

Conservation and Evolutionary Dynamics of the *agr* Cell-to-Cell Communication System across Firmicutes^{∇†}

Arthur Wuster* and M. Madan Babu*

MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 0QH, United Kingdom

Received 18 July 2007/Accepted 28 September 2007

We present evidence that the *agr* cell-to-cell communication system is present across firmicutes, including the human pathogen *Clostridium perfringens*. Although we find that the *agr* system is evolutionarily conserved and that the general functions which it regulates are similar in different species, the individual regulated genes are not the same. This suggests that the regulatory network controlled by *agr* is dynamic and evolves rapidly.

The pathogenic firmicute species *Staphylococcus aureus* has a well-characterized cell-to-cell communication system whose signals are modified peptides (6). This system is encoded in the accessory gene regulator (*agr*) locus (13). The *agr* cell-to-cell signaling system is an important regulator of biofilm formation and virulence factor expression (7). The *agr* locus consists of four genes, which are *agrB*, *agrD*, *agrC*, and *agrA* (Fig. 1A). The *agrD* and *agrB* genes encode the precursor signaling peptide and the enzyme cleaving and modifying the precursor peptide, respectively (Fig. 1B). The resultant signal is called autoinducing peptide (AIP). Systems which are similar to the *agr* system either by homology or by analogy have been identified in firmicutes outside *Staphylococcus*. In some cases they have different names. For example, the *agr*-like system in *Enterococcus faecalis* is known as the *fsr* system (12) and the one in *Lactobacillus plantarum* is called the *lam* system (3). Here we investigate the distribution of *agr*-like systems outside *Staphylococcaceae* in order to understand the conservation and plasticity of their interaction with other cellular components.

In order to define *agr* homologues in taxa other than *Staphylococcus*, we built hidden Markov models of protein domains which are considered to be unique to the *agr* system, namely, the AgrB and AgrD domains. To our knowledge, no function of AgrB apart from processing of AgrD has been reported so far. In a complementary approach, we identified *agr* homologues with PSI-BLAST searches. In the 384 genomes used in this analysis, 33 instances of the AgrB domain and 18 instances of the AgrD domain were found, all of them in firmicutes. The genomic context of the AgrB homologues was mapped onto a phylogenetic tree built from ribosomal protein sequences (Fig. 1C). All identified cases of the AgrD domain were found in close genomic proximity to a gene encoding an AgrB domain, which means that there are no orphan *agrD* homologues in the genomes we searched. All 18 loci which contained both the AgrD and the AgrB homologue were considered to be true

quorum-sensing loci. Four other loci were also considered to be true quorum-sensing loci because they encoded a histidine kinase and a response regulator linked to the AgrB homologue. In the case of *Enterococcus faecalis* V583, the *agrD* homologue *fsrD* is transcribed together and in frame with *fsrB* but is translated independently of it (10).

To investigate whether the *agr* system is an evolutionarily ancient system, we built two phylogenetic trees by use of different methods. The first tree was built from an alignment of AgrB homologues. The second tree was built from an alignment of AgrB, AgrD, AgrC, and AgrA homologues, whose sequences were concatenated (see the supplemental material). The topology of these trees should agree with the topology of the tree built from ribosomal sequences if *agr* is inherited only vertically. Although these trees generally agree with each other and the ribosomal tree, the *agr* locus of *Clostridium acetobutylicum* clusters with *Listeriaceae*. *Clostridium acetobutylicum* is, however, a member of the class *Clostridia*, therefore making its *agr* system a candidate for horizontal gene transfer. This is also supported by AgrD of *Clostridium acetobutylicum* showing remarkable similarity to AgrD of *Listeriaceae* (Fig. 2). To our knowledge, this would be the first reported instance of horizontal gene transfer between *Clostridia* and *Listeriaceae*.

