J. Mol. Microbiol. Biotechnol. (2002) 4(5): 503–513.

JMMB Research article

Regulatory Relationship of Two-Component and ABC Transport Systems and Clustering of Their Genes in the *Bacillus/Clostridium* Group, Suggest a Functional Link Between Them

Pascale Joseph, Gwennaele Fichant, Yves Quentin, and François Denizot*

Laboratoire de Chimie Bactérienne, Institut de Biologie Structurale et Microbiologie, CNRS 31, Chemin Joseph Aiguier, 13402 Marseille Cedex 20, France

Abstract

On the *Bacillus subtilis* chromosome there are five examples of genes encoding two-component systems with response regulators of the OmpR family adjacent to genes encoding sub-family 9 ABC transport systems. Three of these (*yts*, *yvc*, *yxd*) are very similar in gene organization and in sequence. We demonstrate that the TCS and ABC transporter genes do not belong to the same transcriptional unit. The ABC transport and TCS systems are functionally linked, each response regulator controlling the expression of its cognate ABC transporter genes but not its own. Analysis of 48 bacterial genomes revealed that such family clusters only exist in the *Bacillus/Clostridium* group. Evolutionary analyses indicated that almost all clustered OmpR response regulators constitute two groups (''GI'' and ''GII'') whereas almost all clustered sub-family 9 nucleotide-binding domains belong to two other groups (''9A'' and ''9B''). Interestingly, there is a mutually exclusive clustering between genes encoding ''GI'' or a ''GII'' response regulators and genes encoding ''9A'' or a ''9B'' nucleotide binding proteins. We propose that a two-component system and its cognate ABC transporter genes have evolved as a unit in *Bacillus/Clostridium,* both systems participating in a common physiological process.

Abbreviations

nt, nucleotide;
bp, base pair;
TCS, two-component regulatory system;
RR, Response Regulator;
HK, Histidine Kinase;
ABC, ATP binding cassette;
NBD, Nucleotide Binding Domain;
MSD, Membrane Spanning Domain;
SBP, Substrate Binding Protein;
TM, transmembrane domain.

*For correspondence. Email denizot@ibsm.cnrs-mrs.fr; Tel. +33 4 91164141; Fax: +33 4 91718914.

Introduction

The ability of bacteria to survive drastic changes in environmental conditions depends on detection and response, as well as transport systems that allow them to cope with environmental changes.

Detection and transduction of specific signals often involve two-component regulatory systems (TCS) that trigger the cell to an adaptive process (Parkinson and Kofoid, 1992). A typical TCS comprises two proteins: a histidine kinase (HK) and a response regulator (RR). Once the signal has been detected, the HK autophosphorylates at a conserved histidine residue in an ATP-dependent reaction. In a second step, the phosphoryl group is transferred to an aspartyl residue of the RR, which, in most cases, acts as a transcriptional activator/repressor that controls the expression of genes involved in the adaptive response.

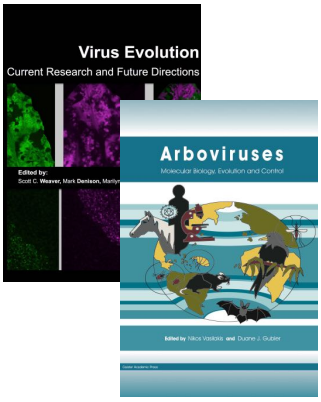
Transport systems, in particular ABC transporters, can either import or export a wide variety of compounds across biological membranes (Ames, 1986; Higgins *et al.*, 1986). The prototypical ABC transporter comprises two membrane-spanning domains (MSD) and two cytoplasmic nucleotide-binding domains (NBD), the latter having an ATPase activity that provides the energy for the translocation of substrates across the membrane. In a bacterial ABC transporter, the four domains are generally expressed as separate proteins. In addition, all uptake systems involve an extra-cytoplasmic partner, the solute binding protein (SBP), which has a high affinity for its substrate and presents it to the import complex (Fath and Kolter, 1993; Higgins, 1992).

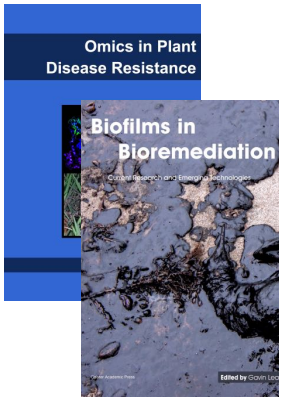
Bacterial genomic sequencing has given rise to a huge amount of data, which after computer analysis for similarities allows unknown genes to be identified. In most cases, they represent a significant fraction of the predicted genes which need to be studied further. Products of such unknown genes sometimes have particular features allowing them to be recognized. This recognition may permit the investigator to characterize other elements of more complex structure. Inventories of TCS (Mizuno, 1997; Fabret *et al.*, 1999; Beier and Frank, 2000; Throup *et al.*, 2000; Rodrigue *et al.*, 2000) and ABC transporter (Linton and Higgins, 1998; Tomii

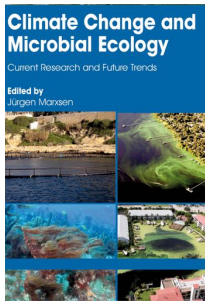
Further Reading

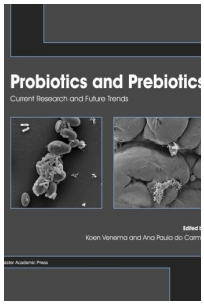
Caister Academic Press is a leading academic publisher of advanced texts in microbiology, molecular biology and medical research. Full details of all our publications at **caister.com**

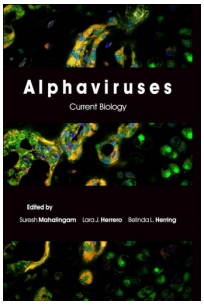
- **MALDI-TOF Mass Spectrometry in Microbiology**
  Edited by: M Kostrzewa, S Schubert (2016)
  www.caister.com/malditof
- ***Aspergillus* and *Penicillium* in the Post-genomic Era**
  Edited by: RP Vries, IB Gelber, MR Andersen (2016)
  www.caister.com/aspergillus2
- **The Bacteriocins: Current Knowledge and Future Prospects**
  Edited by: RL Dorit, SM Roy, MA Riley (2016)
  www.caister.com/bacteriocins
- **Omics in Plant Disease Resistance**
  Edited by: V Bhadauria (2016)
  www.caister.com/opdr
- **Acidophiles: Life in Extremely Acidic Environments**
  Edited by: R Quatrini, DB Johnson (2016)
  www.caister.com/acidophiles
- **Climate Change and Microbial Ecology: Current Research and Future Trends**
  Edited by: J Marxsen (2016)
  www.caister.com/climate
- **Biofilms in Bioremediation: Current Research and Emerging Technologies**
  Edited by: G Lear (2016)
  www.caister.com/biorem
- **Microalgae: Current Research and Applications**
  Edited by: MN Tsaloglou (2016)
  www.caister.com/microalgae
- **Gas Plasma Sterilization in Microbiology: Theory, Applications, Pitfalls and New Perspectives**
  Edited by: H Shintani, A Sakudo (2016)
  www.caister.com/gasplasma
- **Virus Evolution: Current Research and Future Directions**
  Edited by: SC Weaver, M Denison, M Roossinck, et al. (2016)
  www.caister.com/virusevol
- **Arboviruses: Molecular Biology, Evolution and Control**
  Edited by: N Vasilakis, DJ Gubler (2016)
  www.caister.com/arbo
- ***Shigella*: Molecular and Cellular Biology**
  Edited by: WD Picking, WL Picking (2016)
  www.caister.com/shigella
- **Aquatic Biofilms: Ecology, Water Quality and Wastewater Treatment**
  Edited by: AM Romaní, H Guasch, MD Balaguer (2016)
  www.caister.com/aquaticbiofilms
- **Alphaviruses: Current Biology**
  Edited by: S Mahalingam, L Herrero, B Herring (2016)
  www.caister.com/alpha
- **Thermophilic Microorganisms**
  Edited by: F Li (2015)
  www.caister.com/thermophile

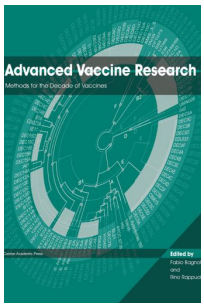


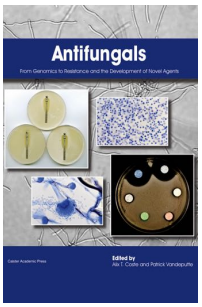












- **Flow Cytometry in Microbiology: Technology and Applications**
  Edited by: MG Wilkinson (2015)
  www.caister.com/flow
- **Probiotics and Prebiotics: Current Research and Future Trends**
  Edited by: K Venema, AP Carmo (2015)
  www.caister.com/probiotics
- **Epigenetics: Current Research and Emerging Trends**
  Edited by: BP Chadwick (2015)
  www.caister.com/epigenetics2015
- ***Corynebacterium glutamicum*: From Systems Biology to Biotechnological Applications**
  Edited by: A Burkovski (2015)
  www.caister.com/cory2
- **Advanced Vaccine Research Methods for the Decade of Vaccines**
  Edited by: F Bagnoli, R Rappuoli (2015)
  www.caister.com/vaccines
- **Antifungals: From Genomics to Resistance and the Development of Novel Agents**
  Edited by: AT Coste, P Vandeputte (2015)
  www.caister.com/antifungals
- **Bacteria-Plant Interactions: Advanced Research and Future Trends**
  Edited by: J Murillo, BA Vinatzer, RW Jackson, et al. (2015)
  www.caister.com/bacteria-plant
- ***Aeromonas***
  Edited by: J Graf (2015)
  www.caister.com/aeromonas
- **Antibiotics: Current Innovations and Future Trends**
  Edited by: S Sánchez, AL Demain (2015)
  www.caister.com/antibiotics
- ***Leishmania*: Current Biology and Control**
  Edited by: S Adak, R Datta (2015)
  www.caister.com/leish2
- ***Acanthamoeba*: Biology and Pathogenesis (2nd edition)**
  Author: NA Khan (2015)
  www.caister.com/acanthamoeba2
- **Microarrays: Current Technology, Innovations and Applications**
  Edited by: Z He (2014)
  www.caister.com/microarrays2
- **Metagenomics of the Microbial Nitrogen Cycle: Theory, Methods and Applications**
  Edited by: D Marco (2014)
  www.caister.com/n2

