High-resolution profiles of the *Streptococcus mitis* CSP signaling pathway reveal core and strain-specific regulated genes

G. Salvadori 1, R. Junges 1, H. A. Åmdal 1, T. Chen 2, D. A. Morrison 3 and F. C. Petersen 1*

**Abstract**

**Background:** In streptococci of the mitis group, competence for natural transformation is a transient physiological state triggered by competence stimulating peptides (CSPs). Although low transformation yields and the absence of a widespread functional competence system have been reported for *Streptococcus mitis*, recent studies revealed that, at least for some strains, high efficiencies can be achieved following optimization protocols. To gain a deeper insight into competence in this species, we used RNA-seq, to map the global CSP response of two transformable strains: the type strain NCTC12261 T and SK321.

**Results:** All known genes induced by ComE in *Streptococcus pneumoniae*, including sigX, were upregulated in the two strains. Likewise, all sets of streptococcal SigX core genes involved in extracellular DNA uptake, recombination, and fratricide were upregulated. No significant differences in the set of induced genes were observed when the type strain was grown in rich or semi-defined media. Five upregulated operons unique to *S. mitis* with a SigX-box in the promoter region were identified, including two specific to SK321, and one specific to NCTC12261 T. Two of the strain-specific operons coded for different bacteriocins. Deletion of the unique *S. mitis* sigX regulated genes had no effect on transformation.

**Conclusions:** Overall, comparison of the global transcriptome in response to CSP shows the conservation of the ComE and SigX-core regulons in competent *S. mitis* isolates, as well as species and strain-specific genes. Although some *S. mitis* exhibit truncations in key competence genes, this study shows that in transformable strains, competence seems to depend on the same core genes previously identified in *S. pneumoniae*.

**Keywords:** *Streptococcus mitis*, Natural transformation, CSP, Competence, Quorum sensing

**Background**

In several streptococci intercellular coordination of gene expression mediated by peptide pheromones is associated with development of competence for transformation [1]. The pheromones activate a signal transduction pathway that regulates natural transformation. This physiological feature provides a selective advantage by allowing competent cells to acquire new characteristics, such as antibiotic resistance, by incorporation of DNA from other cells. In *Streptococcus pneumoniae*, natural competence is a tightly controlled transient state: it spontaneously arises during the early exponential growth phase at a certain cell density and reaches its peak after approximately 20 min, before it quickly shuts down [1, 2]. The regulatory cascade is induced via activation of *comC*, encoding the competence-stimulating peptide (CSP) [1], which is further cleaved and exported by its secretion apparatus ComAB [3]. CSP binds and activates its cognate receptor ComD, which, after phosphorylation, activates the response regulator ComE [4]. Phosphorylated ComE then specifically binds to a target conserved sequence, referred to as the ComE-box, in the promoter region of sigX, a global transcriptional modulator, in addition to *comCDE* itself and *comAB*, creating a positive feedback loop that coordinates an explosive spread of competence among nearby cells. These genes form the core of the quorum sensing apparatus for the induction of competence [5, 6] and are known as ‘early’ genes. SigX then initiates the transcription of a number of late genes involved in DNA uptake, recombination and fratricide by recognizing a
SigX-box consensus sequence (TACGAATA) in their promoter regions [7, 8].

Many of the SigX-regulated genes contribute to various aspects of transformation. A dozen genes in four operons are required for assembly of the DNA transport machinery, while five genes in five operons are considered essential for efficient recombination of donor DNA strands [9]. Competent pneumococcal cells also upregulate at least six genes involved in the production of killing factors and their respective immunity proteins, including the early gene comM, and late genes cibABC, cbpD and lytA [10, 11]. CbpD, more specifically, is a murine hydrolase that plays a key role in the fraticide phenomenon, in which competent cells are able to kill and lyse non-competent sibling cells in the neighboring milieu, in a predatory mechanism most likely wired to acquire their DNA [11]. Interestingly, fraticide is not unique to pneumococci, as it has also been demonstrated in closely related species such as Streptococcus mitis and Streptococcus oralis [10].

*S. mitis* is a pioneer colonizer of the oral cavity, residing on the teeth, tongue and mucous membranes, as well as tonsils and nasopharynx, where it may exist side-by-side in biofilms with *S. pneumoniae* [12–14]. Both *S. mitis* and *S. pneumoniae* are naturally competent for natural transformation, and the exchange of genetic material by homologous recombination between them is recognized as part of their parallel evolution [15]. In addition, transfer of both antimicrobial resistance and virulence genes has been described between the species, evidenced by the presence of mosaic structures in gene sequences [16–18]. Interestingly, it has been suggested that the acquisition of *S. mitis* genes by *S. pneumoniae* has contributed to its evolution, but the opposite has not been proven [15]. Both species share a common ancestor, and it has been suggested that *S. mitis* evolved from a *S. pneumoniae*-like precursor by genome reduction, losing virulence genes and developing mechanisms of adaptation to the human host [15]. *S. mitis* strains represent a wide range of very distinct lineages under the same name, and previous reports have demonstrated that the genetic variability among *S. mitis* strains can be greater than between *S. mitis*, *S. pneumoniae* and *Streptococcus pseudopneumoniae* [19, 20]. Due to its ability to induce oral mucosal antibodies, *S. mitis* has been investigated for its potential as a vaccine vector [21]. Among 18 *S. mitis* strains that have been partially or completely sequenced, the majority of competence genes are widely conserved, with the exception of sigX, which appears in truncated forms in 44% of the strains analyzed [15, 22]. However, the activity, regulation, and possible role of *S. mitis* competence in the flow of genetic information between commensal and pathogenic streptococci remains unknown [22, 23]. Although low transformation yields and the absence of a widespread functional competence system have been reported for this species, recent studies reveal that, at least for some strains, high efficiencies can be achieved following optimization steps in current protocols [22]. Here, we applied such optimal conditions to gain the first detailed insight into the regulatory pathways of competence development and investigate the strain specificity of the CSP signaling response in two *S. mitis* isolates.

**Results**

*S. mitis* transformation efficiency in response to CSP

In contrast to *S. pneumoniae*, most *S. mitis* strains produce different strain-specific CSPs but seem to transform with low efficiency under laboratory conditions. Besides responding to different strain-specific stimulating peptides [22], the *S. mitis* type strain and strain SK321 are located phylogenetically apart (Fig. 1a) [20]. Pairwise comparison of the two strains performed by orthologous clustering using OrthoVenn revealed that the type strain has 1636 annotated proteins, whereas SK321 has 1757. One thousand, three hundred and seventy-nine orthologous clusters are shared, while 203 predicted proteins in the type strain and 352 in SK321 are strain-specific. Altogether, these features make the two *S. mitis* strains an interesting sample of *S. mitis* diversity for initial study of competence gene regulation in this species.

Competence for genetic transformation in pneumococcus occurs during a brief period of highly specialized protein synthesis (10–20 min), coordinated among many or all cells of an actively growing culture. Early genes present strongly increased expression during the period between 5 and 10 min after CSP induction, decreasing nearly to original values by 20 min after initiation of exposure to CSP. “Late” genes display a similar expression pattern, but with a delay of approximately 5 min [9]. Consistent with these timings, transformation kinetics for DNA uptake and recombination in the pneumococcus peaks between 10 to 15 min after CSP induction [24–26]. Evaluation of *S. mitis* type strain and SK321 temporal transformation patterns showed a remarkably similar pattern, with maximal transformation yields after 15 min of CSP induction in both strains (Fig. 1b). Interestingly, while transformation declined substantially after 30 min for the type strain, it declined more slowly in SK321. Based on these data, we chose 15 min after CSP induction as a suitable sampling time for evaluating competence-related gene expression.

Responses to CSP by the *S. mitis* type strain and strain SK321

The competence response in *S. pneumoniae*, which has been described in detail in the literature, comprises three phases, early, late and delayed, which vary in magnitude depending on environmental conditions such as pH,
temperature, and presence of albumin. While the early and late responses depend on well-defined regulons, the delayed response is less well understood, with no specific regulatory mechanisms yet identified. Thus, to better characterize the competence response in *S. mitis*, we investigated the transcriptome profile of the type strain in response to CSP during growth in two contrasting media, the semi-defined medium C + Y YB, and the rich medium TSB. The medium C + Y YB is an optimal medium for competence development for *S. mitis* type strain, supporting higher levels of *sigX* expression when compared to rich medium (TSB) or semi-defined C + Y [22].