In three genomes we were able to identify a putative *agr* system where, to our knowledge, none had been reported before. These genomes are *Moorella thermoacetica* ATCC 39070, *Desulfibacterium hafniense* Y51, and *Thermoanaerobacter tengcongensis*. We propose that these genomes might contain a functional peptide quorum-sensing circuit. Although we could not detect genes which take the role of *agrD* in these species, this could be due to errors in the prediction of small genes or to such genes being nonhomologous and located elsewhere on the genome or fused with other genes, as in the case of *Enterococcus faecalis* (10). Alternatively, this might represent instances of a system that can recognize peptides secreted by other organisms (cross talk). In the genomes, such as those of *Bacillus halodurans*, *Syntrophomonas wolfei*, and the three *Clostridium perfringens* strains, which contain an AgrB homologue that is not linked to a complete two-component system, *agr*-mediated quorum sensing might also be functional. This would require the two-component sensing system to be encoded elsewhere in the genome rather than being linked to *agrB*. However, in the *Clostridium perfringens* strain ATCC

* Corresponding author. Mailing address: MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 0QH, United Kingdom. Phone: 44 (0)1223 402479. Fax: 44 (0)1223 213556. E-mail for Arthur Wuster: awuster@mrc-lmb.cam.ac.uk. E-mail for M. Madan Babu: madanm@mrc-lmb.cam.ac.uk.

† Supplemental material for this article may be found at <http://j.b.asm.org/>.

∇ Published ahead of print on 12 October 2007.

