Mini-Review

Making sense of quorum sensing in lactobacilli: a special focus on Lactobacillus plantarum WCFS1

Mark H. J. Sturme,1,2 Christof Francke,1,4 Roland J. Siezen,1,3,4 Willem M. de Vos1,2 and Michiel Kleerebezem1,2,3

Correspondence
Mark H. J. Sturme
mark.sturme@wur.nl

1TI Food and Nutrition, Nieuwe Kanaal 9A, 6709 PA Wageningen, The Netherlands
2Laboratory of Microbiology, Wageningen University, Dreijenplein 10, 6703 HB Wageningen, The Netherlands
3NIZO Food Research BV, 6710 BA Ede, The Netherlands
4Radboud University Nijmegen Medical Centre/NCMLS, CMBI 260, PO Box 9101, 6500 HB Nijmegen, The Netherlands

In silico identification criteria were defined to predict if genes encoding histidine protein kinases (HPKs) and response regulators (RRs) could be part of peptide-based quorum sensing (QS) two-component regulatory systems (QS-TCSs) in Firmicutes. These criteria were used to screen HPKs and RRs annotated on the completed genome sequences of Lactobacillus species, and several (putative) QS-TCSs were identified in this way. The five peptide-based QS-TCSs that were predicted on the Lactobacillus plantarum WCFS1 genome were further analysed to test their (QS) functionality. Four of these systems contained an upstream gene encoding a putative autoinducing peptide (AIP), of which two were preceded by a double-glycine-type leader peptide. One of these was identical to the plnABCD regulatory system of L. plantarum C11 and was shown to regulate plantaricin production in L. plantarum WCFS1. The third TCS was designated lamBDCA for Lactobacillus agr-like module, where the lamD gene was shown to encode a cyclic thiolactone peptide. The fourth TCS was paralogous to the lam system and contained a putative AIP-encoding gene but lacked the lamB gene. Finally, a genetically separated orphan HPK and RR that showed clear peptide-based QS characteristics could form a fifth peptide-based QS-TCS. The predicted presence of multiple (peptide-based) QS-TCSs in some lactobacilli and in particular in L. plantarum might be a reflection of the ability of these species to persist in a diverse range of ecological niches.

Abbreviations: AIP, autoinducing peptide; HPK, histidine protein kinase; QS, quorum sensing; RR, response regulator; TCS, two-component system; TMS, transmembrane segment.

Three supplementary figures are available with the online version of this paper.

Two-component regulatory systems in lactic acid bacteria

The lactic acid bacteria (LAB) comprise a diverse group of Gram-positive bacteria that are applied in the production of fermented food products such as dairy, meat and vegetable products (Caplice & Fitzgerald, 1999). Several LAB are also found in the gastrointestinal tracts of humans and other animals (Vaughan et al., 2002). Among the LAB the genus Lactobacillus forms a large and diverse group, and several Lactobacillus strains are considered to exert health-promoting effects in man and animals (Ouweland et al., 2002). The genome sequences of a number of Lactobacillus species from different ecological niches are currently available (Makarova et al., 2006; Makarova & Koonin, 2007; Siezen et al., 2004). This enables a comparative genomics analysis of lactobacilli that are restricted to a specific niche and have limited physiological abilities, such as Lactobacillus johnsonii in the human gastrointestinal tract (Priddle et al., 2004) and Lactobacillus delbrueckii subsp. bulgaricus in dairy products (Van de Guchte et al., 2006), with more adaptable species such as Lactobacillus plantarum (Kleerebezem et al., 2003), which is found in fermented food products, on plant material (Caplice & Fitzgerald, 1999) and as a natural inhabitant of the human gastrointestinal tract (Ahrne et al., 1998). To allow for efficient colonization and persistence or effective adaptation to changing environmental conditions, lactobacilli require sensory systems to detect (specific) environmental signals. In bacteria, this function is commonly mediated by two-component regulatory systems (TCSs), which consist of a membrane-located histidine protein kinase (HPK) that monitors one or more environmental factors, and a cytoplasmic response regulator (RR), which modulates expression of specific genes. The HPK and RR
function as a phospho-relay signal-transduction system (Fig. 1), and the genes encoding the cognate HPK and RR are generally organized in an operon structure (Hoch & Silhavy, 1995). TCs monitor and respond to changes in environmental conditions such as osmolarity, nutrient availability (C, N, P) or temperature (Hoch & Silhavy, 1995). However, in Gram-positive species, TCs are also known to respond to specific secreted signaling molecules involved in quorum sensing (QS) (Sturme et al., 2002).

