microbiology

"!...Comment...?"

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. In other instances we may have a useful piece of methodology that we would like to share.

The Editors hope that readers will take full advantage of this section and use it to raise matters that hitherto have been confined to a limited audience.

***Christopher M. Thomas*, Editor-in-chief**

Cohesion of sister chromosomes in *Escherichia coli*

This is a reply to an article published in a previous issue of *Microbiology* (4). We first demonstrated cohesion of replicated *oriC* copies in *E. coli* cells (temperature sensitive *dnaA* mutant cells) synchronized for initiation of chromosomal DNA replication [3; this manuscript is not cited in the associated comment article by Nanninga *et al.* (4)]. In the same manuscript, we also demonstrated that a single *oriC* focus, which was detected by FISH (fluorescence *in situ* hybridization), was located at the middle of the cell (more correctly, at the middle of the nucleoid) just before and after initiation of replication. A single *oriC* focus appeared to separate into two foci late in the chromosome replication cycle. The result suggests that *oriC* is replicated at the middle of the nucleoid and these sister copies remain closely associated for a long time. Replicated sister copies of other chromosomal regions also remain associated after replication until late in replication (11). Sister *oriC* copies tend to locate as a single focus near the middle of the nucleoid until the point that two distinct foci can be seen (3). However, sometimes it appears that a single focus can move to a border of the nucleoid, separate into two foci and these can migrate asymmetrically toward both pole-proximal borders of the nucleoid prior to the cell division as described previously (6). Based on the cohesion model, 'duplication of an *oriC* focus' (6) does not mean 'replication of the *oriC* segment', but 'separation of cohesive *oriC* copies'. Thus the distinction is based on our experimental design in which replication was initiated at a known time by a temperature downshift in the *dnaA*ts strain, and then observation of separate foci was seen only at a much later time.

The second subject is translocation of replication apparatuses during chromosome replication. Paired replication apparatuses that are acting for bidirectional replication of the chromosomal DNA are first closely located with each other at the middle of the nucleoid. However these apparatuses separate and rapidly migrate to 1/4 and 3/4 cell positions at the time that replication forks are estimated to have progressed bidirectionally approximately one third of the distance between *oriC* and terminus. This model, named the 'translocating replication apparatuses model' (3; Fig. 1) or 'replication apparatuses translocation model', is based on analysis of the subcellular localization of SeqA-bound hemimethylated nascent DNA clusters (2, 3, 9, 11; for a review, see 1) and beta-clamp clusters (DNA sliding clump or DnaN) of DNA polymerase III holoenzyme (10). According to this model, copies of *oriC* and *oriC*-proximal genes that are replicated before the translocation of replication apparatuses might remain localized as a single focus in the middle of the nucleoid (Fig. 1B-b and c). On the other hand, other genes that are replicated after the translocation of replication apparatuses might be localized as a single focus around the replication apparatus localized at the 1/4 or 3/4 cell position until release from cohesion (Fig. 1B-e and f). Our previous results by FISH (7) are consistent with the above speculation. After release from cohesion, sister chromosomes are rearranged to form two separated nucleoids, in which *oriC* copies are localized near pole-proximal borders of the nucleoid(s) and *dif* copies are localized near mid-cell, that is, the division site (Fig. 1B-f and g; 6, 7). The migration of *oriC* copies toward pole-proximal borders of the nucleoid is presumably due to an uncharacterized mechanism, similar to the plasmid partition system (1, 5).

It is not so easy to determine the critical time of initiation of chromosome replication in exponentially growing cultures using the baby-machine technique in parallel to analysis of subcellular localization of fluorescent foci. Furthermore, some complicated assumptions are required for judging sister chromosome cohesion in statistical results on the subcellular localization of fluorescent foci in exponentially growing cultures in the steady state. We therefore analysed cells that were synchronized for initiation of replication in order to demonstrate the cohesion phenomenon (3, 11). More recently, we analysed exponentially growing cells that carried a cassette of multiple *lacO* sequences that was inserted near *oriC* of the chromosome, using the following two different methods; (i) FISH in fixed cells to detect *oriC* foci and (ii) fluorescence microscopy in living cells to detect LacI–GFP foci. In both the methods, approximately half of cells with a positive signal cells had one fluorescent focus and the

▶ *GUIDELINES*

Communications should be in the form of letters and should be brief and to the point. A single small Table or Figure may be included, as may a limited number of references (cited in the text by numbers, and listed in alphabetical order at the end of the letter). A short title (fewer than 50 characters) should be provided.

Approval for the publication rests with the Editor-in-Chief, who reserves the right to edit letters and/or to make a brief reply. Other interested persons may also be invited to reply. The Editors of ***Microbiology*** do not necessarily agree with the views expressed in ***Microbiology* Comment**.

Contributions should be addressed to the Editor-in-Chief via the Editorial Office.

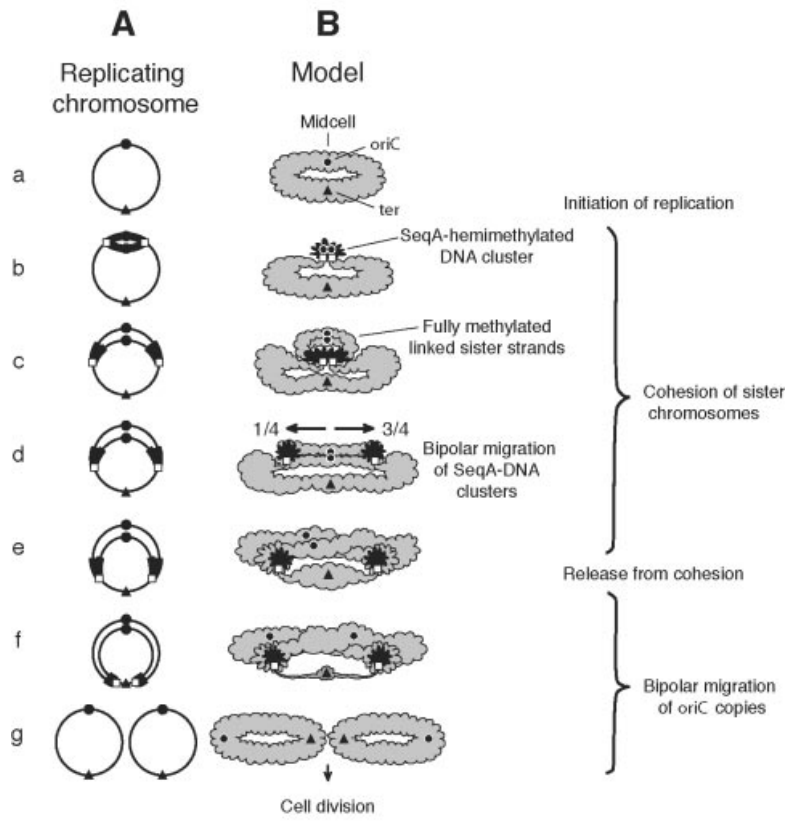

Fig. 1. A model: translocation of replication apparatuses and cohesion of sister chromosomes in *E. coli*. Fine lines and grey masses represent unreplicated or replicated fully methylated chromosomal DNA. Bold lines and black masses represent nascent hemimethylated DNA strands. Black circles, *oriC* sites; black triangles, terminus sites; open squares, replication apparatuses. Horizontal arrows in B-d represent rapid bidirectional migration of SeqA clusters from the mid-cell position to the 1/4 and 3/4 cell positions.

other half had two foci, even though flow cytometry indicated that the majority of newborn cells had one replicating chromosome with two *oriC* copies and that cells had four *oriC* copies prior to cell division (our unpublished data). This discrepancy can be explained by the sister chromosome cohesion model.

My third comment is about the roles of the MukFEB complex (12, 13). MukB seems to participate in sister chromosome cohesion (11). Interestingly, MukB–GFP molecules are recruited to clusters, although part of MukB–GFP molecules diffuse throughout the nucleoid (8). MukB–GFP forms two clusters at 1/4 and 3/4 cell positions prior to migration of SeqA–DNA clusters to the cell quarter positions in minimum glucose medium. After the migration of SeqA–DNA clusters, each MukB–GFP cluster is separated into two, resulting in four MukB–GFP clusters prior to the cell division (8). The cluster formation depends on the presence of both MukF and MukE, suggesting that MukFEB complex forms clusters. Our recent results on the localization of nascent DNA that was pulse-labelled with 5-bromo-2′-deoxyuridine indicate that MukB participates in the correct localization of replication forks at mid-cell or the 1/4 and 3/4 cell positions (our unpublished data). Thus, segregation of sister chromosomes depends on functions of specific gene products.