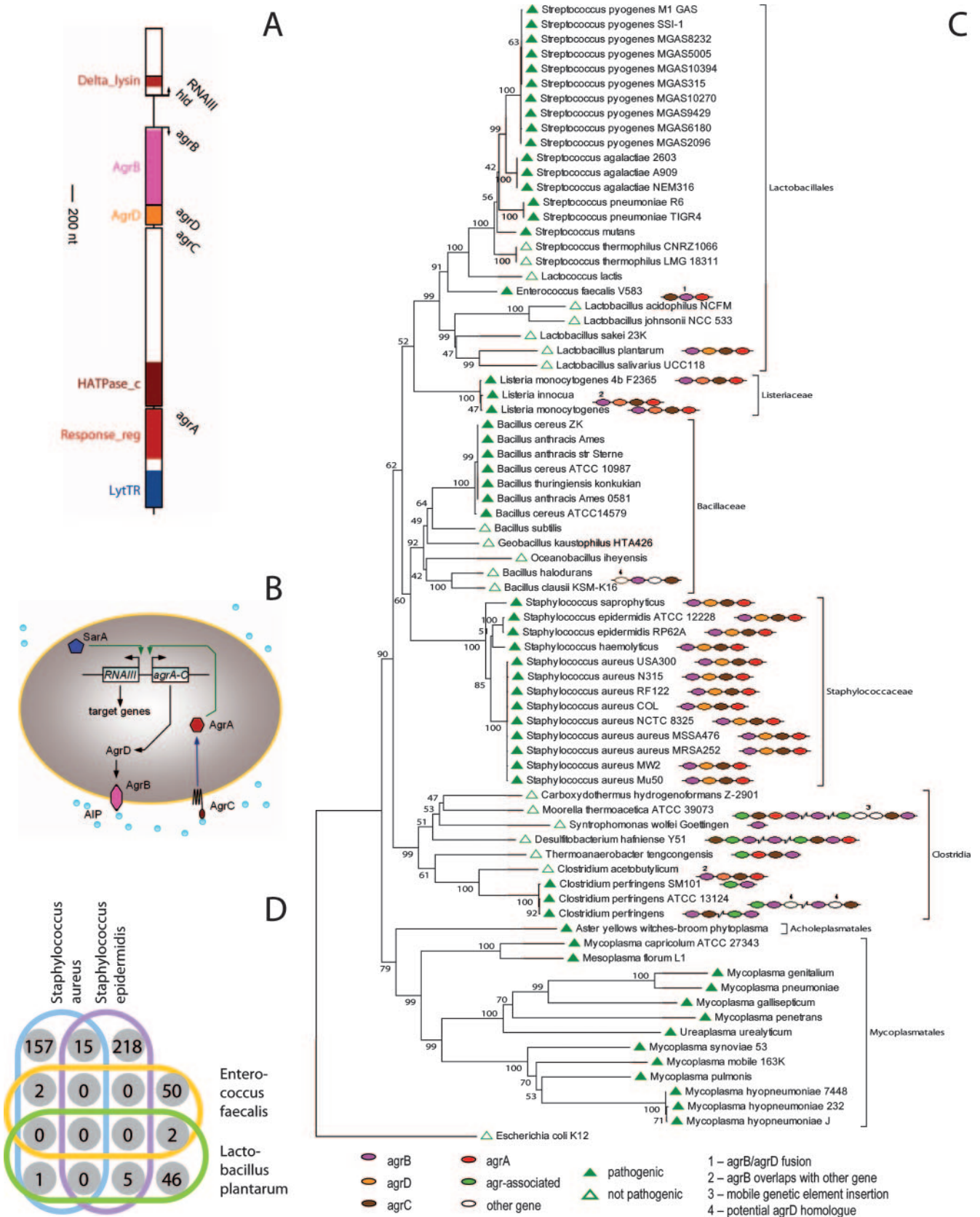


FIG. 1. (A) Organization of the *agr* locus in *S. aureus*, with Pfam domains highlighted in color. nt, nucleotides. (B) Cartoon of the basic interactions of the *agr* system in *S. aureus*. The interactions of the proteins encoded by the *agrBDCA* operon lead to a positive-feedback loop and further *agrBDCA* expression. Blue arrows, phosphorylation; green arrows, transcriptional interactions. RNAIII and *hld*, which encodes δ -lysin,

| | C | PXXP | |
|---|--|----------------------------------|------------------|
| Staphylococcus saprophyticus ATCC 15305 | -MNVIKSISKSISKSISNYFAKVFASIGSISTINPCFG-YTDESEI | PK-ELTDLYE-- | 72494665 |
| Staphylococcus epidermidis ATCC 12228 | -----MENIFNLFIKFFTTILEF | IGTVAGDSVYAS-YFDEPEVPE-ELTKLYE-- | 27316101 |
| Staphylococcus epidermidis RP62A | -----MENIFNLFIKFFTTILEF | IGTVAGDSVYAS-YFDEPEVPE-ELTKLYE-- | 57638027 |
| Staphylococcus haemolyticus JCSC1435 | -----MTVLVDLLIKLFTLLQSIGTIASFTPCTT-YFDEPEVPE-ELTNAK-- | | 68446720 |
| Staphylococcus aureus MRSA252 | -----MKKLLNKVIELLVDFNFIGYRAAYIMEDF-LLDEAEVPK-ELTQLHE-- | | 49242392 |
| Staphylococcus aureus MSSA476 | -----MKKLLNKVIELLVDFNFIGYRAAYIMEDF-LLDEAEVPK-ELTQLHE-- | | 49245275 |
| Staphylococcus aureus MW2 | -----MKKLLNKVIELLVDFNFIGYRAAYIMEDF-LLDEAEVPK-ELTQLHE-- | | 21205132 |
| Staphylococcus aureus COL | -----MNTLNFNLFDFITGILKNIGNIAAYSTDF-IMDEVEVPK-ELTQLHE-- | | 57284895 |
| Staphylococcus aureus NCTC 8325 | -----MNTLNFNLFDFITGILKNIGNIAAYSTDF-IMDEVEVPK-ELTQLHE-- | | 87203491 |
| Staphylococcus aureus USA300 | -----MNTLNFNLFDFITGILKNIGNIAAYSTDF-IMDEVEVPK-ELTQLHE-- | | 87126405 |
| Staphylococcus aureus Mu50 | -----MNTLVNMFDFI IKLAKAIGIVGGVNAAS-LFDEPKVPA-ELTNLYD-K | | 14247810 |
| Staphylococcus aureus N315 | -----MNTLVNMFDFI IKLAKAIGIVGGVNAAS-LFDEPKVPA-ELTNLYD-K | | 13701831 |
| Staphylococcus aureus RF122 | -----MNTLVNMFDFI IKLAKAIGIVGGVNAAS-LFDEPKVPA-ELTNLYD-K | | 82657170 |
| Lactobacillus plantarum WCFS1 | -----MKQKMYEIAHLFKYVGAQQLVMC | VG-IVFETKIPD-ELRK---- | 28272731 |
| Listeria innocua | ----MKNMKSVGKFLSRKLEEQSMK | VADSSMSKACFM-FVYEPKSPFVKMQEKNENK | 16412463 |
| Listeria monocytogenes | ----MKNMKSVGKFLSRKLEEQSMK | VADSSMSKACFM-FVYEPKSPFVKMQEKNENK | 16409408 |
| Listeria monocytogenes str. 4b F2365 | ----MKNMKSVGKFLSRKLEEQSMK | VADSSMSKACFM-FVYEPKSPFVKMQEKNENK | 46879544 |
| Clostridium acetobutylicum ATCC 824 | -MNLKEQLNKKVNDKFKIG-IGKASMK | IGEQANGK-CVLVTLVEPKMPE-ELLKENIDK | 15022901 |
| Clostridium perfringens SM101 | -----MKKLNKLLTLFAALTTVIATTVATSACI | W-FTHQPEEPK-SLRDE---- | 110682078 |
| Clostridium perfringens ATCC 13124 | -----MKKLNKLLTLFAALTTVIATTVATSACI | W-FTHQPEEPK-SLRDE---- | 110674548 |
| Clostridium perfringens ATCC 13124 | MRKSILSMALLTTLTINASTTSFATIGLEEM----- | PESMKKSR----- | 110673345 |
| Enterococcus faecalis V583 | -----DGVGTKPRLNQNSPWFQGNMGQTEPKPKNIE-K | | 29343827 |
| Bacillus halodurans C-125 | -----MKRTAKVISKATLIGSKAFVNASSP-LIYAFKIPENGLKKQK---- | | 10176097 |