In TSB, exposure to CSP in the type strain resulted in the upregulation of 68 genes by > 2-fold, while downregulating 21 genes (Additional file 1: Table S1). When C + Y YB medium was used to examine the CSP response of *S. mitis* type strain, supporting higher levels of *sigX* expression when compared to rich medium (TSB) or semi-defined C + Y [22].

In TSB, exposure to CSP in the type strain resulted in the upregulation of 68 genes by > 2-fold, while downregulating 21 genes (Additional file 1: Table S1). When C + Y YB medium was used to examine the CSP response of *S. mitis* type strain, 79 genes were upregulated > 2-fold, whereas 19 were downregulated (Additional file 2: Table S2). There was a strong similarity in the response to CSP in different media with regards to the number of transcripts upregulated, with an overlap of 53 genes. Interestingly, among the downregulated genes, only two hypothetical proteins coincided between TSB and C + Y YB media. With the exception of two gene clusters corresponding to the yellow circles in Fig. 2, the CSP response profile of *S. mitis* was only slightly affected by the choice of growth medium (Fig. 2). Three genes in one cluster are orthologues of delayed genes in *S. pneumoniae* (SP0785, SP0786 and SP0787) [9], and the other 5-gene cluster highly upregulated in C + Y YB corresponds to a region that is not involved in competence in *S. pneumoniae*. The ComE regulon presented a modestly higher upregulation when exposed to TSB, whereas the SigX regulon seemed to respond similarly in both media. This difference indicates that environmental factor may be important for the regulatory processes in *S. mitis* competence.

In SK321, a strain grouped in a different cluster than the type strain, evaluation of the CSP-induced response was performed using the same cut-off values (> 2-fold). One hundred and sixty-seven genes positively responded to CSP in the strain SK321 when cultured in C + Y YB, while 69 showed significant downregulation (Additional file 3: Table S3). Overall, this isolate presented significantly higher expression levels than the type strain. This was not an isolated observation, since a similarly exacerbated response has been reported, as determined by means of luciferase reporter assays for the *sigX* gene [22]. For comparison of the two *S. mitis* isolates’ expression patterns, we will refer below only to data collected in C + Y YB for both strains.

Since samples for RNA-seq were collected at a single time point (15 min post CSP treatment), it was not possible to classify gene expression by temporal response. However, given the genetic proximity of *S. mitis* to *S. pneumoniae* and similarity in temporal patterns of transformation, we searched for orthologues of differentially expressed genes in a competent *S. pneumoniae* strain derivative of R6 [27] and analyzed their promoter regions for conserved regulatory sequences. By doing so, it was possible to identify orthologous genes, their functions and the expression pattern obtained in this transcriptome analysis, and compare with the DNA microarray data already known for *S. pneumoniae*.

**Early response: Conservation of the ComE regulon**

Genes upregulated by CSP in the *S. mitis* type strain and SK321 are listed with their orthologous pneumococcal
Orthologues of all *S. pneumoniae* ComE-responsive genes acting in regulation of the competence cascade and fratricide were identified as upregulated in both strains. These included 9 transcriptionally activated regions (TARs) accounting for a total of 21 genes responsible for CSP processing and export (*comAB*), competence regulators (*comCDE*) and *sigX* (duplicate copy), and bacteriocin related genes. An early gene encoding an immunity protein involved in competence-regulated self-protection (*comM*) was also upregulated, as well as several other orthologues of *S. pneumoniae* early genes, such as *comW*, *purA*, and *lytR* [9]. The only observable difference in the early response between the two *S. mitis* strains was the upregulation of the orthologue of SP2156 in SK321 (SMSK321_1581), a gene encoding a membrane bound protein of unknown function.

In *S. pneumoniae*, early genes are preceded by a conserved regulatory sequence consisting of two 9 bp imperfect direct repeats (DRs) separated by 12 nucleotides (aCaTTcaG-12-aCAGTTgaG), which is recognized as the phosphorylated ComE-binding site or ComE-box [8]. Alignment of the regions immediately upstream of the orthologues of early genes in the *S. mitis* type strain and SK321 revealed the presence of consensus motifs in all cases, except for the orthologue of SP2156 in the type strain (SM12261_0729), which was not upregulated by CSP (Table 2). During the search for DRs among the upregulated sequences, we noticed a ComE-box upstream of SMSK321_0027, which encodes an uncharacterized Gly-Gly peptide. This gene is orthologous to SP0429, which has never been associated to competence in *S. pneumoniae*, most likely due to a defective regulatory sequence without the first 9 bp DR (Table 2). However, in our observation of transcriptome analysis of *S. pneumoniae* D39 (data not shown) endogenously competent cells showed upregulation of SPD_0391, orthologue of SMSK321_0027. We also detected a DR upstream of SPD_0391; interestingly, this gene has not been related to competence in D39 in a previous report [28]. Although carrying sequences slightly divergent from the consensus for DRs (Table 2), the gene encoding *comW* presented upregulation in both strains independent of medium composition. This suggests that this regulatory site is still responsive despite the presence of a less conserved DR.

**Late response: conservation of the SigX regulon**

Late CSP-induced genes make the largest group of competence-specific products involved in binding, uptake, processing and integration of exogenous DNA, and the production of killing factors. Table 3 shows information on upregulated sequences in the two *S. mitis* strains.
and their orthologue late genes in *S. pneumoniae* [9]. Core genes, under regulation of SigX in the majority of competent streptococcal species [29] are in bold. Overall, upregulated genes were organized in 15 TARs (Fig. 3). For both strains, orthologues of late genes involved in DNA uptake were grouped in two significantly upregulated operons (*comGA-comGG* and *comEA-comEC*). Both operons are composed by genes required for assembly of the DNA transport and uptake machinery and are known as indispensable for transformation in *S. pneumoniae*. In addition, *comFA*, involved in DNA transport, presented a strong induction in both isolates. Four upregulated operons accounted for 5 DNA recombination genes: *ssbB* (NCTC12261<sup>7</sup>, 98.6; SK321, 697.9-fold), *dprA* (NCTC12261<sup>7</sup>, 63.8; SK321, 339.3-fold), *coiA* (NCTC12261<sup>7</sup>, 58.6; SK321, 139.6-fold), and *cina-recA* (NCTC12261<sup>7</sup>, 21.7; SK321, 55.3-fold) and (NCTC12261<sup>7</sup>, 5.3; SK321, 18.8-fold), respectively. These genes are essential for the efficient replacement of donor DNA strands. As mentioned, fratricide has been demonstrated not only for pneumococcus, but also for commensal streptococci [30]. In the present analysis, the gene encoding the murein hydrolase essential for the pneumococcal fratricide mechanism, *cbpD*, was significantly upregulated (NCTC12261<sup>T</sup>, 116-fold; SK321, 458.5-fold). SP0031, a non-core late gene annotated as hypothetical protein in *S. pneumoniae* TIGR4, was also induced in both *S. mitis* isolates (non-annotated gene) (Table 3) and in our observations of *S. pneumoniae* D39 endogenous competence response (data not shown).

