An analysis of LysM domain function in LytE when fulfilling the D,L-endopeptidase requirement for viability in *Bacillus subtilis*

Laura Uelze[1]

Smurfit Institute of Genetics, Trinity College Dublin, Dublin 2. Ireland.

[1] present address: German Federal Institute for Risk Assessment, Berlin, Germany.

Keywords: LytE D,L-endopeptidase, *Bacillus subtilis*, LysM domain function

Abstract

The D,L-endopeptidase requirement states that *Bacillus subtilis* requires either the activity of the LytE or the CwlO enzyme for viability, therefore proving that these two enzymes can complement for each other despite their very different N-terminal domains. Here, we show that another D,L-endopeptidase, LytF, can also fulfill the D,L-endopeptidase requirement for viability, when expressed from the *cwlO* promoter. Both LytE and LytF contain N-terminally located LysM domains, three and five respectively. However, cells expressing another very similar D,L-endopeptidase CwlS, with four LysM domains were not viable. This led us to investigate whether a LytE protein with any one of its three LysM domains permuted can fulfill the D,L-endopeptidase requirement for viability. We found that the three LysM domains are not functionally equivalent and that the N-terminally located LysM domain plays a greater role for functioning of the LytE enzyme than the subsequent domains. Based on an investigation of orthologous enzymes in 19 *B. subtilis* species we propose an evolutionary model describing the development of the LytE-, CwlS- and LytF-type D,L-endopeptidases and their LysM domain repeats. In summary, these results show that the LytE enzyme has been optimized to fulfill the D,L-endopeptidase requirement for cell viability of *B. subtilis* with regard to the number and properties of LysM domains that mediate peptidoglycan-binding.

INTRODUCTION:

Peptidoglycan and teichoic acids are the main structural components of the *Bacillus subtilis* cell wall (Bhavsar and Brown, 2006; Barreteau *et al.*, 2008; Sauvage *et al.*, 2008; Vollmer *et al.*, 2008; Sekiguchi and Yamamoto, 2012). Peptidoglycan is composed of glycan chains formed by polymerization of a disaccharide of *N*-acetylglucosamine (Glc*N*Ac) and *N*-acetylmuramic acid (Mur*N*Ac). Stem pentapeptides composed (proximal to distal order) of amino acids L-alanine (position 1), D-glutamate (position 2), *meso*-diaminopimelate (position 3), D-alanine (position 4) and D-alanine (position 5) extend from the lactyl group of Mur*N*Ac. The glycan strands are linked by formation of a covalent peptide bond between the carboxyl group of D-alanine (position 4) of one strand [with concomitant loss of D-alanine (position 5)] and the amino group of *meso*-diaminopimelate (position 3) of a different strand. Glycan strands can be up to 200 disaccharides in length with a stem peptide linkage index of up to ~33%. Thus peptidoglycan forms a mesh outside the cell to which wall-teichoic acids composed of poly(ribitol- or glycerol-phosphate) are covalently attached (Weidenmaier and Peschel, 2008; Swoboda et al., 2010; Kawai et al., 2011). This rigid mesh structure confers shape on the cell and is capable of withstanding turgor pressures of up to 5 atmospheres (Sekiguchi and Yamamoto, 2012). Interference with peptidoglycan metabolism, either through inhibition of its synthesis by antibiotics (*e.g.* penicillin and vancomycin) or an imbalance between synthesis and turnover leads to aberrant cell structure, cell division and often cell death. Thus maintenance of cell wall integrity is a prerequisite for normal cell viability.

The dramatic changes that occur in cell wall structure during growth and division require exquisite control and coordination of cell wall synthesis and turnover. Autolysins play a key role in sculpting the cell wall during these processes by cleavage of glycosidic linkages in the glycan chains and peptide linkages in the stem peptides (Smith *et al*., 2000; Sekiguchi and Yamamoto, 2012). *B. subtilis* encodes a cohort of 35 enzymes with glucosaminidase, muraminidase, amidase and endopeptidase activities that confer the capability to cleave all covalent linkages of peptidoglycan thereby allowing it to be sculpted, remodeled and turned over during cell growth and division (Smith *et al*., 2000; Sekiguchi and Yamamoto, 2012). A notable feature of the autolysin cohort of *B. subtilis* is the degree of redundancy of individual enzymatic activities (Smith *et al*., 2000; Sekiguchi and Yamamoto, 2012). For example, there are seven enzymes (CwlO, CwlS, LytE, LytF, PgdS, YkfC, YqgT) with D,L-endopeptidase activity encoded in the *B. subtilis* genome, each of which cleaves the stem peptide between D-glutamate (position 2) and *meso*-diaminopimelate (position 3). However these enzymes differ in the type and/or number of domains that interact with cell wall components. Thus the LytE, CwlS and LytF D,L-endopeptidases have 3, 4 and 5 LysM domains respectively located at the amino terminus of the protein, a domain that functions to bind proteins to peptidoglycan (Mesnage *et al*., 2014). In contrast, the N-terminal domain of CwlO functions to bind the enzyme to the FtsEX ATP-type transporter that is located in the cell membrane (Yamaguchi *et al.*, 2004; Dominguez-Cuevas *et al.*, 2013; Meisner *et al.*, 2013). Thus variance in these proteins regions allows differential localization and binding affinity of these enzymes to elements located within the cell envelope.

While none of the seven D,L-endopeptidases are essential, *lytE* and *cwlO* are

synthetically lethal, signifying that the cell requires one or the other of these enzymes for viability (Bisicchia *et al.*, 2007). This indicates that they perform an overlapping essential function for the cell, probably in elongation of the cell wall during growth (Bisicchia *et al.*, 2007; Hashimoto *et al.* 2012). Expression of both CwlO and LytE are positively regulated by the WalRK two-component signal transduction system (TCS). In addition, LytE expression is controlled by the SigI/RsgI sigma factor, which is induced by heat and other cell wall stresses (Tseng *et al.*, 2011; Liu *et al.,* 2017). Therefore the the WalRK and SigI/RsgI regulatory systems combine to fulfill the D,L-endopeptidase requirement during normal growth and stress conditions respectively.

The synthetic lethality of the CwlO and LytE enzymes is interesting on two accounts (BIsicchia *et* al., 2012). Firstly, these enzymes bind to different regions and structures within the cell envelope. The three LysM domains direct binding of LytE to peptidoglycan in the cell wall (Ishikawa *et al.*, 1998; Buist *et al.*, 2008; Mesnage *et al*., 2014). In contrast, the amino terminal region of CwlO binds to the FtsEX protein complex resulting in tethering of the active enzyme to the outer face of the cell membrane (Dominguez-Cuevas *et al.*, 2013; Meisner *et al.*, 2013). The second feature is that the high level of redundancy among the cohort of autolysins in *B. subtilis* has hampered assigning specific functions to individual enzymes. However, the requirement for D,L-endopeptidase activity evident from the synthetic lethality of CwlO and LytE permits investigation of the ability of individual enzymes with this activity to fulfill that requirement.

In this study, the ability of individual enzymes to fulfill the D,L-endopeptidase requirement for viability when individually expressed using the *cwlO* promoter is

investigated. In addition, we tested the LysM domains of LytE to establish which, and how many, are needed to fulfill the D,L-endopeptidase requirement for viability. Our results show that both LytE (3 LysM domains) and LytF (5 LysM domains), but not CwlS (4 LysM domains) can fulfill the D,L-endopeptidase requirement for viability. In addition we show that the three LysM domains of LytE are not functionally equivalent, but that each alone is sufficient to fulfill the D,L-endopeptidase requirement for viability.

MATERIAL AND METHODS

Bioinformatic analysis

To investigate LysM domain-containing proteins, all entries containing the term 'LysM' were extracted from UniProt, (55,014 hits in January 2016). These data were then screened to identify *bona fide* LysM domains using the following criteria: (i) they must be from a bacterial species and be specific to an individual bacterial strain (*e.g. Bacillus subtilis* 168); (ii) putative domains must have a range of 40 – 70 amino acids (Buist *et al.*, 2008; Desvaux *et al.*, 2018); (iiii) each protein must have a unique amino acid. The resulting dataset contained 38,576 LysM domains located in 26,078 proteins that were encoded by 6,971 different bacterial strains in 68 phyla.

A total of 150 LysM domain-containing proteins in this dataset were identified in nineteen closely related strains of the *B. subtilis* species group (NCBI taxonomy). Cluster analysis using ClustalOmega with default settings was performed to identify the subset of proteins with a high level of identify in their catalytic domain, thereby identifying orthologues of CwlS, LytE and LytF D,L-endopeptidases orthologs in these

B. subtilis 168 strains (Sievers *et al.*, 2011; Bodenhofer *et al.*, 2015). The alignment was then used for metric multidimensional scaling conducted with the bios2mds package (Pele *et al.*, 2012) to identify clusters of orthologs. Similarly, the relationships among the 185 LysM domains contained in the cohort of identified orthologs was established based on an alignment of their amino acid sequence (ClustalOmega, substitution matrix = BLOSUM80), followed by construction of a phylogenetic tree using Maximum Likelihood via the phangorn package (Schliep, 2011). The best fitting amino acid substitution models were identified (LG+G+I), the different model parameters subsequently optimized and bootstrapping performed (100 replicates), before the circular phylogenetic tree was visualized with ggtree (Yu *et al.*, 2016).

The phylogeny of the 19 *Bacillus subtilis* strains was inferred through alignment-free genome comparison with feature frequency profiles (FFP) (Sims *et al.*, 2009). Genome sequences (excluding plasmid sequences) of all species were retrieved from GenBank. *E. coli* BL21(DE3) was included as outgroup. FPP was performed essentially as described by Wang and Ash (2015). Neighbour Joining (NJ) trees were constructed based on the Jensen-Shannon divergence matrices. It was found that varying lengths of l between 20 and 25 resulted in highly similar trees. Accounting for the higher degree of relatedness between species investigated in this study compared to Wang and Ash (2015), length was adjusted to l = 24 in the final analysis and the NJ tree visualized with ggtree (Yu *et al.*, 2016). The robustness of the tree was verified with bootstrapping and bootstrapping values of 100 out of 100 were achieved for all branches.

137 **Strains, plasmids and growth conditions**

138 Bacterial strains and plasmids used in this study are listed in Supporting Information
139 Tables S2 and S3 respectively. Oligonucleotides were obtained from IDT, (Belgium) and
140 are listed in the supporting information Table S4. *Escherichia coli* strain TG-1 (Sambrook
141 *et al.*, 1989) was used for propagating plasmids. *E. coli* strain BL21(DE3) was used for
142 protein overexpression and purification. Strains were grown in Luria-Bertani (LB)
143 medium and in salt-free buffered LB (Boylan *et al.*, 1993), to which NaCl was added up
144 to the desired concentration. The high phosphate defined medium (HPDM) and low
145 phosphate defined medium (LPDM) used, are described in Botella *et al*., (2011).
146 Antibiotics were added at concentrations described by Noone *et al*., (2014). OD_{600} was
147 measured using a UVmini-1240 UV-VIS spectrophotometer (Shimadzu Scientific
148 Instruments). *B. subtilis* 168 was transformed according to the procedure of
149 Anagnostopoulos and Spizizen (1960). Growth studies were conducted with the BioTek
150 Synergy™2 system (BioTek Instruments) at 37°C with slow shaking. Three technical
151 repeats were performed for each growth curve.

152 **Strain construction**

153 A *lytE* null mutant strain (LU1) was constructed by replacing the *lytE* gene with a
154 spectinomycin resistance gene (*spc*) The *spc* gene was amplified from plasmid
155 pDG1727 using primer pairs Spec_Fwd / Spec_Rev. Chromosomal regions located
156 upstream and downstream of *lytE* were amplified by PCR using primer pairs
157 LytE_Up_Fwd / LytE_Up_Rev_Spec and LytE_Do_Fwd_Spec / LytE_Do_Rev
158 respectively. These DNA fragments were joined to the fragment encoding the *spc* gene

159 by strand overlap PCR (Horton *et al.*, 1989). This linear DNA fragment was then
160 transformed into *B. subtilis* strain 168 with selection for spectinomycin resistance,
161 thereby generating strain LU1 in which the *lytE* gene is replaced by the *spc* gene.
162 A second *lytE* null mutant strain (LU331) was generated in a similar manner. The strain
163 differentiates from LC01, in that part of the *lytE* signal sequence was preserved in this
164 mutant. The *spc* gene was amplified as described previously. Chromosomal regions
165 located upstream and downstream of *lytE* were amplified by PCR using primer
166 pairs LytE_Up_Fwd_2 / Signal_LytE_Rev_Spec and LytE_Do_Fwd_Spec /
167 LytE_Do_Rev. The DNA fragments were joined to the fragment encoding the *spc* gene
168 by strand overlap PCR and the linear DNA fragment was transformed into WT168, with
169 selection for spectinomycin resistance, to generate strain LU331.
170 Strains were constructed that express the LytE protein from the *cwlO* promoter (LU17,
171 LU21, LU25). These strains were generated through transcriptional fusion between the
172 *cwlO* promoter and the *lytE* gene. To generate strain LU17, four PCR fragments were
173 amplified. First, the DNA fragments located upstream and downstream of *cwlO* locus
174 were amplified with the primer pairs CwlO_UF_Sma / CwlO_UR and QE93_Fwd_Ery /
175 QE_94_Rev respectively. Then the *lytE* open reading frame was amplified with the
176 primer pair LytE_DF and LytE_DR_Ery. Lastly, the erythromycin resistance gene *erm*
177 was amplified with primer pair Mls_Fwd / Mls _Rev from plasmid pDG6464. All PCR
178 fragments were joined together by strand overlap extension PCR and transformed into
179 strain EB35 with selection for erythromycin resistance. The resulting strain LU17
180 encodes two copies of the *lytE* gene (one at the *cwlO* locus and one at the *lytE* locus).
181 Subsequently, the *lytE* gene at the *lytE* locus was replaced with a spectinomycin

resistance gene (*spc*). For this purpose, the spc gene was amplified from LU1 with the primer pair LytE_Up_Fwd / LytE_Do_Rev. The PCR fragment was then transformed into LU17 with selection for spectinomycin resistance. The resulting strain LU21 encodes one copy of the *lytE* gene at the *cwlO* locus.

Another similar strain (LU25) was generated with the additional feature that the leader region of *cwlO* was deleted, thereby stabilizing the *lytE* mRNA and yielding higher LytE protein expression (Noone *et al*., 2014). For this strain, a DNA fragment upstream of *cwlO* from strain which has a 141 bp markerless deletion of the leader region (Noone *et al.*, 2014) was amplified from strain DN3000. This PCR fragment was transformed into LU21 with selection for erythromycin resistance to generate strain LU25.

In order to investigate LytE protein expression levels, the LytE protein was joined to a HA-tag which can be detected by western analysis. Strains LU45, LU47, LU61 and LU63 are derivatives of strains WT168, LS103, LU21 and LU25 in which the *lytE* gene is fused with a C-terminal HA-tag. These strains were constructed by transforming plasmid pLU2 into WT168, LS103, LU21 and LU25 with selection for kanamycin resistance. The successful integration of the plasmid into a *cwlO* mutant strain (LS103) served as proof for the functionality of the HA-tag, given that a failure would lead to a combined *cwlO* / *lytE* mutant, which is synthetically lethal (Bisicchia *et al.*, 2007).

A markerless null *lytF* mutant-was generated by the method of Bloor and Cranenburgh, 2006. The DNA regions upstream and downstream of *lytF* were amplified with primers LytF_Up_Fwd/ ytF_Up_Rev and LytF_Do_Fwd/LytF_Do_Rev respectively. The *cat* gene was amplified using primers Cat_Diff_Fwd/Cat_Diff_Rev and plasmid pDG1661 as template. A *dif* site for recombination was introduced into the 5' end of each primer

(Table S4). The three DNA fragments were then ligated and transformed into WT168 with selection for chloramphenicol resistance. Transformants were then grown without for 24 hours in LB medium without selection, plated onto LB agar plates and screened for loss of chloramphenicol resistance. The null mutation in *lytF* in the resultant strain LU74 (Δ*lytF*) was confirmed by PCR.

In order to generate a strain that expresses the LytF protein from the *cwlO* promoter (with *cwlO* leader – LC111, without *cwlO* leader – LC125) a similar approach as for LU21 and LU25 was followed. Downstream fragments and the erythromycin resistance cassette were amplified as described previously. The upstream fragment was amplified with the primers CwlO_UF and CwlO_UR_LytF from WT168 (containing the *cwlO* leader for LU111) or from DN3000 (no *cwlO* leader for LU125). A middle fragment (containing the *lytF* open reading frame) was amplified with the primers LytF_Fwd / LytF_Rev_Ery. The fragments were joined together by strand overlap extension PCR and the PCR product was transformed into LU74 with selection for erythromycin. The resulting strains LU111 and LU125 encode a copy of the *lytF* gene from the *cwlO* promoter (with / without *cwlO* leader) and the *lytE* gene at the *lytE* locus.

Subsequently, the *lytE* gene at the *lytE* locus was replaced with a spectinomycin resistance gene (*spc*). For this purpose, the spc gene was with the primer pair LytE_Up_Fwd / LytE_Do_Rev from chromosomal DNA of LU. The PCR fragment was then transformed into the parental strains LU111 and LU125 with selection for spectinomycin resistance. The resulting strains LU116 and LU134 encode one copy of the *lytF* gene from the *cwlO* promoter (with / without *cwlO* leader) and are knock-outs for cwlO and *lytE*.