FIG. 2. Alignment of potential AgrD homologues. GenBank identifiers are given in the column to the right. Red, experimentally determined signaling peptides; yellow background, residues strongly conserved across the homologues; green background, residues conserved between *Listeria* and *Clostridium acetobutylicum*. Boldface type shows potential AgrD homologues which were identified by genomic context (linkage to AgrB) rather than by homology. The conserved cysteine, which is required for the formation of the thiolactone ring, and the Pro-X-X-Pro motif are highlighted at the top of the alignment (see text for details).

13124 and strain 13, one of the two paralogues of AgrB co-ocurs with a histidine sensor kinase. Peptide-mediated quorum sensing in the *Clostridium perfringens* species would be of particular interest as this species is the causative agent of the human disease gas gangrene (11). Interestingly, we also identified an uncharacterized gene frequently present in close genomic proximity to the *agr* locus which encodes a transmembrane protein with conserved residues (Fig. 1C; see also Fig. S1 in the supplemental material).

In the *Clostridium perfringens* strains as well as in *Bacillus halodurans*, we were able to identify short open reading frames which might have the function of *agrD* without having detectable sequence similarity to it (marked with the number 4 in Fig. 1C and highlighted in bold in Fig. 2). Although they lack significant sequence similarity to known AgrD homologues, they possess key sequence features connected to AIP function: a central cysteine residue which is required for the formation

of the thiolactone ring (9) and a Pro-X-X-Pro motif which might serve as a recognition site for AgrB (16) (Fig. 2).

Next, we asked whether the genes which are transcriptionally regulated by *agr* are conserved between different species with a known *agr* quorum-sensing system. Differences in gene expression levels between wild-type and *agr* knockout strains have previously been obtained for *Lactobacillus plantarum* WCFS1 (14), *Enterococcus faecalis* OG1RF (1), *Staphylococcus aureus* NCTC 8325 (4), *Staphylococcus aureus* UAMS-1 (2), and *Staphylococcus epidermidis* (15). We used these data to compare genes which are differentially expressed in these *agr* mutants (see Fig. S3 in the supplemental material). According to our criteria (see the supplemental material), 175 genes were differentially expressed in *Staphylococcus aureus*, 54 in *Lactobacillus plantarum*, 54 in *Enterococcus faecalis*, and 238 in *Staphylococcus epidermidis*. After defining clusters of orthologous genes, we asked if there were any orthologues which were