Sota Hiraga

Department of Radiation Genetics, Graduate School of Medicine, Kyoto University, Konoe-cho, Yoshida, Sakyo-ku, Kyoto 606-8501, Japan

Author for correspondence: Sota Hiraga. Tel: +81 75 753 4415. Fax: +81 75 753 4419.
e-mail: hiraga@rg1.rg.med.kyoto-u.ac.jp

1. Hiraga, S. (2000). Dynamic localization of bacterial and plasmid chromosomes. *Annu Rev Genet* **34**, 21–59.
2. Hiraga, S., Ichinose, C., Niki, H. & Yamazoe, M. (1998). Cell cycle-dependent duplication and bidirectional migration of SeqA-associated DNA-protein complexes in *E. coli*. *Mol Cell* **1**, 381–387.
3. Hiraga, S., Ichinose, C., Onogi, T., Niki, H. & Yamazoe, M. (2000). Bidirectional migration of SeqA-bound hemimethylated DNA clusters and pairing of *oriC* copies in *Escherichia coli*. *Genes Cells* **5**, 327–341.
4. Nanninga, N., Roos, M. & Woldringh, C. L. (2002). Models on stickiness of replicated *Escherichia coli oriC*. *Microbiology* **148**, 3327–3328.
5. Niki, H. & Hiraga, S. (1997). Subcellular distribution of actively partitioning F plasmid during the cell division cycle in *E. coli*. *Cell* **90**, 951–957.
6. Niki, H. & Hiraga, S. (1998). Polar localization of the replication origin and terminus in *Escherichia coli* nucleoids during chromosome partitioning. *Genes Dev* **12**, 1036–1045.
7. Niki, H., Yamaichi, Y. & Hiraga, S. (2000). Dynamic organization of chromosomal DNA in *Escherichia coli*. *Genes Dev* **14**, 212–223.
8. Ohsumi, K., Yamazoe, M. & Hiraga, S. (2001). Different localization of SeqA-bound nascent DNA clusters and MukF-MukE-MukB complex in *Escherichia coli* cells. *Mol Microbiol* **40**, 835–845.
9. Onogi, T., Niki, H., Yamazoe, M. & Hiraga, S. (1999). The assembly and migration of SeqA-Gfp fusion in living cells of *Escherichia coli*. *Mol Microbiol* **31**, 1775–1782.
10. Onogi T., Ohsumi, K., Katayama, T. & Hiraga, S. (2002). Replication-dependent recruitment of the β-subunit of DNA polymerase III from cytosolic spaces to replication forks in *Escherichia coli*. *J Bacteriol* **184**, 867–870.
11. Sunako, T., Onogi, T. & Hiraga, S. (2001). Sister chromosome cohesion of *Escherichia coli*. *Mol Microbiol* **42**, 1233–1241.
12. Yamanaka, K., Ogura, T., Niki, H. & Hiraga, S. (1996). Identification of two new genes, *mukE* and *mukF*, involved in chromosome partitioning in *Escherichia coli*. *Mol Gen Genet* **250**, 241–251.
13. Yamazoe, M., Onogi, T., Sunako, Y., Niki, H., Yamanaka, K., Ichimura, T. & Hiraga, S. (1999). Complex formation of MukB, MukE and MukF proteins involved in chromosome partitioning in *Escherichia coli*. *EMBO J* **18**, 5873–5884.

The minimal mobile element

Horizontal transfer of genes is an integral component of bacterial evolution and is particularly associated with processes related to environmental adaptation and virulence. The association of many host adaptive and virulence genes with mobile genetic elements such as transposases and bacteriophages reflects this. Presumably the association of the gene conferring some competitive advantage for the recipient strain facilitates the dissemination of the mobile element. The presence of larger elements such as 'pathogenicity islands' or 'islands of horizontal transfer' is also characteristic of some bacterial species. However, analysis of complete bacterial genomes suggests a previously unrecognized mechanism of mobilization that utilizes natural transformation and homologous recombination, and that is independent of transposases and other mobilization mechanisms.

When the genomes of related bacterial species are compared, sites in which different genes are present are identifiable, located between highly conserved flanking genes. For example, the neisserial *pheS* and *pheT* genes

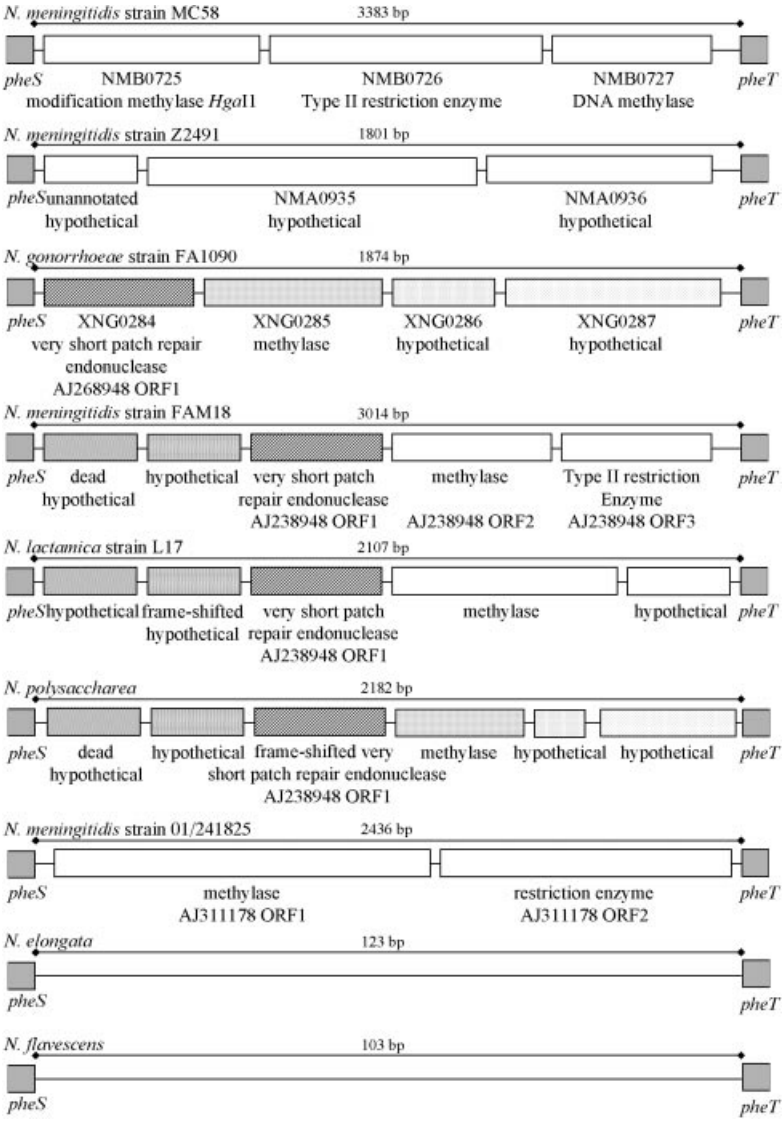

Fig. 1. Nine different intergenic regions between *pheS* and *pheT* identified in the *Neisseria* spp. NMA and NMB numbers refer to the respective annotations (7, 8). *N. gonorrhoeae* strain FA1090 (5) XNG numbers refer to our own annotation: associated coding sequences available on request. *N. meningitidis* strain FAM18: these sequence data were produced by the *N. meningitidis* serogroup C strain FAM18 Sequencing Group at the Sanger Institute and can be obtained from ftp://ftp.sanger.ac.uk/pub/pathogens/nm/. Described coding sequences are available on request. *N. meningitidis* strain 01/241825 sequence is identical to GenBank AJ311178. Other sequences submitted to GenBank: *N. lactamica*, AF542177; *N. polysaccharea*, AF542178; *N. elongata*, AF542173; and *N. flavescens*, AF542175. Homologous coding sequences are indicated. 'Dead' indicates the presence of multiple frame-shifts and/or in-frame termination codons, while 'frame-shifted' indicates a single frame-shift.

(encoding phenylalanyl-tRNA synthetase, alpha and beta chains, respectively) flank two genes of unknown function (NMA0935 and NMA0936) encoding proteins of 354 and 172 aa in length in *Neisseria meningitidis* strain Z2491 (7), and genes coding for two different restriction/modification systems in *N. meningitidis* strain MC58 (8) and *Neisseria gonorrhoeae* strain FA1090 (5) (Fig. 1). These genes are adjacent in both *Helicobacter pylori* strains (1, 9) and *Haemophilus influenzae* (3). In *N. gonorrhoeae* strain FA1090 *glyQ* and *glyS* (encoding glycyl-tRNA synthetase, alpha and beta chains, respectively) are separated by a short intergenic region of 116 nt. However, in the meningococci *glyQ* and *glyS* flank a putative phase-variable gene of unknown function, the sequenced *H. pylori* strains *glyQ* and *glyS* flank genes encoding a nickel-resistance type cation efflux system, and this site in *H. influenza* strain Rd contains two genes of unknown function.