SigX, together with RNAP, recognizes a "cinbox" or "SigX-box" unique DNA element (−10 from the transcription start) and featuring a T-rich region at (−25) in the promoter regions of late genes, initiating their transcription [7]. To search for conserved promoter elements in *S. mitis*, the regions immediately upstream of the potential start sites of orthologues of *S. pneumoniae* late operons were aligned (Fig. 4). Table 4 displays SigX-box sequences with no more than one mismatch detected upstream of 13 SigX core genes; combined with

| Gene ID | Mean fold-change<sup>a</sup> | Description |
|---------|-----------------------------|-------------|
| Orthologues of early genes |
| 0014    | 0014                        | 1251        |
| 0018    | 0016                        | 0337        |
| 0019    | 0017                        | 0338        |
| 0042    | 0048                        | 1310        |
| 0043    | 0049                        | 1311        |
| 2006    | 0613                        | 1651        |
| 2237    | 0788                        | 1635        |
| 2236    | 0787                        | 1634        |
| 2235    | 01116<sup>e</sup>          | 00428<sup>e</sup>  |
| 1110    | 0833                        | 0680        |
| 1945    | 0911                        | 1653        |
| 1944    | 0912                        | 1654        |
| 1943    | 0913                        | 1655        |
| 1942    | 0914                        | 1656        |
| 1717    | 0940                        | 1696        |
| 1716    | 0941                        | 1697        |
| 1549    | 1162                        | 0137        |
| 1548    | 1163                        | 0138        |
| –       | 1164                        | 0139        |
| 1547    | 1165                        | 0140        |
| 2156    | 0729                        | 1581        |
Table 2: Alignment of putative ComE box sequences upstream of clusters of *S. mitis* type strain and SK321 orthologues of early genes

| Function in competence | Annotation | Gene locus | Consensus<sup>b</sup> | aCanTTcaG<sup>a</sup> | aCagTTgaG<sup>a</sup> | −10 site |
|------------------------|------------|------------|------------------------|-----------------------|------------------------|----------|
| Regulation             | comA       | SP0042     | GCCATTGGG             | 12 TCATTGGG          | 32 TAAAGT            |          |
|                        |            | SM12261_0048| GCCATTGGG             | 12 GCATTGGG          | 32 TAAAGT            |          |
|                        |            | SMSK321_1310| GCCATTGGG             | 12 GCATTGGG          | 32 TAAAGT            |          |
|                        | comC       | SP2237     | ACACTTGGG             | 12 ACGTTGGG          | 31 TATAAT             |          |
|                        |            | SM12261_0788| ACACTTGGG             | 12 ACGTTGGG          | 31 TATAAT             |          |
|                        |            | SMSK321_S1634| ACACTTGGG             | 12 ACGTTGGG          | 31 TATAAT             |          |
|                        | comW       | SP0018     | GCATTTTTG             | 10 GAATTCAa          | 38 TATACT             |          |
|                        |            | SM12261_0016| CtttTGAG              | 12 ACAATTCAa         | 31 TATAAG             |          |
|                        |            | SMSK321_0337| CtttTGAG              | 12 GCATTCAa          | 31 TATAAG             |          |
|                        | sigX1      | SP0014     | GCATTTTAG             | 12 ACAaTGG           | 32 TAGACT             |          |
|                        |            | SM12261_0014| GCATTTTAG             | 12 ACAaTGG           | 32 TAGACT             |          |
|                        |            | SMSK321_1651| GCATTTTAG             | 12 ACAaTGG           | 32 TAAAT              |          |
|                        | sigX2      | SP2006     | GCATTTTAG             | 12 ACAaTGG           | 32 TAGACT             |          |
|                        |            | SM12261_0613| GCATTTTAG             | 12 ACAaTGG           | 31 TAAAT              |          |
|                        |            | SMSK321_1251| GCATTTTAG             | 12 ACAaTGG           | 32 TAGACT             |          |
| Fratricide immunity     | comM       | SP1945     | ACATTGAG              | 10 ACaTTCTC          | 13 TATAAT             |          |
|                        |            | SM12261_0911| ACATTGAG              | 12 ACATTGAG          | 32 TATAAT             |          |
|                        |            | SMSK321_1653| ACATTGAG              | 12 ACATTGAG          | 32 TATAAT             |          |
| Unknown functions        | ribF       | SP1110     | ACATTCAat             | 12 ACATTCAa          | 27 TATAG              |          |
|                        |            | SM12261_0833| ACATTCAa              | 12 ACATTCAa          | 27 TATAG              |          |
|                        |            | SMSK321_1087| TCATTGAG              | 12 ACATTGAG          | 27 TATAG              |          |
|                        | Peptide deformylase | SP1549 | ACATTGAG | 11 GCATTATc | 20 TATAAT |
|                        |            | SM12261_1162| ACATTGAG              | 11 GCATTATc          | 20 TATAAT             |          |
|                        |            | SMSK321_0137| ACATTGAG              | 11 GCATTGc           | 20 TATAAT             |          |
|                        | ABC transporter | SP1717 | ACATTCAa | 12 ACATTGAG | 31 TATAAT |
|                        |            | SM12261_0940| ACATTCAa              | 12 ACATTGAG          | 31 TATAAT             |          |
|                        |            | SMSK321_1696| ACATTCAa              | 12 ACATTGAG          | 31 TATAAT             |          |
|                        | SPFH domain | SP2156 | ACATTCAc | 12 ACATTGAG | 42 TATAAT |
|                        |            | SMSK321_1581| ACATTCAc              | 12 ACATTGAG          | 42 TATAAT             |          |
|                        | Hypoth. protein | SP0429 | – | CCATTGAG | 32 TACTAG |
|                        |            | SMSK321_0027| ACATTCAa              | 12 ACATTGAG          | 31 TATAAT             |          |

Bases matching the direct repeat consensus for *S. pneumoniae* are highlighted in bold [8]. Bases divergent from the consensus are represented by lower case letters. A TIGR4 orthologue and its ComE-box sequence is described for each *S. mitis* gene.

<sup>a</sup>Gene number from HOMD.

<sup>b</sup>Consensus sequence in pneumococcus according to [8].

upregulation of downstream genes, these account for a total of 26 SigX controlled genes. Among the nine SigX core genes with unknown functions in competence, eight were upregulated at least 2-fold in SK321 (ccs50, cfb1, yfiA, pepB, pilC, radC, ackA, cinA), and four in the type strain (yfiA, pilC, radC, cinA). In addition, a non-annotated gene upstream of ccs4 (SM12261_0853; SM5K321_1482) coding for a 46-aa long peptide was upregulated in both *S. mitis* strains. Although orthologues of this gene are upregulated during competence in *S. sanguinis* (gene SSA_2233) [31], *S. mutans* UA159 (gene SMU.2076) [29] and *S. pneumoniae* R6 (gene SPR_0181 annotated as orf47) [28], this peptide remains uncharacterized and its role in competence is still unknown. We identified five additional *S. mitis* genes with candidate sites matching the SigX-box consensus – but without orthologues in *S. pneumoniae* - (Table 4), which we discuss further below.
| Function in competence | S. pneumoniae ID | Mean fold-change | S. mitis ID | Mean fold-change | Description |
|------------------------|-----------------|-----------------|-------------|-----------------|-------------|
| SigX Core Orthologues [29] | TIGR4b | R6c | NCTC12261 | SK321 | NCTC12261 | SK321 |
| DNA uptake and recombination | 0023 10 | 0021 0342 | 1.4 | 2.3 | DNA repair protein RadA; radA |
| | 0978 16 | 1411 0150Ba | 58.6 | 139.6 | Competence protein; coIA |
| | 1266 64 | 1607 0695 | 63.8 | 339.3 | DNA protecting protein DprA; dprA |
| | 1908 64 | 0826 1193 | 98.6 | 697.9 | Single-strand binding protein family; ssbB |
| | 1940 16 | 0916 1658 | 5.3 | 188 | Protein RecA; recA |
| | 2208 128 | 0765 1614 | 113.2 | 591.7 | Competence protein ComFA; comFA |
| | 2207 64 | 01089f 00402f | 133.6 | 1167.9 | Competence protein, putative; comFC |
| | 1808 16 | 0438 1050 | 57.0 | 694.6 | Type 4 prepilin peptidase; pilC, pilD |
| | 0954 32 | 1388 0421 | 122.8 | 627.9 | ComE operon protein 1; comEA |
| | 0955 64 | 1389 0422 | 93.6 | 839.4 | DNA internalization-related competence protein ComEC |
| | 2047 64 | 0629 1267 | 78.3 | 218.4 | Methyltransferase small domain superfamily |
| | 2048 64 | 0630 1268 | 187.6 | 992.0 | ComG operon protein 6 |
| | 2049 64 | 01214f 1269 | 141.4 | 1016.1 | Hypothetical protein; cgIE |
| | 2050 64 | 0631 1270 | 164.6 | 852.6 | Competence protein |
| | 2051 64 | 0632 1271 | 192.2 | 1027.7 | Competence protein |
| | 2052 64 | 0633 1272 | 162.5 | 1079.0 | Competence protein CglB; comYB |
| | 2053 64 | 0634 1273 | 146.1 | 929.8 | Putative ABC transporter subunit ComYA; comYA |
| | 2201 64 | 0760 1609 | 134.5 | 458.5 | Choline binding protein D; ctbD |
| Lysis | 0021 10 | 0019 0340 | 1.2 | 1.6 | Deoxyuridine 5′-triphosphate nucleotidohydrolase, dut |
| Unknown functions in competence | 1981 8 | 0563 1204 | 1.3 | 4.3 | Competence-induced protein Ccs50; ccs50 |
| | 1980 8 | 01275f 00562f | 1.5 | 3.9 | CMP-binding factor; yhaM, cbfB |
| | 2206 32 | 0764 1613 | 2.2 | 4.1 | Ribosome-associated factor Y; y6A |
| | 0979 8 | 1412 0443 | 1.5 | 3.4 | Oligopeptidase F; pepB |
| | 0782 8 | 1288 0286 | 7.1 | 472 | Conserved domain protein; pilC |
| | 1088 32 | 1568 0646 | 167.3 | 297.6 | DNA repair protein RadC; radC |
| | 2044 32 | 0627 1265 | 1.6 | 4.5 | Acetate kinase; ackA |
| | 1941 32 | 0915 1657 | 21.7 | 55.3 | Competence/damage-inducible protein CinA; cinA |
| Non-Core Orthologues | Unknown functions in competence | 0024 8 | 0022 0343 | 1.4 | 2.4 | Carbonic anhydrase |
| | 0025 4 | 0023 0344 | 1.3 | 2.2 | Membrane protein, putative |
| | 0031 16 | 00941f 0345 | 6.7 | 1.2 | Hypothetical protein |
| | 0030 32 | 0025 0346 | 6.6 | 26.1 | Competence-induced protein; Ccs16 |
| | 0029 4 | 0026 0347 | 5.0 | 1.9 | Conserved hypothetical protein |
| | 2045 8 | 0628 1266 | 3.3 | 198.7 | Adenine-specific methyltransferase |
| | 2197 16 | 0757 1606 | 3.7 | 86.4 | ABC transporter substrate-binding protein |
| | 2198 16 | 0758 1607 | 3.0 | 118.7 | ABC transporter, permease protein |
Downregulated genes during competence development

