

Genome Update: alignment of bacterial chromosomes

Genomes of the month

There are four new microbial genomes listed in this month's Genome Update, three belonging to Gram-positive bacteria and one belonging to an archaeon that lives at pH 0; all of these genomes are listed in Table 1. The method of genome comparison this month is that of genome alignment and, as an example, an alignment of seven *Staphylococcus aureus* genomes and one *Staphylococcus epidermidis* genome is presented.

The genome of *Picrophilus torridus* strain DSM 9790^T, a member of the *Euryarchaeota*, has been published recently (Futterer *et al.*, 2004). *P. torridus* grows under harsh conditions: optimal growth is at pH 0.7 and 65 °C, making it one of the most thermoacidophilic organisms known. Life under these harsh conditions (*P. torridus* can grow in 1.2 M sulfuric acid!) obviously requires some biochemical tricks. The internal pH is 4.6, but this is still a 10 000 times lower proton concentration than its surrounding environment. Just maintaining this intracellular pH requires a substantial amount of energy. It is perhaps not surprising that the amino acid usage is a bit different for this genome than for other members of the *Euryarchaeota*. There is a strong preference for the aliphatic amino acid isoleucine (about 11 % of the total), which is a trait common to other thermoacidophiles, such as *Thermoplasma acidophilum* and *Thermoplasma volcanium*, but is not seen in other archaeal or bacterial genomes. Amino acid usage plots were discussed in a previous Genome Update (Ussery *et al.*, 2004), and can be found on the web page associated with this article.

The original ancestor Ames strain (Ames 0581) of *Bacillus anthracis* has been sequenced by The Institute for Genomic Research (TIGR) and deposited with GenBank (accession no. AE017334); a publication describing this genome is

anticipated soon. The *B. anthracis* strain Porton (from Porton Down, UK) was used for the previously published 'Ames' strain genome sequence (Read *et al.*, 2003). However, this strain had been cured of its plasmids, and subsequently the integrity of the main chromosomal sequence might have been compromised. In light of this, TIGR has sequenced Ames 0581, isolated from a dead cow in Texas in 1981, which includes the main chromosome and the plasmids pX01 and pX02. Thus, the authors are requesting that this genome be used as the reference strain in genomics studies involving comparison of *B. anthracis* strains.

Two new *S. aureus* genomes from two different clinical isolates have also been published this month (Holden *et al.*, 2004). One of the strains is a methicillin-resistant, hospital-acquired strain (MRSA 252); the other is a community-acquired, methicillin-susceptible strain (MSSA 476). The genome of strain MRSA 252 is quite divergent from that of strain MSSA 476, and also from the other sequenced *S. aureus* genomes. (Noteworthy is the fact that *S. aureus* MRSA 252 is from the EMRSA-16 clonal group responsible for more than half of the multi-drug-resistant *S. aureus* infections in the UK). These two genomes are compared to five other *S. aureus* genomes in the next section.

Table 1. Summary of the published genomes discussed in this Update

Note that the accession number for each chromosome is the same for GenBank, EMBL and the DNA DataBase of Japan (DDBJ).

Genome	Size (bp)	AT content (%)	rRNA operons	tRNAs	CDS	Accession no.
<i>Bacillus anthracis</i> Ames 0581	5 228 310	65	11	95	5821	AE017334
<i>Staphylococcus aureus</i> MRSA 252	2 902 619	67	5	60	2671	BX571856
<i>Staphylococcus aureus</i> MSSA 476	2 799 802	67	6	59	2565	BX571857
<i>Picrophilus torridus</i> DSM 9790 ^T	1 545 895	64	1	47	1535	AE017261

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

Method of the month – alignment of bacterial genomes

There are several different ways of aligning bacterial genomes. One common method is to use ‘MUMmer’ (Delcher *et al.*, 1999). This method is fast and

allows easy visualization of regions of similarity between two chromosomes; for example, an alignment between any two published microbial genomes can be viewed from TIGRs Comprehensive Microbial Resource (CMR) web pages