QS-TCSs in Gram-positive bacteria regulate the expression of genes involved in diverse functions such as virulence, competence or bacteriocin production (Kleerebezem et al., 1997). This modulation is done in a coordinated and cell-density-dependent manner, using specific autoinducing signalling peptides (AIPs) that are often post-translationally modified and exported by dedicated transport systems (Ansaldo & Dubnau, 2004; Håvarstein et al., 1995; Zhang & Ji, 2004), and sensed by responsive cells via dedicated HPKs. Bacteria may contain multiple QS-TCSs, underlining the importance of intercellular communication.

This mini-review describes in silico criteria for the identification of (putative) peptide-based QS-TCSs in lactobacilli and other Firmicutes. In addition, it describes the application of those criteria in the identification of QS-TCSs in the model organism L. plantarum WCFS1 (Kleerebezem et al., 2003). Detailed analysis and experimental confirmation of the functionality of predicted peptide-based QS-TCSs in L. plantarum WCFS1 is also discussed.

Fig. 1. General signal transduction mechanism of two-component regulatory systems. The cellular localization and sequence motifs characteristic of two-component histidine protein kinases of the HPK10 subfamily (membrane-bound) and response regulators (cytoplasmic) are depicted (see also Table 1). TMS, transmembrane segments I–VII. Homology boxes of the HPK10 subfamily as described in the text are: H, X, N and G. Conserved residues in the N-terminal CheY-like receiver domain of response regulators as described in the text are DD, D, and K.

Protein architecture of peptide-based QS-TCSs

A distinctive functional feature of TCs is the ability to transfer a phosphoryl group from ATP to a receptor protein. As a result, the two ‘components’ of the system contain characteristic functional domains. The HPKs have a conserved C-terminal ATP-binding domain in which the phosphoryl-accepting histidine residue is located, and related to that they also have highly conserved clusters of residues called homology boxes (Grebe & Stock, 1999; Parkinson & Kofoid, 1992). Based on the presence and structure of the various homology boxes, a comprehensive classification of HPKs was made by Grebe & Stock (1999). In this classification the vast majority of peptide-based QS-TCSs comprise specific HPKs belonging to the subfamily called HPK10. The only exception known to date is ComP of Bacillus subtilis, which is a QS-HPK that belongs to a different HPK subfamily (HPK7). The HPK10 homology boxes have the following characteristics: the H-box (histidine phosphorylation site) contains a tyrosine two residues downstream from the conserved histidine (characteristic motif F+HDYxN) and lacks an otherwise conserved proline residue at position 5 downstream; the X-box (a hydrophobicity pattern conserved in several subfamilies) is present; the N-box has only one conserved asparagine residue (characteristic motif DNAIE); and the G-box, which plays a critical role in phosphoryl transfer, has a characteristic FSTKxGxGLGL motif (Fig. 1). The D-box that in most subfamilies is part of the nucleotide-binding domain is absent in this subfamily. Furthermore, HPK10 subfamily members commonly possess five to seven N-terminal transmembrane segments (TMSs) (Fig. 1).

In general, RRs contain a C-terminal DNA-binding domain and a N-terminal CheY-like receiver domain (REC: SM00448) that includes a conserved phosphoryl-accepting aspartate residue (Volz, 1993) (Fig. 1). Analogous to the HPKs, the RRs have been classified based on the receiver and the DNA-binding domains (Grebe & Stock, 1999). In both classifications the RRs related to HPKs of the HPK10 subfamily comprise a separate subfamily: the RD and ComE subfamily, respectively. Most established QS-RRs are encompassed within this RD/ComE subfamily of RRs. However, in accordance with the unusual sequence of its cognate HPK, the competence-regulating RR (ComA) of B. subtilis belongs to the RD rather than the RD subfamily of RRs and contains a HTH-LuxR DNA-binding domain (SM00421) (Fuqua et al., 1994). In a more recent analysis most RRs of the RD/ComE subfamily were classified in the LytTR family of response regulators (PF04397), based on a conserved motif in the C-terminal helix–turn–helix (HTH) DNA-binding domain (Nikolskaya & Galperin, 2002). For RRs the presence of a HTH-LytTR DNA-binding domain does not in all cases classify a RR as a peptide-based QS-RR (Nikolskaya & Galperin, 2002). However, the presence of an adjacent HPK with HPK10 subfamily characteristics can be used to classify the RR as such.
Genomic context: linkage of peptide-based QS-TCSs with AIPs and other QS-related functionalities