Genes downregulated during the state of competence in the mitis group are fewer in number than upregulated sequences, and still not well characterized anywhere. SK321 isolate presented 70 genes with a 2- to 10-fold decrease in gene expression (Additional file 3: Table S3), which mostly included genes involved in sugar and amino acid metabolism, alcohol dehydrogenases, and hypothetical proteins, and 10 orthologues of previously reported \textit{S. pneumoniae} CSP-repressed genes [9]. \textit{S. mitis} type strain had fewer downregulated sequences, independently of the medium used (Additional file 1: Table S1 and Additional file 2: Table S2). Only two genes were commonly repressed in TSB and C + Y YB cultured cells, both uncharacterized hypothetical proteins, and the difference may be explained by the different growth conditions. Among the other downregulated sequences in the type-strain, none were orthologues of genes reported at least once as being repressed by CSP in any other streptococci. This is not an isolated observation, since a previous transcriptome analysis of \textit{S. pneumoniae} competent cells has shown strain-specific responses also for downregulated genes [28]. Furthermore, the downregulation of these genes can also be an indirect effect of a general stress response caused by CSP.

Identification of early and late genes in \textit{S. mitis} with no orthologues in \textit{S. pneumoniae}

Despite the close genetic kinship between these species, some \textit{S. mitis} upregulated genes lack orthologues in \textit{S. pneumoniae} strains (Fig. 5a, Additional file 4: Table S4). Four to five small adjacent ORFs located upstream the ABC transporter ComAB were found induced by CSP in both \textit{S. mitis} strains. SM12261_0044–0045 and SMSG321_1305–1306 encode peptides with a GG-type leader sequence and a bacteriocin moiety featuring a GxxxG-like motif [32]. The processing of SM12261_0044–0045 at the Gly-Gly site would give mature peptides of 20 and 32 amino acids residues, while SMSG321_1305–1306 processing would result in 40- and 33-amino-acids mature peptides, respectively. Interestingly, in silico analysis revealed no similarity between active peptide sequences of the two strains. However, the leader sequence is identical for the peptides encoded by SM12261_0045 and SMSG321_1305, which suggests they might be processed and exported by the same bacteriocin transporter. The third gene in each operon codes for another GG-leader peptide, and may play a role as a signaling peptide involved in bacteriocin production. Finally, based on the typical transcriptional pattern of class IIb bacteriocins, the fourth gene possibly codes for an immunity protein. In the promoter region of SM12261_0044 and SMSG321_1305, we identified an usual sequence (TTCGAATA) that matches the SigX box consensus, suggesting that these genes are regulated by SigX (Fig. 5a) [7, 33]. We confirmed this prediction by RT-PCR in a \textit{S. mitis} strain lacking the two copies of \textit{sigX} (Fig. 5b). Thus, this operon appears to be involved in the production and export of competence-related bacteriocins and part of the late CSP response in \textit{S. mitis}.

Both strains possess upregulated operon (SM12261_0749–0750; SMSG321_1598–1599) that codes for a membrane protein and a lipoprotein, respectively. We identified orthologues of SM12261_0750 and SMSG321_1599 in several \textit{S. sanguinis} strains, in addition to a dozen strains of \textit{S. mitis}, while orthologues of

---

### Table 3 \textit{S. mitis} type strain and SK321 orthologues of \textit{S. pneumoniae} late genes in response to CSP in C + Y YB (Continued)

| Function in competence | \textit{S. pneumoniae} | \textit{S. mitis} | Description |
|------------------------|------------------------|------------------|-------------|
|                        | TIGR4\textsuperscript{b} | NCTC12261\textsuperscript{a} | NCTC12261 | SK321 | NCTC12261 SK321 |
| 2199 4                 | 0759                  | 1608             | 3.7 | 92.9 | Conserved hypothetical protein |
| 1939 16                | 0917                  | 1659             | 4.4 | 10.8 | DNA-damage-inducible protein |
| 0980 6                 | 1413                  | 0444             | 1.7 | 3.9  | O-methyltransferase family protein |
| 0981 4                 | 1414                  | 0445             | 1.3 | 2.7  | Foldase protein PrsA |
| 1072 8                 | 1421                  | 0451             | 1.4 | 2.7  | DNA primase |
| 1073 8                 | 1422                  | 0452             | 1.3 | 3.1  | RNA polymerase sigma factor RpoD |
| 1074 8                 | 1423                  | 0453             | 1.3 | 3.0  | N-6 Adenine-specific DNA methylase YitW |
| 0201 8                 | 0853                  | 1482             | 3.2 | 13.5 | Competence-induced protein Ccs4; ccs4 |

\textsuperscript{a}Gene number and product from HOMD
\textsuperscript{b}Gene number and product from [51]
\textsuperscript{c}Mean fold-change induction of CSP in \textit{S. mitis} type strain and SK321 obtained by transcriptome analysis
\textsuperscript{d}Expression pattern during response the of \textit{S. pneumoniae} to CSP pheromone [9]
\textsuperscript{e}No upregulation detected by [9]
\textsuperscript{f}Gene number according to PROKKA annotation [52]
SM12261_0749 and SMSK321_1598 were only detectable in other \textit{S. mitis} strains. In \textit{S. sanguinis} SK36 (SSA_2192), this gene is upregulated during competence, and a SigX-box site marks its promoter region [31]. We found the same SigX-box sequence upstream of SM12261_0750 and SMSK321_1599 (Fig. 5a), and confirmed its activity through RT-PCR (Fig. 5b).