Order from caister.com/order

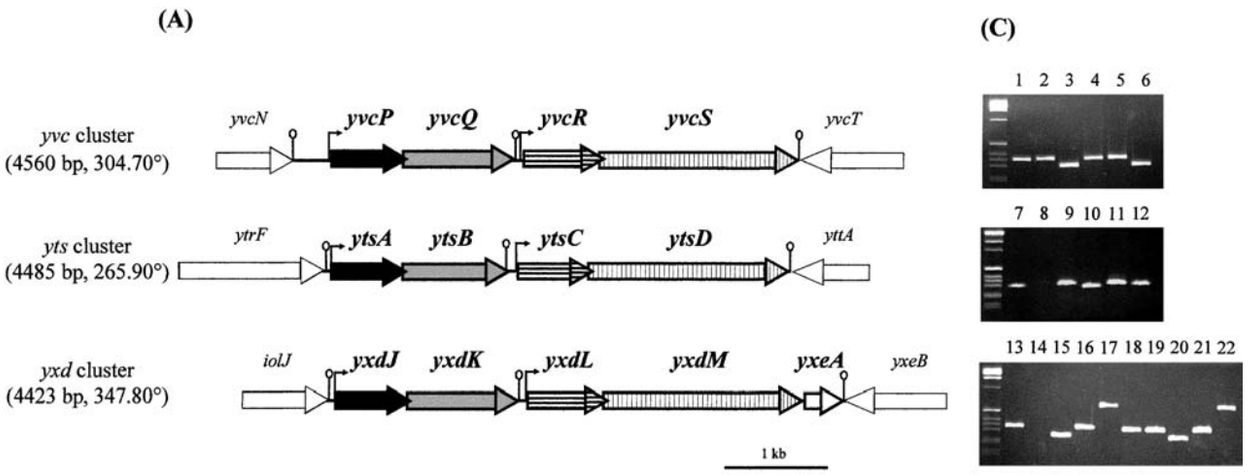
and Kanehisa, 1998; Dassa *et al*., 1999; Quentin *et al*., 1999; Braibant *et al*., 2000; Paulsen *et al*., 2000; Quentin and Fichant, 2000; Saier *et al*., 2002) have been described in this way for several bacteria.

Analyses of the genomic localization of the structural genes in *Bacillus subtilis* revealed five occurrences of very close proximity between genes encoding two-component systems with RR from the OmpR family (Fabret *et al*., 1999) and genes encoding extruders from sub-family 9 of ABC transport systems (Quentin *et al.,*1999). The fact that TCS are very often involved in key physiological processes in bacteria, such as response to environmental changes, prompted us to analyse these genetically clustered partners in more detail. In this paper we present evidence that for the three closely related clusters (*yts, yvc* and *yxd*) a two-component system controls the expression of its cognate ABC transporter genes, suggesting that both systems might be involved in a common physiological process. On the other hand, our results also indicate that TCS does not control the expression of its own genes. A search in other bacterial genomes revealed that such clusters, involving sub-family 9 ABC transporter genes and OmpR response regulator gene, are only found in the *Bacillus/Clostridium* group. Some features of the partners are presented, and an evolutionary model is proposed for their origin and expansion.

Results

Three Conserved Gene Clusters Encode Two-Component and ABC Transport Systems in Bacillus Subtilis

The *yts*, *yvc* and *yxd* gene clusters encoding both TCSs with RR from the OmpR family and ABC transporters from sub-family 9 were found to be very similar. As shown in Figure 1A, they comprise two gene pairs located in tandem and separated by an intergenic region of less than 150 bp. According to the direction of transcription, the same genetic organization—RR, HK, NBD and MSD encoding genes—is observed in all three clusters. In the case of *yxd*, an additional gene (*yxeA*) encoding a small protein (115



(B)

RR	HK	NBD	MSD
YvcP 40 YtsA 48 35 YxdJ	YvcQ 30 YtsB 37 29 YxdK	YvcR 48 YtsC 57 49 YxdL	YvcS 24 YtsD 31 22 YxdM

Figure 1. The three related clusters *yvc*, *yts* and *yxd* in *B. subtilis*. (A) Schematic representation of the three clusters. Black arrows correspond to RR genes, grey arrows to HK genes, horizontally hatched arrows to NBD genes and vertically hatched arrows to MSD genes. The position of putative promoters and terminators are indicated. The size of the clusters and their location on the chromosome are given between brackets. Upstream of *yvcP* is the *crh*-containing operon (Galinier *et al*., 1997), upstream of *ytsA* is the *ytr* operon involved in acetoin utilisation (Yoshida *et al*., 2000b) and upstream of *yxdJ* is the *iol* operon involved in *myo*-inositol catabolism (Yoshida *et al*.,1997). Only the last genes in these operons are represented. (B) Identity scores obtained when comparing the functional related proteins are indicated and expressed as percentages. Similarity searches were done with the gapped version of the BLASTP2 program suite (Altschul *et al.,* 1997). In order to simplify the table, BS/ was omitted in protein names. (C) PCR products generated by RT-PCR. Ethidium bromide-stained 2% agarose gels of PCR products obtained with primers spanning the intergenic region between HK and NBD genes (lanes 2, 5, 8, 11, 14 and 19) and *yxdM-yxeA* (lanes 17 and 22), products obtained with specific primer of individual HK genes (lanes 1, 4, 7, 10, 13 and 18), NBD genes (lanes 3, 6, 9, 12, 15 and 20) and MSD gene (lanes 16 and 21). Templates for PCR were: cDNA prepared from *B. subtilis* 168 total RNA (lanes 1 to 3; 7 to 9; 13 to 17); *B. subtilis* 168 chromosomal DNA (lanes 4 to 6; 10 to 12; 18 to 22). No PCR product with RNA as template was obtained (not shown).

amino acids), is present downstream of the *yxdLM* genes and is predicted to be part of the same transcriptional unit. The similarity found between the nucleotide sequences of the three clusters does not extend to the surrounding sequences (data not shown). The RRs and NBDs show higher percentage of identity (greater than 35% and 48%, respectively, Figure 1B). The HKs and the MSDs, although known to be less conserved, show more than 29% and 22% identity, respectively. Proteins encoded by the *yvc* and *yxd* clusters are more closely related to one another than they are to those encoded by the *yts* cluster. All these results suggest that the three clusters evolved as a unit through two successive duplication events from a common ancestor, the first duplication giving rise to the *yts* cluster, the second generating the *yvc* and *yxd* clusters.

The presence of a putative ρ-independent terminator of transcription downstream of each RR encoding gene and the presence of a putative sigma A promoter within the intergenic region suggest that TCS and ABC transporter genes are not part of the same transcriptional unit. To confirm this suggestion, the expression of the structural genes was monitored using RT-PCR as described in Methods. In all cases DNA bands corresponding to the amplification of each HK cDNA (Figure 1C; lanes 1, 7 and 13) and of each NBD cDNA (Figure 1C; lanes 3, 9 and 15) were obtained indicating that both HK and NBD encoding genes are expressed in *B. subtilis*. However, the lack of amplification of cDNA spanning the *ytsB-ytsC* (Figure 1C; lane 8) or the *yxdK-yxdL* (Figure 1C; lane 14) intergenic region clearly indicates that the corresponding TCS and ABC transporter structural genes are not part of the same transcript, a result consistent with the presence of putative transcriptional terminators downstream *ytsB* and *yxdJ* ($\Delta G^\circ = -15.7$ kcal mol^{-1} and -16.2 kcal mol^{-1}, respectively). In contrast, a DNA band corresponding to the amplification of the cDNA spanning *yvcQ* and *yvcR* was obtained (lane 2), suggesting that in the *yvc* cluster, the TCS and ABC transporter genes might be cotranscribed. This result might be explained by inefficiency of the transcription terminator, located downstream *yvcQ,*which has a low ΔG° value (-6 kcal mol^{-1}). In addition, cDNA amplification was obtained with a primer pair specific to *yxdM-yxeA* (Figure 1C; lane 17), suggesting that *yxeA* is cotranscribed with the ABC transporter genes, in agreement with our predictions.