In Gram-positive bacteria, genes encoding peptide-based QS-TCSs are in general preceded by genes encoding the cognate autoinducing peptide (AIP). Many AIPs have a structure similar to class I bacteriocins (lantibiotics) (McAuliffe et al., 2001) or to class II bacteriocins (non-lantibiotics) (Ennahar et al., 2000). These AIPs contain recognizable leader peptides with conserved residues, such as the double-glycine leader peptides of most class II and some class I bacteriocins (the consensus for residues −12 to −1 of GG-leader is L5xxELxxIXGG) (Nes & Eijsink, 1999). The web-based bacteriocin genome-mining tool BAGEL (http://bioinformatics.biol.rug.nl/websoftware/bagel/bagel_start.php) might be used to identify such bacteriocin-like AIPs that contain leader peptides (de Jong et al., 2006). In addition, some QS-TCS genes are genetically linked to genes encoding AIP transport and/or modification proteins, bacteriocins and bacteriocin-immunity proteins (Kleerebezem et al., 1997). For double-glycine-type AIPs the cognate transporters are ABC transporters that contain a characteristic N-terminal peptidase C39 domain (COG2274) (Håvarstein et al., 1995), while in the case of agr-like systems AgrB-type cysteine proteases are involved in transport and modification of AIPs (Nakayama et al., 2006; Qiu et al., 2005; Zhang & Ji, 2004).

In silico identification of candidate QS-TCSs in lactobacilli

To identify new putative QS-TCSs, protein sequences of experimentally verified QS-TCSs in Staphylococcus aureus (AgrCA) (Ji et al., 1997) and L. plantarum C11 (PnBCD) (Diep et al., 1996) were collected from public databases (http://www.ncbi.nlm.nih.gov/). Potential system homologues were collected from the genomes of lactobacilli and other Firmicutes via iterative BLASTP searches for HPKs or RRs, using default settings (PSI-BLAST, E-value threshold 1 × 10−5) (Altschul et al., 1990).

Then, various peptide-based QS-specific protein characteristics were used to reduce the list of homologues. These characteristics included: (i) for HPKs the presence of HPK10-subfamily domains (Grebe & Stock, 1999) and five to seven N-terminal transmembrane segments (TMSs), and (ii) for RRs the presence of RD/ComE subfamily domains (Grebe & Stock, 1999) or HTH_LytTR DNA-binding domains (Nikolskaya & Galperin, 2002). In addition, (iii) the presence of adjacent AIP-like genes and additional peptide-based QS-related genes was investigated for these putative QS-TCSs. Protein domains were predicted using the HMMs of SMART, including outlier homologues and PFAM domains (Schultz et al., 1998), and membrane topology was predicted using TMHMM 2.0 (Krogh et al., 2001). The remaining HPK and RR sequences were aligned, and bootstrapped neighbour-joining trees were constructed with CLUSTAL_X (Thompson et al., 1997). The trees were analysed and phylogenetic relationships between sequences established (Van der Heijden et al., 2007). The results for lactobacilli are summarized in Table 1 and in Supplementary Figs S1 and S2 (available with the online version of this paper).

The in silico analysis confirmed previously described QS-TCSs and identified novel putative QS-TCSs and AIPs in lactobacilli. The analysis predicted the presence of five QS-TCSs in L. plantarum WCFS1, two QS-TCSs in both the intestinal species Lactobacillus acidophilus NCFM and L. johnsonii NCC533, one QS-TCS in the intestinal species L. salivarius subsp. salivarius UCC118 and the food species L. delbrueckii subsp. bulgaricus ATCC BAA-365, and no QS-TCS in the intestinal species L. gasseri ATCC 33323. In most cases these putative QS-TCSs had adjacent genes encoding class II bacteriocin-like or lantibiotic-like peptides, which might serve as AIPs, as well as ABC transporters with a peptidase C39 domain, which supports a role in QS-regulated bacteriocin production (see Fig. S2). The QS-functionality of abpIPKR in L. salivarius and LBA1798-LBA1800 in L. acidophilus was previously shown (Flynn et al., 2002; Dobson et al., 2007), and the functional analysis of the putative QS-TCSs in L. plantarum is discussed in the next section.

In addition, in each of the food species Lactobacillus brevis ATCC 367, L. casei ATCC 334 and L. sakei subsp. sakei 23K one putative QS-TCS was identified. However, for those systems the functionality of the predicted QS-TCS is doubtful, since their HPK genes seemed to be (i) interrupted (internal deletions) or (ii) incomplete (N-terminus absent), or (iii) the cognate HPHK was apparently absent (i.e. not adjacent to the RR). A good example of this is the sppIPKR system, which was detected and functional in several L. sakei strains, but not functional in L. sakei subsp. sakei 23K (Møretro et al., 2005). This is caused by a 4 bp internal deletion in the HPK gene (sppK) of this strain, resulting in two truncated HPK fragments (sppKNC and sppKKC) and thereby a non-functional QS-TCS. In L. casei ATCC 334 there are two genes present encoding RRs with RD and LytTR domains. One appears to be an orphan gene (LSEI_2389) with a putative AIP downstream (LSEI_2390). The other RR (LSEI_2599) is part of a putative QS-TCS, where the HPK encoded by LSEI_2600 contains the HPK10-subfamily motifs but lacks most of the N-terminus with TMS. Finally, for L. brevis ATCC 367 there are two HPKs present (typical is only one HPK) downstream of a RR with RD and LytTR domains (LVIS_0163), but they lack either a clear H-box (LVIS_0164) or N-box (LVIS_0165), which are typical of the HPK10 subfamily.