Another region composed by two ORFs coding for two hypothetical proteins was upregulated by CSP in both media tested, but only in the type strain. BLAST analyses revealed orthologue sequences of SM12261_0241 in other three \textit{S. mitis} strains (SK1073_0114, SK659_0006 and SK616_1612), and the gene annotation corresponds to a bacteriocin class II with double glycine leader peptide in all of them. Interestingly, the active part of the peptide encoded by SM12261_0241 is identical to the one encoded by isolate SK569. In turn, gene SM12261_0240 is annotated as a conserved hypothetical protein, but its orthologues in SK616_1615, SK569_0169 and SK579_0521 code for Enterocin A immunity proteins. We were not able to identify either ComE-box or SigX-box sequences in the promoter region of SM12261_0241. However, analysis of SM12261_0240-SM12261_0241 mRNA expression in a \textit{Δ}sigX1ΔsigX2 mutant showed upregulation of this transcript (Fig. 5b), confirming its independence from the alternative sigma factor regulation. Indeed, this suggests that this region might be suffering indirect ComE regulation or may simply reflect a broader specificity of ComE than previously thought.

Gene SMSK321_1184, coding for a hypothetical protein, was upregulated more than 100-fold in SK321 but has no orthologue in the type strain. Using BLAST analyses against streptococcal genomes, three orthologues of this gene with at least 70% identity were identified in \textit{S. mitis} SK1073, \textit{Streptococcus salivarius} 57.1 and \textit{S. sanguinis} SK1056. A putative SigX-box sequence with

---

**Fig. 3** Core genes SigX regulon of the mitis group. Synteny conservation of \textit{S. mitis} type strain and SK321, \textit{S. pneumoniae}, \textit{Streptococcus gordonii} and \textit{S. sanguinis}. Red pentagons correspond to genes immediately downstream of a SigX box. Genes induced by competence, black borders; upregulated genes oriented antisense to the gene downstream of a SigX box, dashed black borders; no change in gene expression, borderless faded colors; orthologues in each of the 15 gene groups are represented by similar colors; no orthologues in the regions analyzed, gray pentagons. SM17, \textit{S. mitis} type strain upregulated sequences in TSB and C + YYB media, gene locus tag as in GenBank (accession no. AEDXX0000000); SK321 gene locus tag as in GenBank (accession no. AEDT0000000). SP1, \textit{S. pneumoniae} Rx [9]; gene locus tag as in GenBank (\textit{S. pneumoniae} TIGR4, accession no. AE005672). SP2, \textit{S. pneumoniae} R6 [28]; gene locus tag as in GenBank (\textit{S. pneumoniae} TIGR4, accession no. AE005672). SP3, \textit{S. pneumoniae} G54 [50]; gene locus tag as in GenBank (\textit{S. pneumoniae} G54, accession no. CP001015). SG0, \textit{S. gordonii} Challis [36]; gene locus tag as in GenBank (\textit{S. gordonii} Challis, accession no. CP000725). SSA, \textit{S. sanguinis} SK36 [31]; gene locus tag as in GenBank (accession no. CP000387.1). a, b, c, d Upregulation was not > 2-fold. The image was modified from [29].
Fig. 4 (See legend on next page.)
one mismatch in the first nucleotide was detected upstream of SMSK321_1184 (Table 4), as well as in the promoter regions of all three orthologues. In fact, the presence of this gene in three different species might indicate the importance of this protein in competence throughout the evolution of salivarius and mitis groups of streptococci.

Figure 5c demonstrates that deletion of these unique CSP-responsive genes of S. mitis type strain and SK321 did not affect transformation yields in any of the knockout strains. Thus, although regulated by CSP, these genes seem dispensable for DNA uptake and processing in these strains.

Discussion

Commensal streptococci are pioneer colonizers of the oral cavity that attach to the tooth surfaces, and to which more pathogenic bacteria later adhere to establish a mature multispecies biofilm. This close attachment provides a genetic pool to competent oral streptococci, which become potential recipients for horizontally transferred DNA as well as latent reservoirs of important genetic elements, such as antibiotic resistance genes [15]. More importantly, this attribute may also increase the likelihood of survival for some members of the population during stress conditions. To date, the global transcriptome responses during competence have been studied in only three oral streptococci, S. mutans [29, 34, 35], S. sanguinis [31] and S. gordonii [36], and despite a few reports of competence in S. mitis [22, 30, 37], little or nothing is known about its regulation or response specificity in different strains. Recently, sequence of the genomes of a range of S. mitis revealed that truncation in genes required for competence was a common feature, present in roughly 40% of the strains [20]. However, even when apparently possessing a complete intact competence apparatus, only a few of the remaining strains displayed transformability under laboratory conditions [22]. Interestingly, the sigX is apparently the most common affected gene (absence and/or truncation), while DNA uptake and bacteriocin genes are largely conserved throughout S. mitis strains. Our transcriptome data revealed that the overall response of two competent S. mitis oral isolates to CSP resulted in significant upregulation of genes involved in the modulation of the competence cascade, DNA uptake and recombination, as well as lysis and bacteriocin production. Furthermore, we identified unique upregulated sequences, a fact that highlights the S. mitis singularity despite its close genetic resemblance to S. pneumoniae.

In the present study, the comparison between two S. mitis strains located in different evolutionary branches provided insight into the variation of the global transcriptomes of isolates of the same species under the same growth conditions. When cultured in the competence permissive medium C+YB, approximately 4% of the type strain genome positively responded to CSP, compared to almost 10% in SK321. All orthologues of ComE-responsive genes were upregulated at least 2-fold in SK321, while the type strain displayed a weaker response mainly for comW and comCDE (Table 1). However, this difference might be simply due to a smaller amplification of the CSP signal in this strain, since their transcriptomic maps show immediate expression downstream of their ComE binding sites. A recent study comprising transcriptomic data from at least five species from different phylogenetic groups, showed that streptococci regulate a core of 27 to 30 pan genes under the control of the alternative sigma factor SigX [29]. In the present analysis, the 12 SigX core genes required for DNA uptake in S. pneumoniae were upregulated in both strains, as well as the six genes required for DNA recombination (Table 3 and Fig. 3). Additionally, the SigX regulon comprises genes involved in lytic attack, and the key protein in fratricide, CbpD was strongly upregulated in both S. mitis isolates. Not surprisingly, among the orthologues of early genes and preceded by a ComE binding site was also comM, strongly implying that S. mitis employs a similar immunity mechanism to CbpD as does S. pneumoniae.

CSP-regulated bacteriocin production is a common feature among naturally competent streptococci. S. mutans, one of the most distinguished bacteriocin producers, coordinates the CSP-ComDE regulation with production of mutacins [38], while S. pneumoniae and S. gordonii carry a direct link between SigX regulation and bacteriocin production [39, 40]. The most striking difference when comparing S. mitis orthologues of S. pneumoniae late genes was the absence of the cibABC bacteriocin locus in the two strains studied. Previous reports have shown that some S. mitis strains did not maintain this competence-induced locus throughout their evolution from S. pneumoniae, and that they probably acquired other genes associated with the production
of killing factors [41, 42]. Indeed, we identified a strongly upregulated bacteriocin locus in a single transcriptional unit located upstream of the comAB operon, with a conserved link to the late competence response by carrying two (in the type strain) or even three (in SK321) copies of the SigX box in their promoter regions (Fig. 5b).

Although there are no previous reports, nor clear explanations for the role of multiple SigX box sequences, we hypothesize that recurrent recombination events in this region might have either left or removed additional promoter sequences. Puzzlingly, in S. pneumoniae TIGR4, this region is occupied by the bacteriocin encoding gene

