

Bringing gene order into bacterial shape

Javier Tamames, María González-Moreno, Jesús Mingorance, Alfonso Valencia and Miguel Vicente

A different arrangement of a cluster of genes involved in division and cell-wall synthesis separates bacilli from other bacteria in a phylogenetic analysis. We conclude that the relationships between these genes are not random and might reflect significant events in the evolution of the coupling between growth and division in bacteria.

The shape of a bacterium determines its surface-to-volume ratio. In cocci, which are spherical, this ratio decreases during growth. By contrast, rod-shaped bacilli have a constant surface-to-volume ratio as they grow. Bacilli (e.g. *Escherichia coli*) have genes whose loss (*rodA*, *pbpA*) or increase in function (*bolA*) turn rod-shaped cells into spheres. There are no known instances of genetic alterations turning spherical cells into rod-shaped ones. We detect a significant correlation between the shape of bacterial cells and the bacterial phylogenetic relationship deduced from the arrangement of genes in the *dcw* cluster¹, the main cluster of genes involved in division and cell-wall synthesis (Fig. 1).

The whole process of division is not fully understood^{2,3}, but, in most bacteria, division requires the products of several *fts* genes in the *dcw* cluster¹. Some of the gene products (e.g. *ftsZ*, *ftsA* in *E. coli* and *Bacillus subtilis*) interact and form a division ring in the middle of the cell that is proposed to guide constriction of the membrane and synthesis of the peptidoglycan septum^{2,3}.

The organization of the *dcw* cluster genes has been analyzed in various bacteria. In *E. coli*, the cluster contains 15 closely packed genes; the best known is *ftsZ*. Conservation of the clustering and the relative order of many *dcw* cluster genes is obvious when examining genomes as distantly related as *E. coli* and *B. subtilis*^{1,4} (Fig. 1). To measure the degree of conservation in gene order⁵, we examined the presence of all possible pairs of neighbouring *dcw* cluster genes in eubacterial genomes (mostly fully sequenced) and constructed a phylogenetic tree using a parsimony procedure (Fig. 2).

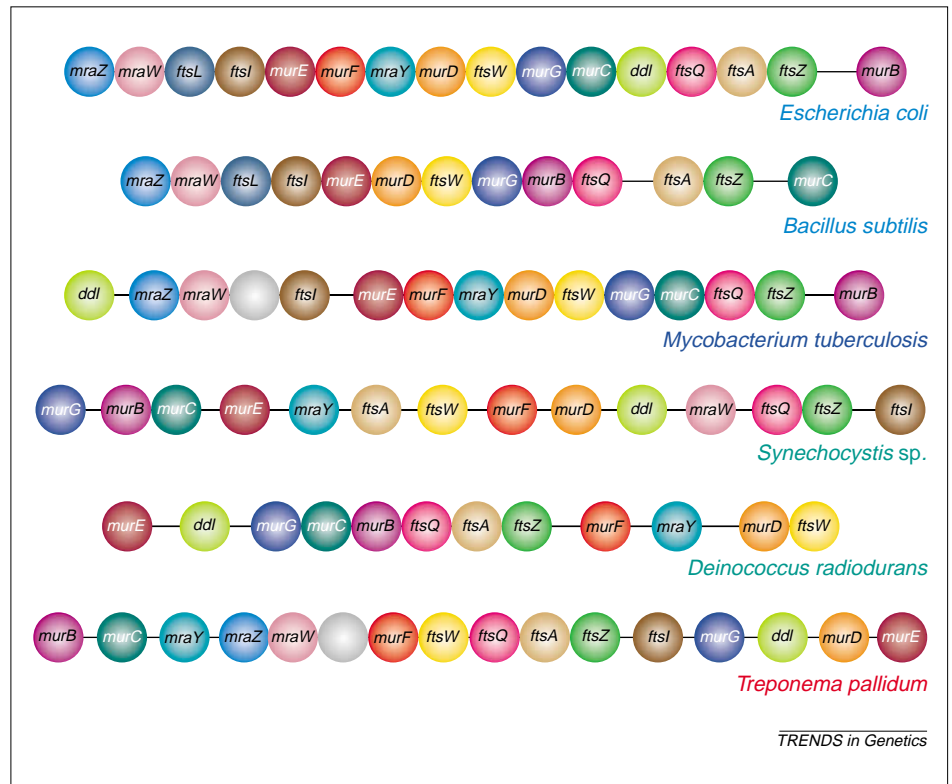


Fig. 1. Organization of the *dcw* cluster in different bacterial species. Each sphere represents one gene. Gray spheres represent a gene that is not involved in division or cell-wall synthesis and therefore is not being considered as part of the cluster. Spheres joined by a line are separated by more than one intervening gene. The coloured species names indicate morphology: light blue, bacilli (rod-shaped); dark blue, actinomycetes; green, cocci (spherical); red, spirochetes. The organization of the *dcw* cluster for the rest of the species can be found at <http://montblanc.cnb.uam.es/tamames/dcw>