To extend this observation, the regions between the *pheS* and *pheT* genes were amplified from a small collection of unrelated neisserial strains and species (Table 1). PCR products were digested with *Dpn*II and *Xba*I, and product sizes and digest patterns were compared on agarose gels. The amplified and genome sequence-derived sequences were sorted into a number of similar groups by these tests (Table 1). An example of each type that differed in size or digest pattern was sequenced (previously unavailable sequences GenBank accession nos: *Neisseria cinerea*, AF542174; *Neisseria elongata*, AF542173; *Neisseria flavescens*, AF542175; *Neisseria lactamica*, AF542177; *Neisseria polysaccharea* AF542178; and *N. subflava*, AF542176), and after accounting for altered digest patterns due to single base polymorphisms, a total of 9 major variants were found between these two *phe* genes (Fig. 1). Previous work on this region in *N. meningitidis* identified a region (AF238948) similar to that in sequence strain FAM18 (2) (Fig. 1). One insert, sequenced from *N. meningitidis* strain 01/241825, was identical to a previously submitted sequence that did not include the *pheS*/*pheT* flanking sequences (AJ311178). Two short intergenic sequences were identified: one of 123 bp was present in *N. elongata* and *N. subflava*, while the other, sequenced from *N. flavescens* and *N. cinerea* was 103 bp in length. These two short intergenic regions are not homologous and do not include any probable coding regions. The variable inserted sequences are predominantly restriction/modification system genes, which are believed to confer self-selectable phenotypes (4, 6), which is consistent with the proposed model.

Additional examples of functionally related, metabolic, gene pairs flanking insertion sites containing different genes were sought by comparing the completed neisserial genome sequences. This revealed five more such sites: the alpha and beta subunits of ribonucleoside-diphosphate reductase, *nrdA* and *nrdB*; the small and large subunits of carbamoyl-phosphate synthase, *carA* and *carB*; two components of an Fe–S cluster associated cysteine metabolism system, *nifU* and *nifS*; LPS biosynthetic genes, *rfaD* and *rfaE*; and the L and M subunits of NADH dehydrogenase, *nuoL* and *nuoM*. This list does not include all genes likely to be mobile by this mechanism within this and related species, the degree of sequence and gene-order conservation being differentially variable between different pairs of strains and species.

In the case of *pheS* and *pheT* sequence comparisons of the ends of these genes in the different insert types showed results consistent with recombination, with divergent bases present in different combinations indicating a mosaic structure associated with recombination, that clustered differently between the two genes (Fig. 2). This is consistent with a

Table 1. *Neisseria* spp. grouped by *pheS–pheT* regions based on PCR amplification and restriction analysis

	Species	Strain	Serogroup	Source
MC58-like				
	N. meningitidis	MC58	B	Sequence strain
	N. meningitidis	2736	X	MRUPHL*
	N. meningitidis	01/241809	X	MRUPHL*
	N. meningitidis	01/241693	X	MRUPHL*
Z2491-like				
	N. meningitidis	Z2491	A	Sequence strain
	N. meningitidis	2730	A	WHORL†
	N. meningitidis	2731	A	WHORL†
	N. meningitidis	2746	W135	WHORL†
	N. meningitidis	2749	Y	WHORL†
	N. meningitidis	2750	Z	WHORL†
	N. meningitidis	2762	A	WHORL†
	N. meningitidis	2765	A	WHORL†
	N. meningitidis	2779	A	WHORL†
	N. meningitidis	2783	A	WHORL†
	N. meningitidis	2798	B	WHORL†
	N. meningitidis	2803	A	WHORL†
	N. meningitidis	2810	A	WHORL†
	N. meningitidis	2812	A	WHORL†
	N. meningitidis	2814	B	WHORL†
	N. meningitidis	2815	B	WHORL†
	N. meningitidis	BZ133	B	WHORL†
	N. meningitidis	98/250521	H	MRUPHL*
	N. meningitidis	00/240794	D	MRUPHL*
	N. meningitidis	97/282675	H	MRUPHL*
	N. meningitidis	97/252572	H	MRUPHL*
	N. meningitidis	00/240868	Z	MRUPHL*
	N. meningitidis	01/241422	Z	MRUPHL*
	N. meningitidis	01/241471	Z	MRUPHL*
FAM18-like				
	N. meningitidis	FAM18	C	Sequence strain
	N. meningitidis	2744	C	WHORL†
L17-like				
	N. lactamica	L17	n/a	R. Moxon‡
	N. lactamica	L13	n/a	R. Moxon‡
	N. lactamica	L18	n/a	R. Moxon‡
	N. lactamica	L19	n/a	R. Moxon‡
	N. lactamica	L20	n/a	R. Moxon‡
N. polysaccharea-like				
	N. polysaccharea		n/a	D. Stephens§
FA1090-like				
	N. gonorrhoeae	FA1090	n/a	Sequence strain
	N. gonorrhoeae	FA19	n/a	W. Shafer§
	N. gonorrhoeae	F62	n/a	W. Shafer§
	N. gonorrhoeae	FA48	n/a	W. Shafer§
	N. gonorrhoeae	FA102	n/a	W. Shafer§
	N. gonorrhoeae	FA171	n/a	W. Shafer§
	N. gonorrhoeae	BR87	n/a	W. Shafer§
	N. gonorrhoeae	EU50	n/a	W. Shafer§
	N. gonorrhoeae	EU75	n/a	W. Shafer§
	N. gonorrhoeae	MS11	n/a	R. Rest¶
	N. gonorrhoeae	150002	n/a	R. Moxon‡
	N. gonorrhoeae	15253	n/a	R. Moxon‡
	N. gonorrhoeae	179008	n/a	R. Moxon‡
	N. gonorrhoeae	27706	n/a	UKNRL#

Table 1. (*cont.*)

	Species	Strain	Serogroup	Source
	N. gonorrhoeae	27728	n/a	UKNRL#
	N. gonorrhoeae	27806	n/a	UKNRL#
	N. gonorrhoeae	27886	n/a	UKNRL#
	N. gonorrhoeae	27921	n/a	UKNRL#
	N. gonorrhoeae	28197	n/a	UKNRL#
	N. gonorrhoeae	28252	n/a	UKNRL#
	N. gonorrhoeae	28386	n/a	UKNRL#
	N. gonorrhoeae	28480	n/a	UKNRL#
	N. gonorrhoeae	28516	n/a	UKNRL#
	N. gonorrhoeae	28539	n/a	UKNRL#
	N. gonorrhoeae	28622	n/a	UKNRL#
	N. gonorrhoeae	28839	n/a	UKNRL#
	N. gonorrhoeae	28897	n/a	UKNRL#
	N. gonorrhoeae	28962	n/a	UKNRL#
	N. gonorrhoeae	29214	n/a	UKNRL#
	N. gonorrhoeae	29528	n/a	UKNRL#
	N. lactamica		n/a	D. Stephens§
29E-like				
	N. meningitidis	01/241825	29E	MRUPHL*
N. elongata-like				
	N. elongata		n/a	D. Stephens§
	N. subflava		n/a	D. Stephens§
	N. meningitidis	2771	B	WHORL†
	N. meningitidis	2804	B	WHORL†

* Meningococcal Reference Unit, Public Health Laboratory, UK

† WHO Reference Laboratory

‡ University of Oxford, UK

§ Emory University, USA

¶ Drexel University, USA

UK National Reference Laboratory

model in which recombination leading to the exchange of sections of these genes is being driven by the phenotypes of the intervening gene(s). While one might predict that the changes in the flanking genes are normally neutral, it is also possible that minor changes in the functions of these genes may sometimes occur.

The essential feature of each case is that the flanking genes are highly conserved metabolic genes that have a conserved order within a transcriptional unit. In the *pheS–pheT* example, the inserted genes are in the same orientation as the conserved flanking genes, which may indicate their inclusion in the transcriptional unit. Insertions that generate polar effects in such a way as to disrupt the function of the flanking genes would probably be selected against. If a gene becomes inserted into such a site it has acquired highly conserved flanking sequences that can act as substrates for homologous recombination between other strains and, depending upon the conservation of the genes, other species. While there may be no specific feature targeting insertion into such sites, once a niche-adaptive or 'selfish' gene has inserted into such a site it will then gain the capacity to be more readily transferred. This probably represents the minimal mobile element and should be considered in analyses of horizontal gene transfer.

Acknowledgements

N.J.S. is supported by a Wellcome Trust Advanced Research Fellowship.

Nigel J. Saunders and Lori A. S. Snyder

Bacterial Pathogenesis and Functional Genomics Group, Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford. OX1 3RE.