228 In order to investigate LytE / LytF protein expression levels, the LytE / LytF protein were
229 joined to a HA-tag which can be detected by western analysis. A C-terminal fusion
230 between either the *lytE* or the *lytF* gene and a HA-tag was achieved by transforming the
231 pLU2 plasmid into strain LU17 to generate strain LU143 and plasmid pLU6 into WT168,
232 LU111, LU116, LU125 and LU134 to generate strains LU147, LU151, LU153, LU155 and
233 LU157 respectively, with selection for kanamycin.
234 To generate strains containing a variant LytE expressed from the *cwlO* promoter in a
235 background with endogenous *lytE* gene, the plasmids pLU16, pLU15, pLU14, pLU17,
236 pLU18, pLU19, pLU20, pLU21, pLU22, pLU23, pLU24, pLU25, pLU26, pLU27, pLU28,
237 pLU30, pLU31, pLU29, pLU37, pLU38 and pLU40, were transformed by double cross-
238 over into WT168 with selection for kanamycin to generate the strains: LU239, LU207,
239 LU212, LU216, LU220, LU223, LU227, LU232, LU235, LU261, LU265, LU269, LU272,
240 LU276, LU280, LU286, LU288, LU312, LU314, LU315 and LU318.
241 To generate strains containing a variant LytE expressed from the *cwlO* promoter in a *lytE*
242 null mutant background, the plasmids pLU16, pLU15, pLU14, pLU17, pLU18, pLU19,
243 pLU20, pLU21, pLU22, pLU23, pLU24, pLU25, pLU26, pLU27, pLU28, pLU30, pLU31,
244 pLU29, pLU38, and pLU40 were transformed by double cross-over into LS089 with
245 selection for kanamycin to generate the strains: LU205, LU209, LU213, LU217, LU222,
246 LU226, LU229, LU233, LU238, LU262, LU266, LU270, LU273, LU277, LU281, LU287,
247 LU289, LU313, LU316 and LU320.
248 Strains were constructed that express a variant LytE from the *lytE* promoter (with intact
249 *cwlO* gene). For the construction of these strains four PCR fragments were amplified.
250 The upstream and the downstream fragments flanking the *lytE* locus were amplified with

251 the primer pairs LytE_Up_Fwd_2 / Signal_LytE_Rev and LytE_Do_Fwd_Ery /
252 LytE_Do_Rev from chromosomal DNA of WT168. The downstream fragment was joined
253 to an erythromycin resistance cassette, which was amplified from pDG646 with the
254 primers Mls_Fwd / Mls_Rev. The upstream fragment was joined to fragments of the
255 variant *lytE* genes, which were amplified with the primers Signal_LytE_Fwd /
256 Catalytic_LytE_Rev_Ery from the plasmids pLU40, pLU23, pLU24, pLU25, pLU20,
257 pLU27, pLU37, pLU16, pLU17, pLU14 and pLU29. After joining of the upstream
258 fragment (upstream homology to the *lytE* locus plus variant *lytE* gene sequence) to the
259 downstream fragment (erythromycin resistance cassette plus downstream homology to
260 the *lytE* locus) by strand overlap extension PCR, the linear PCR fragments were then
261 transformed into strain LU331 with selection for erythromycin resistance, to create the
262 strains LU335, LU337, LU339, LU341, LU372, LU374, LU398, LU399, LU401, LU403
263 and LU405 respectively.
264 In order to disrupt the *cwlO* gene in a series of strains, expressing a variant LytE from
265 the *lytE* promoter, a fragment encompassing the *cwlO* locus was amplified with the
266 primer pair CwlO_UF / QE94_Rev from chromosomal DNA of the *cwlO* null mutant
267 strain EB35. The PCR fragment was then transformed by double cross-over into the
268 strains LU335, LU337, LU339, LU341, LU372, LU374, LU399, LU401, LU403 and
269 LU405 with selection for kanamycin resistance to obtain strains LU343, LU350, LU352,
270 LU354, LU383, LU385, LU407, LU409, LU411 and LU413.
271 In order to investigate variant LytE protein expression levels, the variant LytE proteins
272 were joined to a HA-tag which can be detected by western analysis. To enable a C-
273 terminal fusion of the variant LytE proteins with a HA-tag, plasmid pLU46 with homology

to the *lytE* catalytic domain, followed by the HA-tag and was integrated by single cross-over at the *lytE* locus in the strains LU205, LU213, LU217, LU229, LU262, LU266, LU320, LU343, LU385, LU407, LU411, LU313, LU350, LU270, LU352 and LU413 to generate the strains LU414, LU415, LU416, LU417, LU418, LU419, LU421, LU422, LU424, LU425, LU427, LU432, LU435, LU440, LU442 and LU443. Strains were selected for spectinomycin resistance.

Plasmid construction

Plasmid pLU2, pLU6 and pLU46 enable a C-terminal transcriptional fusion between the either the *lytE* (pLU2, pLU6) or the *lytF* (pLU46) open reading frame and a HA-tag. Plasmids pLU2 and pLU6 were constructed by cloning a fragment, with homologous region to the C-terminus of the *lytE* or the *lytF* gene with the HA-tag sequence inserted before the stop codon, into the parental vector pDG780. The fragments was amplified with the following primer pair LytE_Tag_F_BamHI / LytE_Tag_R_HA_EcoRI; LytF_Tag_F_BamHI / LytF_Tag_R_HA_EcoRI and cloned into pDG780 by use of the BamHI and EcoRI restriction sites. Plasmid pLU46 is a derivative from pLU2 with a spectinomycin resistance cassette. Exchange of the antibiotic resistance cassette was achieved through amplifying the spectinomycin resistance gene from LU1 with Spec_Fwd_EcoRI and Spec_Rev_SalI and cloning it into pLU2 via the SalI and EcoRI restriction sites.

A parental plasmid (pLU10) was constructed for convenient cloning of variant *lytE* open reading frames into a cassette that would allow integration of variant *lytE* genes at the *cwlO* locus, thereby exchanging the *cwlO* against the *lytE* open reading frame. Plasmid

296 pLU10 was generated in a two-step process: First, upstream homology (containing the
297 *cwlO* leader and promoter and part of the *lytE* signal sequence, with a silent BamHI
298 restriction site) was amplified with CwlO_UF_LysM and CwlO_UR_LysM and cloned into
299 pdG780 via the NotI and BamHI restriction sites to create plasmid pLU9. Secondly,
300 downstream homology was amplified with the primer pair CwlO_DF_LysM and
301 CwlO_DR _LysM and cloned into the vector pLU9 via the XhoI and KpnI restriction sites
302 resulting in plasmid pLU10.
303 To obtain plasmid pLU37, harbouring a variant *lytE* open reading frame without a LysM
304 domain, a fragment spanning the signal sequence, the spacer and the sequence of the
305 catalytic domain was amplified with the primers LytE_Spacer_Fwd_Sig /
306 Catalytic_LytE_Rev_XmaI and cloned into pLU10 via the XmaI and BamHI sites.
307 Three plasmids, that harbour a variant *lytE* open reading frame with a single LysM
308 domain each, were obtained in a similar way, following a two-step process. (i) First, the
309 three LysM domains sequences were amplified with the primer pairs
310 L1_Fwd_Sig_BamHI / L1_Rev_Spac, L2_Fwd_Sig_BamHI / L2_Rev_Spac and
311 L3_Fwd_Sig_BamHI / L3_Rev_Spac. (ii) Secondly the *lytE* catalytic domain was
312 amplified with the primer pair Catalytic_LytE_fwd_spac / Catalytic_LytE_rev_XmaI. Then
313 the fragments generated in (i) were joined individually to the PCR fragment (ii), utilizing
314 the spacer sequence as overlapping region. The final fragments were then cloned into
315 pLU10 after digestion of the vector and the fragments with XmaI and BamHI to generate
316 the plasmids pLU23 (*lytE[L1cat]*), pLU24 (*lytE[L2cat]*) and pLU25 (*lytE[L3cat]*).
317 Nine plasmids were generated, harbouring a variant *lytE* open reading frame with two
318 LysM domains each. To do so, the three LysM domains of *lytE* were amplified with two

319 primer pairs each (creating six PCR fragments in total). One primer pair generated part
320 of the *lytE* signal sequence, followed by the LysM domain, followed by the linker
321 sequence (L1_fwd_sig_BamHI / L1_rev_link, L2_fwd_sig_BamHI / L2_rev_link,
322 L3_fwd_sig_BamHI / L3_rev_link). The second primer pair generated the linker
323 sequence followed by the LysM domain, followed by parts of the spacer
324 sequence (L1_fwd_link / L1_rev_spac, L2_fwd_link / L2_rev_spac, L3_fwd_link /
325 L3_rev_spac). For cloning purposes, a silent BamHI restriction site was introduced into
326 the signal sequence and a silent SpeI restriction site was integrated in the linker
327 sequence. The fragments were then combined and joined to the *lytE* catalytic sequence
328 as described before and cloned into the parental vector pLU10, digested with BamHI
329 and XmaI. This process resulted in the plasmids: pLU14 (*lytE[L13cat]*),
330 pLU15 (*lytE[L12cat]*), pLU16 (*lytE[L11cat]*), pLU17 (*lytE[L21cat]*), pLU18 (*lytE[L22cat]*),
331 pLU19 (*lytE[L23cat]*), pLU20 (*lytE[L31cat]*), pLU21 (*lytE[L32cat]*) and
332 pLU22 (*lytE[L33cat]*)
333 Plasmids for the generation of variant LytE with three LysM domains were obtained by
334 cloning PCR fragments of single LysM domains into vectors already harbouring a *lytE*
335 open reading frame with two LysM domains. The three individual LysM domain
336 sequences of *lytE* were amplified with the following primer pairs: L1_fwd_ link / L1_rev_
337 link (L1), L2_fwd_link / L2_rev_link (L2), L3_fwd_ link / L3_rev_link (L3). The individual
338 LysM domains could then be cloned into the parental vectors after digestion of the
339 vectors and the fragments with SpeI. The number and orientation of integrated domains
340 was verified by a screening PCR and sequencing. Cloning the L1 fragment in pLU20,
341 pLU19, pLU18 and pLU14 resulted in the plasmids pLU26, pLU29, pLU31 and pLU38.

Cloning the L2 fragment in pLU18 and pLU14 resulted in the plasmid pLU30 and pLU40. Cloning the L3 fragment in pLU20 and pLU19 resulted in the plasmids pLU27 and pLU28.

To create plasmids that would allow the overexpression and purification of variant LytE proteins in *E. coli*, the respective protein coding sequences (excluding the start codon and the *lytE* signal sequence) were amplified with the forward primers Exp_L1_Fwd, Exp_L3_Fwd, Exp_L1_Fwd, Exp_L1_Fwd, Exp_L3_Fwd and Exp_Spac_Fwd from the plasmids pLU40, pLU20, pLU14, pLU23, pLU25 and pLU37 to generate the plasmids pLU47, pLU49, pLU48, pLU50, pLU51 and pLU52. The same reverse primer (Exp_Cat_Rev) was used in all reactions. PCR fragments were cloned into the NdeI/XhoI sites of the vector pET21b+ thereby re-generating the start codon and generating a C-terminal fusion between the protein and a 6xHis-tag.

RNA extraction and RT-qPCR

RNA was extracted and transcripts were quantified using RT-qPCR as previously described by Salzberg *et* al., 2013). The following primer were used to amplify the transcripts: 16s (16S_Bsu_F / 16S_Bsu_R), *cwlO* (YvcE_5´_qPCR_orf*2 / YvcE_3´_qPCR_orf*2), *lytE* (LytE_qPCR_F2 / LytE_qPCR_R2), *lytF* (LytF_qPCR_fwd / LytF_qPCR_rev), *mreBH* (MreBH_qPCR_F / MreBH_qPCR_R) and *sigI* (SigI_Chip_fwd / SigI_Chip_rev).

Western analysis

Western analysis was performed as described in Noone *et al.* (2012).

363 **Immunofluorescence microscopy**

364 Immunofluorescence microscopy was performed as described previously (Noone *et al.*,
365 2012).

366 **Protein expression and purification**

367 The variant LytE proteins were expressed in *E. coli* strain BL21(DE3) which harboured
368 plasmids pLU47 to pLU53. Protein expression and purification was performed as
369 described previously (Howell *et al.*, 2006).

370

371 **Insoluble peptidoglycan pull down assay**

372 To determine the binding capacity of variant LytE proteins to peptidoglycan an insoluble
373 peptidoglycan pull down assay was performed as described by Wong *et al.* (2014).
374 Purified cell wall of *B. subtilis* 168 was prepared as described in Salzberg and Helmann
375 (2007) (the number of washes was increased to 15). Forty-four pmoles of variant LytE
376 protein were mixed with purified cell wall of *B. subtilis* in assay buffer (50 mM Tris/HCl,
377 pH 8, 0.05 M NaCl, 5 mM ß-mercaptoethanol, 1 % TWEEN20) in a 25 µl reaction
378 volume. Purified cell wall was added until OD600 measurements of the mixture reached
379 a final absorbance of 10. For a negative control, 44 pmole of the wild type LytE protein
380 were mixed with the assay buffer, without added cell wall material. All mixtures were
381 incubated for 10 minutes at 10°C and centrifuged for 10 minutes at 10500 g. Both, the
382 supernatant and the pellet were collected. The pellet was resuspended in 15 µl assay
383 buffer.

The level of LytE protein that bound to the cell wall, and that which remained in the supernated was established on a 12.5 % SDS / PAGE gel. The proteins were visualized by coomassie blue staining and the image was quantified with the ImageJ software (Schneider *et al.*, 2012). To test the influence of salt on the binding of variant LytE proteins to cell walls, NaCl were added to the assay buffer at final concentrations of 0.05 M, 0.1 M, 0.2 M, 0.3 M, 0.4 M, 1 M. Two replicates of the pull down assay at 1 M NaCl were conducted to ensure the reproducibility of the assay. A negative control was included in all experiments.

RESULTS:

Bioinformatic analysis of the LysM domains

There are seven D,L-endopeptidases encoded in the *B. subtilis* genome, three of which contain LysM domains (*cwlS*, *lytE*, *lytF*). These three enzymes are structurally very similarly and their C-terminally located catalytic domains exhibit a high sequence similarity. However, they distinguished by the number of LysM domains with LytE, CwlS and LytF having three, four and five N-terminally located domains respectively. Although the high sequence similarity of their catalytic domains indicates a close evolutionary relationship, the manner in which these three enzymes evolved to have different numbers of LysM domains is unknown. Therefore, we sought to establish the evolutionary relationship between the CwlS, LytE and LytF D,L-endopeptidases and to investigate the relationships between their LysM domains, by identifying orthologs of these enzymes in closely related bacterial species and comparing their LysM domains

through cluster analysis.

Orthologs of CwlS, LytE and LytF were identified in 19 *Bacillus* species by the levels of amino acid identity shown by their catalytic domains (the complete list of proteins is shown in Table S1 of Supplementary Material). Results showed that all strains possessed an ortholog to LytE and LytF. However, only 13 of the 19 Bacillus species encoded a CwlS ortholog. Interestingly, the number of LysM motifs within these orthologs ranged from two to five. By these characteristics the investigated strains could be divided into groups that agreed with their respective phylogenetic lineages.

Group (A) is comprised mainly of *Bacillus licheniformis* and *Bacillus paralicheniformis* strains. Members do not possess a CwlS ortholog, but encode a LytE and a LytF orthologs with 2 and 4 LysM domains respectively. Group (B) is comprised mainly of *Bacillus amyloliquefaciens* strains, each having a CwlS ortholog with 4 LysM domains, and LytE and LytF orthologs with 2 and 3 LysM domains respectively. Group C is completely comprised of *Bacillus subtilis* strains each encoding a CwlS (4 LysM domains), LytE (3 LysM domains) and LytF orthologs (5 LysM domains). Interestingly, *Bacillus subtilis* strain XF-1 is unusual, in that it has a CwlS orthologs with 3 instead of 4 LysM domains (group X). The relationships between the 185 LysM domains in this cohort of LytE, LytF and CwlS enzymes was established based on an alignment of their amino acid sequence, followed by estimation of a phylogenetic tree with Maximum Likelihood method. The results (Figure 1), supported by high bootstrap values show that (i) individual LysM domains cluster according to their position in the protein relative to the catalytic domain; (ii) the N-terminal LysM domain is significantly diverged from those at the other positions; (iii) LysM domains group according to the enzyme type of their

respective catalytic domains (i.e. LytE, LytF, CwlS) and (iv) the LysM domains of LytF and CwlS are more similar to each other than to those of LytE. The three LysM domains of the CwlS orthologue of *Bacillus subtilis* XF-1 (M4KSM2) are in that each clusters with the LysM domains at positions 2, 3, and 4 respectively, rather than with those at positions 1, 2, and 3. Since M4KSM2 is the only CwlS ortholog identified with three instead of four LysM domains, it is likely that this enzyme has undergone a deletion, losing the N-terminal LysM domain.

To further investigate LysM domain duplication, the Uniprot database was interrogated to identify the occurrence of identical LysM domains. Of ~39,000 LysM domains identified, 46 occurred more than once, involving a total of 107 individual domains. However the duplicated domains were located in the same protein in all instances, and were adjacent to each other in 97 of the 107 instances discovered. This suggests that duplication of LysM domains occur within open reading frames with no evidence of horizontal domain transfer between proteins.

These results indicate that LysM domain expansion and reduction contributes to gene divergence and that LysM domains are transferred through stable vertical inheritance.

LytE can fulfil the D,L-endopeptidase requirement for cell viability when expressed from the *cwlO* promoter

As discovered by Bisicchia et al. (2007), LytE and CwlO are synthetically lethal, i.e. the cell requires one or other of these enzymes for viability. It is important to note that although LytE and CwlO possess similar C-terminal domains, they have different N-terminal domains and are expressed under different circumstances. The regulation of

CwlO expression and activity is well described by Noone et al. (2014). In short, the TCS WalRK is the sole activator for CwlO expression, which is therefore only expressed during exponential growth, when the WalRK system is active (Salzberg *et al.*, 2013). LytE expression is also activated by the WalRK system, but LytE expression can also be triggered by active SigI under stress conditions (Tseng and Shaw, 2008; Tseng *et al.*, 2011; Hashimoto *et al.*, 2012).

We sought to investigate whether LytE when expressed from the *cwlO* promoter can complement expression of CwlO. Since it was shown that the *cwlO* leader sequence destabilizes the *cwlO* transcript (Noone *et al.*, 2014) we were also interested in seeing if the same effect applied to the *lytE* transcript.