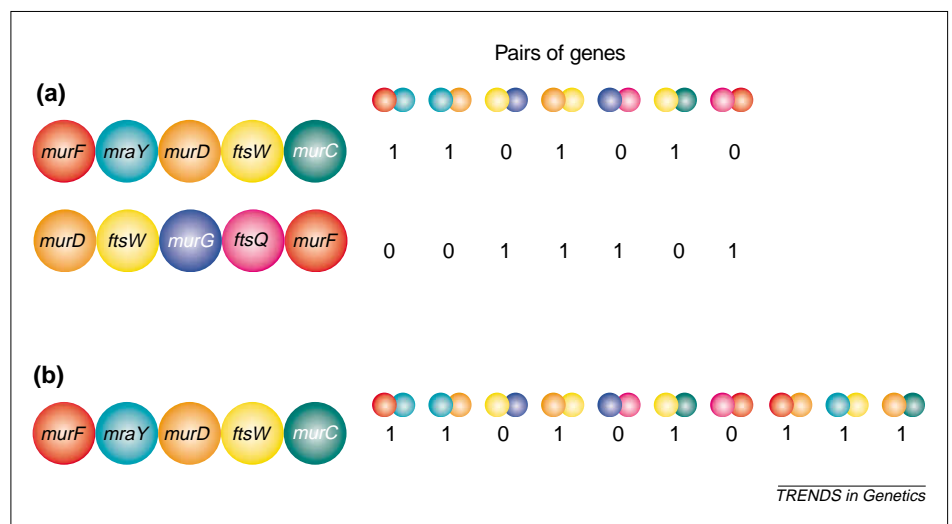


Fig. 2. Procedure for coding gene-order information in vectors. Each species is represented by vectors composed of ones and zeros, corresponding to the presence or absence, respectively, of every possible pair of genes in the *dcw* cluster. Two genes are considered to form a pair if they are located at less than a given distance (expressed in number of intercalated genes between them). For (a) this distance is zero and genes must be adjacent to be a pair. For (b) the distance is one, thus genes separated by one intervening gene are also considered as a pair. The best results (in terms of stability of the branches in the tree) were obtained for a distance of one.

The tree structure (Fig. 3) reveals a correlation of cell shape with the order of the genes and not simply with their presence or absence. Most strikingly, this classification differs from the standard relationships between bacterial species obtained when using traits such as the comparison of small-subunit rRNAs⁶ or gene family composition of genomes^{7,8}. It also varies significantly from the tree obtained when considering simply the presence or absence of the individual *dcw* genes; it is closer to the proposed phylogenies based on rRNA conservation. Finally, we could not reproduce the correlation between gene arrangement and shape when using the gene order from other gene clusters; for instance, the well-conserved cluster of ribosomal proteins.

Unusual groupings in our tree include the split of the gram-positive bacteria, in which *Bacillus subtilis*, *Bacillus anthracis* and *Clostridium acetobutylicum* appear close to the gram-negative bacilli, and the gram-positive cocci are grouped apart. Similarly, the group *Thermus-Deinococcus* is split, with *Thermus thermophilus* well grouped with the bacilli and *Deinococcus radiodurans* more related to other cocci.

How was the *dcw* cluster in bacilli conserved? If horizontal gene transfers (rather than selection) are responsible, they must have occurred recently because gene order tends to be lost rapidly⁹. In this case, it is very likely that the GC contents of *dcw* genes would differ significantly from the average GC content of the harbouring genomes. However, we find that the GC contents of *dcw* genes in different bacilli closely match the global GC content of the species, not exceeding one standard deviation (for the instances analyzed). Therefore, we exclude horizontal gene transfer as a trivial explanation for *dcw* cluster conservation in bacilli.

Conservation of the *dcw* cluster genes (Fig. 4) and of their arrangement could be a paradigm in bacterial phylogeny. On the basis of the lack of an exact correlation between morphology and phylogeny derived from small-subunit rRNAs, Siefert and Fox¹⁰ suggested that bacillar morphology is the common bacterial ancestral shape. They propose that although the rod shape can be lost, giving rise to coccal or other differently shaped organisms, it is not possible to regain it.