Author for correspondence: Nigel J. Saunders. Tel: (+44) 1865 275521. Fax: (+44) 1865 275515. e-mail: saunders@molbiol.ox.ac.uk

1. Alm, R. A., Ling, L. S., Moir, D. T. & 20 other authors. (1999). Genomic-sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*. *Nature* **397**, 176–180.
2. Claus, H., Friedrich, A., Frosch, M. & Vogel, U. (2000). Differential distribution of novel restriction-modification systems in clonal lineages of *Neisseria meningitidis*. *J Bacteriol* **182**, 1296–1303.
3. Fleischmann, R. D., Adams, M. D., White, O. & 37 other authors (1995). Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* **269**, 496–512.
4. Kobayashi, I. (2001). Behavior of restriction-modification systems as selfish mobile elements and their impact on genome evolution. *Nucleic Acids Res* **29**, 3742–3756.
5. Lewis, L., Gillaspy, A. F., McLaughlin, R. E. & 22 other authors. (1997). ftp://ftp.genome.ou.edu/pub/gono/gono-2k.fa; GenBank No AE004969.
6. Nakayama, Y. & Kobayashi, L. (1998). Restrictions-modification gene complexes as selfish gene entities: roles of a regulatory system in their establishment, maintenance, and apoptotic mutual exclusion. *Proc Natl Acad Sci U S A* **95**, 6442–6447.
7. Parkhill, J., Achtman, M., James, K. D. & 25 other authors (2000). Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491. *Nature* **404**, 502–506.
8. Tettelin, H., Saunders, N. J., Heidelberg, J. & 39

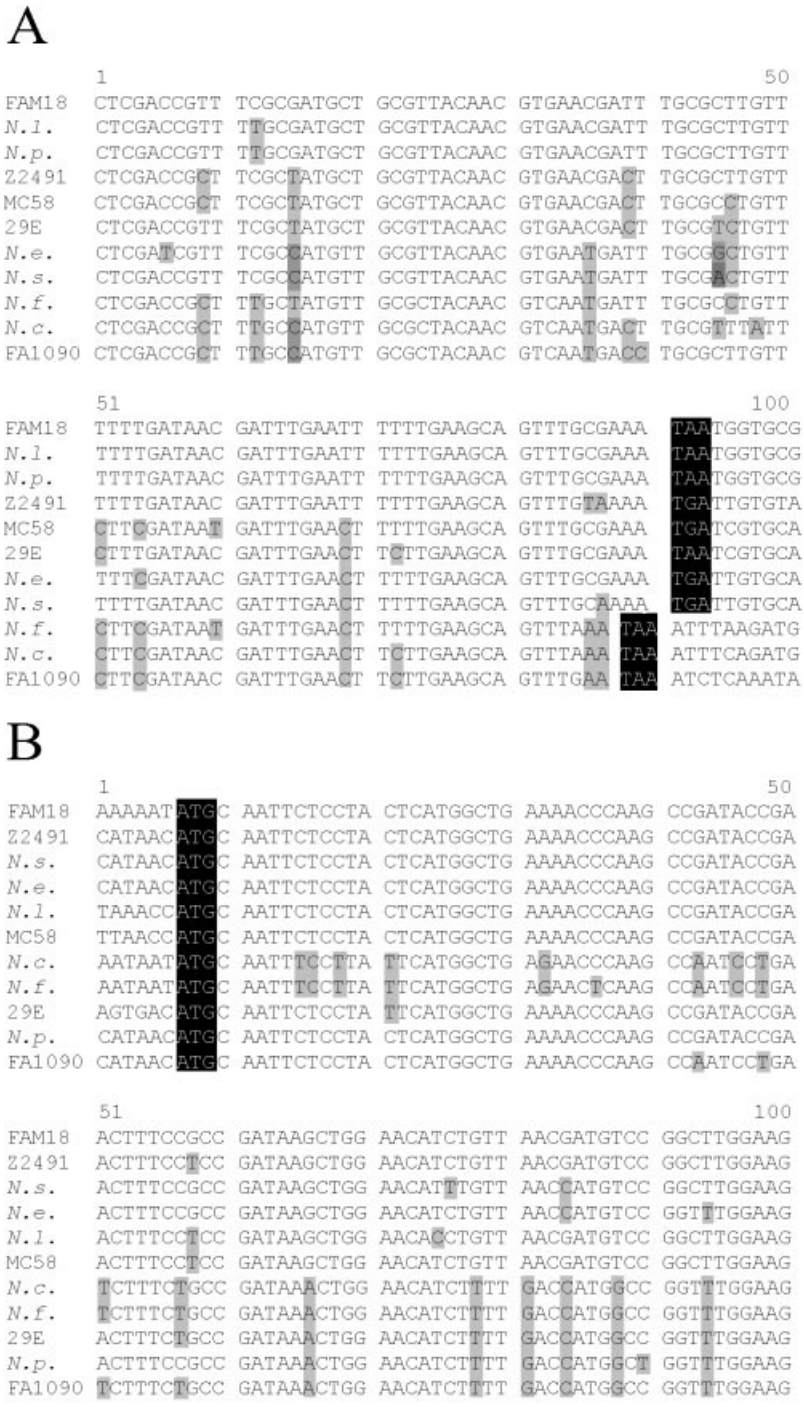

Fig. 2. Regions from 11 strains with different genes between *pheS* and *pheT*. Divergence within the coding region of the flanking genes of the minimal mobile element is indicative of the recombination-generated mosaic structure of these regions. Note that the sequences from these strains have clustered differently at the two ends of this region. A. Region containing the termination codon of *pheS* (white letters). B. Region containing the initiation codon of *pheT* (white letters). FAM18: *N. meningitidis* strain FAM18, sequence from the Sanger Centre; *N.l.*, *N. lactamica*, sequence submitted to GenBank (AF542177); *N.p.*, *N. polysaccharea*, sequence submitted to GenBank (AF542178); Z2491, *N. meningitidis* strain Z2491 (7); MC58, *N. meningitidis* strain MC58 (8); 29E, *N. meningitidis* strain 01/241825, identical *pheS/pheT* region to AJ311178; *N.e.*, *N. elongata*, sequence submitted to GenBank (AF542173); *N.s.*, *N. subflava*, sequence submitted to GenBank (AF542176); *N.f.*, *N. flavescens*, sequence submitted to GenBank (AF542175); *N.c.*, *N. cinerea*, sequence submitted to GenBank (AF542174); FA1090, *N. gonorrhoeae* strain FA1090 (5).

other authors (2000). Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58. *Science* **287**, 1809–1815.
9. Tomb, J. F., White, O., Kerlavage, A. R. & 39 other authors (1997). The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *Nature* **388**, 539–547.

Sequence similarity between the channel-forming domains of voltage-gated ion channel proteins and the C-terminal domains of secondary carriers of the major facilitator superfamily

Integral membrane transport proteins are believed to have evolved as a distinct class of proteins independently of other protein types such as enzymes, structural proteins and regulatory proteins (14). The ancestral elements were small channel-forming oligomeric proteins or peptides analogous in structure and function to the present-day potassium channel of *Streptomyces lividans* KscA (4) and the toxic bacteriocins (5, 12). Genes encoding these one or two transmembrane segment (TMS) peptides duplicated or were fused to other genetic elements to give larger and functionally more complex transport systems (9). Superimposition of catalytic proteins onto channel-forming proteins or carriers gave rise to primary active translocators (14, 17). The specific pathways taken for the evolution of several families of transporters have been elucidated (16). The proposed pathway taken for the evolutionary appearance of different types of transporters (14) is:

channel peptides→channel proteins

→secondary carriers → primary active transporters
→secondary carriers → group translocators

Of the over 300 families of transport proteins currently recognized, three of these predominate (15). These families are the voltage-gated ion channel (VIC) superfamily (TC #1.A.1; 9), the major facilitator superfamily (MFS) of secondary carriers (TC #2.A.1; 10, 18) and the ATP-binding cassette (ABC) superfamily of primary active transporters (1). No previous publication has provided evidence for a structural, functional or evolutionary connection between these three superfamilies. In this article we provide the first evidence for a connection between the VIC superfamily and the MFS.

MFS proteins arose by an internal gene duplication event in which a primordial 6 TMS-encoding element gave rise to a 12 TMS-encoding element (see 10 and references cited therein). Most members of the MFS exhibit 12

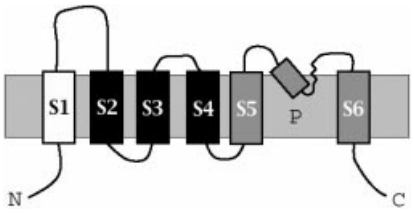

Fig. 1. Model of the 6 TMS channel unit of many VIC family members including the Na^+ channel protein of *Bacillus halodurans*, NaChBac. S1 (white) is the sequence involved in insertion of the protein into the lipid bilayer membrane. S2, S3 and S4 (black) jointly form the voltage sensor; S5, the P-loop and S6 (grey) are components of the pore-forming structure. The transmembrane helices and the short pore (P) helix are displayed as cylinders.