Effect of In Vivo Overproduction of BS/YtsA, BS/YvcP or BS/YxdJ Response Regulators on the Expression of ABC Transporter Genes

The transported substrate and the nature of the induction signal detected by the two-component systems are unknown in all cases. Nevertheless, conservation of physical proximity may indicate that the TCS and its cognate ABC transporter are involved in a common physiological process. If so, the functional link should be reflected in control of the ABC transporter gene expression by its TCS partner. To investigate this possibility, and to compensate for the absence of appropriate induction signal, response regulator overproduction was used to mimic the naturally occurring stimulation conditions (Simon *et al.,* 1994; Olekhnovich e*t al.,* 1999). Each response regulator was cloned into the pDG148-Stu expression vector (Joseph *et al*., 2001) downstream the IPTG inducible pSPAC promoter. Using real-time PCR assays, the mRNA level of both NBD and MSD encoding genes was monitored in the Bs168/pDG148-Stu control strain and in Bs168/pDG148-Stu_*RR* overproducing strains. Using the same approach we also tested the putative autoregulation of the TCS gene expression by monitoring the mRNA level of the HK encoding gene in the same strains. The results summarized in Table 1 are expressed as stimulation indexes (see Methods). When bacteria were grown without IPTG, all observed stimulation indexes were very low (below 6, Table 1), indicating

Table 1. Effect of response regulator overproduction on the expression of the histidine kinase gene and the ABC transporter genes

RR overproduced		BS/YvcP		BS/YxdJ		BS/YtsA	
IPTG		−	+	−	+	−	+
Observed genes	*yvcQ*[a]	0.6 ± 0.1	0.4 ± 0.2				
	yvcR[b]	0.4 ± 0.1	50 ± 10	1 ± 0.3	2 ± 1.4	0.9 ± 0.2	4 ± 0.3
	yvcS[c]	0.7 ± 0.3	50 ± 8				
	yxdK[a]			0.9 ± 0.5	0.9 ± 0.2		
	yxdL[b]	0.7 ± 0.1	10 ± 6	6 ± 3	90 ± 30	0.7 ± 0.3	0.9 ± 0.3
	yxdM[c]			5 ± 4	110 ± 20		
	ytsB[a]					0.9 ± 0.4	0.9 ± 0.2
	ytsC[b]	0.5 ± 0.2	12 ± 9	0.5 ± 0.3	0.4 ± 0.1	0.3 ± 0.1	180 ± 38
	ytsD[c]					0.3 ± 0.04	160 ± 46

Control bacteria (Bs168/pDG148-Stu) and all RR-overproducing bacteria (Bs168/pDG148-Stu_*yvcP*, Bs168/pDG148-Stu_*ytsA*, Bs168/pDG148-Stu_*yxdJ*) were grown in LB medium until OD_{600} reached 0.7. Cells were then incubated either with (+) or without (−) IPTG for 30 minutes and harvested for total RNA preparation. For each sample, the same quantity of total RNA was used to synthesize cDNA. The cDNA preparation was then used in real-time PCR assay as described in Methods to quantify mRNA of each ABC transporter-encoding gene. Total RNA were prepared from three independent cultures and results are expressed as the mean stimulation index ± standard deviation.

[a]HK encoding gene, [b]NBD encoding gene, [c]MSD encoding gene.

that the mRNA level of either HK, NBD or MSD gene was almost the same in the control and the overproducing strains. When IPTG was added to the culture medium, no change in the HK gene expression was detected. Since the RR and HK genes are predicted to constitute an operon, this result suggests the absence of autoregulation. In contrast, in the presence of IPTG, high values of stimulation indexes were obtained (above 50) for both NBD and MSD genes expression, indicating that the mRNA levels of these genes are higher in the overproducing strains than in the control strain. These results were also confirmed by *lacZ* transcriptional fusion in the NBD encoding gene, showing a high β-galactosidase activity in overproducing bacteria compared to the control bacteria (data not shown). As the three response regulators are very similar, we also tested the possibility of cross-talk between the two-component systems. We therefore monitored the effect of the overproduction of each RR on the mRNA level of the NBD encoding genes of non-cognate ABC transporters. The results show that the overproduction of BS/YxdJ does not affect the expression of non-cognate ABC transporters as indicated by low stimulation indexes (below 2, Table 1). In contrast, the overproduction of BS/YtsA slightly affects the expression of the non-cognate ABC transporter encoded by the *yvc* cluster (stimulation index of 4, Table 1) whereas overproduction of BS/YvcP results in an increased expression of both the non-cognate ABC transporters encoded by the *yts* and *yxd* clusters (stimulation indexes of 12 and 10, respectively, Table 1). However, this effect is weaker than that observed on the cognate ABC transporter, suggesting a closer regulatory relationship between cognate partners.

Altogether, these data demonstrate that each of the three response regulators (BS/YtsA, BS/YvcP and BS/YxdJ) does not control gene expression of its corresponding HK partner and therefore is not involved in an autoregulation process and positively controls the expression of its cognate ABC transporter genes supporting the possibility that both systems (ABC transporter and TCS) be involved in the same physiological process.

Gene Clusters Encoding Two-Component Systems (With OmpR Family RR) and Sub-Family 9 ABC Transporter in Other Bacteria

To look for similar gene clusters in other bacteria, 48 bacterial genomes representing the Eubacteria/Archea diversity were analysed (http://ir2lcb.cnrs-mrs.fr/~joseph/2CABC). Some nearly complete genomes from Gram-positive bacteria, with contigs large enough to be analysed, were included in this study to counterbalance for the over-representation of Gram-negative bacterial genomes. The most highly conserved proteins of the systems, NBD from sub-family 9 and RR from OmpR family, were searched first, using characteristic sequence motifs defined as described in Methods. This led to the identification of 213 sequences of OmpR family RR and 149 sequences of family 9 NBD (http://ir2lcb.cnrs-mrs.fr/~joseph/2CABC). Their chromosomal localization was then checked to detect proximity between NBD and RR encoding genes. Finally, the surrounding sequences of these genes were analysed to identify genes encoding the MSD and HK partners, both proteins being identified through BLASTP2 searches using different proteins of both families as query. Many clusters between TCS and ABC transporter encoding genes were detected in other bacterial genomes. Unexpectedly, they were only found within the *Bacillus/Clostridium* group of the low G+C Gram positive bacteria (Table 2), *i.e. Bacillus halodurans*, C*lostridium acetobutilicum, Clostridium difficile, Lactococcus lactis, Staphylococcus aureus, Streptococcus mutans* and *Streptococcus pyogenes*.

Evolution, Comparison and Features of the Proteins Encoded by These Gene Clusters

To understand how the TCSs and their cognate ABC transporters have evolved, evolutionary trees were computed for RRs and NBDs. Among the 213 identified genes encoding RR, 33 are clustered with ABC transporter genes. 27 of them encode RRs which are found in two groups (''GI'' and ''GII'', Figure 2) and six encode RRs, which are spread out on the tree and are indicated ''out'' in Table 2. ''GI'' contains RRs from various low G+C Gram positive bacteria whereas ''GII'' only contains RRs from *Clostridium* species. Two ''GI'' RRs (SPn/5830520 and EF/R_1) have genes unassociated to ABC transporter genes. Subtle differences located in key functional regions of the proteins distinguish the RR from ''GI'' and ''GII'' groups. Some are in the receiver domain within the putative structural loop L3 predicted to interact with the HK (Hoch and Varughese, 2001). As an example, a DG residue pair is found conserved immediately adjacent to a invariant proline in the ''GII'' RR group, whereas these two residues are variable in the ''GI'' group RR. The other differences are located in the DNA-binding domain within the α-loop which has been shown to interact with component of the transcriptional machinery (Igarashi *et al.*, 1991; Pratt and Silavy,1994; Russo *et al.*, 1993); the size of this loop being one amino acid longer in RR from the ''GII'' group than in RR from ''GI'' group.

The HK partners of all the studied RR possess two putative transmembrane domains and therefore are predicted to be membrane-bound.

Among the 149 identified genes encoding NBDs, 35 are clustered with a RR gene. 27 of them encode NBDs which clustered in two groups (''9A'' and ''9B'', Figure 3) and eight encode NBDs, which do not belong to these groups. The latter have been indicated as ''out'' in Table 2. In the ''9A'' and ''9B'' groups, only a few NBDs have a gene that does not cluster with a RR gene. Strikingly, each gene encoding a ''GI'' RR is clustered with a ''9A'' NBD gene, whereas each gene encoding a ''GII'' RR is clustered with a ''9B'' NBD gene (Table 2).