We decided to express LytE under the control of the *cwlO* promoter in a background with a disrupted endogenous *lytE* gene [Figure 2, a].

In comparison to the wild type, *lytE* transcript levels in such a strain were 20 to 30-fold elevated in exponentially growing cells as determined by RT-qPCR [Figure 2, c]. Furthermore a slight increase (~ 2 to 5-fold) in *sigI* and *mreBH* transcript levels was observed. The removal of the *cwlO* leader sequence caused a further increase in *lytE* transcript and a decrease in *sigI* transcript level (~ 2-fold). We confirmed that this can be attributed to the increased longevity of the *lytE* transcript through a mRNA half-life experiment [data not shown]. Elevated *lytE* transcript levels matched increased LytE protein levels [Figure 2, d]. High levels of LytE protein were found both cell-associated and secreted into the medium. However, the LytE protein appeared to be degraded in the medium as cultures approached stationary phase.

Although removal of the *cwlO* leader sequence led to a drop in SigI levels, we observed

473 that overexpressing LytE at this high level affects cell morphology with a fraction of cells
474 showing bulges and thickening [Figure 2, b]. In short, both *lytE* null mutants as well as
475 strains overexpressing LytE, exhibit difficulties to maintain cell shape and longitudinal
476 growth.
477 In summary, this results confirm that the LytE D,L-endopeptidase can complement CwlO
478 when expressed using *cwlO* expression signals - both transcriptional (i.e. promoter) and
479 post-transcriptional (i.e. *cwlO* leader sequence). It is clear however that the two
480 enzymes are not equivalent since the SigI regulon (signalling cell wall stress), is induced
481 when *lytE* is expressed using the normal *cwlO* expression signals (i.e. promoter and
482 leader sequence), but to a lesser extent when the *cwlO* leader sequence is removed.

483 **LytF can also fulfil the D,L-endopeptidase requirement**

484 The synthetic lethality of the CwlO / LytE autolysins allows further investigation of
485 whether any other D,L-endopeptidase encoded in the genome can complement their
486 activity. Thus, it can be investigated which D,L-endopeptidase can support growth and
487 viability when expressed under the control of the *cwlO* promoter.
488 In view of these findings, we sought to establish if the D,L-endopeptidases LytF and
489 CwlS, - which we previously showed to be closely related to LytE (LysM domains), as
490 well as another autolytic enzyme, - YqgT (no LysM domain), can fulfil the D,L-
491 endopeptidase requirement when expressed using the *cwlO* promoter.
492 To establish whether the CwlS, LytF or YqgT D,L-endopeptidases when expressed from
493 the *cwlO* promoter can fulfill the D,L-endopeptidase requirement, we constructed strains
494 in which the *cwlS, lytF* and *yqgT* genes were individually expressed at the *cwlO* locus.

The level of transcript of each gene was varied by expressing each gene both with the *cwlO* leader sequence (giving an unstable transcript) and without the leader sequence (giving a stable transcript). The ability of each enzyme to fulfil the D,L-endopeptidase requirement for viability was then established by attempting to disrupt the *lytE* gene in each of the strains, which was successful for *lytF* but not for *cwlS* or *yqgT*. LytF levels in the strains expressing LytF from the *cwlO* promoter were highly elevated when compared to that observed in strains, in which the enzyme was expressed from its natural promoter [Figure 3, b]. Despite the high level of cellular *lytF* mRNA transcript, only a very low level of full-length LytF protein was cell associated, with a high fraction found to be processed [Figure 3, c]. HA-tagged LytF protein was predominantly localized at the cell septa and cell poles, but was also present along the lateral cell cylinder [Figure 3, d]. Based on the *sigI* transcript levels, it seems that significant cell wall stress is not generated in these strains [data not shown]. However, strains overexpressing LytF from the *cwlO* promoter displayed major growth defects (delayed lag phase) and cell morphology defects [Figure 3, a].

Thus, while cell viability can be achieved by replacing the CwlO and LytE D,L-endopeptidases with expression of LytF, cells are not robust.

The individual LysM domains of LytE are not functionally equivalent

We could show that LytE (3 LysM domains) and LytF (5 LysM domains), but not CwlS (4 LysM domains) can support growth of *B. subtilis*, when expressed from the *cwlO* promoter. Interestingly although the majority of LysM domain containing proteins in the dataset constructed in section one possess only a single LysM domains (~ 67 % of

517 18,246 proteins), all identified orthologs to LytE, CwlS and LytF feature at least 2 LysM
518 domains.
519 There is little known about how individual LysM domains bind to peptidoglycan and the
520 significance of the fact that the identified D,L-endopeptidase orthologs possess a
521 minimum of 2 LysM domains.
522 Therefore we sought to investigate the properties of the individual LysM domains of LytE
523 by designing LytE enzymes with a single LysM domain. We then tested the
524 characteristics of these LytE variants proteins in strains which viability was depended on
525 the permutated autolysin. This was achieved by expressing the variant LytE containing
526 either L1, L2 or L3 in single copy at the *cwlO* or the *lytE* locus and attempting to disrupt
527 either the endogenous *lytE* or the *cwlO* gene [Supplementary Figure 2]. For control
528 purposes, we also included a variant LytE without any LysM domain as well as a
529 reconstituted wild type LytE with the three LysM domains in the wild type combination
530 (L1-L2-L3) in the further analysis.
531 We found that all constructed strains were viable, demonstrating that any single LysM
532 domain of LytE is capable of inferring functionality to the LytE enzyme. A control strain
533 with a LytE protein without a LysM domain could not be constructed, demonstrating that
534 the LytE enzyme requires at least one LysM domain for functionality.
535 Despite the fact that any individual LysM domain enabled a functional LytE enzyme, we
536 detected a significant change in cell morphology with shorter and thicker cells for the
537 variant strains. Interestingly this effect was greater for strains containing a LytE enzyme
538 with L2 or L3 than L1 [Figure 4, a].
539 We suggest that the extent to which a variant LytE protein fulfils the D,L-endopeptidase

requirement can be assessed by determining the extent to which the SigI regulon is induced. This was determined by establishing the mRNA transcript levels of the SigI regulon genes *sigI* and *mreBH* by RT-qPCR. We observed that transcript levels were elevated if a strain contained a variant LytE with a single LysM domain with the combination L2, L3, however no elevation occurred with L1 or with the reconstituted wild type protein [Figure 4, b].

We further decided to test the robustness of these strains by determining their sensitivity to salt stress. As expected, all controls grew normally at NaCl concentrations up to 0.4 M. In contrast, although all variant LytE strains grew comparable to wild type under the absence of salt or at a salt concentration of 0.1 M, growth inhibitions were observed at a salt concentration of 0.2 M and greater [Figure 4, c]. The extent, to which salt stress caused a change in growth behaviour, correlated with the type of LysM domains as well as the expression level (*cwlO* vs *lytE* promoter). Higher expression levels (*cwlO* promoter) improved the resistance of strains to salt stress. Salt tolerance was also increased when the LytE enzyme contained the first LysM domain (L1) as compared to L2 and L3. Since sensitivity to salt stress was not altered in the control strains which were in possession of an intact copy of the endogenous *lytE* or *cwlO* gene, we conclude that it is not the presence of the variant LytE that causes the increased sensitivity to salt stress, but rather a lack of D,L-endopeptidase enzymatic activity.

Our results indicate that although a variant LytE protein with any individual LysM domain is functional, it does not fulfil the D,L-endopeptidase requirement to the same extent as the wild type LytE protein. Cells that express a variant LytE with a single LysM domain, under the absence of the endogenous *cwlO* and *lytE* gene display growth and

morphology defects. Furthermore our experiments show that the three LysM domains of LytE are not functionally equivalent and that L1 is capable of conferring on the catalytic domains the ability to fulfil the D,L-endopeptidase requirement to a greater extent than L2 and L3 as can be seen from the induction of the *sigI* regulon.

LysM domain requirement of LytE

We observed that the three LysM domains are not equal in their ability to confer functionality to the LytE enzyme with the N-terminally located LysM domain L1 being superior to L2 and L3. It is not clear what properties distinguish the first LysM domain from the subsequent LysM domains, other than its location adjacent to the signal sequence domain.

We therefore sought to investigate the effect of the order as well as the number of LysM domains on the functionality of the LytE enzyme.

This was achieved by expressing variant LytE enzymes with two or three LysM domains in a permutated order from either the *cwlO* or the *lytE* promoter [Supplementary Figure 2]. Functionality of these enzymes was then tested by disrupting the endogenous *lytE* or *cwlO* gene. We could show that any LytE variant with two LysM domains in any of the nine possible combinations resulted in a functional enzyme. The same was true for two variant LytE with three LysM domains in the combinations L3-L3-L1 and L2-L1-L3.

Although all strains were viable, we observed that strains expressing variant LytE proteins with three LysM domains produced bigger colonies than those that expressed variant LytE proteins with two LysM domains [Figure 5, a]. Strains containing a variant

585 LytE protein with a single LysM domain displayed the smallest colonies, which also
586 exhibited strong lysis over the course of the experiment. In general, colonies were
587 bigger and experienced less lysis when they were expressed using the stronger *cwlO*
588 promoter rather than the weaker *lytE* promoter. Interestingly the order of LysM domains
589 also mattered. Those strains that expressed a variant LytE with L1 at the first (N-
590 terminal) position displayed larger colonies, than those in which either L2 or L3 were
591 located adjacent to the signal sequence domain. This correlation was especially strong
592 for LytE variants with a set of 2 LysM domains, as can be seen when comparing L1-L3
593 to L3-L1.
594 To address whether the type of variant LytE protein has an effect on the cellular protein
595 level, the level of HA-tagged variant LytE protein was determined by western blot
596 analysis. All variant LytE-HA proteins could be detected and bands migrated according
597 to the predicted MW [Figure 5, b]. The level of variant LytE protein varied significantly in
598 the strains tested. Since the number of LysM domains is proposed to correlate with
599 binding affinity, we expected to find more cell-associated protein in variant proteins with
600 a greater number of LysM domains. Indeed, the level of cell-associated variant LytE
601 protein was higher for proteins with two or more LysM domains, than for proteins with a
602 single LysM domain. If it is assumed that the lower level of cell-associated LytE is
603 caused by lowered binding to the cell wall, and not lowered expression, then we predict
604 the level of variant LytE protein found in the medium to have the reverse profile of that
605 for cell-associated protein. To verify this, we compared the levels of cell-associated
606 variant LytE proteins with those in the culture supernatant. A strong inverse correlation
607 in the level of variant LytE was observed. Variant LytE proteins with one LysM domain

608 were predominantly found in the culture supernatant, while variant LytE proteins with
609 three LysM domains were predominantly found in association with the cell wall.
610 Moreover the distribution of variant LytE proteins with two LysM domains takes an
611 intermediate between that found for those with three and one LysM domains. We
612 confirmed these results by immunofluorescence microcopy, which demonstrated that HA
613 tagged variant LytE with more LysM domains gave a stronger fluorescence signal than
614 those with fewer LysM domains [data not shown]. Interestingly, the pattern of
615 localization along the cell cylinder was not affected.
616 We further tested binding affinity of variant LytE protein in vitro in an insoluble
617 peptidoglycan pull down assay. To investigate the effect of salt on the peptidoglycan
618 binding affinity we added increasing amounts of NaCl to the assay buffer (0.05 M - 1
619 M NaCl) [Figure 5, c]. At a low salt concentration of 0.05 M NaCl, all variant LytE
620 proteins showed an equally strong binding capacity of approximately 100 % (proportion
621 of protein found in the pellet). However, no binding could be detected for the variant
622 LytE protein without LysM domain. Furthermore our results clearly show that the
623 peptidoglycan-binding capacity is negatively correlated to the salt concentration. With an
624 increase in salt concentration (≥ 0.1 M NaCl) binding capacity for the variant LytE
625 proteins is decreased. The rate at which binding capacity is declined varies between
626 proteins and is in negative correlation to the number of LysM domains present. Proteins
627 with more LysM domains exhibited a greater robustness of binding affinity than proteins
628 with a single LysM domain.
629 In summary, this study demonstrates that the number and order of LysM domains are of
630 great relevance for the functionality of the LytE enzyme. Although a LytE enzyme with

631 any one of its three LysM domains is functional and can fulfil the D,L-endopeptidase
632 requirement, increasing the number of LysM domains to two or three has a positive
633 effect on normal colony morphology. The order in which the domains are arranged is
634 also of importance in that variant LytE proteins in which the first LysM domain (L1) is
635 localized in its natural position adjacent to the signal peptide fulfil the D,L-endopeptidase
636 requirement to a greater extent than L2 and L3.
637 Furthermore, the number of LysM domains of a LytE variant protein is positively
638 correlated with peptidoglycan binding affinity and a good determinant of the extent to
639 which the protein is cell-associated or shed into the culture media. We could show that
640 an increase of the number of LysM domains leads to increased binding affinity both *in*
641 *vivo* and *in vitro*. We further demonstrated that binding affinity is negatively affected by
642 salt *in vitro*.

DISCUSSION:

The WalRK and SigIRsgI systems work in tandem to ensure adequate expression of D,L-endopeptidase activity

Expression of D,L-endopeptidase activity is required for normal growth and viability of *B. subtilis* as reported by Bisicchia et al. (2007). Under normal growth condition this activity is provided by CwlO and under stress conditions by the LytE enzyme. We demonstrated in this study that the level of *lytE* transcript and LytE protein is significantly increased when the *cwlO* gene is deleted, probably to compensate for the loss in D,L-endopeptidase activity. Since *lytE* is part of the SigI regulon (Tseng *et al.*, 2011), we suggest that this is achieved through activation of the alternative SigI factor as proposed by Salzberg and colleagues (2013). A similar response was observed by Kasahara and colleagues (2016) who reported that *lytE* transcription was considerably increased in strains with defects in teichoic acid synthesis, notably in a *ltaS* null mutant and who suggest that upregulation is achieved in a SigI dependent manner. Thus, the WalRK and SigIRsgI regulators form a regulatory network that functions to ensure that there is an adequate level of D,L-endopeptidase enzyme present under normal growth (CwlO) and cell wall stress (LytE) conditions.

It was of interest to establish whether LytE can fulfil the D,L-endopeptidase requirement for viability when expressed from the *cwlO* promoter in the absence of the endogenous *lytE* gene (thus removing SigI coupling). It was found that such a strain is viable, signifying that changing the regulation of LytE expression from its natural promoter to the *cwlO* promoter is compatible with viability and indeed normal growth. We observed

665 that LytE expression was higher when driven by the *cwlO* promoter as compared to the
666 endogenous *lytE* promoter (un-induced or induced). LytE expression (stabilized
667 transcript and increased protein level) was further increased by deletion of the *cwlO*
668 leader sequence due to the destabilizing effect of the leader transcript (Noone *et al.*,
669 2014). Surprisingly, we detected a slight increase in *mreBH* and *sigI* when *lytE* was
670 expressed from the *cwlO* promoter with the *cwlO* leader sequence. Based on the
671 premise that *sigI* is released under stress, removal of the *cwlO* leader sequence is
672 beneficial for the cell, since it results in a decrease of *sigI* levels. However, at the same
673 time removal of the *cwlO* leader sequence caused minor deficiencies in cell morphology,
674 indicating that such an elevated level of autolytic activity is detrimental to the cell. The
675 time point at which these phenotypic defects were observed, correlates with the high
676 level of WalRK activity during these growth phases (Prunty *et al.*, 2018)
677 In summary, these data suggest that although expression of LytE from the *cwlO*
678 promoter can fulfil the D,L-endopeptidase requirement for cell viability, it does not
679 complement CwlO function completely and that further elevation of LytE protein levels
680 are required (i.e. by removal of the leader sequence) to achieve a state without a cell
681 wall stress (i.e. no elevation of *sigI* expression).

682 **Expression levels and localization of D,L-endopeptidases within the cell envelope**
683 **are critical**

684 We investigated four of the seven D,L-endopeptidases encoded in the *B. subtilis*
685 genome (LytE, LytF, CwlS, YqgT) and found that CwlS and YqgT cannot complement
686 CwlO / LytE activity when expressed from the *cwlO* promoter. LytE naturally fulfils the

687 D,L-endopeptidase requirement in a *cwlO* mutant, but can also do so when expressed
688 from the *cwlO* locus. LytF cannot complement CwlO / LytE activity when expressed from
689 its natural *lytF* promoter, but gives partial complementation when expressed from the
690 *cwlO* promoter. The fact that LytF is unable to fulfil the D,L-endopeptidase requirement
691 for viability when expressed from the *lytF* promoter, but can do so when expressed from
692 the *cwlO* promoter, indicates that complementation requires either greater than normal
693 LytF protein levels and / or LytF expression at a different time point throughout the
694 growth cycle. However, although a viable strain could be constructed, severe growth
695 and morphology defects were observed. Exacerbation of the aberrant phenotype in
696 strains without *cwlO* leader sequence (causing a further increase in LytF levels)
697 suggests that the level of LytF in this background may be too high for normal growth.
698 We do not know why neither *cwlS* nor *yqgT* were able to fulfil the D,L-endopeptidase
699 requirement. However, we ruled out the possibility of a synthetic lethality between these
700 enzymes by constructing viable double mutants.
701 The fact that the LytE, LytF and CwlS enzymes are highly similar in their C-terminal D,L-
702 endopeptidase catalytic domains, as well as in their N-terminally located LysM domains
703 poses the question as to what causes the inability of CwlS to fulfil the D,L-
704 endopeptidase requirement. We confirmed that there is indeed a high level of CwlS
705 expressed and associated with the cell, when expressed from the *cwlO* promoter [data
706 not shown]. However, it may be of relevance that the CwlS protein is primarily located at
707 the cell septa (Hashimoto *et al.*, 2012), while both LytE and LytF have some presence
708 along the cell cylinder (Hashimoto *et al.*, 2012; Kiriyama *et al.*, 2014; Kasahara *et al.*,
709 2016). This is significant since it has been established that the D,L-endopeptidase

710 requirement is associated with facilitating cell cylinder elongation (Bisicchia *et al.*, 2007).
711 Thus, there may not be a sufficient level of CwlS located along the cylinder to facilitate
712 cell elongation. This is supported by the observations that cells expressing *cwlS* from
713 the *cwlO* promoter appear to be shorter than wild-type cells and to have unusual
714 structures at the poles of the cell (data not shown).
715 Although both LytE and LytF show some binding to the lateral cell wall, LytF is primarily
716 bound to the cell septa and poles, which potentially explains why LytF achieves only
717 partial complementation in comparison to the LytE enzyme. In summary, these results
718 show that beside the level of D,L-endopeptidase expression, their proper subcellular
719 localization are paramount for normal growth.