Functional analysis of candidate QS-TCSs in L. plantarum WCFS1

Previous annotation of the 3.3 Mb genome sequence of L. plantarum WCFS1 revealed the presence of 13 genetically linked TCSs, and one orphan HPK and RR (Kleerebezem et al., 2006). The remaining HPK and RR sequences were aligned, and bootstrapped neighbour-joining trees were constructed with CLUSTAL_X (Thompson et al., 1997). The trees were analysed and phylogenetic relationships between sequences established (Van der Heijden et al., 2007). The results for lactobacilli are summarized in Table 1 and in Supplementary Figs S1 and S2 (available with the online version of this paper).

The in silico analysis confirmed previously described QS-TCSs and identified novel putative QS-TCSs and AIPs in lactobacilli. The analysis predicted the presence of five QS-TCSs in L. plantarum WCFS1, two QS-TCSs in both the intestinal species Lactobacillus acidophilus NCFM and L. johnsonii NCC533, one QS-TCS in the intestinal species L. salivarius subsp. salivarius UCC118 and the food species L. delbrueckii subsp. bulgaricus ATCC BAA-365, and no QS-TCS in the intestinal species L. gasseri ATCC 33323. In most cases these putative QS-TCSs had adjacent genes encoding class II bacteriocin-like or lantibiotic-like peptides, which might serve as AIPs, as well as ABC transporters with a peptidase C39 domain, which supports a role in QS-regulated bacteriocin production (see Fig. S2). The QS-functionality of abpIPKR in L. salivarius and LBA1798-LBA1800 in L. acidophilus was previously shown (Flynn et al., 2002; Dobson et al., 2007), and the functional analysis of the putative QS-TCSs in L. plantarum is discussed in the next section.

In addition, in each of the food species Lactobacillus brevis ATCC 367, L. casei ATCC 334 and L. sakei subsp. sakei 23K one putative QS-TCS was identified. However, for those systems the functionality of the predicted QS-TCS is doubtful, since their HPK genes seemed to be (i) interrupted (internal deletions) or (ii) incomplete (N-terminus absent), or (iii) the cognate HPK was apparently absent (i.e. not adjacent to the RR). A good example of this is the sppIPKR system, which was detected and functional in several L. sakei strains, but not functional in L. sakei subsp. sakei 23K (Møretro et al., 2005). This is caused by a 4 bp internal deletion in the HPK gene (sppK) of this strain, resulting in two truncated HPK fragments (sppKNC and sppKKC) and thereby a non-functional QS-TCS. In L. casei ATCC 334 there are two genes present encoding RRs with RD and LytTR domains. One appears to be an orphan gene (LSEI_2389) with a putative AIP downstream (LSEI_2390). The other RR (LSEI_2599) is part of a putative QS-TCS, where the HPK encoded by LSEI_2600 contains the HPK10-subfamily motifs but lacks most of the N-terminus with TMS. Finally, for L. brevis ATCC 367 there are two HPKs present (typical is only one HPK) downstream of a RR with RD and LytTR domains (LVIS_0163), but they lack either a clear H-box (LVIS_0164) or N-box (LVIS_0165), which are typical of the HPK10 subfamily.
Out of these, five TCSs were identified as candidate QS-TCSs based on the in silico approach described above. They were predicted to comprise five HPKs that showed characteristics of the HPK10 subfamily and six RRs that contained a HTH-DNA-binding domain of the LytTR family. Of the remaining TCSs, which did not fit the general in silico criteria for peptide-based QS-TCSs, we found on closer examination that there were two TCSs where the HPKs and RRs showed similarity to another type of QS-TCS, as found for the ComPA system of B. subtilis. These HPKs showed HPK7-subfamily characteristics (Grebe & Stock, 1999) and the RRs contained a HTH-DNA-binding domain of the LuxR family (Fuqua et al., 1994) (Table 1).