| Table 4 | Alignment of DNA sequences found upstream of S. mitis type strain and SK321 putative late genes. Distances to the start codon of first ORF are shown |
| Function in competence | Annotation | Gene ID | Consensus |
|------------------------|------------|---------|-----------|
| SigX Core Orthologues [29] | DNA recombination | coiA | SM12261_1411 | TTTTT\(\)TACGAATA\(\)−22-\(\)ATG |
| | | SMSK321_na | TTTTT\(\)TACGAATA\(\)−21-\(\)ATG |
| | | dpra | SM12261_1607 | TTTTT\(\)TACGAATA\(\)−21-\(\)ATG |
| | | smsk21_0695 | TTTTT\(\)TACGAATA\(\)−21-\(\)ATG |
| | | sspB | SM12231_0826 | TTT\(\)−1-\(\)TACGAATA\(\)−25-\(\)ATG |
| | | | SM12231_1193 | TTT\(\)−2-\(\)TACGAATA\(\)−20-\(\)ATG |
| | | comGA | SM12261_0634 | TTTTT\(\)−9-\(\)TACGAATA\(\)−25-\(\)ATG |
| | | SMSK321_1273 | TTTTT\(\)−9-\(\)TACGAATA\(\)−25-\(\)ATG |
| DNA uptake | comFA | SM12261_0765 | TTTTT\(\)TACGAATA\(\)−25-\(\)ATG |
| | | SMSK321_1614 | TTTTT\(\)TACGAATA\(\)−25-\(\)ATG |
| | | ciiC | SM12261_0438 | TTTTT\(\)TACGAATA\(\)−25-\(\)ATG |
| | | SMSK321_1050 | TTTTT\(\)TACGAATA\(\)−25-\(\)ATG |
| | | comEA | SM12261_1388 | TTTTT\(\)−8-\(\)TACGAATA\(\)−25-\(\)ATG |
| | | SMSK321_0421 | TTTTT\(\)−8-\(\)TACGAATA\(\)−25-\(\)ATG |
| Fratricide | cbpD | SM12261_0760 | TTTTTTT\(\)TACGAATA\(\)−57-\(\)ATG |
| | | SMSK321_1609 | TTTTTTT\(\)TACGAATA\(\)−49-\(\)ATG |
| Unknown functions | dut | SM12261_0019 | TTTTT\(\)−1-\(\)TACGAATA\(\)−30-\(\)ATG |
| | | SMSK321_0340 | TTTTT\(\)−1-\(\)TACGAATA\(\)−30-\(\)ATG |
| | | ccs50 | SM12261_0563 | TTT\(\)−18-\(\)ACGAATA\(\)−74-\(\)ATG |
| | | SMSK321_1204 | TTT\(\)−18-\(\)ACGAATA\(\)−74-\(\)ATG |
| | | piiC | SM12261_1288 | TTTTT\(\)−9-\(\)TACGAATA\(\)−74-\(\)ATG |
| | | | SMSK321_1657 | TTTTTTT\(\)TACGAATA\(\)−17-\(\)ATG |
| Non-Core Orthologues | Unknown function | S\(^{\prime}\) ccs4 | SM12261_50853 | TTTTT\(\)−2-\(\)TACGAATA\(\)−19-\(\)ATG |
| | | | SMSK321_51482 | TTTTT\(\)−2-\(\)TACGAATA\(\)−19-\(\)ATG |
| | | ccs 16 | SM12261_50025 | TTTTT\(\)−2-\(\)TACGAATA\(\)−19-\(\)ATG |
| | | | SMSK321_50346 | TTTTT\(\)−2-\(\)TACGAATA\(\)−19-\(\)ATG |
| Non-Core Without Orthologues | Putative bacteriocin | Cib-like operon | SM12261_0044 | TTTTT\(\)TACGAATA\(\)−33-\(\)ATG |
| | | | SMSK321_1305 | TTTTT\(\)TACGAATA\(\)−229-\(\)ATG |
| | | Unknown function | Lipoprotein | SM12261_0750 | TTTTT\(\)TACGAATA\(\)−147-\(\)ATG |
| | | | SMSK321_1599 | TTTTT\(\)TACGAATA\(\)−147-\(\)ATG |
| | | Hypot. protein | SMSK321_1184 | TTTTT\(\)−12-\(\)ACGAATA\(\)−13-\(\)GTC |

Bases in agreement to cinbox (SigX-box) consensus in bold [33]. Bases divergent from the consensus are represented by lower case letters. Only the first genes within each induced transcriptionally active region are shown.
blpU preceded by a Blp-box, while in S. pneumoniae D39 there is also a transposase gene transcribed in the opposite direction [41]. This suggests that a shuffling between Blp and competence-induced bacteriocins might have occurred during the parallel evolution of the two species. Besides this bacteriocin operon, we also detected other upregulated sequences without orthologues in S. pneumoniae. These accounted for hypothetical proteins, membrane proteins, lipoproteins and even other bacteriocin-like encoding operon. In fact, by comparative analyses of S. mitis type strain and S. pneumoniae TIGR4, D39 and G54, Kilian et al. [20] showed that 100 S. mitis proteins do not present any homology with any S. pneumoniae strain, and that 83 of the 100 S. mitis proteins without homologues in S. pneumoniae had homologues in S. sanguinis, S. gordonii, S. agalactiae and S. thermophilus and none lacked homologues in other bacteria. While there is no information on whether there are competence-related proteins among these, the presence of S. mitis strain specific genes suggests that they were acquired more recently in evolution, independently by individual S. mitis lineages.

Conclusions
Data gathered throughout the last two decades have provided great understanding about the streptococcal regulation of competence in various groups of this genus, reinforcing the fact that competence for genetic transformation is a conserved trait among streptococci. Overall, our results demonstrated that in two S. mitis isolates CSP induces a global change in gene expression in two S. mitis isolates that not only supports the maintenance of the competent state and the DNA uptake machinery, but also strongly induces expression of genes involved in lysis and bacteriocin production. Most of the previously described competence-induced loci in other streptococci were detected by our method together with several other novel genes, for which functions remain to be elucidated. Furthermore, promoter analysis of the genes not previously known to be induced during S. mitis competence suggests that several of them belong to either the ComE or the SigX regulons. These findings reveal conservation of the competence system in transformable S. mitis strains and highlight the characteristics of strain-specific regulated regions. Particularly, our findings are significant not only from a fundamental understanding of competence in streptococci, but also from a practical perspective, as transformation is an important tool to explore gene functions and to design S. mitis for potential applications such as vaccine development.
Methods

Bacterial strains and growth conditions

All bacterial strains and isogenic derivatives used in this study are listed in Additional file 5: Table S5. Bacterial stocks were stored at −80 °C in Todd Hewitt Broth (THB, Becton Dickinson and Company, Le Pont de Claix, France) or Tryptic Soy Broth (TSB, Soybean-Casein Digest medium, BactoTM) supplemented with 30% glycerol. Pre-cultures were prepared from fresh liquid cultures grown in TSB at 37 °C 5% CO2 until an absorbance of 0.5 at 600 nm (optical density at 600 nm [OD600]). Pre-cultures of NCTC12261 T and SK321 were diluted 100-fold in C + YVB and grown at 37 °C 5% CO2 until an OD600 of 0.04 was reached. The cultures were then divided in two, with one half being treated with 150 nM synthetic CSP and the other half kept untreated. Cultures were incubated for 15 min and pellets were harvested at 8000 g, 4 °C for 10 min. Total RNA was extracted using the High Pure Isolation Kit (Roche, Mannheim, Germany) and treated with Turbo DNase (Ambion, Life Technologies, Carlsbad, California, USA) to clear any DNA contamination. Complementary DNA templates were prepared from RNA First Strand cDNA Synthesis Kit (Thermo Fisher Scientific, Fermentas) according to manufacturer’s protocol. Housekeeping gyrA gene was used to validate the results. The primers for the studied genes were designed by using Primer3Web Platform for uniformity in size (80-150 bp) and melting temperature. The primers sequences are provided in Additional file 4: Table S4. PCR conditions included an initial denaturation at 95 °C for 10 min, followed by a 40-cycle amplification consisting of denaturation at 95 °C for 30 s and annealing and extension at 55 °C for 1 min. Data were collected and analyzed with the software MxPro (Stratagene).

Real time PCR

Pre-cultures of MI014 were diluted 100-fold in C + YVB to a final volume of 10 ml and incubated at 37 °C 5% CO2 until an OD600 of 0.04 was reached. The cultures were then divided in two with one half being treated with 150 nM synthetic CSP and the other half kept untreated. Cultures were incubated for 15 min and pellets were harvested at 8000 g, 4 °C for 10 min. Total RNA was extracted using the High Pure Isolation Kit (Roche, Mannheim, Germany) and treated with Turbo DNase (Ambion, Life Technologies, Carlsbad, California, USA) to clear any DNA contamination. Complementary DNA templates were prepared from RNA First Strand cDNA Synthesis Kit (Thermo Fisher Scientific, Fermentas) according to manufacturer’s protocol. Housekeeping gyrA gene was used to validate the results. The primers for the studied genes were designed by using Primer3Web Platform for uniformity in size (80-150 bp) and melting temperature. The primers sequences are provided in Additional file 4: Table S4. PCR conditions included an initial denaturation at 95 °C for 10 min, followed by a 40-cycle amplification consisting of denaturation at 95 °C for 30 s and annealing and extension at 55 °C for 1 min. Data were collected and analyzed with the software MxPro (Stratagene).