720 **The individual LysM domains of LytE are functionally and structurally distinct**

721 We could show that strains expressing a LytE variants with any single LysM domain are
722 viable. However, we observed cell morphology, colony morphology and growth defects
723 in such strains. These phenotypic defects were more pronounced in LytE variants with a
724 single copy of the second (L2) or third (L3) LysM domains as opposed to the N-terminal
725 located L1, indicating that the individual LysM domains are not functionally equivalent.
726 The increased effectiveness of the first LysM domain (L1) is probably due to the fact that
727 it has evolved to be more suitable for interaction with the adjacent signal sequence
728 domain, which perhaps causes the enzyme to be expressed or exported more
729 efficiently.
730 This finding was confirmed in later experiments when we could show that in strains that
731 express a variant LytE proteins with two or three LysM domains, the order of the

732 domains was relevant for proper colony growth. Variant LytE proteins that contained the
733 first LysM domain (L1) at the N-terminal position resulted in strains that displayed
734 healthier colonies which experienced less lysis than variant LytE in which either L2 or L3
735 were located adjacent to the signal peptide. A possible mechanism by which a LytE
736 protein with impaired functionality could cause colony growth defects was investigated
737 by Mamou and colleagues (2017), who report that LytE is involved in proper colony
738 development through influencing Y arm extension and colony thickening patterns.
739 The fact that the first LysM domain is structurally different from the subsequent domains
740 was further supported through bioinformatic analysis. By comparing the amino acid
741 sequence of the 185 LysM domains contained in the LytE, CwlS and LytF orthologs of
742 19 closely related *B. subtilis* species we found that the N-terminal located LysM domains
743 were the most distinct. This may indicate that due its interaction with the signal peptide
744 there is a greater selection pressure on the N-terminal LysM domain than on
745 subsequent domains.

746 **LysM domains contribute to peptidoglycan binding in vivo and in vitro**

747 We could show that a variant LytE without LysM domains exhibited no binding affinity to
748 purified cell wall *in vitro* and that it could not sustain cell growth. Thus, LysM domains
749 are essential for cell wall binding and therefore functionality of the LytE enzyme. This
750 observation is supported by several previous studies that have shown that a total loss of
751 LysM domains greatly decreases the peptidoglycan-binding capacity in vitro (Steen *et*
752 *al.*, 2005; Eckert *et al.*, 2006; Wong *et al.*, 2014). Increasing the number of LysM
753 domains from one to three, led to increased binding capacity both *in vivo* and *in vitro*,

suggesting that individual LysM domains contribute to binding in a accumulative manner, as was proposed previously (Mesnage *et al.*, 2014; Wong *et al.*, 2014). Surprisingly, strains expressing a LytE variant with any single LysM domain are viable, suggesting that the peptidoglycan-binding affinity of the wild type LytE protein with three LysM domains is excessive. However, our experiments were performed under laboratory conditions and we observed that such strains have elevated stress sensitivity, especially to salt stress. We demonstrated *in vitro* that salt causes a decrease in binding capacity of variant LytE proteins, especially those with fewer than three LysM domains. Therefore, growth defects in the presence of salt are probably due to insufficient binding of the variant LytE D,L-endopeptidase under this circumstances. Given that *B. subtilis* is a soil-dwelling organism, its capability to withstand salt stress is essential for its survival. Therefore, cells that possess a LytE enzyme with less than three LysM domains would not be able to survive natural fluctuations in the salt concentrations in the soil.

Expansion of the LysM domains and evolution of the CwlS, LytE and LytF orthologs

We observed that the LysM domains of the CwlS, LytE and LytF orthologs cluster primarily by position, with the N-terminally located LysM domains forming a remarkably distinct cluster. This finding is contrary to a similar study conducted for fungal proteins. In this study, Cen and colleagues (2017) report that the LysM domains of fungal effectors proteins cluster independently of their sequential position within the parental protein or protein structure family. We suggest that this difference stems from the fact that the LytE, CwlS and LytF proteins have evolved fairly recent, leaving their LysM

domains little time to diverge.

We were further interested in the manner in which D,L-endopetidases acquire new LysM domains. By screening 26.078 bacterial proteins for identical LysM domain amino acid sequences we could show that LysM domains are primarily propagated through duplication events that occur within a gene. Interestingly, the majority of duplicated LysM domains were present within the phylum Firmicutes (79.4 %). This is a much higher percentage than could be expected, provided that domain expansion occurred in all phyla at the same rate (LysM domains of the Firmicutes phylum represent 32.8 % of the total dataset), suggesting that LysM domain expansion is either a more recent or more frequent event in species of the Firmicutes phylum. Interestingly, an uncharacterised protein of *Bacillus subtilis* subsp. *natto* BEST195 (D4G4L0) was found to possess 6 LysM domains, that are of identical amino acid sequence and highly similar (although not identical) on nucleotide level. This indicates that LysM domain duplication events can create several repeats of a single domain at once, leading to the rapid expansion of the domain repeats.

Based on these findings, we propose a model for the evolutionary development of the three D,L-endopeptidases and their LysM domains. Our model is based on the premise that LytE, CwlS and LytF share a common ancestor and that LysM domain repeats expand independently within genes and are passed on through stable vertical inheritance. In this model an ancestral gene with a single N-terminally located LysM domain, undergoes a combination of gene and LysM domain duplication / loss events, which lead to the development of the three orthologs and their observed LysM domain combinations [Figure 6, a]. This model can also be used to infer basic phylogeny

between the investigated *Bacillus* species [Figure 6, b], which is in line with their overall phylogenetic relationships [Supplementary Figure 1].

Conclusion

Overall this study offers further insights into the D,L-endopeptidase requirement of *B. subtilis*. To our knowledge this is the first work that has demonstrated that the D,L-endopeptidase requirement of *B. subtilis* can be fulfilled by a D,L-endopeptidase other than LytE or CwlO. It remains unknown why LytF, but not CwlS is able to complement for CwlO, given their highly similar structure (also localization is a likely suspect).

This is also the first time that is has been investigated whether individual LysM domains are of equal importance for the functionality of an autolytic enzyme in vivo. It was shown that any single LysM domain is sufficient for functionality of the LytE enzyme, however robust growth requires the three LysM domains present in the wild type enzyme. Among these we demonstrated that the first LysM domain is superior over the subsequent domains, especially when located N-terminally, which is probably due to some interaction with the signaling peptide.

An investigation of bacterial LysM domains on a large scale led to some interesting observations, such as that expansion of LysM domains occurs frequently in the *Firmicutes* phylum and that LysM domains are unlikely to be transferred horizontally. Accordingly, a putative model was developed describing the relationship between the LytE, CwlS and LytF D,L-endopeptidases.

It is clear from our results that D,L-endopeptidases must have many features to be able to fulfill the requirement for viability. Among those requirements are the level of enzyme

expressed, the stability of the protein when expressed and the localization of the protein within the cell envelope.

Further investigating these and other aspects of autolytic enzymes and cell wall metabolism will not only deepen our understanding of these processes, but also has the potential to be utilized for the development of new antibiotics and for the optimization of bacterial strains for enzyme and protein production. Furthermore an in-depth understanding of LysM domains and conditions that affect their biding affinity is likely to be of importance for the development of future binding assays and cell surface displays.

References

Anagnostopoulus, C., and Spizizen, J. (1960). Requirements for Transformation in Bacillus subtilis. 81, 741–746. Available at: C:\Users\User\Documents\Citavi 6\Projects\LytE\Citavi Attachments\Anagnostopoulus, Spizizen 1960 - Requirements for Transformation in Bacillus.pdf.

Bisicchia, P., Noone, D., Lioliou, E., Howell, A., Quigley, S., Jensen, T., et al. (2007). The essential YycFG two-component system controls cell wall metabolism in Bacillus subtilis. *Mol. Microbiol.* 65, 180–200. doi:10.1111/j.1365-2958.2007.05782.x.

Bloor, A. E., and Cranenburgh, R. M. (2006). An Efficient Method of Selectable Marker Gene Excision by Xer Recombination for Gene Replacement in Bacterial Chromosomes. *Appl. Environ. Microbiol.* 72, 2520–2525. Available at: C:\Users\

User\Documents\Citavi 6\Projects\LytE\Citavi Attachments\Bloor and Cranenburgh 2006.pdf.

Bodenhofer, U., Bonatesta, E., Horejs-Kainrath, C., and Hochreiter, S. (2015). msa: an R package for multiple sequence alignment. *Bioinformatics* 31, 3997–3999. doi:10.1093/bioinformatics/btv494.

Botella, E., Hübner, S., Hokamp, K., Hansen, A., Bisicchia, P., Noone, D., et al. (2011). Cell envelope gene expression in phosphate-limited Bacillus subtilis cells. *Microbiology* 157, 2470–2484. doi:10.1099/mic.0.049205-0.

Boylan, S. A., Redfield, A. R., Brody, M. S., and Price, C. W. (1993). Stress-induced activation of the sigma B transcription factor of Bacillus subtilis. *J. Bacteriol.* 175, 7931–7937. Available at: C:\Users\User\Documents\Citavi 6\Projects\LytE\Citavi Attachments\Boylan, Redfield et al 1993 - Stress-induced activation of the sigma.pdf.

Buist, G., Steen, A., Kok, J., and Kuipers, O. P. (2008). LysM, a widely distributed protein motif for binding to (peptido)glycans. *Mol. Microbiol.* 68, 838–847. Available at: C:\Users\User\Documents\Citavi 6\Projects\LytE\Citavi Attachments\Buist 2008.pdf.

Cen, K., Li, B., Lu, Y., Zhang, S., and Wang, C. (2017). Divergent LysM effectors contribute to the virulence of Beauveria bassiana by evasion of insect immune defenses. *PLOS Pathog.* 13, e1006604. doi:10.1371/journal.ppat.1006604.

Desvaux, M., Candela, T., and Serror, P. (2018). Surfaceome and Proteosurfaceome in Parietal Monoderm Bacteria: Focus on Protein Cell-Surface Display. *Front. Microbiol.* 9. doi:10.3389/fmicb.2018.00100.

Dominguez-Cuevas, P., Porcelli, I., Daniel, R., and Errington, J. (2013). Differentiated roles for MreB-actin isologues and autolytic enzymes in bacterial morphogenesis. *Mol. Microbiol.* 89, 1084–1098. Available at: C:\Users\User\Documents\Citavi 6\Projects\LytE\Citavi Attachments\Dominguez-Cuevas, Porcelli et al 2013 - Differentiated roles for MreB-actin isologues.pdf.

Eckert, C., Lecerf, M., Dubost, L., Arthur, M., and Mesnage, S. (2006). Functional analysis of AtlA, the major N-acetylglucosaminidase of Enterococcus faecalis. *J. Bacteriol.* 188, 8513–8519. Available at: C:\Users\User\Documents\Citavi 6\Projects\LytE\Citavi Attachments\Eckert, Lecerf et al 2006 - Functional analysis of AtlA.pdf.

Hashimoto, M., Ooiwa, S., and Sekiguchi, J. (2012). Synthetic Lethality of the lytE cwlO Genotype in Bacillus subtilis Is Caused by Lack of D,L-Endopeptidase Activity at the Lateral Cell Wall. *J. Bacteriol.* 194, 796–803. doi:10.1128/JB.05569-11.

Hayhurst, E. J., Kailas, L., Hobbs, J. K., and Foster, S. J. (2008). Cell wall peptidoglycan architecture in Bacillus subtilis. *Proc Natl Acad Sci* 105, 14603–14608. Available at: C:\Users\User\Documents\Citavi 6\Projects\LytE\Citavi Attachments\Hayhurst et al 2008.pdf.

Horton, R. M., Hunt, H. D., Ho, S. N., Pullen, J. K., and Pease, L. R. (1989). Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene* 77, 61–68.

Howell, A., Dubrac, S., Noone, D., Varughese, K. I., and Devine, K. M. (2006). Interactions between the YycFG and PhoPR two-component systems in Bacillus subtilis: the PhoR kinase phosphorylates the non-cognate YycF response regulator upon phosphate limitation. *Mol. Microbiol.* 59, 1199–1215. doi:10.1111/j.1365-2958.2005.05017.x.

Ishikawa, S., Hara, Y., Ohnishi, R., and Sekiguchi, J. (1998). Regulation of a New Cell Wall Hydrolase Gene, cwlF, which affects Cell Separation in Bacillus subtilis. *J. Bacteriol.* 1890, 2549–2555. Available at: C:\Users\User\Documents\Citavi 6\Projects\LytE\Citavi Attachments\Ishikawa 1998.pdf.

Kasahara, J., Kiriyama, Y., Miyashita, M., Kondo, T., Yamada, T., Yazawa, K., et al. (2016). Teichoic Acid Polymers Affect Expression and Localization of dl-Endopeptidase LytE Required for Lateral Cell Wall Hydrolysis in Bacillus subtilis. *J. Bacteriol.* 198, 1585–1594. doi:10.1128/JB.00003-16.

Kiriyama, Y., Yazawa, K., Tanaka, T., Yoshikawa, R., Yamane, H., Hashimoto, M., et al. (2014). Localization and expression of the Bacillus subtilis dl-endopeptidase LytF are influenced by mutations in LTA synthases and glycolipid anchor synthetic enzymes. *Microbiology* 160, 2639–2649. doi:10.1099/mic.0.080366-0.

Mamou, G., Fiyaksel, O., Sinai, L., and Ben-Yehuda, S. (2017). Deficiency in Lipoteichoic Acid Synthesis Causes a Failure in Executing the Colony Developmental Program in Bacillus subtilis. *Front. Microbiol.* 8. doi:10.3389/fmicb.2017.01991.

Meisner, J., Llopis, P. M., Sham, L.-T., Garner, E., Bernhardt, T. G., and Rudner, D. Z. (2013). FtsEX is required for CwlO peptidoglycan hydrolase activity during cell wall elongation in Bacillus subtilis. *Mol. Microbiol.* 89, 1069–1083. Available at: C:\Users\User\Documents\Citavi 6\Projects\LytE\Citavi Attachments\Jeffrey 2013.pdf.

Mesnage, S., Dellarole, M., Baxter, N. J., Rouget, J.-B., Dimitrov, J. D., Wang, N., et al. (2014). Molecular basis for bacterial peptidoglycan recognition by LysM domains. *Nat. Commun.* 5, 1–11. doi:10.1038/ncomms5269.

Noone, D., Botella, E., Butler, C., Hansen, A., Jende, I., and Devine, K. M. (2012). Signal Perception by the Secretion Stress-Responsive CssRS Two-Component System in Bacillus subtilis. *J. Bacteriol.* 194, 1800–1814. Available at: C:\Users\User\Documents\Citavi 6\Projects\LytE\Citavi Attachments\Noone 2012.pdf.

Noone, D., Salzberg, L. I., Botella, E., Basell, K., Becher, D., Antelmann, H., et al. (2014). A Highly Unstable Transcript Makes CwlO D,L-Endopeptidase Expression Responsive to Growth Conditions in Bacillus subtilis. *J. Bacteriol.* 196, 237–247. Available at: C:\Users\User\Documents\Citavi 6\Projects\LytE\Citavi Attachments\Noone 2014.pdf.

Pele, J., Becu, J.-M., Abdi, H., and Chabbert, M. (2012). Bios2mds: an R package for comparing orthologous protein families by metric multidimensional scaling. *BMC Bioinformatics* 13, 1–7. doi:10.1186/1471-2105-13-133.

Prunty, M. P., Noone, D., and Devine, K. M. (2018). The distinct PhoPR mediated responses to phosphate limitation in *Bacillus subtilis* subspecies *subtilis* and *spizizenii* stem from differences in wall teichoic acid composition and metabolism: PhoPR activation in *Bacillus subtilis* subspecies. *Mol. Microbiol.* 109, 23–40. doi:10.1111/mmi.13965.

Rajagopal, M., and Walker, S. (2017). Envelope Structures of Gram-Positive Bacteria. *Curr. Top. Microbiol. Immunol.* 404, 1–44. doi:10.1007/82_2015_5021.

Salamaga, B., Prajsnar, T. K., Jareño-Martinez, A., Willemse, J., Bewley, M. A., Chau, F., et al. (2017). Bacterial size matters: Multiple mechanisms controlling septum cleavage and diplococcus formation are critical for the virulence of the opportunistic pathogen Enterococcus faecalis. *PLoS Pathog.* 13. doi:10.1371/journal.ppat.1006526.

Salzberg, L. I., Powell, L., Hokamp, K., Botella, E., Noone, D., and Devine, K. M. (2013). The WalRK (YycFG) and σ I RsgI regulators cooperate to control CwlO and LytE expression in exponentially growing and stressed Bacillus subtilis cells. *Mol. Microbiol.* 87, 180–195. Available at: C:\Users\User\Documents\Citavi 6\Projects\LytE\Citavi Attachments\Salzberg, Powell et al 2013 - The WalRK YycFG and σ.pdf.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). *Molecular cloning*. 2. ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Schliep, K. P. (2011). phangorn: phylogenetic analysis in R. *Bioinforma. Oxf. Engl.* 27, 592–593. doi:10.1093/bioinformatics/btq706.

Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012). NIH Image to ImageJ. *Nat. Methods* 9, 671–675. doi:10.1038/nmeth.2089.

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 7, 1–6. doi:10.1038/msb.2011.75.