To constitute a functional ABC transporter, an NBD has to interact with an MSD partner, which usually has four to six transmembrane domains (TMs) (Quentin *et al.*, 1999). Interestingly, MSD

Table 2. Summary of ABC/TCS gene clusters found in the Bacillus/Clostridium group

Bacteria	*RR*[a]	Group[b]	NBD[a]	Group[b]	TMs[c]	Gene Organization[d]	ABC/TCS gene distance (nt)
Bacillus subtilis							
	BS/YvcP	I	BS/YvcR	A	10	*RR>HK>NBD>MSD>*	85
	BS/YxdJ	I	BS/YxdL	A	10	RR>HK>NBD>MSD>	146
	BS/YtsA	I	BS/YtsC	A	9	*RR>HK>NBD>MSD>*	104
			BS/YtrE	out	4	NBD>MSD>RR>HK>	99
	BS/YclJ	out	BS/YclH	out	4	<NBD<MSD,RR>HK>	212
Bacillus halodurans							
	BH/0271	I	BH/0273	A	10	RR>HK>NBD>MSD>	115
	BH/0288	I	BH/0286	A	10	NBD>MSD>RR>HK>	34
	BH/0755	I	BH/0753	A	9	RR>HK>NBD>MSD>	149
	BH/2701	I	BH/2699	A	10[e]	RR>HK>NBD>MSD>	61
	BH/3911	I	BH/3913	A	11	RR>HK>NBD>MSD>	116
Clostridium acetobutilicum							
	CA/R_4	I	CA/N_4	A	10	RR>HK>NBD>MSD>	88
	CA/R_6	I	CA/N_6	A	9	RR>HK>NBD>MSD>	91
	CA/R_11	I	CA/N_11	A	10	RR>HK>NBD>MSD>	102
	CA/R_19	II	CA/N_19	B	8	RR>HK>NBD>MSD>	1607
	CA/R_20	II	CA/N_20	B	8	RR>HK>NBD>MSD>	82
	CA/R_2	out	CA/N_2	out	5	NBD>MSD>RR>HK>	79
Clostridium difficile[f]							
	CD/R_1	I	CD/N_1_2	out	4	RR>HK>NBD>MSD>	1345
			CD/N_1_1	A	10	HK>RR><NBD<MSD	287
	CD/R_20	I	CD/N_20	A	10	RR>HK>NBD>MSD>	1202
	CD/R_6	I	CD/N_6	A	10	NBD>MSD>RR>HK>	139
	CD/R_31	I	CD/N_31	A	10	NBD>MSD>RR>HK>	34
	CD/R_21	II	CD/N_21	B	8	RR>HK>NBD>MSD>	79
	CD/R_24	II	CD/N_24	B	8	RR>HK>NBD>MSD>	113
	CD/R_25	II	CD/N_25	B	8	RR>HK><MSD<NBD	181
	CD/R_26	II	CD/N_26	B	8	RR>HK>NBD>MSD>	67
	CD/R_32	II	CD/N_32	B	8	RR>HK>NBD>MSD>	82
	CD/R_33	II	CD/N_33	B	8	RR>HK>NBD>MSD>	81
	CD/R_35	out	CD/N_35	out	4, 4[f]	RR>MSD1>NBD>MSD2>	85
	CD/R_18	out	CD/N_18	out	8	RR>HK>NBD>MSD>	165
Lactococcus lactis							
	LL/LLRG	I	LL/YSAC	A	11	NBD>MSD>RR>HK>	1247
Staphylococcus *aureus*							
	SA/2418	I	SA/2416	A	10	RR>HK>NBD>MSD>	98
	SA/0614	I	SA/VRAF	A	10	RR>HK>NBD>MSD>	180
	SA/2151	out	SA/2149	out	4	<NBD<MSD,RR>HK>	107
Streptococcus Mutans[g]							
	SM/R_1	I	SM/N_1	A	10	NBD>MSD>RR>HK>	71
Streptococcus pyogenes							
	SP/2027	out	SP/2031	out	4	NBD>MSD>RR>HK>	93

[a] See Methods for names of RR and NBD proteins. The sequences of proteins are available on our server (http://ir2lcb.cnrs-mrs.fr/~joseph/2CABC).
[b] Groups correspond to those shown in figure 2 and 3. ''Out'' means that the protein is not included in any of these groups.
[c] Number of MSD transmembrane segments predicted with the TMpred program.
[d] Order of the genes encoding the different proteins on the chromosome. For each gene the direction of transcription is indicated by $>$ or $<$.
[e] The MSDs BH/2698 and BH/2697 are predicted to be encoded by two genes, but their products can be aligned with different parts of the same protein (as well as BS/YvcS than BS/YxdM) suggesting a sequencing error. Thus, the entire expected protein comprises 10 TMs.
[f] The *C. difficile* gene cluster encoding CD/R_35 and CD/N_35 does not encode HK, suggesting that RR either is not functional or recruits an HK from another system. Moreover, the presence of two genes encoding MSDs suggests a more complex structure of the corresponding ABC transporter.
[g] Almost completely sequenced genome.

partners of ''9A'' and ''9B'' NBDs have an unexpected higher number of predicted TMs. Indeed, MSD partners of ''9A'' NBDs have nine to eleven predicted TMs whereas every MSD partners of ''9B'' NBDs has eight predicted TMs. Therefore, in addition to a particular clustering with a ''GI'' or ''GII'' RR gene, the number of predicted TMs in the MSD partner appears to be an another feature of each NBDs groups ''9A'' and ''9B''. As indicated in Table 2, each NBD indicated ''out'' has an MSD partner with the usual number of predicted TMs *i.e.* four or five.

The most frequent genetic organization is RR, HK, NBD, and MSD encoding genes and in most cases, the TCS and ABC transporter structural genes have the same direction of transcription (Table 2). More complex structures are observed in *B. subtilis* and *C. difficile* where TCS genes are flanked on both sides by ABC transporter genes; they will be analysed later (see Discussion). The spacing between the TCS and ABC transporter structural genes is generally less than 300 nucleotides, but in four clusters it is larger (ranging from 1202 to 1607 nucleotides,

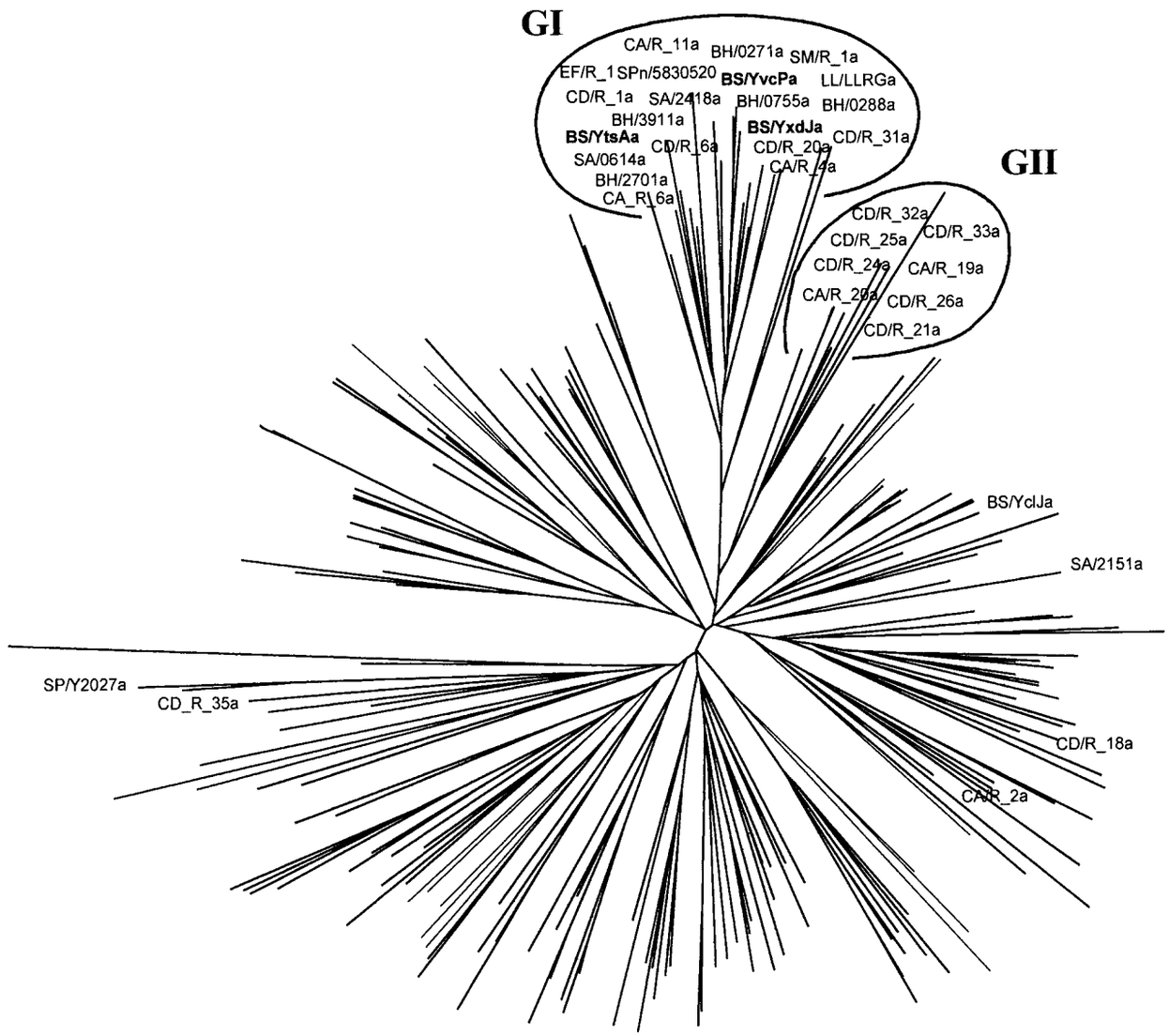


Figure 2. Unrooted tree of the OmpR family response regulators. The branch lengths are proportional to the evolutionary distances. To simplify the figure, RRs are included only if they have interesting characteristics and are discussed in the text. The letter 'a' at the end of the name indicates that RR is encoded by gene clustered with NBD genes. R, response regulator; BS, *Bacillus subtilis*; BH, *Bacillus halodurans*; CA, *Clostridium acetobutilicum;* CD *Clostridium difficle;* EF*, Enterococcus faecalis;* LL*, Lactococcus lactis;* SA*, Staphylococcus aureus;* SM*, Streptococcus mutans;* SPn*, Streptococcus pneumoniae;* SP*, Streptococcus pyogenes.*

Table 2). In two instances, an additional gene is detected in this region. Thus, the sequence between CD/R_20 and CD/N_20 encodes a protein similar to an undecaprenol kinase that confers bacitracin resistance to *Escherichia coli* (Cain *et al.,* 1993). This protein might be involved together with the TCS and ABC transporter in a common physiological process. In the case of cluster encoding LL/LLRG and LL/YSAC, a gene is also present and encodes a protein of unknown function.