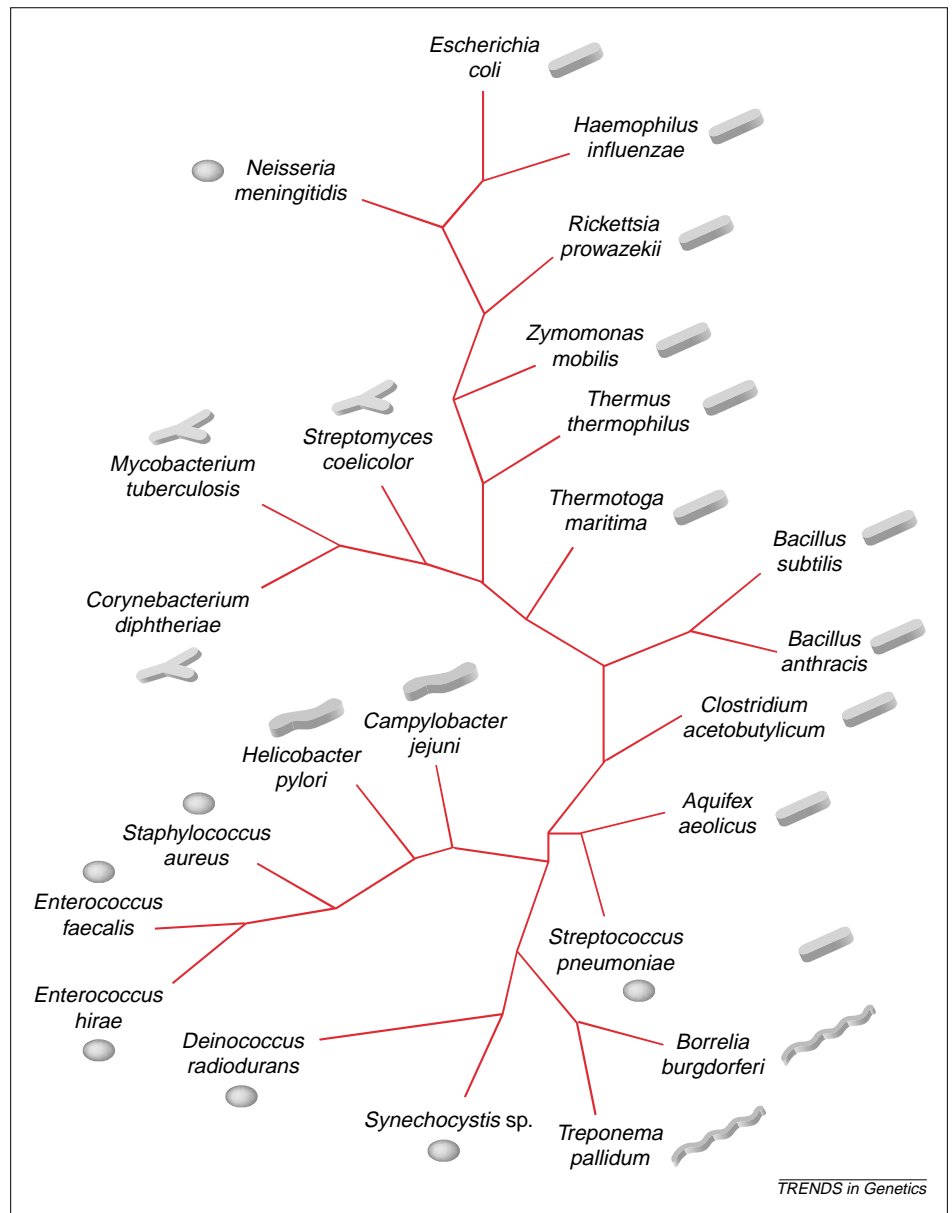


Fig. 3. Comparison of bacterial species based on the conservation of the short-range order of the *dcw* genes. A consensus phylogenetic tree using gene ordering as the studied character is shown. We compared the vectors representing the species (Fig. 2) using a parsimony method¹⁴ to reconstruct the simplest path for the transition between presence and absence of pairs. This tree is stable because the major branches appear in at least 90% of the most parsimonious solutions. As stated in Fig. 2, one intervening gene is admitted to form a pair. Unpublished *Thermus thermophilus* data are courtesy of T. Hartsch (Göttingen Genomics Laboratory, Georg-August-University, Göttingen, Germany).