Sims, G. E., Jun, S.-R., Wu, G. A., and Kim, S.-H. (2009). Alignment-free genome comparison with feature frequency profiles (FFP) and optimal resolutions. *Proc. Natl. Acad. Sci. U. S. A.* 106, 2677–2682. doi:10.1073/pnas.0813249106.

Steen, A., Palumbo, E., Deghorain, M., Cocconcelli, P. S., Delcour, J., Kuipers, O. P., et al. (2005). Autolysis of Lactococcus lactis Is Increased upon d-Alanine Depletion of Peptidoglycan and Lipoteichoic Acids. *J. Bacteriol.* 187, 114–124. doi:10.1128/JB.187.1.114-124.2005.

Tseng, C.-L., Chen, J.-T., Lin, J.-H., Huang, W.-Z., and Shaw, G.-C. (2011). Genetic evidence for involvement of the alternative sigma factor SigI in controlling expression of the cell wall hydrolase gene lytE and contribution of LytE to heat survival of Bacillus subtilis. *Arch. Microbiol.* 193, 677–685. doi:10.1007/s00203-011-0710-0.

Tseng, C.-L., and Shaw, G.-C. (2008). Genetic Evidence for the Actin Homolog Gene mreBH and the Bacitracin Resistance Gene bcrC as Targets of the Alternative Sigma Factor SigI of Bacillus subtilis. *J. Bacteriol.* 190, 1561–1567. Available at: C:\Users\User\Documents\Citavi 6\Projects\LytE\Citavi Attachments\Tseng 2008.pdf.

Wang, A., and Ash, G. J. (2015). Whole Genome Phylogeny of *Bacillus* by Feature Frequency Profiles (FFP). *Sci. Rep.* 5, 13644. doi:10.1038/srep13644.

Wong, J. E. M. M., Alsarraf, H. M. A. B., Kaspersen, J. D., Pedersen, J. S., Stougaard, J., Thirup, S., et al. (2014). Cooperative binding of LysM domains determines the carbohydrate affinity of a bacterial endopeptidase protein. *FEBS J.* 281, 1196–1208. Available at: C:\Users\User\Documents\Citavi 6\Projects\LytE\Citavi Attachments\Wong, Jaslyn E M M, Alsarraf, Husam M A B et al 2013 - Cooperative binding of LysM domains.pdf.

Yamaguchi, H., Furuhata, K., Fukushima, T., Yamamoto, H., and Sekiguchi, J. (2004). Characterization of a New Bacillus subtilis Peptidoglycan Hydrolase Gene, yvcE (Named cwlO), and the Enzymatic Properties of Its Encoded Protein. *J. Biosci. Bioeng.* 98, 174–181. Available at: C:\Users\User\Documents\Citavi 6\Projects\LytE\Citavi Attachments\Yamaguchi 2004.pdf.

Yu, G., Smith, D. K., Zhu, H., Guan, Y., and Lam, T. T.-Y. (2016). ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol. Evol.* 8, 28–36. doi:10.1111/2041-210X.12628.

Figure 1:
phylogentic tree of the LysM motifs

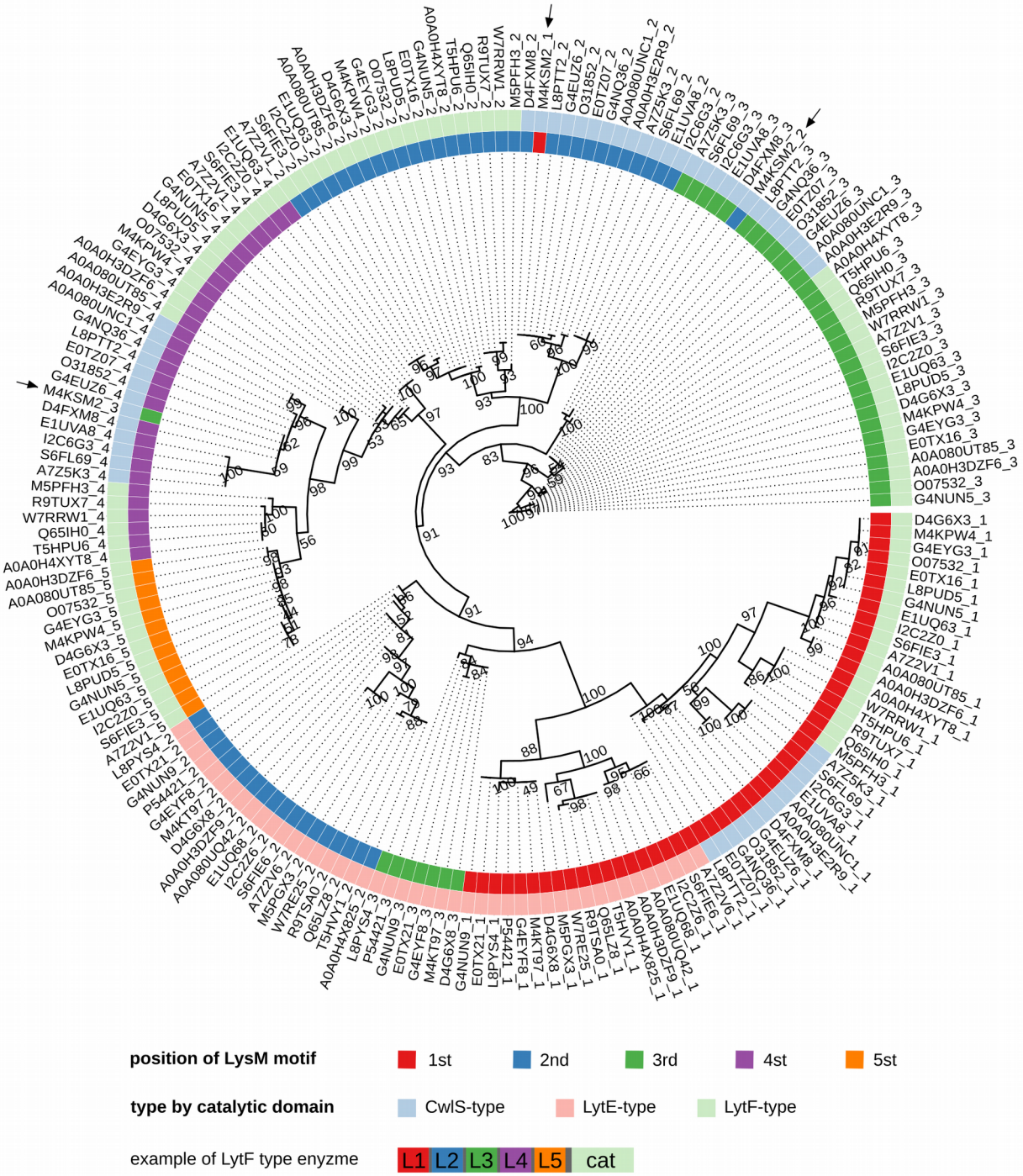
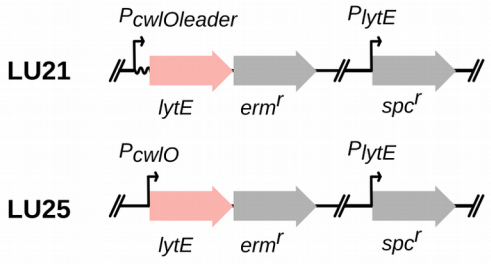
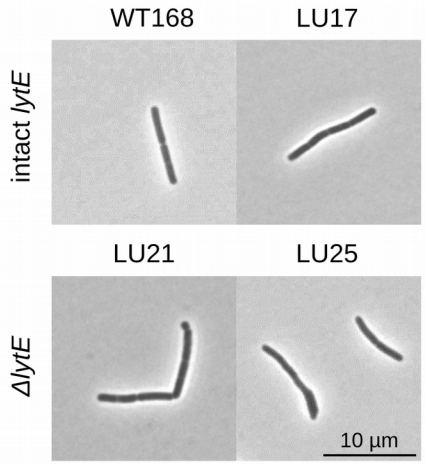


Figure 2:

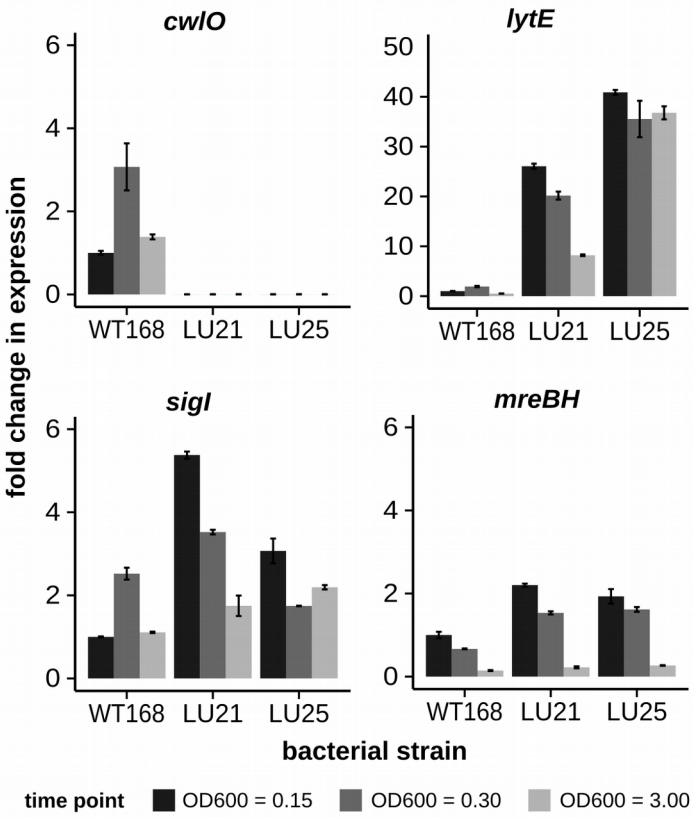
a) strain construction



b) microscopy



c) RT-qPCR



d) western blot

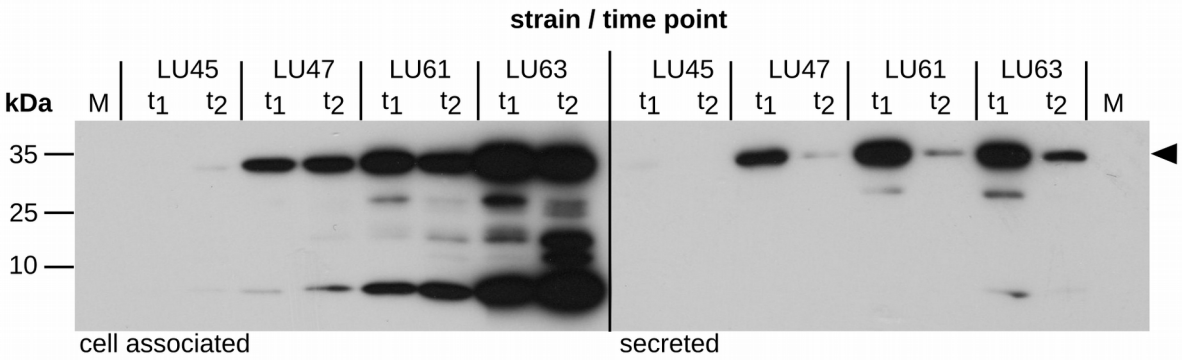
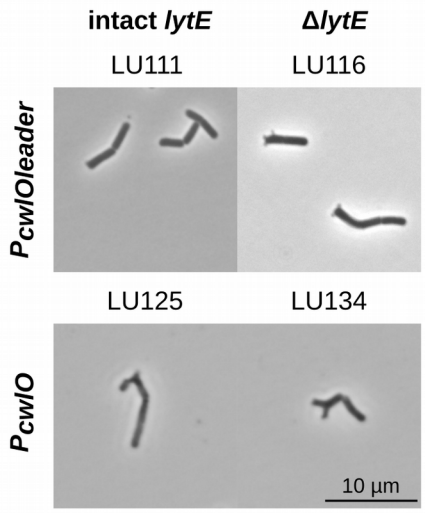
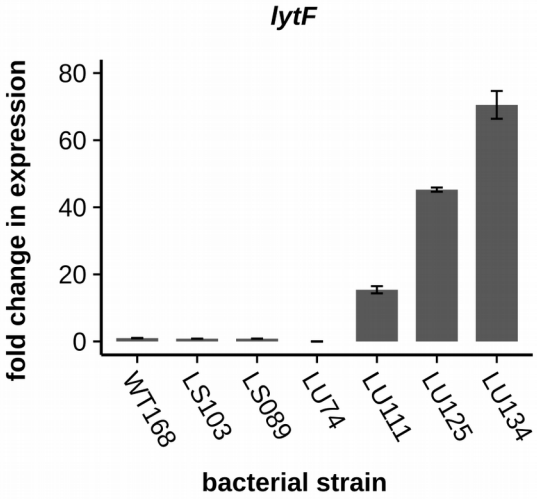


Figure 3:

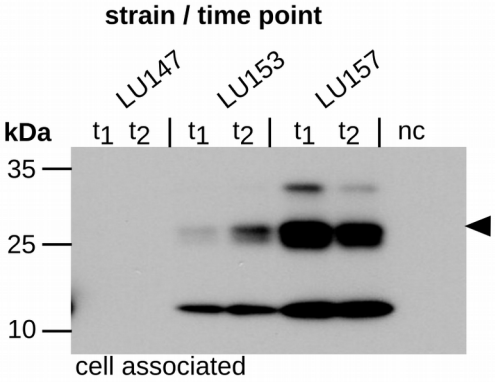
a) microscopy



b) RT-qPCR



c) western blot



d) fluorescence microscopy

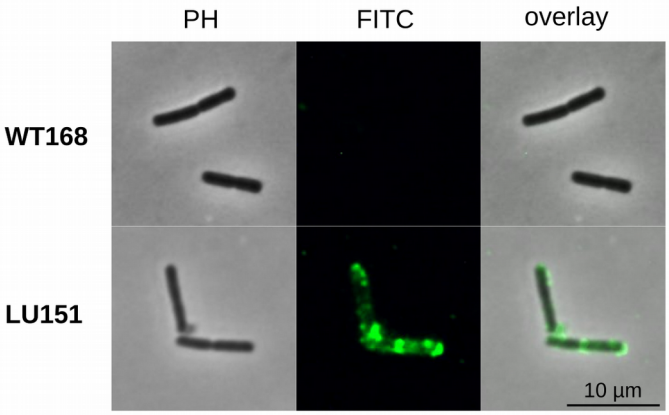
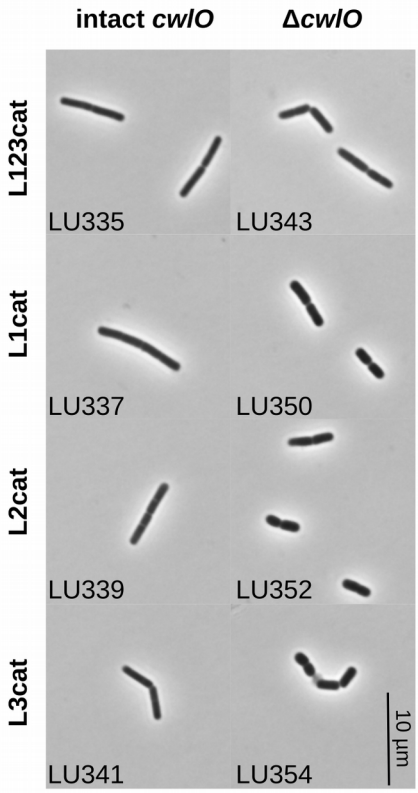
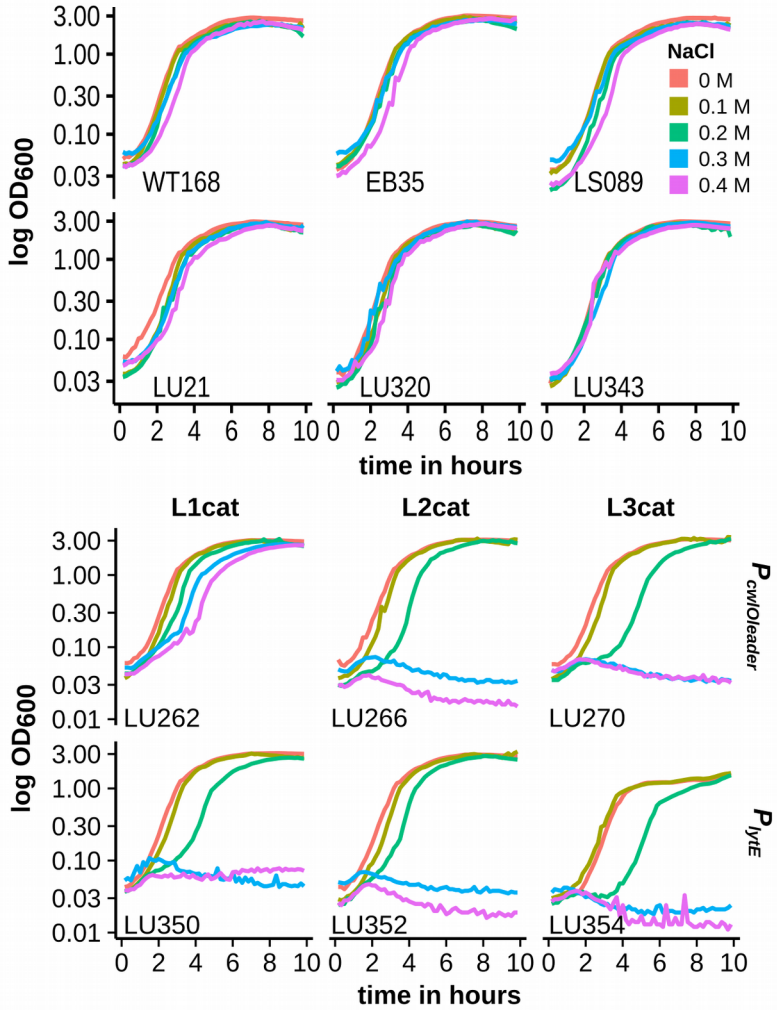