Table 1. General features of HPKs and RRs of candidate QS-TCSs in L. plantarum WCFS1 and related lactobacilli (see also Figs S2 and S3)

<table>
<thead>
<tr>
<th>Gene*/accession no.</th>
<th>Locus†</th>
<th>Size (aa)</th>
<th>TMS‡</th>
<th>HPK10 subfamily§</th>
<th>Gene*/accession no.</th>
<th>Locus†</th>
<th>Size (aa)</th>
<th>LytTR domains</th>
<th>Gene*/accession no.</th>
<th>Locus†</th>
</tr>
</thead>
<tbody>
<tr>
<td>HPK L. plantarum WCFS1 (3.34 Mb)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>RR L. plantarum WCFS1 (3.34 Mb)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>plnB lp_0416</td>
<td>442</td>
<td>7l</td>
<td>+</td>
<td></td>
<td>plnC/plnD lp_0417/0418</td>
<td>247</td>
<td>+</td>
<td>plnA lp_0415</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pltK lp_1355</td>
<td>420</td>
<td>6</td>
<td>+</td>
<td></td>
<td>pltR lp_1356</td>
<td>255</td>
<td>+</td>
<td>pltA lp_1354a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hpk4 lp_1488</td>
<td>343</td>
<td>5</td>
<td>(HPK7)</td>
<td></td>
<td>rrp4 lp_1487</td>
<td>217</td>
<td>(LuxR)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hpk6 lp_1943</td>
<td>367</td>
<td>5</td>
<td>(HPK7)</td>
<td></td>
<td>rrp6 lp_1942</td>
<td>201</td>
<td>(LuxR)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hpk9 lp_3063</td>
<td>422</td>
<td>6</td>
<td>+</td>
<td></td>
<td>rrp8 lp_2665</td>
<td>249</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hpi10 lp_3088</td>
<td>416</td>
<td>6</td>
<td>+</td>
<td></td>
<td>rrp10 lp_3087</td>
<td>248</td>
<td>+</td>
<td>lp_3089</td>
<td></td>
<td></td>
</tr>
<tr>
<td>lamC lp_3581</td>
<td>419</td>
<td>5</td>
<td>+</td>
<td></td>
<td>lamA lp_3580</td>
<td>247</td>
<td>+</td>
<td>lamD lp_3581a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L. acidophilus NCFM (2.0 Mb)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NP_193512 LBA0602</td>
<td>426</td>
<td>6</td>
<td>+</td>
<td></td>
<td>NP_193513 LBA0603</td>
<td>265</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NP_194634 LBA1799</td>
<td>440</td>
<td>7</td>
<td>+</td>
<td></td>
<td>NP_194633 LBA1798</td>
<td>274</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L. johnsonii NCC533 (2.0 Mb)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NP_964473 LJ0448</td>
<td>419</td>
<td>6</td>
<td>+</td>
<td></td>
<td>NP_964474 LJ0449</td>
<td>255</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NP_964617 LJ0764</td>
<td>435</td>
<td>6</td>
<td>+</td>
<td></td>
<td>NP_964619 LJ0766</td>
<td>265</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L. salivarius subsp. salivarius UCC118 (2.13 Mb)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>abpK LSL_1913</td>
<td>429</td>
<td>7</td>
<td>+</td>
<td></td>
<td>abpR LSL_1912</td>
<td>264</td>
<td>+</td>
<td>abpIP LSL_1914</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lactobacillus sakei subsp. sakei 23K (1.9 Mb)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sppKN+sppKC LSA0561</td>
<td>69N/A</td>
<td>0</td>
<td>+</td>
<td></td>
<td>sppIP LSA0561_b</td>
<td>248</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L. brevis ATCC 367 (2.35 Mb)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>YP_794364 LVIS_0164</td>
<td>444</td>
<td>6</td>
<td>± (no H-box)</td>
<td></td>
<td>YP_794363 LVIS_0163</td>
<td>251</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>YP_794365 LVIS_0165</td>
<td>443</td>
<td>7</td>
<td>± (no N-box)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L. delbrueckii subsp. bulgaricus ATCC BAA-365 (1.9 Mb)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>YP_812154 LBU_0022</td>
<td>434</td>
<td>6</td>
<td>+</td>
<td></td>
<td>YP_812153 LBU_0021</td>
<td>260</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L. casei ATCC 334 (2.93 Mb)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>YP_807763 LSEI_2600</td>
<td>303N</td>
<td>2</td>
<td>+</td>
<td></td>
<td>YP_807570 LSEI_2389</td>
<td>268</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>YP_807762 LSEI_2599</td>
<td>234</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* †TMS, transmembrane segments as predicted from TMHMM 2.0 (http://www.cbs.dtu.dk/services/TMHMM/).
||Experimental confirmed data for PlnB in L. plantarum C11 (Johnsborg et al., 2003).
N, Incomplete gene: N-terminal domain (partially) missing.
, 4 bp deletion in sppK gene (Moretto et al., 2005).
Four adjacent HPK- and RR-encoding genes constituted complete TCSs that were classified as candidate peptide-based QS-TCSs (pln, plt, TCS10 and lam). In addition, the orphan HPK and RR could constitute a fifth peptide-based QS-TCS. The relevant features of these HPKs and RRs are summarized in Table 1 and their genetic organization is shown in Supplementary Figs S2 and S3. For all of the complete QS-TCSs a putative AIP was encoded upstream (Fig. 2 and Fig. S3), two of which (pln and plt) contained a putative double-glycine-type leader peptide, while the third and fourth were of a different type. The gene connected to the lam QS-TCS (lamD) encodes a cyclic thiolactone AIP (Sturme et al., 2005), whereas the one connected to TCS10 encodes a putative AIP that shows little similarity to known AIPs. Details of the characterization and functionality of these putative QS-TCSs are discussed below.