Discussion

In *B. subtilis*, three closely related gene clusters — *yts*, *yvc* and *yxd*—encode TCSs with OmpR family RR and ABC transporters involving NBDs from sub-family 9. We demonstrate that in the *yts* and *yxd* clusters, TCS and ABC transporter genes were encoded by separate operon respectively. For the *yxd* cluster, these results fully agree with those obtained by Yoshida *et al.,* (2000a). Indeed, using Northern blot experiments these authors were only able to detect a specific *yxdJK* transcript.

RT-PCR analysis of the *yvc* cluster (Figure 1) suggests the existence of a unique transcript encompassing TCS and ABC transporter genes. However, this is not supported by the results we obtained using real-time PCR assay (Table 1). In fact, when comparing the control strain and the BS/YvcP overproducing strain, we detect no change in the level of the HK

encoding gene whereas we detected a drastic increase of NBD and MSD gene expression in the overproducing strain. This result indicates that at least two types of transcripts exist. The first, corresponding to the ABC transporter genes, is transcribed from an internal *yvcR* upstream transcription promoter, which could be the one predicted (Figure 1A). The second might represent a minor transcript including both the ABC transporter genes and the TCS genes and might, therefore, correspond to a readthrough of the low-efficiency *yvcQ* downstream transcription terminator (see above and Figure 1A).

Using overproduction of each response regulator and real-time PCR, we were able to demonstrate that neither BS/YtsA nor BS/YvcP nor BS/YxdJ controls the expression of its HK partner but positively controls the expression of its cognate ABC transporter genes. When testing the cross-talk between the two-component systems, we showed that i) BS/YtsA affects *yvcR* genes expression slightly, and ii) BS/YvcP increases expression of both *ytsC* and *yxdL* genes. To define the target of several response regulators, Kobayashi *et al.,* (2001) used a global approach with microarray quantification of *B. subtilis* gene expression. They systematically overproduced a RR in a strain having the cognate HK gene disrupted, because known targets of several RR were shown expressed in such conditions but not when RR was overproduced in the wild-type strain (Ogura *et al.*, 2001). The authors proposed several explanations involving the cognate HK which, without any stimulus, might function as a phosphatase on the putatively phosphorylated RR and/or might directly sequester the active multimer form of the unphosphorylated RR. In the present report, we demonstrate that each of the RR we tested is fully active when overproduced in the wild-type strain. This is probably due to a higher RR production level corresponding to a better efficiency of the vector used in the present study. Indeed, we only detected a BS/YvcP overproduction when using this vector (Joseph *et al.*, 2001) and not with the pDG148 vector (data not shown).

When using such disruptant in a cognate HK, another drawback comes from the pMUTIN insertion. This integration causes the truncated part of the sensor gene as well as downstream genes to be under the control of a pSPAC promotor (Vagner *et al.*, 1998). Thus, when inducing overproduction of a RR with IPTG, each of the previous genes will be induced as well, for example the truncated HK. This approach does not therefore allow study of (i) the possible autoregulation of the TCS or (ii) the regulation of genes located downstream of TCS gene if the corresponding gene organization (operon structure or not) is unknown. We analysed the genetic organization of the *yts* and *yvc* clusters and found that they have a two-operon structure (see above). Taking this result into account, the data of Kobayashi *et al.,* (2001) fully agree with those we obtained by real-time PCR assays showing the control of expression of the three ABC transporter genes (*ytsCD*, *yvcRS* and *yxdLM*) by their corresponding clustered RR. Both approaches gave also similar results when monitoring the cross-talk between the two-component systems. As a putative Sigma A promoter is located just upstream of each ABC transporter genes, it is likely that it serves as a target for the cognate activated response regulator for a direct control. This point is under experimental investigation.

In addition, our approach allowed us to show that in no case is there any autoregulation of a TCS.

The close link between TCSs with RR from OmpR family and sub-family 9 ABC transporter both at the regulatory and at the genetic levels, prompted us to search for genetic proximity of these systems in other bacterial genomes. Interestingly, it was only found in the *Bacillus/Clostridium* group. We did not expect such a result because the protein sequence identity shared by RR orthologs or NBD orthologs in the different genomes we have studied rarely reach 50%. Indeed, for such degree of orthologous gene divergence, a low level of conservation of the gene order in genomes is observed (Huynen and Bork, 1998). A few exceptions have been found and always concern proteins with basic cellular functions (Dandekar *et al.*, 1998). These occurrences strongly support the involvement of both systems in a major physiological process that might be a specificity of the *Bacillus/Clostridium* group.

We therefore tried (i) to classify RR and NBD into different subfamilies and (ii) to ''extract'' some particular diagnostic features of these proteins. Evolutionary trees computed with all identified OmpR family RRs, on the one hand, and all identified sub-family 9 NBDs, on the other hand, revealed that: (i) most RRs encoded by genes clustered with NBD genes constitute two distinct groups (''GI'' and ''GII''); these RRs clearly share the same direct ancestor (Figure 2), (ii) most NBDs encoded by genes clustered with RR genes also constitute two distinct groups (''9A'' and ''9B''); however, NBDs from ''9A'' and ''9B'' show independent branches on the tree (Figure 3) suggesting that they did not evolve from the same direct ancestor. (iii) all ''GI'' RR genes are clustered with ''9A'' NBD genes while all ''GII'' RR genes are clustered with ''9B'' NBD genes. In addition, an unexpected elevated number of predicted transmembrane domains in the MSD partners was observed that might constitute an additional characteristic of ABC transporters with ''9A'' or ''9B'' NBDs (Table 2).

We therefore propose the following chronological events for the evolution of the observed clusters. The ancestral TCS gene first duplicated and gave rise to two TCS progenitor genes. These two TCS genes then independently recruited in their vicinity two different ABC transporter genes in such a way that a ''GI'' RR gene was clustered with a ''9A'' NBD gene while a ''GII'' RR gene was clustered with a ''9B'' NBD gene. Subsequently, the TCS/ABC transporter gene tandems were amplified as a unit in bacteria of the *Bacillus*/*Clostridium* group. If we assume that no horizontal genetic transfer has occurred during the process, the origin of the clusters might have been concomitant with the emergence of this lineage. A last step concerned few clusters in which either some genes were lost or local chromosomal modifications resulted in changes in the genetic organisation.