(<http://www.tigr.org>). Another method is the Artemis Comparison Tool (ACT) from the Sanger Centre, which allows the comparison of multiple chromosomes. The software can be downloaded from <http://www.sanger.ac.uk/Software/ACT>

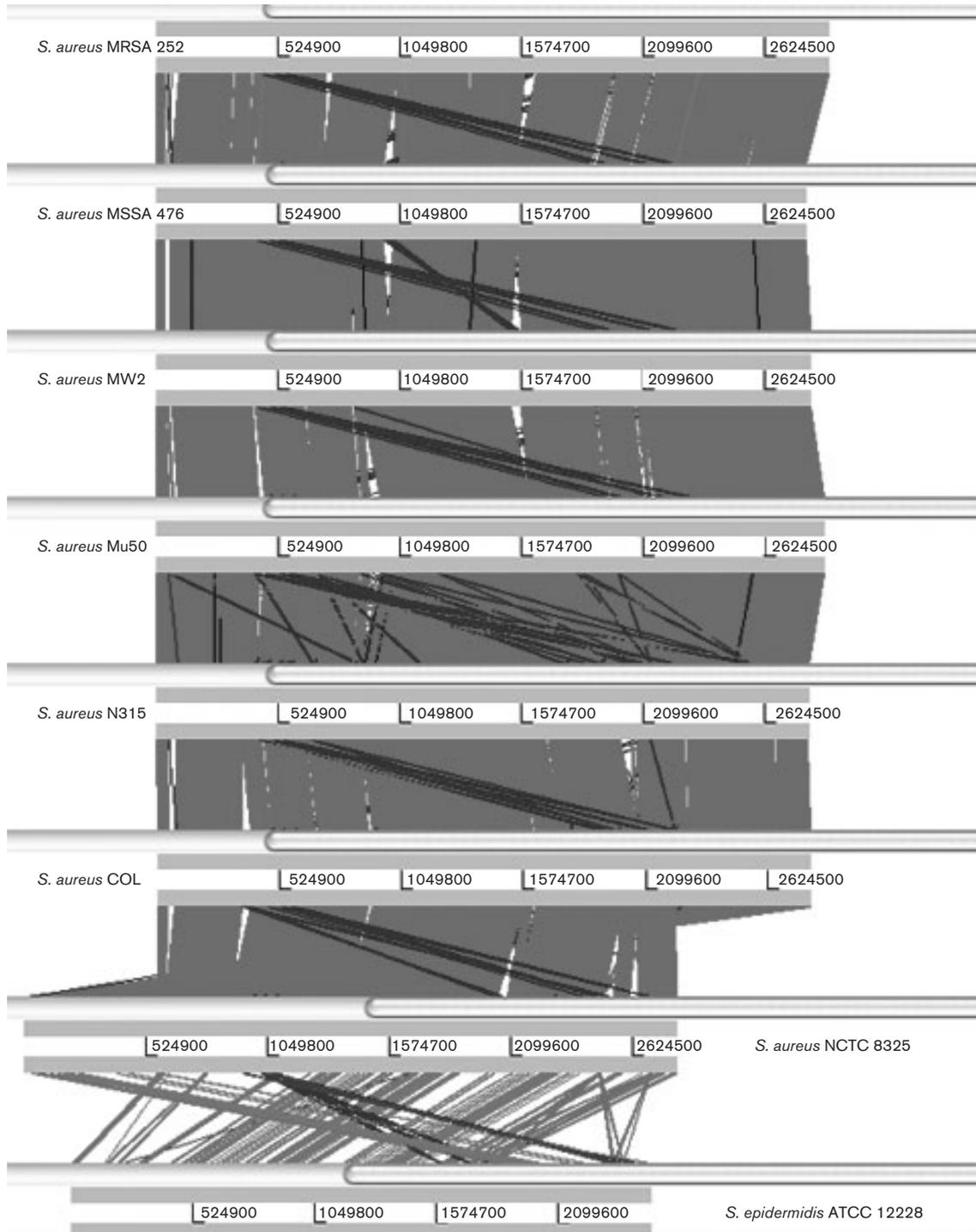


Fig. 1. Alignment of eight *Staphylococcus* genomes, created using the ACT (Artemis Comparison Tool).