**Plantaricin TCS pln (lp_0415 to lp_0418)**

The plantaricin TCS pln (lp_0415 to lp_0418) was identical to the previously described plnABCD system of *L. plantarum* C11, which regulates the production of class II antimicrobial peptides (AMPs) (Diep et al., 1996). This system contains a gene (plnB: lp_0416) encoding a typical HPK of subfamily HPK10 that was shown to have seven TMSs (confirmed in strain C11 by Johnsborg et al., 2003). Downstream of plnB, two RRs are encoded (plnC and plnD) that both contain the RD receiver domains and LytTR HTH-DNA-binding domains. Upstream of plnB, plnA encodes a 48 amino acid double-glycine-type AIP precursor (Nes & Eijsink, 1999). Genes encoding class II bacteriocins (PlnE-PlnF, PlnJ-PlnK and PlnN) were localized near the pln regulatory module (Supplementary Fig. S3). Cleavage of the double-glycine-type leader peptide from the PlnA precursor peptide results in a linear AIP of 26 amino acids without modifications and an amphipathic character (Fig. 2). The pln system was found to be functional in *L. plantarum* WCFS1, as was shown by a bacteriocin agar-well diffusion assay with *L. plantarum* 965 as an indicator strain (Fig. 3). The *L. plantarum* WCFS1 native state was bacteriocin-negative (Bac−), but bacteriocin production could be induced with either bacteriocin-positive supernatant of strain C11 (Bac+) or purified PlnA peptide.

The pln system of WCFS1 is expected to play a role in competition with other bacteria, as in strain C11 it regulates the production of the bacteriocins PlnE-PlnF, PlnJ-PlnK and PlnN. These C11 plantaricins showed activity against closely related species (Andersen et al., 1998), which can be found in the same ecological niches as *L. plantarum* strains WCFS1 and C11 (Vaughan et al., 2002). However, some strain-specific differences in plantaricin activity and target specificity might be expected for various *L. plantarum* strains, considering the differences in pln gene composition and sequence variations identified in the pln operons of strains WCFS1, C11 and NC8 (Maldonado et al., 2004; Molenaar et al., 2005).

**TCS plt (lp_1354a, lp_1355 and lp_1356)**

The plt locus (lp_1354a, lp_1355 and lp_1356) encodes a typical HPK (pltK: lp_1355) of the HPK10 subfamily with five or six predicted TMSs, and a RR (pltR: lp_1356) that contains a RD-type receiver domain and a LytTR HTH-DNA-binding domain. Upstream of pltK, plnA encodes a 48 amino acid double-glycine-type AIP precursor (Nes & Eijsink, 1999). Genes encoding class II bacteriocins (PlnE-PlnF, PlnJ-PlnK and PlnN) were localized near the plt regulatory module (Supplementary Fig. S3). Cleavage of the double-glycine-type leader peptide from the PlnA precursor peptide results in a linear AIP of 26 amino acids without modifications and an amphipathic character (Fig. 2). The plt system was found to be functional in *L. plantarum* WCFS1, as was shown by a bacteriocin agar-well diffusion assay with *L. plantarum* 965 as an indicator strain (Fig. 3). The *L. plantarum* WCFS1 native state was bacteriocin-negative (Bac−), but bacteriocin production could be induced with either bacteriocin-positive supernatant of strain C11 (Bac+) or purified PlnA peptide.

The plt system of WCFS1 is expected to play a role in competition with other bacteria, as in strain C11 it regulates the production of the bacteriocins PlnE-PlnF, PlnJ-PlnK and PlnN. These C11 plantaricins showed activity against closely related species (Andersen et al., 1998), which can be found in the same ecological niches as *L. plantarum* strains WCFS1 and C11 (Vaughan et al., 2002). However, some strain-specific differences in plantaricin activity and target specificity might be expected for various *L. plantarum* strains, considering the differences in pln gene composition and sequence variations identified in the pln operons of strains WCFS1, C11 and NC8 (Maldonado et al., 2004; Molenaar et al., 2005).