RNA sequencing

Pre-cultures of NCTC12261 T and SK321 at OD600 of 0.5 were centrifuged at 8000 g, 4 °C, for 10 min, resuspended 100-fold in TSB or C + YVB and the diluted cultures were incubated at 37 °C 5% CO2 until an OD600 of 0.04 was reached. Then, the total volume of culture was divided in two tubes and grown in the presence or absence of 150 nM of CSP. Following, cells were harvested for RNA extraction at 10000 g, 4 °C, for 10 min. Procedures for RNA extraction, RNA enrichment and preparation of DNA library for sequencing using Illumina® HiSeq were carried out as described elsewhere [47]. Following sequence run, a FASTQ file was derived from each sample. For differential expression analysis of genes, raw read counts for the S. mitis type strain and SK321 transcripts were generated using a Perl script based on the mapped read profiles of the two strains, as previously described [48].

Synthetic peptides

Lyophilized CSPs of S. mitis NCTC12261T (NH2-EIRQTHNFFNFFKR-COOH) and SK321 (NH2-ESRLPKIRFDFFPFPRK-K-COOH) were synthesized by GenScript (GenScript Corporation, NJ), with purity > 95%. Stock solutions were prepared by re-suspending the material in distilled water to a concentration of 10 mM, and stored at −20 °C. Working stocks at a 100 μM concentration were aliquoted and stored at −20 °C.

Transformation

Transformation kinetics was carried out as previously described [22]. Briefly, pre-cultures of NCTC12261 T and SK321 were diluted 100-fold in C + YVB and grown at 37 °C, 5% CO2 until an OD600 of 0.04 was reached. After the addition of 300 nM of CSP, 1 μM recombinant plasmid pVA838 was added at various time points. Cells were incubated at 37 °C for 30 min before addition of 20 U ml−1 DNasel (Roche, DNaseI recombinant, 10 U ml−1), followed by incubation at 37 °C for further 40 min to remove extracellular DNA. Transformants were selected on blood agar plates supplemented with Erythromycin by 24 h of incubation at 37 °C, 5% CO2.

Construction of mutants

For construction of the S. mitis ΔsigX1ΔsigX2 (MI014), two techniques were used. First, the standard PCR ligation mutagenesis strategy was employed [45], with minor modifications, to delete sigX1. Briefly, the sigX1 flanking regions were amplified using primer pairs FP395–FP396 and FP397–FP398. The kanamycin resistance cassette (Km5) was amplified using the primer pairs FP001–FP068. Further, ligation and purification of the PCR products were performed using T4 DNA ligase (Fermentas) and the QIAquick PCR purification kit (Qiagen), respectively. The final product was transformed into S. mitis NCTC12261 T. The specific insertional inactivation of sigX2 was performed with a recombinant integrational vector, pSF152 [46]. An internal fragment of sigX2 was amplified by primer pair FP451-FP452 and was forced-cloned via BamHI and EcoRI 5’ tags of the respective primers, into the corresponding sites of plasmid pSF152. The final products were transformed into the parent strains as previously described [22]. All primers used for mutant constructions are listed in Additional file 5: Table S5.
software package [49] was used for assessment of differential expression levels when comparing samples. In total, 16 transcriptomes were included in the study. The results are derived as means of 12 samples grown in the presence or absence of CSP in three independent biological experiments of S. mitis type strain and SK321 grown in C + YBY. For growth in TSB, the results are derived as means of four samples collected in two independently conducted biological experiments. Genes that exhibited > 2-fold increase in mean read (with a p value < 0.05) were considered differentially expressed. Transcriptome maps were visualized using Jbrowser at http://bioinformatics.forsyth.org/mtd/.

Additional files

**Additional file 1:** Table S1. Type strain upregulated genes (> 2-fold) in response to CSP in TSB [51, 52]. (DOCX 18 kb)

**Additional file 2:** Table S2. Type strain genes up and downregulated (> 2-fold) in response to CSP in C + YBY. (DOCX 19 kb)

**Additional file 3:** Table S3. SK321 genes upregulated (> 2-fold) in response to CSP in C + YBY. (DOCX 33 kb)

**Additional file 4:** Table S4. CSP responses by the type strain and SK321 genes without orthologues in S. pneumoniae. (DOCX 13 kb)

**Additional file 5:** Table S5. Strains, mutants and primers used in this study. (DOCX 19 kb)

Abbreviations

CSP: Competence-stimulating peptide; DR: Direct repeat; GG: Glycine-glycine; HOMD: Human oral microbiome database; OD: Optical density; PCR: Polymerase chain reaction; RNASP: RNA polymerase; RNA-seq: RNA sequencing; RT-PCR: Reverse transcription polymerase chain reaction; TAR: Transcriptionally activated region; THB: Todd-Hewitt Broth; TSB: Tryptic Soy Broth

Acknowledgements

The sequencing service for this work was provided by the Norwegian Sequencing Centre (www.sequencing.uio.no).

Funding

This work received funding from the Faculty of Dentistry, University of Oslo and from the Norwegian surveillance system for antibiotic resistance in microbes (NORM). All the RNA-Seq reads have been deposited to NCBI Sequence Read Archive (SRA) and are available under the BioProject Accession ID PRJNA436941.

Availability of data and materials

All the RNA-Seq reads have been deposited to NCBI Sequence Read Archive (SRA) and are available under the BioProject Accession ID PRJNA436941.

Authors’ contributions

GS, RJ, DM, and FP participated in the conception and design of the work, as well as in the analysis and interpretation of data; GS, RJ and HAA conducted the experiments; TC generated and carried out bioinformatics analyses; all authors participated in drafting and revising the work for intellectual content. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable

Competing interests

The authors declare that they have no competing interests.

Publisher’s Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

1Institute of Oral Biology, Faculty of Dentistry, University of Oslo, Postboks 1052, Blindern, 0316 Oslo, Norway. 2Department of Microbiology, The Forsyth Institute, Cambridge, MA, USA. 3Department of Biological Sciences, College of Liberal Arts and Sciences, University of Illinois at Chicago, Chicago, IL, USA.

Received: 14 March 2018 Accepted: 18 May 2018

Published online: 13 June 2018

References

1. Håvarstein LS, Coomaraswamy G, Morrison DA. An unmodified heptadecapeptide pheromone induces competence for genetic transformation in Streptococcus pneumoniae. Proc Natl Acad Sci U S A. 1995; 92(24):11140–4.

2. Bricker AL, Camilli A. Transformation of a type 4 encapsulated strain of Streptococcus pneumoniae. FEMS Microbiol Lett. 1999;172(2):131–5.

3. Hui FM, Zhou L, Morrison DA. Competence for genetic transformation in Streptococcus pneumoniae: organization of a regulatory locus with homology to two lactococcin A secretion genes. Gene. 1995;153(1):25–31.

4. Pestova EV, Håvarstein LS, Morrison DA. Regulation of competence for genetic transformation in Streptococcus pneumoniae by an auto-induced peptide pheromone and a two-component regulatory system. Mol Microbiol. 1996;21(4):853–62.

5. Luo P, Li H, Morrison DA. ComX is a unique link between multiple quorum sensing outputs and competence in Streptococcus pneumoniae. Mol Microbiol. 2003;50(2):623–33.

6. Håvarstein LS, Martin B, Johnsborg O, Granadel C, Claveyres JP. New insights into the pneumococcal fratricide: relationship to clumping and identification of a novel immunity factor. Mol Microbiol. 2006;59(4):1297–307.

7. Pestova EV, Morrison DA. Isolation and characterization of three Streptococcus pneumoniae transformation-specific loci by use of a lacZ reporter insertion vector. J Bacteriol. 1998;180(10):2701–10.

8. Ween Ø, Gaustad P, Håvarstein LS. Identification of DNA binding sites for ComE, a key regulator of natural competence in Streptococcus pneumoniae. Mol Microbiol. 1999;33(4):817–27.

