Genome Update: rRNAs in sequenced microbial genomes

Genomes of the month – three superkingdoms of life

Once again, this month, three microbial genomes have been published in the 4 weeks since the last Genome Update was written. However, in contrast to last month (Ussery & Hallin, 2004), when all three genomes were bacterial, this time each genome represents a different superkingdom of life. The three genomes include that of a methanogen isolated from a salt-marsh sediment (the archaean Methanococcus maripaludis), that of a soil-dwelling bacillus found to cause cheese spoilage (the bacterium Bacillus cereus) and a yeast genome (the eukaryote Ashbya gossypii).

Methanococcus maripaludis is a mesophilic, methane-producing, nitrogen-fixing member of the Archaea; it is a strict anaerobe and grows in the presence of hydrogen and carbon dioxide. The Methanococcus maripaludis S2 genome (Hendrickson et al., GenBank accession no. BX950229) is 1661 137 bp long (see Table 1), just a few thousand base pairs shorter than that of the thermophile Methanothermobacter jannaschii DSM 2661T, and in general it is one of the smaller archaeal genomes. The Methanococcus maripaludis genome contains few DNA repeats and has three rRNA operons.

The genome of B. cereus ATCC 10987 is 5·2 Mbp long and has about 5600 annotated genes (Rasko et al., 2004); in many ways, the genome is more similar to the sequenced genome of Bacillus anthracis than to that of the other sequenced B. cereus strain. B. cereus, Bacillus thuringiensis and B. anthracis are all members of the B. cereus sensu lato group, and comparison of 16S rRNA gene sequences alone cannot distinguish between members of this group. The genome of B. cereus ATCC 10987 encodes 12 rRNA operons and 98 tRNAs, as shown in Table 1.

The filamentous fungus A. gossypii is used by industry in vitamin production. The genome of A. gossypii ATCC 10895 is the smallest eukaryotic genome sequenced to date (Dietrich et al., 2004), encoding 4718 genes, or close to 1000 genes fewer than the B. cereus genome (see Table 1). Like many eukaryotic genomes, the presence of repeats makes the task of sequencing the whole A. gossypii genome quite difficult, and it is estimated that the genome is 9·2 Mbp long, although only 8·7 Mbp are present in the EMBL files. Furthermore, although it is known that there are about 40 copies of an rRNA repeat in the genome, only a single copy is found in the EMBL file. We have listed two rows in Table 1 for this genome – the first row lists what is reported in the sum of the EMBL files for all seven chromosomes and the second row lists the numbers given in the genome report (Dietrich et al., 2004).

Method of the month – comparative sequence analysis of 16S rRNA genes in sequenced genomes

An unrooted phylogenetic tree of 16S sequenced prokaryotic genomes is shown in Fig. 1. The comparison of 16S rRNA gene sequences to infer the phylogenetic relationships among prokaryotes has been widely used for several decades, and the 16S rRNA molecule is generally accepted as the ultimate molecular chronometer.

<table>
<thead>
<tr>
<th>Genome</th>
<th>Size (bp)</th>
<th>AT content (%)</th>
<th>rRNA operons</th>
<th>tRNAs</th>
<th>CDS</th>
<th>Accession nos</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methanococcus maripaludis S2</td>
<td>1 661 137</td>
<td>66·9</td>
<td>3</td>
<td>37</td>
<td>1722</td>
<td>BX950229</td>
</tr>
<tr>
<td>Bacillus cereus ATCC 10987</td>
<td>5 224 283</td>
<td>64·4</td>
<td>12</td>
<td>98</td>
<td>5642</td>
<td>AE017194</td>
</tr>
<tr>
<td>Ashbya gossypii ATCC 10895</td>
<td>8 741 255</td>
<td>48·3</td>
<td>1</td>
<td>199</td>
<td>4718</td>
<td>AE016814–AE016821</td>
</tr>
<tr>
<td>(Genome estimates) (~9·2 Mbp)</td>
<td>(48)</td>
<td>(40)</td>
<td>(199)</td>
<td>(4718)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Microbiology Comment provides a platform for readers of Microbiology to communicate their personal observations and opinions in a more informal way than through the submission of papers.

Most of us feel, from time to time, that other authors have not acknowledged the work of our own or other groups or have omitted to interpret important aspects of their own data. Perhaps we have observations that, although not sufficient to merit a full paper, add a further dimension to one published by others, or we may have a useful piece of methodology that we would like to share.

Guidelines on how to submit a Microbiology Comment article can be found in the Instructions for Authors at http://mic.sgmjournals.org

It should be noted that the Editors of Microbiology do not necessarily agree with the views expressed in Microbiology Comment.

Chris Thomas, Editor-in-Chief
because it is functionally constant, shows a mosaic structure of conserved and more variable regions and because of its universal presence (Woese, 1987). As can be seen from Fig. 1, complete genomes are available for most major bacterial groups (including the five subdivisions of the Proteobacteria, the Firmicutes, the Actinobacteria and the Cyanobacteria) and for representatives of the two major archaeal groups, the phyla Euryarchaeota and Crenarchaeota. The large distance between the Euryarchaeota and Crenarchaeota can be readily seen in this figure.

The number of rRNA operons per sequenced bacterial genome varies from one to 13, and there is some occasional heterogeneity within the rRNA gene sequences of an organism (Coenye & Vandamme, 2003). The genome of B. cereus ATCC 10987 (Rasko et al., 2004) has 12 rRNA operons, which is a fairly large number when compared to the other sequenced bacterial genomes. The genome of B. cereus ATCC 14579T has 13 rRNA operons, which is the largest number of rRNA operons observed among the sequenced bacterial genomes: the Methanococcus maripaludis genome contains only three rRNA operons. See supplemental web pages for a complete list of all 154 sequenced bacterial genomes, sorted by number of rRNA operons. So far, the prokaryote in which most rRNA operons have been detected is Clostridium paradoxum, which has 15 rRNA operons (Klappenbach et al., 2001).

Next month, the ‘method’ of genome comparison discussed will be a look at tRNA genes in prokaryotic genomes. For the sequenced bacterial genomes, the number of tRNA genes varies from 113 in the genome of Vibrio parahaemolyticus to 29 in that of Mycoplasma pulmonis.

Supplemental web pages
Web pages containing supplemental material related to this article can be accessed from the following url: http://www.cbs.dtu.dk/services/GenomeAtlas/suppl/GenUp004/

Acknowledgements
This work was supported by a grant from the Danish Center for Scientific Computing (DCSC).

Fig. 1. Phylogenetic tree based on 16S rRNA gene sequences from 150 bacterial and 15 archaeal sequenced genomes. The tree was constructed using the neighbour-joining method. All major bacterial and archaeal groups have been indicated, as well as the phylogenetic positions of B. cereus and Methanococcus maripaludis. Bar, 10% sequence dissimilarity. The width of the triangle represents the relative number of rRNA sequences (i.e. the number of bacterial genomes sequenced) for a given phylum.
David W. Ussery,1 Peter F. Hallin,1 Karin Lagesen2 and Tom Coenye3

1Center for Biological Sequence Analysis, Department of Biotechnology, Building 208, The Technical University of Denmark, Lyngby, DK-2800, Denmark

2Department of Molecular Biology, Institute of Medical Microbiology, University of Oslo, The National Hospital, 0027 Oslo, Norway

3Laboratorium voor Microbiologie, Universiteit Gent, Ledeganckstraat 35, B-9000 Gent, Belgium


DOI 10.1099/mic.0.27173-0