---

**Fig. 2.** AIPs encoded on the *L. plantarum* WCFS1 genome. Triangles indicate the cleavage site between the double-glycine leader peptide and the (predicted) mature peptide in PlnA and PltA (shown in bold). The underlined bold residues in the LamD precursor peptide are processed to the mature thiolactone peptide shown on the right. For the putative AIP encoded by lp_3089 the mature peptide sequence and structure are unknown.
expression pattern. The \(plt\) locus does not appear to be genomically directly linked to any ABC transporter or bacteriocin-encoding gene (Supplementary Figs S2 and S3). However, the PltA pre-peptide could be transported by another ABC transporter with a peptidase C39 protease-family domain (Håvarstein et al., 1995), such as the ABC transporter encoded within the \(pbn\) locus (encoded by \(pbnG\) and \(pbnH\); Fig. S3).

**TCS \(lam\) (lp\,3580 to lp\,3582)**

The TCS encoded by \(lp\,3580\) to \(lp\,3582\) was designated \(lam\) for *Lactobacillus agr*-like module, as it showed a similar gene organization to the \(agr\)BDDCA QS system of *S. aureus* (Ji et al., 1995). The *L. plantarum lamBDDCA* locus encodes a HPK (\(lamC: lp\,3581\)) of the HPK10 subfamily with five predicted TMSs, and a RR (\(lamA: lp\,3580\)) that contains a RD-type receiver domain and a LytTR HTH-DNA-binding domain. The locus is organized as an operon and is transcribed as a single transcript in a cell-density-dependent manner (see Fig. 4 and Sturme et al., 2005). Analogous with the staphylococcal \(agr\) system (Ji et al., 1995), the *L. plantarum lamBDDCA* locus is involved in production of a cyclic thiolactone peptide (CVGIW) which is predicted to derive from the LamD precursor (lp\,3581a) that is processed by LamB (lp\,3582) (Fig. 2). Analysis of a \(lamA\) mutant revealed a role for the \(lam\) operon in *L. plantarum* biofilm-forming capacity. Transcriptome profiling of wild-type and \(lamA\) mutant strains in early-, mid- and late-exponential phase uncovered only a small set of clustered genes (2% of all genes) that were significantly modulated by the \(lamA\) mutation. These data confirmed the autoregulation of \(lamBDDCA\), and showed the regulation of a surface polysaccharide biosynthesis gene-cluster, and several cell envelope and sugar utilization functions (Sturme et al., 2005). The same study suggested that a direct repeat sequence within the \(lamB\) promoter region (5’-TCTTTAAT – 12 bp – TCTTTAAA-3’) that displays similarity with previously established cognate cis elements of LamA homologues (Morfeldt et al., 1996; Qin et al.,...
belonging to QS-TCS showed that the lp_3062 gene product might therefore function autonomously (Zulty & Barcak, 2001; Quadri et al., 1997; Risøen et al., 2000) acts as the LamA DNA-binding site.

TCS10 (lp_3087 and lp_3088)

The bioinformatic analysis suggests that the HPK and RR of TCS10 (lp_3087 and lp_3088) are inparalogues of the lamCA-encoded proteins, with the cognate HPKs and RRs showing 55% and 70% identity, respectively, at the amino acid level. This module lacks a lamB homologue and originally seemed to lack a lamD homologue. However, a region somewhat upstream of the hpk10 gene is remarkably similar to that of the region upstream of the lam system. This (promoter) region contains a direct repeat (5'-TCTTGAAT - 12 bp - TCTTAAAG-3'), displaying very high similarity with the proposed regulatory element of the lamB promoter (see above). Interspersed between this regulatory region of TCS10 and the hpk10 gene is a region of about 130 nucleotides that on closer inspection was shown to encode a small peptide (lp_3089) that initially was not annotated (Fig. 2 and Supplementary Fig. S3). Although the peptide shows some conservation with respect to the lamD gene product, it is not clear whether it is a genuine AIP. Taken together these findings support an inparalogous relationship between the RRs and HPKs of the TCS10 and lam systems, and suggests that there could be cross-talk between these two regulatory systems.

Orphan genes rrp8 (lp_2665) and hpk9 (lp_3063)