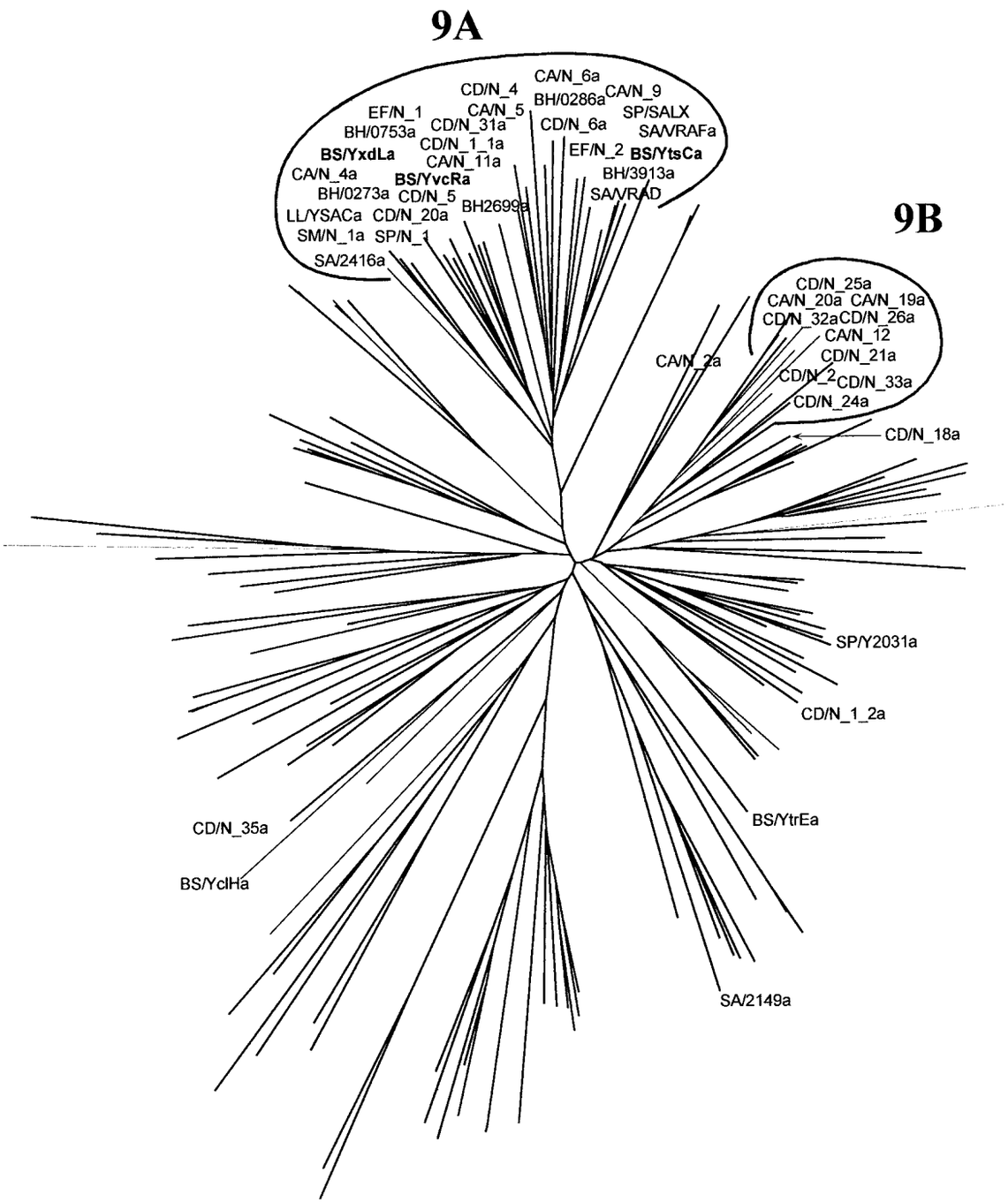


Figure 3. Unrooted tree of the Nucleotide Binding Domains from sub-family 9. The branch lengths are proportional to the evolutionary distances. Group ''9A'' is supported by a bootstrap value of 70%. To simplify the figure, NBDs are included only if they have interesting characteristics and are discussed in the text. The letter 'a' at the end of the name indicates that NBD is encoded by gene clustered with TCS genes. N, NBD; BS, *Bacillus subtilis*; BH, *Bacillus halodurans*; CA, *Clostridium acetobutilicum;* CD *Clostridium difficle;* EF, *Enterococcus faecalis;* LL, *Lactococcus lactis;* SA, *Staphylococcus aureus;* SM, *Streptococcus mutans;* SPn, *Streptococcus pneumoniae;* SP, *Streptococcus pyogenes.*

We have applied the diagnostic features (see above) to the two peculiar complex gene clusters observed in *B. subtilis* and *C. difficile*. In *B. subtilis* the gene encoding the BS/YtsA response regulator is bracketed by the gene encoding the NBD BS/YtsC, on one side, and by the gene encoding the NBD BS/YtrE, on the other side (Table 2). We propose that the correct association is that comprising the BS/YtsA and the BS/YtsC genes, the proximity of the BS/YtrE gene being purely coincidental. This is supported by the following points; (i) according to our proposition, the ''GI'' RR BS/YtsA should be associated with a NBD

belonging to the ''9A'' group and the partner MSD should have nine to eleven predicted transmembrane segments; BS/YtsC has these features while BS/YtrE does not, (ii) BS/YtsA was shown to control BS/YtsC gene expression (this report) but not BS/YtrE gene expression (Kobayashi *et al.*, 2001). In fact the BS/YtrE structural gene is part of an operon that encodes an ABC transporter that is involved in acetoin utilization and is negatively regulated by the operon first gene product (Yoshida *et al.*, 2000b). Applying the same reasoning to the complex gene cluster observed in *C. difficile,* we predict a true clustering between the ''G1'' RR CD/R_1 gene and the ''9A'' NBD gene (Table 2), the proximity of the ''out'' NBD structural gene being coincidental. In *B. subtilis*, we have demonstrated that each of the three related ''G1'' response regulators—BS/YtsA, BS/YvcP and BS/YxdJ—modulates the expression of its cognate ABC transporter genes (with ''9A'' NBD). Interestingly, BS/YclH gene expression has also been shown to be modulated by its cognate partner BS/YclJ (Kobayashi *et al.*, 2001). So, the four clustered TCS (with RR from the OmpR family) and sub-family 9 ABC transporters found in *B. subtilis* (Table 2) are regulatorily linked. We therefore propose that a similar control occurs for the gene expression of all other clustered systems found in the *Bacillus/Clostridium* group identified in this study.

In bacteria, genetic and functional association between two-component and ABC transport systems have already been described. Those for which ABC transporter genes were shown to be controlled by a TCS, both systems having genes in close proximity, are involved in quorum-sensing dealing with peptide antibiotic production (see for review Nes *et al.*, 1996; Kleerebezem *et al.*, 1997) in which the ABC transporter ensure either antibiotic transport or immunity against antibiotic. They include the production systems for nisin (Siegers and Entian, 1995; Kuipers *et al.*, 1995; Ruyter *et al.*, 1996), bacitracin (Neumüller *et al.*, 2001), mersacidin (Altena *et al.*, 2000), sakacin A (Axelsson and Holck, 1995) and a bacteriocin-like peptide (Saizieu *et al.*, 2000).

In the present study, we have identified and characterized new TCS/ABC clustered genes that are only found in the *Bacillus/Clostridium* group. For some of them, TCS regulates ABC transporter gene expression (this report and Kobayashi *et al.*, 2001). However, neither the signal activating the TCSs nor the substrate transported by their cognate ABC transporters is known. It is worthy of note that BS/YxdL, BS/YtsC and BS/YvcR present similarities with other members of the sub-family 9 NBD as SalX from *Streptococcus salivarius* and MbrA from *Streptococcus mutans*, these two proteins being likely involved in salivaricin transport (Upton *et al.*, 2001) and bacitracin resistance respectively. In addition, recent macroarray analysis of *B. subtilis* transcriptome indicated that expression of *ytsD* and *yvcS* genes were slightly modified by glucose; whether this is a physiologically relevant awaits futher studies (Moreno *et al*, 2001; Saier *et al.*, 2002). On the other hand, more attention will be paid to the special case of the *B. subtilis yxd* cluster. Actually, a small ORF, *yxeA*, is cotranscribed with the upstream ABC transporter genes (Figure 1C) and this gene organization is conserved in *L. lactis* (not shown). As *yxeA* is predicted to encode a protein with a cleavable leader peptide, it is of interest to determine which physiological pathway involves this putative extracellular polypeptide and whether it plays a functional role together with the TCS and ABC transporter.

Experimental Procedures

Bacterial Strains and Plasmids

The *B. subtilis* recombinant bacteria used in this study derived from strain 168 *trp* C2. The Bs168/pDG148-Stu control strain contains the expression vector pDG148-Stu (Joseph *et al.,* 2001). The Bs168/pDG148-Stu_*yvcP*, Bs168/pDG148-Stu_*ytsA* or Bs168/pDG148-Stu_*yxdJ* strains contain the plasmid pDG148-Stu in which *yvcP*, *ytsA* or *yxdJ* has been cloned, respectively. The cloning procedure was as in Joseph *et al.,* (2001).

Media

All bacteria were grown at 37°C in Luria-Bertani medium (LB). Recombinant strains were grown in the presence of 20 μg ml^{-1} of kanamycin. When needed, IPTG (Interchim, Montluçon, France) was added to the culture at 1 mM final concentration and cells were harvested after 30 minutes of induction.

Total RNA Purification and cDNA Preparation

Total RNA were prepared from exponentially growing cultures using the High Pure RNA isolation Kit™ from Roche Diagnostics (Mannheim, Germany) according to the supplier's recommendations, except that two steps of DNase treatment were used instead of one. The absence of genomic DNA contamination was checked by PCR in each RNA preparation. cDNA was synthesized using SuperScript II™ from GibcoBRL (Life Technology, Cergy, France) following the protocol described by the manufacturer. For each sample, we used 1μg of total RNA together with 100 ng of random hexamer (GibcoBRL) as primers for each reverse transcription reaction. The cDNA preparation was then used as template for PCR or Real-time PCR as described below.

PCR

For the identification of transcriptional unit, PCR amplifications were carried out using cDNA as template. The primers used for the amplification of HKs were:

yvcQ_fwd, 5′-ATAGGGTCCGCCATGCTGATT-3′;
yvcQ_rev, 5′-TGGACCCATTGGTTGGTGAAG-3′;
yxdK_fwd, 5′-CCGTATCTCGGTTCTTCAGTGTTC-3′;
yxdK_rev, 5′-ACTGATTGGAGAAGCTCGGACAG-3′;
ytsB_fwd, 5′-GGGAGAACAATCTCGATGTGACAG-3′;
ytsB_rev, 5′-AGCTGCTGATCAAGAAGCAGGTG-3′;

for NBDs:

yvcR1_fwd, 5′-CCAAGAGCTGGCTTTGTTTCG-3′;
yvcR2_rev, 5′-TCGAGGTTTCCTGTCGGTTCA-3′;
yxdL1_fwd, 5′-GGTCATTCATAAGCCGTCACTC-3′;
yxdL2_rev, 5′-CCAGCATAGACAGCACATCGAG-3′;
ytsC1_fwd, 5′-GTCCTCAATCGATCAGGTCAGTC-3′;
ytsC2_rev, 5′-CTCTTCCAGCGGATGTTCTCTGC-3′;

for intergenic region between HK and NBD encoding genes:

yvcQR_fwd, 5′AAACAGGTGTGCAGCCGTCTC-3′;
yvcQR_rev, 5′-AAAGCGTGGTTTTTCCGCTTC-3′;
yxdKL_fwd, 5′-CTCAATCATACGGTGGACATCAGC-3′;
yxdKL_rev, 5′-CCGTGAACTCGCCTTCTTCTATTG-3′;
ytsBC_fwd, 5′-ATGTTGAGTCGGAGTTTGGTG-3′;
ytsBC_rev, 5′-CGGAAGCACCCATAATACTGA-3′;

for MSD:

yxdM_fwd, 5′-CTGCCGAGCCTGAGTCTATTTATC-3′;
yxdM_rev, 5′-GAAAAACCCTCCGATGGTGAGAC-3′;

for the amplification of the region spanning *yxdM* and *yxeA* we used yxdM_fwd (see above) as forward primer with yxeA_rev, 5′-TACGCG TTCTTCCGAAGCTC-3′ as reverse primer.