Our observation indicates that the genomic organization of the *dcw* cluster is an important trait, positively selected for in bacilli in which, as shown in Figs 1 and 4, the cluster is more conserved and compact than in bacteria of other shapes. Transitions to other morphologies seem to imply genomic rearrangements and loss of gene-order conservation to a variable degree. Although it is not the main factor determining the morphology of the cells (see below), keeping a compact and ordered *dcw* cluster seems important (and selectable) for bacilli, and this importance

lessens or disappears altogether for other shapes, in which rearrangements of the *dcw* genes would then be allowed.

In agreement with Siefert and Fox, we consider that a transition to bacillar shape from other morphologies is unlikely. Among other difficulties, the theoretical reversion process would include both the reordering of the *dcw* genes to reconstruct the arrangement present in rods and the reversion of any modification introduced in the mechanisms that effect cell division. These mechanisms are known to be different in rods and cocci. The fact that

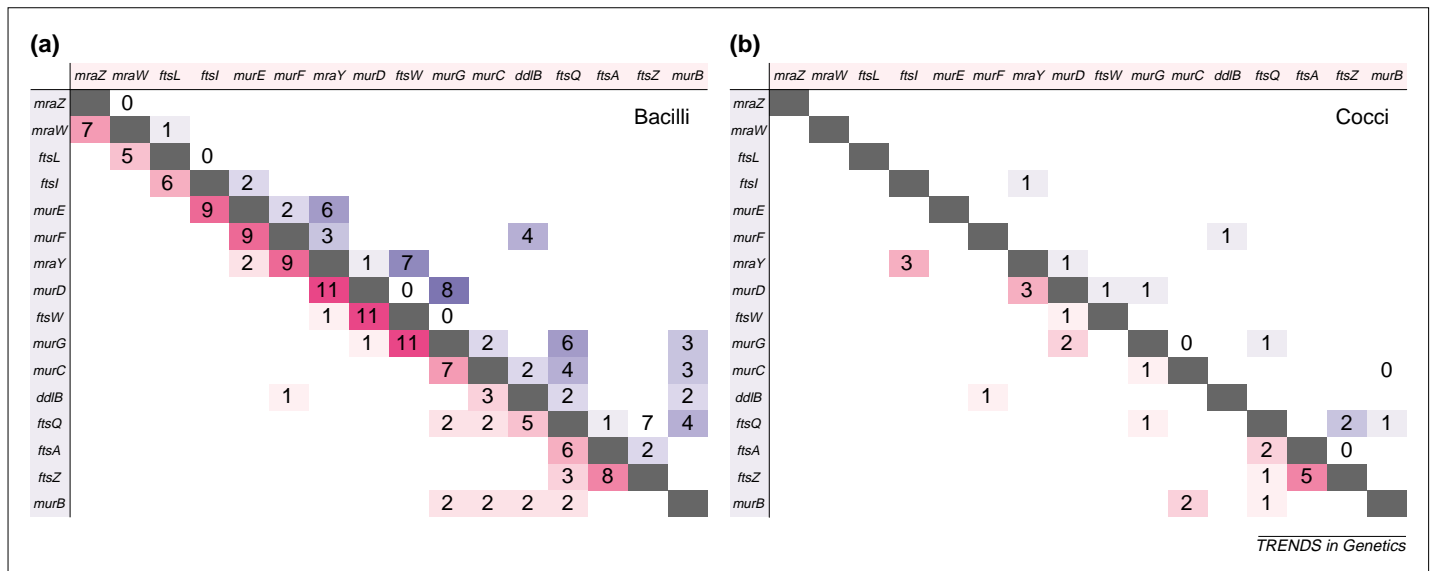


Fig. 4. Conservation of the *dcw* gene composition and order in the genomes of 13 bacilli (a) and six cocci (b) shown in Fig. 3. Red indicates a pair, with intensity proportional to the frequency of the pair in the studied genomes. Blue represents the absence of a pair where the individual genes are present, an indication of the likelihood for a given pair to be present by random association; intensity is proportional to the frequency that these pairs are missing. Pairs that cannot be formed because of the absence of the genes, or that have never been detected are left empty. Numbers within the cells indicate the number of occurrences.

Neisseria meningitidis has a bacilli-like *dcw* cluster even though it is a spherical cell (Fig. 3) can be explained if its transition to spherical shape was too recent to allow shuffling of the *dcw* cluster genes. In support of this view, we note that the *N. meningitidis* *dcw* cluster is less conserved than in closely related bacilli and that the morphologies of other *Neisseria* species (e.g. *Neisseria elongata*) resemble bacilli. Intriguingly, in *Neisseria gonorrhoeae* the cluster is split into four transcriptional units¹¹.