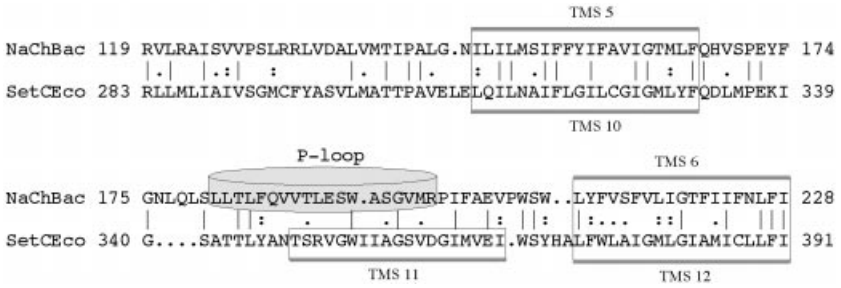

Fig. 2. Alignment of the channel-forming domain of NaChBac with the last three TMSs (TMSs 10–12) in the *E. coli* SetC protein. Positions of TMSs 5 and 6 and the P-loop in NaChBac are as presented in ref. 11. Positions of the last three TMSs in the 12 TMS SetC protein are based on hydropathy profiles but correlate with the positions of established TMSs in other MFS permeases (see 10, 18 and cited references). TMSs are boxed and labelled by number while the P-loop is so labelled. Identities are indicated by vertical lines between identical residues; close and distant similarities are indicated by colons and dots, respectively. The alignment was generated using the GAP program (3).

TMSs although some have 14 due to insertion of the central loop (between TMSs 6 and 7 in the 12 TMS proteins) into the membrane. Further, a 6 TMS MFS homologue is encoded within the genome of *Bacillus subtilis* (YitZ; 19). One family within the MFS, the sugar efflux transporter (SET) family (TC #2.A.1.20), which like other MFS families, includes members with two homologous halves, seems to be primarily concerned with sugar efflux (7, 8). By contrast, the prevalent 6 TMS K^+ channels of the VIC superfamily arose by fusion of a 4 TMS-encoding genetic element (the voltage sensor domain) to a 2 TMS-encoding element (the channel domain) (see Fig. 1; 9). The latter 2 TMSs separated by the P-loop comprise the common element of all members of the VIC superfamily. Gene fusion and gene duplication events have given rise to members of the VIC superfamily with 2, 4, 6, 8, 12, and 24 TMSs per polypeptide chain (16). However, in all cases the basic structure is equivalent to a simple tetramer (4).

Recently, the first voltage-gated Na^+ channel in a prokaryote was identified and characterized (11). This 6 TMS protein, NaChBac from *Bacillus halodurans*, most resembles in sequence animal Ca^{2+} channels and is inhibited by classical calcium blockers (11). However, it has the 6 TMS topology that is characteristic of K^+ channels of the VIC superfamily. While investigating NaChBac, we were surprised to find that the channel-forming domain of this bacterial protein exhibits striking sequence similarity to the C-terminal regions of well-characterized pmf-driven sugar efflux carriers of *E. coli*, SetA, SetB and SetC (7, 8). The degree of sequence similarity between NaChBac of the VIC superfamily and the Set paralogues of the MFS is sufficient to strongly suggest a common evolutionary origin.

Fig. 2 shows a sequence alignment of the C-terminal parts of NaChBac and SetCEco. This alignment of 109 positions exhibited

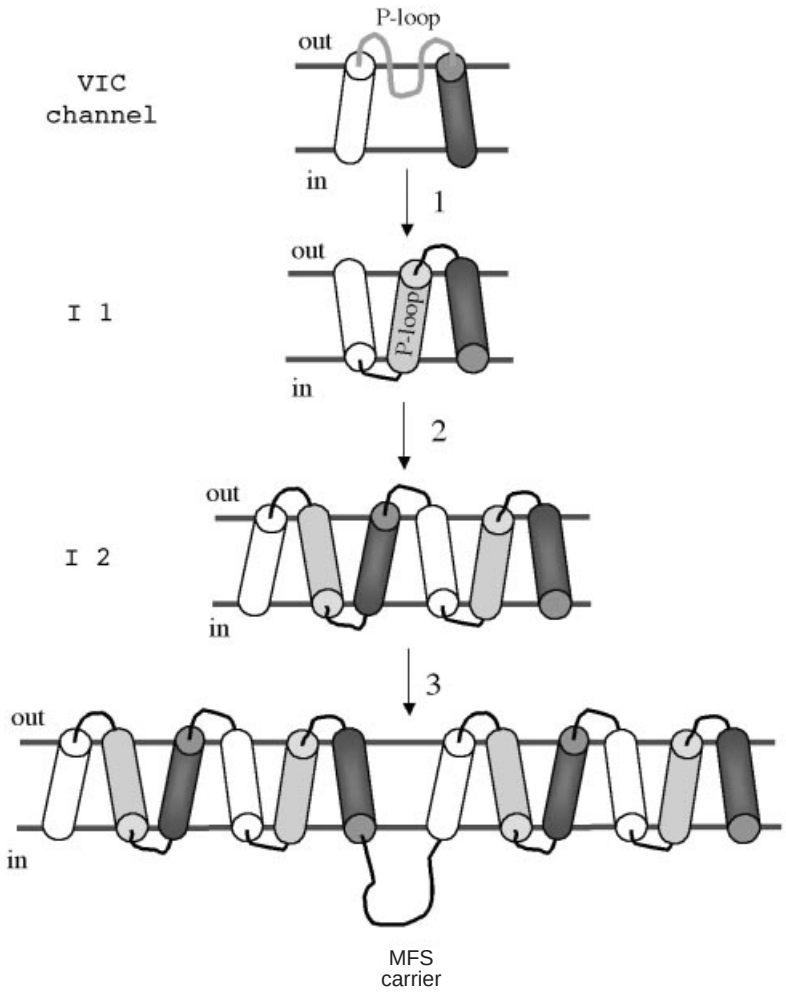

Fig. 3. Proposed pathway for the appearance of a monomeric 12 TMS carrier of the MFS from a primordial tetrameric 2 TMS channel-forming protein of the voltage-gated ion channel (VIC) superfamily involving two intermediate structures, I1 and I2. Note that the two proposed intragenic duplication events shown in steps 2 and 3, give rise to a pseudotetrameric structure, superficially resembling the established tetrameric VIC family structure (4). The evidence for step 1 is presented in this report. While no evidence currently supports the duplication event depicted in step 2, that shown in step 3 is well established (see 10 and references cited therein).

34% identity, 44% similarity and a comparison score using the GAP program (500 random shuffles) of 12·8 SD (3). The probability that this degree of sequence similarity could have occurred by chance is less than 1 in 10^{20} (2, 13). As shown in Fig. 2, putative TMS 12 in SetC (corresponding to the established 12th TMS in many MFS permeases) corresponds in position to TMS 6 in NaChBac. The hydrophilic C-terminal tails of NaChBac and SetC are cytoplasmic, and both TMSs thus have the same orientation in the membrane (out to in). TMS 11 in the aligned sequence of SetC overlaps the residues in NaChBac that comprise the P-loop (Fig. 2). While the fairly hydrophobic P-loop dips into the membrane and re-emerges on the same side of the membrane in all VIC family members, this region is erroneously predicted to be a TMS using several topology predicting computer programs. TMS 11 in SetC, on the other hand, traverses the membrane, based on the established topologies of many MFS permeases (see 10, 18 for reviews). TMS 5 in NaChBac (11) corresponds in position precisely to the predicted TMS 10 in SetC. TMS 5 in NaChBac is of opposite orientation in the membrane compared to TMS 10 of the MFS permeases. It is interesting to note that the regions exhibiting sequence similarity between NaChBac and SetCEco are among the most conserved regions within their respective families. Intragenic duplication of primordial genes encoding membrane proteins with odd numbers of TMSs is known to give rise to homologous integral membrane domains exhibiting opposite orientation in the membrane (6, 13, 20).

Multiple alignments of several SET family proteins of the MFS with several prokaryotic homologues of NaChBac revealed that among the residues conserved between SetCEco and NaChBac (Fig. 2) are many that are well-conserved between both families (data not shown). These observations strengthen the significance of the sequence similarity shown in Fig. 2. They suggest that the evolutionary pathway for the appearance of MFS carriers from a primordial 2 TMS VIC family channel occurred via the pathway shown in Fig. 3. In the proposed pathway, the 2 TMS channel-forming domain of the VIC family gave rise to a primordial 3 TMS element that first underwent an intragenic duplication event to give a 6 TMS protein with the two homologous halves of opposite orientation in the membrane, followed by a later intragenic duplication event to give a 12 TMS protein with these two homologous halves of the same orientation in the membrane. Thus, the origin of the 6 TMS topology for VIC family members, shown in Fig. 1, which resulted from the fusion of two dissimilar elements, and the origin of the 6 TMS repeat unit in the 12 TMS MFS permeases is different. Only S5, the P-loop and S6 are proposed to have been used for construction of the MFS carriers. The simple postulate presented in Fig. 3 should be subject to confirmation or refutation by a combination of experimental and bioinformatic approaches.