The two orphan hpk and rrp genes show typical characteristics of peptide-based QS systems. The orphan gene hpk9 (lp_3063) was predicted to encode a HPK10 subfamily protein containing six TMSs. Interestingly, immediately downstream (7 bp) of hpk9 a small ORF is located (lp_3062) that shows 46% homology to the C-terminal domain of the TlxO protein of Haemophilus influenzae (PF04994). The TlxO protein is proposed to play a key role in cell-density-dependent regulation of genetic competence in H. influenzae, and the C-terminal domain is suggested to function autonomously (Zulty & Barcak, 1995). The lp_3062 gene product might therefore function in co-operation with hpk9. However, no clearly identifiable cognate RR appears to be encoded in the vicinity of hpk9. The downstream-located response regulator lp_3060 probably does not fulfil this function, as it was not transcriptionally linked and lacks a RD-type receiver domain and contains a typical HTH-AraC DNA-binding domain (SM00342) rather than the canonical LytTR or HTH-LuxR DNA-binding domains (see Supplementary Fig. S3). It is possible that the hpk9-encoded HPK in combination with the rrp8-encoded orphan RR forms a functional TCS, which could be involved in peptide-based QS. This possibility is supported by the finding that rrp8 (lp_2665) encodes a RR with a typical LytTR HTH-DNA-binding domain. Moreover, the neighbour-joining trees based on protein sequence alignments of HPKs and RRs belonging to QS-TCS showed that the hpk9- and rrp8-encoded proteins (LPL03063 and LPL02665 in Supplementary Fig. S1) cluster with several HPKs and RRs of other Firmicutes that are adjacent on their respective genomes, such as LEUM_0009/LEUM_0008 from Leuconostoc mesenteroides or EF1820/EF1822 from Enterococcus faecalis. Interestingly, the gene downstream of rrp8 (lp_2664) contains a HDc-superfamily-type phosphohydrolase-domain (SM00471) that might have a role in dephosphorylation of the rrp8 gene product. Overall, the role of hpk9 in QS-mediated regulation and the identity of its eventual partnering transcriptional regulator remain to be established.

Additional candidate QS-TCSs and genes involved in non-peptide-based QS

Besides the typical peptide-based QS-TCSs, TCS4 (lp_1487 and lp_1488) and TCS6 (lp_1942 and lp_1943) both encode HPK7-subfamily-type HPKs (Grebe & Stock, 1999) adjacent to a RR containing a HTH-LuxR DNA-binding domain (see Table 1 and Supplementary Fig. S3). This resembles the protein architecture of the ComPA QS-TCS of B. subtilis that is involved in QS regulation of competence (Weinrauch et al., 1989, 1990). However, both the HPK7 subfamily and the HTH-LuxR DNA-binding domain are not exclusively associated with QS-TCSs (Gray & Garey, 2001; Grebe & Stock, 1999), indicating that on the basis of in silico analysis the involvement of TCS4 and TCS6 in QS can only be tentatively suggested. Interestingly, next to the putative QS-TCS an isolated homologue of the luxS gene was identified (lp_0774), encoding the auto-inducer-2 (AI-2) synthase (Schauder et al., 2001). AI-2 is thought to play a role in QS, although there is still some debate on its actual physiological role (Sun et al., 2004). Associated functions like lsr, which is involved in AI-2 uptake in Salmonella typhimurium, or luxPQ involved in AI-2 sensing and signal transduction in Vibrio harveyi, were not identified in the L. plantarum genome (Sun et al., 2004). AI-2 could be involved in interspecies communication of L. plantarum with other bacteria within the same niche, as the luxS gene was detected in both Gram-negative and Gram-positive bacteria (Xavier & Bassler, 2003). Interestingly, AI-2 activity was detected in rumen samples (Mitsumori et al., 2003), suggesting a natural role of the AI-2 QS system in gastrointestinal ecosystems.

Concluding remarks

L. plantarum WCFS1 contains a relatively high number of (putative) peptide-based QS-TCSs (five, of which at least four can be directly coupled to a putative AIP), as well as other putative QS genes. This could reflect the ecological flexibility of this species, which can be found in plants, fermented foods and the gastrointestinal tract. In comparison, the genomes of related lactobacilli that are more restricted to specific environments seem to encode significantly fewer peptide-based QS-TCSs, despite the differences in genome size (see Table 1). Depending on
niche-specific conditions, different signalling systems might be triggered, resulting in a gene-regulatory network that controls a variety of phenotypic traits in response to environmental conditions in combination with cell density. The potential for cross-talk between the lamBDCA and TCS10 systems might be exemplary for the complexity of a regulatory network that controls surface adherence of _L. plantarum_ under specific conditions (Sturme et al., 2005). Moreover, the presence of competing micro-organisms could activate specific QS-TCSs involved in competition, as has been shown for the plantaricin system in _L. plantarum_ NC8 (Maldonado et al., 2004). Interestingly, the involvement of _agr_-like TCSs in host–microbe interactions of commensal bacteria has recently been suggested for _Roseburia inulinivorans_ (Scott et al., 2006).

Further experimental studies on the regulatory mechanisms of the different QS systems of _L. plantarum_ and other _Lactobacillus_ species and their effects on (global) gene regulation will be necessary, to provide more insight into the role of these systems in the survival of these organisms in their natural environments.

**Acknowledgements**

M. H. J. Sturme would like to acknowledge the support of the Netherlands Organization for Scientific Research (NWO) and C. F. the support of the NBIC/Netherlands Genomics Initiative via the Kluiver Centre for Genomics of Industrial Fermentations.

**References**


AgrA/LytR family.

DNA-binding domain in the transcriptional regulators of the AlgR/transcriptional regulation.

Carnobacterium piscicola from 37