Figure 4:

a) microscopy



c) salt stress



b) RT-qPCR

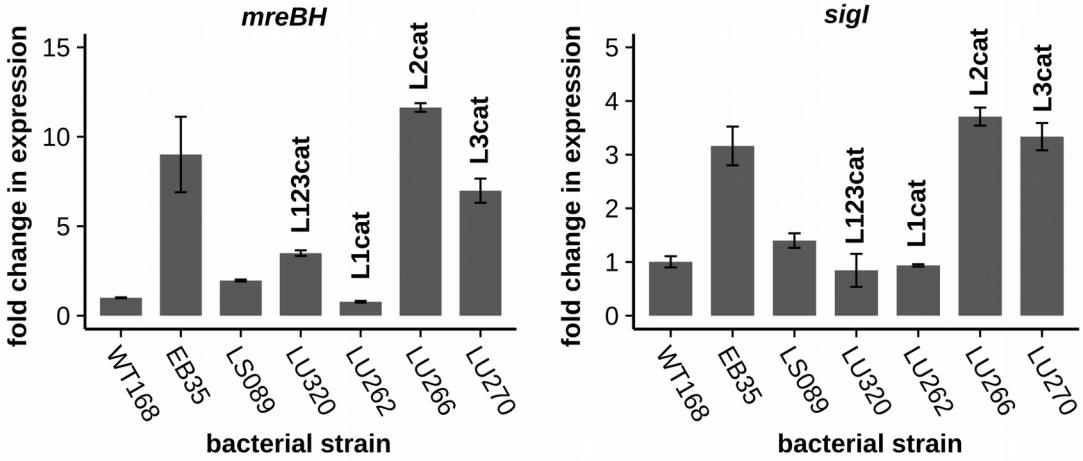
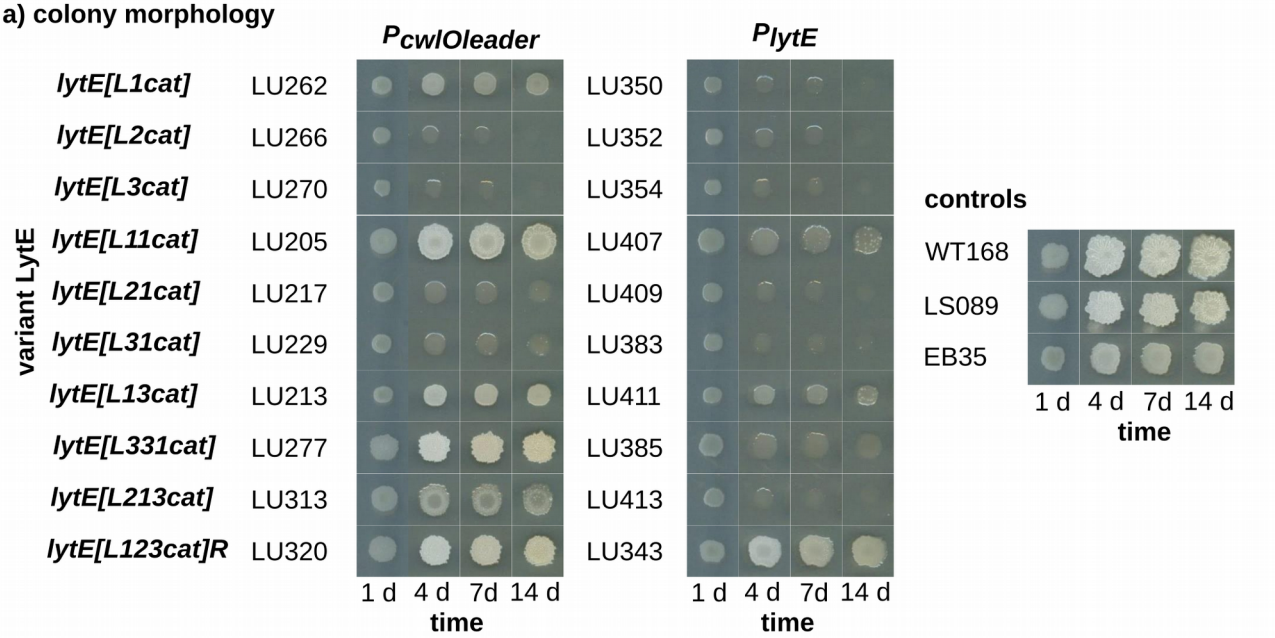
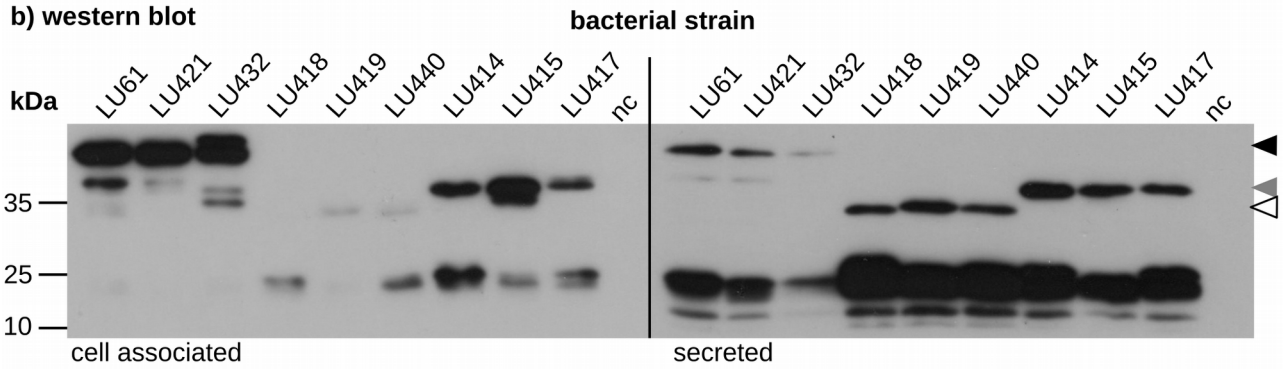


Figure 5:

a) colony morphology



b) western blot



c) pull-down assay

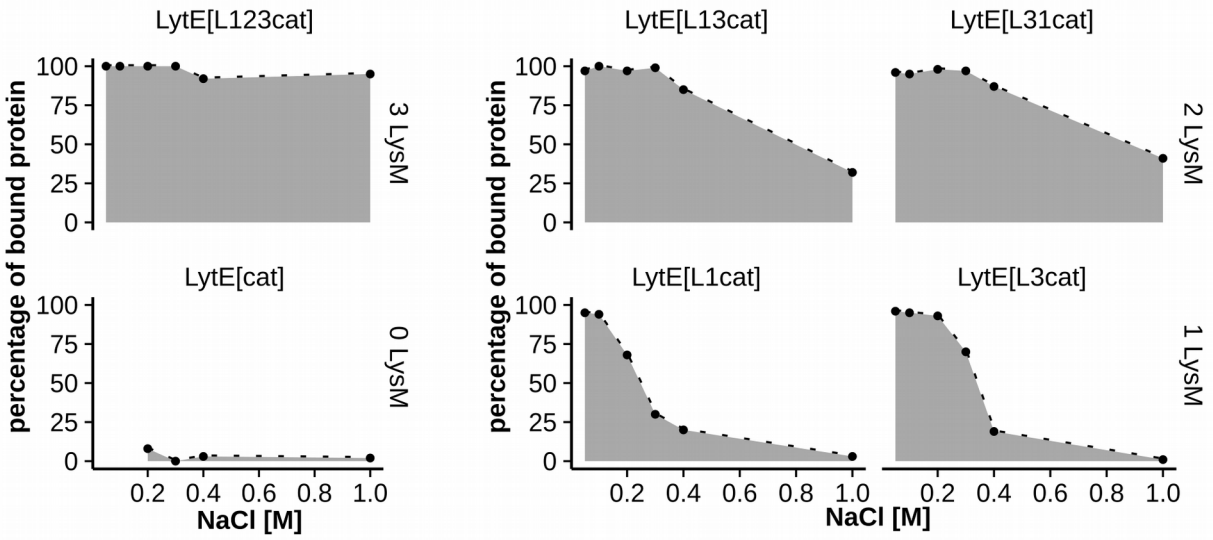
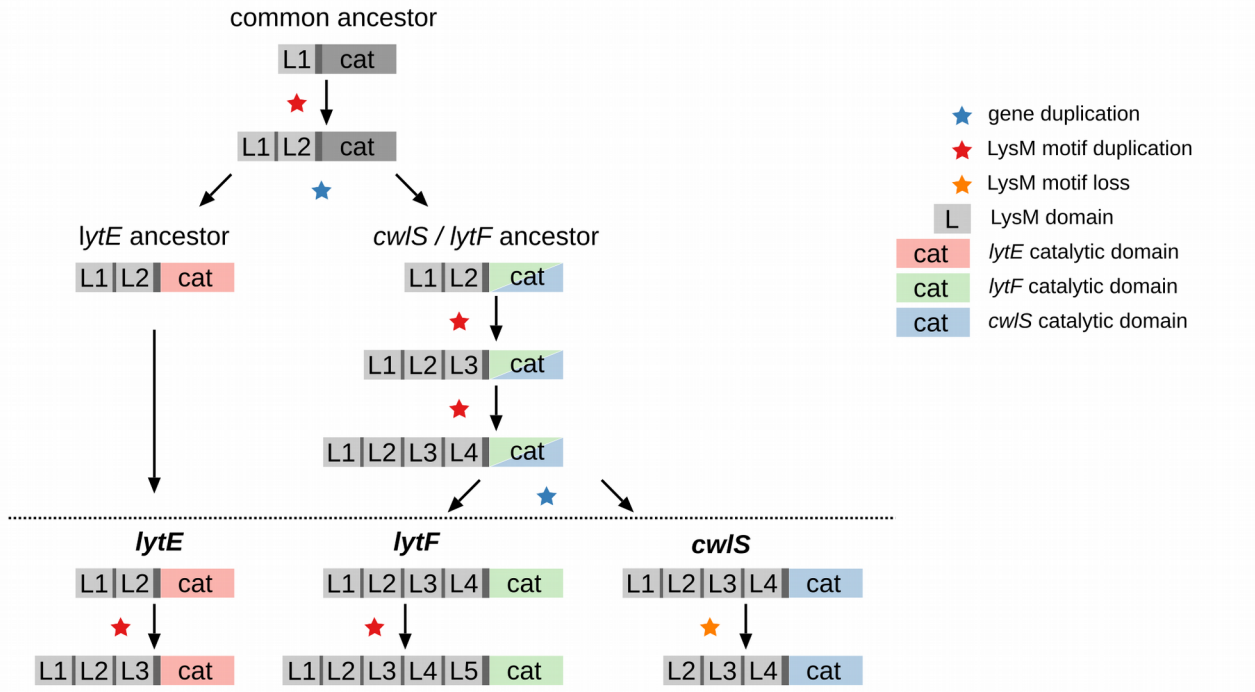
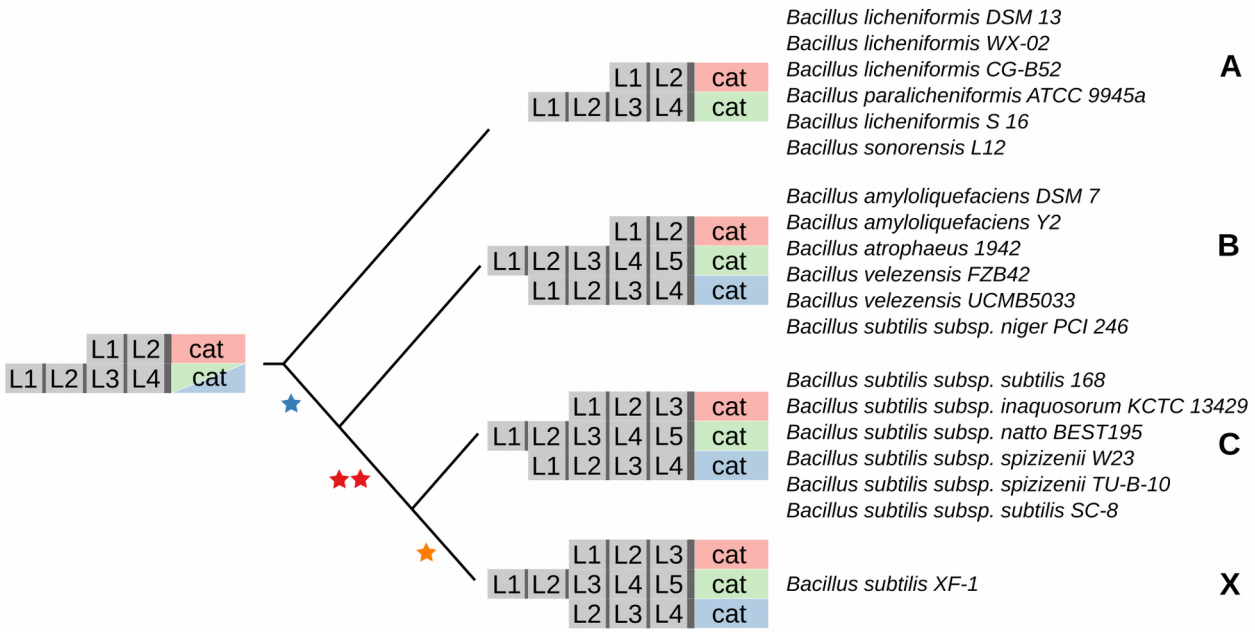


Figure 6:

a) putative model of the evolutionary development of *lytE*, *lytF* and *cwlS*

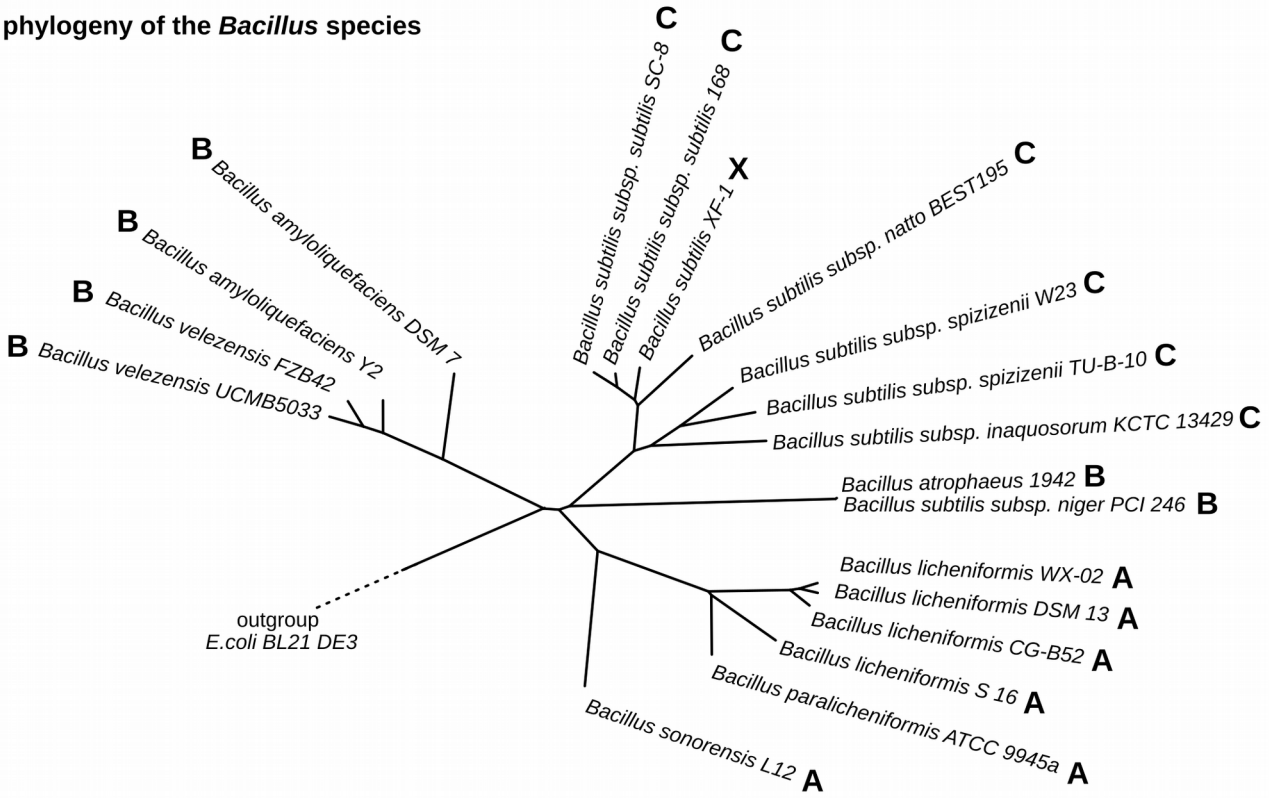


b) proposed model for the evolutionary development of groups with different sets of paralogs



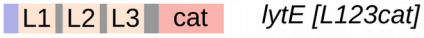
Supplementary Figure 1:

phylogeny of the *Bacillus* species

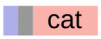


Supplementary Figure 2:

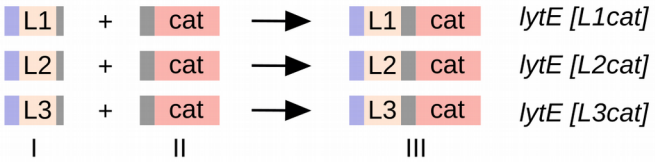
a) schematic structure of the wild type *lytE*



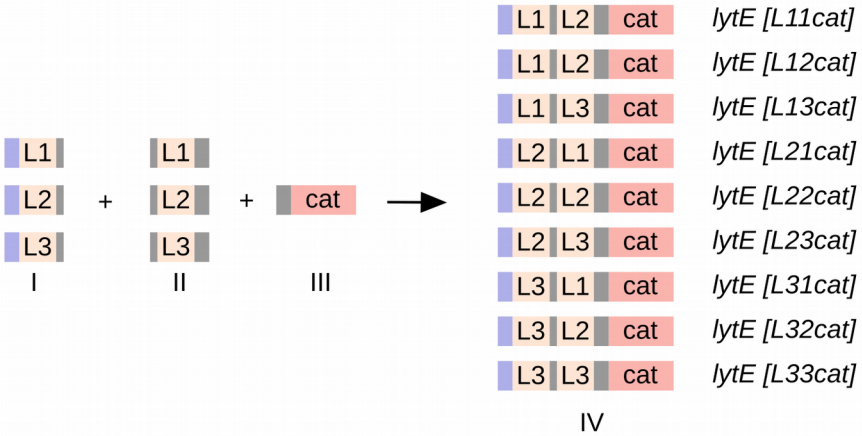
b) a variant *lytE* fragment with no LysM motif