9. Peterson SN, Sung CK, Cline R, Desai BV, Snipersen EU, Luo P, Walling J, Li H, Mintz M, Tsegaye G, et al. Identification of competence pheromone responsive genes in Streptococcus pneumoniae by use of DNA microarrays. Mol Microbiol. 2004;51(4):1051–70.

10. Berg KH, Ohnstad HS, Håvarstein LS. LytF, a novel competence-regulated murein hydrolase in the genus Streptococcus. J Bacteriol. 2012;194(3):627–35.

11. Kaumsløy L, Johnsborg O, Lunde M, Knutsen E, Håvarstein LS. Choline-binding protein D (CbpD) in Streptococcus pneumoniae is essential for competence-induced cell lysis. J Bacteriol. 2005;187(13):4338–45.

12. Whatmore AM, Efratiou A, Pickrell AP, Broughton K, Woodard G, Sturgeon D, George R, Dowson CG. Genetic relationships between clinical isolates of Streptococcus pneumoniae, Streptococcus oralis, and Streptococcus mitis: characterization of “atypical” pneumococci and organisms allied to S. mitis harboring S. pneumoniae virulence factor-encoding genes. Infect Immun. 2000;68(3):1374–82.

13. Pearce C, Bowden GH, Evans M, Fitzsimmons SP, Johnson J, Sheridan MJ, Wintzen R, Cole MF. Identification of pioneer viridans streptococci in the oral cavity of human neonates. J Med Microbiol. 1995;42(1):67–72.

14. L J, Helmerhorst EJ, Leenheer OW, Trowler RF, Yaskell T, Haffajee AD, Socransky SS, Oppenheim FG. Identification of early microbial colonizers in human dental biofilm. J Appl Microbiol. 2004;97(6):1311–8.

15. Kilian M, Ridley DR, Jensen A, Bruggemann H, Tettelin H. Parallel evolution of a low affinity PBP2B in Streptococcus pneumoniae; the role of Streptococcus mitis in pathogenic and mutualistic lifestyles. MBio. 2014;5(4):e01490–8.

16. Kilian M, Ridley DR, Jensen A, Bruggemann H, Tettelin H. Parallel evolution of a low affinity PBP2B in Streptococcus pneumoniae; the role of Streptococcus mitis in pathogenic and mutualistic lifestyles. MBio. 2014;5(4):e01490–8.

17. Dowson CG, Coffey TJ, Kell C, Whiley RA. Evolution of penicillin resistance in Streptococcus pneumoniae: the role of Streptococcus mitis in the formation of a low affinity β-lactamase in S. pneumoniae. Mol Microbiol. 1993;9(3):635–43.

18. Chasenbeeck R, Balmelle N, Weber B, Gardes C, Keck W, de Saizieu A. Mosaic genes and mosaic chromosomes intras- and interspecies genomic variation of Streptococcus pneumoniae. Infect Immun. 2001;69(4):2477–86.

19. Jensen A, Scholz CF, Kilian M. Re-evaluation of the taxonomy of the mitis group of the genus Streptococcus based on whole genome phylogenetic analyses, and proposed reclassification of Streptococcus dentisani as

20. Jensen A, Scholz CF, Kilian M. Re-evaluation of the taxonomy of the mitis group of the genus Streptococcus based on whole genome phylogenetic analyses, and proposed reclassification of Streptococcus dentisani as...
Streptococcus oralis subsp. denticantis comb. nov., Streptococcus tigurinus as Streptococcus oralis subsp. tigurinus comb. nov., and Streptococcus oligofermentans as a later synonym of Streptococcus cristiatus. Int J Syst Evol Microbiol. 2016;66(1):4803–20.

20. Kilian M, Poulsen K, Blomqvist T, Hävarstein LS, Bek-Thomsen M, Tettelin H, Sorensen LB. Evolution of Streptococcus pneumoniae and its close commensal relatives. PLoS One. 2008;3(9):e2683.

21. Daifalla N, Cayabuay MJ, Xie E, Kim HB, Tzipori S, Stashenko P, Duncan M, Campos-Neto A. Commensal Streptococcus mitis is a unique vector for oral mucosal vaccination. Microbes Infect. 2015;17(3):427–42.

22. Salvadori G, Junges R, Morrison DA, Petersen FC. Overcoming the barrier of low efficiency during genetic transformation of Streptococcus mitis. Front Microbiol. 2016;7:1009.

23. Petersen FC, Schiee AA. Natural transformation of oral streptococci. Methods Mol Biol. 2010;666:167–80.

24. Allolio G, Martin B, Granadel C, Claverys JP. Development of competence in Streptococcus pneumoniae: pheromone autoinduction and control of quorum sensing by the oligopeptide permease. Mol Microbiol. 1998;29(1):75–83.

25. Piotrowski A, Luo P, Morrison DA. Competence for genetic transformation in Streptococcus pneumoniae: termination of activity of the alternative sigma factor ComX is independent of proteolysis of ComX and ComW. J Bacteriol. 2009;191(10):3359–66.

26. Berge MJ, Kamgoua A, Martin B, Polard P, Campo N, Claverys JP. Midcell recruitment of the DNA uptake and virulence nucleic acid, EndA, for pneumococcal transformation. PLoS Pathog. 2013;9(9):e1003596.

27. Peterson S, Cline RT, Tettelin H, Sharov V, Morrison DA. Gene expression analysis of the Streptococcus pneumoniae competence regulon by use of DNA microarrays. J Bacteriol. 2000;182(21):6192–20.

28. Dagkessamanskaia A, Moscoso M, Henard V, Guiral S, Overweg K, Reuter M, Martin B, Wells J, Claverys JP. Interconnection of competence, stress and Cia regulons in Streptococcus pneumoniae: competence triggering stationary phase autolysis of ciaR mutant cells. Mol Microbiol. 2004;54(4):1071–86.

29. Khan R, Rukke HV, Hovik H, Amdal HA, Chen T, Morrison DA, Petersen FC. Comprehensive transcriptome profiles of Streptococcus mutans UA159 map Core streptococcal competence genes. mSystems. 2016;1(2):e00038-15.

30. Johnborg O, Eldholm V, Bjørnstad ML, Hävarstein LS. A predatory mechanism dramatically increases the efficiency of lateral gene transfer in Streptococcus pneumoniae and related commensal species. Mol Microbiol. 2008;69(1):245–53.

31. Rodríguez AM, Callahan JE, Fawcett P, Ge X, Xu P, Kitten T. Physiological and molecular characterization of genetic competence in Streptococcus gordonii. Mol Oral Microbiol. 2011;26(2):99–116.

32. Nissen-Meyer J, Oppegard C, Rogne P, Haugen HS, Kristiansen PE. Structure and mode of action of the two-antibody (class-IIb) Bacteriocins. Probiotics Antimicrob Proteins. 2010;2(1):52–60.

33. Campbell EA, Choi SY, Masure HR. A competence regulon in Streptococcus pneumoniae revealed by genomic analysis. Mol Microbiol. 1998;27(5):929–39.

34. Reck M, Tomash J, Wagner-Dobler I. The alternative sigma factor SigX controls Bacteriocin synthesis and competence, the two quorum sensing regulated traits in Streptococcus mutans. PLoS Genet. 2015;11(7):e1005353.

35. Dufour D, Cordova M, Cvikrovich DG, Levesque CM. Regulation of the competence pathway as a novel role associated with a streptococcal bacteriocin. J Bacteriol. 2011;193(2):652–9.

36. Vickerman MM, Lobst S, Jesionowski AM, Gill SR. Genome-wide transcriptional changes in Streptococcus gordonii in response to competence signaling peptide. J Bacteriol. 2007;189(2):7799–807.

37. Gaustad P. Genetic transformation in Streptococcus suis: Effects on genetic transformation by culture filtrates of Streptococcus suis. Int J Syst Bacteriol. 1996;46(3):650–6.

38. Dufour D, Tagg JR, Tompkins GR. Competence-dependent bacteriocin production by Streptococcus gordonii DL1 (challis). J Bacteriol. 2007;189(4):1468–72.

39. Claverys JP, Martin B, Hävarstein LS. Competence-induced fratricide in streptococci. Mol Microbiol. 2007;64(6):1423–33.