In at least two species (nonsporulating *E. coli* and sporulating *B. subtilis*) co-regulation of several *dcw* genes is important for the biology of the cells¹², which would argue in favour of extended gene order conservation in the cluster. There are gene pairs specific for different branches of the tree: the *murD-ftsW* and *ftsW-murG* pairs (defining the higher order *murD-ftsW-murG* structure) are almost exclusively conserved in bacilli, and the *ftsW-ftsQ* pair is present only in spirochetes. Specific gene pairings could be correlated with the evolution of the enzyme complexes that generate the molecular architecture underlying different cell shapes. Several proteins encoded by *dcw* cluster genes interact (e.g. *ftsA-ftsZ* in the division septum) or form adjacent steps of a biochemical pathway (e.g. the products of *murG* and *ftsW* in the translocation of peptidoglycan precursors to the periplasm). The disruption of the gene arrangement in

the *dcw* cluster has not been demonstrated to cause the loss of the *E. coli* bacillar shape¹³. However, a strain in which *ftsZ* is artificially disconnected from its complex natural regulatory signals shows alteration in its division timing resulting in both a change in its aspect ratio (becoming longer) and a high genetic instability¹². We conclude that the neighbor relationships between the *dcw* genes (many of them essential) is not random and that they might reflect significant events in the evolution of the coupling between growth and division in bacteria¹².

Acknowledgements

A.V. is also at Centro de Astrobiología, INTA/CSIC, 28850 Torrejón de Ardoz, Madrid, Spain. M.V. is funded by the European Commission Cell Factory project CT96-0122.

References

- 1 Ayala, J.A. *et al.* (1994) Molecular biology of bacterial septation. In *New Comprehensive Biochemistry: Bacterial Cell Wall* (Vol. 27) (Hakenbeck, R. and Ghuyssen, J.-M., eds), pp. 73–101, Elsevier
- 2 Margolin, W. (2000) Themes and variations in prokaryotic cell division. *FEMS Microbiol. Rev.* 24, 531–548
- 3 Bramhill, D. (1997) Bacterial cell division. *Annu. Rev. Cell Dev. Biol.* 13, 395–424
- 4 Nikolaichik, Y.A. and Donachie, W.D. (2000) Conservation of gene order amongst cell wall and cell division genes in eubacteria, and ribosomal genes in eubacteria and eukaryotic organelles. *Genetica* 108, 1–7

- 5 Tamames, J. *et al.* (1997) Conserved clusters of functionally related genes in two bacterial genomes. *J. Mol. Evol.* 44, 66–73
- 6 Woese, C.R. (1987) Bacterial evolution. *Microbiol. Rev.* 51, 221–271
- 7 Snel, B. *et al.* (1999) Genome phylogeny based on gene content. *Nat. Genet.* 21, 108–110
- 8 Fitz-Gibbon, S.T. and House, C.H. (1999) Whole genome-based phylogenetic analysis of free-living microorganisms. *Nucleic Acids Res.* 27, 4218–4222
- 9 Itoh, T. *et al.* (1999) Evolutionary instability of operon structures disclosed by sequence comparisons of complete microbial genomes. *Mol. Biol. Evol.* 16, 332–346
- 10 Siefert, J.L. and Fox, G.E. (1998) Phylogenetic mapping of bacterial morphology. *Microbiology* 144, 2803–2808
- 11 Francis, F. *et al.* (2000) Organization and transcription of the division cell wall (*dcw*) cluster in *Neisseria gonorrhoeae*. *Gene* 251, 141–151
- 12 Palacios, P. *et al.* (1996) Dependency of *Escherichia coli* cell-division size, and independency of nucleoid segregation on the mode and level of *ftsZ* expression. *Mol. Microbiol.* 20, 1093–1098
- 13 Hara, H. *et al.* (1997) A promoter for the first nine genes of the *Escherichia coli mra* cluster of cell division and cell envelope biosynthesis genes, including *ftsI* and *ftsW*. *J. Bacteriol.* 179, 5802–5811
- 14 Felsenstein, J. (1996) Inferring phylogenies from protein sequences by parsimony, distance and likelihood methods. *Methods Enzymol.* 266, 418–427

J. Tamames*

M. González

Centro de Astrobiología, INTA/CSIC, 28850 Torrejón de Ardoz, Madrid, Spain.

*e-mail: tamames@cnb.uam.es

A. Valencia

J. Mingorance

M. Vicente

Centro Nacional de Biotecnología, CSIC, 28049 Madrid, Spain.