and is fairly easy to use; for example, Fig. 1 was constructed in a single morning, using a portable Macintosh computer. This method is good for a quick visualization of regions of the chromosomes that are similar to each other.

A total of seven *S. aureus* genomes have now been sequenced – more than for any other bacterium to date. Comparison of multiple genomes from the same species can provide evidence for short-term microevolution (Feil, 2004). We have used the ACT method to compare the *S. aureus* genomes to each other, as shown in Fig. 1. One of the most striking features of this figure is how similar the chromosomes are to each other, with a few regions of insertions (and deletions) visible. Chromosomes for many other species do not show such nice alignments; for example, *Escherichia coli* chromosomes vary by more than a million base pairs of insertions/deletions scattered throughout the genome. The top two genomes represent the two strains discussed above. Notice that there are three regions in MRSA 252 (white gaps at the top) that are missing from MSSA 476, whilst there is one large region present in MSSA 476 that is absent from MRSA 252. Strain MW2 (third from the top) is quite similar, in agreement with multi-locus sequence tag (MLST) data (Holden *et al.*, 2004). Also, strains N315 and Mu50 are quite similar to each other, again consistent with previous estimates based on MLST data. The genome of *S. aureus* COL aligns well with that of *S. aureus* N315, except that the COL sequence starts in a different place (the COL sequence is preliminary data from the University of Oklahoma, and does not start at the origin of replication like the other strains shown in the figure). Finally, the *S. epidermidis* strain is quite different (as expected) from the *S. aureus* strains. Thus, in conclusion, use of the ACT allows a fast and powerful method for visualization of the alignment of several bacterial chromosomes.

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/GenUp007/>

Acknowledgements

This work was supported by a grant from the Danish National Research Foundation.

David W. Ussery, Mette S. Jensen, Tine R. Poulsen and Peter F. Hallin

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

Correspondence: David W. Ussery (dave@cbs.dtu.dk)

Delcher, A. L., Kasif, S., Fleischmann, R. D., Peterson, J., White, O. & Salzberg, S. L. (1999). Alignment of whole genomes. *Nucleic Acids Res* **27**, 2369–2376.

Feil, E. J. (2004). Small change: keeping pace with microevolution. *Nat Rev Microbiol* **2**, 483–495.

Futterer, O., Angelov, A., Liesegang, H., Gottschalk, G., Schleper, C., Schepers, B., Dock, C., Antranikian, G. & Liebl, W. (2004). Genome sequence of *Picrophilus torridus* and its implications for life around pH 0. *Proc Natl Acad Sci U S A* **101**, 9091–9096.

Holden, M. T. G., Feil, E. J., Lindsay, J. A. & 42 other authors (2004). Complete genomes of two clinical *Staphylococcus aureus* strains: evidence for the rapid evolution of virulence and drug resistance. *Proc Natl Acad Sci U S A* Epub ahead of print, doi:10.1073/pnas.0402521101

Read, T. D., Peterson, S. N., Tourasse, N. & 49 other authors (2003). The genome sequence of *Bacillus anthracis* Ames and comparison to closely related bacteria. *Nature* **423**, 81–86.

Ussery, D. W., Hallin, P. F., Lagesen, K. & Wassenaar, T. M. (2004). Genome Update: tRNAs in sequenced microbial genomes. *Microbiology* **150**, 1603–1606.

DOI 10.1099/mic.0.27416-0