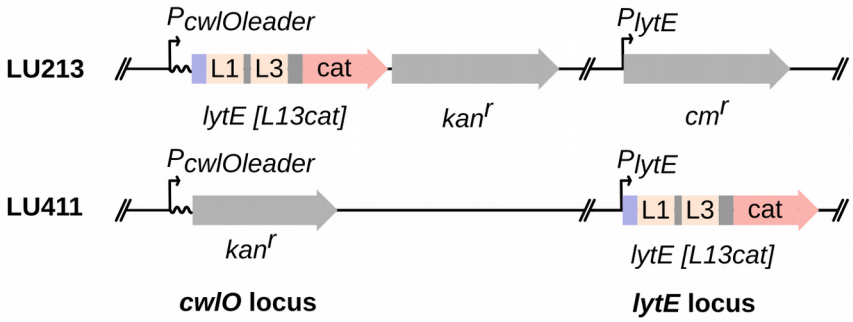
c) construction of variant lytE fragments with one LysM motif



d) construction of variant *lytE* fragments with two LysM motifs



e) integration of a variant *lytE* fragment on either the *cwlO* or the *lytE* locus



Suplementary Material

Table S1: List of the identified orthologs to LytE, CwlS and LytF with their UniProt identifier and NCBI taxonomic identifier

Organism	Taxonomic identifier	LytE ortholog	LytF ortholog	CwlS ortholog
Bacillus subtilis subsp. subtilis 168	224308	P54421	O07532	O31852
Bacillus subtilis subsp. inaquosorum KCTC 13429	1236548	L8PYS4	L8PUD5	L8PTT2
Bacillus subtilis subsp. natto BEST195	645657	D4G6X8	D4G6X3	D4FXM8
Bacillus subtilis subsp. spizizenii W23	655816	E0TX21	E0TX16	E0TZ07
Bacillus subtilis subsp. spizizenii TU-B-10	1052585	G4NUN9	G4NUN5	G4NQ36
Bacillus subtilis subsp. subtilis SC-8	1089443	G4EYF8	G4EYG3	G4EUZ6
Bacillus subtilis XF-1	1233100	M4KT97	M4KPW4	M4KSM2
Bacillus subtilis subsp. niger PCI 246	1495315	A0A080UQ42	A0A080UT85	A0A080UNC1
Bacillus amyloliquefaciens DSM 7	692420	E1UQ68	E1UQ63	E1UVA8
Bacillus amyloliquefaciens Y2	1126211	I2C2Z6	I2C2Z0	I2C6G3
Bacillus velezensis FZB42 / *Bacillus amyloliquefaciens subsp. Plantarum FZB42*	326423	A7Z2V6	A7Z2V1	A7Z5K3
Bacillus velezensis UCMB5033 / *Bacillus amyloliquefaciens subsp. Plantarum UCMB5033*	1338518	S6FIE6	S6FIE3	S6FL69
Bacillus atrophaeus 1942	720555	A0A0H3DZF9	A0A0H3DZF6	A0A0H3E2R9
Bacillus licheniformis DSM 13	279010	Q65LZ8	Q65IH0	none
Bacillus licheniformis CG-B52	1368424	T5HVY1	T5HPU6	none
Bacillus licheniformis S 16	1348578	W7RE25	W7RRW1	none
Bacillus licheniformis WX-02	1126218	A0A0H4X825	A0A0H4XYT8	none
Bacillus paralicheniformis ATCC 9945a	766760	R9TSA0	R9TUX7	none
Bacillus sonorensis L12	1274524	M5PGX3	M5PFH3	none

Table S2: List of strains used in this study

Strain	Genotype	Source
BL21(DE3)	*F ompT hsdSB* (r_B m_B) *gal dcm (DE3)*	Studier and Moffatt, 1986
TG1	*supE hsdΔ5 thi Δ(laC-proAB) F´[traD36 proAB*$^+$ *lacI*^{q} *lacZΔM15]*	Sambrook et al., 1989
TP611	*recBC hsdRM cya610 pcn*	Glaser et al., 1993
EB35	*trpC2 ΔcwlO::kan*^{r}	unpublished, Eric Botella
DN3000	*trpC2 ΔcwlOleader*	Noone et al., 2014
LS089	*trpC2 ΔlytE::cm*^{r}	Salzberg et al. (2013)
LS103	*trpC2 ΔcwlO::cm*^{r}	Salzberg et al. (2013)
LU1	*trpC2 ΔlytE::spc*^{r}	LFH > WT168
LU17	*trpC2 ΔcwlO::*$P_{cwlOleader}$*lytE erm*^{r}	LFH > WT168
LU21	*trpC2 ΔcwlO::*$P_{cwlOleader}$*lytE erm*^{r} *ΔlytE::spc*^{r}	LFH > LU17
LU23	*trpC2 ΔcwlO::*P_{cwlO}*lytE erm*^{r}	LFH > WT168
LU25	*trpC2 ΔcwlO::*P_{cwlO}*lytE erm*^{r} *ΔlytE::spc*^{r}	LFH > LU23
LU45	*trpC2 lytE-HA kan*^{r}	pLU2 > WT168
LU47	*trpC2 lytE-HA kan*^{r} *ΔcwlO::cm*^{r}	pLU2 > LS103
LU61	*trpC2 ΔcwlO::*$P_{cwlOleader}$*lytE-HA kan*^{r} *erm*^{r} *ΔlytE::spc*^{r}	pLU2 > LU21
LU63	*trpC2 ΔcwlO::*P_{cwlO}*lytE-HA kan*^{r} *erm*^{r} *ΔlytE::spc*^{r}	pLU2 > LU25
LU72	*trpC2 ΔcwlS*	LFH dif > WT168
LU74	*trpC2 ΔlytF*	LFH dif > WT168
LU76	*trpC2 ΔyqgT*	LFH dif > WT168
LU106	*trpC2 ΔcwlO::*$P_{cwlOleader}$*cwlS erm*^{r} *ΔcwlS*	LFH > LU72

Strain	Genotype	Source
LU111	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytF erm*^{r} Δ*lytF*	LFH > LU74
LU112	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*yqgT erm*^{r} Δ*yqgT*	LFH > LU76
LU116	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytF erm*^{r} Δ*lytE::spc*^{r} Δ*lytF*	LFH > LU111
LU123	*trpC2* Δ*cwlO::P*$_{cwlO}$*cwlS erm*^{r} Δ*cwlS*	LFH > LU72
LU125	*trpC2* Δ*cwlO::P*$_{cwlO}$*lytF erm*^{r} Δ*lytF*	LFH > LU74
LU134	*trpC2* Δ*cwlO::P*$_{cwlO}$*lytF erm*^{r} Δ*lytE::spc*^{r} Δ*lytF*	LFH > LU125
LU145	*trpC2 cwlS-HA kan*^{r}	pLU7 > WT168
LU147	*trpC2 lytF-HA kan*^{r}	pLU6 > WT168
LU149	*trpC2 yqgT-HA kan*^{r}	pLU8 > WT168
LU151	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytF-HA erm*^{r} *kan*^{r} Δ*lytF*	pLU6 > LU111
LU153	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytF-HA erm*^{r} *kan*^{r} Δ*lytE::spc*^{r} Δ*lytF*	pLU6 > LU116
LU155	*trpC2* Δ*cwlO::P*$_{cwlO}$*lytF-HA erm*^{r} *kan*^{r} Δ*lytF*	pLU6 > LU125
LU157	*trpC2* Δ*cwlO::P*$_{cwlO}$*lytF-HA erm*^{r} *kan*^{r} Δ*lytE::spc*^{r} Δ*lytF*	pLU6 > LU134
LU159	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*cwlS-HA erm*^{r} *kan*^{r} Δ*cwlS*	pLU7 > LU106
LU161	*trpC2* Δ*cwlO::P*$_{cwlO}$*cwlS-HA erm*^{r} *kan*^{r} Δ*cwlS*	pLU7 > LU123
LU163	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*yqgT-HA erm*^{r} *kan*^{r} Δ*yqgT*	pLU8 > LU112
LU205	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE [L11cat] kan*^{r} Δ*lytE::cm*^{r}	LFH > LU239
LU207	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE [L12cat] kan*^{r}	Lin. pLU15 > WT168
LU209	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE [L12cat] kan*^{r} Δ*lytE::cm*^{r}	LFH > LU207
LU212	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE [L13cat] kan*^{r}	Lin. pLU14 > WT168
LU213	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE [L13cat] kan*^{r} Δ*lytE::cm*^{r}	LFH > LU212
LU216	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE [L21cat] kan*^{r}	Lin. pLU17 > WT168
LU217	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE [L21cat] kan*^{r} Δ*lytE::cm*^{r}	LFH > LU216
LU220	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE [L22cat] kan*^{r}	Lin. pLU18 > WT168
LU222	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE [L22cat] kan*^{r} Δ*lytE::cm*^{r}	LFH > LU220
LU223	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE [L23cat] kan*^{r}	Lin. pLU19 > WT168
LU226	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE [L23cat] kan*^{r} Δ*lytE::cm*^{r}	LFH > LU223
LU227	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE [L31cat] kan*^{r}	Lin. pLU20 > WT168
LU229	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE [L31cat] kan*^{r} Δ*lytE::cm*^{r}	LFH > LU227
LU232	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE [L32cat] kan*^{r}	Lin. pLU21 > WT168
LU233	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE [L32cat] kan*^{r} Δ*lytE::cm*^{r}	LFH > LU232
LU235	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE [L33cat] kan*^{r}	Lin. pLU22 > WT168
LU238	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE [L33cat] kan*^{r} Δ*lytE::cm*^{r}	LFH > LU235
LU239	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE [L11cat] kan*^{r}	Lin. pLU16 > WT168
LU261	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE [L1cat] kan*^{r}	Lin. pLU23 > WT168
LU262	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE [L1cat] kan*^{r} Δ*lytE::cm*^{r}	LFH > LU261
LU265	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE [L2cat] kan*^{r}	Lin. pLU24 > WT168
LU266	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE [L2cat] kan*^{r} Δ*lytE::cm*^{r}	LFH > LU265
LU269	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE [L3cat] kan*^{r}	Lin. pLU25 > WT168
LU270	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE [L3cat] kan*^{r} Δ*lytE::cm*^{r}	LFH > LU269
LU272	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE [L311cat] kan*^{r}	Lin. pLU26 > 168
LU273	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE [L311cat] kan*^{r} Δ*lytE::cm*^{r}	LFH > LU272
LU276	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE [L331cat] kan*^{r}	Lin. pLU27 > WT168
LU277	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE [L331cat] kan*^{r} Δ*lytE::cm*^{r}	LFH > LU276
LU280	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE [L233cat] kan*^{r}	Lin. pLU28 > WT168
LU281	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE [L233cat] kan*^{r} Δ*lytE::cm*^{r}	LFH > LU280
LU286	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE [L222cat] kan*^{r}	Lin. pLU30 > WT168
LU287	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE [L222cat] kan*^{r} Δ*lytE::cm*^{r}	LFH > LU286
LU288	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE [L212cat] kan*^{r}	Lin. pLU31 > WT168
LU289	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE [L212cat] kan*^{r} Δ*lytE::cm*^{r}	LFH > LU288

Strain	Genotype	Source
LU312	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE [L213cat] kan*^{r}	Lin. pLU29 > WT168
LU313	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE [L213cat] kan*^{r} Δ*lytE::cm*^{r}	LFH > LU312
LU314	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE [cat] kan*^{r}	Lin. pLU37 > WT168
LU315	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE [L113cat] kan*^{r}	Lin. pLU38 > WT168
LU316	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE [L113cat] kan*^{r} Δ*lytE::cm*^{r}	LFH > LU316
LU318	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE [L123cat]*^{R} *kan*^{r}	Lin. pLU40 > WT168
LU320	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE [L123cat]*^{R} *kan*^{r} Δ*lytE::cm*^{r}	LFH > LU320
LU331	*trpC2* Δ*lytE::spc*^{r}	LFH > WT168
LU335	*trpC2* Δ*lytE::P*$_{lytE}$*lytE [L123cat]*^{R} *erm*^{r}	LFH > WT168
LU337	*trpC2* Δ*lytE::P*$_{lytE}$*lytE [L1cat] erm*^{r}	LFH > WT168
LU339	*trpC2* Δ*lytE::P*$_{lytE}$*lytE [L2cat] erm*^{r}	LFH > WT168
LU341	*trpC2* Δ*lytE::P*$_{lytE}$*lytE [L3cat] erm*^{r}	LFH > WT168
LU343	*trpC2* Δ*lytE::P*$_{lytE}$*lytE [L123cat]*^{R} *erm*^{r} Δ*cwlO::kan*^{r}	LFH > LU335
LU350	*trpC2* Δ*lytE::P*$_{lytE}$*lytE [L1cat] erm*^{r} Δ*cwlO::kan*^{r}	LFH > LU337
LU352	*trpC2* Δ*lytE::P*$_{lytE}$*lytE [L2cat] erm*^{r} Δ*cwlO::kan*^{r}	LFH > LU339
LU354	*trpC2* Δ*lytE::P*$_{lytE}$*lytE [L3cat] erm*^{r} Δ*cwlO::kan*^{r}	LFH > LU341
LU372	*trpC2* Δ*lytE::P*$_{lytE}$*lytE [L31cat] erm*^{r}	LFH > WT168
LU374	*trpC2* Δ*lytE::P*$_{lytE}$*lytE [L331cat] erm*^{r}	LFH > WT168
LU383	*trpC2* Δ*lytE::P*$_{lytE}$*lytE [L31cat] erm*^{r} Δ*cwlO::kan*^{r}	LFH > LU372
LU385	*trpC2* Δ*lytE::P*$_{lytE}$*lytE [L331cat] erm*^{r} Δ*cwlO::kan*^{r}	LFH > LU374
LU398	*trpC2* Δ*lytE::P*$_{lytE}$*lytE [cat] erm*^{r}	LFH > WT168
LU399	*trpC2* Δ*lytE::P*$_{lytE}$*lytE [L11cat] erm*^{r}	LFH > WT168
LU401	*trpC2* Δ*lytE::P*$_{lytE}$*lytE [L21cat] erm*^{r}	LFH > WT168
LU403	*trpC2* Δ*lytE::P*$_{lytE}$*lytE [L13cat] erm*^{r}	LFH > WT168
LU405	*trpC2* Δ*lytE::P*$_{lytE}$*lytE [L213cat] erm*^{r}	LFH > WT168
LU407	*trpC2* Δ*lytE::P*$_{lytE}$*lytE [L11cat] erm*^{r} Δ*cwlO::kan*^{r}	LFH > LU399
LU409	*trpC2* Δ*lytE::P*$_{lytE}$*lytE [L21cat] erm*^{r} Δ*cwlO::kan*^{r}	LFH > LU401
LU411	*trpC2* Δ*lytE::P*$_{lytE}$*lytE [L13cat] erm*^{r} Δ*cwlO::kan*^{r}	LFH > LU403
LU413	*trpC2* Δ*lytE::P*$_{lytE}$*lytE [L213cat] erm*^{r} Δ*cwlO::kan*^{r}	LFH > LU405
LU414	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE-HA [L11cat] kan*^{r} *spc*^{r} Δ*lytE::cm*^{r}	pLU46 > LU205
LU415	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE-HA [L13cat] kan*^{r} *spc*^{r} Δ*lytE::cm*^{r}	pLU46 > LU213
LU416	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE-HA [L21cat] kan*^{r} *spc*^{r} Δ*lytE::cm*^{r}	pLU46 > LU217
LU417	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE-HA [L31cat] kan*^{r} *spc*^{r} Δ*lytE::cm*^{r}	pLU46 > LU229
LU418	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE-HA [L1cat] kan*^{r} *spc*^{r} Δ*lytE::cm*^{r}	pLU46 > LU262
LU419	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE-HA [L2cat] kan*^{r} *spc*^{r} Δ*lytE::cm*^{r}	pLU46 > LU266
LU421	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE-HA [L123cat]*^{R} *kan*^{r} *spc*^{r} Δ*lytE::cm*^{r}	pLU46 > LU320
LU422	*trpC2* Δ*lytE::P*$_{lytE}$*lytE-HA [L123cat]*^{R} *erm*^{r} *spc*^{r} Δ*cwlO::kan*^{r}	pLU46 > LU343
LU424	*trpC2* Δ*lytE::P*$_{lytE}$*lytE-HA [L331cat] erm*^{r} *spc*^{r} Δ*cwlO::kan*^{r}	pLU46 > LU385
LU425	*trpC2* Δ*lytE::P*$_{lytE}$*lytE-HA [L11cat] erm*^{r} *spc*^{r} Δ*cwlO::kan*^{r}	pLU46 > LU407
LU427	*trpC2* Δ*lytE::P*$_{lytE}$*lytE-HA [L13cat] erm*^{r} *spc*^{r} Δ*cwlO::kan*^{r}	pLU46 > LU411
LU432	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE-HA [L213cat] kan*^{r} *spc*^{r} Δ*lytE::cm*^{r}	pLU46 > LU313
LU435	*trpC2* Δ*lytE::P*$_{lytE}$*lytE-HA [L1cat] erm*^{r} *spc*^{r} Δ*cwlO::kan*^{r}	pLU46 > LU350
LU440	*trpC2* Δ*cwlO::P*$_{cwlOleader}$*lytE-HA [L3cat] kan*^{r} *spc*^{r} Δ*lytE::cm*^{r}	pLU46 > LU270
LU442	*trpC2* Δ*lytE::P*$_{lytE}$*lytE-HA [L2cat] erm*^{r} *spc*^{r} Δ*cwlO::kan*^{r}	pLU46 > LU352
LU443	*trpC2* Δ*lytE::P*$_{lytE}$*lytE-HA [L213cat] erm*^{r} *spc*^{r} Δ*cwlO::kan*^{r}	pLU46 > LU413
WT168	*trpC2*	Laboratory stock, Kunst et al., 1997

