E. coli* proteins as query, except for LpxI (for *Caulobacter crescentus*).

<table>
<thead>
<tr>
<th><em>E. coli</em> Protein</th>
<th>UniProt code</th>
<th>Function</th>
<th><em>A. intestini</em> homologue</th>
<th>Amino acid identity (%); E number</th>
</tr>
</thead>
<tbody>
<tr>
<td>LpxA</td>
<td>P0A722</td>
<td>Lipid A biosynthesis</td>
<td>Acin_1765</td>
<td>120/262 (46%); 1x10^-75</td>
</tr>
<tr>
<td>LpxB</td>
<td>P10441</td>
<td>Lipid A biosynthesis</td>
<td>Acin_0625</td>
<td>121/379 (32%); 3x10^-30d</td>
</tr>
<tr>
<td>LpxC</td>
<td>P0A725</td>
<td>Lipid A biosynthesis</td>
<td>Acin_1767</td>
<td>107/284 (38%); 1x10^-32d</td>
</tr>
<tr>
<td>LpxD</td>
<td>P21645</td>
<td>Lipid A biosynthesis</td>
<td>Acin_1770</td>
<td>109/334 (33%); 4x10^-37</td>
</tr>
<tr>
<td>LpxH</td>
<td>P43341</td>
<td>Lipid A biosynthesis</td>
<td>No significant homologue</td>
<td></td>
</tr>
<tr>
<td>LpxI</td>
<td>B8GWR0</td>
<td>Lipid A biosynthesis</td>
<td>Acin_1764</td>
<td>92/283 (33%); 6x10^-35</td>
</tr>
<tr>
<td>LpxK</td>
<td>P27300</td>
<td>Lipid A biosynthesis</td>
<td>Acin_0627 (aa 505-840)</td>
<td>83/341 (25%); 4x10^-22</td>
</tr>
<tr>
<td>KdtA (WaaA)</td>
<td>P0AC75</td>
<td>Lipid A biosynthesis</td>
<td>Acin_0627(aa 13-430)</td>
<td>129/425 (30%); 6x10^-32</td>
</tr>
<tr>
<td>HtrB (LpxL)</td>
<td>P0ACV0</td>
<td>Lipid A biosynthesis</td>
<td>Acin_0632</td>
<td>70/285 (25%); 1x10^-15</td>
</tr>
<tr>
<td>LpxM</td>
<td>C4ZZL2</td>
<td>Lipid A biosynthesis</td>
<td>Acin_0632</td>
<td>61/272 (22%); 7x10^-12</td>
</tr>
<tr>
<td>KdsA</td>
<td>P0A715</td>
<td>Lipid A biosynthesis</td>
<td>Acin_0629</td>
<td>126/268 (47%); 7x10^-31</td>
</tr>
<tr>
<td>KdsB</td>
<td>P04951</td>
<td>Lipid A biosynthesis</td>
<td>Acin_0628</td>
<td>117/239 (49%); 3x10^-36</td>
</tr>
<tr>
<td>KdsC</td>
<td>P0ABZ4</td>
<td>Lipid A biosynthesis</td>
<td>Acin_0631</td>
<td>72/157 (46%); 3x10^-39</td>
</tr>
<tr>
<td>KdsD</td>
<td>P45395</td>
<td>Lipid A biosynthesis</td>
<td>Acin_0630</td>
<td>164/321 (51%); 1x10^-13</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>Gene</th>
<th>Accession</th>
<th>Function</th>
<th>Protein Name</th>
<th>E value</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>LptA</td>
<td>P0ADV1</td>
<td>LPS export (periplasmic Lipid A shuttle)</td>
<td>Acin_2165</td>
<td>39/166</td>
<td>(23%); 0.046</td>
</tr>
<tr>
<td>LptB</td>
<td>P0A9V1</td>
<td>LPS export</td>
<td>Acin_0635</td>
<td>130/237 (55%); 1x10^-30</td>
<td></td>
</tr>
<tr>
<td>LptC</td>
<td>P0ADV9</td>
<td>LPS export</td>
<td>No significant homologue*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LptD</td>
<td>P31554</td>
<td>LPS export (insertion of LPS into OM)</td>
<td>No significant homologue*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MsbA</td>
<td>P60752</td>
<td>Lipid A flippase</td>
<td>Acin_0626</td>
<td>208/572 (36%); 5x10^-121</td>
<td></td>
</tr>
<tr>
<td>BamA</td>
<td>P0A940</td>
<td>Signature protein for OM biogenesis</td>
<td>Acin_1774</td>
<td>137/560 (24%); 1x10^-28</td>
<td></td>
</tr>
<tr>
<td>BamD</td>
<td>P0AC02</td>
<td>OM biogenesis</td>
<td>No significant homologue</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LolA</td>
<td>P61316</td>
<td>OM lipoprotein shuttle</td>
<td>No significant homologue</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LolB</td>
<td>P61320</td>
<td>OM lipoprotein insertion</td>
<td>No significant homologue</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TolC</td>
<td>P02930</td>
<td>Canonical OM protein (type 1 secretion systems)</td>
<td>Acin_1776</td>
<td>103/409 (25%); 5x10^-20</td>
<td></td>
</tr>
<tr>
<td>GspD</td>
<td>P45758</td>
<td>Canonical OM protein (type 2 secretion system)</td>
<td>Acin_0088</td>
<td>74/284 (26%); 1x10^-26</td>
<td></td>
</tr>
</tbody>
</table>

* See main text