Acknowledgements

This work was supported by NIH grant GM64368. We thank Mary Beth Hiller for her assistance in the preparation of this manuscript.

Rikki N. Hvorup and Milton H. Saier, Jr

Division of Biology, University of California at San Diego, La Jolla, CA 92093–0116, USA

Author for correspondence: Milton H. Saier, Jr. Tel: +1 858 534 4084. Fax: +1 858 534 7108.
e-mail: msaier@ucsd.edu

1. Dassa, E. & Bouige, P. (2001). The ABC of ABCs: a phylogenetic and functional classification of ABC systems in living organisms. *Res Microbiol* **152**, 211–229.
2. Dayhoff, M. O., Barker, W. C. & Hunt, L. T. (1983). Establishing homologies in protein sequences. *Methods Enzymol* **91**, 524–545.
3. Devereux, J., Haeberli, P. & Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res* **12**, 387–395.
4. Doyle, D. A., Morais Cabral, J., Pfuetzner, R. A., Kuo, A., Gulbis, J. M., Cohen, S. L., Chait, B. T. & MacKinnon, R. (1998). The structure of the potassium channel: molecular basis of K^+ conduction and selectivity. *Science* **280**, 69–77.
5. Ennahar, S., Sashihara, T., Sonomoto, K. & Ishizaki, A. (2000). Class IIa bacteriocins: biosynthesis, structure and activity. *FEMS Microbiol Rev* **24**, 85–106.
6. Jack, D. L., Yang, N. M. & Saier, M. H., Jr. (2001). The drug/metabolite transporter superfamily. *Eur J Biochem* **268**, 3620–3639.
7. Liu, J. Y., Miller, P. F., Gosink, M. & Olson, E. R. (1999). The identification of a new family of sugar efflux pumps in *Escherichia coli*. *Mol Microbiol* **31**, 1845–1851.
8. Liu, J. Y., Miller, P. F., Willard, J. & Olson, E. R. (1999). Functional and biochemical characterization of *Escherichia coli* sugar efflux transporters. *J Biol Chem* **274**, 22977–22984.
9. Nelson, R. D., Kuan, G., Saier, M. H., Jr & Montal, M. (1999). Modular assembly of voltage-gated channel proteins: a sequence analysis and phylogenetic study. *J Mol Microbiol Biotechnol* **1**, 281–287.
10. Pao, S. S., Paulsen, I. T. & Saier, M. H., Jr (1998). The major facilitator superfamily. *Microbiol Mol Biol Rev* **62**, 1–32.
11. Ren, D., Navarro, B., Xu, H., Yue, L., Shi, Q. & Clapham, D. E. (2001). A prokaryotic voltage-gated sodium channel. *Science* **294**, 2372–2375.
12. Riley, M. A. (1998). Molecular mechanisms of bacteriocin evolution. *Annu Rev Genet* **32**, 255–278.
13. Saier, M. H., Jr (1994). Computer-aided analyses of transport protein sequences: gleaning evidence concerning function, structure, biogenesis, and evolution. *Microbiol Rev* **58**, 71–93.
14. Saier, M. H., Jr (2000a). Vectorial metabolism and the evolution of transport systems. *J Bacteriol* **182**, 5029–5035.
15. Saier, M. H., Jr (2000b). A functional-phylogenetic classification system for transmembrane solute transporters. *Microbiol Mol Biol Rev* **64**, 354–411.
16. Saier, M. H., Jr (2001). Evolution of transport proteins. In *Genetic Engineering: Principles and Methods*, vol. 23, pp. 1–10. Edited by J. K. Setlow. New York: Kluwer Academic/Plenum.
17. Saier, M. H., Jr & Tseng, T.-T. (1999). Evolutionary origins of transmembrane transport systems. In *Transport of Molecules Across Microbial Membranes* (Society for General Microbiology Symposium no. 58), pp. 252–274. Edited by J. K. Broome-Smith, S. Baumberg, C. J. Stirling and F. B. Ward. Cambridge: Cambridge University Press.
18. Saier, M. H., Jr, Beatty, J. T., Goffeau, A. & 11 other authors (1999). The major facilitator superfamily. *J Mol Microbiol Biotechnol* **1**, 257–279.
19. Saier, M. H., Jr, Goldman, S. R., Maile, R. R., Moreno, M. S., Weyler, W., Yang, N. & Paulsen, I. T. (2002). Transport capabilities encoded within the *Bacillus subtilis* genome. *J Mol Microbiol Biotechnol* **4**, 37–67.
20. von Heijne, G. (1999). Recent advances in the understanding of membrane protein assembly and structure. *Q Rev Biophys* **32**, 285–307.

Quorum sensing in context: out of molecular biology and into microbial ecology

Research on intercellular signalling in bacteria has undergone something of a renaissance over the past decade. In part, this can be attributed to the appeal of the 'quorum sensing' hypothesis, which suggests that the selective advantage driving the establishment and conservation of intercellular signalling mechanisms (ISMs) rests in the ability of intercellular signals to reflect cell density (*5*). Largely, however, the explosion in published investigations on ISMs reflects the finding that the expression of virulence factors in a handful of agriculturally or medically significant plant or animal pathogens is dependent on the synthesis of, and response to, diffusible 'signalling' metabolites. At the centre of all this research attention is the gene expression mechanism mediated by *N*-acyl-L-homoserine lactone (AHL), which generically involves the action of an AHL synthase protein (LuxI homologue) and an AHL-responsive transcriptional activator protein (LuxR homologue). The accumulation of AHL produced by LuxI homologues causes LuxR homologues to direct the transcription of specific 'structural' genes. Examples of structural genes regulated in this way include those encoding elastase production in the opportunistic human pathogen *Pseudomonas aeruginosa* and cellulase production in the plant pathogen *Erwinia carotovora* (*5*).

Because the interest in AHL-mediated gene expression has largely arisen from the potential opportunities it offers to control undesirable bacterial activities, many questions relating to the molecular biology and biochemistry of the mechanism have been answered. For example, we now know the substrates required by LuxI homologues to synthesize AHLs and that AHLs interact directly with the N-terminal domain of LuxR homologues, thereby inducing multimeric changes that culminate in DNA binding and recruitment of RNA polymerase to specific promoters with dyad symmetry (*5*). In contrast, fundamental questions pertaining to the ecology and evolution of AHL-mediated gene

Table 1. Phylogenetic association of known AHL systems

Class	Order	No. families with AHL system/ total families	No. genera with AHL system/ total genera	No. species with AHL system*
Alpha *Proteobacteria*	Rhodospirillales	0/2	0/24	0
	Rickettsiales	0/3	0/17	0
	Rhodobacterales	1/1	2/23	2
	Sphingomonadales	0/1	0/9	0
	Caulobacterales	0/1	0/4	0
	Rhizobiales	3/10	5/56	8
	All orders (6)	4/18 (22%)	7/133 (5%)	
Beta *Proteobacteria*	Burkholderiales	2/5	2/33	3
	Hydrogenophilales	0/1	0/2	0
	Methylophilales	0/1	0/3	0
	Neisseriales	1/1	1/14	1
	Nitrosomonadales	1/3	1/4	1
	Rhodocyclales	0/1	0/8	0
	All orders (6)	4/12 (33%)	4/56 (7%)	
Gamma *Proteobacteria*	Chromatiales	0/2	0/30	0
	Acidithiobacillales	0/2	0/2	0
	Zanthomonadales	0/1	0/9	0
	Cardiobacteriales	0/1	0/3	0
	Thiotrichales	0/3	0/15	0
	Legionellales	0/2	0/3	0
	Methylococcales	0/1	0/6	0
	Oceanspirillales	0/2	0/12	0
	Pseudomonadales	1/2	1/28	8
	Altermonadales	0/1	0/12	0
	Vibrionales	1/1	2/6	3
	Aeromonadales	1/2	1/7	2
	Enterobacteriales	1/1	6/41	14
	Pasteurellales	0/1	0/6	0
	All orders (14)	4/22 (18%)	10/180 (5%)	
Delta *Proteobacteria*	All orders (7)	0/18	0/55	0
Epsilon *Proteobacteria*	All orders (1)	0/2	0/6	0
Proteobacteria	All orders (34)	12/72 (17%)	21/438 (4%)	42

* Note that lists of bacterial species with AHL systems are notoriously difficult to keep up-to-date.

expression are only just beginning to be asked. For example, only one hypothesis (quorum sensing) describing the selective advantage of AHL-mediated gene expression is widely known (3). Surprisingly, the quorum sensing hypothesis has never been tested in the environment. Indeed, the role of the AHL ISM in mediating interactions between bacteria and their biotic and abiotic environment globally is poorly understood, if ever considered.