are controlled by the same promoter and share the same transcript. The second promoter controls *agrBDCA*. The *agrC* and *agrA* genes encode the transmembrane kinase which senses the AIP (cyan circles) and a response regulator (red hexagon) which is phosphorylated by AgrC and affects changes in gene expression, respectively. Phosphorylation of the AgrA receiver domain by AgrC leads to AgrA activation and DNA binding. AgrA positively regulates transcription of the *agr* operon. This results in the production of a high level of AIP, leading to even higher *agr* expression. Apart from the *agrBDCA* promoter, another promoter positively regulated by AgrA is the one controlling expression of a small noncoding RNA, RNAIII. The RNAIII promoter is located next to the *agrBDCA* promoter but is transcribed in the opposite direction. RNAIII rather than AgrA is the immediate regulator of most of the genes regulated by the *agr* operon. (C) Tree of firmicutes based on ribosomal sequences with bootstrap values (1,000 iterations). Where applicable, the genomic organization of the *agr* locus homologue is depicted to the right of the operational taxonomic unit. One instantly obvious finding is that homologues of the *agr* system are spread across firmicutes and are not limited to a single phylogenetic class. Furthermore, the genomic organization of the *agr* system is conserved outside the family *Staphylococcaceae*, that is to say, in most cases *agrB* is linked to a two-component signal transduction system. While each genome contained only a single copy of the AgrD domain, some contained more than one copy of the AgrB domain: *M. thermoacetica* (three copies), *D. hafniense* Y51 (three copies), *C. perfringens* (two copies), and *C. perfringens* ATCC 13124 (two copies). (D) Venn diagram showing numbers of differentially expressed genes and how many orthologous clusters are differentially expressed in more than one genome.

differentially expressed in all three genomes. We did not find any cluster which was differentially regulated in all three genomes (Fig. 1D). Therefore, it appears that there is no significant propensity for orthologous genes to be regulated by the *agr* system in the three different species. Using the Infernal program (5) to detect small noncoding RNAs, we found that regulation of genes by *agr* via the global regulator RNAIII (Fig. 1B) occurred only in the *Staphylococcus* species. This suggests that RNAIII is an evolutionary innovation unique to *Staphylococcaceae* and that the downstream regulators under the control of *agr* can evolve rapidly.

In agreement with previous observations, genes which are involved in biofilm formation, such as those encoding exopolysaccharide biosynthesis proteins, are downregulated by *agr* in *Staphylococcus aureus* and *Enterococcus faecalis* (8). Some exopolysaccharide biosynthesis genes, such as those of the *cps2* cluster, are upregulated in *Lactobacillus plantarum*. This seems to contradict the observation that deletion of *lamA*, the *agrA* homologue, has the phenotypic effect of reduced adherence to a glass substratum compared to results for the wild type. The production of virulence proteins is upregulated by the *agr*-like system in *Enterococcus faecalis* and *Staphylococcus aureus* N315, which are both virulent. A clear conclusion is that the *agr* system is not a universal regulator of virulence factors only and is also active in nonpathogenic firmicutes. Furthermore, we conclude that although the functions which are regulated, including biofilm formation, are similar, the individual genes are not. This suggests that the regulatory network controlled by *agr* is flexible and can adapt rapidly to changing environments. Therefore, the *agr* system might be of use in the emerging field of synthetic biology.

A.W. thanks the Medical Research Council for funding.

A.W. and M.M.B. thank Ingmar Schäfer, Aswin Seshasayee, and Subhajoti De for critically reading the manuscript.