Reactions were carried out for 30 cycles at 94°C for 30 s, 55°C for 30 s and 72°C for 30 s.

Real-Time PCR Assays

Real-time PCR assays (Higuchi *et al.,* 1992; Higuchi *et al.,* 1993) were done using a LightCycler™ system (Roche Diagnostics, Mannheim, Germany) according to the manufacturer's instructions. cDNA were amplified in the presence of SYBR-Green™ (Fast Start DNA Master SYBR Green I™) in a 20 µl final volume with optimised primers and $MgCl_2$ concentrations. The amplification procedure comprises a 8 min denaturation at 95°C for Taq DNA polymerase activation and 45 cycles including the three following steps: denaturation at 95°C for 15 s, annealing at 60°C for 6 s and extension at 72°C for 10 s. The set of primers used to amplify cDNA are listed below. They were chosen according to the gene sequence described at the NCBI site.

ytsB_fwd, 5′-AAGGGATCGGTTTTGACATACAG-3′;
ytsB_rev, 5′-ATAAACGCCAGCCATTTTGC-3′;
ytsC_fwd, 5′-TTGCGGGATAAATACCCGAA-3′;
ytsC_rev, 5′-GCTCTTCCAGCGGATGTTCTC-3′;
ytsD_fwd, 5′-TGTCTTCGATCATGTTCAGAATGA-3′;
ytsD_rev, 5′-GAGCGCTGAAACGGTTGTAAT-3′;
yvcQ_fwd, 5′-GATTTCGCCACGGTCTTGATA-3′;
yvcQ_rev, 5′-ACATCAAATGCAAATTCCTCCA-3′;
yvcR_fwd, 5′-CATATTCTCGATCACCGCACA-3′;
yvcR_rev, 5′-AATAATCGCTCTGGCACAAGC-3′;
yvcS_fwd, 5′-CAGTCTGTTTCCTGGGTGGAA-3′;
yvcS_rev, 5′-TCAATGTATTCTCAGCCGTCTCA-3′;
yxdK_fwd, 5′-GCTTCTTATTTTGGCGGGCTA-3′;
yxdK_rev, 5′-AGCTGAGCCAATGATACACGC-3′;
yxdL_fwd, 5′-CAAGCAGCTCGGATTTGTGTT-3′;
yxdL_rev, 5′-GGCAGCATGATATTTTCACCG-3′;
yxdM_fwd, 5′-CCGTGCTTCTGTCCTTATTCG-3′;
yxdM_rev, 5′-CATTGCCTTTAAGAACCATGCC-3′;

Results are expressed as stimulation index *i.e.* mRNA amount of a gene expressed in the Bs168/pDG148-Stu_*RR* strain harbouring the cloned response regulator divided by the mRNA amount of the same gene expressed in the control strain Bs168/pDG148-Stu containing the plasmid alone.

Bioinformatic Analysis

The sequence motifs of OmpR family response regulator were constructed from a sample of OmpR RR sequences retrieved from the PRODOM database (Gouzy *et al.*, 1999) and were computed with the MEME program (Bailey and Elkan, 1994). Motifs characteristic of sub-family 9 NBD were available in the database devoted to ABC transporters (http://ir2lcb.cnrs-mrs.fr/ABCdb; Quentin and Fichant, 2000). The MAST program was used to search for these motifs (Bailey and Gribskov, 1998) in the different proteomes of completely sequenced genomes or in the set of identified ORFs (longer than 50 codons) in the almost completely sequenced genomes. MSDs and HKs were identified with the BLASTP2 program (Altschul *et al.,* 1997) using different proteins of both families as query. For the complete and unfinished genomes, we used data available at the NCBI site (http://www.ncbi.nlm.nih.gov/PMGifs/Genomes/micr.html) and at the TIGR site (http://www.tigr.org/tdb/mdb/mdb.html). The TMpred program (http://www.ch.embnet.org/software/TMPRED_form.html) was used to predict the number and location of transmembrane regions of MSD and HK. Multiple alignments have been obtained with ClustalW (Thompson *et al.*, 1994). The distance matrices were derived from the multiple alignments with the PRODIST method and the PAM option (Felsenstein, 1989). Trees were computed with the neighbour-joining method (Saitou and Nei, 1987) using the PHYLO_WIN program and unrooted trees were drawn with TreeTool (from Mike Maciukenas at http://www.hgmp.mrc.ac.uk/Registered/Option/treetool.html).

In this report we use the following rule to name the proteins of interest: the first two letters characterize the organism (list of abbreviations and organisms available on our site, http://ir2lcb.cnrs-mrs.fr/~joseph/2CABC) and then a slash precedes the name of the protein as listed in databases. For unannotated genomes, a letter characterizes the gene product (R: response regulator, N: NBD) and a number differentiates those identified in the same genome.

Putative RNA secondary structures with their calculated ΔG° values were identified with MFOLD (Mathews *et al.*, 1999) and putative promoters were identified using the PatScan program (http://www-unix.mcs.anl.gov/compbio/PatScan/HTML/patscan.html) with a matrix defined motif. Matrix was derived from a set of known sigma A promoter sequences of *B. subtilis* (Helmann, 1995).

Acknowledgements

The authors gratefully acknowledge Marc Chippaux, Vincent Méjean and Anne Galinier for helpful discussions and Athel Cornish-Bowden for critical reading of the manuscript. We also thank Annick Guiseppi and Jeannine Busuttil for their advice during the course of this study. This work was supported in part by grants from CNRS and MRT fellowships, followed by a fellowships from the Fondation pour la Recherche Médicale to P. J.

References

Altena, K., Guder, A., Cramer, C., and Bierbaum, G. 2000. Biosynthesis of the lantibiotic mersacidin: organization of a type B lantibiotic gene cluster. Appl. Environ. Microbiol. 66: 2565–2571.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25: 3389–3402.

Ames, G.F. 1986. Bacterial periplasmic transport systems: structure, mechanism, and evolution. Annu. Rev. Biochem. 55: 397–425.

Axelsson, L., and Holck, A. 1995. The genes involved in production of and immunity to sakacin A, a bacteriocin from *Lactobacillus sake* Lb706. J. Bacteriol. 177: 2125–2137.

Bailey, T.L., and Elkan, C. 1994. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc. Int. Conf. Intell. Syst. Mol. Biol. 2: 28–36.

Bailey, T.L., and Gribskov, M. 1998. Combining evidence using p-values: application to sequence homology searches. Bioinformatics. 14: 48–54.

Beier, D., and Frank, R. 2000. Molecular characterization of two-component systems of *Helicobacter pylori*. J. Bacteriol. 182: 2068–2076.

Braibant, M., Gilot, P., and Content, J. 2000. The ATP binding cassette ABC transport systems of *Mycobacterium tuberculosis*. FEMS Microbiol. Rev. 24: 449–467.

Cain, B.D., Norton, P.J., Eubanks, W., Nick, H.S., and Allen, C.M. 1993. Amplification of the bacA gene confers bacitracin resistance to *Escherichia coli*. J. Bacteriol. 175: 3784–3789.

Dandekar, T., Snel, B., Huynen, M., and Bork, P. 1998. Conservation of gene order: a fingerprint of proteins that physically interact. Trends Biochem. 23: 324–328.

Dassa, E., Hofnung, M., Paulsen, I.T., and Saier, M.H. Jr. 1999. The *Escherichia coli* ABC transporters: an update. Mol. Microbiol. 32: 887–889.

Fabret, C., Feher, V.A., and Hoch, J.A. 1999. Two-component signal transduction in *Bacillus subtilis*: how one organism sees its world. J. Bacteriol. 181: 1975–1983.

Fath, M.J., and Kolter, R. 1993. ABC transporters: bacterial exporters. Microbiol. Rev. 57: 995–1017.

Felsenstein, J. 1989. PHYLIP - Phylogeny Inference Package version 3.2. Cladistics. 5: 164–166.

Galinier, A., Haiech, J., Kilhoffer, MC., Jaquinod, M., Stulke, J., Deutscher, J., and Martin-Verstraet, I. 1997. The *Bacillus subtilis* crh gene encodes a HPr-like protein involved in carbon catabolite repression. Proc. Natl. Acad. Sci. USA 94: 8439–8444.

Gouzy, J., Corpet, F., and Kahn, D. 1999. Whole genome protein domain analysis using a new method for domain clustering. Computers and Chemistry. 23: 333–340.