As a case in point, we are still unaware of the taxonomic and environmental distribution of AHL-mediated gene expression systems amongst bacteria. Statements suggesting that AHL-mediated gene expression is *a widespread phenomenon in Gram-negative bacteria* often appear in reviews (4) or introductions on the subject (and certainly in grant applications) and yet a simple detailed survey of its distribution amongst bacterial taxa has never appeared in the literature.

A cursory glance at the distribution of known AHL-producing bacterial taxa (Table 1) listed in Bergey's Manual of Systematic Bacteriology (1) reveals that they are constrained to three out of five classes (alpha, beta and gamma *Proteobacteria*) in 1 out of 20 bacterial phyla (*Proteobacteria*). This is generally appreciated. What you don't often hear is that within the alpha *Proteobacteria* only 7 out of 133 (5·3%) genera harbour known AHL-producing species, within the beta *Proteobacteria* only 4 out of 56 (7·1%) genera harbour known AHL-producing species and within the gamma *Proteobacteria* only 10 out of 180 (5·5%) genera harbour known AHL-producing species. That is, 21 bacterial genera (2·2 % of the total number of bacterial genera) contain species known to produce AHLs. At the species level the percentage of AHL producers drops to a fraction of a percent.

We believe this reflects two things. The first is that a comprehensive survey of bacterial species for AHL production has never been conducted. The beta *Proteobacteria* seem to be suffering especial neglect in this regard. Even so, because no AHL-producing species have ever been found outside of the *Proteobacteria* (leading us to assume they will not be found), the true percentage of AHL-producing genera is not likely to deviate far from that based on our current yet incomplete dataset. Note that this assumption ignores the possibility that a large pool of anaerobic or uncultured AHL producers or AHL-responsive non-producers (such as *Escherichia coli*) is lurking undetected by conventional bioassays. Secondly, the above inspection reflects

the fact that AHL production is restricted to a limited number of *Proteobacteria*, let alone Gram-negative bacteria, and should therefore not be considered widespread. Phylogenetic analyses of LuxR and LuxI homologues suggest that these genes were present in the last common ancestor of the *Proteobacteria* (2) and have therefore been discarded by the vast majority of descendants over evolutionary time. These assertions foster critical appraisal of the ecological significance of AHL-mediated gene expression.

Appraising the taxonomic distribution of AHL-mediated gene expression is only the first step in assessing the significance of this mechanism both as a survival strategy and in maintaining ecosystem function. We must also measure the abundance and distribution of the relevant taxonomic groups in the environment and decipher through *in situ* experiments what ecological interactions are mediated by the AHL ISM. Finally, there is a need to establish whether or not all the other proposed bacterial intercellular signalling molecules (AI-2, PAME, peptides, butyrolactones, quinolones, cyclic dipeptides) serve an ecological function analogous to that of AHLs. The widespread use of the term *quorum sensing* suggests that many researchers assume all ISMs do serve the same function. We believe that hypotheses describing the ecological function of these signalling systems rarely receive critical appraisal and even less frequently are tested experimentally. This serves as a final illustration of the need for the field to appraise AHL-mediated gene expression, and indeed intercellular signalling generally from an ecological and evolutionary perspective.

Mike Manefield and Sarah L. Turner

CEH Oxford, Mansfield Road, Oxford OX1 3SR, UK

Author for correspondence: Mike Manefield. Tel: +44 1865 281 630. Fax: +44 1865 281 696.
e-mail: mjma@ceh.ac.uk

1. Garrity, G. M., Winters, M. & Searles. D. B. (2001). *Bergey's Manual of Systematic Bacteriology*, 2nd edn. New York: Springer.
2. Gray, K. M. & Garey J. R. (2001). The evolution of bacterial LuxI and LuxR quorum sensing regulators. *Microbiology* **147**, 2379–2387.
3. Redfield, R. (2002). Is quorum sensing a side effect of diffusion sensing? *Trends Microbiol* **10**, 365.
4. Swift, S., Williams, P. & Stewart, G. S. A. B. (1997). N-acylhomoserine lactones and quorum sensing in Proteobacteria. In *Cell-Cell Signalling in Bacteria*, pp. 291–313. Edited by G. M. Dunny & S. C. Winans. Washington, DC: American Society for Microbiology.
5. Whitehead, N. A., Barnard, A. M. L., Slater, H., Simpson, N. J. L. & Salmond, G. P. C. (2001). Quorum-sensing in Gram-negative bacteria. *FEMS Microbiol Rev* **25**, 365–404.

A putative transcription factor inducing mobility in *Mycoplasma pneumoniae*

Mycoplasma genitalium and *Mycoplasma pneumoniae* are among the smallest free-living cells with only a few hundred genes. Their genomes were among the first ones to be completely sequenced (8, 6) and assumed to comprise a 'minimal gene set' coding for a 'minimal metabolism' (10, 13). By comparative genome analysis (9, 7, 5), for *M. pneumoniae*, for 458 [$= \sim 67\%$; 349 ($= \sim 51\%$) before the recent re-annotation] of the 688 (667) ORFs, functional features have been predicted. *M. genitalium* has a subset of the *M. pneumoniae* genes and a conserved gene order. *M. pneumoniae* and *M. genitalium* show many abnormalities with respect to cellular biology, biochemistry and gene expression. For example, no Rho factor is known and only one σ factor was reported. On one hand this is puzzling since, for example, *Bacillus subtilis* has at least 14 σ factors (2), on the other hand symbionts and parasites seem to require a smaller number of their own transcription factors (12).

Sigma factors induce the binding of core RNA polymerase (CPB) to specific initiation sites at the -35 and -10 regions and then detach. Two groups, σ^{70} and σ^{54}, are known. σ^{70} factors group in *essential primary*, *non-essential primary* and *alternative* σ^{70} factors (17). *Alternative* σ *factors*, such as the flagellar σ^{70} factors, regulate transcription of specific subsets of genes. Sigma factors vary strongly in length from around 150 to 450 aa. Variants within an organism show very little similarity: significant divergence and various domain combinations occur (14, 17, 12). Promoter sequences in *M. pneumoniae* are very heterogeneous and the -35 region is poorly conserved (16). This suggests that there are more groups and thus more σ factors than previously assumed. The lack of knowledge of other σ factors is particularly puzzling since it appears that the pathogenicity of *M. pneumoniae* is due to adaptation of its biochemistry and gene expression in response to external stimuli (1). Consequently, it is conceivable that additional σ factors may exist but have been overlooked by computational genome analysis so far. Therefore, we decided to re-investigate the genome with bioinformatics tools to look for possible further σ factors. We describe here just the most important results while methods, detailed computations and further information about identities and accession numbers of individual proteins are described in the supplement (1).

A search for a helix-turn-helix (HTH) motif against the set of translated potential ORFs from *M. pneumoniae* returned 22 significant hits. (Proteins shorter than 50 residues were not considered further.) The resulting set contained the known σ factor A (P78022). The remaining set was used for Smith-Waterman and BLAST searches (1).

Protein P75169 [172 residues, see (1)] from *M. pneumoniae* has a strong sequence similarity to P47667 (171 residues) from *M. genitalium* and to the σ factor D, P10726 (254 residues) from *B. subtilis*. The CLUSTAL X alignment is shown with secondary structure prediction for P75169 in Fig. 1. P75169 and P10726 are identical at 46 residues (27% or 18% respectively) and have similarity at another 68 residues (40%, 27%). Especially in the putative RNA core polymerase binding region (residues 56–76; box 1 in Fig. 1) and a stretch of 50 residues comprising the DNA binding HTH motif (aa 226–246; box 3) the similarities are around 70%. By contrast, the 'random similarity' to an arbitrarily chosen HTH motif (from P62678), is around 30%. Another stretch of 20 aa at the N terminus, for which we found no hint of a possible function, shows high homology and so does a stretch of 15 aa which is approximately in the other region (aa 101–125; box 2) that is relevant for core binding (17). Since the alignment produced virtually no insertions into the *B. subtilis* σ factor, P75169 and P47667 may well have descended from a common ancester with P10726 by loss of domains and thus may be orthologues. Secondary structure prediction complies well with the helical architecture of σ^{70} factors (see Fig. 1). Only one large gap occurs within a helix and none in either of the two functional association domains (CPB and HTH).

To understand why such strong homologies have not been recognized earlier we carried out pairwise and profile-based database searches. Results are summarized in the supplement (1). In essence, it was possible in direct and reverse searches to pull out all σ factors correctly but no subgroups could be specified despite using different seed alignments or levels of sensitivity. However, this was true for several other σ^{70} factors as well. This agrees with numerous claims about the strong divergence in σ factors (11, 17, 12). Therefore, it is likely that so far the homology has simply been overlooked.