REFERENCES

- Bourgogne, A., S. G. Hilsenbeck, G. M. Dunny, and B. E. Murray. 2006. Comparison of OG1RF and an isogenic *fsrB* deletion mutant by transcriptional analysis: the Fsr system of *Enterococcus faecalis* is more than the activator of gelatinase and serine protease. *J. Bacteriol.* **188**:2875–2884.
- Cassat, J., P. M. Dunman, E. Murphy, S. J. Projan, K. E. Beenken, K. J. Palm, S. J. Yang, K. C. Rice, K. W. Bayles, and M. S. Smeltzer. 2006. Transcriptional profiling of a *Staphylococcus aureus* clinical isolate and its isogenic *agr* and *sarA* mutants reveals global differences in comparison to the laboratory strain RN6390. *Microbiology* **152**:3075–3090.
- Diep, D. B., L. S. Havarstein, J. Nissen-Meyer, and I. F. Nes. 1994. The gene encoding plantaricin A, a bacteriocin from *Lactobacillus plantarum* C11, is located on the same transcription unit as an *agr*-like regulatory system. *Appl. Environ. Microbiol.* **60**:160–166.
- Dunman, P. M., E. Murphy, S. Haney, D. Palacios, G. Tucker-Kellogg, S. Wu, E. L. Brown, R. J. Zagursky, D. Shlaes, and S. J. Projan. 2001. Transcription profiling-based identification of *Staphylococcus aureus* genes regulated by the *agr* and/or *sarA* loci. *J. Bacteriol.* **183**:7341–7353.
- Griffiths-Jones, S., A. Bateman, M. Marshall, A. Khanna, and S. R. Eddy. 2003. Rfam: an RNA family database. *Nucleic Acids Res.* **31**:439–441.
- Ji, G., R. Beavis, and R. P. Novick. 1997. Bacterial interference caused by autoinducing peptide variants. *Science* **276**:2027–2030.
- Kleerebezem, M., L. E. Quadri, O. P. Kuipers, and W. M. de Vos. 1997. Quorum sensing by peptide pheromones and two-component signal-transduction systems in Gram-positive bacteria. *Mol. Microbiol.* **24**:895–904.
- Kong, K. F., C. Vuong, and M. Otto. 2006. *Staphylococcus* quorum sensing in biofilm formation and infection. *Int. J. Med. Microbiol.* **296**:133–139.
- Mayville, P., G. Ji, R. Beavis, H. Yang, M. Goger, R. P. Novick, and T. W. Muir. 1999. Structure-activity analysis of synthetic autoinducing thiolactone peptides from *Staphylococcus aureus* responsible for virulence. *Proc. Natl. Acad. Sci. USA* **96**:1218–1223.
- Nakayama, J., S. Chen, N. Oyama, K. Nishiguchi, E. A. Azab, E. Tanaka, R. Kariyama, and K. Sonomoto. 2006. Revised model for *Enterococcus faecalis* *fsr* quorum-sensing system: the small open reading frame *fsrD* encodes the gelatinase biosynthesis-activating pheromone propeptide corresponding to staphylococcal AgrD. *J. Bacteriol.* **188**:8321–8326.
- Present, D. A., R. Meislin, and B. Shaffer. 1990. Gas gangrene. A review. *Orthop. Rev.* **19**:333–341.
- Qin, X., K. V. Singh, G. M. Weinstock, and B. E. Murray. 2000. Effects of *Enterococcus faecalis* *fsr* genes on production of gelatinase and a serine protease and virulence. *Infect. Immun.* **68**:2579–2586.
- Recei, P., B. Kreiswirth, M. O'Reilly, P. Schlievert, A. Gruss, and R. P. Novick. 1986. Regulation of exoprotein gene expression in *Staphylococcus aureus* by agar. *Mol. Gen. Genet.* **202**:58–61.
- Sturme, M. H., J. Nakayama, D. Molenaar, Y. Murakami, R. Kunugi, T. Fujii, E. E. Vaughan, M. Kleerebezem, and W. M. de Vos. 2005. An *agr*-like two-component regulatory system in *Lactobacillus plantarum* is involved in production of a novel cyclic peptide and regulation of adherence. *J. Bacteriol.* **187**:5224–5235.
- Yao, Y., C. Vuong, S. Kocianova, A. E. Villaruz, Y. Lai, D. E. Sturdevant, and M. Otto. 2006. Characterization of the *Staphylococcus epidermidis* accessory-gene regulator response: quorum-sensing regulation of resistance to human innate host defense. *J. Infect. Dis.* **193**:841–848.
- Zhang, L., J. Lin, and G. Ji. 2004. Membrane anchoring of the AgrD N-terminal amphipathic region is required for its processing to produce a quorum-sensing pheromone in *Staphylococcus aureus*. *J. Biol. Chem.* **279**:19448–19456.