Helmann, J.D., 1995. Compilation and analysis of *Bacillus subtilis* sigma A-dependent promoter sequences: evidence for extended contact between RNA polymerase and upstream promoter DNA. Nucleic Acids Res. 23: 2351–2360.

Higgins, C.F. 1992. ABC transporters: from microorganisms to man. Annu. Rev. Cell. Biol. 8: 67–113.

Higgins, C.F., Hiles, ID, Salmond, G.P., Gill, D.R., Downie, J.A., Evans, I.J., Holland, I.B., Gray, L., Buckel, S.D., Bell, A.W., and Hermodson, M.A. 1986. A family of related ATP-binding subunits coupled to many distinct biological processes in bacteria. Nature. 323: 448–450.

Higuchi, R., Dollinger, G., Walsh, P.S., and Griffith, R. 1992. Simultaneous amplification and detection of specific DNA sequences. Biotechnology. 10: 413–417.

Higuchi, R., Fockler, C., Dollinger, G., and Watson, R. 1993. Kinetic PCR analysis: real-time monitoring of DNA amplification reactions. Biotechnology. 11: 1026–1030.

Hoch, J.A., and Varughese, K.I. 2001. Keeping signals straight in phosphorelay signal transduction. J Bacteriol. 183: 4941–4949.

Huynen, M.A., and Bork, P. 1998. Measuring genome evolution. Proc. Natl. Acad. Sci. USA. 95: 5849–5856.

Igarashi. K., Hanamura. A., Makino. K., Aiba H., Mizuno. T., Nakata. A., and Ishihama. A. 1991. Functional map of the alpha subunit of *Escherichia coli* RNA polymerase: two modes of transcription activation by positive factors. Proc. Natl. Acad. Sci. USA. 88: 8958–8962.

Joseph, P., Fantino, J-R., Herbaud, M-L., and Denizot, F. 2001. Rapid orientated cloning in a shuttle vector allowing modulated gene expression in *Bacillus subtilis*. FEMS Microbiol. Lett. 205: 91–97.

Kleerebezem, M., Quadri, L.E., Kuipers, O.P., and de Vos, W.M. 1997. Quorum sensing by peptide pheromones and two-component signal-transduction systems in Gram-positive bacteria. Mol. Microbiol. 24: 895–904.

Kobayashi, K., Ogura, M., Yamaguchi, H., Yoshida, K., Ogasawara, N., Tanaka T., and Fujita, Y. 2001. Comprehensive DNA microarray analysis of *Bacillus subtilis* two-component regulatory systems. J. Bacteriol. 183: 7365–7370.

Kuipers, O.P., Beerthuyzen, M.M., de Ruyter, P.G., Luesink, E., and de Vos, W.M. 1995. Autoregulation of nisin biosynthesis in *Lactococcus lactis* by signal transduction. J. Biol. Chem. 270: 27299–27304.

Linton, K.J., and Higgins, C.F. 1998. The *Escherichia coli* ATP-binding cassette ABC proteins. Mol. Microbiol. 28: 5–13.

Mathews, D.H., Sabina, J., Zuker, M., and Turner, D.H. 1999. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. J. Mol. Biol. 288: 911–940.

Mizuno, T. 1997. Compilation of all genes encoding two-component phosphotransfer signal transducers in the genome of *Escherichia coli*. DNA Res. 28: 161–168.

Moreno, M.S, Schneider, B.L, Maile, R.R, Weyler, W., and Saier, M.H., Jr. 2001. Catabolite repression mediated by the CcpA protein in *Bacillus subtilis*: novel modes of regulation revealed by whole-genome analyses. Mol. Microbiol. 39: 1366–1381.

Nes, I.F, Diep, D.B, Havarstein, L.S, Brurberg, M.B, Eijsink, V., and Holo, H. 1996. Biosynthesis of bacteriocins in lactic acid bacteria. Antonie Van Leeuwenhoek. 70: 113–128.

Neumuller, A.M., Konz, D., and Marahiel, M.A. 2001. The two-component regulatory system BacRS is associated with bacitracin 'self-resistance' of *Bacillus licheniformis* ATCC 10716. Eur. J. Biochem. 268: 3180–3189.

Ogura, M., Yamaguchi, H., Yoshida, Ki, Fujita, Y., and Tanaka, T. 2001. DNA microarray analysis of *Bacillus subtilis* DegU, ComA and PhoP regulons: an approach to comprehensive analysis of *B.subtilis* two-component regulatory systems. Nucleic Acids Res. 29: 3804–3813.

Olekhnovich, I.N., Dahl, J.L., and Kadner, R.J. 1999. Separate contributions of UhpA and CAP to activation of transcription of the *uhpT* promoter of *Escherichia coli*. J. Mol. Biol. 292: 973–986.

Parkinson, J.S., and Kofoid, E.C. 1992. Communication modules in bacterial signaling proteins. Annu. Rev. Genet. 26: 71–112.

Paulsen, I.T., Nguyen, L., Sliwinski, M.K., Rabus, R., and Saier, M.H., Jr. 2000. Microbial genome analyses: comparative transport capabilities in eighteen prokaryotes. J. Mol. Biol. 301: 75–100.

Pratt, L.A., and Silhavy, T.J. 1994. OmpR mutants specifically defective for transcriptional activation. J. Mol. Biol. 243: 579–594.

Quentin, Y., Fichant, G., and Denizot, F. 1999. Inventory, assembly and analysis of *Bacillus subtilis* ABC transporter systems J. Mol. Biol. 287: 467–484.

Quentin, Y., and Fichant, G. 2000. ABCdb: an ABC Transporter Database. J. Mol. Microbiol. Biotechnol. 2: 501–504.

Rodrigue, A., Quentin, Y., Lazdunski, A., Mejean, V., and Foglino, M. 2000. Two-component systems in *Pseudomonas aeruginosa*: why so many? Trends Microbiol. 8: 498–504.

de Ruyter, P.G., Kuipers, O.P., Beerthuyzen, M.M., van Alen-Boerrigter, I., and de Vos, W.M. 1996. Functional analysis of promoters in the nisin gene cluster of *Lactococcus lactis*.
J. Bacteriol. 178: 3434–3439.

Russo, F.D., Slauch, J.M., and Silhavy, T.J. 1993. Mutations that affect separate functions of OmpR the phosphorylated regulator of porin transcription in *Escherichia coli*. J. Mol. Biol. 231: 261–273.

Saier, M.H., Jr., Goldman, S.R., Maile, R.R., Moreno, M.S., Weyler, W., Yang, N., and Paulsen, I.T. 2002. Transport capabilities encoded within the *Bacillus subtilis* genome. J. Mol. Microbiol. Biotechnol. 4: 37–67.

Saitou, N., and Nei, M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4: 406–425.

de Saizieu, A., Gardes, C., Flint, N., Wagner, C., Kamber, M., Mitchell, T.J., Keck, W., Amrein, K.E., and Lange, R. 2000. Microarray-based identification of a novel *Streptococcus pneumoniae* regulon controlled by an autoinduced peptide. J. Bacteriol. 182: 4696–4703.

Siegers., K, and Entian, K.D. 1995. Genes involved in immunity to the lantibiotic nisin produced by *Lactococcus lactis* 6F3. Appl. Environ. Microbiol. 61: 1082–1089.

Simon, G., Mejean, V., Jourlin, C., Chippaux, M., and Pascal, M.C. 1994. The *torR* gene of *Escherichia coli* encodes a response regulator protein involved in the expression of the trimethylamine N-oxide reductase genes. J. Bacteriol. 176: 5601–5606.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22: 4673–4680.

Throup, J.P., Koretke, K.K., Bryant, A.P., Ingraham, K.A., Chalker, A.F., Ge, Y., Marra, A., Wallis, N.G., Brown, J.R., Holmes, D.J., Rosenberg, M., and Burnham, M.K. 2000. A genomic analysis of two-component signal transduction in *Streptococcus pneumoniae*. Mol. Microbiol. 35: 566–576.

Tomii, K., and Kanehisa, M. 1998. A comparative analysis of ABC transporters in complete microbial genomes. Genome Res. 8: 1048–1059.

Upton, M., Tagg, J.R., Wescombe, P., and Jenkinson, H.F. 2001. Intra and interspecies signaling between *Streptococcus salivarius* and *Streptococcus pyogenes* mediated by SalA and SalA1 lantibiotic peptides. J. Bacteriol. 183: 3931–3938.

Vagner, V., Dervyn, E., and Ehrlich, S.D. 1998. A vector for systematic gene inactivation in *Bacillus subtilis*. Microbiology. 144: 3097–3104.

Yoshida, K.I., Aoyama, D., Ishio, I., Shibayama, T., and Fujita, Y. 1997. Organization and transcription of the myo-inositol operon, *iol*, of *Bacillus subtilis*. J. Bacteriol. 179: 4591–4598.

Yoshida, K., Ishio, I., Nagakawa, E., Yamamoto, Y., Yamamoto, M., and Fujita, Y. 2000a. Systematic study of gene expression and transcription organization in the *gntZ-ywaA* region of the *Bacillus subtilis* genome. Microbiology. 146: 573–957.

Yoshida, K.I., Fujita, Y., and Ehrlich, S.D. 2000b. An operon for a putative ATP-binding cassette transport system involved in acetoin utilization of *Bacillus subtilis*. J. Bacteriol. 182: 5454–5461.