[LFH] long flanking homology fragment; [cm^r] chloramphenicol resistance gene; [emr^r] erythromycin resistance gene; [kan^r] kanamycin resistance gene; [spc^r] spectinomycin resistance gene

Table S3: List of plasmids used in this study

Plasmid	Genotype	Source
pDG780	plasmid with *kan*^{r} cassette	Guerout-Fleury et al., 1995
pDG646	plasmid with *erm*^{r} cassette	Guerout-Fleury et al., 1995
pDG1727	plasmid with *spc*^{r} cassette	Guerout-Fleury et al., 1995
pDG1661	plasmid with *cm*^{r} cassette	Guerout-Fleury et al., 1995
pETLytE	pET21b+ derivative, protein expression vector for expression of the wild-type LytE protein	unpublished, David Noone
pLU2	pDG780 derivative enabling C-terminal HA-tagging of *lytE* over single cross-over, *kan*^{r} cassette	This work
pLU6	pDG780 derivative enabling C-terminal HA-tagging of *lytF* over single cross-over, *kan*^{r} cassette	This work
pLU9	pDG780 derivative with upstream homology to the *cwlO* locus (including *cwlO* promoter and *cwlO* leader sequence)	This work
pLU10	pLU9 derivative with up- and downstream homology to the *cwlO* locus (including *cwlO* promoter and *cwlO* leader sequence)	This work
pLU14	pLU10 derivative enabling integration of variant *lytE [L13cat]* at the *cwlO* locus by double cross-over, *kan*^{r} cassette	This work
pLU15	pLU10 derivative enabling integration of variant *lytE [L12cat]* at the *cwlO* locus by double cross-over, *kan*^{r} cassette	This work
pLU16	pLU10 derivative enabling integration of variant *lytE [L11cat]* at the *cwlO* locus by double cross-over, *kan*^{r} cassette	This work
pLU17	pLU10 derivative enabling integration of variant *lytE [L21cat]* at the *cwlO* locus by double cross-over, *kan*^{r} cassette	This work
pLU18	pLU10 derivative enabling integration of variant *lytE [L22cat]* at the *cwlO* locus by double cross-over, *kan*^{r} cassette	This work
pLU19	pLU10 derivative enabling integration of variant *lytE [L23cat]* at the *cwlO* locus by double cross-over, *kan*^{r} cassette	This work
pLU20	pLU10 derivative enabling integration of variant *lytE [L31cat]* at the *cwlO* locus by double cross-over, *kan*^{r} cassette	This work
pLU21	pLU10 derivative enabling integration of variant *lytE [L32cat]* at the *cwlO* locus by double cross-over, *kan*^{r} cassette	This work
pLU22	pLU10 derivative enabling integration of variant *lytE [L33cat]* at the *cwlO* locus by double cross-over, *kan*^{r} cassette	This work
pLU23	pLU10 derivative enabling integration of variant *lytE [L1cat]* at the *cwlO* locus by double cross-over, *kan*^{r} cassette	This work
pLU24	pLU10 derivative enabling integration of variant *lytE [L2cat]* at the *cwlO* locus by double cross-over, *kan*^{r} cassette	This work
pLU25	pLU10 derivative enabling integration of variant *lytE [L3cat]* at the *cwlO* locus by double cross-over, *kan*^{r} cassette	This work
pLU26	pLU20 derivative enabling integration of variant *lytE [L311cat]* at the cwlO locus by double cross-over, *kan*^{r} cassette	This work
pLU27	pLU20 derivative enabling integration of variant *lytE [L331cat]* at the cwlO locus by double cross-over, *kan*^{r} cassette	This work
pLU28	pLU19 derivative enabling integration of variant *lytE [L233cat]* at the *cwlO* locus by double cross-over, *kan*^{r} cassette	This work
pLU29	pLU19 derivative enabling integration of variant *lytE [L213cat]* at the *cwlO* locus by double cross-over, *kan*^{r} cassette	This work
pLU30	PLU18 derivative enabling integration of variant *lytE [L222cat]* at the *cwlO* locus by double cross-over, *kan*^{r} cassette	This work
pLU31	pLU18 derivative enabling integration of variant *lytE [L212cat]* at the *cwlO* locus by double cross-over, *kan*^{r} cassette	This work
pLU37	pLU10 derivative enabling integration of variant *lytE [cat]* at the *cwlO* locus by double cross-over, *kan*^{r} cassette	This work
pLU38	pLU14 derivative enabling integration of variant *lytE [L113cat]* at the cwlO locus by double cross-over, k*an*^{r} cassette	This work
pLU40	pLU14 derivative enabling integration of variant *lytE [L123cat]*^{R} at the cwlO locus by double cross-over, *kan*^{r} cassette	This work
pLU46	pLU2 derivative enabling C-terminal HA-tagging of *lytE* over single cross-over, *spc*^{r} cassette	This work
pLU47	pET21b+ derivative, protein expression vector for expression of the variant LytE [L123cat]^R	This work
pLU48	pET21b+ derivative, protein expression vector for expression of the variant LytE [L31cat]	This work
pLU49	pET21b+ derivative, protein expression vector for expression of the variant LytE [L13cat]	This work
pLU50	pET21b+ derivative, protein expression vector for expression of the variant LytE [L1cat]	This work
pLU51	pET21b+ derivative, protein expression vector for expression of the variant LytE [L3cat]	This work
pLU52	pET21b+ derivative, protein expression vector for expression of the variant LytE [cat]	This work

[cm^r] chloramphenicol resistance gene; [erm^r] erythromycin resistance gene; [kan^r] kanamycin resistance gene; [spc^r] spectinomycin resistance gene

Table S4: List of oligonucleotids used in this study

Primer name	Sequence
16S_Bsu_F	GTCAGCTCGTGTCGTGAGATG
16S_Bsu_R	CAACTGAATGCTGGCAACTAAGATC
Aaa_LysM_Fwd	cgcggatccacgttatttgcaggagcggcatctgcacaaGTAAAACCGGGTGAATCAG
Aaa_LysM_Rev	atgtagatgacgttttgctgcttgaagacgatgaagatgaagcaggttttgagctgcttgtgcttgttcc TACTTTTAGTACTTGGTTTG
Catalytic_LytE_Fwd_Spac	ttcaagcagcaaaacgtcatctacatcacttaatgtgagcaagctggtttctgatgcaaaagcgttagtc GGAACGCCATATAAATGGGG
Catalytic_LytE_Rev_Ery	attatgtcttttgcgcagtcggcGAAAAAAACCCGTTCCAATGAACGG
Catalytic_LytE_Rev_Xmal	tccccccgggGAAAAAAACCCGTTCCAATGAACGG
CwlO_DF_LysM	ccgctcgagATTAACATATACCCTTTTCGAATA
CwlO_DR_LysM	cggggtaccCCTTCTTCATTCGTAATGCCCAGGT
CwlO_UF	CCTTCATCTGTGGAGCAG
CwlO_UF_LysM	aaggaaaaaagcggccgcATCATTATATAAAACCGGCTT
CwlO_UF_Smal	gcatcccgggCCTTCATCTGTGGAGCAG
CwlO_UR	cgtagctgtaatgatttgctttttcatTTTTTATATCCTCCCTTTTACTGTGAACC
CwlO_UR_CwlS	aagccggctacaatcttctttttcatTTTTTATATCCTCCCTTTTACTGTGAA
CwlO_UR_LysM	taacgtggatcctaaaacaactgctgtcgtagctgtaatgatttgctttttcat TTTTTATATCCTCCCTTTTAC
CwlO_UR_LytF	gccctgctgctaatttctttttcatTTTTTATATCCTCCCTTTTACTGTGAA
CwlO_UR_Yqgt	gtttcaccgtaataaacgccatcatTTTTTATATCCTCCCTTTTACTGTGAA
CwlS_Do_Fwd	tatattatgtaaactaatatgcaagAAAATGAGTATGACAAAAAGCC
CwlS_Do_Rev	TCTGGGGAAATAGGCTGCTCT
CwlS_Fwd	ATGAAAAAGAAGATTGTAGCCGGC
CwlS_LysM_Fwd	cgcggatccacgttatttgcaggagcggcatctgcacaaATAAAAGTAAAAAGCGGTGA
CwlS_LysM_Rev	atgtagatgacgttttgctgcttgaagacgatgaagatgaagcaggttttgagctgcttgtgcttgttcc AGGAACCTTAAGACTTTGTC
Cwls_Rev_Ery	tcgattatgtcttttgcgcagtcggcTAAACAAAAAAGCCTATAAAAGGC
CwlS_Tag_F_BamHl	cgggatccGTCGGAGTGCCTTACCGATGGGGC
CwlS_Tag_R_HA_EcoRl	ggaattctaaacaaaaaagcctataaaaggcttttgtcatactcatttttattatgcataatccg gaacatcatacggataAAAATAACTTCTTGCGCCCAAGTA
CwlS_Up_Fwd	CGGCAAGCATATTAAAACAAGTGAG
CwlS_Up_Rev	tatattctaggaagtaggtattcacTTCATTCAACCTCCTAAACTG
Exp_Cat_Rev	ccgctcgagGAATCTTTTCGCACCGAGGTAACG
Exp_L1_Fwd	ggaattccatatgAGCATTAAGGTGAAAAAAGGCGACA
Exp_L2_Fwd	ggaattccatatgACATACAAAGTAAAGAGCGGGGACA
Exp_L3_Fwd	ggaattccatatgACTTATAAAGTGAAGAGCGGAGACA
Exp_LytF_Fwd	ggaattccatatgATTAAGGTCAAAAGCGGAGATTCTTT
Exp_Spac_Fwd	ggaattccatatgGGAACAAGCACAAGCAGCTCAAAAC
Exp_XlyB_Fwd	ggaattccatatgTATCATGTAAAAAAAGGTGACACACTG
L1_Fwd_Link	ggcaaatctactagttcaaaaagcagcagttcttcttcctcttcttct AGCATTAAGGTGAAAAAAGGC
L1_Fwd_Sig_BamHl	cgcggatccacgttatttgcaggagcggcatctgcacaa AGCATTAAGGTGAAAAAAGGC
L1_Rev_Link	agaagaagaggaagaagaactgctgctttttgaactagtagatttgcc GTTAATCGATAAAGTTTGTCCC
L1_Rev_Spac	atgtagatgacgttttgctgcttgaagacgatgaagatgaagcaggttttgagctgcttgtgcttgttcc GTTAATCGATAAAGTTTGTC
L2_Fwd_Link	ggcaaatctactagttcaaaaagcagcagttcttcttcctcttcttct ACATACAAAGTAAAGAGCGGG
L2_Fwd_Sig_BamHl	cgcggatccacgttatttgcaggagcggcatctgcacaa ACATACAAAGTAAAGAGCGGG
L2_Rev_Link	agaagaagaggaagaagaactgctgctttttgaactagtagatttgcc TTTCAGTTTCAGGACTTGTCC
L2_Rev_Spac	atgtagatgacgttttgctgcttgaagacgatgaagatgaagcaggttttgagctgcttgtgcttgttcc TTTCAGTTTCAGGACTTGTC
L3_Fwd_Link	ggcaaatctactagttcaaaaagcagcagttcttcttcctcttcttct ACTTATAAAGTGAAGAGCGGA
L3_Fwd_Sig_BamHl	cgcggatccacgttatttgcaggagcggcatctgcacaa ACTTATAAAGTGAAGAGCGGA
L3_Rev_Link	agaagaagaggaagaagaactgctgctttttgaactagtagatttgcc TTTCACCTTCAATACTTGGTTG

Primer name	Sequence
L3_Rev_Spac	atgtagatgacgttttgctgcttgaagacgatgaagatgaagcaggttttgagctgcttgtgcttgttcc TTTCACCTTCAATACTTGGT
LytE_DF	ATGAAAAAGCAAATCATTACAGCTACG
LytE_Do_Fwd_Ery	ggttgccagagttaaaggatcATTAGACGGAGCACAGGCGTTTT
LytE_Do_Fwd_Spec	caataaacccttgccctcgctacgTACTGGAAGCCTCGTTACCTCGG
LytE_Do_Fwd_Spec	caataaacccttgccctcgctacgATTAGACGGAGCACAGGCGTTTT
LytE_Do_Rev	GTATTTGCCGTATGCGCTCAGGC
LytE_DR_Ery	attatgtcttttgcgcagtcggcGAAAAAAACCCGTTCCAATGAA
LytE_Probe_F	TATCGATTAACGGCAAATCTACAAGTT
LytE_Probe_T7_R	taatacgactcactatagggGCAGTGCTTGTTCTGCCCACACTT
LytE_qPCR_F2_	GGAGGAAAATATGAAAAAGCAAATCA
LytE_qPCR_R2	TCGCTCTCATACGTATCAACACTCT
LytE_Spacer_Fwd_Sig	cgcggatccacgttatttgcaggagcggcatctgcacaaGGAACAAGCACAAGCAGC
LytE_Tag_F_BamHI	cgggatccTTAGTCGGAACGCCATATAAATGG
LytE_Tag_R_HA_EcoRI	ggaattcgaaaaaaacccgttccaatgaacgggttttctctaaaaattattatgcata atccggaacatcatacggataGAATCTTTTCGCACCGAGGTAACG
LytE_Up_Fwd	AACGGTAGGCTCTTATGTAAGGC
LytE_Up_Fwd_2	AACGGTAGGCTCTTATGTAAGGC
LytE_Up_Rev_Spec	cgttacgttattagcgagccagtcCCTAAAACAACTGCTGTCGTAGC
LytF_Do_Fwd	tatattatgtaaactaatatgcaagTATTAGAGCCTAAAGACGTTGATGC
LytF_Do_Rev	CCGGGATTTGTTTATGTGAGAA
LytF_Fwd	ATGAAAAAGAAATTAGCAGCAGGG
LytF_LysM_Fwd	cgcggatccacgttatttgcaggagcggcatctgcacaaATTAAGGTCAAAAGCGGAG
LytF_LysM_Rev	atgtagatgacgttttgctgcttgaagacgatgaagatgaagcaggttttgagctgcttgtgcttgttcc TGGAATAGTAAGCGTCTGG
LytF_qPCR_Fwd	TGTCGGAGATTTCGTTTTCTTTTC
LytF_qPCR_fwd	TGTCGGAGATTTCGTTTTCTTTTC
LytF_qPCR_Rev	CCCGTTGCCTAGGTAAATTCC
LytF_qPCR_rev	CCCGTTGCCTAGGTAAATTCC
LytF_Rev_Ery	tcgattatgtcttttgcgcagtcggcGCATCAACGTCTTTAGGCTCTAATA
LytF_Tag_F_BamHl	cgggatccCTTGGGGTTCCGTATCGCTGGGGA
LytF_Tag_R_HA_EcoRl	ggaattcgcatcaacgtctttaggctctaataaaaaactgtacagccgttaaggccgcacagtttctg tttttattatgcataatccggaacatcatacggataGAAATATCGTTTTGCACCGAGATA
LytF_Up_Fwd	CTTTTGGTTTATCCTTCGGCCTT
LytF_Up_Rev	tatattctaggaagtaggtattcacCTTAGCTCTCCTTTTTCCCCC
MreBH_qPCR_F	GCAAAACCCGCTTGAATCTG
MreBH_qPCR_R	TGCCTTTTGCAGCTTATCAATC
OE93_Fwd_Ery	ggttgccagagttaaaggatcGAGCAATCCTTACTGGAAAGCAGC
OE94_Rev	ATCGACAAGCGTTACGTTACCGAC
SigI_Chip_Fwd	AAATTCTCTGCTCGCATTTGC
SigI_Chip_Rev	GGAAAGCTCGGCTTCAATCAG
Signal_LytE_Fwd	ATGAAAAAGCAAATCATTACAGCTACG
Signal_LytE_Rev	TTGTGCAGATGCCGCTCCTGCAA
Signal_LytE_Rev_Spec	cgttacgttattagcgagccagtcTTGTGCAGATGCCGCTCCTGCAA
Spacer_Check_Rev	GATGTAGATGACGTTTTGC
Spec_Fwd_Ecorl	ccggaattccggCTAATAACGTAACGTGACTGGCA
Spec_Rev_Sall	ttccgcggccgctatggccgacgtcgacgcgtAGGGTTTATTGTTTTCTAAAATCTG
YocH_Northern_F	GCAGCAAACACAGCTACTGAGAAT
YocH_Northern_R	GCTCTTTTTGCACAGCCTCTT
YqgT_Do_Fwd	tatattatgtaaactaatatgcaagACTGAAACTTTTTCTTTTTTCTGC
YqgT_Do_Rev	TTGGAGCGGCCATTTCGCTGT
YqgT_Fwd	ATGATGGCGTTTATTACGGTG

Primer name	Sequence
YqgT_Rev_Ery	tcgattatgtcttttgcgcagtcggcTCTCATATTTGCTTCGGCAGAA
YqgT_Tag_F_BamHl	cgggatccACGGAGAATCCGCCTGACAGACTG
YqgT_Tag_R_HA_EcoRl	ggaattctctcatatttgcttcggcagaaaaagaaaaagtttcagttattatgcataatccgg aacatcatacggataCTCATGAAAAAATAACGCCCTCCA
YqgT_Up_Fwd	TTACTCGAGGCTGATTACACTG
YqgT_Up_Rev	tatattctaggaagtaggtattcacTCATCCCCCTTTTTTGTTACTATA
YvcE_3qPCR_Orf*2	GGCGGATCTCAAGGATACATAGATG
YvcE_5qPCR_Orf*2	GTCTGCGTCCACAATAGATGAAAC

* lower case = 5´ overhang, upper case = annealing sequence