Phylogenetic trees for the full sequences were indecisive. Again, the groupings for the functional domains (CPB and HTH) were clearer. In particular, the HTH trees grouped P75169 with flagellar σ^{70} factors. Since HTH regions are indicative for the function (17) this further supports the assumption that P75169 and P47667 are involved in the induction of mobility.

We tried to find further evidence for our claim. Location on the genome also suggests involvement in regulation. Other genes such as the known σ^{70} factor P78022 and genes involved in gene regulation are located nearby [see (1) for more details]. Furthermore, the gene appears to be essential since transposon mutagenesis suggests that viable cells must

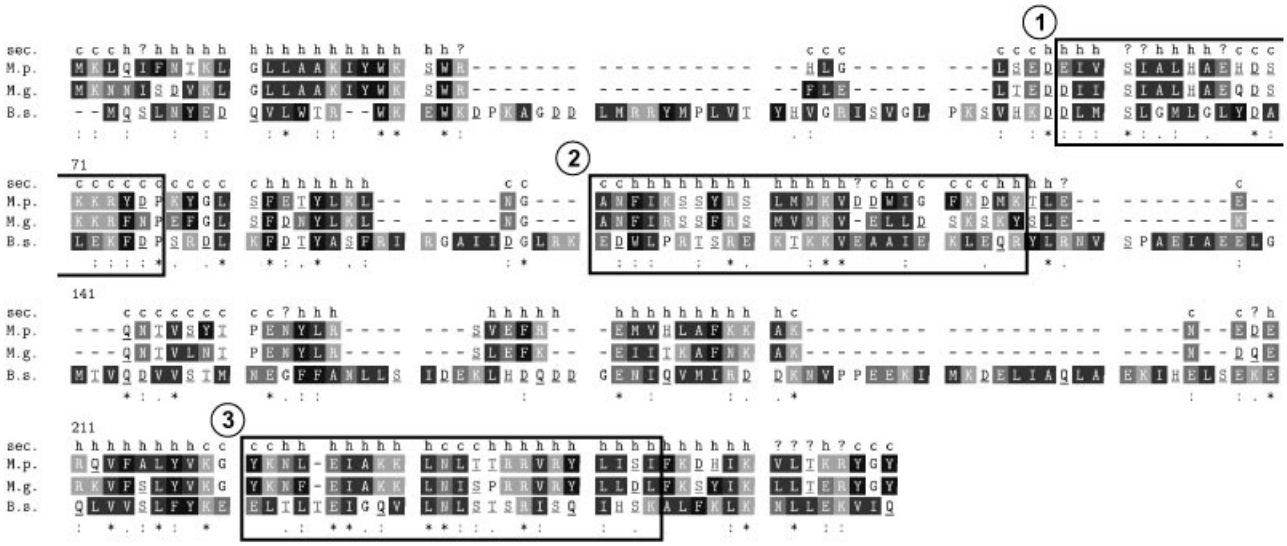

Fig. 1. Alignment of the suspected σ factors from *M. pneumoniae* (*M.p.*; (P75169) and *M. genitalium* (*M.g.*; P47667) with Sig D from *B. subtilis* (*B.s.*; P10726). The first line (sec) shows the secondary structure prediction (h for helix, c for coiled region), the last one the degree of conservation (*, completely conserved; :, almost conserved; ., amino acids of similar properties). The three boxes from N to C terminus indicate the conserved regions. Box 1, potential core binding region; box 2, region involved in core binding; box 3, helix-turn-helix region. See text for further details.

contain P75169 (10). The protein could not be identified in a 2-D gel protein expression study. This is not too surprising since transcription factors often occur only in a small number of copies in the cell and have a relatively short retention time. Furthermore, the protein may be expressed only under very peculiar conditions which are to date largely unknown. P75169 is poorly enhanced on transcription level under standard growth conditions (18). Putative high-scoring promoter sequences can be found with the method described in (16) at −35 (TGCAAA) and −10 (TAAATT).

The *B. subtilis* homologue P10726 belongs to the σ^{70} family group 3 (17), known to influence expression of genes involved in autolysis and flagellar-based motility. It bears strong resemblance to the *fliA* gene from *Escherichia coli*. *M. pneumoniae* has been reported to show some mobility and ability to detach (3, 4) but the mechanism of induction and the corresponding control elements are to date unreported. Thus P75169 is a reasonable candidate for these functions.

Annotation by computations is useful but should be confirmed or disproved in the laboratory (15). Further experiments should, therefore, focus on the precise biochemical role as suggested by our analysis, i.e. finding experimental conditions under which this gene is more strongly expressed, and cloning and functional analysis, for example two-hybrid analysis. We hope that experimentalists will find our suggestions useful and conduct research to verify or falsify our hypothesis.

Acknowledgements

J. W. is supported through the Graduiertenkolleg 'Pathogene Mikroorganismen', a DFG grant (He 780/10-1) and the 'Fonds der Chemischen Industrie'. We thank R. Herrmann and D. Pollack for useful discussions, and J. Ross and R. Gauges for computational help. E.B.-B. is supported by an MRC international recruitment grant and acknowledges support by the KTF during the early stages of the project. We thank C. Zimmermann and H. Goehl for providing unpublished data.

Erich Bornberg-Bauer[1] and January Weiner III[2]

[1]School of Biological Sciences,
The University of Manchester,
2.205 Stopford Building, Oxford Road,
Manchester M13 9PT, UK
[2]ZMBH, University of Heidelberg, D69120
Germany

Author for correspondence: Erich Bornberg-Bauer. Tel: +44 161 275 7396. Fax: +44 161 275 5082. e-mail: ebb@bioinf.man.ac.uk

1. Supplementary data. See http://mic.sgmjournals.org
2. Bairoch, A. & Appweiler, R. (2000). The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic Acids Res* **28**, 45–48.
3. Bredt, W. (1968). Motility and multiplication in *Mycoplasma pneumoniae*: a phase contrast study. *Pathol Microbiol* **32**, 321–326.
4. Bredt, W. (1973). Motility of mycoplasmas. *Ann N Y Acad Sci* **225**, 246–259.
5. Dandekar, T., Huynen, M., Regula, J. T. & 10 other authors (2000). Re-annotating the *Mycoplasma pneumoniae* genome sequence: adding value, function and reading frames. *Nucleic Acids Res* **28**, 3278–3288.
6. Fraser, C., Gocayne, J. D., White, O & 22 other authors (1995). The minimal genome complement of *Mycoplasma genitalium*. *Science* **270**, 397–403.
7. Herrmann, R. & Reiner, B. (1998). *Mycoplasma pneumoniae* and *Mycoplasma genitalium*. *Curr Opin Microbiol* **1**, 572–578.
8. Himmelreich, R., Hilbert, H., Plagens, H., Pirlkl, E., Li, B. C. & Herrmann, R. (1996). Complete sequence analysis of the genome of the bacterium *Mycoplasma pneumoniae*. *Nucleic Acids Res* **24**, 4420–4449.
9. Himmelreich, R., Plagens, H., Hilbert, H., Reiner, B. & Herrmann, R. (1997). Comparative analysis of the genomes of the bacteria *Mycoplasma pneumoniae* and *Mycoplasma genitalium*. *Nucleic Acids Res* **25**, 701–712.
10. Hutchison, C. A., Peterson, S. N., Gill, S. R., Cline, R. T. White, O., Fraser, C. M., Smith, H. O. & Venter, J. C. (1999). Global transposon mutagenesis and a minimal *Mycoplasma* genome. *Science* **286**, 2165–2169.
11. Lonetto, M., Gribskov, M. & Gross, C. A. (1992). The sigma 70 family: sequence conservation and evolutionary relationships. *J Bacteriol* **174**, 3843–3849.
12. Mittenhuber, G. (2002), An inventory of genes encoding RNA polymerase sigma factors in 31 completely sequenced eubacterial genomes. *J Mol Microbiol Biotechnol* **4**, 77–91.
13. Mushegian, A. R. & Koonin, E. V. (1996). A minimal gene set for cellular life derived by comparison of complete bacterial genomes. *Proc Natl Acad Sci U S A* **93**, 10268–10273.
14. Pao, G. M. & Saier, M. H. (1995). Response regulators of bacterial signal transduction systems: selective domain shuffling during evolution. *J Mol Evol* **40**, 136–154.
15. Pollack, J. P. (1997). *Mycoplasma* genes: a case for reflective annotations. *Trends Microbiol* **5**, 413–419.
16. Weiner, J., 3rd, Herrmann, R. & Browning, G. F. (2000). Transcription in *Mycoplasma pneumoniae*. *Nucleic Acids Res* **28**, 4488–4496.
17. Wösten, M. M. (1998). Eubacterial sigma-factors. *FEMS Microbiol Rev* **22**, 127–150.
18. Zimmermann, C. U., Göhlmann, H. W. H., Weiner, J., 3rd & Herrmann, R. (in preparation). Transcription profile of the bacterium *Mycoplasma pneumoniae* grown at different temperatures.